# Bibliografía Trabajo Investigación

**Aclaraciones:**

Referencias dentro del documento: textos en rojo

Links al servidor (libros): números en rojo

Libros (servidor): http://www.4shared.com/dir/9020422/28749742/LIBROS.html

Índices libros HTML: http://acordeon.eresmas.net/meta5/biblio.html

Web: http://www.terra.es/personal/marcos54/

Ir al índice

| | | | | | |
|---|---|---|---|---|---|
| 2009 | 1 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 1 The origins and functions of music<br>**1 The nature of music and its evolution** | Cross, fan | OUP | Inglés |
| 2009 | 2 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 1 The origins and functions of music<br>**2 Universals in music processing** | Stevens , Catherine<br>Byron, Tim | OUP | Inglés |
| 2009 | 3 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 1 The origins and functions of music<br>**3 Music and meaning** | Cross, fan<br>Tofbert, Elizabeth | OUP | Inglés |
| 2009 | 4 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 1 The origins and functions of music<br>**4 The social and personal functions of music in cross-cultural perspective** | Cfayton, Martin | OUP | Inglés |
| 2009 | 5 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 2 Music perception<br>**5 The perception of pitch** | Stainsby, Thomas<br>Cross, fan | OUP | Inglés |
| 2009 | 6 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 2 Music perception<br>**6 Tonal cognition** | Bigand, Emmanuef<br>Poulin-Charronnat, Bénédicte | OUP | Inglés |
| 2009 | 7 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 2 Music perception<br>**7 The perception of musical timbre** | McAdams, Stephen<br>L. Giordano, Bruno | OUP | Inglés |
| 2009 | 8 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 2 Music perception<br>**8 Musical time** | Riess Jones, Mari | OUP | Inglés |
| 2009 | 9 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 2 Music perception<br>**9 Components of melodic processing** | A. Schmuckfer, Mark | OUP | Inglés |
| 2009 | 10 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 2 Music perception<br>**10 Memory for music** | Snyder, Bob | OUP | Inglés |
| 2009 | 11 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 3 Responses to music<br>**11 Bodily responses to music** | A. Hodges, Donafd | OUP | Inglés |
| 2009 | 12 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 3 Responses to music<br>**12 Emotional responses to music** | N. Juslin, Patrik | OUP | Inglés |
| 2009 | 13 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 3 Responses to music<br>**13 The relationship between musical structure and perceived expression** | Gabrielsson, Alf | OUP | Inglés |
| 2009 | 14 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 3 Responses to music<br>**14 Aesthetics** | Huron, David | OUP | Inglés |
| 2009 | 15 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 3 Responses to music<br>**15 Musical preferences** | Lamont, Alexandra<br>Greasley, Alinka | OUP | Inglés |
| 2009 | 16 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 4 Music and the brain<br>**16 The neurobiological basis of musical expectations** | J. Trainor , Laurel<br>J. Zatorre, Robert | OUP | Inglés |
| 2009 | 17 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 4 Music and the brain<br>**17 Disorders of musical cognition** | Colectivo | OUP | Inglés |
| 2009 | 18 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 4 Music and the brain<br>**18 Music, musicians, and brain plasticity** | Schlaug, Gottfried | OUP | Inglés |
| 2009 | 19 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 4 Music and the brain<br>**19 Music and the brain: three links to language** | D. Patel, Aniruddh | OUP | Inglés |
| 2009 | 20 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 5 Musical development<br>**20 Prenatal development and the phylogeny and ontogeny of music** | Parncutt, Richard | OUP | Inglés |
| 2009 | 21 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 5 Musical development<br>**21 Music lessons from infants** | E. Trehub, Sandra | OUP | Inglés |

| | | | | | |
|---|---|---|---|---|---|
| **2009** | 22 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 5 Musical development<br>**22 Music in the school years** | Lamont, Alexandra | OUP | Inglés |
| **2009** | 23 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 5 Musical development<br>**23 The impact of music instruction on other skills** | H. Rauscher, Frances | OUP | Inglés |
| **2009** | 24 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 6 Learning musical skills<br>**24 Musical potential** | McPherson, Gary<br>Hallam, Susan | OUP | Inglés |
| **2009** | 25 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 6 Learning musical skills<br>**25 Practising** | Jorgensen, Harald<br>Hallam, Susan | OUP | Inglés |
| **2009** | 26 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 6 Learning musical skills<br>**26 Individuality in the learning of musical skills** | Gaunt, Helena<br>Hallam, Susan | OUP | Inglés |
| **2009** | 27 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 6 Learning musical skills<br>**27 Motivation to learn** | Hallam, Susan | OUP | Inglés |
| **2009** | 28 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 6 Learning musical skills<br>**28 The role of the family in supporting learning** | Creech, Andrea | OUP | Inglés |
| **2009** | 29 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 6 Learning musical skills<br>**29 The role of the institution and teachers in supporting learning** | Welch, Graham<br>Ockelford, Adam | OUP | Inglés |
| **2009** | 30 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 7 Musical performance<br>**30 Measurement and models of performance** | Windsor, W. Luke | OUP | Inglés |
| **2009** | 31 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 7 Musical performance<br>**31 Planning and performance** | Altenmuller, Eckart<br>Schneider, Sabine | OUP | Inglés |
| **2009** | 32 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 7 Musical performance<br>**32 Sight-reading** | C. Lehmann, Andreas<br>Kopiez, Reinhard | OUP | Inglés |
| **2009** | 33 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 7 Musical performance<br>**33 Performing from memory** | Colectivo | OUP | Inglés |
| **2009** | 34 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 7 Musical performance<br>**34 Movement and collaboration in musical performance** | W. Davidson, Jane | OUP | Inglés |
| **2009** | 35 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 7 Musical performance<br>**35 Emotion in music performance** | N. Juslin, Patrik | OUP | Inglés |
| **2009** | 36 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 7 Musical performance<br>**36 Optimizing physical and psychological health in performing musicians** | T. Kenny, Dianna<br>Ackermann, Bronwen | OUP | Inglés |
| **2009** | 37 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 8 Composition and improvisation<br>**37 Making a mark: the psychology of composition** | Impett, Jonathan | OUP | Inglés |
| **2009** | 38 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 8 Composition and improvisation<br>**38 Musical improvisation** | Ashley, Richard | OUP | Inglés |
| **2009** | 39 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 8 Composition and improvisation<br>**39 Children as creative thinkers in music: focus on composition** | R. Webster, Peter | OUP | Inglés |
| **2009** | 40 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 9 The role of music in everyday life<br>**40 Choosing to hear music: motivation, process and effect** | Colectivo | OUP | Inglés |
| **2009** | 41 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 9 The role of music in everyday life<br>**41 Music in performance arts: film, theatre, and dance** | J. Cohen, Annabel | OUP | Inglés |
| **2009** | 42 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 9 The role of music in everyday life<br>**42 Peak experiences in music** | Colectivo | OUP | Inglés |

| | | | | | |
|---|---|---|---|---|---|
| **2009** | 43 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 9 The role of music in everyday life<br>**43 Musical identities** | Colectivo | OUP | Inglés |
| **2009** | 44 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 9 The role of music in everyday life<br>**44 The effects of music in community and educational settings** | Hallam, Susan<br>MacDonald, Raymond | OUP | Inglés |
| **2009** | 45 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 9 The role of music in everyday life<br>**45 Music and consumer behaviour** | C. North, Adrian<br>J. Hargreaves, David | OUP | Inglés |
| **2009** | 46 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 10 Music therapy<br>**46 Processes of music therapy: clinical and scientific rationales and models** | de l'Etoile, Shannon | OUP | Inglés |
| **2009** | 47 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 10 Music therapy<br>**47 Clinical practice in music therapy** | Hurt-Thaut, Corene | OUP | Inglés |
| **2009** | 48 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 10 Music therapy<br>**48 Research and evaluation in music therapy** | L. Wheeler, Barbara | OUP | Inglés |
| **2009** | 49 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 10 Music therapy<br>**49 Music therapy in medical and neurological rehabilitation settings** | Colectivo | OUP | Inglés |
| **2009** | 50 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 11 Conceptual frameworks, research methods and future directions<br>**50 Beyond music psychology** | Ockelford, Adam | OUP | Inglés |
| **2009** | 51 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 11 Conceptual frameworks, research methods and future directions<br>**51 History and research** | Thaut, Michael | OUP | Inglés |
| **2009** | 52 | **The Oxford Handbook of MUSIC PSYCHOLOGY**<br>Part 11 Conceptual frameworks, research methods and future directions<br>52 Where now? | Colectivo | OUP | Inglés |
| **2006** | 53 | **Music, Motor Control and the Brain**<br>Part 1: History<br>**1. Historical increases in expert music performance skills: optimising instruments, playing techniques and training** | Lehmann, Andreas C. | OUP | Inglés |
| **2006** | 54 | Music, Motor Control and the Brain<br>Part 2: Psychology<br>**2. From cognition to action** | Jäncke, Lutz | OUP | |
| **2006** | 55 | Music, Motor Control and the Brain<br>Part 2: Psychology<br>**3. The nature of memory for music performance skills** | Palmer, Caroline | OUP | |
| **2006** | 56 | Music, Motor Control and the Brain<br>Part 2: Psychology<br>**4. Musical synchronization** | Repp, Bruno | OUP | |
| **2006** | 57 | Music, Motor Control and the Brain<br>Part 3: Movement analysis<br>**5. Hand movements and musical performance** | Jerde, Thomas E.<br>Santello, Marco<br>Soechting, John F. | OUP | |
| **2006** | 58 | Music, Motor Control and the Brain<br>Part 3: Movement analysis<br>**6. Movement analysis in pianists** | Jabusch, Hans-Christian | OUP | |
| **2006** | 59 | Music, Motor Control and the Brain<br>Part 3: Movement analysis<br>**7. Fingering and bowing in violinists** | Wiesendanger, Mario<br>Baader, Andreas<br>Kazennikov, Oleg | OUP | |
| **2006** | 60 | Music, Motor Control and the Brain<br>PArt 3: Movement analysis<br>**8. Movements and analysis of drumming** | Dahl, Sofia | OUP | |
| **2006** | 61 | Music, Motor Control and the Brain<br>PArt 3: Movement analysis<br>**5. Hand movements and musical performance** | Repp, Bruno | OUP | |

| | | | | |
|---|---|---|---|---|
| **2006** | 62 | Music, Motor Control and the Brain<br>Part 4: Representation in the Brain<br>**9. Brain structures of musicians: executive functions and morphological implications** | Schlaug, Gottfried | OUP |
| **2006** | 63 | Music, Motor Control and the Brain<br>Part 4: Representation in the Brain<br>**10. The motor representation in pianists and string players** | Jäncke, Lutz | OUP |
| **2006** | 64 | Music, Motor Control and the Brain<br>Part 4: Representation in the Brain<br>**11. Brain activation during piano playing** | Bangert, Marc | OUP |
| **2006** | 65 | Music, Motor Control and the Brain<br>Part 4: Representation in the Brain<br>**12. Brain activation during string playing** | Nirkko, Arto<br>Kristeva, Rumyana | OUP |
| **2006** | 66 | Music, Motor Control and the Brain<br>Part 4: Representation in the Brain<br>**13. 'Singing in the (b)rain':<br>cerebral correlates of vocal music performance in humans** | Ackermann, Hermann<br>Wildgruber, Dirk<br>Riecker, Axel | OUP |
| **2006** | 67 | Music, Motor Control and the Brain<br>Part 4: Representation in the Brain<br>**14. Sensory-motor networks in singing and speaking:<br>a comparative approach** | Gordon, Reyna Leigh<br>Racette , Amelie Racette & Daniele Schon<br>Schon, Daniele | OUP |
| **2006** | 68 | Music, Motor Control and the Brain<br>Part 4: Representation in the Brain<br>**15. Role of inhibition in motor control of finger functions** | Gerloff , Christian<br>Hummel, Friedhelm | OUP |
| **2006** | 69 | Music, Motor Control and the Brain<br>Part 5: Apollo's curse - the loss of motor control in musicians<br>**16. The end of the song? Robert Schumann's focal dystonia** | Altenmüeller, Eckart | OUP |
| **2006** | 70 | Music, Motor Control and the Brain<br>Part 5: Apollo's curse - the loss of motor control in musicians<br>**17. Epidemiology, phenomenology and therapy of musician's cramp** | Jabusch, Hans-Christian<br>Altenmüeller, Eckart | OUP |
| **2006** | 71 | Music, Motor Control and the Brain<br>Part 5: Apollo's curse - the loss of motor control in musicians<br>**18. The neurophysiology of focal hand dystonia in musicians** | Rosenkranz, Karin | OUP |
| **2006** | 72 | Music, Motor Control and the Brain<br>Part 5: Apollo's curse - the loss of motor control in musicians<br>**19. The development of focal dystonia in musicians as a consequence of maladaptive plasticity:<br>implications for intervention** | Byl, Nancy<br>Priori, Alberto | OUP |
| **2006** | 73 | Music, Motor Control and the Brain<br>Part 5: Apollo's curse - the loss of motor control in musicians<br>**20. Music performance anxiety** | Kesselring, Jurg | OUP |
| **2006** | 74 | **La interpretación musical**<br>Primera parte: Conceptos y preconceptos<br>**1. La interpretación a través de la historia** | Lawson, Colin | Alianza Música |
| **2006** | 75 | La interpretación musical<br>Primera parte: Conceptos y preconceptos<br>**2. La interpretación histórica y el intérprete moderno** | Walls, Peter | Alianza Música |
| **2006** | 76 | La interpretación musical<br>Primera parte: Conceptos y preconceptos<br>**3. Análisis y (¿o?) interpretación** | Rink, John | Alianza Música |
| **2006** | 77 | La interpretación musical<br>Primera parte: Conceptos y preconceptos<br>**4. Comprender la psicología de la interpretación** | Clarke, Eric | Alianza Música |
| **2006** | 78 | La interpretación musical<br>Segunda parte: Aprender a interpretar<br>**5. Sobre la enseñanza de la interpretación** | Ritterman, Janet | Alianza Música |
| **2006** | 79 | La interpretación musical<br>Segunda parte: Aprender a interpretar<br>**6. El desarrollo de la habilidad interpretativa** | Davidson, Jane | Alianza Música |
| **2006** | 80 | La interpretación musical<br>Segunda parte: Aprender a interpretar<br>**7. Preparándose para interpretar** | Reid, Stefan | Alianza Música |

| | | | | | |
|---|---|---|---|---|---|
| **2006** | 81 | La interpretación musical<br>Segunda parte: Aprender a interpretar<br>**8. La memorización de la música** | Williamon, Aaron | Alianza Música | |
| **2006** | 82 | La interpretación musical<br>Tercera parte: Hacer música<br>**9. De la partitura al sonido** | Hill, Peter | Alianza Música | |
| **2006** | 83 | La interpretación musical<br>Tercera parte: Hacer música<br>**10. El lenguaje del cuerpo durante la interpretación** | Davidson, Jane | Alianza Música | |
| **2006** | 84 | La interpretación musical<br>Tercera parte: Hacer música<br>**11. La interpretación en grupo** | Goodman, Elaine | Alianza Música | |
| **2006** | 85 | La interpretación musical<br>Tercera parte: Hacer música<br>**12. El miedo escénico** | Valentin, Elizabeth | Alianza Música | |
| **2006** | 86 | La interpretación musical<br>Cuarta parte: Interpretar la interpretación.<br>**Escuchar la interpretación** | Clarke, Eric | Alianza Música | |
| **2006** | 87 | La interpretación musical<br>Cuarta parte: Interpretar la interpretación.<br>**14. El legado de las grabaciones** | Johnson, Peter | Alianza Música | Español |
| **2006** | 88 | La interpretación musical<br>**Cuarta parte: Interpretar la interpretación.**<br>**15. La crítica de la interpretación musical** | Monelle, Raimond | Alianza Música | |
| **2006** | 89 | La interpretación musical<br>Cuarta parte: Interpretar la interpretación.<br>**16. Los intérpretes y la interpretación** | Dunsby, Jonathan | Alianza Música | |
| **2005** | 90 | **Exploring the musical mind**<br>Cognitive processes<br>The psychology of music reading<br>**Structural cues within notation** | Sloboda, John | OUP | |
| **2005** | 91 | Exploring the musical mind<br>Cognitive processes<br>The psychology of music reading<br>**The role of musical knowledge in reading** | Sloboda, John | OUP | |
| **2005** | 92 | Exploring the musical mind<br>Cognitive processes<br>The psychology of music reading<br>**Teaching music reading** | Sloboda, John | OUP | |
| **2005** | 93 | Exploring the musical mind<br>Cognitive processes<br>Experimental studies of music reading: a review<br>**Review of early reading research** | Sloboda, John | OUP | |
| **2005** | 94 | Exploring the musical mind<br>Cognitive processes<br>Experimental studies of music reading: a review<br>**Experiments varying skill level and structure in text** | Sloboda, John | OUP | |
| **2005** | 95 | Exploring the musical mind<br>Cognitive processes<br>Experimental studies of music reading: a review<br>**Expressive sight-reading** | Sloboda, John | OUP | |
| **2005** | 96 | Exploring the musical mind<br>Cognitive processes<br>The uses of space in music notation<br>**Comparison of text and music notations** | Sloboda, John | OUP | |
| **2005** | 97 | Exploring the musical mind<br>Cognitive processes<br>The uses of space in music notation<br>**Historical trends in notation development** | Sloboda, John | OUP | |
| **2005** | 98 | Exploring the musical mind<br>Cognitive processes<br>The uses of space in music notation<br>**Conceptual issues in the design of notation** | Sloboda, John | OUP | |

| | | | | | |
|---|---|---|---|---|---|
| **2005** | 99 | Exploring the musical mind<br>Cognitive processes<br>The uses of space in music notation<br>**Psychological issues in the design of notation** | Sloboda, John | OUP | |
| **2005** | 100 | Exploring the musical mind<br>Cognitive processes<br>Immediate recall of melodies<br>**Problems of transcription from sung performance** | Sloboda, John | OUP | |
| **2005** | 101 | Exploring the musical mind<br>Cognitive processes<br>Immediate recall of melodies<br>**Analysis of errors in sung recall** | Sloboda, John | OUP | |
| **2005** | 102 | Exploring the musical mind<br>Cognitive processes<br>Immediate recall of melodies<br>**Memory representations for melodies** | Sloboda, John | OUP | |
| **2005** | 103 | Explorin the musical mind<br>Cognitive processes<br>Cognition and real music: the psychology of music comes of age<br>**Nature of scientific paradigms** | Sloboda, John | OUP | Inglés |
| **2005** | 104 | Exploring the musical mind<br>Cognitive processes<br>Cognition and real music: the psychology of music comes of age<br>**Is music psychology paradigmatic?** | Sloboda, John | OUP | |
| **2005** | 105 | Exploring the musical mind<br>Cognitive processes<br>Cognition and real music: the psychology of music comes of age<br>**The contribution of Lerdahl and Jackendoff** | Sloboda, John | OUP | |
| **2005** | 106 | Exploring the musical mind<br>Cognitive processes<br>Psychological structures in music: core research 1980-1990<br>**Citation as a means of assessing significance of research** | Sloboda, John | OUP | |
| **2005** | 107 | Exploring the musical mind<br>Cognitive processes<br>Psychological structures in music: core research 1980-1990<br>**The contribution of Krumhansl** | Sloboda, John | OUP | |
| **2005** | 108 | Exploring the musical mind<br>Cognitive processes<br>Psychological structures in music: core research 1980-1990<br>**Establishing a sense of tonal centre** | Sloboda, John | OUP | |
| **2005** | 109 | Exploring the musical mind<br>Cognitive processes<br>Psychological structures in music: core research 1980-1990<br>**Hierarchies in tonal representation** | Sloboda, John | OUP | |
| **2005** | 110 | Exploring the musical mind<br>Cognitive processes<br>Review of Language, Music and Mind by Diana Raffman<br>**The relationship of philosophy of music to psychology of music** | Sloboda, John | OUP | |
| **2005** | 111 | Exploring the musical mind<br>Cognitive processes<br>Review of Language, Music and Mind by Diana Raffman<br>**Nuance and ineffability in music** | Sloboda, John | OUP | |
| **2005** | 112 | Exploring the musical mind<br>Cognitive processes<br>Review of Language, Music and Mind by Diana Raffman<br>**Different ways in which music might be said to have 'meaning'** | Sloboda, John | OUP | |
| **2005** | 113 | Exploring the musical mind<br>Cognitive processes<br>Does music mean anything?<br>**Is parsing sufficient for understanding in music?** | Sloboda, John | OUP | |
| **2005** | 114 | Exploring the musical mind<br>Cognitive processes<br>Does music mean anything?<br>**Dynamic feelings as precursors of meaning** | Sloboda, John | OUP | |

| | | | | |
|---|---|---|---|---|
| **2005** | 115 | Exploring the musical mind<br>Cognitive processes<br>Does music mean anything?<br>***Thrills* as proto-emotions concerns** | Sloboda, John | OUP |
| **2005** | 116 | Exploring the musical mind<br>Emotion and motivation<br>Music as a language<br>**Musical phonology, syntax, and semantics** | Sloboda, John | OUP |
| **2005** | 117 | Exploring the musical mind<br>Emotion and motivation<br>Music as a language<br>**Peak musical experiences in childhood** | Sloboda, John | OUP |
| **2005** | 118 | Exploring the musical mind<br>Emotion and motivation<br>Music as a language<br>**Motivation for long-term commitment to music** | Sloboda, John | OUP |
| **2005** | 119 | Exploring the musical mind<br>Emotion and motivation<br>Music psychology and the composer<br>**Different models for composer-psychologist dialogue** | Sloboda, John | OUP |
| **2005** | 120 | Exploring the musical mind<br>Emotion and motivation<br>Music psychology and the composer<br>**Relationship of composition to functions of music** | Sloboda, John | OUP |
| **2005** | 121 | Exploring the musical mind<br>Emotion and motivation<br>Music psychology and the composer<br>**Music as an aid to cognitive restructuring** | Sloboda, John | OUP |
| **2005** | 122 | Exploring the musical mind<br>Music in the real world<br>Everyday uses of music listening: a preliminary study<br>**Music in public places** | Sloboda, John | OUP |
| **2005** | 123 | Exploring the musical mind<br>Emotion and motivation<br>Empirical studies of emotional response to music<br>**Verbal responses within forced categories** | Sloboda, John | OUP |
| **2005** | 124 | Exploring the musical mind<br>Emotion and motivation<br>Empirical studies of emotional response to music<br>**Extrinsic and intrinsic antecedents of musical emotion** | Sloboda, John | OUP |
| **2005** | 125 | Exploring the musical mind<br>Emotion and motivation<br>Emotional response to music: a review<br>**Accounting for variation between and within people** | Sloboda, John | OUP |
| **2005** | 126 | Exploring the musical mind<br>Emotion and motivation<br>Emotional response to music: a review<br>**Methods for measuring emotional response** | Sloboda, John | OUP |
| **2005** | 127 | Exploring the musical mind<br>Emotion and motivation<br>Emotional response to music: a review<br>**How musical events elicit emotion** | Sloboda, John | OUP |
| **2005** | 128 | Exploring the musical mind<br>Emotion and motivation<br>Musical performance and emotion: issues and developments<br>**Different expressive roles of performers** | Sloboda, John | OUP |
| **2005** | 129 | Exploring the musical mind<br>Emotion and motivation<br>Musical performance and emotion: issues and developments<br>**Connection between music structure and expressive devices** | Sloboda, John | OUP |
| **2005** | 130 | Exploring the musical mind<br>Emotion and motivation<br>Musical performance and emotion: issues and developments<br>**Evidence for deliberate planning of expressive communication** | Sloboda, John | OUP |

| | | | | |
|---|---|---|---|---|
| **2005** | 131 | Exploring the musical mind<br>Talent and skill development<br>Musical expertise<br>**Social and cultural relativity of musical expertise** | Sloboda, John | OUP |
| **2005** | 132 | Exploring the musical mind<br>Talent and skill development<br>Musical expertise<br>**Musical skill acquisition in non-instructional settings** | Sloboda, John | OUP |
| **2005** | 133 | Exploring the musical mind<br>Talent and skill development<br>Musical expertise<br>**Precursors of musical expertis** | Sloboda, John | OUP |
| **2005** | 134 | Exploring the musical mind<br>Music in the real world<br>The 'sound of music' versus the 'essence of music':<br>dilemmas for music-emotionresearchers<br>**Research which demonstrates the strong contribution of the individual to the n ature of the musical experience** | Sloboda, John | OUP |
| **2005** | 135 | Exploring the musical mind<br>Talent and skill development<br>Musical ability<br>**Expressive playing in sight-reading as a strong test of ability** | Sloboda, John | OUP |
| **2005** | 136 | Exploring the musical mind<br>Talent and skill development<br>Musical ability<br>**How to develop technical and expressive skill togethe** | Sloboda, John | OUP |
| **2005** | 137 | Exploring the musical mind<br>Talent and skill development<br>The acquisition of musical performance expertise:<br>deconstructing the talent' account of individual differences in musical expressivit y<br>**Is musical skill inherited?** | Sloboda, John | OUP |
| **2005** | 138 | Exploring the musical mind<br>Talent and skill development<br>The acquisition of musical performance expertise:<br>deconstructing the talent' account of individual differences in musical expressivit y<br>**The role of practice** | Sloboda, John | OUP |
| **2005** | 139 | Exploring the musical mind<br>Talent and skill development<br>The acquisition of musical performance expertise:<br>deconstructing the talent' account of individual differences in musical expressivit y<br>**Different learning mechanisms for technical and expressive skill** | Sloboda, John | OUP |
| **2005** | 140 | Exploring the musical mind<br>Talent and skill development<br>Are some children more gifted for music than others?<br>**The role of early learning in skill development** | Sloboda, John | OUP |
| **2005** | 141 | Exploring the musical mind<br>Talent and skill development<br>Are some children more gifted for music than others?<br>**Social support mechanisms for skill development** | Sloboda, John | OUP |
| **2005** | 142 | Exploring the musical mind<br>Talent and skill development<br>Are some children more gifted for music than others?<br>**Motivational factors predicting success or dropout** | Sloboda, John | OUP |
| **2005** | 143 | Exploring the musical mind<br>Music in the real world<br>Everyday uses of music listening: a preliminary study<br>**Importance of situation and volition in determining effects of music** | Sloboda, John | OUP |
| **2005** | 144 | Exploring the musical mind<br>Music in the real world<br>Everyday uses of music listening: a preliminary study<br>**Free verbal descriptions of music's role in everyday life** | Sloboda, John | OUP |

| | | | | |
|---|---|---|---|---|
| **2005** | 145 | Exploring the musical mind<br>Music in the real world<br>Music: where cognition and emotion meet<br>**Mismatch between music's value to people and their level of musical skill** | Sloboda, John | OUP |
| **2005** | 146 | Exploring the musical mind<br>Music in the real world<br>Music: where cognition and emotion meet<br>**Tracking everyday uses of music in real time** | Sloboda, John | OUP |
| **2005** | 147 | Exploring the musical mind<br>Music in the real world<br>Music: where cognition and emotion meet<br>**Societal barriers to engagement with music** | Sloboda, John | OUP |
| **2005** | 148 | Exploring the musical mind<br>Music in the real world<br>Music and worship: a psychologist's perspective<br>**Can music have assumed common effects within a defined situation?** | Sloboda, John | OUP |
| **2005** | 149 | Exploring the musical mind<br>Music in the real world<br>Music and worship: a psychologist's perspective<br>**Parallels between different ways of listening to music and different aspects of w orship** | Sloboda, John | OUP |
| **2005** | 150 | Exploring the musical mind<br>Music in the real world<br>Music and worship: a psychologist's perspective<br>**Ineffability as a core concept in both music and worship** | Sloboda, John | OUP |
| **2005** | 151 | Exploring the musical mind<br>Music in the real world<br>Emotion, functionality, and the everyday experience of music:<br>where does music education fit?<br>**Reasons why young adolescents give up instrumental engagement** | Sloboda, John | OUP |
| **2005** | 152 | Exploring the musical mind<br>Music in the real world<br>Emotion, functionality, and the everyday experience of music:<br>where does music education fit?<br>**Cultural trends associated with disengagement from formal music** | Sloboda, John | OUP |
| **2005** | 153 | Exploring the musical mind<br>Music in the real world<br>Emotion, functionality, and the everyday experience of music:<br>where does music education fit?<br>**Responses of the education system to cultural change** | Sloboda, John | OUP |
| **2005** | 154 | Exploring the musical mind<br>Music in the real world<br>The 'sound of music' versus the 'essence of music':<br>dilemmas for music-emotion researchers<br>**Cultural trends which deindividriate the musical experience** | Sloboda, John | OUP |
| **2005** | 155 | Exploring the musical mind<br>Music in the real world<br>The 'sound of music' versus the 'essence of music': dilemmas for music-emotion r esearchers<br>**How science can be misused in the interests of cultural homogenization of the musical experience** | Sloboda, John | OUP |
| **2005** | 156 | Exploring the musical mind<br>Music in the real world<br>Assessing music psychology research: values, priorities, and outcomes<br>**The social benefits of music psychology research psychology** | Sloboda, John | OUP |
| **2005** | 157 | Exploring the musical mind<br>Music in the real world<br>Assessing music psychology research: values, priorities, and outcomes<br>**Differentiating the above from the social benefits of music** | Sloboda, John | OUP |
| **2005** | 158 | Exploring the musical mind<br>Music in the real world<br>Assessing music psychology research: values, priorities, and outcomes<br>**Differing levels of engagement with social issues evident in research** | Sloboda, John | OUP |

| | | | | | |
|---|---|---|---|---|---|
| **2005** | 159 | Exploring the musical mind<br>Emotion and motivation<br>Empirical studies of emotional response to music<br>**Free verbal emotional responses to music** | Sloboda, John | OUP | |
| **2005** | 160 | Exploring the musical mind<br>Talent and skill development<br>Musical ability<br>**Cultural exposure as a primary source of basic ability** | Sloboda, John | OUP | |
| **2004** | 161 | **Musical Excelence**<br>Strategies and techniques to enhance performance | Williamond, Aaron | OUP | Inglés |
| **2004** | 162 | Musical excellence<br>Part I Prospects and limits<br>**A guide to enhancing musical performance** | Williamon, Aaron | OUP | |
| **2004** | 163 | Musical excellence<br>Part I Prospects and limits<br>**General perspectives on achieving Musical excellence** | Chaffi, Roger<br>Lemieux, Anthony F. | OUP | |
| **2004** | 164 | Musical excellence<br>Part I Prospects and limits<br>**Managing the physical demands of musical performance** | Wynn Parry, Christopher B. | OUP | |
| **2004** | 165 | Musical excellence<br>Part I Prospects and limits<br>**Measuring performance enhancement in music** | McPherson, Gary E.<br>Schubert, Emery | OUP | |
| **2004** | 166 | Musical excellence<br>Part II Practice strategies<br>**Strategies for individual practice** | Jorgensen, Harald | OUP | |
| **2004** | 167 | Musical excellence<br>Part II Practice strategies<br>**Strategies for ensemble practice** | Davidson, Jane W.<br>C. King, Elaine | OUP | |
| **2004** | 168 | Musical excellence<br>Part II Practice strategies<br>**Strategies for memorizing music** | Ginsborg, Jane | OUP | |
| **2004** | 169 | Musical excellence<br>Part II Practice strategies<br>**Strategies for sight-reading and improvising music** | Thompson and , Sam<br>Lehmann, Andreas C. | OUP | |
| **2004** | 170 | Musical excellence<br>Part III Techniques and interventions<br>**Physical fitness** | H. Taylor, Adrian<br>Wasley, David | OUP | |
| **2004** | 171 | Musical excellence<br>Part III Techniques and interventions<br>**Alexander technique** | Valentine, Elizabeth | OUP | |
| **2004** | 172 | Musical excellence<br>Part III Techniques and interventions<br>**Mental skills training** | Connolly , Christopher<br>Williamon, Aaron | OUP | |
| **2004** | 173 | Musical excellence<br>Part III Techniques and interventions<br>**Feedback learning of musical expressivity** | Colectivo | OUP | |
| **2004** | 174 | Musical excellence<br>Part III Techniques and interventions<br>**Drugs and musical performance** | West, Robert | OUP | |
| **2004** | 175 | Musical excellence<br>Part III Techniques and interventions<br>**Epilogue: A note on future directions for enhancing musical performance** | Williamon, Aaron | OUP | |
| **2003** | 176 | **The Cognitive Neuroscience of Music**<br>Part I: The origins of music<br>**Musical predisposition in infancy: an update** | Trehub: , Sandra E. | OUP | |
| **2003** | 177 | The Cognitive Neuroscience of Music<br>Part I: The origins of music<br>**The quest for universals in temporal processing in music** | Drake, Carolyn<br>Bertrand, Daisy | OUP | |
| **2003** | 178 | The Cognitive Neuroscience of Music<br>Part I: The origins of music<br>**Mechanisms of musical memory in infancy** | Colectivo | OUP | |
| **2003** | 179 | The Cognitive Neuroscience of Music<br>Part I: The origins of music<br>**Music cognition, culture, and evolution** | Cross, Ian | OUP | |

| | | | | | |
|---|---|---|---|---|---|
| **2003** | 180 | The Cognitive Neuroscience of Music<br>Part I: The origins of music<br>**Is music an evolutionary adaptation?** | Huron, David | OUP | Inglés |
| **2003** | 181 | The Cognitive Neuroscience of Music<br>Part II: The musical mind<br>**The roots of musical variation in perceptual similarity and invariance** | McAdams, Stephen<br>Matzkin, Daniel | OUP | |
| **2003** | 182 | The Cognitive Neuroscience of Music<br>Part II: The musical mind<br>**Tonal cognition** | Krumhansl , Carol L.<br>Toivainen, Petri | OUP | |
| **2003** | 183 | The Cognitive Neuroscience of Music<br>Part II: The musical mind<br>**Learning and perceiving musical structures: further insights from artificial neur al networks** | Colectivo | OUP | |
| **2003** | 184 | The Cognitive Neuroscience of Music<br>Part III: The neurons of music<br>**Neurobiology of harmony perception** | Tramo, Mark | OUP | |
| **2003** | 185 | The Cognitive Neuroscience of Music<br>Part III: The neurons of music<br>**Intracerebral evoked potentials in pitch perception reveal a functional asymmet ry of human auditory cortex** | Colectivo | OUP | |
| **2003** | 186 | The Cognitive Neuroscience of Music<br>Part III: The neurons of music<br>**The neural processing of complex sounds** | Griffiths: , Timothy D. | OUP | |
| **2003** | 187 | The Cognitive Neuroscience of Music<br>Part IV: Musical brain substrates<br>**Music and the neurologist: an historical perspective** | Brust, John C.M. | OUP | |
| **2003** | 188 | The Cognitive Neuroscience of Music<br>Part IV: Musical brain substrates<br>**Brain specialization for music: new evidence from congenital amusia** | Peretz, Isabelle | OUP | |
| **2003** | 189 | The Cognitive Neuroscience of Music<br>Part IV: Musical brain substrates<br>**Cerebral substrates for musical temporal processes** | Samson, Severine | OUP | |
| **2003** | 190 | The Cognitive Neuroscience of Music<br>Part IV: Musical brain substrates<br>**Cerebral substrates of musical imagery** | Halpern: , Andrea R. | OUP | |
| **2003** | 191 | The Cognitive Neuroscience of Music<br>Part IV: Musical brain substrates<br>**Neural specializations for tonal processing** | Zatorre, Robert J. | OUP | |
| **2003** | 192 | The Cognitive Neuroscience of Music<br>Part IV: Musical brain substrates<br>**Exploring the functional neuroanatomy of music performance perception and c omprehension** | Parsons, Lawrence M. | OUP | |
| **2003** | 193 | The Cognitive Neuroscience of Music<br>Part IV: Musical brain substrates<br>**Comparison between language and music** | MBesson , Mireille<br>Schon, Daniele | OUP | |
| **2003** | 194 | The Cognitive Neuroscience of Music<br>Part IV: Musical brain substrates<br>**Musical sound processing: EEG and MEG evidence** | Tervaniemi, Mari | OUP | |
| **2003** | 195 | The Cognitive Neuroscience of Music<br>Part IV: Musical brain substrates<br>**A new approach to the cognitive neuroscience of melody** | Trainor, Laurel | OUP | |
| **2003** | 196 | The Cognitive Neuroscience of Music<br>Part IV: Musical brain substrates<br>**Processing emotions induced by music** | Patel, Aniruddh D.<br>Balaban, Evan | OUP | |
| **2003** | 197 | The Cognitive Neuroscience of Music<br>Part IV: Musical brain substrates<br>**How many music centres are in the brain** | Altenmuller, Eckart O. | OUP | |
| **2003** | 198 | The Cognitive Neuroscience of Music<br>Part V: Musical brain/brain plasticity<br>**Functional organization and plasticity of auditory cortex** | Rauschecker, Joseph P. | OUP | |
| **2003** | 199 | The Cognitive Neuroscience of Music<br>Part V: Musical brain/brain plasticity<br>**The brain of musicians** | Schlaug, Gottfried<br>Chen, Chi | OUP | |

| | | | | | |
|---|---|---|---|---|---|
| **2003** | 200 | The Cognitive Neuroscience of Music<br>Part V: Musical brain/brain plasticity<br>**Representational cortex in musicians** | Colectivo | OUP | |
| **2003** | 201 | The Cognitive Neuroscience of Music<br>Part V: Musical brain/brain plasticity<br>**The brain that makes music and is changed by it** | Pascual-Leone, Alvaro | OUP | |
| **2003** | 202 | The Cognitive Neuroscience of Music<br>Part V: Musical brain/brain plasticity<br>**The sounds of poetry viewed as music** | Lerdahl, Fred | OUP | |
| **2003** | 203 | The Cognitive Neuroscience of Music<br>Part V: Musical brain/brain plasticity<br>**Does exposure to music have beneficial side effects?** | Shellenberg, Glenn | OUP | |
| **2002** | 204 | **The Science and Psychology of Music Performance**<br>Part 1: The Developing Musician<br>1 **Music Potential** | Kemp, Anthony E.<br>Mills, Janet | OUP | Inglés |
| **2002** | 205 | The Science and Psychology of Music Performance<br>Part 1: The Developing Musician<br>2 **Environmental Influences** | Gembris, Heiner<br>Davidson, Jane W. | OUP | |
| **2002** | 206 | The Science and Psychology of Music Performance<br>Part 1: The Developing Musician<br>3 **Motivation** | O'Neill, Susan A.<br>McPherson, Gary E. | OUP | |
| **2002** | 207 | The Science and Psychology of Music Performance<br>Part 1: The Developing Musician<br>4 **Performance Anxiety** | Wilson, Glenn D.<br>Roland, David | OUP | |
| **2002** | 208 | The Science and Psychology of Music Performance<br>Part 1: The Developing Musician<br>5 **Brain Mechanisms** | Altenmuller, Eckart<br>Gruhn, Wilfried | OUP | |
| **2002** | 209 | The Science and Psychology of Music Performance<br>Part 1: The Developing Musician<br>6 **Music Medicine** | Brandfonbrener, Alice G.<br>Kjelland, James M. | OUP | |
| **2002** | 210 | The Science and Psychology of Music Performance<br>Part 2: Subskills of Music Performance<br>7 **From Sound to Sign** | McPherson, Gary E.<br>Gabrielsson, Alf | OUP | |
| **2002** | 211 | The Science and Psychology of Music Performance<br>Part 2: Subskills of Music Performance<br>8 **Improvisation** | Kenny, Barry J.<br>Gellrich, Martin | OUP | |
| **2002** | 212 | The Science and Psychology of Music Performance<br>Part 2: Subskills of Music Performance<br>9 **Sight-Reading** | Lehmann, Andreas C.<br>McArthur, Victoria | OUP | |
| **2002** | 213 | The Science and Psychology of Music Performance<br>Part 2: Subskills of Music Performance<br>10 **Practice** | Barry, Nancy H.<br>Hallam, Susan | OUP | |
| **2002** | 214 | The Science and Psychology of Music Performance<br>Part 2: Subskills of Music Performance<br>11 **Memory** | Aiello, Rita<br>Williamon, Aaron | OUP | |
| **2002** | 215 | The Science and Psychology of Music Performance<br>Part 2: Subskills of Music Performance<br>12 **Intonation** | Morrison, Steven J.<br>Fyk, Janina | OUP | |
| **2002** | 216 | The Science and Psychology of Music Performance<br>Part 2: Subskills of Music Performance<br>13 **Structural Communication** | Friberg, Anders<br>Umberto Battel, Giovanni | OUP | |
| **2002** | 217 | The Science and Psychology of Music Performance<br>Part 2: Subskills of Music Performance<br>14 **Emotional Communication** | Juslin, Patrik<br>Persson, Roland | OUP | |
| **2002** | 218 | The Science and Psychology of Music Performance<br>Part 2: Subskills of Music Performance<br>15 **Body Movement** | Davidson, Jane W.<br>Salgado Correia, Jorge | OUP | |
| **2002** | 219 | The Science and Psychology of Music Performance<br>Part 3: Instruments and Ensembles<br>16 **Solo Voice** | Welch, Graham F.<br>Sundberg, Johan | OUP | |

| | | | | | |
|---|---|---|---|---|---|
| **2002** | 220 | The Science and Psychology of Music Performance<br>Part 3: Instruments and Ensembles<br>17 **Choir** | Ternström, Sten<br>Richard Karna, Duane | OUP | |
| **2002** | 221 | The Science and Psychology of Music Performance<br>Part 3: Instruments and Ensembles<br>18 **Piano** | Parncutt, Richard<br>Troup, Malcom | OUP | |
| **2002** | 222 | The Science and Psychology of Music Performance<br>Part 3: Instruments and Ensembles<br>19 **String Instruments** | Guettler, Knut<br>Hallam, Susan | OUP | |
| **2002** | 223 | The Science and Psychology of Music Performance<br>Part 3: Instruments and Ensembles<br>20 **Wind Instruments** | Fuks, Leonardo<br>Fadle, Heinz | OUP | |
| **2002** | 224 | The Science and Psychology of Music Performance<br>Part 3: Instruments and Ensembles<br>21 **Rehearsing and Conducting** | Price, Harry E.<br>Byo, James L. | OUP | Inglés |
| **2002** | 225 | **Conceptualizing Music**<br>PARTII. ANALYSIS AND THEORY<br>4. Categorization, Compositional Strategy,and Musical Syntax | Zbikowsk, Lawrence M. | OUP | Inglés |
| **2002** | 226 | **Conceptualizing Music**<br>PART I. ASPECTS OF COGNITIVE STRUCTURE<br>1 Categorization | Zbikowsk, Lawrence M. | OUP | Inglés |
| **2002** | 227 | **Conceptualizing Music**<br>PART I. ASPECTS OF COGNITIVE STRUCTURE<br>2. Cross-Domain Mapping | Zbikowsk, Lawrence M. | OUP | Inglés |
| **2002** | 228 | **Conceptualizing Music**<br>PART I. ASPECTS OF COGNITIVE STRUCTURE<br>3. Conceptual Models and Theories | Zbikowsk, Lawrence M. | OUP | Inglés |
| **2002** | 229 | **Conceptualizing Music**<br>PARTII. ANALYSIS AND THEORY<br>5. Cultural Knowledge and Musical Ontology | Zbikowsk, Lawrence M. | OUP | Inglés |
| **2002** | 230 | **Conceptualizing Music**<br>PARTII. ANALYSIS AND THEORY<br>6.Words, Music, and Song: The Nineteenth-Century Lied | Zbikowsk, Lawrence M. | OUP | Inglés |
| **2002** | 231 | **Conceptualizing Music**<br>PARTII. ANALYSIS AND THEORY<br>7. Competing Models of Music: Theories of Musical Form and Hierarchy | Zbikowsk, Lawrence M. | OUP | Inglés |
| **2001** | 232 | **Music and emotion**<br>Introduction<br>**1 Music and emotion: introduction** | Juslin, Patrik N.<br>Sloboda, John A. | OUP | |
| **2001** | 233 | Music and emotion<br>Multidisciplinary Perspectives<br>**2 Philosophical perspectives on music's expressiveness** | Davies, Stephen | OUP | |
| **2001** | 234 | Music and emotion<br>Multidisciplinary Perspectives<br>**3 Musicological approaches to emotion** | Cook, Nicholas<br>Dibben, Nicola | OUP | |
| **2001** | 235 | Music and emotion<br>Multidisciplinary Perspectives<br>**4 Psychological perspectives on music and emotion** | Sloboda, John A.<br>Juslin, Patrik N. | OUP | |
| **2001** | 236 | Music and emotion<br>Multidisciplinary Perspectives<br>**5 Listen to the brain: a biological perspective on musical emotions** | Peretz, Isabelle | OUP | |
| **2001** | 237 | Music and emotion<br>Multidisciplinary Perspectives<br>**6 Anthropological perspectives on music and emotion** | Becker, Judith | OUP | |
| **2001** | 238 | Music and emotion<br>Multidisciplinary Perspectives<br>**7 Aesthetic agency and musical practice: new directions in the sociology of musi c end emotion** | DeNora, Tia | OUP | |
| **2001** | 239 | Music and emotion<br>Multidisciplinary Perspectives<br>**8 Music and emotion: perspectives from music therapy** | Bunt, Leslie<br>Pavlicevic, Mercedes | OUP | |

| | | | | | |
|---|---|---|---|---|---|
| **2001** | 240 | Music and emotion<br>The Composer<br>**9 Emotion and composition in classical music: historiometric perspectives** | Simonton, Dean Keith | OUP | |
| **2001** | 241 | Music and emotion<br>The Composer<br>**10 The influence of musical structure on emotional expression** | Gabrielsson, Alf<br>Lindstrom, Erik | OUP | |
| **2001** | 242 | Music and emotion<br>The Composer<br>**11 Music as a source of emotion in film** | Cohen, Annabel J. | OUP | |
| **2001** | 243 | Music and emotion<br>The Performer<br>**12 The subjective world of the performer** | Persson, Roland S. | OUP | |
| **2001** | 244 | Music and emotion<br>The Performer<br>**13 Negative emotions in music making: the problem of performance anxiety** | Steptoe, Andrew | OUP | |
| **2001** | 245 | Music and emotion<br>The Performer<br>**14 Communicating emotion in music performance: a review and a theoretical f ramework** | Juslin, Patrik N. | OUP | |
| **2001** | 246 | Music and Emotion<br>The Listener<br>**15 Music and emotion: distinctions and uncertainties** | Meyer, Leonard B. | OUP | |
| **2001** | 247 | Music and emotion<br>The Listener<br>**16 Emotional effects of music: production rules** | Scherer, Klaus R.<br>Zentner, Marcel R. | OUP | |
| **2001** | 248 | Music and emotion<br>The Listener<br>**17 Continuous measurement of self-report emotional response to music** | Schubert, Emery | OUP | |
| **2001** | 249 | Music and emotion<br>The Listener<br>**18 Emotions in everyday listening to music** | Sloboda , John A.<br>O'Neill, Susan A. | OUP | |
| **2001** | 250 | Music and emotion<br>The Listener<br>**19 Emotions in strong experiences with music** | Gabrielsson, Alf | OUP | |
| **2001** | 251 | Music and emotion<br>Postlude<br>**20 Music and emotion: commentary** | Sloboda, John A.<br>Juslin, Patrik N. | OUP | |
| **2000 (1988)** | 252 | **Generative processes in music**<br>**Cognitive constraints on compositional systems** | Lerdahl, Fred | OUP | Inglés |
| **2000 (1988)** | 253 | Generative processes in music<br>**Computer synthesis of music performance** | Sundberg, Johan | OUP | |
| **2000 (1988)** | 254 | Generative processes in music<br>**Experimental research into musical generative ability** | Sagi, Maria<br>Vitanyi, Ivan | OUP | |
| **2000 (1988)** | 255 | Generative processes in music<br>**From collections to structure: the developmental path of tonal thinking** | Davidson, Lyle<br>Welsh, Patricia | OUP | |
| **2000 (1988)** | 256 | Generative processes in music<br>**Generative principles in music performance** | Clarke, Eric F. | OUP | |
| **2000 (1988)** | 257 | Generative processes in music<br>**Improvisation: methods and models** | Pressing, Jeff | OUP | |
| **2000 (1988)** | 258 | Generative processes in music<br>**Rehearsal skill and musical competence: does practice make perfect?** | Gruson, Linda M. | OUP | |
| **2000 (1988)** | 259 | Generative Processes in Music<br>**The Psychology of Performance, Improvisation, and Composition** | Colectivo | OUP | |
| **2000 (1988)** | 260 | Generative processes in music<br>**Timing and synchronisation in ensemble performance** | Rasch, Rudolf A. | OUP | |
| **2000 (1988)** | 261 | Generative processes in music<br>**Timing in music performance and its relations to music experience** | Gabrielsson, Alf | OUP | |
| **2000 (1988)** | 262 | Generative processes in music<br>**Tonal structure and children's early learning of music** | Dowling, W. Jay | OUP | |
| **2000 (1988)** | 263 | Generative processes in music<br>**Young children's musical representations: windows on music cognition** | Davidson, Lyle<br>Scripp, Lawrence | OUP | |

| | | | | | |
|---|---|---|---|---|---|
| 1994 | 264 | **Musical Perceptions**<br>Philosophical Perspectives<br>**1. Emotion and Meaning in Music** | Meyer, Leonard B. | OUP | |
| 1994 | 265 | Musical Perceptions<br>**Developmental Perspectives**<br>**4. Songsinging by Young and Old: A Developmental Approach to Music** | Davidson, Lyle | OUP | Inglés |
| 1994 | 266 | Musical Perceptions<br>The Perception of Melody, Tonality, Rhythm And Timing<br>**7. Melodic Contour in Hearing and Remembering Melodies** | Dowling, W. Jay | OUP | |
| 1994 | 267 | Musical Perceptions<br>The perception of musical compositions<br>**12. Can Listening to Music Be Experimentally Studied?** | Aiello, Rita | OUP | |
| 1994 | 268 | Musical Perceptions<br>Developmental Perspectives<br>**5. Coming to Hear in a New Way** | Bamberger, Jeanne | OUP | |
| 1994 | 269 | Musical Perceptions<br>Developmental Perspectives<br>**6. Music Performance: Expression and the Development of Excellence** | Sloboda, John A. | OUP | |
| 1994 | 270 | Musical Perceptions<br>Philosophical Perspectives<br>**2. Music and Language: Parallels and Contrasts** | Aiello, Rita | OUP | |
| 1994 | 271 | Musical Perceptions<br>Philosophical Perspectives<br>**3. Perception: A Perspective from Music Theory** | Cook, Nicholas | OUP | |
| 1994 | 272 | Musical Perceptions<br>The Perception of Melody, Tonality, Rhythm And Timing<br>**8. Describing the Mental Representation of Tonality in Music** | Butler, David<br>Brown, Helen | OUP | |
| 1994 | 273 | Musical Perceptions<br>The Perception of Melody, Tonality, Rhythm And Timing<br>**9. Tonality and Expectation** | Bharucha, Jamshed J. | OUP | |
| 1994 | 274 | Musical Perceptions<br>The Perception of Melody, Tonality, Rhythm And Timing<br>**10. Perception, Production, and Imitation of Time Ratios by Skilled Musicians** | Sternberg, Saul<br>Knoll, Ronald L. | OUP | |
| 1994 | 275 | Musical Perceptions<br>The Perception of Melody, Tonality, Rhythm And Timing<br>**11. The Interpretive Component in Musical Performance** | Shaffer, L. H.<br>McAngus Todd, Neil P. | OUP | |
| 1986/97 | 276 | **La inteligencia hábil**<br>El desarrollo de las capacidades cognitivas<br>**1 La naturaleza de las habilidades**<br>Cognición y psicología<br>¿Qué es una habilidad?<br>La adquisición de la habilidad<br>**La lectura: un estudio de caso de habilidades cognitivas.** | Colectivo | AIQUE | Español |
| 1986/97 | 277 | La inteligencia hábil<br>El desarrollo de las capacidades cognitivas<br>**2 Habilidades de la memoria**<br>La naturaleza de la memoria.<br>Adquisición de habilidades de la memoria<br>Habilidades memorísticas excepcionales<br>**¿Cómo pueden mejorarse las habilidades memorísticas?** | Colectivo | AIQUE | |
| 1986/97 | 278 | La inteligencia hábil<br>El desarrollo de las capacidades cognitivas<br>**3 Habilidades de la comunicación y la comprensión**<br>Las habilidades del hablante<br>Habilidades del oyente<br>Habilidades sociales<br>**¿cómo mejorar las habilidades de aprendizaje?** | Colectivo | AIQUE | |
| 1986/97 | 279 | La inteligencia hábil<br>El desarrollo de las capacidades cognitivas<br>**4 Pensamiento y habilidades**<br>La habilidad en el razonamiento<br>La resolución de problemas<br>La creatividad y las habilidades cognitivas<br>**Las computadoras y la cognición.** | Colectivo | AIQUE | |

| | | | | | |
|---|---|---|---|---|---|
| **1985 (2000)** | 280 | **The musical mind**<br>**1. MUSIC AS A COGNITIVE SKILL**<br>Introduction<br>Chomsky and Schenker<br>Other comparisons between language and music<br>Musical phonology<br>Musical syntax<br>Musical semantics<br>Conclusions | Sloboda, John A. | OUP | Inglés |
| **1985 (2000)** | 281 | The musical mind<br>**2. MUSIC, LANGUAGE, AND MEANING**<br>Introduction<br>Chomsky and Schenker<br>Other comparisons between language and music<br>Musical phonology<br>Musical syntax<br>Musical semantics<br>Conclusions | Sloboda, John A. | OUP | |
| **1985 (2000)** | 282 | The musical mind<br>**3. THE PERFORMANCE OF MUSIC**<br>Introduction<br>Sight reading<br>Rehearsal<br>Expert performance | Sloboda, John A. | OUP | |
| **1985 (2000)** | 283 | The musical mind<br>**4. COMPOSITION AND IMPROVISATION**<br>Introduction<br>Manuscript evidence<br>Composers' writings on their own compositional process<br>Observations of composers at work<br>Improvisation | Sloboda, John A. | OUP | |
| **1985 (2000)** | 284 | The musical mind<br>**5. LISTENING TO MUSIC**<br>Introduction<br>Natural hearing': primitive grouping mechanisms in music<br>Attention in music listening<br>Memory in music listening | Sloboda, John A. | OUP | |
| **1985 (2000)** | 285 | The musical mind<br>**6. MUSICAL LEARNING AND DEVELOPMENT**<br>Introduction<br>Musical enculturation<br>Training and skill acquisition<br>Assessing musical ability | Sloboda, John A. | OUP | |
| **1985 (2000)** | 286 | The musical mind<br>**7. THE MUSICAL MIND IN CONTEXT: CULTURE AND BIOLOGY**<br>Introduction<br>Culture and musical thinking<br>Biology and musical thinking | Sloboda, John A. | OUP | |
| **1982** | 287 | **The Psychology of the Music**<br>**1. The Perception of Musical Ton**<br>*R. A. Rasch and R. Plomp*<br>The Psychoacoustics of Music<br>Perceptual Attributes of Single Tones<br>Perceptual Attributes of Simultaneous Tones<br>Conclusion<br>References | Rasch, R. A.<br>Plomp, R. | Academic Press | Inglés |
| **1982** | 288 | The Psychology of the Music<br>**2. Exploration of Timbre by Analysis and Synthesis**<br>*Jean-Claude Risset and David L. Wessel*<br>Timbre<br>Timbre and the Fourier Spectrum: The Classical View<br>The Shortcomings of the Classical Conception<br>Attack Transients<br>Complexity of Sounds: Importance of Characteristic Features<br>Instrumental and Vocal Timbres: Additive Synthesis<br>Cross Synthesis and Voice Synthesis<br>Additive Synthesis: Percussion Instruments | Risset, Jean-Claude<br>Wessel, David L. | Academic Press | |

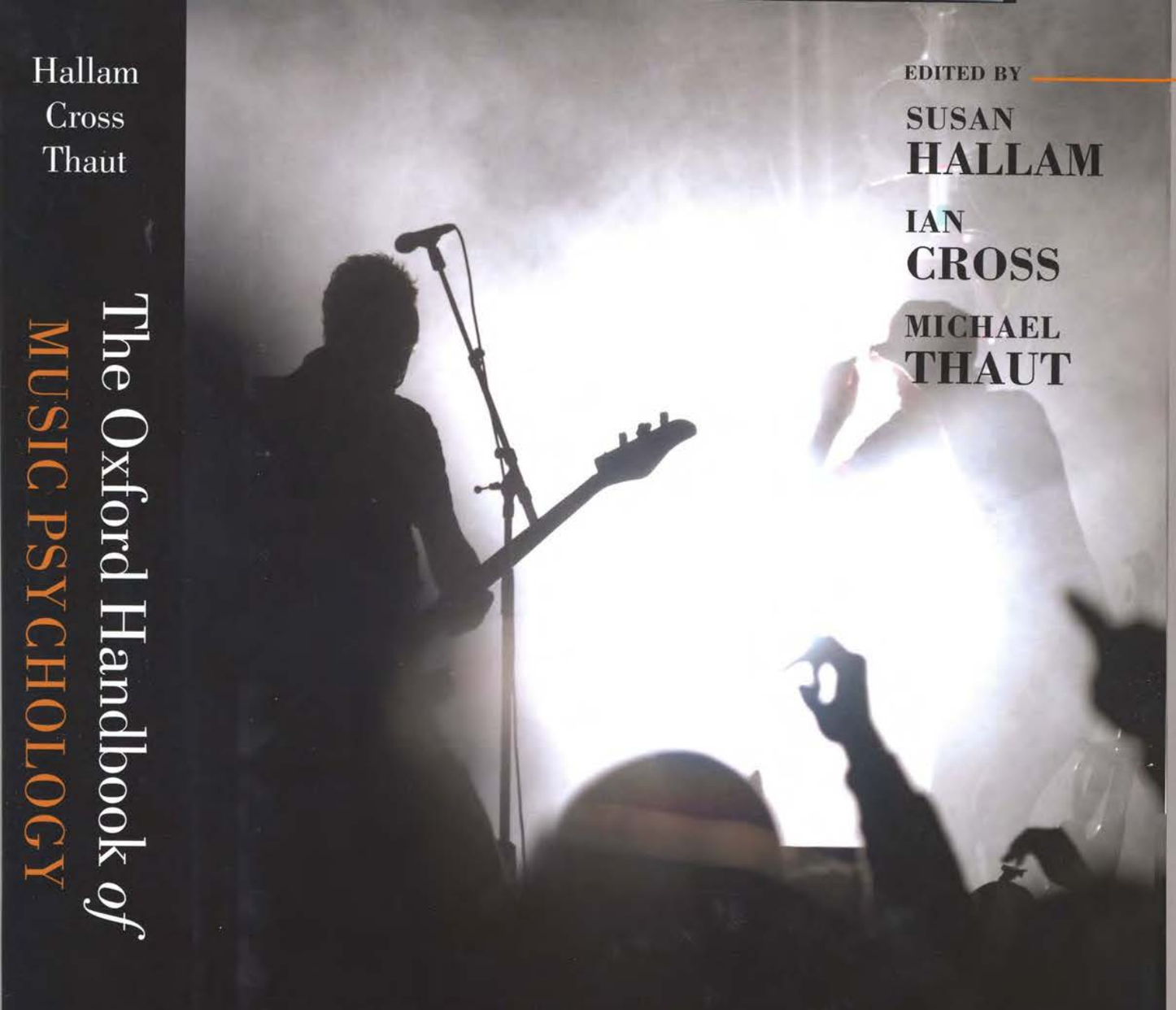
| | | | | |
|---|---|---|---|---|
| **1982** | 289 | The Psychology of the Music<br>**3. Perception of Singing**<br>*Johan Sundberg*<br>Introduction<br>Function of the Voice<br>Resonatory Aspects<br>Phonation<br>Vibrato<br>Pitch Accuracy in Singing Practice<br>Phrasing and Emotion<br>Concluding Remarks | Sundberg, Johan | Academic Press |
| **1982** | 290 | The Psychology of the Music<br>**4. Grouping Mechanisms in Music**<br>*Diana Deutsch*<br>Introduction<br>Grouping Principles<br>Two-Channel Listening to Melodic Sequences<br>Channeling of Rapid Sequences of Single Tones<br>Voluntary Attention<br>Conclusion<br>References | Deutsch, Diana | Academic Press |
| **1982** | 291 | The Psychology of the Music<br>**6. Rhythm and Tempo**<br>*Paul Fraisse*<br>Definitions<br>Rhythm and Spontaneous Tempo<br>Rhythmic Forms<br>The Perception of Musical Rhythms<br>Conclusion<br>References | Fraisse, Paul | Academic Press |
| **1982** | 292 | The Psychology of the Music<br>**5. The Listener and the Acoustic Environment**<br>*R. A. Rasch and R. Plomp*<br>Introduction<br>Methodology<br>Level Effects of Indirect Sound: Loudness<br>Temporal Effects of Indirect Sound: Definition<br>Spatial Effects of Indirect Sound: Spaciousness<br>The Compromise between Definition and Spaciousness<br>Conclusion | Rasch, R. A. Rasch<br>Plomp, R. | Academic Press |
| **1982** | 293 | The Psychology of the Music<br>**7. Timing by Skilled Musicians**<br>*Saul Sternberg, Ronald L. Knoll, and Paul Zukofsky*<br>Perception, Production, and Imitation of Fractions of the Beat<br>Perceptual judgment of Beat Fractions<br>Production of Beat Fractions<br>Imitation of Beat Fractions<br>A Shared-Process Model of the Perception, Production, and Imitation of Beat Fra ctions<br>Further Analysis of Perceptual judgment | Colectivo | Academic Press |
| **1982** | 294 | The Psychology of the Music<br>**8. Intervals, Scales, and Tuning**<br>*Edward M. Burns and W. Dixon Ward*<br>Introduction<br>Are Scales Necessary?<br>Musical Interval Perception<br>Natural Intervals and Scales<br>Conclusions and Caveats<br>References | Burns, Edward M.<br>Ward, W. Dixon | Academic Press |
| **1982** | 295 | The Psychology of the Music<br>**9. The Processing of Pitch Combinations**<br>*Diana Deutsch*<br>Introduction<br>Feature Abstraction<br>Higher Order Abstractions<br>Alphabets and Hierarchies<br>Memory Systems<br>Conclusions<br>References | Deutsch, Diana | Academic Press |

| | | | | |
|---|---|---|---|---|
| **1982** | 296 | The Psychology of the Music<br>**10. Melodic Processes and the Perception of Music**<br>*Burton S. Rosner and Leonard B. Meyer*<br>The Perception and Classification of Two Archetypal Melodic Processes<br>Experimental Findings<br>Implications<br>References | Rosner, Burton S.<br>Meyer, Leonard B. | Academic Press |
| **1982** | 297 | The Psychology of the Music<br>**11. Structural Representations of Musical Pitch**<br>*Roger N. Shepard*<br>Introduction<br>Unidimensional Approaches to Pitch<br>Potentially Multidimensional Approaches to Pitch<br>The Spatial Representation of Pitch<br>Illustrative Analyses of Empirical Data<br>Discussion<br>References | Shepard, Roger N. | Academic Press |
| **1982** | 298 | The Psychology of the Music<br>**12. Musical Ability**<br>*Rosamund Shuter-Dyson*<br>Concepts of Musical Ability<br>Correlational and Factorial Studies of Musical Ability<br>Musical Ability and Other Intellectual Abilities<br>References | Shuter-Dyson, Rosamund | Academic Press |
| **1982** | 299 | The Psychology of the Music<br>**13. Melodic Information Processing and Its Development**<br>*W. Jay Dowling*<br>Introduction<br>Development<br>Adult Memory<br>Contour versus Interval<br>Summary<br>References | Dowling, W. Jay | Academic Press |
| **1982** | 300 | The Psychology of the Music<br>**14. Absolute Pitch**<br>*W. Dixon Ward and Edward M. Burns*<br>Introduction<br>Genesis of AP<br>Measurement of AP<br>Stability of the Internal Standard<br>Learning AP<br>The Value of AP<br>References | Ward, W. Dixon<br>Burns, Edward M. | Academic Press |
| **1982** | 301 | The Psychology of the Music<br>**15. Neurological Aspects of Music Perception and Performance**<br>*Oscar S. M. Marin*<br>Amusia<br>Auditory Agnosia and Verbal Deafness<br>General Comments<br>References | Marin, Oscar S. M. | Academic Press |
| **1982** | 302 | The Psychology of the Music<br>**16. Music Performance**<br>*John A. Sloboda*<br>Introduction<br>The Nature of Performance Plans<br>Acquisition of Performance Plans<br>The Role of Feedback in Performance<br>Social Factors in Performance<br>Summary<br>References | Sloboda, John A. | Academic Press |
| **1982** | 303 | The Psychology of the Music<br>**17. Social Interaction and Musical Preference**<br>*Vladimir J. Konecni*<br>Introduction<br>Effects of Social Stimulation on Aesthetic Choice<br>Effects of Information Load and Arousing Nonsocial Stimulation on Aesthetic C hoice<br>Effects of Listening to Melodies Differing in Complexity on Emotional States an d Social Behavior<br>Listeners' Sequencing and "Chunking" of Musical Materials and the Use of Musi | Konecni, Vladimir J. | Academic Press |

**1982** 304 The Psychology of the Music Erickson, Robert Academic Press
**18. New Music and Psychology**
*Robert Erickson*
Introduction
Music Theory and Music
Understanding Tonality
Music and Perceptual Streaming
Fused Sounds in Music
Music Theory and Experimental Science
References

Hallam
Cross
Thaut

The Oxford Handbook *of*
MUSIC PSYCHOLOGY



EDITED BY
SUSAN
HALLAM
IAN
CROSS
MICHAEL
THAUT

# The Oxford Handbook *of* MUSIC PSYCHOLOGY



OXFORD

The field of music psychology has grown dramatically in the past 20 years, to emerge as one of the mainstream topics within the brain sciences. However, until now, there has been no comprehensive international reference text in the field.

The *Oxford Handbook of Music Psychology* is a landmark text providing, for the first time ever in English, a comprehensive overview of the latest developments in this fast-growing area of research. With contributions from over 50 experts in the field, the range and depth of coverage is unequalled. All the chapters combine a solid review of the relevant literature with well-reasoned arguments and robust discussions of the major findings, as well as original insights and suggestions for future work.

Written by leading experts, the 52 chapters are divided into 11 parts covering both experimental and theoretical perspectives, each edited by an internationally recognized authority on the area.

The first 10 parts each present chapters that focus on specific areas of music psychology:

- the origins and functions of music
- music perception
- responses to music
- music and the brain
- musical development
- learning musical skills
- musical performance
- composition and improvisation
- the role of music in our everyday lives
- music therapy

In each part, expert authors critically review the literature, highlight current issues, and explore possibilities for the future.

The final part examines how, in recent years, the study of music psychology has broadened to include a range of other disciplines. It considers the way that the research has developed in relation to technological advances, fostering links across the field and providing an overview of the areas where the field needs further development in the future.

The *Oxford Handbook of Music Psychology* is the essential reference text for students and researchers across psychology and neuroscience interested in music.

# The Oxford Handbook of Music Psychology

# The Oxford Handbook of Music Psychology

Edited by
**Susan Hallam,**
**Ian Cross**
and
**Michael Thaut**

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP
Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in
Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto
With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam
Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries
Published in the United States
by Oxford University Press Inc., New York
© Oxford University Press, 2009
The moral rights of the authors have been asserted
Database right Oxford University Press (maker)
First published 2009
All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above
You must not circulate this book in any other binding or cover
and you must impose the same condition on any acquirer
British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging-in-Publication Data
The Oxford handbook of music psychology/edied by Susan Hallam, Ian Cross, and Michael Thaut
p. cm
Includes bibliographical references.
ISBN 978-0-19-929845-7
1. Music—Psychological aspects. 2. Music therapy. I. Hallam, Susan. II. Cross, Ian. III. Thaut, Michael, 1952-
ML3830.09 2009
781'.11—dc22
2008046393

Typeset by Cepha Imaging Private Ltd., Bangalore, India
Printed in the United Kingdom
on acid-free paper by CPI Antony Rowe, Chippenham, Wiltshire.

ISBN 978-0-19-929845-7

10 9 8 7 6 5 4 3 2 1

Oxford University Press makes no representation, express or implied, that the drug dosages in this book are correct. Readers must therefore always check the product information and clinical procedures with the most up-to-date published product information and data sheets provided by the manufacturers and the most recent codes of conduct and safety regulations. The authors and the publishers do not accept responsibility or legal liability for any errors in the text or for the misuse or misapplication of material in this work. Except where otherwise stated, drug dosages and recommendations are for the non-pregnant adult who is not breast-feeding.

# Contents

List of Contributors *ix*

## Part 1 The origins and functions of music *1*

*Edited by Ian Cross*

1 The nature of music and its evolution *3*
*Ian Cross*

2 Universals in music processing *14*
*Catherine Stevens and Tim Byron*

3 Music and meaning *24*
*Ian Cross and Elizabeth Tolbert*

4 The social and personal functions of music in cross-cultural perspective *35*
*Martin Clayton*

## Part 2 Music perception *45*

*Edited by Ian Cross*

5 The perception of pitch *47*
*Thomas Stainsby and Ian Cross*

6 Tonal cognition *59*
*Emmanuel Bigand and Bénédicte Poulin-Charronnat*

7 The perception of musical timbre *72*
*Stephen McAdams and Bruno L. Giordano*

8 Musical time *81*
*Mari Riess Jones*

9 Components of melodic processing *93*
*Mark A. Schmuckler*

10 **Memory for music** *107*
*Bob Snyder*

## Part 3 Responses to music *119*

*Edited by Donald A. Hodges*

11 Bodily responses to music *121*
*Donald A. Hodges*

12 Emotional responses to music *131*
*Patrik N. Juslin*

13 The relationship between musical structure and perceived expression *141*
*Alf Gabrielsson*

14 Aesthetics *151*
*David Huron*

15 Musical preferences *160*
*Alexandra Lamont and Alinka Greasley*

**Part 4 Music and the brain** *169*
*Edited by Aniruddh D. Patel*

16 The neurobiological basis of musical expectations *171*
*Laurel J. Trainor and Robert J. Zatorre*

17 Disorders of musical cognition *184*
*Lauren Stewart, Katharina von Kriegstein, Simone Dalla Bella, Jason D. Warren and Timothy D. Griffiths*

18 Music, musicians, and brain plasticity *197*
*Gottfried Schlaug*

19 Music and the brain: three links to language *208*
*Aniruddh D. Patel*

**Part 5 Musical development** *217*
*Edited by David J. Hargreaves*

20 Prenatal development and the phylogeny and ontogeny of music *219*
*Richard Parncutt*

21 Music lessons from infants *229*
*Sandra E. Trehub*

22 Music in the school years *235*
*Alexandra Lamont*

23 The impact of music instruction on other skills *244*
*Frances H. Rauscher*

**Part 6 Learning musical skills** *253*
*Edited by Susan Hallam*

24 Musical potential *255*
*Gary McPherson and Susan Hallam*

25 Practising *265*
*Harald Jørgensen and Susan Hallam*

26 Individuality in the learning of musical skills *274*
*Helena Gaunt and Susan Hallam*

27 Motivation to learn *285*
*Susan Hallam*

28 The role of the family in supporting learning *295*
*Andrea Creech*

29 The role of the institution and teachers in supporting learning *307*
*Graham Welch and Adam Ockelford*

**Part 7 Musical performance** *321*
*Edited by Richard Parncutt*

30 Measurement and models of performance *323*
*W. Luke Windsor*

31 Planning and performance *332*
*Eckart Altenmüller and Sabine Schneider*

32 Sight-reading *344*
*Andreas C. Lehmann and Reinhard Kopiez*

33 Performing from memory *352*
*Roger Chaffin, Topher R. Logan and Kristen T. Begosh*

34 Movement and collaboration in musical performance *364*
*Jane W. Davidson*

35 Emotion in music performance *377*
*Patrik N. Juslin*

36 Optimizing physical and psychological health in performing musicians *390*
*Dianna T. Kenny and Bronwen Ackermann*

**Part 8 Composition and improvisation** *401*
*Edited by Peter Webster*

37 Making a mark: the psychology of composition *403*
*Jonathan Impett*

38 Musical improvisation *413*
*Richard Ashley*

39 Children as creative thinkers in music: focus on composition *421*
*Peter R. Webster*

**Part 9 The role of music in everyday life** *429*
*Edited by John Sloboda*

40 Choosing to hear music: motivation, process and effect *431*
*John Sloboda, Alexandra Lamont and Alinka Greasley*

41 Music in performance arts: film, theatre, and dance *441*
*Annabel J. Cohen*

42 Peak experiences in music *452*
*John Whaley, John Sloboda and Alf Gabrielsson*

43 Musical identities *462*
*Raymond MacDonald, David J. Hargreaves and Dorothy Miell*

44 The effects of music in community and educational settings *471*
*Susan Hallam and Raymond MacDonald*

45 Music and consumer behaviour *481*
*Adrian C. North and David J. Hargreaves*

**Part 10 Music therapy** *491*

*Edited by Michael Thaut*

46 Processes of music therapy: clinical and scientific rationales and models *493*
*Shannon de l'Etoile*

47 Clinical practice in music therapy *503*
*Corene Hurt-Thaut*

48 Research and evaluation in music therapy *515*
*Barbara L. Wheeler*

49 Music therapy in medical and neurological rehabilitation settings *526*
*Anne Kathrin Leins, Ralph Spintge and Michael Thaut*

**Part 11 Conceptual frameworks, research methods and future directions** *537*

*Edited by Susan Hallam, Ian Cross and Michael Thaut*

50 Beyond music psychology *539*
*Adam Ockelford*

51 History and research *552*
*Michael Thaut*

52 Where now? *561*
*Susan Hallam, Ian Cross and Michael Thaut*

Index *569*

# Contributors

**Dr Bronwen Ackermann**
School of Physiotherapy
Faculty of Health Sciences
The University of Sydney
Lidcombe NSW 2141
Australia

**Professor Dr Eckart Altenmüller**
Institut für Musikphysiologie und
Musiker-Medizin
Hohenzollernstr 47
30161 Hannover
Germany

**Dr Richard Ashley**
School of Music
Northwestern University
711 Elgin Road
Evanston IL 60208-1200
USA

**Kristen T. Begosh**
Department of Psychology, U-1020
University of Connecticut
406 Babbidge Road
Storrs CT 06269-1020
USA

**Professor Emmanuel Bigand**
LEAD-CNRS UMR 5022
Université de Bourgogne
Pôle AAFE
Esplanade Erasme, BP 26513
21065 Dijon Cedex
France

**Tim Byron**
MARCS Auditory Laboratories
University of Western Sydney
Locked Bag 1797
Penrith South DC NSW 1797
Australia

**Dr Roger Chaffin**
Department of Psychology,
U-1020
University of Connecticut
406 Babbidge Road
Storrs CT 06269-1020
USA

**Professor Martin Clayton**
The Music Department
Open University
Walton Hall
Milton Keynes MK7 6AA
UK

**Professor Annabel J. Cohen**
Department of Psychology
University of Prince Edward Island
550 University Avenue
Charlottetown PEI
Canada C1A 4P3

**Dr Andrea Creech**
Institute of Education
University of London
20 Bedford Way
London WCIH OAL
UK

**Dr Ian Cross**
Faculty of Music
University of Cambridge
11 West Road
Cambridge
CB3 9DP
UK

**Dr Simone Dalla Bella**
Department of Cognitive Psychology
University of Finance and Management in Warsaw
Pawia 55 Street
01-030 Warsaw
Poland

**Professor Jane W. Davidson**
Department of Music
University of Sheffield
38 Taptonville Road
Sheffield S10 5BR
UK

**Dr Shannon de l'Etoile**
University of Miami
Frost School of Music
P O Box 248165
Coral Gables FL 33124
USA

**Professor Alf Gabrielsson**
Department of Psychology
Uppsala University
Box 1225
SE - 751 42 Uppsala
Sweden

**Dr Helena Gaunt**
Guildhall School of Music & Drama
Silk Street
Barbican
London EC2Y 8DT
UK

**Dr Bruno L. Giordano**
Centre for Interdisciplinary Research in Music, Media and Technology
Schulich School of Music of McGill University
555 Sherbrooke Street West
Montreal QC
Canada H3A 1E3

**Alinka Greasley**
School of Psychology
Keele University
Keele
Staffordshire ST5 5BG
UK

**Professor Timothy D. Griffiths**
Auditory Group
Institute of Neuroscience
The Medical School
Newcastle University
Newcastle upon Tyne NE2 4HH
UK

**Professor Susan Hallam**
The Institute of Education
University of London
20 Bedford Way
London WC1H 0AL
UK

**Professor David J. Hargreaves**
Centre for International Research on Creativity and Learning in Education (CIRCLE)
Roehampton University
Southlands College
Roehampton Lane
London SW15 5SL
UK

**Professor Donald A. Hodges**
Music Research Institute
University of North Carolina at Greensboro
P O Box 26170
Greensboro NC 27402-6170
USA

**Professor David Huron**
School of Music
1866 College Road
Ohio State University
Columbus OH 43210
USA

**Dr Corene Hurt-Thaut**
Assistant Professor of Music Therapy
Department of Music
Colorado State University
Fort Collins CO 80523
USA

**Dr Jonathan Impett**
School of Music
University of East Anglia
Norwich NR4 7TJ
UK

**Dr Mari Riess Jones**
Department of Psychology
Ohio State University and University of Santa Barbara California
5388 Traci Drive
Santa Barbara CA 93111
USA

**Dr Harald Jørgensen**
The Norwegian Academy of Music
PB 5190 Majorstua
N-0302 Oslo
Norway

**Dr Patrik N. Juslin**
Department of Psychology
Uppsala University
Box 1225
SE - 751 42 Uppsala
Sweden

**Professor Dianna T. Kenny**
Director, Australian Centre for Applied Research in Music Performance
Sydney Conservatorium of Music
The University of Sydney
Australia

**Dr Reinhardt Kopiez**
Hochschule für Musik und Theater
Emmichplatz 1
D-30175 Hannover
Germany

**Dr Alexandra Lamont**
School of Psychology
Keele University
Keele
Staffordshire ST5 5BG
UK

**Professor Andreas C. Lehmann**
Hochschule für Musik
Hofstallstrasse 6-8
D-97070 Würzburg
Germany

**Dr Anne Kathrin Leins**
Abendsrothweg 71
D-20251 Hamburg
Germany

**Topher R. Logan**
Community School of the Arts
University of Connecticut, U-5195
3 Witryol Place
Storrs, CT 06269-5195
USA

**Professor Stephen McAdams**
Centre for Interdisciplinary Research in Music, Media and Technology
Schulich School of Music of McGill University
555 Sherbrooke Street West
Montreal QC
Canada H3A 1E3

**Professor Raymond MacDonald**
Department of Psychology
School of Life Sciences
Glasgow Caledonian University
George Moore Building
Cowcaddens Road
Glasgow G4 0BA
UK

**Professor Gary McPherson**
School of Music
University of Illinois
1114 W. Nevada Street
Urbana, IL 61801
USA

**Professor Dorothy Miell**
Department of Psychology
Faculty of Social Sciences
The Open University
Walton Hall
Milton Keynes MK7 6AA

**Professor Adrian C. North**
School of Life Sciences
Heriot Watt University
Edinburgh EH14 4AS
UK

**Professor Adam Ockelford**
School of Education
Roehampton University
Roehampton Lane
London SW15 5PU
UK

**Professor Richard Parncutt**
Centre for Systematic Musicology
University of Graz
Merangasse 70
A-8010 Graz
Austria

**Professor Aniruddh D. Patel**
The Neurosciences Institute
10640 John Jay Hopkins Drive
San Diego CA 92121
USA

**Dr Bénédicte Poulin-Charronnat**
LEAD-CNRS UMR 5022
Université de Bourgogne
Pôle AAFE
Esplanade Erasme, BP 26513
21065 Dijon Cedex
France

**Dr Frances H. Rauscher**
Department of Psychology
University of Wisconsin Oshkosh
800 Algoma Boulevard
Oshkosh WI 54901
USA

**Dr Gottfried Schlaug**
Music and Neuroimaging Laboratory
Beth Israel Deaconess Medical Center
330 Brookline Avenue – Palmer 127
Boston MA 02215
USA

**Dr Mark A. Schmuckler**
Department of Psychology
University of Toronto Scarborough
1265 Military Trail
Scarborough ON
Canada M1C 1A4

**Dr Sabine Schneider**
Institut für Musikphysiologie und
Musiker-Medizin
Hohenzollernstr 47
30161 Hannover
Germnay

**Professor John Sloboda**
School of Psychology
Keele University
Keele
Staffordshire ST5 5BG
UK

**Professor Robert Snyder**
Sound Department
School of the Art
Institute of Chicago
112 South Michigan Ave
Chicago IL 60604

**Professor Ralph Spintge**
Director, Department of Pain Medicine
Sportklinik Hellersen
58515 Ludenscheid
Germany

**Dr Thomas Stainsby**
Department of Experimental Psychology
University of Cambridge
Downing Street
Cambridge CB2 3EB
UK

**Dr Catherine Stevens**
MARCS Auditory Laboratories
University of Western Sydney
Locked Bag 1797
Penrith South DC NSW 1797
Australia

**Dr Lauren Stewart**
Department of Psychology
Whitehead Building
Goldsmiths, University of London
New Cross
London SE14 6NW
UK

**Professor Michael Thaut**
School of the Arts
Center for Biomedical Research in music
Colorado State University
Fort Collins CO 80523
USA

**Elizabeth Tolbert**
Peabody Institute of the Johns
Hopkins University
1 East Mount Vernon Place
Baltimore, MD 21202
USA

**Dr Laurel J. Trainor**
Dept of Psychology, Neuroscience
& Behaviour
McMaster University
1280 Main Street W
Hamilton ON
Canada L8S 4K1

**Professor Sandra E. Trehub**
Department of Psychology
University of Toronto at Mississauga
3359 Mississauga Road N
Mississauga ON
Canada L5L 1C6

**Dr Jason D. Warren**
Dementia Research Centre
Institute of Neurology
University College London
London WC1N 3BG
UK

**Dr Katharina von Kriegstein**
Wellcome Department of Imaging
Neuroscience
Institute of Neurology
University College London
London WCIN 3BG
UK

**Professor Peter R. Webster**
School of Music
Northwestern University
711 Elgin Road, Room 119
Evanston IL 60208-1200
USA

**Professor Graham Welch**
The Institute of Education
University of London
20 Bedford Way
London WC1H 0AL
UK

**John Whaley MSc**
10291 Kismet Lane N
Stillwater, MN 55082
USA

**Professor Barbara Wheeler**
School of Music
University of Louisville
Louisville KY 40292
USA

**Dr Aaron Williamon**
Centre for the Study of Music
Performance (CSMP)
Royal College of Music
Prince Consort Road
London SW7 2BS
UK

**Dr W. Luke Windsor**
School of Music
University of Leeds
Leeds LS2 9JT
UK

**Professor Robert J. Zatorre**
Montreal Neurological Institute
McGill University
3801 University Street
Montreal, QC, Canada H3A2B4
and BRAMS laboratory
Montreal, QC
Canada

# PART 1

# The origins and functions of music

Edited by Ian Cross

# CHAPTER 1

# The nature of music and its evolution

Ian Cross

## The theory of evolution in musicological context

The idea that the theory of evolution has anything to offer to our understanding of music provokes strong reactions (both positive and negative) and has swung in and out of favour amongst those who have thought and written about music over the last hundred and fifty years. For the last few decades of the nineteenth century and the first few of the twentieth, the theory of evolution was of huge significance for music theorists, historians and anthropologists as a means of exploring and explaining musical change, musical difference and musical value. From around the 1940s to the beginning of the 1990s, the theory of evolution more or less vanished from musical and musicological discourse: the term 'evolution' tends to co-occur with musical terms only in the titles of books and articles such as 'The evolution of twelve-tone music', 'of Beethoven's late style'—and so on. However, from around the late 1980s, the theory of evolution began to re-emerge as a valid and prospectively valuable way of thinking about important aspects of music. This chapter will outline aspects of 'classical' evolutionary theory before delineating some of the factors that have led to the fall and rise in the fortunes of evolutionary theory as applied to music. It will give an outline of current conceptions of the relationships between music and evolutionary theory, and will conclude by exploring some of the implications for music psychology of evolutionary views of music.

Dictionary definitions of the term 'evolution' include 'the appearance of events in due succession', and 'the gradual development of something into a more complex or better form'. Only with the publication of Darwin's *Origin of Species* in 1859 did the meaning in terms of which 'evolution' is now most widely understood take clear shape: 'the theoretical process[es] by which all species develop from earlier forms of life'. 'Classical' evolutionary theory starts by noting that not all members of any given population of animals will be identical; there will be minor differences of form and capacity between individual members, and if those differences allow an individual to exploit its environment better than other individuals, they will increase the likelihood that the individual will survive and reproduce. If those differences can be inherited, all members of the population that bear them will be similarly advantaged by being better adapted to their environment, becoming prevalent in the population over time and leading to a change in its make up and perhaps to speciation, the emergence of new species (roughly speaking, a discrete population of animals that do not interbreed with other populations of animals). This process of variation, leading to differential levels of survival and reproduction, leading to speciation, can be termed 'natural selection'. Natural selection, together with sexual selection (evident in processes governing mate choice and hence affecting likelihood of reproduction) constituted the foundations of Darwin's theory of evolution, which was intended to provide an explanation

of the diversity of forms of life and their interrelationships without invoking any creative agency other than the workings of observable physical and biological processes.

The theory of evolution was developed to account for the natural world, but was almost immediately co-opted into theory and debate about human society. For many it appeared to offer a scientific basis on which to erect social theory, particularly insofar as it could be interpreted as giving support to ideas of scientifically grounded distinctions between human races, and of the progressive development of human societies. After all, it seemed to many influential thinkers that the processes of evolution constituted a means of explaining the continual refinement of a species' abilities. What could be more natural, in the Victorian age of progress, than to suppose that differences in degrees of ability between species—with human races interpretable as distinct species, or at least, as being analogous to species—conformed to universal principles that could be summarized in the theory of evolution? With these teleological trappings the theory of evolution came to serve as one of the key features of much writing about music at the beginning of the twentieth century. The 'progressivist' notion of evolution seemed to provide a systematic framework that could account not only for stylistic change and development in music across Western musical history, but also a means of qualifying music from non-Western societies as being more, or less, developed in their relationships to refined and hence 'highly evolved' Western musical practice (for an overview see Rehding 2000).

It should be noted that Darwin's own writings in *The expression of the emotions in man and animals*, and in *The descent of man* on music as an evolved human capacity played almost no role in any of the early twentieth-century musicological writings that espoused evolutionary perspectives. Darwin viewed the human capacity for music as a likely precursor of the capacity for language, having its origins in the vocal expression of emotion and as having had utility in processes of sexual selection (see Cross 2007). Moreover, he was somewhat agnostic concerning whether or not music of different cultures or races had greater or lesser value, but in any case did not agree with the proposition that racial difference corresponded to any difference in innate, human, musical capacity; in other words, humans, to Darwin, were a single species (see Chapters 7 and 19 of *The descent of man* [Darwin 2004]).

As the twentieth century progressed, consideration of origins in the study of music moved away from any exploration of music's relationship to biology to re-focus on the historical relationships between contemporary Western musical theory and practice, on Western musical history, or on music's relationships with abstract domains such as mathematics. For these strands of thought evolution was simply irrelevant to their interests, which were viewed as primarily musicological, concerned with the explication of the historical and ontological roots of Western music. Moreover, the repellent consequences of the racialist theorizing that teleogical interpretations of evolutionary theory had been called upon to sanction were all too evident in the aftermath of the Second World War.

In any case, within anthropology, an increasing tendency to focus on the cultural specificities of societies rather than on pancultural universalities or on any necessary concept of cultural progress diminished the apparent explanatory role of any biological foundation for culture and mind (see Shore 1996). By the middle of the twentieth century, exploration of music beyond the bounds of Western societies had come to concentrate on detailed ethnographic description and on attempts to understand the structures and functions of music in terms derived from the societies' own understandings of their music—in other words, in emic rather than etic terms. There is little evidence of evolutionary thinking as having had any significant impact on understanding music within societies or across cultures. Nevertheless, a vital feature of these ethnomusicological studies is the evidence they provide concerning the heterogeneity of music across different societies, a heterogeneity that problematizes the very notion of 'music' (see Chapter 4 this volume), and any evolutionary approach to understanding music requires at least an operational definition of what might constitute 'music'.

Evolutionary theory is first and foremost a theory about biology. It could be suggested that music has no existence that is independent of

the discourse that frames it: that 'music' is simply a discursive category with an identity that is dependent on its opposition to the other categories of discourse that we employ in the linguistic taxonomies that we have developed—largely in the west—to describe or talk about types of human behaviours. This is a view that seems at least implicit in much musicological thinking of the recent past (see e.g. Korsyn 2003, p. 187). However, most ethnomusicologists would subscribe to the view that there is something that can be described as 'musical' that is evident in the patterns of thought and behaviour of all known societies (see Chapters 2 and 4 this volume). If this is indeed the case, given that humans share a common biology as members of a single species, then whatever constitutes 'music' across cultures may reflect some general processes of thought and behaviour that should, in principle, be a proper focus of exploration in terms of evolutionary theory.

## 'Music' as an object of evolutionary exploration

Music varies from society to society to the extent that one culture's music may not be recognizable as music by members of another culture. This applies both to the structural features of the music and to the functions that it may fulfil. As Stevens and Byron show in Chapter 2 of this volume, the cognitive processing of music across human societies exhibits many common characteristics. Some of these processing commonalities map onto structural musical features, such as the use of discrete pitch levels, unequal-stepped scales, octave equivalence, low-integer frequency ratios, and periodic pulses. Moreover, features such as hierarchical organization and temporal structures that are shaped so as to modulate the expectancies of those engaged with the music, also appear to be generic in music across cultures. It may be that some features, particularly those concerned with the experience of fine-grained pitch structures and periodic temporal structures, should be considered as proper to music. Nevertheless, at least some of these common structural regularities may not be evident in particular instances otherwise identifiable as music (a paradigmatic, though controversial, case being Cage's *4'33'*), and at least some of the universal cognitive processes described in Chapter 2, such as the formation of hierarchical structures, and the implicit learning of expectancies, are operational in domains other than music. While arriving at an empirically grounded account of universal characteristics of musical processing and musical structure is essential in delineating what may constitute the foundations of music, it is not sufficient; it is also necessary to explore the extent to which there may be commonalities between people and across cultures in the ways in which music is manifested as social behaviour.

On the whole, ethnomusicologists have given surprisingly little consideration to the question of musical universals. Amongst the few significant exceptions are Bruno Nettl, Alan Merriam and John Blacking, who adopt somewhat different perspectives. Nettl (2005) takes a pragmatic approach, suggesting that etic (Western) and emic accounts should each feed in to determine what it is that ethnomusicologists should focus on as 'music'. Merriam (1964) suggests that 'music' can best be explored in terms of a tripartite model that embraces music as *sound* (what might conventionally be thought of as constituting music from a Western perspective), as *behaviour* (which embraces the musical—and 'non-musical'—acts of musicians, and the activities in which the production of music is embedded) and as *concept* (how people think about music in terms of its powers and its relations to other domains of human life). Blacking (1995), on the basis of his extensive fieldwork with the Venda peoples of southern Africa, and in particular, on his study of Venda children's music, claims that

> 'Music' is a primary modelling system of human thought and a part of the infrastructure of human life. 'Music'-making is a special kind of social action which can have important consequences for other kinds of social action.
>
> Blacking (1995, p. 223)

Blacking's claims appear to locate music as central to, and in some ways indissociable from, other domains of human behaviour. While the claim for music's centrality is not widely echoed

in the ethnomusicological literature, the conception of music as embedded in broader suites of behaviour is more general, as Bohlman puts it:

> expressive practices do not divide into those that produce music and those that produce something else, say ritual or dance. Music accumulates its identities ... from the ways in which it participates in other activities.
>
> Bohlman (2000, p. 293)

For those engaged in understanding music as it manifests itself across different cultures and historical times, 'music' appears to be protean, and its identification in any consistent manner seems particularly intractable. Certainly, music cannot be defined in the terms in which it is conventionally conceived of in contemporary Western societies, as a consumable commodity constituted of complexly patterned sound that is produced by a class of specialists and engaged with through listening for primarily hedonic reasons. In many, perhaps most, non-Western cultures it involves overt action and active group engagement, and is employed not only in caregiver–infant interaction, entertainment and courtship but also in ritual, particularly at times of significant life transitions (such as the passage from adolescence to adulthood, from season to season, or from life to death). More often than not, music is an integral part of a wider range of everyday activities. If a category of behaviours that can be termed 'music' has any generality across cultures, it seems that it can best be characterized as active, as founded in interaction, and as permeated by many other aspects of social life (see also Chapter 4 this volume).

Given that, in the ethnomusicological view, music and other human activities seem almost amalgamated, are there any features that would serve to distinguish 'music' as a discrete category of human thought and behaviour? At first sight, music seems to possess few characteristics that are not shared with other domains of behaviour, notably dance and language. Music involves patterned action in time, as does dance. Music appears communicative, complex, generative and representational, as does language. The concept of music is amalgamated with that of dance in many—perhaps the majority of—cultures. This fact, together with the stress on music as action in much of the ethnomusicological literature (see Chapter 2 this volume) suggests that it would be parsimonious to treat music and dance either as intrinsically related or simply as different manifestations of the same phenomenon.

Relationships between music and language are more difficult to disentangle, but perhaps the most significant factors that differentiate them are the types of structured interactions that they allow, and their contexts of use. Linguistic interactions are typically structured in time so as to coordinate the temporal succession of participants' contributions. Language possesses a generative complexity that allows for the production and reception of a potentially unlimited set of utterances: it is also directed towards the communication of representations of ideas, states of affairs, attitudes and affects that have relevance to their contexts of production and reception—language can signify or mean, unambiguously. While music seems to share some of these characteristics with language, at least two significant differences are apparent. Music may allow participants to act simultaneously rather than asynchronously as in language. In addition, music's meanings appear less stable and consensual than those of language (see Chapter 3 this volume).

Music's capacity to enable participants to act and to contribute to music-making simultaneously exploits the capacity of entrainment (Clayton *et al.* 2005). Entrainment here refers to the coordination in time of one participant's behaviours with those of another and involves the organization of the perception and behaviour of participants around temporal regularities that are inferred (generally non-consciously) from musical sounds and actions in the form of a periodic pulse or beat that is sensed by all participants (Clayton *et al.* 2005), being evidenced in continual processes of correction of errors by participants in alignment of action in both period and phase (see e.g. Repp 2001). Even engagement with music in apparently passive listening appears to rely on such entrainment processes, evidenced in periodic modulation of attentional load (see Chapter 8 this volume). It is notable that the capacity for entrainment may be unique to humans. While several species of cicadas and frogs do appear to

entrain, they do so in ways that appear different from those implicated in human entrainment and the issue of the origins of entrainment in the hominid or hominin lineage remains to be investigated (see Bispham 2006; Cross *et al.* 2008, in press).

Language has an indisputable efficacy in human interaction, in large part by virtue of its capacities to mean. It is often supposed that music's meanings can be reduced to the emotions it represents, expresses or elicits (see Juslin and Sloboda 2001) which suggests that meaning in music is a poor or natural cousin of meaning in language. However, while it is undoubtedly the case that music is valued for its affective powers in all societies, music's meanings extend beyond its affective value; as Tolbert (2001) notes, music's meanings are equally embodied, natural or affective, and artificial or symbolic. In general, however, meaning in music appears to be less susceptible to consensual determination than is meaning in language; music certainly bears meaning, but the meanings that it can bear are more impenetrable and susceptible to change according to the contexts in which they are experienced than are those of language. As Blacking (1995 p. 237) notes, 'Not only can the "same" patterns of sound have different meanings in different societies; they can also have different meanings within the same society because of different social contexts', an attribute of music characterizable as 'floating intentionality' (Cross 1999).

So music is differentiable from language in its exploitation of the human capacity to entrain, in the probable scope of its generativity and in ways in which it can mean. In some respects this comparison could appear to objectify music as an impoverished version of language. However, music and language coexist in all societies and fulfil different (though perhaps complementary) functions in those societies. While language is capable of expressing semantically decompos-able propositions that have unambiguous reference, music cannot. Nevertheless, there are numerous social situations in which unambiguous reference in communicative acts is not a desideratum as it may precipitate conflict in attitudes or actions. Music's exploitation of the human capacity for entrainment allows participants to experience a sense of 'shared intentionality', whilst under-specifying goals in ways that permit individuals to interact even while holding to personal meanings and goals that may actually be in conflict (Cross 2005). Music has a profoundly social efficacy (see Chapter 4 this volume), and it is possible to delineate music as a medium that is interactive, entraining, and that exhibits floating intentionality. Such a definition seems almost to exclude Western listening, which may often be passive and solitary rather than social. However, the object of listening—the music—can be conceived of as constituting a trace of human activity with which a listener may 'virtually' interact (see Clarke [2005] on 'subject-position' in music).

## Music in recent evolutionary thought

In parallel with the development of broad conceptions of music, over the latter half of the twentieth century theories of evolution underwent a revolution, in part driven by developments in theoretical biology and by the rise of the science of genetics (see e.g. Maynard-Smith and Szathmáry 1995). Neo-Darwinian understandings of evolution elucidated the genetic mechanisms whereby organisms transmitted their characteristics from one generation to the next, as well as processes, in addition to natural and sexual selection, that could operate so as to shape an organism's fitness, or adaptedness, with respect to its environment. Processes of kin selection have been distinguished whereby genetic relatedness can be adduced to account for seemingly altruistic behaviours between members of social species as well as for the emergence of sterile animal castes (see e.g. Hölldobler and Wilson 1994). Processes of genetic drift have also been shown to be operational whereby random changes in the genetic make-up of a population come, over time, to lead to populations with quite different characteristics from the parent population (see Maynard Smith 1998). Other work has shown that processes of selection can *appear* to operate at the level of the group. Theories such as those proposed by Sober and Wilson (1998) support the notion of a role for relationships between

social environment and individual fitness in profoundly social species such as humans. In a social species, behaviours that contribute towards survival and reproductive success may be determined as much by relations between members of the species (and by individual capacities to manage such relations) as by relations between individuals and their physical environments. All these developments provided substantial foundations for postulating that complex individual and social behaviours such as language and music, as well as the mental processes that underlie those behaviours, could be explored and understood in terms of evolutionary theory.

By the beginning of the last decade of the twentieth century, evolutionary psychology was coalescing as an identifiable and significant strand of research within both psychology and evolutionary theory. At first little attention was paid to music, but in 1997 one of its main protagonists, Steven Pinker, produced an evolutionary treatment of music that relegated it to the status of evolutionary by product. In *How the mind works*, Pinker devotes ten pages of his final chapter to considering music, because, in his words, he wishes to consider a mental faculty which 'shows the clearest signs of not being' [an evolutionary adaptation] to set against examples of mental faculties that are self-evidently adaptive, such as language. Pinker starts from the premise that 'as far as biological cause and effect are concerned, music is useless' (p. 528), noting that music is variable in its complexity from culture to culture, that while all tend to enjoy listening to music only a small subset of the population are practitioners, and that music communicates nothing but formless emotion. Hence music shows clear signs of being a 'technology', a human capacity developed and exploited for its own sake and at best evolutionarily neutral, rather than an adaptation. He suggests that this technology developed to exploit capacities that had arisen for largely adaptive reasons, claiming that 'music is auditory cheesecake, an exquisite confection crafted to tickle the sensitive spots of at least six of our mental faculties' (p. 534). However, the significance of Pinker's discussion of music is severely diminished by the lack of fit between his ethnocentric conception of music and that outlined above. Pinker seems to subscribe to a thoroughly Westernized notion of music as a commodified set of complex sound patterns produced by the few and consumed by the many simply for pleasure, rather than as the complex and socially efficacious actions and structures that it is and has been both in the west and in other cultures, places and times.

Subsequently, another prominent evolutionary psychologist, Geoffrey Miller, presented an alternative evolutionary psychological account of music quite at variance with that of Pinker. For Miller, music constitutes a medium that is well-suited to demonstrate the 'protean', unpredictable and creative, properties of an individual, properties that are selectively advantageous and hence desirable in the determination of mate-choice. Music is well-suited to make manifest such properties as it combines 'ritualized rules of tonality, rhythmicity, melody and harmony and protean intentions and variations' (Miller 1997, p. 322). Hence protean individuals might well advertise their creative assets in musical display. Miller (in Wallin *et al.* 2000) seeks to support this argument by reference to a range of Western musical examples (including Jimi Hendrix) claiming that the sexual opportunities afforded by rock-star status are in line with the predictions of his theory. Although music is certainly used for courtship in most, if not all, societies, its roles are always more multifarious. Moreover, Miller's theory would predict that musical ability should exhibit, or at least should have exhibited, a significant sexual dimorphism. While this would be difficult to ascertain in our predecessor species, it is not the case with modern humans; while musical roles are often sexually differentiated, in most cultures musicality appears to be equally exhibited by both males and females. If anything, the manifestation of musicality that is perhaps most culturally widespread, in the form of the use of protomusical and musical forms of interaction between caregiver and infant, is primarily evidenced by females (see Dissanayake in Wallin *et al.* 2000). Moreover, as Fitch (2006) points out, the precocious capacities of human infants in music perception would suggest that musicality is a trait that is unlikely to have been sexually selected for, given that such traits tend to emerge only in sexual maturity in other species.

In contrast to the positions of both Miller and Pinker, the majority of recent evolutionary treatments of music have focused on its effects at the level of the group; music is viewed as having, and as having had, effects on inter-group and intra-group interaction, which are conceived of as having had evolutionary consequences. Hagen and Bryant (2003) propose that group musical displays allow groups to indicate to each other their cohesion and stability, affording groups which can engage successfully in such displays a means of deterring other, more fragmented, groups from attempting to compete for scarce resources. Merker (in Wallin *et al.* 2000) proposes a similar putative role for music as an indicator of coalition strength. Brown (2004) acknowledges music's prospective role in inter-group relations but emphasizes its functionality within groups. He develops the notion of music as efficacious at the level of the group, suggesting that music typically acts at to influence behaviours at the group level. He suggests that music acts to reinforce cooperative behaviour within the group by means of group ritual activities, promoting a sense of 'groupishness' that is likely to enhance prospects of group survival in addition to being effective in situations of inter-group conflict.

This focus on music's significance in promoting cooperative, within-group, behaviours, is mirrored by others. Kogan (1997) views music as significant in human evolution by virtue of its embeddedness in social interaction and its effect of the formation of group identity, as does Dunbar (2004), who suggests that its primary effect is in the consolidation of group bonds. Mithen's (2005) extensive treatment echoes Brown (in Wallin *et al.* 2000) in presenting a view of music and language as having common origins and in suggesting that music had an evolutionary efficacy in the formation of social bonds. Mithen concludes, however, that music is no more than an atavistic relic that may have been functional for earlier species of hominins but whose functions have come to be usurped by language—a view that contrasts significantly with that of most ethnomusicological evidence (see Chapter 4 this volume).

In a series of publications Cross (1999, 2005, 2008, in press) also subscribes to the notion that music has a bonding effect, but suggests in addition that music, by virtue of its polyvalent significances, can contribute positively to individual fitness within the group by facilitating communicative interactions that, were they to be conducted linguistically, might give rise to conflict, a view also presented in Morley (2002). Whilst proposing that music may have an impact on individual fitness outwith social contexts by virtue of its 'metaphorising' powers (Cross 1999), Cross notes that music's powers of entrainment, together with its 'floating intentionality', render it an extremely efficacious communicative medium in managing situations of social uncertainty (Cross 2008, in press, Cross and Woodruff 2008). Hence music can be regarded as possessing attributes that complement those of language, and it can be postulated that music and language are likely to have co-evolved as complementary aspects of the human communicative toolkit, differentiated only by the degree to which they are capable of specifying meaning unambiguously (see also Chapter 3 this volume).

Dissanayake (in Wallin *et al.* 2000) takes a rather more constrained notion of the 'group' that might have constituted the matrix for the evolution of human musicality; the mother–infant dyad. In alignment with the views of Trevarthen (1999), she views infant-directed speech (see Trehub 2003), the infant's responses, and the kinesthetically interactive context of mother–infant communication as forming protomusical behaviours that are oriented towards the (co-)regulation of emotion, as well as forming the basis for the infant's acquisition of prosodic, phonemic and social competence. In this her proposal is in line with that of Falk (2004), who suggests that complex, vocal and gestural mother–infant interactions of the type that can be categorized as motherese are an evolutionarily adaptive response to the emergence of bipedalism, at around the time of the chimp–hominin split (*c.* 5–7 million years ago).

## Music and evolution—the evidence

These diverse evolutionary perspectives on music suggest that it can fruitfully be explored in terms of contemporary evolutionary theory.

The evidence would suggest that music has a significant role in human behaviour—that it is functional—and that it is distinguishable from other domains of human behaviour. The latter point suggests that music may have a distinct evolutionary track record, and its functionality implies that it may have been evolutionarily adaptive in the emergence of the human species (or at least, exaptive—that is, arising only as a consequence of the independent prior emergence of some adaptive capacity but having an impact on evolutionary fitness). This begs the questions of when might the capacities for music have arisen, and how they came to arise. In order to explore these issues we must turn to the archaeological evidence.

Music—and indeed, sound—in prehistory has received little attention in the archaeological literature. D'Errico *et al.* (2003) summarize much of what is known in the context of a re-evaluation of current archaeological approaches, particularly in the light of recent finds in southern Africa. They note that the earliest unequivocally musical artefact found to date is a bone pipe from Geissenklösterle in southern Germany, dating to *c.* 38 000 before the present (BP). A further sequence of bone pipes has been found at Isturitz in southern France (in contexts rich in parietal—cave—art), with dates ranging from 20 000 to 35 000 BP. The earliest of these dates more or less coincides with the arrival of modern humans in Europe and suggests that something like music was of high significance to those peoples. As D'Errico *et al.* (2003) comment, these artefacts are extremely sophisticated, exhibiting subtle design features that are analogous to those found in historic wind instruments; the time, effort and expertize devoted to their manufacture must have been considerable and suggests that music was likely to have been of considerable importance to a people who had just come to inhabit a new and potentially threatening environment. The use of musical artefacts will have been preceded by the expression of musical capacities by voice and body, and would appear likely to have an ancient provenance. The ubiquity of music in native American and Australian societies in forms that are not directly relatable to historic Eurasian or African musics strongly suggests that modern humans brought musicality with them out of Africa. If this is the case we must look to evidence that will enable us to identify the time-depth of the capacity for music, or perhaps of different subcapacities that make up human musicality, which means turning to the prehistoric human fossil record.

A brief account of stages in human evolution might be useful at this point. The hominin (human) lineage last shared a common ancestor with chimpanzees and bonobos some 5–7 million years ago (mya). The earliest unambiguous members of the hominin line are found in Africa: *Ardipithecus ramidus*, at between 7 and 4.5 mya, and a range of *australopithecine* species, from around 4.2 to 2.4 mya). Both lineages were morphologically much more similar to chimps than to modern humans (having similar-size brains at <400 cc), and can be inferred to have had lifeways that closely resembled chimps—other than, in the case of the australopiths, at least partial bipedalism. Subsequently, the homo line proper appears with *Homo habilis* (ca 2.4–1.6 mya) and *Homo ergaster* or *erectus* (ca 1.9–0.2 mya), the latter line having a body build similar to modern humans and possessing brains of at least double the size of the australopiths; both *habilis* and *ergaster* made stone (lithic) tools, the latter making objects of some sophistication. *Homo heidelbergensis* (despite the name a predominantly African species, from around 0.6 to about 0.1 mya) shows a considerable increase in brain size (to around 1200 cc), and evidence of complex lifeways, including the making of non-lithic artefacts such as wooden spears. This last species was the ancestor of both *Homo neanderthalensis*, a mainly northern, highly cold-adapted, species (ca 0.2–0.03 mya) and *Homo sapiens*, ourselves who appear first in Africa some 200 000 years ago.

The fossil record from which we are able to construct this lineage of species suggests strongly that musicality does not emerge as a full-blown capacity, but rather that subcomponents of that capacity emerge at different times and probably in response to different selection pressures. Australopiths had funnel-shaped chests (as do chimps), which do not permit sufficiently fine control over phonation to make the subtly differentiated vocal sounds that would be required of music. Only with *ergaster* does the barrel-shaped chest of modern humans arise

(Frayer and Nicolay in Wallin *et al.* 2000), hence only with this species can we postulate the emergence of a capacity for control of fine-grained vocalizations (see also Morley 2002). However, it is only with *heidelbergensis*, around 0.5 mya, that vocal and auditory capacities characteristic of modern humans can unambiguously be claimed to be in place (Martinez *et al.* 2004). By contrast, the dating of the emergence of a capacity for entrainment, which appears central in musical behaviour, is highly debatable. As mentioned above, it has been postulated that amongst extant species, a capacity for entrainment is specific to humans (Bispham 2006; Cross *et al.* 2008, in press). If this is so then the capacity to entrain must have arisen at some point in the hominin lineage over the last 5–7 million years, but at present we can be no more specific than that.

In view of the close link between music and social behaviour, the dating of the emergence of a capacity for complex sociality is also relevant to the question of music in human evolution. A variety of species, particularly primates, exhibit complex social behaviours, but as Foley (1995) notes, none appears to be as socially flexible and inventive as humans and, probably, our immediate ancestor species. Given the likely continuity in lifeways between chimpanzees and australopithecines, it seems that a complex capacity for sociality comparable to that of modern humans is unlikely to have predated *habilis* or *ergaster*. It is likely that only with the appearance of *heidelbergensis*, with a brain size in the (low) modern human range and some evidence for behavioural complexity of a different order from *ergaster*, that the modern human range of complex social behaviours began to emerge. However, it is only with the arrival of modern humans that we find clear evidence for symbolic behaviours (Henshilwood and Marean 2003), which has been taken to indicate that only with modern humans do we see the appearance of something that would be recognizable as language. Overall, a reasonable initial hypothesis appears to be that many of the capacities for music emerge independently at different times in hominin evolution, but only with modern humans are we likely to find an integrated capacity for music together with language, as suggested by several researchers (including Brown, Cross and Mithen). Whatever its precursors, it seems that modern musicality has very ancient roots.

## Conclusions and directions for research

As should be evident, there is much about the relationships between human musicality and human evolution that remains unknown. If music is truly the product of distinct, and distinctly functional, human capacities, a better understanding of music in evolutionary terms should be of immense value in elucidating the significance and distinctness of music as cognitive process and as social behaviour. Much of what we know or infer is based on very sketchy evidence, and there is immense scope for empirical exploration.

As Stevens and Byron note in Chapter 3, most of the research on which our understanding of the universals of music processing is rooted in Western cultural practice and based on the responses of Western-encultured or formally trained participants. There is a pressing need to extend the scope of research on music cognition into non-Western musical domains. Similarly, ethnomusicological evidence, and the success of music as a therapeutic medium, suggests that music has profound efficacy in social context; research is urgently required that explores the cognitive and socially interactive correlates of that efficacy. Indeed, research into music as interactive behaviour in its own right is very much in its infancy. While a number of studies have explored musical interaction between formally trained western Musicians, and a few are beginning to investigate interaction amongst non-Western musicians (see Chapter 4 this volume and also Moran 2007), very little empirical research into either the generic foundations of musical interaction—though see Himberg (2006), or into musical interactions amongst encultured rather than formally trained participants, has been conducted.

One large-scale research programme that has aimed to identify key components of musicality in ways that might shed light on relationships between music and evolutionary process is that conducted by Peretz and her collaborators

(see e.g. Peretz 2006), which has systematically sought to explore the nature, distribution within the population, and recently, genetic provenance, of the condition known as 'amusia': the incapacity to process musical information coherently. It appears that this condition is associable with an inability to process fine-grained pitch structures and relationships and that it is independent of deficits in other domains such as language and hence proper only to music (though see Patel *et al.* 2008). However, as with other studies that seek to explore any identifiable genetic bases for musicality (see e.g. Drayna *et al.* 2001), the vast majority of research has been conducted in listening tests using Western-encultured participants and musical materials that conform to western common-practice musical principles. As the issues addressed in this chapter and in Chapters 2 and 4, should make clear, it is imperative for the cognitive sciences to engage with music in all its cultural diversity in order to begin adequately to understand the nature of human musicality.

Moreover, as noted above, we know very little about the extent to which aspects of human musicality may be shared with other species. We have very little concrete evidence concerning the cross-specific generality of many capacities that appear central to human musicality, such as the capacity to entrain, or to differentiate between complex sonic and gestural structures, or to use complexly patterned sound and gesture to mediate interactions. There are very substantial ethological literatures (dealing with, for example, inter-individual and social functions of birdsong, and sonic communication amongst primates), as well as a few studies of animal behaviours in the presence of music (for an overview see Rickard *et al.* 2005). However, these literatures and studies require to be explored and extended in the light of questions concerning possible continuities between the capacities of non-human animal species and human musicality that are raised by an evolutionary perspective on music.

Over the last few years there has been a significant increase in the amount of attention devoted to the relationships between music and evolutionary theory. This chapter can only provide a sketch of the field, but those interested to know more should certainly explore Darwin (1998) as well as the most significant publication in the field to date, the volume edited by Wallin *et al.* (2000). A more extended treatment of many of the issues raised in the present chapter is given in Cross (2007).

## References


Bispham J (2006). Rhythm in music: what is it? Who has it? and Why? *Music Perception*, **24**(2), 125–134.

Blacking J (1995). *Music, culture and experience.* University of Chicago Press, Chicago, IL.

Bohlman P (2000). Ethnomusicology and music sociology. In D Greer, ed., *Musicology and sister disciplines*, 288–298. Oxford University Press, Oxford.

Brown S (2004). Evolutionary models of music: from sexual selection to group selection. In F Tonneau, and N Thompson, eds, *Perspectives in ethology 13: behavior, evolution and culture*, 231–281. Plenum Publishers, New York.

Clarke EF (2005). *Ways of listening: an ecological approach to the perception of musical meaning.* Oxford University Press, Oxford.

Clayton M, Sager R and Will U (2005). In time with the music: the concept of entrainment and its significance for ethnomusicology. *ESEM counterpoint*, **1**, 1–45.

Cross I (1999). Is music the most important thing we ever did? Music, development and evolution. In SW Yi, ed., *Music, mind and science*, 10–39. Seoul National University Press, Seoul.

Cross I (2005). Music and meaning, ambiguity and evolution. In D Miell, R MacDonald and D Hargreaves, eds, *Musical communication*, 27–43. Oxford University Press, Oxford.

Cross I (2007). Music and cognitive evolution. In RIM Dunbar and L Barrett, eds, *Handbook of evolutionary psychology*, 649–667. Oxford University Press, Oxford.

Cross I (2008, in press). The evolutionary nature of musical meaning. *Musicae Scientiae.*

Cross I, Bispham J, Himberg T and Swaine J (2008, in press). The evolution of musical rhythm. *Evolutionary Psychology.*

Cross I and Woodruff GE (2008). Music as a communicative medium. In R Botha and C Knight, eds, *The prehistory of language, Vol. 1*, 133–134. Oxford University Press, Oxford.

D'Errico F, Henshilwood C, Lawson G, Vanhaeren M, Tillier A-M, Soressi M *et al.* (2003). Archaeological evidence for the emergence of language, symbolism, and music—an alternative multidisciplinary perspective. *Journal of World Prehistory*, **17**(1), 1–70.

Darwin C (1859). *The origin of species.* John Murray, London.

Darwin C (1998). *The expression of the emotions in man and animals*, 3rd edn. HarperCollins Publishers, London.

Darwin C (2004). *The descent of man and selection in relation to sex*, 2nd edn. Penguin Books, London.

Drayna D, Manichaikul A, de Lange M, Snieder H and Spector T (2001). Genetic correlates of musical pitch recognition in humans. *Science*, **291**, 1969–1972.

Dunbar RIM (2004). *The human story*. Faber and Faber, London.

Falk D (2004). Prelinguistic evolution in early hominins: whence motherese? *Behavioral and Brain Sciences*, **27**(4), 491–541.

Fitch WT (2006). The biology and evolution of music: a comparative perspective. *Cognition*, **100**(1), 173–215.

Foley RA (1995). *Humans before humanity*. Blackwell, Oxford.

Hagen EH and Bryant GA (2003). Music and dance as a coalition signaling system. *Human Nature*, **14**(1), 21–51.

Henshilwood CS and Marean CW (2003). The origin of modern human behavior: critique of the models and their test implications. *Current Anthropology*, **44**(5), 627–651.

Himberg T (2006). Co-operative tapping and collective time-keeping—differences of timing accuracy in duet performance with human or computer partner. In M Baroni AR Addessi R Caterina and M Costa, eds, *Proceedings of the 9th international conference on music perception and cognition*, p. 377. ICMPC 2006, Bologna, Italy.

Hölldobler B and Wilson EO (1993). *Journey to the ants: a story of scientific exploration*. Harvard University Press, Cambridge, MA.

Juslin P and Sloboda JA (eds) (2001). *Music and emotion: theory and research*. Oxford University Press, Oxford.

Koelsch S, Kasper E, Sammler D, Schultze K, Gunter T and Frederici A (2004). Music, language and meaning: brain signatures of semantic processing. *Nature Neuroscience*, **7**(3), 302–307.

Kogan N (1997). Reflections on aesthetics and evolution. *Critical Review*, **11**(2), 193–210.

Korsyn K (2003). *Decentering music: a critique of contemporary music research*. Oxford University Press, Oxford.

Martinez I, Rosa M, Arsuaga J-L, Jarabo P, Quam R, Lorenzo C *et al.* (2004). Auditory capacities in Middle Pleistocene humans from the Sierra de Atapuerca in Spain. *Proceedings of the National Academy of Sciences*, **101(27)**, 9976–9981.

Maynard Smith J (1998). *Evolutionary genetics*, 2nd edn. Oxford University Press, Oxford.

Maynard Smith J and Szathmáry E (1995). *The major transitions in evolution*. Oxford University Press, Oxford.

Merriam AP (1964). *The anthropology of music*. Northwestern University Press, Chicago, IL.

Miller GF (1997). Protean primates: the evolution of adaptive unpredictability in competition and courtship. In A Whiten and RW Byrne, eds *Machiavellian intelligence II: extensions and evaluations*, 312–340. Cambridge University Press, Cambridge.

Mithen S (2005). *The singing Neanderthals: the origins of music, language, mind and body*. Weidenfeld and Nicholson, London.

Moran NS (2007). *Measuring musical interaction*. Unpublished PhD thesis, Open University, Milton Keynes.

Morley I (2002). Evolution of the physiological and neurological capacities for music. *Cambridge Archaeological Journal*, **12**(2), 195–216.

Nettl B (2005). The study of ethnomusicology: thirty-one issues and concepts, 2nd edn. University of Illinois Press, Urbana and Chicago, IL.

Patel AD, Wong M, Foxton J, Lochy A and Peretz I (2008). Speech intonation perception deficits in musical tone deafness (congenital amusia). *Music Perception*, **25**(4), 357–368.

Peretz I (2006). The nature of music from a biological perspective. *Cognition*, **100**(1), 1–32.

Pinker S (1997). *How the mind works*. Allen Lane, London.

Rehding A (2000). The quest for the origins of music in Germany circa 1900. *Journal of the American Musicological Society*, **53**(2), 345–385.

Repp BH (2001). Phase correction, phase resetting, and phase shifts after subliminal timing perturbations in sensorimotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance*, **27**(3), 600–621.

Rickard NS, Toukhsati SR and Field SE (2005). The effect of music on cognitive performance: insight from neurobiological and animal studies. *Behavioral and Cognitive Neuroscience Review*, **4**(4), 235–261.

Shore B (1996). *Culture in mind: cognition, culture, and the problem of meaning*. Oxford University Press, Oxford.

Sober E and Wilson D (1998). *Unto others: the evolution and psychology of unselfish behavior*. Harvard University Press, London.

Tolbert E (2001). Music and meaning: an evolutionary story. *Psychology of Music*, **29**, 89–94.

Trehub SE (2003). The developmental origins of musicality. *Nature Neuroscience*, **6**(7), 669–673.

Trevarthen C (1999). Musicality and the intrinsic motive pulse: evidence from human psychobiology and infant communication. *Musicae Scientiae, special issue: Rhythm, musical narrative, and the origins of human communcation*, 155–215.

Wallin N, Merker B and Brown S (eds) (2000). *The origins of music*. MIT Press, Cambridge, MA.

# CHAPTER 2

# Universals in music processing

Catherine Stevens and Tim Byron

## Introduction and definitions

In this chapter we outline areas of musical processing that may be universal to humans. Music here refers to temporally structured human activities, social and individual, in the production and perception of sound organized in patterns that convey non-linguistic meaning. Music processing refers to the neural contribution in perception, cognition and production of music.

A universal in a domain of human activity such as music implies a static feature of a static environment. Not surprisingly, many ethnomusicologists have railed against a search for similarities at the expense of informative differences, and there is a logical problem with documenting definitive generalities rather than particular instances. The approach adopted here is to emphasize musical processing over content and discuss the psychological processes implicated in understanding and engaging in musical behaviour rather than the content of musical knowledge or action (cf. Harwood 1976; Meyer 1960). Our focus is on the principles by which the brain processes features of musical environments.

To call a process involved in the perception, cognition or production of music 'universal' does not necessarily mean that the process is found in all humans. When we refer to universality, we are using Nettl's (2005) concept of 'statistical universals'. Statistically universal music processing is universal to the great majority of adult humans and cultures. However, there may be statistical outliers—which may take the form of all participants in a particular culture (see Everett [2005] for a culture that may be a statistical outlier in the field of linguistics). Additionally, some people may belong to a different statistical population; for example, those who have been described as having congenital amusia or infants who, for example, lack higher-order cognitive processes not developed until later in life.

Experimental studies of music cognition tend to focus on psychophysical features of Western post-Renaissance tonal music, and samples of participants from a restricted range of countries and cultures. Additionally, most experiments treat music as a purely auditory event and assume implicitly that the use of music in other cultures is similar to the use of music in Western cultures. Both assumptions are inaccurate. Vocal and instrumental music is also visual, motoric and kinaesthetic. All of the arts—oral literature/song poetry, visual arts and performing arts—are integrated, for example, in Australian Aboriginal song (Ellis 1984).

The universal music processes discussed here are hypotheses that require investigation and falsification in as many and varied cultural contexts as possible. The discussion begins with processes of grouping and segmentation, then statistically universal features of musical environments, and finally more general-purpose psychological processes. We illustrate some processes drawing on examples of production of song from particular Australian Aboriginal cultures.

## Perceptual principles of grouping and segmentation

The perceptual organization principles of grouping by proximity, similarity, common fate, good continuation, and figure-ground, which appear to operate in the early stages of music perception (Deutsch 1999; Handel 1984; Trehub and Hannon 2006) are candidates for universals in music processing. Perceptual organization of this kind can be demonstrated in very young animals; general grouping rules override learned knowledge about the shape of objects. For example, in visual perception, camouflage exploits this principle (Bregman 1990). Auditory grouping processes are thought to (a) organize acoustic information into music events (simultaneous grouping); (b) connect events into musical streams (sequential grouping); and (c) chunk event streams into musical units (segmentational grouping). These perceptual processes develop earlier than knowledge about the meaning of the events that have been grouped together (McAdams 1989). The suggestion is that these processes may be a basic property of the mammalian nervous system, and thus a likely contender for a universal in music processing.

At the level of sequential grouping, there are several processes which contribute to the perception of grouping. A sequence of tones is perceived as a group or stream when the pitch distance between temporally adjacent tones is small (pitch proximity); tones with distant pitches are segregated into separate streams (Bregman 1990). Whether events are perceived as a group or stream is also influenced by timbre similarity, grouping by temporal proximity, grouping by good continuation, and grouping by amplitude similarity (Deutsch 1999). At the level of segmentational grouping, we may have a universal preference for musical elements that are grouped correctly, though what is defined as correct grouping may be culture-dependent. Experiments investigating segmentation of Western tonal music in 6- and 4½-month-old infants have shown that infants prefer correctly segmented pieces with relatively long notes and downward pitch contours, over incorrectly segmented pieces (e.g. Krumhansl and Jusczyk 1990; see also Nan *et al.* 2006). Infants' correct segmentation of music may occur because music bears a structural resemblance to speech, which infants have learned through exposure, and/or because downward contours and extended durations naturally mark the end of all auditory signals (Thompson and Schellenberg 2006). Grouping principles could be investigated in broader musical contexts. However, that they work from the bottom up suggests that such processes are universal. Grouping likely aids encoding and storage in memory, enables an understanding of new material, and facilitates transmission to others.

The processing of melodic contour, the ups and downs in pitch of a melody, has the hallmarks of universality. It has been argued that efficient computational recognition of melodies uses melodic contour as chunks (Trehub 2000), and it appears that the rises and falls of a melody are part of the framework by which working memory processes melody. An ability to discriminate between short melodies on the basis of melodic contour is present in Western infants (e.g. Jusczyk and Krumhansl 1993) and adults (e.g. Dowling 1978), and may share cognitive resources with the processing of prosodic information (Patel *et al.* 1998). Ethnomusicologists have used different melodic contour shapes (for example, arch shapes) to categorize different kinds of melodies from a culture into different kinds of song (e.g. Kolinski 1970), and to compare cultures based on their use of different kinds of melodic contours (Lomax 1968). In an exploratory ethnomusicological study by Arom and Fürniss (1993) Aka Pygmy musicians were reported to be able to discriminate between melodies on the basis of melodic contour. As the processing of melodic contour appears to be present in Western infants and adults, and can be used to distinguish music by different cultures, insofar as melody is universal, it is likely that the processing of melodic contour in memory is universal.

## Features of musical environments and perceptual and cognitive constraints

Most discussions of universals in music from a psychological perspective consider the limits

imposed by human perceptual and cognitive processes. The constraints that recur in the literature suggest consensus. However, results on which these accounts are based emanate, in the main, from psychophysical studies involving culturally narrow samples. Regarding pitch first and then time, patterns in musical environments related to human perceptual and cognitive limits that await further cross-cultural scrutiny include the following.

1 Perceptual fusion of harmonic spectra into pitches (Justus and Bharucha 2002).
2 The use of discrete pitch levels (Burns 1999).
3 The semitone as the smallest viable interval. As a semitone is six times larger than the smallest interval that most listeners are able to discriminate (McAdams 1989), human working memory and production limits rather than perceptual discrimination are likely to constrain retention of music constructed from finer-grained musical scales.
4 Musical scales with differently sized steps between consecutive tones. Unequal steps may confer a psychological advantage, allowing tones to have different functions within the scale, such as a reference tone (Balzano 1980).
5 The prevalence of small integer frequency ratios[1] possibly because of their efficiency in auditory perception, i.e., they facilitate encoding, retention of melodies, and enable detection of subtle violations (e.g. Schellenberg and Trehub 1996; but see Balzano 1982 for an alternative view). Such ratios apply not only to fundamental frequencies but also within spectra (Sethares 1998; Tenney 1988).
6 Tones and melodies that are separated by an octave being perceived as similar (Wright *et al.* 2000) although Carterette and Kendall (1999) note that the effect is subtle. The prevalence of octaves in the auditory environment is one possible explanation (Schwartz *et al.* 2003; Trainor 2006); a physiological basis is also debated (Burns 1999).
7 Perception of rhythmic groups of up to 5 seconds in duration—the approximate limit of auditory sensory memory (Darwin *et al.* 1972)—where inter-onset intervals are between 100 and 1500 ms; see Cowan (1984) for the limit exceeding 5 seconds.
8 A regular beat or periodic pulse that affords temporal coordination between performers and elicits a synchronized motor response from listeners (Drake and Bertrand 2001). Importantly, the concept and experience of beat may vary across cultures (Iyer 1998; Stobart and Cross 2000).
9 Small-integer ratios of durations being easier to process than more complex rhythms (Drake and Bertrand 2001). Regular rhythmic patterns may facilitate encoding and retention of gestural, spoken or sung material.

[1] However, Will (1997) gives an example of a cultural context within which frequency differences rather than ratios might constitute the basis for melodic prediction and perception.

## Higher-order processes in music cognition

### Hierarchical organization and relational processing

So far, we have listed perceptual features of music and bottom-up perceptual constraints relating to music. However, higher-level cognitive processes are also engaged in music cognition, composition, improvisation, and performance in diverse cultural settings. The hierarchical organization of music, and relational processes in music, are possible universal processes in music cognition (see also Chapter 6 this volume).

Music is generally complex, generative, multilayered and hierarchical (Bharucha *et al.* 2006) although exceptions to hierarchical organization include the cyclical structures of Balinese gamelan music (Tenzer 2006) and minimalist music (e.g., compositions by Steve Reich). In general, streams of sound are structured hierarchically within and across dimensions; such structure likely confers a processing advantage. Bharucha *et al.* divide hierarchical representations into tonal hierarchies and event hierarchies. Tonal hierarchies organize tones within a key into stable and unstable pitches; some tones are perceived as more stable, thus more important, than others. Event hierarchies extend

upward from smallest subdivisions of a beat, to beat level, then measure, phrase, period, and large-order forms.

### Tonal hierarchies

Evidence of sensitivity to tonal hierarchies has also been established in Balinese music (Kessler *et al.* 1984), Scandinavian North Sami yoiks (Krumhansl *et al.* 2000), and Korean Court music (Nam 1998). These experiments sampled participants from within and outside those musical cultures. The basis of North Indian music, for example, is a set of melodic forms called the rāg that are built on a set of pitch scales called thāts. Tones within rāgs are organized hierarchically by importance, as in Western music. Castellano *et al.* (1984) demonstrated that both Indian and American students are sensitive to the hierarchical ordering of tones in rāgs.

### Temporal hierarchies

The pulsation of some African musical styles is very different from the hierarchical concept of measure in Western classical music and the related concepts of metre, strong, and weak beats (Iyer 1998; Magill and Pressing 1997). Instead, this music is measured and is based on pulsations—a sequence of isochronous temporal units that can be realized as a beat. The beat is the analogue of the tactus of Western tonal music. In west African music, the main beat and its metric grouping are articulated in an indirect fashion—not with accentual reinforcement as in Western music, but with suggestion and complexity (Iyer 1998). The metre is encoded in the rhythm itself; it is unambiguous but culturally specific. Iyer provides examples of a standard bell pattern that would be heard phrased in three different ways by three different cultural groups.

Knowledge of temporal hierarchies is also attuned through experience. For example, Drake and Ben El Heni (2003) compared how Tunisian and French subjects synchronized tapping with music from those two contrasting musical cultures. Participants tapped more slowly and with more hierarchical levels with music of their own culture than with music of an unfamiliar musical culture. Musicians and non-musicians did not differ in a spontaneous synchronization task although musicians could synchronize with significantly more hierarchical levels.

Clayton (1997), examining tal, a concept in North Indian (Hindustani) music broadly equivalent to metre, finds similarities and differences between tal and metre. Tal is a temporal framework acting as a background for rhythmic design; it is a periodic and hierarchic temporal framework which involves the interaction of two or more streams of pulsation. Rhythm in Indian music is interpreted with respect to tal; musically untrained Indian listeners would infer tal much as musically untrained Western listeners infer metre. While the conception of metre in Western culture may not be a musical universal, the hierarchical organization of temporal information may be a musical universal.

### Examples of hierarchical structure in Australian Aboriginal song

In Central Australia the song series or songline is the largest scale traditional musical form (Barwick 1989). A song series describes the Dreaming journey of one or more ancestral beings. A song performance consists of a number of song items, each of about 30 seconds duration, interspersed with periods of informal discussion. Central Australian songs are characterized by several exact repetitions of a fixed word string or text set to an unvarying rhythmic pattern. The textual and rhythmic components are regarded as two facets of the same structure (Barwick 1989). Each rhythmically articulated text relates to a different stage of the journey described in the associated myth. It is set to a melodic contour said by performers to be the same throughout the series and to embody the essence of the ancestral being whose journey through the country creating and naming present-day geographical features is celebrated in the series. For a performance to be classified as a 'song' it must correctly interlock all three elements—melody, rhythm and text (Ellis 1983). Unison group performance is the ideal and is made possible by inexperienced performers joining in when they have grasped the text words and the appropriate rhythm; they follow melodic cues given by the song leader who performs solo for the first few syllables of each item and who cues changes of pitch within the

melody slightly in advance of the group (Barwick 1989). The general melodic shape of the songs in the centre of Australia is, almost without exception, terraced (Ellis 1984); that is, a generally descending melodic line.

Performance involves the structuring of melodic contour according to various textual characteristics. Barwick (1989) argues that the melodic structure and rhythmic/textual structure are conceptually independent despite being performed in a song as an intermeshed sonic stream. Either element is able to be performed in isolation. Some evidence of this includes a) words to describe humming the sound of the Dreaming; b) rhythmically defined humming of verses in order to remember the words; c)wordless definitions of the melody; and d) instances of tapping out the rhythm of the words of a song (Barwick 1989). Thus memory for melody and rhythm is integrated but separable. Barwick describes hierarchical structures in both the melodic and rhythmic/textual dimension of central Australian song referring specifically to the *ngintaka* song series performed by Antakarinya, Yankunytjatjara and Pitjantjatjara people.

Melodic structure refers to the patterns of pitch relationships that recur in the different song items in a series, and the ways in which these recurring tonal patterns are organized into melodic sections, pitch areas, and main tones. In *ngintaka* songlines, there are three related forms of the basic melodic contour: a linear form, cyclical form, and a transposing cyclical form. Barwick describes and illustrates the hierarchical organization of the linear melodic form consisting of an introductory section, preliminary descent, main descent, and concluding section.

The rhythmic organization of the text of *ngintaka* is also hierarchical, reflecting organization of the pattern into text line pairs, text lines, rhythmic segments, and beating cells. The interdependence of melodic and temporal structures in *ngintaka* is reflected in the observation that the decision about which form of the melody to perform, its constitution in melodic sections, and tonal and durational realizations of the sections correlate with different properties of the text to be performed. The choice of linear, cyclical or transposing cyclical melodic form depends on rhythmic organization. Rhythm serves as a mnemonic device for recall of the extramusical information that is associated with a particular song text (Ellis 1984). Barwick (1989) examines the 'point of fit' in combining melodic and rhythmic/textual structures:

> although an item of singing need neither begin nor end at significant boundaries in the rhythmic/textual cyclical, there is always at least one internal point in the sung item at which the major rhythmic/textual and melodic boundaries coincide.
>
> Barwick (1989, p. 19)

This meticulous musicological and structural analysis of *ngintaka* points to the role of *hierarchy* as both an organizational and perceptual process:

> choices relating to the upper levels of the melodic hierarchy (melodic form and section composition), which correlate with structures at the upper levels of the rhythmic/textual hierarchy (overall style of beating accompaniment and repetition organization), have ramifications that extend down to the smaller level decisions about movement around main tones. Performance of the music therefore requires a constant shifting of attention from one level of the system to another.
>
> Barwick (1989, p. 26)

Cognitive processes related to hierarchical organization, grouping and segmentation, expectancy, and entrainment are evident in this example from Australian Aboriginal song.

## Development of musical expectancies and abstract implicit knowledge

### Musical expectancies

A piece of music in a familiar genre generates expectations based on implicit knowledge about common features of the genre; variations in how these expectations are met or violated are important in determining emotional and aesthetic response to music (Meyer 1956; Tillmann *et al.* 2000). In Western tonal music, musical expectancies are shaped by rhythmic or metric structure (Large and Jones 1999), tonal and harmonic structure (Schmuckler 1989) and

melodic structure (Cuddy and Lunney 1995). While it is unclear how broadly these principles apply cross-culturally, the tendency for large intervals to descend is evident in Australian Aboriginal music, Chinese folk songs, traditional Korean music, Ojibwa, Pondo, Venda, and Zulu songs (Huron 2006, although see Vos and Troost 1989), suggesting that melodic expectancies are important in emotional and aesthetic response to music in many cultures.

Differentiation of two kinds of expectations—schematic and veridical—solves a paradox of musical expectation (Bharucha 1994). A familiar piece of music contains no surprises—if the violation of expectations is aesthetically or emotionally important, a piece should become less aesthetically or emotionally interesting with each listen. This is not the case. Schematic expectations are automatic, culturally generic, and develop from assimilation of the music of a genre over years of experience. Veridical expectations refer to the actual next event in a familiar piece even though the event may be schematically unexpected. As schematic expectations are acquired from hearing many individual pieces the two kinds of expectancies will converge but, at times, will diverge, creating the continuing sense of violation in familiar pieces of music as described by Meyer (1956). The capacity to develop expectancies may be a universal process. Huron (2006) posits that implicit, statistical expectations are linked directly to emotional response, eliciting basic neurobiological responses. In his account, statistical learning (e.g., McMullen and Saffran 2004) underlies the development of musical expectations.

A small number of studies have examined expectations in music cross-culturally. Researchers have found listeners are sensitive to melodic expectancies in British folk songs, Webern Lieder, and Chinese pentatonic songs. On occasion, performance of American and Chinese participants has been compared and principles such as 'pitch proximity'—that a second pitch following a first pitch is likely to be close in pitch—and 'pitch reversal'—that listeners expect the direction of the melody to reverse—successfully predicted response patterns regardless of the musical style, or the formal music training, and cultural background of the participants (Schellenberg 1997).

Krumhansl (2000; Krumhansl *et al.* 2000) conducted investigations of an indigenous music of the Scandinavian peninsula, the North Sami yoiks, which are quite distinct from Western tonal music. These experiments used North Sami yoiks as stimuli, with participants including North Sami, Finnish music students who had studied yoiks, and North American musicians unfamiliar with yoiks. Results suggested that Western listeners were most influenced by Western schematic expectations, and that veridical expectations were strongest for Sami participants. However, statistical models of melodic expectancies, specifically melodic continuations, appear robust in explaining melodic expectancies in listeners of North Sami, Finnish or North American cultures. While these findings warrant investigation in other cultures, they nonetheless suggest that melodic expectations are universal, insofar as music is melodic.

### Implicit knowledge of musical structures

Abstract, implicit knowledge of musical structures and conventions is acquired by infants and children through mere exposure to a particular musical environment (Bigand *et al.* 2003). Some people explicitly learn the musical theory of their culture, but most acquire implicit knowledge of musical structures and conventions through incidental exposure to the environment.

### Implicit learning and Australian Aboriginal song

The process of learning via exposure is evident in descriptions of the performance of Australian Aboriginal song. For example, Barwick (2002) discusses the way rhythmic mode in *Church Lirrga* songs of Wadeye in the Northern Territory is brought into relief by the juxtaposition of contrasting songs: 'The pairing of items in different tempo bands is one of the main ways in which learners become aware of the importance of this dimension of performance' (p. 81). The pairing of slow and fast songs is commonly found in northern Central Australia and the Kimberleys (Barwick 2002). Barwick (1989) alludes to the complexity of dimensions,

relations and interrelations that are learned by an Aboriginal song leader:

> in addition to the mass of geographical and mythological detail to be mastered, leading a performance requires a high degree of sophisticated musical awareness, an awareness that is not taught directly, as in western educational practices, but is rather arrived at and internalized through repeated acts of performance.
>
> Barwick (1989, pp. 26–27)

## Temporal expectancies, synchrony, and entrainment

Temporal entrainment is a domain-general mechanism and likely universal process in music perception and production (e.g., Clayton *et al.* 2005; Drake and Bertrand 2001; Fraisse 1982; see also Chapter 8 this volume). Entrainment occurs where two rhythmic processes interact with each other, eventually locking in to a common phase and/or periodicity. Such processes play a role in speech production (Goldstein *et al.* 2006) and models of joint action (e.g. Knoblich and Jordan 2003). Large and Jones (1999) argue that neural oscillators entrain to external events such as metres or rhythms. Because this entrainment process underlies temporal perception, periodic events (i.e., repeating rhythms, metre) facilitate the efficient allocation of limited attentional resources, and make synchronization activity (e.g., tapping along or musical activity) more accurate.

Non-periodic metres are common cross-culturally. 'Free rhythms' such as the Indian *alap* or the Turkish *taksim* are improvised and unmetred, while music with accelerating rhythms is found in Japanese Gagaku music and Tibetan monastic music (Huron 2006). The *aksak* metres (e.g., 3 + 2 + 2) of Bulgarian dances (Moelants 2006) and African rhythms are non-periodic. At first glance, non-periodic rhythms may be problematic for theories of entrainment. Jones argues that when there is no explicit pulse a listener might entrain to what they perceive as the median period length (Barnes and Jones 2000). Periodicity is not necessary for the formation of expectations, according to Huron (2006). Instead, listeners need to be experienced with the temporal structure of an event and some element of its temporal structure must be predictable. Regardless of whether accounts of entrainment or accounts of expectation are better at explaining the perception of temporal structure cross-culturally, some form of allocation of attentional resources based on the expectation of temporal information is likely to be universal.

## Multimodal processes and integration

In many cultures, music and dance are inseparable. An analysis of the didjeridu-accompanied dance-song genre *Lirrga* from Wadeye in the Northern Territory of Australia, has shown that songs fall into named distinct tempo ranges which correlate with different metres in the vocal part resulting in rhythmic mode (Barwick 2002). The significance of rhythmic modes comes from their association with *dance* (Marett 2005). Thus, one universal process in music cognition might be movement perception and its development (Fraisse 1982; Friberg and Sundberg 1999).

Bharucha *et al.* (2006) have distinguished four primary facets of internal experiences of motion in music cognition: a fundamental sense of self-motion through space; perception of the motion of bodies, not necessarily one's own; movement elicited by abstract structures such as tonality; and a metaphorical, synaesthetic sense of motion. Movement in pitch, time, loudness, and timbre has been described as the interplay of tension and relaxation (Jackendoff and Lerdahl 2006). This metaphor of tension and relaxation may emerge as a universal quality in music cognition given further research cross-culturally.

Music is an example of a perception–action cycle (Janata and Grafton 2003; Keller *et al.* 2007); that is, streams of sensory information forming the basis of goal-directed actions. Evidence of neural simulation and mirroring continues to accumulate—seeing *or* hearing an intentional *action* gives rise to neural activity that is comparable with that underlying *performance* of the action (Keysers *et al.* 2003). Studies of music perception and performance are likely to shed new light on universal processes that are the result of tightly coupled sensorimotor systems.

## Conclusion

There is much to do to test assumptions of contemporary theories of music cognition and to identify unique features of musical environments and musical processing. More than 30 years ago, Harwood (1976) called for an emphasis in research on musical processes rather than musical content and suggested that looking at how people learn to listen to and play their community's music is more informative than studying what it is they listen to or play. Researchers will need to be creative in applying a range of appropriate methods to music processing in non-Western cultures; these range from initial ethological and ethnographic phases, to analysis of everyday, multimodal musical events, through to the design of experiments that test particular hypotheses. Clarke and Cook (2004) detail methods for this endeavour. As Cole (1996) notes, such studies will need to 'maintain the integrity of the real-life situations [they are] designed to investigate, be faithful to the larger social and cultural contexts from which the subjects come, and be consistent with ... participants' definitions' (p. 226). Interdisciplinary research that knits together the concerns and methods of the cognitive sciences with those of cultural anthropology, and that involves diverse materials, tasks, and cultural groups, is essential for explanatory and inclusive theories of music cognition, perception, and production.

## Acknowledgments

Supported by an Australian Research Council Discovery Project grant (DP0771890) awarded to the first author and a University of Western Sydney Postgraduate Award held by the second author. We thank Nigel Nettheim, Roger Dean, Barbara Tillmann, Rudi Črnčec, Caroline Jones, and the MARCS Music and Movement Writing Group for helpful comments on an earlier draft.
kj.stevens@uws.edu.au t.byron@uws.edu.au
http://marcs.uws.edu.au

## References

Arom S and Fürniss A (1993). An interactive experimental method for the determination of musical scales in oral cultures: application to the vocal music of the Aka Pygmies of Central Africa. *Contemporary Music Review*, **9**, 7–12.

Balzano GJ (1980). The group-theoretic description of 12-fold and microtonal pitch systems. *Computer Music Journal*, **4**, 66–84.

Balzano GJ (1982). The pitch set as a level of description for studying musical pitch perception. In M Clynes, ed., *Music, mind and brain: the neuropsychology of music*, 321–351. Plenum Press, New York.

Barnes R and Jones MR (2000). Expectancy attention and time. *Cognitive Psychology*, **41**, 254–311.

Barwick L (1989). Creative (ir)regularities: the intermeshing of text and melody in performance of Central Australian song. *Australian Aboriginal Studies*, **1989**/1, 12–28.

Barwick L (2002). Tempo bands, metre and rhythmic mode in Marri Ngarr 'Church Lirrga' songs. *Australasian Music Research*, **7**, 67–83.

Bharucha JJ (1994). Tonality and expectation. In R Aiello and JA Sloboda, eds, *Musical perceptions*, 213–239. Oxford University Press, New York.

Bharucha JJ, Curtis M and Paroo K (2006). Varieties of musical experience. *Cognition*, **100**, 131–172.

Bigand E, Poulin B, Tillmann B, Madurell F and D'Adamo D (2003). Sensory versus cognitive components in harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance*, **29**, 159–171.

Bregman A (1990). *Auditory scene analysis*. The MIT Press, Cambridge, MA.

Burns EM (1999). Intervals, scales, and tuning. In D Deutsch, ed., *The psychology of music*, 215–264. Academic Press, San Diego, CA.

Carterette EC and Kendall RA (1999). Comparative music perception and cognition. In D Deutsch, ed., *The psychology of music*, 725–791. Academic Press, San Diego, CA.

Castellano MA, Bharucha JJ and Krumhansl CL (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General*, **113**, 394–412.

Clarke EF and Cook N (eds) (2004). *Empirical musicology: aims, methods, prospects*. Oxford University Press, Oxford.

Clayton M (1997). Metre and tal in North Indian music. Translated from Le metre et le tal dans la musique de l'Inde du Nord *Cahiers de Musiques Traditionelles*, **10**, 169–189. Trans. G. Goormaghtigh.

Clayton M, Sager R and Will U (2005). In time with the music: the concept of entrainment and its significance for ethnomusicology. *European Meetings in Ethnomusicology (ESEM Counterpoint* **1**), 3–75.

Cole M (1996). *Cultural psychology: a once and future discipline*. The Belknap Press of Harvard University Press, Cambridge, MA.

Cowan N (1984). On short and long auditory stores. *Psychological Bulletin*, **96**, 341–370.

Cuddy LL and Lunney CA (1995). Expectancies generated by melodic intervals: perceptual judgements of melodic continuity. *Perception and Psychophysics*, **57**, 451–462.

Darwin CJ, Turvey MT and Crowder RG (1972). An auditory analogue of the Sperling partial report procedure: evidence for brief auditory storage. *Cognitive Psychology*, **3**, 255–267.

Deutsch D (1999). Grouping mechanisms in music. In D Deutsch, ed., *The psychology of music*, 299–348. Academic Press, San Diego, CA.

Dowling WJ (1978). Scale and contour: two components of a theory of memory for melodies. *Perception and Psychophysics*, **14**, 37–40.

Drake C and Ben El Heni J (2003). Synchronizing with music: intercultural differences. *Annals of the New York Academy of Sciences*, **999**, 429–437.

Drake C and Bertrand D (2001). The quest for universals in temporal processing in music. In RJ Zatorre, I Peretz, eds, *The biological foundations of music: Annals of the New York Academy of Sciences, Vol. 930*, 17–27. The New York Academy of Sciences, New York.

Ellis CJ (1983). When is a song not a song? A study from Northern South Australia. *Bikmaus*, **4**, 136–144.

Ellis CJ (1984). The nature of Australian Aboriginal music. *The International Journal of Music Education*, **4**, 47–50.

Everett DL (2005). Cultural constraints on grammar and cognition in Pirahã: another look at the design features of human language. *Current Anthropology*, **46**, 621–634.

Fraisse P (1982). Rhythm and tempo. In D Deutsch, ed., *The psychology of music*, 149–180. Academic Press, New York.

Friberg A and Sundberg J (1999). Does music performance allude to locomotion? A model of final ritardandi derived from measurement of stopping runners. *Journal of the Acoustical Society of America*, **105**, 1469–1484.

Goldstein L, Byrd D and Saltzman E (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. In MA Arbib, ed., *Action to language via the mirror neuron system*, 215–249. Cambridge University Press, Cambridge.

Handel S (1984). Using polyrhythms to study rhythm. *Music Perception*, **1**, 465–484.

Harwood DL (1976). Universals in music: a perspective from cognitive psychology. *Ethnomusicology*, **20**, 521–533.

Huron D (2006). *Sweet anticipation: music and the psychology of expectation*. The MIT Press, Cambridge, MA.

Iyer VS (1998). Microstructures of feel, macrostructures of sound: embodied cognition in West African and African-American musics. Unpublished doctoral dissertation, Technology and the Arts, University of California, Berkeley.

Jackendoff R and Lerdahl F (2006). The capacity for music: what is it, and what's special about it? *Cognition*, **100**, 33–72.

Janata P and Grafton ST (2003). Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience*, **6**, 682–687.

Jusczyk PW and Krumhansl CL (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception and Performance*, **19**, 627–640.

Justus TC and Bharucha JJ (2002). Music perception and cognition. In S Yantis, volume ed., and H Pashler, series ed. *Stevens' handbook of experimental psychology, Vol. 1: Sensation and perception*, 3rd edn, 453–492. Wiley, New York.

Keller PE, Knoblich G and Repp BH (2007). Pianists duet better when they play with themselves: on the possible role of action simulation in synchronization. *Consciousness and Cognition*, **16**, 102–111.

Kessler EJ, Hansen C and Shepard RN (1984). Tonal schemata in the perception of music in Bali and the West. *Music Perception*, **2**, 131–165.

Keysers C, Kohler E, Umiltà MA, Nanetti L, Fogassi L and Gallese V (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, **153**, 628–636.

Knoblich G and Jordan JS (2003). Action coordination in groups and individuals: learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **29**, 1006–1016.

Kolinski M (1970). Review of 'Ethnomusicology of the Flathead Indians' by Alan P. Merriam. *Ethnomusicology*, **14**, 77–99.

Krumhansl CL (2000). Tonality induction: a statistical approach applied cross-culturally. *Music Perception*, **17**, 461–479.

Krumhansl CL and Jusczyk PW (1990). Infant's perception of phrase structure in music. *Psychological Science*, **1**, 70–73.

Krumhansl CL, Toiviainen P, Eerola T, Toiviainen P, Järvinen T and Louhivuori J (2000). Cross-cultural music cognition: cognitive methodology applied to North Sami yoiks. *Cognition*, **76**, 13–58.

Large E and Jones MR (1999). The dynamics of attending: how people track time-varying events. *Psychological Review*, **106**, 119–159.

Lomax A (1968). *Folk song style and culture*. American Association for the Advancement of Science, Washington, DC.

McAdams S (1989). Psychological constraints on form-bearing dimensions in music. *Contemporary Music Review*, **4**, 181–198.

McMullen E and Saffran JR (2004). Music and language: a developmental comparison. *Music Perception*, **21**, 289–311.

Magill JM and Pressing JL (1997). Asymmetric cognitive clock structures in West African rhythms. *Music Perception*, **15**, 189–222.

Marett A (2005). *Songs, dreamings, and ghosts: the Wangga of North Australia*. Wesleyan University Press, Middletown, CT.

Meyer LB (1956). *Emotion and meaning in music*. Chicago University Press, Chicago, IL.

Meyer LB (1960). Universalism and relativism in the study of ethnic music. *Ethnomusicology*, **4**, 49–54.

Moelants D (2006). Perception and performance of aksak metres. *Musicae Scientiae*, **X**, 147–172.

Nam U (1998). Pitch distributions in Korean Court music: evidence consistent with tonal hierarchies. *Music Perception*, **16**, 243–247.

Nan Y, Knösche TR and Friederici AD (2006). The perception of musical phrase structure: a cross-cultural ERP study. *Brain Research*, **1094**, 179–191.

Nettl B (2005). *The study of ethnomusicology: thirty-one issues and concepts*. University of Illinois Press, Champaign, IL.

Patel AD, Peretz I, Tramo M and Labreque R (1998). Processing prosodic and musical patterns: a neuropsychological investigation. *Brain and Language*, **61**, 123–144.

Schellenberg EG (1997). Simplifying the implication–realization model of melodic expectancy. *Music Perception*, **14**, 295–318.

Schellenberg EG and Trehub SE (1996). Natural musical intervals: evidence from infant listeners. *Psychological Science*, **7**, 272–277.

Schmuckler MA (1989). Expectation in music: investigation of melodic and harmonic processes. *Music Perception*, **7**, 109–150.

Schwartz DA, Howe CQ and Purves D (2003). The statistical structure of human speech sounds predicts musical universals. *The Journal of Neuroscience*, **23**, 7160–7168.

Sethares WA (1998). *Tuning, timbre, spectrum, scale*. Springer-Verlag, Berlin.

Stobart H and Cross I (2000). The Andean anacrusis? Rhythmic structure and perception in Easter songs of Northern Potosí, Bolivia. *British Journal of Ethnomusicology*, **9**, 63–94.

Tenney J (1988). *A history of consonance and dissonance*. Excelsior Music Publishing Co., New York.

Tenzer M (2006). Oleg Tumulilingan. Layers of time and melody in Balinese music. In M Tenzer, ed., *Analytical studies in world music*, 205–236. Oxford University Press, Oxford.

Thompson WF and Schellenberg EG (2006). Listening to music. In R Colwell, ed., *MENC handbook of musical cognition and development*, 72–123. Oxford University Press, Oxford.

Tillmann B, Bharucha JJ and Bigand E (2000). Implicit learning of tonality: a self-organizing approach. *Psychological Review*, **107**, 885–913.

Trainor LJ (2006). Innateness, learning and the difficulty of determining whether music is an evolutionary adaptation. *Music Perception*, **24**, 105–109.

Trehub SE (2000). Human processing predisposition and musical universals. In NL Wallin, B Merker, S Brown, eds, *The origins of music*, 427–448. The MIT Press, Cambridge, MA.

Trehub SE and Hannon EE (2006). Infant music perception: domain-general or domain-specific mechanisms? *Cognition*, **100**, 73–99.

Vos PG and Troost JM (1989). Ascending and descending melodic intervals: statistical findings and their perceptual relevance. *Music Perception*, **6**, 383–396.

Will U (1997). Two types of octave relationships in Central Australian vocal music? *Musicology Australia*, **XX**, 6–14.

Wright AA, Rivera JJ, Hulse SH, Shyan M and Neiworth JJ (2000). Music perception and octave generalization in rhesus monkeys. *Journal of Experimental Psychology: General*, **129**, 291–307.

# CHAPTER 3

# Music and meaning

Ian Cross and Elizabeth Tolbert

## Introduction

What do we mean by 'meaning'? When we say that something has 'meaning', we are claiming that our original something points to, or is attached to, or can be used to infer the existence of, some other thing beyond itself. Some sort of relationship exists between our original entity or event and something beyond itself. Vernacularly, our original something has *significance*. Another simple way of putting it is that our original something *refers* to something else beyond itself. When we engage with music, we tend to feel that it has significance; it appears to mean something, even if that meaning is entirely personal to us. When we look at the ways in which music has been theorized and explored, the issue of musical meaning is often central, particularly when music is being explored in its social context (see Feld and Fox 1994). However, while music may have significance, by comparison with language it is very difficult to say what it is that music refers to, or indeed, whether music refers to anything else at all.

Contemporary methods of approaching the question of musical meaning have been shaped by a range of theories of meaning. Some theories hold that music is a species of art, which can only be understood as a phenomenon that is not susceptible to analysis in terms of general and formally expressible principles. Such theories tend to be predicated on the notion that recent and contemporary Western conceptions of music—in which music is produced by specialists for aural consumption by the many—are representative of music as a generic human phenomenon. Alternative views are found in literatures that aim to explore music as a social phenomenon (whether in Western or non-Western contexts), which tend to be pluralistic in the frameworks that they apply to articulate their understandings of meaning in music, drawing on semiotic, sociological and anthropological theories of meaning. Some influential theories that have emerged within analytical philosophy and that are prevalent in the study of linguistics seem to imply that meaning cannot properly be considered a property of music. In these theories meaning is held to be a property only of (formal) systems that are capable of embodying two types of meaning relation, *reference* and *sense* (see below), and that are thus capable of articulating complex propositions that are definable as true or false. The chapter will present an overview of theories of meaning that have been, and that may be, applicable to investigating music, particularly its cognitive dimensions.

## Theories of meaning

Philosophers and logicians have devoted much attention to the problem of meaning—to analysing how it is that something means. In general, they have focused on meaning as it appears manifested through language. Their ideas, particularly in the form of theories of semantics, have tended to dominate conceptions of meaning within philosophy and within most sciences. Semantic theories generally aim to account for the relationships between the phenomena that constitute the objects of meaning and the expressions and sets of expressions that are used to articulate meanings.

One of the strongest influences on the development of recent and contemporary theories of the semantics of language has been the work of

Frege, who wrote at the end of the nineteenth century. Frege suggested that there are two necessary aspects of meaning that must be considered in any analysis; these he called *reference* and *sense*. Roughly speaking, the *reference* of a term is that which it denotes—that phenomenon in the world onto which it can be mapped. Its *sense* derives from the ways in which that original term relates to other terms that are interpretable as capable of bearing similar kinds of meanings. To give Frege's classic example, the expressions 'Morning Star' and 'Evening Star' have the same single reference—the planet Venus. But their senses differ in an obvious way, in that the sense of one is 'the star that appears in the morning', and that of the other 'the star that appears in the evening'. The senses of the two expressions arise by virtue of the meanings of the terms that are used to articulate them within the language system.

Hence in this view meaning is a consequence of the capacity of terms within a system to refer and at the same time be bounded by their relationships to other terms within a system of terms such as a formal logic (or a language when it is interpreted as having features that of the same type as those of a formal logic). These features enable terms within the system to be used to articulate propositions that may be evaluated as true or false—that have determinate *truth values*. Only in such systems can complex propositions be expressed and can their meaning be fully accounted for in terms of the elements that make up the complex propositions; only such systems have the capacity for *compositionality*, that of expressing semantically decomposable propositions. It has been suggested that only formal semantic theories founded on Frege's ideas, such as those of Montague or of Tarski, constitute 'genuine theories of meaning' (Scruton 1987). While these types of theories do not seem particularly pertinent to music (particularly in view of the difficulty that the ideas of reference, and of compositionality, pose for music), they are nonetheless highly relevant in that they have underpinned the development of theories of meaning throughout the twentieth century that has shaped much of the intellectual context in which any account of meaning in music has to be understood (for more detailed and technical considerations of Frege's ideas, see Grayling 2000, Vol. 1, Chapter 2; Vol. 2, Chapter 12).

A rather different approach to meaning that has been widely applied to music derives from the work of CS Peirce (see Hoopes 1991), who proposed that meaning could be understood in terms of systems of signs, which can be defined as *natural* or *conventional* and can stand in a variety of relationships to that which the sign represents—its meaning. A sign may point beyond itself by virtue of some formal resemblance between the sign and that which is represented, and thus be *iconic*; it may have meaning because of some necessary connection between itself and what it indicates (as in the case of smoke indicating fire) or by virtue of repeated co-occurence, the relationship between sign and its meaning thus being *indexical*; and it may mean by virtue of arbitrary convention, its meaning being rooted in the *symbolic* domain. While Peirce's *semiotic* ideas have largely been neglected or rejected by philosophers and semantic theorists, they have played a role in several influential theories of meaning in music (e.g. Nattiez 1990). Nevertheless, music does not easily map onto the natural–conventional distinction that lies at the root of Peirce's original theories. Music has tended to be viewed as having an ambiguous status somewhere between the natural and the conventional, limiting the explicitness with which the bases for its meanings can be articulated in semiotic terms.

It may seem that meaning is simply a mapping from a term to its referent (and vice versa). However, for Peirce, the source of meaning is in the mind; in his theory, 'All thought is in signs' (Peirce quoted in Violi 1999). The idea that consideration of mind is necessary to theories of meaning lies at the heart of contemporary cognitive science and is central in much contemporary philosophy of meaning and language. Mental states and processes are said to exhibit *intentionality* or aboutness (Dennett 1987, 1995); they represent, or point to, phenomena beyond themselves. The idea of intentionality can be applied to understand the ways in which language can be interpreted as bearing meaning, by taking into account the ways in which we employ language to articulate the aboutness of our mental states.

A more radical account which explicitly sets out to provide an alternative to the idea that meaning can only be correctly defined in relation to the abstract properties of formal systems is found in *experientialist* approaches, of which the most developed is that of Lakoff and Johnson (Johnson 1987; Lakoff 1987). In this view, meaning is a property of the relationships between mind and world; however, it is best understood as mediated not by the principles of formal logic but by the embodied nature of our experiences of being in the world. Meaning is motivated by kinesthetic-image schemas, which are (Lakoff 1987, p. 267) 'relatively simple structures that constantly recur in our everyday bodily experience', and which may take forms such as 'CONTAINERS, PATHS, LINKS etc.' These body-image schemas constitute the preconceptual bases for the formation and combination of concepts, grounding meaning in our bodily experience of the world. A somewhat different approach to an experientialist theory of meaning is proposed by Jackendoff (1987, 2002), although directed more explicitly towards making sense of language structure and use. Experientialist theories, in particular those of Lakoff and Johnson, have been applied to understanding the nature of meaning in music, as we shall see.

Ideas such as reference and sense, natural and conventional signs, and intentionality can help account for different aspects of the ways in which something may be interpreted as bearing meaning. Yet music seems to pose particular problems for these ways of approaching meaning. Usually, we cannot use music to refer to a state of affairs in the world in the ways in which we can use language. In at least this respect, music certainly does not seem capable of bearing the types of meaning borne by language, which would suggest that an analysis of musical meaning in terms of sense and reference is unfeasible. Musical meaning also fits ambiguously with any distinction between natural and conventional signs, and it is also unclear how the notion of intentionality might be applied to clarify the aboutness of the mental states that arise when we engage in music.

In order to elucidate the ways in which meaning in music can be, and has been, explored in ways that are relevant to music (and to the psychology of music), it is necessary to consider what are the kinds of things that can be the objects of meaning. Are they states of affairs in the material world or in possible worlds (for which see Johnson-Laird 1983, p. 172)? Are they states of affairs in some abstract yet immanent domain, such as the theorems of formal logic and their validity—their truth values—within some logical system (Tarski 1956)? Are they states of affairs in some abstract yet immanent domain that *cannot* be described within the theorems of formal logic, such as the domain of the aesthetic (Sibley 1959)? Are they mental phenomena such as the thoughts, beliefs and desires of individuals (Dennett 1987), or the preconceptual schemas proposed by Lakoff and Johnson? Or are they social phenomena, such as cultural groupings, institutions, contracts or shared understandings (Bourdieu 1990)? At different times, theories of meaning have postulated all these types of phenomena as being potential or actual objects of meaning, and in one way or another, all have been applied to understanding the question of meaning in music over the history of Western thought.

## Meaning in music in the western intellectual tradition

For the classical Greek philosophical tradition, music presented two conflicting faces. On the one side, music—and music, for the Greeks, was melodic—moved the passions (for good or ill), and was a fundamentally human activity, affecting our emotions by being stirring and martial, or beautiful and enticing. Its purposes could thus be considered as aligned with those of the science of rhetoric, which aimed to bend human minds and actions to the purposes of the orator. On the other side, music, in the theories erected on the basis of the measurement of aspects of the sounds that comprised it, embodied the natural laws of number and could thus be viewed as reflecting abstract and immanent aspects of the universe such as the principles of natural order, or the workings of the divine (James 1993).

This dichotomous view of music persisted through the first millennium CE, gaining in complexity as it was refracted through the multiple

prisms of early Christian thought. Hence in the medieval world, music could be thought of as having meaning in at least two domains, the human and the divine. The tension between music as mirror of the heavenly world and music as mover of earthly passion was mediated by church doctrine and by the development of musical notations, through which polyphonic music that might be viewed as simply lascivious because of its sonorous beauty could alternatively be interpreted as reflecting the mathematical complexity of divine order (Sparshott and Goehr 2001).

By the later fifteenth century, music came to be seen in part as a foundationally human activity with its values and its meanings requiring to be expressed and interpreted in human rather than divine terms (Palisca 1985). Music's meanings were bound to the products and processes of human reason as well as to music's effects on human emotions, allowing for the possibility that music might express or bear types of meanings that were proper only to music as an art. By the later seventeenth and early eighteenth centuries, music's meanings had come to be largely theorized in terms of human passions or affects. Music's alignment with rhetoric was foregrounded (Sparshott 1998), as was the extent to which its forms mirrored those of the linguistic prosody, though the structures that music could articulate also became more important for their own sake. The values accorded to music and its meanings were largely those of rational pleasure, enjoyment founded on the exercise of reason; ostensibly rational principles (based in part on the development of the science of acoustics) were elaborated and adduced that were intended to account for the extent to which a piece of music had fulfilled its brief by expressing its meaning.

Through the eighteenth century the forms and the meanings of music changed radically. From being bound to the devices of rhetoric and hence to its purposes, music's forms became more and more intelligible in terms of theories of harmony, related to either, and sometimes to both, the findings of physical acoustics, and abstract principles of architectonic structure. Musical meaning thus no longer required any sort of justification by reference to the words that it could convey in song, or to the ways in which it conveyed those words (its 'prosody'); instrumental music came to be conceived of as equally capable of bearing meaning in its own right. However, the effects of music on human passions were not completely displaced from consideration; as Thomas (1995, p. 5) notes, most of the writing on music by the philosophers of the mid- to late-eighteenth century had as a 'governing assumption' the idea 'that music was a kind of language'. Hence by the late eighteenth century, the objects of the meanings of music could be conceived of as private, being in the mind, or as public and social.

## The aesthetic perspective

Towards the end of the eighteenth century a new way of understanding meaning in music emerged, one which had its precursors in the notion that music should be valued for its beauty but which sought to detach music's meaning from its social value and to valorize its experience as art as an end in itself. It developed in tandem with the emergence of the notion of music as being autonomous, or as having value in its own right. The ideas that underpin this approach have come to constitute the prevalent means of addressing questions of music's value and significance within the Western intellectual tradition of the last two centuries or so.

This new approach was first crystallized in ideas that were developed by Hume and by Kant, in proposing that both art and natural beauty can give rise to aesthetic experiences that are exclusive in that they are bound to felt response (in being logically dependent on pleasure and liking) but that are not reducible to any specific set of principles (as they cannot be induced second-hand by verbal description). Kant proposed that aesthetic contemplation is distinguished by being fundamentally 'disinterested', or free from concern with desire or any other purposive attitude. In current approaches the notion of the aesthetic is bound up with the notion that there are phenomena that can only be described as 'art', and various attempts have been made to demarcate the qualities that art must have in order to be art (see e.g. Wollheim 1980). The identity of art has been variously proposed as lying in its capacity for representation, its structural qualities, or in its capacities for expression. In expression-based approaches, works of art are viewed as the loci for aesthetic

experience by virtue of encapsulating both emotional and rational qualities that may be more or less apparent to, or recuperable by, engaged audiences (see Davies 1994). In these theories, music's capacity to engender aesthetic experience does not rely on, and is not expressible in the same terms as, the capacity that language possesses of bearing meaning by expressing complex propositions that have determinable sense and reference (Dempster 1998).

Central to many expressivist analyses of musical meaning has been the idea that music's meaning is fundamentally emotional, though most of those who have sought to develop this approach have treated the relationship between music and emotion in ways that appear to have little connection with the scientific literature on music and affect. Two leading philosophers of music, Roger Scruton and Stephen Davies, both privilege the expressive dimensions of music in seeking to analyse its aesthetic qualities, though taking somewhat diverging paths. Davies (1994, 2003) proposes that a critical component of the aesthetic aspect of music is its engendering of an 'understanding response' in a listener, which requires the capacity to experience predictive coherence in the ongoing flow of a piece of music as heard. In addition to this aspect, music must be recognized as referring beyond itself in being expressive of emotion; music achieves this either by presenting the appearance of emotion (Davies 2003) or because we (1994, p. 277) 'experience the dynamic character of music as like the actions of a person'. The experience of music must involve both types of recognition to qualify as an appropriate aesthetic experience; only if both conditions are met will the music be understood, its 'meaning' grasped. Scruton (1987, 1997) suggests that music does not so much express emotion as embody it in a holistic way such that it is not possible to identify just which features of the music are directly responsible for this embodiment of emotion; he suggests that 'In responding to a piece of music we are being led through a series of gestures which gain their significance from the intimation of community' (1997, p. 357). However, for this aspect of meaning in music to be experienced, audiences must be capable of hearing the music *as* music. Thus for Scruton musical meaning is rooted in the metaphorical experience of movement and is situated in *a* musical culture which aligns the intuitive capacities of audiences, composers and performers.

In theories of musical meaning that are grounded in aesthetic considerations, 'meaning' appears to take quite a different form from that which it holds in the types of semantic theories that can be applied to language. Meaning is immanent as a condition of felt response that depends on the qualities of the music as the object of listening and on the cultural capacities of its audiences. The factors that motivate musical meaning might be expressive, or might have their roots in the structural properties of Western music (particularly in the harmonic domain) that can be analysed in terms of patterns of tension and resolution. Just these types of factors were identified as the proper locus of meaning in music by Hanslick in the nineteenth century, and cognitively oriented theories of music's meaning as rooted in its structural properties have been proposed by theorists such as Lerdahl (see e.g. Lerdahl 2003). Nevertheless, the majority of philosophical approaches to understanding meaning in music have differentiated musical meanings from 'everyday' meanings by binding music's significances solely to the aesthetic domain. As Cook (2001) points out, such approaches stand in stark opposition to those which aim to locate, and to understand, music and its meanings in the social contexts in which it arises. Theories of this latter type have emerged from sociological and ethnomusicological explorations of music; they are central to approaches that adopt anthropological perspectives on the investigation of relationships between musical sounds, practices and concepts (Merriam 1964) in non-Western and, increasingly, in Western cultural contexts, to understand meaning in music.

## Meaning in music in social and cultural context

The idea that music's meanings are inseparable from the social and cultural situations and circumstances in which they arise has become widely accepted in musicological (e.g. Kramer 1995), sociological (e.g. Martin 1995) and ethnomusicological (e.g. Bohlman 2000) circles.

Here, the ideas that underpin much philosophical writing on music—that music constitutes an autonomous and primarily aesthetic domain—are called into question or outright rejected, whether on the basis that such ideas are grounded in ethnocentric assumptions as to what constitutes music, or on the grounds that such approaches cannot deal adequately with the historical and dynamic aspects of engagement with music.

The consensus surrounding the notion that music is explicable only as in its social context, evaporates when addressing the question of what *forms* these meanings may take. As Titon and Slobin (1996, p. 1) note, while 'Music is universal its meaning is not.' Across cultures there is a vast range of ways in which music may be experienced or interpreted as bearing meaning; music's meaning can be understood as foundationally aesthetic, personal or social, or in terms of combinations of any or all of these domains. Within such approaches, as Feld and Fox (1994) point out, it is difficult to disentangle specifically musical meanings from the meanings of the activities of which music is a part. For example, in writing of Venda children's music, Blacking (1967, p. 31) notes that 'Many songs add to the meaning of a social event; they crystallize and confirm certain norms of behaviour'; hence the meaning of musical activities may be thought of as complementary to the specific circumstances that incorporate the music.

From such perspectives, meaning in music is multifarious and cannot be understood solely as a consequence of music's aesthetic dimension. As Clayton (Chapter 4 this volume) indicates, music may fulfil a multiplicity of the functions in different societies; taking this into account, its meanings can be best interpreted as adverting to emic conceptions of social facts and circumstances. Hence Feld (1982) is able to present a compelling account of the music of the Kaluli of Papua New Guinea as displayed in the *gisalo* funerary ceremony, where the musical component of the ceremony possesses a structure, identity, and efficacy in mourning that are inextricably bound both to its sacred connotations (it employs a motif derived from the song of the *muni*, a local bird which is the central numinous entity in the main Kaluli myth of 'the boy who turned into a *muni* bird' and thus joined the departed spirits) and to its place in the ecology of the local environment. Hence the meaning of the music is complex and local, in that it can only be understood by addressing it in the context of the belief systems and cultural practices within which it has a role. In other cultural contexts music's meanings may be interpreted as oriented towards more mundane ends, as when Blacking notes of Venda children's music that 'Knowledge of the children's songs is a social asset and in some cases a social necessity for a child who wishes to be an accepted member of his own age group'(1967, p. 31). Here, music is interpreted as fulfilling an instrumental function in the formation of social groups, its meanings somehow bound up in the emergence, or the maintenance (cf. Slobin 1993), of individual and group identities as social facts.

In parallel with reconceptualizations of the idea of musical meaning emerging from anthropological perspectives are sociologically motivated theories that have sought to ground musical meanings in the social and historical contexts of music, particularly in Western culture. These theories have tended to conceive of music as social construct of which the meanings can only be elucidated by analysis of the processes and conditions that pertain to the contexts within which music is produced and received. Influentially adumbrated by Adorno (1976), subsequent accounts have frequently aimed to address issues of hegemony and resistance (Born and Hesmondalgh 2000), and have led to radical accounts of meaning in familiar works of the classical canon such as the work of McClary (1991) on questions of gender in music, as well as to critical and subjectivist accounts of meaning (Kramer 1995).

Meaning in music has been conceptualized variously by philosophers, musicians, anthropologists and sociologists. While the aesthetic view of music's meaning has been predominant for the last two centuries, other conceptions of the bases for meaning in music have emerged that locate these in the relationship between music and social factors and forces. Many of these theories have been explored in the context of cognitive theories of music although, as we shall see, certain views of musical meaning

present considerable obstacles to scientific theorization and investigation.

## Theoretical and empirical approaches to meaning in the psychology of music

Most attempts to explore meaning empirically by way of music's aesthetic dimension have adopted a severely reduced notion of aesthetic engagement. As noted above, philosophical approaches view aesthetic judgement as dependent on, though not explicable solely in terms of, pleasure and liking; however, experimental approaches have tended to equate aesthetic judgement wholly with evaluative ratings of liking or pleasantness. The most influential theory in the field, that of Berlyne (1971), postulated an inverted U-shape function relating liking to stimulus complexity. Berlyne's theory has been empirically tested and found to account for aspects of the experience of music (North and Hargreaves 1995); however, other studies (Orr and Olsson 2005) have found no consistent relation between liking and complexity when factors such as musical style and level of musical expertize were taken into account.

There are fundamental difficulties in empirically exploring aesthetic experiences. In the first place, there is a lack of consensus as to the phenomenal and phenomenological correlates of aesthetic experience; do these inhere in the representative, the formal, or the expressive capacities of the artistic focus of aesthetic experience, or are they consequences of the capacities of the individual who has the experience? This lack of philosophical agreement is supplemented by the difficulties of implementing experimental designs that deal adequately with the issue of 'felt response' that is at the core of ideas of aesthetic experience. Perhaps some answers might be provided by neuroscientific approaches (see e.g. Blood and Zatorre [2001] in respect of intensely pleasurable responses to music), but even here what is being identified are not the phenomenological states that are the focus of philosophical aesthetics but neurophysiological states, albeit that these may have determinable correlates in behaviour and cognition.

Nevertheless, many recent accounts of aesthetic meaning in music tend to be rooted in the idea that music is expressive, particularly of the emotions. As is evident in the theories and empirical findings described in Chapter 35, most empirical approaches to emotion or affect in music explore the phenomenon within frameworks derived from cognitive and behavioural sciences. Emotions are viewed as dynamic mind–brain–body states that arise in response to the experience of environmental and social events (see e.g. Panksepp and Bernatzky 2002). From the perspective of the cognitive sciences, emotional responses to music tend to be viewed as belonging to subclasses of emotional responses in general rather than as being phenomenally in a category of their own (or even in a category together with responses to the expressivity of other art forms). Hence the types of meaning that psychological approaches understand as arising from emotional engagement with music do not appear to rest on the same types of premises as those that have been postulated by philosophers such as Davies or Scruton as pertinent to the workings of music's expressive powers.

While it might seem that psychological understandings of music's affective meaning should have precedence over the non-empirical ratiocinations of philosophers, it might instead be the case that philosophers have in fact identified good grounds for believing that music's expressivity cannot be understood simply by reference to general theories of emotional response. Indeed, the notion of 'disinterested pleasure' that is at the heart of Kant's conception of aesthetic engagement is one that seems almost to stand in opposition to the necessary interest that is implicit in psychological understandings of the experience of emotion. As yet, questions such as whether the affective states that music appears to elicit arise through empathic processes, as Davies (2001) suggests, or arise directly in response to objective properties of musical structures have not been the sustained focus of experimental investigation.

Affective dimensions of specifically musical meanings have, however, been explored fruitfully by empirical investigation from sociological and social–psychological perspectives. The work of DeNora (2000) suggests that

music can be meaningful in being employed by listeners to do 'emotional work' in regulating their emotions and moods in everyday life. DeNora's findings are reinforced by those of Sloboda *et al.* (2001), who employed an innovative approach (the Experience Sampling Method) to sample the nature and significance of the experience of music in day-to-day life. They found that music was particularly significant in social and individual contexts in which personal choice of music was available, and that the efficacy of music in regulating emotion and mood was greatest in such contexts. Intertwined with the empirical study of the individual and social functionality of music's affective dimensions are explorations of the ways in which music is used in the formation and maintenance of senses of self and of group. A wide range of research is reported in MacDonald *et al.* (2002) that indicates that perhaps the primary motivation for engaging with music in the contemporary Western world is to facilitate the transactions involved in formulating, presenting and affirming the multiple dimensions of individual and group identities.

A very different account of musical meaning seems to emerge from some recent neuroscientific research. In an intriguing experiment, Koelsch *et al.* (2004) found that similar types of brain responses were elicited when words that were semantically incongruous were presented following either linguistic or musical contexts (for the latter type of context, that words that were used had been rated as either congruous or incongruous in respect of the musical excerpts in a preceding experiment). This would seem to indicate that musical meanings were being experienced as directly relatable to the meanings of individual words; in their words (ibid., p. 302) 'music can, as language, determine physiological indices of semantic processing'. This finding suggests that musical meanings may indeed be of a similar type to those of language, in stark contrast to most philosophical accounts of musical meaning. However, in a subsequent experiment (Steinbeis and Koelsch 2008) they suggest that the results of the earlier experiment arose because of imagery or association elicited in response to the music rather than through any intrinsically semantic properties of music. In the latter experiment they find that the semantic unexpectedness of a sentence modulated the neural response to a simultaneously presented musical excerpt that contained an unexpected chord, but that the unexpectedness of the musical event had no effect on the processing of the sentence. They take this finding to indicate that tension–resolution patterns in harmony motivate cognitive and neural responses that are analogous with the contextual integration aspects of language processing, but do not call on representations of semantic knowledge. Hence they suggest that musical 'meaning' is best conceived of as being borne by the patterns of tension and resolution embodied in music, and is of a fundamentally different order from linguistic meaning.

This notion is somewhat problematized by research that applies the experientialist ideas of meaning developed by Johnson and Lakoff. Brower (2000) and Zbikowski (2002) present complex theories of the ways in which body-image schemas may be applied to understand the ways in which music may mean, illustrating their theories by reference to detailed analyses of structures in Western art music. They suggest that while the experience of musical meaning may well be mediated to patterns of tension and resolution, these patterns are themselves motivated by underlying body-image schemas that render the patterns coherent and comprehensible to listeners. Based on the results of a series of experiments and analyses, Larson (1998, 2002, 2004) has explored the extent to which experientialist ideas can account for listeners' experience of aspects of musical structure. Larson and Van Handel (2005) note that 'physical force schemas' appear to account well for listeners' judgements of musical pattern completion, supporting the idea that the types of musical meanings supposed to inhere in tension–relaxation structures in music are in fact grounded in experientialist schemas, a finding supported by subsequent experimental work by Martinez (2007).

## Conclusions

One fact that becomes clear on surveying the literature on musical meaning is that there is no consensus as to how questions of musical

meaning should best be addressed; indeed, certain philosophical perspectives call into question the idea that music can be claimed to have meaning at all. However, the majority cross-disciplinary view is that music is imbued with meaning, whether music is conceived of as aesthetic object, cultural commodity or social process, yet each discipline tends to propagate ideas of musical meaning that tend to lack the robustness to survive outside its borders. Perhaps it is the case that musical meanings are multifarious and that there is no single approach that can claim precedence for its own conceptions of those meanings. This heterogeneity of solutions to the question of musical meaning poses problems for the psychological sciences; as we have seen, quite different depictions of musical meaning emerge from the theoretical and empirical literature surveyed above. Again, it could be suggested that this is an inevitable outcome of the diffuseness of the concept that is being investigated; nevertheless, the diversity of psychological approaches do exhibit some common themes, although it is evident that there are some conceptions of meaning in music that have not been addressed.

Musical meanings in non-Western contexts have remained largely unexplored from psychological perspectives. This is unsurprising, given the relative dearth of studies of music cognition in non-Western cultures compared to the number of studies conducted in Western contexts. Yet the issue of musical meaning can only properly be addressed by the cognitive sciences when the types of meanings highlighted in numerous ethnomusicological studies are made the foci of psychological exploration. In addition, most empirical approaches to meaning in music take as the object of their investigation the processes that are involved in music listening. While this view is consonant with the ideas of music expressed within most philosophical and many sociological approaches (a notable exception in the latter domain being the work of Finnegan [1989]), ethnomusicological research indicates that music might be better conceived of as a mode of interaction rather than as the object of auditory perception. The psychological sciences need to find means of addressing the study of the meanings that are inherent or emergent in processes of musical interaction, with a few studies beginning to indicate possible methods whereby this might be achieved (see e.g. Clayton 2007). Finally, the cognitive sciences of music will need to accommodate new approaches to meaning such as Millikan's (2004) 'teleosemantic' theory, which aims to account for meaning within an evolutionary framework (some theoretical groundwork in this area has been laid in Tolbert [2001] and in Cross and Woodruff [2008]).

The issue of musical meaning is unlikely to be accounted for by one generically applicable theory and is unlikely to be amenable to scientific exploration by any one method. This chapter has laid out some of the principal influences on the study of meaning in music and has suggested that some have had more impact on the scientific exploration of music's significance while others have been unduly neglected. Theoretical and empirical research into musical meaning is still at an early stage within the cognitive sciences; this chapter would suggest that it is important that multiple perspectives on meaning in music, particularly those emerging from the study of music in non-Western contexts, are taken fully into account in future explorations of musical meaning within the cognitive sciences.

## References


Adorno T (1976). *Introduction to the sociology of music.* Seabury Press, New York.

Berlyne DE (1971). *Aesthetics and psychobiology.* Appleton-Century-Crofts, New York.

Blacking J (1967). *Venda children's songs: a study in ethnomusicological analysis.* Witwatersrand University Press, Johannesburg.

Blood AJ and Zatorre RJ (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences*, **98**(20), 11818–11823.

Bohlman P (2000). Ethnomusicology and music sociology. In D Greer, ed., *Musicology and sister disciplines*, 288–298. Oxford University Press, Oxford.

Born G and Hesmondalgh D (eds) (2000). *Western music and others: difference, representation and appropriation in music.* University of California Press, Berkeley, CA.

Bourdieu P (1990). *The logic of practice.* Stanford University Press, Stanford, CA.

Brower C (2000). A cognitive theory of musical meaning. *Journal of Music Theory*, **44**(2), 323–379.

Clayton M (2007). Observing entrainment in music performance: video-based observational analysis of

Indian musicians' tanpura playing and beat marking. *Musicae Scientiae*, **11**(1), 27–60.

Cook N (2001). Theorizing musical meaning. *Music Theory Spectrum*, **23**(2), 170–195.

Cross I and Woodruff GE (2008). Music as a communicative medium. In R Botha, C Knight, eds, *The prehistory of language*, Vol. 1, 113–144. Oxford University Press, Oxford.

Davies S (1994). *Musical meaning and expression*. Cornell University Press, Ithaca, NY.

Davies S (2001). Philosophical perspectives on music's expressiveness. In P Juslin, JA Sloboda, eds, *Music and emotion: theory and research*, 23–44. Oxford University Press, Oxford.

Davies S (2003). *Themes in the philosophy of music*. Oxford University Press, Oxford.

Dempster D (1998). Is there even a grammar of music? *Musicae Scientiae*, **2**(1), 55–64.

Dennett D (1987). *The intentional stance*. MIT Press, Cambridge, MA.

Dennett D (1995). *Darwin's dangerous idea*. Penguin Books, London.

DeNora T (2000). *Music and everyday life*. Cambridge University Press, Cambridge.

Feld S (1982). *Sound and sentiment: birds, weeping, poetics and song in Kaluli expression*. Publications of the American Folklore Society. New series; 5. BZVDW, Philadelphia, PA.

Feld S and Fox AA (1994). Music and language. *Annual Review of Anthropology*, **23**, 25–53.

Finnegan R (1989). *The hidden musicians: music-making in an English town*. Cambridge University Press, Cambridge.

Grayling AC (ed.) (2000). *Philosophy: a guide through the subject*. Oxford University Press, Oxford.

Hoopes J (ed.) (1991). *Peirce on signs: writings on semiotic by Charles Sanders Peirce*. University of North Carolina Press, Chapel Hill, NC.

Jackendoff R (1987). *Consciousness and the computational mind*. MIT Press, Cambridge, MA.

Jackendoff R (2002). *Foundations of language: brain, meaning, grammar, evolution*. Oxford University Press, Oxford.

James J (1993). *The music of the spheres: music, science and the natural order of the universe*. Little, Brown and Co., London.

Johnson M (1987). *The body in the mind: the bodily basis of meaning, imagination and reason*. Chicago University Press, Chicago, IL.

Johnson-Laird PN (1983). *Mental models*. Cambridge University Press, Cambridge.

Juslin P and Sloboda JA (eds) (2001). *Music and emotion: theory and research*. Oxford University Press, Oxford.

Koelsch S, Kasper E, Sammler D, Schultze K, Gunter T and Frederici A (2004). Music, language and meaning: brain signatures of semantic processing. *Nature Neuroscience*, **7**(3), 302–307.

Kramer L (1995). *Music and postmodernist thought*. University of California Press, London.

Lakoff G (1987). *Women, fire and dangerous things*. University of Chicago Press, Chicago, IL.

Larson S (1998). Musical forces and melodic patterns. *Theory and Practice*, **22/23**, 55–71.

Larson S (2002). Musical forces, melodic expectation, and jazz melody. *Music Perception*, **19**(3), 351–385.

Larson S (2004). Musical forces and melodic expectations: comparing computer models and experimental results. *Music Perception*, **21**(4), 457–498.

Larson S and VanHandel L (2005). Measuring musical forces. *Music Perception*, **23**(2), 119–136.

Lerdahl F (2003). Two ways in which music relates to the world. *Music Theory Spectrum*, **25**(2), 367–373.

MacDonald, R Hargreaves D and Miell D (eds) (2002). *Musical identities*. Oxford University Press, Oxford.

Martin P (1995). *Sounds and society: themes in the sociology of music*. Manchester University Press, Manchester.

Martinez I (2007). The cognitive reality of prolongational structure in tonal music. Unpublished PhD thesis, Roehampton University, London.

McClary S (1991). *Feminine endings: music, gender and sexuality*. University of Minnesota Press, Minnesota, MN.

Merriam AP (1964). *The anthropology of music*. Northwestern University Press, Chicago, IL.

Millikan RG (2004). *Varieties of meaning: the 2002 Jean Nicod lectures*. MIT Press, Cambridge, MA.

Nattiez J-J (1990). *Music and discourse: toward a semiology of music*. Princeton University Press, Princeton, NJ.

North AC and Hargreaves DJ (1995). Subjective complexity, familiarity, and liking for popular music. *Psychomusicology*, **14**, 77–93.

Orr MG and Ohlsson S (2005). Relationship between complexity and liking as a function of expertise. *Music Perception*, **22**(4), 583–611.

Palisca C (1985). *Humanism in Italian Renaissance musical thought*. Yale University Press, New Haven, CT.

Panksepp J and Bernatzky G (2002). Emotional sounds and the brain: the neuro-affective foundations of musical appreciation. *Behavioural Processes*, **60**, 133–155.

Scruton R (1987). Analytical philosophy and the meaning of music. *Journal of Aesthetics and Art Criticism*, **46**, 169–176.

Scruton R (1997). *The aesthetics of music*. Clarendon Press, Oxford.

Sibley F (1959). Aesthetic concepts. *Philosophical Review*, **68**, 421–450.

Slobin M (1993). *Subcultural sounds: micromusics of the West*. Wesleyan University Press, Hanover, NH.

Sloboda JA, O Neill SA and Ivaldi A (2001). Functions of music in everyday life: an exploratory study using the Experience Sampling Method. *Musicae Scientiae*, **5**(1), 9–32.

Sparshott F (1998). Reflections on Affektenlehre and dance theory in the eighteenth century. *The Journal of Aesthetics and Art Criticism*, **56**(1), 21–28.

Sparshott F and Goehr L (2001). Philosophy of music: early Christian thought; medieval thought. In S Sadie, ed., *The new Grove dictionary of music and musicians, Vol. 19*, 608–611 Macmillan, London.

Steinbeis N and Koelsch S (2008). Shared neural resources between music and language indicate semantic processing of musical tension–resolution patterns. *Cerebral Cortex*, **18**, 1169–1178.

Tarski A (1956). *Logic, semantics. mathematics: papers from 1923 to 1938*, trans. J. H. Woodger. Oxford University Press, Oxford.

Thomas DA (1995). *Music and the origins of language: theories from the French Enlightenment*. Cambridge University Press, Cambridge.

Titon JT and Slobin M (1996). The music-culture as a world of music. In JT Titon, ed., *Worlds of music: an introduction to the music of the world's peoples*, 1–15. Schirmer Books, New York.

Tolbert E (2001). Music and meaning: An evolutionary story. *Psychology of Music*, **29**, 89–94.

Violi P (1999). Semiotics and cognition. In RA Wilson and FC Keil, eds, *The MIT encyclopedia of cognitive sciences*, 744–745 MIT Press, Cambridge, MA.

Wollheim R (1980). *Art and its objects*, 2nd edn. Cambridge University Press, Cambridge.

Zbikowski LM (2002). *Conceptualizing music: cognitive structure, theory and analysis*. Oxford University Press, Oxford.

# CHAPTER 4

# The social and personal functions of music in cross-cultural perspective

Martin Clayton

## Introduction

In the grounds of an English country house, pop star Robbie Williams teases an audience of thousands before launching into his hit song 'Angels': the thousands sing along as one with his amplified voice, swaying from side to side in synchrony.[1] In the Peruvian Andes, weeks of silence are broken by a festival in which groups of men playing flutes or panpipes attempt to perform as one instrument, while fighting for the attention of the crowds in the town square of Conima (Turino 1993). In a healing ceremony held in Temiar village in Malaysia, a singing medium directs his voice into the back of a patient: he feels the spiritguide singing through him (Roseman 1991, pp. 80, 115). In a small concert hall in an industrial Bengali town, an audience member listens to a singer performing an evening raga and imagines an ideal landscape of gentle breeze, broad river and brilliant red sunset.[2]

Music's uses and contexts are so many and so various that the task of cataloguing its functions is daunting: how can we make sense of this diversity? These functions appear to range from the individual (music can affect the way we feel and the way we manage our lives)—to the social (it can facilitate the coordination of large numbers of people and help to forge a sense of group identity). I will argue that musical behaviour (within which category I include all kinds of listening as well as performance) also covers a vast middle ground in which relationships between self and other or between the individual and the collective are played out. This chapter surveys some of the extant literature on music's functions—referring to literature from ethnomusicology, anthropology, musicology, psychology, and sociology, and discussing a wide variety of musical contexts from around the world—and develops an argument emphasizing music's role in the management of relationships between self and other.

## Approaches to the functions of music

Cross-cultural study of the functions of music reached a peak in the ethnomusicology of the 1950s and 1960s. The most concise and

[1] Robbie Williams, *What We Did Last Summer – Robbie Williams Live At Knebworth* (DVD, EMI, 2003). I would like to thank Laura Leante and Byron Dueck for their comments on a draft version of this chapter.

[2] This example is inspired by research Laura Leante, Tarun Kumar Nayak and I have been pursuing into audience attitudes and responses to raga performance.

comprehensive summary of this approach can be found in Merriam's *The Anthropology of Music* (1964), in which he distinguishes 'uses' from 'functions': '"Use" . . . refers to the situation in which music is employed in human action; "function" concerns the reasons for its employment and particularly the broader purpose which it serves' (p. 210). This distinction is a productive one for my purposes here. A catalogue of the uses of music would list innumerable categories—to lull babies to sleep, to court, to accompany dancing, worship, weddings, funerals, aerobics classes, spirit possession ceremonies, and so on. A discussion of functions should ask what purpose the music serves in these events.

Merriam enumerates no fewer than ten principal functions (pp. 219–227):

1 Emotional expression
2 Aesthetic enjoyment
3 Entertainment
4 Communication
5 Symbolic representation
6 Physical response
7 Enforcing conformity to social norms
8 Validation of social institutions and religious rituals
9 Contribution to the continuity and stability of culture, and
10 Contribution to the integration of society.

His list is a useful device on which to hang a compilation of numerous case studies, but should not be taken as a thoroughly worked-out taxonomy: 'aesthetic enjoyment' and 'entertainment' receive little attention, and can be read as a concession to those who would separate Western music out as a special category; their relationship to the more communicative, expressive and social functions Merriam lists is not examined. Merriam himself admits to finding 'communication' difficult to assess, while his functions 5, 7, 8, 9 and 10 could, as Nettl points out, perhaps be combined into 'the statement that music functions as the symbolic expression of the main values, patterns, or themes of a culture' (1983, p. 15). Attali's take on this was somewhat different, figuring music as 'prophetic' of social change (1985); more generally the issue of whether music *reflects* or *constitutes* a culture, or acts as a site for the *contestation* of cultural practice, has been a matter of debate.

Nettl also offers his own rating of the relative importance of Merriam's different functions in six very different cultures (1983, pp. 150–153). He uses the comparison to make the point that music has a range of different functions in his six chosen contexts, and that the balance between these functions varies cross-culturally. Another issue he raises is whether music has one overarching function or a multiplicity of overlapping functions, a question to which he devotes more thought. Although he has his own stab at the former—'The function of music in human society, what music ultimately does, is to control humanity's relationship to the supernatural, mediating between people and other beings, and to support the integrity of individual social groups' (p. 159)—he has at least as much sympathy for the 'many functions' argument. In this latter position we may find a parallel with Cross's more recent suggestion that music is effective precisely because of its ambiguity or 'floating intentionality', which enables it to 'serve as a medium for the maintenance of human social flexibility' (Cross 2005, p. 36). We might indeed argue that a single musical action can have multiple functions, and that it can perform these functions in different degrees for different participants. Two ethnomusicological examples will serve to reinforce this point. As Turino puts it in his study of Peruvian highland music:

> Music, dance, and drama are especially apt media for *simultaneously* articulating and uniting widely divergent and even conflicting images and meanings. The achachk'umu dancers, for example, represent figures of power who are at once local, sacred, foreign, positive, human, negative, divine, animal, serious, and comical through a performance that is both lightly entertaining and deeply meaningful . . . The complex imagery within Conimeño performing arts and festivals articulates—perhaps more accurately than other semiotic practices, such as everyday speech—the complexity and tensions of history and life itself.
>
> (p. 99, italics in original)

In her study of Prespa Albanian song, Sugarman refers this polysemy more explicitly to the topic of self and subjectivity:

> As a polysemic social practice, singing allows individuals to convey a range of messages that they might wish to make about themselves as social beings. Each rendition of each song thus serves as an embodied performance of *multiple aspects* of that performer's sense of self and of community.
>
> (1997, p. 3, italics added)

The comparative study of music's functions has received little attention in ethnomusicology since Nettl noted that 'the idea that we must find major, overriding functions for music seems to have come to a dead end' (1983, p. 150), although evolutionary, ethological and cognitive approaches have recently reopened the debate, and ethnomusicological literature continues to lend itself to cross-cultural schemes. This is demonstrated by ethologist Dissanayake in her list of 'Social functions served by ritual music', which I abbreviate here:

**(D1)** display of resources;

**(D2)** control and channelling of individual aggression;

**(D3)** facilitation of courtship;

**(D4)** establishment and maintenance of social identity through rites of passage;

**(D5)** relief from anxiety and psychological pain; and

**(D6)** promotion of group cooperation and prosperity.

The first three of these, she argues, can be traced back to the ritualization of animal behaviour, while the others, particularly (D4) and (D6) may be uniquely human (2006, pp. 43–49). These two categories, incidentally, are also the easiest to match with Merriam's list (functions 7–10): otherwise the different factors located by Dissanayake reflect a different disciplinary background more than they do any profound shift in the ethnomusicological research on which they draw.

Ethnomusicological studies have nonetheless made contributions that ought to influence discussions of music's functions. Amongst these are discussions of the role of music in inducing trance, possession and healing, and ecstatic listening (Rouget 1985; Roseman 1991; Friedson 1996; Racy 2003; Becker 2004); increasing awareness of the role of music in the negotiation of ideologies of gender and ethnicity (Koskoff 1989; Stokes 1994; Sugarman 1997; Moisala and Diamond 2000); and a greater focus on the use of music in maintaining and adapting cultural identity in migrant and diasporic cultures (Turino 1993; Shelemay 1998; Slobin 1993, 2003).

Durkheim's influence on ethnomusicology has also has been renewed in recent decades, his theories on the basis of religious thought in collective action inspiring reflections on music. As Durkheim wrote:

> [I]f collective life awakens religious thought on reaching a certain degree of intensity, it is because it brings about a state of effervescence which changes the conditions of psychic activity. Vital energies are over-excited, passions more active, sensations stronger.
>
> (1968, p. 422)

This idea of 'effervescence' or, as Stokes put it, 'the considerable surplus of affect [music-making and dance] generate in performance' (2008), offers a warning against defining music's functions too narrowly in terms of social efficacy. Durkheim suggested further that 'The life thus brought into being even enjoys so great an independence that it sometimes indulges in manifestations with no purpose or utility of any sort, for the mere pleasure of affirming itself' (p. 424). According to this view music is not limited to performing a function such as those on Merriam's and Dissanayake's lists: its power to arouse and excite becomes an end in itself. For some this might throw the whole issue of music's social 'function' into doubt, but it is important to separate motivation from effect here. The fact that participants may be motivated by the promise of a heightened state of arousal does not, for instance, preclude a musical event from influencing the way social relations are imagined: these effects may indeed be more profound if they are not brought to the surface and articulated.

Finally, while sociologists and social psychologists have discussed the deployment of musical resources in the realisation of individual identity

(see below), ethnomusicologists have stressed both the variety of concepts of 'self' cross-culturally and the social constitution of the individual. Roseman draws on the notion of the 'sociocentric self', 'an interactive self that doesn't quite end at the boundaries of the individual' (1991, p. 46). In the case of the Temiar, whose sense of self seems to conform to this model, 'the beauty of interconnectedness with community and nature is offset by concern for the integrity of self', which requires a 'code of ritualized behaviour' for its management. The individual is implicated in the social and the social in the personal, then, and music—however we define it—provides tools for managing these interrelationships.

## Defining musical behaviour

By 'music' we refer to an aspect of *human behaviour*; the creation of *organized sound* is an important part of that behaviour, whether conceived as an aim itself or as a means to another end; and this behaviour is *distinguishable from everyday speech communication*.[3] To start talking about the definition of music at this point may seem an odd strategy: nonetheless, the question of music's functions is inseparably tied up with that of its definition, and therefore it is important to revisit definition in this discussion. If my statement is accepted as a working definition, where do the boundaries of 'musical' behaviour lie? Is the recitation of poetry an instance of music? Auctioneering? Hunting calls? Bell-ringing? Birdsong?

In terms of the definition above the logical distinctions would be human vs non-human; organized sound vs noise; and music vs language. The first of these rules out birdsong—although Feld's analysis of the profound interconnectedness of bird calls, weeping and song in Kaluli culture in Papua New Guinea is a reminder that simple definitions are always challenged by the complexities of real life: non-human sounds are at least implicated in many conceptions of music (1990).

[3] This definition draws, of course, on Blacking's pithy 'humanly organized sound' (1973) and on Nettl's discussion of the issue (1983, pp. 15–25).

The distinction between music and noise lies not in the sound, but in the way human beings make use of it. A team of expert campanologists clearly make music; a bell accidentally sounding in the wind does not (even though the sound may be attractive); when Feld mixes his soundscape recordings into the 'compositions' of *The Time of Bells* (2004), he arguably nonetheless transforms non-musical sound into music.

To continue the same theme, a bell being struck to call people to a meal or meeting is probably not music—at least, not if heard merely as a signal that could have been replaced by a verbal instruction, just as a drum pattern *heard as a verbal message* is not for that listener music, but rather a form of language. Insofar as 'musical' sound can be substituted by words, it can be regarded as a form of language; conversely, almost all forms of vocalization can be said to have 'musical' qualities even when they are primarily linguistic in character. In each case, then, the boundaries of music can be seen to be less than clear cut.

For Rousseau both language and music originated together with the birth of human society (1998): the emergence of a single mode of vocal communication in the early stages of human development is echoed in more recent speculation (see the essays in Wallin *et al.* 2000), and child development studies suggest the same may be true ontogenetically. Vocal utterance is not the only source of musical behaviour, however. Just as important is a human tendency to rhythmic coordination and entrainment (Clayton *et al.* 2005). This can be observed in almost any kind of human activity—including walking and talking—but is particularly noticeable in forms of activity where rhythmic coordination makes physical work more efficient. Many examples of this can be found in the collaborative pounding of roots or grains, in which two or more people alternate strikes—a common activity in many parts of the world, and one which frequently incorporates a 'musical' dimension employing the sounds produced. In most kinds of musical behaviour, cross-culturally, these two elements (vocal utterance and coordinated action) are combined in some way—and if we regard most melodic instrumental music as a kind of extension of the voice (Cone 1974) then much if not

all music-making is founded on these two forms of behaviour.

## Music as communicative medium

One thing that vocalization and coordinated action have in common is that both play a role in connecting individuals: they are both social. One of the ways in which speech and song may be distinguished, in fact, is in the kind of social relationships they presuppose. For Zuckerkandl, the fundamental distinction between speech and song is that speech involves communication from A to B, whereas song presupposes the merging of self and other into a social 'we':

> the tones—singing—essentially express not the individual but the group, more accurately, the individual in so far as he is a member of the group, still more accurately, the individual in so far as his relation to the others is not one of 'facing them' but of togetherness.
>
> (1973, p. 28)

As Miell, Macdonald and Hargreaves point out in their discussion of musical communication, transmission models have nonetheless predominated in music psychology. Their own model attempts 'to specify the main personal, musical and situational variables which give rise to a musical *performance*, [and similarly] to explain the *response* to music in a specific situation' (2005, p. 7): where the domains of performance and response meet they describe the 'spark' of communication (p. 18). In other fields authors have engaged similarly with the notion of 'transmission', whether considered to be from active performer to passive listener, or with active engagement on both sides of the equation. Nattiez's semiotic theory (1990) invokes three levels: instead of a unidirectional arrow from performer to listener, both relate to a 'neutral' middle level (located in the same place as Miell *et al.*'s 'spark' of communication). Feld plays down this neutral level and highlights the active and relational nature of meaning-making, allowing the listener greater autonomy in applying 'interpretive moves' (1994). Recent musicological writing has dealt with the ways in which music affords or contructs a 'subject position' for the listener, in effect constraining this act of interpretation while still allowing the listener some autonomy (Frith 1996, pp. 183ff; Cumming 2000; Dibben 2006). The ways in which real people deploy musical resources in managing their activities, identities and emotional states has been explored in music sociology, social psychology and ethnomusicology (Crafts *et al.* 1993; DeNora 2000; Hargreaves *et al.* 2002,), further elaborating the 'response' side of Miell *et al.*'s model. On the performance side of the equation, Swaine has argued that vocalization is a tool for an individual's emotional self-*regulation*,[4] with expression a secondary effect: the same could no doubt be said of much instrumental performance.

Transmission models clearly have explanatory power, especially when nuanced in the ways described above. No matter how good a fit they may be with some examples of musical behaviour, however, they are equally problematic in many others. I return to my example of Robbie Williams singing 'Angels': many in his audience know the song as well as the singer himself does; there is little, if any, new information to 'transmit'. In this case notions of interaction, intersubjectivity and encounter are surely more productive. If interaction is placed at the heart of our understanding of musical behaviour, the 'transmission' of information or affective states remains as a possibility but not a sine qua non.

Musical interaction can have other goals and other results, and these might include for instance—in the same or in other contexts—a loss or sublimation of ego or the promotion of fellow-feeling. Mithen argues for the evolutionary importance of collective music-making thus: 'Joint music-making served to facilitate cooperative behaviour by advertising one's willingness to cooperate, and by creating shared emotional states leading to "boundary loss"/"we-ness"/"coupling"/"in-group bias"' (218); in ethnomusicology Keil stresses the 'participation' found in many social musical events (1994), while Blacking favoured the term 'fellow feeling', both drawing on Durkheim (see above). Friedson, in his study of spirit healing among the Tumbuka people of Malawi, argues that in this context

[4] See Cross *et al.* 2009, in press.

music can lead to a state of communion or communitas—'a relational quality of full unmediated communication, even communion, between definite and determinate entities'—(Turner and Turner 1978, p. 250, cited in Friedson 1996, p. 125), of intersubjectivity and the annihilation of interpersonal distance.

The notion of 'encounter', too, has been deployed by ethnomusicologists: for Averill, music is an index of encounter (2003), while for Bohlman (2002) encounter is the key organizing concept in his history of 'world music'. Where the notions of communion, group solidarity and fellow-feeling can lead people to write exclusively of social unity, that of 'encounter' leaves space for conflict, for exclusion and for maintaining a sense of alterity, which also have a place in musical performance. Turino's example quoted in my first paragraph shows how group identity and inter-group conflict—albeit of a relatively benign kind—can go hand in hand. Another example of this can be found in Afro-Brazilian Congado ritual, in which individual groups playing simultaneously strive to maintain a tight intra-group rhythmic coordination which indexes their group solidarity, while resisting entrainment with other co-present groups for the same reason (Lucas 2002; Clayton *et al.* forthcoming).

All of these approaches complicate the idea of music as a medium of communication, if the notion of information transfer or transmission is taken as defining 'communication'. Finnegan presents a different view, however, building on Birdwhistell's much broader definition of 'communication' as the 'dynamic aspect of human interconnection' (1968, p. 27, cited in Finnegan 2002, p. 32): following Finnegan's line allows us to continue to address music and language as complementary modes of human communication. Whether music indexes communication, communion or encounter makes a difference, of course: music is either a way for A to transfer information to B; how A and B affirm their collective identity; or simply what happens when A meets B. In fact it can be any and all of these things, and this flexibility is part of the point. If 'communication' encompasses information transfer, communion and encounter of all sorts, then we can freely describe music as part of man's communicatory toolbox, asking how it relates to other related tools such as language and physical gesture.

My discussion above focused on the difficulty of defining boundaries such as those between music and language or between music and noise, but this does music something of a disservice: more positively, what defines music and what makes it a valuable analytical category cross-culturally? What would we lose if we argued the concept away? Ultimately, the phenomenon of people using organized sound in order to connect with each other—in ways which cannot be regarded simply as forms of language or information transfer, or completely explained in terms of verbal concepts—can be regarded as a single, generalizable idea, and a robust one no matter how ill-defined and porous its boundaries are.

## Discussion: the functions of music

Music is not a single form of behaviour any more than it is a single kind of sonic product, but a composite of different forms of sound and behaviour. The behaviours we describe as 'musical' are founded principally on two distinct, yet interrelated phenomena: the human capacity for vocal utterance, and our innate tendency to mutually entrain our actions, especially those involving sound production and quasi-periodic behaviour. Neither of these factors can be reduced to the other, yet in practice most musical behaviour draws on both to varying degrees. Vocal utterance may also be considered as the origin of language, while music as coordinated action relates to other forms of group activity such as physical labour, marching and dance—the last of these frequently as inseparable from music as gesture is from speech. Musical behaviour generates humanly organized sound beyond the limits of normal speech communication: it is distinguishable from non-human sounds, from noise and from language, although each of these boundaries is inevitably fuzzy. In light of these discussions, I propose that musical behaviours tend to perform one or more of the following functions:

1 Regulation of an individual's emotional, cognitive or physiological state. Musical performance has physiological effects—singing

has implications for respiration and body posture, many instruments develop bimanual coordination skills, and so on. Given that physiological changes can influence emotional and cognitive states, it is not surprising that in many cultures musical behaviours are employed in the regulation of these states—involving mild effects such a temporary calming or excitation or more dramatic changes such as trance. Since music can be used in the regulation of emotional and cognitive states, again it is not surprising that performances are frequently taken as indexes of these states, and performers as expressing emotions. Music listening—whether in a live performance situation or to sound recordings—can be used in parallel ways as a tool for self-regulation (for example in managing mood), thanks to the role of mimesis and identification with the music's subjects (if the singer on the record is indicating an excited state, then to the extent that the listener sympathizes with the singer he or she may also become excited).

2 Mediation between self and other. Music has been used for centuries as a tool for interaction in instances where normal speech communication is found to be inadequate. Examples include the use of special forms of song or music to communicate with Gods, spirits or ancestors in ritual contexts. More mundane examples might include the use of song to communicate emotional states felt to be beyond the scope of everyday speech, or to communicate intimately with large numbers of people. Musicking can be both a powerful and an extremely flexible tool for modelling and mediating relations between self and other. These functions can be interpreted either as communication between the group and an external other (such as an ancestral spirit), or as a function of intragroup interaction. The 'other' could be the spirit or other entity felt to be singing through a performer (as in the case of the spirit guide described by Roseman above);[5] it could be another being (human, divinity or spirit) addressed in song; or it could be another entity constructed and represented through song. In the last sense musicologists have drawn attention to the multiplicity of voices created through a musical performance (Cone 1974), and to the role—in opera in particular—of the play of identity and alterity (Abbate 1991; McClary 1992, pp. 29–43).

3 Symbolic representation. Although music seems in some respects to display ambiguity in specifying its referents, it can also be extremely efficient as a semiotic medium. Just as gesture can indicate the direction and velocity of a movement, or the shape, size and spatial disposition of objects with more precision than does the speech it accompanies (McNeill 1992), and facial expression can indicate emotional states more economically than can language, so too musical sound and action can specify aspects both of affect and of movement more precisely than words. Conventional musical signs as deployed in cinema or advertising, moreover, can specify their referents with unmatched precision and economy (Tagg and Clarida 2003). This play of signs has a different emphasis to that in language, perhaps because musical signs tend to be concerned with identity and alterity, bodily motion, and relationships between self and other. If world music is an index of encounter, as Bohlman argues, this music of encounter in fact relies on a degree of ambiguity between the perception of 'self' and 'other' (Leante 2004). Music can clearly evoke issues of alterity and identity without articulating an unambiguous message on the subject.

4 Coordination of action. Human beings display a tendency to entrain their physical actions, either to each other or to an external sound reference (such as a musical performance or recording): the fact that entrainment occurs even when resisted (Clayton 2007; Clayton *et al.* forthcoming) indicates the pervasiveness of the phenomenon in human interactions. The use of music to facilitate coordinated action hardly needs demonstration: what is

[5] Similarly, I have elsewhere quoted Indian singer Veena Sahasrabuddhe's view that her main relationship is an identification with the raga she sings rather than with her audience (Clayton 2005, p. 373).

striking is that in many instances the urge to coordinate action is accompanied by a development of 'fellow feeling', with the emergence or strengthening of bonds of shared mood or emotional state, and in some cases shared ideologies. As Dueck has argued, musical performance not only entertains existing groups, it plays a key role in constituting publics (2007); numerous examples have been described in which music is deployed in order to maintain, revive and/or to create a sense of 'tradition', or to facilitate a sense of identity amongst particular social groups—for instance in diaspora.

There may be examples in which one of these functions outweighs each of the others to such a degree that we may identify it as the principal function of music in that instance. Perhaps, as Nettl suggests, the balance between functions varies between cultures (or indeed between genres, or even between different moments in a single performance). Possibly more common, however, are instances in which several of these functions operate simultaneously. For instance, a popular song such as Robbie Williams's 'Angels' might (1) originate in an urge towards the regulation of an emotional state; and (2) be used in an act of communication between singer and audience that could not have been accomplished by speech alone. It (3) represents a particular pattern of feeling dynamically and precisely, and points to numerous referents, from the singer's public persona to English culture at the end of the twentieth century. Its singing at large public concerts (4) elicits a coordination of both action (singing along, swaying bodies in time with the music) and emotion on a massive scale; and bringing the matter full circle, individuals partaking in this communal experience can be seen to deploy their listening experience and participation as a means of emotional regulation. In this way, this and many other forms of musical behaviour simultaneously perform several distinct yet interrelated functions. This example may also remind us of Durkheim's 'effervescence': it is not simply that multiple functions are performed simultaneously, rather the very intensity of experience provides a basis for its many effects and interpretations.

The four headings above are largely concerned with relations between the personal and the social—identity, alterity and mediation. Music is a tool for the discovery, manipulation and projection of individual identity; an individual's identity construction is however inherently social, implicating a variety of groups to which the individual feels he does or does not belong. Music is a tool for facilitating intimate interactions, and an index of such interactions: it therefore helps to bring social groups into being. Music is also a tool for constituting publics, for allowing social groups larger than the family, clan or village to create themselves and to include or exclude individuals. It can help to dissolve the boundary between self and other in a way speech generally does not; but music can also be deployed to reinforce boundaries and to distance or exclude. This is why it can be described as a flexible tool for managing relationships between self and other.

Like Merriam, I argue that underlying the bewildering variety of uses to which music is and has been put around the world we can define a finite set of global functions. My list is shorter and more synthetic than Merriam's, but this greater concision is offset by more emphasis on the idea that these functions can not be reduced to one: it is not simply that a particular writer, or the language itself, is incapable of specifying *the* function of music; rather, music frequently depends for its efficacy precisely on the indeterminacy of its 'true' or underlying function. We can say, nonetheless, that musical behaviour is deployed in the management of relations between self and other, and that it can and does perform this function at multiple levels simultaneously.

## References

Abbate C (1991). *Unsung voices. Opera and musical narrative in the nineteenth century.* Princeton University Press, Princeton, NJ.

Attali J (1985). *Noise. The political economy of music.* University of Minnesota Press, Minneapolis, MN.

Averill G (2003). *Four parts, no waiting.* Oxford University Press, New York.

Becker J (2004). *Deep listeners: music, emotion and trancing.* Indiana University Press, Bloomington, IN.

Birdwhistell RL (1968). Communication. In DL Sills, ed. *International encyclopedia of the social sciences*, 17 Vols, Vol. 3, 24–29. Macmillan and Free Press, New York.

Blacking J (1973). *How musical is man?* University of Washington Press, Seattle, WA.

Bohlman PV (2002). *World music: a very short introduction.* Oxford University Press, Oxford.

Clayton M (2005). Communication in Indian raga performance. In D Miell, D Hargreaves, R MacDonald, eds. *Musical communication*, 361–381. Oxford University Press, Oxford.

Clayton M (2007). Observing entrainment in music performance: video-based observational analysis of Indian musicians' tanpura playing and beat marking. *Musicae Scientiae*, **11(1)**, 27–60.

Clayton M, Lucas G and Leante L (forthcoming). Inter-group entrainment in Afro-Brazilian Congado ritual.

Clayton M, Sager R and Will U (2005). In time with the music: the concept of entrainment and its significance for ethnomusicology. *European Meetings in Ethnomusicology*, **11** (ESEM Counterpoint 1), 3–75.

Cone E (1974). *The composer's voice.* University of California Press, Berkeley, CA.

Crafts SD, Cavicchi D, Keil C and the Music in Daily Life Project (1993). *My music.* Wesleyan University Press, Hanover, NH and London.

Cross I (2005). Music and meaning, ambiguity and evolution. In D Miell, D Hargreaves, R MacDonald, eds *Musical communication*, 27–43 Oxford University Press, Oxford.

Cross I, Himberg T, Swaine J and Bispham J (2009, in press). Evolution and musical rhythm. *Evolutionary Psychology.*

Cumming N (2000). *The sonic self. Musical subjectivity and signification.* University of Indiana Press, Bloomington, IN.

DeNora T (2000). *Music in everyday life.* Cambridge University Press, Cambridge.

Dibben N (2006). Subjectivity and the construction of emotion in the music of Björk. *Music Analysis*, **25**, 171–197.

Dissanayake E (2006). Ritual and ritualization: musical means of conveying and shaping emotion in humans and other animals. In S Brown, U Volgsten, eds, *Music and manipulation. On the social uses and social control of music*, 31–56. Berghahn, New York and Oxford.

Dueck B (2007). Public and intimate sociability in First Nations and Metis fiddling. *Ethnomusicology*, **51(1)**, 30–63.

Durkheim E (1968) [1915] *The elementary forms of the religious life.* George Allen and Unwin, London.

Feld S (1990) [1982]. *Sound and sentiment: birds, weeping, poetics and song in Kaluli expression*, revised edn. University of Pennsylvania Press, Philadelphia, PA.

Feld S (1994) [1984]. Communication, music and speech about music. In C Keil and S Feld eds, *Music grooves*, 77–95. University of Chicago Press, Chicago, IL.

Feld S (2004). *The time of bells.* Voxlox 104 (audio CD).

Finnegan R (2002). *Communicating. The multiple modes of human interconnection.* Routledge, London and New York

Friedson SM (1996). *Dancing prophets. Musical experience in Tumbuka healing.* University of Chicago Press, Chicago, IL.

Frith S (1996). *Performing rites: on the value of popular music.* Oxford University Press, Oxford.

Hargreaves D, Miell D and Macdonald R (2002). What are musical identities, and why are they important? In R Macdonald, D Hargreaves and D Miell, eds, *Musical identities*, pp 1–20. Oxford University Press, Oxford.

Keil C (1994) [1987]. Participatory discrepancies and the power of music. In C Keil and S Feld eds, *Music Grooves*, 96–108. University of Chicago Press, Chicago, IL.

Koskoff E (ed) (1989). *Women and music in cross-cultural perspective.* University of Illinois Press, Bloomington, IN.

Leante L (2004). Shaping diasporic sounds: identity as meaning in bhangra. *The World of Music*, **46**, 109–122.

Lucas G (2002). Musical rituals of Afro-Brazilian religious groups within the ceremonies of Congado. *Yearbook for Traditional Music*, **34**, 115–127.

McClary S (1992). *Georges Bizet, Carmen.* Cambridge University Press, Cambridge.

McNeill D (1992). *Hand and mind. What gestures reveal about thought.* University of Chicago Press, Chicago, IL.

Merriam A (1964). *The anthropology of music.* Northwestern University Press, Evanston, IL.

Miell D, Hargreaves D and MacDonald R (eds) (2005). *Musical communication.* Oxford University Press, Oxford.

Mithen S (2005). *The singing Neanderthals. The origins of music, language, mind and body.* Weidenfeld and Nicholson, London.

Moisala P and Diamond B (eds) (2000). *Music and gender.* University of Illinois Press, Bloomington, IN.

Nattiez J-J (1990). *Music and discourse. Toward a semiology of music*, trans. C Abbate. Princeton University Press, Princeton, NJ.

Nettl B (1983). *The study of ethnomusicology: twenty-nine issues and concepts.* University of Illinois Press, Bloomington, IN.

Racy AJ (2003). *Making music in the Arab world. The culture and artistry of Tarab.* Cambridge University Press, Cambridge.

Roseman M (1991). *Healing sounds from the Malaysian rainforest. Temiar music and medicine.* University of California Press, Berkeley and Los Angeles, CA.

Rouget G (1985). *Music and trance: a theory of the relations between music and possession.* University of Chicago Press, Chicago, IL.

Rousseau J-J (1998). *Essay on the origin of languages and writings related to music*, trans. J.T. Scott. University Press of New England, Hanover, NH.

Shelemay KK (1998). *Let jasmine rain down: song and remembrance among Syrian Jews.* University of Chicago Press, Chicago, IL.

Slobin M (1993). *Subcultural sounds: micromusics of the West.* Wesleyan University Press, Hanover, NH.

Slobin M (2003). The destiny of 'diaspora' in ethnomusicology. In M Clayton, T Herbert, R Middleton, eds, *The cultural study of music: a critical*

*introduction*, 284–296. Routledge, New York and London.

Stokes M (1994). *Ethnicity, identity and music: the musical construction of place.* Berg, Oxford.

Stokes M (2008) Contemporary theoretical issues: communities and their musics. In C Pegg *et al. Ethnomusicology* §IV, **2**. *Grove Music Online*, http://www.oxfordmusiconline.com/Subscriber/article/grove/music/52178pg.

Sugarman JC (1997). *Engendering song: singing and subjectivity at Prespa Albanian weddings.* Chicago University Press, Chicago, IL.

Tagg P and B Clarida (2003). *Ten little title tunes.* The Mass Media Music Scholar's Press, New York and Montreal.

Turino T (1993). *Moving away from silence. Music of the Peruvian Altiplano and the experience of urban migration.* University of Chicago Press, Chicago, IL.

Turner VW and E Turner (1978). *Image and pilgrimage in Christian culture.* Columbia University Press, New York.

Wallin NL, Merker B and Brown S (eds) (2000). *The origins of music.* MIT Press, Cambridge, MA.

Zuckerkandl V (1973). *Man the musician. Sound and Symbol: Vol. 2*, Trans. Norbert Guterman. Bollingen Series 44/2. Princeton University Press, Princeton, NJ.

# PART 2

# Music perception

Edited by Ian Cross

# CHAPTER 5

# The perception of pitch

Thomas Stainsby and Ian Cross

## Basic psychophysics of pitch perception

### Pitch and frequency

While a musician may find a definition of pitch redundant, psychoacoustics—the scientific investigation of our perception of sound—nonetheless demands one. Although there is an intuitive understanding that pitch is related to frequency, it is important to understand that the two are not directly equatable. Pitch is a percept, measurable only by psychophysical investigation, and frequency is a physical quantity, which describes the periodic properties of a signal. Sounds that contain clearly measurable frequency content may not produce a clear or unambiguous sense of pitch, and the perceived pitch may also depend on other physical quantities such as intensity. The American National Standards Institute offers the following definition:

> Pitch [is] that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high. Pitch depends primarily on the frequency content of the sound stimulus, but it also depends on the sound pressure and the waveform of the stimulus.
>
> ANSI (1994)

However, we require a more restricted definition for the present review. Frequencies below around 30 Hz and above around 5 kHz can lose their identities as *musical* pitches (Attneave and Olson 1971), while still sounding 'low' or 'high', respectively. Additionally, there is something more global about pitch than individual frequencies. Most sounds we hear are mixtures of components with many different simultaneous frequencies, yet our auditory system generally combines these into a single percept of one overall pitch. The mechanisms that group the individual frequency components into an overall pitch are also at work in processing combinations of notes arising from different instruments. Does the combined percept have a single overall sense of pitch, or perhaps a 'timbre' for that combination and instrumental voicing, or alternatively, is there a clear hierarchy with a given 'harmony' consisting of distinct individual pitches?

One solution is to offer definitions of pitch that require the ability to convey a musical melody (e.g. Attneave and Olson 1971; Burns and Viemeister 1976). A combination of this requirement and the ANSI definition meets an intuitive and informal definition that may satisfy most musicians, that 'pitch is the stuff that notes are made of'.

In this chapter's discussion of pitch, we begin with a brief introduction to the function of the auditory system. This considers what limitations there are on our ability to extract frequency information from the sounds reaching our ears, and provides an orientation for some of the mechanisms described in the later sections. A review is then offered of relevant psychoacoustic research—how do we construct a sense of pitch from the physical acoustic input? This leads to a consideration of higher auditory processing functions that include the mechanisms we use to interpret the complicated mixture of sounds around us and the relationships between pitch as it applies to music and how it is manifested in speech and language.

The only sound that has just one frequency component is termed a pure tone or sine tone,

having a quality like that of a tuning fork. Its frequency spectrum has energy at only a single frequency, and its temporal wave form has the shape of a sine function. Although single pure tones are rare in nature, the concept is useful when we consider that real-world complex sounds and mixtures of sounds are made up of the sum of multiple pure tones such as these, termed *frequency components* or *partials*. We can then understand the task of the auditory system as being that of detecting and decoding such mixtures of frequency components, and interpreting this information to make sense of the sound world or *auditory scene* around us (Bregman 1990).

Before describing the mechanics of frequency coding, which occurs primarily in the inner ear, it is necessary to consider briefly the physiology of the auditory system. The acoustic wave from the outside world is captured and filtered by the outer ear, consisting of the pinna (external ear) and the ear canal. This stage of acoustic processing helps amplify the sound and encode information about the direction of sound sources. Subsequently, the sound is transmitted through the middle ear to the inner ear, consisting of the labyrinth (not considered here) and the cochlea, a small fluid-filled coil containing the sensory mechanisms that transform (transduce) mechanical energy into the electrical signals that are processed in the brain. The middle ear, which consists of the eardrum and the auditory ossicles (the three smallest bones in the body), is primarily concerned with the task of matching the impedance of the external air (its resistance to the passage of sound energy) to that of the fluid inside the cochlea. This maximizes the amount of energy transmitted—around 95 per cent would be reflected otherwise. It is not until the sound reaches the inner ear that frequency analysis and coding takes place. The cochlea in the inner ear is situated in the temporal bone of the skull, and is divided along its length (approximately 35 mm) by the basilar membrane (BM). The actual sensory transducers, the hair cells, are aligned along this membrane, in one row of inner hair cells (IHCs)—about 3500 in total—and three or four rows of outer hair cells (OHCs)—about 12 000 in total. The incoming sound sets up a travelling wave of vibration along the basilar membrane (von Békésy 1947). The IHCs in the vicinity of points of significant BM displacement are stimulated by the movement of their stereocilia (tiny hair-like projections grouped in bundles at the top of each hair cell) in the cochlear fluid, which in turn send electrical signals along the auditory nerve fibres that connect with, or *synapse* onto, each IHC. The OHCs fulfil a somewhat different function, which is described below.

The primary method of frequency coding is offered by the structure of the basilar membrane, which varies in width, thickness and stiffness along its length. As a result, each point of the membrane has a unique resonant frequency, i.e. the frequency at which it vibrates most easily. The base of the cochlea (closest to the middle ear) responds maximally to high frequencies (up to around 20 kHz), while the opposite, apical, end responds maximally to low frequencies (down to around 20 Hz). The frequency for which a given place on the basilar membrane shows a maximum response is termed its *characteristic frequency*. This leads us to understanding the first method of coding frequency, termed *place coding* or *tonotopic coding*. For each location on the BM showing maximum response to a characteristic frequency, there are a number of hair cells in the vicinity that will be maximally stimulated. There are roughly 30 000 auditory neurons that synapse onto the 3500 or so IHCs. With tonotopic coding, frequency can be represented by which neurons have been stimulated (i.e. they trace back to a specific localized region of the basilar membrane). Note that because there is a small region of maximum displacement rather than a point source, neural activity is not signalled in a single auditory neuron's response, but in the distribution of the responses of a population of probably hundreds or over a thousand neurons, a pattern of activity that is termed the *neural excitation pattern*. It should be noted that this does not imply that only one frequency will excite a discrete population of neurons; the mapping between frequencies and neurones is many-to-many, with a group of neurons responding to a limited range or band of frequencies, which will vary with signal intensity.

We can see that there is thus a limit to the specificity of frequency information derived from place coding; BM output in response to a single input frequency appears more like the output of

a band-pass filter than the output of a sharply tuned system. This is enhanced by a second, active, tuning process, which is believed to be mediated by the OHCs in the vicinity of the peak of the travelling wave (Rhode 1971; Ruggero 1992; Robles and Ruggero 2001). These cells actively move and change shape in response to stimulation, thus increasing BM displacement, in turn causing greater stimulation of the IHCs. Because the action of the OHCs is quite localized, this spatially restricted amplification of travelling-wave displacement results in a sharpening of the excitation pattern around the characteristic frequency, and hence an improvement in frequency resolution.

Louder sounds generate a larger travelling wave, and hence a greater spread of excitation, than a quiet sound. Higher stimulus levels thus result in decreased frequency resolution, and increase the extent to which loud components 'mask' the quieter ones. This explains the loss of perceptible detail in sound mixtures at loud levels. The extent to which sounds mask each other in this way can be used to elucidate and quantify the concept of frequency selectivity (Fletcher 1940; Patterson 1976; Glasberg and Moore 1990), which can be defined as the ability to hear out the sinusoidal partials or components of a complex sound (Moore 2003). Sounds are said to fall within the same band-pass auditory filter when they are able to mask each other to a sufficient extent. Figure 5.1 shows how auditory filter bandwidth varies with centre frequency

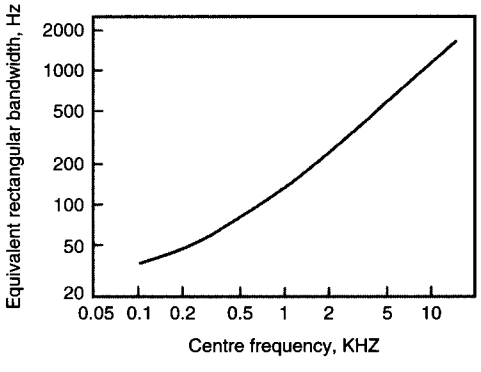



**Fig. 5.1** Average auditory-filter width and the ERB scale (after Glasberg and Moore 1990). The solid line plots the calculated width of the auditory filter for a given centre frequency.

(Glasberg and Moore 1990; Moore 2003). For frequency components to be 'heard out' and processed independently from each other, they need to fall within the passbands of different auditory filters. Once the centre frequency of one auditory filter has been defined, the next adjacent auditory filter needs be a certain distance along the basilar membrane to achieve the requisite independence of frequency selectivity. This is the logic behind the equivalent rectangular bandwidth (ERB) frequency scale (Glasberg and Moore 1990), which can be used to measure frequency in the perceptually relevant scale of auditory-filter bandwidth. Measured in this way, human hearing can be said to cover a range of a little fewer than 40 ERBs. Multiple frequencies falling within the same auditory filter cannot be resolved or processed independently, and only their combined effect is available to the auditory system.

The tonotopic organization and place frequency coding established in the cochlea are maintained up through higher centres in the auditory neural pathways. Progressively less is known about the nature of auditory processing and the function of particular structures as the auditory pathway ascends. A comprehensive review of the neurophysiology of pitch is provided in the chapter by Winter in Plack *et al.* (2005).

The second type of frequency coding, *temporal coding*, is achieved by an analysis of the behaviour of neural firings over time. That is, once the initial place coding of frequency has been achieved from the overall number of firings in a given population of neurons, further frequency information can be extracted from analysis of the period between successive spikes (more-or-less simultaneous bursts of energy produced by a group of neurones). However, such temporal information is only available for the coding of frequencies below approximately 5 kHz. Two factors contribute to this limitation on the frequencies that can be encoded by neural temporal coding. The first is the refractory period of an individual neuron. Once it has fired, a neuron is unable to fire again until it is metabolically 'recharged', which limits its firing to a maximum rate of 1 kHz. However, because stimulus frequency is being coded by populations of fibres, timing information averaged

across multiple neurons can still code frequencies up to 5 kHz. At this point the second limiting factor is reached, that of the breakdown of phase locking (Rose *et al.* 1967; Palmer and Russell 1986). Phase locking is required for the information obtained from the timing of neural spikes to have any meaningful relationship to the stimulus frequency. The term refers to the ability of a neurone to fire at a predictable time point—phase—in the stimulus waveform cycle—e.g. at the maximum downward displacement of the basilar membrane's vibration. Once a neuron can no longer fire at a consistent phase of the stimulus, its response is random and hence noisy. Above 5 kHz, the stimulation period simply becomes too short for any neurons to fire at a predictable phase.

We can thus see that temporal coding should provide robust and reliable frequency information for lower frequencies. By contrast, all coding of frequencies above 5 kHz can only be achieved by place coding. Both types of coding are available over a good portion of the audible frequency range, meaning that the mechanisms of pitch perception can employ both types of frequency information in the majority of instances.

It is relatively straightforward to understand how the auditory system can derive pitch from pure tones. For such stimuli, the pitch is unambiguously related to the frequency. Place and temporal cues are fully congruent for these tones, which is not always the case for more complex stimuli. For these reasons, the pitch of any type of stimulus is often quantified by using the frequency of a pure tone that is subjectively matched to the complex tone as having the same pitch.

The relative contribution of place and temporal frequency coding to the perception of pitch for pure tones has been investigated by examining frequency discrimination data from different task types. This shows that place resolution is more or less constant across the usable frequency range, while for lower frequencies, temporal coding offers superior performance to that achievable from place coding alone (Sek and Moore 1995), yielding frequency discrimination thresholds of *c.* 3–4 Hz at 1 kHz. This observation is supported by the results of Attneave and Olson (1971), who found that while listeners could still detect changes in frequency above 5 kHz, the sense of pitch deteriorated to the point where melodies ceased to be recognizable. Musical practice is consistent with this in that C8, the highest note on a piano and the highest note of an orchestral instrument (the piccolo), has a fundamental frequency of 4176 Hz. Taken together, these facts suggest that temporal-frequency information is necessary for a sense of musical pitch.

The pitch of pure tones is not exclusively determined by frequency however. While an early study by Stevens (1935) found quite large interactions of pitch and level, subsequent studies (e.g. Terhardt 1974; Verschuure and van Meeteren 1975) have found similar but smaller effects. In summary, the pitch of pure tones above around 4000 Hz rises with an increase in level, while it falls with an increase in level for frequencies below about 2000 Hz. The magnitude of the effect in these ranges can be up to 5 per cent, but is less than 1 per cent for intermediate frequencies, and the magnitude and pattern of results is highly variable between individuals. The pitch of pure tones is also affected by duration. Moore (1973) showed that performance on frequency discrimination gets much better at all frequencies as tone duration increases. There is a sharp deterioration in performance for all durations at higher frequencies, where only place-frequency information is available.

## Pitch of complex tones

The pitch of pure tones is a special case, in that it is not of primary importance to consider information across multiple auditory filters or channels. Sufficient temporal and place information can be obtained by attending primarily to only the single filter centred on the signal frequency (though at high levels there is increased spread of excitation to adjacent channels). However, the vast majority of sounds are complex, consisting of multiple frequency components across a number of channels, and it is in such circumstances that one of the most interesting aspects of pitch perception is manifested—the ability to derive an overall unitary percept from a mixture of frequency information. This skill is essential for the appreciation of

music, given that a typical instrument will produce dozens of partials, yet we can assign a single pitch to this, and hence experience melody and harmony. In fact, an argument for the evolution of pitch perception is that it aids the segregation and identification of sound sources, such as the human voice, by helping to group related partials together, and in so doing assists us in interpreting the surrounding auditory scene (see chapter by Darwin, in Plack *et al.* 2005). While many complex sounds have harmonic spectra, ie. their partials have frequencies that are integer multiples of a common fundamental frequency, pitch can also be derived from stimuli that have inharmonic components (Schouten *et al.* 1962) or that contain only broadband noise in their spectra (Cramer and Huggins 1958; Burns and Viemeister 1976). There are also examples of pitches that arise from the input of different stimuli to each ear (Cramer and Huggins 1958; Houtsma 1972; Klein and Hartmann 1981).

A variety of models have been put forward over the years to explain how the auditory system derives a sense of pitch from such a range of stimuli. Related to the duality of place and temporal coding of frequency, pitch perception models tend to be divided into the *pattern matching* and *temporal* families (for overviews of these models see Houtsma 1995, see also the chapter by de Cheveigné in Plack *et al.* 2005). These models principally differ in the relative importance of frequency components that are resolved or unresolved by the auditory system. As described earlier, a component is said to be resolved if its place and temporal frequency information can be interpreted independently of any other components. Figure 5.2 (after Plack and Oxenham 2005) offers an illustration of the relationship between stimulus frequency spectrum, peripheral filtering, and the resultant excitation pattern and BM vibration. A range of studies (Plomp 1964; Plomp and Mimpen 1968; Moore and Ohgushi 1993) suggest that subjects are able to 'hear out' only the first five to eight components of a harmonic complex tone; these studies offer psychoacoustic evidence that beyond these harmonic numbers, the components remain unresolved by the auditory periphery.

The most common sounds we hear in music, such as sustained notes on strings, woodwinds, brass, most keyboard instruments, and sung vowels, have harmonic spectra similar to the one shown in Figure 5.2 (although of course the relative amplitudes of the spectral components differ between instruments). Given that the pitch of such sounds usually corresponds to the fundamental frequency, it might be tempting to assume that the pitch is simply determined by this lowest frequency component. However, the phenomenon of the 'missing fundamental, robustly demonstrated by Schouten (1938), indicates that this is not the case. When higher harmonic components are present but not the fundamental, the pitch may remain the same as if the fundamental were physically present. In the example shown in Figure 5.2, imagine if the lowest component present were at 200 Hz, and all the other multiples of 100 Hz, i.e. 300 Hz, 400 Hz, 500 Hz, etc., remained. The pitch heard would still be 100 Hz, not 200 Hz. Schouten termed the pitch derived from the higher (non-fundamental) components the *residue pitch*, and it has also been called *periodicity pitch, virtual pitch*, or *low pitch* in the literature. Somehow, the global pattern of frequencies physically present is interpreted to create a pitch at a frequency that may not be physically present, and this effect is not a consequence of the existence of combination tones (Licklider 1956). Both pattern-matching and temporal models have been proposed as underlying the sensation of a unitary pitch experienced in respect of a complex tone.

The preceding discussion has considered the case in which a single sense of pitch is derived from simple or complex stimuli. However, much musical practice relies on the ability of the listener to perceive the relationship between simultaneous pitches. A fundamental aspect of this is the perception of consonance and dissonance. While musically trained listeners may have a specific concept of which intervals are more consonant than others, this assessment is heavily influenced by instruction in a particular musical culture. Of particular interest is the fact that even naive listeners have an intuitive sense of consonance as being the degree to which simultaneous notes sound pleasant, harmonious, or euphonious. This unlearned sense of consonance has been termed 'tonal consonance' by Plomp and Levelt (1965). They found that

when simultaneously presented pure tones had very similar frequencies, a high degree of consonance was observed. However, dissonance increased rapidly as frequency difference was increased, reaching a maximum at around 25 per cent of the auditory-filter bandwidth. Further increases in frequency resulted in a progressive increase in consonance until a maximum was reached and maintained once the critical bandwidth was exceeded. Thus, sensory dissonance is related to the amount of interference between two tones within the same auditory filter. This same mechanism can account for the degree of consonance perceived when complex tones are

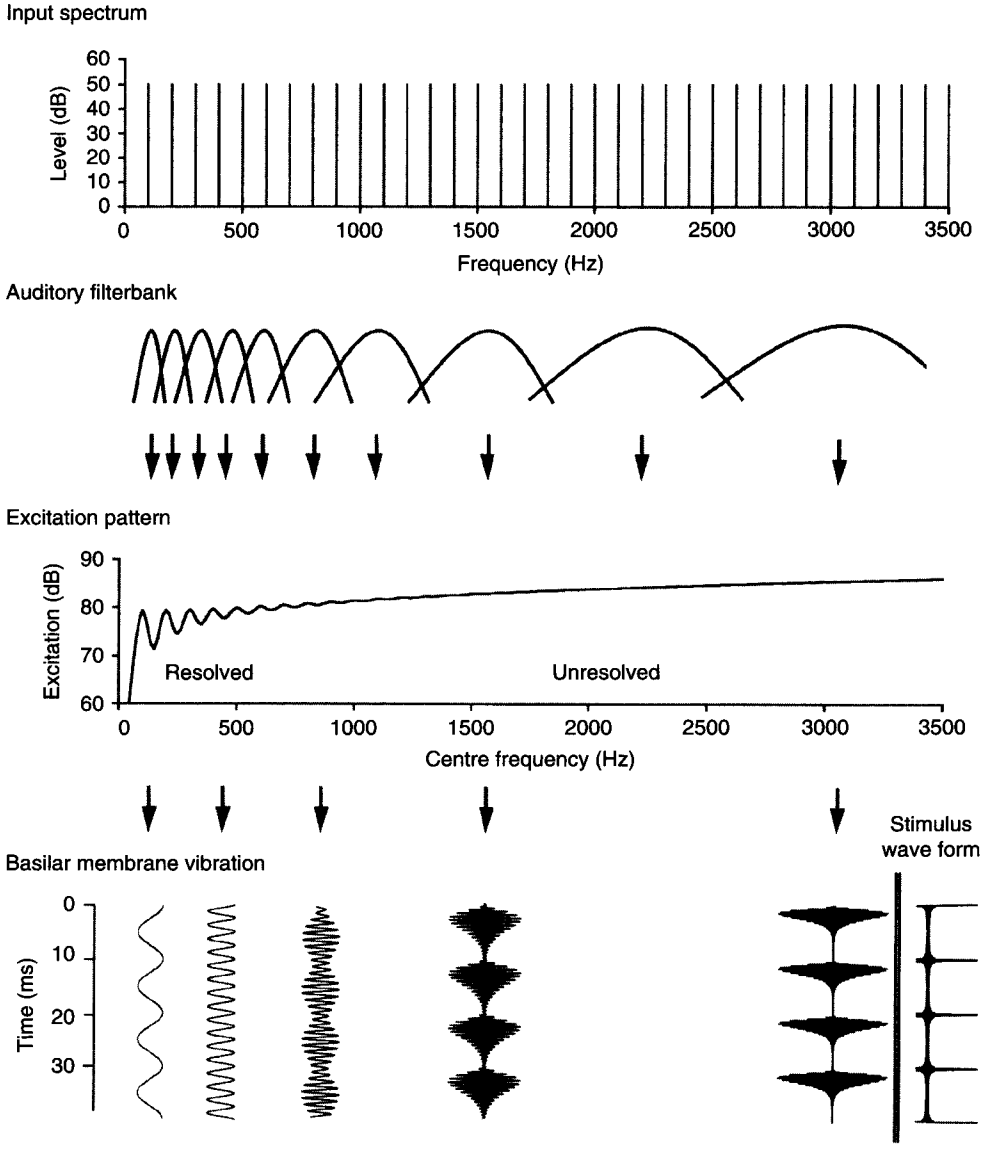



**Fig. 5.2** A schematic illustration of the relationship between a harmonic complex stimulus, its analysis by the auditory filterbank, and the resulting excitation pattern and basilar membrane vibration (from Plack 2005).

when simultaneously presented pure tones had very similar frequencies, a high degree of consonance was observed. However, dissonance increased rapidly as frequency difference was increased, reaching a maximum at around 25 per cent of the auditory-filter bandwidth. Further increases in frequency resulted in a progressive increase in consonance until a maximum was reached and maintained once the critical bandwidth was exceeded. Thus, sensory dissonance is related to the amount of interference between two tones within the same auditory filter. This same mechanism can account for the degree of consonance perceived when complex tones are

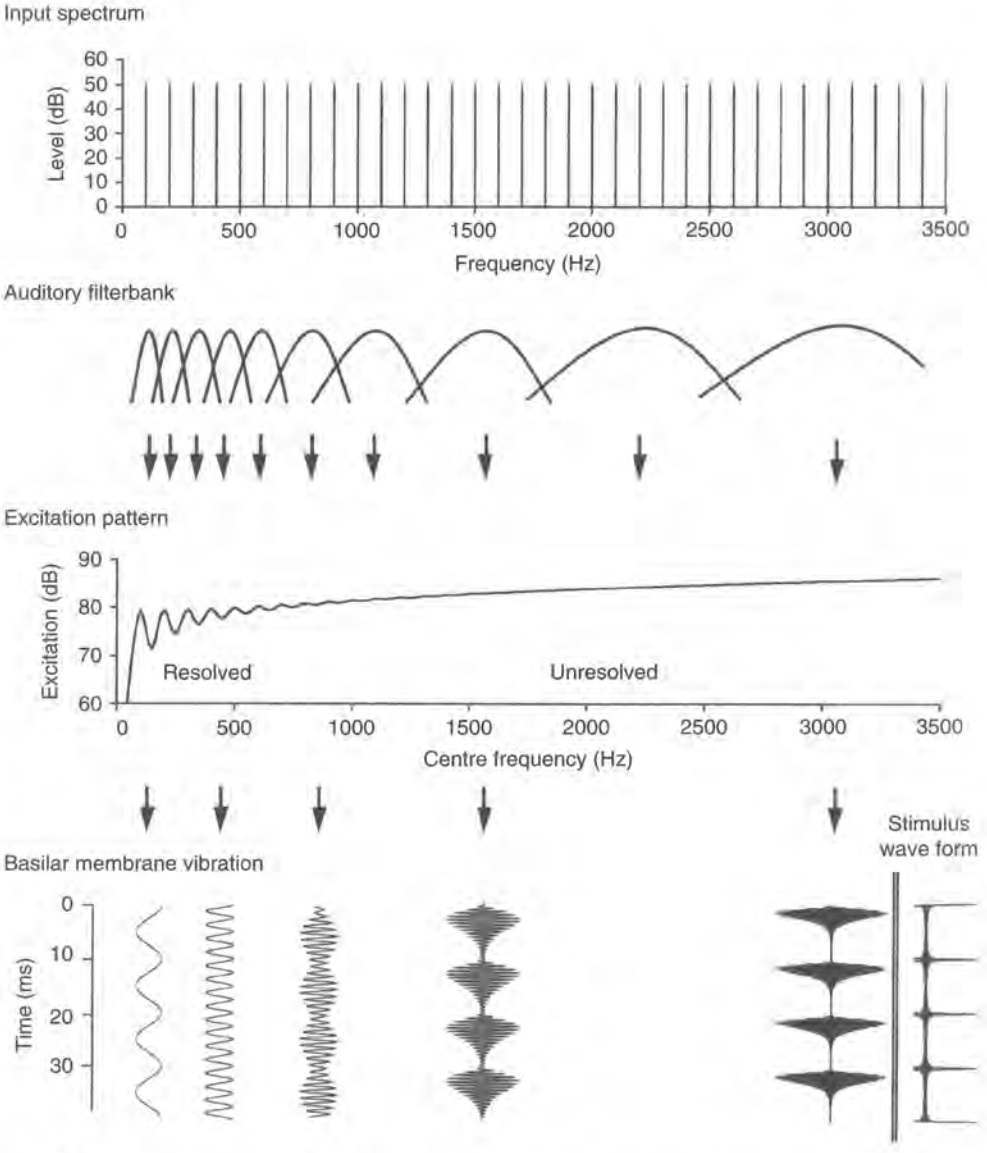



**Fig. 5.2** A schematic illustration of the relationship between a harmonic complex stimulus, its analysis by the auditory filterbank, and the resulting excitation pattern and basilar membrane vibration (from Plack 2005).

presented simultaneously. Unlike the results seen with pure tones however, complex tones do not plateau at a level of maximum consonance once the critical bandwidth is exceeded. Rather, there are peaks in consonance when the fundamental frequencies possess simple integer ratios, e.g. 2:1 (octave), 2:3 (fifth), 3:4 (fourth), etc., as a consequence of the interactions between the individual partials of each tone complex. Dissonance is dependent on the roughness or interference of partials in various auditory filters, even if these are considerably above the filter centred on the fundamental frequency.

## Higher-order processing and organization of pitch

When we experience pitch in music we do not experience single pitches, or arbitrary collections of individual pitches; we experience pitch *patterns*. We can conceptualize such pitch patterns in terms of the vertical dimension producing harmony, while the horizontal dimension produces melody. As Attneave and Olson (1971) put it, musical pitch is a morphophoric—pattern-bearing—medium. In general, the processes outlined above constrain but do not determine the types of patterns that we experience in musical pitch. In order to understand the experience of musical pitch we must take into account processes that operate on, and are in part driven by, the acoustic musical signal, *and* also processes that are mediated by central (largely cortical) mechanisms that have arisen through processes of development and learning. In other words, the perception of musical pitch relies on both bottom-up processes (driven by the musical signal) and top-down processes (arising from our evolutionary history and from individual, culturally situated, learning). The following section will focus on these top-down processes.

### Grouping principles, stream formation and auditory scene analysis

The facility to extract musical patterns from the rich mixture of auditory information that reaches our ears has been termed *auditory scene analysis* by Albert Bregman (Bregman 1990). Although Bregman developed much of his work in terms of the processing of all kinds of auditory stimuli (as evolutionary processes would have shaped us for optimally analysing all events around us), we can take music as a special case in which musical notes are the auditory events of interest. There are two ways in which auditory scene analysis, or auditory grouping, can function. In the case of a solo unaccompanied melody, there is the grouping process whereby we link the successive isolated notes to form one percept of contour. The same processes also allow us to disentangle the mixture of overlapping elements that occurs when multiple sounds are present. This ability to assign some elements to one inferred perceptual auditory source or 'stream', and others to a different stream, has been termed *auditory streaming*. Applied to the vertical dimension, we describe the grouping and parsing of elements that overlap in time as *simultaneous streaming*. It is this process that leads to the perception of harmony, when some frequency components are parsed into different simultaneous notes. The grouping and parsing of temporally non-overlapping notes, termed *sequential streaming*, is what gives rise to a perception of counterpoint. An extensive consideration of how these processes can be shown to generate the commonly accepted rules of harmony and voice-leading in Western music is provided by Huron (2001).

The principle of proximity states that notes are likely to be integrated into the same stream when they are close together in frequency or time. Most melodic writing conforms to a preference for movement by step or small intervals, an adherence to register, and minimal use of rests within a phrase. If other voices or parts are present, the same principles help prevent the notes of one voice being confused with those of others. Where larger melodic leaps do occur, it is very rare for the voice with the large leap to cross over the pitch of an adjacent voice. In the case of unaccompanied melodies, similar principles ensure the internal coherence of a single stream. The breakdown of such coherence, termed *stream segregation*, has been exploited by composers in the use of the technique of *compound melody* or *pseudo-polyphony*, whereby a single instrumental part can produce the percept

of multiple simultaneous streams (a technique popular in much Baroque music: see the works for solo violin or cello by JS Bach). Such examples demonstrate that proximity in the frequency domain is not an independent determinant of streaming, however, as there is a strong interaction with the rate of note presentation. The illusion of multiple interleaved melodies breaks down when the tempo is slowed down sufficiently. Both Miller and Heise (1950) and van Noorden (1975) found that below a minimum pitch proximity of around one semitone, tone sequences were always perceived as integrated, independent of the rate of presentation up to a note inter-onset time of around 500 ms, beyond which larger pitch differences could lead to an integrated percept at these slower rates. The critical pitch difference was termed the *trill threshold* or the *fission boundary* in the two studies, respectively. Van Noorden also identified the *temporal coherence boundary*, which is the pitch difference above which stream segregation is always heard. Between the fission boundary and the temporal coherence boundary, there is a range of ambiguity where tone sequences can be perceived as either integrated or segregated, depending on context and the individual listener.

Two more principles derived from the Gestalt tradition that are relevant to stream formation in the auditory context are those of *common fate* and *good continuation*. These apply both at the level of discrete notes and of spectral components from sounds that may or not be overlapping in time. The principal of common fate groups events that have synchronous features. The principle of good continuation (Bregman and Dannenbring 1973) would lead listeners to group notes together that form a coherent line, and a logical melody. This would work in conjunction with the preference for linear, stepwise or small-interval motion described above.

A further strong streaming or grouping principle in the auditory domain is that of harmonicity. This is the tendency to group together frequency components that all consist of integer multiples of the same fundamental frequency (Moore *et al.* 1985). As many of the sound sources we commonly encounter are harmonic, e.g. sounds deriving from vibrating strings and air columns, it is a very useful cue for grouping components from the same source. The influence of this phenomenon on contrapuntal writing and harmonic voice leading can be seen in the generally accepted practice of avoiding 'parallel fifths and octaves' between voices—such perfect consonances would be instances where all the resolved spectral components of both voices would be harmonically related, and hence an excessive degree or spectral fusion or grouping would result, leading to a loss of independence of the musical voices.

The broad principle of grouping by similarity does not rely on relatedness by pitch; it could work against a particular tendency of a pitch relationship to integrate or segregate a tone sequence. Examples of non-pitch similarity that influence streaming include timbre and temporal fine structure (Roberts *et al.* 2002; Stainsby *et al.* 2004a), level (Stainsby *et al.* 2004b), temporal envelope and apparent spatial location (Hartmann and Johnson 1991). In fact, Moore and Gockel (2002) argue that grouping by similarity can be so strong, that any perceptually salient acoustic attribute can induce streaming. Thus, the final effect of whether integration or segregation is perceived will be the result of the interaction of a number of, possibly competing, factors and the particular sensitivities of the individual listener.

## Pitch in music, music-specific processes and constraints

It is evident that pitch is a continuous percept, which can take any value over a range of more than eight octaves. However, this is not the way musical traditions across many cultures use pitch. Instead, pitch is broken up into discrete piece-wise constant steps or quanta—the degrees of a scale. In Western music, the most common scales are the diatonic modes and the chromatic scale, while other musical cultures have their own individual scale systems, such as the *ragas* of North Indian music, *slendro* and *pelog* scales of Balinese gamelan music, or the *dastgah* of Persian music (Sadie 2001). The fact that musical traditions have adopted scales with discrete steps despite the fact that the most primary of all instruments, the human voice, is not bound by them, suggests that there is a perceptual requirement for such stepwise structuring of pitch. We can perceive frequency differences

that are much smaller than the semitone by which we are bound in Western tonal music (Sek and Moore 1995). Many Arabic scales routinely use quarter-tone steps, yet this does not mean music is composed that uses a 24-step equal-interval scale (Burns 1999). Instead, it uses unequal step sizes, some smaller than the semitone and others larger, so that there are still approximately seven steps to the octave (similar to the Western seven-step diatonic modes). A plausible justification for this is the limited information-handling capacity of the human listener. Across a variety of sensory modes, continua are routinely divided into approximately seven categories (Miller 1956). Humans are intuitively good at having a meaningful 'feel' and reference for approximately seven degrees of a quantity, but errors increase markedly beyond this. The use of musical scales with discrete steps can thus be seen as a method of ensuring the robustness of melodies within a musical tradition.

Linked to the use of discrete steps for a sensory continuum is the phenomenon of *categorical perception*. This refers to the ability to assign the same label to a whole group of broadly similar yet still different stimuli, and the related step-like jump in assignment to the next available label as soon as a stimulus crosses a category boundary. The phenomenon was first described in the perception of speech sounds (Liberman *et al.* 1967), but has since also been demonstrated for non-speech sounds (Locke and Kellar 1973; Miller *et al.* 1976) and for musical pitches (Burns and Ward 1978). A musical example is the tendency to group a range of disparate tunings, some flat and some sharp, to the same note, such as a 'c', while with just a little increase in pitch, a note abruptly changes from being a 'sharp c natural' to being a 'flat c sharp'.

It can readily be seen why a categorical mode of perception is necessary in speech communication, to accommodate the whole range of acoustically diverse sounds that come from different speakers of differing age, sex, and accent. It is plausible that similar forces have shaped the selection of the dominant scales in general musical practice. For example, although Harry Partch (Sadie 2001) developed a whole range of instruments with 43-tone temperament, and Schoenberg advocated the equality of 12 equally tempered semitones, diatonic melodies still dominate the common musical experience in Western performance practice. Admittedly however, there is also a mass of cultural and sociological factors that weigh on any purely perceptually driven evolution of musical fitness and acceptability in these cases.

If it is almost universal for music to be quantized into scale steps, the question arises as to what selects the particular intervals used. Nearly all the examples cited so far use intervals of non-equal size, so it is clear that we do not simply divide the octave neatly into seven equally tempered steps, as may at first appear a mathematically neat solution. The most plausible explanations are that melodic step sizes are selected for intervals that would work well in harmony with other voices, and that unequal step sizes enable listeners clearly to identify and differentiate between serial pattern structures. Harmonic intervals are preferred that maximize consonance, as discussed earlier (Plomp and Levelt 1965), and thus notes that would form such consonant intervals are preferred as scale degrees (Houtsma 1995). In the melodic domain, Trehub *et al.* (1999) showed that infants and adults showed a performance advantage in detecting mistunings in unequal-step scale melodies compared to those presented in equal-step scales.

It is interesting to note how pitch is generally used quite differently in music and language. Speech exhibits a variable pitch contour, with fundamental frequencies ranging over an octave (Fitzsimons *et al.* 2001). Variation in fundamental frequency shapes a phrase and gives clear emphasis to certain words. Meaning can be altered by placement of pitch accent, and whether an utterance is a statement or a question can be cued by the presence or absence of a rising pitch contour at the end of a sentence. This pragmatic function is probably the limit of pitch's contribution to many languages, particularly most European ones (Crystal 1969). However, pitch has explicit lexical significance in tonal languages, such as those found in South East Asia. As an example, Mandarin uses four tones, which when placed on the same combination of consonants and vowels, create words with totally different meanings. This is clearly a specific, categorical and semantic use of pitch,

which goes beyond an auxiliary role of emphasis or clarification. Note however that the four tones of Mandarin are not equatable with four tones of a musical scale. They are categorized as much by contour as by absolute pitch height. The fact that these tones are not arranged in scale steps highlights an important difference between the uses of pitch in music and language. There is still something approximate about pitch in language—there are very broad perceptual categories for pitch and there would need to be a gross difference in pitch before a tone boundary in a tonal language were crossed. Similarly, for the role of intonation and emphasis in European languages, the actual deviation in fundamental frequency contour could be scaled significantly without a change in meaning; it is the shape of the contour that is most significant. Both of these factors clearly do not apply to music in the same way. Sensitivity to tuning is quite precise in music, as only a small deviation from a target pitch is required for a musical performance to be judged as being out of tune. Additionally, melodies cease to be identifiable when the interval range is significantly expanded or compressed (Moore and Rosen 1979). In summary, the use of pitch in music differs from that in speech by articulating complex, set patterns of contour, encoded by the robust categories of scale steps.

## Summary and conclusion

This chapter began with description of the basic physiology and psychophysics that enable pitch to be derived from the acoustic stimulus reaching our ears. The fundamental means of coding frequency in the neural pathways of the auditory system are the dual systems of place coding and temporal coding. While a fairly straightforward relationship can be seen between the physical quantity of frequency and the perceptual quantity of pitch in the case of pure tones, the case is more complicated and can remain ambiguous in the case of complex tones. The auditory system can analyse a range of information from many frequency components to arrive at a unitary percept of overall pitch. The argument was presented that this is an essential skill as part of the process of auditory scene analysis, whereby perceptual representations of distinct sound sources in our environment, so-called auditory streams, can be produced. The processes of auditory scene analysis and auditory streaming were further elucidated by a consideration of other grouping mechanisms such as common onset times, common fate, proximity and similarity. While on one level these processes group the spectral components of individual notes together, on a higher level they also group successive notes together to form melodies and counterpoint.

Finally, a consideration was presented of the different ways pitch is used in music and language. Music relies on the use of discrete, ordered sets of pitches called musical scales. The use of a finite set of discrete steps was argued to facilitate musical memory, melodic recognizability and robustness for transmission. By contrast, pitch in language can be seen to function in two main ways. In non-tonal languages, pitch contour is of primary importance, giving emphasis and inflexion, and distinguishing between a statement and a question. In tonal languages, tones differentiate lexical meaning, and as such absolute frequency values are more fundamental than in non-tonal languages. However, the use of tones in the former still differs significantly from that in music, as the categorical boundaries are significantly broader than in musical scales, and there is not the same use of specifically defined, intervallic steps along a musical scale.

## References

ANSI (1994). *ANSI S1.1–1994. American National Standard Acoustical Terminology*. American National Standards Institute, New York.

Attneave F and Olson RK (1971). Pitch as a medium: a new approach to psychophysical scaling. *American Journal of Psychology*, **84**, 147–166.

Bregman AS (1990). *Auditory scene analysis: the perceptual organization of sound*. Bradford Books, MIT Press, Cambridge, MA

Bregman AS and Dannenbring G (1973). The effect of continuity on auditory stream segregation. *Perception and Psychophyics*, **13**, 308–312.

Burns EM (1999). Intervals, scales and tuning. In D Deutsch, ed., *The psychology of music*, 215–264. Academic Press, San Diego, CA.

Burns EM and Viemeister NF (1976). Nonspectral pitch. *Journal of the Acoustical Society of America*, **60**, 863–869.

Burns EM and Ward WD (1978). Categorical perception—phenomenon or epiphenomenon: evidence from

experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America*, **63**(2), 456–468.

Cramer EM and Huggins WH (1958). Creation of pitch through binaural interaction. *Journal of the Acoustical Society of America*, **30**, 413–417.

Crystal D (1969). *Prosodic systems and intonation in English*. Cambridge University Press, Cambridge.

Fitzsimons M, Sheahan N and Staunton H (2001). Gender and the integration of acoustic dimensions of prosody: implications for clinical studies. *Brain and Language*, **78**(1), 94–108.

Fletcher H (1940). Auditory patterns. *Reviews of Modern Physics*, **12**, 47–65.

Glasberg BR and Moore BCJ (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, **47**, 103–138.

Hartmann WM and Johnson D (1991). Stream segregation and peripheral channeling. *Music Perception*, **9**, 155–184.

Houtsma AJM (1995). Pitch perception. In BCJ Moore, ed., *Hearing*, 267–295. Academic Press, Orlando, FL.

Huron D (2001). Tone and voice: a derivation of the rules of voice-leading from perceptual principles. *Music Perception*, **19**(1), 1–64.

Klein MA and Hartmann WM (1981). Binaural edge pitch. *Journal of the Acoustic Society of America*, **70**, 51–60.

Liberman AM, Cooper FS, Shankweiler DP and Studdert-Kennedy M (1967). Perception of the speech code. *Psychological Review*, **74**, 431–461.

Licklider JCR (1956). Auditory frequency analysis. In C Cherry, ed., *Information theory*, 253–268. Academic Press, New York.

Locke S and Kellar L (1973). Categorical perception in a non-linguistic mode. *Cortex*, **9**, 353–369.

Miller GA (1956). The magical number 7, plus or minus 2—some limits on our capacity for processing information. *Psychological Review*, **63**(2), 1–93.

Miller GA and Heise GA (1950). The trill threshold. *Journal of the Acoustical Society of America*, **22**, 637–638.

Miller JD, Wier CC, Pastore R, Kelly WJ and Dooling RJ (1976). Discrimination and labelling of noise-burst sequences with varying noise-lead times: an example of categorical perception. *Journal of the Acoustical Society of America*, **60**, 410–417.

Moore BCJ (1973). Frequency difference limens for short-duration tones. *Journal of the Acoustical Society of America*, **54**, 610–619.

Moore BCJ (2003). *An introduction to the psychology of hearing*. 5th edn. Academic Press, San Diego, CA

Moore BCJ, Glasberg BR and Peters RW (1985). Relative dominance of individual partials in determining the pitch of complex tones. *Journal of the Acoustical Society of America*, **77**, 1853–1860.

Moore BCJ and Gockel H (2002). Factors influencing sequential stream segregation. *Acta Acustica United with Acustica*, **88**, 320–333.

Moore BC J and Ohgushi K (1993). Audibility of partials in inharmonic complex tones. *Journal of the Acoustical Society of America*, **93**, 452–461.

Moore BCJ and Rosen SM (1979). Tune recognition with reduced pitch and interval information. *Quarterly Journal of Experimental Psychology*, **31**, 229–240.

Palmer AR and Russell IJ (1986). Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hearing Research*, **24**(1), 1–15.

Patterson RD (1976). Auditory filter shapes derived with noise stimuli. *Journal of the Acoustical Society of America*, **59**, 640–654.

Plack CJ (2005). *The sense of hearing*. 1st edn. Lawrence, Erlbaum Associates, Mahwah, NJ.

Plack CJ and Oxenham AJ (2005). The psychophysics of pitch. In CJ Plack, AJ Oxenham, RR Fay and AN Popper, eds, *Pitch: neural coding and perception*, 7–55, Springer, New York.

Plack CJ, Oxenham AJ, Fay RR and Popper AN (eds) (2005). *Pitch: neural coding and preception perception*. Springer, New York.

Plomp R (1964). The ear as a frequency analyzer. *Journal of the Acoustical Society of America*, **36**, 1628–1636.

Plomp R and Levelt WJM (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, **38**, 548–560.

Plomp R and Mimpen AM (1968). The ear as a frequency analyzer II. *Journal of the Acoustical Society of America*, **43**, 764–767.

Rhode WS (1971). Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique. *Journal of the Acoustical Society of America*, **49**, 1218–1231.

Roberts B, Glasberg BR and Moore BCJ (2002). Primitive stream segregation of tone sequences without differences in F0 or passband. *Journal of the Acoustical Society of America*, **112**, 2074–2085.

Robles L and Ruggero MA (2001). Mechanics of the mammalian cochlea. *Physiological Reviews*, **81**, 1305–1352.

Rose JE, Brugge JF, Anderson DJ and Hind JE (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *Journal of Neurophysiology*, **30**, 769–793.

Ruggero MA (1992). Responses to sound of the basilar membrane of the mammalian cochlea. *Current Opinion in Neurobiology*, **2**, 449–456.

Sadie S (ed.) (2001). *The new Grove dictionary of music and musicians*. Macmillan, London. Also available online at http://www.oxfordmusiconline.com.

Schouten JF (1938). The perception of subjective tones. *Proceedings of the Koninklijke Akademie van Wetenschap*, **41**, 1086–1092.

Schouten JF, Ritsma RJ and Cardozo BL (1962). Pitch of the residue. *Journal of the Acoustical Society of America*, **34**, 1418–1424.

Sek A and Moore BCJ (1995). Frequency discrimination as a function of frequency, measured in several ways. *Journal of the Acoustical Society of America*, **97**, 2479–2486.

Stainsby TH, Moore BCJ and Glasberg BR (2004a). Auditory streaming based on temporal structure in hearing-impaired listeners. *Hearing Research*, **192**, 119–130.

Stainsby TH, Moore BCJ, Medland PJ and Glasberg BR (2004b). Sequential streaming and effective level differences due to phase-spectrum manipulations. *Journal of the Acoustical Society of America*, **115**, 1665–1673.

Stevens SS (1935). The relation of pitch to intensity. *Journal of the Acoustical Society of America*, **6**, 150–154.

Terhardt E (1974). Pitch of pure tones: its relation to intensity. In E Zwicker and E Terhardt, eds, *Facts and models in hearing*, 350–357. Springer, Berlin.

Trehub SE, Schellenberg EG and Kamenetsky SB (1999). Infants' and adults' perception of scale structure. *Journal of Experminetal Psychology: Human Perception and Performance*, **25**(4), 965–975.

van Noorden LPAS (1975). Temporal coherence in the perception of tone sequences. Unpublished PhD thesis, Eindhoven University of Technology.

Verschuure J and van Meeteren AA (1975). The effect of intensity on pitch. *Acustica*, **32**, 33–44.

von Békésy G (1947). The variations of phase along the basilar membrane with sinusoidal vibrations. *Journal of the Acoustical Society of America*, **19**, 452–460.

# CHAPTER 6

# Tonal cognition

Emmanuel Bigand and Bénédicte Poulin-Charronnat

THE term *tonal music* can be applied to a large variety of musical styles in the West. This includes that of the four periods (Baroque, Classical, Romantic, and Modern) into which Western art-music is commonly divided, as well as other musical styles from popular traditions such as jazz, rock-pop music, reggae, and salsa. Pieces from these musical styles sound so different that it can be difficult to realize that they share some of the same features. The most basic is that they rest on a single set of 12 pitch classes, referred to as the chromatic scale, *c, c#/db, d, d#/eb, e, f, f#/gb, g, g#/ab, a, a#/bb, b.*[1] These 12 notes are recycled every octave leading to a large number of musical tones. However, the number of pitch classes (12) remains low, (see Chapter 5 this volume for more detailed information about the difference between *pitch height* and *pitch chroma*). This set of pitch classes is organized in subsets of seven tones referred to as diatonic scales. The fact that so many styles in West rest on a low number of pitch classes organized in small subsets of seven notes highlights the combinatory nature of Western music. The purpose of this chapter is to focus on this combinatory aspect. What type of organization in the tonal pitch structure allows producing so many different styles, and so many different pieces inside each style?

In order to address this issue, we start by considering the theoretical distinction between 'tonal' and 'event hierarchies'. The main features of tonal hierarchies are then summarized and we consider how these hierarchies may be parsimoniously represented. The influence of tonal hierarchies on music perception is then reviewed. To conclude this chapter, we briefly consider how tonal hierarchies contribute, along with other musical parameters, to the definition of an event hierarchy specific to a given piece.

[1] For convenience, musical notes will be represented by lower case letters in italics and chords by upper case letters.

## Tonal and event hierarchies in Western music

A critical feature of Western music is its hierarchical organization (Lerdahl and Jackendoff 1983; Meyer 1973; Schenker 1935). This is not to say that all structures in Western music are of hierarchical nature. Associative relations (between thematic cells, for example) also contribute to musical structure. However, the possibility of organizing musical events in a hierarchy of structural importance has deep implications for the listeners. A simple way to understand the concept of hierarchy is to consider that musical events are not all of equal importance: some are structurally important, while others are primarily ornamental. Ornamental events have aesthetic and expressive qualities but they do not contribute to the structure of the piece per se. Removing the ornamental notes does not alter the perceptual identity of a melody, while removing structural tones would modify the melody greatly (Dibben 1994; Mélen and Deliège 1995). If musical events were strictly of equal importance, pieces would sound like monotonous strings of sounds. By contrast, the presence of a hierarchy confers dynamic qualities on musical pieces that have powerful psychological implications: ornamental events seem to 'go towards' events of structural importance, which in turn tend to 'go towards' events

hierarchically more important, and so on across the extent of a piece.

Like linguistic discourses, music is a time-oriented structure that progresses from a beginning to an end. If the music stops at some point before the end, most listeners will realize that it has been irregularly interrupted (Bigand 1993; Palmer and Kruhmansl 1987). In a related vein, playing a piece backwards would strongly alter its perceptual identity because the dynamic relationships between tones would be reversed. Let us consider the set of tones *b-c-d#-e-f#-g*. This set of tones is perceived as a melody in the key of C major when played forwards but as a melody in B major when played backwards (Bharucha 1984). In the former case, the tones *b*, *d#*, and *f#* are perceived as ornamental tones anchored on the structurally important tones of the C major chord (*c-e-g*). Playing the tune backwards reverses the relationships between musical events, and the tones *g*, *e*, and *c* are now perceived as ornamental tones anchored to the structurally more important tones of the B major chord (*f#-d#-b*). This change in perceptual organization may be metaphorically compared with the change in meaning that occurs when a sentence is read in a retrograde way (the boy calls the girl *versus* the girl calls the boy). In the first sentence, *the boy* is anchored to the verb *calls*. In the second sentence, *the girl* is anchored to that verb. In language, the main aspect of syntactic computation is to understand 'who is doing what to who'. Similarly, the core aspect of tonal cognition is to perceive 'which tones anchor which tones'. Solving such a problem reminds us of syntactic computation, because musical events can be deeply anchored to the overall structure of the piece: a given tone can be anchored to another one, which may be, in turn, anchored to another one, and so on. Organizing the pitch events of a piece into a single coherent structure, in such a way that the pitch events are heard in a hierarchy of relative importance, is the most fundamental aspect of tonal cognition for Western music (Lerdahl and Jackendoff 1983).

Where do hierarchies between events come from? Two types of hierarchies have been distinguished in music cognition (see Bharucha 1984; Krumhansl 1990; Lerdahl 1989). The term 'tonal hierarchy' was used to designate an atemporal schema of pitch regularities, specific to Western music, which is stored in long-term memory (see Chapter 2 this volume). It embodies the hierarchical relations that accrue to an entire tonal system beyond instantiation in any particular piece. It is atemporal in that it represents more or less permanent knowledge about the musical system rather than a response to a specific sequence of events. Tonal hierarchies thus designate the regularities in pitch that prevail for all styles of tonal music. By contrast an 'event hierarchy' is a hierarchy of specific pitch–time events inferred from the ongoing temporal sequence of musical events. In the next sections, we will focus on tonal hierarchies however, in conclusion, we will briefly consider how tonal hierarchies influence event hierarchies.

## Tonal hierarchies: intra-key hierarchies and inter-key distances

The main aspects of tonal hierarchies can be summarized in the following way. From a set of 12 pitches, several subsets of seven pitches are defined, each defining a musical key. A musical key is defined in reference to a specific tone which is called the tonic tone and that gives its name to the key. Since there are 12 tones, there are 12 possible tonics. In Western music, some subsets of tones form major keys. Others form minor keys. For example, the C major key is made of the following tones: *c-d-e-f-g-a-b*. The C minor harmonic key is made of the tones, *c-d-e♭-f-g-a♭-b*. In both cases, the pitch *c* is the tonic. Major and minor correspond to two different scales of pitch intervals. For example, the set of pitch intervals separating the tones of the C major key is tone-tone-semitone-tone-tone-tone-semitone. This set defines a scale that is constant for all major keys. In a minor harmonic key, the scale is different: tone-semitone-tone-tone-semitone-one tone and a half-semitone. Applying these major and minor scales to each tone of the chromatic scale defines 24 major and minor keys.

The critical aspect of the tonal pitch system relates to the organization of pitches within keys (referred to as intra-key hierarchies), and to the organization that exists between the 24 major and minor keys (referred to as inter-key distances). Kruhmansl and collaborators (see Krumhansl

1990, for a complete account) have investigated intra-key hierarchies with a 'probe tone' method, which consists of playing a test tone (the probe) after a musical context in a given key. Participants have to evaluate on a seven-point scale how the probe tone fits with the context. The 12 tones of the chromatic scale are presented as probe tones. It was found that all the probe tones did not received the same ratings of 'goodness of fit'. In major key contexts, the highest rating was given to the tonic note (the tone *c* in C major), followed by the fifth degree of the scale (*g*), and then the third degree (*e*), the fourth (*f*) and the remaining tones of the key (*d-a-b*). The tones that do not belong to the key (the 'non-diatonic tones'), received the lowest ratings. This set of values defined the C major key profile (Figure 6.1, top). The reader could easily obtain the profiles of all the other keys by translating these profiles to all the remaining tonics. As illustrated by Figure 6.1 (bottom), the ratings found when minor keys were used as context differed in an interesting way; the tonic remains the tone that fits the best with the key context, but then comes the third scale degree (*e♭*), followed by the fifth and the sixth scale degrees (g and *a♭*, respectively). The other diatonic tones received lower ratings, and the non-diatonic tones, the lowest. The minor key profile remains the same for all other minor keys but it has to be translated towards the tonic of the given key context. Important tones in these key

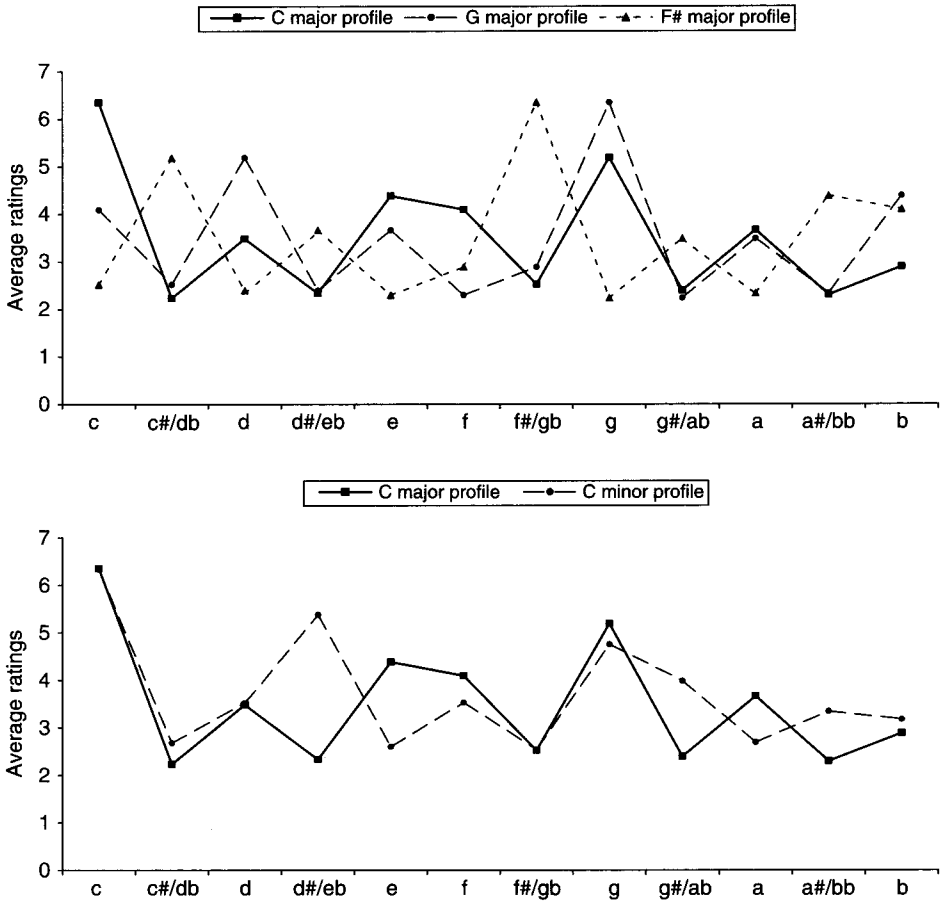



**Fig. 6.1** Key profiles in the keys of C, G and F# major (top) and in the key of C major and C minor (bottom), after Krumhansl (1990).

profiles act as cognitive reference points that anchor all other events. In Western tonal music, the tonic is thus the most important cognitive reference point that anchors all the other tones (Schenker 1935).

Within a key, tones may be grouped to define larger musical units called chords. A chord is the simultaneous or sequential sounding of, at least, three different tones.[2] The three basic component tones of a perfect major or minor chord (also called a triad) are separated by one scale degree. For example, the C major triad is made of the tones *c-e-g*. The E minor triad is made of the tone *e-g-b*. The first interval defines the third of the chord. In a major chord, the third corresponds to a pitch interval of 4 semitones (*c-e*). In a minor chord, the third is made of a pitch interval of three semitones (*c-e♭*). The second interval defines the fifth of the chord. Since there are seven tones per scale, there are seven chords in each key, some being major, others minor, and others diminished.[3] Over all the keys, there are 12 major, 12 minor, and 12 diminished chords. Chords are important events for Western music: they define harmonic frames from which thematic units derive. Vincent d'Indy (1900) metaphorically compared the themes of a musical piece to the actions of a drama, and the relationships between chords to the places where the actions occurred. This points up the fact that the relationships between chords (i.e., the harmony) are of considerable importance in defining the structure of a given piece.

In Western music, the relationships between chords are also of a hierarchical nature; some chords are structurally more important than others. Using the probe tone method, Krumhansl *et al.* (1982) investigated the hierarchies between chords. Goodness of fit ratings varied significantly with the chords: in major key context, chords build on the first scale degree (referred to as the tonic chord) received the highest ratings, followed by the chords build on the fourth and fifth scale degrees (respectively referred to as subdominant and dominant chords). The other chords received lower ratings. Among all the major chords, those that did not belong to the key context received the lowest ratings. In other words, a harmonic hierarchy was found between chords.

The last main aspect of tonal hierarchies relies on the distance between the 12 major and 12 minor keys. Some keys share more tones and chords than others. For example, the key of C and G major share 6 tones and 4 chords, while the key of C major shares only 3 tones and no chord with the key of E major. These relationships are reflected on the chromatic circle of fifths (C-G-D-A-E-B-F#/G♭-C#/D♭- G#/A♭-D#/E♭-A#/B♭-F). The higher the number of steps between two keys, the fewer the common tones and chords shared by the keys. Major keys share also tones and chords with minor keys. For example, C major and A minor share six tones and three chords. The number of tones and chords shared by keys contributes to define musical distances between the keys, with keys sharing numerous tones and chords being closer than others.

From a psychological point of view, the musical distances between keys is more complex that this simple computation. Krumhansl (1990) demonstrated that the perceptual distance between keys is expressed by the correlation that exists between the hierarchical profiles of the keys: the higher the correlation, the closer the keys. This finding suggests that keys are not only related because they share several tones, but also because hierarchically important tones in one key continue to be of importance in others. The hierarchical organization of tones and chords thus changes less between two related keys than between two distant ones. For example, the keys of C major and C minor are musically related although they do not share many tones and chords (four tones, and no chord). However, in both keys, the notes *c* and *g* act as the two most important reference points which anchor the other tones. In a similar way, the key of C major is close to its relative minor key (i.e., A minor). Indeed, the tones *c* and *e* that are hierarchically important in the former key continue to act as referential points in the latter. As a consequence, every major key (C, for example) is close to two minor keys (its parallel and relative minor keys; C

[2] The sounding of two sounds defines an interval, not a chord. More complex chords contained additional tones. For example, the famous dominant seventh chord is made of the tones *c-e-g-b♭* in the F major key, and chords containing two or three ornamental tones are frequent in jazz and bossa nova (as in the major ninth chord *c-e-g-b-d*, or the minor ninth chord *d-f-a-c-e*).

[3] The C major key is made of the following chords: C major (*c-e-g*), D minor (*d-f-a*), E minor (*e-g-b*), F major (*f-a-c*), G major (*g-b-d*), A minor (*a-c-e*), and B diminished (*b-d-f*).

and A minor) and two major keys (those that are separated by one step on the circle of fifths: G and F major). Every major key is also indirectly close to the relative minor keys of the major keys that surround it (C major is close to E minor which is the relative minor of G major). The 12 major and 12 minor keys thus define a complex net of musical relationships (Figure 6.2, bottom).

Moving from one key to another results in a mental reorganization of the hierarchical values of tones and chords. This reorganization is as strong as the new key is distant from the previous one. These changes in cognitive reference points are experienced in an expressive way by listeners. Moreover, the expressive quality of a modulation (i.e., a change in key) is held to depend upon the distances between the two keys: the farther away the keys, the stronger the cognitive reorganization of reference points, and the stronger the expressive effect. As the 12 major and 12 minor keys define a complex net of keys, an almost infinite number of ways exists to modulate from one key to another and to express different feelings for/in listeners.

## Two models of tonal hierarchies

Notes, chords, and keys define three levels of entwined structures that allow Western composers to invent an infinite number of expressive musical pieces. An important issue for cognitive psychology is to capture in a single parsimonious model the creative power of such a musical system. In the present section, we will consider two models of Western pitch structures that derive from different scientific backgrounds.

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | c | | | | | | | | | | | |
| 4 | c | | | | e | | | g | | | | |
| 3 | c | | | | e | | | g | | | | |
| 2 | c | | d | | e | f | | g | | a | | b |
| 1 | c | c#/d♭ | d | d#/e♭ | e | f | f#/g♭ | g | g#/a♭ | a | a#/b♭ | b |

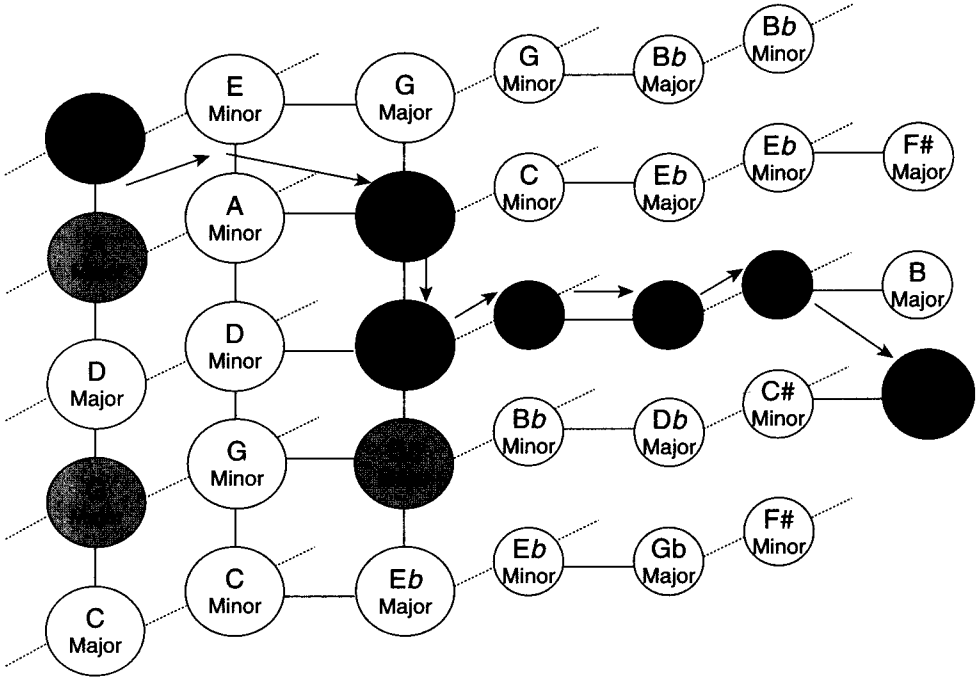



**Fig. 6.2** Pitch class level (top) and regional level (bottom) of Lerdahl's tonal pitch space theory. The arrow illustrates the journey through pitch space of Chopin's E major prelude (adapted from Lerdahl 1991).

and A minor) and two major keys (those that are separated by one step on the circle of fifths: G and F major). Every major key is also indirectly close to the relative minor keys of the major keys that surround it (C major is close to E minor which is the relative minor of G major). The 12 major and 12 minor keys thus define a complex net of musical relationships (Figure 6.2, bottom).

Moving from one key to another results in a mental reorganization of the hierarchical values of tones and chords. This reorganization is as strong as the new key is distant from the previous one. These changes in cognitive reference points are experienced in an expressive way by listeners. Moreover, the expressive quality of a modulation (i.e., a change in key) is held to depend upon the distances between the two keys: the farther away the keys, the stronger the cognitive reorganization of reference points, and the stronger the expressive effect. As the 12 major and 12 minor keys define a complex net of keys, an almost infinite number of ways exists to modulate from one key to another and to express different feelings for/in listeners.

## Two models of tonal hierarchies

Notes, chords, and keys define three levels of entwined structures that allow Western composers to invent an infinite number of expressive musical pieces. An important issue for cognitive psychology is to capture in a single parsimonious model the creative power of such a musical system. In the present section, we will consider two models of Western pitch structures that derive from different scientific backgrounds.

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | c | | | | | | | g | | | | |
| 4 | c | | | | e | | | g | | | | |
| 3 | c | | | | e | | | g | | | | |
| 2 | c | | d | | e | f | | g | | a | | b |
| 1 | c | c#/db | d | d#/eb | e | f | f#/gb | g | g#/ab | a | a#/bb | b |

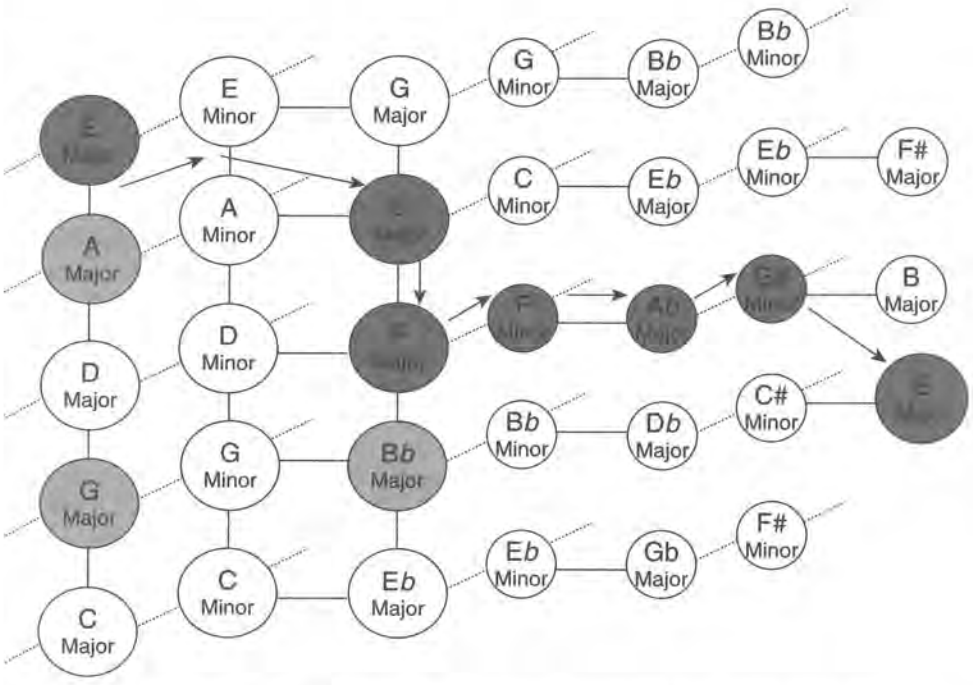



**Fig. 6.2** Pitch class level (top) and regional level (bottom) of Lerdahl's tonal pitch space theory. The arrow illustrates the journey through pitch space of Chopin's E major prelude (adapted from Lerdahl 1991).

The Tonal Pitch Space Theory (TPST), initially described by Lerdahl (1988) and further developed in Lerdahl (1991, 1996, 2001) rests on a long tradition of representing the structure of Western music as a multidimensional space (Longuet-Higgins 1978; Krumhansl 1990; Shepard 1982). In such an approach, tonal hierarchies are represented in a space so that the distances between tones and chords from the instantiated tonic correspond to their relative hierarchical importance: the smaller the distances from the tonic, the stronger the importance. Lerdahl's TPST provides a formal means of quantifying the tonal distances maintained between any two events belonging to any key. Tonal hierarchy is represented in three embedded levels, the first two (pitch class and chordal level) representing the intra-key hierarchy, while the third (regional level) represents the distance between keys. The pitch class level expresses the relationships between the 12 pitch classes. It contains five sublevels corresponding to the chromatic level (1), the diatonic level (2), the triadic level (3), the fifth level (4) and the tonic level (5) (Figure 6.2, top). This basic space expresses Krumhansl and Kessler (1982)'s key profiles of Figure 6.1. In a given context (say, a C major chord in C major), the tonic tone (*c*) will be represented at all five levels. The dominant and the mediant tones (*g* and *e*) will be represented up to levels four and three, respectively. The diatonic tones are represented at level two and non-diatonic tones are represented at level one only. The level at which a given pitch class is represented thus expresses its importance in the tonal hierarchies of the key context. Hierarchies between pitch classes change as a function of context. For example, a C major chord in G major will produce a single change in pitch class (the f# will be now represented at level 2 and the f at level 1). The core idea of TPST is to quantify the tonal stability of each event (i.e., its place in the tonal hierarchy) by computing the number of changes in pitch classes it introduces in the context.

Let us consider the second level of the model, which involves the hierarchies between chords within a given key. The hierarchical values of the chords depend upon their distance to the tonic. This distance is measured in the model by the number of steps that separates the chords on the diatonic circle of fifths (C-G-D-A-E-B-F) and the number of changes in pitch class created by the second chord. In C major, the G major chord stands one step apart from the C major chord on this circle, and it creates four changes in pitch classes by reference to the C major chord. As a consequence, its distance to the tonic chord equals five. Similarly, the tonal pitch space distance of the subdominant chord (F) is five, but the distance of the other chords, hierarchically less important, is higher (e.g., eight for the supertonic chord D minor). These distances express the local changes in hierarchies between tones that prevail when music moves from one chord to another. These changes are experienced by listeners as musical tensions (when the distance increases), or as musical relaxations (when the distance decreases) (Bigand *et al.* 1996). The set of distances traveled through pitch space by a musical sequence thus captures the dynamic quality of the sequence.

The two previous levels would suffice to account for intra-key hierarchies. The third level (regional level) deals with the distances between the 24 keys. The model accounts for key distances by measuring the number of changes in hierarchies between tones and chords consecutive to a shift towards a new key. Not surprisingly, the TPST computes small distances between keys that are one step apart on the chromatic circle of fifths (C, G, and F, for example, which have a distance of 7) and between keys that are related according to parallel and relative minor/major relations (C and C minor, C and A minor). The greatest distance computed by the TPST (30) is found between a given key and its augmented fourth key (C and F# major keys). The distance between two chords belonging to different keys depends upon the way the two chords are interpreted. Given that chords may belong to different keys, this distance can be reduced when the two chords are interpreted as chords of closely related keys. For example, the distance between a C major chord and a C# minor chord equals 30 if the latter chord is analysed as the tonic chord of the C# minor key. The same distance equals 23 if the C# minor chord is analyzed as a member of the E major key. By default, the model would compute the shortest distance. Figure 6.2 (bottom) illustrates how the 24 keys may be organized in a two-dimensional space. Circle of fifth relationships are represented on the vertical axis, and

parallel/relative minor/major relationships are represented on the horizontal axis. Modulating to G, F, E, or C minor from the C major key would result in smaller movements than modulating to a musically less related key (e.g., A♭).

Lerdahl suggests that listening to music corresponds to a journey through pitch space. The arrows of Figure 6.2 illustrate a journey provided by the Chopin prelude in E major. The first part of the prelude remains in E major, with local movements from one chord to another. The second part introduces an intense journey from E to the A♭ major and a fast return to the main key of E. There seems no doubt that the expressive quality of this prelude comes from the velocity at which this journey from E to A♭ occurs during the second phrase (Bigand 2003). The TPST is a musical model of tonal hierarchies that has both theoretical and empirical roots, and that has received support from several empirical studies (Bigand and Parncutt 1999; Bigand *et al.* 1996; Krumhansl 1996; Lerdahl and Krumhansl 2007; Smith and Cuddy 2003; Vega 2003). The TPST is helpful in investigating challenging assumptions about tonal cognition and it has a number of psychological implications (see Bigand 2003). However, it may be viewed as a rather ad hoc complex model, difficult to implement in an automatic way in an artificial system, and of which the architecture is too abstract, as it stands, to correspond point by point to real cognitive processes.

An alternative account of tonal hierarchies that has roots in cognitive psychology and computational sciences has been proposed by Bharucha (1987). Following McClelland and Rumelhart's interactive model of lexical access (McClelland and Rumelhart 1981; Rumelhart and McClelland 1982), Bharucha defines a three-layer interactive model that links tones, chords, and keys. In this model, tone units are connected to all major and minor chords to which they belong (Figure 6.3a). These chords are, in turn, linked to each of the major key units to which they belong. An important musical limitation of the model is that only major and minor triads are represented, and that the model does not contain minor keys. However, the model captures interesting features of tonal hierarchies. For example, when a given triad is sounded (say the tones *c-e-g*), the activated tone units send activation towards all the chord units to which they are connected (Figure 6.3b). At the second iteration, these chord units propagate activation towards the keys to which they are related, and backward to the tones they are connected to (Figure 6.3c). As a consequence, tone units that were not in the stimulus start to be activated. For example, the tone *a* receives some activation from the F and A major chords and the A minor chord that were stimulated by the tones *c* and *e*. At the third iteration, these new activations of the tones propagate towards the chord level. At the same time the activation of the key level spreads backward to the chord units (Figure 6.3d). Since the C major key was active, all the chords of the C major key now start to be stimulated. During the next iterations, all of these activations spread backward and forward up to a point of equilibrium where spreading activation is so small that it does not change the overall state of the network. The interesting point is that this model of MUSical ACTivation (MUSACT) nicely expresses tonal hierarchies. For example, after a C major triad, the most activated chord units correspond to the tonic chord (C), the dominant and subdominant chords (G and F). The activations decrease progressively as long as chords are musically distant from the C major chords. The pattern of activation found for the key units also reflects the inter-key distances. For example, after a perfect cadence in the C major key, the C major key unit is the most activated unit, and this activation decreases as a function of the musical distances from the C major key.

In this model the knowledge of tonal hierarchy is expressed by the pattern of activation that spreads automatically towards tone, chord, and key units. Closely related events lead to a highly correlated pattern of activation, distant events, to negative correlations. In a given context, the amount of activation found for each unit (tones or chords) represents their hierarchical importance in the context. MUSACT thus captures features of tonal hierarchies that are strongly similar to those accounted for by Lerdahl's TPST, but with a completely different architecture. In MUSACT, the tonal hierarchies are represented in a distributed way by the weights of the connections that link tones to major and minor chord units, and chord units to key units. The specific values of these weights were defined ad hoc in the pioneer work by Bharucha (1987),

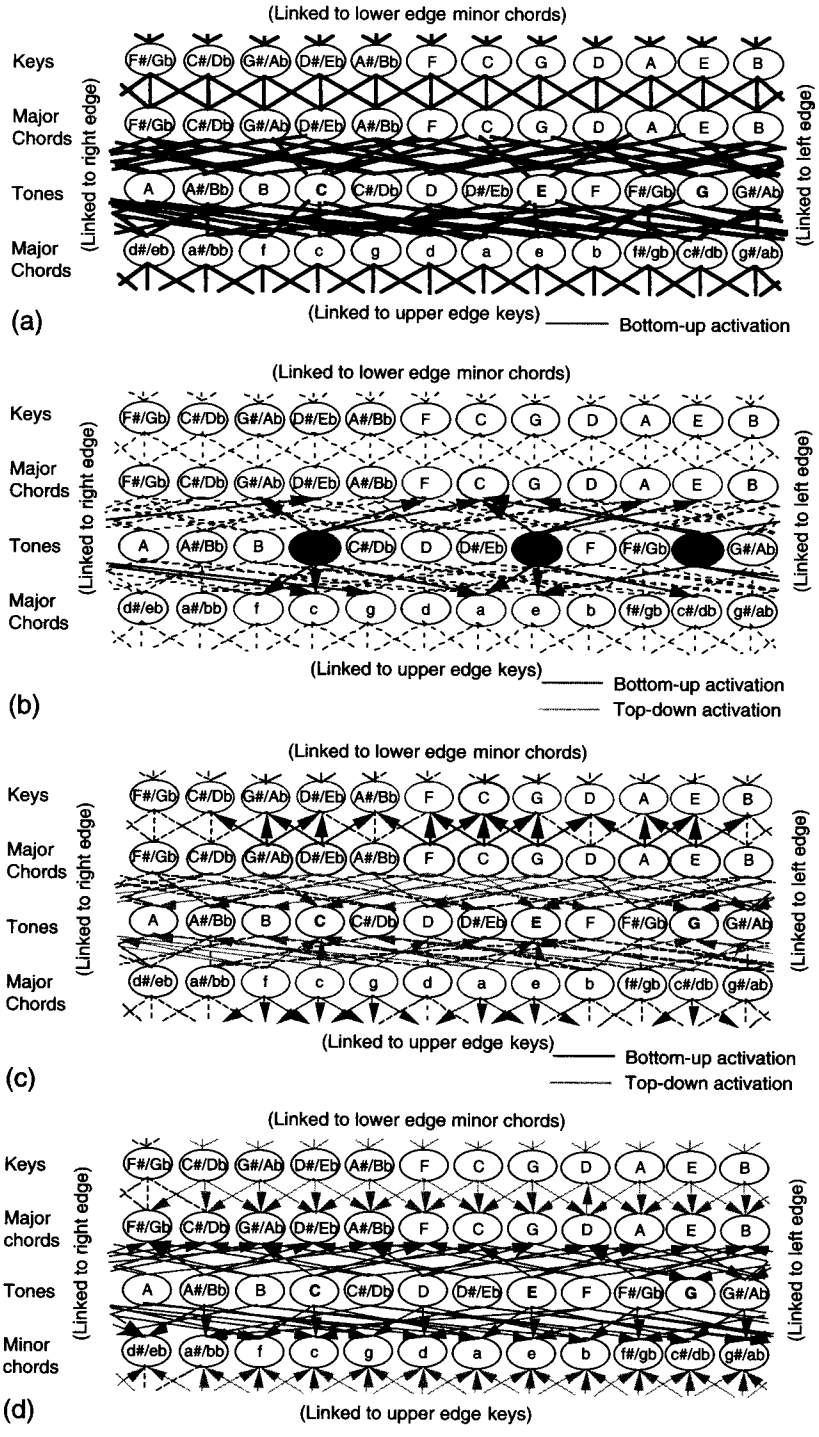



**Fig. 6.3** Bharucha's MUSACT model. (a) When three tones are sounding, activation spread from tones to chord units (b), and then from chord to key and tone units (c), and from key units to chord and from tone to chord units (d) and so on up to equilibrium. See also colour plate section.

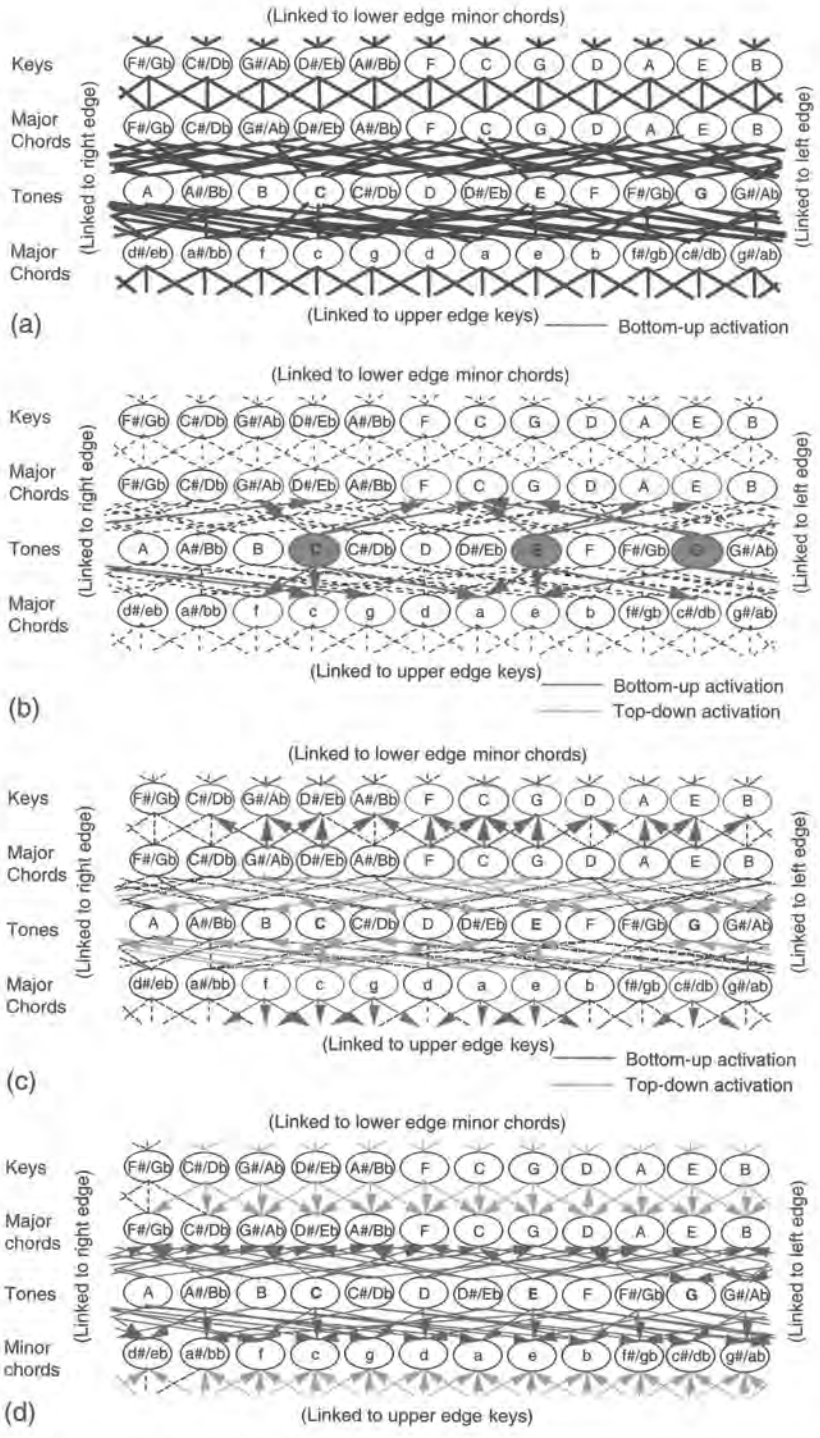



**Fig. 6.3** Bharucha's MUSACT model. (a) When three tones are sounding, activation spread from tones to chord units (b), and then from chord to key and tone units (c), and from key units to chord and from tone to chord units (d) and so on up to equilibrium. See also colour plate section.

but it has been demonstrated that these weights may arise from a learning process that occurs automatically when the neural network is exposed to Western music (Tillmann *et al.* 2000).

## Influences of tonal hierarchies on Western music perception

Much evidence has been collected demonstrating the influence of tonal hierarchies on music perception. For example, Francès (1958) required musicians to detect mistuned notes in piano pieces. This mistuning was performed in different ways. In one condition, some musical notes were mistuned in such a way that the pitch interval between the mistuned notes and those that anchored them was reduced. For example, the leading note (*b* in C major) which is hierarchically weakly important is generally anchored by the tonic note (*c* in C major). Francès mistuned the leading tone by increasing its fundamental frequency (F0) so that the pitch interval between the leading tone and the tonic was reduced. In the other experimental condition, this mistuning was performed in the opposite way. When played without a musical context, participants easily perceived both types of mistuning. Placed in a musical context, only the second type of mistuning (which conflicted with musical anchoring) was perceived. This outcome shows that the perceptual ability to perceive the shift of the F0 of a musical note is modulated by top-down processes that integrate the note into tonal hierarchies. More generally, performers such as violinists or cellists constantly adjust the fundamental frequency of the tones in order to make as clear as possible the way in which these tones are anchored by the hierarchically important tones of the context. This expressive intonation, which is entirely related to the tonal hierarchy, considerably increases a listener's comprehension of the musical discourse.[4]

Tonal hierarchies also influence the perception of melodic contour. Several experiments demonstrate that discriminating a small change in a contour is more difficult when the standard and the comparison melodies are in distant, rather than close, keys (Dowling 1986). Several experiments run by Krumhansl and collaborators (see Krumhansl 1990) demonstrate the influence of tonal hierarchy on the perceptual distance of tones and chords and on their memorization. This influence can be summarized in terms of three contextual principles. The perceptual distance between two events (tones or chords) increases when these events are of hierarchical importance (contextual distance). The perceptual identity of two instances of a given event increases with its hierarchical importance (contextual identity). Two occurrences of the chord C would sound more similar if they occurred in a C major key context rather than in a distant key context (say, in an F# key context). Finally, the perceptual distance of two events depends on their sequential occurrence: the distance decreases when the hierarchically strong important event occurs after the other one (i.e., anchoring principle). It increases in the other case (contextual asymmetry). Krumhansl and collaborators demonstrate that these principles account for confusion errors in memory. For example, it is easier to detect that two occurrences of a seven chord sequence are identical when these chords belong to a single key (diatonic sequence; 69.3 per cent) than when they belong to different keys (55 per cent), illustrating the principle of contextual identity. When the second occurrence of the sequence contains a different chord, it is easier to detect the difference when the chord is changed for a chord belonging to another key than when it remains in the same key (86 per cent *versus* 56.9 per cent). This illustrates the principle of contextual distance. Finally, when the initial sequence is diatonic and the comparison sequence contains a non-diatonic chord, the change is easier to detect (86.8 per cent) than when the first sequence contains a non-diatonic chord that is replaced in the second sequence by a diatonic chord (61.3 per cent). This illustrates the principle of contextual asymmetry.

The influence of tonal hierarchy on the memorization of melody may be very impressive. In one experiment (Bigand 1997), participants were required to listen several times to a target melody

[4] The reader would find a remarkable demonstration of such expressive intonation in Casals' performances of the cello suite by Bach.

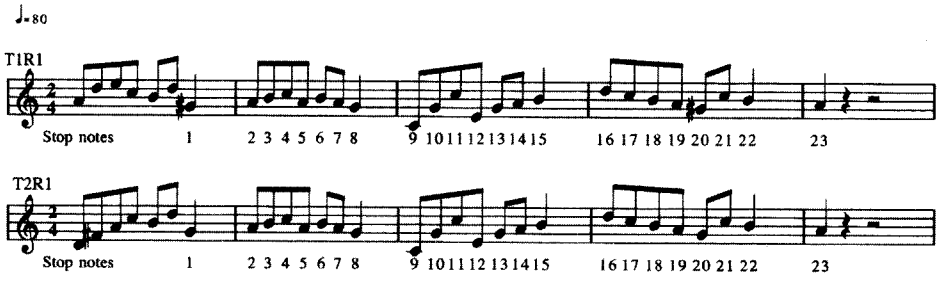



**Fig. 6.4** Examples of standard and comparison melodies used in Bigand (1997).

such as the melody in A minor represented in Figure 6.4 (top). Participants were then presented with new melodies and were asked to evaluate how many pitches had been changed in comparison to the standard. Figure 6.4 (bottom) displays one of the comparison melodies used. As can be seen, it contains most of the tones of the standard. Only a few tones have been changed in order to instill the key of G major at the beginning of the melody. This new key context was assumed to modify the hierarchical weights of the melodic tones. Indeed, the *a* is no longer a tonic tone in this comparison melody, but a weakly important supertonic tone. The reader is invited to use the key profiles of Figure 6.1 to understand how much the hierarchical weights of the tones have been changed in the comparison melody. Not surprisingly, the participants (musically trained and untrained listeners) evaluate that about 66 per cent of the tones have been changed in the comparison melody, that is to say, much more than what has actually been done.

The influence of tonal hierarchies has been reported in several other studies in which participants were asked to evaluate the degree of completion or the degree of musical stability experienced at different points of a musical sequence. Ratings of completion or musical stability were higher when the music stopped on hierarchically important events (Bigand 1993, 1997; Palmer and Krumhansl 1987). A similar result was found for chord sequences (Bigand and Parncutt 1999; Bigand *et al.* 1996). In recent studies, a continuous measure of musical stability has been used, and data were found to correlate significantly with tonal hierarchy (Krumhansl 1996; Lerdahl and Krumhansl 2007). Other studies have investigated the influence of tonal hierarchies on the speed at which some musical features are processed. In Bharucha and Stoeckig's (1986) chord priming experiments, the participants were required to decide accurately but as fast as possible whether a target chord was in tune or out of tune. For the purpose of the experiment one component tone of the chord was slightly mistuned in half of the trials. The target chord was preceded by a chord prime. The critical point of the study was to assess whether the processing of the sensory consonance of the target chord would be modulated by the harmonic relationship of the two chords. The authors provided evidence that consonant target chords were more accurately and more quickly identified as consonant when they were preceded by a related prime (e.g., C and G chords) than by an unrelated prime (e.g., C and F# chords). An opposite finding was found for dissonant targets. Bharucha and Stoeckig argued that a given context primes the processing of chords that are musically related in the context, a finding which is reminiscent of semantic priming effects reported in language (McNamara 2005). Further experiments attempted to demonstrate that this priming effect in music occurs at a cognitive level and does not result from the overlap in harmonic spectra of both chords (Bharucha and Stoeckig 1987; Tekman and Bharucha 1998).

The influence of tonal hierarchies on perceptual expectancy in long musical contexts has also been investigated with the priming paradigm (Bigand and Poulin-Charronnat 2006, for a review). In our experiments, participants were presented with chord sequences (of 8 or 14 chords depending on the studies) that ended on

a target chord. The target acted either as a tonic or as a subdominant chord depending upon the context. The chord preceding the target was always kept constant, except when the local and global harmonic contexts surrounding the target were simultaneously manipulated (Tillmann *et al.* 1998). In most of the studies, participants were required to indicate as quickly and as accurately as possible whether the target chord was consonant or dissonant. Dissonant targets contained a supplementary tone, a half tone above or below a triadic tone. In more recent studies, target chords were played with two different timbres which participants had to identify as fast and accurately as possible (Tillmann and Lebrun-Guillaud 2006). Processing the sensory consonance of the target or its harmonic spectra was always faster and easier when the target acted as a hierarchically important tonic chord. This effect is extremely robust since it has been replicated for relatively long chord sequences organized at three hierarchical levels (Bigand *et al.* 1999), for normal and scrambled musical sequences (Tillmann and Bigand 2001), for chord sequences that have extremely poor voice leading (Poulin-Charronnat *et al.* 2005a), and even when sensory priming was manipulated in order to favour the subdominant target (Bigand *et al.* 2003). Harmonic priming was also found to be stronger than repetition priming (Bigand *et al.* 2005), and to modulate the processing of linguistic information in sung music (Bigand *et al.* 2001; Poulin-Charronnat *et al.* 2005). Finally, it was observed for musically trained and untrained listeners (Bigand and Poulin-Charronnat 2006, for a review), as well as with young children (Schellenberg *et al.* 2005). This large set of consistent data highlights the strength of tonal hierarchies on the formation of musical expectancies.

It is likely that tonal hierarchies influence several other aspects of music perception. In one recent study, we investigated how the tonal hierarchy influences emotional responses to music (Filipic and Bigand 2003). Participants were required to decide as fast and as accurately as possible whether two musical sequences expressed the same or different emotion. Musical sequences that were qualified as either sad or serene in Peretz *et al.*'s (1998) studies were used and the participants were informed that all stimuli fall into one of these two emotional categories. In half of the trials, the musical pieces were in the same keys. In the other half they were in distant keys. We found that emotional responses were faster and more accurate when both pieces were in the same key. This finding suggests that interpreting the tones in a tonal key context is a prerequisite for the recognition of the emotional value of the stimuli.

## From tonal hierarchies to event hierarchies

The critical importance of the tonal hierarchy is to contribute to instilling an event hierarchy specific to a given piece. As formalized in Lerdahl and Jackendoff's Generative Theory of Tonal Music (1983), the tonal hierarchy provides the stability conditions that need to be integrated into rhythmic parameters (notably) in order to define the structural importance of each tone of the piece. The structural importance of a musical event thus depends upon its place in

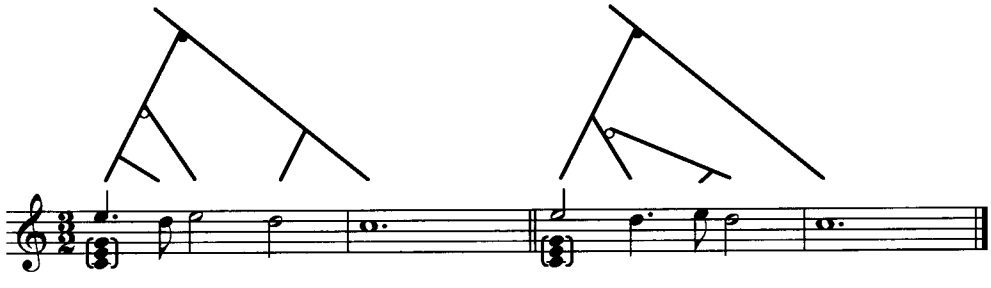


**Fig. 6.5** Influence of tonal hierarchies and rhythm on event hierarchies. Inspired by Lerdahl and Jackendoff (1983) with the help of the first author.

tonal hierarchy and the rhythmic context in which it appears. A tonic note of a given melody would not always act as a structurally important tone. Depending upon the group of tones in which it appears as well as its duration and its metrical position, a tonic note may act as an ornamental event in the event hierarchy. Modifying the rhythm of a set of tones is enough to change the event hierarchy as expressed in the following example (Figure 6.5). The *e*, which is an important tone in the C major key context, may act as an important anchoring tone in some rhythmic contexts (Figure 6.5 left) but not in others (Figure 6.5 right). This point is most crucial as it means that the tonal hierarchy is considerably modulated by numerous other musical parameters such as rhythm, loudness, and timbre. These parameters are adjusted by the composer and the performer in numerous ways, leading to a quasi infinite number of possible event hierarchies. The combination of tonal and event hierarchies creates a great unity in a quasi-infinite variety of Western musical pieces.

## References


Bharucha JJ (1984). Anchoring effects in music: the resolution of dissonance. *Cognitive Psychology*, **16**(4), 485–518.

Bharucha JJ (1987). Music cognition and perceptual facilitation: a connectionist framework. *Music Perception*, **5**(1), 1–30.

Bharucha JJ and Stoeckig K (1986). Reaction time and musical expectancy: priming of chords. *Journal of Experimental Psychology: Human Perception and Performance*, **12**(4), 403–410.

Bharucha JJ and Stoeckig K (1987). Priming of chords: spreading activation or overlapping frequency spectra? *Perception and Psychophysics*, **41**(6), 519–524.

Bigand E (1993). The influence of implicit harmony, rhythm and musical training on the abstraction of 'tension–relaxation schemas' in a tonal musical phrase. *Contemporary Music Review*, **9**, 128–139.

Bigand E (1997). Perceiving musical stability: the effect of tonal structure, rhythm and musical expertise. Journal of Experimental Psychology, *Human Perception and Performance*, **23**, 808–812.

Bigand E (2003). Travelling throught Lerdahl's tonal pitch space theory: a psychological perspective. *Musicae Scientiae*, **7**, 121–140.

Bigand E and Parncutt R (1999). Perceiving musical tension in long chord sequences. *Psychological Research*, **62**(4), 237–254.

Bigand E and Pineau M (1996). Context effects on melody recognition: a dynamic interpretation. *Current Psychology of Cognition*, **15**, 121–134.

Bigand E and Poulin-Charronnat B (2006). Are we 'experienced listeners'? A review of the musical capacities that do not depend on formal musical training. *Cognition*, **100**(1), 100–130.

Bigand E, Madurell F, Tillmann B and Pineau M (1999). Effect of global structure and temporal organization on chord processing. *Journal of Experimental Psychology: Human Perception and Performance*, **25**(1), 184–197.

Bigand E, Parncutt R and Lerdahl F (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception and Psychophysics*, **58**(1), 124–141.

Bigand E, Poulin B, Tillmann B, Madurell F and D'Adamo DA (2003). Sensory versus cognitive components in harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance*, **29**(1), 159–171.

Bigand E, Tillmann B, Poulin B, D'Adamo DA and Madurell F (2001). The effect of harmonic context on phoneme monitoring in vocal music. *Cognition*, **81**(1), B11–B20.

Bigand E, Tillmann B, Poulin-Charronnat B and Manderlier D (2005). Repetition priming: is music special? *Quarterly Journal of Experimental Psychology*, **58A**(8), 1347–1375.

D'Indy V (1900). *Traité de composition musicale*. Durand, Paris.

Dibben N (1994). The cognitive reality of hierarchic structure in tonal and atonal music. *Music Perception*, **12**, 1–25.

Dowling WJ (1986) Context effects on melody recognition: scale-step versus interval representations. *Music Perception*, **3**, 281–296.

Filipic S and Bigand E (2003). Emotion and cognition in music: which comes first ? In R Kopiez, AC Lehmann, I Wolther and C Wolf, eds, *Proceeding of the 5th Triennal ESCOM*, 231–233, 8–13 September, Germany.

Francès R (1958) *La perception de la musique*. Paris: Vrin [1984 The perception of music (Dowling trans.) Earlbaum, Hillsdale, NJ.

Krumhansl CL (1990). *Cognitive foundations of musical pitch*. Oxford University Press, New York.

Krumhansl CL (1996), A perceptual analysis of Mozart's piano sonata, K. 282: segmentation, tension and musical ideas. *Music Perception*, **13**, 401–432.

Krumhansl CL and Kessler EJ. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, **89**(4), 334–368.

Krumhansl CL, Bharucha JJ and Kessler EJ (1982). Perceived harmonic structure of chords in three related musical keys. *Journal of Experimental Psychology: Human Perception and Performance*, **8**(1), 24–36.

Lerdahl F (1988). Tonal pitch space. *Music Perception*, **5**(3), 315–349.

Lerdahl F (1989). Atonal prolongational structure. *Contemporary Music Review*, **4**, 65–87.

Lerdahl F (1991). Pitch-Space journeys in two Chopin Preludes (MR Jones and S Holleran, eds) *Cognitive bases of musical communication*, 171–91 APA, Washington, DC.

Lerdahl F (1996). Calculating tonal tension. *Music Perception*, **13**(3), 319–363.

Lerdahl F (2001). *Tonal pitch space theory*. Oxford University Press, Oxford.

Lerdahl F and Jackendoff R (1983). *A generative theory of tonal music*. The MIT Press, Cambridge, MA.

Lerdahl F and Krumhansl C (2007). Modeling tonal tension. *Music Perception*, **24**(4), 329–366.

Longuet-Higgins HC (1978). The perception of music. *Interdisciplinary Science Reviews*, **3**, 148–156.

McClelland JL and Rumelhart DE (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, **88**, 375–407.

McNamara TP (2005). *Semantic priming: perspectives from memory and word recognition*. [Psychology Press, New York].

Mélen M and Deliège I (1995). Extraction of cues or underlying harmonic structure: which guides recognition of familiar melodies? *European Journal of Cognitive Psychology*, **7**(1), 81–106.

Meyer LB (1973). *Explaining music*. University of California Press, Berkeley, CA.

Palmer C and Krumhansl CL (1987). Pitch and temporal contributions to musical phrase perception: effects of harmony, performance timing, and familiarity. *Perception and Psychophysics*, **41**(6), 505–518.

Peretz I, Gagnon L and Bouchard B (1998). Music and emotion: perceptual determinants, immediacy and isolation after brain damage. *Cognition*, **68**, 111–141.

Poulin-Charronnat B, Bigand E and Madurell F (2005a). The influence of voice leading on harmonic priming. *Music Perception*, **22**(4), 613–627.

Poulin-Charronnat B, Bigand E, Madurell F and Peereman R (2005b). Musical structure modulates semantic priming in vocal music. *Cognition*, **94**(3), B67–B78.

Rumelhart DE and McClelland JL (1982). An interactive activation model of context effects in letter perception: Part 2. The context enhancement effect and some tests and extensions of the model. *Psychological Review*, **89**, 60–94.

Schellenberg EG, Bigand E, Poulin-Charronnat B, Garnier C and Stevens C (2005). Children's implicit knowledge of harmony in Western music. *Developmental Science*, **8**(6), 551–566.

Schenker H (1935). *Der freie Satz*. Universal Edition, Vienna.

Shepard RN (1982). Geometrical approximations to the structure of musical pitch. *Psychological Review*, **89**(4), 305–333.

Smith NA and Cuddy LL (2003). Perceptions of musical dimensions in Beethoven's Waldstein sonata: an application of Tonal Pitch Space theory. *Musicae Scientiae*, **7**, 7–34.

Tekman HG and Bharucha JJ (1998). Implicit knowledge versus psychoacoustic similarity in priming of chords. *Journal of Experimental Psychology: Human Perception and Performance*, **24**(1), 252–260.

Tillmann B and Bigand E (2001). Global context effect in normal and scrambled musical sequences. *Journal of Experimental Psychology: Human Perception and Performance*, **27**(5), 1185–1196.

Tillmann B and Lebrun-Guillaud G (2006). Influence of tonal and temporal expectations on chord processing and on completion judgments of chord sequences. *Psychological Research*, **70**(5), 345–358.

Tillmann B, Bharucha JJ and Bigand E (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, **107**(4), 885–913.

Tillmann B, Bigand E and Madurell F (1998). Local versus global processing of harmonic cadences in the solution of musical puzzles. *Psychological Research*, **61**(3), 157–174.

Vega DA (2003). A perceptual experiment on harmonic tension and melodic attraction in Lerdahl's tonal pitch space. *Musicae Scientiae*, **7**, 35–55.

# CHAPTER 7

# The perception of musical timbre[1]

Stephen McAdams and Bruno L. Giordano

TIMBRE is a misleadingly simple and vague word encompassing a very complex set of auditory attributes, as well as a plethora of psychological and musical issues. It covers many parameters of perception that are not accounted for by pitch, loudness, spatial position, duration, and various environmental characteristics such as room reverberation. This leaves a wealth of possibilities that have been explored over the last 40 years or so. We now understand timbre to have two broad characteristics that contribute to the perception of music: (1) it is a multifarious set of abstract sensory attributes, some of which are continuously varying (e.g. attack sharpness, brightness, nasality, richness), others of which are discrete or categorical (e.g., the 'blatt' at the beginning of a sforzando trombone sound or the pinched offset of a harpsichord sound), and (2) it is one of the primary perceptual vehicles for the recognition, identification, and tracking over time of a sound source (singer's voice, clarinet, set of carillon bells), and thus involves the absolute categorization of a sound (Hajda *et al.* 1997; McAdams 1993; Risset and Wessel 1999). The psychological approach to timbre has also included work on the musical implications of timbre as a set of form-bearing dimensions in music (McAdams 1989).

## Timbre as a multidimensional set of auditory attributes

One of the main approaches to timbre perception attempts to characterize quantitatively the ways in which sounds are perceived to differ. Early research on the perceptual nature of timbre focused on preconceived aspects such as the relative weights of different frequencies present in a given sound, or its 'sound color' (Slawson 1985). A voice singing a constant middle C while varying the vowel being sung, or a brass player holding a given note while varying the embouchure and mouth cavity shape would both vary the shape of the sound spectrum, which represents the level of each sound partial as a function of its frequency (cf. McAdams *et al.* 2004b). Helmholtz (1885/1954) invented some rather ingenious resonating devices for controlling spectral shape to explore these aspects of timbre. However, the real advances in understanding the perceptual representation of timbre had to wait for the development of multidimensional data analysis techniques in the 1960s and signal processing techniques in the 1970s. Plomp (1970) and Wessel (1973) first applied these to timbre perception.

### Timbre space

Multidimensional scaling makes no preconceptions about the physical or perceptual structure of timbre. Listeners simply rate on a scale varying from very similar to very dissimilar all pairs from a given set of sounds. The sounds are usually equalized in terms of pitch, loudness, and duration so that only the timbre varies in order to focus listeners' attention on this attribute. The dissimilarity ratings are then fit to a distance model in which sounds with similar timbres are close together and those with dissimilar

[1] This chapter is an updated version of two previous summaries of musical timbre research (McAdams, 1999, 2003).

timbres are far apart. The graphic representation is called a 'timbre space'. The basic model is expressed in terms of continuous dimensions that are shared among the timbres, the underlying assumption being that all listeners use the same perceptual dimensions to compare the timbres. More complex models also include dimensions or features that are specific to individual timbres (called 'specificities') and different perceptual weights accorded to the dimensions and specificities by individual listeners or classes of listeners (Grey 1977; McAdams *et al.* 1995). Such techniques have been applied to synthetic sounds (Miller and Carterette 1975; Plomp 1970), resynthesized or simulated instrument sounds (Grey 1977; Krumhansl 1989; McAdams *et al.* 1995; Wessel 1979), recorded instrument sounds (Iverson and Krumhansl 1993; Lakatos 2000), and even dyads of recorded instrument sounds (Kendall and Carterette 1991).

Specificities are often found for complex acoustic and synthesized sounds. They are considered to represent the presence of a unique feature that distinguishes a sound from all others in a given context. For example, in a set of brass, woodwind, and string sounds, a harpsichord has a feature shared with no other sound: the return of the hopper which creates a slight 'thump' and quickly damps the sound at the end. This might appear as a strong specificity in the distance model (Krumhansl 1989; McAdams *et al.* 1995).

The models integrate individual and class differences as weighting factors on the different dimensions and the set of specificities. Some listeners pay more attention to spectral properties and ignore temporal aspects, whereas others have the inverse pattern. Such variability may reflect either differences in sensory processing or in listening and rating strategies. Interestingly, no study to date has demonstrated that such individual differences have anything to do with musical experience or training (McAdams *et al.* 1995). It may be that because timbre perception is so closely allied with the ability to recognize sound sources in everyday life, everybody is an expert to some degree.

## Acoustic correlates of timbral dimensions

In many studies, independent acoustic correlates have been determined for the continuous dimensions by correlating the position along the perceptual dimension with a unidimensional acoustic parameter extracted from the sounds (e.g. Grey and Gordon 1978; Krimphoff *et al.* 1994). The most ubiquitous correlates derived from musical instrument sounds include spectral centroid (representing the relative weights of high and low frequencies and corresponding to timbral brightness: an oboe has a higher spectral centroid than a French horn), the logarithm of the attack time (distinguishing continuant instruments that are blown or bowed from impulsive instruments that are struck or plucked), spectral flux (the degree of evolution of the spectral shape over a tone's duration which is high for brass and lower for single reeds), and spectral irregularity (the degree of jaggedness of the spectral shape, which is high for clarinet and vibraphone and low for trumpet). A confirmatory study employing dissimilarity ratings on purely synthetic sounds in which the exact nature of the stimulus dimensions could be controlled was performed by Caclin *et al.* (2005). These authors confirmed the perception of stimulus dimensions related to spectral centroid, attack time and spectral irregularity, but did not confirm spectral flux.

The combination of a quantitative model of perceptual relations among timbres and the psychophysical explanation of the parameters of the model is an important step in gaining predictive control of timbre in several domains such as sound analysis and synthesis and intelligent search in sound databases (Peeters *et al.* 2000). Such representations are only useful to the extent that they are: (a) generalizable beyond the set of sounds actually studied, (b) robust with respect to changes in musical context, and (c) generalizable to other kinds of listening tasks than those used to construct the model. To the degree that a representation has these properties, it may be considered as an accurate account of musical timbre, characterized by an important feature of a scientific model, the ability to predict new empirical phenomena.

Timbre space models have been useful in predicting listeners' perception in situations beyond those specifically measured in the experiments, which suggests that they do in fact capture important aspects of timbre representation. Consistent with the predictions of a timbre

model, Grey and Gordon (1978) found that by exchanging the spectral envelopes on pairs of sounds that differed primarily along the spectral dimension, these sounds switched positions in the space. Timbre space has also been useful in predicting the perception of intervals between timbres, as well as stream segregation based on timbre-related acoustic cues (see below).

### Effects of pitch change on timbre relations

Marozeau and colleagues (2003) have shown that timbre spaces for recorded musical instrument tones are similar at different pitches (B3, C#4, B♭4). Listeners are also able to ignore pitch differences within an octave when asked to compare only the timbres of the tones. When the pitch variation is greater than an octave, interactions between the two attributes occur. Marozeau and de Cheveigné (2007) varied the brightness of a set of synthesized sounds, while also varying the pitch over a range of 18 semitones. They found that differences in pitch affected timbre relations in two ways:

1 Pitch shows up in the timbre space representation as a dimension orthogonal to the timbre dimensions (indicating simply that listeners were no longer ignoring the pitch difference), and
2 Pitch differences systematically affect the timbre dimension related to spectral centroid.

These results suggest a close relation between timbral brightness and pitch height. This link would be consistent with underlying neural representations that share common attributes, such as a tonotopic organization.

## Timbre as a vehicle for source identity

The second approach to timbre concerns its role in the recognition of the identity of a musical instrument or, in general, of a sound-generating event. One reasonable hypothesis is that the sensory dimensions that compose timbre serve as indicators used in the categorization, recognition, and identification of sound events and sound sources (McAdams 1993; Handel 1995).

Research on musical instrument identification is relevant to this issue. Saldanha and Corso (1964) studied identification of isolated musical instrument sounds from the Western orchestra played with and without vibrato. They were interested in the relative importance of onset and offset transients, spectral envelope of the sustain portion of the sound, and vibrato. Identification of isolated sounds is surprisingly poor for some instruments. When attacks and decays were excised, identification decreased markedly for some instruments, particularly for the attack portion in sounds without vibrato. However, when vibrato was present, the effect of cutting the attack was less, identification being better. These results suggest that important information for instrument identification is present in the attack portion, but that in the absence of this information, additional information is still available in the sustain portion (although it is more important for some instruments than others), particularly when vibrato is present. The vibrato may increase our ability to extract information relative to the resonance structure of the instrument (McAdams and Rodet 1988).

Giordano (2005) analysed previously published data on the identification and dissimilarity ratings of musical tones. The goal of this study was to ascertain the extent to which tones generated with large differences in the mechanisms for sound production were recovered in the perceptual data. Across all identification studies, listeners frequently confused tones generated by musical instruments with a similar physical structure (e.g., clarinets and saxophones, both single-reed instruments), and seldom confused tones generated by very different physical systems (e.g., the trumpet, a lip-reed instrument, and the bassoon, a double-reed instrument). Consistently, the vast majority of previously published timbre spaces revealed that tones generated with similar resonating structures (e.g., string instruments vs wind instruments) or with similar excitation mechanisms (e.g., impulsive excitation as in piano tones vs sustained excitation as in flute tones) occupied the same region in the space. These results suggest that listeners can reliably identify large differences in the mechanisms of tone production, focusing on the timbre attributes used to evaluate the dissimilarity of musical sounds.

Several investigations on the perception of everyday sounds extend the concept of timbre

beyond the musical context (see McAdams 1993; Handel 1995; Lutfi 2008 for reviews). Among them, studies on impact sounds provide information on the timbre attributes useful to the perception of the properties of percussion instruments: bar geometry (Lakatos *et al.* 1997), bar material (McAdams *et al.* 2004a), plate material (Giordano and McAdams 2006), and mallet hardness (Freed 1990). The timbral factors relevant to perceptual judgements vary with the task at hand. Spectral factors are primary for the perception of the geometry (Lakatos *et al.* 1997). Spectrotemporal factors (e.g., the rate of temporal change of the spectral centroid and of loudness) dominate the perception of the material of struck objects (McAdams *et al.* 2004a, Giordano and McAdams 2006) and of mallets (Freed 1990).

The perception of an instrument's identity in spite of variations in pitch may be related to timbral invariance, those aspects of timbre that remain constant with change in pitch and loudness. Handel and Erickson (2001) found that musically untrained listeners are able to recognize two sounds produced at different pitches as coming from the same instrument or voice only within a pitch range of about an octave. Steele and Williams (2006) found that musically trained listeners could perform this task at about 80 per cent correct even with pitch differences on the order of 2.5 octaves. These results suggest that there are limits to timbral invariance across pitch, but that they depend on musical training.

Its role in source identification and categorization is perhaps the more neglected aspect of timbre, and brings with it advantages and disadvantages for the use of timbre as a form-bearing dimension in music (McAdams 1989). One of the advantages is that categorization and identification of a sound source may bring into play perceptual knowledge (acquired by listeners implicitly through experience in the everyday world and in musical situations) that helps them track a given voice or instrument in a complex musical texture. Listeners do this easily and some research has shown that timbral factors may make an important contribution in voice tracking (Culling and Darwin 1993; Gregory 1994), which is particularly important in polyphonic settings.

The disadvantages may arise in situations in which the composer seeks to create melodies across instrumental timbres, e.g., the *Klangfarbenmelodien* of Schoenberg (1911/1978). Our predisposition to identify the sound source and follow it through time would impede a more relative perception in which the timbral differences were perceived as a movement through timbre space rather than as a simple change of sound source. For cases in which such timbral compositions work, the composers have often taken special precautions to create a musical situation that draws the listener more into a relative than into an absolute mode of perceiving.

## Timbral intervals

If timbral interval perception can be demonstrated, it opens the door to musical operations on timbre sequences that are commonly used on pitch sequences (Slawson 1985). Another interest of this exploration is that it extends the use of the timbre space as a perceptual model beyond the dissimilarity paradigm.

Ehresman and Wessel (1978; Wessel 1979) took a first step forward in this direction, developing a task in which listeners were asked to make judgements on the similarity of intervals formed between pairs of timbres. The basic idea was that timbre intervals may have properties similar to pitch intervals; that is, a pitch interval is a relation along a well-ordered dimension that retains a degree of invariance under certain kinds of transformation, such as translation along the dimension, or what musicians call 'transposition'. What does transposition mean in a multidimensional space? A timbre interval can be considered as a vector in space connecting two timbres. It has a specific length (the distance between the timbres) and a specific orientation. Together these two properties define the amount of change along each dimension of the space that is needed to move from one timbre to another. If we assume these dimensions to be continuous and linear from a perceptual point of view, then pairs of timbres characterized by the same vector relation should have the same relative perceptual relation and thus embody the same timbre interval. Transposition thus consists of translating the vector anywhere else in the space as long as its length and orientation are preserved.

Ehresman and Wessel tested this hypothesis using a task in which listeners had to compare

two timbral intervals (e.g. A–B vs C–D) and rank various timbre D's according to how well they fulfilled the analogy: timbre A is to timbre B as timbre C is to timbre D. They essentially found that the closer timbre D was to the ideal point defined by the vector model in timbre space (i.e. the ideal C–D vector was a simple translation of the A–B vector), the higher the ranking.

McAdams and Cunibile (1992) subsequently tested the vector model using the 3D space from Krumhansl (1989) (ignoring the specificities). Five sets of timbres at different places in timbre space were chosen for each comparison to test for the generality of the results. Both electroacoustic composers and non-musicians were tested to see if musical training and experience had any effect. All listeners found the task rather difficult to do, which is not surprising given that even professional composers have had almost no experience with music that uses timbre intervals in a systematic way. The main result is encouraging in that globally the data support the vector model, although this support was much stronger for composers than for non-musicians. However, when one examines in detail the five different versions of each comparison type, it is clear that not all timbre comparisons go in the direction of the model predictions.

One confounding factor is that the specificities on some timbres in this set were ignored. These, quite to the contrary, would necessarily distort the vectors that were used to choose the timbres, because the specificities are like an additional dimension for each timbre. As such, certain timbre intervals correspond well to what is predicted because specificities are absent or low in value, whereas others would be seriously distorted and thus not perceived as similar to other intervals due to moderate or high specificity values. What this line of reasoning suggests is that the use of timbre intervals as an integral part of a musical discourse runs the risk of being very difficult to achieve with very complex and idiosyncratic sound sources, because they will in all probability have specificities of some kind or another. The use of timbre intervals may, in the long run, be limited to synthesized sounds or blended sounds created through the combination of several instruments.

## Timbre and musical grouping

An important way in which timbre can contribute to the organization of musical structure is related to the fact that listeners tend to connect perceptually sound events that arise from the same sound source. In general, a given source will produce sounds that are relatively similar in pitch, loudness, timbre and spatial position from one event to the next (cf. Bregman 1990; McAdams and Bregman 1979 for reviews). The perceptual connection of successive sound events into a coherent 'message' through time is referred to as auditory stream integration, and the separation of events into distinct messages is called auditory stream segregation (Bregman and Campbell 1971). One guiding principle that seems to operate in the formation of auditory streams is the following: successive events that are relatively similar in their spectrotemporal properties (i.e. in their timbres) may have arisen from the same source and should be grouped together; individual sources do not tend to change their acoustic properties suddenly and repeatedly from one event to the next. Early demonstrations of auditory streaming on the basis of timbre (Wessel 1979) suggest a link between the timbre–space representation and the tendency for auditory streaming on the basis of the spectral differences that were created (McAdams and Bregman 1979). Early researchers were convinced that it was primarily the spectral aspects of timbre (such as spectral centroid) that were responsible for auditory streaming and that temporal aspects (such as attack time) had little effect (Hartmann and Johnson 1991).

Recently the picture has changed significantly and several studies indicate an important role for both spectral and temporal attributes of timbre in auditory stream segregation (Moore and Gockel 2002). Iverson (1995) used sequences alternating between two recorded instrument tones with the same pitch and loudness and asked listeners to judge the degree of segregation. Multidimensional scaling of the segregation judgments treated as a measure of dissimilarity was performed to determine which acoustic attributes contributed to the impression of auditory stream segregation. A comparison with previous timbre–space work using the

same sounds (Iverson and Krumhansl 1993) showed that both static acoustic cues (such as spectral centroid) and dynamic acoustic cues (such as attack time and spectral flux) were implicated in segregation. Other results consistent with this study have also been reported (Bey and McAdams 2003; Singh and Bregman 1997).

All of these results are important for auditory stream segregation theory on the one hand, because they show that several of a source's acoustic properties are taken into account when forming auditory streams. On the other hand, they are important for music-making (whether it be with computer or acoustic instruments), because they show that many aspects of timbre strongly affect the basic organization of the musical surface into streams. Different orchestrations of a given pitch sequence can completely change what is heard as melody and rhythm, as has been demonstrated by Wessel (1979). Timbre is also an important component in the perception of musical groupings, whether they are at the level of sequences of notes distinguished by changes in timbre (Deliège 1987) or of larger-scale musical sections delimited by marked changes in orchestration and timbral texture (Deliège 1989).

# Timbre as a structuring force in music perception

Timbre perception is at the heart of orchestration, a realm of musical practice that has received relatively little experimental study. Instrumental combinations can give rise to new timbres if the sounds are perceived as blended, and timbre can play a role in creating and releasing musical tension.

## Timbral blend

The creation of new timbres through orchestration necessarily depends on the degree to which the constituent sound sources fuse together or blend to create the newly emerged sound (Brant 1971; Erickson 1975). Sandell (1995) has proposed that there are three classes of perceptual goals in combining instruments:

1 timbral heterogeneity in which one seeks to keep the instruments perceptually distinct,
2 timbral augmentation in which one instrument embellishes another one that perceptually dominates the combination, and
3 timbral emergence in which a new sound results that is identified as none of its constituents.

Blend appears to depend on a number of acoustic factors such as onset synchrony of the constituent sounds and others that are more directly related to timbre, such as the similarity of the attacks, the difference in the spectral centroids, and the overall centroid of the combination.

## Role of timbre in building and release of musical tension

Timbre can also contribute to larger-scale musical form and in particular to the sense of movement between tension and relaxation. This movement has been considered by many music theorists as one of the primary bases for the perception of larger-scale form in music. It has traditionally been tied to harmony in Western music and plays an important role in Lerdahl and Jackendoff's (1983)*A generative theory of tonal music*. Experimental work on the role of harmony in the perception of musical tension and relaxation (or inversely, in the sense of tension that accompanies a moment at which the music must continue and the sense of relaxation that accompanies the completion of the musical phrase) has suggested that auditory roughness is an important component of perceived tension (Bigand *et al.* 1996). Roughness is an elementary timbral attribute based on the sensation of rapid fluctuations in the amplitude envelope. It can be generated by proximal frequency components that beat with one another. Dissonant intervals tend to have more such beating than consonant intervals. As such, a fairly direct relation between sensory dissonance and roughness has been demonstrated (cf. Parncutt 1989; Plomp 1976 for reviews).

As a first step toward understanding how this operates in music, Paraskeva and McAdams (1997) measured the inflection of musical tension and relaxation due to timbral change. Listeners were asked to make judgments on a 7-point scale concerning the perceived degree of completion of the music at several points at

which the music stopped. What results is a completion profile, which can be used to infer musical tension by equating completion with release and lack of completion with tension. Two pieces were tested: a fragment from the Ricercar from the *Musical Offering* for six voices by Bach (tonal) and the first movement of the *Six Pieces for Orchestra* by Webern (non-tonal). Each piece was played in an orchestral version (the Webern instrumentation was used for the Bach), and a direct transcription of this orchestral version for piano on a digital sampler. There were significant differences between the piano and orchestral versions, indicating a significant effect of timbre change on perceived musical tension. However, when they were significantly different, the orchestral version was always more relaxed than the piano version.

The hypothesis advanced by Paraskeva and McAdams (1997) for this effect was that the higher relaxation of the orchestral version might have been due to processes involved in auditory stream formation and the dependence of perceived roughness on the results of such processes (Wright and Bregman 1987). Roughness, or any other auditory attribute of a single sound event, is computed after auditory organization processes have grouped the bits of acoustic information together. Piano sounds have a rather sharp attack. If several notes occur at the same time in the score and are played with a piano sound, they will be quite synchronous. Because they all start at the same time and have similar amplitude envelopes, they will tend to be fused together and the computed roughness will result from the interactions of all the frequency components of all the notes.

The situation may be quite different for the orchestral version for two reasons. The first is that the same timing is used for piano and orchestra versions. In the latter, many instruments are used that have slow attacks whereas others have faster attacks. There could then be a great deal of asynchrony between the instruments in terms of perceived attack time (Gordon 1987). In addition, because the timbres of these instruments are often quite different, several different voices with different timbres arrive momentarily at a given vertical sonority, but the verticality is not perceived because the listener would more likely continue to track individual instruments horizontally. So the attack asynchrony and the decomposition of verticalities into horizontalities would concur to reduce the degree of perceptual fusion. Reduced fusion would mean greater segregation. Thus the roughness in the orchestral version would be computed on each individually grouped auditory event rather than on the whole sound mass. These individual roughnesses in the orchestral version would most likely be much less than that of the piano version. So once again, timbral composition can have a very tight interaction with auditory stream formation processes.

## Conclusion

Musical timbre is a combination of continuous perceptual dimensions and discrete features to which listeners are differentially sensitive. The continuous dimensions often have quantifiable acoustic correlates. The timbre–space representation is a powerful psychological model that allows predictions to be made about timbre perception in situations beyond those used to derive the model in the first place. Timbre intervals, for example, can be conceived as vectors within the space of common dimensions. Timbre space also makes at least qualitative predictions about the magnitude of timbre differences that will provoke auditory stream segregation.

Timbre can play a role in larger-scale movements of tension and relaxation and thus contribute to the expression inherent in musical form. Under conditions of high blend among instruments composing a vertical sonority, timbral roughness is a major component of musical tension. However, it strongly depends on the way auditory grouping processes have parsed the incoming acoustic information into events and streams.

## References

Bey C and McAdams S (2003). Post-recognition of interleaved melodies as an indirect measure of auditory stream formation. *Journal of Experimental Psychology: Human Perception and Performance*, **29**, 267–279.

Bigand E, Parncutt R and Lerdahl F (1996). Perception of musical tension in short chord sequences: the influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception and Psychophysics*, **58**, 125–141.

Brant H (1971). Orchestration. In J Vinton, ed., *Dictionary of contemporary music*, 538–546 EP Dutton, New York.

Bregman AS (1990). *Auditory scene analysis: the perceptual organization of sound*. MIT Press, Cambridge, MA.

Bregman AS and Campbell J (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, **89**, 244–249.

Caclin A, McAdams S, Smith BK and Winsberg S (2005). Acoustic correlates of timbre space dimensions: a confirmatory study using synthetic tones. *Journal of the Acoustical Society of America*, **118**, 471–482.

Culling JF and Darwin CJ (1993). The role of timbre in the segregation of simultaneous voices with intersecting Fo contours. *Perception and Psychophysics*, **34**, 303–309.

Deliège I (1987). Grouping conditions in listening to music: an approach to Lerdahl and Jackendoff's grouping preference rules. *Music Perception*, **4**, 325–360.

Deliège I, (1989). A perceptual approach to contemporary musical forms. *Contemporary Music Review*, **4**, 213–230.

Ehresman D and Wessel DL (1978). Perception of timbral analogies. *Rapports de l'IRCAM*, vol. 13 IRCAM, Paris.

Erickson R (1975). *Sound structure in music*. University of California Press, Berkeley, CA.

Freed DJ (1990). Auditory correlates of perceived mallet hardness for a set of recorded percussive events. *Journal of the Acoustical Society of America*, **87**, 1236–1249.

Giordano BL (2005). Sound source perception in impact sounds. Unpublished doctoral dissertation. University of Padova, Italy.

Giordano BL and McAdams S (2006). Material identification of real impact sounds: effects of size variation in steel, glass, wood and plexiglass plates. *Journal of the Acoustical Society of America*, **119**, 1171–1181.

Gordon JW (1987). The perceptual attack time of musical tones. *Journal of the Acoustical Society of America*, **82**, 88–105.

Gregory AH (1994). Timbre and auditory streaming. *Music Perception*, **12**, 161–174.

Grey JM (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America*, **61**, 1270–1277.

Grey JM and Gordon JW (1978). Perceptual effects of spectral modifications on musical timbres. *Journal of the Acoustical Society of America*, **63**, 1493–1500.

Hajda JM, Kendall RA, Carterette EC and Harshberger ML (1997). Methodological issues in timbre research. In I Deliège and J Sloboda, eds, *Perception and cognition of music*, 253–306. Psychology Press, Hove.

Handel S (1995). Timbre perception and auditory object identification. In B Moore, ed., *Hearing*, 425–462. Academic Press, San Diego, CA.

Handel S and Erickson M (2001). A rule of thumb: the bandwidth for timbre invariance is one octave. *Music Perception*, **19**, 121–126.

Hartmann WM and Johnson D (1991). Stream segregation and peripheral channeling. *Music Perception*, **9**, 155–184.

Helmholtz HLF von (1885/1954). *On the sensations of tone as a physiological basis for the theory of music*. New York, from 1877 trans. by AJ Ellis of 4th German edn, republished 1954 by Dover, New York.

Iverson P (1995). Auditory stream segregation by musical timbre: effects of static and dynamic acoustic attributes. *Journal of Experimental Psychology: Human Perception and Performance*, **21**, 751–763.

Iverson P and Krumhansl CL (1993). Isolating the dynamic attributes of musical timbre. *Journal of the Acoustical Society of America*, **94**, 2595–2603.

Kendall RA and Carterette EC (1991). Perceptual scaling of simultaneous wind instrument timbres. *Music Perception*, **8**, 369–404.

Krimphoff J, McAdams S and Winsberg S (1994). Caractérisation du timbre des sons complexes. II: Analyses acoustiques et quantification psychophysique [Characterization of the timbre of complex sounds. II: Acoustic analyses and psychophysical quantification]. *Journal de Physique*, **4**(C5), 625–628.

Krumhansl CL (1989). Why is musical timbre so hard to understand? In S Nielzén and O Olsson, eds, *Structure and perception of electroacoustic sound and music*, 43–53. Excerpta Medica, Amsterdam.

Lakatos S (2000). A common perceptual space for harmonic and percussive timbres. *Perception and Psychophysics*, **62**, 1426–1439.

Lakatos S, McAdams S and Caussé R (1997). The representation of auditory source characteristics: simple geometric form. *Perception and Psychophysics*, **59**, 1180–1190.

Lerdahl F and Jackendoff R (1983). *The generative theory of tonal music*. MIT Press, Cambridge, MA.

Lutfi R (2008). Human sound source identification. In W Yost, A Popper and R Fay, eds, *Auditory perception of sound sources*, 13–42. Springer, New York.

Marozeau J, de Cheveigné A, McAdams S and Winsberg S (2003). The dependency of timbre on fundamental frequency. *Journal of the Acoustical Society of America*, **114**, 2946–2957.

Marozeau J and de Cheveigné A (2007). The effect of fundamental frequency on the brightness dimension of timbre. *Journal of the Acoustical Society of America*, **121**, 383–387.

McAdams S (1989). Psychological constraints on form-bearing dimensions in music. *Contemporary Music Review*, **4**(1), 181–198.

McAdams S (1993). Recognition of sound sources and events. In S McAdams and E Bigand, eds, *Thinking in sound: the cognitive psychology of human audition*, 146–198. Oxford University Press, Oxford.

McAdams S (1999). Perspectives on the contribution of timbre to musical structure. *Computer Music Journal*, **23**(2), 96–113.

McAdams S (2003). Perception of musical timbre. *Bulletin of Psychology and the Arts*, **4**, 39–42.

McAdams S and Bregman AS (1979). Hearing musical streams. *Computer Music Journal*, **3**(4), 26–43.

McAdams S and Cunibile JC (1992). Perception of timbral analogies. *Philosophical Transactions of the Royal Society, London*, series B, **336**, 383–389.

McAdams S and Rodet X (1988) The role of FM-induced AM in dynamic spectral profile analysis. In H Duifhuis, JW Horst, HP Wit, eds, *Basic issues in hearing*, 359–369. Academic Press, London.

McAdams S, Chaigne A and Roussarie V (2004a). The psychomechanics of simulated sound sources: material properties of impacted bars. *Journal of the Acoustical Society of America*, **115**, 1306–1320.

McAdams S, Depalle P and Clarke E (2004b). Analyzing musical sound. In E Clarke and N Cook, eds, *Empirical musicology: aims, methods, prospects*, 157–196. Oxford University Press, New York.

McAdams S, Winsberg S, Donnadieu S, De Soete G and Krimphoff J (1995). Perceptual scaling of synthesized musical timbres: common dimensions, specificities, and latent subject classes. *Psychological Research*, **58**, 177–192.

Miller JR and Carterette EC (1975). Perceptual space for musical structures. *Journal of the Acoustical Society of America*, **58**, 711–720.

Moore BCJ and Gockel H (2002). Factors influencing sequential stream segregation. *Acustica united with Acta Acustica*, **88**, 320–332.

Paraskeva S and McAdams S (1997). Influence of timbre, presence/absence of tonal hierarchy and musical training on the perception of tension/relaxation schemas of musical phrases. *Proceedings of the 1997 International Computer Music Conference, Thessaloniki*, 438–441. International Computer Music Association, San Francisco, CA.

Parncutt R (1989). *Harmony: a psychoacoustical approach*. Springer-Verlag, Berlin.

Peeters G, McAdams S and Herrera P (2000) Instrument sound description in the context of MPEG-7. *Proceedings of the 2000 International Computer Music Conference, Berlin*, 166–169. International Computer Music Association, San Francisco, CA.

Plomp R (1970). Timbre as a multidimensional attribute of complex tones. In R Plomp and GF Smoorenburg, eds, *Frequency analysis and periodicity detection in hearing*, 397–414. Sijthoff, Leiden.

Plomp R (1976). *Aspects of tone sensation: a psychophysical study*. Academic Press, London.

Risset J-C and Wessel DL (1999). Exploration of timbre by analysis and synthesis. In D Deutsch, ed., *The psychology of music*, 2nd edn, 113–168. Academic Press, San Diego, CA.

Saldanha EL and Corso JF (1964). Timbre cues and the identification of musical instruments. *Journal of the Acoustical Society of America*, **36**, 2021–2126.

Sandell GJ (1995). Roles for spectral centroid and other factors in determining 'blended' instrument pairings in orchestration. *Music Perception*, **13**, 209–246.

Schoenberg A (1911/1978). *Theory of harmony*. Original German publication, 1911 edn. University of California Press, Berkeley, CA.

Singh PG and Bregman AS (1997). The influence of different timbre attributes on the perceptual segregation of complex-tone sequences. *Journal of the Acoustical Society of America*, **120**, 1943–1952.

Slawson W (1985). *Sound color*. University of California Press, Berkeley, CA.

Steele K and Williams A (2006). Is the bandwidth for timbre invariance only one octave? *Music Perception*, **23**, 215–220.

Wessel DL (1973). Psychoacoustics and music: a report from Michigan State University. *PACE: Bulletin of the Computer Arts Society*, **30**, 1–2.

Wessel DL (1979). Timbre space as a musical control structure. *Computer Music Journal*, **3**(2), 45–52.

Wright JK and Bregman AS (1987). Auditory stream segregation and the control of dissonance in polyphonic music. *Contemporary Music Review*, **2**(1), 63–92.

# CHAPTER 8

# Musical time

Mari Riess Jones

THIS chapter presents perspectives on perception of metre and rhythm, with a focus on dynamic attending theory (DAT). Three major sections address, respectively, metre perception, rhythm perception, and the role of time markers.

## Metre perception

Intuitively, metre perception refers to a listener's sensitivity to musical timing regularities, evident when 'keeping time' by tapping in synchrony with musical tones. Sometimes tones are clearly accented and contribute to *metric hierarchies*, as suggested in Figure 8.1(a). Here, an idealized hierarchy features accents that are increasingly stronger (i.e. more salient; thicker bars) on non-adjacent tones separated by increasingly longer time spans. Metric hierarchies comprise overlapping (i.e., *embedded*) time spans; they can be distinguished from time spans between successive tones (i.e., *serial spans*). Whereas metric spans reflect temporal embeddings (Figure 8.1(a)), serial spans invite rhythmic grouping (Figure 8.1(b)). Conventional wisdom holds that metre and rhythm are distinct time structures, engaging different perceptual processes. In this chapter I question this dichotomy.

Let's begin with metre perception. First, we should differentiate metric information in a visual score from that which meets a listener's ears. A typical (Western) score spells out two hierarchical time spans: a measure and its subdivisions. Thus, a notated 4/4 metric signature defines the measure as a time level evenly subdivided into

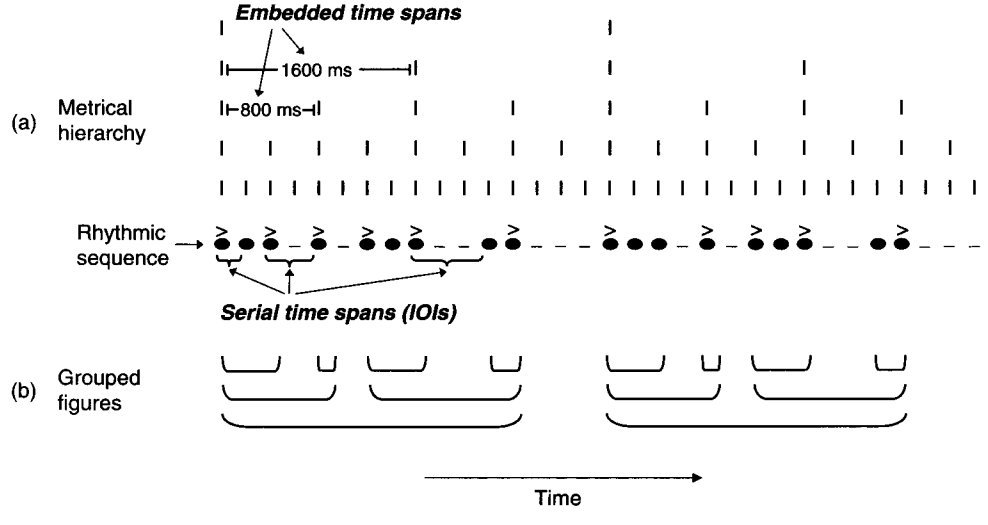



**Fig. 8.1** Metric hierarchy, with embedding time spans (a); and rhythmic groups with serial (IOIs) time spans (b).

four quarter notes (a 4:1 embedded time ratio, Figure 8.1(a)), whereas a triple (3/4) metre implies a subdivision of three (a 3:1 ratio).

Yet, given an unfamiliar melody, a listener often does not 'know' its designated metre; people don't have scored bar lines in their heads. To understand metre perception in sound patterns, we must consider that listeners may infer several metric time levels. From this, challenging issues arise. For instance, listeners may selectively internalize aspects of a metric hierarchy by engaging a series of internal *beats* at one or more time levels. Stronger beats might be felt at points in a sound pattern where multiple objective accents align. Further, if an internal beat pattern persists, then beats are also felt at times where no objective accents exist, i.e., 'subjective accents'. Also, different listeners may focus upon different internal beat patterns as a referent level (tactus). In short, to understand metre perception, we must discover factors in sound patterns and in listeners that bias people to hear a pattern's time structure in a particular way.

Contemporary research on musical time began with classic work of Fraisse (1963), who focused on rhythmic groupings. He proposed that listeners rely on favoured *serial* time ratios to segment groups. Preferred ratios approximated 1:1 for within-group time spans and 2:1 for between-group serial time ratios. However, metric issues entered the picture when Povel (1981) showed that metrical context modulates listeners preferences for 2:1 serial time ratios. Next, I describe theory and research surrounding two current psychological theories of metre perception, an encoding theory and a dynamic attending theory (DAT).

## An encoding theory

Povel and Essens (1985) proposed that metre perception is governed by an internal clock. Metrical sequences, containing regular accents, readily induce a 'good' clock and lead to efficient encoding of serial time intervals, whereas non-metrical sequences, with irregular accents, do not. Examples of metrical and non-metrical monotone sequences, involving various inter-onset time intervals (IOIs), appear in Figures 8.1 and 8.2. Grouping accents (>) putatively occur on:

1 isolated tones;
2 the second tone of a two-tone group; and
3 initial and final tones in groups of three or more tones (Povel and Okkerman 1981).

This clock model has two stages. The first stage entails matching accents with internal clock ticks (i.e., beats). The best clock maximizes accent isochrony via minimizing mismatches (clock violations). Strong metric patterns (patterns 1 and 2, Figure 8.2), have accents (>) that coincide entirely with ticks (●) of the best clock,

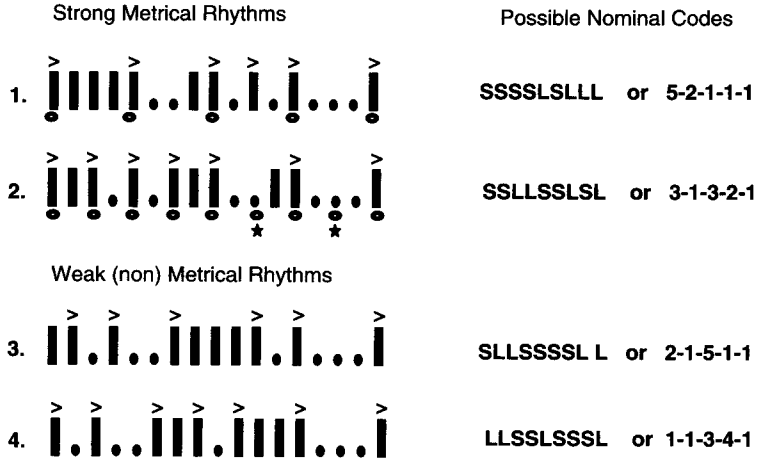



**Fig. 8.2** Four patterns from Povel and Essens (1985). Grouping accents are >; clock ticks are ●; negative clock evidence, s, is *.

namely the clock with fewest violations, i.e., least counter-evidence, C, where:

$$C = W \cdot s + u \quad (1)$$

The W (W > 1.0) weights the number of clock ticks falling on a silence, $s$ (*stars* in Figure 8.2) and $u$ is the number of ticks on unaccented elements. Thus, for pattern 1, $s = 0$ and $u = 0$ whereas for pattern 2, $s = 2$ and $u = 0$. For both patterns, C is low reflecting strong metrical patterns.

The second stage features symbolic memory codes which are more economical for metric than for non-metric patterns. For non-metric patterns, codes reflect grouping properties (cf. Figure 8.2). Thus, although the non-metric pattern 3 lacks a good clock fit, its five groups produce nominal codes of Short or Long IOIs, or symbols of group sizes: 2–1–5–1–1. These codes reflect nominal segmentations, not interval time relationships. Others propose different coding strategies (e.g. Lerdahl and Jackendoff 1983; Temperley 2001).

### Evidence for metric encoding

Povel and Essens required people to reproduce many metric and non-metric sequences. Consistent with clock model predictions, best temporal accuracy occurred with metrical sequences. Since 1985, this model has stimulated much research on metric encoding. However, conflicting reports surround its implications that highly metrical patterns elicit better temporal acuity (Handel 1992; Ross and Houtsma 1994; Handel 1998). Hebert and Cuddy (2002) found that both a metric frame (metrical sequences) and the presence of a rhythmic figure (non-metrical sequences) benefit time-change detection.

Much research on synchronized tapping appeals to motor control theory (see Repp 2005 for review). However, few (e.g. Palmer and Pfordresher 2003) consider metre perception. Some empirical examinations of synchronized and reproduction tapping do directly address clock model assumptions (e.g. McAuley and Semple 1999; Patel *et al.* 2005). McAuley and Semple questioned whether clock induction depends only on negative, i.e., counter-evidence (i.e., ticks at unaccented times). They formulated alternative versions of Equation 1, creating a positive evidence model (P), weighting clock ticks coinciding with accents; and a hybrid model (H), weighting both positive and negative evidence. Musicians and non-musicians tapped to metrical sequences, indexed by C, P, and H. The Hybrid model best predicted musicians' behaviour, whereas Povel's C index best predicted non-musicians' performance. Thus, non-musicians appear to be affected by clock violations, consistent with Povel and Essens model, whereas musicians draw strength from confirmations as well as clock violations.

Patel *et al.* (2005) found that tapping to isochronous patterns at a fixed beat period of 800 ms was less variable with isochronous subdivisions of this period than with non-isochronous (rhythmic) ones. Although consistent with clock predictions, other findings involving subdivisions by 2 or 4 were more difficult for the clock model. These data converge with others (Essens 1986) to suggest that listeners rely on relationships among multiple time levels. Indeed, even very young children appear to use multiple time levels to differentiate duple from triple metre categories (Phillips-Silver and Trainor 2005; Trehub and Hannon 2006; Bergeson and Trehub 2006), although reports of children's preference for duple metre (2:1 ratios) (Drake 1993) were not confirmed.

## Dynamic attending theory and the metric binding hypothesis

Dynamic attending theory (DAT) addresses 'in-the-moment' expectancies in listening. Its relevance to metre perception is discussed in this section, where I propose a new hypothesis: the *metric binding hypothesis*.

Entrainment is a biological process that realizes adaptive synchrony of *internal* attending oscillations with an *external* event. Different event timescales correspond to marked (i.e., accented) metric levels (Figures 8.1, 8.3). Time spans within a metric level can elicit a corresponding neural oscillation, which has a persisting internal periodicity (Pi), manifest as a *temporal expectancy*. It 'tunes into' recurrent time spans at a given level by adjusting its phase in response to temporal *expectancy violations* ($\phi$) at that level.

Various DAT models share four assumptions (Large 1994; McAuley 1995; Large and Kolen 1995;

Large and Jones 1999). First, neural oscillations are *self-sustaining;* they persist over time, extrapolating the induced beat. Second, an oscillator's intrinsic period exhibits *stability;* a perturbation from an ill-timed tone only briefly disrupts a stable oscillation, which returns to its intrinsic period. Third, entraining oscillations exhibit *adaptivity*, the flip side of stability; an oscillator responds appropriately to event-generated expectancy violations by adjusting phase and period. Fourth, *multiple related oscillations* are activated by multiple time levels within metric and rhythmic events.

The metric binding hypothesis expands these assumptions. It adds learning principles to address training and enculturation that contribute to listeners' familiarity with metric categories. It holds: *whenever two or more neural oscillations are simultaneously active, over time their internal entrainments lead to binding and formation of a metric cluster.* A metric cluster comprises sets of co-occurring oscillations with interrelationships that persist due to acquired internal bindings. Entrainments among internal oscillations promote binding, which strengthens as a function of:

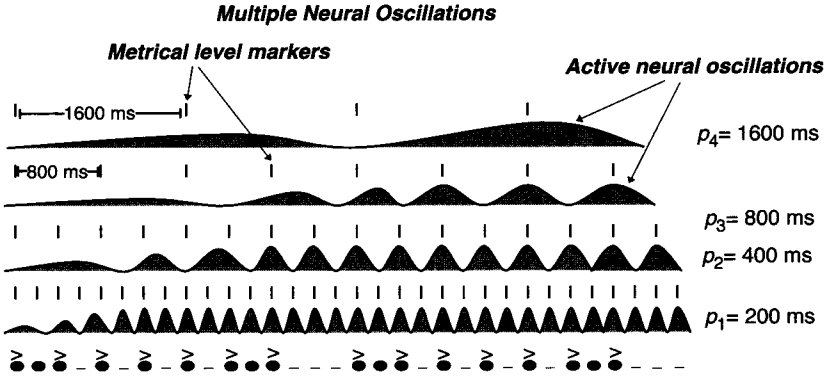
1 Duration of co-occurring oscillatory activity.

2 Phase coincidences, and,

3 Resonance (i.e., relatedness) among oscillator periods.

Figure 8.3 illustrates how metric clusters form. As a rhythmic pattern unfolds, it successively activates neural oscillations with periods (Pi) matching each of the lower-level IOIs. Initially recurrent IOIs, in this case, elicit an oscillation with a period of 200 ms, which is strongly coupled to these IOIs. However, this rhythm subsequently awakens oscillations with periods of 400 and 800 ms due to serial IOIs that follow. Oscillations of periods 400 and 800 ms find added support mainly from higher event time spans between non-adjacent tones. This analysis shows how rhythm and metre co-constrain oscillation activities. Once active, co-occurring internal oscillations mutually entrain.

I propose that, over time, internal entrainment leads to cluster binding and the formation of a persisting metric form (here duple metre). This is a mechanism for bootstrapping learning based on pattern relationships. It is constrained by binding principles such that learning is facilitated not only by longer total times of oscillator co-activity (e.g., Hebbian learning), but also by strong resonance and phase relations among active oscillations. The metric binding




Multiple simultaneously-active oscillations *bind* as a function of: 1. Duration of simultaneous activity 2. Intrinsic period ratios (*pi/pj*)3. Coincidence of pulse peaks

**Fig. 8.3** Multiple metric levels entrain corresponding neural oscillations. Rhythm (serial IOIs) also activates oscillations. This figure illustrates metric binding hypothesis principles.

Large and Jones 1999). First, neural oscillations are *self-sustaining;* they persist over time, extrapolating the induced beat. Second, an oscillator's intrinsic period exhibits *stability;* a perturbation from an ill-timed tone only briefly disrupts a stable oscillation, which returns to its intrinsic period. Third, entraining oscillations exhibit *adaptivity*, the flip side of stability; an oscillator responds appropriately to event-generated expectancy violations by adjusting phase and period. Fourth, *multiple related oscillations* are activated by multiple time levels within metric and rhythmic events.

The metric binding hypothesis expands these assumptions. It adds learning principles to address training and enculturation that contribute to listeners' familiarity with metric categories. It holds: *whenever two or more neural oscillations are simultaneously active, over time their internal entrainments lead to binding and formation of a metric cluster*. A metric cluster comprises sets of co-occurring oscillations with interrelationships that persist due to acquired internal bindings. Entrainments among internal oscillations promote binding, which strengthens as a function of:

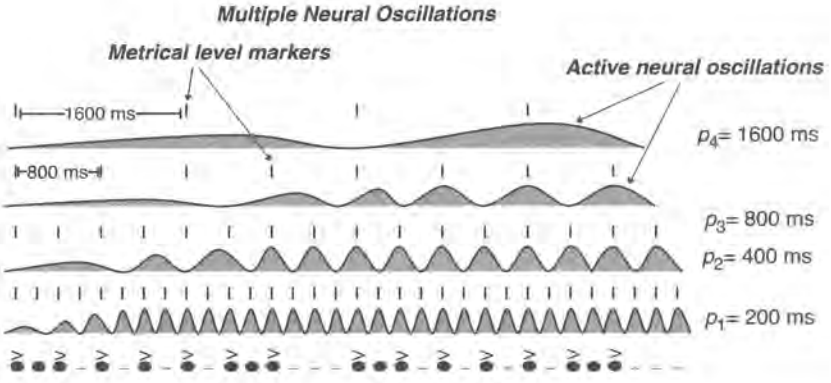
1 Duration of co-occurring oscillatory activity.

2 Phase coincidences, and,

3 Resonance (i.e., relatedness) among oscillator periods.

Figure 8.3 illustrates how metric clusters form. As a rhythmic pattern unfolds, it successively activates neural oscillations with periods (Pi) matching each of the lower-level IOIs. Initially recurrent IOIs, in this case, elicit an oscillation with a period of 200 ms, which is strongly coupled to these IOIs. However, this rhythm subsequently awakens oscillations with periods of 400 and 800 ms due to serial IOIs that follow. Oscillations of periods 400 and 800 ms find added support mainly from higher event time spans between non-adjacent tones. This analysis shows how rhythm and metre co-constrain oscillation activities. Once active, co-occurring internal oscillations mutually entrain.

I propose that, over time, internal entrainment leads to cluster binding and the formation of a persisting metric form (here duple metre). This is a mechanism for bootstrapping learning based on pattern relationships. It is constrained by binding principles such that learning is facilitated not only by longer total times of oscillator co-activity (e.g., Hebbian learning), but also by strong resonance and phase relations among active oscillations. The metric binding




Multiple simultaneously-active oscillations *bind* as a function of: 1. Duration of simultaneous activity 2. Intrinsic period ratios (***pi/pj***)3. Coincidence of pulse peaks

**Fig. 8.3** Multiple metric levels entrain corresponding neural oscillations. Rhythm (serial IOIs) also activates oscillations. This figure illustrates metric binding hypothesis principles.

hypothesis implies that: 'Oscillators that align together, bind together.'

Once acquired, metric clusters grant trained listeners attentional flexibility to activate oscillators for unmarked metric levels (Palmer and Krumhansl 1990) and to flexibly shift focal attending to different marked levels. In fact, Jones and Boltz (1989) proposed two attending modes, future-oriented and analytic attending, to reflect focal attending to higher and lower time levels, respectively.

## Specialized entrainment models

Several DAT models formalize oscillator behaviours for different tasks; typically, a single oscillation exhibits three components: period; pulse, and phase. Some tasks rely on a single oscillation model, others on multiple oscillations.

The simplest oscillator model appears in Figure 8.4 entraining to an isochronous event (Large and Jones 1999). This single oscillation carries a concentration of attending energy, an attentional pulse. With rhythmically simple events, an oscillator's 'tuning' is reflected in a steady narrowing of an attentional pulse about an expected phase point: $\phi = 0$. This entails continuous phase adjustments to minimize differences between momentarily expected phases and observed tone onsets. The goal is an attractor, defined by phase coincidence (synchrony), of expected and observed time points. Once attuned, an entrained oscillation persists to extrapolate beats (attentional pulses), each beat realizing an anticipated region in time (insert).

Metre perception is described with more complex models; at least two oscillations, entraining at different metric levels, are necessary (see Figure 8.3) (Large and Jones 1999; Large and Palmer 2002). Moreover, these oscillations can interact. To illustrate, consider pattern 2 (Figure 8.2). This rhythm differs from the rhythm of Figure 8.3 in containing lower order IOIs that elicit oscillations with periods: 200, 400, 600, 800 ms. That is, three oscillator periods (200, 400, 800 ms) nest neatly with higher metric levels (400, 800 ms, 1600 ms), but one does not (600 ms). Nevertheless, the metric binding hypothesis assumes internal entrainments

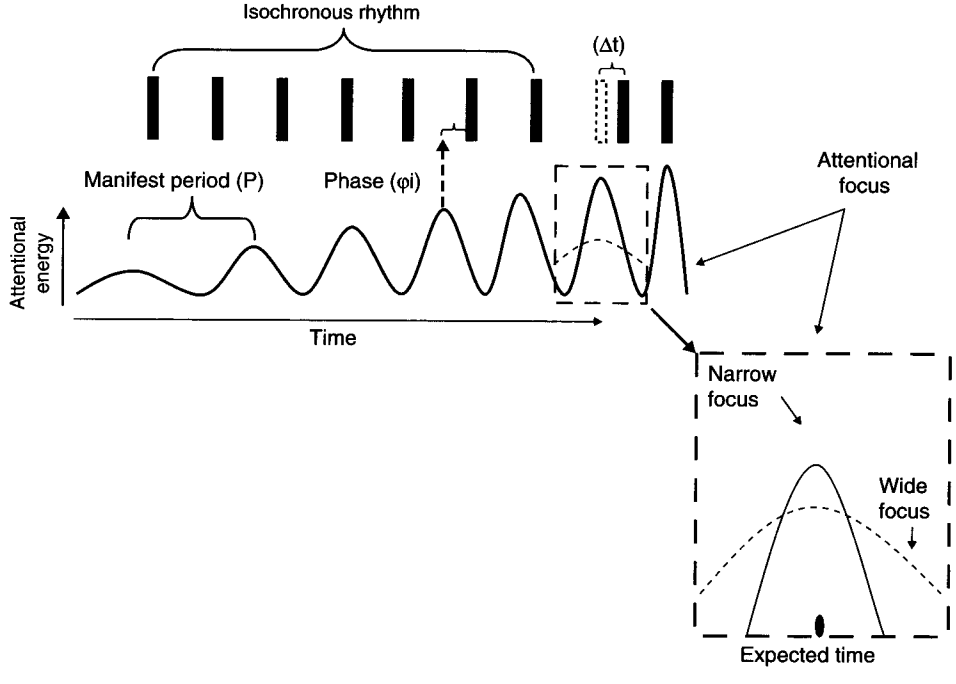



**Fig. 8.4** A single oscillator entrains to an isochronous rhythm. Basic components are period, pulse (insert) and phase.

hypothesis implies that: 'Oscillators that align together, bind together.'

Once acquired, metric clusters grant trained listeners attentional flexibility to activate oscillators for unmarked metric levels (Palmer and Krumhansl 1990) and to flexibly shift focal attending to different marked levels. In fact, Jones and Boltz (1989) proposed two attending modes, future-oriented and analytic attending, to reflect focal attending to higher and lower time levels, respectively.

### Specialized entrainment models

Several DAT models formalize oscillator behaviours for different tasks; typically, a single oscillation exhibits three components: period; pulse, and phase. Some tasks rely on a single oscillation model, others on multiple oscillations.

The simplest oscillator model appears in Figure 8.4 entraining to an isochronous event (Large and Jones 1999). This single oscillation carries a concentration of attending energy, an attentional pulse. With rhythmically simple events, an oscillator's 'tuning' is reflected in a steady narrowing of an attentional pulse about an expected phase point: $\phi = 0$. This entails continuous phase adjustments to minimize differences between momentarily expected phases and observed tone onsets. The goal is an attractor, defined by phase coincidence (synchrony), of expected and observed time points. Once attuned, an entrained oscillation persists to extrapolate beats (attentional pulses), each beat realizing an anticipated region in time (insert).

Metre perception is described with more complex models; at least two oscillations, entraining at different metric levels, are necessary (see Figure 8.3) (Large and Jones 1999; Large and Palmer 2002). Moreover, these oscillations can interact. To illustrate, consider pattern 2 (Figure 8.2). This rhythm differs from the rhythm of Figure 8.3 in containing lower order IOIs that elicit oscillations with periods: 200, 400, 600, 800 ms. That is, three oscillator periods (200, 400, 800 ms) nest neatly with higher metric levels (400, 800 ms, 1600 ms), but one does not (600 ms). Nevertheless, the metric binding hypothesis assumes internal entrainments

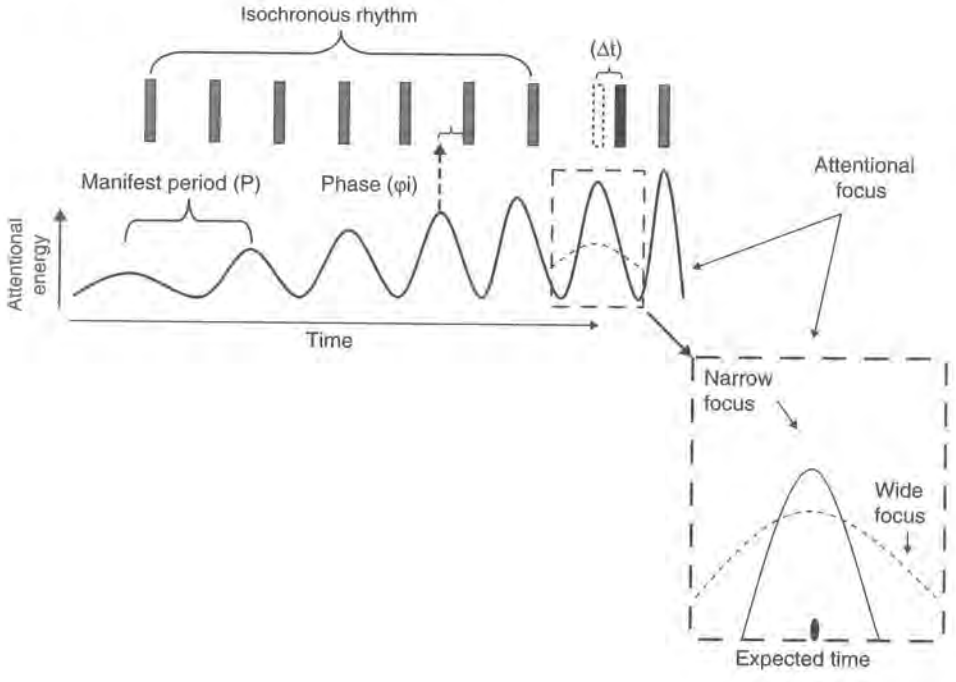



**Fig. 8.4** A single oscillator entrains to an isochronous rhythm. Basic components are period, pulse (insert) and phase.

among active oscillations lead to common metric percepts for different rhythms. Binding among oscillations depends not only on the duration of joint oscillator activities, but also on resonance relations among these oscillations, namely period and phase relationships. Thus, for pattern 2, although the oscillator with an ill-fitting period (e.g., 600 ms), is initially active, it dies out due to poor resonance. The winning metric cluster for pattern 2 turns out to be identical to the metric cluster for the different rhythm of Figure 8.3. This analysis predicts that the 600 ms IOI (pattern 2, Figure 8.2) will be perceptually distorted to 'fit' the duple metre frame common to both rhythms. More generally, metric binding in multi-oscillator systems explains how various rhythmic instances are perceived as members of a common metric category.

The mathematics of multi-oscillator systems is complex. Nevertheless, resonance properties require some mention. When two oscillations are active, their periods form embedding ratios (e.g., 1:1, 2:1, 3:1, 1:2, 2:3; 3:2 etc.) that gauge overall stability of their interaction. Embedding ratios specify resonance states (i.e., attractors) that have different degrees of stability; this, in turn, predicts differential lasting qualities of metric clusters (Large and Kolen 1995; cf. London 2004). For instance, clusters based on duple metre ratios are predicted to be more stable than those based on complex ratios.

Finally, it is common to distinguish metre perception from rhythm perception, in that it reflects an acquired skill, whereas rhythm perception is attributed to innate Gestalt principles (Bregman 1990; Handel 1989). DAT does not incorporate this dichotomy. Rather, dynamical properties offer biological springboards that differentially constrain the learning of metric versus rhythmic time structures. DAT implies that unlearned biological responses (neural oscillations) are fundamental to time perception, whether metric or rhythmic. These biological dispositions (oscillatory entrainments) initially facilitate attending to event structure via activated oscillations. Next, stability differences among simultaneously active oscillations lead to differential learning of metric patterns versus non-metric and rhythmic patterns given the metric binding hypothesis.

## Evidence for DAT

Evidence associated with various entrainment models of metre perception derives from both behavioral and biological sources.

### *Behavioural assessments of DAT*

Experimental research often concerns responses to timing of metrically regular and irregular contexts. The most rigorous tests involve threshold judgements of timing and categorical time judgements (see Jones 2004 for a review).

Single oscillator entrainment models predict that lower-order serial IOIs of monotone rhythms are important, a finding confirmed by Demany and Semal (2002). These models also correctly predict that, in judging time intervals, listeners will distort unexpected time intervals to fit with a contextually expected beat span (Barnes and Jones 2000). Other entrainment models, which involve biologically preferred periods (McAuley 1995), also find support McAuley *et al.* (2006). Two oscillator entrainment models, used to describe metre perception, assume that one oscillation entrains to a higher metric level, marked by more salient accents, while another oscillation, with a shorter period, entrains to embedded event time spans, marked by less salient accents. Such a model was shown to correctly predict that greater metric regularity (within and between trials) enhances temporal acuity (Jones and Yee 1997; Large and Jones 1999).

Acuity thresholds are affected by tempo. London (2004) argued that such acuity limits affect metre identification, biasing listeners toward particular metres at certain tempi. Effectively, a perceived tactus corresponds to a lower metrical level in relatively slow patterns but a higher one in faster patterns (Handel and Oshinsky 1981; Parncutt 1994; Duke 1989).

McAuley *et al.* (2006) found that the limits of tempo perception (e.g. Weber's Law) depend upon entrainment constraints expressed by *entrainment regions.* An entrainment region is a range of tempi, surrounding a preferred event rate, that corresponds to good entrainment. Both the location and width of these regions were found to shift with age: children were shown to prefer narrow entrainment regions, surrounding a fast event rate (active oscillations with brief periods), whereas elderly listeners had

broader entrainment regions centered on slower events (active oscillations with longer periods). Thus, perceived metre may modulate not only with tempo but also with age. Also see Drake and Botte (1993), Miller and McAuley (2005), Jones and McAuley (2005), and Parncutt (1994).

### *Biological assessments of DAT*

Neurophysiological findings buttress entrainment theory (see Zanto *et al.* 2006). Consistent with DAT, electroencephalography (EEG) reveals neural oscillations that synchronize to periodic auditory stimuli in the musical range, i.e. 2 Hz (Will and Berg 2007).

Using event-related potentials (ERPs), Brochard *et al.* (2003) were the first to verify the presence of subjective accents, i.e., internalized beat patterns (Fraisse 1963; Woodrow 1932). They found ERP activity in the parietal cortex of non-musicians that reflected duple metre expectancies (see also Besson and Faita 1995; Janata 2001). Often ERP signal frequencies (0–10 Hz) reveal relatively long latency brain responses (e.g., P300), which may reflect a synthesis of higher frequency oscillations (Makeig *et al.* 2004). Indeed, fronto-cortical ERP recordings revealed two kinds of high frequency activities (Snyder and Large 2005):

1 An induced periodic response, in the Beta/Gamma range (20–30Hz), prior to onsets of metrically expected tones;

2 A phase-locked evoked Gamma (30–60 Hz) response following tone onsets.

These findings confirm DAT distinctions between expectancies (oscillator period) and expectancy violations (phase corrections) (Janata 2001; Zanto *et al.* 2005).

A common question is: 'Are internal oscillations purely event driven?' (Iverson *et al.* 2006). Both behavioural and neurophysiological studies suggest the answer is 'No'. Combined manipulations of event structure with task and instructions (imagery, attentional set) suggest that event structure plays a role in facilitating or inhibiting listeners' compliance with instructions (Palmer and Krumhansl 1990; Klein and Jones 1996; Janata and Grafton 2003; Snyder and Large 2005; Iverson *et al.* 2006). Clearly, people rely upon both guided imagination and event structure to shape metric expectancies.

## Rhythm

Rhythm is a serial figure based on an arrangement of discrete time intervals. It contrasts with metre, which is based upon embedded time intervals. In this section, I outline current limits on our understanding of rhythm perception.

### Theoretical background

Many contemporary approaches to rhythm perception assume that it is psychologically distinct from metre perception (Figure 8.1(b) versus Figure 8.1(a)). Rhythm perception is assumed to depend upon temporal grouping principles, given by Gestalt rules of proximity, similarity, continuity, etc. In turn, Gestalt theory holds that rhythm perception is innate, based upon an automatic, primitive, universal process that is governed by hard-wired, domain-free, grouping principles. Thus, rhythm perception is considered inherently different from metre perception, which is viewed as an acquired skill, reflecting domain-specific musical rules.

The classic dichotomy of rhythm and metre perception is appealing for several reasons. First, it captures an experiential difference between serial (rhythmic) and embedded (metric) time structures. Second, reliance on Gestalt principles leads to coarse coding of serial time spans which correctly captures listeners' often fuzzy percepts of rhythmic time spans. By this account, the rhythmic figure of 400–200–200 ms invites lax encoding of proximal tones as 'groups', segmented by non-proximal tones (i.e., groups of 1 versus 3 tones; cf. Figure 8.2). Nominal codes of time intervals as either short (proximal) or long (non-proximal) yield a coarse rhythmic code, such as long–short–short for this figure. Third, nominal codes often accurately predict perceptual confusions among rhythms. For instance, the 400–200–200 ms figure has the same nominal code (long–short–short) as 500–100–200 ms, leading to observed confusions of such rhythms. In sum, Gestalt rules are intuitively compelling and coarse temporal groupings, nominally encoded, offer explanatory value.

In spite of the appeal of Gestalt theory, a rhythmic/metre dichotomy cloaks pitfalls. First, because Gestalt principles are hard-wired, this approach denies that listeners may blend innate

with acquired responses to time structures. For instance, the two confusable rhythms (above) may be eventually differentiated as listeners acquire sensitivity to different temporal nesting properties. Second, by denying a role for learning, this approach discourages research on this topic. Indeed, Gestalt rules are often mistaken for final explanations when they are simply useful descriptions of phenomena that require explanation. Third, this dichotomy renders it difficult or impossible to address rhythmic priming of metre and vice versa.

DAT assumes that learning builds upon innate oscillatory brain activities. It offers a potential for explaining both metre and rhythm perception based on entrainment constraints associated with their respective time properties. Perceptual learning (i.e. the metric binding hypothesis) depends upon entrainments of innate oscillations elicited by an event's time spans. Although multiple time spans occur in both metric and non-metric rhythms (cf. Figure 8.1), they promote different learning paths. Relationships among saliently marked embedded time spans are typically orderly and aligned to highlight consistent ratio time relationships in metric patterns, whereas in rhythmic figures, time spans from grouping accents can offer unruly, misaligned, embeddings (especially in non-metric patterns) that obscure higher level temporal regularities. Because of this, metric patterns promote quicker binding of oscillators across embedded time levels than do rhythmic patterns. Rhythmic figures that lack consistent higher-order time spans cannot support effective entrainment of higher level oscillators; instead, loosely connected oscillations among IOI resolve to group segmentations. It follows that DAT can address rhythmic priming of metre and vice versa (Desain and Honing 2003). Nevertheless, despite its potential no DAT task model has formalized these ideas to rigorously explain serial segmentation in rhythmic patterns. In this regard DAT and Gestalt approaches share incomplete explanations of rhythm perception.

### Empirical evidence on rhythm perception

Listeners can distinguish among theoretically confusable rhythms if timing differences are sufficiently large (Hebert and Cuddy 2002) or if patterns occur repeatedly (Handel 1992). Although poorly understood, listeners do learn to differentiate confusable rhythms having similar Gestalt grouping codes.

An important difference between DAT and Gestalt accounts is the latter's emphasis on nominal coding of rhythmic time spans. By contrast, DAT features roles for both interval (tempo) and ratio (metre/rhythm) time relationships. Critically, even in brief rhythmic sequences, listeners are sensitive to rhythmic categories with special serial time ratios (e.g., 1:1, 2:1) (e.g., Desain and Honing 2003). Using only two time intervals, incremental lengthening of the first interval (relative to the second) between serial ratios of 1:1 and 2:1 (and vice versa) revealed preferences for the simpler anchoring ratios, consistent with Fraisse's ideas. Not to be overlooked is research showing metric priming of such rhythms (Povel 1981; Clarke 1987; ten Hoopen *et al.* 2006)

Related findings accompany a *time-shrinking* phenomenon (Nakajima *et al.* 1992; Sasaki *et al.* 2002; Nakajima *et al.* 2004). Using brief two-interval patterns, Nakijima and his colleagues find that when a long time interval follows a short one, listeners underestimate it, revealing a gravitation to a preferred serial time ratio of 1:1. However, reversal of the two intervals fails to elicit time expansion. Generally, research on time distortions suggests that perception of rhythmic figures is influenced by stability of certain serial time ratios: 1:1, 2:1, 1:2. Interestingly, brain activities, measured by functional magnetic resonance imaging (fMRI), also appear to depend upon the simpler time ratios produced by listeners (i.e., stable anchoring ratios) and not stimulus ratios (Sakai *et al.* 1999).

## Marking time and accent salience

All time patterns exist because *accents* mark constituent time spans. Accents 'call attention' to onsets of time spans. In this section I focus upon the neglected issue of *accent salience* (for different perspectives, see Krumhansl 2000; Clarke 1999).

## Theoretical background

The dichotomy of rhythm and metre perception holds implications for understanding accents as time markers. Sometimes accents are drawn into this dichotomy through the assumption that codes for metre and rhythm are processed independently and/or stored in respectively separate modules. Thus, metric versus rhythmic accents may be distinguished by linking them, respectively, with different physical dimensions (e.g., time, intensity, pitch etc.). For instance, in some views metre is considered to be marked by intensity accents whereas rhythm is marked by duration accents. A related hypothesis holds that rhythm and metre percepts are not only dichotomous, but that both are independent of melodic features, e.g., pitch accents.

Complicating discussions of independence versus dependence of metre, rhythm and melody is the practice of differentiating musical accents a priori by dimensionality alone, e.g., as intensity, time, or pitch accents. This practice raises certain dilemmas. Consider this: if an accent effectively 'calls' attention to a point in time, then logically accents *cannot* be solely defined by their dimensions. Rather, accents must be contextually defined because they only occur in serial contexts. An accent gains its power as a *serial change*, regardless of dimension. A large relative intensity change in a musical sequence is more attention-getting, i.e., more salient, than a small one. Thus, to compare accents of different types (i.e., dimensions) as effective markers of metre, it is important to equate them for salience.

Dynamic attending theory assumes that accents arise from local serial changes (Jones 1987). This broadens the definition of phenomenal accents of Lerdahl and Jackendoff (1983). Operationally, the salience of an accented tone increases with the:

1 Magnitude of its local serial change along a dimension of variation;
2 Number of simultaneous accents on it (i.e., all else equal, two co-occurring serial changes are stronger than one);
3 Surrounding variability (melodic, rhythmic, etc.) in global serial context (Ellis and Jones in press).

In DAT, various accent types mark time spans of both rhythm and metre. Further, if the perceptual impact of an accent, i.e. its salience, turns on the magnitude of a local serial change, then different accent types (pitch, time, etc.) can have equivalent salience. This idea is formalized in the concept of joint accent structure (JAS) (Jones 1987). A JAS reflects a temporal collaboration of different (salient) accent types that

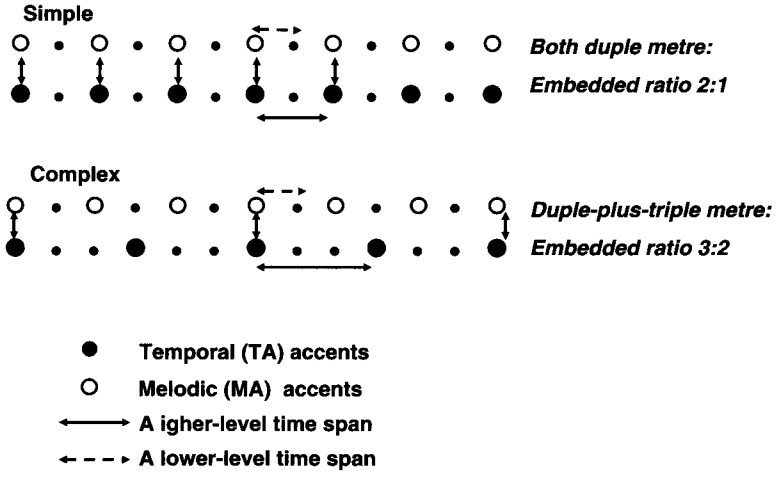



**Fig. 8.5** Upper JAS example: aligned melodic (MA) and duration accents (TA) lead to a duple metre JAS pattern. Lower: misaligned accent sequences form a polyrhythmic JAS-based duple (MA) and triple (TA) patterns.

outline a common, higher order, time structure. As shown in Figure 8.5, structural interdependence of melodic (pitch change) and rhythmic (duration changes) accents is inherent in that a metric structure emerges from time spans marked by both melodic and rhythmic accents. One configuration reflects a simple combination of melodic accent and rhythmic accent sequences (duple metre embedding ratios of 2:1) where both melodic (MA) and temporal (TA) accents align; in the other, a more complex JAS combination results in misaligned accents (polyrhythmic ratios of 3:2). According to a metric binding hypothesis, the former is the more stable due to its resonance properties. Nevertheless, as experimental compositions, both JAS patterns reflect interdependence of melodic, rhythmic and metric structures.

## Experimental evidence on marking time

Experimental evidence offers mixed support for hypotheses about perception of musical events containing melodic and temporal accents. Some favours a hypothesis that melody, rhythm and metre are perceived independently, whereas other evidence favours a perceptual dependency hypothesis. Such issues remain unresolved.

Research which relies upon established compositions is consistent with the idea that melody and metre are perceptually independent. For instance, listeners' inferred beats indicate a reliance on rhythmic (duration) over melodic accents (Snyder and Krumhansl 2001; Hannon *et al.* 2004). Hence, Huron and Royal (1996) question the effectiveness of melodic accents for marking metre. Yet evidence for this conclusion remains inconclusive. In part, this is because the salience of various accents is unknown, and in part it is because such correlational evidence precludes causal inferences. Furthermore, other evidence, based on experimental compositions, favours the alternative view. In these, melodic (pitch) and rhythmic (temporal) accents together appeared to determine listeners' sense of metre (Boltz and Jones 1986; Pfordresher 2003).

A resolution of this issue depends upon insuring comparable salience across accent types. For example, in specified musical contexts we must gauge whether a three semitone pitch leap (melodic accent) is equal in salience to lengthening tone duration by 5 per cent or by 15 per cent (rhythm accents). Overall variability of surrounding melodic and rhythmic contexts must also be controlled.

A few studies have controlled accent salience. Windsor (1993) calibrated intensity accents, and confirmed that larger serial changes in intensity yielded clearer metric identifications. Ellis and Jones (in press) used melodic (MA) and temporal (TA) accents of calibrated equivalence in various sequences to create nine different JAS patterns; to appear in Figure 8.5 (i.e., 2:1 duple; 3:2 polyrhythm). Different JAS conditions combined duple and triple accent patterns in aligned (e.g., duple for TA and MA) and misaligned (e.g., duple TA and triple MA) ways. Listeners rated the metric clarity of all patterns. Results appear in Figure 8.6. In JAS patterns with aligned MAs and TAs, time spans should activate multiple aligned oscillations (for duple or triple metre), and thus here DAT correctly predicts stable metric clusters and high metric clarity. By contrast, in misaligned JAS patterns, with irregular high-level time spans, DAT correctly predicts lower clarity.

These findings reveal the importance of calibrating accent salience. Melodic and rhythmic accents, of comparable salience, lead to JAS

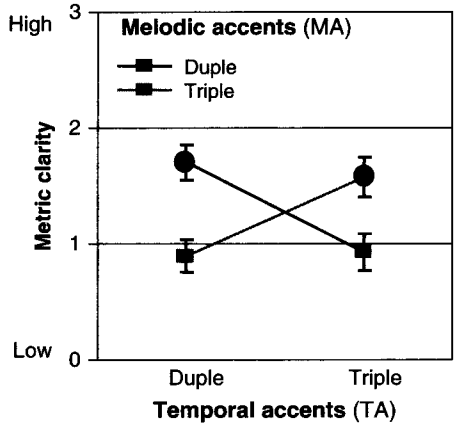



**Fig. 8.6** Metric clarity ratings of average listeners. Melodic (MAs, dark lines) and temporal (TAs, light lines) accent patterns are either aligned (e.g., circles) or mis-aligned (squares) (Ellis and Jones in press).

outline a common, higher order, time structure. As shown in Figure 8.5, structural interdependence of melodic (pitch change) and rhythmic (duration changes) accents is inherent in that a metric structure emerges from time spans marked by both melodic and rhythmic accents. One configuration reflects a simple combination of melodic accent and rhythmic accent sequences (duple metre embedding ratios of 2:1) where both melodic (MA) and temporal (TA) accents align; in the other, a more complex JAS combination results in misaligned accents (polyrhythmic ratios of 3:2). According to a metric binding hypothesis, the former is the more stable due to its resonance properties. Nevertheless, as experimental compositions, both JAS patterns reflect interdependence of melodic, rhythmic and metric structures.

## Experimental evidence on marking time

Experimental evidence offers mixed support for hypotheses about perception of musical events containing melodic and temporal accents. Some favours a hypothesis that melody, rhythm and metre are perceived independently, whereas other evidence favours a perceptual dependency hypothesis. Such issues remain unresolved.

Research which relies upon established compositions is consistent with the idea that melody and metre are perceptually independent. For instance, listeners' inferred beats indicate a reliance on rhythmic (duration) over melodic accents (Snyder and Krumhansl 2001; Hannon *et al.* 2004). Hence, Huron and Royal (1996) question the effectiveness of melodic accents for marking metre. Yet evidence for this conclusion remains inconclusive. In part, this is because the salience of various accents is unknown, and in part it is because such correlational evidence precludes causal inferences. Furthermore, other evidence, based on experimental compositions, favours the alternative view. In these, melodic (pitch) and rhythmic (temporal) accents together appeared to determine listeners' sense of metre (Boltz and Jones 1986; Pfordresher 2003).

A resolution of this issue depends upon insuring comparable salience across accent types. For example, in specified musical contexts we must gauge whether a three semitone pitch leap (melodic accent) is equal in salience to lengthening tone duration by 5 per cent or by 15 per cent (rhythm accents). Overall variability of surrounding melodic and rhythmic contexts must also be controlled.

A few studies have controlled accent salience. Windsor (1993) calibrated intensity accents, and confirmed that larger serial changes in intensity yielded clearer metric identifications. Ellis and Jones (in press) used melodic (MA) and temporal (TA) accents of calibrated equivalence in various sequences to create nine different JAS patterns; to appear in Figure 8.5 (i.e., 2:1 duple; 3:2 polyrhythm). Different JAS conditions combined duple and triple accent patterns in aligned (e.g., duple for TA and MA) and misaligned (e.g., duple TA and triple MA) ways. Listeners rated the metric clarity of all patterns. Results appear in Figure 8.6. In JAS patterns with aligned MAs and TAs, time spans should activate multiple aligned oscillations (for duple or triple metre), and thus here DAT correctly predicts stable metric clusters and high metric clarity. By contrast, in misaligned JAS patterns, with irregular high-level time spans, DAT correctly predicts lower clarity.

These findings reveal the importance of calibrating accent salience. Melodic and rhythmic accents, of comparable salience, lead to JAS

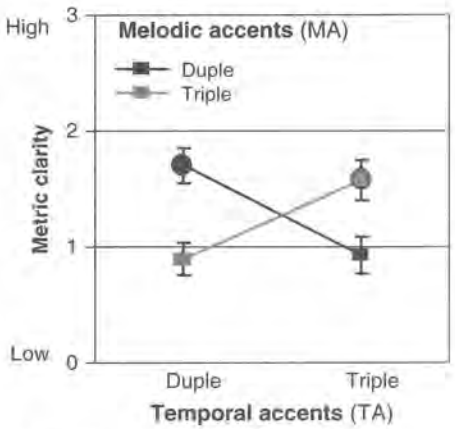



**Fig. 8.6** Metric clarity ratings of average listeners. Melodic (MAs, dark lines) and temporal (TAs, light lines) accent patterns are either aligned (e.g., circles) or mis-aligned (squares) (Ellis and Jones in press).

patterns of different complexity when combined with embedding ratios of 2:1 versus 3:2. Listeners' ratings of these patterns supports the hypothesis that percepts of melody, rhythm and metre are not independent.

## Summary

This chapter selectively reviews psychological research on perception of metre and rhythm. Its focus is upon the dynamics of attending to patterns in time. It considers how event timing, as outlined by salient accents in metre, rhythm, and even melodic structures, may guide attending in time, bootstrap learning, and influence time perception.

## Acknowledgements

The author is grateful for the assistance of Robert Ellis.

## References

Barnes R and Jones MR (2000) Expectancy, attention, and time. *Cognitive Psychology*, **41**, 254–311.

Bergeson TR and Trehub SE (2006) Infants' perception of rhythmic patterns. *Music Perception*, **23**, 345–360.

Besson M and Faita F (1995) An event-related potential (ERP) study of musical expectancy: comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*, **21**, 1278–1296.

Boltz M and Jones MR (1986) Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology*, 18, 389–431.

Bregman A (1990) *Auditory scene analysis*. MIT Press, Cambridge, MA.

Brochard R, Abecasis D, Potter D, Ragot R and Drake C (2003) The 'ticktock' of our internal clock: direct brain evidence of subjective accents in isochronous sequences. *Psychological Science*, **14**, 362–366.

Clarke EF (1987) Levels of structure in the organization of musical time. *Contemporary Music Review*, **2**, 211–238.

Clarke EF (1999) Rhythm and timing in music. In D Deutsch, ed., *The psychology of music*, 473–500. Academic Press, New York.

Demany L and Semal C (2002) Limits of rhythm perception. *Q J Exp Psychol A*, **55**, 643–657.

Desain P and Honing H (2003) The formation of rhythmic categories and metric priming. *Perception*, **32**, 341–365.

Drake C (1993) Reproduction of musical rhythms by children, adult musicians and adult nonmusicians. *Perception and Psychophysics*, **53**, 25–33.

Drake C and Botte M (1993) Tempo sensitivity in auditory sequences: evidence for a multiple-look model. *Perception and Psychophysics*, **54**, 277–286.

Duke RA (1989) Musicians' perception of beat in monotonic stimuli. *Journal of Research in Music Education*, **37**, 61–71.

Ellis RJ and Jones MR (in press) The role of accent salience and joint accent structure in meter perception. *Journal of Experimental Psychology: Human Perception and Performance*.

Essens PJ (1986) Hierarchical organization of temporal patterns. *Perception and Psychophysics*, **40**, 69–73.

Fraisse P (1963) *The psychology of time*. Harper and Row, New York.

Handel S (1989) *Listening: an introduction to the perception of auditory events*. MIT Press, Cambridge, MA.

Handel S (1992) The differentiation of rhythmic structure. *Perception and Psychophysics*, **52**, 497–507.

Handel S (1998) The interplay between metric and figural organization. *Journal of Experimental Psychology: Human Perception and Performance*, **24**, 1546–1561.

Handel S and Oshinsky JS (1981) The meter of syncopated auditory patterns. *Perception and Psychophysics*, **30**, 1–9.

Hannon EE, Snyder JS, Eerola T and Krumhansl CL (2004) The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, **30**, 956–974.

Hebert S and Cuddy LL (2002) Detection of metric structure in auditory figural patterns. *Percept Psychophys*, **64**, 909–918.

Huron D and Royal M (1993). What is melodic accent?. *Converging evidence from musical practice. Music Perception*, **13**, 489–516.

Iverson JR, Repp BH and Patel AD (2006) *Metrical interpretation modulates brain responses to rhythmic sequences*. The Neuroscience Institute, San Diego, CA.

Janata P (2001) Brain electrical activity evoked by mental formation of auditory expectations and images. *Brain Topogr*, **13**, 169–193.

Janata P and Grafton ST (2003) Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nat Neurosci*, **6**, 682–687.

Jones MR (1987). Dynamic pattern structure in music: recent theory and research. *Perception and Psychophysics*, **41**, 621–634.

Jones MR (2004) Attention and timing. In JG Neuhoff, ed., *Ecological psychoacoustics*, 45–89. Academic Press, San Diego, CA.

Jones MR and Boltz M (1989) Dynamic attending and responses to time. *Psychological Review*, **96**, 459–491.

Jones MR and Mcauley JD (2005) Time judgments in global temporal contexts. *Percept Psychophys*, **67**, 398–417.

Jones MR and Yee W (1997) Sensitivity to time change: the role of context and skill. *Journal of Experimental Perception and Psychophysics*, **32**, 693–709.

Klein JM and Jones MR (1996) Effects of attentional set and rhythmic complexity on attending. *Perception and Psychophysics*, **58**, 34–46.

Krumhansl CL (2000) Rhythm and pitch in music cognition. *Psychological Bulletin*, **126**, 159–179.

Large EW (1994) Dynamic representation of musical structure. Unpublished doctoral dissertation. Ohio State University, Columbus.

Large EW and Jones MR (1999) The dynamics of attending: how people track time-varying events. *Psychological Review*, **106**, 119–159.

Large EW and Kolen J (1995) Resonance and the perception of musical meter. *Connection Science*, **6**, 177–208.

Large EW and Palmer C (2002) Perceiving temporal regularity in music. *Cognitive Science*, **26**, 1–37.

Lerdahl F and Jackendoff R (1983) *A generative theory of tonal music*. MIT Press, Cambridge, MA.

London JM (2004) *Hearing in time: psychological aspects of musical meter*. Oxford University Press, New York.

Makeig S, Debener S, Onton J and Delorme A (2004) Mining event-related brain dynamics. *Trends Cogn Sci*, **8**, 204–210.

Mcauley JD (1995) *Perception of time phase: toward an adaptive oscillator model of rhythmic pattern processing*. Indiana University Press, Bloomington, IN.

Mcauley JD and Semple P (1999) The effect of tempo and musical experience on perceived beat. *Australian Journal of Psychology*, **51**, 176–187.

Mcauley JD, Jones MR, Holub S, Johnston HM and Miller NS (2006) The time of our lives: life span development of timing and event tracking. *J Exp Psychol Gen*, **135**, 348–367.

Miller NS and Mcauley JD (2005) Tempo sensitivity in isochronous tone sequences: the multiple-look model revisited. *Percept Psychophys*, **67**, 1150–1160.

Nakajima Y, Ten Hoopen G, Hilkhuysen G and Sasaki T (1992) Time-shrinking: a discontinuity in the perception of auditory temporal patterns. *Percept Psychophys*, **51**, 504–507.

Nakajima Y, Ten Hoopen G, Sasaki T, Yamamoto K, Kadota M, Simons M and Suetomi D (2004) Time-shrinking: the process of unilateral temporal assimilation. *Perception*, **33**, 1061–1079.

Palmer C and Krumhansl CL (1990) Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 728–741.

Palmer C and Pfordresher PQ (2003) Incremental planning in sequence production. *Psychol Rev*, **110**, 683–712.

Parncutt R (1994) A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, **11**, 409–464.

Patel AD, Iversen JR, Chen Y and Repp BH (2005) The influence of metricality and modality on synchronization with a beat. *Exp Brain Res*, **163**, 226–238.

Pfordresher PQ (2003) The role of melodic and rhythmic accents in musical structure. *Music Perception*, **20**, 431–464.

Phillips-Silver J and Trainor LJ (2005) Feeling the beat: movement influences infant rhythm perception. *Science*, **308**, 1430.

Povel D-J and Essens P (1985) Perception of temporal patterns. *Music Perception*, **2**, 411–440.

Povel D-J (1981) Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance*, **7**, 3–18.

Povel DJ and Okkerman H (1981) Accents in equitone sequences. *Percept Psychophys*, **30**, 565–72.

Repp BH (2005) Sensorimotor synchronization: a review of the tapping literature. *Psychon Bull Rev*, **12**, 969–992.

Ross J and Houtsma AJ (1994) Discrimination of auditory temporal patterns. *Percept Psychophys*, **56**, 19–26.

Sakai K, Hikosaka O, Miyauchi S, Takino R, Tamada T, Iwata NK and Nielsen M (1999) Neural representation of a rhythm depends on its interval ratio. *Journal of Neuroscience*, **19**, 10074–10081.

Sasaki T, Suetomi D, Nakajima Y and Ten Hoopen G (2002) Time-shrinking, its propagation, and Gestalt principles. *Percept Psychophys*, **64**, 919–931.

Snyder JS and Krumhansl CL (2001) Tapping to ragtime: cues to pulse finding. *Music Perception*, **18**, 455–489.

Snyder JS and Large EW (2005) Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research*, **24**, 117–126.

Temperley D (2001) *The cognition of basic musical structures*. MIT Press, Cambridge, MA.

Ten Hoopen G, Nakajima Y, Remin G, Massier, B, Rhebergen K and Holleman W (2006) Time-shrinking and categorical temporal ratio perception: evidence for a 1:1 temporal category. *Music Perception*, **24**, 1–22.

Trehub SE and Hannon EE (2006) Infant music perception: domain-general or domain-specific mechanisms? *Cognition*, **100**, 73–99.

Will U and Berg E (2007). Brain wave synchronization and entrainment to periodic acoustic stimuli. *Neuroscience Letters*, **424**, 55–60.

Windsor WL (1993) Dynamic accents and categorical perception of metre. *Psychology of Music*, **21**, 127–140.

Woodrow H (1932) The effects of rate of sequence upon the accuracy of synchronization. *Journal Experimental Psychology*, **15**, 357–379.

Zanto TP, Large EW, Fuchs A and Kelso JAS (2005) Gamma-band responses to perturbed auditory sequences: evidence for synchronization of perceptual processes. *Music Perception*, **22**, 535–552.

Zanot TP, Snyder JS and Large EW (2006) Neural correlates of rhythmic expectancy. *Advances in Cognitive Psychology*, **2**, 221–231.

# CHAPTER 9

# Components of melodic processing

Mark A. Schmuckler

## Introduction

Melody is the most ubiquitous form of musical structure with which listeners come into contact on a daily basis. Mirroring the prevalence and importance of melody, research in music cognition has focused extensively on the processes involved in perceiving and remembering melodic structure. Despite these years of study, however, our understanding of pitch structure in melody can be described simply, with respect to the two components of tonality and pitch contour. Although the importance of these two components has been recognized over the years, it is only recently that workable models of these components have been proposed. This chapter describes such models of tonality and melodic contour, and discusses the role of these components in listeners' perceptions of and memory for melody.

## Components of melodic processing

Melody is, with little doubt, the most ubiquitous form of musical structure with which people come into contact on a daily basis. Everywhere in our environment we are exposed to melody, from the songs we hear on the radio to the tunes we hum to ourselves. Additionally, melody is arguably the first clearly musical structure to which infants are exposed, through the prevalence of parental singing (e.g. Nakata and Trehub 2004, Rock *et al.* 1999; Shenfield *et al.* 2003). Finally, although cultures throughout the world vary in their respective musical structures, virtually all contain some form of melody (Eerola *et al.* 2006; Unyk *et al.* 1992).

Given its central role in music, it is no surprise that investigations of melody have occupied a similarly privileged position in music cognition. Such research has taken a variety of forms, ranging from explorations of how melodies are formed (i.e., what principles underlie the grouping of individual notes into a single perceptual object such as melody), to investigations of how melodies are themselves perceived (e.g., what does a listener truly experience upon hearing a melody) to how melodies are perceived in relation to one another (e.g., melodic categorization or similarity) to how melodies are remembered. Interestingly, although research on such questions has been extensive and insightful, there has been little explicit integration across these topics. Along these lines, one can ask whether there are any fundamental principles or organizational structures that function across these diverse aspects of melodic processing.

Fortunately, a close look at the literature on melodic processing implicates the operation of two basic organizational structures. Specifically, there is strong evidence for two fundamental components of melodic processing—a melody's tonal structure, and its melodic contour. Throughout a wealth of empirical findings these two components have time and again shown themselves to play a principal role in listeners' melodic processing, with tonality and contour continually influencing listeners' perceptions of

and responses to melodies, as well as their subsequent memory for these melodies.

Recognition of the importance of tonality and contour in melody perception is not a new insight (e.g. Dowling 1978). In a classic paper, Dowling (1978) identified just these two components as critical in melodic processing, positing that melodic memory, and by implication melodic perception, is driven by a melody's tonality and contour. Over the years Dowling and others have continually demonstrated the importance of these components in melodic processing, employing a wide array of perceptual encoding, similarity, and memory paradigms.

Although our understanding of melodic processing has grown substantially since Dowling's (1978) original insights, nothing has yet arisen that would undermine the proposed fundamental role played by these two components. As such, tonality and contour are still considered to be the two primary characteristics of listeners' perceptions of melody (e.g. Krumhansl 2000; Schmuckler 2004). One way in which our knowledge has advanced, however, is in the establishment of models for formalizing the structure of these components. Such models allow for an explicit characterization of these parameters, and thus enable the generation of specific predictions for melodic processing based on the structure highlighted by these models.

## Models of tonality and contour

### Models of tonality

The impact of musical tonality, or the hierarchical organization of the chromatic set around a single reference pitch, on the processing of musical passages is one of the most thoroughly studied research topics in music cognition. Approached from musicological (e.g. Lerdahl 2001; Temperley 2001, 2007), psychological (e.g. Krumhansl 1990a, 2000), and neuroscientific perspectives (e.g. Janata *et al.* 2002; Koelsch and Siebel 2005), our knowledge of the impact of tonal structure on musical processing is immense.

In keeping with this focus, there have been a large number of models proposed to account for the psychological organization of tonality (Krumhansl and Kessler 1982; Shepard 1982; Tillman *et al.* 2000), as well as explaining how listeners apprehend tonality in the first place. With reference to this latter goal, models describing the process of key-finding can be roughly divided into two categories (Huron and Parncutt 1993; Schmuckler and Tomovski 2005; Temperley and Marvin 2008)—those that posit the use of structural–functional information in key determination, and those that employ an event distribution strategy to key-finding.

The structural–functional models that have been proposed are diverse (e.g. Abe and Hoshino 1990; Brown 1988; Brown and Butler 1981; Butler 1990; Butler and Brown 1994; Holtzman 1977; Longuet-Higgins and Steedman 1971; Matsunaga and Abe 2005; Yoshino and Abe 2004), although all share an underlying assumption that there are particular local features or components of musical patterns that signal or indicate a specific musical key. Of these structural models, the most well-known are the approaches proposed by David Butler and Helen Brown (Brown 1988; Brown and Butler 1981; Butler 1990; Butler and Brown 1984, 1994) and by Jun-ichi Abe and colleagues (Matsunaga and Abe 2005; Yoshino and Abe 2004). Although a full review of these approaches cannot be considered here, Brown and Butler's 'intervallic rivalry theory' assumes that it is the presence of the rare intervals of the diatonic set (the minor seconds and the tritone), along with an additional disambiguating tone, that indicate the to-be-perceived tonality. Similarly, Abe and colleagues have posited that key-finding arises due to attempts to interpret the tones of a melodic pattern as members of a given diatonic set, and specifically the tonic triad (Abe and Hoshino 1990; Abe and Okada 2004; Matsunaga and Abe 2005; Yoshino and Abe 2004). Thus, both approaches explicitly look for certain structural and/or functional relations within musical events, and use the occurrence of these events to establish a key.

An alternative to this approach are distributional models of key-finding. Probably the best-known such model was proposed by Krumhansl and Schmuckler (1986a; see Krumhansl 1990a), who suggested that key-finding could be accomplished by matching the relative durations of the chromatic set in a piece of music with the hierarchies of perceived tonality stability of the

chromatic set with reference to the 12 major and 12 minor tonalities (Krumhansl and Kessler 1982). Although numerous refinements to this model have been proposed (Huron and Parncutt 1993; Smith and Schmuckler 2004; Temperley 1999, 2001, 2002, 2007; Toiviainen and Krumhansl 2003), the basic premise of this approach has remained invariant.

Both models have, over the years, garnered an impressive degree of empirical support. Multiple studies have demonstrated the importance of certain structural relations in key-finding (Brown 1988; Brown and Butler 1981; Brown *et al.* 1994). In Brown and Butler (1981), for instance, listeners' key judgements in response to melodic trichords accorded well with the predictions of the rare interval hypothesis, with tonal judgements more consistent for trichords containing rare interval information than trichords in which the interval information was consistent with multiple keys. Similarly, multiple studies have demonstrated the importance of distributional information in key-finding (e.g., Schmuckler and Tomovski 2005; Toiviainen and Krumhansl 2003). Schmuckler and Tomovski (2005), for example, found that this model predicted listeners' tonal percepts for a range of materials, including very short initial excerpts of Bach and Chopin preludes, increasingly longer segments of Chopin preludes, and an entire Chopin prelude. In all of these cases, this approach modelled tonal percepts in terms of sensitivity to picking up the intended key of the passage, as well as in predicting errors in key-finding.

The simultaneous success of these different models raises the question of the relation between these approaches. Typically, these models have been viewed as fundamentally inconsistent with one another (Butler 1989; Krumhansl 1990b), although there have surprisingly been no attempts to explicitly compare these approaches. In one exception, reported by Schmuckler and Tomovski (2000), four variants of short musical excerpts were created that systematically manipulated the rare interval and distributional information available in these passages. Samples of these variants appear in Figure 9.1, and presents a series of passages in which distributional and rare interval information imply the same key (variant A), distributional but not rare interval information is specific to a given key (variant B), and distributional and rare interval information implicate different keys (variants C and D).

Using a probe tone and a tonic finding procedure, listeners' tonal percepts in response to these variants were assessed. Figure 9.2 summarizes the results from these studies, presenting the correlation between the averaged probe tone ratings and the C and G major tonal hierarchies and the percentage tonic choices for the target keys of C and G major (all data were ultimately transposed to these keys) for the variants. One clear finding is that tonal percepts matched the note distributions of the passages, a result seen most easily in variants C and D in which ratings and tonic choices aligned with predictions based on note distributions. There was, however, a subtle influence of rare interval information on tonal percepts. Specifically, Figure 9.2 also reveals a systematic weakening of tonal percepts when the two cues do not coincide. Thus, variant A, in which both sets of information indicated the same key, produced the strongest tonal percepts, followed by variant B, in which rare interval information was ambiguous, and finally variants C and D, in which the two sets of information diverged. Clearly the two types of information support each other in tonal perception.

Subsequent analyses support the idea that distributional and rare interval information converge in their tonal implications. For both processes to operate in musical contexts not built solely on the diatonic scale, it must be that the pattern of interval occurrences characterizing the diatonic set also represents the interval frequency of real musical contexts. In keeping with this idea, Table 9.1 presents an interval count for the four variants of the stimulus passages of this study, and reveals that the diatonic interval vector closely matches the interval content of these musical passages. Moreover, in a simulation study, 50 40-note melodies were randomly generated using note frequencies that either matched the tonal hierarchy, or were randomly organized (see Smith and Schmuckler 2004). Table 9.1 also presents the interval counts from this simulation, and demonstrates that randomly generated melodies containing note distributions similar to that found in tonal music also produce interval sets matching the

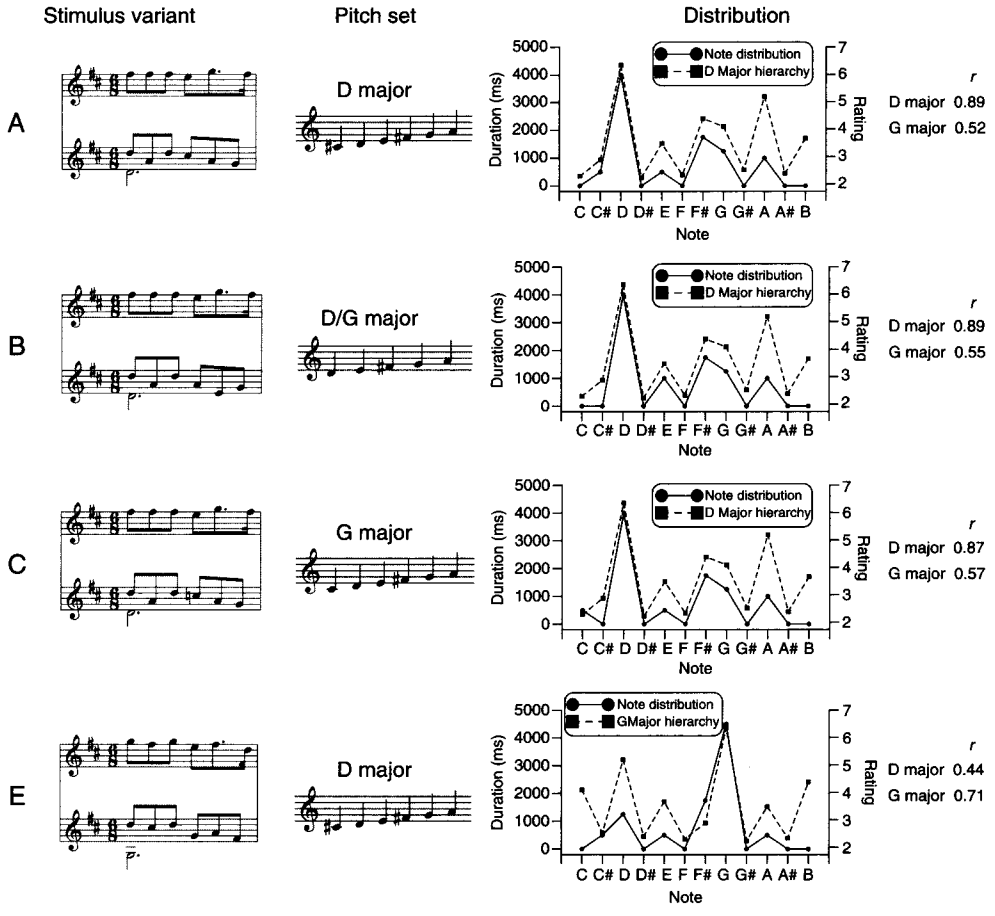



**Fig. 9.1** Sample stimuli from Schmuckler and Tomovski (2000). The four variants correspond to passages in which both the pitch set and the note duration distribution strongly indicate the same key (variant A), the note distribution but not the pitch set indicates a single key (variant B), and the pitch set and note distributions indicate different keys (variants C and D).

diatonic interval vector. Together, these findings suggest that distributional and structural–functional approaches are really part of the same process, and that a complete description of tonality and key-finding will require both global distributional properties, along with more localized, structural–functional information, an idea that has been recognized by other researchers (Temperley and Marvin 2008).

## Models of contour

In contrast to tonality, models of melodic contour have received less attention over the years. Of the contour models that do exist, it is also (roughly) possible to divide them into two categories, the first based on local parameters of contour and the second on global contour information. In the first group are models focusing on the individual interval content of contour (Friedmann 1985, 1987; Marvin 1991, 1995; Marvin and Laprade 1987; Morris 1993; Quinn 1999), whereas the second group consists of procedures for characterizing the rises and falls within the contour, primarily through the use of time series analyes (Eerola and Bregman 2007; Eerola *et al.* 2001, 2006; Schmuckler 1999).

Again, both types of models have generated supporting experimental results. Quinn (1999),

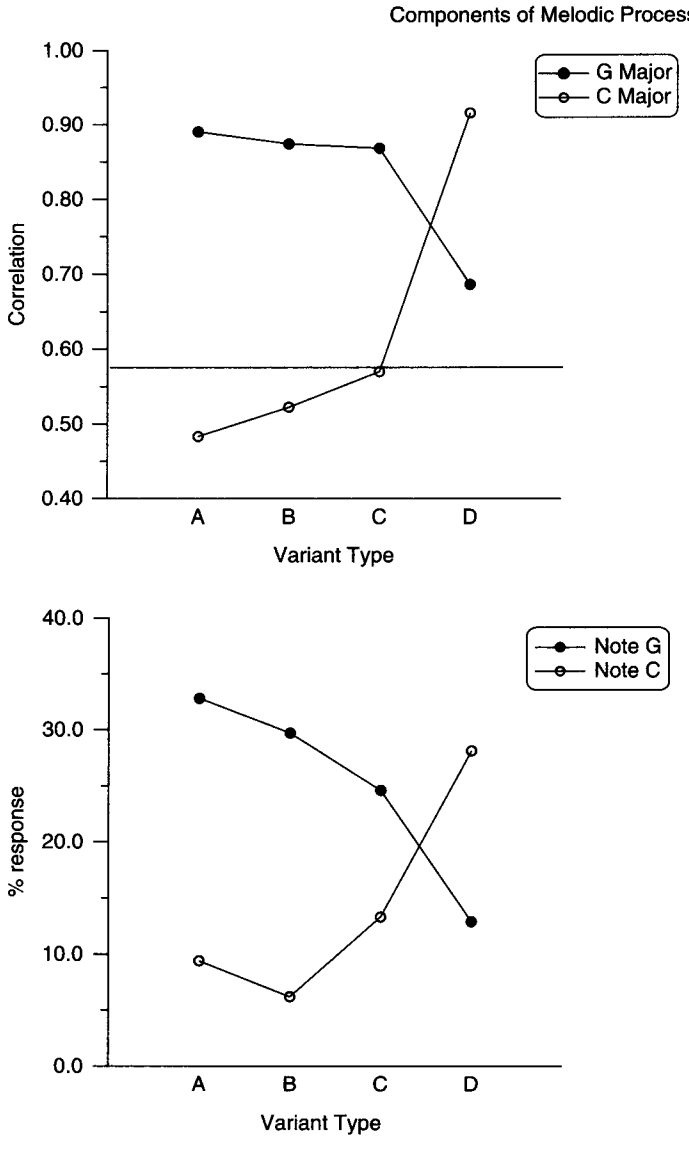



**Fig. 9.2** Results from the two experiments of Schmuckler and Tomovski (2000). The top figure graphs the correlation between the averaged probe tone ratings and the tonal hierarchies for the target keys of G and C major across the four stimulus variants. The bottom figure graphs the percentage of times the notes G and C were chosen as the tonic tone, as a function of the four stimulus variants.

for instance, demonstrated that the perceived similarity of pairs of short melodies was predictable from comparability in the rise/fall relation between adjacent notes, and from non-adjacent notes separated by a single tone. Schmuckler (1999), in contrast, demonstrated that perceived contour similarity for pairs of melodies was predictable from the degree of overlap of the cyclic information of these contours, as quantified by a Fourier analysis of the contour code.

**Table 9.1** The interval content of the diatonic set (Brown and Butler 1981), the stimulus variants of Schmuckler and Tomovski (2000), and of the simulation study using hierarchical and random organizations of pitch information (Smith and Schmuckler 2004). The final column shows the correlation between the interval vector of the diatonic set and the remaining interval vectors

| Source | Interval vector | | | | | | Correlation |
|---|---|---|---|---|---|---|---|
| | m2 | M2 | m3 | M3 | P4 | TT | |
| Diatonic set | 2 | 5 | 4 | 3 | 6 | 1 | – |
| Variant A | 212 | 352 | 303 | 301 | 534 | 58 | $r(4) = 0.95^{**}$ |
| Variant B | 140 | 378 | 325 | 253 | 521 | 0 | $r(4) = 0.99^{**}$ |
| Variant C | 145 | 341 | 328 | 297 | 508 | 36 | $r(4) = 0.96^{**}$ |
| Variant D | 195 | 296 | 244 | 281 | 567 | 57 | $r(4) = 0.90^{**}$ |
| Hierarchical | 98.1 | 131.9 | 120.4 | 114.7 | 170.0 | 43.0 | $r(4) = 0.95^{**}$ |
| Random | 137.6 | 128.7 | 137.7 | 111.4 | 114.2 | 50.1 | $r(4) = 0.52$ |

** $p<.01$

Subsequent work (Schmuckler (2004 submitted) extended these findings to rhythmically complex melodies, using both direct and indirect similarity measures. Together, these findings suggest that melodic contour can be reliably characterized by both the individual interval relations within the contour as well as by more global contour parameters, although it is critical to recognize that these two approaches are not mutually exclusive. Schmuckler (2004) has hypothesized that the former approaches might be most applicable to short melodic contours, whereas the latter approach could be more powerful when applied to longer melodies.

## Tonality and contour in melody perception

One advantage to the availability of tonal and contour models is that they provide a means for predicting aspects of melodic processing such as the perceptual organization of melodic form and melodic memory. The remainder of this chapter focuses on the role of contour and tonality in these aspects of melodic processing.

### Tonality and contour in perceptual organization

Given the interest in tonality and contour, it is surprising that research on the role of these factors in the perceptual organization of melodies per se has been relatively neglected over the years. Of the two, research on perceptual organization has looked more at the role of contour, and has (not surprisingly) found that this factor influences perceptual organization. Probably the most obvious aspect of perceptual organization in which contour has been examined is in stream segregation (Bregman 1990, 2005; Miller and Heise 1950; Van Noorden 1975). Unfortunately, a thorough review of this research is outside the purview of this chapter (see Bregman 1990, 2005; Carlyon 2004; Snyder and Alain 2007, for reviews). Even a cursory look at this topic, however, demonstrates that of the myriad of factors underlying stream segregation, one of the most important involves pitch proximity, and more critically in the current context, that the contour of pitch changes can drive stream segregation. For instance, Bregman and Dannenbring (1973) demonstrated that frequency glides, or tones with a continuous pitch contour, were grouped together into a single auditory stream, as opposed to discrete tones varying in pitch that grouped into separate streams; similar arguments have been advanced for pitch contour in the grouping of speech sounds (Darwin and Bethell-Fox 1977; Remez *et al.*1994).

Two well-known auditory perception illusions—Deutsch's (1975) scale illusion (Butler 1979; Deutsch 1975; Radvanysky *et al.* 1992), and the continuity illusion (Grossberg *et al.* 2004; Miller and Lickliter 1950; Tougas and Bregman 1990)—provide another example of contour influencing perceptual organization. In the scale

illusion, listeners hear simultaneously sounded ascending and descending diatonic scales, with the tones of these scales alternated between the ears. Rather than hearing this pattern organized based on source localization (by ear) or as a prototypical musical pattern (the diatonic scale), listeners hear two simple melodies, the first consisting of an ascending and descending line and the second consisting of a descending and ascending line. In the continuity illusion, a sound of long duration that is interrupted by a sound of short duration is heard as continuing through the noise. However, when the continuous information involves a set of competing glides, listeners sometimes hear a pair of bouncing pitch glides, comparable to the percept in the scale illusion (Grossberg *et al.* 2004; Tougas and Bregman 1990).

In both cases, these percepts are typically ascribed to perceptual organization by pitch proximity, and possibly good continuation. An alternative, however, is that listeners are hearing two simple melodic contours, ones that, according to the Fourier analysis model, contain a large amount of power in the fewest number of cyclic frequency components. This observation suggests that the role of contour in perceptual organization might be more quantitatively predictable, with simplicity of the amplitude spectrum driving a specific organization. If true, this would provide an interesting formalization of the Gestalt law of *prägnanz* (Koffka 1935), or the idea that good patterns have few alternatives (Garner 1970).

Somewhat surprisingly, there are few studies demonstrating an explicit role for tonality on perceptual organization. One line of work related to this question involves research on the perception of polytonality (Krumhansl and Schmuckler 1986b; Thompson and Mor 1992). Krumhansl and Schmuckler (1986b), for instance, examined tonal percepts in Stravinsky's *Petroushka* (which instantiates the two tonal centres of C and F# major) using both selective and divided attention tasks, and demonstrated that listeners were unable to form two independent tonal percepts. Instead, percepts of this passage were a complex amalgamation of the two tonalities, similar in structure to a music-theoretic proposal by Van den Toorn (1983) based on the octatonic collection. Thompson and Mor (1992), in contrast, did find that listeners perceived both keys of bitonal passages by Dubois and Milhaud, with the weights of these tonalities consistent with their relative importance in the musical score. Irrespective of an explanation for these divergent findings (which likely lies in the specifics of the musical contexts employed), this work does imply (indirectly) that tonal information can be a means of perceptually organizing melodic passages.

In a more explicit look at the role of tonality on perceptual organization, Dowling *et al.* (1987) had listeners identify nursery rhymes in which the melody notes were interspersed with distractor tones varying in their pitch overlap, temporal organization, and tonality, with the target. Interestingly, although pitch overlap and temporal organization influenced target identification, the tonal relation between target and distractors did not impact performance. In contrast, Bigand *et al.* (2000) had listeners attend to one of two French nursery rhythms, presented an octave apart, with these melodies played in the same key, a related key, or a distant key. These authors observed that varying tonal relations had no effect on detection of changes to the melodies (hits), although increasing tonal distance did lead to increased false alarms. Together, this work presents a mixed pattern, with some results consistent with the notion that tonality influences the perceptual organization of melodies, whereas other results indicated no such effect.

Recently, Vuvan and Schmuckler (2007) have further explored this question, employing the interleaved melody task of Bey and McAdams (2002) in which listeners hear a target melody followed by a pair of interleaved melodies, and must determine whether one of the interleaved melodies was the same as the target melody. In this study, the interleaved melodies varied in their average pitch spread (6, 9, 12, and 15 semitones), their inter-onset interval between sequential tones (80, 100, 120, and 140 msec), and in their tonalities (same, related, and unrelated; see Figure 9.3 for sample stimuli).

Figure 9.4 shows the averaged d' and bias scores for same/different judgment as a function of pitch spread and tonality (inter-onset interval had no effect in this study). Analyses of discrimination scores revealed a surprising pattern of

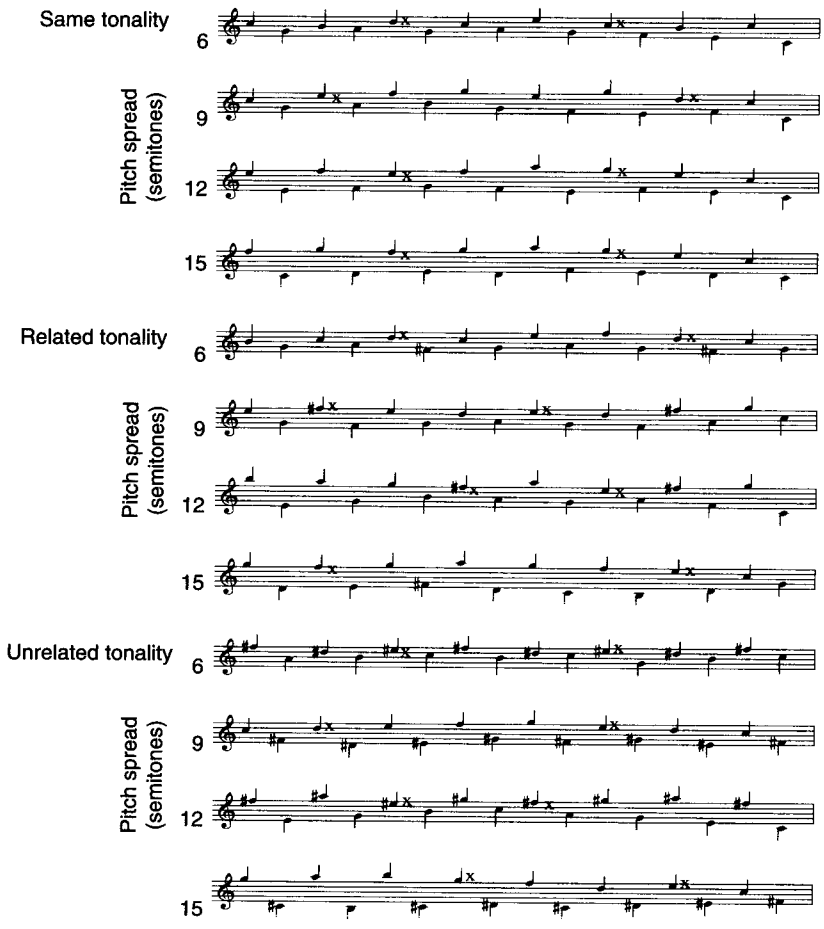



**Fig. 9.3** Sample stimuli from Vuvan and Schmuckler (2007). The two interleaved comparison melodies could be in either the same key (e.g., C and C major), a related key (e.g., C and G major), or an unrelated key (e.g., C and F# major), and could be separated by an average of 6, 9, 12, or 15 semitones. The location and pitch of the changed notes in the target melody is notated by an 'X' adjacent to the note in the musical staff.

decreasing discrimination with increasing pitch spread, although this effect was likely due to a drop in performance at the largest pitch spread. More importantly, analyses of discrimination and bias revealed the worst performance for interleaved melodies having unrelated tonalities, followed by melodies with related tonalities, and finally by melodies with the same tonality. Although counterintuitive, this finding does converge with Krumhansl and Schmuckler (1986b) and Bigand *et al.* (2000), who both observed that divergent tonalities negatively impacted the formation of independent streams.

## Tonality and contour in musical memory

Probably the most central arena in which tonality and contour have been recognized as critical factors in melodic processing involves memory for music. Unfortunately, a comprehensive review of work on this question cannot be presented here. Nevertheless, it is instructive to discuss, even if only superficially, some of the evidence demonstrating the importance of these factors in memory for music.

Research on the role of melodic contour in musical memory has a long history in music

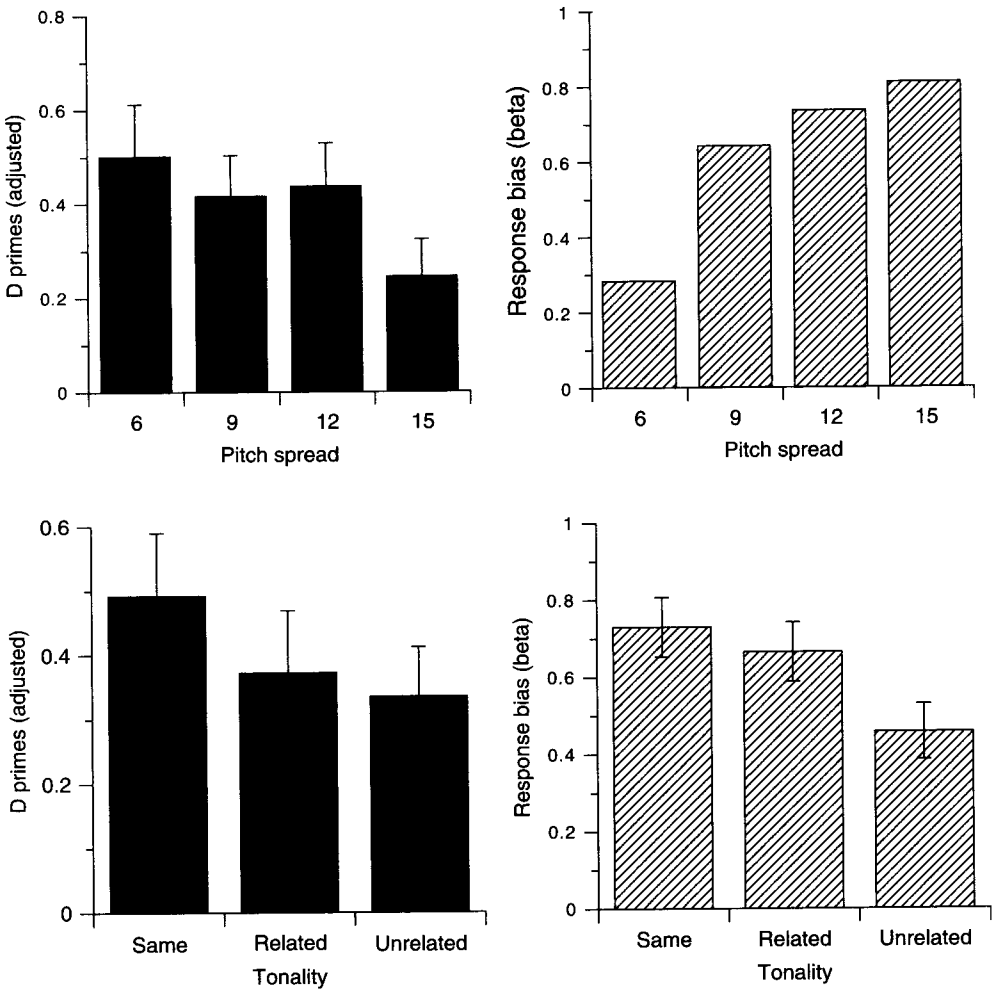



**Fig. 9.4** D' and bias for same/different target-comparison judgements as a function of pitch spread (top) and tonality (bottom) in Vuvan and Schmuckler (2007).

cognition. In early classic work on this topic, Dowling and colleagues (Dowling 1971, 1972; Dowling and Fujitani 1971) demonstrated that listeners' short-term melodic memory was dramatically influenced by the melody's contour, with the ability to distinguish between a standard melody and a comparison melody transposed to a different starting pitch more difficult when the melodies shared a comparable melodic contour. Similarly, Dowling and Hollombe (1977) found that recognition of octave-scrambled familiar melodies (hence melodies available in long-term memory) was significantly enhanced when these melodies retained their original melodic contour, relative to versions that violated this contour. These findings, and others like them, demonstrate that short-term melody recognition is strongly influenced by contour information (Croonen and Kop 1989; Cuddy *et al.* 1981; DeWitt and Crowder 1986; Dowling 1991; Dyson and Watkins 1984; Freedman 1999; Idson and Massaro 1978; Massaro *et al.* 1980; Watkins 1985).

There are, however, important limitations to the use of contour in melody recognition. Edworthy (1982, 1985), for instance, demonstrated that

contour information becomes less precise in memory as melody length increases, a finding attributed to the growth of tonal information in memory. Moreover, multiple studies have demonstrated that contour information dominates memory following short delays, but decreases in potency after longer delays (Dowling 1991; Dowling *et al.* 1995). Dowling (1991), for instance, found that listeners frequently confused same contour, tonal melodies after short delays (e.g., 11 s) but not after long delays (e.g., 39 s). Most importantly, and as summarized by Dowling (1994), it is clear that although contour is crucial for melodic memory, it does not function independently, but rather, it is integrated with other key melodic features, such as tonality.

The evidence for tonal influences on melodic memory is also robust. One way in which tonality influences memory, and an effect comparable to the research just described, is that tonally strong melodies are better remembered than tonally weak or atonal melodies (Boltz 1991; Croonen 1994; Dowling 1991; Dowling *et al.* 1995; Halpern *et al.* 1998; Watkins 1985). Croonen (1994), for instance, demonstrated that listeners' accuracy in a standard-comparison melody task was significantly influenced by the tonality of the stimuli, with strongly tonal sequences better remembered than weakly tonal melodies.

A second, albeit less researched finding, involves the presence of key-distance effects in memory (Bartlett and Dowling 1980; Bharucha and Krumhansl 1983; Krumhansl *et al.* 1982; Krumhansl and Castellano 1983). Bartlett and Dowling (1980), for instance, had listeners' recognize comparison melodies that were exact transpositions of standard melodies from non-exact transpositions when the comparisons were in the same key, a related key, or an unrelated key. These researchers observed that it was harder to distinguish non-transposition comparisons from standards (i.e., higher false alarm rates) when the melodies were in the same or related keys, as opposed to unrelated keys. Interestingly, this effect did not occur for transpositions, suggesting that memory for melodies contains relatively separable stores for tonal information versus interval information.

Other evidence exists for key distance effects in musical memory, although not explicitly for melodic materials (Bharucha and Krumhansl 1983; Krumhansl *et al.* 1982; Krumhansl and Castellano 1983). Krumhansl *et al.* (1982), for instance, found that recognition memory for chords was influenced by the tonal context in which these chords were embedded, such that the key distance of the tonal context and the key of the to-be-remembered chords predicted the probability that a given chord would be recognized, as well as memory confusion errors and asymmetries in confusion errors. Accordingly, there is robust evidence for key distance effects in memory, both for melodic and harmonic sequences.

## Conclusion

Melody has, over the years, been the focus of intensive investigation in research in music cognition. Somewhat ironically, despite these years of study our understanding of melodic structure can be described simply, with respect to two principal components—the underlying tonality of the melody, and the rising and falling pattern in pitch of the melody. The importance of these two components has, of course, been recognized since the beginning of the study of music in an experimental context, although it is only in recent years that formal models of these components have arisen. With the advent of these models, of course, has come the ability to make explicit predictions regarding how these components might predict listeners' melodic processing, including aspects such as perceptual organization and perceived similarity, as well as listeners' memory for melodic materials.

How might these models be applied to studying perceptual organization and melodic memory? One of the principal limitations of previous work examining tonality and contour in musical processing is that such research did not typically provide any formal characterization of the structure of these factors, and so could not make explicit predictions as to when such factors would play a role, or the relative strength of their role in such situations. For instance, contour relations between melodies were often characterized as being simply same or different, and tonality was typically determined on the basis of diatonicism of the melody.

Unfortunately, such characterizations ignore the idea that there are gradations to the degree

to which melodies are related. Contour similarity studies (Eerola and Bregman 2007; Eerola *et al.* 2006; Eerola *et al.* 2001; Quinn 1999; Schmuckler 1999), for instance, have all convincingly demonstrated that one can compare melodic contours on a continuous, as opposed to dichotomous (e.g., same versus different), basis. Similarly, work in tonal perception (e.g., Smith and Schmuckler 2004; Takeuchi 1994) demonstrates that tonality is not simply all-or-none, but can be characterized in terms of degrees of tonal strength, with some passage more tonal than other passages.

As for perceptual organization and memory, the utility of directly assessing the relatedness between stimuli in their contour and tonal structure is obvious. Perceptual organization, for instance, either in the stability of a specific organization, or the ease with which differing groupings can be formed, should be predictable by the level of similarity between competing organizations. In fact, there is already evidence for this finding, in terms of the previous discussed possible explanations for the scale illusion, or in explaining the discrepancy between Krumhansl and Schmuckler's (1986b) and Thompson and Mor's (1992) research on the perception of multiple tonalities.

Similarly, memory for melodies should be predictable based on the measured degree of contour and tonal similarity between melodies. In the typical standard-comparison memory task, for instance, greater contour or tonal similarity between melodies should be reflected in decreased accuracy overall, and possibly in false alarm rates in particular. Again, there are findings indicative of just such effects. Schmuckler (1997, for instance, observed that listeners' memory for melodies was influenced by the expectedness of the melody's ending (better memory for highly expected endings) and the tonal coherence of that ending (better memory for more tonally coherent endings). Although none of this work explicitly employs assessments of tonality and contour as predictors for memory and memory confusions, it is nevertheless consistent with the notion that such factors can be quantified to make such predictions. Clearly, then, further work investigating the utility of tonal and contour models in predicting perceptual organization and melodic memory has a great deal of potential in providing insight into our understanding of melodic processing quite generally.

## Acknowledgements

Preparation of this manuscript was supported by a grant from the Natural Sciences and Engineering Research Council of Canada to the author.

## References

Abe J and Hoshino E (1990). Schema driven properties in melody cognition: experiments on final tone extrapolation by music experts. *Psychomusicology*, **9**, 161–172.

Abe J and Okada A (2004). Integration of metrical and tonality organization in melody perception. *Japanese Psychological Research*, **46**, 298–307.

Bartlett JC and Dowling WJ (1980). Recognition of transposed melodies: a key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance*, **6**, 501–515.

Bey C and McAdams S (2002). Schema-based processing in auditory scene analysis. *Perception & Psychophysics*, **64**, 844–854.

Bharucha JJ and Krumhansl CL (1983). The representation of harmonic structure in music: hierarchies of stability as a function of context. *Cognition*, **13**, 63–102.

Bigand E, McAdams S and Forêt S (2000). Divided attention in music. *International Journal of Psychology*, **35**, 270–278.

Boltz M (1991). Some structural determinants of melody recall. *Memory & Cognition*, **19**, 239–251.

Bregman AS (1990). *Auditory scene analysis: the perceptual organization of sound*. The MIT Press, Cambridge, MA.

Bregman AS (2005). Auditory scene analysis and the role of phenomenology in experimental psychology. *Canadian Psychology*, **46**, 32–40.

Bregman AS and Dannenbring GL (1973). The effect of continuity on auditory stream segregation. *Perception & Psychophysics*, **13**, 308–312.

Brown H (1988). The interplay of set content and temporal context in a functional theory of tonality perception. *Music Perception*, **5**, 219–249.

Brown H and Butler D (1981). Diatonic trichords as minimal tonal cue-cells. *In Theory Only*, **5**, 37–55.

Brown H, Butler D and Jones MR (1994). Musical and temporal influences on key discovery. *Music Perception*, **11**, 371–407.

Butler D (1979). A further study of melodic channeling. *Perception & Psychophysics*, **25**, 264–268.

Butler D (1989). Describing the perception of tonality in music: a critique of the tonal hierarchy theory and a proposal for a theory of intervallic rivalry. *Music Perception*, **6**, 219–241.

Butler D (1990). A study of event hierarchies in tonal and post-tonal music. *Music Perception*, **18**, 4–17.

Butler D and Brown H (1984). Tonal structure versus function: studies of the recognition of harmonic motion. *Music Perception*, **2**, 6–24.

Butler D and Brown H (1994). Describing the mental representation of tonality in music. In R Aiello, JA Sloboda, eds, *Musical perceptions*, 191–212. Oxford University Press, London.

Carlyon RP (2004). How the brain separates sounds. *Trends in Cognitive Sciences*, **8**, 465–471.

Croonen WLM (1994). Effects of length, tonal structure, and contour in the recognition of tone series. *Perception & Psychophysics*, **55**, 623–632.

Croonen WLM and Kop PFM (1989). Tonality, tonal scheme and contour in delayed recognition of tone sequences. *Music Perception*, **7**, 49–38.

Cuddy LL, Cohen AJ and Mewhort DJK (1981). Perception of structure in short melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance*, **7**, 869–883.

Darwin CJ and Bethell-Fox CE (1977). Pitch continuity and speech source attribution. *Journal of Experimental Psychology: Human Perception and Performance*, **3**, 665–672.

Deutsch D (1975). Two-channel listening to musical scales. *Journal of the Acoustical Society of America*, **57**, 1156–1160.

DeWitt LA and Crowder RG (1986). Recognition of novel melodies after brief delays. *Music Perception*, **3**, 259–274.

Dowling WJ (1971). Recognition of inversions of melodies and melodic contours. *Perception & Psychophysics*, **9**, 348–349.

Dowling WJ (1972). Recognition of melodic transformations: inversion, retrograde, and retrograde inversion. *Perception & Psychophysics*, **12**, 417–421.

Dowling WJ (1978). Scale and contour: two components of a theory of memory for melodies. *Psychological Review*, **85**, 341–354.

Dowling WJ (1991). Tonal strength and melody recognition after long and short delays. *Perception & Psychophysics*, **50**, 305–313.

Dowling WJ (1994). Melodic contour in hearing and remembering melodies. In R Aiello, J Sloboda, eds, *musical perceptions* 173–190. Oxford University Press, Oxford.

Dowling WJ and Fujitani DS (1971). Contour, interval and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, **49**, 524–531.

Dowling WJ and Hollombe AW (1977). The perception of melodies distorted by splitting into octaves: effects of increasing proximity and melodic contour. *Perception & Psychophysics*, **21**, 60–64.

Dowling WJ, Kwak S and Andrews MW (1995). The time course of recognition of novel melodies. *Perception & Psychophysics*, **57**, 150–158.

Dowling WJ, Lung KMT and Herrbold S (1987). Aiming attention in pitch and time in the perception of interleaved melodies. *Perception & Psychophysics*, **41**, 642–656.

Dyson MC and Watkins AJ (1984). A figural approach to the role of melodic contour in melody recognition. *Perception & Psychophysics*, **35**(5), 477–488.

Edworthy J (1982). Pitch and contour in music processing. *Psychomusicology*, **2**, 44–46.

Edworthy J (1985). Interval and contour in melody processing. *Music Perception*, **2**(3), 375–388.

Eerola T and Bregman M (2007). Melodic and contextual similarity of folk song phrases. *Musicae Scientiae, Discussion Forum*, **4A**, 211–233.

Eerola T, Himberg T, Toiviainen P and Louhivuori J (2006). Perceived complexity of Western and African folk melodies by Western and African listeners. *Psychology of Music*, **34**, 337–371.

Eerola T, Järvinen T, Louhivuori J and Toiviainen P (2001). Statistical features and perceived similarity of folk melodies. *Music Perception*, **18**, 275–296.

Freedman EG (1999). The role of diatonicism in the abstraction and representation of contour and interval information. *Music Perception*, **16**(3), 365–387.

Friedmann ML (1985). A methodology for the discussion of contour: its application to Schoenberg's music. *Journal of Music Theory*, **29**, 223–248.

Friedmann ML (1987). My contour, their contour. *Journal of Music Theory*, **31**, 268–274.

Garner WR (1970). Good patterns have few alternatives. *American Scientist*, **58**, 34–42.

Grossberg S, Govindarajan KK, Wyse LL and Cohen MA (2004). Artstream: a neural net model of auditory scene analysis and source segregation. *Neural Networks*, **17**, 511–536.

Halpern AR, Bartlett JC and Dowling WJ (1998). Perception of mode, rhythm and contour in unfamiliar melodies: effects of age and experience. *Music Perception*, **15**(4), 335.

Holtzman SR (1977). A program for key determination. *Interface*, **6**, 29–56.

Huron D and Parncutt R (1993). An improved model of tonality perception incorporating pitch salience and echoic memory. *Psychomusicology*, **12**, 154–171.

Idson WL and Massaro DW (1978). A bidimensional model of pitch in the recognition of melodies. *Perception & Psychophysics*, **24**, 551–565.

Janata P, Birk JL, Van Horn JD, Leman M, Tillman B and Bharucha JJ (2002). The cortical topography of tonal structures underling western music. *Science*, **298**, 2167–2170.

Koelsch S and Siebel WA (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*, **9**(12), S78–S84.

Koffka K (1935). *Principles of gestalt psychology*. Harcourt, Brace and World, New York.

Krumhansl CL (1990a). *Cognitive foundations of musical pitch*. Oxford University Press, London.

Krumhansl CL (1990b). Tonal hierarchies and rare intervals in music cognition. *Music Perception*, **7**, 309–324.

Krumhansl CL (2000). Rhythm and pitch in music cognition. *Psychological Bulletin*, **126**, 159–179.

Krumhansl CL, Bharucha JJ and Castellano MA (1982). Key distance effects on perceived harmonic structure in music. *Perception & Psychophysics*, **32**, 96–108.

Krumhansl CL and Castellano MA (1983). Dynamic processes in music perception. *Memory & Cognition*, **11**, 325–334.

Krumhansl CL and Kessler EJ (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, **89**, 334–368.

Krumhansl CL and Schmuckler MA (1986a). *Key-finding in music: an algorithm based on pattern matching to tonal hierarchies*. Paper presented at the 19th annual Mathematical Psychology Meeting, Cambridge, MA.

Krumhansl CL and Schmuckler MA (1986b). The petroushka chord: a perceptual investigation. *Music Perception*, **4**, 153–184.

Lerdahl F (2001). *Tonal pitch space*. Oxford University Press, Oxford.

Longuet-Higgins HC and Steedman MJ (1971). On interpreting Bach. *Machine Intelligence*, **6**, 221–241.

Marvin EW (1991). The perception of rhythm in non-tonal music: rhythmic contours in the music of Edgard Varèse. *Music Theory Spectrum*, **13**, 61–78.

Marvin EW (1995). A generalization of contour theory to diverse musical spaces: analytic applications to the music of dallapiccola and stockhausen. In EW Marvin, R Hermann, eds, *Concert music, rock and jazz since 1945: essays and analytic studies*, 135–171. University of Rochester Press, Rochester, NY.

Marvin EW and Laprade PA (1987). Relating musical contours: extensions of a theory for contour. *Journal of Music Theory*, **31**, 225–267.

Massaro DW, Kallman HJ and Kelly JL (1980). The role of tone height, melodic contour, and tone chroma in melody recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **6**, 77–90.

Matsunaga R and Abe J (2005). Cues for key perception of a melody: pitch set along? *Music Perception*, **23**, 153–164.

Miller GA and Heise GA (1950). The trill threshold. *Journal of the Acoustical Society of America*, **22**, 167–173.

Miller GA and Lickliter JCR (1950). Intelligibility of interrupted speech. *Journal of the Acoustical Society of America*, **22**, 167–173.

Morris RD (1993). New directions in the theory and analysis of musical contour. *Music Theory Spectrum*, **15**, 205–228.

Nakata T and Trehub SE (2004). Infants' responsiveness to maternal speech and singing. *Infant Behavior and Development*, **27**, 455–464.

Quinn I (1999). The combinatorial model of pitch contour. *Music Perception*, **16**, 439–456.

Radvanysky GA, Hartmann WM and Rakerd B (1992). Structural alternations of an ambiguous musical figure: the scale illusion revisited. *Perception & Psychophysics*, **52**, 256–262.

Remez RE, Rubin PE, Berns SM, Pardo JS and Lang JM (1994). On the perceptual organization of speech. *Psychological Review*, **101**, 129–156.

Rock AML,Trainor LJ and Addison TL (1999). Distinctive messages in infant-directed lullabies and play songs. *Developmental Psychology*, **35**, 527–534.

Schmuckler MA (1997). Expectancy effects in memory for melodies. *Canadian Journal of Experimental Psychology*, **51**(4), 292–305.

Schmuckler MA (1999). Testing models of melodic contour similarity. *Music Perception*, **16**, 295–326.

Schmuckler MA (2004). Pitch and pitch structures. In J Neuhoff, ed., *Ecological psychoacoustics*, 271–315. Academic Press, San Diego, CA.

Schmuckler MA (submitted). Melodic contour similarity using folk melodies.

Schmuckler MA and Tomovski R (2000). *Tonal hierarchies and intervallic rivalries in musical key-finding*. Paper presented at the Society for Music Perception and Cognition, Toronto, ON.

Schmuckler MA and Tomovski R (2005). Perceptual tests of an algorithm for musical key-finding. *Journal of Experimental Psychology: Human Perception and Performance*, **31**, 1124–1149.

Shenfield T, Trehub SE and Nakata T (2003). Maternal singing modulates infants arousal. *Psychology of Music*, **31**, 365–375.

Shepard RN (1982). Geometrical approximations to the structure of musical pitch. *Psychological Review*, **89**, 305–333.

Smith NA and Schmuckler MA (2004). The perception of tonal structure through the differentiation and organization of pitches. *Journal of Experimental Psychology: Human Perception and Performance*, **30**, 268–286.

Snyder JS and Alain C (2007). Toward a neurophysiological theory of auditory stream segregation. *Psychological Bulletin*, **133**, 780–799.

Takeuchi A (1994). Maximum key-profile correlation (mkc) as a measure of tonal structure in music. *Perception & Psychophysics*, **56**, 335–346.

Temperley D (1999). What's key for key? The Krumhansl–Schmuckler key-finding algorithm reconsidered. *Music Perception*, **17**, 65–100.

Temperley D (2001). *The cognition of basic musical structures*. The MIT Press, Cambridge, MA.

Temperley D (2002). Bayesian models of musical structure and cognition. *Musicae Scientiae*, **8**, 1061–1069.

Temperley D (2007). *Music and probability*. MIT Press, Cambridge, MA.

Temperley D and Marvin EW (2008). Pitch class distribution and the identification of key. *Music Perception*, **25**, 192–212.

Thompson WF and Mor S (1992). A perceptual investigation of polytonality. *Psychological Research*, **54**, 60–71.

Tillman B, Bharucha JJ and Bigand E (2000). Implicit learning of tonality: a self-organizing approach. *Psychological Review*, **107**, 885–913.

Toiviainen P and Krumhansl CL (2003). Measuring and modeling real-time responses to music: the dynamics of tonality induction. *Perception*, **32**, 741–766.

Tougas Y and Bregman AS (1990). Auditory streaming and the continuity illusion. *Perception & Psychophysics*, **47**, 121–126.

Unyk AM, Trehub SE, Trainor LJ and Schellenberg EG (1992). Lullabies and simplicity: a cross-cultural perspective. *Psychology of Music*, **20**, 15–28.

Van den Toorn PC (1983). *The music of Igor Stravinsky*. Yale University Press, New Haven, CT.

Van Noorden LPAS (1975). Temporal coherence in the perception of tone sequences. University of Technology, Eindhoven.

Vuvan D and Schmuckler MA (2007). *Tonal influences on auditory stream segregation*. Poster presented at the Society for Music Perception and Cognition. Montreal, QC.

Watkins AJ (1985). Scale, key, and contour in the discrimnation of tuned and mistuned approximations to melody. *Perception & Psychophysics*, **37**, 275–285.

Yoshino I and Abe J (2004). Cognitive modeling of key intepretation in melodic perception. *Japanese Psychological Research*, **46**, 283–297.

# CHAPTER 10

# Memory for music

Bob Snyder

## Introduction

This chapter will focus on musical memory: on how adult listeners form mental representations of music, and on how those representations affect music listening. The chapter will be divided into three sections, which will outline general concepts of memory, address research and theory about musical expectations, and consider listeners' ability to remember various aspects of music. The study of memory in music listening is relatively recent, although the scientific study of human memory dates back to the end of the nineteenth century (Ebbinghaus 1885/1964). The ways in which mind and brain are used in the perception and comprehension of music appear to be much in line with the way they are used in processing in other domains, although there are theories proposing that music is partly modularized in cognition and brain organization (Peretz and Zatorre 2005).

## Memory in general

Let us start by defining some of the basic constructs that have been developed in the study of memory. Human memory is said to be *encoded*; events in the world (including events in the body and brain) may cause changes in the detailed micro-structure of the brain that persist over varying amounts of time. These changes are thought to take the form of differences in the strength of connections at gaps (*synapses*) between nerve cells (*neurons*). These connections regulate the flow of electrical charges between neurons. The term *encoding* indicates that these changes assume a form that is different from the things or events that cause them. This is true not only in the sense that neural connections arising through exposure to events in the real world are not the same as those events, but also in the sense that events are encoded in a previously existing context of their meaning to a particular person. In order to persist, a memory must, to some extent, be related to what an individual already knows.

Memory has been defined in different ways that refer to different *modes of functioning*, generally in relation to the time-spans over which it persists. Following research into very brief (250 msec) memory for visual images that yielded the concept of *iconic* memory (Sperling 1960), an analogous concept of a very brief auditory or 'echoic' memory was proposed (Crowder and Morton 1969). Echoic memory is posited as a very brief *sensory* image of an auditory stimulus that persists for a second or two at most; however, the distinction between echoic and auditory short-term memory is not always clear. (Crowder 1993).

A construct with a much more extensive history is that of short-term memory (STM). STM is said to exist on a time scale of seconds, ranging from approximately 4–30 seconds, though it is usually on the order of 4–8 seconds. STM has been referred to as the 'specious present' (James, 1890), and is the window of time within which the present moment of perception and thought are grasped. In addition to this *temporal* limit, STM may also have a *capacity* limit (Cowen 2005) on the actual number of items that can persist as STM at one time. This number was originally proposed by Miller (1956) as $7 \pm 2$, but has recently been scaled down to a usual maximum of 4 or less (Cowen 2005), although capacity can be increased by practice (Ericsson and Kintsch 1995). This is a very small number which seems to be contradicted by everyday experience.

Part of the explanation for this lies in the *associative* nature of human *long-term* memory; memories for items and events contiguous in space or time may become connected, and the occurrence or recall of one such item may *cue* the recall of an associated item.

The idea that memory is associative has a long provenance in western thought, from at least the mid-eighteenth century. Associations are now thought to consist of connections between networks of neurons in the brain (Hebb 1949), in that a group of items related by association can form a single consolidated item in long-term memory. This larger item can then be recalled (activated as STM) in its entirety, a process that can occur on multiple levels. In this way a kind of hierarchical compression occurs, which allows the most efficient use of the limited capacity of conscious STM; George Miller coined the term *chunking* for this process. A memory chunk is a group of 3–4 items related by association; a musical grouping consisting of 3–5 notes would be a chunk, and a phrase consisting of several of these groupings would be a higher-level chunk. It should be noted that the limitations of chunking are implicated in both the recall and the formation of memories. Chunks or prominent items in chunks can be used to cue other chunks, allowing for the recall of longer sequences. Memorability appears related to how 'chunkable' a sequence is, which will depend on the amount of repetition of items, and to boundaries formed by discontinuities or changes in a sequence; the more clearly a sequence can be subdivided in units of chunkable size, the more likely it is to be recalled. Chunking in music is often related to perceptual grouping processes (Bregman 1990; Lerdahl and Jackendoff 1983). The idea of chunking accounts for a great deal of how immediate human memory can have such a limited capacity, yet be so effective.

Everyday use of the term 'memory' refers not to short-term memory, but to *Long-Term memory* (LTM). LTM differs from STM in several ways. Long-term memories persist over a much longer time-period, even a lifetime, as their formation involves lasting structural change in the brain. Moreover, unlike STM, almost all the contents of long-term memory are not conscious at any given moment, although situational associations may cue memories into consciousness. There are distinctions within LTM that must be further considered. First, there is a distinction introduced by Tulving (Tulving 1972) between so-called *episodic* and *semantic* long-term memories: a distinction between memories of situations and events (*episodic* or autobiographical memory), and general knowledge of facts (*semantic* memory), the latter being more abstract and arising through the experience of repeated similar episodes (Schank 1982). Episodic memory is dynamic, and may be changed by the very act of remembering (Roediger *et al.* 2007, Chapter 8). These two types of memory appear to constitute two ends of a continuum extending from recollections of particular episodes through increasingly general models of types of situations to abstract knowledge representations that no longer have roots in any particular experiences. In music cognition research, *episodic* memory is typically tested in recognition or, more rarely, recall tasks. *Semantic* memory tests usually involve requiring participants to make judgments about types of events they judge likely in particular musical situations. This type of semantic memory is referred to as a *schema*, a general expectation about types and distributions of events (Bartlett 1932/1995; also see Gjerdingen 1988, pp. 3–10). Schemas appear to underlie structural regularities in music such as tonality and metre, as well as standardized musical forms.

A further distinction important for LTM is that between *explicit* and *implicit* memory (Cohen and Squire 1980), a distinction between memories that may be accessible to consciousness and those that are not. One of the major discoveries of the last one hundred fifty or so years about the human mind is that much of the activity of the brain and the mind is not available to awareness (Ellenberger 1970) but remains implicit, evidenced particularly in the acquisition of skills. Many kinds of physical skills, such as knowing how to ride a bicycle or produce a clear tone on a musical instrument, require memory for their execution, but those memories are not available to consciousness, and cannot be described verbally. More recently, another form of implicit long-term memory has been proposed—implicit *perceptual* representations (Mandler 2004). This form of memory is involved

in unconscious statistical learning (Reber 1993) that keeps a record of regularities in the environment, and structures unconscious expectations about environmental events. Schemas are describable in terms of this kind of implicit memory, being involved in, e.g., the unconscious generation of expectations about musical events as a piece unfolds. Making such schemas *explicit* is one of the goals of the formal study of music. Implicit perceptual memory is the basis for much of *recognition* memory, though many types of activities involving memory may have both explicit and implicit components which are not easy to tease apart in experiment.

From a biological point of view, it seems that the primary function of LTM is not to construct highly detailed and accurate representations of the past, but to provide a generalized model of the world that will be useful to guide future behavior. Long-term schematic memories tend to be structured in terms of generalized *categories* (McAdams 1989; Snyder 2000). Musical long-term memories and expectations are thus often structured in terms of general categorical schemas such scale-steps, durational categories, etc. This means that listeners tend to not have exact detailed memories of music, but more generalized memories about the *kinds* of events that were heard; hence a listener's repertoire of categories of musical events (knowledge in memory) will affect what they can and do remember.

Short-term and long-term memory appear qualitatively different, yet there must be some mechanisms that allows the integration of information across both domains of memory. The idea of STM has thus been augmented by a more complex notion of 'working memory' (Baddeley 1986). Developing research on STM has revealed that most uses of this type of memory draw on long-term memory in various ways (e.g. Cowen 2005). Generally when we use STM, the contents of awareness draw on already existing knowledge, and this knowledge is in the form of long-term memories. In the most prevalent current theory of working memory (Baddely 1986), such memory consists of several parts, which are thought of as separate memory systems. These are a 'visuo-spatial sketchpad' (used for visualization), a 'phonological loop' (involved in speech production), and a 'central executive'(which sequences thinking and planning). This last is perhaps the least well-understood of the constructs associated with the concept of working memory. The originator of the idea, Baddeley, has himself recently suggested a new component, the episodic buffer (Baddeley 2003). There are probably other forms of working memory as well, for instance for motor movements, non-speech sounds, etc. (Jonides and Smith 1997, pp. 263–265). Some of the aspects of pitch memory mentioned below would qualify as forms of working memory. It may well be that the construct of working memory is an umbrella for a number of distinct memory phenomena that all have limits in the time range of several seconds; indeed, it may simply be a particular kind of persistence in many different processing systems in the brain (Crowder 1993).

At any given time, most of long-term memory is not accessible to consciousness. In order to explain certain kinds of memory phenomena, it has been suggested that there are degrees of unconsciousness, expressed in the construct of *activation* (Cowen 1988). The neural networks that are the basis for long-term memory can be active to a greater or lesser degree; at any particular moment, most will be inactive; some will be in a state of activity but not involved in current consciousness, while some will be active *and* involved in consciousness, a situation that will change from moment to moment. Memory can be *semi-activated* (Chafe 1994); memories at this lower degree of activation may be involved in mental activity, but not in consciousness. Experiments have shown that prior exposure to certain kinds of stimuli can affect later performance on another test, even when a participant is unaware of having experienced the original stimulus (Bornstein 1989). It seems that memory networks established by the original stimulus have been *primed* into semi-activity, which, although remaining below the level of consciousness, can affect ongoing thought and perception, providing a basis for *expectation*. An expectation (which may be more, or less, conscious) is thought to be a group of networks (a *schema*) that have been *cued* and *primed* into semi-activity by current ongoing experience. The concept of expectation is important here because it is a primary mode in which listeners utilize memory in listening to music, and because it is thought to be one of the sources of emotional responses to music (Meyer 1956; Huron 2006).

## Musical expectations

Listeners acquire a variety of schematic expectations (manipulation of which may influence profoundly the affective experience of music) from repeated exposure to statistical relations between various elements in music. Bharucha (1994) makes a distinction between *schematic* expectations (based on a listener's knowledge in semantic memory of the style, genre, etc. of a heard piece) and *veridical* expectations (arising from repeated exposure to a particular piece of music, and hence a form of episodic memory); both types of expectation involve LTM. Huron (2006, p. 228) adds the notion of *dynamic* expectations, arising from local events in the ongoing music in the short-term, which would involve STM (this type of expectation would arise for instance when hearing a piece for the first time). Musical knowledge and/or training would affect all of these kinds of expectation.

In the domain of pitch, most musical cultures use *scales*, which are ways of dividing pitch space into a relatively small number of *degrees* or categories. It has been suggested that one motivation for this is the general limitations of STM mentioned above. In fact, most scales in use in musical cultures worldwide consist of 5–7 different pitches or scale degrees, a number of elements that is consistent with the suggestions about STM capacity made above (Dowling 1978). Within the Western musical traditions, a considerable amount of research has been conducted on listeners' expectations concerning scale degrees. Listeners are frequently able to infer the 'key' or tonal center (central scale degree) of music. This is thought to arise through learning based on the statistical distribution of the occurrence of particular scale-degrees in music (that listeners are able to learn the statistical relations between melodic pitches—even in a relatively short time—has been established in a series of experiments by Saffran *et al.* 1999). Building on classic research by Krumhansl (Krumhansl 1990), Aarden established that listeners' expectations for the occurrence of a particular scale-degree do in fact conform to the distribution of scale degrees in actual music (Aarden 2003 cited in Huron 2006), expectations varying from moment to moment as a piece unfolds.

In addition, listeners have some very general schematic expectations about types of *sequences* of melodic events. Huron and his colleagues have described four patterns of statistical learning in experienced listeners, resulting in several proposed types of very general melodic schemas (Huron 2006); several of these were previously hypothesized by Meyer (1956) and Narmour (1990). Listeners expect *pitch proximity*, expecting a pitched musical event to be followed by an event relatively close to it in pitch. This is thought to be statistical learning based on the fact that the majority of melodic intervals in actual music consist of pitches that are close (within a m3) together (Dowling and Harwood 1986, p. 155). Second, listeners expect *post-skip reversal*, the expectation that after a large melodic interval, there will be a change of direction of melodic motion (Meyer 1956, Narmour 1990). Third, musician listeners expect *step-inertia*, a tendency for a small pitch interval to be followed by another small pitch interval in the same direction (Meyer 1956, Narmour 1990). Fourth, listeners expect late phrase *pitch declination*—that a phrase will move down melodically at the end. Objectively, melodic phrases tend on average to have an arch-like shape, with a rise in pitch at the beginning, and a fall at the end. Interestingly, listeners seem to expect melodic phrases to fall at the end, but not to rise at the beginning.

Rhythmic expectations are rooted in the experience of *pulse* or *tactus*, a regular framework of time-points or *beats* that typically organize our perception of the timing of events in music. Beats will more-or-less coincide with events in the music, but are cognitive entities inferred from those events. Pulse can be defined as a kind of basic expectation about when events in music will happen (see chapter 8). The range of musically usable tempi within which musical pulse is experienced is from about 30 beats per minute (bpm) to about 300 bpm. A listener will typically infer a *basic tempo* that falls within a range of approx. 60–150 bpm (Fraisse 1982), the range within which participants are best able to maintain a steady beat in tapping experiments.

Current theories of pulse suggest that it arises from patterns of oscillation set up in the brain that *entrain* to the pulse of a piece of music,

forming a dynamic memory within which unfolding temporal events are organized (Barnes and Jones 2000), and yielding a framework that guides immediate temporal expectations about when events are experienced as most likely to occur. This type of expectation is very local, existing on the time-scale of working memory. It is also robust, persisting in the face of small timing discrepancies or a temporary lack of confirming events in the music. Pulse may operate at multiple temporal levels that unfold at integer multiples and divisors of the basic pulse (Yeston 1976), but there is usually one level of pulse to which a listener will respond most strongly.

Much music, in addition to being organized around a regular pulse, is also organized at a higher time level into regular cycles of more, and less, accented beats (London 2004). This type of cyclical organization is referred to as *metre*; metrical organization is schematic, and generates expectations not only about when, but also what, kind of events will happen at what points in time. Like pulse, it is a cognitive structure that is inferred from actual events in music. Typical metrical cycles alternate between strong (accented) and weak beats. Metrical cycles in Western music exist within the span of STM, generally starting with a strong beat (downbeat), the most structurally important point in a metrical cycle. This is a point of metrical stability and closure and also the point where an event is most likely to occur (e.g. Huron, 2006, p. 178–9), being a focal point for temporal attending (Palmer and Krumhansl 1990; Jones *et al.* 2002). Just as the timing of events can deviate more or less from the expected beats of a pulse, the grouping of musical events can deviate more or less from the accent pattern of metre (Lerdahl and Jackendoff 1983). One typical way of creating a kind of metrical tension is by placing salient events, such as the first or last events of a grouping, on weak beats of the metre. Metrical organization can be ambiguous to a considerable degree, and many rhythmic sequences can be interpreted in several ways (Lee 1991).

## Music and memory

Experiments conducted by Deutsch (1999) suggest that auditory *working memory* may consist of aspects of musical sounds, outputs of subsystems being combined in later stages of processing. Individual musical pitches use at least one of these subsystems. Listeners presented with two pitches that were the same or a half-step apart accurately judged pitches as the same or different when the pitches were up to 6 seconds apart in time, but accuracy decreased when other pitches were interpolated in the empty time interval. This *interference* effect was strongest when the intervening pitches were chosen from both the octaves above and below the comparison pitches, making the intervening pitch intervals very large, the large pitch intervals rendering difficult the formation of a memory framework to anchor the comparison pitches. Only interpolated sequences varying in pitch led consistently to memory interference , suggesting that memory for pitch is encoded by a specialized memory subsystem (Deutsch 1999). A few individuals (fewer than 1 in 1000) have an ability to accurately identify a pitch outside of any particular context, an ability referred to as 'absolute pitch' (AP). The basis for AP possession is still debatable, but it seems to be learned, learning occurring within a 'critical period' of development and being dependent on early musical training (Ward 1999; also Huron 2006 pp. 110–113, for a discussion). AP possessors are able to effortlessly encode their experience of musical pitch in the form of a (conceptual) pitch name, which resists interference much better than a sensory memory (Siegal 1970; see also Deutsch 2006).

A familiar melody can be presented at almost any tempo and at any pitch level and remain recognizable; hence memory encoding of familiar melodies is not an exact (episodic) copy of particular pitches and time intervals, but a higher-order abstraction (schema) of particular features of the melody. In addition to some surface-level aspects of the music, possible features encoded in memory include *interval*, *contour*, and *scale-step context* (position in a scale). Which of these features are more salient in melody recognition (as opposed to recall) has been tested in several ways. One technique for separating aspects of a remembered melody is octave scrambling, which destroys musical interval structure, but preserves the scale-step order of the pitches in a melody (Deutsch 1972). When familiar melodies are presented with their

pitches randomly scrambled to different octaves, they are generally unrecognizable. However, if the contour (up and down pattern) of the melodic line is preserved, they are more frequently recognized (Dowling 1978).

Listeners are often unable to tell the difference between exact and contour-preserving transpositions of novel melodies (Dowling 1994), for both tonal and atonal melodies. However, the importance of contour versus interval seems to decline the more well-learned a melody is. This is in line with results of studies that have found that processes of memory formation in listening are highly dynamic, not just as a consequence of longer-term learning processes but also in the very process of real-time listening itself. As Dowling *et al.* (2002 p. 273) note, 'The processing of a phrase, once heard, continues automatically [in memory] even while the listener hears new phrases. What the listener can remember having heard continually changes during continued listening.' In a study of melodic recall (Sloboda 1985) it was concluded that recall of simple melodies 'was never note-for note perfect'. In recalling these melodies, participants produced small variations that were harmonically and metrically consistent with the originals (musically trained participants relying somewhat more than did untrained participants on harmonic structure). This study confirms the idea that episodic memory for melody typically consists of an underlying abstract schema in which not all of the surface detail is necessarily retained.

In the perception of melodic intervals, it seems that each culture creates a set of melodic intervals that become standard perceptual categories that detemine the tones occurring in its music (see Chapter 2). These intervals are categorical in that considerable leeway is allowed in what is considered an acceptable rendition of a particular interval. Exactly how much leeway is acceptable is itself a cultural variable (Arom *et al.* 1997), and perception of interval categories is much clearer to musically trained than to untrained listeners (Burns 1999).

Tonality—the organization of music around sets of pitches of which a central pitch is established by its frequent occurence and appearance at salient points—has significant ramifications for musical memory (see Chapter 6). Sets of pitches constitute scales; groups of pitch categories or scale 'steps' to which a passage or piece of music is limited. Both the central pitch and the steps of a scale have varying degrees of perceptual stability, serving as cognitive landmarks in establishing a framework for listeners (Krumhansl 1990). Knowledge of scale step categories is a kind of implicit memory acquired by members of a musical culture through exposure to its music (Burns 1999).

The above-mentioned stability may not be operative in the perception of music that does not conform to tonal principles; in the experience of atonal melodic sequences the anchoring effects of tonality cannot be operative, leading some authors to maintain that atonal music theory has little basis in human perception (Francès 1988; Lerdahl 1988). Experiments by Francès (Francès, 1988) using groups of musicians familiar, and unfamiliar, with atonal idioms, showed that both groups had great difficulty identifying varied repetitions of twelve-tone series, especially in an actual musical context. Further experiments (Krumhansl *et al.* 1987) indicate that the structures of twelve-tone music are slightly more transparent to musically trained listeners familiar with the style.

Rhythm can be defined as the articulation of events within the time limits of working memory. As noted above, to form a perceptible pattern the events that comprise a rhythm must seem connected together in the present. It is not yet entirely clear how rhythmic sequences are encoded in memory, various models proposing different levels of hierarchical coding. As mentioned above, listeners attempt to establish a framework for rhythms by inferring beats at a regular time interval from events in the music (Povel 1984). Most rhythm experiments involving memory are recall tests where listeners are asked to reproduce a heard rhythm. As with pitch, memory for duration appears to be a categorical phenomenon (Clarke 1987); out of many possible durations, memory seems to gravitate towards a few simple durational relationships, with considerable leeway as to what constitutes a particular duration. Foundational research by Fraisse (Fraisse 1982) established that when rhythmic patterns involve more than one duration, listeners tend to hear durations as either long or short, with long durations being

over 400–600 msec. In addition, there appears to be a strong tendency for the long and short durations to be reproduced or remembered in a ratio of 2:1, a relationship that has been described as a durational schema (Povel 1981).

In general, listeners are better able to reproduce sequences of durations when they can be interpreted as having an underlying regular sequence of beats, often involving the use of one particular duration as the basis for the others (Povel 1984). It is also the case that sequences of durations are much easier to remember and reproduce when shorter durations are integral subdivisions of the longer durations. Although the perception and reproduction of rhythms is categorical, small within-category deviations of event onset times (nuances) can give performed rhythms a particular feel (Snyder 2000), referred to as expressive timing (Gabrielsson 1999). This means that a particular rhythmic pattern can be performed many ways without losing its identity as a pattern, as in swing in jazz (Collier and Collier 2002). This type of nuanced information is generally not a part of the episodic memory of rhythm in music; although subtle, such nuances are important parts of what makes rhythms dynamic and emotionally charged.

## Long-term memory and musical form—hierarchical theories of long-term musical representation

Musical form can be defined as the timescale of musical phenomena that require some kind of long-term mental representation for their comprehension. The process of chunking can lead to hierarchical organization in LTM, and most theories of the structure of long-term representations of music use the concept of hierarchy to varying degrees, applied both to episodic memories and abstract schemas. Hierarchical levels in music may range from local groupings up to entire pieces.

When memory passes beyond the limits of STM, it is thought to persist in a more schematic form: to become a kind of reduction. Exactly what remains of actual musical details in musical LTM is an important question. Typically, theories of musical LTM reduction involve ideas of hierarchies of structural importance (Deutsch and Feroe 1981; Lerdahl and Jackendoff 1983), in which certain events in music are structurally more important than others, and these constitute the *gist* of a listener's memory representation (see Chapter 6). These salient structurally important events can be said to constitute the 'deep' level of structure in music, more rapidly changing details forming the musical surface or foreground.

The musical surface is segmented into units of cognitively manageable size (chunks)—the actual groupings, phrases, etc. of the music. Typically in tonal music the proposed deep structural events occur on metrically strong beats, are of longer duration, and are located on an important scale degree (Serafine *et al.* 1989), although the metaphor of hierarchical depth may be more applicable to some types of music than to others (Fink 1999). It should also be noted that although the metaphor of deep and surface structure is also used in generative linguistics, music is not held, in general, to embody specific and definite semantic meaning (though see Chapter 3). Current models of musical representation differ considerably as to the importance, and indeed the cognitive reality of deeper levels of structure.

Two prominent theories of musical mental representation (LTM) are Lerdahl and Jackendoff's (1983) Generative Theory of Tonal Music (GTTM) and the 'Cue Abstraction' theory of Deliège (Deliège and Melen 1997). The GTTM is a hierarchical theory of *tonal* music, and involves four different types of reduction: segmentation analysis, metrical analysis, time-span reduction, and prolongational reduction. The first two act as 'inputs' to the formation of the latter two, which involve larger time-spans and constitute a reduction of the most important tonal events, and a tonal tension-release structure. The mental representations posited by the GTTM represent an idealized 'final state' of largely unconscious (semantic) knowledge by a 'perfect' experienced listener (Lerdahl and Jackendoff 1983, p. 3).

Cue-Abstraction theory is more general, and is proposed as applicable to a wider range of music than to classical tonal music. In this theory, units of memory may be created by any strikingly distinctive features in the surface of the

music (called *cues*) over a range of time-spans, rather than just tonally significant events; these cues act as memory markers that represent larger segments of music. Cue abstraction theory proposes a simpler and looser hierarchality than the GTTM, a set of event hierarchies rather than tonal hierarchies. It privileges prominent surface details over deep structural events in the formation of musical LTM, and the importance of salience over tonal stability, though its operation at a range of time-spans allows it a degree of hierarchicality.

Because of the limitations of working memory, music can be immediately comprehended only on the time-scale of 5–8 sec. Hence the first step in a listener's construction of the form of a piece of music is the segmentation or chunking of the musical surface by identification of boundaries formed by points of change in the music. Many of the same factors that articulate musical units on lower levels such as phrases may also operate on higher structural levels; just as a musical phrase will be articulated by a change in the flow of events, a sectional boundary delineating a larger time-span may be articulated by an even stronger change, usually involving more musical parameters.

Both the GTTM and cue abstraction theory propose that segmentation is a foundation for the establishment of LTM representations, a proposal confirmed empirically by Deliège (1987). Experiments by Clarke and Krumhansl (1990) Krumhansl (1996), Deliège (1989), and Deliège *et al.* (1996) involving both classical and contemporary music established that listeners (both musicians and non-musicians) are often in considerable agreement about the location of major segmentation points in pieces of music.

Although an unconscious memory reduction is a theoretical entity, several experiments have been conducted to try to determine whether listeners have such representations. Building on an earlier experiment by Serafine (Serafine *et al.* 1989), experiments by Dibben (1994) and Bigand (1990) using composed reductions of short musical examples have produced some evidence of the cognitive reality of reductions in tonal music, at least on a modest time-scale. Dibben's results led her to conclude that cognitive representations of atonal music are associational rather than strictly hierarchical, echoing Imberty (1993).

Experiments by Deliège and others (Deliège and Melen 1997) established that listeners, both musicians and non-musicians, use surface features of music as memory cues. These experiments involved organizing randomly reordered segments of a previously heard piece on a timeline representing their original order (Deliège 1993; Clarke and Krumhansl 1990; Deliège and Melen 1997). Unlike segmentation experiments, where the performance of musicians and non-musicians are often comparable, in many of these experiments, musicians tended to do better than non-musicians, though for both groups judgments about segment location were more accurate near the beginning and end of pieces. Other experiments have explored the development of listeners' schemas for musical elements in particular pieces by requiring judgments about the similarity of fragments of music from the same piece. This approach proposes that thematic variants are categorized around abstract *prototypes* (generalized representations of thematic material) and that variants of a theme can be heard as similar (Zbikowski 1999). Deliège (2001) and Krumhansl (1991) performed experiments investigating listeners' ability to identify as yet unheard fragments of a partially heard piece. In both cases listeners were able to use what they had heard to identify unheard examples as being similar and as coming from the same piece. In a further experiment Deliège (1996) asked listeners (musicians and non-musicians) to compare a large set of small (1–2 measure) 'cells' from a Bach violin sonata to two different reference or 'prototype' cells. In general, listeners were able to identify the variants with the correct prototypes, although the non-musicians were found to have a narrower definition of similarity. From results of a similar experiment using pieces by Beethoven and Schoenberg, Lamont and Dibben (2001) concluded that the listeners had used *surface* features to relate fragments.

In summary, it appears that both musicians and non-musicians use surface features in constructing long-term mental representations of music. In effect, the existence and utility of larger-scale hierarchical memory reductions has not been firmly established, though most experiments on tonal structures have investigated relationships between events adjacent in time

rather than those separated by large spans of music (see Chapter 6). The use of deep structural features in ordinary (not analytically oriented) listening has been called into question by the results of several experiments. These experiments have explored the question of the role of memory in the perception of an entire piece of music, in part by addressing the question of what sort of information seems to be available in memory over the total duration of a piece as it unfolds in time. The results of these experiments would suggest that, for most listeners, the types of relationships between constituent parts of a piece that are described in theories such as that of Lerdahl and Jackendoff are not accessible in memory.

Experiments by Cook (1987) and by West Marvin and Brinkman (1999) explored whether the initial key of a piece was still accessible in listeners' memories at the end of the piece. Both studies found some evidence that listeners were able to undertake this task successfully, but only for pieces shorter than ca 30 seconds. This supports the view that tonal closure may be perceived most strongly over short time-spans (Levinson, 1997 but see also Gjerdingen, 1999 for a critique of Cook's study). However, West Marvin and Brinkman also found that participants were unable to indicate that pieces that ended in keys other than the initial key were less 'tonally closed' than pieces which ended in their initial key.

The results of several other studies (e.g. Karno and Konecni 1992) support the view that long-range relationships between musical materials are neither particularly accessible nor stable in memory. Further experiments by Clarke and Krumhansl (1990) Deliège (1993), Deliège *et al.* (1996) and Deliège and Mélen (1997) explored the ability of musically trained and untrained listeners to remember the order in which events had occurred after hearings of classical and post-tonal pieces. Although listeners' judgments of location of segments and segment-order were roughly accurate, they were less so for segments toward the middle of pieces, although musically-trained participants appeared better able to use higher-level schematic knowledge to achieve greater accuracy. It has to be noted that the results of these experiments do not prove that large-scale musical relationships cannot be encoded in memory. It is still hypothetically possible that the memory processes of an expert listeners, or more likely, performers, are capable of dealing with musical relationships over large time-scales. However, for most listeners, it seems likely that memory does not accurately encode these longer-term, larger-scale, longer-duration, hierarchical relationships.

## Conclusions

Much about musical memory still remains to be understood, indeed, explored. New fields of research are continuing to extend what is known, as well as throwing up new questions. The field of neuroscience and the methods of neuroimaging have already shown great promise in identifying the areas and networks in the brain that are implicated in memory for music. Frontal, parietal, and premotor cortical areas, together with the cerebellum, are associated with working memory for musical pitch, while processing and representation of musical meaning appear to rely on areas of the middle temporal gyrus and left anterior temporal areas (Koelsch and Siebel 2005). Similarly, the study of the effect of emotion on cognitive and memory processes is very much at an early stage, but can be expected to provide fruitful insights into the nature of musical memory. Many studies have shown that emotion, or affect, might be one of the most significant factors that determine how and what we remember (see Dolan 2002 for an overview), and given that music is strongly associated with the modulation of emotional state, it can be expected that the study of musical memory will provide critical evidence about the nature of human memory in general (see Juslin and Sloboda 2001).

## References

Aarden B (2003). Dynamic melodic expectancy. PhD dissertation. School of Music, Ohio State University.

Arom S, Leothaud G and Voisin F (1997). Experimental ethnomusicology: an interactive approach to the study of musical scales. In I Deliège and J Sloboda, eds., Perception and cognition of music, 3–30, Psychology Press, Hove, East Sussex.

Baddeley A (1986). *Working memory*. Oxford University Press, New York.

Baddeley A (2003). Working memory and language: an overview. *Journal of Communication Disorders*, **36**, 189–208.

Barnes R and Jones MR (2000). Expectancy, attention, and time. *Cognitive Psychology*, **41**, 254–311.

Bartlett FC (1932). *Remembering*. Cambridge University Press, New York.

Bharucha J (1994). Tonality and expectation. In R Aiello with JA Sloboda, eds, *Musical perceptions*, 213–239. Oxford University Press, New York.

Bigand E (1990). Abstraction of two forms of underlying structure in a tonal melody. *Psychology of Music*, **18**(1), 45–59.

Bornstein RF (1989). Exposure and affect: overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, **106**(2), 265–289.

Bregman A (1990). *Auditory scene analysis*. MIT Press, Cambridge, MA.

Burns E (1999). Intervals, scales and tuning. In D Deutsch, ed., *The psychology of music*, 2nd Edn, 215–264. Academic Press, London.

Chafe W (1994). *Discourse, consciousness, and time*. University of Chicago Press, Chicago, IL.

Clarke E (1987). Categorical rhythm perception: an ecological perspective. In A Gabrielsson, ed., *Action and perception in rhythm and music*, 19–33. Royal Swedish Academy of Music, Stockholm.

Clarke E and Krumhansl C (1990). Perceiving musical time. *Music Perception*, **7**(3), 213–251.

Cohen NJ and Squire LR (1980). Preserved learning and retention of pattern analyzing skill in amnesia: dissociation of knowing how and knowing that. *Science*, **210**, 207–210.

Collier GL and Collier JL (2002). A study of timing in two Louis Armstrong solos. *Music Perception*, **19**(3), 463–483.

Cook N (1987). The perception of large-scale tonal closure. *Music Perception*, **5**(2), 197–206.

Cowen N (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin*, **104**, 163–191.

Cowen N (2005). *Working memory capacity*. Psychology Press, New York.

Crowder R (1993). Auditory memory. In S McAdams, E Bigand, eds, *Thinking in sound*, 113–140. Oxford University Press, Oxford.

Crowder R and Morton J (1969). Precategorical acoustic storage (PAS). *Perception and Psychophysics*, **5**, 365–373.

Deliège I (1987). Grouping conditions in listening to music: an approach to Lerdahl and Jackendoff's grouping preference rules. *Music Perception*, **4**, 325–360.

Deliège I (1989). A perceptual approach to contemporary musical forms. *Contemporary Music Review*, **4**, 213–230.

Deliège I (1993). Mechanisms of cue extraction in memory for musical time. *Contemporary Music Review*, **9**, 191–206.

Deliège I (1996). Cue abstraction as a component of categorization processes in music listening. *Psychology of Music*, **24**, 131–156.

Deliège I (2001) Prototype effects in music listening: an empirical approach to the notion of imprint. *Music Perception*, **18**, 371–407.

Deliège I, Mélen M, Stammers D and Cross I (1996). Musical schemata in real-time listening to a piece of music. *Music Perception*, **14**(2), 117–159.

Deliège I, Melen M (1997). Cue abstraction in the representation of musical form. In I Deliège and J Sloboda, eds, *Perception and cognition of music*, 387–412. Psychology Press, Hove, East Sussex.

Deutsch D (1972) Octave Generalization and tune recognition. *Perception and Psychophyics*, **11**, 411–412.

Deutsch D (1999). The processing of pitch combinations. In D Deutsch, ed., *The psychology of music*, 2nd Edn, 349–412. Academic Press, San Diego, CA.

Deutsch D (2006) The enigma of absolute pitch. *Acoustics Today*, **2**, 11–19.

Deutsch D and Feroe J (1981). The internal representation of pitch sequences in tonal music. *Psychological Review*, **86**, 503–22.

Dibben N (1994). The cognitive reality of hierarchical structure in tonal and atonal music. *Music Perception*, **12**(1), 1–26.

Dolan RJ (2002). Emotion, cognition and behavior. *Science*, **298**, 1191–1194.

Dowling WJ (1978). Scale and contour: two components of a theory for melodies. *Psychological Review*, **85**(4), 341–354.

Dowling WJ (1994). Melodic contour in hearing and remembering melodies. In R Aiello with JA Sloboda, eds, *Musical perceptions*, 173–190. Oxford University Press, New York.

Dowling WJ and Harwood DL (1986). *Music cognition*. Academic Press, San Diego, CA.

Dowling WJ, Tillmann B and Ayers D (2002). Memory and the experience of hearing music. *Music Perception*, **19**(2), 249–276.

Ebbinghaus H (1885) *Uber das Gedachtnis: Untersuchungen zur experimentellen Psychologie (1913) Memory, a contribution to experimental psychology*, trans. HA Ruger and CE Bussenius. New York: Teachers College, Columbia University. Reprint, 1964. Dover, New York.

Ellenberger HF (1970). *The discovery of the unconscious*. Basic Books, New York.

Ericsson KA and Kintsch W (1995). Long-term working memory. *Psychological Review*, **102**, 211–245.

Fink R (1999). Going flat: post-hierarchical music theory and the musical surface. In N Cook, M Everist, eds, *Rethinking music*, 102–137. Oxford University Press, New York.

Fraisse P (1982). Rhythm and tempo. In D Deutsch, ed., *The psychology of music*, 1st Edn, 149–181. Academic Press, San Diego, CA.

Francès R (1988). *The perception of music*, trans. WJ Dowling. Lawrence Erlbaum, Hillsdale, NJ.

Gabrielsson A (1999). The performance of music. In D Deutsch, ed., *The psychology of music*, 2nd Edn, 501–602. Academic Press, London.

Gjerdingen RO (1988). A classical turn of phrase. University of Pennsylvania Press, Philadelphia, PA.

Gjerdingen RO (1999). An experimental music theory? In N Cook and M Everist, eds, *Rethinking music*, 161–170. Oxford University Press, Oxford.

Hebb DO (1949) *The organization of behavior: a neuropsychological theory*. Wiley, New York.

Huron D (2006). *Sweet anticipation, music and the psychology of expectation*. MIT Press, Cambridge, MA.

Imberty M (1993). How do we perceive atonal music? Suggestions for a theoretical approach. *Contemporary Music Review*, **9**, 322–337.

James W (1890/1950) *The principles of psychology*. (Reprint edition, 1950) Dover, New York.

Jones MR, Moynihan H, MacKenzie N and Puente J (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*, **13**(4), 313–319.

Jonides J and Smith EE (1997). The architecture of working memory. In M Rugg, ed., *Cognitive neuroscience*, 243–276. MIT Press, Cambridge, MA.

Juslin P and Sloboda JA (eds (2001)). *Music and emotion: theory and research*. Oxford University Press, Oxford.

Karno M and Konecni V (1992). The effects of structural interventions in the first movement of Mozart's Symphony in G Minor, K. 550 on aesthetic preference. *Music Perception*, **10**(1), 63–72.

Koelsch S and Siebel WA (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*, **9**(12), 578–584.

Krumhansl CL (1990). Cognitive foundations of musical pitch. Oxford University Press, New York.

Krumhansl CL (1991). Memory for musical surface. *Memory and Cognition*, **19**, 401–411.

Krumhansl CL (1996). A perceptual analysis of Mozart's piano sonata K. 282: segmentation, tension, and musical ideas. *Music Perception*, **13**, 401–432.

Krumhansl CL, Sandell GJ and Sergeant DC (1987). The perception of tone hierarchies and mirror forms in twelve-tone serial music. *Music Perception*, **5**, 31–78.

Lamont A and Dibben N (2001). Motivic structure and the perception of similarity. *Music Perception*, **18**(3), 245–274.

Lee C (1991). The perception of metrical structure: experimental evidence and a model. In P Howell, R West and I Cross, eds, *Representing musical structure*, 59–128. Academic Press, London.

Lerdahl F (1988). Cognitive constraints on compositional systems. In JA Sloboda, ed., *Generative processes in music*, 231–259. Oxford University Press, Oxford.

Lerdahl F and Jackendoff R (1983). *A generative theory of tonal music*. MIT Press, Cambridge, MA.

Levinson J (1997). *Music in the moment*. Cornell University Press, Ithica, NY.

London J (2004). *Hearing in time: psychological aspects of musical metre*. Oxford University Press, Oxford.

Mandler J (2004). *The foundations of mind*. Oxford University Press, Oxford.

McAdams S (1989). Psychological constraints on the form-bearing dimensions in music. *Contemporary Music Review*, **4**, 181–198.

Meyer LB (1956). *Emotion and meaning in music*. University of Chicago Press, Chicago, IL.

Miller GA (1956). The magical number seven, plus or minus 2: some limits on our capacity for processing information. *Psychological Review*, **63**(2), 81–93.

Narmour E (1990). *The analysis and cognition of basic melodic structures*. University of Chicago Press, Chicago, IL.

Palmer C and C Krumhansl (1990). Mental representation for musical metre. *Journal of Experimental Psychology, Human Perception and Performance*, **16**(4), 728–741.

Peretz I and Zatorre RJ (2005). Brain organization for music processing. *Annual Review of Psychology*, **56**, 89–114.

Povel D (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance*, **7**(1), 3–18.

Povel D (1984). A theoretical framework for rhythm perception. *Psychological Research*, **45**, 315–337.

Reber A (1993). *Implicit learning and tacit knowledge*. Oxford University Press, Oxford.

Roediger HL, Dudai Y and Fitzpatrick SM (2007). *Science of memory: concepts*. Oxford University Press, Oxford.

Saffran JR, Johnson EK, Aslin RN and Newport EL (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, **70**, 27–52.

Schank R (1982). *Dynamic memory*. Cambridge University Press, Cambridge.

Serafine ML, Glassman N and Overbeeke C (1989). The cognitive reality of hierarchical structure in music. *Music Perception*, **6**, 347–430.

Siegal JA (1970). Sensory and verbal coding strategies in subjects with absolute pitch. *Journal of Experimental Psychology*, **103**, 37–44.

Sloboda J (1985). Immediate recall of melodies. In P Howell, I Cross, R West, eds, *Musical structure and cognition*, 143–168. Academic Press, London.

Snyder B (2000). *Music and memory*. MIT Press, Cambridge, MA.

Sperling G (1960). The information available in brief visual presentations. *Pyschological Monographs*, **74**, 1–29.

Tulving E (1972). Episodic and semantic memory. In E Tulving and W Donaldson, eds, *Organization of memory*, 381–403. Academic Press, New York.

Ward WD (1999). Absolute pitch. In D Deutsch, ed. *The psychology of music*, 2nd edn, 265–298. Academic Press, New York.

West ME and Brinkman A (1999). The effect of modulation and formal manipulation on perception of tonic closure by expert listeners. *Music Perception*, **16**(4), 389–407.

Yeston M (1976). *The stratification of musical rhythm*. Yale University Press, New Haven, CT.

Zbikowski L (1999). Musical coherence, motive, and categorization. *Music Perception*, **17**(1), 5–42.

# PART 3

# Responses to music

Edited by Donald A. Hodges

# CHAPTER 11

# Bodily responses to music

Donald A. Hodges

BODILY responses are among the core, common experiences of music. To observe people at a rock concert or watching a marching band, is to observe heads nodding, feet tapping, and bodies swaying. Along with these observable behaviours are a host of reactions to music that occur beneath the skin, such as changes in heart rate, blood pressure, and blood chemistry. The experience of listening to music involves highly complex interactions among cognitive, affective, and bodily processes that take place within a personal–social–cultural context.

The purpose of this chapter is to explore bodily responses to music. At the outset, it is important to provide some operational definitions, as the terminology in the literature is not used uniformly. Following these definitions, the main body of the chapter will be devoted to reviews of research literature. Although various types of bodily responses to music are discussed in a linear, segmented fashion, it must be remembered that they occur in a holistic fashion.

## Definition of terms

For simplicity, bodily responses are divided into physiological and physical responses. Behavioural responses might also be considered, but as they are dealt with in other chapters, they will not be discussed here. Even as these two categories are being defined, it is important to note that such operational definitions are somewhat arbitrary. Although experiences can be categorized for ease of understanding, the normal experience is one of integrated response, with no sharp boundaries between them. Given this caveat, for the purposes of this chapter the categories are defined as follows:

- *Physiological responses* include internal bodily processes, such as heart rate. Although occasionally these internal processes are reflected in observable changes, for the most part detection requires some type of monitoring device.
- *Physical responses* are external, readily observable, reflexive motor movements such as foot tapping. These responses occur naturally, without specific training. Physical responses that require training, for example training a new music student to tap a foot to keep the beat while performing, will not be covered in this chapter.

## Physiological responses

What follows in the next section are extremely succinct reviews of the literature under the various physiological responses measured. Findings of research related to music listening responses are generally organized into those studies reporting increases, decreases, or no change in responses. This simplistic approach masks an underlying complexity that will be discussed in a concluding section. Because this concluding discussion will deal with general principles affecting more than one type of physiological response, comments in the brief review sections are kept to a minimum.

## The effects of music on heart or pulse rate

A significant amount of research supports the position that listening to music can cause changes in heart rate (e.g. Bernardi *et al.* 2006; Blood and Zatorre 2001; Savan 1999). In general, stimulative music—characterized by fast, loud, staccato passages—tends to cause an increase in heart rate (HR) or pulse rate (PR), while sedative music—characterized by slow, soft, legato passages—tends to cause a decrease. However, some researchers found that any music, whether stimulative or sedative, tends to increase HR or PR (Krumhansl 1997; Rickard 2004).

In only a slightly fewer number of studies did researchers report that music caused no changes in HR or PR (e.g. Davis 1992; Gupta and Gupta 2005; Pujol 1994). In Iwanaga *et al.* (1996), HR did not change in response to stimulative music, but did decrease during sedative music.

The effects of performing music on heart rate are much less studied. Measurements of pulse rate in 24 orchestral musicians of the Vienna Symphony were made during rehearsals and public performances (Haider and Groll-Knapp 1981). Maximum pulse rates during performances reached 151 beats per minute. Pulse rates were up to 8 beats higher in concerts than during rehearsals and fluctuated widely depending on the character of the music being performed, with tempo being a salient feature. For example, during a Rachmaninoff symphony, one trumpet player went from a resting pulse rate of 68 to a peak of over 150. Often, peaks occurred just prior to an entrance. Abel and Larkin (1990) found that student musicians had increased heart rates in anticipation of a graded performance.

## The effects of music on skin conductivity

Galvanic skin responses (GSR), now more often referred to as electrodermal activity or skin conductance, are temporary, primarily negative, fluctuations in the electrical resistance of the skin. Measured by a psychogalvanomoter, usually attached to the fingers and/or palm of the hand, the GSR is a result of mental activity, most often of an affective nature (Venables 1987). An increase in GSR readings indicates a decrease in resistance, which in turn indicates an increase in arousal. Significant changes in GSR in response to music listening were found in numerous studies (e.g. Khalfa *et al.* 2002; Lundqvist *et al.* 2000; Rickard 2004). In other studies, no meaningful changes in GSR were found in response to music listening (Jellison 1975; Keller and Seraganian 1984).

## The effects of music on blood pressure

Stimulative music increased blood pressure in Bernardi *et al.* (2006), Krumhansl (1997) and Pignatiello *et al.* (1989). In a number of studies, music decreased blood pressure; most often sedative music had this effect, though some stimulative music did so too (Lorch *et al.* 1994, Savan 1999). Noteworthy is the fact that self-selected music was found to be effective in lowering blood pressure (Geden *et al.* 1989; Oyama *et al.* 1987a; Updike and Charles 1987). Confounding these results are studies that reported no meaningful change in blood pressure as a result of listening to music (Gupta and Gupta 2005; Strauser 1997), which in some cases was self-selected (Geden *et al.* 1989; Schuster 1985). In a related study, music students experienced an increase in blood pressure in anticipation of a graded performance (Abel and Larkin 1990).

## The effects of music on biochemical responses

A relatively new area of investigation has been the effect of music listening on hormone levels found in blood, urine, or saliva. The basis for this work is the field of psychoneuroimmunology, a field that explores interactions between nervous and immune systems (Hall 1989). The majority of published reports presented in Table 11.1 indicate meaningful changes in biochemicals in response to music listening or musical experiences.

Collectively, this line of research supports the central contention of psychoneuroimmunology, namely that mental and bodily responses are intertwined in a symbiotic relationship. This is well illustrated in three examples. Choral singers had increases in SigA of 159 per cent during rehearsal and 240 per cent during performance,

**Table 11.1** Biochemical responses to music

| Biochemical | Response to music | Sources |
|---|---|---|
| Secretory immunoglobulin A (SIgA) | Increased | Beck *et al.* 2000, Charnetski and Brennan 1998, Kreutz *et al.* 2004, Lane 1991, McCraty *et al.* 1996, Rider *et al.* 1991, Rider and Weldin 1990, Tsao *et al.* 1991 |
| | Decreased | Miluk-Lolasa *et al.* 1994 |
| | No change | Oyama *et al.* 1987a, Rider *et al.* 1985 |
| *Cortisol* | Increased | Beck *et al.* 2000, Gerra *et al.* 1998, VanderArk and Ely 1992, 1993 |
| | Decreased | Bartlett *et al.* 1993, Beck *et al.* 2000, Kreutz *et al.* 2004, Leardi *et al.* 2007, McKinney *et al.* 1997a, Tsao *et al.* 1991 |
| | Increased in music majors but decreased in non-music majors | VanderArk and Ely 1992, 1993 |
| Interleukin-1 | Increased | Bartlett *et al.* 1993 |
| Adrenocorticotropic hormone (ACTH) | Increased | Gerra *et al.* 1998 |
| | Decreased | Oyama *et al.* 1987b |
| Neutrophils and lymphocytes | Decreased | Rider and Achterberg 1989 |
| Norepinephrine | Increased | Gerra *et al.* 1998, Kumar *et al.* 1999, VanderArk and Ely 1992, 1993 |
| Growth hormone | Increased | Gerra *et al.* 1998 |
| Epinephrine | Increased | Kumar *et al.* 1999, VanderArk and Ely 1992, 1993 |
| | No change | Gerra *et al.* 1998 |
| Melatonin | Increased | Kumar *et al.* 1999 |
| Serotonin | No change | Kumar *et al.* 1999 |
| Prolactin | No change | Gerra *et al.* 1998, Kumar *et al.* 1999 |
| Beta-endorphins | Increased | Gerra *et al.* 1998, Goldstein 1980 |
| | Decreased | McKinney *et al.* 1997b |
| Dopamine | The release of dopamine in response to music, indicating a pleasure response, has been implicated by activations of mesolimbic structures involved in reward processing including the nucleus accumbens (NAc) and the ventral tegmental area (VTA), as well as the hypothalamus and insula (Blood and Zatorre 2001, Brown *et al.* 2004, Menon and Levitin 2005). | |
| Oxytocin | Although oxytocin has been mentioned in connection with lullabies and mother–infant bonding (Freeman 1995), such reporting is anecdotal and no published research studies have been identified. | |
| Genetic stress hormone markers | Participants in a recreational music-making programme experienced reversal in 19 of 45 genetic stress hormone markers (Bittman *et al.* 2005). | |

with a corresponding decrease in cortisol of 30 per cent during rehearsal and an average increase of 37 per cent during performance (Beck *et al.* 2000). When patients with Alzheimer's disease underwent four weeks of music therapy sessions, melatonin, norepinephrine, and epinephrine all increased, while serotonin, and prolactin did not (Kumar *et al.* 1999). Among 18–19-year-olds, a meaningful increase was observed in beta-endorphins, ACTH, norepinephrine, growth hormone, and cortisol after listening to techno music (Gerra *et al.* 1998). Classical music induced an improvement in emotional state, but no meaningful changes in hormonal concentrations. This line of research is central to the practice of music medicine and is further explored in Chapter 49.

## The effects of music on respiration

An increase in respiration or breathing rate was found in several studies (Blood and Zatorre 2001; Iwanaga and Moroki 1999). Ries (1969) correlated increases with preference and Haas *et al.* (1986) found that breathing rates entrained with musical rhythm. Researchers found no change in respiration rate as a result of listening to music (Davis 1992; Davis-Rollans and Cunningham (1987) and Iwanaga *et al.* (1996) found that respiration did not change to stimulative music but did decrease to sedative music.

## The effects of music on finger, peripheral skin, or body temperature

Meaningful changes in skin temperature in response to music listening have been reported (Lundqvist *et al.* 2000; Rickard 2004). Beyond the fact that listening to music can affect temperature changes, there is very little consistency in results. In some studies, temperature increased—to sedative music (Kibler and Rider 1983; Peach 1984), to stimulative music (Lundqvist *et al.* 2000; Standley 1991), or to any music (Rickard 2004). In other cases, temperature decreased (Krumhansl 1997; Savan 1999). In only three reports did music not affect skin temperature (Guzzetta 1989; Kibler and Rider 1983; Zimmerman *et al.* 1988).

## The effects of music on muscular tension

Muscular tension is most often measured by electromyography (EMG) and is related to tension-release patterns in music. The majority of studies found that muscular tension changes in response to music listening (Blood and Zatorre 2001; Rickard 2004). Only two published studies report no changes in EMG while listening to music (Davis and Thaut 1989; Scartelli 1984). Nielsen (1983, 1987; cited in Gabrielsson and Lindstrom 2001) reported that the more tension in the music, the harder subjects squeezed on a pair of tongs. Sears (1958) found that sedative music decreased tension more quickly than stimulative music increased tension. Carrick *et al.* (2007) found that listening to certain music had the potential to increase postural stability.

## Chills

The following studies found that music listeners reported chills, often accompanied by shivers, tears, lump in the throat, or pilomotor (goosebump) responses (Craig 2005; Goldstein 1980; Panksepp 1995; Rickard 2004). Blood and Zatorre (2001) discovered that as the intensity of chills increased, so did blood flow to areas of the brain involved in mediating reward, motivation, emotion, and arousal. More than 80% of Sloboda's (1991) subjects reported having chills to music. Guhn *et al.* (2007) found that HR and skin conductance increased in musical passages that elicited the most chills. This confirmed findings by Craig (2005), Rickard (2004), and Panksepp (1995), but contrasted with Blood and Zatorre (2001).

## The effects of music on blood volume

Two studies (Davis and Thaut 1989; Krumhansl 1997) found significant changes in blood volume in response to music listening, while two studies (Jellison 1975; Pignatiello *et al.* 1989) did not.

## The effects of music on gastric motility

More formally known as peristalsis or gastric motility, stomach contractions move food along

the alimentary canal. Presumably, music could speed up or slow down this process. Two studies (Chen *et al.* 2005; Demling *et al.* 1970) support the contention that music can affect gastric activity. Beyond that, there is no consensus on the relationship between type of music and its effect on digestion.

## Miscellaneous studies

### Blood-oxygen

Music had noticeably positive effects on blood oxygen saturation levels (Cassidy and Standley 1995) and varied according to the stimulation rates and intensity (Lovett *et al.* 1952).

### Ankle jerk

Subjects' ankles jerked in response to music (Harrer and Harrer 1977).

### Pupillary reflex

Music had a significant effect on pupillary reflex, with dilation occurring during stimulative music and constriction during sedative music (Slaughter 1954).

### Zygomaticus activity

Happy/positive music elicited increased zygomaticus activity (smile muscle) compared to sad/negative music (Lundqvist *et al.* 2000; Witvliet and Vrana 2007).

## General discussion of physiological responses to music

From this cursory review of research findings, it is apparent that there is overwhelming support for the notion that listening to music affects physiological responses. Furthermore, there is general, though by no means unanimous, support for the notion that stimulative and sedative music tend to increase and decrease physiological responses, respectively. The reason this relationship is not stronger is most likely due to personal preferences.

Music listening experiences are highly idiosyncratic, as each person possesses an individual history (e.g., amount of musical training) and personality, and each listening experience occurs in varying situations. One of Fridja's (1988) laws of emotion is the Law of Situational Meaning: emotions arise in response to the meaning structures of a given situation; different emotions arise in response to different meaning structures (p. 349). Imagine, then, that one hears the same piece of music, say the slow movement of Bach's Concerto for Two Violins in d minor, BWV 1043, performed at a funeral and then a few weeks later at a wedding. Undoubtedly, the reactions would be significantly different.

As is evident from the foregoing reviews, there is not enough research on the effects of personal preference on physiological responses to clarify the issue and some of it is contradictory. For example, Iwanaga and Moroki (1999) found that type of music (i.e., stimulative or sedative) had a differing effect on heart rate, respiration rate, and blood pressure, but music preference did not. Nevertheless, the concept that personal preference is influential is firmly enough accepted to be utilized in the use of music in medical treatment (e.g., anxiolytic music for the treatment of pain and anxiety) (Spintge and Droh 1992; see Chapter 49 this volume for more details on medical applications and Chapter 15 for the role of preferences in general).

Once it is accepted that physiological processes change in response to music listening, what do such changes signify? Does an increase in heart rate, for example, necessarily indicate a positive response or could it just as easily indicate a negative response? Attaching specific physiological responses to specific emotional responses has not received much attention nor had much success. Increased smile muscle activity while listening to happy music as opposed to sad music is one of the few examples (Lundqvist *et al.* 2000). Carefully designed studies, with increased attention to personal preferences and idiosyncrasies, may lead to an enhanced understanding of the influence of music on physiological responses. Currently, the most advanced work is being conducted under the rubric of music medicine.

# Physical responses

Anthropologists have documented that music and movement are linked in peoples all over the world (Blacking 1995). Many activities, such as dancing, marching, exercising, and working to

music exhibit a synchronization and entrainment between auditory input and motor output (Brown *et al.* 2006). Our brains and bodies are wired to respond to sound and to rhythmic sounds in particular (Koepchen *et al.* 1992). Furthermore, entraining our bodies to musical rhythms brings pleasure (Levitin 2006). Listening to music activates the following brain regions in a sequential order:

1 the auditory cortex initially analyses sound;
2 frontal regions process musical structure;
3 the mesolimbic system, involved in arousal and pleasure, is activated and produces dopamine, further activating the nucleus accumbens; and
4 the cerebellum and basal ganglia process rhythm and meter leading to movement (Menon and Levitin 2005).

When we listen to music our bodies naturally respond with largely involuntary gestures, such as head nodding and foot tapping.

Audiomotor systems are interconnected such that musical patterns elicit and entrain strong motor responses (Baumann *et al.* 2007; Thaut 2003; Thaut *et al.* 1999). For example, subjects are able to finger tap synchronously with heard rhythms, even making adjustments when changes in tempo are below the threshold of detection (Tecchio *et al.* 2000). Such effects may be enhanced by mirror neurons (Nelissen *et al.* 2005; Rizzolatti *et al.* 1996). These are neurons in the brain that fire both when an action is performed and when it is observed in someone else. More recent work indicates that mirror neurons may also be triggered by auditory stimuli in what is called 'action-listening'. Lahav *et al.* (2007) demonstrated that non-pianists who learned to perform a brief excerpt and then who heard the same pattern without performing any physical action had activations in bilateral frontoparietal motor-related networks.

In the vast majority of musical experiences around the world, physical responses are a natural phenomenon. Only in a few situations, such as Western classical music concerts, is there an attempt to control these responses. One might speculate that the relatively immobile audience posture during Western classical music concerts, followed by a wild burst of applause after a performance judged to be stellar, might be, in part, a reflection of pent up physical responses.

Physical responses to music begin in the final trimester before birth (Parncutt 2006) and continue throughout infancy with the singing of lullabies, the playing of 'musical' games such as pat-a-cake and peek-a-boo, and the presence of crib mobiles and musical toys. Rhythmic and timing interactions between parent and child can be illustrated by the 'I'm gonna gitcha' game (Stern 1982). The parent chants: 'I'm gonna gitcha I'm gonna gitcha I'm gonna gitcha Gotcha!' At each repetition the pitch and loudness increase, while the amount of time between each repetition progressively elongates. Activities such as this are so important that they form the basis for acquiring cognitive expectancies and for interrelating cognition and affect (Beebe *et al.* 1982; Thaut *et al.* 2005).

The human body consists of multiple oscillators that create the rhythms that drive the body and control behaviour (Brown and Graeber 1982). These interaction rhythms are important to music and dance but are essential to all human interactions (Scheflen 1982). The intentional application of rhythmic entrainment to motor behaviours has been made with Parkinsonian and stroke patients, allowing them to regain walking, reaching, and grasping skills (see Chapter 49 this volume).

## Conclusion

In the book entitled *How Musical is Man?*, Blacking (1973) eloquently captures the symbiotic relationship between music and human beings in the first and last chapters: 'Humanly organized sound' and 'Soundly organized humanity'. In the first chapter, Blacking contends that music is a product of human behaviour resulting from biological processes and cultural agreement. The brain's ability to imagine, create and perceive patterns of sounds combines with social conventions to give us music. Music, for Blacking, is 'sound that is organized into socially accepted patterns' (p. 25). Above all, music exists because human beings impose sonic order rather than accept passively the sounds that nature provides.

Having stressed that music is humanly organized sound, Blacking then counters with 'Soundly organized humanity'. He acknowledges that music affects bodily processes but concentrates

his discussion primarily on the effects of music on social behaviours. Although the focus in the current chapter has, of course, been on the former, Blacking's main point is essentially the main point of this chapter: the music we create in turn affects us in profound ways. Even though the narrower focus of this chapter is on bodily responses, Blacking's larger view is also important to keep in mind. That is, bodily processes are influenced by the larger personal–social–cultural circumstances in which they occur. Lomax (1968), Merriam (1964) and many others have provided considerable field data from around the world to underscore the concept of music embedded in social context.

Main themes in this chapter are that bodily responses to music are among the core experiences of music. They are hugely complex, with a myriad of response types interwoven into the fabric of thoughts, feelings, and social context. Bodily responses are highly idiosyncratic as each person brings a unique self to a music listening situation. Researchers have made significant strides in ferreting out the details of these responses and in collaboration with practitioners have made progress in utilizing this knowledge, particularly in music therapy and music medicine applications. With all this, however, the richness and complexity of the human experience leaves much yet to be discovered.

While recognizing strides already made, the following ideas are offered as a means of advancing our knowledge of how humans react bodily to music to a new level of understanding:

- Greater consensus on terminology and standardization of protocol would be helpful. Some of the contradictory findings reported in the literature review must surely be due to these inconsistencies.
- Another explanation for contradictory findings may be that in some circumstances participants were emotionally responding to the music, while in others they were not (see Chapter 12 this volume for more discussion of this point). This is a confound that needs to be kept in mind in studies of bodily responses to music.
- Moving steadily toward more naturalistic musical experiences would be enlightening. Although it may appear as we as if are very far away from field studies, we may be closer than one might imagine. More than 25 years ago, members of the Vienna Symphony were monitored for heart rate and brain wave activity during rehearsals and public performances (Haider and Groll-Knapp 1981). Since then, ambulatory assessment has become a new and developing field. Using digital technology with hand-held PCs, researchers are now able to monitor such physiological responses as heart rate, blood pressure, blood glucose, and so on while subjects are engaged in routine activities (Fahrenberg 2001). Applying this technology to music listeners in all manner of naturalistic settings should be possible. In fact, in a recent experiment whose results are not yet known, a conductor, five members of a symphony orchestra, and fifteen audience members were all wired with sensors to record heart rate, muscle tension, respiration, and galvanic skin response during a performance (Dyer 2006). Findings may allow for comparisons of physiological responses as the music is expressed in conducting gestures, performer's reactions to those gestures, and audience responses to both.
- Paradoxically, we must pay increased attention to the idiosyncratic nature of responses while also gaining clarity in our understanding of corporate responses. Balancing personalized responses with the group dynamics involved in audience responses is a challenge that has not received much attention. For example, it is easy to imagine that one's bodily responses would be highly influenced by being one of thousands of wildly cheering fans at a rock concert.
- Advancing technologies are giving us an ability to gain increasingly refined views of bodily responses, particularly physiological responses, but this must not let us become myopic. Fine-grained details must be viewed within a social context.

None of these suggestions are made cavalierly. Surmounting inherent problems will be exceedingly difficult. Nevertheless, continuing to measure heart rate responses to 15 different 1-minute musical excerpts during a laboratory experiment is not likely to advance understanding much beyond what we already know. It will

be most interesting to read a review of this literature at some point in the future to see what innovative strategies researchers have devised to bring deeper understanding to these most basic of human responses to music.

## References


Abel J and Larkin K (1990). Anticipation of performance among musicians: physiological arousal, confidence, and state anxiety. *Psychology of Music*, **18**, 171–182.

Bartlett D, Kaufman D and Smeltekop R (1993). The effects of music listening and perceived sensory experiences on the immune system as measured by interleukin-1 and cortisol. *Journal of Music Therapy*, **30**(4), 194–209.

Baumann S, Koeneke S, Schmidt C, Meyer M, Lutz K and Jancke L (2007). A network for audio-motor coordination in skilled pianists and non-musicians. *Brain Research*, **1161**, 65–78.

Beck R, Cesario T, Yousefi A and Enamoto H (2000). Choral singing, performance perception, and immune system changes in salivary immunoglobulin A and cortisol. *Music Perception*, **18**(1), 87–106.

Beebe B, Gerstman L, Carson B, Dolinas M, Zigman A, Rosensweig H, Faughey K and Korma M (1982). Rhythmic communication in the mother-infant dyad. In M. Davis, ed., *Interaction rhythms: periodicity in communication behavior*, 79–100 New York: Human Sciences Press.

Bernardi L, Porta C and Sleight P (2006). Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: the importance of silence. *Heart*, **92**, 445–452.

Bittman B, Berk L, Shannon M, Sharaf M, Westengard J, Guegler K and Ruff D (2005). Recreational music making modulates the human stress response: a preliminary individualized gene expression strategy. *Medical Science Monitor*, **11**(2), BR231–40.

Blacking J (1973). *How musical is man?* University of Washington Press, Seattle WA.

Blacking J (1995). *Music, culture, and experience*. University of Chicago Press, Chicago, IL.

Blood A and Zatorre R (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 11818–11823.

Brown F and Graeber R (eds) (1982). *Rhythmic aspects of behavior*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Brown S, Martinez M and Parsons L (2004). Passive music listening spontaneously engages limbic and paralimbic systems. *NeuroReport*, **15**(13), 2033–2037.

Brown S, Martinez M and Parsons L (2006). The neural basis of human dance. *Cerebral Cortex*, **16**(8), 1157–1167.

Carrick F, Oggero E and Pagnacco G (2007). Posturographic changes associated with music listening. *The Journal of Alternative and Complementary Medicine*, **13**(5), 519–526.

Cassidy J and Standley J (1995). The effect of music listening on physiological responses of premature infants in the NICU. *Journal of Music Therapy*, **32**(4), 208–227.

Charnetski C and Brennan F Jr (1998). Effect of music and auditory stimuli on secretory immunoglobulin A (IgA). *Perceptual Motor Skills*, **87**, 1163–1170.

Chen D, Xu X, Wang Z and Chen J (2005). Alteration of gastric myoelectrical and autonomic activities with audio stimulation in healthy humans. *Scandinavian Journal of Gastroenterology*, **40**(7), 814–821.

Craig D (2005). An exploratory study of physiological changes during 'chills' induced by music. *Musicae Scientiae: The journal of the European Society for the Cognitive Sciences of Music*, **9**(2), 273–285.

Davis C (1992). The effects of music and basic relaxation instruction on pain and anxiety of women undergoing in-office gynecological procedures. *Journal of Music Therapy*, **29**(4), 202–216.

Davis W and Thaut M (1989). The influence of preferred relaxing music on measures of state anxiety, relaxation, and physiological responses. *Journal of Music Therapy*, **26**(4), 168–187.

Davis-Rollans C and Cunningham S (1987). Physiologic responses of coronary care patients to selected music. *Heart-Lung*, **16**, 370–378.

Demling L, Tzschoppe M and Classen M (1970). The effect of various types of music on the secretory function of the stomach. *Digestive Diseases and Sciences*, **15**(1), 15–20.

Dyer R (2006). A new kind of experimental music. *Boston Globe*, 16 April.

Fahrenberg J (2001). Origins and developments of ambulatory monitoring and assessment. In J Fahrenberg, M Myrtek, eds, *Progress in ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies*, 578–614 Hogrefe and Huber, Seattle, WA.

Freeman W (1995). *Societies of brains: a study in the neuroscience of love and hate*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Fridja N (1988). The laws of emotion. *American Psychologist*, **43**(5), 349–358.

Gabrielsson A and Lindström E (2001). The influence of musical structure on emotional expression. In P Juslin, J Sloboda, eds, *Music and emotion: theory and research*, 223–248 Oxford University Press, Oxford.

Geden E, Lower M, Beattie S and Beck N (1989). Effects of music and imagery on physiologic and self-report of analogued labor pain. *Nursing Research*, **38**(1), 37–41.

Gerra G, Zaimovic A, Franchini D, Palladino M, Giucastro G, Reali N, Maestri D, Caccavari R, Deslsignore R and Brambilla F (1998). Neuroendocrine responses of healthy volunteers to 'techno-music': relationships with personality traits and emotional state. *International Journal of Psychophysiology*, **28**(1), 99–111.

Goldstein A (1980). Thrills in response to music and other stimuli. *Physiological Psychology*, **3**, 126–129.

Guhn M, Hamm A and Zentner M (2007). Physiological and music-acoustic correlates of the chill response. *Music Perception*, **24**(5), 473–483.

Gupta U and Gupta B (2005). Psychophysiological responsivity to Indian instrumental music. *Psychology of Music*, **33**(4), 363–372.

Guzzetta C (1989). Effects of relaxation and music therapy on patients in a coronary care unit with presumptive acute myocardial infarction. *Heart Lung*, **18**, 609–616.

Haas F, Distenfeld S and Axen K (1986). Effects of perceived musical rhythm on respiratory patterns. *Journal of Applied Physiology*, **61**(1–3), 1185–1191.

Haider M and Groll-Knapp E (1981). Psychophysiological investigations into the stress experienced by musicians in a symphony orchestra. In M Piperek, ed., *Stress and music*, 15–34 Vienna: Wilhelm Braumüller.

Hall S (1989). A molecular code links emotions, mind and health. *Smithsonian*, **20**(3), 62–71.

Harrer G and Harrer H (1977). Music, emotion, and autonomic function. In M Critchley, R Henson, eds, *Music and the brain*, 202–216 William Heinemann Medical Books, London.

Iwanaga M, Ikeda M and Iwaki T (1996). The effects of repetitive exposure to music on subjective and physiological responses. *Journal of Music Therapy*, **33**(3), 219–230.

Iwanaga M and Moroki Y (1999). Subjective and physiological responses to music stimuli controlled over activity and preference. *Journal of Music Therapy*, **36**(1), 26–38.

Jellison J (1975). The effect of music on autonomic stress responses and verbal reports. In C Madsen, R Greer, C Madsen, eds, *Research in music behavior: modifying music behavior in the classroom*, 206–219 Teachers College Press, New York.

Keller S and Seraganian P (1984). Physical fitness level and autonomic reactivity to psychosocial stress. *Journal of Psychosomatic Research*, **28**(4), 279–287.

Khalfa S, Peretz I, Blondin J and Manon R (2002). Event-related skin conductance responses to musical emotions in humans. *Neuroscience Letters*, **328**, 145–149.

Kibler V and Rider M (1983). The effect of progressive muscle relaxation and music on stress as measured by finger temperature response. *Journal of Clinical Psychology*, **39**, 213–15.

Koepchen H, Droh R, Spintge R, Abel H-H, Klüssenforf D and Koralewski E (1992). Physiological rhythmicity and music in medicine. In R Spintge, R Droh, eds, *Music Medicine*, 39–70 MMB Music, St. Louis, MO.

Kreutz G, Bongard S, Rohrmann S, Hodapp V and Grebe D (2004). Effects of choir singing or listening on secretory immunoglobulin A, cortisol, and emotional state. *Journal of Behavioral Medicine*, **27**(6), 623–635.

Krumhansl C (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology*, **51**(4), 336–352.

Kumar A, Tims F, Cruess D and Mintzer M (1999). Music therapy increases serum melatonin levels in patients with Alzheimer's disease. *Alternative Therapies in Health and Medicine*, **5**(6), 49–57.

Lane D (1991). The effect of a single music therapy session on hospitalized children as measured by salivary immunoglobulin A, speech pause time, and a patient opinion Likert scale. *Pediatric Research*, **29**(4), 2:11A.

Lahav A, Saltzman E and Schlaug G (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *Journal of Neursocience*, **27**(2), 308–314.

Leardi S, Pietroletti R, Angeloni G, Necozione S, Ranalletta G and Del Gusto B (2007). Randomized clinical trial examining the effect of music therapy in stress response to day surgery. *British Journal of Surgery*, **94**(8), 943–947.

Levitin D (2006). *This is your brain on music*. Dutton, New York.

Lomax A (1968). *Folk song style and culture*. Transaction Books, New Brunswick, NJ.

Lorch C, Lorch V, Diefendorf A and Earl P (1994). Effect of stimulative and sedative music on systolic blood pressure, heart rate, and respiratory rate in premature infants. *Journal of Music Therapy*, **31**(2), 105–118.

Lovett Doust J and Schneider R (1952). Studies on the physiology of awareness: the effect of rhythmic sensory bombardment on emotions, blood oxygen saturation, and the levels of consciousness. *Journal of Mental Science*, **98**, 640–653.

Lundqvist L, Carlsson F and Hilmersson P (2000). Facial electromyography, autonomic activity, and emotional experience to happy and sad music. *Journal of Psychology*, **35**(3/4), 225.

McCraty R, Atkinson M, Rein G and Watkins A (1996). Music enhances the effect of positive emotional states on salivary IgA. *Stress Medicine*, **12**, 167–175.

McKinney C, Antoni M, Kumar M, Tims F and McCabe P (1997a). Effects of Guided Imagery and Music (GIM) therapy on mood and cortisol in healthy adults. *Health Psychology*, **16**(4), 390–400.

McKinney C, Tims F, Kumar A and Kumar M (1997b). The effect of selected classical music and spontaneous imagery on plasma beta-endorphin. *Journal of Behavioral Medicine*, **20**(1), 85–99.

Menon V and Levitin D (2005). The rewards of music listening: response and physiological connectivity of the mesolimbic system. *NeuroImage*, **28**(1), 175–184.

Merriam A (1964). *The anthropology of music*. Northwestern University Press, Chicago, IL.

Miluk-Kolasa B, Obminski Z, Stupnicki R and Golec L (1994). Effects of music treatment on salivary cortisol in patients exposed to pre-surgical stress. *Experimental and Clinical Endocrinology*, **102**, 118–120.

Nelissen K, Luppino G, Vanduffel W, Rizzolatti G and Orban G (2005). Observing others: multiple action representation in the frontal lobe. *Science*, **310**(5746), 332–336.

Oyama T, Sato Y, Kudo T, Spintge R and Droh R (1987a). Effect of anxiolytic music on endocrine function in surgical patients. In R Spintge, R Droh, eds, *Music in medicine*, 169–174 Springer-Verlag, Berlin.

Oyama T, Hatano K, Sato Y, Kudo R, Spintge R and Droh R (1987b). Endocrine effect of anxiolytic music in dental patients. In R Spintge and R Droh, eds, *Music in medicine*, 223–226 Springer-Verlag, Berlin.

Panksepp J (1995). The emotional sources of 'chills' induced by music. *Music Perception*, **13**(2), 171–207.

Parncutt R (2006). Prenatal development. In G McPherson, ed., *The child as musician*, 1–31 Oxford University Press, Oxford.

Peach S (1984). Some implications for the clinical use of music facilitated imagery. *Journal of Music Therapy*, **21**(1), 27–34.

Pignatiello M, Camp C, Elder S and Rasar L (1989). A psychophysiological comparison of the Velten and musical mood induction techniques. *Journal of Music Therapy*, **26**(3), 140–154.

Pujol K (1994). The effect of vibrotactile stimulation, instrumentation, and precomposed melodies on physiological and behavioral responses of profoundly retarded children and adults. *Journal of Music Therapy*, **31**(3), 186–205.

Rickard N (2004). Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of Music*, **32**(4), 371–388.

Rider M and Achterberg J (1989). Effect of music-assisted imagery on neutrophils and lymphocytes. *Biofeedback and Self-Regulation*, **14**(3), 247–257.

Rider M, Floyd J and Kirkpatrick J (1985). The effect of music, imagery, and relaxation on adrenal corticosteroids and the re-entrainment of circadian rhythms. *Journal of Music Therapy*, **22**(1), 46–58.

Rider M, Mickey C, Weldin C and Hawkinson R (1991). The effects of toning, listening, and singing on psychophysiological responses. In C Maranto, ed., *Applications of music in medicine*, 73–84 National Association for Music Therapy, Washington, DC.

Rider M and Weldin C (1990). Imagery, improvisation and immunity. *The Arts in Psychotherapy*, **17**, 211–216.

Ries H (1969). GSR and breathing amplitude related to emotional reactions to music. *Psychonomic Science*, **14**(2), 62.

Rizzolatti G, Fadiga L, Gallese V and Fogassi L (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, **3**(2), 131–141.

Savan A (1999). The effect of background music on learning. *Psychology of Music*, **27**(2), 138–146.

Scartelli J (1984). The effect of EMG biofeedback and sedative music, EMG biofeedback only, and sedative music only on frontalis muscle relaxation ability. *Journal of Music Therapy*, **21**(2), 67–78.

Scheflen A (1982). Preface: Comments on the significance of interaction rhythms. In M Davis, ed., *Interaction rhythms: periodicity in communicative behavior*, 13–22 Human Sciences Press, New York.

Schuster B (1985). The effect of music listening on blood pressure fluctuations in adult hemodialysis patients. *Journal of Music Therapy*, **22**(3), 146–153.

Sears W (1958). The effect of music on muscle tonus. In E Gaston, ed., *Music therapy, 1957*, 199–205 Allen Press, Lawrence, KS.

Slaughter F (1954). The effect of stimulative and sedative types of music on normal and abnormal subjects as indicated by pupillary reflexes. In M Bing, ed., *Music therapy*, 246–248 Allen Press, Lawrence, KS.

Sloboda J (1991). Music structure and emotional response: some empirical findings. *Psychology of Music*, **19**, 110–120.

Spintge R and Droh R (1992). Toward a research standard in musicmedicine/music therapy: a proposal for a multimodal approach. In R Spintge, R Droh, eds, *Musicmedicine*, 345–349 MMB Music, St. Louis, MO.

Standley J (1991). The effect of vibrotactile and auditory stimuli on perception of comfort, heart rate, and peripheral finger temperature. *Journal of Music Therapy*, **28**(3), 120–134.

Stern D (1982). Some interactive functions of rhythm changes between mother and infant. In M Davis, ed., *Interaction rhythms: periodicity in communicative behavior*, 101–117 Human Sciences Press, New York.

Strauser M (1997). The effects of music versus silence on measures of state anxiety, perceived relaxation, and physiological responses of patients receiving chiropractic interventions. *Journal of Music Therapy*, **34**(2), 88–105.

Tecchio F, Salustri C, Thaut M, Pasqualetti P and Rossini P (2000). Conscious and preconscious adaptation to rhythmic auditory stimuli: a magnetoencephalographic study of human brain responses. *Experimental Brain Research*, **135**, 222–230.

Thaut M (2003). Neural basis of rhythmic timing networks in the human brain. *Annals of the New York Academy of Sciences*, **999**, 364–373.

Thaut M, Kenyon G, Schauer M and McIntosh G (1999). The connection between rhythmicity and brain function. *IEEE Engineering in Medicine and Biology*, **18**, 101–108.

Thaut M, Peterson D and McIntosh G (2005). Temporal entrainment of cognitive functions: musical mnemonics induce brain plasticity and oscillatory synchrony in neural networks. *Annals of the New York Academy of Sciences*, **1060**, 243–254.

Tsao C, Gordon T, Maranto C, Lerman C and Murasko D (1991). The effects of music and directed biological imagery on immune response S-IgA. In C Maranto, ed., *Applications of music in medicine*, 85–121 National Association for Music Therapy, Washington, DC.

Updike P and Charles D (1987). Music Rx: physiological and emotional responses to taped music programs of preoperative patients awaiting plastic surgery. *Annals of Plastic Surgery*, **19**(1), 29–33.

VanderArk S and Ely D (1992). Biochemical and galvanic skin responses to music stimuli by college students in biology and music. *Perceptual and Motor Skills*, **74**, 1079–1090.

VanderArk S and Ely D (1993). Cortisol, biochemical and galvanic skin responses to music stimuli of different preference values by college students in biology and music. *Perceptual and Motor Skills*, **77**, 227–234.

Venables P (1987). Electrodermal activity. In R Gregory, ed., *The Oxford companion to the mind*, 213–214 Oxford University Press, Oxford.

Witvliet C and Vrana S (2007). Play it again Sam: repeated exposure to emotionally evocative music polarizes liking and smiling responses, and influences other affective reports, facial EMG, and heart rate. *Cognition and Emotion*, **21**(1), 3–25.

Zimmerman L, Pierson M and Marker J (1988). Effects of music on patient anxiety in coronary care units. *Heart-Lung*, **17**, 560–566.

CHAPTER 12

# Emotional responses to music

Patrik N. Juslin

MUSIC has been linked to the emotions at least since Ancient Greece, and emotions do figure prominently in people's reported motives for listening to music. People use music to change emotions, to release emotions, to match their current emotion, to enjoy or comfort themselves, and to relieve stress. Yet, despite the ubiquity of music in everyday life today (see Chapter 40 this volume), emotional responses to music are still regarded as elusive. Why does music touch us so deeply? In this chapter, I will consider the best possible answers to this question that the psychology of music can currently provide. While music philosophers, musicologists, and aestheticians have written a lot about musical emotions (Budd 1985), psychologists have been slow to catch up due to both conceptual and methodological problems (Juslin and Sloboda 2001). Hence, this chapter will focus mainly on research carried out during the last decade.

It is important to make a distinction between *perception* and *induction* of emotions. We may simply perceive an emotion in the music, or we may actually feel an emotion in response to the music. This distinction is often, but not always, made in modern research. It is important to make the distinction for two reasons. First, the underlying mechanisms may be different depending on the process involved. Secondly, the types of emotions usually induced by music may be different from the types of emotions usually perceived in music. This chapter focuses on induction of emotions, whereas perception of emotions is discussed in Chapters 13 and 35.

## The definition and measurement of emotions

A natural point of departure is to define the concept of emotion. Emotions belong to the field of affect. Based on an emerging consensus in the affective sciences (e.g. Davidson *et al.* 2003), *affect* is regarded as an umbrella term, which includes various affective phenomena such as preference, emotion, and mood. Affect simply refers to the fact that these phenomena are valenced—events are evaluated as positive or negative by the individual.

Although researchers do not agree on a precise definition of *emotions*, they agree about the characteristics. Emotions are described as brief (lasting minutes to a few hours) but intense responses to potentially important events or changes in the external or internal environment that involve certain subcomponents:

- cognitive appraisal (e.g., you appraise a situation as 'dangerous'),
- subjective feeling (e.g., you feel 'afraid'),
- physiological response (e.g., your heart starts to beat faster),
- expression (e.g., you shout),
- action tendency (e.g., you run away), and
- regulation (e.g., you try to calm yourself).

In principle, each of these six components can be used to measure emotions, though researchers debate the extent to which the components are 'synchronized' during an emotional response. Many researchers assume that people are always

in *some* affective state. When the states are intense and involve salient stimuli, we tend to call them 'emotions', whereas when the same states are less intense, and their causes are not immediately apparent, we tend to call them 'moods'. The states have been conceptualized both as discrete categories and as dimensions such as valence and arousal.

## Primary issues and methods

The fundamental issues to address in the study of music and emotion are: does music induce emotions? If so, which emotions does music typically induce? Under what circumstances do musical emotions commonly occur? How does music induce emotions? Are musical emotions different from other emotions? Could musical emotions have implications for our health? It is fair to say that music psychologists are only beginning to answer most of these questions.

There are several methods that may be used to investigate musical emotions, such as listening experiments (Waterman 1996), questionnaire studies (Juslin and Laukka 2004), diary studies (Sloboda *et al.* 2001), qualitative interviews (DeNora 2000) and brain imaging (Blood and Zatorre 2001). Since each method has both advantages and disadvantages, the best overall approach might be to combine various methods in a kind of 'method triangulation'. In addition, the componential approach to measurement of emotions is slowly gaining ground in music psychology; that is, music researchers increasingly realize the need to measure several emotion components to be able to draw valid conclusions about emotional responses (Scherer and Zentner 2001). However, such approaches are still available mainly in the laboratory. Field studies continue to rely mostly on self-report, even though there are techniques for ambulatory measurement of physiological responses that could be used (Fahrenberg and Myrtek 1996). As observed already by William James (1884), the physiological response may be the aspect that gives emotions their distinct phenomenological feeling, providing the 'heat' of the emotions: 'What kind of an emotion of fear would be left, if the feelings neither of quickened heartbeats nor of shallow breathing . . . were present, it is quite impossible to think' (pp. 193–194). Thus, it is recommendable to complement self-reports of emotions with physiological measurements (see Chapter 11). Evidence suggests that more intense emotions to music tend to involve more pronounced physiological reactions (Rickard 2004), including 'chills' (Panksepp 1995).

## Does music really induce emotions?

Although most musicians and listeners would seem to take the emotional powers of music for granted, it has been the matter of some controversy whether music really can evoke emotions (e.g. Kivy 1990). Strong empirical evidence has been slow to emerge, though an increasing number of studies have now obtained evidence in terms of different emotion components such as self-reported feeling (e.g. DeNora 2000; Gabrielsson 2001; Juslin and Laukka 2004; Pike 1972; Waterman 1996), physiology (e.g. Davis and Thaut 1989; Krumhansl 1997; Nyklíček *et al.* 1997; Vaitl *et al.* 1993), activation of brain regions similar to those of other emotions (e.g. Blood and Zatorre 2001, Brown *et al.* 2004; Menon and Levitin 2005), emotional expression (e.g. Sloboda 1991, Witvliet and Vrana 2007), action tendency (e.g. Fried and Berkowitz 1979; Harrer and Harrer 1977) and regulation (e.g. Becker 2001; Gabrielsson 2001). There is also preliminary evidence of synchronization among the various emotion components (Lundqvist *et al.* in press). Most of the evidence cited above was collected in Western societies, so cross-cultural studies are urgently needed.

It should be noted that many studies have relied only on self-report, which may be subject to *demand characteristics*, i.e., the total sum of cues that convey the researcher's hypothesis to the participant and thus may influence the participant's behaviour (Orne 1962). It is therefore promising that several studies have found effects of musically induced emotions on so-called implicit measures such as writing speed, word association, and decision time, which may be less sensitive to demand characteristics (Västfjäll 2002). The continued exploration and use of such indirect measures is recommended. Measuring emotions in the voice as a response to music is another hitherto unexplored approach (Juslin and Scherer 2005).

## What is the prevalence of musical emotions?

To estimate the prevalence of musical emotions (i.e., the relative frequency of occurrence in the population of interest), it is necessary to capture the emotions as they spontaneously occur in daily life. The *experience sampling method* (ESM) is a useful tool in this context. Sloboda *et al.* (2001) conducted a pilot study to test the validity of the ESM with regard to music. Eight participants received an electronic pager and a diary. During a week, the pager was activated randomly once every two hours by a central computer. When they were paged, the participants were required to fill out the diary to describe the nature of any experiences of music that had occurred since the last paging. Findings suggested that music was heard during 44 per cent of the recorded episodes, that few of these episodes involved music listening as the main focus, and that music experiences tended to make the listeners more positive, more alert, and more focused in the present—particularly when personal choice over the music was involved. These results were replicated and extended in more extensive study, featuring 32 participants, by Juslin *et al.* (in press). Their participants were required to carry a small palmtop computer, which randomly gave off seven signals each day over a time period of two weeks. Music occurred in 37 per cent of the sampled episodes. In 64 per cent of the episodes that featured music, the participants reported that the music influenced how they felt. This is roughly similar to the retrospective, aggregated estimate provided in a survey study by Juslin and Laukka (2004), in which participants reported that they experienced emotions on average 55 per cent of the time they listened to music. Musical emotions were most frequent in the evening and during weekends. Overall, the results suggest that, although emotional responses to music are common, the occurrence of music does not itself necessarily guarantee that a listener will be moved by the music.

A much debated but little investigated issue is which emotions music can induce. Can music induce the full range of human emotions in listeners? In principle, depending on the specific music, the person listening, and the situation, it would indeed appear possible that music can induce just about any emotion that may be felt in other realms of human life. The wide range of emotions found in Gabrielsson's (2001; Table 19.1) study of strong experiences with music suggests that this is the case. However, given the rare nature of such experiences, a more interesting question is perhaps which emotions music *commonly* evokes in listeners. A handful of survey studies have provided preliminary estimates of the prevalence of particular emotions (Juslin and Laukka 2004 Table 4; Sloboda 1992 Table 1; Wells and Hakanen 1991 Table 1) and suggest that happiness, love, calm, sadness, excitement and nostalgia are relatively frequent emotional responses to music. However, these studies share two problems: they rely on retrospective and aggregated self-reports that may not be reliable and they do not directly compare musical with non-musical emotions, so that the 'uniqueness' of musical emotions can be assessed.

The previously mentioned ESM study by Juslin *et al.* (in press) provided estimates of occurrence of 14 emotional states in response to music as well as other stimuli during a 2-week period. Figure 12.1 shows the relative frequency (%) of emotions caused by music and emotions caused by other stimuli, respectively. As can be seen, emotions such as *calm–contentment*, *happiness–elation*, and *interest–expectancy* were generally common, whereas emotions such as *disgust–contempt* were not so common. A comparison of musical with non-musical emotion episodes indicated that *happiness–elation* and *nostalgia–longing* were significantly more frequent in episodes with musical emotions, whereas *anger–irritation*, *boredom–indifference*, and *anxiety–fear* were significantly more frequent in episodes with non-musical emotions. The remaining differences were not significant. Although the difficulties involved in translating experiences of emotions into verbal labels should be kept in mind, the results from both this and the above studies suggest that music may evoke a range of both basic and complex emotions, something that a satisfactory theory of musical emotions must be able to explain.

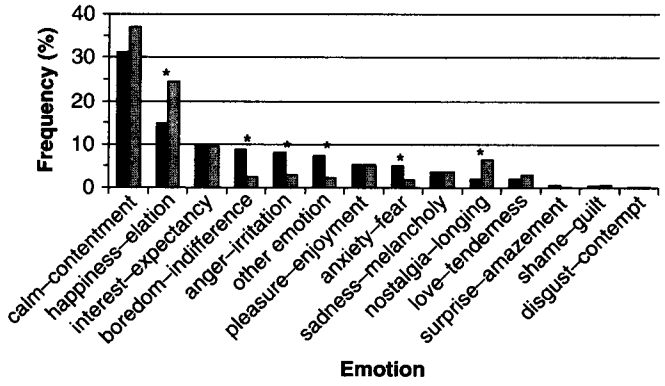



**Fig. 12.1** The relative frequency (in percentages) of felt emotions for (a) non-musical emotion episodes (dark bars) and, (b) musical emotion episodes (light bars) during a 2-week period of experience sampling (from Juslin *et al.* in press). The number of response alternatives was limited by the computer software used. * $p < .05$.

## Which factors influence musical emotions?

Factors that influence emotions in association with musical events have mostly been studied in an explorative fashion, and the factors may be grouped in terms of whether they are related to the music, the individual listener, or the situation.

### Musical factors

Whereas the musical features involved in the expression and perception of different emotions are fairly well mapped by now (see Chapters 13 and 35), the features that induce emotions have been less investigated (though there might be some overlap depending on the induction mechanism, see later this chapter). I predict that it will be difficult to find direct and simple links between musical features and emotions. The musical features must rather be related to possible induction mechanisms. An excellent example is Sloboda's (1991) attempt to link emotional reactions to music to specific events in the music: 83 participants, aged 16–70, most of them performers, were asked to specify particular pieces of music to which they could recall having experienced various physical manifestations frequently associated with an emotion. Having identified such pieces, which came mainly from classical music, they were then asked to specify the exact location within the music that provoked these reactions. Most participants reported whole pieces, movements, or sections of movements, which suggests an emotional response to the 'overall' character of the music (Table 3 in Sloboda 1991). Such a response might reflect 'emotional contagion' based on the emotional expression of the piece (Juslin and Västfjäll in press). However, roughly a third of the participants were able to locate their reaction within a theme or smaller unit. The data showed that musical events associated with 'tears' (i.e., crying, lump in the throat) commonly contained melodic appoggiaturas and melodic or harmonic sequences; events associated with 'shivers' (i.e., goose pimples, shivers down the spine) usually contained a new or unprepared harmony; and events associated with 'heart reactions' contained syncopations and prominent events occurring earlier than prepared for. The results were related to Meyer's (1956) notion that musical expectancies, and violations of these, play an important role in emotional responses to music.

### Individual factors

Individual differences between listeners appear to be larger for induction of emotions than for perception of emotions, though little research has focused on individual differences in musical emotion induction. Still, research has revealed a number of factors in the individual that may

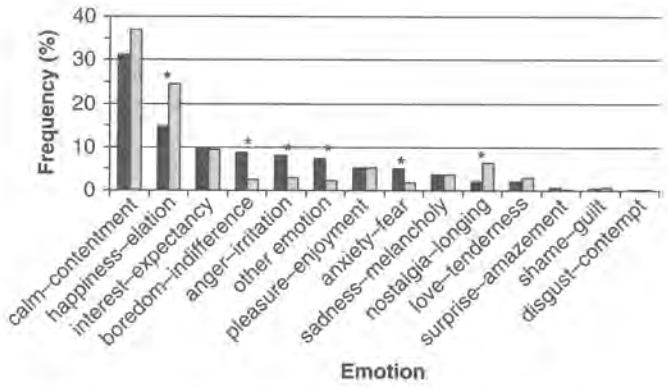



**Fig. 12.1** The relative frequency (in percentages) of felt emotions for (a) non-musical emotion episodes (dark bars) and, (b) musical emotion episodes (light bars) during a 2-week period of experience sampling (from Juslin *et al.* in press). The number of response alternatives was limited by the computer software used. * $p < .05$.

## Which factors influence musical emotions?

Factors that influence emotions in association with musical events have mostly been studied in an explorative fashion, and the factors may be grouped in terms of whether they are related to the music, the individual listener, or the situation.

### Musical factors

Whereas the musical features involved in the expression and perception of different emotions are fairly well mapped by now (see Chapters 13 and 35), the features that induce emotions have been less investigated (though there might be some overlap depending on the induction mechanism, see later this chapter). I predict that it will be difficult to find direct and simple links between musical features and emotions. The musical features must rather be related to possible induction mechanisms. An excellent example is Sloboda's (1991) attempt to link emotional reactions to music to specific events in the music: 83 participants, aged 16–70, most of them performers, were asked to specify particular pieces of music to which they could recall having experienced various physical manifestations frequently associated with an emotion. Having identified such pieces, which came mainly from classical music, they were then asked to specify the exact location within the music that provoked these reactions. Most participants reported whole pieces, movements, or sections of movements, which suggests an emotional response to the 'overall' character of the music (Table 3 in Sloboda 1991). Such a response might reflect 'emotional contagion' based on the emotional expression of the piece (Juslin and Västfjäll in press). However, roughly a third of the participants were able to locate their reaction within a theme or smaller unit. The data showed that musical events associated with 'tears' (i.e., crying, lump in the throat) commonly contained melodic appoggiaturas and melodic or harmonic sequences; events associated with 'shivers' (i.e., goose pimples, shivers down the spine) usually contained a new or unprepared harmony; and events associated with 'heart reactions' contained syncopations and prominent events occurring earlier than prepared for. The results were related to Meyer's (1956) notion that musical expectancies, and violations of these, play an important role in emotional responses to music.

### Individual factors

Individual differences between listeners appear to be larger for induction of emotions than for perception of emotions, though little research has focused on individual differences in musical emotion induction. Still, research has revealed a number of factors in the individual that may

potentially affect emotional responses to music, such as the listener's age, gender, personality, musical training, music preference and current mood (for a review, see Abeles and Chung 1996). Familiarity with the music tends to induce stronger emotional responses (Bartel 1992), which may explain why listeners respond stronger to self-selected than to experimenter-selected music (Harrer and Harrer 1977). Musical training might lead to less emotional involvement due to increasingly analytic listening (Harrer and Harrer 1977; but see also Lehmann 1997 for contradictory findings). Behne (1997) suggested that there may be different individual 'listening styles'. His proposals were based on cluster analyses of data from a questionnaire, which aimed to study musical experiences. Examples of such listening styles were 'emotional' ('I pay attention to what feelings are expressed through the music'), 'distancing' ('I try to follow the musical lines of a particular instrument'), 'sentimental' ('I remember things of the past'), and 'compensating' ('it makes me feel better').

### Situational factors

Factors related to the situation of a musical event could be categorized in different ways. Gabrielsson (2001) divided them into 'physical factors' (e.g., acoustical and visual conditions, time, place), 'social factors' (e.g., listening alone or together with others, type of audience, co-performers), 'special occasions and circumstances' (e.g., vacation), and 'performance conditions' (e.g., music well-rehearsed or not). Only recently have researchers begun to explore the ways that music actually figures in the everyday life of ordinary people (Part 9 of this volume), but it is becoming increasingly clear that key to an understanding of musical emotions is to describe the complex interplay between the music, the listener, and the situation. Juslin *et al.* (in press) noted that emotional responses to music occurred in all kinds of locations, partly because the participants were able to bring their favourite music anywhere (thanks to devices such as mp3 players). However, the prevalence of specific emotions varied depending on the situation. For example, some emotions such as *happiness–elation*, *pleasure–enjoyment*, and *anger–irritation* occurred frequently in 'social' settings, whereas others such as *calm–contentment*, *nostalgia-longing*, and *sadness–melancholy* occurred frequently in more 'solitary' settings. This highlights the need to use representative samples of situations in order to obtain accurate estimates of the prevalence of specific emotions.

## How does music induce emotions?

Most scholars acknowledge that this is the primary issue (e.g., Dowling and Harwood 1986, p. 202). Yet a literature search reveals that few studies make any attempt to test a theory about the psychological mechanism that underlies emotional responses to music. In fact, apart from Meyer's (1956) treatise (arguably the most cited book on this subject ever), most theories of musical emotions focus on the expressive properties of music that allow listeners to *perceive* emotions (e.g., Clynes 1977; Cooke 1959; Juslin 2001; Langer 1957).

One problem with musical emotions is that the conditions for eliciting emotions appear to be different from those in everyday life: in the paradigmatic case, an emotion is evoked when an event is appraised as having the capacity to influence the goals of the perceiver. Since music would not appear to have any capacity to further or block goals in life, researchers have been forced to come up with an alternative explanation that makes more sense in a musical setting. (However, musical induction of emotions through a cognitive appraisal may occur sometimes, such as when our 'goal' to go to sleep at night is 'blocked' by a neighbour playing loud music.)

### A theoretical framework

Building on the work of the pioneers in this field (Berlyne 1971; Meyer 1956) as well as on more recent research (Juslin and Sloboda 2001), Juslin and Västfjäll (in press) presented a theoretical framework featuring six mechanisms through which music can induce emotions. One could conceive of these mechanisms as consisting of a number of (more or less) distinct brain functions that developed gradually and in a specific order during evolution. Because the mechanisms

depend on functions with different evolutionary origins, each mechanism should possess unique characteristics. The framework focuses on the following mechanisms:

*Brainstem reflex* refers to a process whereby an emotion is induced by music because one or more fundamental acoustical characteristics of the music are taken by the brainstem to signal a potentially important and urgent event. All other things being equal, sounds that are sudden, loud, dissonant, or feature fast temporal patterns induce arousal in the listener. Such responses reflect the impact of simple auditory sensations (i.e., music as *sound* in the most basic sense). The perceptual system is constantly scanning the immediate environment in order to discover potentially important changes or events. Sounds that meet certain criteria will thus lead to an increased activation of the central nervous system in the early stages of auditory processing. The precise physiological processes underlying such brainstem responses are not completely understood, although evidence suggests that they occur in close connection with the reticular formation of the brainstem and the intralaminar nuclei of the thalamus, which receive inputs from the auditory system. The brainstem is an ancient structure of the brain that sub-serves a number of sensory and motor functions including, but not limited to, auditory perception and the mediation and control of attention, emotional arousal, heart rate, breathing, and movement (Joseph 2000). It may thus influence physiological processes related to emotions (e.g., pulse, respiration), such that the listener's arousal level increases or decreases.

*Evaluative conditioning* (EC) refers to a process whereby an emotion is induced by a piece of music simply because this stimulus has often been paired with other positive or negative stimuli (De Houwer *et al.* 2001). For example, a particular piece of music may have occurred repeatedly together in time with a specific event that always makes you happy, such as meeting one of your friends. Over time, through repeated pairings, the music itself will eventually evoke happiness even in the absence of the friendly interaction. There are few studies of EC with music (but see Blair and Shimp 1992), though EC has a number of features that are interesting in regard to music. First, EC can occur even if the participant is unaware of the contingency of the two stimuli. In fact, it has been found that EC can be both established and induce emotions without awareness (e.g. Martin *et al.* 1984). Secondly, EC appears to be fairly resistant to extinction, as compared to classical conditioning. Music often occurs in situations where music listening is not the main activity (Sloboda *et al.* 2001) and where subtle conditioning processes outside awareness might easily occur.

*Emotional contagion* refers to a process whereby an emotion is induced by a piece of music because the listener perceives the emotional expression of the music and then 'mimics' this expression internally (Juslin 2001). Emotional contagion has mostly been studied regarding facial expression (Hatfield *et al.* 1994), but Neumann and Strack (2000) also found evidence of contagion from emotional speech. Because music often features expressive acoustical patterns similar to those that occur in emotional speech (for a review, see Juslin and Laukka 2003), it has been argued that we get aroused by the voice-like aspects of music via a process in which a neural 'module' responds quickly and automatically to certain stimulus features, which leads us to mimic the perceived emotion internally (Juslin 2001). While the notion of emotional contagion via music admittedly remains somewhat speculative, a recent functional magnetic resonanace imaging (fMRI) study by Koelsch *et al.* (2006) found that music listening activated brain areas related to a circuitry serving the formation of pre-motor representations for vocal sound production (no singing was observed among the participants). Koelsch *et al.* concluded that this could reflect a mirror-function mechanism, similar to the so-called 'mirror neurons' proposed as a possible explanation of emotional contagion via other non-verbal channels (Preston and de Waal 2002).

*Visual imagery* refers to a process whereby an emotion is induced in a listener because he or she conjures up visual images (e.g., of a beautiful landscape) while listening to the music. The emotions experienced are the result of an interaction between the music and these images. The precise nature of this process remains to be described, but listeners appear to conceptualize the musical structure through a metaphorical non-verbal mapping between the music and

'image-schemata' grounded in bodily experience (see Lakoff and Johnson 1980); for instance, hearing melodic movement as 'upward'. Listeners might respond to mental images much in the same way as they would to the corresponding stimuli in the 'real' world, for instance reacting with positive emotions to a beautiful nature scene. Mental imagery in relation to music has mostly been explored in the context of music therapy (Toomey 1996-1997). Helen Bonny developed a method, *guided imagery and music* (GIM), where a 'traveller' is invited to 'share' his or her images as they are experienced in real time during a programmed musical sequence (Bonny and Savary 1973). Characteristic of visual imagery as an induction mechanism is that the listener is able to influence the process to a considerable extent. Although images might come into the mind unbidden, often a listener may conjure up, manipulate, and dismiss images at will.

*Episodic memory* refers to a process whereby an emotion is induced in a listener because the music evokes a personal memory of a specific event in the listener's life (Baumgartner 1992). This has been referred to as the 'Darling, they are playing our tune' phenomenon (see Davies 1978). Research has revealed that music often evokes memories (Gabrielsson 2001; Sloboda 1992). When the memory is evoked, so too is the emotion associated with the memory. Such emotions may be intense, perhaps because the physiological response patterns to the original events are stored in memory along with the experiential contents. Many listeners use music to remind them of valued past events, suggesting that music may serve an important 'nostalgic' function in everyday life. Music plays a very prominent role in adolescents' lives in regard to the development of a self-identity (e.g., Saarikallio 2007). Hence, one could expect episodic memories associated with music to be particularly emotionally vivid and frequent with regard to music from young adulthood, as indeed seems to be the case. Schulkind *et al.* (1999) found that older adults preferred, knew more about, as well as had stronger emotional responses to music popular during their youth than to music popular later in life.

*Musical expectancy* refers to a process whereby an emotion is induced in a listener because a specific feature of the music violates, delays, or confirms the listener's expectations about the continuation of the music (Meyer 1956; see also Narmour 1992). For instance, the sequential progression of E-F# sets up the expectation that the music will continue with G#. If this does not happen, listeners familiar with the musical idiom may become surprised. The expectations are based on the listener's previous experiences of the same style of music. Although Meyer's theory is highly regarded, it has not stimulated much research on emotions (but see Sloboda 1991), perhaps because the theory is difficult to test. For instance, a musical piece may evoke several different musical expectations at different levels of the music, and these expectations may also be different for different listeners. However, support for Meyer's theory was found by Steinbeis *et al.* (2006), who used subjective and physiological measures to capture emotional responses to unexpected events in music (see Figure 12.2).

Only a few of the above mechanisms have been tested in regard to music, but one thing seems clear: there is no single mechanism that can account for all instances of musical emotion. A crucial goal for future research is to develop a process model of the induction process, though given the large number of mechanisms, such an endeavour is a formidable undertaking. A first step could be to specify the precise characteristics of each mechanism. By synthesizing theory and findings from different domains outside music, Juslin and Västfjäll (in press) were able to provide the first set of hypotheses that can help researchers to distinguish among the proposed mechanisms. The hypotheses concern characteristics such as the information focus, key brain regions, ontogenetic development, cultural impact, induced affect, induction speed, degree of volitional influence, and so forth (see Juslin and Västfjäll, in press, Table 4).

## Implications of the framework

One implication of the novel framework is that it may resolve disagreements in the field. Specifically, apparent contradictions of different approaches may be reconciled by observing that they focus on different mechanisms. Hence, the framework may help to resolve previous disagreements about which emotions music can induce,

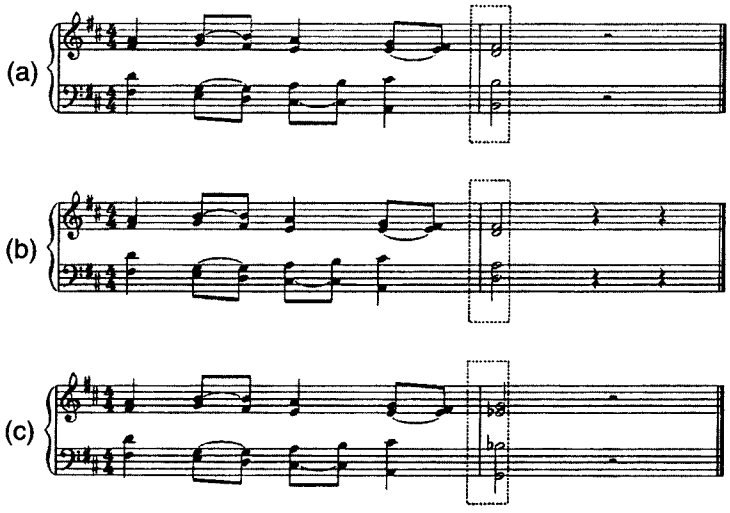



**Fig. 12.2** Examples of musical stimuli used to induce emotions through the musical expectancy mechanism (from Steinbeis *et al.* 2006). (a) is a cadence from an excerpt of an original composition by JS Bach with an unexpected event enclosed within the box. (b) and (c) are identical to version (a) except for the harmonic events within the boxes, which were manipulated to be either more expected (b) or less expected (c) than the original. Reproduced by permission of the MIT Press.

how early musical emotions develop, whether listeners are active or passive in the causal process, how much time it takes to induce an emotion through music, and whether musical emotions are innate or learned responses—it depends on the mechanism concerned.

The most important implication of the proposed framework for future research in the field is that it will not be sufficient to induce and study musical emotions *in general*. In order for data to contribute in a cumulative fashion to our knowledge, researchers must try to specify as far as possible the mechanism involved in each study. Otherwise studies will produce results that are inconsistent or that cannot be given a clear interpretation. One example is provided by the recent flurry of research on the neuropsychological correlates of emotional responses to music (Peretz 2001). Numerous brain regions have been reported in previous studies of music and emotion (e.g., the thalamus, cerebellum, hippocampus, amygdala, prefrontal cortex, Broca's area, nucleus accumbens, visual cortex, and supplementary motor areas). However, different brain regions have been implicated in different studies for reasons that are unclear. One likely explanation is that most studies have simply presented 'emotional music' to listeners without manipulating (or at least controlling for) the underlying induction mechanism. This makes it difficult to understand what the obtained neural correlates reflect in each study. Hence, Juslin and Västfjäll (in press) argue that musical emotions cannot be studied without respect to how they were evoked. An important task for the future is to develop standard paradigms and tasks that can reliably induce specific emotions in listeners through each of the mechanisms above.

## Conclusions

The study of music and emotions is fraught with considerable controversy and disagreement. Current debates in the field revolve around such questions as which emotions music induces, whether there are uniquely musical emotions, and what the relationship is between perceived and induced emotions. An important direction for future research is to consider how emotions might be one of several features of an aesthetic response to music—the appreciation of music

as 'art' (see Chapter 14 this volume). Real-world research has revealed that music is intentionally used by listeners to regulate their moods and emotions, and that positive emotions tend to dominate in their responses. The latter result has stimulated interest in effects of music on physical health and subjective well-being: can the emotional effects of music be used in health interventions that target biochemical substances, such as cortisol, oxytocin, dopamine, and serotonin? The development of such applications would obviously benefit from a better understanding of the underlying mechanisms. The framework discussed earlier may hopefully contribute to more hypothesis-driven approaches to exploring different mechanisms of musical emotions. Other possible applications of research on emotional responses to music include music therapy, film music, marketing, health care, and the gaming industry. In fact, several current applications of music presume its effectiveness in evoking emotions, yet we are only beginning to understand the mechanisms that make this possible. More often than not, music is intentionally designed to stir our emotions, exploiting whatever means available, and we have seen that music might induce emotions in many different ways. Hence, one answer to the question 'why does music move us?' is 'because we *want* music to move us and have designed it so that it will do it with the maximum effect'. However, rather than 'destroying the magic', an understanding of how music evokes emotions will only increase our never-ending fascination with music.

## References


Abeles HF and Chung JW (1996). Responses to music. In DA Hodges, ed., *Handbook of music psychology*, 2nd edn, 285–342. IMR Press, San Antonio, TX.

Bartel LR (1992). The development of the cognitive–affective response test—music. *Psychomusicology*, **11**, 15–26.

Baumgartner H. (1992). Remembrance of things past: music, autobiographical memory, and emotion. *Advances in Consumer Research*, **19**, 613–620.

Becker J (2001). Anthropological perspectives on music and emotion. In PN Juslin and JA Sloboda, eds, *Music and emotion: theory and research*, 135–160. Oxford University Press, New York.

Behne KE (1997). The development of 'Musikerleben' in adolescence: how and why young people listen to music. In I Deliége, JA Sloboda, eds, *Perception and cognition of music*, 143–159. Psychology Press, Hove, UK.

Berlyne DE (1971). *Aesthetics and psychobiology*. Appleton Century Crofts, New York.

Blair ME and Shimp TA (1992). Consequences of an unpleasant experience with music: a second-order negative conditioning perspective. *Journal of Advertising*, **21**, 35–43.

Blood AJ and Zatorre RJ (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of National Academy of Sciences*, **98**, 11818–11823.

Bonny HL and Savary LM (1973). *Music and your mind*. Station Hill, New York.

Brown S, Martinez MJ and Parsons LM. (2004). Passive music listening spontaneously engages limbic and paralimbic systems. *Neuroreport*, **15**, 2033–2037.

Budd M (1985). *Music and the emotions. The philosophical theories*. Routledge, London.

Clynes M (1977). *Sentics: the touch of emotions*. Doubleday, New York.

Cooke D (1959). *The language of music*. Oxford University Press, London.

Davidson RJ, Scherer KR and Goldsmith HH (eds) (2003). *Handbook of affective sciences*. Oxford University Press, New York.

Davies JB (1978). *The psychology of music*. Hutchinson, London.

Davis WB and Thaut MH (1989). The influence of preferred relaxing music on measures of state anxiety, relaxation, and physiological measures. *Journal of Music Therapy*, **26**, 168–187.

De Houwer J, Thomas S and Baeyens F (2001). Associative learning of likes and dislikes: a review of 25 years of research on human evaluative conditioning. *Psychological Bulletin*, **127**, 853–869.

DeNora T (2000). *Music in everyday life*. Cambridge University Press, Cambridge.

Dowling WJ and Harwood DL (1986). *Music cognition*. Academic Press, New York.

Fahrenberg J and Myrtek M (eds) (1996). *Ambulatory assessment: computer-assisted psychological and psychophysiological methods in monitoring and field studies*. Hogrefe, Seattle, WA.

Fried R and Berkowitz L (1979). Music that charms ... and can influence helpfulness. *Journal of Applied Social Psychology*, **9**, 199–208.

Gabrielsson A (2001). Emotions in strong experiences with music. In PN Juslin and JA Sloboda, eds, *Music and emotion: theory and research*, 431–449. Oxford University Press, New York.

Harrer G and Harrer H (1977). Music, emotion, and autonomic function. In M Critchley and RA Henson, eds, *Music and the brain. Studies in the neurology of music*, 202–216. William Heinemann Medical Books, London.

Hatfield E, Cacioppo JT and Rapson RL (1994). *Emotional contagion*. Cambridge University Press, New York.

James W (1884). What is an emotion? *Mind*, **9**, 188–205.

Joseph R (2000). *Neuropsychiatry, neuropsychology, clinical neuroscience*. Academic Press, New York.

Juslin PN (2001). Communicating emotion in music performance: a review and a theoretical framework. In PN

Juslin and JA Sloboda, eds, *Music and emotion: theory and research*, 309–337. Oxford University Press, New York.

Juslin PN and Laukka P (2003). Communication of emotions in vocal expression and music performance: different channels, same code? *Psychological Bulletin*, **129**, 770–814.

Juslin PN and Laukka P (2004). Expression, perception, and induction of musical emotions: a review and a questionnaire study of everyday listening. *Journal of New Music Research*, **33**, 217–238.

Juslin PN, Liljeström S, Västfjäll D, Barradas G and Silva A (2008). An experience sampling study of emotional reactions to music: listener, music and situation. *Emotion*, **8**, 668–683.

Juslin PN and Scherer KR (2005). Vocal expression of affect. In JA Harrigan, R Rosenthal and KR Scherer, eds, *The new handbook of methods in nonverbal behavior research*, 65–135. Oxford University Press, New York.

Juslin PN and Sloboda JA (eds), (2001). *Music and emotion: theory and research*. Oxford University Press, New York.

Juslin PN and Västfjäll D (2008). Emotional responses to music: the need to consider underlying mechanisms. *Behavioral and Brain Sciences*, **31**, 559–575.

Kivy P (1990). *Music alone*. Cornell University Press, Ithaca, NY.

Koelsch S, Fritz T, von Cramon DY, Müller K and Friederici AD (2006). Investigating emotion with music: an fMRI study. *Human Brain Mapping*, **27**, 239–250.

Krumhansl CL (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology*, **51**, 336–352.

Lakoff G and Johnson M (1980). *Metaphors we live by*. University of Chicago Press, Chicago, IL.

Langer SK (1957). *Philosophy in a new key*, 3rd edn. Harvard University Press, Cambridge, MA.

Lehmann AC (1997). Affective responses to everyday life and music listening. *Psychology of Music*, **25**, 84–90.

Lundqvist L-O, Carlsson F, Hilmersson P, and Juslin PN (2008). Emotional responses to music: experience, expression, and physiology. *Psychology of Music*, OnlineFirst, October 15, doi:10.1177/0305735607086048.

Martin DG, Stambrook M, Tataryn DJ and Beihl H (1984). Conditioning in the unattended left ear. *International Journal of Neuroscience*, **23**, 95–102.

Menon V and Levitin DJ (2005). The rewards of music listening: response and physiological connectivity of the mesolimbic system. *Neuroimage*, **28**, 175–184.

Meyer LB (1956). *Emotion and meaning in music*. Chicago University Press, Chicago, IL.

Narmour E (1992). *The analysis and cognition of melodic complexity: the implication–realization model*. University of Chicago Press, Chicago, IL.

Neumann R and Strack F (2000). Mood contagion: the automatic transfer of mood between persons. *Journal of Personality and Social Psychology*, **79**, 211–223.

Nyklíček I, Thayer JF and Van Doornen LJP (1997). Cardiorespiratory differentiation of musically-induced emotions. *Journal of Psychophysiology*, **11**, 304–321.

Orne MT (1962). On the social psychology of the psychological experiment with particular reference to demand characteristics and their implications. *American Psychologist*, **17**, 776–783.

Panksepp J (1995). The emotional sources of 'chills' induced by music. *Music Perception*, **13**, 171–208.

Peretz I (2001). Listen to the brain: a biological perspective on musical emotions. In PN Juslin, JA Sloboda, eds, *Music and emotion: theory and research*, 105–134. Oxford University Press, New York.

Pike A (1972). A phenomenological analysis of emotional experience in music. *Journal of Research in Music Education*, **20**, 262–267.

Preston SD and de Waal FBM (2002). Empathy: its ultimate and proximate basis. *Behavioral and Brain Sciences*, **25**, 1–72.

Rickard NS (2004). Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of Music*, **32**, 371–388.

Saarikallio S (2007). Music as mood regulation in adolescence. *Doctoral dissertation*, University of Jyväskylä, Finland.

Scherer KR and Zentner MR (2001). Emotional effects of music: production rules. In PN Juslin and JA Sloboda, eds, *Music and emotion: theory and research*, 361–392. Oxford University Press, New York.

Schulkind MD, Hennis LK and Rubin DC (1999). Music, emotion, and autobiographical memory: they are playing our song. *Memory and Cognition*, **27**, 948–955.

Sloboda JA (1991). Music structure and emotional response: some empirical findings. *Psychology of Music*, **19**, 110–120.

Sloboda JA (1992). Empirical studies of emotional response to music. In Riess-Jones and S Holleran, eds, *Cognitive bases of musical communication*, 33–46. American Psychological Association, Washington, DC.

Sloboda JA, O'Neill SA and Ivaldi A (2001). Functions of music in everyday life: an exploratory study using the Experience Sampling Method. *Musicae Scientiae*, **5**, 9–32.

Steinbeis N, Koelsch S and Sloboda JA (2006). The role of harmonic expectancy violations in musical emotions: evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, **18**, 380–1393.

Toomey L (1996–1997). Literature review: the Bonny method of guided imagery and music. *Journal of the Association for Music and Imagery*, **5**, 75–103.

Vaitl D, Vehrs W and Sternagel S (1993). Prompts—leitmotif—emotion: play it again, Richard Wagner. In N Birnbaumer and A Öhman, eds, *The structure of emotion: psychophysiological, cognitive, and clinical aspects*, 169–189. Hogrefe and Huber, Seattle, WA.

Västfjäll D (2002). A review of the musical mood induction procedure. *Musicae Scientiae*, Special Issue 2001–2002, 173–211.

Waterman M (1996). Emotional responses to music: implicit and explicit effects in listeners and performers. *Psychology of Music*, **24**, 53–67.

Wells A and Hakanen EA (1991). The emotional uses of popular music by adolescents. *Journalism Quarterly*, **68**, 445–454.

Witvliet CV and Vrana SR (2007). Play it again Sam: repeated exposure to emotionally evocative music polarizes liking and smiling responses, and influences other affective reports, facial EMG, and heart rate. *Cognition and Emotion*, **21**, 3–25.

CHAPTER 13

# The relationship between musical structure and perceived expression

Alf Gabrielsson

THE relationship between structure and expression in music has been discussed by philosophers and music theorists ever since antiquity, but empirical research did not start until about one hundred years ago. Most of this research is limited to Western classical music.

Musical structure is an umbrella term for a host of factors such as tempo, loudness, pitch, intervals, mode, melody, rhythm, harmony and various formal aspects (e.g. repetition, variation, transposition). They are designated by different symbols in the common musical notation. Most symbols represent categories with allowance for performance variation within the respective category. Listeners' perception of expression is thus affected both by the composed/notated structure and by the actual performance. However, performance aspects will not be further discussed here as they are treated elsewhere in this volume (Part 7).

Listeners' perception of expression—for instance, to perceive sadness in the music—should be distinguished from listeners' own reactions (that is, feeling sad themselves); the latter are treated elsewhere (Chapter 12). However, the border between these two alternatives is sometimes blurred, and the relationship between them may vary (Evans and Schubert 2008; Gabrielsson 2002; Juslin and Laukka 2004; Kallinen and Ravaja 2006; Konečni *et al.* 2008).

## Absolutism versus referentialism

Meyer (1956) made a distinction between absolutism and referentialism (other terms are formalism vs. expressionism, or autonomous vs. heteronomous music). Absolutism focuses on intra-musical relations; music is considered self-contained, refers only to itself. Referentialism implies, on the contrary, that music refers to extra-musical phenomena such as emotions, events, or characters.

Music critic Eduard Hanslick claimed that music consists of tone sequences which have no other content than themselves (Hanslick 1989, p. 162), and composer Igor Stravinsky held that 'if music appears to express something, this is only an illusion' (cited in Fisk 1997, pp. 280–281). However, the distinction between absolutism and referentialism is rather vague. Hanslick and Stravinsky later added several modifications to their seemingly dogmatic statements, and 'absolute meanings and referential meanings can and do coexist in one and the same piece of music' (Meyer 1956, p.1).

In the present chapter emphasis will be on *perceived* expression rather than expression somehow inherent in the music. The listener may apprehend music as 'pure' music (absolutism) or as expression of emotions, characters,

events or whatever, and may very well alternate, consciously or unconsciously, between different approaches during the course of a piece. Focus will be on referential meaning, leaving questions about intra-musical relationships and structural listening to other chapters in this volume (Part 2); see Clarke (2005) for further comments on autonomous music and structural listening.

## Expression of emotion

Among referential meanings, expression of emotion is the most common alternative. It has been discussed for centuries, as in the 'doctrine of the affections' (*Affektenlehre*) during the Baroque era and during Romanticism with its emphasis on expression of the subjective and emotional in music. Musicologist Deryck Cooke claimed that music is a language of emotions (1959, pp. 32–33), and philosopher Susanne Langer argued that 'music can *reveal* the nature of feelings with a detail and truth that language cannot approach' (1957, p. 235).

The following is a condensed review of empirical research on emotional expression in music. More complete reviews appear in Gabrielsson and Juslin (2003) and Gabrielsson and Lindström (2001).

### Stimuli

Generally, subjects listened to pieces of music or other tonal stimuli and reported perceived expression. Stimuli were usually recordings of selected pieces of music, sometimes live music (e.g., Hevner 1935; Rigg 1937a), or synthesized tone stimuli. Using real music means good ecological validity but limits conclusions regarding the effects of separate structural factors since they are usually confounded in musical contexts. An alternative that allows more definite conclusions is to systematically vary one or more structural factors in short sound sequences (e.g., Scherer and Oshinsky 1977); however, the ecological validity is then limited. A further, commendable alternative is to systematically vary different structural factors (tempo, mode, form, etc.) in real pieces of music (e.g., Hevner 1935, 1936, 1937; Juslin 1997; Kamenetsky *et al.* 1997; Karno and Konečni 1992; Lindström 2006; Peretz *et al.* 1998; Rigg 1940a, b; Schellenberg *et al.* 2000), and thus try to combine the advantages of the before-mentioned strategies.

### Response format

Subjects reported perceived expression using:

(a) free phenomenological descriptions (e.g. Huber 1923; Tagg 2006) or free choice of descriptive terms/adjectives (Imberty 1979);

(b) choice among descriptive terms provided by the investigator (e.g. Gundlach 1935; Hevner 1935; Wedin 1972a);

(c) combination of free descriptions and choice among descriptive terms (Rigg 1937a);

(d) ratings of how well selected descriptive terms apply to the music in question (e.g. Gabrielsson 1973; Nielzén and Cesarec 1982; Tillman and Bigand 1996;Wedin 1969); Collier (2007) used ranking of emotion words;

(e) nonverbal methods, for instance, pressing a pair of tongs to indicate perceived tension (Nielsen 1983, 1987), finger pressure on a sentograph to study expression of different emotions (Clynes and Nettheim 1982; Gabrielsson and Lindström 1995; de Vries 1991); or

(f) various technical or computerized devices to allow continuous recording of perceived tension (e.g., Fredrickson 2000; Fredrickson and Coggiola 2003; Krumhansl 1996; Madsen 1998; Madsen and Fredrickson 1993) or various emotions (Krumhansl 1997; Schubert 2004).

### Analysis

Free descriptions were subjected to content analysis. Listeners' choices among descriptive terms were analyzed regarding the frequency of chosen terms and inter-subject agreement. Ratings were usually analyzed by multivariate techniques, such as factor analysis, cluster analysis, multidimensional scaling or correspondence analysis in order to find a limited number of fundamental dimensions. Although the labels of these dimensions vary somewhat between different studies, they are in general interpretable in accordance with the valence-arousal model of emotion (Russell 1980; Russell and Feldman

Barrett 1999). There are also indications of a potency dimension (Kleinen 1968; cf. Osgood *et al.* 1957); a dimension related to body posture and gestures (Bigand *et al.* 2005); and a dimension reflecting stylistic differences between classical and popular music (Wedin 1969, 1972b). Further proposals include a distinction between energy arousal and tension arousal (Ilie and Thompson 2006; Schimmack and Grob 2000) and attempts to identify subtle dimensions beyond valence and activation (Collier 2007).

The relationship between the composed structure and perceived expression was studied by

(a) analyzing the musical score in relation to perceived expression (e.g. Gundlach 1935; Imberty 1979; Krumhansl 1996; Nielsen 1983; Thompson and Robitaille 1992);

(b) having musical experts judge the selected pieces with regard to structural properties (e.g. Kleinen 1968; Wedin 1972b);

(c) using various devices for measuring the acoustical properties of the music (e.g. Schubert 2004; Todd 1994); or

(d) using systematic manipulation of the musical stimuli and noting the effects on perceived expression (references given under Stimuli).

## Expression of emotion: effects of separate musical factors

The effects of separate musical factors are summarized in Gabrielsson and Lindström (2001). The following is an abbreviated version of this summary, with some new references added.

### Tempo, note density

Among factors affecting emotional expression, tempo—that is, perceived beat/pulse rate—is usually considered the most decisive.

Fast tempo may be associated with expressions of activity/excitement, happiness/joy/pleasantness, potency, surprise, flippancy/whimsicality, anger, uneasiness and fear.

Slow tempo may be associated with expressions of calmness/serenity, peace, sadness, dignity/solemnity, tenderness, longing, boredom, and disgust.

Both fast and slow tempo may thus be associated with many different expressions dependent on presence and level of other structural factors. However, in terms of the valence-arousal model, fast tempo is generally associated with high(er) activation, slow tempo with low(er) activation, while both of them may be associated with either positive or negative valence.

A related factor is note density, the number of notes per unit of time (e.g. per second). Tempo and note density may sometimes be additive—fast tempo combined with high note density results in still higher activation, slow tempo combined with low note density results in still lower activation—and may sometimes present an ambiguous picture as when high note density appears in combination with slow tempo (Gabrielsson 1988).

### Mode, key

Major mode may be associated with happiness/joy, minor mode with sadness. Major mode may also be associated with expressions as graceful, serene, and solemn, minor mode with expressions as dreamy, dignified, tension, disgust, and anger.

While differences between fast and slow tempo are mainly associated with difference in activation, differences between major and minor mode are thus mainly associated with difference in valence, positive or negative. However, major mode is not a necessary condition for perceived happiness; a piece in minor mode in fast tempo may very well sound happy (e.g., the *Badinerie* in JS Bach's Second Suite for Orchestra). Using scales as stimuli, Collier and Hubbard (2001a) claim that pitch height and direction of pitch movement (contour) may be more important than mode for expression of happiness.

The common belief that certain keys are associated with certain moods has no empirical support (Powell and Dibben 2005).

### Loudness

Loud music may be associated with expressions of intensity/power, excitement, tension, anger, and joy, soft music with softness, peace, tenderness, sadness, and fear. On the whole, loud music seems associated with high activation and potency, soft music with low activation and maybe submissiveness.

Large variations of loudness may suggest fear, small variations happiness or activity. Rapid changes in loudness may be associated with playfulness, pleading or fear, few or no changes with sadness, peace, and dignity.

## Timbre/spectrum

Tones with many higher harmonics may suggest potency, anger, disgust, fear, activity, or surprise, thus usually high activation. Tones with few, low harmonics may be associated with pleasantness, boredom, happiness or sadness, and tones with suppressed higher harmonics may suggest tenderness and sadness, thus usually low(er) activation.

## Pitch

High pitch may be associated with expressions as happy, graceful, serene, dreamy, exciting, surprise, potency, anger, fear, and activity. Low pitch may suggest sadness, dignity/solemnity, vigour, excitement, boredom, and pleasantness. Large pitch variation may be associated with happiness, pleasantness, activity, or surprise, small pitch variation with disgust, anger, fear, or boredom.

## Intervals

For melodic (successive) intervals some results indicate that large intervals sound more powerful than small ones, the octave is perceived as positive and strong and the minor second as the most sad interval. For harmonic (simultaneous) intervals, results concerning consonance and dissonance are similar to corresponding results under Harmony.

## Melody

### Melodic range

Wide melodic range may be associated with joy, whimsicality, uneasiness, and fear, narrow range with expressions as sad, dignified, sentimental, tranquil, delicate, and triumphant.

### Melodic direction (pitch contour)

Ascending melody may be associated with dignity, serenity, tension, happiness, fear, surprise, anger, and potency. Descending melody may be associated with expressions such as exciting, graceful, vigorous, sadness, boredom and pleasantness.

### Melodic motion

Stepwise motion may suggest dullness, intervallic leaps excitement; stepwise motion leading to melodic leaps may suggest peacefulness (Thompson and Robitaille 1992). Activity (sense of instability and motion) may be conveyed by a greater occurrence of minor seconds, tritones, and intervals larger than the octave. Potency (vigour and power) may be expressed by a greater occurrence of unisons and octaves (Costa *et al.* 2004).

## Harmony

Simple, consonant harmony may be associated with expressions as happy/gay, relaxed, graceful, serene, dreamy, dignified, serious, and majestic; complex and dissonant harmony with expression of excitement, tension, vigour, anger, sadness, and unpleasantness.

## Tonality

Melodies composed to sound joyful, dull, and peaceful were tonal, while angry melodies could be atonal. Sad and angry melodies used chromatic harmony (Thompson and Robitaille 1992).

## Rhythm

Regular/smooth rhythm may be perceived to express happiness, dignity, majesty, and peacefulness; irregular/complex rhythm may express amusement, uneasiness, and anger; varied rhythm may express joy. Firm rhythm may be associated with expressions of sadness, dignity, and vigour; flowing/fluent rhythm with expressions such as happy/gay, graceful, dreamy, and serene.

## Articulation

Staccato may be associated with gaiety, energy, activity, fear, and anger, legato with sadness, tenderness, solemnity, longing, and softness.

## Pauses

Perception of rests is dependent on musical context. Silences (rests) following tonal closure

were identified more quickly and perceived as less tense than silences following music lacking such closure (Margulis 2007).

### Musical form

High complexity (melodic/harmonic/rhythmic) is associated with tension or sadness, low complexity with relaxation, joy, or peace. High complexity combined with low dynamism may express melancholy and depression, high complexity combined with high dynamism may express anxiety and aggressiveness. Low complexity and average dynamism may be associated with positive emotions. Repetition, condensation, sequential development, and pauses may mean increased tension. Disruption of global form (changing the order of sections/movements in a musical work) may have little effect on (unknowing) listeners' perception of expression or general evaluation.

### Expression as function of many factors

As seen above, each structural factor may influence many different expressions. This means, conversely, that each expression may depend on many different structural factors as displayed in the extensive Table 26.2 in Gabrielsson and Juslin (2003). This table provides a systematic exposition of how different expressions—happiness/joy/gaiety, sadness/gloom, tension, potency, activity/energy/excitement, relaxation/calm/softness/peace, solemnity/dignity, anger, fear, tenderness/love, boredom, disgust, and surprise—are affected by different structural features. To take but one example: perceived tension is related to dissonance, high sound level, ascending melody, increased note density, harmonic complexity, rhythmic complexity, lack of melody, and various formal properties such as repetition, condensation and sequential development.

## Expression of emotion: summary and implications

The above review shows a complex picture. It seems that any structural factor may indicate a number of different emotional expressions. Results seem most clear-cut regarding effects of tempo/speed, loudness, and timbre/spectrum: increase in any of them results in higher activation, decrease in lower activation (by increase in timbre is meant relatively more higher harmonics). Tempo and loudness, in particular, seem to reign over most other factors, and generally the activation dimension seems more salient and easier to judge than the valence dimension. The effects of pitch height seem more ambiguous than those of loudness or tempo; for instance, while the effects of fast tempo and/or high loudness are associated with high activation, high pitch may be associated with both high and low activation.

The most distinct results thus concern effects of basic variables in human audition (loudness, timbre, pitch) and motion (tempo/speed). With regard to more typical 'musical' variables results are less clear, or there is less research available. The typical major-happy and minor-sad associations may be overruled by interaction with tempo (as exemplified under Mode above) and do not appear until the age of 7–8 years (e.g., Gerardi and Gerken 1995). There are still only tentative results regarding perceived expression of different intervals, and their effects may be modulated by differences in tempo, loudness, and other factors (Gabrielsson and Lindström 2001, p. 243). There is further little research on perceived expression of different kinds of melodic and rhythmic progression and practically none regarding harmonic progression. Aspects of musical form are frequently discussed in relation to structural listening, but their influence on perceived expression is little investigated.

Perceived expression is never determined by a single factor but is a function of many factors which may work in additive or interactive ways. Several studies identify interactions *post hoc*; for instance, intensity x pitch height and pitch height × tempo interactions regarding valence (Ilie and Thompson 2006), scale direction x tempo regarding happiness (Collier and Hubbard 2001b), and triple interaction rhythm x contour x melodic progression in ratings of happiness (Lindström 2006). However, there are only a few studies planned to study the presence (or not) of specific interactions (e.g. Langner and Goebl 2003; Makris and Mullet 2003; Schellenberg *et al.* 2000).

Most studies have utilized music with the 'same' expression throughout, whereas an essential part of music's appeal lies in variation of emotional expression. Many of the most intense expressions are related to (sudden or gradual) changes in, say, loudness (e.g. crescendo, diminuendo), tempo (accelerando, ritardando), timbre, pitch level, rhythm, melodic and harmonic progression, or in various combinations of these factors. Therefore an important topic for continued research is refinement of techniques for continuous recording of expression (Schubert 2001, 2002, 2004; Vines *et al.* 2005). Free phenomenological descriptions may be used as a complement. However, certain subtleties in musical expression may never be captured in a strict scientific manner, nor be adequately verbalized; they may only be accessible in direct experience.

On the other hand some results indicate that listeners are able to perceive the 'correct' expression even in excerpts lasting only one second or less (Bigand *et al.* 2005; Peretz *et al.* 1998). It also seems that, on the whole, perceived emotional expression is not, or only marginally, influenced by differences in gender or musical experience (e.g. Bigand *et al.* 2005; Bigand and Poulin-Charronnat 2006; Fredrickson 2000; Kallinen 2005; Makris and Mullet 2003; Robazza *et al.* 1994; Wedin 1969); there are, however, some indications that women may find music more emotionally expressive than men (e.g. Kamenetsky *et al.* 1997)

## Other referential meanings

Beyond emotions, music may also be perceived to express, reflect, or represent many other phenomena. (For discussion about definitions of expression, representation, or depiction, see Davies 1994; Scruton 1997).

### Program music

Program music flourished during Romanticism, as in works by Berlioz, Liszt, and Richard Strauss. However, composers were aware of music's limited possibilities to represent objects and events and rather referred to the feelings associated with the program (Benestad 1978, p. 233).

Listeners who do not know the program are as a rule unable to provide a 'correct' description of the intended events or situations but are better at perceiving the intended emotional expression (e.g. Brown 1981; Rigg 1937b). However, Osborne's (1989) subjects seemed to capture several aspects of the representations and expressions in Richard Strauss' *Alpine Symphony*.

One may also speak of program music in another sense: programs made up by the listeners themselves. Listeners may perceive a narrative in music they listen to ('this music is about '). Delis *et al.* (1978) argued that listeners construct a story in relation to the music in order to better remember it.

### Human character, personality, identity

In ancient Greece it was believed that music reflected human character and that these expressions would be imitated by the listener. Plato and Aristotle therefore gave advice concerning which modes, instruments, and rhythms ought to be used in education.

Clynes (1987, 1995) hypothesized that composers' character and 'presence' are reflected in unique patterns of timing and dynamics (composers' 'inner pulse'). Empirical tests regarding the existence of such patterns have provided mixed results (Clynes 1995; Repp 1990; Thompson 1989).

Tagg (2006) found striking differences between music perceived as 'male' or 'female'. In comparison with male tunes, female tunes were slower, more legato, had longer phrases, no repeated notes, static bass line, rare offbeats/syncopation, no brass or percussion, and the tonal idiom was classical/romantic compared to rock and jazz in male tunes. A 'male vs.female' factor, as well as some other human characteristics, was also observed in an exploratory study by Watt and Ash (1998). They claimed that music is assigned attributions that normally would be assigned to a person and went on to speculate that 'music creates a virtual person' (*ibid.* p. 49). In fact, many subjects perceive the music to reflect their own personality, how they think and feel; this provides a strong confirmation of their own identity (Gabrielsson and Lindström Wik 2003).

Music may reflect and be used to express not only personal but also social, cultural, and gender

identity (e.g. Dibben 2002; Folkestad 2002; Martin 2006; Ruud 1997; Tarrant *et al.* 2002). Immigrants are eager to preserve their own music in the new country, and teenagers may use certain music to mark their belonging to youth culture in general or to a particular group of people.

## Motion, dynamic forces

It is generally agreed that we may perceive motion in music. However, there is much discussion whether it refers to real motion, self-motion, virtual motion, or is entirely metaphorical. For discussions and overviews, see Bharucha *et al.* (2006, section 4), Clarke (2005), Eitan and Granot (2006), Jackendoff and Lerdahl (2006, section 4.3), and Shove and Repp (1995); see also Gjerdingen (1994) for a neural network model of apparent motion in music.

In music theory it is common to speak of tonal/harmonic motion in the sense that unstable tones/chords 'move' to (resolve to) stable tones/chords, typically in connection with perception of tension and release (e.g. Lerdahl and Krumhansl 2007). The concept of tension is thus related to structural listening but may also mediate emotional expression (Nielsen 1983).

Musicologist Ernst Kurth (1930) thought of music as a play of dynamic forces, gravitational forces between tones and chords, attracting or repelling each other; melody is perceived as a streaming force filled with kinetic energy. Larson and VanHandel (2005) defined three musical forces—musical gravity, musical magnetism, musical inertia—and operationalized them as short melodic patterns that were judged by listeners with regard to experienced 'strength' of the presented pattern completions. Results indicated that listeners' judgements were influenced by musical inertia and musical gravity.

While the above examples are mainly concerned with structural (intra-musical) listening, certain designations in scores (e.g. accelerando, ritardando, alla marcia, grazioso) refer to motion in a more concrete sense, and listeners find it natural to use motion labels as bouncing, dancing, flowing, rocking, swinging, and so on in descriptions of music (Gabrielsson 1973, 1988; Madison 2006). Tempo, articulation, and rhythm patterns seem obvious candidates underlying such motion characters (*cf.* 'galloping rhythm', 'fluent rhythm ', 'uneven rhythm', etc.), but there is still little systematic research on what factors in musical structure affect perceived motion. Eitan and Granot (2006) asked subjects to describe imagined motion of a visualized person when listening to brief melodic figures varying in dynamics, pitch contour, pitch intervals, attack rate, motivic pace, and articulation. The manipulations affected several variables in the listeners' motion imagery, such as lateral and vertical direction, distance change, speed change, and energy level of motion.

It is also generally agreed that there is a close link between motion and emotion: 'Motion is heard in music, and that motion presents emotion characteristics' (Davies 1994, p. 229), 'musical affect arises in large part from its relation to physical patterns of posture and gesture' (Jackendoff and Lerdahl 2006, p. 65). Jackendoff and Lerdahl point out that many terms used for description of emotional expression are also used to describe gestures and postures and discuss possible underlying factors in musical structure and performance. (One may add that it is sometimes felt more natural and easier to describe music in terms of motion rather than emotion.) Here, then, is a domain awaiting pertinent empirical research.

## Social conditions and organization

Philosopher Theodor W. Adorno speculated that music's formal properties may reflect organization and forces of society; for instance, tensions in society would be reflected in rupture of musical form (Benestad 1978, p. 329). While Adorno never sought empirical evidence, Lomax's (1968) large-scale investigation of folk song styles in 233 cultures throughout the world provides many examples of how music can reflect economic and social conditions and attitudes; for comments, see Dowling and Harwood (1986, p. 226).

## Religious faith, transcendental phenomena

Music as expression of religious faith is a worldwide phenomenon. In Western music history, examples range from simple community singing

to masterpieces such as J. S. Bach's passions, Handel's oratories, or masses by Haydn, Mozart, and Beethoven.

Philosopher Arthur Schopenhauer regarded music as an expression of the Will, a never-ending striving toward existence that can never be quite fulfilled and therefore causes human suffering, and Friedrich Nietzsche viewed music as expression of Dionysian character and the will to life and power; see further in Benestad (1978) and Budd (1985).

## Further comments on structure

There is only scattered research on structural variables connected with the referential meanings treated in this section. However, since many of these meanings—program music, motion, character, identity, religious faith—also have obvious emotional connotations, it is reasonable to expect that the structural variables underlying emotional expression are largely pertinent to them as well.

Finally, we should not forget that many people enjoy music simply for its expression of beauty. However, there seem to be no empirical investigations in music psychology directly focusing on structural factors underlying perceived beauty in music.

# References


Benestad F (1978). *Musik och tanke. Huvudlinjer i musikestetikens historia från antiken till vår egen tid [Music and thought. Main lines in the history of musical aesthetics from antiquity to our time]*. Rabén and Sjögren, Stockholm.

Bharucha JJ, Curtis M and Paroo K. (2006). Varieties of musical experience. *Cognition*, **100**, 131–172.

Bigand E and Poulin-Charronnat B. (2006). Are we 'experienced listeners? A review of the musical capacities that do not depend on formal musical training. *Cognition*, **100**, 100–130.

Bigand E, Vieillard S, Madurell F, Marozeau J and Dacquet A (2005). Multidimensional scaling of emotional responses to music: the effect of musical expertise and of the duration of the excerpts. *Cognition and Emotion*, **19**, 1113–1139.

Brown R (1981). Music and language. In *Documentary report of the Ann Arbor symposium. National symposium on the applications of psychology to the teaching and learning of music*, 233–265. Music Educators National Conference, Reston, VA.

Budd M (1985). *Music and the emotions. The philosophical theories*. Routledge, London.

Clarke EF (2005). *Ways of listening. An ecological approach to the perception of musical meaning*. Oxford University Press, New York.

Clynes M (1987). What can a musician learn about music performance from newly discovered microstructure principles (PM, PAS)? In A Gabrielsson, ed., *Action and perception in rhythm and music*, 201–233. Publications issued by the Royal Swedish Academy of Music, No. 55, Stockholm.

Clynes M (1995). Microstructural musical linguistics: composers' pulses are liked most by the best musicians. *Cognition*, **55**, 269–310.

Clynes M and Nettheim N (1982).The living quality of music. Neurobiologic patterns of communicating feeling. In M Clynes, ed., *Music, mind, and brain. The neuropsychology of music*, 47–82. Plenum Press, New York.

Collier GL (2007). Beyond valence and activity in the emotional connotations of music. *Psychology of Music*, **35**, 110–131.

Collier GL and Hubbard TL (2001a). Musical scales and evaluations of happiness and awkwardness: effects of pitch, direction, and scale mode. *American Journal of Psychology*, **114**, 355–375.

Collier GL and Hubbard TL (2001b). Judgements of happiness, brightness, speed and tempo change of auditory stimuli varying in pitch and tempo. *Psychomusicology*, **17**, 36–55.

Cooke D (1959). *The language of music*. Oxford University Press, Oxford.

Costa M, Fine P and Ricci Bitti PE (2004). Interval distribution, mode, and tonal strength of melodies as predictors of perceived emotion. *Music Perception*, **22**, 1–14.

Davies S (1994). *Musical meaning and expression*. Cornell University Press, Ithaca, NY.

Delis D, Fleer J and Kerr P (1978). Memory for music. *Perception and Psychophysics*, **23**, 215–218.

Dibben N (2002). Gender identity and music. In RAR MacDonald, DJ Hargreaves and D Miell, eds, *Musical identities*, 117–133. Oxford University Press, Oxford.

Dowling WJ and Harwood DL (1986). *Music cognition*. Academic Press, Orlando FL.

Eitan Z and Granot RY (2006). How music moves: musical parameters and listeners' images of motion. *Music Perception*, **23**, 221–247.

Evans P and Schubert E (2008). Relationships between expressed and felt emotion in music. *Musicae Scientiae*, **12**, 75–99.

Fisk J (ed.) (1997). *Composers on music*. Northeastern University Press, Boston, MA.

Folkestad G (2002). National identity and music. In RAR MacDonald, DJ Hargreaves and D Miell, eds, *Musical identities*, 151–162. Oxford University Press, Oxford.

Fredrickson WE (2000). Perception of tension in music: musicians versus nonmusicians. *Journal of Music Therapy*, **37**(1), 40–50.

Fredrickson WE and Coggiola JC (2003). A comparison of music majors' and nonmajors' perceptions of tension for two selections of jazz music. *Journal of Research in Music Education*, **51**, 259–270.

Gabrielsson A (1973). Adjective ratings and dimension analysis of auditory rhythm patterns. *Scandinavian Journal of Psychology*, **14**, 244–260.

Gabrielsson A (1988). Timing in music performance and its relations to music experience. In JA Sloboda, ed., *Generative processes in music: the psychology of performance, improvisation, and composition*, 27–51. Clarendon Press, Oxford.

Gabrielsson A (2002). Emotion perceived and emotion felt: same or different? *Musicae Scientiae, Special Issue 2001–2002: Current Trends in the Study of Music and Emotion*, 123–147.

Gabrielsson A and Juslin PN (2003). Emotional expression in music. In RJ Davidson, KR Scherer and HH Goldsmith, eds, *Handbook of affective sciences*, 503–534. Oxford University Press, New York.

Gabrielsson A and Lindström E (1995). Emotional expression in synthesizer and sentograph performance. *Psychomusicology*, **14**, 94–116.

Gabrielsson A and Lindström E (2001). The influence of musical structure on emotional expression. In PN Juslin, JA Sloboda, eds, *Music and emotion: theory and research*, 223–248. Oxford University Press, New York.

Gabrielsson A and Lindström Wik S (2003). Strong experiences related to music: a descriptive system. *Musicae Scientiae*, **7**, 157–217.

Gerardi GM and Gerken L (1995). The development of affective responses to modality and melodic contour. *Music Perception*, **12**, 279–290.

Gjerdingen RO (1994). Apparent motion in music? *Music Perception*, **11**, 335–370.

Gundlach RH (1935). Factors determining the characterization of musical phrases. *American Journal of Psychology*, **47**, 624–644.

Hanslick E (1989/1854). *Vom musikalisch Schönen*. Breitkopf and Härtel, Wiesbaden.

Hevner K (1935). The affective character of the major and minor modes in music. *American Journal of Psychology*, **47**, 103–118.

Hevner K (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology*, **48**, 246–268.

Hevner K (1937). The affective value of pitch and tempo in music. *American Journal of Psychology*, **49**, 621–630.

Huber K (1923). *Der Ausdruck musikalischer Elementarmotive [Expression in elementary musical motives]*. Johann Ambrosius Barth, Leipzig.

Ilie G and Thompson WF (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception*, **23**, 319–329.

Imberty M (1979). *Entendre la musique*. Dunod, Paris.

Jackendoff R and Lerdahl F (2006). The capacity for music: what is it, and what's special about it? *Cognition*, **100**, 33–72.

Juslin PN (1997). Perceived emotional expression in synthesized performances of a short melody: capturing the listener's judgment policy. *Musicae Scientiae*, **1**, 225–256.

Juslin PN and Laukka P (2004). Expression, perception, and induction of musical emotions: a review and a questionnaire study of everyday listening. *Journal of New Music Research*, **33**, 217–238.

Kallinen K (2005). Emotional ratings of music excerpts in the western art music repertoire and their self-organization in the Kohonen neural network. *Psychology of Music*, **33**, 373–393.

Kallinen K and Ravaja N (2006). Emotion perceived and emotion felt: same and different. *Musicae Scientiae*, **10**, 191–213.

Kamenetsky SB, Hill DS and Trehub SE (1997). Effect of tempo and dynamics on the perception of emotion in music. *Psychology of Music*, **25**, 149–160.

Karno M and Konečni VJ (1992). The effects of structural interventions in the first movement of Mozart's Symphony in G minor K. 550 on aesthetic preference. *Music Perception*, **10**, 63–72.

Kleinen G (1968). *Experimentelle Studien zum musikalischen Ausdruck [Experimental studies on musical expression]*. Universität Hamburg, Hamburg.

Konečni VJ, Brown A and Wanic RA (2008). Comparative effects of music and recalled life-events on emotional state. *Psychology of Music*, **36**, 289–308.

Krumhansl CL (1996). A perceptual analysis of Mozart's Piano Sonata K. 282: segmentation, tension, and musical ideas. *Music Perception*, **13**, 401–432.

Krumhansl CL (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology*, **51**, 336–352.

Kurth E (1930/1947). *Musikpsychologie*. Verlag Krompholz, Bern.

Langer SK (1957). *Philosophy in a new key*, 3rd edn. Harvard University Press, Cambridge, Massachusetts.

Langner J and Goebl W (2003). Visualizing expressive performance in tempo–loudness space. *Computer Music Journal*, **27**(4), 69–83.

Larson S and VanHandel L (2005). Measuring musical forces. *Music Perception*, **23**, 119–136.

Lerdahl F and Krumhansl CL (2007). Modeling musical tension. *Music Perception*, **24**, 329–366.

Lindström E (2006). Impact of melodic organization on perceived structure and emotional expression in music. *Musicae Scientiae*, **10**, 85–117.

Lomax A (1968). *Folk song style and culture*. American Association for the Advancement of Science, Washington, DC.

Madison G (2006). Experiencing groove induced by music: consistency and phenomenology. *Music Perception*, **24**, 201–208.

Madsen CK (1998). Emotion versus tension in Haydn's Symphony No. 104 as measured by the two-dimensional continuous response digital interface. *Journal of Research in Music Education*, **46**, 546–554.

Madsen CK and Fredrickson WE (1993). The experience of musical tension: a replication of Nielsen's research using the continuous response digital interface. *Journal of Music Therapy*, **30**, 46–63.

Makris I and Mullet E (2003). Judging the pleasantness of contour-rhythm-pitch-timbre musical combinations. *American Journal of Psychology*, **116**, 581–611.

Margulis EH (2007). Silences in music are musical not silent: an exploratory study of context effects on the experience of musical pauses. *Music Perception*, **24**, 485–506.

Martin PJ (2006). Music, identity, and social control. In S Brown and U Volgsten, eds, *Music and manipulation. On the social uses and social control of music*, 57–73. Berghahn Books, New York.

Meyer LB (1956). *Emotion and meaning in music.* University of Chicago Press, Chicago, IL.

Nielsen FV(1983). *Oplevelse av musikalsk spœnding [Experience of musical tension].* Akademisk Forlag, Copenhagen. (Includes summary in English.)

Nielsen FV (1987). Musical 'tension' and related concepts. In TA Sebeok and J Umiker-Sebeok, eds, *The semiotic web 86. An international yearbook*, 491–513. Mouton de Gruyter, Berlin.

Nielzén S and Cesarec Z (1982). Emotional experience of music as a function of musical structure. *Psychology of Music*, **10**, 7–17.

Osborne JW (1989). A phenomenological investigation of the musical representation of extra-musical ideas. *Journal of Phenomenological Psychology*, **20**, 151–175.

Osgood CE, Suci GJ and Tannenbaum PH (1957). *The measurement of meaning.* University of Illinois Press, Urbana, IL.

Peretz I, Gagnon L and Bouchard B (1998). Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, **68**, 111–141.

Powell J and Dibben N (2005). Key-mood association: a self perpetuating myth. *Musicae Scientiae*, **9**, 289–311.

Repp BH (1990). Further perceptual evaluations of pulse microstructure in computer performances of classical piano music. *Music Perception*, **8**, 1–33.

Rigg MG, (1937a). Musical expression: an investigation of the theories of Erich Sorantin. *Journal of Experimental Psychology*, **21**, 442–455.

Rigg MG (1937b). An experiment to determine how accurately college students can interpret the intended meanings of musical compositions. *Journal of Experimental Psychology*, **21**, 223–229.

Rigg MG (1940a). The effect of register and tonality upon musical mood. *Journal of Musicology*, **2**, 49–61.

Rigg MG (1940b). Speed as a determiner of musical mood. *Journal of Experimental Psychology*, **27**, 566–571.

Robazza C, Macaluso C and D'Urso V (1994). Emotional reactions to music by gender, age, and expertise. *Perceptual and Motor Skills*, **79**, 939–944.

Russell JA (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, **39**, 1161–1178.

Russell JA and Feldman Barrett L (1999). Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of Personality and Social Psychology*, **76**, 805–819.

Ruud E (1997). *Musikk og identitet [Music and identity].* Universitetsforlaget, Oslo.

Schellenberg EG, Krysciak AM and Campbell RJ (2000). Perceiving emotion in melody: interactive effects of pitch and rhythm. *Music Perception*, **18**, 155–171.

Scherer KR and Oshinsky JS (1977). Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion*, **1**, 331–346.

Schimmack U and Grob A (2000). Dimensional models of core affect: a quantitative comparison by means of structural equation modeling. *European Journal of Personality*, **14**, 325–345.

Schubert E (2001). Continuous measurement of self-report emotional response to music. In PN Juslin, JA Sloboda, eds, *Music and emotion: theory and research*, 393–414. Oxford University Press, New York.

Schubert E (2002). Correlation analysis of continuous emotional response to music: correcting for the effects of serial correlation. *Musicae Scientiae, Special issue 2001–2002: Current trends in the study of music and emotion*, 213–236.

Schubert E (2004). Modeling perceived emotion with con tinuous musical features. *Music Perception*, **21**, 561–585.

Scruton R (1997). *The aesthetics of music.* Oxford University Press, New York.

Shove P and Repp BH (1995). Musical motion and performance: theoretical and empirical perspectives. In J Rink, ed., *The practice of performance. Studies in musical interpretation*, 55–83. Cambridge University Press, Cambridge.

Tagg P (2006). Music, moving images, semiotics, and the democratic right to know. In S Brown and U Volgsten, eds, *Music and manipulation. On the social uses and social control of music*, 163–186. Berghahn Books, New York.

Tarrant M, North AC and Hargreaves DJ (2002). Youth identity and music. In RAR MacDonald, DJ Hargreaves, D Miell, eds, *Musical identities*, 134–150. Oxford University Press, Oxford.

Thompson WF (1989). Composer-specific aspects of musical performance: an evaluation of Clynes's theory of pulse for performances of Mozart and Beethoven. *Music Perception*, **7**, 15–42.

Thompson WF and Robitaille B (1992). Can composers express emotions through music? *Empirical Studies of the Arts*, **10**, 79–89.

Tillman B and Bigand E (1996). Does formal musical structure affect perception of musical expressiveness? *Psychology of Music*, **24**, 1–17.

Todd N (1994). The auditory 'primal sketch': a multiscale model of rhythmic grouping. *Journal of New Music Research*, **23**, 25–70.

Vines BW, Nuzzo RL and Levitin DJ (2005). Analyzing temporal dynamics in music: differential calculus, physics, and functional data analysis techniques. *Music Perception*, **23**, 137–152.

deVries B (1991). Assessment of the affective response to music with Clynes's sentograph. *Psychology of Music*, **19**, 46–64.

Watt RJ and Ash RL (1998). A psychological investigation of meaning in music. *Musicae Scientiae*, **2**, 33–53.

Wedin L (1969). Dimension analysis of emotional expression in music. *Swedish Journal of Musicology*, **51**, 119–140.

Wedin L (1972a). Multidimensional scaling of emotional expression in music. *Swedish Journal of Musicology*, **54**, 1–17.

Wedin L (1972b). A multidimensional study of perceptual–emotional qualities in music. *Scandinavian Journal of Psychology*, **13**, 241–257.

# CHAPTER 14

# Aesthetics

David Huron

## Introduction

Aesthetics is commonly defined as the study of beauty, and its opposite, ugliness. Some philosophers conceive of aesthetics as applying solely to the arts or to artistic experience. However, most aesthetic philosophers construe the discipline as applying more broadly to beauty and ugliness in general. The term 'aesthetics' first appeared in a book by Alexander Baumgarten in 1735, yet philosophical discussions of beauty extend back thousands of years. Commentaries on 'good' and 'bad' music can be found in both ancient Greek and ancient Chinese sources.

Beginning in the 1960s, the field of cognitive science became increasingly influential in the philosophy of mind. While much of this influence relates to the nature of thought, reasoning, and consciousness, the impact of cognitive science has expanded to other areas of philosophy, including aesthetics. In this chapter, no effort will be made to provide a comprehensive survey of ideas related to musical aesthetics. Instead, it focuses more narrowly on how cognitive science has influenced—and continues to influence—modern thinking in musical aesthetics. This chapter concludes by arguing that cognitive neuroscience is poised to overtake philosophical aesthetics: rather than *influencing* aesthetic philosophy, aesthetic philosophy is receding to a sideline 'advisory' role, while cognitive science takes an unaccustomed leadership position.

## Philosophical antecedents

In the West, the most influential writer on aesthetics has been the German philosopher Immanuel Kant (1724–1804). Kant laid out two pillars in Western aesthetics. First, he distinguished aesthetic pleasure from other forms of pleasure. Aesthetic pleasure is not some other type of pleasure in disguise: for example, art is not the sublimation of food, sex, warmth, companionship, or some combination of other existing pleasures. Second, Kant argued that aesthetic emotions are 'disinterested'; when we experience an aesthetic pleasure, there are no utilitarian or ulterior motives underlying this experience. A mother's appreciation of the beauty of her daughter cannot be regarded as a purely aesthetic appreciation, since her experience is apt to be tainted by parental pride. Unlike garden-variety emotions, for Kant, aesthetic emotions serve no practical purpose.

In music, Kant's ideas were developed and extended by the famed Austrian music critic Eduard Hanslick (1825–1904). In *Vom Musikalisch-Schönen* (On the beautiful in music) (1854) Hanslick proposed a highly influential view that would later prove to be compatible with a core concept in cognitive science. Hanslick challenged the (then) prevailing beliefs that music somehow represents or expresses feelings. He proposed that sensation is imaginatively interpreted by the listener, and that this aesthetic contemplative process then leads to possible emotions. In short, aesthetic judgement precedes and leads to aesthetic feelings. Until recently, Hanslick's views have defined the principal parameters in debates concerning musical aesthetics. All major philosophers in the aesthetics of music have started by engaging with Hanslick's ideas: see for example Susanne Langer (1942), Peter Kivy (1990), Roger Scruton (1997), Jerrold Levison (1990, 2003) and Stephen Davies (1994).

## Experimental antecedents

Most aesthetics scholars have followed a rationalist approach that emphasizes exegesis of existing theories and critical philosophical discussion. Independent of this philosophical tradition, aesthetics questions have also been addressed by empirically oriented scholars pursuing various avenues of scientific experimentation. In fact, the advent of experimental aesthetics coincides with the start of modern experimental psychology. One of the founders of modern experimental psychology, Wilhelm Wundt (1832–1920) carried out a number of experiments related to aesthetic experience. Wundt (1863) showed that arousal is related to stimulus complexity and proposed that aesthetic pleasure is evoked when the art object is optimally complex—neither too simple nor too complicated.

Aesthetics experiments were similarly carried out by another early experimental psychologist, Gustav Fechner (1801–1887). In his *Vorschule der Ästhetik* (Elementary aesthetics) Fechner suggested that 'half of aesthetics' originates in learned associations (1876, pp. 89–90). A lullaby may evoke feelings of comfort solely because of a learned association between the tune and formative experiences of being comforted by a caregiver. A German émigré to the United States, Max Meyer (1873–1967), carried out a key experiment showing that listeners prefer familiar music over unfamiliar music. Meyer (1903) showed that repeated listening to a novel musical work tends to increase reported satisfaction. While the Gestalt psychologists carried out relatively few experiments, they did produce a number of empirical demonstrations of various perceptual principles and suggested that these principles can be observed in both visual art and music (see, e.g. Koffka 1935).

A long-standing preoccupation in experimental musical aesthetics relates to the perception of consonance and dissonance. The ancient Greeks observed that euphonious sonorities appear to involve component tones whose frequencies are related by simple integer ratios. Both Hermann von Helmholtz (1877) and Carl Stumpf (1883) speculated about possible physiological origins for consonance and dissonance. Robert Lundin (1947) proposed that consonance and dissonance preferences are cultural in origin and arise primarily from the relative frequency of exposure to different sound combinations and the prevailing attitudes toward those sounds within a given social environment. By the 1970s, ample evidence pointing to both physiological and cultural factors influencing judgements of the euphoniousness of a sonority had accumulated. Regarding the physiological basis, a significant breakthrough occurred in the 1960s with the work of Donald Greenwood (1961) and Reinier Plomp and Willem Levelt (1965). These and subsequent researchers showed that at least some of the phenomenal experience musicians call 'dissonance' can be traced to mechanical interference within the organ of hearing. Regarding the influence of enculturation and familiarity on stimulus preference, mainstream psychology remained unaware of the pioneering work of Meyer and Lundin. Experimental research by Wilson (1975, 1979) and others unwittingly re-established what Zajonc (1980) later dubbed the 'mere exposure effect'. Over the past century, over 200 experiments have shown that familiarity has a marked impact on preference—especially when the listener is unaware that the sounds are familiar.

Among several notable empirically grounded volumes on the science of aesthetics, perhaps the most comprehensive attempt is found in the work of the Canadian psychologist Daniel Berlyne. In his 1971 book, *Aesthetics and psychobiology*, Berlyne extended Wundt's observations concerning the relationship between complexity, arousal, and pleasure. In particular, Berlyne linked Wundt's observations with contemporary neurophysiological research on pleasure. Berlyne proposed how the phenomenal experience of pleasure might be related to arousal and complexity. He distinguished two pleasure-inducing effects: one source of pleasure arises from moderate increases in arousal, while a second source of pleasure arises from inhibition or reduction of arousal from an uncomfortably high level. Together, these sources result in an inverted-U function when hedonic value (pleasure) is plotted against arousal level.

More recent experimental investigations by Vladimir Konečni and his colleagues have tested the common assumption that large-scale formal structures contribute to the aesthetic experience. For example, scrambling the order of variations

in JS Bach's *Goldberg Variations* or the sections of a sonata-allegro movement by Mozart does not produce less aesthetically pleasing experiences for listeners (Gotlieb and Konečni 1985; Karno and Konečni 1992). Similarly, Nicholas Cook (1987) carried out an experiment suggesting that beginning and ending in the same key ('tonal closure') fails to evoke a greater sense of completion, coherence or pleasure for passages that are longer than roughly 2 minutes. Where Konečni and Cook have questioned existing ideas about form, other research has identified large-scale structures that occur cross-culturally. For example, Ollen and Huron (2004) carried out an analysis of patterns of musical repetition in music from 50 cultures, and found a cross-cultural preference for early repetition (e.g., AAABAABA rather than ABAABAAA). Huron (2006) noted that this pattern is consistent with the twin goals of pursing increasing predictability while avoiding habituation.

Despite a long history, the experimental tradition in aesthetics has had comparatively little influence among aesthetic philosophers. This is especially true in the case of musical aesthetics, where the influence of empirical approaches has been negligible or irrelevant. Among many music scholars, empirical ideas regarding musical pleasure have been regarded as naive, and indeed, some of the claims justify this wary reception.

## Cognitive revolution

In contrast to the limited impact of experimental aesthetics on mainstream musical aesthetics, the cognitive revolution has proved to be an inspiration to a number of aesthetic philosophers (e.g., Raffman 1993; Madell 2002; Butler 2004; Robinson 2005). Before discussing the relationship between cognitive science and aesthetics, it is useful to distinguish two different conceptions of cognitive science.

Prior to the cognitive revolution, Anglo-American psychology was dominated by behaviourism, a perspective that emphasized sensation, motor behaviour, and learning. Influenced by linguistics and computer science, cognitive psychologists juxtaposed themselves against behaviourism by emphasizing the thoughtful and imaginative aspects of mental life. For example, rather than viewing emotions as types of reflexes, cognitive psychologists tended to view emotions as originating in cognitive appraisals (Arnold 1960; Averill 1980). That is, emotions were evoked by conscious or unconscious evaluations that depend on underlying conceptualizations. According to this view, judgement precedes emotion.

Until the mid 1990s, cognitive science was typically regarded as opposed to more biological approaches to understanding the mind. Early cognitive science also excluded or sidelined the role of affect or emotion. However, over the past two decades, the sciences of the mind have been converging into what Joseph LeDoux has referred to as 'the united kingdom of cognition and emotion' (1996, p. 39). Cognitive science has moved away from a strictly cognitivist perspective to embrace insights from ethology, evolutionary psychology, psychophysiology, genetics, biochemistry, and neuroimaging. The cognitive perspective remains an important component of contemporary cognitive science, but it no longer occupies the core to the exclusion of other perspectives. In short, cognitive science has become increasingly less cognitive.

This transformation of the field has repercussions for understanding the cognitive science of aesthetics. We can talk about two schools: the first might be called the 'old school' or *cognitivist* approach, while the second might be called the 'new school' or *cognitive neuroscience* approach. This distinction is important because there are two contrasting aesthetic traditions that both claim to be influenced by cognitive science.

The old school, or cognitivist view, has been especially congenial with Hanslick's formalist aesthetics. The traditional cognitivist view has been that cognitive appraisal precedes emotion: conscious or unconscious evaluation is the immediate source of evoked emotion. This view is consistent with Hanslick's argument that sensations are imaginatively interpreted, and that a contemplative process subsequently leads to possible emotions. Music evokes emotion only after passing through a cognitive/interpretive filter. Musical meaning and aesthetic evaluation precede musical affect.

This cognitivist view has become especially popular in literary aesthetics. British literary scholar Christopher Butler (2004), for example,

has applied a cognitivist perspective in his compelling account of the pleasures evoked by narrative. However, such cognitivist approaches appear to be more useful in accounting for literature and representative art than for abstract art or instrumental music. For those art forms in which narrative content is foremost, the evoked pleasures do seem to align well with cognitive–appraisalist conceptions of emotion.

However, the cognitivist conception of emotion has been criticized by both psychologists and philosophers. Experimental studies have provided ample demonstrations that at least some emotions can be evoked without cognitive appraisals (Zajonc 1980; see Cornelius 1996, for review). Criticisms of a purely cognitivist perspective on emotions have been echoed by some music philosophers, notably Malcolm Budd (1985) and Geoffrey Madell (2002).

The existing experimental evidence suggests that emotions can be evoked with or without the involvement of cognitive appraisals. On the one hand, emotions can be evoked through conscious ruminations (as, for example, when jealousy is evoked by interpreting a telephone bill as suggesting that one's spouse has been talking to a former love-interest). However, emotions can also be evoked without the intervention of conscious thought (as in a startle response evoked by the slamming of a door or hearing a tone of voice suggestive of aggression).

Such unconscious/automatic responses can also be observed in behaviours that would normally be regarded as 'higher level' mental processes such as sympathy and empathy. Several philosophers (e.g. Roger Scruton) have argued that responding sympathetically to (say) grief is not the same as feeling grief. However, modern neuroscience appears to contradict this claim. Watching someone cut their finger can generate 'gut feelings' in the viewer that closely mimic the negative feelings experienced by the person whose finger is injured. The discovery of so-called 'mirror' neurons suggests that some complex emotions can be experienced with little or no cognitive mediation (Rizzolatti and Craighero 2004). In short, cognitive neuroscience implies that there exist both cognitive and non-cognitive pathways to the evoking of emotion.

## Neo-hedonism

The idea that the principal appeal of art is the evoking of pleasure (hedonism) has not been a popular idea among Western aesthetics philosophers. However, recent cognitive neuroscience has given new life to this old idea. Perhaps the most important achievement of cognitive science has been the discovery that the mind, like the body generally, holds a series of specialized structures (Fodor 1983). A popular metaphor is that the mind is structured like a Swiss army knife. Many of the debates that have preoccupied aesthetics philosophers arise from the empirically incorrect assumption that the mind is unitary and homogeneous (see also Griffiths 1997). Since the mind is capable of propositional thinking, it is easy to suppose therefore that musical experience must arise from propositional thought; since the mind forms associations, musical experience must arise from deciphering associations, and so on.

In modern cognitive neuroscience, such claims are regarded as *cumulative* rather than mutually exclusive. For any given stimulus, the modular mind applies propositional, associational, representational, empathetic, narrative, and other mental processes simultaneously. While the experience of beauty and ugliness can involve cognitive appraisals, not all such experiences *require* cognitive appraisal. In short, the cognitive appraisals of the sort envisioned by Hanslick's followers are sufficient though not necessary conditions for the experience of musical beauty.

## Plural pleasures

The idea of concurrent parallel mental processes has transformed our understanding of the phenomenon of pleasure. There are many behaviours that can evoke pleasure, such as scratching an itch, quenching a thirst, solving a puzzle, successfully predicting a future event, feeling virtuous, emptying one's bladder, conversing with a friend, receiving praise, putting cold hands in warm water, and so on. Physiologists have begun to trace the unique neurological pathways associated with different pleasures, from the pleasure of eating chocolate to the 'runner's high'.

Pleasure centres in the brain were discovered accidentally half a century ago by James Olds and Peter Milner (1954). In addition to these neuroanatomical discoveries, a number of endogenous molecules have been implicated in the experience of pleasure. These include dopamine, oxytocin, serotonin, alpha-, beta-, and gamma-endorphins, alpha-, and beta-neoendorphin, dynorphin A and B, big dynorphin, methionine enkaphalin, leucine enkaphalin, and others (e.g. Pert 1997). Each endogenous molecule evokes a subtly different form of pleasure, and each is released under a number of unique circumstances. For example, oxytocin is released in various interpersonal situations associated with pleasure, including hugging, romantic eye contact, breast feeding, and sex.

In crafting a pleasurable event, people rarely confine themselves to evoking one form of pleasure. The most common pattern of behaviour is to combine several pleasures into a single experience: a person might drink a beer *and* smoke a cigarette *and* converse with friends *and* watch a football game—all at the same time. Current neuroscience indicates that the pleasures evoked by imbibing alcohol, inhaling nicotine, reinforcing social bonds, and prevailing over a perceived enemy have different neurological origins. Yet all may be evoked concurrently in the environment of the sports bar.

A useful metaphor for this hedonic pluralism is the dinner party. You might decide to cook an especially nice meal, but it is unlikely that you will stand alone at the kitchen counter feasting on your carefully prepared Coquille Saint Jacques. Instead, you will probably invite a friend to share in the meal (social pleasure), purchase flowers to embellish the table setting (visual pleasure), light scented candles (olfactory pleasure), put on recorded music (auditory pleasure), and so on. What begins as the crafting of a gustatory pleasure quickly expands into a multifaceted hedonic experience. When given the opportunity, we heap pleasure upon pleasure upon pleasure.

The discovery of multiple pleasure pathways has repercussions for philosophical debates concerning hedonism. Among aesthetic philosophers, the standard rebuttal of hedonism is that it suggests that all pleasurable experiences can be reduced to a single value (see e.g. Kagan 1998). Hedonism implies that pleasures are interchangeable—implying, for example, that a sexual orgasm is equivalent to so many chocolate cookies. The most common argument against hedonism is that the positive emotions evoked by feeling virtuous are simply incommensurate with a sensory pleasure, such as viewing a flower. Geoffrey Madell summarizes this anti-hedonic argument by noting that the pleasure of listening to good music is phenomenologically different from the pleasure of gorging on junk food (2002 p. 89). However, two discoveries have given new life to the hedonic argument: (1) evolutionary psychologists have offered compelling arguments that all emotions (including jealousy, shame, pride, etc.) are evolutionary adaptations that promote survival and procreation, and (2) neurophysiologists have assembled evidence that the brain contains multiple pleasure systems. The neo-hedonic rejoinder to the classic objection to hedonism is that the biology of pleasure is not unidimensional and has multiple independent sources. Pleasures can be equated only when they engage the same endocrine or neurotransmitter pathways.

If pleasures are multidimensional, it is likely that musically evoked pleasure is itself multifaceted. Consider some of the ways by which sounds are thought to evoke pleasure:

- Listeners prefer stereo reproduction over monaural reproduction.
- Familiar sounds are preferred over unfamiliar sounds (Meyer 1903; Zajonc1980).
- Novelty-seeking is a rewarded behaviour in many circumstances (Berlyne 1971).
- Infant-directed singing has many features in common with infant-directed speech and these features are known to be preferred by infants (Unyk *et al.* 1992).
- The traditional practice of voice-leading is thought to capitalize on brain rewards for successful parsing of auditory scenes (Huron 2001).
- People experience pleasure from displays of extraordinary musical skill or virtuosity (Kubovy 1999).
- Experienced listeners regularly take pleasure in recognizing musical quotations or allusions to other works.

- Musically induced 'shivers' or frisson is reported by listeners as distinctly pleasurable (Sloboda 1991; Panksepp 1995; Gabrielsson and Lindstrom 1993; and others). Huron (2006) has suggested how such experiences might evoke pleasure.
- Listeners are consoled by and take pride in music whose style or genre is consistent with a sense of self-identity or social belonging.

This list represents only a partial catalogue of plausible pleasures that might be evoked (in various combinations) by music. As with studies investigating the specific pleasures involved in chocolate consumption or the 'runner's high', it seems likely that future research will trace the particular neurological pathways involved in each of the various forms of musically evoked pleasure. It appears that musical sounds are capable of activating multiple pleasure pathways in the brain. In the manner of the dinner party, musicians can assemble a unique mixture of pleasures into a single musical experience.

Indirect evidence in support of this 'plural pleasures' hypothesis can be found in an experiment by Avram Goldstein (1980). Goldstein exposed listeners to frisson-inducing musical passages and had them rate the pleasantness of the experience. Half of the listeners received an injection of an inert saline solution while the remaining listeners received injections of naloxone, an opiate receptor antagonist. Goldstein's results suggest a reduction in musically induced pleasantness for some of the naloxone-injected listeners compared with the control group. However, the pleasantness of the musical experience was not entirely eliminated, suggesting that musically induced pleasure is not restricted to neural mechanisms that result in the release of endogenous opiates. The implication is that there may be more than one way for music to evoke pleasure.

## Evolutionary aesthetics

For Hanslick, the principal problem in musical aesthetics is explaining musical beauty rather than explaining musical feeling. For the psychologist, however, Hanslick's views imply an 'essentialist' conception of music. For Hanslick, the beauty is somehow *in* the music, rather than evoked *by* the music. For most psychologists, there is nothing in the world that is objectively ugly or beautiful. Humans find darkness threatening and sunshine pleasant, but a bat will have the reverse experience. We enjoy the smell of roses more than the smell of a dead carcass, and are disgusted to discover that our pet dog has the opposite experience. In the words of Donald Symons (1992), 'Beauty is in the adaptations of the beholder.'

Feelings have a deep structure in evolution by natural selection. The feelings evoked on any given occasion can be traced to proximal causes; but the feelings themselves are generated by brain mechanisms that evolved so as to improve the organism's inclusive fitness. We love life and fear death because these feelings contribute to our survival. We fall in love and protect our children because these feelings contribute to reproductive success. According to current orthodoxy in biology and evolutionary psychology, the feelings evoked by art ought to be traceable to one or more underlying evolutionary mechanisms. This logic has led a number of scholars to offer evolutionary accounts of aesthetic experience (see Chapter 1).

Charles Darwin himself launched a history of speculation regarding the possible evolutionary benefits of music and art. In the past half century, new evolutionary aesthetics theories have appeared almost monthly. Evolutionary theories of art are both speculative and controversial. Part of the controversy arises from the ease of 'storytelling'. As Paul Griffiths has noted, 'adaptive hypotheses are too easy to form and too difficult to test' (1997, p. 71). While evolutionary theorizing appears to be an open invitation to unbridled speculation, as Jon Elster has noted, 'The first step toward finding a positive answer is telling a plausible story' (1989, p. 8).

In recent decades, many evolutionary stories regarding art have been proposed. Eibl-Eibesfeldt (1989), for example, has argued that people tend to prefer landscapes that resemble the primordial savanna environments of hominid evolution. Cross-cultural studies suggest that aesthetic preferences favour environmental conditions that have been conducive to survival, not in the contemporary world, but in the Pleistocene world of human evolution. Other theories have been proposed by Ellen Dissanayake (1988),

Nancy Aiken (1998) and others. In the case of music, possible evolutionary origins for music have been discussed by Geoffrey Miller (2000), Ian Cross (2001/2003), David Huron (2001/2003), and Steven Mithen (2006).

Most of the evolutionary accounts that have been offered regarding the origin of music propose a single function (such as sexual selection). This makes sense. If music does indeed have an evolutionary origin, it would have begun by conferring a single pre-eminent adaptive advantage. However, as we have seen above, there is a tendency to amalgamate multiple pleasures into a single human experience. Modern music-making is apt to engage a plethora of pleasure-evoking mechanisms and so it may prove difficult to untangle any presumed original purpose from the agglomerated mix of hedonic mechanisms assembled in modern music-making.

From a life sciences perspective, there are a limited number of stories that can be told that will reconcile art with biology. One story is that music and art might have originated as adaptive pleasures, where the art-related activity increased inclusive fitness in some (non-obvious) way. For example, music might have played an adaptive role through social bonding, sexual selection, or by facilitating language learning. In this case, we would have to conclude that Kant was wrong: aesthetic pleasures are indeed utilitarian. A second alternative story is that music-related behaviours might be non-adaptive forms of pleasure-seeking. Music might simply commandeer a pleasure pathway that is intended to serve some other purpose. In this case, music would be akin to nicotine or heroin addiction: Art is just a fancy way of tickling pleasure mechanisms that exist to reward other behaviors. A third alternative story is that music-related behaviours are biological 'spandrels'. That is, they are incidental artefacts that necessarily accompany other adaptive behaviours. Music might be akin to a benevolent form of sickle-cell anaemia (which is an artifact of a heterozygotic genetic strategy to protect against malaria infection). Music, for example, might simply be a non-functional artefact of brain mechanisms whose purpose is to promote language development.

For many people, none of these ideas is especially appealing. The idea that music is biologically ordained seems far-fetched and problematic. Similarly, the idea that music is a non-adaptive form of pleasure-seeking (like cocaine use) is equally unsavoury. Finally, the idea that music is a physiological accident (like the non-functioning appendix) is anti-climactic. While these ideas invite opinion, the origin or purpose of music is ultimately an empirical question that will be resolved only through future scientific research.

## Conclusion

Cognitive science has brought into relief what appears to be a fundamental disagreement between modern psychology and traditional Western aesthetics. The bedrock of conventional Western aesthetics has been the notion, promulgated by Kant, that there are unique aesthetic pleasures that exist aside from utilitarian pleasures. Evolutionary psychologists and biologists argue that the brain mechanisms that generate emotions represent evolved adaptations.

Over the past two decades, a number of aesthetics philosophers have been inspired by the 'old school' cognitive revolution. In particular, early cognitive science lent credence to the contemplative–appraisalist view of musical aesthetics advocated by Hanslick. However, the subsequent development of cognitive science is raising significant challenges for aesthetic philosophy in general. From the perspective of modern cognitive neuroscience, the disembodied, non-utilitarian notion of aesthetic pleasure posited by Kant cannot easily be reconciled with biology (see also Huron 2006).

Over the course of history, many of the problems addressed under the rubric 'philosophy' have been ceded to newly emerging scientific disciplines. Questions formerly considered part of 'natural philosophy' have shifted to the realm of biology and geology. Cosmology was taken over by physics and astronomy, and questions regarding human behaviour passed to the domains of the social and behavioural sciences. If evolutionary psychologists are correct, then questions concerning the experience of beauty and ugliness may soon slip from the grasp of philosophy. Only time will tell whether we are witnessing the passing of the aesthetics baton from philosophy to empirical science.

## References


Aitken NE (1998). *The biological origins of art.* Praeger, Westport, CT.

Arnold MB (1960). *Emotion and personality: vol. 1. psychological aspects.* Columbia University Press, New York.

Averill JR (1980). A constructivist view of emotion. In R Plutchik and H Kellerman, eds, *Emotion: theory, research and experience*, Vol. 1, 305–339. Academic Press, New York.

Berlyne DE (1971). *Aesthetics and psychobiology.* Appleton-Century-Crofts, New York.

Budd M (1985). *Music and the emotions: the philosophical theories.* Routledge, London.

Butler C (2004). *Pleasure and the arts: enjoying literature, painting, and music.* Oxford University Press, Oxford.

Cook N (1987). The perception of large-scale tonal closure. *Music Perception*, **5**(2), 197–206.

Cornelius R (1996). *The science of emotion: research and tradition in the psychology of emotion.* Prentice Hall, Upper Saddle River, NJ.

Cross I (2001/2003). Music, cognition, culture, and evolution. Annals of the New York Academy of Sciences, Vol. 930, 28–42. Reprinted in I Peretz and R Zatorre, eds, *The cognitive neuroscience of music.* Oxford University Press, Oxford.

Davies S (1994). *Musical meaning and expression.* Cornell University Press, Ithaca, NY.

Dissanayake E (1988). *What is art for?* University of Washington Press, Seattle, WA.

Eibl-Eibesfeldt I (1989). *Human ethology.* Aldine de Gruyter, Hawthorne, NY

Elster J (1989). *Nuts and bolts for the social sciences.* Cambridge University Press, Cambridge.

Fechner GT (1876). *Vorschule der Ästhetik.* Breitkopf and Härtel, Leipzig.

Fodor JA (1983). *Modularity of mind: an essay on faculty psychology.* MIT Press, Cambridge, MA.

Gabrielsson A and Lindstrom S (1993). On strong experiences of music. *Musikpsychologie: Jahrbuch der Deutschen Gesellschaft für Musikpsychologie*, **10**, 118–139.

Goldstein A (1980). Thrills in response to music and other stimuli. *Physiological Psychology*, **3**, 126–129.

Gotlieb H and Konečni VJ (1985). The effects of instrumentation, playing style, and structure in the Goldberg Variations by Johann Sebastian Bach. *Music Perception*, **3**, 87–102.

Greenwood D (1961). Critical bandwidth and the frequency coordinates of the basilar membrane. *Journal of the Acoustical Society of America*, **33**(4), 1513–1523.

Griffiths PE (1997). *What emotions really are: the problem of psychological categories.* University of Chicago Press, Chicago, IL.

Hanslick E (1854). *Vom Musikalish-Schönen.* Leipzig. Trans. Gustav Cohen as *The beautiful in music.* Bobbs-Merrill Co., New York 1957.

Helmholtz H von (1877). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik.* Braunschweig: F. Vieweg. Trans. as *On the sensations of tone as a physiological basis for the theory of music.* P. Smith, New York 1948.

Huron D (2001). Tone and voice: a derivation of the rules of voice-leading from perceptual principles. *Music Perception*, **19**(1), 1–64.

Huron D (2001/2003). Is music an evolutionary adaptation? *Annals of the New York Academy of Sciences*, **930**, 43–61. Reprinted in I Peretz and R Zatorre, eds, *The cognitive neuroscience of music.* Oxford University Press, Oxford.

Huron D (2006). *Sweet anticipation: music and the psychology of expectation.* MIT Press, Cambridge, MA.

Kagan J (1998). *Three seductive ideas.* Harvard University Press, Cambridge, MA.

Karno M and Konečni VJ (1992). The effects of structural interventions in the first movement of Mozart's symphony in G minor K. 550 on aesthetic preference. *Music Perception*, **10**(1), 63–72.

Kivy P (1990). *Music alone: philosophical reflections on the purely musical experience.* Cornell University Press, Ithaca, NY.

Koffka K (1935). *Principles of Gestalt psychology.* Harcourt, Brace, New York.

Kubovy M (1999). On the pleasures of the mind. In D Kahneman, E Diener and N Schwarz, eds, *Well-being: the foundations of hedonic psychology*, 134–154. Russell Sage Foundation, New York.

Langer S (1942). *Philosophy in a new key.* Harvard University Press, Cambridge, MA.

LeDoux J (1996). *The emotional brain: the mysterious underpinnings of emotional life.* Touchstone Books, New York.

Levinson J (1990). *Music, art and metaphysics: essays in philosophical aesthetics.* Cornell University Press, Ithaca, NY.

Levinson J (ed.) (2003). Philosophical aesthetics: an overview. *The Oxford Handbook of Aesthetics.* Oxford University Press, Oxford.

Lundin RW (1947). Toward a cultural theory of consonance. *Journal of Psychology*, **23**, 45–49.

Madell G (2002). *Philosophy, music and emotion.* Edinburgh University Press, Edinburgh.

Meyer M (1903). Experimental studies in the psychology of music. *American Journal of Psychology*, **14**, 456–475.

Miller GF (2000). *The mating mind: how sexual choice shaped the evolution of human nature.* Doubleday, New York.

Mithen S (2006). *The singing Neanderthals: the origins of music, language, mind, and body.* Harvard University Press, Cambridge, MA.

Olds J and Milner P (1954). Positive reinforcement produced by electrical stimulation of the septal area and other regions of the rat brain. *Journal of Comparative and Physiological Psychology*, **47**, 419–428.

Ollen J and Huron D (2004). Listener preferences and early repetition in musical form. In SD Lipscomb, R Ashley, RO Gjerdingen and P Webster, eds, *Proceedings of the 8th International Conference on Music Perception and Cognition*, 405–407. Casual Productions, Evanston, IL.

Panksepp J (1995). The emotional sources of 'chills' induced by music. *Music Perception*, **13**(2), 171–207.
Pert C (1997). *The pleasure molecules*. Scribner, New York.
Plomp R and Levelt WJM (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, **37**, 548–560.
Raffman D (1993). *Language, music, and mind*. MIT Press, Cambridge, MA.
Rizzolatti G and Craighero L (2004). The mirror neuron system. *Annual Review of Neuroscience*, **27**, 169–192.
Robinson J (2005). *Deeper than reason: emotion and its role in literature, music and art*. Oxford University Press, Oxford.
Scruton R (1997). *The aesthetics of music*. Clarendon Press, Oxford.
Sloboda J (1991). Music structure and emotional response: some empirical findings. *Psychology of Music*, **19**(2), 110–120.
Stumpf C (1883). *Tonpsychologie*. S. Hirzel Verlag, Leipzig.
Symons D (1992). What do men want? *Behavioral and Brain Sciences*, **15**(1), 113–114.
Unyk AM, Trehub SE, Trainor LJ and Schellenberg EG (1992). Lullabies and simplicity: a cross-cultural perspective. *Psychology of Music*, **20**(1), 15–28.
Wilson WR (1975). Unobtrusive induction of positive attitudes. PhD dissertation, University of Michigan.
Wilson WR (1979). Feeling more than we can know: exposure effects without learning. *Journal of Personality and Social Psychology*, **37**, 811–821.
Wundt W (1863). *Vorlesungen über die Menschen- und Tierseele*. L. Voss, Leipzig.
Zajonc R (1980). Feeling and thinking: preferences need no inferences. *American Psychologist*, **35**, 151–175.

# CHAPTER 15

# Musical preferences

Alexandra Lamont and Alinka Greasley

THIS chapter explores our current understanding of why we like and choose to listen to the music that we do. We begin by carefully defining terms and considering methods, moving on to discuss the biological influences of arousal and other personality traits on music preference, questions of style discrimination, and finally the cultural influences of experience upon preference. The chapter evaluates existing models of music preference and considers further directions and challenges in the field.

## Introduction

Liking for music in general is a strong human trait which can be as rewarding as food, sex, or drugs (e.g. Blood and Zatorre 2001; Panksepp and Bernatzky 2002). Certain features of music such as consonance are also liked from early infancy (Trainor and Heinmiller 1998), and these uniquely human preferences for music are not found in primates (Lamont 2005).

Research into liking for different kinds of music reveals a multitude of different concepts, sometimes used interchangeably. Over the last two decades, researchers have generally adopted Price's (1986) definition of preference as choosing or giving advantage to one thing over another. A relatively constant distinction has emerged between the concepts of *taste* (a relatively stable valuing) and *preference* (a shorter-term commitment), occupying opposite ends of a continuum (Abeles and Chung 1996). More recently, definitions of preference have also included specific notions of temporality: 'a person's liking for one piece of music as compared with another at a given point in time', while taste is held to reflect 'the overall patterning of an individual's preferences over longer time periods' (Hargreaves *et al.* 2006, p. 135). In practice, shorter-term experiences of preference inform longer-term judgements of taste and vice versa, in a cycle of reciprocal feedback (Hargreaves *et al.* 2006).

A further implicit distinction exists between research focusing on liking for specific pieces of music and that which explores liking for styles of music. Some theoretical explanations attempt to bring these two dimensions together (for example experimental aesthetics: see North and Hargreaves [2000a]). However, most of the research tends to confound responses to the piece and the style level, or to draw broader conclusions about style preference from research using specific pieces.

This chapter will thus adopt a broad definition of musical preference as referring to the music, whether style or piece, that people like and choose to listen to at any given moment *and* over time, highlighting these important dimensions throughout.

## Methods of studying musical preferences

The moment of choice can be measured at a number of levels and using a diversity of measures (see also Abeles and Chung 1996). These measures can take place either in artificial laboratory settings, in more ecologically valid contexts, or somewhere in between (such as imagining a real-life setting and one's likely responses to it in the laboratory).

First, behavioural choices include listeners' psychophysiological responses to a given piece of music, short-term decisions about which piece to listen to, or real-life patterns of

engagement with music over longer time spans as shown by concert attendance or music purchasing. Comparisons between two or more musical stimuli can be studied in infants, children and adults using preferential looking/listening paradigms (Trehub 2006) or variations on the Operant Music Listening Recorder (Greer *et al.* 1974; North and Hargreaves 2000b), which measures the amount of time a participant spends listening to different styles of music playing simultaneously through different channels. A further behavioural method involves playing extracts of music to participants and asking for different kinds of behavioural as well as verbal response (e.g. Marshall and Hargreaves 2007).

Secondly, choices can be expressed verbally in either spoken or written form, using rating scales or semantic differentials (either in response to a range of music provided or more abstractly in response to descriptions or names of types or pieces of music), individual preference nominations, or through interviews. The most common verbal report measure is the rating scale, typically requiring individuals to rate their preference for a list of predetermined musical styles on Likert scales. Several music preference scales have been developed, for example Litle and Zuckerman's (1986) Music Preference Scale, consisting of 60 established music categories from the US record industry, or Rentfrow and Gosling's (2003) Short Test of Musical Preferences, although these have not been used systematically throughout the literature. Another popular technique is in-depth interviewing of participants to uncover the richness and complexity of their everyday musical tastes (e.g. DeNora 2000).

The few studies combining different methods (e.g. Hargreaves 1988) show that different measures have different uses: self-report measures (particularly rating scales) are more suited to describing general long-term preferences, while behavioural measures seem more useful in discriminating between examples within a particular style. There thus tends to be a relatively low correlation between results, and verbal measures do not predict behaviour consistently. It is important to consider the choice of methods alongside research questions, and to consider how far methods limit the generalizability of results.

## Arousal and biological influences on music listening behaviour

At a biological level, music that we prefer seems to affect us differently. Some physiological responses to familiar and unfamiliar music are similar (e.g. Craig 2005; Lai 2004), but Blood and Zatorre (2001) found that when listeners reported more intense 'chills' or highly pleasurable intense experiences, areas of the brain responsible for reward, emotion and arousal were more strongly activated. One explanation for music preference focuses on the notion of arousal as the underlying motivator for music listening behaviour.

### Experimental aesthetics

Berlyne's psychobiological theory (1971), see also Chapter 14 this volume) argues that preference results from the interaction between an individual's level of arousal (held to be relatively stable) and the arousing properties of the music itself (more variable). Researchers have explored the musical characteristics which contribute to its arousal potential, including prototypicality, complexity, familiarity, tempo, and volume (North and Hargreaves 1995a, 1996b; Russell 1986). This research typically presents participants with simple, often artificially contrived and always experimenter-selected, musical stimuli and then measuring their verbal or behavioural preferences (North and Hargreaves 1997a, 2000b). Short-term preference for certain types of unfamiliar music can be consistently related to characteristics of that music. For example, North and Hargreaves (1995a) found a positive linear relationship between liking and familiarity for new age music, and an inverted U-shaped relationship between liking and subjective complexity of the musical examples.

This complements recent neuropsychological evidence about the arousing effects of music on the brain (see Chapter 11 this volume). It also assumes that preference expressed in an experimental setting for a given piece will be reflective of more generalized preference for a given musical style. Using Beatles songs performed in different musical styles, North and Hargreaves (1997b) found style was a more important determinant of liking than song: liking for *Yesterday* in a jazz

style, for example, related more to listeners' liking for jazz than to their liking for *Yesterday*. This suggests there is some merit in using pieces as representative of given styles, but most research has not addressed this explicitly.

One strength of this approach in explaining preference is that listeners' individual ratings of familiarity and subjective complexity with the music are assessed within each study, thereby accounting for the effects of prior experience on liking. However, it is limited through the use of pieces of music that are typically and intentionally unfamiliar to the listeners, together with the assumption that preferences for a piece reflect more general durable style preferences. Furthermore, although Berlyne argued that familiarity and exposure should change individual preference, little research has explored changing preferences over time. The effects of repeated exposure have been studied over relatively short time spans, from hours to weeks (e.g. Peretz *et al.* 1998; Hargreaves 1984), yet developmental evidence suggests that stylistic preferences change in response to a complex set of experiential factors over the lifespan (e.g. Hargreaves and North 1999; see also Chapter 22 this volume). It is important not to over-interpret preferences expressed at a given moment in time as being representative of more enduring patterns of taste (see also Lamont and Webb in press).

## Individual differences

Research has also explored the notion that different individuals have unique, preferred levels of arousal which explain their global music preferences for style. Temperament differences predict differences in preferential listening behaviour, even at 8 months of age. Trehub *et al.* (2002) found that infants who listened longer to a soothing version of a nursery song were rated by their mothers as calm and easy-going, while those who preferred a playful version of the same song were rated as highly active. Research with adults in this individual differences tradition typically employs established personality measures together with questionnaire-based music preference measures such as Litle and Zuckerman's Music Preferences Scale (1986), looking for correlations between the two.

Factors related to extraversion typically relate to preference for particularly arousing styles of music. For example, sensation seekers with high levels of optimal stimulation prefer more intense and/or complex styles of music like hard rock, soft rock, folk and classical music (Litle and Zuckerman 1986). Preference for highly arousing music such as heavy metal, rock, dance and rap correlates with high levels of resting arousal and sensation-seeking (McNamara and Ballard 1999). Preference for hard rock music is linked to excitement-seeking and extraversion (Pearson and Dollinger 2004), high levels of psychoticism and impulsiveness (Rawlings *et al.* 1995), and a relative dislike of other forms of music (Rawlings and Ciancarelli 1997). Conversely, preferences for 'softer' forms of music are associated with lower levels of psychoticism and extraversion (Rawlings *et al.* 1995). Pearson and Dollinger (2004) found that highly intuitive people showed a greater preference for classical, jazz, soul and folk music.

Some research has attempted to explain the connection between music preference and personality in relation to characteristics of the music rather than style labels. For example, Rentfrow and Gosling (2003) presented data indicating that music preferences can be organized into four independent dimensions:

1 Reflective and complex
2 Intense and rebellious
3 Upbeat and conventional, and
4 Energetic and rhythmic.

They then explored correlations between these dimensions and personality, self-views and cognitive ability ratings, finding a number of significant relationships. For example, people who preferred reflective and complex music also had active imaginations, valued aesthetic experiences, and viewed themselves as intelligent and tolerant; people who preferred upbeat and conventional music were more extrovert, agreeable, conservative, and less open to new experiences. They present a number of associations, some of which appear rather spurious (for example, why should people who like energetic rhythmic music be more likely to eschew conservative ideals?), but which they suggest may in time set the groundwork for a comprehensive theory of music preferences.

Rentfrow and Gosling's (2003) study reflects a comprehensive attempt to analyse the relationship between liking for music and aspects of personality. However, they note carefully that cultural and environmental influences also shape the music that an individual will like. Other personality researchers provide evidence that personality traits and music preferences are linked by a third factor of musical experience. For example, Rawlings and Ciancarelli (1997) found that preferences for popular and rock music, associated with extraversion and less openness to experience, were accompanied by a less intense interest in music and less musical training (see also Pearson and Dollinger 2004). This suggests that the influence of experience on personality and on music preference has yet to be fully explored.

The validity of the rating scale approach which has dominated this research depends on sufficient awareness of the differences between musical styles and style labels. Litle and Zuckerman's (1986) scale included specific stylistic examples for guidance, but both their examples and styles are culturally and historically specific (see Rawlings and Ciancarelli 1997). Rentfrow and Gosling (2003) derived their scale from empirical data about spontaneously nominated categories of liking, rated by judges and compared with industry categorization processes, and finally tested for familiarity: 29 out of 30 participants could provide preference ratings for the 14 items, which they suggest confirms their validity. However, no matter how carefully labels are constructed, the fundamental problem remains that rating scales reflect a reductionist approach. Qualitative approaches show that adults typically report preference for many different styles of music, which they often label idiosyncratically (Greasley and Lamont 2006). This sophistication and level of complexity presents an enduring challenge for the field.

### Summary and evaluation of arousal-based explanations

In addition to the issues raised above, both experimental aesthetics and individual differences approaches focus exclusively on the intra-individual level. They thus remain unable to identify the nature of the relationship between temperament and personality dimensions and musical preferences. For example, personality type may be a direct cause of musical preferences, or, as suggested above, may influence individuals' levels of engagement with music and musical activities which, in turn, affects their musical preferences. To our knowledge, none of the research has yet addressed these interactions. Finally, this emphasis on the intra-individual level neglects the social context in which the music listening is taking place.

## Towards a more contextually grounded understanding of musical preferences

Konečni (1982) argued that experimental researchers often treat music as if it existed in a social vacuum, and that it is vital to consider social interactions, emotions, moods, and other environmental factors in order to understand music choices. This raises the necessity of explaining how listening behaviour changes as a function of its immediate social and non-social antecedents, concurrent cognitive activity and resultant emotional states.

More naturalistic research has drawn on real musical stimuli and attempted to simulate real-life situations to account for both the music and the listening context. For example, North and Hargreaves (1996a, b, c, 2000b) investigated the reciprocal relationship between listener and context. Specific musical variables (mostly arousal potential) were manipulated, but the experimental conditions were naturalistic settings such as yoga classes, aerobics classes, and a university cafeteria. Listeners preferred highly arousing music during periods of exercise and arousal-moderating music when relaxing.

However, although causal relationships can be established using experimental methodology, even these more naturalistic investigations appear to be treating the 'social' as an experimental variable. Sloboda (1999) argues that the continuing use of a traditional positivist paradigm (presenting listeners with music chosen by the experimenter, in an environment controlled and constructed by the experimenter) may be responsible for slow progress in the scientific understanding of responses to music. He argues

that music listening is 'intensely situational' (p. 355) and thus context becomes central. This has led to another recent shift in approach towards a focus on capturing people's everyday musical practices and preference behaviour in the contexts in which they naturally occur.

### Preferred music in everyday life

Studies have begun to investigate people's use of music in everyday contexts (Juslin and Laukka 2004; North *et al.* 2004; North and Hargreaves 2007; Sloboda *et al.* 2001). Typically using Experience Sampling Methodology (ESM), participants are contacted (via pagers or mobile telephones) at random intervals during the day and asked to complete response sheets about their music listening. Although these studies predominantly focus on the functions of music in everyday life, the results emphasize the concept of choice. People choose different types of music for different reasons (i.e., they have specific goals and purposes that music engagement fills), and their motivations for music listening are context-dependent. Personal favourites also change over time, and daily favourites reflect situational and emotional 'fit' while long-term favourites are more connected to personal life histories (Lamont and Webb in press). Sloboda *et al.* (2001) also found that greater personal choice was more likely to be associated with positive valued outcomes such as increased arousal, present-mindedness, and positivity. These findings underline the value of researching people's self-chosen uses of music (see also Chapter 40 this volume).

### Reflecting on preferred music

An alternative approach is articulated by research that explores people's own music and the meaning of this to them as individuals, using social constructionist methods such as interviewing and ethnography. From such a perspective, music does not simply act as a stimulus on an individual, but rather its meaning and effects become stabilized through discourse, consumption practice, and patterns of use over time (DeNora 2000).

For example, Batt-Rawden and DeNora (2005) explored the therapeutic uses of music in everyday life using a unique methodology involving both researcher-chosen and participant-chosen music, repeated music listening over long time spans, and repeated in-depth interviews. This study led to a heightened awareness from participants of the ways in which they use music in everyday life (see also Carlton 2006; Sloboda *et al.* 2001). Greasley and Lamont (2006) also found differences between more and less engaged music listeners in terms of listening behaviour, preferences, and self-awareness. Less musically engaged adults lacked a strong commitment to any musical style, and were more likely to listen to an eclectic mix of music acquired from friends. Conversely, more musically engaged adults showed strong commitments to musical styles and a sense of necessity and urgency about buying or obtaining music. They expressed a detailed awareness of the styles they did and did not like listening to, and a thorough and explicit understanding of the effects that different styles of music would have on them.

## Effects of musical preferences

We next consider the impact that musical preferences can have on other areas of life in two ways. The first relates to the use of *any* kind of preferred music to achieve certain non-musical goals. The second relates to the preference for *specific* kinds of music.

Preferred music listening has been shown to be particularly effective in achieving physical and psychophysiological goals, such as pain management and relief. It leads to enhanced control over, and effective distraction from, pain-inducing stimuli under laboratory conditions, when compared with non-preferred or experimenter-selected music (Mitchell *et al.* 2006; Mitchell and MacDonald 2006). Similar effects in reducing pain, anxiety and agitated behaviour have been found both in clinical settings (MacDonald *et al.* 2003; Siedliecki and Good 2006; Sung and Chang 2005) and in chronic pain in everyday life settings (Mitchell *et al.* 2007). Listening to preferred music rather than experimenter-chosen music or silence produces lower heart rate and perceived exertion and fatigue rates (Pothoulaki and Natsume 2006), and improves cognitive performance in

driving simulation tasks (Cassidy 2006). Thus listening to preferred music has powerful effects on aspects of behaviour outside voluntary control as well as on mood and affect. In these cases, the nature of the musical stimulus has no bearing whatsoever on the physical and psychological effects.

Other uses of musical preferences are more closely tied to the particular music that an individual or a group shows preference for. During adolescence, musical preferences play an important role in the formation of identity through processes of in-group behaviour and impression management (Finnäs 1989; Tarrant *et al.* 2004). The social identity effect of musical preferences in bringing people together operates even when participants are unaware of precisely what musical preferences the in- and out-groups have (Bagakiannis and Tarrant 2006). Although specific music often has particular effects on different groups, these strong effects are more marked in adolescence; young adults are more willing to share and tolerate others' music, and this tolerance increases later in adulthood (Carlton 2006; Greasley and Lamont 2006). However, even in adulthood, musical preferences can be used in interpersonal perception to give messages about people's personalities (Rentfrow and Gosling 2006). This catalytic or self-directed effect of music in identity can be contrasted with the emblematic outward-directed effect of music, such as anthems, as symbols of national, ethnic, or cultural identity (Hammarlund 1990, cited in Folkestad 2002). In multicultural situations, the kind of music someone likes can play a significant role in the processes of adjustment to a new culture and retaining links to the old (Ilari 2006; O'Hagin and Harnish 2006).

## Explaining and predicting music preferences

Although there has been a great deal of research exploring different facets of musical preference and taste, only two explicit models of music preference (LeBlanc 1982; Hargreaves *et al.* 2006) have attempted to tie these together.

LeBlanc's interactive theory of music preference (1982) is a complex and comprehensive attempt to represent the influence of 'input information' (the musical stimulus and the listener's cultural environment) and listener characteristics and behaviour. The approach is useful in formally identifying the large number of types of variables that fall into the three broad categories. For example, the listener's cultural environment includes the variables of media, peer group, family, educators and authority figures, and incidental conditioning. The model traces a trajectory through listener characteristics such as attention and mental processing through to a preference decision at a given moment, which then influences subsequent behaviour (e.g., acceptance and then repetition of the stimulus). While the detail is potentially useful, the fact that every variable potentially interacts with every other means, as LeBlanc concedes, that this is unlikely to serve as a usable predictive model. However, subsequent research has attempted to weight the relative contributions of the various factors, and LeBlanc *et al.* (2000) found that musical features accounted for more variation in children's expressed musical preferences, followed by 'culture' and finally age (see Chapter 22 this volume).

Hargreaves *et al.* (2006) developed a far simpler reciprocal feedback model of musical response, consisting of the interactions between the three broad variables of *music*, *listener*, and *situations and contexts* to evoke a given *response*. Drawing on experimental aesthetics, musical features include a reference system (genres, styles, etc.), collative variables (complexity, familiarity) and prototypicality. The listener is characterized in terms of individual differences (gender, age, personality) as well as musical knowledge, preference and taste, and identity. The listener's response to the music is also affected by physiological (engagement, arousal, active listening), cognitive (attention, expectation, discrimination) and affective (emotional, mood, liking) factors. Finally, situations and contexts include social and cultural contexts, everyday situations, and the presence or absence of others.

Both models express a tripartite division between music, listener, and context as well as a large number of interactions both between and within levels of analysis. However, culture should not be treated as a variable but rather as the medium through which all real-life experiences

are mediated (Cole 1996; Lamont 2006). The models also say little about the outcome of musical preferences. LeBlanc's preference decision leading to rejection or acceptance (in the latter case resulting in freely chosen repetition and heightened attention) is simply linked back to the listener's cultural environment and musical stimulus input. Similarly, the concept of reciprocal feedback simply argues that listener and music, listener and situation, and situation and music 'interact'.

## Discussion

Musical preferences serve a range of important functions for individuals and groups, and preferred music can play an important role in physical and psychological well-being. These clearly go beyond the simple behavioural outcome of repeated exposure, and have far-reaching effects ranging from the personal to the cultural.

While experimental research has addressed some important questions in relation to our understanding of musical preferences, there still remain many unanswered issues. The complexities lie largely in the interactive nature of musical preference. Even a single preference expressed at a given moment in time between one of two experimenter-selected pieces of music is likely to be affected by a host of factors, which will vary from individual to individual and may lead to a range of different outcomes. Attempting to isolate and examine these within a positivist approach can be a daunting and potentially fruitless challenge, which may explain why some of the more successful approaches to understanding musical preferences, both experimental and qualitative, appear rather content-free in terms of the *music* that is being preferred.

Furthermore, the complexity and flexibility in the ways that people categorize and label music is a critical issue for the field. As listeners argue about how particular pieces of music, particularly those they like, should be labelled into styles (Greasley and Lamont 2006), research asking participants to tick boxes of music preference categories is not likely to inform us significantly about the underlying meanings of those preferences, except, perhaps, in situations where the 'tick box' approach has validity, such as internet dating, cf. Rentfrow and Gosling (2006). The particular categories employed are also likely to change rapidly along with changes in musical style (Hargreaves and North 1999), limiting comparability between different studies. An approach that prioritizes listeners' own constructions and interpretations of music circumvents some of these problems, and may prove more fruitful in explaining these less stable elements of musical preferences.

The temporal dimension of preference is another central issue, and while we have highlighted the temporal dimensions of decision-making throughout, it is harder to tease out practical implications for a theory of musical preferences. The two models of music preferences reviewed here adopt very different perspectives on temporality: LeBlanc systematically specifies the precise moment of choice but says less about the longer-term concomitants of that choice, whereas Hargreaves and colleagues attempt to capture longer-term dimensions while remaining vague about the choices which are being represented. A more considered explanation of the temporal dimension (where preferences originate, are shaped, grow, and die down—in essence, how reciprocal feedback actually works) is still required (cf. Lamont 2006).

Finally, adopting a cultural psychological approach of treating culture and context more thoroughly as a medium for musical preferences rather than a variable within a model may have the potential to address some of the unresolved issues in this field. Naturalistic and longitudinal methods of enquiry may be more valuable here. For example, interviewing people at home with their music collections (Greasley and Lamont 2006) enables them to interact with music in a far more contextualized manner, encouraging participants to reflect on the wide range of interacting factors influencing preference (see also Batt-Rawden and DeNora 2005).

To conclude, the privileged position of preferred music in individuals' lives is something that future research needs to be sensitive to. A colleague undergoing chemotherapy told us how she actively decided *not* to bring her favourite music into hospital, despite her specialist's exhortations that it would help alleviate her pain. She was concerned that over time listening to her favourite music in this context would lead

to a negative association between the music and the treatment, thus 'spoiling' its potential as a source of pleasure (Chris Banks, personal communication). This kind of real-life engagement with music is not easily explained by inverted U-shapes or artificial distinctions between listener, music and context, yet it is such real-life challenges that future research must find better ways of explaining.

## References


Abeles HF and Chung JW (1996). Responses to music. In DA Hodges, ed., *Handbook of music psychology*, 2nd edn, 285–342. IMR Press, San Antonio, TX.

Bakagiannis S and Tarrant M (2006). Can music bring people together? Effects of shared musical preference on intergroup bias in adolescence. *Scandinavian Journal of Psychology*, **47**, 129–136.

Batt-Rawden K and DeNora T (2005). Music and informal learning in everyday life. *Music Education Research*, **7**, 289–304.

Berlyne DE (1971). *Aesthetics and psychobiology*. Appleton-Century-Crofts, New York.

Blood AJ and Zatorre RJ (2001). Intensely pleasureable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences*, **98**, 11818–11823.

Carlton L (2006). A qualitative analysis of everyday uses of preferred music across the life span. In M Baroni, AR Addessi, R Caterina and M Costa, eds, *Proceedings of the 9th International Conference on Music Perception and Cognition*, 582–583 University of Bologna, Bologna, Italy.

Cassidy G (2006). The effects of preferred music on driving game performance. In M Baroni, AR Addessi, R Caterina and M Costa, eds, *Proceedings of the 9th International Conference on Music Perception and Cognition*, 584–595 University of Bologna, Bologna, Italy.

Cole M (1996). *Cultural psychology: a once and future discipline*. Harvard University Press, Cambridge, MA.

Craig DG (2005). An exploratory study of physiological changes during 'chills' induced by music. *Musicae Scientiae*, **IX**(2), 273–287.

DeNora T (2000) *Music in everyday life*. Cambridge University Press, Cambridge.

Finnäs L (1989). A comparison between young people's privately and publicly expressed musical preferences. *Psychology of Music*, **17**, 132–145.

Folkestad G (2002). National identity and music. In RAR MacDonald, DJ Hargreaves and DE Miell, eds, *Musical identities*, 151–162. Oxford University Press, Oxford.

Greasley AE and Lamont AM (2006). Music preference in adulthood: why do we like the music we do? In M Baroni, AR Addessi, R Caterina and M Costa, eds, *Proceedings of the 9th International Conference on Music Perception and Cognition*, 960–966. University of Bologna, Bologna, Italy.

Greer RD, Dorow LG and Randall A (1974). Music listening preferences of elementary school children. *Journal of Research in Music Education*, **22**, 284–291.

Hargreaves DJ (1984). The effects of repetition on liking for music. *Journal of Research in Music Education*, **32**, 35–47.

Hargreaves DJ (1988). Verbal and behavioural responses to familiar and unfamiliar music. *Current Psychological Research and Reviews*, **6**, 323–330.

Hargreaves DJ and North AC (1999). Developing concepts of musical style. *Musicae Scientiae*, **III**(2), 193–216.

Hargreaves DJ, North AC and Tarrant M (2006). Musical preference and taste in childhood and adolescence. In GE McPherson, ed., *The child as musician*, 135–154. Oxford University Press, Oxford.

Ilari B (2006). Music and identity of Brazilian Dekasegi children and adults living in Japan. In M Baroni, AR Addessi, R Caterina and M Costa, eds, *Proceedings of the 9th International Conference on Music Perception and Cognition*, 123–130. University of Bologna, Bologna, Italy.

Juslin PN and Laukka P (2004). Expression, perception, and induction of musical emotions: a review and a questionnaire study of everyday listening. *Journal of New Music Research*, **33**, 217–238.

Konečni VJ (1982). Social interaction and music preference. In D Deutsch, ed., *The psychology of music*, 497–516. Academic Press, New York.

Lai H-L (2004). Music preference and relaxation in Taiwanese elderly people. *Geriatric Nursing*, **25**(5), 286–291.

Lamont A (2005). What do monkeys' music choices mean? *Trends in Cognitive Science*, **9**(8), 359–361.

Lamont A (2006). *Review of Musical Communication* (eds Miell, MacDonald and Hargreaves). *Musicae Scientiae*, **10**(2), 278–282.

Lamont A and Webb R (in press). Short- and long-term musical preferences: what makes a favourite piece of music? *Psychology of Music*.

LeBlanc A (1982). An interactive theory of music preference. *Journal of Music Therapy*, IXI(1), 28–45.

LeBlanc A, Jin YC, Chen-Hafteck L, Oliviera ADJ, Oosthuysen S and Tafuri J (2000). Tempo preferences of young listeners in Brazil, China, Italy, South Africa, and the United States. *Bulletin of the Council for Research in Music Education*, **147**, 97–102.

Litle P and Zuckerman M (1986). Sensation-seeking and music preferences. *Personality and Individual Differences*, **7**, 575–578.

MacDonald RAR, Mitchell LA, Dillon T, Serpell MG, Davies JB and Ashley EA (2003). An empirical investigation of the anxiolytic and pain-reducing effects of music. *Psychology of Music*, **31**, 187–203.

Marshall N and Hargreaves DJ (2007). Musical style discrimination in the early years. *Journal of Early Childhood Research*, **5**(1), 35–49.

McNamara L and Ballard ME (1999). Resting arousal, sensation-seeking, and music preference. *Genetic, Social and General Psychology Monographs*, **125**, 229–250.

Mitchell LA and MacDonald RAR (2006). An experimental investigation of the effects of preferred and relaxing

music listening on pain perception. *Journal of Music Therapy*, **XLIII**(4), 295–316.

Mitchell LA, MacDonald RAR and Brodie EE (2006). A comparison of the effects of preferred music, arithmetic and humour on cold pressor pain. *European Journal of Pain*, **10**, 343–351.

Mitchell LA, Macdonald RAR, Knussen C and Serpell MG (2007). A survey investigation of the effects of music listening on chronic pain. *Psychology of Music*, **35**(1), 39–59.

North AC and Hargreaves DJ (1995a). Subjective complexity, familiarity, and liking for popular music. *Psychomusicology*, **14**, 77–93.

North AC and Hargreaves DJ (1995b). Eminence in pop music. *Popular Music and Society*, **19**, 41–66.

North AC and Hargreaves DJ (1996a). Situational influences on reported musical preferences. *Psychomusicology*, **15**, 30–45.

North AC and Hargreaves DJ (1996b). The effects of music on responses to a dining area. *Journal of Environmental Psychology*, **16**, 55–64.

North AC and Hargreaves DJ (1996c). Responses to music in aerobic exercise and yogic relaxation classes. *British Journal of Psychology*, **87**, 535–547.

North AC and Hargreaves DJ (1997a). Experimental aesthetics and everyday music listening. In DJ Hargreaves, AC North, eds, *The social psychology of music*, 84–101. Oxford University Press, Oxford.

North AC and Hargreaves DJ (1997b). Liking for musical styles, *Musicae Scientiae*, **I**(1), 109–128.

North AC and Hargreaves DJ (2000a). Collative variables versus prototypicality. *Empirical Studies of the Arts*, **18**, 13–17.

North AC and Hargreaves DJ (2000b). Musical preferences during and after relaxation and exercise. *American Journal of Psychology*, **113**, 43–67.

North AC and Hargreaves DJ (2007). Lifestyle correlates of musical preference: 1. Relationships, living arrangements, beliefs, and crime. *Psychology of Music*, **35**, 58–87.

North AC, Hargreaves DJ and Hargreaves JJ (2004). Uses of music in everyday life. *Music Perception*, **22**, 41–77.

O'Hagin IB and Harnish D (2006). Music as a cultural identity: a case study of Latino musicians negotiating tradition and innovation in northwest Ohio. *International Journal of Music Education*, **24**, 56–70.

Panksepp J and Bernatzky G (2002). Emotional sounds and the brain: the neuro-affective foundations of musical appreciation. *Behavioural Processes*, **60**, 133–155.

Pearson JL and Dollinger SJ (2004). Music preference correlates of Jungian types. *Personality and Individual Differences*, **36**, 1005–1008.

Peretz I, Gaudreau D and Bonnel AM (1998). Exposure effects on music preferences and recognition. *Memory and Cognition*, **26**(5), 884–902.

Pothoulaki M and Natsume M (2006). The effects of preferred music listening in college students and in renal failure patients. In M Baroni, AR Addessi, R Caterina and M Costa, eds, *Proceedings of the 9th International Conference on Music Perception and Cognition*, 586–587 University of Bologna, Bologna, Italy.

Price HE (1986). A proposed glossary for use in affective response. *Journal of Research in Music Education*, **34**, 151–159.

Rawlings D and Ciancarelli V (1997). Music preference and the five-factor model of the NEO Personality Inventory. *Psychology of Music*, **25**, 120–132.

Rawlings D, Hodge M, Sherr D and Dempsey A (1995). Toughmindedness and preference for musical excerpts, categories and triads. *Psychology of Music*, **23**, 63–80.

Rentfrow PJ and Gosling SD (2003). The do re mi's of everyday life: examining the structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, **84**(6), 1236–1256.

Rentfrow PJ and Gosling SD (2006). Message in a ballad: the role of musical preferences in interpersonal perception. *Psychological Science*, **17**(3), 236–242.

Russell PA (1986). Experimental aesthetics of popular music recordings: pleasingness, familiarity and chart performance. *Psychology of Music*, **14**, 33–43.

Siedliecki SL and Good M (2006). Effect of music on power, pain, depression and disability. *Journal of Advanced Nursing*, **54**(5), 553–562.

Sloboda JA (1999). Everyday uses of music listening: a preliminary study. In SW Yi, ed., *Music, Mind and Science*, 354–369. Western Music Institute, Seoul.

Sloboda JA, O'Neill SA and Ivaldi A (2001). Functions of music in everyday life: an exploratory study using the Experience Sampling Method. *Musicae Scientiae*, **V**, 9–32.

Sung H-C and Chang AM (2005). Use of preferred music to decrease agitated behaviours in older people with dementia: a review of the literature. *Journal of Clinical Nursing*, **14**(9), 1133–1140.

Tarrant M, North AC and Hargreaves DJ (2004). Adolescents' intergroup attributions: a comparison of two social identities. *Journal of Youth and Adolescence*, **33**, 177–185.

Trainor LJ and Heinmiller BM (1998). The development of evaluative responses to music: infants prefer to listen to consonance over dissonance. *Infant Behavior and Development*, **21**, 77–88.

Trehub SE (2006). Infants as musical connoisseurs. In GE McPherson, ed., *The child as musician: a handbook of musical development*, 33–49. Oxford University Press, Oxford.

Trehub SE, Nakata T and Bergeson T (2002). Infants' responsiveness to soothing and playful singing. Paper presented at the International Conference on Infant Studies, Toronto.

# PART 4

# Music and the brain

Edited by Aniruddh D. Patel

CHAPTER 16

# The neurobiological basis of musical expectations

Laurel J. Trainor and Robert J. Zatorre

## Introduction

As recently outlined by Huron (2006), we have evolved a brain that rewards accurate prediction. The evolutionary advantages of accurately predicting future events in the world are clear. Over half a century ago, Leonard Meyer (1956) presented his theory that emotional responses to music arise through expectations set up by musical patterns. However, for the most part, music is not about events in the world. What is the link between music and brain mechanisms for prediction? Meyer noted that some chords are more likely to follow other chords in Western musical structure: a melodic fragment sets up expectations for how it will continue. He proposed that low probability realizations give rise to physiological responses, such as changes in heart rate, breathing, and skin responses at an unconscious level, and thus spark emotional responses. Interpretation of the physiological responses in the context of the music gives rise to conscious emotional experience. Meyer proposed that most musical expectations were probably the result of familiarization with a particular musical genre. Indeed, a number of psychological studies have since asked Western listeners to rate their expectations of melodic continuations, and these studies show that people base their expectations on the rules of Western musical theory, whether or not they are explicitly aware of these rules (e.g. Krumhansl 1997; Tillmann *et al.* 2006; Tillmann and Lebrun-Guillaud 2006; Unyk and Carlson 1987). On the other hand, Eugene Narmour (1990) presented his implication–realization model of musical expectation, which included a substantial set of innate, universal, bottom-up principles that should apply to all musical genres regardless of whether or not the listener is familiar with the genre. In testing people's expectations of various continuations of melodic patterns, Schellenberg (1997) was able to simplify Narmour's principle to two main ones. All else being equal, expect the next note to be proximate in pitch to the preceding note, and expect a reversal in pitch contour after relatively large intervals, but a continuation of pitch contour after relatively small intervals.

Recent neuroscientific work on musical processing is beginning to reveal the brain mechanisms by which expectations are set up and by which musical events are evaluated. In this chapter we will explore how the auditory system processes incoming information and generates perceptual representations that allow it to make predictions about future sound events from past context, and how music appears to make use of this general processing mechanism. We will focus on expectation formation in auditory cortex because this is where the most research has been done, but there is also evidence for prediction mechanisms at subcortical levels and at levels beyond sensory areas. We do not provide a complete review of the literature, but rather present a framework for thinking about the

neurological basis of expectation and prediction in musical processing using selected examples.

## Physiological measurement of auditory expectations: predicting *what* will happen next

When a neuron depolarizes or fires, the flow of sodium and potassium ions in and out of the cell membrane results in the creation of electrical field potentials. When a large number of neurons whose axons point in the same direction, as between cortical layers, depolarize at the same time, a field potential is created that is large enough to be measured in a non-invasive way using an array of sensors across the scalp. When the electroencephalogram (EEG) is measured in response to an event such as a sound, it is referred to as an event-related potential (ERP). When occasional changes are made to a stream of repetitions of one sound, or when a change is made to a category of sounds, a frontally negative component is superimposed on the ERP wave form (Figure 16.1) that peaks between 150 and 250 ms, depending on the type of change. This component is called the mismatch negativity, or MMN, because it occurs only when there is an occasional mismatch between the expected stimulus and the presented stimulus (Kujala *et al.* 2007; Näätänen *et al.* 2007; Picton *et al.* 2000). MMN is typically measured by subtracting the average response on the frequent standard trials from the average response on deviant trials (the occasional change trials). Although MMN can be influenced by attention, it occurs automatically and cannot be stopped by top-down conscious processing.

In the present context, MMN is of great interest because it reflects a process in the auditory system for predicting future sound events on the basis of the recent past, and the brain's reaction when those predictions are not fulfilled. We propose here that music fundamentally relies on expectation–realization processes, and that these processes are reflected in the MMN, whose main generators are in secondary auditory cortex (Schönwiesner *et al.* 2007). However, it should be kept in mind that there is a frontal contribution to the MMN response, that lower areas of the nervous system also employ expectation mechanisms (e.g., Csépe *et al.* 1989; Kraus *et al.* 1994; Ruusuvirta *et al.* 1995; Sonnadara *et al.* 2006b), and that some later ERP components reflecting conscious processing of music are also sensitive to expectations (e.g., Trainor *et al.* 2002; Desjardins *et al.* 1999).

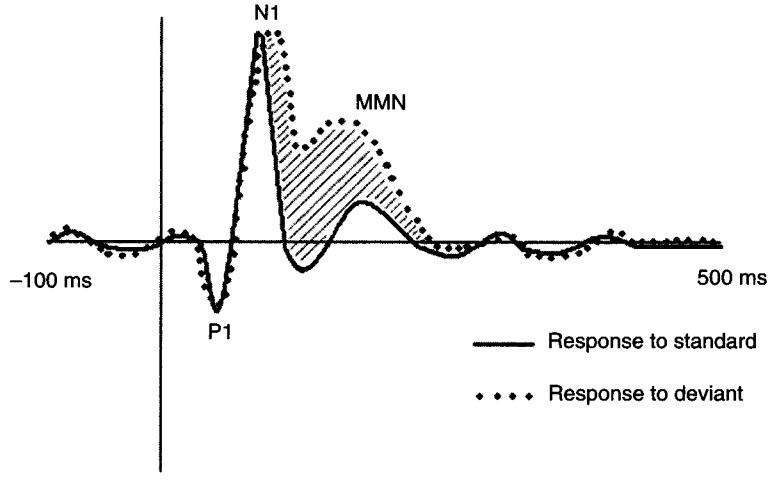



**Fig. 16.1** Stylized rendition of the mismatch negativity event-related potential (MMN). Occasional deviants in a repeating stream of standard tones elicit an extra negativity (MMN) in the wave form of the deviants in comparison to standards. Note that negative deflections are depicted in the upwards direction.

neurological basis of expectation and prediction in musical processing using selected examples.

## Physiological measurement of auditory expectations: predicting *what* will happen next

When a neuron depolarizes or fires, the flow of sodium and potassium ions in and out of the cell membrane results in the creation of electrical field potentials. When a large number of neurons whose axons point in the same direction, as between cortical layers, depolarize at the same time, a field potential is created that is large enough to be measured in a non-invasive way using an array of sensors across the scalp. When the electroencephalogram (EEG) is measured in response to an event such as a sound, it is referred to as an event-related potential (ERP). When occasional changes are made to a stream of repetitions of one sound, or when a change is made to a category of sounds, a frontally negative component is superimposed on the ERP wave form (Figure 16.1) that peaks between 150 and 250 ms, depending on the type of change. This component is called the mismatch negativity, or MMN, because it occurs only when there is an occasional mismatch between the expected stimulus and the presented stimulus (Kujala *et al.* 2007; Näätänen *et al.* 2007; Picton *et al.* 2000). MMN is typically measured by subtracting the average response on the frequent standard trials from the average response on deviant trials (the occasional change trials). Although MMN can be influenced by attention, it occurs automatically and cannot be stopped by top-down conscious processing.

In the present context, MMN is of great interest because it reflects a process in the auditory system for predicting future sound events on the basis of the recent past, and the brain's reaction when those predictions are not fulfilled. We propose here that music fundamentally relies on expectation–realization processes, and that these processes are reflected in the MMN, whose main generators are in secondary auditory cortex (Schönwiesner *et al.* 2007). However, it should be kept in mind that there is a frontal contribution to the MMN response, that lower areas of the nervous system also employ expectation mechanisms (e.g., Csépe *et al.* 1989; Kraus *et al.* 1994; Ruusuvirta *et al.* 1995; Sonnadara *et al.* 2006b), and that some later ERP components reflecting conscious processing of music are also sensitive to expectations (e.g., Trainor *et al.* 2002; Desjardins *et al.* 1999).

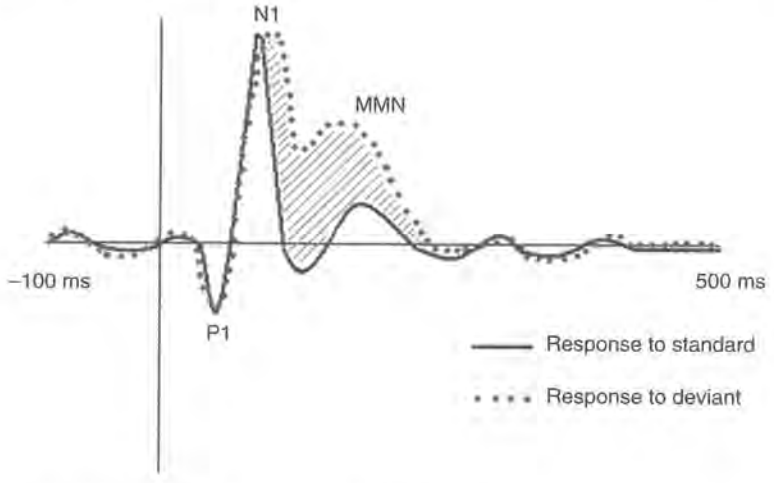



**Fig. 16.1** Stylized rendition of the mismatch negativity event-related potential (MMN). Occasional deviants in a repeating stream of standard tones elicit an extra negativity (MMN) in the wave form of the deviants in comparison to standards. Note that negative deflections are depicted in the upwards direction.

## Acoustic expectations reflected in the MMN response

Mismatch responses reflect a general mechanism in auditory cortex in that MMN has been measured in response to changes in virtually every basic sound feature. Across these different types of change, the larger the change, the larger and earlier the MMN tends to be, and the more rare the deviant, the larger the MMN. Occasional changes in the frequency or pitch of a tone elicit robust MMN (see below). Changes in timbre also elicit MMN (Tervaniemi *et al.* 1997b; Toiviainen *et al.* 1998), as do changes in duration, whether of a few milliseconds (e.g., Desjardins *et al.* 1999; Trainor *et al.* 2001, 2003a; Uther *et al.* 2003) or hundreds of milliseconds (e.g., Kaukoranta *et al.* 1989; Jaramillo *et al.* 2000). Changes in the spatial location of sounds produce MMN (e.g., Paavilainen *et al.* 1989; Sonnadara *et al.* 2006a). Interestingly, both increases and decreases in intensity elicit MMN (Näätänen *et al.* 1989; Näätänen 1992), and even the omission of an expected sound can elicit MMN (Raij *et al.* 1997; Tervaniemi *et al.* 1994), indicating that MMN is not the result of the recruitment of additional neurons required to process a new stimulus, but rather the reflection of expectation violation processes.

What makes the MMN response so interesting is that it is elicited not only in response to the basic sound features indicated above, but also to changes in more complex stimulus features. For example, a change in speech phoneme category in the presence of multiple exemplars of each category elicits MMN (Kujala *et al.* 2001), as does a change in the order of tones in a pattern (e.g., Tervaniemi *et al.* 1997a). Furthermore, the MMN response is sensitive to complex statistical properties of the input, such as the conjunction of simple features. When a number of standard stimuli are presented that vary in simple features such as frequency, intensity and location, the presentation of a sound with a novel combination of these features elicits MMN (Näätänen and Winkler 1999). MMN is also sensitive to abstract rules governing the conjunction of features. When many standard stimuli are presented that vary in frequency and intensity, but follow the rule that stimuli with higher frequency also tend to be more intense, MMN is elicited when this rule is violated (e.g., with a high-frequency, low-intensity sound) (Paavilainen *et al.* 2003).

Interestingly, the mismatch response to each sound feature gives rise to a slightly different pattern of electrical activation across the scalp, suggesting that the location of the neurons generating the MMN is in a slightly different place for each feature (e.g., Giard *et al.* 1995). This suggests that there is not one central MMN mechanism that processes all types of expectations but rather that there are many MMN processors, each specialized for a different sound feature. Thus, expectations for future sound events, and the evaluation of the realization of these expectations, is a widespread type of processing that is applied in many different places in auditory cortex. Furthermore, as will be discussed below, there is evidence that more than one MMN process can occur at the same time. In the next section, we focus on MMN studies probing expectancies in musical contexts.

## Expectations in musical contexts

Melody, the perceptual grouping of successive pitches into a coherent pattern, is central to musical structure. Although melodies are structured somewhat differently across different musical systems, there are universal features of melodies that probably evolved from basic constraints of the auditory system. One perceptually important feature is the contour, or up/down pattern of pitch changes without regard to the exact size of the pitch changes. Pitch contour is not specific to music as it is also important in the perception of linguistic information, playing a role in intonation, stress and, in the case of tone languages, lexical identity. However, melodies are also processed in terms of the exact size of the pitch intervals between tones, and unlike contour, interval processing does seem to be quite specific to music. In any particular musical system, pitch is not a continuous variable, but rather, only a small number of discrete intervals are used. For example, in the Western tonal musical system, the major scale defines one of the most common set of intervals used for composition. Other musical systems use different interval structures in their scales,

but virtually all scales rely on octave equivalence and a small set of tones per octave (Dowling and Harwood 1986).

Melodic expectation arises in several ways. First, with a familiar melody, we can make predictions about the pitch of the next note that is expected based on a long-term memory representation. Second, when we hear music in a particular style, whether familiar or not, we can extract the statistical regularities and make predictions based on those. Third, whether or not a melody is familiar, we can make predictions about next notes that are based on our knowledge of universal melodic features. Fourth, we can use our culture-specific knowledge of what notes are likely to follow other notes in the music style with which we are familiar (Huron 2006; Hannon and Trainor 2007).

How might melodic expectations arise from the cortical system that handles pitch information? In order for expectations to develop, the system must first extract stable representations of the incoming input, and code relevant features in an invariant fashion (Whitfield 1985). The solution to this computational problem seems to lie in the hierarchical arrangement of sensory processing streams, which allow for different levels of abstraction to be computed at each level of processing. In the case of a melody, the interval information must be abstracted from individual tones which may vary in their fundamental frequency from one instance to another. In turn, the pitch of the individual tones must itself be computed in such a way as to achieve perceptual constancy, ignoring irrelevant variation in surface features. Substantial evidence now exists that pitch invariance is computed in an area of belt cortex lateral to primary cortex, in both monkeys (Bendor and Wang 2005), and humans (Zatorre 1988; Johnsrude *et al.* 2000; Griffiths *et al.* 1998; Penagos *et al.* 2004; Krumbholtz *et al.* 2003). This region then feeds information into more distal portions of the processing stream, both anterior and posterior to the pitch-sensitive region, where melodic features are processed (Zatorre 1985; Zatorre *et al.* 1994; Patterson *et al.* 2002). The hierarchy of processing must eventually make contact with long-term memory systems where both implicit knowledge about general patterns, and specific knowledge about a particular musical piece, are stored. Expectations could thus arise from the interplay between the feedforward sensory abstraction mechanism, and the feedback received from stored internal representations.

MMN responses reflect musical expectations at a number of levels. Violation of both contour and interval expectations set up by a melodic context give rise to MMN, even in non-musicians. Saarinen *et al.* (1992) presented standard tone pairs where the pitches of the tones varied from trial to trial, but the pitch always rose from the first to the second tone. In deviant tone pairs the pitch fell from the first to second tone. Thus, despite the fact that standards varied in absolute pitch, the auditory cortex generated an expectation for rising contour. Tervaniemi *et al.* (2001) showed that contour expectations are also generated with a more complex inverted-U-shaped contour. Given that pitch contour is important in both speech and music, it is perhaps not surprising that the brain is set up to automatically predict contour based on current context in musicians and non-musicians alike. However, Trainor *et al.* (2002) showed that non-musicians also form expectations for particular intervals in the absence of absolute pitch information. Specifically, when presented with a standard 5-note melody that was transposed to different pitch levels (keys) on successive repetitions, occasional changes to the last note that altered the interval size but not contour led to robust MMN. Such mismatch responses are typically larger and earlier in musicians than in non-musicians (Figure 16.2), (Fujioka *et al.* 2004), indicating either genetic differences between musicians and non-musicians or the plastic effects of musical training. However, at the same time, these results suggest that even in the absence of musical training, auditory cortex automatically encodes music-specific melodic information and creates expectations for future events.

In polyphonic music, two or more melody lines occur at the same time. Fujioka *et al.* (2005) demonstrated that separate memory traces are formed for each melody, and that predictive processes within each stream can be measured with MMN. On 25 per cent of trials one note was raised by a tone or a semitone in one melody, and on another 25 per cent of trials one tone

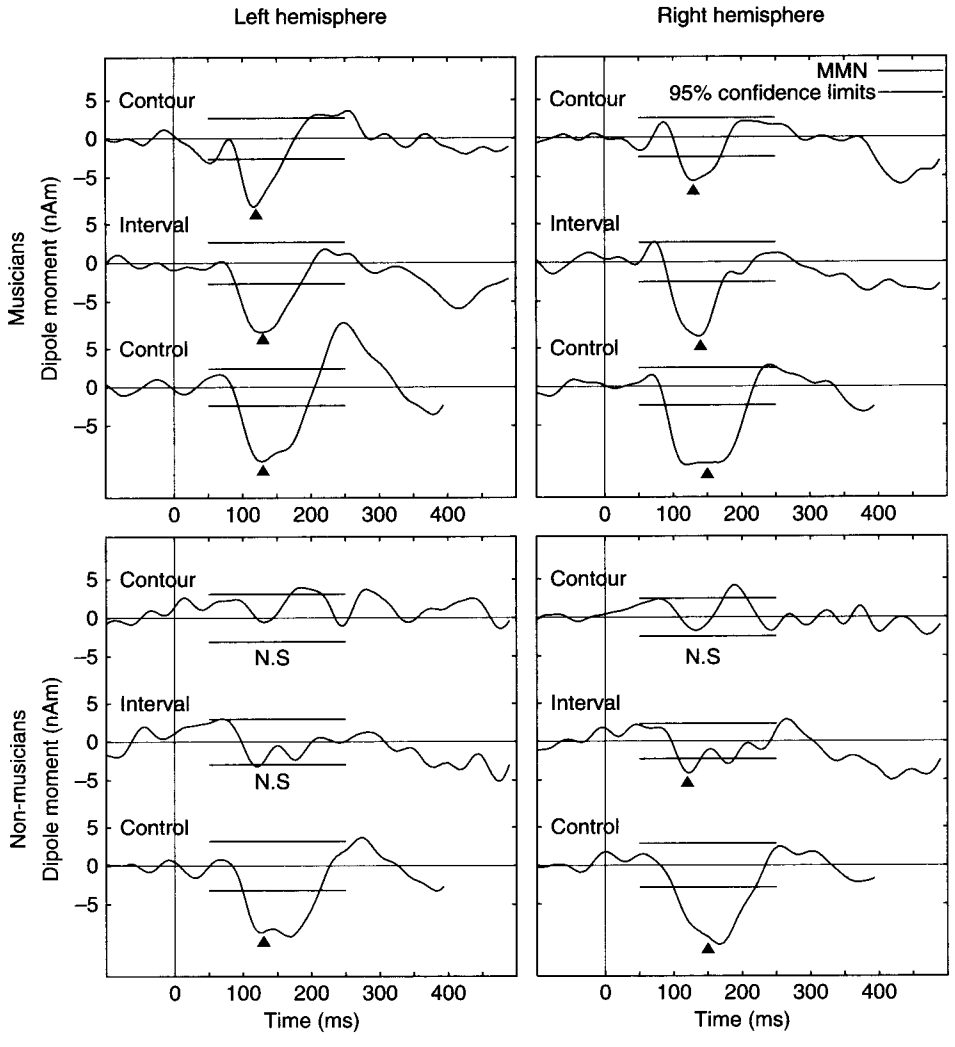



**Fig. 16.2** MMN to occasional changes in the contour or interval of a melody. Source space wave forms in auditory cortex from MEG data are shown. Musicians are shown in the upper panel and non-musicians in the lower panel. In both groups, an MMN response can be seen to changes in pitch interval and to changes in pitch contour, but MMN amplitude is much larger in musicians than in non-musicians. MMN is also seen in the control condition where the pitch of a single repeating tone is occasionally changed. However, in the control condition there are no differences between musicians and non-musicians. Reprinted with permission from Fujioka *et al.* (2004).

was lowered by a tone or a semitone in the other melody. Thus, the overall deviance rate was 50 per cent, so if separate streams were not formed, no MMN response would be expected because standard and deviant trials were equally probable. However, significant MMN was found for changes in both melodies, indicating that the two melodies were encoded separately. In the case of two simultaneous melodies, as with single melodies (Fujioka *et al.* 2004), musicians showed more robust MMN responses than non-musicians. Interestingly, for both groups, high and low melodies were not encoded equally well, as MMN was larger for the higher voice, consist-

ent with the compositional practice of putting the melody in the highest voice, and with people's superior perception of the highest voice (Crawley *et al.* 2002). In a follow up study, Fujioka *et al.* (2008) showed that the superiority of the higher voice is a very general phenomenon, also holding in the situation where each of the two voices consists of a single pitch. Finally, across a number of studies, no difference in MMN amplitude or latency was found for changes that went outside the key of a melody compared to changes that remained within the key (Trainor *et al.* 2001; Fujioka *et al.* 2004, 2005), suggesting that some aspects of culture-specific musical structure are processed beyond the auditory areas responsible for MMN generation.

While we have focused on the MMN as a neural correlate of musical expectancy, it is important to note that there are culture-specific aspects of musical expectancy which are processed in brain areas beyond the auditory regions responsible for MMN generation. For example, in Western musical structure chords follow each other according to genre-specific syntactic rules. Koelsch and his colleagues have demonstrated that a syntactically unexpected chord, such as a Neapolitan chord at the end of a cadence where a tonic chord is expected, gives rise to an early right anterior negativity (ERAN) (Koelsch *et al.* 2000; cf. Patel *et al.* 1998). ERAN is similar to MMN, but occurs somewhat later, is sensitive to harmonic expectations (Koelsch *et al.* 2001), and appears to involve inferior frontolateral cortex (Broca's area) (Maess *et al.* 2001). ERAN is also elicited by unexpected modulations in key (Koelsch *et al.* 2003). As with MMN responses to single and polyphonic melodies (Fujioka *et al.* 2004, 2005), ERAN responses to unexpected chords are present in both musicians and nonmusicians, although larger in musicians (Koelsch *et al.* 2002a). Furthermore, ERAN for unexpected chords is present even when other factors such as simple note repetition and degree of dissonance are controlled for (Koelsch *et al.* 2007; Leino *et al.* 2007). Finally, functional magnetic resonance imaging (fMRI) studies reveal that musical syntax activates a network of brain regions that resembles that for linguistic syntax (e.g., Koelsch *et al.* 2002b; cf. Patel 2003), suggesting further that music makes use of general processing properties of the cortex.

In summary, MMN and ERAN responses demonstrate that, for musical input, the brain is continually using the recent past to predict the future, and the magnitude of these responses reflects the degree to which these expectations are not fulfilled. MMN is sensitive to changes in acoustic features and patterns of sounds, and several MMN processes can occur simultaneously. ERAN is sensitive to culture-specific knowledge that is built up through exposure to a particular music system.

## The early development of musical expectations in infants

ERP responses to sound are very immature in infancy, reflecting the protracted development of the auditory cortex that extends into the teenage years (Huttenlocher and Dabholkar 1997; Moore 2002; Moore and Guan 2001; Ponton *et al.* 2000; Shahin *et al.* 2004; Trainor *et al.* 2003). Despite this great immaturity, robust mismatch responses can be measured from infants, suggesting that infant MMN is among the first developing cortical responses to sound (Näätänen *et al.* 2007). We argued in the introduction that being able to accurately predict the future is of paramount importance to survival. From this perspective, it makes sense, then, that mismatch responses would be among the earliest cortical responses, as they may be essential for optimal wiring of the brain through experience.

In very young infants, the ERP wave form is dominated by a frontally positive, anterior-negative slow wave. Interestingly, when a repeating sound is changed in pitch (e.g., Leppänen *et al.* 1997), fine temporal structure (e.g., Trainor *et al.* 2003), or vowel category (Dehaene-Lambertz and Baillet 1998), the slow wave increases in positivity (see Figure 16.3). An adult-like MMN emerges around 4 months of age or so in response to sound feature changes (He *et al.* 2007a; Trainor *et al.* 2001, 2003a) (see Figure 16.3). The slow wave and adult-like MMN probably represent different processes as both can be seen in infants of intermediate age (He *et al.* 2007a). Furthermore, the MMN in 4-month-olds increases with larger pitch

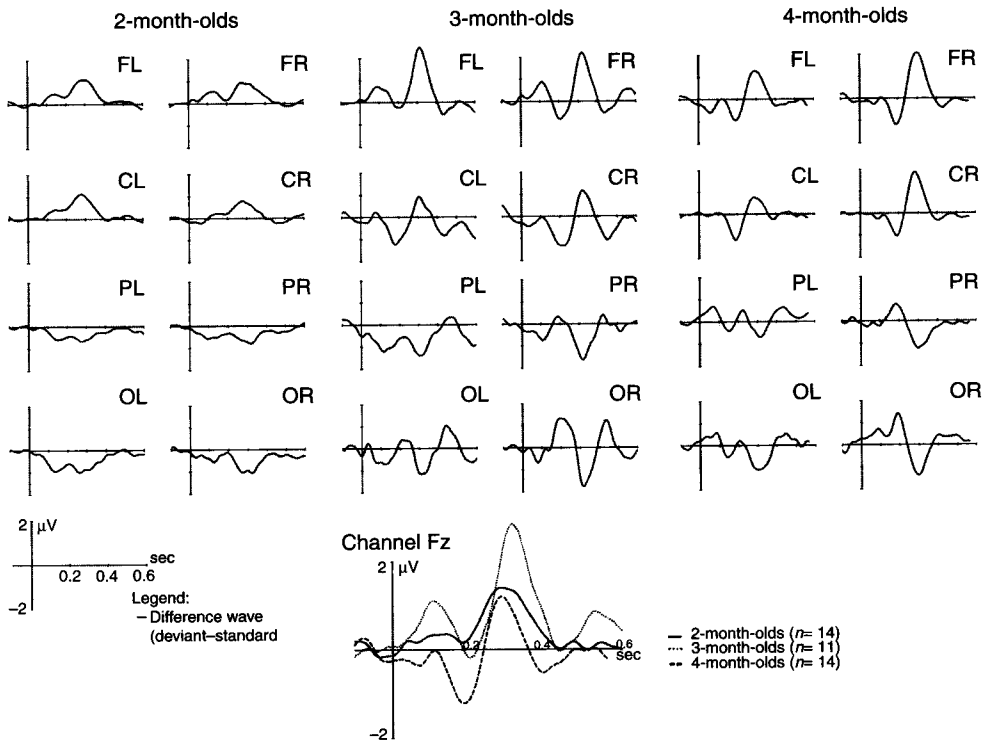



**Fig. 16.3** Difference waves (deviant–standard) to occasional half-octave changes in the pitch of a repeating standard tone in 2-month-olds, 3-month-olds and 4-month-olds at different scalp locations: F, frontal; C, central; P, parietal; O, occipital; L, left; R, right. In 2-month-olds, a slow positive difference wave is seen. In 3- and 4-month-olds a faster negative response resembling adult MMN can be seen followed by a positivity resembling P3a in adults. Both components reverse polarity at the back of the head. MMN becomes larger with increasing age. Reprinted with permission from He *et al.* (2007).

changes and remains robust at fast presentation rates, similar to adult MMN (He *et al.* in press).

These mismatch responses allow investigation of the developmental origins of musical expectations. These studies are just beginning. To date, it is clear that a repeating tone of one pitch sets up expectations for future repetitions of that pitch in infants (e.g., Alho *et al.* 1990; Ceponiené *et al.* 2002; He *et al.* 2007a, in press; Leppänen *et al.* 1997; Morr *et al.* 2002). Furthermore, 4-month-olds show robust MMN responses to the pitch of occasional deviants with missing fundamental components, indicating that they integrate the frequency content of complex tones into a single percept (He and Trainor 2007). Of most interest for musical processing, 4-month-olds also exhibit an MMN response when there is a change in a sequential pitch pattern (He *et al.* 2007b). Specifically, when the order of tones in a repeating two-tone sequence is occasionally reversed, a clear MMN is seen.

In sum, mismatch responses are among the earliest cortical responses to sound. The groundwork has now been established so that future studies can investigate in much more detail how musical expectations develop in the auditory cortex of human infants, and examine questions such as the effects of particular auditory experience on the development of these expectations.

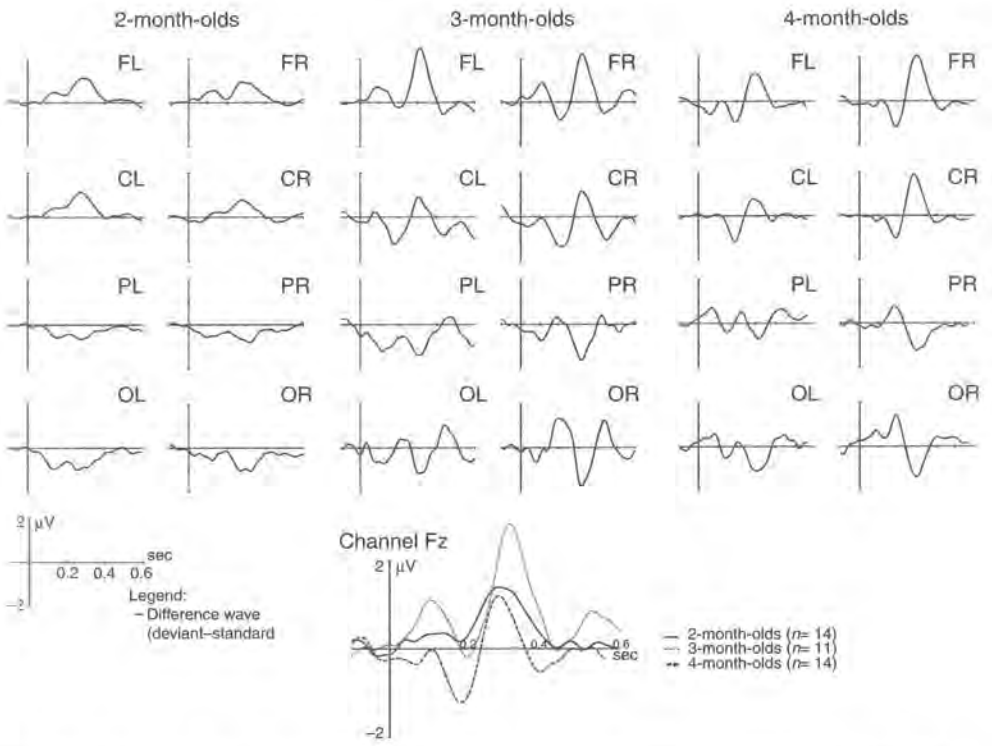



**Fig. 16.3** Difference waves (deviant–standard) to occasional half-octave changes in the pitch of a repeating standard tone in 2-month-olds, 3-month-olds and 4-month-olds at different scalp locations: F, frontal; C, central; P, parietal; O, occipital; L, left; R, right. In 2-month-olds, a slow positive difference wave is seen. In 3- and 4-month-olds a faster negative response resembling adult MMN can be seen followed by a positivity resembling P3a in adults. Both components reverse polarity at the back of the head. MMN becomes larger with increasing age. Reprinted with permission from He *et al.* (2007).

changes and remains robust at fast presentation rates, similar to adult MMN (He *et al.* in press).

These mismatch responses allow investigation of the developmental origins of musical expectations. These studies are just beginning. To date, it is clear that a repeating tone of one pitch sets up expectations for future repetitions of that pitch in infants (e.g., Alho *et al.* 1990; Ceponiené *et al.* 2002; He *et al.* 2007a, in press; Leppänen *et al.* 1997; Morr *et al.* 2002). Furthermore, 4-month-olds show robust MMN responses to the pitch of occasional deviants with missing fundamental components, indicating that they integrate the frequency content of complex tones into a single percept (He and Trainor 2007). Of most interest for musical processing, 4-month-olds also exhibit an MMN response when there is a change in a sequential pitch pattern (He *et al.* 2007b). Specifically, when the order of tones in a repeating two-tone sequence is occasionally reversed, a clear MMN is seen.

In sum, mismatch responses are among the earliest cortical responses to sound. The groundwork has now been established so that future studies can investigate in much more detail how musical expectations develop in the auditory cortex of human infants, and examine questions such as the effects of particular auditory experience on the development of these expectations.

## Temporal expectations and auditory–motor interactions: predicting *when* events will occur

As music unfolds over time, the brain extracts rhythmic regularities and sets up expectancies for when events will occur as well as for what the events will be. Time structure in music involves two aspects. First, there is the sequence of sound events and silence durations, which is referred to as the *rhythmic pattern*. Second, when a pattern contains temporal regularities, listeners infer an underlying regular hierarchical beat structure (what you tap your foot to), called the *metrical structure*. Metrical structure is not given in the stimulus, and there may not actually be sound events on some beats; rather, metre is derived by the brain from temporal regularities in the rhythm pattern. Metrical structure is hierarchical, with the strong beats being spaced progressively further apart in time at higher levels of the hierarchy, such that you can clap or tap, for example, on every beat, every second beat, or every fourth beat of a rhythm pattern organized into binary groups. The strong expectancies set up by metrical structure are evident in that the sense of a beat can continue past the end of a piece of music. One plausible theory as to how this happens is that neural circuits in the brain act as oscillators that can adjust their phase and frequency to some extent in response to the input from rhythmic patterns (Large and Jones 1999). Evidence for the neurological instantiation of temporal expectancy comes from studies of the induced gamma band response. Gamma band activity is in the range of 40–80 Hz, and can be extracted from an EEG recording by doing a Fourier analysis. The evoked gamma-band response occurs shortly after the onset of a sound and is tightly time-locked to the sound onset (e.g., Ross *et al.* 2005). The induced gamma-band response, on the other hand, is only loosely time-locked to the sound, and therefore has a different phase on every trial (e.g., Tallon-Baudry *et al.* 1996). The induced gamma-band response appears to reflect processes of intrinsic temporal expectation as it can be measured at points in time where a sound is expected, even if the sound event is not realized and there is only silence (Snyder and Large 2005). Furthermore, the induced gamma-band response is larger in musicians than in non-musicians, especially for sounds of their instrument of practice, and it also develops to a greater extent over the course of a year in young children taking music lessons compared to children not studying music (Shahin *et al.* 2008).

Temporal expectancies are important both in listening to music and in performing music (Large and Palmer 2002). A number of researchers have noted the close interaction between movement and auditory rhythm (Phillips-Silver and Trainor 2007, 2008; Trainor *et al.* in press; Repp 2005; Todd *et al.* 2007), even in infants (Phillips-Silver and Trainor 2005) and in non-human species (Todd *et al.* 2007). Indeed the close connection between music and dance suggests that musical rhythm might have evolved from rhythmic movement. A number of researchers, working on both speech and music, have postulated that *perception* of timing involves interactions between auditory and motor systems (Hickok and Poeppel 2004; Phillips-Silver and Trainor 2005, 2007, 2008; Repp 2005; Warren *et al.* 2005). This idea is related to research in vision, where similar neural responses exist both to observing an action and to performing the action (Rizzolatti and Craighero 2004).

A number of recent imaging studies dealing with musical perception and production suggest that the two processes may be closely intertwined; for instance, MEG data show that activity in the vicinity of the primary motor cortex could be evoked in pianists when they listened passively to well-known melodies (Haueisen and Knosche 2001). Two other fMRI studies (Bangert *et al.* 2006; Baumann *et al.* 2005) also reported that neural regions overlapped when pianists either listened to music without playing it, or played on a keyboard without any feedback; the overlap included the premotor cortex, supplementary motor area and the posterior auditory cortex. Increased motor excitability in the primary motor cortex of pianists has also been observed as they listened to a known piano piece, compared to a piece they were not trained on (D'Ausilio *et al.* 2006).

The phenomena are not limited to trained musicians, but can also be found amongst non-musicians. For example, a common pattern of EEG activity in auditory and motor areas was found in a task in which non-musicians were trained to play a simple melody on a keyboard (Bangert and Altenmuller 2003). The effect was only apparent when sound–movement mappings were congruent, and not when the learning did not result in a consistent assignment of key-presses to sounds. Similarly Lahav *et al.* (2007) used fMRI to show that premotor cortex and Broca's area were active when non-musicians heard melodies that they had previously learned to play, but not upon hearing equally familiar but motorically untrained melodies. Other studies have also noted premotor cortex activity during music listening under less specific circumstances, such as in melodic discrimination (Brown and Martinez 2007), or while listening to consonant musical excerpts (Koelsch *et al.* 2006). These effects have been interpreted as reflecting possible subvocal rehearsal, a phenomenon also suggested to occur during musical imagery tasks (Zatorre and Halpern 2005). Thus, it is now fairly well established that auditory–motor coactivation can be observed in musicians and in non-musicians, either spontaneously, or more specifically when there is a direct learned mapping between movement and sound.

Of most interest with respect to the involvement of auditory–motor interactions in the prediction of *when* events will happen, is the idea that interactions between posterior auditory cortices and premotor cortices might mediate the cognitive representations that are responsible for integrating feedforward and feedback information during performance and perception (Zatorre *et al.* 2007). Indeed, higher-order temporal organization (metricality) might emerge from the temporal predictions that are enabled by this system. Evidence for this comes from Chen *et al.* (2006) who found that the degree to which auditory input modulates motor behaviour and premotor cortex activity is enhanced by metrical saliency (the intensity of accented beats). In other words, increased predictability of auditory events in music leads to greater auditory–motor interaction. In a related study (Chen *et al.* 2008), a similar region of premotor cortex was also found to be recruited as subjects reproduced progressively more complex rhythmic movements. These studies manipulated the abstract metrical structure of the rhythms, and therefore the organization of the movement sequences (i.e., their predictability). One hypothesis to account for these results is that what modulates premotor activity in these instances is not the direct mapping of sounds to movements, but rather, the selection of movements based on the auditory cue. The dorsal segment of premotor cortex would, according to this view, be involved in extracting higher-order features of the musical sound, including its rhythmic metrical structure, therefore allowing implementation of temporally organized actions. Listening to music may activate motor programmes directly associated with producing the music, for example when a musician knows the piece he or she is listening to; but perhaps more interesting, listening to music also appears to engage auditory–motor interactions that extract higher-order metrical information. This latter mechanism may be critical in setting up the temporal expectancies that are essential for musical understanding (Huron 2006). As such, the interplay between auditory processing streams and the premotor system could be thought of as the temporal counterpart of the pitch-based expectancies discussed above, which rely on the interplay between auditory and memory systems.

## Future directions: understanding the neurobiology behind music expectations and emotional responses

In 1956, Meyer argued convincingly that music induces emotional experiences through physiological responses to musical expectations and the degree to which they are realized. In the present paper, we argue that music makes use of general brain mechanisms devoted to encoding the recent past, predicting the future, and dynamically adjusting the encoding of the recent past when the prediction is not accurate. We suggest, further, that such predictive processes

are essential for adaptation, plasticity, and learning, because evaluation of the success of the prediction provides a learning signal by which neural networks can change in order to more accurately and usefully encode the sensory input. We provided extensive physiological evidence from EEG, fMRI and MEG studies that music makes use of these general-purpose mechanisms, and that musical structure constantly engages the brain in a game of prediction. Some aspects of musical prediction, such as using statistical properties of recent input, are probably relatively independent of specific experience with a musical system whereas others, such as predicting the next chord in a sequence, probably rely on experience with a specific musical system. In particular, the framework that we have presented, in which MMN responses to unexpected events occur automatically and without attention, can explain why a note or chord that is musically unexpected continues to evoke an emotional response even when we are familiar with the piece and know at a conscious level that the unexpected chord is coming.

Huron (2006) suggests that not only does the brain make predictions about the future, but that there are intrinsic reward systems for correct prediction. With less expected musical realizations, a physiological arousal occurs, but, after appraisal, it is evident that there are no 'bad' consequences from the poor prediction because music is not about events in the world. So the emotion we feel tends to be strong and positive. To date, there is little physiological data linking violation of musical expectation and emotional responses. However, one study revealed that subjective ratings of tension and emotion, electrodermal activity, and an early negativity in the ERP all increased more for very unexpected than for somewhat unexpected sound events (Steinbeis *et al.* 2006). It remains for future research, however, to uncover the pathways in the nervous system that link the perception of unexpected events with our emotional responses to them.

## Acknowledgments

The writing of this chapter was supported by grants from the Natural Science and Engineering Research Council of Canada and the Canadian Institutes of Health Research.

## References

Alho K, Sainio K, Sajaniemi N, Reinikainen K and Näätänen R (1990). Event-related brain potential of human newborns to pitch change of an acoustic stimulus. *Electroencephalography and Clinical Neurophysiology*, **77**, 151–155.

Bangert M and Altenmuller EO (2003). Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neurosci*, **4**, 26.

Bangert M, Peschel T, Schlaug G, Rotte M, Drescher D, Hinrichs H *et al.* (2006). Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *Neuroimage*, **30**(3), 917–926.

Baumann S, Koeneke S, Meyer M, Lutz K and Jancke L (2005). A network for sensory-motor integration: what happens in the auditory cortex during piano playing without acoustic feedback? *Annals of the New York Academy of Sciences*, **1060**, 186–188.

Bendor D and Wang X (2005). The neuronal representation of pitch in primate auditory cortex. Nature, 436(7054), 1161.

Brown S and Martinez MJ (2007). Activation of premotor vocal areas during musical discrimination. Brain and Cognition, 63(1), 59.

Ceponiené R, Kushnerenko E, Fellman V, Renlund M, Suominen K and Näätänen R (2002). Event-related potential features indexing central auditory discrimination by newborns. *Cognitive Brain Research*, **13**, 101–113.

Chen J, Penhune V and Zatorre R (2008). Moving in time: brain networks for auditory-motor synchronization are modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience*, **20**, 226–239.

Chen JL, Zatorre RJ and Penhune VB (2006). Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. Neuroimage, 32(4), 1771–1781.

Crawley EJ, Acker-Mills BE, Pastore RE and Weil S (2002). Change detection in multi-voice music: the role of musical structure, musical training, and task demands. *Journal of Experimental Psychology: Human Perception and Performance*, **28**, 367–378.

Csépe V, Karmos G and Molnár M (1989). Subcortical evoked potential correlates of early information processing: mismatch negativity in cats. In E Basar, TH Bullock, eds, *Springer series in brain dynamics*, vol. 2, 278–289. Springer Verlag, Berlin.

D'Ausilio A, Altenmuller E, Olivetti Belardinelli M and Lotze M (2006). Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *European Journal of Neuroscience*, **24**(3), 955–958.

Dehaene-Lambertz G and Baillet SA (1998). A phonological representation in the infant brain. *NeuroReport*, **9**, 1885–1888.

Desjardins RN, Trainor LJ, Hevenor SJ and Polak CP (1999). Using mismatch negativity to measure auditory temporal resolution thresholds. *NeuroReport*, **10**, 2079–2082.

Dowling WJ, Harwood DL (1986). *Music cognition*. Academic Press, Orlando, FL.

Fujioka T, Trainor LJ and Ross B (2008). Simultaneous pitches are encoded separately in auditory cortex: an MMNm study. *NeuroReport*, **19**, 361–366.

Fujioka T, Trainor LJ, Ross B, Kakigi R and Pantev C (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience*, **16**, 1010–1021.

Fujioka T, Trainor LJ, Ross B, Kakigi R and Pantev C (2005). Automatic encoding of polyphonic melodies in musicians and nonmusicians. *Journal of Cognitive Neuroscience*, **17**, 1578–1592.

Giard MH, Lavikainen J, Reinikainen K, Perrin F, Bertrand O, Thévenet M *et al.* (1995). Separate representation of stimulus frequency, intensity, and duration in auditory sensory memory. *Journal of Cognitive Neurosciences*, **7**, 133–143.

Griffiths TD, Buchel C, Frackowiak RS and Patterson RD (1998). Analysis of temporal structure in sound by the human brain. *Nature Neuroscience*, **1**(5), 422–427.

Hannon EE and Trainor LJ (2007). Music acquisition: effects of enculturation and formal training on development. *Trends in Cognitive Sciences*, **11**, 466–472.

Haueisen J and Knosche TR (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience*, **13**(6), 786–792.

He C and Trainor LJ (2007). Finding the pitch of the missing fundamental in infants. Presented at the 8th Conference of the Society for Music Perception and Cognition, Montreal, QC, Canada.

He C, Hotson L and Trainor LJ (2007a). Mismatch responses to pitch changes in early infancy. *Journal of Cognitive Neuroscience*, **19**, 878–892.

He C, Hotson L and Trainor LJ (2007b). Changes in melodic pattern perception over the first few months after birth: electrophysiological evidence. Presented at the 8th Conference of the Society for Music Perception and Cognition, Montreal, QC, Canada.

He C, Hotson L and Trainor LJ (in press). Maturation of cortical mismatch responses to occasional pitch change in early infancy: Effects of presentation rate and magnitude of change. *Neuropsycholgia*.

Hickok G and Poeppel D (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, **92**(1–2), 67–99.

Huron D (2006). *Sweet anticipation: music and the psychology of expectation*. MIT Press, Cambridge, MA.

Huttenlocher PR and Dabholkar AS (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology*, **387**, 167–178.

Jaramillo M, Paavilainen P and Näätänen R (2000). Mismatch negativity and behavioural discrimination in humans as a function of the magnitude of change in sound duration. *Neuroscience Letters*, **290**, 101–104.

Johnsrude IS, Penhune VB and Zatorre RJ (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain*, **123**, 155–163.

Kaukoranta E, Sams M, Hari R, Hämäläinen M and Näätänen R (1989). Reactions of human auditory cortex to changes in tone duration. *Hearing Research*, **41**, 15–22.

Koelsch S, Fritz T, von Cramon DY, Müller K and Friederici AD (2006). Investigating emotion with music: an fMRI study. *Human Brain Mapping*, **27**(3), 239–250.

Koelsch S, Gunter T, Schröger E and Friederici AD (2003). Processing tonal modulations: an ERP study. *Journal of Cognitive Neuroscience*, **15**, 1149–1159.

Koelsch S, Gunter TC and Friederici AD (2000). Brain indices of music processing: 'nonmusicians' are musical. *Journal of Cognitive Neuroscience*, **13**, 520–541.

Koelsch S, Gunter TC, Schröger E, Tervaniemi M, Sammler D and Friederici A (2001). Differentiating ERAN and MMN: an ERP study. *Neuroreport*, **12**, 1385–1389.

Koelsch S, Gunter TC, Yves D, von Cramon DY, Zysset S, Lohmann G and Friederici AD (2002b). Bach speaks: a cortical 'language-network' serves the processing of music. *Neuroimage*, **17**, 956–966.

Koelsch S, Jentschke S, Sammler D and Mietchen D (2007). Untangling syntactic and sensory processing: an ERP study of music perception. *Psychophysiology*, **44**, 476–490.

Koelsch S, Schmidt B-H and Kansok J (2002a). Effects of musical expertise on the early right anterior negativity: an event-related brain potential study. *Psychophysiology*, **39**, 657–663.

Kraus N, McGee T, Littman T and King C (1994). Nonprimary auditory thalamic representation of acoustic change. *Journal of Neurophysiology*, **72**, 1270–1277.

Krumbholz K, Patterson RD, Seither-Preisler A, Lammertmann C and Lutkenhoner B (2003). Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. Cerebral Cortex, 13(7), 765–772.

Krumhansl CL (1997). Effects of perceptual organization and musical form on melodic expectancies. In M Leman, ed., *Music, gestalt, and computing: studies in cognitive and systematic musicology*, 294–320. Springer Verlag, Berlin.

Kujala T, Karma K, Ceponiene R, Belitz S, Turkkila P, Tervaniemi M *et al.* (2001). Plastic neural changes and reading improvement caused by audio-visual training in reading-imparied children. *Proceedings of the National Academy of Sciences*, **98**, 10509–10514.

Kujala T, Tervaniemi M and Schröger E (2007). The mismatch negativity in cognitive and clinical neuroscience: theoretical and methodological considerations. *Biological Psychology*, **74**, 1–19.

Lahav A, Saltzman E and Schlaug G (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience*, **27**(2), 308–314.

Large EW and Jones MR (1999). The dynamics of attending: how people track time-varying events. *Psychological Review*, **106**, 119–159.

Large EW and Palmer C (2002). Perceiving temporal regularity in music. *Cognitive Science*, **26**, 1–37.

Leino S, Brattico E, Tervaniemi M and Vuust P (2007). Representation of harmony rules in the human brain: further evidence from event-related potentials. *Brain Research*, **1142**, 169–177.

Leppänen PH, Eklund KM and Lyytinen H (1997). Event-related brain potentials to change in rapidly presented acoustic stimuli in newborns. *Developmental Neuropsychology*, **13**, 175–204.

Maess B, Koelsch S, Gunter T and Friederici AD (2001). 'Musical syntax' is processed in the area of Broca: an MEG-study. *Nature Neuroscience*, **4**, 540–545.

Meyer LB (1956). *Emotion and meaning in music.* University of Chicago Press, Chicago, IL.

Moore JK (2002). Maturation of human auditory cortex: implications for speech perception. *The Annals of Otology, Rhinology, and Laryngology*, **111**, 7–10.

Moore JK and Guan YL (2001). Cytoarchitectural and axonal maturation in human auditory cortex. *Journal of the Association for Research on Otolaryngology*, **2**, 297–311.

Morr ML, Shafer VL, Kreuzer JA and Kurtzberg D (2002). Maturation of mismatch negativity in typically developing infants and preschool children. *Ear and Hearing*, **23**, 118–136.

Näätänen R (1992). *Attention and brain function.* Lawrence Erlbaum Associates, Hillsdale, NJ.

Näätänen R and Winkler I (1999). The concept of auditory stimulus representation in neuroscience. *Psychological Bulletin*, **125**, 826–859.

Näätänen R, Paavilainen P, Alho K, Reinikainen K and Sams M (1989). Do event-related potentials reveal the mechanism of the auditory sensory memory in the human brain? *Neuroscience Letters*, **98**, 217–221.

Näätänen R, Paavilainen P, Rinne T and Alho K (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, **118**, 2544–2590.

Narmour E (1990). *The analysis and cognition of basic melodic structures: the implication-realization model.* University of Chicago Press, Chicago, IL.

Paavilainen P, Degerman A, Takegata R and Winkler I (2003). Spectral and temporal stimulus characteristics in the processing of abstract auditory features. *NeuroReport*, **14**, 715–718.

Paavilainen P, Karlsson M-L, Reinikainen K and Näätänen R (1989). Mismatch negativity to change in the spatial location of an auditory stimulus. *Electroencephalography and Clinical Neurophysiology*, **73**, 129–141.

Patel AD (2003). Language, music, syntax, and the brain. *Nature Neuroscience*, **6**, 674–681.

Patel AD, Gibson E, Ratner J, Besson M and Holcomb P (1998). Processing syntactic relations in language and music: an event-related potential study. *Journal of Cognitive Neuroscience*, **10**, 717–733.

Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron*, **36**, 767–776.

Penagos H, Melcher JR, Oxenham AJ (2004). A neural representation of pitch salience in nonprimary human auditory ortex revealed with functional magnetic resonance imaging. *Journal of Neuroscience*, **24**(30), 6810–6815.

Phillips-Silver J and Trainor LJ (2005). Feeling the beat in music: movement influences rhythm perception in infants. *Science*, **308**, 1430.

Phillips-Silver J and Trainor LJ (2007). Hearing what the body feels: auditory encoding of rhythmic movement. *Cognition*, **105**, 533–546.

Phillips-Silver J and Trainor LJ (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition*, **67**, 94–102.

Picton TW, Alain C, Otten L, Ritter W and Achim A (2000). Mismatch negativity: different water in the same river. *Audiology and Neuro Otology*, **5**, 111–139.

Ponton C, Eggermont JJ, Kwong B and Don M (2000). Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clinical Neurophysiology*, **111**, 220–236.

Raij T, McEvoy L, Mäkelä JP and Hari R (1997). Human auditory cortex is activated by omissions of auditory stimuli. *Brain Research*, **745**, 134–143.

Repp BH (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychonomic Bulletin and Review*, **12**, 969–992.

Rizzolatti G and Craighero L (2004). The mirror-neuron system. *Annual Review of Neuroscience*, **27**, 169–192.

Ross B, Herdman AT and Pantev C (2005). Stimulus induced desynchronization of human auditory 40-hz steady-state response. *Journal of Neurophysiology*, **94**, 4082–4093.

Ruusuvirta T, Korhonen T, Penttonen M, Arikoski J and Kivirikko K (1995). Hippocampal event-related potentials to pitch deviances in an auditory oddball situation in the cat: experiment I. *International Journal of Psychophysiology*, **20**, 33–39.

Saarinen J, Paavilainen P, Schröger E, Tervaniemi M and Näätänen R (1992). Representation of abstract stimulus attributes in human brain. *NeuroReport*, **3**, 1149–1151.

Schellenberg EG (1997). Simplifying the implication–realization model of musical expectancy. *Music Perception*, **14**, 295–318.

Schönwiesner M, Novitski N, Pakarinen S, Carlson S and Tervaniemi M (2007). Heschl's gyrus, posterior superior temporal gyrus, and mid-ventrolateral prefrontal cortex have different roles in the activation of acoustic change. *Journal of Neurophysiology*, **97**, 2075–2082.

Shahin A, Roberts LE and Trainor LJ (2004). Enhancement of auditory cortical development by musical experience in children, *NeuroReport*, **15**, 1917–1921.

Shahin A, Roberts LE, Chau W, Trainor LJ and Miller LM (2008). *Musical training leads to the development of timbre-specific gamma band activity. NeuroImage*, **41**, 113–122.

Snyder JS and Large EW (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Brain Research: Cognitive Brain Research*, **24**, 117–126.

Sonnadara RR, Alain C and Trainor LJ (2006a). Effects of spatial separation and stimulus probability on event-related potentials elicited by occasional changes in sound location. *Brain Research*, **1071**, 175–185.

Sonnadara RR, Alain C and Trainor LJ (2006b). Occasional changes in sound location enhance middle-latency-evoked responses. *Brain Research*, **1076**, 187–196.

Steinbeis N, Koelsch S and Sloboda JA (2006). The role of harmonic expectancy violations in musical emotions: evidence from subjective, physiological, and neural responses. *J Cognitive Neuroscience*, **18**, 1380–1393.

Tallon-Baudry C, Bertrand O, Delpuech C and Pernier J (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *Journal of Neuroscience*, **16**, 4240–4249.

Tervaniemi M, Ilvonen T, Karma K, Alho K and Näätänen R (1997a). The musical brain: brain waves reveal the neurophysiological basis of musicality. *Neuroscience Letters*, **226**, 1–4.

Tervaniemi M, Rytkönen M, Schröger E, Ilmoniemi RJ and Näätänen R (2001). Superior formation of cortical memory traces for melodic patterns in musicians. *Learning and Memory*, **8**, 295–300.

Tervaniemi M, Saarinen J, Paavilainen P, Danilova N and Näätänen R (1994). Temporal integration of auditory information in sensory memory as reflected by the mismatch negativity. *Biological Psychology*, **38**, 157–167.

Tervaniemi M, Winkler I and Näätänen R (1997b). Pre-attentive categorization of sounds by timbre as revealed by event-related potentials. *NeuroReport*, **8**, 2571–2574.

Tillmann B and Lebrun-Guillaud G (2006). Influence of tonal and temporal expectations on chord processing and on completion judgments of chord sequences. *Psychological Research*, **70**, 345–358.

Tillmann B, Bigand E, Escoffier N and Lalitte P (2006). The influence of musical relatedness on timbre discrimination. *European Journal of Cognitive Psychology*, **18**, 343–358.

Todd NP, Cousins R and Lee CS (2007). The contribution of anthropometric factors to individual differences in the perception of rhythm. *Empirical Music Review*, **2**, 1–13.

Toiviainen P, Tervaniemi M, Louhivuori J, Saher M, Huotilainen M and Näätänen R (1998). Timbre similarity: convergence of neural, behavioral, and computational approaches. *Music Perception*, **16**, 223–241.

Trainor L, McFadden M, Hodgson L, Darragh L, Barlow J, Matsos L *et al.* (2003a). Changes in auditory cortex and the development of mismatch negativity between 2 and 6 months of age. *International Journal of Psychophysiology*, **51**, 5–15.

Trainor LJ, Gao X, Lei J, Lehtovarara K and Harris LR (in press). The primal role of the vestibular system in determining musical rhythm. *Cortex*.

Trainor LJ, McDonald KL and Alain C (2002). Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *Journal of Cognitive Neuroscience*, **14**, 430–442.

Trainor LJ, Samuel SS, Galay L, Hevenor SJ, Desjardins RN and Sonnadara R (2001). Measuring temporal resolution in infants using mismatch negativity. *NeuroReport*, **12**, 2443–2448.

Trainor LJ, Shahin A and Roberts LE (2003b). Effects of musical training on auditory cortex in children. In G Avanzini, C Faienze, D Miciacchi, L Lopez and M Majno, eds, *Annals of the New York Academy of Sciences*, **999**, 506–513.

Unyk AJ and Carlson JC (1987). The influence of expectancy on melodic perception. *Psychomusicology*, **7**, 3–23.

Uther M, Jansen DHJ, Huotilainen M, Ilmoniemi RJ and Näätänen R (2003). Mismatch negativity indexes auditory temporal resolution: evidence from event-related potential (ERP) and event-related field (ERF) recordings. *Cognitive Brain Research*, **17**, 685–691.

Warren JE, Wise RJ and Warren JD (2005). Sounds do-able: auditory-motor transformations and the posterior temporal plane. *Trends in Neuroscience*, **28**(12), 636–643.

Whitfield I (1985). The role of auditory cortex in behavior. In A Peters and E Jones, eds, *Cerebral cortex*. Vol. 4, Association and Auditory Cortices, 329–351. Plenum Press, New York.

Zatorre RJ (1985). Discrimination and recognition of tonal melodies after unilateral cerebral excisions. *Neuropsychologia*, **23**, 31–41.

Zatorre RJ (1988). Pitch perception of complex tones and human temporal-lobe function. *Journal of the Acoustical Society of America*, **84**(2), 566–572.

Zatorre RJ and Halpern AR (2005). Mental concerts: Musical imagery and auditory cortex. *Neuron*, **47**, 9–12.

Zatorre RJ, Chen JL and Penhune VB (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, **8**(7), 547–558.

Zatorre RJ, Evans AC and Meyer E (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience*, **14**, 1908–1919.

CHAPTER 17

# Disorders of musical cognition

Lauren Stewart, Katharina von Kriegstein, Simone Dalla Bella, Jason D. Warren and Timothy D. Griffiths

## Introduction

The study of the neural underpinnings of musical perception and cognition has advanced greatly since Critchley and Henson published their edited volume *Music and the brain* (1977). Technical advances in functional neuroimaging and the development of theoretical models of musical processing (Peretz and Coltheart 2003) have allowed us to gain an understanding how musical processing occurs in the normal human brain. In parallel with this, the development of musical assessment tools such as the Montreal Battery for the Evaluation of Amusia (MBEA) (Peretz 2003) and the use of MRI to precisely define where lesions occur, allow us to determine which parts of these normal networks are critical for individual aspects of musical processing.

Here we present a comprehensive overview of case studies of acquired disorders of musical listening. The precise details of many of these published case studies are given in a previous review of disordered musical listening (Stewart *et al.* 2006). The current chapter draws general principles from those case studies, as well as presenting and commenting on a number of additional cases of disordered musical production.

## Acquired deficits in musical listening

The identification of cases of acquired disorders of musical listening based on symptom profiles in individual patients constitutes the traditional 'symptom-led' approach. An alternative approach to the clinical study of musical deficits has been advocated, based on the study of cases selected on the basis of a particular brain lesion, rather than clinical symptoms: the 'lesion-led' approach. Tables 17.1 and 17.2 refer to symptom-led and lesion-led reports respectively. In the discussion that follows, lesion-led cases will be considered as an adjunct to our primary focus on symptomatic disorders of musical listening.

Examination of the individual cases listed in Table 17.1 demonstrates that a deficit in musical listening can arise as a consequence of a central disturbance of auditory processing: of all the musical listening deficits documented, none could be attributed to a peripheral hearing deficit. As Figure 17.1 shows, the brain lesions that produce deficits in musical listening are widely distributed, with a preponderance of locations in the right hemisphere. However, even though right-sided lesions are more commonly associated with deficits in pitch and other domains, left sided lesions can also produce deficits in these aspects of musical listening. The preponderance of right hemispheric lesions associated with musical listening deficits may, at least partially, reflect a sampling bias: individuals with left hemisphere damage are often aphasic and testing of non-linguistic skills is often difficult and rarely a priority. The majority of cases are attributable to cerebrovascular events, though other pathologies such as focal cerebral degeneration ('progressive amusia') are represented

**Table 17.1** Acquired symptom-led reports

| | |
|---|---|
| Brust 1980 (case 2) (1980) | Mendez and Geehan (MS) (1988) |
| Confavreux *et al.* (1992) | Murayama *et al.* (2004) |
| Di Pietro *et al.* (2004) | Patel *et al.* (CN) (1998) |
| Eustache *et al.* (cases 1, 2) (1990) | Patel *et al.* (IR) (1998) |
| Fries and Swihart (1990) | Peretz *et al.* (CN) (1994) |
| Fujii *et al.* (1990) | Peretz *et al.* (CN) (1996) |
| Griffiths *et al.* (1997) | Peretz *et al.* (IR) (1997) |
| Griffiths *et al.* (2004) | Peretz *et al.* (IR) (1998) |
| Griffiths *et al.* (2006) | Peretz and Gagnon (IR) (1999) |
| Habib *et al.* (1995) | Peretz *et al.* (IR) (2001) |
| Hattiangadi *et al.* (2005) | Piccirilli *et al.* (2000) |
| Hofman *et al.* (1993) | Satoh *et al.* (2005) |
| Johannes *et al.* (1998) | Schon *et al.* (2004) |
| Johkura *et al.* (1998) | Spreen *et al.* (1965) |
| Kohlmetz *et al.* (2003) | Tramo *et al.* (MS) (1990) |
| Lechevalier *et al.* (1984) | Tanaka *et al.* (1987) |
| Levin and Rose (1979) | Terao *et al.* (2005) |
| Mavlov (1980) | Tramo *et al.* (MS) (2002) |
| Mazzoni *et al.* (1993) | Uvstedt (case 9) (1937) |
| Mazzucchi *et al.* (1982) | Wilson *et al.* (2002) |

(Confavreux *et al.* 1992). Because of the nature of these lesions, musical listening disorders are rarely 'pure': over half the cases are associated with disorders of speech perception, and approximately a third of cases with disorders of environmental sound perception. In most cases available data on speech processing are limited, preventing clear comment about the general association of directly related speech deficits (e.g. perceptual dysprosody) or other deficits in the speech domain. There is some evidence to suggest that the earlier stages of an acquired deficit in musical listening (e.g. less than one year after onset) tend to be accompanied by more deficits in listening to other classes of sounds suggesting a disorder in musical listening that can emerge as an isolated deficit following the recovery phase of a more generalized auditory agnosia.

Figure 17.1 shows that many areas are implicated in more than one function. However, it is clear that the necessary bases for music processing are separable: within the domain of music, relatively isolated deficits of pitch (Peretz *et al.* 1994), temporal (Mavlov 1980), timbral (Kohlmetz *et al.* 2003), mnemonic (Peretz 1996) and emotional (Griffiths *et al.* 2004) processing have all described. More fine-grained dissociations also occur: for instance, between pitch interval and pitch contour (Liegeois-Chauvel *et al.* 1998; Peretz 1990) and between rhythm and metre (Di Pietro *et al.* 2004; Wilson *et al.* 2002).

## Pitch: interval

Deficits in the analysis of pitch intervals (the detection of a pitch change and/or the discrimination of the direction of a pitch change) are most strongly associated with lesions involving lateral Heschl's Gyrus (HG) and non-primary auditory cortical areas in Planum Temporale (PT) and the parieto-temporal junction (Figure 17.1). The detection of pitch differences and the discrimination of pitch direction are functionally

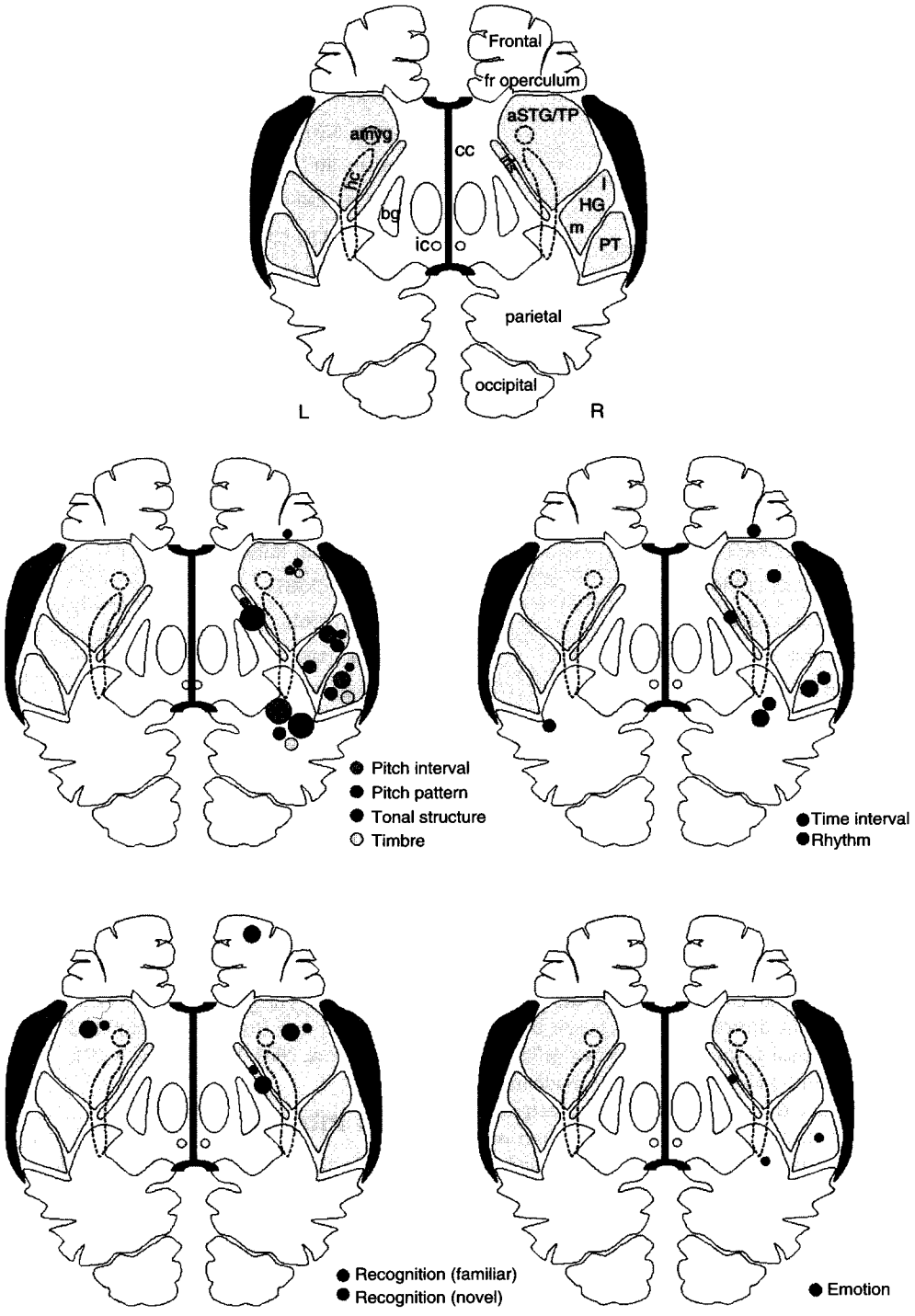



**Fig. 17.1** Critical brain substrates for musical listening disorders. This data is based on the symptomatic cases of musical listening disorder (Table 17.1). The detailed analysis of these cases is

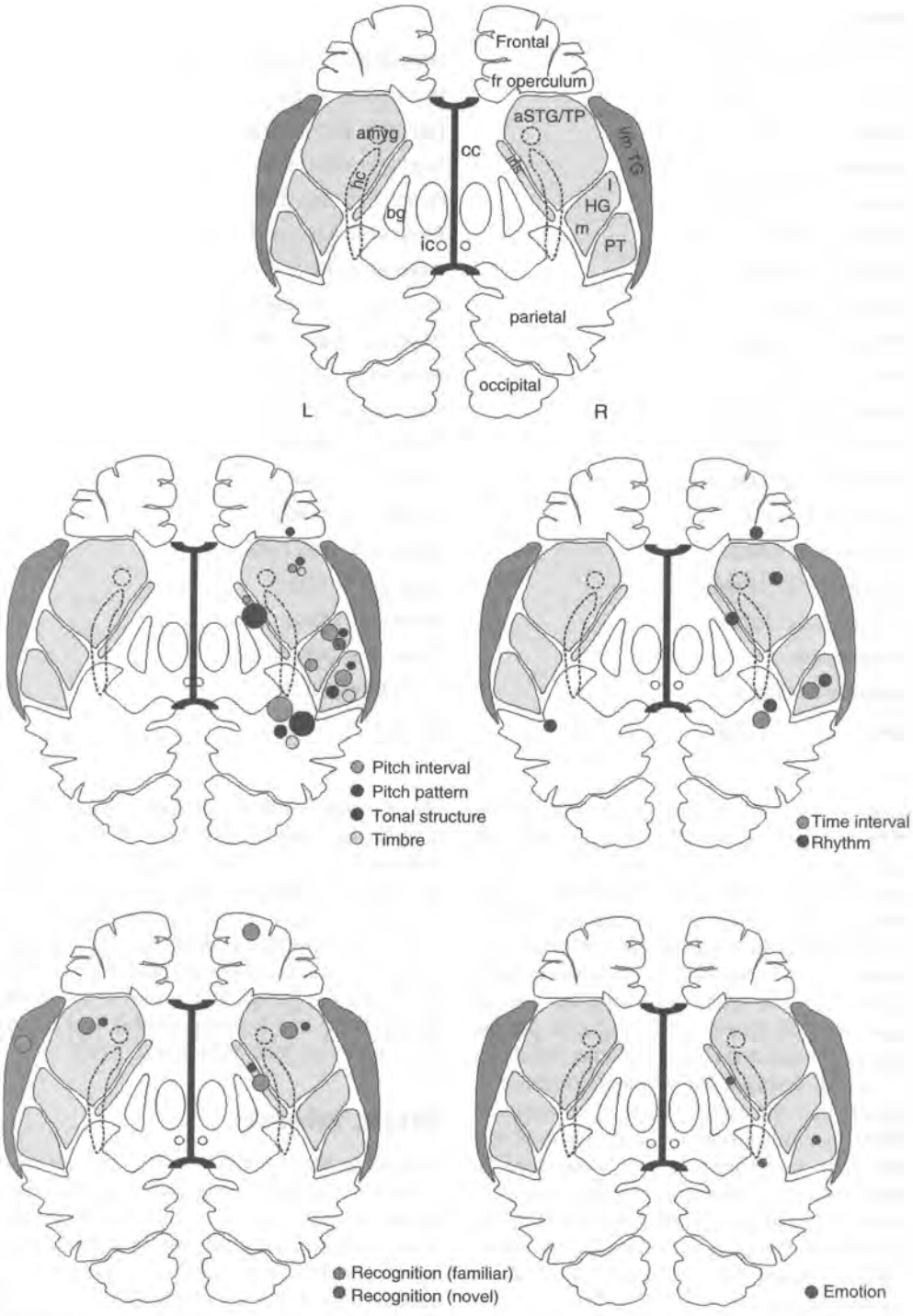



**Fig. 17.1** Critical brain substrates for musical listening disorders. This data is based on the symptomatic cases of musical listening disorder (Table 17.1). The detailed analysis of these cases is

separable, but have been distinguished only infrequently in symptom-led studies. Where it has been specifically assessed, impaired pitch-difference detection is generally associated with involvement of subcortical structures and ascending auditory pathways or primary auditory cortex in medial HG (Habib *et al.* 1995, Hattiangadi *et al.* 2005, Terao *et al.* 2006, Tramo *et al.* 2002) while impaired pitch-direction discrimination is generally associated with involvement of lateral HG (Lechevalier *et al.* 1984, Tanaka *et al.* 1987, Terao *et al.* 2005, Tramo *et al.* 2002).

These data are congruent with evidence from temporal lobectomy series (Table 17.2). Right-lateral HG resection leads to deficits in the detection of pitch-change direction in pure tones (Johnsrude *et al.* 2000) and complex tones (Zatorre 1988) These findings, especially the latter, suggest that aspects of pitch perception, as opposed to the sensory representation of frequency- and time- domain properties of the stimulus, depend on the integrity of cortical areas beyond primary auditory cortex.

## Pitch: pattern

Deficits in the analysis of pitch patterns, such as melodies, are common in acquired disorders of musical listening (Table 17.1). The deficits are most often associated with lesions posterior to HG, in PT and the parieto-temporal junction, and anterior to HG in anterior superior temporal gyrus (STG) (Figure 17.1). These lesions are more common on the right. Melody discrimination has also been widely assessed in the lesion-led literature (Table 17.2): right but not left temporal lobectomy impairs discrimination of pitch pattern, even where HG is not included (Milner 1962). However, inclusion of HG produces deficits in melody perception over and above those due to the resection of the temporal lobe anterior to it (Samson and Zatorre 1988, Zatorre 1985). In a study comparing resections involving posterior versus anterior STG (Liegeois-Chauvel *et al.* 1998), posterior resection was associated with more severe impairment of pitch pattern perception; this effect could not be attributed to involvement of HG. Right anterior temporal lobe resections have also been associated with impaired working memory for pitch (Zatorre and Samson 1991) which is likely to affect perception of long-term structure in melodies, and more specifically, discrimination tasks requiring comparisons between sequential stimuli.

Pitch contour (the pattern of 'ups' and 'downs' in a melody) and the associated absolute pitch values can be considered as corresponding to psychologically distinct 'global' and 'local' levels of pitch processing, respectively (Dowling and Harwood 1985). These levels can be probed by tasks which require discrimination of melodies with different pitch contour ('global' tasks) or different absolute pitch values with the same contour ('local' tasks), respectively. A 'pure' global task would require transposition

presented in supplementary Table 1 of Stewart *et al.* (2006). Five cartoons are shown, each depicting the brain in a schematic axial section that includes all key anatomical areas involved in music listening (identified on the top cartoon); the corpus callosum (black), superior temporal plane (light grey) and middle/inferior temporal gyri (dark grey areas, in exploded view) are coloured for ease of identification. Musical functions have been grouped as follows: pitch processing (pitch interval, pitch pattern, tonal structure, timbre); temporal processing (time interval, rhythm); musical memory (familiar and novel material); and emotional response to music. Each group of functions is assigned to a separate cartoon; individual functions are identified to the right of the corresponding cartoon. Raw data from Stewart *et al.* (2006) have been thresholded; the presence of a coloured circle corresponding to a particular function in a region indicates that at least 50 per cent of studies of the function implicate that region. The size of each circle is scaled according to the proportion of studies of the function implicating that region. Metre is not represented as no brain area was implicated in 50 per cent or more of cases. amyg, amygdala; aSTG, anterior superior temporal gyrus; bg, basal ganglia; cc, corpus callosum; fr, frontal; hc hippocampus; HG, Heschl's gyrus; i, inferior; ic, inferior colliculi; ins, insula; l, lateral; m, medial; PT, planum temporale; TG, temporal gyrus. See colour plate section. This figure is taken from Stewart L, Von Kriegstein K, Warren JD and Griffiths TD. Disorders of musical listening. *Brain* 2006, 129, 2533–2553, by permission of Oxford University Press.

**Table 17.2** Acquired lesion-led reports

| Resection cases | Stroke/other type of lesion |
|---|---|
| Gosselin *et al.* (2005) | Alcock *et al.* (2000) |
| Johnsrude *et al.* (2000) | Ayotte *et al.* (2000) |
| Kester *et al.* (1991) | Grossman *et al.* (1981) |
| Liegeois-Chauvel *et al.* (1998) | Kinsella *et al.* (1988) |
| Milner (1962) | Peretz (1990) |
| Samson and Zatorre (1988) | Prior *et al.* (1990) |
| Samson and Zatorre (1991) | Robin *et al.* (1990) |
| Samson and Zatorre (1992) | Samson *et al.* (2001) |
| Samson and Zatorre (1994) | Schuppert *et al.* (2000) |
| Samson *et al.* (2001) | Shapiro *et al.* (1981) |
| Samson *et al.* (2002) | Sidtis and Volpe (1988) |
| Shankweiler (1966) | Tramo and Bharucha (1991) |
| Warrier and Zatorre (2004) | |
| Zatorre (1985) | |
| Zatorre (1988) | |
| Zatorre and Halpern (1993) | |
| Zatorre and Samson (1991) | |

between the melodies compared (like a shift in key) to prevent the use of any local information. However, this renders the task much harder for non-musicians and the MBEA does not employ it. Peretz (1990) showed that patients with right cerebral hemisphere strokes could process neither global nor local information in melodies, while patients with strokes involving the left hemisphere could process global but not local information. Isolated impairments of local processing were observed, but isolated impairments of global processing were not, leading Peretz to propose an influential hierarchical model of co-operation between the hemispheres. According to this model, the right hemisphere derives pitch contour which is then elaborated upon by the left hemisphere which fills in the detailed pitch structure. In the model, prior contour processing is necessary for absolute pitch values to be processed: lesions involving the right hemisphere therefore compromize the processing of both contour and the absolute values of pitch whereas lesions involving the left hemisphere prevent absolute pitch values being added to the contour provided by the right hemisphere. This model was supported by the findings of a temporal lobectomy series (Liegeois-Chauvel *et al.* 1998). However, other neuropsychological and functional imaging evidence suggests that hemispheric differences are less clearcut (Schuppert *et al.* 2000, Stewart *et al.* 2008).

## Pitch: tonal structure

Tonal structure refers to rule-based patterns of pitch determined by key. While both tonal and atonal pitch patterns exhibit global and local structure, tonal structure is specific to tonal music (i.e., most Western music composed before the twentieth century). The most widely used test of this type of processing is the key violation task in the MBEA, which assesses the ability of subjects to detect notes that are out of key. Deficits in tonal analysis have been associated with damage involving a predominantly right-sided network of non-primary auditory cortical areas including lateral HG, PT, parieto-temporal junction, insula, anterior STG and frontal operculum (Figure 17.1). The frontal opercular involvement is proportionately greater

relative to its involvement in other aspects of pitch processing, but the number of studies is small.

## Timbre

Timbral perception has a number of dimensions that can be related to different acoustic properties of the incoming sound, and lesions affecting any or all of these dimensions could in principle lead to a deficit of timbre perception. Many clinical studies include reports of perceptual alterations in the perceived quality of music (often described as unpleasant, 'flat' or 'mechanical' in nature) or inability to recognize musical instruments, which might represent specific deficits of timbre perception. However, timbre as a distinct functional component of music has only been assessed infrequently in the clinical literature. This evidence suggests that a network of areas in the right superior temporal lobe that overlaps closely with areas implicated in pitch-pattern analysis is critical for normal timbre perception (Figure 17.1). Timbral deficits have generally been observed in conjunction with pitch-pattern deficits, however timbral deficits with spared pitch-pattern perception have also been described after strokes involving right STG (Kohlmetz *et al.* 2003, Mazzucchi *et al.* 1982). The deficits may extend to the discrimination of timbre in voices and environmental sounds (Mazzucchi *et al.* 1982).

These findings from the symptom-led literature are mirrored by studies of temporal lobectomy series (Table 17.2). Right temporal lobectomy leads to deficits in the perception of timbral change when this is in the spectral dimension (Milner 1962, Samson and Zatorre 1994), temporal dimension (Samson and Zatorre 1994), or both (Samson *et al.* 2002). This last study also suggested a subtle effect of left temporal lobectomy deficit on the processing of timbre associated with melodies but not single notes.

## Temporal structure: time interval, rhythm and metre

Like pitch perception, the perception of timing information in music can be analysed hierarchically: the lowest level of temporal processing that we consider here corresponds to the detection of simple durational differences in a tone or detection of a silent interval between two tones. These basic timing elements can be built into more complex structures embedded in metre and rhythm. Schuppert and colleagues (2000) propose a hierarchal scheme where metre and rhythm can be considered as global and local properties in the time domain. The symptom-led evidence implicates predominantly right-sided non-primary auditory cortical areas posterior to HG in the perception of time-interval differences, and bilateral areas widely distributed beyond HG in the perception of rhythm (Figure 17.1). Evidence from lesion-led studies broadly supports a bilateral organization: impaired durational processing has been described following right temporal lobectomy (Milner 1962) and impaired gap detection following left temporo-parietal strokes (Robin *et al.* 1990). Impaired detection of rhythmic violations has been described in left temporo-parietal stroke (Robin *et al.* 1990) and left hippocampal sclerosis (Samson *et al.* 2001), while other studies have not demonstrated laterality differences (Peretz, 1990, Schuppert *et al.* 2000, Shapiro *et al.* 1981). At least some of these apparent discrepancies in lateralization may reflect task effects: for example, the detection of rhythmic errors in familiar tunes (Prior *et al.* 1990, Samson *et al.* 2001) may require musical processing that is not required for the discrimination of simple rhythmic patterns.

Comparatively few studies of metrical processing have been conducted and evidence for a critical brain substrate remains inconclusive. This is likely to be due at least in part to the difficulty of assessing metre reliably, particularly in subjects without formal musical training. In the small symptom-led literature, impairments of metre perception have been associated with individual lesions widely distributed in both cerebral hemispheres. This pattern has been echoed in lesion-led studies. In temporal lobectomy series, Liegeois-Chauvel *et al.* (1998) found metrical impairments following left and right anterior temporal lobe resections, while Kester *et al.* (1991) found a specific decrement in performance following right but not left anterior temporal resection. Ayotte *et al.* (2000) and Peretz *et al.* (1990) both found that stroke patients with heterogeneous left and right hemisphere strokes

were not impaired relative to neurologically normal control subjects, while Schuppert *et al.* (2000) found that both left and right hemispheric stroke patients were impaired relative to controls.

## Memory: familiar and novel material

Impaired recognition of familiar tunes is described in a number of studies (Ayotte *et al.* 2000, Griffiths *et al.* 1997, Johannes *et al.* 1998, Lechevalier *et al.* 1984, Peretz *et al.* 1998, Peretz and Gagnon 1999, Peretz *et al.* 1994, Piccirilli *et al.* 2000, Satoh *et al.* 2005, Wilson *et al.* 2002). In all these cases, recognition problems were accompanied by impaired pitch-pattern perception. However, perception and recognition can sometimes dissociate. Patients may have intact perception but impaired recognition (Eustache *et al.* 1990 (case 1), Peretz 1996) or the converse (Eustache *et al.* 1990 (case 2); Schuppert *et al.* 2000).

Figures 17.1 demonstrates that deficits in the perception and recognition of familiar tunes may occur with damage in either cerebral hemisphere involving the anterior STG and insula (Lechevalier *et al.* 1984, Peretz *et al.* 1998, Peretz and Gagnon 1999, Peretz *et al.* 1994, Satoh *et al.* 2005). In the lesion-led literature, familiar-tune recognition was found to be deficient specifically in association with damage involving the right insula (Ayotte *et al.* 2000).

Impaired recognition of novel material is tested in the incidental memory test from the MBEA, which assesses implicit encoding and retrieval of novel musical material. Clinical impairments of musical incidental memory are associated with damage involving a bilateral network of areas that closely overlaps the network implicated in the recognition of familiar tunes (Figure 17.1), and extends into left middle and inferior temporal cortex. In temporal lobectomy series, incidental memory deficits have been described following both right and left anterior resections (Samson and Zatorre 1992, Zatorre 1985).

## Emotion

Most people listen to music purely for the aesthetic pleasure it brings. A loss of enjoyment in musical listening is a common presenting complaint in clinical disorders of musical listening. In many cases, this loss of pleasure is accompanied by a perceptual derangement: 'like an out of tune child's dulcimer', (Griffiths *et al.* 1997), 'mechanical' (Griffiths *et al.* 2006), or 'instruments [may lose] their distinctive features of timbre and sound dull' (Piccirilli *et al.* 2000). Associated impairments of pitch-pattern perception (Habib *et al.* 1995) and generalized auditory agnosia (Mazzucchi *et al.* 1982) have been documented. The most consistent association of altered emotional response to music across studies is damage involving the right posterior temporal lobe and insula (Figure 17.1). However, few clinical studies have specifically assessed musical emotion.

Just as perception and recognition for musical material may dissociate, so too may perception and emotional response to music. Isolated deficits of musical emotional response have been described: the patient of Griffiths *et al.* (2004) had been used to experiencing an emotional transformation while listening to Rachmaninov preludes (the 'shiver down the spine' phenomenon) but this was lost following an infarction involving left amygdala and insula. In a temporal lobectomy series (Gosselin *et al.* 2005), patients who had undergone resections of the left or right medial temporal lobe, including the amygdala but sparing the STG, found fearful music less scary than a group of matched controls. Conversely, intact emotional response despite impaired music perception has also been observed (Lechevalier *et al.* 1984, Peretz *et al.* 1998, Peretz and Gagnon 1999). The patient of Peretz *et al.* (1998) still derived pleasure from music, and was able to classify tunes as happy or sad and to discriminate tunes based on emotional tone despite severe perceptual and recognition impairments.

## Developmental deficits in musical listening

Case reports of lifelong tone deafness go back more than a century (Grant-Allen 1898), but it has only been in the last five years that the condition has undergone systematic investigation using the same tools that have been applied to

acquired disorders of musical listening. The disorder was first characterized in this way by Peretz and colleagues using the MBEA and given the label 'congenital amusia' (Ayotte *et al.* 2002). This is not the same as what is commonly known as 'tone-deafness', a term which many people apply to themselves, mostly based on the belief that they cannot sing in tune (Sloboda *et al.* 2005, Cuddy *et al.* 2005). Infact, given normal brain development and opportunities for informal singing, the majority of the population can sing in tune (but see Bradshaw and McHenry 2005, Dalla Bella *et al.* 2007 for evidence that singing output may be impaired despite normal perception). Those with congenital amusia, on the other hand, cannot sing in tune, but this is likely to be a consequence of them having a true perceptual agnosia, in which the perception of music is abnormal in the presence of normal hearing and otherwise preserved cognition.

## Pitch

Formal characterization of musical perception using the MBEA (Ayotte *et al.* 2002) demonstrates consistent deficits in the domain of pitch-pattern perception (pitch contour, absolute pitch value and key structure). Foxton *et al.* (2004) carried out tests in which types of pitch pattern that are more basic than melody are assessed. The group of subjects identified as having amusia using the MBEA were found to have abnormal perception of pitch change and pitch direction, the most striking changes being demonstrated for pitch direction. Most subjects in the amusic group had thresholds for the identification of pitch direction that were well above a semitone and likely relevant to the perception of Western music. This finding has been recently replicated in an independent group of amusic individuals (Griffiths *et al.* 2007). Given that pitch direction can be thought of as a 'building block' for contour (the pattern of 'ups' and 'downs' in a melody), this work suggests a fundamental deficit in pitch processing that is below the level of melody processing. However, the presence of such a causative deficit would predict that the deficit could be overcome by creating melodic sequences with large intervals. The fact that this is not the case (Foxton *et al.* 2004) argues against a simple low-level deficit in pitch direction as a single causal mechanism for melodic deficits in amusia. However, it remains possible that an inability to analyse pitch direction leads to a failure to develop normal pitch-pattern perception.

## Temporal structure

Subjects who have been recruited on the basis of lifelong musical difficulty and score below threshold on the pitch subtests of the MBEA may also have problems with 'following the beat' and with dancing, consistent with a deficit in the analysis of metre and rhythm. Hyde and Peretz (2003) found that although subjects with congenital amusia had a deficit in detection of pitch changes in an otherwise monotonic sequence, their detection of a change in the timing of a note in an otherwise regular sequence was normal. A subsequent study in a different group of amusics (Foxton *et al.* 2006) replicated the finding of normal detection of timing deviation within simple rhythms, but showed that when subjects were required to detect the same deviation in a melodic, as opposed to monotonic context, their performance was impaired. The results support a model in which the abnormal perception of pitch pattern in amusia also affects the perception of the temporal structure of music. It is also worth noting that the amusic individuals so far studied have been recruited using adverts that conceptualize musical difficulties in terms of tune rather than time. In contrast, a large number of individuals screened using an internet version of the scale and rhythm subtest of the MBEA (www.delosis.com/listening/home.html) suggest that certain individuals have specific difficulties with time, but not tune (unpublished observations). This is a potentially novel subtype of amusia with a distinct cognitive and neural basis.

## Emotion

A recent questionnaire-based study (McDonald and Stewart 2008) shows that, as a group, those with developmental amusia do not implement music in their everyday lives in the same way as age, gender and training matched controls. They regard the presence of music in public places more negatively than controls and report

significantly fewer emotional responses to music. However, there is considerable individual variation with respect to this. For instance, some amusics report enjoyment of music while others find it unpleasant. The dissociation between the perception of music and its self-reported emotional effects is consistent with evidence from normal functional imaging and clinical studies of acquired lesions that suggest distinct substrates for perception and appreciation but the reasons for such variability are presently unknown.

## Neural substrate

The brain basis for congenital amusia has not been extensively investigated. These individuals do not have a history of neurological damage, and structural brain imaging using MRI reveals no gross structural differences. The technique of voxel-based morphometry allows structural MRI data from two or more groups of individuals (e.g. amusic versus non-amusic individuals) to be interrogated with regard to potential regional differences in grey and white matter density. Recent studies have demonstrated structural differences in the inferior frontal lobe and temporal cortex (Hyde *et al.* 2005, 2007; Mandell *et al.* 2007) and EEG studies of responses to pitch changes have demonstrated normal N1 responses to tones but abnormal N2-P3 responses at longer latency (Peretz *et al.* 2005). The N1 response arises from the PT whilst the longer latency responses have a number of more distributed generators. Congruent with the work based on structural MRI, this work suggests brain abnormalities in amusia that may occur in areas quite distinct from auditory cortex. Further studies will be required to ascertain the locus and nature of the abnormality in these developmental cases.

## Acquired disorders of musical production

Music production deficits such as impaired singing (vocal amusia or oral-expressive amusia), or deficient music performance on an instrument (expressive instrumental amusia or musical apraxia) have been generally referred to as 'expressive amusia' (Benton 1977). Although cases of musical production deficits have been described since the nineteenth century (reviewed in Benton 1977), there is a paucity of studies of instrumental amusia (Botez and Wertheim 1959; McFarland and Fortin 1982). For this reason, the following discussion will focus on acquired deficits in singing.

Impaired vocal performance (e.g., poor singing of well-known songs) consequent to brain damage has been documented in skilled professional singers and in nonmusicians (for reviews, see Ackermann *et al.* 2007; Gordon *et al.* 2006; Marin and Perry 1999). Early case reports indicate that lesions of the right-hemisphere fronto-insular cortex disrupt the ability to sing, hum or whistle a tune (Mann 1898, 1933; Jossmann 1926, 1927, cited in Benton 1977; Botez and Wertheim 1959). For example, Mann (1898 cited in Benton 1977) reported impaired abilities to sing and whistle songs in a professional singer following injury of the right frontal lobe. Despite the dramatic vocal expression disorder, the patient was still able to recognize familiar songs and did not show any signs of aphasia. Similar cases of musicians exhibiting poor singing in absence of language disorders and with relatively spared music perception and recognition were reported by Jossmann (1926, 1927) and Botez and Wertheim (1959). These findings are consistent with the observation that unilateral inactivation of the right hemisphere (i.e., with the Wada test, see Gordon and Bogen 1974) alter the ability to sing, hum or whistle a tune. Unfortunately, however, most of these case descriptions are anecdotal, thus lacking systematic assessment of music production and perception abilities.

More systematic studies of singing deficits (both symptom-led and lesion-led) are included in Tables 17.1 and 17.2. Confavreux *et al.* (1992) report on a patient who suffered focal cerebral degeneration of right-hemisphere regions involving the anterior temporal gyrus and the insula. Music perception abilities were relatively preserved in the presence of impaired singing and expressive aprosody in speech. More recently, Schön and collaborators (2004) presented the case of IP, a tenor singer with right hemisphere lesions distributed in the inferior frontal gyrus, posterior temporal lobe and inferior parietal lobe. IP is a pure case of vocal amusia exhibiting a specific deficit of in the production of musical intervals. In contrast, production of

rhythm and contour were spared, as was musical perception and language abilities. In order to examine the role of the right and left hemispheres in pitch and rhythm vocal production, Alcock *et al.* (2000) examined vocal performance and music perception in a group study of patients with unilateral left- or right-hemisphere lesions. They found that left-hemisphere patients were mostly impaired in perceiving and reproducing rhythms, without difficulties in singing the correct pitch. Right-hemisphere patients, in contrast, had great difficulties in pitch production and perception tasks, with less marked rhythm processing disorders than left-hemisphere patients. Additional evidence confirming that the right hemisphere is engaged in pitch production comes from two recent single case studies in which acoustical measures of singing proficiency were obtained (Murayama *et al.* 2004; Terao *et al.* 2006). In sum, evidence from symptomatic case studies and lesion-led reports, in general, confirm what has been described in early reports, that the right hemisphere (mostly the frontal-insular regions) is central to vocal musical pitch production. In addition, vocal performance disorders can occur in a pure form in absence of perceptual and linguistic deficits, and may concern very specific aspects of music vocal production (e.g., interval production) while leaving other functions intact.

Further evidence indicating a predominant right-hemisphere involvement in vocal pitch performance comes from neuroimaging and TMS studies. In an fMRI study, Riecker *et al.* (2000) found that speech and singing without lyrics engage opposite hemispheres (at the level of anterior insula), with a left-hemisphere dominance for speaking and a right-hemisphere dominance for singing. A similar lateralization pattern was yielded by a PET study in which speaking and singing with words were contrasted (Jeffries *et al.* 2003). In addition, when TMS was applied over left-hemisphere regions traditionally related to speech production (e.g., near Broca's area) speech was disrupted; similar stimulation over homologous brain areas in the right hemisphere impaired singing, at least for some of the participants (Epstein *et al.* 1999; Lo and Fook-Chong 2004; Stewart *et al.* 2001).

A classical interpretation of these findings is that singing familiar songs engages right-hemisphere regions in contrast to propositional speech which engages the left-hemisphere. This account has gained credence from clinical reports that patients suffering from motor aphasia are still able to sing previously learned songs with well-articulated and linguistically intelligible words (e.g., Assal *et al.* 1977; Yamadori *et al.* 1977). However, the evidence is not clear-cut. Poor singing often emerges in association with linguistic deficits resulting from left-hemisphere damage (e.g. Benton 1977). Furthermore, bilateral hemispheric involvement in sung performance is substantiated by evidence that lesions in either of the two hemispheres impair sung performance (Kinsella *et al.* 1988; Prior *et al.* 1990), that both right- and left-hemisphere anesthetization interfere with singing (Borchgrevink 1980; Zatorre 1984), and that singing without words do no elicit any lip-opening asymmetry, as a measure of laterality (e.g., Cadalbert *et al.* 1994; Hough *et al.* 1994). Finally, clinical observations that motor aphasics can articulate words better when singing than when speaking were not fully corroborated by recent systematic studies (Hébert *et al.* 2003; Racette *et al.* 2006). In sum, it is likely that singing is characterized by less strict lateralization than is speech.

## Conclusion

Like any cognitive faculty, music is multifaceted and the identification of the neural basis of any complex faculty must proceed, hand in hand, with an elucidation of its cognitive architecture. The past decade has seen an evolution in the theoretical models of musical processing, allowing the development of theoretically motivated instruments for the systematic evaluation of musical disorders. Such developments have allowed reports of musical disorders to evolve from historical anecdotes to systematic, verifiable accounts which can play a critical role in contributing to our understanding of the cognitive neuroscience of music.

## Acknowledgements

L.S, T.D.G and J.W were supported by the Wellcome Trust. L.S. was additionally supported by the Economic and Social Research Council and J.W. received support from a EC grant to the APOPIS consortium. SDB received an integration grant from the European Commission.

## References


Ackermann H, Wildgruber D and Riecker A (2006). Singing in the (b)rain: cerebral correlates. In E Altenmüller, M Wiesendanger, J Kesselring, eds, *Music, motor control and the brain*, 205–222. Oxford University Press, Oxford.

Alcock KJ, Wade D, Anslow P and Passingham RE (2000). Pitch and timing abilities in adult left-hemisphere dysphasic and right-hemisphere damaged subjects. *Brain and Language*, **75**(1), 47–65.

Assal G, Buttet J and Javet RC (1997). Musical aptitude in aphasics. *Revue Medicale de la Suisse Romande*, **97**(1), 5–12.

Ayotte J, Peretz I and Hyde K (2002). Congenital Amusia. A group study of adults afflicted with a music-specific disorder. *Brain*, **125**, 238–251.

Ayotte J, Peretz I, Rousseau I, Bard C, Bojanowski M (2000). Patterns of music agnosia associated with middle cerebral artery infarcts. *Brain*, **123**, 1926–1938.

Benton AL (1977) The amusias. In M Critchley and RA Henson, eds, *Music and the brain*, 378–397. William Heinemann, London.

Borchgrevink HM (1980) Cerebral lateralization of speech and singing after intracarotid amytal injection. In MT Sarno and O Hook, eds, *Aphasia. Assessment and Treatment*, 186–191. Stockholm: Almqvist and Wiksell

Botez MI and Wertheim N (1959) Expressive aphasia and amusia following right frontal lesion in a right-handed man. *Brain*, **82**, 186–202.

Bradshaw E and McHenry MA (2005) Pitch discrimination and pitch-matching abilities of adults who sing inaccurately. *Journal of Voice*, **14**(3), 431–439.

Brust JC (1980) Music and language: musical alexia and agraphia. *Brain*, **103**, 367–392.

Cadalbert A, Landis T, Regard M and Graves RE (1994) Singing with and without words: hemispheric asymmetries in motor control. *Journal of Clinical and Experimental Neuropsychology*, **16**(5), 664–670.

Confavreux C, Croisile B, Garassus P, Aimard G and Trillet M (1992). Progressive amusia and aprosody. *Archives of Neurology*, **49**, 971–6.

Critchley M and Henson RA (eds) (1977) *Music and the brain. Studies in the neurology of music*. Heinemann, Oxford.

Cuddy LL, Balkwill L-L, Peretz I and Holden RR (2005). Neuropsychological assessment of musical difficulties: a study of 'tone deafness' among university students. Annals of the New York Academy of Sciences, 1060, 311–324.

Dalla Bella, S, Giguère J-F and Peretz I (2007) Singing proficiency in the general population. *Journal of The Acoustical Society of America*, **121**, 1192–1189.

Di Pietro M, Laganaro M, Leemann B and Schnider A (2004). Receptive amusia: temporal auditory processing deficit in a professional musician following a left temporo-parietal lesion. *Neuropsychologia*, **42**, 868–877.

Dowling WJ and Harwood DL (1985). Music and cognition. Academic Press, London.

Epstein CM, Meador KJ, Loring DW, Wright RJ, Weissmann JD, Sheppard S, Lah JJ, Puhalovich F, Gaitan L and Davey KR (1999) Localization and characterization of speech arrest during transcranial magnetic stimulation. *Clinical Neurophysiology*, **110**(6), 1073–1079.

Eustache F, Lechevalier B, Viader F and Lambert J (1990). Identification and discrimination disorders in auditory perception: a report on two cases. *Neuropsychologia*, **28**, 257–270.

Foxton JM, Dean JL, Gee R, Peretz I and Griffiths TD (2004) Characterization of deficits in pitch perception underlying tone deafness. *Brain*, **127**, 801–110.

Foxton JM, Nandy RK and Griffiths TD (2006). Rhythm deficits in tone deafness. *Brain and Cognition*, **62**(1), 24–29.

Fries W and Swihart AA (1990). Disturbance of rhythm sense following right hemisphere damage. *Neuropsychologia*, **28**, 1317–23.

Fujii T, Fukatsu R, Watabe S, Ohnuma A, Teramura K, Kimura I *et al.* (1990) Auditory sound agnosia without aphasia following a right temporal lobe lesion. *Cortex*, **26**, 263–268.

Gordon HW and Bogen JE (1974) Hemispheric lateralization of singing after intracarotid sodium amylobarbitone. *Journal of Neurology, Neurosurgery, and Psychiatry*, **37**, 727–738.

Gordon RL, Racette A and Schön D (2006) Sensory-motor networks in singing and speaking: a comparative approach. In E Altenmüller, M Wiesendanger and J Kesselring, eds, *Music, motor control and the brain*, 223–238. Oxford University Press, Oxford.

Gosselin N, Peretz I, Noulhiane M, Hasboun D, Beckett C, Baulac M *et al.* (2005) Impaired recognition of scary music following unilateral temporal lobe excision. *Brain*, **128**, 628–40.

Grant-Allen (1878). Note-deafness. *Mind*, **10**, 157–167.

Griffiths TD, McDonald C, Kumar S, Deutsch D, Chinnery P and Stewart L (2007). Could a congenital disorder of musical perception ever be explained by a single gene? International Workshop on the Biology and Genetics of Music, Bologna, Italy.

Griffiths TD, Rees A, Witton C, Cross PM, Shakir RA and Green GG (1997). Spatial and temporal auditory processing deficits following right hemisphere infarction. A psychophysical study. *Brain*, **120**, 785–794.

Griffiths TD, Warren JD and Jennings AR (2006). Dystimbria: a distinct musical syndome? Presented at the International Conference for Musical Perception and Cognition.

Griffiths TD, Warren JD, Dean JL and Howard D (2004). When the feeling's gone: a selective loss of musical emotion. *Journal of Neurology, Neurosurgery and Psychiatry*, **75**, 344–345.

Grossman M, Shapiro BE and Gardner H (1981). Dissociable musical processing strategies after localized brain damage. *Neuropsychologia*, **19**, 425–433.

Habib M, Daquin G, Milandre L, Royere ML, Rey M, Lanteri A *et al.* (1995) Mutism and auditory agnosia due to bilateral insular damage—role of the insula in human communication. *Neuropsychologia*, **33**, 327–339.

Hattiangadi N, Pillion JP, Slomine B, Christensen J, Trovato MK and Speedie LJ (2005). Characteristics of auditory agnosia in a child with severe traumatic brain injury: a case report. *Brain and Language*, **92**, 12–25.

Hébert S, Racette A, Gagnon L and Peretz I (2003). Revisiting the dissociation between singing and speaking in expressive aphasia. *Brain*, **126**(8), 1838–1850.

Hofman S, Klein C and Arlazoroff A (1993). Common hemisphericity of language and music in a musician. A case report. *Journal of Communication Disorders*, **26**, 73–82.

Hough MS, Daniel HJ, Snow MA, O'Brien KF and Hume WG (1994) Gender differences in laterality patterns for speaking and singing. *Neuropsychologia*, **32**, 1067–1078.

Hyde KL and Peretz I (2003) Out-of-pitch but still in-time. An auditory psychophysical study in congenital amusic adults. *Annals of the New York Academy of Sciences*, **999**, 173–176.

Hyde KL, Lerch JP, Zatorre RJ, Griffiths TD, Evans AC and Peretz I (2007). Cortical thickness in congenital amusia: when less is better than more. *Journal of Neuroscience* **27**(47), 13028–13032.

Hyde KL, Zatorre RJ, Griffiths TD, Lerch JP and Peretz I (2005). Morphometry of the amusic brain: a two-site study. *Brain*, **129**, 2562–2570.

Jeffries KJ, Fritz JB and Braun AR (2003) Words in melody: an $H_2{}^{15}O$ PET study of brain activation during singing and speaking. *Neuroreport*, **15**(5), 749–754.

Johannes S, Jobges ME, Dengler R and Munte TF (1998). Cortical auditory disorders: a case of non-verbal disturbances assessed with event-related brain potentials. *Behavioural Neurology*, **11**, 55–73.

Johkura K, Matsumoto S, Hasegawa O and Kuroiwa Y (1998). Defective auditory recognition after small hemorrhage in the inferior colliculi. *Journal of Neurological Sciences*, **161**, 91–96.

Johnsrude IS, Penhune VB and Zatorre RJ (2000). Functional specificity in right human auditory cortex for perceiving pitch direction. *Brain*, **123**, 155–163.

Kester DB, Saykin AJ, Sperling MR, O'Connor MJ, Robinson LJ and Gur RC (1991). Acute effect of anterior temporal lobectomy on musical processing. *Neuropsychologia*, **29**, 703–708.

Kinsella G, Prior MR and Murray G (1988) Singing ability after right and left sided brain damage. A research note. *Cortex*, **24**(1), 165–169.

Kohlmetz C, Muller SV, Nager W, Munte TF and Altenmuller E (2003). Selective loss of timbre perception for keyboard and percussion instruments following a right temporal lesion. *Neurocase*, **9**, 86–93.

Lechevalier B, Rossa Y, Eustache F, Schupp C, Boner L and Bazin C (1984). Un cas de surdit, corticale, pargnant en partie la musique. *Revue Neurologique*, **140**, 190–201.

Levin HS and Rose JE. (1979)Alexia without agraphia in a musician after transcallosal removal of a left intraventricular meningioma. *Neurosurgery*, **4**, 168–174.

Liegeois-Chauvel C, Peretz I, Babai M, Laguittin V and Chauvel P (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain*, **121**, 1853–1867.

Lo YL and Fook-Chong S (2004). Ipsilateral and controlateral motor inhibitory control in musical and vocalization tasks. *Experimental Brain Research*, **159**(2), 258–262.

Mandell J, Schulze K and Schlaug G (2007) Congenital amusia: an auditory-feedback disorder? *Restorative Neurology and Neuroscience*, **25**, 323–334.

Marin OSM and Perry DW (1999) Neurological aspects of music perception and performance. In Deutsch D, ed. *Psychology of music*, 653–724. Academic Press, San Diego, CA.

Mavlov L (1980). Amusia due to rhythm agnosia in a musician with left hemisphere damage: a non-auditory supramodal defect. *Cortex*, **16**, 331–338.

Mazzoni (1993). A case of music imperception. *Journal of Neurology, Neurosurgery and Psychiatry*, 56, 322.

Mazzucchi A, Marchini C, Budai R and Parma M (1982). A case of receptive amusia with prominent timbre perception defect. *Journal of Neurology, Neurosurgery and Psychiatry*, **45**, 644–647.

McDonald C and Stewart L (2008) Uses and functions of music in congenital amusia. *Music Perception*, **25**(4), 345–355.

McFarland HR and Fortin D (1982) Amusia due to right temporoparietal infarct. *Archives of Neurology*, **39**(11), 725–727.

Mendez MF and Geehan GR (1988). Cortical auditory disorders:clinical and psychoacoustic features. *Journal of Neurology, Neurosurgery and Psychiatry*, **51**, 1–9.

Milner B (1962). Laterality effects in audition. In VB Mountcastle, ed., *Interhemispheric relations and cerebral dominance*, 177–195. Johns Hopkins University Press, Baltimore, MD.

Murayama J, Kashiwagi T, Kashiwagi A and Mimura M (2004). Impaired pitch production and preserved rhythm production in a right brain-damaged patient with amusia. *Brain and Cognition*, **56**, 36–42.

Patel AD, Peretz I, Tramo M and Labreque R (1998). Processing prosodic and musical patterns: a neuropsychological investigation. *Brain and Language*, **61**, 123–144.

Peretz I (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain*, **113**, 1185–1205.

Peretz I (1996). Can we lose memory for music? A case of music agnosia in a non-musician. *Journal of Cognitive Neuroscience*, **8**, 481–496.

Peretz I, Coltheart M (2003). Modularity of music processing. *Nature Neuroscience*, **6**, 688–691.

Peretz I and Gagnon L (1999). Dissociation between recognition and emotion for melodies. *Neurocase*, **5**, 21–30.

Peretz I, Belleville S and Fontaine F (1997). Dissociations entre musique et langage apres atteine cerebrale; un nouveau cas d'amusie sans aphasie. *Canadian Journal of Experimental Psychology*, **51**, 354–368.

Peretz I, Blood AJ, Penhune V and Zatorre R (2001). Cortical deafness to dissonance. *Brain*, **124**, 928–940.

Peretz I, Brattico E and Tervaniemi M (2005). Abnormal electrical brain responses to pitch in congenital amusia. *Annals of Neurology*, **58**, 478–482.

Peretz I, Gagnon L and Bouchard B (1998). Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, **68**, 111–141.

Peretz I, Champod A-S and Hyde KL (2003). Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences*, **999**, 58.

Peretz I, Kolinsky R, Tramo M, Labrecque R, Hublet C, Demeurize G *et al.* (1994) Functional dissociations

following bilateral lesions of auditory cortex. *Brain*, **117**, 1283–1301.

Piccirilli M, Sciarma T and Luzzi S (2000). Modularity of music: evidence from a case of pure amusia. *Journal of Neurology, Neurosurgery and Psychiatry*, **69**, 541–545.

Prior M, Kinsella G and Giese J (1990). Assessment of musical processing in brain-damaged patients: implications for laterality of music. *Journal of Clinical and Experimental Neuropsychology*, **12**, 301–312.

Racette A, Bard, C and Peretz I (2006). Making non-fluent aphasics speak: sing along! *Brain*, **129**(10), 2571–2584.

Riecker A, Ackermann H, Wildgruber D, Dogil G and Grodd W (2000). Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *Neuroreport*, **11**, 1997–2000.

Robin DA, Tranel D and Damasio H (1990). Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. *Brain and Language*, **39**, 539–555.

Samson S and Zatorre RJ (1988). Melodic and harmonic discrimination following unilateral cerebral excision. *Brain and Cognition*, **7**, 348–360.

Samson S and Zatorre RJ (1991). Recognition memory for text and melody of songs after unilateral temporal lobe lesion: evidence for dual encoding. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **17**, 793–804.

Samson S and Zatorre RJ (1992). Learning and retention of melodic and verbal information after unilateral temporal lobectomy. *Neuropsychologia*, **30**, 815–826.

Samson S and Zatorre RJ (1994). Contribution of the right temporal lobe to musical timbre discrimination. *Neuropsychologia*, **32**, 231–240.

Samson S, Ehrle N and Baulac M (2001). Cerebral substrates for musical temporal processes. *Annals of the New York Academy of Sciences*, **930**, 166–178.

Samson S, Zatorre RJ and Ramsay JO (2002). Deficits of musical timbre perception after unilateral temporal-lobe lesion revealed with multidimensional scaling. *Brain*, **125**, 511–523.

Satoh M, Takeda K, Murakami Y, Onouchi K, Inoue K and Kuzuhara S (2005). A case of amusia caused by the infarction of anterior portion of bilateral temporal lobes. *Cortex*, **41**, 77–83.

Schön D, Lorber B, Spacal M and Semenza C (2004). A selective deficit in the production of exact musical intervals following right hemisphere damage. *Cognitive Neuropsychology*, **21**, 773–785.

Schuppert M, Munte TF, Wieringa BM and Altenmuller E (2000). Receptive amusia: evidence for cross-hemispheric neural networks underlying music processing strategies. *Brain*, **123**, 546–559.

Shankweiler D (1966). Effects of temporal lobe damage on the perception of dichotically presented melodies. *Journal of Comparative Physiology and Psychology*, **62**, 115–119.

Shapiro BE, Grossman M and Gardner H (1981). Selective musical processing deficits in brain damaged populations. *Neuropsychologia*, **19**, 161–169.

Sidtis JJ and Volpe BT (1988). Selective loss of complex-pitch or speech discrimination after unilateral lesion. *Brain and Language*, **34**, 235–245.

Sloboda JA, Wise KJ and Peretz I (2005). Quantifying tone deafness in the general population. *Annals of the New York Academy of Sciences*, **1060**, 255–261.

Spreen O, Benton AL and Fincham RW (1965). Auditory agnosia without aphasia. *Archives of Neurology*, **13**, 84–92.

Stewart L, Overath T, Warren JD, Foxton JM and Griffiths TD (2008) fMRI Evidence for a cortical hierarchy of pitch pattern processing. *PLoS ONE*, **3**(1), e1470. doi:10.1371/journal.pone.0001470.

Stewart L, Von Kriegstein K, Warren JD and Griffiths TD (2006) Disorders of musical listening. *Brain*, **129**, 2533–2553.

Stewart L, Walsh V, Frith U and Rothwell J (2001). Transcranial magnetic stimulation produces speech arrest but not song arrest. *Annals of the New York Academy of Sciences*, **930**, 433–435.

Tanaka Y, Yamadori A and Mori E (1987). Pure word deafness following bilateral lesions. *Brain*, **110**, 381–403.

Terao Y, Mizuno T, Shindoh M, Sakurai Y, Ugawa Y, Kobayashi S *et al.* (2005) Vocal amusia in a professional tango singer due to a right superior temporal cortex infarction. *Neuropsychologia*, **44**, 479–88.

Tramo MJ and Bharucha JJ (1991). Musical priming by the right hemisphere post-callosotomy. Neuropsychologia, 29, 313–325.

Tramo MJ, Bharucha JJ and Musiek FE (1990). Music perception and cognition following bilateral lesions of auditory cortex. *Journal of Cognitive Neuroscience*, **2**, 195–212.

Tramo MJ, Shah GD and Braida LD (2002). Functional role of auditory cortex in frequency processing and pitch perception. *Journal of Neurophysiology*, **87**, 122–139.

Ustvedt HI (1937). Ueber die Untersuchung der musikalischen Funktionen bei Patienten mit Gehirnleiden besonders bei Patienten mit Aphasie. *Acta Medica Scandinavica*, **86**, 1–186.

Warrier CM and Zatorre RJ (2004). Right temporal cortex is critical for utilization of melodic contextual cues in a pitch constancy task. *Brain*, **127**, 1616–1625.

Wilson SJ, Pressing JL and Wales RJ (2002). Modelling rhythmic function in a musician post-stroke. *Neuropsychologia*, **40**, 1494–1505.

Yamadori A, Osumi Y, Masuhara S and Okubo M (1977) Preservation of singing in Broca's aphasia. *Journal of Neurology, Neurosurgery and Psychiatry*, **40**(3), 221–224.

Zatorre RJ (1984). Musical perception and cerebral function: A critical review. *Music Perception*, **2**, 196–221.

Zatorre RJ (1985). Discrimination and recognition of tonal melodies after unilateral cerebral excisions. *Neuropsychologia*, **23**, 31–41.

Zatorre RJ (1988). Pitch perception of complex tones and human temporal-lobe function. *Journal of the Acoustical Society of America*, **84**, 566–572.

Zatorre RJ and Halpern AR (1993). Effect of unilateral temporal lobe excision on perception and imagery of songs. *Neuropsychologia*, **31**, 221–232.

Zatorre RJ and Samson S (1991). Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain*, **114**, 2403–2417.

CHAPTER 18

# Music, musicians, and brain plasticity

Gottfried Schlaug

## Musicians and non-musicians: their brains differ

Playing a musical instrument is an intense, multisensory and motor experience that is typically initiated at an early age and requires the acquisition and practice of a wide range of skills over the course of a musician's lifetime. Considering the early age at which they commence training and the continuous practice that maintaining such exacting skills demands over time, musicians make an ideal human model for studying the brain effects of acquiring highly specialized sensorimotor (Elbert *et al.* 1995; Amunts *et al.* 1997; Gaser and Schlaug 2003; Schlaug 2001; Schlaug *et al.* 2005; Hund-Georgiadis and von Cramon 1999; Meister *et al.* 2005; Bengtsson *et al.* 2005); auditory (Besson *et al.* 1994; Gaab and Schlaug 2003; Keenan *et al.* 2001; Pantev *et al.* 1998, 2001; Schlaug *et al.* 1995a, b; Schneider *et al.* 2002; Zatorre *et al.* 1998); auditory–spatial and visual–motor (Sergent *et al.* 1993; Münte *et al.* 2001); and auditory–motor (Lotze *et al.* 2003; Bangert *et al.* 2006; Zatorre *et al.* 2007) skills.

### Strongest effects are found in primary domain regions

Several studies have found evidence for structural differences in brain regions that are directly involved with some of these skills, such as the primary auditory cortex and auditory association cortex (Schlaug *et al.* 1995a; Zatorre *et al.* 1998; Schneider *et al.* 2002; Gaser and Schlaug 2003) most likely related to important skills of fine spectral and temporal discrimination and pitch categorization. Further examples include structural differences in primary sensorimotor cortex and closely related premotor and supplementary motor regions, which are essential for preparation, execution, and control of precise, independent finger movements (Elbert *et al.* 1995; Amunts *et al.* 1997; Gaser and Schlaug 2003; Bangert and Schlaug 2006), and in the midsagittal side of the corpus callosum (Schlaug *et al.* 1995b; Ozturk *et al.* 2002; Lee *et al.* 2003), which is likely related to interactions between both hemispheres necessary for independent movements of right and left hand. These structural between-group (musicians vs non-musicians) differences mirror structural changes seen in primary motor and cerebellar cortices of experimental animal models that have been subjected to long-term learning of complex motor skills over several months (Anderson *et al.* 2002). Micro-structural changes included increases in the number of synapses and glial cells, increased density of capillaries in primary motor cortex and the cerebellum, and new brain cells in the hippocampus after long-term motor training in adult rats (Black *et al.* 1990; Isaacs *et al.* 1992; Anderson *et al.* 1994; Kleim *et al.* 1996; Kempermann *et al.* 1997; van Praag *et al.* 1999). The sum of these micro-structural changes could amount to structural differences that are detectable on a macro-structural level (see Anderson *et al.* 2002; Bangert and Schlaug 2006).

Most cross-sectional studies comparing the brains of adult musicians with those of matched

non-musicians, have used either age of commencement of musical training or intensity/duration of practice throughout a musician's career as predictors of regional differences suggesting that the longer and the more intensely musicians practiced, the more pronounced the between-group differences were (Elbert *et al.* 1995; Amunts *et al.* 1997; Schlaug 2001; Sluming *et al.* 2002; Schneider *et al.* 2002; Gaser and Schlaug 2003; Bangert and Schlaug 2006). Although these correlations suggest a causal relationship, it is still not a substitute for a longitudinal study proving that the between-group differences are due to training and are not related to genetic causes.

## Relating structural differences to functional imaging correlates

Results of functional imaging studies are not always congruent with structural brain imaging studies. Although several studies have reported that musicians have more grey matter and greater cortical thickness in primary and auditory cortex (Heschl's gyrus) (Schneider *et al.* 2002; Gaser and Schlaug 2003), functional studies have actually shown stronger responses in secondary and tertiary auditory brain regions comparing musicians with non-musicians. Interestingly, non-musicians seem to show stronger activation of primary auditory regions (Besson *et al.* 1994; Trainor *et al.* 1999; Shahin *et al.* 2003; 2004; Gaab and Schlaug 2003; Bosnyak *et al.* 2004). Similar results were found in the motor system. Most functional imaging studies comparing musicians with non-musicians have shown that musicians have either less, or more focused activation in the primary motor cortex than non-musicians and more variable levels of activation in motor association regions such as the premotor and supplementary motor areas (Hund-Georgiadis and von Cramon 1999; Jancke *et al.* 2000; Krings *et al.* 2000; Haslinger *et al.* 2004; Meister *et al.* 2005). This contradiction between the pronounced between-group structural differences and within-group structural changes in primary auditory and motor cortices on one hand and the lower degree of functional activation of these regions on the other hand, has not been fully understood. Some argue that as a system becomes more plastic and expands its representation, it may work more efficiently with fewer neurons and thus has less of a metabolic demand or less of a blood flow demand than a naive subject (a non-instrumentalist) might have while performing a behaviourally matched complex motor sequence (Hund-Georgiadis and Von Cramon 1999; Krings *et al.* 2000; Jancke *et al.* 2000; Koeneke *et al.* 2004; Meister *et al.* 2005).

## Structural differences are also found in an extended network of multimodal sensorimotor integration regions

Structural brain differences comparing musicians with matched non-musician controls have also been found beyond primary domains in regions such as in the inferior frontal gyrus, the superior parietal lobe including the intraparietal sulcus, the inferior lateral temporal lobe, and the cerebellum (Sluming *et al.* 2002; Hutchinson *et al.* 2003; Gaser and Schlaug 2003), yet the functional role of these differences/changes in extra-primary domain regions is not fully understood. It is possible that parietal regions might play a role in multisensory encoding such as in the integration of auditory information with other sensory (e.g. visual) information. The prominent activation of the superior parietal lobule in musical sight-reading (Sergent *et al.* 1992) may be reflective of the sensorimotor transformation of a visuospatial representation (musical notation) to a pattern of intended movements and position of fingers on the keyboard. Stewart and colleagues (2003) were able to show that the superior parietal lobe (SPL) showed activity changes when comparing functional images obtained while musically naive subjects played melodies from notation before and after a 15-week training period that involved reading music and playing the keyboard. This is further support for the notion that the SPL is involved in the sensorimotor translation of visual–spatial representation into a motor programme.

Strong structural differences in the frontal lobe, particularly in the left inferior frontal gyrus (IFG), have been found in two independent studies (Sluming *et al.* 2002; Gaser and Schlaug 2003). The left IFG has also been found to be activated in various functional imaging

studies using musical tasks. The interest in the inferior frontal lobule has been sparked further in recent years, as more studies have found this region to be a critical link in the auditory–motor network. However, the frontal lobe also plays a more general role in music-making which can range from integration of individual auditory events into larger units, making predictions about and recognizing alterations in sequential auditory–perceptual events, and in mapping actions to their associated sounds (Koelsch *et al.* 2002; Levitin and Menon, 2003; Lotze *et al.* 2003; Bangert and Altenmueller 2003; Bangert *et al.* 2006; Lahav *et al.* 2007; Zatorre *et al.* 2007).

It is possible that the specific and continuous engagement of this multimodal sensorimotor integration network and the induced changes in this network across a musician's career may be the neural basis for some of the cognitive enhancements attributed to musical training. A form of cross-modal plasticity, induced by musical training, could affect extra-musical behavioural and cognitive tasks when those tasks engage regions that are altered through long-term music-making. Evidence for this hypothesis does already exist in the form of finding parietal activation with mathematical tasks (Chochon *et al.* 1999) that coincides with regions of activation induced by music making and musical sight-reading in particular (Sergent *et al.* 1992; Stewart *et al.* 2003), or the congruence of activations in the inferior parietal cortex with musical memory task and language tasks, both utilizing network components of the phonological loop (Gaab *et al.* 2003, 2006). Other examples include IFG activation while performing sequential, non-musical mental operations (Sluming *et al.* 2007) or while performing various speech and language tasks but also while performing musical tasks of auditory-motor mapping, integration of individual

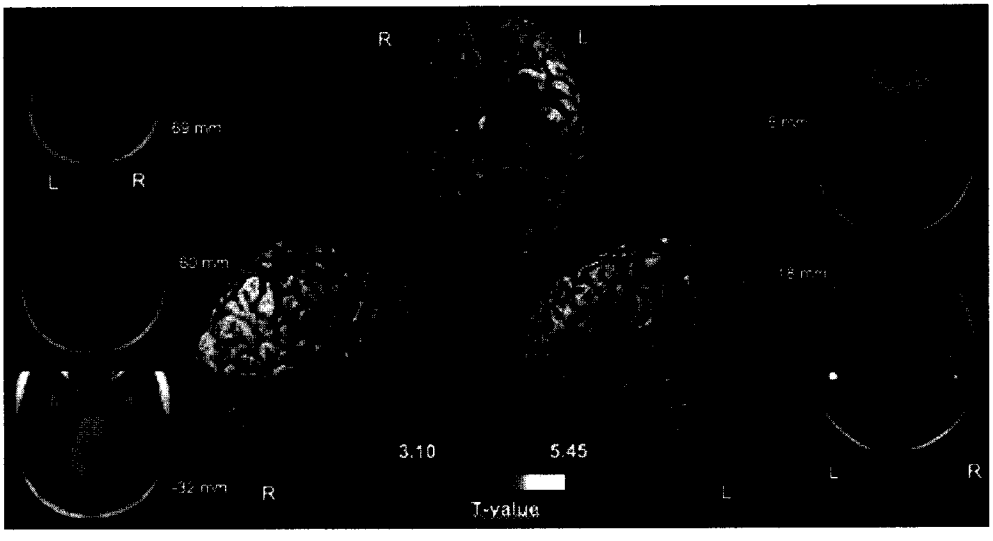



**Fig. 18.1** A voxel-based morphometric study showing grey matter volume differences between three groups. Only voxels exhibiting a significant positive correlation between musician status (professional musician > amateur musician > non-musician) and increased grey matter volume are shown here, and those have been overlaid on the rendered cortical surface of a single subject. The distinction between professional and amateur musicians in this study was based on whether or not the keyboard player's main profession was being a musician (e.g., music teacher, performer). There was also a clear separation between both groups in terms of average practice intensity across their careers, with professional keyboard players having approximately double the amount of practice time than the group of amateur musicians. (This figure was adapted from Figure 1 in Gaser and Schlaug [2003]). See colour plate section.

studies using musical tasks. The interest in the inferior frontal lobule has been sparked further in recent years, as more studies have found this region to be a critical link in the auditory–motor network. However, the frontal lobe also plays a more general role in music-making which can range from integration of individual auditory events into larger units, making predictions about and recognizing alterations in sequential auditory–perceptual events, and in mapping actions to their associated sounds (Koelsch *et al.* 2002; Levitin and Menon, 2003; Lotze *et al.* 2003; Bangert and Altenmueller 2003; Bangert *et al.* 2006; Lahav *et al.* 2007; Zatorre *et al.* 2007).

It is possible that the specific and continuous engagement of this multimodal sensorimotor integration network and the induced changes in this network across a musician's career may be the neural basis for some of the cognitive enhancements attributed to musical training. A form of cross-modal plasticity, induced by musical training, could affect extra-musical behavioural and cognitive tasks when those tasks engage regions that are altered through long-term music-making. Evidence for this hypothesis does already exist in the form of finding parietal activation with mathematical tasks (Chochon *et al.* 1999) that coincides with regions of activation induced by music making and musical sight-reading in particular (Sergent *et al.* 1992; Stewart *et al.* 2003), or the congruence of activations in the inferior parietal cortex with musical memory task and language tasks, both utilizing network components of the phonological loop (Gaab *et al.* 2003, 2006). Other examples include IFG activation while performing sequential, non-musical mental operations (Sluming *et al.* 2007) or while performing various speech and language tasks but also while performing musical tasks of auditory-motor mapping, integration of individual

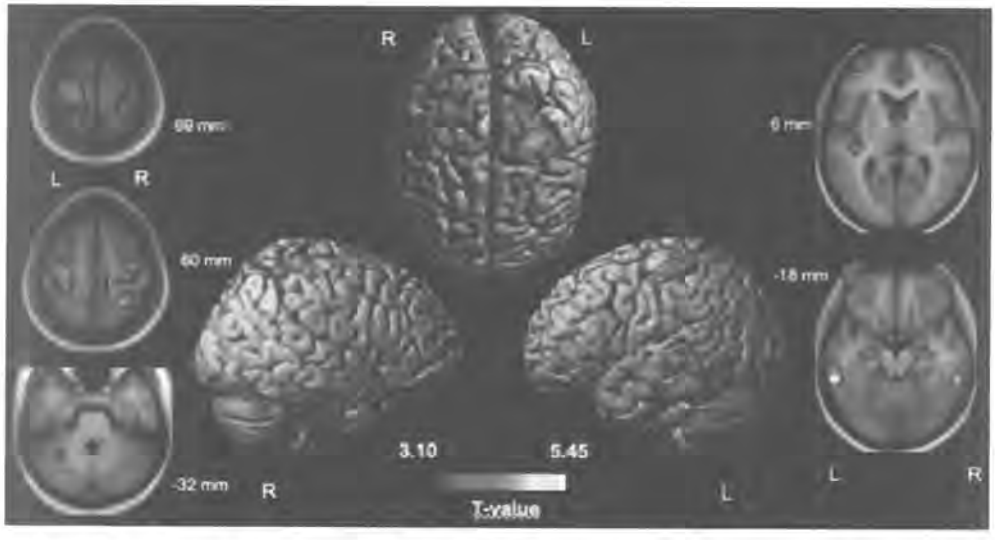



**Fig. 18.1** A voxel-based morphometric study showing grey matter volume differences between three groups. Only voxels exhibiting a significant positive correlation between musician status (professional musician > amateur musician > non-musician) and increased grey matter volume are shown here, and those have been overlaid on the rendered cortical surface of a single subject. The distinction between professional and amateur musicians in this study was based on whether or not the keyboard player's main profession was being a musician (e.g., music teacher, performer). There was also a clear separation between both groups in terms of average practice intensity across their careers, with professional keyboard players having approximately double the amount of practice time than the group of amateur musicians. (This figure was adapted from Figure 1 in Gaser and Schlaug [2003]). See colour plate section.

auditory events into larger units, or making predictions about and recognizing alterations in structured musical sequences (Koelsch *et al.* 2002, 2005; Patel 2003; Tettamanti and Weniger 2006; Lahav *et al.* 2007).

## Within-musician differences—stronger evidence for a causal effect

Further support for brain plasticity as an explanation for the anatomical differences between adult musicians and non-musicians comes from comparisons of musician groups that play different instruments. Pantev and colleagues (2001) found more pronounced cortical responses to trumpet and string tones in the respective players of those instruments, suggesting that perceptual markers of brain activity differ as a function of the instrument played and constantly heard (Pantev *et al.* 2001). In the absence of longitudinal data examining causal relationships, the search for structural brain differences within a heterogeneous group of musicians is a valuable alternative to support the notion of structural brain adaptation. For example, string-players strongly differ from keyboard players in the specific fine motor (hand/finger) skills required for playing their instruments. While string players need highly developed fine motor skills of their left hand and more gross motor skills of the right, keyboard players require highly developed fine motor skills of both hands but with greater facilities in their right hand since the left hand frequently functions as more of an accompaniment in keyboard music. When comparing string and keyboard players, Bangert and Schlaug (2006) found gross anatomical differences in the precentral gyrus (Figure 18.2). Further, the majority of the adult keyboard players had an elaborate configuration (known as 'Omega Sign') of the precentral gyrus on the left more than in the right hemisphere, while most of the adult string players had this 'atypicality' only on the right. Although a genetic cause for the Omega Sign atypicality in this strongly right-handed group of musicians (i.e., a differential hemispheric expression favouring string playing in those who have a prominent

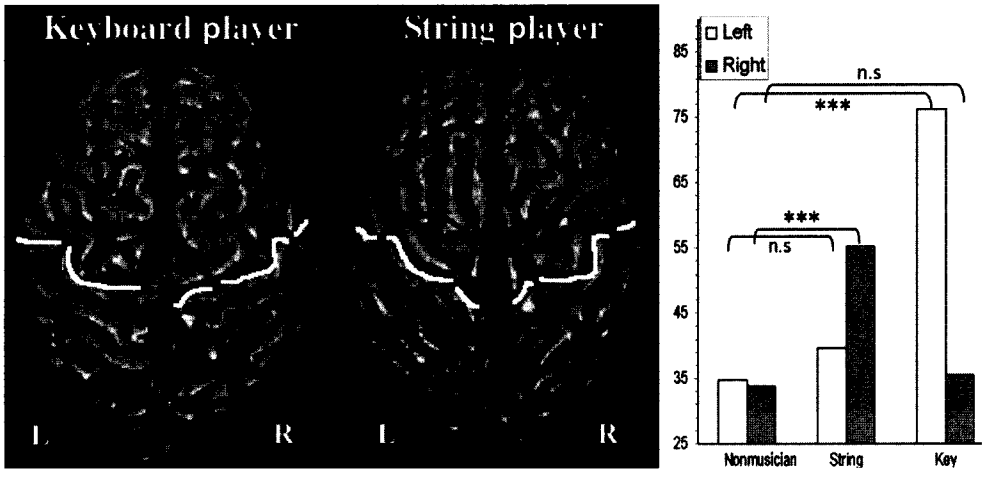



**Fig. 18.2** Brain surface renderings of one typical keyboard and one typical string player. The central sulcus is marked with a white line. The portion of the precentral gyrus containing the gyral configuration similar to that of the inverted Greek letter 'Omega' is found within the red circles. In the two examples, a prominent omega sign can be seen on the left more than the right in the keyboard player, and only on the right in the string player. Across the entire group, we found significantly more prominent 'Omega signs' in the right hemisphere of string players compared to the right hemisphere of non-musicians and in the left hemisphere of keyboard players compared to the left hemispheres of both non-musicians and string players. See colour plate section.

Omega Sign in the right hemisphere and keyboard playing in those who have a prominent Omega Sign in their left hemisphere) cannot be completely ruled out, a correlation between the degree of the Omega Sign's prominence and the musicians' age at commencement of musical training, as well as with the cumulative amount of practice time accrued across the years strongly argues for a structural plasticity mechanism.

We also tested whether or not morphological differences between musicians and non-musicians (more elaborate motor cortex and larger midsagittal corpus callosum) have neurophysiological correlates. The larger midsagittal callosal size might reflect a more elaborate interaction between both hemispheres facilitating independent finger movements of both hands. Using transcranial magnetic stimulation (TMS), Ridding *et al.* (2000) found a reduced transcallosal inhibition and Nordstrom and Butler (2002) reported a reduced intracortical inhibition (ICI) and facilitation (ICF) in musicians. Although finding neurophysiological differences between musicians and non-musicians supports the notion that anatomical differences have neurophysiological correlates, these TMS results were counter-intuitive and contrary to what was expected. We used TMS to assess the interhemispheric interaction in a group of 21 right-handed adults (7 keyboard players, 7 string players and 7 non-musicians). We found that keyboard players have a stronger inhibition from the left to the right (LtoR) hemisphere while string players have a stronger inhibition from the right to the left (RtoL) hemisphere. The degree of inter-hemispheric inhibition (IHI) correlated with the size of the fourth and fifth segment (midbody) of the corpus callosum that contains fibres connecting primary sensorimotor cortex (Figure 18.3 c and d). Non-musicians showed similar IHI from either hemisphere. Since inhibitory interneurons in the primary motor cortex are known to mediate transcallosal inhibitory input to the homologous area in the opposite hemisphere, our results suggest that strengthening the activity of these inhibitory interneurons may be one mechanism by which different regions of the brain exert inhibitory control over those homologous regions in the opposite hemisphere. In addition, in within- and between-group comparisons, string players had significantly greater intracortical facilitation (ICF) in (1) their right hemisphere than their left, and (2) the right hemisphere of non-musicians. The stronger ICF in string players' right hemisphere (controlling the left hand) most likely points to neural processes providing excitatory input to the corticospinal neurons involved in precise, fine motor control of the left hand (Nair *et al.* 2006).

## Cognitive and brain changes when children learn to play a musical instrument

To determine whether the structural and functional differences seen in adult musicians reflect adaptations due to musical training during sensitive periods of brain development, or are instead markers of musical interest and/or aptitude that existed prior to training, it is necessary to examine children and/or adults before the onset of instrumental music training and compare them to a group of control subjects not planning to play and practice regularly on a musical instrument. For the past five years we have been conducting a longitudinal study of the effects of music training on brain development and cognition in young children (Overy *et al.* 2004; Norton *et al.* 2005; Koelsch *et al.* 2005; Schlaug *et al.* 2005). We have tested a large group of 5–7-year-old children prior to initiation of music lessons (baseline), approximately two-thirds of whom chose to take piano lessons and one third chose string lessons. We also tested a slightly smaller, untreated control group matched to the instrumental group in age, socio-economic status (SES), and verbal IQ. Each child underwent a battery of tests (for details see Norton *et al.* 2005). We found no pre-existing cognitive, auditory perceptual, motoric, or structural brain differences between the children that entered the instrumental group and those that entered the control group at baseline (Norton *et al.* 2005), thereby making it unlikely that children who choose to learn an instrument do so because they have an atypical brain. After an average follow-up interval of 15 months, we found significantly greater change scores in the instrumental group than in the control group in skills directly linked to

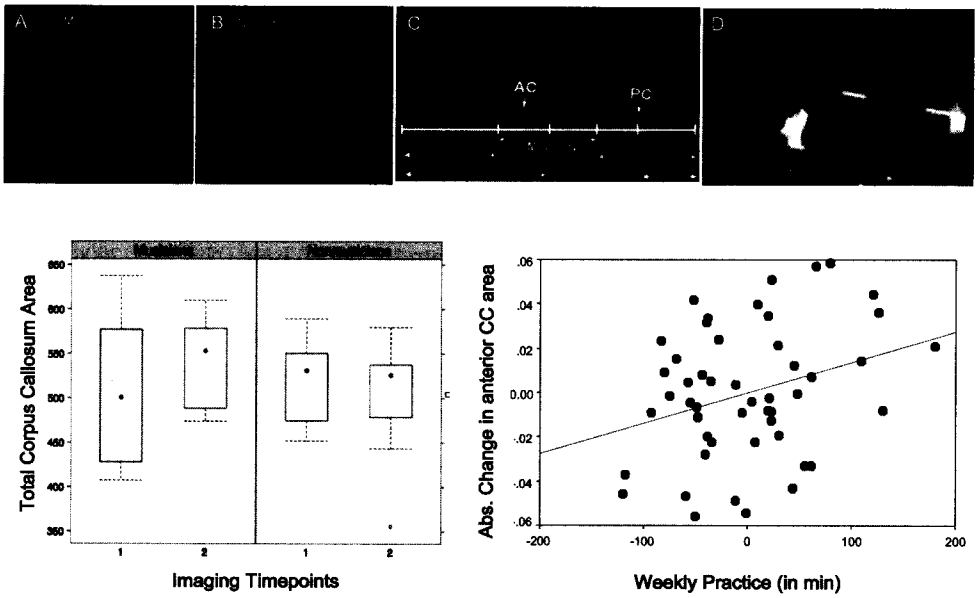



**Fig. 18.3** Corpus callosum differences in adults (musicians vs non-musicians) and changes over time in children. (A) and (B) Show the midsagittal slice of an adult musician (A) and non-musician (B) showing a difference in the size of the anterior and midbody of the corpus callosum. (C) and (D) Show the major subdivisions of the corpus callosum and locations of the interhemispheric fibers connecting the motor hand regions on the right and left hemisphere through the corpus callosum. (E) and (F): The mean (dot inside box plots) midsagittal corpus callosum size of the 5–7-year-old children at baseline (time point 1) and after 15 months of musical training (time point 2) shows slightly opposite trends with the instrumental group showing an increase in the midsagittal corpus callosum area. The anterior portion of the corpus callosum (segments 1–4, see Fig. 18.3c) showed a significant correlation with the amount of weekly practice ($p = 0.034$; partial $r^2 = 0.094$). See colour plate section.

instrumental music training: fine motor skills and auditory discrimination skills including both rhythmic and melodic discrimination. At this time point, no significant between-group differences were seen in non-primary domains such as verbal and visuospatial skills (Forgeard *et al.* submitted). However, our analyses did reveal trends for between-group differences in several of the extra-primary domains or transfer domains (e.g., verbal skills), suggesting that a longer period of observation might lead to significant between-group differences in these transfer domains. Children with higher intensity of practice had stronger trends in these outcome variables. Brain imaging data also support this trend. There was a strong trend for a greater increase in midsagittal corpus callosum size in the instrumental group than in the control group and there was a correlation between practice intensity and change in the anterior corpus callosum (Figure 18.3 e and f).

In addition to our longitudinal study of 5–7-year-old children learning to play a musical instrument, we also have preliminary data on a cross-sectional study of 9–11-year-old instrumentalists with an average of 4–5 years of music training and a group of non-musicians matched in age, handedness, and SES. Preliminary results from this cross-sectional study showed that the instrumentalists also performed significantly better in skills directly linked to instrumental music training: fine motor skills and auditory discrimination skills including both rhythmic and melodic discrimination. In addition, we

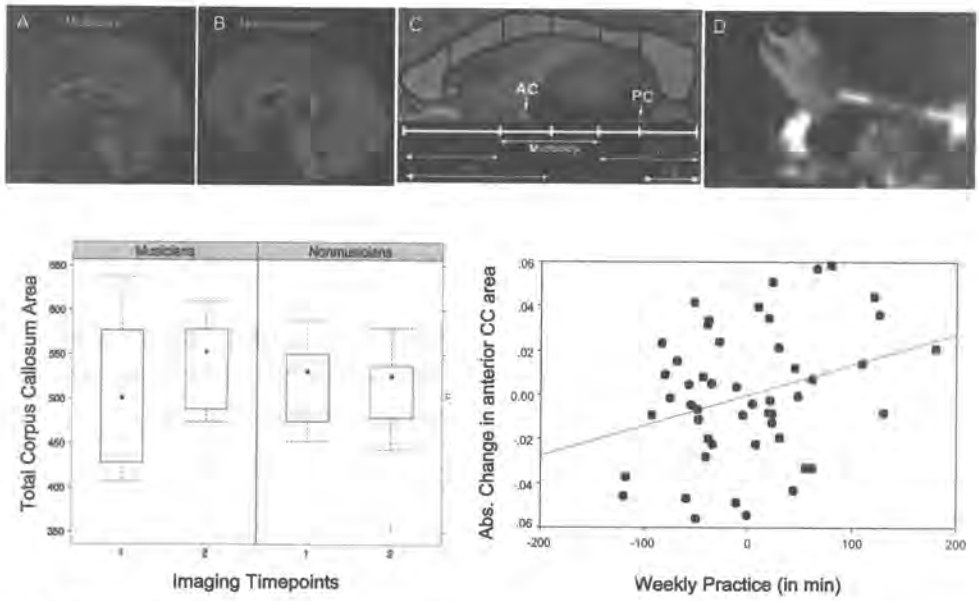



**Fig. 18.3** Corpus callosum differences in adults (musicians vs non-musicians) and changes over time in children. (A) and (B) Show the midsagittal slice of an adult musician (A) and non-musician (B) showing a difference in the size of the anterior and midbody of the corpus callosum. (C) and (D) Show the major subdivisions of the corpus callosum and locations of the interhemispheric fibers connecting the motor hand regions on the right and left hemisphere through the corpus callosum. (E) and (F): The mean (dot inside box plots) midsagittal corpus callosum size of the 5–7-year-old children at baseline (time point 1) and after 15 months of musical training (time point 2) shows slightly opposite trends with the instrumental group showing an increase in the midsagittal corpus callosum area. The anterior portion of the corpus callosum (segments 1–4, see Fig. 18.3c) showed a significant correlation with the amount of weekly practice ($p = 0.034$; partial $r^2 = 0.094$). See colour plate section.

instrumental music training: fine motor skills and auditory discrimination skills including both rhythmic and melodic discrimination. At this time point, no significant between-group differences were seen in non-primary domains such as verbal and visuospatial skills (Forgeard *et al.* submitted). However, our analyses did reveal trends for between-group differences in several of the extra-primary domains or transfer domains (e.g., verbal skills), suggesting that a longer period of observation might lead to significant between-group differences in these transfer domains. Children with higher intensity of practice had stronger trends in these outcome variables. Brain imaging data also support this trend. There was a strong trend for a greater increase in midsagittal corpus callosum size in the instrumental group than in the control group and there was a correlation between practice intensity and change in the anterior corpus callosum (Figure 18.3 e and f).

In addition to our longitudinal study of 5–7-year-old children learning to play a musical instrument, we also have preliminary data on a cross-sectional study of 9–11-year-old instrumentalists with an average of 4–5 years of music training and a group of non-musicians matched in age, handedness, and SES. Preliminary results from this cross-sectional study showed that the instrumentalists also performed significantly better in skills directly linked to instrumental music training: fine motor skills and auditory discrimination skills including both rhythmic and melodic discrimination. In addition, we

found between-group differences in other domains such as verbal skills and visual pattern matching skills with the instrumentalists performing better on these tests. Furthermore, the 9–11-year-old instrumental group had significantly more grey matter volume that was regionally pronounced in the sensorimotor cortex bilaterally. These between-group differences suggest that a longer duration of musical training and also a longer duration of observation in longitudinal studies together with a higher practice intensity makes it more likely to find significant within-group and between-group structural changes.

There is a widespread view that learning to play a musical instrument in childhood stimulates cognitive development and leads to the enhancement of skills in a variety of extra-musical areas which is commonly referred to as transfer (Bangerter and Heath 2004; Bruer 1999). The most commonly observed form of transfer occurs when there is a close resemblance between the training domain and the transfer domain (typically referred to as 'near transfer'), (e.g., fine motor skills that develop while learning to play a musical instrument lead to increased speed and accuracy in typing). While near transfer effects are relatively common, it is notoriously difficult to demonstrate far transfer, where the resemblance between training and transfer domains is much less obvious (e.g., learning to read and perform with precision from musical rhythm notation and understanding fractions in math). Evidence of far transfer from instrumental music training has previously been claimed in the areas of verbal, spatial, mathematical, and IQ performance (Rauscher *et al.* 1993, 1997; Rauscher and Zupan 2000; Chan *et al.* 1998; Vaughn 2000; Hetland 2000; Ho *et al.* 2003; Schellenberg 2004; Fauvel *et al.* 2006). In a randomized experimental study, Schellenberg (2004) showed that a group of 6-year-olds who received keyboard or voice lessons for 36 weeks had significantly larger (although modest) increases in full-scale IQ and standardized educational achievement tests than matched groups of children receiving either drama lessons or no lessons. Schellenberg argued that music lessons simply function as additional schooling—requiring focused attention, memorization, and the progressive mastery of technical skills: it is well established that schooling increases IQ (Ceci and Williams 1997). In our 9–11-year-old cross-sectional comparison, we found differences in several but not all domains. The linear relationship between practice intensity and outcome variables in several far transfer domains strongly suggests a causal relationship, but is not absolute proof of it. The possibility of a causal link between music training and enhanced performance on transfer outcomes was confirmed for near transfer in our longitudinal study after a 15-month observation period. The trends that we observed in other domains suggest that a longer observation period and/or higher intensity of training and practice may be required in order to see effects in far transfer domains.

## Making music specifically engages and changes a seeing-hearing-doing (mirror neuron) network

Musicians and non-musicians show strong differences in a functional brain network that is activated while listening to music. In addition to primary and secondary perceptual regions, this functional network also includes regions in the parietal and frontal lobes. Figure 18.4 shows this extensive functional network in a group comparison of adult musicians and non-musicians, and in 5–7-year-old children prior to music training, all of whom performed the same rhythmic discrimination task while their brain activity was being measured with functional magnetic resonance imaging (fMRI). Normal maturational changes can be seen in the comparison between the adult non-musicians and the musically naive children, while maturational changes modulated by musical experiences result in a different pattern of activation.

The additional regions activated in adult musicians include the parietal lobe (around the intraparietal sulcus), the posterior middle frontal gyrus (Brodmann area 6), and the inferior frontal gyrus (Brodmann areas 44, 45, and 47). In the paragraphs below, I will concentrate on discussing frontal lobe differences, since these regions show the most pronounced between-group differences.

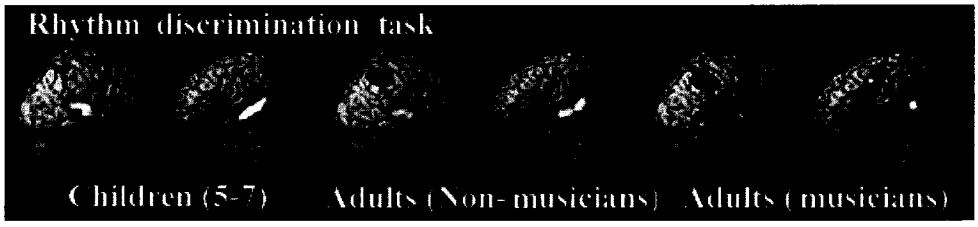



**Fig. 18.4** Cerebral activation pattern of a rhythm discrimination task – modulated by maturation and experience. Statistical parametric images superimposed onto a surface rendering of a standardized anatomical brain depict significant activations during a rhythmic discrimination task in a group of 5–7-year-old, musically naive children, adult non-musicians, and adult musicians. The children showed prominent superior temporal gyrus activation on both sides. The adult groups show an extended pattern of activation involving polar and posterior planar regions of the superior temporal lobe as well as the parietal lobe, parts of the frontal lobe, and the cerebellum. Adult musicians differ from adult non-musicians by having less activation of the primary auditory cortex, but more activation of frontal regions bilaterally, particularly in the inferior frontal gyrus. See colour plate section.

Using functional imaging, it has been found that instrumental practice leads to the establishment of an auditory and visual–sensorimotor network (Bangert and Altenmueller 2003; Bangert *et al.* 2006). Listening to sounds of (and/or seeing) well-learned actions elicits involuntary activity in a network of brain regions (Haueisen and Knösche 2001; Kohler *et al.* 2002; Keysers *et al.* 2003; Hasegawa *et al.* 2004; Haslinger *et al.* 2005; Bangert *et al.* 2006; D'Ausilio *et al.* 2006; Lahav *et al.* 2007). A prominent component of this brain network is the posterior inferior frontal gyrus (IFG) (commonly referred to as Broca's region). It has been shown that the IFG is activated when one sees meaningful, goal-directed actions or hears action-related sounds including grasping food (Rizzolatti *et al.* 1996), imitating hand gestures (Buccino *et al.* 2004a), playing of guitar chords (Buccino *et al.* 2004b), hearing action words (Pulvermuller *et al.* 2005; Aziz-Zadeh *et al.* 2006), hearing instrumental sounds that are associated with motor actions (Lahav *et al.* 2007), as well as when the action is executed. These are functions and properties of a system that is commonly described as mirror neuron system. Mirror neurons were originally described by Rizzolatti *et al.* (1996) in region F5 in monkeys which is considered to be the precursor or homolog of Broca's region or the posterior IFG in humans (Rizzolatti and Arbib, 1998). 'Audiovisual' mirror neurons seem to encode the action abstractly, independent of the information source (auditory or visual) from which evidence regarding the presence of the action is taken (Keysers *et al.* 2003). These 'auditory–visual' mirror neurons exemplify high-level abstraction in the representation of actions—an identical neural system becomes activated whenever a particular action is heard, seen or performed. It is of particular interest that two, independent, voxel-based morphometric studies found more grey matter volume in the inferior frontal gyrus in musicians compared with non-musicians (Sluming *et al.* 2002; Gaser and Schlaug 2003; see also Figure 18.1) which in addition to parietal regions (supramarginal gyrus and superior parietal lobule) and more dorsal premotor regions constitute this network of multimodal–sensorimotor integration. Listening to, seeing notation for, and watching performances of music that one knows how to play, and actually playing that music all appear to engage this putative human mirror neuron system.

## Conclusions

Making music not only engages primary auditory and motor regions and the connections between them, but also engages regions that integrate and connect areas involved in both auditory and motor operations as well as in the integration of other multisensory information.




**Fig. 18.4** Cerebral activation pattern of a rhythm discrimination task – modulated by maturation and experience. Statistical parametric images superimposed onto a surface rendering of a standardized anatomical brain depict significant activations during a rhythmic discrimination task in a group of 5–7-year-old, musically naive children, adult non-musicians, and adult musicians. The children showed prominent superior temporal gyrus activation on both sides. The adult groups show an extended pattern of activation involving polar and posterior planar regions of the superior temporal lobe as well as the parietal lobe, parts of the frontal lobe, and the cerebellum. Adult musicians differ from adult non-musicians by having less activation of the primary auditory cortex, but more activation of frontal regions bilaterally, particularly in the inferior frontal gyrus. See colour plate section.

Using functional imaging, it has been found that instrumental practice leads to the establishment of an auditory and visual–sensorimotor network (Bangert and Altenmueller 2003; Bangert *et al.* 2006). Listening to sounds of (and/or seeing) well-learned actions elicits involuntary activity in a network of brain regions (Haueisen and Knösche 2001; Kohler *et al.* 2002; Keysers *et al.* 2003; Hasegawa *et al.* 2004; Haslinger *et al.* 2005; Bangert *et al.* 2006; D'Ausilio *et al.* 2006; Lahav *et al.* 2007). A prominent component of this brain network is the posterior inferior frontal gyrus (IFG) (commonly referred to as Broca's region). It has been shown that the IFG is activated when one sees meaningful, goal-directed actions or hears action-related sounds including grasping food (Rizzolatti *et al.* 1996), imitating hand gestures (Buccino *et al.* 2004a), playing of guitar chords (Buccino *et al.* 2004b), hearing action words (Pulvermuller *et al.* 2005; Aziz-Zadeh *et al.* 2006), hearing instrumental sounds that are associated with motor actions (Lahav *et al.* 2007), as well as when the action is executed. These are functions and properties of a system that is commonly described as mirror neuron system. Mirror neurons were originally described by Rizzolatti *et al.* (1996) in region F5 in monkeys which is considered to be the precursor or homolog of Broca's region or the posterior IFG in humans (Rizzolatti and Arbib, 1998). 'Audiovisual' mirror neurons seem to encode the action abstractly, independent of the information source (auditory or visual) from which evidence regarding the presence of the action is taken (Keysers *et al.* 2003). These 'auditory–visual' mirror neurons exemplify high-level abstraction in the representation of actions—an identical neural system becomes activated whenever a particular action is heard, seen or performed. It is of particular interest that two, independent, voxel-based morphometric studies found more grey matter volume in the inferior frontal gyrus in musicians compared with non-musicians (Sluming *et al.* 2002; Gaser and Schlaug 2003; see also Figure 18.1) which in addition to parietal regions (supramarginal gyrus and superior parietal lobule) and more dorsal premotor regions constitute this network of multimodal–sensorimotor integration. Listening to, seeing notation for, and watching performances of music that one knows how to play, and actually playing that music all appear to engage this putative human mirror neuron system.

## Conclusions

Making music not only engages primary auditory and motor regions and the connections between them, but also engages regions that integrate and connect areas involved in both auditory and motor operations as well as in the integration of other multisensory information.

Professional instrumentalists learn and repeatedly practice associatiing hand/finger movements with meaningful patterns in sound, and sounds and movements with specific visual patterns (notation) while receiving continuous multisensory feedback. Learning to associate actions with particular sounds leads to functional but also structural changes in frontal cortices. In considering the putative functions of the inferior frontal regions, including its hypothesized role as a supramodal hierarchical processor, auditory–motor mapping region, and hub of the mirror neuron system, it may not be surprising to see that some of the most pronounced differences between musicians and non-musicians, both structurally and functionally, are indeed, in the inferior frontal region. Training this network and changing it as a result of that training has a cross-modal effect on other behavioural/cognitive operations that draw on this network. This cross-modal plasticity in multimodal brain regions may explain some of the cognitive enhancements and/or far transfer effects that have been associated with long term instrumental music training.

## Acknowledgements

The research discussed in this chapter was supported by grants from the National Science Foundation (NSF), the International Foundation for Music Research (IFMR), and The Grammy Foundation. Christian Gaser, Andrea Norton, Katie Overy, Ellen Winner, Andrew Norton, Camilla Rosam, and Marie Forgeard have contributed to the research presented in this chapter.

## References

Amunts K, Schlaug G, Jäncke L, Dabringhaus A, Steinmetz H, Schleicher A and Zilles K (1997). Motor cortex and hand motor skills: structural compliance in the human brain. *Human Brain Mapp*, **5**, 206–215.

Anderson BJ, Eckburg PB and Relucio KI (2002). Alterations in the thickness of motor cortical subregions after motor skill learning and exercise. *Learn Mem*, **9**, 1–9.

Anderson BJ, Li S, Alcantara A, Isaacs KR, Black JE and Greenough WT (1994). Glial hypertrophy is associated with synaptogenesis following motor-skill learning but not with angiogenesis, following exercise. *Glia*, **11**, 73–80.

Aziz-Zadeh L, Wilson SM, Rizzolatti G and Iacoboni M (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Curr Biol*, **16**, 1818–1823.

Bangert M and Schlaug G (2006). Specialization of the specialized in features of external human brain morphology. *Eur J Neurosci*, **24**, 1832–1834.

Bangert M and Altenmüller EO (2003). Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neurosci*, **4**, 26.

Bangert M, Peschel T, Rotte M, Drescher D, Hinrichs H, Schlaug G, Heinze HJ and Altenmüller E (2006). Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *NeuroImage*, **30**, 917–926.

Bangerter A and Heath C (2004). The Mozart effect: tracking the evolution of a scientific legend. *Br J Soc Psychol*, **43**, 605–623.

Bengtsson SL, Nagy Z, Skare S, Forsman L, Forssberg H and Ullen F (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nat Neurosci*, **8**, 1148–1150.

Besson M, Gaita F and Requin J (1994). Brain waves associated with musical incongruities differ for musicians and non-musicians. *Neurosci Lett*, **168**, 101–105.

Black JE, Isaacs KR, Anderson BJ, Alcantra AA and Greenough WT (1990). Learning causes synaptogenesis whereas motor activity causes angiogenesis in cerebellar cortex of adult rats. *Proc Natl Acad Sci USA*, **87**, 5568–5572.

Bosnyak DJ, Eaton RA and Roberts LE (2004). Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. *Cereb Cortex*, **14**, 1088–1099.

Bruer J (1999). *The myth of the first three years*. Free Press, New York.

Buccino G, Binkofskik F and Riggio L (2004a). The mirror neuron system and action recognition. *Brain Lang*, **89**, 370–376.

Buccino G, Vogt S, Ritzl A, Fink GR, Zilles K, Freund HJ and Rizzolatti G (2004b). Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. *Neuron*, **42**, 323–334.

Ceci SJ and Williams WM (1997). Schooling intelligence and income. *Am Psychol*, **52**, 1051–1058.

Chan AS, Ho Y and Cheung M (1998). Music training improves verbal memory. *Nature*, **396**, 128.

Chochon F, Cohen L, van de Moortele PF and Dehaene S (1999). Differential contributions of the left and right inferior parietal lobules to number processing. *J Cogn Neurosci*, **11**, 617–630.

D'Ausilio A, Altenmüller E, Olivetti Belardinelli M and Lotze M (2006). Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *Eur J Neurosci*, **24**, 955–958.

Elbert T, Pantev C, Wienbruch C, Rockstroh B and Taub E (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, **270**, 305–306.

Fauvel J, Flood R and Wilson R (2006). Music and mathematics: from Pythagoras to fractals. Oxford Press, New York.

Forgeard M, Winner E, Norton A, Overy K, Iyengar U, Goldstein TR and Schlaug G. Effects of fifteen months of instrumental music training: evidence for near transfer, submitted.

Gaab N and Schlaug G (2003). The effect of musicianship on pitch memory in performance matched groups. *NeuroReport*, **14**, 2291–2295.

Gaab N, Gaser C and Schlaug G (2006). Improvement-related functional plasticity following pitch memory training. *Neuroimage*, **31**, 255–263.

Gaab N, Gaser C, Zaehle T, Jäncke L and Schlaug G (2003). Functional anatomy of pitch memory—an fMRI study with sparse temporal sampling. *Neuroimage*, **19**, 1417–1426.

Gaser C and Schlaug G (2003). Brain structures differ between musicians and non-musicians. *J Neurosci*, **23**, 9240–9245.

Hasegawa T, Matsuki K, Ueno T, Maeda Y, Matsue Y, Konishi Y and Sadato N (2004). Learned audio-visual cross-modal associations in observed piano playing activate the left planum temporale. An fMRI study. *Cogn Brain Res*, **20**, 510–518.

Haslinger B, Erhard P, Altenmüller E, Hennenlotter A, Schwaiger M, Einsiedel H, Rummeny E, Conrad B and Ceballos-Baumann AO (2004). Reduced recruitment of motor association areas during bimanual coordination in concert pianists. *Human Brain Map*, **22**, 206–215.

Haslinger B, Erhard P, Altenmüller E, Schroeder U, Boecker H and Ceballos-Baumann AO (2005). Transmodal sensorimotor networks during action observation in professional pianists. *J Cogn Neurosci*, **17**, 282–293.

Haueisen J and Knösche TR (2001). Involuntary motor activity in pianists evoked by music perception. *J Cogn Neurosci*, **13**, 786–792.

Hetland L (2000). Learning to make music enhances spatial reasoning. *J Aesthetic Educ*, **34**, 179–238.

Ho YC, Cheung MC and Chan AS (2003). Music training improves verbal but not visual memory: cross-sectional and longitudinal explorations in children. *Neuropsychology*, **17**, 439–445.

Hund-Georgiadis M and von Cramon Y (1999). Motor-learning-related changes in piano players and non-musicians revealed by functional magnetic resonance signals. *Exp Brain Res*, **125**, 417–425.

Hutchinson S, Lee LHL, Gaab N and Schlaug G (2003). Cerebellar volume of musicians. *Cereb Cortex*, **13**, 943–949.

Isaacs KR, Anderson BJ, Alcantara AA, Black JE and Greenough WT (1992). Exercise and the brain: angiogenesis in the adult rat cerebellum after vigorous physical activity and motor skill learning. *J Cereb Blood Flow Metab*, **12**, 110–119.

Jäncke L, Shah NJ and Peters M (2000). Cortical activations in primary and secondary motor areas for complex bimanual movements in professional pianists. *Cogn Brain Res*, **10**, 177–183.

Keenan JP, Halpern AR, Thangaraj V, Chen C, Edelman RR and Schlaug G (2001). Absolute pitch and planum temporale. *Neuroimage*, **14**, 1402–1408.

Kempermann G, Kuhn HG and Gage FH (1997). More hippocampal neurons in adult mice living in an enriched environment. *Nature*, **386**, 493–495.

Keysers C, Kohler E, Umiltà MA, Nanetti L, Fogassi L and Gallese V (2003). Audiovisual mirror neurons and action recognition. *Exp Brain Res*, **153**, 628–636.

Kleim JA, Lussnig E, Schwarz ER, Comery TA and Greenough WT (1996). Synaptogenesis and fos expression in the motor cortex of the adult rat after complex motor skill acquisition. *J Neurosci*, **16**, 4529–4535.

Koelsch S, Fritz T, Schulze K, Alsop D and Schlaug G (2005). Adults and children processing music: an fMRI study. *Neuroimage*, **25**, 1068–1076.

Koelsch S, Gunter TC, v Cramon DY, Zysset S, Lohmann G and Friederici AD (2002). Bach speaks: a cortical 'language-network' serves the processing of music. *Neuroimage*, **17**, 956–966.

Koeneke S, Lutz K, Wüstenberg T and Jäncke L (2004). Long-term training affects cerebellar processing in skilled keyboard players. *Neuroreport*, **15**, 1279–1282.

Kohler E, Keysers C, Umiltà MA, Fogassi L, Gallese V and Rizzolatti G (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, **297**, 846–848.

Krings T, Töpper R, Foltys H, Erberich S, Sparing R, Willmes K and Thron A (2000). Cortical activation patterns during complex motor tasks in piano players and control subjects. A functional magnetic resonance imaging study. *Neurosci Lett*, **278**, 189–193.

Lahav A, Saltzman E and Schlaug G (2007). Action representation of sound: audiomotor recognition network while listening to newly-acquired actions. *J Neurosci*, **27**, 308–314.

Lee DJ, Chen Y and Schlaug G. (2003). Corpus callosum: musician and gender effects. *NeuroReport*, **14**, 205–209.

Levitin DJ and Menon V (2003). Musical structure is processed in language areas of the brain: a possible role for Brodmann Area 47 in temporal coherence. *Neuroimage*, **20**, 2142–2152.

Lotze M, Scheler G, Tan HR, Braun C and Birbaumer N (2003). The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *Neuroimage*, **20**, 1817–1829.

Meister I, Krings T, Foltys H, Boroojerdi B, Mueller M, Töpper R and Thron A (2005). Effects of long-term practice and task complexity in musicians and non-musicians performing cortical motor organization. *Human Brain Mapp*, **25**, 345–352.

Münte TF, Kohlmetz C, Nager W and Altenmüller E (2001). Neuroperception: superior auditory spatial tuning in conductors. *Nature*, **409**, 580.

Nair DG, von Bueren Jarchow A, Pascual-Leone A, and Schlaug G (2006). Differential inter-hemispheric inhibition in pianists and string players: TMS evidence. Human Brain Mapping Meeting, Florence, 2006.

Nordstrom MA and Butler SL (2002). Reduced intracortical inhibition and facilitation of corticospinal neurons in musicians. *Exp Brain Res*, **144**, 336–342.

Norton A, Winner E, Cronin K, Overy K, Lee D and Schlaug G (2005). Are there pre-existing neural, cognitive, or motoric markers for musical ability? *Brain Cogn*, **59**, 124–134.

Overy K, Norton A, Cronin K, Gaab N, Alsop D, Winner E and Schlaug G (2004). Imaging melody and rhythm processing in young children. *NeuroReport*, **15**, 1723–1726.

Oztürk AH, Tascioglu B, Aktekin M, Kurtoglu Z and Erden I (2002). Morphometric comparison of the human corpus callosum in professional musicians and non-musicians by using in vivo magnetic resonance imaging. *J Neuroradiol*, **29**, 29–34.

Pantev C, Oostenveld R, Engelien A, Ross B, Roberts LE and Hoke M (1998). Increased auditory cortical representation in musicians. *Nature*, **392**, 811–814.

Pantev C, Roberts LE, Schulz M, Engelien A and Ross B (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport*, **12**, 169–174.

Patel AD (2003). Language, music, syntax and the brain. *Nat Neurosci*, **6**, 674–681.

Pulvermüller F, Shtyrov Y and Ilmoniemi R (2005). Brain signatures of meaning access in action word recognition. *J Cogn Neurosci*, **17**, 884–892.

Rauscher FH and Zupan MA (2000). Classroom keyboard instruction improves kindergarten children's spatial–temporal performance: a field experiment. *Early Childhood Research Quarterly*, **15**, 215–228.

Rauscher FH, Shaw GL and Ky KN (1993). Music and spatial task performance. Nature, 365, 611.

Rauscher FH, Shaw GL, Levine L, Wrigh E, Dennis W and Newcomb R (1997). Music training causes long-term enhancement of preschool children's spatial–temporal reasoning. *Neurol Res*, **19**, 2–8.

Ridding MC, Brouwer B and Nordstrom MA (2000). Reduced interhemispheric inhibition in musicians. *Exp Brain Res1*, **33**, 249–253.

Rizzolatti G and Arbib MA (1998). Language within our grasp. *TINS*, **21**, 188–194.

Rizzolatti G, Fadiga L, Gallese V and Fogass L (1996). Premotor cortex and the recognition of motor actions. *Cogn Brain Res*, **3**, 131–141.

Schellenberg EG (2004). Music lessons enhance IQ. *Psychol Sci*, **15**, 511–514.

Schlaug G (2001). The brain of musicians: a model for functional and structural adaptation. In R Zatorre and I Peretz, eds, The biological foundations of music. *Ann NY Acad Sci*, **930**, 281–299.

Schlaug G, Jäncke L, Huang Y and Steinmetz H. (1995a). In vivo evidence of structural brain asymmetry in musicians. *Science*, **267**, 699–671.

Schlaug G, Jäncke L, Huang Y, Staiger JF and Steinmetz H. (1995b). Increased corpus callosum size in musicians. *Neuropsychologia*, **33**, 1047–1055.

Schlaug G, Norton AC, Overy K and Winner E (2005). Effects of music training on the child's brain and cognitive development. *Ann NY Acad Sci*, **1060**, 219–230.

Schneider P, Scherg M, Dosch HG, Specht HJ, Gutschalk A and Rupp A (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat Neurosci*, **5**, 688–694.

Sergent J, Zuck E, Terriah S and MacDonald B (1993). Distributed neural network underlying musical sight-reading and keyboard performance. *Science*, **257**, 106–109.

Shahin A, Bosnyak DJ, Trainor LJ and Roberts LE (2003). Enhancemment of neuroplastic P2 and N1c auditory evoked potentials in skilled musicians. *J Neurosci*, **23**, 5545–5552.

Shahin A, Roberts LE and Trainor LJ (2004). Enhancement of auditory cortical development by musical experience in children. *Neuroreport*, **15**, 1917–1921.

Sluming V, Barrick T, Howard M, Cezayirli E, Mayes A and Roberts N (2002). Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *Neuroimage*, **17**, 1613–1622.

Sluming V, Brooks J, Howard M, Downes JJ and Roberts N (2007). Broca's area supports enhanced visuospatial cognition in orchestral musicians. *J Neurosci*, **27**, 3799–3806.

Stewart L, Henson R, Kampe K, Walsh V, Turner R and Frith U (2003). Brain changes after learning to read and play music. *NeuroImage*, **20**, 71–83.

Tettamanti M and Weniger D (2006). Broca's area: a supramodal hierarchical processor? *Cortex*, **42**(4), 491–494.

Trainor LJ, Desjardins RN and Rockel C (1999). A comparison of contour and interval processing in musicians and nonmusicians using event-related potentials. *Aust J Psychol*, **51**, 147–153.

Van Praag H, Christie BR, Sejnowksi TJ and Gage FH (1999). Running enhances neurogenesis, learning, and long-term potentiation in mice. *Proc Natl Acad Sci USA*, **96**, 13427–13431.

Vaughn K (2000). Music and mathematics: modest support for the oft-claimed relationship. *Journal of Aesthetic Education*, **34**, 149–166.

Zatorre RJ, Chen JL and Penhune VB (2007). When the brain plays music. Auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, **8**, 547–558.

Zatorre RJ, Evans AC and Meyer E (1994). Neural mechanism underlying melodic perception and memory for pitch. *J Neurosci*, **14**, 1908–1919.

Zatorre RJ, Perry DW, Beckett CA, Westbury CF and Evans AC (1998). Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. *Proc Natl Acad Sci USA*, **95**, 3172–3177.

CHAPTER 19

# Music and the brain

## Three links to language

Aniruddh D. Patel

## Introduction

A major theme in the neuroscience of music over the past decade has been the existence of specialized cognitive and neural mechanisms involved in music processing (Peretz and Coltheart 2003). Research with brain damaged patients has been particularly important in demonstrating that music involves a neurocognitive system with many components, and that certain components can be selectively damaged without obvious effects on other cognitive abilities. These findings point to some degree of brain specialization for music, specialization that emerges even in the absence of any explicit musical training (Bigand and Poulin-Charronnat 2006).

One can recognize that music is a unique confluence of component abilities, yet still search for connections between music cognition and processes involved in other cognitive domains, such as language, visuospatial thinking, or mathematics. Such comparative work can be revealing about the mechanisms underlying music processing, and can provide new avenues for exploring how these mechanisms work in the brain.

The goal of this chapter is to acquaint the reader with some of the evidence for links between music and language, and to encourage further exploration along these lines (see Patel 2008 for a fuller treatment). The focus is on perceptual processes, and on links between mechanisms involved in the processing of instrumental music and of ordinary, day to day language.

## Sensory encoding of sound

There is growing evidence that musical abilities and training have repercussions for language skills. For example, musical abilities predict unique variance in the ability to perceive and produce subtle phonetic contrasts in a second language (Slevc and Miyake 2006) and in the reading abilities of young children in their first language (Anvari *et al.* 2002). Furthermore, musical training enhances the ability to interpret affective prosody (Thompson *et al.* 2004). Are these benefits simply a non-specific consequence of musical training on attention and executive functioning, or are there more specific links between musical abilities and language processing? Support for the latter view has emerged recently via neuroscientific investigations of brainstem encoding of sound. These studies suggest that musical abilities and/or training sharpen the brain's encoding of linguistic sound at a very early stage of processing.

The human brainstem contains many subcortical processing nuclei, which perform a significant amount of neural processing before sound reaches the cortex. Importantly, these lower brain centres are connected to cortical areas via both bottom-up and top-down connections (cf. Figure 1 of Patel and Iversen 2007). The latter ('corticofugal') connections provide one possible mechanism by which long-term auditory experience may influence early brain responses to sound, via mechanisms of neural plasticity.

Wong *et al.* (2007) first demonstrated an influence of musical abilities on linguistic sound encoding. These researchers examined the

processing of linguistic pitch contours in Mandarin Chinese syllables by musically trained and untrained native English speakers (neither of whom was familiar with Mandarin). The participants listened passively to Mandarin monosyllables with different lexical tones while their brainstem responses were recorded using electroencephalography (EEG). The researchers found that the musicians showed superior brainstem encoding of the linguistic pitch patterns. They also found a positive correlation between the quality of sensory encoding and amount of musical training, suggesting a role for musical experience (rather than innate differences between musicians and non-musicians) in shaping the sensory encoding of linguistic sound.

An impact of musical abilities on linguistic pitch processing seems intuitively plausible because music places strong demands on the control and/or perception of pitch. However, speech also makes extensive use of structured auditory patterns not based on pitch (e.g., timbre-based differences between phonemes). Do musical abilities have any impact on the sensory encoding of timbre-based aspects of language sounds? Recent results by Musacchia *et al.* (2007) address this issue. Like Wong *et al.* (2007), these researchers measured auditory brainstem responses to speech using EEG, but this time to a familiar English syllable ('da') with no salient pitch variation. They found that musicians had earlier brainstem responses than non-musicians to the onset of the syllable. Interestingly, when the syllable was presented as a multimodal event (sound + video of a face saying the syllable), the latency of the brainstem response was shortened for both musicians and non-musicians, but the musicians still showed an earlier response than the non-musicians. The authors suggest that this difference in response latency indicates that musicians had more synchronous neural responses to sound onset, which is a sign of a high-functioning peripheral auditory system.

The superior encoding of linguistic sounds by musicians may be one possible mechanism underlying the linguistic benefits of musical abilities (cf. Tallal and Gaab 2006; Patel and Iversen 2007). These findings may also be relevant to research which has examined the impact of musical training on the cortical processing of linguistic pitch patterns. These studies have examined how well individuals can detect a deviation in voice fundamental frequency (F0) at the end of a spoken utterance. Using EEG, these studies have focused on cortical event-related potentials (ERPs) which occur with a latency of a few hundred milliseconds after the onset of the deviant event. Research with French-speaking adults (Schön *et al.* 2004) and 8-year-old children (Magne *et al.* 2006) has shown that musically trained listeners are better at detecting such deviations, and show enhanced cortical ERPs to these events. In fact, even when listening to sentences in a foreign language (Portuguese), French adults with musical training are better at detecting such deviations, and show substantially shorter latency in their cortical ERPs to those deviations (Marques *et al.* 2007). One notable finding of this line of research is that the influence of musical training on cortical responses emerges remarkably quickly: 8-year-old children with just 8 weeks of musical training differ from their musically untrained counterparts in their cortical ERPs to intonational deviations in speech (Moreno and Besson 2006).

Further work is needed to clarify the relationship between subcortical and cortical enhancements to linguistic sound processing associated with musical training. Since connections in the auditory pathway flow both bottom-up and top-down, experimental studies are needed to disentangle patterns of cause and effect at different levels of the nervous system as the brain changes in response to musical training. Ideally, such studies would measure subcortical and cortical ERP responses to linguistic pitch patterns within the same individuals as they undergo musical training programmes, and relate these measures to behavioural performance on pitch-related language tasks. Baseline measures of brain responses to linguistic sound (prior to musical training) would be desirable, to show that differences between musically trained and untrained groups are not present prior to training (cf. Norton *et al.* 2005).

## Processing of melodic contours

Humans readily recognize a pitch pattern transposed upward or downward in frequency

(e.g., the same melody played on a piccolo or a tuba). This sensitivity to relative (vs absolute) pitch patterns seems commonplace to us, yet a cross-species perspective reveals that this is a rare ability (McDermott and Hauser 2005). Most birds and non-human primates, for example, gravitate to absolute pitch cues in identifying pitch sequences, and extensive training is often needed before sensitivity to relative pitch can be demonstrated.

This specialization in human brains for sensitivity to relative pitch may be related to the importance of relative pitch in spoken intonation. A listener needs to be able to hear the similarity of intonation patterns when spoken in different pitch registers. For example, a child learning English needs to understand that a rising pitch at the end of an intonation contour can signal a question, whether the utterance is produced by another child or by an adult (with a much lower average F0). Similarly, the child needs to learn that a salient pitch excursion on a word in a sentence signals that the speaker is trying to put focus on that word (as in 'give me the RED toy'), even though the precise size of the pitch excursion may differ from one speaker to the next. Thus speech intonation perception requires processing of melodic contour: the general patterns of up and down in a pitch sequence, independent of exact interval sizes or absolute pitch height (Dilley 2005).

Many experiments on music have shown that melodic contour is an important component of music perception. For example, Dowling and colleagues have shown that melodic contour (vs the pattern of precise intervals in a melody) plays a role in immediate memory for unfamiliar melodies among adults (e.g., Dowling *et al.* 1995). Furthermore, melodic contour is one of the first aspects of music to be discriminated by infants (Trehub *et al.* 1984), and 5-year-old children rely heavily on contour in melody perception (Schellenberg and Trehub 1999). Studies of patients with brain damage indicate that the right superior temporal gyrus plays an important role in musical melodic contour perception (Peretz and Zatorre 2005, but cf. Stewart *et al.* 2008).

Given the importance of melodic contour in both speech and music, it is natural to ask whether melodic contours are processed by similar brain mechanisms in the two domains. Early evidence for this idea was provided by Patel *et al.* (1998), who studied two individuals with acquired amusia subsequent to cortical brain damage. The amusics were tested for their ability to discriminate between sentences that differed only in their intonation contour. Two types of sentence pairs were tested. In 'statement–question' pairs, the intonation contours differed at the end of the sentence (e.g., 'He likes to drive fast cars' spoken with a pitch fall or rise on 'cars'). In 'focus–shift' pairs, the intonation contours differed within the sentence (e.g., 'Go in front of the bank, I said' spoken with a salient pitch accent on 'front' or 'bank'). The amusics were also tested for their ability to discriminate tone sequence analogues of intonation patterns. These analogues were created by replacing each syllable in a sentence with a tone whose pitch was fixed at the frequency midway between the highest and lowest F0 of the syllable. Discrete tones were used rather than gliding tones (which would mimic intonation contours more precisely) in order to make the stimuli more music-like. The salient finding from this study was that for both amusics, performance on the linguistic intonation task was very similar to performance on the tone sequence task, suggesting shared processing of melodic contours across the two domains. (One amusic, whose problem was primarily with recognition of familiar music, did well on both tasks, while the other amusic, who had more basic problems with perception and discrimination of musical patterns, did poorly on both tasks.) Examination of the lesion profiles of the amusics suggested that right inferior frontal brain areas were important for the task of remembering and discriminating melodic contours (cf. Zatorre *et al.* 1994).

In contrast to these findings, subsequent research with musically tone-deaf individuals painted a different picture. Before discussing this work, it is worth briefly describing musical tone deafness (or congenital amusia). Congenital amusia (henceforth, 'amusia') is not due to brain damage and appears to be developmental in origin. Amusic individuals report life-long problems with music perception. (Note that it is important to distinguish amusia from the informal label of 'tone-deafness' that many people apply to themselves, usually to indicate that they

do not sing well, cf. Cuddy *et al.* [2005]). For example, amusics have difficulty judging if two melodies are the same or different, in detecting when music is out of key (including their own singing), and in recognizing what should be familiar tunes from their culture (Ayotte *et al.* 2002). These problems cannot be attributed to hearing loss, lack of exposure to music, or to any obvious non-musical social or cognitive impairments. The core deficit in this disorder concerns pitch processing (Foxton *et al.* 2004; Hyde and Peretz 2004). It appears that there is a genetic basis for this disorder (Drayna *et al.* 2001; Peretz *et al.* 2007), and evidence from neuroimaging has revealed specific structural differences between normal and amusic brains in right inferior frontal cortex and right superior temporal gryus, as well as in left frontal and temporal cortical areas (Hyde *et al.* 2007; Mandell *et al.* 2007).

Amusia presents an excellent opportunity to study the relationship between melodic contour processing in speech and music. Ayotte *et al.* (2002) first examined speech intonation perception in congenital amusia, using the stimuli and methods of Patel *et al.* (1998). The amusics had no problem discriminating sentences with different intonation contours. In contrast, they had difficulty discriminating the tone sequence analogs of intonation. These results suggested that amusics were unimpaired in speech intonation perception, and prompted Peretz and Coltheart (2003) to include melodic contour analysis as a domain-specific aspect of musical processing, not shared with speech.

Subsequent work, however, has re-opened the issue of speech intonation perception in amusia. Lochy *et al.* (2004) tested a group of amusics using methods very similar to those of Ayotte *et al.* (2002), but with different results. This time, about 30 per cent of the amusics had significant difficulty discriminating linguistic statements from questions, whereas they did well discriminating sentences with focus on different words. The critical difference between these tasks is that the statement-question task requires discriminating the direction of pitch movement on the same word (up versus down), whereas the focus-shift task simply requires detecting a salient pitch movement within a sentence, because different words bear the large movement in the two members of a focus-shift pair. That is, sensitivity to the direction of pitch movement is irrelevant to the focus-shift task: as long as one can detect a pitch change, and can remember that this change happened on the same or different words in the two sentences, one can solve the task.

These findings were intriguing because psychophysical work on amusia has revealed a salient deficit for the perception of pitch direction (Foxton *et al.* 2004; Griffiths *et al.* 2007). That is, an amusic individual may be able to detect that pitch has changed, but not be able to discriminate an upward from a downward change. Lochy *et al.*'s findings suggest that such direction discrimination deficits may be relevant for speech intonation perception. Given the differences between the findings of Ayotte *et al.* (2002) and Lochy *et al.* (2004), it is important to determine which results replicate. Recently, Patel *et al.* (2008) replicated Lochy *et al.*'s findings with a different group of amusics, finding that about 30 per cent had problems with statement-question discrimination but not focus-shift discrimination. Notably, this replication occurred in a different language (English-speaking amusics, vs the French-speaking amusics of Lochy *et al.*). An important question raised by the results of Lochy *et al.* (2004) and Patel *et al.* (2008) is why only a subset of amusics (~30 per cent) exhibit intonation perception deficits. Does this subset have a qualitatively different kind of deficit from amusics without such problems, or are they simply at the end of a continuum such that all amusics would show intonation perception problems if pitch contrasts are made subtle enough? These questions await further research, but the current findings are sufficient to indicate that melodic contour processing in speech and music may in fact have overlapping neural substrates.

## Syntactic processing

Music and language involve hierarchically structured sequences built from basic units via rich combinatorial principles. That is, both are syntactic systems. Native listeners have implicit knowledge of these combinatorial principles, knowledge which requires no formal training

and which can readily be demonstrated via judgements of the well-formedness of novel sequences and via a range of implicit tasks (Bigand and Poulin-Charronnat 2006).

Despite these general similarities, however, it is also clear that musical and linguistic syntax differ in important ways. For example, fundamental grammatical categories in language (such as nouns and verbs) have no analogue in music, and the function of linguistic syntax in indicating 'who did what to whom' also has no musical parallel. Instead, (Western) musical syntax is based on selecting a set of seven discrete pitches within each octave and creating a musical 'key' by using these pitches in such a way that a hierarchy of structural importance is created among them. This leads to certain pitches and chords being perceived as more stable than others. These stable elements act as cognitive reference points, with one (the tonic note or chord) being the most stable (Krumhansl 1990). This syntax functions to build patterns of tension and resolution in musical sequences (Lerdahl and Jackendoff 1983, Lerdahl and Krumhansl 2007). Thus it is clear that musical syntax is not simply a trivial variant of linguistic syntax.

For many years, the evidence from neuroscience seemed to favour the independence of musical and linguistic syntax in the brain. Specifically, research on the neuropsychology of music provided clear cases of dissociations between linguistic and musical syntactic abilities following brain damage. For example, Peretz (1993) documented a case of a non-aphasic man with bilateral temporal lobe damage (due to strokes) who lost sensitivity to musical key, even though his basic perception of pitch patterns was intact. This is one of several well-documented cases of 'amusia without aphasia' (cf. Marin and Perry 1999).

Yet in direct contrast to these findings, neuroimaging research has pointed to overlap in the syntactic processing of language and music. An early study which directly compared brain responses to musical and linguistic syntactic processing in the same individuals was conducted by Patel *et al.* (1998). The researchers used ERPs to compare brain responses (from musicians) to linguistic syntactic and musical syntactic incongruities. The linguistic incongruities involved manipulations of phrase structure in sentences, and the musical incongruities involved out-of-key chords in chord sequences. The linguistic incongruities generated a late positive ERP component known as the P600, previously known to be associated with syntactic processing in language (Osterhout and Holcomb 1993). The main finding of interest was that the out-of-key chords generated a P600 which was statistically indistinguishable from the 'linguistic' P600. The authors interpreted this to mean that processes involved in syntactic integration were shared across the two domains. In addition to the P600, out-of-key chords also elicited an early right anterior–temporal negativity (RATN) which had not been reported in previous neuroimaging studies of music. The functional significance of this component was not well understood, though it was thought to be involved in musical syntactic processing.

Subsequent ERP work by Koelsch *et al.* (2000) examined brain responses to out-of-key chords in non-musicians, and found an early right anterior negativity (ERAN), which had a shorter latency than the RATN and a more anterior and bilateral scalp distribution. Many subsequent studies by Koelsch and colleagues have found the ERAN to be reliably elicited by structurally unexpected chords, in both musicians and non-musicians, and in the absence of any task related to detecting these chords.[1] Furthermore, elicitation of the ERAN does not require the use of an out-of-key chord: it can also be elicited by a structurally unexpected chord from within the key of the preceding sequence. This rules out auditory sensory incongruity as the driving force behind the elicitation of this component (Koelsch *et al.* 2007).

What is the relationship between the ERAN and the RATN? As suggested by Koelsch and Siebel (2005), it seems likely that they represent the same underlying brain process, and that the earlier latency of the ERAN is related to the temporal predictability of the chord sequences

[1] Another ERP component that is often elicited by out-of-key chords is a negativity peaking around 500 ms after the onset of the target chord (the 'N500'). For possible links between mechanisms involved in this brain response and in language processing, see Steinbeis and Koelsch (2007).

that have typically been used in studies of the ERAN. That is, most such research has used isochronous sequences of chords, rendering the temporal onset of each successive chord quite predictable in time. In contrast, when more temporally variable musical contexts are used (as in Patel *et al.* 1998, and Koelsch and Mulder 2002), an early right negative component is elicited that is more similar to the RATN in latency and scalp distribution. Thus the precise latency and distribution of early right negativities to music is likely to be subject to rhythmic influences.

More interesting from the current standpoint are questions about the relationship of the ERAN to brain mechanisms involved in linguistic syntactic processing. Using MEG, Maess *et al.* (2001) localized the generators of the ERAN to Broca's area and its right hemisphere homologue, suggesting an overlap with language processing. Other research using functional magnetic resonance imaging (fMRI) also implicated Broca's area in the processing of syntactic (harmonic) structure in music (e.g. Tillmann *et al.* 2003). Yet the neuroimaging evidence for overlap between linguistic and musical syntax stood in sharp contrast to evidence from neuropsychology for dissociations between linguistic and musical syntactic abilities following brain damage, as discussed above. Thus it became clear that a theoretical framework was needed which could accommodate both the neuroimaging evidence (for overlap) and the neuropsychology evidence (for dissociation).

Patel (2003) proposed one such framework based on the idea that language and music have distinct and domain-specific syntactic representations (such as words and their syntactic features in language, and chords and their harmonic relations in music), but that activating these representations as part of online processing draws on a common pool of limited neural resources. This idea was termed the 'shared syntactic integration resource hypothesis' (SSIRH). The SSIRH posits that linguistic and musical syntactic representations are stored in distinct brain networks (and hence can be selectively damaged), whereas there is overlap in the networks which provide neural resources for the activation of stored syntactic representations. It was hypothesized that such resources are needed when dealing with difficult structural integrations, because such integrations require the rapid and selective activation of items with low activation levels in representation networks (e.g., structurally unexpected words or chords).

How does this proposal map onto neural architecture? At the moment the answer to this question is not known. In its original formulation, the SSIRH combined the functional proposal outlined above with a rough localizationist proposal, namely that that neural resources reside in frontal brain regions, while syntactic representations reside in posterior regions. Hence syntactic processing is conceived of as involving the reciprocal interaction of frontotemporal regions (cf. Tyler and Marslen-Wilson 2007). Further research using localizationist techniques (such as fMRI) is needed to address this issue.

A principal motivation for developing the SSIRH was to generate predictions to guide future research into the relation of linguistic and musical syntactic processing. One salient prediction regards the interaction of musical and linguistic syntactic processing. In particular, because the SSIRH proposes that linguistic and musical syntactic integration rely on common neural resources, and because syntactic processing resources are limited (Gibson 2000), it predicts that tasks that combine linguistic and musical syntactic integration will show interference between the two. In particular, the SSIRH predicts that difficult harmonic integrations in music will interfere with concurrent difficult syntactic integration in language (Patel 2003). This idea can be tested in paradigms in which a harmonic and a linguistic sequence are presented together and the influence of harmonic structure on syntactic processing in language is studied.

Koelsch *et al.* (2005) conducted one such study, in which short sentences were presented visually in a word-by-word format simultaneously with musical chords, with one chord per word. In some sentences, the final word created a grammatical violation via a gender disagreement. (The sentences were in German, in which many nouns are marked for gender. An example of a gender violation used in this study is: 'Er trinkt den kühlen Bier', 'He drinks the$_{\text{masculine}}$

$cool_{masculine}$ $beer_{neuter}$.) The chord sequences were designed to strongly invoke a particular key, and the final chord could either be the tonic chord of that key or an unexpected out-of-key chord from a distant key (e.g., a D-flat major chord at the end of a C-major sequence). The participants (all non-musicians) were instructed to ignore the music and simply judge if the last word of the sentence was linguistically correct.

Koelsch *et al.* focused on early ERP negativities elicited by syntactically incongruous words and chords. Previous research on language or music alone had shown that the linguistic syntactic incongruities were associated with a left anterior negativity (LAN), while the musical incongruities were associated with an ERAN (Gunter *et al.* 2000; Koelsch *et al.* 2000). For their combined language–music stimuli, Koelsch *et al.* found that when sentences ended grammatically but were accompanied by an out-of-key chord, a normal ERAN was produced. Similarly, when chord sequences ended normally but were accompanied by a syntactically incongruous word, a normal LAN was produced. The question of interest was how these brain responses would interact when a sequence had simultaneous syntactic incongruities in language and music. The main finding was that the brain responses were not simply additive. Instead, there was an interaction: the LAN to syntactically incongruous words was significantly *smaller* when these words were accompanied by an out-of-key chord, as if the processes underlying the LAN and ERAN were competing for similar neural resources. In a control experiment, Koelsch *et al.* showed that this was not due to general attentional effects because the LAN was not influenced by a simple auditory oddball paradigm involving physically deviant tones on the last word in a sentence. Thus the study supported the prediction that tasks which combine linguistic and musical syntactic integration will show interference between the two processes. Further behavioural studies by Slevc *et al.* (in press) and Fedorenko *et al.* (in press) using simultaneous presentation of sentences and musical sequences have provided additional evidence for the interaction of linguistic and musical syntactic processing, even in the absence of any explicit music-related task (cf. Patel in press for descriptions of these studies). Like the study of Koelsch *et al.* (2005), these studies have included control conditions to check that the interactions observed are not simply due to attentional factors.

One line of research which awaits exploration involves within-subjects comparisons of syntactic processing in language and music using fMRI. Such work could help localize shared resource networks for syntactic integration within specific regions of the brain.

## Future directions

This chapter has discussed three points of contact between music and language processing in the brain. Many other areas of possible overlap await investigation, including the neural mechanisms involved in processing phrase boundaries, and the mechanisms involved in the affective appraisal of musical and linguistic sequences (for further discussion of these topics, see Patel 2008). Another little-explored area concerns the relationship between the neural mechanisms involved in the generation of novel linguistic and musical sequences (cf. Brown *et al.* 2006).

## Conclusion

The extent to which music processing overlaps with mechanisms used in other cognitive domains is an active area of research and debate (e.g., Peretz 2006; Patel 2008). One fertile line of work that has grown out of this debate concerns relations between music and language processing. The evidence reviewed in this chapter suggests that music and language may have a number of common processes which act on distinct types of information, e.g., on musical melodies vs linguistic intonation contours, or on chord progressions vs sequences of words. Thus the distinction between the domain specificity of information vs the generality of processing is an essential conceptual tool for research that examines the relationship between music and other cognitive domains (cf. Massaro 1998 for an application of this distinction to the study of speech processing).

Comparative music–language research is still in its infancy. The ultimate promise of such work

is not simply to resolve debates over modularity, but to gain a deeper understanding of the mechanisms humans use to make sense out of sound. This requires doing justice to both the differences and the similarities between the domains.

## Acknowledgments

L. Robert Slevc, Lauren Stewart and Daniele Schön provided helpful comments. Supported by Neurosciences Research Foundation as part of its research programme on music and the brain at The Neurosciences Institute, where ADP is the Esther J. Burnham Senior Fellow.

## References

Anvari S, Trainor LJ, Woodside J and Levy BA (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology*, **83**, 111–130.

Ayotte J, Peretz I and Hyde K (2002). Congenital amusia: a group study of adults afflicted with a music-specific disorder. *Brain*, **125**, 238–251.

Bigand E and Poulin-Charronnat B (2006). Are we 'experienced listeners'? A review of the musical capacities that do not depend on formal musical training. *Cognition*, **100**, 100–130.

Brown S, Martinez MJ and Parsons LM (2006). Music and language side by side in the brain: a PET study of the generation of melodies and sentences. *European J Neuroscience*, **23**, 2791–2803.

Cuddy LL, Balkwill L-L, Peretz I and Holden RR (2005). A study of 'tone deafness' among university students. *Annals of the New York Academy of Sciences*, **1060**, 311–324.

Dilley L (2005). The phonetics and phonology of tonal systems. PhD dissertation, MIT.

Dowling WJ, Kwak S and Andrews MW (1995). The time course of recognition of novel melodies. *Perception and Psychophysics*, **57**, 136–149.

Drayna D, Manichaikul A, de Lange M, Snieder H and Spector T (2001). Genetic correlates of musical pitch recognition in humans. *Science*, **291**, 1969–1972.

Fedorenko E, Patel AD, Casasanto D, Winawer J and Gibson E (in press). Structural integration in language and music: Evidence for a shared system. Memory and Cognition.

Foxton JM, Dean JL, Gee R, Peretz I and Griffiths TD (2004). Characterisation of deficits in pitch perception underlying 'tone deafness'. *Brain*, **127**, 801–810.

Gibson E (2000). The dependency locality theory: a distance-based theory of linguistic complexity. In A Marantaz, Y Miyashita and WO'Neil, eds, *Image, language, brain*, 95–126. MIT Press, Cambridge, MA.

Griffiths TD, Stewart L, McDonald C and Kumar A (2007) A selective deficit in pitch-direction perception in congenital amusia. *Society for Neuroscience Abstracts*, November.

Gunter TC, Friederici AD and Schriefers H (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, **12**, 556–568.

Hyde K and Peretz I (2004). Brains that are out of tune but in time. *Psychological Science*, **15**, 356–360.

Hyde KL, Lerch JP, Zatorre RJ, Griffiths TD, Evans AC and Peretz I (2007). Cortical thickness in congenital amusia: when less is better than more. *Journal of Neuroscience*, **27**, 13028–13032.

Koelsch S and Mulder J (2002). Electric brain responses to inappropriate harmonies during listening to expressive music. *Clinical Neurophysiology*, **113**, 862–869.

Koelsch S and Siebel WA (2005). Toward a neural basis of music perception. *Trends in Cognitive Sciences*, **9**, 578–584.

Koelsch S, Gunter TC, Friederici AD and Schröger E (2000). Brain indices of music processing: 'non-musicians' are musical. *Journal of Cognitive Neuroscience*, **12**, 520–541.

Koelsch S, Gunter TC, Wittforth M and Sammler D (2005). Interaction between syntax processing in language and music: an ERP study. *Journal of Cognitive Neuroscience*, **17**, 1565–1577.

Koelsch S, Jentschke S, Sammler D and Mietchen D (2007). Untangling syntactic and sensory processing: an ERP study of music perception. *Psychophysiology*, **44**, 476–490.

Krumhansl CL (1990). *Cognitive foundations of musical pitch*. Oxford University Press, New York.

Lerdahl F and Jackendoff R (1983). *A generative theory of tonal music*. MIT Press, Cambridge, MA.

Lerdahl F and Krumhansl CL (2007). Modeling tonal tension. *Music Perception*, **24**, 329–366.

Lochy A, Hyde KL, Parisel S Van Hyfte S and Peretz I (April 2004). Discrimination of speech prosody in congenital amusia. Poster presented at the Cognitive Neuroscience Society, San Francisco.

Maess B, Koelsch S, Gunter T and Friederici AD (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience*, **4**, 540–545.

Magne C, Schön D and Besson M (2006). Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, **18**, 199–211.

Mandell J, Schulze K and Schlaug G (2007). Congenital amusia: an auditory-feedback disorder? *Restorative Neurology and Neuroscience*, **25**, 323–334.

Marin OSM and Perry DW (1999). Neurological aspects of music perception and performance. In D Deutsch, ed., *The psychology of music*, 2nd edn, 653–724. Academic Press, San Diego, CA.

Marques C, Moreno S, Castro SL and Besson M (2007). Musicians detect pitch violation in a foreign language

better than nonmusicians: behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience*, **19**, 1453–1463.

Massaro D (1998). *Perceiving talking faces.* MIT Press, Cambridge, MA.

McDermott J and Hauser MD (2005). The origins of music: innateness, development, and evolution. *Music Perception*, **23**, 29–59.

Moreno S and Besson M (2006). Musical training and language-related brain electrical activity in children. *Psychophysiology*, **43**, 287–291.

Musacchia G, Sams M, Skoe E and Kraus N (2007) Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences, USA*, **104**, 15894–15898.

Norton A, Winner E, Kronin K, Overy K, Lee DJ and Schlaug G (2005) Are there neural, cognitive, or motoric markers for musical ability prior to instrumental training? *Brain and Cognition*, **59**, 124–34.

Osterhout L and Holcomb PJ (1993). Event-related potential and syntactic anomaly: evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes*, **8**, 413–437.

Patel AD (2003). Language, music, syntax, and the brain. *Nature Neuroscience*, **6**, 674–681.

Patel AD (2008). *Music, language and the brain.* Oxford University Press, New York.

Patel AD (in press). Language, music, and the brain: a resource-sharing framework. In M Rohrmeir, P Rebuschat, J Hawkins and I Cross, eds, *Language and music as cognitive systems.* Oxford University Press, Oxford.

Patel AD and Iversen JR (2007). The linguistic benefits of musical abilities. *Trends in Cognitive Sciences*, **11**, 369–372.

Patel AD, Gibson E, Ratner J, Besson M and Holcomb P (1998). Processing syntactic relations in language and music: an event-related potential study. *Journal of Cognitive Neuroscience*, **10**, 717–733.

Patel AD, Peretz I, Tramo M and Labrecque R (1998). Processing prosodic and musical patterns: a neuropsychological investigation. *Brain and Language*, **61**, 123–144.

Patel AD, Wong M, Foxton J, Lochy A and Peretz I (2008). Speech intonation perception deficits in musical tone deafness (congenital amusia). *Music Perception*, **25**, 357–368.

Peretz I (1993). Auditory atonalia for melodies. *Cognitive Neuropsychology*, **10**, 21–56.

Peretz I (2006). The nature of music from a biological perspective. *Cognition*, **100**, 1–32.

Peretz I and Coltheart M (2003). Modularity of music processing. *Nature Neuroscience*, **6**, 688–691.

Peretz I and Zatorre RJ (2005). Brain organization for music processing. *Annual Review of Psychology*, **56**, 89–114.

Peretz I, Cummings S and Dubé M-P (2007) The genetics of congenital amusia (or tone-deafness): a family aggregation study. *American Journal of Human Genetics*, **81**, 582–588.

Schellenberg EG and Trehub SE (1999). Redundancy, conventionality, and the discrimination of tone sequences: a developmental perspective. *Journal of Experimental Child Psychology*, **74**, 107–127.

Schön D, Magne C and Besson M (2004). The music of speech: electrophysiological study of pitch perception in language and music. *Psychophysiology*, **41**, 341–349.

Slevc LR and Miyake A (2006) Individual differences in second language proficiency: does musical ability matter? *Psychological Science*, **17**, 675–681.

Slevc LR, Rosenberg JC and Patel AD (in press). Making psycholinguistics musical: Self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bulletin and Review.*

Steinbeis N and Koelsch S (2008). Shared neural resources between music and language indicate semantic processing of musical tension–resolution patterns. *Cerebral Cortex*, doi:10.1093/cercor/bhm149. **18**, 1169–1178.

Stewart L, Overeath T, Warren JD, Foxton JM and Griffiths TD (2008). fMRI evidence for a cortical hierarchy of pitch pattern processing. *PLoS ONE*, **3**(1), e1470. doi:10/1371/journal.pone.0001470.

Tallal P and Gaab N (2006). Dynamic auditory processing, musical experience and language development. *Trends in Neurosciences*, **29**, 382–370.

Thompson WF, Schellenberg EG and Husain G (2004). Decoding speech prosody: do music lessons help? *Emotion*, **4**, 46–64.

Tillmann B, Janata P and Bharucha JJ (2003). Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research*, **16**, 145–161.

Trehub SE, Bull D and Thorpe LA (1984). Infants' perception of melodies: The role of melodic contour. *Child Development*, **55**, 821–830.

Tyler L and Marslen-Wilson W (2008). Fronto-temporal brain systems supporting spoken language comprehension. Philosophical Transactions of the Royal Society, doi:10.1098/rstb/2007.2158. **363**, 1037–1054.

Wong PCM, Skoe E, Russo NM, Dees T and Kraus N (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, **10**, 420–422.

Zattore RJ, Evans AC and Meyer E (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience*, **14**, 1908–1919.

# PART 5

# Musical development

Edited by David J. Hargreaves

CHAPTER 20

# Prenatal development and the phylogeny and ontogeny of music

Richard Parncutt

## Introduction

Prenatal musical development involves the acquisition of perceptual, cognitive, motor and emotional abilities and information that may influence later musical development. I argue in this chapter that research in this area can shed light on the origins of musical behaviours, not only in human prehistory (phylogeny) but also within the lifespan of an individual (ontogeny).

Ontogenetically, the infant is surprisingly sensitive to patterns of sound and movement that adults perceive as musical (Chapter 21 this volume). The origin and evolutionary function (if any) of this sensitivity is unclear. One possibility is that musical patterns are similar to perceptual patterns to which the fetus is regularly exposed before birth: the fundamental frequency trajectory of the mother's voice, its relationship to breathing, and the rhythm of her heartbeat and footsteps (Parncutt 1989, 1993).

In spite of a recent renaissance of research on the phylogeny (origins) of music (summarized by Mithen 2005), no theory currently enjoys broad acceptance. One possibility involves classical conditioning of the fetus during the third trimester by passive exposure to sounds, movements and hormonal changes within the mother's body (Parncutt, in press). Since (changes in) maternal emotional states regularly trigger both (changes in) the patterns of sound and movement to which the fetus is exposed and (changes in) blood hormone levels, the fetus may associate these with each other, giving emotional connotations to patterns of sound and movement.

The importance of the prenatal phase for general psychological development was emphasized by Smotherman and Robinson (1990) and by Hopkins and Johnson (2005). The fetus responds actively to changes in its intrauterine environment in ways that influence its later behavioural and biological development. The prenatal phase may be regarded as a *developmental niche* (cf. Super and Harkness 1986) with its own unique characteristics but also lying on the same developmental continuum as postnatal niches. Although no-one would question that sensation, emotion, cognition and motor abilities have prenatal origins (Hall and Oppenheim 1987), most empirical research about these abilities in the late twentieth century ignored the prenatal period, while at the same time repeatedly asking nature–nurture questions that can only be answered by studying the prenatal epigenesis[1] of behaviour.

[1] The word 'epigenesis' (rather than genesis) refers to the role of both nature (genetics) and nuture (environment), and the interaction between them, in the prenatal emergence of behaviour.

It is problematic to use the term 'music' in conjunction with fetal development. Neither the fetus nor the infant discriminates between music (including singing) and speech; *motherese*[2] (Papousek 1996), the playful vocal-gestural interaction between infants and carers, is a mixture of both. Because motherese is not music but may represent music's origin, the term *protomusic* is appropriate (cf. Fitsch 2006)—by analogy to *protolanguage*, the common ancestor of a group of related languages. For similar reasons, it is misleading to speak of 'womb music'[3] or of 'playing music to cows' (Uetake *et al.* 1997), since the sound patterns presented in such studies are not perceived as music (i.e. not as culturally situated) by the experimental participants. Those patterns may nonetheless have musically interesting physiological, behavioural or cognitive affects.

This chapter focuses on musically relevant psychological aspects of prenatal development: the development of perception, cognition and emotion, the relationships between them, and the musical and musicological implications of those relationships. The chapter begins by surveying relevant fetal sensory abilities: hearing, the vestibular sense of balance and acceleration, and the proprioceptive sense of body orientation and movement. All those senses are relevant for musical development, since in all known cultures music is inseparable from bodily movement and gesture, whether real (Blacking 1995) or implied (Tolbert 2001). The chapter goes on to consider what sounds and other stimuli are available to the fetus: what patterns are the earliest to be perceptually learned? It then considers psychological and philosophical issues of fetal attention, 'consciousness', learning and memory. The chapter closes with speculations about the possible role of prenatal development in the phylogeny of musical behaviours.

[2] I use the term *motherese* rathar than *infant-directed speech* to emphasize the two-way nature of this form of communication (including the active role of the infant) and the musically central role of women (especially in discussions of the phylogeny of music). Of course men (not only fathers) and women other than mothers can and should also speak motherese.

[3] E.g. 'Babies remember womb music', http://news.bbc.co.uk/2/hi/health/1432495.stm, viewed on 11 August 2007.

# The ontogeny of music

## Fetal auditory, vestibular and proprioceptive abilities

All human sensory systems begin to function before birth (Hepper 1992). The acoustical stimulation to which the fetus is exposed is more diverse and carries more information relative to corresponding discriminatory abilities than visual, tactile, olfactory or gustatory (biochemical) stimulation. In that sense, hearing may be regarded as the dominant sensory modality in the prenatal phase, and infancy as a transition from auditory to visual dominance.

The fetus can hear throughout the second half of gestation. Because the fetal inner ear is filled with fluid, much of the sound heard by the fetus is transmitted to the inner ear through the skull by bone conduction (Gerhardt *et al.* 1996; Sohmer *et al.* 2001). The cochlea begins to process sounds at about around 20 weeks gestation; the cochlea reaches adult size at 25 weeks, but continues to develop until birth (Bibas *et al.* 2008). Motor responses to sound begin during the same period (Joseph 2000). At 20 weeks, the fetus is sensitive to a narrow band of spectral frequencies around 300 Hz (Hepper and Shahidullah 1994). As the fetus develops, its auditory abilities—including the perceptible range of spectral frequencies, the discrimination of frequencies and rapid sequences of events, and the storage and recognition of pitch–time patterns—gradually improve and may approach adult levels at birth (Joseph 2000), even though brain myelination is not complete until several years later.

It is physically impossible for the fetus to localize sound sources (Parncutt 2006). First, the fetal head does not cast an acoustic shadow at the relatively low frequencies that are transmitted into the amniotic fluid. Second, interaural time differences are smaller for the fetus than for an adult due to the smaller head size and the faster speed of sound in liquid. The fetus therefore has no access to information about the direction from which sounds emanate. Prenatal sound is monophonic and omnidirectional.

The vestibular system (or organ of balance) begins to function at about the same time as the cochlea. The cochlea and vestibular system

lie at opposite ends of the (bony/osseous) labyrinth in the temporal bone or inner ear. The vestibular system consists of three semicircular canals that are sensitive to angular accelerations (rotation), and the saccule and utricle which each contain otoliths (dense structures) that are sensitive to gravity and linear accelerations (horizontal movement in the utricle, vertical in the saccule). The membranous labyrinth grows quickly and attains adult size by the middle of the gestation period; the otic capsule ossifies between 18 and 24 weeks (Nemzek *et al.* 2000).

## Fetal motility and proprioception

An overview of fetal motor development was given by de Vries and Hopkins (2005). Fetal motility begins at about eight weeks' gestation, when the fetus begins to move through the amniotic fluid and to extend and flex its limbs. From then on, the movement repertoire of the fetus gradually expands. Some movements may be regarded as adaptations to the prenatal environment, while others prepare the fetus for postnatal life. In the third trimester, different sleep–wake states are associated with different movement repertoires and different sizes and frequencies of movement. Movements are not confined to the limbs, but also include startles, breathing movements, jaw openings, sucking, swallowing, facial, tongue and laryngeal movements.

As movements develop, so too does *motor control*. Motor control generally involves an interaction between the neural signals that control muscle activity and sensory feedback about the limb location and muscle tension. The spinal reflex arc begins to function during the first trimester, enabling motor signals to bypass the brain, speeding up motor reflexes.

Development of motor control occurs in parallel with development of *proprioception*—the sense of the relative position and motion of parts of the body. Since musical meaning involves the perception of gesture (Tolbert 2001), which in turn depends on proprioception, the prenatal development of proprioception may be relevant for the later emergence of musical abilities.

Gestures also play an important role in motherese (Trevarthen 1985). As they get older, infants increasingly understand the meaning of the physical and vocal gestures of their carers and learn to imitate them. It would be surprising if this striking postnatal development did not depend in some way on prenatal exposure to patterns of sound, movement and emotion.

## The fetal auditory environment

The fetal auditory environment is rich and varied, and provides many opportunities for prenatal perceptual learning. The fetus is exposed to sounds from both inside and outside the mother's body. Internal sounds include her voice, breathing, heartbeat, digestion, body movements, and footsteps (Lecanuet 1996). Of these, the voice is most often audible (Fifer and Moon 1994). The internal sounds tend to be louder than the external sounds: Richards *et al.* (1992) found the mother's voice to be about 8 dB louder to the fetus than the voices of her conversational partners. The fetus is also exposed to patterns of movement that are coupled to sound patterns, such as when the mother walks.

Both internal and external sounds are muffled (low-pass filtered) as they pass through the mother's body and amniotic fluid. Spectral components in the approximate range 100–1000 Hz are attenuated relatively little and may even be slightly amplified, even if their original is external (Richards *et al.* 1992). When the fetus is exposed to speech, either internally from the mother or from an external source, muffling makes vowels more salient (audible) than consonants and the fundamental frequency contour more salient than spectral information (timbre, phonemes) (cf. Smith *et al.* 2003)—consistent with the important role of pitch contour in music perception (Dowling and Fujitani 1971).

## Fetal 'consciousness'

The fetus is not conscious in the everyday sense of *reflective* consciousness, that is, the ability to reflect upon perceptual experience. It does however have a range of abilities that may be considered part of a broad definition of consciousness, including perception, cognition, and emotion; sleep–wake cycles (Nijhuis *et al.* 1982); preferences; and attention. Newborns are capable of demonstrating preferences by the rate of sucking on a pacifier (DeCasper and Fifer 1980),

suggesting that the fetus could also demonstrate preferences if a suitable empirical method could be developed. Kisilevsky *et al.* (2004) investigated the heart rate and movement of fetuses in response to a musical stimulus and observed a change at about 33 weeks, suggesting an ontogeny of attention.

According to the *levels of consciousness* approach (Zelazo 2003), adult reflective consciousness comprises separable components that are acquired at specific ages. The ability to label objects (including people) is acquired at the age of about 1 year, to distinguish self and from others at 2, to reflect about ideas or theories and apply rules at 3, and to reflect about self and others at 4. Since neither the newborn nor the fetus has any of these abilities, we may assert that the fetus is not conscious in the everyday sense. Attempts to ascribe consciousness to the fetus may be regarded as projections of adult reflective consciousness—animistic projection, anthropomorphism (Parncutt 2006).

## Prenatal learning

Learning may be defined as acquisition of information that affects later behaviour. Fetal learning has been repeatedly demonstrated by the empirical paradigms of *habituation* and *classical conditioning*.

In habituation, an organism is exposed several times to the same stimulus and gradually stops reacting to it; in everyday language we would say the organism gets used to the stimulus or the stimulus becomes uninteresting for the organism. Leader *et al.* (1982) observed fetal habituation to a repeated vibrotactile stimulus at 22–30 weeks gestation. In a different habituation paradigm, Shahidullah and Hepper (1994) found that the fetus can discriminate between sounds at 35 weeks better than at 27 weeks.

Other empirical studies have demonstrated that the fetus is capable of associating stimuli with each other by classical conditioning. In that paradigm, stimuli are paired with each other in a temporal sequence, and after several such pairings the organism begins to expect the second stimulus as soon as the first is presented.

Classical conditioning may be regarded as the basic mechanism underlying *statistical learning*. All higher organisms are sensitive to statistical properties of their environment: given many occurrences of event A in different contexts, they learn the probability that it will be accompanied by (or co-occur with) event B, that is, the probability that event A predicts event B: conditional probability (Fiser and Aslin 2002). Saffran *et al.* (1996) demonstrated that 8-month-old infants can learn statistical properties of nonsense speech during only two minutes' exposure. Since statistical learning is such a fundamental means of picking up information about the environment for all organisms, we may safely assume that humans begin to learn statistically before birth.

## Transnatal memory

Transnatal memory is postnatal retention of prenatally acquired information. If the above arguments concerning fetal (lack of) consciousness are correct, transnatal memory is always implicit, that is, not under any form of conscious control—unlike an adult's memory for a telephone number, which normally requires conscious effort.

Transnatal memory for stimulus patterns presented repeatedly to different fetal sensory modalites can last for weeks or months and may therefore be considered to be a form of long-term memory (Hepper 1991, 1992; Hopkins and Johnson 2005; Mastropieri and Turkewitz 2001). In experiments to measure the duration of transnatal memory, a given stimulus pattern is presented repeatedly before birth (e.g. a specific piece of music), or a pattern to which the fetus is naturally exposed (e.g. the mother's voice) is used as an experimental stimulus. When the same pattern is presented for the first time after birth, or in a specific new way, the reaction of the infant is observed and compared with its reaction to unfamiliar control stimuli.

*Episodic* memory involves memory for a specific event. The existence of *transnatal episodic memory* is suggested by reports of therapeutic patients in trance states who seem able to recall prenatal events or the experience of their own birth. Hartogh (2003) questioned the validity of such evidence. In general, memories are affected by relevant knowledge and by cultural and social context, and can easily be constructed (Harris et al. 1989). Since episodic memory in humans

normally depends on language, and since prenatal episodic memory has no known evolutionary function, prenatal episodic memory is unlikely to exist. That does not necessarily prevent prenatal regression therapy from having a useful therapeutic function.

## The phylogeny of musical behaviour

Recent research about the origins of music (summarized by Mithen 2005) has considered many scenarios. One involves prenatal exposure to the internal sounds and movements of the mother's body, which are associated with her changing physical and emotional state (Parncutt in press). The claim is that the associations between sound, movement and emotion upon which music is based are ultimately of prenatal origin.

### Emotion

Strong emotions are generally associated with reflexes, instincts or drives that promote survival and reproduction (Tinbergen 1989), such as hunger, sex, fear, pain, disgust, jealousy, surprise, and anger. Music is exceptional in that it communicates strong emotions that are marked by changed states of consciousness and spiritual experiences (Gabrielsson and Lindström Wik 2003), although it has no clear survival value. A plausible theory of the origins of music should be consistent with this fundamental contradiction.

It is often supposed that music is emotional because it is associated with social behaviour and group survival (e.g., Dean and Bailes 2006). However, evolutionary theory primarily explains the behaviour of isolated pairs of individuals (e.g. reciprocal altruism) rather than groups considered as a whole (Boyd and Richerson 1988; Trivers 1971), and such an approach cannot easily explain music's *strong* emotionality and spirituality.

Musical emotions are associated with learned patterns of sound and movement. A possible source of those associations is motherese (Dissanayake 2000), whose vocabulary of gestural meanings evidently includes a universal component (Kuhl *et al.* 1997; Papousek 1996). The question arises as to the origin of those gestural meanings.

### Infant musical skills: innate or learned?

Another surprising thing about music is that all humans, including infants, possess basic musical skills. A plausible theory of the origins of music should be consistent with the documented musical abilities of infants. Trehub and Hannon (2006) proposed that

> infants' music perception skills are a product of general perceptual mechanisms that are neither music- nor species-specific. Along with general-purpose mechanisms for the perceptual foundations of music, we suggest unique motivational mechanisms that can account for the perpetuation of musical behavior in all human societies.
>
> Trehub and Hannon (2006, Abstract)

Prenatal associations between patterns of sound, movement and emotion could underlie such 'general perceptual mechanisms' and 'unique motivational mechanisms'. If so, they may explain the nature and origin of musical emotion.

Several sources of evidence suggest that the musical abilities of infants are at least partially inborn. Amusia can be inborn or genetically determined in a small proportion of the population (Ayotte *et al.* 2002). Children of deaf parents prefer infant-directed singing over adult-directed singing, suggesting an innate component—or at least an inborn preference for exaggeration (Masataka 1999). The development of musical ability is affected by prenatal testosterone levels (Sluming and Manning 2000).

Models of the interplay between genes and environments (Garcia Coll *et al.* 2004) suggest that behaviours and skills such as musicality are neither inborn nor innate, but a mixture of the two. Several empirical studies cited in this chapter are consistent with the prenatal learning of protomusical skills. For example, newborns respond to the emotional content of speech, but only in their maternal language, suggesting that their ability to recognize that emotion was acquired before birth (Mastropieri and Turkewitz 2001). The claim that musical skills

are *largely* learned may be valid throughout the lifespan. In early life, the gestural vocabulary of motherese may be largely learned from prenatal exposure to the internal sounds of the mother's body. Later, musical expertise may depend primarily on the total accumulated duration of practice (Howe *et al.* 1998).

## Fetal–maternal communication

Bonding (secure attachment) between primary caregiver and infant plays an important psychological and physiological role in early development (Schore 2001). The idea that maternal–infant bonding is an evolutionary adaptation is consistent with high rates of infant mortality among both non-human primate and human hunter-gatherer populations (>50%: Denham 1974). Maternal–infant bonding generally increases the chance of infant survival, but not necessarily of surviving infanticide or of preventing abandonment when chances of survival are particularly low (Hausfater 1984). Prenatally acquired knowledge about maternal emotional states may promote postnatal bonding and survival by helping the infant to communicate its needs appropriately (cf. Broad *et al.* 2006) and may in that sense be adaptive. Other factors being equal, the chance that an infant will survive to reproductive age will increase if infant demands on the mother or other carers do not radically exceed their momentary capabilities or resources.

Infant-mother bonding is two-way and reciprocal (Lee 2006): each party is at some level sensitive to the physical and emotional state of the other. Empirical research is beginning to document the infant's active perception of the mother: compare Stern's (2002) cognitive, psychoanalytic approach with Trevarthen's (1980) concept of intersubjective communication. To successfully monitor the mother's physical and emotional state, the infant must have prior knowledge about the relationship between maternal state and behaviour. The fetus has constant access to two reliable sources of information about the physical and emotional state of its mother: behavioural (sound and movement) and biochemical (blood hormone concentrations).

Regarding behavioural information, all patterns of sound and movement that are audible within the body in everyday situations, including vocalization, respiration, circulation, movement, footsteps and digestion, depend on physical and emotional state (Mastropieri and Turkewitz 2001). The human fetus has access to three behavioural sources of information about maternal state: sound patterns, linear and rotational acceleration of the fetal body, and relative movement of the fetal limbs. These are perceived by the fetal auditory, vestibular and proprioceptive systems respectively.

Regarding biochemical information, the hormones involved in the maternal–fetal interaction arise from three different sources: the placenta, maternal organs and fetal organs (Power and Schulkin 2005). The placenta and fetal membranes produce a large number of steroids that regulate and balance both maternal and fetal physiology. They include progesterone and oestrogen, which play a role in maintenance of pregnancy and support of the embryo/fetus (Albrecht *et al.* 2000); testosterone, which affects fetal development (Matt and MacDonald 1984); oestrogen, related to female secondary sex characteristics (Nelson and Bulun 2001); corticotropin-releasing hormone, which influences the duration of pregnancy (Hillhouse and Grammatopoulos 2002); relaxin, which facilitates birth (Klonisch *et al.* 1999); and placental lactogen (somatomammotropin), which influences nutrient (carbohydrate, lipid) levels in the maternal blood (Walker *et al.* 1991). In an evolutionary approach, hormonal manipulation of maternal nutrient supply represents an early stage of parent–offspring conflict (Wells 2003).

Hormone production in the mother involves a mixture of regular adult hormonal processes and processes specific to pregnancy. Regarding the former, externally caused stressors (flight-fight reaction) lead to stimulation of the adrenal gland, which secretes adrenaline (epinephrine) and noradrenaline (norepinephrine) into the blood, which in turn increases blood sugar, muscle tension, and blood pressure. Stress also causes the release of corticotropin-releasing hormone (CRH) and the production of glucocorticoids that affect immune responses (Elenkov *et al.* 1999). CRH is also associated with anxiety and depression (Steckler and Holsboer 1999), and cortisol is associated with fear and stress (Kalin *et al.* 1998). Oxytocin is

more specific to reproduction and is associated with orgasm, birth and breastfeeding (Newton 1978).

That the fetus is sensitive to hormone concentrations in the maternal blood is clear from studies that demonstrate the effect of those concentrations during pregnancy on postnatal development. For example, postpartum concentrations of testosterone, estradiol, androstenedione and cortisol correlate with the children's later levels of physical aggression (Susman *et al.* 2001). Molecules that pass the blood–brain barrier include oxygen, carbon dioxide, alcohol, and steroid hormones, of which there are five main groups: progestagens, glucocorticoids, mineralocorticoids, androgens, and oestrogens (Pardridge and Mietus 1979). The permeability of the blood–brain barrier to steroid hormones depends on the molecule and involves different temporal delays (Zloković *et al.* 1988). The placenta is permeable to nutrients, oxygen, alcohol, antibodies and steroid hormones with different temporal delays of seconds, minutes or hours (Bajoria *et al.* 1998; Bajoria and Fisk 1998).

Thus, both evolutionary–biological and developmental–psychological approaches predict that the fetus perceives changes in patterns of sound and movement within the mother's body (behavioural information) following everyday changes in maternal state. The corresponding changes in maternal hormone levels are delayed by passage through the placental and fetal blood–brain barriers (biochemical information). The behavioural change thus predicts the biochemical change, allowing classical conditioning to occur—just as, in Pavlov's famous experiment, the footsteps of a master bringing food to a dog predict the appearance of the food.

Any stimulus can be associated with any other by classical conditioning, and all animals are capable of classical conditioning. The human fetus is no exception (Smotherman and Robinson 1990). The fetus must therefore learn to associate the behavioural information described above with the corresponding biochemical information. The theory of classical conditioning predicts that after several repetitions of such a sequence the fetus will begin to respond emotionally to the behavioural patterns—before the arrival of the biochemical information. While the time interval between behavioural and biochemical information in this scenario is presumably of the order of minutes, classical conditioning may also occur for longer interstimulus intervals of the order of hours (Garcia *et al.* 1974).

## Prenatal influences on postnatal behaviour

Prenatally established associations can influence postnatal behaviours and the development of musical culture in the context of *motherese*, *play* and *ritual*. In all three cases, *operant conditioning* (Skinner 1938) may be the underlying mechanism. Patterns of sound and movement that occur by accident in these behaviours may be similar to patterns of sound and movement that were prenatally linked to emotion. The triggering of associated emotions may reinforce the actions or behaviours that produced the patterns of sound, increasing their frequency of occurrence. Since motherese, play and ritual are social activities whose participants were subject to similiar prenatal conditioning processes, this theory predicts that such behaviours and associated emotions will generally be socially shared, enabling the development of music as a form of social behaviour.

On this basis, we might expect to find associations between sound, movement and emotion in all prenatally hearing animals. The reason why non-human animals are not musical in the human sense is presumably that only humans possess reflective consciousness, which emerged at least 60 000 years ago and enabled a cultural explosion (Mithen 2005). Reflective consciousness may be regarded as a co-requisite for the ability to conceptualize the past and future in relation to the present, which enables deliberate/intentional planning and action (Noble and Davidson 1996). According to this view, prenatal associations between sound, movement and emotion became 'music' when humans acquired the ability to deliberately manipulate shared emotions using sound—that is, to perform.

# Conclusions

Recent research in music psychology has tended to avoid questions of prenatal development due to the practical and ethical difficulties associated

with empirical investigations and the dubious quality of much of the available research literature. This period of restraint may be coming to an end as researchers realize that the prenatal period could be a source of answers to central questions in music psychology, and as developing empirical technologies make it increasingly possible to carry out methodologically sound empirical investigations.

The relevant empirical literature is expanding rapidly, but it is spread over many different disciplines, and many central issues remain to be critically addressed by independent research groups. The present theory presented on the origins of music may be internally consistent and logical, but further empirical work will be necessary to investigate the details of the hypothesized prenatal associations and their effect on postnatal behaviour.

Since our knowledge of music's emergence in the context of ritual will always be limited, any theory of music's phylogeny will always be speculative. The ontological question of how prenatal learning affects postnatal musical development will become increasingly accessible to empirical investigation, as non-invasive observational techniques improve. An improved understanding of the interaction between genes and environment in psychological development (e.g. Bakshi and Kalin 2000), combined with new approaches to fetal behaviour (such as preferences) and prenatal influences on postnatal behaviour (including transnatal memory), will lead to new insights that will confirm, complement, challenge or overthrow the ideas presented in this chapter.

## References


Albrecht ED, Aberdeen GW and Pepe G J (2000). The role of estrogen in the maintenance of primate pregnancy. *American Journal of Obstetrics and Gynecology*, **182**, 432–438.

Ayotte J, Peretz I and Hyde K (2002). Congenital amusia: a group study of adults afflicted with a music-specific disorder. *Brain*, **125**, 238–251.

Bajoria R and Fisk NM (1998). Permeability of human placenta and fetal membranes to thyrotropin-stimulating hormone in vitro. *Pediatric Research*, **43**, 621–628.

Bajoria R, Peek MJ and Fisk NM (1998). Maternal-to-fetal transfer of thyrotropin-releasing hormone in vivo. *American Journal of Obstetrics and Gynecology*, **178**, 264–269.

Bakshi VP and Kalin NH (2000). Corticotropin-releasing hormone and animal models of anxiety: gene–environment interactions. *Biological Psychiatry*, **48**, 1175–1198.

Bibas AB, Xenellis J, Michaels L, Anagnostopoulou S, Ferekidis E and Wright A (2008). Temporal bone study of development of the organ of Corti: correlation between auditory function and anatomical structure. *Journal of Laryngology and Otology*, http://journals.cambridge.org/.

Blacking J (1995). *Music, culture and experience: selected papers of John Blacking* (ed. R Byron). University of Chicago Press, London.

Boyd R and Richerson PJ (1988). The evolution of reciprocity in sizable groups. *Journal of Theoretical Biology*, **132**, 337–356.

Broad KD, Curley JP and Keverne EB (2006). Mother–infant bonding and the evolution of mammalian social relationships. *Philosophical Transactions of the Royal Society London: B Biological Sciences*, **361**(1476), 2199–2214.

de Vries JIP and Hopkins B (2005). Fetal movements and postures: what do they mean for postnatal development? In B Hopkins and SP Johnson, eds, *Prenatal development of postnatal functions*, 177–220. Greenwood, Westport, CT.

Dean RT and Bailes F (2006). Toward a sociobiology of music. *Music Perception*, **24**, 83–84.

DeCasper AJ and Fifer WP (1980). Of human bonding: newborns prefer their mothers' voices. *Science*, **208**, 1174–1176.

Denham WW (1974). Population structure, infant transport and infanticide among Pleistocene and modern hunter-gatherers. *Journal of Anthropological Research*, **30**, 191–198.

Dissanayake E. (2000). *Art and intimacy: how the arts began*. University of Washington Press, Seattle, WA.

Dowling WJ and Fujitani DS (1971). Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, **49**/2B, 524–531.

Elenkov IJ, Webster EL, Torpy DJ and Chrousos GP (1999). Stress, corticotropin-releasing hormone, glucocorticoids and the immune/inflammatory response: acute and chronic effects. *Annals of the New York Academy of Sciences*, **876**, 1–13.

Fifer WP and Moon CM (1994). The role of the mother's voice in the organization of brain function in the newborn. *Acta Paediatrica Supplement*, **397**, 86–93.

Fiser J and Aslin RN (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences*, **99**(24), 15822–15826.

Fitsch WT (2006). The biology and evolution of music: a comparative perspective. *Cognition*, **100**, 173–215.

Gabrielsson A and Lindström Wik S (2003). Strong experiences related to music: a descriptive system. *Musicae Scientiae*, **7**, 157–217.

Garcia J, Hankins WG and Rusiniak KW (1974). Behavioral regulation of the milieu interne in man and rat. *Science*, **185**(4154), 824–831.

Garcia Coll C, Bearer EL and Lerner RM (eds) (2004). *Nature and nurture: the complex interplay of genetic and environmental influences on human behavior and development*. Erlbaum, Mahwah, NJ.

Gerhardt KJ, Huang X, Arrington KE, Meixner K, Abrams RM and Antonelli PJ (1996). Fetal sheep in utero hear through bone conduction. *American Journal of Otolaryngology*, **17**(6), 374–379.

Hall WG and Oppenheim RW (1987). Developmental psychobiology: prenatal, perinatal, and early postnatal aspects of behavioral development. *Annual Review of Psychology*, **38**, 91–128.

Harris RJ, Sardarpoor-Bascom F and Meyer T (1989). The role of cultural knowledge in distorting recall for stories. *Bulletin of the Psychonomic Society*, **27**, 9–10.

Hartogh T (2003). Prä- und perinatale Erinnerungen und ihr musiktherapeutischer Zugang: Eine kritische Analyse. *Musik-, Tanz- und Kunsttherapie*, **14**(4), 167–176.

Hausfater G (1984). Infanticide: comparative and evolutionary perspectives. *Current Anthropology*, **25**(4), 500–502.

Hepper PG (1991). An examination of fetal learning before and after birth. *Irish Journal of Psychology*, **12**, 95–107.

Hepper PG (1992). Fetal psychology: an embryonic science. In JG Nijhuis, ed., *Fetal behaviour*, 129–156. Oxford University Press, Oxford.

Hepper PG and Shahidullah BS (1994). Development of fetal hearing. *Archives of Disease in Childhood*, **71**, F81–F87.

Hillhouse EW and Grammatopoulos DK (2002). Role of stress peptides during human pregnancy and labour. *Reproduction*, **124**, 323–329.

Hopkins B and Johnson SP (2005). Prenatal development of postnatal functions. Greenwood, Westport, CT.

Howe MJA, Davidson JW and Sloboda JA (1998). Innate talents: reality or myth? *Behavioral and Brain Sciences*, **21**, 339–407.

Joseph R (2000). Fetal brain behavior and cognitive development. *Developmental Review*, **20**, 81–98.

Kalin NH, Larson C, Shelton SE and Davidson RJ (1998). Asymmetric frontal brain activity, cortisol, and behavior associated with fearful temperaments in Rhesus monkeys. *Behavioral Neuroscience*, **112**, 286–292.

Kisilevsky BS, Hains SM J, Jacquet A-Y, Granier-Deferre C and Lecanuet JP (2004). Maturation of fetal responses to music. *Developmental Science*, **7**, 550–559.

Klonisch T, Hombach-Klonisch S, Froehlich C, Kauffold J, Steger K, Huppertz B and Fischer B (1999). Nucleic acid sequence of feline preprorelaxin and its localization within the feline placenta. *Biology of Reproduction*, **60**, 305–311.

Kuhl PK, Andruski JE, Chistovich IA, Chistovich LA, Kozhevnikova EV, Ryskina VL *et al.* (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, **277**, 684–686.

Leader LR, Baillie P, Martin B and Vermeulen E (1982). The assessment and significance of habituation to a repeated stimulus by the human fetus. *Early Human Development*, **7**, 211–219.

Lecanuet J-P (1996). Prenatal auditory experience. In I Deliège, JA Sloboda, eds, *Musical beginnings*, 3–34. Oxford University Press, Oxford.

Lee SY (2006). A journey to a close, secure, and synchronous relationship: infant–caregiver relationship development in a childcare context. *Journal of Early Childhood Research*, **4**, 133–151.

Masataka N (1999). Preference for infant-directed singing in 2-day-old hearing infants of deaf parents. *Developmental Psychology*, **35**, 1001–1005.

Mastropieri D and Turkewitz G (2001). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental Psychobiology*, **35**, 204–214.

Matt DW and MacDonald GJ (1984). In vitro progesterone and testosterone production by the rat placenta during pregnancy. *Endocrinology*, **115**, 741–747.

Mithen S (2005). *The singing Neanderthals: the origins of music, language, mind and body*. Weidenfeld and Nicholson, London.

Nelson LR and Bulun SE (2001). Estrogen production and action. *Journal of the American Academy of Dermatology*, **45**(3 Suppl), S116–124.

Nemzek WR, Brodie HA, Hecht ST, Chong BW, Babcook CJ and Seibert JA (2000). MR, CT, and plain film imaging of the developing skull base in fetal specimens. *American Journal of Neuroradiology*, **21**, 1699–1706.

Newton N (1978). The role of the oxytocin reflexes in three interpersonal reproductive acts: coitus, birth and breastfeeding. In L Carenza, L Zichella and P Pancheri, eds, *Clinical psychoneuroendocrinology in reproduction*, 411–418. Academic Press, New York.

Nijhuis JG, Prechtl HFR, Martin CB jr and Bots RSGM (1982) Are there behavioural states in the human fetus? *Early Human Development*, **6**, 177–195.

Noble W and Davidson I (1996). *Human evolution, language and mind*. Cambridge University Press, Cambridge.

Papousek M (1996). Intuitive parenting: a hidden source of musical stimulation in infancy. In I Deliège and J Sloboda, eds, *Musical beginnings*, 88–112. Oxford University Press, Oxford.

Pardridge WM and Mietus LJ (1979). Transport of steroid hormones through the rat blood–brain barrier. *Journal of Clinical Investigation*, **64**, 145–154.

Parncutt R (1989). *Harmony: a psychoacoustical approach*. Springer-Verlag, Berlin.

Parncutt R (1993). Prenatal experience and the origins of music. In T Blum, ed., *Prenatal perception, learning, and bonding*, 253–277. Leonardo, Berlin.

Parncutt R (2006). Prenatal development. In GE McPherson, ed., *The child as musician*, 1–31. Oxford University Press, Oxford.

Parncutt R (in press). Prenatal conditioning, the mother schema, and the origins of music and religion. *Musicae*

*Scientiae* (Special issue on Music and Evolution, Ed. O. Vitouch).

Power ML and Schulkin J (2005). Birth, distress and disease: placenta–brain interactions. Cambridge University Press, Cambridge.

Richards DS, Frentzen B, Gerhardt KJ, McCann ME and Abrams RM (1992). Sound levels in the human uterus. *Obstetrics and Gynecology*, **80**, 186–190.

Saffran JR, Aslin RN and Newport EL (1996). Statistical learning by 8-month-old infants. *Science*, **274**(5294), 1926–1928.

Schore AN (2001). Effects of a secure attachment relationship on right brain development, affect regulation, and infant mental health. *Infant Mental Health Journal*, **22**, 7–66.

Shahidullah S and Hepper PG (1994). Frequency discrimination by the fetus. *Early Human Development*, **36**, 13–26.

Skinner BF (1938). *The behavior of organisms*. Appleton-Century-Crofts, New York.

Sluming VA and Manning JT (2000). Second to fourth digit ratio in elite musicians: evidence for musical ability as an honest signal of male fitness. *Evolution and Human Behavior*, **21**, 1–9.

Smith SL, Gerhardt KJ, Griffiths SK, Huang X and Abrams RM. (2003). Intelligibility of sentences recorded from the uterus of a pregnant ewe and from the fetal inner ear. *Audiology and Neuro-Otology*, **8**, 347–353.

Smotherman WP and Robinson SR (1990). The prenatal origins of behavioral organization. *Psychological Science*, **1**, 97–106.

Sohmer H, Perez R, Sichel J-Y, Priner R and Freeman S (2001). The pathway enabling external sounds to reach and excite the fetal inner ear. *Audiology and Neurotology*, **6**, 109–116.

Steckler T and Holsboer F (1999). Corticotropin-releasing hormone receptor subtypes and emotion—suppression of pituitary ACTH release and peripheral inflammation. *Biological Psychiatry*, **46**, 1480–1508.

Stern DN (2002). *The first relationship: infant and mother*. Harvard University Press, Cambridge, MA.

Super CM and Harkness S (1986). The developmental niche: a conceptualization at the interface of child and culture. *International Journal of Behavioral Development*, **9**, 545–569.

Susman EJ, Schmeelk KH, Ponirakis A and Gariepy JL (2001). Maternal prenatal, postpartum, and concurrent stressors and temperament in 3-year-olds: a person and variable analysis. *Development and Psychopathology*, **13**, 629–652.

Tinbergen N (1989). *The study of instinct*. Oxford University Press, Oxford.

Tolbert E (2001). Music and meaning: an evolutionary story. *Psychology of Music*, **29**, 84–94.

Trehub SE and Hannon EE (2006). Infant music perception: domain-general or domain-specific mechanisms? *Cognition*, **100**, 73–99.

Trevarthen C (1980). The foundations of intersubjectivity: development of interpersonal and cooperative understanding in infants. In D Olson, ed., *The social foundations of language and thought*, 316–342. Norton, New York.

Trevarthen C (1985). Facial expressions in mother–infant interaction. *Human Neurobiology*, **4**, 21–32.

Trivers RL (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, **46**, 35–57.

Uetake K, Hurnik JF and Johnson L (1997). Effect of music on voluntary approach of dairy cows to an automatic milking system. *Applied Animal Behaviour Science*, **53**, 175–182.

Walker WH, Fitzpatrick SL, Barrera-Saldana HA, Resendez-Perez D and Saunders GF (1991). The human placental lactogen genes: structure, function, evolution and transcriptional regulation. *Endocrine Reviews*, **12**, 316–328.

Wells JCK (2003). Parent–offspring conflict theory, signaling of need, and weight gain in early life. *Quarterly Review of Biology*, **78**, 169–202.

Zelazo PD (2003). The development of conscious control in childhood. *Trends in Cognitive Sciences*, **8**, 12–17.

Zloković BV, Lipovac MN, Begley DJ, Davson H and Rakić L (1988). Slow penetration of thyrotropin-releasing hormone across the blood–brain barrier of an in situ perfused guinea pig brain. *Journal of Neurochemistry*, **51**, 252–257.

# CHAPTER 21

# Music lessons from infants

Sandra E. Trehub

WHAT can we learn about music and about musicality from infants? Sceptics may question the possibility of deriving fruitful answers to such questions from immature beings whose hearing is deficient (relative to adults) and whose exposure to 'good' music, even conventional music, is limited. This chapter is aimed at less sceptical readers—those who are open to the possibility of *nature* making some contribution to our musical beginnings and to our subsequent development. It is organized as a series of lessons that emerge from empirical research over the past few decades.

## Their world is alive with the sound of music

Preverbal infants hear a steady stream of incomprehensible but highly melodious speech. This speech, known variously as *motherese, parentese*, or infant-directed speech, is characterized by exaggerated prosody, including elevated pitch, expanded pitch contours, large dynamic range, and rhythmic regularity (Fernald 1991). The most striking aspects of maternal speech—its effusiveness and positive affective quality—are also the most difficult to quantify. At periodic breaks in the flow of mellifluous speech, infants sigh, yawn, gurgle or coo, all of which function as contributions to the dyadic conversation.

Research to date has largely emphasized the common melodic and rhythmic features in speech to infants, even across cultures (Fernald *et al.* 1989). It has become clear, however, that unique or person-specific musical features are also evident in speech to infants. In one study (Bergeson and Trehub 2007), sequences of intervals, or tunes, were transcribed from recordings of several mothers as they interacted with their infants (4–7 months of age) on different occasions. The observed intervals were unrelated to conventional musical intervals, but highly trained listeners agreed on the transcriptions. Mothers used a handful of tunes consistently, usually with different content, on different occasions. Moreover, each mother's tunes were distinct from those of every other mother, justifying their designation as *signature tunes* (Bergeson and Trehub 2007). It seems, then, that infants receive regular exposure to specific speech melodies in the context of interactions with their primary caregiver.

Aside from the musical speech that infants hear, they also hear a good deal of singing (Trehub *et al.* 1997). Mothers (and many fathers and grandparents) the world over sing to infants in the course of providing care (Trehub and Trainor 1998). For the most part, they use a distinct musical genre for this purpose—lullabies and play songs. Although music for adults often differs substantially from one culture to another, music for infants has many cross-cultural similarities. For example, lullabies from foreign musical cultures are readily recognizable as such (Trehub *et al.* 1992), perhaps by virtue of common features such as simplicity, repetitiveness, and falling pitch contours (Unyk *et al.* 1993).

Even more distinctive than the musical material for infants is the manner of singing to infants. Infant-directed singing is generally part of intimate interactions whose goals are playful,

soothing, or emotion-sharing. Although songs have prescribed pitch intervals and rhythms, performances for infants share some features with maternal speech, such as elevated pitch, slurred articulation of words, and positive vocal tone (Trainor *et al.* 1997).

Casual listeners readily distinguish infant-directed from non-infant-directed versions of the same song by the same singer (Trehub *et al.* 1993b, 1997). Surprisingly, each mother seems to perform the same songs at nearly identical pitch level and tempo on different occasions (Bergeson and Trehub 2002), qualifying, perhaps, as *signature performances.*

## Maternal performances are irresistible

Presumably, the maternal style of speech and singing would not have persisted across cultures and historical periods if it did not achieve its intended goals, for example, soothing or accelerating sleep in the case of lullabies and joyful engagement in the case of play songs. Over the past couple of decades, investigators have sought infants' perspective on the sounds they hear. To do so, they devised procedures for ascertaining whether infants prefer one type of speech or music to another. The essence of these procedures is to present contrasting material that is linked to infants' gaze toward one of two loudspeakers. For example, infants' gaze at one loudspeaker results in the presentation of one sound pattern, which continues to play until they look away. Their gaze at the other loudspeaker results in the presentation of a contrasting sound pattern until infants look away. This procedure continues for a specified period, after which total gaze duration for each selection is compared. If infants look significantly longer to hear one of the selections, the presumption is that they preferred it to the other. In this manner, one can estimate their relative interest in and attention to the material. In some cases, investigators videotape the sessions to document observable affective responses.

Procedures such as these have revealed greater infant attention and positive affect to melodious motherese than to conventional adult speech (Cooper and Aslin 1990; Fernald 1985; Werker and McLeod 1989). In principle, familiarity with the maternal style could play a role in these preferences. Despite the popular notion that familiarity breeds contempt, the familiarity of music enhances its appeal (Szpunar *et al.* 2004).

Infants also prefer to listen to infant-directed than to non-infant-directed versions of the same song by the same singer (Trainor 1996). Evidence that 2-day-old infants with limited exposure to music or speech exhibit the same preference (Masataka 1999) rules out experiential factors. Prenatal exposure would have been minimal or absent for these particular infants because their deaf parents communicated by means of sign language. The findings are consistent with a *natural* or inborn preference for the maternal singing style and, in all likelihood, for the maternal speech style as well.

Young infants are poor at regulating their own state or arousal, so fussing and crying are frequent occurrences. At such times, caregivers intervene with non-verbal sounds (e.g., shush), singing, rocking, jiggling, feeding, or some combination of these. Singing is particularly effective at arresting, even preventing, such fussiness. Indeed, the effects of singing are evident in hormonal changes that reflect alterations in arousal or stress (Shenfield *et al.* 2003). Unfortunately, uninformed media reports and deceptive advertising—none of it based on solid scientific evidence—have convinced many parents that Mozart recordings are superior to their own humble performances. Such parents know that their songs 'work' in the sense of bringing comfort and joy to their offspring, but they worry that their performances lack the nutrients for the developing brain that are present in 'good' music.

## More than meets the ear

The prevalence of recorded music makes it seem that music is a mere auditory phenomenon. For most of history, however, and for many remote communities in developing nations, music is commonly experienced through live performances in which the listener plays an active role. In such contexts, the performer's visual gestures are as critical as the sounds produced, and the audience typically moves in time with the music. Visual gestures (head and body motion), touch (e.g., holding, stroking), and movement (rocking,

swaying) are also integral to mothers' sung and spoken performances. There are suggestions that motherese and maternal singing evolved as ways of keeping infants content when mothers' eyes and hands were required for food gathering or other work (Falk 2004). One prediction arising from this perspective is that mothers should intensify their vocal expressiveness when expressive means in other modalities are unavailable. In fact, there is evidence that mothers use a higher pitch level and wider pitch range (i.e., prosodic enhancement) when they are unable to touch their infants (Nakata and Trehub in preparation).

The consequences of hearing *and* seeing recordings of their mother talking or singing are potent (Nakata and Trehub 2004). Under these circumstances, infants become quite still and stare at their mother's image for extended periods. Interestingly, they remain engaged significantly longer during bimodal singing episodes than during comparable speech episodes.

Infants and adults often integrate information from different modalities, which influences their perception of what they hear, see or feel. For example, when infants are bounced to an ambiguous drumming rhythm (i.e., no accented beats) on every second beat, they subsequently prefer listening to an accented version in duple metre rather than triple metre (Phillips-Silver and Trainor 2005). If they are bounced on every third beat, they subsequently prefer the rhythm in triple metre. Simply watching an adult move on every second or third beat does not induce the preference. Similarly, adults' encoding of rhythms is influenced by their own movement but not by observations of others' movement (Phillips-Silver and Trainor 2007). Perhaps mothers' typical practice of moving with infants as they sing helps infants encode the rhythms of their songs.

## Consonance matters

Much of the research on infant music perception has been done with auditory patterns that lack the expressive qualities of maternal performances, including the demonstrably potent qualities of the human voice (Vouloumanos and Werker 2004, 2007). Undoubtedly, instrumental and synthesized music are less interesting for infants but such materials enable precise control over various features, which makes it possible to specify the features noticed or preferred. Such research has revealed that, from the newborn period on, infants prefer consonant sequences of harmonic intervals to dissonant sequences (Masataka 2006; Trainor and Heinmiller 1998; Trainor *et al.* 2002). Infants remain calm while listening to consonant instrumental music, but they show a variety of negative reactions when dissonant intervals replace many of the consonant intervals (Zentner and Kagan 1996). Infants are also proficient at detecting melodic changes, big and small, but they are considerably more accurate in the context of sequences with consonant melodic intervals than those with dissonant intervals (Schellenberg and Trehub 1996). The perfect fifth interval, which is ubiquitous across musical cultures (Sachs 1943), seems to be especially important for infant listeners (Trainor and Trehub 1993).

## Musical memories

There have been no attempts to assess infants' memory for the songs that mothers sing at home, but there is evidence that infants retain some aspects of music heard regularly. When infants are exposed to a Mozart sonata periodically during a 2-week period, they subsequently distinguish it from another Mozart sonata (Saffran *et al.* 2000). More limited at-home exposure to synthesized folk melodies enables infants to distinguish those melodies from others (Trainor *et al.* 2004). In circumstances such as the latter, infants retain information about relative pitch but not absolute pitch (Plantinga and Trainor 2005). The salience of the musical selection and performing style may affect what infants remember. For example, infants retain absolute features from expressively sung lullabies after 2 weeks of periodic exposure (Volkova *et al.* 2006). Their short-term memory for music is influenced by a number of factors including its tonality (Trehub *et al.* 1990), temporal regularity (Bergeson and Trehub 2006), and degree of consonance (Schellenberg and Trehub 1996, Trainor and Trehub 1993).

## From universal to cultural tuning

For the first several months of life, infants approach music with open or universal ears.

Like their adult counterparts, they perceive the equivalence of melodies across changes in pitch level, or transposition (Chang and Trehub 1977; Trehub *et al.* 1987) and across changes in tempo (Trehub and Thorpe 1989). Unlike their adult counterparts, who are differentially sensitive to familiar diatonic scales, infants experience no more difficulty detecting changes to melodies based on the Javanese pelog scale than to those based on Western diatonic scales (Lynch *et al.* 1990). Similarly, they perform as well on invented scales as on the major scale, provided the invented scales incorporate the principle of unequal step size (Trehub *et al.* 1999). Although infants are surprisingly proficient at detecting melodic changes (Trehub 2000), they are unlike adults in their insensitivity to the implications of key membership or harmony (Trainor and Trehub 1992). The available evidence indicates that such sensitivity requires years of culture-specific exposure to music, emerging between 5 and 7 years of age (Trainor and Trehub 1994) and exhibiting further development thereafter (Lamont and Cross 1994).

Infants also exhibit universal principles of rhythm perception such as the grouping of notes based on similarities in pitch, loudness, timbre, and temporal proximity (Thorpe *et al.* 1988; Thorpe and Trehub 1989. When the flow of music is interrupted briefly, infants prefer such interruptions between musical phrases rather than within phrases, indicating that they grasp aspects of musical phrasing (Jusczyk and Krumhansl 1993).

Simple metres with isochronous timing (i.e., equal durations between strong beats) such as duple and triple metre are thought to be inherently simpler than the complex metres that are common in many parts of the world (e.g., Eastern Europe, South Asia, Africa). Recent evidence suggests otherwise. For example, 6-month-old infants detect timing changes as readily in Balkan music with simple metres as in Balkan music with complex metres (Hannon and Trehub 2005a). When North American adults are exposed to the same materials, they detect the changes only in the context of simple metres, but adults of Balkan origin detect both types of changes (Hannon and Trehub 2005a). These findings negate the view of inborn predispositions or preferences for simple metres. Instead, they suggest that adults' long-term experience with Western rhythms interferes with their perception of foreign rhythms. By the time infants are 12 months of age, they exhibit adult-like difficulties with the complex metres (Hannon and Trehub 2005b). Unlike their adult counterparts, however, limited listening experience with complex metres is sufficient to overcome those difficulties. Not only does adults' experience interfere with their perception of foreign rhythms, it also interferes with their ability to learn such rhythms. Although passive listening is sufficient for 12-month-olds, adults may require more active listening, perhaps accompanied by movement. Guided moving to music, which influences adults' encoding of simpler rhythmic materials (Phillips-Silver and Trainor 2007), may help adults master the nuances of metrically complex music.

## In sum

The story that emerges from infancy involves a rich musical environment, with mothers delivering performances that match the inclinations of their infants. Moreover, infants have predispositions or inborn preferences for musical features that are common across the world's cultures. Because musical systems across the world differ in notable respects, it makes sense that infants are open to the available alternatives. With increasing exposure to music, they gain expertise as listeners, but that expertise comes at the cost of diminished sensitivity to features that are irrelevant or infrequent in their own musical culture. Even when their listening skills begin to narrow in culturally appropriate ways, they remain flexible listeners and learners, which enables them to attain listening proficiency with foreign rhythms.

Much has been written about young children's ease of mastering the sounds of foreign languages and adults' difficulty in this regard. There may be parallels with regard to music, but the window of opportunity for music perception is unknown. Finally, there are hints that musical enculturation may occur more rapidly in the temporal domain than in the pitch domain. Evidence on the issue is limited, but it is consistent with the view that temporal processing is at

the heart of listening to music and other auditory events (Large and Jones 1999).

## Acknowledgements

The preparation of this paper was assisted by grants from the Social Sciences and Humanities Research Council and the Natural Sciences and Engineering Research Council of Canada.

## References

Bergeson TR and Trehub SE (2002). Absolute pitch and tempo in mothers' songs to infants. *Psychological Science*, **13**, 72–75.

Bergeson TR and Trehub SE (2006). Infants' perception of rhythm patterns. *Music Perception*, **23**, 345–360.

Bergeson TR and Trehub SE (2007). Signature tunes in mothers' speech to infants. *Infant Behavior and Development*, **30**, 648–654.

Chang HW and Trehub SE (1977). Auditory processing of relational information by young infants. *Journal of Experimental Child Psychology*, **24**, 324–331.

Cooper RP and Aslin RN (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, **61**, 1584–1596.

Falk D (2004). Prelinguistic evolution in early hominins: whence motherese? *Behavioral and Brain Sciences*, **27**, 491.

Fernald A (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, **8**, 181–195.

Fernald A (1991). Prosody in speech to children: prelinguistic and linguistic functions. *Annals of Child Development*, **8**, 43–80.

Fernald A, Taeschner T, Dunn J, Papousek M, Boysson-Bardies B and Fukui I (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, **16**, 477–501.

Hannon EE and Trehub S (2005a). Metrical categories in infancy and adulthood. *Psychological Science*, **16**, 48–55.

Hannon EE and Trehub SE (2005b). Tuning in to musical rhythms: infants learn more readily than adults. *Proceedings of the National Academy of Sciences*, **102**, 12639–12643.

Jusczyk PW and Krumhansl CL (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception and Peformance*, **19**, 627–640.

Lamont A and Cross I (1994). Children's cognitive representations of musical pitch. *Music Perception*, **12**, 27–55.

Large EW and Jones MR (1999). The dynamics of attending: how people track time-varying events. *Psychological Review*, **106**, 119–159.

Lynch MP, Eilers RE, Oller DK and Urbano RC (1990). Innateness, experience, and music perception. *Psychological Science*, **1**, 272–276.

Masataka N (1999). Preference for infant-directed singing in 2-day-old hearing infants of deaf parents. *Developmental Psychology*, **35**, 1001–1005.

Masataka N (2006) Preference for consonance over dissonance by hearing newborns of deaf parents and of hearing parents. *Developmental Science*, **9**, 46–50.

Nakata T and Trehub SE (2004). Infants' responsiveness to maternal speech and singing. *Infant Behavior and Development*, **27**, 455–464.

Nakata T and Trehub SE (in preparation). Melodious means of keeping in touch with infants.

Phillips-Silver J and Trainor LJ (2005). Feeling the beat: movement influences infant rhythm perception. *Science*, **308**, 1430.

Phillips-Silver J and Trainor LJ (2007). Hearing what the body feels: auditory encoding of rhythmic movement. *Cognition*, **105**, 533–546.

Plantinga J and Trainor LJ (2005). Memory for pitch: infants use a relative pitch code. *Cognition*, **98**, 1–11.

Sachs C (1943). *The rise of music in the ancient world: East and West*. Norton, New York.

Saffran JR, Loman MM and Robertson RRW (2000). Infant memory for musical experiences. *Cognition*, **77**, B15–B23.

Schellenberg EG and Trehub SE (1996). Natural intervals in music: a perspective from infant listeners. *Psychological Science*, **7**, 272–277.

Shenfield T, Trehub S and Nakata T (2003). Maternal singing modulates infant arousal. *Psychology of Music*, **31**, 365–375.

Szpunar KK, Schellenberg EG and Pliner P (2004). Liking and memory for musical stimuli as a function of exposure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **30**, 370–381.

Thorpe LA and Trehub SE (1989). Duration illusion and auditory grouping in infancy. *Developmental Psychology*, **25**, 122–127.

Thorpe LA, Trehub SE, Morrongiello BA and Bull D (1988). Perceptual grouping by infants and preschool children. *Developmental Psychology*, **24**, 484–491.

Trainor LJ, (1996). Infant preferences for infant-directed versus noninfant-directed playsongs and lullabies. *Infant Behavior and Development*, **19**, 83–92.

Trainor LJ and Heinmiller BM (1998). The development of evaluative responses to music: infants prefer to listen to consonance over dissonance. *Infant Behavior and Development*, **21**, 77–88.

Trainor LJ and Trehub SE (1992). A comparison of infants' and adults' sensitivity to Western musical structure. *Journal of Experimental Psychology: Human Perception and Performance*, **18**, 394–402.

Trainor LJ and Trehub SE (1993). What mediates infants' and adults' superior processing of the major over the augmented triad? *Music Perception*, **11**, 185–196.

Trainor LJ and Trehub SE (1994). Key membership and implied harmony in Western tonal music: developmental perspectives. *Perception and Psychophysics*, **56**, 125–132.

Trainor LJ, Clark ED, Huntley A and Adams BA (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior and Development*, **20**, 383–396.

Trainor LJ, Tsang CD and Cheung VHW (2002). Preference for sensory consonance in 2- and 4-month-old infants. *Music Perception*, **20**, 187–194.

Trainor LJ, Wu L and Tsang CD (2004). Long-term memory for music: infants remember tempo and timbre. *Developmental Science*, **7**, 289–296.

Trehub SE, (2000). Human processing predispositions and musical universals. In NL Wallin, B Merker and S Brown, eds The origins of music, 427–448. MIT Press, Cambridge MA.

Trehub SE and Thorpe LA (1989). Infants' perception of rhythm: categorization of auditory sequences by temporal structure. *Canadian Journal of Psychology*, **43**, 217–229.

Trehub SE and Trainor LJ (1998). Singing to infants: lullabies and play songs. *Advances in Infancy Research*, **12**, 43–77.

Trehub SE, Schellenberg EG and Kamenetsky SB (1999). Infants' and adults' perception of scale structure. *Journal of Experimental Psychology: Human Perception and Performance*, **25**, 965–975.

Trehub SE, Thorpe LA and Morrongiello BA (1987). Organizational processes in infants' perception of auditory patterns. *Child Development*, **58**, 741–749.

Trehub SE, Thorpe LA and Trainor LJ (1990). Infants' perception of *good* and *bad* melodies. *Psychomusicology*, **9**, 5–19.

Trehub SE, Unyk AM and Trainor LJ (1993a). Adults identify infant-directed music across cultures. *Infant Behavior and Development*, **16**, 193–211.

Trehub SE, Unyk AM and Trainor LJ (1993b). Maternal singing in cross-cultural perspective. *Infant Behavior and Development*, **16**, 285–295.

Trehub SE, Unyk AM, Kamenetsky SB, Hill DS, Trainor LJ, Henderson JL, *et al.*, (1997). Mothers' and fathers' singing to infants. *Developmental Psychology*, **33**, 500–507.

Unyk AM, Trehub SE, Trainor LJ and Schellenberg EG (1992). Lullabies and simplicity: a cross-cultural perspective. *Psychology of Music*, **20**, 15–28.

Volkova A, Trehub SE and Schellenberg EG (2006). Infants' memory for musical performances. *Developmental Science*, **9**, 583–589.

Vouloumanos A and Werker JF (2004). Tuned to the signal: the privileged status of speech for young infants. *Developmental Science*, **7**(3), 270–276.

Vouloumanos A and Werker JF (2007). Listening to language at birth: evidence for a bias for speech in neonates. *Developmental Science*, **10**, 159–164.

Werker JF and McLeod PJ (1989). Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology*, **43**, 230–246.

Zentner MR and Kagan J (1996). Perception of music by infants. *Nature*, **383**, 29.

# CHAPTER 22

# Music in the school years

Alexandra Lamont

## Introduction

Music is ubiquitous in young children's experiences (see Chapter 21 this volume), but as children get older their experiences become more diverse. Defining the trajectories of musical development is thus complex; explaining them is still more challenging. Beginning by considering definitions of development, this chapter first provides a selective overview of existing evidence on children's responses to musical elements and secondly evaluates our understanding of musical development in culture and context.

### What is musical development?

Despite much research interest, there is still no overarching consensus on what musical development is. Sloboda (1985) distinguished between two types of musical development: enculturation (generalized and effortless) and training (specialized, deliberate and conscious). Many researchers now consider the development of musical competence (Deliège and Sloboda 1996), further subdefined into activities such as 'singing, graphic representation of music, melodic perception, and composition' (Hargreaves 1996 p. 145) or 'skills, understandings and attitudes' (McPherson 2006, p. v). Runfola and Swanwick (2002) argue that any explanation of musical development must incorporate the three areas of production, performance, and perception (see also Hargreaves and Zimmerman 1992).

Other forms of development have more clearly defined goals: for example, literacy development has an end point of being able to read. Becoming a musician is a more nebulous concept (MacDonald *et al.* 2002; Chapter 42 this volume), and not all children think of themselves as equally musical (Lamont 2002). Even research which adopts a relatively straightforward 'expertise' model of adult musicians has difficulty accounting for the diversity of findings in the musically more naive (Davidson 1994), and this may explain the imprecision in definitions. This chapter will take an inclusive approach to what musical development might encompass.

## Developing musical understanding

First I review research on how children understand separate elements of music, individually (pitch/harmony, rhythm/metre, timbre) and in combination (structure, form, style), drawing on research that isolates and explores these experimentally and evidence from more ecologically valid tasks such as singing and composing.

### Pitch, tonality, melody and harmony

Pitch includes tone chroma (individual notes), contour (high/low and patterns), and melodic and harmonic relationships between tones (tonality/tonal hierarchy) (see also Chapter 5 this volume). One early pitch capacity is absolute pitch, typically demonstrated by labelling individual pitches by note names. Absolute pitch is more easily acquired in early childhood (Takeuchi and Hulse 1993) and can be trained

more rapidly in children aged 4–5 than those aged 13–15 (Crozier 1997). Crozier suggests that through music training which emphasizes relative pitch skills, children unlearn absolute pitch. A second capacity relates to the psychophysical dimension of pitch height. Young children are unable to label pitch as 'high' and 'low', and much of their difficulty results from problems in analogical mapping between the musical and verbal concepts (Costa-Giomi and Descombes 1996). Although differences in pitch height can be perceived by infants (see Chapter 21), the terminology to describe the dimension of pitch height must be learned and takes time to stabilize.

Considering relationships between pitches, research has explored children's developing sense of tonality. Some studies find marked age differences: for example, Krumhansl and Keil (1982) showed an increasing sophistication in differentiation between and among diatonic and non-diatonic notes from 6 to 11 years (cf. Lamont and Cross 1994). Other studies find stable representations across this age range (Cuddy and Badertscher 1987), with children performing with less sophistication than musically trained adults. Training appears to accelerate the patterns of development found with age (Morrongiello and Roes 1990; Lamont 1998). Melodic expectancy studies using both perception and singing tasks show that for children aged 5–11 the best next note in a melody is one close in pitch, whereas a pitch reversal only emerges as a good next note at the age of 11 (Schellenberg *et al.* 2002). Preschoolers can detect mistunings to conventional tonal melodies (Trehub *et al.* 1986), but often fail to notice transpositions and contour-preserving changes as 'different' (Morrongiello *et al.* 1989), while out-of-harmony mistunings become detectable around the age of 7 (Trainor and Trehub 1994). In priming tasks, 6- and 11-year-olds make faster judgements about which vowel sound or instrument a target tone is sung or played on, or whether it is consonant or dissonant, when the target is a tonic chord (Schellenberg *et al.* 2005).

Converging evidence for developing pitch understanding comes from musical production tasks. Davidson (1985, 1994) shows a progression from contour schemes in songs at 3 years of age to fixed pitch reference points with older children. Bamberger's (1986, 1991) research also highlights qualitatively different ways of understanding pitch in context. 6–11 year olds without musical training organize melodies figurally into shapes, while highly trained 11- and 12-year-olds can use pitch formally in an abstract manner. Children's melodic improvisations show a similar developmental sequence: children aged 6–7 only use the first five diatonic tones, at 8–9 prefer tones presented in the initial stimulus, and at 10–11 create improvisations emphasizing the tonic triad; the oldest children are most likely to end with a tonic chord (Paananen 2006a, b).

In summary, young children possess certain fundamental pitch capacities, such as the ability to respond to absolute pitch, which become more refined and culturally specified with age. By 6 or 7, children have developed an implicit understanding of Western tonality almost equivalent to that of adults, and use this in listening judgements and decisions about creating music. Furthermore, a shift from context-dependency to context-independency in pitch relationships occurs between 6 and 11 (Lamont and Cross 1994; Paananen 2006a). Musical training accelerates the development of understanding of pitch and tonality, although often does not change it radically. Finally, although some studies find differences between 11-year-olds and adults, very little research has explored development beyond the age of 11.

## Tempo, timing, rhythm and metre

There has been less research on children's understanding of the temporal aspects of music, and existing studies use many different methodologies. In listening tasks, little difference is found according to age or music training in children's abilities to detect temporal irregularities (Drake *et al.* 2000) or to segment according to changes in tone duration or pause duration (Drake 1993) between 4 and 12 years. In general, children perform slightly less consistently or accurately than adults (Drake and Bertrand 2003). Production tasks also indicate little developmental difference in children's ability to synchronize (Drake *et al.* 2000) or to reproduce simple binary and regular rhythmic patterns (Drake and Gérard 1989) between 5 and 12 years, although they are less

accurate than adults with ternary ratio rhythms and irregular patterns.

Evidence from children's invented notations of rhythms shows differences in understanding in middle childhood, moving from enactive action drawings to figural patterns and shapes; only children with training are able to focus on absolute durations and inter-onset intervals (Bamberger 1991). Children's compositions also show shifts from conventional metric patterns at age 5–8 to more speculative use of rhythm and metre at 9–11 (Swanwick and Tillman 1986; Brophy 2002; Paananen 2006c).

Certain fundamental features of temporal understanding are demonstrated relatively consistently across childhood, although others such as segmentation and the understanding (both perceptual and compositional) of more complex rhythmic ratios show gradual improvements with age and training. Again, although differences in understanding are found between children and adults, development beyond the age of 12 has not been explored.

## Timbre

Very little research has been conducted on the development of timbre perception. One direct investigation of timbral sensitivity (Lowther 2004) finds rapid development between the ages of 3–8. Younger children make easy discriminations very accurately (e.g. comparing a triangle sound to a ratchet, or tuba to celeste), but are much less successful on more difficult comparisons (e.g. flute/clarinet). From children's qualitative responses, Lowther concludes that although performance increases with age, specific experiences of particular sounds are responsible for helping children make sense of timbral information.

Indirect information about timbre perception skills comes from research such as the priming paradigm used by Schellenberg *et al.* (2005, Experiment 2), finding that children aged 8–11 can discriminate between a piano and a trumpet. Despite the lack of concrete empirical evidence, timbre identification is often used as one indicator of exceptional musical ability in children, being presumed to indicate a high level of auditory discrimination (Shuter-Dyson 1999) linked to potential for learning.

## Combinations of elements: musical structure and form

Brand (2000) has explored children's errors in learning to sing a new and unfamiliar song between the ages of 6 and 12. While older children learn the song more rapidly, all children irrespective of age or musical training make plausible errors in terms of song organization (e.g. more symmetrical phrases). Thus children attempt to organize the music they encounter into meaningful sections and familiar gestures, as in Davidson's 'pot-pourri' songs (1985). Oura and Hatano (1988) find 9–10-year-olds with around 5 years of musical experience can reproduce a novel melody as rapidly and accurately as trained adults (and more so than inexperienced adults). Thus familiar organizing principles such as tonality or repetition are seen to have been internalized and used in order to make sense of unfamiliar music.

Children's similarity judgements in terms of pitch and duration improve in accuracy between 5 and 11 years (Stevens and Gallagher 2004), although higher accuracy levels are found for pitch than duration. When given the opportunity to organize musical stimuli along the dimensions of pitch, contour, tempo, rhythm, timbre and loudness (Schwarzer 1997), young children focus on melody-unspecific features such as loudness (5–6-year-olds), tempo, and timbre (6–7-year-olds), while adults use more melody-specific features like contour. Children categorize short melodies analytically (using one musical element) rather than holistically (using combinations), and none of the younger children use pitch or contour relations. However, sorting tasks using real musical stimuli show that surface features such as dynamics and contour are commonly used, as well as underlying features such as tonal structure, by 10–11-year-olds (Koniari *et al.* 2001). Tonality can also be used to organize excerpts by younger children with musical training (Martin *et al.* 2003).

In a study directly comparing melody/tonality and rhythm/metre, Paananen (2006a) shows that children's improvisations at age 6–7 either focus on the surface (melodic/rhythmic figures) or deep (tonal or metrical hierarchy) structures, passing through a substage of beginning coordination between the two at 8–9 to a final substage

of integration at 10–11. Children's invented notations also become more specific and detailed with age in relation to various musical dimensions (pitch, rhythm, timbre), illustrating how symbolization is tied to musical understanding (Fung and Gromko 2001; Bamberger 2006).

This evidence is somewhat conflicting because of the variety of tasks employed. When making simple comparison judgements to artificial musical stimuli (e.g. Stevens and Gallagher 2004), young children can focus on isolated musical elements (cf. Paananen 2006a). However, when using more complex stimuli or real music (e.g. Koniari *et al.* 2001), children are better able to combine musical elements. Although children aged 5–6 can use different musical elements and structures in complex tasks, older children are better able to deal with more complex underlying structures of music, and this is also sometimes dependent on formal training. It may take until adulthood to fully integrate different musical features (Schwarzer 1997), but research has not yet explored this beyond the age of 11.

## Musical structure and emotion

Adults' perception of emotion relies on the structural features of tempo and mode in Western music (see Chapter 11 this volume). With happy and sad musical excerpts from real music contrasting in tempo and mode, 3–4-year-olds are unable to distinguish them on any basis; 5-year-olds recognize emotions based on tempo changes, while 6–8-year-olds perform similarly to adults in recognizing emotional changes due to both tempo and mode (Dalla Bella *et al.* 2001). In more naturalistic studies, children aged 3–5 can recognize music that adults judged as happy, sad, angry and fearful (Nawrot 2003), although they sometimes confuse fear and anger (Terwogt and van Grinsven 1991).

While Dalla Bella *et al.* (2001) note that it is remarkable that 6-year-olds demonstrate full knowledge of the rules governing happy and sad emotions in music, it should be recalled that these are 'basic' emotions and there is room for further sophisticated development. More research is required, along the lines of the recent study by Stachó (2006) indicating that children aged 3–6 are significantly less able to decode the emotional content of different musical performances.

## Style discrimination

Musical preferences can be shown even before birth (see Chapter 20 this volume). Three-year-olds mainly prefer child-oriented music like nursery rhymes or television themes (Lamont 2008), and can also discriminate between different styles of classical and popular music given an appropriately sensitive methodology (Marshall and Hargreaves 2007). However, little research has directly explored the development of children's understanding of style (cf. Hargreaves *et al.* 2006).

Children appear to experience phases of 'open-earedness', such as middle childhood, where they tolerate a large amount of musical styles (Hargreaves 1982), and 'closed-earedness', such as adolescence, where their preferences are far more constrained (LeBlanc 1991). For example, Boal-Palheiros *et al.* (2006) find children aged 9–11 like very unfamiliar and complex music (Boulez, Ligeti and Stockhausen) more than those aged 12–14. In adolescence, a narrow range of musical preference forms an important part of self-identity (North *et al.* 2000; Stålhammar 2003). The 'rebound of open-earedness' in early adulthood (LeBlanc 1991, p. 37) coincides with a phase of seriously acquiring music (Greasley and Lamont 2006). Music preferences still convey messages about other individuals, although more subtle judgements can be made (Rentfrow and Gosling 2006).

The rapidly changing nature of popular musical styles presents an empirical challenge (cf. Hargreaves *et al.* 2006). Furthermore, developmental effects are understood to be moderated by a large number of other factors influencing preference (LeBlanc 1982) which relate to the music itself, to the listener, and to the context (Hargreaves *et al.* 2006) (see Chapter 15 this volume).

## Production and performance skills

Although all children sing, children's singing skill takes time to stabilize (Welch 2006). Their vocal ranges at the start of school are relatively constrained, with large individual differences in

skill (Leighton and Lamont 2006). Children's voices continue to change throughout childhood and adolescence, both continuously, for example the ability to sing in tune (Welch 1996) and qualitatively, for example physical changes in production capacity and vocal development in adolescence (Welch 1998).

Some children also learn to play an instrument, and the component skills are considered elsewhere (see Part 6 this volume; McPherson 2006). While research has revealed much about what is required to make progress with learning an instrument, there is still no real developmental perspective on this topic. McPherson and Davidson (2006) outline a set of learning principles for children of different ages (up to 5, 6–9, and 10 and above) but these are based more on common-sense suggestions (e.g. beginning at an early age, encouraging parents to sing and play musical games) than principled theorizing.

The process and product of children's compositions has been used to generate models of development (e.g. Barrett 1996; Davies 1992). Generally, children become more sophisticated and abstract in their musical compositions with age (Swanwick and Tillman 1986). Although it is often argued that creativity is not age-dependent or phase-related but rather dependent on context, some general changes can be charted (Burnard 2006). For example, early childhood is characterized by the importance of play, whereas middle childhood has wider social and educational influences, and adolescence prioritizes personal identity.

Finally, learning to read and write music is a highly culture-specific form of development (see Mills and McPherson 2006). Converting children's intuitive and invented notations into conventional Western staff notation is complex, and children often experience difficulties in matching sounds to symbols (McPherson and Gabrielsson 2002). Pedagogic approaches using meaningful wholes rather than individual isolated pitches are more effective in the earlier stages, perhaps reflecting children's musical understanding (Mills and McPherson 2006). However, arguably certain elements of music are better perceived analytically rather than holistically even by young children (cf. Schwarzer 1997) and so this approach may not be universally successful.

## Theorizing development

A critical issue when considering all this research is the *causes* of differences in musical understanding. Adopting an individual differences approach, we would expect differences according to age, gender, and prior musical experience (over short and long time spans). Most research tends to seek, and generally finds, age-related differences. Some studies also include the effects of formal musical training. However, the field has not yet produced theoretical models of musical understanding which explain the range of musical behaviours coherently (cf. North and Hargreaves 2008; Runfola and Swanwick 2002). I next consider some of the challenges to theorizing musical development.

### Musical development: one thing or many?

First, including different types of musical behaviour within a single theory presents a major theoretical challenge. As illustrated above, many differences in research findings can be ascribed to the particular tasks: building a composition requires different skills and poses different challenges to listening to two notes and deciding if they are the same or different, and explicit tasks also show slower rates of development than implicit ones (Schellenberg *et al.* 2005). To make progress on the relative 'rates' of development of different elements of music even within the field of musical understanding requires studies which explicitly include and compare these in real musical settings. For example, Stevens and Gallagher (2004) show that children can make better judgements based on pitch than duration, yet Schwarzer (1997) illustrated that when more potential musical elements are added children aged 5–6 choose neither of these dimensions. It is thus important that experimental demands do not unnecessarily constrain children's responses, and that researchers seek opportunities to allow children to demonstrate the full range of their understanding.

Should a theory of musical development be able to account for many different kinds of musical behaviour, including understanding alongside activities such as singing, playing instruments and composing? In developmental

psychology, existing grand theories of general development (Piaget 1953; Vygotsky 1978) have largely been abandoned in favour of more specific models such as theories of cognitive development (e.g. Halford 1992; Siegler 1996). In music, the generalizability of theories or models hinges on a balance between precision and explanatory power. As in mainstream developmental psychology, the current state of the field is disparate and lacking coherence (see also North and Hargreaves 2008). Broader approaches, like Hargreaves and Galton's model of artistic competence (1992) or Swanwick and Tillman's developmental spiral (1986), can provide a useful description of similarities in stages of understanding (loosely tied to age) across different musical activities. However, these remain at a purely descriptive level. More constrained musical models of specific musical skills or understanding (e.g. Paananen 2006a, b; Stevens and Gallagher 2004; Welch 1996) seem to be more successful in explaining musical behaviour and predicting development in a limited number of domains. It may be that an overarching theory of musical development is still out of reach—at any rate, such a theory will need to be a great deal more flexible than those currently in circulation.

## Where is musical development going, and how does it get there?

Considering the process of development itself, as illustrated above, very little research has systematically attempted to map cross-sectional patterns over large enough age spans or fine enough levels of detail to be able to explain the directions or the motors of change. There are two main reasons for this. First, from a theoretical perspective, debate continues over whether development progresses linearly or in a stepwise fashion towards one or multiple end points (Bamberger 2006; Hargreaves 1996; Mills 2005; Swanwick and Tillman 1986). Secondly, most research is not methodologically designed in such a way to enable these questions to be answered. Most studies compare children of different ages (sometimes including adults). Differences imply either quantitative change resulting from exposure or qualitative change resulting from different ways of thinking, while similarities indicate no development. Furthermore, the period between 11–12 years of age and adulthood is rarely studied, and music training is rarely added as a developmental mechanism. As shown above, results seem highly dependent on the musical domain, the tasks administered, and the particular ages selected.

However, more interactive studies following children over short time spans highlight micro-level changes in their musical understanding. For example, Bamberger's (1991) detailed case study of an individual child's progress from a focus on the figural properties of melodies through to an awareness of formal features provides an insight into the process of change. This illustrates that the motor of change is the child's engagement and re-engagement with the musical material, in this case largely stimulated by a research/teaching intervention. Applying this kind of approach to more finely age- and experience-graded samples over longer time spans is necessary in order to be able to disentangle the theoretical questions of the relative contributions of age and experience which are becoming recognized as being so critical in musical development.

## Natural development and cultural environments

The relative contribution of innately specified and culturally acquired aspects of musical development is another hotly debated issue. Much recent evidence supports the notion of certain aspects of musical understanding being innately specified, particularly considering elements like pitch in isolation. The cultural environment has been particularly weakly theorized in musical development: some simple approaches draw on connectionist principles to suggest that exposure predicts behaviour (Gruhn and Rauscher 2002), supported by evidence for linear patterns of development with age, for example in tonal sensitivity (Krumhansl and Keil 1982).

However, both natural and cultural approaches are rooted in a Western philosophical and psychological tradition where the individual child is either the 'motor' or the 'recipient' of development (cf. Burman 1994). For example, McPherson's explicit scope in exploring how biological, environmental, social, cultural and historical factors have shaped the acquisition of

musical skill, understanding and attitudes (2006) treats the child as a recipient of culture. Vygotsky's (1978) approach placed a far greater emphasis on the specific social and cultural experiences that individuals encounter and which shape their individual development. Given the fundamentally social nature of music in society, a more complex formulation of the learning process is required which accounts for the child as an active participant in culture. This aspect of cultural environment has been weakly theorized even in models which attempt to account for experience (e.g. Swanwick and Tillman 1986; see Runfola and Swanwick 2002). Such a goal may be achieved following Gruhn and Rauscher's suggestion to focus research on the 'micro-structure of cognition' (2002, p. 446).

## Concluding remarks: the importance of context

The research carried out to date on the development of musical understanding (pitch, tonality, rhythm, metre, form, structure and style) and of musical activities (singing, instrumental learning, composing, and improvising) reveals a complex pattern of evidence. The principal debates in the field hinge around the types and breadth of development being considered, and the influences of nature and culture, age and experience. I conclude with some suggestions for future directions relating to the importance of context.

Elsewhere (Lamont 2002) I have argued for the usefulness of contextually grounded theories such as Bronfenbrenner's (1979) ecological systems theory in understanding children's developing spheres of influence and levels of engagement with music, ranging from the micro contexts of home and school up to the macro contexts of dominant beliefs in society. In similar vein, Burnard (2006) adopts a systems view in understanding influences on children's developing creativity. She identifies a 'superculture' of children's musical creativity, incorporating the overlapping spheres of culture, society, in- and out-of-school contexts to specify various microcultures to be investigated. These approaches provide more formalized and thus practical ways of identifying and exploring the significant contexts in children's and adolescents' lives which support and motivate many kinds of musical development.

A fuller consideration of context may also help address the empirical challenges of engagement and motivation for music. Not all children engage with music at school, with music tuition, or with music from different cultures or subcultures (e.g. Lamont *et al.* 2003; Stålhammar 2003). This provides a different impetus for developmental research. In addition to exploring the individual features of age and experience more systematically, it becomes vital to investigate the effects of different situations and settings on musical development of every kind. More flexible yet sensitive theories of musical development may offer solutions to these challenges, as well as informing us better about the ways in which children develop in and through music and the diversity of musical goals to which they aspire.

## References


Bamberger J (1986). Cognitive issues in the development of musically gifted children. In RJ Sternberg and JE Davidson, eds, *Conceptions of giftedness*, 388–413. Oxford University Press, Oxford.

Bamberger J (1991). *The mind behind the musical ear*. Harvard University Press, Cambridge, MA.

Bamberger J (2006). What develops in musical development? In GE McPherson, ed., *The child as musician: a handbook of musical development*, 69–91. Oxford University Press, Oxford.

Barrett M (1996). Children's aesthetic decision-making: an analysis of children's musical discourse as composers. *International Journal of Music Education*, **28**, 37–62.

Boal-Palheiros G Ilari B and Monteiro F (2006). Children's responses to 20th century 'art' music, in Portugal and Brazil. In M Baroni, AR Addessi, R Caterina, M Costa, eds, *Proceedings of the 9th International Conference on Music Perception and Cognition*, 588–595. University of Bologna, Bologna, Italy.

Brand E (2000). Children's mental musical organisations as highlighted by their singing errors. *Psychology of Music*, **28**(1), 62–80.

Bronfenbrenner U (1979). *The ecology of human development*. Harvard University Press, Cambridge, MA.

Brophy TS (2002). The melodic improvisations of children aged 6–12: a developmental perspective. *Music Education Research*, **4**(1), 73–92.

Burman E (1994). *Deconstructing developmental psychology*. Routledge, London.

Burnard P (2006). The individual and social worlds of children's musical creativity. In GE McPherson, ed., *The child as musician: a handbook of musical development*, 353–374. Oxford University Press, Oxford.

Costa-Giomi E and Descombes V (1996). Pitch labels with single and multiple meanings: a study with French-speaking children. *Journal of Research in Music Education*, **44**(3), 204–214.

Crozier J (1997). Absolute pitch: practice makes perfect, the earlier the better. *Psychology of Music*, **25**(2), 110–119.

Cuddy LL and Badertscher BD (1987). Recovery of the tonal hierarchy: some comparisons across age and levels of musical experience. *Perception and Psychophysics*, **41**, 609–620.

Dalla Bella S, Peretz I, Rousseau L and Gosselin N (2001). A developmental study of the affective value of tempo and mode in music. *Cognition*, **80**(3), B1–B10.

Davidson L (1985). Tonal structures of children's early songs. *Music Perception*, **2**(3), 361–373.

Davidson L (1994). Song singing by young and old: a developmental approach to music. In R Aiello, with JA Sloboda, eds, *Musical perceptions*, 99–130. Oxford University Press, Oxford.

Davies C (1992). *Listen to my song. A study of songs invented by children aged 5–7 years. British Journal of Music Education*, **9**, 19–48.

Deliège I and Sloboda JA (eds) (1996). *Musical beginnings: origins and development of musical competence*. Oxford University Press, Oxford.

Drake C (1993). Influence of age and experience on timing and intensity variations in the reproduction of short musical rhythms. *Psychological Belgica*, **33**, 217–228.

Drake C and Bertrand D (2003). The quest for universals in temporal processing in music. In I Peretz, R Zatorre, eds, *The cognitive neuroscience of music*, 21–31. Oxford University Press, Oxford.

Drake C and Gérard C, (1989). A psychological pulse train: how young children use this cognitive framework to structure simple rhythms. *Psychological Research*, **51**, 16–22.

Drake C, Jones MR and Baruch C (2000). The development of rhythmic attending in auditory sequences: attunement, reference period, focal attending. *Cognition*, **77**, 251–288.

Fung CV and Gromko JE (2001). Effects of active versus passive listening on the quality of children's invented notations and preferences for two pieces from an unfamiliar culture. *Psychology of Music*, **29**(2), 128–138.

Greasley AE and Lamont A (2006). Music preference in adulthood: why do we like the music we do? In M Baroni, AR Addessi, R Caterina, M Costa, eds, *Proceedings of the 9th International Conference on Music Perception and Cognition*, 960–966. University of Bologna, Bologna, Italy.

Gruhn W and Rauscher F (2002). The neurobiology of music cognition and learning. In R Colwell, C Richardson, eds, *New handbook of research on music teaching and learning*, 445–460. Oxford University Press, Oxford.

Halford GS (1992). *Children's understanding: the development of mental models*. Erlbaum, Hillsdale, NJ.

Hargreaves DJ (1982). The development of aesthetic reactions to music. *Psychology of Music, Special Issue*, 51–54.

Hargreaves DJ (1996). The development of musical and artistic competence. In I Deliège, JA Sloboda, eds, *Musical beginnings: origins and development of musical competence*, 145–170. Oxford University Press, Oxford.

Hargreaves DJ and Galton M (1992). Aesthetic learning: psychological theory and educational practice. In B Reimer, RA Smith, eds, *National Society for the Study of Education yearbook on the arts in education*, 124–150. NSSE, Chicago, IL.

Hargreaves DJ, North AC and Tarrant M (2006). Musical preference and taste in childhood and adolescence. In GE McPherson, ed., *The child as musician: a handbook of musical development*, 135–154. Oxford University Press, Oxford.

Hargreaves DJ and Zimmerman MP (1992). Developmental theories of music learning. In R Colwell, ed., *Handbook of research on music teaching and learning*, 377–391. Schirmer, New York.

Koniari D, Predazzer S and Mélen M (2001). Categorization and schematization processes used in music perception by 10- to 11-year old children. *Music Perception*, **18**(3), 297–324.

Krumhansl CL and Keil FC (1982). Acquisition of the hierarchy of tonal functions in music. *Memory and Cognition*, **10**, 243–251.

Lamont A (1998). Music, education, and the development of pitch perception: the role of context, age, and musical experience. *Psychology of Music*, **26**(1), 7–25.

Lamont A (2002). Musical identities and the school environment. In RAR MacDonald, DJ Hargreaves, DE Miell, eds, *Musical identities*, 41–59. Oxford University Press, Oxford.

Lamont A (2008). Young children's musical worlds: musical engagement in three-year-olds. Journal of Early *Childhood Research*, **6**(3), 247–261.

Lamont A and Cross I (1994). Children's cognitive representations of musical pitch. *Music Perception*, **12**, 27–55.

Lamont A, Hargreaves DJ, Marshall NA and Tarrant M (2003). Young people's music in and out of school. *British Journal of Music Education*, **20**(3), 229–241.

LeBlanc A (1982). An interactive theory of music preference. *Journal of Music Therapy*, **19**, 28–45.

LeBlanc A (1991). Effect of maturation/aging on music listening preference: a review of the literature. Paper presented to the Ninth National Symposium on Research in Music Behavior, Cannon Beach, Oregon, March 7–9.

Leighton GL and Lamont A (2006). Exploring children's singing development: do experiences in early schooling help or hinder? *Music Education Research*, **8**(3), 311–330.

Lowther D (2004). An investigation of young children's timbral sensitivity. *British Journal of Music Education*, **21**(1), 63–80.

MacDonald RAR, Hargreaves DJ and Miell DE (eds) (2002). *Musical identities*. Oxford University Press, Oxford.

Marshall N and Hargreaves DJ (2007). Musical style discrimination in the early years. *Journal of Early Childhood Research*, **5**(1), 35–49.

Martin N-J, Lamont A and Dibben N (2003). Children's perception of similarity relations in music. In R Kopiez, AC Lehmann, I Wolter and C Wolf, eds, *Proceedings of the 5th Triennial ESCOM Conference*, Hanover,

Germany, 586–590. Hanover University of Music and Drama, Hanover.

McPherson GE (ed.) (2006) *The child as musician: a handbook of musical development*. Oxford University Press, Oxford.

McPherson GE and Davidson JW (2006). Playing an instrument. In GE McPherson, ed., *The child as musician: a handbook of musical development*, 331–351. Oxford University Press, Oxford.

McPherson GE and Gabrielsson A (2002). From sound to sign. In R Parncutt, GE McPherson, *eds, The science and psychology of musical performance: creative strategies for music teaching and learning*, 99–115. Oxford University Press, Oxford.

Mills J (2005) *Music in the school*. Oxford University Press, Oxford.

Mills J and McPherson GE (2006). Musical literacy. In GE McPherson, ed., *The child as musician: a handbook of musical development*, 155–171. Oxford University Press, Oxford.

Morrongiello BA and Roes CL (1990). Developmental changes in children's perception of musical sequences: effects of musical training. *Developmental Psychology*, **26**, 814–820.

Morrongiello BA, Roes CL and Donnelly F (1989). Children's perception of musical patterns: effects of music instruction. *Music Perception*, **6**, 447–462.

Nawrot ES (2003). The perception of emotional expression in music: Evidence from infants, children, and adults. *Psychology of Music*, **31**(1), 75–92.

North AC and Hargreaves DJ (2008). *The social and applied psychology of music*. Oxford University Press, Oxford.

North AC, Hargreaves DJ and O'Neill SA (2000). The importance of music to adolescents. *British Journal of Educational Psychology*, **70**, 255–272.

Oura Y and Hatano G (1988). Memory for melodies among subjects differing in age and experience in music. *Psychology of Music*, **16**, 91–109.

Paananen P (2006a). Melodic improvisation at the age of 6–11 years: development of pitch and rhythm. *Musicae Scientiae*, XI(1), 89–119.

Paananen P (2006b) Harmonizing a tonal melody at the age of 6–15 years. In M Baroni, AR Addessi, R Caterina and M Costa, eds *Proceedings of the 9th International Conference on Music Perception and Cognition*, 484–488. University of Bologna, Bologna, Italy.

Paananen P (2006c). The development of rhythm at the age of 6–11 years: non-pitch rhythmic improvisation. *Music Education Research*, **8**(3), 349–368.

Piaget J (1953). *The origins of intelligence in children*. Routledge and Kegan Paul, London.

Rentfrow PJ and Gosling SD (2006). Message in a ballad: the role of musical preferences in interpersonal perception. *Psychological Science*, **17**(3), 236–242.

Runfola M and Swanwick K (2002). Developmental characteristics of music learners. In R Colwell, C Richardson, eds, *New handbook of research on music teaching and learning*, 373–397. Oxford University Press, Oxford.

Schellenberg EG, Adachi M, Purdy KT and McKinnon MC (2002). Expectancy in melody: tests of children and adults. *Journal of Experimental Psychology: General*, **131**(4), 11–537.

Schellenberg EG, Bigand E, Poulin-Charronat B, Garnier C and Stevens C (2005). Children's implicit knowledge of harmony in Western music. *Developmental Science*, **8**(6), 551–566.

Schwarzer G (1997). Analytic and holistic modes in the development of melody perception. *Psychology of Music*, **25**(1), 35–56.

Shuter-Dyson R (1999). Musical ability. In D Deutsch, ed. *The psychology of music*, 2nd edn, 627–651. Academic Press, New York.

Siegler RS (1996). *Emerging minds: the process of change in children's thinking*. Oxford University Press, Oxford.

Sloboda JA (1985). *The musical mind: the cognitive psychology of music*. Clarendon Press, Oxford.

Stachó L (2006). Interpretation of the emotional content of musical performance by 3- to 6-year-old children. In M Baroni, AR Addessi, R Caterina, M Costa, eds, *Proceedings of the 9th International Conference on Music Perception and Cognition*, 504–511. University of Bologna, Bologna, Italy.

Stålhammar B (2003). Music teaching and young people's own musical experience. *Music Education Research*, **5**(1), 61–68.

Stevens C and Gallagher M (2004). The development of mental models for auditory events: relational complexity and discrimination of pitch and duration. *British Journal of Developmental Psychology*, **22**, 569–583.

Swanwick K and Tillman J (1986). The sequence of musical development: a study of children's composition. *British Journal of Music Education*, **6**, 305–339.

Takeuchi AH and Hulse SH (1993). Absolute pitch. *Psychological Bulletin*, **113**(2), 345–361.

Terwogt MM and van Grinsven F, (1991). Musical expression of moodstates. *Psychology of Music*, **19**, 99–109.

Trainor LJ and Trehub SE (1994). Key membership and implied harmony in Western tonal music: developmental perspectives. *Perception and Psychophysics*, **56**, 125–132.

Trehub SE, Cohen AJ, Thorpe LA and Morrongiello BA (1986). Development of the perception of musical relations: semitone and diatonic structure. *Journal of Experimental Psychology: Human Perception and Performance*, **12**, 295–301.

Vygotsky LS (1978). *Mind in society: the development of higher psychological processes*. Harvard University Press, Cambridge, MA.

Welch GF (1996). The developing voice. In GF Welch, P White, and D Sergeant, *Singing development in early childhood. Final report to the Leverhulme Trust*. The Centre for Advanced Studies in Music Education, Roehampton Institute, London.

Welch GF (1998). Early childhood musical development. *Research Studies in Music Education*, **11**, 27–41.

Welch GF (2006). Singing and vocal development. In GE McPherson, ed. *The child as musician: a handbook of musical development*, 311–329. Oxford University Press, Oxford.

CHAPTER 23

# The impact of music instruction on other skills

Frances H. Rauscher

## Overview

As one can easily see from the chapters in this section of the Handbook, the area of children's musical development is diverse and exciting. The associated research programmes address some basic issues of human cognition and development, as well as some more applied concerns about children's musical understanding. The purpose of this chapter is to examine how making music may actually enhance children's abilities in other domains of reasoning. The goal is to integrate what is known about the area and to provide a substantive conclusion about the issue.

## The Mozart effect

Our journey into the extra-musical benefits of music instruction must begin by clarifying what has been erroneously referred to as the 'Mozart effect'. The term 'Mozart effect' has been used to refer to the effects of music on children's behaviour, brain function, and overall intelligence. The phrase was initially introduced by the media in relation to a study using college students as subjects, not children (Rauscher *et al.* 1993). The study found that a composite score of three spatial tasks was higher for college students who first listened to a portion of a Mozart Sonata compared to 10 minutes of spoken relaxation instructions or silence. This facilitation lasted only 10–15 minutes. After a series of subsequent studies were unable to replicate the research using a variety of spatial tasks (for review see Chabris 1999), further analyses revealed that the effect was carried by the spatial–temporal task only (Rauscher and Shaw 1998). A meta-analysis that included 36 studies concluded that the Mozart effect 'is limited to a specific type of spatial task ... It is a moderate effect, and it is robust' (Hetland 2000a, p. 136). It is important to note that the original study made no mention of children's spatial abilities or general intelligence. To my knowledge, only three studies have examined the effect of listening to music on children's cognitive abilities, and the results are equivocal: one study found a significant effect (Ivanov and Geake 2003), whereas two studies did not (McKelvie and Low 2002; Schellenberg and Hallam 2005). The original report (Rauscher *et al.* 1993) was followed by an onslaught of media attention, with several articles extending the work to children and intelligence. Before long, the music industry began exploiting the findings. Compact discs with titles such as 'The Mozart Effect for Babies,' and even 'Mozart for Mutts' began appearing in music stores throughout North America and abroad. Indeed, there is very little research to suggest that children who listen to Mozart score higher on ability or intelligence tests than other children (see, for example, Rauscher and Hinton 2006). In fact, there is even evidence to support the notion that children who listen to popular music may score higher on a spatial–temporal task than children who listen to Mozart

(Schellenberg and Hallam 2005). From this and other studies, it appears that the Mozart effect is largely due to arousal or mood rather than to Mozart's music per se.

## The effects of music instruction

This chapter, however, is not about the effects of listening to music. It is about the possible effects of music *instruction* on children's cognitive abilities. Before I begin to review the basic findings, however, two words of caution are warranted. First, several of the studies included in this chapter are *correlational.* A correlation between an intervention (e.g., music instruction) and an outcome measure (e.g., academic performance) offers no information about causality. Of course, two factors that are correlated may be causally related, although one cannot know how: A may cause B (music instruction may cause enhanced academic performance) or B may cause A (children with strong academic abilities may choose to study music). Two correlated factors may also be causally unrelated since a third factor may cause both. I therefore recommend a conservative approach to interpreting the findings of correlational studies.

Second, music should not be taught to students in order to improve their visuospatial or mathematical skills. As Hetland and Winner state,

> If the arts are given a role in our schools because people believe the arts cause academic improvement, then the arts will quickly lose their position if academic improvement does not result . . . The arts must be justified in terms of what the arts can teach that no other subject can teach.
>
> (2001, p. 3)

The studies presented here should not be used to advocate for school music programmes.

Let us now turn to the available research. Studies conducted over the past several decades have provided a wealth of information regarding the non-musical benefits of early music instruction. This line of research fits within the context of earlier discussions on the impact of early experience on children's intellectual development. During the eighteenth century the British empiricists, such as John Locke's *tabula rasa* view of the brain, linked intelligence to the natural powers of reason, implying that intelligence arises from experience and could be fostered in the right environment. In the nineteenth century, however, with the rise of rationalist science and, in particular, notions of natural selection and the study of heredity and genetics, attention was directed instead to the biological bases of intelligence. It was during this period that determinism and biological–maturational theories of intelligence took precedence. The 'nature vs nurture' debate was first articulated in the early twentieth century in terms rather similar to the framework it takes even today. Today, the debate concerning the nature of intelligence has focused on the complex effects of experience on the developing brain. New technologies, such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) have taken current debate beyond the historically oppositional framework of nature or nurture, emphasizing instead that intellectual potential and achievement are the result of complex, only partially understood interactions between the physical body and experience. Thus, in many ways, the study of the effects of music instruction on intelligence is only the latest episode in a long-running controversy, and one with extensive historical roots.

One of the earliest studies to experimentally examine the effects of instrumental instruction on children's cognitive abilities was performed by Irving Hurwitz and his colleagues (Hurwitz *et al.* 1975). First-grade children were assigned to one of two groups. The experimental group received Kodály music lessons for five days each week for seven months. The control group did not receive music instruction. The Kodály method emphasizes the development of rhythmical and motor skills. The purpose of the study was to examine the effects of this method on children's sequencing and spatial skills and reading ability. Children were tested using a variety of sequencing and spatial tasks, as well as two tasks measuring general verbal intelligence. The experimental group scored significantly higher than the control group on three of the five sequencing tasks and four of the five spatial tasks. No significant differences between groups

were found for either of the verbal measures. However, when the data were analysed according to sex, it was determined that the effect was carried by the boys in the study. Girls in the experimental and control groups did not differ on any of the items administered. To determine if reading ability was affected by Kodály music instruction, the authors compared the school records of children who received the instruction in the first grade to children who did not receive the instruction. Although there was no difference between the two groups on reading readiness at the beginning of the first grade, at the end of the first grade, after one academic year of Kodály instruction, the reading achievement scores of the children in the music group were significantly higher than those in the comparison group. When tested again at the end of the second grade, after two academic years of Kodály instruction, the Kodály group continued to score higher than the control group on reading achievement. Spatial and sequencing skills were not retested. Although the authors acknowledge that these positive results may have been due to the presence of a Hawthorne effect (wherein performance improves following any new intervention), they concluded that further research examining the mechanisms of transfer from music instruction to spatial abilities and reading achievement is warranted.

A more recent study also found that music instruction may improve spatial task performance (Rauscher *et al.* 1997). We assigned students from three preschools to music, computer, or no-instruction groups. The instruction groups received several months of instruction in either piano keyboard coupled with group singing sessions, group singing sessions only, or computer lessons. The purpose of the computer group was to control for the Hawthorne effect. Keyboard and computer lessons were provided individually each week for ten minutes, whereas group singing sessions occurred daily for 30 minutes. One hour each day was reserved for keyboard or computer practice. The children were tested using one spatial–temporal reasoning task and three spatial recognition tasks taken from the Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R) (Wechsler 1989) before and after instruction began. The spatial–temporal task required the children to arrange pieces of a puzzle to create a meaningful whole. Placing the pieces together in the correct order resulted in a higher score. The spatial recognition tasks involved matching, categorizing, and recognizing similarities of objects that were physically presented to the child. Sequential order is not relevant for these types of tasks. Although the pre-test scores of the four groups did not differ, children in the keyboard group scored significantly higher on the spatial–temporal task only. There were no differences between groups on the spatial–recognition tasks. The singing, computer and no-instruction groups did not improve significantly on any of the tests administered. It should be noted, however, that the musical experience of the children in the singing group was quite different from that of the keyboard group. The singing group did not receive individual instruction. A music instructor played the piano while the children simply sang along, whereas the keyboard group learned how to read music. One therefore cannot conclude from this study that vocal instruction would not affect spatial–temporal task performance.

In a related study, Gromko and Poorman (1998) examined the spatial–temporal task performance of children following seven months of weekly singing and songbell instruction. The children's spatial abilities were pre- and post-tested using five subtests of the WPPSI. A composite score was then computed from these five subtests. Results indicated that the music group's gain scores (post-test minus pre-test) were significantly higher than the control group's gain scores. The researchers concluded that music training affects the development of spatial intelligence in young children.

Several other studies have found improvement in children's spatial skills following music instruction (Bilhartz *et al.* 2000; Costa-Giomi 1999; Graziano *et al.* 1999; Orsmond and Miller 1999; Rauscher 2002; Rauscher and Zupan 2000; Zafranas 2004). A relatively recent meta-analysis synthesized the results of 15 independent studies and concluded that 'music instruction has ... been clearly shown to enhance at least one kind of spatial skill and may in the future be shown to enhance others' (Hetland 2000b, p. 226). However, other areas of intelligence may also be affected by music instruction. For example, Gardiner

and his colleagues provided children with seven months of Kodály music instruction along with visual arts training (experimental group) (Gardiner *et al.* 1996). Their reading and maths scores were then compared to children who received the school's standard arts curriculum (control group). Although the pre-test scores of the experimental group were lower than the control group's grade averages for reading and maths, after the special arts training they scored equal to the control group in reading and above them in maths. The same result was found when the children were tested following an additional seven months of special arts training. Unfortunately, because the music and arts instruction were provided together, it is not possible to determine if the results were due to the music curriculum.

A recent study found that early music instruction emphasizing different musical skills produced different effects on cognitive performance (Rauscher *et al.* 2007). At-risk preschool children received piano, singing, rhythm, computer, or no instruction for two years. All children were pre- and post-tested using standardized intelligence, visuospatial, and auditory–perceptual tests. Although pre-test scores of the five groups did not differ, the three music groups scored significantly higher following instruction than the control groups on mental imagery tasks. The scores of the rhythm group, however, were significantly higher than all other groups on tasks requiring temporal cognition and mathematical ability. This suggests that the type of music instruction children receive has measurably different effects on cognition, with instruction in rhythm having the strongest impact on temporal and sequencing tasks.

Other studies have found advances in other abilities following music instruction. A positive relationship between reading and music instruction has been reported by several researchers (Anvari *et al.* 2002; Butzlaff 2000; Douglas and Willatts 1994). However, most of these studies were correlational rather than experimental and therefore cannot determine anything about causality. See Hurwitz *et al.* (1975), described above, for an experimental approach to understanding music's effects on reading. In two separate studies, Overy and her colleagues reported that fifteen weeks of rhythm-based music instruction improved the phonemic awareness and spelling abilities of dyslexic children (Overy 2000, 2003). In addition, one study reported that children who received music lessons for two or more years scored higher than controls on a mathematics achievement test (Cheek and Smith 1999). Two studies have found relationships between music instruction and verbal memory (Chan *et al.* 1998; Ho *et al.* 2003).

These diverse findings led Schellenberg (2004) to wonder if music instruction might affect general intelligence. This indeed appears to be the case: 144 six-year-olds were randomly assigned to four groups: two music-lesson groups (keyboard or voice) and two control groups (drama lessons or no lessons). Lessons were provided over a period of 36 weeks in groups of six. The Wechsler Intelligence Scale for Children (WISC-III) (Wechsler 1991), a test of general intelligence, was administered before and after the 36-week period. The keyboard and voice groups showed small but significant increases in IQ compared to children in the drama and no-lesson control groups combined. Schellenberg then examined differences across groups in individual WISC-III subtests, such as verbal comprehension and perceptual organization, and found significant advantage for the consolidated music group on 10 of 12 scales. Based on the number of experimental studies showing effects of keyboard instruction on spatial abilities, Catterall and Rauscher (2008) theorized that the students who received music instruction in Schellenberg's study might show greater improvement in the scales related to spatial ability than in other scales. To test this, we compared the verbal IQ and performance (i.e., spatial) IQ scores of the music and control groups separately. Since the numerical scales for verbal and performance IQ differ, we calculated the effect sizes associated with participating in music by dividing the music group's gain scores by the respective standard deviations of the control group's pre-test scores on the same measure. Music participation in Schellenberg's experiment showed an effect size of 0.45 on verbal IQ and a somewhat larger effect size (0.55) on performance IQ. An effect size of 0.30 is considered moderate but significant. Effect sizes above 0.50 are considered robust. These data suggest that the increase in general intelligence affected by

music lessons was due to gains in visual–spatial intelligence more than to gains in verbal intelligence. We concluded that the impact of music on general intelligence may be driven by gains in spatial intelligence.

## Neurophysiological mechanisms

Playing a musical instrument requires instantaneous examination of what has already occurred in the performance (e.g., up-bow vs down-bow; fingering) as well as thinking ahead to prepare for future challenges. Researchers over the past few decades have examined the brain bases for these highly specialized skills. One of the most remarkable features about the brain is its ability to reorganize as a function of experience. Based on William James' suggestion at the end of the nineteenth century that learning may alter synaptic connectivity (James 1890), Donald Hebb formulated an important and innovative theory that views the brain as a dynamic system (Hebb 1949). Hebb's theory is often paraphrased as 'neurons that fire together wire together', although this oversimplification of the nervous system should not be taken literally. The theory is commonly evoked to explain learning in which simultaneous activation of cells leads to pronounced increases in synaptic strength—a phenomenon known as Hebbian learning. The advent of brain imaging technology has permitted remarkable gains in our understanding of how an individual's brain develops and produces uniquely human capacities. Imaging studies suggest that different amounts and types of environmental input can alter cortical structure and function. Specifically, an increasing amount of research suggests that early music instruction influences the brain, perception, and cognition.

The effects of music instruction are evident in anatomical differences in the brains of musicians versus non-musicians. Studies have shown that extensive practice on a musical instrument can induce cortical reorganization, perhaps due to Hebbian learning rules. This reorganization may consequently produce functional changes in how the brain processes information relevant to the particular area of expertise. If the extensive practice occurs early in development, the alterations in brain function may become hardwired and produce permanent changes in the way such information is processed. Such changes may represent the neural substrate of behavioural expertise. People who begin music instruction later may still develop considerable expertise, but perhaps without the cortical reorganization that occurs in those who begin earlier in life.

A study using structural magnetic resonance imaging reported that musicians who began instruction before age 7 had a larger cross-section of the anterior corpus callosum than non-musicians and musicians who began instruction after age 7 (Schlaug *et al.* 1995). No differences were found between non-musicians and musicians who began instruction after age 7, suggesting the importance of early music instruction for callosal plasticity. Schlaug and his colleagues proposed that the differences found reflect the need for greater interhemispheric communication of individuals performing complex bimanual motor sequences.

To examine differences in the motor cortex of musicians and non-musicians, researchers have examined the intrasulcal length of the posterior bank of the precentral gyrus (ILPG), a marker that roughly corresponds to the area of the motor cortex (Amunts *et al.* 1997). The researchers found a greater symmetry of the ILPG in the dorsal subregion of the motor cortex, a brain region that corresponds to the location of the functional hand motor area, of right-handed musicians (mostly pianists) compared to right-handed non-musicians. This symmetry is possibly due to a disproportionate increase in ILPG of the non-dominant (right) hemisphere as a function of the bimanual motor movements required for playing the piano. The ILPG was larger in both the left and right hemispheres in musicians than in non-musicians. The size of the ILPG in both hemispheres correlated positively with the age that music training commenced: pianists who began instruction at the earliest ages had the largest ILPGs.

A study using magnetoencephalography (MEG) provides further evidence for the influence of music instruction on plasticity. Dipole moments of the digits of the left hand were significantly larger in violinists compared to non-musicians (Elbert *et al.* 1995). Greater dipole

moments can indicate either a larger cortical representation for that function or enhanced coherence of nerve impulses in that region. Right-hand representation did not differ between string players and non-musicians. The largest effects were found for those who began before age 12. A follow-up MEG study found auditory cortex dipole moments for piano tones were enlarged in musicians relative to non-musicians (Pantev *et al.* 1998). Again, there was a positive correlation between effect size and when participants initiated instruction.

Taken together, these studies and others suggest that early music training alters brain structure and/or function relevant to cognition. However, it remains unclear whether these differences between the brains of musicians and non-musicians existed prior to music instruction, or whether they occurred as a function of music instruction. To determine whether these differences reflect plasticity due to music training or instead are markers of musical interest or aptitude, it would be informative to conduct an imaging study of children before music instruction began and again following the onset of instruction. The brain scans of those children could then be compared with those of children who are not studying a musical instrument. This study is currently underway (Norton *et al.* 2005). Using magnetic resonance imaging (MRI), the brains of young children were scanned prior to the onset of instrumental music instruction. Children who would not be receiving music instruction were also scanned. Spatial, verbal, motoric, and music perception tasks were also administered. Preliminary data revealed no pre-existing differences of any kind between the experimental and control groups. The children will continue to be examined on a yearly basis. The lack of any pre-existing differences in brain structure and cognitive task performance suggests that the morphological and cognitive dissimilarities between musically trained and untrained individuals may be an outcome of music instruction.

## Cognitive transfer

Transfer occurs when a person applies knowledge or skills that have been learned in one context to new contexts. Researchers interested in transfer were initially guided by theories that emphasized the similarities between the initial learning experience and later learning. Thorndike (1913) proposed that the amount of transfer that could occur between two domains was dependent upon the similarity of the elements of the domains. The more equivalent the elements of the two domains, the greater the likelihood of positive transfer. Thus, transfer is always a function of the relationship between what is learned and what is tested.

Psychologists draw a distinction between near and far transfer. Near transfer involves applying the same knowledge or skill in very similar circumstances, such as using one's car-driving skills to drive a truck. Far transfer implies a big leap. A chess master carrying the notion of control of the centre from chess to political or military settings would count as far transfer. Not only are political and military settings quite different in general from chess despite chess's underlying military metaphor, but the meaning of 'centre' plays out differently. In chess, it is the centre area of the board. In the military context, centre might mean a centre of command, communications, munitions, or supply, for example. Any of these would make tempting targets.

Salomon and Perkins (1989) point out two distinct mechanisms of transfer, the low road and the high road. Low-road transfer occurs when perceptual similarities of one situation to another trigger the making of a connection. For example, consider a professional figure-skater who switches to roller-blading. The familiar action of pushing off on one foot, then gliding, and then pushing off with the other foot evoke all the old figure-skating habits. Low-road transfer depends on abundant practice with the skill or knowledge in question to set up the perceptual triggering. It is a phenomenon of experience. Consequently, there is not much far low-road transfer. Far low-road transfer can occur, however, when the same skill gets practiced over a variety of circumstances, gradually stretching from one context to another until it achieves high generality.

High-road transfer, on the other hand, depends upon the learners' deliberate and mindful abstraction of a principle. The chess master that applies chess skills to a military setting is engaging in high-road transfer from chess to

military strategy. In contrast to low-road processing, which is relatively spontaneous, high-road processing requires reflection. The transfer of musical skills to, for example, spatial–temporal tasks may therefore be an example of low-road far transfer. To determine if this is indeed the case one would need to somehow measure the overlap between the original domain of learning (e.g., playing a musical instrument) and the novel one (e.g., performing a spatial–temporal task). This undertaking would require a theory of how knowledge is represented and conceptually mapped across the domains.

Singley and Anderson (1989) argue that transfer between tasks is a function of the degree to which the tasks share *cognitive* elements. This hypothesis is hard to test experimentally until the task components are identified. Thus, a complete understanding of spontaneous transfer from music to another domain of reasoning (e.g., spatial–temporal reasoning or arithmetic) is possible only to the extent that the cognitive elements of the two domains can be identified. For example, the part–whole concept is a very important construct for many mathematical problems. This concept requires understanding the relationship between parts to wholes, such as when learning percentages, decimals, and fractions. In music, the part–whole concept is especially relevant in the conceptualization of rhythm. A literate musician is required to continually mentally subdivide the beat to arrive at the correct interpretation of rhythmic notation. The structure of the problem is essentially the same as any part–whole problem posed mathematically. Perhaps this relationship helps explain the finding that children who received instruction on rhythm instruments scored higher on part–whole mathematics problems than those who received piano or singing instruction (Rauscher *et al.* 2007). Further investigation into the components common to musical and mathematical knowledge will aid in the understanding of these possible transfer effects.

Schellenberg (2003) suggests that

> the ability to attend to rapidly changing temporal information, skills relevant to auditory stream segregation, the ability to detect temporal groups, sensitivity to signals of closure and other gestalt cues of form, emotional sensitivity and fine motor skills ... should be particularly likely to transfer to a variety of nonmusical domains.
>
> (p. 444)

Likewise, Norton *et al.* (2005, p. 131) state that the

> skills such as decoding visual information into motor activity, memorizing extended passages of music, learning music structures and rules, learning to make fine auditory spectral and temporal discriminations, and learning to perform skilled bimanual finer movements

may all contribute to the transfer of musical knowledge to other cognitive abilities. For example, both music and mathematics employ and manipulate symbols, they both investigate and develop patterns (numerical in the case of math, and tonal or rhythmic for music), and they are both abstract constructions. A trained musician finds order in what is essentially arbitrarily organized sound, and has conceptually constructed a system of patterns, relationships, regularities, series, proportions, fractions, subdivisions, and ratios, represented and understood with a complex array of symbols. Thus, a deep structural level overlap between musicianship and mathematical knowledge could be expected in a number of areas. Research grounded in transfer theory has indeed shown relationships between music instruction and a variety of cognitively related skills (see, for example, Gromko 2004). Although studies specifically testing transfer as a mechanism are extremely difficult to implement due to an insufficient understanding of the overlap of the cognitive components inherent in the two domains, I suggest that transfer remains a potential explanation for improved cognitive abilities following music instruction.

## Conclusion

Traditionally, school curricula have been characterized by hard boundaries between disciplines. The studies presented in this chapter suggest that curriculum reform could aim at the development of more permeable boundaries between disciplines, so that spontaneous transfer

of skills and knowledge is facilitated. Zohar (1994) argues that thinking skills are likely to be enhanced if students receive explicit instruction to recognize the underlying logical structures of domains that underpin transfer. In addition to explicit teaching, it is possible that students' thinking and learning could be enhanced through analysis of the structural elements of disciplines that appear distinct on the surface but that are conceptually and logically similar at deep structural levels.

I conclude this chapter as I began it: with a word of caution. Although this area of study certainly seems worthy of further investigation in educational settings, we must be careful when applying the findings of music transfer to educational practice. Indeed, this research represents an excellent example of what has become known as a 'design experiment'. Design experiments are described as educational research experiments carried out in a complex learning context to determine how an innovation affects student learning and educational practice (Cobb *et al.* 2003). Real-life educational contexts are, in turn, excellent settings for experimental tests of an intervention. I believe such experiments represent a crucial research approach within the broader context of partnerships involving teachers, educational researchers, and scientists. I do not, however, advocate teaching music to students in order to improve their visuospatial or mathematical skills. The research base on studies showing enhancement of cognitive function following music instruction is still in its infancy, and there is certainly more to learn. The studies presented here are of scientific importance because they suggest that music and other areas of intelligence share common elements and may be psychologically and neurologically related. I believe researchers should continue to search for links between music instruction and cognitive performance, but that these studies should not be placed at the top of the advocacy agenda. In a recent *Boston Globe* article on arts education, Winner and Hetland (2007) state that

> We don't need the arts in our schools to raise mathematical and verbal skills—we already target these in maths and language arts. We need the arts because in addition to introducing students to aesthetic appreciation, they teach other modes of thinking we value.
>
> (2007, p. E1)

Music education should not have to prove its value by showing that it has non-musical outcomes.

## References


Amunts K, Schlaug G, Jancke L, Dabringhaus A, Steinmetz H, Schleicher A and Zilles K (1997). Motor cortex and hand motor skills: structural compliance in the human brain. *Human Brain Mapping*, **5**, 206–215.

Anvari SH, Trainor LJ, Woodside J and Levy BZ (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology*, **83**, 111–130.

Bilhartz TD, Bruhn RA and Olson JE (2000). The effect of early music training on child cognitive development. *Journal of Applied Developmental Psychology*, **20**, 615–636.

Butzlaff R, (2000). Can music be used to teach reading? *Journal of Aesthetic Education*, **34**, 167–178.

Carstens CB, Huskins E and Hounshell GW (1995). Listening to Mozart may not enhance performance on the Revised Minnesota Paper Form Board Test. *Psychological Reports*, **77**, 111–114.

Catterall JS and Rauscher FH (2008). Unpacking the impact of music on intelligence. In W Gruhn, FH Rauscher, eds, *Neurosciences in music pedagogy*, 171–201. Nova Science Publishers, New York.

Chabris C., (1999). Prelude or requiem for the 'Mozart effect'? *Nature*, **400**, 826–827.

Chan AS, Ho YC and Cheung MC (1998). Music training improves verbal memory. *Nature*, **396**, 128.

Cheek JM and Smith LR (1999). Music training and mathematics achievement. *Adolescence*, **34**, 759–761.

Cobb P, Confrey J, diSessa A, Lehrer R and Schauble L (2003). Design experiments in educational research. *Educational Researcher*, **32**(1), 9–13.

Costa-Giomi E (1999). The effects of three years of piano instruction on children's cognitive development. *Journal of Research in Music Education*, **47**, 198–212.

Douglas S and Willats P (1994). The relationship between musical ability and literacy skills. *Journal of Research and Reading*, **17**, 99–107.

Elbert T, Pantev C, Wienbruch C, Rockstroh B and Taub E (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, **270**, 305–307.

Gardiner MF, Fox A, Knowles F and Jeffrey D (1996). Learning improved by arts training. *Nature*, **381**, 284.

Graziano AB, Peterson M and Shaw GL (1999). Enhanced learning of proportional math through music training and spatial–temporal training. *Neurological Research*, **21**, 139–152.

Gromko JE (2004). Predictors of music sight-reading ability in high school wind players. *Journal of Research in Music Education*, **52**, 6–15.

Gromko J and Poorman A (1998). The effect of music training on preschoolers' spatial–temporal task performance. *Journal of Research in Music Education*, **46**, 173–181.

Hebb DO (1949). *The organization of behavior*. John Wiley and Sons, Chichester, UK.

Hetland L (2000a). Listening to music enhances spatial–temporal reasoning: evidence for the Mozart effect. *Journal of Aesthetic Education*, **34**, 105–148.

Hetland L, (2000b). Learning to make music enhances spatial reasoning. *Journal of Aesthetic Education*, **34**, 179–238.

Hetland L and Winner E (2001). The arts and academic achievement: what the evidence shows. *Arts Education Policy Review*, **102**, 3–6.

Ho YC, Cheung MC and Chan AS (2003). Music training improves verbal but not visual memory: cross-sectional and longitudinal explorations in children. *Neuropsychology*, **17**, 439–450.

Hurwitz I, Wolff PH, Bortnick BD and Kokas K (1975). Nonmusical effects of the Kodály music curriculum in primary grade children. *Journal of Learning Disabilities*, **8**, 167–174.

Ivanov VK and Geake JG (2003). The Mozart effect and primary school children. *Psychology of Music*, **31**(4), 405–413.

James W (1890). *Principles of psychology*. Holt, New York.

McKelvie P and Low J (2002). Listening to Mozart does not improve children's spatial ability: final curtains for the Mozart effect. *British Journal of Developmental Psychology*, **20**, 241–258.

Norton A, Winner E, Cronin K, Overy K, Lee DJ, Schlaug G., (2005). Are there pre-existing neural, cognitive, or motoric markers for musical ability? *Brain and Cognition*, **59**, 124–134.

Orsmond GI and Miller LK (1999). Cognitive, musical, and environmental correlates of early music instruction. *Psychology of Music*, **27**, 18–37.

Overy K (2000). Dyslexia, temporal processing, and music: the potential of music as an early learning aid for dyslexic children. *Psychology of Music*, **28**, 218–229.

Overy K (2003). Dyslexia and music: from timing deficits to musical intervention. *Annals of the New York Academy of Science*, **999**, 497–505.

Pantev C, Oostenveld R, Engelien A, Ross B, Roberts LE and Manfried H (1998). Increased auditory cortical representation in musicians. *Nature*, **392**, 811–813.

Rauscher FH (2002). Mozart and the mind: factual and fictional effects of musical enrichment. In J Aronson, ed., *Improving academic achievement: impact of psychological factors on education*, 269–278. Academic Press, New York.

Rauscher FH and Hinton SC (2006). The Mozart effect: music listening is not music instruction. *Educational Psychologist*, **41**, 233–238.

Rauscher FH and Shaw GL (1998). Key components of the 'Mozart effect'. *Perceptual and Motor Skills*, **86**, 835–841.

Rauscher FH and Zupan M (2000). Classroom keyboard instruction improves kindergarten children's spatial–temporal performance: a field experiment. *Early Childhood Research Quarterly*, **15**, 215–228.

Rauscher FH, LeMieux M and Hinton SC (2007). Lasting improvement of at-risk children's cognitive abilities following music instruction. Manuscript submitted for publication.

Rauscher FH, Shaw GL and Ky KN (1993). Music and spatial task performance. *Nature*, **365**, 611.

Rauscher FH, Shaw GL, Levine LJ, Wright EL, Dennis WR and Newcomb R (1997). Music training causes long-term enhancement of preschool children's spatial–temporal reasoning abilities. *Neurological Research*, **19**, 1–8.

Salomon G and Perkins DN (1989). Rocky roads to transfer: rethinking mechanisms of a neglected phenomenon. *Educational Psychologist*, **24**, 113–142.

Schellenberg EG, (2003). Does exposure to music have beneficial side effects? In R Peretz, RJ Zatorre, eds, *The cognitive neuroscience of music*, 430–448. Nova Science Press, New York.

Schellenberg EG, (2004). Music lessons enhance IQ. *Psychological Science*, **15**, 511–514.

Schellenberg EG and Hallam S (2005). Music listening and cognitive abilities in 10- and 11-year-olds: the Blur effect. *Annals of the New York Academy of Science*, **1060**, 1–8.

Schlaug G, Jancke L, Huang Y, Staiger JF and Steinmetz H (1995). Increased corpus callosum size in musicians. *Neuropsychologia*, **33**, 1047–1055.

Singley K and Anderson JR (1989). *The transfer of cognitive skill*. Harvard University Press, Cambridge, MA.

Thorndike EL, (1913). *Educational psychology*. Columbia University Press, New York

Wechsler D, (1989). *Wechsler Preschool and Primary Scale of Intelligence-Revised*. The Psychological Corporation, San Antonio, TX.

Wechsler D (1991). *Wechsler Intelligence Scale for children*, 3rd edn. The Psychological Corporation, San Antonio, TX.

Winner E and Hetland L (2007). *Art for our sake. Boston Globe*, 2 September.

Zafranas N (2004). Piano keyboard training and the spatial–temporal development of young children attending kindergarten classes in Greece. *Early Child Development and Care*, **174**, 199–211.

Zohar A (1994). Teaching a thinking strategy: transfer across domains and self-learning versus class-like setting. *Applied Cognitive Psychology*, **8**, 549–563.

# PART 6

# Learning musical skills

Edited by Susan Hallam

# CHAPTER 24

# Musical potential

Gary McPherson and Susan Hallam

## Introduction

There is general agreement that music is a universal trait of humankind (Blacking 1995), that *Homo sapiens* as a species has the propensity for musical development, and that musical potential is as universal as linguistic ability (Wallin *et al.* 2000). An ongoing controversy persists however, concerning the extent of individual variability in musical potential and the extent to which observable differences in acquiring musical skills result from social contexts that facilitate learning, genetic factors, or interactions between the two. This chapter outlines key elements of these debates and also considers how 'musical potential' has been assessed.

## The nature–nurture debate

Francis Galton (1876) pioneered the study of genetic influences on learning and development using evidence from twin studies to argue that traits leading to eminence were largely inherited. However, recent research suggests that there are complex interactions between the environment and genetic factors which influence observable behaviour, with genetic factors having a closer association with physical attributes than psychological factors, and that many dimensions of an individual's development, such as memory, language development and intelligence, can be enhanced through systematic practice and learning (Gross 2005). Research attempting to establish the extent of the heritability of musical potential has reached similar conclusions with no decisive evidence showing that it is directly dependent on aural acuity, intelligence, or other types of artistic ability (for reviews see Hodges 1996; Shuter-Dyson 1999). A more plausible explanation is that musical development is the result of a range of gene combinations interacting with environmental stimulation in an interactive rather than additive manner (Ceci 1990). Evidence that the cerebral cortex has an amazing ability to self-organize in response to stimuli such as music supports this view (Rauschecker 2003). Cortical activation during music processing reflects personal musical experiences accumulated over time including listening to music, learning to play an instrument, formal instruction, and professional training resulting in multiple mental representations of music that are, in part, interchangeable and rapidly adaptive (Altenmuller 2003). While self-selection for musicianship by individuals with innate functional and structural brain differences cannot be completely ruled out, the evidence indicates that it is musical training that leads to changes in brain function and structure (Schlaug 2003). As genetic inheritance can clearly be enhanced by a musically enriched social context, considerable research is now focused on identifying the environmental factors that facilitate or impede musical development.

### Musical savants and Williams syndrome

The most difficult phenomena to explain without resorting to some notion of inherited differences in musical potential are children at the extremes of neurodevelopment, for instance, musical savants, children with Williams syndrome, and child prodigies.

Savants have generally low cognitive functioning but are able to achieve at normal levels in some activities, especially those related to 'non-symbolic, artistic, visual and motor abilities'

such as 'music, art, maths and various forms of calculating (such as calendar-counting)' (Gross 2005, p. 685). Many musical savants exhibit absolute pitch (Miller 1989) enabling them to make confident, rapid judgements about individual pitches and complex chords. They are also sensitive to rules reflecting harmonic relationships and the structure of musical compositions (Young and Nettelback 1995; Sloboda *et al.* 1985). Explaining these skills without resorting to genetic explanations is difficult. However, environmental influences should not be underestimated. Many savants have limited sight and language disorders, which may lead to increased development of auditory processing skills and the use of music as a means of communication. They also spend considerable time practising their skills.

Individuals with Williams syndrome have low measured intelligence and difficulties with mathematical and spatial reasoning but are more adept than might be expected in language and music, the development of the latter depending on access to appropriate musical opportunities. Levitin and colleagues (Levitin and Bellugi 1998; Levitin *et al.* 2003, 2007) have shown that these children are typically as musically accomplished, engaged and interested as ordinary children but display greater emotional responses to music, become interested at a younger age, spend more time listening to music, instinctively experience music much more fully than others, and possess a highly sensitive emotional attachment to music.

## Musical prodigies

Prodigies are children who, from an early age, display exceptional talent. Famous examples include Mozart, Bach, Beethoven and Mendelssohn whose significant success in later life depended on having undertaken considerable focused and well-directed effort during their early years (McPherson and Williamon 2006). Ruthsatz and Detterman (2003) identified a recent example of a prodigy, a 6-year-old who despite having had no formal tuition acquired considerable musical skill by imitating other performers and improvising his own musical pieces. He could sing in two languages, had taught himself to play numerous instruments, had an Intelligence Quotient (IQ) of 132, an extraordinary memory, and attained a high score on Gordon's (1982) music aptitude measure. His exceptional musical behaviours were self-motivated and spontaneous and he particularly liked entertaining people. His musical abilities were closely aligned with his extraordinary memory and high IQ, more so than with the time or type of practice he undertook. As we will see below, such early, specific situational behaviours that spark changes in cognitive development to allow a child to be able to direct extremely high levels of attention toward music are becoming of great interest to researchers.

The overt musical behaviours of savants and Williams syndrome individuals share some similarities with the behaviour of prodigies, even though their neurodevelopmental trajectories differ. Neuroconstructivists suggest that typical and atypical development can be viewed as different trajectories in a continuum of possibilities. An atypically developing trajectory affects the interactions of others with the child and the kind of experiences that the child seeks out, which further impact on the trajectory (Mareschal *et al.* 2007). For example, when parents believe that their child has musical ability, they are more inclined to provide musical resources and reward musical activity, which in turn supports increasing levels of expertise as the child engages more fully with music (McPherson, in press). These social dynamics result in the child developing particular neural structures that make further musical development much easier (Altenmüller and Gruhn 2002; Hodges 2006). Familial responses of this type may occur in relation to savants, Williams syndrome children, and prodigies.

Other research suggests that the typical 'rage to master' which characterizes prodigies can be explained as a result of domain-specific high attentional control that begins in infancy to produce a spontaneous version of deliberate practice (Vandervert and Liu in press). This view suggests that the high attentional control of prodigies originates and then accelerates connections between the cerebral cortex (where mental modelling construction and repetition occur) and the cerebellum (where model formation occurs), such that cerebellar control models feedback to the working memory areas

of the cortex. In this way the child prodigy's working memory becomes faster, more concentrated, and more efficient (Vandervert 2007). This explains the behaviour of these individuals in terms of 'the *reciprocal* learning relationships between the anticipatory, *adaptive* cognitive-affective and attentional modelling functions of the cerebellum and those of the cerebal cortex' (Vandervert and Liu in press).

Shavinina (in press) proposes that extreme levels of giftedness occur as a result of stimulation and activation early in life when the developing child selectively responds in ways that heighten his or her cognitive, emotional, and social sensitivities. Such sensitive periods in these children's early years provide the foundations for giftedness in that they accelerate the gifted child's mental development through the actualization of intellectual potential and cognitive experience (Shavinana 2007). Cognitive experience of this type provides the psychological basis from which highly gifted children are able to develop their creative, metacognitive, and extracognitive (i.e., feelings, beliefs, intellectual values, intuition) abilities.

The above explanations support an interactive, dynamic model of how exceptional achievement in music develops as a result of environmental forces acting together with innate potentials at critical moments in a child's development. The following sections attempt to frame these conceptions within specific areas of musical engagement.

## The role of learning in the development of musical expertise

Research undertaken within the expertise paradigm has also challenged previously accepted notions that high-level achievement depends exclusively on inherited ability. The basic premise of this theory is that time spent on 'deliberate practice' underpins the development of high-quality expert performance. For instance, it has been established that classical Western musicians need to have accrued up to 16 years of practice to achieve levels that will lead to international standing in playing an instrument. Such individuals usually begin playing at a very early age and over succeeding years increase the amount of practice undertaken, sometimes up to as much as 50 hours a week by adolescence (Sosniak 1985).

Ericsson and colleagues (1993) have suggested a monotonic relationship between 'deliberate practice' (which they define as goal-oriented, structured and effortful practice that is influenced by motivation, resources and attention) and an individual's acquired performance (for review see Chapter 25 this volume, Lehmann and Gruber 2006). This is supported by evidence that musicians with the highest levels of expertise accumulate considerably more hours of practice than their less successful peers, although there are substantial individual (Jørgensen 2002; Sloboda *et al.* 1996), instrumental and genre differences (Kopiez 1998; Gruber *et al.* 2004; Jørgensen 2002).

While researchers agree that practice is important in facilitating the development of expertise, several studies question the simplicity of a monotonic relationship. Sloboda and Howe (1991) found that students identified as having greater ability by their teachers had spread their practice across each of their instruments and therefore undertaken less practice on their main instrument. Wagner (1975) found that increased practice did not lead to any greater improvement in performance over an eight-week period, and Zurcher (1972) found no relationship between total practice time and performance achievement.

Reported correlations between achievement and time spent practising vary between 0.25 (Doan 1973 and 0.67 (Hallam 1998a). In the Hallam study, the correlation rose to 0.84 when years of time learning was correlated with achievement, as opposed to time spent practising. It seems therefore, that the overall length of time over which learning has taken place may be as important as the actual amount of practice in determining level of expertise. This was especially evident in a causal model developed by McPherson *et al.* (1997) which shows a strong association between the length of time learning and taking lessons and high school musicians' ability to sight-read and perform music that they had rehearsed over the previous weeks and months.

Accumulated practice from the time of beginning learning to the present does not seem to

predict the quality of performance at any point in time (Hallam 1998a; Williamon and Valentine 2000). Other factors such as teachers' ratings of musical ability, self-esteem, and involvement in extra-curricular music activities are better predictors of achievement (Hallam 2004). A longitudinal study with beginning instrumentalists also showed that accumulated practice only partly explains children's ability to perform rehearsed music and sight-read and none of their ability to memorize music, play by ear, or improvise. McPherson (2005) showed that accumulated practice explained between 9 and 32 per cent of the variance in the learner's ability to perform rehearsed repertoire over their first three years of learning, and even less for their sight-reading ability. Other skills, such as the sophistication of the mental strategies which the young players adopted to guide their playing, were more important (see also Chapter 25 this volume).

To date, much of the research has failed to take account of the amount of time spent acquiring musical skills through listening to music, engaging in playful musical activity, and participating in group activities where learning and consolidation of skills occurs in an informal learning context (e.g. Kokotsaki and Hallam 2007). In addition, research has neglected those who may have undertaken considerable amounts of practice but have dropped out of music instruction. There are complex relationships between prior knowledge, motivation, effort and perceived efficacy which influence decisions to continue or discontinue learning (Hallam 1998a; Sloboda *et al.* 1996). When a child begins to learn an instrument, prior musical knowledge affects facility of learning and the time needed to achieve mastery. While undertaking additional practice may compensate for a lack of prior knowledge, this has a time cost and requires perseverance. If a task proves too challenging, then a child may perceive that the effort required to succeed is too great and may give up learning altogether (Hurley 1995). Difficulties may also be evident when a child perceives that he or she does not have sufficient ability. Such perceptions often lead to a loss of self-esteem, loss of motivation, less practice, and a downward spiral leading to the termination of lessons (Austin *et al.* 2006; Chandler *et al.* 1987).

Another way to understanding individual differences is to focus on the personal beliefs held by learners and their parents. Parental and child ability conceptions are recognized as having a major impact on motivation and the desire of children to continue learning, especially when faced with obstacles (Austin *et al.* 2006). Indeed, McPherson's (in press) review of literature on mother–child interactions shows that parental ability conceptions can be self-fulfilling. McPherson and Davidson (2002) interviewed mothers before and after their child commenced learning an instrument. Those who held fixed views that their child may not have sufficient ability to succeed musically tended to provide less support for practice than mothers whose view was more malleable. They were also more likely to encourage their child to pursue other activities when they came to believe that their child was not coping. These fixed views of musical potential explained why some of the unsuccessful learners came to feel that they did not have sufficient ability to cope with learning music. Some mothers actually gave up on their child as a potential musician much sooner than the child.

## The development of general musical skills

The impact of parents and the home environment is of profound importance in the development of children's musical potential (McPherson, in press). The general milieu of the environment to which a child is exposed and the opportunities parents and significant others provide are among the most critical factors for realizing children's musical potential (McPherson and Williamon 2006)(see also Chapter 28 this volume).

The seeds of musical potential are sown early because the human auditory system is functional 3–4 months before birth. After 28 to 30 weeks of gestation, foetuses can reliably react to external sounds, such that their heart rates vary as a result of exposure to music (Parncutt 2006). The process of musical enculturation therefore begins from the point at which brain development starts to become influenced by auditory stimulation. These processes gain momentum in the minutes after birth when a mother and infant will imitate each other's vocalizations in

ways that show a shared emotional experience that some believe is the very basis of musicality (Trevarthen and Malloch 2002).

In the immediate months after an infant is born, the complex skills required for understanding and analysing music within any particular culture start to develop as a result of ongoing exposure to music. Even though each infant will experience many different levels of exposure, the perceptual and motor control needed to sing an intended pitch and the self-monitoring necessary to notice pitch or rhythmic differences can be accelerated by training (Trehub 2006). Wide individual differences exist regarding the extent to which preschool children engage in singing. For some, it is a part of almost all activities, while others sing only occasionally (Sundin 1997). Even though their exposure to music will have been markedly different, all children will have typically developed a surprising array of internal schemata for particular types of music long before they reach school or begin formal music instruction.

## Measuring musical potential

Historically, the developers of musical aptitude tests have held varying views regarding the heritability of musical ability. Seashore (1938a, Seashore *et al.* 1938) believed that musical ability was a set of loosely related basic aural discrimination skills, which had a genetic basis and did not change over time. Wing (1981), Drake (1957) and Bentley (1966) shared Seashore's view of musical ability as being inherited, although they differed in their conceptions of the nature of that ability and how it might be assessed.

More recent tests, based on measurements involving tonal (melody, harmony), rhythm (tempo, metre), and preference (phrasing, balance, style) aptitudes, have been devised by Gordon (1965, 1979, 1982). Gordon (2007) suggests that students rarely display high (or low) aptitudes in all seven aptitudes and that all are based on the ability to 'audiate'; a term he has coined to describe how individuals give meaning to music that is heard or imagined.

The rationale underlying all of these approaches is that 'musicality' has its basis in aural perception. However, the predictive reliability of all of these measures is generally low (Hodges and Haack 1996). While alternative, more active measures for selecting pupils for learning to play an instrument have been adopted by teachers, these have tended not to be formalized. Perhaps the most common of these has been selection based on the child's ability to sing. However, the relationship between developmental tonal aptitude and use of the singing voice may also be very small (Rutkowski 1996). Researchers now generally recognize that aural skills alone are insufficient to predict future success across the full range of musical activities, especially those involving motor skills (Gilbert 1981) and creativity (Vaughan 1977; Webster 1988).

## Recent conceptions of musical potential

McPherson and Williamon (2006) adapted Gagné's (2003) *differentiated model of giftedness and talent* to music as a means of defining the natural innate abilities, intrapersonal factors, and environmental catalysts that might impact on the development of musical skills. This conception defines *gifts* (e.g., intellectual, creative, socio-affective, sensorimotor) as natural innate potentials to achieve and *talent* as observable skills. The framework proposes that at least eight distinct types of musical talents (performing, improvising, composing, arranging, analysing, appraising, conducting, music teaching) can be developed through systematic practice and training. Moreover, although physical and mental dexterity, musicality, motor memory and auditory memory are all evident in the first few weeks of formal musical training, each needs to be refined and developed further through extensive practice and learning for children to develop their musical talents. Another key element of the theory is that some types of talents can go unnoticed or, as in the case of composing, develop later than others. By including a range of non-performance based outcomes of musical involvement, the model represents a broader conception of musical potential than has been evident in the past.

As an extension, McPherson (1993, 1996) distinguishes between visual (i.e., sight-reading, performing rehearsed music from notation), aural (i.e., playing from memory and by ear)

and creative (i.e., improvising) aspects of music performance. When considering these in relation to groups of children of differing ages and abilities, he found that different musical skills are involved in developing each of these ways of performing music, that there is not an automatic transfer between the three orientations, and that each needs to be developed separately and in combination to maximize potential. This is rarely the case in most formal learning situations, where visual forms of performance often dominate at the expense of other orientations.

From an even wider perspective, Hallam (1998b) suggests that learning to play a musical instrument depends on the development of a wide range of different professional and personal skills. Some of these may be required for all musical activities, and others are applied more selectively to particular tasks (see Table 24.1). In order to become a successful musician, individuals need to develop social skills (being able to work with other musicians, promoters, the public); planning and organizational skills (planning practice schedules, programmes, travel

**Table 24.1** Musical skills

| | |
|---|---|
| Aural skills required for: | Developing rhythmic accuracy and a sense of pulse; |
| | Good intonation; |
| | The facility to know how music will sound without having to play it; |
| | Improvisational skills. |
| Cognitive skills required in the processes of: | Reading music; |
| | Transposition; |
| | Understanding keys; |
| | Understanding harmony; |
| | Understanding the structure of music; |
| | The memorization of music; |
| | Composing; |
| | Understanding different musical styles and their cultural and historic contexts. |
| Technical skills required for developing: | Instrument specific skills; |
| | Technical agility; |
| | Articulation; |
| | Expressive tone quality. |
| Musicianship skills are concerned with: | Being able to play expressively; |
| | Being able to project sound; |
| | Developing control; |
| | Conveying meaning. |
| Performance skills include: | Being able to communicate with an audience; |
| | Communicating with other performers; |
| | Being able to coordinate a group; |
| | Presenting to an audience. |
| Learning skills are concerned with: | Being able to learn, monitor and evaluate progress independently. |

arrangements); and time management skills (being punctual, meeting deadlines). These are clearly required for developing expertise in a range of professions and while necessary are not exclusively 'musical'.

## Actualization of musical potential

In the modern world, children have greater access to music through the media and are able to learn music in a multitude of different ways from the past. Technological developments have resulted in changes to the way music is perceived and valued within society, such that mere aural perception is no longer regarded as the defining aspect of musical ability.

Haroutounian (2000), in analysing the level of importance attached to particular criteria in identifying musically able children, suggests that the general behaviours of 'sustained interest' and 'self-discipline' are more important than music-specific characteristics which are normally regarded as indicative of musical ability. Similarly, Hallam and Prince (2003) asked a sample of musicians, student musicians, educators, and the general public to define musical ability. They reported that 71 per cent of the respondents viewed musical ability as being able to play a musical instrument or sing. This finding suggests that musical ability is often identified on the basis of developing practical skills. Overall, 28 per cent of the sample mentioned aural skills as indicative of musical ability, 32 per cent included listening and understanding, 24 per cent having an appreciation of music, and 15 per cent being responsive to music. The integration of a range of skills was cited by only 9 per cent of respondents. Personal qualities including motivation, personal expression, immersion in music, total commitment, and metacognition (being able to learn how to learn) were cited most often by musicians. Unsurprisingly, the musicians gave more complex responses, with many more elements in their explanations.

Further work by Hallam and Shaw (2003) using rating scales to illicit responses to a set of statements about musical ability showed that it was conceptualized in relation to rhythmic ability, organization of sound, communication, motivation, personal characteristics, the integration of a range of complex skills, and performing in a group. Having a musical ear ranked lower in responses than might have been expected given its prominent position with regard to musical ability historically. The high ratings given to motivation and personal commitment demonstrate their importance in developing high-level skills. Overall, the conceptions of musical ability generated by the research were complex and multifaceted, and they reflected the wide range of expert achievement that occurs in the music professions of the developed world.

## Conclusions

The extent to which genetic endowment underpins or limits all subsequent musical development has and continues to be fiercely debated (see Hallam 2006; McPherson and Williamon 2006; *High Ability Studies* Volume 18; *Behavioral and Brain Sciences* Volume 21), although there is general consensus that human beings as a species are pre-programmed to acquire a wide range of musical skills. We argue that what children are born with *enables* rather than *constrains* what they will eventually be able to achieve. While a range of generalized abilities may come into play when learning music, a host of environmental and personal catalysts work in combination with teaching and learning processes to develop particular types of talent. These talents form the basis of the many varied ongoing professional, amateur, and informal forms of meaningful engagement that individuals can have with music.

In developed countries, where formal schooling has taken over some of the traditional roles of the family, tests of musical aptitude have been devised to facilitate the identification of children who might benefit from music instruction, or to provide a base line for catering for individual differences after instruction has commenced. In our view, the rationale that underpins approaches based exclusively on aural acuity is questionable, especially given current methodologies and technologies which do not enable us to state with any certainty whether observed differences in musical ability in children are the result of genetic inheritance or learning.

As noted by Kemp and Mills (2002), musical potential is a complex phenomenon that involves many factors. While aural abilities are undeniably important, they do not provide the basis from which to accurately assess a child's current or future musical potential. Instead, musical potential is best thought of as malleable and ever-changing, and a dimension of human experience that takes many forms and occurs at many different levels. As a species-specific behaviour, music is inextricably linked to our basic human design (Welch and Adams 2003), therefore all children are inherently musical and deserve access to the types of informal and formal experiences that will maximize their own, individual musical potential.

## References


Altenmüller E and Gruhn W (2002). Brain mechanisms. In R Parncutt and GE McPherson, eds, *The science and psychology of music performance: creative strategies for teaching and learning*, 63–81. Oxford University Press, New York.

Altenmuller EO (2003). How many music centres are in the brain? In I Peretz and R Zatorre, eds, *The cognitive neuroscience of music*, 346–356. Oxford University Press, Oxford.

Austin J, Renwick J and McPherson GE (2006). Developing motivation. In GE McPherson, ed., *The child as musician: a handbook of musical development*, 213–238. Oxford University Press, Oxford.

Bentley A. (1966). *Measures of musical abilities*. NFER-NELSON, Windsor, England.

Blacking J (1995). *Music, culture, and experience*. University of Chicago Press, Chicago, IL.

Ceci SJ (1990). *On intelligence ... more or less: a bio-ecological treatise on intellectual development*. Prentice Hall, Englewood Cliffs, NJ.

Chandler TA, Chiarella D and Auria C (1987). Performance expectancy, success, satisfaction and attributions as variables in band challenges. *Journal of Research in Music Education*, **35**, 249–258.

Doan GR (1973). An investigation of the relationships between parental involvement and the performance ability of violin students. Unpublished Doctoral dissertation, Ohio State University, Columbus, OH.

Drake RM (1957). *Manual for the Drake musical aptitude tests*, 2nd edn. Science Research Associates, Chicago, IL.

Ericsson KA, Krampe RT and Tesch-Romer C (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, **100**, 363–406.

Gagné F (2003). Transforming gifts into talents: the DMGT as a developmental theory. In N Colangelo and GA Davis, eds, *Handbook of gifted education*, 3rd edn, 60–74. Allyn and Bacon, Boston, MA.

Galton F (1876). The history of twins as a criterion of the relative powers of nature and nurture. *Royal Anthropological Institute of Great Britain and Ireland Journal*, **6**, 391–406.

Gilbert JP (1981). Motoric music skill development in young children: a longitudinal investigation. *Psychology of Music*, **9**(1), 21–25.

Gordon EE (1965). *Musical aptitude profile*. GIA, Chicago, IL.

Gordon EE (1979). *Primary measures of music audiation*. GIA, Chicago, IL.

Gordon EE (1982). *Intermediate measures of music audiation*. GIA, Chicago, IL.

Gordon EE (2007). *Learning sequences in music: a contemporary music learning theory*. GIA: Chicago, IL.

Gross E (2005). *Psychology: the science of mind and behaviour*, 5th edn. Hodder Arnold, London.

Gruber H, Degner S and Lehmann AC (2004). Why do some commit themselves in deliberate practice for many years—and so many do not? Understanding the development of professionalism in music. In M Radovan and N Dordevic, eds, *Current issues in adult learning and motivation*, 222–235. Slovenian Institute for Adult Education, Ljubljana.

Hallam S (1998a). The predictors of achievement and drop out in instrumental music tuition, *Psychology of Music*, **26**, 116–132.

Hallam S (1998b). *Instrumental teaching: a practical guide to better teaching and learning*. Heinemann, Oxford.

Hallam S (2004). How important is practicing as a predictor of learning outcomes in instrumental music? In SD Lipscomb, R Ashley RO Gjerdingen and P Webster, eds, *Proceedings of the 8th international conference on music perception and cognition*, 3–7 August 2004, 165–168. Northwestern University, Evanston, IL.

Hallam S (2006). Musicality. In G McPherson, ed., *The child as musician: a handbook of musical development*, 93–110. Oxford University Press, Oxford.

Hallam S and Prince V (2003). Conceptions of musical ability. *Research Studies in Music Education*, **20**, 2–22.

Hallam S and Shaw J (2003). Constructions of musical ability. *Bulletin of the Council for Research in Music Education*, Special Issue, 19th International Society for Music Education Research Seminar, Gothenburg, Sweden, School of Music, University of Gothenberg, August 3–9 2002, **153/4**, 102–107.

Haroutounian J (2000). Perspectives of musical talent: a study of identification criteria and procedures. *High Ability Studies*, **11**, 137–160.

Hodges DA (1996). Human musicality. In DA Hodges, ed., *Handbook of music psychology*, 29–68. IMR Press, San Antonio, TX.

Hodges DA (2006). The musical brain. In GE McPherson, ed., The child as musician: a *handbook of musical development*, pp 51–68. Oxford University Press, Oxford.

Hodges DA and Haack PA (1996). The influence of music on human behavior. In DA Hodges, ed., *Handbook of music psychology*, 469–555. IMR Press, San Antonio, TX.

Hurley CG (1995). Student motivations for beginning and continuing/discontinuing string music tuition. *The*

*Quarterly Journal of Music Teaching and Learning*, **6**, 44–55.

Jørgensen H. (2002). Instrumental performance expertise and amount of practice among instrumental students in a conservatoire. *Music Education Research*, **4**, 105–119.

Kemp AE and Mills AE (2002). Musical potential. In R Parncutt and GE McPherson, eds, *The science and psychology of music performance: creative strategies for teaching and learning*, 3–16. Oxford University Press, Oxford.

Kokotsaki D and Hallam S (2007). Higher education music students' perceptions of the benefits of participative music making. *Music Education Research*, **9**, 93–109.

Kopiez R. (1998). Singers are late beginners. Sangerbiographien aus Sicht der Expertiseforschung. Ein Schwachstellenanalyse. [Singers biographies from the perspective of research on expertise. An analysis of weaknesses] In H Gembris R Kraemer and G Maas, eds, *Singen als Gegenstand der Grundlagenforschung*, 37–56. Wissner, Augsberg.

Lehmann AC and Gruber H (2006). Music. In KA Ericsson, N Charness, PJ Feltovich and RR Hoffman, eds, *The Cambridge handbook of expertise and expert performance*, 457–470. Cambridge University Press, Cambridge.

Levitin DJ, Bellugi U. (1998). Musical abilities in individuals with Williams syndrome. *Music Perception*, **15**, 357–89.

Levitin DJ, Cole K, Chiles M, Lai Z, Lincoln A and Bellugi U (2007). Characterizing the musical phenotype in individuals with Williams syndrome. *Child Neuropsychology*, **10**, 223–247.

Levitin DJ, Menon V, Schmitt JE, *et al.* (2003). Neural correlates of auditory perception in Williams syndrome: an iMRI study. *Neuroimage*, **18**, 74–82.

Mareschal D, Johnson MH, Sirois S, Spratling MW, Thomas MSC and Westerman G (2007) *Neuroconstructivism: how the brain constructs cognition. Volume 1*. Oxford University Press, Oxford.

McPherson GE (1993). Factors and abilities influencing the development of visual, aural and creative performance skills in music and their educational implications. Doctor of Philosophy, University of Sydney, Australia. Dissertation Abstracts International, 54/04-A, 1277. (University Microfilms No. 9317278.)

McPherson GE (1996). Five aspects of musical performance and their correlates. *Bulletin of the Council for Research in Music Education*, **127**, 115–121.

McPherson GE (2005). From child to musician: skill development during the beginning stages of learning an instrument. *Psychology of Music*, **33**, 5–35.

McPherson GE (in press). The role of parents in children's musical development. *Psychology of Music*.

McPherson GE and Davidson JW (2002). Musical practice: mother and child interactions during the first year of learning an instrument. *Music Education Research*, **4**, 143–158.

McPherson GE and Williamon A (2006). Giftedness and talent. In GE McPherson, ed., *The child as musician: a handbook of musical development*, 239–256. Oxford University Press, Oxford.

McPherson GE, Bailey M and Sinclair K (1997). Path analysis of a model to describe the relationship among five types of musical performance. *Journal of Research in Music Education*, **45**, 103–129.

Miller LK (1989). *Musical savants: exceptional skill in the mentally retarded*. Erlbaum, Hillsdale, NJ.

Parncutt R (2006). Prenatal development. In GE McPherson, ed., *The child as musician: a handbook of musical development*, 1–32. Oxford University Press, Oxford.

Rauschecker JP (2003). Functional organization and plasticity of auditory cortex. In I Peretz and R Zatorre, eds, *The cognitive neuroscience of music*, 357–365. Oxford University Press, Oxford.

Ruthsatz J and Detterman DK (2003). An extraordinary memory: the case study of a musical prodigy. *Intelligence*, **31**, 509–518.

Rutkowski J (1996). The effectiveness of individual/small group singing activities on kindergartners' use of singing voice and developmental music aptitude. *Journal of Research in Music Education*, **44**, 353–368.

Schlaug G (2003). The brain of musicians. In I. Peretz and R. Zatorre, eds, *The cognitive neuroscience of music*, 366–381. Oxford University Press, Oxford.

Seashore CE (1938, reprinted 1960). *Psychology of music*. Dover, New York.

Seashore CE, Lewis L and Saetveit JG (1938). *Seashore measures of musical talents*. Psychological Corporation, New York.

Shavinina LV (2007). On the advancement of the expert performance approach via a deep understanding of giftedness. *High Ability Studies*, **18**, 79–82.

Shavinina LV (in press). When child prodigies, unique representations, and the extracognitive combine: toward a cognitive–developmental theory of giftedness. In LV Shavinina, ed., *The international handbook on giftedness*. Amsterdam: Springer Science and Business Media.

Shuter-Dyson R (1999). Musical ability. In D Deutsch, ed., *The psychology of music*, 627–651. Harcourt Brace and Company, New York.

Sloboda J, Hermelin B and O'Connor N (1985). An exceptional musical memory. *Musical Perception*, **3**, 155–170.

Sloboda JA and Howe MJA (1991). Biographical precursors of musical excellence: an interview study. *Psychology of Music*, **19**, 3–21.

Sloboda JA, Davidson JW, Howe MJA and Moore DG (1996). The role of practice in the development of performing musicians. *British Journal of Psychology*, **87**, 287–309.

Sosniak LA (1985). Learning to be a concert pianist. Developing talent in young people. In BS Bloom, ed., *Developing talent in young people*, 19–67. Ballantine, New York.

Sundin B (1997). Musical creativity in childhood—a research project in retrospect. *Research Studies in Music Education*, **9**, 48–57.

Trehub SE (2006). Infants as musical connoisseurs. In GE McPherson, ed., *The child as musician: a handbook of musical development*, 33–49. Oxford University Press, Oxford.

Trevarthen C and Malloch S (2002). Musicality and music before three: human vitality and invention shared with pride. *Zero to Three*, September, 10–18.

Vandervert LR (2007). Cognitive functions of the cerebellum explain how Ericsson's deliberate practice produces giftedness. *High Ability Studies*, **18**, 89–92.

Vandervert LR and Liu H (in press). How working memory and the cognitive cerebellum collaboratively produce the child prodigy. In LV Shavinina, ed., *The international handbook on giftedness*. Amsterdam: Springer Science and Business Media.

Vaughan MM (1977). Measuring creativity: its cultivation and measurement. *Bulletin of the Council for Research in Music Education*, **50**, 72–77.

Wagner MJ (1975). The effect of a practice report on practice time and musical performance. In CK Madsen, RD Greer and CH Madsen Jr, eds, *Research in music behaviour*, 125–130. Teachers College Press, New York.

Wallin N, Merker B and Brown S (2000). *The origins of music*. The MIT Press, Cambridge, MA.

Webster PR (1988). New perspectives on music aptitude and achievement. *Psychomusicology*, **7**, 177–194.

Welch GF and Adams P. (2003). *How is music learning celebrated and developed?* British Educational Research Assocation (BERA), Southwell, UK.

Williamon A and Valentine E (2000). Quantity and quality of musical practice as predictors of performance quality. *British Journal of Psychology*, **91**, 353–376.

Wing HD (1981). *Standardised tests of musical intelligence*. National Foundation for Educational Research, Windsor.

Young L and Nettelbeck T (1995). The abilities of a musical savant and his family. *Journal of Autism and Developmental Disorders*, **25**, 231–247.

Zurcher W (1972). The effect of model-supportive practice on beginning brass instrumentalists. In CK Madsen, RD Greer and CH Madsen, eds, *Research in music behavior: modifying music behavior in the classroom*, 125–130. Teachers College Press, New York.

# CHAPTER 25

# Practising

Harald Jørgensen and Susan Hallam

## Introduction

Practice is central to the development of all aspects of musical expertise. The musician not only needs to consider the development of technical skills but must also develop musical interpretation, may have to play or sing from memory, rehearse and perform in cooperation with other musicians, improvise, and contend with stage fright. These elements require aural, technical, cognitive, communication, performance and learning skills. These complex skills cannot be acquired, improved and maintained by simple repetitious practice.

Effective practice has been defined by Hallam (1997c) as 'that which achieves the desired end-product, in as short a time as possible, without interfering negatively with longer-term goals' (p. 181). This definition assumes that effective practice might take many forms and implies that the musician requires considerable metacognitive skills to facilitate the completion of task requirements or, in the case of the novice, appropriate support. Practising may be addressed from a psychological viewpoint as an act of learning, where theories of psychomotor learning and motor programmes are relevant, and it may also be viewed as 'self-teaching' (Jørgensen 2004). The multifaceted nature of practice has been encapsulated in models which provide a framework for understanding its relationships with creativity and performance (see Hallam 1997c; Chaffin and Lemieux 2004).

Empirical research on practice has a history that dates back to the beginning of the twentieth century, although the majority of research has been undertaken in the last 25 years. Its focus has been almost exclusively in relation to the training of classical musicians and the individual practitioner.

## The quantity of practice

Practitioners and researchers agree that there are two important variables relating to practice that determine progress and attainment: the *quality* and *quantity* of practice in interaction with prior knowledge and skills. In the following section we will concentrate on the quantity of practice, returning to quality issues in the section on practice strategies.

### Time spent practising

Research on time spent practising has addressed three different aspects: the initial starting age; the accumulated amount of practice from initial starting age to the present; and the amount of practice at one particular time or during a limited period of time. A major challenge to this type of research is the difficulty in obtaining reliable and valid measures of the amount of practice undertaken (Madsen 2005).

Despite these difficulties some broad trends have emerged. First, most of those who reach a high level of expertise on an instrument have made an *early start*, on either their major instrument or another instrument. This phenomenon has been demonstrated on a broad range of instruments (Jørgensen 2001; Sosniak 1985). Pianists and violinists tend to be particularly early starters, aged from 3–8 years old, while brass and woodwind performers start a little later with their major instrument (Jørgensen 2001). Starting to play at an early age, when physically the body is more flexible, may be important. Certainly, Wagner (1988) has demonstrated that pianists' hands can change physically if they begin playing when very young.

Another trend is that time spent in practice usually increases as age and expertise develops,

with most young people practising almost every day (Sloboda *et al.* 1996). There is evidence that the increase in practice time is greater for those who go on to become professional rather than amateur musicians. However, after entrance to the profession, duration of regular practice time is observed to decrease as pressures of rehearsals and public performance increase (Krampe 1994).

Since many students start early and gradually increase the amount of time that they practise, it follows that many of them have *accumulated a large amount of practice time* by their late teens, and that expert performers have invested several thousand hours of practice over a period of 15–16 years before reaching a high performance level in their twenties (Sosniak 1985). Accumulated practice time has also been found to relate to the performance of rehearsed music in novice players aged 7–9 years, although use of specific strategies seems to be more important when children engage in sight-reading, playing from memory, playing by ear, or improvising (McPherson 2005).

Most research on the amount of practice has not differentiated between instruments or has focused on only one instrument. An exception to this is a study by Jørgensen (1997). He found that the keyboard students in a conservatoire invested most time in practice, 25–30 hours a week, followed by strings, woodwind, brass and voice. There were also differences between specific instruments within these groups. Violinists, for instance, tended to practise more than double bass players, and trumpet players practised more than tuba players. Lammers and Kruger (2006), in a study of American and Japanese students, reported similar results. Physiological restrictions related to the instrument's physical and technical demands (Jørgensen 1997) and the nature and extent of the repertoire are probably important factors determining amount of practice.

Assessment, whether formal in examinations, or informal in a lesson, has an impact on practice time. Hallam (2001) reported that 95 per cent of the novices and advanced students in her study increased their practice time in weeks preceding examinations. Practice activity increases as the number of lessons received increases (Sloboda *et al.* 1996), and there is an increase in practice time the day after a weekly lesson (Lehmann and Ericsson 1998). Not surprisingly the quantity of practice decreases during holidays, even for students in a specialized music school (Sloboda *et al.* 1996).

Research on the relationship between *amount of practice at a particular time and general achievement level* at this time has considered a range of different instruments and age groups with varying outcomes (see Chapter 24 this volume and Jørgensen and Hallam in preparation). The different age groups and levels of expertise, lengths of time period studied, as well as the variety of instruments included in the studies may partly explain why the results are different from study to study. All of the studies have found a positive correlation between amount of practice at a particular point of time and general achievement, although Ericsson *et al.* (1993) found no difference in length of weekly practice between the 'best' and the 'good' violin students. The difference between the 'best' and 'good' violinists was in amount of accumulated practice.

Ericsson and colleagues (1993) suggested a monotonic relationship between 'deliberate practice' and an individual's acquired performance. Similarly, Sloboda *et al.* (1996, p. 308) stated that 'We believe that we have established, beyond any reasonable doubt, that amount of relevant practice is a key variable in determination of music performance expertise.' However, there are substantial individual differences in the relationship between the quantity of practice and attainment (Ericsson *et al.* 1993; Sloboda *et al.* 1996; Jørgensen 2002), suggesting that attainment is not exclusively a question of quantity of practice, but also of quality, which is a result of individual engagement with and knowledge of practice strategies.

Two studies have related amount of current practice to more specific areas of achievement, with conflicting results. Williamon and Valentine (2000) looked at practice and performance among piano students from under 11 to more than 24 years in four levels of skill on one composition, and found that pianists at higher levels of expertise spent more time in each practice session, but that quantity of practice was not significantly related to quality of performance as rated by experienced teachers. Wagner (1975), with college students, assessed performance on 'a selection which best represented their

level of musicianship at that time', and found a positive relationship between amount of practice and 'level of musicianship'.

A specific issue is the time spent practising a single composition. To date, research has focused on the memorization of piano music by professional musicians, students with high levels of expertise (Chaffin and Imreh 1997), or pianists with a broad range of skill levels (Williamon and Valentine 2000). The findings illustrate how time-consuming the memorization process can be, depending on the complexity of the piece. Lehmann and Ericsson (1998) studied a university student preparing for her degree recital, memorizing eight unfamiliar pieces by Haydn, Prokofiev and Debussy. She spent a total of 531 hours practising to prepare for a concert with a total playing time of 37 minutes.

### Motivation and drop out

Evidence from students who have dropped out suggests that there are complex relationships between time spent practising, prior knowledge, motivation, and perceived efficacy (Hurley 1995). The amount of practice undertaken predicts whether students will drop out (Sloboda *et al.* 1996; Hallam 1998), although other factors such as socio-economic status, self-concept in music, reading achievement, scholastic ability, measured musical ability, maths achievement, and motivation are also predictors (McCarthy 1980; Klinedinst 1991; Hallam 1998).

McPherson and McCormick (2006) found that self-efficacy was the best predictor of performance. The role of psychological characteristics (rather than musical talent or expertise) was also observed by McNamara *et al.* (2006) in their study of a group of internationally renowned musicians. They found that some characteristics were generic (e.g. dedication, planning and commitment), while others were particularly evident at transition points in their career, i.e. perseverance and adaptability in the transition from study to entry into the music profession.

## The quality of practice

Quality is an elusive matter. The concept of 'deliberate practice' defined as goal-oriented, structured, and effortful, was introduced by Ericsson and colleagues (1993) to address the issue of quality. They also outlined constraints which might determine the quantity and quality of practice—motivation, resources, attention. Much of the research exploring issues relating to the quality of practice has focused on the strategies that musicians adopt when practising. Jørgensen (2004) has proposed four strategy types, i.e. planning strategies; strategies for the conduct (execution) of practice; strategies to evaluate practice; and meta-strategies. A similar conception is that of practice as self-regulated learning (McPherson and Zimmerman 2002), where the practitioner is recommended to engage in forethought, performance/volitional control, and self reflection.

### Planning strategies

#### The organization of practice

Instrumentalists and singers are expected to *practise regularly*. Several management strategies have been observed in relation to this. Some students in higher music education practise at the same time every day (Duke *et al.* 1997), while others integrate practice into a daily or weekly plan (Jørgensen 1997). Most students, however, try to fit in practice sessions between other activities without any preconceived plan. The morning may be the best time for high levels of concentration (Lehmann and Ericsson 1998). Ericsson and colleagues (1993) found that conservatoire students at the highest levels of expertise practised in the morning, took naps in the afternoon, and then put in more practice in the evening. For novice students the regularity of practice may be more related to one specific day a week (Hallam 2001), or a specific period of the day, i.e. 'before bedtime' (Pitts *et al.* 2000b). Sloboda *et al.* (1996) showed that the students in a selective specialist music school distributed their repertory practice evenly to morning, afternoon and evening sessions.

Practice may be most effective when it is organized in sequential and logical manner (Barry 1992). At the start of practice sessions many musicians use warm-up exercises, although there is considerable individual variation in the extent to which these are perceived to be necessary (Hallam 1995a). Technical exercises

often follow with repertory work left until last (Duke *et al.* 1997). Sloboda *et al.* (1996) showed that the students in a selective specialist music school practised more than 40 per cent of their scales in the morning session. For conservatoire students, there are pronounced differences between instruments in the relative amount of warming up exercises, technical work, and repertory practice (Jørgensen 1998).

### Setting goals and adopting appropriate approaches for particular tasks

Ericsson *et al.* (1993) concluded that a well-defined task was one of four requirements for effective learning through practice, although the evidence suggests that novices and more accomplished students often fail to formulate goals for practice activities and mastering specific tasks (Jørgensen 1998). The nature of particular learning tasks inevitably affects the goals set and the ways in which they are undertaken. McPherson (2005) stressed the importance of helping students to develop task-appropriate strategies for such diverse tasks as practising notated repertoire, playing from memory, sight-reading, playing by ear, and improvising.

Musicians seem to approach practising particular repertoire in different ways. Miklaszewski (1995) observed that professional musicians spent a much shorter time learning a late romantic miniature than three contemporary variations. Lehmann and Ericsson (1995) suggest that the increasing technical demands of twentieth-century music have influenced the nature of practice. Research on sight-reading and improvisation suggests that time spent engaged in these activities is the key element in their development to expert levels (see Chapter 32 this volume).

### Developing interpretation

A specific task in formulating goals is the 'performance plan', and the way interpretation is planned and developed through practice sessions. Some musicians plan interpretation at the outset, based on a study of the score or from ideas gleaned from listening to a wide range of music and different interpretations of the same piece (Hallam 1995b; Lisboa *et al.* 2005); primarily letting the expressive ideas guide the technical work (Chaffin *et al.* 2003). A second approach is to develop a performance plan after mastering most of the technical challenges (Nielsen 2001). Experienced performers also report taking the audience perspective into account when developing expressive ideas (Oura and Hatano 2001).

## Strategies for the conduct of practice

### Variable practice

Schema theory suggests that motor programmes, including those required for playing a musical instrument, are strengthened by increased variability in practice (for instance, practising a passage with different articulations, at different tempos, or a technique using different examples), rather than repetition of the same actions. This facilitates transfer to other tasks (Schmidt 1976). The evidence for the effectiveness of this is mixed (see Hallam 1997c for a review).

### Part–whole strategies

Particularly relevant to the acquisition of musical skill is the question of *part–whole transfer* of training. Given a reasonably short piece of music to practise, observational studies have shown that many novice pupils play through the music without stopping to focus on difficult sections, and usually repeat the whole piece several times (see Hallam 1997b; Renwick and McPherson 2002). Some novice and advanced students use a combined approach, starting with the whole and stopping to practise difficult sections en route (Hallam 1997b; Miklaszewski 1989). This gives the performer an overview of the music, and the opportunity to identify and select parts which require more intense work while relating the parts to longer sections or the whole (Chaffin and Imreh 1997; Nielsen 1997). Sections for concentrated work are selected on the basis of a range of criteria including those relating to the formal structure and motor aspects of the performance (Miklaszewski and Sawicki 1992); new or related elements (Nielsen 1999), and the visual layout of the music and its harmonic progression (Williamon and Valentine 2000). As practice progresses and the music is increasingly mastered technically, the sections worked on become longer (Chaffin and Imreh 1997; Nielsen 1999), although attention to detail and work on small sections may continue throughout practice sessions. It is clear from these examples

that the relationship between practising a piece in its entirety, focusing on parts, and mastery is complex. For different tasks whole or part strategies may be more appropriate.

### Transfer of learning

Exercises are sometimes used to address *specific challenges* within a given composition (Hallam 1995; Nielsen 1999), but the alternative and more common solution is to practise difficult sections within the music being learnt. Transfer of learning in music needs to be considered in relation to particular tasks and different timescales. In the short term, practice seems to be most effective when it relates specifically to the task being undertaken, with the conditions for learning and performance being as similar as possible.

## Strategies for increasing tempo

When passages need to be played at speed, Drake and Palmer (2000) observed three approaches adopted by students at different levels of expertise. Beginners tended to stick to one tempo throughout practice sessions, novices increased tempo gradually until they reached a limit, while the most accomplished students gradually increased tempo over each practice session. Research on the efficiency of these approaches has been equivocal probably due to the ambiguity of concepts such as 'slow', 'fast', and 'in performance tempo'; differences between beginners and experts; the way strategies relating to tempo are often combined with other strategies; and the length and complexity of music involved. Since the adoption of these strategies involves motor and muscular considerations, the observation by Winold *et al.* (1994), that fast and slow speeds of movement seem to be controlled in different ways is important.

### Mental practice

Mental practice, where the learner thinks through the procedures without actually playing, has been compared with playing practice in several studies, with conflicting results (Kopiez 1990; Ross 1985). This is hardly surprising, taking into account that:

- the studies differ in the length, familiarity and complexity of the music used;
- 'mental strategy' has been operationalized differently;
- the length of time using the strategies has differed; and
- familiarity with the use of mental strategies has varied between those taking part in the studies.The conclusion by Ross (1985), that a combination of mental and physical practice is most effective because mental practice allows concentration on the cognitive aspects of music performance without the distractions of exercising motor control, is probably sound.

### Strategies for preparing for performance

Even if the music is well prepared, performance anxiety may jeopardize all prior efforts, particularly where the performer has to play from memory. The most common strategy for overcoming performance nerves is to be well-prepared and over-learn, investing more time than is required for basic mastery (Lehman and Ericsson 1998) (see Chapters 31, 33 and 36 this volume for issues relating to planning, memorization and anxiety).

## Strategies to evaluate practice

### Monitoring the effectiveness of practice

To monitor the effectiveness of practice it is particularly important that appropriate schema against which to evaluate progress are developed. Hallam (1997a) found that 60 per cent of beginners, novices and advanced students when learning a new piece left errors uncorrected. She observed that errors once made tended to become permanent and were left uncorrected. Only when students had considerable expertise and had well-developed schema were errors consistently corrected. Whitaker (1996), studying performers, arrangers, conductors and composers also established the necessity of possessing a mental template that served as the focus for all learning and performance activities. In contrast to professionals, many students have been found to be unable to generate auditory schemata from written notation (Grøndahl 1987). These skills need to be developed until at expert level current information about progress can be utilized to develop more sophisticated mental representations. Several external remedies and techniques have been used in the process of

developing schemata. Some studies have reported that to use a tape with a recording of the music is an efficient strategy (Puopolo 1971; Zurcher 1975), although others have found no effect of a recorded model (Anderson 1981; Linklater 1997).

### Meta-strategies

Metacognitive strategies are concerned with the planning, monitoring and evaluation of learning. There are considerable differences between beginners, novices and experts in their knowledge and deployment of different practising and self-regulating strategies (Hallam 1997b; Pitts *et al.* 2000a, b) as well as individual differences among musicians and novices at the same level of competence (Nielsen 1997, 1999, 2001; Austin and Berg 2006). Hallam (1997b) demonstrated that professional musicians had well-developed metacognitive skills, including self-awareness of strengths and weaknesses, extensive knowledge regarding the nature of different tasks and what would be required to complete them satisfactorily, and strategies which could be adopted in response to perceived needs. This not only encompassed technical matters, interpretation and performance but also issues relating to learning itself, concentration, planning, monitoring and evaluation. Novices demonstrated less metacognitive awareness, the amount and structure of their practice tending to be determined by external commitments, e.g. examinations. There are differences in students' practising strategies and metacognitive skills, the former being inextricably intertwined with the acquisition of expertise. Skilled musicians are more likely to monitor and control their playing by focusing their attention on what they are practising and how it can be improved.

Among monitoring strategies, the monitoring of concentration is regarded by many music students as crucial to practice. Attempting to improve undergraduate students' concentration, Madsen and Geringer (1981) required them to keep a record of their practice, rate each practice session for productivity over an 8-week period and complete a distraction index. This increased observed attentiveness and performance although the index was time-consuming to complete. Nielsen (2001) suggests that students can enhance self-guidance by covertly or overtly describing how to proceed, giving comments on progress, noting concentration lapses and changes in motivation. Focus is crucial to avoid mindless repetition. The regulation of strategies requires deliberate effort, task selection, speed and intensity.

## Rehearsing in small groups

Research on rehearsals in small groups has shown that there is no single best strategy for rehearsing repertoire (Davidson and King 2004). Individual ensembles find their own best ways of working. Berg (2000) found that high school student ensembles adopted four main activities, initiating, performing, orienting and assisted learning. Professional groups usually have a draft plan of the material to be covered across a series of rehearsals with room for extra sessions should they be needed (Goodman 2000). Group cohesiveness centres around the music (Davidson and Good 2002) but trust and respect are crucial for the functioning of groups working together continuously over long periods of time, e.g. string quartets (Young and Colman 1979). To be successful in the long term rehearsals have to be underpinned by strong social frameworks (for a full review see Davidson and King 2004).

## The role of the learning environment in practice

The impact that institutional learning environments have on practice has been little studied, although research at conservatoire level in Norway (Jørgensen 1997) reported that music education students and church music students practiced more than expected, probably due to the predominant performance values in a conservatoire.

Instrumental teachers are an important influence on practice. Research on the way that they teach about practising has had mixed results. Jørgensen (2000), in research with beginning conservatoire students in Norway, reported that 40 per cent indicated that their previous teachers had invested 'little' or 'no' effort in teaching them how to practise. However, in the USA, teachers have reported that they always or

almost always include instructions about practice in their lessons (Barry and McArthur 1994). Students do seem to be able to learn how to use expert practising strategies (Barry 1992) and having done so report more positive attitudes towards practising, are more likely to engage in practice planning and problem identification, are better able to select appropriate performance goals, and are able to formulate more cognitively complex goals (Kenny 1992). This suggests that there can be benefits in teaching about practice.

## Conclusions

Over the last 25 years our understanding of the nature of practice and its importance in the development of expertise has increased enormously. We know that both the quantity and quality of practice contribute to the level of expertise attained, and that the individual's ability to adopt more effective practising strategies is inextricably linked with their level of expertise. There has also been a recognition that musicians exhibit considerable diversity in the ways that they practise and that these can lead to equally successful outcomes. Despite the considerable progress made there are still areas where we know little. In relation to the quantity of practice there is a need for more studies that address the interaction between instrument, age, level of expertise and amount of practice within a range of different contexts. In relation to the quality of practice there is a need for studies exploring:

- how practice plans and goals are formulated and the way that these influence practice and subsequently performance;
- skill transfer between warming up exercises, technical studies, and repertory practice and the impact of these on performance;
- the effectiveness of aural models of what is to be learned and other types of feedback;
- concentration in practice, and how self-regulating techniques and meta-cognitive skills can be developed;
- the relationships between learning approaches, motivation and practice efficiency;
- how best to teach practice strategies at all levels of expertise;
- the way that social interaction in groups, including non-verbal communication, affects performance;
- the ways in which practice is undertaken on a variety of different tasks, e.g. sight reading, improvisation and in different genres, e.g. popular music, jazz, and world musics.

Such research will not only increase our understanding but also contribute towards enhancing learning and teaching.

## References


Anderson JN (1981). Effects of tape recorded aural models on sight-reading and performance skills. *Journal of Research in Music Education*, **29**, 23–30.

Austin JR and Berg MH (2006). Exploring music practice among sixth grade band and orchestra students. *Psychology of Music*, **34**, 535–558.

Barry NH (1992). The effects of practice strategies, individual differences in cognitive style, and gender upon technical accuracy and musicality of student instrumental performance. *Psychology of Music*, **20**, 112–123.

Barry NH and McArthur V (1994). Teaching practice strategies in the music studio: a survey of applied music teachers. *Psychology of Music*, **22**, 44–55.

Berg MH (2000). Thinking for yourself: the social construction of chamber music experience. In RR Rideout, SJ Paul, eds. *On the sociology of music: vol 2. Papers from the Music Education Symposium at the University of Oklahoma*, 91–112. University of Massachusetts Press, Amherst, MA.

Chaffin R and Imreh G (1997). Pulling teeth and torture: musical memory and problem solving. *Thinking and Reasoning*, **3**, 315–336.

Chaffin R, Imreh G, Lemieux A and Chen C (2003). Seeing the big picture: piano practice as expert problem solving. *Music Perceotion*, **20**, 465–490.

Chaffin R and Lemieux AF (2004). General perspectives on achieving musical excellence. In A Williamon, ed., *Musical excellence: strategies and techniques to enhance performance*, 19–40. Oxford University Press, Oxford.

Davidson JW and Good JMM (2002). Social and musical co-ordination between members of a string quartet: an exploratory study. *Psychology of Music*, **30**, 186–201.

Davidson J and King EC (2004). Strategies for ensemble practice. In A Williamon, ed. *Musical Excellence*, 105–122. Oxford University Press, Oxford.

Drake C and Palmer C (2000). Skill acquisition in music performance: relations between planning and temporal control. *Cognition*, **74**, 1–32.

Duke RA, Flowers PJ and Wolfe DE (1997). Children who study with piano with excellent teachers in the United States. *Bulletin of the Council for Research in Music Education*, **132**, 51–84.

Ericsson KA, Krampe RT and Tesch-Romer C (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, **100**, 363–406.

Goodman E (2002). Ensemble performance. In J Rink, ed., *Musical performance: a guide to understanding*, 153–167. Cambridge University Press, Cambridge.

Grøndahl D (1987). *Thinking processes and structures used by professional pianists in keyboard learning.* Hovedfagsewksamen in Music Education, The Norwegian State Academy of Music, Oslo.

Hallam S (1995a). Professional musicians' orientations to practice: implications for teaching. *British Journal of Music Education*, **12**, 3–19.

Hallam S (1995b). Professional musicians' approaches to the learning and interpretation of music. *Psychology of Music*, **23**, 111–128.

Hallam S (1997a). The development of memorisation strategies in musicians: implications for instrumental teaching. *British Journal of Music Education*, **14**, 87–97.

Hallam S (1997b). Approaches to instrumental music practice of experts and novices: implications for education. In H Jorgensen and A Lehman, eds, *Does practice make perfect? Current theory and research on instrumental music practice*, 89–108. NMH-publikasjoner 1997:1, Norges musikkhgskole, Oslo.

Hallam S (1997c).What do we know about practising? Towards a model synthesising the research literature. In H Jorgensen and A Lehman, eds, *Does practice make perfect? Current theory and research on instrumental music practice*, 179–231. NMH-publikasjoner 1997:1. Norges musikkhgskole, Oslo.

Hallam S (1998). Predictors of achievement and drop out in instrumental tuition. *Psychology of Music*, **26**, 116–132.

Hallam S (2001). The development of expertise in young musicians: strategy use, knowledge acquisition and individual diversity. *Music Education Research*, **3**, 7–23.

Hurley CG (1995). Student motivations for beginning and continuing/discontinuing string music tuition. *The Quarterly Journal of Music Teaching and Learning*, **6**, 44–55.

Jørgensen H (1997). Time for practicing? Higher level students' use of time for instrumental practicing. In H Jørgensen and AC Lehmann, eds, *Does practice make perfect? Current theory and research on instrumental music practice*, 123–140. Norges musikkhøgskole, Oslo.

Jørgensen H (1998). *Planlegges øving?* (Is practice planned?). Norwegian Academy of Music, Oslo.

Jørgensen H (2000). Student learning in higher instrumental education: who is responsible? *British Journal of Music Education*, **17**, 67–77.

Jørgensen H (2001).Instrumental learning: is an early start a key to success? *British Journal of Music Education*, **18**, 227–239.

Jørgensen H (2002). Instrumental performance expertise and amount of practice among instrumental students in a conservatoire. *Music Education Research*, **4**, 105–119.

Jørgensen H (2004). Strategies for individual practice. In A Williamon, ed., *Musical excellence*, 85–104. Oxford University Press, Oxford.

Jørgensen H and Hallam S (in preparation) How important is the quantity of practice in attainment?

Kenny WE (1992). The effect of metacognitive strategy instruction on the performance proficiency and attitude toward practice of beginning band students. Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign, IL.

Klinedinst RE (1991). Predicting performance achievement and retention of fifth-grade instrumental students. *Journal of Research in Music Education*, **39**, 225–238.

Kopiez R (1990). *Der Einfluss kognitiver Strukturne auf das Erlernen eines Musikstucks am Instrument.* (The influence of cognitive structures on the learning of instrumental music.) Peter Lang, Frankfurt.

Krampe RT (1994). *Maintaining excellence: cognitive-motor performance in pianists differing in age and skill level* (Studien und Berichte/MPI fur Bildungsforschung 58). Sigma, Berlin.

Lammers M and Kruger M (2006). Brass and woodwind student practice habits in Norway, Japan and the United States. *NACWPI Journal*, **54**, 4–13.

Lehmann AC and Ericsson KA (1995). The relationship between historical constraints of musical practice and increase of musicians' performance skills. Paper presented at the 7th European conference on Developmental Psychology, Poland. 1995, 23–27 August.

Lehmann AC and Ericsson KA (1998). Preparation of a public piano performance: the relationship between practice and performance. *Musicae Scientae*, **2**, 67–94.

Linklater F (1997). Effects of audio- and videotape models on performance achievement of beginning clarinettists. *Journal of Research in Music Education*, **45**, 402–414.

Lisboa T, Williamon A, Zicari M and Eiholzer H (2005). Mastery through imitation: a preliminary study. *Musicae Scientiae*, **19**, 75–110.

Madsen C (2005). A 30-year follow-up study of actual applied music practice versus estimated practice. *Journal of Research in Music Education*, **52**, 77–88.

Madsen C K and Geringer JM (1981). The effect of a distraction index on improving practice attentiveness and musical performance. *Bulletin of the Council for Research in Music Education*, **66–67**, 46–52.

McCarthy JF (1980). Individualised instruction, student achievement and drop out in an urban elementary instrumental music program. *Journal of Research in Music Education*, **28**, 59–69.

MacNamara A, Holmes P and Collins D (2006). The pathway to excellence: the role of psychological characteristics in negotiating the challenges of musical development. *British Journal of Music Education*, **23**, 285–302.

McPherson GE (2005). From child to musician: skill development during the beginning stages of learning an instrument. *Psychology of Music*, **33**, 5–35.

McPherson GE and McCormick J (2006) Self-efficacy and performing music. *Psychology of Music*, **34**, 321–336.

McPherson GE and Zimmerman BJ (2002). Self-regulation of musical learning. In R Colwell and C Richardson, eds, *The new handbook of research on music teaching and learning*, 348–372. Oxford University Press, Oxford.

Miklaszewski K (1989). A case study of a pianist preparing a musical performance. *Psychology of Music*, **17**, 95–109.

Miklaszewski K (1995). Individual differences in preparing a musical composition for public performance. In M Manturzewska, K Milaszewski and A Bialkowski, eds, *Psychology of music today: proceedings of the international seminar of researchers and lecturers in the psychology of music*, 138–147. Fryderyk Chopin Academy of Music, Warsaw.

Miklaszewski K and Sawicki L (1992). Segmentation of music introduced by practicing pianists preparing compositions for public performance. In R Dalmonte and M Baroni, eds, *Secudo convegno Europeo di analasi musicale* [Proceedings of the second European conference on musical analysis], 113–121. University of Trento Press, Trento.

Nielsen SG (1997). Self-regulation of learning strategies during practice: a case study of a church organ student preparing a musical work for performance. In H Jorgensen and AC Lehmann, eds, *Does practice make perfect? Current theory and research on instrumental music practice*, 109–122. Norges Musikkogskole, Oslo.

Nielson SG (1999). Learning strategies in instrumental music practice. *British Journal of Music Education*, **16**, 275–91.

Nielsen SG (2001). Self-regulating learning strategies in instrumental music practice. *Music Education Research*, **3**, 155–167.

Oura Y and Hatano G (2001). The constitution of general and specific mental models of other people. *Human Development*, **44**, 144–159.

Pitts S, Davidson J and McPherson GE (2000a). Developing effective practising strategies: case studies of three young instrumentalists. *Music Education Research*, **2**, 45–56.

Pitts S, Davidson JW and McPherson GE (2000b). Models of success and failure in instrumental learning: case studies of young players in the first 20 months of learning. *Bulletin of the Council for Research in Music Education*, **146**, 51–69.

Puopolo V (1971). The development and experimental application of self-instructional practice materials for beginning instrumentalists. *Journal of Research in Music Education*, **19**, 342.

Renwick JM and McPherson GE (2002). Interest and choice: student-selected repertoire and its effect on practising behaviour. *British Journal of Music Education*, **19**, 173–188.

Ross SL (1985). The effectiveness of mental practice in improving the performance of college trombonists. *Journal of Research in Music education*, **33**, 221–230.

Schmidt RA (1976).The schema as a solution to some persistent problems in motor learning theory. In GE Stelmach, ed., *Motor control: issues and trends*, 41–65. Academic Press, New York.

Sloboda JA, Davidson JW, Howe MJA and Moore DG (1996). The role of practice in the development of performing musicians. *British Journal of Psychology*, **87**, 287–309.

Sosniak LA (1985). Learning to be a concert pianist. Developing talent in young people. In BS Bloom, ed., *Developing talent in young people*, 19–67. Ballantine, New York.

Wagner MJ (1975). The effect of a practice report on practice time and musical performance. In CK Madsen, RD Greer and CH Madsen Jr, eds, *Research in music behaviour*, 125–130. Teachers College Press, New York.

Wagner C (1988). The pianist's hand: anthropometry and biomechanics. *Ergonomics*, **31**, 97–131.

Whitaker NL (1996). A theoretical model of the problem solving and decision making of performers, arrangers, conductors and composers. *Bulletin of the Council for Research in Music Education*, **128**, 1–14.

Williamon A and Valentine E (2000). Quantity and quality of musical practice as predictors of performance quality. *British Journal of Psychology*, **91**, 353–376.

Winold H, Thelen E and Ulrich BD (1994). Coordination and control in the bow arm movements of highly skilled cellists. *Ecological Psychology*, **6**, 1–31.

Young VM and Coleman AM (1979).Some psychological processes in string quartets. *Psychology of Music*, **7**, 12–16.

Zurcher W (1975). The effect of model-supportive practice on beginning brass instrumentalists. In CK Madsen, RD Greer and CH Madsen Jr, eds, *Research in Music Behaviour*, 125–30. Teachers College Press, New York.

CHAPTER 26

# Individuality in the learning of musical skills

Helena Gaunt and Susan Hallam

## Introduction

Historically, research concerned with individual differences in music has tended to focus on how musicians as a group differ from non-musicians, and the key characteristics which differentiate between musicians. This chapter outlines what we know about the way that individual characteristics including sex differences, physiological characteristics, age, personality, and cognitive and learning styles impact on engagement with and approaches to learning music. Issues relating to musical ability are explored in Chapter 24. The way that elements of the musical and wider environment influence the nature of engagement with music are also considered, excepting the role of the family, which is the focus of Chapter 28, and the school, which is the focus of Chapter 22. This chapter concludes by presenting theoretical frameworks which may serve to guide future research in elucidating the way that different environments interact in a reiterative manner with the individual leading, to different musical engagement outcomes.

## Individual characteristics

### Sex differences

In the past women have adopted roles in music which have been subservient to those of men, despite the fact that no reliable sex differences have been found in a range of measures of musical ability (Shuter-Dyson and Gabriel 1981; Hallam 2004), that girls tend to perform better in school music examinations (DES 1991; Agak 2002), and that girls report more positive attitudes towards music and musical activities (Crowther and Durkin 1982), and more positive competence beliefs and values for instrumental music than boys (Eccles *et al.* 1993).

Paradoxically, despite the dominance of men in the music professions music is regarded as a 'feminine' subject (Colley *et al.* 1994), although boys are more interested in and confident about music when it is linked to technology (Hanley 1998; Ho 2001) or when musical instruments depend on technology (Hallam *et al.* 2006). Girls are perceived to be interested and successful in singing, playing classical music and in dealing with notation while boys are perceived to have greater confidence in improvisation and composition. Girls are also perceived to be more persistent and more successful with instrumental study, have a broader listening repertoire, and are open to a wider range of musical styles (Green 1997; Hargreaves *et al.* 1995). Singing, in particular, is designated as a feminine activity (Green 1997; Hanley 1998) and although the proportions vary across studies and cultures, in most countries girls are more likely to learn to play a musical instrument than boys (Sheldon and Price 2005; Hallam *et al.* 2006). There are also gender differences in the particular instruments that boys and girls play. Early studies in Western cultures revealed that drums, trombone

and trumpet tended to be played more by boys, while flute, violin and clarinet have tended to be played by girls (Abeles and Porter 1978). Despite the increasing equality of women in Western cultures, stereotypical choice of instruments continues (Sinsel *et al.* 1997; Sheldon and Price 2005; Hallam *et al.* 2006) and exists in non-Western cultures (Ho 2001), although girls are selecting a wider variety of instruments to play along the feminine–masculine continuum than boys (Zervoudakes and Tanur 1994). These gender instrumental stereotypes are shared by parents (Abeles and Porter 1978; Delzell and Leppla 1992).

Presenting instruments to children aurally and visually without players can encourage boys to select more feminine instruments (Abeles and Porter 1978) as can changing the sex-role model playing the instrument (Bruce and Kemp 1993). However, where children choose to play an instrument which is considered inappropriate for their gender they may experience bullying or loss of popularity in school (Howe and Sloboda 1992). Children have clear ideas about which instruments are likely to lead to bullying if they are played (O'Neill and Boulton 1995).

Gendered engagement with different music genres changes over the lifespan. McKeage (2004) found that more girls participated in playing jazz in high school than in college. The women's decisions to discontinue were affected by primary instrument selection, institutional obstacles that narrowed participation options, feeling more comfortable in traditional ensembles, and an inability to connect jazz participation to career options.

## Physiological demands of different instruments

Physiology plays an important role in the nature of individual engagement with music. For instance, there are considerable differences in relation to the human voice (Thurman and Welch 2000; Harris *et al.* 1998). The pitch range of a human voice varies with sex and age, the voice tending to become stronger and less agile somewhere in the third decade of life, as the laryngeal cartilages start to ossify (Welch and Sundberg 2002). Different instruments also make different physiological demands both in terms of how a basic sound is produced and the coordination required to play patterns of notes. For instance, singers and wind players rely to a considerable extent on refined muscular control of the respiratory system to make and vary the dynamic and timbre of the sound they make, while string players use both arms and hands for independent tasks which then have to be coordinated to control pitch, articulation and the quality of sound. These physiological requirements may inhibit the take up of some instruments by some individuals and also have implications for the stage of physical development at which learning them can begin.

Whilst instruments are versatile in their potential for musical expression, some within particular genres take on specific roles which impact on the individual player's experience of the music and participation. For example, in jazz, the role of rhythm section players is considered substantially different from 'front line' players who take the lead and play solos, but are not taking the 'whole tune' (MacDonald and Wilson 2005).

## Age

Age affects the development of musical skills in a number of ways, through physical development, levels of cognitive skill and approaches to learning, and changing musical taste and motivation. Gembris and Davidson (2002) suggested that between the ages of 0 and 10, musical development in instrumental lessons may be limited by general sensorimotor as well as mental and emotional developmental processes. This is less the case with singing, which plays an almost universal role in childhood and is closely linked to the acquisition of language (Davidson 1994), although a fully professional singing voice in the Western classical tradition develops more slowly than proficiency in instrumental disciplines, and is rarely mature before the age of 30.

Sensorimotor abilities peak in adolescence and early adulthood, but then start to decrease, making facility harder to maintain as an instrumentalist. A crucial period of development amongst professional musicians (other than singers) has been identified as between 12 and 16 years. Sensorimotor ability is high, and cognitive skills are sufficiently developed to enable

independent practice to be increased significantly (Ericsson *et al.* 1993). Taking up a complex instrument as an adult is more difficult physically and tends not to yield the same results in terms of dexterity and instrumental control. On the other hand, knowledge and experience of music continue to grow throughout life, and professional classical musicians have articulated how they have used this resource to compensate for reduced technical facility as the speed of cognitive processing and kinaesthetic response decrease (Coffman 2002). The ability to use and manipulate musical symbols also changes over time, gradually becoming more sophisticated (Davidson and Scripp 1992; Gembris 2002). Coffmann (2002) outlines the different learning strategies adopted at different ages, with adult learners often showing particularly high levels of motivation including personal motivations (self-expression, recreation, self-improvement, use of leisure time), musical motivations (professed love of music, performing for one's self and others, learning more about music) and social motivations (meeting new people, being with friends, having a sense of belonging). Adopting an alternative classification, Gates defines adult learners as dabblers, recreationists, hobbyists, amateurs, apprentices, and professionals (Gates 1991).

Adult beginners may have higher expectations of themselves than children, and experience greater disappointment if they do not achieve their aspirations (Cope 2002). Some prefer informal and self-directed, critically reflective and collaborative modes of learning (Brookfield 1983, 1987), although in formal lessons they may rely on the teacher's expertise and analysis of their playing (Chen 1996). These differences seem to depend on the particular context of learning and may not be related to age per se.

The culture of our childhood determines the types of musical traditions with which we engage and as a result the specific musical schemata which develop and are reinforced over time. Those brought up within a Western culture may be able to distinguish fine pitch differences in Western tonality but be unable to distinguish similar pitch differences within the Javanese pelog scale (Lynch *et al.* 1990). Similarly, children exposed to Western music show no preference for consonance over dissonance up to the age of five, but by the age of nine or ten have developed this (Zenatti 1993).

## Personality

The musical temperament of Western classical musicians has been reported to be dominated by introversion, independence, sensitivity and anxiety. Kemp (1996) describes the professional musician as a bold introvert who directs energy inwards and appears outwardly reserved. The nature of solitary practice encourages autonomy, and independence of thought, although introversion while supporting the long hours of practice required may also induce anxiety. High levels of imagination, sensitivity and intuition are reflected in high levels of pathemia (extreme sensitivity) (Kemp 1996) while the presence of suspiciousness and low self-sentiment have been reported as potential contributing factors for the increase of anxiety. Kemp (1996) has suggested that musicians become so committed to music that many fail to separate their personal identities from their identity as musicians. Student performers tend to relate their self-esteem to how they perform (Dews and Williams 1989; Tobacyk and Downs 1986) and in musicians it is difficult to discriminate between self-esteem and self-efficacy (Smith *et al.* 2000). Kemp (1996) has argued that identification of musical achievement with personal value results in either 'prima donna' behaviour observed in high self-esteem performers that perceive themselves to be successful, or negative self-perceptions in performers with low self-esteem who have heightened apprehension, self-criticism and perceptions of incompetence. The competitive nature of the classical music profession provides a context which may encourage these trends. However, the personal characteristics of non-Western musicians in non-competitive environments have not been studied making it impossible to draw conclusions about causality, although comparisons have been made between those engaged in different aspects of Western music.

Research has been undertaken in relation to orchestral players, keyboard players, composers and singers (Davies 1978; Wilson 1984) as well as popular musicians (Wills and Cooper 1988). The findings suggest that string players are introverted, imaginative and radical, while brass

players are more extraverted and have lower levels of self-discipline and intelligence in comparison to other performing groups. Percussionists also tend towards extroversion. These findings suggest that the extent of practice required for these different instruments (typically more for string players) and their role in the orchestra (more prominent for brass, wind and percussion players) either attracts people with personality characteristics suited to these roles or that these characteristics develop in response to the particular demands being made of them by their chosen roles in the music profession. Similar issues are raised with regard to keyboard players who have been described as extroverted, warm-hearted, group-dependent, highly conscientious and self-sentimental, conservative and low in anxiety, and singers who are generally considered as extraverted, independent and sensitive (Kemp 1996).

There may also be complex relationships between personality, practising and performance. Kemp (1996) suggested that the relatively high levels of introversion exhibited by musicians in comparison with the general population might assist them in undertaking repetitive practice, while Hallam (2001a) observed that there appeared to be a relationship between concentration in practice and arousal levels in performance in professional musicians. Those who described themselves as effective, focused learners experienced higher arousal levels in performance which were often problematic.

## Cognitive and learning styles, and approaches to learning

In adult learners it has generally been accepted that some individual characteristics are more stable than others. Personality and cognitive styles are perceived as more consistent over time, while learning styles and approaches to learning have tended to be perceived as more responsive to particular learning contexts (Curry 1987). The relationships between these have been relatively little researched in music. Such research as there is has tended to explore whether there are differences between musicians in the way that they learn, with a particular focus on practice and interpretation. Differences have been found in learning styles, conceptions of and approaches to practice, and the length of time taken by professional musicians to learn repertoire.

Two main approaches to interpretation have been identified: analytic holist (taking a wide view of what is to be learned, considering the connections between disparate ideas, making wide use of analogy and illustration, developing ideas from listening extensively to music, planning interpretation in advance of learning) and intuitive serialist (developing interpretation as music is learned, using intuition rather than analysis, avoiding extensive listening to minimize undue influence on ideas), with versatile learners adopting either serialist or analytic strategies as necessary (Hallam 1995a). Musicians also adopt differential approaches to practising, adopting either technical, musical or mixed approaches, the latter depending on the nature of the specific task. Approaches to practising relate to some extent to the styles of interpretation described above. No musicians adopting an analytic holist approach adopted a technical approach to practice, and no musicians adopting an intuitive serialist approach adopted a musical approach to practice. Versatile learners tended to adopt a mixed approach (Hallam 1995b).

Changes in approach have also been identified as expertise develops. Reid (2001) identified five conceptions of learning amongst students learning to play an instrument or sing in a higher education context. At level 1 (instruments) the students focused their attention on the technical aspects of learning and learned by copying their teachers. At level 2 (elements) the students shared similar concerns with those at level 1 but also focused on some musical elements. Students whose conceptions were at the level of musical meaning (level 3) focused on the meaning found in the music and learned by reflecting and adapting their teacher's advice. At level 4 (communicating) the focus of learning shifted to learning to communicate musical meaning and the students learned by experimenting with different styles of playing music using the teacher as only one source. At the final level (express meaning) students focused on expressing personal meanings through the music. These findings reflect those of Hallam (1995a) who identified nine levels of intellectual development in relation to the development of interpretation.

Some of the research has suggested that the adoption of particular approaches to learning leads to particular learning outcomes, while other findings suggest that similarly high-quality performing standards can be attained regardless of the approach adopted (Hallam 1995a, b; Miklaszewski 1995). These differences may depend on the context of performance.

## Environmental influences on individuality

### Characteristics of musical cultures

Historically, ethnicity determined the nature of the cultural background within which the individual grew up, although with increasing globalization this is now less the case. The impact of ethnicity on the nature of musical participation is determined by the complex interactions between the particular ethnic group to which the individual belongs and the extent to which its cultural traditions have remained intact, adapted, or been superseded by those of the wider global culture. Differences between musical cultures are evident in three ways, each of which can impact on the differences between musicians. These differences concern the music itself including genre, the kinds of behavior associated with making, teaching and listening to the music (the use of notation, aural traditions, relationship between composition and performance, the role of improvisation, formal or informal learning and teaching, group versus individual tuition) and ideas about music and its place in society (Lundquist 2002). Within each musical culture these dimensions form a dynamic system.

Different genres of music are associated with different expectations of performance. In some genres participation in performance is regarded as normal for all players, and the performance context is often dynamic and unpredictable (Cope 2002). In classical Western music the ultimate goal of learning to play an instrument has been viewed as being able to perform in public, although this can be interpreted in many ways from playing in a school concert to performing as an international soloist. Few players achieve the latter (Cope and Smith 1997). Within this tradition technical competence is highly prized, whereas in folk cultures, musicians can regard themselves as competent to play within their chosen musical and social context with relatively little technical expertise (Cope 2002). In popular music, the use of regular recording and listening back to it as a means of reflecting on and evaluating progress and the musical product is increasingly used, and a relationship between this and the development of critical reflection and self-evaluation of musical skills has been suggested (Boespflug 1999; Lebler 2006). This contrasts with the process of exchange between teacher and student which dominates apprenticeship in Western and Indian classical music.

In musical cultures where participation in music-making through singing, playing and dance is integral to social participation, the nature of the support provided for learning in music is a normal part of daily life, embedded in the community, and resulting in young and old alike acquiring an extensive repertoire and often able to play several instruments (see Blacking 1967; Walker 2006). Within the Western classical tradition, where an individual typically spends more time in isolated practice of an instrument, support for learning tends to be more dependent on individual family members, teachers and particular contexts of music-making. In this environment, lifelong commitment to music (involving intrinsic motivation) in adults has been shown to depend on childhood conditions which were supportive of music (see Chapter 28).

Within Western, post-industrial societies, the role of music has changed dramatically in the last 150 years as a result of changing patterns in work, leisure and social interaction; the development of technology, the recording industry and mass media; and improvements in the standard of living. In the UK, for example, many opportunities for music-making in community centres and within family settings have been replaced by a more solitary consumption of recorded music (Zillmann and Gan 1997), although alongside this there are often more opportunities for engagement with music within education (Hallam *et al.* 2005).

### Socio-economic status

Socio-economic status is an important determinant of the type of music with which individuals

engage (North and Hargreaves 2007) and plays a key role in musical enculturation. It also impacts on the extent of active engagement in music (Brandstrom 2000) and dropout from instrumental music lessons (McCarthy 1980; Klinedinst 1991). This may be due in part to financial considerations as it can be expensive to provide instruments, pay for tuition, support concert attendance, and create appropriate facilities for practice. Children from middle class backgrounds are more likely to engage with classical music than those from poorer backgrounds.

## The way forward: theoretical approaches conceptualizing interactions between the individual and the environment

Research to date exploring individuality in relation to engagement with and learning in music has tended to neglect consideration of the complex interactions which occur between the individual and their environment. Key to developing understanding of these interactions are biosocial models and systems theory. Below we describe two such models: an ecological systems approach to human development and a biosocial model of interaction.

### An ecological systems model of human development

Bronfenbrenner (1979) proposed an ecological model of human development, in which the notion of situations or environments is conceptualized as a series of nested structures, each structure itself being a system interacting with other systems (see Figure 26.1). At the centre Bronfenbrenner conceptualized a micro-system consisting of an individual in their immediate environment. For a teenager teaching himself the guitar informally in a pop group the micro-system might include the teenager, the group, and those family members having a direct impact on the individual's musical development. The meso-system refers to the individual's interactions in a wider group of settings.

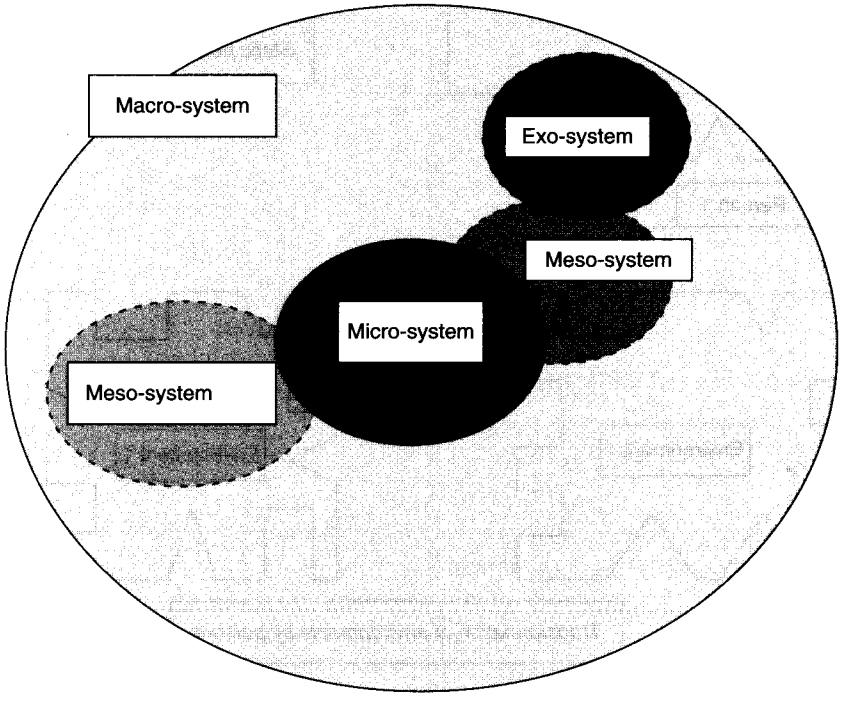



**Fig. 26.1** Bronfenbrenner's nested systems.

engage (North and Hargreaves 2007) and plays a key role in musical enculturation. It also impacts on the extent of active engagement in music (Brandstrom 2000) and dropout from instrumental music lessons (McCarthy 1980; Klinedinst 1991). This may be due in part to financial considerations as it can be expensive to provide instruments, pay for tuition, support concert attendance, and create appropriate facilities for practice. Children from middle class backgrounds are more likely to engage with classical music than those from poorer backgrounds.

## The way forward: theoretical approaches conceptualizing interactions between the individual and the environment

Research to date exploring individuality in relation to engagement with and learning in music has tended to neglect consideration of the complex interactions which occur between the individual and their environment. Key to developing understanding of these interactions are biosocial models and systems theory. Below we describe two such models: an ecological systems approach to human development and a biosocial model of interaction.

### An ecological systems model of human development

Bronfenbrenner (1979) proposed an ecological model of human development, in which the notion of situations or environments is conceptualized as a series of nested structures, each structure itself being a system interacting with other systems (see Figure 26.1). At the centre Bronfenbrenner conceptualized a micro-system consisting of an individual in their immediate environment. For a teenager teaching himself the guitar informally in a pop group the micro-system might include the teenager, the group, and those family members having a direct impact on the individual's musical development. The meso-system refers to the individual's interactions in a wider group of settings.

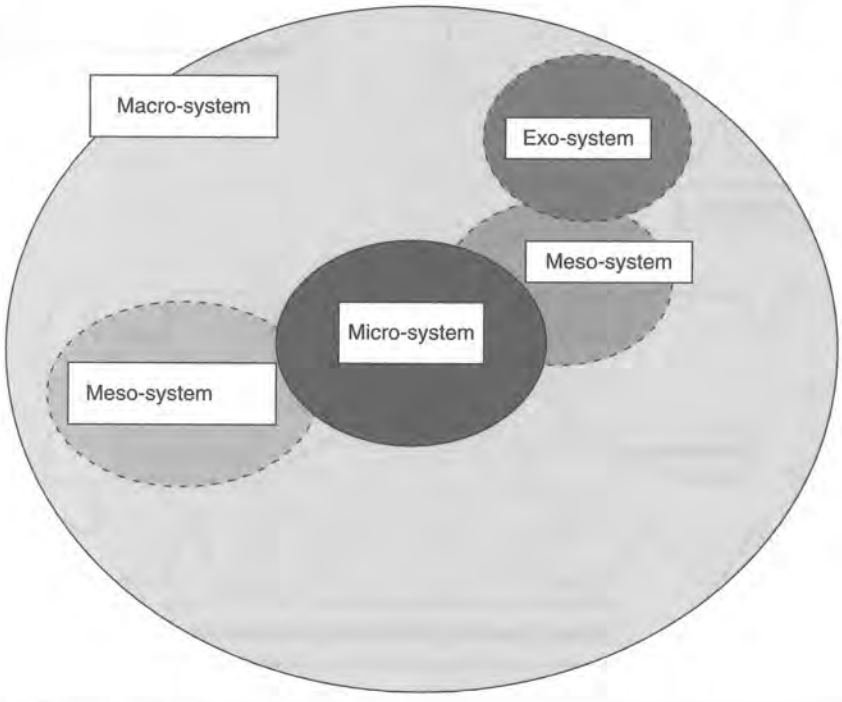



**Fig. 26.1** Bronfenbrenner's nested systems.

In the example described above, a meso-system could consist of the teenager's experiences relating to extra-curricular activities at school or through participation in a youth group which facilitated musical activities. Bronfenbrenner also suggests that the individual might be influenced by activity in what he termed an exo-system, a system in which the individual does not interact with others directly but in which another person close to the individual interacts. In the example above, this might be the influence of a parent's working life on a child. For instance, a parent working long hours in an entrepreneurial role might stimulate the teenager to become more committed to the pop group because it offered the potential for commercial success, while also providing an activity in the parent's absence. Finally, all three systems are nested within a macro-system defined as a particular subculture in which particular beliefs, values and ideologies of the lower order systems will be embodied, for instance, a Western culture. An important characteristic of the ecological model is the premise that the interconnections between the different systems within the model are critical to individual development and as important as events taking place within any one system.

## A biosocial model of interaction

Adopting a different perspective, Hettema and Kenrick (1992) outline six categories of interaction: static person–environment mesh (the individual is situated in an unchanging environment); choice of environments by persons (the individual selects new environments which meet their needs); choice of persons by environments (typified by a variety of selection processes, for instance in education and the work place);

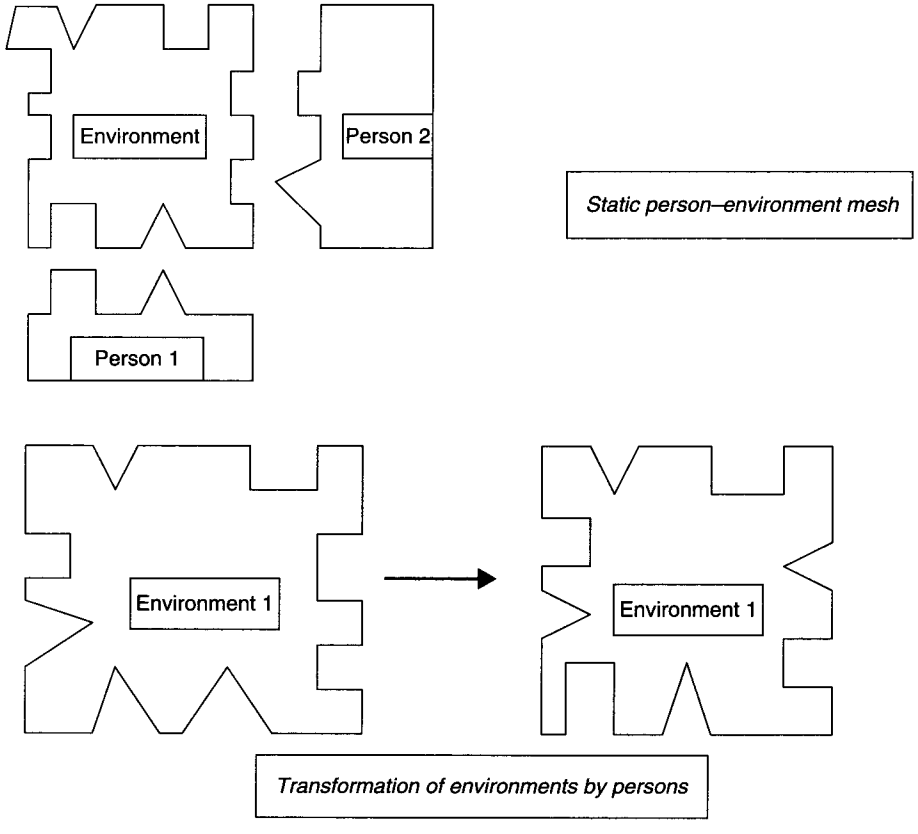



**Fig. 26.2** A static person–environment mesh and transformation of environment by person mesh (adapted from Hettema and Kenrick 1992).

transformation of environments by persons (individuals through their actions change environments, for instance, through leadership or disruption); transformation of persons by environments (individuals are socialized into new environments); and person–environment transactions or mutual transactions in which both persons and environments change over time (Hettema and Kenrick 1992). Each category represents different degrees of fit between the individual and the environment. Two examples are given in Figure 26.2. Such interactions influence the development of individuals and the environments within which they find themselves.

### A combined model

An example of how the two models may operate simultaneously is shown in Figure 26.3. The presented scenario is based on a teenager teaching themselves the guitar in a pop group formed with a group of friends. Combining the two models highlights the dynamic relationships between the individual and the environment, and suggests the need for research approaches which attempt to develop a framework for understanding the multilayered factors involved in the acquisition of musical skills and the complex relationships between them.

## Conclusions

Much of the research exploring individuality has neglected consideration of the dynamic interactions between the individual and their environment which determine the nature of particular behaviours. To increase our understanding of musical learning studies need to take account of the interactions between different elements, map the relationships between micro-, meso-, exo- and macro-systems, and explore the ways that individuals negotiate their relationships with the musical environments to which they have access. Ideally, this would be through longitudinal research which documented individuals' engagement with music through the lifecourse in different learning and performance

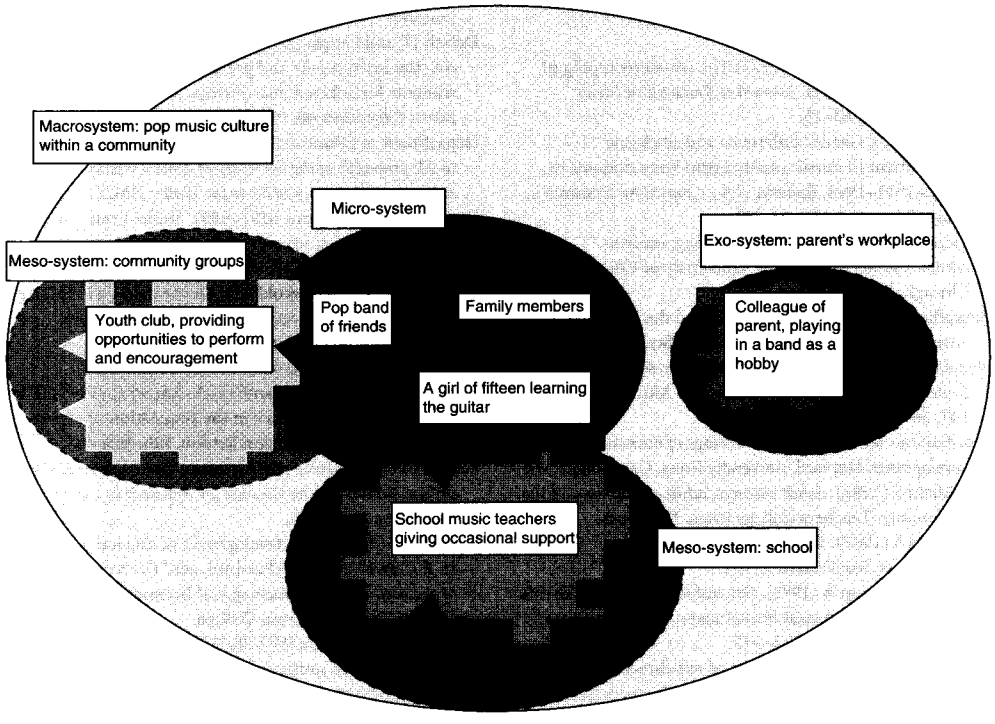



**Fig. 26.3** A biosocial systems model of individual musical development.

transformation of environments by persons (individuals through their actions change environments, for instance, through leadership or disruption); transformation of persons by environments (individuals are socialized into new environments); and person–environment transactions or mutual transactions in which both persons and environments change over time (Hettema and Kenrick 1992). Each category represents different degrees of fit between the individual and the environment. Two examples are given in Figure 26.2. Such interactions influence the development of individuals and the environments within which they find themselves.

### A combined model

An example of how the two models may operate simultaneously is shown in Figure 26.3. The presented scenario is based on a teenager teaching themselves the guitar in a pop group formed with a group of friends. Combining the two models highlights the dynamic relationships between the individual and the environment, and suggests the need for research approaches which attempt to develop a framework for understanding the multilayered factors involved in the acquisition of musical skills and the complex relationships between them.

## Conclusions

Much of the research exploring individuality has neglected consideration of the dynamic interactions between the individual and their environment which determine the nature of particular behaviours. To increase our understanding of musical learning studies need to take account of the interactions between different elements, map the relationships between micro-, meso-, exo- and macro-systems, and explore the ways that individuals negotiate their relationships with the musical environments to which they have access. Ideally, this would be through longitudinal research which documented individuals' engagement with music through the life-course in different learning and performance

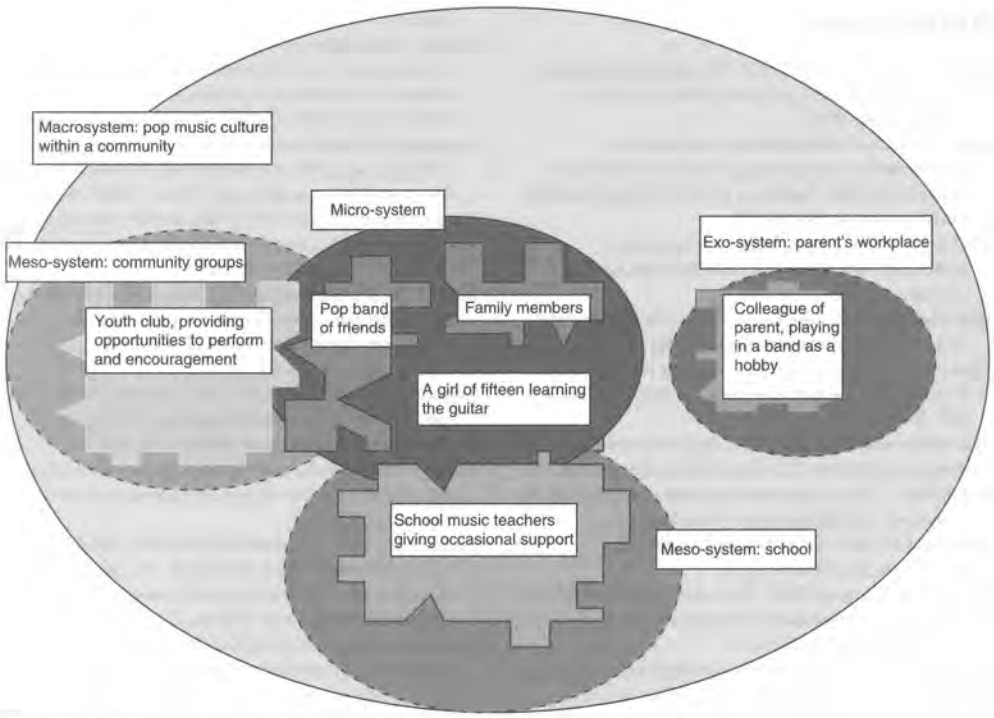



**Fig. 26.3** A biosocial systems model of individual musical development.

cultures and contexts. However, findings drawn from small-scale cross-sectional studies could be drawn together to map the impact of a range of factors contributing to our understanding of commonality and difference in the way that systems influence learning and development.

Research also needs to contextualize findings within recent global changes. The rapid development of technology and sound recording, and the expansion of the mass media have created a huge increase in music consumption, and have dictated standards of production and of 'performance' which make no allowances for the errors which may occur in 'live' performance. There has also been increased recognition of the value of active engagement with music to the individual's intellectual functioning and well-being (see Chapters 23 and 44). As a result, in the developed world, through educational and community initiatives, there has been an emphasis on music for all rather than a select few. In the longer term this will provide greater research opportunities to explore systemic interactions and their influence on a range of learning outcomes.

## References


Abeles HF and Porter SY (1978).The sex-stereotyping of musical instruments. *Journal of Research in Music Education*, **26**, 65–75.

Agak H (2002). Gender difference and academic achievement in music among Form Four students in Kenya 1991–1994. *Bulletin of the Council for Research in Music Education*, **153**, 94–101.

Blacking J (1967). *Venda children's song: study in ethnomusicological analysis*. University of Chicago, Chicago, IL.

Boespflug G (1999). Popular music and the instrumental ensemble. *Music Educators Journal*, **May**, 33–37.

Brandstrom S (2000). For whom is music school intended? *Bulletin of the Council for Research in Music Education*, **147**, 36–39.

Bronfenbrenner U (1979). *The ecology of human development*. Harvard University Press, Cambridge, MA.

Brookfield S (1983). *Adult learners, adult education and the community*. Teachers College Press, New York.

Brookfield S (1987). *Developing critical thinkers*. Jossey-Bass, San Francisco, CA.

Bruce R and Kemp A (1993). Sex stereotyping in children's preferences for musical instruments. *British Journal of Music Education*, **10**, 213–217.

Chen H (1996). An investigation of self-directed learning among non-music major adult piano learners in one-to-one piano instruction. Dissertation abstracts international 57(7), 2929A (UMI No. AAT96–35962).

Coffman D (2002). Adult education. In R. Colwell, C. Richardson, eds, *The new handbook of research on music teaching and learning*, 199–209. Oxford University Press, Oxford.

Colley A, Comber C and Hargreaves DJ (1994). Gender effects in school subject preferences: a research note. *Educational Studies*, **20**, 13–18.

Cope P (2002). Informal learning of musical instruments: the importance of social context. *Music Education Research*, **4**, 93–104.

Cope P and Smith H (1997). Cultural context in musical instrument learning. *British Journal of Music Education*, **14**, 283–289.

Crowther R and Durkin K (1982). Sex- and age-related differences in the musical behaviour, interests and attitudes towards music of 232 secondary school students. *Educational Studies*, **8**, 131–139.

Curry L (1987). *Integrating concepts of cognitive or learning style: a review with attention to psychometric standards*. Canadian College of Health Service Executives, Ottawa, Ontario.

Davidson L (1994). Song singing by young and old: a developmental approach to music. In R Aiello, ed., *Musical perceptions*, 99–130. Oxford University Press, New York.

Davidson L, Scripp L (1992). Surveying the coordinates of cognitive skills in music. In R Colwell, C Richardson, eds, *The new handbook of research on music teaching and learning*, 392–413. Schirmer Books, New York.

Davies JB (1978). *The psychology of music*. Hutchinson, London.

Delzell JK and Leppla DA (1992). Gender association of musical instruments and preferences of fourth-grade students for selected instruments. *Journal of Research in Music Education*, **40**, 93–103.

Department of Education and Skills (1991). *Music for ages 5 to 14: proposals of the Secretary of State for Education and Science and Secretary of State for Wales*. HMSO, London.

Dews CLB and Williams MS (1989). Student musicians' personality styles, stresses and coping patterns. *Psychology of Music*, **17**, 37–47.

Eccles J, Wigfield A, Harold RD, Blumenfeld P (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, **64**, 830–47.

Ericsson KA, Krampe RT and Tesch-Romer C (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, **100**, 363–406.

Gates JT (1991). Music participation: theory, research, and policy. *Bulletin of the Council for Research in Music Education*, **109**, 1–35.

Gembris H (2002). The development of musical abilities. In R Colwell and C Richardson, eds, *The new handbook of research on music teaching and learning*, 487–508. Oxford University Press, Oxford.

Gembris H., Davidson JW (2002) The role of the environment in musical development. In R Parncutt, GE McPherson, eds, *The science and psychology of musical performance*, 17–30. Oxford University Press, New York.

Green L (1997). *Music, gender and education.* Cambridge University Press, New York.

Green L (2001). *How popular musicians learn: a way ahead.* Ashgate, Aldershott.

Hallam S (1995a). Professional musicians' approaches to the learning and interpretation of music. *Psychology of Music*, **23**, 111–128.

Hallam S (1995b). Professional musicians' orientations to practice: implications for teaching. *British Journal of Music Education*, **12**, 3–19.

Hallam S (2001a). The development of metacognition in musicians: implications for education. *The British Journal of Music Education*, **18**, 27–39.

Hallam S (2004). Sex differences in the factors which predict musical attainment in school aged students. *Bulletin of the Council for Research in Music Education*, **161**, 107–117.

Hallam S, Rogers L and Creech A (2005). *Survey of local authority music services 2005. Research report 700.* Department for Education and Skills, London.

Hallam S, Rogers L and Creech A (2008). Gender differences in musical instrument choice. *International Journal for Music Education*, **26**, 7–19.

Hanley B (1998). Gender in secondary music education in British Columbia. *British Journal of Music Education*, **15**, 51–6.

Hargreaves DJ, Comber C and Colley A (1995). Effects of age, gender and training on musical preferences of British secondary school students. *Journal of Research in Music Education*, **43**, 242–250.

Harris T, Harris S, Rubin JS, Howard DM (1998). *The voice clinic handbook.* Whurr, London.

Hettema J and Kenrick DT (1992). Models of person-situation interactions. In GV Caprara, GL Van Heck, eds, *Modern personality psychology: critical reviews and new directions*, 393–417. Harvester Wheatsheaf, New York.

Ho W-C (2001). Musical learning: differences between boys and girls in Hong Kong Chinese co-educational schools. *British Journal of Music Education*, **18**, 41–54.

Howe MJA and Sloboda J (1992). Problems experienced by talented young musicians as a result of the failure of other children to value musical accomplishments. *Gifted Education*, **8**, 16–18.

Kemp AE (1996). *The musical temperament: psychology and personality of musicians.* Oxford University Press, Oxford.

Klinedinst RE (1991). Predicting performance achievement and retention of fifth-grade instrumental students. *Journal of Research in Music Education*, **39**, 225–238.

Lebler D (2006). Student as master? Reflections on a learning innovation in popular music pedagogy. Paper presented at the International Society for Music Education World Conference, Kuala Lumpur, Malaysia, 17–21 July, 2006.

Lundquist BR (2002). Music, culture, curriculum and instruction. In R Colwell and C Richardson, eds, *The new handbook of research on music teaching and learning*, 626–647. Oxford University Press, Oxford.

Lynch MP, Eilers RE, Oller DK and Urbano RC (1990). Innateness, experience, and music perception. *Psychological Science*, **1**, 272–276.

MacDonald RAR and Wilson GB (2005). Musical identities of professional jazz musicians: a focus group investigation. *Psychology of Music*, **33**, 395–417.

McCarthy JF (1980). Individualised instruction, student achievement and drop out in an urban elementary instrumental music program. *Journal of Research in Music Education*, **28**, 59–69.

McKeage KM (2004). Gender and participation in high school and college instrumental jazz ensembles. *Journal of Research in Music Education*, **52**, 343–356.

Miklaszewski K (1995). Individual differences in preparing a musical composition for public performance. In M Manturzewska, K Miklaszewski and A Biatkowski, eds, *Psychology of music today*, 138–147. Fryderyk Chopin Academy of Music, Warsaw.

North AC and Hargreaves DJ (2007). Lifestyle correlates of musical preference: a. Relationships, living arrangements, beliefs and crime. *Psychology of Music*, **35**, 58–87.

O'Neill SA and Boulton MJ (1995). Is there a gender bias towards musical instruments: a function of gender? *Psychology of Music*, **24**, 171–183.

Reid A (2001). Variation in the way that instrumental and vocal students experience learning music. *Music Education Research*, **3**, 25–40.

Sheldon DA and Price HE (2005). Sex and instrumentation distribution in an international cross-section of wind and percussion ensembles. *Bulletin of the Council for Research in Music Education*, **163**, 43–51.

Shuter-Dyson R and Gabriel S (1981). *The psychology of musical ability*, 2nd edn. Methuen, London.

Sinsel TJ, Dixon WE and Blades-Zeller E (1997). Psychological sex type and preferences for musical instruments in fourth and fifth graders. *Journal of Research in Music Education*, **45**, 390–401.

Smith A M, Maragos A and Van Dyke A (2000). Psychology of the musician. In R Tubiana, PC Amadio, eds, *Medical problems of the instrumentalist musician*, 135–170. Martin Dunitz Ltd, London.

Thurman L and Welch G (2000). *Body mind and voice: foundations of voice education.* The VoiceCare Network, St. John's University, Minnesota, MN.

Tobacyk JJ and Downs A (1986). Personal construct threat and irrational beliefs as cognitive predictors of increases in musical anxiety. *Journal of Personality and Social Psychology*, **51**, 779–82.

Walker R (2006) Cultural traditions. In GE McPherson, ed., *The child as musician: a handbook of musical development*, 439–460. Oxford University Press, Oxford.

Welch GF and Sundberg J (2002). Solo voice. In R Parncutt, G McPherson, eds, *The science and psychology of music performance: creative strategies for teaching and learning*, 253–268. Oxford University Press, Oxford.

Whitchurch GG and Constantine LL (1993). Systems theory. In PG Boss, WJ Doherty, R Laroosa, WR Schumm, S Steinmetz, eds, *Source book of family theories: a contextual approach*, 325–352. Plenum Press, New York.

Wills G and Cooper CL (1988). *Pressure sensitive: popular musicians under stress.* Sage, London.

Wilson GD (1984). The personality of opera singers. *Personality and Individual Differences*, **5**, 195–201.

Zenatti A (1993). Children's musical cognition and taste. In TJ Tighe, WJ Dowling, eds, *Psychology of music: the understanding of melody and rhythm*, 177–196. Erlbaum, Hillsdale, NJ.

Zervoudakes J and Tanur J (1994). Gender and musical instruments: winds of change. *Journal of Research in Music Education*, **42**, 58–67.

Zillmann D and Gan S-L (1997). Musical taste in adolescence. In DJ Hargreaves, AC North, eds, *The social psychology of music*, 161–187. Oxford University Press, Oxford.

# CHAPTER 27

# Motivation to learn

Susan Hallam

## The historical perspective

Human motivation is extremely complex. Historically, numerous theories have developed in attempts to explain it. These, to varying degrees, emphasize motivation as deriving from within the individual, within the environment, or as a complex interaction between the two mediated by cognition. The most recent theories emphasize the way that our perceptions of events are determined by our construction of them, these interpretations subsequently influencing our self-esteem, self-efficacy and motivation. They acknowledge the capacity of individuals to determine their own behaviour, whilst also recognizing the role of the environment in rewarding or punishing particular behaviours influencing subsequent cognitions and later actions. There has also been increasing recognition that motivation operates at different levels and across different timescales (for more detail of generic theories of motivation and earlier reviews of motivation in relation to music see Asmus 1994; O'Neill and McPherson 2002; Hallam 2002; Austin *et al.* 2006).

Much of the early research on motivation in music was not embedded within any specific motivational research paradigm or theoretical position, although expectancy-value models which have been proposed to explain motivation for particular tasks in education (e.g., Eccles 1983) have provided a framework for some recent work. The research has also focused almost exclusively on motivation to learn and continue to play an instrument. There has been little interest in motivation to listen to music, compose, or engage in other musical activities.

## A framework for understanding motivation in music

In this chapter, I set out a model which attempts to integrate the various theoretical approaches to understanding motivation, embedded within a broadly systemic approach as proposed by Bronfenbrenner (1979), which suggests that the process of human development depends on mutual accommodation which occurs throughout the life-course between an individual and the various systems which they or others close to them encounter in their environment (see Chapter 26 for details).

Figure 27.1 sets out a framework which illustrates the complex interactions which occur over time in relation to motivation. Certain aspects of our individuality are predetermined, for instance, our biological temperament, our sex, and our age. These are shaped through interaction with the environment to develop our personality, gender identity, cognitive processes, and our self-perceptions. We are motivated because we desire social approval, particularly from those we admire and respect. Such praise from others is internalized, raises self-esteem, and enhances confidence. Some environmental influences are internalized to such an extent that they come to affect the individual's functioning over time in a fairly consistent way. Individuals set themselves goals, which determine their behaviour. These goals are influenced by individual and self-perception characteristics as well as environmental factors. Where the environment satisfies individual needs and

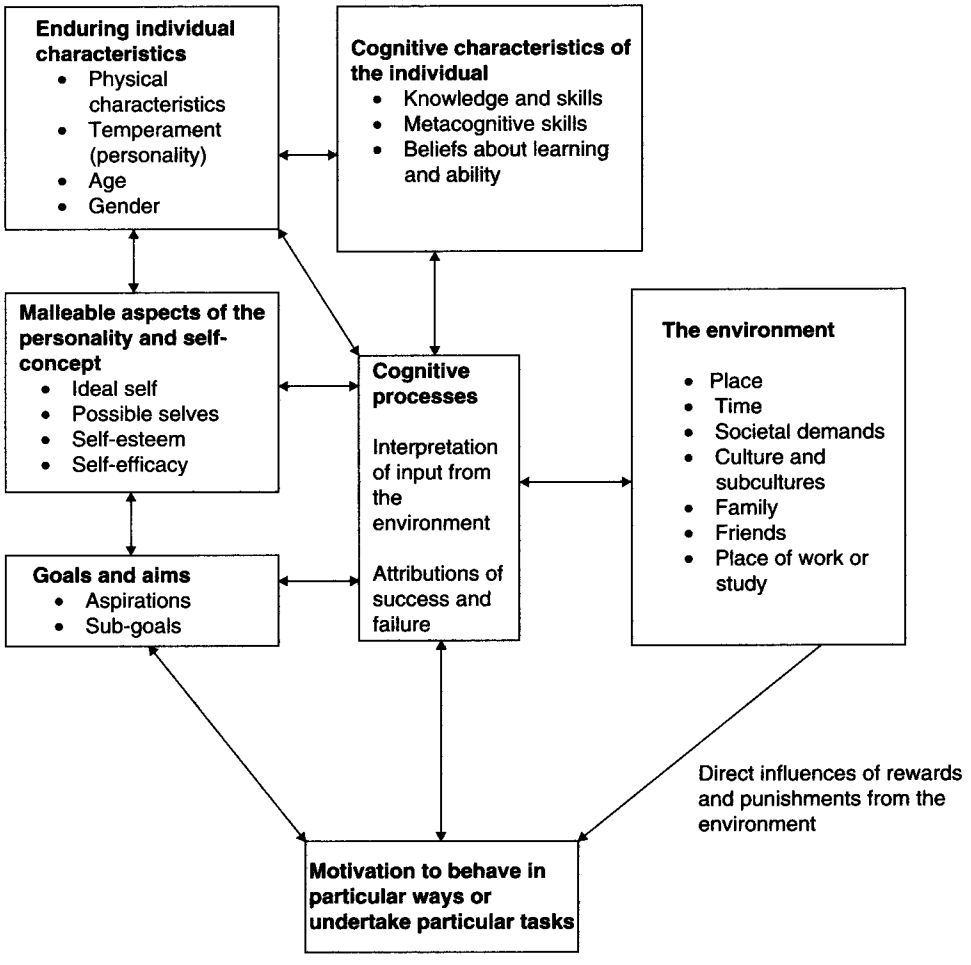



**Fig. 27.1** Interactions between individual and environmental factors in determining motivation.

facilitates personal goals motivation is likely to be enhanced. Where the environment presents obstacles, the individual may give up or be spurred on to greater efforts to overcome them, perhaps by finding a more conducive environment. Behaviour is the end link in the chain, but at the time of enactment it too can be influenced and changed by environmental factors. There is interaction between the environment and the individual at every level and in the long and short term. Individuals can act upon the environment to change it, or seek out new environments more conducive to their needs (Hettema and Kenrick 1992). The model recognizes the importance of cognitive factors and self-determination in behaviour. While we have needs and desires we are aware that we need to consider the consequences of our actions before we attempt to satisfy them. Cognition plays a role in the ways in which we attempt to enhance our self-esteem, leading us to attribute our success or failure to causes which will allow us to maintain a consistent view of ourselves. When a learner has completed a learning task successfully this will have an impact on self-esteem and motivation which will be carried forward to subsequent learning tasks. Conversely, when learning outcomes are negative, motivation may be impaired. In the following sections I provide a more detailed account of what we know about each of the elements

outlined in the model as they relate to motivation in music.

## Individual characteristics

### Enduring individual characteristics: physical, gender and personality factors

Chapter 26 considers many elements of individuality which may influence the nature of engagement with music including physical, gender, and personality characteristics. These are not discussed further here. The focus in this section is the complex and multifaceted nature of motivation to become actively involved in music-making, which includes the way that it intrinsically acts to meet needs which vary between individuals in relation to their personality characteristics. For instance, Motte-Haber (1984) has suggested a range of internal motives including the desire for achievement, curiosity, and self-actualization, while Nagel (1987) stressed the need for personal fulfilment met by the emotion-inducing quality of music, satisfaction of a positive social response within performance settings, exploration of aggressive drives through the exploitation of the motor skills entailed in musical performance, and some voyeuristic and exhibitionist desires. Persson (1993; Persson *et al.*1996) studying pianists also emphasized the importance of hedonic drive, while acknowledging the role of social and achievement motives. Gellrich and colleagues (1986) identified three achievement-related motives: a general achievement motivation; a specifically music-oriented achievement orientation; and a sensual–aesthetic motive, the pleasure and joy of playing certain pieces of music. Musicians it seems derive considerable personal fulfilment from the act of making music, in addition to the social rewards that it offers. The balance between these motives may change over time as individuals progress through their musical careers (Sosniak 1985; Harnischmacher 1997; Manturzewska 1990) and may contribute to determining individual career trajectories.

The diversity of individual motivation has been explored in relation to amateur musicians in addition to professionals. Some amateurs view music as personal amusement, while for others it constitutes serious leisure only distinguishable from the work of professionals because it is not their livelihood (Gates 1991). For some music is a key element of their identity (Pitts 2005) and they invest huge energy and creativity in it (Finnegan 1997). Their amateur status suggests that they find their musical activities intrinsically motivating, although social interaction can also be an important motivator.

### Malleable aspects of the personality and self-concept

An individual's identity or self-concept represents the way they thinks about themself and their relationships with others (Rogers 1961) (see Chapter 43) and plays a crucial role in motivation. Behaviour is influenced by the individual's interpretation of situations and events; their expectations; and the goals that they have set for themselves related to their identity. Music self-concept has been shown to be linked to motivation, interest, and participation in school and out of school musical activities (Austin 1991) and develops early as quite young children are able to assess how well they can complete particular musical tasks (Greenberg 1970). However, being good at something does not necessarily transfer to interest in it. Asmus and Harrison (1990) working with non-music major college students found no relationship between music motivation and aptitude. They argued that the drive for music stems from love of music.

Children's self-concept in relation to music at school and the value that they place on music appear to decline as they get older, the latter more than for reading and mathematics (Wigfield *et al.* 1997; Mota 1999). However, those adolescents who are actively engaged in music-making through playing an instrument favour this activity over almost all others. Being a musician becomes an important element of their identity. As children progress through school they become increasingly aware of their own capabilities through making comparisons and receiving feedback from others. Where comparisons are made with high-attaining others self-concept is likely to be deflated. This process is influenced by the context of learning, whether it is supportive or critical.

Self-efficacy beliefs based on our evaluation of the likelihood of our success in relation to

particular tasks or areas of work also play a part in determining whether particular goals will be pursued and subsequently achieved (Bandura 1982). McPherson and McCormick (1999, 2006) studied 190 pianists aged 9–18 before taking an examination and found that self-efficacy prior to entering the examination predicted their examination result more strongly than any other factors. Bandura (1989) suggested that motivation for an activity is at its peak when strong self-efficacy beliefs are combined with some moderate uncertainty about the outcome, i.e. when a person feels competent but challenged. However, for intrinsic motivation to flourish, feelings of self-determination are also necessary. So even if an activity is challenging and a person feels themself to be highly competent, they will not display high levels of intrinsic motivation unless the activity is perceived as self-determined (Ryan and Deci 2000). Evidence of the links between intrinsic motivation, self-determination, and the use of self-regulating practising strategies support this (Austin *et al.* 2006). Renwick and McPherson (2002) demonstrated how a 12-year-old practised with heightened attention, persistence and enhanced strategy use when working on a piece which she had chosen herself as opposed to one chosen by the teacher. The extent to which learners value what they are doing predicts engagement across many different facets of musical practice (O'Neill 1999) and contributes to prediction of examination results more strongly than weekly practice or anxiety (McPherson and McCormick 2000).

### Goals and aims

The goals and aims that individuals strive to attain are related to their identity, self-concept, self-efficacy and what they believe is possible for them. These take account of both context and cognition. If an individual perceives themself as successful and attributes this success to high ability they may come to include in their self-concept a 'positive possible future self in that domain' (Markus and Ruvolo 1989). Possible selves can be powerful motivators, providing long-term goals and encouraging the setting up of interim goals which need to be achieved en route. If an individual does not have a positive possible self as a musician (professional, amateur or listener) in the long, medium, or short term they are unlikely to maintain their interest in music. This is illustrated by McPherson and McCormick (2000) who found that for many children playing an instrument was no different from participating in a team sport, taking up a hobby, or pursuing other interests. Very few saw it as something that could possibly lead to a future career. MacNamara and colleagues (2006) also showed that career planning was in evidence in the very earliest stages of learning for nine professional musicians along with dedication, commitment, determination and a willingness to make sacrifices. Self-belief became increasingly important as they encountered more critical appraisal and greater competition as they progressed (MacNamara *et al.* 2006). However, goals can sometimes conflict with each other and their fulfilment can be disrupted by others (Harnischmacher 1995). Individuals may have to make trade-offs between goals at different levels and undertake some activities to attain a particular goal which they may not find particularly enjoyable. For instance, young people wishing to become professional musicians do not always enjoy practising alone and may require parental support and encouragement to do so (Howe and Sloboda 1991).

## Cognition: individual characteristics and processes

All modern theories of motivation take account of cognition—an acceptance that much of our behaviour is mediated by our thoughts about and perceptions of events (Kelly 1955). This process is influenced, in part, by locus of control, the extent to which individuals perceive that they have control over situations (Rotter 1966). While prior knowledge and skills in a domain are powerful determinants of performance in that domain, our beliefs about our current capabilities also play a part. Also important are our beliefs about the nature of learning. In mainstream education a distinction has been made between performance and learning goals (Elliott and Dweck 1988), the former concerned with gaining positive judgements of competence as compared with others and avoiding negative ones,

the latter with increasing mastery, reflecting the desire to learn new skills, master new tasks or understand new things.

In learning in music, the relationship between these goals is complex. For instance, in studies of band students, Austin (1988, 1991) found that those assigned to a competitive condition exhibited levels of motivation and performance achievement comparable to those in a non-competitive condition, while Schmidt (2005) reported that instrumental students defined their success in relation to mastery and cooperative orientations, placing less emphasis on competitive and ego orientations. The context of active participation in music seems to influence the goals adopted, some contexts supporting the development of both types of goals. Sandene (1998) found that students' goal orientations were related to the perceived motivational climate in the classroom with overt teacher behaviours, in particular the ratio of positive to negative feedback, being particularly important.

To succeed in music clearly requires mastery behaviour to sustain the motivation for many hours of practice. For instance, O'Neill (1997), studying beginner instrumentalists, found that children who experienced mastery-oriented behaviour after experiencing failure on a problem-solving task made more progress on their instrument than children who initially displayed helpless behaviour. Research from other domains has shown a relationship between students' theories of intelligence and their goal choices. Where students hold an entity theory of intelligence (fixed and immutable) they are more likely to adopt performance goals, while those holding an incremental view of intelligence are more likely to choose a learning goal (Dweck and Leggett 1988). There is potential for exploring these relationships in novice, expert, amateur and professional musicians (see also Austin *et al.* 2006).

Individuals are motivated to establish, maintain and promote a consistent and usually positive self-image, so they develop a variety of coping strategies to maintain self-worth, some of which may be self-defeating, for instance, reducing effort (Covington 1984). How individuals attribute successes and failures is important in maintaining self-esteem (Weiner 1986). The causes of success or failure can be seen as stable or unstable; controllable or uncontrollable and internal or external. Overall, five major attribution categories have been found in music: effort, musical background, classroom environment, musical ability, and affect for music (Asmus 1986a), although findings specifically related to performance in an examination also included effort in preparation, effort in the examination, nervousness, luck, and task difficulty (McPherson and McCormick 2000). Musical ability and effort are the most frequently cited attributions by music students (Austin and Vispoel 1998). Highly motivated music students tend to make effort attributions, while students with low motivation cite ability (e.g. Austin and Vispoel 1998). These findings seem to be broadly consistent across grade levels (Asmus 1986b; McPherson and McCormick 2000), school settings, and music populations, although there is some evidence that ability attributions become more frequent as children get older (Arnold 1997; Asmus 1986b). Harter (1985) proposes that what is important in maintaining self-esteem and motivation is beneffectance. This involves attributing successful outcomes to internal causes and unsuccessful outcomes to external ones. This is likely to be important in sustaining motivation in musical activities. Considering the effect of attributing success or failure to the use of particular learning strategies, Vispoel and Austin (1993) found that explaining failure in terms of the adoption of less than optimal learning strategies was effective in improving these and increasing effort.

Metacognition (the term given to our knowledge of our own learning) is relevant for motivation insofar as it indicates our awareness of our own strengths and weaknesses and the ways in which we learn best, and may be implicated in the way that we manage our attributions. Metacognitive strategies are concerned with the planning, monitoring and evaluation of learning and performance. For most musicians life is dominated by public performance and preparation for it. As practice is not always intrinsically motivating, developing strategies for managing motivation is crucial. Preparing for public performance necessitates giving priority to practice (Manturzewska 1969), mobilizing arousal specifically for performance (Bochkaryov 1975) and managing anxiety (Hamann 1982). These all

depend on the development of appropriate metacognitive skills.

## The environment

The environment is crucial in determining the opportunities that individuals have to engage with music and the extent to which they will be supported while doing so. Music is not valued equally in all cultures. In some it is viewed as decadent and is forbidden. In others it is highly valued and those involved in its composition or execution are highly revered members of society. Economic, demographic and political factors can have a major impact on the opportunities that may be available for musical engagement (Simonton 1997). Over time the value placed on music can change, and within any particular culture different types of music may be valued by specific subgroups. We know very little about how these cultural and societal factors mediate motivation to engage with music or the type of activity selected. In the Western world some individuals learn to play an instrument because it is expected by their family or school. Others start by chance because tuition is on offer in school or their friends have decided to play.

Although most individuals are involved in full time education when they take up a musical instrument, there has been almost no research on the relationship between the institutional learning environment and motivation to engage with music. There is some evidence that in a conservatoire environment there are relationships between institutional and departmental expectations and the amount of time spent practising which are mediated by the dominating value system (Jorgensen 1997), and that in mainstream education support within schools from generalist classroom teachers and the senior management team impacts on the extent to which instrumental teaching is effective (Hallam and Prince 2000). Teachers play an important role in motivating students and can contribute towards pupils' attitudes and subsequent attainment (Szubertowska 2005). Where teachers motivate pupils to engage with music, identities as musicians develop leading to more positive attitudes towards school music and teachers (Lamont 2002). In adolescence the peer group is very powerful and can bring negative pressure to bear in relation to engagement with some types of music (Finnas 1987; 1989). To withstand this musical identities need to be well developed. The influence of early teachers who are viewed as warm and sympathetic seems to be particularly important in this respect (Sosniak 1985; Sloboda and Howe 1991). Relatively uncritical encouragement in the early stages of engagement with music encourages the development of a positive musical identity. Once this is established later teachers provide high-status role models with whom the young musician can identify and emulate (Manturzewska 1990).

There is considerable evidence that parents have a vital influence on children's motivation for involvement with music and the acquisition of musical skill (see Chapter 28). In the early years the family are likely to be the main source of musical stimulation and may be key to engendering future interest in music. The age at which children first sing is related to the number of musical behaviours initiated by the parents (Howe *et al.* 1995) and the development of perfect pitch seems to occur with particularly systematic exposure to music in early childhood (Sergeant 1969). The influence of parents can also contribute to the choice of instrument to be played (Fortney *et al.* 1993) and their ongoing support plays a crucial role in whether children persist and commit to musical engagement in the long term. While the literature as a whole indicates that having a highly supportive and encouraging home background is important, it is clearly not essential. Highly effective individuals in a range of domains have been found to have histories marked by severe frustration, deprivation and traumatic experiences (MacKinnon 1965). Another important contributor to motivation is the relationship between teachers, parents and learners (see Chapter 28). The most motivated students are those who work in harmony with their teachers and parents.

## Direct influences of rewards and punishments from the environment

Intrinsic motivation is a crucial aspect of developing self-identity as a musician. This may take

several forms, but one key element is enjoyment of the experience of engagement with music. Tasks that are intrinsically motivating share certain structural and emotional characteristics (Csikszentmihalyi 1990), offering a level of challenge that is in balance with a person's current skills (Good and Brophy 1991). When this occurs an individual derives pleasure from the work and tends to continue with it. This is known as a state of flow. If the task is too easy the person becomes bored. If the work demands skills beyond the capabilities of the individual anxiety is created. Csikszentmihalyi and colleagues (1993) have shown that students experience a state of flow most often when participating in chosen activities such as music as opposed to academic-oriented activities or when interacting with peers. Custodero (1999) observed 4- and 5-year-old children in a music classroom over 8 weeks and flow experience was associated with high self-concept or skill, perceived challenge, and active engagement.

Not all musical activities are intrinsically motivating for all people. For instance, novice and professional musicians exhibit diversity in motivation to practise. Many students require parental encouragement to practise (Howe and Sloboda 1991) or other rewards including playing favourite pieces of music or pleasing teachers. Practising for love of the instrument is rare (Harnischmacher 1995). A study of the motivation of beginner instrumentalists found that where the children remained enthusiastic about continuing to play their involvement centred around the instrument itself and the repertoire that they were learning, with less motivated children referring to participation in a band, or the opinions of their parents and friends as shaping their own attitudes (Pitts *et al.* 2000).

## Lack of motivation

For a full understanding of musical motivation, the study of drop-outs is important. Those ceasing to play tend to do less practice, have attained less (Sloboda *et al.* 1996; Hallam 1998) and perceive themselves as less musically able, receiving less family encouragement and having greater strengths in other recreational activities (Frakes 1984). The time costs of playing an instrument are too great in relation to the rewards that they receive (Hurley 1995). Overall, no single explanatory factor emerges in explaining drop-out but several have been identified including lower socio-economic status, poor ability to understand instructions, and poor self-concept in relation to reading music, maths, reading achievement, and general scholastic attainment (Klinedinst 1991; Hallam 1998). Research to date has tended to focus on participation in formal music tuition in the Western classical tradition. Different factors may apply to those dropping out of other musical activities.

## Future research directions

The evidence presented above suggests that motivation to be involved in active music-making is determined by complex interactions between the individual and the environment within which they find themselves. Some of the environmental effects, in particular those relating to early musical experiences, learning outcomes, self-efficacy and subsequently self-esteem are internalized by the individual in such a way that they become part of that individual's characteristics, rendering it both impossible and pointless to disentangle them. Once internalized they impact on motivation to continue to be involved in music. The individual's commitment to and involvement in music-making can also affect their environment and the people in it. Families may make changes to support their musical offspring, friends may be influenced to participate in making music. The transaction is bidirectional. The nature of the music profession, which is extremely competitive, means that only the highly motivated will have sufficient determination to succeed, although many more may become highly skilled amateurs.

The research to date has mainly, although not exclusively, related to school-aged students and the types of tuition on offer in those contexts. There is a need for research exploring issues of motivation in different learning contexts, informal and formal, and for learners of different ages and across different types of music. Related to this is the question of drop-out at higher levels of expertise than have so far been considered, for instance, those who study music in higher education, who do not then go on to make their living from musical activities, or indeed those

who become professional musicians and then leave the profession early. In addition, most of the research has focused on motivation to become a musician. Listening to music plays an important part in the daily lives of most people in the Western world but there is little research relating to those individuals for whom listening to music is a passion, who have extensive collections of recorded music and are extremely knowledgeable about music without necessarily actively participating in making it. There is also a need to explore why music plays little part in the lives of some people.

## References


Arnold JA (1997). A comparison of attributions for success and failure in instrumental music among sixth-, eighth-, and tenth grade students. *Update: Applications of Research in Music Education*, **15**, 19–23.

Asmus EP (1986a). Achievement motivation characteristics of music education and music therapy students as identified by attribution theory. *Bulletin of the Council for Research in Music Education*, **86**, 71–85.

Asmus EP (1986b). Student beliefs about the causes of success or failure in music: a study of achievement motivation. *Journal of Research in Music Education*, **34**, 262–278.

Asmus EP and Harrison CS (1990). Characteristics of motivation for music and musical aptitude of undergraduage nonmusic majors. *Journal of Research in Music Education*, **38**, 258–268.

Asmus EP (1994). Motivation in music teaching and learning. *The Quarterly Journal of Music Teaching and Learning*, **5**, 5–32.

Austin JR (1988). The effect of music contest format on self-concept, motivation, and attitude of elementary band students, *Journal of Research in Music Education*, **36**, 95–107.

Austin JR (1991). Competitive and non-competitive goal structures: an analysis of motivation and achievement among elementary band students. *Psychology of Music*, **19**, 142–158.

Austin JR, Renwick J and McPherson GE (2006). Developing motivation. In GE McPherson, ed., *The child as musician: a handbook of musical development*, 213–238. Oxford University Press, Oxford.

Austin JR and Vispoel WP (1998). How American adolescents interpret success and failure in classroom music: relationships among attributional beliefs, self-concepts and achievement. *Psychology of Music*, **26**, 26–45.

Bandura A (1982). Self-efficacy mechanism in human agency. *American Psychologist*, **37**, 122–147.

Bandura A (1989). Self-regulation of motivation and action through internal standards and goal systems. In LA Pervin, ed., *Goal concepts in personality and social psychology*, 19–86. Erlbaum, Hillsdale, NJ.

Bochkaryov LL (1975). The psychological aspects of musicians' public performance. *Voprosy Psikhologii*, **21**, 68–79.

Bronfenbrenner U (1979). *The ecology of human development*. Harvard University Press, Cambridge, MA.

Covington M. (1984). The motive for self-worth. In R Ames and C Ames, eds, *Research on motivation in education: student motivation*, 77–111. Academic Press, New York.

Csikszentmihalyi M (1990). *Flow: the psychology of optimal experience*. Harper Row, New York.

Csikszentmihalyi M, Rathunde K and Whalen S (1993). *Talented teenagers: the roots of success and failure*. Cambridge University Press, Cambridge.

Custodero LA (1999). Construction of musical understandings: the cognition–flow interface. Paper presented at the Cognitive Processes of Children Engaged in Music Activity Conference, Champaign, Urbana, IL, June, 1999.

Dweck CS and Leggett EL (1988). A social cognitive approach to motivation and personality. *Psychological Review*, **95**, 256–373.

Eccles J (1983). Expectancies, values and academic behaviours. In JT Spence, ed., *Achievement and achievement motives*, 75–146. Freeman, San Francisco, CA.

Elliott ES and Dweck CS (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, **54**, 5–12.

Finnas L (1987). Do young people misjudge each other's musical taste? *Psychology of Music*, **15**, 152–166.

Finnas L (1989). A comparison between young people's privately and publicly expressed musical preferences. *Psychology of Music*, **17**, 132–145.

Finnegan R (1997). Music, performance and enactment. In H Mackay, ed., *Consumption and everyday life*, 114–158. Sage/Open University, London.

Fortney PM, Boyle JD and DeCarbo NJ (1993). A study of middle school band students' instrumental choices. *Journal of Research in Music Education*, **41**, 28–39.

Frakes L (1984). Differences in music achievement, academic achievement and attitude among participants, dropouts and non-participants in secondary school music. Unpublished Phd thesis, University of Iowa.

Gates JT (1991). Music participation: theory, research and policy. *Bulletin of the Council for Research in Music Education*, **109**, 1–35.

Gellrich M, Osterwold M and Schulz J (1986). Leistungsmotivation bei Kindern im Instumentalunterricht. Bericht uber eine erkundungsstudie. (Children's performance motivation in instrumental teaching.) *Musikpsychologie*, **3**, 33–69.

Good TL and Brophy JE (1991). *Educational psychology*. Longman, New York.

Greenberg M (1970). Musical achievement and self-concept. *Journal of Research in Music Education*, **18**, 57–64.

Hallam S (1998). Predictors of achievement and drop out in instrumental tuition. *Psychology of Music*, **26**, 116–132.

Hallam S (2002). musical motivation: towards a model synthesising the research. *Music Education Research*, **4**, 225–244.

Hallam S and Prince V (2000). *Research into instrumental music services, research report RR229*. Department for Education and Employment, London.

Hamann DL (1982). An assessment of anxiety in instrumental and vocal performances. *Journal of Research in Music Education*, **30**, 77–90.

Harnischmacher C (1995). Spiel oder Arbeit? Eine Pilotstudie zu, instrumentalen Ubeverhalten von Kindern und Jugendlichen. In Gembris H, Kraemer RD and G Maas, eds, *Musikpadagogosche Forschungsberichte 1994*, 41–56. Wisner, Augsburg.

Harnischmacher C (1997) The effects of individual differences in motivation, volition, and maturational processes on practice behaviour of young instrumentalists. In Jorgensen H and Lehman A, eds, *Does practice make perfect? Current theory and research on instrumental music practice*, 71–88. Norges musikkh gskole, Oslo.

Harter S (1985) Competence as a dimension of self-evaluation: toward a comprehension model of self-worth. In R Leahy, ed., *The development of the self*, 55–118. Academic Press, New York.

Hettema J and Kenrick DT (1992). Models of person-situation interactions. In G-V Caprar and GL Van Heck, eds, *Modern personality psychology: critical reviews and directions*, 393–417. Harvester Wheatsheaf, New York.

Howe MJA, Davidson JW, Moore DM and Sloboda JA (1995). Are there early childhood signs of musical ability? *Psychology of Music*, **23**, 162–76.

Howe MJA and Sloboda JA (1991). Young musicians' accounts of significant influences in their early lives 1. The family and the musical background. *British Journal of music Education*, **8**, 39–52.

Hurley CG (1995). Student motivations for beginning and continuing/discontinuing string music tuition. *The Quarterly Journal of Music Teaching and Learning*, **6**, 44–55.

Jorgensen H (1997). Time for practising? Higher level music students' use of time for instrumental practising. In H Jorgensen and A Lehman, eds, *Does practice make perfect? Current theory and research on instrumental music practice*, 123–140. Norges musikkh gskole, Oslo.

Kelly GA (1955). *The psychology of personal constructs*. Norton, New York.

Klinedinst RE (1991). Predicting performance achievement and retention of fifth-grade instrumental students. *Journal of Research in Music Education*, **39**, 225–238.

Lamont A (2002). Musical identities and the school environment. In RAR MacDonald, DJ Hargreaves and D Miell, eds, *Musical identities*, 41–59. Oxford University Press, Oxford.

MacKinnon DW (1965). Personality and the realization of creative talent. *American Psychologist*, **20**, 273–281.

MacNamara A, Holmes P and Collins D (2006). The pathway to excellence: the role of psychological characteristics in negotiating the challenges of musical development. *British Journal of Music Education*, **23**, 285–302.

Manturzewska M (1969). *Psychologiczne Warunki Osiagniec Pianistycznych (Psychological conditions and piano achievement)*. Ossolineum, Warsaw.

Manturzewska M (1990). A biographical study of the life-span development of professional musicians. *Psychology of Music*, **18**, 112–139.

Markus H and Ruvolo A (1989). Possible selves: personalized representations of goals. In LA Pervin, ed., *Goal concepts in personality and social psychology*, 211–242. Lawrence Erlbaum Associates, Hillsdale, NJ.

McPherson GE and McCormick J (1999). Motivational and self-regulated learning components of musical practice. *Bulletin of the Council for Research in Music Education*, **141**, 98–102.

McPherson GE and McCormick J (2000). The contribution of motivational factors to instrumental performance in a performance examination. *Research Studies in Music Education*, **15**, 31–39.

McPherson GE and McCormick J (2006). Self-efficacy and performing music. *Psychology of Music*, **34**, 322–336.

Mota G (1999). Young children's motivation in the context of classroom music: an exploratory study about the role of music content and teaching style. *Bulletin of the Council for Research in Music Education*, **141**, 119–123.

Motte-Haber H de la (1984). Die Bedeutung der Motivation fur den Instrumentalbericht (The significance of motivation in instrumental reports). *Zeitschrift fur Musikpadagogik*, **51**, 51–54.

Nagel JJ (1987). An examination of commitment to careers in music: implications for alienation from vocational choice. Unpublished doctoral dissertation, University of Michigan, Ann Arbor, MI.

O'Neill SA (1997). The role of practice in children's early musical performance achievement during the early years of learning a musical instrument. In H Jørgensen and A Lehmann, eds, *Does practice make perfect? Current theory and research on instrumental music practice*, 3–70. Norges Musikkhogskole, Oslo.

O'Neill SA (1999). The role of motivation in the practice and achievement of young musicians. In SW Yi, ed., *Music, mind and science*, 420–433. Seoul National University Press, Seoul.

O'Neill SA and McPherson GE (2002). Motivation. In R Parncutt and GE McPherson, eds, *The science and psychology of musical performance: creative strategies for teaching and learning*, 31–46. Oxford University Press, Oxford.

Persson RS (1993). The subjectivity of musical performance: an exploratory music–psychological real world enquiry into the determinants and education of musical reality. Unpublished doctoral dissertation, School of Human and Health Sciences, Huddersfield University, Huddersfield, UK.

Persson RS, Pratt G, Robson C (1996). Motivational and influential components of musical performance: a qualitative analysis. In AJ Cropley and D Dehn, eds, *Fostering the growth of high ability: European perspectives*, 287–301. Ablex, Norwood, NJ.

Pitts SE (2005). *Valuing musical participation*. Ashgate, Aldershot.

Pitts SE, Davidson JW and McPherson GE (2000). Models of success and failure in instrumental learning: case studies of young players in the first 20 months of learning. *Bulletin of The Council for Research In Music Education*, **146**, 51–69.

Renwick JM and McPherson GE (2002). Interest and choice: student-selected repertoire and its effect on practising behaviour. *British Journal of Music Education*, **19**, 173–188.

Rogers CR (1961). *On becoming a person*. Houghton Mifflin, Boston, MA.

Rotter JB (1966). Generalised expectancies for internal versus external control of reinforcement. *Psychological Monograph*, 80 (Whole No. 609).

Ryan RM and Deci EL (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well being. *American Psychologist*, **55**, 68–78.

Sandene BA (1998). An investigation of variables related to student motivation in instrumental music. *Dissertation Abstracts International*, **58**, 3870A. (UMI No. 9811178.)

Schmidt CP (2005). Relations among motivation, performance achievement, and music experience variables in secondary instrumental music students. *Journal of Research in Music Education*, **53**, 134–147.

Sergeant D (1969). Experimental investigation of absolute pitch. *Journal of Research in Music Education*, **17**, 135–43.

Simonton DK (1997). Products, persons, and periods: histrometric analyses of compositional creativity. In DJ Hargreaves and AC North, eds, *The social psychology of music*, 107–122. Oxford University Press, Oxford.

Sloboda JA, Davidson JW, Howe MJA and Moore DG (1996). The role of practice in the development of performing musicians. *British Journal of Psychology*, **87**, 287–309.

Sloboda JA and Howe MJA (1991). Biographical precursors of musical excellence: an interview study. *Psychology of Music*, **19**, 3–21.

Sosniak LA (1985). Learning to be a concert pianist. In BS Bloom, ed., *Developing talent in young people*, 19–67. Ballentine, New York.

Szubertowska E (2005). Education and the music culture of Polish adolescence. *Psychology of Music*, **33**, 317–330.

Vispoel WP and Austin JR (1993). Constructive response to failure in music: the role of attribution feedback and classroom goal structure. *British Journal of Educational Psychology*, **63**, 110–129.

Weiner B (1986) *An attributional theory of motivation and emotion*. Springer-Verlag, New York.

Wigfield A, Eccles JS, Yoon KS, Harold RD, Arbreton AJA, Freedman-Doan C and Blumenfield PC (1997). Changes in children's competence beliefs and subjective task values across the elementary school years: a 3-year study. *Journal of Educational Psychology*, **89**, 451–469.

CHAPTER 28

# The role of the family in supporting learning

Andrea Creech

## Introduction

Powerful images of musicians' parents have been immortalized in accounts of the lives of many iconic figures in Western music. Parents have been depicted as exerting enormous influence on their children's musical development, as in the cases of Mozart (Solomon 1994), Clara Schumann (Galloway 2002), and more recently, Yehudi Menuhin (Menuhin 1977) and Jacqueline du Pré (Easton 1989), to name but a few. At the other end of the spectrum are exceptional accounts of musicians who attained high levels of expertise and musical intelligence without parental support, as in the case of Louis Armstrong (Collier 1983). Kyle Pruett (2003, p. 154) invites us to conjure up our 'most enduring stereotype of the musician's parent. Then consider the opposite pole. Next, reflect on the intermediaries.' This chapter will attempt to respond to Pruett's challenge by considering these 'intermediaries' and suggesting the ways in which parents may most constructively support their children's musical development.

Years of educational research, theory and wisdom sustain the view that parents play a key role in their children's academic achievement, educational aspirations, well-being and motivation (Baker 1997; Grolnick *et al.* 2002; Gonzalez De Hass *et al.* 2005). Researchers have concluded that young children and adolescents alike experience improved student outcomes when their parents are actively involved at home and in the school context (Henderson and Berla 1994).

In the domain of music, and specifically relating to the acquisition of expertise on musical instruments, the question of how families can best support their children's interest has been a preoccupation of many researchers. Positive relationships between musical home environments and the musical responsiveness of children from these homes have been reported (Kirkpatrick 1962; Shelton 1965; Wermuth 1971), and these findings have been elucidated by more recent research that has found children's musical development to be influenced by parental musical background (Bloom and Sosniak 1981; Klinedinst 1991), socio-economic background (Klinedinst 1991), parent support for practice and lessons (Doan 1973; Brokaw 1982; Sloboda and Howe 1991; Zdzinski 1992; Davidson *et al.* 1995), parental goals, aspirations and values (Sosniak 1985; Addison 1990; Davidson and Scott 1999), parental self-efficacy (Creech 2001), family interaction patterns (Davidson and Borthwick 2002), and parent–teacher–pupil relationships (Manturzewska 1990; Creech 2006; Hallam 1998). It has been acknowledged that the conditions which facilitate the development of musical gifts and talents involve considerable investment of time and financial resources as well as parental belief and philosophical commitment (Chadwick 1996). The process of facilitating developing talents is seen by Feldhusen (2001) as a long-range one in which parents work together with pupils and teachers. This growing body of research suggests that the ways parents might support their children in persisting with learning musical instruments and developing musical expertise are diverse and complex.

The advent of the Japanese Suzuki method of violin teaching, which came to the West during

the 1960s holding as a central tenet the importance of the parent as 'home teacher', played a powerful role in highlighting the issue of parent participation in the realm of children's instrumental learning. The Suzuki method, however, cannot claim sole ownership of the concept of parent–teacher–pupil partnership in instrumental learning; it has been demonstrated through biographical evidence (Lochner 1950; Weschler-Vered 1986; Milstein and Volkov 1990; Lewis and Staryk 2000; Stern and Potok 1999; Menuhin 1977) that parents, across cultures and historical time frames, have chosen to involve themselves integrally in the process of their children's acquisition of musical expertise. Empirical studies into the role of parents in instrumental learning (Doan 1973; Brokaw 1982; Sosniak 1985; Sloboda and Howe 1992; Davidson *et al.* 1996; Creech 2001, 2006; McPherson and Renwick 2001; Welch *et al.* 2006) have indicated that parental involvement in their children's musical development is not unusual; parents from diverse backgrounds and with children learning by a range of teaching methods and representing a number of instruments and musical genres, have supported their children in pursuit of musical excellence.

Parent involvement has been defined as 'the dedication of resources by the parent to the child within a given domain' (Grolnick 1997, p. 538). Grolnick conceptualizes parent involvement as comprising behavioural support, cognitive/intellectual support, and personal support. This chapter will begin with a brief discussion of the ways in which these three types of parental support have been found to influence children's musical development, with an emphasis given to the latter category of personal support, an area of parental involvement that has been shown in the domain of music to have far-reaching consequences for pupils, parents and teachers alike (Pruett 2003; Creech 2006). A discussion of whether interpersonal relating styles determine the extent to which parents engage in various types of involvement will follow, paying particular attention to the findings of a recent investigation into the influence of interpersonal relationships on learning outcomes (Creech 2006). Based on these findings a typological approach will be taken, with parental support of their children's musical development discussed within the framework of six distinct types of parent–pupil–teacher partnerships in instrumental learning.

## Behavioural support

Overt manifestations of parental support, including participatory activities and modelling the importance of the subject area, are included under the umbrella of *behavioural support* (Grolnick 1994). Parent participation in school activities has been found to be a positive predictor of primary school-aged children's school engagement (Izzo *et al.* 1999), and helping with homework is one of a number of parent practices that has been found to have a positive relationship with secondary students' academic engagement (Steinberg *et al.* 1992).

In the domain of instrumental learning parents may offer behavioural support in the form of monitoring and participating in practice, attending lessons and adopting the role of home teacher. In an experimental study involving beginning band students (aged 11–12) Brokaw (1982) found that

> while it was not surprising to discover a strong relationship between the amount of time a student spends practising and the student's achievement in performance ... the amount of time spent by parents in supervising home practice is even a better predictor of successful achievement in the initial stages of development.
>
> (1982, p. 97)

Previously, Doan (1973) had produced similar results, in his study of 11-and 12-year-old violin students. Amongst a number of factors, parental supervision of practice and parent attendance at the child's concerts were identified by Doan as being significantly correlated to student achievement on the violin (1973, p. 79).

Other research has helped to elucidate Doan and Brokaw's findings. In a study of American concert pianists, Sosniak (1985) found that although many of the parents of her cohort had little musical background, their role of stimulating and supporting practice had been vital in sustaining their children's growth in musical competence throughout a first phase of early formal instruction and second phase of systematic

acquisition of knowledge and skills. Sloboda and Howe (1991) concurred with Sosniak when they found that high-achieving 8–18-year-old students in a specialist music school had benefited from the support and encouragement of parents who, with little formal knowledge of music, took responsibility for helping with home practice and for encouraging their children to gain and maintain good practice habits. Beginning band pupils, aged 7–9, were found by McPherson and Renwick (2001) to be helped by parents in developing self-regulated learning, particularly during the first year of learning. Davidson *et al.* (1995) demonstrated that parental commitment to assisting, encouraging and supporting the child in the early stages of learning was a more important predictor of successful musical outcomes than any specialist knowledge on the part of the parent. 'Without the positive involvement of the parent in the process, the highest levels of achievement are likely to remain unattainable' (1995, p. 44).

Zdzinski (1992), in a study involving teenage woodwind players, demonstrated that the effects of behavioural support upon musical achievement may differ with student age. This view was supported by the findings from a more recent study involving 337 violin pupils and their parents (Creech 2006) where a considerable drop in monitoring and assisting with practicing, providing feedback during practice sessions and attending instrumental lessons was found amongst the parents of 13–14-year-olds, as compared to those with children up to the age of 12. Furthermore, when parents were grouped according to levels of behavioural support it was evident that high levels of participatory activities during adolescent years was significantly associated with inhibited pupil–teacher accord and with limitations on the child's autonomous learning. These findings support the evidence reported by Ginsburg *et al.* (1993) who found decreased pupil autonomy amongst early adolescents whose parents operated a regime of homework surveillance.

The ages of 12–18 have been described as the mid-life crisis of young musicians, when the need to acquire or disown the interest in music becomes paramount (Bamberger 1987). This is the point at which behavioural support may become less helpful, while an increasing emphasis on cognitive/intellectual support and personal support may become more valuable to the developing musician.

Parents who have engaged in considerable behavioural support during the early years of their children learning musical instruments face a special challenge of managing the transition to pupil–parent independence, in relation to the instrumental learning. Creech (2006) reported that this transition to independent pupil learning and the shift of emphasis from behavioural support to cognitive/intellectual and personal support sometimes involved the difficult risk that their child would make the choice to discontinue, a choice that carried implications of rejection of parental values. In addition, for many parents this shift entailed the loss of a role they had constructed for themselves, and despite being supportive of their children becoming independent learners many reluctantly surrendered their participation in practice and lessons. Teenage children often took the lead, 'encouraging' parents to relinquish their home teacher role by making it clear that this type of parental input was not welcome. Teachers too sometimes took the initiative, asking parents to cease attending lessons in order to allow a more independent relationship to flourish between pupil and teacher. The most successful transitions were experienced by parents whose children had established strong and positive relationships with their teachers, remaining receptive to parental interest yet taking responsibility for learning without external motivation provided by parents.

## Cognitive/intellectual support

Exposing the child to cognitively stimulating activities and materials and engaging in intellectually domain-specific activities in the home have been found to comprise an important area of parental support (Grolnick 1994). Kulieke and Olszewski-Kubilius (1989) demonstrated that families of gifted children engaged in this form of support by espousing values relating to persistence and achievement in the subject area and by facilitating the progress of their children along particular domain-specific paths. Some evidence suggests that children internalize their parents' values and educational aspirations;

Marchant (2001) reported that the perceptions held by early adolescents of their parents' values had a strong association with motivation and attainment. Whether adolescents are likely to incorporate the parents' orientation and persevere on these paths has been found by Smith (1991) to be dependent on parent–adolescent communication regarding educational aims, together with perceived agreement between the two parents.

According to Csikszentmihalyi *et al.* (1993, p. 174) high levels of cognitive/intellectual support and challenge have a positive effect on teenagers across all talent areas. Parents of these accomplished children typically:

- Devote great amounts of time and energy to meeting the needs of their children
- Set high standards
- Encourage productive use of time
- Provide challenging opportunities
- Make sure lessons and materials are available
- Set aside areas of the home where child can work privately.

In the context of instrumental learning, parents offering this type of support provide the opportunities and materials that will assist the development of their children's musical intelligence, including arranging instrumental lessons, attending professional concerts with their children, listening to and discussing music in the home, encouraging participation in extra-curricular musical activities, and providing a quality instrument.

Evidence from a study investigating parental influence on the musical development of musically involved young Australians (Chadwick 1996) suggests that parents who perceived their children to be exhibiting characteristics of musical giftedness provided high levels of cognitive support for home-based musical activities. It has been suggested that children flourish when parents follow their lead and provide the wherewithal for in-depth exploration of their particular interests (Silverman 1992). Chadwick (2000) elucidates this view, suggesting that opportunities to engage in challenging and rigorous musical activities from an early age enhance normative musical development and provide a foundation for the development of musical expertise.

Creech (2006) reported that parents of violin pupils aged 13–18 who had persisted with learning for seven or more years and had attained relatively high levels of expertise on their instruments (beyond grade five, practising six or more hours per week) engaged in higher levels of music-specific cognitive/intellectual support than those with children below the age of 13 and in the beginning stages of playing. In the same study, when 337 parent–pupil pairs were grouped according to parent scores for cognitive/intellectual support, increased pupil enjoyment and parent satisfaction were found in the groups where the parents offered the greatest amount of this type of support. Cognitive/intellectual support was also found to be associated with increased professional satisfaction on the part of the teachers and with pupil–teacher influence, whereby pupils demonstrated independent learning, contributing to setting goals and effecting changes in the learning agenda.

## Personal support

A growing body of evidence demonstrates that high levels of personal support represents a key area of parent involvement associated with sustaining a child's musical well-being (Pruett 2003; Creech 2006).

Education researchers have found that children in all age groups, including secondary school students, value their parents' help, interest and support, and that parental influence on children's behaviour remains extensive in adolescence (Brown *et al.* 1993; Crozier 1999). However, Crozier emphasizes the need for pupils to have some control over parents' involvement, highlighting the importance of negotiation rather than imposition of psychological control characterized by intrusive or manipulative controlling/surveillance measures (Crozier 1999; Baumrind 2005). Crozier here touches on the issue, so stark in adolescence, of the delicate balance between dimensions of 'agency' (the drive for independence) and 'communion' (the need to be engaged with others), which has been identified both in the literature relating to parenting style (Baumrind 1989; Brown *et al.* 1993; Noack 1998) and that concerned with interpersonal style (Van Tartwijk *et al.* 1998) and relationships (Birtchnell 1993; Tubbs 1984; Noller *et al.* 2001).

Steinberg *et al.* (1992) demonstrated that teenage students were most likely to benefit from parental involvement activities that included attending school programmes and extra-curricular activities and maintaining an interest in student progress when these forms of parental support were offered within a parenting style characterized by acceptance, warmth, democracy, and some degree of student autonomy. While Ginsburg and Bronstein (1993) found that behavioural support in the form of surveillance and monitoring of homework was associated with an extrinsically motivated orientation on the part of their children, they also noted that when parents provided encouragement and praise the children were more likely to demonstrate curiosity and an intrinsic interest in challenging learning tasks. It has been suggested that emotional support offered by parents in the form of interest and enthusiasm for their children's learning establishes 'a foundation for socializing children's motivation to learn' (Gonzalez-DeHass *et al.* 2005, p. 111). Furthermore, there is strong evidence that regardless of social class parental involvement in the form of home interest, support and discussion has a major impact on educational outcomes (Sui-Chu and Willms 1996; McNeal 2001).

Researchers and theorists argue that 'people function most cohesively and confidently in contexts in which they experience significant others as being both caring and autonomy-supportive' (Noack 1998, p. 227) and achieving a successful balance in this respect is a key factor in providing personal support for children's musical development.

In a musical context, biographical evidence gathered from finalists in the German 'Jugend Musiziert' music competition demonstrated that these young expert musicians had typically benefited from supportive and motivating relationships with their mothers (Bastian 1989). More recently, personal support was found to be associated with persistence with learning and musical attainment as well as with increasing pupil enjoyment of music and motivation (Creech 2006). Those parents with the highest scores for personal support were interested in understanding their child's point of view and prepared to compromise when their personal goals or expectations conflicted with those of the child, subscribing to the stated aim of supporting the child in whatever course of action was chosen (i.e. to continue or not with musical studies). These parents were deeply interested in their children's learning, and were involved with their child learning an instrument to the extent that they considered this endeavour to have been a life-changing experience for themselves. A predominant view amongst these parents was that their children should develop independent learning skills and yet continue to feel supported by their parents. Strong pupil–teacher rapport was valued and this together with admiration of the teacher appeared to make it relatively easy for parents to become progressively less directly involved in the learning partnership (Creech 2006).

## The potential for conflict

Considering the potential for conflict amongst those involved with musical instrument learning, particularly over practice, musical preferences, and time commitment expectations, it is perhaps rather surprising that little research to date has been directly concerned with this issue.

The emotional demands made on parents by their musical children can be considerable. The parent–child relationship is particularly vulnerable when adolescents reach the aforementioned musical mid-life crisis (Bamberger 1987) and as young musicians become increasingly susceptible to performance anxiety and the fear of negative judgement (Robson 1987). Clearly it is important for children to sustain a sense of being emotionally supported by their parents even in the face of disagreements. Research with gifted children and their families advocates family meetings as a forum where conflicts and power struggles may be resolved via shared responsibility and negotiation (Silverman 1992).

While parents of young musicians have been found to make frequent references to reluctant practising and to parent–pupil conflict these issues did not, as many parents fear, inevitably diminish either their children's persistence with learning or their own personal satisfaction (Creech 2006). Strategies that parents have identified as being successful in dealing with this potential battleground include simply leaving the room, discussions at times other than practice

time, lowering expectations of daily practice, challenging the child, and offering praise as a reward. Parents did not acknowledge that material rewards would be a useful strategy to resolve conflict over practice, and considered the most successful strategy to have been allowing the child to choose when and how much to practise, within parameters negotiated with the teacher (Creech 2006).

Pruett identifies transcultural qualities of 'good enough parenting' (Pruett 2003, p. 155), amongst which the challenges of personal support are encapsulated. Included on his list are sensitivity to children and their ever-changing needs, the ability to make children feel loved, adored and enjoyed, devotion to sustaining strong values, affirmation of the child's uniqueness while expecting competence, and sustaining an abiding presence through thick and thin. Research in the domain of music supports the enduring importance of these qualities, suggesting that pupils function best when they perceive the adults as both caring and supportive of autonomy and when they are able to engage in ongoing mutual interaction with adults who continue to have a stake in their development and to act as their advocate (Noack 1998; Manturzewska 1990; Creech 2006).

## Parent–pupil–teacher interaction: a typological approach

Any discussion of the parental role in children's musical development must take into account the interpersonal context. Systems theory provides a framework in which this context, comprising parents, teachers and pupils, may be conceptualized as a micro-system characterized by *holism*, whereby a human system is understood as the sum of its individual members, and *exchange*, whereby any change in one member of the system produces change throughout the system (Becvar and Becvar 1996). O'Neill (1996) argues that the parent–pupil–teacher relationship in the context of musical instrument learning can justifiably be conceptualized as a micro-system because all three participants experience new patterns of action and communication as a direct result of their interaction, and because many motivational issues can be understood and possibly resolved when considered as a function of the micro-system. From this perspective parental support may be seen as part of the functioning of the parent–teacher–pupil system, influenced by the parent's position on a *control–responsiveness* axis, in relation to teacher and pupil.

### Control and responsiveness

Baumrind (2005) suggests that the extent to which parents engage in supportive behaviour is associated with interpersonal qualities she labels as responsiveness and demandingness (interpreted here as *control*). A study by Marchant (2001) involving 230 early adolescents measured parental style using these two dimensions and provided evidence that children's school achievement is indeed influenced by interpersonal relationships amongst teachers, parents and pupils. Trusty and Lampe (1997) investigated the dimension of parent control and suggest that this interpersonal dynamic, when manifested in the context of relationships where parents offer cognitive and personal support in the form of discussing school, current events or troubling issues with their teenagers, is associated with an internal locus of control on the part of the young people whereby they experience pride and satisfaction for their successes. However, Ginsburg and Bronstein (1993) investigated the impact of behavioural support on student motivation and found that when parents monitored, enforced or helped with homework in a controlling fashion their early adolescent children showed less autonomy, satisfaction, persistence and intrinsic motivation related to their schoolwork. Garland (2005) discuss control as a dimension that varies from controlling to autonomy-supportive behaviour, and puts forth the view that while the former encourages children's obedience and compliance the latter promotes choice, self-reliance and participation in decision-making.

Baumrind's model of demandingness—responsiveness interpersonal relating style is reflected in Birtchnell's relating theory (2001) whereby interaction is conceptualized on a horizontal closeness–distance axis intersecting with a vertical upperness–lowerness axis. Birtchnell does not

priviledge different positions on his interpersonal model, pointing out that while closesness holds people together, distance provides the space to become autonomous, and while upperness allows the opportunity for people to exert influence on others, lowerness enables individuals to benefit from the care and leadership of others.

## Categories of parent–teacher–pupil interaction

The models proposed by Baumrind and Birtchnell resemble Leary's model for interpersonal interaction (1957) that served as the basis for the examination of parent control and responsiveness in the context of their relationships with their children who were learning the violin and with their children's instrumental teachers (Creech 2006). Parent control, on the vertical axis, was found to comprise underlying dimensions that were (1) perceived teacher leadership, (2) parent–teacher communication, (3) parent isolation within the learning partnership, (4) parent ambition and (5) parent preponderance. Responsiveness, on the horizontal axis, comprised (1) perceived teacher approachability, (2) teacher–parent intimidation, (3) parent–pupil reciprocity and (4) parent–teacher acquiescence.

Six distinct clusters of parent–teacher–pupil interaction types, determined by dimensions of control and responsiveness, were found amongst the parent–teacher–pupil trios. A model (Figure 28.1) representing these six types of learning partnerships demonstrates that clusters one, two and three may be conceptualized as primary dyad plus a third party, while cluster four is represented as two primary dyads connected by one common member. Cluster five is characterized by very little communication between any two of the three individuals, while cluster six is characterized by reciprocity amongst all three participants.

## Differences amongst the clusters of parent–teacher–pupil interaction types

### Parent support

Significant differences were found between the parents of each cluster with respect to the types

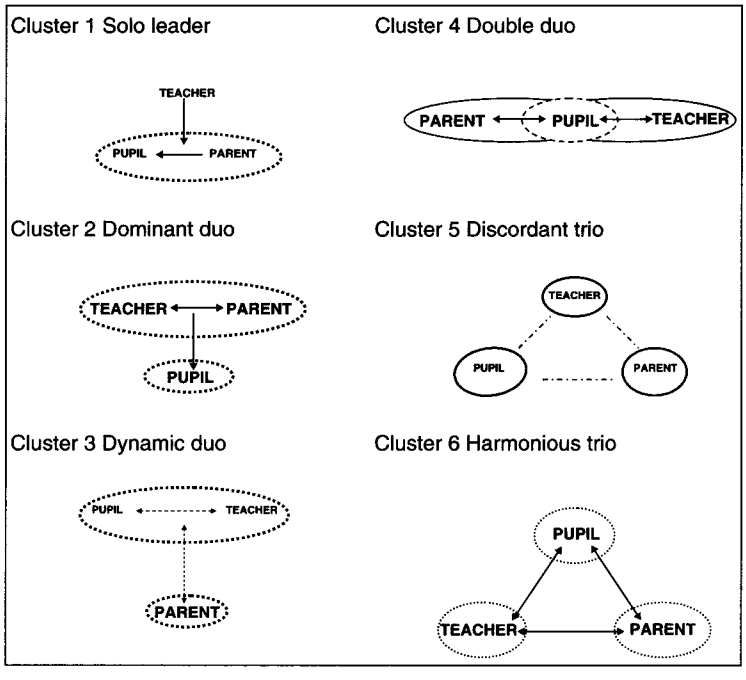



**Fig. 28.1** Parent–teacher–pupil interaction types.

of support they were most likely to engage in. Parents in the *dynamic duo* and *discordant trio* clusters, characterized by a distant and powerless parent–teacher relation, offered the least amount of support, overall. Those occupying the *solo leader* cluster, where the parent took responsibility for ensuring that parent and pupil together followed the directive teacher, were found to offer the most behavioural support. The highest levels of cognitive/intellectual support were offered by parents occupying the *dominant duo* cluster, characterized by parents who adopted a predominant and controlling role in relation to both pupil and teacher. The *harmonious trio* cluster, where parents balanced ambitions for their children with responsiveness in relation to both teacher and pupil, had parents who engaged in the most personal support (Figure 28.2).

### Pupil age

As pupils matured parent–pupil–teacher trios tended to gravitate towards the *dynamic duo* (pupil–teacher relationship becoming the primary dyad and parents becoming less influential), adding to the body of evidence showing that while parent–child relationships are typically hierarchical, family relationships undergo transformations towards more egalitarian patterns during adolescence (Noack 1998). While younger pupils were more likely to occupy the categories of *solo leader* and *dominant duo* where the basis for teacher–parent cooperation was the perceived dependency of the child, during the adolescent years pupils ceased to be dependent on the parent for motivation and structure (Johnson 1991). It was evident that some parents became isolated third parties (*dynamic duo*) while others shifted from an emphasis on behavioural

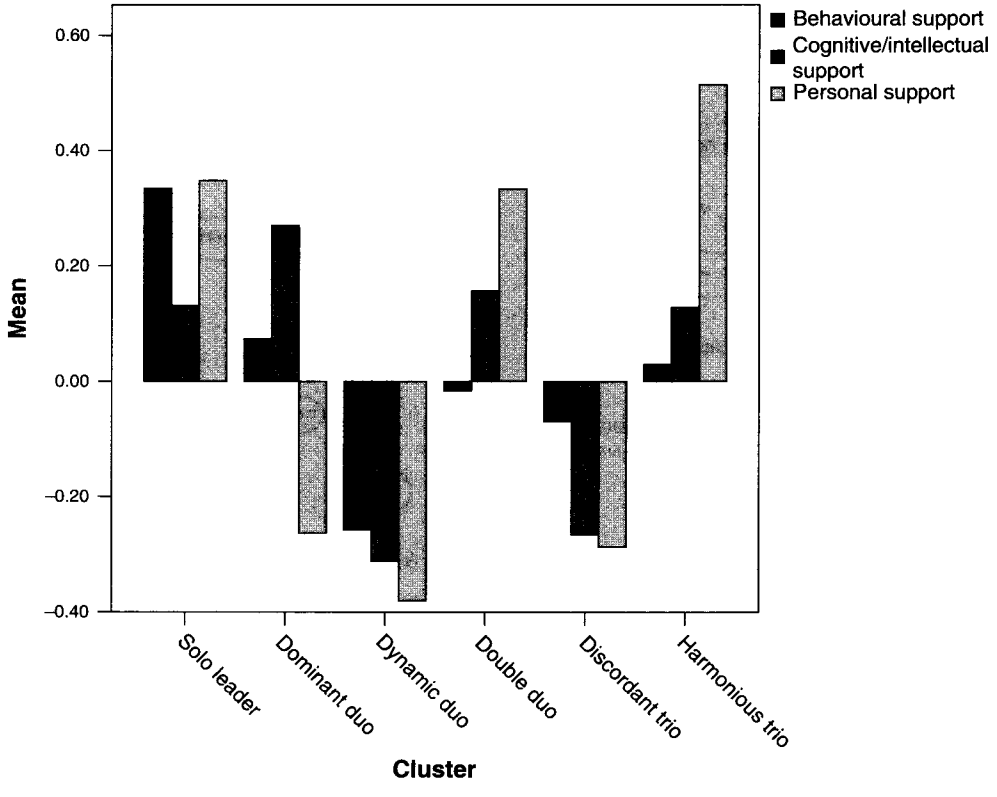



**Fig. 28.2** Parent support according to parent–teacher–pupil interaction type (standardized scores).

of support they were most likely to engage in. Parents in the *dynamic duo* and *discordant trio* clusters, characterized by a distant and powerless parent–teacher relation, offered the least amount of support, overall. Those occupying the *solo leader* cluster, where the parent took responsibility for ensuring that parent and pupil together followed the directive teacher, were found to offer the most behavioural support. The highest levels of cognitive/intellectual support were offered by parents occupying the *dominant duo* cluster, characterized by parents who adopted a predominant and controlling role in relation to both pupil and teacher. The *harmonious trio* cluster, where parents balanced ambitions for their children with responsiveness in relation to both teacher and pupil, had parents who engaged in the most personal support (Figure 28.2).

### Pupil age

As pupils matured parent–pupil–teacher trios tended to gravitate towards the *dynamic duo* (pupil–teacher relationship becoming the primary dyad and parents becoming less influential), adding to the body of evidence showing that while parent–child relationships are typically hierarchical, family relationships undergo transformations towards more egalitarian patterns during adolescence (Noack 1998). While younger pupils were more likely to occupy the categories of *solo leader* and *dominant duo* where the basis for teacher–parent cooperation was the perceived dependency of the child, during the adolescent years pupils ceased to be dependent on the parent for motivation and structure (Johnson 1991). It was evident that some parents became isolated third parties (*dynamic duo*) while others shifted from an emphasis on behavioural

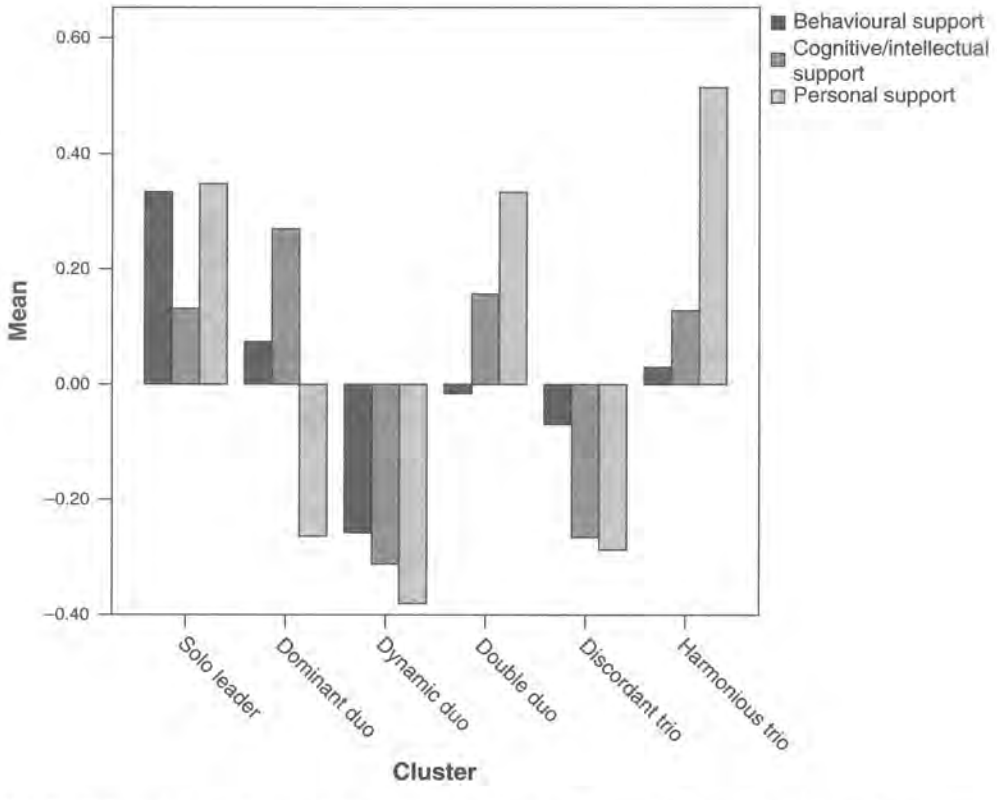



**Fig. 28.2** Parent support according to parent–teacher–pupil interaction type (standardized scores).

support to personal and cognitive/intellectual support, continuing to support their children at home but becoming remote from the teacher (*double duo*). Still others (likely to be found in the *harmonious trio*) achieved a balance of agency and communion with their children and teachers within a relationship characterized by continued moderate behavioural support together with an emphasis on both cognitive/intellectual and personal support.

### Persistence with learning

Numbers occupying the *discordant trio* category decreased sharply after three years of study, suggesting that characteristics of this type may have led to pupil drop-out, or alternatively to changes to other teachers. Conversely, there was a large increase in numbers occupying *harmonious trios*, after three years of study. The numbers in the *dynamic duo* category remained stable and were not sensitive to years of study, while numbers occupying *dominant duo* and *double duo* categories declined after six years of study. Thus it would seem that the balance of parent support achieved by those in *harmonious trios*, which included moderate behavioural support, higher levels of cognitive/intellectual support and a great deal of personal support, may be associated with their children's perseverance on the instrument.

Few pupils occupying the *discordant trio* cluster progressed beyond grade five, while the *dominant duo* and *harmonious trio* clusters contained greater proportions of pupils who had attained a minimum of grade four. Parents occupying these two clusters, while differing in respect of personal support, were similar in that they demonstrated high parent–teacher acquiescence and perceptions of strong teacher leadership but low teacher–parent intimidation and isolation.

The *harmonious trio* and the *solo leader* were the clusters where the greatest proportion of pupils did in excess of three hours practice per week. Parents in these two clusters were most similar in that they perceived strong leadership on the part of the teachers and also offered a great deal of personal support to their children. Again, parental perception of teacher leadership seems to be a key factor in sustaining persistence with learning.

### Pupil learning outcomes

Differences amongst the clusters of interaction types were evident in respect of pupil learning outcomes including enjoyment of music, personal satisfaction, motivation, self-efficacy and self-esteem. Overall the *discordant trio* produced the least positive outcomes while the *harmonious trio* produced the most consistently positive outcomes for pupils. With reference to the former this finding adds to earlier research that suggests where parents place a low value on the subject matter, have low expectations of success, do not have the wherewithal to help their children at home and/or are intimidated by teachers the result can be a downward spiral of mutual distrust, lack of communication and absence of shared purpose amongst parents, teachers and pupils alike (Bandura 1997; Hurley 1995). In contrast, the findings in respect of the latter elucidate earlier research that has proposed a model of parent–professional–child partnership whereby parents lie at the heart of a system which advances the child's development while professionals take primary responsibility for advancement of knowledge and skills (Henry 1996). Furthermore, these findings support the view that pupils function best when they perceive the adults as both caring and supportive of autonomy and when they are able to engage in ongoing mutual interaction with adults who continue to have a stake in their development and to act as their advocate (Noack 1998; Manturzewska 1990).

### Summary: the versatile parent

Effective and supportive parental involvement in instrumental learning requires parents to be versatile, adept at moving between the close and distant positions on the responsiveness axis and between directive and acquiescent positions on the control axis of the model for interpersonal dimensions. This involves providing much practical assistance and personal support during the early years of learning yet seeking and following the teacher's advice in musical matters and allowing the child and teacher the space to develop an autonomous relationship. It also involves remaining resilient in the face of reluctant practising while remaining as the child's interested and supportive advocate, long after

practical help has ceased to be appropriate or welcomed by the teacher and pupil. Most importantly, parents should not become uninvolved in their children's learning in the name of agency, nor disempower their children in the name of communion. Specifically, positive outcomes may be achieved when parents (a) elicit their children's views regarding appropriate parental involvement, (b) negotiate with their children over practising issues, within parameters set by the teacher, (c) provide a structured home environment for practice, (d) take an interest in promoting good teacher-pupil rapport, (e) communicate with the teacher in relation to the child's progress, and (f) remain as a supremely interested audience.

Returning to the systems framework, the versatile parent may be conceptualized as one part of a dynamic interpersonal system, influenced by and having a reciprocal influence on the characteristics of both teacher and pupil. There is much scope for future research in relation to this point, particularly concerning (a) how parent–teacher–pupil systems evolve over time and (b) the processes by which these interpersonal systems may underpin teaching and learning outcomes.

## References


Addison R (1990). parents' views on their children's musical education in the primary school: a survey. *British Journal of Music Education*, **7**, 133–141.

Baker A (1997). Parent involvement in children's education: a critical assessment of the knowledge base. *Annual Meeting of the American Education Research Association*, Eric Document Reproduction Service, New York.

Bamberger J (1987). The mind behind the musical ear. Presented at the Biology of Music Making Conference, Denver, CO.

Bandura A (1997). *Self-efficacy: the exercise of control*. W.H. Freeman and Company, New York.

Bastian HG (1989). *Leben fur musik. Eine Biographie-Studie uber musikalisch (Hoch-) Begabungen*. Schott, Mainz.

Baumrind D (1989). Rearing competent children. In W Damon, ed., *Child development today and tomorrow*, 349–378. Jossey-Bass, London.

Baumrind D (2005). Patterns of parental authority and adolescent autonomy. *New Directions for Child and Adolescent Development*, **108**, 61–69.

Becvar D and Becvar R (1996). *Family therapy: a systemic integration*. Allyn and Bacon, Boston, MA.

Birtchnell J (1993). *How humans relate: a new interpersonal theory*. Psychology Press, Hove, East Sussex.

Birtchnell J (2001). Relating therapy with individuals, couples and families. *Journal of Family Therapy*, **23**, 63–84.

Bloom B and Sosniak L (1981). Talent development. *Educational Leadership*, November, 86–94.

Brokaw JP (1982). The extent to which parental supervision and other selected factors are related to achievement of musical and technical-physical characteristics by beginning instrumental music students. Unpublished PHD Dissertation; Music, University of Michigan, Michigan, USA.

Brown BB, Mounts N, Lamborn SD and Steinberg L (1993). Parenting practices and peer group affiliation in adolescence. *Child Development*, **64**, 467–482.

Chadwick F (1996). *Gifted education: proceedings from the 1996 national conference in Adelaide, South Australia*. Australian Association for the Education of the Gifted and Talented, Adelaide.

Chadwick F (2000). An Australian perspective on talent development in music: the influence of environmental catalysts upon the provision of opportunities for learning, training, and practice in the musical domain. Unpublished PhD thesis; University of New South Wales, Sydney.

Collier JL (1983). *Louis Armstrong: an American genius*. Oxford University Press, Oxford.

Creech A (2001). Play for me: An exploration into motivations, issues and outcomes related to parental involvement in their children's violin study. Unpublished MA dissertation; Music, University of Sheffield, Sheffield.

Creech A (2006). Dynamics, harmony and discord: a systems analysis of teacher-pupil-parent interaction in instrumental learning. Unpublished PhD thesis; Institute of Education, University of London, London.

Crozier G (1999). Parent involvement: who wants it? *International Studies in Sociology of Education*, **9**, 111–130.

Csikszentmihalyi M, Rathunde K and Whalen S (1993). *Talented teenagers: the roots of success and failure*. Cambridge University Press, Cambridge.

Davidson J and Borthwick SJ (2002). Family dynamics and family scripts: a case study of musical development. *Psychology of Music*, **30**, 121–136.

Davidson J, Howe M, Moore D and Sloboda J (1996). The role of parental influences in the development of musical performance. *British Journal of Developmental Psychology*, **14**, 399–412.

Davidson J, Howe M and Sloboda J (1995). the role of parents and teachers in the success and failure of instrumental learners. *Bulletin of the Council for Research in Music Education*, **127**, 40–44.

Davidson J and Scott S (1999). Instrumental learning with exams in mind: a case study investigating teacher, student and parent interactions before, during and after a music examination. *British Journal of Music Education*, **16**, 79–95.

Doan G (1973). An investigation of the relationships between parental involvement and the performance

ability of violin students. Unpublished PhD Dissertation; Music, Ohio State University, Ohio, USA.
Easton C (1989). *Jacqueline du Pré: a biography*. Da Capo Press, Cambridge, MA.
Feldhusen JF (2001). *Talent development in gifted education*. ERIC Clearinghouse on Disabilities and Gifted Education, Arlington, VA.
Galloway J (2002). *Clara*. Jonathan Cape, London.
Garland S (2005). Perceived threat, controlling parenting, and children's achievement orientations. *Motivation and Emotion*, **29**, 103–120.
Ginsburg GS and Bronstein P (1993). Family factors related to children's intrinsic/extrinsic motivational orientation and academic performance. *Child Development*, **64**, 1461–1474.
Gonzalez-DeHass AR, Willems PP and Doan Holbein MF (2005). Examining the relationship between parental involvement and student motivation. *Educational Psychology Review*, **17**, 99–123.
Grolnick W (1994). Parents' involvement in children's schooling: a multidimensional conceptualization and motivational model. *Child Development*, **65**, 237–252.
Grolnick W (1997). Predictors of parent involvement in children's schooling. *Journal of Educational Psychology*, **89**, 538–548.
Grolnick W, Gurland S, DeCourcey W and Jacob K (2002). Antecedents and consequences of mothers autonomy support: an experimental investigation. *Developmental Psychology*, **38**, 143–155.
Hallam S (1998). The predictors of achievement and dropout in instrumental tuition. *Psychology of Music*, **26**, 116–132.
Henderson AT and Berla N (1994). *A new generation of evidence: the family is critical to student achievement*. National Committee for Citizens in Education, Columbia, MD.
Henry M (1996). *Young children, parents, and professionals: enhancing the links in early childhood*. Routledge, London.
Hurley CG (1995). Student motivations for beginning and continuing/discontinuing string music instruction. *The Quarterly Journal of Music Teaching and Learning*, VI, 44–55.
Izzo CV, Weissberg RP, Kasprow WJ and Fendrich M (1999). A longitudinal assessment of teacher perceptions of parent involvement in children's education and school performance. *American Journal of Community Psychology*, **27**, 817–839.
Johnson D (1991). Parents, students and teachers: a three-way relationship. *International Journal of Educational Research*, **15**, 171–181.
Kirkpatrick W (1962). Relationships between the singing ability of pre-kindergarten children and their home environment. Unpublished PhD thesis; University of Southern California, Los Angeles.
Klinedinst R (1991). Predicting performance achievement and retention of fifth-grade instrumental students. *Journal of Research in Music Education*, **39**, 225–238.
Kulieke MJ and Olszewski-Kubilius P (1989). The influence of family values and climate on the development of talent. In JL VanTassel-Baska and P Olszewski-Kubilius, eds, *Patterns of influence on gifted learners*, 40–59. Teachers College Press, New York.
Leary T (1957). *Interpersonal diagnosis of personality: a functional theory and methodology for personality evaluation*. Ronald Press Company, New York.
Lewis T and Staryk S (2000). *Fiddling with life: the unusual journey of Steven Staryk*. Mosaic Press, Canada.
Lochner LP (1950). *Fritz Kreisler*. Paganiniana Publications Inc., Neptune City, NJ.
Manturzewska M (1990). A biographical study of the life-span development of professional musicians. *Psychology of Music*, **18**, 112–139.
Marchant GJ (2001). Relations of middle school students' perceptions of family and school contexts with academic achievement. *Psychol Schools*, **38**, 505–519.
McNeal RB (2001). Differential effects of parental involvement on cognitive and behavioural outcomes by socioeconomic status. *Journal of Socio-Economics*, **30**, 171–179.
McPherson G and Renwick J (2001). A longitudinal study of self-regulation in children's musical practice. *Music Education Research*, **3**, 169–186.
Menuhin Y (1977). *Unfinished journey*. Macdonald and Jane's Publishers Ltd, London.
Milstein N and Volkov S (1990). *From Russia to the West*. Henry Holt and Co., New York.
Noack P (1998). School achievement and adolescents' interactions with their fathers, mothers, and friends. *European Journal of Psychology of Education*, XIII, 503–513.
Noller P, Feeney JA and Peterson C (2001). *Personal relationships across the lifespan*. Psychology Press, Hove, East Sussex.
O'Neill S (1996). Factors influencing children's motivation and achievement during the first year of instrumental music tuition. Unpublished PhD thesis; Keele University, Keele, UK.
Pruett K (2003). First patrons: parenting the musician. Presented at the 21st Annual Symposium on the Medical Problems of Musicians and Dancers, Aspen, Colorado.
Robson B (1987). Post-performance depression in arts students. *Medical Problems of Performing Artists*, **2**, 137–141.
Shelton J (1965). The influence of home musical environment upon musical response of first grade children. Unpublished PhD dissertation; George Peabody College for Teachers, Vanderbilt University, Nashville TN.
Silverman L (1992). *How parents can support gifted children*. Council For Exceptional Children, Reston, VA.
Sloboda J and Howe M (1991). Biographical precursors of musical excellence: an interview study. *Psychology of Music*, **19**, 3–21.
Sloboda J and Howe M (1992). Transitions in the early musical careers of able young musicians: choosing instruments and teachers. *Journal of Research in Music Education*, **40**, 283–294.

Smith TE (1991). Agreement of adolescent educational expectations with perceived maternal and paternal educational goals. *Youth and Society*, **23**, 155–174.

Solomon M (1994). *Mozart: a life*. Hutchinson, London.

Sosniak LA (1985). Learning to be a concert pianist. In BS Bloom, ed., *Developing talent in young people*, 19–67. Ballantine, New York.

Steinberg L, Lamborn SD and Darling N (1992). Impact of parenting practices on adolescent achievement: authoritative parenting, school involvement, and encouragement to succeed. *Child Development*, **63**, 1266–1281.

Stern I and Potok C (1999). *My first 79 years*. Alfred A. Knopf, New York.

Sui-Chu EH and Willms JD (1996). Effects of parental involvement on eighth-grade achievement. *Sociology of Education*, **69**, 126–141.

Trusty J and Lampe RE (1997). Relationship of high-school seniors' perceptions of parental involvement and control to seniors' locus of control. *Journal of Counselling Development*, **75**, 375–384.

Tubbs SL (1984). *A systems approach to small group interaction*. Random House, New York.

Van Tartwijk J, Brekelmans M and Wubbels T (1998). Students' perceptions of teacher interpersonal style: the front of the classroom as the teacher's stage. *Teaching and Teacher Education*, **14**, 607–617.

Welch G, Duffy C, Potter J and Whyton T (2006). *Investigating musical performance (IMP): comparative studies in advanced musical learning*. Institute of Education, University of London, Funded by the ESRC/TLRP, grant reference RES-139–25–0258.

Wermuth R (1971). Relationship of musical aptitude to family and student activity in music, student interest in music, socioeconomic status, and intelligence among caucasian and negro middle school students. Unpublished PhD Dissertation; Ohio State University, Ohio, USA.

Weschler-Vered A (1986). *Jascha Heifetz*. Schirmer Books, New York.

Zdzinski S (1992). Relationships among parental involvement, music aptitude, and musical achievement of instrumental music students. *Journal of Research in Music Education*, **40**, 114–125.

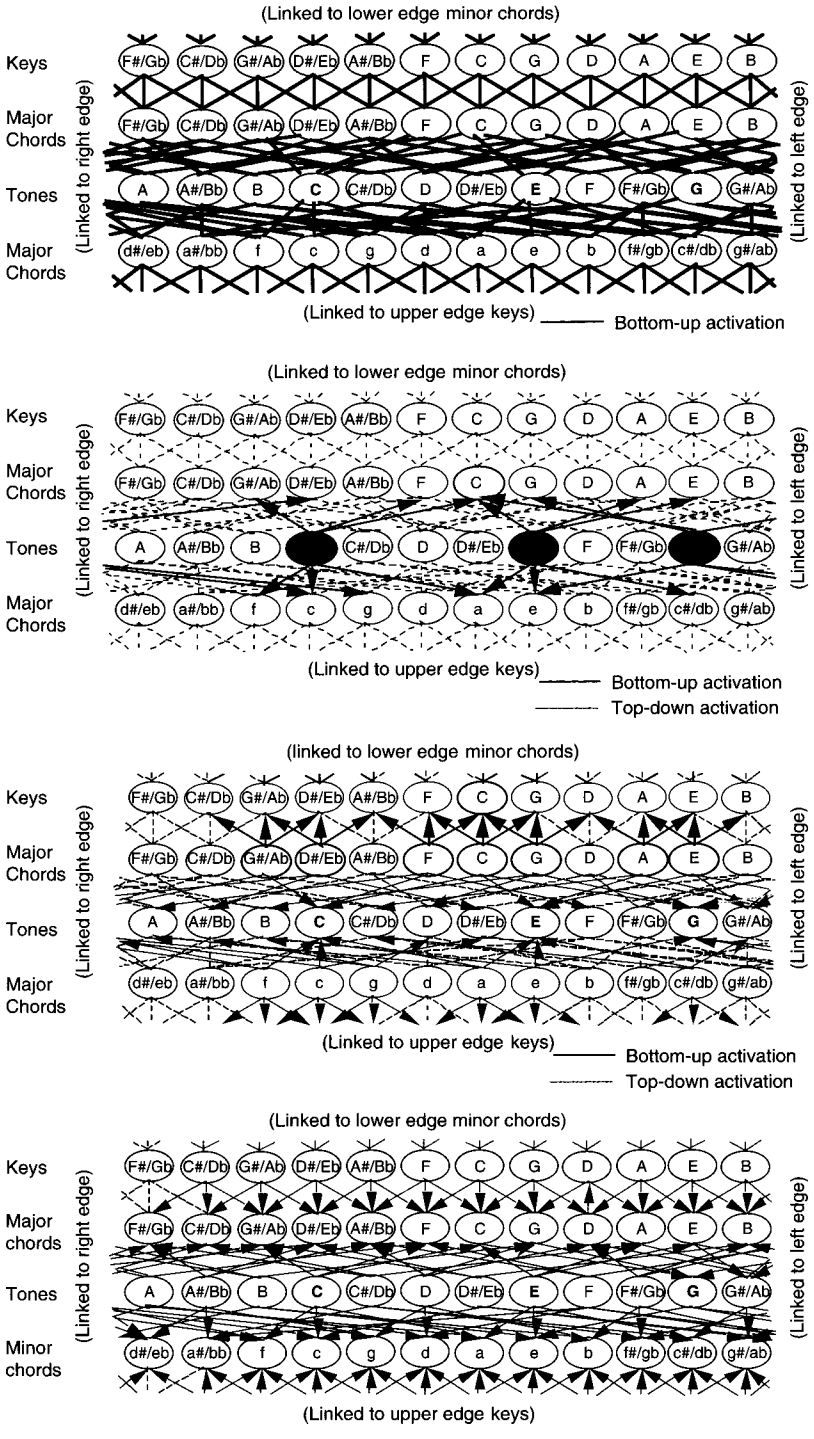



**Plate 1** Bharucha's MUSACT model. (a) When three tones are sounding, activation spread from tones to chord units (b), and then from chord to key and tone units (c), and from key units to chord and from tone to chord units (d) and so on up to equilibrium. See Chapter 6, p 66, for further details.

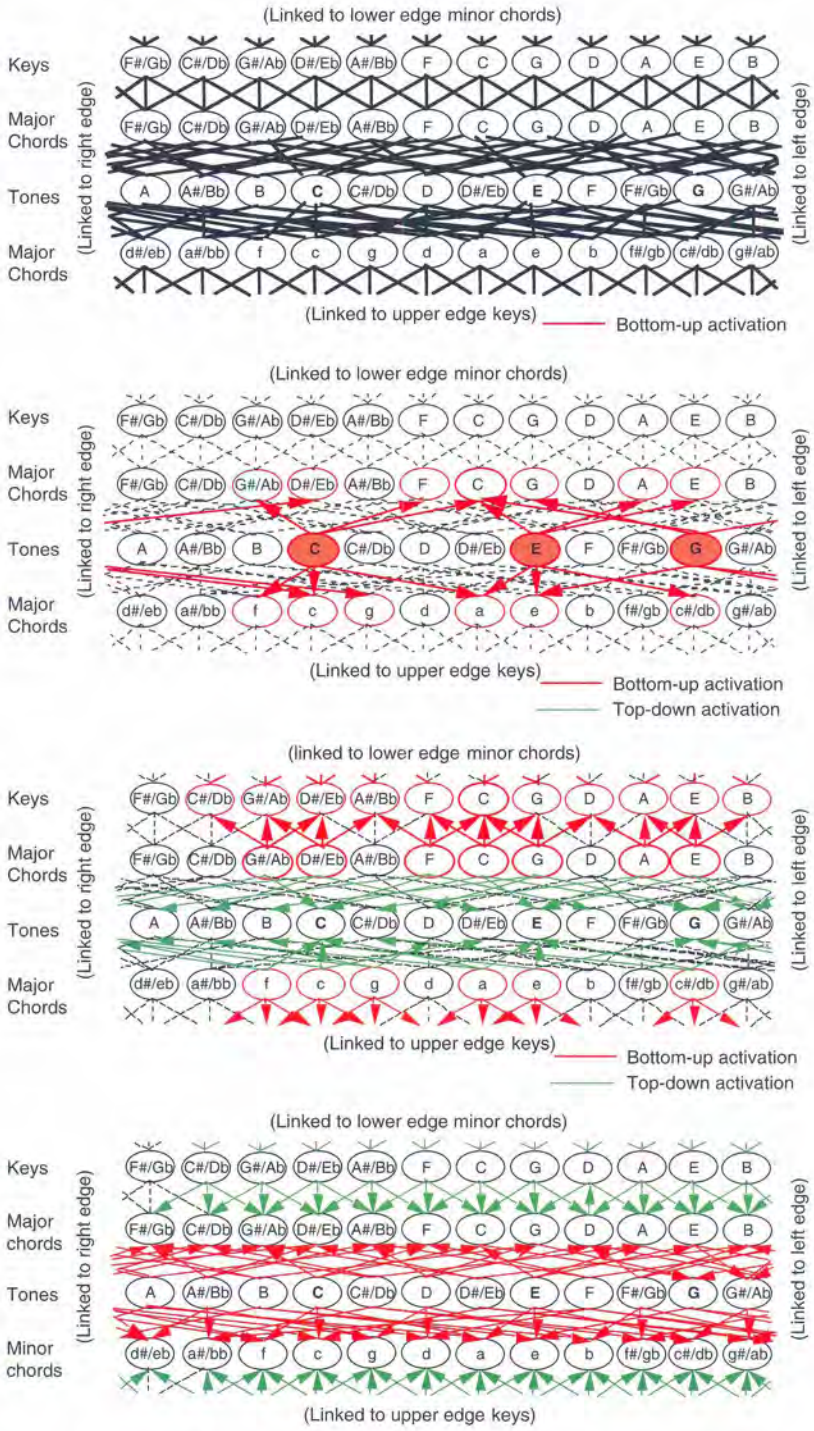



**Plate 1** Bharucha's MUSACT model. (a) When three tones are sounding, activation spread from tones to chord units (b), and then from chord to key and tone units (c), and from key units to chord and from tone to chord units (d) and so on up to equilibrium. See Chapter 6, p 66, for further details.

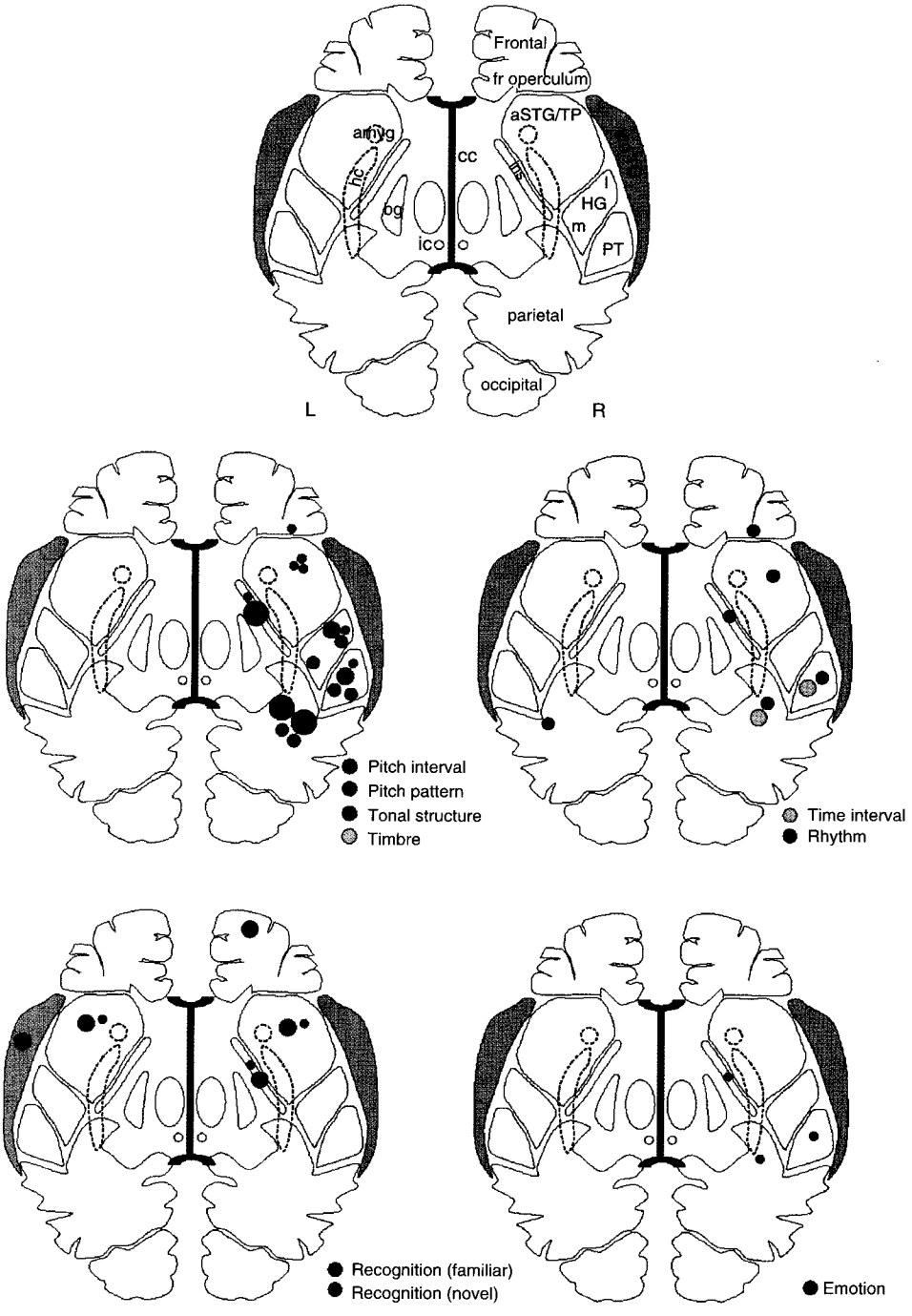



**Plate 2** Critical brain substrates for musical listening disorders. Schematic cartoons based on symptomatic cases of musical listening disorder. For each musical listening function, the presence of a circle indicates that at least 50% of studies reviewed involved damage to the corresponding brain region. The size of the circle scales with the proportion of studies in which such a link was evident. See Chapter 17, p 186, for further details. Key: amyg, amygdala; aSTG, anterior superior temporal gyrus; bg, basal ganglia; cc, corpus callosum; fr, frontal; hc hippocampus; HG, Heschl's gyrus; i, inferior; ic, inferior colliculi; ins, insula; l, lateral; m, medial; PT, planum temporale; TG, temporal gyrus

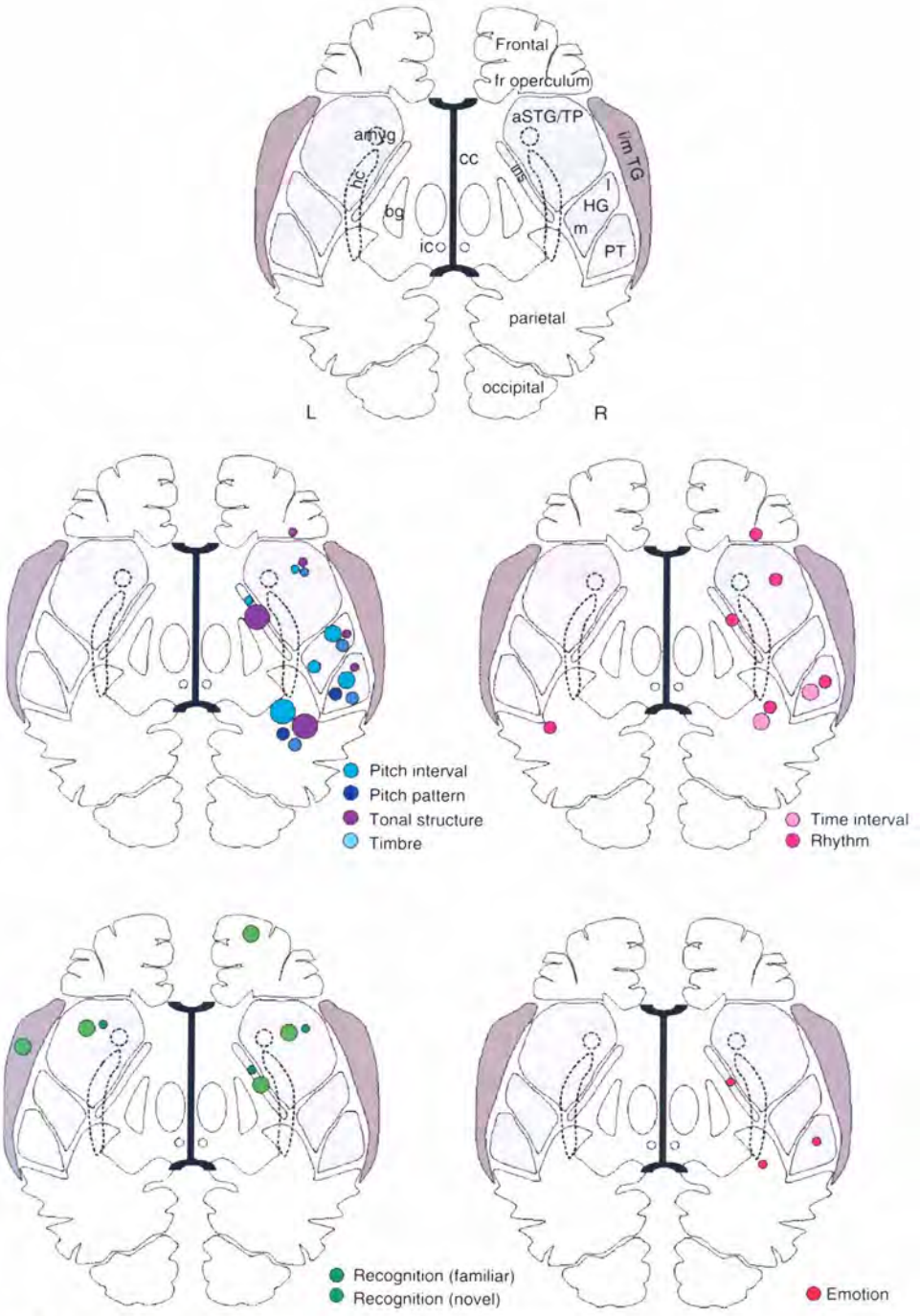



**Plate 2** Critical brain substrates for musical listening disorders. Schematic cartoons based on symptomatic cases of musical listening disorder. For each musical listening function, the presence of a circle indicates that at least 50% of studies reviewed involved damage to the corresponding brain region. The size of the circle scales with the proportion of studies in which such a link was evident. See Chapter 17, p 186, for further details. Key: amyg, amygdala; aSTG, anterior superior temporal gyrus; bg, basal ganglia; cc, corpus callosum; fr, frontal; hc hippocampus; HG, Heschl's gyrus; i, inferior; ic, inferior colliculi; ins, insula; l, lateral; m, medial; PT, planum temporale; TG, temporal gyrus

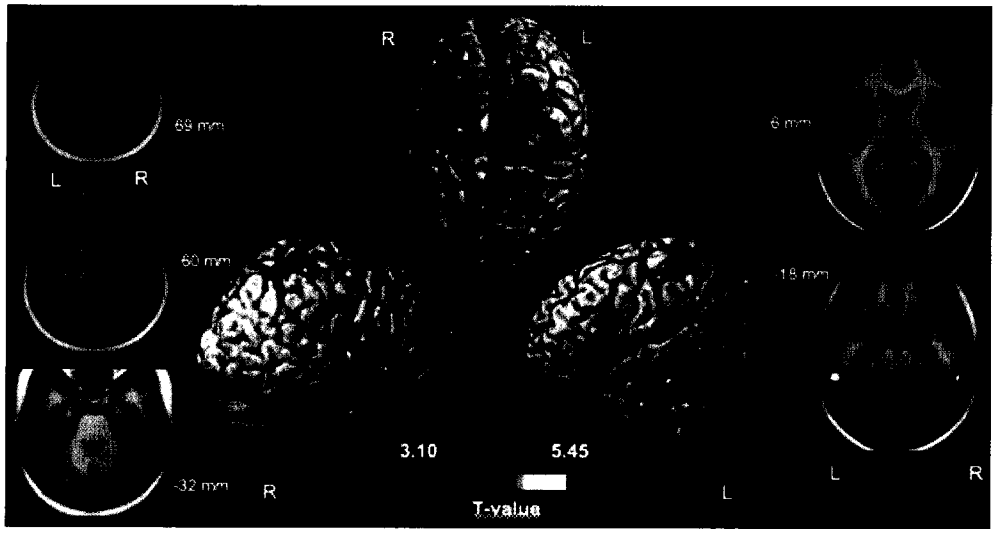



**Plate 3** A voxel-based morphometric study showing grey matter volume differences between three groups. Only voxels exhibiting a significant positive correlation between musician status (professional musician > amateur musician > non-musician) and increased grey matter volume are shown here, and those have been overlaid on the rendered cortical surface of a single subject. (This figure was adapted from Figure 1 in Gaser and Schlaug [2003]). See Chapter 18, p 199, for further details.

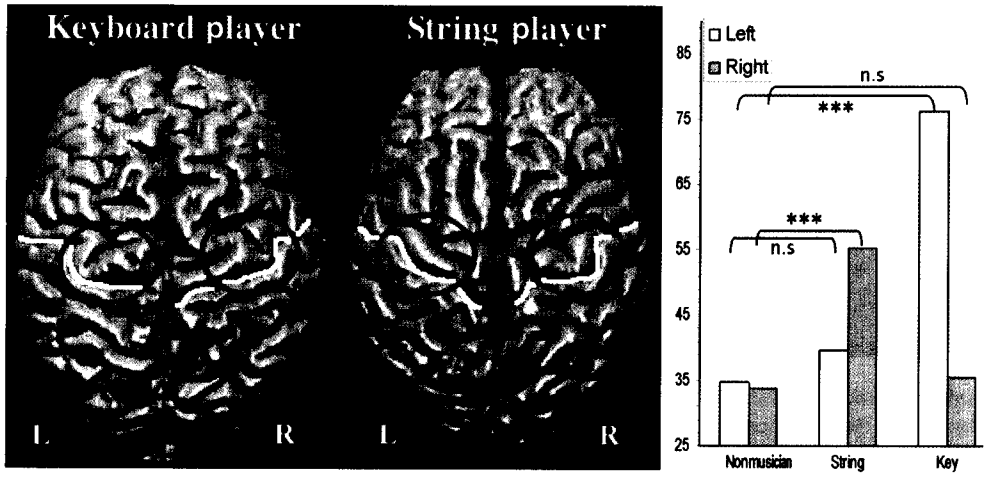



**Plate 4** Brain surface renderings of one typical keyboard and one typical string player. The central sulcus is marked with a white line. The portion of the precentral gyrus containing the configuration similar to that of the inverted Greek letter 'Omega' is found within the red circles. In the two examples, a prominent omega sign can be seen on the left more than the right in the keyboard player, and only on the right in the string player. See Chapter 18, p 200, for further details.

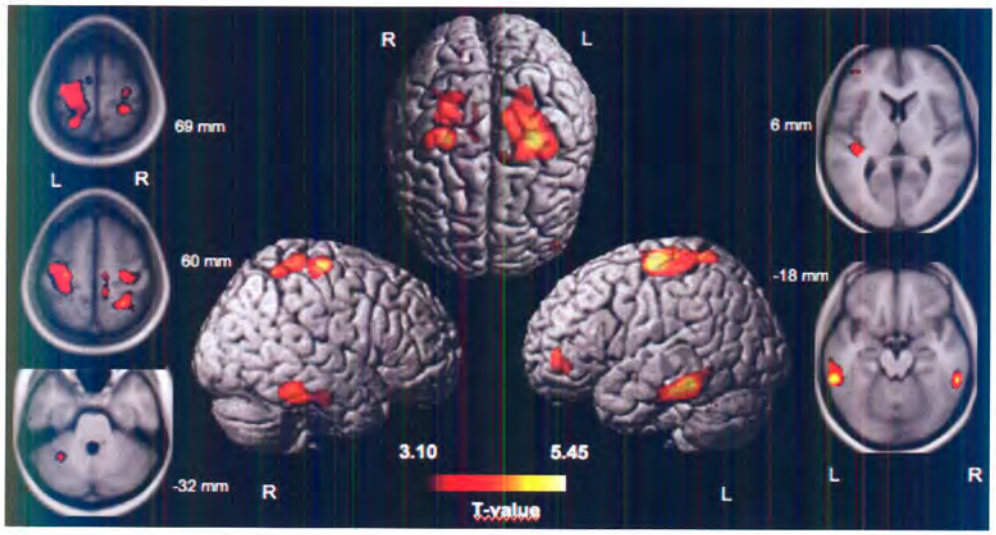



**Plate 3** A voxel-based morphometric study showing grey matter volume differences between three groups. Only voxels exhibiting a significant positive correlation between musician status (professional musician > amateur musician > non-musician) and increased grey matter volume are shown here, and those have been overlaid on the rendered cortical surface of a single subject. (This figure was adapted from Figure 1 in Gaser and Schlaug [2003]). See Chapter 18, p 199, for further details.

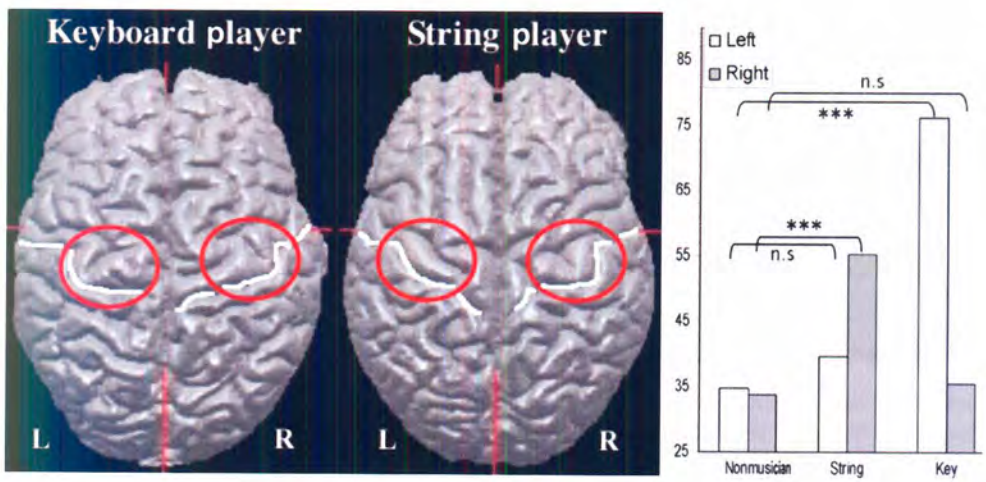



**Plate 4** Brain surface renderings of one typical keyboard and one typical string player. The central sulcus is marked with a white line. The portion of the precentral gyrus containing the configuration similar to that of the inverted Greek letter 'Omega' is found within the red circles. In the two examples, a prominent omega sign can be seen on the left more than the right in the keyboard player, and only on the right in the string player. See Chapter 18, p 200, for further details.

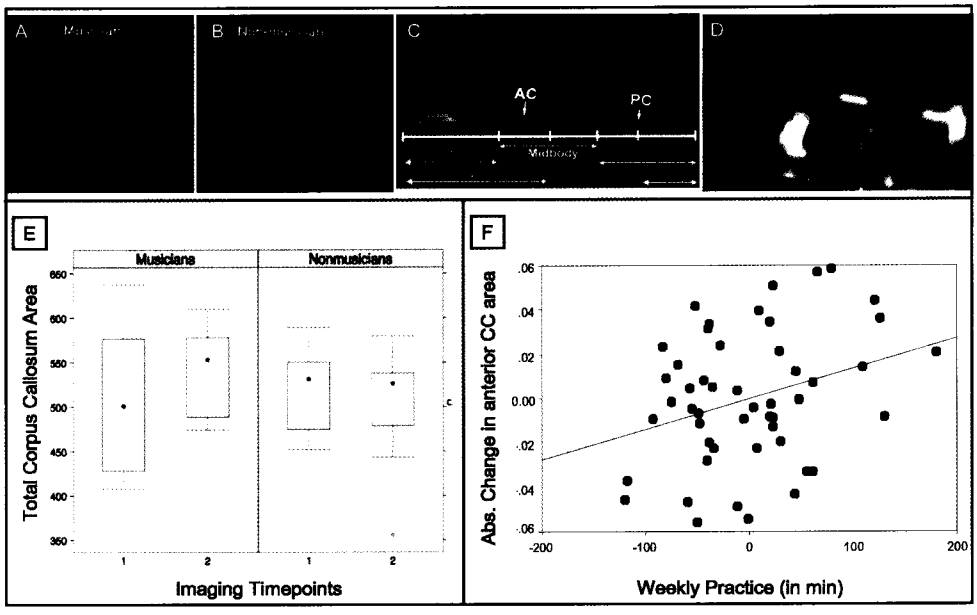



**Plate 5** Corpus callosum differences in adults (musicians vs non-musicians) and changes over time in children. (A) and (B) show the midsagittal slice of an adult musician (A) and non-musician (B) showing a difference in the size of the anterior and midbody of the corpus callosum. (C) and (D) show the major subdivisions of the corpus callosum and locations of the interhemispheric fibers connecting the motor hand regions on the right and left hemisphere through the corpus callosum.
(E) and (F) show the mean (dot inside box plots) midsagittal corpus callosum size of the 5–7-year-old children at baseline (time point 1) and after 15 months of musical training (time point 2) shows slightly opposite trends with the instrumental group showing an increase in the midsagittal corpus callosum area. The anterior portion of the corpus callosum (segments 1–4, see Fig. 18.3c) showed a significant correlation with the amount of weekly practice ($p = 0.034$; partial $r2 = 0.094$). See Chapter 18, p 202, for further details.

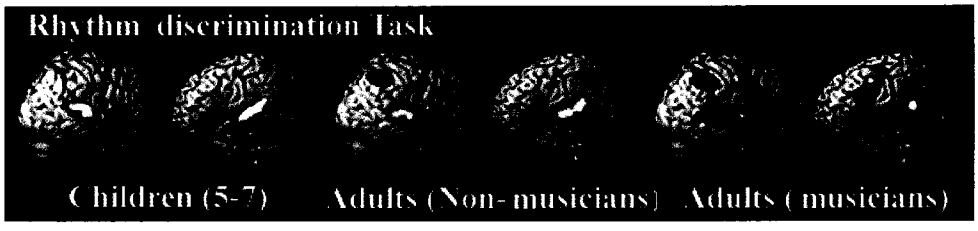



**Plate 6** Cerebral activation pattern of a rhythm discrimination task – modulated by maturity and experience. Statistical parametric images superimposed onto a surface rendering of a standardized anatomical brain depict significant activations during a rhythmic discrimination task in a group of 5–7-year-old, musically naive children, adult non-musicians, and adult musicians. The children showed prominent superior temporal gyrus activation on both sides. The adult groups show an extended pattern of activation involving polar and posterior planar regions of the superior temporal lobe as well as the parietal lobe, parts of the frontal lobe, and the cerebellum. Adult musicians differ from adult non-musicians by having less activation of the primary auditory cortex, but more activation of frontal regions bilaterally, particularly in the inferior frontal gyrus. See Chapter 18, p 204, for further details.

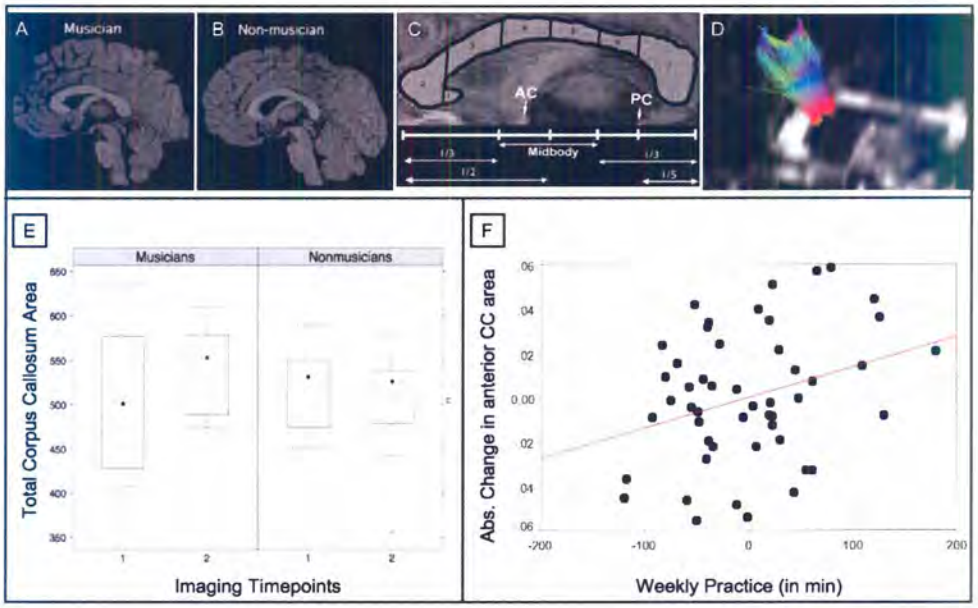



**Plate 5** Corpus callosum differences in adults (musicians vs non-musicians) and changes over time in children. (A) and (B) show the midsagittal slice of an adult musician (A) and non-musician (B) showing a difference in the size of the anterior and midbody of the corpus callosum. (C) and (D) show the major subdivisions of the corpus callosum and locations of the interhemispheric fibers connecting the motor hand regions on the right and left hemisphere through the corpus callosum.
(E) and (F) show the mean (dot inside box plots) midsagittal corpus callosum size of the 5–7-year-old children at baseline (time point 1) and after 15 months of musical training (time point 2) shows slightly opposite trends with the instrumental group showing an increase in the midsagittal corpus callosum area. The anterior portion of the corpus callosum (segments 1–4, see Fig. 18.3c) showed a significant correlation with the amount of weekly practice ($p = 0.034$; partial $r2 = 0.094$). See Chapter 18, p 202, for further details.

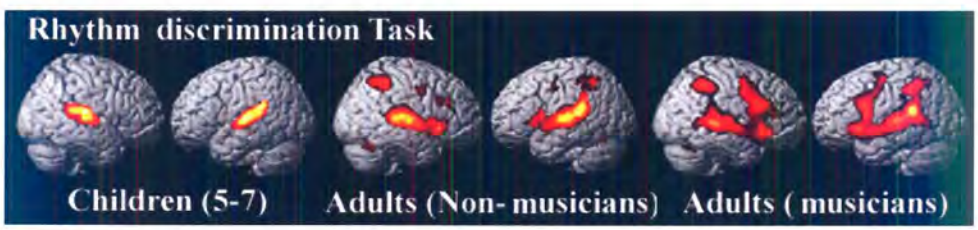



**Plate 6** Cerebral activation pattern of a rhythm discrimination task – modulated by maturity and experience. Statistical parametric images superimposed onto a surface rendering of a standardized anatomical brain depict significant activations during a rhythmic discrimination task in a group of 5–7-year-old, musically naive children, adult non-musicians, and adult musicians. The children showed prominent superior temporal gyrus activation on both sides. The adult groups show an extended pattern of activation involving polar and posterior planar regions of the superior temporal lobe as well as the parietal lobe, parts of the frontal lobe, and the cerebellum. Adult musicians differ from adult non-musicians by having less activation of the primary auditory cortex, but more activation of frontal regions bilaterally, particularly in the inferior frontal gyrus. See Chapter 18, p 204, for further details.

CHAPTER 29

# The role of the institution and teachers in supporting learning

Graham Welch and Adam Ockelford

## Introduction

Musical behaviour and development are natural by-products of the interface between intrapersonal maturational processes and an individual's enculturation into locally dominant sound worlds (e.g. Hallam and Lamont 2004; Welch 2006a) and are basic to human design. They are not dependent on the input of an institution or a 'teacher' as such, but are related to the evolution of generative skills in a sonic environment—our natural propensity to 'continually create, recreate and develop new ideas and materials' (Hallam and Lamont 2004, p. 243). Within such environments, there are various 'institutions' (including cultural settings as well as educational establishments) and 'teachers', i.e. people who have a role in musical learning, with both exemplified in the peer-to-peer popular music skill development of musicians and adolescents (Green 2001; Tarrant *et al.* 2002). Other examples are found in the interweaving of indigenous musics with the rituals of daily life, such as in the *iorram* or rowing songs of the Isle of Mull (Macnab 1970) and as practised by the Northern Ewe children of Eastern Ghana (Agawu 1995).

> *6.00 a.m.* The day's work has begun in earnest. In one home, girls are pounding dried cassava in a mortar to make *kokonte*. In another, they are pounding recently harvested rice in order to remove the husks. As with other forms of daily pounding ... the work ... is made a little less routine by incorporating some rhythmic interest. In place of a regularly spaced alternation between two pounders ... a variant may be introduced. Pounder 1 keeps a steady pace while Pounder 2 pushes her strokes closer to Pounder 1's. Here as elsewhere in Northern Ewe culture, work merges into play and reemerges into work.
>
> Agawu (1995, p. 12)

Contemporary explanations for the mechanism for the process of learning in social and cultural settings often draw on the work of Vygotsky, Luria and Leont'ev in the first half of the twentieth century (cf. Bannon 1997; Cole 1999), as well as being related to systems theory—the interlinking of relationships within some form of organization—(von Bertalanffy 1968), and social ecology theory—the nurturing of development within social contexts (Bronfenbrenner 1979). The early Russian investigators and those who developed their work subsequently (such as the Finnish researcher Engeström) explored how learning and development are the product of inter- and intrapersonal behaviours that are shaped by cultural artefacts (e.g., literature), alongside tools (including psychological tools, e.g., language and other symbol systems), expectations, 'rules'/conventions and norms. The internalization of artefacts is also seen to facilitate the agency of the individual, such that the artefacts themselves are modified through personal use, enabling the possibility of consequent change within the culture. Thus there is an

ongoing mediation process in how individuals interact with the world around them and make sense of their reality.

A key concept in this view of culturally based learning is 'activity', which has been defined as 'the engagement of a subject toward a certain goal or objective' (Ryder 2005). One widely cited model of activity within a system is provided by Engeström (1999, 2001) (see Figure 29.1). In his interrelated system of elements as applied to education, the 'subject' (the learner) is supported in reaching the 'object' (the intended learning outcome) through interaction with various 'mediating artifacts' (such as language and other symbol systems) (a conceptualization strongly associated with the work of Vygotsky). This process is seen as being embedded in a social context that provides support for the activity through the subject and intended outcome being located within a 'community' that has 'rules' (expectations for behaviour) and also the likelihood of a 'division of labour' (diversity of effort). Because of the possibility of tensions within the activity system, it may be that the actual outcome is at variance with the intended object (i.e. that there are unintended outcomes—as explored below).

If the activity system is seen in relation to learning within an institution (such as a school or social collective, for example, the family) in which certain people either adopt or are expected to have, roles as 'teachers', then it is possible to envisage how the theorized activity system might be able to support the intended learning in music (or fail to). The world of music has certain characteristics:

- There are many different musical genres and subgenres;
- Each has its own customary view of what counts as musical learning, or at least the outcomes of musical learning in relation to performance, as well as the traditions (custom and practice) in how learning is usually fostered in relation to the genre's characteristic features, as exemplified in the 'institutions' of India, for example the tradition of musical households or *Gharanas* (Farrell 2001) and Japan, for example the culture schools managed by major business and voluntary organizations (Murao and Wilkins 2001);
- In many diverse musics, there is evidence of the high status accorded to the accomplished expert who demonstrates solo mastery over the sonic material;
- Certain cultures, such as China, Afghanistan and India, continue to have strong traditional music genres. These are characterized by expertise transmission within families across several (or many) generations (Jones 1995; Doubleday and Baily 1995; Farrell 1997). In such traditional cultures, the 'teacher' is in

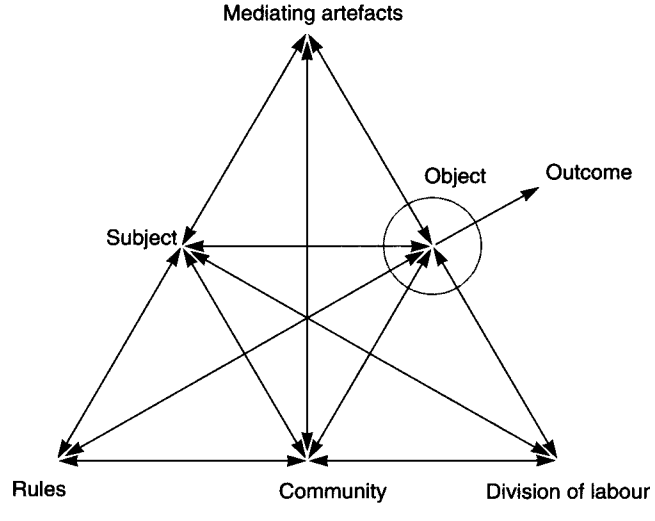



**Fig. 29.1** The structure of a human activity system (Engeström 2001b, p. 136).

a master/apprentice relationship with the 'pupil' within a strong local community, often fostered by pupils playing for each other in the presence of the master.

Aspects of the role of institutions and teachers in the process by which musical expertise is learned and demonstrated are illustrated in the four examples that follow (see below). The first relates to musical development in a special education context, the second to advanced music learning in higher education, the third to the relatively new music education world of the female cathedral chorister and the fourth to pupils' experiences of lower secondary school classrooms in England. In the first three cases, the intention is to suggest how the teacher and institution support the activity of learning within a social context, whilst the last demonstrates that not all music learning in within an institution is necessarily positive or as intended.

## The role of institutions and teachers in supporting musical learning: four examples

### Example one: supporting musical learning in a young prodigious musical savant

Derek Paravicini is a musical savant, now in his late twenties, who is well known to the public through success on the concert platform as a jazz and popular pianist on both sides of the Atlantic. His early life was very difficult, and the initial development of his musical expertise was highly unconventional (Ockelford 2007).

Derek was born premature at 26 weeks, weighing a little over 700 g. In the fight for survival that followed, he lost his sight through retinopathy of prematurity, and developed unspecified neurological impairments that meant he grew up with severe learning difficulties. His family were upper-class and employed the services of an experienced Nanny to care for him and to oversee his upbringing. Neither she, nor anyone else, expected Derek to develop musically in a way that was at all out of the ordinary (there was no history of exceptional musical development in the family). However, following the diagnosis of Derek's blindness a few weeks after returning home from hospital, Nanny decided that sound was likely to play an important part in his life. So she sang to him constantly and surrounded him with sound-making toys. Just like any other baby's, Derek's environment was also perfused with music from the TV, radio and other incidental sources.

Derek was attracted to music as a potential source of stimulation and order in the world around him, and unbeknown to Nanny or his family, his ability to process musical sounds developed rapidly and precociously. Desperate to find an activity that would keep the 1-year-old Derek gainfully occupied, Nanny gave him his grandfather's little electric keyboard to play with. From the start, Derek loved the sounds that it produced and discovered that he could imitate some of the musical sounds that he had came into contact with. With no intervention on the part of Nanny or his family, and with no visual model to guide him, Derek taught himself how to get his hands and fingers in the right places at the right times to recreate some of the snatches of melody and harmony that were familiar to him. This situation—Stage 1 in the journey of Derek's musical learning—is summarized in terms of Engeström's model in Figure 29.2.

One day, after about six months of self-directed exploration at the keyboard, Nanny heard Derek play a version of 'Cockles and Mussels'—using both hands, with a tune and rudimentary accompaniment. His repertoire soon widened, and Derek's relationship with Nanny, his extended family and friends took a new course with the addition of this unexpected but welcome dimension. Still there was no *formal* intervention or guidance, though, from those around Derek in terms of supporting his learning. In the months and years that followed, Derek continued to chart his own, unique, autodidactic course (Stage 2: Figure 29.3).

At the age of five, Derek came to the attention of the second author (AO), who was then teaching in a school for the blind in south-west London. Derek's raw talent was evident, as were his technical eccentricities—including the use of his knuckles, hands and even the occasional

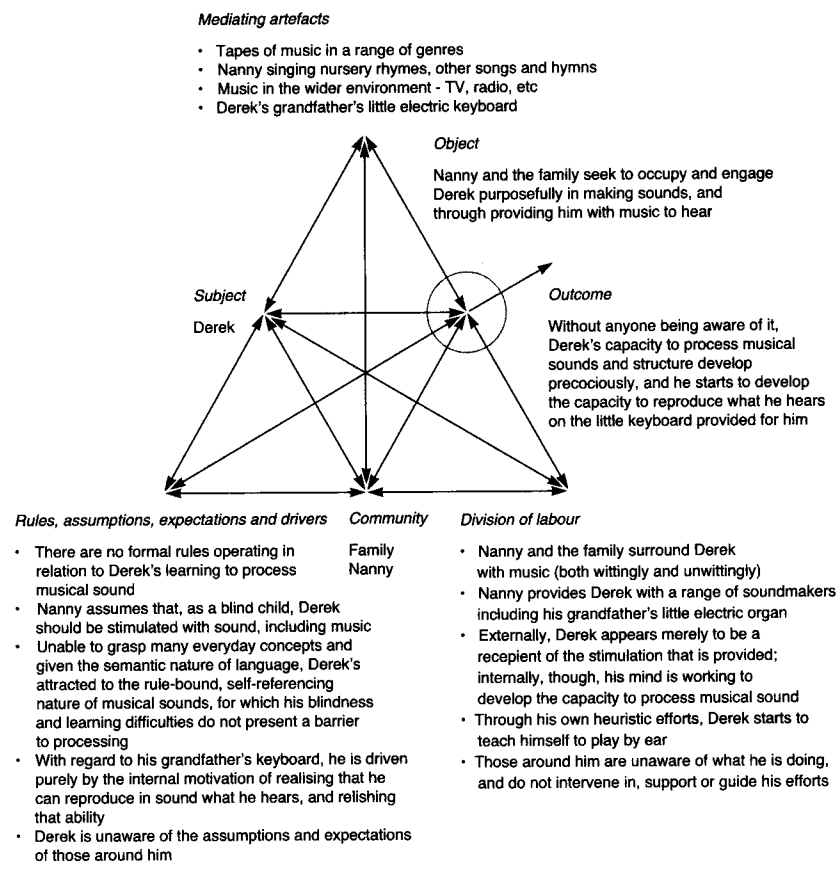



**Fig. 29.2** Stage 1, 0–21 months.

judicious dip of his nose which enabled him to reproduce sonorities that were beyond the reach of his limited span (around a fifth on the standard-sized keyboard). In AO's view, while Derek's achievements up to that point were remarkable, particularly given his blindness and severe developmental delay, it was essential that he should come to accept the intervention of another, to guide his further musical development and realize his creative intent. However, Derek had never experienced anything approaching a conventional 'teacher–pupil' relationship, and he was not remotely inclined even to share *his* piano with anyone else, let alone engage in a structured learner–teacher dyad. Hence, a large part of AO's effort for the first six months of working with Derek was directed towards showing him that interaction through music could be productive and, above all, enjoyable. Gradually, involving an initial degree of physical intervention, Derek did come to appreciate that discourse through the medium of musical improvisation was possible and could indeed be a source of great pleasure.

Engaging in social discourse through sound could not alone solve Derek's technical challenges, however, and a final stage in AO's early relationship with him was necessary, whereby Derek would allow AO to show him—physically—how to hold his hands and use his fingers in conventional patterns that would facilitate his technical development. Clearly, Derek had no concept of the goals that AO was pursuing, so it was critical that he enjoyed being guided

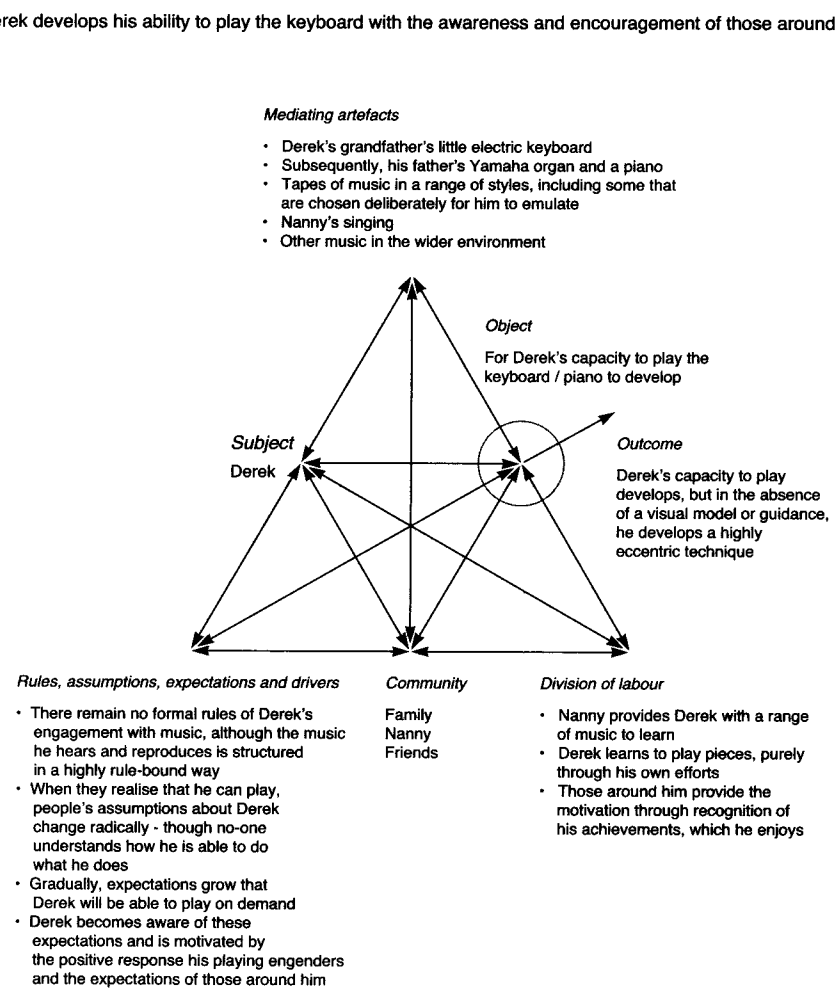



**Fig. 29.3** Stage 2, 21–66 months.

through the daily ritual of technical exercises, scales and arpeggios that AO devised for him. Luckily, Derek relished the one-to-one attention and the orderliness of his practice routine, and the pattern of highly formalized intervention with AO continued throughout his childhood on a daily basis (Stage 3; Figure 29.4).

There is much to learn from Derek's story: most obviously, the fact that an individual can be motivated to pursue musical learning to a highly advanced level at an early age with no direct intervention or encouragement on the part of others, and with no global sense of moving towards the goal of becoming a competent performer (see also Ockelford 2008, Ockelford *et al.* 2006). Beyond this, however, it is also the case that Derek, as someone with severe learning difficulties, could, initially through a discourse comprising nothing but musical sounds, develop a relatively conventional teacher–pupil relationship, that eventually enabled him as an adult to have a career as an internationally recognized musician.

## Example two: supporting musical learning in higher education

Higher education (HE), on the other hand, is a relatively (often highly) selective educational environment that seeks to advance already competent

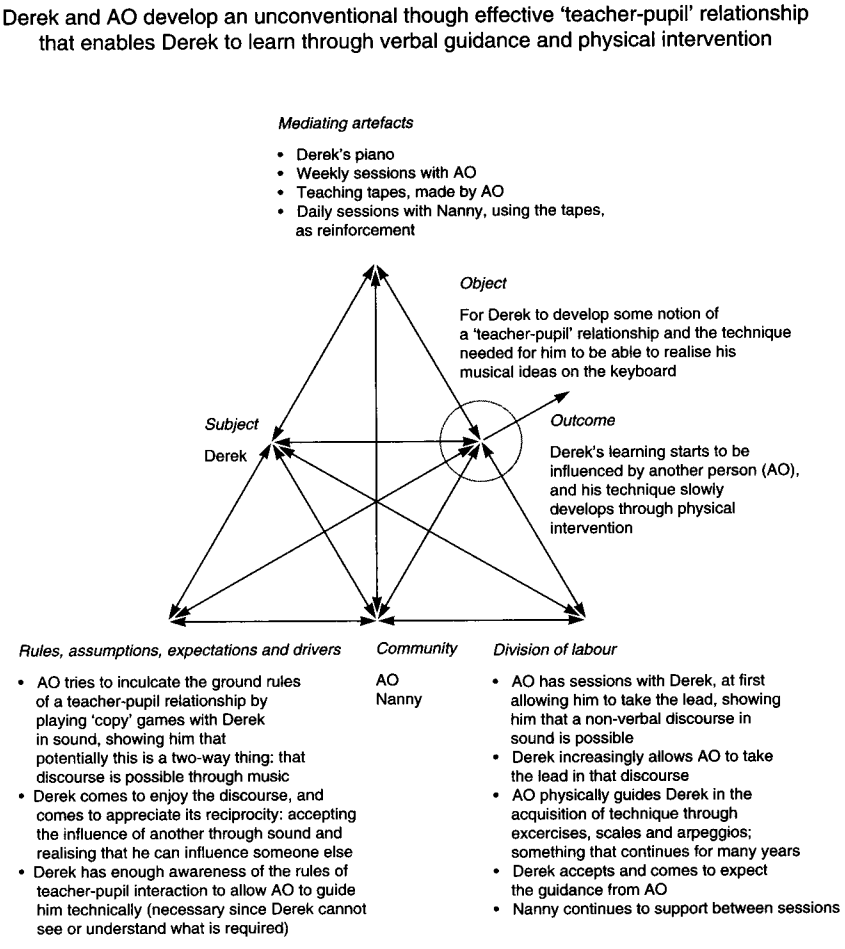



**Fig. 29.4** Stage 3, from 66 months.

musical skill levels in young people who normally exhibit (or report) little or no disability (though see for example Lerner and Straus 2006).

Within the spectrum of HE in music, there are observable similarities and differences in the way that the activity of music learning is processed. These are particularly related to context, such as the age and gender of performers, their principal musical genre and the particular HE location. For example, an ongoing study of advanced music learning and teaching in four UK higher education institutions (HEIs) is investigating how classical, popular, jazz and Scottish traditional musicians deepen and develop their learning about performance in undergraduate, postgraduate and wider music community contexts (Welch *et al.* 2006). In the first year of the study, a specially designed, web-based questionnaire was used to survey 244 musicians across the four HEIs. In addition to demographic information, participants provided self-reports about their earliest engagement with music (including first instrumental or vocal lessons), secondary and tertiary education, as well as significant musical experiences and influences. The participants were questioned about the perceived relevance of a range of musical skills and activities, experience of performance and general life anxiety, how they spent their time in an average week, the pleasure that they derived from engagement in musical activities, various self-views (musical self-efficacy, self-esteem,

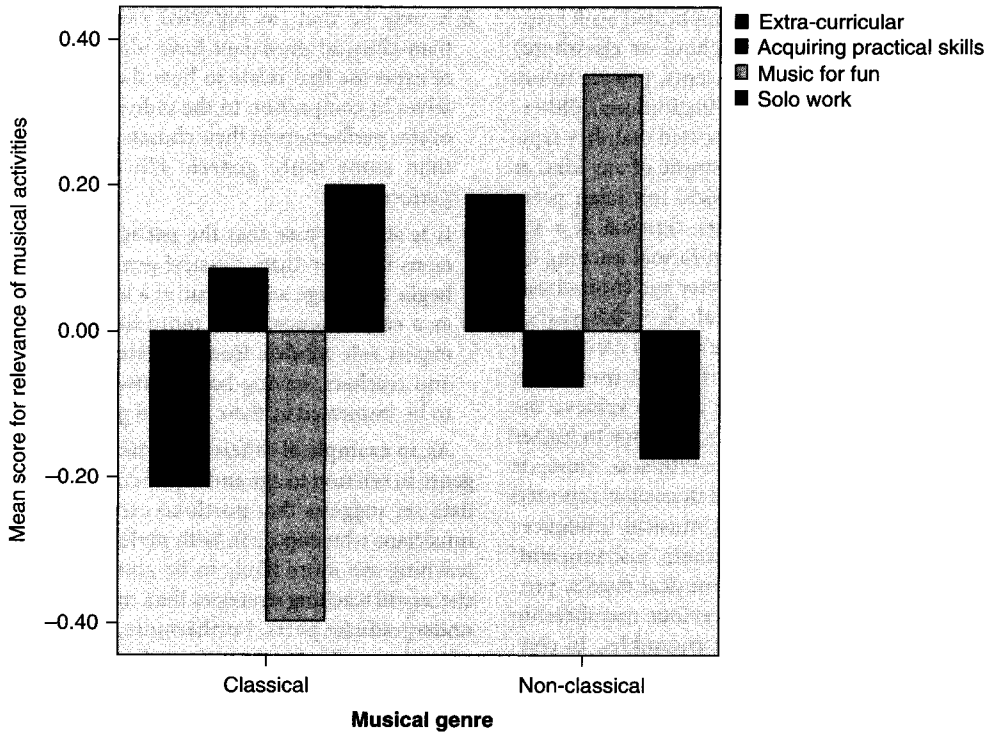



**Fig. 29.5** Classical and other-than-classical-musicians' mean scores for categories of musical activities (*extra-curricular*: listening to music from own and outside of own genre, acquiring general musical knowledge, engaging in professional conversations, networking; *acquiring practical skills*: practising alone, practising with others, taking lessons, solo and group performance, listening to music from own genre; *music for fun*: playing for fun alone or with other; *solo work*: mental rehearsal, giving lessons, solo performance). (Creech *et al.* 2008, p223)

self-regulation), group membership and their beliefs about the nature of expertise in musical performance and teaching.

The resultant data analyses suggest that all musicians attached great importance to achieving a high overall standard of performance, although they had different perspectives on the processes by which this might be achieved (Creech *et al.* 2008). For example, in terms of the performance expectations of their particular musical 'community' and its requisite 'tools' and 'rules' (*pace* Engeström), classical musicians ranked the ability to improvise as the least important musical skill, but perceived sight-reading to be very important. In contrast, in relation to the performance expectations of *their* particular communities, the other-than-classical musicians (jazz, popular, Scottish traditional) tended to assign the least importance to the ability to sight-read, but placed greater emphasis on playing from memory and improvisation. Although all musicians recognized the value of practice, the community of classical musicians tended to place greater emphasis on practising alone, whereas other-than-classical musicians attached greater relevance to making music for fun, networking and extra-curricular activities such as listening to a diverse range of musics and engaging in professional conversations with peers (see Figure 29.5). Nevertheless, both classical and other-than-classical groups considered musical expertise to involve the possession of global musical skills that could be transferred to other musical genres (Papageorgi *et al.* in press).

An exploration of underlying processes that might explain the similarities and differences

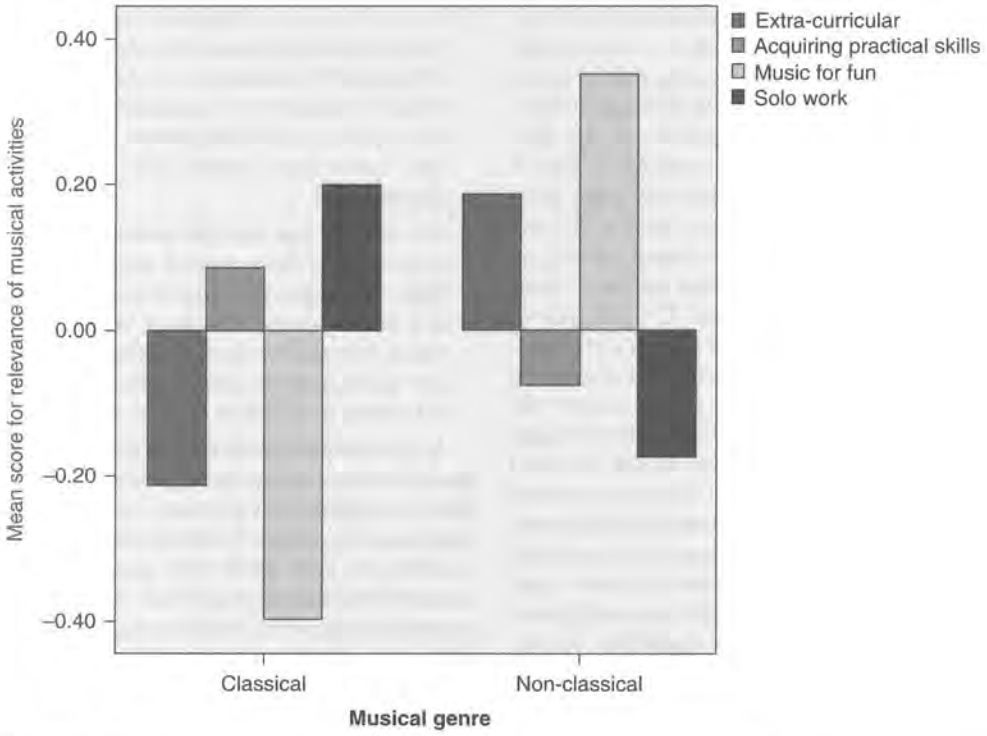



**Fig. 29.5** Classical and other-than-classical-musicians' mean scores for categories of musical activities (*extra-curricular*: listening to music from own and outside of own genre, acquiring general musical knowledge, engaging in professional conversations, networking; *acquiring practical skills*: practising alone, practising with others, taking lessons, solo and group performance, listening to music from own genre; *music for fun*: playing for fun alone or with other; *solo work*: mental rehearsal, giving lessons, solo performance). (Creech *et al.* 2008, p223)

self-regulation), group membership and their beliefs about the nature of expertise in musical performance and teaching.

The resultant data analyses suggest that all musicians attached great importance to achieving a high overall standard of performance, although they had different perspectives on the processes by which this might be achieved (Creech *et al.* 2008). For example, in terms of the performance expectations of their particular musical 'community' and its requisite 'tools' and 'rules' (*pace* Engeström), classical musicians ranked the ability to improvise as the least important musical skill, but perceived sight-reading to be very important. In contrast, in relation to the performance expectations of *their* particular communities, the other-than-classical musicians (jazz, popular, Scottish traditional) tended to assign the least importance to the ability to sight-read, but placed greater emphasis on playing from memory and improvisation. Although all musicians recognized the value of practice, the community of classical musicians tended to place greater emphasis on practising alone, whereas other-than-classical musicians attached greater relevance to making music for fun, networking and extra-curricular activities such as listening to a diverse range of musics and engaging in professional conversations with peers (see Figure 29.5). Nevertheless, both classical and other-than-classical groups considered musical expertise to involve the possession of global musical skills that could be transferred to other musical genres (Papageorgi *et al.* in press).

An exploration of underlying processes that might explain the similarities and differences

between these two groups implicated both institutions (whether home, school or elsewhere) and teachers (including parents, private tutors and teachers in educational institutions). Other-than-classical musicians reported that they typically began to engage with music of any kind at a later age than their classical musician peers (non-classical: $\bar{X}$ = 8.4 years, classical: $\bar{X}$ = 6.6 years). Similarly, they began formal learning on their first instrument at a later age (non-classical: $\bar{X}$ = 12 years, classical: $\bar{X}$ = 8.8 years). Notwithstanding the nature of their early musical experiences, all the participant musicians, irrespective of genre, were able to achieve the requisite baseline skill levels for entry to higher education and beyond. Nevertheless, contexts were not identical. Classical musicians reported that their most important musical influences (past and present) were parents, instrumental/vocal teachers, significant musical events, professional colleagues and previous membership of county (regional) music ensembles. In contrast, other-than-classical musicians claimed to be particularly influenced by well-known performers, as well as university or college lecturers and informal groups of friends (Creech *et al.* 2007). Across all musicians, irrespective of genre, higher education tutors, whether lecturers in particular aspects of music or specialist solo instrumental and vocal teachers, were reported to be significant agents in a communal process of advanced music learning.

Overall, participant classical musicians rated themselves higher in terms of perceived musical expertise. The basis for this difference is likely to relate to (1) the comparative longevity of classical musical cultures in HE, (2) other aspects of participants' group-based self-views and (3) differences in participants' cumulative years of study. For example, interviews with senior academics in the participant HEIs revealed that:

1 Classical music had been established for much longer in their academic programmes compared to the three selected other-than-classical genres (jazz, popular and Scottish traditional). Concomitantly, classical music teaching and learning and assessment practices were reported to be more firmly embedded, rehearsed, formalized and understood within their particular communities;
2 It may be that (as reported earlier) other-than-classical musicians have idealized views of expertise that relate to how they see themselves in comparison to the individual quality of star performers in their chosen genre rather than some more generic HE measure of performance;
3 It is also the case that the participant musicians in other-than-classical genres typically begin to engage with music at a later age and, as a consequence, were more likely to have expert role models from outside their peers and teachers, having had relatively less time to be immersed in their musical genre.

As an example of differences *within* a musical genre in relation to age and experience, the same data set suggests that portfolio career classical musicians who engage in both performance and teaching are more likely to be able to identify successful teaching strategies than their younger, undergraduate peers. Furthermore, the activity of teaching, allied to extensive solo performance experiences, are likely to reduce levels of performance anxiety.

## Example three: supporting female music learning in a UK cathedral setting

Male choristers have participated in UK cathedrals since their inception in 597AD at Canterbury. In comparison, it was not until 1991 that Salisbury became the first old cathedral foundation to admit girls on the same basis as boys. The political impact and success of their initiative (although foreshadowed by other religious institutions earlier in the twentieth century but without the same publicity) has led to a growing (sometimes grudging) acceptance of female choristers within the previously all-male culture. By 2006, a majority of cathedrals had choristers of both sexes for the first time in their long history, even though it continues to be relatively rare for the two sexes to sing together, other than at special festival events (Welch 2007).

One of the cited reasons for the longevity of the all-male cathedral music tradition was that young females were regarded as being unable to sing with the same 'pure' quality of vocal timbre demonstrated by the young male voice in the

performance of the cathedral sacred music repertoire. This belief does have some basis in the physical realities of child voice acoustics, even though research has demonstrated its fallibility. First, there are slight differences in the relative sizes of girls' and boys' vocal anatomy, the male being slightly larger throughout childhood and into adolescence (cf. Welch and Howard 2002) and these could be expected to generate perceptible disparities in acoustic outputs. Secondly, with regard to untrained children's voices, there is increasingly perceptible psycho-acoustic differentiation between the sexes as they progress through childhood (Sergeant *et al.* 2005; Sergeant and Welch in press), with observable gender-related differences in their sung spectra. However, the power of the musical activity system in the cathedral is such that formal induction of girls into its performance expectations can generate changes in their basic vocal behaviour such that any gender differences are reduced significantly and often become imperceptible. A series of perceptual studies from the mid 1990s onwards has demonstrated that membership of a cathedral choir can allow girls to be trained to produce sounds that are 'boy-like' in character in order to match the customary, male-biased, performance expectations of the musical repertoire (Sergeant and Welch 1997; Howard and Welch 2002; Welch 2006; Welch and Howard 2002).

Detailed longitudinal case study data from one cathedral suggests that the activity of becoming a female chorister is closely linked to the customary tripartite relationship in music (Small 1999) between the physical setting, people (performers and listeners) and the way that the musical soundscape constrains the variety

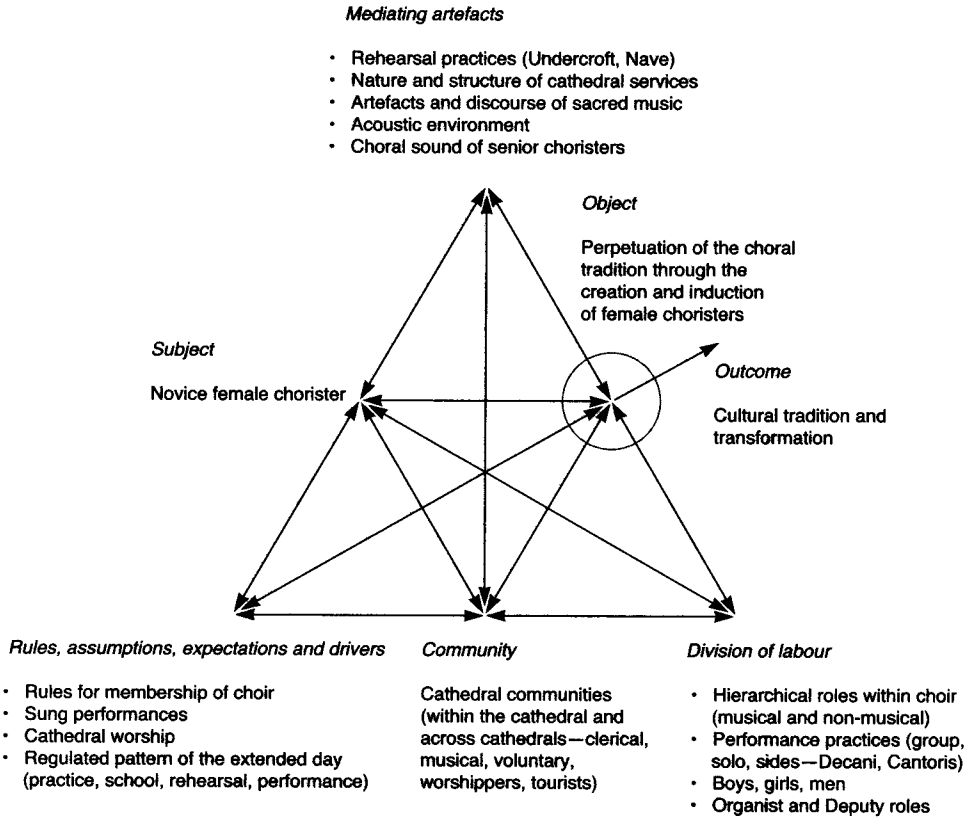



**Fig. 29.6** An example of the activity system that frames the development of the novice (female) cathedral chorister (Welch 2007, p. 29).

of possible musical outcomes. The induction process for both young males and females involves the novice chorister (probationer) being required to learn, practise, rehearse and perform music systematically many times each week across the school and ecclesiastical year, whilst standing between and listening to the vocal models of established senior performers (head chorister, deputy head chorister, senior corner girls), supported by group-based (communal) teaching by a highly experienced church musician (organist and choir director or deputy), in order to master a ritualized repertoire (such as the *Introit, Psalms, Response Settings* and *Anthem* for Evensong) that involves both choral and solo performance: see Figure 29.6 for how this experience can be framed within an activity system (Welch 2007).

The acoustic features of the rehearsal and performance spaces within the chancel and choir of the cathedral also shape the learning experience, with features of the choristers' sound amplified by the high stone vaulting of the nave and adjacent spaces and fed back to the performers. The collective and ritual nature of the performance is reinforced by the addition of adult male voices (for most services) and the organ. It is not surprising, therefore, that females are able to perform the established religious repertoire in customary fashion, given the interrelated expectations of the musical learning process and the longevity of the performance practices within the culture that are handed down across generations of cathedral musicians (and where it is extremely rare for a new director to be appointed from outside the tradition).

## Example four: supporting music learning in the lower secondary school

In contrast to the above, it is possible that institutions and teachers can also hinder as well as support learning in music. For example, notwithstanding recent improvements in the quality of music education across all four curriculum 'stages' in English schools, embracing the ages 5–16 (Welch 2006b), there is often a particular disparity in the individual experiences of pupils in the lower secondary school (ages 12–14). Although the independent school inspection body, OfSTED,[1] has reported a steady improvement in lower secondary school music teaching since 2001, the overall quality in the other school age phases (both older and younger) has often been reported as higher. The reasons for these systematic differences are likely to relate to:

1 The organization of the timetable in primary schools where one class teacher tends to take all curriculum subjects and, therefore, is in a better position to know each individual child across a school year and to match music teaching to individual need,
2 English National Curriculum music learning expectations of primary school children are lower (in terms of the complexity of musical behaviours) than in secondary schools; this may encourage more positive judgements from school inspectors (who are themselves unlikely to be primary music specialists);
3 Upper secondary school music (ages 14–16) is optional, catering for a small minority of pupils (between 7–9 per cent) who have opted to study music for a further two years to examination at age 16; and
4 Lower secondary school music is compulsory for all pupils—a period of intense personal development in their musical identity (Tarrant *et al.* 2002).

Nevertheless, some secondary schools have been able to counter this trend, even in contexts where 'success' in school examinations may not be the norm. For example, a recent analysis of data from secondary schools in one part of inner London with a high Black African-Caribbean (BAC) population found that up to 62% of such pupils were opting to study music (with an average across schools of 24%). This was much higher than the non-BAC groups in the same schools (average 8%) (Spence 2006). Furthermore, these pupils went on to achieve higher grades in their 16+ examinations for music than in their other

[1] The Office for Standards in Education (OfSTED) assesses the quality of music teaching on behalf of the Government. This is a non-ministerial Government department that reports directly to the UK Parliament. OfSTED is headed by a senior civil servant and is required to inspect and report on the quality of education in schools and in initial teacher education.

subjects, such as English, mathematics and science. A recurrent theme from interviews with these young people was that lesson content had greater enjoyment when it drew on music that they knew, liked and/or could play. They also reported greater liking of teachers who showed respect for them and their music. Such positive emotional engagement appears to be reflected in their relative success in school music.

Further insights in why some young people are 'successful' at school music and some are not are reported in a study of lower secondary school music (Saunders 2006). Pupils in the final year of compulsory music education in England (Year 9, age 13+) have to make a decision about whether or not to continue their school music studies. Coding and analyses of the transcripts of interviews with 44 pupils in one school revealed that pupils tended to see themselves in one of five categories (see Table 29.1).

Only one of the five groups is fully engaged with school music (termed the 'traditional Western musician'), in part because of their formal instrumental performance skills and an ability to identify successfully with the content of the school music curriculum. However, four other groups of pupils have alternative, less positive perspectives. Some see themselves as non-musicians; some have self-taught instrumental skills, but do not identify with school music, even though they are accomplished musicians ('alternative Western musician'); some are accomplished musicians in a traditional sense, but believe that school music is irrelevant and a possible threat to their musician status ('disengaged traditional Western musician'); and some find that the mismatch between their musical skills and the skills needed for the school curriculum is intimidating to their self-concept as a musician ('disengaged alternative Western

**Table 29.1** Pupils' self-identified grouping related to classroom music in Year 9 (age 13+) of lower secondary school

| Pupil identity related to school music | Example key characteristics |
|---|---|
| Traditional Western musician | Strong engagement with all music; good family support |
| | Successful experience of formal learning of an instrument |
| | Peer group status as a 'musician' |
| Western non-musician | Strong engagement with music outside school; limited family support |
| | Limited instrumental skill; immediate peer group do not identify with school music; peer group recognition as a 'non' musician |
| Alternative Western musician | Strong engagement with all music; family support |
| | Informally/self-taught instrumental skills; school music seen as irrelevant; peer group status as a 'musician' |
| Disengaged traditional Western musician | Strong engagement with music outside school |
| | Formally taught instrumental skills, family support variable |
| | Peer group do not identify with school music |
| | Music curriculum seen as a 'threat' to musician status within peer group |
| Disengaged alternative Western musician | Strong engagement with music outside school |
| | Informally or self-taught instrumental skills; family support limited |
| | Peer group share counter culture of musical identity |
| | 'Musician' status threatened by lack of success |

musician'). When interviewed, a significant number of young people consider themselves (wrongly) to be unsuitable for continuing music studies. Yet virtually all connect successfully with music outside school, as consumers, listeners, and/or practitioners. Unless pupils identify and find success with school music (e.g. Spence 2006), they are likely to disengage. Subsequent research and analyses of pupil interviews in three other schools has confirmed these findings (Saunders 2008). Comments from individual pupils demonstrate the kinds of diversity in their experiences of music in school:

> It's my favourite lesson of the whole week and well, I wish we could do a bit more of it really.

> I think the stuff at lunchtimes and for the shows is really good, but that lessons aren't anywhere near as exciting as that. I suppose that's because we can't all sing and play like them though.

> Music at school is pants; it's just not any good. It's all duff music. I mean, who listens to this stuff?

> I'm learning from my brother and his mate taught him, so when we get together, it's a bit scrappy, but we're getting better. I don't do it that much in class cos Miss wants us working on the keyboards and I'm not great at that.

Collectively, these comments illustrate the underlying positive connection that these young people have with music, although not necessarily the kinds of music (and musical practices) found in school. Adolescence is a period characterized by the emergence of strong musical identities (Zillman and Gan 1997), often with a clear perception of boundaries between different musics as young people use music to negotiate positions for themselves within their group identities (O'Neill and Green 2001). It is not surprising, therefore, that school music has such a mixed reception during this age phase.

## Conclusion

Overall, the various examples cited above are an indication of how learning and teaching in music are shaped by processes outside the individual, not least because of the influences of group membership (allied to age and gender), performance expectations and practices, and professional and institutional cultures. The process of individual induction into the characteristics of a particular musical culture by teachers and institutions influences the formation of identities in music (cf. MacDonald *et al.* 2002), for better or for worse, at least in terms of dominant models within the culture. Indeed, the development of music teachers themselves can be seen within an activity system, i.e. the teacher's understanding of their role is developed both by informal personal reflection of the experience of performance and their own learning and, more systematically, through their own induction process by attendance at a specialist pedagogically focused institution.

## References


Agawu K (1995). *African rhythm.* Cambridge University Press, Cambridge.

Bannon L (1997). *Activity theory.* Retrieved 24 October 2005 from http://wwwsv.cict.fr/cotcos/pjs/TheoreticalApproaches/Actvity/ActivitypaperBannon.htm

Bronfenbrenner U (1979). *The ecology of human development: experiments by nature and design.* Harvard University Press, Cambridge, MA.

Cole M (1999). Cultural psychology: some general principles and a concrete example. In Y Engeström, R Miettinen and R-L Punamäki, eds, *Perspectives on activity theory*, 87–106. Cambridge University Press, Cambridge.

Creech A, Papageorgi I, Duffy C, Morton F, Haddon L, Potter J, de Bézenac C, Whyton T, Himonides E and Welch GF (2008). Investigating musical performance: commonality and diversity amongst classical and other-than-classical musicians. *Music Education Research*, **10**(2), 215–234.

Doubleday V and Baily J (1995). Patterns of musical development among children in Afghanistan. In EJ Fernea, ed., *Children in the muslim Middle East*, 431–444. University of Texas Press, Austin, TX.

Engeström Y (1999). Activity theory and individual and social transformation. In Y Engeström, R Miettinen, R-L Punamäki, eds, *Perspectives on activity theory*, 19–38. Cambridge University Press, Cambridge.

Engeström Y (2001). *Expansive learning at work: toward an activity-theoretical reconceptualisation. With a commentary by Michael Young.* Learning Group, Occasional Paper No 1. Institute of Education, London.

Farrell G (1997). *Indian music and the West.* Clarendon Press, Oxford.

Farrell G (2001). India. In DJ Hargreaves and AC North, eds, *Musical development and learning: the international perspective*, 56–72. Continuum, London.

Green L (2001). *How popular musicians learn.* Ashgate, Aldershot.

Hallam S and Lamont A (2004). Learners: their characteristics and development. *Psychology of Music*, **32**(3), 243–252.

Howard DM and Welch GF (2002). Female chorister development: a longitudinal study at Wells, UK. *Bulletin of the Council for Research in Music Education*, **153**(4), 63–70.

Jones S (1995). *Folk music of China: living instrumental traditions*. Clarendon Press, Oxford.

Lerner N and Straus JN (eds) (2006). *Sounding off: theorizing disability in music*. Routledge, New York.

Mcnab PA (1970). *The Isle of Mull*. (David and Charles), Newton Abbot.

MacDonald R, Hargreaves DJ and Miell D (eds) (2002). *Musical identities*. Oxford University Press, Oxford.

Murao T and Wilkins B (2001). Japan. In DJ Hargreaves, AC North, eds, *Musical development and learning: the international perspective*, 87–101. Continuum, London.

Ockelford A (2007). *In the key of genius: the extraordinary life of Derek Paravicini*. Hutchinson, London.

Ockelford A (2008). *Music for children and young people with complex needs*. Oxford University Press, Oxford.

Ockelford A, Pring L, Welch GF and Treffert D (2006). *Focus on music: exploring the musical interests and abilities of blind and partially-sighted children with septo-optic dysplasia*. Institute of Education/RNIB, London.

O'Neill S and Green L (2001). Social groups and learning in music education. In BERA Music Education Review Group, *Mapping music education research in the UK*, 26–31. British Educational Research Association, Macclesfield.

Papageorgi I, Creech A, Duffy C, Potter J, Whyton T, Morton F, Haddon L, de Bézenac C, Himonides E and Welch GF (in press). Investigating musical performance: the development and prediction of expertise in advanced musical learners. *Psychology of Music*.

Ryder M (2005) *What is activity theory?* Retrieved 24 October 2005 from http://carbon.cudenver.edu/~mryder/itc_data/act_dff.html.

Saunders J (2006). Music learning in year 9: the pupils' perspective. *NAME Magazine*, **18**, 9–11.

Saunders J (2008). The music classroom: pupil experience and engagement during adolescence. Unpublished PhD thesis. Institute of Education, University of London, London.

Sergeant DC, Sjölander PJ and Welch GF (2005). Listeners' identification of gender differences in children's singing. *Research Studies in Music Education*, **24**, 28–39.

Sergeant DC and Welch GF (1997). Perceived similarities and differences in the singing of trained children's choirs. *Choir Schools Today*, **11**, 9–10.

Sergeant DC and Welch GF (in press). Age-related changes in long-term average spectra of children's voices. *Journal of Voice*. Published online July 2007, available at: http://www.jvoice.org.inpress.

Small C (1999). Musicking—the meanings of performing and listening. A lecture. *Music Education Research*, **1** (1), 9–21.

Spence S (2006). Black Caribbean children and school music: expectation and achievement. *NAME Magazine*, **18**, 12–15.

Tarrant M, North AC and Hargreaves DJ (2002). Youth identity and music. In R MacDonald, DJ Hargreaves and D Miell, eds, *Musical identities*, 134–150. Oxford University Press, Oxford.

Welch GF (2006a). The musical development and education of young children. In B Spodek, O Saracho, eds, *Handbook of research on the education of young children*, 251–267. Lawrence Erlbaum Associates Inc., Mahwah, NJ.

Welch GF (2006b). What research into music teacher education tells us about the contexts and challenges for teacher education. *Proceedings, Beijing International Forum on Music Education* 2006, 84–97. 15–17 May, Beijing: NAMM/CSME/ISME/CNU. [Chinese version, 98–113.]

Welch GF (2006c). Singing and vocal development. In G McPherson, ed., *The child as musician: a handbook of musical development*, 311–329. Oxford University Press, Oxford.

Welch GF (2007). Addressing the multifaceted nature of music education: an activity theory research perspective. *Research Studies in Music Education*, **28**, 23–38.

Welch GF, Duffy C, Potter J and Whyton T (2006). *Investigating musical performance (IMP): comparative studies in advanced musical learning*. Institute of Education, University of London, funded by the ESRC/TLRP, grant reference RES-139–25–0258. http://www.tlrp.org/proj/Welch.html.

Welch GF and Howard DM (2002). Gendered voice in the cathedral choir. *Psychology of Music*, **30**, 102–120.

Welch GF, Ockelford A, Zimmermann S-A, Carter F-C and Himonides E (2006). Sounds of intent: initial mapping of musical behaviours and development in profoundly disabled children. In M Baroni, AR Addessi, R Caterina and M Costa, eds, *Proceedings, 9th International Conference on Music Perception and Cognition*, 274–275. Bologna University Press, University of Bologna, Bologna.

Welch GF, Ockelford A, Carter F-C, Zimmermann S-A and Himonides E (in press). Sounds of intent: mapping musical behaviour and development in children and young people with complex needs. *Psychology of Music*.

von Bertalanffy L (1968) *General system theory: foundations, development, applications*. George Braziller, New York.

Zillman D and Gan S (1997). Musical taste in adolescence. In DJ Hargreaves and AC North, eds, *The social psychology of music*, 161–187. Oxford University Press, Oxford.

# PART 7

# Musical performance

Edited by Richard Parncutt

CHAPTER 30

# Measurement and models of performance

W. Luke Windsor

## Introduction

### Rationale and overview

This chapter describes and evaluates the ways in which both the measurement and modelling of performances have contributed to the emerging discipline of music psychology. What is intended is a tutorial guide covering major issues, techniques and findings in this area, along with enough context and exemplification to meet this end. Therefore, only a representative sample of references to the literature will be made, and although it is hoped that the chapter might serve as a bibliographical starting point for the researcher, apologies are made in advance for any omissions of literature that might have been central to another's view of this field. Moreover, although the discriminability of measured changes in expressive parameters is an important issue for music psychology this chapter must restrict itself to consideration of measurements themselves, and direct comparisons between such measurements and the predictions made by models.

## Measuring performance

### Hardware for measuring performances

#### Mechanical devices for measurement

It is not until the invention of the Piano Camera and its application to a psychology of performance by Seashore and colleagues (see Seashore 1938) that an accurate manner of registering the action of piano keys and pedals in relation to both timing and velocity was obtained. Shaffer, who published a number of studies of music performance in the early 1980s (see e.g., Shaffer 1981, 1984; Shaffer *et al.* 1985), used a modified grand piano using photo-electric cells to register the movement of the hammers. The development of a number of commercial MIDI compatible grand and upright pianos which also utilize photocells has meant that specialist instrumentation is no longer needed for such research. Although the only study (Goebl and Bresin 2003) which addresses recording accuracy on recent Yamaha and Bosendorfer pianos finds considerable evidence for timing errors (especially for the Yamaha instrument) many of the errors are smaller than the errors associated with MIDI itself. However, if exceptionally fine-grained timing microstructure in performance is an issue for the researcher care must be taken to ensure that such recording inaccuracy is not, for example, mistaken for performance variance: some of the errors measured by Goebl and Bresin were five times larger than the errors one might expect from MIDI transmission of data.

#### Movement and physiological transduction

Recent published work on body movement in performance (and its control) has begun to demand and apply sophisticated techniques that measure and analyse bodily or instrumental movement (see e.g., Askenfelt 1986; Clarke and Davidson 1998; Wanderley and Vines 2006), muscular activity (see e.g. Fjellman-Wiklund *et al.* 2004a, b), or even electrical activity in the

central or peripheral nervous system (see e.g., Kelso *et al.* 1992).

### Acoustic instrumentation

Hardware for the acquisition and digitization of sound is available within most computer architectures, either off the shelf, or using third-party solutions. Given that analysing representations of the acoustic wave form of sounds allows for the analysis of timing, intensity, frequency and derivatives of these parameters (such as spectral envelope) is potentially of interest to anyone researching the psychology of performance who wishes to eschew keyboard instruments as data sources, and for the analysis of acoustic recordings of historical interest, such hardware is a vital resource. Moreover, although it may create other measurement issues, it avoids the potentially distorting influence of mechanical measurement.

## Software and techniques for measuring and analysing performances

The main purposes of computer software in this context are to (1) transform and record measurements into a particular numerical form, (2) convert between different data formats, (3) apply mathematical functions to deliver visual, numerical or statistical analyses of a data set. To illustrate different software solutions (and demonstrate how they are designed to integrate with hardware choices and particular research questions) two examples will be offered. Whilst not intended as comprehensive, these will give some indications of the range of issues faced when choosing measurement and analysis software.

The first example measures timing and dynamics in commercial recordings of piano performance (Repp 1998, 1999). The main research questions addressed here are what features of inter-onset timing and dynamic variation are shared between performances of the same piece, and how these vary between individual performers, groups of performers and whether sociocultural factors account for this variance. Repp's measurement approach here is fairly direct: the timing of onsets was obtained through the 'visual and auditory cues' provided by a commercial wave form editor. Where there were performance asynchronies (as in chordal events) the onset of the highest pitch was chosen (Repp 1998). Although a partly perceptual measure, this is in direct contrast to some musicological studies of tempo variability (such as Bowen 1996) which involve analysis of the researcher's finger taps at the beat or bar level (indeed there are no psychological studies which use this approach, perhaps because it may suffer from any number of known of unknown perceptuomotor biases or inaccuracies). Neither does Repp utilize automated or semi-automated onset detection software such as that developed by Moelants and Rampazzo (1997) or Dixon (2001), which other researchers have found useful when analysing timing in audio recordings (see e.g., Moelants 2000). For dynamics, measurements of peak sound level in dB were obtained using the automatic peak-picking function in the Signalyze program, using the prior analysis of timing as a guide. Repp then uses a combination of general spreadsheet and statistical software to organize, display and analyse these measurements, converting onset times into inter-onset intervals, and applying standard statistical techniques such as principle components analysis (to identify a small number of common traits), and correlation or analysis of variance (to test whether these vary with the chosen sociocultural variables).

The second example (Windsor *et al.* 2006) analyses inter-onset timing in a dataset of MIDI encoded piano performances, provided at a single sitting by an expert pianist. Here, the extraction of onset timings was achieved by a Yamaha Disklavier, which was connected to a computer running a commercial sequencer package. These MIDI files were converted into a textual format with time expressed in seconds, and were annotated with structural information (such as the beginnings and endings of bars and phrases at different hierarchical levels) in order that the onset timing could be modelled, using the POCO performance research environment (see Honing 1990). Statistical techniques were applied to fit the timing data optimally to the musical structure using a third-party add-on to POCO (SAPA) which shared the same LISP programming environment, and a commercial statistics package was used to display the results in tabular and graphical form.

Future developments in analysis techniques in the form of functional data analysis of time-varying parameters, which might include measurements of body movement and tempo, might offer useful opportunities to relate such continuous time-varying measures to each other (see e.g., Vines *et al.* 2005, 2006). Such techniques can be applied where sampling rates in the two domains to be compared are unrelated, or where one wishes to compare measurements taken from performances with different global tempi.

## Analytical methods and discoveries

Psychological studies of performance that derive from measurements exhibit a strong bias towards the measurement and analysis of onset timing. Dynamic variation has also been studied fairly extensively, but perhaps because of the relative ease with which onset times can be derived from performances, most studies are happy to focus on one or more aspects of timing. It is also the case that despite some attempts to model the relationship between parameters such as timing and dynamics (see e.g., Todd 1992) the relationship between the different variables that make up a performance is still little understood (although see Sloboda 1983; Drake and Palmer 1993).

## Contributions of measurement to a psychology of performance

### Motor programming

Music provides a rich context for studying the control of timing: the performer must often follow instructions from a score or match their movements to a co-performer (or ensemble), and the timing of their actions must be rhythmically organized to reflect the musical structure. The main observation in relation to timing control in music performance is that a serial timekeeper mechanism controls the production of fine-grained units in performance, whereas the timing of larger hierarchical units may be continuously modulated by a rate parameter (Shaffer *et al.* 1985; Vorberg and Wing 1996). These directly controlled units are often argued to be at the beat level in the metrical hierarchy: Shaffer (e.g. 1981, 1984) applied covariance analysis at different hierarchical levels (note, beat, bar) to attempt to identify negatively covarying intervals. Although such analyses normally rely on stationary time-series and may be distorted by drift or systematic tempo changes—which is often removed (Vorberg and Wing 1996, pp. 197–199) which are present in musical performances of this kind, Palmer (1997, p. 130) concludes that such results 'suggest that temporal precision in performance is influenced by the structure of the sequence—in particular, the salience of the beat level or tactus'.

The next issue that arises from this is how such a timing mechanism interacts with the actual events to be performed. The theory of motor programming (see e.g., Schmidt 1985; Gentner 1987) suggests that in a musical context an individual performance is derived from a mental representation of sequence of actions, often organized hierarchically: if one adds to this (as in Shaffer *et al.* 1985) the concept of a controlling internal timekeeper one has a recipe for generating an infinite number of performances from a single motor programme, with the timing of one hierarchical level directly controlled. Given that it is clear that performances by the same performer are surprisingly consistent in timing microstructure (see e.g., Gabrielsson 1987), even when recorded at different times (see e.g. Shaffer 1984; Clynes and Walker 1986), can show remarkable durational similarity, and that patterns of expressive timing reflect the hierarchical metrical and phrase structure of the music (see e.g. Shaffer 1981; Clarke 1988; Palmer 1989), it has become common to assume that for memorized performance at least, each performance given by an individual is related to a common motor programme, and that such motor programmes are structured in such a way as to reflect musical concerns. One theoretical question this raises is the extent to which such programmes are invariant given different overall rates (or global tempi). Deviations from proportional invariance are widely observed (Clarke 1982; MacKenzie and van Eerd 1990; Desain and Honing 1994; Windsor *et al.* 2001; Timmers *et al.* 2002; Repp *et al.* 2002; Windsor *et al.* 2006), and may reflect changes in perceptual or rhythmic grouping at different tempi or failures to transfer learning across tempi (Palmer 1997).

### Coordination: within and between performers

Although there is an unfortunate concentration on solo performance in the music–psychological literature, there is a growing body of empirical data which relates to the psychological processes governing within- and between-performer coordination. Evidence from studies of chord asynchrony (see e.g. Rasch 1988; Shaffer 1984; Palmer 1989) shows that melody lead is widespread, again providing evidence that timing in music performance is governed by structural understanding.

Shaffer (1984) argued that synchronization in expert ensemble performance is not achieved by coordinating actions at the level of individual events or even the lowest level of nominal isochrony (beat) or tactus, but at the level of the bar. His data certainly suggest that timing control, whether in a solo or ensemble situation, is most clearly observable at these intermediate structural levels. He (1984) argues that just as a single pianist may coordinate two separate timekeepers for each hand using a single motor programme, so too may two or more performers integrate their performance using such a shared plan. It is also further evidence that musical structures and expressive strategies are formally linked: being able to predict patterns of timing relies on a shared conception not only of musical structure but on the resulting interpretative decisions.

### Rhythm and timing

It has long been observed that musical performances of notated scores do not preserve their canonic durations (Seashore 1938), and there are a number of hypotheses that claim to explain this discrepancy. Some researchers explain such deviations from the metronome as expressive, either of underlying musical structures (Clarke 1988) or of extra-musical content (such as emotion or mood) (see e.g., Palmer 1989; Juslin and Madison 1999). Others attempt to show that many such deviations are the result of compensation for perceptual biases or due to motor noise (Penel and Drake 1998). Even simple rhythmic patterns are not performed metronomically, with adjacent intervals exhibiting assimilation (such that intervals are played more equal than they are notated) or contrast (such that similar intervals are made more different) and also context-dependent effects relative to the position of long and short intervals within a sequence and the number of events to be performed (Gabrielsson 1974; Repp *et al.* 2002). Interestingly, the latter of these two studies (which partly confirms earlier work by Fraisse (1956) shows that in a musical context, three note rhythms are rather unlike two note rhythms in that the patterns of assimilation change with tempo, and these changes bring increasing ratio-simplification of the two long-er intervals (and little change to the shorter interval).

### Expression, communication and signification

The deviations from nominal score durations (and from small whole number ratios) evident in measurements of performances, along with patterns of dynamics, timbre and vibrato, have suggested that performers may be using such means to express, communicate or signify either intra-musical or extra-musical content. Such an approach is best exemplified by Palmer (1989) and the work of Shaffer, Clarke and Todd (see above). Within this perspective it is possible to find in the literature copious examples of measurements which suggest rule-following expressive behaviour. Expressive timing and dynamics seem to be linked to phrase and metrical structure in predictable ways: many studies (see Clarke 1988; Palmer 1997; Clarke 1999; Gabrielsson 1999 for reviews) have argued on the basis of empirical evidence that the location, and extent of deviations from metronomic performance can be predicted from analyses of phrase structure, metrical structure and voice prominence: hence, it has become accepted that such features may be generative in nature, mapping structural features through a system of rules onto patterns of articulation, rubato, and the like.

Such a generalizing tendency begs the question of the extent to which performances differ, and the underlying reasons for these differences. It is clear that different structural interpretations or intended moods or emotions leads to different performance outcomes, and this is observed in work on emotion (Juslin 2000) as it

is in work which varies structural interpretation (such as Clarke 1988; Palmer 1989; Clarke and Windsor 2000). This is precisely what a generative rule-based explanation of expression in performance would predict. However, large-scale comparison of multiple performances of the same music (e.g., Repp 1998, 1999) suggests that even where performers may share a structural interpretation (and leaving aside differences in intended characterization) there may be significant differences between performers which are as yet hard to explain in any formal manner. Using principal components analysis, Repp here identifies traits in interpretation that suggest different strategies in relation to the same musical structure. It is clear that here is another way in which performers can choose to differ from one another. Precisely how such choices are made is little understood, although work which looks at the development of performances in a more longitudinal manner (such as Chaffin and Imreh 2002) might help deliver some answers.

# Algorithmic and statistical modelling of performance

## Hypothesis testing and simulation

Models of performance fall into two main categories by motivation: either they seek to test a hypothesis about behaviour, or they seek to simulate that behaviour. As Desain *et al.* (1998) argue, it is possible to fit the output of models with very different assumptions to the same data, and it may be necessary to pay close attention to the architecture of the individual models to judge whether they can be distinguished in terms of their psychological veracity. Moreover, as Honing (2006) demonstrates persuasively, models that fit data well do not always tell us more about the behaviour modelled: they may fit trivial aspects of the data precisely because they do not account for important features in the data which are hard to model. Honing gives two additional criteria to goodness of fit: *simplicity* and *surprise*, arguing that whereas some models will over-generalize to all sorts of behaviour, some will only fit a narrow range of behaviour and are therefore a priori better as they are more falsifiable.

## Data-driven and theory-driven simulation

Whereas Gabrielsson (1999, pp. 550–557) distinguishes between models based on measurements and models based in intuitions, it is perhaps more useful here to propose a methodological distinction between data-driven and theory-driven simulations of human data (or a continuum, given that models nay combine the two approaches). Although the one example he gives of a model based on measurements (Todd 1985, 1989, 1992, 1995; see also Windsor and Clarke 1997; Clarke and Windsor 2000) was partly based on measurements it is more accurate to suggest that this is a theory-driven model which was then compared and refined through comparison with measurements of human performances. The model is actually driven by an extremely rigid application of a generative understanding of music performance. Similarly theoretically driven, although based on a rather loose set of intuitions (based largely on informal empirical investigation characterized as 'analysis by synthesis') is the model developed by Friberg and others (see e.g. Friberg 1991; Sundberg 1988).

The alternative to starting from rule-like algorithmic approaches which are then tested against data is to take measurements and to attempt to fit them to hypotheses using statistical means. This is the approach taken by Palmer (e.g. 1989) where hypotheses about the effects of musical structure and performance intentions on rubato, asynchrony and dynamics are tested using standard linear modelling techniques familiar to experimental psychology. Such approaches (based as they are on small numbers of performers and one or two pieces) may present problems of generalization, and do not require the researcher to make specific predictions about the precise effect that, for example, voicing has on asynchrony.

## Performance modelling: psychological relevance

In this final subsection on modelling an attempt will be made to evaluate the extent to which such models not only capture general principles of performance, but also whether their architecture in any way represents psychologically plausible

processes (see Desain *et al.* 1998). To this end, four examples will be reviewed briefly. Unfortunately, space does not allow for detailed consideration of the performance modelling approaches taken by Clynes (Clynes 1983) nor that of Mazzola and colleagues (Mazzola 2002): both share some features with one or more of the approaches described below, and are important contributions in their own right.

The first model considered here is most commonly know as *Director Musices*, and has been described in detail in many publications (see e.g. Sundberg 1988; Friberg 1991). Interestingly, just as Todd's approach (e.g., 1992 and see above) both benefits from and is limited by its theoretical simplicity, its ad hoc construction is probably closer to psychological reality than many more systematic models. It is clear that performers combine many different expressive strategies in a performance (see e.g., Palmer, 1989) and that these are often independent (Repp 1999). Nonetheless, such diverse rules make evaluation of the model as a whole, or its independent parameters, extremely tricky, although it has been fit to real performance data with some success (see e.g. Sundberg *et al.* 2003; Zanon and de Poli 2003a, b; Widmer and Goebl 2004). The recent attempt to complement the generative rules in *Director Musices* with models of various types of noise, emotional intent and physical motion (Juslin *et al.* 2002) may indeed make this problem even more pronounced (Widmer and Goebl 2004, p. 206).

Although there have been some successful attempts to fit the different versions of Todd's generative model to actual data (e.g. Todd 1989), 1992) and other more equivocal results (Windsor and Clarke 1997; Clarke and Windsor 2000; see also Widmer and Goebl 2004), Todd's admission in later work that two different models of tempo change might be required for the model to fit different rubato depths (Todd 1995 rather weakens this model's elegance and explanatory power. The model, especially in its earliest form, is an extremely powerful example of how to translate psychologically plausible theories into a coherent algorithm. This model takes a small number of assumptions (about the representation of musical structure and the mapping of structure to expression) which relate well to empirical and theoretical ideas of the time.

Widmer and Tobudic (2003) combine some aspects of curve-fitting, which are similar to Todd's approach, to machine-learning techniques. Widmer and Goebl (2004) suggest that this is a way of combining 'local timing and dynamic deviations' with tempo and dynamic curves associated with phrasing. Such an approach is elegant and clearly effective in terms of fitting real data, accounting for these different levels and types of expression, and doing so without excessive numbers of parameters. Instead it adapts to a training set of performances using fitting and machine-learning. Both of these features make this questionable as a psychologically plausible model, although the rules that it discovers may reflect or describe the outcomes of underlying psychological processes. Hence, this model reflects its background in artificial intelligence (AI), where the outcomes may offer more in terms of the development of machine simulation than they do to the realm of psychology.

Such potential problems are also evident for the final and most recent model to be considered, described in Windsor *et al.* (2006). However, here it is an analytical bias that is evident. The overall degree of fit achieved is high, but the authors focus on the way in which the fitting technique decomposes a complex time-series into different components, and how this can be used to answer empirical questions related to proportional invariance over tempo. The model itself might have some degree of psychological plausibility, but only if the components have an independence in the motor or memory systems of the performer, and it is not at present clear how this might be assessed. In conclusion, this approach, although it looks on the surface to resemble that of Todd, is much closer to the approaches taken in standard statistical testing in experimental psychology.

## Measurement and modelling of performance in context

Psychological work on performance such as that described above has had little impact on research and practice within the field of music. Where measurement of performance is evident there seems to be little interest in or account taken of psychological findings. Repp's work on diversity

and commonality (1998, 1999) is hardly known outside the music–psychological community, despite its potential musicological importance), different measurement techniques are employed (as in Bowen 1996), and modelling of performance often seems irrelevant to either musicological or pedagogical research. In some ways this is unsurprising, as music psychology comes with many concerns foreign to the wider study or practice of music. However, there are some indications that a psychological approach to measuring and modelling performance might offer insights and techniques of wider value.

For example, a number of recent projects involving measurement and or modelling have gained funding or produced results which are relevant to, or aimed at, performance pedagogy. In Uppsala, Juslin and colleagues have attempted to use empirical and modelling techniques as part of a project aiming to develop methods of teaching expressivity (see e.g., Juslin *et al.* 2004). Similarly, Desain and colleagues are currently working on a project which both utilizes existing understanding of expression and real-time visualization techniques for learning expression (Brandmeyer *et al.* 2006). Similar goals underlie work by Goebl and Widmer (2006). In a rather different way, research in the development of expert performance has begun to adopt the empirical measurement techniques described here: for example Chaffin and Imreh (2002) have addressed issues of learning and memory in performance through detailed analysis of recorded performances.

Measurement and modelling have provided the psychology of music with substantial insights into the processes underlying music performance. One can safely assume that these insights will be extended given the attention paid to this area by researchers. However, it is disappointing to note that work on expression in music, and one might take as an example that on tempo rubato, is being simultaneously studied both by musicologists and psychologists, without much reference to each others' work. Although the goals of psychology and musicology may be different, both seek to explain and describe musical behaviours in their different ways, and one might hope for more cross-fertilization of methodologies and knowledge. This might simply extend to seeing more reference to historical and cultural context in work on performance within a psychological context (such as in Repp 1998, 1999; Windsor *et al.* 2001) or more methodological exactitude in work carried out in a musicological context (although see Bowen 1996; Fabian 2003). More than this, however it calls for researcher collaborations between these two disciplines (see e.g., Fabian and Schubert 2003) or indeed researchers that are sufficiently experienced in both fields.

## References


Askenfelt A (1986). Measurement of bow motion and bow force in violin playing. *Journal of the Acoustical Society of America*, **80**, 1007–1015.

Bowen Jose A (1996). Tempo, duration and flexibility: techniques in the analysis of performance. *Journal of Musicological Research*, **16**(2), 1–47.

Brandmeyer A, Hoppe D, Sadakata M, Timmers R and Desain P (2006) Practice space: a platform for real-time visual feedback in music instruction. *Proceedings of the 9th International Conference on Music Perception and Cognition* (ICMPC9), Bologna, Italy.

Chaffin R and Imreh G (2002). Practicing perfection: piano performance as expert memory. *Psychological Science*, **13**(4), 342–349.

Clarke EF (1982). *Timing in the performance of Erik Satie's Vexations. Acta Psychologica*, **50**, 1–19.

Clarke EF (1988). Generative principles in music performance. IN JA Sloboda, ed., *Generative processes in music: the psychology of performance, improvisation, and composition*, 1–26. Clarendon Press, Oxford.

Clarke EF (1999). Rhythm and timing in music. In D Deutsch, ed., *The psychology of music*, 473–500. Academic Press, New York.

Clarke EF and Davidson JW (1998). The body in performance. In W Thomas, ed., *Composition–performance–reception. Studies in the creative process in music*, 74–92. Ashgate Press, Aldershot.

Clarke EF and Windsor WL (2000). Real and simulated expression: a listening study. *Music Perception*, **17**(3), 1–37.

Clynes M (1983). Expressive microstructure in music, linked to living qualities. In J Sundberg, ed., *Studies of music performance*, 76–181. Royal Swedish Academy of Music, Stockholm.

Clynes M and Walker J (1986). Music as time's measure. *Music Perception*, **4**, 85–119.

Desain P and Honing H. (1994). Does expressive timing in music performance scale proportionally with tempo? *Psychological Research*, **56**, 285–292.

Desain P, Honing H, Van Thienen H and Windsor L (1998). Computational modeling of music cognition: problem or solution? *Music Perception*, **16**(1), 151–166.

Dixon S (2001). Automatic extraction of tempo and beat from expressive performances. *Journal of New Music Research*, **30**(1), 39–58.

Drake C and Palmer C (1993). Accent structures in music performance. *Music Perception,* **10**(3), 343–378.

Fabian D (2003). *Bach performance practice 1945–1975: A comprehensive review of sound recordings and literature.* Ashgate, Aldershot.

Fabian D and Schubert E (2003). Expressive devices and perceived musical character in 34 performances of Variation 7 from Bach's Goldberg Variations. *Musicae Scientiae,* Special Issue 2003–2004, 49–68.

Fjellman-Wiklund A, Grip H, Karlsson JS and Sundelin G (2004a). *EMG trapezius muscle activity pattern in string players: Part I—is there variability in the playing technique? International Journal of Industrial Ergonomics,* **33**, 347–356.

Fjellman-Wiklund A, Grip H, Karlsson JS and Sundelin G (2004b). *EMG trapezius muscle activity pattern in string players: Part II—influences of basic body awareness therapy on the violin playing technique. International Journal of Industrial Ergonomics,* **33**, 357–367.

Fraisse P (1956). *Les structures rhythmiques.* Editions Universitaires, Louvain.

Friberg A (1991). Generative rules for music performance: a formal description of a rule system. *Computer Music Journal,* **15**(2), 56–71.

Gabrielsson A (1974). Performance of rhythm patterns. *Scandinavian Journal of Psychology,* **15**, 63–72.

Gabrielsson A (1987). Once again: the theme from Mozart's piano Sonata in A Major (k. 331). In A Gabrielsson, ed., *Action and perception in rhythm and music,* 81–104. Royal Swedish Academy of Music, Stockholm.

Gabrielsson A (1999). The performance of music. In D Deutsch, ed., *The psychology of music,* 501–602. Academic Press, New York.

Gentner DR (1987). Timing of skilled motor performance: tests of the proportional duration model. *Psychological Review,* **94**(2), 255–276.

Goebl W and Bresin R (2003). Measurement and reproduction accuracy of computer-controlled grand pianos. *Journal of the Acoustical Society of America,* **114**(4), 2273–228.

Goebl W and Widmer G (2006). Unobtrusive practice tools for pianists. *In Proceedings of the 9th International Conference on Music Perception and Cognition* (ICMPC9), 209–214. Bologna.

Honing H (1990). POCO: an environment for analysing, modifying, and generating expression in music. In *Proceedings of the 1990 International Computer Music Conference,* 364–368. International Computer Music Association, San Francisco,CA.

Honing H (2006). Computational modeling of music cognition: a case study on model selection. *Music Perception,* **23**(5), 365–376.

Juslin PN (2000). Cue utilization in communication of emotion in music performance: relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance,* **26**, 1797–1813.

Juslin PN and Madison G (1999). The role of timing patterns in recognition of emotional expression from musical performance. *Music Perception,* **17**(2), 197–221.

Juslin PN, Friberg A and Bresin R (2002). Toward a computational model of expression in performance: the GERM model. *Musicae Scientiae,* Special Issue 2001–2002, 63–122.

Juslin PN, Friberg A, Schoonderwaldt E and Karlsson J (2004). Feedback learning of musical expressivity. In A Williamon, ed., *Enhancing musical performance: a resource for performers, teachers, and researchers,* 247–270. Oxford University Press, New York.

Kelso JAS, Bressler SL, Buchanan S, Deguzman GC, Ding M, Fuchs A and Holroyd T (1992). A phase transition in human brain and behavior. *Physics Letters A,* **169**(3), 134–144.

MacKenzie CL and Van Eerd DL, (1990). Rhythmic precision in the performance of piano scales: motor psychophysics and motor programming. In M. Jeannerod, ed., Attention and performance XIII, 375–408. Lawrence Erlbaum Associates, Hillsdale, NJ.

Mazzola G (2002). *The topos of music, geometric logic of concepts, theory, and performance.* Birkhäuser, Basel.

Moelants D (2000). Statistical analysis of written and performed music. A study of compositional principles and problems of coordination and expression in punctual serial music. *Journal of New Music Research,* **29**(1), 37–60.

Moelants D and Rampazzo C (1997). A computer system for the automatic detection of perceptual onsets in a musical signal. In A. Camurri, ed., *KANSEI—the technology of emotion,* 141–146. AIMI-DIST, Genovo.

Palmer C (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance,* **15**(12), 331–346.

Palmer C (1997). Music performance. *Annual Review of Psychology,* **48**, 115–138.

Penel A and Drake C (1998). Sources of timing variations in music performance: a psychological segmentation model. *Psychological Research,* **61**, 12–32.

Rasch RA (1988). Timing and synchronisation in ensemble performance. In JA Sloboda, ed., *Generative processes in music,* 70–90. Oxford University Press, Oxford.

Repp BH (1998). A microcosm of musical expression: I. Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America,* **104**, 1085–1100.

Repp BH (1999). A microcosm of musical expression: II. Quantitative analysis of pianists' dynamics in the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America,* **105**(3), 1972–1988.

Repp BH Windsor WL, Desain PWM., (2002). Effects of tempo and timing of simple musical rhythms. *Music Perception,* **19**(4), 565–594.

Schmidt RA (1985). The search for invariance in skilled movement behavior. *Research Quarterly for Exercise and Sport,* **56**(2), 188–200.

Seashore CE (1938). *Psychology of music.* McGraw-Hill, New York.

Shaffer LH (1981). Performances of Chopin, Bach and Bartok: studies in motor programming. *Cognitive Psychology,* **13**, 326–376.

Shaffer LH (1984). Timing in solo and duet piano performances. *The Quarterly Journal of Experimental Psychology*, **36A**, 577–595.

Shaffer LH, Clarke EF and Todd NPM (1985). Metre and rhythm in piano playing. *Cognition*, **20**, 61–77.

Sloboda JA (1983) The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology A*, **35**, 377–396.

Sundberg J (1988). Computer synthesis of music performance. In JA Sloboda, ed., *Generative processes in music: the psychology of performance, improvisation, and composition*, 52–69. Clarendon Press, Oxford.

Sundberg J, Friberg A and Bresin R (2003). Attempts to reproduce a pianist's expressive timing with director musices performance rules. *Journal of New Music Research*, **32**(3), 317–325.

Timmers R, Ashley R, Desain PWM, Honing HJ and Windsor WL (2002). Timing of ornaments in the theme of Beethoven's Paisiello variations: empirical data and a model. *Music Perception*, **20**(1), 3–33.

Todd NP (1985). A model of expressive timing in tonal music. *Music Perception*, **3**, 33–58.

Todd NP (1989). Towards a cognitive theory of expression: the performance and perception of rubato. *Contemporary Music Review*, **4**(1), 405–416.

Todd NP (1992). The dynamics of dynamics: a model of musical expression. *Journal of the Acoustical Society of America*, **91**, 3540–3550.

Todd NP McA. (1995). The kinematics of musical expression. *Journal of the Acoustical Society of America*, **97**, 1940–1949.

Vines BW, Krumhansl CL, Wanderley MM and Levitin DJ (2006). *Cross-modal interactions in the perception of musical performance. Cognition*, **101**, 80–113.

Vines BW, Nuzzo RL and Levitin DJ (2005). Analyzing temporal dynamics in music: differential calculus, physics, and functional data techniques. *Music Perception*, **23**, 139–154.

Vorberg D and Wing AM (1996). Modeling variability and dependence in timing. In H Heuer and SW Keele, eds, *Handbook of perception and action II: motor skills*, 181–262. Academic Press, London.

Wanderley MM and Vines B (2006). Ancillary gestures of clarinettists. In A Gritten and E King, eds, *Music and gesture: new perspectives on theory and contemporary practice*, Ashgate Press, Aldershot.

Widmer G and Goebl W (2004). Computational models of expressive music performance: the state of the art. *Journal of New Music Research*, **33**(3), 203–216.

Widmer G and Tobudic A (2003). Playing Mozart by analogy: learning multi-level timing and dynamics strategies. *Journal of New Music Research*, **32**(3), 259–268.

Windsor WL and Clarke EF (1997). Expressive timing and dynamics in real and artificial musical performances: using an algorithm as an analytical tool. *Music Perception*, **15**(2), 127–152.

Windsor WL, Aarts R, Desain P, Heijink H and Timmers R (2001) The timing of grace notes in skilled musical performance at different tempi: a preliminary case study. *Psychology of Music*, **29**, 149–169.

Windsor WL, Desain PWM, Penel A and Borkent M (2006). A structurally guided method for the decomposition of expression in music performance. *Journal of the Acoustical Society of America*, **119**(2), 1182–1193.

Zanon P and De Poli G (2003a). Time-varying estimation of parameters in rule systems for music performance. *Journal of New Music Research*, **32**(3), 295–316.

Zanon P and De Poli G (2003b). Estimation of parameters in rule systems for expressive rendering in musical performance. *Computer Music Journal*, **27**(1), 29–46.

# CHAPTER 31

# Planning and performance

Eckart Altenmüller and Sabine Schneider

## Introduction

Performing music at a professional level is probably the most demanding of human accomplishments. Making music requires the integration of multimodal sensory and motor information and precise monitoring of the performance via auditory feedback. In the context of Western classical music, musicians have to reproduce highly controlled movements almost perfectly with a high reliability. These specialized sensorimotor skills require extensive training periods over many years, starting in early infancy and passing through stages of increasing physical and strategic complexities. The superior skills of musicians are mirrored in plastic adaptations of the brain on different timescales.

In the first section of this chapter we introduce essential general information for musical readers concerning the organization of cortical, subcortical and cerebellar motor systems in the brain. The electrophysiological correlates of motor planning and motor expectation will be briefly mentioned, as they provide a deeper understanding of the time course of anticipation and retrieval of motor programmes in music performance. In the second section, brain processes during acquisition of skilled movements in music-making will be addressed and the dynamics of neuronal networks will be demonstrated. Because these processes rely on plastic adaptations of brain networks and anatomical brain structures, the interplay of increasing precision of movements and plastic changes in the brain will be explained. In the third section, new findings on practice strategies and performance quality will be reported. Brain imaging measures collected during mental practice or listening tasks suggest that both motor and auditory cortical areas are active during musical thought processes. Motor-based brain representations are found in behavioural studies on performers' musical interpretations, transfer of learning from one musical task to another, mental practice effects, and anticipatory movements. Implications from these behavioural tasks suggest that an accurate auditory and motor representation underlies successful performance from memory. In the fourth section, the causes of degradation of skilled movements in professional musicians will be addressed. This disorder, termed focal dystonia, is due to maladaptive brain plasticity with fusion of brain representations of adjacent digits in somatosensory brain regions. Such a fusion and blurring of receptive fields of the digits results in a loss of control, because skilled motor actions are necessarily bound to intact somatosensory feedback input. Prolonged practice and pain syndromes due to overuse can precipitate dystonia, which is developed by about 1 per cent of professional musicians and usually ends their career. Finally, we will end the chapter with an outlook and add some comments concerning the significance of results of brain research in order to improve practice habits and performance in musicians.

## Introduction

There can be no doubt that making music is one of the most demanding tasks for the human central nervous system. It involves the precise

execution of very fast and, in many instances, extremely complex physical movements that must be coordinated with continuous auditory feedback. Practice is required to develop these skills and carry out these complex tasks. Perhaps the most important study on practice to emerge during the past couple of decades was undertaken by Ericsson and his colleagues in 1993 with students at the Berlin Academy of Music. Ericsson *et al.* proposed the concept of 'deliberate practice' as a means of studying goal-oriented, structured and effortful facets of practice in which motivation, resources and attention determine the amount and quality of practice undertaken. They argued that a major distinction between professional and amateur musicians (and perhaps successful versus un-successful learners) is the amount of deliberate practice undertaken during the many years required to develop instrumental skills to a high level (Ericsson and Lehmann 1996). They proposed that highly skilled musicians exert a great deal more effort and concentration during their practice than less skilled musicians, and are more likely to plan, image, monitor and control their playing by focusing their attention on what they are practising and how it can be improved.

Motor skills are best acquired by massed practice involving countless repetitions, whereas aural skills are typically refined through a broad variety of listening experiences. Both types of skills are not represented in isolated brain areas however, but rather depend on the multiple connections and interactions established during training within and between the different regions of the brain. The general ability of our central nervous system to adapt to changing environmental conditions and newly imposed tasks during its entire life span is referred to as *plasticity*. In music, planning, learning through experience and training are accompanied by development and changes which not only take place in the brain's neuronal networks as a result of a strengthening of neuronal connections but also in its overall gross structure. Unfortunately, it is still not completely understood how practice habits and sensorimotor maturation influence each other. With respect to brain plasticity it is known that music practice enhances myelination, grey matter growth and fibre formation of brain structures involved in the specific musical task (for a review see Münte *et al.* 2002).

There are two main reasons why researchers believe that these effects on brain plasticity are more pronounced in instrumental music performers than in other skilled activities. First, musical training usually starts very early, sometimes before age six when the adaptability of the central nervous system is highest, and second, musical activities are strongly linked to positive emotions, which are known to enhance plastic adaptations. We would be wise to keep in mind however, that the methodologies currently used in contemporary brain research might produce a bias. As an example, it could be argued that the results demonstrated for group investigations of classical instrumentalists are due to these musicians having a similar acculturation due to the canonical nature of their training. Classical pianists tend to study etudes of Hanon, Czerny and Chopin, and the similarity of their training may produce uniform brain adaptations which in turn then dominate any individual changes. In other pursuits such as the visual arts, creative writing, architecture and composing music, individualized training may produce more diverse effects that may be masked within group statistics.

## Neuroanatomy and neurophysiology of motor systems involved in planning and performance

Playing a musical instrument requires highly refined motor skills that are acquired over many years of extensive training, and that have to be stored and maintained as a result of further regular practice. Auditory feedback is needed to improve and perfect performance. Performance-based music-making therefore, relies primarily on a highly developed auditory–motor integration capacity, which can be compared to the phonological loop in speech production. In addition, somatosensory feedback constitutes another basis of high-level performance. Here, the kinaesthetic sense, which allows for control and feedback of muscle- and tendon-tension as well as joint positions which enable continuous monitoring of finger-, hand- or lip-position in

the frames of body and instrument coordinates (e.g., the keyboard, the mouthpiece), is especially important. In a more general context, the motor system of music performance can be understood as a subspecialty of the motor systems for planned and skilled voluntary limb movements.

Planned voluntary skilled limb movements involve four cortical regions in both hemispheres: the *primary motor area* (M1) located in the *precentral gyrus* directly in front of the central sulcus; the *supplementary motor area* (SMA) located anterior to the M1 of the frontal lobe and the inner (medial) side of the cortex; the *cingulate motor area* (CMA) below the SMA and above the corpus callosum on the inner (medial) side of the hemisphere; and the *premotor area* (PMA), which is located adjacent to the lateral aspect of the primary motor area (see Figure 31.1).

SMA, CMA and PMA can be described as *secondary motor areas*, because they are used to process movement patterns rather than simple movements. In addition to the cortical regions, the motor system includes the subcortical structures of the basal ganglia, and the cerebellum. The sensory areas are necessary in order to maintain the control of movements. Their steady kinaesthetic feedback information is required for any guided motor action. The sensory areas are located in the *primary somatosensory area* (S1) behind the central sulcus in the parietal lobe. This lobe is involved in many aspects of movement processing. It is an area where information from multiple sensory regions converges. In the posterior parietal area, the body coordinates in space are monitored and calculated and visual information is transferred into body coordinates. As far as musicians are concerned, this area is prominently activated during tasks involving multisensory integration, for example during sight-reading and the playing of complex pieces of music (Haslinger *et al.* 2005).

The *primary motor area* (M1) represents the movements of body parts in a separate but systematic order. The representation of the leg is located on the top and the inner side of the hemisphere, the arm in the upper portion, and the hand and mouth in the lower portion of M1. This representation of distinct body parts in corresponding brain regions is called *somatotopic* or *homuncular order*. Just as the *motor homunculus* is represented upside down, so too is the *sensory homunculus* on the other side of the central sulcus. The proportions of both—the motor

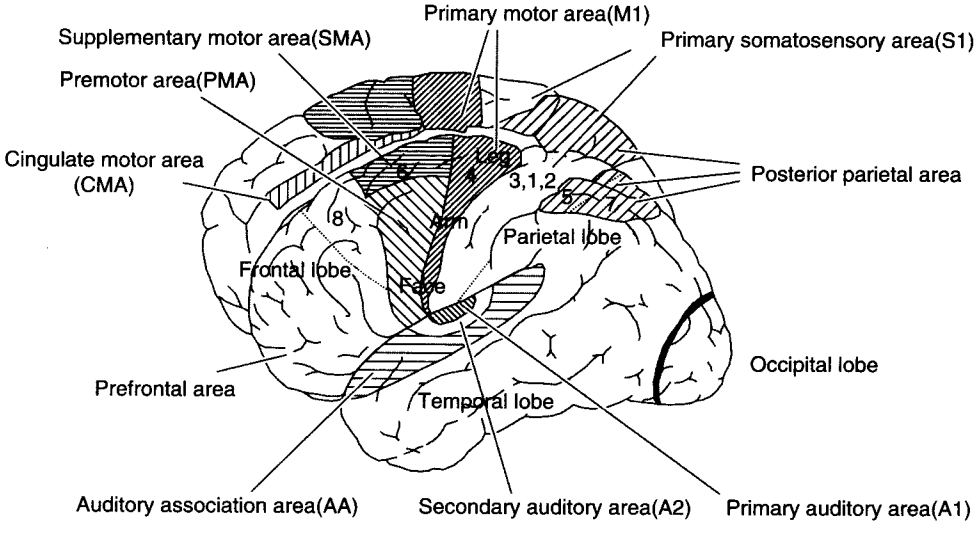



**Fig. 31.1** Brain regions involved in sensory and motor music processing. The left hemisphere is shown in the forground (lower right); the right hemisphere in the background (upper left). The numbers relate to the respective Brodmann's areas, a labelling of the cortical areas according to the fine structure of the nervous tissue.

and the sensory homunculus—are markedly distorted because they are determined by the density of motor and sensory innervations of the respective body parts. For example, control of fine movements of the tongue requires many more nerve fibres transmitting the information to this muscle as compared to the muscles in the back. Therefore, the hand, the lips and the tongue require almost two-thirds of the neurons in this area. However, as further explained below, the representation of the body parts may be modified by usage. Moreover, the primary motor area does not simply represent individual muscles: multiple muscular representations are arranged in a complex way so as to allow the execution of simple types of movements rather than the activation of a specific muscle. This is a consequence of the fact that a two-dimensional array of neurons in M1 has to code for three-dimensional movements in space (Gentner and Classen 2006). Put more simply, our brain does not represent muscles but rather movements.

The *supplementary motor area* (SMA) is mainly involved in the coordination of the two hands, in the sequencing of complex movements and in the triggering of movements based on internal cues. It is particularly engaged when the execution of a sequential movement depends on internally stored and memorized information. The SMA can be subdivided into two distinct functional areas. In the anterior SMA, it would seem that the planning of complex movement patterns is processed. The posterior SMA seems to be predominantly engaged in two-handed movements and, in particular, in the synchronization of both hands during complex movement patterns.

The function of the *cingulate motor area* (CMA) is still under debate. Electrical stimulation and brain imaging studies demonstrate its involvement in movement selection in situations when movements are critical to obtain reward or punishment. This points towards close links between the cingulate gyrus and the emotion-processing limbic system. From what we know therefore, it would seem that the CMA plays an important role in mediating cortical cognitive functions and limbic–emotional functions. The *premotor area* (PMA) is primarily engaged when externally stimulated behaviour is being planned and prepared. It is involved in the learning, execution and recognition of limb movements and seems to be particularly concerned with processing of visual information which is necessary for movement planning.

The *basal ganglia*, located deep inside the cerebral hemispheres, are interconnected reciprocally via the thalamus to the motor and sensory cortices, thus constituting a loop of information flow between the cortex and the basal ganglia. They are indispensable for any kind of voluntary actions that are not highly automated. Their special role consists in the control of voluntary action by selecting appropriate motor actions and by comparing the goal and course of those actions with previous experience. In the basal ganglia, the flow of information between the cortex and the limbic emotion system, in particular the amygdala, converges. It is therefore assumed that the basal ganglia process and control the emotional evaluation of motor behaviour in terms of expected rewards or punishment. Finally, the *cerebellum* contributes essentially to the timing and accuracy of fine-tuned movements.

## Observing planning in the brain

During the last two decades, knowledge of brain regions involved in complex tasks such as playing a musical instrument has increased enormously. This is mainly due to the development of novel technologies that allow non-invasive assessment of the intact brain's function. A pioneering step was the observation of movement-related brain potentials reflecting planning and movement preparation in the brain. These brain potentials can be extracted from the ongoing electrical activity of neuronal populations in the cerebral cortex using electroencephalography (EEG). The most prominent activation is the so callled Bereitschaftspotential (see Figure 31.2).

The Bereitschaftspotential (BP) is a ramp-like brain activation which precedes any self-paced voluntary motor activity, starting 2000 to 1000 ms prior to movement onset. There is still some debate on the structures in the brain generating the BP. It seems that the first part of the ramp-like shift is produced in the SMA, reflecting the planning of a movement. The subsequent part of the shift is probably generated in the primary

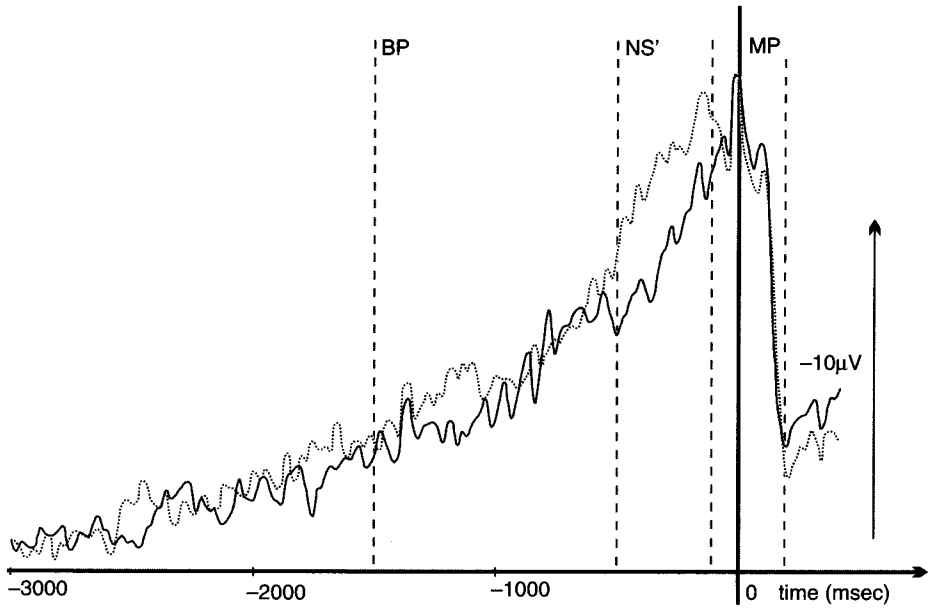



**Fig. 31.2** Typical Bereitschaftspotential (BP) in healthy pianists and in pianists suffering from focal dystonia (dotted line) prior to scale playing (Time 0). Brain activation starts about 3 seconds prior to movement onset, reflecting planning and preparation. The proper BP starts 1.5 prior to movement onset and is generated in the supplementary motor cortex. The negative slope (NS') 500 ms prior to movement onset is generated in the anterior part of the primary motor cortex. The motor potential (MP) finally corresponds to activation of the motor neurons in the pyramidal tract. After movement onset, brain activation is reset, yielding a sudden positive shift in brain activation. Note that pianists suffering from pianists' dystonia have larger amplitudes especially in the NS' component of the Bereitschaftspotential (modified from Peschel and Altenmüller 2004).

motor areas, reflecting the activation of motor neurons directly linked to muscles via the spinal cord (for a concise review see Altenmüller *et al.* 2004).

Traditionally, the BP was related to the intentional decision processes of willed action. In intriguing experiments Libet *et al.* (1983) demonstrated that the BP starts about 350 ms prior to conscious awareness of the intention to act. When the subjects 'vetoed' their decision to act, the BP, which had normally developed prior to this 'veto', collapsed and no movement occurred. Libet (1985) concluded that voluntary acts can be initiated by unconscious cerebral processes before conscious intention appears but that conscious control over the actual motor performance of the acts remain possible. This experimental design has had a long and often controversial history. After all, it has remained unclear whether the urge to act, and the action itself, represent actual differences in brain states (Eagleman 2004).

EEG measures such as the BP-recordings have an excellent temporal resolution, reflecting the electrical activity of neurons in the range of milliseconds. Therefore these methods are suitable to investigate the rapid neuronal interactions which constitute the basis of motor planning and performance.

Other imaging tools such as positron emission tomography (PET), and functional magnetic resonance imaging (fMRI) allow the functional assessment and the precise localization of active brain regions. However, these methods have the disadvantage of a relatively poor time resolution, allowing the monitoring of neuronal activation during planning and performance in the range of seconds, but not of milliseconds.

Additionally, new imaging techniques, derived from MRI) technology, demonstrate minute changes in brain structure with precision. Voxel-based morphometry (VBM) for example, provides detailed information of the thickness of the grey matter in the layers of neurons in the cerebral cortex. Using this technique, cross-sectional studies have demonstrated differences in grey matter volume in the range of cubic millimeters as a result of musical training (Gaser and Schlaug 2003). Diffusion tensor imaging (DTI), on the other hand, is a way to assess direction and volume of fibre tracts in the white matter of the brain. In pianists this method has shown changes in myelination of the callosal body which connects the two brain hemispheres (Bengtsson *et al.* 2005).

## Learning to plan: the acquisition of fine motor skills

Our knowledge concerning the regions and mechanisms of the brain involved in sensorimotor learning is still incomplete. Overall, musicians appear to process new incoming stimuli more effectively compared to non-musicians. According to newly emerging evidence (for a review see Halsband and Lange 2006) all structures involved in motor control participate in the acquisition of new sensorimotor skills. The cerebellum is involved in the selection, the sequence and the timing of movements and the basal ganglia play a crucial role in procedural learning and automation of movements.

It has been known for some time that the activity in the SMA and in the premotor area of the brain are enhanced as a result of increasing complexity of finger movement sequences (Roland *et al.* 1980). Using fMRI, Karni and colleagues (1995) investigated adult subjects' learning of complex finger sequences which are similar to those necessary for piano playing. After 30 minutes of practice the representation of the fingers in the primary motor cortex increased. However, without further training, this effect diminished after one week with the hand representation returning to its previous size. In contrast, continuous practice resulted in a stable enlargement of the hand area in the primary motor cortex. This effect was specific for the daily trained sequence of complex finger movements, and did not occur when the subjects improvised complex finger movements which were not subsequently repeated. Parallel to the enlargement of the hand area in the primary motor cortex, the size of the cerebellar hand representation diminished, suggesting that the cerebellum plays an important role only in the initial phase of motor learning.

The above mentioned study does not take into account one special quality of musicianship, namely the strong coupling of sensorimotor and auditory processing required for performing music. Practicing an instrument involves assembling, storing, and constantly improving complex sensorimotor programmes through prolonged and repeated execution of motor patterns under the controlled monitoring of the auditory system. In a cross-sectional experiment, strong linkages between auditory and sensorimotor cortical regions as a result of many years of practice have been reported (Bangert *et al.* 2006). Using fMRI, professional pianists were asked to listen to simple piano tunes without moving their fingers or any other body part. Figure 31.3(a) demonstrates the increase in activation of professional pianists in comparison to non-musicians. There is an impressive activation of the motor cortex demonstrating the subconscious or automated auditory–motor co-activation.

Furthermore, in a longitudinal study, it was possible to show that the formation of such neuronal multisensory connections needs less than six weeks of regular piano training (Bangert and Altenmüller 2003). This demonstrates how dynamically brain adaptations accompany musical learning processes.

Activation of motor co-representations can occur in trained pianists not only by listening to piano tunes, but also by observing a pianist's finger movements while watching a video. In Figure 31.3(b) the increases in brain activation of trained pianists are shown whilst they are observing video sequences of a moving hand at the piano as compared to the activation of musically naive subjects (Haslinger *et al.* 2005). Besides the motor hand area in the primary motor cortex, secondary auditory cortices in the temporal lobe and the cerebellum are activated. This neuronal network corresponds to a '*mirror*

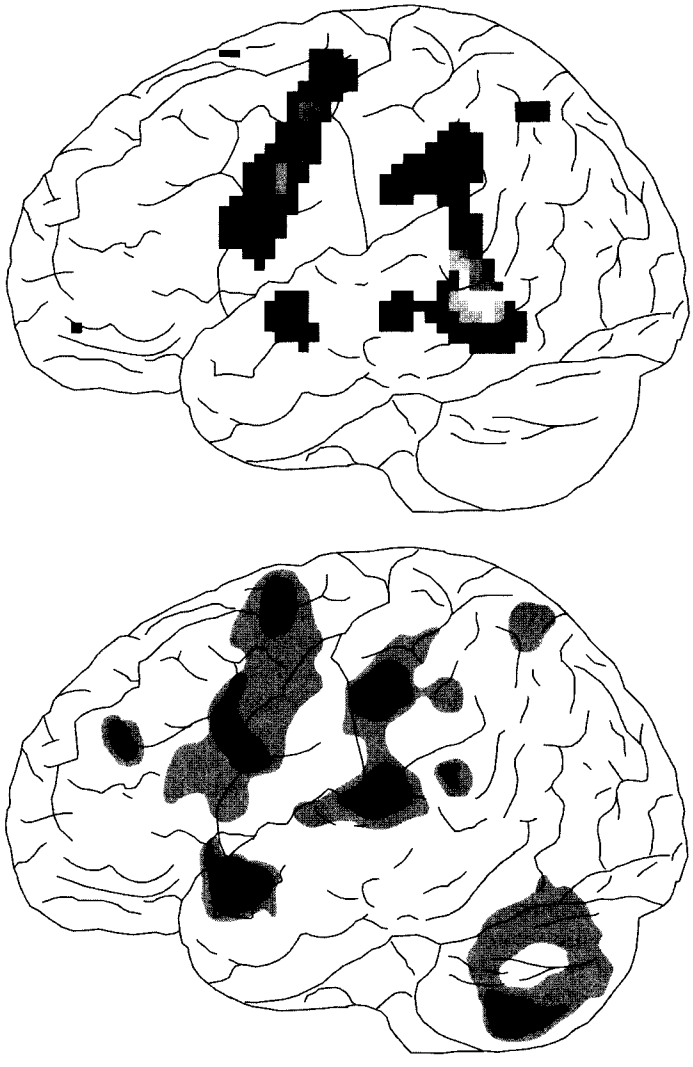


**Fig. 31.3** (a) Additional brain activity (grey zones) of skilled pianists compared to non-pianists when listening to piano tunes without moving their fingers. The primary motor cortex of the precentral area and auditory association areas are lighting up, demonstrating an unconscious co-representation of heard tunes as movement patterns (modified from Bangert *et al*. 2006). (b) Additional brain activity (grey zones) of skilled pianists compared to non-pianists when observing pianist movements in a soundless video. Again, the precentral area and auditory association areas are lighting up demonstrating the mirror system: the observed movements are unconsciously activated, albeit the subjects did not move their fingers. Furthermore the auditory areas are activated demonstrating a visual–auditory co-representation of seen movements. This effect demonstrates impressively the powerful human imitation system, and may be utilized by teachers when demonstrating at the instrument (modified from Haslinger *et al*. 2005).

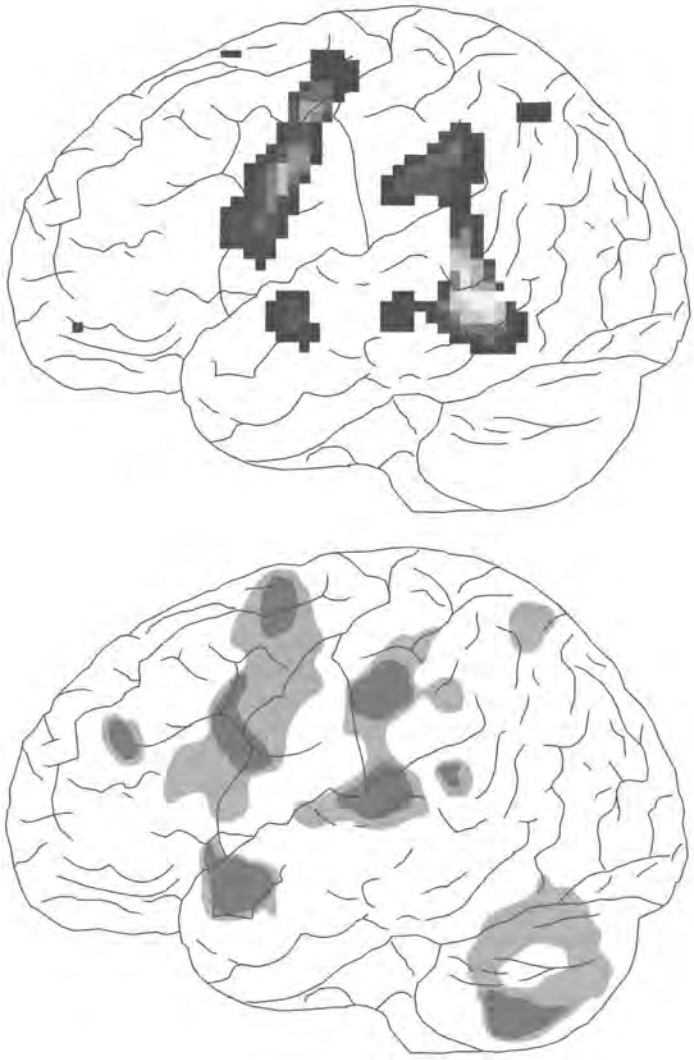


**Fig. 31.3** (a) Additional brain activity (grey zones) of skilled pianists compared to non-pianists when listening to piano tunes without moving their fingers. The primary motor cortex of the precentral area and auditory association areas are lighting up, demonstrating an unconscious co-representation of heard tunes as movement patterns (modified from Bangert *et al*. 2006). (b) Additional brain activity (grey zones) of skilled pianists compared to non-pianists when observing pianist movements in a soundless video. Again, the precentral area and auditory association areas are lighting up demonstrating the mirror system: the observed movements are unconsciously activated, albeit the subjects did not move their fingers. Furthermore the auditory areas are activated demonstrating a visual–auditory co-representation of seen movements. This effect demonstrates impressively the powerful human imitation system, and may be utilized by teachers when demonstrating at the instrument (modified from Haslinger *et al*. 2005).

*neuron network*'. As a consequence of musical practice, it follows that careful demonstration at the instrument may enhance learning. Such a teaching method based on demonstration and imitation is widely used at all levels of musical training, and would appear to be particularly effective in cases where teachers demonstrate an action or series of actions that are carefully and methodically observed by the student.

Practicing through listening and/or observation can be considered as special cases of *mental training*. Narrowly defined, mental training is understood as the vivid imagination of movement sequences without physically performing them. As with observation of actions, principally the same brain regions are active as if the imagined action is performed; that is, the primary motor cortex, the supplementary motor cortex and the cerebellum (Kuhtz-Buschbeck *et al.* 2003). In a study investigating mental training of finger movement sequences of different complexities, brain activation increased along with the degree of difficulty of the imagined motor task. Furthermore, when continuing mental practice over a period of several days, the involved brain regions showed plastic adaptations. Although these adaptations are less dramatic than if the motor tasks were practiced physically, mental training produced a clear improvement in task performance as assessed in finger tapping tests.

## Plasticity of sensory motor systems: musicians' brains are different

During the past decade, brain imaging has provided important insights into the enormous capacity of the human brain to adapt to complex demands. These adaptations are referred to as *brain plasticity* and do not only include the connections or firing rates of neurons—the 'software' of our brain—but also the 'hardware', namely the fine structure of nervous tissue and even the visible gross structure of brain anatomy. Brain plasticity is best observed in complex tasks with high behavioural relevance for the individual such that they cause strong emotional and motivational activation. Plastic changes are more pronounced in situations where the task or activity is intense and the earlier in life it has been developed. Obviously, the continued activities of accomplished musicians provide the prerequisites of brain plasticity in an ideal manner. It is therefore not astonishing that the most dramatic brain plasticity effects have been demonstrated in professional musicians (for a review see Münte *et al.* 2002).

Our understanding of the molecular and cellular mechanisms underlying these adaptations is far from complete. Brain plasticity may occur on different time axes. For example, the efficiency and size of synapses may be modified in a time window of seconds to minutes, the growth of new synapses and dendrites may require hours to days. An increase in grey matter density, which mainly reflects an enlargement of neurons, needs at least several weeks. White matter density also increases as a consequence of musical training. This effect is primarily due to an enlargement of myelin cells: the myelin cells, wrapped around the nerve fibres (axons) are contributing essentially to the velocity of the electrical impulses travelling along the nerve fibre tracts. Under conditions requiring rapid information transfer and high temporal precision these myelin cells grow and as a consequence nerve conduction velocity increases. Finally, brain regions involved in specific tasks may also be enlarged after long-term training due to the growth of structures supporting the nervous function, for example, in the blood vessels that are necessary for the oxygen and glucose transportation to sustain nervous function.

Comparison of the brain anatomy of skilled musicians with that of non-musicians shows that prolonged instrumental practice leads to an enlargement of the hand area in the motor cortex (Amunts *et al.* 1997) and to an increase in grey matter density corresponding to more and/or larger neurons in the respective area (Gaser and Schlaug 2003). These adaptations appear to be particularly prominent in all instrumentalists who have started to play prior to the age of ten and correlate positively with cumulative practice time. Furthermore, in professional musicians, the normal anatomical difference between the larger, dominant (mostly right-) hand area and the smaller, non-dominant (left-) hand area is less pronounced when compared to non-musicians. These results suggest that functional adaptation

of the gross structure of the brain occurs during training at an early age.

Similar effects of specialization have been found with respect to the size of the corpus callosum. Professional pianists and violinists tend to have a larger anterior (front) portion of this structure, especially those who have started prior to the age of seven (Schlaug *et al.* 1995). Since this part of the corpus callosum contains fibres from the motor and supplementary motor areas, it seems plausible to assume that the high demands on coordination between the two hands, and the rapid exchange of information may either stimulate the nerve fibre growth—the *myelination* of nerve fibres that determines the velocity of nerve conduction—or prevent the physiological loss of nerve tissue during ageing.

In summary, when training starts at an early age (before about seven years), these plastic adaptations of the nervous system affect brain anatomy by enlarging the brain structures that are involved in different types of musical skills. When training starts later, it modifies brain organization by rewiring neuronal webs and involving adjacent nerve cells to contribute to the required tasks. These changes result in enlarged cortical representations of, for example, specific fingers or sounds within existing brain structures. In the following section, the behavioural correlates of the maladaptive plastic changes, leading to a loss of motor control of highly skilled movements, will be focused on.

## Focal dystonia: when planning goes wrong

Approximately 1 in 100 professional musicians suffer from a loss of voluntary control of their extensively trained, refined, and complex sensorimotor skills—a condition generally referred to as focal dystonia, violinist's cramp, or pianist's cramp. In most cases, focal dystonia is so disabling that it prematurely ends the artist's professional career (Altenmüller 2003). Subtle loss of control in fast passages, finger curling (cf. Figure 31.4), lack of precision in forked fingerings in woodwind players, irregularity of trills, sticking fingers on the keys, involuntary flexion of the bowing thumb in strings, impairment of control of the embouchure in woodwind and brass players in certain registers, are the various symptoms that can mark the beginning of the disorder. At this stage, most musicians believe that the reduced precision of their movements is due to a technical problem. As a consequence, they intensify their efforts, but this often only exacerbates the problem.

Males, classical musicians of a younger age and instrumentalists such as guitarists, pianists and woodwind players are among the most commonly affected by focal dystonia. The majority of patients have solo positions and often they have a perfectionist, control-type personality. About 20 per cent of such patients report a history of chronic pain syndromes or over-use injury. Preventing these musicians from developing chronic over-use and tendinitis will most probably prevent them from developing focal dystonia (Jabusch and Altenmüller 2006). However, once focal dystonia is established, the cure of the pain syndrome will generally not eliminate the pathological movement pattern.

Thus far the aetiology of focal hand dystonia is not completely understood, but is probably multifactorial. Without going into the details, most studies of focal dystonia reveal abnormalities in three main areas: (a) reduced inhibition in the motor system at cortical, subcortical and spinal levels (b) reduced sensory perception and integration; and (3) impaired sensorimotor integration. The latter changes are mainly believed to originate from dysfunctional brain plasticity. There is growing evidence for an abnormal cortical processing of sensory information as well as degraded representation of motor functions in patients with focal dystonia. A study with trained monkeys demonstrated that chronic over-use and repetitive strain injury in highly stereotyped movements can actively degrade the cortical representation of the somatosensory information that guides the fine motor hand movements in primates (Byl *et al.* 1996). A similar degradation of sensory feedback information and concurrent fusion of the digital representations in the somatosensory cortex was confirmed in a brain activation study conducted in musicians with focal dystonia, although these musicians had no history of chronic pain (Elbert *et al.* 1998). Therefore, additional factors such as a genetic predisposition and certain susceptibility appear to play an



**Fig. 31.4** Typical patterns of dystonic posture in a pianist, a violinist, a flutist and a trombone player.

important role in the development of focal dystonia (Schmidt *et al.* 2006). Interestingly, in musicians suffering from focal dystonia the BP is markedly larger as compared to healthy musicians (see Figure 31.2). This 'overshoot' in brain activation prior to movement execution seems to be linked to the core deficit in focal dystonia, the defective inhibition of motor output.

Unfortunately, there is no simple cure for the condition. Retraining may be successful in a minority of cases, but usually requires several years to succeed. Symptomatic treatment with temporary weakening of the cramping muscles by injecting Botulinum-toxin has proven to be helpful in other cases; however, since the injections need to be applied regularly every three to five months during the professional career, it presents no solution for young patients. Thus, the challenge is to prevent young musicians from such a disorder. Reasonable practice schedules, economic technique, prevention of over-use and pain, mental practice, avoidance of exaggerated perfectionism and psychological support with respect to self-confidence are the components of such a prevention programme.

## Conclusions: some implications for practice

In the preceding paragraphs we have demonstrated the neurobiological foundations of planning, motor learning and practice. Here we will summarize the data in order to formulate some practical rules which might be useful for the daily work of instrumentalists.

As with all skilled human motor activities, effective planning, movement preparation and practising are largely based on procedural knowledge. How to practice and when to stop practising is best learned by experience. Practising can be considered as a self-organizing process, which frequently starts with uneconomical activation of large neuronal pools in the sensorimotor brain regions. Optimizing the movement patterns occurs under continuous sensory feedback from the ears, the eyes, the muscles, tendons, joints, and from the skin. The integration of this information into movement patterns is the most important step in procedural learning. It is mainly based on the formation of neuronal networks; for example, the connections between



**Fig. 31.4** Typical patterns of dystonic posture in a pianist, a violinist, a flutist and a trombone player.

important role in the development of focal dystonia (Schmidt *et al.* 2006). Interestingly, in musicians suffering from focal dystonia the BP is markedly larger as compared to healthy musicians (see Figure 31.2). This 'overshoot' in brain activation prior to movement execution seems to be linked to the core deficit in focal dystonia, the defective inhibition of motor output.

Unfortunately, there is no simple cure for the condition. Retraining may be successful in a minority of cases, but usually requires several years to succeed. Symptomatic treatment with temporary weakening of the cramping muscles by injecting Botulinum-toxin has proven to be helpful in other cases; however, since the injections need to be applied regularly every three to five months during the professional career, it presents no solution for young patients. Thus, the challenge is to prevent young musicians from such a disorder. Reasonable practice schedules, economic technique, prevention of over-use and pain, mental practice, avoidance of exaggerated perfectionism and psychological support with respect to self-confidence are the components of such a prevention programme.

## Conclusions: some implications for practice

In the preceding paragraphs we have demonstrated the neurobiological foundations of planning, motor learning and practice. Here we will summarize the data in order to formulate some practical rules which might be useful for the daily work of instrumentalists.

As with all skilled human motor activities, effective planning, movement preparation and practising are largely based on procedural knowledge. How to practice and when to stop practising is best learned by experience. Practising can be considered as a self-organizing process, which frequently starts with uneconomical activation of large neuronal pools in the sensorimotor brain regions. Optimizing the movement patterns occurs under continuous sensory feedback from the ears, the eyes, the muscles, tendons, joints, and from the skin. The integration of this information into movement patterns is the most important step in procedural learning. It is mainly based on the formation of neuronal networks; for example, the connections between

auditory and motor areas and in a step-wise reduction of cortical activity and augmentation of subcortical activity in the basal ganglia and the cerebellum.

When playing a musical instrument, the central nervous system is mainly involved in processing a huge amount of incoming information from the ears and eyes, and from the sensory organs in muscles, tendons, joints, and skin. The consolidation of the networks necessary for programming movement sequences occurs mainly in the breaks after playing and during sleep. As a consequence, the more complex a task is, the shorter the practice time that should be scheduled in one session and the longer the breaks should be planned. Sleep is another factor supporting procedural memory formation. Therefore, sufficient sleep should be encouraged, especially when an instrumentalist of any ability level is working hard to master new repertoire.

Generally, a practice session should be terminated when signs of fatigue appear. It is important to consider that over-practice (practice into bodily or mental fatigue) not only leads to no improvement, but to an active worsening of motor programmes. This is due to a blurring of central nervous system sensorimotor representations, when muscular fatigue appears. Furthermore, a lack of attention causes a higher probability of uneconomical movements or production of false notes which, as a consequence, are stored in the procedural memory.

The human mirror system is a powerful tool to facilitate skill learning. Auditory and visual cues presented to students activate their sensorimotor representations and can lead directly to the formation of motor programmes. This is the basis of imitation learning. On the other hand, sloppy and careless demonstrations may produce a negative effect for students, decreasing their sensorimotor programmes as they adopt bad habits modelled by their teacher. Teachers should therefore demonstrate skills in a variety of ways in order to ensure that their students are able to comprehend the difference between effective and ineffective performance techniques.

We would like to conclude our chapter with a general remark: processes involved with instrumental musical training are probably the most complex of all human activities. Importantly, these are not restricted to the sensorimotor brain circuits but also involve emotion, memory and imagination. The best trained musicians with the best working sensorimotor networks will not move their listeners if imagination, colour, fantasy and emotion are not a part of their artistic expression. These qualities are often not trained solely within a practice studio, but depend on and are possible linked to experiences from daily life, to human relationships, to a rich artistic environment and to empathy and emotional depth. These factors, which profoundly influence the aesthetic quality of a music performance, are at present far from being accessible to any neuroscientific research.

## Acknowledgments

The authors thank Gary McPherson for his valuable comments on an earlier version of the manuscript. Both authors are generously supported by a grant of the German Research Foundation (Al 269/7–1).

## References

Altenmüller E (2003). Focal dystonia: advances in brain imaging and understanding of fine motor control in musicians. *Hand Clin*, **19**, 523–538.

Altenmüller E, Münte TF and Gerloff Ch (2004). Neurocognitive functions and the EEG. In E Niedermeyer and F Lopes da Silva, eds, *Electroencephalography*, 5th Edn, 661–682. Lippincott Williams, Baltimore, MD.

Amunts K, Schlaug G, Jäncke L, Steinmetz H, Schleicher A, Dabringhaus A and Zilles K (1997). Motor cortex and hand motor skills: structural compliance in the human brain. *Human Brain Mapping*, **5**, 206–215.

Bangert M and Altenmüller E (2003). Mapping perception to action in piano practice: a longitudinal DC-EEG-study. *BMC Neuroscience*, **4**, 26–36.

Bangert M, Peschel T, Rotte M, Drescher D, Hinrichs H, Schlaug G, Heinze HJ and Altenmüller E (2006). Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *NeuroImage*, **15**, 917–926.

Bengtsson SL, Nagy Z, Skare S, Forsman L, Forssberg H and Ullen F (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nature Neuroscience*, **8**, 1148–1150.

Byl NN, Merzenich MM and Jenkins WM (1996). A primate genesis model of focal dystonia and repetitive strain injury: I. Learning-induced dedifferentiation of the representation of the hand in the primary somatosensory cortex in adult monkeys. *Neurology*, **47**, 508–520.

Eagleman DM (2004). The where and when of intention. *Science*, **303**, 1144–1146.

Elbert T, Candia V, Altenmüller E, Rau H, Rockstroh B, Pantev C and Taub E (1998). Alteration of digital representations in somatosensory cortex in focal hand dystonia. *NeuroReport*, **16**, 3571–3575.

Ericsson KA, Krampe RT and Tesch-Römer C (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, **100**, 363–406.

Ericsson KA, Lehmann AC (1996). Expert and exceptional performance: evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, **47**, 273–305.

Gaser C and Schlaug G (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience*, **23**, 9240–9245.

Gentner R and Classen J (2006). Modular organization of finger movements by the human central nervous system. *Neuron*, **52**, 731–742.

Halsband U and Lange RK (2006). Motor learning in man: a review of functional and clinical studies. *J Physiol Paris*, **99**, 414–424.

Haslinger B, Erhard P, Altenmüller E, Schroeder U, Boecker H and Ceballos-Baumann AO (2005). Transmodal sensorimotor networks during action observation in professional pianists. *Journal of Cognitive Neuroscience*, **17**, 282–293.

Jabusch HC and Altenmüller E (2006). Epidemiology, phenomenology and therapy of musician's cramp. In E Altenmüller, J Kesselring and M Wiesendanger eds, *Music, motor control and the brain*, 265–282. Oxford University Press, Oxford.

Karni A, Meyer G, Jezzard P, Adams MM, Turner R and Ungerleider LG (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature*, **377**, 155–158.

Kuhtz-Buschbeck JP, Mahnkopf C, Holzknecht C, Siebner H, Ulmer S and Jansen O (2003). Effector-independent representations of simple and complex imagined finger movements: a combined fMRI and TMS study. *European Journal of Neuroscience*, **18**, 3375–3387.

Libet B (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *The Behavioural and Brain Sciences*, **8**, 529–566.

Libet B, Gleason CA, Wright EW and Pearl D (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). *Brain*, **106**, 623–642.

Münte TF, Altenmüller E and Jäncke L (2002). The musician's brain as a model of neuroplasticity. *Nature Neuroscience*, **3**, 473–478.

Peschel T and Altenmüller E (2004). Pathologische bewegungskorrelierte kortikale Potentiale (MRCPs) bei Musikern mit aktionsinduzierter fokaler Dystonie. *Klin Neurophysiologie*, **34**, 96–106.

Roland PE, Larsen B, Lassen NA and Skinhoj E (1980). Supplementary motor area and other cortical areas in the organization of voluntary movements in man. *Journal of Neurophysiology*, **43**, 118–136.

Schlaug G, Jäncke L, Huang Y and Steinmetz H (1995). Increased corpus callosum size in musicians. *Neuropsychologia*, **33**, 1047–1055.

Schmidt A, Jabusch HC, Altenmüller E, Hagenah J, Brüggemann N, Hedrich K, Saunders-Pullman R, Bressman SB, Kramer PL and Klein C (2006). Dominantly transmitted focal dystonia in families of patients with musicians' cramp. *Neurology*, **67**, 691–693.

# CHAPTER 32

# Sight-reading

Andreas C. Lehmann and Reinhard Kopiez

## Introduction

Anyone who wants to perform the works of traditional Western music, and approach other musical styles using a similar performance practice, will most likely have to master music notation. Many musical cultures rely on a system of symbols to store and teach complex musical styles that are not, or only partly, grounded in improvisation. While the beginning music reader has to overcome the same problems as all readers do when learning to read other texts, namely going from a tedious matching of symbols to sounds to meaning, the expert reader has automatized the process of encoding and transforming the signs into embodied action.

For our purposes, we will call sight-reading the execution—vocal or instrumental—of longer stretches of non- or under-rehearsed music at an acceptable pace and with adequate expression. Some people also label this 'playing by sight' or 'prima vista'. Another related set of activities might be called music or note reading: for example, the following of a score with the aim of studying a piece while the music is playing, or the studying of a new piece of music away from the instrument prior to physically practising it. Both activities may even be accompanied by the sounding out of some notes on an instrument. However, the characteristic goal of sight-reading is the authentic performance, or as Mozart once phrased it, 'to play the piece . . . so as to make believe that it had been composed by the one who plays it' (letter from 1778 cited in Crofton and Fraser 1985, p. 111). Similar to improvisation, sight-reading requires the instant adaptation to new constraints, which places it among those that motor scientists refer to as open skills (as opposed to closed skills that require reproduction of well-rehearsed motions, such as swimming, figure skating or playing a well-rehearsed piece of music). Today, studio musicians and accompanists must be able to sight-read, and many orchestra musicians have done it on a regular basis for centuries. It is this kind of un- or under-rehearsed performance that we are concerned with because it forms a well-defined and discrete skill comparable to that of playing by ear or improvising.

When and under what conditions does sight-reading emerge as a skill in a culture? Music archaeology tells us about some forms of notation from ancient Egypt and Greece and other cultures which have developed coding systems for music that mainly serve mnemonic purposes (e.g. Bent *et al.* 2008). The Western model of notation started to develop in the ninth century in order to code multivoice vocal and later instrumental music. In a musical culture (regardless of geographical place and historical time) in which music is solely improvised, sight-reading is not necessary, yet music reading skills may still be used for learning to play an instrument (e.g. in India). When composers and performers assume specialized roles and performers are expected to perform a fixed repertoire, then sight-reading may be required to limit rehearsal times or if the repertoire changes frequently (e.g. in the Baroque era). While the nineteenth century still saw renowned performers play other composers' pieces from the score, canonization of the repertoire, the developing art of interpretation, and rising audience expectations has since then led to a performance practice based on memorized performance by solo performers of the piano and violin. Our modern performance traditions have come to favour

polished performances and relegated sight-reading to a useful craft, generally not worthy of public notice or competition (we know only of the Karl Bergemann sight-reading competition, Hanover, Germany).

Obviously, most orchestral, chamber, and studio musicians play from the score and so do many performers of contemporary art music. There are a few known solo pianists who have used scores even for standard repertoire. It is unclear whether musicians perform better with or without the score, but the audience is likely to expect memorized performances. Also, using the score requires page-turning or the presence of a music stand which may be disturbing for the audience and the performer and might detract from the sounding music. A music stand may obstruct the view and hinder gestural freedom and the positive influence of expression that is conveyed through the body (cf. Lehmann *et al.* 2007, pp. 173–174).

In marked contrast to the public neglect and low prestige of sight-reading among performers stands the steady interest of music psychologists and pedagogues in this skill (for reviews see Sloboda 1984; Lehmann and McArthur 2002; Lehmann *et al.* 2007, Chapter 6). Starting with the early music psychology experiments of the 1920s (published in Jacobsen 1941) and the development of sight-reading tests (Watkins 1942, and the Watkins–Farnum Performance Scale), all aspects of skills relating to sight-reading attracted renewed interest in the 1970s (e.g. Sloboda 1974, 1976, 1977) and have continued to do so (cf. Lehmann and Ericsson 1993, 1996; Kopiez and Lee 2006, 2008 for the acquisition of sight-reading skills) up until recent psychophysiological studies (e.g. Schon and Besson 2005; Yumoto *et al.* 2005). It is important for research that we measure individuals' ability to perform at first sight under standardized conditions. Optimally, those conditions should mirror real-life conditions encountered by expert sight-readers (Lehmann and Ericsson 1993).

In this chapter we will briefly look at how notation is perceived and then move on to the structure of sight-reading while taking into account the real-time conditions under which it takes place. This will include a discussion of perceptual and problem-solving issues. Finally we will outline the course of skill acquisition with its characteristic differences between novices and experts, and present a model of sight-reading performance.

## Perception of music notation and sight-reading

To begin, we have to understand how the eye operates when we try to acquire information in everyday life. Contrary to what most people believe, the eye does not function like a movie camera. Rather, its operation can be likened to that of a flashlight in the dark being turned on and off at short intervals. Roughly four to five times a second the eye moves around the visual field in discrete jumps (saccades) with short resting points (fixations). The saccades take about 15–50 msecs, the fixations about 150–200 msecs. At this point it becomes clear that a conscious attending to every eighth note in a piece at MM = 120 would be almost impossible. During each fixation, the external image is projected on the retina at the back of the eye. Wile the retina is comparably large, we are able to receive a sharp image only from a narrow part in the middle. This round central part is called the fovea centralis, and whatever surrounds it will produce the somewhat blurry parafoveal image. Hence, the field of vision that will be perceived in great clarity only averages 0.5–2°, which corresponds to the size of an inch at a reading distance of about 30 inches (75 cm), or the area covered by a pointed-up thumb with extended arm. The parafoveal vision includes 10° of the field of vision (e.g. Rayner and Pollatsek 1989). The role of preview benefits and parafoveal-on-foveal effects are currently a hot topic in word-reading research. It is from such individual snapshots that our brain fashions what we experience as a large and steady picture of the outside world. Unfortunately, the eye movements cannot be allocated wilfully but are in fact guided by preconscious processes and drawn by outside stimuli. For example, movements and boundaries in the visual field attract attention, just as do human faces, but our cognition also guides the eye movements. For example, when a car disappears behind another we are likely to scan the plausible location of its

reappearance, and we search the face of a person for cues to his or her mood. Today, we know that information gathered from one or several fixations is integrated in meaningful units or chunks of information which are the basis for further processing. Since the location and duration of fixations is indicative of the processing underlying music reading, eye movements offer important insights into the workings of the musical mind. To explain the structure of sight-reading, we have to account for how much and which information is retrieved from the page and how it is assembled into meaningful units that are sequentially programmed and executed.

The problem in surveying the results of eye movement in sight-reading is that the research methodologies are not standardized with regard to complexity of stimulus, tempo of performance, and so forth. Unlike in text-reading research where many studies appear within a few years using the same paradigm, the time lag between publications on sight-reading is large and findings are often difficult to integrate. By and large, the earliest studies (Jacobsen 1941) established what subsequent studies have confirmed, namely that eye movement patterns are dependent on the level of expertise: beginners had many fixations, long pauses during fixations, and unsystematic reading of note combinations; intermediate musicians had about as many fixations as there were notes, and they read chords in systematic fashion from bottom to top; experts showed fewer fixations than notes and also systematic reading of chords (from top to bottom). Saccades can point forward in reading direction but also backwards (regressively)—for example to the current point of performance. This is most likely done to double-check things that have been read already or result from attention being detracted to performance errors (not a very efficient strategy). With increasing experience, the sight-reader experiences a reduction in the number of regressive fixations. Kinsler and Carpenter (1995), who strangely claimed that 'a thorough search of the literature failed to find any account whatsoever of the eye movements used to read music' (p. 12), studied eye movements during performance of notated rhythms. Their results showed that slower tempi lead to more and shorter fixations. One problem in research with self-selected performance tempo is that slower tempi necessarily result in more fixations. Lannert and Ullman (1945) found a 0.45 correlation between tempo and accuracy, and we can never be sure if participants have traded off faster tempi for more fixations and thus ensure a more accurate performance. Equal speeds can only be achieved by using a pacing-voice methodology (Lehmann and Ericsson 1993).

That notational input has an influence on the eye movements was mentioned previously (cf. systematic vertical reading of chords). Notational variants (e. g., eighth notes with or without connecting bars) resulted in person-specific eye movements, and eighth notes (with connecting bars) tended to be looked at in pairs while quarter notes were attended to individually (Kinsler and Carpenter 1995). Truitt *et al.* (1997, also Goolsby 1994) questioned why fixations often landed between notes, and we suggest that readers tend to construct intervals rather than reading every note. Weaver (1943) found that polyphonically structured music was read in horizontal zigzagging patterns that tended to follow melodic lines in the different voices, whereas homophonic music resulted in zigzagging up and down motions. However, a critical look at his stimuli unveils that notation and structure were hopelessly confounded, e.g. no attempt was made to notate polyphonic stimuli in alternative (more homophonic looking) ways. none of the stimuli was polyphonic, yet it notated as if homophonically structured. Regardless of the shortcomings, we can say that experience and structure of input modify viewing patterns.

Once the information has been retrieved during one or several fixations, it is stored and assembled in meaningful units in anticipation of the motor performance. Here also it was found that experience allowed for larger temporal range of planning (Drake and Palmer 2000). The extent and nature of the buffering of information is part of the memory system to be discussed in the next section.

## Memory processes

The amount of information stored temporarily from a particular sequence of fixations or during a certain timespan can be assessed by experiments that allow sight-readers only limited visual

access to the score by either very brief (tachistoscopic) display, by limiting the period of time during which the score is visible, or by using a computer that follows the fixations with a 'moving window' technique that permits variable preview. By this we can measure aspects of memory, namely the perceptual and the eye–hand span. The perceptual span denotes the distance between the current point of performance and the farthest distance the person is looking ahead. Using a moving window technique, Truitt *et al.* (1997) found that a preview of two beats leads to a slower tempo, larger variability in note durations, and errors. Subjects performed better with previews between two and four beats or, ideally, with previews to the end of the next bar. Furneaux and Land (1999) found the number of notes to be about four for experts and two for novices. Unlike Truitt *et al.* (1997) who found the time between fixation and performance to stand at 0.5 secs, Furneaux and Land (1999) documented between 0.7–1.3 seconds, dependent on the tempo. Similar results have been found in studies on typewriting (Gentner 1988). This narrow preview is at odds with the phenomenological experience of sight-readers who claim a much larger preview. With multiple fixations that can go anywhere in the piece, musicians construct motor programmes that rely on more than mere visual input of the foveal area (see below for further details). This leads to the larger eye–hand spans that can be measured when withdrawing the notation unexpectedly (Sloboda 1977). Sloboda found that meaningful musico-structural units influenced the length of the eye–hand span. For example, a larger distance from the next phrase boundary tended to stretch the eye–hand span; a shorter caused it to shrink. Hiding a longer piece of music at arbitrary points might still allow for cumulative effects of the previously sight-read material, whereas the method of briefly displaying disjointed snippets of information for several hundred milliseconds does not. We know that the reading context influences patterns of fixation (Bekkering and Neggers 2002). If we consider that repeated trials lead to better sight-reading accuracy and that better sight-readers have a better recall for material after a single trial (Lehmann and Ericsson 1993), we have to consider the effects of long-term working memory typically found among experts (Ericsson and Kintsch 1995). This privileged access to long-term memory allows expert readers to store briefly presented material in long-term memory without extensive rehearsal.

## Inner hearing

Some authors have claimed that inner hearing and audiation processes may be important in sight-reading, and independent tests of audiation, imagery, and pattern matching are positively associated with sight-reading ability (e.g. Kornicke 1995; Waters *et al.* 1998; Kopiez and Lee 2006). These processes would suggest that the mental representation of music notation involves the building of melodic and other expectancies by the performer. Recent electrophysiological studies have confirmed this point. In studies in which listeners followed a visually presented score that was accompanied by the corresponding sounds, discrepancies between the printed score and auditory events resulted in a recorded mismatch negativity about 150 msec after the dissociation in the vicinity of Heschl's gyrus, where auditory pitch detection is located (Schon and Besson 2005; Yumoto *et al.* 2005). This means that performers are likely able to know if their sounds match the score or not. Whether or not auditory images are used for planning of movements may depend on the musician. For example, Banton (1995) found that sight-reading without auditory feedback led to only slightly more mistakes than normal feedback; however, omitting visual access to the keys resulted in markedly poorer performance. A classic experiment by Allport *et al.* (1972) revealed that pianists could repeat words that were presented while sight-reading at the piano, which suggests that auditory feedback is not necessary. However, it may be used to create expectations.

## Sight-reading as problem-solving

We already mentioned that not all notes can be focused on and that problem-solving processes will have to complement the incomplete visual input. In fact, everyday experiences teach us that some pieces are able to be sight-read more easily

than others, which suggests that those processes may be more or less easily accomplished. Ortmann in 1934 (cited in Clifton 1986) showed with brief presentation times of stimuli, 400 msec and 2 sec, respectively, that diatonically organized music, smaller intervals, and sequences that were congruent with tonal expectations were more easily read than others. This suggests—and this conclusion is backed by much research in other domains—that we tend to form meaningful units that are influenced by our previous knowledge, and hence, expectations. Kinsler and Carpenter (1995) found that repeated renditions of the same piece were accompanied by a reduced amount of ocular movements. Presumably, larger chunks were formed, more previous knowledge was brought to the task, and the visual input functioned as a retrieval cue to a known motor programme (just like playing well-rehearsed music from notation).

What happens when our expectations are not met by the printed score? In a clever study (Sloboda 1976), pianists performed a classically sounding piece of music in which several notes had been altered by a half or whole step to violate tonal expectations. These violations were either introduced in the beginning, middle or the end of a phrase, and they were evenly distributed across the left- and right-hand part. Participants were asked to play exactly what was written. As the researcher expected, many of the artificial alterations were erroneously corrected to sound tonal again, middle positions being more likely corrected, and a repeated trial led to even more such corrections. These alterations may be termed proof-reader's errors. Sloboda (1974) asked music experts to judge mistakes made by sight-readers with regard to their musical adequacy. In keeping with our expectations, the errors were plausible alternatives to or reductions of the written music. Expectations are powerful: one can ask pianists to fill in blanks in an unfamiliar piece while sight-reading it, and they will generate successful inferences (i.e. improvise) based solely on the context of the piece, their stored knowledge, and current expectations (Lehmann and Ericsson 1996). Fine *et al.* (2006) showed that singers also made more mistakes when sight-reading tonally modified Bach chorals compared to unmodified ones and that they were also hindered by modified notes in other voices' parts.

Sight-reading is more than pressing the right notes at the right time; it also involves adding musical expression. Although no specific studies have been done so far, we suggest that expression is added algorithmically according to a likely grammar of musical expression (e.g. Friberg *et al.* 2006). Most interestingly, merely counting up the number of correct notes in a sight-reading task correlates highly with expert ratings of the same performance (Lehmann and Ericsson 1993), suggesting that better sight-readers integrate musical expression and correct notes on the fly. This may be due to a larger temporal range of planning. Only when the performer knows what is ahead can meaningful musical expression be added (e.g. slowing at the end of a phrase).

In essence, problem-solving or reconstructive processes during sight-reading are considerable if we consider how few fixations are available and how much material needs to be covered. Obviously, those processes function more effectively in better sight-readers. While plausible expectations are constantly being constructed and will usually facilitate performance, in rare instances they may lead the sight-reader astray and cause errors that unveil the underlying mediating processes.

## Acquiring the skill

Sight-reading is such a specialized skill that it is futile to look for specific inborn traits that may cause individual differences in sight-reading skill. Among musicians observable individual differences in sight-reading ability are great and need to be explicated. Several variables have attracted the interest of researchers: all sorts of training variables, intelligence, other musical skills, musical ability, general indicators of memory, and reading performance.

The most promising predictors so far have been training variables. Kornicke (1995) and Banton (1995) found that higher self-rated experience as an accompanist or more regular sight-reading practice was related to more accurate sight-reading in the authors' studies. In Kornicke's study, a cumulative index of experience from several scales correlated reliably ($r = 0.4$) with sight-reading achievement. In a regression analysis the best predictor was the estimated number of pieces sight-read. Lehmann and Ericsson

(1996) assessed first, the size of repertoire that pianists possessed to accompany soloists and ensembles and second, accumulated duration of accompanying experience. Both indicators accounted independently for individual differences in sight-reading: a larger accompanying repertoire and more accompanying experience led to better sight-reading performance under standardized conditions. One could conclude from this that experience with the representative situation (real-time demands, short or no preparation time, etc.) as well as the knowledge base acquired through the performance of many pieces facilitates future performance. In general, among the participants of the study the experience as accompanist started a few years after the onset of piano training. They reported playing progressively more difficult accompanying piano parts, commensurate with their parallel increase in pianistic skills. Counter to many musicians' intuitions, this explains nicely why positive correlations are found between general instrumental skills and sight-reading at pre-professional levels (McPherson 1995, 2005).

A similar study on the acquisition of sight-reading skills included different levels of musical complexity (Kopiez and Lee 2006, 2008). In their experiment, subjects sight-read pieces of five levels of difficulty. Sight-reading performance was assessed along with indicators of psychomotor speed (tapping speed, trill speed), elementary cognitive skills (reaction time, speed of information processing [number connection test]), general cognitive skills (tests of short-term and working memory and general mental capacity), and expertise-related indicators (inner hearing, practice times for piano alone and accompanying, starting ages). The best single predictors for the overall score across all five levels of difficulty were trill speed between the third and fourth fingers, sight-reading experience up to age 15, scores on the number connection test, and inner hearing (Kopiez and Lee 2008). While those results emphasize the impact of experience and efficient perceptual encoding, they also suggest that pianistic abilities play a role. The authors also analysed their results separately for the individual levels of complexity (Kopiez and Lee 2006, see Figure 32.1) and showed convincingly that performance at different levels of difficulty was mediated by different combinations of predictors. Hence, those factors that contribute to better performance at lower levels of difficulty may not be the same as those which mediate superior performance at higher levels.

Many educational studies and reviews have offered suggestions on how to improve sight-reading (e.g. Lehmann and McArthur 2002, pp. 147–148; McPherson 2005) and insightful teachers have developed graded material to train sight-reading systematically. However, some advice might be more useful than other: for example, training eye movements by rolling your eyes or similar exercises are certainly futile—you would simply get better at rolling your eyes. Rather, young performers should get accustomed to playing their instrument without constantly looking at it (e.g. piano) to free up their vision to look at the score while still finding their way on the instrument. Performing under real-time conditions also precludes stuttering, i.e. stopping at every mistake and correcting it, but rather 'faking' one's way through the score, i.e. trying to infer plausible content. This can only be done if the student has ample experience with a certain style of music and can build up suitable expectations about how the music might continue. It is here also that knowledge of music theory can be applied. One has to acknowledge that a stable and deliberate interpretation may not be possible at first sight but that attending to dynamic and articulatory signs along with applying simple rules of expression (e.g. creating phrase arches of tempo and loudness) will generate a musically sounding first impression.

## Summary

Sight-reading provides a complex problem-solving situation with an intricate interplay of bottom-up mechanisms (driven by the input stimulus of the score and auditory feedback) and top-down processes (driven by expectations and cognitions). It is conceivable that limitations on the general playing of an instrument or a lack of technical proficiency exist that may consequentially impact the ability to sight-read: One can never sight-read beyond the level of rehearsed performance, but how close to it one sight-reads seems to be very much a matter of training. Sight-reading ability at lower skill levels may partly emerge from general instrumental skill

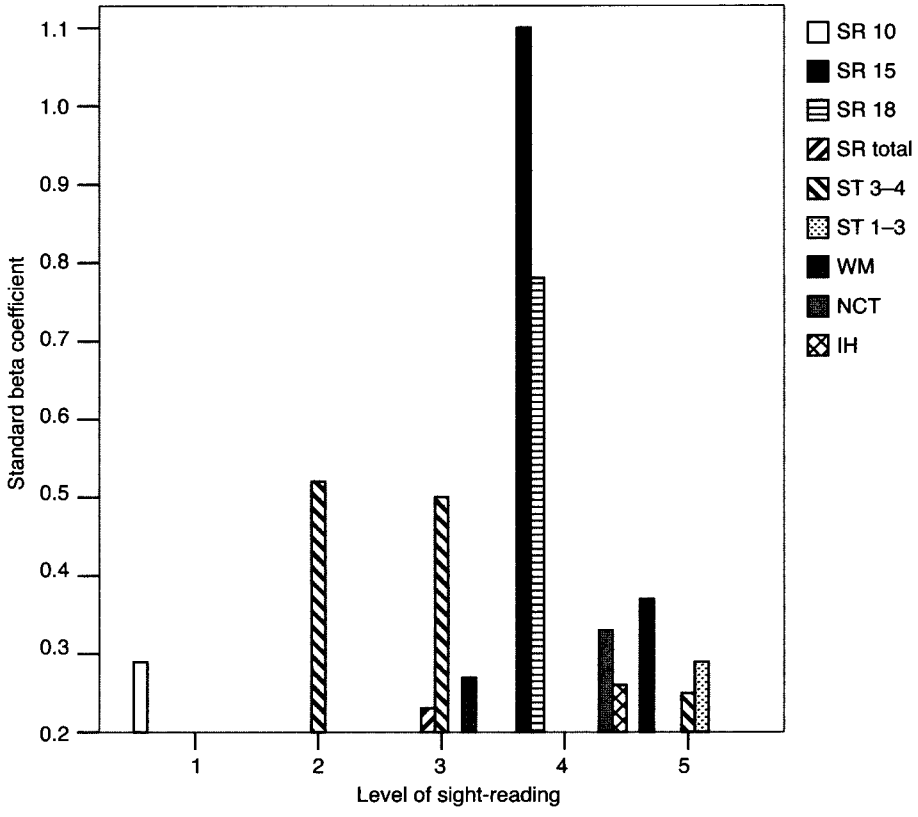



**Fig. 32.1** Significant predictors of sight-reading performance, separate for five levels of stimulus difficulty 1=easy, 5=difficult (data from Kopiez and Lee 2006). Predictors: SR, sight-reading experience up to age 10, 15, 18 or total up to the time of the experiment; ST, speed trilling between fingers 1–3 or 3–4 of right hand; WM, working memory in a number task; NCT, speeded number connection test; IH, inner hearing using an embedded melody paradigm.

increase whereas expert sight-reading necessitates extensive deliberate efforts to improve performance. By engaging in many hours of related experience, for example as an accompanist (in the case of pianists), sight-readers develop particular cognitive adaptations, such as efficient encoding, building of expectations and plausible inferencing, and memory skills. These help them cope with the real-time demands of reconstructing on the fly the score along with a preliminary expressive interpretation.

## References

Allport DA, Antonis B and Reynolds P. (1972). On the division of attention: a disproof of the single channel hypothesis. *Quarterly Journal of Experimental Psychology*, **24**(2), 225–235.

Banton LJ (1995). The role of visual and auditory feedback during the sight reading of music. *Psychology of Music and Music Education*, **23**, 3–16.

Bekkering H and Neggers FW (2002). Visual search is modulated by action intentions. *Psychological Science*, **13**(4), 370–374.

Bent ID *et al.* (22 Apr. 2008). Notation. Grove music online, http://www.oxfordmusiconline.com/subscriber/article/grove/music/20114pg1.

Clifton JV (1986). Cognitive components in music reading and sight-reading performance. Doctoral dissertation, University of Waterloo, Ontario.

Crofton I and Fraser D (1985). *Dictionary of musical quotations*. Schirmer, New York.

Drake C and Palmer C. (2000). Skill acquisition in music performance: relations between planning and temporal control. *Cognition*, **74**(1), 1–32.

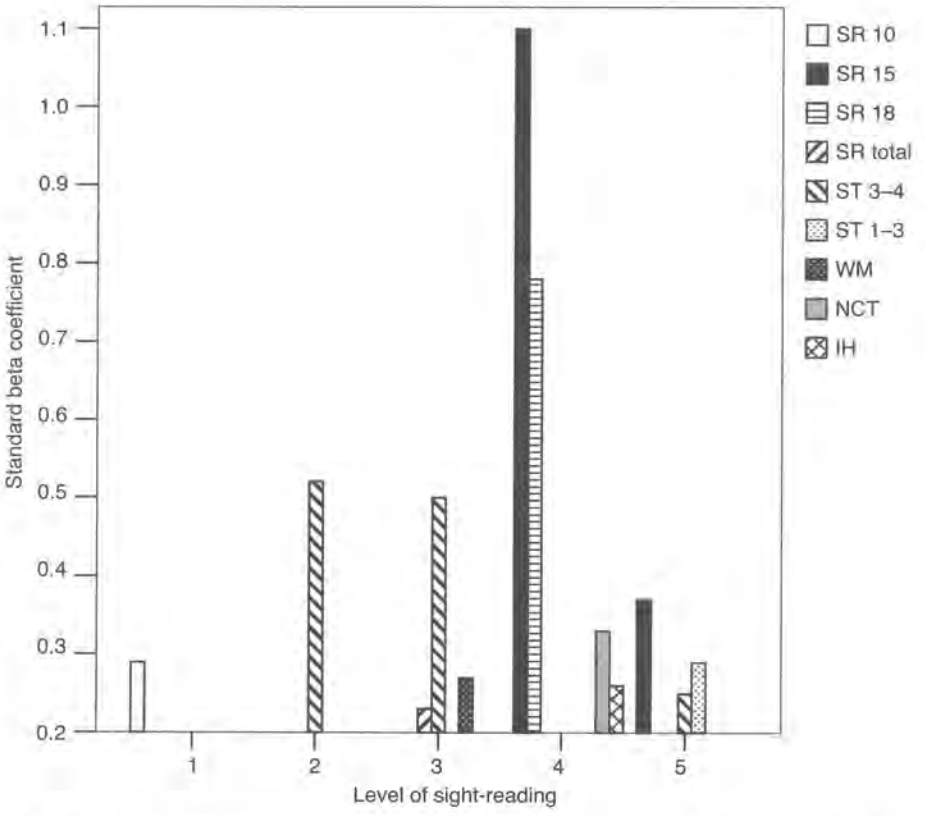



**Fig. 32.1** Significant predictors of sight-reading performance, separate for five levels of stimulus difficulty 1=easy, 5=difficult (data from Kopiez and Lee 2006). Predictors: SR, sight-reading experience up to age 10, 15, 18 or total up to the time of the experiment; ST, speed trilling between fingers 1–3 or 3–4 of right hand; WM, working memory in a number task; NCT, speeded number connection test; IH, inner hearing using an embedded melody paradigm.

increase whereas expert sight-reading necessitates extensive deliberate efforts to improve performance. By engaging in many hours of related experience, for example as an accompanist (in the case of pianists), sight-readers develop particular cognitive adaptations, such as efficient encoding, building of expectations and plausible inferencing, and memory skills. These help them cope with the real-time demands of reconstructing on the fly the score along with a preliminary expressive interpretation.

## References

Allport DA, Antonis B and Reynolds P. (1972). On the division of attention: a disproof of the single channel hypothesis. *Quarterly Journal of Experimental Psychology*, **24**(2), 225–235.

Banton LJ (1995). The role of visual and auditory feedback during the sight reading of music. *Psychology of Music and Music Education*, **23**, 3–16.

Bekkering H and Neggers FW (2002). Visual search is modulated by action intentions. *Psychological Science*, **13**(4), 370–374.

Bent ID *et al.* (22 Apr. 2008). Notation. Grove music online, http://www.oxfordmusiconline.com/subscriber/article/grove/music/20114pg1.

Clifton JV (1986). Cognitive components in music reading and sight-reading performance. Doctoral dissertation, University of Waterloo, Ontario.

Crofton I and Fraser D (1985). *Dictionary of musical quotations*. Schirmer, New York.

Drake C and Palmer C. (2000). Skill acquisition in music performance: relations between planning and temporal control. *Cognition*, **74**(1), 1–32.

Ericsson K and Kintsch W (1995). Long-term working memory. *Psychological Review*, **102**(2), 211–245.

Fine P, Berry A and Rosner B (2006). The effect of pattern recognition and tonal predictability on sight-singing ability. *Psychology of Music*, **34**(4), 431–447.

Friberg A, Bresin R and Sundberg J (2006). Overview of the KTH rule system for musical performance. *Advances in Cognitive Psychology*, **2**(2–3), 145–161.

Furneaux S and Land MF (1999). The effects of skill on the eye–hand span during musical sight-reading. *Proceedings of the Royal Society. Biological Sciences*, **266**(1436), 2435–2440.

Gentner D (1988). Expertise in typewriting. In M Chi, R Glaser and M Farr, eds, *The nature of expertise*, 1–21. Erlbaum, Hillsdale, NJ.

Goolsby TW (1994). Profiles of processing: eye movements during sight reading. *Music Perception*, **12**(1), 97–123.

Jacobsen OI (1941). An analytical study of eye-movements in reading vocal and instrumental music. *Journal of Musicology*, **3**, 1–22.

Kinsler V and Carpenter RH (1995). Saccadic eye movements while reading music. *Vision Research*, **35**, 1447–1458.

Kopiez R and Lee JI (2006). Towards a dynamic model of skills involved in sight reading music. *Music Education Research*, **8**(1), 97–120.

Kopiez R and Lee JI (2008). Towards a general model of skills involved in sight reading music. *Music Education Research*, **10**(1), 41–62.

Kornicke LE (1995). An exploratory study of individual difference variables in piano sight-reading achievement. *Quarterly Journal of Music Teaching and Learning*, **6**(1), 56–79.

Lannert V and Ullman M (1945). Factors in the reading of piano music. *American Journal of Psychology*, **58**, 91–99.

Lehmann AC and McArthur VH (2002). Sight-reading. In R Parncutt and G McPherson, eds, *Science and psychology of music performance*, 135–150. Oxford University Press, Oxford.

Lehmann AC and Ericsson KA (1996). Performance without preparation: structure and acquisition of expert sight-reading and accompanying performance. *Psychomusicology*, **15**(1/2), 1–29.

Lehmann AC and Ericsson KA (1993). Sight-reading ability of expert pianists in the context of piano accompanying. *Psychomusicology*, **12**(2), 182–195.

Lehmann AC, Sloboda JA and Woody RH (2007). *Psychology for musicians: understanding and acquiring the skills*. Oxford University Press, New York.

McPherson GE (1995). The assessment of musical performance: development and validation of five new measures. *Psychology of Music*, **23**, 142–161.

McPherson GE (2005). From child to musician: skill development during the beginning stages of learning an instrument. *Psychology of Music*, **33**, 5–35.

Rayner K and Pollatsek A (1989). *The psychology of reading*. Erlbaum, Hillsdale, NJ.

Schon D and Besson M (2005). Visually induced auditory expectancy in music reading: a behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, **17**(4), 694–705.

Sloboda JA (1974). The eye–hand span. An approach to the study of sight reading. *Psychology of Music*, **2**, 4–10.

Sloboda JA (1976). The effect of item position on the likelihood of identification by interference in prose reading and music reading. *Canadian Journal of Psychology*, **30**, 228–236.

Sloboda JA (1977). Phrase units as determinants of visual processing in music reading. *British Journal of Psychology*, **68**, 117–124.

Sloboda JA (1984). Experimental studies of music reading: a review. *Music Perception*, **2**, 222–236.

Truitt FE, Clifton C, Pollatsek A and Rayner K (1997). The perceptual span and the eye–hand span in sight-reading music. *Visual Cognition*, **4**(2), 134–161.

Waters AJ, Townsend E and Underwood G (1998). Expertise in musical sight-reading: a study of pianists. *British Journal of Psychology*, **89**, 123–149.

Watkins JG (1942). *Objective measurement of instrumental performance*. Columbia University Teachers College, New York.

Weaver HE (1943). A study of visual processes in reading differently constructed musical selections. *Psychological Monographs*, **55**, 1–30.

Yumoto M, Matsuda M, Itoh K, Uno A, Karino S, Saitoh O, Kaneko Y, Yatomi Y and Kaga K (2005). Auditory imagery mismatch negativity elicited in musicians. *Neuroreport*, **16**(11), 1175–1178.

# CHAPTER 33

# Performing from memory

Roger Chaffin, Topher R. Logan and Kristen T. Begosh

WHAT is the difference between 'learning' a new piece of music and 'memorizing' it? Both involve memory, but of different kinds. The memories that develop spontaneously while learning a new piece take the form of *associative chains* in which each passage cues the memory of what comes next. Deliberate memorization transforms the motor and auditory chains, making them *content addressable*. A memory is content addressable if you can ask yourself, e.g., 'How does the third repetition of the main theme go?', and the music comes to mind. Associative chains have a major weakness: to reach any link in the chain you have to start at the beginning. For a musician, this becomes a problem when something goes wrong in performance. Besides the embarrassment of starting over, there is the agony of wondering whether memory will fail again in the same place. Content-addressable memories avoid this problem. They can be located directly by thinking of the relevant location in the piece. In a memorized performance, content-addressable memory provides a safety net that permits recovery in case the associative chain breaks and the performance is disrupted.

Associative chains and content-addressable memories are learned in different ways and have different properties. Content-addressable memories are more likely to be explicit (conscious) and to involve declarative (language-based) knowledge *that* such-and-such is the case, whereas associative chains are more likely to be implicit (unconscious) and to involve procedural (motor-based) knowledge of *how* to do something. To memorize a piece of music for performance, the musician must smoothly integrate the two kinds of memory.

English has only the one term 'memory' to refer to these two very different mental processes. In everyday talk, musicians make the distinction by referring to 'learning' and 'memorizing'. There is the potential for confusion here. For example, how should we understand musicians who say that they do not memorize, for whom memorization is 'something that just happens' (André-Michel Schub), 'a subconscious process' (Harold Bauer), that is 'very simple' (Walter Gieseking), 'like breathing' (Jorge Bolet); how should we understand Jorge Bolet, when he says that he memorized Liszt's *Mephisto Waltz* in 75 minutes (Chaffin *et al.* 2002, Chapter 3)?

For the musician, the relevant question is whether memory will be reliable on stage. Jorge Bolet probably did not mean that he was ready to go on stage and perform. Professional performers may sometimes find themselves in the position of having to perform at short notice, but they do not normally choose to. It is risky. What happens if something goes wrong? If the memory is in the form of an associative chain, then the only recourse is to start again at the beginning of the chain. This kind of catastrophic memory failure is an unfortunate staple of student recitals. Students often make the mistake of assuming that because they can get through a piece without the score in the studio, they can do the same in live performance. They do not appreciate that the associative chain is just the first step; much more work is

needed to create a reliable, content-addressable memory.

Experienced performers know better; they give themselves a safety net. Memory failures are inevitable in live performance. A performer may go for years without a problem but, with enough performances, eventually it will happen. The important thing is to recover gracefully. Experienced performers do not stop and go back to the beginning. They go on. They have a mental map of the piece that allows them to keep track of where they are as the performance unfolds. The map provides landmarks where they can restart the performance if necessary (Chaffin *et al.* 2002, Chapter 9). When something goes wrong, the expert jumps to the next landmark and the performance continues. Most of the time, the audience is not even aware that anything went wrong. Memorization is the creation of this safety net.

Our account of memory for performance builds on the standard view of memory described by Bob Snyder in Chapter 10 of this volume (also Ginsborg 2004). We focus on the role of serial chaining and content addressability. Two areas of the episodic memory literature are particularly relevant to our discussion: oral traditions and expert memory. In oral traditions, materials such as children's rhymes and folk songs are handed down from one generation to another without the benefit of written records, often for hundreds of years. We will draw on David Rubin's (1995, 2006) analysis of this phenomenon and on his *basic system theory of episodic memory* to describe the role of different types of memory (auditory, motor, visual, emotional, structural, and linguistic) in associative chaining.

The second area of psychological research that we will draw on is the study of expert memory. The history of music is filled with examples of extraordinary feats of memory and these are often exhibited as evidence that the musician in question possessed some special gift or talent. For example, the young Mozart's writing out of Allegri's *Miserere* from memory was seen, at the time and ever since, as evidence of his genius (Chaffin *et al.* 2002, p. 66). The conclusion of careful study by psychologists is that such feats are *not* the product of a special talent for memorization but are the entirely predictable result of years of training and the effective use of retrieval schemes (Ericsson and Charness 1994). Expert memorists develop retrieval strategies to make their memories content addressable so that they can find the information they need when they need it (Ericsson and Kintsch 1995). We will use Anders Ericsson's theory of expert memory to explain how experienced performers memorize, as opposed to simply learn, a new piece.

## Associative chaining

Music performance relies heavily on associative chaining: what you are playing reminds you of what comes next. In this respect, memory for music is similar to memory for rhymes, songs, and poems. In each case, the task of memorization is made easier by the fact that what comes next is heavily constrained by what precedes it (Rubin 1995, 2006). For example:

There was a young man of Japan,

Whose limericks never would ____.

The possibilities for the next word in this little verse are constrained by syntax to verbs, by semantics to verbs that can take 'limericks' as agent, and by rhyme to words that end in '__an'. These multiple constraints narrow the possibilities. One does not have to have heard this limerick before to know that the missing word is 'scan'.

## The role of schemas[1]

How do we know that the second line must rhyme with the first? We recognize it as a limerick. Even if the word 'limerick' were not explicitly mentioned in the second line, we would recognize the characteristic formulaic opening and 'Te dee-ya, te dee-ya, te-dum' rhythm. Rapid recognition of this sort is a normal feature of memory: it is the same when we recognize a strawberry, a rain storm, or a birthday party. Information in long-term memory is represented in the form of *schemas* that summarize our previous experience and tell us what to expect. The schema for

[1] The Latin root of 'schema' dictates a plural form 'schema' or 'schemata'. We prefer the more regular, colloquial form 'schemas'. All three forms are in common use in the psychological literature.

limericks tells us that the second line will rhyme with the first. Similarly, schemas tell us to eat strawberries, use umbrellas when it rains, and give presents at birthday parties.

Contrary to popular belief, memory is not a vast storehouse containing exact records of a myriad of original events (Mandler 1984). Memories for specific events (episodic memories) are *reconstructed* at each remembering on the basis of schematic (semantic) knowledge representing generic memories. Schemas allow us to economically recall our past in enormous detail, but this ability has a price. When we take the trouble to check, many of the details turn out to be wrong. The same schematic frameworks that allow us to remember are also a source of distortion. We remember the gist, and fill in the details, systematically misremembering in the process.

Given the general fallibility of memory, musicians' routine reliance on rote memory seems remarkable. How is accurate recall possible, if memory for a piece must be reconstructed from generic musical schemas each time it is played? We would expect performances to be full of mistakes as the musician replaces the exact notes provided by the composer with the musical gist based on generic knowledge of harmonic, melodic, metric, and rhythmic patterns.

The answer comes from studies of how memory functions in oral traditions (Rubin 1995, 2006). In non-literate cultures, oral traditions such as ballads, epic poems, and religious enactments often remain stable across centuries, indicating that memories for their performance have been transmitted, more or less verbatim, from one generation to the next across many years. How is this done? Like everyone else, the memories of bards, minstrels, and storytellers are reconstructed at each performance (Rubin 1995, 2006). Their performances *do* vary, but they are sufficiently consistent that the distortion is minimal, even across generations. This surprising level of accuracy is a product of *multiple constraints*. In every oral tradition studied, the material follows strict formal constraints on rhyme, rhythm, and alliteration. Our example of the limerick illustrates how these constraints operate. The multiple constraints of grammar, meaning, metre and rhyme reduce the possibilities available.

In music, similar constraints are provided by melody, harmony, metre, and rhythm. In addition, repetition is normally much more pronounced in music than in language and provides additional local constraints (Huron 2006, p. 229–231). All of these constraints combine to make the task of memory reconstruction easier. Knowing how different composers use the various conventions of each musical genre makes constraints more specific, and so memorization is easier for experts than for novices (Williamon and Valentine 2002). In song, the constraints of the musical and literary forms combine, making memorization easier when words or music are learned together rather than separately (Ginsborg and Sloboda 2007).

## Multiple memory systems

Another feature that music shares with the materials transmitted in oral traditions is that it is recalled as part of a *performance*. Performance calls on the many different cognitive and bodily systems involved in action, each of which lays down its own memory traces, subject to its own schemas. These provide multiple retrieval cues, making memory for performance more robust than memory for text (Rubin 1995, 2006). On the basis of cognitive and neurological evidence, Rubin (2006) has proposed a model of memory in which multiple memory systems contribute to episodic memories. We will restrict our description to those systems most relevant to musical performance: auditory, motor, visual, emotional, narrative, and linguistic memory. A memorized performance is generated through the interaction of the information available in each system. For example, musicians find it easier to play a piece of music they have memorized than to write it out because playing provides memory cues from the motor system that are absent when writing out the score (Chaffin and Logan 2006). Just as in the limerick example, the multiple retrieval cues from the different systems interact to reduce the range of possibilities.

### Auditory memory

The history of Western music is replete with stories of musicians who were able to hear entire works in their heads (Deutsch and Pierce 1992). Psychological studies confirm this ability in people with or without musical training, and

have begun to specify the form in which the auditory information is stored (Halpern 1992). These studies confirm that people can 'hear' a melody in their heads, usually without accompanying imagery from other modalities, suggesting that the ability is based on an independent auditory memory (Reisberg 2001, Chapter 11). Neuropsychological studies provide further confirmation of the existence of a separate auditory subsystem (e.g. Fornazzari *et al.* 2006). In performance, auditory memory tells the musician what comes next, providing cues to elicit the music from memory, while also letting the musician know that things are on track (Finney and Palmer 2003).

We demonstrated one important way that auditory memory helps constrain recall when we showed how the rhythmic and rhyming schema for limericks narrow the range of possibilities for recall. In a similar vein, Rubin (2006) argued that the organization of ballads into stanzas with invariant metric and rhyming schemas has been largely responsible for the preservation of an oral tradition in North Carolina that is directly traceable to European ballads of the Middle Ages. In similar ways, schemas for standard rhythmic, melodic and harmonic patterns allow musicians to remember music better than non-musicians (Halpern and Bower 1982). Auditory memory appears to contain information about both pitch contour (relative pitch) as well as pitch category (absolute pitch), since people can sometimes recall music in the same key as the original (Dowling 1978; Halpern 1989).[2]

## Motor memory

Motor memory allows actions to be executed automatically by providing kinaesthetic memory of the sensory feedback from joints, muscles, and touch receptors. Although motor skills have been studied since the earliest days of experimental psychology (Adams 1987), the contribution of the motor system to memory was neglected by early cognitive theories of memory, entering mainstream cognitive theorizing only with the distinction between procedural (motor) and declarative (conceptual) memory (Anderson 1978; Squire 1987). More recently, the discovery of 'mirror neurons' in the motor system that respond to seeing the corresponding action performed by another (Rizzolatti and Craighero 2004) has reaffirmed earlier claims that the motor and sensory systems are intimately linked (Liberman and Mattingly 1985). The motor system is still largely treated as a completely separate system (Rosenbaum 2005), however, and the study of its contribution to memory is still in its infancy, under the rubric of 'embodied cognition' (Glenberg 1997).

Musicians talk about motor memory as being 'in the hands'. Perhaps the most important feature of motor memory for musicians is that it is implicit (unconscious). Musicians know *that* they can play a particular piece (declarative knowledge), but the knowledge of *how* to play can only be exhibited by actually playing (procedural knowledge). This is a source of anxiety, and may lead to over-practice. Playing seems to be the only way to reassure oneself that memory for a piece is intact. Mental practice provides an alternative but requires explicit memory. To make motor memory explicit, actions must be recoded in propositional form so that they can be rehearsed in working memory as a thought of the general form, 'Next, do this'. This kind mental instruction is a form of *linguistic memory*, discussed below.

Motor memory provides the clearest examples of associative chaining in memory; each action in the series cues the next. This is what makes motor memories implicit: to be accessed, they must be performed. Actions can, of course, be cued in other ways. People stand up for national anthems, shake hands when introduced, and remember to stop at the grocery store on the way home from work. This last example is different from the others, because the action is directed by a cue that we provide for ourselves. The cue is a thought in working memory, e.g., 'Take this exit'. This is the same kind of self-cuing that makes mental rehearsal possible. Chaffin, Imreh & Crawford (2002) introduced the term *performance cues* to refer to the use of this kind of cue in music performance. Setting up performance cues is the main work involved in memorizing for performance and is described below in the sections on *expert memory* and *performance cues.*

[2] Remembering pitch contour requires the use of spatial imagery which Rubin (2006) identifies as a separate basic system.

## Visual memory

Visual memory of the score is used mainly in the early stages of memorizing, while visual memory of the hands on the instrument becomes more important in the later stages. The role of visual memory for the score is evident in the difficulty that some musicians experience when working with a different edition of a score from the one they used to initially learn a piece (e.g. Chaffin *et al.* 2002, p. 37). A new score is difficult to work with because the visual information it provides is different from the musician's visual memory. It is common to remember the location of a passage on the page, a form of spatial imagery. Student musicians frequently use the spatial organization of music by pages rather than the formal structure of the piece to organize their practice (Williamon and Valentine 2000).

As in the general population, there are large individual differences between musicians in their subjective experience of visual memory. Some musicians report having 'photographic' memories, while others say that their visual memories are poor or unhelpful. For example, Myra Hess described how she could 'see' and 'read' the printed page when playing from memory, whereas Alfred Brendel reported that his memory was 'not visual at all' (Chaffin *et al.* 2002, pp. 37–41). These reports probably reflect real differences in the detailed information available in visual memory (Reisberg 2001, Chapter 11). Reports of visual images are misleading, however, in two ways. First, mental images are *not* pictures. Images are not neutral, objective depictions of reality but are organized interpretations that reflect the way that the original was understood. To discover whether you misread a note you cannot inspect your mental image of the score, you have to go back and look at the real thing. Second, people who report having no visual memory still have spatial memories which are stored in a separate system. While spatial memory does not provide a vivid subjective experience, it does provide information about the location of notes on the page.[3] So a person could experience no visual imagery but still be disrupted by using a different edition of a score.

## Emotional memory

Memories for emotional events are formed more easily and are less likely to be forgotten than non-emotional memories (Bower 1981; Talmi *et al.* 2007). This is as true for music as any other material (Schulkind *et al.* 1999). The positive effects of emotion on memory are disrupted by damage to neural areas involved in emotion (Greenberg and Rubin 2003). Together, these findings are the basis for identifying emotional memory as a separate system.[4] It seems clear that the performer's visceral response to the music contributes to musical memory. We have observed that musicians find it difficult to play from memory when asked to perform without expression and surmise that playing without expression eliminates emotional cues that normally contribute to the retrieval of the music from memory.

## Structural memory

We suggest that structural memory is the musical equivalent of Rubin's *narrative memory*: memory for the overall sequential organization and goal structure of a story or biography. Memories for events are organized by schemas that connect temporal series of discrete actions through narrative structures based on the goals of the actors involved (Mandler 1984). Although often expressed in language, narrative structure can be expressed in a variety of forms including pictures, cartoons, silent films, dreams, dance, and mime (Rubin 2006). In the Western classical tradition, the same kind of narrative structure is responsible for the hierarchical organization of a piece into sections and subsections based on melodic, harmonic, and metrical structures. In preparing a piece of music, experienced musicians analyse these structural properties and use them to organize both their practice and their memories (Chaffin and Imreh 1997,

[3] We have not followed Rubin (2006) in singling out spatial imagery as a separate form of memory because in music performance it appears to operate across modalities, binding together representations in the auditory, motor, and visual systems

[4] This is an oversimplification since emotion is a complex and varied phenomenon that draws on multiple neural systems (Rubin 2006).

2002; Chaffin *et al.* 2002; Hallam 1995; Williamon and Valentine 2002).

In 'programme' music, narrative organization is explicitly applied to music. Despite the resistance of critics to providing a storyline for every piece of music, the ease with which music lends itself to this kind of treatment suggests that musical and narrative structure share common roots. We suggest that they stem from the same cognitive system; musical form and the storyline of a musical programme are both manifestations the underlying ability to identify large-scale structural relations between events.

An important difference between narrative and musical structure is that the former seems to be easier to perceive. Even young children are sensitive to narrative structure (Nelson and Fivush 2004), whereas sensitivity to musical structure develops slowly with musical training (Williamon and Valentine 2002) and is not always found, even with experienced musicians (Chapter 10 this volume, pp 113–115). The difference may be due to the fact that in our culture people generally have a lot more experience telling stories than they do playing music.

### Linguistic memory

The mental instructions that experienced performers use to remind themselves what to do at key points in a performance are a form of *linguistic memory* (Chaffin *et al.* 2002).[5] These instructions do not necessarily involve words. They are stored in an abstract 'subject-predicate' (propositional) form that usually points to other modalities (motor, auditory, visual, and emotional memories). However, their propositional form means that they can normally be glossed in words, e.g., 'Hold back' or 'Now, like this' (Englekamp 2001).

An important characteristic of linguistic memories is that they can be rehearsed in working memory, where they can serve to direct other mental processes. When the activity of other cognitive systems is re-described in language, the inner speech that results provides a means of mental control that can be used to implement plans and strategies (Reisberg 1992, p. viii; Rubin 2006).[6] Rehearsing a mental instruction in working memory broadcasts it throughout the nervous system, automatically activating other systems and coordinating their activity (Barrs 1988).[7] As we noted above, this ability can be used for mental rehearsal or to recover if the associative chain of a memorized performance breaks.

## Content-addressable memory

Associative chaining works well so long as the chain is intact. If the performance stops, however, the chain is broken, and then memory failure is complete and catastrophic. The performer can only go back to the beginning and start over. To avoid such ignominy, experienced performers prepare a safety net that provides other options; they prepare multiple starting points.

When you want to sing happy birthday, you simply think, 'Happy Birthday', and start singing. The verbal label acts as a retrieval cue for the start of the song and the rest is then cued by associative chaining. Now imagine that you want to start at the last line. Most of us cannot do this immediately. We have to start at the beginning and run through. Once we have the last line in working memory, however, we can easily set up a new starting point by thinking, 'Start of the last line' as we sing. A few repetitions to strengthen the associative link between the new cue and singing the last line and we have a new starting point. Any time we want to start at the last line, we can now simply think, 'Start of the last line' and start singing.

We have set up a new performance cue, making this place in the music content addressable (Chaffin *et al.* 2002). Simply thinking of the cue now activates the memories needed to start singing.

[5] The first author has previously referred to both linguistic and structural memory as 'conceptual' or 'declarative' memory (e.g., Chaffin *et al.* 2002; Chaffin and Imreh, 2002). The present terminology represents a refinement of that classification.

[6] The important role of inner speech in mental control has been noted by many psychologists including Pavlov, Watson, Vygotsky, and Piaget.

[7] The process of directing and monitoring our own mental operations in this way may be responsible for uniquely human qualities of conscious experience (Dennett 1991).

The performance cue lacks the multidimensional richness of the associative chain, where the next link was cued by sound, action, and emotion. What it lacks in richness, it makes up for in flexibility. You can now think of the passage at any time, without running through the whole piece from the beginning.

We will focus on two aspects of this strategy. First, when applied to a long piece of music, the strategy of creating multiple starting points has many similarities with how experts memorize in other domains that have nothing to do with music. We will describe these similarities in the next section. Second, there is a risk involved in setting up additional starting points. Thinking about what you are doing can interfere with skilled performance, a phenomenon known as *choking* (Beilock and Carr 2001). We will describe how experienced musicians avoid this problem in the section on *performance cues.*

## Expert memory

Experts in any domain memorize with a facility that seems superhuman (Gobet and Simon 1996). Musicians are no exception; as we have already noted, their biographies are full of tales of amazing memory feats. The abilities of other expert memorists have been attributed to the use of highly practiced retrieval strategies by skilled memory theory (Chase and Ericsson 1982) and its extension, long-term working memory theory (Ericsson and Kintsch 1995). These theories are based on the study of domains such as chess boards (Chase and Simon 1973), digit strings (Thompson *et al.* 1993), and dinner orders (Ericsson and Oliver 1989) that are very different from music performance: structural and linguistic memory are primary and associative chaining of motor and auditory memories play minor roles. Despite the differences, the principles of expert memory established in these domains apply to music performance because experienced musicians also rely on structural and linguistic memory to provide a safety net in case the chain of motor and auditory memories breaks (Chaffin and Logan 2006).

The feats of expert memorists can be explained in terms of three principles: meaningful encoding of novel material, use of a well-learned retrieval structure, and extended practice to decrease the time needed for retrieval from long-term memory (Ericsson and Kintsch 1995). The same three principles apply to expert music performance (Chaffin *et al.* 2002; Krampe and Ericsson 1996). First, experts' knowledge of their domain of expertise allows them to make use of schematic knowledge already stored in memory to organize information into larger chunks (Tulving 1962). For a musician, these include familiar patterns like chords, scales, and arpeggios, whose practice forms an important part of every musician's training (Halpern and Bower 1982). Second, expert memory in any domain requires a retrieval scheme to organize the cues that provide access to the chunks of information in long-term memory (Ericsson and Oliver 1989). For a musician, the formal structure of the music conveniently provides a ready-made hierarchical organization to serve as a retrieval scheme. For example, Figure 33.1 shows how the hierarchical organization of JS Bach's Italian Concerto (Presto) into movements, sections, subsections, and bars was used by a pianist to organize her memory for the piece (Chaffin *et al.* 2002). The third principle of expert memory is that prolonged practice is required to bring the speed of operation of a memory retrieval scheme like the one in Figure 33.1 up to the speed needed to guide behaviour (Ericsson and Kintsch 1995). For the musician, this involves practising memory retrieval until it is rapid and reliable enough to keep pace with the performance.

Rapid memory retrieval is important in music performance to prevent the hands from 'running away' as the retrieval of procedural knowledge by associative chaining outpaces the slower, content-addressable retrieval of declarative knowledge. The smooth integration of the two systems creates 'long-term working memory' (Ericsson and Kintsch 1995). Practice is needed so that the performance cue for what comes next arrives in working memory at just the right moment, before the corresponding motor sequences, but not so far in advance as to distract from the execution of the preceding passage and cause 'choking'.

The interplay of the two retrieval systems is illustrated in Figure 33.2. The figures shows the two routes by which memory for a piece of music

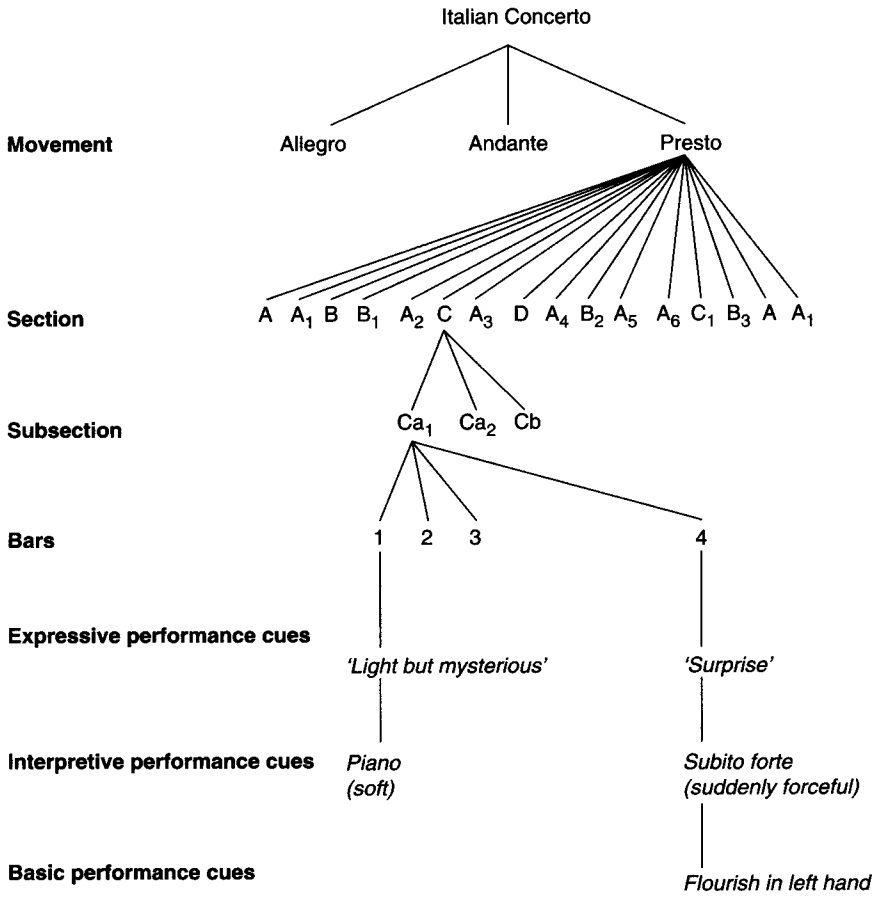



**Fig. 33.1** Hypothetical hierarchical retrieval scheme 'unpacked' for Section C of the Presto from the Italian Concerto by JS Bach. Main themes (sections) are represented by capital letters. Section C is 'unpacked' into subsections (Ca1, Ca2, Cb). Subsection Ca1 is further 'unpacked' into its performance cues. (From Chaffin R., Imreh G., & Crawford M. (2002). *Practicing perfection: Memory and piano performance*. Erlbaum, Mahwah NJ, p. 200. Copyright © 2002. Adapted with permission).

can be retrieved. At the bottom of the figure are the serial associations set up while learning to play the piece. These associations, based on schema for rhythm, meter, harmony, and melody, directly link each passage with the next. Each passage is cued only by the preceding passage. Direct, content-addressable access is provided by a second retrieval system, shown at the top of the figure. Here, a hierarchical retrieval organization, similar to that in Figure 33.1, provides direct access to any section of the piece. Performance cues embedded in this organization provide possible 'starting points' in case things go wrong in performance.

## Performance cues

One of the main challenges in memorizing for performance is to integrate the two retrieval systems. As one pianist put it in talking about her learning the Italian Concerto (Presto) by JS Bach:

> My fingers were playing the notes just fine. The practice I needed was in my head. I had to learn to keep track of where I was. It was a matter of learning exactly what I needed to be thinking of as I played, and at exactly what point so that as I approached a switching point I would automatically think

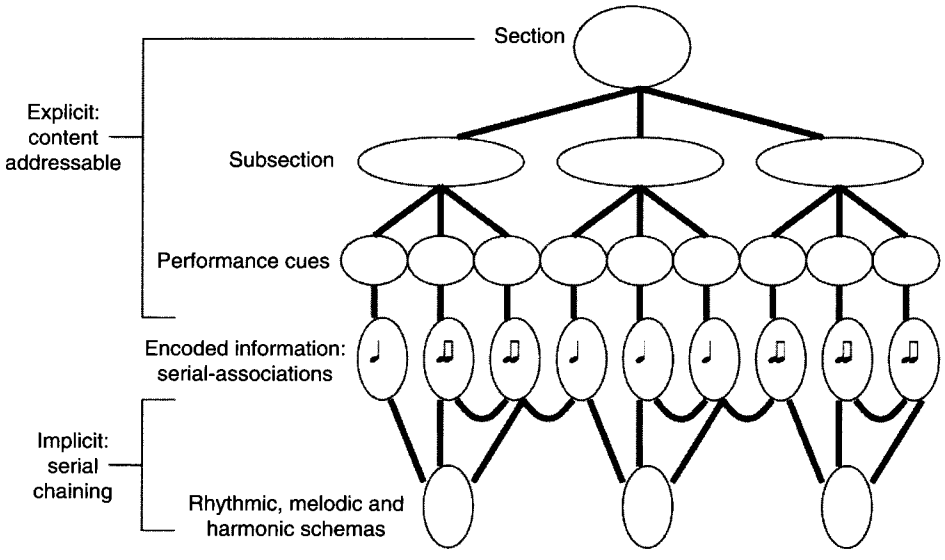



**Fig. 33.2** Model of episodic memory for music performance showing separate retrieval organizations for access by content address and by serial associative chaining. (From Ericsson K. A., & Kintsch W., Long-term working memory. Psychological Review, 102, 211–245. Copyright © 1995, The American Psychological Association. Adapted with permission).

about where I was, and which way the switch would go.

Chaffin *et al.* (2002, p. 224)

The musician was talking about the need to practice performance cues. Performance cues are landmarks in the mental map of a piece that an experienced musician maintains in working memory during performance. Because they can be accessed both by serial cuing and directly, by address, they provide a safety net in case serial cuing breaks down. Careful preparation of performance cues makes it possible for soloists to reliably perform challenging works from memory on the concert stage. By repeatedly paying attention to performance cues during practice, the musician ensures that that they become an integral part of the performance, coming to mind effortlessly as the music unfolds. The performer remains mindful of these aspects of the performance while allowing others to be executed automatically. When things go smoothly, performance cues are a source of spontaneity and variation in highly polished performances (Chaffin *et al.* 2007). When things go wrong, they provide places at which the soloist can recover and go on.

Performance cues point to different types of memory according to which aspect of the music they address.[8] *Structural* cues are critical places in the formal structure, such as section boundaries. *Expressive* cues represent musical feelings, e.g., excitement.[9] *Interpretive* cues refer to musical gestures, such as changes of tempo or dynamics. *Basic* cues point to motor memory for critical details of technique, e.g., a fingering that sets the hand up for what follows. Musicians are likely to agree on the musical structure of a piece.[10] They are likely to differ, however, on other cues that are more specific to the performer or instrument. For example, basic performance cues for a cellist include decisions about intonation, bowing, and changing string that are not relevant for a pianist (Chaffin *et al.*

[8] The organization of cues into types here is descriptive and somewhat arbitrary. Other descriptive organizations are possible.

[9] Musical structure and expression are necessarily linked, but expressive cues do not always coincide with structural boundaries (Chaffin *et al.* 2008).

[10] The musical structure of some pieces may, however, be understood in more than one way, e.g., Ginsborg *et al.* (2006).

2008). For solo works, the only performance cues required are those for the individual musician, while for ensemble performance the musicians must also establish shared performance cues to coordinate their actions (Ginsborg *et al.* 2006).

Performance cues are content addressable through their location in the hierarchical organization of the piece (see Figure 33.1). In learning a new piece, the musician moves up and down the hierarchy, attending to each level of organization and each type of cue in turn (Chaffin *et al.* 2006; Williamon *et al.* 2002). Like experts in other fields, who approach a new problem by looking at the 'big picture', experienced musicians approach the task of learning a new piece by getting an overall artistic image of how the music should sound, focusing on structural and expressive cues (Neuhaus 1973; Chaffin *et al.* 2003). Beyond this commonality, the order in which different types of performance cue are practiced appears to depend on the individual, piece, and situation. What all the experienced performers that have been studied to date have in common is that they practice performance cues. This provides them with a safety net for times when associative cuing fails (Chaffin and Logan 2006).

## Conclusion

Though it has a long history in Western classical music, playing from memory is often a source of great anxiety for performers. Pianist Janina Fialkowska talked of the 'terror of forgetting'. Lazar Berman reported that, 'Every time I play in front of an audience, it is a very important and difficult affair, both physically and spiritually. I am never sure that it is going to end well' (Chaffin *et al.* 2002, Chapter 3). Anton Rubenstein wrote that fear of memory failure 'inflicted upon me tortures only to be compared with those of the Inquisition' (Rubenstein 1969, p. 18).

Given the costs involved in playing from memory and its long tradition, stretching back 175 years, it might be expected that musicians would have developed a systematic understanding of the problems involved. This has not happened (Aiello and Williamon 2004). Individual musicians know a great deal about strategies for memorization (Hallam 1995, 1997), but this knowledge is conveyed from teacher to student through an apprenticeship system of training that makes it unavailable for systematic analysis. Memorization is viewed as an individual and mysterious process. It is up to each person to find their own method (Ginsborg 2002). This is regrettable and unnecessary. Memory varies no more from one person to another than any other trait or capacity. Beneath a superficial diversity, the cognitive and neurological systems involved in memory are common to all human beings. In this chapter we have described those aspects that are most relevant to performing music from memory.

## Acknowledgements

We thank Mary Crawford, Jane Ginsborg, and Aaron Williamon for helpful comments on previous versions of this chapter.

## References

Adams J (1987). Historical review and appraisal of research on the learning and transfer of human motor skills. *Psychological Bulletin*, **101**, 41–74.

Aiello R and Williamon A (2004). Memory. In R Parncutt and GE McPherson, eds, *The science and psychology of music performance*, 167–182. Oxford University Press, New York.

Anderson JR (1978). Arguments concerning representations for mental imagery. *Psychological Review*, **85**, 249–277.

Baars BJ (1988). *A cognitive theory of consciousness*. Cambridge University Press, Cambridge University Press.

Beilock SL and Carr TH (2001). On the fragility of skilled performance: what governs choking under pressure? *Journal of Experimental Psychology: General*, **130**, 701–725.

Bower GH (1981). Mood and memory. *American Psychologist*, **36**, 129–148.

Chaffin R and Imreh G (1997). 'Pulling teeth and torture': musical memory and problem solving. *Thinking and reasoning*, **3**, 315–336.

Chaffin R and Imreh G (2002). Practicing perfection: piano performance as expert memory. *Psychological Science*, **13**, 342–349.

Chaffin R, Imreh G and Crawford M (2002). *Practicing perfection: memory and piano performance*. Erlbaum, Mahwah, NJ.

Chaffin R, Imreh G, Lemieux AF and Chen C (2003). 'Seeing the big picture': piano practice as expert problem solving. *Music Perception*, **20**, 465–490.

Chaffin R, Lemieux A and Chen C (2006). Spontaneity and creativity in highly practiced performance. In I Deliège and GA Wiggins, eds, *Musical creativity: multidisciplinary research in theory and practice*, 200–218. Psychology Press, London.

Chaffin R, Lemieux A and Chen C (2007). 'It's different every time I play': spontaneity in highly prepared musical performance. *Music Perception*, **24**, 455–472.

Chaffin R, Lisboa T, Logan T and Begosh KT (in press). Preparation for cello performance: The role of performance cues. *Music Psychology*.

Chaffin R and Logan T (2006). Practicing perfection: how concert soloists prepare for performance. *Advances in Cognitive Psychology*, **2**, 113–130.

Chase W and Ericsson KA (1982). Skill and working memory. In GH Bower, ed., *The psychology of learning and motivation*, 1–58. Academic Press, New York.

Chase W and Simon H (1973). Perception in chess. *Cognitive Psychology*, **4**, 55–81.

Dennett DC (1991). *Consciousness explained*. Little, Brown and Co., New York.

Deutsch D and Pierce JR (1992). The climate of auditory imagery and music. In D Reisberg, ed., *Auditory imagery*, 237–260. Earlbaum, Hillsdale, NJ.

Dowling WJ (1978). Scale and contour: two components of a theory of memory for melodies. *Psychological Review*, **85**, 341–354.

Englekamp J (2001). *Memory for action*. Psychology Press, Hove, England.

Ericsson KA and Charness N (1994). Expert performance: its structure and acquisition. *American Psychologist*, **49**, 725–747.

Ericsson KA and Kintsch W (1995). Long-term working memory. *Psychological Review*, **102**, 211–245.

Ericsson KA and Oliver WL (1989). A methodology for assessing the detailed structure of memory skills. In AM Colley, JR Beech, eds, *Acquisition and performance of cognitive skills*, 193–215. John Wiley and Sons, Oxford.

Finney SA and Palmer C (2003). Auditory feedback and memory for music performance: sound evidence for an encoding effect. *Memory and Cognition*, **31**, 51–64.

Fornazzari L, Nadkarni S and Miranda D (2006). Preservation of episodic musical memory in a pianist with Alzheimer disease. *Neurology*, **66**, 610–611.

Ginsborg J (2002). Singing by heart: memorisation strategies for the words and music of songs. In JW Davidson, ed., *The music practitioner: exploring practices and research in the development of the expert music performer, teacher and listener*, 149–160. Ashgate Press, London.

Ginsborg J (2004). Strategies for memorizing music. In A Williamon, ed., *Musical excellence: strategies and techniques to enhance performance*, 123–142. Oxford University Press, Oxford.

Ginsborg J, Chaffin R and Nicholson G (2006). Shared performance cues in singing and conducting: a content analysis of talk during practice. *Psychology of Music*, **34**, 167–194.

Ginsborg J and Sloboda J (2007). Singers' recall for the words and melody of a new, unaccompanied song. *Psychology of Music*, **35**, 421–440.

Glenberg AM (1997). What memory is for? *Behavioral and Brain Sciences*, **20**, 1–55.

Greenberg DL and Rubin DC (2003). The neuropsychology of autobiographical memory. *Cortex*, **39**, 687–728.

Hallam S (1995). Professional musicians' approaches to the learning and interpretation of music. *Psychology of Music*, **23**, 111–128.

Hallam S (1997). The development of memorization strategies in musicians: implications for education. *British Journal of Music Education*, **14**, 87–97.

Halpern AR (1989). Memory for the absolute pitch of familiar songs. *Memory and Cognition*, **17**, 572–581.

Halpern AR (1992). Musical aspects of auditory imagery. In D Reisberg, ed., *Auditory imagery*, 1–27. Erlbaum, Hillsdale, NJ.

Halpern AR and Bower GH (1982). Musical expertise and melodic structure in memory for musical notation. *American Journal of Psychology*, **95**, 31–50.

Huron D (2006). *Sweet anticipation: music and the psychology of expectation*. MIT Press, Cambridge, MA.

Krampe RT and Ericsson KA (1996). Maintaining excellence: deliberate practice and elite performance in young and older pianists. *Journal of Experimental Psychology: General*, **125**, 331–359.

Liberman AM and Mattingly IG (1985). The motor theory of speech perception revised. *Cognition*, **21**, 1–36.

Mandler G (1984) *Stories, scripts, and scenes: aspects of schema theory*. Erlbaum, Hillsdale, NJ.

Nelson K and Fivush R (2004). The emergence of autobiographical memory: a social cultural developmental theory. *Psychological Review*, **111**, 486–511.

Neuhauss H (1973). *The art of piano playing*. Praeger Publishers Inc., New York.

Reisberg D (1992). *Auditory imagery*. Erlbaum, Hillsdale, NJ.

Reisberg D (2001). *Cognition: exploring the science of the mind*, 2nd edn. WW Norton, New York.

Rizzolatti G and Craighero L (2004). The mirror-neuron system. *Annual Review of Neuroscience*, **27**, 169–192.

Rosenbaum DA (2005). The Cinderella of psychology: the neglect of motor control in the science of mental life and behavior. *American Psychologist*, **60**, 308–317.

Rubin DC (1995). *Memory in oral traditions: the cognitive psychology of epic, ballads, and counting-out rhymes*. Oxford University Press, New York.

Rubin DC (2006). The basic-system model of episodic memory. *Perspectives on Psychological Science*, **1**, 277–311.

Rubinstein A (1969). *Autobiography of Anton Rubinstein: 1829–1889*, trans. A. Delano. Haskell House Publishers, New York (Original work published 1890).

Schulkind MD, Hennis LK and Rubin DC (1999). Music, emotion, and autobiographical memory: they're playing your song. *Memory and Cognition*, **27**, 948–955.

Squire LR (1987). *Memory and brain*. Oxford University Press, New York.

Talmi D, Schimmack U, Paterson T and Moscovitch M (2007). The role of attention and relatedness in emotionally enhanced memory. *Emotion*, **7**, 89–102.

Thompson CP, Cowan TM and Frieman J (1993). *Memory search by a memorist*. Erlbaum, Hillsdale, NJ.

Tulving E (1962). Subjective organization in free recall of 'unrelated' words. *Psychological Review*, **69**, 344–354.

Williamon A and Valentine E (2000). Quantity and quality of musical practice as predictors of performance quality. *British Journal of Psychology*, **91**, 353–376.

Williamon A and Valentine E (2002). The role of retrieval structures in memorizing music. *Cognitive Psychology*, **44**, 1–32.

Williamon A, Valentine E and Valentine J (2002). Shifting the focus of attention between levels of musical structure. *European Journal of Cognitive Psychology*, 493–520.

CHAPTER 34

# Movement and collaboration in musical performance

Jane W. Davidson

## Introduction

The body has a crucial role in the production and perception of musical performance that has been recognized for centuries. Indeed, the Ancient Greeks recognized the centrality of the body in all artistic performance expression, revealing that there was little to distinguish between the physical coordination required in music and dance. Speaking of all human activity, ethnomusicologist John Blacking (1977) argued in *The Anthropology of the Body* that a detailed understanding of the body set within the individual's cultural context can reveal subtle details about that person and how they interact within society. Increasingly since the 1970s scholarship has paid more attention to the role of the body in artistic activity, providing useful theories relating to how embodied thoughts and actions allow us to know and understand the world we inhabit through artistic performance activities like music and dance. For example, by the 1990s, in the area of critical musicology, Susan McClary's (1991) *Feminine endings* usefully addressed the function of the body through groundbreaking studies of popular artists, especially female artists like Madonna and kd Lang. She illustrated how it was as much our knowledge of these women as pop stars that influenced our experience of them, as it was their physical appearance and stage persona alongside the expressive affect experienced in their musical expressions that contributed to the performance: their bodies were recognized as a very special site of social exchange, thus highlighting the physical, social and cultural nexus.

Research in the field of music psychology on the body has reflected some of the recent social anthropology and critical musicology trends, and so has developed a strand of socially focused enquiry. These ideas have been central to the work of the current author and will be explored in this chapter. Generally, however, the origins of modern psychology of music work on the body has tended to focus more on the bodily control aspects of playing, such as how movements assemble to execute a musical task. Intriguingly, these motor programming studies have been interfaced with some studies on music learning so that evidence for the emergence of these skills has occurred. Further perceptual enquiries and movement tracking studies have explored the body movements used for expressive musical effects. Also, applied research projects for pedagogical results have explored the use of bodily movement and movement metaphors to enhance understanding of musical expressivity for performers, with parallels being drawn between movement effort and flow to depict qualities in the musical semantics. Finally, within the group music context, studies of the identification of types of bodily movement used for interpersonal collaboration, including the coordination of musical and extra-musical material (musical and general social dynamics), and how this information helps to

develop efficient collaborations have been undertaken. Thus, there has been a broad range of investigation, although the overall volume of work to date has been relatively small. In this chapter all the areas of existing enquiry will be explored. Though not comprehensive, the aim of this chapter is to explore the key existing music psychology investigations on this topic. It begins with research on motor programming, moving to more social aspects of performance and bodily movement, and finishing by considering musical collaborations.

## Motor programming

We know that the assemblage of all the subcomponent actions required for a musician's movement activity is a highly complex process. Recent writing by Jäncke (2006) explains that the movements musicians use emerge from 'a concatenation of mechanical and neural factors'. Though such movements have been of interest for many decades, understanding how the movement plans are organized has been a challenge to researchers (see Lashley 1951; Bernstein 1967; Wing 2002). There is an underlying 'degrees of freedom' problem that relates to understanding the vast potential array of muscle and nerve impulses that create actions. Since Bernstein (1967) suggested that motor programmes (memory representations) operate for classes of movements rather than individual movements, theories to account for movement systems have ranged from the brain operating as a central executive to control movements, to proposals where movement control arises in an open system where the movements are themselves self-organizing (see Dahl 2004 for a more detailed explanation of these ideas). Ideas have been broad-ranging: for example, it could be that movements trigger one another in a *chain* or a *chain complex* (sequential or in parallel), a *radial net* (this is, where one action may drive several others of equal weighting), a *hierarchy* (here one or more actions trigger more actions at the next level down through the branches of the hierarchy) or a *heterarchy* (multiple feedback loops across the hierarchical structure) (see Wing 2002 for details). The current reality is that we do not have a clear single theory to understand how the movements are organized.

For the musician, an understanding motor programming itself is not essential, however, knowing that the activity of the body involved in playing a piece of music depends on generating and embedding these motor programmes in memory by rehearsal is essential. Indeed, musicians need to rehearse their physical actions until they achieve fluency, or an incompressible minimum time to execute them. We know that it can take years for a professional musician to reach fluency. Sloboda and colleagues (1994) found this to be in the range of 3 hours of practice a day across the 10 childhood years of learning. Intriguingly, and offering support for the positive effects of rehearsal, researchers have revealed that practice results in enlarged representations of somatosenory and auditory cortex (Pantev *et al.* 2001), as well as the motor areas in their brains (Altenmüller and Gruhn 2002). Moreover, since these programmes are established in memory, thought practice can have benefit on physical skill, though it is not as profitable as actual physical practice (Palmer 2006). Additionally, Dalla Bella and Palmer (2004) have shown that pianists move their fingers approximately three to four events ahead of time, suggesting that an early retrieval for action is required, motor activity being an anticipatory and unfurling behaviour in performance.

The development of the motor programmes for economy leads to automaticity. In a complex physical activity like musical performance, automaticity is absolutely necessary because it offers fluency. This automaticity varies from the basic assemblage of the body to execute a single musical note through to rapid sequences of notes, often containing subtle musical effects such as counterpoint or inner voicings to be played out for expressive effects. In essence, the expert's performance movements are seemingly effortless and necessarily effortless: in achieving a high degree of automation (unconscious processing) in the action, the expert performer's conscious thought is relatively 'free' to deal with moment by moment modifications that may be necessary as the music is being performed. The novice, by contrast, requires full conscious attention on the biomechanical activity. Movements often seem cumbersome, the novice not being able to achieve the actions to fluent effect. This means that the

novice is not able to consider other performance-related concerns, for example adapting to new situations such as encountering a piano with a sticking key (see Lehmann and Davidson 2002 for more details).

Of course automaticity is required not only for movement fluency, but for the link between eye and hand if the musician is reading music, or between co-performers and timing, dynamic and overall expressive markings or ideas if the performance is to be coordinated and varying along the same dimensions for all players at all times. Clearly automaticity is a complex and multilayered skill. The layers of complexity were demonstrated in terms of musical timing by Shaffer (1984), who showed that when carrying out an activity like keyboard performance, there are specific timing effects that result as a consequence of the task demand. He noted, for example, that the typing performance by touch typists had specific timing profiles for certain combinations of letters. It is certain that the individual performer's body and instrument itself adds to this situation. For instance, a small pianist will have to develop his or her representations for playing a sequence of loud, large hand-span chords in a manner that will be slightly different from a larger person, even if the technicalities of playing the piano are based on the same principles and similar thought processes. Additionally, the musical systems developed within a specific culture and the types of sonic outputs created (e.g., Indian rag scales or Western tonality) are related to the ergonomics of the musical instrument used (Baily 1985). Wiesendanger *et al.* (2006) have shown that the bowing arm of the string player is highly constrained by the dynamical principles that operate between the bow and string and the arm controlling the bowing action. We can see the interaction of bodily constraints, ergonomics and cultural context within Western art music when a composition is played on different instruments of the same family; for instance, on a harpsichord versus the piano, or even from one piano to another. The differences in the size and shape of the instrument, plus the force required to play it, inevitably shape the physical approach required, and this in turn is influenced by the stylistic requirements of the music to be played—e.g., Baroque versus contemporary.

Shaffer also examined piano keyboard performances, noting that the timing profiles were additionally related to musical structure features, with, for example, a slowing always occurring at phrase boundaries. So, besides the requirements of achieving motor programming automaticity in order to execute sounds on instruments, it has been demonstrated that each music performer does not play in a purely mechanically efficient manner. The performer also has a set of representations that draw on knowledge and experience of musical style. Furthermore, as will be shown in the next section and throughout this chapter, cultural appropriateness also impacts significantly on how the movements are assembled and presented.

## Body movements for musically expressive effects

### Solo performers

A study by Davidson (1993) demonstrated that when the same piece of music was played with different expressive intentions (deadpan or withheld expression, with usual expression, and with exaggerated expression), the performers moved their bodies in identifiably different ways. Detailed tracking of a pianist's movements (Davidson 1994) showed that although the hands, arms, head and torso followed similar movement contours (those required to execute the music) across performances, there were significant differences in the scale of the movements. This suggested that the more highly expressive the piece, the larger and more ample the movements. The lesser the expressive intention, the smaller the movement. This finding seems to be rather like those seen in the actions of a speaker who frames and articulates speech, where big ideas or highly expressive or emotional states use much larger gestures than more constrained states (Kendon 1980).

From movement tracking data collected, it seems that it was impossible for the performers to use their bodies in a completely 'deadpan' manner. The same was found to be true of the musical sound, with some expressive feature of the music still being articulated, e.g., making slower and more pronounced movements at the boundary points, and surging forwards at a rising crescendo.

Such a finding would correspond with the idea that a musician's body is both generating and reacting to the sounds it is producing.

It is important to note that the performers observed in a series of studies from 1991 to 2007 (see Davidson 1993, 1994, 1995, 1997, 2002, 2006, 2007) demonstrated that the performers bodies moved in a rhythmical pulsing/swaying movement. Sometimes this was more or less apparent, but once movement tracks were made, the swaying could be traced across whatever piece was being played. For pianists, cellists, and tuba players this swaying occurred around the balancing point (fulcrum) of their sitting position; and for instrumentalists who were standing, this swaying was articulated from the balancing position of the weight over the feet, knees, and hips. An explanation for this movement was that the players were pivoting their bodies around a fulcrum through which all expressive information was being articulated. Such conclusions were drawn from extensive observations and detailed analyses of the performance movements generated. Indeed, it was discovered that within the overall swaying there were more local indications of the same sort of action (e.g., a rotating wrist) which conveyed very similar types of information (observers could detect from any single body part the expressive intention). It was concluded, therefore, that a physical centre or fulcrum to express mental representations relating to expressive intention may exist, and that these are then revealed in the body a hierarchical manner: overall movement sway, local body part movement (see in particular, Davidson and Correia 2002; Davidson 2005).

However, in addition to the overall performance movement—which seemed to be a combination of basic biomechanical action integrated with overall information about the music's general expressive intention—it was also discovered that musicians peppered their performances with many specific postures and gestures that had qualities akin to the types of non-verbal communication that accompany speech, e.g., those metaphorical and illustrative features which both generate and react to ideas being expressed (Kendon 1980). For example, head nodding in a series of rapid gestures that seem to mirror the music as it repeats a specific figure or sequence, potentially illustrating the repetition within the musical material. Other gestures seemed to have iconic value, that is, where some feature of the musical action was being described. For example, pianists creating the contour of the music being played by tracing this form in the air with the left hand whilst playing the notes with the right hand. This arguably is done to 'draw out' the smooth legato line that is being attempted in the music.

The specifically identifiable gestures used by individual instrumentalists seems to be rather limited in number, yet what these gestures had the potential to express appeared to be endlessly variable, depending on the context in which they are used. Indeed, Davidson's (2007) assessment of the pianist's movement vocabulary showed that the movements were limited so less than 20 movement types, and these happened to appear across all sorts of musical styles. Indeed, one style of music (Beethoven, for example) an emphatic 'wiggle' of the torso might illustrate an ornament in the music in one context, yet in another context (CPE Bach) this could signal the start of a long legato passage.

These gestures were person-specific, the 'wiggle' being specific to a case study pianist. Some of the other movements seemed more related to illustrative gestural purposes—e.g., nodding for emphasis—and there was a high degree of individual variability in how many of these gestures were used (Davidson 2005).

Since there are some schools of pedagogy which appear to focus only on the efficiency of the action of playing, it might be asked if these postures and gestures are really necessary for the optimal expression of the musical ideas in performance. A study by Davidson and Dawson (1995) requested that pianists should learn an especially composed piece which did not require any range movement as it was constrained to the central two octaves, yet when completely free to move, the pianists made quite extravagant movements in their performances, including both sways and identifiable gestures. Some were asked to learn the piece in a restraining harness to restrict them from moving around, though they were free to reach the keyboard and play all the notes required. In this condition, the performances were never as musically expressive or indeed as visually 'pleasing' as the freely

learned pieces. This indicates that freedom of movement is obviously important in the generation of musical expression.

So far, the indications have been that the gestural and postural movements used help to articulate the musical expression, and that movement gestures have roles of metaphorical function to articulate the musical sounds. It is possible to take these ideas one step further to suggest that musical material itself operates like a virtual person—a social inter-actor—with the performer. So, in a score-indicated fortissimo, a violinist may nod her head as she plays loudly. This action could be an illustration of the force of the movement required, or she may be 'agreeing' with the force of the musical sounds by nodding in 'agreement', as if interacting with another person. These are all speculative ideas, but it is feasible that the music itself operates on the performer like a partner in a conversation, except that in the case of a musical performance the performer both generates and reacts to the musical narrative.

## Music pedagogy

It is important to note that the idea of movement and metaphor being crucial to musical meaning is featured in several schools of music pedagogy. Dalcrozian Eurythmics, for example, uses movement coordination such as groups of children bouncing balls in unison as a means of illustrating to them how musical material needs to be coordinationed and musical phrases shaped (see Gell 1997, 2005). Guile (in preparation) has undertaken a Laban technique for dance training and has employed his concepts, such as Effort, into the music learning context. Undertaking extensive training programmes with children, she has found that demonstrating actions with effort characteristics such as 'dabbing' and 'punching', children are able to imbue their musical performance with expressive/emotional characters not otherwise present. At a more advanced level, Pierce (1994) has drawn on movement concepts to assist tertiary level and professional performers to optimize the expressive features of playing.

Thus far, the gestures and movements we have considered all relate to musical material itself, and deal only with the soloist. This ignores the significant influence that co-actors and audience may have on performance movement.

## Co-performance

An attempt to understand which postures and gestures are used by co-performers, Davidson and Coulam (2006) studied singers working with an accompanist. It was shown that a co-performing accompanist preferred collaborating with the singers who coordinated musical detail through non-verbal means using the sorts of metaphoric devices mentioned above. However, in addition to these, illustrative gestures such as clicking fingers to set a tempo, and using a series of regulatory head nods and arm gestures to control when and how the players should develop an idea or stop were also used. Such non-verbal means were preferred over talking (Davidson 2005). Indeed, when interviewed after rehearsal and performance activity, the co-performers spoke about 'reading the signs' between them as being crucial to the successful ensemble. Thus, in order to optimize the non-verbal material, gaze and eye contact were very important. Intriguingly, the performers were conscious that they needed to look at one another frequently to make sure that their bodies could coordinate to make the precise musically timed expressive effects.

Durrant (1994) has researched conductors, especially choral conductors, and noted that they use culturally embedded emblems quite extensively, but these are often very elaborate codes (e.g. we need to 'know that a "V" sign with the index and middle finger means "Victory"), and part of the rehearsal period is concerned with choristers learning and then speeding up on decoding what these emblems signify. As the emblems are usually complex signs to short-circuit a lot of spoken instruction, they typically combine technical and expressive instruction. For example, using the hand at the side of the head in a rotational gesture to indicate the notion of creating 'space' for the singing voice to sound and accompanying this with a sharp downward pointed gesture to indicate the vocal attack on the sound. Although there is very limited formal research, Davidson (1997) has hypothesized that the most successful conductors on the international professional circuit

are those who not only focus on non-verbal interaction, but who are also efficient in their use of illustrators and regulators and keep emblems to a minimum, given the time involved in learning their meanings.

Clearly, co-performers need to be able to decipher intention within and between musical parts in order to achieve a fluent performance. Therefore it is necessary to understand the nature of their collaboration and the effect it has on creating a musical performance.

## Co-performer collaboration: from rehearsal to performance

Van Knippenberg and Hogg (2003) discuss that any form of group interaction necessarily gives rise to individual as well as group outcomes. It is these outcomes that motivate people to join and continue in groups. Davidson and Good (2002) discovered that collaborative activities between string quartet members depended in almost equal measure upon: first, the dynamics of musical content and its coordination (achieving entrances and exits etc.) in order to get the performance to cohere; and, secondly, those dynamics of a more personal nature (for example, the first violinist making a head nodding gesture of approval to the second violinist after executing a particularly well-accomplished passage) in order to make the social group cohere. In their data, Davidson and Good saw a power struggle emerging in the rehearsal activity observed. This was based principally on sexual politics—one flirtatious male amongst three more or less admiring females—and it was noted that the flirtation was apparent in the use of non-verbal gestures. The male second violinist teased the first violinist about her lack of clarity, imitating in a bold and exaggerated bowing arm the sudden jerkiness of her regulatory gesture to say 'start the music, now!' The example served also to show how multifunctional the role of non-verbal communication can be in the rehearsal. Additionally, the example shows that individual agendas—even if at odds with the musical goals—can still be expressed, and a coherent musical outcome achieved. Indeed, we are all aware of some famous ensembles who do not get along socially, yet are able to rehearse and perform in an excellently coordinated manner by sticking solely to the musical task.

Recalling the string quartet study and the video data obtained, it was intriguing to observe that when there was a lull in the playing activities of the rehearsal, the cellist moved away from the group sitting arrangement to place herself at the edge of the room, away from a flirtatious social interaction between the first and second violinists. The viola player sat in her traditional position but looked away, shifting her body position as if to exclude herself from the interaction. This type of behaviour has been examined within non-verbal communication research and is known as 'proxemics'—the use of personal space.

The field of proxemics was first studied by Hall (1963), and it was soon discovered that humans have a desire to maintain certain spatial 'zones' or safe distances from one another dependent on circumstance, cultural norm and individual variability. At an everyday level, we are all sensitive to someone who stands too close for comfort or is too far away or avoids eye contact. Music-making is no exception. Finding an appropriate place to stand or sit and assessing performance distance seems crucial.

Williamon and Davidson (2002) found that in the collaboration of a piano duo, where the players sit within one another's personal space (the brief overlapping arms are often required), practice brought an increase in the quantity of non-verbal interaction from first meeting through to a performance. The most used movement was a slight head and eye indication for regulation. Also, and again intriguingly, during the course of the rehearsals, the movement styles of the two individuals altered. The player who moved the most moderated his movements by reducing the quantity of his bodily movements somewhat, while the more conservative player produced more movement. It was as if they began to move as one, rather than two individuals, once they were within the same zone. Evidently, the social dimension was found in that they sensed that they were performing this 'together', as one 'unit'.

Davidson and Good (2002) also noticed this movement unifying effect in the string quartet especially when it played Britten's 'Rhapsody'; the players found it much easier to achieve the

canonical musical entries if they were all ebbing and flowing in exactly the same manner, leaning across into one another's movement pathway.

The current author has researched many types of musical interaction and collaboration, and for her one of the most revelatory experiences has been to work with blind musicians. Over a number of years she has run workshops on rehearsal, performance and stage presentation classes in association with the Royal National Institute for Blind People UK (RNIB). In this work, she has become highly sensitive to the importance of proxemics. Clearly, for practical reasons, blind people may need to pace out distances between stage activities, or be advised how best to position themselves in relation to co-players and audience for the advantage of all. These practicalities aside, blind people also demonstrate how important 'safe' distance zones for interactions are for all. As opera singing student Victoria Oruwari comments in a personal communication, blind people can feel 'invaded' as sighted people stand too close for comfort ('they often stand close in attempt to be better understood'). Sighted people also employ a range of non-verbal touching behaviours in their everyday social encounters as they chat (tapping on the shoulder, nudging with the elbow) which are behaviours blind people do not participate in ('Why do they do all that poking and prodding?'). This information has important implications for rehearsal and performance. Achieving a good distance between sighted and blind players for extra-musical feedback, for instance picking up cues from breath and the sounds of movements (shift of foot position, turn of torso *etc.*) all provide information for the blind musician, without engaging in verbal explanation, or worse still uninvited touch. However, the same could be said of any social interaction. These are codes that we need to respect. Keeping a respectful or appropriate distance is essential, just as it is to be sufficiently close to share in musical and necessary social exchange.

## Actions including and beyond words

As explored above, movements in musical rehearsal and performance clearly aid in the achievement of musical coherence and can enhance social unity. Studying chamber musicians, Goodman (2002) and Murninghan and Conlon (1991) found that non-verbal communication during rehearsal often solved problems where talk had failed. By simply playing and sorting out coordination 'on the hoof', drawing on non-verbal means of communication, there was often no need to discuss a particular point at all. Or alternatively, when a verbal disagreement about musical interpretation emerged, playing the piece through in several different manners would often lead to a decision being made, again without verbal discussion. These findings are supported in the much larger context of a Western symphony orchestra. It is well-reported that in rehearsals, where there is a conductor, this person leads (Atik 1994). The musical ensemble research indicates that too much talking either from the conductor or in the form of questions from an individual player disrupts the flow of the music in rehearsals (see Weeks 1996; Yarborough 1975; Durrant 1994; Price and Byo 2002). That is, if too much questioning is permitted from one individual, it is likely to lead to interference as other members may become irritated or distracted.

Focusing on the key requisites of ensemble performance, Davidson and Goodman (2004) highlighted that specific moment by moment information needs to be processed and responded to in an ongoing manner, in order to accommodate a co-player's sudden change of tempo or coping with a memory slip. Such a capacity to deal with these matters clearly depends on former exposure to such situations and so also relies on previous knowledge of similar situations. Most powerfully, it depends on an opening of ears and eyes to hear and see cues.

What of these cues? Are they of the form and quantity normally encountered in social conversation? Goodman's work (see Goodman 2002; Davidson and Goodman 2004) has drawn on Bales's (1950, 1999) small group behaviour framework. It shows that in terms of both speech and gesture used in rehearsals of two cello and piano duos, the high amount of agreement indicated in the performers' discourses exceeded a suggested 'normal' social upper limit, suggesting that the performers gave mainly positive socio-emotional reactions (perhaps stronger

than witnessed in other small group scenarios). Indeed, in addition to the illustrative, regulatory and emblematic gestures described above, it has been both observed and discussed (Davidson and Coulam 2006; Davidson 2007) that performers—whether singers or instrumentalists—who make non-verbal signals which revealed unconscious processes of self-stimulation, e.g. rubbing the ear lobe in a gentle manner, or flicking the finger tips in a particularly soft but repetitive manner. These—completely unconsciously produced movements—have been regarded as important factors in determining performance quality, for those performers who were rated highest by their collaborating accompanist used more of these self-stimulating adaptors, the movements seeming to offer knowledge of the performer's 'intimate' socio-emotional states.

## Are there elite examples of exceptional performance communication?

In building relations, rehearsing and developing an ensemble, it seems that there are elite profiles. Murninghan and Conlon (1991) studied more or less successful string quartets. They noted that the most successful took account specifically of the opinions and behaviours of the second violinist, who could easily feel like the lesser partner to the first violinist. This finding fits well with group behaviour theory that individuals need to adapt to meet the group demand (see Davidson 1997).

The ensemble achieved between siblings might offer insights or examples of special collaborative and communicative advantage, all members having a deep knowledge of one another in terms of abilities, personalities and movement style (adaptive and personal as well as external and more projected behaviours). Examining the Irish pop band, The Corrs, it was surprising to find that the three sisters who perform in the ensemble were quite idiosyncratic, being influenced by the types of musical instruments, the musical tasks they occupied within the band and their personalities (see Kurosawa and Davidson 2005). For example, Andrea was much more extrovert, and employed many more illustrative and display behaviours. Sharon focused on communicating the narrative of the song, her manner was smaller and less demonstrative than Andrea. Caroline was the most reticent and rather than communicative with her sisters and the audience, most of her movements had a regulatory function. Even when performing the same song, Caroline and Sharon use their bodies differently, with Caroline using more emblematic gestures, gaze and general facial expression, and Sharon using regulators and illustrators.

To try to gain insight into these differences, biographical information was collected. It emerged that Caroline is reported to be the most shy, as well as being the youngest sibling. Also, she plays drums, instruments which fully engage the limbs and which operate within the tight constraint of 'keeping the musical pulse', so she may have developed a general performance style which relies entirely on coordinating and regulating rather than exploring expressive content. By contrast, the violin, played by Sharon, has far more movement possibilities, the arms, hands and fingers arguably being used through their mechanical action through the bow and on the strings to create the musical expression. Finally, Andrea is the most clearly defined as a lead vocalist, and perhaps has more freedom for bodily expression as well as being more used to having attention focused on her. So, according to the case study of The Corrs, it would seem that the observed collaboration and movements of the performers reveal aspects of personality and reflect the type of instrument as well as the hierarchy of prominence of position within their ensemble.

Performers certainly seem to manage multiple social tasks: they have to construct the correct ambience in the rehearsal situation, they have to develop their musical material for the audience: whilst they have to interact with their co-performers to coordinate the performance task, they also have to communicate with and sometimes interact with the audience. Frith (1996) observes that a performance involves thoughts and actions to create the performance which relate to: the occasion itself (e.g., 'star-like' postures and gestures of being the 'stage persona'); the music to be performed (the musical material's narrative content; e.g., 'living' out

the qualities of the character depicted in a song); and the individual performer's true states (current mood, sense of self, etc.). This suggests that there is the potential for a tension between the information contained in the music (its own narrative content), the performer's real state (the individual performers on stage, presenting their own personalities), and behaviours they may engage in to present the music and themselves to their co-performers and audience. It would seem that that performers would need to be efficient at managing these aspects in combination.

Successive studies of popular, jazz, classical and traditional Cantonese opera singers (Davidson 2001, 2002; Davidson and Lai 2007 revealed that above all other musical performers, singers clearly adopt the different roles Frith outlines. The roles were revealed, as indicated with The Corrs, in the very different types of behaviours and communication, much being revealed in bodily movement. For instance, Annie Lennox literally 'showed off' to her audience using provocative sexual body postures. The countertenor Michael Chance created a sense of poise and 'performance occasion' by making a large and slow sweeping forward arm movements over the course of the whole introduction to a Bach solo cantata. The Chinese opera singer Amy Wu held an emblematic stance drawing a focused attention to her raised hand which then flickered, illustrating the 'flame of the red candle' (the title of the song) and presenting the story in a bold and dramatic manner to her audience, yet showed her personal shyness and vulnerability in timid small bows and sideways glances after she had finished performing.

With few existing music research models in this area, it is necessary to turn to another field of psychological research, sports psychology, to pursue the implications of such findings in more detail. There has been some investment of research in sports to understand the personal attributes of elite-level performing athletes (see Vealey *et al.* 1998). Subtle theoretical work by Beauchamp *et al.* (2003) has shown that these performers display an acute understanding of 'self', coupled with the capacity to establish very quickly the patterns of behavioural preference of their collaborators—being highly empathic, they are able to embrace the collaborator's (team member) or audience's perspective in performance. So, working within specific traditions, the 'good' music performer is likely to be someone who can indeed play to the demands of both co-performer and audience. The Jungian concepts of being able to think, feel, sense and intuit are presented by Beauchamp and colleagues as being of central importance to the elite performer, with extraverted attitudes of thought (taking action and seeking order) being associated with feelings of being outgoing, and with sensing preferences for the here and now combining with intuitions including creative work, but also having those reflective, sensitive characteristics apparent. Ronglan (2007), who undertook a season-long in-depth case study of an elite women's handball team to look at the process within the team to build for cooperation and collaboration, showed the same types of characteristics to be crucial to success, with personal enthusiasm, willpower, and persistence being stressed as highly desirable attitudes—but also empathy and patience towards others. In sports, such research has been used to assist in developing suitable training programmes to build performer strategy for collaborative team work and solo excellence. There is, of course, such scope in music though frameworks have not been established to undertake such investigations or training programmes. We can look to extremely successful music performer examples, however, to see whether their case fits the theoretical propositions.

Consider Robbie Williams, arguably one of the most loved performers in the world today. His concerts are hugely popular, and investigating the content of his performer–co-performer and audience exchanges in some detail, it is possible to see that a principal concern for him is to create a coherence between himself and the musicians and to engage the audience actively in the performance. An analysis of Williams' singing the song 'She's the one' at the Knebworth Music Festival in 2004 by Davidson (2006) shows that he gestures warmly to his band, making empathic movements in time with their accompaniment and spends time engaging with them on stage. He also runs to greet his audience on a special performance 'catwalk' which stretches out into the crowd. Crucially, the audience 'know Robbie' through extensive exposure on video. The audience 'knows' how he moves and gestures when he adopts his stage persona.

The crowd even begins to make Robbie's own emblematic gestures ahead of him, anticipating the unfurling narrative of the song he is performing. Despite the thousands in the crowd, they whole audience is found to sway, swirl, dance and sing in perfect unison with their idol. Thus, the pop audience has a special kind of attunement to Robbie's stage persona in the performance. Of course much has been learned in viewing preceding music videos, but it is clear that all of this has been translated into the audience's own expectation and knowledge of Williams' and his stage movements. Yet there is more to Robbie than this the 'cheeky chappy' persona. The crowd also seems to have an empathy with Robbie himself—his authenticity and vulnerability. How is this achieved? Robbie connects with individual members within the whole audience, and without showing self-consciousness, he carries out small personal adjustments behaviours such as fiddling with his earpiece. He also speaks to the audience as it if were a single person. Robbie seems to be the consummate musical communicator.

Of course, each individual performance is a unique confluence of new interactive elements, and familiarity with such experiences prepares co-performers for likely scenarios and outcomes: but does the presence of an audience change the quality of musician behaviour?

## Movements for audience effect

We know that physical gestures are not prerequisites to understanding either speech or music, such as in a phone conversation or music played on a CD recording. However, since we make fewer gestures when there is no face-to-face contact in both speech and music, it would seem that whilst some gestures are specifically oriented towards generating the music, a proportion would seem to be used for audience communication effects. Differences between musical and audience concerns are demonstrated in the evidence found in the biography of case the Canadian pianist, Glenn Gould.

Early in his career, Gould performed in all the top classical concert venues including Carnegie Hall. Later in his career, he decided to give up public performance, focusing solely on recording studio performances. Delalande (1990) made a study of Gould's performances taken from rare film footage of both periods. What is striking is that in the recording of the studio playing, Gould's movements are highly repetitive and often quite disturbing to watch. By contrast, those of the public recitals have smooth, flowing movements. Thus, there is some evidence that in the studio case, Gould's concerns were entirely focused on the music, whereas in the public context, he was taking into account the audience's presence.

Not every social tradition permits the same social concerns to develop in collaborative performance, however. In Japanese classical music, for example, performers are not supposed to make any extraneous gestures or facial expressions to detract from the emotional content of the music, thus giving relatively few social cues to the audience or indeed their co-performers about their outer expression (Malm 2000).

Studying Indian classical music, Clayton *et al.* (2004) have shown that performers not only use movement gestures to highlight features of the music in a metaphorical way, but they use gestures that are often almost immediately adopted by both their co-performers and audiences. Fascinatingly, a performer may initiate one gesture, such as a head shaking movement, which is then mirrored by another player in the ensemble, and then passed on to members of the audience. The result is that all participants in the event mutually 'tune-in' to the shared moment. It is as if by collaborating in the production of the gestures they are being active participants in the making of the music itself.

Generally, in Western classical music performance, audiences are far more passive, sitting quietly and acknowledging the performance activity with hand clapping, but never as part of the performance itself. In jazz, audiences do take on more of a performance-focused role, toe-tapping and dancing in response to the music. Similar evidence occurs in other culture—Balinese music drama (e.g., McIntosh 2006a, b). However, as the earlier example of Robbie Williams demonstrated, pop performance offers one of the greatest forums where performer and audience can share in a directed but often mutual 'tuning-in' social interaction. We see in the case of Williams that he offers both the self-stimulating

adaptors—the movements seeming to offer knowledge of the performer's 'intimate' states—alongside the projected 'public' states of 'display' (showing off) and emblematic gesture.

## Conclusions

It has been demonstrated throughout this chapter that playing a piece of music requires a high-level interaction of developed cognitive and action processes; thus a range of internal representations in memory are necessary. The ability to use mental representations is determined by knowledge and experience. The capacity to generate representations seems to be limitless, with opportunities for learning permitting refinement, adaptation and addition. It is possible to summarize the key points explored as follows:

For performers:

- The body and bodily movement is inherently bound up with the mental representations required to develop musical performance skills
- Musical performance skills involve the biomechanical aspects of playing the music fluently and also expressively
- Technical achievement and expression combine to produce movements with overall fluency, and also local specifically identifiable gestures which are not necessary for note execution, but are essential for the communicative aspects of performance
- These gestures have functions, a first level being to give the music an immediate and communicative purpose for the performer as they create the music product (generating and responding to the musical sounds in a interactional manner between themselves and the music)
- A second level of function is to provide co-performer cues for regulation of musical content and expression of immediate and perhaps idiosyncratic expressive content
- The movements used to generate these stimuli draw from individual movement repertoires, some of which reveal intimate personal states, others of which are for much more direct communicative ends, e.g., end the phrase *now!*
- The original of these movements is likely to be similar to that of non-verbal communication in speech
- The social and cultural origin of the performance tradition out of which a performance is being produced influences the extent and types of communicative movement produced.

For audiences:

- The movements permit accurate perception of musical intention
- The movements offer possibilities for sharing and participation.

Needless to say, understanding bodily movement is a key area for further enquiry for both performers and audiences. Much research still needs to be done to detail systematically the types and functions of all movement in musical performance. Yet however limited our present knowledge, it is obvious that psychological investigations have given musicians and their audiences insights into how crucial bodily movement is to the execution and interpretation of musical performance, and this research area continues to grow.

## References

Altenmüller E and Gruhn W (2002) Brain mechanisms. In R Parncutt and GE McPherson, eds, *The science and psychology of music performance: creative strategies for teaching and learning*, 63–82. Oxford University Press, Oxford.

Atik Y (1994). The conductor and the orchestra: interactive aspects of the leadership process. *Leadership and Organisation Development Journal*, **13**, 22–28.

Bailes RF (1950) A set of categories for the analysis of small group interaction. *American Sociological Review*, **15**, 257–263.

Bales RF (1999). *Social interaction systems: theory and measurement*. Transaction Publishers, London.

Baily J (1985) Music structure and human movement. In P Howell, I Cross and R West, eds, *Musical Structure and Cognition*, 237–258. Academic Press, London.

Beauchamp M, Lee KE, Haxby JV and Martin A (2003) fMRI response to video and point-light displays of moving humans and manipulable objects. *Journal of Cognitive Neuroscience*, **15**(7), 991–1001.

Bernstein N (1967) *Coordination and regulation of movement*. Pergamon, London.

Blacking J (1977) *The anthropology of the body*, ASA Monograph 15, edited volume. Academic Press, London.

Clayton M, Sager R and Will U (2004) In time with the music: the concept of entrainment and its significance for ethnomusicology. *ESEM Counterpoint*, **1**, 1–45, 70–74.

Dahl S (2004). Playing the accent—comparing striking velocity and timing in an ostinato rhythm performed by four drummers. *Acta Acustica*, **90**(4), 762–776.

Dalla-Bella S and Palmer C (2004). Tempo and dynamics in piano performance: the role of movement and amplitude. In SD Liscomb, R Ro Gjerdingen and P Webster, eds, *Proceedings of the international conference on music perception and cognition*, 256–257. Causal Productions, Adelaide.

Davidson JW (1993). Visual perception of performance manner in the movements of solo musicians.*Psychology of Music*, **21**, 103–113.

Davidson JW (1994). What type of information is conveyed by the body movements of solo musician performers? *Journal of Human Movement Studies*, **6**, 279–301.

Davidson JW (1997). The social psychology of performance. In DJ Hargreaves, AC North, eds, *The social psychology of music*, 209–226. Oxford University Press, Oxford.

Davidson JW (2001) The role of the body in the production and perception of solo vocal performance: a case study of Annie Lennox. *Musicae Scientiae*, **V**(2), 235–256.

Davidson JW (2002) The performer's identity. In R MacDonald, D Miell and DJ Hargreaves, eds, *Musical identities*, 97–116. Oxford University Press, Oxford.

Davidson JW (2005) Bodily communication in musical performance. In D Miell, DJ Hargreaves and R Macdonald, eds, *Musical communication*, 215–238. Oxford University Press, New York.

Davidson JW (2006) 'She's the One': multiple functions of body movement in a stage performance by Robbie Williams. In A Gritten and E King, eds, *Music and gesture*, 208–226. Ashgate, Aldershot.

Davidson JW (2007). Qualitative insights into the use of expressive body movement in piano performance.*Psychology of Music*, **35**(3), 381–401.

Davidson JW and Correia JS (2002) Body movement in performance. In R Parncutt and GE McPherson, eds, *The science and psychology of music performance: creative strategies for teaching and learning* 237–250. Oxford University Press, Oxford

Davidson JW and Coulam A (2006) Exploring jazz and classical solo singing performance behaviours: a preliminary step towards understanding performer creativity. In G Wiggins and I Deliège, eds, *Musical creativity: current research in theory and practice*, 181–199. Oxford University Press, New York

Davidson JW and Dawson JC (1995). The development of expression in body movement during learning in piano performance. *Conference Proceedings of Music Perception and Cognition Conference*, p. 31. University of California, Berkeley, June.

Davidson JW and Good JMM. (2002) Social and musical coordination between members of a string quartet: an exploratory study. *Psychology of Music*, **30**, 86–201.

Davidson JW and Goodman E (2004) E. Strategies for ensemble performance. In A Williamon, ed., *Musical excellence: strategies and techniques to enhance performance*, 105–122. Oxford University Press, Oxford.

Davidson JW and Lai V (2007) Shanghai street: art in and of a city. *International Journal of Arts in Society*, **1**(7), 109–119.

Delalande F (1990) Human movement and the interpretation of music. Paper presented at the Second International Colloquium on the Psychology of Music, Ravello, Italy.

Durrant C (1994). Towards an effective communication: a case for structured teaching of conducting. *British Journal of Music Education*, **11**, 56–76.

Frith S (1996) *Performance rites*. Oxford University Press, Oxford.

Gell H (1997) *Lessons in music through movement*. CIRCME, University of Western Australia.

Gell H (2005) *Dalcroze Eurythmics: music through movement: a hundred lessons and thousands of ideas for early childhood*. Callaway Centre, University of Western Australia.

Goodman E (2002) Ensemble performance. In J Rink, ed., *Musical performance: a guide to understanding*, 153–167. Cambridge University Press, Cambridge.

Guile LM (in preparation) Laban's concept of effort applied to musical expression. Work for PhD dissertation, University of Sheffield, UK.

Hall ET (1963). A system for the notation of proxemic behaviour. *American Anthropologist, Selected papers in methods and techniques*, **63**(5), 1003–1026.

Jäncke L (2006) From cognition to action. In E Altenmüller, M Wiesendanger and J Kesselring, eds, *Music, motor control and the brain*, 25–37. Oxford University Press, Oxford.

Kendon A (1980) Gesticulation and speech: two aspects of the process. In MR Key, eds, *The relation between the verbal and nonverbal communication*, 207–228. Mouton, The Hague.

Kurosawa K and Davidson JW (2005) Non-verbal interaction in popular performance: The Corrs. *Musicae Scientiae*, **9**(1), 111–136.

Lashley KS (1951) The problem of serial order in behaviour. In LA Jeffress, ed., *Cerebral mechanisms in behaviour*, 118–142, Wiley, New York.

Lehmann AC and Davidson JW (2002) Taking an acquired skills perspective on music performance. In R Colwell, C Richardson, eds, *Second handbook on music teaching and learning*, 542–560. Oxford University Press, Oxford.

Malm WP (2000) *Traditional Japanese music and musical instruments*. Kodansha International, Tokyo.

McClary S (1991) *Feminine endings: music, gender and sexuality* University of Minnesota Press, Minneapolis, MN.

McIntosh J (2006a). Moving through tradition: children's practice and performance of dance, music and song in South-Central Bali, Indonesia. Unpublished PhD dissertation, Queen's University Belfast.

McIntosh J (2006b). How singing, dancing and playing shape the ethnographer: research with children in a

Balinese dance studio. *Anthropology Matters*, **8**(2), 1–17. Available at: http://www.anthropologymatters.com/journal/2006 2/mcintosh_2006_how.pdf.

Murninghan JK and Conlon DE (1991). The dynamics of intense work groups: a study of British string quartets. *Administrative Science Quarterly*, **36**, 165–186.

Palmer C (2006) The nature of memory for music performance skills. In E Altenmüller, M Wiesendanger and J Kesselring, eds, *Music, motor control and the brain*, 109–123. Oxford University Press, New York.

Pantev C, Engelien A, Candia V and Elebert T (2001). Representational cortex of musicians: plastic alterations in response to music. *Annals of the New York Academy of Sciences*, **930**, 300–314.

Pierce A (1994) Developing Schenkerian hearing and performing. *Integral*, **8**, 51–123.

Price HE and Byo JL (2002). Rehearsing and conducting. In R Parncutt and GE McPherson, eds, *The science and psychology of musical performance: ceative strategies for teaching and learning*, 335–351. Oxford University Press, Oxford.

Ronglan LT (2007). Building and communicating collective efficacy: a season-long in-depth study of an elite sport team. *The Sport Psychologist*, **21**, 8–93.

Shaffer LH (1984) Timing in solo and duet piano performance. *Quarterly Journal of Experimental Psychology*, **36**, 577–595.

Sloboda JA, Davidson JW, Howe MJ A and Moore DG (1994). The role of practice in the development of expert musical performance. *British Journal of Psychology*, **87**, 287–309.

Van Knippenberg D and Hogg MA (2003). A social identity model of leadership effectiveness in orgnisations. *Research in Organisational Behaviour*, **25**, 243–296.

Vealey RS, Hayashi SW, Garner-Holman M and Giacobbi P (1998). Sources of sport-confidence: conceptualisation and instrument development. *Journal of Sport and Exercise Psychology*, **20**, 54–80.

Weeks P (1996). A rehearsal of a Beethoven passage: an analysis of correction talk. *Research on Language and Social Interaction*, **29**, 247–290.

Wiesendanger M, Baader A and Kazennikov O (2006) Fingering and bowing in violinists: a motor control approach. In E Altenmüller, M Wiesendanger, J Kesselring, eds, *Music, motor control and the brain*, 109–123. Oxford University Press, New York.

Williamon RA and Davidson JW (2002) Exploring co-performer communication. *Musicae Scientiae*, **VI** (1), 1–17.

Wing A (2002) Voluntary timing and brain function: an information processing approach. *Brain and Cognition*, **48**, 7–30.

Yarborough C (1975). Effect of magnitude of conductor behaviour on students in mixed choruses. *Journal of Research in Music Education*, **23**, 134–146.

CHAPTER 35

# Emotion in music performance

Patrik N. Juslin

THERE are several features that we have come to expect from an expert performance: technical mastery, confidence, originality, flexibility, and a true understanding of the musical style. Yet the feature that both performers and listeners appear to regard as the most important is that the performer is expressive. The most loved artists are commonly the ones that are able to express and evoke emotions in listeners.

It should be noted at the outset that there are several ways in which emotion might enter into the performance of music: emotions might influence a performer's motivation, choice of repertoire, interpretation, and concentration, as well as the precise nature of the performance. However, there is little systematic knowledge about most of these processes. Previous studies have mainly concerned how performers express emotions, and this chapter will thus focus on this question, though with brief detours to related topics.

First, I provide working definitions of key concepts (e.g., expression, communication), and consider how performers conceive of these issues. Then I review up-to-date evidence on how performers express emotions. Finally, I propose directions for future research.

## Conceptual foundations

Music performance, virtually by definition, requires a performer and a piece of music. The term *interpretation* usually refers to the individualistic shaping of the piece according to the musical ideas of the performer (Palmer 1997). This may or may not involve an intention to 'express' something (e.g., an idea, an emotion) beyond the musical structure, although such expression is usually achieved through the way in which the structure is articulated.

*Expression* will be used in this chapter to refer to a particular set of perceptual qualities that reflect psychophysical relationships between 'objective' properties of the music (patterns of information) and 'subjective' impressions of the listener (e.g., perceiving certain emotions). Sometimes the perception of an expressive performance evokes an emotion in the listener, but this is not required for the listener to hear the music as expressive. We may be able to perceive that a performer is playing expressively, yet we remain untouched by the performance.

Expression has often been treated as a homogeneous category of which there is more or less (Marchand 1975). However, a careful review of the literature suggests that performance expression is better conceived of as a multidimensional phenomenon that can be decomposed into various subcomponents. Drawing on previous research, I have proposed that performance expression derives from five primary sources, referred to as the GERMS model (Juslin 2003):

- *Generative rules (G)* that mark the structure in a musical manner (Clarke 1988). By means of variations in such parameters as timing, dynamics, and articulation, a performer is able to convey group boundaries, metrical accent, and harmonic structure.
- *Emotional expression (E)* that serves to communicate emotions to listeners (Juslin 1997a).

By manipulating overall features of the performance such as mean tempo, a performer is able to play the same structure with different emotional characters.

- *Random fluctuations (R)* that reflect human limitations in motor precision (Gilden 2001). It has been revealed in several studies that even expert performers who try to play perfectly even time intervals show small, involuntary fluctuations in the timing of their performance.
- *Motion principles (M)* that hold that tempo changes should follow natural patterns of human movement, or biological motion, in order to obtain a pleasing shape (Shove and Repp 1995).
- *Stylistic unexpectedness (S)* that reflects a performer's deliberate attempts to 'deviate' from stylistic expectations concerning performance conventions to add tension and unpredictability to the performance (Meyer 1956, p. 206).

In reality, all components occur together in complex interactions, though for specific purposes (e.g., research, teaching), it can be useful to consider them separately (Juslin 2003).

In this chapter, I will focus on emotional expression (for a working definition of emotion, see Chapter 12 this volume), whilst acknowledging that the other aspects are also important (see Chapters 30–34 this volume). In one sense the term 'emotional expression' is misleading: it is only occasionally that performers are truly expressing their own emotions during the performance—perhaps because optimal performance requires a certain psychological state (e.g., relaxed concentration) that is incompatible with experiencing certain emotions. More typically, what the performer presents in a music performance is not the emotion itself, but rather its expressive form—derived from other forms of non-verbal communication. Still, the term 'emotional expression' is now widely established, and has been used in both cases where the expression is 'spontaneous' (genuinely felt) and where it is 'symbolic' (portrayed). I will therefore retain the term in this chapter.

The notion of expression does not require that there is a correspondence between what a listener perceives in the performance and what the performer intends to express. In contrast, the concept of *communication* requires that there is both a performer's intention to express a specific emotion and recognition of this emotion by a listener. The listener may, additionally, come to experience or feel the emotion in question, but this is not required for it to qualify as a case of communication. For a discussion of emotional reactions to music, see Chapter 12.

## Performers' views

Issues concerning expression, communication, and emotions tend to invite controversy. However, while philosophers, musicologists, psychologists, and educators commonly express strong views on these issues, it is surprisingly rare that the performers or listeners themselves are consulted. Thus, before reviewing findings from basic research on these issues, it may be appropriate to consider recent findings on performers' views. Lindström *et al.* (2003) carried out a questionnaire study, featuring 135 expert performers from music conservatoires in three countries (England, Italy, Sweden). Open-ended responses to the question 'In your view, what does it mean to play expressively?' were content-analysed and divided into categories. The results revealed that the musicians (both classical and popular) defined 'playing expressively' largely in terms of 'communicating emotions' (44 per cent) and 'playing with feeling' (16 per cent). The two response categories were not always easy to distinguish, although the first focused more on conveying something to an audience, whereas the second focused more on the performer's own feelings. A third group (34 per cent) provided answers in terms of 'a focus on the music itself' (e.g., 'conveying the structure'). Of particular interest is that 83 per cent of the performers reported that they consciously try to express specific emotions in their music performances 'always' or 'often'.

These results were confirmed and extended by Minassian *et al.* (2003), who conducted a survey study, including 53 expert performers of classical music, in order to explore what factors were associated with an 'optimal' performance. Performances judged as optimal tended to be those where the performer had a clear intention to communicate (usually an emotional 'message'), was emotionally engaged with the music, and believed the message had been received by

the audience. Interview studies have further indicated that performers in pursuit of a professional performing career are more inclined to think of music as a vehicle for communication than performers in pursuit of a non-music career (Burland and Davidson 2004), and that many performers value music as a form of emotional outlet (Gullberg and Brändström 2004).

## Research paradigms

There is an immense number of historical treatises that feature very detailed descriptions or prescriptions concerning the performance practices that may be used to enhance emotional expression (Bach 1778/1985; Buelow 1983; Mattheson 1739/1954). Most of these writings, however, are based on speculation, intuition, folk theory, or personal experience –which may or may not be correct.

Systematic attempts to investigate emotional expression—as manifested in performance features (e.g., timing, timbre) rather than features of particular pieces of music (e.g., melody, mode)—developed only in the mid-1990s. Most investigations used the so-called *standard paradigm* (a term borrowed from studies of vocal expression of affect; e.g., Juslin and Scherer 2005): this means that musicians are asked to play brief melodies to express various emotions chosen by the researcher. The resulting performances are first recorded and then evaluated in listening experiments to check whether listeners can recognize the intended expression. Each performance is further analysed to study what acoustic means each performer used to achieve each emotional expression. The basic assumption is that because the melody remains the same in different emotional expressions, whatever effects that are found in listeners' judgements or acoustic measures should mainly be the result of the performer's expressive intention. Having performers play the same piece with different emotional expressions might appear 'unnatural' from a musical point of view. However, this design is necessary to secure the internal validity of the experiments: if different emotions are expressed by different melodies, it is impossible to know whether the obtained effects on listener judgements or performance measures are due to the melody, the performance, or some interaction between the two.

However, there is no denying that this is a fairly artificial set-up, even though musicians sometimes do try alternative interpretations of the same piece (Lindström *et al.* 2003). Hence, it is paramount to try other approaches to investigating emotional expression in performances, such as analysing existing recordings of performances (Siegwart and Scherer 1995), observing a performer from the preparation of a piece to a concert performance (Clarke *et al.* 2005), or investigating listeners' responses to performance in a real concert using questionnaire (Thompson 2006).

## Can performers communicate emotions accurately to listeners?

It might appear strange to talk about accuracy of communication in a musical context. However, it is reasonable to assume that most performers are concerned about whether their interpretation is perceived by the audience in the manner that they intended it. The performer may, for instance, wish to highlight an emotional character in the piece. The degree to which performer and listeners 'agree' about the expression of the performance can pragmatically be seen as a measure of the accuracy of the communication. Listeners' judgements can be indexed in terms of adjective ratings (Juslin 1997a), forced choice (Juslin 1997b, Experiment 1), free labelling (Juslin 1997c), or continuous response (Sloboda and Lehmann 2001).

Juslin and Laukka (2003) reviewed 41 studies of emotional expression in performance. The studies covered a wide range of musical styles such as classical music, folk music, Indian ragas, jazz, pop, rock, children's songs, and free improvisations. The most common style was classical music (17 studies = 41 per cent). The number of emotions studied ranged from 3 to 9 (M = 4.98), and typically included happiness, sadness, anger, fear, and tenderness. Twelve musical instruments and twelve nationalities were included. Most studies used professional performers and the performances were mostly monophonic to facilitate acoustic measurements.

Juslin and Laukka conducted a meta-analysis of communication accuracy featuring 29 studies, which indicated that performers are able to

communicate 'basic emotions' (happiness, anger, sadness, fear, and tenderness) to listeners with an accuracy approximately as high as in facial and vocal expression of emotions. Overall decoding accuracy was equivalent to a 'raw' proportion correct of $P_c = 0.70$, in a forced-choice task with five response alternatives. Across studies, sadness and anger were the emotions that were communicated with the highest level of accuracy, and this pattern was found for both music performance and vocal expression. The available evidence suggests that the communication proceeds in terms of fairly broad emotion categories, whereas finer distinctions within these categories are hard to communicate reliably—at least, without some additional context. In general, commonly postulated 'basic' emotions are easier to communicate than 'complex' emotions. For explanations, see p. 385–386.

The accuracy depends on a number of variables (e.g., the piece, the response format, the procedure), yet perception of emotions in music is robust in that listeners' judgements are only marginally influenced by musical training, age, and gender of the listener (see Gabrielsson and Juslin 2003). That musical training is not required to express (Yamasaki 2002) or recognize (Juslin 1997a) emotion in music performance suggests that general mechanisms of perception of emotions are involved—which is supported by the finding that abilities to decode emotions in music performance are correlated with measures of emotional intelligence (e.g., Resnicow *et al.* 2004). Developmental studies of perception of emotions in music have not clearly separated performance features (e.g., tempo, timbre) from compositional features (e.g., melody, harmony), but preliminary evidence from several studies indicates that that children as young as 3 or 4 years old are able to decode 'basic emotions' from music with better than chance accuracy (for a review, see Juslin and Laukka 2003, p. 788). The ability appears to improve with age, but declines again in old adults (Laukka and Juslin 2007).

## How are different emotions communicated?

Most of the processes that underlie communication of emotions in music are implicit—which helps to 'mystify' the nature of expression (see Juslin *et al.* 2004). One of the pioneers in research on emotional expression in music more generally (e.g., Chapter 13), Kate Hevner (1935), observed:

> If the great artist could speak to the audience verbally as effectively as he does musically, our efforts would be unnecessary, but seldom he expresses himself except through the medium of his art, and when he does, it is usually not in the terms calculated to be most useful and helpful.
>
> (p. 204)

Hence, psychologists are forced to explore the mechanisms by means of acoustic measurements, commonly carried out using computer software. Table 35.1 presents a summary of the most important acoustic parameters and also explains how they are measured. In the following, I shall refer to these parameters as *cues* (bits of information) that together make up the code used to communicate various emotions. Researchers have recently developed computer algorithms that can help to speed up the otherwise complicated and time-consuming process of manually analysing the cues (Friberg *et al.* 2007).

### Code description

Table 35.2 shows a summary of cues that are correlated with specific emotions in musical expression. For completeness, the table includes cues that are commonly part of the musical structure (for further discussion of structural features, see Chapter 13). However, those cues that can usually be controlled or modulated by the performer are set in italics. While Table 35.2 is limited to a few emotion categories (i.e., those that have been explored most frequently in previous research), one can easily imagine how the categories can be combined or 'mixed' in various ways, and also how the expression can be altered during a performance. Some of the cues shown in Table 35.2 may be common knowledge to most performers, whereas other cues are less obvious. The number of cues available depends on the instrument used. Other things being equal, the communication will be more reliable the more cues that are available. The results suggest that tempo, sound level, timbre and

**Table 35.1** Definition and measurement of primary acoustic cues in emotional expression in music performance

| Acoustic cues | Perceived correlate | Definition and measurement |
|---|---|---|
| **Pitch** | | |
| Fundamental frequency (F0) | Pitch | Acoustically, F0 is defined as the lowest periodic cycle component of the acoustic wave form. One can distinguish between the macro pitch level of particular musical pieces, and the micro intonation of the performance. The former is often given in the unit of the semitone, the latter is given in terms of deviations from the notated macro pitch (e.g., in cents). |
| F0 contour | Intonation contour | Sequence of F0 values. In music, intonation refers to manner in which the performer approaches and/or maintains the prescribed pitch of notes, in terms of deviations from precise pitch. |
| Vibrato | Vibrato | Periodic changes in the pitch (or loudness) of a tone. Depth and rate of vibrato can be measured manually from the F0 trace (or amplitude envelope). |
| **Intensity** | | |
| Intensity | Loudness | Intensity is a measure of the energy in the acoustic signal. Usually measured from the amplitude of the acoustic wave form. The standard unit used to quantify intensity is a logarithmic transformation of the amplitude called the decibel (dB). |
| Attack | Rapidity of tone onsets | Attack refers to the rise-time or rate of rise of the amplitude of individual notes. Usually measured from the acoustic wave form. |
| **Temporal features** | | |
| Tempo | Velocity of music | The mean tempo of a performance is obtained by dividing the total duration of the performance until the onset of its final note by the number of beats, and then calculating the number of beats per minute (bpm). |
| Articulation | Proportion of sound to silence in successive notes | The mean articulation of a performance is typically obtained by measuring two durations for each tone—the duration from the onset of a tone until the onset of the next tone (dii), and the duration from the onset of a tone until its offset (dio). These durations are used to calculate the successive dio:dii ratio (the articulation) of each tone. These values are averaged across the performance and expressed as a percentage. A value around 100% refers to legato articulation; a value around 70% or lower refers to staccato articulation. |
| Timing | Tempo and rhythm variation | Timing variations are usually described as 'deviations' from the nominal values of a musical notation. Overall measures of the amount of deviations in a performance may be obtained by calculating the number of notes whose deviation is less than a given percent of the note value. Another index of timing changes concerns so-called durational contrasts between 'long' and 'short' notes in rhythm patterns. Contrasts may be played with 'sharp' durational contrasts (close to or larger than the nominal ratio) or with 'soft' durational contrasts (a reduced ratio). |

*Continued*

**Table 35.1** Definition and measurement of primary acoustic cues in emotional expression in music performance (*continued*)

| Acoustic cues | Perceived correlate | Definition and measurement |
|---|---|---|
| **Timbre** | | |
| High-frequency energy | Timbre | Refers to the relative proportion of total acoustic energy above versus below a certain cut-off energy in the frequency spectrum of the sound wave. In music performance, timbre is in part a characteristic of the specific instrument. However, different techniques of playing may also influence the timbre of many instruments, such as the guitar. |
| The singer's formant | Timbre | The singer's formant refers to a strong resonance around 2500—3000 Hz, and adds brilliance and carrying power to the voice. It is attributed to a lowered larynx and widened pharynx, which forms an additional resonance cavity. |

**Table 35.2** Summary of cues correlated with specific emotions in musical expression

| Emotion | Cues |
|---|---|
| Happiness | *fast tempo*, *small tempo variability*, major mode, simple and consonant harmony, *medium-high sound level*, *small sound level variability*, high pitch, much pitch variability, wide pitch range, ascending pitch, perfect 4th and 5th intervals, *rising micro intonation*, *raised singer's formant*, *staccato articulation*, *large articulation variability*, smooth and fluent rhythm, *bright timbre*, *fast tone attacks*, *small timing varibility*, *sharp contrasts between 'long' and 'short' notes*, *medium-fast vibrato rate*, *medium vibrato extent*, *micro-structural regularity* |
| Sadness | *slow tempo*, minor mode, dissonance, *low sound level*, *moderate sound level variability*, low pitch, narrow pitch range, descending pitch, *'flat' (or falling) intonation*, small intervals (e.g., minor 2nd), *lowered singer's formant*, *legato articulation*, *small articulation variability*, *dull timbre*, *slow tone attacks*, *large timing variability* (e.g., rubato), *soft contrasts between 'long' and 'short' notes*, *pauses*, *slow vibrato*, *small vibrato extent*, *ritardando*, *micro-structural irregularity* |
| Anger | *fast tempo*, *small tempo variability*, minor mode, atonality, dissonance, *high sound level*, *small loudness variability*, high pitch, moderate pitch variability, ascending pitch, major 7th and augmented 4th intervals, *raised singer's formant*, *staccato articulation*, *moderate articulation variability*, complex rhythm, sudden rhythmic changes (e.g., syncopations), *sharp timbre*, *spectral noise*, *fast tone attacks/decays*, *small timing variability*, *accents on tonally unstable notes*, *sharp contrasts between 'long' and 'short' notes*, *accelerando*, *medium-fast vibrato rate*, *large vibrato extent*, *micro-structural irregularity* |
| Fear | *fast tempo*, *large tempo variability*, minor mode, dissonance, *low sound level*, *large sound level variability*, *rapid changes in sound level*, high pitch, ascending pitch, very wide pitch range, large pitch contrasts, *staccato articulation*, *large articulation variability*, jerky rhythms, *soft timbre*, *very large timing variability*, *pauses*, *soft tone attacks*, *fast vibrato rate*, *small vibrato extent*, *micro-structural irregularity* |
| Tenderness | *slow tempo*, major mode, consonance, *medium-low sound level*, *small sound level variability*, low pitch, fairly narrow pitch range, *lowered singer's formant*, *legato articulation*, *small articulation variability*, *slow tone attacks*, *soft timbre*, *moderate timing variability*, *soft contrasts between long and short notes*, *accents on tonally stable notes*, *medium-fast vibrato*, *small vibrato extent*, *micro-structural regularity* |

pitch are the most powerful cues in terms of their effects on listeners' ratings (see Juslin 1997b; Juslin and Madison 1999; Scherer and Oshinsky 1977). Certain instruments might be more effective in communicating certain emotions than others depending on their specific timbre (Behrens and Green 1993; Bunt and Pavlicevic 2001). The focus on 'basic emotions' is sometimes criticized, though findings from survey studies of both music performers (Lindström *et al.* 2003) and listeners (Juslin and Laukka 2004) suggest that they regard these emotions as the ones that are easiest to express via music, and they thus represent a natural point of departure from which more 'complex' emotions may gradually be investigated. See Juslin *et al.* (2004, Table 13.2) for examples of complex emotions.

## Expressive patterns and ornamentation

Most studies in the field so far have focused on cues that are applied in much the same way, and to much the same extent, throughout a piece. Indeed, experiments suggest that 'overall' levels of cues, such as mean tempo, explain considerably more variance in listeners' emotion ratings than does variability (e.g. Madison 2000). However, there exist patterns of changes in tempo and dynamics that are characteristic of particular emotions. Listening tests using systematic manipulations of synthesized (Juslin 1997b) or real performances (Juslin and Madison 1999) indicate that such patterns may be used by a listener to decode the emotional expression. What is it in these patterns that conveys the relevant information? One clue comes from the study by Lindström (1999), who found that performers accentuated different aspects of the melodic structure of a piece depending on the intended expression; thus, for example, if a certain note in the melodic structure was regarded as especially 'happy', then the performers emphasized this note in happiness expressions, but de-emphasized it in sadness expressions.

Another structure-related aspect of emotional expression in performances that has been little investigated is *ornamentation*, the use of expressive features that are not necessary to the melodic or harmonic line, but only serve to decorate that line. Historical performance treatises feature extensive discussions of ornaments, and they may be especially important in Baroque music (e.g. Bach 1778/1985; Neumann 1978). Standard ornaments include the trill (a rapid alternation between two notes), the mordent (a rapid, single alternation between two notes), the glissando (a continuous slide upwards or downwards between two notes), and the appoggiatura (a grace note which delays the next note of the melody, taking half or more of its written time value). A first attempt to investigate the use of ornaments in the expression of specific emotions was made by Timmers and Ashley (2007). They analysed performances of fragments from Handel's G-minor sonata for recorder and basso continuo (HWV 360) by a flutist and a violinist and found that the use of ornaments varied according to the intended expression. Figure 35.1 shows examples of the ornaments use to express the various emotions.

## Synthesis of emotional expression

Performance analyses have shown that performers can use a number of cues to express particular emotions. However, these analyses do not prove that listeners actually *use* the same cues in their judgements. To take just one example, the above study of ornaments showed that a performer used ornaments differently to express various emotions, but it did not prove that a listener could actually utilize any of these ornaments to decode the intended expression. Thus, to test the validity of hypotheses about expressive cues derived from performance studies, it is necessary to conduct listening experiments with synthesized performances. Note that there are two different approaches to synthesis of emotional expression in performance that can answer different questions. Based on a distinction introduced by Brunswik (1956), I have called them *representative design* and *systematic design* (Juslin 1997b).

Representative design means that one tries to recreate the emotional expressions of real music performances by programming a computer to perform in accordance with the emotion-specific patterns of acoustic cues obtained in previous studies (Table 35.2). Ideally, the patterns used to synthesize the emotional expression should be 'representative' of real performances in the

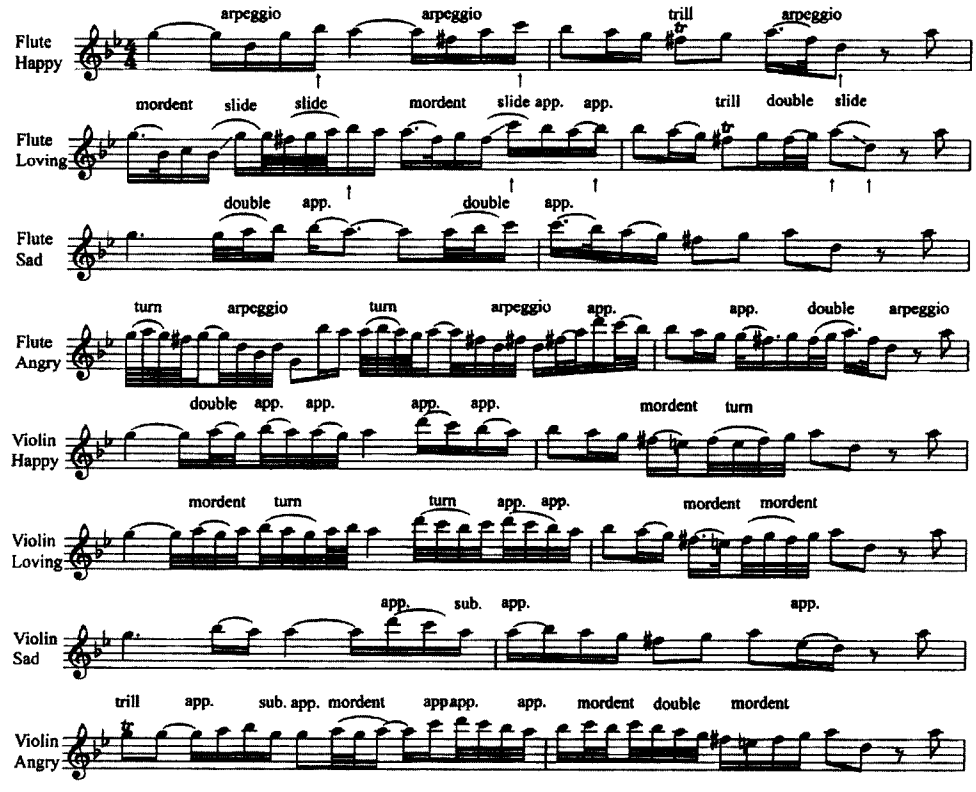



**Fig. 35.1** Examples of ornaments used by a flutist and a violinst in music performances intended to express 'basic emotions' (from Timmers and Ashley 2007). Reprinted by permission of the University of California Press.

sense that they display similar statistical characteristics (e.g., cue intercorrelations, means, and standard deviations). Preliminary studies have shown that one can program a computer to communicate emotions with the same level of accuracy as a human performer (Juslin 1997b, Experiment 1; see also Bresin and Friberg 2000). However, such synthesized performances do not sound as musically satisfying as human performances do, perhaps because they lack other expressive features, such as marking of structure, biological motion, and random fluctuations, that are also important (Juslin *et al.* 2002). A representative design is useful when comparing the decoding accuracy of real and synthesized performances, although it cannot prove that the implementation of an individual cue is correct, or even that listeners are using the cue.

To be able to unambiguously attribute variance in a listener's judgements of emotions to individual acoustic cues, one needs to use a systematic design, or more specifically a factorial design. Such a design removes all the intercorrelations among cues, and renders it possible to ascertain the effectiveness of individual cues. Notably, only some of the performance cues in Table 35.2 have been tested in this way (Juslin 1997b, Experiment 2; see also Scherer and Oshinsky 1977). The lack of independent cue manipulation may explain why some early studies of emotional expression in music reported that some cues were associated with opposing emotional expressions.

## A theoretical approach

### Origin of the code

Juslin (1997a, 2000, 2001) proposed that the acoustic code used in emotional expression in

music performance reflects two primary factors. The first factor is innate patterns for vocal expressions of 'basic emotions'. According to this notion—partly inspired by Herbert Spencer (1857)—the origin of the 'expressive code' is to be found in involuntary and emotion-specific physiological changes associated with emotional reactions, which strongly influence different aspects of voice production (for a review of the relationships among emotion, physiology, and voice, see Juslin and Scherer 2005). This notion was later named 'Spencer's law' by Juslin and Laukka (2003), who also offered strong empirical support from an extensive meta analysis of parallels between vocal and musical expression of emotions. One implication of this finding is that 'basic emotion' categories will be privileged, because of their biological preparedness for effective communication. This is not to deny that more precise emotional contents can also be conveyed under certain circumstances (e.g., through social or musical convention).

The second factor governing emotional expression in performance is social learning of various forms. This is a life-long process, beginning with the early interaction between mother and infant. (In fact, this 'imprinting' process might begin even prior to birth; see Chapter 20.) For example, when mothers talk to their infants, if they want to calm their infant, they reduce the tempo and intensity of their speech, and talk with slowly falling pitch contours, whereas if they want to express disapproval towards some unfavourable activity, they employ brief, sharp, and staccato-like contours (Papoušek 1996). Although the code used by mothers seems to be universal, the expressive style of the mother modulates the expressive style of the infant. This modulation of expressive skills continues throughout life, as one accumulates experience and learns links between cues and extra-musical phenomena, such as movement (see Chapter 34).

## Nature of the code: modelling the communicative process

Studies of emotional expression in music performance present some puzzling findings, which can only be explained if the nature of the communication process is explored in greater detail. For instance, how can the communicative process usually be successful despite the fact that different instruments offer different cues, and that there are so wide individual differences among performers in expressive style? One way of capturing the crucial characteristics of the communicative process is to conceptualize it in terms of a variant of Egon Brunswik's (1956) *lens model*, as first suggested and implemented by Juslin (1995, 2000).

The modified lens model (Figure 35.2) captures how a performer 'encodes' (i.e., expresses) emotions by means of a set of acoustic cues (e.g., variations of tempo, sound level, or timbre) that are probabilistic (i.e., uncertain) and partly redundant. The emotions are 'decoded' (i.e., recognized) by listeners, who utilize these same cues to 'infer' the expression. The cues are probabilistic in that they are not perfectly reliable indicators of the intended expression. For example, a fast tempo is not a perfectly reliable indicator of happiness expressions, because a fast tempo occurs also in anger expressions. Thus, relationships among cues and emotions are merely correlational—they do not have a one-to-one mapping. Consequently, performers and listeners have to combine the cues for reliable communication to occur. However, this is not just a matter of pattern matching, because cues contribute in an additive fashion to listeners' judgements—each cue is neither necessary, nor sufficient, but the larger the number of cues used, the more reliable the communication. The redundancy among cues partly reflects how sounds are produced on musical instruments (e.g., a harder string attack produces a tone that is both louder and sharper in timbre). A mathematical formulation of the relationships among the various parts of the lens model, *the lens model equation*, was first formulated by Hursch *et al.* (1964). When applied to music performance (Juslin 2000), it allows one to explain the success (or not) of the communicative process by decomposing it into different components (e.g., the consistency of cue utilization, the degree to which the performer's cue utilization is matched to the listeners' cue utilization). Such a decomposition is highly useful if one wants to improve the communicative process (for evidence, see Juslin *et al.* 2006).

Although the lens model may appear simple, it has a crucial implication for research on

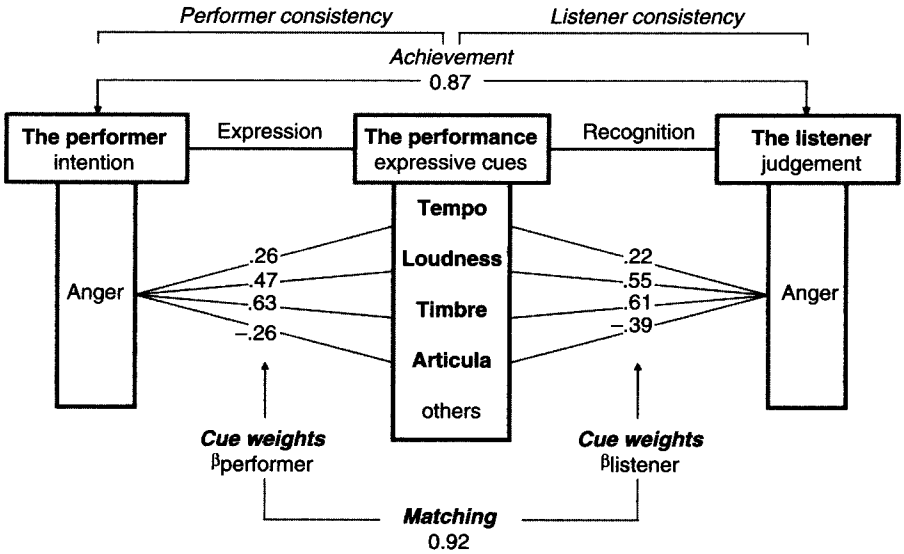



**Fig. 35.2** Modified lens model of communication of emotions in music performance (adapted from Juslin 1995). *Achievement* ($r_a$) refers to the relationship between the performer's intention (e.g., intending to express sadness) and the listener's judgement (e.g., perceiving sadness). It is a measure of how well the performer succeeds in communicating an emotion to listeners. *Cue weight* ($\beta_1$, $\beta_2$, $\beta_3$ ... ) refers to the strength of the relationship between a cue (e.g., tempo), on the one hand, and a performer's intentions or listeners' judgements, on the other. *Matching* (G) refers to the degree of similarity between the performer's and the listeners' use of acoustic cues, respectively. *Consistency* ($R_e$ and $R_s$) refers to the degree of consistency with which the performer and listeners, respectively, are able to use the cues.

communication: if the cues are redundant to some extent, more than one way of using the cues may lead to a similarly high level of decoding accuracy, because different cues may substitute for one another—so-called *vicarious functioning*. Hence, Brunswik's lens model may explain why there is accurate communication of emotions even when the cues are used inconsistently across different performers or pieces of music. Multiple cues that are partly redundant yield a robust communicative system that is forgiving towards deviation from optimal cue utilization. Therefore, researchers should not expect acoustic cues in expressive performances to conform rigidly to the patterns shown in Table 35.2; the communicative system is so robust that perfect consistency is not required. Robustness comes with a price, however. The redundancy of the cues means that the same information is conveyed by many cues. This limits the information capacity, which is another reason why 'basic emotion' categories are the easiest to convey.

The 'original' version of the lens model was limited to performance cues (Juslin 2000), but listeners' perception of emotions may also be affected by interactions between composer and performer cues. Hence, we have suggested an *expanded lens model* (Juslin and Lindström 2003; see Figure 35.3) in which both composer cues and performance cues are included to make it possible to explore their relative contributions. Furthermore, important interactions between performer and composition cues are included as separate predictors in the regression model. An experimental study based on this framework, using synthesized and systematically varied pieces of music, showed that about 75–85 per cent of the variance in listeners' emotion ratings could be explained by a linear combination of the main effects alone. Moreover, results suggested that the interactions between

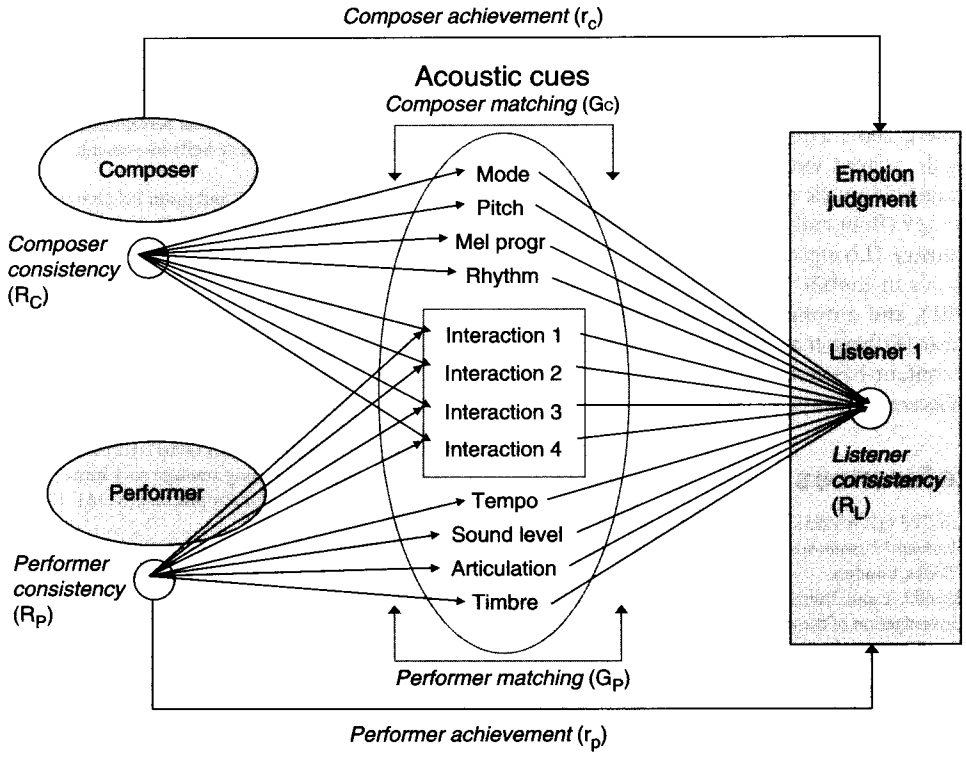



**Fig. 35.3** Extended lens model of communication of emotions in music (from Juslin and Lindström 2003).

composer and performer cues made small—although not negligible—contributions to the predictive power. That the cues contribute primarily in an additive fashion to judgements is partly why the communicative process is so robust, because additive cues may compensate for one another in ways that multiplicative cues cannot.

## Directions for future research

As revealed by this review, researchers have made much progress in understanding how music performers are able to communicate emotions via music performances. However, there are several limitations that need to be addressed in future research. One limitation is that the studies thus far have focused mainly on Western music (but see Balkwill and Thompson 1999; Clayton 2005). Hence, it remains to be seen how much of the results will generalize to other musical cultures. Even within the current culture, more research is required to examine more subtle, style-specific aspects of expression that contribute to the communicative process (e.g., ornamentation). Another important future direction is to study more closely how the different aspects of expression (e.g., structure marking, emotional expression, motion) interact with one another (Juslin *et al.* 2002). So far, the components have been studied either separately, or as if they all consisted of a homogeneous single entity. Closer collaboration among performance researchers within different subspecialties represents a promising avenue towards a deeper understanding of performance expression in all its complexity. One further topic that currently receives increasing attention is the influence of other perceptual channels. While most studies of expression in music performance have focused

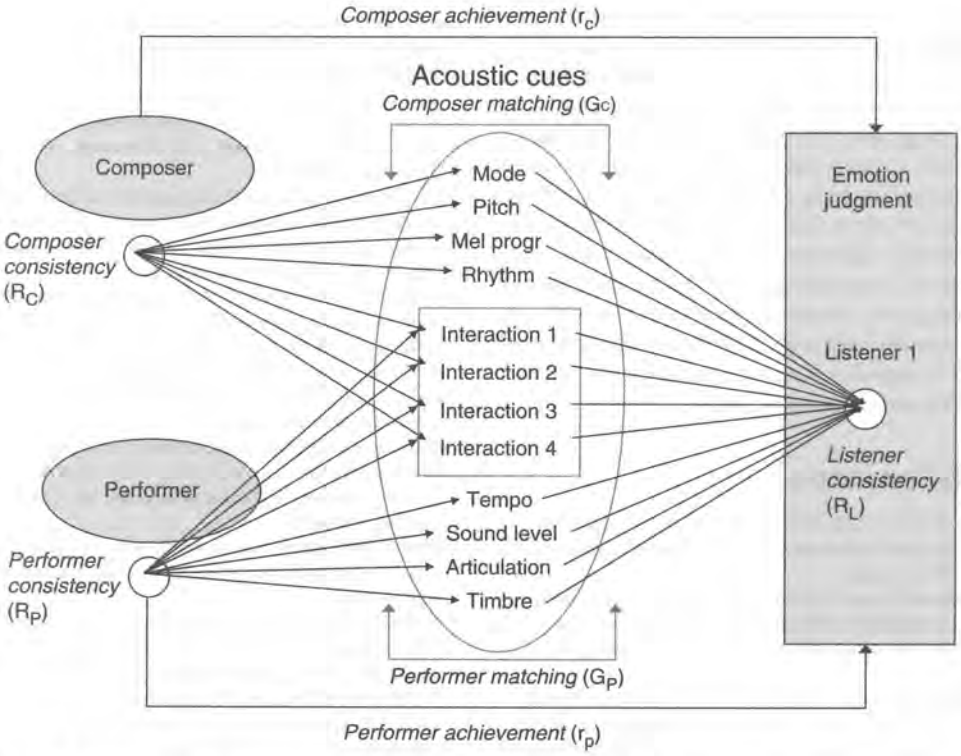



**Fig. 35.3** Extended lens model of communication of emotions in music (from Juslin and Lindström 2003).

composer and performer cues made small—although not negligible—contributions to the predictive power. That the cues contribute primarily in an additive fashion to judgements is partly why the communicative process is so robust, because additive cues may compensate for one another in ways that multiplicative cues cannot.

## Directions for future research

As revealed by this review, researchers have made much progress in understanding how music performers are able to communicate emotions via music performances. However, there are several limitations that need to be addressed in future research. One limitation is that the studies thus far have focused mainly on Western music (but see Balkwill and Thompson 1999; Clayton 2005). Hence, it remains to be seen how much of the results will generalize to other musical cultures. Even within the current culture, more research is required to examine more subtle, style-specific aspects of expression that contribute to the communicative process (e.g., ornamentation). Another important future direction is to study more closely how the different aspects of expression (e.g., structure marking, emotional expression, motion) interact with one another (Juslin *et al.* 2002). So far, the components have been studied either separately, or as if they all consisted of a homogeneous single entity. Closer collaboration among performance researchers within different subspecialties represents a promising avenue towards a deeper understanding of performance expression in all its complexity. One further topic that currently receives increasing attention is the influence of other perceptual channels. While most studies of expression in music performance have focused

only on auditory features, it seems that visual features of the performer can also make important contributions to the perceived expression of the music in live settings (see Dahl and Friberg 2007; Thompson *et al.* 2005). Finally, much current work concerns applications of findings in music education (Juslin *et al.* 2006), therapy (Bunt and Pavlicevic 2001), the gaming industry (Livingstone and Brown 2005), ring signals in mobile phones (Bresin and Friberg 2001), and automatic analysis of emotions in music (Friberg *et al.* 2007) which might be used in content-based music searches at the Internet (Wieczorkowska *et al.* 2005).

## References


Bach CPE (1778/1985). *Essay on the true art of playing keyboard instruments*, trans. WJ Mitchell. Eulenburg Books, London.

Balkwill L-L and Thompson WF (1999). A cross-cultural investigation of the perception of emotion in music: psychophysical and cultural cues. *Music Perception*, **17**, 43–64.

Behrens GA and Green SB (1993). The ability to identify emotional content of solo improvisations performed vocally and on three different instruments. *Psychology of Music*, **21**, 20–33.

Bresin R and Friberg A (2000). Emotional coloring of computer-controlled music performance. *Computer Music Journal*, **24**, 44–62.

Bresin R and Friberg A (2001). Expressive musical icons. In J Hiipakka, N Zakarov and T Takala, eds, *Proceedings of the International Conference on Auditory Display (ICAD) 2001*, 141–143. ICAD, Espoo, Finland.

Brunswik E (1956). *Perception and the representative design of experiments*. University of California Press, Berkeley, CA.

Buelow GJ (1983). Johann Mattheson and the invention of the Affektenlehre. In GJ Buelow and HJ Marx, eds, *New Mattheson studies*, 393–407. Cambridge University Press, Cambridge.

Bunt L and Pavlicevic M (2001). Music and emotion: perspectives from music therapy. In PN Juslin and JA Sloboda, eds, *Music and emotion: theory and research*, 181–201. Oxford University Press, Oxford.

Burland K and Davidson JW (2004). Tracing a musical life transition. In JW Davidson, ed., *The music practitioner*, 225–249. Ashgate, Aldershot.

Clarke EF (1988). Generative principles in music performance. In JA Sloboda, ed., *Generative processes in music: the psychology of performance, improvisation, and composition*, 1–26. Clarendon Press, Oxford.

Clarke EF, Cook N, Harrison B and Thomas P (2005). Interpretation and performance in Bryn Harrison's *être-temps*. *Musicae Scientiae*, **9**, 31–74.

Clayton M (2005). Communication in Indian raga performance. In D Miell, DJ Hargreaves and R MacDonald, eds, *Musical communication*, 361–381. Oxford University Press, Oxford.

Dahl S and Friberg A (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception*, **24**, 433–454.

Friberg A, Schoonderwaldt E and Juslin PN (2007). CUEX: an algorithm for automatic extraction of expressive tone parameters in music performance from acoustic signals. *Acta Acustica united with Acustica*, **93**, 411–420.

Gabrielsson A and Juslin PN (2003). Emotional expression in music. In RJ Davidson, KR Scherer and HH Goldsmith, eds, *Handbook of affective sciences*, 503–534. Oxford University Press, Oxford.

Gilden DL (2001). Cognitive emissions of 1/f noise. *Psychological Review*, **108**, 33–56.

Gullberg A-K and Brändström S (2004). Formal and non-formal music learning amongst rock musicians. In J Davidson, ed., *The music practitioner*, 161–174. Ashgate, Aldershot.

Hevner K (1935). Expression in music: a discussion of experimental studies and theories. *Psychological Review*, **42**, 186–204.

Hursch CJ, Hammond KR and Hursch JL (1964). Some methodological considerations in multiple-cue probability studies. *Psychological Review*, **71**, 42–60.

Juslin PN (1995). Emotional communication in music viewed through a Brunswikian lens. In G Kleinen, ed., *Music and expression: Proceedings of the Conference of DGM and ESCOM, Bremen, 1995*, 21–25. University of Bremen, Bremen, Germany.

Juslin PN (1997a). Emotional communication in music performance: a functionalist perspective and some data. *Music Perception*, **14**, 383–418.

Juslin PN (1997b). Perceived emotional expression in synthesized performances of a short melody: capturing the listener's judgment policy. *Musicæ Scientiæ*, **1**, 225–256.

Juslin PN (1997c). Can results from studies of perceived expression in musical performances be generalized across response formats? *Psychomusicology*, **16**, 77–101.

Juslin PN (2000). Cue utilization in communication of emotion in music performance: relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance*, **26**, 1797–1813.

Juslin PN (2001). Communicating emotion in music performance: a review and a theoretical framework. In PN Juslin and JA Sloboda, eds, *Music and emotion: theory and research*, 309–337. Oxford University Press, Oxford.

Juslin PN (2003). Five facets of musical expression: a psychologist's perspective on music performance. *Psychology of Music*, **31**, 273–302.

Juslin PN, Friberg A and Bresin R (2002). Toward a computational model of expression in music performance: the GERM model. *Musicæ Scientiæ, Special Issue 2001–2002*, 63–122.

Juslin PN, Friberg A, Schoonderwaldt E and Karlsson J (2004). Feedback-learning of musical expressivity. In A Williamon, ed., *Musical excellence: strategies and techniques for enhancing performance*, 247–270. Oxford University Press, Oxford.

Juslin PN, Karlsson J, Lindström E, Friberg A and Schoonderwaldt E (2006). Play it again with feeling: Computer feedback in musical communication of emotions. *Journal of Experimental Psychology: Applied*, **12**, 79–95.

Juslin PN and Laukka P (2003). Communication of emotions in vocal expression and music performance: different channels, same code? *Psychological Bulletin*, **129**, 770–814.

Juslin PN and Laukka P (2004). Expression, perception, and induction of musical emotions: a review and a questionnaire study of everyday listening. *Journal of New Music Research*, **33**, 217–238.

Juslin PN and Lindström E (2003). Musical expression of emotions: modeling composed and performed features. Paper presented at 5th Conference of the European Society for the Cognitive Sciences of Music (ESCOM), Hanover, Germany, September 2003.

Juslin PN and Madison G (1999). The role of timing patterns in recognition of emotional expression from musical performance. *Music Perception*, **17**, 197–221.

Juslin PN and Scherer KR (2005). Vocal expression of affect. In JA Harrigan, R Rosenthal and KR Scherer, eds, *The new handbook of methods in nonverbal behavior research*, 65–135. Oxford University Press, New York.

Juslin PN and Sloboda JA (eds) (2001). *Music and emotion: theory and research*. Oxford University Press, Oxford.

Laukka P and Juslin PN (2007). Similar pattern of age-related differences in emotion recognition from speech and music. *Motivation and Emotion*, **31**, 182–191.

Lindström E (1999). Expression in music: interaction between performance and melodic structure. Paper presented at the Meeting of the Society for Music Perception and Cognition, Evanston, USA, 14–17 August, 1999.

Lindström E, Juslin PN, Bresin R and Williamon A (2003). 'Expressivity comes from within your soul': a questionnaire study of music students' perspectives on expressivity. *Research Studies in Music Education*, **20**, 23–47.

Livingstone SR and Brown AR (2005). Dynamic response: real-time adaption for music emotion. In Y Pisan, ed., *Proceedings of the Second Australasian Conference on Interactive Entertainment*, 105–111. Creativity and Cognition Studios Press, Sydney.

Madison G (2000). Properties of expressive variability patterns in music performance. *Journal of New Music Research*, **29**, 335–356.

Marchand DJ (1975). A study of two approaches to developing expressive performance. *Journal of Research in Music Education*, **23**, 14–22.

Mattheson J (1739/1954). *Der volkommene Capellmeister*. Bärenreiter Verlag, Basel.

Meyer LB (1956). *Emotion and meaning in music*. University of Chicago Press, Chicago, IL.

Minassian C, Gayford C and Sloboda JA (2003). Optimal experience in musical performance: a survey of young musicians. Paper presented at the Meeting of the Society for Education, Music, and Psychology Research, London, March 2003.

Neumann F (1978). *Ornamentation in Baroque and post-Baroque music: with special emphasis on JS Bach*. Princeton University Press, Princeton, NJ

Palmer C (1997). Music performance. *Annual Review of Psychology*, **48**, 115–138.

Papoušek M (1996). Intuitive parenting: a hidden source of musical stimulation in infancy. In I Deliége and JA Sloboda, eds, *Musical beginnings. Origins and development of musical competence*, 89–112. Oxford University Press, Oxford.

Resnicow JE, Salovey P and Repp BH (2004). Is recognition of emotion in music performance an aspect of emotional intelligence? *Music Perception*, **22**, 145–158.

Scherer KR and Oshinsky JS (1977). Cue utilisation in emotion attribution from auditory stimuli. *Motivation and Emotion*, **1**, 331–346.

Shove P and Repp BH (1995). Musical motion and performance: theoretical and empirical perspectives. In J Rink, ed., *The practice of performance*, 55–83. Cambridge University Press, Cambridge.

Siegwart H and Scherer KR (1995). Acoustic concomitants of emotional expression in operatic singing: the case of Lucia in Ardi gli incensi. *Journal of Voice*, **9**, 249–260.

Sloboda JA and Lehmann AC (2001). Tracking performance correlates of changes in perceived intensity of emotion during different interpretations of a Chopin piano prelude. *Music Perception*, **19**, 87–120.

Spencer H (1857). The origin and function of music. *Fraser's Magazine*, **56**, 396–408.

Thompson S (2006). Audience responses to a live orchestral concert. *Musicae Scientiae*, **10**, 215–244.

Thompson WF, Graham P and Russo FA (2005). Seeing music performance: visual influences on perception and experience. *Semiotica*, **156**, 203–227.

Timmers R and Ashley R (2007). Emotional ornamentation in performances of a Handel sonata. *Music Perception*, **25**, 117–134.

Wieczorkowska A, Synak P, Lewis R and Zbigniew WR (2005). Extracting emotions from music data. *Lecture Notes in Computer Science*, **3488**, 456–465.

Yamasaki T (2002). Emotional communication in improvised performance by musically untrained players. In T Kato, ed., *Proceedings of the 17th International Congress of the International Association of Empirical Aesthetics*, 521–524. IAEA, Osaka.

CHAPTER 36

# Optimizing physical and psychological health in performing musicians

Dianna T. Kenny and Bronwen Ackermann

## Introduction

Performing musicians face a number of physical, social and psychological challenges that must be mastered if their musical career is to be both rewarding and sustainable. However, musicians are at high risk of physical and psychological strain and injury in the execution of their art. Physical and psychological stressors exert reciprocal and synergistic effects on the musician, and careful analysis of the intrinsic characteristics of the performer and the extrinsic demands on the musician must be made in order to develop appropriate interventions. In this chapter we provide an overview of the risks and challenges facing musicians with the aim of developing awareness and understanding of how to prevent and manage these challenges. The chapter is divided into two sections: physical challenges and psychological challenges, focusing on music performance anxiety. In each section, we outline the key issues and then provide an overview of evidence-based treatment.

## Physical health issues

Musicians are like athletes because both require superior sensorimotor integration (Alternmuller *et al.* 2000), neuro-musculoskeletal skill and many hours of training and practice to achieve mastery (Tubiana 2000). Several large epidemiological studies have shown high physical injury rates among musicians (Manchester 2006). Performer-related risk factors for injury include poor posture, poor physical condition, inadequate instrument set-up, long hours of playing, insufficient rest breaks and inefficient movement patterns (or poor technique) (Ackermann and Adams 2004a).

Although musicians may suffer injury from non-performance related causes such as lifting and carrying awkward or heavy instruments and suitcases (when on tour), demanding work schedules, sitting on poorly designed orchestral chairs, temperature variations (Manchester 2006), demanding repertoire and poor visibility of music scores (Horvath 2002), the majority of musicians' injuries are over-use injuries (Dawson *et al.* 1998) with soft-tissue symptoms predominating (Pascarelli and Hsu 2001).

Musicians of all ages and levels of skill are vulnerable to injury, but the risk increases as hours of playing increases. The neck, upper limbs and lower back (Fjellman-Wiklund *et al.* 2003) are vulnerable with upper limb over-use injuries comprising 75–85 per cent of all injuries (Slade *et al.* 1999). Strains of the muscle–tendon unit predominate, with other common problems including inflammatory disorders such as tenosynovitis, arthritic problems and hypermobility (also called double-jointedness) i.e. joints that stretch more than normal (Dawson 2002). People with hypermobile joints may be more easily injured and may be more at risk of developing problems from muscle over-use, as muscles must work harder to control joint movement.

Injury types vary according to instrument type, gender (Engquist *et al.* 2004), years of performing, repertoire, hours of practice and age (Warrrington *et al.* 2002). Older musicians more typically develop degenerative conditions (i.e. conditions that have a gradual deterioration in the structure of a body part with a consequent loss of the part's ability to function) while younger musicians suffer more from performance-related musculoskeletal pain (Warrington *et al.* 2002).

Effective training for any athletic pursuit, including music performance, involves a fine balance between working hard enough to continually improve performance, while simultaneously resting enough to avoid incurring an over-use injury, a task assisted by specifically designed cross-training programmes (Marieb 2001). Muscle fatigue, which may arise from central (brain) or peripheral (muscle) fatigue (Gandevia 2001) due to over-practising without adequate rest breaks (Ackermann and Adams 2004a) may be a factor in developing altered movement patterns that can be a precursor to injury. Exercise aimed at conditioning muscles may reduce the effect of muscle fatigue by increasing muscle cross-sectional area and creating neural adaptations that lead to an improved ability to recruit motor units (Herbert 1993).

## Types of injuries

### Muscle/tendon injuries

Muscles are the primary source of force needed for the performance of motor skills such as instrumental playing. An occupational over-use injury may result from insufficient muscle tolerance to cope with the quality, quantity and rate of task demands (Hagberg *et al.* 1997). Dynamic postures are less likely to cause pain than static postures due to the short rests that occur between the bursts of activity of muscles involved in dynamic movements (Vergara and Page 2002). Some strategies for musicians within task-constraints are possible for small-scale movements that can allow these dynamic motions to occur (Rolland 1974). Pablo Casals intentionally practised relaxing between phrases, no matter how difficult the material, to minimize the build up of adverse muscle tension (Ma 1986).

### Joint injuries

Joint injuries in musicians are degenerative, likely to be related to repetitive use, but more specifically to regional overload. For example, the right thumb of a clarinettist carries the whole weight of the instrument and shows early degenerative changes (Chesky *et al.* 2000). Ergonomic devices such as a neck strap to carry the weight of the clarinet may effectively reduce strain on the thumb, although long-term effects of transmitting this load through the neck are unclear (Chesky *et al.* 2000).

### Nerve compression disorders

Nerve compression syndromes in musicians are related to the demands and nature of musicians' work, such as sustained awkward positions, sustained muscle contraction pressure, or the compression arising from having to support the instrument itself (Spinner and Amadio 2000). Common nerve-entrapment disorders include symptoms from compression of the ulna nerve at the elbow, the median nerve at the wrist, cervical radiculopathies, occasional digital neuropathies and symptomatic thoracic outlet syndrome (Schuele and Lederman 2004).

### Central nervous system disorders

The most common performance-related condition of the central nervous system (CNS) affecting musicians is focal dystonia, a movement dysfunction syndrome thought to be due to disruptions between central sensory processing and motor output (Chen and Hallett 1998). Focal dystonias involve abnormal, often twisting movements that tend to affect performers in a similar way, with the left ring finger the main reported site of the dysfunction in violinists (Hochberg and Hochberg 2000). While a definitive cause remains unclear, the amount of movement repetition is a major factor in developing dystonias (Hochberg and Hochberg 2000), and these are potentially more potent if abnormal biomechanical factors are present (Wilson 2000).

## Treatment

Treatment, assessment and management procedures are still largely based on clinical experience rather than scientific research (Schuele and Lederman 2004). A primary goal of treatment is

not only relief of symptoms but restoration of function, addressing both local and general effects of an injury (Herring and Kibler 1998). Psychological factors may be involved in the genesis or maintenance of physical problems. Spahn *et al.* (2001) found significant rates (25%) of somatization and somatoform disorders in musicians presenting to hand surgery clinics. Most of their sample (75%) did not attribute their current physical problems to psychosocial factors, even though there was no identified organic cause for their physical problems. The authors concluded that psychosomatic aspects play a decisive role in somatic problems of musicians and that these should be addressed in treatment to avoid unwarranted medical interventions.

### Rest and rehabilitation

There is no benefit from prolonged rest for soft tissue injuries in the majority of cases (Nash *et al.* 2004). Scott (1997) defines the term rest for soft tissue injuries with the mnemonic 'Resume Exercise below Soreness Threshold', and emphasizes the importance of continuing to play in a reduced and careful fashion wherever possible without causing aggravation of the injury. Guidelines for rehabilitating sports overuse injuries stress the importance of relative rest, where overall fitness is maintained even in the acute stage of an injury while rehabilitating the injured part back to performance requirements (Herring and Kibler 1998). Relative rest can be achieved if necessary with the assistance of splints or ergonomic aids.

Poor muscle balance around the shoulder girdle is frequently considered to be a cause of upper limb symptoms in the musician, and restoring a good balance of muscle condition as well as strengthening postural muscles in this area is an important part of rehabilitation (Chamagne 2000).

### Retraining

Retraining is used for the treatment of both neurological and musculoskeletal conditions occurring in musicians. Some treatment approaches for focal dystonia include splinting with specific exercises (Candia *et al.* 1999), postural and movement retraining strategies (Chamagne 2000), proprioceptive and sensory retraining techniques (Ackermann and Adams 2005), and the slow-down exercise regime involving task-specific music drills (Sakai 2006). Movement patterns can be re-trained by restoring good muscle balance and functioning of the arm as a whole kinetic chain unit to correct any underlying mechanical inefficiencies (Dreyer and Boden 1999).

### Manual therapy

Trigger point therapy may be useful for soft tissue problems, particularly when palpation of active trigger points reproduces the musicians' symptoms (Davies 2002). Joint mobilization techniques of the cervical spine, thoracic spine, forearm, wrist and hand are necessary to restore full functional range of movement for musicians following an injury (Kember 1998).

## Prevention of playing-related injuries

Prevention is the best form of management for occupational over-use injuries (Melhorn 1998). Key factors in injury prevention include awareness of correct postural requirements, technique and biomechanics involved in playing one's instrument, and maintaining overall good physical condition that is achieved by warming up, stretching (Zaza 1994) and strength and endurance training (Marieb 2001).

### Posture

Correct posture minimizes stress applied to each joint. Poor posture, defined as a 'faulty relationship of the various segments of the body, producing increased stress on supporting structures' (Aaras *et al.* 2001), is a major risk for injury because musicians have to maintain awkward, relatively static postures over extended time periods (Brandfonbrener 2000). Musicians must maintain their posture within physiological boundaries, even with asymmetrical instruments such as the violin and flute, with good proximal muscle support and weight balance to avoid injury and to allow optimal fine control of movements during performance (Tubiana *et al.* 1989). The musical instrument should be considered as an extension of the musicians' body (Dommerholt 2000).

Ideal sitting posture alignment occurs with hips and knees at 90° of flexion, with a 10° backward

inclination of a supportive back on the chair (Kendall *et al.* 1993), or a level seat and a backrest set back 20° to minimize lumbar loadings (Bonney and Corlett 2002). Sitting should also be balanced between sides by weight-bearing evenly on both sides, as excessive weight-bearing on one side may lead to lateral stresses on the lumbar discs (Cailliet 1990).

### Technique, biomechanics and physical condition

Within a technique or movement pattern on a musical instrument, early more rigid movements are replaced by efficient movements within anatomical and task constraints as the skill is mastered (Sparrow and Newell 1998). Musicians working under pressure may be so focused on achieving a musical goal that a distortion of technique or posture occurs, and these aberrations may then become established in the motor programme, causing a subconscious alteration in technique (Wilson 2000). These physical adaptations do not represent the players' normal technique and may lead to maladaptive changes including muscle misuse or more serious neurological disorders such as focal dystonia (Wilson 2000). Excessive muscle tension also accompanies mechanical inefficiencies in performance (Ma 1986).

### Warming up and stretching

Regular stretching improves flexibility (Wilkinson 1992), and musicians benefit from instrument-specific stretching programmes (Markison 1998). Stretching should be performed regularly to maintain adequate range for performance demands (Norris 1993). Stretching prior to performance only may not be sufficient to minimize injury risk (Pope *et al.* 2000). General warming-up and cooling-down routines are recommended (Markison 1998). Players using instrumental warm-up strategies may protect themselves against the development of a playing-related injury (Zaza and Farewell 1997).

### Strength or endurance conditioning

Poor physical condition is a predisposing factor to playing-related injuries (Ackermann and Adams 2004a). Instrument-specific strength and endurance training is effective in reducing injury frequency and intensity as well as reducing the perceived exertion associated with practising their instrument (Ackermann *et al.* 2002). In contrast, participating in unsupervised general sports was not found to provide any benefits in terms of injury prevention for musicians (van Hees 1997).

### Ergonomic aids and advice

Various ergonomic interventions aimed at reducing physical load include hand splint adaptations for the trombone to assist with reach difficulties (Quarrier and Norris 2001), development of polymer drumsticks with reduced vibration characteristics (Zaza *et al.* 2000), a neck strap to carry the weight of the clarinet and bassoon may effectively reduce strain on the thumb (Chesky *et al.* 2000), and many other designs such as angle-headed flutes, key extensions on wind instruments, and remodelling of viola or guitar bodies (Norris 2000). Instruction on good lifting technique is important for musicians who may injure their lower back as a result of carrying heavy or awkward-shaped instruments (Fjellman-Wiklund *et al.* 2003).

## Psychological health issues

### Stressors experienced by musicians

Sternbach (1995) described the working conditions of professional musicians as generating a 'total stress quotient' that far exceeds that observed in other professions. Like elite athletes, performing artists must maintain their skills at peak form, endure many hours of solitary, repetitive practice, constantly self-evaluate their performances and subject their public performances to close scrutiny. They are required to work in a pattern akin to shift work, be available to travel to performance venues, leave their families while on tour, adjust to changing time zones, live at close quarters with colleagues and peers, and cope with financial insecurity. For these reasons, it can be difficult to differentiate between the occupational and physical stressors discussed in the previous section and psychological problems that can arise in individual musicians that may require individualized psychological intervention.

Frequently reported psychosocial issues while on tour or working on contract with interstate

and overseas orchestras include loneliness, homesickness, sexual frustration and relationship breakdown. Occupational issues include language barriers, unfamiliar backstage arrangements at concert venues and variable quality of dressing rooms. In addition to these psychological and occupational stressors, there is the physical stress associated with moving instruments and luggage, setting up on different stages, adjusting to differently shaped chairs at every venue, sleeping in different beds with different pillows, coping with jet lag, general fatigue and lack of sleep. New injuries or pains are frequently reported by professional musicians on tour, as a direct result of these factors (Ackermann 2002).

Individuals vary in their capacity to cope with such stressful working conditions. However, since not all performers suffer the same degree of psychological distress or indeed report the same levels of occupational stress, individual differences in a range of psychological characteristics are likely to account for variations in the degree to which musicians experience symptoms. For example, the difficulty in coping may be compounded for those who are also highly anxious, who lack confidence in their abilities and who engage in unhelpful strategies to deal with their anxieties, such as the regular consumption of alcohol, and licit (e.g. beta blockers) or illicit (e.g. marijuana) drugs. Since music performance anxiety is one of the most commonly reported psychological stressors in musicians, the remainder of this chapter will focus on music performance anxiety, its manifestations, consequences and treatment.

## Anxiety in public performance

Performance anxiety is a group of disorders that affect individuals in a range of performance settings, such as examinations, competitions and public speaking. Solo and orchestral instrumentalists (van Kemenade *et al.* 1995) and solo and choral vocal artists (Kenny *et al.* 2004) all report experiencing music performance anxiety. Music performance anxiety is also observed in young musicians (Kenny and Osborne 2006), and children and adolescents show a similar constellation of symptoms to college level music students and professional musicians (Osborne *et al.* 2005). Students of classical music report higher levels of performance anxiety than students of Jazz (Kaspersen and Gotestam 2002). Like all performance anxieties, music performance anxiety occurs on a continuum of severity from 'normal everyday healthy aspects of stress and anxiety that are intrinsic to the profession' (Brodsky 1996, p. 91) to the severely debilitating symptoms of stage fright.

## The phenomenology and determinants of music performance anxiety

Performance anxiety may occur as an isolated disorder, affecting only one specific part of a person's life, such as public speaking, test-taking or music performance. However, for a significant minority of those suffering performance anxieties, other comorbid disorders may be present, the most common of which is generalized anxiety disorder, which appears to co-occur in about one third of those presenting with severe performance anxiety (Sanderson *et al.* 1990). Generalized anxiety disorder is characterized by excessive, uncontrollable and often irrational worry about everyday concerns and is disproportionate to the actual source of worry. People with generalized anxiety disorder may have a long history of chronic worry and apprehension in most facets of their lives, not just in performance situations (American Psychological Association 2000).

Others may qualify for a diagnosis of social phobia (social anxiety) if the performer demonstrates significant impairment in interactions with others as well as in the performance setting and who otherwise meet the criteria for social phobia presented in DSM-IV-TR (American Psychological Association 2000). About 10–15% of those with a social phobia also meet criteria for clinical depression (Kessler *et al.* 1999). For a subgroup of music performance anxiety sufferers, there may be underlying psychological conflicts that need to be identified and resolved before the symptoms abate (Lazarus and Abramovitz 2004).

Two distinct aspects of performance anxiety have been identified—cognitive anxiety and somatic anxiety (Martens *et al.* 1990). High cognitive anxious individuals generally display a

consistent style of thinking about their performance that includes the following characteristics:

1 stronger negative expectancies before the event
2 stronger negative bias in their retrospective self-evaluations of performance
3 stronger expectation that their performance will be judged negatively by their examiners/audience
4 stronger concerns about the consequences of a poor performance
5 heightened responsiveness to changes in reactions of judges or audience
6 failure to derive comfort from evidence that they have handled the situation skilfully (Wallace and Alden 1997).

Somatic anxiety refers to the experiencing of a cluster of physical symptoms during an anxiety-provoking activity such as performing in front of an audience. It is characterized by muscle tension, agitation, and other phenomena such as trembling, sweating, dry mouth, shallow breathing and 'buttlerflies in the stomach' that are associated with the 'fright–fight–flight' response, first described by Cannon (1915) and subsequently by many researchers (Friedman and Silver 2007). These symptoms occur as a result of arousal of the sympathetic nervous system via the release of the hormone epinephrine (adrenaline) and to a lesser extent norepinephrine from the medulla of the adrenal glands (Gleitman *et al.* 2004).

## Music performance anxiety and performance quality

Performance quality is determined by a number of interacting factors, including the ability of the performer and the level of achievement attained as a musician (Fortune 2007), the degree to which the performance repertoire has been mastered (Wilson 2002), the fit between technical ability and task difficulty (Fehm and Schmidt 2006), the circumstances of the performance, for example, whether the performance will be evaluated by expert judges, audience characteristics (Brotons 1994), and the type and severity of anxiety experienced (Wang 2002).

To date, Wilson (2002) has offered the most comprehensive model of music performance anxiety that incorporates the performer's trait anxiety, or their constitutional and learned tendency to become anxious in response to socially stressful situations; the degree of task mastery achieved; and degree of situational stress, such that high anxiety is more likely to be experienced in situations where social or environmental pressures are high. Performance anxiety may exert either an enhancing or detrimental effect on performance depending upon the interaction between these three factors (Fehm and Schmidt 2006). For example, an individual with high trait anxiety will perform best with an easy, well-prepared piece in a relaxed environment, whereas an individual with low trait anxiety will perform better if the piece is challenging and performed in an evaluating environment, such as an exam or competition.

## Optimizing music performance

An optimal performance is determined by a complex interaction between person characteristics, task characteristics and performance setting. These include adequate preparation, achievement of task mastery such that the complex motor tasks required to perform the task have been (over) learnt to the point of being automatic (Oliveira and Goodman 2004), familiarization with the performance venue and adequate rehearsal with other performers in the case of ensemble performance. When all of these characteristics occur at an optimal level, the performer is said to be 'in the zone' (Young and Pain 1999) or to have achieved a state of 'flow' (Marr 2000). Another construct to describe peak performance, borrowed from sport psychology, is the 'individualized zone of optimal functioning' (IZO) (Hanin 1986), that is, the performer has achieved the optimal level of pre-performance anxiety that results in a peak performance. 'Optimal' pre-performance anxiety is a good predictor of performance quality (Turner and Raglin 1991).

Although one would expect highly anxious individuals to experience performance breakdown or impaired performance quality more often than low-anxiety performers, this is not usually the case (Strahan and Conger 1998).

Even in situations where the highly anxious do perform less well than the less anxious, social or artistic performance catastrophes are rare. There are a number of possible explanations for this. First, people display highly individual ways of responding to stressful situations. Very small changes in context or task-expectancy can change a person's appraisal of a situation as anxiety-provoking or not (Bandura 1991). Secondly, those situations that could produce a catastrophe are indeed rare, even for highly anxious individuals. Thirdly, highly anxious professional performers are likely to engage in a number of pre-performance compensatory activities, such as over-learning, additional rehearsals, or visiting the venue before the performance to ensure that their performance can withstand the additional anxiety they know they will experience during the performance (Kenny *et al.* 2004).

## Treatments for music performance anxiety

Many treatment programmes have been developed to assist the anxious or stressed musician. However, most of the available treatments have not been adequately assessed as to their effectiveness. These include:

1 Prescription pharmacological interventions such as antidepressants, benzodiazepines, beta-adrenergic receptor blockers, and busipone.
2 Meditative interventions (autogenic training, (self-) hypnosis, meditation, yoga)
3 Physiological and physically based interventions (aerobic exercise, Alexander technique, biofeedback, Feldenkrais, massage)
4 Relaxation therapies e.g. progressive muscle relaxation training
5 Cognitive and cognitive behavioural interventions: assertiveness training, attention-focusing techniques, cognitive behaviour therapy (cognitive restructuring), multimodal behavioural therapy, coping skills training, exposure therapy, goal setting, lifestyle changes (e.g. development of non-musical hobbies and interests), imagery (distraction and focused), mental rehearsal, stress inoculation, systematic desensitization and systematic rehearsal
6 Music therapy (music enhanced relaxation techniques; group therapy for musicians)
7 Psychotherapy (Nagel 2004).

Space permits only a brief overview, update and summary of the most commonly used and researched treatments and their effectiveness—the cognitive and behavioural therapies. A detailed review and analysis of other treatments for performance anxiety can be found in Kenny (2005).

### Cognitive, behavioural and cognitive behaviour therapies

Behaviour is determined by a combination of thoughts, feelings and past and present behaviours (Turkington *et al.* 2006). Three groups of therapies—behavioural (van de Wiel *et al.* 2007), cognitive (Willner, 2006) and cognitive behavioural (Butler *et al.* 2006)—are all based on the same principles, but use the available therapeutic techniques in different amounts. Behavioural therapies focus primarily on changing the dysfunctional behaviours that arise when people feel anxious. One of the main targets of behavioural therapies for anxiety disorders is excessive muscle tension, which is treated with deep muscle relaxation training (Conrad and Roth 2007) and systematic desensitization (Pagoto *et al.* 2006), a procedure in which the person is encouraged to imagine the feared or anxiety provoking situation in graded steps, called the fear hierarchy, until they can visualize the situation without experiencing the muscle tension that used to accompany the visualizations. Once the fear hierarchy has been mastered in the therapist's office (imaginal desensitization), people are encouraged to apply their new skills in the actual, anxiety-provoking situation (called *in vivo* desensitization) (Choy *et al.* 2007). This allows for behavioural exposure, i.e. repeatedly practising the task in the feared situation until the associated anxiety is reduced to manageable levels. For this to be effective, the task must have been mastered so that the performer can ensure successful performance in the more stressful situation (Rauch and Foa 2006).

Cognitive therapy focuses on mental states such as thoughts, feelings and images (Willner 2006). Dysfunctional cognitions, those that create emotional distress or maladaptive responses,

can be the result of cognitive deficits (i.e. difficulties with problem solving) (van Winkel *et al.* 2006), and/or errors or biases in information processing (Joormann *et al.* 2007). Cognitive therapy changes faulty thinking patterns that give rise to maladaptive behaviours, such as excessive muscle tension, avoidance of the feared situation, or impaired performance. In this therapy, people learn a skill called cognitive restructuring, a process whereby people replace negative, unproductive, catastrophic thinking with more rational, useful ways of understanding their problem situations (Murphy *et al.* 2007). Based on changed thinking patterns, people are often able to reassess or reappraise their feared situations in ways that reduce the perceived threat.

Cognitive behavioural therapy (CBT) is a combination of behavioural and cognitive interventions (Turkington *et al.* 2006). CBT uses educational and psychological interventions that are based on the idea that changing negative thinking patterns and behaviours can have a powerful effect on a person's emotions, which in turn can change behaviour in situations in which the negative emotions arose. Like all new learning, CBT requires commitment, practice and application in situations outside the therapy office (Yovel and Safren 2007). CBT is focused and directive, usually of short duration and is action-oriented,—it relies on the client's record-keeping, active participation, application and evaluation (Rees *et al.* 2005).

A more recent variant of CBT, multimodal behavioural therapy (Lazarus and Abramovitz 2004) proposes a multimodal approach to treatment that involves the assessment and management of any combination of seven components: behaviour, affect, sensations, imagery, cognition, interpersonal relationships, and drugs/biological factors. Research is needed to assess whether this more comprehensive approach will produce better results for the anxious musician.

## Prevention of music performance anxiety

Parenting style is important in the prevention of anxiety and depression in children (Barlow 2000). Parents who are responsive, unobtrusive, not overly protective, and who give children age-appropriate opportunities to exercise control over their environment foster a sense of control and mastery in their children that provides protection against the development of anxiety (Gar *et al.* 2005).

Most forms of performance anxiety are difficult to treat and anxiety levels after treatment rarely reduce to those of non-anxious people (Kenny 2005). The best form of treatment is to prevent its occurrence. Sound pedagogy, appropriate parental support and expectations, and the learning of self-management strategies early in one's musical education can help to mitigate the effects of entering a highly stressful profession.

Children should be offered frequent, low-stress opportunities to perform almost from the beginning of their musical training. These performances should be presented in a positive, non-judgemental way, so that young performers can learn that performance is an integral, enjoyable and manageable part of their musical education. Children should not be prematurely thrust into competitive environments whose focus is evaluation (such as auditions or competitions) and when they are, students need to be well prepared for the performance both musically and psychologically. Repertoire should be well within the technical capacity and interpretive abilities of the student and the material should be well-learned. Sensible pre-performance routines should be established that attend to the performer's physical well-being (having adequate diet and sleep), psychological well-being (developing positive self-statements and realistic self-appraisal capacity) and musical demands (adequate practice, appropriate level of complexity, cognitive and physical capacity commensurate with the musical demands). These strategies will enhance the student's sense of competence and control so that when confronted with critical performances, a strong sense of a competent self will guide a self-actualized performance.

# References

Aaras A, Horgen G, Bjorset HH, Ro O and Walsoe H (2001). Musculoskeletal, visual and psychosocial stress in VDU operators before and after multidisciplinary ergonomic interventions. A 6-year prospective study—part II. *Applied Ergonomics*, **32**, 559–571.

Ackermann B (2002). Managing the musculoskeletal health of musicians on tour. *Medical Problems of Performing Artists*, **17**(2), 63–67.

Ackermann B and Adams R (2004a). Perceptions of causes of performance-related injuries by music health experts and injured violinists. *Perceptual and Motor Skills*, **99**, 669–678.

Ackermann B and Adams R (2005). Finger movement discrimination in focal hand dystonia: a cellist case study. *Medical Problems of Performing Artists*, **20**(1), 77–81.

Ackermann B, Adams R and Marshall E (2002). Strength or endurance training for undergraduate music majors at a university. *Medical Problems of Performing Artists*, **17**(1), 33–41.

Altenmueller E, Gruhn W, Liebert G and Parlitz D (2000). The impact of music education on brain networks: evidence from EEG studies. *International Journal of Music Education*, **35**, 47–53.

American Psychiatric Association (2000). *Diagnostic and Statistical Manual (DSM-IV-TR)*. American Psychiatric Association, Washington, DC.

Bandura A (1991). Self-efficacy conception of anxiety. In S Schwarzer, RA Wicklund, eds, *Anxiety and self-focused attention*, 89–110. Harwood Academic Publishers, Amsterdam, Netherlands.

Barlow H (2000). Unravelling the mysteries of anxiety and its disorders from the perspective of emotion theory. *American Psychologist*, **55**(11), 1245–1263.

Bonney R and Corlett E (2002). Head posture and loading of the cervical spine. *Applied Ergonomics*, **33**, 415–417.

Brandfonbrener A (2000). Epidemiology and risk factors. In R Tubiana and P Amadio, eds, *Medical problems of the instrumentalist musician*, 171–194. Martin Dunitz Ltd, London.

Brodsky W (1996). Music performance anxiety reconceptualised: a critique of current research practices and findings. *Medical Problems of Performing Artists*, **11**, 88–98.

Brotons M (1994). Effects of performing conditions on music performance anxiety and performance quality. *Journal of Music Therapy*, **31**, 63–81.

Butler AC, Chapman JE, Forman EM and Beck AT (2006). The empirical status of cognitive-behavioral therapy: a review of meta-analyses. *Clinical Psychology Review*, **26**, 17–31.

Cailliet R (1990). Abnormalities of the sitting postures of musicians. *Medical Problems of Performing Artists*, **5**(4), 131–135.

Candia V, Elbert T, Altenmueller E, Rau H, Schafer T and Taub E (1999). Constraint-induced movement therapy for focal hand dystonia in musicians. *The Lancet*, **353**(9146), 42.

Cannon W (1915). *Bodily changes in pain, hunger, fear and rage: an account of recent researches into the function of emotional excitement*. Appleton, New York.

Chamagne P (2000). Functional assessment and rehabilitation of musician's focal dystonia. In R Tubiana and PC Amadio, eds, *Medical problems of the instrumentalist musician*, 343–362. Martin Dunitz Ltd, London.

Chen R and Hallett M (1998). Focal dystonia and repetitive motion disorders. *Clinical Orthopaedics and Related Research*, **351**, 102–106.

Chesky KS, Kondraske G and Rubin B (2000). Effect of elastic neck strap on right thumb force and force angle during clarinet performance. *Journal of Occupational and Environmental Medicine*, **42**(8), 775–776.

Choy Y, Fyer AJ and Lipsitz JD (2007). Treatment of specific phobia in adults. *Clinical Psychology Review*, **27**(3), 266–286.

Conrad A and Roth WT (2007). Muscle relaxation therapy for anxiety disorders: it works but how? *Journal of Anxiety Disorders*, **21**(3), 243–264.

Davies C (2002). Musculoskeletal pain from repetitive strain in musicians: insights into an alternative approach. *Medical Problems of Performing Artists*, **17**(1), 42–49.

Dawson WJ, Charness ME, Goode DJ, Lederman RJ and Newmark J (1998). What's in a name? Terminologic issues in performing arts medicine. *Medical Problems of Performing Artists*, **13**(2), 45–50.

Dawson WJ (2002). Upper-extremity problems caused by playing specific instruments. *Medical Problems of Performing Artists*, **17**(3), 135–139.

Dommerholt J (2000). Posture. In R Tubiana and PC Amadio, eds, *Medical problems of the instrumentalist musician*, 399–420. Martin Dunitz Ltd, London.

Dreyer SJ and Boden SD (1999). Natural history of rheumatoid arthritis of the cervical spine. *Clinical Orthopaedics and Related Research*, **366**, 98–106.

Engquist K, Orbaek P and Jakobsson K (2004). Musculoskeletal pain and impact on performance in orchestra musicians and actors. *Medical Problems of Performing Artists*, **19**(2), 55–61.

Fehm L and Schmidt K (2006). Performance anxiety in gifted adolescent musicians. *Journal of Anxiety Disorders*, **20**(1), 98–109.

Fjellman-Wiklund A, Brulin C and Sundelin G. (2003). Physical and psychosocial work-related risk factors associated with neck–shoulder discomfort in male and female music teachers. *Medical Problems of Performing Artists*, **18**(1), 33–41.

Fortune, J. (2007). Performance-related musculoskeletal disorders in tertiary level flute players and relationships with muscle tension, music performance anxiety, musical task complexity and musical ability. Unpublished Masters thesis, Sydney Conservatorium of Music, University of Sydney.

Friedman HS and Silver RC (eds) (2007). *Foundations of health psychology*. Oxford University Press, New York.

Gandevia SC (2001). Spinal and supraspinal factors in human muscle fatigue. *Psychological Reviews*, **81**(4), 1725–1789.

Gar NS, Hudson JL and Rapee RM (2005). Family factors and the development of anxiety disorders. In Hudson JL and Rapee RM, eds, *Psychopathology and the family*, 125–145. Elsevier Science, New York.

Gleitman H, Fridlund AJ and Reisber D (2004). *Psychology*, 6. Norton, New York.

Hagberg M, Christiani D, Courtney TK, Halperin W, Leamon TB and Smith TJ (1997). Conceptual and definitional issues in occupational injury etiology. *American Journal of Industrial Medicine*, **32**, 106–115.

Hanin YL (1986). State-trait anxiety research on sports in the USSR. In CD Spielberger and R Diaz, eds, *Cross-cultural anxiety*, 45–64. Hemisphere Publishing Corp/ Harper and Row Publishers, New York.

Herbert RD (1993). Human strength adaptations-implications for therapy. In J Crosbie and J McConnell, eds, *Key issues in musculoskeletal physiotherapy*, 142–171. Butterworth-Heinemann Ltd, Oxford.

Herring SA and Kibler WB (1998). A framework for rehabilitation. In WB Kibler, SA Herring, JM Press and PA Lee, eds, *Functional rehabilitation of sports and musculoskeletal injuries*, 98–108. Aspen Publishers Inc., Gaithersburg, MD.

Hochberg FH and Hochberg NS (2000). Occupational cramps/focal dystonias. In R Tubiana and PC Amadio, eds, *Medical problems of the instrumentalist musician*, 295–310. Martin Dunitz, London.

Horvath J. (2002). *Playing (less) hurt*. Morris Publishing, Kearney, NE.

Joormann J, Talbot L and Gotlib IH (2007). Biased processing of emotional information in girls at risk for depression. *Journal of Abnormal Psychology*, **116**(1), 135–143.

Kaspersen M and Gotestam KG (2002). A survey of music performance anxiety among Norwegian music students. *European Journal of Psychiatry*, **16**(2), 69–80.

Kember J (1998). The physical therapists contribution. In I Winspur and CB Wynn Parry, eds, *The musician's hand. A clinical guide*, 136–142. Martin Dunitz Ltd, London.

Kendall FP, McCreary EK and Provance PG (1993). *Muscles. Testing and function*, 4th edn. Williams and Wilkins, Baltimore, MD.

Kenny DT (2005). A systematic review of treatment for music performance anxiety. *Anxiety, Stress and Coping*, **18**(3), 183–208.

Kenny DT and Osborne MS (2006). Music performance anxiety: new insights from young musicians. In WF Thompson, ed., *Advances in cognitive psychology*. Retrieved 16 June 2006, from http://www.ac-psych.org.

Kenny DT, Davis P and Oates J (2004). Music performance anxiety and occupational stress amongst opera chorus artists and their relationship with state and trait anxiety and perfectionism. *Journal of Anxiety Disorders*, **18**, 757–777.

Kessler RC, Stang P, Wittchen HU, Stein M and Walters EE (1999). Lifetime comorbidity between social phobia and mood disorders in the US National Comorbidity Survey. *Psychological Medicine*, **29**, 555–567.

Lazarus AA and Abramovitz A. (2004). A multimodal behavioral approach to performance anxiety. *Journal of Clinical Psychology*, **60**, 831–840.

Ma (1986) cited in AG Brandfonbrener (2005). Interview with Yo Yo Ma (parts I and II). *Medical Problems of Performing Artists*, **20**(3), 140–148.

Manchester R (2006). Toward better prevention of injuries among performing artists. *Medical Problems of Performing Artists*, **21**(1), 1–2.

Marieb EN (2001). *Human anatomy and physiology*, 5th edn. Benjamin Cummins, New York.

Markison RE (1998). Adjustment of the musical interface. In I Inspur and C Wynn-Parry, eds, *The musician's hand. A clinical guide*, 149–159. Martin Dunitz, London.

Marr J (2000, October). Commentary: flow, intrinsic motivation, and second generation cognitive science. *The Online Journal of Sport Psychology*, **2**(3). Retrieved 17 May 2006 from http://www.athleticinsight.com/Vol2Iss3/Commentary_2.htm.

Martens R, Burton D, Vealey R, Bump L and Smith D (1990). The development of the Competitive State Anxiety Inventory-2 (CSAI-2). In R Martens, RS Vealey and D Burton, eds, *Competitive anxiety in sport*, 117–190. Human Kinetics, Champaign, IL.

Melhorn MJ (1998). Cumulative trauma disorders and repetitive strain injuries: the future. *Clinical Orthopaedics and Related Research*, **351**, 107–126.

Murphy WP, Yaruss JS and Quesal RW. (2007). Enhancing treatment for school-age children who stutter I. Reducing negative reactions through desensitization and cognitive restructuring. *Journal of Fluency Disorders*, **32**(2), 121–138.

Nagel JJ. (2004). Performance anxiety theory and treatment: one size does not fit all. *Medical Problems of Performing Artists*, **19**(1), 39–43.

Nash CE, Mickan SM, Del Mar CB and Glasziou PP (2004). Resting injured limbs delays recovery: a systematic review. *The Journal of Family Practice*, **53**(9), 706–712.

Norris R (1993). *The musician's survival manual: a guide to preventing and treating injuries in instrumentalists*. MMB Music Inc, Saint Louis, MO.

Norris RN (2000). Applied ergonomics. In R Tubiana and PC Amadio, eds, *Medical problems of the instrumentalist musician*, 595–613. Martin Dunitz Ltd, London.

Oliveira FTP and Goodman D (2004). Conscious and effortful or effortless and automatic: a practice performance paradox. *Perceptual and Motor Skills*, **99**, 315–324.

Osborne MS, Kenny DT and Holsomback R (2005). Assessment of music performance anxiety in late childhood: a validation study of the Music Performance Anxiety Inventory for Adolescents (MPAI-A). *International Journal of Stress Management*, **12**(4), 312–330.

Pagoto SL, Kozak AT, Spates CR and Spring B (2006). systematic desensitization for an older woman with a severe specific phobia: an application of evidenced-based practice. *Clinical Gerontologist*, **30**(1), 89–98.

Pascarelli EF and Hsu YP (2001). Understanding work-related upper extremity disorders: clinical findings in 485 computer users, musicians, and others. *Journal of Occupational Rehabilitation*, **11**(1), 1–21.

Pope R P, Herbert RD, Kirwan JD and Graham BJ (2000). A randomized trial of pre-exercise stretching for prevention of lower limb injury. *Medicine and Science in Sports and Exercise*, **31**(2), 271–277.

Quarrier NF and Norris RN (2001). Adaptations for trombone performance: ergonomic interventions. *Medical Problems of Performing Artists*, **16**(2), 77–80.

Rauch S and Foa E (2006). Emotional processing theory (EPT) and exposure therapy for PTSD. *Journal of Contemporary Psychotherapy*, **36**(1), 61–65.

Rees CS, McEvoy R and Nathan PR (2005). Relationship between homework completion and outcome in cognitive behaviour therapy. *Cognitive Behaviour Therapy*, **34**(4), 242–247.

Rolland P (1974). *The teaching of action in string playing.* Urbana String Research Associates, Urbana, IL.

Sakai N (2006). Slow-down exercise for the treatment of focal hand dystonia in pianists. *Medical Problems of Performing Artists*, **21**(1), 25–28.

Sanderson WC, DiNardo PA, Rapee RM and Barlow DH (1990). Symptom comorbidity in patients diagnosed with DSM-III-R anxiety disorders. *Journal of Abnormal Psychology*, **99**, 308–312.

Schuele SU and Lederman RJ (2004) Occupational disorders in instrumental musicians. *Medical Problems of Performing Artists*, **19**(3), 123–128.

Scott WA (1997). Overuse injuries. In RE Sallis and F Massimino, eds, *Essentials of sports medicine*, 517–527. Mosby Publishers, Baltimore, MD.

Slade JF, Mahoney JD, Dailinger JE and Baxamusa TH (1999). Wrist and hand injuries in musicians: management and prevention. *The Journal of Musculoskeletal Medicine*, **16**(9), 542.

Spahn C, Ell N and Seidenglanz K (2001). Psychosomatic findings in musician patients at a department of hand surgery. *Medical Problems of Performing Artists*, **16**(4), 144–151.

Sparrow WA and Newell KM (1998). Metabolic energy expenditure and the regulation of movement economy. *Psychonomic Bulletin and Review*, **5**(2), 173–196.

Spinner RJ and Amadio PC (2000). Compression neuropathies of the upper extremities. In R Tubiana and PC Amadio, eds, *Medical problems of the instrumentalist musician*, 273–294. Martin Dunitz Ltd, London.

Sternbach DJ (1995). Musicians: a neglected working population in crisis. In SL Sauter and LR Murphy, eds, *Organizational risk factors for job stress*, 283–302. American Psychological Association, Washington, DC.

Strahan E and Conger AJ (1998). Social anxiety and its effects on performance and perception. *Journal of Anxiety Disorders*, **12**, 293–305.

Tubiana R (2000). Functional anatomy. In R Tubiana and PC Amadio, eds, *Medical problems of the performing instrumentalist*, 1–4. Martin Dunitz Ltd, London.

Tubiana R, Chamagne P and Brockman R (1989). Fundamental positions for instrumental musicians. *Medical Problems of Performing Artists*, **4**(4), 73–76.

Turkington D, Dudley R, Warman DM and Beck AT (2006). Cognitive-behavioral therapy for schizophrenia: a review. *Focus*, **4**, 223–233.

Turner PE and Raglin JS (1991). Anxiety and performance in track and field athletes: a comparison of ZOFR and inverted-U hypothesis. *Medial Science in Sport and Exercise*, **23**, 119.

vande Wiel NMH, Matthys W, Cohen-Kettenis PT, Maassen GH, Lochman JE and van Engeland H (2007). The effectiveness of an experimental treatment when compared to care as usual depends on the type of care as usual. *Behavior Modification*, **31**(3), 298–312.

van Hees OS (1997). Physical exercise as prevention for musculoskeletal problems in musicians. A panacee demasque. *Health and the musician conference proceedings*, A1.20–A1.30. 23–27 March, York, England. British Performing Arts Medicine Trust, London.

van Kemenade JF, van Son M and van Heesch NC (1995). Performance anxiety among professional musicians in symphonic orchestras: a self-report study. *Psychological Reports*, **77**, 555–562.

van Winkel R, Myin-Germeys I, Delespaul P, Peuskens J, De Hert M and van Os J (2006). Premorbid IQ as a predictor for the course of IQ in first onset patients with schizophrenia: a 10-year follow-up study. *Schizophrenia Research*, **88**, 47–54.

Vergara, M and Page A (2002). Relationship between comfort and back posture and mobility in sitting-posture. *Applied Ergonomics*, **33**, 1–8.

Wallace S and Alden L (1997). Social phobia and positive social events: the price of success. *Journal of Abnormal Psychology*, **106**, 416–424.

Wang J (2002). Developing and testing an integrated model of choking in sport. Unpublished PhD dissertation, Victoria University, Melbourne, Australia.

Warrington J, Winspur I and Steinwede D (2002). Upper-extremity problems in musicians related to age. *Medical Problems of Performing Artists*, **17**(3), 131–140.

Wilkinson A (1992). Stretching the truth. A review of the literature on muscle stretching. *Australian Journal of Physiotherapy*, **38**, 283–287.

Willner P (2006). Readiness for cognitive therapy in people with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, **19**(1), 5–16.

Wilson FR (2000). Current controversies on the origin, diagnosis and management of focal dystonia. In R Tubiana and PC Amadio, eds, *Medical problems of the instrumentalist musician*, 311–327. Martin Dunitz Ltd, London.

Wilson GD (2002). *Psychology for performing artists*, 2nd edn. Whurr, London, UK.

Young JA and Pain MD (1999). The zone: evidence of a universal phenomenon for athletes across sports. *The Online Journal of Sport Psychology*, **1**(3). Retrieved 17 May 2006, from http://www.athleticinsight.com/Vol1Iss3/Empirical_Zone.htm.

Yovel I and Safren SA (2007). Measuring homework utility in psychotherapy: cognitive-behavioral therapy for adult attention-deficit hyperactivity disorder as an example. *Cognitive Therapy and Research*, **31**(3), 385–399.

Zaza C (1994). Research-based prevention for musicians. *Medical Problems of Performing Artists*, **9**, 3–6.

Zaza C and Farewell VT (1997). Musicians' playing-related musculoskeletal disorders: an examination of risk factors. *American Journal of Industrial Medicine*, **32**, 292–300.

Zaza C, Fleiszer MS, Main FW and Mechefske C (2000). Beating injury with a different drumstick: a pilot study. *Medical Problems of Performing Artists*, **15**(1), 39–44.

# PART 8

# Composition and improvisation

Edited by Peter Webster

# CHAPTER 37

# Making a mark

## The psychology of composition

Jonathan Impett

## Introduction

Given the central place of creativity as a defining component of the self-image of humanity, it is extraordinary how until recently it has received little attention from psychologists (Albert and Runco 1999). Studies have tended to consider the general case of the generation of new ideas or the process of invention within a particular constrained environment. While both biography and analysis can shed light on the process of composition, musicology has generally been more concerned with the context or fruits of the process. Artistic creativity poses considerable challenges to the experimental psychologist—in gaining access to the creative process itself other than through self-reporting, and in evaluating the quality, significance and uniqueness of work. The formulation of questions concerning an activity which is both cultural and personal naturally tends to reflect current modes of explanation.

There is little agreement on the cultural definition of composition, even within the context of contemporary Western culture. In conventional terms, it may be understood to be an activity bounded by the artefact of the musical work and by the persona of the composer. Discussion of the psychology of composition calls into question fundamental assumptions about the nature and role of musical activity and the musical artefact, and such assumptions must be acknowledged in any investigation. Taking Western masterworks as our sole evidence might be equivalent to understanding architecture only on the basis of the pyramids; objects of wonder but perhaps eccentric to any general case of human behaviour. The modern Western composition could be seen as the distillation of a social activity described by Blacking as 'humanly organized sound' (Blacking 1973). In an economic interpretation, Attali suggests that technology might return composition to its socially distributed state (Attali 1985, pp. 133–148). The cultural environment surrounding such works values them as exceptions, works of unique genius, expressions of individuality. As evidence from which to generalize they are thus to some extent self-invalidating. Biography and autobiography are likewise inseparable from context. The romantic, romanticized or modernist artist is inseparable from his work; Berlioz is an often-cited example. From late Bach to Boulez individual works stand as autobiographical statements, acts of musical self-exploring and self-situating. From at least the late sixteenth century, the advancement of the art has been a conscious obligation on most composers, and must therefore be understood as having substantially informed the relevant processes. Schoenberg's writings are perhaps the clearest expression of this awareness (Schoenberg 1975). In other cultural contexts, the role of the musician may be precisely to practice an art such that it may be transmitted without technical corruption or personal intervention. Guercio (2006) investigates the ways in which an awareness of public biography has fed back into artists' practice.

The concept of the work, the object of composition, has come increasingly into question, particularly in terms of its identity, autonomy, and the intentionalities it may be seen to embody. Goehr (1992) presents a case for the

historical contextualizing of the work concept. Davies (2001) proposes a measure of *thickness* of the work to reflect the degree to which its instantiation depends on culturally distributed knowledge. We can identify two characteristics of the activity of composition: a degree of identity between subject and object—composer and composed—and the structuring of music over a temporal scale beyond that of unstructured short-term memory or the doggerel patterning of simple song or dance forms.

## Motivation

> I believe that a real composer writes music for no other reason than that it pleases him. Those who compose because they want to please others, and have audiences in mind, are not artists.
>
> Schoenberg (1975, p. 54)

The question of the motivation to compose is likewise beyond the scope of experimental psychology. Where observer, observed and reader inhabit the same cultural milieu this is less problematic; in his view of the artwork as the manifestation of an artist's particular biography and psychological development, Freud is reinforcing the Romantic aesthetic of self-expression native to his own context. He concluded that 'whence it is that the artist derives his creative capacity is not a question for psychology' (Freud 2001a, p. 187). His observations on the relationship between artist, work and public have been more influential. He suggests that the artist is driven to the imagination by dissatisfaction with his reality and returns to reality via the work, which then serves the same function for others (Freud 2001b, p. 224).

Some measure of social empathy is at play, therefore. Heidegger extends the notion of phenomenological reality to one of truth, relating aesthetic success to the ethical behaviour of the artist in respect of his material (Heidegger 1971, pp. 15–86). Badiou's development of this view of artistic behaviour as ethical also in respect of society emphasizes the search for commonality over the assertion of difference, an approach that resonates with theories of creativity discussed below (Badiou 2001). The technical researches of both Schoenberg and Cage could be interpreted in this light; technique mediates between the banality of the imagination and the aspiration of critical reflection. Freud's insistence on the ludic nature of artistic creativity has been enthusiastically taken up by contemporary critical theorists. The concept of *play* is central to Gadamer's explanation of the ontology of art and of aesthetic experience (Gadamer 1989, pp. 102–171). Lacan's related notion of *jouissance* serves as the basis for his distinction between desire and drive. Following the work of Lacan and Foucault, contemporary critical theory is fundamentally informed by psychoanalysis. Its central topics of the relationships between subject and object, desire and drive, bear directly on the metaphysical and psychological relationship between composer and work and the urge to compose.

> Desire is grounded in its constitutive lack, while drive circulates around a hole, a gap in the order of things ... drive is quite literally *the very drive to break the 'all' of continuity in which we are embedded*, to introduce a radical imbalance into it.
>
> Žižek (2006, pp. 61–63)

## Psychology and composition

Given the primacy of composition in Western music, the silence of psychology in this respect is deafening. Sloboda (1985, p. 103) observed 'that composition is the least studied and least well understood of all musical processes, and that there is no substantial literature to review'. Brown still notes that 'Although there is growing interest in music cognition, little progress has been made in understanding composition per se' (Brown 2003, p. 2). As psychology has developed through the twentieth century, the authority of earlier psycho-biographical narrative explanations has been called into question. Meanwhile composition has turned consciously to self-reflection as other norms and constraints have dissolved. Already in 1937 Cage suggested that the future principle of the organization of musical sound would be 'man's common ability to think' (Cage 1961, p. 6). In the light of new technologies, Minsky has suggested more recently that art and psychology must inevitably converge (Minsky 1999). Musical expressionism is the stylistic embodiment of unmediated expression

in the romantic sense; the conscious embodiment of psychological states. Adorno saw Schönberg's *Erwartung* as heroically confronting the compositional challenge of the dissolution of tonality—by no coincidence a product of the same milieu as psychoanalysis and phenomenology. Subsequent approaches to post-tonal language included Hindemith's *Unterweisung*, based on perceived intervallic tension in melodic and harmonic space (Hindemith 1942, pp. 56–58). Few have attempted to work through the implications of music psychology to derive a set of aesthetic constraints. A notable exception is Lerdahl (1988), for whom only functional tonality satisfies the requirements. Implicit in Huron's theorizing of musical experience in terms of expectation is an understanding of composition as the prediction, management and thus presumably internal modelling of that expectation (Huron 2006, p. 240).

The incapacitating vertigo of the composer faced with infinite possibility is described by Ferneyhough (1995, p. 22). Sloboda suggests that a new dialogue between psychology and the composer can help in reducing this search space (Sloboda 2005, p. 192). The sonic and music–organizational possibilities of new technologies in the late twentieth century and the perception of composition as a potential research activity together informed the shape of institutions such as IRCAM in Paris, incorporating research into the psychology of music; Born (1995) gives an account of the co-evolution of institution and individual compositional practice. Writers such as Roads (1996) and Cook (2001) acknowledge psychoacoustics as an essential part of the electroacoustic composer's necessary knowledge. The close link between compositional technologies and research into psychoacoustics is a strong indicator of the extent to which conventional modes of music-making embody both innate understanding and cultural experience of music psychology. Technological approaches to composition and music-generation have been based on varying degrees of cognitive verisimilitude in the underlying mechanisms. Rowe (2001) presents a broad range of such research, largely developed in imitation of human compositional abilities. Cope's *Experiments in Musical Intelligence* (Cope 2001) generates new compositions within a given style based on the analysis and recombination of individual compositional *signatures*. Hofstadter points out that irrespective of similarity of psychological process, the inability of a culture to distinguish must lead us to question assumptions and understanding (Cope 2001, pp. 33–82).

## Theories of creativity

> Surely it's premature to ask how great composers write great symphonies before we know how ordinary people think of ordinary tunes. I don't believe there is much difference between ordinary and 'creative' thought. Right now, if asked which seems the more mysterious, I'd have to say the ordinary kind.
>
> Minsky (1985, p. 80)

Study of the general case of creativity begins in earnest as a reaction to the twin myths of romantic aesthetics and mid-twentieth-century information theory. Koestler (1964) was among the first to address the human practice of creativity, to demystify the process. Boden (1990) examines the paradoxes and conceptual difficulties in dealing with such a charged cultural concept. The earlier notion of genius as embodying the distance between conventional experience and that of the great artist—somewhat problematic in psychological terms, in that by definition it cannot be more widely understood—has been revisited. Simonton (1999) proposes an evolutionary explanation; that genius typically explores a wide range of potential behaviours, often in parallel, searching for complexity, novelty and emergent structure.

Most theories of creativity acknowledge that the everyday transmission and reception of ideas is in some sense creative. To varying degrees they incorporate the exploration of conceptual or physical spaces, the drawing of new connections between concepts or spaces, the satisfying of rules or constraints and the solving of problems. Holyoak and Thagard (1995) propose a model of creative thought around the drawing of analogy between different mental spaces within interacting constraints of similarity, structure and purpose. Boden (1990) describes the spontaneous reflexive process of the mind's generation of maps of its own conceptual spaces. In the absence of coherent bodies of data,

theories of creativity have been largely informed by the challenges of artificial intelligence. In Copycat, a computational model of analogy-seeking, Hofstader shows how one concept slips into another under different kinds of pressure or attractor, producing the mental fluidity fundamental to creative thought (Hofstadter 1995, pp. 205–267). Gärdenfors' notion of multidimensional *conceptual spaces* avoids the symbolic/connectionist division of computational models (Gärdenfors 2000).

Brown (2003) narrates the composition of Debussy's *Ibéria* as a process of knowledge-based problem-solving. He acknowledges the multidimensionality of the undertaking: 'music historians, music theorists, and cognitive scientists can join forces in trying to understand [how composers make their choices]' (Brown 2003, p. 162). Problem-solving and space-searching models generally rely on some form of heuristic to direct the process. This might be seen as where the craft of composition lies and why it requires such time to develop. Weisberg (1999) examines the relationship between creativity and knowledge through the development of expertise, citing particularly the '10-year rule' proposed by Hayes (1989) following examination of the creative trajectories of 76 composers. Explanations of this kind also assume some definable end-state. The model of composition proposed by Sloboda (1985, p. 118) takes account of the symbiotic evolution of work and work-definition. Brown's amended version incorporates historical context (Brown 2003, p. 9) (Figure 37.1).

Byrne (2005) shows that rational thought and counterfactual imagination are predicated on common cognitive mechanisms. She describes the probing of possible space by the imagining of alternatives to a given reality and how imagination works by exploring the parameters along which it might be otherwise. This seems a useful model for understanding the selecting of musical material according to its potential in context, as well as its subsequent treatment.

A mechanism of 'conceptual blending' is proposed by Fauconnier and Turner, to model the production of analogical links and inference transfers by mappings between mental spaces. 'Blending ... operates on two Input mental spaces to yield a third space, the *blend*. The blend *inherits partial structure* from the input

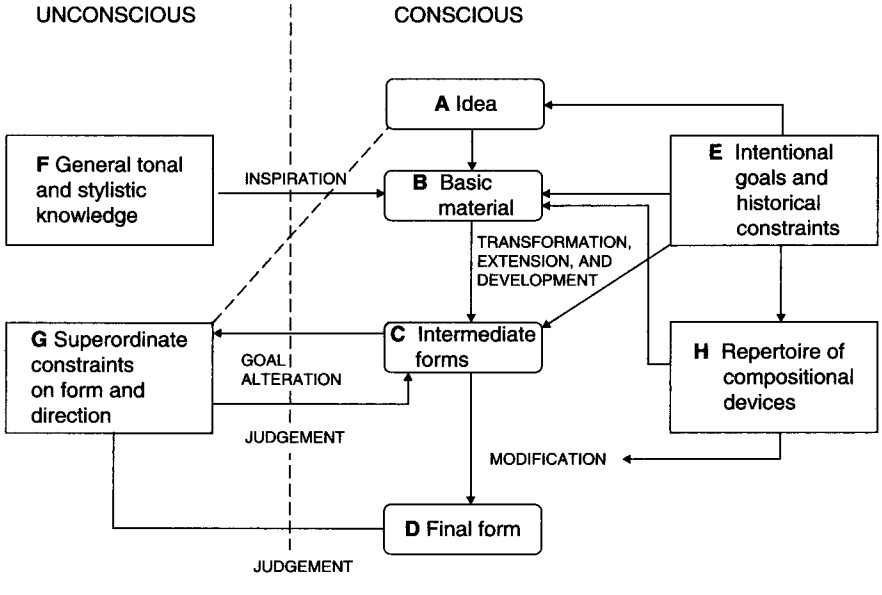



**Fig. 37.1** Diagram of typical compositional resources and processes. Brown (2003, p. 9); developed from Sloboda (1985, p. 118).

spaces and *has emergent structure* of its own' (Fauconnier 1997, p. 149). Blending through the multimodality, multiple spaces and scales and potential conflicts of human experience is achieved in *conceptual integration networks* (Fauconnier and Turner 2002, pp. 309–353). Turner describes the integration over time of a vast perceptual diversity, most clearly embodied in the activity of the artist, as *compression* (Turner 206, p. 94). Applying this approach to music, Zbikowski identifies three fundamental cognitive mechanisms: categorization, cross-domain mapping and the use of conceptual models (Zbikowski 2006, p. 116). Complex cultural and contextual factors can thus be incorporated to account for the emergence of meaning: 'Compositional strategies ... assume an alliance between syntax and processes of meaning construction: composers arrange musical materials with relatively specific expressive or communicative goals in mind' (Zbikowski 2002, p. 138). Specific musical detail derives from broader strategies through templates of categories and attribute values (Zbikowski 2002, pp. 149–153).

## Metaphors and models

Psychology and aesthetics are interlinked. The possibility of the composer having some responsibility to art, music, truth or any higher authority has material implications for the conceptual model within which composition takes place. The cosmology of Zarlino, Bach's theology or the Romantic aesthetics of Wagner in some respects have equivalent implications. In acknowledging the limits on understanding the creative process from sketches, Nottebohm adduces Goethe's metaphor of a plant, of unity in diversity, that runs through Romantic accounts of art and can be presumed to inform the self-understanding and conceptual models of composers from Beethoven to Webern (Webern 1963, p. 40). As mid-twentieth century composers questioned assumptions about musical language and aesthetics, they sought structuring metaphors in other areas of thought: natural phenomena (Xenakis), science fiction (Ligeti), semiotics (Berio), social theory (Cardew). Lachenmann describes his process explicitly as a sequence of metaphors: reflection, instrument-building and the freeing of imagination (Lachenman 2004, p. 45).

The material role of metaphor in thought has been shown by Lakoff and Johnson:

> Most of our normal conceptual system is metaphorically structured; that is, most concepts are partially understood in terms of other concepts ... The heart of metaphor is inference. Conceptual metaphor allows inferences in sensory-motor domains ... to be used to draw inferences about other domains (e.g., domains of subjective judgement.
>
> Lakoff and Johnson (2003, pp. 56, 244)

They demonstrate how abstract conceptual thought—which might include musical reasoning—is grounded in metaphors derived from human physicality and physical context (Lakoff and Johnson 1999). The reported synesthesia of Scriabin or Messiaen could be such metaphorical transfer; a case in which both spaces map onto physical experience. The range of personal, technical and cultural metaphor used by composers means that their self-reporting requires sensitive interpretation for use as evidence of psychological process. The generality of Bahle's study of 1947, based on the self-reporting of both living and past composers, illustrates the difficulty in drawing conclusions. He lists a range of intra- and extra-musical factors, composing environment, the urge to imitate or explore and artistic compulsion (Cooper 1990, pp. 14–15).

If the ability to model certain aspects of the physical and social world is to some extent innate (Pinker 1999, pp. 316–333), this can interact with cultural experience such that a composer might be able, for example, to model the performer's physicality and audience's impression of a particular violin phrase. The constraints of physical causality are effectively removed in working with new technologies (Lury 1998), and much contemporary technology-based creativity is concerned with the search for structural and conceptual metaphors to focus unconstrained imagination.

## Imagining

What we imagine does not necessarily take on a concrete form and may remain in a state of

> virtuality, whereas invention is not conceivable apart from its actual being worked out.
>
> Stravinsky (1942, p. 53)

Composition is an activity simultaneously of aural and aesthetic imagining, in some kind of feedback relationship. Both physical and artistic percepts are generated as hypotheses. The relative roles of emotional and intellectual elements in the latter are a proper topic for aesthetics, although in brain-science terms emotion and subjectivity are both real and rational (Damasio 1999; Thagard 2006). The compositional behaviour of a particular individual may be characterized by the scope, precision, innovation and relationship of aural and aesthetic imagination.

In many art forms, modes of scaling or reduction retain sufficient salient features of the whole such that the reduced form can be meaningfully understood or manipulated. A sketch for a painting may already display essential stylistic or representational properties while affording compositional reflection. A literary narrative can be outlined in the same language as its full version. Music does not afford the same reduction of scale. There is little invariance; material, surface and 'meaning' are lost. Some established schema—sonata form, for example—might allow the composer to concentrate on appropriate material, development and deviations from norms. Sloboda suggests that the modular process of working simultaneously on different levels of detail is akin to that of architecture (Sloboda 2005, pp. 194–195). Ligeti describes the repeated process of such formal macro planning in his Violin Concerto (Steinitz 2003, p. 335); Berg follows a similar procedure (Hall 1996, p. 21). The composition becomes an autonomous artefact as the composer discovers its nature from multiple perspectives.

In an often-quoted (and doubted) passage of reflection on his own process, Mozart describes imagining a composition as a simultaneity. The act of consciously contemplating a unitary concept takes time. Notation is itself a process of reflection and decision-making, and combined with the time of imagining is unlikely to happen in real time or less. The variables at play here seem to be the degrees of consciousness and assimilation of the processes involved. Following examination of Mozart's sketches, Konrad has concluded that he experimented at the piano and on paper, and that 'he prepared himself consciously for a particular compositional task and sought the working environment necessary for its fulfilment' (Konrad 1992, p. 77).

The relationship between imagined time, musical time as the object of reflection and the experiential time of musical performance is therefore critical to the process of composition; it acts on the temporal plane of memory. Berio describes the process of emergence from the interplay of imagining and memory of 'forms which stir and question memory but at the same time deny it' (Berio 2006, p. 41). A memory-based approach to contemporary composition is proposed by Snyder (2000). Kramer suggests that musical time works not only in a uniform speed and direction, but also non-linearly, retrospectively and in multiple layers (Kramer 1988). Some such principle must be fundamental to composition. Its precise nature is a highly individual question, and closely related to a particular composer's modes of experimentation and representation.

## Improvisation

Sloboda (1985, pp. 138–150) suggests that studying improvisation might provide insight into composition. Cultural understandings of improvisation vary; absence of fully prescriptive notation and a degree of ephemerality of the phenomenon are common features. Nettl and Melinda bring ethnomusicological perspectives to bear on 'music created in the course of performance' to investigate the musicians' own understanding of their behaviour (Nettl with Melinda 1998, p. 1). Isolating difficult-to-translate metaphorical descriptions can be misleading; many are technique- or theory-specific. Two have particular resonance, however. Sabah Fahkri describes the development of an ecstatic state through a process of feedback: 'I consider the audience to be me and myself to be the audience' (Racy 1998, p. 96). Jnan Prakash Ghosh describes his role in Hindustani music as 'keeping it going' (Slawek 1998, p. 336), implying an acknowledgment of some autonomy of the musical phenomenon and abstraction of its possibility. In contrast, Boulez dismisses improvisation as irrelevant to composition 'instant

imagination misfires more often than it fires; and anyway such freedom has less to do with invention in the proper sense than with the practicalities of invention' (Boulez 1991, p. 32).

Jazz and its successors provide instructive examples of longer-term musical design emerging through the practice of improvisation. Armstrong's *West End Blues* presumably grew out of his own improvisational style, from a continuous, repetitive practice, and having emerged in 1927 remains quite stable through subsequent versions. In Coltrane's solos on *Giant Steps* or *A Love Supreme* we hear the fruit, however spontaneous, of sustained intellectual and technical engagement with his material. Cecil Taylor prefers to communicate his more abstract structures orally: 'That gives the lie to the idea that the only structured music possible is that music which is written. Which is a denial of the whole of human expression' (Spellman 2004, p. 71). The musico-cultural theory and graphical scores of Anthony Braxton demonstrate the role of musical intelligence and imagination in the feedback loop of reflective practice, and the inevitable generation of higher-level structures (Braxton 1985). The technical achievement and musical power of the limit-case work of guitarist Derek Bailey in insisting on an entirely 'non-idiomatic' mode of improvisation is evidence of the irrepressibility of the urge to compose, of the difficulty in suppressing memory, association and imaginative projection (Watson 2004, p. 212). Evan Parker describes his own work as composition through improvisation and points to the impossibility of creativity denying structure, memory and repetition (Parker 1992). As recording technologies evolved from being simply the documentation of a performance, a new situation developed in which compositional exploration may be audibly inscribed. Thus in the recorded activity of composing Miles Davis's *Jack Johnson* or Mick Jagger's *Sympathy for the Devil* we can follow the distribution of musical ideas, the absorption of the input and techniques of others and the pattern of technological intervention (Davis 2005; Godard 2006).

Passages designed to give the impression of successful improvisation are instructive. Presumably improvisation contributed to Beethoven's generation of ideas, for example. His cadenzas to his own piano concertos could be seen as composed illustrations of his own creative process. Consideration of the cognitive mechanisms of improvisation gives some indication of its relationship to composition. Surveying the skills and components of improvisation, Pressing (2001) proposes a feedback and feedforward problem-solving model based on the generation of and selection from arrays of elements having musical, acoustic and motoric features. Sarath points to time as the essential parameter of difference between composition and improvisation (Sarath 1996). He derives a concept of retensive–protensive temporality from Husserl to describe the relationship of the musical thought of the improviser with the continuously moving present.

# Traces

## Sketches

A composer's sketches provide a trace of his conscious thought processes. While early sketch studies sought to investigate specific genius rather than composition in general, the field has more recently developed methodologies that afford broader interpretation (Kerman 1982; Brown 2003). Deductions can be made about the chronology of composition and possibly the models and constraints they embody and the problems they address. One might observe the 'harried' nature of Beethoven's writing (Hall 1996, p. 9) or the considered recording of musical thoughts in Webern's sketchbooks. The distribution of ideas on the page can itself reveal patterns of thought (Hall 1996, p. 13); Sloboda notes how Stravinsky uses a single simple idea to generate a variety of material (Sloboda 1985, pp. 108–112). Less certain is any induction as to the precise mental processes of which they are a trace. In reconsidering the discipline in 1982, Kerman observed that 'one could go on to study musical creativity in general, at least in theory' (Kerman 1982, p. 177). In an early case study, Nottebohm points out that Beethoven's sketches demonstrate both top-down and bottom-up approaches; sometimes a movement is repeatedly reformed in its entirety, sometimes a piece of thematic material is worked in different directions. Beethoven's comment that he often works on

several compositions simultaneously suggests a mutually reinforcing feedback between projects (Nottebohm 1979, p. 5).

## Technology

Modes of representation and the manipulations they afford are essentially a question of technology. So too are issues of dissemination and likely audience; notation, printing and the Internet transform the composer's concepts of self and agency. Banchieri's *Contrapunto alla Mente* (1608) is a parody of the limits of rule-based contrapuntal composition without the use of notation. Luzzaschi and Monteverdi describe the process (Owens 1997, pp. 65–66). The use of the wax tablet to work out points of polyphonic imitation must inform the grain of thought—the relationship between local contrapuntal texture and broader architecture. Changes in paper production through the eighteenth century meant that while we effectively have no sketches by Bach, whose supply of paper had to be approved and accounted for, Beethoven could presume a limitless supply. He could therefore experiment and store freely, reusing, reworking and re-mediating over his whole life—a random access memory allowing the process of composition to be highly distributed through time and contexts. In 1025, Guido d'Arezzo presented his new form of notation as addressing questions of memory. Symbol manipulation lies at the heart of Western composition. The symbolic representation and manipulation afforded by technologies from wax tablet to computer constitute a form of conceptual prosthesis. Virtual, mental quasi-external representations are both more ephemeral and more plastic than their material counterparts. The points at which a current state is externalized constitutes a unique signature in the compositional process. Instruments themselves constitute music computers; the piano is the ideal small-scale calculator for tonal music, with its octave equivalence, registral mapping, unique shapes for keys and scales and consistent but discrete relationships between them.

The multiple and hybrid representational environments of computer-enabled composition generate less transparent traces. Ferneyhough sees the computer as enhancing creative liberty, affording a more immediate and intuitive understanding of a complex compositional situation. 'It can generate a "musical archaeology", accessing and linking vastly different types of material and degrees of evolution from moment to moment' (Ferneyhough 1998). Lyotard describes the conventional modes of memory in the creation of art as *breaching* and *scanning*, to which he adds *anamnesis*—the remembering of the previously unknowable—in the case of technological creativity (Lyotard 1991, p. 48). In his study of digital craft, McCullough (1996, p. 99) describes computational environments as 'dense notational contexts for action'. Technology-informed changes in aesthetic stance also condition the nature of the traces remaining. For some of the late works of Nono, for example, there are copious sketches but no authoritative score. Compositional methodologies as set out in various historical manuals may reflect an understanding of composition which informs any conforming behaviour. They are not identical with an individual's psychological process, however.

## Conclusion

Composition is a reflexive, iterative process of inscription. The work, once named as such and externalizable to some degree, passes circularly between inner and outer states. It passes through internal and external representations—mostly partial or compressed, some projected in mental rather than physical space, not all necessary conscious or observable—and phenomenological experience real or imagined. At each state-change the work is re-mediated by the composer, whose decision-making process is conditioned by the full complexity of human experience. This entire activity informs the simultaneous development of the composer's understanding of the particular work in its autonomy, of their own creativity and of music more broadly. Environment (culture, technology) and agents (composer, work) co-evolve at different rates. Important areas of future research include the perception of form, the imagining of music 'off-line' and the process of identifying novelty and potential in that imagining.

Donald (2006) proposes six cognitive principles to art, all fundamental to the activity of composition. They may be paraphrased thus:

1 Composition should be regarded as a specific kind of cognitive engineering, intended to influence the minds of an audience.
2 Composition always takes place in the context of distributed cognition.
3 Composition is constructivist in nature, aimed at the deliberate refinement and elaboration of mental models and world views.
4 Composition is meta-cognitive in nature.
5 Composition is a technology-driven aspect of cognition.
6 The role of the composer, viewed as a component in a distributed cognitive system, is not necessarily fixed.

While the urge to compose—to invent, structure and define sound and musical behaviour—may be to some degree innate, modes of conceiving, representing and realizing are the product of a situated process. Even if some or all of that activity is so well assimilated personally or culturally that it remains hidden from experimental view, it remains a behaviour in respect of an emerging object. That object is an artefact embodying the composer's own simulation of its situated performance. Stravinsky describes the process:

> The very act of putting my work on paper, of, as we say, kneading the dough, is for me inseparable from the pleasure of creation. So far as I am concerned, I cannot separate the spiritual effort from the psychological and physical effort; they confront me on the same level and do not present a hierarchy.
>
> Stravinsky (1942, p. 51)

## References


Albert RS and Runco MA (1999). A history of research on creativity. In RJ Sternberg, ed., *Handbook of creativity*, 16–31. Cambridge University Press, Cambridge.

Attali J (1985). *Noise: the political economy of music*, trans B. Massumi. University of Minnesota Press, Minneapolis, MN.

Badiou A (2001). *Ethics: an essay on the understanding of evil*, trans. P. Hallward. Verso, London.

Berio L (2006). *Remembering the future*. Harvard University Press, Cambridge, MA.

Blacking J (1973). *How musical is man?* University of Washington Press, Seattle, WA.

Boden M (1990). *The creative mind: myths and mechanisms*. Weidenfeld and Nicholson, London.

Boulez P (1991). *Stocktakings from an apprenticeship*, trans. S. Walsh. Clarendon Press, Oxford.

Born G (1995). *Rationalising culture*. University of California Press, Berkeley, CA.

Braxton A (1985). *Tri-axium writings*, 3 vols. Synthesis Music, Oakland, CA.

Brown M (2003). *Debussy's Ibéria*. Oxford University Press, Oxford.

Byrne, R (2005). *The rational imagination: how people create alternatives to reality*. MIT Press, Cambridge MA.

Cage J (1961). The future of music: credo. In *Silence*, 3–6. Wesleyan University Press, Hanover, NH.

Cook P (2001). *Music, cognition and computerized sound: an introduction to psychoacoustics*. MIT Press, Cambridge, MA.

Cooper B (1990). *Beethoven and the creative process*. Clarendon Press, Oxford.

Cope D (2001). *Virtual music: computer synthesis of musical style*. MIT Press, Cambridge, MA.

Damasio A (1999). *The feeling of what happens: body and emotion in the making of consciousness*. Harcourt Brace, New York.

Davies S (2001). *Musical works and performances*. Oxford University Press, Oxford.

Davis M (2005). *The complete Jack Johnson sessions* (CD). Sony Records.

Donald M (2006). Art and cognitive evolution. In M Turner, ed., *The artful mind: cognitive science and the riddle of human creativity*, 3–20. Oxford University Press, New York.

Fauconnier G (1997). *Mappings in thought and language*. Cambridge University Press, Cambridge.

Fauconnier G and Turner M (2002). *The way we think: conceptual blending and the mind's hidden complexities*. Basic Books, New York.

Ferneyhough B (1995). *Collected writings*, In J. Boros and R. Toop (eds). Harwood Academic Publishers, Amsterdam.

Ferneyhough B (1998). Talk on Adorno given at Goldsmiths College, London, 21 February 1998. http://www.entretemps.asso.fr/Adorno/Informel/Ferneyhough.htm

Freud S (2001a). *The complete psychological works*, vol. 13. Penguin, London.

Freud S (2001b). *The complete psychological works*, vol. 12. Penguin, London.

Gadamer HG (1989). *Truth and method*, 2nd edn. Continuum, New York.

Gärdenfors P (2000). *Conceptual spaces: the geometry of thought*. MIT Press, Cambridge MA.

Godard JL (2006). *Sympathy for the devil* (DVD). Fremantle Home Entertainment.

Goehr L (1992). *The imaginary museum of musical works*. Oxford University Press, Oxford.

Guercio G (2006). *Art as existence: the artist's monograph and its project*. MIT Press, Cambridge, MA.

Hall P (1996). *A view of Berg's Lulu through the autograph sources*. University of California Press, Berkeley, CA.
Hayes JR (1989). Cognitive processes in creativity. In JA Glover, RR Ronning and CR Reynolds, eds, *Handbook of creativity*, 135–146. New York: Plenum.
Heidegger M (1971). *Poetry, language, thought*, trans. A. Hofstadter. Harper and Row, New York.
Hindemith P (1942). *The craft of music composition*, trans. A. Mendel. Schott, London.
Hofstadter D (1995). *Fluid concepts and creative analogies*. Basic Books, New York.
Holyoak KJ and Thagard P (1995). *Mental leaps: analogy in creative thought*. MIT Press, Cambridge, MA.
Huron D (2006). *Sweet anticipation: music and the psychology of expectation*. MIT Press, Cambridge, MA.
Kerman J (1982). Sketch studies. *19th-Century Music*, **6**(2), 174–180.
Koestler A (1964). *The act of creation*. Hutchinson, London.
Konrad U (1992). *Mozarts Schaffensweise. Studien zu den Werkautographen, Skizzen und Entwürfen*. Vandenhoeck and Ruprecht, Göttingen.
Kramer JD (1988). *The time of music*. Macmillan, New York.
Lachenmann H (2004). Philosophy of composition. In P Dejans, ed., *Identity and difference: essays on music, language and time*, 55–70. Leuven University Press, Leuven.
Lakoff G and Johnson M (1999). *Philosophy in the flesh: the embodied mind and its challenge to Western thought*. Basic Books, New York.
Lakoff G and Johnson M (2003). *Metaphors we live by*, revised edn. Chicago University Press, Chicago, IL.
Lerdahl F (1988). Tonal pitch space. *Music Perception*, **5**(3), 315–335.
Lury C (1998). *Prosthetic culture: photography, memory and identity*. Routledge, London.
Lyotard JF (1991). *The inhuman*, trans. G. Bennington and R. Bowlby. Polity Press, Cambridge.
McCullough M (1996). *Abstracting craft: the practiced digital hand*. MIT Press, Cambridge, MA.
Minsky M (1985). *The society of mind*. Simon and Schuster, New York.
Minsky M (1999). The future merging of art, science and psychology. In T Druckrey, ed., *Arts electronica: facing the future*, 229–233. MIT Press, Cambridge, MA.
Nettl B with Melinda R (eds) (1998). *In the course of performance*. University of Chicago Press, Chicago, IL.
Nottebohm G (1979). *Two Beethoven sketchbooks*, trans J. Katz. Gollancz, London.
Owens JA (1997). *Composers at work: the craft of musical composition 1450–1600*. Oxford University Press, New York.
Parker E (1992). Introduction to *'De Motu' for Buschi Niebergall*. http://www.efi.group.shef.ac.uk/fulltext/demotu.html.
Pinker S (1999). *How the mind works*. Penguin, London.
Pressing J (2001). Improvisation: methods and models. In J Sloboda, ed., *Generative processes in music:the psychology of performance, improvisation, and composition*, 129–178. Oxford University Press, Oxford.
Racy AJ (1998). Improvisation, ecstasy and performance dynamics in arabic music. In B Nettl, R Melinda, eds, *In the course of performance*, 95–112. University of Chicago Press, Chicago, IL.
Roads C (1996). *The computer music tutorial*. MIT Press, Cambridge, MA.
Rowe R (2001). *Machine musicianship*. MIT Press, Cambridge, MA.
Sarath E (1996). A new look at improvisation. *Journal of Music Theory*, **40**(1), 1–38.
Schoenberg A (1975). *Style and idea: selected writings*, L. Stein, (ed.). Faber, London.
Simonton DK (1999). *Origins of genius: Darwinian perspectives on creativity*. Oxford University Press, New York.
Slawek S (1998). Terms, practices and processes of improvisation in Hindustani instrumental music. In B. Nettl, ed., *In the course of performance*, 335–368. University of Chicago Press, Chicago, IL.
Sloboda J (1985). *The musical mind: the cognitive psychology of music*. Oxford University Press, Oxford.
Sloboda J (2005). *Exploring the musical mind*. Oxford University Press, Oxford.
Snyder B (2000). *Music and memory: an introduction*. MIT Press, Cambridge, MA.
Spellman AB (2004). *Four jazz lives*. University of Michigan Press, Ann Arbor, MI.
Steinitz R (2003). *György Ligeti: music of the imagination*. Faber, London.
Stravinsky I (1942). *Poetics of music*. Harvard University Press, Cambridge, MA.
Thagard P (2006). *Hot thought: mechanisms and applications of emotional cognition*. MIT Press, Cambridge, MA.
Turner M (2006). The art of compression. In M Turner, ed., *The artful mind: cognitive science and the riddle of human creativity*, 93–114. Oxford University Press, New York.
Watson B (2004). *Derek Bailey and the story of free improvisation*. Verso, London.
Webern A (1963). *The path to the new music*, trans. L. Black. Theodor Presser, Bryn Mawr, PA.
Weisberg RW (1999). Creativity and knowledge: a challenge to theories. In RJ Sternberg, ed., *Handbook of creativity*, 226–250. Cambridge University Press, Cambridge.
Zbikowski LM (2002). *Conceptualizing music: cognitive structure, theory and analysis*. Oxford University Press, New York.
Zbikowski LM (2006). The cognitive tango. In M Turner, ed., *The artful mind: cognitive science and the riddle of human creativity*, 115–132. Oxford University Press, New York.
Žižek C (2006). *The parallax view*. MIT Press, Cambridge MA.

# CHAPTER 38

# Musical improvisation

Richard Ashley

## Introduction

Musical improvisation is, to many in the Western world, an activity shrouded in mystery. Most listeners are familiar with some genres of music in which improvisation is a commonplace, such as rock and other popular styles, jazz, or perhaps 'ethnic' musics—that is to say, composed or improvised 'traditional' musics falling outside the typical Western canons. Therefore listeners are aware that many musicians can, and routinely do, produce novel musical utterances in real time. The question for most listeners is not 'Can music be improvised?' but rather 'How is improvisation carried out?' With this formulation of the question, musical improvisation becomes a suitable topic for psychological investigation, focusing on cognitive, physical, and interpersonal processes, and on the musical structures on which these processes operate. Viewed in this way, there are parallels that can be seen with regard to processes of verbal production. No one would think it unusual to find native speakers of a language producing new utterances on the spur of the moment; indeed, to find someone without such an ability is the unusual case (as pointed out in Blacking [1976]).

This chapter seeks to bring together the literature on musical improvisation which will be of interest and benefit to those wishing to know more about it from a cognitive perspective. There is currently a modest but growing literature dealing with the psychology of musical improvisation; the curious might compare the number of articles cited herein with those in other chapters of this volume. One major strand of the music–psychological literature will not be dealt with here, namely the use of improvisation in music-therapeutic, clinical, or psychotherapeutic settings. The interested reader is referred to the appropriate sections of this volume for references on those topics.

## Constraints on the processes of musical improvisation

Before moving on to a review of some dozens of studies dealing with improvisation, it is helpful to set a minimal theoretical framework for these discussions. The core notion here is that of *constraint*: of the scarcity of resources, in real time, that humans can utilize for making music when they are not following a predetermined plan of action (whether notated or not, as insightfully noted in Baily [1999]). To improvise appropriately in a given musical style—especially if the improviser is concerned with significant, productive interaction with other musicians, and with communicating with an audience—is a kind of tightrope act, where there is always the chance that something will go wrong (the recorded performances of master musicians show that even they make mistakes sometimes). Making music demands considerable resources from people; that is, no doubt, part of why music-making is enjoyable, in that it brings us into some more optimal relationship between our capabilities and our actions, along the lines of Csíkszentmihályi's notion of 'flow' (Csíkszentmihályi 1990, 1996). The use of one's abilities inside the constraints of one's body and its limits in performance, the timing of one's actions with external events, and retrieving and utilizing one's knowledge promptly in improvisation provide a powerful framework for a sense of personal achievement, and is thus a great

proving ground for Csíkszentmihályi's concepts of optimal experience and creativity.

What, more precisely, are significant constraints on musical improvisation? We shall consider three here, not closing the door to others. These are: the limits of the body (physical constraints); the limits of time (temporal constraints); the limits of knowledge (cognitive constraints). From these we shall then move into an examination of the psychological literature regarding improvisation.

## Constraint 1: the body

Improvisers work with their hands, feet, and voices to make music in the moment. Unlike a composer, who can consider and reconsider at length the different possibilities which a diverse ensemble might provide for the expression of a given musical idea, the improviser is left with only his or her own physical abilities to make the music when needed. The complex relationship of the human body's abilities to real-time music-making has rarely been explored (but see Eric Clarke's *oeuvre* for a stimulating, Gibsonian approach to the issue). In some ways, this can be seen as comparable to processes of speech production and articulatory phonology (Levelt 1989) in language, but is still more complex, in that these musical utterances must typically be articulated in the context of larger-scale structures, like discourse structure in language, and in tight temporal concert with other performers, unlike speech, where the norm is one speaker at a time.

## Constraint 2: real time

The proverb says, 'Time and tide wait for no man' ... and, we might add, time waits for no one improvising music. Sometimes the improvising musician is without any other collaborators (many blues recordings fall into this category), and in that case the improviser is free to alter the timing of their performance as they see fit, having no need to to coordinate with other performers. However, in many other situations, the improviser has other musicians with whom they are interacting in the moment and so very close timing constraints apply. These are often in the range of 1/25th of a second (40 msec or less)—see Iyer (2002, 2004)— and thus are of critical concern, since the margin for error is so slight. Why might this matter? because, as Stravinsky allegedly said, 'Music is the art that defines Man's relationship to Time.' Music provides the fine grid against which our ongoing lives are measured, more finely than with language or with clocks; it is humanity's great measure of time, down to the finest levels we can determine—and our auditory systems can detect the minutest levels of timing, well under 1/1000 of a second.

## Constraint 3: limits on what we know

Typically, people can only act on what they know; music is no exception to this rule. For the case of improvisation, the knowledge one uses should be encoded in procedural (know-how-to) rather than declarative (know-about) form. There are several reasons for this, mostly having to do with speed. Declarative knowledge is generally applicable, but slow: it must be interpreted in order to be used. Procedural knowledge, on the other hand, is typically viewed as recoded or 'compiled' into a tighter, more efficient form, from that initially given in its declarative form. Thus, improvisers must have their knowledge encoded in procedural form, to be used quickly enough, but must also make use of the implicit knowledge of musical structure possessed by listeners in order to make their in-the-moment compositions coherent and stylistically appropriate.

Thus, structure, pattern, and bodily engagement emerge as key elements of improvisational technique, and this has been reflected in psychological research into improvisation. This approach attempts to go beyond the limitations of interview and self-report; as Johnson-Laird writes:

> [musicians] can articulate only a limited answer, because the underlying mental processes are largely unconscious. If you ask yourself how you are able to speak a sequence of English sentences that make sense, then you will find that you are consciously aware of only the tip of the process.
>
> Johnson-Laird (2002, p. 417)

making plain why the discipline of psycholinguistics exists as well as the science of music cognition.

## Studies on improvisation in jazz

Much of the psychological literature dealing with musical improvisation focuses on jazz improvisation of some kind. The reasons for this are several: some of the researchers involved are themselves jazz musicians; improvisation lies at the heart of jazz, rather than being a more occasional garnishing of precomposed musical structures (as is the case, for example, with guitar solos in many rock songs); and jazz musicians display improvisational skill at a very high level. Since investigations of jazz are centrally placed in psychological studies, we turn now to them.

### Where do ideas come from in jazz?

One of the most common popular misconceptions of jazz improvisation is that it involves 'picking notes out of the air' or some similar statement. Jazz musicians themselves know that this is not the case, but have in general done little to explicate the mental and musical processes involved in improvisation. There are notable exceptions, such as Berliner's *Thinking in jazz* (1994), or Monson's *Saying something* (1996), deep ethnomusicological investigations into the workings of jazz improvisation. These studies, which are musicological rather than psychological in nature, provide a rich background for understanding some basics of jazz improvisation. It is a skill which takes thousands of hours of effortful practice to develop; it develops in community rather than with isolated individuals; it is intensely physical in real time; and it involves acquiring and becoming fluent with a kind of musical vocabulary or lexicon of patterns, which form the basis of musical expression the same way words or phrases do in spontaneous production of language.

### Improvisation, pitch, and tonality

To many listeners, the facility with which improvising musicians produce the melodies and harmonies they play is something almost magical. How can these notes be found so quickly—and so accurately? Clearly the tones are not selected at random or the results would not be so coherent and compelling. The answers to this question have been sought by researchers in a variety of ways.

The first, and quite well-documented, aspect of pitch organization by improvisers is their use of a set of pre-existing materials which serve as a vocabulary or lexicon from which they can select and on which transformations may be effected. This has long been noted in improvisation outside the Western tradition, for example in Balkan epic song performance (Lord 1960). In the early- to mid-twentieth century, two scholars—Milman Perry and Albert Lord—developed an idea of oral composition in Homeric poetry and in epic bards from the Balkans which has influenced much subsequent thought into the nature of improvisation. This work brought to the fore the notion of formulas as key elements of the oral composer's craft. These formulas can be thought of as recurring turns of phrase, such as 'the rosy fingers of dawn' in the Homeric epics. Researchers into jazz improvisation have adopted this notion using different terms, including 'licks', 'crips', and 'schemas' (Berliner 1994). Building a store of these patterns, which then can be reworked, varied and refined over time, is one of the primary ways in which young improvisers learn their craft and manage the time demands of improvisation.

However, there are deeper organizing principles at work, as well. There is solid evidence that jazz musicians' improvisations obey many of the principles of tonal organization found in the Western art-music tradition, as has been shown in a number of ways. Järvinen (1995) conducted an in-depth analysis of improvised saxophone solos by Charlie Parker. Järvinen's question was: in what ways do 'bebop' jazz improvisations' pitch content reflect the principles of tonal organization that have been found in other, composed repertoires (cf. Krumhansl 1990)? He found that the tonal profile—the prominence of scale degrees relative to the tonic pitch class—of the tones of Parker's solos fit quite well with Krumhansl's results This indicates that the broad statistical properties of tonality serve to guide the improviser at a deep level. This finding

is buttressed by other researchers, such as Larson (2002), whose work has long demonstrated that fundamental aspects of tonal organization as shown in melodic and harmonic teleology are shared between jazz and 'classical' music. Another Finnish researcher, Petri Toiviainen (1995), has used computational models of neural networks as a tool for exploring the tonal and rhythmic organization of improvisations, and has demonstrated ways in which the patterns and techniques, such as the improviser's use of centrally important 'target tones', can be simulated with these models, and also the subtle and powerful interplay between pitch and rhythm in improvisation (Järvinen and Toiviainen 2000).

## Improvisation, time, and rhythm

Rhythm in its different aspects is the lifeblood of improvisation (Pressing 2002; Iyer 2002). The improviser must be fully in command of rhythm. With regard to jazz, one recurring question is that of 'swing' or 'groove', at varying tempos, especially the extremely fast ones jazz musicians sometimes use. The literature shows (e.g. Friberg and Sundstrom 2002) that improvisers have a very fine control of timing and great accuracy, within the limits of what the body can do. Hence, at very fast tempi, while maintaining an impressively steady beat, the 'swing ratio' of unequal eighth-notes approaches a 1:1 ratio, perhaps because of physical limitations on moving hands and drumsticks.

However, not all improvisations take place under the kind of high real time demands that fast tempo jazz imposes. Ashley (2002) investigated the more small-scale improvisational aspects of master jazz musicians playing the melodies of ballad. These are slow, commonly played songs, and the variations in performance, which are enormously large relative to those found in the Classical tradition, should be considered acts of improvisation rather than simple re-creation and performance nuance. The primary findings of this study were related to performer's responses to musical structures. First, the performers observe the motivic patterning of the melodies, in that musical passages sharing the same basic figures were altered in similar ways, preserving the categorical nature of the melodies no matter how extreme the differences in timing might be from some original or nominal form of these motives. At a level below this, there were extensive deviations from the nominal rhythms of these materials, either as notated or as performed on prior recordings which served as the model for the improvisers. Above the level of the motive, at the rhythmic level of the phrase, there was a strong tendency for the soloists's expressive deviations to cease at cadence points, thereby anchoring the melody in the larger formal structure through reinforcing cadences. Expressive variation in timing was thus seen as operating on two levels: one structural, where motivic identity is and phrase rhythm is preserved, and one micro-expressive, where individual variation is valued. Microtiming variations are well-documented throughout the literature on jazz performance, as shown in Iyer (2002) and Collier and Collier (2002) from vastly different literatures and points of view.

# Improvisation and the body

Unlike musical composition, which can be mostly construed as a more-or-less purely cognitive undertaking, improvisation is an activity deeply steeped in and connected with the human body. To improvise means that one is producing music, audibly and in real time; this necessarily means either playing an instrument or singing. We approach this topic through the work of the psychologist-turned-ethnomusicologist John Baily, noting that imilar issues are raised, for example, by Sudnow (1993) in his narrative of learning to play the piano. Baily was trained as an experimental psychologist in the area of motor behaviour. When he turned to issues of music performance he found a particularly interesting topic in which to work: issues of performance on a kind of long-necked Afghani lute, the dutâr. This instrument underwent significant changes in its form over a relatively short span of years, and the techniques needed for playing it and the kinds of musical figures which were thereby idiomatic on it were necessarily altered. The study of these issues, and others related to them, led Baily (and his associate Blacking) to reconceptualize the relationship between the body and musical structure. In this

'motor grammar' approach, the sequences of sounds which are typically the object of musical study are not seen as abstract structures which performers then recreate, but are instead understood to be deeply informed and shaped by the physiological processes necessary for their performance (Baily and Blacking 1991). The patterns of sound are the result of motor behaviours, and thus bodily movement can be seen as the ultimate source of musical structures. Baily's work, both on the dutâr (Baily 1985) and on the guitar (Baily and Driver 1992), can be seen as having many parallels to articulatory phonology in linguistics, where the way in which language is shaped as a series of sounds is deeply rooted in the dynamic behaviour of the vocal tract (Browman and Goldstein 1989). It is interesting, and gratifying, to consider speech and music—the two main ways in which humans use sound for communication and for pleasure—as fundamentally connected to the embodied nature of human cognition.

## Computational and grammatical models of improvisation

Given the complexity of the issues which have been covered so far, it is no surprise that researchers might seek to tie them together in a more unified model. A trio of scientists serve as examples of the approach to understanding improvisation through compuational modelling: Johnson-Laird, Jeff Pressing and Mark Steedman. We consider each briefly in turn.

Philip Johnson-Laird, who is both a psychologist and an improvising musician, has dealt off and on for years with questions of creativity, touching on jazz improvisation more than once. In his 2002 article, he sets out a computational system which can produce some kinds of jazz-based improvisations (walking bass lines) given a chord progression. The program Johnson-Laird outlines attempts to deal with a primary psychological constraint facing jazz improvisers: limited short-term memory, operating in real time. His intriguing solution is to use the constraints occurring in musical structures as ways of limiting the set of options faced by a player at any given instant, so that the loads on memory and decision are lessened considerably. From the choices presented in the set of options yielded by the constraint, quite simple and efficient selection mechanisms suffice. This approach is quite different from those based on over-learned pitch-patterns or formulas, and the dialogue between these two viewpoints continues.

Jeff Pressing (1984, 1988, 1998, 2001), a polymath whose musical focus was improvisation, spent years developing a sophisticated, comprehensive approach to musical improvisation which has as the time of this writing no parallel. Its attempt to draw in physiological, cognitive, cultural, and structural dimensions of improvisation is impressive. In Pressing's model, structural and physical representations of improvisational technique are all represented, and work together to inform the musical result. The physical, embodied considerations in improvisation are fundamental to Pressing's thought: he notes 'the historical fact that music, of all art and sport forms, has developed improvisation to the highest degree', due he thinks to the high time resolution of the auditory and motor systems. Pressing's model of improvisation is based on notions of the interaction between the physicality of performance and the materials being performed, and deals in its own way with the notion of constraints (physical or music–structural) as important elements of a control structure for efficient, real-time improvisational behaviours. The model is explained in significant detail in Pressing (1988) and cannot be adequately described in brief terms, but basically generates novel 'utterances' through the production of musical figures ('event clusters') which may then either be continued based on one or more features of the events currently underway, or may be turned in new directions by a process of interruption. It is in part this duality of control structure, which enables both continuity and surprise, which is responsible for the system's quite impressive performance.

Mark Steedman's work has concerned itself with musical structure and its formal representation since his student days. His work, while not properly speaking a model of improvisation, has been very influential in understanding how improvisers might generate new versions of harmonic progressions from simpler underlying

prototypes. In Steedman (1984) he proposes a grammar of blues chord progressions which involves the use of typical transformation rules or rewrite rules, allowing a simple progression to be elaborated recursively into much more sophisticated, and yet stylistically appropriate, sequences. Since then he has continued to work on how such a grammar might be best constructed, and Steedman (1996) presents a solution which is more elegant in many ways but presents problems for the reader not already familiar with formal grammars. Steedman's work shows how one of the core issues in jazz improvisation, harmony, can be modelled efficiently for inclusion into a comprehensive model of improvisation.

## Improvisation, interaction, and identity

To be an improviser is, most often, to be a member of a community of musicians who share a common purpose, set of skills, and musical vocabulary. Different researchers have pursued this matter in a variety of ways. Some, like the musicologist Ingrid Monson (1996), have focused on the processes of interaction between jazz musicians while they are playing together, finding that the interplay between performers is essential to the nature of the art and to the skills and behaviours of individuals as well as groups. Others, like the communication theorist Keith Sawyer (1999, 2003, 2006) have connected musical improvisation with other collaborative activities, notably in the domain of theatre. Sawyer identifies three main elements of importance: improvisation, collaboration, and emergence. Improvisation, of course, refers to the production of new ideas 'on the spot' rather than recreating already composed ones; collaboration locates creativity not in any individual but in the group as a whole; and emergence, the notion that somehow the 'whole is greater than the sum of its parts,' and has dynamic properties which are hard to predict ahead of time. Seddon (2005) identified six modes of communication between student improvisers, both verbal and non-verbal, and proposes the notion of empathetic attunement and creativity as a key element in successful group improvisation. It is clear that group process is one of the areas in improvisation research which is ripe for additional study, and the growing field of dynamic systems and complex organization has much to offer in this regard.

Connected to the idea of group activity are the notions of community and the improviser's identity. This topic has produced a few interesting studies (e.g. Duranti and Burrell 2004) but has been most seriously investigated by a pair of researchers, Raymond MacDonald and Graeme Wilson (MacDonald and Wilson 2005, 2006; Wilson and MacDonald 2005). The focus of their research has been the ways in which membership in a community of improvisers serves to define, for the members, both important aspects of their individual 'identities' or self-understandings, and the very concepts they use in understanding what kind of music they see themselves as playing. MacDonald and Wilson make the case that group interviews, rather than interviews with individuals, are the best method for investigating these matters, as issues of identity and cognitive constructs such as the meaning of 'jazz' are in fact developed and shaped by the social group's interactions, which is most clearly seen in group discussions.

## Improvisation, children, and music learning

Since the publication of Hargreaves (1986) there has been a significant growth in the research literature surrounding developmental aspects of improvisation. We can only touch on the matter in this chapter but it is clear that improvisation can be a very significant aspect of childrens' musical development and an important avenue of creativity (cf. Webster 1987, 2002). A number of researchers have investigated ways in which children develop as improvisers. Paananen (2007) demonstrated the steady growth of hierarchic concepts in childrens' improvisations from the age of 6 to 11; these hierarchies were reflected in tonal structures as well as in rhythm, metre, and motivic organization. Brophy (2005) undertook a longitudinal study of children from the age of 7 through to the age of 9, recording their pentatonic improvisations on xylophones. His primary interest was in phrase and motivic

structure and metre; he found that there was increasing use of motive and antecedent-consequent phrase structure in the first two years of his sample, another sign of increasing sophistication at quite an early age. Kratus (1989) showed that children too young to be skillful in more structured composition (those up to the age of 7) could be productive as improvisors, suggesting that improvisation with younger children would be a good path for music learning. Other research (such as Azzara 1993) has demonstrated that using improvisation as a regular aspect of a fifth-year elementary school music learning curriculum had positive significant transfer to other musical activities and abilities.

Other researchers have looked at the influence of context on children's improvisations. Guilbault (2004) found that 5- and 6-year-olds are sensitive to the accompaniments played while they improvise; when a harmonically oriented accompaniment was provided, the children made more use of tonal function and implied harmony in their improvisations. Young (2003) dealt with 3- and 4-year-olds who were improvising together with an adult partner, and identified social interaction in the form of play to be an important source of children's creative behaviours, resulting in an enriched musical output. Finally, Burnard (1999, 2002) has shown that 12-year-olds engage in quite sophisticated group improvisational behaviours, involving interactions between their bodies, movements, and choices of instruments to facilitate improvisation, and valuing group collaboration and interaction in their music-making.

## Conclusions

The value of research into improvisation is clearly demonstrated by the studies mentioned here, and many others as well. Improvisation is not an isolated side issue with regard to human music-making; it connects musical structure, our bodies, and our sense of ourselves as individuals and as members of social units in powerful ways. Much work remains to be done but a strong foundation exists at this time.

## References


Ashley R (2002). Do[n't] change a hair for me: the art of jazz rubato. *Music Perception*, **19**(3), 311–332.

Azzara CD (1993). Audiation-based improvisation techniques and elementary instrumental students' music achievement. *Journal of Research in Music Education*, **41**(4), 328–342.

Baily J (1999) Ethnomusicological perspectives on Sawyer's ideas. *Psychology of Music*, **27**(2), 208–211.

Baily J and Blacking J (1991). Some cognitive aspects of motor planning in musical performance. *Psychologica Belgica*, **XXXI**(2), 147–162.

Baily J and Driver P (1992). Spatio-motor thinking in playing folk blues guitar. *World of Music*, **34**(3), 57–71.

Baily J (1985). Music structure and human movement. In P Howell, I Cross and R West, eds, *Musical structure and cognition*, 237–258. Academic Press, London.

Berliner P (1994). *Thinking in jazz: the infinite art of improvisation*. University of Chicago Press, Chicago, IL.

Blacking J (1976). *How musical is man*? University of Washington Press, Seattle, WA.

Brophy TS (2005). A longitudinal study of selected characteristics of children's melodic improvisations. *Journal of Research in Music Education*, **53**(2), 120–133.

Browman CP and Goldstein L (1989). Articulatory gestures as phonological units. *Phonology*, **6**, 201–251.

Burnard P (1999). Bodily intention in children's improvisation and composition. *Psychology of Music*, **27**(2), 159–174.

Burnard P (2002). Investigating children's meaning-making and the emergence of musical interaction in group improvisation. *British Journal of Music Education*, **19**(2), 157–172.

Collier GL and Collier JL (2002). A study of timing in two Louis Armstrong solos. *Music Perception*, **19**(3), 463–483.

Csíkszentmihályi M (1990). *Flow: the psychology of optimal experience*. Harper and Row, New York.

Csíkszentmihályi M (1996). *Creativity: flow and the psychology of discovery and invention*. Harper Perennial, New York.

Duranti A and Burrell K (2004). Jazz improvisation: a search for hidden harmonies and a unique self. *Ricerche di Psicologia*, **27**(3), 71–101.

Friberg A and Sundström A (2002). Swing ratios and ensemble timing in jazz performance: evidence for a common rhythmic pattern. *Music Perception*, **19**(3), 333–349.

Guilbault DM. (2004). The effect of harmonic accompaniment on the tonal achievement and tonal improvisations of children in kindergarten and first grade. *Journal of Research in Music Education*, **52**(1), 64–76.

Hargreaves DJ (1986). *The developmental psychology of music*. Cambridge University Press, New York.

Iyer V (2002). Embodied mind, situated cognition, and expressive microtiming in African-American music. *Music Perception*, **19**(3), 387–414.

Iyer V (2004). Improvisation, temporality and embodied experience. *Journal of Consciousness Studies*, **11**(3–4), 159–173.

Järvinen T (1995). Tonal hierarchies in jazz improvisation. *Music Perception*, **12**(4), 415–437.

Järvinen T and Toiviainen P (2000). The effect of meter on the use of tones in jazz improvisation. *Musicae Scientiae*, **IV**(1), 55–74.

Johnson-Laird PN (2002). How jazz musicians improvise. *Music Perception*, **19**(3), 415–442.

Kratus J (1989). A time analysis of the compositional processes used by children ages 7 to 11. *Journal of Research in Music Education*, **37**(1), 5–20.

Krumhansl C (1990). *Cognitive foundations of musical pitch*. Oxford University Press, Oxford and New York.

Larson S (2002). Musical forces, melodic expectation, and jazz melody. *Music Pereption*, **19**(3), 351–385.

Levelt WJM (1989). *Speaking: from intention to articulation*. MIT Press, Cambridge, MA.

Lord A (1960). *The singer of tales*. Harvard University Press, Cambridge, MA.

MacDonald R and Wilson G (2005). Musical identities of professional jazz musicians: a focus group investigation. *Psychology of Music*, **33**(4), 395–417.

MacDonald RAR and Wilson GB (2006). Constructions of jazz: how jazz musicians present their collaborative musical practice. *Musicae Scientiae*, **10**(1), 59–83.

Monson I (1996). *Saying something: jazz improvisation and interaction*. University of Chicago Press, Chicago, IL.

Paananen PA (2007) Melodic improvisation at the age of 6–11 years: development of pitch and rhythm. *Musicae Scientiae*, **11**(1), 89–119.

Pressing J (1984). Cognitive processes in improvisation. In R Crozier and A Chapman, eds, *Cognitive processes in the perception of art*, 345–363. North Holland, Amsterdam.

Pressing J (1988). Improvisation: methods and models. In J Sloboda, ed., *Generative processes in music*, 129–178. Clarendon, Oxford.

Pressing J (1998). Psychological constaints on improvisational expertise and skill. In B Nettl, ed., *In the course of performance*, 47–67. University of Chicago Press, Chicago, IL.

Pressing J (2001). Improvisation: methods and models. In J Sloboda, ed., *Generative processes in music: the psychology of performance, improvisation, and composition*, 129–178. Clarendon Press/Oxford University Press, New York.

Pressing J (2002). Black Atlantic rhythm: its computational and transcultural foundations. *Music Perception*, **19**(3), 285–310.

Sawyer RK (1999). Improvised conversations: music, collaboration, and development. *Psychology of Music*, **27**(2), 192–205.

Sawyer RK (2003). *Group creativity: music, theater, collaboration*. Lawrence Erlbaum Associates Publishers, Mahwah, NJ.

Sawyer RK (2006). Group creativity: musical performance and collaboration. *Psychology of Music*, **34**(2), 148–165.

Seddon FA (2005). Modes of communication during jazz improvisation. *British Journal of Music Education*, **22**(1), 47–61.

Steedman MJ (1984). A generative grammar for jazz chord sequences. *Music Perception*, **2**, 52–77.

Steedman MJ (1996). The blues and the abstract truth: music and mental models. In A Garnham and J Oakhill, eds, *mental models in cognitive science*, 305–318. Erlbaum, Mahwah, NJ.

Sudnow H (1993). *Ways of the hand*. MIT Press, Cambridge, MA.

Toiviainen P (1995). Modeling the target-note technique of bebop-style jazz improvisation: an artificial neural network approach. *Music Perception*, **12**(4), 398–413.

Webster PR (1987). Conceptual bases for creative thinking in music. In JC Peery, IC Peery and TW Draper, eds, *Music and child development*, 158–174. Springer-Verlag, New York.

Webster PR (2002). Creative thinking in music: advancing a model. In T Sullivan and L Willingham, eds, *Creativity and music education*, 16–33. Canadian Music Educators Association, Edmonton.

Wilson GB and MacDonald RAR (2005) The meaning of the blues: musical identities in talk about jazz. *Qualitative Research in Psychology*, **2**(4), 341–363.

Young S (2003) The interpersonal dimension: a potential source of musical creativity for young children? *Musicae Scientiae, Special Issue*, 175–191.

CHAPTER 39

# Children as creative thinkers in music

## Focus on composition

Peter R. Webster

THE study of creative thinking in music involves a complex combination of cognitive and affective variables, often executed at the highest levels of human thinking and feeling. This is such a complicated set of long-term engagements (composition, repeated music listening, or decisions about previously composed music in performance) or 'in the moment' engagements (improvisation and one-time listening), that it becomes quickly apparent why this field has not attracted more music researchers and why many feel the topic is hopelessly impregnable.

There are many reasons for optimism, however, as we are starting a new century of scholarship. The changes in education and the role of music in formal learning demands that we address creative thinking as best we can. The serious study of creative thinking from a psychological perspective is helpful to review as well. This chapter is designed to offer some perspective on these matters and to place in context the growing research literature in music completed by musicians. While it is generally acknowledged that children's creative thinking in music occurs as part of many music experiences such as listening, performance, conducting, and improvising (Reimer 2003), the focus here is on composition.

## Education context

Since the beginning of the twentieth century in North America, music educators have seen themselves as largely teachers of music performance, especially in the secondary schools where music study becomes centred on the 'talented' few (Mark 1996, p. 11). Many school systems require music instruction in the elementary and middle school grades for the entire school population with the curriculum often centring on singing, movement activities, and some listening. Attempts to include a wider variety of music experiences, including composition and improvisation at all grade levels, are only now beginning, inspired in part by the National Standards (Consortium of National Arts Education 1994) and arguments from music education philosophers and framers of curriculum (Reimer 2003). The situation in other English-speaking countries such as Australia and England has been qualitatively different. For some time, the emphasis in curriculum has slanted more toward compositional experiences or at least a more balanced approach that has addressed composition together with performance.

The English-language research literature on creative thinking in music for children, therefore, has been influenced in large part by these teaching traditions. Music psychologists interested in children's generative behaviour and unfamiliar with developments in music teaching and learning will find it necessary to understand this context.

One particularly powerful view of learning that has influenced the research is constructivism (Kafai and Resnick 1996). Constructionist thinking is not new to educational theory, with roots that can be traced to Dewey, Piaget, and others. This line of thought has been given a fresh perspective in writings on school reform

(Gardner 1999). The basic goals of constructivism are to place emphasis on creativeness and to motivate learning through activity. Learning is seen as more effective when *situated in activity* rather than received passively. In other words, children learn best when actively involved in creating things and not when asked to only memorize facts for later recall. Such memorization is important of course and must not be eliminated; rather, such learning should be partnered with creative activities that allow students to demonstrate command through action. Each child should be allowed to *construct* their own understanding with the expert help of the teacher. At the heart of these ideas is a shift away from thinking about education as being centred solely in the mind of the teacher to more of a partnership between teacher and student, with the teacher as the major architect of learning. Project-centred learning is celebrated with students solving problems.[1]

## Bases in psychological literature

J. Paul Guilford's 1950 keynote address to the American Psychological Association (Guilford 1950) is often cited as the commencement of the modern-day study of creativity. In the address, he noted the lack of attention paid to divergent thinking[2] and called for more systematic study. His work would evolve into the formation of a factor-of-intellect model of human intelligence that celebrated the intersection of product, operations and content (Guilford 1967). His subsequent factor analytic studies brought attention in psychology to a multiple intelligence theory that was meaningful. The specific model proposed came under fire by the research community due to the problems inherent in factor analysis as an empirical methodology (Sternberg 1999), but the spirit and logic behind Guilford's work lives on in many guises.

Since that time, the growth of formal study of creativity has been slow, at least until most recently. There are two edited volumes published in the last few years that are excellent compilations on the subject of creativity. The first is by Finke *et al.* (1996). The book has 14 chapters aimed at offering a contemporary view of creativity in a cognition context. Topics such as insight, problem-solving, memory, and incubation are included, as well as an interesting opening chapter by the editors on cognitive processes in creative contexts. Attention is also paid to machine intelligence and on connectionism and neural nets.

A more recent collection of writings was published by Cambridge University Press, edited by Sternberg (1999). This volume is perhaps the most comprehensive and definitive, single volume in the field of general creativity and contains 22 chapters written by many notable scholars in the field today. The book is important because it documents the recent upturn in interest among cognitive psychologists in the study of this difficult topic. Sternberg (2006) has also recently completed a retrospective analysis of his own work on creative thinking and it is highly recommended.

There are now major journals in psychology devoted to the topic of creativeness, including the *Journal of Creative Behavior* and the more recent *Thinking Skills and Creativity*. Two books on adult creativity have been written—one by Gardner (1997) and the other by Csikszentmihalyi (1997)—and each applies recent work to the explanation of genius.

Despite these developments, creativity as a construct (or as a set of complicated constructs) has largely been avoided in modern psychology. Sternberg and Lubart (1999) offer six possible reasons:

1 Mystic and spiritual roots of this topic which tends to put off the more scientific community
2 Its pragmatic, commercial nature exploited by those who offer popular accounts of the creative thinking process which are not based on theory and research

[1] For an extensive treatment of constructivism and its role in music teaching and learning, see Wiggins *et al.* (2006) and the accompanying response by Webster.

[2] The concepts of 'divergent' and 'convergent' thinking are at the heart of much of my writing and thinking about creativeness. Divergent thinking is simply that kind of thinking for which the result has no single goal and a number of products may result—a kind of personal brainstorming. Convergent thinking is work that focuses on a final result. Creative work involves both kinds of thinking many times and in many complex ways.

3 Early work on the subject that was not theoretically or methodologically central to the field of psychology, and, as such, not respected
4 Problems with definition and criteria that 'put off' the researcher who is looking for easier and perhaps more conceptually understood topics for tenure and promotion committees to understand
5 Approaches that have viewed creativity as an extraordinary part of an ordinary thing so as not to really need separate study, and
6 Unidisciplinary approaches to creativity that have tended to view a part of creativity as the whole phenomenon, trivializing or marginalizing the study (Sternberg 1999, p. 4).

Those researchers in general psychology that have been brave or inspired enough to deconstruct the general creative experience through empirical study have taken the following approaches (Mayer 1999):[3]

- *Psychometric.* Assessment work aimed at the creation of tools to measure specific traits or evaluate overall creative ability (Guilford, Torrance, McKinnon)
- *Experimental.* Traditional empirical paradigms designed to seek cause and effect relationships (Sternberg and Davidson, Collins and Amabile)
- *Biographical.* Studies that use historical data to understand the creative process and creative thinking (Wallace and Gruber, Gardner, Simonton)
- *Psychodynamic.* Writings based on clinical evidence and philosophical/psychological speculation about creativeness (Freud, Kris, Kubie)
- *Biological.* Data derived from physiological data (Martindale and Hines, Hudspith)
- *Computational.* Conceptual work based on mathematical and computer-based models and simulations (Boden, Shank)
- *Contextual.* More qualitative work based on the realities of social context (Csikszentmihalyi)

Each approach has strengths and inherent weakness. Selected reading of studies in each approach is highly recommended to gain a sense of the contemporary scene in the general literature before reviewing specific studies in music and music composition with children. In drawing implications for music teaching and learning, this literature helps to bolster the aspects of enabling conditions (both personal and culturally based) and enabling skills (personal competence) that are so critical for creative thought. Much of this literature, too, underscores the vital importance of divergency of thought and imagination in context with more convergent thinking that often involves just plain hard work. The vital role of social context is also apparent from this general literature.

## Promising new research for music

When comparing the approaches in the general literature to music, the psychometric, experimental, and contextual approaches from Mayer's listing are noticeable. A more 'descriptive' approach is emerging in music which places emphasis on the content analyses of the creative music experiences themselves. Hickey's review (2002) of creativity research in music and other arts stressed the importance of what she called 'confluence studies' which bring together social/cultural contexts with historical forces, events and trends. These studies, which will be noted below, often focus on the musical decision-making processes and on more personal issues outside of the music itself such as confidence, peer-interaction, and self efficacy.

Over fifteen years ago, I published a review of the literature on creative thinking in music education (Webster 1992). I have continued to maintain an annotated bibliography that attempts to cover the field of music teaching and learning.[4] My organizational model for this

[3] A complete review of each approach is not possible here but see Mayer (1999) for a more detailed description of each approach and for references to the scholar's work whose names appear in this listing.

[4] For the most recent annotated bibliography, consult the following website: http://musicalcreativity.com/?cat=3.

literature includes studies organized in three major categories:

1 *Theoretical*, works based on philosophical or psychological arguments as well as review, standards and historical writings;
2 *Practical Application*, literature designed to inform praxis but not derived fundamentally from empirical evidence; and
3 *Empirical*, work from numerical or observational data.

This empirical category is the most complex, with studies that examine teaching strategies, assessment design, technology, relationship between variables in and outside music, and the actual creative experience. Work on collaboration and social context has gained a great deal of ground in the last few years. Work with technology and teaching strategies is growing quickly as well. The early literature model was based on less than 200 writings. A current review of this literature will reveal nearly three times this amount.

There are a number of trends that can be seen in the literature in the last 15 years. Here is a summary of the major developments noted in this new literature in terms of children's composition. Key references are noted to give the reader a start in exploring the body of work. A brief summary of this work follows.

- Adoption of the postmodern tendency to question the assumptions made by previous generations and to be concerned more completely with social context and more naturalistic settings and qualitative methodologies in studying data (Burnard 2006).
- Empowering children's voices about the composition experience and valuing the meaning that children describe (Barrett 2003; Burnard 2000; Stauffer 2003).
- Heighten interest in the young child's work with invented music notation and the child's discussion of the notation as a window to understanding knowledge (Gromko 1994).
- New approaches to assessment, including consensual techniques and peer assessment (Byrne *et al.* 2003; Hickey 2001; Hickey and Lipscomb 2006).
- Attention to the role of collaboration and group composition (Kaschub 1997; Odam 2000; Miller 2004; Faulkner 2003; Seddon 2006).
- Emergent thinking on the pedagogy of composition teaching (Hewitt 2002; Webster 2003; Stephens 2003; Berkley 2004; Strand 2006; Odena and Welch 2007).
- New speculation on and experimentation with the role of music technology (Hickey 1997; Nilsson and Folkestad 2005; Crow 2006; Ruthman 2007; Pachet 2006; Burnard 2007).
- Models of creative thinking in music from a variety of perspectives (Webster 2002; Burnard and Younker 2002; Hickey 2003; Wiggins 2003; Espeland 2003).

## Shift in research paradigm

Burnard (2006) presented an informative retrospective of research approaches as a prelude to a chapter on children's meaning-making as composers. It characterizes a shift from the more 'positivist, large-scale studies aiming to measure creativity in children's composition towards ethnographic, qualitative approaches and to research focusing on the actual site of operations and practice' (p. 111). She argued that, with the development of socially centred theories, the need to study children in more natural settings doing more natural activities is more desirable. She further argued that more attention must be paid to the children's views, perspectives and accounts of process and product. Burnard's accounting of this shift in paradigm is reinforced as one studies many of the most recent research studies on children's composition noted in this chapter; however, it should be noted that more theoretical and positivist approaches continue as researchers seek better answers to assessment, teaching practice, and the description of what children actually produce musically.

## Empowering children's voices

A good example of the literature that moves more toward Burnard's characterization is the work honouring the naturally expressed, child perspective on what it means to compose. Barrett (2003) has written eloquently about children's descriptions of composition and the function that it plays in the lives of children.

Of interest is less the musical content itself and more the way children internalize and deal with the request to compose music and the way they make meaning from the experience. It is reasoned that the more we understand this process, the more we understand learning. Burnard (2000) has explored this theme effectively in her examination of how children blend improvisation with composition in the classroom. Stauffer (2003) offered a perspective on identity and voice in her review of several years of compositional data among children using computers to compose music. The accent in her work is on observation of evolving musical thinking with minimal direction from the teacher/researcher.

## Invented notation

One window into the minds of children when dealing with music composition is invented notation. Musical drawings offer a means for non-verbal communication of cognitive processing and representation. Gromko (1994) and others such as Barrett (1997) have explored this approach. Typically, children without formal notation understanding are asked to 'notate' their composed work so that others might understand it. After the invented notation is created, children might be asked to tell the researcher what the notation represents. This often leads to a better understanding of what drives or organizes the music composition. Such work leads to a better understanding of how children think in sound and how this changes with age and experience.

## New approaches to assessment

Invented notation is one way to deconstruct the complexities of children's composition, but other approaches to assessment of both product and process have been explored. Byrne *et al.* (2003) have reported positive and significant correlations between levels of optimal experience of Csikszentmihalyi's 'flow' levels and quality of group compositions as measured by creativity ratings. Such work is very tentative to be sure and fraught with difficulties, but such experiments are worthy additions to the literature as a partner to more qualitative work.

Hickey (2001) and Hickey and Lipscomb (2006) have reported some success using the consensual assessment technique made famous by Amabile. This approach is based on a theory that creativeness is best measured by experts in the domain of a field using global approaches and that the articulation of specific criteria is less valid.

## Collaboration and group composition

One interesting facet of recent study of children's composition is the effectiveness of collaborative and group composition. Kaschub (1997) profiled the processes of two composition projects, one with six sections of sixth-grade general music students and one with a high school choir. In each case, a composer worked with the groups to create a process cooperatively. The article is rich with examples of how children gained a stronger understanding of music. Seddon (2006) reported success in using composition as a group activity with computer-mediated collaboration over a distance. Similarly positive experiences were reported by Faulkner (2003) in working with group composition involving 6–16-year-olds in a rural school in Iceland; however, his study also identified some causes for concern in terms of group work. Odam (2000) published a critical perspective on group composition as it was implemented in certain schools in England and Wales as part of obligatory Key Stage 3 curriculum; he noted that too many teachers used inappropriate methods and much time was wasted. Miller (2004) contributed an action research study that featured group composition in an elementary school. This study is noteworthy because it was based on a careful review of the literature and featured the story of a teacher who was employing compositional experiences for the first time.

## Pedagogy of composition

Of interest to those that teach or are planning to teach, this recent development in the literature deserves special attention. Examples include the work by Hewitt (2002) and Odena and Welch (2007) which demonstrated a direct link between success in teaching composition and past experience with music and compositional thinking. Implications for teacher-preparation programmes seem obvious. Berkley (2004) contributed data on approaching the teaching

of composition as a problem-solving activity that should be presented in a positive atmosphere of student autonomy. Stephens (2003) presented several techniques that have been shown to be effective in teaching and stressed the importance of melodic ideas, technique, musical structure, and personal voice. In writing about the judicious use of revision, I have tried to argue for a balance between complete freedom in teacher direction and dominate control (Webster 2003). An extensive survey study by Strand (2006) of music teachers in Indiana revealed no consistent definition of composition and no real consistency in activities that purported to be compositionally based. Much work remains to be done on the preparation of music teachers to lead successful teaching sequences in composition that is inspired by both the more qualitative and quantitative studies described here.

## Role of music technology

Advances in music technology have profoundly affected the research and practice of children's compositional engagement. We have already mentioned the work of Seddon and Stauffer in this regard. The case studies by Hickey (1997) of individuals working with computer technology are important models. The work by Nilsson and Folkestad (2005) is yet another example of a set of investigations of how computer-based technology can be used to study the creative process; noteworthy was their desire to study product and process as an organic whole and to do so with a focus on social/process issues together with music content. On the other hand, Crow (2006) provided a more critical perspective on the implementation of certain music technology and fault was found not only with the materials but the techniques that employed them. Ruthman (2006) documented the use of social computing software applications and online collaboration techniques which promise to make a major advancement in music teaching. Pachet (2006) provided a more pessimistic view of common hardware and software applications and favoured interactive, reflexive music systems. One such system, The Continuator-I, developed at the Sony Computer Science Laboratories in France, allows the user to perform gestures on a music keyboard with an answer returned in such a way that a musical dialogue can be established. Implications for this system in encouraging and studying creative thinking in music are great.

One of the most interesting contributions to the literature concerning the nexus of technology and creative thinking in music is from Burnard (2007). She provided a set of theoretical constructs for the consideration of technology as a pedagogic change agent.

## Models of creative thinking

Finally, model building related to children's creative thinking in music continues to flourish. My own model (Webster 2002) has been newly revised based on much of the literature cited here. The model argues for a central process of staged thinking with music materials that represent the interplay of divergent and convergent processes—all informed by enabling skills and conditions. The work by Burnard and Younker (2002) on the pathways that children follow in the creative process is supportive of this model and extends it in fascinating ways. Espeland (2003) has contributed a model for compositional process that features personal and compositional actions in the context of school. Hickey (2003) has offered a componential model based in part on the work of Amabile; it stresses the critical role of social and intrinsic motivation and the nature of instructional set. Inspired by her own extensive teaching and the recent socially based literature on learning, Wiggins (2003) has provided a powerful 'frame' for understanding creative process for individuals and groups. Embedded in the frame is a carefully conceived interplay between teacher and student that stresses a safe and rich environment and a shared understanding of culture, curriculum, and compositional problem.

# Conclusion

This chapter has taken a decidedly 'teaching and learning' approach in summarizing the many studies on creative thinking in music. This has been purposeful on my part because some of the best work in this field is done by researchers interested in the developing minds of children

and the best way to educate them in this rich and complex century. Chapters such as this are difficult to write as much good work is not mentioned. The hope is that this sampling of recent studies may convince the reader to delve deeper into this topic, design and execute research, and apply the ideas in practice.

## References


Barrett M (2007). Invented notations: a view of young children's musical thinking. *Research Studies in Music Education*, **8**, 2–14.

Barrett M (2003). Freedom and constraints: constructing musical worlds through the dialogue of compositions. In M Hickey, ed., *Why and how to teach music composition: a new horizon for music education*, 3–31. MENC, Reston, VA.

Berkley R (2004). Teaching composing as creative problem solving: conceptualising composing pedagogy. *British Journal of Music Education*, **21**(3), 239–264.

Burnard P (2000). Examining experiential differences between improvisation and composition in children's music-making. *British Journal of Music Education*, **17**(3), 227–245.

Burnard P (2006). Understanding children's meaning-making as composers. In I Deliège and G Wiggins eds, *Musical creativity*, 111–133. Psychology Press, New York.

Burnard P (2007). Reframing creativity and technology: promoting pedagogic change in music education. *Journal of Music, Technology and Education*, **1**(1), 37–55.

Burnard P and Younker B (2002). Mapping pathways: fostering creativity in composition. *Music Education Research*, **4**(2), 245–261.

Byrne C, MacDonald R and Carlton L (2003). Assessing creativity in musical compositions: flow as an assessment tool. *British Journal of Music Education*, **20**(3), 277–290.

Csikszentmihalyi M (1997). *Creativity: flow and the psychology of discovery and invention*. Harper Collins, New York.

Crow B (2006). Musical creativity and the new technology. *Music Education Research*, **8**(1), 121–130.

Espeland M (2003). The African drum: the compositional process as discourse and interaction in a school context. In M Hickey, ed., *Why and how to teach music composition: a new horizon for music education*, 167–192. MENC, Reston, VA.

Faulkner R (2003). Group composing: pupil perceptions from a social psychological study. *Music Education Research*, **5**(2), 101–124.

Finke R Ward T and Smith S (1996). *Creative cognition: theory, research, and applications*. MIT Press, Cambridge, MA.

Gardner H (1997). *Extraordinary minds: portraits of exceptional individuals and an examination of our extraordinariness*. Basic Books, New York.

Gardner H (1999). *The disciplined mind: what all students should understand*. Simon and Schuster, New York.

Gromko J (1994). Children's invented notations as measures of musical understanding. *Psychology of Music*, **22**(2), 136–147.

Guilford J (1950). Creativity. *American Psychologist*, **5**, 444–454.

Guilford J (1967). *The nature of human intelligence*. McGraw-Hill, New York.

Hewitt A (2002). A comparative analysis of process and product with specialist and generalist pre-service teachers involved in a group composition activity. *Music Education Research*, **4**(1), 25–36.

Hickey M (1997). The computer as a tool in creative music-making. *Research Studies in Music Education*, **8**(1), 56–70.

Hickey M (2001) An application of Amabile's consensual assessment technique for rating the creativity of children's musical compositions. *Journal of Research in Music Education*, **49**(3), 234–244.

Hickey M (2002). Creativity research in music, visual art, theatre, and dance. In R Colwell and C Richardson, eds, *The new handbook of research on music teaching and learning*, 398–415. Oxford University Press, New York.

Hickey M (2003). Creative thinking in the context of music composition. In M Hickey, ed., *Why and how to teach music composition: a new horizon for music education*, 31–53. MENC, Reston, VA.

Hickey M and Lipscomb S (2006). How different is good? How good is different? The assessment of children's creative musical thinking. In I Deliège and G Wiggins, eds, *Musical creativity*, 97–110. Psychology Press, New York.

Kaschub M, (1997). A comparison of two composer-guided large group composition projects. *Research Studies in Music Education*, **8**, 15–28.

Kafai Y and Resnick M (eds) (1996). *Constructionism in practice: designing, thinking, and learning in a digital world*. Lawrence Erlbaum, Mahwah, NJ.

Mark M (1996). *Contemporary music education*, 3rd edn. Schirmer Books, New York.

Mayer R (1999). Fifty years of creativity research. In R Sternberg, ed., *Handbook of creativity*, 449–460. Cambridge University Press, New York.

Miller B (2004). Designing compositional tasks for elementary music classrooms. *Research Studies in Music Education*, **22**, 59–71.

Consortium of National Arts Education, (1994). *National Standards for Arts Education*. MENC, Reston, VA.

Nilsson B and Folkestad G (2005). Children's practice of computer-based composition. *Music Education Research*, **7**(1), 21–38.

Odam G (2000). Teaching composing in secondary schools: the creative dream. *British Journal of Music Education*, **17**(2), 109–128.

Odena O and Welch G (2007). The influence of teachers' backgrounds on their perceptions of musical creativity: a qualitative study with secondary school music teachers. *Research Studies in Music Education*, **28**, 71–82.

Pachet F (2006). Enhancing individual creativity with interactive musical reflexive systems. In I Deliège and G Wiggins, eds, *Musical creativity*, 359–375. Psychology Press, New York.

Reimer B (2003). *A philosophy of music education: advancing the vision*, 3rd edn. Prentice Hall, Upper Saddle River, NJ.

Ruthman A (2007). Strategies for supporting music learning through online collaborative technologies. In J Finney and P Burnard, eds, *Music education with digital technology*, 131–141. Continuum, London.

Seddon F (2006). Collaborative computer-mediated music composition in cyberspace. *British Journal of Music Education*, **23**(3), 273–284.

Stauffer S (2003). Identity and voice in young composers. In M Hickey, ed., *Why and how to teach music composition: a new horizon for music education*, 91–112. MENC, Reston, VA.

Stephens J (2003). Imagination in education: strategies and models in the teaching and assessment of composition. In M Hickey, ed., *Why and how to teach music composition: a new horizon for music education*, 113–140. MENC, Reston, VA.

Sternberg R and Lubart T, (1999) The concept of creativity: prospects and paradigms. In R Sternberg, ed., *Handbook of creativity*, 3–15. Cambridge University Press, New York.

Sternberg R (1999). *Handbook of creativity*. Cambridge University Press, New York.

Sternberg R (2006). Creating a vision of creativity: the first 25 years. *Psychology of Aesthetics, Creativity, and the Arts*, **S**(1), 2–12.

Strand K (2006). Survey of Indiana music teachers on using composition in the classroom. *Journal of Research in Music Education*, **54**(6), 154–167.

Webster P (1992). Research on creative thinking in music: the assessment literature. In R Colwell, ed., *Handbook of research on music teaching and learning*, 266–279. Schirmer Books, New York.

Webster P (2002). Creative thinking in music: advancing a model. In T Sullivan and L Willingham, eds, *Creativity and music education*, 16–33. Canadian Music Educators' Association, Edmonton, AB.

Webster P (2003). What do you mean, 'Make my music different?' Encouraging revision and extension in children's music composition. In M Hickey, ed., *Why and how to teach music composition: a new horizon for music education*, 55–65. MENC, Reston, VA.

Wiggins J (2003). A frame for understanding children's compositional processes. (M Hickey, ed.), *Why and how to teach music composition: a new horizon for music education*, 141–167. MENC, Reston, VA.

Wiggins J, Blair D, Ruthman A and Shively J, (2006). Constructivism. In E Wing, ed., *Mountain lake reader—Spring 2006*, 82–93. Middle Tennessee State University, Murfreesboro, TN.

# PART 9

# The role of music in everyday life

Edited by John Sloboda

CHAPTER 40

# Choosing to hear music

## Motivation, process, and effect

John Sloboda, Alexandra Lamont and Alinka Greasley

## Introduction

A substantial amount of music listening in contemporary Western society is deliberately chosen. This chapter reviews what is known about the psychology of self-chosen exposure to musical performances of others (recorded or live).

The cultural contexts studied in this research are characterized by three broad trends. First is the growing prevalence of 'musical consumers' (i.e. people who do not participate in musical events as performers, or rarely do so). Second is the increasing availability of recorded music of all genres and periods (through methods of production, marketing and distribution—especially Internet-based). Finally, the increasing miniaturization, portability, and flexibility of music delivery systems (e.g. iPods) means that the choice to hear specific music can be exercised in more and more situations. This has led to fundamental changes in the ways in which people use music (Liebowitz 2004).

These trends present a paradox. Although musical engagement is increasingly passive, because it is not linked to performance skill or participation, music use has never provided more opportunities for active agency. The choice of what music to listen to, when, and how has proliferated.

We organize the research reviewed here by the functional niche that the music is chosen to be part of. Music often accompanies some non-musical activity, chosen to enhance that activity in some way by affecting a psychological state which impacts on desired outcomes. In these contexts the music may not be the primary focus of attention or concern—the focus is rather on its effects.

Six main niches appear in the literature. These are:

1 travel (e.g. driving a car, walking, using public transport);
2 physical work (everyday routines like washing, cleaning, cooking, and other forms of manual labour);
3 brain work (e.g. private study, reading, writing, and other forms of thinking);
4 body work (e.g. exercise, yoga, relaxation, pain management);
5 emotional work (e.g. mood management, reminiscence, presentation of identity); and
6 attendance at live music performance events as an audience member.

Within these niches, we also identify four recurring functions of self-chosen music use. *Distraction* is a way of engaging unallocated attention and reducing boredom. *Energizing* is a means of maintaining arousal and task attention. In *entrainment*, the task movements are timed to coincide with the rhythmic pulses of the music, giving the task or activity elements of a dance. Finally, *meaning enhancement* is where the music draws out and adds to the significance of the task or activity in some way. Evidence for these comes predominantly from qualitative and ethnographic research, where music listeners are asked to talk or write freely about their music uses (e.g. DeNora 2000; Sloboda 1999; Sloboda *et al.* 2001), but we also present experimental findings as appropriate.

# Travel

Most everyday travel is made alone (for instance, to work). Music is chosen to accompany travel more consistently than in almost any other setting. This is made possible through mobile sound delivery systems that can accompany the traveller. Around 90 per cent of travel episodes are accompanied by self-chosen music (North *et al.* 2004; Sloboda *et al.* 2001; Stutts *et al.* 2001).

## Driving a car

Driving is a highly skilled task, requiring constant vigilance, which—if not appropriately applied—can result in death or serious injury. Dibben and Williamson (2007) review evidence that in-car music can decrease driving performance through its distracting and masking qualities: music makes it more difficult to attend to the key visual and auditory signals that must be constantly monitored to avoid accidents (e.g. engine noise, external warning signals). An additional accident-related distractor is adjusting controls on in-car audio-systems, thus taking attention off the road (Stutts *et al.* 2001).

Music while driving generally enhances performance through its capacity to assist with achieving or maintaining appropriate levels of arousal and concentration. Long and boring drives (e.g. on rural highways) promote drowsiness. Use of a sound system has been found to reduce accidents in such situations (Cummings *et al.* 2001), although the drowsiness-reducing effect is most marked in the first 30 minutes of driving, and drivers believe that music reduces their drowsiness more than it actually does as measured by EEG (Reyner and Horne 1998).

Some studies have looked at the effects of music on driver behaviour, often in simulated driving environments (e.g. Beh and Hirst 1999; Brodsky 2002; Turner *et al.* 1996). The benefits of music on reaction times and avoidance of hazards are found to depend on the nature, speed, and intensity of the music. Whilst music enhances reaction times to signals in the centre of vision, high levels of music-induced or situation-induced arousal can narrow the field of attention, causing people to drive faster and decreasing their driving safety.

Are drivers aware of the potentially deleterious effects of the music they choose to drive to? A survey of 2473 UK drivers by Dibben and Williamson (2007) showed that most drivers (87 per cent) listen to the same music while driving as at home. When asked about the effect of music in the car, many (62 per cent) cited music's capacity to soothe them, making them calmer and more relaxed. Rather fewer (around 25 per cent) claimed that music helped them concentrate whilst driving. The safest drivers (as indicated by having retained their no-claims insurance discount) were most likely to prefer silence while driving.

In sum, drivers listen to music in their cars for primarily motivational and enjoyment reasons. They do not typically self-consciously design an auditory environment designed to minimize accidents, but rather listen to the music they like in general in the car.

## On public transport

Around 60 per cent of journeys on UK public transport involve listening to chosen music via portable sound systems (North *et al.* 2004). Selecting reasons for listening in this context, respondents noted that 85 per cent of music listening episodes 'helped to pass the time', and 73 per cent chose 'I enjoyed it'. However, only 21 per cent of music-listening episodes helped participants concentrate/think. This confirms that, in general, music listening on public transport is used to enhance (or distract from) a routine, even boring, low-demand experience.

One obvious consequence of private music listening via headphones is that the music isolates the listener from other people (Bull 2001). This choice may thus allow music to mask unwanted noise or conversation, or to function as a defence mechanism against the anxiety-provoking elements of being in close proximity with strangers.

# Physical work

Domestic chores (such as washing, cooking, cleaning, gardening) are the most prevalent everyday physical tasks. These simple and often-repeated tasks have low interest and attentional demands, and music is a very frequent accompaniment.

'Singing while you work' has been a core activity in cultures over time and space (e.g. Gioia 2006). Music is frequently described as an accompaniment to mundane (mainly solitary and domestic) everyday activities (Sloboda 1999), and North *et al.* (2004) found that self-chosen music accompanied housework on 90 per cent of occasions.

Qualitative ethnographic evidence suggests that all four functions of music use—distracting, energizing, entrainment, and meaning enhancement—arise in relation to this activity (DeNora 2000; Sloboda *et al.* 2001). The following example highlights how music energizes, entrains and adds meaning to the routine of cleaning (from Greasley and Lamont 2006):

> I've got a CD that I **always** clean my house to, which is Swing When You're Winning by Robbie Williams [laughs]... it just kinda makes it a bit easier, a spoonful of sugar I suppose helps the medicine go down, and it's just, it's lively as well, most of it is quite lively and if you're hoovering, it's yeah, you can dance around to it while you're doing it.
>
> (Female 26 years old)

The interviewee later noted that she was 'not a massive Robbie Williams fan'. This exemplifies the functionality of the chosen music. Individuals choose music in conjunction with concurrent tasks and activities to help them achieve goals (see also DeNora 2000; Hargreaves and North 1999; Sloboda 1992, 1999).

# Brain work

## Desk work in occupational settings

Technological advances mean that individual workers can increasingly make self-directed decisions about whether and how to listen to music. In an office environment, Lesiuk (2005) provided a selection of CDs for individual use via headphones over a 2-week period. Music was found to improve a number of work-related measures, including overall mood: those who listened to music for more time had a more positive mood (see also Oldham *et al.* 1995).

Haake and Dibben (2006) surveyed employees' personal use of music in a variety of office settings. They found a high degree of music listening at work from their volunteers (80 per cent reported listening at work, for an average of 36 per cent of total working time). Respondents tended to listen to music while doing routine solitary tasks, word processing, web-surfing and emailing. They tended not to listen during tasks involving interaction with colleagues. Considering the functions of music listening, respondents cited task-related functions (improving concentration and focus, or blocking out unwanted noise) and more general positive effects (reducing stress, contributing to well-being and improving the perceived quality of the working environment). Some respondents suggested that music stimulated positive social interaction between colleagues by providing a talking point. Finally, particularly for those undertaking low-demand tasks, music was a way of relieving boredom. In common with other studies (e.g. Sloboda 1999) the main disadvantages of music in the workplace related to the involuntary overhearing of music chosen by others. Another problem raised was the difficulty of attracting a colleague's attention when they were wearing headphones.

## Private study

Very often children undertake homework in a room where a TV or radio is playing. Pool and colleagues (2003) found that background music from TV or radio had no significant effect on homework performance or duration. In contrast, soap operas caused a significant decline in performance, probably as their explicit narrative and verbal content caused cognitive interference with the task. Beentjes and van der Voort (1997) found 64 per cent of Dutch children claimed to play pop music 'often' or 'always' while doing writing assignments for homework. Most children claimed that this benefited their performance, although these results can imply that if there must be background media, music is less disruptive than spoken media.

Music is frequently used by adults as an accompaniment to private study at home (as recently confirmed for university students by Kotsopoulou and Hallam [2006]). However, there appears to be no systematic published research on the use of music for adult private study.

## Body work

Music has the capacity to alter bodily processes, such as physiological states and behavioural movements (see Chapter 11 this volume) and coordination and motivational levels (Chapter 44 this volume), and is used to accompany activities such as exercise, relaxation, and pain management. Below we review studies exploring factors that influence people's choices of music to accompany such activities.

### Exercise, yoga and relaxation

The key issue in the choice of music to accompany physical activities such as exercise or relaxation is the appropriateness of the music for the activity, and the functions of energizing and entrainment are central. Most experimental research uses experimenter-selected music (e.g. North and Hargreaves 1996, 2000), although some include a limited choice within this; one recent study surveyed the preferred music of their population before selecting music for use in their experiment (Karageorghis *et al.* 2006). Results indicate that choice and preference is affected by the appropriateness of the arousal potential of the music for the activity. For example, people prefer listening to high arousal music during exercise and low arousal music during relaxation (North and Hargreaves 2000), and faster music is preferred for higher intensity exercise (Karageorghis *et al.* 2006). Studies with restricted music choices also show that the perceived appropriateness of the music for the exercise affects liking (North and Hargreaves 1996), and over the long term this might affect independent choices made in real-life settings.

Belcher and DeNora (1999; see also DeNora 2000) explored the function of entrainment in relation to music and exercise in some detail. Using ethnographical methods (including in-depth interviews, questionnaires, and observation), weekly aerobics sessions were studied over a year. They found that beats per minute was one important feature of music choice for different exercise settings. Successful entrainment was partly influenced by participants' perceptions of the music. Valued sessions had music which varied in rhythm, lyrics, melodies and orchestration, facilitating appropriate levels of energy, motivation and endurance for the aerobics routines. Despite evidence for a consistent relationship between features of music (especially tempo) and exercise settings, there has been virtually no research looking at wholly self-chosen music that accompanies body work activities like exercise and relaxation.

### Pain management

As music is held to have active and structuring properties on and for the body (DeNora 2000), it has been explored as a tool for pain management. Studies have examined the effects of music on pain perception and management in both clinical and experimental settings (Good 1996; Mitchell and MacDonald 2006; Mitchell *et al.* 2006). Beneficial effects of repeated music listening over a week have been found for chronic pain sufferers in terms of higher levels of perceived control over the pain and lower levels of reported pain, depression and disability (Siedliecki and Good 2006).

Experimental studies show that choosing from a range of experimenter-selected music is more beneficial than listening to entirely experimenter-chosen music (Siedliecki and Good 2006), but listening to one's own preferred music has the most significant effect on pain (Mitchell and MacDonald 2006; Mitchell *et al.* 2006). Survey research illustrates that chronic pain sufferers regularly listen to music (75 per cent of the sample in Mitchell *et al.* 2007) and are also aware of music's potential for attention diversion and relaxation. Music which is entirely under the listener's control can thus be a very effective tool in pain management for a number of different reasons, including attention diversion and perceived control over the situation (Mitchell *et al.* 2007).

## Emotional work

Having evaluated the use of music in relation to non-musical tasks and goals, we next consider some more intrinsic outcomes of music listening by focusing on the use of music to manage mood, self, and emotions.

### Reminiscence

A common function of self-chosen music listening is reminiscence (reminding oneself of past events), which is particularly common in older adults (Juslin and Laukka 2004). Sloboda (1999) found that adult participants who were asked to recall early memories involving music were able to report detailed and highly emotional memories despite the considerable time which had elapsed. Many adults use their musical autobiographies to reconstruct their life story (DeNora 2000; Greasley and Lamont 2006), illustrating that music is intertwined in the construction of autobiographical narrative.

### Mood management

Listeners also regularly report mood management as a common function of self-chosen music listening, as indicated from surveys (Laukka 2007; Sloboda 1999) and qualitative studies (DeNora 2000; Greasley and Lamont 2006; Hays and Minichiello 2005). Unsurprisingly, given that people choose to listen to music that they like and that makes them feel good, they predominantly experience positive emotions as a consequence of music listening (Juslin and Laukka 2004). Individuals have considerable awareness of how they can use music for personal functions such as mood management. Batt-Rawden and DeNora (2005) describe this practical knowledge as 'lay' therapeutic practice, which, they note, is often subsumed by other goals or tasks rather than an independent act of behaviour (see earlier sections).

There are some differences in this awareness between individuals. For example, Sloboda (1999) found that women's free responses relating to mood were more detailed and articulate, and North *et al.* (2000) found adolescent girls to be more concerned than boys with how music could aid their emotional needs. Greasley and Lamont (2006) show that less musically engaged listeners seem less aware of the range of emotional functions that music has, while more engaged listeners have this 'lay' practice at their fingertips. For example, one male respondent noted:

> There's a different mood attached to all of them, there's a different feeling attached to all of them, um, I know every single one of them inside out, so I know what I want and where it'll be and I always have a choice.

Research itself can also stimulate the process of self-reflection, making listeners more aware of the emotional potential of music listening in everyday settings. In an innovative participatory study, Batt-Rawden and DeNora (2005) explored links between music listening, well-being and health. Participants reported the process of listening, choosing, and responding to different CDs to be therapeutic in its own right, sometimes heightening awareness of how and what to listen to for purposes of self-care.

### Presentation or confirmation of social identity

As well as using music for reminiscence and emotional self-regulation, music can also be used in the presentation or confirmation of identity. Preferences for particular types of music may carry a message to others regarding a range of attitudes and values (Zillmann and Gan 1997; Rentfrow and Gosling 2006; North and Hargreaves 2007). Adolescents' music preferences in particular have been shown to reinforce both how they see themselves and how they are seen by others (Tarrant *et al.* 2001). Listeners are also aware of the messages they may send by expressing particular musical preferences. Finnäs (1989) found marked differences between the music preferences young people stated publicly and those that they privately endorsed. Young adults also believe music preferences reveal information about their personal qualities (Rentfrow and Gosling 2003). In subscribing to a music taste culture, people are often associating with a particular lifestyle and wider set of values (Russell 1997). (For a more detailed account of musical identity, see Chapter 43 this volume).

## Attending live events

Despite the technological advancements in availability of recordings and media portability mentioned at the outset, attending live music events is still popular. For example, the Arts Council of England reports that around 10 per cent of the

population attend classical concerts in a year, while 36–39 per cent attend any live musical event (cited in Riley and Laing 2006; see also DCMS 2006). However, despite the sustained popularity of attending live musical events, this facet of everyday musical experience has received very little research attention.

## Motivations for attending live musical events

Socio-economic factors have long been recognized as influential in determining patterns of arts attendance. While most people report enjoyment and seeing a specific performer or event as their primary motivations for attending arts events (DCMS 2006; Pitts 2005a), it seems that more people would attend classical concerts if tickets were free of charge (Knight Foundation 2002). Moreover, in the US, only 4 per cent of the population regularly attend live events out of 27 per cent 'potential' classical music consumers (Brown and Bare 2003). This suggests an economic barrier (absolute or perceived) to concert-going amongst the less affluent.

Economic approaches to culture provide complex analyses of how people choose to spend their leisure time. Linder (1970) suggested that economic growth leads to *less* time spent on culture, since the opportunity cost of attending a concert rises alongside wages. However, increased income leads those people who choose to engage in cultural activities to do so through relatively goods-intensive experiences such as CD buying rather than time-intensive experiences such as concert attending (Ekelund and Ritenour 1999). Bourdieu's (1984) sociological concept of cultural capital explains why certain social groups engage in particular kinds of cultural behaviour, relating patterns of cultural consumption to sociological variables such as class and gender. Considering music concert attendance, classical music concert audiences are typically professionals with higher incomes, while folk music attendees are typically middle- and lower-class (DiMaggio and Useem 1978; Pegg 1984).

These large-scale approaches (see also North and Hargreaves 2007) are useful in informing us of high-level social factors influencing patterns of engagement, and powerfully demonstrate that many people decide that classical music in particular is 'not for them'. However, they tell us little from a psychological perspective about motivations for attending live events or the experiences that audiences have at such events, and there has been very little research directly addressing these questions.

In her exploration of audience participation at a chamber music festival, Pitts (2005a) provided some insight into why, having chosen to attend the festival, audiences selected particular concerts within it. Reasons broadly related to musical preferences (for example, avoidance of vocal concerts), the performers, and the audience's familiarity or unfamiliarity with the music (for example, many deliberately chose concerts with some unfamiliar elements). Reasons for not selecting a concert related mainly to musical preference and familiarity, especially relating to over-familiarity with certain works. While Pitts' research emphasizes musical factors that influence participation and attendance at events, non-musical factors like the mood of the listener and the physical environment (e.g. traffic noise, uncomfortable seating) also affect listeners' predicted enjoyment of concerts (Thompson 2007). This might also influence audiences' choice to attend a particular event. The reasons for attending live events are thus complex and extend well beyond the content of the programme or the star performer.

## What kinds of experiences do concert-goers have?

The expressive and emotional effects that a performer or group can have on a single listener are now well understood (e.g. Juslin 2001; Davidson 2005), but much less is known about the experiences of a group of listeners at real musical events.

Some research has focused on the retrospective nature of the listener's experience at live events. Gabrielsson and Lindstrom-Wik (2003) found around 75 per cent of strong experiences related to music arose in listening rather than performing situations, and a similar proportion in live rather than recorded experiences. Free written descriptions show that these experiences are diverse in nature, encompassing physical reactions, perception, cognition, feeling and

emotion, existential and transcendental aspects and personal and social aspects. A recent UK study (Lamont, in preparation) indicated that 54 per cent of university students' strongest experiences of music were at pop or rock festivals and concerts and 27 per cent related to live classical events, while less than 20 per cent of the strong experiences resulted from listening to recorded music.

Less research has focused directly on the live performance situation as it occurs. Pitts (2005a) reported informal conversations with audience members at a chamber music festival, although her analysis focused mainly on retrospective accounts provided by audiences through written and interview methods. Using a more structured yet highly ecologically valid method, Thompson (2006) asked a concert audience to rate their experience, enjoyment and emotional response to the performance. Listeners were able to differentiate between the quality of the performance and their own affective responses. Moreover, their enjoyment of the event was better predicted by their own degree of emotional engagement than the perceived quality of the performance, although much variance was left unexplained. Thompson's research highlights some important questions for further research, including the importance of looking at both musical and non-musical factors (see also Thompson 2007).

### Is attending a live event different to listening to recorded music?

Little research has directly considered the relationship of live experiences in comparison with listening to recorded music. Gabrielsson and Lindstrom-Wik (2003) did not present analysis of the differences between live and recorded experiences, although Lamont's (in preparation) study highlights some striking differences:

> Listening to them [Radiohead] on CD is one thing, but when thousands of people surround you, singing to every word like you, the atmosphere electric, there's no other feeling as strong, or intense, as that.

Thus concert attending provides a greater sense of community than solitary listening (see also Bogucki *et al.* 2005), and for the popular music audience, concerts complement rather than duplicate the experience of listening to pre-recorded music. Similar responses were found from some classical music festival attendees who suggest that CDs complement the live listening situation (Pitts 2005a). However, other audience members reported avoiding recordings of pieces they had experienced live, noting that 'nothing can replicate what we have experienced this evening' (Pitts 2005b, p. 99). This selective approach to listening to recordings may be more common amongst classical music listeners, although more research is required.

In summary, the live performance situation and the audience's response to it are only beginning to be understood from a psychological perspective. Biographical retrospective data highlights that the live performance situation can have a substantial impact on individuals' lives, and research undertaken closer to the time of experiencing the event emphasizes the importance of emotional engagement with the experience. Future research will need to continue to explore this complex but highly significant aspect of music listening in everyday life.

## Conclusions

There is an idealized view of music listening which is implicit in much of the traditional music appreciation literature (see, for instance, Cook 1998). This places 'pure art music' as an object for the reverent attention of the ideal listener, silent, still, totally focused on the music and its aesthetic or spiritual value.

The psychological research presented here paints a much more varied picture, where the surrounding contexts of music not only intrude on the act of hearing, but in many cases shape and control the very purpose, nature, and effect of that hearing. Music may, indeed, be listened to with total silent attentiveness, but for many people this appears to be the exception rather than the norm. Instead, music is more often the chosen accompaniment to all sorts of activities which have no direct connection to music, and could mainly be accomplished without music being present at all. Far from requiring reverent attention, much music experience could be described as background—like the soundtrack of a film, never defining of the focus

of attention, yet psychologically powerful nonetheless.

Two areas of research are ripe for development. One is a fuller characterization of the factors which shift music experiences from the periphery to the centre. Under what conditions does the urge to hear and re-hear specific music become a dominant one, and what kind of changes in thought and feeling come about as one becomes very familiar with a piece of music, through choice? To put it another way, how does music move between 'background' and 'foreground' in a person's everyday life? A second, related question, is the extent to which individuals develop expertise in music choice, in the sense of selecting specific pieces of music with the reliably fulfilled intention of achieving certain psychological outcomes (just as taking an analgesic is done to reliably ease pain). Are some people more expert in their self-chosen administration of musical 'doses' than others, and what makes them more expert? Placing the study of self-chosen music listening within a theoretical framework that acknowledges the high degree of agency involved may well be an important next step in advancing understanding a phenomenon which, despite its ubiquity, still proves elusive to characterize.

## References


Batt-Rawden K and DeNora T (2005). Music and informal learning in everyday life. *Music Education Research*, **7**, 289–304.

Beentjes JW J and van der Voort THA (1997). The impact of background media on homework performance: students' perceptions and experimental findings. In P Winterhoff-Spurk and THA van der Voort, eds, *New horizons in media psychology: research co-operation and projects in Europe*, pp 175–189. Westdeutscher Verlag, Wiesbaden.

Beh HC and Hirst R (1999) Performance on driving-related tasks during music. Ergonomics, 42, 1087–1098.

Belcher S and DeNora T (1999). Music is part of the equipment that makes your body hard! *eXercise (Publication of the National Governing Body for Exercise and Fitness in England)*, **5**, 4–5.

Bogucki Duncan N and Fox MA (2005). Computer–aided music distribution: the future of selection, retrieval and transmission. *First Monday*, **10**(4), http://firstmonday.org/issues/issue10_4/duncan/index.html.

Bourdieu P (1984). Distinction: a social critique of the judgement of taste. Routledge, London.

Brodsky W (2002) The effects of music tempo on simulated driving performance and vehicular control. *Transportation Research Part F*, **4**, 219–241.

Brown AS and Bare J (2003). *Bridging the gap: orchestras and classical music listeners* (Issues Brief No. 2.) John S. and James L. Knight Foundation, Miami, FL.

Bull M (2001). The world according to sound: investigating the world of Walkman users. *New Media Society*, **3**, 179–197.

Cook N (1998). *Music: a very short introduction*. Oxford University Press, Oxford.

Cummings R, Kopesell TD, Moffat JM and Rivara FP (2001). Drowsiness, counter-measures to drowsiness, and the risk of a motor vehicle crash. *Injury Prevention*, **7**, 194–199.

Davidson JW (2005). Bodily communication in musical performance. In D Miell, RAR MacDonald and DJ Hargreaves, eds, *Musical communication*, 215–237. Oxford University Press, Oxford.

DeNora T (2000). *Music in everyday life*. Cambridge University Press, Cambridge.

DCMS (Department for Media, Culture and Sport) (2006). *Taking part: the national survey of culture, leisure and sport*. Statistical release: Department for Culture, Media and Sport. http://ww.culture.gov.uk/global/research/taking_part_survey/survey_outputs.htm.

Dibben N and Williamson VJ (2007). An exploratory survey of in-vehicle music listening. *Psychology of Music*, **35**(4), 571–589.

DiMaggio P and Useem M (1978). Social class and arts consumption: the origins and consequences of class differences in exposure to the arts in America. *Theory and Society*, **5**, 141–161.

Ekelund RB Jr and Ritenour S (1999). An exploration of the Beckerian theory of time costs: symphony concert demand. *The American Journal of Economics and Sociology*, http://www.findarticles.com/p/articles/mi_m0254/is_4_58/ai_58496764/pg_1.

Finnäs L (1989). A comparison between young people's privately and publicly expressed musical preferences. *Psychology of Music*, **17**, 132–145.

Gabrielsson A and Linström Wik S (2003). Strong experiences related to music: a descriptive system. *Musicae Scientiae*, **VII**(2), 157–217.

Gioia T (2006). *Work songs*. Duke University Press, Durham, NC.

Good M (1996). Effects of relaxation and music on post-operative pain: a review. *Journal of Advanced Nursing*, **24**, 905–914.

Greasley AE and Lamont A (2006). Music preference in adulthood: why do we like the music we do? In M Baroni, AR Addessi, R Caterina and M Costa, eds, *Proceedings of the 9th international conference on music perception and cognition*, 960–966. University of Bologna, Bologna, Italy.

Haake AB and Dibben NJ (2006). Music-listening practices in workplace settings in the UK. In M. Baroni, AR Addessi, R Caterina and M Costa, eds, *Proceedings of the 9th international conference on music perception and cognition*, p. 1094. University of Bologna, Bologna, Italy.

Hargreaves DJ and North AC (1999). The functions of music in everyday life: redefining the social in music psychology. *Psychology of Music*, **27**, 71–83.

Hays T and Minichiello V (2005). The meaning of music in the lives of older people: a qualitative study. *Psychology of Music*, **33**, 437–451.

Juslin PN (2001). Communicating emotion in music performance: a review and a theoretical framework. In PN Juslin, JA Sloboda, eds, *Music and emotion: theory and research*, 309–337. Oxford University Press, Oxford.

Juslin PN and Laukka P (2004). Expression, perception and induction of musical emotions: a review and a questionnaire study of everyday listening. *Journal of New Music Research*, **33**, 217–238.

Karageorghis CI Jones L and Low DC (2006). Relationship between exercise, heart rate and music tempo preference. *Research Quarterly for Exercise and Sport*, 77(2), 240–244.

Knight Foundation (2002). *Classical music consumer segmentation study: how Americans relate to classical music and their local orchestras*. John S. and James L. Knight Foundation, Miami, FL.

Kotsopoulou A and Hallam S (2006). Age differences in listening to music while studying. In M Baroni, AR Addessi, R Caterina and M Costa eds, *Proceedings of the 9th international conference on music perception and cognition*, University of Bologna, Bologna, Italy.

Lamont A (in preparation). University students' musical biographies and strong emotional experiences of music.

Laukka P (2007). Uses of music and psychological well-being among the elderly. *Journal of Happiness Studies*, **8**, 215–241.

Lesiuk T (2005). The effect of music listening on work performance. *Psychology of Music*, **33**(2), 173–191.

Liebowitz S (2004). Will MP3 downloads annihilate the record industry? The evidence so far. *Advances in the Study of Entrepreneurship, Innovation, and Economic Growth*, **V**(15), 229–260.

Linder S (1970). *The harried leisure class*. Columbia University Press, New York.

Mitchell LA and MacDonald RAR (2006). An experimental investigation of the effects of preferred and relaxing music listening on pain perception. *Journal of Music Therapy*, **XLIII**(4), 295–316.

Mitchell LA, MacDonald RAR, Knussen C and Serpell MG (2007). A survey investigation of the effects of music listening on chronic pain. *Psychology of Music*, **35**(1), 39–59.

Mitchell LA, MacDonald RAR and Brodie EE (2006). A comparison of the effects of preferred music, arithmetic and humour on cold pressor pain. *European Journal of Pain*, **10**, 343–351.

North AC and Hargreaves DJ (1996). Responses to music in aerobic exercise and yogic relaxation classes. *British Journal of Psychology*, **87**, 535–547.

North AC and Hargreaves DJ (2000). Musical preferences during and after relaxation and exercise. *American Journal of Psychology*, **113**, 43–67.

North AC and Hargreaves DJ (2007). Lifestyle correlates of musical preference: 1. Relationships, living arrangements, beliefs, and crime. *Psychology of Music*, **35**(1), 58–87.

North AC, Hargreaves DJ and Hargreaves JJ (2004). Uses of music in everyday life. *Music Perception*, **22**, 41–77.

North AC, Hargreaves DJ and O'Neill SA, (2000). The importance of music to adolescents. *British Journal of Educational Psychology*, **70**, 255–272.

Oldham G, Cummings A, Mischel L, Schmidtke J and Zhou J (1995). Listen while you work? Quasi-experimental relations between personal-stereo headset use and employee work responses. *Journal of Applied Psychology*, **80**(5), 547–564.

Pegg C (1984). Factors affecting the musical choices of audiences in East Suffolk, England. In R Middleton and D Horn, eds, *Popular music 4: performers and audiences*, 51–73. Cambridge University Press, Cambridge.

Pitts S (2005a). What makes an audience? Investigating the roles and experiences of listeners at a chamber music festival. *Music and Letters*, **86**(2), 257–269.

Pitts S (2005b). *Valuing musical participation*. Ashgate, Aldershot.

Pool MM, Koolstra CM and van der Voort THA (2003). The impact of background radio and television on high school students' homework performance, *Journal of Communication*, **53**(1), 74.

Rentfrow PJ and Gosling SD (2003). The do re mi's of everyday life. The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, **84**(6), 1236–1256.

Rentfrow PJ and Gosling SD (2006). Message in a ballad: the role of musical preferences in interpersonal perception. *Psychological Science*, **17**(3), 236–242.

Reyner LA and Horne JA (1998). Evaluation of 'in-car' countermeasures to sleepiness: cold air and radio. *Sleep*, **21**(1), 46–50.

Riley M and Laing D (2006). *The value of jazz in Britain*. Arts Council England/Jazz Services, London.

Russell PA (1997). Musical tastes and society. In DJ Hargreaves and AC North, eds, *The social psychology of music*, 161–187. Oxford University Press, Oxford.

Siedliecki SL and Good M (2006). Effect of music on power, pain, depression and disability. *Journal of Advanced Nursing*, **54**(5), 553–562.

Sloboda JA (1992). Empirical studies of emotional response to music. In M Riess-Jones and S Holleran, eds, *Cognitive bases of musical communication*, 33–46. American Psychological Association, Washington, DC.

Sloboda JA (1999). Everyday uses of music listening: a preliminary study. In SW Yi, ed., *Music, mind and science*, 354–369. Western Music Institute, Seoul.

Sloboda JA, O'Neill SA and Ivaldi A (2001). Functions of music in everyday life: an exploratory study using the Experience Sampling Method. *Musicae Scientiae*, **V**, 9–32.

Stutts JC, Reinfurt DW, Staplin L and Rodgman EA (2001) *The role of driver distraction in traffic crashes*. Report prepared for the AAA foundation for traffic safety,

Washington, DC. www.aaafoundation.org/projects/index.cfm?button=distraction (consulted 14 August 2006).

Tarrant M, North AC and Hargreaves DJ (2001). Social categorisation, self-esteem, and the musical preferences of male adolescents. *Journal of Social Psychology*, **141**(5), 565–581.

Thompson S (2006). Audience responses to a live orchestral concert. *Musicae Scientiae*, **X**(2), 215–244.

Thompson S (2007). Determinants of listeners' enjoyment of a performance. Psychology of Music, 35(1), 20–36.

Turner ML, Fernandez JE and Nelson K (1996). The effect of music amplitude on the reaction to unexpected visual events. *Journal of General Psychology*, **123**, 51–62.

Zillmann D and Gan S (1997). Musical taste in adolescence. In DJ Hargreaves and AC North, eds, *The social psychology of music*, 161–187. Oxford University Press, Oxford.

CHAPTER 41

# Music in performance arts

## Film, theatre and dance

Annabel J. Cohen

## Introduction

### Music in performance arts: music that is not alone

Music is rarely alone (Cook 1998). Throughout civilization, it has typically been performed in conjunction with dance, poetic text, theatre, and, more recently, film and television (Cohen 1999). This chapter considers the psychology of music in the contexts of performance arts, in particular the arts of the moving image, drama, and dance. Research in the psychology of music far exceeds psychological research in any of the other arts. Within music psychology, research on the role of music in film and television constitutes a small but vibrant subdomain (e.g., Cohen 1994; Lipscomb 2008). The growing research on the psychology of film music reveals that the role of music in the context of other performance arts is amenable to psychological investigation. Similar progress can be envisioned for a psychology of music in theatre and dance where foundations are fortunately beginning to emerge (e.g., for dance Grove *et al.* 2005; for opera Boerner 2004). The present chapter may help to further set the stage for such work.

### Nomenclature

The psychology of music usually focuses on the listener, but in fact visual and kinaesthetic senses usually play a role, be it watching the motion of the performer, sensing the ambiance of the venue, or tapping one's foot to a beat. Music for film, video, theatre and dance more obviously requires not only a *listener* but also a *viewer*. For participatory dance, a *mover* sensitive to kinaesthetic information, that is bodily motion, must also be considered. In the present chapter, the term *audience* refers to the generic member of an audience. Thus the statement *the mind of the audience* refers to the mind of a typical member of the group attending a performance (notwithstanding individual differences).

## Music as part of a multimedia experience

The following section begins with a review of the role of music in film and television, taking a historical approach and considering the recent research in the field of music psychology. Next we will consider the role of music in dramatic contexts and finally dance performance and participatory social dance.

### Moving images and the psychology of music

The art of moving images began in the early twentieth century just after the beginnings of experimental psychology which originated only a few decades earlier, and, like film, owed its progress to twentieth century technologies. Hugo Münsterberg of Harvard University was the first psychologist to take film seriously.

As indicated in his book *The Photoplay*, published in 1916, he believed that of all the arts, film was most like music from a structural and psychological standpoint, although theatre and photography more strongly resembled film on the surface than did music. Münsterberg also noted the value of music in film for relieving tension, maintaining interest, providing comfort, reinforcing emotion, and contributing to the aesthetic experience. Of course his views applied only to silent film as he predated the talking film. Although there is no psychological research on the role of music in the original silent film context, the conditions of the silent film (i.e., without dialogue and sound effects) simplify psychological studies of the role of music in film, and several researchers have taken advantage of this to study effects of music on interpretation, beginning with Marshall and Cohen (1988).

Marshall and Cohen (1988) asked whether two contrasting examples of background music would alter observers' attitudes toward characters in a short (2-minute) animation. The animation featured two triangles and a circle and had been used in social psychology research by Heider and Simmel (1944), who showed that typically viewers ascribe personality traits to these geometric 'characters' in systematic ways. For example, the larger triangle was regarded as a bully. Using semantic differential judgements (bipolar adjectives scales such as fast–slow, strong–weak), Marshall and Cohen showed that the ratings of the film overall and of the individual characters differed under the two types of music. For example, the rated activity of the small triangle was greater for one than for the other of the two types of music. Further examples of the role of music on the interpretation of silent film have been shown in studies of sense of closure (Thompson *et al.* 1994), the perceived aggressiveness–friendliness of wolves (Bolivar *et al.* 1994), and realistic clips of ambiguous human interactions (Bullerjahn and Güldenring 1994; Cohen 1993; Shevy 2007; Tan *et al.* 2007).

It was assumed that the advent of the talking film in 1927 would drive music for film into obsolescence. However, surprisingly, films without music were found lacking (Kalinak 1992, p. 45; Kracauer 1960, p. 138), and a new industry for recorded music for the talking film replaced the old industry of live musical accompaniment for the silent film. The movie industry soon recognized the importance of the music soundtrack. Recently, Simonton (2007) analysed the relation between success of over 400 award-winning films (as measured by ticket sales, awards, and critic ratings) and success (as indicated by awards) of film songs and film scores. He sampled data from seven professional organizations which give annual awards for film including film music, either best score or song or both, for example, the Academy of Motion Picture Arts and Sciences (Oscars), the British Academy of Film and TV Arts (BAFTAs), and the Online Film Critics Society. Success of score, but not film song, was positively correlated with receipt of the best picture award. The positive influence of the film score was as high as that for visual effects (e.g., cinematography, costumes, etc.) and for technical aspects, though all of these were less effective than dramatic components (e.g., best actor). Simonton (2007) suggested that a successful score contributes to the film narrative without drawing attention to itself, while a successful song may actually disrupt the narrative.

Tan and colleagues (2008) compared the effects of background and foreground music. In their study, a 1.25 minute action sequence from Spielberg's *Minority Report* (2002) was paired with three soundtracks: (1) the original soundtrack which belonged to the shopping mall environment of the action, (2) the same soundtrack presented as a background score, and (3) a second soundtrack presented as background score. There were 245 participants who completed rating scales following the presentation of each condition. Background music as compared to foreground music in a film led to different interpretations with respect to characters' emotions, relationship between characters, and mood of the scene. The study is the first to show that different specific effects of the music arise depending on whether the music is part of the foreground drama (known as the *diegesis*), or part of the background (*non-diegesis*).

Music in a film is typically non-diegetic, having no basis in the reality of the story (cf. Gorbman 1987). This non-diegetic function of music raises an important question for the psychology of music. How does the audience use this musical information and why does it add value to a film, as suggested by Simonton's (2007) correlational study of award-winning films and

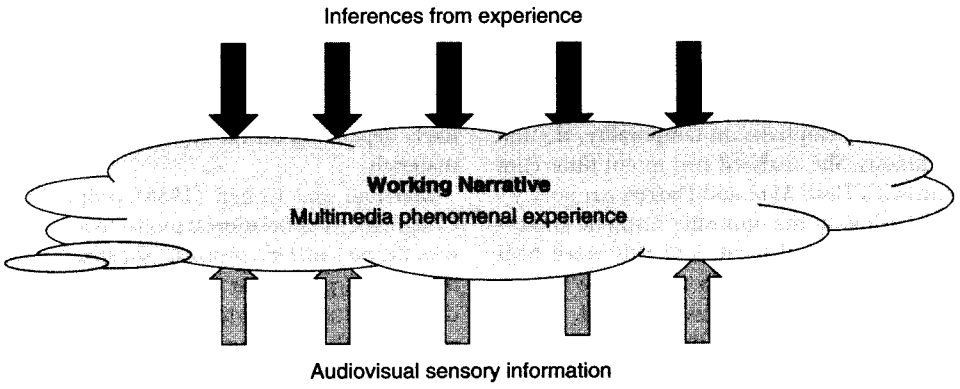



**Fig. 41.1** Sensory and cognitive sources of information that contribute to the audience's creation of the experience of an arts performance

as consistent with the experience of most filmgoers? The answer arises by considering that the audience does not aim to record all of the information impinging on the eye and ear. Rather, the audience seeks sufficient information to support engagement and entertainment, making use of two sources of information: sensory information arising from the real world presentation and past experience residing in long-term memory (see Figure 41.1). The audience is constantly synthesizing a story, a working narrative, derived jointly from sensory information and hypotheses or expectations based on long-term memory. According to this model, then, the audience takes what is needed from music and discards or ignores the rest. Given the capacity limitations of the mind (Miller 1956), including in multimedia settings (Fisch 2000; Mayer 2005; Sweller 2005), the audience has really no other choice. This approach is basic to Cohen's Congruence–Association Model of music in multimedia contexts, to be described later.

Film information other than music, such as the film's visual images, speech or sound effects, have been regarded as less effective than music in communicating emotional meaning (cf. discussion of music as the most efficient code for emotional expression in film by film-music theorist Kalinak [1992, p. 87]). The actual sounds of the music used to convey emotion are irrelevant to the story and consequently to the audience. Thus, acoustical information from the music can be discarded as far as engagement in film is considered. The sound of the music may be registered but not as an essential component of the story.

The integration of film and music within the constraints of mental capacity may also be viewed from the perspective of two analytic processes (Cohen 2005; Marshall and Cohen 1988). The first focuses on the *structural* aspects of the media: how patterns of information change in time and space (for example, rhythms of loudness and pitch, visual movement, visual intensity, colour, and contour). Marshall and Cohen (1988) argued that the temporal congruence of musical and visual patterns directed attention to specific elements of animation, and thus visual attention patterns could differ as a function of the particular pattern of music–visual congruency.

Several researchers have explored structural congruency of music and film. Lipscomb (2005) showed influences of temporal audiovisual alignment on judged effectiveness of simple visual animations. Kendall (2005, 2008) defined specific iconic archetypal forms (e.g., arch and ramp) which apply in both musical and spatial domains and has noted that judged congruence of visual and auditory versions of these forms is higher for some forms (e.g. arch) than others. Kim and Iwamiya (2008) examined the effect of formal congruency between sound patterns and animated text (Telops) using materials from television programmes as well as computer synthesis. They distinguished two kinds of formal congruency: similarity in temporal accent pattern and similarity in other changes (e.g., augmentation). In a related follow-up study, Iwamiya (2008) noted that increased formal congruency led to increased subjective congruency and as

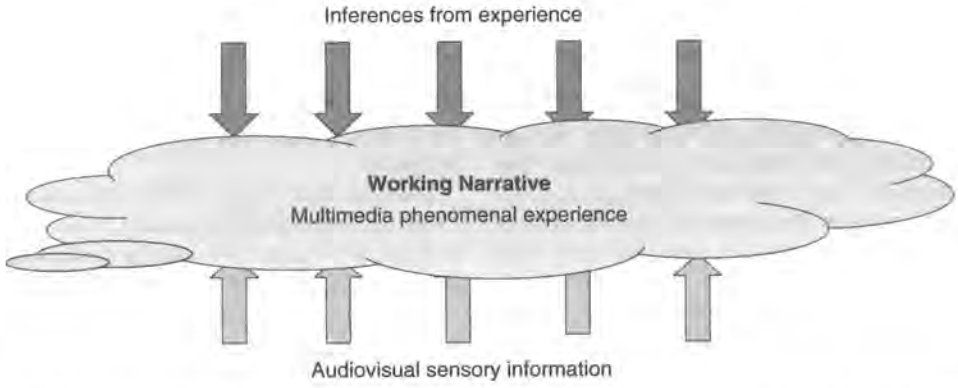



**Fig. 41.1** Sensory and cognitive sources of information that contribute to the audience's creation of the experience of an arts performance

as consistent with the experience of most filmgoers? The answer arises by considering that the audience does not aim to record all of the information impinging on the eye and ear. Rather, the audience seeks sufficient information to support engagement and entertainment, making use of two sources of information: sensory information arising from the real world presentation and past experience residing in long-term memory (see Figure 41.1). The audience is constantly synthesizing a story, a working narrative, derived jointly from sensory information and hypotheses or expectations based on long-term memory. According to this model, then, the audience takes what is needed from music and discards or ignores the rest. Given the capacity limitations of the mind (Miller 1956), including in multimedia settings (Fisch 2000; Mayer 2005; Sweller 2005), the audience has really no other choice. This approach is basic to Cohen's Congruence–Association Model of music in multimedia contexts, to be described later.

Film information other than music, such as the film's visual images, speech or sound effects, have been regarded as less effective than music in communicating emotional meaning (cf. discussion of music as the most efficient code for emotional expression in film by film-music theorist Kalinak [1992, p. 87]). The actual sounds of the music used to convey emotion are irrelevant to the story and consequently to the audience. Thus, acoustical information from the music can be discarded as far as engagement in film is considered. The sound of the music may be registered but not as an essential component of the story.

The integration of film and music within the constraints of mental capacity may also be viewed from the perspective of two analytic processes (Cohen 2005; Marshall and Cohen 1988). The first focuses on the *structural* aspects of the media: how patterns of information change in time and space (for example, rhythms of loudness and pitch, visual movement, visual intensity, colour, and contour). Marshall and Cohen (1988) argued that the temporal congruence of musical and visual patterns directed attention to specific elements of animation, and thus visual attention patterns could differ as a function of the particular pattern of music–visual congruency.

Several researchers have explored structural congruency of music and film. Lipscomb (2005) showed influences of temporal audiovisual alignment on judged effectiveness of simple visual animations. Kendall (2005, 2008) defined specific iconic archetypal forms (e.g., arch and ramp) which apply in both musical and spatial domains and has noted that judged congruence of visual and auditory versions of these forms is higher for some forms (e.g. arch) than others. Kim and Iwamiya (2008) examined the effect of formal congruency between sound patterns and animated text (Telops) using materials from television programmes as well as computer synthesis. They distinguished two kinds of formal congruency: similarity in temporal accent pattern and similarity in other changes (e.g., augmentation). In a related follow-up study, Iwamiya (2008) noted that increased formal congruency led to increased subjective congruency and as

well influenced the evaluation (goodness) of the visual form. On a larger scale, musicologist Kulezić-Wilson (2005) has recently explored the notion of the musicality of the structure of film, with specific emphasis on temporality, rhythm and kinesis. She analysed two recent films (Jim Jarmusch's *Dead Man* and Darren Aronofsky's *Pi*) focusing on the opposing aesthetic choices of the shot and the cut, and this work begs for psychological research to test the human sensitivity to the structures described. Cohen (2002) noted the rondo-like structure of *The Red Violin* and suggested comparable music and film research to explore its cognitive and aesthetic effects.

The second aspect of film and music for analysis is *meaning*. Earlier in this section, studies of film music were referred to that showed the effect of music on interpretation of a film, film characters or depicted world. The music can also lead to the feeling of emotions, which is an aspect of meaning that can often be best measured physiologically. Thayer and Levenson (1983) showed that background film scores were successful in both reducing and increasing electrodermal (skin conductance) responses to a stressful film. Baumgartner *et al.* (2006) have shown that classical music (as compared to no music) enhanced the emotional meaning of and physiological response (skin conductance, heart rate, respiration, and EEG) to emotional visual images taken from the International Affective Picture System (IAPS). Related work has been reported by Spreckelmeyer *et al.* (2006) who measured brain evoked potentials (ERP) to brief (350–550 ms) presentations of happy or sad emotional IAPS scenes and happy or sad single tones sung by trained opera singers. Participants focused on either the visual or auditory information and were required to rate the emotional meaning. Influence of integration of the two modalities was seen in ERP waves as early as 150 ms. The brain-wave recordings were more sensitive than were the ratings of the emotional valence. The visual scenes had a greater influence on the judgements while focusing on the audio than did the audio stimuli while focusing on the visual modality. These studies pave the way to research with actual video (as opposed to pictorial) and music materials so as to reveal the underlying music–video integration process. The research suggests that integration occurs quickly and that the brain treats the emotional meaning of visual and auditory material similarly, with sensory (audio or visual) dominance likely depending on the relative salience of the materials.

Marshall and Cohen (1988) proposed the Congruence–Associationist model to accommodate the two analytic processes for structure and meaning, arguing that structural congruencies direct attention to particular elements of a film scene, and that the associations (meaning) from the music are then ascribed to this direction of attention. In light of research on comprehension in other fields (e.g., Kintsch 1998, for text comprehension), Cohen (2001) expanded the model to emphasize the audience's narrative purpose in viewing a media presentation. She proposed that the audience is constantly engaged in generating a *working narrative*, a story of the unfolding drama. The working narrative is the audience's experience of a film. Cues from music as well as from the other sources of information contribute to the working narrative. The second expansion of the model accommodated a notion of the influential French film theorist Christian Metz. Metz gave prominence to music as one of five channels relevant to the experience of film: text, visual scenes, speech, sound effects (noise), and music (described by film theorist Stam 2000). The Congruence–Associationist model thus included these five channels as seen in Figure 41.2.

Marilyn Boltz (2004) conducted research on the role of music in television, and interpreted her results with respect to the Congruence–Associationist Model which she referred to as the C-A Model (to be referred to hereafter as CAM). In Boltz'study, 72 participants were presented with short (approximately 4 min) audiovisual television clips for which the music was presented as either mood congruent or incongruent. They were asked to attend to the audio track, the video track, or to both tracks. Subsequently, they were tested for their memory of the visual film clips and the music excerpts. Performance in the mood-congruent condition was superior to that of the incongruent condition. Performance in the incongruent condition was good for the attended audio or visual medium but did not extend to the unintended

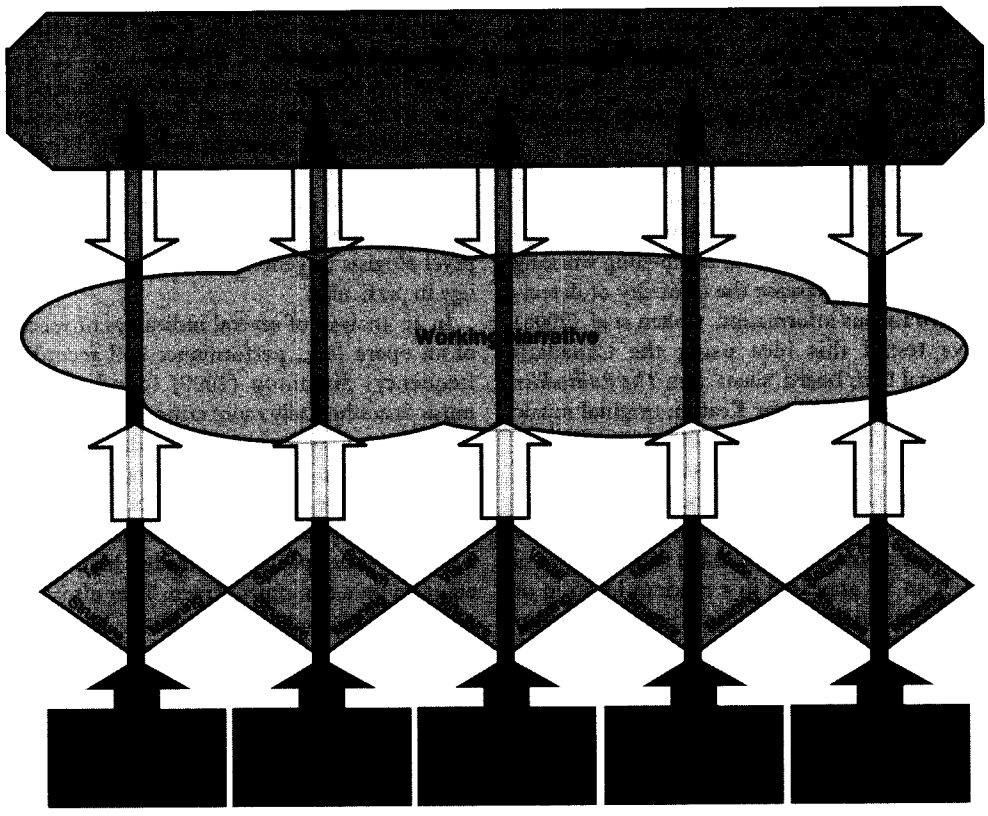


**Fig. 41.2** Congruence–Associationist framework for understanding film-music communication (see text for explanation).

medium as it did when the mood of the music and video clip coincided. According to Boltz, congruence in mood of the music and film elicited a search for similar structural properties in the music and film, and the success of the search led to jointly encoding the music and visual material and an integrated music–video memory. Thus, memory for music information would mean memory for the visual information. However, without mood congruencies, attention to the entire music visual complex suffers as the focus is on only one of the modalities (as per instructions in the experiment) and attention is not automatically directed to the other modality which conveys incongruent meaning.

Arguing that an effective working narrative leads to audience absorption or involvement in the film, our laboratory has aimed to determine the role of music in engaging the audience in the film, a phenomenon with which most moviegoers can identify. Two procedures were developed. One is a direct method in which the audience is asked to rate their level of absorption in a film clip presented under different conditions of music. Cohen *et al.* (2005) have shown higher absorption in the presence versus absence of music in short excerpts of several types of genre (e.g., animation, soap opera, drama). As predicted, more realistic genres (such as news) did not show the effect, although the relation between impact of music on absorption and realism of the genre was non-linear, no doubt a result of numerous variables arising from the specific examples chosen. Cohen *et al.* (2006), also explored the relative contributions of the three audio channels–sound effects, dialogue, and music–on self-rated absorption using two contrasting 1-minute clips (from *Witness* and *Day of the Jackal*) presented with only one type

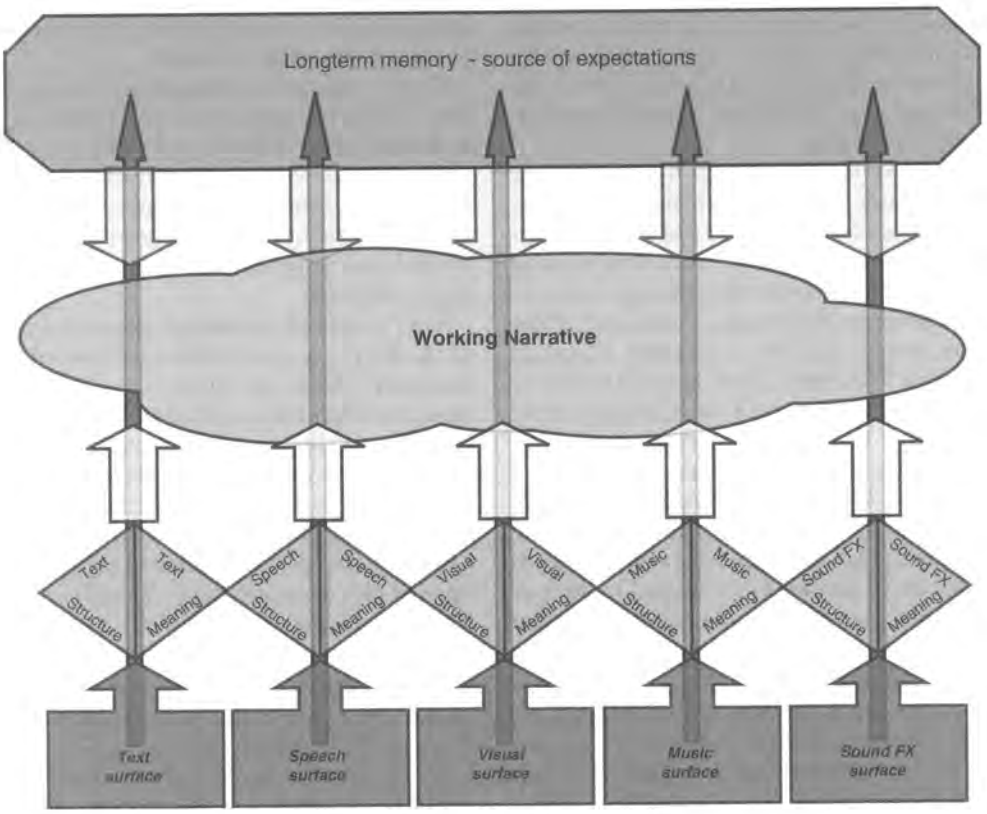



**Fig. 41.2** Congruence–Associationist framework for understanding film-music communication (see text for explanation).

medium as it did when the mood of the music and video clip coincided. According to Boltz, congruence in mood of the music and film elicited a search for similar structural properties in the music and film, and the success of the search led to jointly encoding the music and visual material and an integrated music–video memory. Thus, memory for music information would mean memory for the visual information. However, without mood congruencies, attention to the entire music visual complex suffers as the focus is on only one of the modalities (as per instructions in the experiment) and attention is not automatically directed to the other modality which conveys incongruent meaning.

Arguing that an effective working narrative leads to audience absorption or involvement in the film, our laboratory has aimed to determine the role of music in engaging the audience in the film, a phenomenon with which most moviegoers can identify. Two procedures were developed. One is a direct method in which the audience is asked to rate their level of absorption in a film clip presented under different conditions of music. Cohen *et al.* (2005) have shown higher absorption in the presence versus absence of music in short excerpts of several types of genre (e.g., animation, soap opera, drama). As predicted, more realistic genres (such as news) did not show the effect, although the relation between impact of music on absorption and realism of the genre was non-linear, no doubt a result of numerous variables arising from the specific examples chosen. Cohen *et al.* (2006), also explored the relative contributions of the three audio channels–sound effects, dialogue, and music–on self-rated absorption using two contrasting 1-minute clips (from *Witness* and *Day of the Jackal*) presented with only one type

of auditory background. Heightened absorption in the presence of music, as opposed to speech and sound effects, was noted for the *Witness* clip that used music which had originally accompanied it in the film.

The second methodology for studying the role of music on absorption in film is less direct, and is based on the rationale that higher absorption (arising from a more compelling working narrative) will increase the difficulty of detecting extraneous information. Cohen *et al.* (2008) have tested this idea using the Canadian National Film Board 'silent' film *The Railrodder* (1965) (starring Buster Keaton; original music composed by the eminent Eldon Rathburn). There were 72 participants assigned to one of three music background conditions (original, inappropriate, no music). While watching the 20-minute film, they responded with a key-press when an 'X' appeared in a random location of the computer screen on 20 occasions. The film was presented twice. Consistent with the hypothesis, the original music, as compared to no music, led to slower response time. On the second presentation, the original music also led to slower responses than inappropriate music. Individual scores on the Tellegen (1982) Absorption Scale correlated negatively with hit rate, validating the detection task as a measure of absorption.

The paradigms that have been developed for studying music perception and cognition in the context of film have thus begun to reveal the effects of music on meaning, memory, structural judgments, physiological variables, and absorption (involvement). The Congruence–Associationist model has been proposed to accommodate these data and suggest testable hypotheses. All of this may be useful in addressing the role of music in other performance art forms of theatre and dance.

## Theatre

The role of music in opera, musical theatre, and other dramatic genres has received little attention from researchers in music psychology, in contrast to that which has focused on film and video. Part of the reason for this arises from the live nature of theatre, which makes replication of experimental conditions difficult. Nevertheless, many questions raised by music in theatrical contexts in fact could be studied.

Opera is among the most ambitious and complex of the performance art forms. It entails an orchestral score, soloists, ensemble, libretto, sets, dance, as well as epic dramatic themes. How the mind integrates and appreciates such rich information should be of great interest to psychologists in general and to music psychology in particular.

In an analysis of several indicators of success of an opera (e.g., performance and recording frequency) Simonton (2000) found that the music is aesthetically more critical to opera than is the libretto. Gregory (1998) conducted a study of opera music in which listeners showed the changes in the strength of their emotional response to a selection of opera arias by moving a mouse whose position indicated high or low emotion on a computer screen. They also completed emotional adjective checklists. Consistency among listeners was shown, and higher emotion was elicited by voice than by the orchestra alone. Similar studies with video recording or in a live situation beg to be conducted.

Boerner (2004) proposed a hierarchical componential framework of performance quality in opera, emphasizing a distinction between the musical and staging aspects. The audience is viewed as judging the congruence of musical and staging dimension, congruence within these separate dimensions (the relation between soloists, choir, and orchestra), congruence within the individual components (relation among soloists). The author admits the concept congruence here, based on Adorno, is vague and in need of operationalism. A questionnaire study of audience assessment of live opera (Verdi's *La forz del destino*) supported Boerner's componential model, but suggested that staging contributed more to the judgments than did the music (Boerner *et al.* 2008). While this seems paradoxical, particularly in the face of Simonton's evidence for the importance of the music in opera, the authors suggest that high quality of the music is assumed by the audience, while the quality and type of staging is more variable and thus gains greater conscious attention. The authors note the importance of further related research that would explore different operas and control for individual differences.

It would also be of interest to apply this methodology to musical theatre, which bears similarities to opera yet has notable differences which might be regarded in terms of reduced cognitive load (e.g., use of popular musical style and native language). Music is also occasionally found in dramatic works. For example, Mendelssohn composed music specifically for Shakespeare's *A Midsummer Night's Dream*. Frankly, however, the psychology of music knows little about the role of music in musical theatre and in drama, and the field is wide open to exploration on the more microanalytic level using paradigms from film-music research or at the level of audience appreciation as recently developed for opera.

## Dance

With the advent of film and video recorders, progress in dance research seems more secure than in the past, now that dance performances and choreography can be preserved. Jordan (2000, p. 102) however emphasizes that no single source of information about dance provides an adequate basis for research, as each form of representation highlights some aspects at the expense of others. Notation systems for dance, such as Labanotation and Benesh, have both selective and subjective components.

With few exceptions, classical ballet entails dancing to music. Choreographers have varied in the significance they attach to the structure of the music and the function of ballet in representing this structure. The Russian choreographer Fedor Lopukhov (1886–1973) influenced the development of dance as a non-narrative art form (cf. Jordan 2002) which gave full importance to the structure of music. Lopukhov prescribed matching dance to music on the dimensions of emotional climaxes, curve, colour, and key changes. Major and minor keys were to be treated differently, 'en dehors' and 'en dedans' respectively. Lopukhov was an influence on Ballanchine, who danced in one of Lopukhov's famous works, 'Dance Symphony' (Jordan 2002, p. 3). Contemporaneous with these two eminent choreographers is Jacques-Dalcroze, the Swiss pedagogue who shared the basic principle of the structural similarities of music and human motion. He believed that the best way to understand and learn about music was to experience the elements of music with the body. His curriculum, known as Dalcroze Eurhythmics, has been taught to generations of young people worldwide. Recent attention has been directed to Dalcroze from the perspectives of psychology, developmental education, and African musical pedagogy respectively (Seitz 2005; Juntunen and Hyvönen 2004; Phuthego 2005) supporting the significance of Dalcroze' ideas, and implying the need for empirical research that will validate intuitions about the connections between bodily motion and musical understanding.

A pioneering psychological study of the interaction between music and dance is provided by Krumhansl and Schenk (1997) who used a film of Balanchine's choreography of Mozart's Divertimento No. 15 to test whether dance could represent the structural and expressive qualities of music. Three groups of participants were asked to watch the choreography alone, listen to the music without seeing the dance, or watch the dance and listen to the music simultaneously. All three groups then rated the tension and emotion throughout the piece as well as indicated the beginnings and endings of phrases. Results showed agreement among all three groups and additivity of the information from the separate channels of music and dance: the response in the 'music and dance' condition could be predicted as a combination of the responses to the 'music only' and 'dance only' conditions. The generality of these results to other works of dance would be worth exploring in view of the fact that the particular work chosen by Ballanchine was one in which the choreography was treated with great sensitivity to music-theoretic structural parallels.

Mitchell and Gallaher (2001) examined the ability to detect a 'match' between a piece of music and its intended choreography. The musical pieces were by John Cage, Peter Gabriel, and David Lanz, and the dances were choreographed and performed by a university-level dance educator. Participants were presented with a piece of choreography in silence followed by three selections of music. They were then asked to choose the musical selection that best matched the choreography. The reverse condition, in which the music was presented followed by three examples of dance, was also tested. Participants were also presented with a piece of

choreography paired with a piece of music, either intended or not intended to go with it. In both the sequential and simultaneous test format Mitchell and Gallaher's results show consistently high ratings for perceiving a 'match' between a piece of music and choreography intended to go with it. Participants' answers to a questionnaire indicated that their matches were influenced by similar characteristics perceived in the media, 'including emotion, fluidity, an African or Middle Eastern quality, and temporal characteristics such as rhythm and pace'. The results indicated that memory for dance was sufficient to enable judgements of congruence with musical excerpts presented later, and that the basis for judged music–dance congruence generalized across participants in the study and was multidimensional.

Based on both the work of Mitchell and Gallaher (2001) and that of a dance theorist, Fogelsanger and Afanador (2006) propose four parallels between music and dance that may contribute to a sense of congruency:

1 matching or intertwining pulse or rhythm
2 alignment of coincident structural temporal aspects other than pulse
3 analogous cross-modal qualities
4 complementary referents.

It is notable that the first two items refer to structural characteristics, while the latter refer to associations or meanings, which echoes a distinction made earlier between structure and meaning in the discussion of film and music. Dance can represent structure for its own sake (e.g., choreography of George Ballanchine) and be independent of typical narrative plot but can also tell a story through pantomime and other symbolic movement, such as the story of *Coppelia*, *les Sylphides*, *the Nutcracker*, or a *Midsummers Night's Dream*.

As such the proposal of Folgelsanger and Afanador (2006) fits within the CAM framework and is amenable to psychological test, for example, to determine whether in fact these parallel aspects of music and dance lead to perceived congruency, and whether some dimensions of congruency are more salient than others. While it is appreciated that music–dance structural congruence is hardly a universal choregraphic goal, it is still important to know what congruencies are discernable by an audience and what impact they have on appreciation. In regard to audience appreciation of choreography, Renee Glass completed a doctoral thesis which developed a qualitative and quantitative questionnaire (Audience Response Tool, ART) to determine audience responses to contemporary dance. The dances she examined entailed sound effects but not music in the traditional sense, but the tool in theory could be applied to dances in which music played a role.

Ivar Hagendoorn (2004) has recently proposed that the experience of watching dance creates pleasure for us through the fulfilling and undermining of our expectations. Viewers of a dancer to some extent sense the dancer's movement as if they were doing it themselves. Hagendoorn's proposal of a neurobiological system for a kinaesthetic response to dance aligns with recent evidence for mirror neurons which respond equally to an action carried out by the organism or simply from viewing such action (Rizzolatti and Craighero 2004). However, other research on biological motion detection by Troje (2002) accounts for perception of biological motion without having to call into play mirror neurons. Brown *et al.* (2006) in a unique if not extraordinary positron emission tomography (PET)-scan study of experienced dancers of tango localized a portion of the brain known as the cerebellar vermis that enables entrainment of dance steps to music as opposed to self-paced movement.

Whereas classical ballet and modern dance are performance art for the stage, social dance is an avenue open to and taken by most individuals by their adolescence or early adulthood. In spite of both the prevalence of social dance as an aspect of human behaviour and the reliance on music for this activity, there has been almost no attention of music psychology directed to the social dance situation. Music in the context of social dance in a sense legitimates patterns of human interaction that might seem ridiculous or embarrassing otherwise. Music is at least one of the components in social settings that allows people to behave differently than they otherwise would. The permission-giving power of music may result from effects of music that engage more emotional as opposed to analytic processing

(Niedenthal *et al.* 1999). Cross (2007, p 660) has suggested that

> music's powers of entrainment, together with its 'floating intentionality', fit it for use as a medium for communicative interactions in which meanings are under-determined to the extent that participants are free to develop their own interpretations of the significance of their own, and others', contribution to the collective musical behaviour.

## Concluding comments: future research, trans-art congruencies, and lifespan experience

Because music occurs frequently in the context of other arts, a model of music perception and cognition must take these other domains into account. The next decades may well see research programmes on the promising frontiers of the psychology of opera, musical theatre, and dance that will match the recent research activity of the psychology of music and the moving image. The psychology of music and the moving image may also do well to take note of recent research in opera and dance which aims at a global understanding of what the audience appreciates.

For all of the performance contexts for music, a recurrent concept is that of congruence: congruence between music structures and structure within the visual or kinaesthetic dimensions, congruence of meaning, and congruence that refers to the gestalt or sense of the whole. Thus, questions remain about the perceptual and cognitive responses elicited by these similarities between the music, visual, and kinaesthetic domains, and the temporal course of any integration processes.

Yet another aspect that will need attention is that of the role of individual experience of the audience member and when in the life course this experience occurs. A final diagram is therefore offered to reflect both the chronometric and synchronic dimensions of a general model of the audience engaged in music and the performance arts (Figure 41.3). This model, though very abstract, conceptualizes the mind of the audience as actively engaged in performances in

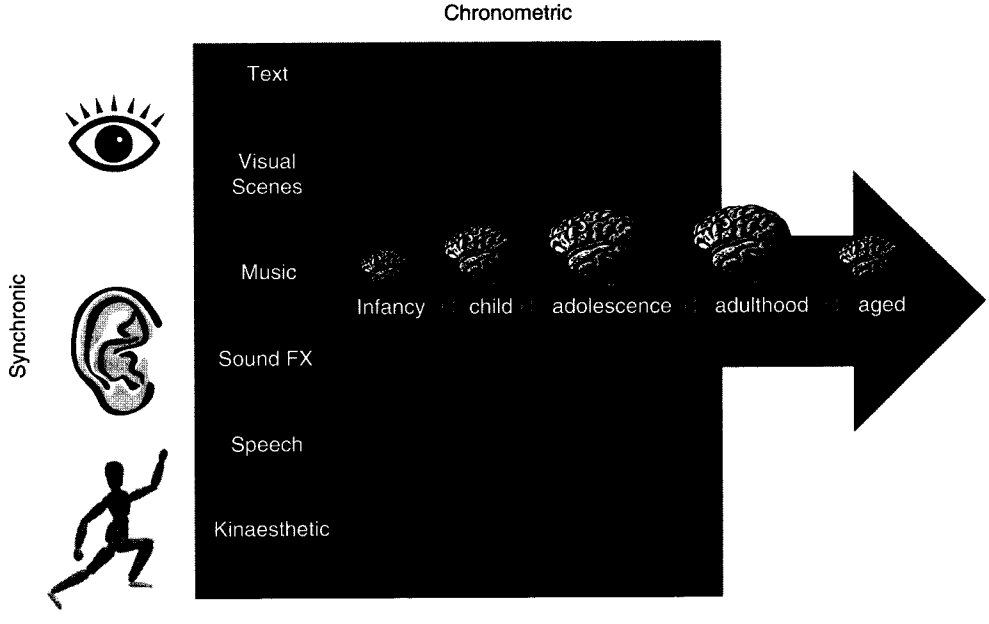



**Fig. 41.3** The Congruence–Associationist Model (CAM) in a developmental context. The kinaesthetic sense has been added to accommodate its significance as referred to in the text.

(Niedenthal *et al.* 1999). Cross (2007, p 660) has suggested that

> music's powers of entrainment, together with its 'floating intentionality', fit it for use as a medium for communicative interactions in which meanings are under-determined to the extent that participants are free to develop their own interpretations of the significance of their own, and others', contribution to the collective musical behaviour.

## Concluding comments: future research, trans-art congruencies, and lifespan experience

Because music occurs frequently in the context of other arts, a model of music perception and cognition must take these other domains into account. The next decades may well see research programmes on the promising frontiers of the psychology of opera, musical theatre, and dance that will match the recent research activity of the psychology of music and the moving image. The psychology of music and the moving image may also do well to take note of recent research in opera and dance which aims at a global understanding of what the audience appreciates.

For all of the performance contexts for music, a recurrent concept is that of congruence: congruence between music structures and structure within the visual or kinaesthetic dimensions, congruence of meaning, and congruence that refers to the gestalt or sense of the whole. Thus, questions remain about the perceptual and cognitive responses elicited by these similarities between the music, visual, and kinaesthetic domains, and the temporal course of any integration processes.

Yet another aspect that will need attention is that of the role of individual experience of the audience member and when in the life course this experience occurs. A final diagram is therefore offered to reflect both the chronometric and synchronic dimensions of a general model of the audience engaged in music and the performance arts (Figure 41.3). This model, though very abstract, conceptualizes the mind of the audience as actively engaged in performances in

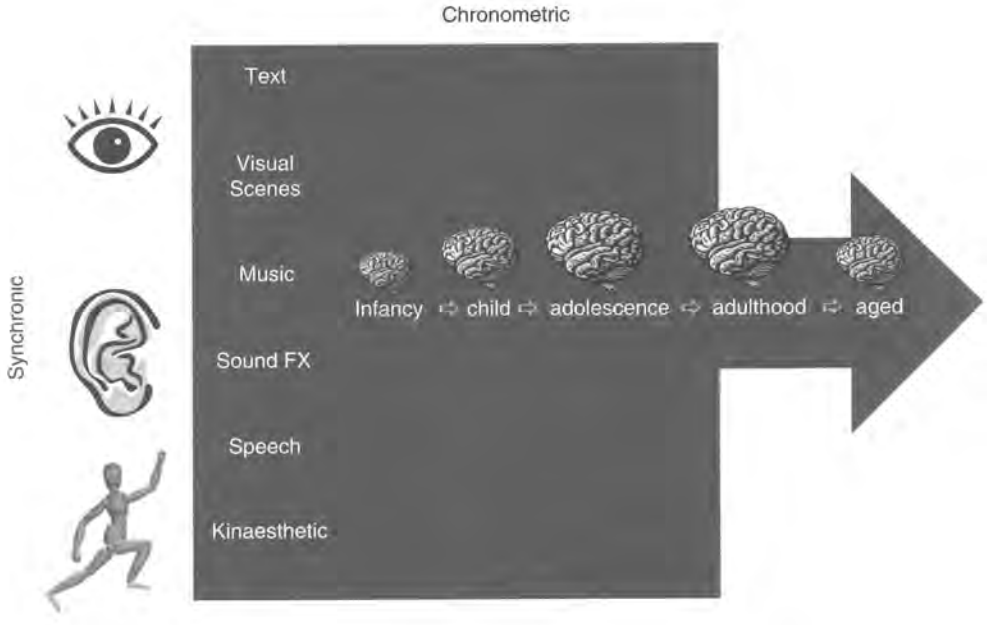



**Fig. 41.3** The Congruence–Associationist Model (CAM) in a developmental context. The kinaesthetic sense has been added to accommodate its significance as referred to in the text.

the arts and affected by such engagement in different ways across the lifespan.

## Author note and acknowledgements

The Social Sciences and Humanities Research Council is acknowledged for its support of the author's program of research in the psychology of film music.

## References

Baumgartner T, Esslen M and Jäncke L (2006). From emotion perception to emotion experience: emotions evoked by pictures and classical music. *International Journal of Psychophysiology*, **60**, 34–43.

Boerner S (2004). Artistic quality in an opera company: toward the development of a concept. *Nonprofit Management and Leadership*, **14**, 425–436.

Boerner S, Neuhoff H, Renz S and Moser V (2008). Evaluation of music theatre: empirical results on content and structure of the audience's quality judgment. *Empirical Studies of the Arts*, **26**, 15–35.

Bolivar VJ, Cohen AJ and Fentress JC (1994). Semantic and formal congruency in music and motion pictures: effects on the interpretation of visual action. *Psychomusicology*, **13**, 28–59.

Boltz M (2004). The cognitive processing of film and musical soundtracks. *Memory and Cognition*, **32**, 1194–1205.

Brown S, Martinez MJ and Parsons LM (2006). The neural basis of human dance. *Cerebral Cortex*, **16**, 1157–1167.

Bullerjahn C and Güldenring M (1994). An empirical investigation of effects of film music using qualitative content analysis. *Psychomusicology*, **13**, 99–118.

Cohen AJ (1993). Associationism and musical soundtrack phenomena. *Contemporary Music Review*, **9**, 163–178.

Cohen AJ (ed) (1994). The psychology of film music. Special volume, *Psychomusicology*, **13**, 1–165.

Cohen AJ (1999). The functions of music in multimedia: a cognitive approach. In SW Yi, ed., *Music, mind, and science*, 53–69. Seoul National University Press, Seoul, Korea.

Cohen AJ (2001). Music as a source of emotion in film. In PN Juslin, JA Sloboda, eds, *Music and emotion*, 249–272. Oxford University Press, Oxford.

Cohen AJ (2002). Music cognition and the cognitive psychology of film structure. *Canadian Psychology*, **43**, 215–232.

Cohen AJ (2005). How music influences the interpretation of film and video: approaches from experimental psychology. In RA Kendall and RW Savage, eds, *Selected reports in ethnomusicology: special issue in systematic musicology*, **12**, 15–36.

Cohen AJ, Macmillan KA and Drew R (2005). The role of music on absorption in film and video: effects of genre. Poster presented at Neuroscience and Music II. Leipzig.

Cohen AJ, MacMillan KA and Drew R (2006). The role of music, sound effects and speech on absorption in a film: the congruence–associationist model of media cognition. *Canadian Acoustics*, **34**, 40–41.

Cohen AJ and Siau, Y-M (2008). The Narrative Role of Music in Multimedia Presentations: The Congruence-Association Model (CAM) of Music and Multimedia. In K. Miyazaki, M. Adachi, Y. Hiraga, Y. Nakajima and M. Tsuzaki (eds), Proceedings of the 10th International Conference on Music Perception & Cognition (ICMPC10), August 2008, Sapporo Japan. Adelaide: Causal Productions. CD-ROM.

Cook N (1998). *Analysing musical multimedia*. Clarendon Press, Oxford.

Cross I (2007). *Music and cognitive evolution*. In RIM Dunbar, L Barrett, eds, *Oxford handbook of evolutionary psychology*, 649–667. Oxford University Press, Oxford.

Fisch SM (2000). A capacity model of children's comprehension of educational content on television. *Media Psychology*, **2**, 63–91.

Fogelsanger A and Afanador K (2006). Parameters of perception: vision, audition, and twentieth-century music an d dance. 38th Congress on Research in Dance, Tempe, Arizona.

Gorbman C (1987). *Unheard melodies: narrative film music*. Indiana University Press, Bloomington, IL.

Gregory AH (1998). Tracking the emotional response to operatic arias. In SW Yi, ed., *Fifth ICMPC Proceedings*, 265–270. Seoul National University, Seoul.

Grove R, Stevens C and McKechnie S (eds) (2005). *Thinking in four dimensions: creativity and cognition in contemporary dance*. Melbourne University Press, Melbourne, AU.

Hagendoorn I (2004). Some speculative hypotheses about the nature and perception of dance and choreography. *Journal of Consciousness Studies*, **11**, 79–110.

Heider F & Simmel M (1944). An experimental study of apparent behavior. *American Journal of Psychology*, **57**, 243–259.

Iwamiya S (2008). Subjective congruence between moving picture and sound. Paper submitted to S Lipscomb (organizer-invited session) Experiencing musical multimedia: empirical investigations of cross-modal perception and cognition organized by S Lipscomb for the 10th International Conference on Music Perception and Cognition, Hokkaido, Japan.

Jordan S (2000). *Moving music: dialogues with music in twentieth-century ballet*. Dance Books, London.

Jordan, S (2002) *Writings on Ballet and Music, Fedor Lopukhov, Studies in Dance History*, Madison, Wisconsin: University of Wisconsin Press.

Juntunen M-L and Hyvönen L (2004), Embodiment in musical knowing: How body movement facilitates learning with Dalcroze Eurythmics, *British Journal of Music Education* **21** (2), 199–214.

Kalinak K (1992). *Settling the score*. University of Wisconsin Press, Madison, WI.

Kendall RA (2005). Empirical approaches to musical meaning. In RA Kendall and R Savage (eds). *Selected Reports in Ethnomusicology*, **12**, 69–102.

Kendall R (2008). Stratification of musical and visual structures ii: visual and pitch contours. Paper submitted to S Lipscomb (organizer-invited session) Experiencing musical multimedia: empirical investigations of cross-modal perception and cognition organized by S Lipscomb for the 10th International Conference on Music Perception and Cognition, Hokkaido, Japan.

Kim K-H and Iwamiya S-I (2008). Formal congruency between telop patterns and sounds effects. *Music Perception*, **25**, 429–448.

Kintsch W (1998). *Comprehension: a paradigm for cognition*. Cambridge University Press, Cambridge.

Kracauer S (1960). *Theory of film: the redemption of physical reality*. Oxford University Press, Oxford.

Krumhansl CL and DL Schenck (1997). Can dance reflect the structural and expressive qualities of music? A perceptual experiment on Balanchine's choreography of Mozart's Divertimento No. 15. *Musicae Scientiae*, **1**, 63–85.

Kulezić-Wilson D (2005). Composing on screen: the musicality of film. Unpublished Dissertation. Faculty of Art, Design and Humanities of the University of Ulster.

Lipscomb S (2005). The perception of audio-visual composites: accent structure alignment of simple stimuli. *Selected Reports in Ethnomusicology*, **12**, 37–67.

Lipscomb S (2008). (Organizer-invited session) Experiencing musical multimedia: empirical investigations of cross-modal perception and cognition. 10th International Conference on Music Perception and Cognition, Hokkaido, Japan.

Marshall S and Cohen AJ (1988). Effects of musical soundtracks on attitudes to geometric figures. *Music Perception*, **6**, 95–112.

Mayer RE (2005). Cognitive theory of multimedia learning. In R. E. Mayer (ed.) *Cambridge handbook of multimedia learning* (31-48). New York, NY: Cambridge.

Miller GA (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, **63**, 81–97.

Mitchell RW and MC Gallaher (2001). Embodying music: matching music and dance in memory. *Music Perception*, **19**, 65–85.

Münsterberg H (1916/1970). *The photoplay: a psychological study*. Arno, New York.

Niedenthal PM, Halberstadt JB and Innes-Ker AH (1999). Emotional response categorization. *Psychological Review*, **106**, 337–361.

Phuthego M (2005). Teaching and learning African music and Jaques-Dalcroze's eurhythmics. *International Journal of Music Eduction*, **23**, 239–248.

Rizzolatti G and Craighero I (2004). The mirror-neuron system. *Annual Review of Neuroscience*, **27**, 169–92.

Seitz J (2005) Dalcroze, the body, movement and musicality. *Psychology of Music*, **33**, 419–435.

Shevy M (2007). The mood of rock music affects evaluation of video elements differing in valence and dominance. *Psychomusicology*, **19**(2), 57–78.

Simonton DK (2000). The music or the words? Or how important is the libretto for an opera's aesthetic success? *Empirical Studies of the Arts*, **18**, 105–118.

Simonton DK (2007). Film music: are award-winning scores and songs heard in successful motion pictures? *Psychology of Aesthetics, Creativity, and the Arts*, **1**(2), 53–60.

Spreckelmeyer K, Kutas M and Urbach T (2006). Combined perception of emotion in pictures and musical sounds. *Brain Research*, **1070**, 160–170.

Stam R (2000). *Film theory: an introduction*. Blackwell, Malden, MA.

Sweller J (2005). Implications of cognitive load theory for multimedia learning. In RE Mayer, ed., *Cambridge handbook of multimedia learning*, 19–30. Cambridge University Press, New York.

Tan S-L, Spackman MP and Bezdek MA (2007). Viewers' interpretation of film characters' emotions: effects of presenting film music before or after a character is shown. *Music Perception*, **25**, 135–152.

Tan S-L, Spackman MP and Wakefield EM (2008). Source of film music (diegetic or non-diegetic) affects viewers' interpretation of film. In Miyazali et al. (eds.), Proceedings of the 10th International Conference on Music Perception & Cognition (ICMPC10), August 2008, Sapporo, Japan. Adelaide: Causal Productions. CD-Rom.

Tellegen A (1982). *Brief manual for the Differential Personality Questionnaire*. Unpublished manuscript, University of Minnesota, Minneapolis.

Thayer JF and Levenson R (1983). Effects of music on psychophysiological responses to a stressful film. *Psychomusicology*, **3**, 44–54.

Thompson WF, Russo FA and Sinclair D (1994). Effects of underscoring on the perception of closure in filmed events. *Psychomusicology*, **13**, 9–27.

Troje NF (2002). Decomposing biological motion: a framework for analysis and synthesis of human gait patterns. *Journal of Vision*, **2**, 371–387.

CHAPTER 42

# Peak experiences in music

John Whaley, John Sloboda and Alf Gabrielsson

## Introduction

Many people experience intensely positive affective states in response to music. The intensity of these experiences often brings about a sense of transcendence and transformation, and their relative rarity makes them greatly prized in the minds and lives of the individuals experiencing them. People tend to retain vivid memories of such experiences many years after they have occurred, and they are often cited as a major reason for continued involvement with music (to increase the chances of such an experience occurring again: see e.g. Sloboda [2005]). Yet these are 'everyday' experiences in the sense that they seem to require no special context in which to occur. They may occur in the home, while travelling, in places of work or study, or in the more specialized contexts of concert hall or place of worship. They may be triggered by any genre of music, and during both listening and performance. Peak experiences also seem to require no specific musical training or expertise; they have been described by young and old, by the learned and the musically unlettered. Moreover, these experiences are not conjured by acts of will or intention—they usually come unannounced and unexpected—and thus cannot be the subject of skill acquisition or pedagogy.

Maslow (1959) was one of the first to formalize the study of what he called Peak Experiences, or 'moments of highest happiness and fulfilment' (1999, p. 85), and 'a spurt in which the powers of the person come together in a particularly efficient and intensely enjoyable way' (1999, p. 106). Maslow spent many years studying peak experiences, and his findings laid the groundwork upon which all peak literature is built. While his investigations were not limited to musical triggers, he discovered very early that 'the two easiest ways of getting peak experiences [are] through music and sex' (1976, p. 169), and it is upon the former that this chapter is based.

What follows highlights research on the peak experience, emphasizing literature focusing on music-specific peaks. After outlining four studies fundamental to the study of peaks in music, a section will discuss precursors to peaks and proposed differences between those who have achieved peaks and those who have not. A section on the cognitive, perceptual, emotional, and physical phenomena associated with peak experiences will then be followed by an investigation on the after-effects of peaks. Next, a section discussing methodologies for the investigation of musical peak experiences highlights the possibilities and difficulties of this work. Finally, a brief section summarizes the contents of the chapter and looks towards the future of research in this field.

## Fundamental peak experience in music literature

Four studies represent the empirical backbone of the study of the musical peak experience. As these will be heavily referenced in each subsequent section, a brief description here of their methods and findings is helpful.

Maslow coined the term 'peak experience' after analysing 190 written and 80 oral responses, having asked participants for descriptions of

> the most wonderful experience of your life; happiest moments, ecstatic moments, moments of rapture, perhaps from being in love, or from listening to music, or suddenly 'being hit' by a book or painting, or from some great creative moment.
>
> (1999, p. 83)

The contents of the experiences were individually unique, and in reporting his findings the author 'added together all the partial responses to make a "perfect" composite syndrome' (1999, p. 83). Beyond arguing that music is a reliable trigger of peak experiences, Maslow describes many of the phenomena experienced both during and after the peak, and asserts that: the peak experience is a self-validating, self-justifying moment with its own intrinsic value. It fills an individual with a sense of wonder and awe. Peaks are never negative, unpleasant, or evil, and cause a characteristic disorientation of time and space as well as a loss of fear, anxiety, doubts, and inhibitions. Frequently eliciting sudden feelings of happiness and well-being, peaks also bring about a heightened sense of an individual's physical and 'existential' state of being, a sort of positive hyper self-awareness. After a peak the individual may experience a number of effects, including a more positive view of the self, other people, and the world, various therapeutic effects (both physical and mental), even a sense that life itself is worth living.

In an effort to analyse art-specific peaks, Panzarella (1980) gathered reports of an 'intense joyous experience of listening to music or looking at visual art' (p. 71) from a pool largely composed of musicians and visual artists. Roughly half of responses were music-related. Content analysis revealed 11 response categories that were factored into four major dimensions of the experience: renewal ecstasy, motor-sensory ecstasy, withdrawal ecstasy, and fusion-emotional ecstasy: renewal ecstasy concerned a new or altered perception of the world and a desire to engage further with the artistic medium. Any physical elements like tears, changed heart rhythms, chills, or quasi-physical elements like a floating sensation were categorized as motor-sensory ecstasy. Withdrawal ecstasy involved a perceptual narrowing, during which there was a sense that everything disappeared except the object (i.e. music or art) and that the object was in perfect focus. Finally, fusion-emotion ecstasy referred to a 'fusion' with the medium, like a sense that the music or painting merged and became 'one' with the individual. Panzarella (1980) also argues that the peak experience has three distinct 'temporal stages', beginning with 'cognitive responses and loss of self; climaxing with continued loss of self and motor responses; and subsiding with emotional responses, [and] self-transformations' (p. 69). Various precursors and aftereffects of the 'joyous experience' are also reported.

Ecstasy was also a central theme in an earlier study by Laski (1961), who collected 63 incidents of 'transcendent ecstasy'. Triggers of such experiences included childbirth, sex, art, beauty, science, and religion, though the author concludes that 'of all the more common triggers to ecstasy, music would be the most rewarding to study in any attempt to find a relation between the qualities of the triggers and the effects produced' (p. 190). The contents of these episodes were collapsed into four categories: feelings of loss (e.g. sense of time, space, self, sorrow), feelings of gain (e.g. joy, salvation, perfection, unity with music or mankind, knowledge), feelings of ineffability (the experience eludes description or analysis), and quasi-physical feelings (e.g. peace, pain, light, heat).

To date, the most comprehensive attempt to understand exceptional experiences with music is the Strong Experiences of Music (SEM) project conducted by Gabrielsson and Lindström Wik (1993, 2000, 2003). This project makes use of over 1000 free descriptions of strong experiences in connection with music, collected and content-analysed over more than 10 years. Participants were asked, in as much detail as they could, to describe 'the strongest, most intense experience of music that [they had] ever had' (2003, p. 163). The resulting wealth of data culminated in the Strong Experiences of Music Descriptive System (SEM-DS), which lists the most common attributes of strong experiences in music and includes seven categories: general characteristics, physical reactions, perception, cognition, feeling/emotion, existential/transcendental, and

personal/social. These attributes, which will be further discussed on p. 455–7 of this chapter, provide significant insight into the nature of the strong experience, what occurs *during* the experience. Various other publications using SEM data discuss other aspects of the experience, including therapeutic after-effects (Gabrielsson and Lindström 1995), emotions (Gabrielsson 2001), experiences reported by old people (Gabrielsson 2002), and musical triggers (Gabrielsson 2006).

A note about nomenclature is necessary. Maslow's (1999) peak experience, Laski's (1961) ecstasy, Panzarella's (1980) joyous experience, and Gabrielsson and Lindström-Wik's (2003) strong experience are not identical concepts. For instance, peak experiences are, by definition, positive, whereas some strong experiences are negative. In addition, one could argue for a separation between *exceptional* experiences (like strong, ecstatic, or joyous) and true *peak experiences* which, thanks to Maslow, have a slightly more stringent definition.

Recognizing these inherent problems, we focus this chapter on 'peak experience in music' for two reasons: First, the term peak experience is a recognized term within the broad psychological literature. A term such as joyous experience is not in common use. Second, although there are negative experiences triggered by music, we would argue that positive experience is the most frequent and prototypical case, from which negative experiences may be considered a deviation, or a special case. However, we do not intend to restrict this discussion to any single, strictly defined term. We recognize the overlap amongst these terms, and accept the fact that no single term or definition fully captures the experience discussed here, be it peak, strong, optimal, ecstatic, or otherwise.

## Precursors and predictors of the peak experience

Maslow (1976) concluded that 'any experience of real excellence, of real perfection. ... tends to produce a peak experience', but immediately and frustratingly added 'though not always' (p. 17). Peak experience literature is full of such qualifications, and has shown mixed results with regard to the demographic, situational, and musical differences between those who have achieved peaks (henceforth referred to as 'peakers') and those who have not ('non-peakers').

Maslow stated that only the self-actualized, i.e. those who had achieved the highest level of his hierarchy of needs (1943), were capable of enjoying a peak experience, as it was a 'momentary intensification of their general tendency to apprehend reality more directly and clearly' (1999, p. 19). Lacking a strict definition of self-actualization, subsequent research identified some general personality characteristics of peakers which include openness, tender-mindedness, flexibility, an experimental nature, an interest in creative work and a holistic, interpersonal, and emotional approach (Masluk 1999; Mathes 1982; McClain and Andrews 1969; Paffard 1970). Peakers appear to display a general willingness to explore and experience states of consciousness differing from their usual states of being and are 'open' in the broadest sense of the word (Masluk 1999).

Within music-specific peak literature, Panzarella (1980) discovered only minor, insignificant personality differences between peakers and non-peakers and argues that 'peak experiences represent an unusual mode of functioning for the individual rather than an intensification of the person's usual mode of functioning' (p. 85).

Research has also shown mixed results with regard to the predictive power of demographic differences between peakers and non-peakers. Studies comparing the peak reports of various demographic sample groups have shown that despite significant differences between groups' age, gender, education, race, and religion, there are no significant differences observed in their descriptions of peak experiences (Allen *et al.* 1964; Gordon 1985; Yeagle *et al.* 1989; Lanier *et al.* 1996). Nevertheless, some older studies found significant differences in peak descriptions between Caucasians and blacks (Mathis and McClain 1968), non-psychotics and psychotics (Margoshes and Litt 1966), and low and high verbal abilities (McClain and Andrews 1969).

Research on where a musical peak experience is most likely to occur has shown a few tendencies, though the situational unpredictability

remains similar to that of personality and demographic differences. Having collected a sample of experiences with music during childhood, one study found that peak experiences were most likely to occur during relaxed, informal music listening, while surrounded by positive social and environmental factors (like friends and family), and in casual non-evaluative surroundings, like listening to the radio (Sloboda 2005). Although this finding is confirmed by other studies (e.g. Gabrielsson and Lindström Wik 2003), research has yet to systematically examine the peak-inducing power of various settings. With regard to music performance, Gabrielsson and Lindström Wik (2003) found that, despite the large number of musicians in their sample, over 80 per cent of all SEM were in music listening. A survey of performing musicians revealed a wide variety of circumstantial factors in performance-specific peak experiences, including perception and involvement with audience, relationships among band members, and connection with music being performed (Boyd and George-Warren 1992).

With regard to musical genre, Maslow found that classical music, specifically 'the great classics' (1976, p. 170) were the most reliable trigger of peak experiences. Panzarella (1980) and Laski (1961) also list 'serious/classical music' as the primary peak-inducing music in their work, though each include descriptions of experiences induced by rock, folk, jazz, and other previously less well-represented genres. A recent analysis of Gabrielsson's (2006) SEM data produced 15 separate categories/genres, which included all varieties of classical, folk, and popular pieces. This analysis also revealed significant demographic differences between participants and their triggers, where older participants were more likely to list classical music and younger participants more likely to list popular music.

Do peak-producing musical works have structural or sonic features that predict their ability to evoke peaks? Very little research has been undertaken on this topic. In analysing the capacity of music to evoke 'chills' (one physical component of some peak experiences), Sloboda (1991) discovered that certain musical devices were statistically associated with the precise temporal location of a chill. These devices, including appoggiaturas and syncopations, all seem to involve violations of listener expectancy. His findings confirm those of Gabrielsson (2006) in that the responses collected were not specific to any one musical genre, and were found in a wide range of genres.

## The nature of the peak experience

> I was filled by a special feeling that the music began to take command of my body. I was charged in a way. A tremendous feeling of harmony which made me really enjoy the music, and I found it difficult to stand still ... the mystery and power really gripped me. I was filled by an enormous warmth and heat. I really swallowed all the notes that were streaming out in the air, not a single note, effect or sequence missed my hungry ears. I was captivated by each of the instruments and what they had to offer me. Nothing else existed!
>
> Gabrielsson and Lindström Wik (2003, p. 166)

The research literature contains much data, exemplified by excerpt above, about the nature of the peak experience drawn from retrospective accounts of how it felt *while* it was happening. Such accounts demonstrate key features of the peak: their positivity, their powerfulness, rarity, non-voluntary nature, and some emotional, physical, perceptual, and cognitive correlates (Gabrielsson and Lindström-Wik 2003; Maslow 1976). In short, the peak is the antithesis of mundane, it goes beyond the everyday, and is not easily forgotten. The peak has even been called indescribable, and many find their expressive palate insufficient to properly elucidate the experience (Thorne 1963; McClain and Andrews 1969). Peaks are somewhat unpredictable; they are non-volitional and cannot be 'generated' (Maslow 1976).

Verbal reports provide insights into the subjective content of a peak experience, what happens during it, and what it feels like. There is no single element a peak experience *must* contain—rather there exists a diverse set of elements, each of which could be said to be indicative or typical. Gabrielsson and Linstrom Wik (2003) categorize

these elements into a three-level hierarchy, the top and most general of which includes seven categories: general characteristics, physical reactions, perceptions, cognitions, feelings/emotions, existential/transcendental and personal/social aspects. These categories are proposed as capturing the most common elements of the experience.

The following excerpt, taken from Gabrielsson and Lindström (1993, p. 123), is from a woman's description of how she experienced Tchaikovsky's Symphony No. 6, *Pathétique*, when she listened to it for the first time. Note the various elements—physical, perceptual, cognitive, emotional, transcendental and personal/social—in her description:

> I have had similar [strong] experiences of other music but none so terribly deep as *Pathétique*. In certain passages it evokes sobs, and I feel totally crushed. My listening is fully concentrated, the rest of the world disappears in a way, and I become merged with the music or the music with me, it fills me completely. I also get physical reactions, wet eyes, a breathing that gets sobbing in certain passages, a feeling of crying in my throat and chest. Trying to find words for the emotion themselves, I would like to use words as crushed, shaken, tragedy, maybe death, absorption, but also tenderness, longing, desire in vain, a will to live, prayer. The whole experience also has the character of a total standstill, a kind of meditative rest, a last definite and absolute end, after which nothing else can follow.
>
> After the experience I am sort of 'gone'—the people in the foyer and their murmur are at a distance, like a stage set rather than something real. I have a hard time trying to talk to them, to 'get started' and to 'return' to the ordinary reality. The most difficult part is to talk about the music experience itself, it is possible only after a while, when it has faded out a little. One of my concert friends is very sensitive to classical music and he too can have wet eyes. It has happened several times that we are just standing there, shaking our heads and looking down on the floor. Nobody finds any words, one cannot add anything to what the music already has told.

Descriptions like these testify to both the uniquely powerful nature of the strong experience and its ability to include many experiential elements in a single moment. This woman's experience is not only physical ('wet eyes ... breathing... . sobbing'), emotional ('tenderness, longing'), or perceptual ('the rest of the world disappears, and I become merged with the music'), but *all* of these, and more. It is these unique but overlapping elements that Gabrielsson and Lindström Wik's (2003) SEM-DS attempts to categorize, and in what follows these categories will be elaborated with reference to other research studies where available.

Physical reactions during the peak experience may include such physiological responses as chills/thrills, tears, changes in breathing, heart rate, and body temperature. Beyond internal physical manifestations, overt actions range from closing of the eyes to singing, shouting, jumping or dancing around. Contrarily, some people experience total physical immobility, or a sudden desire to be alone. Finally, included with physical reactions are a set of quasi-physical reactions, like a sense that the body has been dissolved into the music or merged with it, or that the body has actually transcended time or space, an out-of-body experience.

Some of the physical reactions which typically accompany peaks also occur in the context of emotional reactions to music which are less intense. This is particularly true of chills/thrills (e.g. Konecni *et al.* 2006; Panksepp 1995; Sloboda 1991). These are phenomena such as pilo-erection, 'shivers down the spine', tears and lump in the throat. These phenomena are ubiquitous—in some people they can occur several times a day, or even several times within the same piece of music. On their own they may be of rather little lasting significance, but when accompanied by other psychological experiences may become part of a memorable and significant peak.

Perceptual phenomena during peaks may include auditory, visual, tactile, kinaesthetic, synaesthetic, intensified/multimodal perception, and musical perception–cognition elements. The auditory–musical perception may be accompanied by strong visual impressions, tension or relaxation of the muscles, or simply an intensified sense of perception, among others. Performers, like listeners, frequently

describe a perceptual sense of being 'embedded' in the sound (Boyd and George-Warren 1992).

The cognitive category includes elements like changed attitude, a changed experience of situation, body and mind, time and space, or wholeness, a loss of control, a changed attitude to music, a connection to old associations, memories, and thoughts, cognitive imagery, and musical cognition–emotion. Cognitive elements may relate to some of the more lasting effects of the experience (further discussed in the next section) when, for example, an individual achieves a new understanding of some element in their life, like a relationship or their appreciation of the art.

Emotional elements include strong/intense emotions, both positive and negative emotions, and mixed emotions. Positive emotions predominate, especially joy and happiness, while typical low arousal feelings include a sense of peace and calm, and high arousal feelings include rapture and euphoria. In addition to the prototypical positive emotions, there are examples of negative emotions such as loneliness, sadness, anxiety, or even anger during the experience. Typically such negative feelings are related to various personal and other circumstances rather than to the music itself. Mixed emotions often result from the mix of positive affect and the negative social or personal circumstances, such as hearing music that reminds one of a lost loved one. For the majority of people, however, peaks are accompanied by unusually intense positive emotions.

Existential and transcendental aspects include, among other things, reflections on human life and existence, cosmic experience, experience of other worlds, religious visions, and encounters with the divine. While these items are easily categorized and listed, their experiential impact and post-experience implications can be immense. Many, for example, describe their peak experience as having resulted in a better understanding of the human condition, or offered a glimpse at God, Heaven, or similar 'other worldly' elements.

Finally, personal and social aspects include elements such as feeling liberated, uplifted or cleansed, getting new insights, hope, power and increased self-esteem, and further various therapeutic effects.

## Effects of the peak experience

Maslow (1999) likened the peak to a 'personally defined Heaven from which the person then returns to earth. Desirable after-effects of such an experience, some universal and some individual, are ... very probable' (pp. 110–111) and discovered that some 'after-effects may be permanent, [though] the high moment itself is not' (p. 14). His analysis of peak reports revealed seven broad categories of after effects, not all of which result from every experience. These categories are supported by a later peak experience literature (Frick 1982; Lanier *et al.* 1996; Warmoth 1965).

In relation to music, Panzarella (1980) reported that 90 per cent of participants attributed long-lasting, often permanent, effects to their peaks. Through his analysis, he derived a list that closely resembles Maslow's (1999) and includes:

1 General appreciation of music or visual art
2 Altering of self-appreciation
3 Altering of relationships with other people
4 Altering of attitudes toward life or world in general
5 Increased aesthetic sensitivity
6 Enhanced perception of everyday reality
7 Long-lasting mood effects.

The first element, 'general appreciation of music or visual art', is notably absent from previous lists of after-effects, including Maslow's (1999). Participants in this study who achieved joyous experiences through either art or music emerged with a renewed appreciation for the medium. A similar study found that participants reported a desire to continue pursuing music either as a hobby or profession, and a desire to continue listening and engaging with music (Gabrielsson and Lindström Wik 2003).

Gabrielsson and Lindström (1995) document the ability of strong experiences to elicit positive therapeutic effects. Having analysed hundreds of SEM reports, the authors mention therapeutic implications including: relief from physical pain, grief, and depression, a release of personal and social barriers, increased self-confidence and sense of one's ego, even a total reorientation of one's attitude and value system. This report,

the most specific and thorough of its kind, provides substantial music-specific evidence of one of Maslow's original effects.

A final after-effect that warrants discussion is the desire to continue pursuing peaks. Research has shown that peak experiences in music, particularly in early childhood, have long-lasting effects on the likelihood of engagement with music (Sloboda 2005). Having collected descriptions of experiences with music before age ten, Sloboda (2005) found that positive experiences early in life were likely to spur continued involvement in music or, in other words, continued peak-seeking behaviour. Similar behaviour has been observed in adult music performers, who demonstrate a desire to repeat the strong feelings they experience while performing, so much so that a survey of pop and rock musicians uncovered a tendency to turn to drugs to induce the same feelings (Boyd and George-Warren 1992). Interestingly, veteran skydivers showed similar peak-seeking rationale for their continued desire to engage in their work (Lipscombe 1999), which suggests that peaks are of great consequence to the continued pursuit of both professions.

Despite the preceding evidence in support of Maslow's (1999) list of after-effects, there is a significant methodological pitfall in peak research, namely the attempt to measure an effect without a baseline. As a result of the unpredictability of the peak, it is hard to establish any independent, pre-peak data. Thus when a study discovers a 'more positive view' or an 'altered perception', these are derived from retrospective comparisons provided in the narratives of participants. One method for overcoming this limitation would be testing a large number of people who all attend the same musical event, monitoring experience before, during, and after. This would permit statistical comparison of pre and post-experience scores among peakers. This approach is currently being pursued by two of the current authors.

In conclusion, existing research provides evidence that music-related peak experiences have meaningful after-effects, and that most people who experience peaks report such effects (Ebersole 1972). The fact that reports in Gabrielsson and Lindström Wik (2003) contain all of Maslow's (1999) and Panzarella's (1980) after effects further supports this conclusion. The effects of the musical peak experience, it seems, do not end after the fleeting moment, but carry on in the lives of the peaker.

## Methodology

The majority of the music-related peak experience literature is based on what Sloboda (2005) calls the 'autobiographical memory technique', which he finds both 'rich and workable'(p. 188). Relying on participants' memories for these significant moments, researchers using this technique simply ask participants to produce narratives about experiences with music. The prompt can be as simple as asking participants to describe 'the ... most intense experience of music that you have ever had' (Gabrielsson and Lindström Wik 2003, p. 163), or more generic like that used by Maslow (1999), who asked participants to describe 'the most wonderful experience of [their lives]'(p. 83). Variations on these prompts are found in numerous articles (Laski 1961; Panzarella 1980; Lanier *et al.* 1996; Minassian *et al* 2003; Sloboda 2005).

The autobiographical memory technique is not, however, without its methodological limitations. There is some doubt about participants' ability to recall specific details about an experience, particularly when recall is many years after the episode. There are also disclosure issues, as strong verbal and expressive skills are critical to describing the experience (Thorne 1963; McClain and Andrews 1969). This could mean that many potential peakers, including children, are expressively incapable of describing their peak experiences. Finally, the qualitative methods of this technique afford somewhat limited quantitative analyses.

As a response to the latter, peak experience researchers have developed hybrid methodologies which combine the autobiographical memory technique with a set of quantitative methods. After asking participants to freely 'describe one incident in your life characterized by highest happiness', the Experience Questionnaire (Privette 1983, 1984; Privette and Bundrick 1987) presents a series of 42 Likert-scale questions and five scaled descriptions of the peak experience, to which participants rate their agreement. Reliability of this

technique is well within reasonable limits (Privette and Sherry 1986), and the quantitative data produced permits more aggressive statistical analyses.

This method has been successfully applied to music-specific experiences. Participants in Gabrielsson and Lindström Wik (2003) not only produced free descriptions of their strong experiences but also rated how well their reactions corresponded with a large set of statements on strong experiences with music. Across participants, statements indicating positive feelings and other salient characteristics in SEM (e.g. absorption, confirmation, transcendence) received high ratings, whereas statements indicating negative feelings received low ratings. Factor analysis revealed up to 14 factors accounting for 64 per cent of the total variance. Most of these factors corresponded to (and thus confirmed) various categories in the SEM-DS, though there were also factors revealing co-occurrence of different categories (e.g, one factor included transcendent state, perfection, peace and happiness, healing experience), which may be difficult using solely content analysis of free descriptions. Notable, however, is the fact that these 14 factors fail to account for over one-third of the total variance—a testament to the extremely multifaceted character of SEM. Nevertheless, using these hybrid qualitative/quantitative methods, this study is able to collect bias-free, ecologically valid narratives and use quantitative techniques to test for the presence of correlated elements. For further discussion, see Gabrielsson and Lindström Wik (2003, p. 186 onwards).

Despite the benefits of these newer techniques, the ability to derive sound conclusions from qualitative data should not be underestimated. In his autobiographical memory research, for instance, Sloboda (2005) coded qualitative statements for the presence or absence of certain situational elements. These included internal significance (where the significance of the event was attributed to some factor internal to the music—such as a particular sound event or quality) and external significance (where the significance was attributed to some factor external to the music—such as the social context). Analysis showed a statistically significant relationship between internal and external significance. There were no cases of internal significance when the external significance was negative. This qualitative data thus clearly reveals something about the necessary conditions for a peak.

In sum, the study of musical peak experience poses substantial methodological challenges which require empirical ingenuity to overcome. Progress is likely to be furthered by a broader combination of qualitative and quantitative methods. Such progress should never blind us, however, to the fact that a full appreciation of peak experiences requires us to keep hold of the uniqueness and unrepeatability of each individual experience. Such uniqueness is not 'irrelevant noise' to be filtered out of analyses, but is a core component of what makes these experiences what they are.

## Conclusions

Musical peak experiences are a significant component of the lives of many people. They are powerful, valued, have lasting effects, and—for some—are a reason for continued engagement with music. They involve elements of life that could be described as transcendent, transformational, even spiritual. Indeed, it is perhaps no coincidence that music has been such an essential part of many religious traditions and rituals, aimed at encouraging various states of 'uplift', be they contemplative or ecstatic (e.g. Becker 2001). Nonetheless, psychology rightly and necessarily concentrates on the inner characteristics of such experiences rather than their wider social and cultural meanings.

The research reviewed in this chapter shows that, despite significant methodological difficulties and a relatively small literature, a number of facts about the musical peak experience are known with some certainty:

1 Not all individuals are equally prone to peak experiences. Some elements of personality (specifically flexibility and openness) seem to encourage these experiences. On the other hand, demographic factors such as age, gender, and education seem to be of little importance.

2 All types and genres of music can engender peaks, though perception of aesthetic value seems important.

3 Many different settings and environmental circumstances can be found in descriptions of peaks, including both formal and informal music listening and performance.

4 Peaks contain strong and distinctive perceptual, cognitive, physical, and emotional components which make them highly salient to the person experiencing them.

5 Peaks are often remembered, and ascribed causal significance for various after-effects, long after their occurrence.

Relative to other areas of music psychology (e.g. cognition/perception, developmental topics), the study of peaks has not attracted as large a body of research. As a result, the topic has not benefited to the full extent from recent methodological developments in the field, for instance, the use of multivariate statistical methods, or innovative experience-sampling methods. With such methods, it should be possible to make progress on some of the unanswered questions in the field, such as whether there are distinct syndromes within the larger phenomenon, and whether specific enabling conditions can be more precisely characterized, thus potentially helping people to find their way more reliably to these high-value and life-enhancing experiences.

## References


Allen RM, Haput TD and Jones RW (1964). Analysis of peak experiences reported by college students. *Journal of Clinical Psychology*, **20**, 207–212.

Becker J (2001). Anthropological perspectives on music and emotion. In PN Juslin, JA Sloboda, eds, *Music and emotion: theory and research*, 135–160. Oxford University Press, Oxford.

Boyd J and George-Warren H (1992). *Musicians in tune: seventy-five contemporary musicians discuss the creative process*. Simon and Schuster, New York.

Ebersole P (1972). Effects and classification of peak experience. *Psychological Reports*, **30**, 631–635.

Frick WB (1982). Conceptual foundations of self-actualization: a contribution to motivation theory. *Journal of Humanistic Psychology*, **22**, 33–52.

Gabrielsson A (2001). Emotions in strong experiences with music. In PA Juslin, JA Sloboda, eds), *Music and emotion: theory and research*, 431–449. Oxford University Press, Oxford.

Gabrielsson A (2002). Old people's remembrance of strong experiences related to music. *Psychomusicology*, **18**, 103–122.

Gabrielson A (2006). Strong experiences elicited by music—what music? In P Locher, C Martindale and L Dorfman, eds, *New directions in aesthetics, creativity and the arts*, 251–267. Baywood Publishing, New York.

Gabrielsson A and Lindström S (1993). On strong experiences of music. *Jahrbuch der Deutschen Gesellschaft für Musikpsychologie*, 10, 118–139.

Gabrielsson A and Lindström S (1995). Can strong experiences of music have therapeutic implications? In R Steinberg, ed., *Music and the mind machine. The psychophysiology and psychopathology of the sense of music*, 195–202. Springer-Verlag, Berlin.

Gabrielsson A and Lindström Wik S (2000). Strong experiences of and with music. In D Greer, ed., *Musicology and sister disciplines: past, present and future*, 100–108. Oxford University Press, Oxford.

Gabrielsson A and Lindström Wik S (2003). Strong experiences related to music: a descriptive system. *Musicae Scientiae*, **7**(2), 157–217.

Gordon RD (1985). Dimensions of peak communication experiences: an exploratory study. *Psychological Reports*, **57**, 824–826.

Konecni VJ, Wanic RA and Brown A (2007). Emotional and aesthetic antecedents and consequences of music-induced thrills. *American Journal of Psychology*, **120**, 619–643.

Lanier LS, Privette G, Vodanovich S and Bundrick CM (1996). Peak experiences: lasting consequences and breadth of occurrences among realtors, artists, and a comparison group. *Journal of Social Behavior and Personality*, **11**(4), 781–791.

Laski M (1961). *Ecstasy: a study of some secular and religious experiences*. Cresset Press, London.

Lipscombe N (1999). The relevance of the peak experience to continued skydiving participation: a qualitative approach to assessing motivations. *Leisure Studies*, **18**(4), 267–288.

Margoshes A and Litt S (1966). Vivid experience: peak and nadir. *Journal of Clinical Psychology*, **22**, 175.

Maslow AH (1943). A theory of human motivation. *Psychological Review*, **50**, 370–396.

Maslow AH (1959). Cognition of being in the peak experience. *Journal of Genetic Psychology*, **94**, 43–66.

Maslow AH (1976). *The farther reaches of human nature*. Penguin Books, New York.

Maslow AH, (1999). *Toward a psychology of being*, 3rd edn. John Wiley and Sons, Princeton, NJ.

Masluk TJ (1999). Reports of peak and other experiences during a neuro-technology-based training program. *Journal of the American Society for Psychical Research*, **93**, 1–98.

Mathes EW (1982) Peak experience tendencies: scale development and theory testing. *Journal of Humanistic Psychology*, **22**(3), 92–108.

Mathis WJ and McClain EW (1968). Peak experiences of white and Negro college students. *Journal of Clinical Psychology*, **24**, 318–319.

McClain EW and Andrews HB (1969). Some personality correlates of peak experiences: a study in self-actualization. *Journal of Clinical Psychology*, **25**(1), 36–38.

Minassian C, Gayford C and Sloboda JA (2003). Optimal experience in music performance: a survey of young musicians. Papers presented at a meeting of the society for education, music, and psychology research. London, March 2003.

Panksepp J (1995). The emotional sources of 'chills' induced by music. *Music Perception*, **13**, 171–207.

Panzarella R (1980). The phenomenology of aesthetic peak experiences. *Journal of Humanistic Psychology*, **20**(1), 69–85.

Paffard MK (1970). Creativity and peak experiences. *British Journal of Educational Psychology*, **40**, 283–290.

Privette G (1983). Peak experience, peak performance, and flow: a comparative analysis of positive human experiences. *Journal of Personality and Social Psychology*, **45**(6), 1361–1368.

Privette G (1984). *Experience questionnaire*. The University of West Florida, Pensacola, FL.

Privette G and Bundrick CM (1987). Measurement of experience: construct and content validity of the experience questionnaire, *Perceptual and Motor Skills*, **65**(1), 415–332.

Privette G and Sherry D (1986). Reliability and readability of questionnaire: peak performance and peak experience. *Psychological Reports*, **58**, 491–494.

Sloboda JA (1991). Musical structure and emotional response: some empirical findings. *Psychology of Music*, **19**(2), 110–120.

Sloboda JA (2005). Music as a language. In Sloboda JA, *Exploring the Musical Mind*, 175–189. Oxford University Press, Oxford.

Thorne FC (1963). The clinical use of peak and nadir experience reports. *Journal of Clinical Psychology*, **19**, 248–50.

Warmoth A (1965). The peak experience and the life history. *Journal of HumanisticPsychology*, **3**, 86–91.

Yeagle E, Privette G and Dunham F (1989). Highest happiness: an analysis of artists' peak experience. *Psychological Reports*, **65**, 523–530.

# CHAPTER 43

# Musical identities

Raymond MacDonald, David J. Hargreaves and Dorothy Miell

## Introduction

Many chapters within this Handbook have highlighted just how important music is for us all, and the extent to which music is ever-present within modern society. This ubiquitous aspect of music signals many features about how we as humans respond to music and highlights the universal importance of musical communication. Indeed, the notion that we all respond emotionally to music is not a vague utopian ideal, but the conclusion drawn by an increasing number of academic researchers interested in researching the foundations of musical behaviour (Juslin and Sloboda 2001; Trevarthen 2002). We suggest, given the importance of music within our lives, that it influences how we see ourselves and how we relate to the world around us. In short, we contend that music crucially influences our identity.

This chapter presents a number of key themes relating to the concept of musical identities. We provide a definition of identity, with a discussion of why identity is a timely topic for consideration. Also presented is a specific definition of musical identities, and we implicate music as an important feature of personal identity. We then move on and consider a number of specific research examples to shed further light upon the importance of musical identities. We present an overview of a series of studies investigating musical identities of jazz musicians. These highlight the utility of qualitative techniques, and in particular focus group and semi-structured interview methods, for understanding how professional musicians construct their identities in relation to both their musical activities and wider psychological and cultural issues. The chapter looks next at how theories of motivation and the self can help to explain some of the behavioural aspects of musical identities. We provide evidence that children's self-concepts, and in particular their levels of confidence (both of which are related to musical identities), can influence the rate of musical development and musical achievement, drawing briefly on a study which compares the views of pupils, parents and teachers about what it is to be 'good at music'.

## Defining identities

Our contention is that music is not only important for us all, but that it plays a fundamental role in the development, negotiation and maintenance of our personal identities (MacDonald *et al.* 2002; Roe 1999; Stålhammar 2006). Any individual involved in musical activities (i.e. both listening and performing) develops aspects of personal identity that are inextricably linked to these musical behaviours. Moreover, for young people at least, it seems that music may play an absolutely fundamental role in the development and maintenance of identity: a number of researchers have provided evidence to suggest that music is the most important recreational activity in which young people engage (Zillman and Gan 1997; Roe 1999). In terms of signalling to the world fundamental aspects about a young person's sense of self, music is often used as a kind of 'badge of identity' (Tarrant *et al.* 2002; MacDonald and Miell 2000). Furthermore, there is growing evidence that music remains a fundamental part of our identity across the life span (Applegate and Potter 2002; Carlton 2006).

Turning to identity research at a more general level, Anthony Giddens asserts that identity, within contemporary society, is the most

important issue with which humans must engage (Giddens 1991). We all have to make choices about our lives; from mundane questions about what to wear when we get up in the morning, to life-changing decisions about family, friendships and career paths. A key point here is that in earlier times, when traditional social order with clearly defined roles gave much more specific guidance regarding our life choices, these decisions were much easier to make and in some ways made for us. However, we now have many more options open to us and we have to make these decisions, what Giddens calls 'identity work', in a much more individualized and detailed manner. As Giddens states:

> What to do? How to act? Who to be? These are focal questions for everyone living in circumstances of late modernity—and ones which, on some level or another, all of us answer, either discursively or through day-to-day social behaviour.
>
> (1991, p. 70)

Consequently, identity research is now established as a multidisciplinary field, taking place in a wide variety of academic contexts, to such an extent that there exists a international organization called *The Society for Research on Identity Formation*, and a related journal, *Identity*, whose aim is 'to provide a forum for identity theorists and researchers around the globe to share their ideas and findings regarding the problems and prospects of human self-definition'. We find identity research not only in all branches of the arts and humanities, but also within the much more positivistic worlds of medicine and science. No single discipline has exclusive rights to identity research, nor can each identity research project be informed by all current thinking on identity (MacDonald *et al.* 2004). Of course, each research project will highlight different facets of identity processes and, in many cases, may involve differing definitions of identity. For example, the studies of jazz musicians outlined below utilize a discursive definition of identity, and emphasize how language shapes and constructs our sense of self, whereas the motivational theories presented later on in the chapter take a more social cognitive approach to the self. While there is no agreed definition of identity, researchers generally emphasize aspects that relate to how individuals construct themselves in relation to one or more phenomena. There is also an acknowledgement that each of us has many identities relating to the different roles we adopt (mother, musician, friend, colleague etc.). These identities are not discrete, will often merge and are constantly being negotiated and evolving. Identity is currently a very topical research issue for academics from many disciplines, perhaps because it is a fundamental issue for individuals worldwide in this period of postmodernity.

## Musical identities

Moving on to the specific domain of music, we contend that music plays an important role in the negotiation, construction and maintenance of identities. For example, an individual who is a keen Bob Dylan fan will maintain a sense of self that is influenced by this liking for Bob Dylan: many of his friends may also be Bob Dylan fans, and his choice of clothes and political beliefs may be influenced by his taste in music. His linguistic devices may be influenced by his musical tastes, and in conversation he may choose to signal these musical tastes as a marker of his individuality. Thus, music plays a central role in this individual's identity. Zillman and Gan (1997), Tarrant *et al.* (2002), and DeNora (2000) all investigate the role that music plays in identity construction and maintenance. Zillman and Gan (1997) provide evidence to suggest that music is the primary leisure activity and, in terms of identity construction, the most important recreational activity undertaken young people. Tarrant *et al.* (2002) highlight the importance that shared musical tastes have in forming and maintaining friendship groups. This demonstrates how music can be an important marker of identity and now young people in particular want to have musical tastes in common with their friends. De Nora (2000) takes a sociological perspective in demonstrating how enmeshed music is within everyday life, and quite clearly draws a direct link between the importance of music and the construction of the self.

Individuals who are involved in music-making (be that as a professional opera singer or

an occasional singer in the bath tub) develop an identity that is crucially influenced by these activities. For example, a young child going to piano lessons for the first time may be praised by her parents for practising and, over a period of time, sees her piano playing as one of many important activities in which she is engaged. Thus, when talking about her routine and weekly activities to friends and family her view of herself as a piano player influences how she talks and thinks about herself: as such, she develops an identity as a musician in the broadest sense of the word.

Another important point to note is that the identity of being a musician is a socially and culturally defined concept. We don't acquire the label 'musician' simply by attaining advanced technical skills. By this we mean that viewing ourselves as musical is not directly linked to the levels of technical ability or knowledge that we have. Similarly, there is no legal control over who can or can not call themselves musicians. It appears that whether or not we view ourselves as 'musicians' or 'musical' is crucially influenced by our social and cultural surroundings and by the ways in which we relate to people around us (Green 2002). The notion of being a musician is influenced by certain non-musical factors, which involves what we have termed *identity paradoxes*. Lamont (2002) highlights how the school setting plays a role in a child's developing musical identity by taking a detailed look at the music education environment, and demonstrating how the organization of lessons within a school has a crucial effect upon a child developing musical identity. Borthwick and Davidson (2002), on the other hand, look at more informal family environments, and how communication within the family influences children's developing sense of musicality. Specifically, these authors show how one sibling adopts the identity of *musician* while other siblings, who may be musically experienced, view themselves as not musical in comparison to this sibling. O'Neill (2002) investigates how young people learn to view themselves as musicians and uses a qualitative social constructionist approach to investigate the relationship between a young person's sense of musicality and wider social influences.

Gender issues are also important to take into consideration when thinking about how musical identities develop and are maintained. A number of researchers have highlighted how cultural expectations and hegemonic influences affect the developing musical identities of males and females differently, both in terms of musical tastes and participation (Dibben 2002; Green 1997; McClary 1991; O'Neill 1997a; Welch 2006). For example, Dibben (2002) and McClary (1991) discuss how women are underrepresented within the professional music world and Green (1997) provides evidence to support the notion that hegemonic masculine identities facilitate the development of a composer's identity. This may help explain why there are proportionally far fewer women, in comparison to men, working as composers within Western classical music. O'Neill (1997a) shows how gender stereotypes influence musical instrument choices made by boys and girls while Welch (2006) points out that the exclusion of women from cathedral choirs dates back to AD 597.

Crucially, all of the work highlighted above demonstrates how environmental and wider cultural influences are brought to bear on self-perceptions of musicianship and musical tastes. In particular, this work also demonstrates how essentially non-musical variables fulfill fundamental roles in an individual's sense of developing musical identity.

## Identities in jazz

One particular approach to studying musical identities is to consider how musicians talk about their musical lives. This approach was utilized in series of studies that investigated how jazz musicians construct, maintain and negotiate their musical identities in a variety of different contexts (Wilson and MacDonald 2005; MacDonald and Wilson 2005, 2006). This approach is presented as an exemplar of how research into musical identities can be conducted from a qualitative perspective. The key methodological points are that this research can utilize both focus group and interview methodology. Additionally, the data can be analysed using a range of theoretical perspectives including thematic analysis, interpretative phenomenological analysis, discourse analysis and narrative analysis. The jazz context provides

a particular useful and timely way to study identity. On the one hand the performance of this music represents a unique form of collaborative creativity (Sawyer 1992; Meadows 2003; Berliner 1994). Secondly, while there is a growing literature investigating the particular psychological and musical processes of jazz music (Cooke and Horne 2002), there is a need for research to focus upon the social aspects of identity maintenance, and in particular the multifaceted ways in which the collaborative context of jazz performance influences these identity projects (Sawyer 2003).

The first study utilized a thematic analysis of focus group data to investigate how jazz musicians talked about improvisation (Wilson and MacDonald 2005). What emerged from focus group data was a negotiated social construct rooted in the understanding of the participants themselves and an excellent method for investigating musical identities. From the analysis two main themes were identified. The first theme was termed '*Jazz is*'. While there was considerable debate around the topic of what jazz music is, the discussions centred around: *improvising and composing; swing elements of jazz* and *collective processes involved in jazz playing.* Improvisation was seen as a stronger expectation in jazz than in other forms of music, and was discussed both as the elaboration of a tune and as a looser form of collective creativity. *Swing*, the maintenance of a common and consistent pulse of a particular quality, was widely seen as a central yet intangible quality in the collaboration of musicians playing jazz. A perceived tension between collective and individual creativity, such as that identified by Sawyer (1992), also emerged in the accounts presented, however; achieving the requisite collective interaction was voiced as more important than individual goals of self-expression. Individual and collective creativity were often viewed as existing in balance with each other in jazz. Social interaction therefore becomes extremely important in its execution. The second broad theme that emerged from the data was termed '*Jazz life is*' and related to wider social and cultural issues pertaining how jazz musicians live. Within this category discussion centred around: *social context, professional context, being in a group* and *self and others.* These four themes gave detailed accounts of what the participants reported as being the key features of a jazz musician's life style.

This focus group analysis presented a thematic account highlighting the importance of shared understandings and practices in musical, social and cultural terms and features of identity for these participants. This thematic analysis coded what participants said into descriptive categories, and this is a common approach within qualitative studies of musical identity (e.g. Monson 1996). However, discourse analysis seeks to move beyond the categorical grouping of thematic content analysis to investigate structural features in discourse and to explore how particular speech acts perform a range of functions for a speaker and move beyond a thematic account (Edwards and Potter 1992).

With this issue in mind we examined discursive features of dialogue between jazz practitioners during the focus group interviews to reveal the relationship between the construction of specific meanings for the music and musicians and the negotiation of musical identities (Wilson and MacDonald 2005). This paper reported three key elements of identity work from a discursive analytic position: interpretative repertoires for jazz; the resolution of uncertainties created in handling these repertoires within the conversation; and ways in which the data are characterized by the focus group context. Two distinct interpretative repertoires (cf. Gilbert and Mulkay 1984) were employed in these data, which we termed the 'expert' and the 'naïve'. These were used flexibly to support participants' claims to identity as jazz musicians, while positioning others as separate from this identity. In the expert repertoire, jazz was accounted for as music produced by groups of individuals whose hard-won knowledge and skills gave them a particular insight into the music. In the 'naïve' repertoire participants emphasized the ineffable and soulful aspects of jazz music that were seen as being beyond complete comprehension.

The two papers discussed above give accounts of focus group data analyses and suggested a number of themes and features that we wished to follow up outwith the group context. In particular, we wanted to gather the views of both male and female practitioners (all focus group participants had been men) who were not

previously known to the interviewer (both researchers had known the focus group participants personally and professionally). A series of ten individual in-depth interviews with professional jazz musicians in London was therefore conducted to provide rich data on individual identities, social contexts, and the experience of improvising (MacDonald and Wilson 2006). The findings highlighted how identity as a jazz musician was negotiated with other individual identities. For example, the relationship between gendered identities and identity as a jazz musician was examined as an instance of how conflict between these identities may allow hegemonic influences to operate.

In summary, these three papers highlight the importance of talking about music in terms of negotiating the constantly evolving musical identities of professional musicians. For these participants, musical identities were implicated as central to identity construction across a range of social situations. In the next section we move from analysing talk about music to investigating how motivation is related to musical participation and musical identities.

## Musical identity and motivation

In our earlier discussion we cited the example of a young child, going to piano lessons for the first time, who gradually develops a view of herself as a pianist, or more generally as a musician. This development can be encouraged by praise from her parents and wider family about practising, or about particular musical achievements, and by talking to her friends and schoolmates about her progress. The positive feedback that she gets from these significant others in her social world will play an important role in her developing identity as a musician, and will thereby motivate her to higher levels of practice and achievement, which in turn will lead to further strengthening of the 'musician' role: and so a 'virtuous cycle' can develop in which growing levels of musical achievement and musical identity are interdependent. In other words, children's motivation to succeed in music is inextricably linked with aspects of their musical identity: the ways in which they think about their own abilities have a direct influence upon their motivation to engage in activities which develop those abilities, and vice versa. It therefore makes sense to consider how theories of motivation and the self can explain some of the effects of musical identities.

Motivation to learn in music is the topic of Chapter 27 in this volume, by Susan Hallam, and we refer the reader to that chapter: our distinctive emphasis in this section is on how musical identities give rise to certain aspects of motivation. An extensive body of theory and research has been conducted on motivation for educational attainment more generally—in subjects such as mathematics and reading—and this also has been applied to motivation in music learning by O'Neill and McPherson (2002), Hallam (2002) and Austin *et al.* (2006).

Self-theories form a central part of the theoretical models of educational and musical motivation which have emerged in recent decades, and Austin *et al.* (2006) illustrate this point very clearly in proposing a process model of motivation which is based on the work of Connell (1990). The model has four components—the *self system* (e.g. perceptions, beliefs, thoughts, emotions); the *social system* (e.g. teachers, peers, siblings); *actions* (e.g. motivated behaviours including learning and self-regulation); and *outcomes* (e.g. learning, achievements). Each of these four components is seen to have a reciprocal causal relationship with each of the others, such that the motivational system develops and changes as learning proceeds, and as new challenges are sought by the learner. Hallam (Chapter 27) goes further than this, proposing a wide-ranging model of musical motivation which sees 'malleable aspects of the personality and self-concept' feeding, along with various environmental influences, into motivation, interaction with the cognitive characteristics of the individual, and with the different cognitive processes involved.

Most writers in this area agree that expectancy –value theory (e.g., Eccles *et al.* 1983; Pintrich and Schunk 2002) is the most well established and useful theoretical approach, and models based on this approach have three main components. These are *value* components—the extent to which learners view a particular task, activity or domain as being important, and as being of value to them: *expectancy* components—learners' beliefs about their abilities to succeed

in the activity: and *affective* components, namely how learners feel about themselves in relation to the activity. Thus, our young pianist's motivation to progress to higher levels will depend on the extent to which she sees playing the piano as being important in her life, and in having future benefits and pay-offs: on whether or not she perceives she is capable of achieving higher levels of success; and on whether or not she simply enjoys playing the piano in relation to all her other interests and activities.

Eccles *et al.* (1983) proposed that there are four main aspects of the value component, namely its *attainment value* (the importance to the individual of success in the task); its *utility value* (its perceived usefulness to the individual); its *intrinsic interest* (the absorption in and enjoyment of the activity for its own sake); and its *perceived cost* (the loss of time spent on other activities as a result of engagement in musical activities, and the consequent loss of interest in them. Many parents, including the present authors, have faced the 'Saturday morning dilemma'—school sports team or band/orchestra?!). The third of these aspects raises the important distinction between intrinsic and extrinsic motivation: the former is evident when interest in an activity is natural, unforced and high, and the latter occurs when external rewards and punishments (such as praise and criticism) are involved.

Amabile's (1996) social psychological theory of creativity emphasizes the importance of intrinsic motivation, proposing that 'the intrinsically motivated state is conducive to creativity, whereas the extrinsically motivated state is detrimental' (p. 107). The *perceived cost* of engaging in an activity may involve both intrinsic and extrinsic motivation: children may decide that the time and work involved in practising an instrument is not worth the effort, for example, and consciously decide to give up in favour of other pursuits (e.g. Eccles *et al.* 1998). It is also worth pointing out that external rewards do not always lead to improvements in performance: if intrinsic motivation is already high, rewards can have the effect of undermining performance rather than of improving it.

The decision to either persist in music, or to give up, is clearly interdependent with the second, *expectancy* component of the expectancy–value model, which refers to learners' beliefs in their own ability to succeed on the task, and this is what Bandura (1997) calls *self-efficacy*. Learners' prior view of their own potential ability to succeed on a task—their competency beliefs—can be an important determinant of their actual performance on that task, and this leads to some further conceptual distinctions. Dweck (2000) suggests that people differ in the extent to which they display *mastery-oriented* as distinct from *helpless* behaviour, which in turn relates to their *locus of control.* Learners with an *internal* locus of control are likely to display mastery-oriented patterns of behaviour, for example in trying to persist with a task even when they meet difficulties or setbacks. Those with an *external* locus of control, however, feel that circumstances are beyond their own control, and may display helpless patterns of behaviour. In Chapter 27, Hallam draws on these concepts in describing how learners' competency beliefs can be an important determinant not only of whether they engage in that task as well as of their actual performance on that task.

O'Neill (1997b) investigated the longer term effects of these motivational patterns on children who were about to begin instrumental tuition by giving them problem-solving tasks which were arranged such that they were bound to fail. She found that those who displayed mastery-oriented behaviour after failing on the task made more progress after one year of instrumental music lessons than those who initially displayed helpless behaviour. This result leads us to a more general consideration of the third component of expectancy–value theory, namely the *affective* component: people's beliefs about and reactions to success and failure can be explained in terms of attribution theory (e.g. Weiner 1985, 1992). According to this, people attribute their success or failure on a task to four main factors: *ability* ('I am good/poor at this task'): *effort* ('I practised hard/insufficiently'); *task difficulty* ('this particular task was easy/difficult'); and *luck* ('I had a lucky/unlucky day'). The first and second of these are internal in the sense that they are under the control of the learner, and the third and fourth are subject to external influences.

These theoretical analyses of educational motivation, along with Hallam's discussion in Chapter 27, reveal some of the mechanisms by which musical achievement and development can be determined by musical identities. We referred earlier to a hypothetical 'virtuous cycle' according to which increasing levels of musical achievement give rise to higher levels of musical self-esteem, and vice versa; children who feel that they are competent musicians are likely to achieve higher levels of success than those who do not, as are those who find music intrinsically interesting, and value it positively (Eccles *et al.* 1983). This can work in both directions, of course: some children may get the idea that they are 'unmusical', perhaps because of an unwitting remark by a teacher, parent or another pupil, and this perception could correspondingly lead on to a downward spiral of not trying, therefore becoming less able, therefore trying even less, and so on.

The *content* of musical ability self-perceptions were investigated as part of a series of studies conducted within the Curriculum Development Project in the Arts and Music Monitoring Programme, sponsored by the Qualifications and Curriculum Authority (QCA) in England. Hargreaves and colleagues (2004) compared the views of pupils, parents and teachers about 'what it is to be good at music' along with parallel questions about art and design, dance, and drama, from the points of view of pupils, teachers and parents. Focus groups of pupils and parents, and individual teachers, were asked to discuss three questions: (a) what does it mean to be good at music (what does it involve)?; (b) can one 'be good at' some aspects of music and not others?; and (c) how can one get better at music, and how would others know whether or not this had occurred?

Performing, and knowledge about music, were both seen as especially important aspects, particularly by pupils and teachers. Music was seen as comprising a range of different skills and capacities, including non-musical skills such as learning quickly, leading a group, and having confidence. Teachers tended to emphasize the requirements of the English National Curriculum, and to conceive their views of musical ability and achievement in those terms. Parents tended to focus on skills and activities which are visible, notably those involving public performance. All three groups' responses suggested that music is no longer seen as a special talent which individuals either do or do not possess, and that it is perfectly possible to be good at some parts of it and not others. Qualifications and exams assumed lesser importance than might have been expected, and pupils' views were no less detailed and insightful than those of teachers and parents.

The pupils' views were explored in detail in a more recent report by Lamont *et al.* (2007) on a focus group investigation of 134 8–14-year-olds' understandings of what it means to be a musician, and of the differences between their descriptions of musicians in general with their own self-concepts as musicians. The results clearly showed that they defined musicians in terms of playing a musical instrument, making up music, and that closely related to this were being a singer, and being a music teacher. These children also emphasized effort, practice and hard work as being important characteristics of musicians. Their beliefs about themselves as musicians were more varied, however, and were more likely to include factors beyond their control (such as musical ability, and critical periods in their lives). There was a general acknowledgment of the role and importance of music in society, and of the possibility of music as a career and leisure choice. The motivation to be active and to succeed in music seems to be closely bound up with the development of positive musical identity.

## Conclusions

In summary, recent developments within the psychology of music advocate a theoretically eclectic position for academics endeavouring to develop our understanding of the psychology of music (Miell *et al.* 2005). Much of the earlier research utilized a predominately quantitative empirical epistemology which focused on isolating discrete variables and was less concerned with the wider social and cultural influences on music perception (Hargreaves and North 1997; MacDonald *et al.* 2002). However, a more multidisciplinary and pluralistic position is entirely in keeping with postmodern research priorities

within the wider academic community, and so researchers interested in investigating musical identities including psychologists, music therapists, musicologists, music educationalists and others, can adopt quite different theoretical stances on the research processes. The current situation within musical identity research reflects these multidisciplinary features, and we find research that is empirical (using both qualitative and quantitative approaches), and review based.

In this chapter we have focused on key themes and approaches within musical identities research. We have highlighted these issues with examples from published material and from our own current research, and in doing so have concentrated on two broad areas: jazz identities and motivations and identity. These two areas highlight the pluralistic manner in which researchers are currently investigating musical identities. In particular the section above on jazz identities highlights how the way in which musicians talk about their musical experiences helps to construct their musical identities. Related to these observations, the section on motivation draws together recent thinking in this area to highlight the importance of motivational aspects of musical development and also how self-perceptions influence the development of music skills and, through these, the development of musical identities. This chapter highlights the crucial role of wider cultural and social variables in determining not only the extent to which we view ourselves as musical or not, but also the extent to which we engage in musical activities (both listening and playing).

Musical identities are an important consideration for researchers interested in the psychology of music, not least because we all have musical identities, but also because identity 'work' is a fundamental psychological process in which we all engage. This identity work can take the form of musicians reflecting on their own practice and how this practice is related to their lifestyle, or it can be how children view their own sense of musical skill and development. In short, we are suggesting that talking about music and, through this, developing and negotiating musical identities, influence our engagement with music in important ways.

## References


Applegate C and Potter P (2002). *Music and German national identity*. University of Chicago Press, Chicago, IL.

Austin J, Renwick J, and McPherson GE (2006). Developing motivation. In G. E. McPherson (ed), *The child as musician: A handbook of musical development*, 213–238. Oxford: Oxford University Press.

Bandura A (1997). *Self-efficacy: the exercise of control*. W.H. Freeman, New York.

Becker H (2000). The etiquette of improvisation. *Mind, Culture, and Activity*, 7(3), 171–176.

Berliner P (1994). *Thinking in jazz: the infinite art of improvisation*. University of Chicago Press, Chicago, IL.

Borthwick SJ and Davidson JW (2002). Personal identity and music: a family perspective. In RAR MacDonald, DJ Hargreaves and D Miell, eds, *Musical identities*, 60–79. Oxford University Press, Oxford.

Carlton (2006) A qualitative analysis of everyday uses of preferred music across the life span. Paper presented at ICMPC 9, Bologna, Italy, August 2006.

Connell JP (1990). Context, self and action: A motivational analysis of self-system processes across the life-span. In D. Cicchetti & M. Beeghly (eds) *The self in transition: Infancy to childhood*, 61–97. Chicago: University of Chicago Press.

Cooke M and Horne D (2002). *The Cambridge companion to jazz*. Cambridge University Press, Cambridge.

DeNora T (2000). *Music in everyday life*. Cambridge University Press, Cambridge.

Dibben N (2002). Gender identity and music. In RAR MacDonald, DJ Hargreaves and D Miell, eds, *Musical identities*, 117–134. Oxford University Press, Oxford.

Dweck CS (2000). *Self-theories: their role in motivation, personality and development*. Psychology Press, Hove, Sussex.

Eccles JS, Adler TF, Futterman R, Goff SB, Kaczala CM, Meece JL and Midgley C (1983). Expectancy, values, and academic behaviours. In JT Spence, ed., *Achievement and achievement motives: psychological and sociological approaches*, 75–146. W.H. Freeman, San Francisco.

Eccles JS, Wigfield A and Scheinfele U. (1998). Motivation to succeed. In W Damon and N Eisenberg, eds, *Handbook of child psychology: vol 4. Social, emotional and personality development*, 5th edn, 215–225. Wiley, New York.

Edwards D and Potter J (1992). *Discursive psychology*. Sage, London.

Giddens A (1991). *Modernity and self-identity. Self and society in the late modern age*. Polity Press, Cambridge.

Gilbert N and Mulkay (1984). *Opening Pandora's box: a sociological analysis of scientists' discourse*. Cambridge University Press, Cambridge.

Green L (1997) *Music, gender, education*. Cambridge University Press, Cambridge.

Green L (2002) *How popular musicians learn*. Ashgate, Aldershot.

Hallam, S (2002). *Musical Motivation: Towards a Model Synthesising the Research Music Eduction Research*, **4**(2), 225–244.

Hargreaves DJ, Lamont A, Marshall N and Tarrant M (2004). What is 'being good at music'? *NAME (National Association of Music Educators) Magazine*, **13**, 4–7.

Hargreaves DJ & North AC (eds) (1997). *The social psychology of music*. Oxford: Oxford University Press.

Juslin PN and Sloboda JA (eds) (2001). *Music and emotion: theory and research*. Oxford University Press, New York.

Lamont A (2002). Musical identities and the school environment. In RAR MacDonald, DJ Hargreaves and D Miell, eds), *Musical identities*, 49–55. Oxford University Press, Oxford.

Lamont A, Hargreaves DJ, Marshall N and Tarrant M. (2007). *Musical identities at school*. (Submitted).

MacDonald RAR and Miell D (2000). Creativity and music education: the impact of social variables. *International Journal of Music Education*, **36**, 58–68.

MacDonald RAR and Wilson GB (2006). Constructions of jazz: how jazz musicians present their collaborative musical practice. *Musicae Scientiae*, **10**(1), 59–85.

MacDonald RAR and Wilson GB (2005). The musical identities of professional jazz musicians: a focus group investigation. *Psychology of Music*, **33**(4), 395–419.

MacDonald RAR, Hargreaves DJ and Miell D (2004). The sounds of ideologies clashing. *Action, Criticism and Theory for Music Education*, **3**(1), 3–9. http://www.siue.edu/MUSIC/ACTPAPERS/v3/EditorResponse04.pdf.

MacDonald RAR, Hargreaves DJ and Miell D (eds) (2002). *Musical identities*. Oxford University Press, Oxford.

McClary S (1991) *Feminine endings: music, gender and sexuality*. University of Minnesota Press, Minneapolis, MN.

Meadows ES (2003) *Bebop to cool: context, ideology*, and musical identity. Greenwood, Westport CT.

Miell D, MacDonald RAR and Hargreaves DJ (eds) (2005). *Musical communication*. Oxford University Press, Oxford.

Monson I (1996). *Saying something: jazz improvisation and interaction*. University of Chicago, Chicago, IL.

O'Neill S (1997a) Gender and music. In D J Hargreaves and AC North, eds, *The social psychology of music*, 46–60. Oxford University Press, Oxford.

O'Neill S (1997b). The role of practice in children's early musical performance achievement. In H Jørgensen and AC Lehmann, eds, *Does practice make perfect? Current theory and research on instrumental practice*, 53 Norges Musikhøgskole, Oslo.

O'Neill S (2002) The self-identity of young musicians. In RAR MacDonald, DJ Hargreaves and D Miell, eds), *Musical identities*, 79–97. Oxford University Press, Oxford.

O'Neill S and McPherson GE (2002). Motivation. In R Parncutt & GE McPherson, eds, *The science and psychology of musical performance: Creative strategies for music teaching and learning*, 31–46. Oxford: Oxford University Press

Pintrich PR and Schunk DH, (2002). *Motivation in education: theory, research, and applications*, 2nd edn. Merrill, Upper Saddle River, NJ.

Roe K (1999) Music and identity among European youth: music as communication. Soundscape.info, *Journal on Media Culture*, **2**, http://www.icce.rug.nl/~soundscapes/HEADER/colophon.shtml.

Sawyer K (1992). Improvisational creativity: an analysis of jazz performance. *Creativity Research Journal*, **5**(3), 253–263.

Sawyer RK (2003) *Group creativity: music*. Erlbaum, Theater, Collaboration, Mahwah, NJ.

Stålhammar B, (2006) *Musical identities and music education*. Shaker Verlag, Aachen.

Tarrant M, North AC and Hargreaves DJ (2002). Youth identity and music. RAR MacDonald, DJ Hargreaves and D Miell, eds), *Musical identities*, 134–150. Oxford University Press, Oxford.

Trevarthen C (2002) Origins of musical identity: evidence from infancy for musical social awareness. In RAR MacDonald, DJ Hargreaves and DE Miell, eds, *Musical identities*, 22–46. Oxford University, Oxford.

Weiner B (1985). An attributional theory perspective of achievement motivation and emotion. *Psychological Review*, **92**, 548–573.

Weiner B (1992). *Human motivation: metaphors, theories, and research*. Sage, Newbury Park, CA.

Welch G (2006) Singing and vocal development. In GE McPherson, ed., *The child as musician: a handbook of music development*, 311–331. Oxford University Press, Oxford.

Wilson GB and MacDonald RAR (2005). The meaning of the blues: musical identities in talk about jazz. *Qualitative Research in Psychology*, **2**, 341–363.

Zillman D and Gan S (1997) Musical taste in adolescence. In DJ Hargreaves and AC North, eds, *The social psychology of music*, 161–188. Oxford University Press, Oxford.

CHAPTER 44

# The effects of music in community and educational settings

Susan Hallam and Raymond MacDonald

## Introduction

It has been recognized through the ages that music can have a very powerful influence on our emotions, moods and behaviour. In recent years technology has led to easy access to music listening, and individuals and organizations have taken advantage of this to utilize music to influence how we feel and act in many areas of our lives (see Chapters 40 and 45). Young people in particular spend a great deal of time playing music and often use it to accompany their studying. Alongside the increase in and ease of music listening, there has been an acknowledgement that there are wider benefits to active engagement with music beyond those emanating from its value as an art form. In the Western world, the latter part of twentieth century saw a huge growth in opportunities for making music within formal educational settings and more recently, through what is termed 'community music', playing, teaching and learning music outside of formal education (Elliot 2004). Research assessing the impact of such engagement has been used by music educators and community musicians to support the continued funding of such activities by policy makers.

In this chapter we will review the literature which has explored the ways that music may have benefits beyond those associated with our enjoyment of active engagement or listening in community and educational settings. We will outline developments in community music, consider research which has assessed its impact, explore the way that music has been used to attain educational aims beyond those relating to music education itself, and consider issues and findings relating to the use of background music when studying.

## The impact of active engagement with making music in community and educational settings

In general, individuals actively engage with making music because they enjoy it, although where it is a compulsory element of the school curriculum this may not be the case. In recent years, perhaps in part because of the need to justify the place of music in the school curriculum and the continuance of funding for informal community music-making opportunities, there has been an increasing focus on the wider benefits of active engagement in music.

### The impact of music in the community

Over the past 15 years there has been an explosion of interest in music making outside of formal educational settings. Much of this has come to be termed 'Community Music' and it is now an established area of research and practice with a dedicated journal, *The International Journal of*

*Community Music* and a commission of The International Society for Music Education. Veblen (2004) presented an overview of the many different ways in which community music is practiced around the world which highlighted the key concern of community music practitioners to increase access to music-making for all members of the public. The growing recognition of community music as an important musical activity is very much in keeping with current ideas of multidisciplinary practices that signal the integration of social, educational, medical, and therapeutic practices and thinking (Ruud 2004).

The growth of community education has in part been an acknowledgement that every human being has a social and biological guarantee of musicianship (Trevarthen 2002; MacDonald *et al.* 2002) and evidence that everybody, regardless of social, educational, psychological or medical aspects can communicate through music (Elliot 2004; Ansdell and Pavlicevic 2004; Miell *et al.* 2005). Musical skill develops through opportunities for musical engagement in childhood, adolescence and early adulthood, the influence of family, school and wider cultural factors being crucial in the individual's developing musicality (Borthwick and Davidson 2002; Lamont 2002; MacPherson in press; see also Chapters 22, 27, 28, and 29). Adult participation in community music is an extension of active music-making in childhood (e.g. Conda 1997), the pattern of engagement changing over the life course, diminishing in the middle years and increasing in retirement (Larson 1983), life-changing events sometimes providing an impetus for re-engagement (Conda 1997). The reasons for adults' participation in music have been grouped into three main categories: personal motivations (self-expression, recreation, self-improvement, and use of leisure time); musical motivations (love of music, performing for oneself and others, learning more about music); and social motivations (meeting new people, being with friends, and having a sense of belonging). No single reason consistently emerges as the most important (Coffman 2002). Hinkle (1988) suggests a further category—spirituality.

There is a growing body of evidence that involvement in music-making in community contexts has positive psychological or social benefits for the participants (Davidson 2004). While it is important to distinguish community music from music therapy (see Chapters 46 and 47), much has been written about the parallels and contrasts between community music-making, music education and music therapy. In particular Ansdell and Pavlicevic's edited text *Community music therapy* (2004) outlines how recent advances in community music-making and music therapy can facilitate a greater cross-fertilization of ideas between these areas of practice. Ockelford (2000) also discusses the need for greater communication between practitioners and researchers involved in music education and music therapy in order to develop our knowledge of the kinds of effects music may have.

The benefits of adult participation in music, in a professional or amateur capacity, are many and varied. Choral singing seems to have positive emotional, social, physical, and creative outcomes (Beck *et al.* 2000) which may also be spiritual (Clift and Hancox 2001). Bailey and Davidson (2002), researching members of choirs for homeless men and middle class participants, found that group singing and performance even at low levels of musicality yielded considerable emotional, social and cognitive benefits to participants. The emotional effects were similar regardless of training or socio-economic status, but interpersonal and cognitive components had different meanings for marginalized and middle-class singers. The marginalized individuals appeared to embrace all aspects of the group singing experience, while the middle class choristers were inhibited by prevalent social expectations of musicianship. Social and musical components appear to contribute equally to commitment to and enjoyment of membership of amateur operatic societies, the musical and acting components providing challenge, achievement, and opportunity to escape temporarily the frustrations of everyday living (Pitts 2004).

Community music interventions can provide senior citizens with a venue to make friends and encourage participation in intergenerational music activities between children, teenagers, and senior citizens, all of which have considerable social value (Bowers 1998; Darrow *et al.* 1994).

Music, through community education programmes, can also help adults move toward socio-political transformation (Kaltoft 1990). Engaging with music has a range of positive personal and social outcomes which do not seem to diminish with age. In older people music assists in developing self-identity, connecting with others, maintaining well-being, and experiencing and expressing spirituality, providing ways for people to maintain positive esteem, feel competent, independent, and avoid feelings of isolation or loneliness (Hays and Minichiello 2005). Although older adults require more time to encode and select information and have slower reaction times these can be minimized through task familiarity, sufficient structuring of information and regular exercise (Cavanaugh 1997). Primary mental abilities decline with age but higher-order mental activities do not, and adults may maintain or even increase mental abilities in areas of expertise including music (Pieters 1996).

There are direct health benefits for active musical participation. Choral performance increases secretory immunoglobulin A and reduces cortisol levels of professional (Beck *et al.* 2000) and amateur (Kreutz *et al.* 2004) choristers before and after singing. Playing the piano exercises the heart as much as a brisk walk (Parr 1985) and music-making contributes to perceived good health, quality of life, and mental well-being (Coffman and Adamek 1999; Vanderark *et al.* 1983).

Community music-making for adults with learning difficulties can facilitate both musical and psychological developments. MacDonald and colleagues (1999) reported that 12 weeks of community-based music workshops focused on playing a Javanese gamelan produced significant improvements in musical ability, communication skills and self-perception of musical ability. The improvement in communication skills may be linked to development in joint attention (O'Donnell *et al.* 1999). This is defined as a shared focus of attention on the same object by two individuals and is similar to the concept of 'shared social reality' (Nameera *et al.* 2007). MacDonald and Miell (2002) highlighted various ways in which community music-making was perceived by participants with special needs to be of benefit to them. The participants provided detailed accounts of how involvement was empowering and indicated that other people's expectations of their abilities changed positively. Important for them was that the music-making they were involved in was perceived to be of a professional standard. Other research has shown how improvisation workshops, percussion workshops, instrumental lessons, and singing workshops delivered to a variety of groups of adults with dementia, and adolescents with mental health disorders, can have beneficial effects on particular aspects of cognitive function (De Simone and MacDonald 2007).

## The wider benefits of music in formal education

The impact of active engagement with music in the classroom is usually explained through transfer effects (see Chapter 23), the emphasis frequently being on intellectual rather than personal and social outcomes, although there has been some research considering the latter. One strand of research has focused on the impact of different types of music tuition, e.g. Dalcroze, Orff, Kodaly, while others have explored the way that music can be used to directly enhance learning materials.

### Effects of particular teaching approaches

Over time, there has been considerable interest in the impact of the Kodaly method of teaching music on children's intellectual development. Positive effects have been found on temporal and spatial abilities (Hurwtiz *et al.* 1975), reading and arithmetic (Gardiner *et al.* 1996) and IQ (Schellenberg 2004), while activities involving music and movement have been found to have a positive impact on a range of outcomes including delayed development (Moore 1984) and motor performance (Zachopoulou *et al.* 2004) in comparison with controls (see also Chapter 23).

### The role of music in enhancing language and literacy skills

Some researchers have focused on the impact of more general musical activities on language and literacy skills (e.g. McCarthy 1985). Using rhyme, rhythm and repetition has been shown to facilitate the learning of vocabulary (Baechtold

and Algier 1986), particularly in remedial readers (Newsom 1979), and has subsequently led to improvement in reading, comprehension, attitudes and motivation. Musical activities have also been shown to assist in developing second language learning in children (e.g. Kennedy and Scott 2005) and in improving memory for a range of other materials including multiplication tables (Clauson and Thaut 1997), and stories (e.g. Allen and Butler 1996).

The direct impact of the use of music to improve literacy through the use of songs and singing has not been clearly demonstrated (e.g. Ebisutani *et al.* 1991), although music lessons designed to develop auditory, visual and motor skills have benefited reading (Douglas and Willats 1994), as have studies aimed at improving rhythmic skills (Long 2007), and there is evidence that there is a relationship between phonemic awareness, simply reading ability and pitch discrimination (e.g. Anvari *et al.* 2002). Early skill with music might enhance reading acquisition to the extent that reading depends on the same basic auditory analysis skills.

### Social and personal learning outcomes

In formal education, increasing the number of classroom music lessons has been shown to have a positive impact on social behaviour. In a large-scale European study, children receiving additional music lessons demonstrated increased social cohesion within class, greater self-reliance, better social adjustment and more positive attitudes. These effects were particularly marked in low ability, disaffected pupils (Spychiger *et al.* 1993). In the UK, Harland and colleagues (2000) showed that the most frequent overall influences of engagement with music on pupils were related to their personal and social development, in particular confidence and self-esteem.

Music has also been used to change problem behaviours in particular groups of students. DeCarlo (2001) used rap music, with some success, with young urban African American adolescents with behaviour problems to promote the development of appropriate social skills related to morality, identity, judgement, decision-making, anger management, impulse control, and crime and punishment. Also in the USA, after school music programmes have been shown to be especially beneficial for reducing the likelihood of disaffection in at-risk children (e.g. Otterbourg 2000).

## The impact of background music in educational settings

### Issues in studying the impact of background music

There are a number of methodological and theoretical issues which have bedevilled research exploring the impact of background music on learning and behaviour in educational settings: finding robust ways of categorizing the nature of the music itself; considering how to systematically measure its impact; and how to take account of individual differences in response.

Early research tended to ignore the characteristics of the particular music being played, assuming that all music would have a similar impact. More recent research has attempted to address this issue by differentiating music on the basis of genre, its perceived potential to stimulate or relax, whether it is vocal or instrumental, and in relation to its cognitive complexity. While all of these approaches have some merit, none have been able to capture the complexity of music as it occurs in real time. This particularly applies to Western classical music with its frequent changes of mood, tempo, timbre, complexity, and volume.

The second key methodological issue relates to the way that impact is assessed. A range of methods has been adopted including physiological and neurological measures, observation of behaviour, assessment of task performance, and the use of rating scales. However, the relationships between outcomes assessed by these various measures are complex. While arousing music tends to lead to an increase in a variety of physiological responses in children (e.g. Furman 1978) this is not always the case with adults. For instance, Iwanaga and colleagues (1996) found that university students reported increased excitation when listening to the Sacrifice Dance from the Rite of Spring but their heart rate and respiration rates did not change, although

calming effects reported accompanying listening to Satie's Gymnopedie No. 1 were matched with a decrease in heart rate and respiration. Favourite music, whether stimulating or relaxing, also tends to lower subjective tension without necessarily having a similar impact on physiological responses (Iwanaga and Moroki 1999).

The third methodological issue relates to the subjective aspects of music perception. Individuals respond to the same music in very different ways, depending on their musical preferences and their individual characteristics. The structural features of the music (e.g. tempo, modality, instrumentation, genre); cultural factors (aspects of our environment including tonality and the way that musical associations are culturally shaped and learned); and associative factors (personal and subjective meanings we place on a particular piece of music depending on our musical experiences) all play a part in our responses to music. Where associative factors come into play, the structural and cultural aspects of the music are superseded by the personal and associative aspects (see Figure 44.1). Preference may therefore render very different types of music *functionally equivalent.* For example, the music which young people may chose to play while studying may differ widely but lead to similar physiological effects.

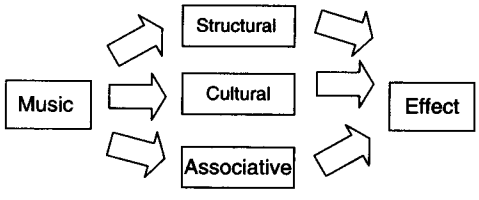



**Fig. 44.1** Influences on music perception.

Explanations for the differential impact of background music include those invoking music's ability to affect arousal and mood, interfere with processing, and reduce distraction. The Yerkes–Dodson law provides one explanation stating that the arousal level of the individual increases performance up to an optimal level beyond which over-arousal leads to deterioration. This occurs more quickly when the task to be performed is complex or under-learned. Completing a simple task requires a higher level of arousal for concentration to be maintained, complex tasks require lower arousal levels (Konecni and Sargent-Pollock 1976). Personality factors are also implicated in optimal arousal levels. Introverts have higher resting levels of arousal than extroverts and are more susceptible to over arousal. Closely related to arousal theories are those related to mood and emotions, there being considerable evidence that music can affect mood (e.g. Niedenthal and Setterland 1994) and alertness (Sousou 1997).

An alternative explanation for the impact of background music is that it reduces the cognitive processing capacity available for other material. Concentrated listening to music requires cognitive effort for processing, analysis and extracting meaning (Berlyne 1971). Listening to complex, arousal-evoking music may therefore reduce the attentional space available for task performance (e.g. Broadbent 1971). However, when people play background music they do not attend to both the music and the task simultaneously. Attention switches between the two (Madsen 1987). Depending on their interest, their focus may be greater on the task or the music. A further complication is the role of meta-cognition—participants being aware that the music is interfering with task performance and consciously adopting strategies to prevent this. This may explain age-related differences in the actual and perceived impact of music on performance (Kaniel and Aram 1998; Kotsopoulou and Hallam 2006).

## The impact of background music on behaviour and task performance

Research on background music in educational contexts has explored the impact on children at different ages, with special educational needs, and undertaking a range of different tasks. For instance, in young children there is evidence that arousing music increases activity (Reiber 1965) and interactive play (Gunsberg 1991), while in older children calming music can enhance reported altruistic behaviour (Hallam *et al.* 2002). Calming music has also been found to have a positive impact on the behaviour of children with emotional and behavioural difficulties, reducing their stress and anxiety in a

variety of settings (Hallam and Price 1998; Savan 1998). Sedative music can result in a greater increase in activity level in children with learning difficulties than stimulating music, the latter replacing the children's need for activity and self-stimulation. These differences in response mean that music interventions aimed at changing behaviour need to be tailored to the requirements of specific groups of children (Reardon and Bell 1970).

Much of the research on the impact of background music has lacked a clear theoretical basis for predicting outcomes, and findings often appear to be contradictory. For instance, in research relating to memorization, when music has been played concurrently with material to be remembered aurally it has not facilitated memory (Furnman 1978), paired associate recall (Myers 1979), or phonological short-term memory (Salame and Baddeley 1989), but when recall is of written sentences presented visually primary-aged children perform better when calming music is playing as opposed to silence (Hallam *et al.* 2002). Music may interfere with aural presentation and recall because of shared cognitive structures reducing processing capacity. Written presentation does not interfere in this way so arousal effects come into play. Similarly, where background music has words it may have a greater negative impact on reading comprehension and other literacy tasks (Martin *et al.* 1988).

Much of the research exploring the impact on tasks in other domains, for instance, mathematics, can be explained through the impact of music on arousal levels (e.g. Mayfield and Moss 1989), although other factors may also be important, for instance, the extent to which learners are used to working with music playing in the background (e.g. Etaugh and Michal 1975), the familiarity of the particular music being played (Hilliard and Tolin 1979), or whether the music is vocal or instrumental (Donlan 1976). Well-learned skills are less affected by high levels of arousal than those requiring creativity or high levels of concentration (Hallam and Godwin 2000) and older students may also be better able to adopt strategies which minimize the negative impact of the music on task outcomes (Kotsopoulou and Hallam 2006). Overall, the impact of background music on performance on any particular task depends on many interacting factors.

### A framework for future research

Figure 44.2 sets out a model of possible contributory factors including the nature of the music itself: its genre; whether it is stimulating or relaxing; its complexity; whether it is familiar; liked; and has been selected by the individual listening to it or imposed by others. The model suggests that the effects of music are mediated by the characteristics of the individual; their age, ability; personality; meta-cognitive skills; musical expertise; familiarity with the music being played; and the frequency with which they normally listen to music when they are studying. The current emotional arousal and mood state of the individual may also be influenced by individual characteristics and recent life events. Individual characteristics also have a direct effect on learning outcomes and a further indirect effect through meta-cognitive activity. The environment within which the activity is taking place may also be important, for instance, whether the individual is alone, in a familiar place, whether there are other distractions. The characteristics of the task, for instance, the nature of the processing required, its difficulty, whether it is perceived as interesting or boring will also play a part. Currently little research takes account of these factors.

## Issues for future research

Many similar factors are important in relation to assessing the impact of active music-making as in assessing the impact of background music, for instance, taking account of the characteristics of the individual, the nature of the music itself, and the context within which the activity takes place. However, there are important differences. Crucially, when making music the focus is the musical task, which of itself may be rewarding. In addition, playing background music to support non-musical task completion has an immediate, time-limited impact restricted to that task which is unlikely to bring about long-term change in the individual. In contrast, active engagement with music making may have a longer-term impact on the individual's

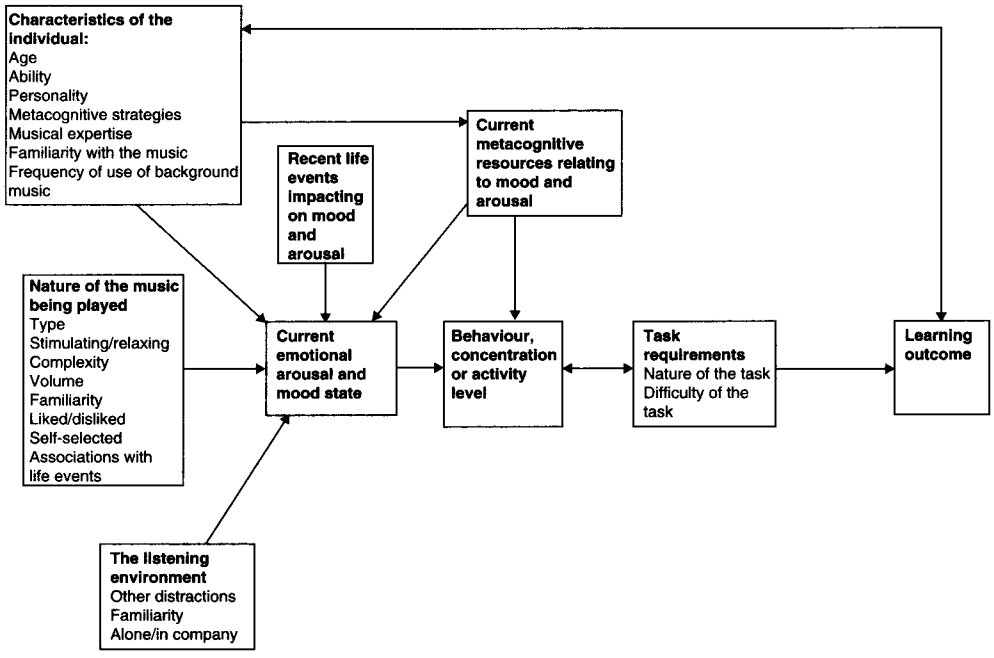



**Fig. 44.2** Model of the effects of background music and learning (derived from Hallam and Goodwin 2000).

intellectual, personal, social and spiritual development. A range of different issues therefore need to be considered by researchers including:

- over what timescale does the impact occur?
- how long is it maintained?
- does participation have to be voluntary for there to be a positive impact?
- to what extent is the musical activity as opposed to any concurrent social activity important in determining the impact?

## Conclusions

In the modern world all kinds of music are available to most people 24 hours a day. It pervades our everyday lives and influences our behaviour. Children are exposed to it at home and in the wider community. In contrast to this, in schools, music is rarely used to enhance behaviour and subsequently learning, although it can have powerful effects and create an environment which is more conducive to on-task behaviour. It may be that educators feel that it is inappropriate to use music in this way or that it may interfere with the learning of some students, even though it may benefit others. Music can also support learning in other subjects through transfer effects and the way in which it can provide additional cues for remembering. This may be particularly beneficial for young children.

There are now more opportunities than ever before for all sections of society to engage in organized music-making through the growth of community music interventions around the world including signing groups for older adults, percussion workshops for disadvantaged young people, gamelan groups for individuals with special needs, and song-writing classes. There is growing evidence to suggest that these community music interventions not only help people develop musical skills but that they also are psychologically beneficial in a variety of ways.

However, there is a need for more systematic research in community and formal education contexts which establishes the conditions under which music can enhance a variety of learning, intellectual, personal, and social outcomes. Such research would be invaluable in helping to develop our knowledge of the processes and

outcomes of music listening and participation, increasing our understanding of the ways in which music impacts on learning and behaviour, and enabling us to guide musicians working in education and the community as to how to maximize the positive benefits.

## References


Allen BA and Butler L (1996). The effects of music and movement opportunity on the analogical reasoning performance of African American and white school children: a preliminary study. *Journal of Black Psychology*, **22**, 316–328.

Anvari SH and Trainor LJ, Woodside J and Levy BA, (2002). Relations among musical skills, phonological processing and early reading ability in preschool children. *Journal of Experimental Child Psychology*, **83**, 111–130.

Ansdell G and Pavlicevic M (2004). *Community music therapy*. Jessica Kingsley, London.

Baechtold S. and Algier A (1986). Teaching college students vocabulary with rhyme, rhythm and ritzy characters. *Journal of Reading*, **30**, 240–253.

Bailey BA and Davidson JW (2002). Adaptive characteristics of group singing: perceptions from members of a choir for homeless men. *Musicae Scientiae*, **6**, 221–256.

Beck R, Cesario T, Yousefi S and Enamoto H (2000). Choral singing, performance perception and immune system changes in salivary immunoglobulin and cortisol. *Music Perception*, **18**, 87–106.

Berlyne DE (1971). *Aesthetics and psychobiology*. New York, Appleton-Century Crofts.

Borthwick SJ and Davidson JW (2002). Personal identity and music: a family perspective. In R MacDonald, D Miell and DJ Hargreaves, eds, *Musical identities*, 60–78. Oxford University Press, Oxford.

Bowers J (1998). Effects of an intergenerational choir for community-based seniors and college students on age-related attitudes. *Journal of Music Therapy*, **35**, 2–18.

Broadbent D (1971). *Decision and stress*. Academic Press, London.

Cavanaugh JC (1997). *Adult development and aging*, 3rd edn. Brooks/Cole, Pacific Grove, CA.

Clauson DW and Thaut MH (1997). Music as a mnemonic device for children with learning disabilities. *Canadian Journal of Music Therapy*, **5**, 55–66.

Clift S and Hancox G (2001). The perceived benefits of singing: findings from preliminary surveys of a university college choral society. *The Journal of the Royal Society for the Promotion of Health*, **121**, 248–256.

Coffman DD (2002). Adult education. In R Colwell and C Richardson, eds, *The new handbook of research on music teaching and learning*, 199–209. Oxford University Press, Oxford.

Coffman DD and Adamek M (1999). The contribution of wind band participation to quality of life of senior adult band members. *Dialogue in Instrumental Music Education*, **20**, 25–34.

Conda JM (1997). The late bloomers piano club: a case study of a group in progress. Unpublished Doctoral Dissertation, University of Oklahoma, 1997. Dissertation Abstracts International, 58, 409A.

Darrow AA, Johnson CM and Ollenberger T (1994). The effect of participation in an intergenerational choir on teens' and older persons' cross-age attitudes. *Journal of Music Therapy*, **31**, 119–134.

Davidson JW (2004). What social psychology of music offers music therapy. In G Ansdell and M Pavlicevic, eds, *Community music therapy*, 47–68. Jessica Kingsley, London

DeCarlo A (2001). Rap therapy? An innovative approach to groupwork with urban adolescents. *Journal of Intergroup Relations*, **27**, 40–48.

De Simone J and MacDonald RAR (2007). An investigation of the effects of a structured music programme on cognition in the treatment of patients with dementia. Paper presented at The 7th European Music Therapy Congress, Eindhoven, August 2007.

Donlan D (1976) The effect of four types of music on spontaneous writings of high school students. *Research in the Teaching of English*, **10**, 116–126.

Douglas S and Willats P (1994). The relationship between musical ability and literacy skills. *Journal of Research in Reading*, **17**, 99–107.

Ebisutani K, Donlan D and Siebers E (1991). The effects of music on reading, oral language, and writing abilities: a review of literature. ERIC database document 333 356.

Etaugh C and Michal D (1975). Effects on reading comprehension or preferred music and frequency of studying to music. *Perceptual and Motor Skills*, **41**, 553–554.

Elliott D (2004). Editorial. *The International Journal of Community Music*, **1**, 1.

Furman CE (1978). The effect of musical stimuli on the brainwave production of children. *Journal of Music Therapy*, **15**, 108–117.

Gardiner MF, Fox A, Knowles F and Jeffrey D (1996). Learning improved by arts training. *Nature*, **381**, 284.

Gunsberg AS (1991). Play as improvisation: the benefits of music for developmentally delayed young children's social play. *Early Child Development and Care*, **66**, 85–91.

Hallam S and Godwin C (2000). The effects of background music on primary school pupils' performance on a writing task. Paper presented at the annual conference of the British Educational Research Association, University of Wales, Cardiff, 7–9 September, 2000.

Hallam S and Price J (1998). Can the use of background music improve the behaviour and academic performance of children with emotional and behavioural difficulties? *British Journal of Special Education*, **25**, 87–90.

Hallam S, Price J and Katsarou K (2002). The effects of background music on primary school pupils' task performance. *Educational Studies*, **28**, 111–122.

Harland J, Kinder K, Lord P, Stott A, Schagen I and Haynes J (2000). *Arts education in secondary schools: effects and effectiveness*. NFER/The Arts Council of England, RSA, London.

Hays T and Minichiello V (2005). The meaning of music in the lives of older people: a qualitative study. *Psychology of Music*, **33**, 437–452.

Hilliard OM and Tolin A (1979). Effect of familiarity with background music on performance of simple and difficult reading comprehension tasks. *Perceptual and Motor Skills*, **49**, 713–714.

Hinkle LB (1988). The meaning of choral experience to the adult membership of the German singing societies comprising the United Singers Federation of Pennsylvania. Unpublished doctoral dissertation, Pennsylvania State University, 1987. Dissertation Abstracts International, 48, 2568A.

Hurwitz I, Wolff PH, Bortnick BD and Kokas K (1975). Non-musical effects of the Kodaly music curriculum in primary grade children. *Journal of Learning Disabilities*, **8**, 45–52.

Iwanaga M and Moroki Y (1999). Subjective and physiological responses to music stimuli controlled over activity and preference. *Journal of Music Therapy*, **xxxvi**, 26–38.

Iwanaga M, Ikeda M and Iwaki T (1996). The effects of repetitive exposure to music on subjective and physiological responses. *Journal of Music Therapy*, **xxxiii**, 219–230.

Kaltoft G (1990). Music and emancipatory learning in three community education programs. *Dissertation Abstracts International*, **51**(7), 2239A. (UMI No. AAT90–33861.).

Kaniel S and Aram D (1998). Influence of task difficulty and background music on working memory activity: developmental considerations. *Alberta Journal of Educational Research*, **44**, 342–345.

Kennedy R and Scott A (2005). A pilot study: the effects of music therapy interventions on middle school students' ESL skills. *Journal of Music Therapy*, **42**, 244–261.

Konecni VJ and Sargent-Pollock D (1976). Choice between melodies differing in complexity under divided-attention conditions. *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 347–356.

Kotsopoulou A and Hallam S (2006). Age differences in listening to music while studying. In M Baroni, AR Addessi, R Caterina and M Costa, eds, *Proceedings of the 9th International Conference on Music Perception and Cognition*. University of Bologna, 22–26 August 2006. Available at: http://www.icmpc2006.org.

Kreutz G, Bongard S, Rohrmann S, Hodapp V and Grebe D (2004). Effects of choir singing or listening on secretory immunoglobulin A, cortisol, and emotional state. *Journal of Behavioural Medicine*, **27**, 623–635.

Lamont A (2002). Musical identities and the school environment. In RAR MacDonald, DJ Hargreaves, D Miell, eds, *Musical identities*, 49–55. Oxford University Press, Oxford.

Larson P (1983). An exploratory study of lifelong musical interest and activity: case studies of twelve retired adults. Unpublished Doctoral dissertation, Temple University, 1982. *Dissertation Abstracts International*, **44**, 100A.

Long M (2007). The effect of a music intervention on the temporal organisation of reacing skills. Unpublished PhD Thesis, Institute of Education, University of London.

MacDonald RAR, JB Davies and O'Donnell PJ (1999). Structured music workshops for individuals with learning difficulty: an empirical investigation. *Journal of Applied Research in Intellectual Disabilities*, **12**, 225–241.

MacDonald RAR, Hargreaves DJ and Miell D (eds) (2002). *Musical identities*. Oxford University Press, Oxford.

MacDonald RAR and Miell D (2002). Music for individuals with special needs: a catalyst for developments in identity, communication and musical ability. In RAR MacDonald, DJ Hargreaves and DE Miell, eds, *Musical identities*, 163–179. Oxford University, Oxford.

Madsen CK (1987). Background music: competition for focus of attention. In CK Madsen, CA Prickett, eds, *Applications of research in music behaviour*, 315–325. The University of Alabama Press, Tucaloosa, AL.

Martin RC, Wogalter MS and Forlano JG (1988). Reading comprehension in the presence of unattended speech and music. *Journal of Memory and Language*, **27**, 382–398.

Mayfield C and Moss S (1989). Effect of music tempo on task performance. *Psychological Reports*, **65**, 1283–1290.

McCarthy WG (1985). Promoting language development through music. *Academic Therapy*, **21**, 237–242.

McPherson GE (in press). The role of parents in children's musical development. *Psychology of Music*

Miell D, MacDonald RAR and Hargreaves DJ (eds) (2005). *Musical communication*. Oxford University Press, Oxford.

Moore J (1984). Rhythm and movement: an objective analysis of their association with music aptitude. *Dissertation Abstracts International*, **45**, 1328 A.

Myers EG (1979) The effect of music on retention in a paired-associate task with EMR children. *Journal of Music Therapy*, **16**, 190–198.

Nameera A, Morton A and Gernsbacher R (2007). Joint attention and vocabulary development: a critical look. *Language and Linguistics Compass*, **1**, 195–207.

Newsom SD (1979). Rock 'n' roll 'n' reading. *Journal of Reading*, **22**, 726–730.

Niedenthal PM and Setterland MB (1994). Emotion congruence in perception. *Personality and Social Psychology Bulletin*, **20**, 401–411.

Ockelford A (2000). Music in the education of children with severe or profound difficulties: issues in current UK provision, a new conceptual framework and proposals for research. *Psychology of Music*, **28**, 197–218.

O'Donnell PJ, MacDonald RAR and Davies JB (1999). Video analysis of the effects of structured music workshops for individuals with learning difficulties. In D Erdonmez and RR Pratt, eds, *Music therapy and*

*music medicine: expanding horizons*, 219–228. MMB Music, Saint Louis, MO.

Otterbourg SD (2000). *How the arts can enhance after-school programs.* National Endowment for the Arts (ERIC Document Reproduction Service No ED446829), Washington, DC.

Parr SM (1985). The effects of graduated exercise at the piano on the pianist's cardiac output, forearm blood flow, heart rate, and blood pressure. *Dissertation Abstracts International*, **46**(6), 1436A. (UMI No. AAT85–18673.).

Pieters JM (1996). Psychology of adult education. In AC Tuijnman, ed., *International encyclopedia of adult education and training*, 2nd edn, 150–158. Elsevier Science, New York.

Pitts SE (2004). Everybody wants to be Pavarotti: the experience of music for performers and audience at a Gilbert and Sullivan festival. *Journal of the Royal Musical Association*, **129**, 149–167.

Reardon DM and Bell G (1970). Effects of sedative and stimulative music on activity levels of severely retarded boys. *American Journal of Mental Deficiency*, **75**, 156–159.

Reiber M (1965). The effect of music on the level of activity of children. *Psychonomic Science*, **3**, 325–326.

Ruud E (2004). Introduction: the ripple effect. In G Ansdell, M Pavlicevic, eds, *Community music therapy*, i–vi. Jessica Kingsley, London.

Salame P and Baddeley A (1989). Effects of background music on phonological short term memory. *The Quarterly Journal of Experimental Psychology*, **41**, 107–122.

Savan A (1998). A study of the effect of background music on the behaviour and physiological responses of children with special educational needs. *The Psychology of Education Review*, **22**, 32–36.

Schellenberg EG (2004). Music lessons enhance IQ. *Psychological Science*, **15**, 511–514.

Sousou SD (1997). Effects of melody and lyrics on mood and memory. *Perceptual and Motor Skills*, **85**, 31–40.

Spychiger M, Patry J, Lauper G, Zimmermann E and Weber E (1993). Does more music teaching lead to a better social climate? In R Olechowski and G Svik, eds, *Experimental research in teaching and learning*, 322–336. Peter Lang, Bern.

Trevarthen C (2002). Origins of musical identity: evidence from infancy for musical social awareness. In RAR MacDonald, DJ Hargreaves and DE Miell, eds, *Musical identities*, 22–46. Oxford University Press, Oxford.

Vanderark SD, Newman I and Bell S (1983). The effects of music participation on quality of life in the elderly. *Music Therapy*, **3**, 71–81.

Veblen KK (2004). The many ways of community music. *The International Journal of Community Music*, **1**, 8–16.

Zachopoulou E, Tsapakidou A and Derri V (2004). The effects of a developmentally appropriate music and movement program on motor performance. *Early Childhood Research Quarterly*, **19**, 631–642.

# CHAPTER 45

# Music and consumer behaviour

Adrian C. North and David J. Hargreaves

THE past 20 years has seen a massive increase in research on the effects of music on consumers in both retail and commercial leisure settings such as restaurants, hotels, bars, gyms, and the like. There are at least two reasons for this, namely an increasing interest among music psychologists in social psychological factors and the application of findings to practice (North and Hargreaves 2008), and an increasing interest among consumer psychologists (and people in cognate disciplines) in how aspects of the service environment can influence customers. Indeed, the quantity of research has reached such a level that North and Hargreaves (2008) were able to devote an extended chapter to a detailed review, and Garlin and Owen (2006) were able to conduct a meta-analysis based on only a portion of the number of studies published. Garlin and Owen's analysis showed 'small-to-moderate, yet quite robust effects in terms of background music and the dependents: value returns, behaviour duration and affective response' (p. 755): North and Hargreaves' (2008) review concurred. Indeed in addition to our extended chapter on the subject in *The social and applied psychology of music*, we have been invited to contribute chapters to edited volumes on four separate occasions over recent years (North and Hargreaves 1997, 2005, 2006; North *et al.* 2002) which suggests that, if nothing else, the field has become a recognized part of the psychology of music.

Given the recency of these other reviews it would be a wasted opportunity to merely reiterate the same arguments here. Instead we have opted to take this opportunity to focus solely on our own research on the subject that has been carried out over the past decade. Each of these studies has had its own particular theoretical perspective of course, but as a whole our research in this area has had a broader ambition, namely to begin mapping out the *range* of influences that music might have. In short, the question we have asked is 'What effects can music have on consumers?'

There is both a theoretical and a practical motivation behind this question. From a theoretical perspective, the existing research points to three psychological processes that underlie many effects. The first of these processes concerns the effect of music on physiological arousal. As we shall see shortly, several studies have argued that faster music speeds up customer activity, and explain this by referring to (now fairly antiquated) evidence from experimental aesthetics concerning how factors such as musical tempo might increase activity in the autonomic nervous system. The second process concerns the ability of music to prime certain thoughts. Several specific variations of the theory have been proposed (e.g. Hansen and Hansen 1990; Martindale and Moore 1989; North *et al.* 2004), but all of them describe how hearing a piece of music can activate several related ideas that are either expressed directly or indirectly in the music and lyrics or expressed indirectly as a result of shared cultural understanding of the meaning of the music in question. For example, hearing the *Sergeant Pepper* album by The Beatles primes thoughts related to 'love' as a

result of not just those lyrics that directly address the subject, but also because, on a cultural level, we associate the album with the 'flower power' movement of the late 1960s. Finally, the research suggests that the emotional effects of music can also influence consumers. There are many specific manifestations of this approach, although most involve some aspect of emotional response to music having a subsequent influence on the processing of commercially relevant information.

Of course three of the key explanations evoked by over a century of experimental psychology have also concerned arousal-based effects, priming effects, and the impact of emotion on behaviour. For example, classic research on skilled performance (e.g. Yerkes and Dodson 1908) has focused on how the arousal and information-processing load evoked by a stimulus can mediate performance across numerous domains. Similarly one of the central tenets of social cognition is that each new stimulus is processed in the light of stimuli encountered previously, such that responses to the latter prime responses to the former. Furthermore, large sections of social psychology (e.g. prosocial behaviour, environmental psychology) have considered the impact of emotional states on the ways in which people behave in everyday situations. Since arousal-, priming-, and emotion-based effects have such fundamental relevance to a range of human behaviours and can all result from hearing music, it seems reasonable to make the starting assumption that arousal-, priming-, and emotion-based effects of music might influence a broad range of consumer behaviours.

It also makes practical sense to map out the range of influences of music on consumer behaviour. In short, a store manager may use a certain type of music because they have read reports suggesting that it increases the speed with which customers shop. However, the music could be extremely counter-productive from a commercial standpoint if it nonetheless decreases customers' propensity to spend money, distracts their attention from special offers, makes the store feel crowded, or annoys customers so much that they decide to leave, to name but a few of the possible unintended outcomes. In short, the store manager can only use music optimally if they understand *all* the effects that a given piece of music will have. In particular, the commercial world should be interested in purely theoretical research, as we have long-subscribed to the maxim that 'There is nothing so practical as a good theory': in a commercial context, this means that if a store manager understands how music has the effects that it does, then they can control those effects and use them to their best advantage. A theoretical and a practical understanding of the effects of music in commercial environments go hand in hand with one another.

We begin the chapter with a brief overview of two particular effects of music that have received a considerable amount of attention. From here we go on to use our own research to illustrate the many other commercially relevant processes that can be influenced by music.

## Speed of activity and time perception

Two particular topics have dominated research on music and consumer behaviour. These are the effect of music on the speed with which customers behave, and the impact of music on time perception.

Several studies indicate that the speed of customer activity is related positively to the tempo and volume of background music. Many researchers have drawn on the findings of research on experimental aesthetics from the 1960s and 1970s (e.g. Berlyne 1971) which argued that louder or faster music lead to greater levels of physiological arousal, whereas slower or quieter music lead to correspondingly lower levels of physiological arousal. For example, Milliman (1982) found that slow music led to customers in a supermarket shopping 15 per cent more slowly than when fast music was played. Furthermore, customers in the slow music condition also spent around 33 per cent more money, and this is probably because the act of shopping more slowly also gave customers more opportunity to browse a range of products. Smith and Curnow (1966) produced similar results when they manipulated musical volume rather than tempo in a supermarket. Similarly Milliman (1986) studied 1392 groups of restaurant customers over 8 weekends.

Slow music led to customers eating more slowly, completing their meal in 56 minutes on average compared to 45 minutes when fast music was played. This also affected spending. Customers in the slow music condition, who ate more slowly, ending up spending more money on drinks ($30.47 vs $21.62 in the fast music condition). Roballey *et al.* (1985) adopted a dependent variable that is surely unique in the entirety of published psychology, namely bites per minute, finding that fast music in a cafeteria led to 4.40 bites per minute whereas slow music led to only 3.83 bites per minute.

Effects such as these highlight a theme that we will return to several times in this chapter, namely that almost all uses of music in commercial environments reflect a trade-off in which practitioners must decide what type of music they want to use in order to achieve what specific commercial gain. In this case, there is a clear trade-off between the speed of customer activity and spending. For example, if someone is the manager of a café then on Saturday lunchtime all the tables will be full. It is consequently in their interest to play fast music, free up some tables, and therefore maximize profit. If it is Tuesday morning and the café is quite empty then it makes more sense to play slower music in the hope that Milliman's (1986) findings would be replicated and customers would spend more money on drinks.

The second topic that has received considerable research attention has concerned the possible impact of music on time perception. There are numerous, slightly differing, explanations of this effect. However, common to many is the notion, drawn from mainstream research on time perception, that the perception of the passage of time is related to the amount of information encoded during the period in question. This explanation states that the mind invokes an heuristic to the effect that 'if more has happened then more time must have passed'. Research using this approach has, for example, suggested that louder music should therefore lead to the perception of more time having passed. Louder music imposes a greater information-processing load, 'tricks' people into believing that more has happened, and therefore leads to the perception that more time has passed (e.g. Kellaris and Altsech 1992). Another version of this argument reasons that if music imposes a greater processing load on the listener (i.e. by being loud, fast, or complex) then this distracts attention from an internal cognitive timer that monitors the passage of time. Other research has gone on from here and related time perception to whether people like the music they hear. This research employs the intuitive notion that people pay more attention to stimuli that they like. So if somebody likes music that they hear then they pay more attention to it. Because they pay more attention to it, the music imposes a greater information processing load, and so liked music leads to the perception that more time has passed (e.g. Kellaris and Mantel 1994). Although there is evidence to support the arguments set out here, unfortunately this neat theorizing has not led to consistently neat findings. If correct, however, this line of reasoning again implies the existence of another trade-off for practitioners. Followed through to the logical conclusion, research on time perception implies that if a business aims solely to reduce the amount of time that people think they have been waiting for then it might be better to play music that the customers don't like! Alternately, disliked music will nonetheless annoy customers of course, so if perceptions of service quality are more important than waiting time customers should hear music that they like.

## The range of effects of music

In 1996 we began a series of studies that have aimed to outline the range of consumer behaviours that might be influenced by music. The remainder of this chapter is devoted to an overview of seven of these influences.

### Visiting

North and Hargreaves (1996a) used an emotion-based approach to investigate whether music could influence where people went. We set up an 'advice stall' in an on-campus cafeteria. In effect, the stall was a couple of tables pushed together on which were leaflets offering advice on a range of typical student welfare issues such as safe sex, recommended maximum levels of alcohol consumption, how to find good housing, and similar issues. From the stall we played

pop music of three levels of complexity. Melodic complexity is the extent to which the music is erratic, varied, and difficult to predict. Several lab-based studies in experimental aesthetics (see Berlyne 1971) have shown that moderately complex music seems to be liked most: music of low and high complexity tends to be unpopular.

We measured the impact of the music on the number of people visiting the stall. 'Visiting' was measured in two ways. First, diners in the cafeteria were asked to complete a questionnaire asking them to state how likely they were to visit the stall. Measured this way, people's propensity to visit the advice stall was not influenced by the music. However, a behavioural measure of visiting produced a very different conclusion. On the floor, about one metre from the stall, were a series of small stickers that marked the boundary of an 'approach zone'. Someone was counted as having visited the stall if they put one foot inside this zone and took another step towards the stall. When moderately complex pop music was played then 10.71 people visited the stall every 10 minutes, and this represented significantly more people than when no music (4.01 people) and high complexity pop music (3.67 people) was played.

This has several implications. First, people were predisposed to visit the physical source of liked music. This suggests strongly that playing liked music, in a shop for instance, would encourage people to enter the premises, or could attract people to areas of the premises that they tend not to visit otherwise (e.g. the point farthest from the door or upstairs). Second, the effect of the no music condition relative to the moderate and high complexity pop was interesting. The 'right' music was better at attracting people to the stall than was playing nothing: the 'wrong' music was worse than playing nothing. In short, the real question facing practitioners is not 'Should I play music or not?' but rather 'What is the right kind of music to play?': music that is liked is commercially 'better' than no music, whereas music that is disliked is commercially 'worse' than no music. Again, practitioners face a decision. If a manager is not confident that customers will like the music that is played then they should turn it off: if they are confident that customers will like the music then this should bestow a considerable commercial advantage. Third, these findings show that complexity is one of the variables that allows managers to determine what the 'right' kind of music is: laws of musical preference derived from the lab can be applied in commercial premises to produce music that customers like. Indeed, when completing the questionnaire diners were asked to say how much they liked the music (see North and Hargreaves 1996b), with data showing that moderately complex music was most popular. Diners' answers to the questionnaire also showed however that the more they disliked the music so the more noticeable it became: if a business plays the 'right' kind of music then it will get no thanks, but if it plays the 'wrong' kind of music then everyone will notice! This probably goes a long way to explaining the negative stereotype that the general public holds concerning piped music.

## Arousal-based goals

North and Hargreaves (2000) adopted an arousal-based approach to investigate further what the 'right' kind of music might be for commercial premises. As noted earlier, research on experimental aesthetics has argued that people like moderately complex music the most. Researchers such as Berlyne (1971; see also e.g. North and Hargreaves 1996c) have argued that this is because moderately complex music brings about a moderate level of arousal in the autonomic nervous system: Berlyne's theory says that high complexity music brings about too much arousal, and low complexity music brings about too little arousal. Berlyne's theory also argues that factors other than complexity can also cause music to have arousal-evoking properties. For example, louder or faster music is claimed to bring about more arousal than quieter or slower music.

However, if Berlyne's theory is correct and people always use musical complexity, tempo, and volume to bring about a state of moderate arousal, then why do gyms and nightclubs play music that is loud, fast, and otherwise designed apparently to increase an already very *high* level of arousal? According to Berlyne's theory, this music ought to be disliked. Indeed, North and

Hargreaves (1996d) found that when people were asked to imagine themselves in a nightclub then they claimed that they would want to hear arousing music, which suggests that this type of music is not simply being foisted upon an ungrateful public but is instead what people want to hear. In short, there is a problem when attempting to apply Berlyne's theory to common practice in commercial premises.

North and Hargreaves (2000) investigated this by using an experimental simulation of a gym and also a relaxing situation. In our first experiment we asked people to ride an exercise bike or lie down and relax. *After* doing these tasks, people could choose between a loud, fast (i.e. arousing) piece of music or a slow, quiet (i.e. much less arousing) version of the same piece. Consistent with Berlyne's theory, people who had just finished exercising chose to listen to the slow, quiet music that would reduce the level of arousal brought about by cycling to a moderate level: people who had just finished relaxing chose to listen to the fast, loud music that would raise their level of arousal from the low level brought about by relaxation back up to a moderate level. However, in the second experiment we asked people to choose between the two types of music *while* they either exercised or relaxed. This time we got the opposite pattern of results. People who were cycling chose to listen to loud, fast music that would further increase their level of arousal: people who were relaxing chose slow, quiet music that would further reduce their level of arousal. In other words, although most of the time (e.g. when sitting in a café at lunchtime) people like moderately arousing music, we sometimes choose to put ourselves in a state of very high or very low arousal and, under these circumstances, we use music to help achieve these arousal-based goals. In particular, people in gyms and nightclubs want to be in a state of high arousal and loud, fast music can help them to achieve this goal: people in hotels, health spas, or on long-haul flights want to be in a state of low arousal, and would want access to slow, quiet music that helps to achieve this goal.

## Store image and purchasing

North and Hargreaves (1998) used a priming-based approach to investigate the effect of music on the atmosphere or 'image' of a store and also actual purchasing. We returned to the same café that we used in our 'visiting' study. Rather than focus on complexity, this time we played different musical styles. Over four days we played pop music, classical music, easy listening music (that was modelled on the 'sweeping strings' stereotype of piped music), and no music. Customers were asked to complete a questionnaire. First they were given a list of 14 items on sale in the café such as a can of Diet Coke and a slice of vegetarian pizza. Customers had to say the *maximum* that they would be prepared to spend on each item in turn. They also had to rate the atmosphere in the café according to 20 different adjectives.

Different musical styles gave rise to very different 'atmospheres'. Pop music led to the perception that the café was lively and youthful, classical music led to the perception that the café was upmarket and sophisticated, whereas the stereotypical piped music led to the perception that the café was downmarket. This is quite interesting in its own right. However, where store atmosphere really seemed to make itself felt was in customers' statements of how much they were prepared to spend in the café. When no music was played, customers were prepared to spend a total of £14.30 on the 14 items. When stereotypical piped music was played, customers were prepared to spend a total of £14.51 on the 14 items. However, when pop music was played (which made the café seem lively and youthful), customers were then prepared to spend £16.61 on the 14 items. Furthermore, when classical music was played (which made the café seem upmarket and sophisticated) then customers were prepared to spend £17.23 on precisely the same items in precisely the same place. Remember also that the customers were exclusively students who would prefer pop music over classical music: although the effect of pop music on the amount that customers were prepared to spend can probably be attributed to liking for that music, the greater effect of classical music cannot be. Rather it seems that customers were tapping in to the atmosphere portrayed by the music and acting accordingly. To put it bluntly, the upmarket and sophisticated atmosphere promoted by classical seemed to prime customers to feel more affluent.

This idea of using music to predispose people to behave in a particular way has been referred to sometimes as 'musical fit', and has been studied by several researchers over the past few years. For example, Areni and Kim (1993) played classical music and pop music in a wine cellar. They found that although the two different types of music did not lead to customers buying any more wine, classical music led to customers buying more *expensive* wine, and they argue that this was because the classical music primed customers to feel more affluent and act accordingly. This notion of musical fit was the subject of a follow-up study that we carried out. North *et al.*'s (2003) study was carried out in a restaurant in the town of Market Bosworth in Leicestershire. Classical music, pop music, and no music were played over three weeks, and customers' spending seemed to be influenced by the apparent effects of musical fit. When classical music was played, customers spent £32.52 on average, whereas when pop music was played customers spent £29.46, and when no music was played customers spent £29.73. Since the restaurant management, quite understandably, did not want us to approach customers we were unable to measure their degree of liking for the music and rule this out as an explanation for the results. However, in apparent support for the 'musical fit' explanation of diners' spending, the positive effects of classical music on spending were particularly strong in the case of starters and coffee; in other words, items that it is possible for customers to avoid while nonetheless being able to tell themselves that they had 'eaten out'. Did they spend more on starters and coffee because the classical music made them feel more affluent?

## Product choice

North and colleagues (1997, 1999a) also used a priming-based approach to investigate another aspect of musical fit, namely product choice. We used an aisle-end display in a local supermarket that featured several French and German wines. Each shelf on the display featured one type of French wine and one type of German wine, with the two wines on that shelf representing a similar degree of 'dryness' and being of very similar price. In other words, the main distinction between the two types of wine on each shelf (at least for non-connoisseurs) was arguably the country of origin. On the top shelf of the display was a small tape player. On alternate days this played either French or German music. Note that this manipulation of the music did not require customers to know that for example, Debussy was French or that Beethoven was German. Rather the manipulation went for the lowest common denominator such that 'French music' featured accordions heavily (as well as other stereotypically 'French' features), whereas the 'German music' featured an oompah band performance of songs concerning, among other issues, the pleasures of drinking white wine on the Rhein. The display was in view of the checkouts, and we sat behind these, next to a taxi phone, with carrier bags full of shopping to create the impression that we were normal customers waiting for our taxi to arrive. In reality we were observing the number of bottles of wine sold.

The results were consistent with the notion of musical fit. When French music was played then French wine outsold German by five bottles to one. When German music was played then German wine outsold French by two bottles to one. We investigated this further using a questionnaire. Once customers left the wines section (and entered the frozen pizza aisle!) they were approached by a researcher who asked them to state whether the type of wine they had bought had been influenced by the music that was playing. The sales data showed that they ought to answer that it had been. Unfortunately only 44 customers agreed to answer this question. However, although the sales data indicated that wine choices had been influenced by the music, only one customer was prepared to admit this on the questionnaire.

Findings such as these have obvious implications. First and most obviously they suggest that musical fit can influence product choice as well as spending. Second, the difference between actual wine sales and responses to the questionnaire could be taken as implying that customers were in some way unaware of the effect of musical fit on their behaviour. If this were so, it would of course have serious ethical implications (and indeed one national newspaper headline at the time claimed that psychologists had

discovered a way of making consumers buy things that they didn't want). We suspect strongly, however, that a far more mundane explanation is correct. We noted earlier that the music played from our display was crass. Consequently we find it hard to believe that, when asked midway through a shopping trip, people would be prepared to say that 'in effect, yes, all it took to get me to buy French wine was a bit of music played from a tinny speaker': they had a far stronger motivation to tell the researcher, in effect, that 'no, I was not influenced, and I can assure you that I always buy this type of wine'. Instead the music probably worked in just the same way as an advert placed by the entrance, advertising the fact that French (or German) products were on sale and raising their salience. Indeed we would actually go a step further than this and argue that the effects of musical fit will probably occur only under conditions where customers are otherwise undecided about which course of action to take. For example, if the supermarket study had replaced wines with cars, then it seems almost inconceivable that music could have affected people choosing between buying a Renault and a Volkswagen (although this does of course remain to be shown by data). When other information is available to demonstrate which of the available options is best, or when people otherwise have a very strong motivation to think about the options in detail, we expect that musical fit would become a relatively unimportant factor in their decision-making. Balanced against this of course is the fact that consumers do make many decisions on the spur of the moment with little thought: a shopper may have pasta on their shopping list but no firm opinion about which of two brands is best, and be even less inclined to stand in the supermarket for a few moments and think about it. There is tremendous scope for the application of musical fit.

## Musical 'fit' in advertising

If musical fit does operate by raising the salience of ideas related to the music, then this greater salience ought to have implications for memory. North *et al.* (2004) studied musical fit in radio advertising. We reasoned that if music in the adverts 'fitted' with the advertised product then, by raising the salience of the product, the music should help people to remember it. We prepared radio adverts for five brands, an online bank, a people carrier, a bathroom cleaner, a chocolate bar, and an energy (i.e. high-sugar) drink. On top of the voiceover in each advert we played music that either did or did not fit with the characteristics of each brand. After hearing the adverts, people were asked to read a short story before then being asked to recall what they had heard. Recall was indeed higher when the adverts had featured music that 'fitted' the advertised product. This is consistent with the notion that musical fit had primed listeners' thought processes, such that this priming aided recall. Rather less profoundly, these results show that musical fit also operates in the context of advertising. Also, if musical fit improves memory by raising the salience of certain ideas or causing a greater degree of thought, then it may also be the case that very low levels of congruity between music and product achieve the same effect: people would be expected to try to resolve the extreme incongruity between music and product, and the greater amount of thought devoted subsequently to the product may also have the effect of improving memory for it.

## Music on-hold

Several other studies have considered time-related effects of music. North *et al.* (1999b) investigated a simple question, namely whether on-hold phone music should work. The research was set up to look into two competing hypotheses. On the one hand, as described earlier, the research on time perception indicates that liked music should lead to longer time duration estimates. This means that hearing liked music while waiting on hold ought to produce longer estimated waiting times and an increased propensity to hang up. On the other hand, an emotion-based approach led us to reason that if callers like the music they hear while on-hold then they ought to be more inclined to stay on the line: if North and Hargreaves (1996a) showed that that liked music could attract people to an advice stall then why shouldn't liked music attract people to a phone line? Alternately, another emotion-based possibility was that liking for the

music was less relevant than the potential ability of music to simply calm callers down and counteract their irritation at being kept waiting.

To test these ideas we placed an advert in our local newspaper. This asked people to call a phone number in order to take part in a survey. When people called we automatically put them on hold. While waiting they heard one of three things. Some people heard 'liked music' which we operationalised as excerpts of three songs by The Beatles, namely *Yesterday*, *Eleanor Rigby*, and *Hey Jude*. Other people heard the 'weapon of choice' of most on-hold phone systems, namely pan-pipe music, which in this case featured cover versions of the same excerpts of the three Beatles' songs. A final group of people heard nothing at all apart from a brief, repeating spoken message encouraging them to keep waiting. At no point in the experiment did we actually answer to the phone: instead we simply timed how long people waited for before they lost patience and hung up.

The spoken message led to callers waiting, on average, for 197.76 seconds; and the original songs by The Beatles led to callers waiting for 229.50 seconds. However, the most successful condition was the pan-pipe cover versions, which led to callers waiting for an average of 257.28 seconds. Indeed, the success of the pan-pipes is best illustrated by pointing out that, in order to keep the line clear for new callers, we automatically cut people off once they had waited for 300 seconds. Given that some people hung up the moment they realized they had been put on hold, the average waiting time for the pan-pipe condition indicates in effect that we simply couldn't get rid of people in this condition!

In an attempt to explain the success of the pan-pipes (and arrange to pay people for their trouble) we used caller identification to call people back. Callers were asked to rate our telephone service on several dimensions. The Beatles and the pan-pipes conditions were liked as much as one another, and so liking for the music could not explain the effects of the pan-pipes on waiting time. However, when people were asked to rate the phone service according to five adjectives only one, 'peaceful', gave rise to significant differences between the conditions, with the pan-pipe condition receiving the highest ratings. In short, on-hold phone music can persuade people to wait for longer than can a spoken message, but it is important that this music should reduce stress and irritation in order to be successful.

### Priming store image

Much of the research reviewed above highlights the importance of using music to control the impression that customers have of a particular set of commercial premises. This then raises the issue of how precisely music influences these impressions, and this was investigated by North *et al.* (2000). We studied music in a bank branch and a local bar. In the city centre bar we played classical music, pop music, and no music over nine days. Customers rated the extent to which the bar could be described in terms of 11 different adjectives. Factor analysis of these ratings indicated that the adjectives fell into three factors, namely 'dynamic/upbeat', 'cerebral', and 'aggressive/downmarket'. The bank study played classical music, easy listening music, and no music over three weeks. Customers rated the extent to which the bank could be described in terms of 20 different adjectives, and a factor analysis showed that these adjectives fell into four factors that were remarkably similar to those in response to the bar, namely 'dynamic/upbeat', 'dignified', 'assertive/aggressive', and 'inspirational'. In both the bar and the bank, different musical styles gave rise to different scores on these factors. More simply, any attempt to use music to control the impression that customers form of commercial premises should focus primarily on the factors described here, since these subsume many of the other, more fine-grained responses that customers might have.

## Conclusions

The findings reviewed here provide initial positive signs that music can have a wide range of positive commercial benefits. Music can influence the places that customers go to, customers' ability to achieve a desired level of arousal, the atmosphere of commercial premises, the amount that customers are prepared to spend,

the amount they actually spend, the products they buy, their memory for advertising, and the amount of time they wait on hold. In addition to research by other academics on time perception and the impact of musical tempo and volume on the speed of customer behaviour, our own studies described here begin to map out the range of effects that music might have on consumers. Furthermore, the 10 effects described here seem to be based on fundamental psychological processes concerning arousal, priming, and the effects of emotions on behaviour, such that they may represent only the beginning of a list rather than a final definitive account.

Furthermore the research described here highlights an important issue for practitioners. If music can have many different effects it is important to prioritize those that correspond best with the business' marketing goals. Following from this, there can be no single 'right' type of music that is a universal commercial panacea, and the real issue is how to select music that addresses the most important marketing goals and does not hamper others. The specific answer to the question 'What music should my business play?' involves consideration of the psychological effects of music that are known to exist, the importance of various marketing goals, the musical preference of the customers (or at least careful steps to avoid annoying customers), along with numerous practical constraints such as budget or the need to install appropriate equipment.

Note also that we say that the evidence here provides only 'initial signs' of a positive effect of music because of the obvious difficulty of devising a detailed model of the process and testing all the possible aspects of this. The typical customer is engaged in a complex decision-making task, weighing up at least two competing alternative products along numerous cognitive and social psychological dimensions (e.g. a consideration of price vs benefit, product quality, relevance to self-concept, prevailing fashions etc.). Furthermore, some of these dimensions will only have become salient very recently, whereas others have been evaluated in detail over extended periods of time (as a result of prior experience of the products, advertising, or the opinions of friends, to name just three). At the same time, the customer will be managing several competing attentional demands (e.g. the presence of others, other time pressures etc.). As if understanding this were not complicated enough, music itself of course varies along numerous dimensions, all of which affect a customer in some way, with some of these effects being complementary and others less so. Devising an explanation of how all these factors interact is a daunting task, but at least work has begun on the basic building blocks, namely an outline of the various effects of music on consumers.

## References


Areni CS and kim D (1993). The influence of background music on shopping behavior: classical versus top-forty music in a wine store. *Advances in Consumer Research*, **20**, 336–340.

Berlyne DE (1971). *Aesthetics and psychobiology*. Appleton-Century-Crofts, New York.

Garlin FV and Owen K (2006). Setting the tone with the tune: a meta-analytic review of the effects of background music in retail settings. *Journal of Business Research*, **59**, 755–764.

Hansen CH and Hansen RD (1990). Rock music videos and antisocial behavior. *Basic and Applied Social Psychology*, **11**, 357–369.

Kellaris JJ and Altsech MB (1992). The experience of time as a function of musical loudness and gender of listener. *Advances in Consumer Research*, **19**, 725–729.

Kellaris JJ and Mantel SP (1994). The influence of mood and gender on consumers' time perceptions. *Advances in Consumer Research*, **21**, 514–518.

Martindale C and Moore K (1989). Relationship of musical preference to collative, ecological, and psychophysical variables. *Music Perception*, **6**, 431–455.

Milliman RE (1982). Using background music to affect the behavior of supermarket shoppers. *Journal of Marketing*, **46**, 86–91.

Milliman RE (1986). The influence of background music on the behavior of restaurant patrons. *Journal of Consumer Research*, **13**, 286–289.

North AC and Hargreaves DJ (2008). *The social and applied psychology of music*. Oxford University Press, Oxford.

North AC and Hargreaves DJ (1996a). The effects of music on responses to a dining area. *Journal of Environmental Psychology*, **16**, 55–64.

North AC and Hargreaves DJ (1996b). Responses to music in a dining area. *Journal of Applied Social Psychology*, **24**, 491–501.

North AC and Hargreaves DJ (1996c). Responses to music in aerobic exercise and yogic relaxation classes. *British Journal of Psychology*, **87**, 535–547.

North AC and Hargreaves DJ (1996d). Situational influences on reported musical preference. *Psychomusicology*, **15**, 30–45.

North AC and Hargreaves DJ (1997). Music and consumer behaviour. In DJ Hargreaves and AC North, eds, *The social psychology of music*, 268–289. Oxford University Press, Oxford.

North AC and Hargreaves DJ (1998). The effect of music on atmosphere and purchase intentions in a cafeteria. *Journal of Applied Social Psychology*, **28**, 2254–2273.

North AC and Hargreaves DJ, (2000). Musical preference during and after relaxation and exercise. *American Journal of Psychology*, **113**, 43–67.

North AC and Hargreaves DJ (2005). Musical communication in commercial contexts. In D Miell R MacDonald and DJ Hargreaves, eds, *Musical communication*, 405–422. Oxford University Press, Oxford.

North AC and Hargreaves DJ (2006). Music in business environments. In S Brown and U Volgsten, eds, *Music and manipulation: on the social uses and social control of music*, 103–125. Berghahn, Oxford.

North AC, Hargreaves DJ and McKendrick J (1997). In-store music affects product choice. *Nature*, **390**, 132.

North AC, Hargreaves DJ and McKendrick J (1999a). The effect of music on in-store wine selections. *Journal of Applied Psychology*, **84**, 271–276.

North AC, Hargreaves DJ and McKendrick J (1999b). Music and on-hold waiting time. *British Journal of Psychology*, **90**, 161–164.

North AC, Hargreaves DJ and McKendrick J (2000). The effects of music on atmosphere and purchase intentions in a bank and a bar. *Journal of Applied Social Psychology*, **30**, 1504–1522.

North AC, Hargreaves DJ, MacKenzie L and Law R (2004). The effects of musical and voice 'fit' on responses to adverts. *Journal of Applied Social Psychology*, **34**, 1675–1708.

North AC, Hargreaves DJ and Tarrant M (2002). Social psychology and music education. In RJ Colwell and CP Richardson, eds, *The new handbook of research on music teaching and learning*, 604–625. Oxford University Press, New York.

North AC, Shilcock A and Hargreaves DJ (2003). The effect of musical style on restaurant customers' spending. *Environment and Behavior*, **35**, 712–718.

Roballey TC, McGreevy C, Rongo RR, Schwantes ML, Steger PJ, Wininger MA and Gardner EB. (1985). The effect of music on eating behavior. *Bulletin of the Psychonomic Society*, **23**, 221–222.

Smith PC and Curnow R (1966). 'Arousal hypothesis' and the effects of music on purchasing behavior. *Journal of Applied Psychology*, **50**, 255–256.

Weyant JM (1978). Effects of mood states, costs, and benefits on helping. *Journal of Personality and Social Psychology*, **36**, 1169–1176.

Yerkes RM and Dodson JD (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurological Psychology*, **18**, 459–482.

# PART 10

# Music Therapy

Edited by Michael Thaut

CHAPTER 46

# Processes of music therapy

## Clinical and scientific rationales and models

Shannon de l' Etoile

> To apprehend music as an essential form of human behavior is to make more secure the foundations of music in therapy. Music therapy has long needed such a platform for its theoretical constructs, one that would be in accord with biological as well as psychological concepts.
>
> Gaston 1968

## Introduction

These groundbreaking words clarified the need for a theoretical basis for music therapy practice; however, decades later, a consistent theory has yet to be recognized. The problem may stem from poorly organized research that has little relevance to clinical practice, thus limiting its applicability (Aigen 1991; Asmus and Gilbert 1981; Gfeller 1995; Nicholas and Gilbert 1980; Wheeler 1983). Furthermore, music therapists have traditionally borrowed research paradigms from other disciplines, such as psychology or the biological sciences (Aigen 1991). While such efforts might have initially helped the profession to gain recognition in scientific or academic settings, they do not clarify the unique role of music in changing human behaviour. Consequently, these external paradigms may not be appropriate for guiding music therapy practice.

To explore this idea, Gfeller (1987) conducted a content analysis of 243 articles published in the *Journal of Music Therapy (JMT)* between 1964 and 1985 that related to theory or practice in music therapy. Her results revealed that no single theory or philosophy appeared to be prominent in music therapy practice. Rather, 48 per cent of the articles utilized various theories in psychology, including psychoanalytic (11 per cent), behavioural (26 per cent) or other prominent schools (11 per cent), including cognitive, humanistic and developmental theories.

Without a common theory, music therapists may struggle to produce consistent clinical outcomes and to communicate effectively with other health care professionals (Scovel and Gardstrom 2002). A unified theory may best emerge from a review of previous theories. While it would not be within the purview of this chapter to review all theories ever applied in music therapy practice, Gfeller's research (1987) provides a logical point of reference. Based on her findings, this chapter will review behavioural, psychoanalytic and humanistic music therapy. This discussion will culminate in the presentation of two newly developed theories that demonstrate a mutually beneficial relationship between research and clinical practice.

## Behavioural music therapy

Behaviourism grew out of B.F. Skinner's work, which dominated the practice of psychology between 1920 and the middle of the twentieth century (Scovel and Gardstrom 2002). Skinner claimed that learning happens when events are paired with consequences. Originally, behaviourism emphasized the study of objective behaviour of individuals while interacting with the environment and relied on experimental methods, thus minimizing a focus on mental processes or emotions.

Since its inception, behaviourism has evolved and become less mechanistic in its current clinical applications. Contemporary behaviour therapy offers three basic approaches, starting with operant conditioning, in which behaviour is a function of its consequences (Standley *et al.* 2004). A second approach is classical conditioning, which involves manipulation of the environment to alter cognitive processes. More recently, social learning theory has emerged as another approach that incorporates strategies from social, personality and developmental psychology, such as vicarious learning and self-regulation (Standley *et al.* 2004).

Behavioural music therapy was most prominent during the 1960s and 1970s; peaked from 1976–1978, and slightly declined after that time (Gfeller 1987; Standley *et al.* 2004). Results from a more recent content analysis of the *JMT* (Gregory 2002) further indicate that behaviourism was emphasized in the literature during the 1970s and 1980s, and continued throughout the 1990s. The approach has most often been reported with children in educational settings, but has also been used with a variety of other clinical populations, including premature infants, adults with developmental disabilities and older adults with Alzheimer's disease (Standley *et al.* 2004).

A behavioural music therapist manipulates musical stimuli to facilitate a change in the client's observable, measurable behaviour (Scovel and Gardstrom 2002). A well-known behavioural approach is Applied Behaviour Analysis (ABA), which utilizes principles of operant conditioning. With ABA, the therapist first establishes a baseline of the client's behaviour, implements a strategy for change, and then documents the client's response (Brownell 2002; Hanser 1995). Another common application of behavioural music therapy is the use of music as contingent reinforcement. The opportunity to play a favourite instrument, for example, could reinforce a desired behaviour (Scovel and Gardstrom 2002).

Standley (1996) conducted a meta-analysis of 98 studies that used music as a contingent reinforcer for educational and therapeutic behaviours. Results indicated that contingent music is more effective than other, non-musical contingent reinforcers, and is more effective than continuous music at facilitating positive changes in behaviour. While Standley's findings provide strong evidence of music's effectiveness as a contingent reinforcer, other researchers may disagree. In an earlier study by Wolfe and Hom (1993), typically developing preschool-age children participated in a number of musical and non-musical interventions designed to help them learn phone numbers. Results revealed that the contingent use of music (i.e., being able to play a preferred musical instrument) was not effective in helping the children to memorize the numbers. Rather, the children were most successful at learning during the intervention involving singing the phone numbers to the melodies of familiar songs.

These contrasting findings may cast some doubt on the behaviourism's appropriateness as a comprehensive theory to explain music's therapeutic mechanism. The fact that contingent music is not consistently effective points to the idea that the behavioural principle of contingency is not the agent of change. Rather, the clients' behaviour is likely to change due to their response to the perceived, inherent value of the music. Understanding the reward value of music should then become the central focus of both research and therapy.

To explain further, while behaviourism has benefited the music therapy profession, the theory does not explain how music is functioning to change behaviour. For example, certain behavioural principles are necessary and effective in a number of therapeutic settings, including the use of contingent reinforcement (musical or otherwise), adherence to group rules and the ability to self-regulate. These constructs; however, may be applied in any

therapeutic situation (musical or otherwise) and may have little to do with the music itself. Furthermore, within schools of psychology, clinical applications of behaviourism tend to fall under the 'psychotherapy' heading. While music therapy certainly does address many psychotherapeutic goals, it can also focus on other clinical needs, such as language development or gait training, which clearly would be misplaced under the psychotherapy heading. Aligning music therapy strictly with behaviourism could potentially be an inaccurate and restrictive manoeuvre.

While behavioural research approaches (i.e., single subject and multiple baseline designs) can effectively identify treatment outcomes, they must be substantiated by a scientific understanding of how the music is functioning within the design. Critical questions remain unanswered, such as: *why* does the behaviour change only when the music is present? Is the music rewarding due to timbre, style, familiarity, tempo, or pitch range? Does simple pattern repetition provide an appropriate level of sensory stimulation that the client finds innately rewarding?

Music therapists should not operate from blind assumptions about how the music is functioning. As Gfeller (1987) states, recognizing music as an effective reinforcement in behavioural therapy may not adequately explain music's influence on human behaviour. Rather, the therapist's application of music must be based on an extensive and systematic examination of the involved musical elements (i.e., novelty, complexity, redundancy, etc.), and the ways that these elements interact with the individual and the environment (Gfeller 1987). Without this knowledge, generalization of techniques will be limited at best. In summary, despite its widespread usage and demonstrated efficacy, behavioural music therapy does not offer a complete theoretical explanation for practice and research in music therapy.

## Psychoanalytic music therapy

Psychoanalysis is based on the work of Sigmund Freud, whose groundbreaking ideas transformed the practice of psychotherapy during the early twentieth century (Scovel and Gardstrom 2002). More contemporary adaptations of Freud's ideas have since been developed by Alfred Adler, Carl Jung, Erik Erikson, Melanie Klein, and others (Hadley 2002; Scovel and Gardstrom 2002). Freud claimed that people function at various levels of awareness, including unconscious, preconscious and conscious. For some individuals, unresolved emotional conflicts may reside in the unconscious and thus create personality abnormalities. Consequently, the personality must be reconstructed to achieve optimum health. According to Freud, repressed unconscious material must be brought into the client's awareness and then processed through the experiences of transference and countertransference (Bruscia 1998a; Scovel and Gardstrom 2002). Common psychoanalytic techniques include free association, dream analysis and interpretation.

To avoid confusion, various terms may be used interchangeably with psychoanalytic psychotherapy, including 'psychodynamic psychotherapy' and 'insight-oriented psychotherapy' (Isenberg-Grzeda *et al.* 2004). For the purposes of this chapter, all of these terms pertain to clinical techniques that have evolved from Freud's classical psychoanalytic technique.

Gfeller's (1987) content analysis of the *JMT* showed that from 1964 to 1984, 11 per cent of the articles published were linked to psychoanalytic theory. In addition, psychoanalytic theory dominated the *JMT* within its first six months of publication before declining in the mid 1970s. In psychoanalytic music therapy, music experiences may be used to supplement or replace traditional verbal psychoanalysis (Bruscia 1998a). The music therapist creates experiences to stimulate thoughts and feelings related to the client's past and present life, thus eliciting repressed unconscious material. One of the most common techniques used in psychoanalytic music therapy is vocal and/or instrumental improvisation between therapist and client (Bruscia 1998a). Improvisation gives the client the opportunity to engage in free expression of all thoughts, feelings and sensations; thus, a musical free association.

Mary Priestley and colleagues expanded these fundamental ideas in the 1970s by establishing the Analytical Music Therapy (AMT) approach,

which combines musical improvisation with movement and verbal processing (Scovel and Gardstrom 2002). In AMT, therapist and patient improvise together, and subsequently use psychoanalytic techniques to determine what happened or what they were thinking about during the improvisation (Eschen 2002). According to Priestley, improvised music essentially provides a gateway to the unconscious, as it allows for non-verbal expression of unconscious processes (Hadley 2002). The therapist works as a non-directive partner with the patient and maintains an openness to the processes of transference and countertransference (Eschen 2002). Transference occurs when the client projects unfinished emotional business onto the therapist or the music, while countertransference pertains to the therapist's emotional response to the patient (Priestley 1975). The technique is most often used with patients in psychiatric settings and in other areas of health care. While AMT has not gained strong acceptance in the United States, it is widely practised in Europe.

Another common technique in psychoanalytic music therapy is imaging to music, which is often achieved through the Bonny Method of Guided Imagery and Music (GIM) (Scovel and Gardstrom 2002). With this technique, the therapist provides carefully sequenced classical music selections to help the client achieve altered states of consciousness. In these altered states, repressed emotional themes may emerge through imagery, which can be processed with the therapist through verbal dialogue (Bonny 1994).

While AMT and GIM offer guidelines for how a clinician should proceed during a therapy session, psychoanalytic music therapy does not typically utilize specific techniques (Isenberg-Grzeda *et al.* 2004). Rather, the clinician will use any number of strategies, such as vocal or instrumental improvisation, songwriting, lyric analysis or imaging to music, based on his understanding of the therapeutic process as it unfolds. While this flexibility might seem appealing initially, the lack of consistent, standardized techniques can hinder general understanding and application of the theory.

Furthermore, some disagreement exists among psychoanalytic clinicians regarding the agent of change in the therapeutic process: is it the musical experience, the verbal exchange, or the combination of the two? (Isenberg-Grzeda *et al.* 2004). The fact that at least part of the technique's success depends on the nature of the relationship between patient and therapist also raises some concerns, as this highly subjective variable is not easily replicated, generalized or measured. Additionally, much of the explanation for the therapeutic mechanism in psychoanalytic music therapy seems to result from the therapist's *interpretation* of the client's behaviour (i.e., the meaning of the instrument that the client selects, the way that the instrument is played, recurring themes within melodic and rhythmic patterns, etc.). This interpretation is also subject to bias, including the clinician's level of training and amount of clinical experience.

The subjective nature of psychoanalytic music therapy also does not lend itself well to empirical examination. Consequently, most of the research conducted in this area has been qualitative with some documentation of treatment effects published via case studies (e.g., Bruscia 1991). Certainly qualitative research can guide clinical practice and inform subsequent quantitative investigation. Additionally, a recent proliferation of literature on psychoanalytic music therapy indicates a growing interest in this application and belief in its efficacy (Isenberg-Grzeda *et al.* 2004). Overall, however, little empirical evidence exists pertaining to the effectiveness of psychoanalytic music therapy or its direct generalizability to other clients and settings. This lack of data limits the theory's growth as well as its acceptance and recognition within the medical community.

Another issue regarding psychoanalytic music therapy pertains to music's ability to elicit unconscious material, as well as to create opportunities for transference and countertransference. These processes are considered fundamental to the client's experience in psychoanalysis, yet the theory lacks a scientific description of the precise way in which music facilitates these events. If a therapist does not understand how a technique works, he will be restricted in his ability to use it effectively and risks using the technique inappropriately.

Like behavioural music therapy, the psychoanalytic approach is considered a psychotherapeutic modality. Limiting music therapy to this one application may be detrimental to the profession's growth and viability. Even casual observation reveals that musical experiences can improve numerous skills outside the psychotherapeutic realm, such as short-term memory, or movement accuracy. Despite the many clients who have benefited from this approach, psychoanalytic music therapy does not offer standardized techniques, nor does it provide a scientific explanation for music's therapeutic mechanism. For these reasons, psychoanalytic music therapy does not provide a comprehensive theory for clinical practice and research.

## Humanistic music therapy

In the 1960s and 1970s, therapists developed alternatives to the behavioural and psychoanalytic approaches, thus producing several new theories with a humanistic approach, including existential therapy (Viktor Frankl, Rollo May), person-centred therapy (Carl Rogers) and Gestalt therapy (Fritz Perls) (Corey 2005). While each approach includes unique tenets and techniques, all humanistic approaches share common elements, including a respect for the client's experience and a belief in the client's ability to make constructive choices (Corey 2005). Perhaps at the heart of the humanistic movement is the concept that each human being has an innate potential that can be actualized and through which one can find meaning.

Abraham Maslow was one of the earliest and most influential proponents of the humanistic movement (Scovel and Gardstrom 2002). He described a hierarchy of human needs, ranging from biological requirements to self-actualization, or the need to move toward wholeness and fulfillment. Mental disorders result when one fails to progress through the hierarchy, thus inhibiting the ability to find meaning and be responsible for oneself and others. By providing unconditional acceptance and maintaining a focus on the here-and-now, the humanistic therapist helps the client to identify and address factors that are blocking his self-actualization and to take greater responsibility for his choices and actions.

In her content analysis of the *JMT* from 1964–1984, Gfeller (1987) discovered that 11 per cent of the articles pertained to theories borrowed from various schools of psychology, including humanistic, and that these articles were most prominent after 1970. The humanistic music therapist uses music to identify the client's needs and facilitate the process of self-actualization (Scovel and Gardstrom 2002). A variety of methods may be used, ranging from active (i.e., improvisation) to receptive (i.e., listening) (Bruscia 1998b).

One of the most well-known approaches within humanistic music therapy is Creative Music Therapy (CMT), developed by Paul Nordoff and Clive Robbins (Aigen *et al.* 2004; Scovel and Gardstrom 2002). Both Nordoff and Robbins were drawn to Maslow's humanistic concepts, which they adopted and integrated with the ongoing development of their approach in the 1970s (Robbins 1993). In particular, they appreciated Maslow's emphasis on developing a client's strengths and potentials, rather than focusing on his deficiencies (Aigen *et al.* 2004).

With CMT, the clinician uses vocal and/or instrumental improvisation to encourage growth of the client's inner creative drive, thus accessing unknown developmental potentials that are blocked due to disability (Nordoff and Robbins 1977; Robbins 1993). Through the improvisation, the therapist can describe the client's personality, match the client's mood, mirror the client's facial expression or physical bearing, as well as imitate sounds that the client is making (Aigen *et al.*, 2004). Consequently, a client should experience increased emotional awareness as well as improved personality organization, and be able to interact with the environment in a more meaningful way (Nordoff and Robbins 1977; Turry 1998). CMT has been used primarily with children who have cognitive and motor deficits, and also adults in medical, psychiatric and geriatric settings (Andsell 1995; Scovel and Gardstrom 2002).

CMT does not prescribe a format or procedural sequence for the improvisations (Aigen *et al.* 2004). Rather, the therapist matches and enhances the client's emotions from moment to moment and the process is one of trial and error. A wide spectrum of beliefs exists among CMT clinicians regarding application of the technique

and whether or not it should be combined with various modes of psychotherapy. Some practitioners are more interpretive regarding client behaviour, and others are more music-centred, believing that interpretation can distort the musical interaction (Aigen *et al.* 2004; Turry 1998). Ultimately, CMT stems from a music-centred philosophy, in that music functions as the primary agent of change and offers unique therapeutic benefits that cannot be achieved verbally.

Similar to psychoanalytic music therapy, the humanistic approach is highly subjective, and thus difficult to examine using traditional scientific methods. Consequently, most research in humanistic music therapy has been qualitative, using the case study format (Aigen *et al.* 2004; Aigen 1995, 1998). To document treatment outcome, CMT clinicians typically do not collect data pertaining to specific target behaviours (Aigen *et al.* 2004). Rather, sessions are recorded and then reviewed in an extensive process known as indexing, which identifies ongoing changes in the client as well as in the music.

The humanistic approach has benefited the music therapy profession in several ways. Philosophically, a belief in the client's ability to change and grow should underlie any effective approach to treatment. Musically, the CMT focus on the aesthetic quality of the music and artistic acumen of the therapist certainly elevates standards of professional competence. Despite these benefits, the humanistic approach offers no consistent, standardized techniques, as would be expected of a therapeutic intervention in a medical setting. Data collection within CMT, albeit extensive, appears to focus more on musical outcomes as opposed to functional outcomes that may translate to improved behaviour beyond the therapy session.

Additionally, humanistic music therapy lacks a scientific explanation for how music functions as the agent of change. For instance, how does music access the inner creative drive? How does music facilitate personality organization, or improved perceptual processes? While client responses are evident from behavioural observation, a thorough understanding of music therapy's effect must include scientific knowledge of internal processes. Without this information, a clinician is essentially 'reinventing the wheel' in every therapy session, and will struggle to produce reliable outcomes. These issues undermine the strength of humanistic music therapy as an all-encompassing paradigm for the profession.

## Paradigm shift: from social science to neuroscience

The above theories reveal that music therapy practice has been based largely on a social science framework; that is, music has been valued for its ability to facilitate emotional expression and social integration (Thaut 2005). While these musical effects continue to have merit in contemporary practice, more recent evidence has prompted the profession toward a paradigm shift. Since the early 1990s, the advent of modern research techniques in cognitive neuroscience, such as brain imaging, brain-wave recordings and kinematic motion analysis, have provided a greater understanding of the neural processes involved in the perception and production of music. These data collection methods provide an understanding of how music 'works' in a therapy session. With this knowledge, music therapists can operate from a scientific framework and produce stronger, more specific outcomes than general well-being. Consequently, music therapy has evolved from a social science model to a neuroscience model of clinical practice and research.

This evolution has resulted in the formation of a new theory for research and clinical practice, known as Neurologic Music Therapy (NMT) (Clair and Pasiali 2004; Thaut 2005). NMT techniques address cognitive, sensory, and motor dysfunction resulting from disease of the human nervous system. Such techniques are designed to produce functional outcomes and are evidence-based, drawing from scientific data in basic and clinical research. Furthermore, while NMT techniques should be modified to meet each patient's unique needs, they are standardized in terminology and application, thus enhancing communication and clinical efficacy among music therapists and with other health care providers. Specific details regarding NMT techniques are provided elsewhere in this volume. The purpose of this chapter is to elucidate the theoretical components of the NMT model.

## The Rational-Scientific Mediating Model

NMT theory is founded in a neuroscience model of music perception, known as the Rational-Scientific Mediating Model (R-SMM), which explains how music functions as a mediating stimulus (Clair and Pasiali 2004; Thaut 2000, 2005). To explain further, aesthetic stimuli such as music typically evoke two sequential responses (Berlyne 1971). The initial response produces some type of feedback or internal stimulation in the recipient. This response joins with the stimulation coming directly from the music, thus producing a mediating response that determines what the recipient will do. Consequently, overt behaviour results from the stimulation coming from the music itself, as well as the stimulation arising from one's initial reactions to the music. Mikiten (1996) eloquently describes how a mediating response to music produces a desirable outcome:

> It is hard to frame a hypothesis and conduct research into how playing a certain song makes a cancer patient 'feel' better. Further, different songs affect patients differently. However, it does make sense to hypothesize that sound structures are detected by brain structures involved in memory. The act of recognition by brain circuits triggers endorphin release that modulates the brain structures responsible for pain. This point of view facilitates the design of hypotheses for experimental verification.
>
> Mitiken (1996, p. 21)

The mediating response in this case is recognition of sound structures and subsequent endorphin release. With knowledge of these internal processes, the music therapist understands how the intervention works and can apply it most effectively. The R-SMM explains how music functions to elicit a mediating response that eventually produces a desired, overt, behavioural response. Understanding of the mediating response is based on scientific evidence and not on speculation inferred from behavioural observations.

The R-SMM provides a framework for exploring and understanding mediating responses to music within the domains of affect, cognition and sensorimotor processes (Clair and Pasiali 2004; Thaut 2000, 2005). The model is rooted in the idea that the scientific basis for music therapy is found in the neurological, physiological and psychological foundations of music perception and production. Application of the R-SMM proceeds in the following four steps:

1 Musical response models: research in music therapy must begin with a solid understanding of musical behaviour, including its neurological, physiological and psychological components. At this step, the researcher might ask, how do humans perceive, produce or respond to music? What systems of the body, underlying neural pathways or perceptual processes are involved?

2 Non-musical parallel models: research at this level examines processes in non-musical brain and behaviour function. Through comparative analysis with step 1, this step establishes a meaningful link between musical behaviour and non-musical behaviour. A parallel process between musical and non-musical functioning must be established, before suggesting that music can influence non-musical behaviour.

3 Mediating models: if a parallel process has been identified at step 2, the researcher can move forward by examining or conducting research to explore the influence of music on non-musical brain and behaviour function. Mediating research does not involve investigation of long-term treatment effects; rather, it explores music's immediate effect on both typical and clinical populations. Results clarify the scientific mechanisms of music's effect on behaviour, or mediating response, thus providing the foundation for clinical research.

4 Clinical research models: when positive mediating effects have been established at step 3, the researcher can design and implement a clinical protocol examining the long-term, therapeutic effects of music. The focus at this level is on lasting functional behaviour change and carry-over after treatment.

As an example of the R-SMM, a researcher may explore music therapy for memory deficits in children with learning disabilities. In Step 1,

musical response models, the researcher gathers evidence regarding how music is memorized. In perusing the literature, he may discover that when learning music, Gestalt laws of perceptual grouping are commonly used to combine similar musical elements, thus reducing the memory load and allowing for more effective storage and recall of musical information. In Step 2, non-musical parallel models, the researcher explores how non-musical information is typically memorized, and looks for any shared strategies with memory for music. Examination of previous literature reveals that when memorizing visual or textual information (i.e., lists of words or phone numbers), Gestalt laws of perception are also utilized to group items based on similar features, thus facilitating efficient storage and recall.

With this parallel process established, the researcher can proceed to Step 3, mediating models. Here the researcher examines the effect of the musical behaviour (i.e., memory strategies for music) on the non-musical behaviour (i.e., memory strategies for textual information). The mediating response is the innate human tendency to perceive order and groupings in music stimuli. Perceptual grouping then provides a framework or vehicle for organizing and rehearsing the new information, thus leading to improved storage and recall.

The researcher might first implement his idea with a typical population, such as teaching a new phone number to preschool-age children by singing it to the melody of a familiar song. If positive results are obtained, then the researcher may apply the technique to children with memory deficits due to learning disabilities. Promising evidence at this stage would logically support systematic research at step 4, clinical research models. The researcher could then design a long-term treatment study, investigating the effect of a musical mnemonic device as a memory strategy for children with learning disabilities.

## The Transformational Design Model

The R-SMM provides clear guidelines for conducting research regarding music's therapeutic effects. A supplemental model is needed; however, to assist the clinician in translating research findings from the R-SMM into everyday practice (Thaut 2005). The Transformational Design Model (TDM) meets this need by providing a systematic, step-by-step approach to designing, implementing and evaluating clinical interventions (Clair and Pasiali 2004; Thaut 2000, 2005). The five steps of the TDM include:

1 Diagnostic and functional assessment of the patient
2 Development of therapeutic goals and objectives
3 Design of functional, non-musical therapeutic exercises and stimuli
4 Translation of step 3 into functional therapeutic music experiences
5 Transfer of therapeutic learning to functional, non-musical, real world applications.

Steps 1, 2, 3 and 5 of the TDM are common to all therapeutic disciplines (Thaut 2000, 2005). Step 1 involves a diagnostic and etiological assessment of the client which leads to Step 2, establishment of functional goals with corresponding measurable objectives. This information informs Step 3, the design of therapeutic exercises and activities to address the goals and objectives. A critical component of Step 3 is designing such exercises according to the functional behaviours of the patient and not on musical considerations. This approach to treatment planning ensures a functional outcome, not a musical outcome, and enhances collaborative efforts with other disciplines, thus producing more specific and meaningful treatment outcomes.

Step 4 makes the TDM unique to neurologic music therapy practice. Here the music therapist translates the functional and therapeutic exercises from Step 3 into musical experiences that are isomorphic; that is, they share the same functional structure (Thaut 2000, 2005). For example, if a patient has recently experienced a right hemisphere stroke, he may have weakness in his left arm and hand, including difficulty with independent finger movement. The functional exercise in this case would be repetitive movement of the affected arm, hand and fingers, for an appropriate duration of time and

at an adequate pace, in order to increase range of motion and to increase strength and endurance. The musical translation of this functional exercise could consist of playing simple fingering exercises on the piano, while being guided by the therapist's accompaniment which determines tempo, frequency and duration of movement, in addition to providing aesthetic motivation.

To ensure the isomorphic nature of the therapeutic music experience (TME), three types of logic must be incorporated (Thaut 2005):

1 Functional logic: the TME must serve as a musical analogue of the non-musical experience. Furthermore, the TME must give the client multiple or sustained opportunities to practice the non-musical behaviour. By incorporating functional logic, the therapist can facilitate positive changes in a functional behaviour that the client will utilize in everyday life.

2 Musical logic: the musical component of the TME must meet the aesthetic and artistic principles of good musical forms. To explain further, the music must be aesthetically pleasing, whether it is provided in active (i.e., improvisation) or receptive (i.e., listening) forms. Regardless of complexity, only with optimal musical patterns can the beneficial effects of music be manifest.

3 Scientific logic: the most effective TME will draw from the scientific evidence established using the R-SMM. Whenever possible, the music therapist must design TMEs based on a solid understanding of the parallel process between musical and non-musical experiences in order to elicit a mediating response to influence behaviour. By utilizing the research evidence, the therapist need not reinvent the wheel in each therapy session, and can instead, select the most appropriate technique from a validated taxonomy of strategies.

## Summary

The R-SMM and the TDM give music therapists two critical tools for research and clinical practice. The R-SMM integrates the understanding of musical and non-musical behaviours into a framework for effective therapy and rehabilitation (Thaut 2005). Its dynamic and open-ended nature allows for incorporation of future research efforts to continuously establish new evidence and rationales for the music's therapeutic effects. Meanwhile, the TDM provides a practical guide for music therapists to create goal-oriented interventions based on scientific evidence. Together, the models offer an approach to research and practice based on an understanding of how music influences behaviour to produce functional outcomes. Clinicians cannot assume that music magically or automatically brings about a desired change in behaviour. Rather, interventions must be founded in empirical knowledge of the internal processes and mechanisms at work during musical experiences.

By providing a scientific understanding of music's influence on behaviours in all domains of functioning, as well as standardized techniques, NMT addresses the profession's needs. Simultaneously, NMT fulfills Gaston's original appeal for a theory that incorporates both biological and psychological constructs. Furthermore, NMT provides an approach to practice and research that is unique to the music therapy profession and not borrowed from other disciplines. As a theoretical foundation, NMT will secure the profession's acceptance, credibility and future viability as a valid therapeutic modality.

## References

Aigen K (1991). The roots of music therapy: towards an indigenous research paradigm. *Dissertation Abstracts International*, **52**(6), 1933A. (University Microfilms No. DEY91–34717.)

Aigen K (1995). Cognitive and affective processes in music therapy with individuals with developmental delays: a preliminary model for contemporary Nordoff–Robbins practice. *Music Therapy*, **13**, 13–45.

Aigen K (1998). *Paths of development in Nordoff–Robbins music therapy*. Barcelona, Phoenixville, PA.

Aigen K Miller CK, Kim Y, Pasiali V, Kwak EM and Tague DB (2004). Nordoff–Robbins music therapy. In AA Darrow, ed., *Introduction to approaches in music therapy*, 63–77. American Music Therapy Association, Silver Spring, MD.

Andsell G (1995). *Music for life: aspects of creative music therapy with adult clients*. Jessica Kingsley, London.

Asmus EP and Gilbert JP (1981). A client-centered model of therapeutic intervention. *Journal of Music Therapy*, **18**, 41–51.

Berlyne DE (1971). Perception. In DE Berlyne, ed. *Aesthetics and psychobiology*, 96–114 Meredith Corporation, New York.

Bonny H (1994). Twenty-one years later: a GIM update. *Music Therapy Perspectives*, **12**, 70–74.

Brownell MD (2002). Musically adapted social stories to modify behaviors in students with autism: four case studies. *Journal of Music Therapy*, **39**, 117–144.

Brusica K (1991). *Case studies in music therapy*. Barcelona, Phoenixville, PA.

Bruscia K (1998a). An introduction to music psychotherapy. In K Bruscia, ed., *The dynamics of music psychotherapy*, 1–15 Barcelona, Gilsum, NH.

Bruscia K (1998b). *Defining music therapy*. Barcelona, Gilsum NH.

Clair AA and Pasiali V (2004). Neurologic music therapy. In AA Darrow, ed., *Introduction to approached in music therapy*, 143–158. American Music Therapy Association, Silver Spring, MD.

Corey G (2005). Person-centered therapy. In G Corey, ed. *Theory and practice of counseling and psychotherapy*, 7th edn, 162–189. Brooks/Cole–Thomson Learning, Belmont, CA.

Eschen JT (2002). Analytical music therapy—introduction. In JT Eschen, ed., *Analytical music therapy*, 17–33. Jessica Kingsley Publishers, London.

Gaston ET (1968). Man and music. In ET Gaston, ed., *Music in therapy*, 7–29. Macmillan Publishing, New York.

Gfeller KE (1987). Music therapy theory and practice as reflected in research literature. *Journal of Music Therapy*, **24**, 178–194.

Gfeller KE (1995). Status of music therapy research. In BL Wheeler, ed., *Music therapy research: quantitative and qualitative perspectives*, 29–63. Barcelona Publishers, Phoenixville, PA.

Gregory D (2002). Four decades of music therapy behavioral research designs: a content analysis of *Journal of Music Therapy* articles. *Journal of Music Therapy*, **39**, 56–71.

Hadley S (2002). Theoretical bases of analytical music therapy. In JT Eschen, ed., *Analytical music therapy*, 34–48. Jessica Kingsley Publishers, London.

Hanser SB (1995). Applied behavior analysis. In BL Wheeler, ed., *Music therapy research: quantitative and qualitative perspectives*, 149–164. Barcelona Publishers, Phoenixville, PA.

Isenberg-Grzeda C, Goldberg F and Dvorkin JM (2004). Psychodynamic approach to music therapy. In AA Darrow, ed., *Introduction to approaches in music therapy*, 79–101. American Music Therapy Association, Silver Spring, MD.

Mikiten T (1996). A method for research in music medicine. In RR Pratt and R Spintge, eds, *Musicmedicine*, Vol. 2, 14–23. MMB Music, St. Louis, MO.

Nicholas MJ and Gilbert JP (1980). Research in music therapy: a survey of music therapists' attitudes and knowledge. *Journal of Music Therapy*, **17**, 207–213.

Nordoff P and Robbins C (1977). *Creative music therapy*. John Day, New York.

Priestley M (1975). *Music therapy in action*. MMB Music, St. Louis, MO.

Robbins C (1993). The creative processes are universal. In M Heal and T Wigram, eds, *Music therapy in health and education*, 7–25. Jessica Kingsley, London.

Scovel MA and Gardstrom SC (2002). Music therapy within the context of psychotherapeutic models. In RF Unkefer and MH Thaut, eds, *Music therapy in the treatment of adults with mental disorders: theoretical bases and clinical interventions*, 2nd edn, 117–132. MMB Music, St. Louis, MO.

Standley JM (1996). A meta-analysis on the effects of music as reinforcement for education/therapy objectives. *Journal of Research in Music Education*, **44**, 105–133.

Standley JM, Johnson CM, Robb SL, Brownell MD and Kim SH (2004). Behavioral approach to music therapy. In AA Darrow, ed., *Introduction to approaches in music therapy*, 103–123. American Music Therapy Association, Silver Spring, MD.

Thaut MH (2000). *A scientific model of music in therapy and medicine*. IMR Press, San Antonio, TX.

Thaut MH (2005). Music in therapy and medicine: from social science to neuroscience. In MH Thaut, ed. *Rhythm, music and the brain: scientific foundations and clinical applications*, 113–136. Taylor and Francis Group, New York.

Turry A (1998). Transference and countertransference in Nordoff-Robbins music therapy. In K Bruscia, ed., *The dynamics of music psychotherapy*, 161–212. Barcelona, Gilsum, NH.

Wheeler BL (1983). Prologue. *Music Therapy*, **3**, 2–3.

Wolf DE and Hom C (1993). Use of melodies as structural prompts for learning and retention of sequential verbal information by preschool students. *Journal of Music Therapy*, **30**, 100–118.

CHAPTER 47

# Clinical practice in music therapy

Corene Hurt-Thaut

## Introduction

### What is music therapy?

Music therapy is a health care profession in which music is used as a therapeutic medium to address developmental, adaptive, and rehabilitative goals in the areas of psychosocial, cognitive, and sensorimotor behaviour of individuals with disabilities. A music therapist designs therapeutic music interventions which address the client's functional goals and are adapted to the client's functional level. Over 5000 music therapist are employed throughout the United States, serving a wide range of clients, including the elderly, developmentally disabled adults and children, mental health populations, physically disabled, school-aged children, early childhood, substance abuse, sensory impaired, neurologically impaired, and terminally ill (AMTA 2006).

There are three critical components when considering the use of music in therapy: musical creativity and proficiency, scientific logic, and therapeutic logic. First, the music used in therapy must come from a trained professional who has a high quality of understanding of music theory, history, and performance. Therefore, a music therapist background involves extensive training in music, including proficiency in voice, guitar, and piano. Second, in order to use music as a therapeutic tool, a therapist must understand the scientific foundations of the influence of music on functional changes in non-musical brain and behaviour functions. Although course work at different universities varies considerably, music therapist are exposed to, and in some cases, take extensive course work in the sciences in order to gain an understanding of anatomy and physiology, neuroanatomy, biology, and psychological functions. Thirdly, in order to use music in therapy, music interventions must make sense therapeutically. It is essential for music therapist to be trained in group and individual counselling skills, therapeutic design, and selection of music interventions based on functional goals and objectives.

### The music therapy treatment process

Music therapists typically work in conjunction with a team of professionals including physicians, nurses, social workers, physical therapist, occupational therapist, and speech therapist. A music therapist must receive a referral or request for music therapy services, upon which they are responsible for completing an assessment, developing a treatment plan including goals and objectives, documenting the progress of a client throughout treatment, evaluating progress toward goals and objectives, and developing a termination plan for their clients.

Referrals typically come from the physician in a hospital setting, but in other settings they may come from the psychologist, another therapist, teachers, parents, social workers, and even from the client themselves.

Once a referral is received, the therapist must do an assessment, of the person's strengths, needs and problems, including information about the client's medical history, cognition,

social abilities, physical abilities, vocational or educational background, emotional status, communication skills, family, and leisure skills. A music therapy assessment is typically non-musical in nature, and designed to identify whether a client is appropriate for music therapy and then assist the music therapist in treatment selection based on need areas.

Once an assessment has determined the strength and need areas, a treatment plan can be established. The treatment plan identifies the goals and objectives that will be addressed with the client during treatment. The goal identifies and area that is in need of improvement (e.g. to improve fine motor skills). The objective identifies the specific parameters of the behaviour that the client needs to accomplish, the criterion and qualifiers for that behaviour, and the target date of achievement.

Documentation of treatment progress is essential not only to monitor the quality, cost effectiveness, and efficiency of treatment, but also as a means to communicate with other members of the treatment team and to justify charges to insurance companies. Documentation requirements and formats vary greatly depending on the type of facility that a therapist is working at.

Once a client has met their treatment goals, or is recommended for discharge, the music therapist must complete a final evaluation of treatment. This evaluation includes a review of the entire music therapy process, from the initial goals and objectives to the final progress made toward them.

### The history of music therapy

The profession of music therapy was established in the United States in 1950, as a result of work being done with patients in veterans' hospitals following the Second World War. Music was being used with returning soldiers to boost the morale, and help with the rehabilitation of physical, social and emotional health. However, evidence of the powerful value of music has been documented for thousands of years in all cultures. David played the harp to sooth Saul; in Egypt, certain songs were legally ordained in the education of youth to promote virtue and morality; Pythagoras directed certain mental disorders to be treated by music; Xenocrates cured maniacs by melodious sounds, and Asclepiades conquered deafness with a trumpet (Davis *et al.* 1999). Music therapy has grown substantially in the research and clinical knowledge since 1950, and today is recognized in the medical community as a solid, thriving profession.

## Application of music therapy to clinical populations

### Music therapy in psychiatric populations

During the early development of music therapy, most therapists were working with psychiatric populations. Although the use of music in therapy has become much more diverse since then, mental health accounts for 21 per cent of the populations served by music therapist (AMTA 2006). Treatment most often addresses thought disorders, mood disorders, personality disorders and anxiety disorders.

Wheeler (1983) classified the treatment of adults with mental disorders into three levels: (1) supportive music therapy as an activity therapy: (2) insight music therapy with re-educative goals, and (3) insight music therapy with reconstructive goals. Each level of classification requires a higher degree of insight and participation from the client.

On the level of supportive music therapy, the therapist main goals are to promote mood-lifting stimulation and healthy behaviours. Music therapy experiences at this level require active involvement and awareness of the here and now.

On the level of insight and process-oriented music therapy with re-educative goals, the client is now at a higher level of insight and is ready to reflect and process interpersonal relationships and emotions with a greater level of verbal participation. A client at this level is well oriented to reality, demonstrates appropriate interaction skills and is ready to focus on conscious conflicts and associated unhealthy defence mechanisms.

The last level of intervention involves reconstructive, analytically and catharsis-oriented

music therapy. At this level, the music therapy interventions are used to uncover and process subconscious conflicts resulting from past traumatic situations. Therapist working at this level require advanced training in psychotherapeutic techniques.

In addition to different levels of treatment, music therapy clinical practice with mental disorders is guided within the framework of several different psychotherapeutic models. The therapist typically chooses different treatment approaches based on the needs of a particular client or the philosophy of a treatment facility. Some of the most frequently used models include: psychodynamic, cognitive, humanistic/existential, cognitive neuropsychiatric, biomedical, behavioural, and holistic.

### Interventions

A wide range of music therapy programmes and techniques are currently used in the treatment of mental disorders, including music performance, music psychotherapy, music and movement, music combined with other expressive arts, recreational music, and music for relaxation.

Music performance addresses goal areas such as affective self-expression, identification of feelings, appropriate outlet for negative feeling processes, reality testing, retraining auditory and perceptual skills and memory function, executive functions, motivation, self-esteem, and socialization. Music therapy techniques may be process- or product-oriented, depending on the therapeutic goal, and often include instrumental group improvisation, instrumental and vocal performance ensembles, group singing therapy, individual instrumental or vocal instruction, and individual music improvisation (Unkefer and Thaut 2002).

Music psychotherapy implements music therapy techniques such as supportive, interactive, and catalytic music groups and individual therapy. A supportive music group would address goals in the areas of identification and expression of feeling responses, expanding range of expressiveness, appropriate outlets for negative feelings, reality focus, task-oriented activities, attention-focusing activities and social participations. An intervention may involve guided music listening techniques to focus attention, provide theme-centred structure for verbal interaction, and evoke feeling responses or alter feeling states. In an interactive music group or individual therapy, the main goal areas become focused on the identification and expression of one's own feelings, problem-solving, awareness of one's own behaviours, and facilitation of behaviour changes (Unkefer and Thaut 2002). Music therapy interventions may frequently consist of guided music listening which leads to discussion of lyric content, mood of the music, and relevance of the music to the client's personal life. The catalytic music group encourages the the client to become more aware of feelings, thoughts, and experiences that have previously been repressed, through music therapy interventions such as guided listening techniques, guided imagery and music (GIM), and psychodrama.

Music and Movement can be used on many different levels in order to address a variety of goals ranging from awareness and expression of feelings and emotions, to social interaction and physical exercise. *Movement awareness* is a technique which encourages clients to move and express themselves on a non-threatening level through beginning music and movement exercises in a group setting. Another technique called *Movement exploration* uses elements of music and movement to help the client develop a repertoire of body movements as they improve their body image and begin to feel competent in moving effectively. *Movement interaction* is a technique which uses music and movement to provide the opportunity to experience social and emotional concepts in a non-verbal modality. At this level of movement, the focus is on the client and the client's relationship to others. *Expressive movement* is a technique using music and movement to help a client become aware of feelings and emotions that relate to their personal functioning and coping skills in daily life. Musical themes are used to encourage the experience and expression of feeling concepts such a love, loss, grief, depression, social isolation and withdrawal, interpersonal conflict, hope, and joy. *Dance* is another form of music and movement which uses prestructured dance forms, steps, and styles with music to encourage social interaction, self-confidence, and recreational skills. *Music and exercise* can be used with any

age group and level of movement ability to address strength, endurance, muscle tone, flexibility, agility, body control, vital capacity, and cardiovascular efficiency. Music becomes not only a motivator for movement, but also provides a rhythmic structure to pace and cue movement patterns (Thaut 2002).

Music combined with other expressive arts is frequently used in the treatment of mental disorders as a multisensory external stimulus to provide pleasant and motivating reality experiences. The technique *music and fine arts* integrates music with other mediums such as drawing, drama and sculpting in order to allow for expression of feelings and emotions. *Music and writing* pairs music with writing such as poetry or prose, which also provides experiences for expression of feelings and cognitive responses. Both of these techniques can be used in a group or individual setting.

Recreational music programmes can address goals such as non-threatening reality orientation, involvement in goal-directed activity, completion of task, and social participation.

Music therapy interventions may include music games, recreational music, music appreciation awareness, and leisure-time skill development.

Music and relaxation programmes can be very effective at providing an outlet for negative feeling processes, providing feedback and support for movement control, pairing physical exercise with relaxation, increasing frustration tolerance, promoting coping skills, and focusing attention in a structured and calming environment. Numerous music therapy techniques are currently used for relaxation. *Music and progressive muscle relaxation training* pairs relaxing music with the relaxing and tensing of different muscle groups, in order to help a client discriminate between relaxation and tension responses. *Music for surface relaxation* allows a client's preferred music to provide a temporary respite from anxiety-provoking or stressful situations. *Music imagery* encourages a client to explore altered states of consciousness for the purpose of allowing imagery, symbols, and latent feelings to surface from the inner self. *Music-centred relaxation* is a technique which uses music to divert a client's attention from states of psychological and physical tension. Through this technique the client's focus on the music helps them gain a safe, pleasant perceptual experience.

## Future direction

Currently, music therapy practice in mental health has primarily been based on borrowed theories from other disciplines. Thaut (2002) purposed that a major mechanism of the therapeutic effectiveness is embedded in music's unique ability to access affective/motivational systems in the brain. Therefore, a theoretical paradigm of music therapy in psychiatry must integrate musical response models in music perception and music cognition with the understanding of music's influence on non-musical human behaviour, psychologically and neurobiologically, and with concepts of behavioural learning and therapeutic change (Gaston 1968, Thaut 2000).

Research in psychology and the neurosciences has increasingly pointed to the important relationship between affect and cognition. According to Izard *et al.* (1984), emotions are an essential part of cognitive processing. He stated that emotion is the main force for motivation, and motivation is and essential force that drives behaviour and cognition. Rachman (1981) also believed that emotions accompany all cognition, but emotional reactions occur before cognitive processing. The associative network theory purposed by Bower (1981) states that the human memory can be compared to an associative network of events and emotions. In this network, events are considered to be memory units that become associated with certain emotional states experienced while the event is occurring. Thus, when memory units feed in to our emotional state they become attached. Therefore, during cognitive learning, people associate emotion with certain elements of environmental context. These three theories are important in establishing the relationship between mood, behaviour, and cognition.

Other research has compared the effects of musical and non-musical mood-induction techniques. Studies by Clark (1983), Sutherland (1982), Teasdale and Spencer (1984), and Thaut and de L'Etoile (1993) have shown music induction techniques to produce stronger and longer

lasting results than other non-musical techniques in affect modification.

Nielzen and Cesarec (1982) and Steinberg *et al.* (1985), compared the emotional experience of music in normal and psychiatric subjects. Findings of their studies indicated that music communicated the same emotional experience to both normal and psychiatric groups.

Martin *et al.* (1988) and Eifert *et al.* (1988) examined affect modification through music. Martin looked at the relationship between the intensity of a headache and mood, finding that a positive change in affect decreased the intensity of headaches while the negative change in affect increased headache intensity. Eifert, when looking at affect modification through evaluative conditioning with music, identified that liked and disliked music provided strong positive and negative conditioning, respectively.

A cognition–affect model in neuropsychiatric music therapy, based on scientific mechanisms relating the perceptual experience in music to brain and behaviour, could lead to meaningful therapeutic applications within an understanding of the biological basis and the role of affect–cognition processes in mental disorders (Halligan and David 2001).

## Music therapy with children

Music therapy addresses three goal areas in order to increase cognitive, motor, social, communicative, musical and emotional functioning in disabled children: (a) educational, (b) rehabilitative, and (c) developmental (Thaut1999). A music therapist works closely with the interdisciplinary treatment team and must be familiar with developmental, psychological, and medical information that pertains to each disability in order to address these goals appropriately. Music therapists work with children, ranging from early childhood to school-aged. Diagnosis can include autism, cerebral palsy, developmental delays, burn injuries, traumatic brain injury, and other neurological impairments.

Educational goals focus on the academic development of a child. These goals may address social, cognitive, or physical skills. Examples of this may be using a musical mnemonic to teach a child his phone number and address, practicing learning colours and shapes by filling in the blanks of a song, doing creative movement to music, or practicing social skills while participating in a group instrumental playing experience.

Rehabilitative goals work toward restoration or compensatory strategies to improve movement, respiration, posture and sensory perception. For example, using rhythmic auditory stimulation to address gait in a 10-year-old who has developmentally already learned to walk, but recently suffered a traumatic brain injury which has left him with ataxic gait.

Developmental goals strive toward enhancing the normal development of a child by providing normal social, emotional, and sensorimotor experiences through music. Musical experiences are used to submerge the child in normal recreational and leisure experiences that meet the child on their existing functional level. For example, teaching a child how to play a musical instrument, using adaptive equipment, in order for the child to play in the school orchestra.

### Interventions

#### *Motor function*

Music therapy offers a variety of musical interventions and experiences to address motor function in disabled children. The possibilities are limitless for creatively challenging children to reach, balance, stretch, and utilize their upper and lower extremities in functional movements. Musical instruments can provide a specific target for the children to aim towards, defining the parameters of the desired movement. The music therapist uses the music to help facilitate the client's participation while simplifying and readily adapting to the functional level of the child. Simple, repetitive melodies, familiar or not, have been most successful with children. The music therapist's ability to distil the lyrics to an appropriate vocabulary level, or simply create their own, more appropriate musical compositions, leads to much more effective interaction with children of all functional levels.

#### *Communication skills*

Music therapy can play a large role in the development and rehabilitation of both verbal and non-verbal communication skills in children. Music therapist can address disorders such

as apraxia; fluency disorders such as stuttering and cluttering; aphasia; and voice disorders which may result in abnormal pitch, loudness, timbre, breath control, or prosody of speech. Goals in the area of speech and communication address issues such as functional and spontaneous speech, speech comprehension, motor control and coordination essential for articulation, fluency of speech, vocal production and sequencing of speech sounds, and rate and intelligibility. When working with children it is especially important to be aware of their developmental stage when choosing an intervention that would best address their current needs. Developmental Speech and Language Training Through Music (DSLM) is one neurologic music therapy technique which uses developmentally appropriate musical materials and experiences to enhance speech and language development through singing, chanting, playing musical instruments, and combining music, speech and movement.

### *Cognitive learning*

Several music therapy interventions are used to address cognitive learning in children in order to aid in memory, attention, and executive function skills. Musical mnemonics can be an organizational tool, using grouping or chunking through melodic and rhythmic patterns to assist in text recall with children. Many studies have looked at the relationship between music and the performance of autistic children on cognitive tasks such as attention (Kostka 1993), auditory processing (Heaton *et al.* 1988, 2001), affective perception (Heaton *et al.*1999), improvisation (Hermelin *et al.* 1989), local and global processing of music (Mottron *et al.* 2000), auditory discrimination (McGovern *et al.* 1991), and learning and retention of information (Wolfe 1993).

### *Emotional and social development*

Children with disabilities have the same need for healthy emotional and social development and opportunities to express their feelings as do normal children. Because of the experiences of their disabilities, these children may actually have an increased need to cope with feelings such as grief, depression, or loneliness (Thaut 1999). On the other hand, they may also be withdrawn socially, causing them to miss out on the support systems that develop through peer relationships. A combination of a lack of ability to participate in some of the more traditional channels to express emotions and a decreased social support system can become a real dilemma for these children.

Music therapy interventions can provide an effective medium to create rewarding social and emotional experiences for children. In addition to working on physical, sensory, cognitive or speech goals, music can also help to increase self-esteem and provide an outlet for emotional expression through success-oriented experiences at any level of functioning. Music can also be used in a group setting to provide opportunities for peer interaction and support. Learning how to enjoy the recreational use of music or even developing a lifelong musical skill such as playing an instrument are also important in the overall strategy to normalize the life of a disabled child.

## Music therapy in neurological rehabilitation

Neurological rehabilitation is an area in music therapy that is strongly substantiated by the basic science and clinical research coming from scientists in and out of the field of music therapy since the mid-1990s. Advanced training is strongly recommended when working with this population due to the knowledge of neuroanatomy, brain pathology, and medical terminology, that is required by the therapist. A wide variety of techniques are used to address motor control, cognition, and psychosocial behaviour with neurological disorders such as stroke, Parkinson's disease, Huntington's disease, traumatic brain injury, multiple sclerosis, and spinal cord injuries.

### Interventions

The following intervention are based on neurologic music therapy (NMT), a research-based system of standardized clinical techniques for sensorimotor training, speech and language training, and cognitive training. NMT is used in neurologic rehabilitation, neuropaediatric therapy, neurogeriatric therapy, and neurodevelopmental therapy. It is based on the Rational

Scientific Mediating Model (R-SMM), a neuroscience model of music perception and production and the influence of music on functional changes in non-musical brain and behaviour function (Thaut 2005).

### *Sensorimotor rehabilitation*

Music therapy offers a variety of musical interventions and experiences to address motor skills in neurological rehabilitation. Common goals that are addressed include gait and mobility, strength and endurance, coordination, balance and posture and range of motion. These goals are addressed through techniques such a: Rhythmic Auditory Stimulation (RAS), Patterned Sensory Enhancement (PSE), and Therapeutic Instrumental Music Performance (TIMP).

Rhythmic auditory stimulation (RAS) is a neurologic technique used to facilitate the rehabilitation of movements that are intrinsically biologically rhythmical, most importantly gait. RAS uses the physiological effects of auditory rhythm on the motor system to improve the control of movement in rehabilitation of functional, stable and adaptive gait patterns in patients with significant gait deficits due to neurological impairment. RAS can be used in two different ways: (1) as an immediate entrainment stimulus providing rhythmic cues during movement, and (2) as a facilitating stimulus for training in order to achieve more functional gait patterns.

Patterned sensory enhancement (PSE) is a technique which uses the rhythmic, melodic, harmonic and dynamic–acoustical elements of music to provide temporal, spatial, and force cues for movements which reflect functional exercises and activities of daily living. PSE is broader in application than RAS, because it is (a) applied to movements that are not rhythmical by nature (e.g., most arm and hand movements, functional movement sequences such as dressing or sit-to-stand transfers) and (b) it provides more than just temporal cues. PSE uses musical patterns to assemble single, discrete motions (e.g., arm and hand movements during reaching and grasping), into functional movement patterns and sequences. PSE cues movements temporally, spatially, and dynamically during training exercises (Thaut *et al.* 1991).

Therapeutic instrumental music performance (TIMP) is the playing of musical instruments in order to exercise and stimulate functional movement patterns. Appropriate musical instruments are selected in a therapeutically meaningful way in order to emphasize range of motion, endurance, strength, functional hand movements, finger dexterity, and limb coordination (Elliot 1982, Clark and Chadwick 1980). During TIMP, instruments are not typically played in the traditional manner, but are placed in different locations to facilitate practice of the desired functional movements.

### *Speech rehabilitation*

Music therapy can play a large role in the development and rehabilitation of both verbal and nonverbal communication skills. Techniques in music therapy can be used to address disorders such as apraxia; fluency disorders such as stuttering and cluttering; aphasia; and voice disorders which may result in abnormal pitch, loudness, timbre, breath control, or prosody of speech. Goals in the area of speech and communication address issues such as functional and spontaneous speech, speech comprehension, motor control and coordination essential for articulation, fluency of speech, vocal production and sequencing of speech sounds, and rate and intelligibility.

The following eight techniques address speech disorders, based on the current research in this area:

1 Melodic intonation therapy (MIT) is a treatment technique developed for expressive aphasia rehabilitation which utilizes a patient's unimpaired ability to sing, to facilitate spontaneous and voluntary speech through sung and chanted melodies which resemble natural speech intonation patterns (Sparks *et al.* 1974). When using MIT with aphasia, the emphasis is in increasing the linguistic or semantic aspects of verbal utterances.

2 Speech stimulation (STIM) is the use of musical materials such as songs, rhymes, chants, and musical phrases simulating prosodic speech gestures to stimulate non-propositional speech. STIM uses completion or initiation of over learned familiar song lyrics, association of words with familiar tunes, or musical phrases to elicit

functional speech responses (Basso *et al.* 1979). For example, spontaneous completion of familiar sentences is stimulated through familiar tunes or obvious melodic phrases (e.g., 'You are my'..., or 'How are you ...?').

3 Rhythmic speech cuing (RSC) is the use of rhythmic cuing to control the initiation and rate of speech thru cuing and pacing. The therapist may use the client's hand, a drum, or possibly a metronome to prime speech patterns or pace the rate of speech. This technique can be useful to facilitate motor planning for an apraxic patient, cue muscular coordination for dysarthria, or assist in pacing with fluency disorders.

4 Vocal intonation therapy (VIT) is the use of intoned phrases simulating the prosody, inflection, and pacing of normal speech. This is done through vocal exercises which train all aspects of voice control including: inflection, pitch, breath control, timbre, and dynamics. An example would be to sing a five-note scale and gradually move the starting pitch up or down by half steps with a child who has a limited pitch range in their normal speaking voice. This exercise could be further expanded by adding a functional sentence i.e., 'Let's go out and play.'

5 Therapeutic singing (TS) is a technique which involves the unspecified use of singing activities to facilitate initiation, development, and articulation in speech and language as well as to increase functions of the respiratory apparatus. Therapeutic singing can be used with a variety of neurological or developmental speech and language dysfunctions (Glover *et al.* 1996; Jackson *et al.* 1997).

6 Oral motor and respiratory exercises (OMREX) involves the use of musical materials and exercises, mainly through sound vocalization and wind instrument playing, to enhance articulatory control and respiratory strength and function of the speech apparatus. This technique would be used with such populations as developmental disorders, dysarthria, and muscular dystrophy (Hass and Distenfield 1986).

7 Developmental Speech and Language Training Through Music (DSLM) is the specific use of developmentally appropriate musical materials and experiences to enhance speech and language development through singing, chanting, playing musical instruments, and combining music, speech, and movement.

8 Symbolic communication training through music (SYCOM) is the use of musical performance exercises using structured instrumental or vocal improvisation to train communication behaviour, language pragmatics, appropriate speech gestures, emotional communication in non-verbal language systems, that is sensory structured, has strong affective saliency, and can simulate communication structures in social interaction patterns in real time.

### *Cognitive rehabilitation*

Several music therapy interventions have been developed to address cognitive learning, based on research evidence providing clinical support in the role of music to aid in memory, attention, and executive function training. Three areas are addressed in cognitive training techniques: auditory attention and perception training, memory training, and executive function training.

1 Musical sensory orientation training (MSOT) is the use of music, presented live or recorded, to stimulate arousal and recovery of wake states and facilitate meaningful responsiveness and orientation to time, place, and person. In more advanced recovery of developmental stages, training would involve active engagement in simple musical exercises to increase vigilance and train basic attention maintenance with emphasis on quantity rather than quality of response (Ogata 1995).

2 Musical neglect training (MNT) involves active performance exercises on musical instruments, which are structured in time, tempo, and rhythm, with an appropriate spatial configuration of instruments to focus attention to neglected or unattended visual field. It may also involve receptive music listening to stimulate hemispheric brain arousal while engaging in exercises addressing visual neglect or inattention (Hommel *et al.* 1990; Frasinetti *et al.* 2002; Anderson and Phelps 2001).

3 Auditory perception training (APT) is the use of musical exercises to discriminate and identify different components of sound, such as time, tempo, duration, pitch, timbre, and rhythmic patterns, as well as speech sounds. Integration of different sensory modalities such as visual, tactile, and kinaesthetic input are used during active musical exercises such as playing from symbolic or graphic notion, using tactile sound transmission, or integrating movement to music (Bettison 1996; Gfeller *et al.* 1997; Heaton *et al.* 1988).

4 Musical attention control training (MACT) involves structured active or receptive musical exercises, using pre-composed performance or improvisation, in which musical elements cue different musical responses in order to practice sustained, selective, divided, and alternating attention functions (Thaut 2005).

5 Musical mnemonics training (MMT) is the use of musical exercises to address various memory encoding and decoding/recall functions. Immediate recall of sounds or sung words using musical stimuli may be used to address echoic functions. Musical stimuli may be used as a mnemonic device or memory template in a song, rhyme, chant, or to facilitate learning of nonmusical information by sequencing and organizing the information in temporally structured patterns or chunks (Deutsch 1982; Gfeller 1983; Wallace 1994; Claussen and Thaut 1997; Maeller 1996).

6 Associative mood and memory training (AMMT) involves musical mood induction techniques to instate (a) a mood-congruent mood state to facilitate memory recall, or (b) to access associative mood and memory function through inducing a positive emotional state in the learning and recall process (Bower 1981; Dolan 2002).

7 Musical executive function training (MEFT) is the use of improvisation and composition exercises in a group or individually to practice executive function skills such as organization, problem-solving, decision-making, reasoning, and comprehension, within a social context that provides important therapeutic elements such as performance products in real time, temporal structure, creative process, affective content, sensory structure, and social interaction patterns (Dolan 2002).

## Music therapy in the geriatric population

Due to advances in research and increased medical knowledge, the life expectancy of the average American has continues to grow, resulting in an estimated 21 per cent of the population over the age of 65 by 2030. The growing number of elderly has forced an increased interest in the unique challenges associated with aging, particularly among health care professionals.

Music therapy has demonstrated great success with the geriatric population, and is currently used in hospitals, nursing homes, assisted living facilities, psychiatric settings, adult day care, and wellness programmes, in order to address cognitive, sensorimotor, and psychosocial skills. A vast range of both rehabilitative and adaptive goals are addressed through music therapy techniques, including increasing upper and lower extremity range of motion, strength and endurance, and mobility; increasing independence in activities of daily living; improving attention, long- and short-term memory and executive functions; increasing reality orientation; improving verbal and non-verbal communication skills; decreasing agitation and wandering; promoting relaxation; leisure skill training; increasing social interaction and participation; and enhancing reminiscence.

### Interventions

Numerous music therapy interventions are used with the geriatric population to help validate or orient them to their environment. Reality orientation through music can help a disoriented client remember information pertaining to where they are and why, what day it is, and what their schedule will be.

Validation therapy may also be used with a late stage Alzheimer's or dementia patient who is unable to orient to current information, but is perhaps living in a memory of the past. With this technique, the music therapist's role is to meet the patient where they are at, and try to listen and validate their point of view.

Reminiscence in a group or individual setting can be a powerful structural tool in therapy to

help clients remember past events and experiences in their life. Music can effectively set the mood for a discussion, stimulate or validate past memories, and create a theme for dialogue. Reminiscence also promotes social interaction, participation and positive sense of self.

Leisure skill reintegration or training can be another effective use of music in therapy. Many elderly people no longer participate in past leisure interest due to lack of motivation, or decreased physical or cognitive functioning. Reintegrating a client into a past interest or teaching them a new skill can provide opportunities for self-expression, intellectual stimulation, and social integration. A music therapist may teach a client compensatory strategies in order to play a musical instrument that they once played, or utilize their past musical knowledge to teach them a new instrument which is more suited to their current level of functioning. Another form of leisure integration may involve assisting a client in finding community resources in order to pursue their leisure interest (ie. singing in a church choir or attending a performance).

Intergenerational music programmes are frequently integrated in residential care facilities for older adults. Music can be an excellent tool to promote interaction between generations through movement, instrument playing, singing, and performance experiences.

Elderly people frequently experience decrease activity levels due to difficulty with mobility, lack of motivation or even lack of transportation. Research has shown that exercise is an important factor in promoting health and well-being. Lack of exercise can contribute to heart disease, decreased energy levels, and increased risk of falls in the elderly. Music paired with exercise in a group setting can be an effective tool to help motivate, improve coordination and timing of movement patterns, and increase endurance, as well as promote socialization.

## Summary and future directions

The field of music therapy has grown substantially since it was founded in 1950. The advances in research and medical knowledge continue to help explain the therapeutic effects of music on behaviour based on scientific evidence; providing the framework to systematically and creatively transform musical responses into therapeutic responses.

When looking at the use of music in therapy for mental disorders, Thaut (2002) purposed that a major mechanism of the therapeutic effectiveness is embedded in music's unique ability to access affective/motivational systems in the brain. Research in psychology and the neurosciences has increasingly pointed to the important relationship between affect and cognition. Halligan and David (2001), addressed the movement toward a cognition–affect model in neuropsychiatric music therapy, leading to meaningful therapeutic applications within an understanding of the biological basis and the role of affect–cognition processes in mental disorders.

In addition to working on physical, sensory, cognitive and speech goals, music therapy interventions can also help to increase self-esteem and provide an outlet for emotional expression. Music can be used in an individual or group setting to provide and teach overall strategies to normalize the life of a disabled child.

The use of music in therapy in neurological rehabilitation continues to grow, as scientific models provide evidence which explains the therapeutic effects of music on behaviour. Based on the evidence provided in the areas of cognition, speech and language, and motor behaviours, a taxonomy of music therapy techniques can be used to select applications of music in therapy that have predictable therapeutic outcome and benefits for this population.

In 2006, geriatrics accounted for about 15 per cent of the populations served by music therapist. As the elderly population continues to grow, the demand for music therapy will also grow.

Research has shown that music can be a very effective tool in the treatment of the elderly. Interventions such as reality orientation, reminiscence, leisure skill reintegration and training, intergenerational music programmes, and music and exercise, have been implemented with excellent success with this populations.

In the past 15 years, new approaches to scientific research based on a neuroscience model

of music perception and production and the influence of music on functional changes in non-musical brain and behaviour functions, have helped begin to answer the questions that need to be addressed in order for future growth of scientific and medical acceptance of music therapy. In many areas of motor control, cognition, and affective behaviour, research clearly explains the therapeutic effect of music on behaviour based on scientific evidence, therefore providing the framework to systematically and creatively transform music into therapy.

## References


AMTA (2006). *AMTA member sourcebook*. AMTA, Silver Springs, MD

Anderson AK and Phelps EA (2001). Lesions of the human amygdale impair enhanced perception of emotionally salient events. *Nature*, **411**, 305–309.

Basso A, Capitani E and Vigndo LS (1979). The influence of rehabilitation on language skills in aphasic patients. *Archives of Neurology*, **36**, 190–196.

Bettison S (1996). The long-term effects of auditory training on children with autism. *Journal of Autism and Developmental Disorders*, **26**, 361–375.

Bower GH (1981). Mood and memory. *American Psychologist*, **36**(2), 129–148.

Clark C and Chadwick D (1980).*Clinically adapted instruments for the multiply handicapped*. Magnamusic-Baton, St. Louis, MO.

Clark DM (1983). On the induction of depressed mood in the laboratory; evaluation of the Velten and musical procedures. *Advances in Behavior Research and Therapy*, **5**, 27–49.

Claussen DW and Thaut MH (1997). Music as a mnemonic device for children with learning disabilities. *Canadian Journal of Music Therapy*, **5**, 55–56.

Davis, WB, Gfeller, KE and Thaut MH (eds) (1999). *An introduction to music therapy, theory and practice*. Boston: McGraw-Hill.

Deutsch D (1982). Organizational processes in music. M Clynes, ed., Music, *mind and brain*, 119–131. Plenum Press, New York.

Dolan RJ (2002). Emotion, cognition, and behavior. *Science*, **298**, 1191–1194.

Eifert G, Craill L, Carey E and O'Connor C (1988). Affect modification through evaluative conditioning with music. *Behavior Research and Therapy*, **26**, 321–330.

Elliott B (1982). *Guide to the selection of musical instruments with respect to physical ability and disability*. MMB Music, Inc., Saint Louis, MO.

Frasinetti F, Pavani F and Ladavos E (2002). Acoustical vision of neglected stimuli: interaction among spatially convergent audio-visual imputs in neglect patients. *Journal of Cognitive Neuroscience*, **14**, 62–69.

Gaston ET (ed.) (1968). *Music in therapy*. Macmillan, New York.

Gfeller K, Woodworth G, Robin DA, Witt S and Knutson JF (1997) Perception of rhythmic and sequential pitch patterns by normally hearing adults and adult cochlear implant users. *Ear and Hearing*, **18**, 252–260.

Gfeller KE (1983). Musical mnemonics as an aid to retention with normal and learning disabled students. *Journal of Music Therapy*, **20**, 179–189.

Glover H, Kalinowski J, Rastatter M and Stuart A (1996) Effect of instruction to sing on stuttering frequency at normal and fast rates. *Perceptual and Motor Skills*, **83**, 511–522.

Haas F, Distenfeld S and Axen K (1986). Effects of perceived music rhythm on repiratory patterns. *Journal of Applied Physiology*, **61**, 1185–1191.

Halligan PW and David AS (2001). Cognitive neuropsychiatry; towards a scientific psychopathology. *Nature Reviews Neuroscience*, **2**, 209–214.

Heaton P, Hermelin B and Pring L (1988). Autism and pitch processing: a precursor for savant musical ability? *Music Perception*, **15**, 291–305.

Heaton P, Hermilin B and Pring L (1999). Can children with autistic spectrum disorders perceive affect in music? An experimental investigation. *Psychological Medicine*, **29**(6), 1405–1410.

Heaton P, Pring L and Hermelin B (2001). Musical processing in high functioning children with autism. *Annals of the New York Academy of Sciences*, **930**, 443–444.

Hermelin B, O'Connor N, Lee S and Treffert D (1989). Intelligence and musical improvisation. *Psychological Medicine*, **19**(2), 447–457.

Hommel M, Peres B, Pollak P, Memin B *et al.* (1990). Effects of passive tactile and auditory stimuli on left visual neglect. *Archives of Neurology*, **47**, 573–576.

Izard C, Dagan J and Zajonc RB (eds) (1984). *Emotions, cognition, and behavior*. Cambridge University Press, New York.

Jackson SA, Treharne DA and Boucher J (1997). Rhythm and language in children with moderate learning difficulties. *European Journal of Disorders of Communication*, **32**, 99–108.

Kostka M (1993). A comparison of selected behaviors of a student with autism in special education and regular music class. *Music Therapy Perspectives*, **11**, 57–60.

Maeller DH (1996). Rehearsal strategies and verbal working memory in multiple sclerosis. Unpublished master thesis, Colorado State University.

Martin P, Nathan P, Milech D and Van Kappel M (1988). The relationship between headaches and mood. *Behavior Research and Therapy*, **26**, 353–356.

McGovern RF, Berka C, Languis ML and Chapman D (1991). Detection of deficits in temporal pattern discrimination using the Seashore rhythm test in young children with reading impairments. *Journal of Learning Disabilities*, **24**, 58–62.

Mottron L, Peretz I and Menard E (2000). Local and global processing of music in high-functioning persons with

autism: beyond central coherence. *Journal of Child Psychology and Psychiatry*, **41**(8), 1057–1065.

Nielzen S and Cesarec Z (1982). Emotional experience of music by psychiatric patients compared with normal subjects. *Acta Psychiatrica Scandinavica*, **65**, 450–460.

Ogata D (1995). Human EEG responses to classical music and simulated white noise: effects of a musical loudness component on consciousness. *Perceptual and Motor Skills*, **80**, 779–790.

Rachman S (1981). The primacy of affect: some theoretical implications. *Behavior Research and Therapy*, **19**, 279–290.

Sparks RW, Helm N and Albert M (1974). Aphasia rehabilitation resulting from melodic intonation therapy. *Cortex*, **10**, 313–316.

Steinberg R, Raith L, Rossnagl G and Eben E (1985). Music psychopathology; III. Musical expression and psychiatric disease. *Psychopathology*, **18**, 274–285.

Sutherland G, Newman B and Rachman S (1982). Experimental investigations of the relations between mood and intrusive, unwanted cognitions. *British Journal of Medical Psychology*, **55**, 127–138.

Teasdale and Spencer (1984). Induced mood and estimate of past success. *British Journal of Clinical Psychology*, **23**, 149–150.

Thaut MH (1999). Music therapy for children with physical disabilities. In WB Davis, KE Gfeller and MH Thaut, eds, *An introduction to music therapy, theory and practice*, 148–162. McGraw-Hill, Boston, MA.

Thaut MH (2000). *A scientific model of music in therapy and medicine*. MMB Music Inc., Saint Louis, MO.

Thaut MH (2002). Toward a cognitive-affective model in neuropsychiatric music therapy. In RF Unkefer and MH Thaut, eds, *Music therapy in the treatment of adults with mental disorders: theoretical bases and clinical interventions*, 86–99. MMB Music Inc., Missouri, MO.

Thaut MH and De L'Etoille SK (1993). The effect of music on mood-state dependent recall. *Journal of Music Therapy*, **30**, 70–80.

Thaut MH, Schleiffers S and Davis WB (1991). Analysis of EMG activity in biceps and triceps muscle in a gross motor task under the influence of auditory rhythm. *Journal of Music Therapy*, **28**, 64–88.

Thaut MM (2005). *Rhythm, music and the brain*. Taylor and Francis Group, Abingdon, UK.

Unkefer RF and Thaut MH (2002). *Music therapy in the treatment of adults with mental disorders: theoretical bases and clinical interventions*. MMB Music Inc., Missouri, MO.

Wallace WT (1994). Memory for music: effect of melody on recall of text. *Journal of Experimental Psychology: Learning, Memory, Cognition*, **20**, 1471–1485.

Wheeler B (1983). A psychotherapeutic classification of music therapy practices: a continuum of procedures. *Music Therapy Perspectives*, **1**, 8–16.

Wolfe DE and Hom C (1993). Use of melodies as structural prompts for learning and retention of sequential verbal information by preschool students. *Journal of Music Therapy*, **30**, 100–118.

# CHAPTER 48

# Research and evaluation in music therapy

Barbara L. Wheeler

RESEARCH has been an integral part of music therapy since the inception of the National Association for Music Therapy (NAMT) in 1950. The first of the original six objectives of NAMT was 'to encourage and report research projects' (Gilliland 1952, p. v). Gfeller (1995) states:

> Research activities were included immediately in the publications of this fledgling organization. *Music Therapy 1951* (Gilliland 1952), the first Book of Proceedings of NAMT, included a small offering of research abstracts by music therapists and a bibliography of pertinent research from fields such as psychology and physiology. This bibliography was no doubt intended to raise professional awareness of and competence in scientific research. Each subsequent Book of Proceedings (*Music Therapy 1952–1962*) included a section entitled Research, which contained not only abstracts but also philosophical addresses by music therapists, psychologists, and physicians espousing the need for scientific experiments to justify music therapy.
>
> Gfeller (1995, p. 38)

Research was also included prominently in the Certificate of Incorporation of the American Association for Music Therapy (AAMT), founded in 1971, including:

- To stimulate, guide, and direct research of music therapists nationally.
- To publish the results of such research in scientific papers and in professional journals and to authorize publication of the results of such research by other educational institutions without compensation payable to the corporation. ...
- To furnish grants in aid to investigators who propose to undertake study and research in the field of music therapy.
- To promote the exchange of information among the various individuals and institutions engaged in the practice and research of music therapy nationally (Dena Condron, personal communication, 27 January 1993, in Gfeller 1995, pp. 41–42).

The Bylaws of the American Music Therapy Association (AMTA), formed when NAMT and AAMT unified in 1998, include this statement: 'The Association serves as the primary organizational agency for the advancement of education, clinical practice, research, and ethical standards in the music therapy profession' (AMTA 2005b, p. 11). Research was adopted as a strategic priority by the AMTA Board of Directors in 2005, including the development of an operational plan that:

> (a) addresses the direction of research in support of evidence-based music therapy practice and improved workforce demand; and, (b) recognizes and incorporates, where necessary, federal, state and other entity requirements

**Table 48.1** Refereed journals in music therapy published in English*

| Journal | Country of publication | First year of issue |
|---|---|---|
| *Journal of Music Therapy* (JMT) | US | 1964 |
| *Canadian Journal of Music Therapy* (CAM) | Canada | 1973 |
| *Music Therapy: Journal of the American Association for Music Therapy* (MT) | US | 1981(–96) |
| *Music Therapy Perspectives* (MTP) | US | 1982 |
| *British Journal of Music Therapy* (BJMT)† | UK | 1987 |
| *Australian Journal of Music Therapy* (AJMT) | Australia | 1990 |
| *Nordic Journal of Music Therapy* (NJMT) | Norway | 1992 |
| *Annual Journal of the New Zealand Society for Music Therapy* (NZ) | New Zealand | 1994 |

*This list does not include the many creative arts therapy journals and bulletins published internationally nor online publications in music therapy. *The South African Journal of Music Therapy*, no longer in publication, was not consulted.
†Until 1994, this was titled the *Journal of British Music Therapy*.
From Edwards, J. (2005). Developments and issues in music therapy research. In B. L. Wheeler (ed.), *Music Therapy Research*, 2nd edn. Barcelona, Gilsum, NH. Reprinted with permission.

> for evidence-driven research as it relates to practice policy and reimbursement.
>
> Hairston (2005, p. 8)

The AMTA Professional Competencies (AMTA 2003) include:

- Interpret information in the professional research literature;
- Determine if conclusions drawn in a study are supported by the results;
- Demonstrate basic knowledge of the purpose and methodology of historical, quantitative, and qualitative research;
- Perform a data-based literature search;
- Apply selected research findings to clinical practice.

NAMT began publication of the *Journal of Music Therapy* in 1964. This journal has served as an important venue for disseminating music therapy research in the United States and internationally. Other publications that include some research are *Music Therapy: Journal of the American Association for Music Therapy* (published from 1981 to 1996) and *Music Therapy Perspectives.*

Music therapy research has, of course, not been limited to the United States. Edwards (2005) summarizes the refereed journals in music therapy in Table 48.1. This shows the remarkable growth in the number of publications, growing from one in 1970 to two in 1980 and six in 1990 and continuing to the present.

Numerous analyses of the music therapy literature have been done (see Edwards 2005 for a summary). Several analyses (Brooks 2003; Edwards 2005; Gregory 2002) illuminate recent changes in the focus of music therapy research.

Brooks (2003) examined 1521 articles from nine music therapy journals over a 37-year period. She tallied the numbers of research articles classified as quantitative, qualitative, clinical, historical, philosophical/theoretical, and professional.[1] The number of articles published in each of these journals is, of course, related to the length of time that the journal has been

[1] Since the focus of this chapter is music therapy research, clinical and professional articles are not included in this summary, although they were part of Brooks' (2003) analysis.

published and the number of issues a year. Brooks found quantitative research articles to be the predominant category across all journals, with 542 articles. She found 55 historical articles and 136 philosophical/theoretical articles.

Of course, only a portion of music therapy research is published in journals for which music therapy is the primary focus. Many music therapy studies are found in medical publications, including those devoted to nursing, rehabilitation, and neuropsychology. Music therapists also publish in special education, psychology, and psychotherapy publications journals and books and in other interdisciplinary areas. Systematic music therapy reviews are also conducted and published as part of the Cochrane reviews (see http://www.cochrane.org/reviews/index.htm).

## Examples of music therapy research

Music therapy research is approached from various perspectives and thus employs a number of methodologies. The value and use of quantitative and qualitative research has been debated by music therapists (Aigen 1995; Bruscia 1995a; Wheeler 1995). Questions about when each is appropriate and whether and/or when they can be combined are of interest.

One of the considerations when thinking of the differences between quantitative and qualitative research is the purpose of the research. Research conducted within a positivistic paradigm is concerned with cause and effect and uses the natural science model of discovery, while research conducted within a non-positivistic paradigm aims to increase understanding of people's multiple constructions of reality. These are very different positions and goals. The distinctions between quantitative and qualitative research thus go far beyond the type of data with which they deal to encompass differences in philosophy, research interests, and methodology. These different purposes should be kept in mind in the discussion that follows.

Bruscia (2005) suggests that there are three broad topical areas of music therapy research: discipline, profession, and foundational. Much of the discussion that follows will be organized around these three broad topical areas. Because music therapy is an applied discipline, most music therapy research is related to clinical or translational research (National Advisory Mental Health Council 2000; Nunes *et al.* 2002), although some research may be more strictly clinical than others.

*Discipline research*, according to Bruscia (2005), 'includes all those studies that examine the myriad facets of music therapy practice. As such, it deals with clinical topics such as assessment, treatment, and evaluation' (p. 81), areas that are at the heart of music therapy. The AMTA definition of music therapy, 'the clinical and evidence-based use of music interventions to accomplish individualized goals within a therapeutic relationship by a credentialed professional who has completed an approved music therapy program' (AMTA 2005a), points to the importance of data and research for providing evidence-based interventions.

*Profession research* includes

> studies that deal with music therapists and what they do collectively to establish and promote music therapy as a healthcare service. It deals with psychological, socioeconomic, political, legal, and educational aspects of music therapy as an organized profession.
>
> Bruscia (2005, p. 81)

Bruscia suggests that topics of profession music therapy research include employment practices, music therapists (characteristics, attitudes, etc.), professional education and training, professional standards, legislation and public relations, and history and culture.

*Foundational research* includes

> all those studies that relate partially but not completely to the discipline or profession of music therapy. It deals with topics that emanate from related fields (for example, psychology, music, medicine, education) but have important implications for music therapy.
>
> Bruscia (2005, p. 81)

Because music therapy is an interdisciplinary endeavor, topics in these related areas are important. Bruscia says, 'Such research can be considered foundational because the findings often provide empirical or theoretical support for practices, principles, or constructs used in music therapy, but, strictly speaking, were not originally focused on music therapy matters' (p. 81).

Some foundational research is basic research (Borg and Gall 1989) in that it is done primarily to increase knowledge without necessarily having in mind an application of the research findings, but much of it is applied research in areas other than music therapy. Foundational research may study either the discipline or the profession of music therapy.

This section will consider quantitative and qualitative music therapy research.

## Quantitative research

As research is so important to music therapy, it is not surprising that it covers a number of topics. With increased reliance on evidence-based practice,[2] the amount and quality of the research that is available has increased in importance.

One way of making sense of a large number of research studies is through a review. Dileo (2005) suggests that there are two primary types of reviews, qualitative and systematic. A qualitative review involves a narrative presentation of relevant literature, while a systematic review 'involves the application of scientific strategies, in ways that limit bias, to the assembly, critical appraisal, and synthesis of all relevant studies that address a specific clinical question' (Cook *et al.* 1997, p. 376). Quantitative systematic reviews use statistical techniques to combine or evaluate studies, while qualitative systematic reviews summarize studies but do not combine them statistically.

The Cochrane Database of Systematic Reviews (http://www.cochrane.org) includes numerous reviews of various healthcare interventions. While the results of Cochrane reviews of some medical areas (e.g., drug trials) are accepted with little controversy, there is considerable discussion of their applicability to some clinical areas,[3] as well as the implications of their exclusion of certain ways of knowing.[4] Cochrane reviews have been done of music therapy for people with dementia (Vink *et al.* 2003/2006), autistic spectrum disorder (Gold *et al.* 2006), and schizophrenia (Gold *et al.* 2005), and music for pain relief (Cepeda *et al.* 2006).

Some systematic reviews are meta-analyses, in which the researcher combines the numerical results of studies with various research methods and findings and looks for consistencies among these findings (meta-analyses are also often used in Cochrane reviews). The researcher who performs a meta-analysis calculates a standard effect size for each study, giving an indication of the size and variability of the phenomenon under investigation and allowing the studies to then be compared among themselves across all measures and variables. Although concerns have been raised about several aspects of meta-analyses (Bailar 1997; Eysenck 1994; Naylor 1988), meta-analysis is generally considered to be a useful technique for discerning patterns among disparate studies and reaching conclusions. Meta-analyses in music therapy have been done of the following populations and types of treatment: medical and dental treatment (Standley 1986, 2000), medical treatment (Dileo and Bradt 2005), medical treatment of paediatric patients (Standley and Whipple 2003), premature infants in neonatal intensive care units (NICU) (Standley 2002), dementia (Koger *et al.* 1999), children and adolescents with autism (Whipple 2004), children and adolescents with psychopathology (Gold *et al.* 2004), symptoms of psychosis (Silverman 2003), neurologic rehabilitation of upper and lower limbs (Chandra 2005), and stress reduction (Pelletier 2004).

In general, meta-analyses have shown quite positive results for music therapy while the Cochrane reviews, which tend to have more stringent requirements for inclusion of studies and thus generally include far fewer studies, have been less positive. Some results from meta-analyses and Cochrane reviews are discussed below.

### Discipline research

Research and results will be reviewed for autistic spectrum disorders, neurological disorders, and medical problems of children and adults. These are a few of the areas in which music therapy research exists.

[2] Evidence-based practice (EBP) refers to 'integrating individual clinical expertise with the best available external clinical evidence from systematic research' (Sackett *et al.* 1996, p.71). It is also referred to as evidence-based medicine (EBM).

[3] See Hömberg's (2005) discussion which asks, 'How far can principles of EBM practice be applied to training techniques in neurological rehabilitation?' (p. 3).

[4] Morse (2006) discusses the implications of the lack of acceptance of qualitative research in the evidence-based movement.

### *Autistic spectrum disorder*

There is a great deal of clinical literature on music therapy with children with autistic spectrum disorders (see Alvin and Warwick 1991; Nordoff and Robbins 2007) but much less research. Two systematic analyses of music therapy research with autism have been done. Whipple (2004) analysed 12 dependent variables from nine quantitative studies that compared music to no-music conditions during treatment of children and adolescents with autism and concluded that all music interventions, regardless of purpose or implementation, have been effective. A Cochrane review by Gold *et al.* (2006) of all randomized controlled trials or controlled clinical trials that compared music therapy or music therapy added to standard care to placebo therapy, no treatment, or standard care included three small studies. It found music therapy to be superior to placebo therapy with respect to verbal and gestural skills, but without significant effects on behavioural problems. It is clear that additional research on music therapy and autism is needed.

### *Neurological disorders*

A large amount of empirical research has been done on music therapy with neurological problems, including people who have had cerebral vascular accidents (CVA, stroke), traumatic brain injuries (TBI), Parkinson's disease, and cerebral palsy. Many of the studies in this area have been of Neurologic Music Therapy (NMT), a model developed at the Center for Biomedical Research in Music at Colorado State University (http://www.colostate.edu/depts/cbrm/) and based in findings from empirical research that have demonstrated the effects of music on sensorimotor training as well as on speech and language training and cognitive training. As a model that integrates scientific research into the biological effects of music with parallel models of non-music therapy treatment (Thaut 2005), it tries to build a systematic path from basic science research to various stages of translational and clinical research. NMT has been accepted as a proven effective intervention for this population. One of the methods of NMT, Rhythmic Auditory Stimulation (RAS), is listed as one of 10 higher-level evidence-based motor therapies in a critical review of this area (Hömberg 2005). Chandra's (2005) meta-analysis on neurologic rehabilitation of upper and lower limbs, including many NMT techniques, provided evidence for the effectiveness of these procedures.

### *Medical problems*

Meta-analyses have provided important information on the effectiveness of music and music therapy in medicine. Standley did the first of these in 1986 and has updated them several times. In the most recent meta-analysis (2000), she found an overall average effect size (ES) of 0.84 for music versus non-music conditions in all of the 92 studies and 232 dependent variables that were included. When including only the primary dependent variables of the 92 studies, she found an average ES of 1.17 and an average ES of 1.40 in the studies that utilized patients' preferred music. Both are considered large effect sizes (Cohen 1977).

Standley and Whipple (2003) did a meta-analysis of empirical research studies that contrasted music versus no music conditions during medical treatment of paediatric patients. They found a medium effect size (Cohen 1977) in an analysis of 29 studies, indicating that including music is better than not including music in paediatric medical treatment. Standley (2002) also performed a meta-analysis of 10 research studies on the effectiveness of music with premature infants in neonatal intensive care units. She concluded that music has significantly positive effects that remain consistent across a number of variables.

Most recently, Dileo and Bradt (2005) completed a meta-analysis of 184 studies involving music in medical treatment. They included only studies that had a no-music control group and only those dependent variables for which two or more study results were available, and included 47 dependent variables. Effects of music and music therapy were reported for each dependent variable, grouped according to 11 medical specialty areas: premature infants, fetal responses to music, paediatric, obstetrics/gynaecology, cardiology/intensive care, oncology and terminal illness, general hospital, surgery, rehabilitation, dementia, and dentistry. Nearly all showed statistically significant effect sizes, although with a lack of consistency across studies.

## Profession research

Profession research topics to be presented here include ethical issues and employment practices, a sample of the topics for which music therapy research exists. The studies analysing music therapy research trends, discussed earlier, are also examples of profession research. To this author's knowledge, no meta-analyses or systematic reviews of the quantitative profession music therapy literature have been done.

### *Ethical issues*

Most of the research on ethical issues in music therapy has been done by Dileo and her colleagues and has included a broad range of topics. Dileo (2000) wrote a textbook on music therapy ethics, built on her research and that of others, that makes the study of ethics in music therapy accessible to students and others.

### *Employment practices*

Employment practices of music therapists have been investigated periodically for many years, primarily through surveys. These surveys focus on several areas.

Some surveys are of the clinical practice of music therapists. Recent surveys have been of music therapists in specific settings, including school settings (Smith and Hairston 1999), correctional psychiatry (Codding 2002), and private practice (Silverman and Hairston 2005); of music therapy methods in the treatment of early elementary school children with ADHD (Jackson 2003); and of the collaboration and consultation of music therapists (Register 2002). A comparative analysis of surveys focused on hospice administrators' knowledge of music therapy (Hilliard 2004).

Other surveys have focused on the educational process, including clinical training. Maranto and Bruscia (1988) conducted separate surveys of music therapy educators, music therapy clinical training supervisors, music therapy clinicians, and university catalogues to gather information on a number of areas. Jensen and McKinney (1990) studied university curricula by studying the catalogues to compare the requirements in music therapy and related areas with the curricular requirements of NAMT. Groene and Pembrook (2000) surveyed collegiate faculty for their views of new knowledge and skills needed, competency-based assessment, and clinical training practices and potential changes.

Some surveys sought music therapists' opinions of the usefulness of their educational and clinical training for their later work as music therapists. These include surveys on the perceptions of the quality of collegiate training by Gault (1978) and of the relationship of applied music requirements to clinical practice by Cohen *et al.* (1997).

Some surveys were of the tasks and competencies required of music therapists. Some of these have been done and are done on a regular basis by the music therapy organizations and by the Certification Board for Music Therapists (CBMT) and the results are often published in their professional newsletters and other publications, while others have been done by individuals.

Several surveys of professional practices of music therapists have been conducted in the United Kingdom. Hills *et al.* (2000) used the Maslach Burnout Inventory to study how burnout varied among people working in multidisciplinary teams and those who worked independently. Stewart (2000) surveyed UK music therapists to assess personal qualities of music therapists, working models, support networks, and job satisfaction.

## Foundational research

Many studies that are done in disciplines outside of music therapy, with no direct relationship to music, can be foundational to music therapy. In addition, much of the content of the psychology of music that is presented in this book, when it is research-based, can be foundational to music therapy. An overview of types of research in the psychology of music, psychology, music, medicine, and education that can be foundational to music therapy will be given here, although it is beyond the scope of this chapter to provide specific examples.

### *Psychology of music*

Studies of the perception and cognition of music would be included as foundational to music therapy. Similarly, studies of people's responses to music, including physiological, behavioural, emotional, cognitive, and aesthetic, would be in

this area, as would studies of the neural processing of music. Many other topics in the psychology of music could also be included.

### *Music*

Research on various types of music can also be foundational to music therapy. An analysis of a particular piece of music as done by a musicologist might be useful for a music therapist to use to enhance the understanding of music used in music therapy.

### *Psychology*

The same areas that were listed as possible foundational studies for the psychology of music, but without application to music, could begin this section. Thus, we include perception and cognition; physiological, behavioural, emotional, cognitive, and aesthetic responses; and research on neural processing of stimuli. We might also add studies of personality and of the social behaviour of humans. Any area of psychological research can potentially be foundational to music therapy.

### *Medicine*

Research on medical areas is foundational to music therapists who work in those specialties. Examples include research on medical areas such as cancer, cardiac problems, respiratory problems, neurological disorders, and so forth, and issues surrounding medical care such as pain, pre-surgical anxiety, and family support.

### *Education*

Both regular education and special education can be foundational to music therapy. Development and teaching methods are among the areas that may be relevant from education research. Examples of special education research that is foundational to music therapy includes research on the assessment, needs, and education of children who are enrolled in special education classes and on the benefits of inclusion or of children's responses to children with special needs who are in their classes.

## Qualitative research

Qualitative research generally has a different focus than does quantitative research, and the assumptions that underlie it are different. As stated earlier, a general goal of qualitative research is to increase understanding of people's multiple constructions of reality. The purpose of qualitative research is often to lead to a deeper understanding of the music therapy process. Qualitative research goes beyond the use of qualitative data to reflect the beliefs of its followers

> that not all that is important can be reduced to measurements [and] it is essential to take into account the interaction between the researcher and the participant(s) being studied. Findings cannot be generalized beyond the context in which they are discovered, and values are inherent in and central to any investigation.
>
> Wheeler (2005b, p. 13)

Although qualitative research is newer in most disciplines than is quantitative research, its regular use in music therapy began in the 1980s and has grown greatly in the last few decades. A recent book on music therapy research (Wheeler 2005a) includes 12 chapters on types of qualitative research, reflecting the growth of this type of research.

### Discipline research

Qualitative discipline research in music therapy focuses on understanding the client and therapy better. This improved understanding of therapy is important for understanding the music therapy process. A number of studies in this area examine the therapy process and thus contribute to understanding music therapy. Some of these, selected to give an overview of some of the research in this area, are shared here. No attempt is made to summarize their 'results', as results have a different meaning in qualitative research, and a summary of any of the findings would not do justice to the research.

One of the early music therapy qualitative studies led to a theory of a field of play as underlying the improvisational music therapy process (Kenny 1989, 2006). In 'A suspiciously cheerful lady', Aasgaard (2000), looked at what happened to a song after it was composed. Aigen (1997) studied a group of adolescents over a year, including videotaping and analysing the sessions and interviewing the therapists about

the sessions and, in a separate study (Aigen 2002) examined the music therapy process of a young adult with a developmental disability. Gardstrom (2004) and Skewes (2001) studied adolescents in improvisationally based music therapy groups, while Langenberg and her colleagues (Langenberg *et al.* 1993, 1995) developed a set of research procedures to study psychoanalytically informed music therapy.

Several studies have looked at aspects of the music therapist's experience or interaction in the music therapy process; some studies look at both therapists' and clients' experiences. Because the focus of these studies is on the therapy, they are considered examples of discipline research. Bruscia (1995b) examined several modes of consciousness in Guided Imagery and Music (GIM), Forinash (1992) examined therapists' experience of clinical improvisation, Amir (1993, 1996) examined meaningful moments in music therapy, Grocke (1999) studied pivotal moments in GIM, and Comeau (2004) studied music therapists' experiences of being effective and ineffective. Wheeler (1999) looked at her experience of pleasure when working with children with multiple severe disabilities.

### Profession research

The only completed profession qualitative research study of which the author is aware is an investigation of students' experiences during music therapy practice (Wheeler 2002).

### Foundational research

Foundational research can be found in many related disciplines. It has a long history in anthropology; more recent extensive applications in education, psychology, and nursing; and has been used for research in numerous other disciplines, many of which can potentially be useful as foundations for music therapy.

One study that was done by a music therapist but could be considered foundational is Racette's (2004) study of the experience of listening to music when upset. Ferrara's (1984) 'Phenomenology as a tool for musical analysis', a study and set of research methods that focused on music but influenced music therapy research, served as the basis for adaptations of phenomenological research methods to music therapy (Forinash and Gonzalez 1989; Ruud 1987).

## Other types of research

Additional types of research make important contributions to the music therapy literature. These are listed in a recent book on music therapy research (Wheeler 2005a) as music research, philosophical inquiry, and historical research, with an additional chapter that focuses on developing theory. Examples of these types of music therapy research are available in the literature.

# Summary

Music therapy research is an integral part of music therapy and has been since the beginning of the US music therapy associations. Numerous journals, both in music therapy and related disciplines, include music therapy research. Quantitative, qualitative, and other types of research are important in music therapy, each fulfilling different purposes.

# References


Aasgaard T (2000). 'A suspiciously cheerful lady': a study of a song's life in the paediatric oncology ward, and beyond. *British Journal of Music Therapy*, **14**, 70–82.

Aigen K (1995). Principles of qualitative research. In BL Wheeler, ed., *Music therapy research: quantitative and qualitative perspectives*, 283–311. Barcelona, Gilsum, NH.

Aigen K (1997). *Here we are in music: one year with an adolescent creative music therapy group*. MMB, St. Louis, MO.

Aigen K (2002). *Playin' in the band: a qualitative study of popular music styles as clinical improvisation*. Barcelona, Gilsum, NH.

Alvin J and Warwick A (1991). *Music therapy for the autistic child*, 2nd edn. Oxford University Press, Oxford.

Amir D (1993). Moments of insight in the music therapy experience. *Music Therapy*, **12**, 85–100.

Amir D (1996). Experiencing music therapy: meaningful moments in the music therapy process. In M Langenberg, K Aigen and J Frommer, eds, *Qualitative music therapy research: beginning dialogues*, 109–130. Barcelona, Gilsum, NH.

AMTA (American Music Therapy Association) (2003). *AMTA professional competencies*. Available from http://www.musictherapy.org/competencies.html (Accessed 18 July 2006).

AMTA (American Music Therapy Association)(2005a). *Frequently asked questions about music therapy*. Available from:http://www.musictherapy.org/faqs.html#WHAT_IS_MUSIC_THERAPY (Accessed 18 July 2006).

AMTA (American Music Therapy Association) (2005b). *AMTA bylaws*. Available at: http://www.musictherapy.org/membersonly/official/bylaws.html (Accessed 18 July 2006).

Bailar JC (1997). The promise and problems of meta-analysis. *The New England Journal of Medicine*, **337**(8), 559–561.

Borg WR and Gall MD (1989). *Educational research: an introduction*, 5th edn. Longman, White Plains, NY.

Brooks D (2003). A history of music therapy journal articles published in the English language. *Journal of Music Therapy*, **40**, 151–168.

Bruscia KE (1995a). Differences between quantitative and qualitative research paradigms: implications for music therapy. In BL Wheeler, ed., *Music therapy research: quantitative and qualitative perspectives*, 65–76. Barcelona, Gilsum, NH.

Bruscia KE (1995b). Modes of consciousness in Guided Imagery and Music (GIM): a therapist's experience of the guiding process. In CB Kenny, ed., *Listening, playing, creating: essays on the power of sound*, 165–197. State University of New York Press, Albany, NY.

Bruscia KE (2005). Research topics and questions in music therapy. In BL Wheeler, ed., *Music therapy research*, 2nd edn, 81–93. Barcelona, Gilsum, NH.

Cepeda MS, Carr DB, Lau J and Alvarez H (2006). Music for pain relief [Cochrane Review]. *The Cochrane Library*, 2, no pagination. John Wiley and Sons, Chichester.

Chandra P (2005). The effect of sound stimuli on neurologic rehabilitation of upper and lower limbs: a meta analysis. Unpublished master's thesis, Florida State University, Tallahassee, FL. Available at: http://etd.lib.fsu.edu/theses/available/etd-07112005–131201/ (Accessed 8 January 2007).

Codding PA (2002). A comprehensive survey of music therapists practicing in correctional psychiatry: demographics, conditions of employment, service provision, assessment, therapeutic objectives, and related values of the therapist. *Music Therapy Perspectives*, **20**, 56–68.

Cohen J (1977). *Statistical power analysis for the behavioral sciences*, 2nd edn. Academic Press, New York.

Cohen NS, Hadsell NA and Williams SL (1997). The perceived quality of applied music requirements in the vocational practices of professional music therapists. *Music Therapy Perspectives*, **15**, 67–72.

Comeau P (2004). A phenomenological investigation of being effective as a music therapist. In B Abrams, ed., *Qualitative inquiries in music therapy*, Vol. 1, 19–36. Barcelona, Gilsum, NH.

Cook DJ, Mulrow CD and Haynes RB (1997). Systematic reviews: synthesis of best evidence for clinical decisions. *Annals of Internal Medicine*, **126**, 376–380.

Dileo C (2000). *Ethical thinking in music therapy*. Jeffrey Books, Cherry Hill, NJ.

Dileo C (2005). Reviewing the literature. In BL Wheeler, ed., *Music therapy research*, 2nd edn, 105–111. Barcelona, Gilsum, NH.

Dileo C and Bradt J (eds) (2005). *Music therapy and medicine: a meta-analysis of the literature according to medical specialty*. Jeffrey Books, Cherry Hill, NJ.

Edwards J (2005). Developments and issues in music therapy research. In BL Wheeler, ed., *Music therapy research*, 2nd edn, 20–32. Barcelona, Gilsum, NH.

Eysenck HJ (1994). Systematic reviews: meta-analysis and its problems. *British Medical Journal*, **309**, 789–792.

Ferrara L (1984). Phenomenology as a tool for musical analysis. *The Musical Quarterly*, **7**, 355–373.

Forinash M (1992). A phenomenological analysis of Nordoff-Robbins approach to music therapy: the lived experience of clinical improvisation. *Music Therapy*, **11**, 120–141.

Forinash M and Gonzalez D (1989). A phenomenological perspective of music therapy. *Music Therapy*, **9**, 35–46.

Gardstrom SC (2004). An investigation of meaning in clinical music improvisation with troubled adolescents. In B Abrams, ed., *Qualitative inquiries in music therapy*, Vol. 1, 77–160. Barcelona, Gilsum, NH.

Gault AW (1978). An assessment of the effectiveness of clinical training in collegiate music therapy curricula. *Journal of Music Therapy*, **15**, 36–39.

Gfeller K (1995). The status of music therapy research. In BL Wheeler, ed., *Music therapy research: quantitative and qualitative perspectives*, 29–63. Barcelona, Gilsum, NH.

Gilliland E (1952). The development of music therapy as a profession. In E Gilliland, ed., *Music therapy 1951*, v–xvi. North Shore Printers, Waukegan, IL.

Gold C, Heldal TO, Dahle T and Wigram T (2005). Music therapy for schizophrenia and schizophrenia-like illnesses [Cochrane Review]. *The Cochrane Library*, **2**, no pagination. John Wiley and Sons, Chichester.

Gold C, Voracek M and Wigram T (2004). Effects of music therapy for children and adolescents with psychopathology: a meta-analysis. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, **45**, 1054–1063.

Gold C, Wigram T and Elefant C (2006) Music therapy for autistic spectrum disorder [Cochrane Review]. *The Cochrane Library*, **2**, no pagination. John Wiley and Sons, Chichester.

Gregory D (2002). Four decades of music therapy behavioral research designs: a content analysis of *Journal of Music Therapy* articles. *Journal of Music Therapy*, **39**, 56–71.

Grocke DE (1999). A phenomenological study of pivotal moments in Guided Imagery and Music therapy. Unpublished doctoral dissertation, University of Melbourne, Melbourne, Australia. Available at http://www.musictherapyworld.net.

Groene RW and Pembrook RG (2000). Curricular issues in music therapy: a survey of collegiate faculty. *Music Therapy Perspectives*, **18**, 92–102.

Hairston M (2005) Presidential perspectives. In American Music Therapy Association, *Music Therapy Matters*, **8**(3and4), 8. Available at: http://www.musictherapy.org/membersonly/news/MTM8(3).pdf (Retrieved 18 July 2006).

Hilliard RE (2004). Hospice administrators' knowledge of music therapy: a comparative analysis of surveys. *Music Therapy Perspectives*, **22**, 104–108.

Hills B, Norman I and Forster L (2000). A study of burnout and multidisciplinary team—working amongst professional music therapists. *British Journal of Music Therapy*, **14**(1), 32–40.

Hömberg V (2005). Evidence-based medicine in neurological rehabilitation—a critical review. *Acta Neurochirurgica Supplement*, **93**, 3–14.

Jackson NA (2003). A survey of music therapy methods and their role in the treatment of early elementary school children with ADHD. *Journal of Music Therapy*, **40**, 302–323.

Jensen KL and Mckinney CH (1990). Undergraduate music therapy education and training: current status and proposals for the future. *Journal of Music Therapy*, **26**, 158–178.

Kenny CB (1989). *The field of play: a guide for the theory and practice of music therapy*. Ridgeview, Atascadero, CA.

Kenny C (2006). *Music and life in the field of play: an anthology*. Barcelona, Gilsum, NH.

Koger SM, Chapin K and Brotons M (1999). Is music therapy an effective intervention for dementia? A meta-analytic review of the literature. *Journal of Music Therapy*, **36**, 2–15.

Langenberg M, Frommer J and Tress W (1993). A qualitative approach to analytical music therapy. *Music Therapy*, **12**, 59–84.

Langenberg M, Frommer J and Tress W (1995). From isolation to bonding: a music therapy case study of a patient with chronic migraines. *The Arts in Psychotherapy*, **22**, 87–101.

Maranto C Dileo, and Bruscia K (1988). *Methods of teaching and training the music therapist*. Temple University, Philadelphia, PA.

Morse JM (2006). The politics of evidence. *Qualitative Health Research*, **16**, 395–404.

National Advisory Mental Health Council (2000). *Translating behavioral science into action: Report of the National Advisory Mental Health Council Behavioral Science Workgroup*. National Institute of Mental Health, Bethesda, MD.

Naylor CD (1988). Two cheers for meta-analysis: problems and opportunities in aggregating results of clinical trials. *Canadian Medical Association Journal*, **138**, 891–895.

Nordoff P and Robbins C (2007). *Creative music therapy: a guide to fostering clinical musicianship*, 2nd edn, revised and expanded. Barcelona, Gilsum, NH.

Nunes EV, Carroll KM and Bickel WK (2002). Clinical and translational research: introduction to the special issue. *Experimental and clinical psychopharmacology*, **10**, 155–158.

Pelletier CL (2004). The effect of music on decreasing arousal due to stress: a meta-analysis. *Journal of Music Therapy*, **41**, 192–214.

Racette K (2004). A phenomenological analysis of the experience of listening to music when upset. In B Abrams, ed., *Qualitative inquiries in music therapy*, Vol. 1, 1–18. Barcelona, Gilsum, NH.

Register D (2002). Collaboration and consultation: a survey of board certified music therapists. *Journal of Music Therapy*, **39**, 305–321.

Ruud E (1987). Musikk som kommunikasjon og samhandling [Music as communication and interaction]. Institutt for musikk og teate. Unpublished doctoral dissertation, Oslo University, Oslo Norway.

Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB and Richardson WD (1996). Evidence-based medicine: what it is and what it isn't. *British Medical Journal*, **312**(7023), 71–72.

Silverman MJ (2003). The influence of music on the symptoms of schizophrenia: a meta-analysis. *Journal of Music Therapy*, **40**, 27–40.

Silverman MJ and Hairston MJ (2005). A descriptive study of private practice in music therapy. *Journal of Music Therapy*, **42**, 262–271.

Skewes K (2001). The experience of group music therapy for six bereaved adolescents. Unpublished doctoral dissertation, University of Melbourne, Melbourne, Australia. Available at: http://www.musictherapyworld.net.

Smith DS and Hairston MJ (1999). Music therapy in school settings: current practice. *Journal of Music Therapy*, **36**, 274–292.

Standley J (1986). Music research in medical/dental treatment: meta-analysis and clinical implications. *Journal of Music Therapy*, **23**, 56–122.

Standley JM (2000). Music research in medical treatment. In *Effectiveness of music therapy procedures: documentation of research and clinical practice*, 3rd edn, 1–64. American Music Therapy Association, Silver Spring, MD.

Standley JM (2002). A meta-analysis of the efficacy of music therapy for premature infants. *Journal of Paediatric Nursing*, **17**(2), 107–113.

Standley JM and Whipple J (2003). Music therapy for premature infants in the neonatal intensive care unit: HEALTH and developmental benefits. In SL Robb, ed., *Music therapy in paediatric healthcare: research and evidence-based practice*, 1–30. American Music Therapy Association, Silver Spring, MD.

Stewart D (2000). The state of the UK music therapy profession: personal qualities, working models, support networks and job satisfaction. *British Journal of Music Therapy*, **14**(1), 13–31.

Thaut MH (2005). *Rhythm, music, and the brain: scientific foundations and clinical applications*. Routledge, New York.

Vink AC, Birks JS, Bruinsma MS and Scholten RJS (2003/2006). Music therapy for people with dementia [Cochrane Review]. *The Cochrane Library*, 1, no pagination. John Wiley and Sons, Chichester.

Wheeler BL (ed.) (1995). *Music therapy research: quantitative and qualitative perspectives.* Barcelona, Gilsum, NH.

Wheeler BL (1999). Experiencing pleasure in working with severely disabled children. *Journal of Music Therapy*, **36**, 56–80.

Wheeler BL (2002). Experiences and concerns of students during music therapy practica. *Journal of Music Therapy*, **39**, 274–304.

Wheeler BL (ed.) (2005a). *Music therapy research*, 2nd edn. Barcelona, Gilsum, NH.

Wheeler BL (2005b). Overview of music therapy research. In BL Wheeler, ed., *Music therapy research*, 2nd edn, 3–19. Barcelona, Gilsum, NH.

Whipple J (2004). Music in intervention for children and adolescents with autism: a meta-analysis. *Journal of Music Therapy*, **41**, 90–106.

CHAPTER 49

# Music therapy in medical and neurological rehabilitation settings

Anne Kathrin Leins, Ralph Spintge and Michael Thaut

SINCE the 1990s treatment models for the functional use of music in specific clinical pictures were developed under the label 'musicmedicine' or 'medical music therapy'. The new understanding of the biological basis of music and the influence of music on non-musical behaviour and brain function has led to a redefinition of music and its working mechanisms as a therapeutic tool.

Focus so far has been on the theoretically founded application of music in pain therapy and neurological rehabilitation, while the term 'neurological rehabilitation' is not understood in the narrow sense that its semantic meaning might imply. It applies not only to the actual restoration of function, e.g. post-stroke or traumatic brain injury, but also to degenerative illnesses and disorders such as Alzheimer's or Parkinson's disease (Thaut 2005).

## Music therapy in neurological rehabilitation settings

### Introduction

Research of the last 20 years on neuroplasticity and cortical reorganization has led to the development of innovative, neuroscientifically founded treatment models for neurologic rehabilitation. Music as a complex, temporally structured 'sound language' arouses the human brain on a sensory, motor, perceptive–cognitive and emotional level simultaneously and stimulates and integrates neuronal pathways in a music-specific way. This stimulation can enhance and modify behaviour, processing and perception processes by specific and well-directed use in therapeutic approaches, which are based on active training (Thaut 2005).

The therapeutic application of music to cognitive, sensory and motor dysfunctions due to neurologic disease of the human nervous system has been subsumed under the term 'neurologic music therapy (NMT)'.

NMT encompasses a wide variety of standardized treatment techniques which are evidence-based on scientific and clinical research and directed toward functional non-musical therapeutic goals.

Clinical applications of NMT are subdivided into three domains of rehabilitation:

- sensorimotor rehabilitation,
- speech and language rehabilitation and
- cognitive rehabilitation.

In each domain NMT can be applied to treat patients within different clinical fields and disciplines, such as inpatient and outpatient neurologic rehabilitation, neurogeriatrics, neuropaediatrics and neurodevelopmental therapies. Depending on the patient's needs, therapeutic goals are directed towards functional rehabilitative, developmental or adaptive efforts.

### Clinical outcome studies

Clinical research provides evidence that NMT training has a positive effect on relevant outcome criteria. Focus of research so far has been

NMT in motor rehabilitation. In several studies it has shown to significantly improve gait and other movement parameters (e.g. upper extremity function) in hemiparetic motor rehabilitation (cf. Luft *et al.* 2004; Schauer and Mauritz 2003; Thaut *et al.* 1993, 1997, 2002, 2007; Whitall *et al.* 2000) as well as gait training in Parkinson's disease (cf. Bernatzky *et al.* 2004; Del Olmo and Cudeiro 2003; Howe *et al.* 2003; Lim *et al.* 2005; McIntosh *et al.* 1997; Morris *et al.* 2004; Pacchetti *et al.* 2000; Rochester *et al.* 2005; Thaut *et al.* 1996;).

A smaller number of studies has found positive results also in cerebral palsy (Malherbe 1992; Thaut *et al.* 1998), traumatic brain injury (Hurt *et al.* 1998; Kenyon and Thaut 2000) and developmental disabilities (Robertson *et al.* 2002a, b). NMT in gait rehabilitation can be used as a self-contained training protocol but its principles of rhythmic cuing and temporal regulation can also be integrated into other intervention methods.

Research efforts on the effect of rhythm on motor facilitation started with the evaluation of the effect of rhythmic cuing on stride and muscle activation patterns through use of electromyography (EMG) in the gait of healthy individuals (Thaut *et al.* 1992). Improvement in stride symmetry and EMG patterns, particularly the amplitude variability in muscle contractions across the stride cycle, was observed (Thaut *et al.* 1992). Based on these results the immediate effect of rhythm on gait ability in stroke patients was explored (Thaut *et al.* 1993). Results of this study showed that most patients were clearly able to synchronize their step times to the rhythm. In the course of synchronizing gait movements to the rhythm, stride time symmetry and stride length symmetry improved significantly, as did weight-bearing time on the paretic side. Stride time variability decreased to a significant degree. Subjects also showed a more balanced muscular activation pattern on EMG between the paretic and non-paretic leg, and EMG patterns on the paretic side became significantly more consistent.

In order to evaluate the long-term use of NMT in gait rehabilitation, a two-armed parallel group design was used comparing the NMT-method 'rhythmic auditory stimulation' (RAS) (for details see Thaut 2005) to conventional physical therapy gait training (Thaut *et al.* 1997a). Results demonstrated a statistically significant superiority of the RAS-group compared to the physiotherapy group with regard to the kinematic parameters gait velocity and stride length. This positive outcome has been undermined in a multicentre study (Thaut *et al.* 2007). Thaut *et al.* (2007) found a significant superiority of the RAS-group as compared to the physiotherapy group concerning gait velocity, stride length and frequency.

The immediate effect of RAS on repetitive flexion–extension–reaching movements of the paretic arm in hemiparetic stroke patients was the subject of another study (Thaut *et al.* 1998b). Results revealed immediate improved performance with RAS.

As to Parkinson's disease (PD) patients, results of an initial study revealed that patients could synchronize their step patterns to auditory rhythm even when not taking medication and this effect produced significant improvements in mean gait velocity, cadence and stride length over walking without RAS. Based on this data, it was shown in a controlled study design that PD patients profit from a music-based 3-week home-based walking programme of 30 minutes each day (Thaut *et al.* 1996a) and that the therapeutic effect lasts for 3–4 weeks (Mc Intosh *et al.* 1998).

Data on the differential use of metronome vs rhythmic–auditory stimulation in a music context are not consistent. Thaut *et al.* (1997a) have provided evidence in an experimental setting that the temporal synchronization of a sensorimotor task is more precise with rhythmically accentuated musical stimulation than with metronome clicks in healthy subjects. In relation to the improvement of the quality of gait in Parkinson's disease patients, Enzensberger *et al.* (1997), however, found a superiority of metronome stimulation as compared to stimulation by rhythmically accentuated music.

In the course of the development of innovative methods of rhythmical auditory stimulation in neurologic motor rehabilitation, the therapeutic use of music was also applied to speech and language training. Thaut *et al.* (2001) provided evidence for a significant improvement of PD patients with dysarthria as to intelligibility after 'rhythmic speech cuing' (see Thaut 2005). There is also empirical evidence for 'melodic

intonation therapy' with patients suffering from Broca aphasia (Sparks and Holland 1976).

NMT techniques in motor therapy have found their way into the corpus of scientific medical survey literature (cf. Adler and Ahlskog 2000; Hömberg 2005; Hummelsheim 1999; Jeffrey and Gould 1995).

As to the application of music in the therapy of cognitive functions—focusing on memory and attention training—a number of studies have been published within the last years (cf. Chan *et al.* 1998; Claussen and Thaut 1997; Foster and Valentine 2001; Haslam and Cook 2002; Knox and Jutai 1996; Ma *et al.* 2001). These studies broaden the evidence-base for NMT as a therapeutic technique in the rehabilitation of neurologic diseases.

## Neuroscientific foundation of NMT: a working model

Thaut (2005) suggests a preliminary working model of specific mechanisms of music as a rehabilitative stimulus. The aim is to summarize evidence as a step toward a clearer understanding of the biomedical properties embedded in music perception. He advises that these mechanisms cannot be considered as working linearly in a mutually exclusive manner.

Thaut (2005) identifies four such mechanisms or underlying rationales for biomedical application of music and rhythm:

1 *Rhythmic stimulation and entrainment.* Rhythmic stimulation has profound effects as coordinative sensory input to entrain timing functions (cf. Paltsev and Elner 1967; Rossignol and Melvill 1976; Thaut *et al.* 1999; Thaut 2003), especially in motor control and speech. Rhythm provides temporal structure through metrical organization, predictability and patterning. Rhythm regulates physiological and behavioural functions via entrainment mechanisms (cf. Aschersleben and Prinz 1995; Molinari *et al.* 2003; Stephan *et al.* 2002, 2003; Tecchio *et al.* 2000; Thaut and Schauer 1997b; Thaut *et al.* 1998e). Rhythmic entrainment provides immediate time regulation but can also be accessed to enhance long-term training effects. Temporal regulation not only affects timing issues but also provides templates for optimization of spatial and force dynamics (cf. Thaut *et al.* 1996).

2 *Patterned information processing.* Timing is a key component in neural information processing in regard to perception and learning (cf. Galaretta and Hestrin 2001). Rhythm and synchronization are critical parameters in this processing, creating temporal structure in neural network activation (cf. Ikegaya *et al.* 2004; Fries *et al.* 2001). Time cues in the objects and stimuli of perception and learning themselves may be important facilitators to enhance the underlying physiology of temporal pattern formation, thus enhancing critical aspects of cognitive processes (cf. Lee and Blake 1999). Music is a very complex temporal stimulus. In all dimensions of its temporal architecture—from the spectral patterns of its psychoacoustical foundations to the most complex compositional principles of rhythm and polyphony—we find pattern structures of timing and synchronization embedded as building principles (see Kenyon and Thaut 2003; Thaut 1998c). Research suggests that music can uniquely engage the brain as a language of time (cf. Buonomano and Merzenich 1995), providing temporal structure to enhance learning and perception, especially in the areas of cognition, language and motor learning (Thaut 2005).

3 *Differential neurological processing.* The neuroanatomy and neurophysiology of music processing show shared, parallel and distinct neural processing systems for non-musical functions (Thaut 2005). For example, speech and singing are mediated by different neural networks (Patel 2003). Music regulates attention and arousal in the brain in a complex, bilaterally distributed fashion (Coull *et al.* 2004). Learning verbal material through song accesses different neural network configurations than learning through verbal presentation (Thaut and Peterson 2003). Auditory rhythm is a powerful sensory cue that can regulate motor timing and coordination in the presence of a deficient internal timing system in the brain. These are examples of how music can be used as an alternative modality to access functions or provide alternative transmission routes for information processing in the brain.

4 *Affective–aesthetic response: arousal, motivation, emotion.* Music communicates emotion

and meaning through the perception of its intrinsic symbol structure of musical elements as well as through emotional responses that have become connected to it through an associative learning process (cf. Berlyne 1971; Juslin and Sloboda 2001; Meyer 1956). Many theories of the meaning of music see the primary purpose of music as expressing and representing emotions and feelings. Regardless of its precise nature, music is a powerful stimulus that induces emotions (cf. Juslin and Sloboda 2001). As such, it has a strong influence on affective states. Arousal, motivation and emotions have important regulatory functions in behaviour organization, behaviour change and learning. Attention, perception, memory, executive function, physical response and learning can be effectively influenced and enhanced by appropriate affective states (cf. Thaut 2005).

## Summary

Neurologic music therapy encompasses evidence-based methods of the functional application of music in the rehabilitation of neurologic disorders.

One of the most widely implemented clinical areas is its application in sensorimotor therapy, focusing on gait training. These methods are based on neurophysiological and neuroanatomic research on auditory–motor synchronization. It has been demonstrated that rhythm can serve as an anticipatory and continuous time reference onto which the movement is mapped within a stable temporal template. Furthermore, a clinical research basis is rapidly developing in the application of neurologic music therapy in speech/language rehabilitation and cognitive rehabilitation.

## Music therapy and pain therapy

### Introduction

Meta-analyses (Bunt 1997; Standley 1986) underline the effectiveness of music therapy in acute pain. Dileo (2003) finds small to medium effect sizes in children and medium to large effect sizes in adult pain management. A Cochrane review (Cepeda *et al.* 2006) on music for pain relief finds that listening to music reduces pain intensity levels and opioid requirements, but evaluates the overall magnitude of the benefits as small. Spintge (cf. 2000) has provided extensive evidence on the effectiveness of music therapy in pain management an on its psychophysiological effects.

There is anecdotal evidence for the use of both receptive and active music therapy in pediatric pain management, but research studies in this field are scarce. Bradt (2001) finds music therapy to be effective in postoperative pain management of orthopaedic paediatric patients. Results of a pilot study from Barrera *et al.* (2002) reveal that interactive music therapy is effective in reducing anxiety and pain of hospitalized children with cancer. Earlier studies of music therapy during medical invasive procedures provide inconsistent results (cf. Arts *et al.* 1994; Fowler-Kerry and Lander 1987). It seems that with children interactive music therapy approaches during painful medical invasive procedures are preferable to the use of receptive music therapy, because the use of headphones during these procedures may enhance feelings of insecurity.

Music therapy is also applied effectively in chronic pain therapy (Chesky *et al.* 1997; Hillecke 2005; Leins 2006; Müller-Busch 1997; Nickel *et al.* 2005; Rider 1985, 1987; Sedei-Godley 1987; Shorr 1993). It is usually administered within an interdisciplinary framework and aims at improving pain sensation, but also coping with pain and comorbid disorders.

### Concepts and effects of anxioalgolytic music against acute pain and stress in medical settings

Musical stimuli alleviating psychophysiological responses to stress and pain are comprised as anxioalgolytic music (AAM) (Spintge 1983). Medico-functional applications of AAM include, for instance, perioperative stress in surgery and dental care, anaesthesia and pain medicine, intensive care, behavioural disorders, obstetrical care, geriatrics, neurological and psychological motor dysfunctions, and palliative care (Avancini *et al.* 2003, 2005; Droh and Spintge 1987; Leins 2006; Pratt and Spintge 1996, Spintge 1998, 2000, 2007; Spintge *et al.* 1999; Spintge and Droh 1991, 1992). The psychophysical condition of a patient in such situations is characterized by emotional distress, anxiety, and, very often, pain (spintge 1991). The traditional

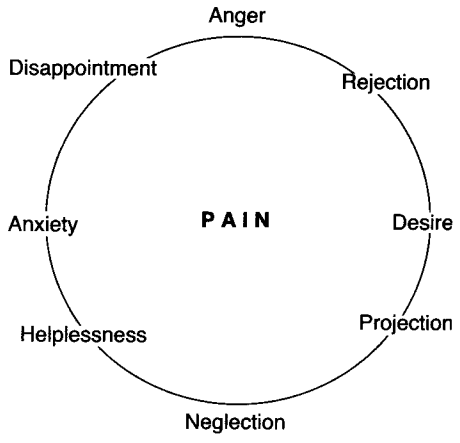



**Fig. 49.1** Circulus vitiosus of pain.

medical interventions, using analgesics and sedatives, for example, often fall short when drugs prove insufficient or reduce the ability of a patient to cooperate. In the following we focus upon acute pain and stress.

The so-called circulus vitiousus of stress and pain (Figure 49.1) describes the psychophysiological situation. Conscious perception, self-esteem, and subjective feeling are as significantly impaired as the central nervous (neurovegetative) regulation of cardiovascular and cardiorespiratory systems. Hormonal control, motor control, sensory and perceptive information processing are disturbed.

Over the course of 30 years in more than 160 000 patients, the routine application of AAM demonstrates that music, as the most intense emotional means of communication, has its special significance as part of a multimodal treatment regime. In this instance, multimodal means to consider illness as a holistic phenomenon, demanding a therapeutic approach, which includes emotional, psychological, social and biological (somatic) aspects (see the so-called biopsychosocial model of pain and the definition of health as given by the WHO).

At present medico-functional music research has identified several target areas for AAM, as explained in Table 49.1. Musical rhythm is an especially effective structural parameter (Koepchen *et al.* 1992; Spintge 1996).

Within a multimodal treatment regime of both acute and chronic distress and pain, music can be used both as a complementary and therapeutic means itself, with impacts upon different levels of neurophysiological processing of pain and stress.

On the one hand, the influence of music (a very complex stimulus) on consciousness is very powerful compared to other environmental stimuli. Averse stimuli are less or even no longer perceived and processed 'under music'. On the other hand, subcortical centres of pain-processing and emotional control, such as the limbic system, are directly influenced. In addition, descending pain-inhibiting structures of the central nervous system are activated in the same way as Melzack and Wall (1965) describe in their (adapted version of the) gate-control-theory. Also, the new concept of neuroplasticity explains the genetical re-programming of dorsal root nerve cells that have been sensitized against pain input before (Zieglgaensberger 1998). These concepts force us to reconsider the role of so-called complementary therapies in general, because we find that the genome of nerve cells can be reprogrammed, thereby changing their functional capacities. Furthermore, the continuous input of similar stimuli for instance through a piece of music enhances this reprogramming.

Another target area for the influence of music upon pain perception lies in musicogenic muscle detonization. Proprioception from skeletal muscles influences emotional state as well as pain perception, and vice versa (David 1985; Gellhorn *et al.* 1964; Simons 1988).

Also, there is a direct pain-decreasing impact of music on reticular formation and thalamus in the central nervous system (Avanzini *et al.* 2003, 2005; David 1985). This process reduces the sensitivity of other areas in the brain to incoming pain signals as well as decreasing the release of stress hormones and catecholamines.

Last, but not least, subjective perception of ongoing, pain-accompanying vegetative activation can be changed through the parallel input of external musical stimuli. By listening to music a situation can be re-construed as being less threatening and less stressful and painful.

Research findings lead to what we call the 'missing link concept' of rhythmicity as a bridge between AAM on one side and physiology and medicine on the other side (Spintge 1996).

Music stimuli are processed through interdependent mechanisms that are as yet not fully understood, by themselves or together. It is clear, however, that the aggregate of all the mechanisms

involved is responsible for the observable psychophysiological effects of AAM music.

Through methodological and technical effort and with the help of up-to-date central nervous system (CNS) monitoring equipment such as functional magnetic resonance imaging (fMRI) and positron emission tomography ( PET), it is possible, however, to scientifically correlate quantitatively and qualitatively physiological parameters and subjective measures with the structure of medico-functional music stimuli (cf. Avanzini *et al.* 2003, 2005; Hodges *et al.* 2005; Sergent 1993).

The practical relevance of all these findings and concepts expresses itself in many ways (see Table 49.1) and leads for instance to improved motor coordination in rehabilitation programmes for low back pain, Parkinson's disease, stroke, etc. (McIntosh *et al.* 1996; Thaut 1997, 2003).

## Music therapy with chronic pain patients

In earlier studies predominantly receptive music therapy approaches are used (Chesky *et al.* 1997; Rider 1985, 1987; Sedei-Godley 1987; Shorr 1993) but within the last ten years active music therapy approaches with chronic pain patients have gained more attention (Hillecke 2005; Leins 2006; Müller-Busch 1997).

Müller-Busch (1997) proved the effectiveness of music therapy for patients with muscle-related pain. Hillecke (2005) found significant effects in the pre-post-comparison and the

**Table 49.1** Music in pain therapy - psychophysiological effects and treatment outcome

| System/Parameter | Reaction |
|---|---|
| cardiovascular | decrease of heartrate<br>decrease of arterial bloodpressure<br>anti arrhythmic |
| respiration | decrease of respiratory minute-volume<br>decrease of oxygen-consumption<br>synchronisation/harmonisation of rhythm |
| inner secretion & metabolism | reduced release of:<br>catecholamines<br>ACTH<br>Cortisol<br>Prolactin<br>β-Endorphin<br>decrease in basal metabolic rate<br>sleep induction |
| immune system | psychomeuronal modulation of neuroendocrine receptors for catecholamines, endorphins, substance P, dopamine leading to:<br>rise in IgA-salive-levels<br>rise in T-cell activity<br>relief of inflammatory pain |
| external secreation & excretion | reduced salivation and perspiration |
| reception & perception | rise in pain tolerance |
| motor control | reduced restlessness<br>reduced muscle tonus<br>muscle spasms dissolved<br>psychomotor facilitation |
| drug consumption | 50% to 100% reduced premedication |
| hospital treatment period | reduced:<br>premature infants intensive care: 3 days<br>surgery in the elderly: 1 day |

group comparison condition in a sample of chronic pain patients. Music therapy enhanced effectiveness in a clinically significant way in pain and psychological measures.

Leins (2006) investigated the effectiveness of music therapy with children with migraine headache in a randomized, placebo-controlled, three-armed parallel group design. A specific music therapy concept has been developed (Nickel *et al.* 2002) which takes particular developmental stages of children into account and integrates the patients family. It focuses on the re-establishment of well-being and on coping with pain and pain-inducing situations/emotions through specific music therapeutic interventions. After therapy and in the follow-up measure the music therapy condition showed a significant placebo superiority with regard to relative reduction of headache frequency (Leins 2006).

## Summary

Music therapy is widely applied in pain therapy. Meta-analyses underline the effectiveness of music therapy in acute pain of adult pain patients. Research focuses on clinical outcome, but also investigates possible underlying mechanisms such as the psychophysiological effects of music. There is large anecdotal evidence for the use of receptive but also active music therapy in paediatric pain management, research studies in this field are however scarce. Music therapy is also applied effectively in chronic pain therapy. It is usually administered within an interdisciplinary framework and aims at improving pain sensation, but also coping with pain and comorbid disorders.

## References


Adler H and Ahlskog J (2000). *Parkinson's disease and movement disorders*. Humana Press, Totawa, NJ.

Arts SE, Abu-Saad HH, Champion GD, Crawford MR, Fisher RJ, Juniper KH and Ziegler JB (1994). Age-related response to lidocaine-prilocaine (EMLA) emulsion and effect of music distraction on the pain of intravenous cannulation. *Pediatrics*, **93**, 797–801.

Aschersleben G and Prinz W (1995). Synchronizing actions with events: the role of sensory information. *Perception and Psychophysics*, **57**, 305–317.

Avanzini G, Faienza C, Minciacchi D, Lopez L and Majno M (2003). The neurosciences and music. *Annals of the New York Academy of Sciences*, **999**, whole issue.

Avanzini G, Lopez L, Koelsch S and Majno M (2005). The neurosciences and music vol 2. From perception to performance. *Annals of the New York Academy of Sciences*, **1060**, whole issue.

Barrera ME, Rykov MH and Doyle SL (2002). The effects of interactive music therapy on hospitalized children with cancer: a pilot study. *Psycho-Oncology*, **11**, 379–388.

Berlyne DE (1971). *Aesthetics and psychobiology*. Appleton, Century and Crofts, New York.

Bernatzky G, Bernatzky P, Hesse HP, Staffen W and Ladurner G (2004). Stimulating music increases motor coordination in patients afflicted by Morbus Parkinson. *Neuroscience Letters*, **361**, 4–8.

Bradt J (2001). The effect of music entrainment on postoperative pain perception in pediatric patients. Dissertation submitted to the Temple University Graduate Board.

Bunt L (1997). Clinical and therapeutic uses of music. In DJ Hargraeves and AC North, eds, *The social psychology of music*, 249–267. Oxford University Press, New York.

Buonomano DV and Merzenich MM (1995). Temporal information transformed into a spatial code by a neural network with realistic properties. *Science*, **267**, 1028–1030.

Cepeda MS, Carr DB, Lau J and Alvarez H (2006). Music for pain relief. *The Cochrane Database Systematic Reviews*, 2.

Chan AS, Ho YC and Cheung MC (1998). Music training improves verbal memory. *Nature*, **396**, 128.

Chesky KS, Russel J, Lopez Y and Kondraske GV (1997). Fibromyalgia tender point pain: a double-blind, placebo-controlled pilot study of music vibration using the music vibration table. *Journal of Musculosceletal Pain*, **5**, 33–42.

Claussen D and Thaut MH (1997). Music as a mnemonic device for children with learning disabilities. *Canadian Journal of Music Therapy*, **5**, 55–66.

Coull JT, Vidal F, Nazarian B and Macar F (2004). Functional anatomy of the attentional modulation of time estimation. *Science*, **303**, 1506–1508.

David E (1985). Clickevozierte Potentiale während der Darbietung unterschiedlicher Musikprogramme am Menschen. In R Spintge, R Droh, Eds, *Music in medicine*, 61–86. Editiones Roche, Grenzach.

Del Olmo MF and Cudeiro J (2003). A simple procedure using auditory stimuli to improve movement in Parkinson's disease: a pilot study. *Neurology and Clinical Neurophysiology*, **2**, 1–7.

Dileo C (2003). A meta-analysis of the literature in medical music therapy and MusicMedicine with an agenda for future research. *Abstractband des VIII. Symposium for Music in Medicine of the International Society of Music in Medicine* (ISMM), 100–101.

Droh R and Spintge R (1987). *Musik in der Medizin* (Music in medicine). Heidelberg-New York: Springer.

Enzensberger W, Oberländer U and Stecker K (1997). Metronomtherapie bei Parkinson Patienten. *Nervenarzt*, **12**, 972–977.

Foster N and Valentine E (2001). The effect of auditory stimulation on autobiographical recall in dementia. *Experimental Aging Research*, **27**, 215–228.

Fowler-Kerry S and Lander JR (1987). Management of injection pain in children. *Pain*, **30**, 169–175.

Fries P, Reynolds JH, Rorie AE and Desimone R (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*, **291**, 1560–1563.

Galaretta M and Hestrin S (2001). Spike transmission and synchrony detection in networks of GABAergic interneurons. *Science*, **291**, 1560–1563.

Gellhorn E (1964). Motion and Emotion. *Psychological Review*, **71**, 457–472.

Gembris H (1997). Psychovegetative Zustände des Musikhörens. *Zeitschrift für Musikpädagogik*, **2**, 59–65.

Haslam C and Cook M (2002). Striking a chord with amnesic patients: evidence that song facilitates memory. *Neurocase*, **8**, 453–465.

Hillecke TK (2005). *Heidelberger Musiktherapiemanual: Chronischer, nicht maligner Schmerz.* uni-edition, Berlin.

Hodges DA, Hairston WD and Burdette J (2005). Aspects of multisensory perception: the integration of visual and auditory information in musical experiences. *Annals of the New York Academy of Science*, **1060**, 175–185.

Hömberg V (2005). Evidence-based medicine in neurological rehabilitation—a critical review. *Acta Neurochir Suppl*, **93**, 3–14.

Howe TE, Lovgreen B, Cody FW, Ashton VJ, Oldham JA, (2003). Auditory cues can modify the gait of persons with early-stage Parkinson's disease: a method for enhancing Parkinsonian walking performance? *Clinical Rehabilitation*, **17**(4), 363–367.

Hummelsheim H (1999). Rationales for improving motor function. *Current Opinion in Neurology*, **12**(12), 697–701.

Hurt CP, Rice RR, McIntosh GC and Thaut MH (1998). Rhythmic auditory stimulation in gait training for patients with traumatic brain injury. *Journal of Music Therapy*, **35**(4), 228–241.

Ikegaya Y, Aaron G, Cossart R, Aronov D, Lampl I, Ferster D and Yuste R (2004). Synfire chains and cortical songs: temporal modules of cortical activity. *Science*, **302**, 559–564.

Jeffrey DR and Gould DC (1995). Rehabilitation of the stroke patient. *Current Opinion in Neurology*, **8**, 62–28.

Juslin PN and Sloboda JA (2001). *Music and emotion: Theory and research.* Oxford: Oxford University Press.

Kenyon GP and Thaut MH (2000). A measure of kinematic limb instability modulation by rhythmic auditory stimulation. *Journal of Biomechanics*, **33**, 1319–1323.

Kenyon GP and Thaut MH (2003). Rhythm-driven optimization of motor control. *Recent Research and Developments in Biomechanics*, **1**, 29–47.

Knox R and Jutai J (1996). Music-based rehabilitation of attention following brain injury. *Canadian Journal of Rehabilitation*, **9**, 169–181.

Koepchen HP, Droh R, Spintge R Abel HH, Kluessendorf D and Koralewski E (1992). Rhythmicity and music in medicine. In R Spintge and R Droh, Eds, *MusicMedicine*, 39–70. MMB, Saint Louis, MO.

Lee SH and Blake R (1999). Visual form created solely from temporal structure. *Science*, **284**, 1165–1167.

Leins AK (2006). *Heidelberger Musiktherapiemanual: Migräne bei Kindern* (Heidelberg Music Therapy Manual: Migraine in Children). uni-edition, Berlin.

Lim I, Van Wegen E, de Goede C, Deutekom M, Nieuwboer A, Willems A and Jones D (2005). Effects of external rhythmical cueing on gait in patients with Parkinson's disease. *Clinical Rehabilitation*, **19**(7), 695–713.

Luft AR, McCombe-Waller S, Whitall J, Forrester LW, Macko R, Sorkin JD, Schulz JB, Goldberg AP and Hanley DF (2004). Repetitive bilateral arm training and motor cortex activation in chronic stroke: a randomized controlled trial. *JAMA*, **292**(15), 1853–1861.

Ma Y, Nagler J, Lee M and Cabrera I (2001). Impact of music therapy on the communication skills of toddlers with pervasive developmental disorder. *Annals of the New York Academy of Sciences*, **930**, 445–447.

Malherbe V, Breniere Y and Bril B (1992). How do cerebral palsied children with hemiplegia control their gait? In M Woollacott and F Horak, eds, *Posture and Control Mechanisms, vol 2*, 102–105. University of Oregon Books, Eugene, OR.

McIntosh GC, Brown SH, Rice RR and Thaut MH (1997). Rhythmic auditory-motor facilitation of gait patterns in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, **62**, 122–126.

McIntosh GC, Rice RR, Hurt CP and Thaut MH (1998). Long-term training effects of rhythmic auditory stimulation on gait in patients with Parkinsonian's disease. *Movement Disorders*, **13**(2), 212.

McIntosh GC, Thaut MM and Rice RR (1996) Rhythmic auditory stimulation as an entrainment and therapy technique: effects on gait of stroke and Parkinson's patients. In RR Pratt, R Spintge, eds, *MusicMedicine*, 145–152. MMB, Saint Louis, MO.

Melzack R and Wall P (1965). Pain mechanisms. *Science*, **150**, 971–979.

Meyer LB (1956). *Emotion and meaning in music.* The University of Chicago Press, Chicago, IL.

Molinari M, Leggio MM, De Martin M, Cerasa A and Thaut MH (2003). Neurobiology of rhythmic motor entrainment. *Ann NY Acad Sci*, **999**, 313–321.

Morris GS, Suteerawattananon M, Etnyre BR, Jankovic J and Protas EJ (2004). Effects of visual and auditory cues on gait in individuals with Parkinson's disease. *Journal of the Neurological Sciences*, **219**, 63–69.

Müller-Busch C (1997). *Schmerz und Musik—Musiktherapie bei Patienten mit chronischen Schmerzen.* Gustav Fischer, Stuttgart.

Nickel AK, Hillecke T, Argstatter H and Bolay HV (2005). Outcome research in music therapy—a step on the long road to an evidence-based treatment. *Annals of the New York Academy of Science*, **1060**, 283–293.

Nickel AK, Hillecke TK, Oelkers R, Resch F and Bolay HV (2002). Musiktherapie mit Kindern mit Migräne. *Psychotherapeut*, **47**, 285–290.

Pacchetti C, Mancini F, Aglieri R, Fundaro C, Martignoni E and Nappi G (2000). Active music therapy in Parkinson's disease: an integrative method for motor and emotional rehabilitation. *Psychosomatic Medicine*, **62**(3), 386–393.

Paltsev YI and Elner AM (1967). Change in the functional state of the segmental apparatus of the spinal cord under the influence of sound stimuli and its role in voluntary movement. *Biophysics*, **12**, 1219–1226.

Patel AD (2003). Language, music, syntax and the brain. *Nature Neuroscience*, **6**(7), 674–681.

Pratt RR and Spintge R (1996). *MusicMedicine* vol. 2. MMB, Saint Louis, MO.

Rider MS (1985). Entrainment mechanisms are involved in pain reduction, muscle relaxation, and music-mediated imagery. *Journal of Music Therapy*, **22**, 183–192.

Rider MS (1987). Treating chronic disease and pain with music-mediated imagery. *Arts in Psychotherapy*, **14**, 113–120.

Robertson SD, Chua R, Maraj BK, Kao JC and Weeks DJ (2002a). Bimanual coordination dynamics in adults with Down syndrome. *Motor Control*, **6**, 388–407.

Robertson SD, Van Gemmert AW and Maraj BK (2002b). Auditory information is beneficial for adults with down syndrome in a continuous bimanual task. *Acta Psychologica*, **110**, 213–229.

Rochester L, Hetherington V, Jones D, Nieuwboer A, Willems AM, Kwakkel G and Van Wegen E (2005). The effect of external rhythmic cues (auditory and visual) on walking during a functional task in homes of people with Parkinson's disease. *Archives of Physical Medicine and Rehabilitation*, **86**(5), 999–1006.

Rossignol S and Melvill J (1976). Audio-spinal influence in man studied by the h-reflex and its possible role on rhythmic movements synchronized to sound. *Electroencephalography and Clinical Neurophysiology*, **41**, 83–92.

Schauer M and Mauritz KH (2003). Musical motor feedback (MMF) in walking hemiparetic stroke patients: randomized trials of gait improvement. *Clinical Rehabilitation*, **17**, 713–722.

Sedei-Godley CA (1987). The use of music therapy in pain clinics. *Music Therapy Perspectives*, **4**, 24–28.

Sergent J (1993). Mapping the musician's brain. *Human Brain Mapping*, **1**, 20–38.

Shorr JA (1993). Music and pattern change in chronic pain. *Advances in Nursing Science*, **15**, 27–36.

Simons DG (1988). Myofascial pain syndromes of head, neck and low back. In R Dubner, GF Gebhart and MR Bond, Eds, *Pain research and clinical management*, 186–200. Blackwell, New York.

Juslin PN and Sloboda JA (2001). *Music and emotion: theory and research*. Oxford University Press, Oxford.

Sparks RW and Holland AL (1976). Melodic intonation therapy for aphasia. *Journal of Speech and Hearing Disorders*, **41**, 287–297.

Spintge R (1983). Psychophysiologische Operations-Fitness mit und ohne anxiolytische Musik (Psychophysiological fitness for surgery with, without anxiolytic music). In R Droh and R Spintge Eds, *Angst, Schmerz, Musik in der Anaesthesie* (Anxiety, pain, music in anesthesia), 77–88. Editiones Roche, Grenzach.

Spintge R (1991). The neurophysiology of emotion and its therapeutic applications in music therapy and MusicMedicine. In C Maranto,ed., *Applications of music in medicine*, 59–72. National Association for Music Therapy, Washington, DC.

Spintge R (2000). Musik in Anaesthesie und Schmerztherapie. (Music in anaesthesia and pain therapy). *Anaesthesie Intensivtherapie Notfallmedizin Schmerztherapie AINS*, 254–261.

Spintge R (2007). MusikMedizinische Forschung Heute und Morgen (Musicmedicine research today, tomorrow). In H.-H Decker-Voigt, ed., *Lexikon Musiktherapie* (Encyclopedia of music therapy), 37–42. Hogreve, Stuttgart.

Spintge R and Droh R (1987). *Music in medicine*. Springer, Heidelberg–New York.

Spintge R and Droh R (1991). Ergonomic approach to treatment of patient's perioperative stress. *Canadian Anesthetist Society Journal*, **35**(3), 104–106.

Spintge R, Besser-Siegmund C, Hodges DA, Nichols C and Schlaefke M (1999). *Schlafstörungen* [Sleep disorders]. Polymedia, Hamburg.

Spintge R. (1996). Physiology, mathematics, music, and medicine: definitions and concepts for research. In RR Pratt, R Spintge, Eds, *MusicMedicine*, 3–13. MMB, Saint Louis, MO.

Spintge R. (1998a). *Verspannungsschmerz* (Myofascial pain). Polymedia, Hamburg.

Standley JM (1986). Music research in medical/dental treatment: meta-analysis and clinical applications. *Journal of Music Therapy*, **23**, 56–122.

Stephan KM, Thaut MH, Wunderlich G, Schicks W, Tellmann L, Herzog H, McIntosh GC, Seitz RJ and Hömberg V (2003). Distinct cerebro-cerebellar circuits underlie temporal adjustment of motor behavior. *Proc Soc Neurosci*, **462**, 8.

Stephan KM, Thaut MH, Wunderlich G, Schicks W, Tian B, Tellmann L, Schmitz T, Herzog H, McIntosh GC, Seitz RJ and Hömberg V (2002). Conscious and subconscious sensorimotor synchronization—prefrontal cortex and the influence of awareness. *NeuroImage*, **15**, 345–352.

Tecchio F, Salustri C, Thaut MH, Pasqualetti P and Rossigni PM (2000). Conscious and preconscious adaptations to rhythmic auditory stimuli: a magnetoencephalographic study of human brain responses. *Exp Brain Res*, **135**, 222–230.

Thaut MH (1996c). Physiological and motor responses to music stimuli. In RF Unkefer, ed., *Music therapy in the treatment of adults with mental disorders: theoretical basis and clinical interventions*, 33–49. Schirmer Books, New York.

Thaut MH (2003). Neural basis of rhythmic timing networks in the human brain. *Annals of the New York Academy of Sciences*, **999**, 364.

Thaut MH (2005). *Rhythm, music, and the brain*. Routledge, New York.

Thaut MH and Peterson DA (2003). A role of theta and alpha synchronization in verbal learning with a musical template. *Proceedings of the Society for Neuroscience*, **194**, 21.

Thaut MH and Schauer ML (1997b). Weakly coupled oscillators in rhythmic motor synchronization. *Proc Soc Neurosci*, **298**, 20.

Thaut MH, Hoemberg V, Kenyon G and Hurt CP (1998b). Rhythmic entrainment of hemiparetic arm movements in stroke patients. *Proceedings of the Society for Neuroscience*, **653**, 7.

Thaut MH, Hurt CP, Dragan D and McIntosh GC (1998a). Rhythmic entrainment of gait patterns in children with cerebral palsy. *Developmental Medicine and Child Neurology*, **40**(78), 15.

Thaut MH, Kenyon GP, Hurt CP, McIntosh GC and Hömberg V (2002). Kinematic optimization of spatiotemporal patterns in paretic arm training with stroke patients. *Neuropsychologia*, **40**, 1073–1081.

Thaut MH, Kenyon GP, Schauer ML and McIntosh GC (1999a). The connection between rhythmicity and brain function. *IEEE Engineering in Medicine and Biology*, **18**, 101–108.

Thaut MH, McIntosh GC and Rice RR (1997a). Rhythmic facilitation of gait training in hemiparetic stroke rehabilitation. *Journal of Neurological Sciences*, **151**, 207–121.

Thaut MH, McIntosh GC, McIntosh KW and Hömberg V (2001). Auditory rhythmicity enhances movement and speech motor control in patients with Parkinson's disease. *Functional Neurology*, **16**, 163–172.

Thaut MH, McIntosh GC, Rice RR and Prassas SG (1992). Effect of auditory rhythmic cueing on temporal stride parameters and EMG patterns in normal gait. *J Neurol Rehab*, **6**, 185–190.

Thaut MH, McIntosh GC, Rice RR and Prassas SG (1993). The effect of auditory rhythmic cuing on stride and EMG patterns in hemiparetic gait of stroke patients. *Journal of Neurologic Rehabilitation*, **7**, 9–16.

Thaut MH, McIntosh GC, Rice RR, Miller RA, Rathburn J and Brault JM (1996a). Rhythmic auditory stimulation in gait training with Parkinson's disease patients. *Movement Disorders*, **11**, 193–200.

Thaut MH, Miller RA and Schauer LM (1998c). Multiple synchronization strategies in rhythmic sensorimotor tasks. Phase vs. period correction. *Biological Cybernetics*, **73**, 241–250.

Thaut M, Leins AK, Rice R, Argstatter H, Kenyon GP, McIntosh GC, Bolay HV and Fetter M (2007). Rhythmic Auditory Stimulation improves gait more than NDT/Bobath training in near-ambulatory patients early post stroke: a single-blind, randomized trial. *Neurorehabilitation and Neural Repair (NNR)*, 5, American Society of Neurorehabilitation, online DOI 10.1177/1545968307300523.

Whitall J, McCombe WS, Waller S, Silver KH and Macko RF (2000). Repetitive bilateral arm training with rhythmic auditory cueing improves motor function in chronic hemiparetic stroke. *Stroke*, **31**(10), 2390–2395.

Zieglgänsberger W and Schadrack J (1998). Neuronal plasticity and pain signalling in the central nervous system. *Acta Anaesthesiologica Scandinavica*, **42**, 9–10.

# PART 11

# Conceptual frameworks, research methods and future directions

Edited by Susan Hallam, Ian Cross and Michael Thaut

# CHAPTER 50

# Beyond music psychology

Adam Ockelford

ALTHOUGH music psychology in the latter part of the twentieth century was dominated by the principles of cognitive–psychological thinking, these never completely drove nor, indeed, constrained this domain of research: even in 1970s, 80s and 90s, before the explosion of activity related to music and the cognitive sciences that has characterized the early years of the twenty-first century, music psychology did not function as a tightly defined academic discipline operating within a single, clear-cut epistemological framework. One can speculate as to the reasons why this should have been so. Partly, no doubt, different views as to what constituted valid music–psychological 'evidence' arose as a consequence of the heterogeneous and constantly evolving nature of musicological discourse, which provided a starting point for some who were seeking to adopt a psychological approach. The lack of an homogenous depth to the corpus of music–psychological research may be attributable to the relatively few participants who were active in the area, and their diverse disciplinary and institutional locations. Also, the field may have attracted visitors whose concerns lay primarily at the periphery of mainstream music–psychological work because they could find no other home for their theoretical or empirical endeavours. Finally, the very interdisciplinary nature of music psychology itself meant that as a conceptual territory it had extensive borders that offered many opportunities for intellectual interlopers to cross.

It is notable that the first journal to be devoted exclusively to the subject (*Psychology of Music*, which dates from 1973)[1] was founded by the (then) Society for Research in Psychology of Music *and Music Education*,[2] conferring on the publication a multidisciplinary focus that it has maintained ever since. Today, the stated aim of *Psychology of Music* is to 'increase scientific understanding of all psychological aspects of music and music education', a desideratum which 'includes studies on listening, performing, creating, memorizing, analysing, learning and teaching as well as applied social, developmental, attitudinal and therapeutic studies'.[3] This breadth of intention is fully reflected in the journal's content. Of the 471 papers published between 1973 and the first half of 2007, an informal classification suggests that only 207 (44 per cent) can reasonably be defined as falling within the realm of cognitive psychology (including the measurement of musical abilities, the perception of musical sounds, the cognition of musical structures, learning, memory and the development of music-related skills). Sixty papers (12.5 per cent) are concerned with issues pertaining to performing (including improvisation and performance anxiety). Notwithstanding the journal's stated aims, only 55 articles (11.5 per cent) relate directly to music education, although a further

[1] Until that time publication was solely in mainstream psychology journals.
[2] Italics added; from 2003 known as 'SEMPRE'—the Society for Education, Music and Psychology Research.
[3] From the journal's 'aims and scope', *Psychology of Music*, 35(2), 2007.

15 (3 per cent) investigate the use and potential effect of music in wider educational or developmental contexts. Fifty-four (11.5 per cent) explore aesthetics or affective response to music in one form or another. Thirteen articles (3 per cent) are primarily ethnomusicological in content. Eleven (2.5 per cent) engage with issues of epistemology or methodology (matters to which we shall return). Eight (1.5 per cent) involve research in music therapy. Seven (1.5 per cent) have a music–theoretical focus. The remaining 41 (9 per cent) are not readily categorizable, by virtue of small numbers or idiosyncratic content—ranging from the philosophy of melody to a consideration of the potential relevance of music to extra-terrestrials!

*Music Perception*, which first appeared in 1983 under the editorship of Diana Deutsch at the University of California, while having cognitive–psychological research as its primary focus, was always more explicit than *Psychology of Music* about its wide-ranging academic purview, which has changed little over the last quarter of a century. Today, *Music Perception* is overtly interdisciplinary in nature, with an editorial policy that seeks to publish 'theoretical and empirical papers, methodological articles and critical reviews concerning the study of music', incorporating articles from a broad range of disciplines, including 'psychology, psychophysics, neuroscience, music theory, acoustics, artificial intelligence, linguistics, philosophy, anthropology and cognitive science'.[4] Similarly, *Musicae Scientiae*, the journal of the European Society for the Cognitive Sciences of Music (ESCOM), which first appeared in 1997, is interdisciplinary in nature too, although cognitive psychology is, once again, the predominant strand in its fabric. *Musicae Scientiae* accepts 'empirical, theoretical and critical articles directed at increasing understanding of how music is perceived, represented and generated'. Consideration is given to any 'systematic work within the domains of psychology, philosophy, aesthetics, music analysis, historic musicology, cognitive science, education, artificial intelligence, modelling and neuropsychology'[5,6]

The broad church espoused by the three specialist journals currently active in the field of music psychology is reflected in the conferences and seminars organized by its proponents. Chief among these, the International Conference of Music Perception and Cognition (ICMPC), which was first staged in 1989, and since 1992 has occurred biennially, attracts a remarkable breadth of contributions.[7] For example, ICMPC 9, which was held in Bologna in 2006, called for presentations on the following topics: pitch and tonal perception; rhythm, metre, and timing; aesthetic perception and response; computational models; timbre and orchestration; emotion in music; memory and music; neuroscience; development; education; music, meaning, and language; performance and composition; the singing voice; acoustics and psychoacoustics; the social psychology of music; cognitive musicology; and music therapy. These were realized in some 526 papers, posters, workshops and symposia (a fivefold growth over the 17 years since ICMPC 1).

Surely (it could be argued), such eclecticism is unreservedly a good thing, presumably resulting in a rich cross-fertilization of ideas and approaches, yielding new, broad-based research that is rooted in a range of epistemologies and utilizing a variety of methodologies? Actually the position turns out to be much more complicated than this: as we shall see, conceptual cross-pollination often proves to be difficult to pull off; and where some see epistemological hybridization as a strength, others are wary of diluting the relative purity of purpose and procedure characteristic of a single discipline.

Given such issues, it is hardly surprising that the manner in which music psychology could

[4] From 'Information for authors', *Music Perception*, **24**(4), 2007.

[5] From 'Information for authors', *Musicae Scientiae*, **11**(1), 2007.

[6] The interdisciplinary output of *Psychology of Music*, *Music Perception* and *Musicae Scientiae* contrasts with that of the fourth specialist publication in the field, *Psychomusicology*, which, styling itself as a 'journal of music cognition', first appeared in the US in 1981 and continued to appear until 1997. Throughout these 16 years, *Psychomusicology* adhered closely to its initial editorial brief.

[7] See http://www.icmpc.org/organisation.html#history for a brief history of the ICMPC series.

and should relate to its sister disciplines has exercised a number of those working in these fields over the years. An early and perhaps inevitable tension that emerged in *Psychology of Music*, which was articulated by John Sloboda (then editor) in his 'open letter' of 1986,[8] was between music psychology and music education. In particular (to make explicit what was implicit in Sloboda's statement), one could sense a certain discomfort in his having to consider for publication certain music *education* research within the context of a music *psychology* journal, with its rather different perceptions of what constituted an appropriate level of rigour and objectivity. With their distinct epistemologies, could the two disciplines every work together productively? Was the latter ever likely to be of value to the former? A number of responses to Sloboda's letter were made, including one by David Hargreaves,[9] who, in looking back over the first decade of *Psychology of Music*'s output, acknowledged that 'Contributions tend to be *either* psychological *or* educational, and those which combine theoretical and practical concerns tend to be few and far between.' He concluded that 'the end result is something of a shotgun wedding'. Hargreaves suggested that the most fruitful way forward might lie in the *developmental* psychology of music, an area to which he gave a focus and impetus in the mid-1980s, and which has flourished ever since—particularly in relation to the early years (see, for example, Deliège and Sloboda 1996; the special issue of *Musicae Scientiae* 1999/2000; McPherson 2006) and, more recently, special educational needs (for example, Pring and Ockelford 2005; Ockelford *et al.* 2006; Ockelford 2008). Beyond developmental psychology, in the 1990s, a number of new fronts opened up between music psychologists and educationists in what has proved to be a highly fruitful union (Parncutt and McPherson 2002; Williamon 2004). For example, research of relevance to those learning to perform has been undertaken in the areas of practice and the acquisition of expertise (Hallam 1995, 2001; Jørgensen 2004); memorization (Ginsborg 2007); sight-reading (Lehmann and McArthur 2002; Thompson and Lehmann 2004) and improvisation (Kenny and Gellrich 2002).

Returning, for a moment, to the 1980s, further contributions followed in *Psychology of Music* on other interdisciplinary issues, including the report of a seminar held at The City University, London, in 1987 concerning the relationship between music therapy and music psychology (whose discussants were Leslie Bunt, Ian Cross, Eric Clarke and Sarah Hoskyns).[10] With echoes of the perennial quantitative/qualitative dichotomy, key elements in the debate were the relevance to therapy of what was measurable in psychological terms, and, conversely, the psychological status of the inferences that therapists were content to accept as admissible evidence. It was agreed that what mattered was whether the insights so gained—which, ideally, should be couched within a theoretical framework—were relevant and interesting rather than ultimately 'provable'. The long-term impact thinking such as this can be felt today in collected editions such as *Microanalysis in Music Therapy* edited by Thomas Wosch and Tony Wigram (2007). Here, as Barbara Wheeler notes in the Foreword (p. 11), the common thread linking the contributions is a new-found rigour in analysing therapy sessions, involving the study of 'specific responses and experiences and precise musical and behavioral responses and interactions' (see also Ockelford 2008).

During the 1990s, comparable interdisciplinary discourse rumbled on in other areas too, leading, for example, to Elizabeth Tolbert's (2001) exploration of the evolution of musical meaning through an attempt to bridge ethnomusicological and psychological approaches (between which she considered there had been little rapprochement up to that time). In particular, she pointed to differing perspectives on the relative importance of individual as opposed to collective meaning, and divergent views as to the significance of universal rather than culturally embedded musical processes and structures.

Arguably the interface about which there has been most vociferous debate in recent years,

[8] *Psychology of Music*, **14**(2), 144–145.
[9] *Psychology of Music*, **14**(2), 83–96.
[10] *Psychology of Music*, **16**(1), 62–70.

however, is that between music psychology and music *theory* (as it is generally known in the US) or music *analysis* (the term most widely used in the UK). Why the contention? Perhaps because proponents from the two camps feel that they have an equal claim over a common territory: an understanding of how music 'works'. Their aims, though, are quite different, and this appears to be where the difficulties arise. Writing in 1989, Eric Clarke put it like this:

> Broadly speaking, the aim of musicologists and composers in tackling issues of musical structure can be characterized as the attempt to formulate theories of the structural relations within and between musical works, and their origins, development and effectiveness as formal devices. A correspondingly brief summary of the aim of psychologists of music is the development of theories of the mental processing of musical events, or the relationship between the listener, performer or composer and the musical environment. In a number of respects these aims are quite complementary, but the different disciplines that they represent come into conflict in the way in which they describe their material, and in what they extract and evaluate as significant findings.
>
> Clarke (1989, pp. 1–2)

At times this conflict has become quite heated. In his 'Fortenotes' that appeared in *Music Analysis*, 17(2) (a tribute volume to the Yale music theorist Allen Forte, who introduced 'pitch-class set analysis' to musicology in his seminal book *The structure of atonal music* [1973]), Jonathan Dunsby satirizes music psychologists as would-be enforcers of an imaginary law which dictates that musicological validity necessarily equates to perceptibility:

> I cannot be alone in having taught, *StrAMly* [*The structure of atonal music*], surreptitiously, 'hunt the hexachord'. That's the way you make a set complex work, asking a student to interrogate whether that embarrassing challenger-set really mattered so much and could not perhaps be excluded as a feature of the music, or whether there were not many more lurking hexachords that s/he had heard/seen (I am almost tempted to add '/played', but presumably in this forum I can write shielded from the Perception Police).
>
> Dunsby (1998, p. 179)

In order to understand how academic discourse on music could have become so polarized, one has to appreciate that groups within the music-theoretical community had felt themselves under attack from music psychologists who had shown, for example, that the perception of octave equivalence in pitch sequences—a 'given' in the composition of serial music and pitch-class set analysis—was by no means a 'given' to the musical ear operating in the absence of a score (Deustch and Boulanger 1984),[11] and that the measures of similarity between pitch-class sets as developed by Robert Morris (1979–1980)—measures that lay at the heart of this music-theoretical enterprise—did not accord with similarity judgements made aurally.[12]

However, the traffic was not all one way. Still on the subject of differing perceptions of similarity, in 1994, Nicholas Cook launched a scathing attack on Rita Wolpert's (1990) research, in which musicians and 'non-musicians' (so-called) were asked, among other things, to compare a tune and accompaniment played (a) on a different instrument and (b) on the same instrument as the original, but with the accompaniment transposed down a fifth. The musicians consistently chose option (a)—for them, playing the accompaniment in the wrong key made a bigger difference than playing the music on a different instrument—whereas the non-musicians almost exclusively opted for (b): for them, the identity of the instrument outweighed any changes they noticed in the accompaniment. According to Wolpert, these findings show that musicians do not listen in the same way as non-musicians: their choice of instrumentation over correct harmonic accompaniment 'suggests a profound overestimation of what most listeners hear'. As Cook points out, though, this is a far-fetched conclusion: what Wolpert's experiment actually reveals is that listeners with different backgrounds respond in different ways to questions

[11] See also the psychologically inspired critique of serialism by Fred Lerdahl (1992).
[12] Cheryl Bruner (1984).

as to whether one musical extract is more or less like another (1994, p. 68).[13]

These two examples bring sharply into focus just how different the aims and values of music psychology and music theory continued to be in the years that followed Eric Clarke's (1989) exhortation to researchers to seek to establish a rapport between them.[14] In 2003, David Temperley summarized the ongoing division thus:

> music psychology tends to focus on how people typically hear (or play or compose) pieces, tending towards generalities or commonalities; whereas music theory and analysis usually seek to discover what listeners could (or should) hear, and bear largely on specific compositions.[15]

In similar vein, in reviewing Kevin Korsyn's *Decentering Music* of 2003, Elizabeth Margulis (2005) wrote:

> Music cognition tends to explore those aspects of musical experience that are relatively robust and shared across large populations (betraying a dependence on what Korsyn sees as the problematic construct of 'normalcy'), rather than those that are unique and more amenable to the committed introspection of a single listener. ... Music analysts who rely on introspection as a methodology might manifest a commitment to music as an individual experience, constructed as fully by the listener as by the composer and performer. This vision elevates the specialist, and promotes the importance of training. Researchers who rely on empirical methodologies might reveal a commitment to music as more of a shared experience, with invariant features that characterize the hearing of a neophyte as much as a person with decades of training.
>
> Margulis (2005, pp. 334–335)

Joshua Mailman (2007), in his review of Adam Ockelford's *Repetition in music: theoretical and metatheoretical perspectives* (2005a) refers to the music-cognitive approach as 'populist', as opposed to the 'progressive' tack taken by theorists and analysts, which seeks to extend the boundaries of musical understanding, rather than defining what may be 'typical' or 'usual'. Mailman cites Joseph Dubiel (1999) in support:

> The crucial condition for any increase in musical knowledge is to keep yourself ready to be struck by aspects of sound that you weren't listening for. [This means that] the value of analyses will ultimately be their value as ear-openers.
>
> (p. 274)

In accordance with this sentiment, Mailman takes Ockelford to task for his critique of David Lewin's analysis of the opening of the development section of Mozart's Symphony No. 40, K. 550, which had appeared in Lewin's landmark text *Generalized musical intervals and transformations* (1987)—a mathematically based theory of musical structure. The passage in question, Lewin had observed, can be interpreted as a chain of retrograde inversions (RICH); see Figure 50.1.

However, the ontological status of this pattern is unclear. As Ockelford (2005a, p. 99) asserts: 'There is no evidence that Mozart conceived the passage in this way, nor that listeners perceive it so, nor even that analysts typically construe the sequence as being structured thus,[16] although there are precedents.'[17] Ockelford hypothesizes that listeners would be more likely to make sense of the passage by (subconsciously) modelling its structure as a series of transpositions: a more direct interpretation than Lewin's, which requires less cognitive manipulation of the musical information that is available.

Mailman, however, considers Ockelford's approach to the passage to be 'too narrowly conceived' (op. cit, p. 369), contending that he

> overlooks what we gain from Lewin's discovery of RICH in Mozart's symphony: when we notice RICHs in Webern's 12-tone works, we

[13] Ironically, Cook himself comes in for comparable methodological criticism in relation to his foray into experimental psychology which investigated the perception of large-scale tonal closure (1987); see Gjerdingen (1999).

[14] In 'Mind the gap'—see note 10.

[15] Personal communication to Adam Ockelford, cited in Ockelford (2005b).

[16] See, for example, Saint-Fox (1947), Dearling (1982) and Abert (1990).

[17] For instance, Keller (1966, p. 97).

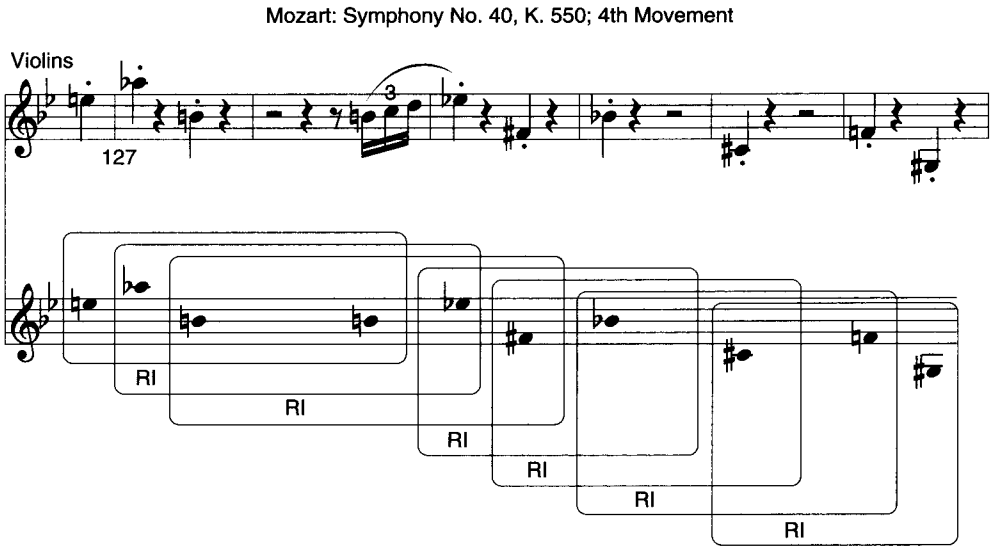



**Fig. 50.1** 'RICH' in Mozart, K. 550 (after Lewin 1987, p. 230).

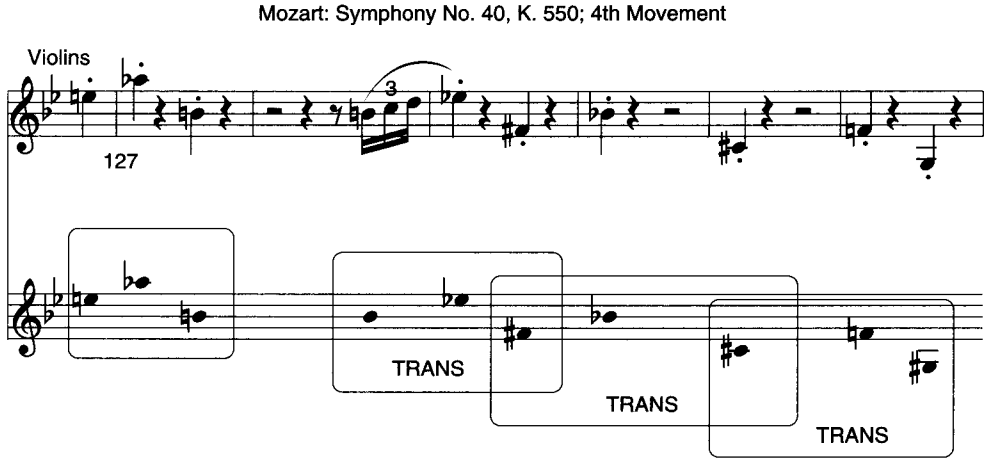



**Fig. 50.2** More straightforward interpretation of the structure of the passage as a series of transpositions (after Ockelford, 2005a, p. 102).

> can relate these conceptually to the RICHs in Mozart—a lucrative inter-repertoire link ... Abstractions like RICH push beyond the everyday experience of repertoire that the 'populist' approach often assumes.

However, as Ockelford points out, in a conceptual thread that weaves its way through *Repetition in music*, there is a serious problem with this line of thinking. To understand why, let us first take a step back to the theory that lies at the heart of the volume: namely, that all internal musical structure relies ultimately on repetition, in one manifestation or another. For example, as Sloboda noted in his first book *The musical mind: the cognitive psychology of music* (1985), for music perception to 'get off the ground', there is a need for a framework of discrete and re-identifiable locations in pitch and perceived time, without which the dialectics of tension/resolution and motion/rest could not

exist (pp. 154 and 259). That is to say, in order for our perceptual and cognitive processing abilities not to be overwhelmed, composers have to work within tight constraints, whereby the number of different categories of pitch, interval, and the time between the onsets of successive notes is limited. This means that a high degree of 'background' repetition is inevitable in any piece. Furthermore, while the burden of the musical message tends to be conveyed by characteristic combinations of pitch and rhythm, further background restrictions typically apply to other qualities of perceived sound too, such as timbre and loudness. These almost invariably fulfil a secondary role as 'carriers' of the principal stream of information, and as a consequence tend towards coherence based on uniformity or incremental change—features which, once more, are founded on repetition.[18] Hence, behind the creation of every work lie constraints that mean that many musical events, and the relationships between them, will be the same, regardless of the subsequent choices of the composer. A key issue, therefore, for those trying to fathom how we make sense of musical structures is not so much about the discovery of sameness and similarity per se, but of analysing the *significance of commonality in different contexts.*

To give an idea of the scale of this issue, take, for example, Chopin's Prelude in B minor, Op. 28, No. 6, which comprises 403 separate notes that are played on the piano, typically within a period of two to three minutes. If one considers the relationships between (any) pairs of pitches the same as being of potential structural significance, then the analyst is faced with around 13 000 candidates. If it is the relationships between pairs of *intervals* the same that are thought to be of possible structural relevance, then she or he would have 500 *million* to choose from. And this is in just one domain: pitch.

Arresting statistics, one might think, but of no possible musicological value, since they ignore two key factors in the creation of musical structure: the sequence of events and their associated rhythm. Yet that is exactly what Allen Forte's set theory, alluded to above, does. The theory was originally intended to offer a mechanism through which the structure of the atonal music of Schoenberg, Webern and Berg and other composers written at the beginning of the twentieth century could be explained systematically. In undertaking set-theoretical analysis, one makes the initial assumption that all pitches and intervals (unlike those in tonal music) are potentially equal in structural terms. Hence the immediate analytical challenge posed is one of scale: how to reduce the available data to manageable proportions. To tackle this problem, a process of reduction was devised that borrowed a number of concepts from Schoenberg's 'serial' compositional procedures. This holds that one set of pitches can be regarded as equivalent to another, irrespective of transposition or inversion, the octave in which values are realized, whether or not they are repeated and, additionally, the order in which they occur. Ferreting out equivalent sets (as the citation from Dunsby above indicates) is fundamental to this type of analysis. But how is it to be done? Although some sets can be isolated as units 'by conventional means, such as a rhythmically distinct melodic figure' (Forte 1973, p. 83), such techniques do not necessarily 'adequately reveal structural components', since methods of segmentation may be 'concealed' (ibid.). In order to uncover the hidden organization of pitch, Forte recommends a procedure termed 'imbrication': 'the systematic (sequential) extraction of subcomponents of some configuration'—that is, listing all the pitch-class sets contained within a passage in the hope that this will reveal relationships that were not otherwise apparent.[19]

While it is conceivable that such a process may uncover pitch structures of musical interest that listening alone would have failed to detect (and which may inform subsequent audition), there remain a number of problems with this technique. For example, sets may not have been apparent in the first instance for the very good reason that they could not be extracted perceptually from large and complex aggregations of notes—and how, therefore, could they *ever* be structurally relevant to any listener, novice or expert? Again, the scale of the enterprise presents a huge problem. Take, for example, Schoenberg's

[18] See, for example, Boulez (1963/71, p. 37); Erickson (1975, p. 12).

[19] See, for example, John Roeder's (1988) analysis of Webern's Piece for Cello and Piano, Op. 11, No. 3.

*Three Piano Pieces*, Op. 11, No. 1—one of the first atonal works to be composed. The opening 4½ bars comprise 24 events in the form of a right-hand melody supported with discrete chords and traces of counterpoint, together lasting no more than 10 seconds. Forte (1981) undertakes an analysis, in which he identifies 14 different pitch-class sets, occurring in total 28 times. However, by adopting a systematic approach, using Forte's recommended technique of imbrication, it becomes apparent that all of the 208 possible pitch-class sets are actually present! These occur with mind-numbing frequency: for example, there are 262 appearances of 3-note sets, 884 4-note sets, 3152 5-note sets and (preliminary analysis suggests) over 10000 6-note sets. How is the analyst ever to make sense of these data? Presumably by reverting to the musical intuitions that were abandoned in the first place in the interests of scientific rigour!

Although analysis of this type occupies only one region in the broad domain of music theory—arguably in an area that is furthest away in epistemological and methodological terms from music psychology—the issues of perceptibility and relevance for different groups of listeners apply to other music-theoretical approaches too. For example, Heinrich Schenker's multilayered analyses, which are based on the notion that, in the context of tonal music, some pitches and harmonies can structurally 'prolong' others, are not contentious in music-psychological terms when the prolongations are relatively near to the musical 'surface' and last only a few seconds.[20] However, Schenker's own application of this principle deep into the structural foundations of a piece, whereby it is asserted that single chords may in some sense control events lasting many minutes, has found no empirical support (see note 13). Unlike much pitch-class set analysis, however, it is conceivable, particularly for those with 'absolute pitch', that pieces *could* be heard as Schenker analyses them—as long-term prolongations of the tonic and dominant chords—if expert listeners chose to do so. The same applies to other music-analytical approaches too. Take, for example, Rudolph Réti's modus operandi, as set out in *The thematic process in music* (1951), through which he seeks to unearth (typically) 'hidden' motivic relationships that he claims exist within and between the movements of pieces, subconsciously unifying them in the ears of listeners. Again, despite the feeling that Réti may be confusing the 'background' repetition that, as we have seen, is inevitable in any comprehensible piece of music, with the 'foreground' that a composer may choose to overlay upon it, it is conceivable for expert listeners to hear music in line with Réti's analyses if they elect to do so.

Where do approaches such as this leave the relationship between music psychology and music theory? Let us return to Eric Clarke's proposed 'rapport' of 1989, which, he suggests, could be achieved by developing a kind of description that recognizes the mutual relationship between a perceiver and his or her environment.

> The aim of such an approach would be to describe musical events for a particular kind of perceiver [ranging, we may surmise, in Mailman's terminology, from a 'populist' listener to a 'progressive' theorist], taking account of the stimulus material, the perceptual systems that exist, and the cultural systems within which evaluations of musical function are made. This is in essence an argument for an ecological description, since it proposes that while there is an indefinite number of possible descriptions of the same state of affairs from a variety of different perspectives, and at a number of different levels, the kind of description that is of primary interest to us will be at a level, and of a breadth appropriate to human beings, their musical artifacts and activities, and the natural and cultural environment within which they are situated.
>
> Clarke (1989, p. 12)

Ockelford (2005b) proposes a way in which such a description could be modelled. He suggests that all potential music–structural relationships[21] can be considered to exist on a continuum with

[20] As the recent empirical work of Isabel Martinez shows (The cognitive reality of prolongational structures in tonal music, unpublished Ph.D. dissertation, University of Roehampton, 2007.)

[21] That is, in terms of Ockelford's 'zygonic' theory, between events the same, where one could conceivably be deemed to derive from the other.

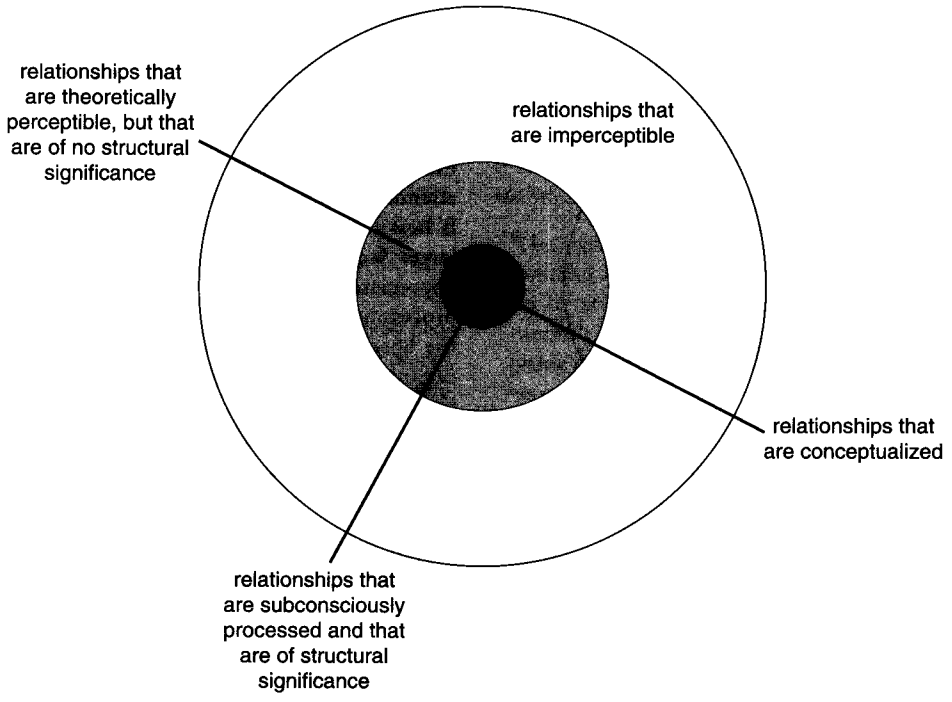



**Fig. 50.3** Representation through different shading of a set of potentially structural relationships as they are hypothesized to exist in relation to mental processing.

three distinct sectors: those that are (1) imperceptible or non-cognizable; (2) perceptible but of no direct significance to musical structure (arising, for example, by chance, as a result of 'background' repetition); and (3) subconsciously processed and of direct relevance to the cognition of structure. Inevitably, the boundaries between these sectors are fuzzy, since which potential relationships actually become reified in cognition, and the significance of these, is, as Clarke's ecological standpoint indicates, liable to vary from listener to listener, and even with the same listener on different occasions. However, where a given relationship is likely to reside on the continuum can in broad terms be predicted in relation to cognitive constraints and preferences that could be empirically tested (Ockelford 2002). A fourth condition—structural relationships that are consciously processed/conceptualized (by composers and theorists or analysts, for example)—is subject to even greater variation, having the potential to be overlaid anywhere on the other three regions.

This model can be represented graphically as shown in Figure 50.3. In order to interpret the diagram correctly, one has to imagine an exponential growth in the number of potential relationships as one moves outwards from the centre. Using the model, it is possible to capture visually the epistemological issues raised above and clarify some of the misunderstandings that are reported—a necessary step in seeking to resolve them.

First, we will use the model to depict the case of a listener relatively familiar with a piece and its broader stylistic context, though not approaching matters with a mindset to conceptualize what is being heard (an example of Mailman's 'populist' perceiver). Intuitively, the music makes sense to the person concerned, and so we can surmise that he or she must be processing structural relationships subconsciously—a situation represented in Scenario A (shown in Figure 50.4). Turning next to Margulis's neophyte, one could imagine that she or he would pick up on rather fewer structural relationships,

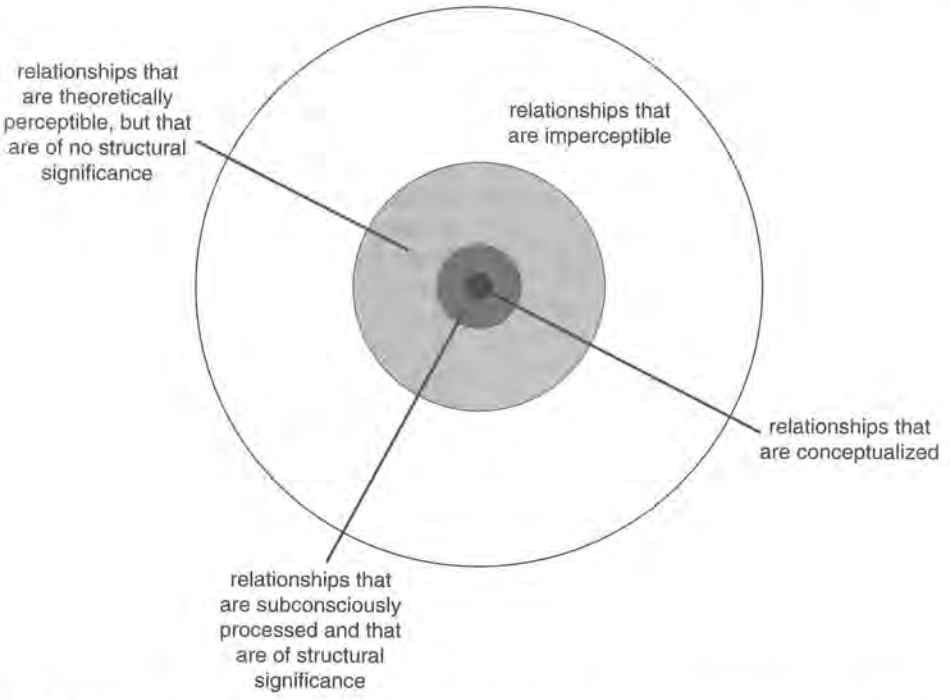



**Fig. 50.3** Representation through different shading of a set of potentially structural relationships as they are hypothesized to exist in relation to mental processing.

three distinct sectors: those that are (1) imperceptible or non-cognizable; (2) perceptible but of no direct significance to musical structure (arising, for example, by chance, as a result of 'background' repetition); and (3) subconsciously processed and of direct relevance to the cognition of structure. Inevitably, the boundaries between these sectors are fuzzy, since which potential relationships actually become reified in cognition, and the significance of these, is, as Clarke's ecological standpoint indicates, liable to vary from listener to listener, and even with the same listener on different occasions. However, where a given relationship is likely to reside on the continuum can in broad terms be predicted in relation to cognitive constraints and preferences that could be empirically tested (Ockelford 2002). A fourth condition—structural relationships that are consciously processed/conceptualized (by composers and theorists or analysts, for example)—is subject to even greater variation, having the potential to be overlaid anywhere on the other three regions.

This model can be represented graphically as shown in Figure 50.3. In order to interpret the diagram correctly, one has to imagine an exponential growth in the number of potential relationships as one moves outwards from the centre. Using the model, it is possible to capture visually the epistemological issues raised above and clarify some of the misunderstandings that are reported—a necessary step in seeking to resolve them.

First, we will use the model to depict the case of a listener relatively familiar with a piece and its broader stylistic context, though not approaching matters with a mindset to conceptualize what is being heard (an example of Mailman's 'populist' perceiver). Intuitively, the music makes sense to the person concerned, and so we can surmise that he or she must be processing structural relationships subconsciously—a situation represented in Scenario A (shown in Figure 50.4). Turning next to Margulis's neophyte, one could imagine that she or he would pick up on rather fewer structural relationships,

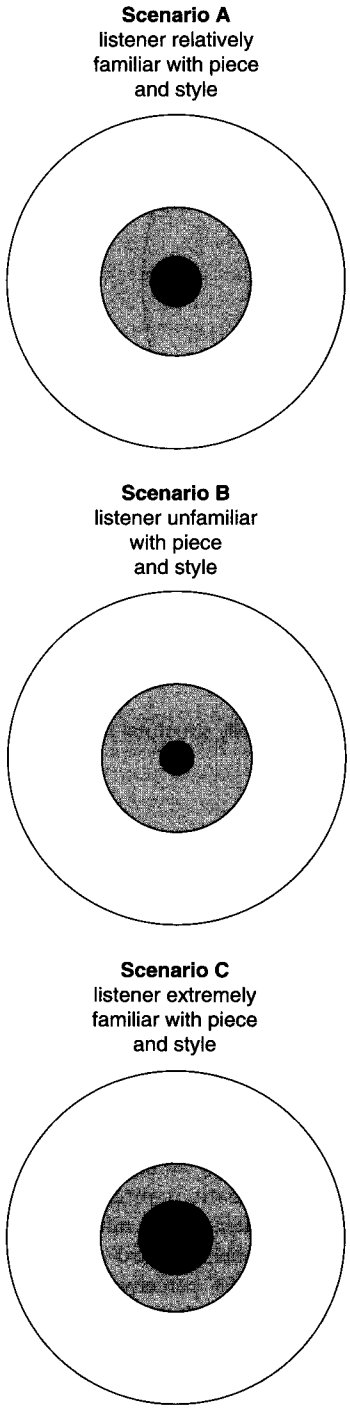



**Fig. 50.4** Varying state of relationships hypothesized to exist in different scenarios pertaining to listening.

though still presumably enough for the music to be recognizable as abstract patterns in sound that in this case convey a certain, rather limited, meaning. This is depicted in Scenario B. Notice that Sector 1 is diminished in size. Conversely, the situation with an expert listener, who is attending to piece without conceptualizing what is heard, may be illustrated with a somewhat larger Sector 1: Scenario C. It is important to appreciate that in Scenarios A, B and C, the processing of musical structure can occur within the absence of description or analysis of the organization that is present—indeed, without the listener having had any formal musical education—the recognition of perceived sonic patterns and the ascription of meanings to them being enabled purely through repeated exposure.

Naturally, when composers, performers or listeners, whether functioning as psychologists, theorists, educators, therapists or others, consciously *think about* music—or ask others to—the situation changes. Take, for example, Wolpert's experiment, mentioned above. Her results show that the musicians in her experiment were able to conceptualize structural relationships in the domains of pitch and perceived time that they would otherwise have been likely to hear subconsciously. Hence, the situation arose that is depicted in Scenario D (see Figure 50.5). However, it appears that Wolpert's so-called non-musicians, when required to effect the same comparisons, were drawn by the nature of the question they were asked: to conceptualize *different* perceptible relationships, this time in the domain of timbre, that are not generally considered by musicians to be 'structural' (rather, residing in the musical 'background'). Hence we have Scenario E. Clearly, both Scenarios D and E have a certain intrinsic validity (since they both represent sets of empirical findings), but the situations they model are quite different, and the danger occurs when the two are subjected to a common interpretational framework.

We move next to the scenarios engendered by the music analyses that have been discussed, beginning with Forte's account of Schoenberg's Op. 11, No. 1. Now it may be that some of the relationships between pitch-class sets of three notes, which we can assume would pass the great majority of listeners by unnoticed (and which

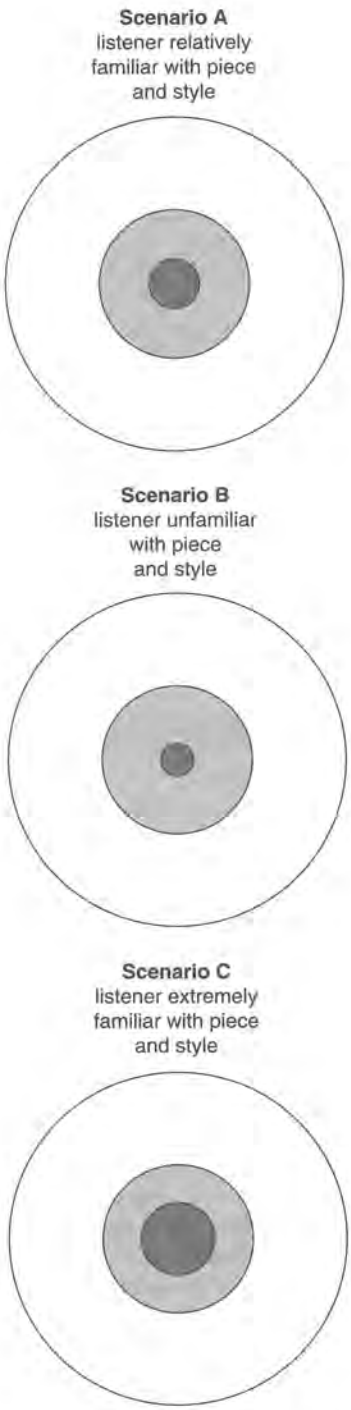



**Fig. 50.4** Varying state of relationships hypothesized to exist in different scenarios pertaining to listening.

though still presumably enough for the music to be recognizable as abstract patterns in sound that in this case convey a certain, rather limited, meaning. This is depicted in Scenario B. Notice that Sector 1 is diminished in size. Conversely, the situation with an expert listener, who is attending to piece without conceptualizing what is heard, may be illustrated with a somewhat larger Sector 1: Scenario C. It is important to appreciate that in Scenarios A, B and C, the processing of musical structure can occur within the absence of description or analysis of the organization that is present—indeed, without the listener having had any formal musical education—the recognition of perceived sonic patterns and the ascription of meanings to them being enabled purely through repeated exposure.

Naturally, when composers, performers or listeners, whether functioning as psychologists, theorists, educators, therapists or others, consciously *think about* music—or ask others to—the situation changes. Take, for example, Wolpert's experiment, mentioned above. Her results show that the musicians in her experiment were able to conceptualize structural relationships in the domains of pitch and perceived time that they would otherwise have been likely to hear subconsciously. Hence, the situation arose that is depicted in Scenario D (see Figure 50.5). However, it appears that Wolpert's so-called non-musicians, when required to effect the same comparisons, were drawn by the nature of the question they were asked: to conceptualize *different* perceptible relationships, this time in the domain of timbre, that are not generally considered by musicians to be 'structural' (rather, residing in the musical 'background'). Hence we have Scenario E. Clearly, both Scenarios D and E have a certain intrinsic validity (since they both represent sets of empirical findings), but the situations they model are quite different, and the danger occurs when the two are subjected to a common interpretational framework.

We move next to the scenarios engendered by the music analyses that have been discussed, beginning with Forte's account of Schoenberg's Op. 11, No. 1. Now it may be that some of the relationships between pitch-class sets of three notes, which we can assume would pass the great majority of listeners by unnoticed (and which

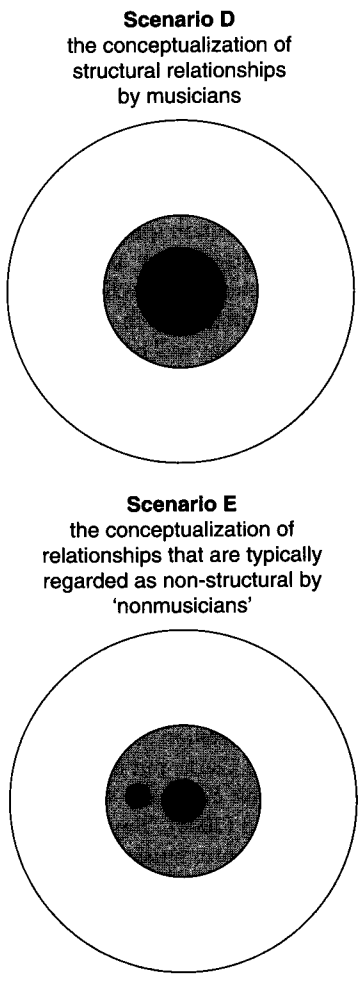



**Fig. 50.5** Interpretation and representation of Wolpert's findings (1990).

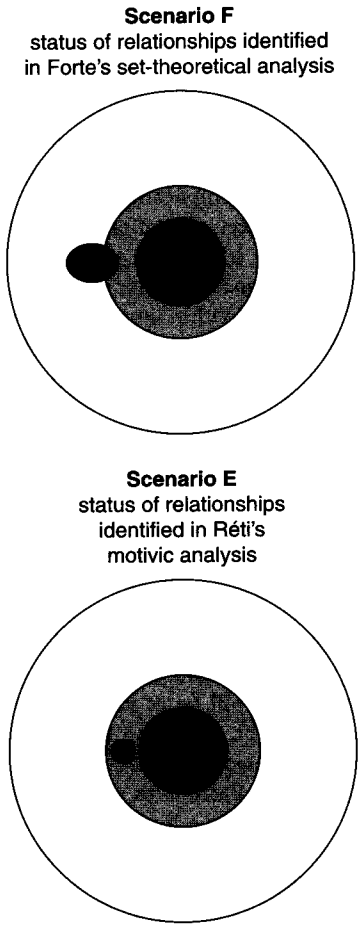



**Fig. 50.6** Interpretation and representation of Forte's and Réti's styles of analysis.

would therefore in most circumstances be deemed structurally non-essential), could be perceived with practice, although there are others that would appear to be inaudible, no matter how well-prepared the listener were to be: for example, when the analysis decrees that a pitch within a four-note chord should be heard as functioning in three different four-note sets simultaneously. Hence it seems that Forte's analysis spans Sectors 2 and 3—see Scenario F (in Figure 50.6). Réti's analyses, on the other hand, would on the whole be perceptible once they have been drawn to listeners attention: hence Scenario G. However, these arguably imbue the regular warp and the weft of the musical fabric with significance over and above the melodic and harmonic patterns that composers have woven through it, and they have failed to gain acceptance among musicologists or musicians.

A similar situation holds in relation to Mailman's critique of Ockelford's assessment of Lewin's late-Mozart analysis, which is discussed above. In terms of the three-sector model, Lewin's approach, using the transformation 'RICH', conceptualizes relationships that would typically be in Sector 2, whereas it is possible to assign the same musical structure to transpositional relationships, which exist in Sector 1. To return to Mailman's argument, however, if we

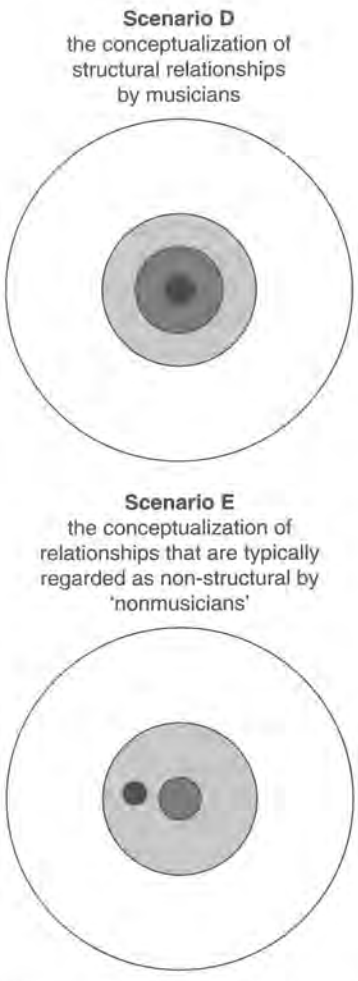



**Fig. 50.5** Interpretation and representation of Wolpert's findings (1990).

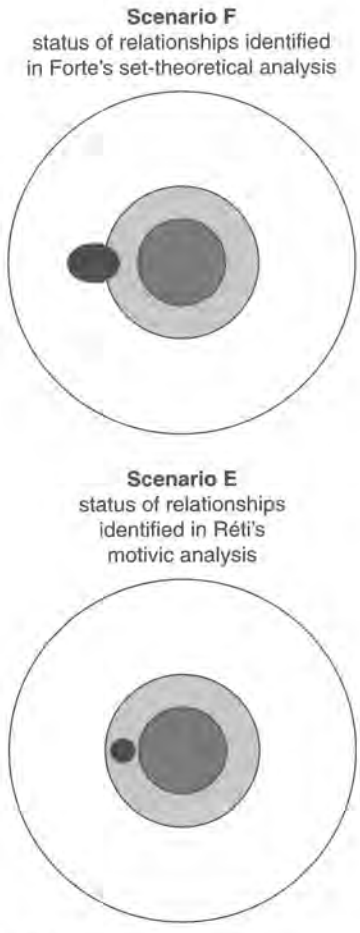



**Fig. 50.6** Interpretation and representation of Forte's and Réti's styles of analysis.

would therefore in most circumstances be deemed structurally non-essential), could be perceived with practice, although there are others that would appear to be inaudible, no matter how well-prepared the listener were to be: for example, when the analysis decrees that a pitch within a four-note chord should be heard as functioning in three different four-note sets simultaneously. Hence it seems that Forte's analysis spans Sectors 2 and 3—see Scenario F (in Figure 50.6). Réti's analyses, on the other hand, would on the whole be perceptible once they have been drawn to listeners attention: hence Scenario G. However, these arguably imbue the regular warp and the weft of the musical fabric with significance over and above the melodic and harmonic patterns that composers have woven through it, and they have failed to gain acceptance among musicologists or musicians.

A similar situation holds in relation to Mailman's critique of Ockelford's assessment of Lewin's late-Mozart analysis, which is discussed above. In terms of the three-sector model, Lewin's approach, using the transformation 'RICH', conceptualizes relationships that would typically be in Sector 2, whereas it is possible to assign the same musical structure to transpositional relationships, which exist in Sector 1. To return to Mailman's argument, however, if we

are ever to *expand* Sector 1—to be 'progressive', which he sees as a crucial part of music theorists' role—then we need to push out the boundaries beyond those which the ear would otherwise determine unassisted. That is, there is virtue in consciously seeking to claim structural territory from Sector 2. The implication is that conceptualization may affect perception and, in effect, a migratory effect is possible, whereby relationships that were once in Sector 2 metaphorically shift to Sector 1. Within limits, this seems reasonable enough. The author can recall such an experience, when, having become acquainted with Schoenberg's analysis of Brahms's fourth symphony, motivic relationships became apparent that were subsequently embedded in listening experiences (1947/75, pp. 405–408). However, this is a long way from saying that much set-theoretical analysis could ever have intrinsically *musical* relevance, no matter now practised the listener. As we have seen, because all music is supersaturated with repetition, any piece contains a virtually infinite number of potential patterns, which will be of varying degrees of interest, no doubt, according to one's point of view, but the majority of these should not be intellectually conflated with *musical* structure. Rather, they exist as a by-product of the way that music is put together in order for it to be comprehensible. It is surely at this point that the music psychology's relationship with music theory reaches a boundary (and arguably where music theory itself becomes something rather different).

## Summary

In summary, then, music psychology overlaps with a number of other disciplines, including music education, therapy, ethnomusicology, and music theory and analysis. There are tensions in each case, but, as the citations above have indicated, benefits too for those who are prepared to explore with an open mind. Ultimately, however, music psychology cannot be extended beyond the boundaries of its epistemological box, always granted that the sides are flexible and subject to change: indeed, such movement is likely to come about through the influence of adjacent disciplines. This has been shown, for example, in the groundbreaking work of Aaron Williamon and his colleagues at the Royal College of Music in London—the first UK conservatoire to establish a music–psychological research centre, which has aimed to support students in improving their performance skills and managing the high levels of stress that are often induced by performing in public. In terms of the common space occupied by elements of music psychology and music theory, further developments may well mean having to accept that the initial research questions and the *evaluation* of data may well be guided by musical intuitions, but that the gathering and *analysis* of data should be rigorous and undertaken with a 'scientific' detachment. Two decades on from Clarke's appeal to 'Mind the gap', there is now a greater clarity as to what this conceptual intersection might look like, how it might function, and even what it should be called: 'empirical musicology'. According to Honing (2006), empirical musicology, which now has its own journal (*Empirical Musicology Review*) 'grew out of a desire to ground theories on empirical observation and to construct theories on the basis of the analysis and interpretation of such observations'. And as Cook and Clarke put it in the introduction to their book *Empirical musicology: aims, methods, prospects* (2004), 'Empirical musicology. can be thought of as musicology that embodies a principled awareness of both the potential to engage with large bodies of relevant data, and the appropriate methods for achieving this' (p. 5). Hence one senses that in this sphere of activity the influence of music psychology may be keenly felt in the next decade.

## References

Abert H (1990). *W.A. Mozart: Zweiter Teil, 1783–1791*. Breitkopf and Härtel Musikverlag, Leipzig.

Boulez P (1963/71). *Boulez on music today*, trans. S Bradshaw and RR Bennett. Faber and Faber, London.

Bruner C (1984). The perception of contemporary pitch structures. *Music Perception*, **2**(1), 25–39.

Clarke (1989). Mind the gap: formal structures and psychological processes in music. *Contemporary Music Review*, 3, 1–13.

Cook N (1987). The perception of large-scale tonal closure. *Music Perception*, 5(2), 197–206.

Cook N (1994). Perception: a perspective from music theory. In R Aiello, J Sloboda, eds, *Musical perceptions*, 64–95. Oxford University Press, Oxford.

Cook N and Clarke E (eds) (2004). *Empirical musicology: aims, methods, prospects*. Oxford University Press, Oxford.

Dearling R (1982). *The music of Wolfgang Amadeus Mozart*. Associated University Presses, East Brunswick, NJ.

Deliège I and Sloboda J (1996). *Musical beginnings: origins and development of musical competence*. Oxford University Press, Oxford.

Deutsch D and Boulanger R (1984). Octave equivalence and the immediate recall of pitch sequences. *Music Perception*, **2**(1), 41–51.

Dubiel J (1999). Composer, theorist, composer/theorist. In N Cook and M Everist, eds, *Rethinking music*, 262–283. Oxford University Press, Oxford.

Dunsby J (1998). Fortenotes. *Music Analysis*, **17**, 171–181.

Erickson R. (1975). *Sound structure in music*. University of California Press, Berkeley, CA.

Forte A (1973). *The structure of atonal music*. Yale University Press, New Haven, CT.

Forte A (1981). The magical kaleidoscope: Schoenberg's first atonal masterwork, Op. 11, No. 1. *Journal of the Arnold Schoenberg Institute*, **5**, 127–168.

Ginsborg J (2007). Strategies for memorizing music. In A Williamon, ed., *Musical excellence: strategies and techniques to enhance performance*, 123–141. Oxford University Press, Oxford.

Gjerdingen R (1999). An experimental music theory? In N Cook and M Everist, eds, *Rethinking music*, 161–170. Oxford University Press, Oxford.

Hallam S (1995). Professional musicians' orientations to practice: implications for teaching. *British Journal of Music Education*, **12**(1), 3–19.

Hallam S (2001). The development of expertise in young musicians: strategy use, knowledge acquisition and individual diversity. *Music Education Research*, **3**(1), 7–23.

Honing H (2006). On the growing role of observation, formalization and experimental method in musicology. *Empirical Musicology Review*, **1**(1), 2–6.

Jørgensen H (2004). Strategies for individual practice. In A Williamon, ed., *Musical excellence: strategies and techniques to enhance performance*, 85–103. Oxford University Press, Oxford.

Keller H (1966). Wolfgang Amadeus Mozart. In R Simpson, ed., *The symphony: volume one*, 50–103. Penguin Books Ltd, Harmondsworth, Middlesex.

Kenny B and Gellrich M (2002). Improvisation. In R Parncutt, G McPherson, eds, *The science and psychology of music performance: creative strategies for teaching and learning*, 117–134. Oxford University Press, Oxford.

Korsyn K (2003). *Decentering music: a critique of contemporary musical research*. Oxford University Press, New York.

Lehmann A and McArthur V (2002). Sightreading. In R Parncutt, G McPherson, eds, *The science and psychology of music performance: creative strategies for teaching and learning*, 135–150. Oxford University Press, Oxford.

Lerdahl F (1992). Cognitive constraints on compositional systems. *Contemporary Music Review*, **6**(2), 97–121.

Lewin D (1987). *Generalized intervals and transformations*. Yale University Press, New Haven, CT.

Mailman J (2007). Review article. *Psychology of Music*, **35**(2), 363–375.

Margulis E (2005). Review of the book *Decentering music*. *Psychology of Music*, **33**(3), 331–337.

McPherson G (2006). ed. *The child as musician: a handbook of musical development*. Oxford University Press, Oxford.

Morris R (1979–1980). A similarity index for pitch-class set. *Perspectives of New Music*, **18**, 445–460.

Ockelford A (2002). The magical number two, plus or minus one: some limitations on our capacity for processing musical information, *Musicae Scientiae*, **6**, 177–215.

Ockelford A (2005a). *Repetition in music: theoretical and metatheoretical perspectives*. Ashgate, London.

Ockelford A (2005b). Relating musical structure and content to aesthetic response: a model and analysis of Beethoven's Piano Sonata, Op. 110. *Journal of the Royal Musical Association*, **130**(1), 74–118.

Ockelford A (2008). *Music for children and young people with complex needs*. Oxford University Press, Oxford.

Ockelford A, Pring L, Welch G and Treffert D (2006). *Focus on music: exploring the musical interests and abilities of blind and partially-sighted children with septo-optic dysplasia*. Institute of Education, London.

Parncutt R and McPherson G (eds) (2002). *The science and psychology of music performance: creative strategies for teaching and learning*. Oxford University Press, Oxford.

Pring L and Ockelford A (2005). Children with septo-optic dysplasia—musical interests, abilities and provision: the results of a parental survey, *British Journal of Visual Impairment*, **23**(2), 58–66.

Réti R (1951). *The thematic process in music*. Greenwood Press, Westport, CT.

Roeder J (1988). A declarative model of atonal analysis. *Music Perception*, **6**(1), 21–34.

Sloboda J (1985). *The musical mind: the cognitive psychology of music*. Clarendon Press, Oxford.

Saint-Fox G de (1947). *The symphonies of Mozart*, trans. L Orrey. Dennis Dobson Ltd, London.

Schoenberg A (1947/75). Brahms the progressive. In L Stein, Ed., *Style and idea: selected writings of Arnold Schoenberg*, 398–448. University of California Press, Berkeley, CA.

Thompson S and Lehmann A (2004). Strategies for sight-reading and improvising music. In A Williamon, ed., *Musical excellence: strategies and techniques to enhance performance*, 143–159. Oxford University Press, Oxford.

Tolbert E (2001). Music and meaning: an evolutionary` story. *Psychology of Music*, **29**(1), 84–94.

Williamon A (ed.) (2004). *Musical excellence: strategies and techniques to enhance performance*. Oxford University Press, Oxford.

Wolpert R (1990). Recognition of melody, harmonic accompaniment, and instrumentation: musicians vs. non-musicians. *Music Perception*, **8**, 95–105.

Wosch T and Wigram T (2007). *Microanalysis in music therapy: methods, techniques and applications for clinicians, researchers, educators and students*. Jessica Kingsley, London.

# CHAPTER 51

# History and research

Michael Thaut

THIS chapter will provide an introductory synopsis of the history of research developments and research themes in music psychology from its inception in the nineteenth century until the early twenty-first century. The synopsis has to be broad to fit into the limitations of a book chapter, but a broader approach also offers the advantage to lay out more clearly discernible lines of evolution of large research themes over time. That is of particular interest in music psychology as it is a research-based field with an intrinsically inter- or multidisciplinary character. A bibliography of books rather than a reference list of specific papers should provide the appropriate data support for this synopsis.

Psychology of music as a new and independent scientific discipline started its development in the middle of the nineteenth century. It was a component part of the arrival of psychology as a new discipline studying human behaviour through observation, measurement, and testing using scientific methods (experiments, hypotheses, statistics). Previously, the scholarly discourse about music as an artistic and human phenomenon was mostly dominated by philosophy and—perhaps surprisingly to the modern reader—the discourses in the early natural sciences. The first work that is usually associated with the original impetus for studies in the psychology of music was Helmholtz' book *On the sensation of tones as a physiological basis for the theory of music* (Die Lehre von den Tonempfindungen als physiologische Grundlage fuer die Theorie der Musik, 1863), a primarily physiological study on tone perception. Helmholtz was a physician, physiologist, and physicist whose landmark work on sensory perception (visual and auditory) refuted at that time the still dominating schools of Naturphilosophie (natural philosophy), a speculative way of thinking and interpreting the physical world. Helmholtz was an empiricist and a very influential scientific thinker of his times. Interestingly, Wilhelm Wundt—who is considered with William James to be one of the founders of modern experimental psychology—was for a while a student in Helmholtz' laboratories in Berlin.

It is important to remember some of the beginnings of the psychology of music, in order to understand the fundamental and extraordinarily radical change and determination of the future directions in studying music that was brought on by this new discipline. Only a few decades earlier, the German philosopher Schopenhauer had called music the purest incarnation of the human will, the embodiment and expression of human feelings (sadness, joy, love) in their abstract interpretation as metaphysical ideas. Even while a new experimental science of music as human behaviour emerged, contemporary philosophers such as Nietzsche still interpreted music philosophically, for example, as transcending reality, redemptive to the human condition, and either destructive (Dionysiac) or idealizing (Apollonic) reality and dissolving individuality to become one with the governing forces of the universe.

The new psychology of music diverged quickly from such centuries-old traditions of philosophical interpretation of music as human behaviour. The physicist Carl Stumpf was in 1883 the first to utilize modern scientific methodologies (experimental testing, hypotheses, statistical analysis) applied to investigations of musical–psychological phenomena—in his particular case to perceptions of consonance and dissonance—in his publication *Tonpsychologie* (Psychology of Tones) and other works.

The early and foundational works in psychology of music were mostly driven by experimental investigations into sound perception, linking psychology and psychophysics—or specifically psychoacoustics in regard to music—to understanding human responses to music. Interestingly, psychoacoustics remains to this day an important part of research agendas and teaching curricula in psychology of music. Carl Seashore, who in many ways is considered the primary pioneer in American psychology of music, developed the first objective test of musical aptitude in 1919, conceptually it was strongly connected to the psychophysical origins of the psychology of music. He designed a test, consisting of acoustical stimulus pairs, for which the test subject had to decide if both stimuli were same or different. The test items measured, e.g., differences in the perception of pitch, timbre, duration, or rhythmic pattern structure. Test items were produced by tone generators to avoid bias and influence for the test subjects on test performance through differences in musical exposure and familiarity. Seashore's ideas were rigorously scientific and revolutionary for his times: however, in many follow-up studies his test showed fairly low predictive power for future musical success and achievement in people. Seashore's ideas were so rigorously perception-based in psychophysics that his test is nowadays still considered an objective measure of acoustical acuity, but not a good measure of the more complex nature of musical abilities. Interestingly, the rhythm section of the Seashore test has survived and continues to show its validity as part of a neuropsychological assessment test for brain injuries (Halstead–Reitan), testing in all probability abstract auditory temporal perception as a measure of cognitive-perceptual functioning in an injured brain.

In the 1920s and 30s, a new focus entered the original research focus of psychology of music. This shift is often referred to as a shift from a psychology of tone to a more analytical and penetrating psychology of 'music', emphasizing the comprehensive perception and evaluation of complex pattern structures of melody, harmony, form, and rhythm as constituent building blocks of the language of music. Ernst Kurth's work *Musikpsychologie* (1931) is often cited as one of the first works to proceed with this new shift. This new shift was well supported experimentally and theoretically by the emergence of a new direction in the psychology of perception, the new school of Gestalt perception which holds—in basic tenet—that the sum of individual parts in a complex pattern structure assume a new perceptual quality as a 'whole' that is different from simply viewing a stimulus pattern as collection of individual parts. One of the simplest experiments to illustrate this concept was based on the so-called phi phenomenon or phenomenon of apparent movement. The rapid alteration of blinking lights being turned on and off periodically and in sequence creates the optical illusion of light moving back and forth along a line from point to point. Musical forms follow a complex architecture of sound patterns, assembled in rule-based systems, and acoustically and perceptually organized in five dimensions simultaneously and sequentially: rhythm/time—pitch/melody/polyphony—space—timbre—loudness. It is an interesting observation, that music is the only sensory language code that proceeds sequentially and simultaneously at the same time within its rhythm–polyphony dimensions and thus presents unique challenges to complexity thinking in the perception/cognition of information processing. Contrarily, verbal language, for example, only proceeds in sequentiality. A large amount of research has followed and continues to follow this direction of investigating experimentally complex music perception, following gestalt principles of perception. The vitality of this field shows the continued benefits of this approach, also in light of modern brain research within the models of cognitive neuroscience. We will return to this subject again later in this chapter.

Roughly in tandem with the emergence of Gestalt psychology and its extensions on the study of music perception, a third strand of research became subject of new investigations: the affective response to music. Verbal responses to quantify and categorize affective responses to music appeared in the early twentieth century as a new study focus, differentiating affective responses into studies of emotional responses, mood states, and perceived feeling states.. One of the first and widely used such verbal tests was Hevner's Adjective Circle (1936), a checklist of

adjectives characterizing and categorizing subjective feeling responses to music. Researchers tried to find useful categories in music that would allow them to study potential links how musical variables would influence affective responses. Among the first categorizations of music al variables were consonance vs dissonance and minor vs major modes. Soon other variables were examined, such as tempo, pitch, and variations in melody, harmony, and rhythm, frequently subcategorized under descriptors like sedative or stimulative music. Other variables investigated included repetition, familiarity, and various investigations into the function of musical taste. Many leading music psychologists of the first half of the twentieth century, such s Farnsworth, Hevner, Schoen, and Seashore, engaged in studies of the affective response in music. An impressive array of data demonstrated overwhelming evidence that music can evoke deep and complex affective responses, but direct linear cause–effect linkages between music variables and specific affective responses were difficult to pinpoint. Predictive power in such studies remained very low. Data showed on the contrary that affective responses to music were often highly idiosyncratic to individual listeners and their musical enculturation as well as to the specific conditions in which the listening and testing took place, e.g., mood states or arousal needs the subjects brought into the listening condition. Studies of the affective response in music were also complicated by the difficulties in the general psychological studies of affect, mood, and emotion to conceptualize affective states. Two theoretical approaches have dominated and continue to shape such discussion: a categorical approach to defining emotional states, using a limited number of emotion states, such as happiness, anger, disgust, surprise, sadness, etc. A second school of thinking is often called a dimensional approach in which emotion states are not categorically defined but emotions occur only within two or more dimensions, such as levels of perceived pleasure and levels of perceived intensity. A dimensional approach has obviously had more attraction to music research as music is void of connotative semantic meaning, and therefore a dimensional response to affect in music would bypass the dilemma of exact verbal descriptors of felt emotion in music and definitions of what musical variables are associated with exactly what categorical states in affect and emotion. However, dimensional approaches have the disadvantage for emotion research in music that the determination of affective states remains somewhat diffuse, and causal or predictive linkage with musical variables is difficult to conceptualize. Not until the second half of the twentieth century did several new landmark theories in aesthetics—influential for re-approaching affective response research in music—move this area of research in psychology of music decidedly forward.

In conjunction with studying verbal responses in affective responses to music, a very active line of physiological investigations emerged in this particular area, holding with the original impetus of the new psychology of music to study perception objectively as close to the human nervous system as possible. The research into the affective response to music as a physiological response spanned almost the full circle of the twentieth century, until it was superseded by breakthrough developments in brain research technology that allowed researchers to study brain responses *in vivo* under complex affective and cognitive stimulus conditions that began to include music at a very early stage of using these new technologies.

Prior to this new development, starting in the first years of the twentieth century, a huge amount literature begin to be amassed for the next 70 to 80 years, studying mostly physiological responses of the autonomic nervous system such as cardiac (heart rate), respiratory function, vascular changes (blood pressure), or galvanic skin response (GSR: changes in electrical skin conductance). This research showed many fascinating insights into the reactivity and sensitivity of physiological responses to music. The research has shown unequivocally that music elicits physiological changes that are indicative of changes in the perception and interpretation of the music by the listener. However, after those basic insights the data picture quickly became more complicated. Changes in physiological responses to musical stimuli did not seem to follow clear linear relationships with characteristics and changes in musical variables. For example, slower tempi in music do not always show slower heart rates, higher loudness levels or brighter timbres do not consistently create

increases in vascular constriction or other measures. To create better interpretations of the physiology, studies quickly went to design paradigms combining verbal responses with physiological measures. Also in such a two-pronged assessment approach, simple cause–effect relationships remained elusive. For example, studies have shown that increases in self-reported relaxation measures may coexist with increases in physiological measures, e.g., accelerated heart rates.

Data seemed to suggest that the physiological response per se appears contentless, i.e., similar changes in heart rate or GSR can be associated with very different interpretations as to behavioural meaning. Secondly, the data suggested that interpretations of musical pieces or changes in musical elements (e.g., tempo) are highly idiosyncratic and contingent upon a listener's complex personal interpretation system. Therefore, in later developments many studies began to move away from music that was pre-selected or pre-categorized by the researcher to musical stimuli that—often self-selected by the experimental participants—could be very different in structure, style, and genre but had consistent meanings across listeners. In addition to recognizing complex idiosyncratic response models that made generalized investigations and interpretations of responses in physiology and verbal reports to pre-categorized music styles almost meaningless, further insight came from arousal theories that pointed out that changes in physiological reactions are highly dependent in their meaning and interpretations on existing arousal states preceding the stimulus-elicited changes. For example, temporary boosts to high arousal may be perceived as pleasant inductors and precursors of subsequent relaxation (the gym effect). In states of low arousal, moderate boosts of arousal (e.g., increases in heart rates) may be perceived as pleasant and relaxing.

Considerable breakthroughs in aesthetic theory began to influence psychology of music research in the second half of the twentieth century. Two highly influential theories—that still continue to influence most new developments in theories of meaning and music—were proposed respectively by Leonard Meyer in the 1950s and Daniel Berlyne in the 1960s and 70s.

Meyer proposed a model mechanism how music elicits meaningful affective responses in the music recipient—be it performer or listener. Based on Dewey's conflict theory of emotion which states that emotional responses are triggered when a tendency to respond is inhibited, expectations are built in music through compositional schemes that are continuously delayed, inhibited, and resolved—creating a complexly woven sound architecture based on building musical structures of anticipation–tension–resolution. In the interplay between the build-up of anticipation, the interjection of tension, and the creation of resolutions, the suspense evokes arousal—especially during the tension phase, i.e., the temporary inhibition of expectations—accompanied by a search for resolutions that is subsequently resolved within the musical structure. The resolution may then serve dialectically as the next springboard to renew the sequence of events since the resolution sets up a new set of expectations. The stronger the build-up of suspense or tension, the stronger will be the arousal build-up, and consequently the stronger will be the emotional release upon resolution. In order to have an appropriately structured experience, the music must have a well-balanced mixture of elements of novelty or uncertainty and redundancy or certainty in order to build meaningful tension-release schemes. Meyer's model, which was initially a purely theoretical proposal, has received considerable empirical support and is at the basis of many newer aesthetic models using tension–release mechanisms to explain musical meaning and affective experience in music elicited by the perception of music structures. Proponents of this view are usually called absolute expressionists because they believe that although musical meaning is communicated intrinsically through the perception of musical elements, i.e., the syntax of music, the perception of the melodic, rhythmic, and harmonic patterns and structures of the music can evoke meaningful affective experiences, however, not in the categorical semantic sense of emotion as previously discussed. Meyer's model constituted significant progress in explanatory accounts of understanding the nature and causation of meaning in music, including affective responses. His model and later models, connected to him, have received a slow but steady appreciation in the cognitive neuroscience of music, influencing study designs and hypothesis development for assessing brain responses to music.

Berlyne was the main proponent of another new development in aesthetic theory which is still highly relevant for modern brain research in music and aesthetic theories in the psychology of music. Berlyne founded the movement of 'experimental aesthetics' which was in many ways ahead of its time. He laid out a framework for how the perception of elements and patterns in artworks contain arousal-inducing attributes that constitute physiological arousal as the basis for affective–aesthetic behaviour in art perception. Berlyne—whose views on how we respond and how we perceive artworks was clearly influenced by gestalt theories of perception—showed at least in principle how artworks can have a profound effect on brain and behaviour by linking perceptual and arousal processes that are intrinsic to the human brain. At the time, his physiological elaborations on how aesthetically ordered visual and sound patterns in the visual arts and music have unique arousal-inducing properties that—when perceived as 'Good Gestalts'—create meaningful and affectively rich perceptual experiences, did not have the benefit of new research models and deeper knowledge of human brain function in cognitive neuroscience. However, his basic framework has been not so much rejected but simply vastly expanded by modern human brain imaging and neurophysiology. Berlyne was perhaps the first scholar in music psychology close to suggesting a basis for a new biological aesthetics that includes music as a biological core function of the human brain. We will come back to this point later.

In the last 25 years of the twentieth century, music psychology underwent its most fundamental change with the advent of brain imaging as an entirely new technology of studying the human *in vivo*, involved in complex cognitive operations. Up that point, human brain research could learn much from animal models in regard to motor system function or the neuroanatomy of various perceptual systems, but higher cognitive functions could only be diffusely approximated through animal brain research. Other techniques had major limitations in creating large samples, establishing testable hypotheses, or studying brain processes during complex cognitive tasks. For example, posthumous brain research from persons who had known and well described behavioural dysfunctions, research with persons who had undergone well-defined neurosurgical interventions (e.g., split brain surgery), or dichotic listening tests linking ear preference with hemispheric preference in perceptual processing, all had helped contribute knowledge on cognitive brain function and were widely used in music research, but could never accomplish what the development of comprehensive knowledge on brain function measuring a human brain *in vivo* could deliver.

The two most prevalent brain imaging techniques, positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), allowed unprecedented access to studying the brain basis of human cognition and helped establish a new field of scientific inquiry, appropriately referred to as cognitive neuroscience. Cognitive neuroscience in many ways has begun to change psychology as a mostly behaviourally based research discipline. Interestingly, music became rather quickly—even in the pioneer stages of these new technologies—a subject of immediate interest, and music studies are among some of the first human brain imaging studies carried out. The cognitive neuroscience of music is now an integral part music psychology, its research efforts and the knowledge base of the biology underlying music behaviour. Brain imaging has profoundly altered the traditional nature of research in music psychology.

Initially, research paradigms using brain imaging in music were relatively basic and looked for simple contrasts. For example, typical early studies compared listening to music vs listening to a narrated text, or compared musical responses of professional musicians vs non-musicians. Based on such baseline designs, some of the early findings focused on differential hemispheric activations in music compared to non-musical tasks, where the right hemisphere seemed particularly involved in music processing compared to the left brain hemisphere. Such findings created a certain amount of publicity and notoriety in considering that the right hemisphere may be the site of artistic abilities in the brain. As research designs became more sophisticated and more complex designs investigated specific musical responses, such notions turned out to be highly oversimplified. Music as a complex sound language that engages multiple processing systems

in the brain in the area of perceptual, motor, and cognitive functions, actually turned out to be a stimulus that is best characterized by very complex activations patterns in highly distributed patterns across the whole brain on cortical and subcortical levels. The brain imaging research of the past 20 years has evolved in roughly four categories: music perception—broader cognitive processes, e.g., regarding emotional saliency—music learning and plasticity—biomedical applications of music.

In music perception, a wide array of studies have investigated the functional neuroanatamy of music perception (and production) relative to specific musical elements or specific production tasks. Perception of smaller elements such as pitch, harmony, rhythm, timbre, and perception of larger pattern elements such as melodies or scales, vs full musical excerpts, and other structural components of music have been studied in respect to underlying brain function. Additionally, aspects of music production such as motor control in instrumental and vocal (singing vs speech) performance have been extensively studied. Investigations into rhythm perception and rhythm production—especially issues in rhythmic sensorimotor synchronization—have received particular attention in brain imaging research as an important component aspect of music perception.

Broader aspects of music responses—for example more interpretative responses in regard to meaning and emotional saliency—have been studied less to date than more structural music perception questions. However, recent attempts are beginning to make up for this noticeable lag. Imaging studies in regard to emotion and music may still face age-old challenges in regard to the previously discussed difficulties in appropriate conceptualizations and theoretical underpinnings of meaning and emotional behaviour in music. Also, brain research into the meaning and emotion in music must be based on solid theoretical foundations in aesthetic theory and may require more intense and integrative collaborative efforts between musicologists, music psychologists, and cognitive neuroscientists in the future. However, a very important and fertile field for future research seems wide open in this area.

New findings in the neurobiology of learning have motivated considerable efforts in musical brain imaging. The discovery that experience, learning, and training can drive brain plasticity—i.e., changes in the neural organization and connectivity of the brain—led to a number of studies investigating brain plasticity induced by music learning and music training. The evidence shows impressively that music changes the brain and that certain areas, e.g., motor areas involved in instrumental training (keyboard, violin, etc.) or visuospatial and cognitive areas trained through intense music training can become enlarged and preferentially activated via the establishment of new and well-developed neural circuits. There is a broader impact implied by many of these studies because changes in brain plasticity through music learning—especially if they involve brain areas that also mediate critical non-musical cognitive and perceptual functions—may suggest that music learning can confer a wider cognitive benefit to general non-musical cognition and intellectual functions. Such notions have considerable importance for all forms of music training as well as the role of music education in public education.

Finally, since the early 1990s a combination of brain imaging and clinical behavioural studies has investigated the biomedical effect of music in a neurorehabilitation environment, where music is used as a therapeutic stimulus to retrain motor, language, and cognitive function in persons with brain injuries and neurological disorders, such as stroke, Parkinson's disease, traumatic brain injury, dementia, cerebral palsy, and other dysfunctions of the human nervous system. Studies have shown impressively that music is an effective sensory language to facilitate the recovery of neurologic function when used in a therapeutic context that emphasizes learning and training models to re-educate the brain via mechanisms of brain plasticity. In that respect, the new neuroscience research in music has led to a fairly radical reformulation of concepts and principles for the role of music in therapy and medicine, away from traditional social science-based concepts of music therapy—traditionally emphasizing notions of well-being and cultural values in relationship building and human interaction—to neuroscience-based models, emphasizing music as a language that can actively address the rehabilitation of perception, cognition, and motor function.

Brain imaging has brought about enormous changes in research directions and virtually reinvented the cognitive science of music embedded in music psychology form its early origins. However, as a technique it has several serious limitations that need to be kept in mind when assessing its importance. Interestingly these limitations are particularly relevant in music research. The current techniques in neuroimaging provide a detailed look at brain activation based on topography, i.e., maps of highly developed spatial resolution. Brain imaging creates maps of brain activity, measured over a relatively long period of recording time in the second range (e.g., 20–30 secs). In other words, the time resolution—i.e., when something happens rather than where—has limits in assessing with exactitude the timing of brain activations. This is important in music research because music is physically based on vibration patterns of air molecules, and as such the perception of sound is based on time parameters of vibrating objects. Music is a temporally based sensory language and brain activations in music perception must respond to information processing in the range of milliseconds (or below). To that end, the auditory system which processes music is a highly sensitive system in regard to speed and precision in responding to very small fluctuations in temporal changes in the vibrations patterns that constitute sound. Thus, neuroimaging can make important contributions to the mapping of music perception and musical responses but can only add limited information to the temporal aspects of music processing in the brain.

For this reason, two other techniques measuring brain activations—one new and one older—have received considerable attention in music research, providing an excellent complement to brain imaging via PET or fMRI. The techniques are called electroencephalography (EEG) and magnetoencephalography (MEG). Their respective mechanisms of signal recording are somewhat different although they record the same brain signal. EEG records brain waves—usually via sensors attached to the scalp—by detecting the electrical currents that are emitted by the electrochemical potentials of neurons during their firing activity. MEG records the same electrochemical signal but measures it by the measuring the magnetic field created by the electrochemical signal. EEG is a technique going back at least 50 years that was developed initially to detect any abnormalities in the patterns of neuronal firing activity which occur in an oscillatory 'rhythmic' fashion: hence the somewhat misleading term brain 'waves' was created. Data recordings and analysis of brain rhythms have become exceedingly more detailed and complex, so that EEG research has experienced a very fruitful renaissance. EEG measures have very high temporal resolution down to 1 millisecond—however, they have poor spatial precision because surface sensors receive and record signals from widely distributed neuron ensembles and exact spatial determination of the source of the signals is not possible. MEG has a major advantage over EEG because a technique called dipole analysis in the magnetic field allow the creation of fairly good spatial locators for the brain signal source. Both EEG and MEG can record brain activity associated with a certain behavioural task with high accuracy of timing of the signal processing in the brain. In music, EEG and MEG can be used to study the time course of music perception across different brain areas in millisecond ranges which give an excellent picture of the brain's temporal processing of music. In conjunction with PET or fMRI imaging, a very complete spatial and temporal map of music processing in the brain can be established. However, combining techniques requires complex equipment interfaces that are not easy to carry out technologically.

One of the most innovative current approaches to studying music perception is to compare music perception on analogue variables to verbal language. The comparison makes sense in many ways as both music and verbal language are auditory-based communication systems, hold vital function in human society, and have several systems of structural organization in parallel, such as phonological sound category systems and syntactical rule systems. There are also some obvious differences, of course. For instance, both systems have pitch and timbre as sound categories, but in music pitch is the primary categorical system to organize sounds into musical patterns, whereas in verbal language timbre is a primary organizer of speech sounds (vowels, consonants). One of the

most critical differences lies in the absence of semantics in music, whereas linguistic semantics is a critical component of verbal language. The semantic meaning of a word is the reference to objects or experiences in reality and to the mental representations of such reality. Semantic meaning is referential, making propositions about reality. Those are fairly philosophical concepts that show how deeply linguistics is embedded in concepts of mind and thought. Musical statements or communication symbols—melodic motifs, rhythmic phrases, harmonic sequences—are non-semantic in the sense that they only refer to themselves within the rules of the musical syntax used. That brings up another important point for consideration. Both music and verbal language can be said to be based on rules of how structures of sounds or words are constructed, via linguistic grammar or compositional grammar (in music). Both grammars are generative in that they have rules that can generate and predict meaningful 'correct' statements in many combinatorial ways in a given language. They also have, however, decisive differences. Violations of linguistic grammar can be determined to be incorrect and rendering statements meaningless. Violations of musical 'grammar' rules are soft and not absolute. Violations of musical grammar are common and often deliberate in order to discover new musically meaningful 'grammars'. Therefore, direct comparisons between musical syntax violations and linguistic syntax violations are not entirely possible and reveal little about shared or parallel brain processing for different language systems. However, carefully constructed comparisons between analogue language processes in music and linguistics have been recently attempted in brain imaging, and can yield insight not only into music-specific cognitive brain processes but also into aspects of cognitive processing systems in the brain that may operate cross-modally for different 'language' systems.

A new model may push investigations in music research one step further within the language model. Instead of investigating music in comparison to verbal language, a new model may investigate music as an autonomous biological language of the brain in relationship to other verbal and non-verbal language systems. In this model, the brain is considered a multiple language generator. The brain generates and operates in multiple languages of thought. Thoughts as mental representations occur in verbal language form but also other forms of languages in non-verbal systems come easily to mind, e.g., the language of quantities and size expressed through numbers. The arts and music would be a specific case within the languages of images and percepts. Music as a language is based on the ability to think in sound, to perceive relationships between sounds and to generate rules to build sound events into complex sound scapes. The whole architecture of music as such would then be considered based on the ability of complex non-verbal temporal auditory thinking, which is by definition a deeply cognitive process. Human beings are clearly born with that ability, just as they are born with a biologically innate knowledge of verbal language. It is of great interest in this matter to study pre-historic artworks which show figurines and paintings of highly modern design, shape, and with highly evolved aesthetic qualities of abstraction and symbolism. We find artifacts like that dated 70 000 or more years old—tens of thousands of years before the first emergence of written documents for word languages and basic numerical concepts. Artefacts of musical instruments can currently be dated as far back as 35–45 000 years. Recent brain investigations regarding music processing show interestingly many areas involved that are also involved in complex linguistic (verbal language) processing and other cognitive functions. These findings may provide support for a 'music as a language' model within a comprehensive model of multiple brain languages.

Modern psychology of music has engaged in other areas of investigation and analysis besides the ones discussed so far. Social communication in regard to cultural values, uses, and function of music have always been an important aspect of research in music psychology. Music psychology and music anthropology have interacted in studies of music perception as well as historical roles of music in human culture. Music education and music therapy often consider psychology of music as a foundational field of study for their disciplines, and music psychology has indeed contributed important knowledge to such those fields. However, it is probably fair to

say that the majority of research effort in music psychology has not been in applied or translational research for educational, developmental, or therapeutic purposes. However, such applied or translational focus may become one of many new future directions in the psychology of music.

## Bibliography


Avanzini G, Faienza C, Minciacchi D, Lopez L and Majno M (eds) (2003). *The neurosciences and music.* New York Academy of Sciences, New York.

Avanzini G, Lopez L, Koelsch S and Majno M (eds) (2005). *The neurosciences and music II.* New York Academy of Sciences, New York.

Berlyne DE (1971). *Aesthetics and psychobiology.* Appleton Century Crofts, New York.

Bregman A (1990). *Auditory scene analysis: the perceptual organization of sound.* MIT Press, Cambridge, MA.

Clynes M (ed.) (1982). *Music, mind, and brain: the neuropsychology of music.* Plenum Press, London, New York.

Cook PR (ed.) (1999). *Music, cognition, and computerized sound: an introduction to psychoacoustics.* MIT Press, Cambridge, MA.

Critchley M and Henson RA (1977). *Music and the brain: studies in the neurology of music.* CC Thomas, Springfield IL.

Deutsch D (1999). *The psychology of music.* Academic Press, New York.

Dowling WJ and Harwood DL (1986). *Music cognition.* Academic Press, New York.

Fauconnier G and Turner M (2002). *The way we think.* Perseus Basic Books, New York.

Galaburda AM, Kosslyn SM and Christen Y (eds) (2002). *The languages of the brain.* Harvard University Press, Cambridge, MA, London.

Hodges DA (ed) (1999). *Handbook of music psychology.* IMR Press, San Antonio, TX.

Jackendoff R (2002). *Foundations of language.* Oxford University Press, Oxford.

Juslin PN and Sloboda JA (eds) (2001). *Music and emotion: theory and research.* Oxford University Press, Oxford.

Lundin RW (1967). *An objective psychology of music.* The Ronald Press Company, New York.

Merriam AP (1964). *The anthropology of music.* Northwestern University Press, Evanston IL.

Meyer LB (1956). *Emotion and meaning in music.* University of Chicago Press, Chicago, IL.

Miell D, MacDonald R and Hargreaves DJ (2005). *Musical communication.* Oxford University Press, Oxford.

Pierce JR (1992). *The science of musical sound.* WH Freeman, New York.

Patel AD (2008). *Music, language, and the brain.* Oxford University Press, Oxford, New York.

Radocy RE and Boyle JD (1979). *Psychological foundations of musical behavior.* Charles C Thomas, Springfield IL.

Riess Jones M and Holleran S (eds) (1990). *Cognitive bases of musical communication.* American Psychological Association, Washington, DC.

Sefarine ML (1988). *Music as cognition.* Columbia University Press, New York.

Sethares WA (2007). *Rhythm and transforms.* Springer, London.

Shaw-Miller S (2002). *Visible deeds of music: art and music from Wagner to Cage.* Yale University Press, New Haven, CT.

Snyder B (2000). *Music and memory.* MIT Press, Cambridge, MA.

Thaut MH (2005). *Rhythm, music, and the brain.* Routledge, New York, London.

Wallin NL, Merker B and Brown S (2000). *The origins of music.* MIT Press, Cambridge, MA.

Zatorre RJ and Peretz I (eds) (2001). *The biological foundations of music.* New York Academy of Sciences, New York.

# CHAPTER 52

# Where now?

Susan Hallam, Ian Cross and Michael Thaut

THE purpose of this chapter is to draw together emerging themes from across all sections of the book. The themes identified are not exhaustive but capture broad trends, ideas, issues and questions which have resonance across the whole of music psychology. The themes are not mutually exclusive. There is overlap between them. Each theme has implications for many areas of music psychology. The first theme concerns the power of music, the second the way that modern technology has enhanced music's influence through increasing access and enabling individuals to listen to the music of their choice at any time or place. Subsequent themes relate to our understanding of basic perceptual and cognitive processes, and music as language, communication and interaction. A key issue is the need to take greater account of culture and context, with more research being undertaken in non-Western cultures. The chapter concludes with a consideration of methodological advances and the need for those from different disciplines with an interest in music psychology to work more closely together.

## The power of music

Music has a multiplicity of functions which operate at several levels: that of the individual, the social group and society in general. In relation to each of these music can have a very powerful impact influencing behaviour, emotions, and moods in a variety of settings frequently without our conscious awareness (see Chapters 11–15 and 40–45). The exceptional power of music may be related to it being able to stimulate us aurally, visually, intellectually, emotionally and physically, simultaneously.

Background music selected by others and music accompanying visual images and verbal messages can have a very immediate impact through its power to invoke associations and change moods and arousal levels. This has led to its use in a variety of leisure, commercial and employment settings (see Chapters 40, 41 and 45). The ever-growing body of research in this area has made considerable progress in understanding the nature of the relationships between musical stimuli and behaviour but there is still much work to do to establish the particular conditions, contexts and musical characteristics which lead to specific outcomes and to identify those characteristics of the individual which may ameliorate the impact. We also need to establish the factors that are important in determining negative responses to music.

Concentrated engagement with music, usually through listening, can evoke peak emotional experiences which can have permanent and powerful therapeutic effects enhancing relationships, attitudes, general appreciation of art or music, aesthetic sensitivity and perception of everyday reality (see Chapter 42). To understand these powerful responses better we need to be able to identify the specific conditions, individual and contextual, which pertain at the time. There is also a growing body of evidence of the longer-term effects which active engagement with music in educational, community, therapeutic and medical contexts can have on health, well-being and intellectual development (see Chapters 23, 44, 46–49). Traditionally, active engagement with music and music therapy, informed by cultural values and social science models, has been assigned a rather global and diffuse role in bringing about long-term change

in behaviour, relying on concepts of well-being, social relationship building and the development of transferable skills. Recent developments in the neurobiology of music have enabled more concrete evidence about what music can achieve. By studying the brain basis of musical behaviour and learning, it has become clear how music learning reciprocally effects changes in the brain (Chapters 16 and 18). The concept of neuroplasticity applied to musical learning in relation to the initial education of the brain and re-educating the injured brain (Chapters 17, 18, 23, 46–49) has created the conditions for new scientific research on the effects of music in educational, medical, therapeutic, and rehabilitation settings that can be transferred and generalized to a wide range of non-musical functions, for instance, motor and speech training and cognitive functions such as memory, attention, and executive functions. In music therapy and medicine much promising research has been accumulated and has begun to be translated into practice (Chapters 46–49). In educational and community contexts progress is slower (see Chapters 23 and 44), but overall these developments will continue to present a powerful research agenda in the future, creating a new form of neurologically and evidence-based music research in education, medicine and related disciplines, and therapy for the twenty-first century.

Music also has a powerful impact on social interactions, the identities individuals create for themselves, the social groups with which they engage, and the meanings that they construct and experience in and through music (see Chapters 1–4, 27 and 43). A limitation of research to date is that most has been undertaken within Western musical cultures. Its validity needs to be tested within other musical contexts and cultures to explore whether similar phenomena occur elsewhere. Other relatively neglected issues relate to the extent to which identification with particular musical genres determines behaviour beyond the musical. For instance, with regard to engendering violent, antisocial or rebellious behaviour, are particular musical genres preferred because they reflect already existing behavioural tendencies, or does the music itself encourage particular behaviours? We also need to explore the relationships between the extent, quality and antecedents of the individual's involvement with music and their identification with particular genres, taking account of the experiences of individuals for whom music is relatively unimportant. Understanding why some individuals do not engage with music may shed light on a range of issues relating to musical ability, musical preferences and motivation.

## The impact of technology

Technological advances in the twentieth century changed the ways in which music could be experienced. They not only made music more accessible but made it possible for each individual to select the music that they wished to listen to in any combination of genres in any environment. People now use music to meet a range of psychological needs: as a distraction, to focus attention or energize, to manipulate moods, and to work through emotional issues (see Chapter 40). Music plays a key role in people's lives as they travel, and engage in domestic and social activities in the home and community, and at leisure. Considerable progress has been made in identifying the ways in which music is used at the individual level, but we know little about how it is selected for particular purposes, or how successful this process is. The way that music impacts on and is used in our daily lives will clearly continue to be a major focus of research in the future.

The easy availability of music and its commodification challenges earlier idealized conceptions of music listening as a concentrated process with a focus on the spiritual or aesthetic. Currently, we lack knowledge about different approaches to listening, and the way that music used as background moves from the periphery to the centre of attention. This is an area for future development.

Technology has not only led to the increased availability of music but also its globalization and commercialization. These changes will inevitably have impacted on those involved in the creation and performance of music. As yet we know relatively little about this. Research which explored the nature of the work of professional performers and composers in the twenty-first century and the impact of this on learning, motivation and performance would have major implications for the future training of musicians (see Chapters 30–38).

## Understanding of basic perceptual and cognitive processes

Although there now exists a considerable body of empirical research on music perception and cognition, many 'primitive' perceptual processes are not yet known in sufficient detail to allow us to feel that we fully understand the processes of music perception and cognition (see Chapters 5–10). Many fundamental aspects of musical sound are not yet well enough understood in terms of their physics in order for us to be confident that we can develop predictive theories about how we experience such sounds. For example, the ways in which our perception of a violin's sound may be shaped by the pattern of the radiation of sound from the instrument is constrained by the fact that, as yet, there exists no complete physical account of how the instrument's structure, and mode and physical context of performance, might shape that pattern of radiation. The nature of the factors that influence how we hear sounds as having particular timbral qualities is not wholly clear (see Chapter 7), nor do we know how judgements about sound quality relate to issues of attention and memory (see Chapter 10). These are all questions for future research.

There is also the question of the contribution to music perception and cognition of sensory modalities other than the auditory (see Chapter 3). Several researchers have theorized and explored ways in which music perception and cognition appear to involve cross-domain mappings between the auditory attributes of music and features of other domains of human experience (visual, haptic, motional, etc., as well as cognitively higher-order constructs such as body-image schemas). At present, it is not clear how integral these cross-domain mappings might be for our experience of music and there is immense scope for future exploration.

## Music and language/music as language

Recently, there has been a regeneration of interest in the relationship between music and language as communicative systems, and as manifesting common properties as representation systems within the brain. The parallel and shared components between these two auditorily based systems of communication and mental representation offer new opportunities for research and early findings are promising, with notions of generativity, syntax and grammar suggested as accounting for the processes involved in the experience of musical structures (see Chapters 6 and 19). The extent to which syntactical aspects of music can be regarded as being of the same order—in terms of scope and complexity—as syntax in language is still unclear, as is the nature of the factors that enable complex serial structures to be perceived and remembered in both language and music. If music can be conceived of and perceived as though it conforms to grammatical principles, what types of grammar are entailed, and do these constitute 'true' grammars of the type that form foci of linguistic research, or are they more constrained in their scope and power? If the grammars of music and language were to be found to be closely comparable, the question then arises of how music can be differentiated from language as a domain of human thought and behaviour (see Chapters 1–4). There is also the issue of how different musical practices have arisen in different cultures.

This research agenda may be strengthened by studying music—within a broader concept of all the arts—as an autonomous language itself, i.e., a representation of thought in non-verbal aesthetic images and percepts. Such conceptualization may create new and productive insights into music as an innate biological language of the brain, its origins, its role and function in the cognitive and intellectual development of the human brain, and its unique contributions to the complexity of human brain function—not as an enriching cultural phenomenon, but as a core modality of human cognition.

## Music as communication and interaction

A key emerging theme is that of music as communication. Advances in technology have enabled more detailed consideration of the assessment and modelling of performance processes providing new insights into performance practices, communication issues relating to playing from

memory, the role of movement, collaboration between performers and issues relating to performance anxiety and the medical problems that can arise from a lengthy career in the music profession (see Chapters 30–36). This research has focused almost exclusively on Western classical music, and we know little about these issues as they relate to other genres and musical cultures.

There is also a dearth of research on the way that audiences participate in the communicative process. The relationship between music in the individual mind and music as an interactive behaviour is not well understood, we know little about how our private musical cognitions relate to our public musical interactions. Musical performance may appear as the exercise of individual musical skills, but equally as the outcome of joint musical endeavour. To date we have few insights regarding the way in which the musical behaviours of others, our perceptions of those behaviours, and our conceptions of the mental processes that we infer as underlying them, contribute to the experience of music as performance or as interaction. While some empirical research has begun to explore these issues and to develop theories of music cognition in interaction (see Chapters 1, 4, 34 and 38), the vast majority of research has focused on individual music cognitions and perceptions. Exploration of capacities that are likely to be central to interactive musicality, such as that for entrainment (see Chapters 4 and 8) is in its infancy, and questions of how such capacities relate to broader human capacities for complex social communicative interaction have barely been addressed. Indeed, the question of whether or not at least some types of musical judgement may be exercised as social and interactive behaviours rather than being evidenced in acts of individual judgement (as much music sociology would suggest) remain to be explored, as does the problem of how such situations might be investigated scientifically.

## The interface between the biological and cultural

An ongoing issue is the relationship between what is biologically determined at the species or individual level and what is determined by the cultural environment in which the individual develops. Are aspects of the human capacity for music found in the behavioural and mental capacities of any non-human animal species (see Chapter 1)? If so, are those human capacities interpretable as evolutionarily homologous to those of non-human species (which would indicate common evolutionary descent), or are they best construed as analogous (likely to have arisen because of similar selection pressures)? Answers to these questions would greatly advance our understanding of the significance of music as a human faculty, and require to be pursued by music psychologists in collaboration with ethnologists, geneticists, and evolutionary biologists and psychologists. Investigations of the 'social mind' and the 'social brain' are now major foci for cognitive and neuroscientific investigation, in part precipitated by the realization that the complexity of human sociality is one of the principal defining characteristics that appears to differentiate us from other species. Behavioural research has focused on issues such as theory of mind, social reasoning, and decision-making, and has often adverted to evolutionary approaches in seeking explanations for the scope of, and the constraints on, human social behaviours. Neuroscientific research has also begun to turn its attention to these and other aspects of human social capacities, and exploration of the place of music in human social capacities must be a primary focus for future research.

There are also unresolved issues relating to the propensity for individual differences in musical ability (see Chapter 24). Advances in research on genetics, to date, have had little impact on the debate. Few human behaviours or traits have been traced to specific gene pairs and it is likely that those who exhibit musical skills are drawing on a range of different gene combinations, which exert an influence on physical make-up in addition to cognitive and emotional development. Interactive rather than additive models of the relationship between the environment and genetic inheritance are now generally accepted. These interactions begin in the womb and continue throughout the lifespan (see Chapters 20–22).

Developing empirical methodologies are making it increasingly possible to carry out

methodologically sound investigations relating to the prenatal period. As our understanding of prenatal influences increases we need to be able to relate these to postnatal behaviour. Cross-cultural longitudinal research following individuals from birth would enable us to identify those factors which contribute to the development of different musical skills in particular contexts. This would enable us to explore how children acquire musicality in non-Western, particularly traditional cultures and establish the factors which are common across cultures in child rearing and other social practices which encourage the acquisition of musical skills (see Chapters 20–22 and 28 and 29). While the family plays an important role in the early stages of engagement with music in Western cultures, this may not be the case elsewhere. Cultural influences on families, for instance the extent to which music is valued in the subculture to which the family belongs and the nature of the value placed on music in society, may be important, for instance, whether it is viewed as related to religion, as entertainment, as a form of protest, or as a way of life. Cross-cultural research would also enable us to establish if what we know about practice and performance applies equally to all cultures and musical genres (see Chapters 25 and 30–36). Currently there is little research on learning, practice, memorization processes, and performance outside the Western classical tradition. Realistically, longitudinal cross-cultural research is unlikely to be funded (though highly desirable) but it would be possible to undertake cross-cultural, cross-sectional research which would offer some insights. This would elaborate our knowledge of the impact of different cultural and familial traditions on perception, learning, cognition, and responses to music. The research would need to acknowledge that musical potential manifests itself in a range of different ways, creating further challenges for researchers in developing appropriate methodologies for exploring the processes of learning, assessing change, and attempting to identify the contribution of particular environmental influences. A number of systemic interactive models have been developed which would assist this process (see Chapters 26–29).

While existing research has made a major contribution to understanding the way in which specialist musical skills are acquired in a Western musical context, there still remain a number of key questions. For instance, how important is the child's choice of instrument in determining motivation and the level of expertise attained? To what extent do opportunities to engage in a wide range of musical activities contribute to the development of expertise as opposed to deliberate practice on a single instrument? To what extent do other factors, apart from those relating to musical expertise, contribute to success in the music profession (see Chapters 24–29)? There is also a developing research agenda in relation to amateur engagement with music. With the population in the Western world having more time for leisure and living longer, more people are developing musical skills as a hobby. We know relatively little about the factors that contribute to individuals returning to or taking up musical activities in later life, whether their motivation is musical, social or a combination of these, and how it is sustained.

The processes of composition and improvisation are situated within particular musical contexts which have their own conventions and rules. Existing research has largely taken place within a Western context, with a focus on professional musicians or the activities of young people working within the school curriculum (see Chapters 37–39). The relative lack of research in non-Western cultures means that we do not know whether the approaches and strategies identified in expert and novice composers and improvisers are specific to Western musical contexts or whether they have relevance elsewhere. We also know relatively little about the long-term development of composers and the way that their thinking, skills and careers change over time and may differ between genres. For instance, has the increasing commercialization of music had an impact on their working practices and output? The other ongoing challenge for research in this area relates to assessment and developing ways of judging the relative value or creativity of particular compositions or improvisations.

The majority of music cognition research has been laboratory-based, using Western-encultured listeners and Western–generally common-practice period–musical materials. While these studies, with the tight controls and

replicable methodologies required by a scientific approach, have yielded many insights into music in mind, very few studies have managed to investigate non-Western musical perceptions with the same degree of rigour. In some ways this is not surprising; the exigencies of conducting empirical research in field contexts are not particularly consonant with the need to be able to control experimental variables and exclude those that cannot be controlled. Nevertheless, it is vital that future research should aim to investigate the ways in which non-Western musical perceptions and cognitions are aligned with, or differ from, those of the Western participants from whom our current knowledge is principally derived, ideally by using means that will enable the exploration of a wide range of emic perceptions and conceptions.

## Methodological limitations and advances

All research is limited by the methodologies available to it, and music psychology is no exception. A wide range of approaches and methodologies have been adopted across the field in attempts to develop our understanding of musical practices and processes many of which are unavailable to conscious awareness. There has been much reliance on self-report (interviews, diaries, rating scales), either concurrent or retrospective, and experiments or observations the outcomes of which are used to infer musical processes. The advent of *in vivo* brain imaging technology in the field of cognitive neuroscience has emerged as a new force complementing existing methods, enabling the integration of behavioural research designs in psychology with concomitant studies into the functional neuroanatomy underlying these cognitive behaviour functions (see Chapter 51). An important and highly fruitful trend is likely to be the fusing of many areas of research. We see the beginnings of this already, for instance, when more and more sophisticated and musicologically sound behavioural studies in the perception of musical structure are combined with investigations into the underlying neurobiology of music. Advances in technology and the use of innovative experimental methods might also enable exploration and comparison of the brain states of participants in musical interaction, which could greatly illuminate the dynamics of music as social process.

## Collaboration in research

Emerging throughout the book are issues relating to the need for collaboration between those working in different disciplines. The historic synopsis of research developments in music psychology illustrates the high degree of diversity as well as the unique dynamics of forming interdisciplinary research agendas that have continuously driven the field forward in a number of different directions (see Chapters 50 and 51). Music psychology has never presented itself as a uniform field of investigation and research goals. Various influences from different fields have converged on a scientific study of music and shaped the research, including aesthetics, cognitive psychology, physics, physiology, social sciences, anthropology, and most recently, cognitive neuroscience. Music psychology also overlaps with a number of other disciplines, including music education, therapy, ethnomusicology, and music theory and analysis. To further our understanding in many areas of work greater collaboration is required. For instance, in developing understanding of musical learning and development, music psychologists need to join forces with researchers in music education and with ethnomusicologists, as the types of information, theory and research practice that the questions demand require that developmental theories of music be elaborated and tested in a wide range of ethnographic contexts.

A continual multidisciplinary integration of and collaboration with musicology, i.e., music theorists, composers, and historians, into scientific research agendas where musicologists, musicians, psychologists, and other sciences work shoulder to shoulder in equal partnerships would also greatly benefit music psychology. For instance, there is much research by both musicologists and psychologists on expressive performance, particularly rubato. However, these two strands of work rarely make reference to each other. Yet both seek to explain and describe musical behaviours and cross-fertilization

of methodologies and knowledge would aid a deeper understanding.

In relation to developing our understanding of the perceptual and cognitive processes in music, it is essential for music psychologists to collaborate closely with engineers, physicists, psychoacousticians, linguists, and others in order to develop the field. This is necessary not only to ensure that theory and experiment in music psychology is informed and guided by well-founded understandings of physical and psychoacoustical facts and processes, but also to make clear to those with whom we are collaborating the ways in which the complexities of the questions raised by music perception and cognition may be informative in the development of their own disciplines.

There would also be benefits in closer collaboration with researchers working in relation to other art forms. The study of creativity in performance, improvisation and composition would benefit from collaboration with those working in the visual arts, drama and dance. This would enable communalities and differences to be identified and shared issues, for instance in relation to assessment, to be better understood.

## Endnote

Music psychology does not function as a tightly defined academic discipline operating within a single, focused epistemological framework. It is interdisciplinary. This has led to contributions from a wide range of related fields, although from time to time there are tensions between them. While some may argue that the interdicsplinary nature of music psychology is a weakness leading to a lack of focus, there have been clear benefits to the quantity and quality of research generated when compared with the other visual and performing arts. From this perspective interdisciplinarity is a cause for celebration.

# Index

Please note that page references to Figures or Tables are in *italic* print


## A

A minor key 62, 63
AAM (anxioalgolytic music) 529–531
AAMT (American Association for Music Therapy) 515
Aarden B 110
Aasgaard T 521
ABA (Applied Behaviour Analysis) 494
aboutness 25
absolute pitch (AP) 111
absolutism, versus referentialism 141–142
accent salience, and marking time 88–90
acoustic instrumentation, measuring performance 324
acoustic processing 48
action, coordination of 41–42
action-listening 126
activation, memory 109
actualization, of musical potential 261
Adorno TW 29, 147, 405, 446
adrenaline 395
advertising, musical 'fit' in 487
aesthetic perspective 151–157
  cognitive revolution 153–154
  evolutionary aesthetics 156–157
  experimental aesthetics 161–162
  experimental antecedents 152–153
  history and research 555
  meaning and music 27–28, 30
  neo-hedonism 154–156
  philosophical antecedents 151
*Aesthetics and psychobiology* (Berlyne) 152
Afanador K 448
affect 131
*Affektenlehre* (doctrine of the affections) 142
Agawu K 307
age factors 275–276, 302–303
AI (artificial intelligence) 328
*alap*, Indian 20
Alcock KJ 193
algorithmic and statistical modelling 327
*Alpine Symphony* (Strauss) 146
Alzheimer's disease, patients with 124
Amabile TM 467
American Association for Music Therapy (AAMT) 515
American National Standards Institute (ANSI) 47
AMT (Analytical Music Therapy) 495, 496
amusia
  congenital 191, 210
  defined 12
  infant musical skills 223
  MBEA 184, 188, 190, 191
  and melodic contours 211
Analytical Music Therapy (AMT) 495, 496
Anderson JR 250
'Angels' (Williams) 35, 39, 42
angioalgolytic music (AAM) 529–531
ankle jerk, effects of music 125
ANSI (American National Standards Institute) 47
anthropology, and evolutionary theory 4
*The Anthropology of Music* (Merriam) 36
*The Anthropology of the Body* (Blacking) 364
anxiety, performance *see* performance anxiety
AP (absolute pitch) 111
Applied Behaviour Analysis (ABA) 494
Arabic scales 55
Armstrong L 409
Arom S 15
Aronofsky D 444
art, and aesthetic perspective 27–28
articulation, emotional expression 144
artificial intelligence (AI) 328
*asak* metre 20
Ash RL 146
Ashley R 383, 416
Asmus EP 287
assessments, dynamic attending theory
  behavioural 86–87
  biological 87
associative chaining 352, 353
attainment value 467
Attali J 36
Attneave F 53
audience/audience effect 373–374, 441
auditory memory 354–355
auditory perception illusions 98–99
auditory scene analysis 53–54
auditory streaming 53–54, 76
auditory system, physiology 48
Austin JR 289
Australian Aboriginal song, hierarchical structure 14, 17–18
  and implicit learning 19–20
australopiths, and evolutionary theory 10
autistic spectrum disorders, music therapy research 519
autobiographical memory 108, 458
Averill G 40
Ayotte J 189

## B

B major 60
Bach JS 54, 78
  Concerto for Two Violins in d minor 125
  *Goldberg Variations* 153
  Italian Concerto 359–360
  preludes 95

background music, impact
on behaviour and task performance 475–476
future research 476–477
on learning and behaviour 474–475
Baily J 416
Balinese music, tonal hierarchies 17
ballet 448
Bamberger J 240
Banchieri A 410
Bandura A 288, 467
Bangert M 200
Banton LJ 347, 348
Baroque music 54
Barrett M 424, 425
Bartlett JC 102
Barwick L 18, 19–20
basal ganglia 335
basic cues 360
basilar membrane (BM) 48–49
Batt-Rawden K 164, 435
Baumgarten A 151
Baumrind D 300, 301
beat 16, 17, 110
*The Beatles* 488
Beethoven L van 409–410
behaviour, music as 5, 6
*see also* musical behaviour
behaviourism, music therapy 494–495
*see also* CBT (cognitive behavioural therapy)
Behne KE 135
Belcher S 434
Ben El Heni J 17
Bereitschaftspotential (BP) 335
Berg MH 270
Berkley R 425–426
Berlyne DE 30, 161, 162, 484, 485
*Aesthetics and psychobiology* 152
Bernstein N 365
Bey C 99
Bharucha JJ 20, 65, *66*, 68, 110
Bigand E 68, 99, 114
biochemical responses, effects of music on 122, *123*, 124
biology, and evolutionary theory 4–5
biosocial interaction model 280–281, 393
Birchnell J 300–301
Birdwhistell RL 40
Blacking J 5–6, 7, 29, 39, 126–127
*Anthropology of the Body* 364
blending, conceptual 406–407
Blood A 124
blood pressure, effects of music on 122
blood volume, effects of music on 124
blood-oxygen studies, effects of music 125
BM (basilar membrane) 48–49
Boden M 405
bodily responses to music
*see also* emotional responses to music
definition of terms 121
physical responses 125–126
physiological *see* effects of music, physiological
studies 125
body, and improvisation 414, 416–417
body movements, musically expressive effects 366–368
body work, and listening to music 434
Boerner S 446
Bohlman P 6, 40, 41
Bolet J 352
Boltz M 85, 444
bonding 224
bone pipes, and evolutionary theory 10
Bonny H 137
Borthwick SJ 464
Bosendorfer pianos 323
Botez MI 192
Bourdieu P 435
Bower GH 506
BP (Bereitschaftspotential) 335
Bradt J 519
brain
*see also* brain differences, musicians and non-musicians; musicians; *specific parts*, such as Broca's area
changes where learning to play instruments 201–203
cortical regions 334
melodic contours, processing 209–211
mirror neuron network 126, 203–204
multimodal sensorimotor integration regions 198–200
and music 208–215
planning in, observing 335–337
pleasure centres in 155
primary domain regions 197–198
sensory encoding of sound 208–209
syntactic processing 211–214
brain differences, musicians and non-musicians 197–201
brain anatomy 339–340
extended network of multimodal sensorimotor integration regions 198–200
functional imaging correlates, and structural differences 198
structural 198–200
brain work, and listening to music 433
brainstem reflex 136
Braxton A 409
breathing, effects of music on 124
Bregman, Albert 53
Bregman AS 98
Brendel A 356
Broca's area 179, 198, 199, 204
Brokaw JP 296
Bronfenbrenner U 279–280, 285
Bronstein P 299, 300
Brooks D 516
Brophy TS 418
Brown H 95
Brown M 404, 406
Brown S 9
Brunswik 383, 386
Bruscia KE 517, 520, 522
Bryant GA 9
Burnard P 419, 424, 425, 426
Butler C 95, 153–154
Byrne C 425
Byrne R 406
Byron T 5, 11

# C

C major key 64, 65, 70
intra-key hierarchies and inter-key distances 60, *61*, 62–63
tonic choices 95
C minor key 60, *61*, 63
Cage J 404, 447
CAM framework 448
Cannon W 395
Carpenter RH 346, 348
Carterette EC 16
Casals P 391
Castellano MA 17
categorical perception 55
Catterall JS 247
CBT (cognitive behavioural therapy) 397
central nervous system disorders, musicians 391
Certification Board for Music Therapists (CBMT) 520
Cesarec Z 507
Chaffin R 329, 355
chain complex 365
Chance M 372
characteristic frequency 48
children
*see also* education; infants, learning from; musical development; prenatal development
assessments, new approaches 425
cognitive learning 508
collaboration and group composition 425
communication skills, and music therapy 507–508
creative musical thinking 421–427

education context 421–422
emotional and social development 508
empowering children's voices 424–425
and improvisation 418–419
models of creative thinking 426
motor function, and music therapy 507
music technology, role 426
musical therapy with 507–508
new research 423–426
notation 425
pedagogy of composition 425–426
psychological literature 422–423
research paradigms 424
school years 316–318
chills, effects of music 124, 156
choking, and expert memory 358
Chopin F *63*, 65, 95
choral singers 122, 124, 473
chords
*see also* keys, tonal hierarchies; tonal hierarchies; tones
and musical expectations 171
priming paradigm 68–69
target 69
triads 62
choreography 447
chromatic scale 54, 59
chunking, memory 108
*Church Lirrga* songs, Wadeye 19, 20
cingulate motor area (CMA) 334, 335
Clarke EF 21, 326, 414
classical conditioning 222
Clayton M 17, 29
clinical behavioural studies 557
clock model, encoding theory 82–83
Clynes M 146, 327
CMA (cingulate motor area) 334, 335
CMT (Creative Music Therapy) 497–498
CNS (central nervous system) disorders, musicians 391
cochlea 48
Cochrane Database of Systematic Reviews 518
cognition disorders, musical
acquired, of musical production 192–193
and emotion 190
emotion 191–192
listening deficits 184–185, *186*, 190–191
and memory 190
neural substrate 192
pitch 185, 187–188, 191
temporal structures 189–190, 191
timbre 189
tonal structures 188–189
cognitive, behavioural and cognitive behaviour therapies 396–397
cognitive anxiety 394
cognitive psychology, and tonal hierarchies 65
cognitive rehabilitation, music therapy 510–511
cognitive sciences, of music 32
cognitive transfer, music instruction 249–250
cognitivist/cognitive neuroscience approaches 153
Cohen AJ 442, 443, 444, 445, 446
Cole M 21
Coltheart M 211
Coltrane J 409
common fate principle 54
communication
*see also* emotional expression
code description 383
concept 378
emotions 380, *381–382*, 383–384
expressive patterns and ornamentation 383
functions of music 39–40
intersubjective 224
to listeners 379–380
measurement of performance 326–327
modelling communicative process 385–387
music as communication 563–564
performance 371–373
synthesis of emotional expression 383–384
community and educational setting,
impact of music in 471–473
formal education 473–474
notion of 'community' 418
composition
compositionality 25
creativity theories 405–407
group, and collaboration 425
imagining 407–408
improvisation 408–409
metaphors and models 407
motivation 404
pedagogy of 425–426
and psychology 404–405
sketches, of composers 409–410
technology 410
traces of conscious thought processes 409–410
compound melody/melodies 53–54
compression 407
concept, music as 5
conceptual integration networks 407
conceptual spaces 406
concert-goers, experiences 436–437
conflict theory of emotion (Dewey) 555
confluence studies 423
Congado ritual (Afro-Brazilian) 40
congenital amusia 191, 210
Congruence-Association Model 443, 444, 446
Conlon DE 371
consciousness
and composition 409–410
fetal 221–222
levels of 222
consonance 51, 52–53, 231
constraint 413
constructivism 421–422
consumer behaviour 481–489
advertising, musical 'fit' in 487
arousal-based goals 484–485
music on-hold 487–488
priming store image 488
product choice 486–487
range of effects of music on 483–488
speed of activity and time perception 482–483
store image and purchasing 485–486
visiting 483–484
content-addressable memory 352, 357–358
contextual distance principle 67
Continuator-I 426
*Contrapunto alla Mente* (Banchieri) 410
Cook N 21, 28, 115, 153
Cook P 405
Cope D 405
co-performance 368–369
collaboration 369–370
The Corrs 371
cortical activation, music processing 255
corticotropin-releasing hormone (CRH) 224
Coulam A 368
Craig D 124
Creative Music Therapy (CMT) 497–498
creativity theories
composition 405–407
creative musical thinking, in children 421–427
models of creative thinking 426
Creech A 297, 298
CRH (corticotropin-releasing hormone) 224
Critchley M 184
Cross I 9, 36, 449
Crow B 426
Crozier G 298

Csikszentmihalyi M 291, 298, 413, 425
Cuddy LL 211
Cue-Abstraction theory 113–114
cues
  basic 360
  expressive 360
  interpretive 360
  performance 355, 358, 359–361
cultural capital 435
cultures, musical, environmental influences on individuals 278–279
Cunibile JC 76
Curnow R 482

# D

Dalla Bella S 238, 365
dance 20, 447–449
  Laban technique 368
  modern 448
Dannenbring GL 98
d'Arezzo G 410
Darwin C 3–4, 155
  and neo-Darwinian understandings 7
DAT *see* dynamic attending theory (DAT)
Davidson JW
  learning of musical skills 275, 297
  musical identities 464
  performance 366, 367, 368–369, 370
Davies, Stephen 28, 30
Davies S 404
Davis M 409
Dawson JC 367
de Cheveigné A 74
de Vries JIP 221
*Decentering Music* (Korsyn) 543
declarative knowledge 414
Delalande F 373
deliberate practice, concept 257, 267, 333
Deliège I 114, 115
demand characteristics 132
demandingness model 300–301
DeNora T 30–31, 164, 434
D'Errico F 10
Desain P 327
desk work, and listening to music 433
detection software 324
Deutsch D 98, 111
development, musical *see* musical development
developmental niche 219
Dewey J 555
diatonic modes, scales 54, 55
Dibben N 114, 432, 433
diegesis/non-diegesis 442
Dileo C 519
direction, melodic 144
discipline research, music therapy 518, 521–522
discordant trio clusters 302, 303
disinterested pleasure, notion of 30
Dissanayake E 9, 37
dissonance 51, 52, 53
distraction 431
Doan G 296
Dollinger SJ 162
dominant duo clusters 302, 303
Donald M 411
double duo clusters 303
double-jointedness 390
Dowling WJ 99, 101, 102
Drake C 17, 269
Dreaming journey, Australian Aboriginal song 17, 18
driving, and listening to music 432
du Pré J 295
Dunbar RIM 9
Durkheim E 37, 39
Durrant C 368
dynamic attending theory (DAT)
  behavioural assessments 86–87
  biological assessments 87
  evidence for 86–87
  and metric binding 83–87
  specialized entrainment models 85–86
dynamic duo clusters 302
dystonia 332, 340–341

# E

E major key *63*, 65
E minor key 62
early right anterior negativity (ERAN) 176, 212, 213, 214
EBP (evidence-based practice) 518
EC (evaluative conditioning) 136
Eccles JS 467
echoic memory 107
ecological systems model, human development 279–280
ecstasy 453
education
  formal 473–474
  higher, supporting learning in 311–314
  music learning 421–422
  quantitative research, music therapy 521
Edwards J 516
EEG (electroencephalography) 172, 178, 209, 335, 336, 558
effective practice, defined 265
effects of music, physiological *see also* bodily responses to music
  on biochemical responses 122, *123*, 124
  on blood pressure 122
  on blood volume 124
  chills 124, 156
  on finger, skin or body temperature 124
  on gastric system 124–125
  on heart or pulse rate 122
  on muscular tension 124
  on respiration 124
  on skin conductivity 122
  studies 125
effervescence 37
Ehresman D 75–76
Eibl-Eibesfeldt I 156
Eitan Z 147
electroencephalography (EEG) 172, 178, 209, 335, 336, 558
electromyography (EMG) 124
Ellis RJ 90
embedded time spans, metric hierarchies 81
embodied cognition 355
EMG (electromyography) 124
emotion, in music performance 377–388
  accurate communication to listeners 379–380
  conceptual foundations 377–378
  future research 387–388
  nature of code 385–387
  origin of code 384–385
  performers' views 378–379
  research paradigms 379
  theoretical approach 384–387
emotional contagion 134, 136
emotional expression
  absolutism versus referentialism 141–142
  different emotions, communicating 380, *381–382*, 383–384
  dynamic forces 147
  GERMS model 377–378
  human character/personality 146–147
  implications 145–146
  in music performance *see* emotion, in music performance
  musical structures *see* musical structures
  program music 146
  religious faith 147–148
  response format 142
  social conditions and organization 147
  synthesis of 383–384
emotional memory 356
emotional responses to music

*see also* bodily responses to music
ability of music to induce emotions 132
conditions for eliciting
implications of framework 137–138
theory 135–137, *138*
emotional expression *see* emotional expression
factors influencing
individual 134–135
musical 134
situational 135
and musical expectations 179–180
musical preferences 190
prevalence of emotions 133, *134*
primary issues and methods 132
tonal hierarchies 69
emotion/emotions
*see also* emotional expression; emotional responses to music
cognition disorders, musical 190, 191–192
conflict theory of 555
definition and measurement of 131–132
musical listening 190
and musical structures 238
perception and induction of 131
and performance *see* emotion, in music performance
phylogeny of musical behaviour 223
empathy 404
employment practices, quantitative research, music therapy 520
encoding, and memory 107
encounter, notion of 40
energizing 431
Engeström Y 307, 308
ensemble performance 370
entrainment
body work 434
defined 83
and evolutionary theory 6–7
higher-order processes, music cognition 20
and listening to music 431
entrainment regions 86
epigenesis 219
epinephrine 395
episodic memory 108, 137, 222, 353
ERAN (early right anterior negativity) 176, 212, 213, 214
ERB (equivalent rectangular bandwidth) frequency scale 49
*ergaster* (homo line), in evolutionary theory 10–11
ergonomic aids and advice 391, 393
Erickson R 75
Ericsson KA
expert memory 353
musical potential 257
planning and performance 333
practising 266, 267, 268
sight-reading 348–349
ERPs (event-related potentials) 87, 172, 176, 209, 444
*Erwartung* (Schönberg) 405
ESM (Experience Sampling Method) 31, 133, 164
Essens PJ 82, 83
ethical issues, quantitative research, music therapy 520
ethnomusicology
and Durkheim 37
and evolutionary theory 5, 6
and functions of music 35–36
functions of music 37
grouping and segmentation principles 15
and 'self' 38
evaluative conditioning (EC) 136
event clusters 417
event hierarchies
definitions 60
higher-order processes, music cognition 16–17
and tonal hierarchies 59, 69–70
in Western music 59–60
event-related potentials (ERPs) 87, 172, 176, 209, 444
evolutionary theory
aesthetics 156–157
classical 3
definitions of 'evolution' 3
development 4
evidence, music and evolution 9–11
and music 3–12
recent thought 7–9
universal music processes 5–6
exchange 300
excitation pattern, neural 48, 49
exercise, and listening to music 434
expectation, concept 109
expectations, musical
acoustic and MMN response 173
auditory, measurement 172–176
contexts 173–176
definition of musical expectancy 137
early development, in infants 176–177
and emotional responses 179–180
and melody 173
memory 110–111
mismatch negativity 172, 173
neurobiological basis 171–180
temporal, and auditory-motor interactions 178–179
universals 18–19
Experience Sampling Method (ESM) 31, 133, 164
experientialist approaches 26, 31
*Experiments in Musical Intelligence* (Cope) 405
expert memory 355, 358–359
expressive cues 360
expressive movement 505
expressivist analyses, musical meaning 28
eye-hand span 347


# F


F major key *61*
family, role in supporting learning 295–304
behavioural support 296–297
cognitive/intellectual support 297–298
conflict, potential for 299–300
parent-pupil-teacher interaction 300–304
personal support 298–299
Fauconnier G 406
Faulkner R 425
Fechner G 152
Feld S 29, 38, 39
Feldhusen JF 295
*Feminine endings* (McClary) 364
Ferneyhough B 405
Ferrara L 522
fetal auditory, vestibular and proprioceptive abilities 220–221
fetal 'consciousness' 221–222
fetal motility and proprioception 221
Fialkowska J 361
Finnegan R 32, 40
fission boundary 54
Fitch WT 8
fitness, and music 9
fixations 345
flow, state of 291, 413, 425
fMRI (functional magnetic resonance imaging)
emotional responses to music 136
instruction, effects of 245
mirror neuron network 203
and musical expectations 176
planning and performance 336, 337
and rhythm 88
syntactic processing 213
focal hand dystonia 332, 340–341
Fogelsanger A 448
Folkestad G 426
fossil records, evolutionary theory 10
foundational research, music therapy 520, 522
Fourier analysis, contour code 97

Fox AA 29
Fraisse P 82
Francès R 67
Frege Gottlob 24–25
frequency, and pitch 47–50
frequency coding 48–49
frequency components 48
Freud S 404, 495
Fridja N 125
Friedson SM 39–40
Fujioka T 174
functional magnetic resonance imaging *see* fMRI (functional magnetic resonance imaging)
functions of music
  approaches to 35–38
  balance between 42
  communication 39–40
  in cross-cultural perspective 35–42
  defining musical behaviour 38–39
  and uses of music 36
Furneaux S 347
Fürniss A 15

# G

G major key *61*, 68, 95
Gabriel P 447
Gabrielson A
  emotional responses to music 133, 135
  listening to music 436, 437
  musical structures and perceived expression 142, 143
  peak experiences 454, 455, 456
  performance 327
  Strong Experiences of Music (SEM) project 453
Gadamer HG 404
Gagné F 259
Gallaher MC 447, 448
Galton F 255
galvanic skin responses (GSR) 122
Gan S 463
Gärdenfors P 406
Gardstrom SC
Garland S 300
Garlin FV 481
gastric motility, effects of music on 124–125
Gellrich M 287
Gembris H 275
gender factors 29, 274–275, 287
generalized anxiety disorder 394
generative rules, GERMS model 377
Generative Theory of Tonal Music (GTTM) 69, 77, 113, 114
genes, and environments 223
genetic drift processes, evolutionary theory 7
genius, notion of 405
geriatric population, music therapy 511–512
Geringer JM 270
GERMS model 377–378
Gestalt tradition
  perceptual grouping 500
  and pitch 54
  *prägnanz* 99
  and rhythm 87–88
gestures, motherese 221
Gfeller KE 493, 495
Ghosh JP 408
*Giant Steps* (Coltrane) 409
Giddens A 462, 463
GIM (Guided Imagery and Music) 137, 496, 505, 522
Ginsburg GS 297, 299, 300
Giordano BL 74
Glass R 448
Gockel H 54
Goebl W 328, 329
Goehr L 403–404
Goethe J W von 407
Gold C 519
*Goldberg Variations* (Bach) 153
Goldstein A 155
good continuation principle 54
Good JMM 369
Goodman E 370
goodness of fit ratings, chords 62
Gordon JW 74
Gosling SD 162, 163, 166
Gould G 373
Granot RY 147
Greasley AE 164, 435
Greek philosophical tradition, meaning in music 26
Greenwood D 152
Gregory AH 446
Grey JM 74
Griffiths TD 190
Grolnick W 296
Gromko J 246, 425
grouping
  musical, and timbre 76–77
  and perception 15, 53–54, 500
groups
  collaboration and group composition 425
  group performance, unison 17
  rehearsing in 270
GSR (galvanic skin responses) 122
GTTM (Generative Theory of Tonal Music) 69, 77, 113, 114
Guhn M 124
Guided Imagery and Music (GIM) 137, 496, 505, 522
Guilbault DM 419
Guilford JP 422

# H

Haake AB 433
Haas F 124
*habilis* (homo line), in evolutionary theory 10, 11
habituation 222
Hagen RH 9
Hagendoorn I 448
hair cells 48, 49
Hallam S 260, 261, 269, 270, 277
Handel S 75
Hanslick E 28, 141, 151, 154
  formalist aesthetics of 153, 157
hardware, measuring musical performance
  acoustic instrumentation 324
  mechanical devices 323
  movement and physiological transduction 323–324
Hargreaves D
  consumer behaviour 483, 484, 485
  functions of music 39
  improvisation, musical 418
  musical preferences 161, 163, 165
harmonic priming, and repetition priming 69
harmonic spectra 16, 51
harmonicity 54
harmonious trio clusters 302, 303
harmony 31, 144
Haroutounian J 261
Harrison CS 287
Harter S 289
Hartogh T 222
Harwood D 21
health optimization, musicians
  injury types 391
  manual therapy 392
  physical health issues 390–393
  playing-related injuries, prevention 392–393
  psychological health issues 393–397
  rest and rehabilitation 392
  retraining 392
  treatment of injuries 391–392
heart/pulse rate, effects of music on 122
Hebb D 248
hegemony 29
*heidelbergensis* (home line), in evolutionary theory 10, 11
HEIs (higher education institutions) 312, 314
Heise GA 54
Helmholtz HLF von 72
Henson RA 184
Hepper PG 222
Heschl's gyrus 198, 347
Hess M 356

heterarchy 365
Hetland L 245
Hettema J 280
Hevner K 380
Hewitt A 425
Hickey 423
Hickey M 425, 426
hierarchical organization
  and Australian Aboriginal song 14, 17–18, 19–20
  concept of 'hierarchy' 59
  event hierarchies *see* event hierarchies
  and long-term musical representation 113–115
  and relational processing 16–18
  tonal hierarchies *see* tonal hierarchies
  in Western music 59
hierarchical organization, music 5
higher education, supporting learning in 311–314
higher-order processes 16–18
  and pitch organization 53–56
highland music, Peruvian 36
Hills B 520
Hindustani music 17, 408
Hinton SC 244
history and research
  aesthetic theory 555
  affirmative response to music 554
  clinical behavioural studies 557
  imaging techniques 556, 558
  new psychology of music 552
  physiological responses 555
Hofstadter D 405, 406
Hogg MA 369
holism 300
Hollombe AW 101
Holyoak KJ 405
Hom C 494
hominin (human) lineage, and evolutionary theory 10–11
homuncular order 334
Honing H 327
Hopkins B 219, 221
*How Musical is Man?* (Blacking) 126
*How the mind works* (Pinker) 8
Howe M 297
HR (heart rate) 122, 124
humanistic music therapy 497–498
Hume, David 27
Huron D
  aesthetic perspective 153
  composition 405
  musical expectations 171, 180
  musical memory 110
  musical time 90
  universal music processes 19
Hursch 385
Hyde KL 191
hypotheses 14, 327

# I

IAPS (International Affective Picture System) 444
ICF (intracortical facilitation) 201
ICI (intracortical inhibition) 201
ICMPC (International Conference of Music Perception and Cognition) 540
iconic memory 107
iconic signs 25
identity paradoxes 463
IFG (left inferior frontal gyrus) 179, 198, 199, 204
IHCs (inner hair cells) 48
IHI (inter-hemispheric inhibition) 201
ILPG (intrasulcal length of posterior bank of precentral gyrus) 248
imaging techniques
  functional magnetic resonance imaging *see* fMRI (functional magnetic resonance imaging)
  history and research 556, 558
  magnetic resonance imaging 192
  planning and performance 336
  rhythm 88
imagining 407–408
improvisation, musical
  and body 414, 416–417
  children and music learning 418–419
  and composition 408–409
  computational and grammatical models 417–418
  constraint on processes of 413–414
  interaction and identity 418
  in jazz 415–416
  knowledge limitations 414–415
  real time constraints 414
Imreh G 329
indexical signs 25
infants
  and consonance 231
  early development of musical expectation 176–177
  learning from 229–233
  and maternal performances 230
  and musical memory 231
  and tuning 231–232
infants, learning from
injuries, in musicians 391
inner ear 48
inner hair cells (IHCs) 48
instruction, music
  cognitive transfer 249–250
  effects 245–248
  impact on other skills 244–251
  Mozart effect 244–245
  neurophysiological mechanisms 248–249
instrumental music, and meaning 27
integer frequency ratios, hierarchical organization and relational processing 16
intelligence testing 247
intentionality, and meaning 25
interaction, musical 39, 564
inter-hemispheric inhibition (IHI) 201
inter-key distances, and intra-key hierarchies 60–63
International Affective Picture System (IAPS) 444
International Conference of Music Perception and Cognition (ICMPC) 540
inter-onset time intervals (IOIs) 82, 84, 86
interpretive cues 360
intervals
  dissonant 77
  emotional expression 144
  and memory 111
  pitch 185, 187
  temporal structures 189–190
intracortical facilitation (ICF) 201
intracortical inhibition (ICI) 201
intra-key hierarchies, and inter-key distances 60–63
intrinsic interest 467
IOIs (inter-onset time intervals) 82, 84, 86
IQ (intelligence quotient) 247
isochronous timing 232
Iverson P 76
Iwamiya S-I 443
Iwanaga M 474
Iyer VS 17
Izard C 506
IZO (individualized zone of optimal functioning) 395

# J

Jackendoff R 26, 69, 77, 113
Jagger M 409
James W 132, 248
Jäncke L 365
Jarmusch J 444
Järvinen 415
JAS (joint accent structure) 89–90
jazz
  improvisation 409
  improvisation in 415–416
  musical identities in 464–466
Jensen KL 520

Johnson M 26, 31, 407
Johnson SP 219
Johnson-Laird P 414, 417
joint accent structure (JAS) 89–90
joint issues, musicians 391
Jones MR 20, 85, 90
Jordan S 447
Jørgensen H 266, 267, 270
jouissance 404
Juslin PN
  emotional expression 379–380, 384, 385
  emotional responses to music 133, 135, 137, 138
  musical structures and perceived expression 142
  performance 329

# K

Kalinak K 443
Kaluli of Papua New Guinea, music of 29, 38
Kant I 27, 30, 151, 157
Kaschub M 425
Kemp AE 262, 276, 277
Kendall RA 16
Kenrick DT 280
Kerman J 409
Kessler EJ 64
Kester DB 189
keys
  *see also* tones
  emotional expression 143
  major and minor 60–63
  tonal hierarchies
    inter-key distances and intra-key hierarchies 60–63
    models 63–67, *66*
Kim K-H 443
kin selection processes, evolutionary theory 7
kinaesthetic-image schemas, and meaning 26
kinaesthetic sense 333, 441
Kinsler V 346, 348
Kodály music method 245, 246
Koelsch S 31, 212, 213
Koestler A 405
Kogan N 9
Konečni V 152–153, 163
Korniche LE 348
Korsyn, Kevin 543
Kratus J 419
Kruhmansl 60–61
Krumhansl CL
  melodic processing 100, 102, 103
  performance arts 447
  tonal cognition 60–61, 64, 67
  universal music processes 19
Kulezic-Wilson D 444
Kulieke MJ 297
Kurth E 147

# L

Laban technique, dance training 368
labyrinth, membranous 221
Lacan J 404
Lachenmann H 407
Lahav A 179
Lakoff G 26, 31, 407
Lamont AM 164, 435, 468
Lampe RE 300
LAN (left anterior negativity) 214
Land MF 347
Langer S 142
language
  *see also* speech
  enhancement of, in formal education 473–474
  linguistics, study of 24
  and music *see* language and music
  origins 38
  protolanguage 220
  semantics of 24–25
  syntactic computation 60
  tonal 55–56
  vocal utterance as origin of 40
language and music
  comparative research 214–215
  and evolutionary theory 6, 7
  musical meaning 26, 31
  pitch 56
    *see also* pitch; pitch perception
Lanz D 447
Larson S 31
Laski M 453, 454, 455
Laukka P 133, 379, 385
Law of Situational Meaning 125
learning
  background music, impact of 474–475
  family, role in supporting 295–304
  female, in UK cathedral setting 314–316
  implicit 19–20
  from infants 229–233
  institutions/teachers, role in supporting 307–318
  instrument playing 201–203
  musical expertise, development of 257–258
  outcomes 303, 474
  persistence with 303
  prenatal 222
  statistical 222
  transfer of 269
learning environment, role 270–271
learning musical skills
  age factors 275–276
  cognitive and learning styles 277–278
  environmental influences 278–279
  innate or learned 223–224
  instruments, physiological demands 275
  interactions, individual and environment 279–281
  musical cultures 278–279
  personality 276–277
  sex differences 274–275
  socio-economic status 278–279
Leary T 301
LeBlanc A 165, 166
left anterior negativity (LAN) 214
left inferior frontal gyrus (IFG) 179, 198, 199, 204
Lehmann AC 268, 348–349
Lennox A 372
lens model equation/expanded lens model 385, 386
Leont'ev AN 307
Lerdahl F
  composition 405
  meaning and music 28
  musical memory 113
  timbre 77
  tonal cognition 63–64, 65, 69
Lesiuk T 433
Levelt WJM 51–52, 152
Levenson R 444
Levitin DJ 256
lexical access interactive model 65
Libet B 336
Liegeois-Chauvel C 189
Ligeti G 408
Likert scales 161
Lindström E 142, 143, 378, 383, 456, 457
Lindström-Wik S 436, 437, 453, 454, 455
linguistic memory 355, 357
linguistics, study of 24
Lipscomb S 425
*Lirrga* songs, Wadeye 19, 20
listening tests 383
listening to music 431–438
  *see also* preferences, musical
  arousal and biological influences on behaviour 161–163
  and body work 434
  and brain work 433
  concert-goers, experiences 436–437
  deficits, acquired 184–185, *186*
  and desk work 433
  developmental deficits 190–191
  emotion conveyed to listeners 379–380
  emotional work 434–435

and functions of music 41
live events 435–437
live performances versus recorded music 437
and pain management 434
and physical work 432–433
and pitch space 65
and private study 433
styles of listening 135
and travelling 432
Litle P 163
live events, attending 435–437
motivation for 436
live performances versus recorded music 437
Lochy A 211
long-term memory (LTM) *see* LTM (long-term memory)
Lopukhov F 447
Lord A 415
loudness, emotional expression 143–144
LTM (long-term memory) 108
and musical form 113–115
primary function 109
LtoR (left to the right hemisphere) 201
Lubart 422–423
lullabies 229
Lundin R 152
Luria AR 307
Luzzaschi L 410


# M


M1 (primary motor area) 334
MA (melodic accent) 90
MacDonald R 31, 39, 418
MacDonald RAR 473
MacNamara A 288
Madell G 155
Madsen CK 270
magnetoencephalography (MEG) 248–249, 558
major and minor keys 60–63
*see also* keys
Mandarin, tones of 55–56
Maranto C 520
Marchant GL 297–298, 300
marking time and accent salience 88–91
experimental evidence 90–91
theory 89–90
Marozeau J 74
Marshall S 442, 443, 444
Martinez I 31
Maslach Burnout Inventory 520
Maslow A 452, 453, 454, 458, 497
MBEA (Montreal Battery for the Evaluation of Amusia) 184, 188, 190, 191
McAdams S 76, 77, 78, 99
McAuley JD 86
McClary S 29, 364
McClelland JL 65
McCormick J 267
McCullough M 410
McKeage KM 275
McKinney CH 520
McNamara A 267
McPherson GE 258, 259–260, 267, 288
meaning and music
aesthetic perspective 27–28, 30
affective dimensions 30–31
definition of 'meaning' 24
empirical and theoretical approaches, in psychology of music 30–31
expressivist analyses 28
non-Western contexts 32
performance arts 444
in social and cultural context 28–30
theories of 30–31
theories of meaning 24–26
Western intellectual tradition 24, 26–27
meaning enhancement 431
measure, hierarchical concept of 17
measurement of performance
*see also* performance
algorithmic and statistical modelling 327
analytical methods/discoveries 325
in context 328–329
coordination with/between performers 326
expression, communication and signification 326–327
hardware 323–324
motor programming 325
and psychology of performance 325–327
rhythm and timing 326
software 324–325
media 443
medieval period, meaning in music 27
MEG (magnetoencephalography) 248–249, 558
Mélen M 115
melodic accent (MA) 90
melodic contours
and amusia 211
and memory 111
models 96–98
and perception 15, 98–100
pitch pattern 187
processing of 209–211
structuring 18
tonal hierarchies 67
melodic processing 94–6
melodic structure 18
melody/melodies
compound 53
contour *see* melodic contours
emotional expression 144
Greek philosophical tradition 26
and memory 103
multiple interleaved 54
musical development 235, 236
and musical expectations 173
and octaves 16
processing *see* melodic processing
standard and comparison, examples *68*
tonal and atonal 112
unfamiliar 82
memory
activation 109
auditory 354–355
autobiographical 108, 458
content-addressable 352, 357–358
emotional 356
episodic 108, 137, 222
expert 355, 358–359
explicit and implicit 108
failure 352
familiar and novel material 190
in general 107–109
linguistic 355, 357
long-term *see* LTM (long-term memory)
motor 355
multiple systems 354–358
musical *see* musical memory
narrative 356
perceptual representations 108–109
short-term *see* STM (short-term memory)
sight-reading 346–347
structural 356–357
transnatal 222–223
visual 356
Mendelssohn F 447
mental training 339
Menuhin Y 295
Merriam, Alan 5, 36, 37, 42
meta-analyses, music therapy 519
metacognition 289–290
metre perception, encoding theory 82–83
metre/metres
*see also* metric encoding
*aksak* 20
metric binding, and dynamic attending theory *see* dynamic attending theory (DAT)
musical development 236–237
non-periodic 20
perception *see* metre perception, encoding theory
temporal hierarchies 17
temporal structures 189–190

metric binding, and dynamic attending theory 83–87
metric clusters 85, 86
metric encoding 82–83
metric hierarchies 81
metrical structure 178
Meyer LB 19, 134, 137, 171, 179
Meyer Max 152
middle ear 48
MIDI pianos 323
Miell D 39, 473
Mikiten T 499
Miklszewski K 268
Miller, Geoffrey 8, 9
Miller, George 108
Miller GA 54
Millikan RG 32
Mills AE 262
Milner P 155
Minassian C 378
Minsky M 404
mirror neuron network (seeing-hearing-doing) 126, 203–204
mismatch negativity (MMN) 172, 173, 175, 176, 180
mistuned notes, piano pieces 67
Mitchell LA 434
Mitchell RW 447, 448
Mithen S 9, 39
MMN (mismatch negativity) 172, 173, 175, 176, 180
modulation (key changes) 63
Moelants D 324
Monson I 418
Monteverdi C 410
Montreal Battery for the Evaluation of Amusia (MBEA) 184, 188, 190, 191
mood management, and listening to music 435
Moore BCJ 54
Mor S 103
Morley I 9
motherese 220, 225
mother-infant communication 9
motion, melodic 144
motion principles, GERMS model 378
motivation
  and composition 404
  environmental influences 290
  future research 291–292
  goals and aims 288
  historical perspective 285
  individual characteristics 287–290
  lack of 291
  live events, attending 436
  in music, framework for understanding 285–287
  and musical identities 466–468
  physical, gender and personality factors 287
  and practising 267
  rewards and punishments, from environment 290–291
  self-concept 287–288
motor control 221
motor memory 355
motor programming 325, 365–366
motor skills, acquisition 337, *338*, 339
motor systems
  neuroanatomy and neurophysiology 333–335
  sensory, plasticity 339–340
Motte-Haber H de la 287
movement 505
movement tracking data 366
moving images, and psychology of music 441–446
Mozart effect 244–245
MRI (magnetic resonance imaging) 192
multimedia and music 441–449
Münsterberg H 441, 442
Murninghan JK 371
MUSACT (MUSical ACTivation) model 65, *66*
muscle/tendon injuries, musicians 391
music
  affirmative response to 554
  and brain 208–215
  as communication 563–564
  as communicative medium 39–40
  definitions 6
  emotional responses *see* emotional responses to music
  and evolutionary theory *see* evolutionary theory
  and fine arts 506
  functions *see* functions of music
  as interaction 564
  as/and language 563
  listening to *see* listening to music
  meaning in *see* meaning: in music
  and musical memory 111–113
  on-hold 487–488
  ontogeny 220–223
  origins of study 4
  physiological effects *see* effects of music, physiological
  power of 561–562
  psychology of *see* psychology of music
  as time-oriented structure 60
music analysis 542
*Music and the brain* (Critchley and Henson) 184
*Music Perception* (journal) 540
music therapy
  application to clinical populations
    children 507–508
    geriatric population 511–512
    neurological rehabilitation 508–512
    psychiatric 504–507
  behavioural 494–495
  chronic pain, patients with 531–532
  clinical practice 503–513
  defined 503
  history 504
  humanistic 497–498
  neurologic *see* NMT (neurologic music therapy)
  and pain therapy 529–532
  processes 493–501
  psychoanalytic 495–497
  rehabilitation
    cognitive 510–511
    interventions 508–509
    sensorimotor 509
    speech 509–510
  Relational-Scientific Mediating Model 499–500
  research and evaluation 515–522
  Transformational Design Model 500–501
  treatment process 503–504
MUSical ACTivation (MUSACT) model 65, *66*
musical behaviour
  defining 38–39
  and functions of music 40–41
  phylogeny of 223–225
musical development
  context 241
  defined 235
  developing musical understanding 235–239
  melody 235, 236
  metre 236–237
  musical structures and form 237–238
  natural development and cultural environments 240–241
  pitch 235–236
  production and performance skills 238–239
  rhythm 236–237
  style discrimination 238
  tempo 236–237
  theory 239–241
  timbre 237
  timing 236–237
  tonal hierarchies 235
musical expectations *see* expectations, musical
musical form, emotional expression 145
musical identities
  defining 462–463
  in jazz 464–466
  and motivation 466–468

musical meaning *see* meaning and music
musical memory
*see also* memory
associative chaining 352, 353
expectations 110–111
infants, learning from 231
long-term, and musical form 113–115
melodies 103
and music 111–113
performing from 352–361
schemas, role 353–354
tonality and contour in 100–102
musical potential *see* potential, musical
musical preferences *see* preferences, musical
musical skills *see* skills, musical
musical structures
*see also* temporal structures
absolutism versus referentialism 141–142
definitions 141
and emotion/emotional expression 142–144, 238
and form 237–238
implicit knowledge of 19
musical development 237–238
and perceived expression 141–148
referential meanings 146–148
musical tension *see* tension, musical
musical time *see* time, musical
musicians
between-group/within-group structural changes 198
cross-sectional studies 197–198
health optimization *see* health optimization, musicians
learning to play instruments, cognitive and brain changes 201–203
and non-musicians, brain differences 197–201, 339–340
within-musician differences, causal effect 200–201
myelination 340

# N

Nagel JJ 287
Nakajima Y 88
NAMT (National Association for Music Therapy) 515, 516
Narmour E 171
narrative memory 356
National Association for Music Therapy (NAMT) 515, 516
Nattiez J-J 39
natural selection 3–4
nature-nurture debate, musical potential 255
neo-hedonism 154–156
nerve compression disorders, musicians 391
Nettl, Bruno 5, 14, 37
and functions of music 36, 42
neural excitation patterns 48, 49
neural substrate 192
neuroanatomy 447
neurobiological factors, musical expectations *see* expectations, musical
neurobiology, history and research 557
neurologic music therapy *see* NMT (neurologic music therapy)
neurological disorders, music therapy research 519
neurological rehabilitation
cognitive 510–511
interventions 508–509
sensorimotor 509
speech 509–510
neurons 107
neurophysiological mechanisms, music instruction 248–249
*ngintaka*, song series 18
Nielson SG 270
Nielzen S 507
Nilsson B 426
NMT (neurologic music therapy) 498
neuroscientific foundation 528–529
and R-SMM 499
noise, and music 38
North AC
aesthetic perspective 161, 163
consumer behaviour 483, 484, 485, 487
listening to music 435
notation
invented 425
and sight-reading 345–346
Western model 344
note density, emotional expression 143

# O

occupational setting, listening to music in 433
Ockleford A 472
octaves 16
Odena O 425
OHCs (outer hair cells) 48, 49
Olds J 155
Ollen J 153
Olson RK 53
Olszewski-Kubilius P 297
Omega Sign 200, 201
O'Neill S 289, 300, 464
ontogeny of music 220–223
opera 446
operant conditioning 225
*Origin of Species* (Darwin) 3
ornamental events 59
ornamentation 383
Oruwari V 370
oscillatory activity, dynamic attending theory 84–85, 86
otoliths 221
outer hair cells (OHCs) 48, 49
over-use injuries 390
Owen K 481
oxytocin 224–225

# P

Paananen P 237, 418
Pachet F 426
pain management, and listening to music 434
pain therapy, and music therapy 529–532
chronic pain 531–532
Palmer C 269, 326
Panksepp J 124
Pantev C 200
Panzarella R 453, 454, 455, 457
Paraskeva S 77, 78
Paravicini D 309–311
parentese 229
parenting style 397
parent-pupil-teacher interaction
categories 301
control and responsiveness 300–301
parental support 301–302
persistence with learning 303
pupil age 302–303
pupil learning outcomes 303
versatility, parental 303–304
violin teaching, Suzuki method 296
Parker C 415
Parker E 409
Parkins DN 249
Partch H 55
partials (frequency components) 48
part-whole concept 250, 268–269
Patel AD 83, 211, 213
pattern matching 51
pauses, emotional expression 144–145
peak experiences 452–460
effects of 457–458
fundamental, in music literature 452–454
methodology 458–459
nature of 455–457
precursors and predictors 454–455
prevalence of emotions 133

Pearson JL 162
Peirce CS 25
perceived cost 467
perception, music
  categorical 55
  functional neuroanatomy 557
  grouping and segmentation 15
  and melodic contours 15, 98–100
  notation and sight-reading 345–346
  and pitch *see* pitch perception
  rhythm, empirical evidence 88
  and timbre 77–78
  tonal hierarchies, influences on 67–69
  tonality and contour 98–100
  understanding perceptual and cognitive processes 563
perception-action cycle 20
percussionists 277
Peretz I 11–12, 69, 189, 191, 211
performance
  body movements, musically expressive effects 366–368
  communication 371–373
  co-performance 368–369
  emotion in *see* emotion, in music performance
  ensemble 370
  measuring *see* measurement of performance
  motor programming 365–366
  movement and collaboration in 364–374
  movements for audience effect 373–374
  music pedagogy 368
  musical development 238–239
  optimizing 395–396
  and planning *see* planning and performance
  preparation strategies 269
  psychology of 325–327
  public, anxiety caused by 394
  sole performers 366–368
  words, actions including and beyond 370–371
performance anxiety
  *see also* psychological health issues, musicians
  phenomenology and determinants of 394–395
  prevention 397
  public performance 394
  and quality of performance 395
  treatments for 396–397
performance arts
  dance 20, 447–449
  future research 449
  lifespan experience 449–450
  multimedia and music 441–449
  theatre 446–447
  trans-art congruencies 449
  visual and kinaesthetic senses 441
performance cues 355, 358, 359–361
performance modelling 327–328
  in context 328–329
periodicity/periodicity pitch 20, 51
peristalsis 124
Perry M 415
personality 276–277, 287
PET (positron emission tomography) 245, 336, 448
phase coincidences, dynamic attending theory 84
phase locking 50
phylogeny of music 219
  musical behaviour 223–225
physical responses to music 125–126
physical work, and listening to music 432–433
physiological responses 121–125, 555
Piano Camera 323
piano/piano playing
  classical pianists 333
  mistuned notes, detection 67
  skilled pianists *338*
  Sudnow 416
  types of piano 323
Pinker, Steven 8, 9
pitch
  absolute 111
  change in, and timbre relations 74
  cognition disorders 191
  defined 47
  emotional expression 144
  and evolutionary theory 5
  and frequency 47–50
  high 144
  and higher-order processing 53–56
  intervals 185, 187
  and jazz 415
  in language 56
    *see also* language
  musical development 235–236
  pattern 187–188
  and perceptual/cognitive restraints 16
  and speech 55, 56
  and timbre 74
  tonal structure 188–189
pitch chroma 59
pitch classes 59, *63*, 64
pitch declination 110
pitch height 59
pitch perception
  complex tones 50–53, *52*
  higher-order processing and organization of pitch 53–56
  music-specific processes and constraints 54–56
  psychophysics 47–53
pitch proximity 19, 53, 54, 110
pitch reversal 19
Pitts S 435, 437
place coding 48, 50
planning and performance 332–342
  neuroanatomy and neurophysiology of motor systems 333–335
  observing planning in brain 335–337
plasticity 199, 333, 339–340
play, concept 404
playing-related injuries, prevention 392–393
Plomp R 51–52, 152
plural pleasures, neo-hedonism 154–156
PMA (premotor area) 334, 335
POCO performance research environment 324
polyphonic music 174–175
Pool MM 433
Poorman A 246
positron emission tomography (PET) 245, 336, 448
postnatal behaviour, prenatal influences on 225
posture, injury prevention 391, 392–393
potential, musical 255–262
  actualization of 261
  general skills, development 258–259
  learning, role of 257–258
  measuring 259
  nature-nurture debate 255
  prodigies 256–257
  recent thought 259–261
  savants 255–256
Povel D-J 82, 83
PR (pulse rate) 122
practising
  appropriate approaches, adopting for particular tasks 268
  conduct of practice, strategies 268–269
  effective practice, defined 265
  evaluation of practice 269–270
  goal setting 268
  interpretation, developing 268
  learning environment, role 270–271
  mental practice 269
  meta-strategies 270
  monitoring effectiveness of 269–270
  motivation and drop out 267
  organizing of practice 267–268
  performance 269

planning and performance 341–342
planning strategies 267–268
quality of 267–270
quantity of 265–267
*prägnanz*, Gestalt law of 99
preferences, musical
*see also* listening to music
arousal and biological influences, on listening behaviour 161–163
contextually grounded understanding 163–164
effects 164–165
emotion 190
in everyday life 164
explaining/predicting 165–166
individual differences 162–163
methods of study 160–161
social identity effect 165
Preistley M 495–496
premotor area (PMA) 334, 335
premotor cortex 179
prenatal development 219–233
*see also* infants
fetal auditory environment 221
fetal auditory, vestibular and proprioceptive abilities 220–221
fetal 'consciousness' 221–222
fetal motility and proprioception 221
fetal-maternal communication 224–225
innate or learned infant musical skills 223–224
learning 222
ontogeny of music 220–223
phylogeny of musical behaviour 223–225
prenatal influences on postnatal behaviour 225
transnatal memory 222–223
Prespa Albanian song 37
Pressing J 409, 417
Price HE 160
primary domain regions, brain 197–198
primary motor area (M1) 334
priming paradigm, chords 68–69
Prince V 261
principal components analysis 327
probe tones 95
problem-solving, sight-reading as 347–348
procedural knowledge 414
prodigies, musical 256–257
production skills
acquired disorders 192–193
musical development 238–239
profession research, music therapy 520, 522
program music, emotional expression 146
proprioception, prenatal development 220–221
prosody 27
protean individuals 8
protolanguage 220
protomusic 220
proximity, pitch 19, 53, 54
Pruett K 300
pseudo-polyphony 53–54
psychiatry, application of music therapy 504–507
future directions 506–507
interventions 505–506
psychoacoustics 47
psychoanalysis, music therapy 495–497
psychological health issues, musicians
anxiety, in public performance 394–395
cognitive, behavioural and cognitive behaviour therapies 396–397
performance anxiety *see* performance anxiety
stressors experienced 393–394
psychology
and composition 404–405
evolutionary 8
of performance 325–327
psychology of music
*see also* music
and meaning 30–31
and moving images 441–446
new 552
quantitative research, music therapy 520–521
*Psychology of Music* (journal) 539, 541
psychoneuroimmunology 122
public transport, and listening to music 432
pulse 110–111
pupillary reflex, effects of music 125
pure tones 47–48, 50


# Q


qualitative research, music therapy 521–522
quantitative research, music therapy
autistic spectrum disorders 519
discipline 518
education 521
employment practices 520
ethical issues 520
foundational 520
medical problems 519
medicine 521
neurological disorders 519
profession 520
psychology of music 520–521


# R


RA (rhythmic accent) 90
rāgs, tones in 17
radial net 365
Rampazzo C 324
random fluctuations, GERMS model 378
range, melodic 144
RAS (Rhythmic Auditory Stimulation) 519
RATN (right anterior-temporal negativity) 212, 213
Rauscher FH 244, 247
reference, and meaning 24, 25, 26
referentialism, versus absolutism 141–142
rehabilitation, neurological
clinical outcome studies 526–528
cognitive 510–511
interventions 508–509
sensorimotor 509
speech 509–510
rehearsing, in small groups 270
Reid A 277
relational processing, and hierarchical organization 16–18
Relational-Scientific Mediating Model (R-SMM), music therapy 499–500
relaxation, and listening to music 434
religious faith 37, 147–148
reminiscence, and listening to music 435
Rentfrow PJ 162, 163, 166
Renwick JM 288
repetition priming, and harmonic priming 69
Repp BH 324, 327
representative design 383
research
biological and cultural 564–566
collaboration in 566–567
and history *see* history and research
methodological limitations and advances 566
and music therapy 518–522
qualitative 521–522
quantitative 518–521
residue pitch 51
resistance 29
resonance, dynamic attending theory 84
respiration, effects of music on 124
reversal, pitch 19
rhetoric, meaning in music 27

rhythm
defined 112
emotional expression 144
'free' 20
and improvisation 416
measurement of performance 326
as mnemonic device 18
musical development 236–237
perception, empirical evidence 88
pulse 110
temporal structures 189–190
theory 87–88
rhythmic accent (RA) 90
Rhythmic Auditory Stimulation (RAS) 519
rhythmic pattern 178
Rickard N 124
Riecker A 193
Ries H 124
right anterior-temporal negativity (RATN) 212, 213
Roads C 405
Robinson SR 219
rock stars, and sexuality 8
Romanticism 146
Roseman M 38, 41
roughness, auditory 77
Rousseau J-J 38
Royal M 90
R-SMM (Relational-Scientific Mediating Model), music therapy 499–500
Rubenstein A 361
Rubin D 353, 354, 355
Rumelhart DE 65
Ruthman A 426

# S

S1 (primary somatosensory area) 334
Saarinen J 174
saccades 345, 346
Salomon G 249
Sandene BA 289
savants, musical 255–256, 309–311
Sawyer K 418
Sawyer RK 465
scales, musical
chromatic 54, 59
degrees of 54, 55
hierarchical organization and relational processing 16
minor harmonic keys 60
multidimensional scaling 72, 76
scale-step context 111
Schellenberg EG 171, 247, 250
schemas 26, 108, 109, 268
schematic expectations 19
Schenk DL 447
Schlaug G 200
Schmidt CP 289
Schmuckler MA 95, *96*, *97*, 98, 99, 100, 103
Schön D 192
Schönberg A 405
school years, supporting learning in 316–318
*see also* musical development
Schopenhauer Arthur 148
Schouten JF 51
Schulkind MD 137
Schumann C 295
Scott WA 392
Scruton, Roger 28, 30
Sears W 124
Seashore, Carl 259, 323, 553
Seddon F 425, 426
segmentation, perceptual principles 15
self-actualization 454
self-concepts/self-efficacy 38, 267, 287–288
self-theories 466
SEM (Strong Experiences of Music) project 453, 454, 455, 459
semantic long-term memory 108
semantic theories 24–25
SEM-DS (Strong Experiences of Music Descriptive System) 453
semiotic ideas 25, 39
semitone 16, 55
sense, and meaning 24, 25, 26
sensorimotor abilities 275–276
sensorimotor rehabilitation, music therapy 509
sensory homunculus 334
sensory motor systems, plasticity 339–340
sequential grouping, perceptual principles 15
sequential streaming 53
Serafine ML 114
serial chaining 353
serial time spans/ratios 81, 82
sexuality, and evolutionary theory 8
Shaffer LH 326, 366
Shahidullah S 222
Shavinina LV 257
Shaw J 261
short-term memory (STM) *see* STM (short-term memory)
sight-reading 344–350
inner hearing 347
memory processes 346–347
and music notation, perception 345–346
as problem-solving 347–348
skills, acquiring 348–349
signature performances 230
signature tunes 229
significance, and meaning 24
signs 25, 26, 41
silence, weeks of in Peruvian Andes 35
Simonton DK 446
simulation, algorithmic and statistical modelling 327
simultaneous streaming 53
Singley K 250
Skewes K
skills, musical
innate or learned 223–224
instruction, impact 244–251
learning *see* learning musical skills
and sight-reading 348–349
skin conductivity, effects of music on 122
Skinner BF 494
Slobin M 29
Sloboda JA
chills 124
cognitive and intellectual support 297
Experience Sampling Method of 31, 133
factors influencing emotions 134
imagining 408
improvisation 408
and listening to music 435
on motor programming 365
musical preferences 163
peak experiences 455, 458, 459
practising 268
psychology and composition 404
and sight-reading 347
SMA (supplementary motor area) 334, 335, 337
Smith PC 482
Smotherman WP 219
Sober E 7–8
social behaviour, and music 11
social efficacy of music, and evolutionary theory 7
social identity
musical preferences 165
presentation and confirmation 435
social species, behaviours 8
sociocentric self 38
socio-economic status, learning musical skills 278–279
solo leader clusters 302, 303
somatic anxiety 394, 395
somatization disorders 392
somatotopic order 334
songs
*see also* Australian Aboriginal song, hierarchical structure; Prespa Albanian song
Australian Aboriginal 17
fast and slow, pairing of 19
and speech 39
Sosniak LA 296

sound/sounds
common, in music 51
complex tones 50–53, *52*
and fetal development 220, 221
and frequency 49
music as sound 5
sensory encoding of 208–209
Spahn C 392
specialized entrainment models, dynamic attending theory 85–86
specificities 73
speech
*see also* language; vocal utterance
and pitch 55, 56
and song 39
speech rehabilitation, music therapy 509–510
Speilberg S 442
'Spencer's law' 385
SPL (superior parietal lobe) 198
SSIRH (syntactic integration resource hypothesis) 213
staccato 144
standard paradigms 379
Standley JM 494, 519
statistical learning 222
statistical universals, concept 14
Stauffer S 426
Steedman M 417–418
Steele K 75
Steinberg R 507
step sizes, scales 55
Stephens J 426
stepwise motion 144
Stern DN 224
Sternbach DJ 393
Sternberg 422–423
Stevens C 5, 11
stimulative music 122
stimuli, emotional expression 142
stimulus complexity 30
stimulus frequency 49–50
STM (short-term memory) 107, 108, 109, 110
Stoeckig K 68
Stokes M 37
store image, consumer behaviour 485–486, 488
Strand K 426
strategies, musical practising
*see also* practising
evaluation of practice 269–270
meta-strategies 270
part-whole 250, 268–269
and performance 269
tempo, increasing 269
transfer of learning 269
variable 268
Strauss, Richard 146
Stravinsky Igor 99, 409
stream formation 53–54
stream segregation 53
strength or endurance conditioning, injury prevention 393
stretching 393
Strong Experiences of Music Descriptive System (SEM-DS) 453
Strong Experiences of Music (SEM) project 453, 454, 455, 459
structural cues 360
structural memory 356–357
studying, and listening to music 433
Stumpf C 152, 552
style discrimination, musical development 238
stylistic unexpectedness, GERMS model 378
Sudnow H 416
Sugarman JC 37
superior parietal lobe (SPL) 198
supplementary motor area (SMA) 334, 335, 337
Suzuki method of violin teaching 295–296
Swaine J 39
symbol manipulation 410
symbolic representation 41
symbolism 25
synapses 107
synchronization 326
synchrony, higher-order processes, music cognition 20
syntactic computation, language 60
syntactic integration resource hypothesis (SSIRH) 213
syntactic processing, brain 211–214
systemic design 383
systems, terms within 25

# T

TA (temporal accents) 90
tactus 110
Tagg P 146
*taksim*, Turkish 20
tal and metre 17
tapping, synchronized 83, 126
Tarrant M 463
Tarski A 25
Tchaikovsky P 456
TDM (transformational Design Model), music therapy 500–501
teachers/institutions, role in
supporting learning 307–318
enhancement of language and literacy skills 473–474
female learning, UK cathedral setting 314–316
in higher education 311–314
in lower secondary school 316–318
in musical savant 309–311
particular teaching approaches 473
social and personal learning outcomes 474
technique, injury prevention 393
technology
composition 410
impact 562
music, role 426
Tellegen Absorption Scale 446
Temiar village (Malaysia), healing ceremony in 35, 38
temperature, effects of music on 124
tempo
fast and slow 143
musical development 236–237
strategies for increasing 269
temporal accents (TA) 90
temporal coding 49–50
temporal coherence boundary 54
temporal expectancies 20, 83
temporal families 51
temporal hierarchies, higher-order processes, music cognition 17
temporal structures 5, 189–190, 191
tension, muscular 124
tension, musical 77–78
tension-relaxation structures, in music 31
terraced melodic shape 18
Tervaniemi M 174
Thagard P 405
Thayer JF 444
theatre 446–447
therapeutic music experience (TME) 501
Thomas DA 27
Thompson WF 103
timbral blend 77
timbral intervals 75–76
timbre
and acoustic correlates 73–74
and auditory attributes 72–74
cognition disorders, musical 189
emotional expression 144
and music perception 77–78
musical development 237
and musical grouping 76–77
and musical tension 77–78
and pitch change 74
as source identity vehicle 74–75
space models 73–74
timbre-space work 76–77
time, musical
and improvisation 416
marking time and accent salience 88–91
metre perception 81–87
musical development 236–237
rhythm 87–88

time-shrinking phenomena 88
Timmers R 383
*The Time of Bells* (Feld) 38
Titon JT 29
TME (therapeutic music experience) 501
TMS (transcranial magnetic stimulation) 201
Tobudic A 328
Todd NPM 326, 328
Toiviainen P 416
Tolbert E 7
Tomovski 95
Tomovski R *96*, *97*, *98*
tonal cognition, Western music, tonal and event hierarchies 59–60
tonal consonance 51–52
tonal hierarchies
  and cognitive psychology 65
  definitions 60
  and event hierarchies 59, 69–70
  higher-order processes, music cognition 16, 17
  intra-key hierarchies and inter-key distances 60–63
  models 63–67, *66*
  in Western music 59–60, 67–69
tonal languages 55–56
tonal music, defined 59
Tonal Pitch Space Theory (TPST) 63–64, 65
tonal profile 415
tonality
  defined 112
  emotional expression 144
  and jazz 415–416
tonality models, melodic processing 94–96
tone-deafness 190, 191, 210
tones
  complex, pitch of 50–53, *52*
  emotional expression 144
  grouping and segmentation perceptual principles 15
  musical development 235
  and octaves 16
  pure 47–48, 50
  set of 60, 70
tonic finding procedures 95
tonotopic coding 48, 49
TPST (Tonal Pitch Space Theory) 63–64, 65
training, part-whole transfer of 268
Trainor L 174
transcranial magnetic stimulation (TMS) 201
Transformational Design Model (TDM), music therapy 500–501
transmission models 39
travelling, and listening to music 432
Trehub SE 55, 162
Trevarthen C 9, 224
trigger point therapy 392
trill threshold 54
Truitt FE 346
Trusty J 300
truth values 25
Tulving E 108
Tumbuka people, Malawi 39–40
tuning, universal to cultural 231–232
Turino T 36, 40
Turner M 406

# U

universal music processes
  definitions 14
  and evolutionary theory 5
  higher-order processes *see* higher-order processes, music cognition
  implicit knowledge of musical structures 19
  multimodal processes and integration 20–21
  musical expectancies 18–19
  perceptual and cognitive constraints 15–16
uses of music 36

# V

Valentine E 266
Van den Toorn PC 99
Van Handel L 31
Van Knippenberg D 369
van Noorden LPAS 54
Västfjäll D 135, 137, 138
VBM (voxel-based morphometry) 337
Venda children's music 29
veridical expectations 19
vestibular system 220–221
vicarious functioning 386
Vienna Symphony musicians 122
violin teaching, Suzuki method 295–296
violinists 266
virtual pitch 51
visiting, and effects of music 483–484
Vispoel WP 289
visual imagery 136–137
visual memory 356
vocal utterance 38, 40
*Vom Musikalisch-Schönen* (Hanslick) 151
von Helmholtz H 152
*Vorschule der ästhetik* (Fechner) 152
voxel-based morphometry (VBM) 337
Vuvan D 99
Vygotsky LS 241, 307

# W

Wagner C 265, 266–267
warming up and stretching, injury prevention 393
Watt RJ 146
Weaver HE 346
Webern A 409
Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R) 246, 247
Weisberg RW 406
Welch G 425
Wertheim N 192
Wessel DL 75–76, 77
*West End Blues* (Armstrong) 409
Western intellectual tradition, musical meaning 24, 26–27, 32
Western music
  emotional expression 386
  event hierarchies 59–60
  expectations 110
  metrical cycles existing within 111
  notation 344
  tonal hierarchies 59–60, 67–69
Whipple 519
Whitaker NL 269
Widmer G 328, 329
Wiesendanger M 366
Will U 16
Williamon A 259, 266, 369
Williams A 75
Williams R 35, 39, 42, 372–373
Williams syndrome 256
Williamson VJ 432
Wilson D 7–8
Wilson GD 395, 418
Windsor WL 90
Winner E 245
Winold H 269
Wolfe DE 494
Wong PCM 208–209
working memory 111
world music 40
WPPSI-R (Wechsler Preschool and Primary Scale of Intelligence-Revised) 246, 247
Wu A 372
Wundt W 152

# X

Xenocrates 504

# Y

Yamaha pianos 323
Yerkes-Dodson law 475
*Yesterday* (McCartney) 161, 162
yoga, and listening to music 434
yoiks, North Sami 19
Young S 419
Younker B 426

# Z

Zajonc R 152
Zatorre R 124
Zdzinski S 297
Zillman D 463
Zuckerkandl V 39
Zuckerman M 163
zygomaticus activity, studies on 125

The *Oxford Handbooks* series is a major initiative in academic publishing. Each volume offers an authoritative and state-of-the-art survey of current thinking and research in a particular subject area. Specially commissioned essays from leading international figures in the discipline give critical examinations of the progress and direction of debates. *Oxford Handbooks* provide scholars and graduate students with compelling new perspectives upon a wide range of subjects in psychology.

**Susan Hallam** is Dean of the Faculty of Policy and Society at the Institute of Education, University of London, UK.

**Ian Cross** is Director of the Centre for Music & Science in the Faculty of Music, University of Cambridge, UK.

**Michael Thaut** is Dean, School of the Arts, and Director of the Center for Biomedical Research in Music at the Colorado State University, USA.

Related titles published by
OXFORD UNIVERSITY PRESS

**The Social and Applied Psychology of Music**
A North & D Hargreaves

**Communicative Musicality: exploring the basis of human companionship**
Edited by S Malloch & C Trevarthen

**Music, Language and the Brain**
A Patel

**The Natural Musician: on abilities, giftedness and talent**
D Kirnarskaya

# The Oxford Handbook *of* MUSIC PSYCHOLOGY

- Presents an exhaustive overview of the field
- Includes 52 chapters written by people at the forefront of their fields
- Written to be accessible to both students and researchers
- Essential for anyone with an interest in music, psychology, and brain function

CONTRIBUTORS

Bronwen Ackermann, Eckart Altenmüller, Richard Ashley, Kristen R. Begosh, Emmanuel Bigand, Tim Byron, Roger Chaffin, Martin Clayton, Annabel J. Cohen, Andrea Creech, Ian Cross, Simone Dalla Bella, Jane W. Davidson, Shannon de l'Etoile, Alf Gabrielsson, Helena Gaunt, Bruno L. Giordano, Alinka Greasley, Timothy D. Griffiths, Susan Hallam, David J. Hargreaves, Donald A. Hodges, David Huron, Corene Hurt-Thaut, Jonathan Impett, Harald Jørgensen, Patrik N. Juslin, Dianna T. Kenny, Reinhardt Kopiez, Alexandra Lamont, Andreas C. Lehmann, Anne Kathrin Leins, Topher R. Logan, Stephen McAdams, Gary McPherson, Raymond MacDonald, Dorothy Miell, Adrian C. North, Adam Ockelford, Richard Parncutt, Aniruddh D. Patel, Bénédicte Poulin-Charronnat, Frances H. Rauscher, Mari Reiss Jones, Gottfried Schlaug, Mark A. Schmuckler, Sabine Schneider, John Sloboda, Robert Snyder, Ralph Spintge, Thomas Stainsby, Catherine Stevens, Lauren Stewart, Michael Thaut, Elizabeth Tolbert, Laurel J. Trainor, Sandra E. Trehub, Katharina von Kriegstein, Jason D. Warren, Peter R. Webster, Graham Welch, John Whaley, Barbara Wheeler, Aaron Williamon, W. Luke Windsor, Robert J. Zatorre

OXFORD
UNIVERSITY PRESS

**www.oup.com**

ISBN 978-0-19-929845-7
9 780199 298457



Eckart Altenmüller
Mario Wiesendanger
Jürg Kesselring

# MUSIC, MOTOR CONTROL AND THE BRAIN

# Music, Motor Control and the Brain

Edited by

Eckart Altenmüller
Mario Wiesendanger
Jürg Kesselring

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

© Oxford University Press 2006

The moral rights of the author have been asserted
Database right Oxford University Press (maker)

First published 2006

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose the same condition on any acquirer

British Library Cataloguing in Publication Data

Data available

Library of Congress Cataloging in Publication Data

Data available

Typeset by Newgen Imaging Systems (P) Ltd., Chennai, India
Printed in Great Britain
on acid-free paper by
Biddles Ltd., King's Lynn, Norfolk

ISBN 0–19–853000–5 978–0–19–853000–8 (Hbk.)
ISBN 0–19–929872–6 978–0–19–929872–3 (Pbk.)

10 9 8 7 6 5 4 3 2 1

# Preface

In a delightful short scene about the protagonists Mozart and Salieri, Alexander Pushkin (1799–1837) described the different characters of the two contemporary composers. In the opening monologue, Salieri contemplates his career, how he mastered initial difficulties in executing the 'dry finger movements' until they were 'adequate for his ears'. But now, having met Mozart, Salieri develops a great jealousy toward his young, brilliant competitor, lamenting that he remained an artisan of music: 'I strangled and dissected my music—it was algebra, not music'. The scene ends with Salieri poisoning Mozart, now considered as a most unlikely story! What Pushkin is telling us in this paradigm, however, has to do with the two aspects of mastering music: the technicalities against the full flavour of music creation and performance. There is no doubt that music research deals with many of the difficult technical problems, especially of music performance. Helmholtz (1821–1894) was trained as a medical doctor in Berlin, but was actually most interested in mathematics, physics, as well as in the emerging physiology of the time. Moreover, he was also a passionate piano player (apparently of concert quality). His interest and love for music induced him to write his famous *opus magnum*, 'Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik' (Helmholtz 1863). His was probably the first physical–physiological approach to the world of music, incorporating knowledge that emerged from the period of the Enlightenment.

Since Leopold Mozart's school of violin playing, numerous books of musical education have provided information on how to tackle the practical-technical problems. Scales and etudes have to be practiced again and again over many years. But even mastering tricky and difficult exercises is no guarantee of becoming a great musician. Many factors are, however, more difficult to prescribe in details, for example modulations and articulations of tone sequences, timbre, and intentional variations of tempo that all contribute to the essential emotional feeling. But these factors contribute strongly to the excellence of the individual performing artist.

The chapters of this book deal much with the nevertheless essential physiological mechanisms and technical rules of the musician's performance—Salieri's contribution in Pushkin's sketch! A number of chapters provide new knowledge on the extraordinary execution of complex, slow and fast movements on the musician's instrument, or vocal cords, and about the integration of performance with the auditory return. Our choice to emphasize motor performance was motivated by the fact that technology for efficient measurement of biomechanical, auditory, and neural representations of music performance has recently become available. Chapters on timing and rhythm, important ingredients of music performance, are also included. The selected body segments, notably hand,

fingers, and airway effectors, have to be coordinated in complex synergistic actions, often also with different temporal constraints between the two limbs. Virtuoso music requires incredibly high speeds, particularly of finger movements. The multiple spatial–temporal selection of effectors acting on the various instruments have their physical and physiological constraints that the brain has to deal with. In turn, auditory feedback is also transmitted to the brain after a short delay. Thus, two major streams (motor and sensory) are running almost in parallel.

We are now in the era of brain imaging, the great challenge being to learn about representational 'hot spots' of music, i.e. the locations and constellations of activated neuronal assemblies during passive listening to music as well as during its (more difficult!) execution. It has become clear that *modules* are variably activated as interconnected neural networks—there is no single large centre for music in the brain.

Professional musicians play for thousands of hours during their lives. Unfortunately, as a result of overuse, musicians may develop various pathologies such as musician's finger cramps (hand dystonia). Health problems in performing artists ('Apollo's curse') are not rare and may tragically end a musician's career. Fortunately, new insights into pathophysiology shed light on the cause of musician's cramp. Mapping studies have revealed brain representations of the effectors involved (notably fingers) that tend to be less discretely delimited compared with the normal situation. As a consequence the independent use of effectors, like two fingers, is disturbed or lost. Medical therapies and novel approaches to retraining over prolonged periods have led to improvements. Examples of remarkably increased cortical representations in professional musicians have been observed. This plasticity depends on the behavioural use: there is an ongoing give and take! Brain plasticity in professional musicians appears to be outstanding, probably because the specific effectors, like the fingers, are used on a daily basis and often for hours to control the sounds of the music instrument.

We thank Oxford University Press and the contributing authors for this publication. Most of the chapters are based on an international conference that took place on the Monte Verità, Ascona, Switzerland, which was organized by Mario Wiesendanger with the help of Eckart Altenmüller and Jürg Kesselring. Our thanks go to the Foundation Stefano Franscini of the Eidgenössische Technischen Hochschule, Zürich, Switzerland.

Helmholtz, H. (1863). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. F. Vieweg & Sohn, Braunschweig.

Pushkin, A. (1998). Mozart und Salieri—Szene aus dem 'Faust' (translated by Alexander Nitzberg from Russian into German). Grupello Verlag, Düsseldorf.

# Contents

Contributors *ix*

Part 1 **History**

**1** Historical increases in expert music performance skills: optimizing instruments, playing techniques, and training *3*
*Andreas C. Lehmann*

Part 2 **Psychology**

**2** From cognition to action *25*
*Lutz Jäncke*

**3** The nature of memory for music performance skills *39*
*Caroline Palmer*

**4** Musical synchronization *55*
*Bruno H. Repp*

Part 3 **Movement analysis**

**5** Hand movements and musical performance *79*
*Thomas E. Jerde, Marco Santello, Martha Flanders, and John F. Soechting*

**6** Movement analysis in pianists *91*
*Hans-Christian Jabusch*

**7** Fingering and bowing in violinists: a motor control approach *109*
*Mario Wiesendanger, Andreas Baader, and Oleg Kazennikov*

**8** Movements and analysis of drumming *125*
*Sofia Dahl*

Part 4 **Representation in the brain**

**9** Brain structures of musicians: executive functions and morphological implications *141*
*Gottfried Schlaug*

**10** The motor representation in pianists and string players *153*
*Lutz Jäncke*

**11** Brain activation during piano playing *173*
*Marc Bangert*

**12** Brain activation during string playing *189*
*Arto Nirkko and Rumyana Kristeva*

**13** 'Singing in the (b)rain': cerebral correlates of vocal music performance in humans *205*
*H. Ackermann, D. Wildgruber, and A. Riecker*

**14** Sensory–motor networks in singing and speaking: a comparative approach *223*
*Reyna Leigh Gordon, Amélie Racette, and Daniele Schön*

**15** The role of inhibition in the motor control of finger function *239*
*Christian Gerloff and Friedhelm Hummel*

Part 5 **Apollo's curse—loss of motor control in musicians**

**16** The end of the song? Robert Schumann's focal dystonia *251*
*Eckart Altenmüller*

**17** Epidemiology, phenomenology, and therapy of musician's cramp *265*
*Hans-Christian Jabusch and Eckart Altenmüller*

**18** The neurophysiology of focal hand dystonia in musicians *283*
*Karin Rosenkranz*

**19** The development of focal dystonia in musicians as a consequence of maladaptive plasticity: implications for intervention *293*
*Nancy N. Byl and Alberto Priori*

**20** Music performance anxiety *309*
*Jürg Kesselring*

# Contributors

**Hermann Ackermann,**
Department of General Neurology,
Hertie Institute for Clinical Brain
Research, University of Tübingen,
Hoppe-Seyler-Strasse 3,
D-72076 Tübingen,
Germany.
E-mail: hermann.ackermann@uni-tuebingen.de

**Eckart Altenmüller,**
Institut für Musikphysiologie und
Musiker-Medizin, Hochschule für
Musik und Theater Hannover,
Hohenzollernstrasse 47,
30161 Hannover,
Germany.
E-mail: altenmueller@
hmt-hannover.de

**Andreas Baader,**
Institute of Physiology,
University of Berne,
Switzerland.

**Marc Bangert,**
Department of Neurology,
Beth Israel Deaconess Medical Center
and Harvard Medical School,
Boston MA, USA and
Institute of Music Physiology and
Musicians Medicine,
Hanover University of Music and Drama,
Hohenzollernstrasse 47, D-30161
Hanover, Germany.
E-mail: bangert@gmx.com

**Nancy N. Byl,**
Department of Physical Therapy and
Rehabilitation Science,
University of California,
San Francisco, 1318 7th Ave., Box 0736,
San Francisco, CA 94122–0736, USA.
E-mail: byl@itsa.ucsf.edu

**Sofia Dahl,** KTH,
School of Computer Science and
Communication,
Speech, Music, and Hearing
SE-100 44 Stockholm, Sweden.
E-mail: sofiadah@kth.se

**Martha Flanders,**
Department of Neuroscience,
University of Minnesota,
6–145 Jackson Hall,
321 Church Street SE,
Minneapolis, MN 55455,
USA.

**Christian Gerloff,**
Klinik and Poliklinik für Neurologie
Neurozentrum
Universitätsklinikum
Hamburg-Eppendorf
Martinistr. 52
20246 Hamburg
Germany.

**Reyna Leigh Gordon,**
Institut de Neurosciences Cognitives de la
Mediterranée (INCM), CNRS, Marseille,
France.

**Friedhelm Hummel,**
Klinik und Poliklinik für Neurologie
Neurozentrum
Universitätsklinikum
Hamburg-Eppendorf
Martinistr. 52
20246 Hamburg
Germany.
E-mail: hummel@uke.uni-hamburg.de

**Hans-Christian Jabusch,**
Institut für Musikphysiologie und Musiker-Medizin,
Hochschule für Musik und Theater Hannover, Hohenzollernstrasse 47, 30161 Hannover, Germany.
E-mail: jabusch@hmt-hannover.de

**Lutz Jäncke,**
University Zurich, Institute of Psychology, Department of Neuropsychology.
Treichlerstrasse 10,
8032 Zurich,
Switzerland.

**Thomas E. Jerde,**
Neurosciences Program,
Stanford University,
School of Medicine,
300 Pasteur Drive,
Alway Building, M105-A
Stanford, CA 94305-5121.
E-mail:tjerde@stanford.edu

**Oleg Kazennikov,**
Institute of Problems of Information Transmission,
Russian Academy of Sciences, Moscow, Russia.

**Jürg Kesselring,**
Department of Neurology, Rehabilitation Centre,
CH-7317 Valens, Switzerland.
E-mail: kesselring.klival@spin.ch

**Rumyana Kristeva,**
Department of Neurology,
Albert-Ludwigs-University, Breisacher Strasse 64,
D-79106 Freiburg,
Germany.
E-mail: kristeva@nz11.ukl.uni-freiburg.de

**Andreas C. Lehmann,**
Hochschule für Musik Würzburg,
Hofstallstrasse 6–8, D-97070 Würzburg, Germany. E-mail: ac.lehmann@hfm-wuerzburg.de

**Arto Nirkko,**
Neurology, University Hospital,
CH-3010 Bern, Switzerland.
E-mail: anirkko@insel.ch

**Caroline Palmer,**
Department of Psychology,
McGill University,
1205 Dr Penfield Ave,
Montreal H3A 1B1, Canada.
E-mail: caroline.palmer@mcgill.ca

**Alberto Priori,**
Movement Disorders and Deep Brain Stimulation Unit,
Department of Neurological Sciences,
University of Milan Medical School,
IRCCS Ospedale Maggiore Policlinco, Via Francesco Sforza 35, Milano, 20122 Italy.
E-mail: alberto.priori@unimi.it

**Amélie Racette,**
Department of Psychology,
University of Montreal,
Canada.

**Bruno H. Repp,**
Haskins Laboratories,
270 Crown Street, New Haven,
CT 06511-6695, USA.
E-mail: repp@haskins.yale.edu

**Axel Riecker,**
Department of Neurology,
University of Ulm,
Oberer Eselsberg 45,
D-89081 Ulm,
Germany.
E-mail: axel.riecker@uni-ulm.de

**Karin Rosenkranz,**
Sobell Department of Motor Neuroscience and Movement Disorders,
Institute of Neurology,
Queen Square,

London WC1N 3BG, UK.
E-mail: k.rosenkranz@ion.ucl.ac.uk

**Marco Santello,**
Department of Kinesiology,
Arizona State University, USA.

**Gottfried Schlaug,**
Department of Neurology,
Music and Neuroimaging Laboratory,
Beth Israel Deaconess Medical Center
and Harvard Medical School,
330 Brookline Avenue,
Boston, MA 02215, USA.
E-mail: gschlaug@bidmc.harvard.edu;
website, http://www.musicianbrain.com/

**Daniele Schön,**
Institut de Neurosciences Cognitives
de la Mediterranée (INCM),
CNRS, Marseille, France.
E-mail: daniele.schon@incon.cnrs-mrs.fr

**John F. Soechting,**
Department of Neuroscience,
University of Minnesota,
6-145 Jackson Hall,
321 Church Street SE,
Minneapolis, MN 55455,
USA.
E-mail: soech001@umn.edu

**Mario Wiesendanger,**
Institut de Physiologie,
Département de Médicine,
Rue du Musée 5,
CH-1700 Fribourg,
Switzerland.
E-mail: neuro@freesurf.ch
(also at Department of Neurology,
University of Berne,
Switzerland).

**Dirk Wildgruber,**
Department of Psychiatry and
Psychotherapy,
University of Tübingen,
Osianderstrasse 24,
D-72076 Tübingen,
Germany.
E-mail: dirk.wildgruber@med.
uni-tuebingen.de

Part1

# History

Chapter 1

# Historical increases in expert music performance skills: optimizing instruments, playing techniques, and training*

Andreas C. Lehmann

## Introduction

Throughout history, some human achievements have gone beyond what contemporaries considered possible at the time and have changed our cultural heritage forever, often in a revolutionary fashion. Adult soloists and especially child prodigies surprise—'*surprenieren*' as the term was in Mozart's time—their audiences by their musical wizardry. There is no doubt that musical performance features among the most comprehensive cultural practices. The seemingly effortless display of star pianists, singers, violinists, rock guitarists, and percussionists cloaks the fact that the skills involved rank among the most exacting that the human brain and body is capable of.

Experts in all domains have to acquire certain skills; they need incentives to acquire those and subsequently make creative contributions to the domain. For example, the desire to overcome previous performance limitations or to compete successfully with other musicians will motivate individuals to search for new solutions to technical problems. In the quest for these solutions, they may take advantage of changing circumstances that offer possibilities which were hitherto unknown. We can assume that a field such as music tends to stimulate ever higher and more complex levels of performance as long as there are incentives for performers to be drawn into the business. Incentives to choose a domain such as music are, as Csikszentmihalyi (1996) points out, strongly related to the possibility of earning a living.

* This chapter is based in part on Lehmann A.C. and Ericsson K.A. (1998). The historical development of domains of expertise: performance standards and innovations in music. In: A. Steptoe (ed.) *Genius and the mind: studies of creativity and temperament in the historical record*, pp. 64–97. Oxford University Press, Oxford.

In the twentieth century many philosophers and scientists discussed the issue of progress in Western civilization (e.g. Nisbet 1980). Although everyone agrees that scientific and technological advances have led to higher productivity and affluence, there is considerable controversy about whether quality of life has improved as a result of these changes. Likewise, evaluative judgements about what type of performance or what type of music is more enjoyable to listeners in a given historical period cannot be made, as the aesthetic criteria are likely to differ across periods. It is therefore likely that technical artistry (virtuosity) is unrelated to the aesthetic enjoyment of a piece of music. Still, the ongoing musicological discussion attests to the problematic status of technical prowess in music, and recent physiological insights give rise to some concerns (see elsewhere in this volume).

In this chapter we propose that the skills of high-level performing musicians have gradually improved over the last three centuries—at least for instruments with possibilities for solos and for those in great demand. This is a result of the efforts of individual musicians to surpass their contemporaries. We first assess changes in public music performance across historical periods and contemplate the problem of objectively assessing the performance skills for past times. Aspects of performance are then identified that allow us to study musical performance objectively, irrespective of style or period. Another factor concerns the emergence of instrumental music for individual performers or soloists, which has allowed audiences to compare the performances of different artists. The next section discusses the increases in preparation and specialization over time that have led to the achievement of higher levels of performance. We then present the strong empirical evidence for our claim that the skills required of musicians have increased. In addition, we describe other mechanisms that allow musicians to excel, such as innovative playing techniques and improvements made to musical instruments. We go on to discuss the assimilation of these changes into a body of shared knowledge which has allowed later generations of performers to acquire superior techniques and thus surpass their predecessors. The final section summarizes and integrates our findings into a model of competition and increase of performance skills.

## How can we assess skills retrospectively?

Our quantifying view of skill is new, but it is gaining increased attention in musicology because scholars are traditionally more concerned with the structure of composed music and the composers themselves than with the performers. On the contrary, the notion of increase in skill is obvious in sports, another domain in which motor performance is important. Achievements here have always been objectively recorded in terms of times, distances, or scores. New records have

naturally been accepted as temporary. An example from competitive swimming should make this obvious (see Fig. 1.1). When, in 1924, Johnny Weissmuller a.k.a. "Tarzan", was the first to swim the 100 m freestyle in less than 1 min, the media were frantic, and he earned himself a role in the first Tarzan movies. At the time women were far from reaching similar times. Forty years later, Dawn Fraser broke the 1-min time also, creating great excitement. Today, 1 min is a good amateur time, as a casual internet search will disclose. 'Higher' and 'faster' are clear objectives in sports, whereas in music the situation is much more complicated.

Since we lack sound recordings of musicians from previous centuries, the level of skilled performance of earlier musicians has to be inferred from other sources. Our analysis focuses on the piano and the time period over which it was developed. Similar analyses for more recent instruments (e.g. the saxophone in classical music, double bass in jazz music) where evidence from recordings can be used are currently under way and appear to confirm the results for the piano. For the piano we can assess the approximate levels of skill required for performance. Since historically most music was performed in public, we think it safe to assume that past musicians were able at least to play what was written down more or less accurately. Unfortunately, this type of inference is only appropriate for certain periods and musical cultures. In other periods, performers added extra-notational features such as ornaments, and the complexity and skill demands of such pieces are clearly underestimated by the notes that appear on the manuscript or printed page. Furthermore, published music may have been composed for execution by contemporary amateurs and may, therefore, hardly reflect the skills of professional musicians (see Küster (1993) for an insightful discussion). Although objective data about the expressive playing skills and artistic interpretations of past musicians are virtually impossible to obtain, the capacity of

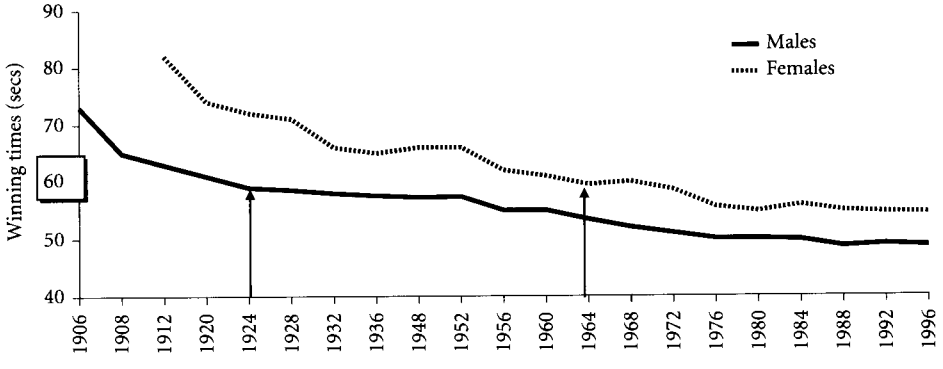



**Fig. 1.1** Decreasing swimming times for the 100 m freestyle during the twentieth century (1924, Johnny Weissmuller; 1964, Dawn Fraser). The data show Oympic gold-medal-winning times.

instruments at any given historical time gives clues to the skills required for effective performance.

## Reasons for historical increases in expert music performance over time

Despite numerous historical changes in performance conditions, the musician's job of performing music for the entertainment of the public has remained the same well-defined task. The most extreme situation is that of the solo performance. Although most of our examples here relate to traditional Western art music, it is plausible to assume that our model could also be modified to apply to jazz, popular music, or music of non-Western cultures.

The organization of society during a particular historical period will dictate in a rather unpredictable fashion how, where, and when music is made. The performance conditions for instrumental music have therefore undergone significant changes throughout history (e.g. Seaton 1991). At the risk of simplification, we can say that by gaining independence from singing, instrumental music has acknowledged its unique characteristics. Also, classical musicians have come more and more to play notated music that is not necessarily their own, relying less on their improvisational skills. Today, hardly any classically trained pianist improvises in public as did Ludwig van Beethoven or Franz Liszt. Expert musicians have to adapt to the performance conditions and task demands they encounter: it is within these limits that they strive to improve their performance.

Today, the situations that stimulate public performances of the greatest technical sophistication and involve the highest personal motivation are probably international competitions for pianists and violinists. Winning a prestigious competition is necessary for a career as an international soloist (Sosniak 1985), so these competitions are probably the best available objective measure of outstanding music performance today.

The biographies of famous performers attest to the fact that today's elite musicians spend months and years preparing and perfecting their performance of few pieces before they are presented at a competition or concert. In contrast, the preparation for a public performance was generally shorter in earlier times. Contemporary records indicate that many pieces were more or less sight-read by orchestral musicians: Beethoven's orchestral works, which were extremely difficult for the time, were premiered after only a few rehearsals by orchestras that included amateurs (Cook 1993; McVeigh 1993). Keyboard and other performers often relied on their improvisational skills when working from short-hand notation, most obviously in the seventeenth and eighteenth centuries

when the system of thorough-bass (harmonies being indicated by figures) was commonplace (Bailey 1992).

Because sight-reading or *ad hoc* improvisations are carried out under severe time constraints, performers may not be able to take full advantage of their technical and artistic prowess. By allowing more time for preparing an improvisation or rehearsing and memorizing notated music, a performance of higher quality with fewer mistakes, faster speeds, and greater technical difficulty becomes possible. Although there may be individual differences in the biomechanical dispositions and correlates of the many components involved in skilled performance, optimizing motor performance still requires excessive training (Kopiez, 2004). The axiom that public music-making can be improved through prior preparation is important because the audience cannot know how much preparation preceded a performance. A related point is the reduction of variability through preparation. Present-day audiences expect musicians to entertain them with performances of reproducible quality (maybe close to the CD recording available at the concert). Informal conversations with artists confirm this heightened expectation. Of course, we do not know what expectations audiences had in former times. However, reports dating from the early twentieth century about famous musicians such as Anton Rubinstein, Alfred Cortot, and Fritz Kreisler suggest that large day-to-day variability in performance was more acceptable than it is now (Dubal 1990; Elder 1993).

Today's performers seek out the best musical instruments that they can afford. Although to an untrained eye some instruments, such as the violin, look very similar to instruments used centuries ago, they have actually changed considerably (see, for example, Küster (1993) for examples regarding the horn, piano, and violin). Virtually all of the changes have involved increases in the instruments' capacity for changes in loudness (dynamics) and have tended to give the musician more flexibility and control when playing technically difficult passages. Over the last three centuries, the violin has been fitted with a chin and shoulder rest, a slightly longer neck, metal strings, and fine-tuners (Dilworth 1992). The extended possibilities of instruments, which will be documented later in this chapter for the piano, have led to more difficult playing techniques being possible.

There are at least two other significant differences between modern players and musicians of the past, namely specialization and training. During the last three centuries, the role of classical musicians has changed to allow them more time to develop their instrumental skills at the expense of other music-related skills they used to possess. Virtually all of today's elite performers specialize on a single instrument and perform a small number of rehearsed music pieces

(repertoire) for that instrument. In contrast, musicians in earlier historical periods usually played several different instruments (for example Wolfgang Amadeus Mozart performed his piano and violin concertos proficiently), they composed, and they taught. Furthermore, today's elite performers start their music training at an early age, often between the ages of 4 and 5. Highly experienced teachers design training programmes to allow these students to make maximum progress from the start (Sosniak 1985; Ericsson *et al.* 1993). In contrast, many musicians of earlier times acquired their initial music skills as apprentices at considerably older ages. Antonio Salieri, Mozart's contemporary, started in his teens, while Bach's students in Leipzig gained access to instrumental music instruction only after the age of 12 (Wolff *et al.* 1983). A recent analysis comparing Mozart to a sample of lesser known composers of the same age showed that the comparison group started training (not necessarily instrumental but vocal) at 6.8 years of age (Lehmann *et al.* 2002). But who had access to high-quality training? Explicit instruction and formal practice activities may have been more scarce than today. However, when appropriate instructional materials became available, scholars have noted an associated surprisingly rapid increase in level of performance (e.g. Woodfield (1984) for the viol and Baldauf-Berdes (1993) for singing).

## Musical performance in a cultural context

The music that audiences have wished to hear and the physical location of performances have varied throughout history in response to changing philosophical ideas and social conditions (see Parakilas (2002) for examples regarding piano music, Salmen (1988) for the concert and Seaton (1991) for a general description). One of the foremost functions of music is the entertainment of an audience, and professional musicians have fulfilled this function differently in different times. For instance, in the Baroque period, music conveyed static emotional states (*Affekte*) using idiomatic ornamentations, figures, and phrases. Since the early Classical period, there has been an increase in the dramatic expressiveness of music that is paralleled by developments both in other arts and in philosophy. Improvements in instrument design have supported these trends. For example, in contrast to the harpsichord, the development of Cristofori's fortepiano allowed for variability in loudness and permitted increased expressivity (Komlos 1995). The need to amplify the soft guitar in twentieth-century jazz in larger ensembles that included brass instruments led to the invention of the electric guitar (and its virtuoso player). Thus, public music performance has changed as a function of preferred styles and playing conditions, and successful musicians have adapted to the demands of their times.

The majority of musicians accompany others most of the time, but more exposure is gained in solos. Over time, many instruments have developed a solo function (a good recent example is the viola). Every musical culture has its own way of selecting and allowing certain individuals to march to the forefront and display themselves. Soloists sometimes show off their skills, and on many occasions virtuosi have been accused of devaluing their performances for the purpose of self-enhancement (Baillot 1835/1991; Rowland-Jones, 1992). We will not delve into a musicological or aesthetic discussion on the pros and cons of the virtuosos—the fact is that they exist.

## Competition among musicians or 'who is the best?'

Whenever employers or audiences are provided with choices among performing musicians, they will select the one who is viewed as better or has the better reputation. Therefore, in order to make a living, musicians have to compete with each other to attract audiences, fame, and patronage. The situation is even more dramatic in rock and popular music than it is in classical music. As a result, local musicians may not always be the first choice, and musicians from distant cultural centres may be favoured (think of the globalization and dominance of English-language popular music). In former times, the opportunities for such comparisons arose when musicians began to travel with nobility, or when musicians migrated in search of work. Such was the case in the fifteenth century when German musicians were sought after in Italy because of their superior skills (Polk 1994). Once they had taught their new skills to local musicians, their stranglehold on employment decreased and the 'importation' of musicians was no longer the norm (Polk 1991). Conversely, aspiring musicians, including the young Mozart, travelled to places with advanced musical cultures, such as Italy, to work with eminent musicians and learn new styles.

Musicians of different periods in history are known to have tried to outperform each other in veritable 'playoffs' (i.e. like Eurovision song contests but actually involving skills). The evidence is anecdotal and therefore not always reliable, but its mere existence is certainly telling. Here are some 'piano finals': Liszt versus Thalberg (Walker 1990), Mozart versus Clementi (Komlos 1995), Handel versus Scarlatti (Blume 1956), Beethoven versus Gelinek, and Bach versus Marchand (questionable; Breig 1998). Even today, jazz players continue to engage in friendly 'cutting contests' (Berliner 1994). Thus, despite the fact that there are aspects of musical performance difficult to evaluate objectively, playing in public is partly an exhibition of skills in which musicians compete for their audiences and for the respect of their peers.

At all times there have been members of the audience who could judge the performance. Today, through widespread musical training and the availability of

recordings, many concert-goers have attained a considerable level of (listening and) playing skills themselves, especially on solo instruments such as the piano and the violin. Along with their appreciation for the highest levels of music performance audiences have gained a greater ability to detect imperfections. This effect is increased through a strong canonization of the repertoire. Listening to recordings by the world's foremost musicians on high-fidelity stereo systems creates high expectations for the concert soloist—a problem that earlier musicians certainly encountered to a lesser degree.

## Specialization and deliberate practice

Although specialization is a powerful instrument for increasing performance, it can also narrow the range of possible performances, thus changing the domain. Using keyboard players as an example, we will illustrate this process of professional specialization, which, incidentally, can also be observed in other domains of expertise such as science and sport (Ericsson and Charness 1994; Ericsson 1996). Beginning in the Middle Ages and continuing during the Baroque period, most musicians were relatively diversified. J. S. Bach, for example, was a composer, performer, music director, improviser, music teacher, and Latin teacher. As a rule, instrumentalists of the past were able to perform on more than one instrument in public, and keyboard players customarily played all possible instruments including the organ, harpsichord, and clavichord (e.g. Bowers 1992). Initially, music for keyboards was interchangeable, but during the seventeenth century composers started taking advantage of the specific characteristics of each individual instrument (Komlos 1995). During the Classical period, musicians such as the pianist Muzio Clementi started to specialize in public performances on a single instrument. As musicians specializing on single instruments, their careers also became more focused. By the beginning of the twentieth century, most pianists had completely specialized in the performance of the standard repertoire. Others, such as Carl Czerny, concentrated exclusively on teaching. With an emphasis on memorized performance, sight-reading in public declined, and opportunities for playing less rehearsed music became restricted to pianists who specialized in accompanying (coaching, chamber music, ballet). Some twentieth-century performers even specialized in one particular era or a limited number of composers, or became known for their interpretation of individual works (e.g. Vladimir Horowitz for Rachmaninov's Third Piano Concerto). Similar trends toward specialization can be observed in performers of orchestral instruments, such as the flute and violin (Bowers 1992; McVeigh 1993).

Why does specialization work? In general, training is limited by the starting age for a particular historical period and by the age at which the individual enters

the profession. Furthermore, higher levels of performance in several domains have been related to larger amounts of lifetime practice (see Ericsson and Lehmann (1996) for a review). Therefore, when future performers devote all their training exclusively to those aspects of their task that are most desired (marketable?), they should attain a higher final level of skill; alas, often at the expense of other skills. Not only are the years of training limited but also the amount of concentrated practice that can be sustained over extended periods of time, which appears to be limited to 4 or 5 hours a day for adults. As a result, today, by the age of 20, the most promising musicians have engaged in over 10,000 hours of deliberate practice (Ericsson *et al.* 1993). Attempts to maintain practice beyond this optimal limit may lead to overuse injuries and burnout, making it less likely that the musician will reach and maintain the required level. Although it might be tempting to think that earlier musicians simply worked longer and harder or had fewer distractions than today's musicians, the explanation above should dispense with this thought: today's music students practice at the limit of what is possible (often beyond!).

## Empirical evidence for changes in performance expertise

At this point it is necessary to present empirical evidence for the increase in performance expertise in piano music (see Lehmann and Ericsson 1998*a*). Music teachers are frequently faced with the problem of assessing the difficulty of particular pieces they assign to their students. Therefore, experienced educators have devised repertoire lists that contain complexity ratings for specific pieces (e.g. the Syllabi of the Toronto Conservatory or the Associated Board Examinations) and that extensively cover the common repertoire for the instrument.

Depending on the intended starting age and duration of instruction, the different educational plans contain between 10 and 15 grades. The criteria applied by music experts for grading music are not explicitly stated and are most likely based on their experience as performers and teachers. Technical problems, such as speed, seem to be important criteria. For example, the demands on playing speed for scales increase from less than 150 notes per minute at level 1 to over 500 notes per minute at level 11 (Piano Syllabus, 1986). Interpretational complexity may also play a role, but there is less indication as to how this type of complexity is assessed. In a recent diary study with an advanced pianist who was preparing for a public performance, we found that complexity ratings for a given set of pieces correlated 0.71 ($P < 0.05$) with the accumulated practice time for these same pieces (Lehmann and Ericsson 1998*b*). Since the correlation between lists is also large ($r > 0.8$), one is tempted to view them as a useful source of data on the complexity of pieces. By choosing a modern perception of performance

complexity as the standard for cross-historical comparisons, we cannot capture additional technical difficulties that stemmed from performance practices in the past. However, we can analyse the difficulty that pieces from different times would pose for today's performers and determine when in history the more difficult techniques (by today's standards) were invented and introduced. We will now exemplify our procedure using the piano.

We assembled complexity ratings by Clementi, Haydn, Mozart, Beethoven, and Schubert for sonatas known to most amateur piano players, spanning roughly 100 years starting from 1750. We found a significant correlation between compositional dates and complexity ratings (from Pierce and Fuszek 1987) of $r = 0.68$ ($P < 0.01$) which shows that piano sonatas became more technically difficult over time. Further validation for our statistical results comes from the comments of the composers themselves. Joseph Haydn, for instance, realized that his piano sonatas were becoming increasingly difficult as he grew older (Rosenblum 1988). Interestingly, analysing only Haydn's sonatas confirms this statement from today's perspective ($r = 0.58$, $P < 0.01$, shown in Fig. 1.2). Similar analysis could probably be done with other genres and instruments. Our present results allow us to infer that the pianistic skills necessary to perform sonatas and other pieces must have increased.

Based on the above-mentioned educational plans, it is also possible to infer the difficulty of particular playing techniques by the grade level at which they are introduced. For a number of these techniques there is an approximate historical date at which they were invented and included in compositions. Some techniques are, of course, linked to the development of the instrument, while others emerged after the instrument had reached its modern form. Not surprisingly, dynamic change, which was impossible on the harpsichord, is taught to piano students from the very beginning of instruction. The complexity of the left hand

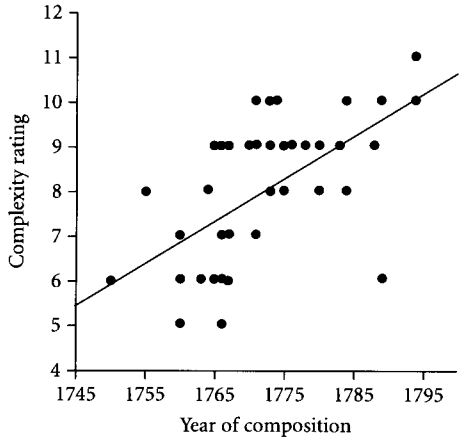



**Fig. 1.2** Increase in technical complexity (from Pierce and Fuszek 1987) in the piano sonatas by Joseph Haydn.

of Mozart's (1756–1792) works has increased compared with previous works by Haydn (1732–1809) (Wier 1940). And starting with Czerny's (1791–1857) studies for the left hand a whole range of extremely difficult left-hand repertoire has been composed (e.g. by Ravel (1875–1937)), some of which was specially dedicated to the pianist Paul Wittgenstein who returned from World War I without his right arm. We cannot rule out that in some cases of practicing the non-dominant hand (oftentimes the left hand) extensively, and in combination with a disposition for decreased lateralisation, unexpectedly large increases in performance may be possible (Kopiez and Galley, in press). Alas, not much is known about the handedness and lateralisation of famous left-hand injured performers (e.g., Scrijabin, Brahms, most recently Rebstein). The standard fingering for double octave scales, codified by Johann Cramer (1771–1858), appears at level 7. Scales in double thirds and sixths, which appear at level 12 or higher, became necessary for the works of the nineteenth-century virtuosi. The modern use of the damper pedal was introduced by John Field around 1800, and very complex pedalling is necessary to perform the music of our times. In summary, the playing techniques that require more training for modern music students were introduced later in history. This finding is one of the strongest pieces of evidence for our claim that the technical mastery of musicians has increased over time.

## Why has complexity increased over time?

Complexity has increased over time because musical instruments have been optimized, new and innovative playing techniques have been developed by leading performers, and music training has been made more widely accessible and efficient (see Lehmann and Ericsson (1998*a*) for additional details). Essentially, at this point research into expert performance intersects with research into creativity and motivation. We assume that elite experts want to contribute to their domain in creative ways by perfecting existing art or changing the domain (see Gardner 1997), and these elite experts can be found among the instrument makers as well as among the performers, composers, and teachers.

Increases in skill occur when individual elite performers go beyond the existing strictures and invent new techniques that make the most efficient use of instrumental capabilities. The success of instrument makers depends on their ability to produce superior musical instruments and to win endorsement of their instruments by famous performers. Admittedly, changes in instrument design involve trade-offs, as certain sound effects, nuances, and playing techniques are replaced by new ones (Baines 1992). The number of keys on the piano, for example, has nearly doubled since its invention. In the course of its development, the piano increased in dynamic (loudness) range, obtained the possibility of extremely fast repetition of notes, and gained many additional notes in the high

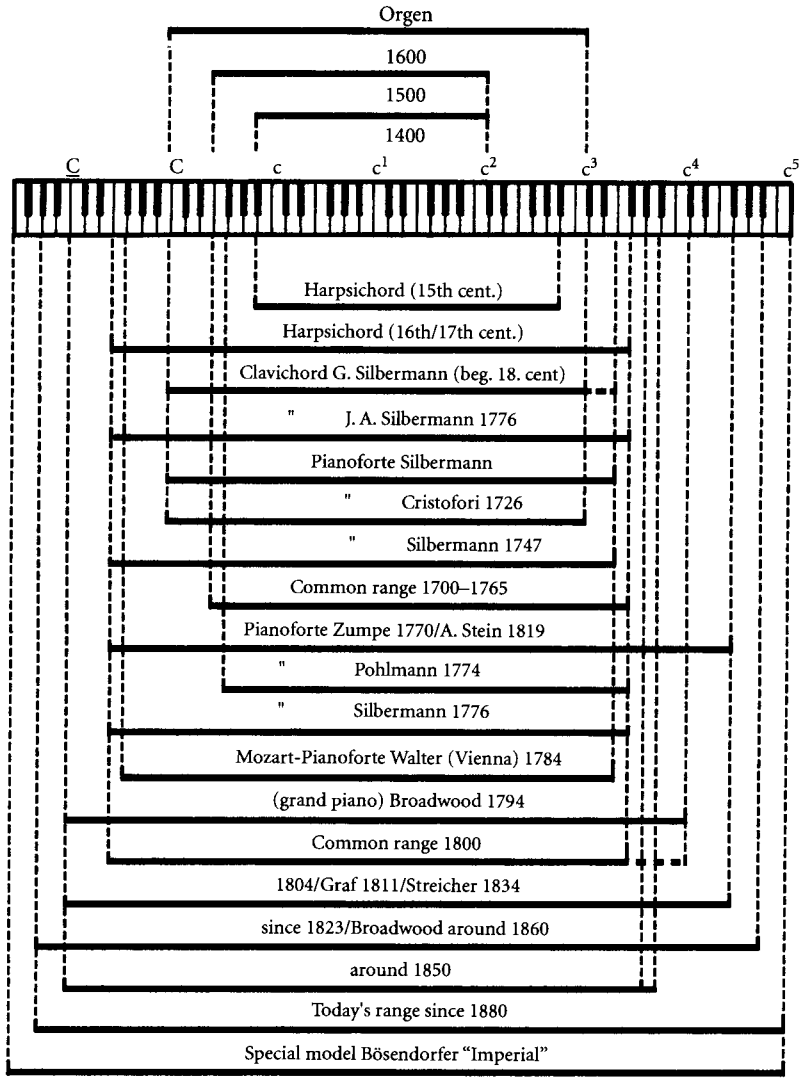



**Fig. 1.3** Tonal ranges for different pianos (keyboard instruments) have increased over time from 4 octaves (48 notes) to 7.5 octaves (95 notes). (From Philip 2003, p. 764). Adapted with kind permission from G. Phillip (2003), Klavierspiel und Improvisation, Altenburg: Kamprad, p. 274 (Fig. 130).

and low register (see Marcuse (1964) for details; Fig. 1.3). An increase in tonal range over time is frequently found in the history of instruments. Very often performers and composers collaborate on such improvements. Successful mastery of the increased tonal range required the playing of larger horizontal leaps, and the dexterity to cover a broad span with scales and patterns (arpeggios) despite sometimes awkward hand positions or even hand crossings, while at the same time keeping control over shades of loudness.

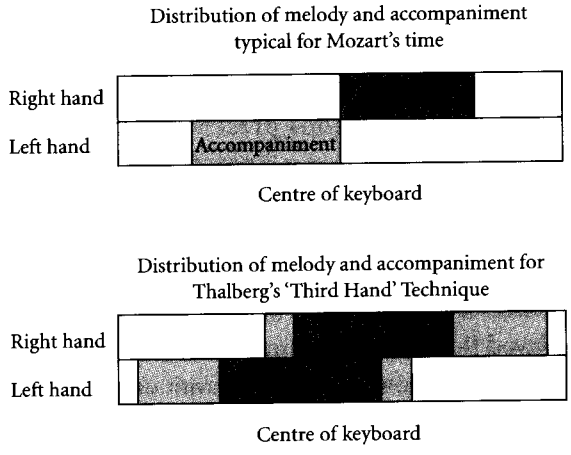



**Fig. 1.4** Mapping of hands and dynamics to the keyboard layout for Mozart's time (top) and for Thalberg's 'third hand' technique (bottom). The horizontal range of the keyboard is shown schematically from above. The centre of the keyboard suggests middle C.

During many periods of musical history, virtuoso players have made creative contributions by incorporating their personal playing experiences and styles into their written compositions. Wolfgang Amadeus Mozart was known for his agile left hand and is said to have increased the complexity of the left hand in his compositions. Frederic Chopin's style took into account ergonomic factors related to the location of the black and white keys, and his music is consequently full of black keys, or accidentals. One specific technique, sometimes referred to as 'third hand' impression, was developed by the pianist Sigismund Thalberg in the 1830s (Hitchcock 1971). Instead of playing the melody with one hand, and the accompaniment with the other, the melody notes are distributed between hands in the middle of the keyboard, while the accompaniment is played by both hands to the left and right of the melody (see Fig. 1.4). This advanced technique destroys the classical mapping of hands onto the keyboard (right hand = melody, left hand = accompaniment). Maintaining a dynamic difference between the melody (louder) and the accompaniment (softer) on one keyboard, often within one hand, imposes great skill requirements on the performer.

New playing techniques and possibilities also emerge from the interaction between performers and composers. Composers have at all times sought to incorporate the most advanced contemporary skills into their works and have consulted with elite musicians to identify the full range of possibilities of particular instruments. The following examples derive from several centuries: Vivaldi is known to have tailored his compositions to the particular strengths of his soloists in eighteenth-century Venice (Baldauf-Berdes 1993); Bach consulted with some of the finest flute players of his time (Powell and Lasocki 1995); Mozart worked with the oboe virtuoso Ramm and the clarinetist Anton Stadler (Haynes 1992); Beethoven consulted the double bass virtuoso Dragonetti (Blume 1954, vol. 3,

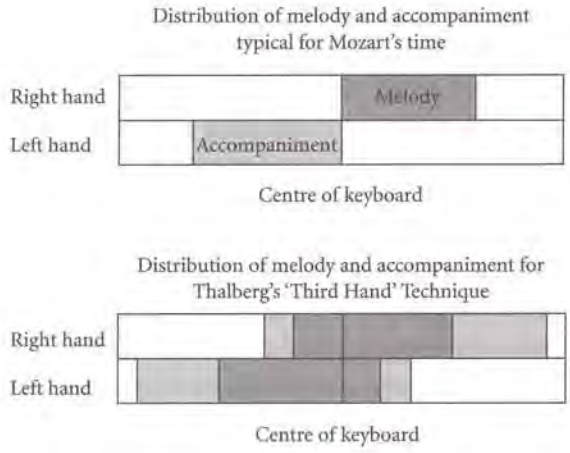



**Fig. 1.4** Mapping of hands and dynamics to the keyboard layout for Mozart's time (top) and for Thalberg's 'third hand' technique (bottom). The horizontal range of the keyboard is shown schematically from above. The centre of the keyboard suggests middle C.

During many periods of musical history, virtuoso players have made creative contributions by incorporating their personal playing experiences and styles into their written compositions. Wolfgang Amadeus Mozart was known for his agile left hand and is said to have increased the complexity of the left hand in his compositions. Frederic Chopin's style took into account ergonomic factors related to the location of the black and white keys, and his music is consequently full of black keys, or accidentals. One specific technique, sometimes referred to as 'third hand' impression, was developed by the pianist Sigismund Thalberg in the 1830s (Hitchcock 1971). Instead of playing the melody with one hand, and the accompaniment with the other, the melody notes are distributed between hands in the middle of the keyboard, while the accompaniment is played by both hands to the left and right of the melody (see Fig. 1.4). This advanced technique destroys the classical mapping of hands onto the keyboard (right hand = melody, left hand = accompaniment). Maintaining a dynamic difference between the melody (louder) and the accompaniment (softer) on one keyboard, often within one hand, imposes great skill requirements on the performer.

New playing techniques and possibilities also emerge from the interaction between performers and composers. Composers have at all times sought to incorporate the most advanced contemporary skills into their works and have consulted with elite musicians to identify the full range of possibilities of particular instruments. The following examples derive from several centuries: Vivaldi is known to have tailored his compositions to the particular strengths of his soloists in eighteenth-century Venice (Baldauf-Berdes 1993); Bach consulted with some of the finest flute players of his time (Powell and Lasocki 1995); Mozart worked with the oboe virtuoso Ramm and the clarinetist Anton Stadler (Haynes 1992); Beethoven consulted the double bass virtuoso Dragonetti (Blume 1954, vol. 3,

p. 739), and from the piano virtuoso Steibelt he probably borrowed the special pianistic effect of rapid alteration between two notes (Stillings 1972); and Haydn worked with the cellist Kraft (Loesch 1992). Because composers work with the leading performers, the resulting compositions present difficulties during rehearsal and performance for lesser musicians (e.g. known examples arise from Richard Wagner's scores). Today these orchestral parts are standard fare for instrumentalists. Furthermore, there are numerous anecdotes for virtually all instruments about works that were regarded as unplayable at the time of their composition. Virtuoso performers impressed their audiences by playing exactly those works, e.g. when Liszt played Beethoven's Hammerklavier sonata.

Other innovations result from adaptations of methods from one instrument to another. For example, seventeenth-century texts suggest that expert singers and instrumentalists imitated each others' techniques and ornamentation (Brown 1976). In the fifteenth century, the organist Conrad Paumann transferred the playing of several simultaneous voices (polyphonic playing) from the keyboard to the lute (Polk 1994). Sigismund Thalberg's 'third hand' technique was inspired by the English harpist Parish-Alvars (Walker 1990), while Niccolo Paganini is known to have transferred his extensive knowledge of the guitar to the violin. Liszt in turn attempted to transfer Paganini's virtuoso violin technique to the piano. Modern bassoonists and oboists use a breathing technique (circular or continuous breathing) originally also used by players of the Australian didgeridoo. The jazz artist Bobby McFerrin vocally imitates every instrument. Transferred techniques are extremely taxing because they are likely to violate some standard ergonomic constraints of the instrument, thus adding to the level of difficulty.

The increased complexity of music performance at the highest levels has implications for training musicians because it forces music teachers to find ways to facilitate (accelerate) the acquisition of new techniques. Although detailed historical records of the relationship between instruction and levels of performance are virtually non-existent, child prodigies offer the possibility to look at individuals who are maximally immersed in learning and performing music. We can therefore compare the change in the rate at which piano playing prodigies have attained certain levels of skill in musical performance and find out if the acquisition of skills has accelerated over time. For this analysis, we have listed solo pieces played by child prodigies and the ages at which they performed these pieces in public. The musician had to be under age 13 at the time of performance to qualify as a child prodigy. Our sample is hardly random, but consists of some of the best known keyboard prodigies from Johann Sebastian Bach (born 1671) onwards. After finding the earliest public performance known for each individual, we assessed the difficulty level of the piece that was performed by looking it (or a comparable piece) up in the standard lists of graded repertoire mentioned

earlier. We then computed an index by dividing the number of years of study necessary to master the piece of music today by the actual number of years of study the prodigy had received prior to playing the piece in public. This fraction was then multiplied by 100. For example, if a modern-day student played Schubert's Moments Musicaux (rated level 7 by Pierce and Fuszek, 1987) after 7 years of instruction, this student would receive a score of 100 [(7/7) × 100]. If another student performed the same pieces after only 3.5 years of instruction, a score of 200 would result [(7/3.5) × 100]. Thus, the precocity of performers of the past relative to today's standards is tangible in numbers exceeding 100. In fact, there is a significant tendency for prodigies of later generations to play more difficult pieces after shorter periods of training than did earlier prodigies ($r = 0.67$, $P < 0.05$; see Fig. 1.5, filled squares). This index reflects an acceleration of acquisition of skills once training is initiated.

A second index was then computed to look at the age when certain pieces were performed. This was done by dividing the projected age of performance from modern curricula by the age at which a given prodigy actually performed the piece. Again, we find a significant correlation ($r = 0.85$, $P < 0.01$; see Fig. 1.5, open circles) demonstrating that by today's standards, more recent prodigies have been more advanced in their performance skills than prodigies from earlier times. But given that the difficulty of music in the repertoire has also increased over time, it is possible that the relationship we observed is merely due to more recent prodigies playing music that was unavailable to earlier generations. However, even after statistically controlling for the level of difficulty of the performed piece, we found that precociousness has increased reliably over time. The empirical evidence presented here is endorsed by a host of anecdotes that

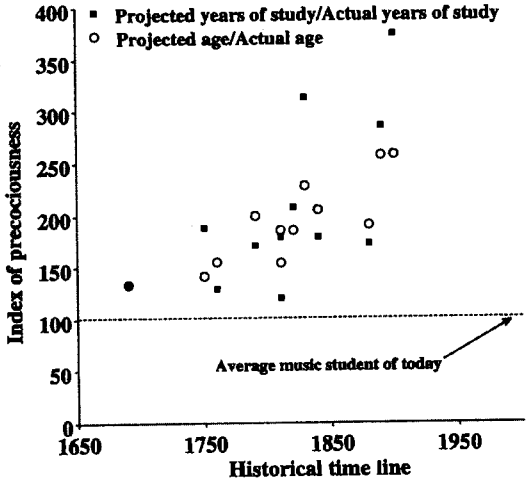



**Fig. 1.5** Evidence for accelerated skill acquistion of 12 piano prodigies (J. S. Bach, W. A. Mozart, M. A. Mozart (Nannerl), C. Czerny, F. Mendelssohn, F. Liszt, Artur Rubinstein, Anton Rubinstein, C. Saint-Saens, E. Nyiregyhazi, J. Hofmann).

point to an increase in performance skills over time (e.g. McVeigh 1993, p. 221; cf. Lehmann and Ericsson (1998*a*) for extensive references). Recent media reports on young Chinese pianists suggest that this historical trend is unbroken.

Although we have carefully searched for contradictory evidence stating declines in skill, we have only been able to locate anecdotal evidence of a decline in florid singing skills (Henderson 1921), in improvisation (Donington 1963), and in sight-reading skills. But we have not found any examples contradicting the increase in technical proficiency of the performance of rehearsed music.

The factors that contribute to optimized training are manifold (see Lehmann and Ericsson (1998*a*) for details) and include the installation of systematic instruction which is often assisted by didactic materials (method books; e.g. Türk 1982). These manuals are generally written by the foremost music educators of their time, but once they are published, they can be used even by mediocre teachers to train large numbers of music students. Moreover, better performers receive lessons with fewer and more qualified teachers than less good performers (see Sosniak (1985) and Howe and Sloboda (1991) for modern-day examples), usually preferring individual over group instruction (Lehmann *et al.* (2002) for Mozart's time). It also appears that for young performers, external monitoring in the form of supervised practice is necessary (see Lehmann 1997, for historical prodigies, Davidson *et al.* (1996) for modern-day music students and Frederickson (1995) for a famous contemporary example). Ericsson *et al.* (1993) have suggested that an earlier start in training leads to higher performance at younger ages. Winning a competition or having a debut at a young age is likely to result in extra resources and access to the foremost teachers. Some modern-day teachers encourage earlier starting ages, and these are facilitated by instrument-makers who provide child-sized instruments (even Wolfgang Amadeus Mozart is known to have used a small violin as a child). For instance, Edler-Busch (1993) suggests that improvements in flute construction in the nineteenth century allowed for a younger starting age than was previously customary, while today adding a curved head-joint may reduce starting age even further (Rainey 1985). These ergonomic adaptations also enable longer practice and playing times, especially for children.

## Summary and discussion

Many aspects of playing music in public have changed dramatically over the past several centuries. The repertoire has increased substantially through the steady accumulation of compositions, and—at least for the piano—the upper end of more recent compositions tends to be more complex and technically challenging than that of older ones. These increases have been attributed to improvements in instrumental design and to innovations and advances in performance technique.

Concurrently, the increased time spent preparing for performances, coupled with growing professional specialization, have given the performer the opportunity to master more complex pieces of music. Finally, the optimization of training has allowed young pianists to attain the high level of skills necessary for modern virtuosi. Ever increasing standards continue to determine the limits of adult performance until they too are surpassed. New standards and achievements are assimilated into education and training, allowing future musicians to match and to surpass the previously final level of high-quality performance.

At any given time in history, a small number of performers achieve the highest levels of adult skills for that era. We have looked at the performance of rehearsed music for the piano, but similar analyses could be done for any other instrument, such as the classical violin, or for improvisational skills in jazz, or for instrumental performance in rock music. Although one would have to look at different historical time spans, all would lead to similar conclusions. Using recordings it is now even possible to assess whether expert performers of primarily orally transmitted performance practices in jazz, non-Western music, or popular music have become more skilled, thereby broadening the scope of our claims (such work is currently under way).

To clarify again, this analysis is not meant to make any evaluative statements about the aesthetic achievements and importance of specific musical compositions. Regarding the expressive aspects of music performance, we follow Gerig (1974) and Ferrucio Busoni (1968) who argued that complete mastery of technical problems is a prerequisite for the production of expressive interpretations. Although beautiful music can be rendered with little virtuosity, musical performance undeniably allows for competitiveness and athleticism that has been a constant source of controversy among practitioners and aestheticians alike, an issue which is beyond the scope of this chapter.

We can speculate that the motor behind the increase in expert performance is the motivation of individuals to receive rewards and recognition (e.g. commercial success, media attention, filled concert halls). Eminent performers of their time will distinguish themselves by fully exploring the instrument's capacities, or by surpassing existing skills in creative ways. Eventually, these achievements are taught to others by the performers themselves or by specialized pedagogues. Teachers are eager to transmit new techniques because their reputations rest on the students who reach the highest levels of performance. Similarly, students will strive to display the greatest possible skills compared with their peers, so as to gain access to the best available teachers and resources. Ultimately, the implementation of new techniques through instruction leads the entire community of elite musical performers to a higher level of music performance skill. For many performance traditions, instruments, and repertoires of written music, the

historical developments have started to asymptote (like in many athletic events, see Fig. 1.1). This is understandable, since ever-decreasing returns require exponentially augmenting efforts during training, and given a highly optimized training situation there is simply not much room for further improvement. Moreover, for some instruments, certain physiological limits are starting to loom large (cf. Elbert *et al.* 1998; Münte *et al.* 2002; Kopiez 2004). Still, considering the unpredictable nature of past achievements in musical performance, it is reasonable to assume that as long as society offers appropriate incentives and elite performers and musical clowns keep striving to surpass the achievements of current and past masters, creative innovations will never cease to be made. Our guess would be that even physiological limits may stimulate even more creative solutions—in music as they would in other domains such as science and sports.

## References

**Bailey, D.** (1992). *Improvisation: its nature and practice in music.* Dacapo, New York.

**Baillot, P.** (1991). *The art of the violin* (ed. and trans. L. Goldberg). Northwestern Illinois Press, Evanston, IL. (Original published in 1835.)

**Baines, F.** (1992). Musical instruments and performers of the past. In: Paynter, T. Howell, R. Orton, and P. Seymour (eds) *Companion to contemporary musical thought*, Vol. 2, pp. 961–976. Routledge, London.

**Baldauf-Berdes, J.L.** (1993). *Women musicians of Venice: musical foundations 1525–1855.* Clarendon, Oxford.

**Berliner, P.F.** (1994). *Thinking in jazz: the infinite art of improvisation.* University of Chicago Press, Chicago, IL.

**Blume, F.** (ed.) (1949–1986). *Musik in Geschichte und Gegenwart*, Vols 1–17. Barenreiter, Kassel.

**Breig, W.** (1998). Bach und Marchand in Dresden. *Bach Jahrbuch*, 7.

**Brown, H.M.** (1976). *Embellishing 16th-century music.* Oxford University Press, Oxford.

**Busoni, F.** (1968). *Exercises and studies for the piano: Part I* (ed. F. Goebels). Breitkopf and Hartel, Wiesbaden.

**Cook, N.** (1993). *Beethoven symphony no. 9.* Cambridge University Press, Cambridge.

**Csikszentmihalyi, M.** (1996). *Creativity, flow, and the psychology of discovery and invention.* Harper Collins, New York.

**Davidson, J.W., Howe, M.J.A., Moore, D.G., and Sloboda, J.A.** (1996). The role of parental influences in the development of musical ability. *Br J Dev Psychol* **14**, 399–412.

**Dilworth, J.** (1992). The violin and bow-origins and development. In: R. Stowell (ed.) *The Cambridge companion to the violin*, pp. 1–29. Cambridge University Press, Cambridge.

**Donington, R.** (1963). *The interpretation of early music.* Faber, London.

**Dubal, D.** (1990). *The art of the piano.* Tauris, London.

**Edler-Busch, E.** (1993). Trompeten waren auch mal gerade: Zum Anfangsalter für das Querflötenspiel. *Üben und Musizieren*, **10**, 26–29.

**Elbert, T., Candia, V., Altenmüller, E., Rau, H., Sterr, A., Rockstroh, B., Pantev, C., and Taub, E.** (1998). Alteration of digital representations in somatosensory cortex in focal hand dystonia. *NeuroReport*, **9**, 3571–3575.

Elder, D. (1993). Edward Kilenyi remembers Dohnanyi. *Clavier*, **32**, 11–18.

Ericsson, K.A. (1996). The acquisition of expert performance: an introduction to some of the issues. In: K A. Ericsson (ed.) *The road to excellence: the acquisition of expert performance in the arts, sciences, sports, and games*, pp. 1–48. Erlbaum, Mahwah, NJ.

Ericsson, K.A. and Charness, N. (1994). Expert performance. Its structure and acquisition. *Am Psychol*, **49**, 725–747.

Ericsson, K.A. and Lehmann, A.C. (1996). Expert and exceptional performance: evidence for maximal adaptations to task constraints. *Ann Rev Psychol*, **47**, 273–305.

Ericsson, K.A., Krampe, R.T., and Tesch-Römer, C. (1993). The role of deliberate practise in the acquisition of expert performance. *Psychol Rev*, **100**, 363–406.

Frederickson, D. (1995). Van Cliburn remembers his remarkable mother. *Clavier*, **35**, 6–9.

Gardner, H. (1997). *Extraordinary minds*. Basic Books, New York.

Gerig, R.R. (1974). *Famous pianists and their technique*. Luce, Washington, DC.

Haynes, B. (1992). Mozart and the oboe. *Early Music*, **20**, 43–62.

Henderson, W.J. (1921). *Early history of singing*. Longmans Green, New York.

Hitchcock, D. (1971). Sigismund Thalberg: an evaluation of the famous composer-pianist, and the 100th anniversary of his death. *Piano Q*, **20**(77), 12–16.

Howe, M.A. and Sloboda, J.A. (1991). Young musicians' accounts of significant influences in their early lives. 1. The family and the musical background. *Br J Music Educ*, **8**, 39–52.

Komlos, K. (1995). *Fortepianos and their music: Germany, Austria, England, 1760–1800*. Clarendon Press, Oxford.

Kopiez, R. (2004). Virtuosität als Ergebnis psychomotorischer Optimierung [Virtuosity as a result of psychomotor optimisation]. In: H. v. Loesch, U. Mahlert, and P. Rummenhöller (eds) *Virtuosität*. Schott, Mainz. 205–231.

Kopiez, R., Galley, N., and Lee, J. The advantage of a decreasing right-hand superiority: The influence of laterality on a selected musical skill (sight reading achievement). *Neuropsychologia* (doi: 10.1016/j.*neuropsychologia*.2005.10.023).

Küster, W. (1993). *Das Konzert: Form und Forum der Virtuosität*. Bärenreiter, Kassel.

Lehmann, A.C. (1997). Acquisition of expertise in music: efficiency of deliberate practice as a moderating variable in accounting for sub-expert performance. In: J. Sloboda and I. Deliege (eds) *Perception and cognition of music*, pp. 165–191. Taylor and Francis, London.

Lehmann, A.C. and Ericsson, K.A. (1998*a*). The historical development of domains of expertise: Performance standards and innovations in music. In: A. Steptoe (ed.) *Genius and the mind: studies of creativity and temperament in the historical record*, pp. 64–97. Oxford University Press, Oxford.

Lehmann, A.C. and Ericsson, K.A. (1998*b*). Preparation of a public piano performance: the relation between practice and performance. *Musicae Scientiae*, **2**, 69–94.

Lehmann, A.C., Ericsson, K.A., and Hetzer, J. (2002). How different was Mozart's music education and training? A historical analysis comparing the music development of Mozart to that of his contemporaries. In: K. Stevens, D. Burnham, G. McPherson, E. Schubert, and J. Renwick (eds) *Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney, 2002*, pp. 426–429. Causal Productions, Sydney (CD-ROM).

Loesch, H.V. (1992). *Das Cellokonzert von Beethoven bis Ligeti*. Lang, Frankfurt.

Marcuse, S. (1964). *Musical instruments; a comprehensive dictionary*. Doubleday, Garden City, NY.

McVeigh, S. (1993). *Concert life in London from Mozart to Haydn.* Cambridge University Press, Cambridge.

Münte, T.F., Altenmüller, E., and Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nat Rev Neurosci*, **3**, 473–478.

Nisbet, R. (1980). *History of the idea of progress.* Basic Books, New York.

Parakilas, J. (ed.) (2002). *Piano roles. A new history of the piano.* Yale University Press, New Haven, CT.

Phillip, G. (2003). *Klavierspiel und Improvisation.* Kamprad,

Piano Syllabus (1986). *Royal Conservatory of Music Toronto: Piano Syllabus.* Royal Conservatory, Toronto.

Pierce, R. and Fuszek, R. (1987). *The P-F guide: a conceptual approach to piano instruction.* R. Pierce Music, West Covina, CA.

Polk, K. (1991). Patronage and innovation in instrumental music in the 15th century. *Historical Brass Soc J*, **3**,151–178.

Polk, K. (1994). Innovation in instrumental music 1450–1510: the role of German performers within European culture. In: J. Kmetz (ed.) *Music in the German renaissance*, pp. 202–214. Cambridge University Press, Cambridge.

Powell, A. and Lasocki, D. (1995). Bach and the flute: the players, the instruments, the music. *Early Music*, **23**, 9–30.

Rainey, J.E. (1985). *The flute manual.* University Press of America, Lanham, MD.

Rosenblum, S. (1988). *Performance practices in classic piano music.* Indiana University Press, Bloomington, IN.

Rowland-Jones, A. (1992). *Playing recorder sonatas: interpretation and technique.* Clarendon, Oxford.

Salmen, W. (1988). *Das Konzert: eine Kulturgeschichte.* Beck, Munich.

Seaton, D. (1991). *Ideas and styles in the Western musical tradition.* Mayfield, Mountain View, CA.

Sloboda, J.A., Davidson, J.W., Howe, M.J.A., and Moore, D. (1996). The role of practise in the development of expert musical performance. *Br J Psychol*, **87**, 287–309.

Sosniak, L.A. (1985). Learning to be a concert pianist. In: B.S. Bloom (ed.) *Developing talent in young people*, pp. 19–67. Ballantine, New York.

Stillings, F.S. (1972). *The sonatas of Beethoven as he played and taught them.* Music Teachers National Association, Cincinnati, OH.

Türk, D.G. (1982). *School of clavier playing* (trans. R.H. Haggh). University of Nebraska Press, Lincoln, NE. (Originally published in 1789.)

Walker, A. (1990). *Franz Liszt*, Vol. 1. Knopf, New York.

Wier, A.E. (1940). *The piano: its history, makers, players, and music.* Longmans and Green, London.

Wolff, C., Emery, W., Jones, R., Helm, E., Warburton, E., and Derr, E. (1983). *The New Grove: Bach family.* Norton, New York.

Woodfield, I. (1984). *The early history of the viol.* Cambridge University Press, Cambridge.

Part 2

# Psychology

Chapter 2

# From cognition to action

Lutz Jäncke

Musical abilities are traditionally linked to various cognitive fields like attention, memory, or visuospatial functions. In this context it is often argued that there is a kind of positive transfer between musical performance and the above-mentioned cognitions. In this chapter recent studies exploring the relationship between general cognitions and musical skills in musicians and non-musicians are reviewed. It will be shown that there are intimate links between musical ability and various cognitions including perception, sight-reading, memory, action planning and organizing, attention, and visuospatial functions. We will also discuss whether music might have an impact on the development of intellectual functions.

## Introduction

In Western culture musical expertise and performance has traditionally been associated with various so-called non-musical cognitions (e.g. memory and attention or even intelligence). In former times it was conjectured that musicians are musically gifted because of a different cognitive structure (and the underlying cortical organization supporting these cognitive functions) furnishing them with extraordinary musical performance. Other lines of theory and evidence have proposed that the cognitive structure of musicians undergoes substantial changes due to the intensive musical training required to become a musician. This latter approach has attracted much interest within the last two decades because of its potential implications for education. Recently, it has also been debated that the presumed positive effect of musical performance on cognition (and thus, intelligence) could be exploited in neuropsychological therapies. Although the above-mentioned associations between cognition and musical expertise have received much attention, especially in the lay press, there are astonishingly few substantial scientific publications addressing this topic. In the following we will describe and discuss scientific findings which may help to elucidate the potential relationship between cognition and musical performance.

There is no doubt that specific cognitive and motor skills are required to perform music. For example, playing a musical instrument makes it necessary to

program, execute, and sequence various movements. In case of performing music from a musical score it is necessary to read printed music, which is different from reading letters. Thus, it is necessary to understand particular notes and to translate these visual forms into appropriate motor commands, usually very rapidly. If the musician is only reading the musical score without performing it (for example in a situation when s/he is composing musical pieces) s/he reads the score and transforms the notes into auditory representations which help them to imagine how the musical piece will sound. Another issue which emphasizes the multitude of cognitive functions associated with performing music is the fact that musicians often learn to play a musical piece by heart, thus relying strongly on memory functions. Finally, learning to play a musical instrument requires attention, concentration, discipline, and motivation. Although this short summary emphasizes the specific cognitive functions necessary to perform music, relatively little scientific work has been published on this topic. In the following we will focus on the particular cognitive functions believed to be strongly associated with musical expertise. We will first describe cognitive functions that are closely related to musical expertise (sight-reading, memory for retrieving musical information, auditory attention, and auditory perception), and will also discuss those so-called non-musical skills which are thought to be indirectly related to musical expertise.

## Perception

For non-musicians, hearing of music is mostly associated with strong emotional reactions (pleasure and aesthetic feelings, but also negative emotions in situations where music is not appropriate). However, music can automatically evoke motor actions, even in non-musicians. The most obvious example is dancing or swaying to music. Thus, there must be a strong link between perception and action. For skilled musicians this sensorimotor association is very important. For example, playing music in an orchestra requires continuous ongoing adaptation and comparison of one's own performance with the auditory templates from the rest of the orchestra (feedback control). The musician has to play with the appropriate sound quality, register, rhythm, or tempo. In the case of composing, it is necessary to play a short musical phrase, evaluate the generated sound pattern by ear, and then put it on record by writing it up in musical notation.

From studies of human brain lesions and brain imaging it is known that specific brain structures specialize in analysing pitch, timbre, and contour relations, as well as musical timing and rhythm (Peretz and Zatorre, 2004). Most of these neural networks are functional in non-musicians, suggesting that these structures have evolved in the human brain independently of intensive training

or musical experience (for a summary see Münte *et al.* 2002). However, several studies have shown that these structures can be shaped by musical experience. For example, the timbre of a particular note (overtone structure) is differently processed in musicians and non-musicians. This has been shown in a number of behavioural experiments in which musicians and non-musicians had to decide which of two tones were similar or different in terms of pitch. If the pitch of the two tones remained the same but the timbre varied (e.g. a guitar tone versus a piano tone), non-musicians made significantly more errors than musicians. In one study there was even a strong correlation between accuracy of discrimination and performance in pitch–timbre discrimination ($r = 0.66$), indicating that musicians use pitch information to a greater extent than non-musicians (Pitt and Crowder 1992). Beal (1985) tested musically trained and untrained subjects on their ability to discriminate changes in timbre. Akin to the aforementioned experiment, Beal found that discrimination accuracy was almost perfect (98%) when timbre and pitch information remained constant. However, when pitch and timbre information varied (e.g. comparing tones of similar pitch but different timbres; guitar versus piano) the performance of the non-musicians dropped substantially; they were half as accurate as musicians. Pitt and Crowder made the claim that timbre information, and not pitch, is the more salient dimension for non-musicians because timbre delivers more relevant information about environmental events. However, analysis of pitch should therefore require more formal musical training. Thus, in non-musicians differences in timbre between two tones should be distracting even if the pitches are identical, causing problems with pitch discrimination.

Recent brain imaging studies have revealed that timbre processing is different in the auditory cortex depending on the familiarity of the presented tone. Specifically, it has been demonstrated that auditory cortical representations of tones of different timbre (violin and trumpet) are enhanced compared with sine wave tones in violinists and trumpeters, preferentially when they hear timbre tones of their own instrument (Pantev *et al.* 2001). These findings suggest that experience-dependent plasticity optimizes the auditory system in musicians.

While the above-cited experiments describe a characteristic formation and specialization of a musician's auditory system for processing timbre and pitch, another principle of learning is associated with the improved cognition of musicians. From the standpoint of cognitive psychology, optimization of musical perception is associated with the processing of larger perceptual chunks on lower processing levels with less processing capacity. Thus, during the course of musical training musical perception is optimized by making many steps more automatic. It has recently been shown that perception of pitch contour and

interval deviations in simple melodies is associated with a marked event-related brain response, termed mismatch negativity (MMN), which reflects automatic processing of incoming stimuli in skilled musicians. Thus it is suggested that musical training enhances the ability to automatically register abstract changes in the relative pitch structure of melodies (van Zuijen *et al.* 2004).

Besides improved musical perception, musicians also show some kind of changed memory processes for recognizing familiar tunes. This has recently been examined in an elegant experiment published by Bella *et al.* (2003). They used a specifically adapted recognition paradigm by sequentially presenting the tunes of a melody and requiring the subjects to provide a familiarity judgement or recognize the melody. Interestingly, musicians judged the familiarity of the musical pieces within fewer notes than did non-musicians (after four notes in musicians and five notes in non-musicians). Thus, familiarity evolves a little bit earlier in musicians, suggesting that perceptual memory (mostly working automatically) for melodies works faster in musicians.

Taken together, musicians outperform non-musicians in several aspects of musical perception, including perception of pitch, timbre, and timing. In addition, they also show improved performance in recognition of melodies. Altogether, these enhanced aspects of cognition are necessary for the performance of music even at moderate levels.

## Sight-reading

Reading musical scores is in many ways similar to reading words. The special ability to read musical scores while playing an instrument is called sight-reading. In order to sight-read music, a musician must translate notes on a page to appropriate motor commands subserving specific movements. In addition, these notes should also be interpreted and translated into aesthetically appealing signals. Consider the really complicated task that a pianist has to perform while playing music from a score. At a basic level (the level of a beginning student) there is a direct translation between a note and the striking of a corresponding key. For example, reading the note F results in striking the key for note F on the keyboard. At a virtuoso level, the pianist looks at a series of notes and merges that immediate information with knowledge held in long-term memory about musical harmonies, global rhythm phrases, and the interpretation of the overall musical line. In addition, the experienced musician may relate the current piece of music to his or her musical schema. This may be fused with the emotional theme of the music. And these are only a few of the more obvious components inherent in musical playing and interpretation.

In order to accomplish all these cognitive processes during sight-reading it is basically impossible to play a musical piece 'note by note'. Therefore, skilled musicians must use a strategy allowing them to look ahead in the score so as to anticipate what may come next, in a similar way to which an experienced reader might read a book. Although the above given explanation of the sight-reading process in musicians is somehow plausible, only a few experiments have studied this phenomenon empirically. Sophisticated experiments have been carried out by Truitt *et al.* (1997). Using recordings of eye movement and sophisticated analysis of key strokes they analysed the so-called eye–hand span and the perceptual span. The eye–hand span is a measure of the distance between the note being looked at and the note being played. The perceptual span is the region around the note being looked at from which useful information is extracted. For example if one looks at a particular letter in a word the perceptual span is the number of neighbouring letters which are also perceived. The authors found that the perceptual span of skilled and less skilled pianists was approximately the same, ranging between two and four beats. However, the eye–hand span differed between skilled and less skilled pianists. For the less skilled pianists, the eye–hand span was only about half a beat, indicating that the fixation point was less than one beat ahead of the hands. The eye–hand span for the skilled pianists was about on average two to three beats; thus, the fixation point was much further ahead of the hands than in non-skilled pianists. In a few instances the skilled pianists focused behind the note currently being played, suggesting that the player may have been reflecting on the note just played. In a way these findings are astonishing in that they showed that the visual perceptual span is similar for skilled and less skilled pianists and that the eye–hand span of skilled pianists is not much different between skilled and less skilled pianists, although skilled pianists show a larger eye–hand span. Although the eye–hand span is larger in skilled pianists it does not explain the virtuoso level at which skilled pianists play their instrument. Thus, there is a discrepancy between these experimental data and the conventional wisdom that pianists might extract information from a wider region of vision than non-musicians. The reason for the apparent similarity (or small difference between skilled and less skilled musicians) is possibly due to the constraints of short-term memory. If the encoding process gets too far ahead of the motor output (i.e. when the eye–hand span becomes too large), there is probably a loss of information that is stored in short-term memory. However, the evidence that skilled musicians can perform at virtuoso levels emphasizes that some kind of optimization has been implemented not on the visual but rather on the motor side of this complex sensorimotor process. Obviously, skilled pianists have developed the ability to translate complex visual information (e.g. several notes) into a complex motor program subserving the fast movements associated with piano playing.

## Cognitive control of musical performance

Playing a musical instrument or singing even at a low level of skill requires a highly adapted organization of many submovements. By preparing and executing a particular movement, an organism normally has great latitude in combining component movements in order to reach a goal. When the movement is actually executed, only a small number of the possible movements are used. The degrees of freedom problem refers to the number of distinct ways in which the movement can be performed. A movement can vary in many ways, including its angular direction, extent of muscle contraction, and motor neuron discharge. Thus, when practicing a particular movement various movement parameters have to be optimized. Rosenbaum (1991) identified at least three factors that are optimized during practising movements: efficiency, interaction, and mechanics:

1. Efficiency means the avoidance of extremes in movement parameters because it is known that extreme movement parameters are difficult to control. These parameters include rate changes in the velocity of movements, in muscle torque (the force to produce rotational motion), and in jerk (the rate of change of acceleration).
2. Interactions between limbs: the movements of our limbs tend to be coupled so that the movement of one limb or joint limits the movement of the other. For example, finger, hand, and arm movements are coupled. These and other interdependences between the limbs reduce the number of independent choices for the motor system when executing movements. Thus, programming a movement with one limb will inevitably cause particular movements in the coupled limb, a phenomenon reducing the degrees of freedom of movement.
3. The mechanical properties of the organism and of the environment aid in selecting and guiding movements without having to plan and to coordinate them. For example, moving a finger upwards needs muscle contractions. However, moving the finger downwards basically needs less muscle force because gravity takes over to bring the fingers back to the initial position.

Several models have tried to explain how the cognitive system accomplishes optimization in the context of skilled movements. Rosenbaum *et al.* (1993) proposed a knowledge-based model to explain the degrees-of-freedom problem. According to this model target postures (e.g. final movement positions including various constellations of submovements) are stored in movement memory. Each of the postures receives a weight that is based on the likelihood of the success of the posture. For each posture each degree of freedom also has a particular weight. The weights capture the costs of the posture, both in terms of the accuracy and

the effort of execution. Once a posture has been selected, a trajectory is chosen by the motor system incorporating biomechanical constraints. In relation to musicians one might speculate that as the level of musical skill increases the degrees of freedom are reduced to the control of particular music-related movements. Thus, musicians can control more complex movements with less control effort. This might explain the existence of the enlarged neural networks found in the sensorimotor system of professional musicians (Münte *et al.* 2002).

Motor learning is also related to adjustments of mental plans of movements. Mental plans are cognitive representations of movements which are independent of the limbs with which the movement can be performed. For example, mental plans contain conceptual relations of actions (e.g. the sequence of submovements). Whether mental plans and lower-level motor actions are learned independently and whether this learning is dependent on the level of skill is a long-standing matter of dispute. Palmer and Meyer (2000) have addressed this issue in more detail. They required skilled adult music performers and novice child pianists to learn a motor sequence task under speeded up conditions. After this initial practice phase they asked the subjects to perform a second sequence with the same or different motor assignments (hand or finger) and the same or different conceptual relations (pitch sequence). The results they found bear most important implications for understanding the relationship between cognition and movements in general and provide invaluable insight into the way in which skilled musicians organize and represent their movement control in particular. They found that with increasing skill, mental representations of performance become dissociated from the movements required to produce a musical sequence. Thus, it is plausible that the mental plans of advanced performers are based on abstract, conceptual pitch relations and are independent of the hand or finger movements. In contrast, novice child pianists showed transfer of learning reflecting the movements as much as the abstract pitch relationship. These findings suggest that the motor and conceptual dimensions become distinct within a sequence representation as learning to perform progresses. In addition, these findings throw a new light on the definition of motor skill in showing that cognitive *and* motor functions determine the level of motor skill.

## Memory functions

Perception and production of music imposes specific demands on the various memory structures. One important reason for this assumption is based on the fact that musical sounds (like all auditory signals) unfold over time. Thus, it is necessary for the auditory system to integrate the sequentially ordered sounds into a coherent musical perception. This series-to-parallel transformation can be

considered a mechanism of working memory, temporarily storing auditory units and combining them to a single percept (e.g. a sound pattern). In addition, recognizing and generating musical tunes makes it necessary that some invariant properties of musical pieces are stored in long-term memory. Most researchers believe that a specific, more perceptually organized musical memory system represents information about the form and structure of musical events (perceptual representation). Thus, unlike in language, music is thought to be mainly perceptually driven while language memory systems are strongly related to a rigid semantic system. However, although musical information is not directly related to semantic information (name of the singer, title of the piece, date of recording, etc.) it could be associated with emotional and semantic information (associative memory). A very specific feature of the perceptually based music memory system is that the stored information is relatively abstract, allowing the recognition despite change in instrumentation, loudness, tempo, or register. Lesion studies as well as recent brain imaging studies have shown that this (perceptual) memory system is located bilaterally in the auditory cortex (including the supramarginal gyrus). In addition, inferior frontal and inferior temporal brain areas have been shown to be strongly involved in recognizing familiar tunes.

It is relatively simple for musicians to recall and recognize complex musical works for performance. This is not only true for performers playing an instrument or singers, it is also most evident for conductors who have to memorize not only the musical structure but a lot of detailed information about instrumentation, tempo, and many other aspects of the musical piece. Thus, musicians have to establish a kind of expert memory system for musical pieces. Recently, much work has been devoted to identifying the characteristics of the expert musical memory system. Several authorities argue that musical memory is a prototypical example for other kinds of expert memory shown by chess players, mnemonists, or scientists (Chase and Ericsson 1981; Williamon and Valentine 2002; Chaffin *et al.* 2003). The skilled memory theory has been proposed on theoretical grounds to account for the remarkable memory abilities observed in experts across several domains including musical performance. This theory is based on three core aspects. First, experts are able to use their existing knowledge structures in semantic memory to store new information during skilled performance of a given task. Second, this information is incorporated into a specific memory mechanism in which cues, item information, and associative information are used to facilitate encoding and retrieval of information in and from long-term memory. This memory mechanism is called the retrieval structure. Third, the time required for encoding and retrieval operations decreases with extended practice.

Ericsson and Kintsch (1995) have recently extended the skilled memory theory into the long-term working memory (LTWM) theory. Of crucial importance for the LTWM framework is the way in which domain-specific information is stored and accessed. Ericsson and Kintsch (1995) propose that information is encoded and retrieved by means of (1) a retrieval structure (as mentioned above) and/or (2) through knowledge-based associations connecting items to other items or to LTM patterns and schemas. According to this, musicians will use their 'individual' understanding of the music to organize the retrieval cues and the retrieval structure.

But how do musicians strengthen their retrieval structure compared with non-skilled musicians? All experts strengthen retention by mapping the to-be-remembered information in terms of something that is familiar, sensible, or important to them. Often the to-be-remembered information is also divided into structurally important and less (subjective) important information. The structurally important information provides the cue for the retrieval structure. Recently, Williamon and Egner (2004) showed that particular bars within a complex piece of music serve as structural markers for encoding to and retrieval from memory. Thus, not each specific bar or note is individually encoded or retrieved; only specific bars (the so-called structural bars) are preferentially processed during encoding and retrieval. However, this information is sufficient for superior memory processing in the context of a very efficient retrieval structure.

As well as the fact that expert musicians have a specific and most efficient retrieval structure which helps them to memorize musical pieces, non-professional (but also professional) musicians have consistently demonstrated an efficient verbal memory system superior to that of non-musicians—at least in tonal languages like Chinese. This has been shown in a series of experiments conducted by Chan *et al.* (1998). Recently, Sluming *et al.* (2002) found that in musicians there is an increase in grey matter in the frontal cortex known to accommodate neural networks involved in several important memory processes. Thus, one might conclude that a kind of positive transfer between musical performance and verbal memory functions has taken place. Further experiments are needed to clarify to what extent this transfer is indeed a transfer between different functions (music to verbal). Recent studies have elegantly demonstrated a considerable overlap of neural structures and processes underlying the perception of language and music (Schuppert *et al.* 2000; Sluming *et al.* 2002; Koelsch *et al.* 2002, 2004; Patel 2003). Thus, performing one domain will necessarily influence the other domain because both music and language are processed by the same network and/or similar processes.

## Attention and perception

Playing music requires specific attentional abilities. For example, a performing orchestral musician has to simultaneously separate and integrate self-generated from externally generated musical stimuli while the motor actions necessary to play his/her instrument have to be executed. Attentional abilities are most evident in conductors, who can identify specific musicians within a multiplayer section. Münte *et al.* (2003) elegantly showed that conductors are endowed with exceptional auditory spatial abilities, enabling them to localize sounds even in positions of extremely eccentric peripheral auditory space. However, it is not only conductors who are endowed with superior auditory spatial attention but also musicians playing instruments at various levels of skill. For example, it has been shown that musicians, relative to non-musicians, pre-attentively extract more information from musically relevant stimuli (Rosenbaum *et al.* 1993). Since effects of long-term experience on pre-attentive auditory processing have so far only been reported for language-specific phonemes, these results indicate that sensory memory mechanisms can be modulated by training on a more general level.

## Visuospatial functions

That visuospatial functions might be related to musical performance seems to be a little bit unusual at first sight. However, this issue has been raised by the seminal publication of Rauscher *et al.* (1993). In this paper the authors reported that listening to a Mozart sonata for several minutes enhanced visuospatial functions in kindergarten children. From then on this effect has found its way into the literature as the 'Mozart effect'. Based on this and similar findings the authors suggested a strong relationship between visuospatial performance and listening to or processing music. One of the authors (Shaw 2000) even argued that visuospatial and musical processing might be controlled by similar neural assemblies and neuronal firing rate. Although these publications have attracted much attention, debate, and follow-up studies, attempts to replicate this finding have been mostly unsuccessful in studies conducted by independent laboratories. Meanwhile, several studies have shown that emotional state (mostly positive ones) or a certain level of arousal should be considered a most important moderating variable in determining this effect. However, although the 'Mozart effect' has been shown to be fragile and not directly related to music, one can speculate that there might exist some (admittedly not well understood) relationships between visuospatial functions and musical performance. For example, many ways of performing musical pieces require spatial functions. For example, pianists need three-dimensional spatial processing to find the correct position of

the keys while they are playing. Recently, Lutz *et al.* (2005) showed in a functional magnetic resonance imaging study that simple finger sequencings in non-musicians (similar to those used during piano playing) evoked strong activations in parietal areas (including the intraparietal sulcus) known to be involved in spatial processing (Jordan *et al.* 2001). Thus, it might indeed be plausible that professional musicians automatically practice spatial functions while performing music. However, this is an issue which needs further research in the future.

## Conclusion

This chapter has described several cognitive functions for which musicians (and in particular professional musicians) have developed specific cognitive functions. Most of these functions are specific to musical proficiency but some of these functions are not directly related to musical performance (e.g. verbal working memory). Recently, Schellenberg (2001, 2004) has demonstrated that the IQ is slightly enhanced in those children who have participated in musical lessons for about a year. Although the IQ-enhancing effect is small and within the normal variability range of IQ values (six IQ points), this finding is of special importance because it is the first report published in a reputable scientific journal demonstrating a direct positive transfer between different cognitive functions in musicians in the context of a longitudinal study. Although many might take this finding as support for the notion that musical training might be an ideal tool for educating pupils, this finding more strongly supports the notion that cognition is strongly associated with action, because it is not passive hearing of musical pieces that enhances the IQ, it is rather the active musical performance.

## References

**Beal, A.L.** (1985). The skill of recognizing musical structures. *Mem Cognit*, **13**, 405–412.

**Bella, S.D.**, Peretz, I., and Aronoff, N. (2003). Time course of melody recognition: a gating paradigm study. *Percept Psychophys*, **65**, 1019–1028.

**Chaffin, R.**, Imreh, G., Lemieux, C., and Chen, C. (2003). Seeing the big picture: piano practice as expert problem solving. *Music Percept*, **20**, 465–490.

**Chan, A.S., Ho, Y.C.**, and Cheung, M.C. (1998). Music training improves verbal memory. *Nature*, **396**, 128.

**Chase, W.G. and Ericsson, K.A.** (1981). Skilled memory. In: J.R. Anderson (ed.) *Cognitive skills and their acquisition*, pp. 141–189. Erlbaum, Hillsdale, NJ.

**Ericsson, K.A. and Kintsch, W.** (1995). Long-term working memory. *Psychol Rev*, **102**, 211–245.

**Ho, Y.C., Cheung, M.C., and Chan, A.S.** (2003). Music training improves verbal but not visual memory: cross-sectional and longitudinal explorations in children. *Neuropsychology*, **17**, 439–450.

**Jordan, K., Heinze, H.J., Lutz, K., Kanowski, M., and Jäncke, L.** (2001). Cortical activations during the mental rotation of different visual objects. *NeuroImage*, **13**, 143–152.

**Koelsch, S., Gunter, T.C., Cramon, D.Y., Zysset, S., Lohmann, G., and Friederici, A.D.** (2002). Bach speaks: a cortical 'language-network' serves the processing of music. *NeuroImage*, **17**, 956–966.

**Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., and Friederici, A.D.** (2004). Music, language and meaning: brain signatures of semantic processing. *Nat Neurosci*, **7**, 302–307.

**Lutz, K., Wüstenberg, T., Heinze, H.J., and Jäncke, L.** Timing and response selection in short finger tapping sequences investigated with fMRI. *Exp Brain Res.* (Unpublished manuscript)

**Lutz, K., Koeneke, S., Wüstenberg, T., and Jäncke, L.** (2005)Asymmetry of cortical activation during maximum tapping in right and left handers. *Neurosci,* **373**, 61-66.

**Münte, T.F., Altenmüller, E., and Jäncke, L.** (2002). The musician's brain as a model of neuroplasticity. *Nat Rev Neurosci*, **3**, 473–478.

**Münte, T.F., Nager, W., Beiss, T., Schroeder, C., and Altenmüller, E.** (2003). Specialization of the specialized: electrophysiological investigations in professional musicians. *Ann NY Acad Sci*, **999**, 131–139.

**Palmer, C. and Meyer, R.K.** (2000). Conceptual and motor learning in music performance. *Psychol Sci*, **11**, 63–68.

**Pantev, C., Roberts, L.E., Schulz, M., Engelien, A., and Ross, B.** (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *NeuroReport*, **12**, 169–174.

**Patel, A.D.** (2003). Language, music, syntax and the brain. *Nat Neurosci*, **6**, 674–681.

**Peretz, I. and Zatorre, R.J.** (2004). Brain organization for music processing. *Annu Rev Psychol.*

**Pitt, M.A. and Crowder, R.G.** (1992). The role of spectral and dynamic cues in imagery for musical timbre. *J Exp Psychol Hum Percept Perform*, **18**, 728–738.

**Rauscher, F.H., Shaw, G.L., and Ky, K.N.** (1993). Music and spatial task performance. *Nature*, **365**, 611.

**Rosenbaum, D.A.** (1991). *Human motor control.* Academic Press, San Diego, CA.

**Rosenbaum, D.A., Engelbrecht, S.E., Bushe, M.M., and Loukopoulos, L.D.** (1993). Knowledge model for selecting and producing reaching movements. *J Mot Behav*, **25**, 217–227.

**Schellenberg, E.G.** (2001). Music and nonmusical abilities. *Ann NY Acad Sci*, **930**, 355–371.

**Schellenberg, E.G.** (2004). Music lessons enhance IQ. *Psychol Sci*, **15**, 511–514.

**Schuppert, M., Münte, T.F., Wieringa, B.M., and Altenmüller, E.** (2000). Receptive amusia: evidence for cross-hemispheric neural networks underlying music processing strategies. *Brain*, **123**(3), 546–559.

**Shaw, G.L.** (2000). *Keeping Mozart in mind.* Academic Press, San Diego, CA.

**Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes, A., and Roberts, N.** (2002). Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *NeuroImage*, **17**, 1613–1622.

**Truitt, F.E., Clifton, C., Pollatsek, A., and Rayner, K.** (1997). The perceptual span and eye-hand span in sight reading music. *Vis Cogn*, **4**, 143–161.

**van Zuijen, T.L., Sussman, E., Winkler, I., Naatanen, R., and Tervaniemi, M.** (2004). Grouping of sequential sounds–an event-related potential study comparing musicians and nonmusicians. *J Cogn Neurosci*, **16**, 331–338.

**Williamon, A. and Egner, T.** (2004). Memory structures for encoding and retrieving a piece of music: an ERP investigation. *Brain Res Cogn Brain Res*, **22**, 36–44.

**Williamon, A. and Valentine, E.** (2002). The role of retrieval structures in memorizing music. *Cogn Psychol*, **44**,1–32.

Chapter 3

# The nature of memory for music performance skills

Caroline Palmer

Although most studies of memory for music performance focus on dimensions of musical structure, recent studies suggest that motor factors of performance are also represented. Brain imaging measures collected during mental practice or listening tasks suggest that both motor and auditory cortical areas are active during musical thought processes. Four types of behavioural study with performers reveal motor-based representations: performers' musical interpretations, transfer of learning from one musical task to another, mental practice effects, and anticipatory movements. These sources of evidence implicate distinct motor and non-motor (melodic) information in performers' memory for music. Implications from these behavioural tasks suggest that an accurate auditory and motor representation underlies successful performance from memory.

## Introduction

Theories of memory often cite music performance as an example of expert memory, and many studies have examined the factors that influence performers' memory (for review see Palmer (1997) and Gabrielsson (1999)). Music performance can be depicted as a cognitive skill with large memory demands, or as a motor skill with large physical execution demands. Cognitive dimensions of performers' memory for music are often described in terms of musical structure, such as harmony, tonality, phrasing, and meter (Palmer 1997; Gabrielsson 1999; Palmer and Pfordresher 2003). Motor aspects of performance are measured in physical dimensions such as fluency, speed, rhythmic precision, and hand coordination (Krampe and Ericsson 1996; Drake and Palmer 2000). Musicians' memory that allows them to perform a specific musical piece, referred to as 'memory for performance', is distinguished here from other memories of performance (by oneself or someone else).

Music performance is of much interest to psychologists precisely because it is demanding both cognitively and in terms of motor skills. Memory for music entails sequences on the order of thousands of pitch events that are produced at

rates as high as 8–12 events per second, with less than 3% error (Palmer and van de Sande 1993; Repp 1996; Finney and Palmer 2003). In addition to categorical pitch events, listeners (infants as well as adults) remember fine expressive nuances of performances, including timing, intensity, and articulation information that distinguish different performances of the same music (Palmer *et al.* 2001). A more accurate conceptualization of the problem of describing memory for performance may be: what are the underlying features of music performance that are represented in memory? Which features are described in motor coordinates? How do factors such as maturation, musical acculturation, and training influence the representation of those features?

This chapter addresses these questions in the domain of piano performance. Piano is one of the most-often chosen instruments of study by children in North American schools (Johnson 1994). Piano has consistently been among the most-often performed instrument by amateur musicians in American households (Gallup Poll 1994, 2000) and piano performance is one of the most easily measured instruments with the advent of MIDI (music instrument digital interface). Thus, research in performance is based largely on measures of pianists' behaviours. Despite the rapid growth of research into piano performance during the past 20 years, less attention has been paid to the motor aspects of memory for performance (such as finger and hand movements) than to auditory (intensity, duration, etc.) or structural aspects (phrase structure, rhythm, etc.).

## Motor representations of piano performance

The motor demands of piano performance are quite large, especially in terms of finger movements. One of the few models to attempt to capture aspects of motor difficulty is an ergonomic model of fingering in piano performance (Parncutt *et al.* 1997). This model, based on a series of rules, reflects specific ergonomic difficulties such as weak versus strong fingers and spans between finger pairs. Measures of the pianists' fingering consistency and performance accuracy for musical exercises correlated negatively with the model's predictions of fingering difficulty (Sloboda *et al.* 1998). The model compared favourably with performances that used the least difficult fingers; however, the model sometimes made predictions for simpler fingerings that were not preferred by the pianists.

Imagery techniques applied to brain organization also support motor and non-motor features of performers' memory for music. Pantev *et al.* (2001) found enlarged representations of the auditory and somatosensory cortex with musical practice. The somatosensory cortex showed neuronal changes specific to fingers used during practice. Pascual-Leone and colleagues (reviewed in Pascual-Leone

1995) found changes in cortical motor areas as non-pianists practised finger exercises; physical practice as well as a combination of mental and physical practice established those changes. With magnetoencephalograph (MEG) techniques, Haueisen and Knösche (2001) reported that pianists displayed motor activity in the contralateral primary motor cortex while listening to familiar music (music they had performed). In addition, functional magnetic resonance imaging (fMRI) measures in skilled violinists demonstrated activity in primary auditory regions while the violinists silently tapped the opening section of a violin concerto; the amateur violinists' fMRI results did not indicate auditory stimulation (Lotze *et al.* 2003). To address whether co-activation in auditory and motor cortical areas arises from years of joint auditory–motor practice or over shorter time spans, Bangert and Altenmüller (2003) used direct current electroencephalography (DC-EEG) with non-musicians who were trained to associate a purely auditory (passive listening) task with a purely motor (silent finger movement) task. One group had consistent mappings of auditory events to motor events, while another group learned with altered mappings. Only the consistent mapping group exhibited additional EEG activity in right anterior regions, which the authors took to indicate an audiomotor interface that developed in the first weeks of practice. Of interest is how far the behavioural changes following pairings of motor and auditory activity extend both in time and across tasks.

## Interpretive influences on performance memory

The occasional errors that performers make (hitting the wrong key or chord) often display active memories for other to-be-performed events. For example, pitch errors tend to reflect tones intended for elsewhere in the same piece (Palmer and van de Sande 1993). Such performance errors also indicate motor features. Palmer and van de Sande (1993) documented pianists' errors in two-handed piano music; left-handed errors were more common than right-handed errors (regardless of the pianists' handedness), consistent with Peters (1985) findings of greater right-hand coordination skills among pianists. Harmonically related errors (a substituted pitch that arises from the same chord as the intended pitch) were more common than harmonically unrelated errors, even when their production required different hand or finger movements (Palmer and van de Sande 1993). Likewise, hand differences in errors occurred for both harmonically related and harmonically unrelated errors.

Do motor factors interact with non-motor features of memory for performance? Palmer and van de Sande (1993) asked pianists to perform the same music with different melodic interpretations. The melody or primary (most important) voice is not directly marked in musical notation, and performers often interpret

the voice intended as melody as more important (Palmer 1989). The voice interpreted as melody was less likely to contain errors, even when pianists changed their melodic interpretation for the same musical piece. Performers also made more errors in parts controlled by the left hand, regardless of hand dominance. Motor factors (hands and fingers used) did not interact with interpretive factors (melody being emphasized) in the likelihood of errors in the pianists' performance, suggesting that interpretive and motor factors are represented independently in memory for performance (Palmer and van de Sande 1993).

## Transfer of performance skills

Another source of evidence for motor dimensions of memory for performance is the conditions under which performers generalize what they know from one performance situation to another. In typical transfer of learning tasks, participants learn one task and then perform a second task. The ease with which they perform the second task is thought to reflect what was learned in the first task. The mirror symmetry of hand and finger movements in piano performance provides a convenient format in which to test transfer of motor learning. Consider the sequence of finger movements 5–4–3–2–1 in the right hand (where thumb is 1), used to press adjacent keys on a piano; this sequence becomes 1–2–3–4–5 when the same keys are pressed by the left hand. Thus, the same melody can be played with different hand and finger movements in piano performance. In addition, different melodies (different pitch sequences) can be performed with the same hand and finger movements. Although the transfer of musical skills to math, spatial reasoning, and other tests of intelligence has been studied (Rauscher *et al.* 1995; Schellenberg 2004), fewer studies have examined the transfer of learning from one musical task to another musical task. Are well-learned motor movements transferred from one melody to another? Palmer and Meyer (2000) measured the transfer of pianists' hand and finger movements from one practised melody to a novel melody. When the second (novel) melody required the same finger movements as the initial melody performed, pianists were able to play it more quickly. Comparisons across age and skill levels indicated that the more advanced performers showed the greatest transfer across motor movements, suggesting that they were able to generalize from one hand and finger set to another. The least skilled pianists showed no ability to generalize beyond the particular finger sequence learned.

Does knowledge of a set of motor movements transfer across temporally distinct patterns in music performance? For example, can a pianist performing one rhythmic pattern with a set of finger movements such as 1–2–3–4–5 generalize those same finger movements to produce a different rhythmic pattern as fluently

as the first pattern? Using a similar transfer of learning task, Meyer and Palmer (2003) showed that pianists' finger movements did transfer across different rhythmic patterns: pianists could perform novel rhythms with the same finger movements as quickly as the first-learned rhythms. In addition, there was rhythmic transfer: pianists could perform well-learned rhythms with different hand and finger movements as quickly as with the first-learned movements. There was no interaction between rhythmic transfer and motor transfer; the times at which keys were pressed were remembered independently of the motor features that produced them. Similar findings were obtained when pianists transferred from one melody to another that differed in meter and motor movements. No interactions were observed among the temporal (meter and rhythm) and motor structures: retaining temporal structure from one melody to the next facilitated speeded performance more than retaining motor movements, and motor features played a smaller role in the transfer of knowledge across melodies. These findings are consistent with the general view that representations of timing in sequence production are not defined primarily in terms of motor features (MacKay 1982, 1987; Semjen and Ivry 2001), at least for performers with moderate to high levels of musical experience.

## Mental and physical practice

Practice may be the single most important factor that influences memory for performance. Musicians' verbal reports of practice goals incorporate many levels of analysis, including structure, interpretation, and motor aspects of technique (Chaffin *et al.* 2003). Several studies suggest that deliberate rehearsal accounts for skill differences among music performers (Ericsson *et al.* 1993; Sloboda *et al.* 1996). More recently, comparisons have been made of the efficacy of physical practice with mental practice. When musicians practice the motor movements for performance in the absence of their physical instrument, they often make other overt movements such as drumming fingers on a tabletop or tapping feet under the table, suggesting that the motor features of performance are important for practice. If mental practice has effects on performance similar to those of physical practice, is it due to motor or non-motor components of mental practice? Are the thought processes underlying mental practice similar to those underlying physical practice?

Mental practice has been defined as the mental rehearsal of a specific task in the absence of actual physical movement (Coffman 1990; Driskell *et al.* 1994). Mental practice is distinguished from other mental techniques such as analytical study, general mental imagery, imitation, or self-arousal. Measurement of the effects of mental practice typically includes improvement in accuracy or time to

complete a task relative to some control task (such as no practice or normal practice). Meta-analyses of mental practice effects (Feltz and Landers 1983; Driskell *et al.* 1994) indicate two consistent findings. First, normal (physical) practice exceeds mental practice alone, which exceeds no practice (Coffman 1990). Second, physical practice plus mental practice instructions exceed physical practice alone (Rubin-Rabson 1937; Ross 1985). These findings suggest that improvement with practice in general is due to two components: a physical (motor) component and a mental (non-motor) component. The meta-studies interpret differences among mental practice findings as indicating that tasks with greater cognitive requirements (memory, attention, symbol manipulation) show greater effects of mental practice than tasks with high-motor requirements (coordination, endurance, strength). Examples of tasks with a high cognitive load include maze-learning and card-sorting. Examples of tasks with a low cognitive load include balancing and dart throwing.

Two common theoretical explanations of the effects of mental practice include the cognitive (symbolic) hypothesis and the psychoneuromuscular hypothesis. The cognitive hypothesis states that the effects of mental practice apply to the cognitive components of a skill: those that can be represented symbolically or visuospatially (Feltz and Landers 1983). In contrast, the psychoneuromuscular hypothesis (which grew out of the ideomotor view) holds that the effects of mental practice apply to the physical components of a skill, such as low-gain innervation of muscles used in the physical enactment of the skill (Jacobson 1930; Shaw 1938). However, potential problems with studies of mental practice make conclusions difficult. Instructions to participants that define mental practice often vary or are absent. In addition, mental practice conditions are often accompanied by some kind of physical practice, either overt, such as foot-tapping (Wollner and Williamon 2004), or covert, such as movements of throat muscles (Jacobson 1932). As a result, the components of mental practice that are considered cognitive or motor are often determined *a priori* by experimenters.

Although the effects of mental practice have been examined in many behavioural tasks, fewer studies have compared types of mental practice in the context of music performance. Coffman (1990) showed that mental practice by pianists improved their performance compared with no practice. Pianists' mental practice with an auditory model showed advantages over mental practice alone (Lim and Lippman 1991; Theiler and Lippman 1995). Rubin-Rabson (1937) showed that analytical pre-study of the score often aided performers' memorization of unfamiliar music; this analytical study may have involved auditory or motor imagery (Lim and Lippman 1991). Mental practice may help musicians learning unfamiliar music by facilitating the creation of an auditory and/or motor image. Comparison of different practice conditions showed that listening to a performance

was an effective aid to learning to perform unfamiliar music (Rosenthal 1984; Rosenthal *et al.* 1988). Not all studies show a facilitation of mental practice, however; Rosenthal *et al.* (1988) found that silent analysis (similar to mental practice, without explicit instructions) followed by sight-reading was no more effective than sight-reading alone.

An indirect source of evidence for mental practice is the lack of any detrimental effect that removal of auditory feedback causes to performers once they have learned a musical piece well. Finney and Palmer (2003) measured the amount of time performers took to play a familiar piece from memory on a silent electronic piano; total performance durations were within 5% of the durations when the pianists played the same piece with normal auditory feedback. This result was not specific to well-learned pieces; after playing a novel musical piece 10 times, the removal of auditory feedback caused no change in the duration of the performance. However, when auditory feedback was removed during the initial practice session, pianists' later performances (with auditory feedback) were significantly slower and contained more errors (Finney and Palmer 2003). Thus, the absence of auditory feedback during practice of an unfamiliar piece did not affect pianists' accuracy while the music notation was in front of them, but the absence of auditory feedback during practice did impair their later performance from memory. Ross' (1985) study of trombonists indicated that normal practice conditions with both auditory and kinesthetic feedback present allowed musicians to correct and adjust their performances; that ability decreased as feedback was removed during practice. In sum, these studies suggest that pianists can substitute mental feedback for auditory feedback once they have practised the music sufficiently to form a mental representation.

Highben and Palmer (2004) contrasted mental and physical practice, in terms of auditory and motor feedback, to determine their role in how pianists learn to perform unfamiliar music. Comparisons were made of different types of mental practice by replacing auditory or motor feedback with instructions to imagine the missing feedback: how the piece sounds, or how the finger movements feel, during practice. In a 'normal' practice condition, pianists moved their fingers on the keys and heard themselves play over headphones during practice. In a 'motor only' practice condition, the pianists moved their fingers on the keys but auditory feedback was removed; they were told to imagine what the piece would sound like. In an 'auditory only' practice condition, motor feedback was removed (the pianists held their fingers in loose fists) and auditory feedback was present in the form of a computer-generated recording of the piece, and pianists were told to imagine what the finger movements would feel like. In a 'covert' practice condition, both motor feedback and auditory feedback were removed during practice (pianists held their fingers in loose fists during silence), and they were

given both sets of instructions. Each participant performed in each condition with a different musical piece; a within-subject design was considered important to control for individual differences in mental practice abilities. After performers practised from a musical score, the score was removed, and pianists performed from memory under normal feedback conditions. Two independent measures of ability for mental imagery were collected: one for motor imagery and one for auditory imagery, for comparison with the effects of mental practice.

Both auditory and motor forms of practice facilitated pianists' subsequent performance from memory of unfamiliar music. Removal of auditory or motor feedback at practice caused significant memory deficits in later performance, despite the presence of both types of feedback at test. Physical practice conditions (in which auditory and/or motor feedback were present) led to better performance recall than conditions with mental practice. Recall was best following normal practice conditions, and worst when both auditory and motor feedback were removed during practice. Furthermore, pianists who scored higher on the test of aural skills performed better from memory following the absence of auditory feedback during practice, compared with pianists who scored lower on the test of aural skills. Thus, it is likely that performers with high aural skills were better able to use auditory imagery during learning than other performers (Highben and Palmer 2004). Whereas previous studies demonstrated the overall efficacy of mental practice in music performance (Ross 1985; Coffman 1990), Highben and Palmer's (2004) findings suggest specifically that auditory and motor forms of mental practice can aid the learning of unfamiliar music by performers. In addition, the presence of a motor component of mental practice that facilitates memory for performance indicates that explanations of the effects of mental practice based solely on symbolic, videospatial, or other cognitive (non-motor) forms of representation are not sufficient to explain memory for performance.

Individual differences in imagery abilities that are related to memory differences have implications for brain imaging studies as well as for behavioural studies. Most comparisons of methods of mental practice rely on cross-group comparisons (see Feltz and Landers 1983; Driskell *et al.* 1994), for which any correlated memory differences are not measured. The presence of individual differences suggests that within-subject designs—those that allow comparisons across all conditions within individual performers—combined with independent measures of behavioural correlates may be important controls for memory differences that result from differences in mental imagery abilities.

## Anticipatory behaviour in music performance

One of the hallmarks of memory for performance is anticipatory planning: the preparation of events prior to their execution (Rosenbaum 1991). Anticipatory

behaviour is evidenced in occasional errors that reveal events intended for the future, and also in movements during the production of correct events that reveal trajectories toward future events. Studies of speech and music performance show anticipatory behaviour in the types of errors people make (such as a speaker producing 'I took the store . . . ' instead of the intended 'I took the car to the store' (Garrett 1980)). One of the main factors influencing anticipatory behaviour is practice. Drake and Palmer (2000) compared the anticipatory behaviours of child and adult pianists of various skill levels. The percentage of anticipatory pitch errors (compared with perseveratory errors, or produced pitches that were intended for earlier in the sequence) increased with both age and experience (Palmer and Drake 1997; Drake and Palmer 2000). Practice effects on anticipatory behaviour are found in many domains. Speech errors suggested that with more practice, speakers were more likely to anticipate a phoneme that was intended for later in the utterance (Dell *et al.* 1997). As the overall error rate decreased, speakers' percentages of anticipatory errors increased. Palmer and Pfordresher (2003) also found a consistent increase in anticipatory behaviour with practice in piano performance, and a general relationship between overall error rate and anticipatory behaviour: as pianists' pitch error rates decreased, the percentage of anticipatory errors increased. Furthermore, pianists' pitch errors were likely to span sequence distances of three or four events, termed the 'range' of planning: the faster the performance, the smaller the range over which anticipatory behaviour was evidenced (Palmer and Pfordresher 2003).

Analyses of finger movements in piano performance also display evidence of anticipatory behaviour. One of the earliest studies of musical motion (Ortmann 1929), using photographic techniques, documented finger movements during piano performance. Ortmann's records indicated anticipatory interactions among finger movements (movement of the second finger before the third finger strikes a key), as well as anatomical measures (arm weight, finger length, hand position) relevant for piano performance (such as the role of forearm bone length on rotation during production of tremolo). Ortmann (1929) documented general principles that are still the focus of movement research today: there are multiple routes to reach any key on the piano (the degrees of freedom problem), and the way in which each key is struck is influenced by the ways in which preceding and subsequent keys are struck (co-articulation).

Current techniques of measuring motion rely on optoelectronic systems. One type includes active markers or sensors that are placed on joints with wires and emit infrared signals 'captured' by cameras. Another motion capture technique uses passive markers that reflect light generated from a separate source that is detected by cameras. Both systems pinpoint the three-dimensional coordinates of each marker at each point in time. Engel *et al.* (1997) used an optoelectronic

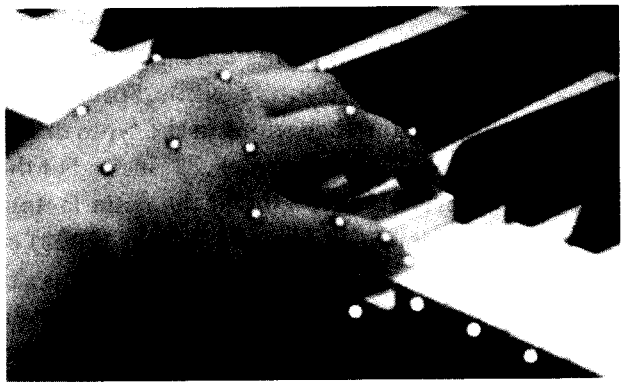


**Fig. 3.1** Picture of reflective markers on a pianist's right-hand fingers and piano keys. (From Palmer and Dalla Bella 2004.)

system to measure the degree of anticipatory behaviour of fingers as pianists performed short musical excerpts that began with the same pitches but diverged in the middle of the excerpt. Pianists' finger motions changed about 160–500 ms prior to the point of notated divergence in the musical score. Because performances differed in tempo, they were normalized with respect to time and anticipatory times were not reported in number of events.

Palmer and Dalla Bella (2004) measured anticipatory movements in piano performance of simple melodies performed at a range of specified tempi. Pianists' finger motions were recorded with a Vicon-8 system with passive 3 mm markers placed on a pianist's right hand, as shown in Fig. 3.1, and 14 cameras placed around the pianist recorded light reflected from the markers. This system, also applied to violin bowing (Visentin and Gongbing, 2003), has the advantage of requiring no wires on the pianists' hands. An example of a pianist's trajectory of motion of the fifth finger of the right hand is shown in Fig. 3.2, during a performance of the simple melody shown at the top. The position, velocity, and acceleration graphs below refer to the position of the marker placed on the tip of the fingernail (the finger position of greatest motion) in the vertical plane (height above the keyboard). The top panel shows the finger height; minimum values indicate where the piano key was pressed. The middle panel shows the velocity of the finger motion, and the bottom panel shows the acceleration. In this melody, the pianist's fifth finger pressed the keys on note events 6, 8, and 10. The pianist's trajectories of motion indicated changes in velocity and acceleration patterns of each finger prior to its arrival on a key. The key arrival (indicated by minimum finger height, top panel) is marked by peak finger acceleration (bottom panel). Anticipatory motion is evidenced in the finger heights (top panel) during the event prior to arrival on a key (events 6, 8,10).

As shown in Fig. 3.2, the pianist's fingers reached peak amplitudes usually within one event prior to a keypress. However, the trajectories of each finger



**Fig. 3.1** Picture of reflective markers on a pianist's right-hand fingers and piano keys. (From Palmer and Dalla Bella 2004.)

system to measure the degree of anticipatory behaviour of fingers as pianists performed short musical excerpts that began with the same pitches but diverged in the middle of the excerpt. Pianists' finger motions changed about 160–500 ms prior to the point of notated divergence in the musical score. Because performances differed in tempo, they were normalized with respect to time and anticipatory times were not reported in number of events.

Palmer and Dalla Bella (2004) measured anticipatory movements in piano performance of simple melodies performed at a range of specified tempi. Pianists' finger motions were recorded with a Vicon-8 system with passive 3 mm markers placed on a pianist's right hand, as shown in Fig. 3.1, and 14 cameras placed around the pianist recorded light reflected from the markers. This system, also applied to violin bowing (Visentin and Gongbing, 2003), has the advantage of requiring no wires on the pianists' hands. An example of a pianist's trajectory of motion of the fifth finger of the right hand is shown in Fig. 3.2, during a performance of the simple melody shown at the top. The position, velocity, and acceleration graphs below refer to the position of the marker placed on the tip of the fingernail (the finger position of greatest motion) in the vertical plane (height above the keyboard). The top panel shows the finger height; minimum values indicate where the piano key was pressed. The middle panel shows the velocity of the finger motion, and the bottom panel shows the acceleration. In this melody, the pianist's fifth finger pressed the keys on note events 6, 8, and 10. The pianist's trajectories of motion indicated changes in velocity and acceleration patterns of each finger prior to its arrival on a key. The key arrival (indicated by minimum finger height, top panel) is marked by peak finger acceleration (bottom panel). Anticipatory motion is evidenced in the finger heights (top panel) during the event prior to arrival on a key (events 6, 8,10).

As shown in Fig. 3.2, the pianist's fingers reached peak amplitudes usually within one event prior to a keypress. However, the trajectories of each finger

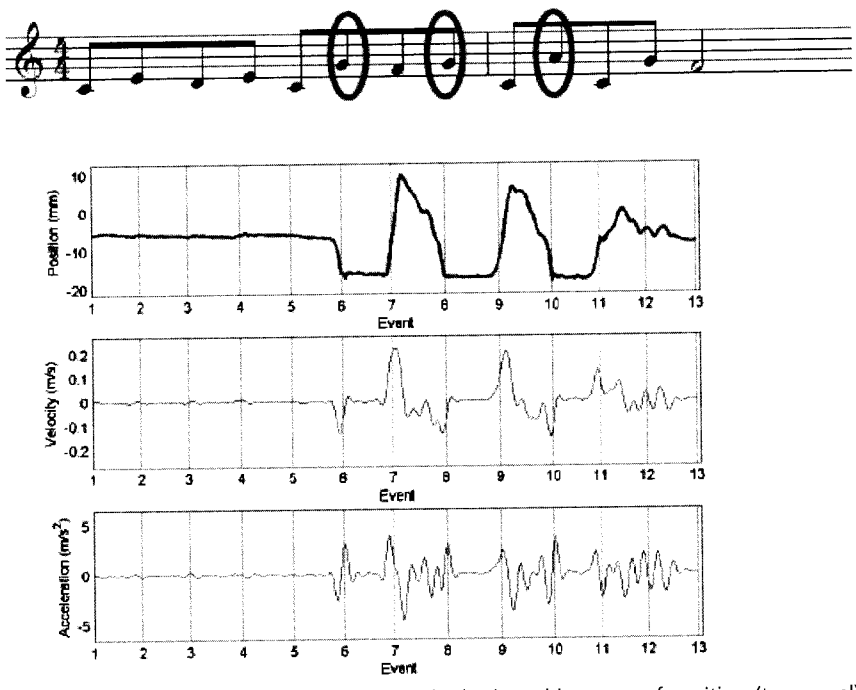



**Fig. 3.2** Height of a pianist's fifth finger above the keyboard in terms of position (top panel), velocity (middle panel), and acceleration (bottom panel) during a single performance of the melody shown at the top, at a moderate tempo (120 beats/min). Vertical lines indicate the time of each keypress (as recorded on a MIDI keyboard).

began to change in velocity and acceleration one to three events prior to the anticipated arrival at the key by that finger; by four events prior, the trajectories showed the same amount of change in velocity and acceleration as when the finger had no upcoming keypress (resting level) (Dalla Bella and Palmer 2004). This anticipatory behaviour in finger trajectories is consistent with findings of memory retrieval occurring three to four events before the keypress (Palmer and Pfordresher 2003); finger trajectories toward keypresses must require some information about the arrival location prior to the execution of the movement. Furthermore, the faster the tempo, the less time there is for anticipatory movements, consistent with the memory retrieval model. Although statistical (morphometric) techniques that identify consistencies in shape and structural time patterns have not yet been applied to musical movement, motion capture techniques offer promise for rigorous measurement of co-articulation properties (how finger movements are influenced by subsequent and preceding fingers) and other shape/time constancies in musicians' movements that were first identified long ago (Ortmann 1929).

## Summary

Research in music performance is beginning to document the nature of memory for the motor aspects of performance. Performers' hand and finger movements, as well as conceptual intentions, are encoded in memory for performance and tend to have independent effects on pitch accuracy. Performers' memory for melodies (specific pitch sequences) and finger/hand movements generalize in transfer tasks; furthermore, motor and melodic information transfer independently. Skilled performers show more transfer of learning across melodies that require different motor movements than do novices. Mental practice shows evidence of motor components that facilitate memory for performance. Finally, motion capture techniques for measuring music performance are beginning to document the time course of anticipatory movements.

One ramification of these findings is that memory for performance is flexible; performers can apply what they know about motor movements to different performance situations, and behavioural and neural changes result—as seen in the plasticity with which non-musicians learn aural–motor associations. Flexibility of motor movements is essential for the interpretive nature of music performance; otherwise, significant additional practice would be necessary before a musician could perform a familiar piece with an alternative interpretation. A second ramification is that performers differ in their individual abilities, as evidenced in interpretive effects on memory, in transfer of learning from one melody to another, and in mental practice benefits. Mental practice is appropriate for the study of brain states, measured in EEG, fMRI, and MEG studies, because of its avoidance of motion 'artefacts'. Scientific interest in applying imaging methods and motion capture techniques to music performance suggests that answers may soon be found to the interesting question of how motor aspects of music performance are represented in memory.

## Acknowledgements

This research was supported in part by the Canada Research Chairs Program, the Canada Foundation for Innovation, and NSERC Grant 298173. Simone Dalla Bella, Sean Hutchins, and Peter Pfordresher provided helpful comments on an earlier draft.

## References

**Bangert, M. and Altenmüller, E.O.** (2003). Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neurosci*, **4**, 26–40.

**Chaffin, R., Imreh, G., Lemieux, A.F., and Chen, C.** (2003). 'Seeing the big picture': piano practice as expert problem solving. *Music Percept*, **20**, 465–490.

**Coffman, D.D.** (1990). Effects of mental practice, physical practice, and knowledge of results in piano performance. *J Res Music Educ*, **38**, 187–196.

**Dalla Bella, S. and Palmer, C.** (2004). Tempo and dynamics in piano performance: the role of movement amplitude. In: S.D. Lipscomb, R. Ashley, R.O. Gjerdingen, and P. Webster (eds) *Proceedings of the International Conference on Music Perception and Cognition*, pp. 256–257. Causal Productions, Adelaide.

**Dell, G.S., Burger, L.K, and Svec, W.R.** (1997). Language production and serial order: a functional analysis and a model. *Psychol Rev*, **104**, 123–147.

**Drake, C. and Palmer, C.** (2000). Skill acquisition in music performance: relations between planning and temporal control. *Cognition*, **74**, 1–32.

**Driskell, J.E., Copper, C., and Moran, A.** (1994). Does mental practice enhance music performance? *J Appl Psychol*, **79**, 481–492.

**Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., and Taub, E.** (1995). Increased cortical representation of the fingers of the left hand in strong players. *Science*, **270**, 305–307.

**Engel, K.C., Flanders, M., and Soechting, J.F.** (1997). Anticipatory and sequential motor control in piano playing. *Exp Brain Res*, **113**, 189–199.

**Ericsson, K.A., Krampe, R.T., and Tesch-Römer, C.** (1993). The role of deliberate practice in the acquisition of expert performance. *Psychol Rev*, **100**, 363–406.

**Feltz, D.L, and Landers, D.M.** (1983). The effects of mental practice on motor skill learning and performance: a meta-analysis. *J Sport Psychol*, **5**, 25–27.

**Finney, S.A. and Palmer, C.** (2003). Auditory feedback and memory for music performance: sound evidence for an encoding effect. *Memory Cognition*, **31**, 51–64.

**Gabrielsson, A.** (1999). The performance of music. In: D. Deutsch (ed.) *The psychology of music*, 2nd edn, pp. 501–602. Academic Press, San Diego, CA.

**Gallup, G.H.** (1994). *The Gallup poll.* Scholarly Resources, Wilmington, DE.

**Gallup, G.H.** (2000). *The Gallup poll.* Scholarly Resources, Wilmington, DE.

**Garrett, M.F.** (1980). Levels of processing in sentence production. In: B. Butterworth (ed.) *Language production: speech and talk*, pp. 177–220. Academic Press, London.

**Haueisen, J. and Knösche, T.R.** (2001). Involuntary motor activity in pianists evoked by music perception. *J Cogn Neurosci*, **13**, 786–792.

**Highben, Z. and Palmer, C.** (2004). Effects of auditory and motor mental practice in memorized piano performance. *Bull Council Res Music Educ.* **159**, 58–65.

**Jäncke, L., Schlaug, G., and Steinmetz, H.** (1997). Hand skill asymmetry in professional musicians. *Brain Cogn*, **34**, 424–432.

**Jacobson, E.** (1930). Electrical measurement of neuromuscular states during mental activities. *Am J Physiol*, **94**, 22–34.

**Jacobson, E.** (1932). Electrophysiology of mental activities. *Am J Psychol*, **44**, 677–694.

**Johnson, Y.P.** (1994). Preferred general music classroom activities among low-income, urban-minority middle school students. Unpublished doctoral dissertation, Ohio State University.

**Krampe, R. and Ericsson, K.A.** (1996). Maintaining excellence: deliberate practice and elite performance in young and older pianists. *J Exp Psychol: Gen*, **125**, 331–359.

**Lim, S. and Lippman, L.G.** (1991). Mental practice and memorization of piano music. *J Gen Psychol*, **118**, 21–30.

**Lotze M., Scheler , G., Tan H-RM, Braun, C., and Birbaumer, N.,** (2003). The musician's brain: functional imaging of amateurs and professionals during performance and imagery. NeuroImage, **20**: 1817–1829.

MacKay, D.G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychol Rev*, **89**, 483–506.

MacKay, D.G. (1987). *The organization of perception and action: a theory for language and other cognitive skills*. Springer-Verlag, New York.

Meyer, R.K. and Palmer, C. (2003). Temporal and motor transfer in music performance. *Music Percept*, **21**, 81–104.

Ortmann, O. (1929). *The physiological mechanics of piano technique*. Dutton, New York.

Palmer, C. (1989). Mapping musical thought to musical performance. *J Exp Psychol: Hum Percept Perform*, **15**, 331–346.

Palmer, C. (1997). Music performance. *Ann Rev Psychol*, **48**, 115–138.

Palmer, C. and Dalla Bella, S. (2004). Movement amplitude and tempo change in piano performance. *J Acoust Soc Am*, **115**, 2590.

Palmer, C. and Drake, C. (1997). Monitoring and planning capacities in the performance of musical skills. *Can J Exp Psychol*, **51**, 369–384

Palmer, C. and Meyer, R.K. (2000). Conceptual and motor learning in music performance. *Psychol Sci*, **11**(1), 63–68.

Palmer, C. and Pfordresher, P.Q. (2003). Incremental planning in sequence production. *Psychol Rev*, **110**, 683–712.

Palmer, C. and van de Sande, C. (1993). Units of knowledge in music performance. *J Exp Psychol: Learn Mem Cog*, **19**, 457–470.

Palmer, C., Jungers, M.K., and Jusczyk, P.W. (2001). Episodic memory for musical prosody. *J Mem Lang*, **45**, 526–545.

Pantev, C., Engelien, A., Candia, V., and Elbert, T (2001). Representational cortex in musicians: Plastic alterations in response to musical practice. *Ann NY Acad Sci*, **930**, 300–314.

Parncutt, R., Sloboda, J.A., Clarke, E.F., Raekallio, M., and Desain, P. (1997). An ergonomic model of keyboard fingering for melodic fragments. *Music Percept*, **14**, 341–382.

Pascual-Leone, A. (1995). The brain that plays music and is changed by it. *Ann NY Acad Sci*, **930**, 315–329.

Peters, M. (1985). Performance of a rubato-like task: when two things cannot be done at the same time. *Music Percept*, **2**, 471–482.

Rauscher, F.H., Shaw, G.L., and Ky, K.N. (1995). Listening to Mozart enhances spatial-temporal reasoning: toward a neurophysiological basis. *Neurosci Lett*, 44–47.

Repp, B.H. (1996). The art of inaccuracy: why pianists' errors are difficult to hear. *Music Percept*, **14**, 161–183.

Rosenbaum, D.R. (1991). *Human motor control*. Academic Press, San Diego, CA.

Rosenthal, R.K. (1984). The relative effects of guided model, model only, guided only, and practice only on the accuracy of advanced instrumentalists' musical performance. *J Res Music Educ*, **32**, 265–273.

Rosenthal, R.K, Wilson, M., Evans, M., and Greenwalt, L. (1988). Effects of different practice conditions on advanced instrumentalists' performance accuracy. *J Res Music Educ*, **36**, 250–257.

Ross, S.L. (1985). The effectiveness of mental practice in improving the performance of college trombonists. *J Res Music Educ*, **33**, 221–230.

**Rubin-Rabson, G.** (1937). The influence of analytical pre-study in memorizing piano music. *Arch Psychol*, **20**, 3–53.

**Schellenberg, E.G.** (2004). Music lessons enhance IQ. *Psychol Sci*, **15**, 511–514.

**Semjen, A. and Ivry, R.** (2001). The coupled oscillator model of between-hand coordination in alternate-hand tapping: a reappraisal. *J Exp Psychol: Human Percept Perform*, **27**, 251–265.

**Shaw, W.** (1938). The distribution of muscular action potentials during imaging. *Psychol Rec*, **2**, 195–216.

**Sloboda, J.A., Clarke, E.F., Parncutt, R., and Raekallio, M.** (1998). Determinants of finger choice in piano sight-reading. *J Exp Psychol: Human Percept Perform*, **24**, 185–203.

**Sloboda, J.A., Davidson, J.W., Howe, M.J.A., and Moore, D.G.** (1996). The role of practice in the development of performing music. *Br J Psychol*, **87**, 287–309.

**Theiler, A.M. and Lippman, L.G.** (1995). Effects of mental practice and modeling on guitar and vocal performance. *J Gen Psychol*, **122**, 329–343.

**Visentin, P. and Gongbin, S.** (2003). The kinetic characteristics of the bow arm during violin performance: an examination of internal loads as a function of tempo. *Med Prob Perform Artists*, **18**, 91–97.

**Wollner, C. and Williamon, A.** (2004). Expressive timing and intensity profiles in mental performances. In: S.D. Lipscomb, R. Ashley, R.O. Gjerdingen, and P. Webster (eds) *Proceedings of the International Conference on Music Perception and Cognition*. Causal Productions, Adelaide.

Chapter 4

# Musical synchronization

Bruno H. Repp

The ability to coordinate rhythmic movement with rhythmic sound is fundamental to music performance and dance. This chapter begins by summarizing some of the most important findings about the synchronization of simple movements with simple rhythmic sequences, which constitutes the bulk of synchronization research. Three musical scenarios are then distinguished: (1) synchronization of simple movements with complex sequences (e.g. tapping to the beat of music), (2) synchronization of complex movements with a simple sequence (e.g. playing music with a metronome), and (3) synchronization of complex movements with a complex sequence (e.g. playing music in an ensemble). Research on each of these topics is reviewed, and possible future directions for research are considered.

## Introduction

Sensorimotor synchronization (SMS), the coordination of rhythmic movement with rhythmic sensory stimuli, is a universal human skill that is fundamental to music and dance (Merker 1999–2000). Musical performance in a group requires the precise coordination of sound-producing movements with the sounds and movements produced by other musicians. Dance involves the coordination of (typically, quiet) movements with music and/or with the movements of other dancers. Some forms of dance, such as tap dancing and marching, also involve production of sound. Solo performers sometimes practice with a metronome, which is a form of SMS. Listeners to music and observers of dance may tap or sway along, and if so they are engaging in SMS.

Even if listeners are sitting still, their internal attentional rhythms may be entrained to the sensory input (Large and Jones 1999). Although this is an instance of synchronization and may even involve imaginary actions, it is not an instance of SMS because there is no overt movement. A solo pianist must coordinate the two hands and the fingers of each hand to achieve synchrony between a melody and its accompaniment, and between the tones of chords (Palmer 1996; Repp 1996). Although this involves both movement and sensory feedback, it is not an instance of SMS because there is no external rhythm that serves as a referent.

These examples delineate the subject matter of the present chapter, which is SMS in musical contexts. Most studies of SMS in the literature make no pretense of being concerned with music. They employ simple movements, typically finger tapping, and simple stimulus sequences, typically a metronome. Nevertheless, their results are relevant to music because they reveal the basic mechanisms underlying SMS. The next section summarizes the most important results of that research and gives selected references. (For a more detailed review and more complete references see Repp (2005c).)

## Synchronization of simple movements with simple sequences

### The mean negative asynchrony

One of the oldest findings in SMS research is that during tapping in synchrony with an auditory metronome the taps tend to precede the tones by a few tens of milliseconds (Miyake 1902). This is called the anticipation tendency or mean negative asynchrony (MNA). (An asynchrony is defined to be negative when the tap precedes the tone, positive otherwise.) It has been observed in nearly all simple SMS studies in the literature. Participants are generally unaware of this tendency. When they are trained by means of visual feedback to achieve a mean asynchrony close to zero, they feel they have to delay their taps in order to achieve this goal (Aschersleben 2003). Thus, the MNA corresponds to the subjective point of synchrony. In a classic study, Fraisse and Voillaume (1971) tricked unsuspecting participants in an SMS task by making the metronome tones contingent on their taps, thereby eliminating the MNA. The participants steadily accelerated their tapping tempo, as if trying to re-establish the MNA.

Many studies have been concerned specifically with the MNA because it seems to tell us something important about how perception and action are coordinated in time (see Aschersleben (2002) for a review). Various explanations for the MNA have been suggested, including different neural transmission times for different kinds of sensory information (Fraisse 1980), different central accumulation times for different amounts of sensory information (Aschersleben *et al.* 2004), an asymmetric cost function for asynchronies (Vos and Helsper 1992), minimization of variance through shortening of the internal tapping period (Vorberg and Wing 1996), and perceptual underestimation of empty intervals between tones (Wohlschläger and Koch 2000). Although each explanation has some merit, none of them seems to account for the full range of empirical findings.

Several of these findings suggest that the MNA is largely specific to simple SMS tasks and therefore not particularly relevant to musical SMS. First, the MNA decreases when auditory feedback (i.e. a tone) is made contingent on each

tap (Aschersleben and Prinz 1995), so that asynchronies can be gauged within the auditory modality. While such explicit auditory feedback is uncommon in SMS studies, it is the standard situation in music performance. In large part, the MNA seems to be a consequence of comparing auditory input with tactile and proprioceptive feedback from the movement. Second, the MNA tends to disappear when additional tones are inserted between the metronome tones or when additional movements intervene between taps (Wohlschläger and Koch 2000). This, too, occurs commonly in music, where rhythms are usually more complex than a simple beat. Third, musically trained individuals generally seem to show smaller MNAs than untrained individuals (Aschersleben 2002).

Nevertheless, the tendency to anticipate external referents may account for some phenomena observed in music performance, such as rushing when playing in an ensemble or coming in too soon after a rest.

## Variability

Sensorimotor synchronization inevitably involves uncontrolled variability. Not only are movements inherently variable, but also the perception of a perfectly regular metronome is subject to variability arising in the neural circuitry. Wing and Kristofferson (1973) proposed an influential model according to which the variance of rhythmic movement has two additive components: one generated by a central timekeeper and the other contributed by more peripheral motor processes. Timekeeper variance increases with interval duration, whereas motor variance is relatively constant and thus makes less and less of a contribution to the total variance as the tempo decreases (Wing 1980). Perceptual variability likewise increases with interval duration and thus causes an increase in the perceptual threshold for changes in duration, which approximately follows Weber's law between about 300 and 2000 ms (Friberg and Sundberg 1995). Within these limits, a typical value for the coefficient of variation (the standard deviation as a percentage of the mean interval duration) of both movement timing and perception of timing is 4%, although musicians can show values less than half as large. The Wing–Kristofferson model was proposed to account for unpaced tapping, but it can easily be extended to SMS by adding error correction (see section on Error correction below).

Models of SMS typically assume variability to be random (i.e. white noise), but there is now strong evidence that biological noise does not have a flat spectrum but rather an increasing energy at low frequencies ($1/f$ or fractal noise; Gilden 2001). In unpaced tapping, this is reflected in fluctuations in tempo which occur at many time scales (Madison 2004). In SMS, such fluctuations in tempo are much smaller because of error correction with reference to the pacing sequence, but the asynchronies fluctuate in a similar manner (Chen *et al.* 1997,

2002). This is believed to reflect the workings of a complex dynamic system underlying performance.

## Rate limits

Sensorimotor synchronization has both upper and lower rate limits, which roughly correspond to the interval durations beyond which Weber's law does not hold and may indeed be a consequence of a rapid increase in relative variability. The slowest rate at which SMS is possible corresponds to interval durations of about 2 s (Mates *et al.* 1994; Engström *et al.* 1996). At intervals longer than 2 s it becomes difficult to anticipate future events, and taps tend to follow rather than precede metronome tones. It is likely that this limit can be extended somewhat by mental subdivision strategies, such as musicians are likely to employ, but this has not yet been investigated.

The fastest rate at which synchronization with a metronome is possible depends strongly on the manner of synchronization. At least three different rate limits can be distinguished. When the task is to make one tap with each tone, the rate limit is set simply by the fastest rate at which the finger can move, which corresponds to inter-tap intervals of 150–200 ms for most people (Peters 1980). An even faster limit is evident when the task is to tap only with selected metronome tones, such as every fourth tone (Repp 2003) or only a single tone (Bartlett and Bartlett 1959), so that the biomechanical limit is avoided. In that case, synchronization is possible up to interval durations of 100–125 ms between tones. The eventual breakdown in coordination (which is evident as phase drift) suggests that the tones can no longer be attended to as individual events at such fast rates. A much slower limit is obtained when the task is off-beat (anti-phase) tapping, that is to tap at the midpoints of the intervals between tones. That limit is typically near interval durations of 350 ms (Repp 2005*a*, *b*). These values are for musically trained individuals; musically untrained participants may show considerably higher values.

The greater difficulty of anti-phase than in-phase coordination has been of particular interest to researchers who take a dynamic systems approach to SMS (e.g., Kelso *et al.* 1990). Their research usually focuses on continuous oscillatory limb movement, not on tapping. When the tempo is increased, a phase transition from anti-phase to in-phase coordination is commonly observed, which reflects a difference in the relative stability of these coordination modes. This difference is only evident when the system is stressed, however. Anti-phase tapping with a metronome tends to be *less* variable than in-phase tapping when the interval duration exceeds 400 ms (Semjen *et al.* 1992) because the taps subdivide the intervals, and shorter intervals afford lower variability as long as they exceed 200 ms (Repp 2003).

Empirical findings such as the ones reviewed here support London's (2002, 2004) contention that 200 ms is about the shortest possible interval between beats in music (cf. van Noorden and Moelants 1999), whereas 100 ms is the shortest possible interval between subdivisions of a beat.

## Error correction

Sensorimotor synchronization cannot be achieved and sustained without error correction. Error correction prevents the inherent variability of the movement from having a cumulative effect on the times of occurrence of successive actions (Vorberg and Wing 1996). Two distinct error correction processes have been hypothesized to exist: phase correction and period correction (Mates 1994). Phase correction adjusts the time of occurrence of each successive action but does not affect the period of the underlying timekeeper or oscillator that governs the rate of the action. Period correction adjusts the interval of the timekeeper or period of the oscillator. A conceptual parallel can be seen between the two sources of variability postulated by the Wing–Kristofferson model and the two forms of error correction: In each case, a more central process and a more peripheral process are distinguished.

Both error correction processes occur against a background of uncontrolled variability, so that they cannot be observed easily in raw data. Two general approaches have been taken to investigate error correction: mathematical modelling and perturbation methods. The first approach estimates error correction parameters from raw time-series data (asynchronies and/or inter-tap intervals), whereas the second approach manipulates the timing of the pacing sequence and averages data across repeated trials in which the same timing perturbation occurs, thereby reducing the noise in the data.

Each error correction process is assumed to be based on sensory information about the temporal relationship between the taps and the pacing sequence. Phase correction is usually assumed to be based on perception of asynchronies. A linear phase correction model (Mates 1994; Pressing 1998; Vorberg and Schulze 2002) assumes that each tap is adjusted by a fixed proportion of the asynchrony generated by the preceding tap. This model has been quite successful in accounting for synchronization with a metronome, where there is little reason for period correction to occur. Period correction is assumed to be based on perception of a mismatch between the period of the internal timekeeper and the intervals of the pacing sequence, and the correction is assumed to be a linear function of that mismatch (Mates 1994). There are alternative ways of conceptualizing the perceptual information on which phase and period corrections are based, but discussion of this issue would lead too far here (see Hary and Moore 1985,

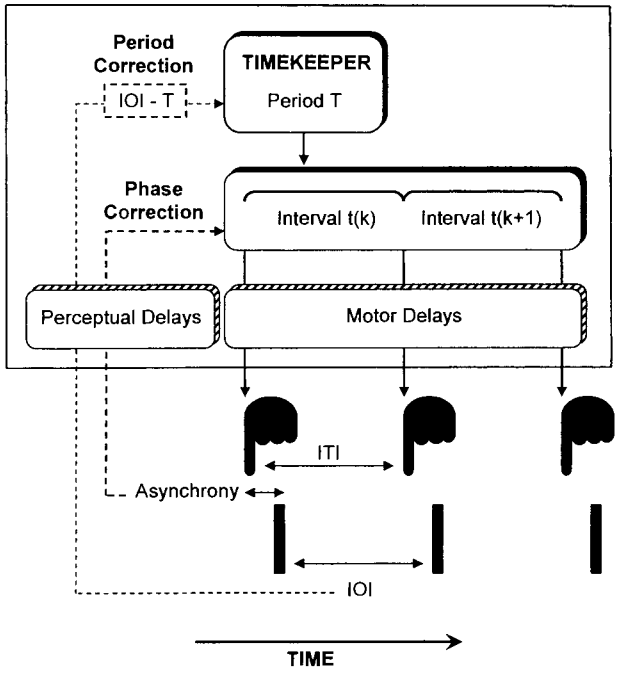



**Fig. 4.1** Schematic diagram of the two-process error correction model for sensorimotor synchronization (after Mates 1994). A central timekeeper with period *T* is assumed to control the intervals [*t*(*k*), *t*(*k* +1), . . .] between successive motor commands. The period *T* can be adjusted on the basis of perceived discrepancies between its duration and the tone inter-onset interval (IOI) in the sequence (*period correction*). The intervals [*t*(*k*), *t*(*k* + 1), . . .] can further be adjusted individually on the basis of registered asynchronies between taps and tones (*phase correction*). The frame indicates internal processes; only inter-tap intervals (ITIs) and asynchronies are directly observable. They are affected by both period correction and phase correction. (From Repp and Keller 2004.) Reproduced with permission of The Experimental Psychology Society (http://www.psypress.co.uk/journals.asp)

1987; Repp 2005c). The dual-process model of error correction is illustrated schematically in Fig. 4.1.

Phase correction and period correction are hypothetical *internal* processes that affect the observable behaviour, namely the adaptation of the relative tapping phase (asynchronies) and tapping period (inter-tap intervals) to the pacing sequence. Each internal process affects *both* observable variables (Repp 2001*b*), and therefore the processes are difficult to separate when they occur in tandem. For that reason, statistical models have focused mainly on phase correction in SMS with a metronome, where period correction can be assumed to be minimal. Timing perturbations in the pacing sequence, however, may elicit both phase and period correction, most obviously when the perturbation is a persistent tempo change. An estimate of period correction may be obtained by stopping the pacing sequence at various points and asking participants to continue tapping at what they consider the final tempo (Repp and Keller 2004).

The period of continuation tapping provides a retrospective estimate of the period of the internal timekeeper at the point where synchronization changed to self-paced tapping.

There is a good deal of evidence that phase correction is relatively automatic whereas period correction is in large part consciously controlled. Phase correction is the process that maintains synchrony with a metronome, but people are rarely aware of the continuous corrective adjustments they are making. When a phase perturbation is introduced in the metronome sequence, phase correction occurs in response to the perturbation regardless of whether or not the perturbation itself or the resulting asynchrony is detectable (Repp 2000, 2001*a*). In fact, phase correction seems to be most effective in response to small perturbations. Phase correction is also unaffected by diversion of attention (Repp and Keller 2004). When participants are instructed not to react to a detectable perturbation, they can reduce their phase correction response but are unable to suppress it completely (Repp 2002*a*; Repp and Keller 2004). By contrast, period correction in response to a tempo change in the pacing sequence is more effective when the tempo change is detectable than when it is not (Repp 2001*b*), is affected negatively by diversion of attention, and can be suppressed voluntarily (Repp and Keller 2004). Thus, whereas phase correction is a relatively low-level process that serves to stabilize SMS, period correction is a higher-level strategic process that may be involved not only in SMS but in other activities that require modulation of an internal period, such as expressive timing in music performance. Most likely, phase correction and period correction also engage different neural networks (Stephan *et al.* 2002; Praamstra *et al.* 2004), although much more needs to be learned about their neural underpinnings. A reasonable hypothesis based on existing data is that phase correction involves the cerebellum whereas period correction involves a corticothalamic circuit.

We are now ready to consider studies of SMS in situations that are somewhat closer to music performance. The next three sections divide these studies into three groups, according to the relative complexity of the pacing sequence and/or the motor activity. We first review studies of simple tapping to complex rhythms (i.e. more complex than a metronome) or to music. Then we discuss the performance of complex rhythms or of music in synchrony with a metronome. Finally, we consider tasks in which both the motor activity and the pacing sequence are complex, as they are in performance in a musical ensemble.

## Synchronization of simple movements with complex sequences

### Tapping a beat

One of the most commonly observed forms of spontaneous SMS is tapping a foot to the beat of music. Not surprisingly, therefore, isochronous tapping is frequently

used as a way of determining the period and phase of perceived metrical beat(s) in rhythmic patterns or in music (e.g. Oshinsky and Handel 1978; Handel and Oshinsky 1981; Handel and Lawson 1983; Parncutt 1994; Jones and Pfordresher 1997; van Noorden and Moelants 1999; Drake *et al.* 2000). However, these *beat-finding* studies are usually concerned only with the approximate position of the taps, not with asynchronies, variability, or error correction. Their main goal is not to study SMS but simply to locate the beat in the rhythmic structure, and often also to compare human performance with that of a computer algorithm. A few recent studies, however, have looked at synchronization with a beat in some detail.

Snyder and Krumhansl (2001) required participants to tap to the beat of piano ragtime music which was played under computer control with metronomic timing and equalized tone amplitudes. The sequences presented were either the original musical excerpts or only the highly syncopated right-hand part of the piano score, and each version was also reduced to a monotone pitch. The pitch manipulation had no effect on synchronization, but variability was greater when participants tapped with the right-hand part alone than when they tapped with the combined left- and right-hand parts. There was no significant MNA, which is consistent with the finding of Wohlschläger and Koch (2000) that tones interpolated between beats eliminate the MNA.

Large (2000) modelled the Snyder and Krumhansl data successfully with a computer algorithm based on a network of coupled oscillators. The oscillators were coupled not only to each other but also to the auditory input (by means of a coupling function in the model), and this resulted in the entrainment of an appropriate oscillator to the beat. Coupling to input is a dynamic systems implementation of phase correction (cf. Pressing 1999). Period correction was not included in Large's model; instead, a large number of oscillators with fixed frequencies were employed, so that switching between oscillators with similar periods was equivalent to period correction.

Toiviainen and Snyder (2003) conducted a beat tapping study in which excerpts from pieces by J. S. Bach in 4/4 meter were presented in three versions: integral, left hand only, or right hand only. Some excerpts started on a downbeat, whereas others started on the second beat. Again, no MNA was found; in fact, the taps lagged slightly behind the beat. Tapping variability was slightly greater for the separate parts than for the integral version, and slightly greater when participants tapped on the third beat of each bar than when they tapped on the first beat (the nominal downbeat). Whether participants tapped on the first beat or on the third beat depended in large part on whether the excerpt started on a downbeat or on the second beat; in the latter case, the second beat was usually interpreted as an upbeat (i.e. as the fourth beat in the measure).

A recent study by Patel *et al.* (2005) was not concerned with beat finding but with tapping variability as a function of metrical structure when the period and

phase of the beat were prescribed. Strongly metrical (SM) and weakly metrical (WM) rhythms were constructed by concatenating monotone sequences used previously by Povel and Essens (1985). Povel and Essens had shown that SM rhythms induce the feeling of a strong beat, whereas in WM rhythms a beat is more difficult to perceive. The main structural difference between the two is that SM sequences have events in every beat location, whereas in WM sequences some beat events are missing. Patel *et al.* prefixed each type of sequence with an isochronous induction sequence marking the beat period. Participants started tapping with the induction sequence and continued tapping the beat after the rhythmic pattern started. The results showed that taps were significantly more variable in WM than in SM sequences.

Several other conditions were included in that study. In one condition, participants simply tapped with an isochronous sequence of beats. Variability in that condition was similar to that of tapping with SM sequences. Thus, the variously subdivided beats of the SM sequences did not stabilize tapping any more than did a simple sequence of beats. However, tapping was less variable when participants tapped with every other tone of an isochronous sequence that was twice as fast. Thus, simple regular subdivision did improve the precision of SMS (cf. Repp 2003). In another condition, participants tapped with a sequence that had been generated from an isochronous beat sequence by deleting the beats that corresponded to missing beats in the WM sequences. Taps were more variable in that condition than in tapping with an isochronous sequence of beats, which suggests that the greater variability of tapping with WM sequences than with SM sequences was due to the absence of a number of beat events in the WM sequences. Indeed, when—in yet another condition—the missing beats in WM sequences were filled in, the tapping results were just like those with SM sequences because the manipulation had turned the WM into SM sequences. The importance of beat events can be understood by considering that taps falling on empty beats do not generate any error feedback (i.e. no asynchrony). Therefore, there is no phase correction following such taps (cf. Repp 2002*c*), and this increases the variability of the subsequent tap and thereby increases the overall variability of the asynchronies.

Similarly, when a tone occurs without an accompanying tap—a likely occurrence in tapping to the beat of a complex rhythm—there is no asynchrony at that point. Nevertheless, such subdivision tones are monitored perceptually and do affect phase correction (Large *et al.* 2002). Intervening subdivision tones reduce the phase correction response to a perturbation of the beat, and they elicit a shift of the next tap when their own timing is perturbed (Repp 2002*d*). Thus, additional structure in a pacing sequence (which could be music) is readily used for guiding simple movements, but additional movements (which could be those of music making) do not necessarily help synchronization with a simple beat

(although they do reduce the MNA in non-musicians, as Wohlschläger and Koch (2000) have shown). The latter situation requires more detailed study, however.

## On-beat versus off-beat tapping with rhythms

The task in beat tapping studies is generally to tap in phase with the perceived beat, although occasionally participants may prefer to tap in anti-phase. In music performance, however, various other forms of coordination can occur, such as playing a note on the second and/or third beats of a rhythm in 3/4 meter, or on the second and/or fourth beats of a rhythm in 4/4 meter. Such tasks have been investigated only very recently in the laboratory (Repp 2005 *a*, *b*).

In these studies, simple monotone rhythms were presented, consisting of cyclically repeated groups of two or three tones, with the between-group interval being twice as long as the within-group interval(s). Participants were asked to tap with a particular tone in each cycle or in the middle of the 'gap' (i.e. the long interval). In several experiments, the dependent variable was the fastest tempo at which synchrony could be maintained. The results showed that it was easier to tap with a group-initial or group-final tone (which, according to Povel and Essens (1985) carry a rhythmic grouping accent and therefore are likely carriers of the metrical beat) than with a group-medial tone or in a gap. When one of the tones was accentuated by means of higher intensity and pitch, it was easier to tap with that tone than with other tones. When the (musically trained) participants were instructed to intentionally 'hear' the metrical beat on a particular tone while tapping in various positions, this dissociation of the metrical beat from the taps had surprisingly little effect on task difficulty as measured by the rate limit of SMS. In other words, participants found it just as easy to tap off the beat as on the beat, as long as they were tapping in the same group position. However, participants found it quite difficult to dissociate the metrical beat from their taps at fast tempi. To facilitate this dissociation, another experiment used slower tempi and focused on the variability of the asynchronies. Metrical interpretation still had no effect. Surprisingly, however, the effect of accentuation also went away, and the effect of grouping accent changed as well. Clearly, further research along these lines is needed, as the pattern of results is more complicated than expected.

## Tapping to slow music: effects of musical structure

In the research considered so far in this section, the number of tones in the rhythm typically exceeded the number of taps made. In other words, the beats were subdivided in various ways. Several studies have investigated instead tapping in synchrony with the lowest metrical level of very slow music, so that there were few or no subdivisions between the tones that served as synchronization targets. This is

like tapping with a metronome, except that the pacing sequence is real piano music. The question of interest is whether and how the musical structure (melody, harmony, meter, perhaps even expressive features) affects the timing of the taps.

Repp (1999*a*, *b*) asked participants to tap along with the sixteenth notes in the first few measures of Chopin's Etude in E major, a very slow piece. The music was performed metronomically under computer control. Interestingly, the asynchronies and inter-tap intervals, when averaged across a number of trials and participants, exhibited a systematic timing pattern, even though the music was in strict time. Thus it appears that some aspects of musical structure have an involuntary influence on the precise timing of coordinated movements. A later study (Repp 2002*b*) found exactly the same pattern of tap timing when participants synchronized their taps with an isochronous sequence of clicks that was superimposed on the music (Fig. 4.2b), and even when the participants merely imagined the music in synchrony with the clicks, although the variations were smaller in that case (Fig. 4.2c). When synchronizing with the clicks without imagining music, the average inter-tap intervals were essentially constant (Fig. 4.2a).

The original hypothesis leading to this research (Repp 1999*a*) was that the patterning of the taps would reflect expectations of expressive timing. However, there was only a weak resemblance between the timing pattern of the taps and the typical expressive timing profile of the music. The timing of the taps seemed to reflect more the metrical structure of the music, whereas expressive timing primarily reflects the melodic–rhythmic grouping structure. Similar conclusions were reached in a study that investigated tapping with two complete Preludes by Chopin (Repp 1999*c*). However, the structural determinants of tap timing in this situation have not been fully unravelled and are in need of further investigation.

Repp (1999*b*, 2002*b*) used the same Chopin Etude excerpt to investigate participants' ability to synchronize with expressively timed music. In the earlier study (Repp 1999*b*), the participants were pianists who had just played the excerpt on a digital piano. When they were asked to tap along with one of their own recorded performances or with a synthesized performance having a typical expressive timing pattern, they were clearly able to anticipate the expressive timing variations, but not perfectly: large deviations from regularity were usually underestimated, and the resulting asynchrony had to be corrected on the next tap(s). The later study (Repp 2002*b*) compared synchronization of taps with three kinds of sequences: expressively timed music; a monotone sequence with the same timing pattern; and music with an atypical or structurally inappropriate (phase-shifted) expressive timing pattern. Synchronization with the expressively timed music was relatively most accurate, and interestingly this advantage was already evident on the first trial. Thus, hearing the musical structure helps predict appropriate variation of expressive timing.

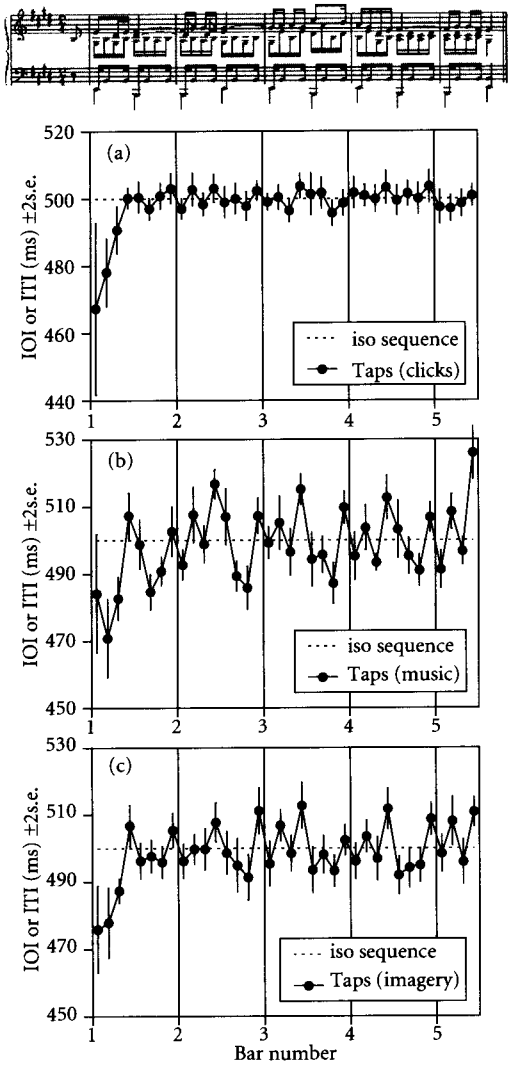



**Fig. 4.2** Average inter-tap interval (ITI) profiles with double standard errors in three synchronization conditions: (a) tapping with an isochronous (iso) sequence of clicks; (b) tapping with the same sequence of clicks accompanied by isochronous music; (c) tapping with the same sequence of clicks while merely imagining the music in synchrony with the clicks. The initial three ITIs reflect the 'tuning in' to the sequence tempo. (From Repp 2002*b*.)

Synchronization with all three sequence types improved with practice (Repp 2002*b*). Initially, participants' prediction of the timing variation was relatively poor, and this was reflected in a large lag-1 correlation and a small lag-0 correlation between the sequence inter-onset intervals and the inter-tap intervals. In other words, participants mainly reacted to the variation in timing, tracking it rather than predicting it. In the course of 10 trials, however, the lag-1 correlation gradually decreased and the lag-0 correlation increased, which reflects increasing prediction. These changes can be interpreted in terms of the two error correction mechanisms, phase correction and period correction (see section on Error

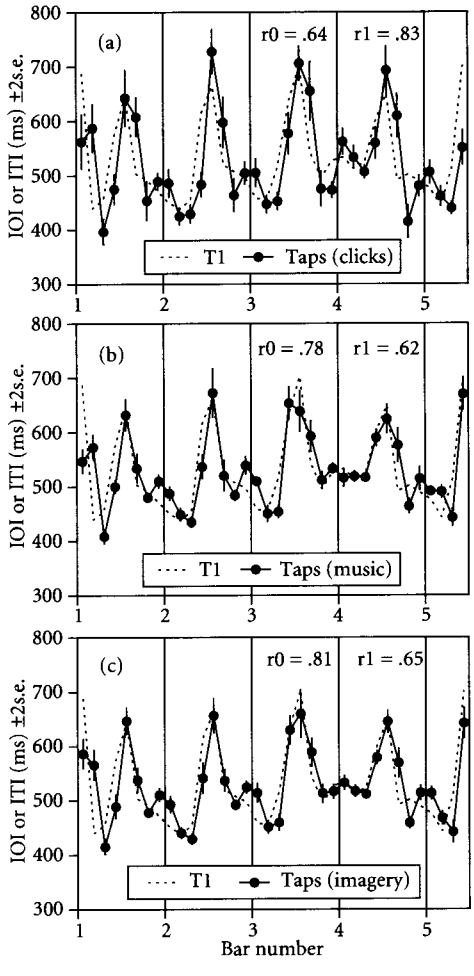



**Fig. 4.3** Average inter-tap intervals (ITIs) with double standard errors in three synchronization conditions: (a) tapping with a sequence of clicks whose inter-onset intervals (IOIs) vary according to an 'expressive' timing pattern (T1); (b) tapping with the same sequence of clicks accompanied by music having the same (fairly typical) expressive timing pattern; (c) tapping with the same sequence of clicks while merely imagining the music in synchrony with the clicks. The lag-0 and lag-1 correlations between the IOIs and ITIs are also given (r0 and r1, respectively). Note the increase in r0 and the decrease in r1 in the latter two conditions relative to the first condition. (From the electronic appendix to Repp 2002*b*.)

correction above). Phase correction is the automatic reactive mechanism that causes tracking, whereas period correction is the predictive mechanism. Expectation and active prediction of a learned timing pattern presumably involve an intentional modulation of the period of the internal timekeeper. As learning proceeds, prediction via period correction improves, so that there are smaller asynchronies for phase correction to react to.

Figure 4.3 illustrates the relative accuracy of synchronizing with an expressive timing pattern that is fairly typical for performances of the excerpt from the Chopin Etude. Inter-tap intervals averaged across a number of trials and participants are shown for three conditions: synchronization with (a) a sequence of clicks with 'expressive' timing but without accompanying music, (b) the clicks accompanied by the synchronous music, and (c) the clicks while the music is merely being

imagined. The target timing pattern is shown as the dotted line. Note that the initial underprediction of large deviations is smaller in the second and third conditions than in the first condition, and this is reflected in an increased lag-0 correlation (prediction) and a decreased lag-1 correlation (tracking).

### Marching and dancing to music

There have been very few systematic studies of activities other than finger tapping in synchrony with music. It can be safely assumed, however, that the basic principles revealed by finger-tapping studies also apply to other forms of synchronized movement, especially if they involve periodic contacts with a surface. That is true for the feet in both marching and dancing, whereas the arms and the body move more freely. Direct comparisons between continuous (free) and discontinuous (contact) movements in an SMS task remain to be conducted.

In this context, it should be noted that music has been used in a clinical setting to improve the gait patterns of stroke victims and patients with Parkinson's disease (see Thaut *et al.* (1999) for a summary). These patients readily entrained their steps to the music and derived greater benefits from rehabilitation training than did control groups. Thaut *et al.* hypothesized that there is a direct pathway from the auditory system to spinal motor neurons, which enables rhythmic input to drive motor action.

## Synchronization of complex movements with simple sequences

### Playing music in synchrony with a metronome

Playing music with a metronome is a task familiar to all music students, especially pianists. Surprisingly, however, it has not been studied in detail. In what seems to be the only study in the literature, Repp (1999*c*) asked pianists to play two Chopin Preludes in synchrony with a metronome and compared the results to playing without a metronome but in strict time (i.e. without any expressive timing). The main difference was that the correlation between successive inter-beat intervals was negative when playing with a metronome but positive when playing merely in strict time. The Wing–Kristofferson model predicts a negative lag-1 autocorrelation in self-paced rhythm production. The reason why a positive correlation was obtained instead is that there is residual expressive timing variation in performances intended to be in strict time (Penel and Drake 1998; Repp 1999*b*), and this systematic variation entails positive covariation of successive intervals. The emergence of a negative correlation in the metronome-paced condition was attributed to phase correction in synchronization. In both conditions, the timing of beats was much more regular than in normal expressive performance.

Interestingly, however, the key depressions between beats (i.e. rhythmic subdivisions) showed large deviations from regularity that were hardly affected by either synchronization with a metronome or by the mere intention to play in strict time. Thus, regularity at the beat level does not imply more regular timing at a lower metrical level.

## Tapping rhythms in synchrony with a metronome or rhythm template

A recent study (Repp *et al.* 2005) investigated the tapping of 'uneven' rhythms at various tempi. In such rhythms, the inter-beat intervals are not divided evenly. The rhythms included three sets: 2 + 3, 3 + 2; 2 + 2 + 3, 2 + 3 + 2, 3 + 2 + 2; 2 + 3 + 3, 3 + 2 + 3, 3 + 3 + 2. The distinction among the rhythms within each set rested solely on which element was considered as the metrical (down)beat. The rhythms were produced in four conditions: self-paced; synchronized with a metronome, such that the downbeat taps coincided with the tones; synchronized with a precise rhythm template; and synchronized with a rapid isochronous sequence of subdivisions. The last condition served to determine the rate limit of synchronization (see section on Variability above), which was found to be considerably higher than for tapping with every fourth tone of a rapid sequence (Repp 2003), presumably due to the greater cognitive demands of uneven rhythms.

The conditions of main interest here are the other two synchronization tasks. There were four noteworthy findings. First, a characteristic distortion of the 3:2 interval ratios in the direction of 2:1 was equally present in free tapping and in synchronization with a precise rhythm template. In other words, synchronization did not increase the precision of rhythm production. Second, in tapping with a template, the downbeat tap was not synchronized more accurately than the subdivision taps. Rather, the taps generally seemed to be timed so as to minimize the mean asynchrony per rhythm cycle. Third, the asynchronies of the downbeat taps were very different in synchronization with a simple metronome than in synchronization with a rhythm template. In the former condition, they were the only taps that generated asynchronies and thus were the only asynchronies to be minimized, whereas in the latter condition, two or three asynchronies per cycle had to be balanced. Finally, which of the rhythm elements was considered the downbeat had little effect on timing and synchronization. However, the position of a tap in a rhythmic group had a strong effect on its asynchronies with a pacing sequence.

Franêk *et al.* (1988) also found that participants could not synchronize very accurately with cyclically repeated rhythms when they embodied complex interval ratios. Various other studies (e.g. Semjen and Ivry 2001; Semjen

and Vos 2002) have used a metronome to pace bimanual tapping or other periodic bimanual movements having complex interval ratios. However, those studies focused on the bimanual interaction, which is beyond the scope of this chapter.

## Synchronization of complex movements with complex sequences

### Playing music in an ensemble

One important topic still in need of investigation is synchronization between two individuals, even with a task as simple as tapping simultaneously or in alternation. In such a situation, each participant provides the pacing sequence for his or her partner, and phase correction (as well as perhaps period correction) will occur in both sequences. Preliminary work on such dyadic interactions has been reported by Himberg and Cross (2004) for finger tapping and by Boker and Rotondo (2003) for dancing.

Rasch (1979) measured the mean 'asynchronization' occurring between the players of three professional trios playing classical music. Asynchronization was defined as the 'root mean square of the standard deviation of the onset difference times for all pairs of voices' (p. 123) and was found to range from about 30 to 50 ms. Synchronization was better at fast than at slow tempi. Shaffer (1984) used Rasch's measure to examine the coordination between two pianists in duet playing and found asynchronization of similar magnitude, despite considerable expressive timing variation. Evidently, the pianists were able to predict each other's expressive timing to a considerable extent. In both studies, however, the musicians were able to take advantage of visual as well as auditory cues for coordination.

A recent study (Keller *et al.* in press) examined synchronization accuracy in pianists playing duets with recorded performances of one part of the duet, so that visual cues were not available. The main question was whether pianists would be more accurate synchronizing with performances recorded by themselves than with performances recorded by other pianists. This was indeed found to be the case, although the difference was small. Keller *et al.* attributed it to implicit action knowledge which enables players to predict their own expressive timing patterns better than those of other players.

Friberg and Sundström (2002) measured the relative timing of tone onsets in excerpts from recorded performances by great jazz musicians. They found that the 'swing ratio', a characteristic long–short pattern, was more pronounced in the playing of drummers than in soloists. This implies asynchronies between soloist and drummer. In contrast to the finding by Repp *et al.* (2005) that the mean asynchrony per cycle tends to be minimized when synchronizing a rhythm with a

template, Friberg and Sundström found that soloists lagged behind drummers on the downbeat but were synchronized precisely on the off-beat. This may reflect an expressive strategy of 'playing behind the beat' as well as the structural importance of off-beats in jazz. Other preliminary observations about synchronization in jazz performance have been reported by Prögler (1995) and Schögler (1999–2000).

Ensemble playing requires not only in-phase synchronization with other players but also precise alternation (anti-phase coordination) and other, more complex, forms of rhythmic coordination, as in African drumming or Balinese gamelan. Keller (1999, Keller and Burnham 2005) required percussionists to produce a memorized rhythm concurrently with a different, computer-controlled rhythm in the course of a study of 'prioritized integrative attending' (i.e. attending to one's own part as well as to the integrated rhythm of several parts (Keller 2001)). Synchronization accuracy was assessed only in a qualitative way. More studies of this degree of sophistication and realism are needed to close the gap between laboratory SMS research and actual music performance.

### Following a conductor's gestures

Another interesting area awaiting investigation is synchronization of music performance with a conductor's movements, which are often far more complex than a simple beat pattern. Some preliminary work has been conducted by Luck (2002, 2004). He recorded the arm movements of several conductors when they gave a downbeat, reproduced the movement trajectory on a computer screen, and asked participants to make a single tap when they perceived the downbeat as occurring. By manipulating the shape and velocity pattern of the trajectory, he determined that velocity and acceleration were important in conveying a downbeat, but the relative importance of different variables varied from conductor to conductor.

## Conclusions

A large number of SMS studies using simple movements and simple auditory sequences has laid the theoretical and empirical groundwork for a more detailed study of the synchronization skills required in music performance, dance, and some athletic disciplines. The SMS research in these latter areas is still very limited. The present chapter (which makes no claim to be comprehensive) has mentioned some relevant studies and has pointed out topics awaiting more careful study. Clearly, here is a potentially rich area for research, the results of which will be of both theoretical interest and of practical relevance.

## Acknowledgements

During preparation of this chapter the author received support from NIH grants MH-51230, DC-03663, and HD-01994.

## References

Aschersleben, G. (2002). Temporal control of movements in sensorimotor synchronization. *Brain Cogn*, **48**, 66–79.

Aschersleben, G. (2003). Effects of training on the timing of repetitive movements. In: S.P. Shohov (ed.) *Advances in psychology research*, Vol. 23, pp. 15–30. Nova Science, Huntington, NY.

Aschersleben, G. and Prinz, W. (1995). Synchronizing actions with events: the role of sensory information. *Percept Psychophys*, **57**, 305–317.

Aschersleben, G., Gehrke, J., and Prinz, W. (2004). A psychophysical approach to action timing. In: C. Kaernbach, E. Schröger, and H. Müller (eds) *Psychophysics beyond sensation: laws and invariants of human cognition*, pp. 117–136. Erlbaum, Mahwah, NJ.

Bartlett, N.R. and Bartlett, S.C. (1959). Synchronization of a motor response with an anticipated sensory event. *Psychol Rev*, **66**, 203–218.

Boker, S.M. and Rotondo, J.L. (2003). Symmetry building and symmetry breaking in synchronized movement. In: M. Stamenov and V. Gallese (eds) *Mirror neurons and the evolution of brain and language*, pp. 163–171. John Benjamin, Amsterdam.

Chen, Y., Ding, M., and Kelso, J.A.S. (1997). Long memory processes ($1/f^{\alpha}$ type) in human coordination. *Phys Rev Lett*, **79**, 4501–4504.

Chen, Y., Repp, B.H., and Patel, A.D. (2002). Spectral decomposition of variability in synchronization and continuation tapping: comparisons between auditory and visual pacing and feedback conditions. *Hum Move Sci*, **21**, 515–532.

Drake, C., Penel, A., and Bigand, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Percept*, **18**, 1–23.

Engström, D.A., Kelso, J.A.S., and Holroyd, T. (1996). Reaction–anticipation transitions in human perception–action patterns. *Hum Move Sci*, **15**, 809–832.

Fraisse, P. (1980). Les synchronisations sensori-motrices aux rythmes [Sensorimotor synchronization to rhythms]. In: J. Requin (ed.) *Anticipation et comportement*, pp. 233–257. Centre National, Paris.

Fraisse, P. and Voillaume, C. (1971). Les repères du sujet dans la synchronisation et dans la pseudo-synchronisation [The subject's references during synchronization and during pseudo-synchronization]. *L'Année Psychol*, **71**, 359–369.

Franêk, M., Radil, T., and Indra, M. (1988). Tracking irregular acoustic patterns by finger tapping. *Int J Psychophysiol*, **6**, 327–330.

Friberg, A. and Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *J Acoust Soc Am*, **98**, 2524–2531.

Friberg, A. and Sundström, A. (2002). Swing ratios and ensemble timing in jazz performance: evidence for a common rhythm pattern. *Music Percept*, **19**, 333–349.

Gilden, D.L. (2001). Cognitive emissions of $1/f$ noise. *Psychol Rev*, **108**, 33–56.

Handel, S. and Lawson, G.R. (1983). The contextual nature of rhythmic interpretation. *Percept Psychophys*, **34**, 103–120.

**Handel, S. and Oshinsky, J.S.** (1981). The meter of syncopated auditory polyrhythms. *Percept Psychophys*, **30**, 1–9.

**Hary, D. and Moore, G.P.** (1985). Temporal tracking and synchronization strategies. *Hum Neurobiol*, **4**, 73–77.

**Hary, D. and Moore, G.P.** (1987). Synchronizing human movement with an external clock source. *Biol Cybern*, **56**, 305–311.

**Himberg, T. and Cross, I.** (2004). Interaction in musical time. In: S.D. Lipscomb, R. Ashley, R.O. Gjerdingen, and J.P. Webster (eds) *Proceedings of the Eighth International Conference on Music Perception and Cognition* (CD-ROM). Causal Productions, Adelaide.

**Jones, M.R. and Pfordresher, P.Q.** (1997). Tracking musical patterns using joint accent structure. *Can J Exp Psychol*, **51**, 271–290.

**Keller, P.E.** (1999). Attending in complex musical interactions: the adaptive dual role of meter. *Aust J Psychol*, **51**, 166–175.

**Keller, P.E.** (2001). Attentional resource allocation in musical ensemble performance. *Psychol Music*, **29**, 20–38.

**Keller, P.E. and Burnham, D.K.** (2005). Musical meter in attention to multipart rhythm. *Music Perception*, **22**, 629-661.

**Keller, P.E., Knoblich, G., and Repp, B.H.** (in press). Pianists duet better when they play with themselves: On the possible role of action simulation in synchronization. *Cognition and Consciousness.*

**Kelso, J.A.S., DelColle, J.D., and Schöner, G.** (1990). Action-perception as a pattern formation process. In: M. Jeannerod (ed.) *Attention and performance XIII*, pp. 139–169. Erlbaum, Hillsdale, NJ.

**Large, E.W.** (2000). On synchronizing movements to music. *Hum Move Sci*, **19**, 527–566.

**Large, E.W. and Jones, M.R.** (1999). The dynamics of attending: how we track time-varying events. *Psychol Rev*, **106**, 119–159.

**Large, E.W., Fink, P., and Kelso, J.A.S.** (2002). Tracking simple and complex sequences. *Psychol Res*, **66**, 3–17.

**London, J.** (2002). Cognitive constraints on metric systems: some observations and hypotheses. *Music Percept*, **19**, 529–550.

**London, J.** (2004). *Hearing in time*. Oxford University Press, New York.

**Luck, G.** (2002). Conductors' gestures: perception of, and synchronization with, visual beats. In: C. Stevens, D. Burnham, G. McPherson, E. Schubert, and J. Renwick (eds) *Proceedings of the Seventh International Conference on Music Perception and Cognition* (CD-ROM), p. 638. Causal Productions, Adelaide.

**Luck, G.** (2004). The relative contribution of four movement-related variables to the induction of a visual beat. In: S.D. Lipscomb, R. Ashley, R.O. Gjerdingen, and J.P. Webster (eds) *Proceedings of the Eighth International Conference on Music Perception and Cognition* (CD-ROM). Causal Productions, Adelaide.

**Madison, G.** (2004). Fractal modeling of human isochronous serial interval production. *Biol Cybern*, **90**, 105–112.

**Mates, J.** (1994). A model of synchronization of motor acts to a stimulus sequence. I. Timing and error corrections. *Biol Cybern*, **70**, 463–473.

**Mates, J., Radil, T., Müller, U., and Pöppel, E.** (1994). Temporal integration in sensorimotor synchronization. *J Cogn Neurosci*, **6**, 332–340.

**Merker, B.** (1999–2000). Synchronous chorusing and the origins of music. *Musicae Scientiae* (special issue), 59–73.

**Miyake, I.** (1902). Researches on rhythmic activity. *Stud Yale Psychol Lab*, **10**, 1–48.

**Oshinsky, J.S. and Handel, S.** (1978). Syncopated auditory polyrhythms: discontinuous reversals in meter interpretation. *J Acoust Soc Am*, **63**, 936–939.

**Palmer, C.** (1996). On the assignment of structure in music performance. *Music Percept*, **14**, 23–56.

**Parncutt, R.** (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Percept*, **11**, 409–464.

**Patel, A.D., Iversen, J.R., Chen, Y., and Repp, B.H.** (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*, **163**, 226–238.

**Penel, A. and Drake, C.** (1998). Sources of timing variations in music performance: a psychological segmentation model. *Psychol Res*, **61**, 12–32.

**Peters, M.** (1980). Why the preferred hand taps more quickly than the non-preferred hand: three experiments on handedness. *Can J Psychol*, **34**, 62–71.

**Povel, D.-J. and Essens, P.** (1985). Perception of temporal patterns. *Music Percept*, **2**, 411–440.

**Praamstra, P., Turgeon, M., Hesse, C.W., Wing, A.M., and Perryer, L.** (2003). Neurophysiological correlates of error correction in sensorimotor synchronization. *NeuroImage*, **20**, 1283–1297.

**Pressing, J.** (1998). Error correction processes in temporal pattern production. *J Math Psychol*, **42**, 63–101.

**Pressing, J.** (1999). The referential dynamics of cognition and action. *Psychol Rev*, **106**, 714–747.

**Prögler, J.A.** (1995). Searching for swing: participatory discrepancies in the jazz rhythm section. *Ethnomusicology*, **39**, 21–54.

**Rasch, R.A.** (1979). Synchronization in performed ensemble music. *Acustica*, **43**, 121–131.

**Repp, B.H.** (1996). Patterns of note onset asynchronies in expressive piano performance. *J Acoust Soc Am*, **100**, 3917–3932.

**Repp, B.H.** (1999*a*). Detecting deviations from metronomic timing in music: effects of perceptual structure on the mental timekeeper. *Percept Psychophys*, **61**, 529–548.

**Repp, B.H.** (1999*b*). Control of expressive and metronomic timing in pianists. *J Motor Behav*, **31**, 145–164.

**Repp, B.H.** (1999*c*). Relationships between performance timing, perception of timing perturbations, and perceptual-motor synchronization in two Chopin preludes. *Aust J Psychol*, **51**, 188–203.

**Repp, B.H.** (2000). Compensation for subliminal timing perturbations in perceptual-motor synchronization. *Psychol Res*, **63**, 106–128.

**Repp, B.H.** (2001*a*). Phase correction, phase resetting, and phase shifts after subliminal timing perturbations in sensorimotor synchronization. *J Exp Psychol: Hum Percept Perform*, **27**, 600–621.

**Repp, B.H.** (2001*b*). Processes underlying adaptation to tempo changes in sensorimotor synchronization. *Hum Move Sci*, **20**, 277–312.

**Repp, B.H.** (2002*a*). Automaticity and voluntary control of phase correction following event onset shifts in sensorimotor synchronization. *J Exp Psychol: Hum Percept Perform*, **28**, 410–430.

**Repp, B.H.** (2002*b*). The embodiment of musical structure: effects of musical context on sensorimotor synchronization with complex timing patterns. In: W. Prinz and B. Hommel (eds) *Common mechanisms in perception and action: attention and performance XIX*, pp. 245–265. Oxford University Press, Oxford. Electronic appendix at http://www.haskins.yale.edu/MISC/REPP/Chopin_appendix.pdf.

**Repp, B.H.** (2002*c*). Phase correction following a perturbation in sensorimotor synchronization depends on sensory information. *J Motor Behav*, **34**, 291–298.

**Repp, B.H.** (2002*d*). Effects of metrical structure on phase resetting in sensorimotor synchronization. In: C. Stevens, D. Burnham, G. McPherson, E. Schubert, and J. Renwick (eds) *Proceedings of the Seventh International Conference on Music Perception and Cognition* (CD-ROM), pp. 572–574. Causal Productions, Adelaide.

Repp, B.H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: the synchronization threshold and the benefits and costs of interval subdivision. *J Motor Behav*, **35**, 355–370.

Repp, B.H. (2005a) Rate limits of on-beat and off-beat tapping with simple auditory rhythms:1. Qualitative observations. *Music Percept.* **22**, 479–496.

Repp, B.H. (2005b) Rate limits of on-beat and off-beat tapping with simple auditory rhythms: 2.The roles of different kinds of accent. *Music Percept.* **23**, 165–187.

Repp, B.H. (2005c). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review*, **12**, 969–992.

Repp, B.H. and Keller, P.E. (2004). Adaptation to tempo changes in sensorimotor synchronization: effects of intention, attention, and awareness. *Q J Exp Psychol*, **57A**, 499–521.

Repp, B.H., London, J., and Keller, P.E. (2005). Production and synchronization of uneven rhythms at fast tempi. *Music Perception*, **23**, 61–78.

Schögler, B. (1999–2000). Studying temporal co-ordination in jazz duets. *Musicae Scientiae* (special issue), 75–91.

Semjen, A. and Ivry, R.B. (2001). The coupled oscillator model of between-hand coordination in alternate-hand tapping: a reappraisal. *J Exp Psychol: Hum Percept Perform*, **27**, 251–265.

Semjen, A. and Vos, P.G. (2002). The impact of metrical structure on performance stability in bimanual 1:3 tapping. *Psychol Res*, **66**, 50–59.

Semjen, A., Schulze, H.-H., and Vorberg, D. (1992). Temporal control in the coordination between repetitive tapping and periodic external stimuli. In: *Proceedings of the Fourth Rhythm Workshop: rhythm perception and production*, pp.73–78. Imprimérie Municipale, Bourges.

Shaffer, L.H. (1984). Timing in solo and duet piano performances. *Q J Exp Psychol*, **36A**, 577–595.

Snyder, J. and Krumhansl, C.L. (2001). Tapping to ragtime: cues to pulse finding. *Music Percept*, **18**, 455–489.

**Stephan, K.M., Thaut, M.H., Wunderlich, G. *et al.*** (2002). Conscious and subconscious sensorimotor synchronization—prefrontal cortex and the influence of awareness. *NeuroImage*, **15**, 345–352.

**Thaut, M.H., Kenyon, G.P.,** Schauer, **M.L., and McIntosh, G.C.** (1999). The connection between rhythmicity and brain function. *IEEE Eng Med Biol*, *18*, 101–108.

**Toiviainen, P. and** Snyder, **J.S.** (2003). Tapping to Bach: resonance-based modeling of pulse. *Music Percept*, **21**, 43–80.

van **Noorden, L. and Moelants, D.** (1999). Resonance in the perception of musical pulse. *J New Music Res*, **28**, 43–66.

**Vorberg, D. and Schulze, H.-H.** (2002). A two-level timing model for synchronization. *J Math Psychol*, **46**, 56–87.

**Vorberg, D. and Wing, A.** (1996). Modeling variability and dependence in timing. In: H. Heuer and S. W. Keele (eds) *Handbook of perception and action*, Vol. 2, pp. 181–262. Academic Press, London.

**Vos, P.G. and Helsper, E.L.** (1992). Tracking simple rhythms: on-beat versus off-beat performance. In: F. Macar, V. Pouthas, and W.J. Friedman (eds) *Time, action, and cognition: towards bridging the gap*, pp. 287–299. Kluwer, Dordrecht.

**Wing, A.M.** (1980). The long and short of timing in response sequences. In: G.E. Stelmach and J. Requin (eds) *Tutorials in motor behavior*, pp. 469–486. North-Holland, Amsterdam.

Wing, **A.M. and Kristofferson, A.B.** (1973). Response delays and the timing of discrete motor responses. *Percept Psychophys*, **14**, 5–12.

Wohlschläger, **A. and Koch, R.** (2000). Synchronization error: an error in time perception. In: P. Desain and L. Windsor (eds) *Rhythm perception and production*, pp. 115–127. Swets and Zeitlinger, Lisse.

Part 3

# Movement analysis

Chapter 5

# Hand movements and musical performance

Thomas E. Jerde, Marco Santello, Martha Flanders, and John F. Soechting

## Introduction

Hand movements, like all movements, are subject to biomechanical and neural constraints. Accordingly, not all movements are possible and others are subject to distortions and errors. The design of musical instruments and the structure of musical compositions, as human inventions, undoubtedly conform to these constraints. Nevertheless, musical performance is a skilled and demanding task, often requiring years of practice; thus it may test the limits of these constraints. It follows that the study of motor control in music should provide a unique opportunity for delineating the neural and biomechanical constraints that also govern more mundane tasks and for describing how practice and learning shape and modify these constraints. Unfortunately, however, little is known about movement kinematics in musical performance. In this chapter we will discuss the influence of these constraints by extrapolating from observations on hand movements in other tasks.

Musical performance involves the serial execution of a sequence of notes, and as such it is a specific example of the serial organization of behaviour (Lashley 1930). It has been hypothesized that such complex movements are generated by the concatenation of simpler, basic elements, so-called movement primitives (e.g. Mussa-Ivaldi and Bizzi 2000). With practice, such movement sequences become fluent and almost automatic (cf. Hikosaka *et al.* 1995) and individual elements in the sequence may overlap temporally, a phenomenon termed 'co-articulation'. Here, we will also review some of our results on the extent of co-articulation in highly practiced behaviours such as typing, piano playing and finger spelling in American Sign Language. Finally, we will close with some general observations on the expressive requirements of musical performance and how they may affect motor performance. A schematic overview of this chapter is provided in Fig. 5.1.

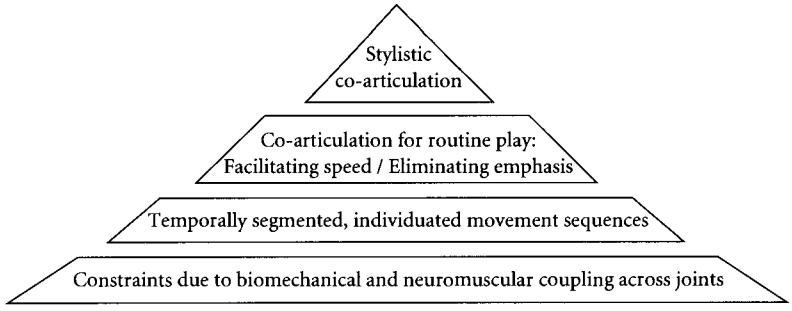



**Fig. 5.1** Levels of organization affecting motor execution during musical performance. The influence of each of the levels is discussed in the text.

## Biomechanical and neural constraints on hand movements

The major constraint that governs finger movements is that it is virtually impossible to restrict motion to a single finger. This phenomenon has been termed a lack of 'individuation' (Schieber 1995) or 'enslavement' (Zatsiorsky *et al.* 2000). It has been demonstrated in the somewhat artificial situation in which a subject is asked to exert isometric force with different combinations of fingers. However, it is also present during highly practiced movements such as typing, piano playing, and finger spelling (Flanders and Soechting 1992; Engel *et al.* 1997; Soechting and Flanders 1997; Jerde *et al.* 2003*a*).

An example is provided for the kinematics of typing in Fig. 5.2. The traces show the motion of proximal interphalangeal (pip) joints of the four fingers as touch typists typed each of the letters normally typed with the right hand. The heavier traces correspond to the finger actually striking the key (e.g. the *P* for the little finger) and the traces are arranged according to the layout of the keyboard. It is clear that each of the fingers is in motion for each of the keys, irrespective of whether or not the finger actually strikes a key or not. Furthermore, the motion at each of the pip joints had a similar time course, and an amplitude that showed a spatial gradient (largest for the *Y* and smallest for the *P* and the *L* for the little, ring, and middle fingers).

The kinematics of finger movements during piano playing exhibit similar features (Engel *et al.* 1997) although they were not studied in as detailed a fashion. A lack of individuation is also present during finger spelling (Jerde *et al.* 2003*a*). Many of the postures corresponding to the letters of the alphabet in finger spelling are at the limits of joint motion (e.g. one or more fingers fully flexed or extended), to facilitate their visual recognition. That is to say, finger spelling should foster individuation of finger movements. Nevertheless, we found that we could identify, with a high degree of reliability, the letter that was spelled from the posture of only 4 of the 15 measured degrees of freedom of the hand (using

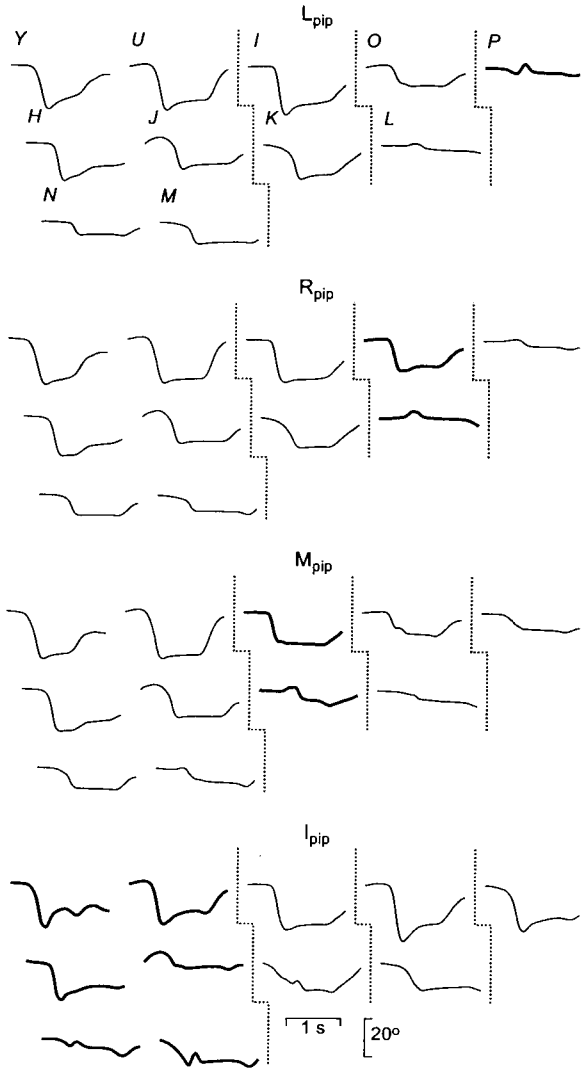



**Fig. 5.2** Motions of the proximal interphalangeal joints (pip) of each of the four fingers (index, I; middle, M; ring, R; little, L). Traces depict finger motions spanning the keypresses of each of the letters typed with the right hand, a downward deflection corresponding to extension. Each of the letters was embedded in words in which the surrounding letters were typed with the left hand. Traces corresponding to the finger actually striking the key are emphasized with a heavier line style. Note the similarity in the pattern of motion at all of the joints. (Adapted from Soechting and Flanders (1997).)

one joint of the thumb and one each of the ring, middle, and index fingers). For example, we could differentiate the *I* and the *S* (which ideally differ only in the posture of the little finger, fully flexed for *S* and extended for *I*) on the basis of the postures of the three other fingers and the thumb.

Tasks requiring the use of multiple digits such as whole-hand grasping are also characterized by a few, well-defined patterns of motion regardless of object geometry (Santello and Soechting 1998; Santello *et al.* 1998) and sensory modality used to shape the hand (Santello *et al.* 2002; Winges *et al.* 2003). The predominant pattern underlying all hand postures, independent of the number of digits used

for grasping, consisted of simultaneous flexion/extension of all digits (Santello *et al.* 1998, 2002) resulting from consistent covariations in the angular excursion of the metacarpal and proximal interphalangeal joints. This, in turn, leads to closing and opening of the hand as a unit. However, using a principal components analysis, we showed that higher-order principal components, which accounted for only a small amount of the postural variability, nevertheless contributed a significant amount of information about the object to be grasped. The results suggest fine control of finger movements superimposed on these simpler patterns.

These simple kinematic patterns emerge from a concatenation of mechanical and neural factors. The mechanical factors include the coupling among the muscle compartments and tendons of long flexors and extensors of the fingers (Fahrer 1981; von Schroeder and Botte 1993; von Schroeder *et al.* 1999; Keen and Fuglevand 2003). These muscles also cross several joints within each finger, further increasing the mechanical coupling between movements occurring at these joints. These passive linkages limit the extent to which we can move one finger independently of others in tasks involving the production of individuated finger force (Hager-Ross and Schieber 2000; Zatsiorki *et al.* 1998, 2000).

The neural factors include the organization of inputs to the motor neuron pools of hand muscles. The fibres of a given motor unit can exert force on more than one tendon (Kilbreath *et al.* 2002) and activation of such motor units would inevitably lead to forces being produced by digits that should not be involved in the task. Furthermore, motor units controlling distinct digit compartments of flexor and extensor muscles receive common input during production of individuated finger force (Reilly *et al.* 2004) as well as whole-hand grasping (Winges *et al.* 2003). This also leads to forces being generated at non-instructed digits.

The selective activation of hand muscles to produce motion of single or multiple digits probably plays an even more important role in the dexterous control of the hand. Due to the biomechanical architecture of hand muscles, complex co-activation patterns of many muscles are required to produce individuated finger movement as well as to prevent motion at non-instructed digits (Schieber 1995; Reilly and Schieber 2003). These observations suggest that the neuromuscular architecture of the hand might have evolved to allow the coordinated action of *multiple* digits rather than individuated finger control. Hence it is not surprising that extensive practice is required to attain the exquisite dexterity and individuation of finger movement necessary to play a musical instrument.

## Movement sequences and co-articulation

Co-articulation, or the influence of preceding and following movements on posture in movement sequences, is a key aspect of skilled motor planning, and has been demonstrated in speech (cf. Kent and Minifie 1977). We initially sought

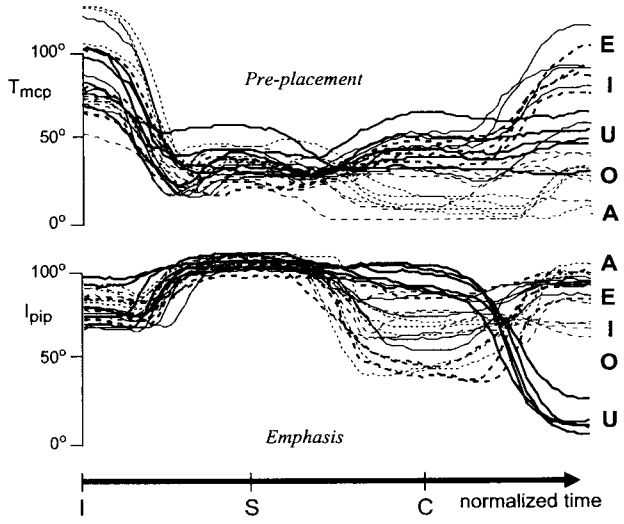



**Fig. 5.3** Assimilation (pre-placement) and dissimilation (emphasis) during finger spelling. Traces depict the motions at the metacarpophalangeal joint of the thumb ($T_{mcp}$) and of the index pip joint ($I_{pip}$) as a subject finger spelled words containing the letters ISC followed by one of the vowels (e.g. DISCOVER). Note that the posture at the two joints at the time of the 'C' depends on the upcoming vowel. Posture at the $I_{pip}$ emphasizes the distinction between the C and the next letter, while postural changes at the $T_{mcp}$ are minimized. (Adapted from Jerde *et al.* (2003*b*).)

to find evidence for this phenomenon in sequences of finger movements in typing (Soechting and Flanders 1992). However, that task is characterized by a strictly serial organization of movements; the finger movements that characterize one keypress are unaffected by the subsequent letter in the sequence up until the time of the keypress for the first letter. This was true when the next letter was typed with the opposite hand or with the same finger or a different finger of the same hand. In a more recent study of finger spelling in American Sign Language (Jerde *et al.* 2003*b*), we did find that in fluent finger spelling of individual words the hand posture spelling a given letter contains significant information about preceding and following letters (Fig. 5.3). These forward and reverse influences can be characterized as assimilation and dissimilation. In assimilation, the difference in joint angles between succeeding letters is reduced, which we interpret as pre-placing or leaving behind certain fingers. On the contrary, in dissimilation, the difference in joint angles is increased, interpreted as emphasizing the difference in hand shape between the letters. Instances of assimilation and dissimilation occurred in the same movement sequence, at the different joints (emphasizing the fact that individual joints can be controlled independently, the evidence presented in the previous section notwithstanding). Importantly, the joints which showed the greatest dissimilation were those that had been previously shown to

contain the most information about the letter being spelled (Jerde *et al.* 2003*a*). Thus it appears that in finger spelling co-articulation occurs not just as a result of optimization of kinematic profiles but also as a communicative strategy to enhance intelligibility.

Musical performance is another instance of highly practiced, skilled sequential movement in which we might expect co-articulation to play a role. Like speech or finger spelling, a highly practiced task with specific constraints on the timing of execution, the playing of a musical instrument generally involves the assembly of discrete movement units in sequence, in which movements corresponding to specific notes might be adjusted depending on what notes precede or follow. Engel *et al.* (1997) showed that certain manoeuvres in piano playing (such as the thumb-under manoeuvre when an ascending scale is played) do elicit measurable anticipatory modifications in hand kinematics, but only in specific instances, unlike the general phenomenon of co-articulation observed in speech and finger spelling. Thus it appeared that a basic level of music performance may consist of temporally segmented movement sequences (see Fig. 5.1). However, as we continue to investigate organization of movement in musical performance we may better understand the role of co-articulation, and in what ways it may be manifested, by considering more generally how a highly skilled musician translates a musical representation into a sequence of movements according to specific artistic demands.

## Motor control and musical expression

To understand more fully the role that co-articulation plays in musical performance, we should first understand its role in the music itself. Music can be defined broadly as temporally organized sounds, where each individual sound draws its significance from its relation to the others preceding and following it. In modal and tonal music, representing the majority of developed music in most cultures, a piece or passage of music is written in a certain key. This means that a specific pitch, called the 'tonic', is the foundation for the scale, and every note derives salience (and significance in analysis) from its relationship with the tonic pitch. With regard especially to the past four centuries of Western tonal music, the strongest function of a pitch is the force with which it leads back to the tonic pitch. This return to the tonic is called a cadence, and is usually perceived with a sense of finality or satisfaction; cadences of various types comprise the harmonic backbone that provides structure to the music. The pitch most predominant in the music immediately before the cadence is typically the fifth note of the scale, also called the 'dominant'; when this pitch is prominent in the underlying harmony, a cadence is usually expected. Other pitches in the scale have similarly

important and distinctive roles: the seventh pitch is the 'leading tone', which usually also leads strongly back to the tonic, several other pitches have similar leading tendencies establishing a complex vocabulary of harmonic progression, and the third pitch, depending on how wide the interval is between it and the tonic, defines the mode of the scale as major or minor, which greatly affects the mood of the music. Besides these harmonic considerations, a melody by its very nature derives its distinct character from the ordering of its pitches, and how each pitch relates to the harmony underlying it. Thus the context relating a note to others surrounding it, and the particular transitions that join them, are fundamental aspects of musical structure.

Because of these interdependences in the roles of individual notes in a piece of music, a sensitive, artistic performance depends on the performer's awareness of each note in the context of others surrounding it—for example, a pitch of C# acting as the third in an A major chord might be played differently from one acting as the leading tone in D major, even in the same passage. This phenomenon can be observed clearly in expressive melodic devices such as appoggiaturas and suspensions, which are notes that form dissonance with expected harmony. A performer can greatly heighten dramatic tension in a musical phrase by emphasizing these unexpected pitches, setting them apart from the surrounding notes by differences in volume, articulation, and timbre. In such cases, a musician must strive consciously to alter the performance of a note to draw attention to its musical role. The effect can be enhanced by altering the surrounding notes as well, to increase contrast with the note to be emphasized.

Several gross movements preceding the altered note may require adjustment to accommodate these alterations: a violinist may need to take more bow, and a wind player may need to draw extra breath several measures in advance. These types of movements represent a kind of reverse influence reminiscent of that observed in co-articulation, but it is hardly surprising that one can predict a horn player is going to play something louder because she takes a large breath. However, more subtle motor co-articulation might be observed in adjustments made on a more local level to facilitate the emphasis of particular notes. For example, a pianist can apply greater force in striking a key if the finger (or the entire arm) is raised, which requires upward movement in advance of the key-press. On a conscious level, the pianist is probably thinking about the note to be accented and the extra force necessary (the stroke), and not the preparation that must be accomplished during preceding notes (the lift). In fact, the pianist Vladimir Horowitz recommended that students attempting to learn to accent one note independently of others in a chord should at first consciously alter their hand and finger position prior to striking the chord, allowing the behaviour to become automatic with practice (Horowitz 1928). If such a technique is indeed

employed in performance, it ought to be possible to measure the modification of posture or muscle activation prior to an accented note.

In contrast to conscious applications of context-dependent nuance, musicians require practice to develop the co-articulation required to eliminate unwanted emphases from a phrase brought about by biomechanical constraints (Fig. 5.1). For example, in playing a simple scale, pianists must perform a 'thumb-under' manoeuvre to reach all the notes in sequence with the fingers of one hand, and the larger movement necessary to accomplish this imparts a natural tendency for these notes to be different from the surrounding ones in timing and in volume. It takes considerable practice to iron out these timing irregularities and unwanted volume changes by anticipatory compensation, so that the musician can play a scale with every note having equal emphasis. It is only from a baseline of smooth, unaccented performance that a musician can employ musical taste with precision, to overlay the basic phrase with nuances of articulation and intensity.

Thus a musician must undergo thorough training to be able, at an automatic level, to eliminate differences between notes, while simultaneously, at a conscious level, to emphasize differences between notes for stylistic effect (Fig. 5.1). In physical terms, both these effects must be achieved by anticipatory adjustments to the movement sequence. In general, the overlapping of multiple layers of nuance over a musical phrase (involving both steadiness and change in volume, articulation, and tempo flexibility) where these expressions are carried out by a kinematically connected physical system subject to biomechanical constraints as well as simple inertia, suggests that even if strict serial ordering of discrete movements were possible in performing a musical phrase, the results would be artistically inferior to what might be obtained were the control of each movement to be tailored to reflect the dynamic musical and muscular context in which it occurs. As Horowitz said: 'So if I play all the notes steadily along, without graduation, or without relating one to another, and without climactic or guiding emphasis, I have said nothing, even though I may have played the notes correctly and in correct time' (Horowitz 1932). Conceivably, the mechanisms used to diminish or emphasize differences in physical execution between notes in a musical phrase might be related to the assimilation and dissimilation demonstrated in finger spelling.

The technical proficiency that is required to play evenly and to layer nuances onto a musical phrase does assume the ability to play the notes in that phrase accurately and in the correct order, but this is in itself no small feat, as anyone who has tried to play an instrument can tell you. From a motor control perspective, the effortlessness with which experienced musicians accomplish this translation from notes on a page to a complicated movement sequence is perhaps even more interesting than the refinements and adjustments that set

apart a truly remarkable performance from a mediocre one. On this basic level, the importance of note-to-note transition movements becomes even more apparent, and is reflected in the standard pedagogy of musical performance. Taking the example of a wind instrument player, step one in learning the technique of the instrument is memorizing which keys must be closed to produce a given note. But mastering these static fingerings is only the lowest level of competence. Serious music students pursue practice routines of technical exercises designed expressly to isolate the most important finger movement sequences. Musicians practice endless varieties of scales and arpeggios, because in Western music if you know them in all 12 keys, you have already trained to automaticity a rather high percentage of all the note-to-note transitions and sequences you might encounter in the repertoire (Wye 2003). It is the systematic overlearning of these transitions that allows effortless fluency in performance, like acquiring a comprehensive vocabulary of note-combination transitions. Musicians also practice these patterns with varieties in tempo, volume, and articulation, attempting to internalize all combinations of conditions that might be required in a performance situation. Linda Gruson notes: 'Musical practising may be viewed as a sequence of transitions from controlled to automatic processing in which larger and larger chunks of musical information are built up from more basic sub-components' (Gruson 1988).

In this sense, musical performance is very similar to speech: the low-level selection of movements must become an automatic process so that it does not interfere with conveying the intended meaning. As an orator does not have to consciously decide how to conjugate an upcoming verb, neither does an accomplished pianist have to waste cognitive resources deciding at what point in a long scale to perform a thumb-under manoeuvre. Like an orator might decide beforehand to use a particularly cogent word or choice of word order, a pianist may consciously pre-select a particularly elegant fingering for a tricky passage. And just as an orator, despite needing command of an enormous working vocabulary, need not have practiced a specific sentence in order to say it fluently, likewise a good pianist can sight-read unfamiliar music by breaking it down into chunks and piecing them together in accordance with a standard syntax of well-studied harmony.

Even the best pianists (as well as orators) do occasionally make mistakes, however. But again, like in language, there is a tendency for musicians to make plausible mistakes that fit within the correct harmonic framework, such as playing the wrong note of a chord, rather than playing a note which is not even a part of the chord (Schaffer 1981). This tendency may be a result of both conscious and unconscious processes. Consciously, if a performer feels hesitation regarding the proper note to play, making a very quick decision is made simpler by having a

constricted range of options, limited to only those that are harmonically acceptable; this reflects an effective cognitive strategy to deal with performance error. However, an alternative explanation may involve not a conscious cognitive strategy but instead reliance on the musician's highly practiced motor programs. In a moment of crisis during performance, it is conceivable that the performer unconsciously reverts to better-established motor programs that correspond with the given harmonic situation.

The facility with which musicians learn movement sequences can also be seen in the tradition of violinists and pianists in particular to play entire recitals from memory. Here, musicians can rely not only on specific motor programs established during practice, but on other forms of memory. Visual memory plays a substantial role for some, both of the printed score and of the hands on the keyboard, and auditory memory as well. But one of the most important techniques available to highly trained musicians is analytic memory. As Aiello and Williamon (2002) describe, 'while inexperienced pianists might strive to remember all the individual notes that occur in the right and left hands in a certain measure, experienced pianists might easily recognize that those notes constituted a cadential formula in a particular tonality. Their knowledge of harmony, together with their advanced keyboard skills, would enable them to play these notes with ease.... This process of chunking allows for rapid categorization of domain-specific patterns and accounts for the speed with which experts recognize the key elements in a problem situation'. They also note several examples of professional pianists who avoid relying on kinesthetic or 'muscle' memory to perform memorized recitals, claiming it is less reliable than analytic memory. This method not only demonstrates the highly developed analytical skills that experienced musicians have available to them, but also underscores the efficiency with which they can assemble units of their musical motor vocabulary into fluent sequences. It appears that for those who have achieved true mastery of their instrument, translation from a musical representation in memory to specific motor sequences does not require conscious attention.

## Acknowledgments

This work was supported by USPHS Grants NS-15018 and NS-27484.

## References

**Aiello, R. and Williamon, A.** (2002). Memory. In: R. Parncut and G. McPherson (eds) *The science and psychology of music performance: creative strategies for teaching and learning.* Oxford University Press, New York.

**Engel, K.C., Flanders, M. and Soechting, J.F.** (1997). Anticipatory and sequential motor control in piano playing. *Exp Brain Res*, **113**, 189–199.

Fahrer, M. (1981). Interdependent and independent actions of the fingers. In: R Tubiana (ed.) *The hand*, pp. 399–403. W.B. Saunders, Philadelphia, PA.

Flanders, M. and Soechting, J.F. (1992). Kinematics of typing: parallel control of the two hands. *J Neurophysiol*, **67**, 1264–1274.

Gruson, L. M. (1988). Rehearsal skill and musical competence: does practice make perfect? In: J. Sloboda (ed.) *Generative processes in music: the psychology of performance, improvisation, and composition*, pp. 91–112. Clarendon Press, Oxford.

Hager-Ross, C.K. and Schieber, M.H. (2000). Quantifying the independence of human finger movements: comparisons of digits, hands and movement frequencies. *J Neurosci*, **20**, 8542–8550.

Hikosaka, O., Rand, M.K., Miyachi, S., and Miyashita, K. (1995). Learning of sequential movements in the monkey: process of learning and retention of memory. *J Neurophysiol*, **74**, 1652–1661.

Horowitz, V. (1928). Interview with J. Eisenberg. Noted Russian pianist urges students to simplify mechanical problems so that thought and energy may be directed to artistic interpretation. *The Musician*, June, 11.

Horowitz, V. (1932). Technic the outgrowth of musical thought. *The Etude*, March, 163–164.

Jerde, T.E., Soechting, J.F., and Flanders, M. (2003*a*). Biological constraints simplify the recognition of hand shapes. *IEEE Trans Biomed Eng*, **50**, 265–269.

Jerde, T.E., Soechting, J.F., and Flanders, M. (2003*b*). Coarticulation in fluent fingerspelling. *J Neurosci*, **23**, 2383–2393.

Keen, D.A. and Fuglevand, A.J. (2003). Role of intertendinous connections in distribution of force in human extensor digitorum. *Muscle Nerve*, **28**, 614–622.

Kent, R.D. and Minifie, F.D. (1977). Coarticulation in recent speech production models. *J Phonetics*, **5**, 115–133.

Kilbreath, S.L., Gorman, R.B., Raymond, J., and Gandevia, S.C. (2002). Distribution of the forces produced by motor unit activity in the human flexor digitorum profundus. *J Physiol (Lond)*, **543**, 289–296.

Lashley, K.S. (1930). Basic neural mechanisms in behavior. *Psychol Rev*, **37**, 1–24.

Mussa-Ivaldi, F.A. and Bizzi, E. (2000). Motor learning through the combination of primitives. *Phil Trans R Soc Lond B Biol Sci*, **355**, 1755–1769.

Reilly, K.T. and Schieber, M.H. (2003). Incomplete functional subdivision of the human multitendoned finger muscle flexor digitorum profundus: an electromyographic study. *J Neurophysiol*, **90**, 2560–2570.

Reilly, K.T., Nordstrom, M.A., and Schieber, M.H. (2004). Short-term synchronization between motor units in different functional subdivisions of the human flexor digitorum profundus muscle. *J Neurophysiol*, **92**, 734–742.

Santello, M. and Soechting, J.F. (1998). Gradual molding of the hand to object contours. *J Neurophysiol*, **79**, 1307–1320.

Santello, M., Flanders, M., and Soechting, J.F. (1998). Postural hand synergies for tool use. *J Neurosci*, **18**, 10105–10115.

Santello, M., Flanders, M., and Soechting, J.F. (2002). Patterns of hand motion during grasping and the influence of sensory guidance. *J Neurosci*, **22**, 1426–1435.

**Schaffer, L.H.** (1981). Performances of Chopin, Bach and Bartok: studies in motor programming. *Cogn Psychol*, **13**, 326–376.

**Schieber, M.H.** (1995). Muscular production of individuated finger movements: the roles of extrinsic finger muscles. *J Neurosci*, **15**, 284–297.

**Soechting, J.F. and Flanders, M.** (1992). Organization of sequential typing movements. *J Neurophysiol*, **67**, 1275–1290.

**Soechting, J.F. and Flanders, M.** (1997). Flexibility and repeatability of finger movements during typing: analysis of multiple degrees of freedom. *J Comp Neurosci*, **4**, 29–46.

**von Schroeder, H.P. and Botte, M.J.** (1993). The functional significance of the long extensors and juncturae tendinum in finger extension. *J Hand Surg Am*, **18**, 641–647.

**von Schroeder, H.P., Botte, M.J., and Gellman, H.** (1990). Anatomy of the juncturae tendinum of the hand. *J Hand Surg Am*, **15**, 595–602.

**Winges, S.A., Weber, D.J., and Santello, M.** (2003). The role of vision on hand pre-shaping during reach to grasp. *Exp Brain Res*, **152**, 489–498.

**Wye, T.** (2003). *Practice book for the flute 2, technique*. Novello, London.

**Zatsiorsky, V.M., Li, Z.M., and Latash, M.L.** (1998). Coordinated force production in multi-finger tasks: finger interaction and neural network modeling. *Biol Cybern*, **79**, 139–150.

**Zatsiorsky, V.M., Li, Z.M., and Latash, M.L.** (2000), Enslaving effects in multi-finger force production. *Exp Brain Res*, **131**, 187–195.

Chapter 6

# Movement analysis in pianists

Hans-Christian Jabusch

## Introduction

Performing instrumental music at a professional level is arguably among the most complex of human accomplishments. As is the case with other instrument families, playing on keyboard instruments requires movements with highest spatiotemporal precision. Piano music such as that from the late nineteenth century is often characterized by a hypercomplex structure and a high density of notes. In virtuoso Late Romantic music, pianists have to bimanually coordinate the production of up to 1800 notes per minute, as for example in the 6th Paganini Etude by Franz Liszt (Münte *et al.* 2002). Visualization and analysis of movements associated with such virtuoso keyboard playing as well as with simple motor tasks on the keyboard have been conducted for different purposes. Two different principles have been applied: (1) the analysis of motion of the upper extremities of the musician or (2) the investigation of the processes taking place in the instrument such as key movements or hammer movements as indirect parameters describing the musician's actions. In the following section a brief historical overview is given of the development of both directions of study.

## Historical overview

Early approaches of visualizing mechanical processes during keyboard playing resulted from attempts in the eighteenth century to record keyboard music. In 1780, Joseph Merlin constructed a so-called Klavier-Cembalo which recorded key movements as graphite lines on a clockwork-driven paper roll (Kopiez 1996). In 1895, Binet and Courtier installed a rubber tube between the keybed and the keys of a piano. Depression of a key resulted in a change in air pressure within the tube which was graphically visualized on a paper roll and indicated the duration and loudness of a tone (Binet and Courtier 1895). The temporal resolution of this system was about 1 ms (Kopiez 1996).

Piano teachers made early attempts to understand the hand motion of pianists in order to find an efficient method for playing the piano. However, the ideas of different teachers were frequently contradictory. As an example, the influential

piano pedagogue and composer Reinecke and his students suggested the principle of a fixed and a quiet hand, the keys being struck by flexion of the fingers applying only minimal motion of the forearm and metacarpus (Schultz 1962). In contrast, based on anatomical knowledge and pianistic experience, Matthay and Breithaupt advocated a touch using the arm weight which should be produced by relaxing the supporting muscles. This was argued to be the best touch for playing on a modern concert piano (Matthay 1903; Breithaupt 1912). The limitation of such statements was the fact that they were based on the teachers' personal experiences rather than on objective data or scientific observations. They did not make a clear distinction between perceptual reality and physical reality.

In the 1920s Ortmann was the first investigator to conduct extensive research on movement analysis of the piano-playing hand using a variety of technical devices which we shall outline briefly (Ortmann 1929). Direct lever recording was applied for assessment of simple motor tasks on the piano: aluminium levers were attached to the fingers closely behind the proximal interphalangeal (PIP) joints and one additional aluminium lever was attached to the metacarpus. The levers created lines on a paper roll of a so-called kymograph, recording movements of the hand as well as movements of the fingers in the metacarpophalangeal (MCP) joints. In order to detect small movements, adjustment of the levers offered three-times magnification. Attaching the levers to different parts of the movement apparatus of the hand allowed for the recording of motion of various joints or parts of the upper extremity, and visualization of their interactions during piano playing became possible. Their reaction to the resistance of the piano key as well as to finger–key–surface impact and finger–key–bed impact could be detected by simultaneous recording of the key movement using an additional lever system. Contraction of large muscles like the pectoralis major was assessed by an apparatus consisting of an airtight chamber and a plunger which was attached to the muscle. Contraction of the muscle resulted in pressure on the air chamber thus moving a piston which was connected to a writing lever. Using a combination of these devices, relationships between the time of muscle contractions and movements as well as the course of contraction and relaxation of different muscles could be illuminated during simple movements on the piano. Ortmann investigated the impact of different tempi and dynamics on muscle contractions in simple motor tasks and demonstrated that the patterns of muscle activation changed with each change of tempo or dynamics. Lateral arm movements were recorded using a pantograph which consisted of a lever system with a tracing point attached to the wrist and a recording point drawing its motion on a plain paper surface. Recording of complex movements was achieved using a photographic plate under subdued ambient light conditions: small electric bulbs were attached to those parts of the finger, hand, or arm, the movements of

which were to be recorded. Long-term exposure during complex movements on the piano allowed photographic recording of these movements. Photographic plates were arranged in two or three different planes and displayed the movements in two or three dimensions. Ortmann documented complex finger and hand movements such as in slow and rapid scales, arpeggios (chords with the individual notes being played subsequently), and other movements, and again he found the crucial impact of speed on the coordination of movements in a given motor task (Fig. 6.1). As a conclusion for piano practice, he recommended that from the beginning each passage should be practiced at that tempo and dynamics at which it is finally to be played. Moreover, Ortmann gave detailed descriptions of hand and finger movements in different conditions such as different touch forms (e.g. legato: notes are played in a smooth, connected manner; staccato: notes are played short and detached), different levels of pianistic ability, and inter-individual anatomical differences, and he gained insights relevant for piano technique, piano pedagogy, and piano practice which are still valuable today.

Bernstein and Popova undertook an impressive study with the aim of finding phenomena that were reproducibly seen in professional pianists and of

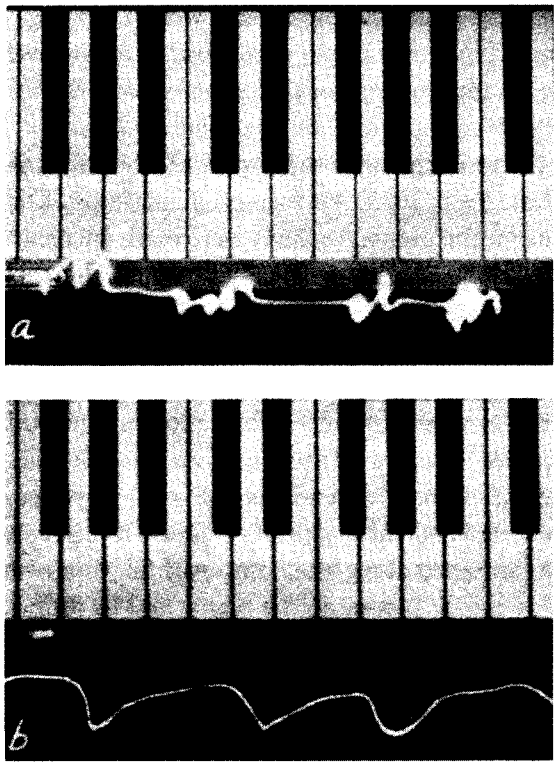



**Fig. 6.1** Movements of the metacarpus in a very slow (a) and a rapid (b) C-major scale, played on the piano with the right hand in the upward direction. Movements are displayed as bright lines and were recorded photographically from above with a light bulb attached to the metacarpus. (From Ortmann 1929.)

which were to be recorded. Long-term exposure during complex movements on the piano allowed photographic recording of these movements. Photographic plates were arranged in two or three different planes and displayed the movements in two or three dimensions. Ortmann documented complex finger and hand movements such as in slow and rapid scales, arpeggios (chords with the individual notes being played subsequently), and other movements, and again he found the crucial impact of speed on the coordination of movements in a given motor task (Fig. 6.1). As a conclusion for piano practice, he recommended that from the beginning each passage should be practiced at that tempo and dynamics at which it is finally to be played. Moreover, Ortmann gave detailed descriptions of hand and finger movements in different conditions such as different touch forms (e.g. legato: notes are played in a smooth, connected manner; staccato: notes are played short and detached), different levels of pianistic ability, and inter-individual anatomical differences, and he gained insights relevant for piano technique, piano pedagogy, and piano practice which are still valuable today.

Bernstein and Popova undertook an impressive study with the aim of finding phenomena that were reproducibly seen in professional pianists and of

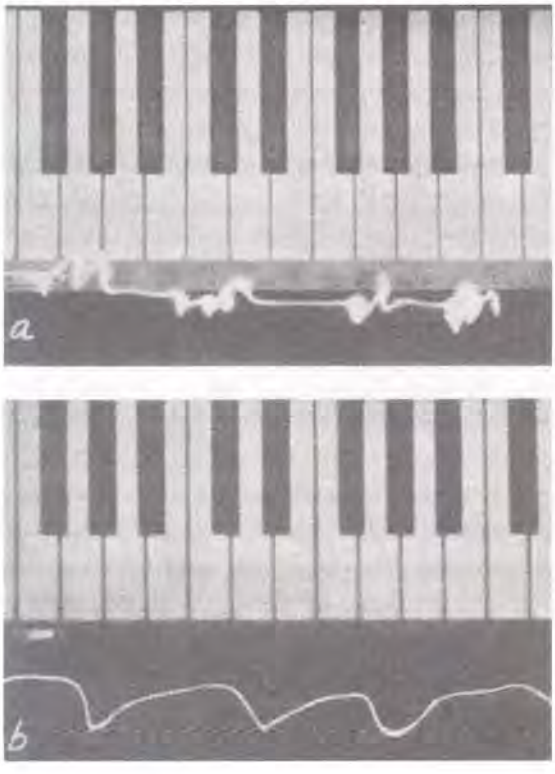



**Fig. 6.1** Movements of the metacarpus in a very slow (a) and a rapid (b) C-major scale, played on the piano with the right hand in the upward direction. Movements are displayed as bright lines and were recorded photographically from above with a light bulb attached to the metacarpus. (From Ortmann 1929.)

understanding basic mechanisms underlying pianistic movements (Bernstein and Popova 1929; Kay *et al.* 2003). In order to investigate the influence of the variables of tempo and strike force on the coordination of a certain movement, they focused on one pianistic task which consisted of octave repetitions played with the right hand under standardized conditions. Fourteen professional pianists performed the task several times at different dynamic levels ranging between pianissimo (very soft) and fortissimo (very strong) and in different tempi up to 500 strikes per minute. A kymocyclographic camera allowed visual recording of movements of light bulbs attached to the skull, the shoulder, the elbow joint, the wrist, and the metacarpus according to a principle similar to Ortmann's but as moving pictures and with a temporal resolution of up to 520 images per second. Graphodynamic methods were used to define motions of the centres of mass of the different segments of the upper extremity as well as their accelerations and the inertial forces at these points. Combining these sophisticated methods, Bernstein and Popova described physical laws underlying movements in octave repetitions. During slow tempi, movements consisted of isolated impulses; at medium tempi, movements corresponded to those of a complex pendulum; at the fastest tempi, movements transformed into forced elastic oscillations, similar to those of a simple pendulum with the wrist being kept passively elastic. Concerning the frequently discussed topic of 'arm weight playing', they reported that at very slow tempi falling of the arm purely under its own weight was seen occasionally, whereas at tempi over three strikes per second downward movements of the arm were always accompanied by muscular activity.

In 1932, Seashore conducted experiments with the aim of an 'objective analysis of musical performance' (Seashore 1936, 1938). He focused on patterns of timing and loudness in interpretations on the piano. Analyses were made by recording the hammer movements in the piano using the so-called 'Iowa piano camera' which translated piano tones into a bar code. This was created by light beams which recorded the hammer movements on a film and allowed the analysis of temporal structures of piano playing with a resolution of 10 ms. Additionally, the system recorded the hammer speed, providing indirect information on the loudness of playing. Motion of the upper extremities of the musician was not assessed. Seashore investigated temporal variations within a piece (agogics) played several times by the same pianist. He reported high intra-individual similarities in patterns of timing during repeated performance of a given composition. Moreover, he analysed patterns of loudness of chords played on the piano, and he found that chords were played with an accentuation of the highest note. Much later, these findings were confirmed by computer-based performance research.

During the following decades, no further progress was made in movement analysis in piano playing. Similar to other authors, Gát suggested various principles

of piano methodology based on his anatomical knowledge and pianistic experience (Gát 1956). He underlined his ideas with several series of photos of piano-playing hands. These photos were of accomplished pianists and demonstrated successful movement patterns to piano students and teachers. Beyond visual inspection for pedagogical purposes, however, they did not allow for an identification of general principles of motion on the piano.

In the late 1960s Wagner initiated a new era in the analysis of piano playing by applying the latest technical innovations to his studies. He investigated standardized motor tasks in order to understand basic timing phenomena in piano playing. He used an electronic device which was attached to the mechanics of the piano and produced electrical impulses when a hammer touched a string (Wagner 1968). The measuring error was 1 ms. Wagner recorded the inter-onset intervals in simple scale-playing tasks within a five-tone range under standardized conditions and in different pianists. He described reproducible individual patterns of irregularities in inter-onset intervals (Wagner 1968) and reported on the influence of tempo on the regularity of inter-onset intervals in scale playing. Standard deviations of inter-onset intervals were lowest, demonstrating a low degree of irregularity, for a medium tempo of between 6.1 and 8 strokes per second (Wagner 1971). The irregularity of inter-onset intervals increased in the ranges of higher as well as of lower tempi. In 1973, Wagner and colleagues used a computer-based feedback system in order to provide assistance in the process of piano practice. The pianist was informed about the precision of timing parameters after the performance of different exercises. The inter-onset intervals were used as a primary target variable, and feedback was given, especially with regard to temporal evenness as well as rhythmic precision in playing. Wagner objectified inter-individual differences in the stability of inter-onset intervals (Fig. 6.2) in scale playing and introduced so-called performance diagrams which provided graphical representation of timing precision and its improvement during the process of practising (Wagner *et al.* 1973).

## Current development

Since the early 1980s there has been growing interest in the analysis of movements and performance in piano playing, from different perspectives: (1) music psychology focused on performance research and expertise research in piano-playing; (2) in neuroscience, piano playing was chosen as a paradigm for research into sensory-motor control; and (3) in the field of musicians' medicine, movement analysis in piano playing became a promising diagnostic tool and provided insights into the pathophysiology of playing-related disorders. Current developments shall be outlined in the fields of both movement analysis of the upper extremities and analysis of indirect parameters such as key movements or hammer movements.

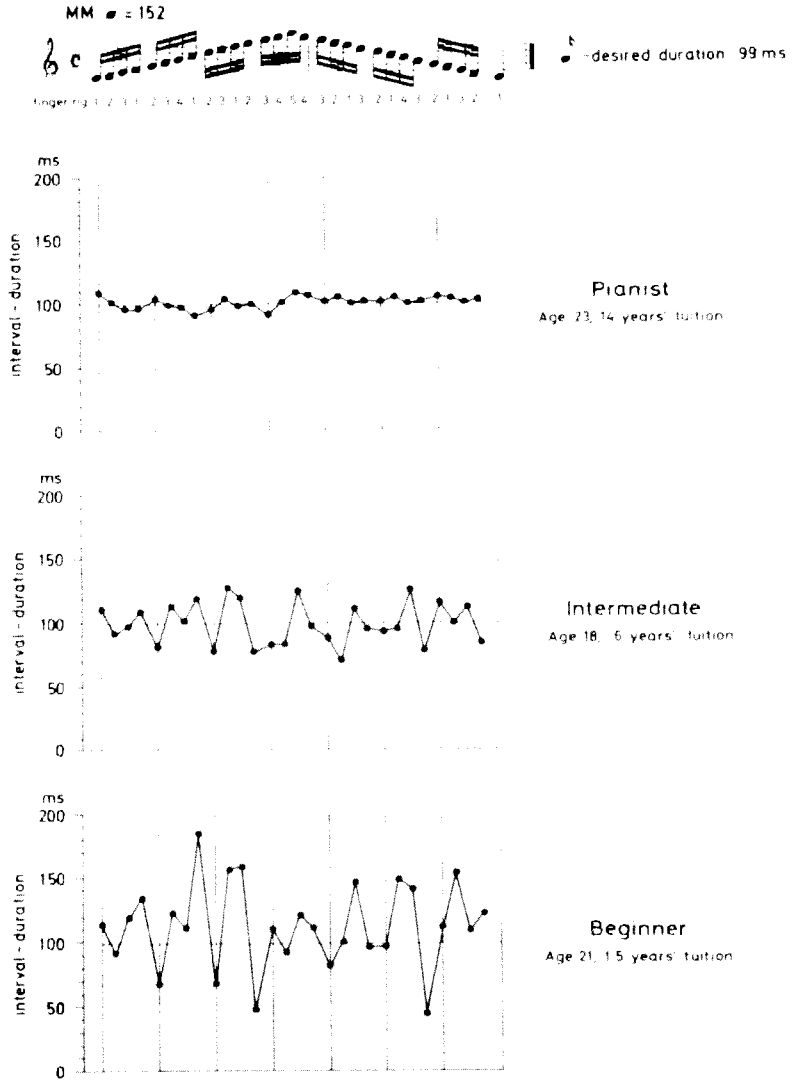



**Fig. 6.2** Inter-onset intervals (termed interval duration) of subsequent notes in C-major scales played by a beginner, an intermediate student, and a professional pianist and recorded using a computer-based device attached to a Bechstein piano. (From Wagner *et al.* 1973.)

## Analysis of keyboard-related parameters

After the establishment of the 'music instrument digital interface' (MIDI) protocol as a standard of digital processing of music, the first MIDI keyboards were introduced in 1983. MIDI allowed a computer-based analysis of performance parameters such as loudness and timing parameters, the latter with a precision in the range of milliseconds. MIDI keyboards provided information about onset

and offset times of tones, about tone pitch, and about key velocities as an indirect measure of loudness. MIDI became a valuable tool for investigation of topics in music psychology. In the following, several examples are given for this application. Using a MIDI keyboard, Wilson recorded trills and a segment of Chopin's B-minor Sonata played by different pianists. He visualized tone durations and inter-onset intervals and described individual playing patterns (Wilson 1992). Krampe investigated the role of deliberate practice in the acquisition and maintenance of high-level skills in pianists (Krampe 1994). Using a MIDI keyboard, he focused on the repeatability of musical performance in 48 pianists of different ages and levels of skill who were repeatedly playing J. S. Bach's Prelude No.1 in C major from the first volume of the Well-Tempered Clavier. Consistency of interpretation was used as a marker of musical skill. Krampe found that variations of tempo and dynamics were more consistent in professional pianists than in amateurs. There was no age-related decline in the variation of speed and loudness. Additional information was obtained by means of diaries, by retrospective estimates about past amounts of training, and by the analysis of other skill-related tasks. As a result, Krampe described a model based on the relation between intensity of practice during various life phases and performance which explained group differences in skill-related tasks.

MIDI-based analysis was used for the monitoring of practice effects after mental practice (Highben and Palmer 2004), for transfer-of-learning experiments (Meyer and Palmer 2003), and for error analyses (Repp 1996*b*; see also Chapter 3). Several MIDI-based studies with the focus on temporal and dynamic features of piano performance were published by Repp. In two pianists playing Schumann's 'Träumerei' at different tempi he found that major temporal and dynamic features of the performances changed roughly in proportion with the tempo (Repp 1994). In MIDI recordings of 10 pianists playing from the Romantic and Impressionistic piano literature, analysis of note-onset times revealed a pervasive tendency for the highest-pitched notes to lead the lower-pitched notes. Strong correlations were found between average MIDI velocity differences and average lead time (Repp 1996*a*). MIDI recordings were utilized to determine the relation between performance parameters and the structure of the composition. In MIDI recordings of the Classical and Romantic piano literature, timing patterns and dynamic characteristics served to underline the structure of the music (Repp 1990, 1996*c*). Similarly, Shaffer (1995) found that performance both elucidated structure and invented a musical character which reflected patterns of expression. In earlier studies, Shaffer used a modified Bechstein grand piano for examination of timing and dynamics and their relation to the musical structure. The piano was equipped with optical sensors to allow recording of timing and dynamics, based on hammer velocity, with a measuring error of less

than a millisecond (Shaffer 1981). Analysis of timing patterns in a Bach fugue played by an expert pianist led to the assumption that timing was determined by some form of underlying clock system. Timing in polyrhythmic structures of three against four played with both hands in a Chopin etude (No. 1 in F minor from the 'Trois Nouvelles Etudes') was analysed with respect to independence of movement in different hands and fingers. There was great use of rubato (flexible tempo), and there was evidence that the metrical polyrhythm was played under two levels of deformation: one was a rubato common to both hands, the other was a temporal displacement of one hand relative to the other. Shaffer suggested that an abstract representation of motor output is mapped into two command arrays, one for each hand, each having a clock timing with a flexible coupling between both clocks (Shaffer 1981).

MIDI-based analysis of piano performance as an indirect measure of finger motion was reported by Jabusch *et al.* (2004). In a study with 13 professional pianists, the evenness of scale playing in the legato style was investigated, with special focus on hand and finger coordination in the thumb-under movements in the ulnar playing direction as well as finger cross-over manoeuvres in the radial playing direction. The pianists played sequences of 10 to 15 C-major scales over two octaves in both playing directions and with each hand separately. Scales were played using the regular C-major fingering (1–2–3–1–2–3–4–1–2–3–1–2 –3–4–5 and reverse). The tempo was metronome-paced and standardized at 120 beats per minute for a quarter note, and scales were played in sixteenth notes. As an example, scale analysis results are reported for the right-hand tests. Over all scales of all pianists the mean standard deviation of inter-onset intervals was 8.1 ms in the ulnar playing direction and 8.9 ms in the radial playing direction, indicating an extraordinarily high level of evenness in playing. This was achieved by a peculiarity in the coordination of the thumb-under movements (3–1 and 4–1 fingering in the ulnar playing direction) and finger cross-over manoeuvres (1–3 and 1–4 fingering in the radial playing direction). In the preparation phase of thumb-under movements in the ulnar playing direction, there was a consistent finding of significantly shortened tone durations of notes preceding the tones played by the thumb. This resulted in minimal gaps (mean 13.0 ms) between the notes at the 3–1 and 4–1 fingering positions (Fig. 6.3). In the radial playing direction, such significant shortening of tone durations in the preparatory phase of finger cross-over manoeuvres was seen in the 2–1 fingering positions, indicating an early pronation movement of the hand facilitating the 1–3 as well as 1–4 finger cross-over manoeuvres. Impaired auditory perception in the temporal proximity of key strokes due to auditory forward and backward masking phenomena may explain that such irregularities in the note offsets are not perceived and are therefore not a focus for practice (Yost 2000).

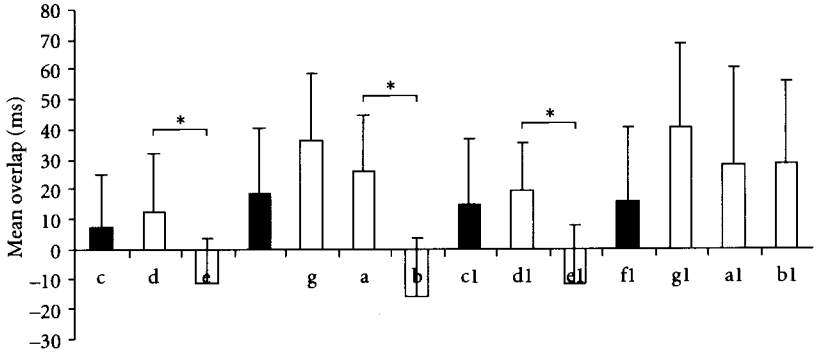



**Fig. 6.3** MIDI-based scale analysis (Jabusch *et al.* 2004). Overlaps of subsequent notes of an upward C-major scale played in the legato style with the right hand. Bars display the mean values of repeated performances of 13 healthy professional pianists (error bars: standard deviations). Black bars indicate notes played with the thumb. Negative overlaps indicate short leaps at the positions of the thumb-under movements (*$P < 0.01$).

## Upper extremity movement analysis

Technical progress in three-dimensional movement analysis during the 1990s allowed optoelectronic recording of upper extremity motion with a high spatial and temporal resolution. The ELITE system (B.T.S., Milan, Italy), for example, consists of several infrared-sensitive cameras which after illumination by infrared light capture the position of reflective markers. These can be attached to the fingers or other parts of the upper extremity. Three-dimensional coordinates of the markers are traced automatically during motion at a sampling rate of 100 Hz and with a spatial resolution of about 0.3 mm (Engel *et al.* 1997; Sforza *et al.* 2003). Engel *et al.* (1997) used this system in combination with MIDI-based performance analysis and recorded the movements of the right wrist and fingers of pianists. The pianists played scales with identical beginnings and different continuations, either including or excluding thumb-under movements. The aim of the study was to ascertain in which way hand and finger kinematics diverged prior to the depression of the last common note. As a result, playing an ascending scale with the requirement of a thumb-under manoeuvre could evoke an anticipatory modification as much as 500 ms in advance of the last common note as compared with scale playing without thumb-under movement (Engel *et al.* 1997). In view of these findings, it was suggested that a strict serial execution of a movement sequence was favoured as long as this was compatible with the demands of the tasks. Sforza *et al.* (2003) used the ELITE system in order to investigate the repeatability of finger movements during scale playing in different pianists with different experience. Repeatability of movements of different recordings was analysed for each finger by superimposing the traces of the

vertical coordinate of the finger tip of the different recordings as a function of time. Comparison of the areas under these superimposed graphs revealed a repeatability of between 65.8 and 81.4% (mean superimposition) in five different pianists. Repeatability was found to be lower in experienced pianists than in learners and teachers. As a possible explanation, the experts were suggested to be more focused on other aspects of musical performance than hand and finger movements. This might be in contrast to less experienced pianists who might be more focused on the movements required for execution of a basic task such as scale playing (Sforza *et al.* 2003). Palmer and Dalla Bella (2004) used a combination of MIDI and three-dimensional movement analysis using video cameras and passive markers on pianists' hands and fingers to investigate the relation between tempo and movements. Pianists repeatedly performed short melodies with increasing tempi until they made errors. Most tempo-related changes in the variability of finger movements were observed in the plane perpendicular to the keyboard. Peak amplitudes of motion before striking the keys increased as tempo increased. It was concluded that increased movement amplitudes at faster rates may reduce or compensate for trade-offs between speed and accuracy.

## Movement analysis in musicians' medicine

Movement analysis in keyboard playing has frequently been reported in the context of musicians' medicine. Pain and overuse syndromes are frequent among pianists, and many musculoskeletal problems in pianists appear to be due to technique, working conditions, or posture (Brandfonbrener 2000). There is an early report by Salmon and Newmark (1989) about the applicability of MIDI in analysing technical aspects of musical skills in the context of musicians' medicine. They suggested using MIDI to assess the impact of hand and finger injuries and to evaluate the effects of rehabilitation programmes. Fry *et al.* (1998) investigated motor control in pianists suffering from musculoligamentous overuse in the arms and hands. They compared parameters such as tone durations, inter-onset intervals, key velocities, and time off the metronome beat in MIDI recordings of pianists with overuse syndromes and a matched group of pianists with no history of overuse. They found more skill-based errors in affected pianists and concluded that a disturbance of coordination could be observed in pianists with an overuse syndrome.

Movement analysis in pianists has been used to find risk factors for overuse syndromes and to identify 'protective behaviour' on the piano with the aim of prevention of overuse. In order to elucidate the influence of finger positions on forces in joints, tendons, ligaments, and muscles, Harding *et al.* (1989) developed a two-dimensional mathematical model of a pianist's finger. Different pianists played on a force-calibrated MIDI keyboard, and recording data were analysed to

determine those finger positions that minimize biomechanical tensions in tendons and joint forces during piano playing. Multiple key strikes played with the same key velocity (loudness) and with different finger positions were analysed. Compared to other positions, the use of a curved finger position with a large flexion angle of the metacarpophalangeal joint was found to reduce tensions in the flexor tendons and resultant forces in the finger joints. It was concluded that the described model might help pianists to minimize tendon tension and joint forces with the particular aim of prevention of injury. In a subsequent study, Wolf *et al.* (1993) used the same method and an additional video system consisting of two cameras. One camera recorded the pianists' hands in the vertical view, another recorded the horizontal view in a sagittal plane. Eight injury-free pianists with different levels of experience were recorded playing a passage from one of Mendelssohn's Songs Without Words (opus 19/2). Data for the force of the key strikes correlated inversely with piano-playing experience. In keeping with Harding *et al.*'s (1989) observation, a curved finger position with an almost perpendicular contact between the distal phalanx and the key resulted in the lowest force coefficients in the distal interphalangeal (DIP) joints as well as in the tendons of the flexor digitorum profundus muscle (Wolf *et al.* 1993). Bejjani *et al.* (1989) also investigated piano playing with this curved finger position and compared it with a flat hand position with extended fingers as well as with a roof-like position characterized by a right-angle flexion in the MCP joints. Using three-dimensional video analysis and surface electromyography of eight muscle groups of the right upper extremity, they found that the curved finger position required a maximum of abduction–adduction movements in the MCP joints, and the flat hand position with extended fingers required the minimum of such movements. The latter position was associated with significantly less activity in the flexor carpi radialis muscle than the other positions. Activity of the anterior deltoid muscle was independent of the hand position. The roof-like hand position with a right-angle flexion in the MCP joints was associated with greatest range of ulnar–radial deviation (Bejjani *et al.* 1989). The impact of playing technique on wrist kinematics was also studied by Chung *et al.* (1992). They examined 'weight playing' pianists who were playing with increased usage of intrinsic hand muscles and forearm rotation and limited use of long flexors and extensors, and compared them with pianists playing in the 'traditional style'. Using a biaxial electrogoniometer, they found that in exercises and in classical music 'weight playing' was associated with less wrist motion than 'traditional playing'. The opposite was found for flexion–extension activity in the wrist during arpeggios and trills. Traditional pianists tended to exhibit increased radioulnar activity (Chung *et al.* 1992).

Biomechanical load in different joints depending on different motor tasks was investigated by Sakai *et al.* (1996) using a passive movement analysis system

(Expert Vision System, Motion Analysis Corp., Santa Rosa, CA, USA) which provided an accuracy to within a 0.1% error of the field of view. Ten pianists were examined during scale playing and during the playing of repeated chords. Specific joints were identified which displayed different biomechanical loads during playing of different motor tasks. In scale playing, flexion was mainly observed in the MCP joints whereas the proximal interphalangeal (PIP) joints were extended. In contrast, playing of repeated chords was associated with repeated flexion, predominantly at the wrist joint. The maximum extension angle of the wrist was significantly larger in chords than in scales; the maximum extension angle in the MCP and PIP joints was significantly larger in scales than in chords. Maximum flexion in the DIP joints was larger in scales than in chords (Sakai *et al.* 1996).

Hand ergonomics and their influence on touch control were studied by Lee (1990). Biomechanical variables such as hand and finger length, hand width, finger span, and ulnar deviation were measured in 13 pianists who played a 'tied-finger' exercise and an arpeggio exercise over more than two octaves. Pianists with greater wrist mobility played the tied-finger exercise with a shorter note duration and in a faster tempo, and they had less sustaining control of the tied finger. No impact of biomechanical variables was seen on arpeggio playing.

Dynamic finger forces in pianists with different levels of skill were investigated by Parlitz *et al.* (1998) using a modified grand piano with an f-scan sensor-matrix foil (Tekscan Inc., Boston, MA, USA) connected to a computer. Two groups of pianists, one amateur and one expert group, played three tied-finger exercises by Dohnanyi with different levels of difficulty. From dynamic force measurements, the mean pulse per touch as well as the mean touch duration were analysed. Playing in a standardized tempo and loudness, amateurs applied significantly higher forces to the keys, with a longer duration. Finger forces and duration increased with higher demands in both amateurs and expert pianists. Since the resolution of force detection was limited, a more sophisticated system was developed. In order to identify 'economy' in piano playing, Tiedemann and colleagues used a grand piano with modified keys allowing for continuous force measurement during all stages of key strike and release (Drescher *et al.* 1999; Tiedemann *et al.* 2000). The keys were constructed as double transverse beams, each with four strain gauge foils which detected the forces acting on different positions of the key and allowed precise force measurement independent of the finger position (Fig. 6.4a). Using this device, forces acting on the keys were obtained before and after onset of the tones in expert pianists, amateurs, and novices. A computer program was developed which provided real-time visual feedback of applied forces (Fig. 6.4b). Novices were playing in a 'force-uneconomical' way,

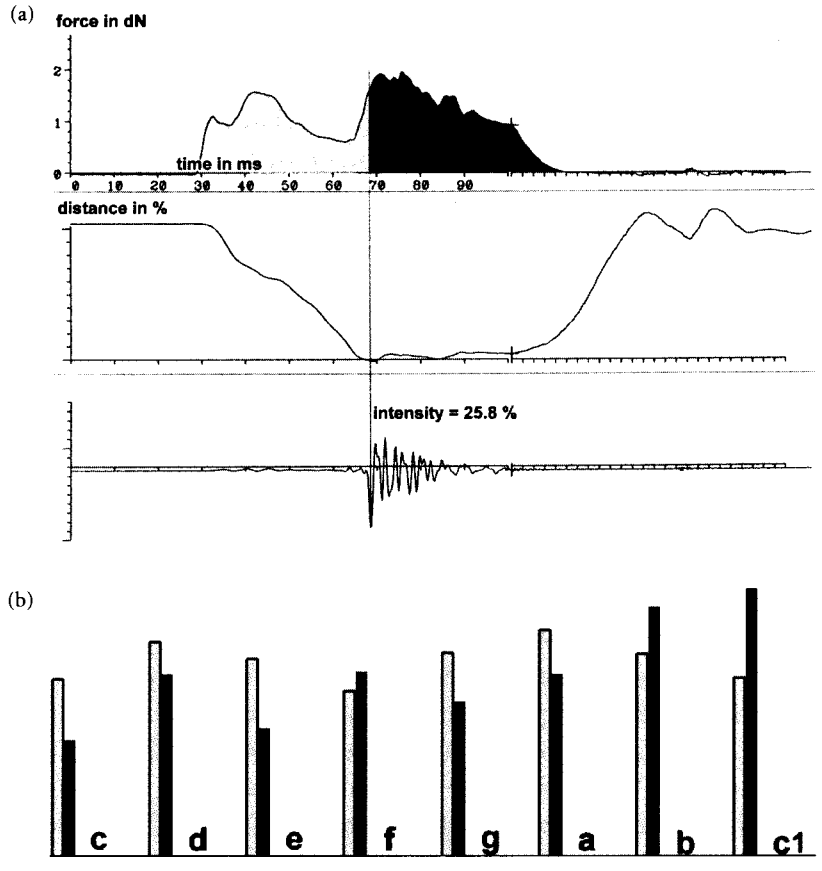



**Fig. 6.4** (a) Graphical representation of the force applied to a piano key (upper graph), the distance of key travel (middle graph) and the negative acceleration of the hammer while striking the strings (lower graph) during a single key stroke and release. The force/time area prior to note onset is shaded in grey; the force/time area after note onset is shaded in black. (b) Visual feedback screen after playing a C-major scale. The column pairs indicate force/time areas prior to note onset (grey) and after note onset (black) for all notes of the scale. Especially after onset of the notes f and b and c1, a lot of force was acting on these keys, demonstrating playing with 'uneconomic force'. (Adapted from Drescher 1999.)

applying high key forces especially during the period after onset of the tone and before key release. The forces acting on the key in this period were significantly lower in expert pianists and amateurs and could even be reduced after 5 min of playing with visual feedback of key forces in both groups. Playing *con espressione* (with musical expression) was associated with higher after-onset forces acting on the keys, even in expert pianists. It was concluded that advanced and expert pianists applied strategies to reduce unnecessary forces in piano playing and that

in emotionally intense playing even experts preferred expressive gestures over 'economical' use of forces (Tiedemann *et al.* 2000).

Jabusch *et al.* (2004) described the application of the aforementioned MIDI-based scale analysis for quantification of focal dystonia in pianists. In focal dystonia (see Chapter 17), involuntary flexion or extension of individual fingers fundamentally impairs technical skills on the instrument. Because focal dystonia is task-specific, quantification has to be carried out while the patient is executing the task; thus a pianist has to be evaluated while playing the piano (Altenmüller 2003). Scale playing is affected early during onset of pianist's dystonia. Jabusch *et al.* (2004) conducted MIDI-based scale analysis of standardized and metronome-paced C-major scales in eight patients with pianist's dystonia and eight healthy pianists matched for age, gender, handedness, and pianistic expertise. The individual duration of the disorder in the patient group was between 6 months and 8 years (2.9 $\pm$ 2.5 years) at the time of the study. Mean standard deviations of inter-onset intervals (sIOI) were calculated as a measure of irregularity in scale playing. Consistently, higher values were found for sIOI in affected hands of patients (26.3 $\pm$ 9.9 ms) compared with the hands of the healthy group (9.5 $\pm$ 2.1 ms; $P < 0.001$, Mann–Whitney U-test) displaying a higher level of irregularity in patients' scales. By calculating the mean sIOI plus two standard deviations in healthy reference hands a range of normality was identified (sIOI $<$ 13.7 ms) for the healthy subjects of this study. All sIOI findings in dystonic hands were above this range of normality. Thus, sIOI was found to be a highly discriminative parameter for the playing impairment in dystonic pianists. One patient was treated with injections of botulinum toxin A (see Chapter 17). Treatment effects were monitored by a series of 11 follow-up examinations within a time period of 151 days. After a significant initial decrease of symptoms, on day 32 post-injection, the sIOI value was inside the range of normality yielded in this study. At this time, the patient was able to play technically demanding pieces which had not been possible before treatment. Subsequently, slow fading out of treatment effects could be monitored by gradually increasing irregularities in inter-onset intervals. A second injection of botulinum toxin A on day 147 resulted in significantly lowered sIOI values 4 days later (Fig. 6.5). All subjects were recorded playing scales on audio/video. Three raters blinded to the state of health of the pianists rated the recordings according to the four-step Arm Dystonia Disability Scale (ADDS) (Fahn 1989). Subjective rating scores correlated with sIOI values (Spearman $r = 0.79$, $P < 0.01$). sIOI values, however, provided a higher resolution than the ADDS scores. MIDI-based scale analysis, especially the sIOI parameter, turned out to be a precise and reliable tool for indirect assessment of motor control in pianists with and without movement disorders.

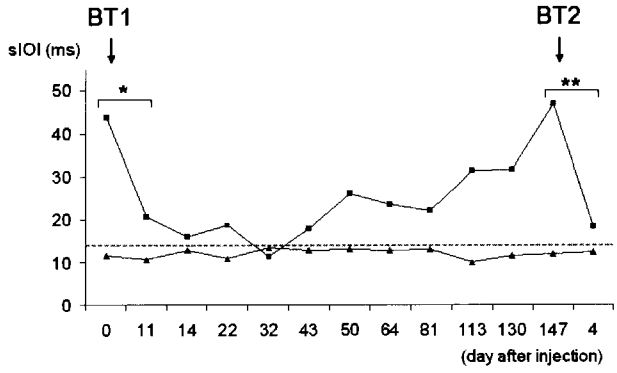



**Fig. 6.5** MIDI-based scale analysis results for a pianist with focal dystonia of the right hand during therapy with botulinum toxin A (BT). The degree of irregularity is reflected by mean standard deviations of inter-onset intervals (sIOI). Black squares: sIOI values for the affected hand. Black triangles: sIOI values for the unaffected hand. The area under the dashed line is the range of normality. BT1, date of first injection of BT; BT2, date of re-injection of BT (*$P < 0.01$; **$P < 0.001$). (From Jabusch *et al.* 2004.)

## Conclusion

Movement analysis in piano playing has been at the focus of different disciplines such as music psychology, piano pedagogy, neuroscience, and musicians' medicine. Analyses were based on two principles, either the detection of motion of the upper extremity of the pianist or the investigation of keyboard-related parameters indirectly describing the pianist's actions. As early as the 1920s, and with limited technical equipment, physiologists had conducted impressive studies and drawn conclusions about the basic physical principles of pianists' hand movements, some of which are still valid today. Technical progress in three-dimensional movement analysis and in MIDI-based techniques has helped to improve spatial and temporal resolution of movement detection and performance measurements to ranges of millimetres and milliseconds. Other techniques, such as force measurements and electromyography, have been described in order to examine a variety of additional parameters simultaneously. All these sophisticated methods have allowed illumination of the principles of pianists' motor behaviour and motor learning and the role of practice and expertise. For medical problems such as overuse syndromes and focal dystonia these techniques have helped to find risk factors or have served as diagnostic tools.

Although the technical possibilities are great and the numerous studies described above have been performed with high precision and thoroughness, our understanding of movements in pianists is still limited. We have to admit that we are far from understanding the complex phenomena and processes taking place when a pianist masters a piece with maximum demands, such as Franz Liszt's 6th

Paganini Etude. It remains enigmatic how a pianist manages to produce up to 1800 key strokes per minute, every note being played with a temporal precision in the range of milliseconds and every note embedded in a grand musical context which we experience as musical beauty.

## Acknowledgement

The author wishes to express his appreciation to Reinhard Kopiez, PhD, from the Institute for Research in Music Education (University of Music and Drama, Hannover, Germany) for valuable discussion of this topic.

## References

**Altenmüller, E.** (2003). Focal dystonia: advances in brain imaging and understanding of fine motor control in musicians. *Hand Clin*, **19**(3). 523–538.

**Bejjani, F.J., Ferrara, L., Xu, N.** *et al.* (1989). Comparison of three piano techniques as an implementation of a proposed experimental design. *Med Probl Perform Art*, **4**, 109–113.

**Bernstein, N.** and **Popova, T.** (1929). Untersuchung über die Biodynamik des Klavieranschlags. *Arbeitsphysiologie*, **1**, 396–432.

**Binet, A.** and **Courtier, J.** (1895). Recherches graphiques sur la musique. *L'Année Psychol*, **2**, 201–222.

**Brandfonbrener, A.G.** (2000). Epidemiology and risk factors. In: R. Tubiana and P.C. Amadio (eds) *Medical problems of the instrumentalist musician*, pp. 171–194. Martin Dunitz, London.

**Breithaupt, R.M.** (1912). *Die natürliche Klaviertechnik. Volume 1: Handbuch der modernen Methodik und Spielpraxis*, 3rd edn. C.F. Kahnt Nachfolger, Leipzig.

**Chung, I.S., Ryu, J., Ohnishi, N., Rowen, B.,** and **Headrich, J.** (1992). Wrist motion analysis in pianists. *Med Probl Perform Art*, 7, 1–5.

**Drescher, D., Parlitz, D., Tiedemann, J.,** and **Altenmüller, E.** (1999). A new measuring instrument for determining and depicting the forces of the fingers applied to the piano keys. *Europiano*, **39**(2), 23–42.

**Engel, K.C., Flanders, M.,** and **Soechting, J.F.** (1997). Anticipatory and sequential motor control in piano playing. *Exp Brain Res*, **113**, 189–199.

**Fahn, S.** (1989). Assessment of primary dystonias. In: T.L. Munsat (ed) *Quantification of neurologic deficit*, pp. 241–270. Butterworths, Boston, MA.

**Frucht, S.J., Fahn, S.,** and **Greene, P.E.** (2001). The natural history of embouchure dystonia. *Mov Disord*, **16**(5), 899–906.

**Fry, H.J.H., Hallett, M., Mastroianni, T., Dang, N.,** and **Dambrosia, J.** (1998). Incoordination in pianists with overuse syndrome. *Neurology*, **51**, 512–519.

**Gát, J.** (1956). *Die Technik des Klavierspiels*. Corvina, Budapest.

**Harding, D.C., Brandt, K.D.,** and **Hillberry, B.M.** (1989). Minimization of finger joint forces and tendon tensions in pianists. *Med Probl Perform Art*, **4**, 103–108.

**Highben, Z.** and **Palmer, C.** (2004). Effects of auditory and motor mental practice in memorized piano performance. *Bull Council Res Music Educ*, **159**, 58–65.

**Jabusch, H.C., Vauth, H., and Altenmüller, E.** (2004). Quantification of focal dystonia in pianists using scale analysis. *Mov Disord*, **19**(2), 171–180.

**Kay, B.A., Turvey, M.T., and Meijer, O.G.** (2003). An early oscillator model: studies on the biodynamics of the piano strike (Bernstein & Popova, 1930). *Motor Control*, **7**, 1–45.

**Kopiez, R.** (1996). Aspekte der Performance-Forschung. In: H. de la Motte-Haber (ed.) *Handbuch der Musikphysiologie*, 2nd edn, pp. 505–587. Laaber-Verlag, Laaber.

**Krampe, R.T.** (1994). Maintaining excellence: cognitive-motor performance in pianists differing in age and skill level. *Studien und Berichte/Max-Planck-Institut für Bildungsforschung*, 58. Edition Sigma, Berlin.

**Lederman, R.J.** (1991). Focal dystonia in instrumentalists: clinical features. *Med Probl Perform Art*, **6**, 132–136.

**Lee, S.H.** (1990). Pianists' ergonomics and touch control. *Med Probl Perform Art*, **5**, 72–78.

**Matthay, T.** (1903). *The act of touch in all its diversity*. Longmans, Green and Co., London.

**Meyer, R.K. and Palmer, C.** (2003). Temporal and motor transfer in music performance. *Music Percept*, **21**, 81–104.

**Münte, T.F., Altenmüller, E., and Jäncke, L.** (2002). The musician's brain as a model of neuroplasticity. *Nat Rev Neurosci*, **3**(6), 473–478.

**Neuhaus, H.** (1967). *Die Kunst des Klavierspiels*, pp. 73–74. Musikverlage Hans Gerig, Cologne.

**Ortmann, O.** (1929). *The physiological mechanics of piano technique*. E.P. Dutton & Co., New York. (Reprinted 1962)

**Palmer, C. and Dalla Bella, S.** (2004). Movement amplitude and tempo change in piano performance. *J Acoust Soc Am*, **155**, 2590.

**Parlitz, D., Peschel, T., and Altenmüller, E.** (1998). Assessment of dynamic finger forces in pianists: effects of training and expertise. *J Biomech*, **31**(11), 1063–1067.

**Repp, B.H.** (1990). Patterns of expressive timing in performances of a Beethoven minuet by nineteen famous pianists. *J Acoust Soc Am*, **88**(2), 622–641.

**Repp, B.H.** (1994). Relational invariance of expressive microstructure across global tempo changes in music performance: an exploratory study. *Psychol Res*, **56**(4), 269–284.

**Repp, B.H.** (1996*a*). *Patterns of note onset asynchronies in expressive piano performance. J Acoust Soc Am*, **88**(2), 622–641.

**Repp, B.H.** (1996*b*). The art of inaccuracy: why pianists' errors are difficult to hear. *Music Percept*, **14**(2), 161–184.

**Repp, B.H.** (1996*c*). The dynamics of expressive piano performance: Schumann's 'Träumerei' revisited. *J Acoust Soc Am*, **100**(1), 641–650.

**Sakai, N., Liu, M.C., Su, F.C., Bishop, A.T., and An, K.N.** (1996). Motion analysis of the fingers and wrist of the pianist. *Med Probl Perform Art*, **11**, 24–29.

**Salmon, P. and Newmark, J.** (1989). Clinical applications of MIDI technology. *Med Probl Perform Art*, **4**, 25–31.

**Schultz, A.** (1962). Introduction. In: Ortmann, O. The physiological mechanics of piano technique, Reprint, XV–XXIX.

**Seashore, C.E.** (ed.) (1936). *Objective analysis of musical performance*. The University Press, Iowa City, IA.

Chapter 7

# Fingering and bowing in violinists: a motor control approach

Mario Wiesendanger, Andreas Baader, and Oleg Kazennikov

## Overview

The degree of motor control required by professional violin players is staggering and derives from many years of continuing practice, typically beginning at a young age. The constraints of the musical score as well those due to the 'handling' of the violin demand continued efforts to master rather awkward positioning and the fast and asymmetric movements of different bowing speeds and the required temporal and spatial precision in fingering. The main aim of our work described in this chapter was to assess the timing of inter-limb coordination during playing of a simple tune on the violin. With an optoelectronic video system, we recorded the trajectories and their time derivatives of individual fingers (left), and bow strokes (right). Fingering was characterized by serial and parallel (anticipatory) mechanisms. Furthermore, timing and pitch (the spatial precision of contact between the finger and the string) were often determined by a combined action of two fingers of the left hand. Surprisingly, we found that synchronization between bow and finger actions could deviate by about 50ms, rarely up to 70 ms, from ideal synchronization (0 ms). This 'error' was, however, not perceptible. We explain this finding as being due to a delayed initiation of the string vibrations and the build-up to a full spectrum of sound after initiation or reversal of bow strokes. A number of delaying mechanisms intervene during playing, such as movements of the whole hand along the fingerboard (position changes), finger changes to another string, bow reversal, and bow-string changes. These irregularly occurring events need extra time (of the order of 20–30 ms), and yet these factors do not disturb the flow of the music. It has also been established that string vibrations occur about 50 ms after the initiation of a bow stroke (the sticking effect). We therefore concluded that the physical properties of vibrating strings allow for a smooth perception of the flow of sound, even with deviations of about 50 ms from perfect bow–fingering synchronization. Music also needs instances of sharp accentuations, e.g. with beat and syncopation. In preliminary

investigations with force gauges we measured bowing forces, i.e. grip force at the frog with the right thumb and the middle finger and the pressure exerted on the bow with the right index finger. These forces control tonal intensity, from the softest pianissimo (*pp*) to the loudest fortissimo (*ff*). Gradual fluctuations in intensity (crescendo and decrescendo), as well as quick changes in intensity for accentuations of the melody are reflected in the kinetics of bowing. In addition the pressure executed mainly by the right index finger also needs to compensate for the ever-changing bow lever: when the bow moves toward the tip of the bow (the down-stroke), the bow pressure has to increase progressively in order to maintain a constant intensity of tone. In the case of an up-stroke, the bow has to be unloaded progressively. We conclude that even relatively simple investigations may provide interesting insights into the rich field of biomechanics and motor control in music performance.

## Some history of motor control in music performance

A major emphasis in music research has been on perceptive aspects including rhythm, timbre, harmony, syntax, and semantics, and recently also much on its representation in the brain. Auditory perception is, however, preceded by, and the result of, an extraordinarily high level of skillful motor control. To play virtuoso music, like Paganini's Capricci, requires exceptional skills in bowing and fingering, sometimes combined with awkward postures of the fingering and bowing limbs. Execution implies an exquisite temporal and ordinal precision of finger and bow movements, together with a fine-grained force control. It therefore imposes considerable difficulty for beginners and also advanced players.

Acknowledged violinists typically start their career at a young age, i.e. when the brain has a high plastic capacity (Elbert *et al.* 1995; Münte *et al.* 2002). In view of the complexity and spatiotemporal precision of musical skills it is no surprise that research into music performance lags behind studies of perception, with the exception of rhythm studies (e.g. Palmer 1997; Repp 1999). One of the pioneers in this field was W. Trendelenburg, a physiologist in Tübingen and an accomplished cellist, who published a book (Trendelenburg 1925) on playing stringed instruments from a physiological point of view. Among many aspects of playing stringed instruments, he discovered that the initiation of bow movement precedes the development of string vibration by about 50 ms, a problem that became relevant to our work and to which we will turn our attention below. Nicolai Bernstein, a well-known Russian pioneer of biomechanical studies of natural goal-directed movements, developed sophisticated instrumentation to study and assess the complex biomechanics of motor routines in workers. However, an interest in arts and music led Bernstein, with Popova (Bernstein

and Popova 1929), to embark on an extensive study of 14 concert pianists to explore the kinematics and kinetics of piano key strokes. The task was to play, one-handed and at different metronome speeds, some fragments of piano music. Four light bulbs were fixed to a pianist's arm (including the back of the hand) and one on the head. In addition, one bulb was fixed in the workspace serving as a stationary reference point. By means of a specially designed movie camera and a mirror system placed at an angle from the camera, they managed to obtain three-dimensional coordinates of trajectories and their temporal derivatives (chronocyclography). The movie camera had a rhythmic shutter to sample discrete points on a sensitive film. In addition, load impacts on the keys were also measured. The aim was to calculate the interplay of gravity and inertial forces of arm–hand acceleration as a function of the dynamics of the key strokes and of imposed tempo changes. This was a truly novel way of assessing the biomechanics of music performance. Their paper is now little known, probably because it was published in German in an obscure periodical. Later Braitenberg (1986), a neuroscientist and viola player, had similar reflections about the interactions of passive and active forces in bowing, as well as the learned co-planar arm segments of rectilinear bow strokes and bow tilts for string changes.

The many types of bowing, with the large varieties of tempo, loudness, timbre, and accentuations, can be as demanding as the fingering of the left hand. A challenging problem is the asymmetry and also, to a large extent, the independence of bimanual actions in string players. In contrast to keyboard instruments and the harp (with separate strings for each playable note), the violin has only four strings. This means that the full exploitation of tones requires whole-hand translations along the violin neck (termed position changes) as well as frequent string changes. In order to select a new string, the bow has to rotate around the string–bow contact, mainly effected by shoulder elevations and adductions. As well described by Braitenberg (1986), bow strokes, executed parallel to the bridge and co-planar with the upper arm, are constrained by dynamic adjustments of hand abductions and adductions and flexion–extension of the elbow. Bow strokes result in dynamic changes of the bow leverage and therefore require a parallel 'fine-tuning' of compensatory forces (the bow becomes heavier when the hand is nearing the bridge in an up-stroke). Since the loudness of a tone depends on the pressure exerted on the bow (mainly by the index finger), the production of bow forces must integrate both constraints, leverage and loudness.

String players have to 'catch' the correct spatial intervals on the fingerboard. The spatial intervals, corresponding to half-tone intervals along the fingering board, are non-linear and not marked (in contrast to the guitar). Correct positioning with the fingertips for the intended pitch (also called fingerstops and double stops when two fingers are placed simultaneously on two different

strings) is a major stumbling block for amateur string players and requires intensive learning and memorizing over years of practice for correct intonations. Superimposed on the above technicalities one finally, and most importantly, has to add the musical features, i.e. intentional modulations of tempo, loudness, timbre, and even pitch (vibrato), features that essentially mark the individual artist.

In the following we report on our initial biomechanical study in violinists who were asked to repeatedly play a simple tune on a single string. The primary objective was to assess the temporal features of bimanual coordination, i.e. the synchronization between bowing and fingering. In a second step, we also measured the dynamics of bowing and the consistency of the fingering in more complex tone sequences, requiring changes of string and position. These aspects do not include all the features of 'real' professional music. Nonetheless, it is a basic prerequisite to learn about simpler mechanisms. We observed various mechanisms that were present in all players. In addition it was also clear that subtle differences occur among players in terms of strategies, as will be shown below. Achievement of the same goal with the use of variable strategies is a typical feature of volitional goal-directed movements and reflects the principle of motor equivalence (Lashley 1930; Hebb 1949).

## A new approach to assessing the timing of fingering and bowing

We have succeeded for the first time in recording the trajectories of bow and finger movements in violin players (Baader *et al.* 2005). For this we used a movement analysis system with the objective of measuring synchronization of the two limbs. To obtain trajectories of the playing fingers (index, middle, ring, and little), reflecting markers were affixed dorsally, close to the fingertips. With respect to bowing, we did not record the multi-articulate movements of the right arm segments but rather the resulting bow trajectories. Therefore, one marker, or three markers for three-dimensional recordings, were fastened on the bow to obtain bow trajectories. Initially, the reflecting moving markers were recorded with one video camera (Elite, 100 Hz, infrared strobe, example shown in Fig. 7.2); more recently, a three-camera-system (Vicon, 200 Hz) was used for improved three-dimensional recordings of markers (Plate 1 and Fig. 7.3). Unavoidable whole-body movements distort the proper finger and bow trajectories. Therefore, additional reflecting markers were placed on the violin in all recordings in order to transform room-centred coordinates into violin-centred coordinates. We constrained the musical passage to a simple tune played on one string, without position changes of the fingering hand. Each tone was produced by a new *détaché* bow stroke in an ascending or descending sequence. This

enabled us, in a relatively simple way, to approach our goal of assessing bimanual coordination in the context of an asymmetric, skillful, and goal-oriented musical task. Nevertheless, the present task was still based on long-term learning of bowing and fingering and should be distinguished from frequently used bimanual rhythms, such as simple in-phase/anti-phase finger tapping (see Obhi 2004).

Figure 7.1(A) shows the markers on the finger tips providing trajectories of the left-hand fingers, three markers on the bow for bow trajectories, and three markers on the violin for the coordinate transformation. Figure 7.1(B) shows the corrected finger and bow trajectories for a simple sequence played on the D string. The quasi-straight bow strokes and the movements of the fingers, as they tap on the string, are clearly distinguishable.

Figure 7.2(A) illustrates the finger and bow movements of two amateur players who repeatedly played the tune on the D string; the displacement profiles (position versus time) from 10 repetitions were superimposed and aligned at mid-sequence. In an ascending sequence, one can clearly discern when a given finger produced a tone by depressing the string, also termed the fingerstop, i.e. when a short horizontal profile coincides with the horizontal lines indicating the string–fingerboard position. Following the small horizontal bars (below string position), one can easily verify the correct tone sequence corresponding with the musical notation (7.2B). The notes E, F, G and A were played, respectively, by index, middle, ring, and little fingers, whereas tone D was played on the empty D string, i.e. with all fingers lifted (shaded intervals). As in all players, the records of both these players were remarkably consistent, even in the smallest details of the profiles.

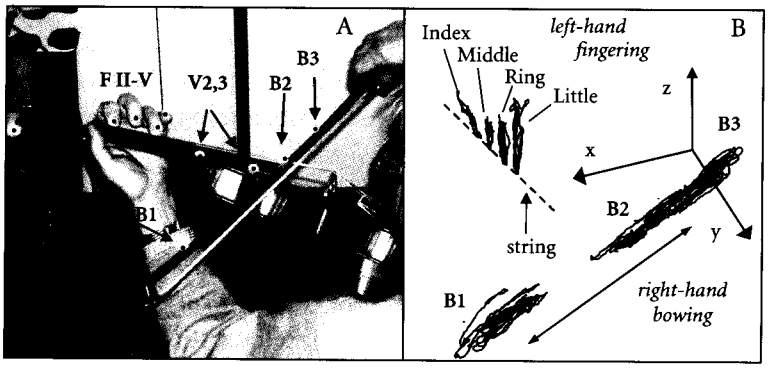



**Fig. 7.1** (A) Set-up of reflecting markers for three-dimensional recordings of bow movements (B1, B2, B3) and of finger movements (F II–V). Markers V1, V2, and V3 are used as reference for transforming finger and bow coordinates into violin-centred coordinates. (B) Finger and bow trajectories after coordinate transformation.

enabled us, in a relatively simple way, to approach our goal of assessing bimanual coordination in the context of an asymmetric, skillful, and goal-oriented musical task. Nevertheless, the present task was still based on long-term learning of bowing and fingering and should be distinguished from frequently used bimanual rhythms, such as simple in-phase/anti-phase finger tapping (see Obhi 2004).

Figure 7.1(A) shows the markers on the finger tips providing trajectories of the left-hand fingers, three markers on the bow for bow trajectories, and three markers on the violin for the coordinate transformation. Figure 7.1(B) shows the corrected finger and bow trajectories for a simple sequence played on the D string. The quasi-straight bow strokes and the movements of the fingers, as they tap on the string, are clearly distinguishable.

Figure 7.2(A) illustrates the finger and bow movements of two amateur players who repeatedly played the tune on the D string; the displacement profiles (position versus time) from 10 repetitions were superimposed and aligned at mid-sequence. In an ascending sequence, one can clearly discern when a given finger produced a tone by depressing the string, also termed the fingerstop, i.e. when a short horizontal profile coincides with the horizontal lines indicating the string–fingerboard position. Following the small horizontal bars (below string position), one can easily verify the correct tone sequence corresponding with the musical notation (7.2B). The notes E, F, G and A were played, respectively, by index, middle, ring, and little fingers, whereas tone D was played on the empty D string, i.e. with all fingers lifted (shaded intervals). As in all players, the records of both these players were remarkably consistent, even in the smallest details of the profiles.

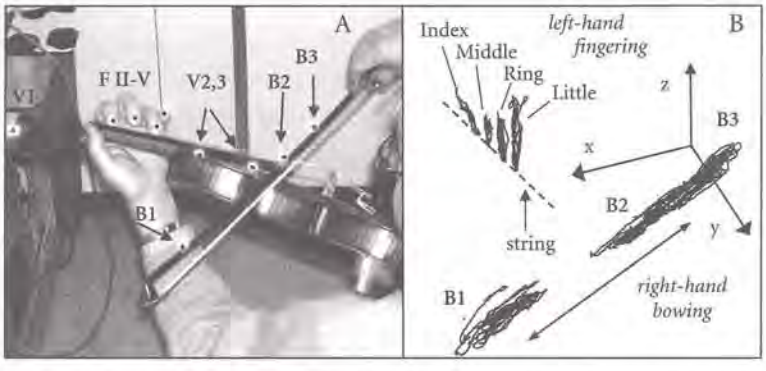



**Fig. 7.1** (A) Set-up of reflecting markers for three-dimensional recordings of bow movements (B1, B2, B3) and of finger movements (F II–V). Markers V1, V2, and V3 are used as reference for transforming finger and bow coordinates into violin-centred coordinates. (B) Finger and bow trajectories after coordinate transformation.

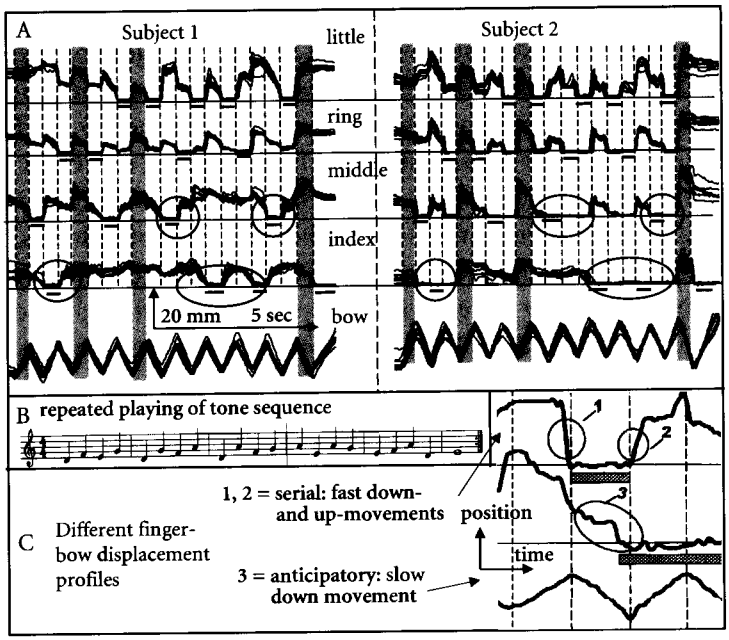



**Fig. 7.2** Finger and bow displacement profiles (position versus time) in two subjects playing a simple tune (bottom). All notes are played on the D string; the board/string position is shown separately for each finger as a horizontal line. Fingerstops are underlined (black bars) and allow one to identify the correct sequence of the tune. Horizontal segments of finger displacement traces without string contact reflect the 'enslaving' of neighbouring fingers momentarily not engaged in tone production. The vertical dashed lines mark the sequence of bow reversals and of tone changes, and the vertical grey columns indicate that the bow stroke was on an empty string (note D, i.e. all fingers lifted).

A number of additional features were observed: for example the first two flat intervals of the little finger did not coincide with the string position (Fig. 7.2A), which means that no tone was produced; rather it reflects the tendency of the fingers to move and stop together, especially with the neighbouring tone-producing action finger (in this case the ring finger). This is a common feature, termed 'enslaving' (Zatsiorsky *et al.* 2000), which also occurs in other manipulatory actions.

Whereas there was a considerable consistency in the repeated sequencing of the fingering in a given subject, comparison between individuals revealed differences in strategies, as highlighted in Fig. 7.2A. In contrast to subject 1, subject 2 remained on the string whenever possible when another finger above the string had already occupied the string (see marked index finger). This difference does not affect the tone sequence, but to be prepared on the string for a renewed action is a useful strategy, often required in etudes for string instruments.

Another general feature, illustrated in Fig. 7.2(C), is that, in a step from a higher tone (A, little finger) to a lower tone (F, middle finger), the latter finger is placed in *anticipation*, i.e. before the upper finger is lifted. Thus, it is the *lift* of the

upper finger that determines the timing of tone *onset*, whereas the anticipating finger below controls *pitch*, but not timing. Combined actions of two fingers also occur when, in an ascending sequence on the same string, a tone is followed by a higher tone (e.g. index to middle finger). In this purely serial action, the duration of the first tone is curtailed by the next tone. We concluded that, depending on the sequence, timing and pitch selection may consist of a combination of actions of two fingers and that pure serial organization coexists with parallel anticipatory mechanisms. Serial and parallel anticipatory processes also occur in speaking. Thus, in pronouncing the word '*construe*', the lips become rounded well in anticipation to the uttering of the vowel 'u', a phenomenon that had been called *co-articulation* (Rosenbaum 1991). The intervals between anticipatory placing of one finger and bow reversal were measured separately and were significantly longer than the true synchronization intervals of serial steps (described separately below). The mean synchronization intervals were typically below 50 ms whereas co-articulation intervals (both intervals measured between finger placing and bow reversal) varied much more, and were typically between 50 and 100 ms.

Our principal aim was to assess how well the bow action was synchronized with the tone-initiating finger actions. Since we used one bow stroke for each tone, we could measure the interval between bow reversal and the fingerstop (for a clear detection of the latter event, we used the velocity peak that occurred just at impact of the finger on the board). In short, we found that synchronization did not correlate with the proficiency of a player. The 'error' interval (the time difference between bow reversal and the peak finger velocity for each fingerstop) varied among subjects ($-10$ ms to $+70$ ms). However, the standard deviation was less variable among subjects and was independent of the mean. It was therefore concluded that the observed standard deviation (SD) of 10–30 ms was an adequate marker of finger–bow synchronization. The SD about each subject's mean time interval, also termed the variable error (VE; see, for example, Schmidt (1988)), was regarded as the time measure of most interest.

What might be the reason for the observed temporal imprecision? Trendelenburg (1925) measured string vibrations by means of a string galvanometer. He observed that onset of vibration lagged behind initiation of the bow stroke by about 40–50 ms. Also, it is noteworthy that the full acoustic spectrum of a tone (responsible for the timbre) generated by a bow stroke takes about 30–40 ms to develop (Deutsch 1985). Synchronization among players in a professional string ensemble was found to deviate by a mean of 50 ms (Rasch 1988). Similarly, when subjects had to tap at a given metronome rhythm, departures from synchronization, up to about 50 ms, were not perceived (Fraisse 1980). Taken together, these observations suggest that, at least in string players, bimanual synchronization errors as small as 50–60 ms are tolerated in terms of auditory perception.

## Bowing movements

### A short résumé of previous work

The essay of Braitenberg (1986) on the issue of the principles of motor control in bowing has been mentioned briefly above. Nelson (1983) analysed kinematic bowing data from a professional violinist, discussing how velocity data can demonstrate a minimum energy demand in bowing. Other studies were made to serve as a database for overuse syndromes in musicians (Tulchinsky and Riolo 1994; Kihira *et al.* 1995; Turner-Stokes and Reid 1999; Shan and Visentin 2003; Visentin and Shan 2003). Of particular interest are observations made in professional cellists playing short excerpts from pieces by Schubert and Brahms (Winold *et al.* 1994). From bow, wrist, and elbow movements, the authors analysed the temporal derivatives to obtain some dynamic aspects of music execution, such as musical accentuations (*articulations*), which were prominent in the rapid succession of notes in the Schubert piece, and were reflected particularly in prominent acceleration peaks in the elbow and wrist at directional changes of the bow. Interestingly, the slower tempo of the Brahms fragment produced an acceleration peak in the middle of the long strokes. It was suggested that this pulse might be a signal for temporally coordinating the finger action of the left hand. This led the authors to suggest '. . . that some aspects of the coordination and control of cello bowing were constrained by the mechanical and energetic demands of the bowing'. The idea was that temporal coordination of the musical aspects was constrained by dynamic rules.

### Principal types of bow movements

Among the vast variety of musical bow strokes, such as forte, piano, *détaché*, staccato, and spiccato, we analysed the strokes for setting the strings into vibration and the tilts for string changes. Plate 1 (A1, A2) illustrate, respectively, the two principal types of bow movements. Three reflecting markers were stuck along the length of the wooden part of the bow, near the frog (grasped by the right hand), at about the middle of the bow, and near the tip. Trajectories of the first two markers (red and green) overlap, whereas the third marker (blue) does not in this case. The trajectories are parallel to the bridge and are the result of a complex, learned coordination of forces acting on the shoulder, elbow, and wrist (see also Braitenberg 1986). The strings are supported by the curved bridge. String changes therefore necessitate a discrete rotational movement around the contact point between the bow and the string. For changing from a lower string to a higher string, the tip end of the bow goes up and the frog end down with the respective movement of the arm (mainly shoulder down in this case until

contact with the desired string). The two principal types of movements are well distinguished in Plate 1: the straight bow strokes (at a right angle to the strings) are best seen from the top (A1), whereas the curved rotational movements, resulting in angular changes of bow strokes around the rotation point, are best seen in a front view (A2).

In Plate 1(B), the kinematics of slow *détaché* bow strokes (i.e. separate strokes for each note) are shown as triangular displacement profiles. Confirming the work of Winold *et al.* (1994), the velocity profiles showed marked oscillations, mainly at bow reversals. Typical frequencies of the oscillations were 6–8 Hz, as revealed by fast Fourier transforms. Note that for these slow bow movements velocity profiles were nearly horizontal, but with fast bow reversals. Correspondingly, in *phase-plane plots* of slow bow movements, the stroke phases were near horizontal (near constant velocity), changing into fast, i.e. near-vertical, reversal phases. With increasingly faster bow strokes, the phase-plane progressively changed into circular profiles (not shown). As concluded by Winold *et al.* (1994), the phase-plane plots characterize rather well the dynamics of bowing. In Plate 1(C1), rotational movement displacements are shown in terms of bow angle versus time. The scales of bow angles were from the highest string E, to strings A, D, and to the lowest G string. The steepness of the slopes was similar (except perhaps for the greatest angle, from string E to string G). The time used for these angular movements was less than 100 ms (C1) and of course was shorter with jumps over fewer strings.

We have mentioned above that a fingerstop was synchronized with the onset of a new stroke (the stroke reversal point). Now the question arises whether coordination of the two hands during playing is also influenced by bow/string changes. Preliminary data show that contact between the finger and the string typically occurred at the middle of the tilt movement, as shown in Plate 1(C2). Perceptually, there was no noticeable auditory sound gap. This is not surprising given the fact that vibrations fade out gradually as the bow leaves the string. The interval at which the vibrations of the previous string stop and those of the next string are initiated was obviously sufficiently short (50 ms or less) not to disturb the flow of the melody. It remains to be seen whether the synchronization between fingerstop and bow reversal is influenced by an additional tilt movement of the bow. Finally, it is worth mentioning that in legato playing (several tones played within one bow stroke), bimanual coordination, as described above, is not taking place. However, peaks of bowing pressure undoubtedly contribute in the accentuation of some of the fingerstops, as suggested by Winold *et al.* (1994). There can be no doubt that kinetic mechanisms of bowing play a considerable role in producing musical accentuations and therefore also in bimanual coordination.

## Bow movements and forces

The large variety of finger and bow movements, as well as posture, is driven by a wide range of forces. The fine control of these forces is an important and determining factor of the character of the music. For example, the loudness of tones depends on the pressure exerted on the string by the bow stroke. We have recently developed a kinetic approach to bowing. As alluded to earlier, bowing requires an exquisite force control for the grip exerted at the frog, mainly by the thumb and middle finger. The index finger has the task of adapting the pressure exerted on the bow to modulate the loudness and accentuations, but simultaneously has also to take into account the ever-changing bow lever, demanding compensatory pressure changes by the index finger in front of the rotation point at the frog and by the little finger behind the rotation point (Galamian 1983). For a down-stroke of the bow of constant loudness (from the frog end to the tip of the bow), the index finger has to progressively recruit more pressure on the bow and to diminish the pressure for an up-stroke. The handling of bow forces has been well described (but not measured) from the point of view of motor control by Trendelenburg (1925).

Our aim was to investigate the interplay of the pressure exerted on the bow by the index finger ('*bow pressure*') and the grip force ('*bow grip force*') exerted mainly between thumb and the middle finger. In addition, we introduced electromyographic (EMG) recordings from the first dorsal interosseus muscle of the right bowing hand. Without going into details of our method, the principle of measurement will be briefly summarized. The kinematics of fingering and bowing was processed by a Vicon system (three cameras, 200 Hz). For bow pressure (Bow-PR), we fastened three fine, soft air-filled rubber tubes onto the dorsal bow surface at the frog end. These were tightly sealed and connected to a low-pressure silicon sensor (piezoresistor/Wheatstone bridge), fastened to a small plate on the back of the hand. Another fine rubber tube was wrapped three times around the frog for sensing the grip force (Bow-GF) exerted mainly by the thumb and middle finger. This rubber tube was also tightly sealed at the end and connected to a second pressure-sensitive gauge on the back of the hand. The electrical output signals were led via one flexible tube to the respective inputs for amplification and for half-wave rectification of EMG signals (100–5000 Hz). All calibrated and conditioned signals were processed by an SC/ZOOM interface and special software for signal capture and post-processing (developed at the Physiology Institute, Umeå). The sampling rate was 800 Hz for EMG and 400 Hz for force.

A number of musical passages were tested, varying from slow *détaché* bow strokes on one string to fast and *martelé* strokes over four strings. In Fig. 7.3, part of an Allegro vivace, shown at the bottom of Fig. 7.3, was taken from a violin

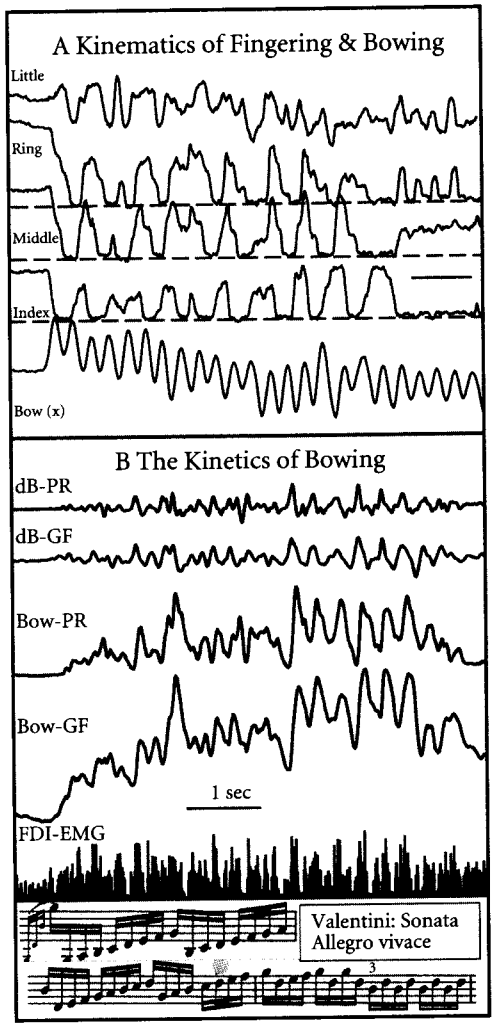



**Fig. 7.3** Comparison of the kinematics (A) and the kinetics (B) while playing the initial bars of a Valentini sonata consisting of fast semi-quavers (4 × 4 semi-quavers per bar); all notes played with separate *détaché* bow strokes with string and finger position changes. The instruction is to start piano and to gradually build up the tone intensity (crescendo) to a maximum, i.e. fortissimo. FDI-EMG, electromyogram of the first dorsal interosseus of the bow hand, which contributes to the pressure exerted by the index finger on the bow. Bow-PR, total pressure exerted by index finger on the wooden part of the bow near the frog. Bow-GF, grip force exerted on the frog by the middle finger and the thumb; dB-PR and dB-GF are the rates of the respective forces. Parts (A) and (B) were not recorded simultaneously and are thus not exactly timed equally.

sonata of Giuseppe Valentini (Idea per Camera, *c.*1705) in order to illustrate the kinematics of fingering and bowing (A), as well as the kinetics of bowing (B, recorded separately from A). The musical segment was played repeatedly and consisted of 48 rapid bow strokes and tones, following in groups of 4 × 4

semi-quavers (1/16). The kinematics illustrate the sequencing of fingers and bow strokes. The sequence starts with an up-stroke on the empty string (tone G) and is followed by tones A, B, C, D with, respectively, the actions of the index, middle, and ring fingers. Note that in this initial ascending sequence the index finger remains on the string when the middle finger is already tapping the string. This is again a clear demonstration of serial processing, with a new finger action simultaneously terminating the previous one, and also of an economic strategy to leave the lower fingers on the string (as described earlier). Another example of 'economy' can be observed toward the end when the index finger remains flat on the string while the ring finger goes down and up four times, oscillating between tones B–D–B–D, and so on. It is the ring finger alone that is responsible for the timing of the two tones (this is a similar situation to oscillating between two neighbouring fingers to produce trills). The regular timing of bow strokes contrasts with some irregularities in amplitude that are subtle expressions of accentuations. The little finger was rarely used in this sequence. However, it demonstrates clearly the phenomenon of simultaneous movements without tapping the string, mentioned before as 'enslaving'.

In the kinetics of bowing, there is a slow build-up of force corresponding to the gradual increase from piano to forte. While playing forte, the accentuations of single bow strokes (and thus of notes) were expressed as prominent dynamic peaks. A parallel increased segmentation can be seen in the EMG of the first dorsal interosseus muscle (FDI) that plays a prominent role in adjusting bow pressure for loudness and musical accentuation.

## Conclusions and outlook

A biomechanical approach to musical performance is challenging and does not provide quick results. Taking stock of the present results, we nevertheless suggest that even relatively simple musical sequences can reveal interesting principles of motor control, especially about inter-limb coordination and the kinetics of bowing.

A large number of non-musical manipulations, including gestures, are based on the actions of both arms (Wiesendanger and Serrien 2005). With the use of an instrumented drawer manipulandum and weighted objects, we succeeded in obtaining some insight into the construction of bimanual actions. The performance of human skills includes both learned and object-related bimanual operations, like eating with a knife and fork and many other daily actions. This motor behaviour is strongly related to goal achievement (Wiesendanger *et al.* 1994; Perrig *et al.* 1999; Wiesendanger and Serrien 2001*a*,*b*, 2005; Kazennikov *et al.* 2002). We observed the trajectories of individual arms to vary considerably, depending on the constraints (for example the normal light of the workspace versus complete

darkness). However, *goal achievement* was much less variable. Invariant goal achievement is a typical feature of goal-directed behaviour and has been termed *motor equivalence* (Lashley 1933; Abbs and Cole 1987). Progress in understanding some of the rules of bimanual behaviour encouraged us to investigate musical bimanual skills as well.

Playing music, which is deeply rooted in human behaviour, is a particularly interesting, mostly bimanual, sensorimotor behaviour, or rather a higher-order skill, which rests on many years of intensive practice. Three processes are tightly coupled: reading the musical notation, translating it to motor performance, and listening to the resulting music. It is only in advanced players that the three processes 'come together' in an ideal musical experience. Among the frequent and typical shortcomings of amateur players are incorrect intonation, slowing of tempo during difficult fingering, and bad handling of the many bowing characteristics. It is an incredible achievement of the human brain to reach a highly professional level. With the emergence of many technical improvements, the time is ripe for exploring the motor control of musical performance. Thus, we show for the first time the feasibility of assessing the temporal structure of the coordinated, but asymmetric, fingering and bowing of violinists. We also discuss arguments that deviations of up to about 50 ms from ideal synchronization (0 ms) are not perceived. This may well be different in players of other instruments, like drummers.

Another promising start concerned the kinetics of bowing: how do we handle the most intricate control of grip and bow pressure forces? We observed adjustments to the lever changes during bow strokes, as well as the additional controls required for the tonic dynamic adjustments for forte and piano instructions and for the peaks of fast accentuation in the musical flow. Still other basic motor mechanisms may be investigated that may potentially influence the timing, like the translation of left-hand positions along the fingerboard. Understanding music performance in terms of motor control may also be helpful in finding new ways in dealing with the growing number of professional musicians suffering from overuse syndromes. Objective long-term study of motor control in music performance might also be helpful in assessing the therapeutic benefits of treatments for such syndromes.

## References

**Abbs, J.H. and Cole, K.J.** (1987). Neural mechanisms of motor equivalence and goal achievement. In: S.P. Wise (ed.) *Higher brain functions*, pp. 15–43. Wiley, New York.

**Baader, A.P., Kazennikov, O., and Wiesendanger, M.** (2005). Coordination of bowing and fingering in violin playing. *Cogn Brain Res*, **23**, 436–443.

**Bernstein, N. and Popowa, T.** (1929). Untersuchung über die Biodynamik des Klavieranschlags. *Arbeitsphysiologie*, **1**, 396–432.

**Braitenberg, V.** (1986). The cerebellum and the physics of movement: some speculations. In: M. Glickstein and C. Yeo (eds) *Cerebellum and neuronal plasticity*, pp. 193–207. Plenum, New York.

**Deutsch, W.A.** (1985). Musik und Computer. In: H. Bruhn and R. Oerter (eds) *Musikpsychologie. Ein Handbuch in Schlüsselbegriffen*, pp. 107–119. Urban & Schwarzenberg, Munich.

**Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., and Taub, E.** (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, **270**, 305–307.

**Fraisse, P.** (1980). Les synchronisations sensori-motrices aux rythmes. In: J. Requin (ed.) *Anticipation et comportement*, pp. 233–257. CNRS, Paris.

**Galamian, I.** (1983). *Principles of violin playing and teaching*. Prentice-Hall Inc., Englewood Cliffs, NJ.

**Hebb, D.O.** (1949). *Organization of behavior*. Wiley, New York.

**Kazennikov, O., Perrig, S., and Wiesendanger, M.** (2002). Kinematics of a coordinated goal-directed bimanual task. *Behav Brain Res*, **134**, 83–91.

**Kihira, M., Ryu, J., Han, J.S., and Rowen, B.** (1995). Wrist motion analysis in violinists. *Med Probl Perform Art*, **10**, 79–85.

**Lashley, K.S.** (1930). Basic neural mechanisms in behavior. *Psych Rev*, **37**, 1–24.

**Lashley, K.S.** (1933). Integrative functions of the cerebral cortex. *Physiol Rev*, **13**, 1–42.

**Münte, T.F., Altenmüller, E., and Jäncke, L.** (2002). The musician's brain as a model of neuroplasticity. *Nat Rev Neurosci*, **3**, 473–478.

**Nelson, W.L.** (1983). Physical principles for economics of skilled movements. *Biol Cybern*, **46**, 135–147.

**Obhi, S.S.** (2004). Bimanual coordination: an unbalanced field of research. *Motor Control*, **8**, 111–120.

**Palmer, C.** (1997). Music performance. *Ann Rev Psychol*, **48**, 115–138.

**Perrig, S., Kazennikov, O., and Wiesendanger, M.** (1999). Time structure of a goal-directed bimanual skill and its dependence on task constraints. *Behav Brain Res*, **103**, 95–104.

**Rasch, R.A.** (1988). Timing and synchronization in ensemble performance. In: J.A. Sloboda (ed.) *Generative processes in music*, pp. 70–90. Clarendon Press, Oxford.

**Repp, B.H..** (1999). Control of expressive and metronomic timing in pianists. *J Mot Behav*, **31**, 145–164.

**Rosenbaum, D.A.** (1991). Chapter 9: *Human Motor Control*, Speaking and singing. In (ed.) pp. 296–338. Academic Press, San Diego, CA.

**Schmidt, R.A.** (1988). *Motor control and learning*. Human Kinetics, Champaign, IL.

**Shan, G. and Visentin, P.** (2003). A quantitative three-dimensional analysis of arm kinematics in violin performance. *Med Probl Perform Artists*, **18**, 3–10.

**Trendelenburg, W.** (1925). *Die natürlichen Grundlagen der Kunst des Streichinstrumentenspiels*. Springer, Berlin.

**Tulchinsky, E. and Riolo, L.** (1994). A biomechanical motion analysis of the violinist's bow arm. *Med Probl Perform Artists*, **9**, 119–124.

**Turner-Stokes, L. and Reid, K.** (1999). Three-dimensional motion analysis of upper limb movement in the bowing arm of string-playing musicians. *Clin Biomech*, **14**, 426–433.

**Visentin, P. and Shan, G.** (2003). The kinetic characteristics of the bow arm during violin performance: an examination of internal loads as a function of tempo. *Med Probl Perform Artists*, **18**, 91–97.

**Wiesendanger, M. and Serrien, D.J.** (2001*a*). Neurological problems affecting hand dexterity. *Brain Res Rev*, **36**, 161–168.

**Wiesendanger, M. and Serrien, D.J.** (2001*b*). Toward a physiological understanding of human dexterity. *News Physiol Sci*, **16**, 228–233.

**Wiesendanger, M. and Serrien, D.J.** (2005). Bimanual coordination and its disorders. In: H..J. Freund, M. Hallett, M. Jeannerod, and R. Leiguarda (eds) *Higher-order motor disorders—from neuroanatomy and neurobiology to clinical neurology*, pp. 193–235. Oxford University Press, Oxford.

**Wiesendanger, M., Wicki, U., and Rouiller, E.M.** (1994). Are there unifying structures in the brain responsible for interlimb coordination? In: S.P. Swinnen and H. Heuer (eds) *Interlimb coordinaton: neural, dynamical and cognitive constraints*, pp. 179–207. Academic Press, San Diego, CA.

**Winold, H., Thelen, E., and Ulrich, D.B.** (1994). Coordination and control in the bow arm movements of highly skilled cellists. *Ecol Psychol*, **6**, 1–31.

**Zatsiorsky, V.M., Li, Z.M., and Latash, M.L.** (2000). Enslaving effects in multi-finger force production. *Exp Brain Res*, **131**, 187–195.

Chapter 8

# Movements and analysis of drumming

Sofia Dahl

Players of percussion instruments strive to acquire playing techniques that can be adapted to the different feedbacks from the various instruments that are played. Analyses of drummers' performances show flexible wrist movements, moving the stick in a 'whiplash' motion. Skilled players appear to maintain their movement strategy consistently for different playing conditions, but how much of the arm that is involved in the movement differs between players. The dynamic level influences both the height to which the stick is lifted in preparation for a stroke and also the timing of strokes. Intervals beginning with accented strokes tend to be prolonged more at soft than at loud dynamic levels.

## Introduction

It is not hard to get an object to resonate; one only has to hit it. Compared with the amount of practice necessary to get a reasonable sound out of a trumpet or a violin, the act of striking a drum seems almost trivial. And drumming would be trivial if it were not for timing. Music performance is, after all, considered to be one of the most complex forms of skilled serial actions. If striking something is easy, doing so at the right time is not.

While it is outside the scope of this chapter to compare the playing of different types of instruments to each other, it is worth considering that percussion instruments share some properties that distinguish them from most other instruments (like woodwind or strings). It is these very properties that pose some difficulties for the player.

### Diversity—not one instrument but many

Percussion includes a huge variety of instruments, both membranophones (e.g. drums) and idiophones (e.g. cymbals, wood blocks, etc.). Normal percussion playing requires that the player performs the same rhythm on different instruments with differing physical properties (for instance, world-renowned solo percussionist Evelyn Glennie has a collection of 1400 instruments and travels with, on average, 1.7 tons of equipment). A change of instrument changes the kinesthetic

feedback to the player, who must adapt to the conditions at hand. This will lead to many compromises, a single drumstick or mallet may not be equally suitable for all the instruments that are to be played.

### Short contact times

Percussion instruments generally produce sounds with impulse-like characteristics. Normally the onset of the note is well defined, the duration is short, and in general the player has little control over the tone, once initiated. The note can be shortened by dampening (e.g. by forcing the mallet to stay in contact with the drumhead after the hit), but it cannot be lengthened. While, for instance, players of wind instruments have close control of the vibrating air column during the full duration of a note, the percussion player's direct contact with the instrument is limited to a few milliseconds. This implies that whatever the resulting striking force and dampening effect the percussionist wants to induce would need to be integrated in the entire striking gesture. The mallet will strike the drumhead (or some other object) with the velocity and effective mass supplied by the player's movement, and the same striking gesture will also determine the duration of contact.

### Changing pitch and timbre is awkward

In order to modify the timbre or the fundamental frequency a percussion player can strike at different positions, or choose a mallet with a different mass, hardness, and/or shape. The stiffness and tension of a drumhead can also be adjusted. However, in a normal playing situation there is seldom time for such types of adjustment. A change of timbre or pitch is typically achieved by striking another instrument, with its own specific characteristics and limitations. There is a need to plan ahead in order to reach the instruments that are positioned at varying distances from the player.

### Time constraint and need for compromises

The characteristics and properties of percussion instruments can be summarized by three challenges that have to be met. In order to pass as a good, or even acceptable, percussionist a player must:

1. Be able to master the range of expression needed, given the tempo, in terms of dynamics and sound qualities.
2. Strike the right spot at the right moment. This becomes increasingly difficult when there are many different instruments to reach, or when quick changes are necessary between drumsticks and brushes, for example.

3. Be so relaxed that it is possible to maintain the same type of movements for several minutes without fatigue—while still performing with as close control of the stick as possible. For instance, to maintain an illusion of a 'prolonged' tone, a roll, each hand should be able to reproduce close to equal sound levels and timbre.

As for any musician, time constraint is crucial. This can lead to conflicts in several cases, for instance it might not be possible to reach the optimal position in the optimal time. In many respects, percussion playing can be described as a continuous search for a compromise that allows the desired expression within the time given. It could be hypothesized that skilled players could take advantage of the conditions at hand to maximize the desired effects (such as expression, time that can be used for preparation, etc.) and minimize the undesired effects (performance errors, strain, etc.).

## Playing techniques

Players of percussion instruments strive to acquire playing techniques that can be adapted to the feedback from the instrument. Most techniques share a common ground regarding the quality of 'efficient' playing. This general consensus can be formulated as three guidelines on how to facilitate playing:

1. Let the drumstick/mallet do most of the work.
2. Take advantage of the rebound that is given by the instrument.
3. Plan stick positions ahead.

All these points rely on the wrist to work as a hinge, relaxed and flexible. The acquisition of relaxed and flexible movements is one of the main issues in learning drumming techniques. The striking movement should be 'like waving goodbye or bouncing a ball' (Cook 1988). With a cramped grip the drumstick does not have enough freedom to move and accelerate. As a result, the command 'relax!' frequently occurs in instruction manuals.

### Cracking a whip

The resulting excitation of the instrument ('shock spectrum') is related to the history of the contact force. A high velocity of the mallet at impact results in a high and short force pulse. This corresponds to a high sound level and a rich spectrum with strong high-frequency partials. A wooden stick can help a player to excite an instrument more vigorously by increasing the striking velocity at low physiological cost. At low dynamic levels this is easy enough; there will be a sound just by letting the stick fall. As dynamic levels increase more force is needed, something that can be achieved by providing a sufficient 'runway' during which the stick increases its velocity.

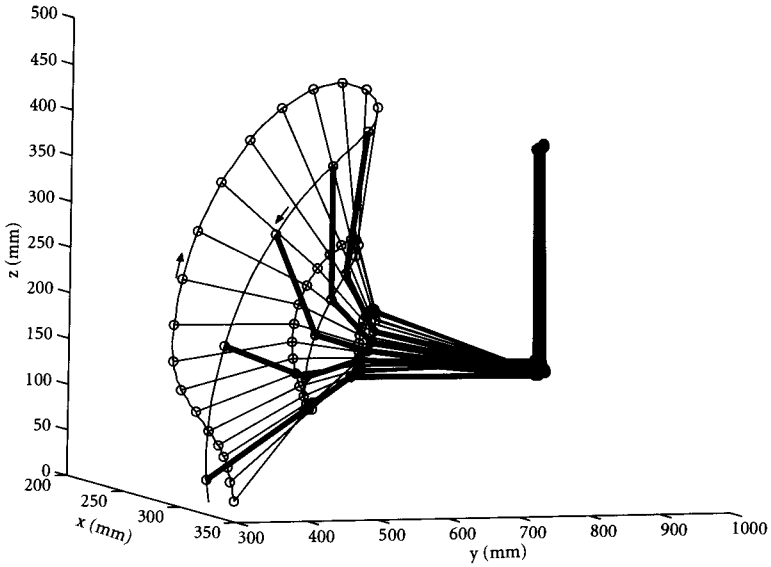



**Fig. 8.1** Schematic representation of a player's arm movement during a stroke. The circles mark the positions of five markers on the drummer's stick, hand, wrist, elbow, and shoulder throughout the movement. The leftmost, largest loop describes the trajectory of the tip of the drumstick, while the two inner trajectories represent, respectively, the path of the hand and wrist. The preparation for the stroke starts with an upward movement of the wrist, dragging the drumstick upward. The stick is flicked upward, surpassing the hand and taking a vertical orientation. Also in the actual down-stroke (here indicated by thicker lines) the wrist leads the downward movement. The drumstick follows in a whiplash manner, gaining high velocity from the elongated arc produced by the preparatory motion. (The figure was generated using three-dimensional data from Waadeland (2003). Time separation between the frames is about 25 ms.)

An example of how a single stroke is played is shown by the 'stick figure' in Fig. 8.1. The figure displays the three-dimensional positions of five markers on a player's stick, hand, wrist, elbow, and shoulder during a stroke. The three loops mark the trajectories of the tip of the drumstick (leftmost loop), the hand (middle loop), and the wrist joint (right loop). The figure shows how the hand and wrist lead the drumstick. The preparation for the stroke is initiated by a raising of the wrist with the stick lagging behind, tip pointing down. The stick is first flicked up to a vertical orientation and then flung down to the drumhead, increasing velocity over the full height of the stroke.

The strategy for prolonging the arc during which acceleration occurs can also be found in other contexts. Players of tennis, baseball, and golf all display similar striking movements when they have time to do so (Wickstrom 1983; Abernethy *et al.* 1997).

## Letting it bounce

A relaxed grip of the drumstick also opens up the possibility of taking advantage of the rebound from the surface. In drumming, particularly in snare drum

playing, the player tries to take advantage of the fact that a normal drumhead is elastic and therefore will, so to speak, 'send the stick back to where it came from' (Shivas 1957). The rebound vouches for pushing the stick away from the surface quicker than the player would him/herself. By learning control of the rebound it is also possible to play more than one stroke per movement gesture, an energy efficient technique that is used for fast rolls.

The use of the rebound leads to a kind of 'working area' with combinations of tempi and dynamics that are easier to master than others. As when bouncing a rubber ball, certain heights and tempi fit well with the natural periodicity of the ball. However, as one forces the ball closer to the ground it becomes increasingly difficult to maintain a slow tempo. Similarly, at soft dynamic levels, the rebound from the drumhead can be a bit too soft to give the desired 'flow' of movement at some tempi. On the other hand, loud dynamic levels give a rebound which can be too strong to master. Both these cases impose difficulties that have to be dealt with within the given time constraint. At slow tempi, at *pp* level, the drumstick must be held steady and away from the drumhead before it is time for the next stroke, while in the *ff* case, the upward recoil of the stick may need to be restrained if the next downward movement is to begin in time. Therefore very fast playing at loud sound levels becomes difficult. And, as a result, pushing the boundaries to increase the working area of playing demands more control from the player.

## Planning ahead

By initiating a transition between one position and another as soon as possible, the player strives to maximize the time between events. One example of how this is done can be found in the playing of accented strokes. In drumming, one of the most commonly used features is to accent a stroke by increasing loudness, and much time and effort is spent mastering the playing of accented rhythm patterns.

Techniques are taught to facilitate the playing of strokes at different dynamic levels by specifying at what height the stick should be *at the end* of the stroke. The terms, tap, up-stroke, down-stroke, and full stroke are commonly used to help the performer plan and carry out the right movements (see, for example, Moeller 1956; Famularo 1999). Soft strokes are generally played just by snapping the wrist without raising the arm ('tapping'). The hand typically stays in a resting position 10–20 cm above the drumhead and the drumstick barely rises above it. 'The tap differs from the stroke in that it is a *wrist* action, and does not involve the arms to any appreciable extent. A large percentage of snare drum music is played using only alternated taps, the strokes being reserved for very heavy sounds and accents. (Holloway and Bartlett 1973).

Prior to a strong blow, the hand must lift the drumstick after the preceding tap. Here the rebound from the drumhead is a welcome assistance when the hand and drumstick need to rise to a higher position ('upstroke' or 'pull-out'), from where the stroke is delivered. When strong strokes are repeatedly played the stick is usually allowed to bounce up to full height ('full' or 'free' stroke). If the following stroke is to be a soft tap the stick must end up in a suitably lower position ('downstroke' or 'control stroke').

Using these basic strokes the player can execute sequential strokes of different heights. This is a prerequisite for playing various combinations of strokes at different dynamic levels, including accents.

## Data from empirical studies

Up until recently, the study of drummers' movements has not been a very active area. At present, there are only a few publications available. Nevertheless, those that exist provide information about some of the prominent features of drumming performance and they will be presented here in relation to the guidelines for playing discussed above.

### Movement strategies among players

To begin with, there are clear differences between skilled and unskilled drummers. Trappe *et al.* (1998) compared the performances of non-drummers, beginners, students, and professional drummers. The motion patterns of the professionals were found to be flexible and whiplash-like. The students showed similar patterns, but the calculated angles between segments (drumstick, hand, lower and upper arm) showed less control of the stick compared with the professionals. Compared with the drummers, the wrist movements of the non-drummers were stiffer and less flexible. Among the non-professionals there seemed to be a tendency for reducing the degrees of freedom in different joints by co-activating agonists and antagonists.

In studies of interleaved accents for different playing conditions Dahl (2000, 2004) found that players used different movement strategies. The overall patterns of the movement strategies were maintained consistently for different tempi, surfaces and dynamic levels. However, the height to which the stick was lifted in preparation for a stroke and the vertical velocity of the stick marker at impact were both strongly linked to the dynamic level. There was also some influence of the type of surface that was struck; a soft pillow, a runner disk similar to a normal practice pad, or a hard, rigid metal plate (covered with a thin sheet of rubber).

Figure 8.2 shows the movement trajectories for three of the four players in Dahl's study performing a repeated pattern with every fourth stroke accented. The figure displays an excerpt of approximately four measures at *mf*, 200 beats

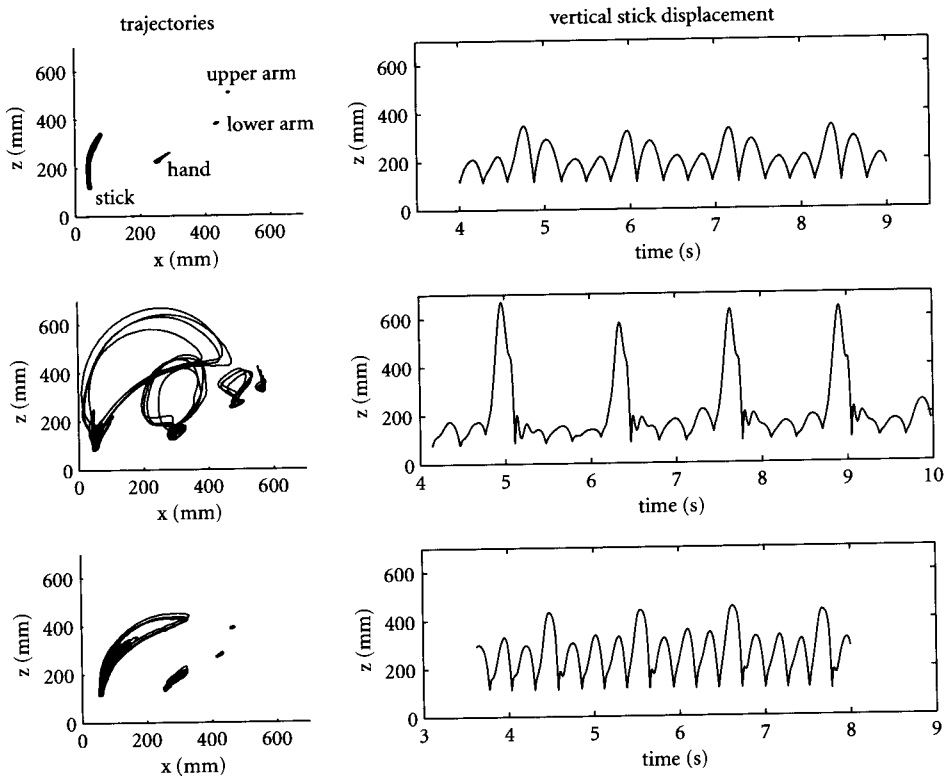



**Fig. 8.2** Measurements of movement strategies for playing a repeated pattern of single strokes with interleaved accents every fourth note. The panels in the left column display the movement trajectories for four markers on three players' sticks, hands, and lower and upper arms. The trajectories are viewed from the players' left side, vertical direction (*z*-axis), and horizontal direction (*x*-axis). The accented strokes are initiated from a greater height than are the unaccented strokes. These increased preparatory heights can be seen as larger loops in the trajectories. The right column shows the vertical displacements of the stick marker versus time. The hits are seen as sharp dips in the curves. The increased preparatory heights before every fourth, accented stroke are clearly seen. The three players have their main backgrounds from different musical contexts: symphonic percussion (top row), jazz drumset playing (middle), and military drumming (bottom). The movement strategies of the three players differ, not only in how much of the arm is involved in the accented stroke, but also how high the stick is allowed to rebound after the accented stroke (Dahl, 2004).

per minute (bpm), played on the rubber disc surface. The panels in the left column display the sideways view of the motion trajectories, while the right column panels show the vertical component of the stick trajectory versus time.

The preparation for the accents can be seen both as larger loops in the three-dimensional trajectories and as higher peaks in the vertical component, compared with those of the unaccented strokes. The players' drumsticks, hands, and lower and upper arms are involved to different extents in the movements. The

three professional players have their main experiences from different musical backgrounds: Symphonic percussion (S1, top row), jazz drumset playing (S2, middle), and military drumming (S3, bottom). All the players performed with their right dominant hand.

In particular three things are to be noted from the trajectories:

1. All players initiate the accented stroke from a greater height than the unaccented. However, the interpretation of the accent, in terms of how strongly it should be emphasized, differs between players.
2. The players handle the strong rebound after the accented stroke in different ways. Player S1 (top) allows the stick to rebound fully, incorporating the rebound in the following preparation, causing it to be almost as high as that for the accented stroke. Players S2 and S3 (middle and bottom) let the rebound after the accent be dampened by the hand, resulting in some oscillation. Restraint of the stick is necessary if it is to end up in a suitable position before the subsequent stroke (see 'down-stroke' in the previous section).
3. In order to reach the high position in preparation for the accent, player S2 (middle) actually initiates the upward motion before the preceding stroke. The angling of the stick at the hit before the accent can be seen as an upward shift in the displacement curve. The dip in the curve appears as sharp as for the other unaccented strokes, but because the marker was placed some distance from the tip of the drumstick it becomes evident that the preparation for the accented stroke has already begun.

The technique used with the large preparatory motion of player S2 is sometimes referred to as a 'Moeller stroke' (Moeller 1956; Famularo 1999). Even though the movement involves the whole arm, raising the stick well above the shoulder, the accented stroke is not delayed. By starting the upward movement before the preceding (unaccented) stroke the player ensures that there is ample time to reach the higher position. The player lets the tip of the stick fall down on the drumhead, while at the same time the end of the stick is moved up by the hand. To drag the stick up to such a height may seem elaborate, but it clearly serves a purpose in producing high velocities at impact for the accented strokes. In the vertical displacement, it can be seen how the stick quickly accelerates in the downstroke from the turning point.

## Striking velocities

The characteristics of the players' individual movement patterns in Dahl's study (Dahl 2004) were reflected in the measured striking velocity. The heights to which the stick was lifted in preparation for the strokes were highly correlated with the vertical velocity of the drumstick marker at impact (correlation coefficients between 0.92 and 0.96).

The condition with most influence on the movement patterns was the dynamic level, resulting in higher striking velocities at higher dynamic levels. When comparing different striking surfaces, the players tended to increase striking velocity when playing on the soft surface, and to decrease striking velocity for the hard surface.

Both the height from which the strokes were initiated and the striking velocities showed that the main difference between the playing styles of the drummers was the emphasis on the accented stroke as compared with the unaccented stroke. For instance, players S1 and S2 produced similar average striking velocities for the unaccented strokes, but while S1 played the accented strokes on average with a velocity 1.5 times higher than for unaccented strokes, the striking velocity for S2's accented strokes was almost five times that for unaccented.

Figure 8.3 shows a comparison between how players S1 (top) and S2 (bottom) emphasized the accented compared with the unaccented stroke, at different tempi and dynamic levels. The figure shows linear planes fitted to the measured striking velocities for unaccented strokes (bottom plane) and accented strokes (top plane) for the two players performing on the rubber surface. As illustrated by the different inclination of the planes in the figure, tempo and dynamic level have different influences on the players' emphasis on the accented strokes.

One possible explanation for the large differences in the players' interpretation of the accent can be sought in the different backgrounds of the players. Players S1 and S3 are mainly active in the symphonic and military orchestral traditions, while S2 mainly plays drumset in the Afro-American music tradition. In orchestral playing, an accent, although emphasized, should not be *over* emphasized. In contrast, many genres using drumsets often encourage large dynamic differences between accented and unaccented beats. In fact, unaccented notes are sometimes played as soft as possible 'ghost notes' (Dahl 2004).

## Relationship to timing

Since the purpose of many of the movement strategies in the playing of percussion is to arrive on time, analysis of the timing performance is an important, but precarious, concern. As in all musical performance, it is difficult to separate the timing variations intended for expressive reasons from unintended variations such as drift in tempo and occasional displacements of individual tones (e.g. Gabrielsson 1999; Clarke 1999).

In analysis of the timing of isochronous tapping sequences performed by three professional drummers, Madison (2000) observed an average standard deviation of 2.8% of the mean inter-onset interval (IOI). Trappe *et al.* (submitted) reported differences in temporal accuracy corresponding to the level of professional expertise. While the standard deviations for the expert players were about 3% (or 6 ms for an IOI of 198 ms), the beginners displayed deviations of more

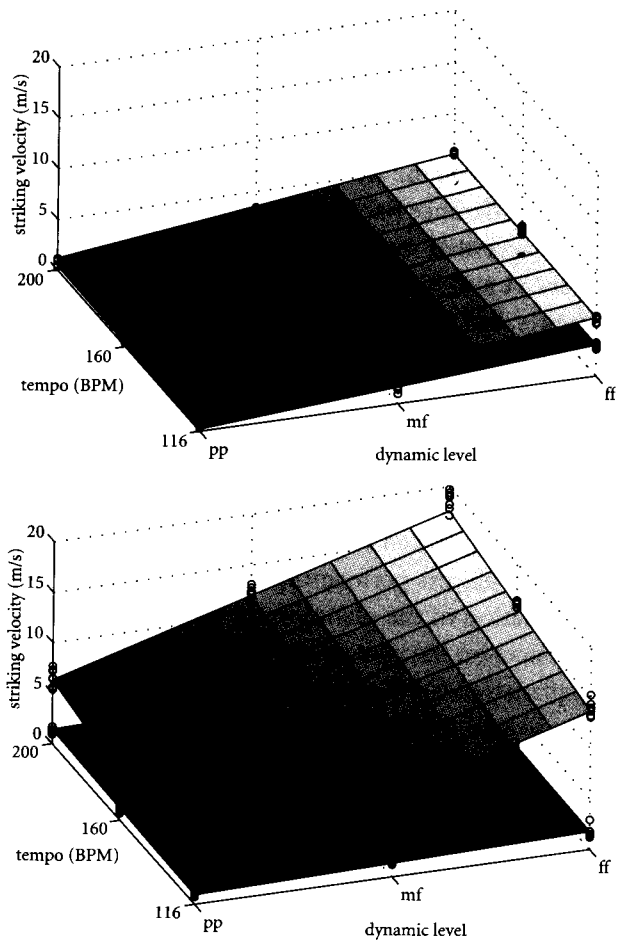



**Fig. 8.3** Striking velocities of a marker on the drumstick for player S1 (top) and S2 (bottom panel) playing at three tempi and three dynamic levels on a rubber surface. The graphs show a linear plane fitted to the measured striking velocities for unaccented strokes (bottom plane) and accented strokes (top plane). The measured data points are indicated by circles. The fit of a linear plane to the data is not optimal, but made to simplify comparisons. The velocities are calculated from the displacement data of a marker on the drumstick and, as the marker was placed some distance in on the stick, the velocity at the very tip of the drumstick would be about 30% higher.

than double this figure (7% or 14 ms for the same IOI). Between students and expert players there were only slight differences in the timing performance.

For the playing of accented patterns, Dahl (2004) reported that the timing deviations were generally higher for playing at soft dynamic levels compared with loud playing. Although the accented strokes were initiated from a greater height than unaccented strokes, they arrived 'on time'. Instead, the following stroke was delayed, resulting in a lengthening of the interval beginning with the accented stroke. The average lengthening reached 3% (across all subjects and

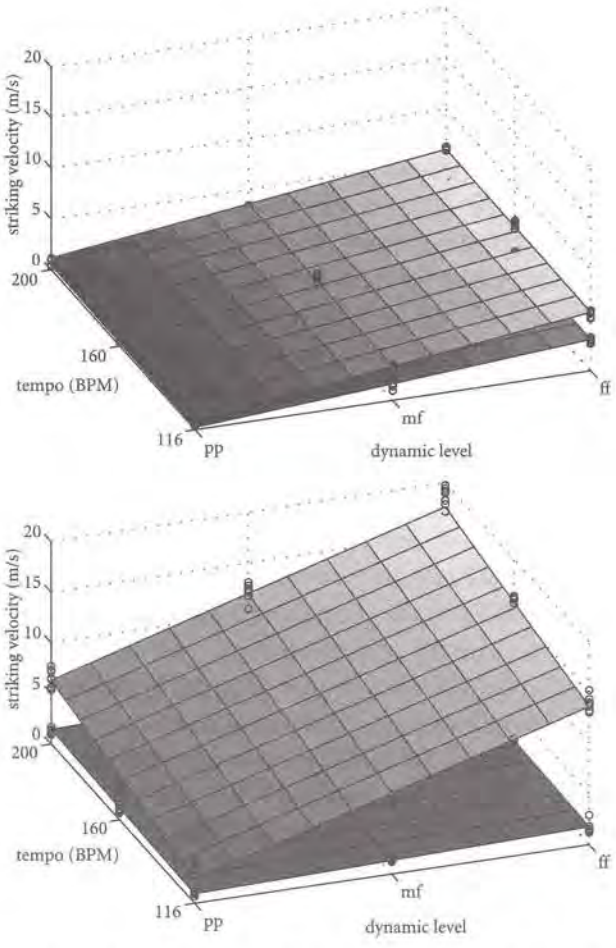



**Fig. 8.3** Striking velocities of a marker on the drumstick for player S1 (top) and S2 (bottom panel) playing at three tempi and three dynamic levels on a rubber surface. The graphs show a linear plane fitted to the measured striking velocities for unaccented strokes (bottom plane) and accented strokes (top plane). The measured data points are indicated by circles. The fit of a linear plane to the data is not optimal, but made to simplify comparisons. The velocities are calculated from the displacement data of a marker on the drumstick and, as the marker was placed some distance in on the stick, the velocity at the very tip of the drumstick would be about 30% higher.

than double this figure (7% or 14 ms for the same IOI). Between students and expert players there were only slight differences in the timing performance.

For the playing of accented patterns, Dahl (2004) reported that the timing deviations were generally higher for playing at soft dynamic levels compared with loud playing. Although the accented strokes were initiated from a greater height than unaccented strokes, they arrived 'on time'. Instead, the following stroke was delayed, resulting in a lengthening of the interval beginning with the accented stroke. The average lengthening reached 3% (across all subjects and

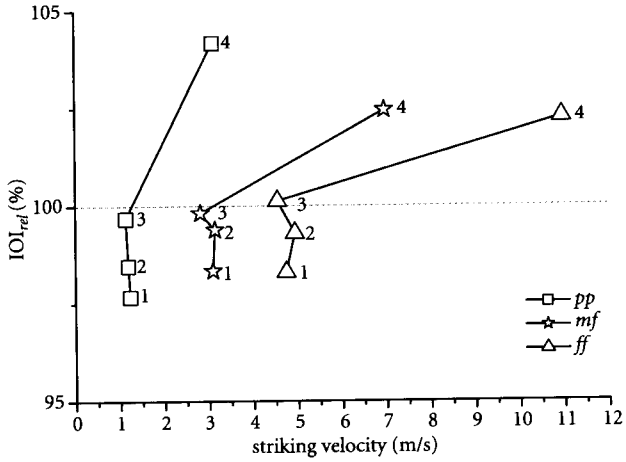



**Fig. 8.4** Relation between striking velocity and relative inter-onset interval ($IOI_{rel}$) for different metric locations and dynamic levels. The figure shows the grand average according to the position of the strokes in the measure (numbered 1, 2, 3, and 4) at *pp* (squares), *mf* (stars), and *ff* (triangles). The data points designated 1 represent all first strokes in each measure averaged across all tempi, subjects, and surfaces, and so forth. The intervals beginning with the accented strokes (4) are lengthened at the expense of the two following, unaccented, intervals. The striking velocity, here used as an indicator of excitation strength (dynamic level), is defined by the motion of the drumstick marker. The actual velocity at the tip of the drumstick would be about 30% higher. The increase in striking velocity for the accented strokes is by about a factor of two for all dynamic levels. The lengthening in $IOI_{rel}$, on the other hand, decreases with increasing dynamic level (Dahl 2004).

playing conditions). A prolonged interval after an accented note is a common feature in musical performance (e.g. Drake and Palmer 1993; Gabrielsson 1999), but has also been found for accents in finger tapping (Billon *et al.* 1996). For the drummers, the lengthening not only emphasizes the accent but also allows the oscillations of the rebounding stick to decay (cf. Fig. 8.2).

Figure 8.4 illustrates the influence of accented and unaccented strokes on striking velocity and relative inter-onset interval ($IOI_{rel}$, defined as IOI divided by the mean IOI across eight measures) for different dynamic levels (Dahl 2004). The figure, in which each data point is averaged across 288 strokes, shows the mean $IOI_{rel}$ versus striking velocity of the drumstick marker for the three unaccented strokes (numbered 1, 2, and 3) and the accented stroke (4) at three dynamic levels. As seen in the figure, the lengthening of the interval beginning with the accented stroke is more pronounced at a soft dynamic level. At *pp* (squares) the average lengthening is about 4%, compared with just above 2% at *mf* (stars) and *ff* (triangles). The interval ending with the accented stroke (stroke 3) is of the same length as the mean IOI across the sequence (100%). However, it is played with slightly less striking velocity, an effect of the early initiated upward movement of the stick some of the players use in preparation for the accented blow.

In a preliminary analysis of drummers performing jazz swing patterns Waadeland (2003) reported that the short notes on a cymbal were played using only one striking movement of the hand. Again, this points towards a strategy of letting a loosely held stick rebound and fall back to strike the cymbal again a second time. As tempo increases, the grip of the stick needs to be adjusted so that the time interval between the two strokes decreases, but only to a certain limit. Tones that are played too close after one another 'blur' together and are difficult to perceive as separate events. In measurements of cymbal in jazz recordings, Friberg and Sundström (2002) observed that the inter-onset intervals between the shorter notes in the swing pattern tended to be no shorter than 100 ms, even for very fast tempi (320 BPM).

## Conclusion and discussion points

Skilled playing of percussion involves adjusting to and, where possible, utilizing the kinesthetic feedback from the instrument in question. Because the kinesthetic feedback between instruments is different, players seem to settle for a playing technique they can use for a number of playing conditions. The shape of the movement trajectories for a performer is maintained regardless of tempo or struck surface. The dynamic level, however, affects to what extent the arm is involved and also affects the timing of strokes (Dahl 2004). The movement strategies used in drumming can be described as whiplash-like and aim at achieving high stick velocities on impact. The reported values of more than 20 $ms^{-1}$ at the moment just before hit can be compared with the hammer velocity in piano playing, where a striking velocity of 3 $ms^{-1}$ corresponds to loud and 5 $ms^{-1}$ to extremely loud playing (cf. Askenfelt and Jansson, 1991).

In some respects the piano can also be said to be a percussion instrument. Piano players experience short contact times and limited access to timbre and note duration. Many of the challenges and limitations involved in percussion playing are similar to those experienced by pianists. However, in the Western musical tradition there seems to be more well-established and widespread playing techniques for the piano than for percussion. Most commonly, instruction books on drumming do not address the actual playing movements but halt at descriptions of how to hold the stick in a matched grip (both hands gripping the sticks as if just picked up from a table) or traditional grip (left hand playing with the palm up). A lot is left for the teacher to explain, or for the student to find out.

The awareness that students of drumming and percussion have of their own movements is likely to have changed with the inclusion of instruction videos. However, even if the faster movements of the drumstick were visible on a video (standards of 25–30 frames per second give a temporal resolution of 40–30 ms),

the fine-tuned components of a relaxed grip and a flexible movement remain hidden. More importantly, the 'right' feeling cannot be explained if the instructor him/herself is not aware of what strategies are used.

For some musicians too much awareness can sometimes be seen as a threat, something that demands attention and hinders a natural performance. Others believe a greater knowledge of acoustics and human performance helps them expand their understanding of the instrument and that this understanding facilitates playing. Still other musicians merely want to know that their playing technique works and will not cause them injury. We all learn through different channels, some visually (reading, observing), others by hearing (listening, being instructed). For instructors the challenge can be to provide useful knowledge through different means, whichever means best suits the student. The challenge for researchers is then to discover what 'useful knowledge' is. At present, most movement studies of drumming have focused on the playing of patterns with regular single strokes played with one hand. More complicated interactions between the player, stick, and surface (such as how a single playing gesture is used for the playing of many fast strokes in rolls) are yet to be looked into. Hopefully in the future such issues can be dealt with in collaboration between musicians, teachers, and researchers.

## Acknowledgements

The author is indebted to Carl Hakon Waadeland for providing the data for Fig. 8.1, and to Anders Askenfelt and Peta Sjölander for valuable comments on the manuscript.

## References

**Abernethy, B., Kippers, V., MacKinnon, L.T., Neal, R.J., and Hanrahan, S.** (1997). *The biophysical foundations of human movement*. Human Kinetics, Champaign, IL.

**Askenfelt, A. and Jansson, E.** (1993). From touch to string vibrations II: the motion of the key and hammer. *J Acoust Soc Am*, **90**, 2383–2393.

**Billon, M., Semjen, A., and Stelmach, G.E.** (1996). The timing effects of accent production in finger-tapping sequences. *J Motor Behav*, **28**, 198–210.

**Clarke, E.** (1999). Rhythm and timing in music. In: D. Deutsch (ed.) *The psychology of music*, pp. 473–500. Academic Press, New York.

**Dahl, S.** (2000). The playing of an accent—preliminary observations from temporal and kinematic analysis of percussionists. *J New Music Res*, **29**(3), 225–233.

**Dahl, S.** (2004). Playing the accent—comparing striking velocity and timing in an ostinato rhythm performed by four drummers. *Acta Acust*, **90**(4), 762–776.

**Drake, C. and Palmer, C.** (1993), Accent structures in music performance. *Music Percept*, **10**, 343–378.

**Famularo, D.** (1999). *It's your move. Motions and emotions*. Warner Brothers Publications.

**Friberg, A. and Sundström, A.** (2002). Swing ratios and ensemble timing in jazz performance: evidence for a common rhythmic pattern. *Music Percept*, **19**(3), 333–349.

**Gabrielsson, A.** (1999). The performance of music. In: D. Deutsch (ed.) *The psychology of music*, pp. 501–602. Academic Press, New York.

**Holloway, R.A. and Bartlett, H.R.** (1973). *Guide to teaching percussion*, 2nd edn. Wm C Brown Company Publishers, Dubuque, IA.

**Madison, G.** (2000). On the nature of variability in isochronous serial interval production. In: P. Desain and L. Windsor (eds) *Rhythm perception and production*, pp. 95–113. Swets & Zeitlinger, Lisse.

**Moeller, S.A.** (1956). *The Moeller book*. Ludwig Music Publishing Co., Grafton, OH.

**Shivas, A.A.** (1957). *The art of the timpanist and drummer*. Dobson Books, London.

**Trappe, W., Katzenberger, U., and Altenmüller, E.** Expertise related difference in cyclic motion patterns in drummers: a kinematic analysis. (Submitted.)

**Trappe, W., Parlitz, D., Katzenberger, U., and Altenmüller, E.** (1998). 3-d measurement of cyclic motion patterns in drummers with different skill. In: *Proceedings of the Fifth International Symposium on the 3-d Analysis of Human Movement*, pp. 97–99. University of Tennessee, Chattanooga, TN.

**Waadeland, C.H.** (2003). Analysis of jazz drummers' movements in performance of swing grooves—a preliminary report. In: R. Bresin (ed.) *Proceedings of Stockholm Music Acoustic Conference 2003*, Vol. II, pp. 573–576. Stockholm.

**Wickstrom, R.L.** (1983). *Fundamental motor patterns*, 3rd edn. Lea & Feibiger, Philadelphia, PA.

Part 4

# Representation in the brain

Chapter 9

# Brain structures of musicians: executive functions and morphological implications

Gottfried Schlaug

For many years researchers have questioned whether the structure of the brain changes in response to long-term stimulation and/or long-term practice of simple and complex skills. While experiments on long-term motor training in animals have provided substantial support for structural plasticity, similar data in humans is still missing. A comparison between musicians and non-musicians is an ideal model for examining whether structural plasticity occurs in the human brain. Musicians learn complex motor and auditory skills as well multimodal skills (e.g. the translation of visually perceived musical symbols into motor commands with simultaneous auditory monitoring of output) in childhood, and continuously practice them into adulthood and throughout their entire professional careers. Several studies have found evidence for structural differences in regions of the brain that may be directly involved in the perception, multimodal integration, motor control, and execution of musical skills. Studies have also found evidence for a correlation between structural/functional brain differences and the intensity/duration of practice as well as the type of instrument played. This chapter summarizes the data on structural differences found between the brains of musicians and non-musicians, with an emphasis on regional differences beyond the primary sensorimotor region and the functional role that these regions might play, particularly regions in the middle and inferior frontal lobes.

## Brain differences between adult musicians and non-musicians

Instrumental training is an intense, multi-sensory and motor experience typically initiated at an early age. Playing an instrument requires reading musical notation and translating it into sequential, bimanual motor activity, developing fine motor skills with multi-sensory feedback, memorizing long musical passages, and improvizing music. Studies have explored the basis in the brain of these highly specialized sensorimotor (Elbert *et al.* 1995; Amunts *et al.* 1997;

Hund-Georgiadis and von Cramon 1999; Schlaug 2001; Gaser and Schlaug 2003), auditory (Besson *et al.* 1994; Schlaug *et al.* 1995; Pantev *et al.* 1998; Zatorre *et al.* 1998; Keenan *et al.* 2001; Schneider *et al.* 2002; Gaab and Schlaug 2003), and auditory–spatial (Münte *et al.* 2001) skills. As shown in Plate 2, in a voxel-based morphometric analysis of non-musicians compared with amateur and professional musicians, professional keyboard players have significantly more grey matter in several regions of the brain than amateur musicians or non-musicians (Gaser and Schlaug 2003): these regions include the primary sensorimotor cortex as well as the adjacent superior pre-motor and anterior superior parietal cortex bilaterally, the mesial Heschl's gyrus (primary auditory cortex), the cerebellum, and the left inferior frontal gyrus, and part of the lateral inferior temporal lobe. The differentiation between professional and amateur musicians was based on whether or not a keyboard player's main profession was being a musician (e.g. music teacher, performer). There was also a clear separation between both groups in terms of average practice intensity across their professional life: professional musicians had done about twice the amount of practice as the amateur musicians (Gaser and Schlaug 2003). There are further anatomical differences in musicians' brains. Musicians with absolute pitch have shown greater left-sided asymmetry of the planum temporale (Schlaug *et al.* 1995; Zatorre *et al.* 1998; Keenan *et al.* 2001).

While it may not be as surprising that structural differences are found in those brain regions that subserve skills directly linked to musical training (such as the primary sensorimotor cortex that supports the sequential, independent fine motor skill in both hands or the primary auditory cortex that supports auditory discrimination skills), structural differences outside these primary regions, such as those found in the inferior frontal gyrus, are of particular interest since they may indicate that plasticity can occur in regions that either have control over primary musical functions or serve as multimodal integration regions for musical skills. Functional correlates of differences in music processing between musicians and non-musicians typically take the form of greater functional lateralization and stronger activation of secondary and tertiary auditory brain regions in musicians, whereas non-musicians might have stronger activation of primary auditory regions (Gaab *et al.* 2003). These effects have also been shown in short-term training studies both in adult non-musicians and in young children, using auditory evoked potentials (Besson *et al.* 1994; Trainor *et al.* 1999; Shahin *et al.* 2003, 2004; Bosnyak *et al.* 2004). Furthermore, studies have shown that there is even further functional specialization within the musician group itself. This has been shown in functional experiments by Pantev *et al.* (2001) who found more pronounced cortical responses to trumpet and string tones in the respective players of these instruments, demonstrating that functional brain differences are associated with

the particular musical instrument played. Similarly, when comparing string and keyboard players, Bangert *et al.* (2005) found gross anatomical differences in the pre-central gyrus. The majority of the adult keyboard players had an elaborate configuration of the pre-central gyrus on the left hemisphere, while most of the adult string players had this 'atypicality' on the right (Plate 3). Ongoing work is aimed at determining the frequency this phenomenon also occurs in those not playing an instrument and in children of varying age groups.

There is evidence suggesting that the atypicalities in musicians' brains are not innate but are instead the result of training, since these atypicalities are greater in musicians who began study at a younger age (Elbert *et al.* 1995; Schlaug *et al.* 1995) and who undertook more intense practice (Schneider *et al.* 2002; Gaser and Schlaug 2003; Hutchinson *et al.* 2003). Animal studies also support the view of training-associated brain plasticity (e.g. Greenough and Black 1992; Dawson *et al.* 2000; Anderson *et al.* 2002). Microstructural changes (increases in the number of synapses, glial cells, and capillary density within the cerebellum and primary motor cortex, as well as new brain cells in the hipppocampus) occur after long-term motor training in rats (Black *et al.* 1990; Isaacs *et al.* 1992; Anderson *et al.* 1994; Kleim *et al.* 1996; Kempermann *et al.* 1997; van Praag *et al.* 1999). The sum of these changes could amount to structural differences detectable on a macrostructural level (see Anderson *et al.* 2002).

While these studies suggest that the atypicalities in musicians' brains are due to intense training rather than pre-existing differences, they do not definitively rule out the hypothesis that these differences are innate. The critical test has not yet been carried out. To determine whether the structural and functional differences seen in musicians reflect adaptations due to musical training during sensitive periods of brain development or are instead markers of musical interest and/or aptitude that existed prior to training it is necessary to examine children and/or adults before the onset of instrumental music training and compare them with a group of control subjects not planning to play and practice regularly on a musical instrument. We have some preliminary data to support the practice hypothesis both in a single adult subject as well as in a group of children (see below). Plate 4 shows the results of a voxel-by-voxel statistical analysis to determine a change in volume of both grey and white matter in a single subject who practised a stringed instrument intensely for several months but had not played his instrument for 4 years prior to that. This subject was scanned repeatedly before, during, and after an intense period of practice (1–2 hours per day) for approximately 3 months. The image displays all areas that showed a change in volume over time. The cross-hair is centred on the left motor cortex. Multiple areas within primary motor, pre-motor, parietal, temporal, and occipital regions showed a change in volume over time. There is some overlap between the volume

changes in this single subject and the structural differences between a group of professional keyboard players and controls (Plate 2).

## A cross-sectional comparison of 5–7-year-old children prior to instrumental music training

For the past 2 years we have been conducting a pilot longitudinal study of the effects of music training on brain development and cognition in young children (Schlaug *et al.* 2003; Overy *et al.* 2004; Norton *et al.* 2005; Koelsch *et al.* 2005) The major questions addressed were: (1) whether there are pre-existing differences in brain structure/function and/or cognitive skills in children who choose to study a musical instrument compared with those who do not and (2) whether instrumental training initiated at the age of 5–6 leads to cognitive enhancement and stimulates regional brain growth in areas previously shown to be enhanced in adult musicians (Schlaug *et al.* 1995; Schneider *et al.* 2002; Gaser and Schlaug, 2003; Hutchinson *et al.* 2003; Lee *et al.* 2003). We tested 50 5–7-year-old children at baseline prior to music lessons. Approximately two-thirds of the children chose to take piano and one-third string lessons. We also tested a smaller, untreated control group (currently $n = 25$) matched to the instrumental group in age, socioeconomic status (SES), and verbal IQ. Each child underwent a battery of tests including the Object Assembly, Block Design, and Vocabulary subtests from either the Wechsler Intelligence Scale for Children (WISC-III) (for children aged 6 years and older) or the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III) (for children under the age of 6), the Raven's Colored Progressive Matrices (CPM) and Raven's Standard Progressive Matrices (SPM), the Auditory Analysis Test (Rosner and Simon 1971) as a measure of phonemic awareness, Gordon's Primary Measures of Music Audiation (PMMA) as a measure of musical skill/aptitude, and two motor tests (a finger-tapping test and a finger-sequencing task described in detail by Walker *et al.* (2002)) as a measure of speed and dexterity in both right and left hands.

Children also underwent a structural and functional magnetic resonance (MR) scan of their brain using a child-appropriate magnetic resonance imaging (MRI) protocol. MR images were acquired on a 3 T General Electric MRI scanner. The structural MR sequence had a spatial resolution of $1 \times 1 \times 1.5$ mm. We used a fully automatic technique for computational analysis of differences in local grey and white matter (Ashburner and Friston 2000; Gaser and Schlaug 2003).

We found no pre-existing cognitive, music, motor, or structural brain differences between children that entered the instrumental group and children that entered the control group at baseline (Norton *et al.* 2005), thereby making it

**Table 9.1** Morphometric Brain Measures of 5–7 year old Children

| | Total brain volume ($cm^3$) | Grey matter volume ($cm^3$) | White matter volume ($cm^3$) | Corpus callosum size volume ($mm^2$) |
|---|---|---|---|---|
| Instrumental group | 1033 (94) | 683 (67) | 348 (31) | 528 (79) |
| Non-instrumental group | 1032 (116) | 681 (88) | 344 (32) | 512 (75) |

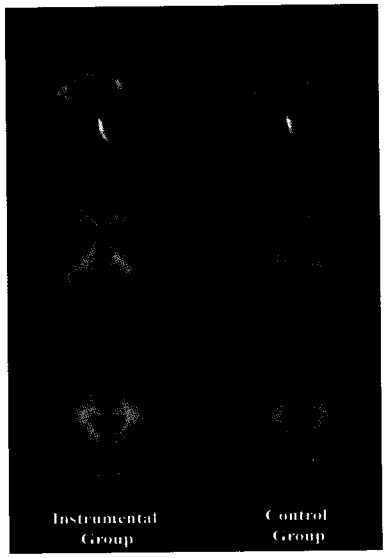



**Fig. 9.1** Three views of averaged T1-weighted brains of the instrumental and non-instrumental group.

unlikely that children who choose to learn an instrument do so because they have an atypical brain, and suggesting that the atypicalities seen in the brains of adult musicians are most likely the product of intensive music training rather than pre-existing biological markers of musicality. Table 9.1 shows the mean brain measures. There were no differences in the absolute brain volume, volume of grey matter, volume of white matter, or the size of the mid-saggital corpus callosum. Figure 9.1 shows the averaged brains after spatial standardization for the group of children planning to learn an instrument and the group of children who do not plan to study an instrument. A voxel-based analysis did not show any significant differences in volume of grey matter between the two groups.

In our preliminary analyses of the effects of 1 year of music training (only half of the children have so far completed their second round of testing), we found significantly greater change scores in the instrumental group than in the control group in skills directly linked to music training: fine motor skills (mean of 10% for the instrumental group compared with 5% for the non-instrumental group) and auditory discrimination skills as measured by Gordon's PMMA (1986)

**Table 9.1** Morphometric Brain Measures of 5–7 year old Children

| | Total brain volume ($cm^3$) | Grey matter volume ($cm^3$) | White matter volume ($cm^3$) | Corpus callosum size volume ($mm^2$) |
|---|---|---|---|---|
| Instrumental group | 1033 (94) | 683 (67) | 348 (31) | 528 (79) |
| Non-instrumental group | 1032 (116) | 681 (88) | 344 (32) | 512 (75) |




**Fig. 9.1** Three views of averaged T1-weighted brains of the instrumental and non-instrumental group.

unlikely that children who choose to learn an instrument do so because they have an atypical brain, and suggesting that the atypicalities seen in the brains of adult musicians are most likely the product of intensive music training rather than pre-existing biological markers of musicality. Table 9.1 shows the mean brain measures. There were no differences in the absolute brain volume, volume of grey matter, volume of white matter, or the size of the mid-saggital corpus callosum. Figure 9.1 shows the averaged brains after spatial standardization for the group of children planning to learn an instrument and the group of children who do not plan to study an instrument. A voxel-based analysis did not show any significant differences in volume of grey matter between the two groups.

In our preliminary analyses of the effects of 1 year of music training (only half of the children have so far completed their second round of testing), we found significantly greater change scores in the instrumental group than in the control group in skills directly linked to music training: fine motor skills (mean of 10% for the instrumental group compared with 5% for the non-instrumental group) and auditory discrimination skills as measured by Gordon's PMMA (1986)

(9% versus 6%). We have not yet demonstrated transfer effects, although in several transfer domains (e.g. verbal and mathematics) the instrumental group improved more than did the control group (not siginifcant). Brain data also support this trend. There was a non-significant greater increase in volume of grey matter in the instrumental group than in the control group, but no significant change in size of the corpus callosum has yet emerged.

## A cross-sectional comparison of 9–11-year-old children (instrumentalists versus non-instrumentalists)

We recently conducted a cross-sectional comparison between a group of 9–11-year-old instrumentalists with an average of 3–4 years of training and a group of non-instrumental children matched in age, handedness, and SES. Preliminary results from this cross-sectional study showed that the instrumentalists had a significantly larger volume of grey matter which was regionally pronounced in the sensorimotor cortex, but also in the occipital lobe bilaterally (Schlaug *et al.* 2004). Plate 5 shows the regional distribution of differences in grey matter volume when the instrumental group was compared with the non-instrumental group on a voxel-by-voxel basis. Table 9.2 shows the average (standard deviation, SD) volume data for total brain, grey matter, and white matter volume.

The instrumental group also outperformed the well-matched control group on the Vocabulary subtest of the WISC-III, a phonemic awareness test (Auditory Analysis), and the Key Math test.

We also investigated whether correlations existed between the performance in music discrimination skills and interindividual differences in grey matter distribution. The composite score of Gordon's PMMA was correlated with the individual grey matter images of all 5–7-year-old children. We found significant associations between the performance score and regional grey matter concentrations in the temporal lobe and the inferior parietal lobe as well as the cerebellum. The same correlation analysis in the 9–11-year-old group including children with significant musical experience revealed a significant change in the maps. Now, there were significant correlations that included the inferior frontal lobe on both sides in addition to the middle to anterior part of the temporal lobe (see Plate 6).

**Table 9.2** Morphometric Brain Measures of 9–11 year old Children

| | **Total brain volume ($cm^3$)** | **Grey matter volume ($cm^3$)** | **White matter volume ($cm^3$)** |
|---|---|---|---|
| Instrumental group | 1105 (99) | 747 (75) | 357 (39) |
| Non-instrumental group | 1015 (133) | 661 (82) | 354 (36) |

## The functional significance of frontal 'auditory' and 'auditory–motor' regions

Plate 7 shows a comparison between two groups of adults (musicians and non-musicians) and a group of 5–7-year-old children, all performing the same rhythm discrimination task while activity is being measured with functional MR imaging, normal maturational changes in the processing of this information can be seen in the comparison between the adult non-musicians and the children, but additional modulations of these maturational changes can be seen in several regions when the group of adult musicians is compared with both the children's group and the adult non-musician group. Those additional regions include the parietal lobe (mainly around the intra-parietal sulcus), the posterior middle frontal gyrus (Brodmann area 6), and the inferior frontal gyrus (Brodmann areas 44, 45, and 47). In this chapter we will concentrate on the middle and inferior frontal gyrus, since those regions appear to show the greatest difference between adult musicians and non-musicians. The activation of the parietal lobe can be seen in both groups of adults. This activation may indicate that with increased experience (either implicit or explicit), regions in the parietal lobe may become active due either to poly-sensory integration of information or to increased global processing strategies in adults as compared with children who may apply more local processing strategies.

Several studies have found activations or activity changes in a dorsal posterior pre-motor region with purely perceptual discrimination tasks, combined perceptual-memory tasks, and auditory–motor association tasks (Zatorre *et al.* 1998; Ohnishi *et al.* 2001; Gaab *et al.* 2003; Bangert and Altenmüller 2003; Haslinger *et al.* 2004; Hasegawa *et al.* 2004; Meister *et al.* 2004). Across studies, it appears that musicians showed a more pronounced activation of the middle and inferior frontal gyrus regions. It is of particular interest that two independent voxel-based morphometric studies found a larger volume of grey matter in the inferior frontal gyrus in musicians compared with non-musicians (Sluming *et al.* 2003; Gaser and Schlaug 2003; see also Fig. 9.1), although the functional significance of this frontal 'auditory' region is still unclear. It has been speculated that the inferior frontal gyrus might play a role in performing tasks that have to do with the sequencing of behaviourally relevant auditory information (Kölsch *et al.* 2002; Gaab *et al.* 2003; Kölsch *et al.* 2005). Furthermore, they may become active when the integration of information, particularly in the temporal domain, is important for solving a task (Levitin et al., 2004). Research has also shown that these frontal regions are part of the mirror neuron system. The mirror neuron system might be important for associating actions with sounds.

The ability to develop associations between actions and sounds is essential for the production and comprehension of music, or any kind of behaviourally

relevant sound information. Several research studies have implicated parts of the frontal cortex as important regions for integrating information from multiple sensory modalities and brain regions (Bremmer *et al.* 2001; Kohler *et al.* 2002; Keysers *et al.* 2003). It is therefore likely that the frontal cortex plays an important role in developing associations across modalities, including action–sound mappings (Bangert and Altenmüller 2003). Research with monkeys has revealed a neural mechanism that may be responsible for representing associations between actions and sensory stimulation. Rizzolatti and colleagues identified a class of neurons in area F5 of monkey cortex, the presumed homologue of Brodmann area (BA) 44 and BA 6, that is tuned to the actions of others as well as actions of the self (Rizzolatti *et al.* 2002; Buccino *et al.* 2004). These so-called 'mirror' neurons respond both when an action is observed and when that same action is performed. In addition to the sight and performance mirror neurons, a subset of mirror neurons in monkeys also responds to the sound of an action (Kohler *et al.* 2002; Keysers *et al.* 2003). These 'auditory–visual' mirror neurons exemplify high-level abstraction in the representation of action—an identical neural system becomes activated whether a particular action is heard, seen, or performed.

The action representation maintained by mirror neurons has been shown to evolve through learning. The visual stimulus that elicits activity in a particular neuron can change with exposure to new stimuli involving the identical action goal (Rizzolatti and Arbib 1998). It may be inferred that associations between sound and action are similarly malleable. Schubotz and colleagues hypothesized that Hebbian learning drives a somatotopic organization of pre-motor brain activity for processing meaningful temporal sequences (Schubotz and von Cramon 2002; Schubotz *et al.* 2003). They found that for an identical task involving the anticipation of temporal order, stimuli associated with vocal production (sine tones) generated activation of the inferior ventral pre-motor cortex whereas stimuli associated with hand manipulation (circles of varying diameter) generated activation of the superior ventral pre-motor cortex (Schubotz and von Cramon 2002).

This is of immediate relevance for plasticity research using musicians. Music practice and development exemplifies neural plasticity for evolving representations of actions and sounds. Professional singers learn and continuously practice to associate vocal articulation/pharyngeal movements with meaningful patterns in sound. Similarly, professional keyboard players learn and continuously practice to associate hand/finger movements with meaningful patterns in sound. There is some evidence for the second case, since Bangert and Altenmüller (2003) have shown that learning associations between actions and piano tones will lead to changes in activity in the dorsal frontal/pre-motor

cortex. More research is needed to determine whether parts of the frontal/premotor cortex are a general-purpose locus for processing sequences of sound events that are meaningfully related. This hypothesis would certainly explain a majority of frontal activations that have been seen with pure auditory perceptual tasks, auditory memory tasks, auditory discrimination tasks, and auditory–motor association tasks (Zatorre *et al.* 1994; Kölsch *et al.* 2002; Levitin and Menon 2003; Gaab *et al.* 2003; Bangert and Altenmüller 2003). Thus, considering the putative functions of the inferior frontal region, it might not be surprising to see that some of the most pronounced differences between musicians and non-musicians, both structurally and functionally, are in the inferior frontal region.

## Acknowledgements

The research mentioned in this chapter was supported by grants from the National Science Foundation (NSF), the IFMR, and the Grammy Foundation. Christian Gaser, Andrea Norton, Katie Overy, Marc Bangert, Karl Cronin, Ellen Winner, D. J. Lee, and Stefan Koelsch have contributed to the research presented in this chapter.

## References

**Amunts, K., Schlaug, G., Jäncke, L.** *et al.* (1997). Motor cortex and hand motor skills: structural compliance in the human brain. *Hum Brain Mapp*, **5**, 206–215.

**Anderson, B.J., Li, S., Alcantara, A., Isaacs, K.R., Black, J.E., and Greenough, W.T.** (1994). Glial hypertrophy is associated with synaptogenesis following motor-skill learning, but not with angiogenesis exercise. *Glia*, **11**, 73–80.

**Anderson, B.J., Eckburg, P.B., and Relucio, K.I.** (2002). Alterations in the thickness of motor cortical subregions after motor skill learning and exercise. *Learn Memory*, **9**, 1–9.

**Ashburner, J. and Friston, K.J.** (2000). Voxel-based morphometry—the methods. *NeuroImage*, **11**, 805–821.

**Bangert, M. and Altenmüller, E.O.** (2003). Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neurosci*, **4**, 26

**Bangert, M., Nair, D., and Schlaug, G.** (2005). Structural specialization of the specialized. Human Brain Mapping Meeting, Toronto.

**Besson, M., Gaita, F., and Requin, J.** (1994). Brain waves associated with musical incongruities differ for musicians and non-musicians. *Neurosci Lett*, **168**, 101–105.

**Black, J.E., Isaacs, K.R., Anderson, B.J., Alcantra, A.A., and Greenough, W.T.** (1990). Learning causes synaptogenesis whereas motor activity causes angiogenesis in cerebellar cortex of adult rats. *Proc Natl Acad Sci USA*, **87**, 5568–5572.

**Bosnyak, D.J., Eaton, R.A., and Roberts, L.E.** (2004). Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. *Cereb Cortex*, **14**, 1088–1099.

**Bremmer, F., Schlack, A., Jon Shah, N. *et al.*** (2001). Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron*, **29**, 287–296.

**Buccino, G., Binkofski, F., and Riggio, L.** (2004). The mirror neuron system and action recognition. *Brain Lang*, **89**, 370–376.

**Cronin, K., Overy, K., Norton, A., Marchina, S., Winner, E., and Schlaug, G.** (2004). The effects of experience and brain maturity on neural correlates of music processing. Poster presented at Human Brain Mapping, Budapest, 2004 (published in *NeuroImage*, **22**).

**Dawson, G., Ashman, S.B., and Carver, L.J.** (2000). The role of early experience in shaping behavioral and brain development and its implications for social policy. *Dev Psychopathol*, **12**, 695–712.

**Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., and Taub, E.** (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, **270**, 305–306.

**Gaab, N. and Schlaug, G.** (2003). The effect of musicianship on pitch memory in performance matched groups. *NeuroReport*, **14**, 2291–2295.

**Gaab, N., Gaser, C., Zaehle, T., Jäncke, L., and Schlaug, G.** (2003). Functional anatomy of pitch memory—an fMRI study with sparse temporal sampling. *NeuroImage*, **19**, 1417–1426.

**Gaser, C., and Schlaug, G.** (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience*, **23**, 9240–9245.

**Greenough, W. T. and Black, J.E.** (1992). Induction of brain structure by experience: Substrates for cognitive development. In: M.R. Gunnar and C.A. Nelson (eds) *The Minnesota Symposia on Child Psychology: Vol. 24. Developmental behavioral neuroscience*, pp. 155–200. Erlbaum, Hillsdale, NJ.

**Hasegawa, T., Matsuki, K., Ueno, T. *et al.*** (2004). Learned audio-visual cross-modal associations in observed piano playing activate the left planum temporale. An fMRI study. *Cogn Brain Res*, **20**, 510–518.

**Haslinger, B., Erhard, P., Altenmüller, E.** *et al.* (2004). Reduced recruitment of motor association areas during bimanual coordination in concert pianists. *Hum Brain Mapp*, **22**, 206–215.

**Hund-Georgiadis, M. and von Cramon, Y.** (1999). Motor-learning-related changes in piano players and non-musicians revealed by functional magnetic-resonance signals. *Exp Brain Res*, **125**, 417–425.

**Hutchinson, S., Lee, L.H.L., Gaab, N., and Schlaug, G.** (2003). Cerebellar volume of musicians. *Cereb Cortex*, **13**, 943–949.

**Isaacs, K.R., Anderson, B.J., Alcantara, A.A., Black, J.E., and Greenough, W.T.** (1992). Exercise and the brain: angiogenesis in the adult rat cerebellum after vigorous physical activity and motor skill learning. *J Cereb Blood Flow Metab*, **12**, 110–119.

**Keenan, J.P., Thangaraj, V., Halpern, A.R., and Schlaug, G.** (2001). Absolute pitch and planum temporale. *NeuroImage*, **14**, 1402–1408.

**Kempermann, G., Kuhn, H.G., and Gage, F.H.** (1997). More hippocampal neurons in adult mice living in an enriched environment. *Nature*, **386**, 493–495.

**Keysers, C., Kohler, E., Umiltà, M.A., Nanetti, L., Fogassi, L., and Gallese, V.** (2003). Audiovisual mirror neurons and action recognition. *Exp Brain Res*, **153**, 628–636.

**Kleim, J.A., Lussnig, E., Schwarz, E.R., Comery, T.A., and Greenough, W.T.** (1996). Synaptogenesis and fos expression in the motor cortex of the adult rat after complex motor skill acquisition. *J Neurosci*, **16**, 4529–4535.

Kölsch, S., Gunter, T.C., von Cramon, D.Y., Zysset, S., Lohmann, G., and Friederici, A.D. (2002). Bach speaks: a cortical 'language-network' serves the processing of music. *NeuroImage*, 17, 956–966.

Kölsch, S., Fritz, T., Schulze, K., Alsop, D., and Schlaug, G. (2005). Adults and children processing music: an fMRI study. *NeuroImage*. **25**, 1068–76.

Kohler, E., Keysers, C., Umiltà, M.A., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, **297**, 846–848.

Lee, D.J., Chen, Y., and Schlaug, G. (2003) Corpus callosum: musician and gender effects. *NeuroReport*, **14**, 205–209.

Levitin, D.J. and Menon, V. (2003). Musical structure is processed in language areas of the brain: a possible role for Brodmann Area 47 in temporal coherence. *NeuroImage*, **20**, 2142–2152.

Meister, I.G., Krings, T., Foltys, H. *et al.* (2004). Playing piano in the mind—an fMRI study on music imagery and performance in pianists. *Brain Res Cogn Brain Res*, **19**, 219–228.

Münte, T.F., Kohlmetz, C., Nager, W., and Altenmüller, E. (2001). Neuroperception: superior auditory spatial tuning in conductors. *Nature*, **409**, 580.

Norton, A., Winner, E., Cronin, K., Overy, K., Lee, D., and Schlaug, G. (2004). Are there pre-existing neural, cognitive, or motoric markers for musical ability? Brain Cogn, **59**, 124–34.

Ohnishi, T., Matsuda, H., Asada, T. *et al.* (2001). Functional anatomy of musical perception in musicians. *Cereb Cortex*, **11**, 754–760.

Overy, K., Norton, A., Cronin, K. *et al.* (2004). Imaging melody and rhythm processing in young children. *NeuroReport*, **15**, 1723–1726.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L.E., and Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, **392**, 811–814.

Patel, A.D. (2003). Language, music, syntax and the brain. *Nat Neurosci*, **6**, 674–681.

Van Praag, H., Christie, B.R., Sejnowksi, T.J., and Gage, F.H. (1999). Running enhances neurogenesis, learning, and long-term potentiation in mice. *Proc Natl Acad Sci USA*, **96**, 13427–13431.

Rizzolatti, G. and Arbib, M.A. (1998). Language within our grasp. *Trends Neurosci*, **21**(5), 188–194.

Rizzolatti, G., Fogassi, L., and Gallese, V. (2002). Motor and cognitive functions of the ventral premotor cortex. *Curr Opin, Neurobiol*, **12**, 149–154.

Rosner, J. and Simon, D. (1971). Auditory Analysis Test. *J Learn Disab*, **4**, 40–48.

Schlaug, G. (2001). The brain of musicians: a model for functional and structural adaptation. In: R. Zatorre and I. Peretz (eds) *The biological foundations of music* (*Ann NY Acad Sci*, **930**, 281–299).

Schlaug, G., Jäncke, L., Huang, Y., and Steinmetz, H. (1995*a*). *In vivo* evidence of structural brain asymmetry in musicians. *Science*, **267**, 699–701.

Schlaug, G., Jäncke, L. Huang, Y., and Steinmetz, H., (1995*b*). Increased corpus callosum size in musicians. *Neuropsychologia*, **33**, 1047–1055.

Schlaug, G., Winner, E., Norton, A., Cronin, K., and Overy, K.(2003). Effects of music training on children's brain and cognitive development. Poster presented at Society for Neurosciences, New Orleans, November, 2003.

**Schlaug, G., Lee, D.J., Overy, K., Cronin, K., Norton, A., and Winner, E.** (2004). Does brain anatomy predict musicianship? Poster presented at Human Brain Mapping, Budapest, 2004 (published in *NeuroImage*, **22S**).

**Schneider, P., Scherg, M., Dosch, H.G., Specht, H.J., Gutschalk, A., and Rupp, A.** (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat Neurosci*, **5**, 688–694.

**Schubotz, R.I. and von Cramon, D.Y.** (2002). Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: an fMRI study. *NeuroImage*, **15**, 787–796.

**Schubotz, R.I., von Cramon, D.Y., and Lohmann, G.** (2003). Auditory what, where, and when: a sensory somatotopy in lateral premotor cortex. *NeuroImage*, **20**, 173–185.

**Shahin, A., Bosnyak, D.J., Trainor, L.J., and Roberts, L.E.** (2003). Enhancemment of neuroplastic P2 and N1c auditory evoked potentials in skilled musicians. *J Neurosci*, **23**, 5545–5552.

**Shahin, A., Roberts, L.E., and Trainor, L.J.** (2004). Enhancement of auditory cortical development by musical experience in children. *NeuroReport*, **15**, 1917–1921.

**Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes, A., and Roberts, N.** (2002). Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *NeuroImage*, **17**, 1613–1622.

**Trainor, L.J., Desjardins, R.N., and Rockel, C.** (1999). A comparison of contour and interval processing in musicians and nonmusicians using event-related potentials. *Aust J Psychol*, **51**, 147–153.

**Walker, M.P., Brakefield, T., Morgan, A., Hobson, J.A., and Stickgold, R.** (2002). Practice with sleep makes perfect: sleep-dependent motor skill learning. *Neuron*, **35**, 205–211.

**Wiesel, T.N. and Hubel, D.H.** (1963). Single-cell responses in striate cortex of kittens deprived of vision in one eye. *J Neurophysiol*, **26**, 1003–1007.

**Zatorre, R.J., Evans, A.C., and Meyer, E.** (1994). Neural mechanism underlying melodic perception and memory for pitch. *J Neurosci*, **14**, 1908–1919.

**Zatorre, R.J., Perry, D.W., Beckett, C.A., Westbury, C.F., and Evans, A.C.** (1998). Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. *Proc Natl Acad Sci*, **95**, 3172–3177.

Chapter 10

# The motor representation in pianists and string players

Lutz Jäncke

This chapter summarizes current findings about the neuroanatomical and neurophysiological peculiarities of musicians' brains. It is argued that musicians are a paradigm case for studying the functional plasticity of the central nervous system on three levels (function, neurophysiology, and neuroanatomy). Because most musicians start their musical training very early in life and continue to practise throughout their entire life the brain structures involved in the control of musical functions are constantly stimulated and thus shaped.

## Introduction

Musicians form an ideal and interesting subject pool for the study of structural and functional adaptations associated with exceptional cognitive and motor performance (several authors have argued this, for example Schlaug (2001) and Münte *et al.*(2002)). In order to understand this, one has to imagine the typical career of a professional musician. One outstanding feature of musically talented children who progress to become outstanding performers is their conscientious approach to practising their instrument at ages which could obviously influence brain development. This was highlighted by Ericsson *et al.* (1993*a*,*b*), who found that accumulated practice time for the best violin students at the Berlin Academy was 7400 h by the time they were 18 years old, and for expert pianists it was 7600 h. This is in contrast to 3400 h for violinists who were student music teachers and to 1600 h for amateur pianists. Ericsson *et al.* also reported that the highest achievers practised regularly at the same time of day and for a consistent amount of time. These findings are supported by those of Sloboda and Davidson (1996), who reported a positive correlation between the amount of practice and grade achievement in the practical examinations of the Associated Board of the Royal Schools of Music. Even with novice musicians, the amount of practice can predict musical outcome, as reported in a study of 7–9-year-old children at the beginning of instrumental training (O'Neill 1997). This author found that performance outcome depended on the amount of practice and general level of

motivation. It was also shown that intelligence and musical aptitude had no influence on the outcome. It is thought that this extraordinary intensity of musical training might shape the brain and its functions. Peculiarities in the brains of musicians have been reported on the grounds of the morphometric analysis of post-mortem brains, lesion data, and neuropsychological tests. In the following, we shall discuss the peculiarities of musicians' motor systems. First we will refer to differences in motor functions and then discuss the anatomical peculiarities of the motor system, and finally the neurophysiological differences. This presentation is mainly based on the peculiarities of the motor system of pianists and string players because the motor functions of brass players or singers have not been examined in detail so far.

## Specific motor functions in pianists and string players

Performing music at a professional level is arguably among the most complex of human accomplishments. A pianist, for example, has to bimanually coordinate the production of up to 1800 notes per minute. Similar motor control demands are placed on violinists who additionally have to cope with unusual biomechanical constraints to hold the violin. A further point demonstrating the superiority of the motor system of professional musicians is their extraordinary ability to use both hands (or arms) in a coordinated manner. For example, during piano playing one hand plays the melody (mostly the dominant right hand manipulating the high pitches) while the other hand (the subdominant hand) plays the rhythm. For the violin the case is different. While the left hand is manipulating the strings in a way partly similar to playing the piano, the right arm is swinging the bow. Here the right arm has to generate a fast spatial–temporal pattern coordinated with fast finger movements of the left hand. Thus, both instruments place different demands on bimanual coordination (see Chapter 7). Many researchers have tried to delineate specific motor functions in musicians on the basis of sophisticated behavioural experiments. One of the simplest motor tasks is tapping with the index finger and measuring the *maximum tapping frequency* with one target finger.

To obtain the maximum tapping frequency, subjects are required to tap as fast as possible with one finger (e.g. the index finger) within a pre-determined period (e.g. 20 s). The intertap interval as well as the maximum number of taps is counted using simple computer software and keys registering the taps. The maximum tapping speed has been shown to indicate basic neurophysiological properties of the primary motor cortex controlling the particular finger. It has been shown that the maximum tapping speed with the index finger is significantly faster for the dominant hand. That is, right-handers reveal faster tapping speeds

for the right than for the left hand while left-handers show the opposite pattern. This asymmetry in hand skills is relatively stable during the course of short-term hand-skill training (Peters 1976; Peters and Durding 1978). It is argued that maximum tapping speed is mainly controlled by the hand motor area in the cortex. Several positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) experiments have shown that the rate of finger tapping correlates with the intensity or spatial extent of activation within the sensorimotor cortex (including M1 and S1). In addition to that, single-cell recording studies in monkeys have shown that the discharge activity of M1 motoneurons maximally correlates with the force and velocity of the movement (Hepp-Reymond 1988). Since fast finger tapping requires high movement velocities, one might conclude that fast finger tapping is mainly controlled by M1 motoneurons.

Pianists and string players differ substantially from non-musicians with respect to the fact that they tap much faster (Fig. 10.1). The tapping rate of the dominant hand is increased by about 14% and the rate of the non-dominant hand by about 20% compared with normal controls (right hand; musicians 128 taps/20 s, 6.4 Hz; non-musicians 110 taps/20 s, 5.5 Hz; left hand, musicians 120 taps/s., 6 Hz; non-musicians 95 taps/s, 4.75 Hz). The non-dominant left hand demonstrates a slightly larger performance gain compared with the dominant hand, a fact resulting in a decreased asymmetry between both hands in musicians. Typical hand asymmetry scores are calculated by dividing the performance difference of both hands by the performance sum of both hands, $(R - L)/(R + L)$. The obtained asymmetry score is significantly lower in musicians than in non-musicians (Jäncke *et al.* 1997). Interestingly, there was also a

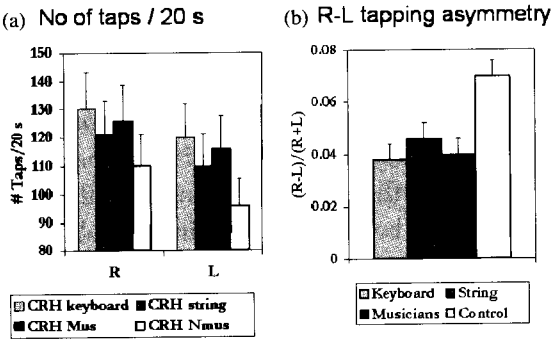



**Fig. 10.1** (a) Tapping performance (number of taps/20 s) for musicians and non-musicians and also for string versus keyboard players. (b) Tapping asymmetry scores, $(R - L)/(R + L)$, for musicians and non-musicians and also for string versus keyboard players. Vertical bars indicate standard errors of the mean. Data are taken from Jäncke *et al.* (1997).

for the right than for the left hand while left-handers show the opposite pattern. This asymmetry in hand skills is relatively stable during the course of short-term hand-skill training (Peters 1976; Peters and Durding 1978). It is argued that maximum tapping speed is mainly controlled by the hand motor area in the cortex. Several positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) experiments have shown that the rate of finger tapping correlates with the intensity or spatial extent of activation within the sensorimotor cortex (including M1 and S1). In addition to that, single-cell recording studies in monkeys have shown that the discharge activity of M1 motoneurons maximally correlates with the force and velocity of the movement (Hepp-Reymond 1988). Since fast finger tapping requires high movement velocities, one might conclude that fast finger tapping is mainly controlled by M1 motoneurons.

Pianists and string players differ substantially from non-musicians with respect to the fact that they tap much faster (Fig. 10.1). The tapping rate of the dominant hand is increased by about 14% and the rate of the non-dominant hand by about 20% compared with normal controls (right hand; musicians 128 taps/20 s, 6.4 Hz; non-musicians 110 taps/20 s, 5.5 Hz; left hand, musicians 120 taps/s., 6 Hz; non-musicians 95 taps/s, 4.75 Hz). The non-dominant left hand demonstrates a slightly larger performance gain compared with the dominant hand, a fact resulting in a decreased asymmetry between both hands in musicians. Typical hand asymmetry scores are calculated by dividing the performance difference of both hands by the performance sum of both hands, $(R - L)/(R + L)$. The obtained asymmetry score is significantly lower in musicians than in non-musicians (Jäncke *et al.* 1997). Interestingly, there was also a

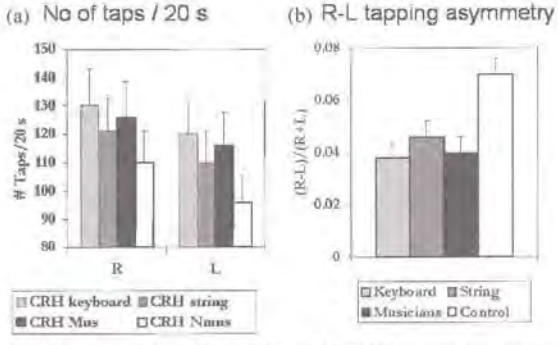



**Fig. 10.1** (a) Tapping performance (number of taps/20 s) for musicians and non-musicians and also for string versus keyboard players. (b) Tapping asymmetry scores, $(R - L)/(R + L)$, for musicians and non-musicians and also for string versus keyboard players. Vertical bars indicate standard errors of the mean. Data are taken from Jäncke *et al.* (1997).

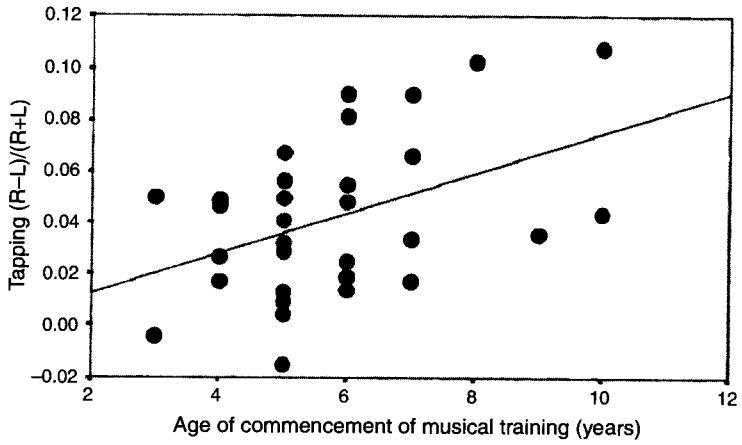



**Fig. 10.2** Tapping asymmetry scores (R − L)/(R + L) obtained for musicians related to age of commencement of musical training.

difference between pianists and string players, with pianists outperforming string players with a tapping speed approximately 8% faster, although string players are far better than normal controls. Thus, we conclude this performance gain should be due to the extraordinary training musicians (and pianists) get from practising their particular instruments. This argument of use-dependent plasticity is supported by the fact that the tapping asymmetry in musicians is related to the age at which they begin their musical training. Musicians starting their musical training very early in life revealed the smallest tapping asymmetry compared with musicians starting later in life (see Fig. 10.2). Thus, there is strong evidence that this simple motor measure indicating basic motor control mechanisms depends on practice effects and on the age when practising started. Obviously, the motor system (here M1) is tuned to effectively control finger movements.

Besides the fact that pianists and string players tap with a faster speed, several studies have shown that pianists generate the taps with decreased variability; thus, their tapping is more stable. This stability has been found not only for regular beats but also for rhythms generated unimanually or bimanually. From mathematical models, many believe that this timing stability is due to superior peripheral motor implementations (inserting a motor program from central processing centres to the executing organs) and timekeeper executions (e.g. using an internal clock more efficiently) (Krampe 2002; Krampe *et al.* 2002). A further superior aspect of movement control is that with increasing familiarity with a specific piece of music musicians tap more slowly when they are asked to tap along with the rhythm of the music. This slowing of sensorimotor response is

thought to be associated with a different way to integrate motor and sensory functions. In fact, it is thought that with increasing familiarity with the musical piece, musicians tend to synchronize their tapping with the perceived rhythm, relying on higher hierarchical levels (Drake and Ben El 2003). Recent work by Drost *et al.* (2004) uncovered an additional very important peculiarity of the motor system of musicians. These authors found evidence from behavioural experiments that musicians implicitly anticipated a specific motor program when they were required to play a musical chord in response to an imperative visual stimulus congruent with the chord. However, when this association between visual stimulus and motor response was disturbed by a concurrent task (e.g. playing an auditory stimulus which is congruent with the chord) the performance in playing the chords (reaction time) dropped only in the musicians. Thus, the musicians seem to anticipate implicitly the appropriate motor program (playing a chord) when seeing the imperative stimulus and when hearing the concurrent auditory stimulus. Obviously, both stimuli have their own associations with motor programs which might interfere. Taken together, these (and other) behavioural experiments clearly demonstrate that musicians (in particular pianists) have developed a specialized and highly tuned sensorimotor system enabling them to perform at a high level of skill.

An interesting and impressive example of the adaptive nature of the sensorimotor system of pianists has been described by the Harvard psychologists Bruno Laeng and Anne Park (1999). They examined how right- and left-handed novice pianists could adapt to play a reversed piano with the high pitches on the left and the low on the right. Interestingly, left-handers had fewer problems learning on a reversed piano while right-handed learners faced some problems. However, when experienced pianists were tested playing a reversed piano, they were basically unable to play on reversed keyboards even if they were strong and consistent left-handers. Thus, this study demonstrates that pianists acquire strong associations between their motor responses and visual cues (arrangement of keys on the keyboard) which are difficult to rearrange or break when they have been established. Maybe long-lasting and intensive training is required to build up new visual–motor associations.

However, we had the opportunity to study a well-known pianist who is able to play on reversed and regular keyboards (the theoretical arguments about the extraordinary underlying neural mechanisms are given in the theoretical paper by Jäncke (2002)). Interestingly, this pianist started his career by playing on regular keyboards. Simply by accident he noticed that he could also play on a reversed piano. Meanwhile he performs exclusively on reversed pianos, motivated by the subjective impression that he is playing with more emotional emphasis and precision using the reversed keyboard. We believe that he is a rare exception of

a professional pianist exhibiting the capacity to play on both keyboard orientations. Hopefully, we shall uncover the extraordinary organization of his neural network controlling music when he performs in forthcoming experiments.

## Anatomical peculiarities of musicians' brains

Since the age of phrenology, neuroscientists have tried to relate extraordinary skills to changes in anatomy of the brain. For example, at the beginning of the twentieth century Auerbach reported that the middle and posterior thirds of the superior temporal gyrus were larger than normal in several post-mortem studies of the brains of famous musicians (Meyer 1977). Modern brain-imaging techniques, such as high-resolution MRI, allow us to study anatomical details in the brains of living humans. Studies in which these techniques have been used have shown in particular that the motor system of musicians is differently structured. The first systematic study demonstrating an anatomical difference in the motor system was published by Schlaug *et al.* (1995). Using MRI they measured the size of the mid-saggital area of the corpus callosum of musicians and non-musicians. Previous zytoarchitectonical studies have shown that this measure correlates with the number and diameter of axons crossing through the mid-line. Thus, the larger the area of the corpus callosum the more efficient is the interhemispheric communication due to more or larger axons crossing through the mid-line. In fact, Schlaug *et al.* found a larger anterior corpus callosum in musicians who started their musical training early in life (before the age of 7). There was no difference between musicians starting later in life and normal control subjects; in addition, there was also no difference between the size of the posterior corpus callosum in either group. The fact that the anterior corpus callosum was larger in musicians starting musical training early in life was taken as evidence for superior interhemispheric communication between the anteriorly located motor areas including the pre-motor and primary motor areas. Although this study was the first to demonstrate a particular anatomical difference in the motor system of musicians, it was unclear whether these findings indeed indicated more efficient interhemispheric communication with respect to excitatory or inhibitory information. On the basis of current research using neurophysiological methods to study functional differences between musicians and non-musicians and current models of interhemispheric communication (Cook 1984), it is assumed that inhibitory influences are more important in coordinating the motor areas of hemispheres. In view of these suggestions, both hemispheres can inhibit themselves more efficiently because of the superior interhemispheric connection. Although this study has elegantly reopened the debate about whether these structural differences are related directly to musical ability, it was only an indirect

indicator of anatomically different organization of motor systems in musicians. At the same time, Elbert *et al.* (1995) published a remarkable paper in which they showed that the cortical somatotopic representation of the left hand was changed in professional violinists. Violinists showed larger cortical representations of the subdominant hand than non-musicans.

Using MRI, Amunts *et al.* (1996) measured the intrasulcal length of the pre-central gyrus, i.e. the depth of the central sulcus (CS), as an estimate of the size of the motor area for the hand. This measure was obtained from horizontal slices of linearly normalized brains. The intrasulcal length of the posterior contour of the pre-central gyrus was measured as an indicator of the size of the primary motor cortex., i.e. Brodmann's area 4 in the hand representation area in a dorso-ventral sequence of 35 horizontal sections. The sections extended from Talairach coordinates $z = 69$ to $z = 35$, thus including the functional hand representation. In summary, these MRI studies revealed the following results:

1. Right-handed males showed a clear left-ward asymmetry of the depth of the CS for the superior horizontal slices from $z = 70$ to about 50 while left-handed males showed a right-ward asymmetry.
2. Professional right-handed keyboard players who had received intensive bimanual training from early childhood were more symmetrical both in hand skill tests and in the sulcal depths.
3. The absolute length of the CS was related to the age of commencement of musical training in as much as those musicians who had started early in life with musical training revealed the largest CS on both sides.

These findings suggest that the anatomical asymmetry in the depth of the CS may depend on handedness and also on the amount or duration of hand skill training (see Fig. 10.3). These findings fit well with data from a magnetoencephalograph (MEG) study analysing dipoles during simple hand and finger movements (Volkmann *et al.* 1998). The authors found larger dipole source intensities in the primary motor cortex of the preferred hand than of the non-preferred hand both in right and left handers. Further studies provide evidence that handedness may have anatomical correlates not only in the asymmetry of the motor cortex for the primary hand but also in the spinal cord (Melsbach *et al.* 1996). There is also evidence for an asymmetric volume of the anterior and posterior cerebellum correlating with handedness (Snyder *et al.* 1995).

Two more recent studies using voxel-based morphometry (VBM) or classical region of interest analysis to detect structural differences between pianists and non-musicians found larger volumes of grey matter in the musician group in the motor network including primary motor (M1), the somatosensory areas (S1), the pre-motor areas (PMC), the superior parietal lobe (SPL), and the left

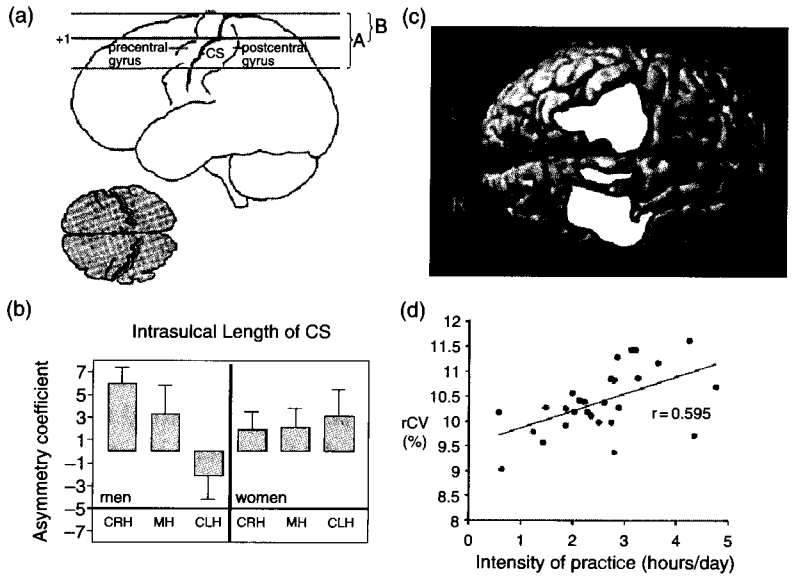



**Fig. 10.3** Schematic description of macrostructural differences between musicians and non-musicians with respect to motor areas. (a) Schematic drawing of the slice position at which the intrasulcal length of the central sulcus (CS) was measured. (b) Bar graphs demonstrating size differences between right and left ILCS measures and differences between musicians and non-musicians. (c) Redrawing of the findings of Gaser and Schlaug showing the anatomical locations where stronger grey matter densities are found in musicians. (d) Scattergram showing the relation between cerebellar volume (expressed as per cent of total brain volume) and intensity of daily practice in musicians.

cerebellum (CB) (Gaser and Schlaug 2003; Hutchinson *et al.* 2003). Thus, with respect to M1 these studies support the previous findings of anatomical peculiarities in the primary hand motor region. The findings with respect to the pre-motor, parietal, and cerebellar regions are new. Motor-related regions such as the pre-motor and cerebellar cortex have been shown to play a critical role in planning, preparation, execution, and control of bimanual sequential finger movements. The cluster of differences in cerebellar grey matter found in these studies are located in the area of the cerebellar finger–hand representation as shown in some functional imaging studies (Grodd *et al.* 2001). Further studies have found activation in this area with auditory–musical tasks and during cognitive learning of skills (Gaab and Schlaug 2003). Thus, these aspects also could contribute to the structural differences in the cerebellum.

A further important finding of one of these studies was that musical status (expressed as level of musical skill) and practice intensity (number of hours a day spent practising the instrument) correlated with the measures of grey matter

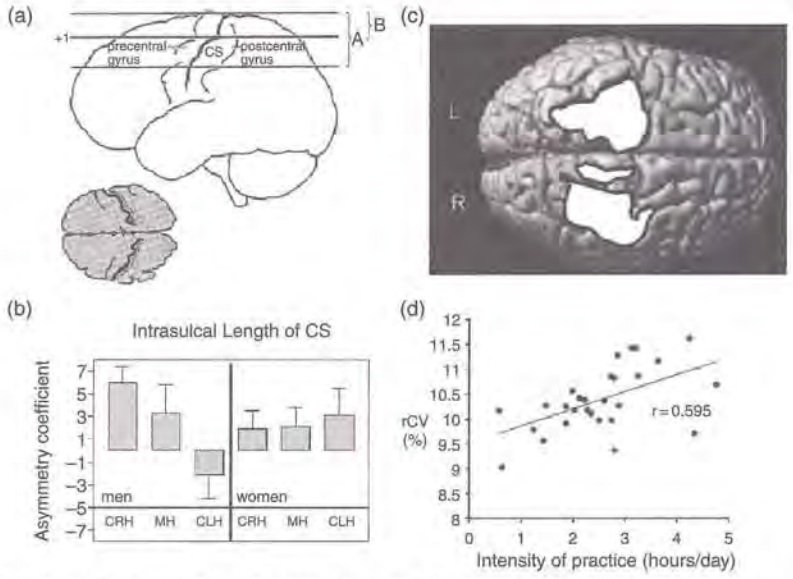



**Fig. 10.3** Schematic description of macrostructural differences between musicians and non-musicians with respect to motor areas. (a) Schematic drawing of the slice position at which the intrasulcal length of the central sulcus (CS) was measured. (b) Bar graphs demonstrating size differences between right and left ILCS measures and differences between musicians and non-musicians. (c) Redrawing of the findings of Gaser and Schlaug showing the anatomical locations where stronger grey matter densities are found in musicians. (d) Scattergram showing the relation between cerebellar volume (expressed as per cent of total brain volume) and intensity of daily practice in musicians.

cerebellum (CB) (Gaser and Schlaug 2003; Hutchinson *et al.* 2003). Thus, with respect to M1 these studies support the previous findings of anatomical peculiarities in the primary hand motor region. The findings with respect to the pre-motor, parietal, and cerebellar regions are new. Motor-related regions such as the pre-motor and cerebellar cortex have been shown to play a critical role in planning, preparation, execution, and control of bimanual sequential finger movements. The cluster of differences in cerebellar grey matter found in these studies are located in the area of the cerebellar finger–hand representation as shown in some functional imaging studies (Grodd *et al.* 2001). Further studies have found activation in this area with auditory–musical tasks and during cognitive learning of skills (Gaab and Schlaug 2003). Thus, these aspects also could contribute to the structural differences in the cerebellum.

A further important finding of one of these studies was that musical status (expressed as level of musical skill) and practice intensity (number of hours a day spent practising the instrument) correlated with the measures of grey matter

density in the motor areas. Thus, musicians with high musical status or with long practise times showed the strongest changes in the grey matter of the motor areas.

Interestingly, there are only a few findings demonstrating structural differences between musicians and non-musicians with respect to the density of white matter. One study using diffusion tensor imaging (DTI) techniques showed that musicians displayed significantly greater fractional anisotropy (FA) in the genu of the corpus callosum, while significantly less FA was found in the corona radiata and the internal capsule bilaterally. Because FA is thought to indicate the efficiency of white matter tracts, these changes in white matter are thought to reflect the anatomical constraint on increased interhemispheric communication (via the corpus callosum), but reduced intracortical communication (via the corona radiata) (Schmithorst and Wilke 2002).

Overall, the anatomical studies mentioned above provide strong arguments for a relatively strong relation between specialized skills (here motor skills relevant to music) and the size, composition, and shape of particular brain structures (especially the motor areas). Indeed, these studies are correlational in nature; thus, they do not permit direct inference as to whether these relationships are due to innate predispositions or to influences dependent on use. If these structural differences are innate, individuals exhibiting such differences in brain anatomy might be (unconsciously?) drawn or even forced to become musicians and, therefore, face fewer obstacles in mastering a musical instrument. Although self-selection for musicianship by individuals with innate structural differences in their brain cannot be completely ruled out, the strong relationship between structural differences and musical status (as shown in the study of Gaser and Schlaug (2003)), as well as data supporting the relationship between age of commencement of musical training and the size of the motor area for the hand (as shown by the study of Amunts *et al.* (1997)) support the proposal that volumetric structural differences seen in musicians might actually be a consequence of long-term music-related motor training. Maybe these use-dependent influences may reveal their entire strength at early ages (1–6 years), when the human brain increases its size from 400 to 1200 g accompanied by an enormous increase in the capacity for many psychological functions (cognitive and motor functions). However, only future work exploiting longitudinal study designs might help to disentangle the relative contribution of predisposition and practice.

In order to accept the effect of use-dependent influences on brain anatomy one has to refer to findings coming from the animal literature. For example, Anderson *et al.* (2002) have shown regional growth and structural adaptation in cerebral grey matter in response to intense environmental demands during

a critical period of brain maturation. In addition, they have shown microstructural changes in the cerebellum, primary motor cortex, and the hippocampus related to the learning of motor skills and continuous motor activity (Anderson *et al.* 1994). These microstructural changes comprise an increase in the number of synapses per neuron and a greater number of glial cells in cortical areas as well as increased glial volume per Purkinje cell in the cerebellar cortex. Interestingly, these anatomical changes were only found in rats trained to perform acrobatic movements. A further animal group which was required to perform routine exercises only had a predominant increase in capillary density, as well as smaller changes in synapse and glial cell density.

## Neurophysiological peculiarities in the sensorimotor system

In the previous sections we have discussed findings from classical behavioural studies and from more recent anatomical studies. Although these studies have considerably increased our understanding of the processes involved in superior musical performance, the possible underlying anatomical changes and the constraints of the musician's brain, the gap between function and anatomy still needs to be bridged. Interestingly, there are only a few studies which have directly investigated neurophysiological functions in musicians during various music-related tasks and which have compared these measures with normal controls. One of the earliest studies examining musicians on the basis of electroencephalography (EEG) measurements was published by Lang *et al.* (1990). They recorded direct current (DC) potentials in 15 musicians before and during the execution of bimanual motor sequences. Unfortunately, they did not examine normal control subjects, thus no direct comparison between musicians and control subjects is possible. The musicians either tapped with their two index fingers in synchrony, or they tapped against triplets ('2 against 3'). Either the right or the left finger started tapping the tapping sequences and after about 4 s the other finger joined in. Ss were free to start the sequences, i.e. to determine the onset times $t_1$ and $t_2$. Shifts of cortical DC potentials were averaged twice: (1) time-locked to $t_1$ and (2) time-locked to $t_2$. When moving in synchrony (easy rhythm) shifts in DC potential and their likely electrical generators indicated activations of the two primary motor cortices (M1). When bimanually tapping different rhythms, there was not only an activation of M1, but also a very large activation of the mesial, central cortex. It was suggested that this mesial motor area (including the supplementary motor area (SMA)) has the function of controlling the initiation of movements in the difficult sequence which have to fit into a very precise timing plan. However, because there is no direct comparison with normal control subjects, it is difficult to delineate the music-related peculiarity of these findings.

A more recent study using a somewhat similar technique has been published by Kristeva *et al.* (2003). These authors made EEG measurements from 58 scalp positions to investigate the temporal sequence and the time course of activation of the activated brain areas while violin players (two advanced musical students and five professional players) executed or imagined a musical sequence on a violin. Using modern methods to analyse and to cortically map current source densities (CSD), the authors showed that three of seven violinists revealed strong activation bilaterally in the frontal opercular region earlier than the motor areas (M1) while a further violinist showed simultaneous activations in these regions. The frontal opercular regions were also strongly activated throughout musical execution or imagining. The supplementary motor area was involved in both preparation for the sequence and during execution and imagination of the sequence. The left primary sensorimotor area was also involved in the preparation and termination of the musical sequence, both for execution and imagination. The main finding of this study is that the frontal operculum (which is part of the ventral pre-motor cortex) is bilaterally involved in the preparation, execution, and imagination of playing a musical piece with the violin. Thus, the ventral pre-motor cortex, which has been suggested to house the human homologue of the so-called 'mirror neurons', is strongly active during complicated motor actions like playing a violin. This suggestion complements findings obtained from studies conducted with non-musicians in which the ventral pre-motor cortex has been shown to be involved in the preparation and execution of various (non-music-related) complicated motor tasks. Most of these tasks have investigated the observation, imagination, and execution of every-day activities like grasping, squeezing, modelling, and tapping (Jäncke *et al.* 2001; Binkofski *et al.* 1999*a*,*b*; Buccino *et al.* 2001, 2004). The new finding of the Kristeva *et al.* (2003) study is that the ventral pre-motor cortex is also involved in music-related activities which have to be acquired due to long-lasting training. Thus, the ventral pre-motor cortex can be trained to participate in the control of complex movements like violin playing.

Lotze *et al.* (2003) examined professional and amateur violinists during actual and imagined performance of Mozart's violin concerto in G major (KV216) using fMRI. Because it is impossible to use a violin within a fMRI scanner the authors asked their subjects to perform their finger movements on their chests. According to the authors the violinists could easily use this task as a substitute for real violin playing after a short training session. The main finding of this study is that professional musicians revealed focused cerebral activations in the contralateral primary sensorimotor cortex, the bilateral superior parietal lobes, and the ipsilateral anterior cerebellar hemisphere. In addition, they also found that professional musicians exhibited higher activity of the right primary auditory

cortex during execution. Interestingly, a similar pattern of cortical activation was found during the imagination of violin-playing movements. However, there was no activation within the auditory cortex during imagination of movement. These findings support the assumption that musicians execute music-relevant finger sequences with higher motor control economy as amateur musicians. The finding of an absence of activation of the auditory cortex during imagined movements was interpreted as evidence that activation of the auditory cortex is only present if the interconnected motor areas are activated to control real music-relevant movements.

A very recent study employed a similar logic to examine musicians during motor tasks which only musicians can perform (Meister *et al.*, 2004). In this fMRI study, the authors investigated the cortical network mediating music performance compared with music imagery in 12 music academy students. As task, the subjects were required to execute or imagine playing the right-hand part of a Bartok piece. There were two important findings in this study: first, during imagery and execution a similar bilateral frontoparietal network was active comprising the pre-motor areas, the pre-cuneus, and the medial part of the superior parietal lobe (SPL); secondly, during music performance (but not during imagery) the contralateral primary motor cortex and the posterior parietal cortex (PPC) were active bilaterally. Thus, both imagery and actual execution of a musical piece partly evoke activation in a similar cortical network. However, imagination of movement does not activate the primary motor cortex and the PPC; thus these areas are probably more involved in the actual execution of music performance.

As mentioned above, although these three studies have shown that distributed brain areas are involved in the control of playing a musical instrument, they do not directly help us to understand what is different or special in the motor system of musicians. Although this is an interesting question it is not easy to examine, because many motor activities relevant to musical performance are so special that only musicians are in a position to generate them. Thus, one has to find tasks which are very close to musical tasks but which also can be performed (in principle) by non-musicians. Three studies have been published so far in which pianists have been compared with non-musicians during different motor tasks in the context of fMRI measurements. Krings *et al.* (2000) used an overlearned unimanual sequential finger opposition task for the dominant right hand and found significant differences in the extent of cerebral activation between both musicians and non-musicians, with piano players (four right-handed professional piano players with an average 4 ± 8 h of training per day) having a smaller number of activated voxels in the primary motor cortex, the supplementary motor area, the pre-motor cortex, and the superior parietal lobule.

In the Jäncke *et al.* (2000) study two professional pianists were compared with two carefully matched non-musicians in terms of their cortical activation pattern during a bimanual tapping task. The bimanual tapping task was a variant of the so-called 'rubato-task' which comes very close to piano playing (a similar task was used by the early EEG-study of Lang *et al.* (1990)). The subjects are required to tap with one hand while the other hand makes two taps. After a short training session the control subjects were basically able to perform similarly to the pianists; however, the cortical activation pattern in the pianists and non-musicians was considerably different. Non-musicians revealed stronger haemodynamic responses within the left primary (M1) motor area and bilaterally in the SMA proper. However, when the sub-dominant hand tapped at the faster rate non-musicians revealed very strong activations in mesial motor areas including the pre-SMA and the rostral and caudal CMA. There were also stronger activations in the right M1, LPi, and in a posterior paracentral region including the pre-cuneus. In recent studies it has been shown that these motor areas are activated when motor control becomes more complex and unpredictable for both unimanual and bimanual movements. In addition, brain-imaging studies in humans have shown that the region presumably corresponding to the CMA is partly active when a variety of motor tasks require subjects to voluntarily select movements (Koeneke *et al.* 2004). Related to these findings, one may assume that non-musicians rely more strongly on explicit and voluntarily motor control strategies while professional piano players control their movements in a more automatic fashion.

Thus, these two studies suggest that professional pianists use smaller neural networks within the primary and secondary motor areas in order to control unimanual and bimanual movements. If this contention is correct, one might assume that professional pianists or violinists control their movements much more efficiently. One reason for an increased efficiency of motor control might be related to the 'degrees of freedom problem'. According to this theoretical assumption, different muscles are functionally linked together and controlled conjointly. In this context, it is speculated that with increasing levels of motor skill more and more effectors are linked together, thus reducing the number of 'degrees of freedom' to be controlled via motor commands. In this sense, highly trained pianists most likely control fewer 'degrees of freedom' for these tasks, thus enabling them to control uni- and bimanual movements much more efficiently with smaller neural networks than non-musicians. This in turn would increase the control capacity of pianists because they now can control more 'degrees of freedom' or more motor programs with a given network. Considering the fact that highly trained musicians also have larger hand motor areas, this would suggest that they have an efficiently organized neural network as well as a generally larger network at their disposal in the hand motor area.

A study which goes beyond the simple comparison between pianists and non-musicians during different motor tasks is the study published by Hund-Georgiadis and von Cramon (1999). In this study the authors investigated haemodynamic responses in several motor areas during short-term motor learning of the dominant right hand in 10 piano players (PPs) and 23 non-musicians (NMs), using a complex finger-tapping task. In summary they found the following results:

1. All subjects achieved considerable increases in tapping frequency during the training session of 35 min in the scanner.
2. PPs, however, performed significantly better than NMs and showed increasing activation in the contralateral primary motor cortex throughout motor learning in the scanner. At the same time, the involvement of secondary motor areas, such as the bilateral supplementary motor area, the pre-motor, and the cerebellar areas, diminished in relation to the NMs throughout the training session.
3. Extended activation of primary and secondary motor areas in the initial training stage (7–14 min) and rapid attenuation were the main functional patterns underlying short-term learning in the NM group; attenuation was particularly marked in the primary motor cortices as compared with the PPs.
4. When tapping of the rehearsed sequence was performed with the left hand, transfer effects of motor learning were evident in both groups. Involvement of all relevant motor components was smaller than after initial training with the right hand. Ipsilateral pre-motor and primary motor contributions, however, showed slight increases of activation, indicating that dominant cortices influence complex sequence learning of the non-dominant hand.

The very new finding of this study is that experience (motor skill) influences the contribution of secondary motor areas in the motor learning of simple sequential unimanual tasks. In concordance with the studies mentioned above, this study also demonstrates minor involvement of the SMA in the PP group and shows attenuation effects in both groups during motor learning. Modern brain imaging studies have demonstrated that the mesial motor areas (especially the SMA proper and the pre-SMA) are strongly involved in the acquisition and execution of complex unimanual and bimanual tasks. Thus, the stronger the activation in these areas the more complex is the movement control or the more demanding is the movement control. With increasing level of skill the complexity of movement control and the demands of movement control diminish. Therefore with increasing levels of skill fewer neurons in these areas will be involved during the control of unimanual or bimanual movements, as in the case of professional musicians.

The above-mentioned fMRI studies have shown that the motor system of musicians has specifically adapted to the demands of playing a musical instrument. Although these studies have considerably furthered our understanding of these adaptive processes there is one particular disadvantage associated with fMRI data: it is basically impossible to infer from increased or decreased haemodynamic responses whether these changes are due to cortical excitation or inhibition. In order to know more about the particular neurophysiological processes underlying these haemodynamic changes more direct neurophysiological data are needed.

A typical example of this kind of study is the work of Nordstrom and Butler (2002). They studied intracortical inhibition (ICI) and intracortical facilitation (ICF) within the motor cortex ipsilateral to the moving finger in 15 musicians and 15 non-musician controls using transcranial magnetic stimulation (TMS). TMS was used to elicit muscle evoked potentials (MEPs) from left first dorsal interosseous (FDI) muscle at rest and during voluntary abduction of the index finger. A paired-pulse TMS technique was used to test for ICI or ICF. The basic principle is that a subthreshold TMS stimulus is given followed by a TMS pulse evoking a MEP. Depending on the ISI and the neurophysiological condition of the stimulated motor cortex, the amplitude of the MEP will vary in size. For example, if the MEP is evoked shortly after the first subthreshold TMS stimulus (1–5 ms), it is normally lower in amplitude, indicating cortical inhibition. MEPs evoked 8–15 ms after the first TMS pulse are normally enhanced with respect to the amplitude. Thus, this enhancement effect is taken as evidence of cortical facilitation effects. Using this technique the authors found that ICI was less effective in musicians with FDI at rest and during activity, but only with an ISI of 3 ms. In addition, ICF was less effective in musicians under conditions of both rest and activity. The authors argue that the reduced ICI and ICF in musicians could be due to reduced excitability of the neurons responsible for ICI or ICF or to an altered balance of excitatory inputs to corticospinal neurons which favours neurons that are not acted upon by the ICF circuit. The reduced influence of ICI and ICF circuits on the excitability of corticospinal neurons in musicians is likely to reflect a training-induced adaptation. Although this interpretation is somewhat speculative and the underlying neurophysiological aspects are not fully understood, these findings demonstrate that neurophysiological activation within the primary motor cortex is different in musicians.

A further study by the same group (Ridding *et al.* 2000) also using TMS tested for possible differences with respect to interhemispheric communication. Here they used paired-pulse TMS to assess interhemispheric inhibition in six adult professional musicians who began musical training at an early age. Conditioning TMS was applied to the hand area of the motor cortex of one hemisphere,

followed 4–16 ms later by a test stimulus applied to the other hemisphere. Tests were performed at rest, and with the first dorsal interosseous muscle contralateral to the conditioning hemisphere voluntarily active. Conditioning TMS in musicians was 29% less effective at reducing the size of the test MEP at rest, and 63% less effective in the active condition, compared with control subjects. Thus, there is obviously less transcallosal interhemispheric inhibitory influence in musicians than in controls. This finding is quite important for interpreting the anatomical findings with respect to the size of the corpus callosum because this study shows that there is less interhemispheric inhibition in musicians, at least for the hand motor areas. Given that the corpus callosum is larger in musicians and that large areas of corpus callosum normally indicate more and larger myelinated axons crossing through the midline, this study would indicate that more axons support less interhemispheric inhibition. However, although interesting, more work is needed to explain this finding more sufficiently.

The above-mentioned studies have shown that there are peculiar differences between musicians and non-musicians in terms of simple motor measures, macrostructural peculiarities of the motor system, and neurophysiological activation patterns of the motor areas. Whether these differences are due to some kind of pre-wired or genetic prerequisites supporting musical performance or whether these differences are simply a consequence of the use-dependent stimulations associated with learning an instrument is still a matter of dispute. Although many data support the notion that musical training shapes the brain both in terms of anatomy and functional organization (Münte *et al.* 2002; Schlaug 2001) direct evidence for this use-dependent influence is still missing. An interesting contribution supporting the idea that use-dependent stimulation in fact stimulates cortical organization has been published recently by Bangert and Altenmüller (2003). They recorded changes in cortical activation patterns (DC-EEG potentials) induced by short- (20 min) and long-term (5 week) piano learning using auditory and motor tasks. Two groups both new to piano playing were trained. One group was allowed to learn the standard piano key-to-pitch map (the 'map' group). For the other group (the 'no-map' group), random assignment of keys to tones prevented such a map. From the EEG obtained during the training session, the authors calculated the EEG coherence and DC intensity. They found marked changes in the EEG (both for EEG coherence and EEG DC intensity) even after a short training period of only 20 min. This effect was enhanced after 5 weeks of training. The increased DC changes were most prominent in the left central and right anterior regions if the training was conducted with the right hand. The right anterior activation was only present for the 'map' group which was in the position to associate auditory cues (the notes) with a particular motor action. Obviously, the frontal EEG activity captures the

neural activations of pre-motor areas which are known to associate sensory cues with motor programs. Interestingly, these associations are generated very early during motor learning.

## Conclusions

As demonstrated in this chapter there are several neuroanatomical and neurophysiological studies supporting the idea that intense practice on a musical instrument stimulates cortical adaptations. These adaptations can be seen at macroanatomical levels, with increased volumes and grey matter densities of many motor and sensorimotor areas. They can also be seen on the level of neurophysiological activation indexed by changed haemodynamic responses or EEG measures attributed to the motor areas. These data show that the entire motor network, including the pre-motor, primary sensorimotor, parietal, and cerebellar regions, acts differently in musicians. Interestingly, these differences can also be found in experimental situations when subjects imagined playing a musical instrument. In addition, several of these studies support the notion that early and/or intensive musical training are the main variables determining the strength and impact of the reorganization of the motor system. Future studies have to show whether these plastic changes are possible during the entire life or are restricted to critical periods of life.

## References

**Amunts, K., Schlaug, G., Schleicher, A.** *et al.* (1996). Asymmetry in the human motor cortex and handedness. *Neuroimage*, **4**, 216–222.

**Anderson, B.J., Eckburg, P.B., and Relucio, K.I.** (2002). Alterations in the thickness of motor cortical subregions after motor-skill learning and exercise. *Learning Memory*, **9**, 1–9.

**Anderson, B.J., Li, X., Alcantara, A.A., Isaacs, K.R., Black, J.E., and Greenough, W.T.** (1994). Glial hypertrophy is associated with synaptogenesis following motor-skill learning, but not with angiogenesis following exercise. *Glia*, **11**, 73–80.

**Bangert, M. and Altenmüller, E.O.** (2003). Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neurosci*, **4**, 26.

**Binkofski, F., Buccino, G., Posse, S., Seitz, R.J., Rizzolatti, G., and Freund, H.** (1999*a*). A fronto-parietal circuit for object manipulation in man: evidence from an fMRI-study. *Eur J Neurosci*, **11, 3276–3286.**

**Binkofski, F., Buccino, G., Stephan, K.M., Rizzolatti, G., Seitz, R.J., and Freund, H.J.** (1999*b*). A parieto-premotor network for object manipulation: evidence from neuroimaging. *Exp. Brain Res*, **128**, 210–213.

**Buccino, G., Binkofski, F., Fink, G.R.** *et al.* (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur J Neurosci*, **13**, 400–404.

**Buccino, G., Vogt, S., Ritzl, A.** *et al.* (2004). Neural circuits underlying imitation learning of hand actions; an event-related FMRI study. *Neuron*, **42**, 323–334.

Cook, N.D. (1984). Callosal inhibition: the key to the brain code. *Behav Sci*, *29*, 98–110.

Drake, C. and Ben El, H.J. (2003). Synchronizing with music: intercultural differences. *Ann NY Acad Sci*, **999**, 429–437.

Drost, U.C., Rieger, M., Brass, M., Gunter, T.C., and Prinz, W. (2004). Action-effect coupling in pianists. *Psychol Res*

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., and Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, **270**, 305–307.

Ericsson, K.A., Krampe, R.T., and Heizmann, S. (1993*a*). Can we create gifted people? *Ciba Found Symp*, **178**, 222–231.

Ericsson, K.A., Krampe, R.T., and Tesch-Roemer, C. (1993*b*). The role of deliberate practice in the acquisition of expert performance. *Psychol Rev*, **100**, 363–406.

Gaab, N. and Schlaug, G. (2003). The effect of musicianship on pitch memory in performance matched groups. *Neuroreport*, **14**, 2291–2295.

Gaser, C. and Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *J Neurosci*, **23**, 9240–9245.

Grodd, W., Hulsmann, E., Lotze, M., Wildgruber, D., and Erb, M. (2001). Sensorimotor mapping of the human cerebellum: fMRI evidence of somatotopic organization. *Hum Brain Mapping*, **13**, 55–73.

Hepp-Reymond, M.C. (1988). Functional organization of motor cortex and its participation in voluntary movements. In: H.D.Seklis and J. Erwin (eds) *Comparative primate biology*, pp. 501–624. Liss, New York.

Hund-Georgiadis, M. and von Cramon, D.Y. (1999). Motor-learning-related changes in piano players and non-musicians revealed by functional magnetic-resonance signals. *Exp Brain Res*, **125**, 417–425.

Hutchinson, S., Lee, L.H., Gaab, N., and Schlaug, G. (2003). Cerebellar volume of musicians. *Cereb Cortex*, **13**, 943–949.

Jäncke, L. (2002). The case of a left-handed pianist playing a reversed keyboard: a challenge for the neuroscience of music. *Neuroreport*, **13**, 1579–1583.

Jäncke, L., Schlaug, G., and Steinmetz, H. (1997). Hand skill asymmetry in professional musicians. *Brain Cogn*, **34**, 424–432.

Jäncke, L., Shah, N.J., and Peters, M. (2000). Cortical activations in primary and secondary motor areas for complex bimanual movements in professional pianists. *Brain Res Cogn Brain Res*, **10**, 177–183.

Jäncke, L., Kleinschmidt, A., Mirzazade, S., Shah, N.J., and Freund, H.J. (2001). The role of the inferior parietal cortex in linking the tactile perception and manual construction of object shapes. *Cereb Cortex*, **11**, 114–121.

Koeneke, S., Lutz, K., Wüstenberg, T., and Jäncke, L. (2004). Bimanual versus unimanual coordination: what makes the difference? *Neuroimage*, **22**, 1336–1350.

Krampe, R.T. (2002). Aging, expertise and fine motor movement. *Neurosci Biobehav Rev*, **26**, 769–776.

Krampe, R.T., Engbert, R., and Kliegl, R. (2002). The effects of expertise and age on rhythm production: adaptations to timing and sequencing constraints. *Brain Cogn*, **48**, 179–194.

**Krings, T., Topper, R., Foltys, H.** *et al.* (2000). Cortical activation patterns during complex motor tasks in piano players and control subjects. A functional magnetic resonance imaging study. *Neurosci Lett*, **278**, 189–193.

**Kristeva, R., Chakarov, V., Schulte-Monting, J., and Spreer, J.** (2003). Activation of cortical areas in music execution and imagining: a high-resolution EEG study. *Neuroimage*, **20**, 1872–1883.

**Laeng, B. and Park, A.** (1999). Handedness effects on playing a reversed or normal keyboard. *Laterality*, **4**, 363–377.

**Lang, W., Obrig, H., Lindinger, G., Cheyne, D., and Deecke, L.** (1990). Supplementary motor area activation while tapping bimanually different rhythms in musicians. *Exp Brain Res*, **79**, 504–514.

**Lotze, M., Scheler, G., Tan, H.R., Braun, C., and Birbaumer, N.** (2003). The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *Neuroimage*, **20**, 1817–1829.

**Meister, I.G., Krings, T., Foltys, H.** *et al.* (2004). Playing piano in the mind-an fMRI study on music imagery and performance in pianists. *Brain Res.Cogn Brain Res*, **19**, 219–228.

**Melsbach, G., Wohlschläge, A., Spiess, M., and Güntürkün, O.** (1996). Morphological asymmetries of motoneurons innervating upper extremities: clues to the anatomical foundations of handedness? *Int J Neurosci*, **86**, 217–224.

**Meyer, A.** (1977). Music and the brain. In: C.Macdonald and R.A. Henson (eds) Heinemann Medical Books, London.

**Münte, T.F., Altenmüller, E., and Jäncke, L.** (2002). The musician's brain as a model of neuroplasticity. *Nat Rev Neurosci*, **3**, 473–478.

**Nordstrom, M.A. and Butler, S.L.** (2002). Reduced intracortical inhibition and facilitation of corticospinal neurons in musicians. *Exp Brain Res*, **144**, 336–342.

**O'Neill, S.** (1997). The role of practice in early musical performance achievement. In: H.Jorgensen and A. Lehmann (eds) *Does practice make perfect? Current theory and research*, pp. 53–70. Norges Musikhogskole, Oslo.

**Peters, M.** (1976). Prolonged practice of a simple motor task by preferred and non-preferred hands. *Percept Mot Skills*, **42**, 447–450.

**Peters, M. and Durding, B.M.** (1978). Handedness measured by finger tapping: a continuous variable. *Can J Psychol*, **32**, 257–261.

**Ridding, M.C., Brouwer, B., and Nordstrom, M.A.** (2000). Reduced interhemispheric inhibition in musicians. *Exp Brain Res*, **133**, 249–253.

**Schlaug, G.** (2001). The brain of musicians. A model for functional and structural adaptation. *Ann NY Acad Sci*, **930**, 281–299.

**Schlaug, G., Jäncke, L., Huang, Y., Staiger, J.F., and Steinmetz, H.** (1995). Increased corpus callosum size in musicians. *Neuropsychologia*, **33**, 1047–1055.

**Schmithorst, V.J. and Wilke, M.** (2002). Differences in white matter architecture between musicians and non-musicians: a diffusion tensor imaging study. *Neurosci Lett*, **321**, 57–60.

**Sloboda, J. and Davidson, J.** (1996). The young performing musician. In: I.Deliege and J. Sloboda (eds) *Musical beginnings*. Oxford University Press, Oxford.

**Snyder, P.J., Bilder, R.M., Wu, H., Bogerts, B., and Lieberman, J.A.** (1995). Cerebellar volume asymmetries are related to handedness: a quantitative MRI study. *Neuropsychologia*, **33**, 407–419.

**Volkmann, J., Schnitzler, A., Witte, O.W., and Freund, H.** (1998). Handedness and asymmetry of hand representation in human motor cortex. *J Neurophysiol*, **79**, 2149–2154.

Chapter 11

# Brain activation during piano playing

Marc Bangert

The chapter gives an overview of brain imaging in piano playing. Because the piano keyboard, more so than other musical instruments, offers a versatile tool that integrates well into parameter-controlled experimental laboratory set-ups, a large number of studies have been carried out addressing issues not only of instrumental performance but also general questions about music processing, working memory, and learning and neuroplasticity. This chapter examines in detail an electroencephalography (EEG) study covering the topics of auditory–sensorimotor maps and rapid motor plasticity. The last part of the chapter provides a meta-review of research uncovering areas of the brain that contribute to various music-processing tasks. The review speculates on a possible core network for music, and points out its relationship to the network dedicated to speech and language.

## Introduction

> Attunement of ears, mind and hands. Most of us possess these pieces of anatomy but sometimes it would seem that they are not 'wired together properly.' Each has an important function to perform by itself, but if each is not attuned to the other two, then one function cannot pass to the next, and we are momentarily out of order. When properly 'wired', the ear will hear something, either in the mind's ear or from an outside source (another member in the group), the aural sensation being forwarded to the mind, where intellectual computations and translations take place. Then the mind issues directives to the muscles of the hands, which hopefully will respond obediently, quickly and cleanly. Now the ears hear the result and coordinately with the mind decide what to hear and play next, repeating the whole cycle again and again.
>
> Coker (1997)

What Jerry Coker, the renowned and inspired jazz instructor, outlines above is (maybe not accidentally) the most precise description of the challenges with which the central nervous system is faced when making music at a professional level. Professional musicians excel in one of the most complex and demanding sensorimotor tasks—music performance requiring high-speed motor control under a rigorous and relentless auditory feedback.

How is this extremely rapid exchange of information between auditory and motor areas coordinated? How is motor control fine-tuned by auditory feedback, and how is the feedback system prepared for what is to be expected?

Brain networks are known to be subject to short- and long -term plastic changes. Changes due to the training of motor skills can occur very rapidly, in the range of minutes (Pascual-Leone *et al.* 1995; Classen *et al.* 1998), and lifelong practice might even alter macrostructural anatomy (Schlaug *et al.* 1995). Highly trained pianists often state that while listening to piano music their fingers feel like they are irresistibly itching to play along, whereas while practising silent drills, e.g. 'tabletop dexterity', these musicians can vividly 'hear' what they are playing.

When playing musical instruments, auditory information has to be processed together with sensorimotor data. Thus, the same music is represented in multiple ways: as sounds, as movements (for example on a piano keyboard), as symbols (notes on a score), and so on.

## Brain activation during performance: a distributed network

The production and performance of music is not a matter for professional musicians alone. Almost everyone can sing, hum, and whistle. When it comes to playing a musical instrument, the more natural process of vocal motor control is substituted for by (mostly) hand and finger movement, and now additional areas of the brain have to be recruited. In general, the majority of studies dealing with the production of music in professional musicians reveal a common network, comprising pre-motor areas (PM), the supplementary motor area (SMA), primary motor areas (M1), the dorsolateral pre-frontal cortex (DLPFC) including Broca's area in both hemispheres, and the posterior parietal cortex (PPC).

In one of the early functional neuroimaging studies on piano performance (Sergent *et al.* 1992), positron emission tomography (PET) was used to measure regional cerebral blood flow in pianists during right-handed performance of a sight-read musical score. Among the main findings of this study was the observation that cortical areas distinct from, but adjacent to, those underlying language operations were activated during sight reading. A contribution of areas that have classically been related to speech to instrumental performance has been confirmed many times over the subsequent 12 years, in a variety of paradigms and with several brain imaging techniques.

In professional piano players there are learning-related changes in cerebral motor representation; thus it was found that professional piano players show less cerebral activation in pre-motor and motor areas when compared with non-musicians performing the same movement (Jäncke *et al.* 2000; Krings *et al.* 2000).

Itoh *et al.* (2001) demonstrated that the areas activated during a pianist's performance included the PPC and the primary sensorimotor areas (SM1). Because the paradigm elegantly distinguished performance with the right hand, left hand, or both hands, these authors could disentangle the effector-specific (hand) and task-specific (music) portions of the ongoing brain activity. While unilateral SM1 activation was correlated with motion of the corresponding contralateral hand, PPC activation was correlated with piano performance irrespective of hand modality.

Using functional magnetic resonance imaging (fMRI), Meister *et al.* (2004) compared music performance and music imagery in 12 musicians performing the right-hand part of a Bartok piece. In both conditions they found activation of a bilateral frontoparietal network comprising the pre-motor areas, the pre-cuneus, and the medial part of BA 40. During music performance, but not during imagery, the contralateral M1 and PPC were active bilaterally. This was interpreted to reflect the role of the primary motor cortex in motor execution and the requirements for visuomotor integration during music performance. The only cortical activation with a clear lateralization in their study was found in the left posterior part of the dorsolateral pre-frontal cortex (Meister *et al.* 2004).

A right-sided dominance of the dorsolateral/inferior parts of the frontal lobe was found in several studies. The activation of the right inferior frontal gyrus was suggested to be related specifically to the production of music (Langheim *et al.* 2002). Carter *et al.* (1995) showed activation of this area during selective attention tasks and working memory. An fMRI study conducted by Rao *et al.* (1997) also suggested a role for the right inferior frontal gyrus in the production of music. These results indicate that activation of the right inferior frontal gyrus supports the retrieval and rehearsal of internal non-linguistic auditory representations of time intervals. The right-hemisphere Broca homologue was also shown to be involved in explicit learning of motor sequences and associative motor learning (Seitz and Roland 1992).

## Brain activation during imagined performance

It is known that excellent musicians not only rehearse their musical skills by daily practice but also use strategies of mental imagery. For example, Horowitz used to practice mentally before playing in concerts in order to avoid the feedback of a piano other than his own. Rubinstein did so in order to make his daily practice more efficient (Schoenberg 1987).

The main component of mental rehearsal is motor imagery, i.e. a mental simulation of action which involves many parts of the sensorimotor system that are active during execution. It therefore provides a possibility for the central nervous system to evaluate the consequences of future actions and to prepare

properly for the actual execution (Jeannerod 2001). The relationship between motor imagery and motor action can be deduced from the observation that the timing patterns of both processes are similar (Sirigu *et al.* 1995) and that the changes in corticospinal excitability involve the same muscles (Fadiga *et al.* 1999). Functional imaging studies on motor imagery versus performance in non-musicians have shown that imagery and performance activate essentially the same cortical regions (Porro *et al.* 1996; Lotze *et al.* 1999; Gerardin *et al.* 2000), with the exception, however, of the primary motor cortex: the results regarding the involvement of the primary motor cortex in motor imagery differ between studies.

Using fMRI, Meister *et al.* (2004) compared music performance with music imagery in professional musicians (see above). During imagery they found bilateral activation in the pre-motor areas, the pre-cuneus, and the medial part of BA 40. The only missing brain response (when compared with actual execution) was the contralateral M1 and bilateral PPC.

## Maps, links, and plasticity

Years of practice may establish a neuronal correlate of an audiomotor connection; this has recently been shown by brain imaging studies for both directions, auditory-to-motor and motor-to-auditory. Professional musicians often report that listening to a well-performed piece of music can involuntarily trigger the respective finger movements. In a magnetoencephalography (MEG) experiment, Haueisen and Knösche (2001) demonstrated that when listening to well-performed piano music pianists exhibit involuntary motor activity involving the contralateral primary motor cortex (M1). To investigate motor-to-auditory processing as a possible feedforward projection, Lotze *et al.* (2003) collected fMRI brain scans of violinists and amateurs as they silently tapped out the first 16 bars of Mozart's violin concerto in G major. The expert performers had significant activity in the primary auditory regions, which was missing in the amateurs.

Using electrophysiological or neuroimaging methods, adaptive changes in the cortical organization similar to those found after stimulation or injury (Merzenich *et al.* 1988; Jacobs and Donoghue 1991; Wang *et al.* 1995) have been found in humans after skill learning and as an adjustment after injury (Seitz *et al.* 1990; Pascual-Leone *et al.* 1994, 1995; Karni *et al.* 1995; Hund-Georgiadis and von Cramon 1999; Charness and Schlaug 2000). Although some studies have pointed out the rapid reversal of representational changes, other studies have found persistent alterations in response to the early acquisition of fine sensorimotor skills such as there being a larger sensory finger representation in the left hand of string players (Elbert *et al.* 1995).

In order to clarify which of the peculiarities of musicians' brains are actually induced by practice, it is necessary to observe plastic processes from the very

beginning of learning to play an instrument. This provides a unique opportunity to disentangle the different parameters that form the network required for playing an instrument. These contributing parameters are sight-reading, visual–auditory transfer (knowing how a musical score will sound), anticipatory auditory imagery, motor planning and execution, audiomotor feedback control, visuomotor control, and analysis of various somatosensory feedback channels.

Pascual-Leone *et al.* (1995) used focal transcranial magnetic stimulation (TMS) to map the motor cortical areas targeting long finger flexor and extensor muscles bilaterally. They showed that as subjects learned a five-finger exercise on the piano over the course of 5 days under auditory feedback, the area of cortical representation for the long finger flexor and extensor muscles enlarged significantly (Pascual-Leone *et al.* 1995; Pascual-Leone 2001). However, this increase could only be demonstrated when the cortical mapping studies were conducted following a 20- to 30-min rest after the practice (and test) session. No such modulation in the cortical output maps was noted when maps were obtained before each daily practice session (Pascual-Leone *et al.* 1999).

Karni *et al.* (1998) provided evidence that a few minutes' daily practice of a finger opposition sequence induced large, incremental performance gains over a few weeks of training, which were associated with changes in cortical movement representation within M1. The authors suggested that these changes reflected the initial set-up of a task-specific motor program, while the subsequent changes indicated the consolidation of this motor program.

Stewart *et al.* (2003) focused on the acquisition of sight-reading skills. They investigated musically naïve participants before and after a period of 15 weeks during which they were taught to read music and play the keyboard. When participants played melodies from musical notation after training, fMRI activation was seen in the superior parietal cortex bilaterally. These authors suggest that reading music involves the automatic sensorimotor translation of a spatial code (written music) into a series of motor responses (key presses).

Sight-reading, however, is not necessarily at the core of music performance—as illustrated by playing by heart or by musical improvisation. In the following, an experiment is presented that sets out to specify the role of the right frontal activity found in several studies mentioned above.

## EEG correlates of auditory sensorimotor integration during piano learning

In professional musicians, the cortical activation patterns during performance display widely distributed networks far beyond a simple representation of the modalities involved (audible tones and the digits of the hands; Bangert *et al.* 2005).

Bangert, Altenmüller, and Parlitz at the Institute of Music Physiology and Musicians' Medicine (Hannover, Germany), have investigated the question of how this apparent multimodal and sensorimotor representation is acquired during practice. They invited 17 musically naïve right-handed subjects to be trained twice a week, over a period of 5 weeks, to play the piano with an interactive computer system. They listened to short (3 s) piano melodies and practiced replaying these melodies with the five digits of their right hands as accurately as possible. The system was adaptive: an online performance analysis of the replayed melodies determined the tempo and complexity of subsequently presented targets—the better the performance, the harder the challenge. The subjects' sole feedback on whether a trial was successful was an implicit one: they heard themselves playing.

The EEG acquisition paradigm differed from the practice settings. Whereas the training procedure aimed at sensorimotor binding, for the actual measurement the auditory and motor features of the practiced task were dissociated, as outlined by the basic idea above (Plate 9). The subjects' event-related slow DC-EEG potentials were measured at 30 electrode sites *either* while *passively listening* to short melodic piano sequences (similar to the practice melodies) *or* while arbitrarily *pressing keys on a mute piano keyboard*. The term 'passive listening' refers to the instruction just to 'take the music in', without any attempt to analyse or memorize it.

## Automatic co-activation of the cortical representations of 'ears' and 'hands'

Before the first practice (Plate 9), the average non-musician's brain exhibited textbook behaviour: auditory musical processing activated more frontal and central areas, and plain motor execution was accompanied by activity in the left primary motor area. (Note that all the training was performed with the right hand.)

The first signs of plasticity occurred after just 20 min of practice (Plate 9). The passive listening led to a widespread additional activation immediately following the very first session at the piano, but in the vicinity of the central sulcus clearly lateralized to the left (again, the right hand). The subsequent measurements (Plate 9) were collected after five and ten training sessions.

The co-activation of the left sensorimotor cortex was eventually focused in a spatially distinct region. Moreover, in the case of the subjects described here, the frontal activity showed a salient focus of activation at a frontotemporal area of the right cerebral hemisphere (in the following, abbreviated rFTA). The rFTA comprises an anterior part of the superior temporal gyrus (STG) and the inferior frontal cortex. Information has been published on the localization and function of rFTA components using various methods: lesion studies (Samson and Zatorre

1988; 1992; Zatorre and Samson 1991; Zatorre and Halpern 1993; Liegeois-Chauvel *et al.* 1998), PET (Zatorre *et al.* 1994, 1996; Griffiths *et al.* 1998; Halpern and Zatorre 1999), fMRI (Langheim *et al.* 2002), and MEG (Tervaniemi *et al.* 1999).

The right panel of Plate 9 depicts the simultaneous development of cortical activation for the silent motor movement paradigm. Activity in the parietal lobe again increased during the first half of the training period (until week 3), but subsequently decreased by the end of the 5-week session. The most remarkable change can be seen on the side ipsilateral to the moving hand: it looks as if the sensorimotor cortex of the unused left hand had been inhibited during the first session and remained in that state ever after. Again, in addition to these effects, development of some prominent area of activation around the rFTA was observed. Since this area has been demonstrated to be involved in the *perception* of pitch sequences, this effect could provide an explanation for why skilled pianists often state that silent practice creates an apparent auditory impression (or illusion) of the music they are playing.

## Two sides of a coin

During data analysis an increasing similarity of the brain activation patterns for the auditory and motor tasks was noticed, until in the expert performers it became hard to tell which was which. The scalar product of two vectors of unit length gives a measure of the similarity of the two activation patterns. Nearly identical patterns would then result in a value close to unity, inverse patterns would yield –1, and differently shaped (orthogonal) patterns would dot-multiply to zero. In contrast to the Euclidean distance, the normalized dot product is robust against non-specific effects with small potentials or differing potentials, and accounts for 'qualitative' similarity in pattern only.

Figure 11.1 illustrates what this method yields when applied to the 'passive auditory'/'silent motor' resemblance. As expected from the appearance of the patterns, the grand average of the beginners' 'auditory-by-motor' scalar product turned out to increase with accumulated practice time (see Fig. 11.2). For the expert performers, perception and action seem to be just two aspects of one integrative skill; in one case, a 38-year-old pianist with 32 years of instrumental practice revealed a staggering value of $d(t) = 0.98$ (average in the time window from 0–2000 ms after onset of music/movement).

## Learning a sensorimotor map

The next step set out to clarify the role of the rFTA, which, strikingly, appeared active in the auditory as well as in the silent motor task. Therefore, a control group of another eight naïve subjects was established. They were administered

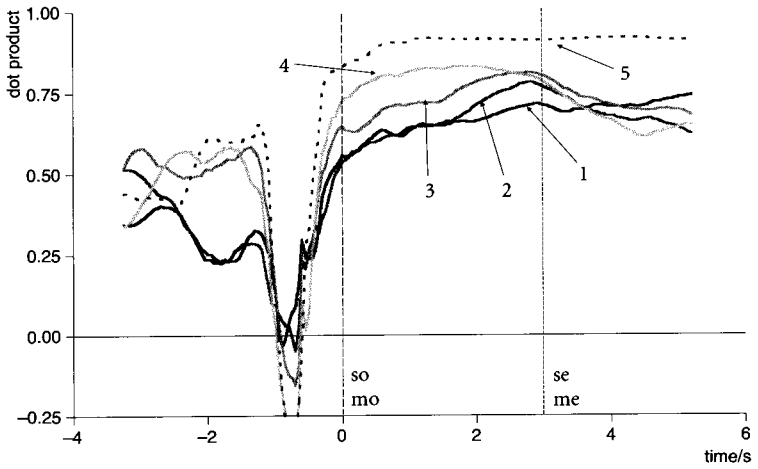



**Fig. 11.1** Scalar product of the normalized passive auditory/silent motor brain activation vectors, as a function of time. so = stimulus onset, se = stimulus end, mo = movement onset, me = movement end. Curve 1, novices ($n$ = 9) before the first session; curve 2, novices after 20 min of practice; curve 3, novices after about 21 days of practice; curve 4, novices after about 39 days of practice; curve 5, professionals ($n$ = 3).

the same 5-week training programme (reproducing acoustically presented melodies as precisely as possible with respect to note order, rhythmic timing, and loudness of the piano keystrokes) with one slight modification: the assignment of the five relevant notes (pitches) to the five relevant keys, or fingers, respectively, was shuffled after each single training trial, so that the control group was not given any chance to figure out any coherence between fingers and notes except the temporal coincidence of keystroke and sound. In other words they had no opportunity to establish an internal 'map' between motor events and auditory targets.

Still, the control subjects had to practice proper timing and control of fine finger forces for the reproduction of rhythm and loudness. In fact, as the behavioural data indicated, the control subjects learned to perform these aspects of playing an instrument better than the 'regular learners' (Fig. 11.2a).

Figure 11.2b shows the activation differences between the two groups. We can therefore see a neural vestige of what the regular learners had been trained to make use of, while the control subjects had not.

One intriguing difference can be observed in the localization of the sensorimotor representation of the right hand: whereas a clear co-activation of the sensorimotor hand area (while listening) had developed in the regular learner group, such an effect was absent in the control subjects. In these subjects, no common representation of ear and hand was established. A consistent organization of

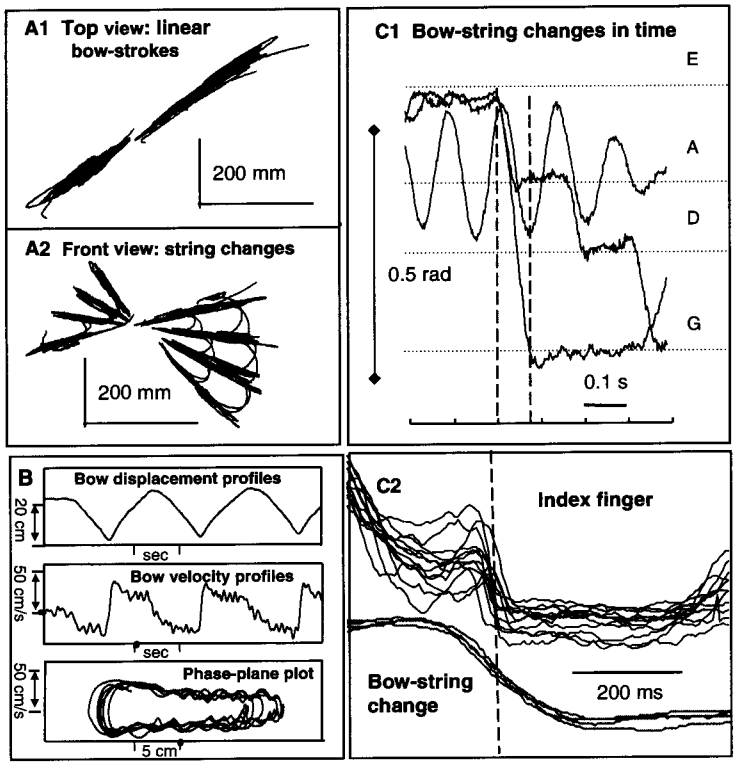



**Plate 1** (A1, A2) The two principal bow movements: rectilinear bow strokes and rotating movements during string changes. (B) Slow *détaché* bow strokes: displacement profiles (above), velocity profiles (middle), and phase-plane plots (below). (C1) Bow rotations for string changes (as in A2), but transformed in displacement profiles (1 radian ~57°); bow tilts over one, two, and three strings. The jump from the highest string (E) to the lowest string (G) amounts to about 0.5 radians or ~28° and required about 80 ms in this example. (C2) string change together with displacement profiles of the playing finger. Ten superimposed traces each.

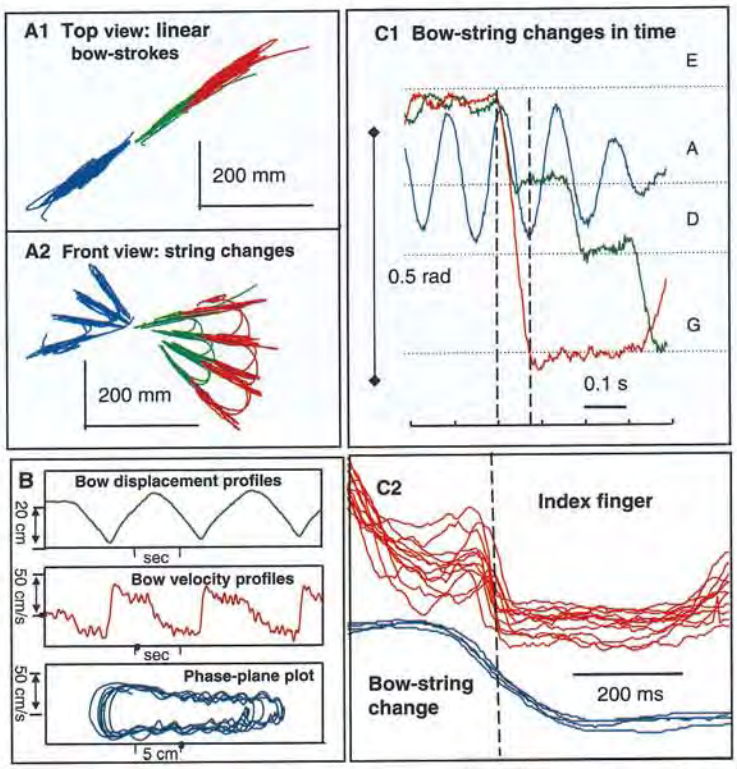



**Plate 1** (A1, A2) The two principal bow movements: rectilinear bow strokes and rotating movements during string changes. (B) Slow *détaché* bow strokes: displacement profiles (above), velocity profiles (middle), and phase-plane plots (below). (C1) Bow rotations for string changes (as in A2), but transformed in displacement profiles (1 radian ~57°); bow tilts over one, two, and three strings. The jump from the highest string (E) to the lowest string (G) amounts to about 0.5 radians or ~28° and required about 80 ms in this example. (C2) string change together with displacement profiles of the playing finger. Ten superimposed traces each.

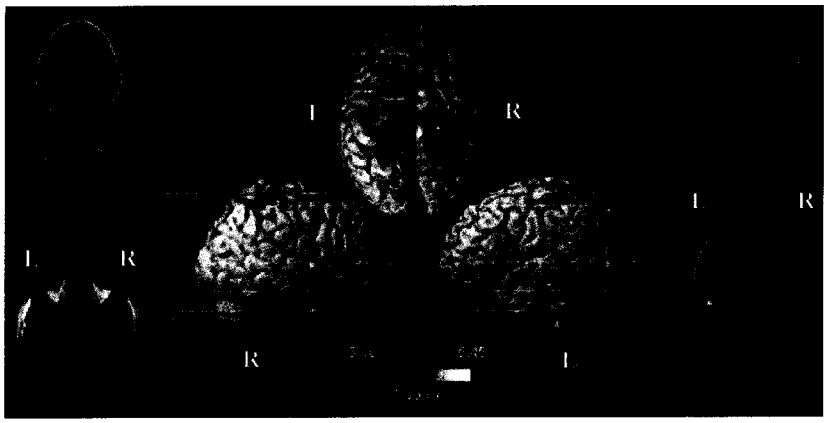



**Plate 2** A voxel-based morphometric study comparing grey matter differences between professional musicians, amateur musicians, and non-musicians. The musician status was modeled as a three-level gradation in which professional musicians were ranked highest, amateur musicians were ranked intermediate, and non-musicians were ranked lowest. Only those voxels with a significant positive correlation between musician status and increase in grey matter volume are shown, overlaid on the rendered cortex surface of a single subject.

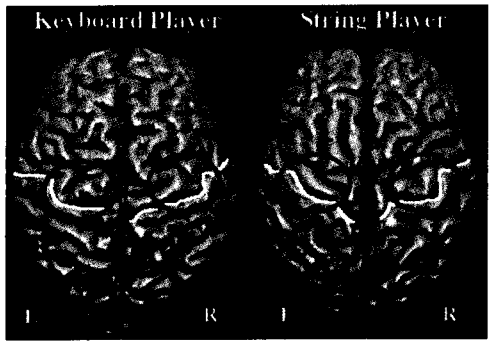



**Plate 3** Within-musician, instrument-typical, gross-anatomical differences in the configuration of the pre-central gyrus, the so-called Omega Sign (OS). The majority of the adult keyboard players had an elaborated Omega configuration in the left hemisphere, while most of the adult string players had a pronounced Omega in the right hemisphere.

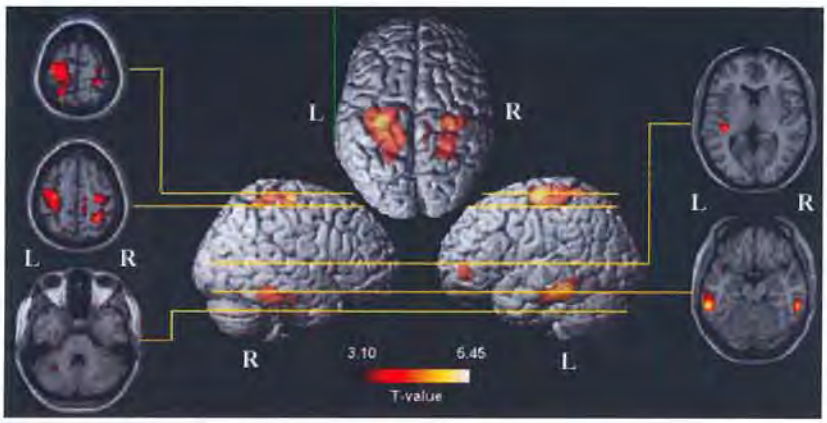



**Plate 2** A voxel-based morphometric study comparing grey matter differences between professional musicians, amateur musicians, and non-musicians. The musician status was modeled as a three-level gradation in which professional musicians were ranked highest, amateur musicians were ranked intermediate, and non-musicians were ranked lowest. Only those voxels with a significant positive correlation between musician status and increase in grey matter volume are shown, overlaid on the rendered cortex surface of a single subject.

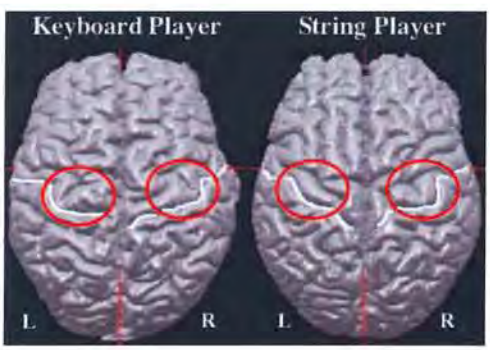



**Plate 3** Within-musician, instrument-typical, gross-anatomical differences in the configuration of the pre-central gyrus, the so-called Omega Sign (OS). The majority of the adult keyboard players had an elaborated Omega configuration in the left hemisphere, while most of the adult string players had a pronounced Omega in the right hemisphere.



**Plate 4** A voxel-based morphometric study examining grey matter volume changes over time in a single former musicians comparing high resolution T1-weighted MRI scans before, during, and after an intense period of practice period over several months. Someof the significant grey matter volume changes correspond to the group differences seen in the study shown in Plate 2.

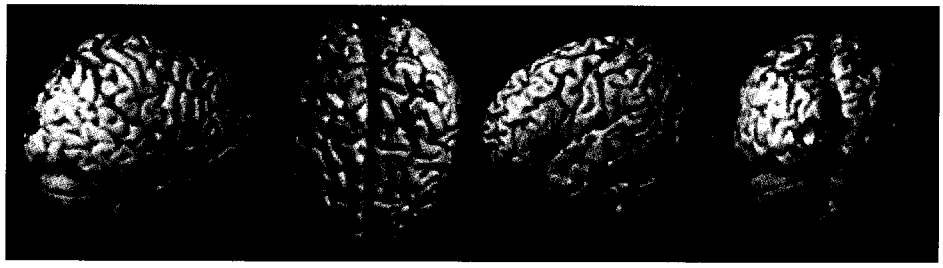


**Plate 5** A voxel-based morphometric study comparing grey matter differences between a group of 9–11 year old instrumentalists and age- as well as gender-matched non-musician controls. Results are overlaid on the rendered cortex surface of a single subject.

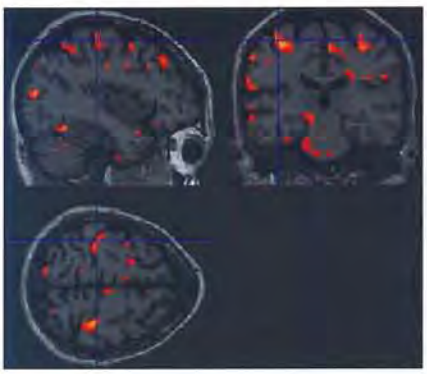


**Plate 4** A voxel-based morphometric study examining grey matter volume changes over time in a single former musicians comparing high resolution T1-weighted MRI scans before, during, and after an intense period of practice period over several months. Someof the significant grey matter volume changes correspond to the group differences seen in the study shown in Plate 2.



**Plate 5** A voxel-based morphometric study comparing grey matter differences between a group of 9–11 year old instrumentalists and age- as well as gender-matched non-musician controls. Results are overlaid on the rendered cortex surface of a single subject.

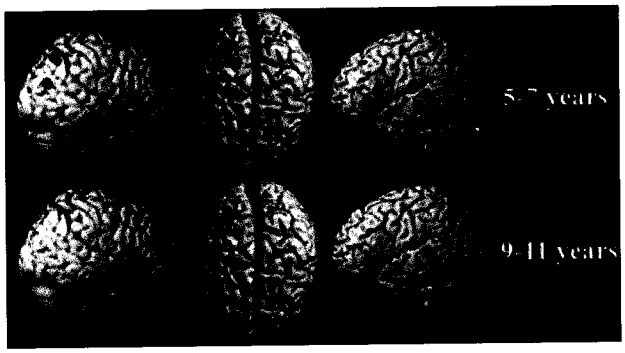



**Plate 6** Correlating behavioral performance and gray matter volume differences. The top row shows significant correlations between performance in the Primary Measures of Music Audiation (PMMA) test and voxel-by-voxel grey matter density values in a group of 5–7 year old, musically-naïve children. The pattern shows a bi-hemispheric, temporal lobe dominance. The bottom row shows significant correlations between performance in the Intermediate Measures of Music Audiation (IMMA) test and voxel-by-voxel grey matter density values in a group of 9–11 year old instrumentalists with 3–4 years of instrumental training. Now the pattern shows an anterior temporal and inferior frontal lobe dominance. Statistical results are overlaid on the rendered cortex surfaces of a single subject.




**Plate 7** Statistical parametric images superimposed onto a surface rendering of a standardized anatomical brain depicting significant activations during a rhythmic discrimination tasks in a group of 5–7 year-old, musically-naïve children, adult non-musicians, and adult musicians.

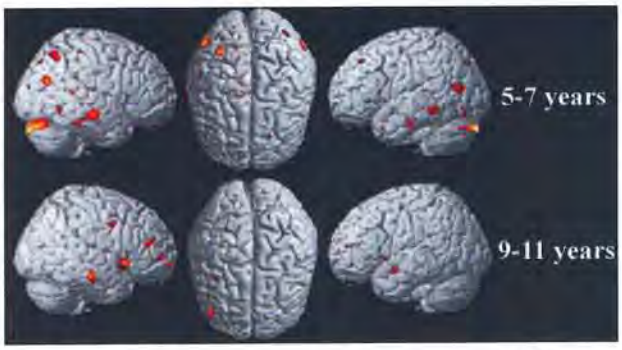



**Plate 6** Correlating behavioral performance and gray matter volume differences. The top row shows significant correlations between performance in the Primary Measures of Music Audiation (PMMA) test and voxel-by-voxel grey matter density values in a group of 5–7 year old, musically-naïve children. The pattern shows a bi-hemispheric, temporal lobe dominance. The bottom row shows significant correlations between performance in the Intermediate Measures of Music Audiation (IMMA) test and voxel-by-voxel grey matter density values in a group of 9–11 year old instrumentalists with 3–4 years of instrumental training. Now the pattern shows an anterior temporal and inferior frontal lobe dominance. Statistical results are overlaid on the rendered cortex surfaces of a single subject.




**Plate 7** Statistical parametric images superimposed onto a surface rendering of a standardized anatomical brain depicting significant activations during a rhythmic discrimination tasks in a group of 5–7 year-old, musically-naïve children, adult non-musicians, and adult musicians.

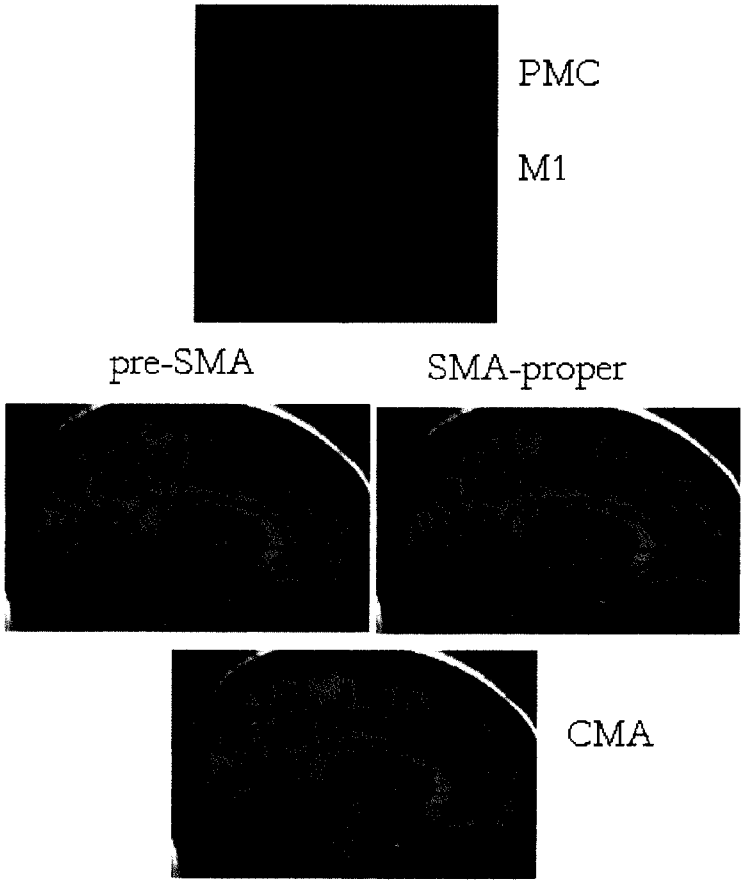



**Plate 8** Brain areas in which musicians typically show decreased haemodynamic responses during different unimanual or bimanual movements.

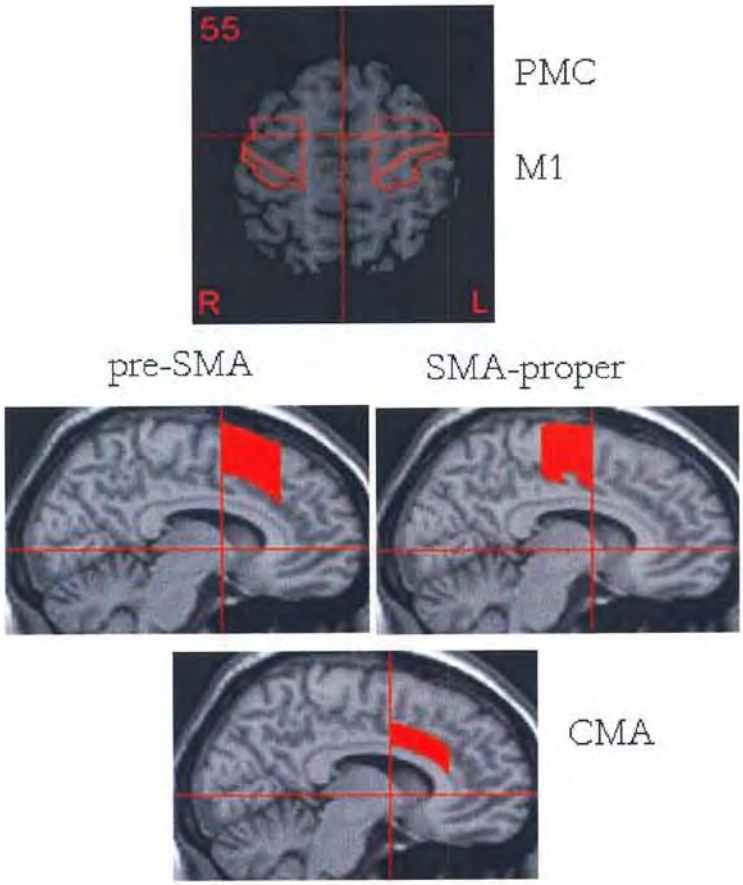



**Plate 8** Brain areas in which musicians typically show decreased haemodynamic responses during different unimanual or bimanual movements.

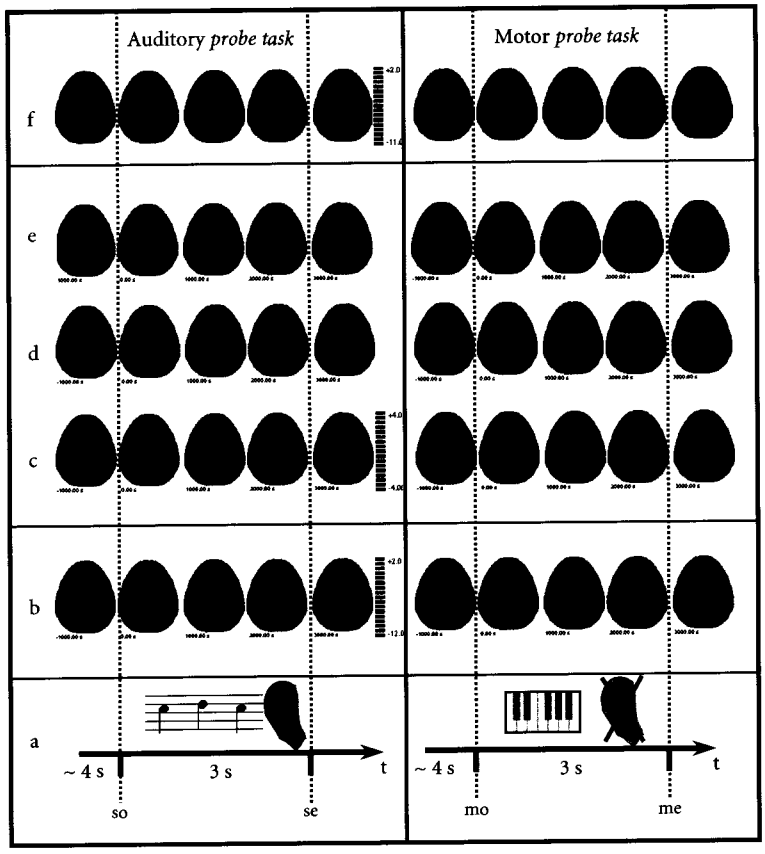



**Plate 9** Changes in cortical DC potentials induced by training. The displayed activation patterns are 'electrical top views' onto the unfolded surface of the head (nose pointing upward, the lateral borders covering subtemporal electrodes at the axial level of the eyes and ears). In the difference patterns (c–e), blue = activation increase, red = activation decrease. The left half of the figure relates to the passive auditory task, the right half relates to the silent motor task. so = stimulus onset, se = stimulus end, mo = movement onset, me = movement end. (a) Experimental paradigm. (b) Initial DC-EEG prior to first practice. (c) Activation changes (additional negative potential compared with the baseline (a)) after the first 20-min practice. (d) Activation changes after about 21 days of practice (five sessions). (e) Activation changes after about 39 days of practice (10 sessions). (f) DC-EEG of seven professional pianists.

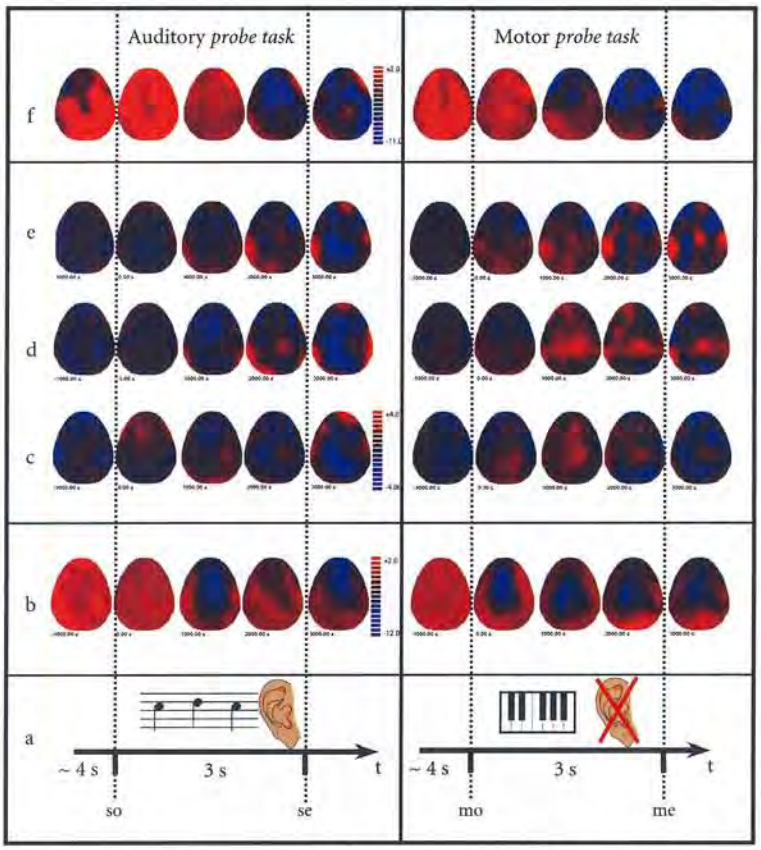



**Plate 9** Changes in cortical DC potentials induced by training. The displayed activation patterns are 'electrical top views' onto the unfolded surface of the head (nose pointing upward, the lateral borders covering subtemporal electrodes at the axial level of the eyes and ears). In the difference patterns (c–e), blue = activation increase, red = activation decrease. The left half of the figure relates to the passive auditory task, the right half relates to the silent motor task. so = stimulus onset, se = stimulus end, mo = movement onset, me = movement end. (a) Experimental paradigm. (b) Initial DC-EEG prior to first practice. (c) Activation changes (additional negative potential compared with the baseline (a)) after the first 20-min practice. (d) Activation changes after about 21 days of practice (five sessions). (e) Activation changes after about 39 days of practice (10 sessions). (f) DC-EEG of seven professional pianists.

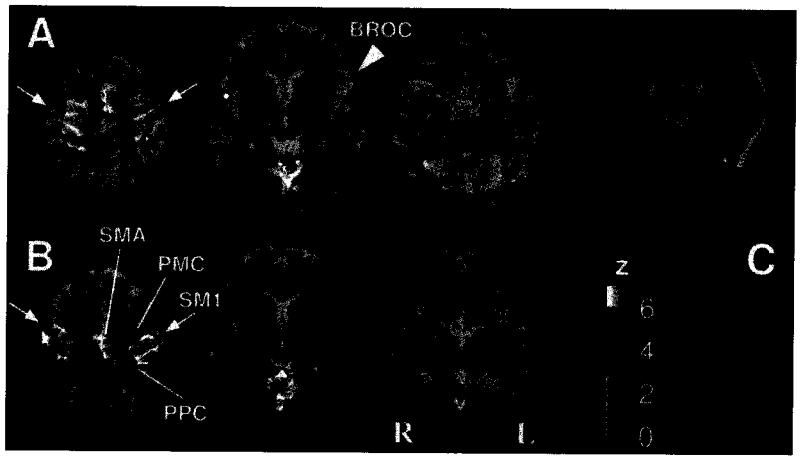



**Plate 10** Functional MRI activations: (A) violin playing on a sham instrument (without auditory feedback), (B) bimanually coordinated non-musical reference task (repetitive pulling of a drawer with the left hand, synchronized grip of an object within the drawer with the right hand), (C) whole brain volume and slice angulation. Baseline was the resting state for both tasks. For each task, three oblique slices (slightly tilted with respect to the transverse AC–PC orientation) at the level of the motor cortex, at the level of the frontal operculum, and at the level of the cerebellum, are shown. (A) Prominent bilateral activation is seen in the sensorimotor cortex (SM1; arrows = central sulcus), pre-motor cortex (PMC), supplementary motor area (SMA), posterior parietal cortex (PPC), thalamus and cerebellum. A slight asymmetry with higher activation of the right motor cortex (i.e. contralateral to the fingering hand) can be seen, whereas there is no asymmetry with the bimanual reference task. The main difference in activation between the two tasks was the prominent activation in Broca's area in the violin task (arrowhead), as compared with its lack (in the present slice) or only marginal activation of this region in the reference task. The violin task also produced a weak activation of the right-sided homologue of Broca's area. This asymmetry is opposite to that in the sensorimotor areas and thus most likely related to hemispheric dominance.

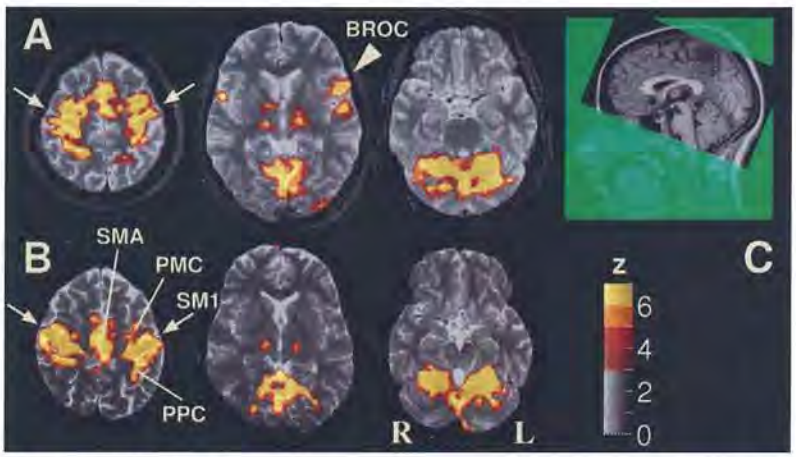



**Plate 10** Functional MRI activations: (A) violin playing on a sham instrument (without auditory feedback), (B) bimanually coordinated non-musical reference task (repetitive pulling of a drawer with the left hand, synchronized grip of an object within the drawer with the right hand), (C) whole brain volume and slice angulation. Baseline was the resting state for both tasks. For each task, three oblique slices (slightly tilted with respect to the transverse AC–PC orientation) at the level of the motor cortex, at the level of the frontal operculum, and at the level of the cerebellum, are shown. (A) Prominent bilateral activation is seen in the sensorimotor cortex (SM1; arrows = central sulcus), pre-motor cortex (PMC), supplementary motor area (SMA), posterior parietal cortex (PPC), thalamus and cerebellum. A slight asymmetry with higher activation of the right motor cortex (i.e. contralateral to the fingering hand) can be seen, whereas there is no asymmetry with the bimanual reference task. The main difference in activation between the two tasks was the prominent activation in Broca's area in the violin task (arrowhead), as compared with its lack (in the present slice) or only marginal activation of this region in the reference task. The violin task also produced a weak activation of the right-sided homologue of Broca's area. This asymmetry is opposite to that in the sensorimotor areas and thus most likely related to hemispheric dominance.

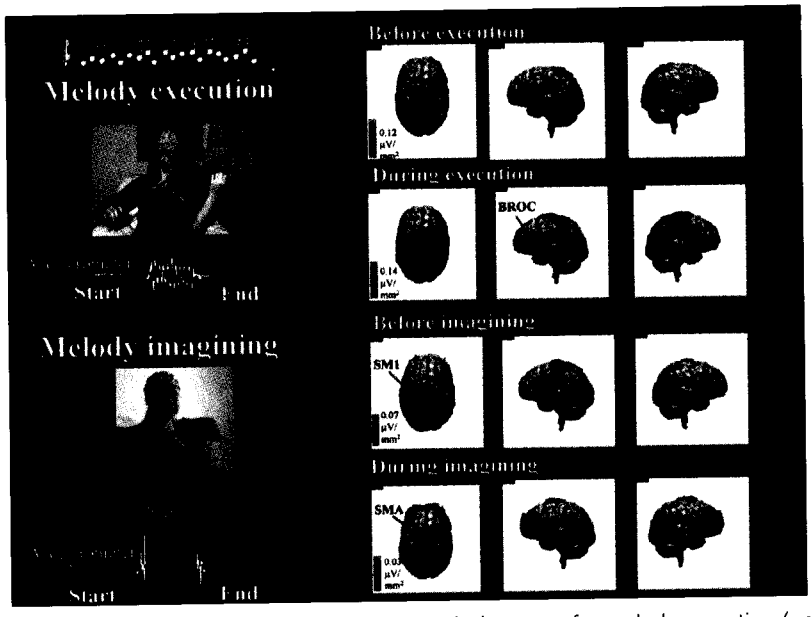



**Plate 11** Experimental design and reconstructed cortical sources for melody execution (upper panel, on the right) and imagining (lower panel, on the right) for one of the subjects (S7). The melody is shown at the top of the upper panel (left). High-resolution EEG recorded with an electrode cap from 58 scalp positions during execution of a musical sequence on a violin and during imagining the same sequence (left part of the figure). The start and the end of the musical sequence are defined by recording the accelerogram. The start and the end of imagining the musical sequence are marked by short abrupt movements of the right hand also recorded by accelerogram. The reconstructed CCD maps are superimposed on the segmented brain of the subject (top, left and right view of the brain). The CCD maps, which are displayed with the individual current vectors, show the current flow distribution on the cortex, which can account for the potentials measured on the scalp. Reconstructed cortical sources for melody execution: the CCD maps before the melody execution suggest distributed sources in SM1s, in the SMA, and in the inferior frontal gyrus. During execution activation of networks in the left and right SM1s, in the SMA, in the frontal inferior gyrus, and in the superior temporal gyrus is observed. Reconstructed cortical sources for imagining: before imagining there is activation of SM1s, the SMA, the pars opercularis of the inferior frontal gyrus, and the gyrus occipitalis. During imagining there is activation of the SMA, the gyrus supramarginalis, and the frontal polar areas but no activation of the SM1s and the gyrus temporalis superior. (Modified from Kristeva *et al.* (2003) with the kind permission of *NeuroImage*.)

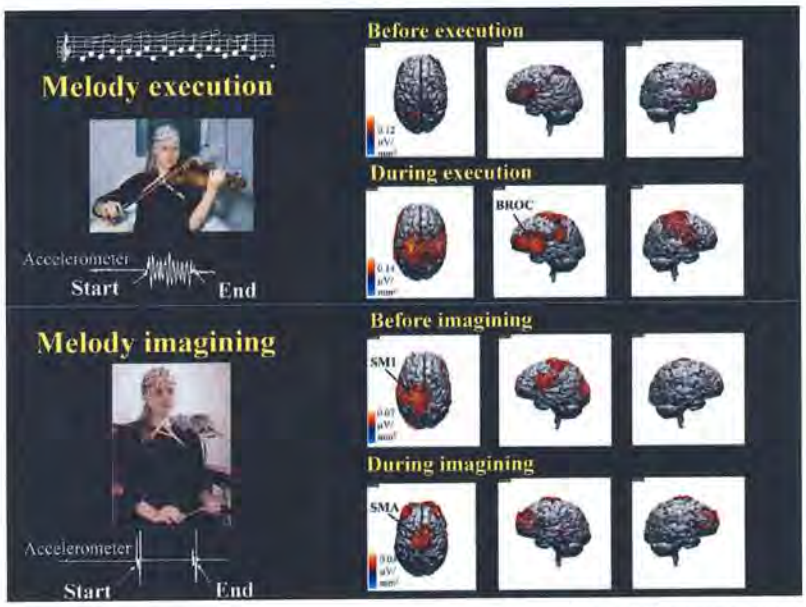



**Plate 11** Experimental design and reconstructed cortical sources for melody execution (upper panel, on the right) and imagining (lower panel, on the right) for one of the subjects (S7). The melody is shown at the top of the upper panel (left). High-resolution EEG recorded with an electrode cap from 58 scalp positions during execution of a musical sequence on a violin and during imagining the same sequence (left part of the figure). The start and the end of the musical sequence are defined by recording the accelerogram. The start and the end of imagining the musical sequence are marked by short abrupt movements of the right hand also recorded by accelerogram. The reconstructed CCD maps are superimposed on the segmented brain of the subject (top, left and right view of the brain). The CCD maps, which are displayed with the individual current vectors, show the current flow distribution on the cortex, which can account for the potentials measured on the scalp. Reconstructed cortical sources for melody execution: the CCD maps before the melody execution suggest distributed sources in SM1s, in the SMA, and in the inferior frontal gyrus. During execution activation of networks in the left and right SM1s, in the SMA, in the frontal inferior gyrus, and in the superior temporal gyrus is observed. Reconstructed cortical sources for imagining: before imagining there is activation of SM1s, the SMA, the pars opercularis of the inferior frontal gyrus, and the gyrus occipitalis. During imagining there is activation of the SMA, the gyrus supramarginalis, and the frontal polar areas but no activation of the SM1s and the gyrus temporalis superior. (Modified from Kristeva *et al.* (2003) with the kind permission of *NeuroImage*.)

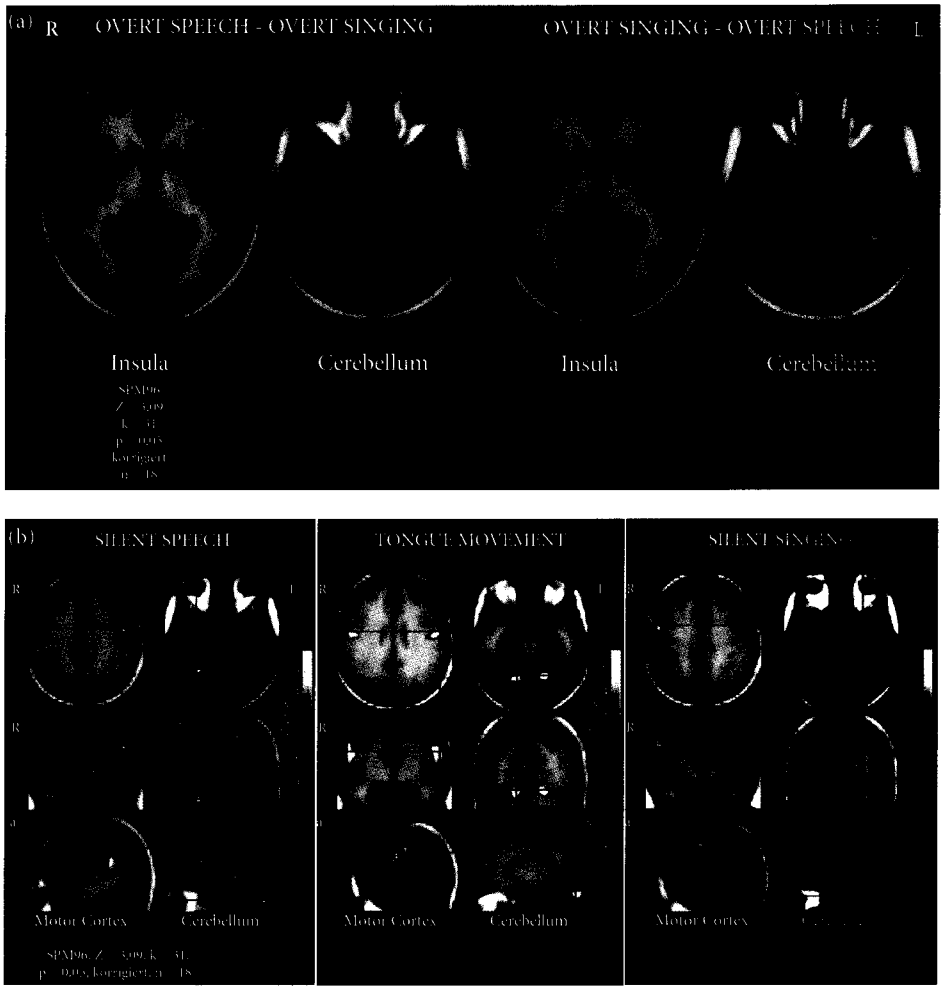



**Plate 12** (a) Haemodynamic activation patterns at the level of the anterior insula during recitation of a highly overlearned word string (speaking) and reproduction of a non-lyrical tune (singing). Cognitive subtraction approach: overt versus covert speaking, overt versus covert singing. (b) Left panel = 'inner speech' (silent recitation of a highly overlearned word string); mid panel = non-verbal vertical tongue movements; right panel = 'inner singing' (silent reproduction of a non-lyrical tune). fMRI activation maps superimposed on the averaged $T_1$-weighted anatomical images (Talairach space) across subjects at the level of the supplementary motor area (SMA), motor cortex, temporoparietal junction, and cerebellum (upper row = transverse, middle row = coronal, lower row = sagittal slices). For (a) and (b) $Z > 3.09$, $k > 31$, $P < 0.05$ (corrected), $n = 18$.

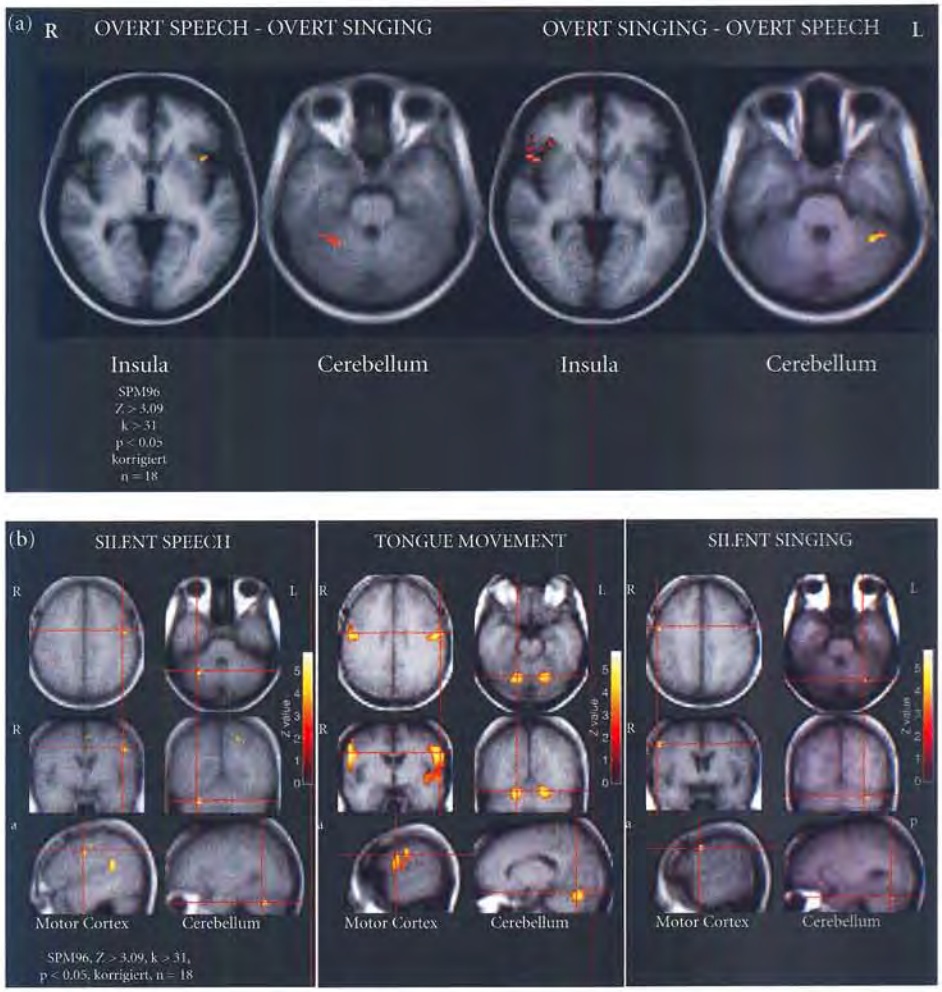



**Plate 12** (a) Haemodynamic activation patterns at the level of the anterior insula during recitation of a highly overlearned word string (speaking) and reproduction of a non-lyrical tune (singing). Cognitive subtraction approach: overt versus covert speaking, overt versus covert singing. (b) Left panel = 'inner speech' (silent recitation of a highly overlearned word string); mid panel = non-verbal vertical tongue movements; right panel = 'inner singing' (silent reproduction of a non-lyrical tune). fMRI activation maps superimposed on the averaged $T_1$-weighted anatomical images (Talairach space) across subjects at the level of the supplementary motor area (SMA), motor cortex, temporoparietal junction, and cerebellum (upper row = transverse, middle row = coronal, lower row = sagittal slices). For (a) and (b) $Z > 3.09$, $k > 31$, $P < 0.05$ (corrected), $n = 18$.

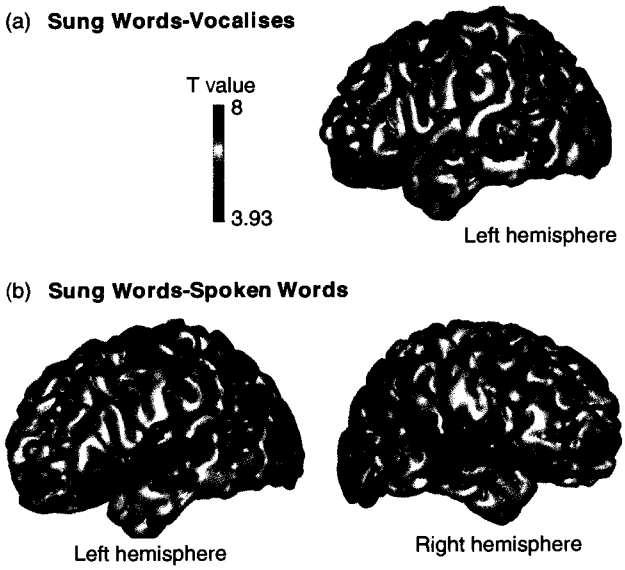



**Plate 13** (a) Brain areas more activated in the sung words than vocalises conditions. (Statistical map of significant activations superimposed on 3D reconstruction of the brain.) Preliminary results revealed clusters of voxels in the left hemisphere only, mainly in the left middle temporal gyrus (MTG) and inferior temporal gyrus (ITG). (b) Brain areas more activated in the sung words than spoken words conditions. Results revealed clusters of voxels in the anterior and posterior portions of the superior temporal gyrus (STG), MTG, transverse temporal gyrus (TTG) and insula bilaterally, but to a greater extent in the right hemisphere. Important clusters of voxels were also activated in the right pre-central gyrus and right middle frontal operculum, in addition to smaller clusters in the left inferior frontal gyrus.

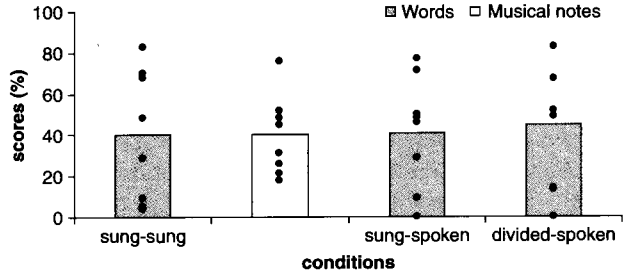



**Plate 14** Mean scores obtained by aphasic patients while repeating words and musical notes of unfamiliar songs in three conditions. Each dot represents the scores of one patient.

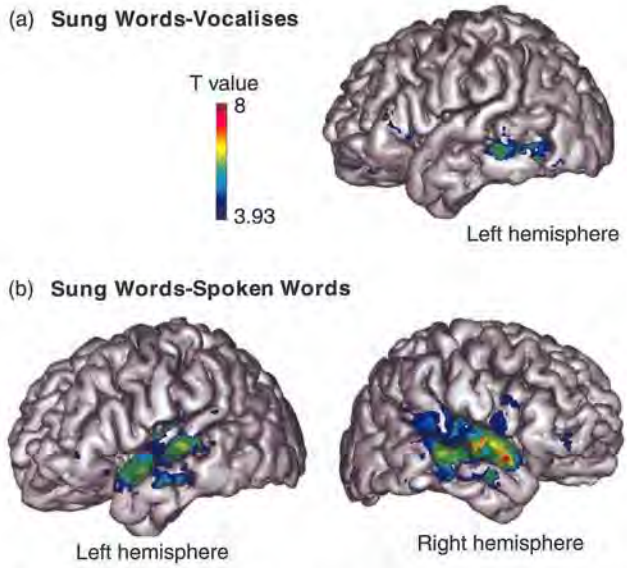



**Plate 13** (a) Brain areas more activated in the sung words than vocalises conditions. (Statistical map of significant activations superimposed on 3D reconstruction of the brain.) Preliminary results revealed clusters of voxels in the left hemisphere only, mainly in the left middle temporal gyrus (MTG) and inferior temporal gyrus (ITG). (b) Brain areas more activated in the sung words than spoken words conditions. Results revealed clusters of voxels in the anterior and posterior portions of the superior temporal gyrus (STG), MTG, transverse temporal gyrus (TTG) and insula bilaterally, but to a greater extent in the right hemisphere. Important clusters of voxels were also activated in the right pre-central gyrus and right middle frontal operculum, in addition to smaller clusters in the left inferior frontal gyrus.

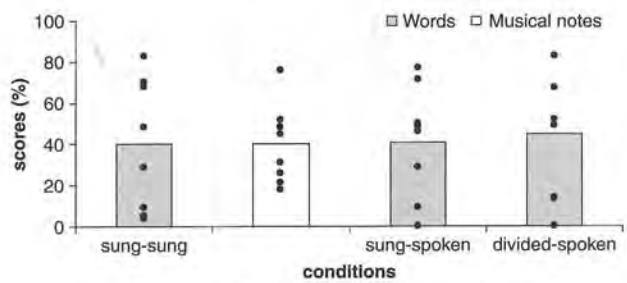



**Plate 14** Mean scores obtained by aphasic patients while repeating words and musical notes of unfamiliar songs in three conditions. Each dot represents the scores of one patient.

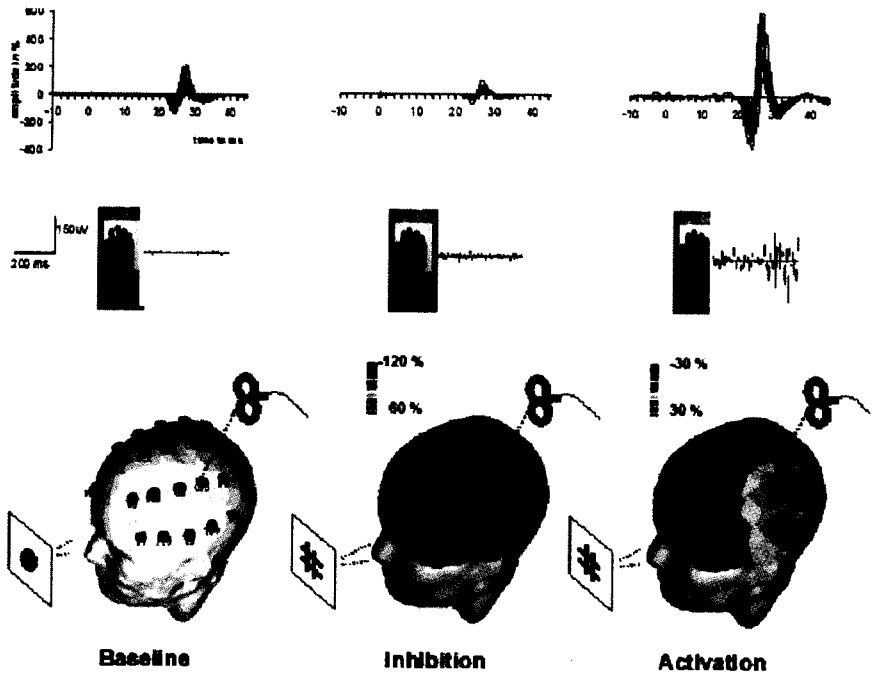



**Plate 15** TMS and EEG data of a representative subject for all relevant conditions. Top: Motor evoked potentials (MEP) induced by transcranial magnetic stimulation (TMS) over the primary motor cortex [over the optimal point for stimulation of the first digital interosseus (FDI) muscle] during the baseline (left), inhibition (middle), and activation (right) conditions. Middle: Subject's hand placed on the custom-made keyboard. The EMG activity of the FDI is shown during the performance of the baseline (left), inhibition (middle), and activation conditions (right). Bottom left: Three-dimensional digitized individual head shape with realistic (co-registered) placement of the surface electrodes. Bottom middle: Inhibition condition. Bottom right: Activation condition. Topographic spectral power maps are in the frequency range of 11–13 Hz. Maps are projected to the individual three-dimensional digitized head shape of the subject. Task-related increases of oscillatory activity ($\alpha$-TRPI) are given in blue. Decreases of oscillatory activity ($\alpha$-TRPD) are given in red. In the inhibition condition (middle), i.e. when visual cues were presented on the screen but no movement was required, there was a topographically distinct task-related increase of oscillatory activity ($\alpha$-TRPI) bilaterally over the sensorimotor areas. During the activation conditions, there was a task-related decrease of oscillatory activity ($\alpha$-TRPD) bilaterally over the sensorimotor areas. Note the topographic similarity of $\alpha$-TRPI and $\alpha$-TRPD.

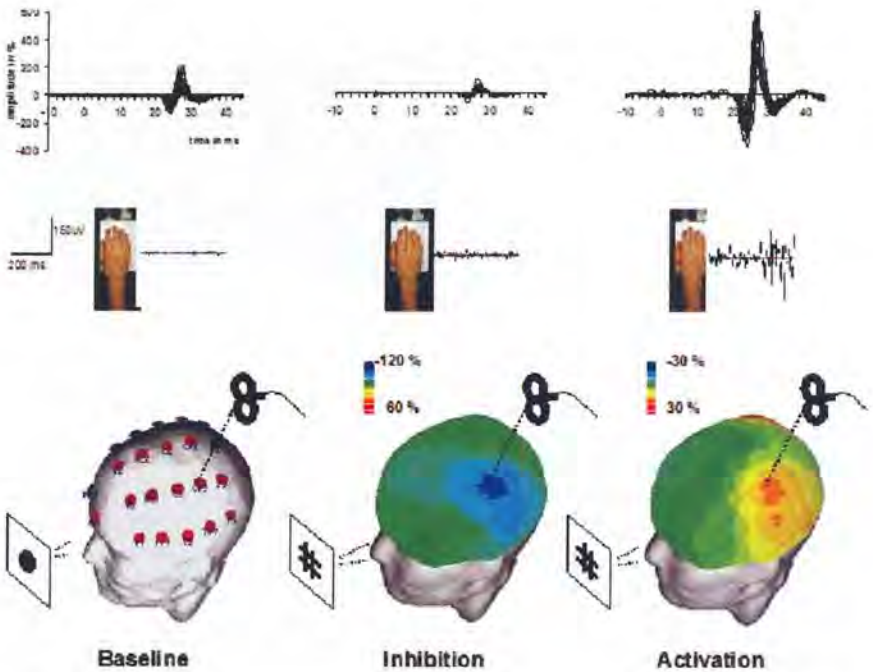



**Plate 15** TMS and EEG data of a representative subject for all relevant conditions. Top: Motor evoked potentials (MEP) induced by transcranial magnetic stimulation (TMS) over the primary motor cortex [over the optimal point for stimulation of the first digital interosseus (FDI) muscle] during the baseline (left), inhibition (middle), and activation (right) conditions. Middle: Subject's hand placed on the custom-made keyboard. The EMG activity of the FDI is shown during the performance of the baseline (left), inhibition (middle), and activation conditions (right). Bottom left: Three-dimensional digitized individual head shape with realistic (co-registered) placement of the surface electrodes. Bottom middle: Inhibition condition. Bottom right: Activation condition. Topographic spectral power maps are in the frequency range of 11–13 Hz. Maps are projected to the individual three-dimensional digitized head shape of the subject. Task-related increases of oscillatory activity ($\alpha$-TRPI) are given in blue. Decreases of oscillatory activity ($\alpha$-TRPD) are given in red. In the inhibition condition (middle), i.e. when visual cues were presented on the screen but no movement was required, there was a topographically distinct task-related increase of oscillatory activity ($\alpha$-TRPI) bilaterally over the sensorimotor areas. During the activation conditions, there was a task-related decrease of oscillatory activity ($\alpha$-TRPD) bilaterally over the sensorimotor areas. Note the topographic similarity of $\alpha$-TRPI and $\alpha$-TRPD.

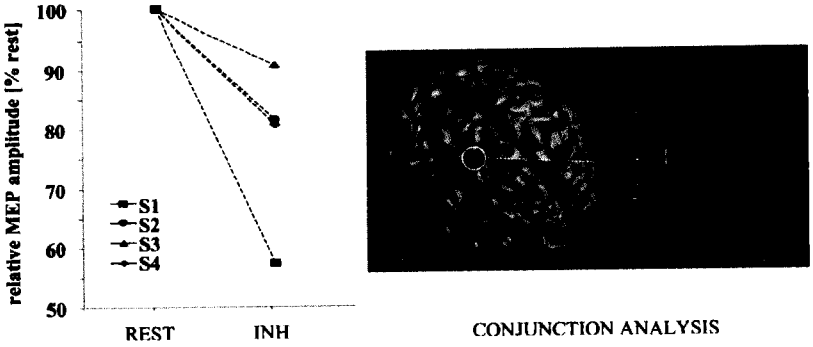



**Plate 16** Functional MRI and TMS data. Right: fMRI conjunction analysis. Red/orange colour coding identifies areas with significant decrease of the metabolic BOLD response during motor inhibition compared with baseline and significant increase during retrieval of the learned movement sequences. The curves on the right represent the temporal evolution of the BOLD response during activation (retrieval) (green) and inhibition (blue). Left: Corresponding TMS data in four of the six subjects studied with fMRI. The MEPs during motor inhibition (INH) are given as a percentage of the MEPs at baseline (REST). Note the consistent suppression of the motor evoked TMS responses during INH, and the accompanying negative BOLD response in this condition (blue curve).

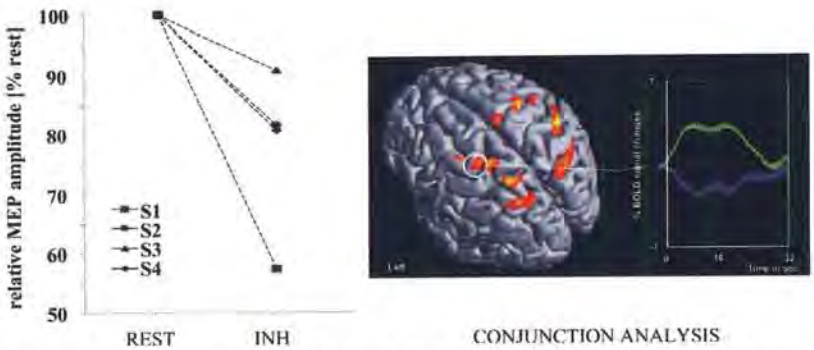



**Plate 16** Functional MRI and TMS data. Right: fMRI conjunction analysis. Red/orange colour coding identifies areas with significant decrease of the metabolic BOLD response during motor inhibition compared with baseline and significant increase during retrieval of the learned movement sequences. The curves on the right represent the temporal evolution of the BOLD response during activation (retrieval) (green) and inhibition (blue). Left: Corresponding TMS data in four of the six subjects studied with fMRI. The MEPs during motor inhibition (INH) are given as a percentage of the MEPs at baseline (REST). Note the consistent suppression of the motor evoked TMS responses during INH, and the accompanying negative BOLD response in this condition (blue curve).

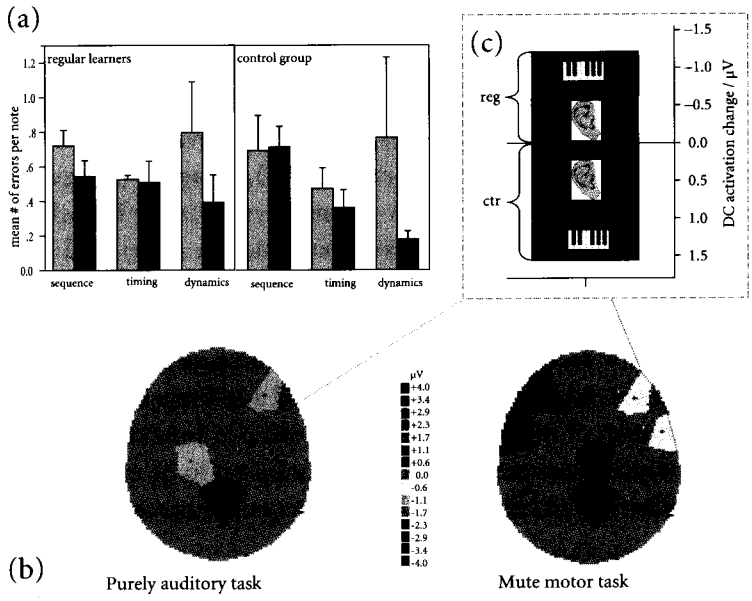



**Fig. 11.2** (a) Average error incidences per note of a given melody, split up into sequence, timing and dynamics errors. Regular learner group on the left, control group on the right. Light grey bars, first training session; dark gray bars, last session. (b) Changes of cortical activity induced by 10 sessions of training. Only significant differences between the two groups are shown: red = more negative potential (activation increase), blue = more positive potential (activation decrease), green = no change. (c) Potential changes for a right frontotemporal selection of electrode sites. reg = regular learner group, ctr = control group: green bars, auditory task; red bars, silent movement condition.

perceptual and motor events seems to be a prerequisite for the development of a common sensorimotor representation: It is likely, then, that the sensorimotor map has 'to make sense in order to make synapses'.

Another group difference appeared in the rFTA region (Fig. 11.2b). This was basically due to two mutual effects (Fig. 11.2c). First, in the regular learners, the activity of the area concerned had increased. Second, in the control group, this activity even decreased during training. The subjects had acquired high timing accuracy for the reproduction of fast complex rhythmic sequences and a 'feeling' for the piano keyboard without putting the rFTA into action at all. Obviously, the actual function of the rFTA is to process the sequential order of *pitch patterns* feeding them into working memory, and also to retrieve them from working memory for execution. That means that the rFTA is likely to have properties of a supramodal neural network suitable for translating music into movement, i.e. the rFTA may provide an interface map.

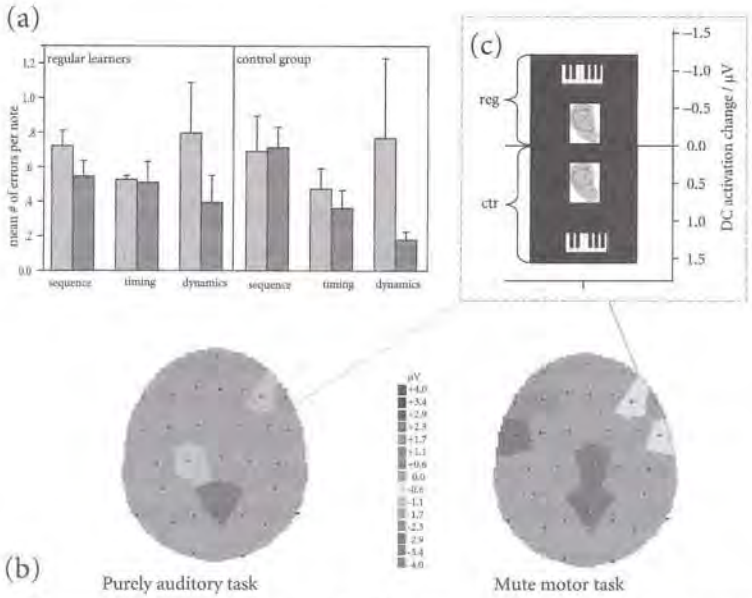



**Fig. 11.2** (a) Average error incidences per note of a given melody, split up into sequence, timing and dynamics errors. Regular learner group on the left, control group on the right. Light grey bars, first training session; dark gray bars, last session. (b) Changes of cortical activity induced by 10 sessions of training. Only significant differences between the two groups are shown: red = more negative potential (activation increase), blue = more positive potential (activation decrease), green = no change. (c) Potential changes for a right frontotemporal selection of electrode sites. reg = regular learner group, ctr = control group: green bars, auditory task; red bars, silent movement condition.

perceptual and motor events seems to be a prerequisite for the development of a common sensorimotor representation: It is likely, then, that the sensorimotor map has 'to make sense in order to make synapses'.

Another group difference appeared in the rFTA region (Fig. 11.2b). This was basically due to two mutual effects (Fig. 11.2c). First, in the regular learners, the activity of the area concerned had increased. Second, in the control group, this activity even decreased during training. The subjects had acquired high timing accuracy for the reproduction of fast complex rhythmic sequences and a 'feeling' for the piano keyboard without putting the rFTA into action at all. Obviously, the actual function of the rFTA is to process the sequential order of *pitch patterns* feeding them into working memory, and also to retrieve them from working memory for execution. That means that the rFTA is likely to have properties of a supramodal neural network suitable for translating music into movement, i.e. the rFTA may provide an interface map.

## A shared core network?

Recent research on musicians and music processing and performance has repeatedly revealed the following to be involved in musical tasks: motor preparation and execution (Parsons 2001; Hickok *et al.* 2003; Kristeva *et al.* 2003; Lotze *et al.* 2003; Haslinger *et al.* 2004; Meister *et al.* 2004), imagery (Lotze *et al.* 2003; Meister *et al.* 2004), music perception (Ohnishi *et al.* 2001; Kölsch *et al.* 2002; Gaab *et al.* 2003; Hickok *et al.* 2003; Levitin and Menon 2003), and instrumental mirror systems (Hasegawa *et al.* 2004; Haslinger *et al.* 2005).

We have deliberately pooled the aforementioned very different studies covering a whole variety of facets of music perception and production. From the fMRI and PET data these studies provided, we collected and meta-analysed all available Talairach coordinates of the significant task-specific or musician-specific images, and evaluated which areas are most commonly found to be active in this variety of tasks. The following had by far the highest frequency across the studies:

1. In the left hemisphere:
   - frontal operculum (Broca's area/BA 44/46/47) (Kölsch *et al.* 2002; Kristeva *et al.* 2003 (EEG-no coordinates available); Levitin and Menon 2003; Lotze *et al.* 2003; Hasegawa *et al.* 2004; Haslinger *et al.* 2004; Bangert *et al.* 2005);
   - frontal BA 9 (Ohnishi *et al.* 2001; Kölsch *et al.* 2002; Gaab *et al.* 2003; Hasegawa *et al.* 2004; Meister *et al.* 2004; Bangert *et al.* 2005);
   - superior temporal gyrus (BA 22/Wernicke's area) (Parsons 2001; Hasegawa *et al.* 2004; Bangert *et al.* 2005);
   - supramarginal gyrus (BA 40) (Meister *et al.* 2004; Hasegawa *et al.* 2004; Bangert *et al.* 2005).
2. In the right hemisphere:
   - frontal BA 9 (Kölsch *et al.* 2002 (non-musicians); Lotze *et al.* 2003);
   - frontal operculum (Broca's area/BA 44/BA 46) (Kölsch *et al.* 2002; Hasegawa *et al.* 2004);
3. Bilaterally:
   - (SMA/PMA/BA 6) (Gaab *et al.* 2003; Meister *et al.* 2004; Hasegawa *et al.* 2004; Bangert *et al.* 2005);
   - supramarginal gyrus (BA 40) (Meister *et al.* 2004; Hasegawa *et al.* 2004; Bangert *et al.* 2005).

Regardless of whether a music study paradigm involves perception or execution, the ensemble of four main components, resembling Brodmann's areas 9, 44/46, 20–22, and 40, appear to be crucial contributors to a musician's skills—possibly a 'core network' of musicianship. It has to be added, though, that

this network could be related to auditory–sensorimotor co-processing only. The important visuomotor and visuo-auditory aspects of music processing (music notation and sight-reading) seem to employ a substantially different network, with, surprisingly, almost no overlap with any of the audiomotor-focused studies listed above (Stewart *et al.* 2003).

## Musical speech or musical language?

It is interesting to speculate why Broca's and Wernicke's areas, though traditionally language-related, should be part of the cortical network dedicated to musical expression in skilled instrumentalists. The concept of language has recently become more and more generalized and is no longer confined to auditory–vocal verbal language. Direct comparison of music and language processing has revealed considerable parallels (Hickok *et al.* 2003) and the small literature review above indicates a substantial role for these areas in music, just as in language. The importance of Broca's area for musicians has recently been shown not only functionally but also structurally by Sluming *et al.* (2002), who used voxel-based morphometry to demonstrate an increased density of grey matter in Broca's area in musicians.

It may be an important distinction to employ the term 'speech' rather than (or additionally to) 'language' when thinking about the function of the Broca–Wernicke network. 'Language' usually rather implies its structure and logical framework, like syntax and semantics, while 'speech' includes the actual process of articulatory execution, and is thereby action-related. The 'language' framework often confines possible interpretations to the peculiarities of logical structure of the experimental paradigm and stimulus material (e.g. musical 'syntax'), although the actual function of Broca's area is much more general, in terms of 'speech' and not only language. Following up on the 'Bach speaks' catchphrase coined by Kölsch *et al.* (2002), we would rephrase this as 'Bach may know the language, but it is rather Gould who does the speaking'.

A bottom-up approach towards the clarification of the functional relevance of the inferior frontal cortex comes from the study of mirror neurons in monkeys. By definition, 'mirror neurons' discharge when a monkey makes a specific action *and* when it observes another individual making a similar action (Rizzolatti and Fadiga 1998). In monkeys, mirror neurons are found in the rostral ventral pre-motor cortex (area F5), originally for the visual domain. However, Kohler *et al.* (2002) recently reported that a population of neurons called audiovisual mirror neurons additionally responds when only the *sound* of the effective action is presented to the monkey. Multimodal neurons have been described in several cortical areas and subcortical centres. These neurons, however, responded to specific stimulus locations or directions of movement. The difference from the neurons described by Keysers *et al.* (2003) is that they do not code space, or some

spatial characteristics of stimuli, but actions when they are only heard. These 'audiovisual mirror neurons' therefore represent actions independently of whether these actions are performed, heard, or seen. Another remarkable property of audiovisual mirror neurons is that about half of them respond with a similar intensity of discharge whether the action is only heard, only seen, or both heard and seen. This finding is important, as it suggests that the neurons code the action in an abstract way, which does not depend on the source of information (auditory or visual) from which the evidence about the presence of the action is taken (Keysers *et al.* 2003). It is not only visual and auditory–visual mirror neurons that have been reported, but also an exclusively auditory responsive domain with high selectivity of these neurons for acoustic stimuli (Romanski and Goldman-Rakic 2002).

Additional support for this physiological finding comes from cytoarchitectural studies (Petrides and Pandya 1999, 2002; Romanski *et al.* 1999). Although the areas in question have a complex multimodal input, they could be differentiated in terms of some of their inputs. In fact, the 'mirror' functionality of area F5 has led researchers to consider this region to be the precursor of Broca's area in the monkey (Rizzolatti and Arbib 1998). Three studies in musicians point towards a musical mirror system in musicians, one using auditory action–observation (Ohnishi *et al.* 2001) and the other two using visual action–observation (Hasegawa *et al.* 2004; Haslinger *et al.* 2005). What makes the hypothesis of an auditory mirror system particularly interesting is that the classic notion of Broca's area being dedicated to language and speech processing has been expanded to include a broad range of stimuli including music.

## Conclusions

The cortical structures providing a musician's audiomotor performing skills are always activated together, even when one of the physical aspects of performing (instrumental sound, somatosensory input, movement output) is missing. This substrate of the 'inner ear' ranks among the very first implicit acquisitions during the process of learning to play an instrument. The years of practice provide just the ability to be aware/make conscious use of these neuronal circuits that have already been established. The same structures are likely to be directly involved in musical performance imagery, hence contributing elements of both perception and action to a musician's mental representation of music.

## References

**Aboitiz, F. and Garcia, R.** (1997). The anatomy of language revisited. *Biol Res*, **30**(4), 171–183.

**Bangert, M. and Altenmüller, E.O.** (2003). Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neurosci*, **4**, 26.

**Bangert, M., Parlitz, D., and Altenmüller, E.O.** (1998). Audio-motor integration in beginner and expert pianists: acombined DC-EEG and behavioural study. In: N. Elsner and R. Wehner (eds.) *The Göttingen neurobiology report II*, p. 753. Thieme, Stuttgart.

**Bangert, M., Peschel, T., Rotte M.** *et al.* (2005). Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. NeuroImage, in press. Doi:10.1016/j.neuroimage.2005.10.044

**Besson, M. and Faita, F.** (1995). An event related potential (ERP) study of musical expectancy: comparison of musicians with non-musicians. *J Exp Psychol Human Percept Perform*, **21**, 1278–1296.

**Carter, C.S., Mintum, M., and Cohen, J.D.** (1995). Interference and facilitation effects during selective attention: an H2150 PET study of Stroop task performance. *NeuroImage*, **2**, 264–272.

**Charness, M. and Schlaug, G.** (2000). Cortical activation during finger movements in concert pianists, dystonic pianists, and non-musicians. *Neurology*, **54**, A221.

**Classen, J., Liepert, J., Wise, S.P., Hallett, M., and Cohen, L.G.** (1998). Rapid plasticity of human cortical movement representation induced by practice. *J Neurophysiol*, **79**(2), 1117–1123.

**Coker, J.** (1997). *Jerry Coker's complete method for improvisation: for all instruments*, revised edn. Warner Bros, Miami, FL.

**Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., and Taub, E.** (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, **270**, 305–307.

**Fadiga, L., Buccino, G., Craighero, L., Fogassi, L., Gallese, V., and Pavesi, G.** (1999). Corticospinal excitability is specifically modulated by motor imagery: a magnetic stimulation study. *Neuropsychologia*, **37**, 147–158.

**Gaab, N., Gaser, C., Zaehle, T., Jancke, L., and Schlaug, G.** (2003). Functional anatomy of pitch memory–an fMRI study with sparse temporal sampling. *NeuroImage*, **19**, 1417–1426.

**Gerardin, E.A., Sirigu, S., Lehericy, J.B.** *et al.* (2000). Partially overlapping neural net-works for real and imagined hand movements. *Cereb Cortex*, **10**, 1093–1104.

**Griffiths, T.D., Büchel, C., Frackowiak, R.S., and Patterson, R.D.** (1998). Analysis of temporal structure in sound by the human brain. *Nat Neurosci*, **1**, 422–427.

**Halpern, A.R. and Zatorre, R.J.** (1999). When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. *Cereb Cortex*, **9**(7), 697–704.

**Hasegawa, T., Matsuki, K., Ueno, T.** *et al.* (2004). Learned audio-visual cross-modal associations in observed piano playing activate the left planum temporale. An fMRI study. *Cogn Brain Res*, **20**, 510–518.

**Haslinger, B., Erhard, P., Altenmüller, E.** *et al.* (2004). Reduced recruitment of motor association areas during bimanual coordination in concert pianists. *Hum Brain Mapp*, **22**(3), 206–215.

**Haslinger, B., Erhard, P., Altenmüller, E., Schroeder, U., Boecker, H., and Ceballos-Baumann, A.O.** (2005). Transmodal sensorimotor networks during action observation in professional pianists. *J Cogn Neurosci*, **17**, 282–293.

**Haueisen, J. and Knösche, T.R.** (2001). Involuntary motor activity in pianists evoked by music perception. *J Cogn Neurosci*, **13**(6), 786–792.

**Hickok, G., Buchsbaum, B., Humphries, C., and Muftuler, T.** (2003). Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *J Cogn Neurosci*, **15**(5), 673–682.

**Hund-Georgiadis, M. and von Cramon, Y.** (1999). Motor-learning-related changes in piano players and non-musicians revealed by functional magnetic-resonance signals. *Exp Brain Res*, **125**, 417–425.

**Itoh, K., Fuji, Y., Suzuki, K., and Naked, T.** (2001). Asymmetry of parietal lobe activation during piano performance: a high field functional magnetic resonance imaging study. *Neurosci Lett*, **309**, 41–44.

**Jacobs, K.M. and Donoghue, J.P.** (1991). Reshaping the cortical motor map by unmasking latent intracortical connections. *Science*, **251**, 944–947.

**Jäncke, L., Shah, N.J., and Peters, M.** (2000). Cortical activations in primary and secondary motor areas for complex bimanual movements in professional pianists. *Brain Res Cogn Brain Res*, **10**(1–2), 177–183.

**Jeannerod, M.** (2001). Neural simulation of action: a unifying mechanism for motor cognition. *NeuroImage*, **14**, S103–S109.

**Karni, A., Meyer, G., Jezzard, P.** *et al.* (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature*, **377**, 155–158.

**Karni, A., Meyer, G., Rey-Hipolito, C.** *et al.* (1998). The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. *Proc Natl Acad Sci* USA, **95**, 861–868.

**Keysers, C., Kohler, E., Umilta, M.A., Nanetti, L., Fogassi, L., and Gallese, V.** (2003). Audiovisual mirror neurons and action recognition. *Exp Brain Res*, **153**, 628–636.

**Kohler, E., Keysers, C., Umilta, M.A., Fogassi, L., Gallese, V., and Rizzolatti, G.** (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, **297**, 846–848.

**Kölsch, S., Gunter, T.C., von Cramon, D.Y., Zysset, S., Lohmann, G., and Friederici, A.D.** (2002). Bach speaks: a cortical 'language-network' serves the processing of music. *NeuroImage*, **17**(2), 956–966.

**Krings, T., Topper, R., Foltys, H.** *et al.* (2000). Cortical activation patterns during complex motor tasks in piano players and control subjects. A functional magnetic resonance imaging study. *Neurosci Lett*, **278**, 189–193.

**Kristeva, R., Chakarov, V., Schulte-Monting, J., and Spreer, J.** (2003). Activation of cortical areas in music execution and imagining: a high-resolution EEG study. *NeuroImage*, **20**(3), 1872–1883.

**Langheim, F.J., Callicott, J.H., Mattay, V.S., Duyn, J.H., and Weinberger, D.R.** (2002). Cortical systems associated with covert music rehearsal, *NeuroImage*, **16**, 901–908.

**Levitin, D.J. and Menon, V.** (2003). Musical structure is processed in 'language' areas of the brain: a possible role for Brodmann area 47 in temporal coherence. *NeuroImage*, **20**, 2142–2152.

**Liberman, A.** (1995). *Speech: a special code*. MIT Press, Cambridge, MA.

**Liegeois-Chauvel, C., Peretz, I., Babaï, M., Laguitton, V., and Chauvel, P.** (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain*, **121**, 1853–1867.

**Lotze, M., Montoya, P., Erb, M.** *et al.* (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: an fMRI study. *J Cogn Neurosci*, **11**, 491–501.

**Lotze, M., Scheler, G., Tan, H.R., Braun, C., and Birbaumer, N.** (2003). The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *NeuroImage*, **20**(3), 1817–1829.

MacKay, D.G. (1992). Constraints of theories of inner speech. In: D. Reisberg (ed.) *Auditory imagery*, pp. 121–149. Lawrence Erlbaum Associates, Hillsdale, NJ.

Meister, I.G., Krings, T., Foltys, H. *et al.* (2004). Playing piano in the mind—an fMRI study on music imagery and performance in pianists. *Cogn Brain Res*, **19**, 219–228.

Merzenich, M.M., Recanzone, G., Jenkins, W.M. *et al.* (1988). Cortical representational plasticity. In: P. Rakic and W. Singer (eds) *Neurobiology of neocortex*, pp. 41–47. John Wiley and Sons, New York.

Ohnishi, T., Matsuda, H., Asada, T. *et al.* (2001). Functional anatomy of musical perception in musicians. *Cereb Cortex*, **11**(8), 754–760.

Parsons, L.M. (2001). Exploring the functional neuroanatomy of music performance, perception, and comprehension. *Ann NY Acad Sci*, **930**, 211–231.

Pascual-Leone, A. (2001). The brain that plays music and is changed by it. *Ann NY Acad Sci*, **930**, 315–329.

Pascual-Leone, A., Grafman, J., and Hallett, M. (1994). Modulation of cortical motor output maps during development of implicit and explicit knowledge. *Science*, **263**, 1287–1289.

Pascual-Leone, A., Nguyet, D., Cohen, L.G., Brasil-Neto, J.P., Cammarota, A., and Hallett, M. (1995). Modulation of muscle responses evoked by transcranial magnetic stimulation during the acquisition of new fine motor skills. *J Neurophysiol*, **74**(3), 1037–1045.

Pascual-Leone, A., Tarazona, F., and Catala, M.D. (1999). Applications of transcranial magnetic stimulation in studies on motor learning. *Electroencephalogr Clin Neurophysiol Suppl*, **51**, 157–161.

Patel, A.D., Gibson, E., Ratner, J. *et al.* (1998). Processing syntactic relations in language and music: an event-related potential study. *J Cogn Neurosci*, **10**, 717–733.

Penhune, V.B., Zatorre, R.J., and Evans, A.C. (1998). Cerebellar contributions to motor timing: a PET study of auditory and visual rhythm. *J Cogn Neurosci*, **10**(6), 752–765.

Petrides, M. and Pandya, D.N. (1999). Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *Eur J Neurosci*, **11**, 1011–1036.

Petrides, M. and Pandya, D.N. (2002). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *Eur J Neurosci*, **16**, 291–310.

Porro, C.A., Francescato, M.P., Cettolo, V. *et al.* (1996). Primary motor and sensory cortex activation during motor performance and motor imagery: a functional magnetic resonance imaging study. *J Neurosci*, **16**, 7688–7698.

Price, C.J., Wise, R.J., Warburton, E.A. *et al.* (1996). Hearing and saying. The functional neuro-anatomy of auditory word processing. *Brain*, **119**, 919–931.

Rao, S.M., Harrington, D.L., Haaland, K.Y., Bobholz, J.A., Cox, R.W., and Binder, J.R. (1997). Distributed neural systems underlying the timing of movements. *J Neurosci*, **17**, 5528–5535.

Rizzolatti, G. and Arbib, M.A. (1998). Language within our grasp. *Trends Neurosci*, **21**, 188–194.

Rizzolatti, G. and Fadiga, L. (1998). Grasping objects and grasping action meanings: the dual role of monkey rostroventral premotor cortex (area F5). *Novartis Found Symp*, **218**, 81–95, discussion 95–103.

Romanski, L.M. and Goldman-Rakic, P.S. (2002). An auditory domain in primate prefrontal cortex. *Nat Neurosci*, **5**, 15–16.

**Romanski, L.M., Bates, J.F., and Goldman-Rakic, P.S.** (1999), Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. *J Comp Neurol*, **403**, 141–157.

**Samson, S. and Zatorre, R.J.** (1988). Melodic and harmonic discrimination following unilateral cerebral excision. *Brain Cogn*, **7**, 348–360,

**Samson, S. and Zatorre, R.J.** (1992). Learning and retention of melodic and verbal information after unilateral temporal lobectomy. *Neuropsychologia*, **30**, 815–826.

**Schlaug, G., Jäncke, L., Huang, Y., Staiger, J.F., and Steinmetz, H.** (1995). Increased corpus callosum size in musicians. *Neuropsychologia*, **33**(8), 1047–1055.

**Schoenberg, H.** (1987). *Great pianists*. Fireside Books, St Louis, MO.

**Seitz, R.I. and Roland, P.E.** (1992). Learning of sequential finger movements in man: a combined kinematic and positron emission tomography (PET) study. *Eur J Neurosci*, **4**, 154–165.

**Seitz, R.J., Roland, P.E., Bohm, C.** *et al.* (1990). Motor learning in man: a positron emission tomographic study. *NeuroReport*, **1**, 57–60.

**Sergent, J., Zuck, E., Terriah, S., and MacDonald, B.** (1992). Distributed neural network underlying musical sight-reading and keyboard performance. *Science*, **257**, 106–109.

**Sirigu, A., Cohen, L., Duhamel, J.R.** *et al.* (1995). Congruent unilateral impairments for real and imagined hand movements. *NeuroReport*, **6**, 997–1001.

**Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes, A., and Roberts, N.** (2002). Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *NeuroImage*, **17**(3), 1613–1622.

**Smith, J.D., Reisberg, D., and Wilson, M.** (1992). Subvocalization and auditory imagery: interactions between the inner ear and inner voice. In: D. Reisberg (ed.) *Auditory imagery*, pp. 95–120. Lawrence Erlbaum Associates, Hillsdale, NJ.

**Sokolov, Y.N.** (1972). *Inner speech and thought*. Plenum, New York.

**Stewart, L., Henson, R., Kampe, K., Walsh, V., Turner, R., and Frith, U.** (2003). Brain changes after learning to read and play music. *NeuroImage*, **20**(1), 71–83.

**Tervaniemi, M., Kujala, A., Alho, K., Virtanen, J., Ilmoniemi, R.J., and Näätänen, R.** (1999). Functional specialization of the human auditory cortex in processing phonetic and musical sounds: a magnetoencephalographic (MEG) study. *NeuroImage*, **9**, 330–336.

**Wang, X., Merzenich, M.M., Sameshima, K.** *et al.* (1995). Remodeling of hand representation in adult cortex determined by timing of tactile stimulation. *Nature*, **378**, 71–75.

**Zatorre, R.J. and Halpern, A.R.** (1993). Effect of unilateral temporal-lobe excision on perception and imagery of songs. *Neuropsychologia*, **31**(3), 221–232.

**Zatorre, R.J. and Samson, S.** (1991). Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain*, **114**, 2403–2417.

**Zatorre, R.J., Evans, A.C., and Meyer, E.** (1994). Neural mechanisms underlying melodic perception and memory for pitch. *J Neurosci*, **14**, 1908–1919.

**Zatorre, R.J., Halpern. A.R., Perry, D.W., Meyer, E., and Evans, A.C.** (1996). Hearing in the mind's ear: a PET investigation of musical imagery and perception. *J Cogn Neurosci*, **8**, 29–46.

Chapter 12

# Brain activation during string playing

Arto Nirkko and Rumyana Kristeva

## Introduction

Playing music with any type of instrument requires precise spatial and accurate temporal, mostly bimanually coordinated, motor control. The level of sophistication needed to produce an enjoyable musical performance can only be reached by long-term learning of these specific motor skills, usually over many years. In addition to motor performance, musical perception abilities—consciously or not—are also trained in parallel, as they are necessarily involved in the feedback processes which serve to fine-tune the motor output in the long term: the speed at which music is produced easily exceeds both the central and peripheral conduction times, making closed feedback control impossible. Direct feedback is limited by nerve conduction velocities, already delaying afferent proprioceptive motor input by tens of milliseconds in the periphery. The possibility for closed loop auditory feedback is additionally hampered by the fact that the music must be 'perfect' in the first place during the performance, and cannot be adjusted during the performance until the auditory feedback is perfect. For these two reasons feed-forward control is required (Wolpert and Ghahramani 2000; Wolpert and Flanagan 2001). Continuous mid- or long-term adjustment must rely on both proprioceptive and auditory feedback, especially during the learning phase but also later to ensure stability: even professional musicians need to continue practising.

Thus, the end product of musical performance is based on highly over-learnt sensorimotor skills, which interact with perceptive auditory skills (which are also over-learnt) and continuous adjustment. Further (in terms of temporal coordination) more independent interactions with other sensory, cognitive, and motor systems are evident from the simultaneous requirements for reading notes, watching, and interpreting the conductor (vision and control of eye movement). These processes interact with working memory and the recall of musical sequences from long-term memory, and also with processes such as postural motor control. The latter is not directly involved in producing tone sequences from the musical instrument, but enables the body and instrument to

be kept in a relative position suitable for producing these tone sequences. In addition, a component of actually enjoying the music is likely to be present in the players as well as the listeners. In the player, this is complemented by a probable component of enjoying the playing itself.

With these considerations in mind, it seems quite evident that large parts of the brain must be involved in playing music, many or all of them shared with other tasks, not necessarily of a musical nature. This poses specific problems for the assessment of different components, like 'pure playing' (or 'music production'), as opposed to 'listening' (or 'music perception'), and assessment of some other components involved. Principally, planning and production could be segregated in time from perception, because of the temporal differences between motor preparation, execution, physical tone generation, and auditory perception. This is possible when using methods with high temporal resolution, such as electroencephalography (EEG), which directly assess neuronal activity. The slow negative shifts of the cortical direct current (DC) potential have been shown to reflect activation of the underlying areas. Thus the preparation of a task is accompanied by task-preceding negativity and the execution of a task is accompanied by task-processing negativity. In contrast to EEG, current imaging methods with a high spatial resolution, such as functional magnetic resonance imaging (fMRI), depend on a relatively sluggish haemodynamic response: they are thus biologically limited to an inherently low temporal resolution of the order of several seconds, which is insufficient for direct observation of the temporal sequence of different processes following each other with delays in the range of (tens to hundreds of) milliseconds. An alternative solution to the problem is the segregation of processes related to production from processes related to perception by task design, using a set-up to assess only one component per fMRI experiment. Another intriguing question is whether executing and imagining music involve the same areas of the brain.

In this chapter we attempt to integrate data from both fMRI and EEG to elucidate the activation of the cortical areas in musical performance for both execution and imagination of music.

## Stringed instruments and other instruments

Whereas many processes during string playing may be assumed to share similarities with production of music on other instruments, one evident difference is the highly asymmetric quality of the function of each hand in playing a stringed instrument. In many instruments, such as the piano, the fingers of both hands also have to perform asymmetric and intermanually well-coordinated movements, but the functional significance of these movements is similar (selecting or

producing a specific tone) and can even sometimes be performed by either one hand or the other, depending on the spatiotemporal context in the musical sequence. In contrast to this, in string playing each hand has multiple separate, fundamentally different and non-interchangeable functions: one hand (usually the left) has both a postural function in holding the instrument and its main musical function in 'selecting' specific tones, whereas the other hand (usually the right) initiates and sustains production and modulation of the tone. This hand also has a role in selecting the tone on a coarser level (selecting the active strings), but even this selection is fundamentally different from what the left hand does, both regarding different finger, hand, and arm movements and a different musical function. Thus, hemispheric asymmetries have to be specially taken into account in the evaluation of the brain representation during string playing.

## The paradigms

In both fMRI and EEG experiments we used the same musical sequence as used by Wiesendanger and colleagues (see Fig. 12.3 and Chapter 7).

Four musicians were investigated in the fMRI experiments. To segregate the production of music from its perception we used pure motor 'music playing' on an MRI-compatible 'sham violin' made of wood, without strings or any other means to produce sound, thus eliminating auditory feedback processes. Further, visual processing was avoided by having the subjects practice and then play the short sequence by heart, with their eyes closed. Due to spatial restrictions in the MRI scanner, both the pseudo-violin and the pseudo-bow were small and were not held in the normal position, but rather were rested on the abdomen when lying in a supine position. The resulting movement excursions were distal, acted from the wrist and not with the whole arm. With this set-up, only a minimum of postural functions were involved and the paradigms focused on motor sequences with direct musical meaning (tone selection with the fingers of the left hand, and bowing with the right hand). To assess regions specifically involved in music production and not only in motor control, the musical task was compared with a bimanual motor sequencing task without musical meaning: the left hand had to repetitively pull a drawer on an MRI-compatible manipulandum, while the right hand had to grip an object in the drawer, in a bimanually synergistic and synchronized manner. Regions known to be mainly concerned with motor execution (the primary sensorimotor cortex, SM1) served as a control region: These regions were not expected to exhibit relevant components specific to music, and were used as an internal control to assess whether the demands of the different tasks on the motor system were similar. A block design with continuous repetition of the same sequence (either with movements executed or only

imagined) was used as the activation task, and with rest (holding the instrument) as the baseline task.

In the EEG experiments we recorded high-resolution EEG data from six healthy right-handed professional violin players (mean musical practice for 15.7 (standard deviation, SD 4.1) years) (Kristeva *et al.* 2003). The EEG was recorded before, during, and after playing single runs of the musical sequence on real violins and were averaged time-locked to the start and to the end of the musical sequence. Averaging time-locked to the start of the musical sequence allowed us to demonstrate the onset of the music-preceding negativity, i.e. the negative slope of the cortical DC potential, which reflected the sequence of activation of the different brain areas during the preparation for execution or imagination of the music. Averaging time-locked to the end of the musical sequence allowed us to show the music-processing negativity, i.e. the sustained negativity which reflected the activation of the cortical areas during execution or imagination of the music.

In both fMRI and EEG experiments, the playing the music was compared with imagining the music. This allowed separation of the areas mainly involved in motor execution from those involved in imagining, planning, and working memory, thus differentiating musical from purely motor areas.

## Methods

### Functional MRI (fMRI)

Functional MRI was performed as previously described (Nirkko *et al.* 2001): in short, blood oxygenation level-dependent (BOLD) fMRI data were acquired on a clinical 1.5 T scanner (Magnetom Vision, Siemens, Erlangen, Germany) with the standard circularly polarized head coil, using a high-resolution ($128 \times 128$ matrix, 1.56 mm pixel size) whole-brain (30 slices, one every 4 mm, interleaved acquisition) echo planar imaging (EPI) sequence with a flip angle of 90° and an echo time (TE) of 82 ms, at a rate of one whole-brain volume every repetition time (TR) of 6 s, for a total of 68 volumes in 408 s. The slices were tilted parallel to a line connecting the caudal outline of the genu of the corpus callosum and the occipital protuberance/confluens sinuum. This angulation was found to cover the whole brain including the cerebellum with a minimum of slices, to be more reproducible than the AC–PC line even on slightly non-median scouts, and to reduce susceptibility artefacts due to pneumatized areas below the frontobasis and the thalamic regions (Nirkko *et al.* 2001). Evaluation was performed on a Sun/Sparc/Solaris workstation (Sun Microsystems, Palo Alto, CA, USA) with custom-written software (Nirkko 2000). After motion correction (which was minimal due to fixation of the subject's head with a U-shaped dental plate in

addition to the lateral pads on the head coil), *Z*-score maps selecting clusters of at least 15 pixels size and a peak *Z* score of at least 8 (corresponding to a Bonferroni correction for all acquired pixels in the whole brain volume) were colour coded and overlaid on the original EPI images for evaluation with precise anatomical localization free of the possibilities of misregistration, and on higher-resolution structural spin-echo images with identical positioning parameters for visualization in the figures.

A quantitative measure for activation as a percentage of signal change was measured in structurally defined volumes of interest (VOI), which were manually defined again on the original EPI slices: in addition to the primary and secondary sensorimotor areas described in Nirkko *et al.* (2001), the pre-sylvian frontal opercular area (corresponding to part of Broca's region) and the anterior and posterior insular regions were assessed bilaterally. The mean percentage activation and standard deviation in each region was calculated from the 68 samples per experiment and used for statistical comparison between groups and between regions.

## Electroencephalography (EEG)

Electrical potentials (bandpass filter 0–100 Hz; sampling rate 500 Hz) were recorded from 58 scalp positions. An accelerometer was placed on the right index finger to define the start and the end of the musical sequence.

The musical sequence was executed or imagined in a self-paced mode, at irregular intervals of 12 to 24 s, starting from complete relaxation (Bereitschaftspotential paradigm (Kornhuber and Deecke 1965)). Each subject executed or imagined the musical sequence 90 times in total per condition (playing on the real violin and imagining). The analysis time for the potentials time-locked to the start of the melody (for investigating the music-preceding negativity) was set to 5000 ms before to 1500 ms after the start trigger. After that the signal was averaged and baseline-corrected using the time between 5 s and 4 s prior to the start trigger.

The analysis time for the potentials time-locked to the end of the melody (for investigating the music-processing negativity) was set to 8 s before the end trigger to 1 s after it. After that the signal was averaged and baseline-corrected between 8 and 7.5 s before the end trigger.

To describe the start of the preparatory brain activity we measured the time of onset of the music-preceding negativity. To describe the preparatory brain activity at the start of the sequence, we measured the amplitude of the music-preceding negativity calculated as the mean negativity over the time between 89 ms prior to and 30 ms after the start of the musical sequence. To investigate brain activity during the musical sequence, we calculated the mean negativity

over the time between 2000 ms and 1000 ms before the end of the sequence, i.e. of the music-processing negativity.

In one of the subjects we performed advanced source reconstruction analysis, taking into consideration the individual brain morphology from MRI (for more details see Kristeva *et al.* (2003)). Thus we were able to show the cortical sources during *preparation* for execution and imagination of music and during execution and imagination of the music itself.

## Brain activation patterns

In fMRI, in both the violin task and the bimanual reference task, a prominent activation pattern corresponding to the previously described sensorimotor network of both hands was seen (Nirkko *et al.* 2001), consisting of the primary sensorimotor cortex around the central sulcus (SM1), the secondary cortical areas of the pre-motor cortex (PMC) and the posterior parietal cortex (PPC), the supplementary motor area (SMA), and the cerebellum. In the violin task only, activation was consistently evident in the pre-sylvian frontal operculum ('BROC', part of Broca's language area) (Plate 10) on standard $Z$-score maps (six of eight sides in the four violin experiments, as compared with four of 20 sides in the 10 bimanual experiments). Quantitative evaluation yielded significant signal increases of 0.86% ($P << 0.001$, always corrected) in the left BROC during violin playing, and 0.44% ($P << 0.001$) in the right BROC. Conversely, the bimanual reference task also yielded a highly significant, but quantitatively much smaller, signal increase of 0.24% in the left BROC and 0.18% in the right BROC. The violin task thus lead to a 3.6 times higher activation in the left BROC than the bimanual reference task, and the difference between the two tasks was highly significant ($P << 0.001$, corrected).

The SM1 (left and right sides pooled) showed no difference between the tasks (difference 0.10%, not significant). Pooling both sides was necessary because by its very nature the violin task put greater demands on fingering with the left hand, and thus led to significantly asymmetric activation. Notably, this asymmetry during violin playing in SM1 (right more than left) is opposite to the asymmetry in BROC (left more than right).

During the music imagining task, a significant fMRI activation of the bilateral BROC and of the SMA without activation of the left and right SM1 was observed.

In the EEG experiment the execution of the melody lasted 3.2 $\pm$ 0.6 s and imagining the melody lasted 3.1 $\pm$ 0.4 s.

Figure 12.1 (upper panel) shows the music-preceding negativity for two representative subjects for execution and for imagining the sequence. This music-preceding negativity is seen over the left and right BROC, over the SMA, and over the left and right SM1s.

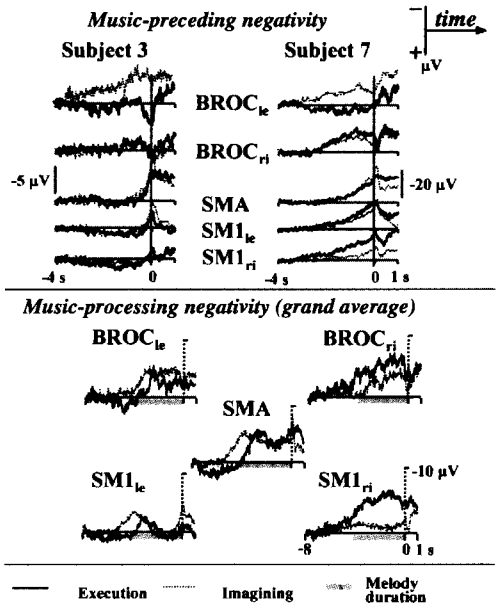



**Fig. 12.1** EEG results. Upper panel: Music-preceding negativity (averaged potentials time-locked to the start of execution (solid line) and to the start of imagining (dotted line)). Negativity up. Averaged potentials for Subject 3 (S3, on the left) and Subject 7 (S7, on the right) recorded from electrodes overlying the left and right frontal opercular regions ($BROC_{le}$, $BROC_{ri}$), supplementary motor area (SMA), left primary sensorimotor ($SM1_{le}$), and right primary sensorimotor ($SM1_{ri}$) areas. Averages from 70 trials for execution and 64 for imagining conditions for S3. Averages from 44 trials for execution and 66 for imagining conditions for S7. Note the earlier onset over the left and right frontal opercular regions for S3. The onset times for S7 are similar for bilateral frontal opercular regions and for left and right SM1. Lower panel: Music-processing negativity (grand average curves for potentials time-locked to the end of the musical sequence for execution (solid line) and for imagining (dotted line)) for $BROC_{le}$, and $BROC_{ri}$, SMA, $SM1_{le}$, and $SM1_{ri}$. Negativity up. The duration of the sequence is marked by a grey line. Note the high music-processing negativity over bilateral frontal opercular regions and SMA for both execution and imagining the sequence. Note also that the processing negativity over the right SM1 is very high and it occurs only during execution. Note also that the left SM1 is activated only around the start and the end of the sequence for both conditions. (Modified from Kristeva *et al.* (2003) with the kind permission of *NeuroImage*.)

While the beginning of the preparatory activity of the brain for the SMA, the left SM1, and the right SM1 was very consistent, this was not the case with the preparatory activity of the 'more cognitive' areas, the left and right BROC: in three of the musicians the bilateral BROC became active earlier (about 3 s before the start) than the motor areas (about 1.5 s prior to the start). The music-preceding negativity for one of these three subjects (S3) is shown in Fig. 12.1 (upper panel, on the left). In one of the musicians the bilateral BROC became active simultaneously with the motor areas at approx. 3 s prior to the start.

The music-preceding negativity for this subject is shown in Fig. 12.1 (upper panel, on the right). In the remaining two subjects the music-preceding potential prior to execution and imagining showed a very high variability and the objective criterion for defining the onset time was not met, thus making it impossible to make a numeric statement.

Higher preparatory brain activity at the start of the execution of the melody was found for the right sensorimotor cortex, i.e. contralateral to the fingering hand.

The music-processing negativity reflecting the activation of the different cortical areas during music execution and imagining is shown in Fig. 12.1 (lower panel). Music-processing negativity over the SMA is seen for both execution and imagining. A large sustained music-processing negativity during both conditions can also be seen, as over the left and right BROC with stronger involvement of the left BROC during imagining and of the right BROC during real violin playing.

Large differences between playing and imagining are seen in the activity of the right SM1. During execution of the melody the activity of the right SM1 is characterized by a large processing negativity. During the imagination of the melody, brain activity is observed only shortly preceding the movement marking the end of imagining the sequence. Brain activity in the left SM1 only occurs before the end of the execution and before the end of imagining the sequence.

While the above analysis of the music-specific activity is based on the form and time course of the cortical DC potential recorded from five cortical areas only, Plate 11 shows the reconstructed cortical sources for one of the subjects on the basis of the potential distribution from 58 electrode positions. Before both execution and imagination of the melody, the reconstructed cortical current density (CCD) maps suggest distributed sources in primary sensorimotor areas, the SMA and bilaterally in the BROC. During execution of the melody additional activation is observed in the superior temporal gyrus.

During imagination of the melody a strong activation is present in the SMA, the gyrus supramarginalis, and the frontal polar areas. However, no activation of the superior temporal gyrus or of the primary sensorimotor areas is observed.

## Broca's area and its right-hemispheric homologue

Our fMRI and EEG results both indicate that the fronto-opercular part of the Broca language regions is involved in the production of music. This region has already been implicated in motor sequencing and manipulation (Binkowski *et al.* 1999), and our data indeed demonstrate a significant but weak activation with a bimanual reference task as well. However, the fMRI activation during music

production was 3.6 times higher ($P < 0.001$ corrected), suggesting a relevant music-specific contribution.

The stronger involvement of the right BROC during execution of the melody in the EEG experiment is well in line with findings of Bangert *et al.* (1998). Using DC-EEG these authors showed pre-frontal-temporobasal activation during piano playing which was interpreted to reflect audiomotor integration. The results of Rao *et al.* (1997) are also in support of the role of the right inferior frontal gyrus in music production. Using fMRI these authors suggested that the activation of the right inferior frontal gyrus supports the retrieval and rehearsal of internal non-linguistic auditory representations of time intervals.

The EEG results show a considerably larger area of activation than the fMRI results, covering not only the fronto-opercular areas corresponding to the Broca language centres, but also parts of the temporal lobe. This is presumably because the auditory feedback during the EEG experiments with real violins yielded an additional activation of the auditory cortex in the superior temporal lobe, which is adjacent to the frontal operculum as seen from the surface of the skull. The superior temporal gyrus was also shown to play a role in analysing higher-order patterns formed by the features of the individual notes (Griffiths 2001). In contrast, the fMRI experiments were performed without auditory feedback, yielding only areas related to pure motor performance.

Notably, the asymmetry in the motor cortex with preferential activation of the cortex contralateral to the more active hand was opposite to the asymmetry in Broca's area, which was concordant with the dominance of the left hemisphere in the language areas.

Traditionally, the language areas have been functionally divided between the temporal area (Wernicke) for language analysis and understanding, and the frontal area (Broca) for production of speech. Our fMRI results regarding music production, together with other studies assessing musical perception, suggest that a similar division applies to music. Interestingly, contrary to popular scientific belief that music is a more emotionally involved modality which might be represented in the right hemisphere, the fMRI results show that the left hemisphere is also dominant for music production, at least when pure production is assessed in professional musicians. However, our EEG results during real violin playing show that the right hemisphere is more strongly involved if musical perception is included, suggesting that the perception of the music produced is processed mainly in the right hemisphere.

The stronger involvement of the left fronto-opercular cortex during the imagination of music in the EEG experiment and during playing on the sham violin in the fMRI experiment suggest that the hemispheric preponderance of the music-specific activity is related to the presence or absence of auditory

feedback: the auditory perceptive component seems to involve the right hemispheric homologous region to a greater degree, while the performance/production of music alone (without perception) is more lateralized to the left. This interpretation is well in line with studies on musical perception using magnetoencephalography (MEG) (Maess *et al.* 2001), positron emission tomography (PET) (Halpern 2001), or fMRI (Ohnishi *et al.* 2001), where the temporal cortex (including the Wernicke receptive language area) was predominantly involved in the processing or imagination of music.

In the EEG experiment the left and right BROC displayed much larger interindividual differences in terms of beginning of the preparatory activity. This favours the notion that the cortical areas involved in music processing have individually formed patterns of temporal organization during music preparation (Altenmüller 2001). This finding argues against the common opinion that populations of musicians are rather homogeneous in their behaviour. The high variability may result from the high individualization of the musicians' brains during long years of intensive learning of music skills. However, one cannot exclude that the high variability may result from different brain anatomy.

## The sensorimotor cortical network

In the fMRI experiments there was no significant difference between musical and non-musical bimanual tasks regarding the activation of the pooled bilateral primary motor cortex (SM1) serving as the control region. This confirms that in terms of motor activation, the two tasks were equivalent. The asymmetry of SM1 activation, with stronger involvement of the right SM1 in the fMRI experiment during the violin task and in the EEG experiment during melody execution, is explained by the inherently asymmetric motor demands of violin playing on each hand and was to be expected. The right SM1 controls the left hand, which performs the more complex sequencing motor task. Further, there is a use-dependent cortical reorganization of the right somatosensory (Elbert *et al.* 1995) and primary motor area which has to overcome 'spill-over' and enslaving to organize the individuated distal movements of the left hand fingers learned during a long practice (Amunts *et al.* 1996; Slobounov *et al.* 2002).

The strong SMA activation and the absence of sources in the left and right SM1s when imagining the melody in both fMRI and EEG experiments is well in line with results of Stephan *et al.* (1995), Seitz and Roland (1992), and Binkowski *et al.* (2000). Functional MRI experiments with imagined movements showed activation of the SMA, especially its 'intermediate' part (Stephan *et al.* 1995). Previous work of ours (Ball *et al.* 1999) has also suggested an inhibitory/disinhibitory function for the intermediate SMA. This would offer an explanation

for its activation during imagination of the melody, because this specific task requires inhibition of overt movements, i.e. suppression of the left and right SM1. Lack of activation of both the left and right SM1 during imagination of a musical sequence was also shown in fMRI experiments (Nirkko *et al.* 2001; Langheim *et al.* 2002). Another explanation for the strong activation of the SMA during the imagining task could be the implication of this area in covert vocalization (Zatorre *et al.* 1996).

The activation of the pre-frontal areas shares similarities with verbal and visuospatial working memory activations as shown by fMRI (Janowsky *et al.* 1989; Milner *et al.* 1991; Wheeler *et al.* 1995; Cabeza and Nyberg 2000; Haxby *et al.* 2000; Fletcher and Henson 2001). It is possible that the activations of the pre-frontal areas seen by us are related to musical working memory.

## Conclusions

These observations about bilateral activation of the frontal opercula during preparation for, imagining, and execution of music, and the differences in hemispheric lateralization regarding perception and production extend the hypotheses about the functional role of these regions: The Broca area subserves language production and comprehension (Sasaki *et al.* 1995; Friederici 1998; Bookheimer *et al.* 2000; Caplan *et al.* 2000; Friederici *et al.* 2000; Caplan 2001; Caplan and Dapretto 2001; Fiez 2001; Moro *et al.* 2001; Blank *et al.* 2002; Heim *et al.* 2002; Roder *et al.* 2002; Tettamanti *et al.* 2002), musical syntax (Koelsch *et al.* 2001, 2002; Maess *et al.* 2001), action observation (Galese *et al.* 1996; Rizzolatti and Arbib 1998; Rizzolatti *et al.* 1998; Umilta *et al.* 2001) and imitation of actions (Iakoboni *et al.* 1999; Heiser *et al.* 2003), execution, recognition, and imagery of skilled forelimb movements and generally 'higher-order movement control' (Binkofski *et al.* 2000), selective attention and working memory, as well as acoustic learning (Schäffler *et al.* 1993). Our fMRI and EEG findings indicate that frontal opercular regions also subserve music production and imagining. It would be of interest to investigate how the different subareas of frontal opercular regions differentially contribute to the different functions.

The beginning of the preparatory activity of the Broca areas shows large inter-individual differences. This argues against the common opinion that populations of musicians (violin players) are rather homogeneous in their behaviours.

Our findings also allow some speculations about the important role of the bilateral frontal opercular networks in preparation for and during musical execution and imagining. It is possible that these networks have 'mirror neuron' properties that subserve not only action, observation, imagining, and imitation of other individual's actions, but also observation of one's own performance or

imagining (Kristeva *et al.* 2003), and together with their role in manipulation (Binkowski *et al.* 1999) are thus part of a feedback circuit subserving adjustment and fine-tuning of sequenced motor performance.

## Acknowledgements

We thank M. Wiesendanger for the inspiration for this work, and our colleagues who were involved in parts of the reported research: V. Chakarov, J. Schulte-Mönting, J. Spreer, H. Hampel (EEG), K.-O. Lövblad, and S. M. Redmond (fMRI). We gratefully acknowledge financial support received for the EEG part from DFG (grant Kr 1392/7–3) and from the Research Fund of the Albert-Ludwigs-University, and for the fMRI part from Swiss National Science Foundation (SNF grant 4038–044053) and the research funds of the neuroradiology and neurology departments of the University Hospital Berne.

## References

**Altenmüller, E.O.** (2001). How many music centers are in the brain? *Ann NY Acad Sci*, **930**, 273–280.

**Amunts, K., Schlaug, G., Schleicher, A.** *et al.* (1996). Asymmetry in the human motor cortex and handedness. *NeuroImage*, **4**, 216–222.

**Ball, T., Schreiber, A., Feige, B.** *et al.* (1999). The role of higher-order motor areas in voluntary movement as revealed by high-resolution EEG and fMRI. *NeuroImage*, **10**, 682–694.

**Bangert, M., Parlitz, D., and Altenmüller, E.** (1998). Cortical audiomotor corepresentation in piano learning and expert piano performance. *Soc Neurosci*, 658.

**Binkofski, F., Buccino, G., Posse, S., Seitz, R.J., Rizzolatti, G., and Freund, H.J.** (1999). A fronto-parietal circuit for object manipulation in man: evidence from an fMRI study. *Eur J Neurosci*, **11**, 3276–3286.

**Binkofski, F., Amunts, K., Stephan, K.M.** *et al.* (2000). Broca's region subserves imagery of motion: a combined cytoarchitectonic and fMRI study. *Hum Brain Map*, **11**, 273–285.

**Blank, S.C., Scott, S.K., Murphy, K., Warburton, E., and Wise, R.J.** (2002). Speech production: Wernicke, Broca and beyond. *Brain*, **125**, 1829–1838.

**Bookheimer, S.Y., Zeffiro, T.A., Blaxton, T.A., Gaillard, P.W., and Theodore, W.H.** (2000). Activation of language cortex with automatic speech tasks. *Neurology*, **55**, 1151–1157.

**Cabeza, R. and Nyberg, L.** (2000). Neural bases of learning and memory: functional neuroimaging evidence. (review). *Curr Opin Neurol*, **13**, 415–421.

**Caplan, D.** (2001). Functional neuroimaging studies of syntactic processing. *J Psycholinguist Res*, **30**, 297–320.

**Caplan, D., Alpert, N., and Waters, G.** (2000). PET studies of syntactic processing with auditory sentence presentation. *NeuroImage*, **9**, 343–351.

**Caplan, R. and Dapretto, M.** (2001). Making sense during conversation: a fMRI study. *NeuroReport*, **12**, 3625–3632.

**Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., and Taub, E.** (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, **270**, 305–307.

**Fiez, J.A.** (2001). Neuroimaging studies of speech: an overview of techniques and methodological approaches. *J Commun Disord*, **34**, 445–454.

**Fletcher, P.C. and Henson, R.N.** (2001). Frontal lobes and human memory: insights from functional neuroimaging. (review) *Brain*, **124**, 849–881.

**Friederici, A.D.** (1998). The neurobiology of language comprehension. In: *Language comprehension: a biological perspective*, pp. 263–301. Springer-Verlag, New York.

**Friederici, A.D., Wang, Y., Herrmann, C.S., Maess, B. and Oertel, U.** (2000). Localization of early syntactic processes in frontal and temporal cortical areas: a magnetoencephalographic study. *Hum Brain Map*, **11**, 1–11.

**Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G.** (1996). Action recognition in the premotor cortex. *Brain*, **119**, 593–609.

**Griffiths, T.D.** (2001). The neural processing of complex sounds. (review). *Ann NY Acad Sci*, **930**, 133–142.

**Haxby, J.V., Petit, L., Ungerleider, L.G., and Courtney, S.M.** (2000). Distinguishing the functional roles of multiple regions in distributed neural systems for visual working memory. (review) *NeuroImage*, **11**, 380–391.

**Halpern, A.R.** (2001). Cerebral substrates of musical imagery. *Ann NY Acad Sci*, **930**, 179–192.

**Heim, S. Opitz, B., and Friederici, A.D.** (2002). Broca's area in the human brain is involved in the selection of grammatical gender for language production: evidence from event-related functional magnetic resonance imaging. *Neurosci Lett*, **328**, 101–104.

**Heiser, M., Iacoboni, M., Maeda, F., Marcus, J., and Mazziotta, J.C.** (2003). The essential role of Broca's area in imitation. *Europ J Neurosci*, **17**, 1123–28.

**Iacoboni, M., Woods, R.P., Brass, M.** *et al.* (1999). Cortical mechanisms of human imitation. *Science*, **286**, 2526–2528.

**Janowsky, J.S., Shimamura, A.P., Kritchevsky, M., and Squire, L.R.** (1989). Cognitive impairment following frontal lobe damage and its relevance to human amnesia. *Behav Neurosci*, **103**, 548–560.

**Koelsch, S., Gunter, T., Cramon, D.** *et al.* (2002). Bach speaks: a cortical 'Language-Network' serves the processing of music. *NeuroImage*, **17**, 956–966.

**Koelsch, S., Maess, B., Gunter, T.C., and Friederici, A.D.** (2001). Neapolitan chords activate the area of Broca. A magnetoencephalographic study. *Ann NY Acad Sci*, **930**, 420–421.

**Kornhuber, H.H. and Deecke, L.** (1965). Hirnpotentialänderungen bei Willkürbewegungen und Passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflügers Arch*, **284**, 1–17.

**Kristeva, R., Chakarov, V., Schulte-Monting, J., and Spreer, J.** (2003). Activation of cortical areas in music execution and imagining: a high-resolution EEG study. *NeuroImage*, **20**(3), 1872–1883.

**Kristeva-Feige, R.** (2003) Distributed source modeling in the analysis of movement-related activity. In: M. Jahanshahi and M. Hallett (eds) *The Bereitschaftspotential. Movement-related cortical potentials*, pp. 79–93. Kluwer Academic/Plenum Publishers, New York.

**Langheim, F., Callicott, J.H., Mattay, V.S., Duyn, J.H., and Weinberger, D.R.** (2002). Cortical systems associated with covert music rehearsal. *NeuroImage*, **16**, 901–908.

Maess, B., Koelsch, S., Gunter, T.C., and Friederici, A.D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nat Neurosci*, **4**, 540–545.

Milner, B., Corsi, P., and Leonard, G. (1991). Frontal-lobe contribution to recency judgements. *Neuropsychologia*, **29**, 601–618.

Moro, A., Tettamanti, M., Perani, D. *et al.* (2001). Syntax and the brain: disentangling grammar by selective anomalies. *NeuroImage*, **13**, 110–118.

Nirkko, A.C. (2000). A small software utility for fully automated download and evaluation of fMRI data. *NeuroImage*, **11**, S919.

Nirkko, A.C., Ozdoba, C., Redmond, S.M. *et al.* (2001*a*). Different ipsilateral representations for distal and proximal movements in the sensorimotor cortex: activation and deactivation patterns. *NeuroImage*, **13**(5), 825–835.

Nirkko, A.C., Baader, A.P., Loevblad, K.O., Milani, P., and Wiesendanger, M. (2001*b*). Cortical representation of music production in violin players: behavioral assessment and functional imaging of finger sequencing, bimanual coordination and music specific brain activation. *NeuroImage*, **5**, 106.

Ohnishi, T., Matsuda, H., Asada, T., *et al.* (2001). Functional anatomy of musical perception in musicians. *Cereb Cortex*, **11**, 754–760.

Rao, S.M., Harrington, D.L., Haaland, K.Y. *et al.* (1997). Distributed neural systems underlying the timing of movements. *J Neurosci*, **17**, 5528–5535.

Rizzolatti, G. and Arbib, M.A. (1998). Language within our grasp. *Trends Neurosci*,**21**, 188–194.

Rizzolatti, G., Luppino, G., and Matelli, M. (1998). The organization of the cortical motor system: new concepts. (Review) *Electroencephalogr Clin Neurophysiol*, **106**, 283–296.

Roder, B., Stock, O., Neville, H., Bien, S., and Rösler, F. (2002). Brain activation modulated by the comprehension of normal and pseudo-word sentences of different processing demands: a functional magnetic resonance imaging study. *NeuroImage*, **15**, 1003–1014.

Sasaki, K., Kyuhou, S., Nambu, A. *et al.* (1995). Motor speech centres in the frontal cortex. *Neurosci Res*, **22**, 245–248.

Schäffler, L., Luders, H.O., Dinner, D.S., Lesser, R.P., and Chelune, G.J. (1993). Comprehension deficits elicited by electrical stimulation of Broca's area. *Brain*, **116**, 695–715.

Seitz, R.J. and Roland, P.E. (1992). Learning of sequential finger movements in man: a combined kinematic and positron emission tomography (PET) study. *Eur J Neurosci*, **4**, 154–165.

Slobounov, S., Chiang, H., Johnston, J., and Ray, W. (2002). Modulated cortical control of individual fingers in experienced musicians: an EEG study. *Clin Neurophysiol*, **113**, 2013–2024.

Stephan, K.M., Fink, G.R., Passingham, R.E. *et al.* (1995). Functional anatomy of the mental representation of upper extremity movements in healthy subjects. *J Neurophysiol*, **73**, 373–386.

Tettamanti, M., Alkadhi, H., Moro, A. *et al.* (2002). Neural correlates for the acquisition of natural language syntax. *NeuroImage*, **17**, 700.

Umilta, M.A., Kohler, E., Gallese, V. *et al.* (2001). I know what you are doing. A neurophysiological study. *Neuron*, **19**, 155–165.

**Wolpert, D.M. and Flanagan, J.R.** (2001). Motor prediction. (review). *Curr Biol*, **11**, R729–R732.

**Wolpert, D.M. and Ghahramani, Z.** (2000) Computational principles of movement neuroscience. (review). *Nat Neurosci*, **3**(Suppl), 1212–1217.

**Wheeler, M.A., Stuss, D.T., and Tulving, E.** (1995). Frontal lobe damage produces episodic memory impairment. *J Int Neuropsychol Soc*, **1**, 525–536.

**Zatorre, R.J., Halpern, A., Perry, D.W., and Meyer, E.** (1996). Hearing in the mind's ear: a PET investigation of musical imagery and perception. *J Cogn Neurosci*, **8**, 29–46.

Chapter 13

# 'Singing in the (b)rain': cerebral correlates of vocal music performance in humans

H. Ackermann, D. Wildgruber, and A. Riecker

## Introduction

In the first issue of the 1871 edition of *The Descent of Man, and Selection in Relation to Sex* Darwin suggested that musical vocalizations had evolved into human speech: '. . . some early progenitor of man probably used his voice largely, as does one of the gibbon-apes at the present day, in producing true musical cadences, that is in singing; . . . serving to express various emotions . . . . The imitation by articulate sounds of musical cries might have given rise to words expressive of various complex emotions' (p. 56). By contrast, the nineteenth-century philosopher Herbert Spencer, an adherent of Darwin's theory of natural selection as the driving force of biological evolution, proposed that 'the cadences used in emotional speech afford the foundation from which music has been developed' (Darwin 1981/1871, p. 336). As an alternative to these assumptions of a successive emergence of vocal music performance and speech production, in either direction, other models maintain that both capabilities arose together, from a common ancestral facility ('musilanguage stage of hominid vocalizations'; Brown 2000). In addition to structural similarities between speech utterances and singing behaviour, this suggestion has been based upon functional imaging data: 'the cortical substrates for both language and music depend upon widely distributed networks that in some cases overlap; use both sides of the brain, but are dominated by opposite hemispheres' (Falk 2000, p. 197). So far, the hypothesis of differential lateralization effects of these two modalities of human acoustic communication mainly rests upon data bound to music perception and musical imagery. The present chapter addresses the motor domain of vocal music performance trying to elucidate whether the neural substrates of human singing capabilities show opposite lateralization effects compared with speech and language functions.

A large number of studies addressing the cerebral correlates of motor and cognitive aspects of speech production have accumulated so far. By contrast, considerably fewer investigations of the brain circuits subserving vocal music performance are available. This chapter will review the relevant data obtained in both clinical populations and healthy subjects with a focus on functional haemodynamic imaging as a more recent approach to the study of the relationship between brain and behaviour. Since a close connection has been suggested between human singing capabilities and sound production in other species, especially the 'musical cadences' observed in birds and lesser apes (Darwin 1981/1871; Richman 1993), this chapter begins with a discussion of song-like vocalizations in non-human animals.

## Song production in birds and non-human primates: behavioural and neurobiological data

Like humans, birds are highly vocal creatures and acoustic communication plays a central role in their lives (Catchpole and Slater 1995). Songs are defined as 'long, complex vocalizations produced by males in the breeding season' whereas calls tend to be shorter and simpler, pertaining to the behavioural repertoire of both sexes throughout the year (p. 10). However, the line of demarcation between calls and songs is to some extent arbitrary. About 4000 out of a total of 9000 or so avian species known to exist belong to the category of song birds (oscines). Although not undisputed, the two main functions of male song appear to be female attraction and territorial defence.

The syrinx, i.e. the bird's equivalent to the human larynx, contains special membranes set into vibration when air from the lungs is forced over them (Suthers 1999). Some, but not all, oscine species exhibit a functional asymmetry of the innervation of this sound-generating device. In both canaries and chaffinches, transsection of the hypoglossal branches supplying the right half of the syrinx hardly compromises song production whereas the same procedure at the contralateral side severely disrupts singing behaviour. However, other species exhibit right-sided dominance or fail to show any lateralization effects of syringeal innervation.

An elaborate hierarchy of brain structures controls vocal performance in birds (e.g. Nottebohm 1999). Besides the peripheral part of the song control pathway, there is some evidence for functional asymmetries at the cerebral level as well. Thus, unilateral right-hemisphere damage to the higher vocal centre, i.e. a forebrain nucleus giving rise to the principal motor pathway for song production, was found to compromise singing behaviour more than contralateral lesions in the zebra finch. Furthermore, perceptual functions such as song discrimination also seem to be characterized by hemispheric lateralization effects that may, however, vary from task to task (for a review see Catchpole and Slater 1995).

Bird song and human speech production show a variety of striking similarities, for example the shaping of vocalizations by sensory experience and the way in which a critical period of development influences learning abilities (Doupe and Kuhl 1999). It is less clear how far these constraints act upon human singing behaviour as well. To infer any analogies between the neural mechanisms of singing capabilities in birds and humans would, thus, be premature.

Primate species show a rich and elaborate repertoire of vocalizations, subserving a variety of communicative functions such as the display of affective states or the advertisement of predators. Alarm calls even may show a referential signature (Hauser 1996). However, songs in terms of long and uninterrupted vocal bouts have been observed in four primate genera only, *Indri, Tarsius, Callicebus*, and *Hylobates*, characterized by rather distant phylogenetic relationships to each other (Geissmann 2000). As a consequence, singing behaviour appears to have developed several times during primate evolution. Our nearest relatives among the 'singing' primates are the gibbons or lesser apes (*Hylobates*). Among others, these animals show a distinct type of locomotion, called brachiation, and produce loud morning vocalizations of an average duration of up to several tens of minutes, characterized by an inherited stereotypical acoustic structure. Mated pairs of most *Hylobates* species often engage in duets.

It has been proposed that singing behaviour in non-human primates always evolved from alarm or territorial calls because both kinds of vocalizations show commonalities in their acoustic structure and communicative functions. The same developmental pattern has also been applied to production of song within the hominid lineage (Geissmann 2000). However, vocal music performance of humans differs in several significant respects from the alarm and territorial calls of non-human primates. For example, melodies are usually based on a restricted sample of tones and singing is characterized by a steady beat or pulse. Thus, again, any speculations about common brain mechanisms of song production across primate species must be proposed with some caution.

## Brain structures subserving singing capabilities: clinical data obtained from subjects with acquired focal cerebral lesions

### Preserved vocal music performance in patients suffering from motor or global aphasia

Subjects with severely restricted speech production due to motor or global aphasia subsequent to a left-hemisphere lesion sometimes, nevertheless, show strikingly preserved musical capabilities, being able to sing the lyrics of overlearned songs in a rather fluent and well-articulated manner or at least to hum familiar tunes (e.g. Yamadori *et al.* 1977: 21 out of a total of 24 Broca aphasics). Within

the medical literature, Trousseau (1865), Falret (1867) and Hughlings Jackson (1871) first reported this kind of functional dissociation within the domain of human acoustic communication (references in Edgren 1895). In order to facilitate speech production, melodic intonation therapy, therefore, advises patients with acquired expressive language disorders to impose short phrases on 'simple, non-linguistically loaded melody patterns' (Albert *et al.* 1973). Aphasic syndromes may even be accompanied by compulsive singing and whistling activities (hypermusia; Jacome 1984). The latter behavioural abnormalities seem to occur in parallel with the recovery of verbal skills. Furthermore, even complete removal of the language-dominant hemisphere may fail to compromise a patient's ability to sing (Gott 1973; Burkland and Smith 1977). The right half of the brain has, thus, been assumed to predominantly subserve vocal music performance. As a consequence, damage to that hemisphere must be expected to disrupt singing capabilities.

## Cerebral localization of oral–expressive (vocal/motor) amusia: the role of the right-hemisphere fronto-insular cortex

Knoblauch (1888) suggested a classification of the 'disorders of musical capabilities subsequent to brain lesions' in strict parallel to the Wernicke–Lichtheim schema of aphasic syndromes, encompassing a constellation of motor amusia ('Kernamusie'). Benton (1977) called this syndrome oral–expressive or vocal amusia and restricted it to the 'loss of the capacity to sing, hum or whistle a tune'. Instrumental apraxia ('loss of the ability to play one's instrument despite the absence of appropriate motor defects') and musical agraphia ('inability to transcribe a series of heard notes') are considered separate varietes of amusia. In accordance with this taxonomy, more recent case studies have also noted a dissociation between vocal performance, on the one hand, and the ability to play an instrument (McFarland and Fortin 1982) or to write down melodies in musical notation (Brust 1980, case 1), on the other.

Mann (1898) was the first to perform a post-mortem brain examination in a case of 'vocal motor amusia'. The patient had not exhibited any signs of an aphasic disorder. Autopsy revealed a cystic lesion within the posterior half of the right-hemispheric middle frontal gyrus (F2), encroaching upon adjacent pre-central regions. A review by Henson (1985) identified just two further cases of oral–expressive amusia (Jossmann 1926, 1927: ligation of right internal carotid artery, rhythm uncompromised; Botez and Wertheim 1959: deficient intonation and rhythm during singing and whistling subsequent to right-hemisphere tumour, posterior third of F1 and F2). In addition, several clinical studies reported a loss of the ability to (re-)produce melodies subsequent to gunshot or shrapnel wounds associated with a lesion of posterior aspects of F2 and adjacent parts of the

pre-central gyrus as determined on the basis of the skull defect (Mendel 1916; Mann 1917; Foerster 1918; Rohardt 1919; references in Jossmann 1927).

However, Henson (1985) reviewed a further instance of expressive amusia embedded into a more widespread constellation of cognitive deficits (Pötzl and Uiberall 1937). The multiple bilateral lesions detected at post-mortem examination spared right-hemisphere frontal areas but extended to the ipsilateral insula. In line with this observation, more recent case studies of amusia syndromes affecting, among other things, vocal performance also revealed a lesion of opercular and insular areas of the right hemisphere sparing ipsilateral F2. Thus, Peretz *et al.* (1994, case C.N.) observed impaired discrimination and recognition of tunes concomitant with an inability to sing familiar melodies in a patient with bilateral temporal lesions extending to the right anterior insula and, to a limited extent, to the right inferior frontal gyrus (see Patel *et al.* 1998, case I.R., for a further example). And, Griffiths *et al.* (1997) were able to document, among other things, a receptive musical deficit concomitant with normal perception of environmental sounds and speech in a patient who had suffered a right-hemisphere temporoparietal infarction extending to the insular cortex. Most notably, the subject 'felt that his ability to sing had been impaired'. Rather than damage to a circumscript part of F2 and the adjacent pre-central gyrus, lesions within a more widespread network of intra- and anterior perisylvian regions may, thus, give rise to motor amusia.

As a further aetiological variant of oral–expressive amusia, Confavreux *et al.* (1992) observed a patient with focal degeneration of right-hemisphere perisylvian regions, including marked insular atrophy. Initially, the patient's complaints were restricted to a loss of the ability to sing or to hum even highly familiar tunes as well as a low, flat voice. Otherwise, clinical and neuropsychological examinations were, by and large, unremarkable (see Ghacibeh and Heilman 2003 for a further case report).

Besides damage to right-hemisphere pre-motor/opercular/insular cortex, lesions of the ipsilateral temporal lobe sporadically yielded predominant motor amusia (Brain 1941, case 3; Takeda *et al.* 1990). Even a tumour restricted to Heschl's gyrus was found to impair singing capabilities (Russell and Golfinos 2003). As a rule, however, lesions bound to temporal areas of the right or both sides appear to spare vocal performance giving rise to syndromes of 'pure' receptive amusia (see Henson 1985), to more subtle perceptual deficits such as impaired appreciation of timbre and compromised discrimination/recognition of tonal melodies (Zatorre 2003) or, finally, to various syndromes of verbal and non-verbal auditory agnosia (for a review see Ackermann and Mathiak 1999).

Since, as a rule, syndromes of 'pure' motor amusia are bound to right-sided frontal lesions, several authors assume that anterior parts of the 'minor' hemisphere subserve

singing capabilities. However, these deficits more frequently emerge in association with aphasic syndromes subsequent to a left-hemisphere lesion (Edgren 1895; Henschen 1926; Benton 1977). Therefore, Henschen (1926) assigned 'the power of singing' to the left third frontal convolution in front of Broca's area. In line with these suggestions, a recent clinical study found singing of novel tunes to be more severely impaired in patients who had suffered left-hemisphere cerebrovascular events than in subjects with contralateral lesions (Prior *et al.* 1990). By contrast, reproduction of familiar melodies turned out to be compromised in both groups. As an alternative to models of lateralized vocal music performance, either side of the brain has been assumed to mediate the generation of tunes (Kleist 1928).

### Summary

A considerable number of patients with motor aphasia subsequent to a left-hemisphere lesion show preserved vocal muscial performance, or at least spared generation of non-lyrical tunes. Nevertheless, oral–expressive amusia in association with right-sided cerebral disorders represents a rather rare event. As a consequence, therefore, singing behaviour must be characterized by less strict lateralization effects at the group level than speech production. Conceivably, the various lesion sites found to be associated with oral–expressive amusia each give rise to different pathomechanisms of disrupted song production. For example, impaired vocal performance could reflect degradation of tonal perception and memory systems in the case of damage to the right temporal lobe (Jellinek 1956, case 2; Ayotte *et al.* 2002). In some instances at least, impairments in musical expression following left-hemisphere lesions might primarily depend upon dysphasic deficits (Brust 1980, case 2) or compromised rhythm processing, a disorder more frequently observed in left-hemisphere dysfunctions (Jellinek 1933; Mavlov 1980). The few available cases of oral–expressive amusia subsequent to right-hemisphere damage must be expected to reflect, in view of the rarity of these constellations, an idiosyncratic rather than a general pattern of functional cerebral organization of human singing behaviour. As a consequence, clinical case studies are inherently limited with respect to any further elucidation of the brain's network underlying vocal musical performance.

## Brain structures subserving singing capabilities: experimental behavioural and electrophysiological data obtained from subjects without disorders of acoustic communication

### Unilateral transient inactivation of cerebral hemispheres (Wada test)

Application of sodium amylobarbitone (amobarbital) into either internal carotid artery gives rise to rather instantaneous, but transient functional depression of

the ipsilateral cerebral hemisphere for a period of 3 to 5 min (Wada test). By contrast, the contralateral half of the brain remains uncompromised. Besides speech/language functions, this procedure has also been used for the determination of hemispheric lateralization of singing capabilities. In seven out of a group of eight epileptic patients considered for cerebral commissurotomy, Gordon and Bogen (1974) observed monotonous song production 'devoid of correct pitch rendering' following a right-sided Wada test. Subjects had been asked to complete a tune on the syllable 'la' practised prior to the injection of amobarbital. Rhythm turned out to be much less disrupted. In contrast to the production of speech, vocal music performance was found to be largely preserved after left-carotid application of the narcotic drug. Based on these observations, a dissociation of singing and speaking functions with respect to hemispheric lateralization was assumed.

## Behavioural laterality measures (time-sharing interference, lip-opening asymmetries) of singing capabilities

Within some limits, functional interference between two concurrent sensorimotor or cognitive activities depends upon the topographic relationships of the underlying cerebral control structures ('time-sharing interference paradigm'). Thus, a task characterized by well-established functional brain asymmetry, e.g. finger tapping, may give some clues about the hemispheric lateralization effects of vocal music performance. Based on this approach, a variety of studies have reported singing to predominantly depend upon the right hemisphere or to be bound to bilateral processes of a rather symmetrical distribution (for a review see Hough *et al.* 1994). Extending these earlier investigations, Hough *et al.* (1994) coupled finger tapping to three vocal tasks, i.e. speaking (continuous repetition of a short sentence), singing a rote song ('Happy birthday to you') as well as ascending/descending production of a diatonic five-note scale ('verbal/manual time-sharing design'). Finger tapping in the absence of any (overt) vocalizations served as a baseline. Comparison of task performance across conditions allows for the calculation of a laterality score of singing and speaking capabilities, respectively. As expected, speaking predominantly recruited left-hemisphere resources whereas, by contrast, singing appeared to involve bilateral interactions of the brain or to be primarily bound to right-sided structures. Most notably, however, performance significantly depended upon gender: females exhibited more pronounced side differences during sentence recitation compared with the two singing tasks whereas rote song elicited the largest lateralization effects in males. In contrast to these findings, a preceding study of exclusively male subjects based on another time-sharing interference paradigm, i.e. balancing a dowsing rod on the left or right finger, yielded discrepant data: humming selectively compromised

the right-handed task, indicating more salient left-hemisphere processing of this kind of vocal performance (Hicks 1975).

Measurements of lip-opening asymmetries by, for example, single-frame photographic analysis provide a further means to assess the hemispheric lateralization effects of motor control. Singing a well-known nursery rhyme as well as reciting its lyrics yielded enlarged right-sided oral movements, indicating predominant left-hemisphere control of these activities ('right mouth advantage'; Cadalbert *et al.* 1994). By contrast, production of the respective melody on the syllable 'ma', i.e. in the absence of the song's words, failed to elicit any mouth-opening asymmetries. Thus, both cerebral hemispheres appear to participate in the generation of tunes.

## Electrophysiological recordings during musical tasks

Apart from direct recordings at the cortical surface during, for example, surgical interventions, electro- (EEG) or magnetoencephalographic (MEG) techniques represent the most direct approach available for investigating the neural correlates of distinct sensorimotor or cognitive tasks in humans. Measurements of evoked electrical potentials or magnetic field strengths at the scalp are a widely used approach to the study of the various components of speech and music perception. By contrast, rather few electrophysiological studies addressing singing capabilities are available so far. Among other things, functional lateralization effects give rise to cerebral asymmetries in EEG alpha power in terms of reduced activity within the 8–12 Hz frequency band over the hemisphere more engaged in a given task. Using this approach, McKee *et al.* (1973) first reported lateralization effects towards the right side during music performance. However, a later study failed to detect any EEG alpha asymmetries during production of highly familiar tunes overlearned during childhood on the syllable 'la' (Galin *et al.* 1982). Conceivably, the extent of functional asymmetry during vocal musical tasks depends upon preceding training and, thus, processing strategies: enhanced right-hemisphere activation during whistling the melody of a song (versus recitation of its lyrics) was found to be restricted to musically unsophisticated participants (e.g. Davidson and Schwartz 1977).

## Summary

Several behavioural studies point at a predominant contribution of the right hemisphere to vocal music performance in humans. However, these lateralization effects appear to be less robust or pronounced than functional asymmetry of speech production. In addition, right but not left-sided Wada tests, as a rule, interrupt singing behaviour in human subjects. *Actual* vocal music performance, therefore, predominantly must depend upon the 'minor' hemisphere. Persistent

damage to right-sided cerebral structures then could give rise to functional recovery via the contralateral side since the available clinical data indicate that in principle both cerebral hemispheres are able to subserve the generation of tunes (see the last section). By contrast, the time span of the Wada test might be too short for any compensation processes to emerge.

## Haemodynamic functional brain imaging of singing capabilities: opposite hemispheric lateralization effects at the level of the anterior insula

Local variations of neural activity give rise to an 'automatic' increase of regional cerebral blood flow (neurovascular coupling). Thus, registration of these haemodynamic effects should provide a feasible means for the identification of the neural structures engaged in distinct sensorimotor or cognitive tasks. The two or so early investigations based on the xenon clearance technique failed to detect any significant inter- or intra-hemispheric changes in cerebral blood flow during singing and/or humming (Ryding *et al.* 1987; Formby *et al.* 1989). Positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) represent the two most important brain imaging techniques based on neurovascular coupling mechanisms available today. So far, there has been only a single PET study aimed at the delineation of the cerebral network subserving vocal music performance (Perry *et al.* 1999). Participants were asked to repetitively produce vowel /ä/ across a time interval of 60 s at a pre-specified target pitch, pausing for breath as needed. In contrast to passive listening to complex tones presented at a similar rate and transposed to the same frequency range, this task elicited, as expected, haemodynamic responses of various structures bound to motor control (mesiofrontal regions, bilateral face area of the pre-central gyrus, both sides of the insula, and the cerebellum). Two asymmetries could be noted within these areas:

(1) higher activation of the right pre-central gyrus compared with the contralateral side,

(2) bilateral responses at the level of the intrasylvian cortex, but at slightly different locations.

Repetitive vowel productions at the same pitch address some kind of 'rudimentary singing' (Perry *et al.* 1999) and, thus, do not necessarily represent a valid probe of the cerebral network underlying the generation of melodies. A recent fMRI investigation carried out by our group asked subjects to reproduce a tune drawn from a serenade on the syllable 'la' [1.5 T whole-body scanner (Siemens Vision), echo-planar imaging (EPI) sequence across the complete brain volume, anatomic reference = $T_1$ weighted 3D turbo-flash sequence, $n$ = 18 healthy

right-handed speakers (nine females) of standard German, age 22–63 yrs; Riecker *et al.* 2000]. Non-lyrical musical material was selected in order to avoid any retrieval of propositional data structures during functional imaging. This task served as a control condition for an experiment that aimed at the determination of the cerebral network mediating speech motor processes based on the continuous recitation of the months of the year (automatic speech; e.g. Ryding *et al.* 1987). Rather than repetition of single lexical items, highly overlearned word strings can be considered a paradigm of fluent verbal output, thus being more similar to melody production. Furthermore, these test materials should elicit a rather monotonous mode of speech production and thus focus on segmental aspects of verbal utterances and avoid the eventual confounding influences of intonational patterns. Since motor imagery has been assumed to provide a 'window' into the mechanisms of action planning, i.e. preparatory activities preceding actual performance of movement sequences (Jeannerod 1994), both overt (loud) and covert (silent) task performance was considered for analysis. Subtraction of the haemodynamic activation patterns bound to these two conditions was expected to allow for the isolation of the brain structures bound to motor execution processes.

Subjects performed speaking and singing at their habitual tempo, and were asked to refrain from verbal thought during the rest periods in between. Each task was applied 12 times in a counter-balanced order (onset-to-onset intervals between activation phases = 24 s). Production rates of speaking and singing had been determined during the training phase (overt generation of the word string 'January–December': mean = 5.6 s, standard deviation = 1.3 s; overt production of part of a serenade: mean = 6.9 s, standard deviation = 1.1 s). Among other things, a cognitive subtraction procedure was performed: overt speech versus overt singing and vice versa. Analysis of the group data revealed rather bilateral activation at the level of the pre-central gyrus and the cerebellum both during speaking and singing concomitant with (non-significantly) higher responses at the right (singing; see above and Perry *et al.* 1999) or left side (speaking). As a new aspect of functional brain asymmetry, opposite lateralization effects emerged within the intrasylvian cortex (speaking aloud = left anterior insula, singing aloud = right anterior insula; Plate 12). This pattern of insular activation turned out to be preserved after subtraction of the respective haemodynamic responses obtained during the silent modes of speaking and singing (overt versus covert speech, overt versus covert singing). Thus, any differences between overt and covert performance do not just reflect threshold effects. Furthermore, the robust double dissociation at the level of the anterior insula was confirmed by the second cognitive subtraction procedure (singing aloud minus speaking aloud and vice versa).

Based on neuroradiological data obtained from a large sample of patients, Dronkers (1996) found the anterior insula to represent the area of maximum overlap of lesion sites in apraxia of speech. This constellation is characterized by effortful and groping articulatory movements, dysprosody, inconsistent distortions of speech sounds across repeated productions, and obvious difficulties initiating verbal utterances. However, non-speech functions of laryngeal and orofacial muscles are found to be unimpaired. Several preceding PET studies documented predominantly left-sided activation of the intrasylvian cortex during naming (line drawings depicting concrete objects), reading, and repetition (auditorily or visually displayed nouns) tasks (e.g. Wise *et al.* 1999; for a review see Ackermann *et al.* 2004). Considering these findings, left-sided intrasylvian structures have been assumed to operate as a centre of speech motor coordination. In a similar vein, activation of the right-hemisphere anterior insula during reproduction of a non-lyrical tune might reflect the sequencing of vocal tract musculature in association with vocal music performance.

## Cerebral substrates of musical imagery

Besides listening to a live or recorded performance, music including songs can be re-experienced by imagining it in one's head (musical imagery). Imaginal and perceptual processes show a broad overlap of the respective underlying neural substrates at the level of the central auditory system. Thus, Zatorre and Halpern (1993) performed a pitch discriminaton task in patients who had undergone excision of either the left- or right-sided temporal lobe for the control of medically intractable epilepsy. As a first task, subjects heard a familiar song over headphones, concomitant with a written version of its text on a screen, and were asked to judge whether the second of two capitalized lyrics was higher or lower in pitch than the first one. During a subsequent condition, no song was presentend and the paricipants had to generate an auditory image of the tune, everything else being equal. Most notably, patients with right-sided lesions performed significantly worse on both tasks than subjects suffering from left hemisphere temporal damage or normal controls. Using similar test materials, a subsequent PET study further specified the neural network associated with both music perception and musical imagery: (the bilateral secondary auditory regions, both frontal lobes, the left parietal corex, and the supplementary motor area; Zatorre *et al.* 1996). Subtraction of the haemodynamic responses bound to the two tasks from each other revealed significantly increased activation of bilateral inferior frontopolar regions, the right posterior thalamus, and the subcallosal gyrus during the imagery condition as compared with the discrimination of heard pitch. Conceivably, this circuit is bound to the retrieval of auditory information from memory stores. Using non-verbal tunes as test materials, a subsequent PET study

found lateralization effects towards the right side within the temporal lobe, frontal areas, and thalamus (Halpern and Zatorrre 1999). Functional brain asymmetry, therefore, emerges once words have been removed from imagery tasks.

Besides overt speaking and singing, the fMRI study of our group referred to above also included covert performance of the same tasks, i.e. conditions which also represent some kind of 'auditory imagery' (see Ivry and Fiez 1999). As compared with baseline (rest), activation of two complementary cerebral networks emerged (covert reproduction of a non-lyrical tune = right pre-central gyrus and left cerebellar hemisphere, 'inner speech' = the opposite response pattern). Most notably, no significant activation of the intrasylvian cortex could be detected. These findings are in line with two preceding investigations that included vertical non-speech tongue movements as a control conditions (Wildgruber *et al.* 1996; Ackermann *et al.* 1998). Based on these data, it could be ascertained that frontal haemodynamic activation during silent speaking and singing is bound to vocal tract motor areas (Plate 12).

The studies by Halpern and Zatorre (Zatorre *et al.* 1996; Halpern and Zatorre 1999) addressing the discrimination of melodies had failed to detect significant haemodynamic activity of the sensorimotor cortex during musical imagery. Since our fMRI study focused on the *generation* of 'inner' speech and singing, respectively, areas engaged in motor execution might have been recruited as well or to a larger extent. The cerebral networks bound to processes of imagination might thus depend, within some limits, upon the demands of the task.

## Summary: vocal music performance and the right hemisphere

Based on a variety of clinical and functional imaging data, Zatorre (e.g. Zatorre 2003) assumed the right hemisphere to have a higher proficiency in pitch processing. Our fMRI data indicate that lateralized pitch representation 'spills over' to cortical areas engaged in motor control during singing:

1. Reproduction of a familiar non-lyrical tune yielded right-sided activation at the level of the anterior insula concomitant with bilateral haemodynamic responses of the pre-central gyus and cerebellar hemispheres whereas speaking elicited an opposite lateralization effect within the intrasylvian cortex.
2. Silent reproduction of the same tune (musical imagery) resulted in a more limited response pattern, encompassing the right sensorimotor cortex and left cerebellum, but sparing intrasylvian areas (silent recitation of word strings = opposite activation pattern).

Inner speech is considered to be a pre-articulatory, but otherwise fully parsed, representation of sound structure (for further details see Ackermann *et al.* 2004).

Therefore, covertly produced verbal utterances must be expected to have a temporal organization similar to their spoken counterparts. Considering our fMRI data, left motor cortex and right cerebellar hemisphere (both structures are interconnected by reciprocal fibre systems) seem to support these pre-articulatory motor control processes. Based on clinical data obtained from patients with speech motor deficits and on fMRI data registered during syllable repetition and speech perception tasks, this network might support the sequencing of mental syllables into a smooth, i.e. co-articulated, representation of verbal utterances at a speaker's habitual speaking tempo (Ackermann *et al.* 2004). In some analogy to covert speech production, 'inner singing' can be considered a pre-articulatory, but otherwise fully parsed, representation of the tonal and rhythmic organization of a tune. Whereas inner speech requires the sequencing of retrieved mental syllables, the observed right-motor/left-cerebellar network bound to 'inner singing' might underlie the specification of melody templates retrieved from a right-hemisphere storage site with respect to pitch level and duration of its tonal elements.

Haemodynamic activation of the anterior insula was observed during overt task performance only. Thus, the intrasylvian cortex seems to operate 'downstream' of the motor cortex/cerebellar networks bound to inner speech and musical imagery, and might support the transformation of pre-articulatory/pre-phonatory verbal and musical codes into an elaborate spatiotemporal pattern of motor commands driving the muscles of the vocal tract. Insular regions are involved in the coordination of midline musculature (head, neck, trunk) during respiration and swallowing. Conceivably, speaking and singing capabilities have been superimposed during the course of language evolution on these phylogenetically older motor activities via intrasylvian structures (Ackermann and Riecker 2004).

Apart from functional imaging data, a variety of clinical and experimental data indicate singing capabilities to be predominantly bound to the right hemisphere. In comparison with speech production, these lateralization effects are less robust by far (see above). Compared with expressive speech/language functions, vocal music performance must thus be characterized by more 'fluid' hemispheric lateralization effects concomitant with significantly better possibilities of recovery.

## References

**Ackermann, H. and Mathiak, K.** (1999). Symptomatology, neuroanatomic correlates and pathomechanisms of central hearing disorders (pure word deafness, verbal/nonverbal auditory agnosia, cortical deafness): a review. *Fortschr Neurol Psych*, **67**, 509–523 (in German).

**Ackermann, H. and Riecker, A.** (2004). The contribution of the insula to motor aspects of speech production: a review and a hypothesis. *Brain Lang*, **89**, 320–328.

**Ackermann, H., Wildgruber, D., Daum, I., and Grodd, W.** (1998). Does the cerebellum contribute to cognitive aspects of speech production? A functional magnetic resonance imaging (fMRI) study in humans. *Neurosci Lett*, **247**, 187–190.

**Ackermann, H., Mathiak, K., and Ivry, R.B.** (2004). Temporal organization of 'internal speech' as a basis for cerebellar modulation of cognitive functions. *Behav Cogn Neurosci Rev*, **3**, 14–22.

**Albert, M.L., Sparks, R.W., and Helm, N.A.** (1973). Melodic intonation therapy for aphasia. *Arch Neurol*, **29**, 130–131.

**Ayotte, J., Peretz, I., and Hyde, K.** (2002). Congenital amusia: a group study of adults afflicted with a music-specific disorder. *Brain*, **125**, 238–251.

**Benton, A.L.** (1977). The amusias. In: M. Critchley and R.A. Henson (eds) *Music and the brain*, pp. 378–397. William Heinemann, London.

**Botez, M.I. and Wertheim, N.** (1959). Expressive aphasia and amusia following right frontal lesion in a right-handed man. *Brain*, **82**, 186–202.

**Brain, W.R.** (1941). Visual disorientation with special reference to lesions of the right cerebral hemisphere. *Brain*, **64**, 244–272.

**Brown, S.** (2000). The 'musilanguage' model of music evolution. In: N.L. Wallin, B. Merker, and S. Brown (eds) *The origins of music*, pp. 271–300. MIT Press, Cambridge, MA.

**Brust, J.C.M.** (1980). Music and language: musical alexia and agraphia. *Brain*, **103**, 367–392.

**Burkland, C.W. and Smith, A.** (1977). Language and the cerebral hemispheres: observations of verbal and nonverbal responses during 18 months following left ('dominant') hemispherectomy. *Neurology*, **27**, 627–633.

**Cadalbert, A., Landis, T., Regard, M., and Graves, R.E.** (1994). Singing with and without words: hemispheric asymmetries in motor control. *J Clin Exp Neuropsychol*, **16**, 664–670.

**Catchpole, C.K. and Slater, P.J.B.** (1995). *Birdsong: biological themes and variations.* Cambridge University Press, Cambridge, UK.

**Confavreux, C., Croisile, B., Garassus, P., Aimard, G., and Trillet, M.** (1992). Progressive amusia and aprosody. *Arch Neurol*, **49**, 971–976.

**Darwin, C.R.** (1981/1871). *The descent of man, and selection in relation to sex* (photoreproduction of the 1871 edition). Princeton University Press, Princeton, NJ.

**Davidson, R.J. and Schwartz, G.E.** (1977). The influence of musical training on patterns of EEG asymmetry during musical and non-musical self-generation tasks. *Psychophysiology*, **14**, 58–63.

**Doupe, A.J. and Kuhl, P.K.** (1999). Birdsong and human speech: common themes and mechanisms. *Ann Rev Neurosci*, **22**, 567–631.

**Dronkers, N.F.** (1996). A new brain region for coordinating speech articulation. *Nature*, **384**, 159–161.

**Edgren, J.G.** (1895). Amusie (musikalische Aphasie). *Deut Z Nervenheilk*, **6**, 1–64.

**Falk, D.** (2000). Hominid brain evolution and the origins of music. In: N.L. Wallin, B. Merker, and S. Brown (eds) *The origins of music*, pp. 197–216. MIT Press, Cambridge, MA.

**Formby, C., Thomas, R.G., and Halsey, J.H. Jr** (1989). Regional cerebral blood flow for singers and nonsingers while speaking, singing, and humming a rote passage. *Brain Lang*, **36**, 690–698.

**Galin, D., Ornstein, R., Herron, J., and Johnstone, J.** (1982). Sex and handedness differences in EEG measures of hemispheric specialization. *Brain Lang*, **16**, 19–55.

**Geissmann, T.** (2000). Gibbon songs and human music from an evolutionary perspective. In: N.L. Wallin, B. Merker, and S. Brown (eds) *The origins of music*, pp. 103–123. MIT Press, Cambridge, MA.

**Ghacibeh, G.A. and Heilman, K.M.** (2003). Progressive affective aprosodia and prosoplegia. *Neurology*, **60**, 1192–1194.

**Gordon, H.W. and Bogen, J.E.** (1974). Hemispheric lateralization of singing after intracarotid sodium amylobarbitone. *J Neurol Neurosurg Psych*, **37**, 727–738.

**Gott, P.S.** (1973). Language following dominant hemispherectomy. *J Neurol Neurosurg Psych*, **36**, 1082–1088.

**Griffiths, T.D., Rees, A., Witton, C., Cross, P.M., Shakir, R.A., and Green, G.G.R.** (1997). Spatial and temporal auditory processing deficits following right hemisphere infarction: a psychophysical study. *Brain*, **120**, 785–794.

**Halpern, A.R. and Zatorre, R.J.** (1999). When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. *Cereb Cortex*, **9**, 697–704.

**Halpern, A.R.** (2003). Cerebral substrates of musical imagery. In: I. Peretz and R.J. Zatorre (eds) *The cognitive neuroscience of music*, pp. 217–230. Oxford University Press, Oxford, U.K.

**Hauser, M.D.** (1996). *The evolution of communication*. MIT Press, Cambridge, MA.

**Henschen, S.E.** (1926). On the function of the right hemisphere of the brain in relation to the left in speech, music and calculation. *Brain*, **49**, 110–123.

**Henson, R.A.** (1985). Amusia. In: J.A.M. Frederiks (ed.) *Handbook of clinical neurology, vol. 1(45): Clinical neuropsychology*, pp. 483–490. Elsevier, Amsterdam.

**Hicks, R.E.** (1975). Intrahemispheric response competition between vocal and unimanual performance in normal adult human males. *J Comp Physiol Psychol*, **89**, 50–60.

**Hough, M.S., Daniel, H.J., Snow, M.A., O'Brien, K.F., and Hume, W.G.** (1994). Gender differences in laterality patterns for speaking and singing. *Neuropsychologia*, **32**, 1067–1078.

**Ivry, R.B. and Fiez, J.A.** (2000). Cerebellar contributions to cognition and imagery. In: M.S. Gazzaniga (ed) *The new cognitive neurosciences*, 2nd edn, pp. 999–1011. MIT Press, Cambridge, MA.

**Jacome, D.E.** (1984). Aphasia with elation, hypermusia, musicophilia and compulsive whistling. *J Neurol Neurosurg Psych*, **47**, 308–310.

**Jeannerod, M.** (1994). The representing brain: neural correlates of motor intention and imagery. *Behav Brain Sci*, **17**, 187–245.

**Jellinek, A.** (1933). Zur Phänomenologie der Amusie: expressive Amusie und Aphasie eines Lautensängers. *Jahrb Psych Neurol*, **50**, 115–141.

**Jellinek, A.** (1956). Amusia: on the phenomenology and investigation of central disorders of the musical functions. *Folia Phoniatr*, **8**, 124–149.

**Jossmann, P.** (1926). Motorische Amusie. *Zbl Neurol Psych*, **44**, 260.

**Jossmann, P.** (1927). Die Beziehungen der motorischen Amusie zu den apraktischen Störungen. *Monatsschr Psych Neurol*, **63**, 239–274.

**Kleist, K.** (1928). Gehirnpathologische und lokalisatorische Ergebnisse über Hörstörungen, Geräuschtaubheiten und Amusien. *Monatsschr Psych Neurol*, **68**, 853–860.

**Knoblauch, A.** (1888). Ueber Störungen der musikalischen Leistungsfähigkeit infolge von Gehirnläsionen. *Deut Arch Klin Med*, **43**, 331–352.

**Mann, L.** (1898). Ein operierter Fall von traumatischer Herderkrankung des rechten Frontallappens mit Sectionsbefund (vocale motorische Amusie). *Monatsschr Psych Neurol*, **4**, 369–378.

**Mavlov, L.** (1980). Amusia due to rhythm agnosia in a musician with left hemisphere damage: a non-auditory supramodal defect. *Cortex*, **16**, 331–338.

**McFarland, H.R. and Fortin, D.** (1982). Amusia due to right temporoparietal infarct. *Arch Neurol*, **39**, 725–727.

**McKee, G., Humphrey, B., and McAdam, D.W.** (1973). Scaled lateralization of alpha activity during linguistic and musical tasks. *Psychophysiology*, **10**, 441–443.

**Nottebohm, F.** (1999). The anatomy and timing of vocal learning in birds. In: M.D. Hauser and M. Konishi (eds) *The design of animal communication*, pp. 63–110. MIT Press, Cambridge, MA.

**Patel, A.D., Peretz, I., Tramo, M., and Labreque, R.** (1998). Processing prosodic and musical patterns: a neuropsychological investigation. *Brain Lang*, **61**, 123–144.

**Peretz, I., Kolinsky, R., Tramo, M., Labreque, R., Hublet, C., Demeurisse, G., and Belleville, S.** (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain*, **117**, 1283–1301.

**Perry, D.W., Zatorre, R.J., Petrides, M., Alivisatos, B., Meyer, E., and Evans, A.C.** (1999). Localization of cerebral activity during simple singing. *NeuroReport*, **10**, 3979–3984.

**Pötzl, O. and Uiberall, H.** (1937). Zur Pathologie der Amusie. *Wien Klin Wochenschr*, **50**, 770–775.

**Prior, M., Kinsella, G., and Giese, J.** (1990). Assessment of musical processing in brain-damaged patients: implications for laterality of music. *J Clin Exp Neuropsychol*, **12**, 301–312.

**Richman, B.** (1993). On the evolution of speech: singing as the middle term. *Curr Anthropol*, **34**, 721–722.

**Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G., and Grodd, W.** (2000). Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *NeuroReport*, **11**, 1997–2000.

**Russell, S.M. and Golfinos, J.G.** (2003). Amusia following resection of a Heschl gyrus glioma: case report. *J Neurosurg*, **98**, 1109–1112.

**Ryding, E., Bradvik, B., and Ingvar, D.H.** (1987). Changes of regional cerebral blood flow measured simultaneously in the right and left hemisphere during automatic speech and humming. *Brain*, **110**, 1345–1358.

**Suthers, R.A.** (1999). The motor basis of vocal performance in songbirds. In: M.D. Hauser and M. Konishi (eds) *The design of animal communication*, pp. 37–62. MIT Press, Cambridge, MA.

**Takeda, K., Bandou, M., and Nishimura, Y.** (1990). Motor amusia following a right temporal lobe hemorrhage: a case report. *Rinsho Shinkeigaku*, **30**, 78–83.

**Wildgruber, D., Ackermann, H., Klose, U., Kardatzki, B., and Grodd, W.** (1996). Functional lateralization of speech production at primary motor cortex: a fMRI study. *NeuroReport*, **7**, 2791–2795.

**Wise, R.S.J., Green, J., Büchel, C., and Scott, S.K.** (1999). Brain regions involved in articulation. *Lancet*, **353**, 1057–1061.

**Yamadori, A., Osumi, Y., Masuhara, S., and Okubo, M.** (1977). Preservation of singing in Broca's aphasia. *J Neurol Neurosurg Psych*, **40**, 221–224.

**Zatorre, R.J.** (2003). Hemispheric asymmetries in the processing of tonal stimuli. In: K. Hugdahl and R.J. Davidson (**eds**) *The asymmetrical brain*, pp. 411–440. MIT Press, Cambridge, MA.

**Zatorre, R.J. and Halpern, A.R.** (1993). Effect of unilateral temporal-lobe excision on perception and imagery of songs. *Neuropsychologia*, **31**, 221–232.

**Zatorre, R.J., Halpern, A.R., Perry, D.W., Meyer, E., and Evans, A.C.** (1996). Hearing in the mind's ear: a PET investigation of musical imagery and perception. *J Cogn Neurosci*, **8**, 29–46.

Chapter 14

# Sensory–motor networks in singing and speaking: a comparative approach

Reyna Leigh Gordon, Amélie Racette, and Daniele Schön

Sheshell ebb music wayriver she flows
James Joyce, *Finnegan's Wake*

## Introduction

The comparative analysis of different cognitive functions, using different investigative techniques whenever possible, provides a more powerful means for understanding how the brain processes information and computes operations than studying either domain alone. Indeed, concentration on a single brain region and a single cognitive function with one single investigation method could yield an incomplete view of the functioning of the brain. As noted by Bates (1994), this type of approach most often calls for domain specificity and, as a consequence, assumes a modular explanation of the brain function in question. By coupling the study of differences with the study of similarities, a comparative approach may provide a more complete picture of which computations are specific and which ones rely on more general cognitive principles.

The choice of comparing music and language is far from being arbitrary (see also reviews by Besson and Schön (2001) and Patel and Peretz (1997)). Music, like language, is a universal trait specific to humans. Both require analysis of the spectrum and frequency of sound, streaming, sequencing, extracting regularities, memorization, attentional focus, and learning. It is not by chance that music and language combine beautifully in song!

### From perception to production

Most of the literature on music and speech has been centred on perception; production has received much less attention. This can mostly be explained in two ways. First, it is much easier to measure the perception of the listener than the production of the performer. Think, for example, how easily one can measure

pitch or phoneme discrimination (e.g. Is [ba] different from [da]?) but how difficult it is to record, analyse, and quantify the difference between two movements and the sounds they generate. In addition most brain imaging techniques are very sensitive to movements, which tend to cause artefacts in the data.

The second reason is more theoretical. Music and speech performance require the translation of a mental representation into action, such as playing an instrument, singing, or speaking. Moreover, performance is subject to frequent and controlled variations in acoustic parameters, such as intensity, spectral, and temporal changes, from one performance to another. In other words, the plan required to play or speak can be realized in an infinite number of ways. This diversity of outputs, particular to each individual and situation, closely resembles the diversity of input from which a person can recognize the same melody or words.

Even in the case of dissociations between comprehension and production deficits in patients with brain lesions, it is now considered that the two systems are not physically separate; rather, a central system is at work, in which the operations apply at both the access and retrieval levels (Semenza, 1999; Rapp, 2001). Thus, similar constraints apply to theories of comprehension and production (Schön *et al.* 2004).

The study of music and speech production is also conducive to obtaining a better understanding of the essence of musical and linguistic structures. Although it is commonly believed that the structural characteristics of music and language exist independently of the output systems, this is not necessarily the case. On the contrary, the extent to which movements constrain musical syntax and linguistic phonology is considerable. For example, the fact that small intervals are easier to sing (and play on most instruments) heavily influences the melodic structure of music, in that small intervals are extremely frequent while large intervals occur much less often. Similarly, in language, the fact that certain sounds are more difficult to pronounce than others (e.g. /lgma/ versus /ma/), influences the concatenation of different phonemes to form words and sentences. Therefore, the motor systems allowing the production of music and speech enjoy a privileged role in shaping the structure of music and language, and the study of the production of speech and music can thus be important for the study of the origins of music and language.

## Why study song cognition?

One approach to comparing the neural bases of language and music is through the use of song, which is a unique combination of these two cognitive domains. In song, language and music are merged into *one* acoustic signal with *two* salient dimensions. By manipulating either the linguistic or the musical dimension of

song, a unitary stimulus, the study of the relationship between these two dimensions may reveal important information about the neural networks underlying language and music cognition. Furthermore, singing is a universal human activity that can be used not only to compare language and music but also to gain insight into the one or the other in diverse contexts.

Despite the fact that song provides an ideal means to compare how the brain processes language and music, it has been relatively little exploited by cognitive neuroscientists until recently (Peretz *et al.* 2004*a*). In this chapter we will cite several studies that have investigated the cognition of song and discuss the findings relevant to the comparison of language and music processing. We seek to show how these studies have addressed the interaction or independence of language and music, the modulation of attention, the functional mapping of brain areas involved in singing, memorization processes, and production of song in patients with language disorders. The intrinsic limits of the model and directions for future research will be discussed in the conclusion.

## Relationship between the linguistic and musical dimensions of song

While some studies of the cognition of song use an approach that aims to study the linguistic and musical aspects of song separately by comparing normal speech and vocalises (Gordon and Bogen, 1974; Ryding *et al.* 1987; Riecker *et al.* 2000) or sung language and spoken language (Cadalbert *et al.* 1994; Epstein *et al.* 1999; Jeffries *et al.* 2003; Stager *et al.* 2003), other studies compare how the musical and linguistic dimensions of song are processed, by manipulating these dimensions as they occur in sung language only (Bartholomeus 1974; Besson *et al.* 1998; Bigand *et al.* 2001; Bonnel *et al.* 2001; Lidji 2003; Poulin-Charronnat *et al.* 2005). On the one hand, evidence for independent processing of language and music in song suggests that separate neural substrates are responsible for processing these two different types of information. On the other hand, evidence for interactive processing between the linguistic and musical dimensions provides support for the idea that common neural resources are utilized to process language and music.

For instance, interactions between these two dimensions in song have been observed by Bigand *et al.* (2001), who showed that a subtle change in harmonic processing (from tonic to subdominant) interferes with phoneme monitoring in the perception of choral music, sung with non-words. Similar musical material, this time sung with sentences, was then used to demonstrate that harmonic processing also interferes with semantic priming, measured by reaction time and error rates in a lexical decision task (Poulin-Charronnat *et al.* 2005). These results support the theory that language and music in song are intertwined,

especially in terms of implicit processing, and are thereby interactively processed. Additional evidence for interactive processing between the linguistic and musical dimensions of song has been found for other levels of processing, such as between phonology and melodic intervals in sung non-words (Lidji 2003).

However, evidence for independent processing was found by Besson *et al.* (1998) in an event-related brain potential (ERP) study that aimed to study the relationship of text to music in the perception of opera excerpts, sung a capella. Results showed distinct ERP components for semantic (N400) and harmonic (P300) violations. Moreover, detailed analysis revealed that an additive model of semantic and harmonic processing accounted very well for the observed effects, leading the authors to propose that lyrics and melody in song are processed independently. Further support for independence of semantics and harmony comes from another experiment using the same stimuli, in which a dual-task paradigm was utilized to show there was no decrease in performance when simultaneously attending to language and music in song perception, compared with attending to only one dimension at a time (Bonnel *et al.* 2001).

## Attention and levels of processing in song

In view of the number of seemingly contradictory results on the status of language and music in song, it is important to note that since each of these cognitive faculties encompasses several levels of processing (in language: phonology, semantics, syntax, pragmatics; in music: rhythm, melody, harmony, timbre), any experimental results obtained on the relationship between language and music in song may depend on which levels of each system are being examined. While some of the subprocesses involved in the cognition of song may be localized and even lateralized according to the same trends that characterize the perception of speech and music (e.g. spectral characteristics processed in the right auditory cortex and temporal characteristics in the left; see Zatorre *et al.* (2002)), other subprocesses (e.g. working memory) may take place in neuronal networks that are widely distributed. The dynamic combination of all the levels of processing involved in song could therefore account for interactions between certain levels, while other levels may be relatively autonomous.

There is a growing body of evidence indicating that the relationship between language and music in song cognition is intrinsically dependent on attentional processes, as can be seen from the occurrence of task-related effects caused by a modulation of attention between the linguistic and musical dimensions of song.

This point has been illustrated in the past through the use of dichotic listening in order to investigate lateralization effects related to the music or linguistic dimensions of song (Bartholomeus 1974; Goodglass and Calderon 1977). In fact, the allocation of attentional resources to different dimensions of a sung stimulus

could be one of the most important factors in determining the relationship between language and music in song. In addition to ERP evidence for the effects of attentional load on lower-level auditory processing (Alain and Izenberg 2003), an ERP study by Levy *et al.* (2003) further demonstrates the importance of attention and task demands in determining how neural resources are allocated to process musical and linguistic/vocal stimuli. Using an oddball paradigm, listeners were first instructed to respond when they heard a piano tone presented among a series of sustained pitches played by various instruments or sung by different voices, on a neutral vowel. A voice-specific response (VSR) was observed in the ERPs as a positive component, peaking at 320 ms (P320), which was found in response only to sung tones even when subjects were attending a piano sound (Levy *et al.* 2001). Interestingly, the VSR occurred when participants were paying attention to timbre but not when they were paying attention to another acoustic parameter, duration (Levy *et al.* 2003). Since neuroimaging studies have outlined the functional anatomy of human voice-selective areas in the superior temporal sulcus (see Belin *et al.* (2004) for a review), these recent studies by Levy *et al.* indicate that the voice-specific effects observed are task-dependent and partially reflect the allocation of attention in a system that is somewhat specialized in the identification of human voices.

## Functional mapping of brain structures involved in song processing

Understanding the nature of the relationship between the linguistic and musical dimensions of song should give an insight into whether music and language are processed by separate or integrated brain structures. Recent brain imaging studies of song production have attempted to shed light on the neural bases of language and music in song; in particular, certain patterns of localization have been uncovered which coincide well with the literature on language and music processing (Tervaniemi *et al.* 2000; Zatorre *et al.* 2002).

Some new evidence for dominance of the right hemisphere in singing comes from studies conducted with transcranial magnetic stimulation (TMS), a technique that provokes temporary reversible lesions of localized cortical zones (Epstein *et al.* 1999; Stewart *et al.* 2001; Lo and Fook-Chong 2004). In accordance with neuropsychological data, speech is inhibited by applying TMS to certain areas of the left cortex; conversely, when TMS is performed on the right hemisphere, some participants are unable to produce melody in song. Speech and non-lyrical melodic singing were also compared by Riecker *et al.* (2000), using functional magnetic resonance imaging (fMRI) to show activation of opposite hemispheres with divergent networks for speaking (left-lateralized) and singing (right-lateralized) (see Chapter 13 for a more detailed discussion). Thus it seems

that, compared with speech, the production of sung vocalises utilizes more right hemisphere resources. However, in this pioneering study a more ecological condition of sung language was not included in the protocol, so it is difficult to assess whether the results obtained can be interpreted in terms of real song.

In contrast, sung language was studied by Jeffries *et al.* (2003) utilizing positron emission tomography (PET), in which non-musicians performed under two conditions: singing an 'overlearned' song and speaking the text of the same song. Opposite lateralization effects were found here as well; the authors concluded that the production of words in song encompasses right-hemisphere-dominant systems known to play a role in the processing of music. Nevertheless, since there was no 'vocalises' condition (non-linguistic melodic singing), it is impossible to determine whether these right-hemisphere activations reflect music processing or linguistic production specific to song, or a combination of the two. In this sense, the use of song to explore music cognition is confirmed, while the status of sung language remains unclear.

The cerebral structures corresponding to different levels of non-linguistic musical processing involved in singing vocalises were differentiated in a recent PET study of song production in musicians (Brown *et al.* 2004). They found that the planum polare (BA38) is more active in complex musical tasks (harmonization and melody repetition) than simpler ones (monotonic vocalization); in addition, the role of the right and left frontal operculum in musical imitation and vocal learning was discussed. However, the phonological and semantic elements of song were not examined, so it is difficult to know if the same musical vocalization network would be mobilized to sing songs with words.

To summarize, studies of the cognition of song have demonstrated that song is a useful model for elucidating the roles of language and music in the brain. Some evidence for left-lateralized language processing and right-lateralized musical processing has been found, but overall it remains unclear as to whether the relationship between these two dimensions of song is of an interactive or independent nature. Task demands and attention modulations seem to have an influence on the results; different results are also observed as a function of which levels of processing are being examined. Furthermore, song perception has not been studied outside the context of a production task with either fMRI or PET. Most importantly, no single study has directly compared spoken language, sung language, and vocalises.

In ongoing research at our laboratory in Marseille, we are conducting a series of experiments (Gordon *et al.* 2004) that aim to determine, first, the cerebral regions involved in the perception of song and, second, the nature of the relationship between the musical and linguistic dimensions of song. In the first experiment the fMRI method was used to establish a functional mapping of the

brain areas involved in the perception of song, compared with speech and vocalises. Thirteen participants (non-musicians) were scanned while listening to stimuli from each of the following three conditions: spoken words (tri-syllabic French words, spoken); sung words (tri-syllabic words sung on three-note melodies); and vocalises (the syllable 'vi' sung on three-note melodies). In all conditions, participants performed a same–different task, in which they responded as quickly and accurately as possible as to whether the two stimuli in the pair were the same or different.

Preliminary results of a random-effects analysis of the data showed tendencies towards left lateralization for spoken words and right lateralization for vocalises and sung words, in agreement with the literature on song production (Riecker *et al.* 2000; Jeffries *et al.* 2003; Brown *et al.* 2004) and language and music processing (Zatorre *et al.* 2002). The contrast between sung words and vocalises (Plate 13) showed left temporal activations most likely reflecting aspects of language processing (e.g. complex phonology, semantics). Conversely, melodic processing was reflected in the contrast between sung words versus spoken words, which revealed bilateral temporal and frontal activity with right-hemisphere dominance.

Interestingly, a comparison of the two main contrasts reveals a greater similarity of activations in the sung words and vocalises conditions than in the sung and spoken words conditions. While the trend of these preliminary results remains to be confirmed by testing additional participants, it could indicate that the perception of song depends more on a musical 'singing voice network' than a linguistic network. However, herein lies the difficulty intrinsic to the study of cognition of song, of creating an experimental design that achieves the delicate balance between linguistic and musical elements. Indeed, the salience of one element or another certainly interacts with the kind of factor, such as task difficulty and allocation of attentional resources, which may modulate brain activity in song cognition, as discussed earlier.

## Song memory in healthy musicians and non-musicians and aphasics

As the status of online processing of language and music in song continues to be controversial, this relationship remains complex in terms of how the linguistic and musical dimensions of song are encoded in memory. In particular, the notion that music may serve as a mnemonic device for learning verbal material has a long history. In oral traditions, minstrels transmitted stories through songs. Today, children learn the alphabet as a song, and students try to remember laws of physics sung to popular songs through karaoke (Dickson and Grant 2003).

It is still unclear why music should facilitate word recall. One hypothesis is based on the fact that production of words is slowed down when we sing compared with when we speak. Kilgour *et al.* (2000) have indeed shown that there is no more advantage for word recall in hearing a text sung versus hearing it spoken when a sung text is compressed to be as short as a spoken text. Speed, or syllable duration, would thus be a crucial variable in the effect of music on word recall. Repetition and simplicity of a melody have also been brought forth as explanations of the better word recall of songs (Wallace 1994). Conversely, no more words were recalled when hearing a text on a complex and changing melody than when hearing a spoken text. Another hypothesis originates from studies using recognition paradigms; these results have suggested that the lyrics and melody of the same song are strongly associated in memory (Serafine *et al.* 1984, 1986; Crowder *et al.* 1990; Morrongiello and Roes 1990; Peretz *et al.* 2004*b*). A melody could therefore give access to the words with which it is learned. Finally, it is also possible to argue that there are structural constraints in music, as well as in song lyrics, that can combine to improve recall (Rubin 1995). For instance, the metrical structure of music, or the number of musical notes in a line, limit the possibilities in word recall, as do the rhymes, assonances, and alliterations that frequently characterize the lyrics.

Experimental studies addressing the question of the effect of music on verbal recall have mainly looked at the presence or absence of music at presentation of the text (e.g. Rainey and Larsen 2002), which can only indicate whether music helps to *encode* the words, and not whether music helps to *produce* more words. Only Jellison and Miller (1982) have systematically manipulated the response (sung or spoken) and the presentation (sung or spoken), and found no advantage (and even a disadvantage) of learning words or digits by singing on a simple melody instead of speaking. Although Gingold and Abravanel (1987) did not find any differences in children's performance between texts learned by reciting or singing, they did show that the musical material can interfere with memorization of a text, since bad musical settings disrupted learning of words in song.

In aphasics, there is a classic observation that they can sing words they are not able to produce otherwise (Keith and Aronson 1975; Yamadori *et al.* 1977; Assal *et al.* 1977; Jacome 1984). These kinds of anecdotal reports led to the use of musical support in speech therapy. This is the case for melodic intonation therapy, which is based on an exaggeration of the melodic line of linguistic phrases, accentuating simple pitches and rhythm (Sparks *et al.* 1974). However, recent studies have surprisingly been unable to show an advantage of singing over speaking in brain-damaged patients suffering from speech disorders (Hébert *et al.* 2003; Peretz *et al.* 2004a). These single-case studies have shown that

an equivalent number of words are produced when the same words have to be repeated by singing or by speaking. Moreover, the same types of errors were committed in both modalities. The idea that non-fluent aphasics are able to sing would stem from the observation that their rote memory of automatic material, such as songs, is superior to their performance in generative tasks, such as spontaneous speech. Thus, music *per se* would not explain this difference.

In order to answer the question of whether singing is accompanied by a more efficient memory formation, two studies were conducted, based on oral learning of unfamiliar songs, in three different conditions (Racette and Peretz, in press; Racette *et al.* in revision). In the first condition (sung–sung) participants listened to the sung version of the lyrics and sung them back. In the second condition, the sung–spoken condition, participants listened to the sung version of the lines and recalled only the lyrics by pronouncing them in a natural way. Comparison of these two conditions allowed us to measure the effect of music on response. In the third condition, the divided–spoken condition, the participant listened to the spoken lyrics and recalled only the lyrics. In this condition, the presentation was defined as divided because, as the lyrics were spoken, the corresponding melody was hummed on /la/ in the background. Consequently, when the lyrics were spoken, the context was equivalent to when lyrics were sung, since the melody was presented in both cases. Thus, the divided–spoken condition allowed us to control for the difficulty of extracting text from a song while maintaining the presence of the melody during encoding.

## Neurologically intact participants

Thirty-six university students participated in a first study (Racette and Peretz, in revision, a). Half were non-musicians and half were enrolled in the Music Faculty. Participants learned one different song in each of the three conditions following an adaptive procedure. Songs selected for the study were songs that already existed, composed by a folk singer, but that were unfamiliar to all participants. Songs were composed of eight lines. The percentage of correctly recalled words in each condition, as well as musical notes for the sung–sung condition, were calculated when participants recalled the song.

The first surprising finding was that the musicians were no better than the non-musicians in recalling words and musical notes. This might be explained by the fact that musicians usually learn a song with the score. In the procedure used here, participants could only rely on the auditory code of the song, which is also familiar to non-musicians. Therefore, musicians had no advantage over non-musicians in an oral learning procedure in which they were not allowed to use a visual support. When the data for the two groups were pooled together, as

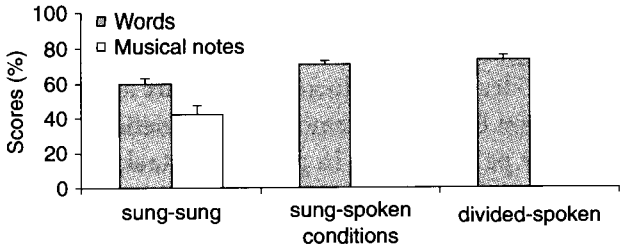



**Fig. 14.1** (a) Mean scores obtained by neurologically intact participants while recalling words and musical notes of unfamiliar songs in three conditions.

can be seen in Fig. 14.1, the sung recall was significantly lower than both spoken recalls. This means that music interfered with word recall at response. Moreover, because there were no more words recalled in the sung–spoken condition than in the divided–spoken conditions, there was no facilitation of music at presentation. There was no interference, either. Moreover, there were more words recalled than musical notes when singing. These results suggest that the lyrics and the melody of a song are represented in independent codes in memory. Participants would have to access both the melody and the lyrics, making singing a dual task, at least in the first steps of learning.

## Aphasic patients

What happens when one of the codes is altered? It is now known that there is a double dissociation between production of the words (altered in CC (Hébert et al. 2003), spared in IR (Peretz *et al.*1997)) and production of the melody (altered in IR, spared in CC) of songs in patients. Aphasic patients could thus use (spared) music to access words.

In this second study (Racette and Peretz, in revision), eight French-speaking patients were recruited who suffered from a variety of expressive aphasias following left-hemisphere brain lesions. They learned songs in the same three conditions as the university students, but following a simplified learning procedure. The songs were also reduced to four lines. The percentage of words correctly repeated after every line was calculated.

As is shown in Plate 14, there was no difference between conditions. Thus, singing did not improve word production compared with speaking. This was true for the number of words produced but also for the types of errors in the different conditions. Moreover, aphasic patients did not produce more words than notes. The same profile was obtained in subsequent experiments using familiar material. Hence, there were as few words of familiar songs repeated by singing as by speaking. Even if songs are naturally sung with both the lyrics and the melody, there was no cost of producing the lyrics alone. This again suggests that the words

and the melody are accessed separately in memory for songs. In an additional experiment, spoken and sung (on a familiar melody) repetitions of prayers and proverbs again showed that no more words were produced by singing than speaking. These results show that, in equivalent settings, there is no advantage of singing versus speaking the lyrics. This phenomenon applied both to the learning of unfamiliar songs and the production of highly familiar and automatic material.

The only way singing could be more conducive to the production of language than speaking is when patients are asked to sing along, with a slow speed. This is often the first step in speech therapy, where the patient produces words in unison with the therapist. The effect of speed might be magnified by singing, which is also slower than speaking. After extensive training, patients might be able to produce more words when singing along than when speaking in unison. However, they would have to learn to produce words alone at some point, since singing is not a functional way of communication in everyday life. Moreover, when singing along, participants heard what they had to produce at the same time they were producing it. Thus, musical notes might have to be heard to be able to cue recall.

Music and language have to be compared on similar levels to determine if they share the same underlying processes or not. When comparing the production of sung and spoken words, it was suggested that there is a unique code for words, be they sung or spoken. This verbal code would be separated from the musical code for production of musical notes. Therefore, oral production of words and music in both neurologically intact participants and aphasic patients provides additional support for independent representations of music and language in memory for songs.

## Discussion

We have seen that the study of song is particularly interesting because it allows us to investigate both musical and linguistic processing as well as their relationship. Indeed, the study of singing is not restricted to musicians, as is the case for instrumental playing, but can be done with so-called non-musicians. In this regard we would like to point out that in Western society the category of 'musician' is defined by very strict criteria, such as musical literacy, extensive theoretical musical education, and instrumental training. However, in doing this we often forget the fact that music, like language, does not necessarily require explicit learning. Indeed, implicit learning takes place for music too (Bigand, 2003), and, as for language, it is not just limited to comprehension. In other words, as most people can speak, most people can sing. This consideration should encourage researchers to better investigate the ability of 'non-musicians' to sing, which may not be so different from the singing abilities of 'musicians' (of course excluding 'singers'!).

The results presented above are not encouraging concerning the effects of musical rehabilitation techniques. However, this null effect might be due to several reasons. First of all, it is important to note that the effect of music training on linguistic or other cognitive abilities has been mostly found with *extensive* musical training (Schellenberg, 2001). In other words, the short-term effects of music, such as the famous Mozart effect, usually seem to be an artefact of mood, arousal, or attentional processes. Therefore, only long rehabilitation procedures may benefit from the use of song or other musical material. Indeed, when possible, the rehabilitation of aphasic patients is a rather long process that takes place over several months. Second, the controversial results of the literature concerning the facilitation of song over speech on aphasic patients might be due to the diversity of the aphasic syndrome. In other words, song facilitation might be effective for only a certain type of aphasic patient, although the study described above did not show any benefit for different types of speech disorder.

This brings us to a more theoretical reflection about the beneficial effect of music in speech disorders. Rather than asking whether music could be beneficial, a first step to progress in this field would be to ask what in music is beneficial to speech production. As already stated above, music has different levels of processing and requires different types of computations that are performed by different neural networks. Therefore, the treatment should put the emphasis on the correct level of processing. For instance, a treatment emphasizing the metric/rhythmic support of speech may improve repetition to a greater degree than using only the intonational elements of melody (Boucher *et al.* 2001). Indeed, a strong metric and rhythmic structure of a musical phrase, if it matches the prosodic structure of the words, may facilitate word retrieval, especially if rhythm in both language and music shares common neural resources (Patel and Daniele, 2003). When music is used as a mnemonic device in learning verbal material, the metric structure constrains the linguistic information, while the melodic structure does so to a lesser degree (although this might be different in the case of tone languages (Chan 1991)). It should also be noted that in song the musical structure takes precedence in determining the realization of a large number of the acoustic parameters, and somewhat dictates, or at least transforms, the linguistic prosody. Since this transformation may heavily influence the intelligibility of the lyrics, one should exercise caution when studying language processing in song.

Additional benefits may be gained from other aspects of music therapy. Since aphasic patients do not necessarily suffer from a music disorder (Hébert *et al.* 2003), producing musical notes alone is still a good way for them to produce vocal sounds. While this would probably not be beneficial to agrammatic aphasic patients, it may help in the rehabilitation process of patients with more vocal/motor dysfunctions (e.g. dysarthria).

Moreover, it seems that most aphasic patients enjoy singing. While this is related to the general effects of music on mood and arousal (Schellenberg 2001; Grape *et al.* 2003), this will engender greater motivation in rehabilitation sessions, which may turn into both greater and more effective participation during the single session and longer endurance over the rehabilitation period.

To conclude, more theoretically driven research is needed to better clarify the intriguing possibility that speech disorders may benefit from musical rehabilitation techniques. With this aim, the clinical research will need to be coupled with fundamental research in order to refine the theoretical basis of the rehabilitation methods in use. Several important questions that also go beyond the issue of rehabilitation remain open: what musical information is most likely to improve word retrieval? For instance, concerning word retrieval, would a strong metric structure play a more important role than melodic structure? Will memory storage in general benefit from any type of musical information or training? Moreover, to what extent is musical practice better than passive listening? Will listening to a song or singing it have similar or different effects on mood or arousal? More action is needed to better understand the relationship between speech and music.

## References

**Alain, C. and Izenberg, A.** (2003). Effects of attentional load on auditory scene analysis. *J Cogn Neurosci*, **15**(7), 1063–1073.

**Assal, G., Buttet, J., and Javet, R.C.** (1977). [Musical aptitude in aphasics]. *Rev Med Suisse Romande*, **97**(1), 5–12.

**Bartholomeus, B.** (1974). Effects of task requirements on ear superiority for sung speech. *Cortex*, **10**(3), 215–223.

**Bates, E.** (1994). Modularity, domain specificity and the development of language. *Discuss Neurosci*, **10**, 136–149.

**Belin, P., Fecteau, S., and Bedard, C.** (2004). Thinking the voice: neural correlates of voice perception. *Trends Cogn Sci*, **8**(3), 129–135.

**Besson, M., Faita, F., Peretz, I. Bonnel, A.-M., and Requin, J.** (1998). Singing in the brain: independence of lyrics and tunes. *Psychol Sci*, **9**(6), 494–498.

**Besson, M. and Schön, D.** (2001). Comparison between language and music. *Ann NY Acad Sci*, **930**, 232–258.

**Bigand, E.** (2003). More about the musical expertise of musically untrained listeners. *Ann NY Acad Sci*, **999**, 304–312.

**Bigand, E., Tillmann, B., Poulin, B., D'Adamo, D.A., and Madurell, F.** (2001). The effect of harmonic context on phoneme monitoring in vocal music. *Cognition*, **81**(1), 11–20.

**Bonnel, A.M., Faita, F., Peretz, I., and Besson, M.** (2001). Divided attention between lyrics and tunes of operatic songs: evidence for independent processing. *Percept Psychophys*, **63** (7), 1201–1213.

**Boucher, V., Garcia, L.J., Fleurant, J., and Paradis, J.** (2001). Variable efficacy of rhythm and tone in melody-based interventions: implications for the assumption of a right-hemisphere facilitation in non-fluent aphasia. *Aphasiology*, **15**, 131–149.

**Brown, S., Martinez, M.J., Hodges, D.A., Fox, P.T., and Parsons, L.M.** (2004). The song system of the human brain. *Cogn Brain Res*, **20**(3), 363–375.

**Cadalbert, A., Landis, T., Regard, M., and Graves, R.E.** (1994). Singing with and without words: hemispheric asymmetries in motor control. *J Clin Exp Neuropsychol*, **16**(5), 664–670.

**Chan, S.Y.** (1991). Improvisation in a ritual context. In: *The music of Cantonese opera*. The Chinese University Press, Hong Kong.

**Crowder, R.G., Serafine, M.L., and Repp, B.** (1990). Physical interaction and association by contiguity in memory for the words and melodies of songs. *Memory Cogn*, **18**(5), 469–476.

**Dickson, D. and Grant, L.** (2003). Physics karaoke: why not?. *Phys Educ*, **38**(4), 320–323.

**Epstein, C.M., Meador, K.J., Loring, D.W.** *et al.* (1999). Localization and characterization of speech arrest during transcranial magnetic stimulation. *Clin Neurophysiol*, **110**(6), 1073–1079.

**Gingold, H. and Abravanel, E.** (1987). Music as a mnemonic: the effects of good- and bad-music settings on verbatim recall of short passages by young children. *Psychomusicology*, 7, 25–39.

**Goodglass, H. and Calderon, M.** (1977). Parallel processing of verbal and musical stimuli in right and left hemispheres. *Neuropsychologia*, **15**(3), 397–407.

**Gordon, H.W. and Bogen, J.E.** (1974). Hemispheric lateralization of singing after intracarotid sodium amylobarbitone. *J Neurol Neurosurg Psychiat*, **37**(6), 727–738.

**Gordon, R.L., Schön, D., Magne, C.** *et al.* (2004). An fMRI study of the neural basis of song perception. Poster presented at *Conference on Interdisciplinary Musicology (CIM04), Graz, Austria, 15–18 April, 2004.*

**Grape, C., Sandgren, M., Hansson, L.O., Ericson, M., and Theorell, T.** (2003). Does singing promote well-being?: an empirical study of professional and amateur singers during a singing lesson. *Integ Physiol Behav Sci*, **38**(1), 65–74.

**Hébert, S., Racette, A., Gagnon, L., and Peretz, I.** (2003). Revisiting the dissociation between singing and speaking in expressive aphasia. *Brain*, **126**(8), 1838–1850.

**Jacome, D.E.** (1984). Aphasia with elation, hypermusia, musicophilia and compulsive whistling. *J Neurol Neurosurg Psychiat*, **47**(3), 308–310.

**Jeffries, K.J., Fritz, J.B., and Braun, A.R.** (2003). Words in melody: an $H_2{}^{15}O$ PET study of brain activation during singing and speaking. *NeuroReport*, **14**(5), 749–754.

**Jellison, J.A. and Miller, N.L.** (1982). Recall of digit and word sequences by musicians and nonmusicians as a function of spoken or sung input and task. *J Music Ther*, **XIX**(4), 194–209.

**Keith, R.L. and Aronson, A.E.** (1975). Singing as therapy for apraxia of speech and aphasia: report of a case. *Brain Lang*, **2**(4), 483–488.

**Kilgour, A.R., Jakobson, L.S., and Cuddy, L.L.** (2000). Music training and rate of presentation as mediators of text and song recall. *Mem Cogn*, **28**(5), 700–710.

**Levy, D.A., Granot, R. and Bentin, S.** (2001). Processing specificity for human voice stimuli: electrophysiological evidence. *NeuroReport*, **12**(12), 2653–2657.

**Levy, D.A., Granot, R., and Bentin, S.** (2003). Neural sensitivity to human voices: ERP evidence of task and attentional influences. *Psychophysiology*, **40**(2), 291–305.

Lidji, P. (2003). Relation between lyrics and tunes in the perception and memorisation of songs. Unpublished undergraduate thesis, Université Libre de Bruxelles.

Lo, Y.L. and Fook-Chong, S. (2004). Ipsilateral and contralateral motor inhibitory control in musical and vocalization tasks. *Exp Brain Res.* (In press.)

Morrongiello, B.A. and Roes, C.L. (1990). Children's memory for new songs: integration or independent storage of words and tunes? *J Exp Child Psychol*, **50**(1), 25–38.

Patel, A.D. and Daniele, J.R. (2003). An empirical comparison of rhythm in language and music. *Cognition*, **87**(1), B35–B45.

Patel, A.D. and Peretz, I. (1997). Is music autonomous from language? A neuropsychological appraisal. In: I. Deliège and J. Sloboda (eds) *Perception and cognition of music*, pp. 191–215. Erlbaum Psychology Press, London.

Peretz, I., Belleville, S., and Fontaine, S. (1997). [Dissociations between music and language functions after cerebral resection: A new case of amusia without aphasia]. *Can J Exp Psychol*, **51**(4), 354–368.

Peretz, I., Gagnon, L., Hébert, S., and Macoir, J. (2004*a*) Singing in the brain: Insights from cognitive neuropsychology. *Music Percept*, **21**(3), 373–390.

Peretz, I., Radeau, M., and Arguin, M. (2004*b*). Two-way interactions between music and language: evidence from priming recognition of tune and lyrics in familiar songs. *Memory Cogn*, **32**(1), 142–152.

Poulin-Charronnat, B., Bigand, E., Madurell, F., and Peereman, R. (2005). Musical structure modulates semantic priming in vocal music. *Cognition*, **94**(3), B67–B78.

Racette, A. and Peretz, I. (in press). Learning Lyrics: To sing or not to sing?, Memory & Cognition.

Racette, A., Peretz, I., and Bard, C. (in revision). Making non-fluent aphasics speak: Sing Along!.

Rainey, D.W. and Larsen, J.L. (2002). The effect of familiar melodies on initial learning and long-term memory for unconnected text. *Music Percept*, **20**(2), 173–186.

Rapp, B. (2001). *The handbook of cognitive neuropsychology*. Psychology Press, Philadelphia, PA.

Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G., and Grodd, W. (2000). Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *NeuroReport*, **11**(9), 1997–2000.

Rubin, D.C. (1995). *Memory in oral traditions*. Oxford University Press, New York.

Ryding, E., Bradvik, B., and Ingvar, D.H. (1987). Changes of regional cerebral blood flow measured simultaneously in the right and left hemisphere during automatic speech and humming. *Brain*, **110**(5), 1345–1358.

Schellenberg, E.G. (2001). Music and nonmusical abilities. *Ann NY Acad Sci*, **930**, 355–371.

Schön, D., Lorber, B., Spacal, M., and Semenza, C. (2004). A selective deficit in the production of exact musical intervals following right-hemisphere damage. *Cogn Neuropsychol*, **27**, 773–784.

Semenza, C. (1999). Lexical-semantic disorders in aphasia. In: G. Denes and L. Pizzamiglio (eds) *Handbook of neuropsychology*, pp. 215–244. Psychology Press, Hove.

Serafine, M.L., Crowder, R.G., and Repp, B.H. (1984). Integration of melody and text in memory for songs. *Cognition*, **16**(3), 285–303.

**Serafine, M.L., Davidson, J., Crowder, R.G., and Repp, B.H.** (1986). On the nature of melody-text integration in memory for songs. *J Mem Lang*, **25**, 123–135.

**Sparks, R., Helm, N., and Albert, M.** (1974). Aphasia rehabilitation resulting from melodic intonation therapy. *Cortex*, **10**(4), 303–316.

**Stager, S.V., Jeffries, K.J., and Braun, A.R.** (2003). Common features of fluency-evoking conditions studied in stuttering subjects and controls: an $H_2{}^{15}O$ PET study. *J Fluency Disord*, **28**(4), 319–335 (quiz 336).

**Stewart, L., Walsh, V., Frith, U., and Rothwell, J.** (2001). Transcranial magnetic stimulation produces speech arrest but not song arrest. *Ann NY Acad Sci*, **930**, 433–435.

**Tervaniemi, M., Medvedev, S.V., Alho, K.** *et al.* (2000). Lateralized automatic auditory processing of phonetic versus musical information: a PET study. *Hum Brain Mapp*, **10**(2), 74–79.

**Wallace, W.T.** (1994). Memory for music: effect of melody on recall of text. *J Exp Psychol: Learn Memory Cogn*, **20**(6), 1471–1485.

**Yamadori, A., Osumi, Y., Masuhara, S., and Okubo, M.** (1977). Preservation of singing in Broca's aphasia. *J Neurol Neurosurg Psychiat*, **40**(3), 221–224.

**Zatorre, R.J., Belin, P., and Penhune, V.B.** (2002). Structure and function of auditory cortex: music and speech. *Trends Cogn Sci*, **6**(1), 37–46.

Chapter 15

# The role of inhibition in the motor control of finger function

Christian Gerloff and Friedhelm Hummel

Successful performance of music requires contextual modulation of learned motor 'programs', that is, the retrieval or non-retrieval (inhibition) of behavioural elements depending on the situation. This chapter describes recent advances in our understanding of the neural correlates of these elementary aspects of behaviour, as identified with transcranial magnetic stimulation (TMS), electroencephalography (EEG), and functional magnetic resonance imaging (fMRI). Inhibition of pre-learned behavioural programs (e.g. sequences of finger movements) is represented in the brain by a reduction of net synaptic activity in the cerebrocerebellar pathway and by specific modulation of oscillatory activity of the cortical neurons, leading to a suppression of excitability of the motor cortex. The metabolic correlate of this inhibition is a decrease in the blood oxygenation level-dependent (BOLD) signal to below the resting state (negative BOLD). The reverse is true for retrieval (i.e. an increase in the BOLD signal with enhanced excitability of the motor cortex). These data highlight the importance of a finely tuned interplay of excitatory and inhibitory neuronal activities for the successful control of complex motor behaviour as needed for excellent musical performance.

## Introduction

The performance of music virtually always requires skilled motor behaviour. Successful motor behaviour depends on learning and the appropriate adaptation of learned actions to different situations or contexts. Adequate contextual control of behaviour requires retrieval and non-retrieval, or suppression, of learned actions, that is, activation *and* deactivation of brain regions representing these actions. For example, in a concert different musicians play their parts at different times, and only if this timing is precise will the performance be good. Precise timing implies that each musician knows exactly when to retrieve the learned motor sequence and when to 'wait'. Waiting, however, can occur in situations that are very similar to situations requiring action, for example in the context of

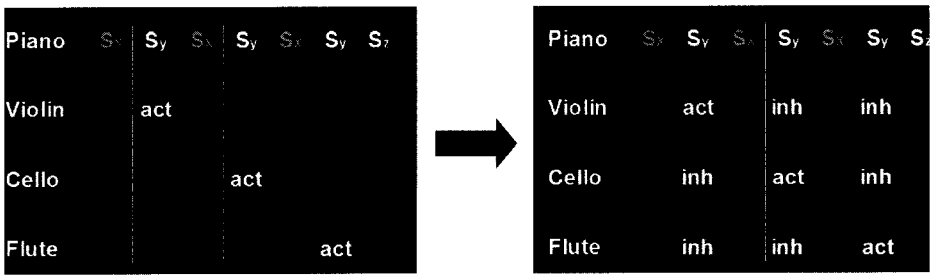



**Fig. 15.1** Schematic of the concept of active inhibition in the context of a musical performance. Left: The piano plays three different sequences Sx, Sy, and Sz. On the first Sy, the violin joins in, on the second Sy the cello, and on the third Sy the flute. In each case, Sy is the stimulus triggering the start of the instrument. Commonly, this is interpreted as a series of excitatory events. Right: The same 'concert', but highlighting the fact that the same stimulus (beginning of Sy) prompts the violin first to play and then 'not to play' twice. The same holds true for the other instruments. Each square labelled 'inh' indicates the necessity of inhibiting the retrieval of the learned motor program because the context has changed (although Sy is always identical).

variations of the same theme. In this situation, waiting might be more of an active process than is commonly thought. Waiting might be active inhibition rather than passive 'idling'. Figure 15.1 illustrates this concept.

What are the neural correlates of this 'bidirectional' (excitatory and inhibitory) contextual control? In particular, what is the neural correlate of effective inhibition in the sense of suppression of behaviour? The concept of inhibition is pervasive in human experiments on complex motor behaviour but is commonly ignored. We have studied this question in a paradigm of finger movement sequences which were performed on custom-made MRI-compatible keyboards.

The importance of this aspect of motor control is emphasized by the clinical observation that disturbances of contextual modulation (especially deficient inhibition) can be of pathophysiological relevance in patients who have diseases with known or proposed deficits of inhibitory neuronal circuitry, such as dystonia, Tourette syndrome, or compulsive disorders (Berardelli *et al.* 1998; Greenberg *et al.* 2000; Jankovic 2001). In certain forms of dystonia (e.g. musician's or occupational dystonia, writer's cramp), the occurrence of neurological symptoms depends strongly on the context of the situation (Sheehy and Marsden 1982; Cohen and Hallett 1988; Fahn 1988).

## The paradigm

The basic principle of the paradigm used is that identical visual stimuli are presented to prompt the subject to act (condition ACT) or not to act (condition INH), depending on the context (Hummel *et al.* 2002). The motor act consists of finger movement sequences that were learned prior to the actual experiment. In

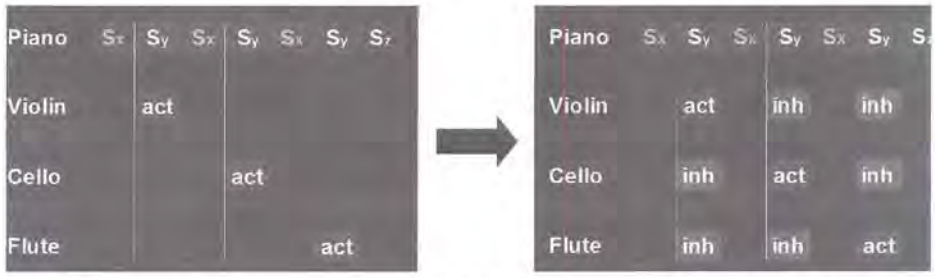



**Fig. 15.1** Schematic of the concept of active inhibition in the context of a musical performance. Left: The piano plays three different sequences Sx, Sy, and Sz. On the first Sy, the violin joins in, on the second Sy the cello, and on the third Sy the flute. In each case, Sy is the stimulus triggering the start of the instrument. Commonly, this is interpreted as a series of excitatory events. Right: The same 'concert', but highlighting the fact that the same stimulus (beginning of Sy) prompts the violin first to play and then 'not to play' twice. The same holds true for the other instruments. Each square labelled 'inh' indicates the necessity of inhibiting the retrieval of the learned motor program because the context has changed (although Sy is always identical).

variations of the same theme. In this situation, waiting might be more of an active process than is commonly thought. Waiting might be active inhibition rather than passive 'idling'. Figure 15.1 illustrates this concept.

What are the neural correlates of this 'bidirectional' (excitatory and inhibitory) contextual control? In particular, what is the neural correlate of effective inhibition in the sense of suppression of behaviour? The concept of inhibition is pervasive in human experiments on complex motor behaviour but is commonly ignored. We have studied this question in a paradigm of finger movement sequences which were performed on custom-made MRI-compatible keyboards.

The importance of this aspect of motor control is emphasized by the clinical observation that disturbances of contextual modulation (especially deficient inhibition) can be of pathophysiological relevance in patients who have diseases with known or proposed deficits of inhibitory neuronal circuitry, such as dystonia, Tourette syndrome, or compulsive disorders (Berardelli *et al.* 1998; Greenberg *et al.* 2000; Jankovic 2001). In certain forms of dystonia (e.g. musician's or occupational dystonia, writer's cramp), the occurrence of neurological symptoms depends strongly on the context of the situation (Sheehy and Marsden 1982; Cohen and Hallett 1988; Fahn 1988).

## The paradigm

The basic principle of the paradigm used is that identical visual stimuli are presented to prompt the subject to act (condition ACT) or not to act (condition INH), depending on the context (Hummel *et al.* 2002). The motor act consists of finger movement sequences that were learned prior to the actual experiment. In

this learning session, external stimuli presented at a rate of one per second are associated with the required movements. In the recording session, during ACT these movements are retrieved paced by visual stimuli. During the INH condition, the subject is instructed not to retrieve the movements although the same stimulus is presented. The information about the context (ACT or INH) is given several seconds before presentation of the pacing stimuli. In contrast to the classical go/no-go paradigms, this produces a shift in context without creating time pressure in the millisecond range and it can be applied to highly complex motor acts. Figure 15.2 illustrates the experimental design which was used with TMS, EEG, and fMRI experiments.

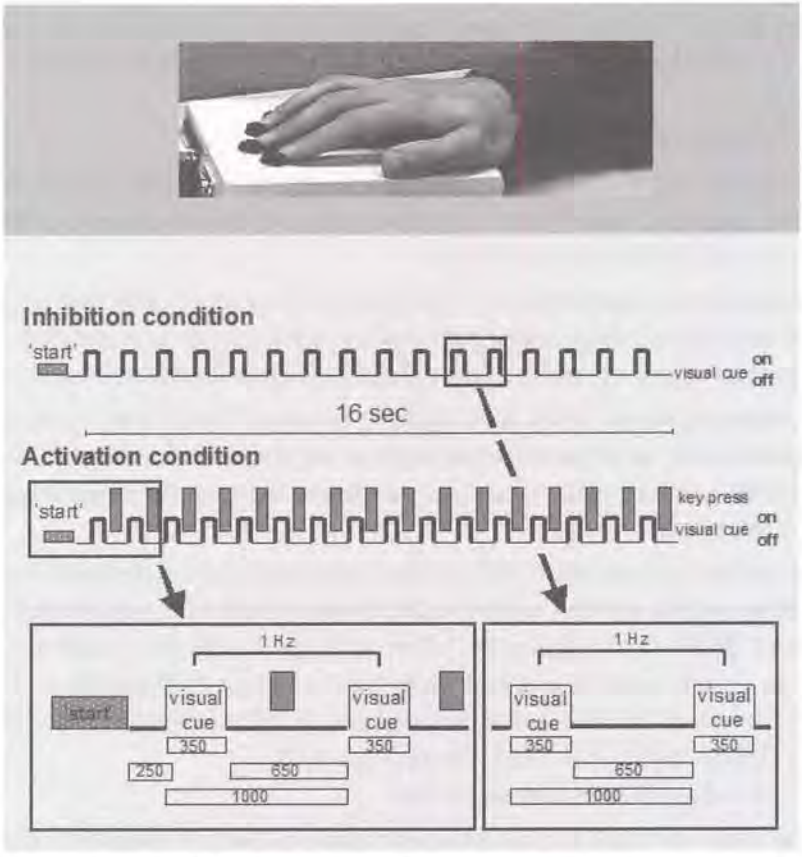



**Fig. 15.2** Top: Custom-made keyboard. Bottom: Temporal structure of stimulus presentation. In the inhibition condition, the subjects observed the visual cues but were instructed not to retrieve the learned finger movements. In the activation condition, they responded to visual cues with retrieval of individual finger movements. Solid blocks indicate key presses, high and low on the black line responds to visual cue on and off. The duration of intervals is given in millisecond in the zoomed boxes below.

this learning session, external stimuli presented at a rate of one per second are associated with the required movements. In the recording session, during ACT these movements are retrieved paced by visual stimuli. During the INH condition, the subject is instructed not to retrieve the movements although the same stimulus is presented. The information about the context (ACT or INH) is given several seconds before presentation of the pacing stimuli. In contrast to the classical go/no-go paradigms, this produces a shift in context without creating time pressure in the millisecond range and it can be applied to highly complex motor acts. Figure 15.2 illustrates the experimental design which was used with TMS, EEG, and fMRI experiments.

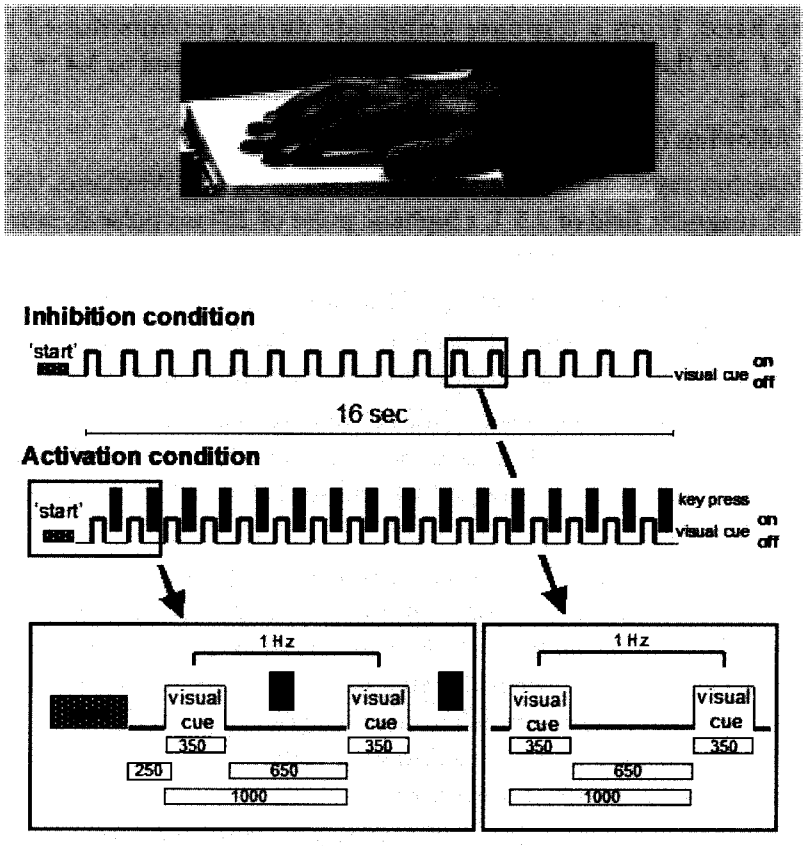



**Fig. 15.2** Top: Custom-made keyboard. Bottom: Temporal structure of stimulus presentation. In the inhibition condition, the subjects observed the visual cues but were instructed not to retrieve the learned finger movements. In the activation condition, they responded to visual cues with retrieval of individual finger movements. Solid blocks indicate key presses, high and low on the black line responds to visual cue on and off. The duration of intervals is given in millisecond in the zoomed boxes below.

The following description will focus on the results and their implications for motor control. Methodological information is included where it is considered to be relevant.

## Neural correlates of movement retrieval

All subjects overlearned the finger movements during the training session, which was confirmed by their performance (in correct responses) of 97.4 ± 1.4% (mean ± SD) ($n$ = 18 normal volunteers for TMS/EEG ) and 98.0 ± 3.7% ($n$ = 6 for TMS/fMRI) during fMRI in the experimental session (ACT condition).

When retrieving the learned movement sequences, TMS documented significant facilitation of the primary motor cortex (M1) contralateral to the moving hand. From a response amplitude during unconstrained rest of 1.4 ± 1.1 mV, the motor evoked potential (MEP) increased to 3.4 ± 2.1 mV during ACT.

The following *TMS methods* were used:

- A 'Magstim rapid' stimulator (Magstim Company Ltd, Whitland, UK) with a peak magnetic field of 2.5 T, a biphasic pulse of 250 ms pulse duration, and a magnetic field rise time of 60 ms.
- A magnetic coil shaped like an 'eight' (inner diameter of each loop 5 cm) was used with the handle pointing backwards and laterally at an angle of about 45° to the sagittal plane in order to elicit MEPs in the first dorsal interosseus muscle.
- For each condition, 20–30 trials (MEPs) were recorded. Based on theoretical considerations and typical latencies of cortical responses to visual stimulation (e.g. P100), a fixed delay of 120 ms was chosen between the visual trigger and the TMS pulse.

The facilitation seen with TMS was accompanied by focal desynchronization of oscillatory EEG activity, in line with previous studies on task-related power decrease (TRPD) in the alpha (7–13 Hz) and beta (14–30 Hz) frequency range. TRPD as a parameter for cortical activation is similar to the well-established concept of event-related desynchronization (ERD) of oscillatory neuronal activity (Pfurtscheller *et al.* 1994; Gerloff *et al.* 1998).

The following *EEG methods* were used:

- Focal TRPD or ERD are accepted surrogate markers of excitatory neuronal activation. Continuous EEG was recorded from 28 (Ag/AgCl) surface electrodes (Fig. 1b), mounted in a cap (Electro-Cap International, Inc., Eaton, OH, USA).
- Impedance was kept below 5 kΩ. Data were sampled at 250 Hz, the upper cut-off was 50 Hz, and the time constant was set to DC (Synamp amplifiers

and software by NeuroScan Inc., El Paso, TX, USA). Linked earlobe electrodes served as reference.

- The analysis in the frequency domain was based on FFT algorithms and normalization as described elsewhere (Gerloff *et al.* 1998; Andres *et al.* 1999). After stabilizing the variance by log transformation, the data were entered into a factorial ANOVA design. The significance level was set to $P < 0.05$ after correction for multiple comparisons.

Topographical EEG analysis revealed activity not only in the contralateral sensorimotor region but also in the central region of the hemisphere ipsilateral to the moving hand (Hummel *et al.* 2002). Bilateral activation of sensorimotor areas during performance of complex finger movements has been a characteristic finding in various studies (Sadato *et al.* 1996; Gerloff and Hummel 2003; Hummel *et al.* 2003). This pattern was confirmed and complemented by the subsequent fMRI experiment. The structures activated during ACT involved M1 (bilateral BA4), secondary motor areas (right dorsal pre-motor cortex, PMd; supplementary motor area, SMA; both BA6), primary somatosensory cortex (S1; left BA3), posterior parietal cortex (i.e. bilateral pre-cuneus and superior parietal lobe, SPL; both BA7), the cingulate cortex (BA24, BA31), and the cerebellum. In addition to these motor-related areas, visual areas (BA19, BA37) were activated—consistent with visual cueing.

The following *functional MRI methods* were used:

- For fMRI data, functional images of the BOLD responses (Ogawa *et al.* 1990) were acquired with a 1.5 T scanner (Magnetom Vision, Siemens, Germany) with a gradient echo planar imaging (EPI) sequence in which one volume is equivalent to 28 axial slices (4 mm thickness, 1 mm gap) with repetition time (TR), 3000 ms; echo time (TE), 39 ms; a 64 × 64 matrix; flip angle 90°; and field of view of 192 × 192 mm. Serving as the anatomical reference were high-resolution images based on a $T_1$-weighted three-dimensional turbo-flash sequence. The commercially available BrainVoyager (Version 4.9, Brain Innovation, R. Goebel, Maastricht, The Netherlands) was used for fMRI analysis.
- Pre-processing of data included three-dimensional motion correction, slice scan-time correction, spatial smoothing of EPI images with full-width at half-maximum (FWHM) of 4 mm and temporal smoothing with trend removal by temporal fast Fourier transform-based high-pass filtering, co-registration, and transformation of scans according to Talairach and Tournoux. Details of these procedures are described in Trojano *et al.* (2000).
- A general linear model (fixed effect) with predictors for the experimental conditions of ACT and INH and a further predictor for instruction was

computed for multiple regression analysis. This design allowed us to compare the ACT condition and the INH condition with baseline. As our main hypothesis was that some of the same areas involved in activation (excitation) during ACT can be inhibited during INH, a conjunction analysis was computed. With this analysis it is possible to determine areas that are significantly activated during ACT and significantly 'deactivated' during INH (= 'bidirectional modulation of metabolic activity' above and below baseline).

## Neural correlates of motor inhibition (non-retrieval)

The key finding prompting us to pursue this line of experiments was a significant reduction in the MEP amplitude when subjects were asked not to move even though the visual cues (learned to be associated with movement) were presented. From a baseline value of 1.4 ± 1.1 mV, the MEP amplitude decreased to 0.8 ± 0.7 mV in the inhibition condition.

Cortical oscillatory activity showed a distinct task-related power increase (TRPI), similar to event-related synchronization (ERS) in the alpha frequency range (maximum 12–13 Hz). As this effect was quite restricted to the alpha frequency, it will be referred to as α-TRPI. As during execution of the motor sequences (ACT), during INH the task-related change in cortical oscillations was not restricted to the M1 contralateral to the moving hand. Also α-TRPI exhibited a bilateral topographic distribution over the sensorimotor regions of both hemispheres (Hummel *et al.* 2002). Previously, alpha power increases had been largely interpreted as correlates of cortical 'idling'. The coincidence of MEP suppression and α-TRPI, however, provides evidence that regional enhancement of alpha oscillations might be operational in down-regulating cortical excitability (active inhibition).

The observed α-TRPI was specific to the situation of previous motor learning. Without prior training and thus establishment of a cue–response association, no α-TRPI occurred. It was also absent if neutral visual material was presented (a video clip without connection to motor performance or other elements of the experimental setting) instead of the visual cues.

We therefore conclude that in this experimental situation α-TRPI is a correlate of contextual inhibitory modulation of a learned motor program. It is tempting to speculate that α-TRPI operates as a mechanism capable of suppressing recently established increases in synaptic gains by long-term potentiation (LTP). Preliminary data on the duration of α-TRPI after a single session support this view, as α-TRPI shows reversibility, which is also one of the characteristics of LTP. Further studies are necessary in order to clarify the exact mechanisms underlying α-TRPI during contextual inhibition. Plate 15 summarizes TMS and EEG results for the ACT and INH condition. The similarity of the topography of

α-TRPI and α-TRPD was one major observation that lead us to assess the fMRI data not only with traditional parametric or differential analyses but also with a conjunction analysis which allows for identification of areas showing both inhibitory *and* excitatory changes.

The fMRI data allowed for a finer anatomical characterization of the neuronal network involved in contextual motor inhibition, including the cerebellum (Hummel *et al.* 2004). We identified as areas with bidirectional modulation of activity the bilateral M1 (BA4), the secondary motor areas (right dorsal premotor cortex, PMd; supplementary motor area, SMA; both BA6), the primary somatosensory cortex (S1; left BA3), the posterior parietal cortex (i.e. bilateral pre-cuneus and superior parietal lobe, SPL; both BA7), the cingulate cortex (BA24, BA31), and the cerebellum. That is, an extended cerebrocerebellar network showed enhanced metabolic activity during ACT but suppressed activity (below baseline) during INH. Besides the implication for motor physiology, this was also one of the first experiments to prove a clear association of inhibition of electrophysiological markers and negative BOLD responses. It is necessary to point out that inhibitory neuronal processes are not well reflected in the neuroimaging literature. For example, a literature search in the international database PubMed in January 2004 revealed 3389 articles for the search terms 'functional imaging' and 'activation', but only 294 for 'functional imaging' and 'inhibition'. In contrast, 30–40% of cortical neurons are inhibitory Golgi type II cells, and the divergence of synaptic connections from one inhibitory interneuron to target neurons is much higher than the one of excitatory neurons. In essence, all the evidence points to a functionally relevant (if not dominant) role of inhibitory neuronal activities in human behaviour. The fMRI results are given in Plate 16.

What does it mean at the theoretical level that the BOLD signal turns negative, i.e. below baseline? This question has been controversially debated for several years now. In the framework of this chapter it is important to note that there are theoretical considerations, modelling data (Tagamets and Horwitz 2001), and recent experimental studies of the visual system of humans and non-human primates (Shmuel *et al.* 2002) that provide grounds for the concept that neuronal inhibition can be reflected by negative BOLD (for a more detailed discussion of this topic see Hummel *et al.* 2004), and that negative BOLD is not merely a consequence of local vascular autoregulation or capillary steal effects.

## Implications for the musician's brain

There are two main points which we would like to stress. First, neuroimaging studies on motor physiology in musicians need to take into account the essential role of inhibitory neuronal circuits. In many instances it might make sense to

compute contrasts that allow for the assessment of reductions in activity rather than only evaluating putative excitatory phenomena. Second, the present data might help to further our understanding of neurological diseases frequently affecting professional musicians, such as focal dystonia. To this end, we have studied a group of six patients (five of them with musician's dystonia, one with writer's cramp), and the most striking finding was an absence of the typical inhibitory $\alpha$-TRPI (Hummel *et al.* 2002). This result extends previous observations of compromised inhibitory circuits in dystonia, because it shows for the first time the absence of proper contextual inhibition during performance of a complex motor task. The multimodal investigation of such phenomena in patients is a promising approach to the pathophysiology of dystonia.

## Conclusions

The study described here identifies neural correlates of controlling motor behaviour. We propose the following model for the contextual control of acquired motor behaviour. During acquisition of a sensorimotor program (e.g. association of an external cue with individual finger movements), neural connections between brain areas involved in the processing of motor behaviour are strengthened so as to warrant improved behaviour. The external cues activate not only the visual cortex but also sensorimotor areas. To account for contextual changes, which may require motor behaviour during the appearance of the cue in one situation but not in another, the sensorimotor network must be activated or deactivated (suppressed) depending on these requirements. This *bidirectional* neuronal control is represented by modulation of cortical excitability (TMS), modulation of cortical alpha oscillations, and positive and negative BOLD signal changes in fMRI. These modulations might be controlled by pre-frontal top-down processes.

## Acknowledgements

We wish to thank E. Altenmüller for referral of the patients with musician's dystonia. This research was supported by a grant (SFB 550 / C5) from the Deutsche Forschungsgemeinschaft to CG and a grant from the A. v. Humboldt Foundation (Feodor-Lynen) to FH.

## References

**Andres, F.G., Mima, T., Schulman, A.E., Dichgans, J., Hallett, M., and Gerloff, C.** (1999). Functional coupling of human cortical sensorimotor areas during bimanual skill acquisition. *Brain*, **122**, 855–870.

**Berardelli, A., Rothwell, J.C., Hallett, M., Thompson, P.D., Manfredi, M., and Marsden, C.D.** (1998). The pathophysiology of primary dystonia. *Brain*, **121**, 1195–1212.

Cohen, L.G. and Hallett, M. (1988). Hand cramps: clinical features and electromyographic patterns in a focal dystonia. *Neurology*, **38**, 1005–1012.

Fahn, S. (1988). Concept and classification of dystonia. *Adv Neurol*, **50**, 1–8.

Gerloff, C. and Hummel, F. (2003). Inhibitory control of acquired motor programs in the human brain. *Suppl Clin Neurophysiol*, **56**, 170–174.

Gerloff, C., Richard, J., Hadley, J., Schulman, A.E., Honda, M., and Hallett, M. (1998). Functional coupling and regional activation of human cortical motor areas during simple, internally paced and externally paced finger movements. *Brain*, **121**, 1513–1531.

Greenberg, B.D., Ziemann, U., Cora-Locatelli G. *et al.* (2000). Altered cortical excitability in obsessive-compulsive disorder. *Neurology*, **54**, 142–147.

Hummel, F., Andres, F., Altenmüller, E., Dichgans, J., and Gerloff, C. (2002). Inhibitory control of acquired motor programmes in the human brain. *Brain*, **125**, 404–420.

Hummel, F., Kirsammer, R., and Gerloff, C. (2003). Ipsilateral cortical activation during finger sequences of increasing complexity: representation of movement difficulty or memory load? *Clin Neurophysiol*, **114**, 605–613.

Hummel, F., Saur, R., Lassoga, S. *et al.* (2004) To act or not to act. Neural correlates of executive control of learned motor behavior. *NeuroImage*, **23**, 1391–1401.

Jankovic, J. (2001). Tourette's syndrome. *N Engl J Med*, **345**, 1184–1192.

Ogawa, S., Lee, T.M., Kay, A.R., and Tank, D.W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci USA*, **87**, 9868–9872.

Pfurtscheller, G., Pragenzer, M., and Neuper, C. (1994). Visualization of sensorimotor areas involved in preparation for hand movement based on classification of mu and central beta rhythm in single EEG trials in man. *Neurosci Lett*, **181**, 43–46.

Sadato, N., Campbell, G., Ibanez, V., Deiber, M., and Hallett, M. (1996). Complexity affects regional cerebral blood flow change during sequential finger movements. *J Neurosci*, **16**, 2691–2700.

Sheehy, M.P. and Marsden, C.D. (1982). Writers' cramp—a focal dystonia. *Brain*, **105**(3), 461–480.

Shmuel, A., Yacoub, E., Pfeuffer, J. *et al.* (2002). Sustained negative BOLD, blood flow and oxygen consumption response and its coupling to the positive response in the human brain. *Neuron*, **36**, 1195–1210.

Tagamets, M.A. and Horwitz, B. (2001). Interpreting PET and fMRI measures of functional neural activity: the effects of synaptic inhibition on cortical activation in human imaging studies. *Brain Res Bull*, **54**, 267–273.

Trojano, L., Grossi, D., Linden, D.E. *et al.* (2000). Matching two imagined clocks: the functional anatomy of spatial analysis in the absence of visual stimulation. *Cereb Cortex*, **10**, 473–481.

Part 5

# Apollo's curse—loss of motor control in musicians

Chapter 16

# The end of the song? Robert Schumann's focal dystonia

Eckart Altenmüller

Robert Schumann is one of the most prominent composers of the early romantic period. He was born in Zwickau, Saxony in 1810. Early in his adolescence he displayed extraordinary skills in piano playing and attempted to become a concert pianist. After an initial success, increasing technical difficulties hampered his pianistic progress in the years 1831 and 1832. Finally, he developed a task-specific loss of voluntary control of the middle finger of his right hand. By means of a finger-stretching device, Schumann tried to improve the situation. In parallel, he composed the Toccata, Opus 7, a piano work which allowed high-level virtuoso performance without the use of the middle finger of the right hand. However, from 1833, he almost completely gave up playing the piano literature, but continued to improvise. Robert Schumann was suffering from a focal, task-specific dystonia of the right hand, also referred to as pianist's cramp. This disorder is characterized by a painless loss of skilled motor control in a task-specific context. Its neurobiological origin is seen in maladaptive plasticity of neuronal networks with blurring of afferent and efferent receptive fields of adjacent finger representations in the cerebral cortex and basal ganglia. The general basis of such a blurring may consist of a deficient lateral inhibition of synaptic pathways. Risk factors for developing musician's dystonia are male gender, extensive cumulative practice time, extreme motor workload concerning the temporal and spatial quality of the affected movements and personality traits such as proneness to anxiety and perfectionism. All these factors can be demonstrated in Robert Schumann's early life.

## Preamble

> I offer up myself to make your son Robert, with his talent and imagination, within three years to one of the greatest living piano players, who will play warmer than Moscheles and greater than Hummel.
>
> Friedrich Wieck in a letter to Johanne Christiane Schumann, Robert's mother, on 9 August 1830, quoted according to Burger (1999, p 87)

Who was this promising young pianist who, according to the judgement of a person with intimate knowledge of the piano world, after only 3 years of diligent study could be better than the great virtuosos of the time, Hummel and Moscheles, who competed with the young piano stars Franz Liszt, Frederic Chopin, and Felix Mendelssohn-Bartholdy. Was he a child prodigy? Had he already made a name for himself with his abilities? Had he been noticed through numerous concerts in the great musical centres of Europe? Not at all—Robert Schumann was a young provincial from Saxony, who had studied law, not music, at Heidelberg. He had been noticed in a few private salons because of his free improvising on the piano, but was completely unknown to the music world outside the area. His only noteworthy public appearance took place on 24 January 1830 in the great hall of the Heidelberg 'Museum'. Robert, then 20 years old, played the 'Variations on the Alexander March' by Ignaz Moscheles, a concert piece which was quite popular at the time and also very demanding. The appearance was a tremendous success, although Robert did not come across as a giant at the keys but rather as a shy young man. He was plagued by stage fright. His school friend Theodor Töpken reported:

> To our relief and because he was used to my presence while playing the piano, he asked me to turn pages for him during concerts, to which I agreed, in spite of my anxiety due to the difficulty of the page turning
>
> Töpken quoted according to Burger (1999, p 74)

His piano playing must have been impressive. The Heidelberger society celebrated him, and the accounts of friends and contemporaries speak for themselves. In December 1829, Schumann noted, in a mix of pride and irony, twenty-one 'collected discussions about my piano playing' in his diary; to mention only a few:

> Music teacher Faulhaber: I recognise in him an extraordinary master.—Music director Hoffmann: Bravo: such a thing makes one glad again.—Professor Morstadt: Superb.—Dr. Wüstenfeldt: Heavenly
>
> Quoted according to Robert Schumann *Tagebücher*, in Eismann (1971)

Obviously, Friedrich Wieck's prognosis was not without foundation, but Robert Schumann's piano-playing career was short. One single public appearance, a few evenings of salon music, passing esteem from Heidelberg and Leipzig society and then—silence at the piano. His falling silent is the subject of our research. Which circumstances led to the termination of his pianistic career? What did it have to do with the puzzling suffering of his hand? What creative solutions did Robert seek? And was it really a silencing, or were his hand ailments a necessary aid in the birth of the composer? The biographical study presented here concentrates on the years 1829 to 1833 and on the development of a disorder of the fine motor skills in his right hand that took place during this time. Schumann's psychological crises of

that time only have a place at the boundaries of the discussion. They have been reviewed in the past by several biographers (Ostwald 1985; Lederman 1999; Steinberg 1999). Finally, Schumann's hand ailment will be more precisely classified than has been possible in earlier writings according to new medical insights.

## Childhood scenes

Robert Schumann was born on 6 June 1810 in Zwickau as the fifth child of the bookseller, author, and publisher Friedrich Schumann. His mother, Christiane Schumann, was the daughter of a city surgeon from Zeitz. In the middle-class atmosphere of the home, Robert's musical and literary tendencies were mostly nurtured by his father. At the age of seven, Robert received his first piano lessons from the baccalaureate Kuntsch, who had taught himself to play the organ. Father Schumann purchased a very valuable grand piano and allowed his son to play for him daily after lunch. After only a few years baccalaureate Kuntsch declared himself incapable of giving Robert further lessons, so that as of 1825 Schumann was left to his own devices with regards to his pianistic education. His school friend Flechsig reported:

> I found the young Robert already to be a finished pianist in his 13th year, who also allowed himself to be heard in public concert
>
> Emil Flechsig, *Erinnerungen an Schumann*, around 1875, quoted in Burger (1999, p. 36)

In his 1846 note 'Materials for a Resume', Schumann reported about the years 1820–1826:

> Free improvising (several hours daily) . . . Sick longing for music and piano when I have gone a long time without playing . . . Cello and flute with the city music director Meißner . . . I am strongest at free improvisation . . . entrancing fire of my performance . . . Complete lack of guidance noticeable: Ear training, Technique especially,—Theory
>
> From Eismann (1956, p. 18)

Schumann's great pianistic skill and excellent sight-reading together with his considerable technical shortcomings were noticed by Friedrich Wieck, who, beginning on 1 August 1828, gave lessons to the fledgling law student in Leipzig.

In the beginning, Schumann was required to study scales and finger exercises, much to the chagrin of his roommate Flechsig. The school and college peer made the following remark about this time:

> He took piano lessons from Wieck, who always called him ' "enrage" [maniac] at the piano', and had to do finger exercises like a beginner again, desperately one-sided to listen to
>
> Emil Flechsig, *Erinnerungen an Schumann*, around 1875, quoted in Burger (1999, p. 36)

The readiness for the development of a systematic piano technique was the basis under which Wieck agreed to take on Schumann after his decision to study music in 1831:

> But it is true, that for Robert, the greatest difficulty is the calm, cold, prudent, and lasting conquest of the mechanics as the basis of all piano playing
>
> Letter from Friedrich Wieck's to Schumann's mother 9 August 1832, quoted from Burger (1999, p. 87)

## Upswing

In his time at Leipzig and Heidelberg, Schumann was an extremely ambitious piano player, who practised between 6 and 7 hours daily. Wieck knew how to motivate him, and Schumann was certainly also receptive to his praise:

> Piano lesson with Wieck—Scales and praise [20 November 1828]

but on the same day he entrusts this to his diary:

> One says: people wonder how I can still play scales, seeing as how I play excellently otherwise;—I would not want to receive the first mark from anyone, when I cannot give it to myself

It becomes clear that more than the outside encouragement the actual motivating factor is his high self-determination. The work at the piano then took up a large portion of the daily routine. Schumann visited his relatives in Zwickau for Christmas in 1828, made a lot of music and celebrated a great deal. On 12 December, after ample enjoyment of Burgundy, he fell in a ditch while walking home in the dark. At home the carousing continued—this succinct entry followed:

> dislocation of the arm

Then on 27 December 1828, as the first entry:

> The arm hurts,—the sick arm and the toccata

obviously without any substantial hindrance of the piano playing, since he played four-handed variations from Czerny with his friend Josephine Carus on the same day. Incidentally, the toccata mentioned is not the later drafted Toccata Opus 7, rather a work from the composer Charles Mayer (1799–1862), which Schumann was practising during the Christmas season in 1828. The arm injury appears to remain without repercussion, because in the winter and spring of 1829, Schumann worked intensively, without complaint of pain, on the improvement of his technique:

> Finger exercises from Hummel from 7–12 o'clock [18 February 1829]

Also on the programme were the Hummel A minor Concerto and the Alexander Variations by Ignaz Moscheles—a technically very demanding bravura piece. After the move to Heidelberg in May, the systematic finger exercises and the work on the Alexander Variations is carried on even more intensely. It is from this time that the first drafts of the Toccata later published as Opus 7 originate. Whether or not the pain and disturbances in the fine motor skills which appeared in the winter of 1829/1830, as referred to retrospectively in a letter to Agnes Carus, were related to strain from piano playing is questionable, since no entries regarding this can be found in the regularly kept diaries:

> yes, it came to the point, that when I crossed over the fourth finger, my whole body distorted convulsively and after six minutes of finger exercises, I felt the most unending pain, in short, was broken. In summer I magnificently came back, and I have practiced regularly three or four hours every day with good results.
>
> Letter to E. A. Carus, 25 September 1830 in Kross (1978, p. 28)

It is unusual that Schumann would not mention such devastating symptoms in his diary entries. In these years, the diary entries and the contents of letters to relatives and friends differ more frequently from each other, because Schumann omitted his potentially disturbing plans, especially those concerning his lack of motivation in his career, from the letters. It remains unclear, however, which function the dramatic representation to Agnes Carus of the hand ailment in the winter of 1829/1830 fulfilled.

Undeniably, the progress desired at the time was not always attainable:

> 2 hours finger exercises—Toccata 10 times—finger exercises 6 times—variations 20 times myself—and it still didn't work with the Alexander Variations in the evening—frustration over this—deep frustration [4 January 1830]

Schumann did, however, play the same Alexander Variations with considerable effect 3 weeks later, on 24 January 1830, in his first and only larger public appearance on the piano. In a letter addressed to his brother Julius, but not sent, he wrote:

> You can hardly believe how loved and, really without complimenting myself, respected and honored I am. I have been given the Epithet of a 'Favorite of the Heidelberger Public.' The cornerstone for this was lain naturally by—a concert, in which I played the Alexander Variations by Moscheles. The Bravo and Da capo calling would by God take no end and I became quite warm and took on a glow
>
> Unsent letter to Julius Schumann, dated February 1830, quoted from Schumann (1898, p. 102)

Apparently, even that does not go completely without self-criticism:

> my stumbling at the beginning—the last variation played consummately—unending applause [21 January 1830]

Only 2 days later, Schumann notes after an alcohol excess

my numbed finger [26 January 1830]

This singular remark should not be given too much weight, since only 3 days later 'respectable piano studies' and a successful house concert for an English family in Heidelberg with works by Hummel and the Piano Trio in E-flat major Opus 100 from Schubert were noted, and over several months no mention of pain or strange sensation occurs. The representation of his success at the Heidelberg concert in the letter to his brother and the silence about his piano-playing activities in a letter to his mother on the same day indicate that Schumann is trying to gain approval from his brothers and sisters for his decision, which he had most likely made by this point, to pursue a career as a musician. A few months later, his mother gives her consent for what she considered to be a dangerous step, due not in small part to the good prognosis quoted at the beginning with regards to Robert's career as a pianist.

## In the Night

Schumann returned to Leipzig for a second time in order to continue his studies, and from the very beginning he encountered many more problems. Robert was under enormous pressure to succeed. In order to pacify his mother, he had agreed to a 6-month trial period, after which Wieck would decide whether or not Robert was suited to a career as a pianist. Financial worries forced him to write ever more desperate letters to his home. He finally moved in to a small apartment in the house of his piano teacher Friedrich Wieck. The proximity to the personality of Wieck probably caused a change in Schumann's perception of his teacher. Schumann discovered, much to his disappointment, that Wieck was less interested in his own training than that of his daughter Clara. It remains unclear whether or not Wieck, as early as the end of 1830, had noticed the physical limitations and especially the diminution of Robert Schumann's fine motor skills, and due to this changed his prognosis. In the diary entries beginning in the spring of 1831 there are a noticeably frequent number of entries concerning playing technique, hand position, and relaxation. It is a mix of frustration mingled with the occasional, ever-diminishing experiences of success at the piano. The month of May 1831 seems to be a key month in this regard, in which the technical difficulties of playing are revealed:

Much piano played, Field's Rondeau, Moscheles' third etude, my middle movement—very relaxed in the etudes. [12 May]

Got up early—My soberness is rewarded; played very well—soft, pearl-like tune and fantasy. [13 May]

Piano playing yesterday was quite satisfactory and made progress. Should Wieck be right about my studies? [14 May]

> Piano bad—the Moscheles Etude timid and unsure—Where does that come from? Been playing on it for fourteen days, attentively and persistently studied. [25 May]
>
> The Chopin is going excellently; today is the fifth day on which I have studied four hours each day. If only there will be no relapses! Protect me, my genius and never disappoint me! [5 July]
>
> My dear Robert, don't lose your courage if it is not flowing and going so well, like in the last eight days; practice patience, lift your fingers quietly, hold your hand still and play slowly: and everything must come back together. [9 July]
>
> The piano didn't want to work yesterday; it was as if someone was holding me by the arm. [13 July]
>
> It has been sincerely miserable on the piano the past several days; I cried yesterday with rage. [21 July]
>
> Now I want to proceed in such a manner with my quiet art since I know where it is, it must also be reachable; if only I had no fingers and could play with my heart for others! [14 August]

Schumann worked on a change in his hand position, which temporarily brings an improvement in his attacks:

> It is going quite well with the piano, excellently in the past few days. The flexibility is shocking and the tunes are flowing and progressing as in old times . . . . I am holding my wrist a little higher, approximately like the Belleville [famous female piano virtuoso], even though the graceful, wavelike line is missing. [13 October]

While Clara Wieck and her father were on a concert tour in Paris from September 1831 until April 1832, Robert developed an apparatus to improve the strength of his middle finger, and questionably that of his index finger as well. The doctor and friend Dr. M. Reuter remembered 10 years later in a letter from 1841:

> He mentioned first in his youth that the pointer and middle fingers had noticeably less strength and flexibility than the other fingers. The long term usage of a machine, with which the aforementioned fingers were pulled closer to the back of the hand, led to a state similar to laming, to the degree that they first, only had a weak sense of feeling and secondly, with regards to movement, could no longer be controlled by will.
>
> Rothe (2002)

Schumann christened this device 'Cigar mechanics'. Later he revealed to friends that he had constructed a sling that was attached to the piano, which pulled the middle finger upwards and held it there. The other fingers remained moveable at the keyboard.

> It is going pretty good with the third finger through the 'cigar mechanic'. [7 May 1832]

Now and then, Robert became hopeful:

> Yesterday I composed and played; the new method is the only one that works; it is the one. [11 May 1832]

But this attempt at self-therapy does not help for long:

> The third finger seems really irreparable. [22 May 1832]
> The third finger is completely stiff. [14 June 1832]

After this entry, there are no more clues to be found about Robert's hand ailment in his diaries. The subject seems to be closed. Apparently, Robert had finally given up the plan to become one of the greatest piano virtuosos of the time. This relieving decision was certainly made easier through the first successes of his compositional work—he had switched from one life's goal to another. He formulates this in a letter to his mother, just 2 years after giving up his piano studies:

> . . . Do not worry about the finger! I can still compose, and I would hardly be any happier as a travelling virtuoso—for that, I was spoiled at home. It doesn't bother me when I improvise. Even my old courage to improvise in front of people came back.
>
> Schumann (1898, p. 176)

A special work for piano accompanies the difficult years from 1829–1833, and was quite possibly inspired by the movement disorder. This is the 'study in double notes', the later Toccata Opus 7 (Fig. 16.1). In the figuration of the right hand, the middle finger can be largely left out. The third passages starting in measure 32 are somewhat uncomfortable, but by alternating the ring and index fingers, the upper voice is quite playable. It is not far fetched to say that this is an attempt to find a creative solution to the movement disorder by avoiding the use of the middle finger. Schumann drafted the piece as an etude, and was proud of

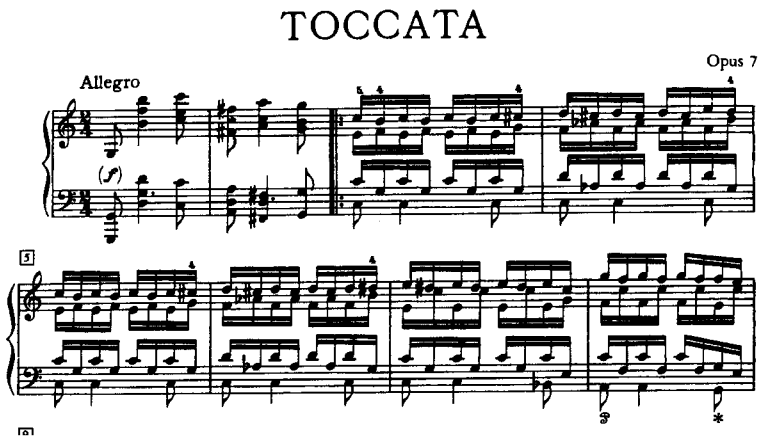



**Fig. 16.1** Beginning of the Toccata Opus 7. This extremely difficult piece can be played without the middle finger of the right hand. (From Hallbergers Prachtausgaben der Romantiker, Verlag Eduard Hallberger, Stuttgart, 1890.)

the horrendous difficulties which the Toccata offered. He wrote to his mother on 2 July 1834:

> Take the attached piece (the Toccata) as a proof of my continuing efforts. Anybody in Zwickau will hardly ever make it.
>
> Schumann (1898, p. 240)

The Toccata Opus 7 is probably the most original of Schumann's early compositions. It is closest to the romantic, demonizing style of Liszt and Paganini. The piece had an enormous impact on his contemporaries. In a review of one of Clara Wieck's piano evenings on 9 November 1834 it was said:

> The last piece made a wonderful impression, a Toccata from Schumann—the work is a mold of originality and novelty, and in spite of its strict style, it worked with a deeply gripping magic on all the listeners.
>
> Boetticher (1984)

## The end of the song?

Even though Schumann had inwardly distanced himself from virtuosity, he still tried, from the spring of 1832 to the summer of 1833, to improve his situation through medical measures. The treatment included rest, diet, electricity, bathing the hand in animal blood, and homeopathy. All new and expensive treatments of the day that were available were—unsuccessfully—utilized (Neumayr 1989). It is little noticed, however, that Schumann did not become completely silent at the piano. Improvization was still part of his daily musical activities. Even in the spring and summer of 1832, the time of the critical worsening of his movement disorder (6 June 1832 'the finger is completely stiff'), Robert tells about incredible experiences at the piano:

> . . . as I came home, close to nine o'clock, I sat down at the piano and it seemed to me that utter flowers and gods were coming out of my fingers, such was the stream of thoughts. [29 May 1832]
>
> How long and overflowing did I fantasize yesterday. [4 July 1832]

It is surely the freedom of the choice of musical conduits which made it possible for him to compensate for his movement problems while improvizing, similar to the case of the Toccata. Furthermore it can be assumed that the degree disturbance of the fine motor skills varied. Schumann still played the piano literature, mostly to get to know particular works. He preferred to perform his own, in part very demanding, pieces in small private contexts. Hieronymus Truhn describes Schumann's piano playing in the year 1837 as quite impressive:

> He moved his fingers with an almost frightening speed, as if ants were crawling around on the piano; he played his own things—I honestly never heard anything else from

> him—with only very little accentuation, but with lots of usage of both pedals. In the last instance, one naturally must not find a lacking of taste; he only played with as much pedal as needed to help a not so grand piano.
>
> Jansen (1883)

## Why?

There has been much speculation about the causes of Robert Schumann's hand problems. Even the finger affected was under discussion. In a letter to Agnes Carus in September 1830, Schumann reveals pain in the ring finger of the right hand in the winter of 1829/1830. In the retrospective report of Robert's friend Dr. Moritz Reuter, the index and middle fingers are discussed. A loss of control in the middle finger of the right hand is described in the diaries beginning in May 1832. On the whole, these precise and repeated diary entries from the critical time period around 1832 speak clearly for an isolated and painless loss of control in the middle finger of the right hand. It is noticeable that the extent of the pain symptoms was overestimated by many biographers. Indications of pain in the hand or arm existed only after an arm injury in December 1828. These were sufficiently explained by a fall after too much wine and did not really affect Schumann after the fact. Several diagnoses are discussed in the literature. Edler (1982/2002) diagnosed a 'complicated tenosynovitis or tendinitis followed by stiffening', Fahrer (1992) an 'augmentation of the extensor tendon of the middle finger through the cigar mechanics', and Franken (1997) the 'result of an arsenic treatment in 1831'. None of these three diagnoses can be applied. A complicated tendonitis would have caused a great amount of pain if it had occurred at close to the same time as the movement disorder. Such pain was not mentioned in the diaries or letters. An augmentation of the extensor tendon due to the 'cigar mechanics' can be eliminated because the causality was reversed: the stretching mechanism was supposed to improve the movement disorder. The completely uncharacteristic symptoms speak against the diagnosis as a result of an arsenic treatment for syphilis. Arsenic poisoning leads to pain, stomach and intestinal cramps, and numbness and paralysis of the feet and hands. A paralysis would also have affected the hand in areas other than piano playing. Neither arsenic treatment nor any symptoms of poisoning were mentioned in the diaries or letters in May 1831, the assumed time frame of infection and potential therapy. After sexual contact with his lover Christel (Charitas), Schumann reports in his diary on 12 May 1831 about a wound on his penis, which could be considered a syphilis infection. It is, however, untypical that the wound hurt, and that treatment with quicksilver or arsenic was not recommended but instead bathing with narcissus water, which was supposed to reduce inflammation. Schumann especially, who tended to be a hypochondriac, would most likely have mentioned a

dangerous arsenic therapy and the resulting side-effects in his diary. Robert Schumann suffered from a task-specific focal dystonia, a musician's cramp. It is the only diagnosis which can sufficiently explain all of the symptoms and the progress of the sickness. Already Merriman *et al.* (1986) suspected a focal dystonia as the underlying pathology, but without providing supporting empirical evidence in detail. Merriman's viewpoint was adopted by Lederman (1999). The entries from his diaries, the attempts to compensate for the disorder by composing piano music extremely well suited for the disorder in question, and the circumstances under which the illness broke out do not leave any doubt concerning this diagnosis. Beyond that, according to new epidemiological studies, Robert Schumann had the typical profile of an 'at-risk patient' (Altenmüller 2003). In summary, Schumann's medical history in the years 1829–1832 is the first convincing documented case of this illness anywhere. Musician's cramp is a neurological disorder characterized by the loss of fine motor control of long-practised skilled movements during instrumental playing. It is related to writer's cramp, but seems to be more frequent in the population at risk. On average, 1 in every 100 musicians in Germany will develop musician's cramp. Men are affected approximately six times more frequently than women. The movement disorder is usually task specific limited to instrumental playing and does not extend to other movements. There are no indications that Schumann's writing skills were affected, even though his editorial and compositional work required many hours' writing each day over a long period of time. There was no pain or



**Fig. 16.2** Presumed hand position of Robert Schumann's dystonic right hand when playing the first bars of the Toccata. Note the middle finger in hyperflexion.

dangerous arsenic therapy and the resulting side-effects in his diary. Robert Schumann suffered from a task-specific focal dystonia, a musician's cramp. It is the only diagnosis which can sufficiently explain all of the symptoms and the progress of the sickness. Already Merriman *et al.* (1986) suspected a focal dystonia as the underlying pathology, but without providing supporting empirical evidence in detail. Merriman's viewpoint was adopted by Lederman (1999). The entries from his diaries, the attempts to compensate for the disorder by composing piano music extremely well suited for the disorder in question, and the circumstances under which the illness broke out do not leave any doubt concerning this diagnosis. Beyond that, according to new epidemiological studies, Robert Schumann had the typical profile of an 'at-risk patient' (Altenmüller 2003). In summary, Schumann's medical history in the years 1829–1832 is the first convincing documented case of this illness anywhere. Musician's cramp is a neurological disorder characterized by the loss of fine motor control of long-practised skilled movements during instrumental playing. It is related to writer's cramp, but seems to be more frequent in the population at risk. On average, 1 in every 100 musicians in Germany will develop musician's cramp. Men are affected approximately six times more frequently than women. The movement disorder is usually task specific limited to instrumental playing and does not extend to other movements. There are no indications that Schumann's writing skills were affected, even though his editorial and compositional work required many hours' writing each day over a long period of time. There was no pain or



**Fig. 16.2** Presumed hand position of Robert Schumann's dystonic right hand when playing the first bars of the Toccata. Note the middle finger in hyperflexion.

deformation associated with the loss of control. Probably beginning in May 1831, Schumann's middle finger involuntarily drew itself in. In a possibly unconscious attempt to compensate for this, Robert Schumann composed the final version of the Toccata, Opus 7. Figure 16.2 shows just such a hand position with a curling middle finger during the playing of the first measures of the Toccata.

Although the neurobiological origins of this disorder have not yet been completely clarified, it is probable that musician's dystonia is in most cases due to dysfunctional (or maladaptive) brain plasticity. Support for this theory comes from a magnetoencephalographic (MEG) study performed in musicians with focal dystonia. Compared with healthy musicians, the dystonics showed a fusion of the digital representations in the somatosensory cortex, reflected in the decreased distance between the representation of the index finger and the little finger when compared with healthy control musicians (for a review see Münte *et al.* and Chapter 18 (2002) and Chapter 19). Such a blurring of receptive fields of the digits may well result in a loss of control, since skilled motor actions are necessarily bound to intact somatosensory feedback input. The pathological mechanism of the blurring may be based on an impaired lateral inhibition of adjacent neuronal networks processing afferent information from single fingers. With a corresponding genetic predisposition, this disintegration can be accelerated by excessive and intensive practising. The emotions play an important role, since the fixation of incorrect motor programs happens especially intensely under the influence of anxiety and stress hormones (Jabusch and Altenmüller 2004). Musicians with focal dystonia very often suffer from anxiety disorders and from perfectionist tendencies (Jabusch *et al.* 2004; see Chapter 17). These personality traits are present long before the condition becomes manifest, and can also be recognized in Robert Schumann. According to more recent epidemiological studies, the high-risk group is considered to be young to middle-aged men who play classical music, with guitarists and pianists being affected most often. Within the instrumental groups, those who spent the longest time practising their instrument on average during their studies were most at risk. Practice time, professional position, and personality traits are all closely connected (Jabusch *et al.* 2005).

The therapy for focal dystonia remains problematic to this day. In Robert Schumann's case, the result of the treatment with the 'cigar mechanics' and the other attempts at therapy is clear—a career as a pianist remained unattainable for him. On the whole, though, his problems must be seen as a mild and non-progressive form of the condition since he was still quite able to play the piano. Improvization at the piano must have been a return to the roots of music for him, connected to an inner freedom from fear, doubt, and from the extreme pressure of pianistic perfection. The energetic diversion of his creative potential

to composition shows the tremendous will for creation, which had already set him apart as an adolescent. For us, Schumann's decision to follow a career as a composer was a blessing, because it allowed his creative talent be developed to masterful perfection.

## References


**Altenmüller, E.** (2003). Focal dystonia: advances in brain imaging and understanding of fine motor control in musicians. *Hand Clin*, **19**, 1–16.

**Boetticher, W.** (1984). Rezension im 'Kometen' Leipzig 1834, 26 September. *Robert Schumann's Klavierwerke*, Part II, p 22. Heinrichshofen-Verlag, Wilhelmshafen.

**Burger, E.** (1999). *Robert Schumann*, p. 87. Schott Publishers, Mainz.

**Edler, A.** (1982). *Robert Schumann und seine Zeit*. Laaber-Verlag, Düsseldorf (reprinted 2002).

**Eismann, G.** (1956). *Robert Schumann—Ein Quellenwerk*, Vol. 1, p. 18. VEB Breitkopf und Härtel, Leipzig.

**Eismann, R.** (1971). *Robert Schumann 'Tagebücher'*, Vols 1–5. Stroemfeld/RoterStern Verlag.

**Fahrer, M.** (1992). The right hand of Robert Schumann. *Ann Hand Surg*, **11**, 237–240.

**Franken, F.H.** (1997). Robert Schumann in der Irrenanstalt in Bonn Endenich. Zum aufgefundenen ärztlichen Verlaufsbericht 1854–1856 von Doktor Franz Richarz. In: *Brahms-Studien*, Vol. 11, pp. 107–120. Hans Schneider-Verlag, Tutzing.

**Jabusch, H.C. and Altenmüller, E.** (2004). Anxiety as an aggravating factor during onset of focal dystonia in musicians. *Med Probl Perform Art*, **19**, 75–81.

**Jabusch, H.C., Müller, S.V., and Altenmüller, E.** (2004). High levels of perfectionism and anxiety in musicians with focal dystonia. *Mov Disord*, **19**, 990–991.

**Jabusch, H.C., Zschucke, D., Schmidt, A., Schuele, S., and Altenmüller, E.** (2005) Focal dystonia in musicians: clinical manifestation and long term outcome in 144 patients. *Mov Disord*. **20**, 1623–1626.

**Jansen, F.G.** (1883). *Die Davidsbündler*, pp. 74–75. Breitkopf & Härtel, Leipzig.

**Kross, S.** (ed.) (1978). *Robert und Clara Schumann, Briefe und Notizen*, p. 28. Bonn.

**Lederman, R.** (1999). Robert Schumann. *Semin Neurol*, **19**, 17–24.

**Merriman, L., Newmark, J., Hochberg, F.H., Shahani, B., and Leffert, R.** (1986). A focal movement disorder of the hand of six pianists. *Med Probl Perform Art*, **1**, 17–19.

**Münte, T.F., Altenmüller, E., and Jäncke, L.** (2002). The musician's brain as a model of neuroplasticity. *Nat Rev Neurosci*, **3**, 473–478.

**Neumayr, A.** (1989). *Musik und Medizin*, Vol. 2, pp. 151–183. Wiener Verlag, Himberg.

**Ostwald, P.** (1985). *Schumann—the inner voices of a musical genius*. Northeastern University Press, Boston, MA.

**Rothe, H.J.** (2002). Neue Dokumente zur Schumann Forschung, p. 5 quoted from Arnfried Edler. In: *Robert Schumann und seine Zeit*, p. 297. Laaber-Verlag, Laaber.

**Schumann, R.** (1898). *Jugendbriefe* (ed. C. Schumann). Breitkopf & Härtel, Leipzig.

**Steinberg, R.** (1999). Robert Schumann's illness in the eyes of his pathographers. In: E.J. Engstroem, M.M. Weber, and P. Hoff (eds) *Knowledge and power. perspectives in the history of psychiatry*, pp. 65–78. Verlag für Wissenschaft und Bildung, Berlin.

Chapter 17

# Epidemiology, phenomenology, and therapy of musician's cramp

Hans-Christian Jabusch and Eckart Altenmüller

## Introduction

Focal dystonia in musicians, also known as musician's cramp, is a task-specific movement disorder which presents itself as a painless muscular incoordination or loss of voluntary motor control of extensively trained movements while a musician is playing the instrument (Jankovic and Shale 1989; Lederman 1991; Brandfonbrener 1995; Frucht *et al.* 2001; Altenmüller 2003). For those who are affected focal dystonia is highly disabling, and in many cases terminates musical careers. According to estimations, 1% of all musicians are affected (Altenmüller 2003). In this chapter the phenomenology and epidemiology of musician's dystonia are addressed on the basis of data from 144 affected musicians diagnosed and followed-up at the Institute of Music Physiology and Musicians' Medicine of the Hannover University of Music and Drama. Data will be compared with those of other reports and discussed with regard to risk factors for the development of musician's dystonia. Treatment strategies and results will be outlined based on a retrospective inquiry and self-evaluation of affected musicians treated with currently available therapies.

All 144 patients were professional musicians diagnosed with focal dystonia at our out-patient clinic between 1994 and 2001. The diagnostic procedure during their first visit included a complete neurological examination as well as visual inspection while patients were playing their instruments. None of the patients were suffering from secondary dystonias or other neurological disorders. For assessment of the phenomenological and epidemiological data, a retrospective chart review was performed. Data were statistically analysed using $\chi^2$-tests and Fisher's exact tests to identify disproportionate frequencies in the clinical manifestation of subgroups of patients. The level of significance was set at $P < 0.05$.

## Demographics, epidemiology, and phenomenology

The patient group consisted of 116 male (81%) and 28 female (19%) musicians. The mean age at onset of symptoms was 33 years (range 17–63 years) and the

mean duration of dystonic symptoms at the time of the first visit was 5.1 years (range 0.1–28 years). At the time of the onset of symptoms, 74 patients (51%) had professional positions as soloists and 24 (17%) were tutti players in orchestras, 25 patients (17%) had teaching positions, and 21 (15%) were students. Concerning the genre, 137 patients (95%) were classical musicians, 2 patients (1%) were jazz musicians, and 5 patients (4%) were pop musicians. Nine patients (6%) had a history of either writer's cramp or musician's dystonia in their family. The distribution of instrument groups was as follows: 22 patients (15%) were bowed string players, 40 patients (28%) were keyboard instrumentalists, 29 patients (20%) were playing plucked instruments, 37 patients (26%) were woodwind players, and 16 patients (11%) were brass players. The distribution of individual instruments is displayed in Fig. 17.1. In order to compare the distribution of instruments among dystonic musicians with that in a normal population of instrumentalists, the respective numbers of members in the German Orchestra Union were used according to the March 2004 statistics of the German Orchestra Union (representing over 90% of German orchestral musicians). In our patient group, 20 high string players (violin, viola) and 2 low string players (violoncello, double-bass) were affected. This ratio was higher than expected according to the distribution of 3810 high and 1464 low string players in the German Orchestra Union ($\chi^2 = 3.81$, d.f. $= 1$, $P = 0.05$). One hundred and twenty-four patients were suffering from focal upper limb dystonias (86%), which presented themselves mostly as hand dystonias in the

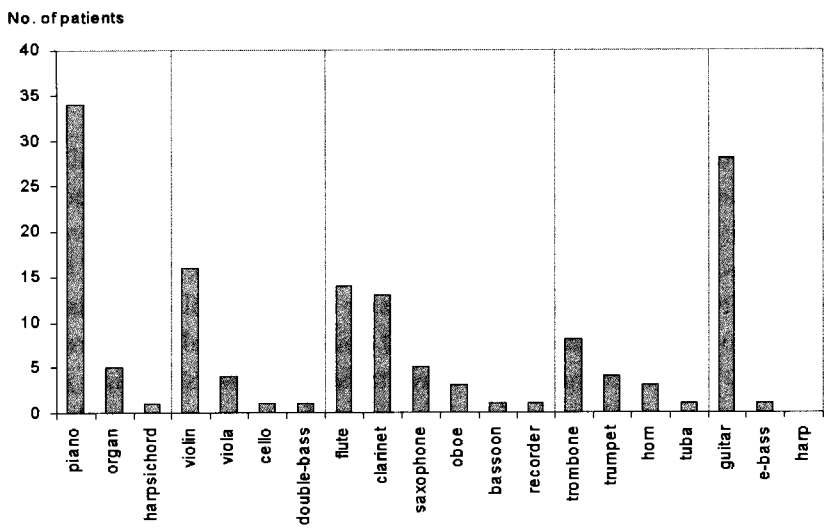



**Fig.17.1** The instruments played by 144 patients with musician's dystonia. Instrument families are separated by vertical lines.

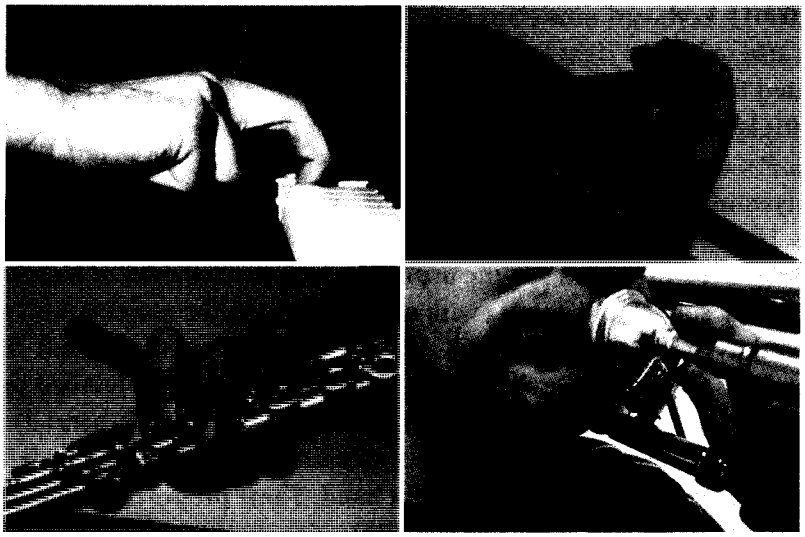


**Fig.17.2** Typical patterns of dystonic posture in a pianist, a violinist, a flutist, and a trombone player.

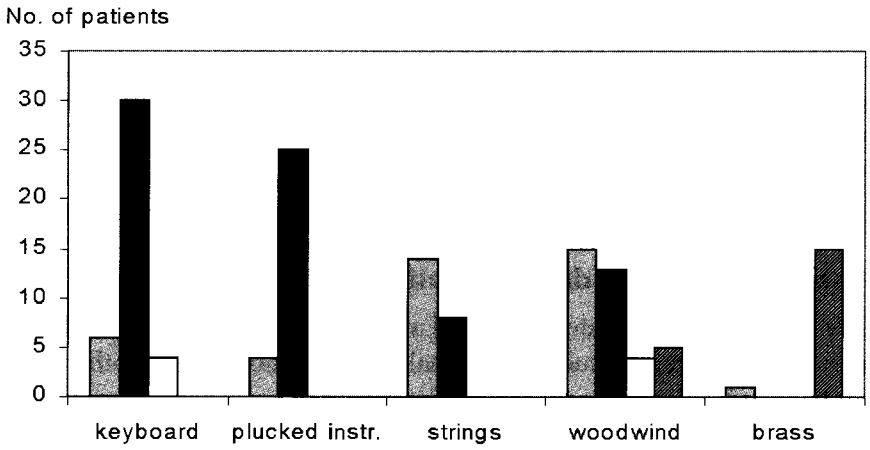



**Fig.17.3** Localization of musician's dystonia as distributed in different instrument families: grey bars, left hand affected; black bars, right hand affected; white bars, both hands affected; hatched bars, embouchure dystonia.

typical manner with involuntary cramping of one or more fingers while patients were playing; one trombone player displayed dystonic movement patterns in his left upper arm. Embouchure dystonia was present in 20 patients (14%). Typical patterns of dystonic posture are displayed in Fig. 17.2. Details of the localization of focal dystonia in different instrument families are given in Fig. 17.3. $\chi^2$ -tests of 115 patients with unilateral hand dystonia revealed a different laterality of focal hand dystonia in different instrument families ($\chi^2 = 23.5$, d.f. $= 3$,



**Fig.17.2** Typical patterns of dystonic posture in a pianist, a violinist, a flutist, and a trombone player.

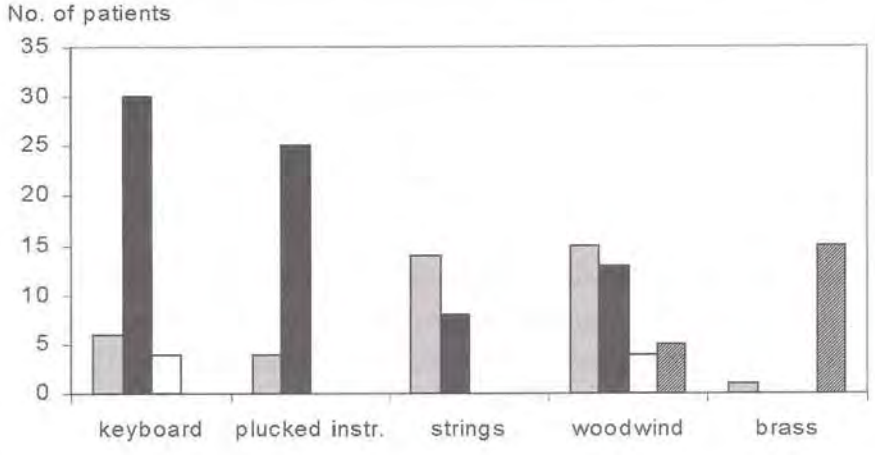



**Fig.17.3** Localization of musician's dystonia as distributed in different instrument families: grey bars, left hand affected; black bars, right hand affected; white bars, both hands affected; hatched bars, embouchure dystonia.

typical manner with involuntary cramping of one or more fingers while patients were playing; one trombone player displayed dystonic movement patterns in his left upper arm. Embouchure dystonia was present in 20 patients (14%). Typical patterns of dystonic posture are displayed in Fig. 17.2. Details of the localization of focal dystonia in different instrument families are given in Fig. 17.3. $\chi^2$ -tests of 115 patients with unilateral hand dystonia revealed a different laterality of focal hand dystonia in different instrument families ($\chi^2 = 23.5$, d.f. $= 3$,

$P < 0.0001$): keyboard musicians and those with plucked instruments were primarily affected in the right hand and string players in the left hand (Bonferroni-corrected follow-up $\chi^2$-tests; keyboard vs. string players: $\chi^2 = 13.3$, d.f. = 1, $P < 0,01$; keyboard musicians vs. musicians with plucked instruments: $\chi^2 = 13.6$, d.f. = 1, $P < 0,01$). No such lateralization was seen in woodwind players. Bilateral hand dystonia was seen only in keyboard instrumentalists and in woodwind players. Handedness was categorized as right-handed/ left-handed/ambidextrous according to the preferences in the use of hands in the activities of handwriting, eating with a spoon, and cutting with scissors. Among the patients with unilateral hand dystonia there were 108 right-handed patients and 5 left-handed patients; two patients were ambidextrous. Seventy-five right-handed patients had right-hand dystonia and 33 right-handed patients had left-hand dystonia. This pattern was inverse in left-handed patients: in four of these the left hand was affected and in one patient the right hand. These proportions were significantly different (Fisher's exact test; $P < 0.05$).

Ninety-five patients (66%) had dystonic symptoms only while they were playing their instrument, whereas 49 patients (34%) had additional difficulties in other activities such as writing on the computer keyboard or in everyday activities. Local pain preceded focal dystonia in 13 (9%) patients: 7 patients (5%) experienced local pain only before onset of dystonic symptoms while 6 (4%) patients had pain in the affected region before and after onset of dystonic symptoms. Twenty-five (17%) patients reported local pain only after onset of dystonic symptoms, whereas 106 (74%) patients never experienced pain in the affected region.

## Risk factors for the development of musician's dystonia

The demographic data demonstrated a preponderance of male musicians in the sample of the present study, with a male:female ratio of 4:1. In other reports the male:female ratio of patients with musician's dystonia was between 2:1 (Lederman 1991) and 5:1 (Lim and Altenmüller 2003). Since female musicians form the majority of the musicians in Germany, the latter ratio was corrected to 6:1 (Lim and Altenmüller 2003). Similarly, this effect has to be taken into account for the male:female ratio of 4:1 of the present sample of patients, which additionally underlines the increased risk of male musicians developing focal dystonia. A positive family history of dystonia, as seen in 6% of our patient group, has already been described in other reports as a risk factor for development of musician's dystonia (Jankovic and Shale 1989; Altenmüller 2003; Lim and Altenmüller 2003). The mean age at onset of dystonic symptoms (33 years) observed in the present study harmonizes with other studies demonstrating a peak in the fourth decade of life (Lederman 1991; Brandfonbrener 1995; Brandfonbrener and Robson 2002; Altenmüller 2003).

Our results demonstrated a correspondence between the instrument group and the localization of focal dystonia. This is obvious for brass players who suffer predominantly from embouchure dystonia. Additionally, among the musicians with unilateral hand dystonia, the correspondence between the instrument group and the localization of focal dystonia was significant. Keyboard musicians (piano, organ, harpsichord) and those with plucked instruments (guitar, e-bass; no harpists) were primarily affected in the right hand. All these instruments are characterized by a higher workload in the right hand. Correspondingly, bowed string players who have a higher workload and complexity of movements in the left hand were predominantly affected in the left hand. These observations agree with other reports (Tubiana and Chamagne 2000; Brandfonbrener and Robson 2002; Schuele and Lederman 2004). Moreover, focal hand dystonia in both hands was only seen in those musicians with instruments requiring similar movement patterns in both hands (although with a different workload), such as woodwind and keyboard instruments. Within the string family, musicians who were playing the high string instruments (violin, viola) were more often affected than those with low string instruments (violoncello, double-bass). The observed occurrence in high string players was significantly higher than expected according to the overall distribution of high and low string players among German orchestral musicians. The demands on spatial sensorimotor precision required for violin and viola playing are higher than for the low string instruments. Thus, those musicians with instruments requiring highest spatial sensorimotor precision were primarily affected by dystonia. Furthermore, in our sample of musicians with unilateral hand dystonia a context was observed between the handedness and the side affected by dystonia. The majority of right-handed patients was affected in the right hand and the majority of left-handed patients was affected in the left hand. Although the number of left-handed patients was small, this difference was significant and indicated that the hand with a higher total workload tended to be affected. It is noteworthy that there was no string player in the group of left-handed patients. The musical genre of the overwhelming majority of patients was classical music. In contrast to pop or jazz music with improvised structures and great freedom of interpretation, musical constraints are most severe in classical music. The latter requires a maximum of temporal accuracy in the range of milliseconds which is scrutinized by the performing musician as well as by the audience at any moment of the performance. This, as a consequence, combines the situation of public performance in classical music with a high level of social pressure: the gap between success and failure is minimal in this genre.

Taken together, epidemiological findings indicate that the workload of the respective body part, the complexity of movements, and the degree of spatial and

temporal sensorimotor precision as well as the level of social constraints associated with the musical performance were related to the occurrence of musician's dystonia.

## Pathophysiological findings

The pathophysiology of focal dystonia remains unclear. Beside of alterations in the circuitry of the basal ganglia (Naumann *et al.* 1998; Preibisch *et al.* 2001) and dysfunctional plasticity in the sensory thalamus (Lenz and Byl 1999; Hua *et al.* 2000), there is growing evidence for abnormal cortical processing of sensory information as well as degraded representation of motor function in patients with focal dystonia (Deuschl *et al.* 1995; Hallett 1998; Elbert *et al.* 1998; Bara-Jimenez *et al.* 1998; Pujol *et al.* 2000). Cortical inhibition was found to be decreased in dystonic patients (Deuschl *et al.* 1995; Toro *et al.* 2000; Hummel *et al.* 2002). In animal studies, repetitive movements induced symptoms of focal hand dystonia and a distortion of the cortical somatosensory representation in monkeys (Byl *et al.* 1996) suggesting that practice-induced alterations in cortical processing may play a role in focal hand dystonia. Recently, behavioural treatment in patients with focal hand dystonia resulted in clinical amelioration and in reversed distortions of the somatosensory map, giving additional evidence for a practice-induced alteration of the cortical map (Candia *et al.* 2003). The importance of practice and the use of the limb in musician's dystonia was additionally underlined by the finding of improvement in symptoms after immobilization of the affected limb in dystonic musicians (Priori *et al.* 2003; see also Chapter 19). An impressive finding was reported by Rosenkranz and colleagues who used focal vibratory input to individual hand muscles to produce sensory input whilst the excitability of corticospinal outputs to the vibrated and other hand muscles was evaluated with transcranial magnetic stimulation (Rosenkranz *et al.* 2005, see also Chapter 18). In musicians with dystonia, focal vibratory input to hand muscles resulted in abnormally suppressed short-latency intracortical inhibition of all other hand muscles examined, independent of their functional connectivity with the vibrated muscles. In healthy non-musicians, such vibratory input increased short-latency intracortical inhibition of neighbouring muscles. In healthy musicians, however, vibratory input resulted in a suppressed short-latency intracortical inhibition only of those neighbouring muscles functionally connected with the vibrated muscles. In patients with writer's cramp, short-latency intracortical inhibition of neighbouring muscles was unchanged. It was concluded that the pathophysiology of musician's dystonia differs from that of writer's cramp. Furthermore, it was hypothesized that long hours of practising complex movements first produce the modulation of

sensorimotor interaction seen in healthy musicians, and that this later progresses into the non-focal pattern of short-latency intracortical inhibition in musician's dystonia (Rosenkranz *et al.* 2005).

The findings of an increased workload and complexity of movements as well as increased spatial and temporal sensorimotor requirements being related to musician's dystonia are in agreement with the hypothesis of a pathomechanism which includes practice-induced alterations. A higher workload in affected musicians might additionally be reflected by the observation of a more perfectionist attitude in dystonic musicians—even before onset of dystonia—compared with other musicians (Jabusch *et al.* 2004*a*) and by the finding of a preponderance of soloists (51%) in our sample of dystonic musicians. A perfectionist attitude as well as an outstanding professional position might point at a different and more intense working behaviour resulting in a higher workload in those musicians who develop focal dystonia.

Local pain and intensified sensory input for various reasons like nerve entrapment or trauma or overuse have been described as potential triggers of dystonia (Jankovic and Shale 1989; Lederman 1991; Charness *et al.* 1996; Altenmüller 2003). The parallels of abnormal cortical processing of sensory information and cortical reorganization in patients with focal dystonia and those with chronic pain have already been reported (Flor *et al.* 1997; Tinazzi *et al.* 2000; Jabusch *et al.* 2004*a*). Local pain preceded focal dystonia in 9% of the patients in the present study.

## Possible interaction between predisposition and extrinsic and intrinsic factors

In view of the neurophysiological, epidemiological, and psychological findings in musicians with focal dystonia, predisposing factors have been described, such as male gender (Lederman 1991; Altenmüller 2003; Lim and Altenmüller 2003) as well as a positive family history (Jankovic and Shale 1989; Altenmüller 2003; Lim and Altenmüller 2003) which might constitute a particular vulnerability or susceptibility to musician's dystonia. Additional extrinsic and intrinsic factors may trigger the manifestation of musician's dystonia on the basis of a given susceptibility. Intrinsic triggering factors are physical disorders resulting in local pain and/or intensified somatosensory input (trauma, nerve entrapment, overuse injury) (Lederman 1991; Charness *et al.* 1996; Altenmüller 2003) as well as psychological conditions such as perfectionism and anxiety which were found to be related to musician's dystonia (Jabusch *et al.* 2004*a*). According to Rosenkranz *et al.* (2005), intracortical disinhibition in musician's dystonia might be an acquired, practice-induced phenomenon. Extrinsic triggering

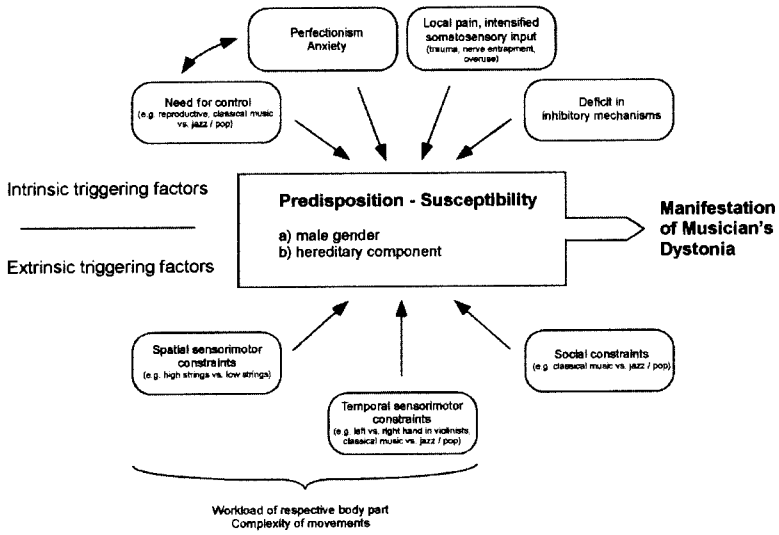



**Fig. 17.4** The possible interaction between predisposition and intrinsic and extrinsic triggering factors in the manifestation of musician's dystonia. Further explanations are given in the text.

factors, according to epidemiological findings (Tubiana and Chamagne 2000; Brandfonbrener and Robson 2002; Schuele and Lederman 2004), are spatial and temporal sensorimotor constraints as well as musical and social constraints typical of the performance situation in classical music. The possible interaction between predisposition and intrinsic and extrinsic triggering factors in the manifestation of musician's dystonia is displayed in Fig. 17.4.

## Treatment strategies and results

For a survey of treatment results, 160 consecutive professional musicians and music students with musician's dystonia were recruited (Jabusch *et al.* 2005). Treatment strategies included: medication with trihexyphenidyl; injection therapy with botulinum toxin; ergonomic changes; pedagogical retraining; non-specific exercises on the instrument. Outcome data were assessed using a standardized questionnaire which was mailed to patients. Those patients who did not respond were contacted by phone, and a standardized interview was conducted using the questionnaire. Patients were asked to estimate (1) the cumulative treatment response on a four-step scale (free of symptoms, improved, no change, deteriorated); (2) the treatment response to individual therapies (improvement, no effect, deterioration). They were asked about (3) the nature and duration of therapies not applied by the authors (e.g. exercises on the instrument) and (4) their musical activity and professional situation at that time.

Outcome data for the aforementioned 144 patients could be obtained. Data were analysed using $\chi^2$-tests for outcome measures and multiple regression analyses for predictability of treatment responses. Percentage rates were calculated based on the total number of available answers.

## Long-term outcome

The mean age at the time of the completion of the questionnaire was 42 years (range 23–69 years). Outcome was revealed on average 8.4 years after onset of dystonic symptoms (range 1–30 years) and 3.4 years after the initial visit (range 1–8 years). At the time of the survey, two patients (1%) were free of symptoms, 75 patients (52%) reported an alleviation of symptoms as compared with the time of their first visit, 50 patients (35%) did not notice any changes and 16 patients (11%) reported a deterioration. Regardless of treatment, a smaller fraction of patients with embouchure dystonia showed an improvement (3 of 20 patients) compared with those with limb dystonia (62 of 123 patients; $\chi^2 = 8.7$, d.f. $= 1$, $P < 0.01$). A change of profession was reported by 35 patients (29%). A larger fraction of music students changed their profession (12 of 21 patients, 57%) compared with those who were already in their professional careers at onset of symptoms (23 of 100 patients; $\chi^2 = 9.84$, d.f. $= 1$, $P < 0.01$) (Jabusch *et al.* 2005).

## Trihexyphenidyl

Trihexyphenidyl was given to patients with all kinds of musician's cramps when no contraindication was present. Trihexyphenidyl was administered beginning with 1 mg/day, and the dosage was slowly increased until therapeutic effects were observed. Adjustment of the dosage was made depending on beneficial effects and side-effects. Patients with side-effects and no improvement of symptoms were withdrawn from trihexyphenidyl treatment.

Trihexyphenidyl was applied in 69 patients (48% of all) with an average maximum dosage of 11 mg/day (range 1–30 mg/day). Within 2 months the medication was interrupted in 20 patients due to side-effects or unsatisfactory treatment response. The other 49 patients (34% of all) received trihexyphenidyl on average for 16 months (range 3–67 months). The most frequent side-effects were dry mouth (29 patients; 42% of patients receiving trihexyphenidyl), tiredness (15 patients; 22%), dizziness (13 patients; 19%), agitation (10 patients; 14%), memory impairment (8 patients; 12%), drowsiness (7 patients; 10%), depression (7 patients; 10%), loss of concentration (5 patients; 7%), nausea (4 patients; 6%), hyperkinesia (3 patients; 4%), impaired visual accommodation (2 patients; 3%), and tremor (2 patients; 3%). The average minimum dosage at which side-effects occurred was 9 mg/day (range 1–20 mg/day). Due to

side-effects, 16 patients (23%) were withdrawn from trihexyphenidyl. Side-effects were not related to age or gender. After trihexyphenidyl, an improvement was reported by a total of 23 patients (33%) and by 19 patients with limb dystonia who received trihexyphenidyl for more than 2 months (42% of these; average duration 26 months). Patients' rating of treatment results is displayed in Fig. 17.5. None of five patients with embouchure dystonia showed an improvement after trihexyphenidyl. The occurrence of side-effects was not significantly different in patients who reported an improvement after trihexyphenidyl (side-effects present in 22 out of 23 patients) than in patients with no improvement (side-effects present in 32 out of 44 patients). The average maximum dosage of trihexyphenidyl was 11 mg/day (range 4–30 mg/day) in patients who had an improvement and 11 mg/day (range 1–20 mg/day) in patients with no improvement. Multiple regression analysis revealed a model predicting 49% of the variance of patients' rating of trihexyphenidyl effects ($R$-square = 0.52; $R$-square adjusted = 0.49). The following variables predicted a positive rating: limb localization of dystonia (versus embouchure dystonia: $\beta = 0.59$; d.f. = 1; $F = 37$; $P < 0.001$), a high maximum dosage of trihexyphenidyl ($\beta = 0.28$; d.f. = 1; $F = 6.8$; $P = 0.01$), a low number of concomitant treatments ($\beta = -0.35$; d.f. = 1; $F = 11$; $P < 0.01$). The variables age, gender, localized versus non-localized dystonia (localized dystonia affecting one or two fingers versus non-localized dystonia), type of dystonia (flexion/extension/embouchure), task specificity at first visit, and duration of dystonia were not sufficient predictors of the rating of the effects of trihexyphenidyl.

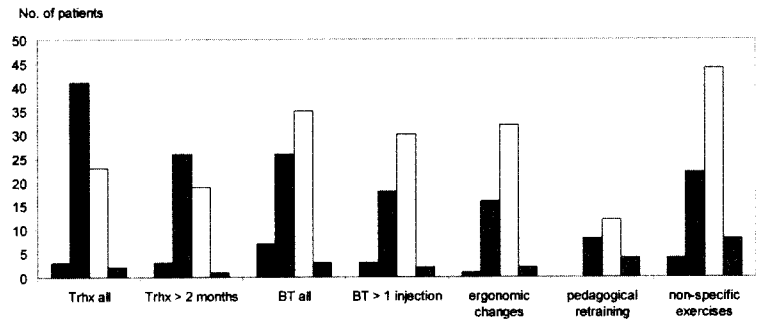



**Fig.17.5** Patients' rating of treatment results: black bars, deterioration of dystonic symptoms; grey bars, no change in dystonic symptoms; white bars, alleviation of dystonic symptoms; hatched bars, no answer. Trihexyphenidyl (Trhx): results are given separately for all patients who received Trhx versus for those who received Trhx for more than 2 months. Botulinum toxin (BT): results are given separately for all patients who received BT versus for those who received more than one injection. (From Jabusch *et al.* 2005.)

side-effects, 16 patients (23%) were withdrawn from trihexyphenidyl. Side-effects were not related to age or gender. After trihexyphenidyl, an improvement was reported by a total of 23 patients (33%) and by 19 patients with limb dystonia who received trihexyphenidyl for more than 2 months (42% of these; average duration 26 months). Patients' rating of treatment results is displayed in Fig. 17.5. None of five patients with embouchure dystonia showed an improvement after trihexyphenidyl. The occurrence of side-effects was not significantly different in patients who reported an improvement after trihexyphenidyl (side-effects present in 22 out of 23 patients) than in patients with no improvement (side-effects present in 32 out of 44 patients). The average maximum dosage of trihexyphenidyl was 11 mg/day (range 4–30 mg/day) in patients who had an improvement and 11 mg/day (range 1–20 mg/day) in patients with no improvement. Multiple regression analysis revealed a model predicting 49% of the variance of patients' rating of trihexyphenidyl effects ($R$-square = 0.52; $R$-square adjusted = 0.49). The following variables predicted a positive rating: limb localization of dystonia (versus embouchure dystonia: $\beta = 0.59$; d.f. = 1; $F = 37$; $P < 0.001$), a high maximum dosage of trihexyphenidyl ($\beta = 0.28$; d.f. = 1; $F = 6.8$; $P = 0.01$), a low number of concomitant treatments ($\beta = -0.35$; d.f. = 1; $F = 11$; $P < 0.01$). The variables age, gender, localized versus non-localized dystonia (localized dystonia affecting one or two fingers versus non-localized dystonia), type of dystonia (flexion/extension/embouchure), task specificity at first visit, and duration of dystonia were not sufficient predictors of the rating of the effects of trihexyphenidyl.

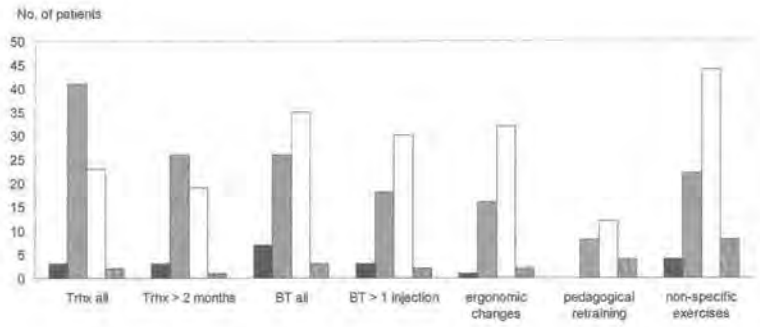



**Fig.17.5** Patients' rating of treatment results: black bars, deterioration of dystonic symptoms; grey bars, no change in dystonic symptoms; white bars, alleviation of dystonic symptoms; hatched bars, no answer. Trihexyphenidyl (Trhx): results are given separately for all patients who received Trhx versus for those who received Trhx for more than 2 months. Botulinum toxin (BT): results are given separately for all patients who received BT versus for those who received more than one injection. (From Jabusch *et al.* 2005.)

A limb localization of dystonia, a high maximum dosage, and a low number of concomitant treatments were found to predict a positive outcome of treatment with trihexyphenidyl. The low number of concomitant treatments may be a consequence of a successful treatment with trihexyphenidyl in some patients which limits its value as a predictor variable. A high maximum dosage was a sufficient, but not necessary, condition for an improvement. One patient who tolerated a maximum dosage of 30 mg/day had an improvement, but in other patients an improvement was seen after low dosages. However, there was no patient with high maximum dosages in the group of non-responders to trihexyphenidyl. There was no difference in the occurrence of side-effects between responders and non-responders. All patients with embouchure dystonias were non-responders. As a conclusion, a treatment attempt with trihexyphenidyl appears to be an option for musicians with limb dystonia when no contraindications are present (Jabusch *et al.* 2005).

## Botulinum toxin

Botulinum toxin A (BT) injections were recommended only to those patients in whom primary dystonic movements could be clearly distinguished from secondary compensatory movements. A lyophilized botulinum toxin A powder (Dysport®, Ipsen Ltd, Berkshire, UK) was injected using an electromyographic (EMG)-guided technique (Karp *et al.* 1994; Schuele *et al.* 2005). Target muscles were identified by visual inspection of the dystonic movements while patients were playing their instruments. In patients with flexion or extension dystonia of individual fingers, injections in the forearm muscles were preferred. Additional injections in the hand muscles were performed in patients with dystonic flexion in the metacarpophalangeal (MCP) joints and an extension component in the proximal interphalangeal (PIP) and distal interphalangeal (DIP) joints.

Botulinum toxin injections were applied in 71 patients (49%), three of whom had embouchure dystonia. Eighteen patients had only one injection due to unsatisfactory treatment results. Five patients who reported an improvement were not satisfied enough to continue BT treatment. Fifty-three patients (37%) received more than one injection (average 5.7 injections per patient, range 2–25) with an average treatment duration of 16 months (range 1–58 months) and an average interval of 4.2 months (range 1–22 months) between treatment sessions. The average dosage per treatment session was 128 units (range 9–428). The most injected muscles were the flexor digitorum superficialis and the flexor digitorum profundus (50 patients each; 70% of patients treated with BT), followed by the flexor carpi radialis (13 patients; 18%), the flexor pollicis longus, the extensor digitorum and the extensor indicis (7 patients each; 10%), and the interosseus palmaris (5 patients; 7%). The average dosage per muscle group was 112 units

(range 88–150) in the upper arm and shoulder muscles, 38 units (range 5–85) in forearm extensors, 65 units (range 10–175) in forearm flexors, and 26 units (range 5–84) in hand muscles. After BT, an improvement was reported by 35 patients (49%) and by 30 patients who received more than one injection (57% of these; average duration 26 months). Of the latter group, 21 of 33 patients with injections given only in the forearm muscles reported an improvement (64%), and 7 of 13 patients with injections in the forearm muscles and hand muscles (54%) had an improvement: two out of two keyboard players, two of four woodwind players, and three of four guitarists with the right hand affected. No improvement was seen in the affected left hands of one guitarist and two string players after injections in the forearm muscles and hand muscles. Neither of two patients with injections in the upper arm/shoulder muscles experienced an improvement and none of three patients with embouchure dystonia experienced an improvement after BT. Multiple regression analysis did not reveal any of the following variables as predictors of outcome after BT: gender, age, duration of dystonia, localized versus non-localized dystonia, type of dystonia, target muscles, mean BT dosage, task specificity at first visit, number of concomitant treatments. Positive rating of BT effects was correlated to the number of BT treatment sessions (Spearman $r = 0.37$, $P < 0.01$) and the duration of BT treatment (Spearman $r = 0.56$, $P < 0.01$).

According to these results, BT treatment was successful in those patients in whom primary dystonic movements could be clearly distinguished from secondary compensatory movements. This was difficult when compensatory movements were more pronounced than primary dystonic movements. The best outcome was reported after injections in the forearm muscles. Additional injections in hand muscles were useful in patients with a dystonic flexion in the MCP joints and an extension component in the PIP and DIP joints. Botulinum toxin injections in hand muscles appear to be recommendable only in instrumentalists with little lateral finger motion such as in woodwind players and guitarists (right hand). Keyboard players may benefit from this option when they avoid repertoire requiring a wide hand span and extreme lateral finger motion. Botulinum toxin injections in hand muscles were not successful in musicians who needed to perform lateral finger movements such as in the left hands of one guitarist and of two string players. Musicians with embouchure dystonia and with dystonia affecting the upper arm and shoulder muscles did not benefit from BT treatment (Jabusch *et al.* 2005).

## Ergonomic changes

Ergonomic changes were recommended whenever applicable. The aim of ergonomic changes was either a blocking of dystonic movements, for example by

attaching splints to the affected fingers, or circumvention of dystonic movements. This was achieved by modifications of the instrument, e.g. the repositioning of individual keys in woodwind instruments or the replacement of ring keys by plateau keys (Altenmüller 2003). Support systems (belts, stands, tripods) were recommended when the dystonic hand had to carry the instrument while playing such as in oboe, clarinet, and bassoon players. Ergonomic changes were made in 51 patients (35% of all), 32 patients (63% of these) experienced improvement and used the ergonomic changes for an average of 35 months (range 3–125 months). Two patients with embouchure dystonia used an ergonomic aid and reported no improvement. As a conclusion, ergonomic changes should be considered in all patients with limb dystonia. They are not helpful in embouchure dystonia (Jabusch *et al.* 2005).

## Pedagogical retraining and unspecific exercises

Pedagogical retraining was applied in patients with all forms of musician's dystonia. They comprised a variety of behavioural approaches used under the supervision of instructors and they included elements based on the following principles reported previously (Byl and McKenzie 2000; Boullet 2002; Candia *et al.* 2003; see also Chapter 19):

1. Movements of affected body parts were limited to a level of tempo and force at which the dystonic movement would not occur.
2. Compensatory movements (e.g. of adjacent fingers) were avoided, partially under the application of splints.
3. Instant visual feedback with mirrors or monitors helped patients to recognize dystonic and non-dystonic movements.
4. Body awareness techniques (e.g. Feldenkrais®) were used to increase the patient's perception of non-dystonic movements.

Pedagogical retraining was given in 24 patients (17% of all), 12 patients (50% of these) experienced improvement. Patients reporting an improvement had undergone this treatment for an average of 28 months (range 3–72months). Two patients with embouchure dystonia took part in pedagogical retraining and reported no improvement.

Besides the reported therapies recommended by the authors, patients additionally or alternatively performed unspecific technical exercises on their instruments such as are usually practiced by instrumentalists to improve their technique. Seventy-eight patients (54% of all) were practising such unspecific technical exercises, 44 patients (56% of these) experienced improvement. Six out of 11 patients with embouchure dystonia reported an improvement after practising such exercises (Jabusch *et al.* 2005).

These results provide evidence for the benefit of a behavioural approach, either by pedagogical retraining or by exercises on the instrument. Pedagogical retraining was applied in patients with all forms of musician's dystonia; however, only patients with limb dystonia improved. More than half of the patients who were practising unspecific exercises reported improvement, among them were six out of 11 patients with embouchure dystonia who applied such exercises. It is noteworthy that the results of these exercises were rated better than those of medical therapies. These positive results may be biased by a particular attitude among musicians and the perceived self-control through exercises. On the other hand positive results of retraining and of exercises underline the benefit of active involvement of patients in the treatment process. At present, the existing behavioural approaches are heterogeneous (Byl and McKenzie 2000; Boullet 2002; Candia *et al.* 2003). Instrument-specific retraining programmes ought to be elaborated which include the aforementioned principles. All approaches to retraining therapy require the time and patience of affected musicians which seem to be the limiting factors with these therapies. Behavioural aspects, however, should be included in all treatment approaches, at least as an adjunct.

## Findings in the light of the literature

In the overall outcome which was recorded on average 8.4 years after onset of symptoms, 71% of patients managed to stay in the musical profession. In other studies, focal dystonia had led to the end of their musical career in more than half of the patients (Schuele and Lederman 2003, 2004). Outcome measures were conducted 13.8 years (string instrumentalists) (Schuele and Lederman 2004) and 8.5 years (woodwind players) (Schuele and Lederman 2004) after onset of symptoms in these reports. The shorter duration of dystonia might be an explanation for the better outcome of our patients group compared with the report on the string players. In the study on woodwind players, the percentage of patients with embouchure dystonia was 25%; in our group it was 13.9%. The generally worse outcome in embouchure dystonia might explain the different outcome after similar durations of dystonia in both studies. In another report 17 out of 33 musicians with focal dystonia had stopped playing at the end of the follow-up period (Brandfonbrener and Robson 2002). The duration of dystonia and total follow-up time were, however, not mentioned, which did not allow a direct comparison. This was also the case in a report on 145 musicians with an average rehabilitation time of 24 months, after which 35 musicians returned to playing in public (Tubiana and Chamagne 2000).

## Instrumental students

The subgroup of instrumental students deserve some additional considerations. Focal dystonia is regarded as incurable at present. Although there are some

patients with encouraging follow-up results, therapy remains difficult and requires years and sometimes decades of patience, effort, and stress. Most patients who are at the zenith of their career have no choice and have to undergo treatment in order to improve their situation. Instrumental students, on the contrary, are young and flexible enough to change their profession. In our group, 12 out of 21 affected students changed their profession during the follow-up period. According to our impression, the results of therapy do not justify keeping students busy with treatment attempts over years in the most flexible and productive period of their lives. We conclude that strategies for students with musician's dystonia should include early support for a change to another profession.

### Limitations

Beside of the reported inclusion/exclusion criteria, the choice of treatment was strongly influenced by preferences of patients. The time investment required appeared to influence the choice of treatment. Older patients who desired a quick improvement preferred BT injections. In contrast, several patients found the injection therapy 'unnatural' and preferred a retraining therapy. A major limitation of the study was the lack of objective assessment tools for musician's dystonia which were not available at the beginning of the follow-up. Recently, reliable and precise objective methods have been developed for quantification of musician's dystonia (Jabusch and Altenmüller 2004; Jabusch *et al.* 2004*b*). In future studies such methods will have to be applied for reliable monitoring of treatment effects. Although we were limited in our assessment to a retrospective and subjective rating, several factors may support the validity of our result. Consistent negative results reported by patients with embouchure dystonia after treatment with BT, trihexyphenidyl, and retraining seem to indicate a low placebo sensitivity of musician's dystonia. Most previous studies observed that a significant number of musicians discontinued treatment within a year if they felt the response did not meet their needs. The adherence seen in our study, with an average treatment duration in responders between 2 and 3 years (and up to 10 years), seems unlikely if the result did not improve their instrumental performance significantly.

## Conclusion

Focal dystonia in musicians is still the main challenge in the field of musicians' medicine. With the available therapies the majority of patients manage to stay in their profession, many of them, however, with substantial compromises. Therapy of patients with embouchure dystonia remains problematic. The available medical approaches are ineffective, and other options have yet to be developed.

There is a pressing need for novel therapies for musician's dystonia. Several new methods have been described, such as immobilization therapy (Priori *et al.* 2001; see Chapter 19). These have to be investigated in larger numbers of patients in order to confirm beneficial effects. Behavioural therapies and interdisciplinary strategies combining pharmacological and pedagogical methods are promising, but the different approaches need to be evaluated. Since phenomenological and epidemiological data and results from electrophysiological and follow-up studies imply a behavioural component in the development as well as in the treatment of musician's dystonia, future research is required to identify 'beneficial behaviour' on the instrument. This might, possibly, also be of help for finding strategies with the particular aim of prevention of musician's dystonia.

## Acknowledgement

We would like to express our appreciation to Werner Linge, Director of Administration of the German Orchestra Union, for providing us with the March 2004 statistics on members of the German Orchestra Union.

## References

**Altenmüller, E.** (2003). Focal dystonia: advances in brain imaging and understanding of fine motor control in musicians. *Hand Clin*, **19**(3), 523–538.

**Bara-Jimenez, W., Catalan, M.J., Hallett, M., and Gerloff, C.** (1998). Abnormal somatosensory homunculus in dystonia of the hand. *Ann Neurol*, **44**(5), 828–831.

**Boullet, L.** (2002). Is there a cure for focal dystonia? Experiences with a new retraining therapy for pianists. *Proc German Soc Music Physiol Perform Arts Med*, **9**(4), 17.

**Brandfonbrener, A.G.** (1995). Musicians with focal dystonia: a report of 58 cases seen during a ten-year period at a performing arts medicine clinic. *Med Probl Perform Art*, **10**(4), 121–127.

**Brandfonbrener, A.G. and Robson, C.** (2002). A review of 111 musicians with focal dystonia seen at a performing artist's clinic 1985–2002. *Mov Disord*, **17**(5), 1135.

**Byl, N.N. and McKenzie, A.** (2000). Treatment effectiveness for patients with a history of repetitive hand use and focal hand dystonia: a planned, prospective follow-up study. *J Hand Ther*, **13**(4), 289–301.

**Byl, N.N., Merzenich, M.M., and Jenkins, W.M.** (1996). A primate genesis model of focal dystonia and repetitive strain injury: I. Learning-induced dedifferentiation of the representation of the hand in the primary somatosensory cortex in adult monkeys. *Neurology*, **47**(2), 508–520.

**Candia, V., Elbert, T., Altenmüller, E., Rau, H., Schäfer, T., and Taub, E.** (1999). A constraint-induced movement therapy for focal hand dystonia in musicians. *Lancet*, **353**(9146), 42.

**Candia, V., Wienbruch, C., Elbert, T., Rockstroh, B., and Ray, W.** (2003). Effective behavioral treatment of focal hand dystonia in musicians alters somatosensory cortical organization. *Proc Natl Acad Sci USA*, **100**(13), 7942–7946.

**Charness, M.E., Ross, M.H., and Shefner, J.M.** (1996). Ulnar neuropathy and dystonic flexion of the fourth and fifth digits: clinical correlation in musicians. *Muscle Nerve*, **19**(4), 431–437.

**Cole, R.A., Cohen, L.G., and Hallett, M.** (1991). Treatment of musician's cramp with botulinum toxin. *Med Probl Perform Art*, **6**(4), 137–143.

**Deuschl, G., Toro, C., Matsumoto, J., and Hallett, M.** (1995). Movement-related cortical potentials in writer's cramp. *Ann Neurol*, **38**(6), 862–868.

**Elbert, T., Candia, V., Altenmüller, E.** *et al.* (1998) Alteration of digital representations in somatosensory cortex in focal hand dystonia. *NeuroReport*, **9**(16), 3571–3575.

**Flor, H., Braun, C., Elbert, T., and Birbaumer, N.** (1997). Extensive reorganization of primary somatosensory cortex in chronic back pain patients. *Neurosci Lett*, **224**(1), 5–8.

**Frucht, S.J., Fahn, S., and Greene, P.E.** (2001). The natural history of embouchure dystonia. *Mov Disord*, **16**(5), 899–906.

**Hallett, M.** (1998). The neurophysiology of dystonia. [review]. *Arch Neurol*, **55**(5), 601–603.

**Hua, S.E., Garonzik, I.M., Lee, J.I., and Lenz, F.A.** (2000). Microelectrode studies of normal organization and plasticity of human somatosensory thalamus. *J Clin Neurophysiol*, **17**(6), 559–574.

**Hummel, F., Andres, F., Altenmüller, E., Dichgans, J., and Gerloff, C.** (2002). Inhibitory control of acquired motor programmes in the human brain. *Brain*, **125**, 404–420.

**Jabusch, H.C. and Altenmüller, E.** (2004). Three-dimensional movement analysis as a promising tool for treatment evaluation of musicians' dystonia. *Adv Neurol*, **94**, 239–245.

**Jabusch, H.C., Müller, S.V., and Altenmüller, E.** (2004*a*). Anxiety in musicians with focal dystonia and those with chronic pain. *Mov Disord*, **19**(10), 1169–1175.

**Jabusch, H.C., Vauth, H., and Altenmüller, E.** (2004*b*). Quantification of focal dystonia in pianists using scale analysis. *Mov Disord*, **19**(2), 171–180.

**Jabusch, H.C., Zschucke, D., Schmidt, A., Schuele, S., and Altenmüller, E.** (2005). Focal dystonia in musicians: treatment strategies and long-term outcome in 144 patients. *Mov Disord.* **20**(12), 1623–1626.

**Jankovic, J. and Shale, H.** (1989). Dystonia in musicians. *Semin Neurol*, **9**(2), 131–135.

**Karp, B.I., Cole, R.A., Cohen, L.G., Grill, S., Lou, J.S., and Hallett, M.** (1994). Long-term botulinum toxin treatment of focal hand dystonia. *Neurology*, **44**, 70–76.

**Lederman, R.J.** (1991). Focal dystonia in instrumentalists: clinical features. *Med Probl Perform Art*, **6**(4), 132–136.

**Lenz, F.A. and Byl, N.N.** (1999). Reorganization in the cutaneous core of the human thalamic principal somatic sensory nucleus (ventral caudal) in patients with dystonia. *J Neurophysiol*, **82**(6), 3204–3212.

**Lim, V.K. and Altenmüller, E.** (2003). Musicians' cramp: instrumental and gender differences. *Med Probl Perform Art*, **18**(1), 21–26.

**Naumann, M., Pirker, W., Reiners, K., Lange, K.W., Becker, G., and Brucke, T.** (1998). Imaging the pre- and postsynaptic side of striatal dopaminergic synapses in idiopathic cervical dystonia: a SPECT study using [123I] epidepride and [123I] beta-CIT. *Mov Disord*, **13**(2), 319–323.

**Preibisch, C., Berg, D., Hofmann, E., Solymosi, L., and Naumann, M.** (2001). Cerebral activation patterns in patients with writer's cramp: a functional magnetic resonance imaging study. *J Neurol.* **248**(1), 10–17.

**Priori, A., Pesenti, A., Cappellari, A., Scarlato, G., and Barbieri, S.** (2001). Limb immobilization for the treatment of focal occupational dystonia. *Neurology*, **57**(3), 405–409.

**Pujol, J., Roset-Llobet, J., Rosines-Cubells, D.** *et al.* (2000). Brain cortical activation during guitar-induced hand dystonia studied by functional MRI. *NeuroImage*, **12**(3), 257–267.

**Rosenkranz, K., Williamon, A., Butler, K., Cordivari, C., Lees, A.J., and Rothwell, J.C.** (2005). Pathophysiological differences between musician's dystonia and writer's cramp. *Brain*, **128**, 918–931.

**Schuele, S. and Lederman, R.J.** (2003). Focal dystonia in woodwind instrumentalists: long-term outcome. *Med Probl Perform Art*, **18**(1), 15–20.

**Schuele, S. and Lederman, R.J.** (2004). Long-term outcome of focal dystonia in string instrumentalists. *Mov Disord*, **19**(1), 43–48.

**Schuele, S., Jabusch, H.C., Lederman, R.J., and Altenmüller, E.** (2005). Botulinum toxin injections in the treatment of musician's dystonia. *Neurology*, **64**(2), 341–343.

**Tinazzi, M., Fiaschi, A., Rosso, T., Faccioli, F., Grosslercher, J., and Aglioti, S.M.** (2000). Neuroplastic changes related to pain occur at multiple levels of the human somatosensory system: a somatosensory-evoked potentials study in patients with cervical radicular pain. *J Neurosci*, **20**(24), 9277–9283.

**Toro, C., Deuschl, G., and Hallett, M.** (2000). Movement-related electroencephalographic desynchronization in patients with hand cramps: evidence for motor cortical involvement in focal dystonia. *Ann Neurol*, **47**(4), 456–461.

**Tubiana, R. and Chamagne, P.** (2000). Prolonged rehabilitation treatment of musician's focal dystonia. In: R. Tubiana and P.C. Amadio (eds) *Medical problems of the instrumentalist musician*, pp. 369–378. Martin Dunitz, London.

Chapter 18

# The neurophysiology of focal hand dystonia in musicians

Karin Rosenkranz

## Introduction

The classification of musician's dystonia as a form of focal hand dystonia is based on the fact that involuntary abnormal contractions appear that are fairly similar to those observed, for example, in writer's cramp. But compared with the latter there are differences in clinical features that are especially characteristic of musician's dystonia:

1. Musician's dystonia is often highly task specific and highly focal, for example a pianist experiences 'curling in' of the fourth finger while playing scales downwards but not upwards.
2. The age of onset of musician's dystonia is earlier than in writer's cramp—33 years for musician's dystonia (see Chapter 17) versus 38.5 years in writer's cramp (O'Riordan *et al.* 2004).
3. The prevalence of musician's dystonia amongst professional musicians is higher than that of writer's cramp in the general population (Lim *et al.* 2001).
4. The consequences of being affected by focal hand dystonia are, for musicians, often more devastating.

Another distinguishing factor is the clear association of musician's dystonia with excessive practise of complex hand (or mouth) movements. Whereas most patients with other forms of focal hand dystonia, such as writer's cramp, have no history of excessive hand use above that seen in the general population, in musician's dystonia it is often the movements that are practised most that are those where the dystonic symptoms manifest.

As outlined in Chapter 9, there are structural and functional changes in the brains of healthy musicians that have been interpreted as a beneficial adaptation to musical training supporting the development of a unique capacity for skilful performance (Amunts *et al.* 1997; Bangert *et al.* 2001; Münte *et al.* 2002). The interesting question is whether or not practise of a movement and the plastic changes in the brain play an important role in the pathophysiology of musician's

dystonia, therefore distinguishing it from other forms of focal hand dystonia. Over the last decades many studies of focal hand dystonia have been undertaken which have increased our understanding of this condition. However, many studies focused on focal/segmental upper limb dystonia and there is very little information specifically on patients with musician's dystonia.

The intention of this chapter is to give an overview of the neurophysiology of focal hand dystonia in musicians. So as not to miss out on important findings that certainly apply to all forms, several studies done in a more 'general' population of patients with focal hand dystonia are presented. At the end of this chapter we present the results of our own study describing neurophysiological differences between writer's cramp and musician's dystonia and discuss the implications of these findings for the pathophysiology of these two forms of focal hand dystonia. This might be the first step towards an answer to the above question.

## Pathophysiology of focal hand dystonia

Most studies of focal hand dystonia reveal abnormalities in three main areas: (1) reduced excitability of inhibitory connections at all levels (spinal, brainstem, and cortical) of the motor system; (2) reduced sensory perception and integration; and (3) impaired sensorimotor organization/integration. However, it should be noted that although each of these alone could potentially be the cause of dystonic contractions—i.e. reduced inhibition could lead to co-contraction of antagonist muscles and overflow of activity to muscle not involved in a task; impaired sensory and sensorimotor integration could lead to poor selection of movement parameters and excess muscle contraction—none are in fact a necessary cause of dystonia. Thus, similar abnormalities are often observed in the unaffected arm if the clinical symptoms are unilateral (Deuschl *et al.* 1991; Ridding *et al.* 1995; Meunier *et al.* 2001). Our hypothesis is that some combination of these factors seems to be necessary for clinical symptoms and the amount of reorganization that is necessary depends on the individual level of abnormalities in inhibitory circuits and in sensorimotor integration.

## Excitability of spinal, brainstem, and cortical inhibitory circuits

This has been summarized in the review by Berardelli *et al.* (1998). In brief, spinal inhibition can be probed by testing the excitability of the reciprocal inhibitory circuits between extensor and flexor muscles in the forearm (Nakashima *et al.* 1989; Panizza *et al.* 1989; Chen *et al.* 1995). Brainstem inhibition is tested by examining the recovery curve of the electrically elicited blink reflex (Berardelli *et al.* 1985), and motor cortical inhibition is tested by measuring the duration of

the silent period following a transcranial magnetic stimulation (TMS) pulse (Rona *et al.* 1998). In almost all cases there is less inhibition in patients with dystonia than in healthy controls.

However, two recent studies testing intracortical inhibition using a paired-pulse TMS protocol have shown that in patients with action-induced and task-specific dystonia, who were able to keep their arm completely relaxed during the experiment, the level of intracortical inhibition is normal (Stinear and Byblow 2004; Rosenkranz *et al.* 2005).

More recent studies have examined the time course of excitability in cortical inhibitory systems prior to focal hand movements. The intracortical inhibitory circuits targeted by short-interval paired-pulse TMS decrease the excitability of their projections to the prime moving muscle whereas excitability increases in those that project to nearby muscles that must stay relaxed during the task (Reynolds and Ashby 1999). This is compatible with an organization in which inhibition is normally withdrawn from the focal target muscle and at the same time is increased in muscles that are not involved in the task. Patients with focal dystonia have a deficit in this focusing mechanism (Stinear and Byblow 2004).

## Sensory perception, integration, and mapping

Temporal and spatial discrimination, being the ability to perceive two stimuli as temporally or spatially separate, are impaired in patients with generalized dystonia (Tinazzi *et al.* 1999) and focal hand dystonia (Bara Jimenez *et al.* 2000*a*,*b*; Sanger *et al.* 2001). A recent study compared spatial sensory discrimination abilities in patients with primary generalized DYT1 dystonia and other forms of dystonia, especially focal hand dystonia. Interestingly, only the group with focal hand dystonia showed an increased spatial discrimination threshold, pronounced in the dominant hand, whereas no abnormalities of spatial discrimination were found in DYT1 cases (Molloy *et al.* 2003). This finding suggests that altered sensory processing might play a different pathophysiological role in these conditions.

Electrical stimulation of either ulnar or median nerves at the wrist gives rise to SEPs in the sensory cortex. If both nerves are stimulated together, then the SEP is smaller than the expected combined sum of each alone. Tinazzi *et al.* (2000) suggested that this might be due to lateral inhibition between inputs from each source, which might under normal circumstances focus their central effects within the brain. This effect was reduced in patients with focal hand dystonia (Tinazzi *et al.* 2000). As this paradigm gives an indication of a sensory 'between-field' or surround-inhibition, the recovery of SEPs in response to paired median nerve stimuli is an indicator for the 'in-field' inhibition. Frasson *et al.* (2001) showed that this 'in-field' suppression was reduced in dystonia patients.

Mapping studies of the sensory cortex using SEP/magnetoencephalographic (MEG) techniques found that the locus of sensory inputs from individual fingers overlap more in patients with focal hand dystonia than in healthy controls (Bara Jimenez *et al.*1998; McKenzie *et al.* 2003). Similar observations have been made in patients with musician's dystonia in a selective study using MEG mapping. Elbert *et al.* (1998) showed that there is an overlap of the representational zones of the digits in the primary somatosensory cortex for the affected hand of dystonic musicians compared with the representations of the digits in non-musician control subjects. In healthy musicians, an increase in sensory finger representations has been described and interpreted as changes to conform to the current needs and experiences of the individual (Elbert *et al.* 1995). The question arising here is whether these changes develop too far in musician's dystonia, from being beneficial to a being a maladaptation.

However, these mapping studies done at rest might have to be interpreted with care. Braun *et al.* (2003) investigated dynamic alterations in the functional organization of the primary somatosensory cortex (SI) during writing in patients with writer's cramp and healthy controls with neuromagnetic source imaging. Similar to the studies cited above, they found a reduced distance between cortical representations of the thumb and little finger of the affected hand in patients when tested at rest. However, similar to healthy controls, patients showed normal modulation of the functional organization of SI as induced by writing: the cortical distances between representations of the thumb and little finger increased. Thus, despite alterations in the organization of the primary somatosensory cortex in writer's cramp, the capability of SI to adapt dynamically to different tasks is, astonishingly, not impaired (Braun *et al.* 2003).

## Interaction of sensory input and motor output

Under normal conditions, motor outflow from the cortex 'gates' sensory input from the moving part of the body. This 'gating' starts before the onset of motion (Shimazu *et al.* 1999). Murase *et al.* (2000) tested SEPs to median nerve stimulation during preparation for and execution of movement in patients with writer's cramp. Differences were found between healthy subjects and sufferers from writer's cramp, especially in the frontal N30, a component that is thought to be generated by sequential processing of sensory input after arrival at the primary somatosensory cortex. Whereas during the pre-movement phase healthy subjects showed a significant frontal decrease of the amplitude of the N30, writer's cramp patients did not. However, during active and passive movements the N30 component was equally attenuated in writer's cramp and healthy controls. During the preparatory period of a movement the sensitivity of cells that respond to afferent input might be altered to allow an undisturbed starting of the

actual movement. This process may be compromised in writer's cramp and contribute to a failure of the motor program before it has even been properly started.

Other authors investigated the integration of sensory inputs and motor outputs with a variety of protocols that use TMS to probe the excitability of the motor cortex. Tamburin *et al.* (2001) showed in healthy subjects that sensory input from one finger has a particular time course of effects on the amplitude of MEPs evoked in hand muscles by the TMS pulse. In addition, the effects are usually spatially segregated, at least to some extent, with the effects being greatest on muscles near the stimulated finger and least on those further away. Both the time course and the spatial specificity are abnormal in dystonia (Tamburin *et al.* 2002). In a similarly designed study, this time using electrical median nerve stimulation, Abbruzzese *et al.* (2001) examined healthy subjects as well as patients with hand cramps, two of them musicians, and patients with torticollis. In healthy subjects the amplitude of the motor-evoked-potential (MEP) in the hand muscles is reduced by preceding peripheral stimulation. This was the case for torticollis patients too. But the patients with hand cramps showed a facilitation of MEPs instead of an inhibition. Thus, the central integration of sensory input evoked from the site of manifestation of symptoms seems to be specifically defective.

## Pathophysiological observations in musician's dystonia

In a recent study we compared the sensorimotor organization of the hand in healthy non-musicians, healthy professional musicians in training, and in two groups of patients with musician's dystonia or with writer's cramp (Rosenkranz *et al.* 2005). The protocol employed focal vibration of a single hand muscle to evoke proprioceptive input from the activation of muscle spindles and TMS of the motor cortex to evoke MEPs in three different hand muscles (first dorsal interosseus (FDI), abductor pollicis brevis (APB), and adductor digiti minimi (ADM)) (Rosenkranz and Rothwell 2003). We tested how MEPs and paired-pulse (short interval) intracortical inhibition (SICI) in each muscle was modulated when vibration was applied to each muscle in turn. In healthy non-musicians vibration of one muscle increased the amplitude of MEPs and decreased SICI in the vibrated muscle, whilst having the opposite effect on the non-vibrated hand muscles (Fig. 18.1). This pattern of sensorimotor interaction probably reflects an interaction of the sensory input with intracortical inhibitory circuits in or near to the motor cortical representations of the hand muscles.

Previously we showed that in healthy non-musicians the spatial pattern of interaction between sensory input and motor output could be changed by a prolonged period of sensory stimulation (Rosenkranz and Rothwell 2004). Thus,

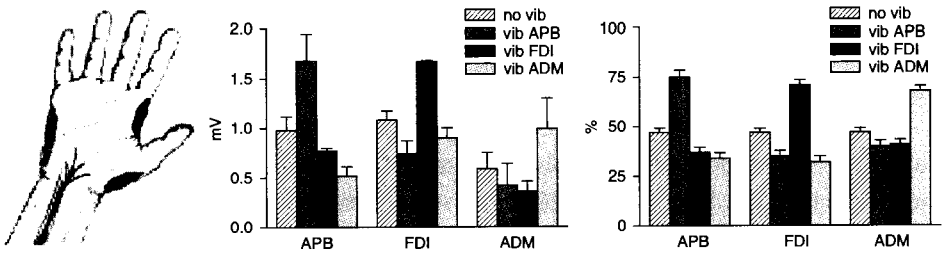



**Fig. 18.1** Schematic illustration of the protocol used to test sensorimotor interactions in small hand muscles. EMG recordings are made from three different muscles (APB, ADM, FDI) and low-amplitude vibration is applied to each of them in turn for a period of 1.5 s. The vibration amplitude is such that no tonic vibration reflex is evoked, nor any perceptions of illusory movement. Transcranial magnetic stimulation is applied to the contralateral cortex to evoke MEP in the hand 1 s after onset of vibration. The bar plots on the left show the motor-evoked potentials (MEPs) in each of these hand muscles without vibration, during vibration of APB, during vibration of FDI, and during vibration of ADM. In the vibrated muscle, MEPs are increased and decreased in the non-vibrated muscles. In other trials, short interval intracortical inhibition (SICI) is estimated by using a paired-pulse TMS protocol (Kujirai *et al.* 1993). This decreases the size of the control MEP by about 50%. The bar plots on the right show the results when SICI is evaluated during vibration: it is reduced to about 75% in the vibrated muscle and increased to around 25% in the other muscles (Rosenkranz *et al.* 2005).

15-min continuous vibration of both the APB and FDI muscles led to a situation in which subsequent input from FDI had the same, rather than the opposite, effect on APB. Similarly, after the intervention, input from APB began to facilitate FDI. As expected from previous work in animal experiments, the effects did not occur if the subject did not focus their attention on the input during the prolonged period of vibration. Our conclusion was that a period of training with changed sensory inputs can alter the usual pattern of sensorimotor interaction in the cortex.

In the experiments with musicians we found that in healthy musicians the pattern of sensory–motor interaction was similar to what we had observed in non-musicians after a 15-min sensory intervention (Rosenkranz *et al.* 2005). This would be consistent with the idea that prolonged practice in the musicians had led to a permanent reorganization in the way in which sensory input from individual fingers interacts with motor output. We speculated that this might be advantageous in that it would favour cooperation between finger movements in rapid sequences of movements.

When we examined the patients with musician's dystonia, we observed a pattern that was even further removed from normal (Fig. 18.2): input from any of the three hand muscles facilitated all the other three hand muscles, as if the pattern of spatial organization in healthy musicians had been extended until it led to a situation in which there was no distinction between inputs from any of

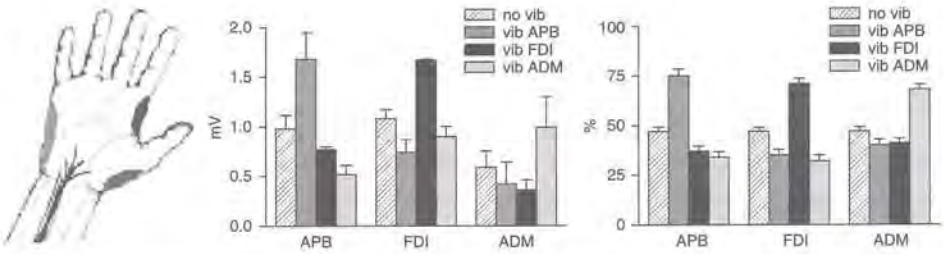



**Fig. 18.1** Schematic illustration of the protocol used to test sensorimotor interactions in small hand muscles. EMG recordings are made from three different muscles (APB, ADM, FDI) and low-amplitude vibration is applied to each of them in turn for a period of 1.5 s. The vibration amplitude is such that no tonic vibration reflex is evoked, nor any perceptions of illusory movement. Transcranial magnetic stimulation is applied to the contralateral cortex to evoke MEP in the hand 1 s after onset of vibration. The bar plots on the left show the motor-evoked potentials (MEPs) in each of these hand muscles without vibration, during vibration of APB, during vibration of FDI, and during vibration of ADM. In the vibrated muscle, MEPs are increased and decreased in the non-vibrated muscles. In other trials, short interval intracortical inhibition (SICI) is estimated by using a paired-pulse TMS protocol (Kujirai *et al.* 1993). This decreases the size of the control MEP by about 50%. The bar plots on the right show the results when SICI is evaluated during vibration: it is reduced to about 75% in the vibrated muscle and increased to around 25% in the other muscles (Rosenkranz *et al.* 2005).

15-min continuous vibration of both the APB and FDI muscles led to a situation in which subsequent input from FDI had the same, rather than the opposite, effect on APB. Similarly, after the intervention, input from APB began to facilitate FDI. As expected from previous work in animal experiments, the effects did not occur if the subject did not focus their attention on the input during the prolonged period of vibration. Our conclusion was that a period of training with changed sensory inputs can alter the usual pattern of sensorimotor interaction in the cortex.

In the experiments with musicians we found that in healthy musicians the pattern of sensory–motor interaction was similar to what we had observed in non-musicians after a 15-min sensory intervention (Rosenkranz *et al.* 2005). This would be consistent with the idea that prolonged practice in the musicians had led to a permanent reorganization in the way in which sensory input from individual fingers interacts with motor output. We speculated that this might be advantageous in that it would favour cooperation between finger movements in rapid sequences of movements.

When we examined the patients with musician's dystonia, we observed a pattern that was even further removed from normal (Fig. 18.2): input from any of the three hand muscles facilitated all the other three hand muscles, as if the pattern of spatial organization in healthy musicians had been extended until it led to a situation in which there was no distinction between inputs from any of

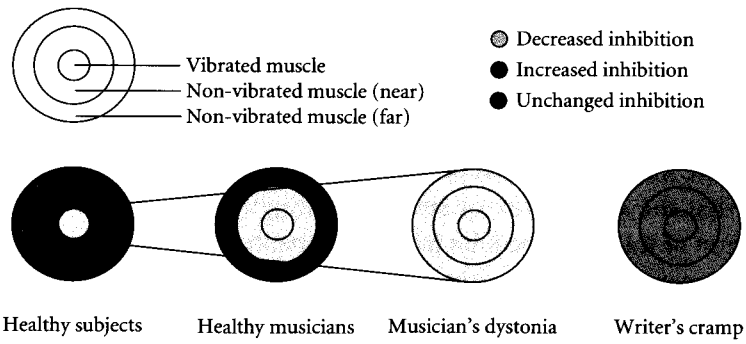



**Fig. 18.2** Schematic summary of the effects of focal vibratory input on short interval intracortical inhibition (SICI) in three hand muscles. In this diagram, the hand muscle representations are drawn as circles with the vibrated muscle in the centre, and the 'near' and 'far' non-vibrated muscles surrounding it. This illustration represents the effective functional connectivity between muscles, rather than anatomical distances between muscle representation in the cortex. Shading represents the level of SICI: a reduction of SICI, an increase, and an unchanged SICI compared with the non-vibration condition (see key). In healthy subjects, focal vibratory input to one muscle reduces SICI in that muscle and increases it in other muscles. In healthy musicians, SICI in both 'near' muscles is reduced by vibration applied to either of them, whilst it is increased in the 'far' one. In musician's dystonia, vibration of any muscle reduces SICI in all of them. Patients with writer's cramp show no modulation of SICI by vibration of any muscle; the SICI remains unchanged compared with baseline levels.

the muscles. Interestingly, this was only the case in musician's *hand* dystonia; the two musicians with embouchure dystonia tested here showed a pattern similar to healthy musicians.

In writer's cramp, the pattern of organization differed. In these subjects, input from the muscles failed to have any influence at all on motor excitability. It was as if in these individuals the usual effect of sensory input was completely absent.

The data are consistent with a model in which musical practice in healthy musicians leads to changes in organization of the motor cortex that are probably beneficial to performance. However, in musician's dystonia these progress too far and perhaps they then begin to interfere with movement rather than assist it. This raises the possibility that musician's dystonia is a form of training-induced dystonia comparable to that described in the animal model of Byl *et al.* (1996): in susceptible individuals an initially beneficial adaptation of sensorimotor organization progresses too far and leads to problems in targeting motor commands.

With regard to the epidemiology of musician's dystonia (see Chapter 17), one might assume that those musicians who are known to practise a lot and from an early age might drive their brain's facility for plastic changes too far and exhaust the capacity of regulatory mechanisms, like for example cortical inhibition, that would normally balance the system. Since, fortunately, not all highly performing musicians suffer from dystonia, it could be hypothesized that those who develop

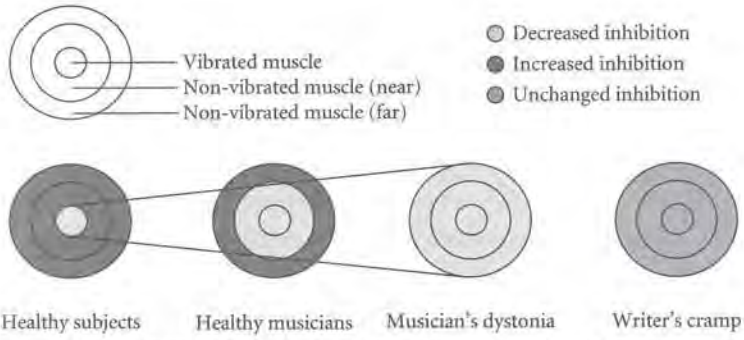



**Fig. 18.2** Schematic summary of the effects of focal vibratory input on short interval intracortical inhibition (SICI) in three hand muscles. In this diagram, the hand muscle representations are drawn as circles with the vibrated muscle in the centre, and the 'near' and 'far' non-vibrated muscles surrounding it. This illustration represents the effective functional connectivity between muscles, rather than anatomical distances between muscle representation in the cortex. Shading represents the level of SICI: a reduction of SICI, an increase, and an unchanged SICI compared with the non-vibration condition (see key). In healthy subjects, focal vibratory input to one muscle reduces SICI in that muscle and increases it in other muscles. In healthy musicians, SICI in both 'near' muscles is reduced by vibration applied to either of them, whilst it is increased in the 'far' one. In musician's dystonia, vibration of any muscle reduces SICI in all of them. Patients with writer's cramp show no modulation of SICI by vibration of any muscle; the SICI remains unchanged compared with baseline levels.

the muscles. Interestingly, this was only the case in musician's *hand* dystonia; the two musicians with embouchure dystonia tested here showed a pattern similar to healthy musicians.

In writer's cramp, the pattern of organization differed. In these subjects, input from the muscles failed to have any influence at all on motor excitability. It was as if in these individuals the usual effect of sensory input was completely absent.

The data are consistent with a model in which musical practice in healthy musicians leads to changes in organization of the motor cortex that are probably beneficial to performance. However, in musician's dystonia these progress too far and perhaps they then begin to interfere with movement rather than assist it. This raises the possibility that musician's dystonia is a form of training-induced dystonia comparable to that described in the animal model of Byl *et al.* (1996): in susceptible individuals an initially beneficial adaptation of sensorimotor organization progresses too far and leads to problems in targeting motor commands.

With regard to the epidemiology of musician's dystonia (see Chapter 17), one might assume that those musicians who are known to practise a lot and from an early age might drive their brain's facility for plastic changes too far and exhaust the capacity of regulatory mechanisms, like for example cortical inhibition, that would normally balance the system. Since, fortunately, not all highly performing musicians suffer from dystonia, it could be hypothesized that those who develop

it might have a certain susceptibility due to the presence of subclinical deficits in motor inhibition and/or sensory processing that might not even have become apparent had they chosen a different profession.

From a practical point of view, this emphasis on the importance of pre-symptomatic hand training in provoking expression of (sub-)clinical deficits suggests that the appearance of symptoms might be delayed or prevented by modification of training schedules to reduce repetition and increase the variability of the movements that are practised.

Finally, the data show that the situation in the patients with writer's cramp may be rather different since they show a complete lack of any sensory effects on the excitability of the hand area of the motor cortex. In them, we hypothesize that their central nervous system reorganizes in a different way because they don't share the same history of hand movement with musicians and their affected movements do not need to rely on such close interaction between sensory input and motor output to the hand as in musicians. Thus, repetitive sensory input driven by movement practice plays only a minor (if any) role. A primary motor incoordination probably occurs due to an intrinsic deficit of the inhibition mechanisms that subsequently induces a dis-organization of sensorimotor interaction.

## Acknowledgements

The author thanks Professor John Rothwell for helpful discussion. This work was supported by the Deutsche Forschungsgemeinschaft (DFG), Action Medical Research, the Medical Research Council UK, and the Dystonia Research Foundation.

## References

**Abbruzzese, G., Marchese, R., Buccolieri, A., Gasparetto, B., and Trompetto, C.** (2001). Abnormalities of sensorimotor integration in focal dystonia. A transcranial magnetic stimulation study. *Brain*, **124**, 537–545.

**Amunts, K., Schlaug, G., Jancke, L.** *et al.* (1997). Motor cortex and hand motor skills: structural compliance in the human brain. *Hum Brain Mapp*, **5**, 206–215.

**Bangert, M., Haeusler, U., and Altenmüller, E.** (2001). On practice: how the brain connects piano keys and piano sounds. *Ann NY Acad Sci*, **930**, 425–428.

**Bara-Jimenez, W., Catalan, M.J., Hallett, M., and Gerloff, C.** (1998). Abnormal somatosensory homunculus in dystonia of the hand. *Ann Neurol*, **44**, 828–831.

**Bara-Jimenez, W., Shelton, P., Sanger, T.D., and Hallett, M.** (2000*a*). Sensory discrimination capabilities in patients with focal hand dystonia. *Ann Neurol*, **47**, 377–380.

**Bara-Jimenez, W., Shelton, P., and Hallett, M.** (2000*b*). Spatial discrimination is abnormal in focal hand dystonia. *Neurology*, **55**, 1869–1873.

**Berardelli, A., Rothwell, J.C., Day, B.L., and Marsden, C.D.** (1985). Pathophysiology of blepharospasm and oromandibular dystonia. *Brain*, **108**, 593–608.

**Berardelli, A., Rothwell, J.C., Hallett, M., Thompson, P.D., Manfredi, M., and Marsden, C.D.** (1998). The pathophysiology of primary dystonia. *Brain*, **121**, 1195–1212.

**Braun, C., Schweizer, R., Heinz, U., Wiech, K., Birbaumer, N., and Topka, H.** (2003). Task-specific plasticity of somatosensory cortex in patients with writer's cramp. *NeuroImage*, **20**, 1329–1338.

**Byl, N.N., Merzenich, M.M., and Jenkins, W.M.** (1996). A primate genesis model of focal dystonia and repetitive strain injury: I. Learning-induced dedifferentiation of the representation of the hand in the primary somatosensory cortex in adult monkeys. *Neurology*, **47**, 508–520.

**Chen, R.S., Tsai, C.H., and Lu, C.S.** (1995). Reciprocal inhibition in writer's cramp. *Mov Disord*, **10**(5), 556–561.

**Deuschl, G., Michels, R., Berardelli, A., Schenck, E., Inghilleri, M., and Lucking, C.H.** (1991). Effects of electric and magnetic transcranial stimulation on long latency reflexes. *Exp Brain Res*, **83**, 403–410.

**Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., and Taub, E.** (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, **270**, 305–307.

**Elbert, T., Candia, V., Altenmüller, E.** *et al.* (1998). Alteration of digital representations in somatosensory cortex in focal hand dystonia. *NeuroReport*, **16**, 3571–3575.

**Frasson, E., Priori, A., Bertolasi, L., Mauguiere, F., Fiaschi, A., and Tinazzi, M.** (2001). Somatosensory disinhibition in dystonia. *Mov Dis*, **16**, 674–682.

**Kujirai, T., Caramia, M.D., Rothwell, J.C.** *et al.* (1993). Corticocortical inhibition in human motor cortex. *J Physiol*, **471**, 501–519.

**Lim, V., Altenmüller, E., and Bradshaw, J.L.** (2001). Focal dystonia: current theories. *Hum Mov Sci*, **20**, 875–914.

**McKenzie, A.L., Nagarajan, S.S., Roberts, T.P., Merzenich, M.M., and Byl, N.N.** (2003). Somatosensory representation of the digits and clinical performance in patients with focal hand dystonia. *Am J Phys Med Rehabil*, **82**, 737–749.

**Meunier, S., Garnero, L., Ducorps, A.** *et al.* (2001). Human brain mapping in dystonia reveals both endophenotypic traits and adaptive reorganization. *Ann Neurol*, **50**, 521–527.

**Molloy, F.M., Carr, T.D., Zeuner, K.E., Dambrosia, J.M., and Hallett, M.** (2003). Abnormalities of spatial discrimination in focal and generalized dystonia. *Brain* **126**, 2175–2182.

**Münte, T.F., Altenmüller, E., and Jaencke, L.** (2002). The musician's brain as a model of neuroplasticity. *Nat Rev*, **3**, 473–478.

**Murase, N., Kaji, R., Shimazu, H.** *et al.* (2000). Abnormal premovement gating of somatosensory input in writer's cramp. *Brain*, **123**, 1813–1829.

**Nakashima, K., Rothwell, J.C., Day, B.L., Thompson, P.D., Shannon, K., and Marsden, C.D.** (1989). Reciprocal inhibition between forearm muscles in patients with writer's cramp and other occupational cramps, symptomatic hemidystonia and hemiparesis due to stroke. *Brain*, **112**, 681–697.

**O'Riordan, S., Raymond, D., Lynch, T.** *et al.* (2004). Age at onset as a factor in determining the phenotype of primary torsion dystonia. *Neurology*, **63**, 1423–1426.

**Panizza, M.E., Hallett, M., and Nilsson, J.** (1989). Reciprocal inhibition in patients with hand cramps. *Neurology*, **39**, 85–89.

**Reynolds, C. and Ashby, P.** (1999). Inhibition in the human motor cortex is reduced just before a voluntary contraction. *Neurology*, **53**, 730–735.

**Ridding, M.C., Sheean, G., Rothwell, J.C., Inzelberg, R., and Kujirai, T.** (1995). Changes in the balance between motor cortical excitation and inhibition in focal, task-specific dystonia. *J Neurol Neurosurg Psychiat*, **59**, 493–498.

**Rona, S., Berardelli, A., Vacca, L., Inghilleri, M., and Manfredi, M.** (1998). Alteration of motor cortical inhibition in patients with dystonia. *Mov Dis*, **13**, 118–124.

**Rosenkranz, K. and Rothwell, J.C.** (2003). Differential effect of muscle vibration on intracortical inhibitory circuits in humans. *J Physiol*, **551**, 649–661.

**Rosenkranz, K. and Rothwell, J.C.** (2004). The effect of sensory input and attention on the sensorimotor organization of the hand area of the human motor cortex. *J Physiol*, **561**, 307–320.

**Rosenkranz, K., Williamon, A., Butler, K., Cordivari, C., Lees, A.J., and Rothwell, J.C.** (2005). Pathophysiological differences between musician's dystonia and writer's cramp. *Brain*, **128**(4), 918–931.

**Sanger, T.D., Tarsy, D., and Pascual-Leone, A.** (2001). Abnormalities of spatial and temporal sensory discrimination in writer's cramp. *Mov Dis*, **16**, 94–99.

**Shimazu, H., Kaji, R., Murase, N.** *et al.* (1999). Pre-movement gating of short-latency somatosensory evoked potentials. *NeuroReport*, **10**, 2457–2460.

**Stinear, C.M. and Byblow, W.D.** (2004). Impaired modulation of intracortical inhibition in focal hand dystonia. *Cereb Cortex*, **14**, 555–561.

**Tamburin, S., Manganotti, P., Zanette, G., and Fiaschi, A.** (2001). Cutaneomotor integration in human hand motor areas: somatotopic effect and interaction of afferents. *Exp Brain Res*, **141**, 232–241.

**Tamburin, S., Manganotti, P., Marzi, C.A., Fiaschi, A., and Zanette, G.** (2002). Abnormal somatotopic arrangement of sensorimotor interactions in dystonic patients. *Brain*, **125**, 719–730.

**Tinazzi, M., Frasson, E., Bertolasi, L., Fiaschi, A., and Aglioti, S.** (1999). Temporal discrimination of somesthetic stimuli is impaired in dystonic patients. *NeuroReport*, **10**(7), 1547–1550.

**Tinazzi, M., Priori, A., Bertolasi, L., Frasson, E., Mauguiere, F., and Fiaschi, A.** (2000). Abnormal central integration of dual somatosensory input in dystonia. Evidence for sensory overflow. *Brain*, **123**, 42–50.

Chapter 19

# The development of focal dystonia in musicians as a consequence of maladaptive plasticity: implications for intervention

Nancy N. Byl and Alberto Priori

## Introduction

Focal dystonia in musicians, or musician's cramp, is a disabling disorder presenting as a painless muscular incoordination or loss of voluntary motor control. This problem most commonly affects extensively trained movements while a musician is playing the instrument (Sheehy and Marsden 1982; Rosenbaum and Jankovic 1988; Lockwood 1989). Involuntary twisting, end-range postures originating from co-contractions of agonists and antagonists interfere with the necessary finely graded, sequenced performance of musical tasks. The aetiology of focal hand dystonia is not completely understood, but it is probably multifactorial (see Fig. 19.1 and Chapter 17). It may develop in individuals with a genetic history of dystonia (Ozelius *et al.* 1997) or physiological dysfunction of the basal ganglia (Black *et al.* 1998) or the sensory motor cortex (Elbert *et al.* 1998). Other aetiologies include a traumatic injury (Jankovic 2001), nerve compression (Charness *et al.* 1996), or a response to excessive, stereotypical, aberrant motor behaviours (Byl *et al.* 1996; Sanger and Merzenich 2000). Interestingly, behaviours that frequently precede the onset of focal dystonia usually include stressful repetitive hand movements (e.g. extensive practice on a musical instrument or overpractice using a new instrument or a new technique) under conditions of high speed, high force, or demanding productivity (Barr and Barbe 2002; Comella *et al.* 2003).

## Cortical plasticity in acquisition of skills

The brain is topographically mapped by sensation, motor performance and function (Merzenich *et al.* 1983) (see Fig. 19.2). When learning new sensory and motor skills, functional, topographical, and structural changes occur in the central nervous system (e.g. Jenkins *et al.* 1990). In the acquisition of fine motor

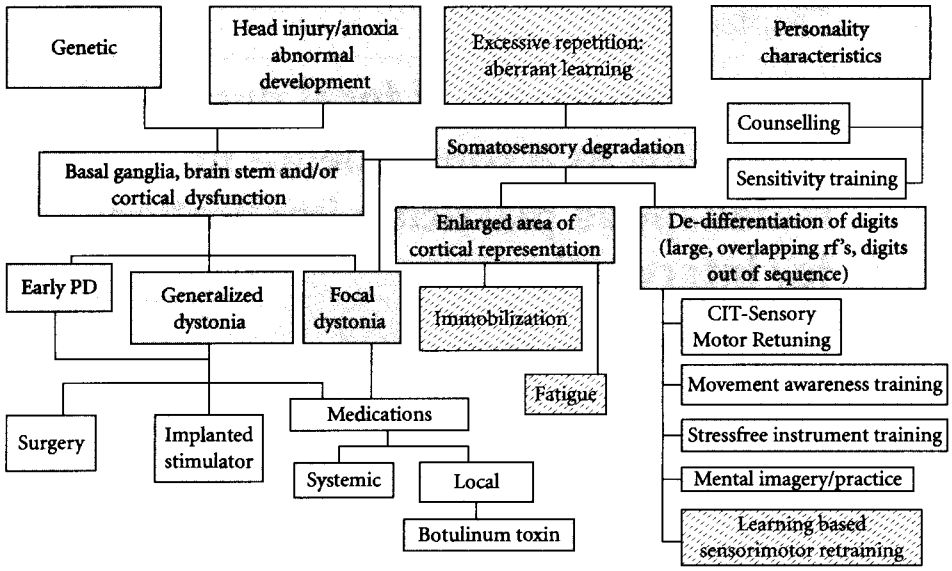



**Fig. 19.1** Aetiology pathophysiology ad management of focal hand dystonia. Although focal hand dystonia is considered an idiopathic movement disorder, there are multiple risk factors that are associated with the development of the condition, including repetitive overuse of the hand. There is evidence of degradation of the somatosensory topography of the hand.

skills, such as learning to play an instrument, specific skin locations are stimulated at different moments during the task, engaging competitive cortical processes that drive further differentiation of skin surface representations. Receptive fields in the sensory cortex become smaller, and are represented by more coherent activities generated by smaller populations of neurons (Recanzone *et al.* 1992; Xerri *et al.* 1999). The overall area of cortex representing the skin surfaces used commonly increases, as it is driven by the task to be represented in the cortex in progressively finer grain. Similar competitive training-driven plasticity outcomes apply for proprioceptive, kinesthetic, and motor representations. The overall result is cortical representations become more selective and refined. Time constants and space constants of cortical processes are *specialized* for the task at hand. This progressive, learning-driven specialization, emerging in parallel in multiple cortical areas, *is* skill acquisition (Recanzone *et al.* 1992).

The size of the receptive fields (i.e. the bandwidths of the filters representing information from limited skin regions or specific muscles) of cortical neurons may increase or decrease based on attended, rewarded practice (Xerri *et al.* 1999). The time constants governing dynamic processing of sensory inputs or controlling movement sequencing may be progressively shortened or lengthened, depending on the parameters of the behavioural tasks driving progressive changes

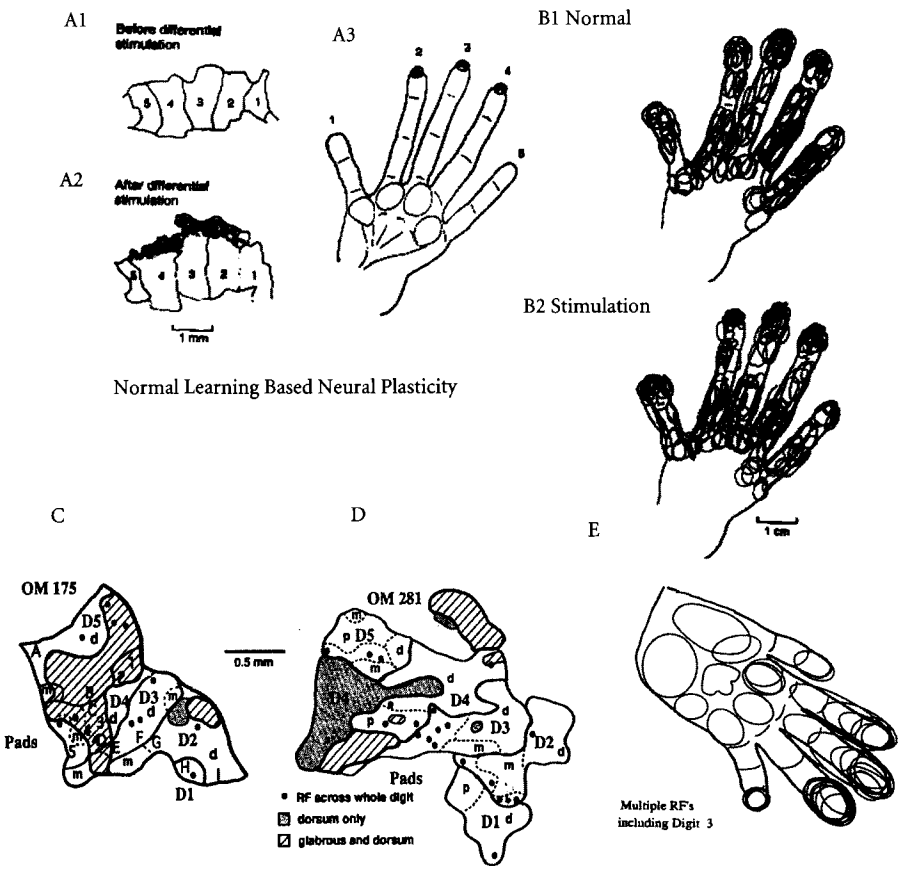



**Fig. 19.2** With learning based training, the cortical representation of the hand increases in size (A1 to A2) and the receptive fields on the digits become smaller and more dense (B1–B2). When there is excessive, near simultaneous repetitive movements, the topographical representation of the hand is degraded and the digits overlap with adjacent digits (C). In addition, the representation of the digits overlaps the ventral and dorsal surface overlap(D). On the digits, the receptive fields are large and a single digit has multiple receptive fields on adjacent digits (E).

(Xerri *et al.* 1998). When a skill is well learned, it can be performed automatically without conscious attention. At that point, there is no new learning and there is representational consolidation and a reduction of the size of the area of representation. Also, when stimuli are received near simultaneously or coincident in time, it can lead to a reduction in the size of the topographical representation and a loss of differentiation of receptive fields and overlapping receptive fields across an extended cortical representation, across multiple digits, or across dorsal and glabrous surfaces (Wang *et al.* 1995; Blake *et al.* 2002). These physical neural changes driven by repetitive, attended rewarded behaviours are referred to as 'neuroplasticity' or 'cortical plasticity'.

At least three specific training conditions must be met to drive large-scale cortical plasticity in sensorimotor control systems in the adult brain. First, learning activities must be progressive and repetitive to be enduring (Byl and Merzenich 2001). Physical motor or sensory exercises alone do not drive cortical changes; *learning* is required. Plasticity in the adult brain is strongly modulated as a function of behavioural state and context. Learning also depends on levels of attention and motivation as well as on perceived goal achievement, expectation, and 'reward or punishment'. These behavioural states, which are common conditions in musicians' professional lives, engage the release of modulatory transmitters delivered from cholinergic, dopaminergic, and adrenergic modulatory control system nuclei that regulate the processes of adult cortical plasticity (Juliano and Eslin 1991).

Second, learning must be *progressive*. Given the nature of the competitive plasticity processes in cortical networks, large-scale changes commonly require the use of progressive training exercises, moving from distinctions that an animal or human can achieve (i.e. that its cortical networks can resolve, from strongly correlated activities, in terms of the distribution of spatial and temporal activity), to fine, overlapping steps to ultimately achieve a 'higher' or substantially different level of discrimination or resolution (Kilgard and Merzenich 2001). However, the scale of plasticity is limited in any given cortical zone by the anatomical spread, temporal integration, and combination of inputs (Byl and Merzenich 2001). Processes of change at a given location in a given cortical area are controlled by the strengths and time distributions of inputs feeding that cortical network sector. Because synaptic plasticity is dependent on the coincidence of inputs, or 'Hebbian' (Hebb 1949), the biggest changes are driven by inputs that generate strongly correlated responses in cortical neurons and networks.

Third, enduring, large-scale plasticity of the capacity of the adult cortex to process elements of information requires *repetition and effort*. There is some trade-off between the number of trial repetitions and the strength of the

modulatory control effects that enable plastic change. Efficiently implemented training must amplify the enabling inputs as much as possible. At the same time, large-scale plastic changes in time and space require a heavy schedule of stimulus representation or movement. The strongest neural structural effects are driven by 'paced learning', particularly where the subject is heavily engaged in the same behavioural task over a series of successive days (Recanzone *et al.* 1992).

If the conditions necessary for plasticity are met, specific changes also depend on the precise timing of input events. The Hebb rule is: 'Neurons that fire together wire together'. For example, sensory inputs that arrive in a cortical representational sector at consistently separated moments in time generate separate, segregated, differentiated, competing representations in the cortex (Fig. 19.2). In contrast, sensory inputs that arrive in cortical networks nearly simultaneously in time are mutually co-strengthened, integrated, summarized, and *de*-differentiated; that is combined in representation in the cortex. Based upon this simple principle, cortical plasticity may result in either the refinement or the degradation (de-differentiation) of cortical representations of sensory feedback or movement (Wang *et al.* 1995; Byl *et al.* 1996; Byl and Merzenich 2001).

## Models of the origin of focal dystonia based on maladaptive plasticity

Beginning in the mid 1990s, Byl *et al.* (1996) trained monkeys in an attended, rewarded, highly stereotyped behavioural context with a task that simultaneously engaged relatively large hand surfaces with impulsive inputs, analogous to those delivered from a stereotypical power grip similar to that used to manipulate tools. These primates performed a repetitive, daily task over a number of successive months with daily training that demanded rapid opening/closing of the hand or rapid, precise spreading of the thumb and the index finger to make contact with two defined targets. In these experiments, early in the attended repetitive training paradigm, animals developed signs of repetitive strain injury that meant they would not work for a week and behaved as if the hand was painful. Some of the monkeys developed a classic focal hand dystonia, not uncommonly after being nursed through one or more bouts of repetitive strain injury. In the monkeys with dystonic posturing, there was profound degradation of the representation of the skin of the trained hand in the primary somatosensory cortex. The monkey performing the pointing task had an electrode implanted over the hand area, specifically representing digits one and two. This allowed the investigators to map the cortex during training. With the onset of the movement dysfunction, there was an increase in the overlap of digits one and two and the receptive field size increased to 87.68 $mm^2$ (compared with 8 $mm^2$

for normal owl monkeys). Signs of cortical columnar substitution were also noted. After 6 weeks, a number of sites could evoke responses from the glabrous skin of the hand and the lower face. At the time of the complete cortical mapping, the implanted electrode was found to be stable and the size of the receptive fields of the first two digits had increased to 101.41 $mm^2$ (Blake *et al.* 2002).

Degradation of the sensory feedback 'maps' was also seen in the untrained cortical hemisphere. Anatomically, the tendons, ligaments, and nerves were intact; there were no signs of nerve entrapment and no signs of acute inflammation (Topp and Byl 1999). The clinical expressions of focal hand dystonia observed in these monkeys were similar to those described in human patients. They included uncontrollable movements of the digits, co-contractions of flexors and extensors, difficulty releasing the flexed posture of the digits, or inability to perform delicate fine motor movements quickly and accurately. The trained primates who did not develop the dystonic hand postures were those working less intensively with fewer repetitions, more breaks, and shorter work periods.

The animal studies have been substantiated by studies of human subjects where degraded somatosensation has been reported in patients with focal hand dystonia. Patients with focal hand dystonia have poor spatial resolution thresholds over affected hand areas with reduced accuracy of localization (Bara-Jimenez *et al.* 2000). They have less salient sensations evoked by incident stimulation, degraded form perception (Byl *et al.* 1996), reduced kinesthetic and haptic abilities (Sanger *et al.* 2002), and problems with temporal spatial processing (Fiorio *et al.* 2003). All of these deficits are consistent with a *de*-differentiated representation of somatosensory feedback information.

Through cortical response mapping and imaging studies, several studies have also *directly* shown that somatosensory cortical representations of affected hand surfaces are de-differentiated in patients with focal hand dystonia. Representations of the separate digits in these individuals lose their topographic order (Fig. 19.3) and the spread of evoked activities are consistent with highly undifferentiated representations of sensory feedback from affected hand surfaces. There was a significant reduction in the amplitude in the early phase of the somatosensory evoked response when integrated over time for subjects with focal hand dystonia compared with controls (Fig. 19.4). However, within the subjects with focal hand dystonia, the somatosensory parameters were different in severity (Byl *et al.* 2002). It also appears that the somatosensory de-differentiation, excitatory, and inhibitory responses may vary by type of dystonia and, for example, differ in musician's cramp and writer's cramp (Rosenkranz *et al.* 2005). Further, the somatosensory, sensorimotor, and motor fields were all found to be abnormal compared with healthy controls (Lin *et al.* 2005).

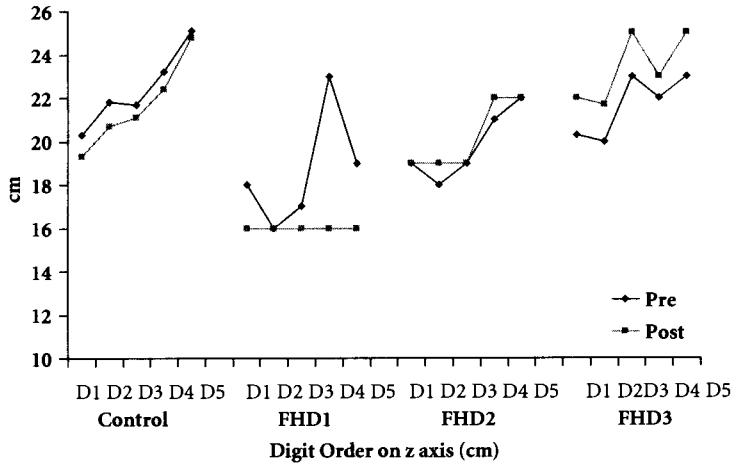



**Fig. 19.3** Digit representation in patients with focal hand dystonia (FHD). In healthy subjects, the representation of digits 1–5 as measured by magnetoencephalography should confirm an orderly progression from inferior to superior on the vertical axis (*z*-axis). In musicians with focal hand dystonia, the area of representation of the involved hand is smaller than the uninvolved hand and the precise and orderly differentiation of the digits is compromised. After participating in learning-based sensorimotor training (once a week for 12 weeks), three musicians increased the size of the area of the hand representation, but the digits still remained out of sequence.

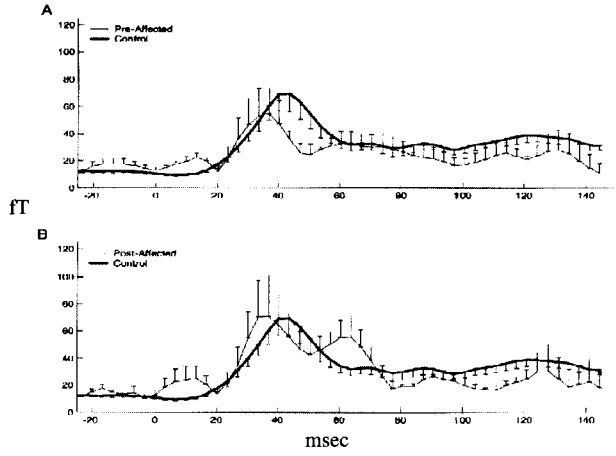



**Fig. 19.4** Somatosensory evoked field response: amplitude integrated over time. Based on the somatosensory evoked field response to light touch, the amplitude of the neural firing is reduced in patients with focal hand dystonia compared with controls. After learning-based sensorimotor training, the amplitude of the field response increases (similar to healthy subjects); however, there is some irregular neural firing in the early and late phases of the evoked response.

## Current treatment strategies for focal hand dystonia

Many of the musicians diagnosed with focal hand dystonia are forced to end their professional careers in music performance (Schuele and Lederman 2004). Although a review of the currently available pharmacological treatments of focal hand dystonia is outside the scope of this chapter, local injection therapy with botulinum toxin in the cramping muscles is only a symptomatic treatment and does not address the central origin of the disorder. Based on the hypothesis of aberrant learning, effective treatment must accomplish a reversal of the degradation of the central nervous system that accounts for one origin and expression of hand dystonia. Any effective therapy must re-differentiate the hand representations. That can only be achieved by intensive, goal-oriented, progressive learning-based training.

Conservative, non-pharmacological interventions such as exercise, retraining on the instrument, desensitization training and counselling, constraint-induced therapy (sensory motor retuning), immobilization, and fatigue and learning-based sensorimotor training have been tried. The first step towards successful rehabilitation is to stop the abnormal movement. The next step is to establish a

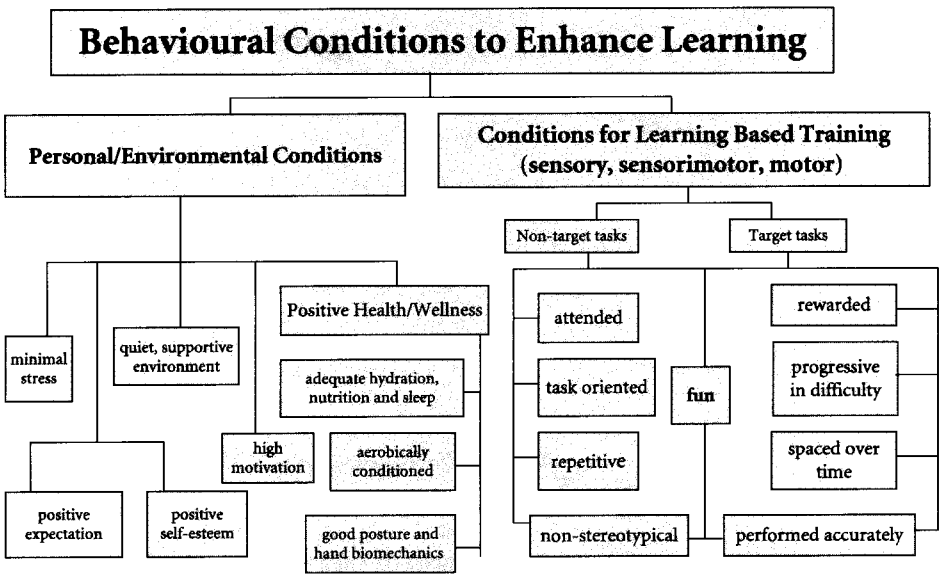
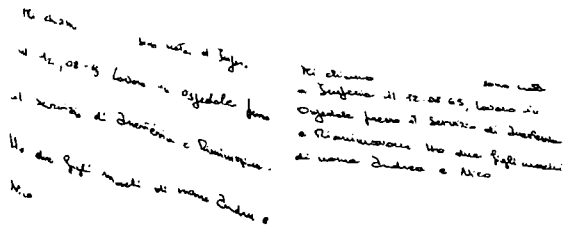



**Fig. 19.5** Principles of learning based training to maximize neural adaptation. For patients with focal hand dystonia, the goal is to restore normal fine motor control of the hand. This requires a re-differentiation of the somatosensory representation of the hand. A strong foundation of good health provides a good environment for learning-based training. Learning activities must be attended, repetitive, rewarded, fun, performed normally, progressive in difficulty, and spaced over time.

foundation of positive health and wellness where learning can take place. This includes good nutrition, hydration, stress management, aerobic exercise, good ergonomic hand use, and appropriate instrumental modifications.

The principles of neuroplasticity are most closely aligned to constraint-induced training, retraining at the instrument, immobilization, and learning-based sensorimotor training. In a study (Candia *et al.* 2002) of 11 musicians with focal hand dystonia, the unaffected fingers of the dystonic hands were constrained with splints. For 2 weeks, 2 to 3 hours a day, the patients were trained to use the dystonic fingers on fine motor tasks emphasizing instrumental play. When control of the most involved digits improved, the splints on the less involved digits were gradually removed and practice at the instrument was increased. Eight out of eleven musicians showed an improvement of their motor skills (Candia *et al.* 2002). In another study (Candia *et al.* 2003) patients with focal hand dystonia showed restored cortical organization after treatment.

Based on the evidence that topographical representation of the hand is abnormal in patients with focal hand dystonia, immobilization of the hand has been proposed as an appropriate treatment at the first signs of dystonia. During the past two centuries, neurologists have observed that rest or immobilization might be helpful in treating focal occupational dystonia (Lanska 2002) and several therapeutic devices have been proposed for this purpose. Limb immobilization influences several physiological and anatomical factors from changes in the muscle tissue, tendons, and bone to changes in the cerebral cortex, including effects on spinal motoneurons. In non-dystonic individuals, prolonged immobilization of a limb leads to a shrinkage of the cortical representation of the immobilized limb (Liepert *et al.* 1995). Interestingly, motor fatigue, which transiently reduces and reshapes motor cortical representation in normal subjects, transiently improves motor control in patients with focal task-specific dystonia (Pesenti *et al.* 2001). The transient improvement of focal occupational dystonia after a maximal voluntary fatiguing contraction is specific for dystonia and has been proposed as a diagnostic test (the hand grip test). Thus, the objective of immobilization is to place the finger, hand, wrist, and forearm into a static splint to stop all movement in order to decrease the size of the representational area of the involved hand (Priori *et al.* 2001). In that work the treatment consisted of the application of a static splint that immobilized the finger and the wrist joints on the dystonic side. Patients wore the splint for 24 h a day for 4–5 weeks. Although only a small number of subjects were included in the first trial (eight subjects), after 5–6 weeks of immobilization in a splint for 24 h per day (with the splint being removed only to clean the limb), all of the subjects showed significant improvement (up to 80%) in motor control of the hand in the target task several weeks after the splint was removed (Fig. 19.6). All patients had some transient clumsiness and the



(a) (b)

**Fig. 19.6** Writing samples before and after immobilization. Handwriting in a patient with writer's cramp before splinting (a). Handwriting 4 months after splint removal (b). There was improvement in both speed and quality (from Priori *et al.* 2001, with permission).

unusual sensation of being unable to control their hand when the splint was removed (e.g. at 1–3 weeks). Clumsiness prevented them from performing independent finger movements. None of the patients showed a worsening of their dystonia. This gain in function was retained for over 6 months in seven of the eight subjects participating in the study. Specific training activities designed to reorganize the cortical representation were not included in this paradigm to specifically evaluate the effects of limb immobilization without confounding factors.

Pesenti *et al.* (2001) carried out a similar study with a larger group of subjects with chronic focal hand dystonia, including some musicians, taking into account several clinical variables (Priori *et al.* 2003). In one group of patients there was variable improvement ('responders') and in the other group no improvement ('non-responders'). Despite the variability, most of the patients reported that immobilization induced at least some improvement. Immobilization seemed effective in young patients with a recent onset (< 1 year) of severe symptoms related to an overuse syndrome, whose dystonic symptoms improved transiently by more than 20% in the hand grip test. It could be hypothesized that patients with severe, established hand dystonia have more extensive degradation in the representation of the digits. Immobilization alone would not be sufficient to re-differentiate the digits and the integration of other rehabilitation protocols like learning-based sensorimotor training could produce better results

Learning-based sensorimotor training (LBST) is another approach to rehabilitation for patients with focal hand dystonia (Byl *et al.* 2000, 2001, 2002)(Fig. 19.5). The objective of LBST is to redefine spatial and temporal processing capacities in the sensory and motor cortex in order to restore the performance of target-specific skills. This strategy is consistent with repetitive task practice but with an initial concentration on improving sensory differentiation and not just graded, fine motor movement. Sensory conditions are designed to stimulate separate skin locations with non-stereotypical hand engagement trial by trial, in order to drive competitive Hebbian network plasticity to once again shrink the

cortical receptive fields (to give narrow input array bandwidths). Learning-based sensorimotor training is started with the least involved digits on the affected hand and then the most involved digits. If the sensory stimulation causes abnormal muscle contractions, training is moved to the unaffected side or the unaffected digits. Different exercise strategies address subtasks of the target tasks emphasizing different aspects of sensory feedback (i.e. somaesthesia, proprioception, kinesthesia, haptics). The patients initially make relatively simple distinctions about relatively large stimuli, with the difficulty of the task increasing as each patient mastered progressively more difficult distinctions. One set of tasks included reading Braille or embossed letters, making estimations of roughness, and performing discriminations of grating orientation. A second set of sensory tasks included flutter discrimination tasks and the recognition of small surface asperities. Sensorimotor and fine motor tasks designed to activate proprioceptive and sensorimotor feedback as well as cutaneous inputs were included. A series of haptic tasks required the subjects to manipulate and identify surface features and shapes of three-dimensional objects held in their hands.

The tasks are crudely progressive, beginning by using sensory inputs to shape the hand and performing partial tasks using normal movements. Then the sensorimotor training progresses to target fine motor tasks, beginning with tasks similar to the target task and progressing to the target task. The training may even begin in an atypical position (e.g. lying on the back to practice drumming). Initially, the patient simply learns to interface the target task by placing the hand on the instrument and then works at performing simple movements normally (e.g. dropping one finger on a key and maintaining the functional shape of the hand). An attempt is made to sustain the task at a level of difficulty at which errors could be made, i.e. requiring the subject's focused attention. Training stimuli are presented on a heavy, repetitive, daily schedule, meeting general standards for effective 'paced' learning. Feedback is provided by self-observation, verbal comments, auditory feedback from muscle contractions (Deepak and Behari 1999), mirror imagery (Sathian *et al.* 2000) and mental imagery (McBride and Rothstein 1979).

This training protocol has been tested in a series of patients with focal hand dystonia (Byl *et al.* 2000, 2003, Byl and McKenzie 2002). Subjects met a physical therapist for 1–2 h per week, and were trained on multiple tasks designed to engage the broad types of sensory inputs that provide feedback for fine hand motor control. The difficulty of each of these tasks was increased as subjects improved their performances. The task required constant attention. After appropriate instruction and demonstration in the clinic, subjects were instructed to practice the tasks for 30–60 min a day at home. Each patient received a handbook of instructions for home use. As a rule, to achieve effective reversals of the focal hand dystonia to a level at which the trainee could return to work,

training was continued for 6–12 weeks. In the first study, all 12 trained patients showed improvements in performance ability on a variety of tests of sensory discrimination, fine motor accuracy and speed, strength, flexibility, and functional independence. Motor control improved to an average of 70–94% of normal and all but three returned to their usual work.

## Summary

In this chapter we have summarized the evidence from animal and human studies suggesting that focal hand dystonia is due to multiple risk factors, one of which may be maladaptive plasticity. There is consistent evidence for somatosensory degradation as well an imbalance between inhibition and excitation at multiple levels of the central nervous system. Given a hypothesis of aberrant neural adaptation, the good news is that the central nervous system is clearly adaptable. Thus, one way to reverse learning-induced, negative behavioural changes underlying focal hand dystonia is for the patient to *learn* their way out of it by retraining the brain.

In the early stages of neural adaptation, when hand dystonia is developing, immobilization of the hand to stop the abnormal movement may shrink the size of the cortical hand representation and be associated with improved function. However, actual cortical mapping studies have not been carried out to confirm this in musicians with focal dystonia. Post-immobilization, the patient should be instructed in healthy biomechanical, stress-free hand-use strategies with learning-based retraining to restore sensory discrimination and fine motor control. In the more established and severe cases of focal hand dystonia, the individual must not only stop the abnormal posturing of the hand, but engage in learning-based sensory and motor retraining.

It is also clear that not all patients with focal hand dystonia are the same. There may be unique cortical changes and responses in musicians that differ from those with writer's cramp. Even those with the same type of focal dystonia may also demonstrate differences in response. Further, those with severe symptoms may differ from than those with mild symptoms. This variability may explain why not all patients respond the same to specific behavioural intervention strategies. Each patient needs to be carefully examined to try and match the intervention to the performance dysfunction.

We could not locate any blinded, large, randomized clinical trials that compared the efficacy of different intervention strategies. The effectiveness of clinical training can be limited by a number of problems. Patient compliance and the behavioural contexts of training-driven plasticity are difficult to monitor. In addition, attentional control is not assured, progressing of task difficulty is variable, and tasks are performed with minimal refined regulation. It is difficult

to deliver consistent and careful behavioural rewards for 'correct' performance trials and tasks can become dull and uninteresting rather than engaging and fun. Thus, research is still needed to determine the most efficient training strategies to maximize the effectiveness of learning. Training programs need to become efficient, possibly automated, where individuals not only have greater control of their own training process but where continuously adaptive challenges of stimulus discrimination/identification can be presented to enable a re-differentiation of cutaneous, proprioceptive, kinesthetic, haptic, and fine motor representations in the brains of affected patients. Despite the limitations of the evidence, recovery has been reported by some highly motivated individuals who are willing to perform the necessary dynamic, repetitive, attended, rewarded, progressive task practice that can sufficiently re-differentiate the sensory and motor cortex to restore the performance of fine motor skills.

## References

**Bara-Jimenez, W., Shelton, P., and Hallett, M.** (2000). Spatial discrimination is abnormal in focal hand dystonia. *Neurology*, **55**(12), 1869–1873.

**Barr, A.E. and Barbe, M.F.** (2002). Pathophysiological tissue changes associated with repetitive movement: a review of the evidence. *Phys Ther*, **82**(2), 173–187.

**Black, K.J., Ongur, D., and Perlmutter, J.S.** (1998). Putamen volume in idiopathic focal dystonia. *Neurology*, **51**(3), 819–824.

**Blake, D.T., Cheung, S., Bedenbaugh, P., Nagarajan, S., Lamb, M., and Merzenich, M.M.** (2002). Sensory representation abnormalities that parallel focal hand dystonia in a primate model. *Somatosens Mot Res*, **19**(4), 347–357.

**Breakefield, S.O.** (1997). The early onset of torsion dystonia gene (DYT1) encodes an ATD binding protein. *Nat Mol Genet*, **40**, 40–48.

**Byl, N.N. and McKenzie, A.** (2000). Treatment effectiveness for patients with a history of repetitive hand use and focal hand dystonia: a planned, prospective follow-up study. *J Hand Ther*, **13**(4), 289–301.

**Byl, N.N. and McKenzie, A.** (2002). The effect of sensory discrimination training on structure and function in patients with focal hand dystonia: 3 case series. *Arch Phys Med Rehab*, **84**, 1505–1514.

**Byl, N.N. and Merzenich, M.M.** (2001). Principles of neuroplasticity: implications for neurorehabilitation and learning. In: E. Gonzalez, S.J. Myers, J.E. Edelstein, J.S. Lieberman, and J.A. Downey (eds) *Downey and Darling's physiological basis of rehabilitation medicine*, pp. 609–628. Butterworth-Heineman, Woburn, MA.

**Byl, N.N., Merzenich, M.M., and Jenkins, W.M.** (1996). A primate genesis model of focal dystonia and repetitive strain injury: I. Learning-induced dedifferentiation of the representation of the hand in the primary somatosensory cortex in adult monkeys. *Neurology*, **47**(2), 508–520.

**Byl, N., Nagarajan, S.S., Newton, N., and McKenzie, A.L.** (2000). Effect of sensory discrimination training of structure and function in a musician with focal hand dystonia. *Phys Ther Case Rep*, **3**, 94–113.

**Byl, N.N., Nagarajan, S.S., Merzenich, M.M., Roberts, T., and McKenzie, A.** (2002). Correlation of clinical neuromusculoskeletal and central somatosensory performance: variability in controls and patients with severe and mild focal hand dystonia. *Neural Plast*, **9**(3), 177–203.

**Candia, V., Schäfer, T., Taub, E.** *et al.* (2002). Sensory motor retuning: a behavioral treatment for focal hand dystonia of pianists and guitarists. *Arch Phys Med Rehab*, **83**(10), 1342–1348.

**Candia, V., Wienbruch, C., Elbert, T., and Rockstroh, B.** (2003). Effective behavioral treatment of focal hand dystonia in musicians alters somatosensory cortical organization. *Proc Natl Acad Sci USA*, **200**, 7942–7946.

**Charness, M.E., Ross, M.H., and Shefner, J.M.** (1996). Ulnar neuropathy and dystonic flexion of the fourth and fifth digits: clinical correlation in musicians. *Muscle Nerve*, **19**(4), 431–437.

**Comella, C.L., L.S., Wuu, J., Stebbins, G.T., and Chmura, T.** (2003). Rating scales for dystonia: a multicenter assessment. *Mov Disord*, **18**, 303–312.

**Deepak, K. and Behari, M.** (1999). Specific muscle EMG biofeedback for hand dystonia. *Appl Psychophysiol Biofeedback*, **24**, 267–280.

**Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., and Taub, E.** (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, **270**(5234), 305–307.

**Elbert, T., Candia, V., Altenmüller, E., Rau, H., Sterr, A., and Rockstroh, B.** (1998). Alteration of digital representations in somatosensory cortex in focal hand dystonia. *NeuroReport*, **9**(16), 3571–3575.

**Fiorio, M., Tinazzi, M., Bertolasi, L., and Aglioti, S.M.** (2003). Temporal processing of visuotactile and tactile stimuli in writer's cramp. *Ann Neurol*, **53**(5), 630–635.

**Hebb, D.** (1949). *The organization of behavior*. Wiley, New York.

**Jankovic, J.** (2001). Can peripheral trauma induce dystonia and other movement disorders? Yes. *Mov Disord*, **16**, 7–11.

**Jenkins, W.M., Merzenich, M.M., Oachs, M., Allard, T., and Buic-Robles, E.** (1990). Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation. *J Neurophysiol*, **63**(1), 82–104.

**Juliano, S.L., M.W., and Eslin, D.** (1991). Cholinergic depletion prevents expansion of topographic maps in somatosensory cortex. *Proc Natl Acad Sci USA*, **88**, 780–784.

**Kilgard, M.P. and Merzenich, M.M.** (2001). Sensory input directs spatial and temporal plasticity in primary auditory cortex. *J Neurophysiol*, **86**(1), 326–338.

**Lanska, D.** (2002). Limb immobilization for the treatment of focal occupational dystonia (letter). *Neurology*, **58**, 991.

**Liepert, J., Tegenthoff, M., and Malin, J.P.** (1995). Changes in cortical motor area size during immobilization. *Clin Neurophysiol*, **97**, 382–386.

**Lin, P., Byl, N.N., and Nagarajan, S.S.** (2005). Abnormalities of sensorimotor magnetic fields in focal hand dystonia (abstract). In: *Annual Conference, California Chapter of the American Physical Therapy Association*, October 2005. Submitted to *NeuroImage*.

**Lockwood, A.** (1989). Medical problems in musicians. *N Engl J Med*, **320**, 221–227.

**McBride, E. and Rothstein, A.L.** (1979). Mental and physical practice and the learning and retention of of open and closed skills. *Percept Motor Skills*, **49**, 359–365.

**Ozelius, L., Hewett, J.W., Page, C.E.** *et al.*

**Pesenti, A., Priori, A., Scarlato, G., and Barkieri, S.** (2001). Transient improvement induced by motor fatigue in focal occupational dystonia: the handgrip test. *Mov Disord*, **16**(6), 1143–1147.

**Priori, A.** (2003). Immobilization for focal hand dystonia. In: *International Meeting of the French Neurological Society: Dystonia, October, 2003, Paris, France.*

**Priori, A., Pesenti, A., Cappellari, A., Scarlato, G., and Barbieri, S.** (2001). Limb immobilization for the treatment of focal occupational dystonia. *Neurology*, **57**(3), 405–409.

**Recanzone, G.H., Jenkins, W.M., Hradek, G.T., and Merzenich, M.M.** (1992). Progressive improvement in discriminative abilities in adult owl monkeys performing a tactile frequency discrimination task. *J Neurophysiol*, **67**(5), 1015–1030.

**Rosenbaum, F. and Jankovic, J.** (1988). Focal task-specific tremor and dystonia; categorization of occupational movement disorders. *Neurology*, **38**, 522–527.

**Rosenkranz, K., Williamon, A., Butler, K., Cordivari, C., Lees, A.J., and Rothwell, J.C.** (2005). Pathophysiological differences between musician's dystonia and writer's cramp. *Brain*, **128**, 918–931.

**Sanger, T.D. and Merzenich, M.M.** (2000). Computational model of the role of sensory disorganization in focal task-specific dystonia. *J Neurophysiol*, **84**(5), 2458–2464.

**Sanger, T.D., Tarsy, D.M., and Pascual-Leone, A.** (2001). Abnormalities of spatial and temporal sensory discrimination in writer's cramp. *Mov Disord*, **16**, 94–99.

**Sathian, K., Greenspan, A., Wolf, S.L.** (2000). Doing it with mirrors- a novel approach to stroke rehabilitation. *J Neural Repair Neurosci*, **14**, 73–76.

**Schuele, S. and Lederman, R.J.** (2004). Long-term outcome of focal dystonia in string instrumentalists. *Mov Disord*, **19**, 43–48.

**Sheehy, M.P. and Marsden, C.D.** (1982). Writers' cramp—a focal dystonia. *Brain*, **105**(3), 461–480.

**Topp, K.S. and Byl, N.N.** (1999). Movement dysfunction following repetitive hand opening and closing: anatomical analysis in owl monkeys. *Mov Disord*, **14**(2), 295–306.

**Wang, X., Merzenich, M.M.., Sameshima, K., and Jenkins, W.M.** (1995). Remodelling of hand representation in adult cortex determined by timing of tactile stimulation. *Nature*, **378**(6552), 71–5.23.

**Xerri, C., Merzenich, M.M., Jenkins, W., and Santucci, S.** (1998). Plasticity of primary somatosensory cortex paralleling sensorimotor skill recovery from stroke in adult monkeys. *J Neurophysiol*, **79**(4), 2119–2148.

**Xerri, C., Merzenich, M.M.., Jenkins, W.M., and Santucci, S.** (1999). Representational plasticity in cortical area 3b paralleling tactual-motor skill acquisition in adult monkeys. *Cereb Cortex*, **9**(3), 264–276.

Chapter 20

# Music performance anxiety

Jürg Kesselring

## Definition

Music performance anxiety (MPA) may be defined as 'a state of arousal and anxiety occurring before or while a person is performing non-anonymously in front of an audience producing a valuable or evaluated task touching on his/her self-esteem'. A description by a sufferer is particularly realistic and dramatic: 'my heart is beating up to the neck, I can hardly breathe, my knees are trembling, my hands are wet, I cannot think clearly, I am disorientated, I feel helpless and abandoned. Everybody will laugh at me, they will hate me, even kill me—and they will be right . . . now it is clear that I had completely overestimated myself. What a shame . . . !'. This description (most probably by someone who is not actually performing because this would impede his or her finding the words) contains all the reactions one would expect from heightened activity of the sympathetic nervous system, an adrenergic response which is usually set up in situations of fear and danger which induce 'fight or flight' reactions for survival. Music performance anxiety is a problem of the whole person and cannot be reduced to a disorder of motor control.

## Symptoms

Music performance anxiety is a special form of emotional behaviour (Damasio 1999) with aspects of:

1. Motor expressive behaviour: e.g. trembling or shaky voice, hands, arms, knees, or feet; irregular breathing, changes of facial expression (mimics), chewing finger nails, etc.
2. Reactions of the autonomous nervous system (sympathetic as well as parasympathetic): e.g. palpitation (rapid heart rate), sweating, flushing; shortness of breath or hyperventilation (clammy hands), fidgeting, pacing, loss of sensitivity or numbness in fingers, gastrointestinal disturbances (dry mouth, loss of appetite, nausea, vomiting), frequent trips to the toilet, increased urination, diarrhoea, insomnia.
3. Subjective feelings (which may in part be influenced by the motor responses and the reactions of the autonomic nervous system): e.g. feeling tense,

increased irritability and vulnerability, depression, anxiety, feelings of helplessness, being overwrought, panic, loss of control, anguish (feeling faint or dizzy), fear, fright, shock, numbness, feeling inadequate or worthless.

4. Cognitive appraisal (thoughts and reflections concerning feelings and bodily symptoms): e.g. concentration and memory problems (decreased concentration and increased distractibility, mind blanks, blocks, freezes, memory lapses), circular thinking (catastrophizing), confusion, difficulties making decisions, interfering negative thoughts.
5. Behavioural changes: e.g. insomnia, loss of appetite, avoidance reactions, hectic or paralyzed movement, talkative or mute, exhaustion.

## Causes of music performance anxiety

Is music performance really a situation so dangerous as to provoke these 'fight, flight and fright' reactions and why does the individual concerned feeling him- or herself to be in such a situation when performing music in public? Music performance anxiety may be present during performance and when not mastered may render it impossible, or it may be present for hours or days or even weeks before a performance. This is clear evidence for the fact that it is the imagined situation of performance which induces the same bodily reactions as does an acute danger in the real world. The imagination of a real frightening situation is apparently equally as effective in producing these bodily reactions as such a situation itself. The changes observed would have an adaptive function in relation to a physical threat but the after effects of the alarm system can, when not overcome before and during the performance, greatly interfere with its accomplishment (Wilson and Roland 2002). What is perceived as the danger leading to the anxiety reactions is mainly a fear of being judged and possibly devalued in one's self-esteem. The activity of the sympathetic nervous system, however, is Janus faced: it acts equally as a positive stimulant increasing alertness and focusing concentration. The quality of arousal is related to arousal with an inverted U-shaped curve: low levels of arousal are insufficiently motivating while excessive arousal interferes with performance because concentration is disrupted, memory blocks occur, and there is loss of steadiness in the hands and voice (the Yerkes–Dodson law; see Wilson and Roland 2002)

Considerable research has focused on MPA and coping strategies (Smith *et al.* 2000) but only a few studies have monitored physiological parameters before, during, and after performance, demonstrating increased heart rate, pulse, and blood pressure (and lower confidence) particularly during auditions (Fredrikson and Gunnarsson 1992). The latter study allowed the differentiation of a high-anxiety (tremor often experienced) from a low-anxiety (no tremor) group, and

importantly performance quality was rated higher in the latter group by an attending expert professor. Interestingly, however, raters blind to group assignment who listened to tape recordings and only had auditory cues made the reverse judgement, suggesting that quality judgement (at least in the tasks that were investigated) made by raters present at the performance was influenced more by factors associated with visual cues than by auditory ones—probably by behaviours like self-confidence and body posture.

Biological reactions such as fighting, fleeing, or avoidance make sense in a strictly biological realm but may become counterproductive in the social context in which one of the most feared situations is the possibility of being laughed at or ridiculed. Stage fright results from an imbalance between self-set claims and expectations and coping capabilities, and it arises particularly in situations which may endanger self-esteem, the affection and sympathy of others, or the continuation of a career.

Musicians perform complex psychomotor skills under conditions of high arousal when outcome is critical, performance is being evaluated, and competition is fierce. A person's appraisal of the demands of a performance, the perceived coping resources (e.g. ability and preparedness), and the 'meaning' or consequences (Smith *et al.* 2000) are all influenced by personality (self-esteem, confidence, auditory style) and motivational factors (perfectionism, practice time).

Music performance as a social situation gains significance by the presence and (mainly silent) participation of spectators and audience. In order to be able to participate the audience expects to encounter fulfilment of norms and standards but equally to be led by the person on stage to novel aesthetic experiences. Performance is an exchange of emotional values. Anxiety in such situations may arise through the suspected expectations of the audience. It is not just the fear of threatening events in nature but the fear of the reaction of other, largely unknown and anonymous, people. A sense of reality and autonomy must be balanced against dependency on expectations of the evaluation and appreciation by others. Pathological MPA may therefore be understood as a form of social phobia (Steptoe and Fidler 1987; Cox and Kennardy 1993). It is a mirror-world in which a person finds themself alone, dependent on the pre-conceived imagined estimations of others without being able to confront them in reality.

Performance is a conflicting mixture of attraction and avoidance: a challenge on the one hand of being heard and seen and on the other hand a fear of being exposed with possible failure. Such fear is more marked when a pre-defined role has to be performed (as is the case in music performance). There is general agreement that auditions are the most stressful performance situations because they combine scrutiny and evaluation with a socially inferior status and may have a direct influence upon career progression (Wilson and Roland 2002). The degree

of anxiety increases in direct proportion to the rigidity of the role, and this has to be taken into account by the person in that role well before the actual performance. The more we specialize and restrict our competence, the more vulnerable we become, and music performance is one of the highest possible specializations with narrowly defined stereotypes of bodily, mental, and emotional functions. The quality of a performance may be described, predicted, and analysed in a three-dimensional model of stress factors interacting with each other: Trait anxiety (susceptibility to stress), situation (environmental pressure), and task mastery (Wilson 2002). Expectations in our fantasy may grow into veritable monsters and become more dangerous than the ones we encounter in real daily life.

Performance anxiety is most often related to a kind of judge or censor, internal or external, real or imagined, a consequence of previous life experiences. Unfortunately, very often the 'inner censor' is critical, doubtful, and cowardly—the very opposite of the qualities we strive for when performing, namely self-confidence, persuasive power, and courage. Reconciliation of these contradictory tendencies is therefore a pre-requisite for coming to terms with MPA. The 'worry-factor' (Wise 1971) leads only to increased self-centredness and is attracting just what we are most afraid of.

## Emotional evaluation and self-esteem

At the base of the problem of MPA lies the anxiety of being unmasked in front of the public, to lose value (and power) in failure. It is the conflict between the way by which a person is valued by others, by the values and validations experienced previously, and between the self-given evaluation, i.e. between expectations, the way one would like to be seen, and between self-evaluation. Most importantly this private conflict takes place in public.

Emotions are created by inner engagement and involvement with the task. In MPA, shame, anger, and disorientation predominate. Anxiety is an emotion of the whole person, expressing itself as restlessness, irritability, insecurity, tension, clumsiness, sweating, freezing, etc. The problem in MPA is not anxiety itself but rather our attitude towards it. The narrow boundaries of anxiety may be overcome by courage on the basis of trust, which is the opposite of anxiety. A reason for trust is serious and adequate (not perfect) preparation. Practising means to make the difficult easy (or at least easier). Anxiety about failure most often results from projecting one's own self-destructive views and ideas onto others (the audience, jury, teachers, rivals). Shame is another emotional factor closely related to anxiety and always involved in MPA. Performing inevitably means exposing parts of oneself which are usually protected. The ambivalence of wanting to be heard and seen on the one hand and of wanting to protect intimacy and identity on the other often renders performance so difficult. Anger arises if

we are not able to fulfil the expectations of the inner censor or when perfectionism demands total control, which is never completely possible. Distractibility arises from inadequate concentration and lack of focus on the task and it may influence all cognitive domains even leading to complete disorientation.

## Prevalence

In 2002 a survey was conducted in three orchestras in Switzerland: the Zürich Chamber Orchestra (ZKO) (full-time professional musicians), the Glarner Musikkollegium (GMK) (part-time professional musicians), and the Orchester Gesellschaft Zürich (OGZ) (good amateurs) to which 50 musicians responded (for details see the Appendix).

When asked whether they ever experienced anxiety while performing, 25 of the 32 women declared to have MPA sometimes and 6 often; and of the 18 men responding 12 had MPA sometimes and 4 often. Only three amateurs declared that they never experienced MPA.

Stage fright was experienced more often during solo performance than when playing chamber music or in an orchestra concert, as is to be expected (Wilson and Roland 2002). More than half of the responders considered MPA to be disturbing often, 7/50 very much so. Symptoms began with equal frequency long before and shortly before performance, when entering stage and before difficult passages, respectively. These figures correspond well with those found among 155/650 professional orchestral players in The Netherlands who responded in a survey (Van Kemenade *et al.* 1995): in 59% MPA was considered serious enough to affect professional life, in 32% to a slight degree, in 32% to a moderate degree, in 28% it was very distinct, and in 9% of strong intensity. Anticipation of anxiety lasted for days (36%), weeks (10%), and even months (5%) and the conclusion was: 'that their lives would have been merrier, happier, more relaxed and less stressed and more satisfying'. Among the symptoms were (in decreasing frequency) palpitation, intestinal problems, tremor, difficulty concentrating, dry mouth, wet hands, difficulty with intonation, unreliable memory, increased sweating, and irregular breathing.

Similar figures have been found by Krawehl and Altenmüller (2000) with symptoms among 40 music students considered to be very disturbing in 30% and sometimes disturbing in 60%, and by Wesner *et al.* (1990).

When asked about personality type (on a scale from 0–10), 12 considered themselves to be introverted (7 women, 5 men) and 16 (11 women and 5 men) to be clearly extroverted while 22 (17 women, 5 men) ranged in between. Interestingly of those reporting MPA sometimes, twice as many considered themselves as being extroverted versus introverted, whereas those who experienced MPA often were equally distributed. On a scale for emotional stability,

MPA was reported with equal frequency in those who considered themselves to be stable and in those who considered themselves to be unstable. On a scale for curiosity and openness (ranging from hesitant/anxious to open/courageous) MPA occurred twice as often in those who considered themselves to be open and courageous. All of those with a feeling of psychological weakness had MPA.

When asked about the ability to relax, 10 rated themselves as being rather tense (4 women, 6 men) and 20 (11 women and 9 men) to be relaxed. Neither this question nor the questions regarding leisure satisfaction nor whether one would consider oneself as being a nervous person helped to predict the occurrence of MPA. Thirty participants could not agree to 'feel well in (even unfamiliar) society', phobias (height, open spaces, animals) were an issue in 11 women and 3 men, panic attacks occurred in 4, mainly among those who said they had MPA sometimes. This was rarer than in the study by Clark and Agras (1991) in which all participants fulfilled DSM-III criteria for social phobia.

## Treatment strategies

Cognitive-behavioural strategies are used by many musicians to cope with MPA: positive thought, visualization and distraction, muscle relaxation, deep breathing, proper nutrition and exercise (Gratto 1998). In our survey, when asked 'What do you do against MPA?' the answers were (in decreasing frequency) relaxation, exercise, autogenous training, sport, beta-blockers, and tranquillizers (only amateurs!). Half of them had the impression that their coping behaviour would be very helpful, a quarter indicated that it would help only a little.

Although the effects of coping behaviours on health and performance are of central importance, little research has been conducted on the efficacy of specific strategies (Smith *et al.* 2002). Experiences in sports medicine suggest that exposure to intermittent stressors may condition the body to respond positively to adrenergic stimulation. Physically fit individuals have a higher threshold for stressor intensity, a more rapid recovery to baseline levels, and an elevated endorphin concentration. Gradual incremental exposure to stressful situations may promote resilience and effort without distress.

### Beta-blockers

Beta-blockers block the receptors for the physical effects of the natural fight or flight response mentioned above. The ability of beta-blockers to help anxiety is related only to their blockade of beta receptors outside the brain; they will not help the emotional symptoms of stage fright (e.g. sleep problems or negative inner voices). It is important to find, under medical supervision, the ideal dosage by experimenting several times and under different circumstances before a major performance to get to know the personal effects and adverse reactions.

Beta-blockers are not approved for use in anxiety or stage fright, but they may be prescribed for this purpose because there is scientific literature to support this use. They do not make playing better just by relieving physical problems resulting from the fight or flight response, and some musicians feel that adrenalin helps their performance, giving them an edge that adds intensity to the performance. Beta-blockers do not directly improve emotional states, except to the extent that some musicians feel better when their physical problems are relieved, namely interacting with 'the obligate body-relatedness of feeling' (Damasio 1999). If performance anxiety manifests itself mainly in psychological ways (e.g. negative inner voices), beta-blockers will not help.

Scientific studies clearly show that beta-blockers significantly reduce symptoms that can hinder some people's playing. The musicians in the studies said they felt better about their performance after taking beta-blockers, and music critics consistently judged their performances to be better. There is, however, the danger that beta-blockers only help technical aspects of performance, but may in fact diminish emotional elements. A study of 2122 musicians in major US symphony orchestras showed that 27% reported taking beta blockers (72% of these mainly for auditions, 52% for solo recitals, 50% for difficult orchestral performances, 42% for concerto performances, and 4% before every performance) (Fishbein *et al.* 1988).

Negative self-talk often mediates performance (Wilson and Roland 2002) mostly resulting from fear of failure, of getting worried and tense, of making mistakes such as playing wrong notes, losing the timing, of not being perfect, of disapproval (negative past performance experiences), of losing control, of problems with distraction or memory problems, and lack of confidence in ability (not feeling good enough, believing others are better).

Instead one should try to focus on realistic, positive self-talk with clear, precise goals, perform regular mental rehearsals, and use many performance simulations. A voice of trust must be actively sought ('Whether you think you can do it or whether you think you cannot do it—you will be right!'), and curiosity must be maintained: the pleasure of adventures and discoveries, free from prejudices ('Link to the inner child').

Successful professional performers view anxiety or nervousness before a performance as a normal and even beneficial part of preparation (Wilson and Roland 2002). A very experienced teacher recommends creating a network of channels through which the player can transmit a potential flow of communication, *in spite of* the difficulties surrounding him. MPA does not occur in gypsies because they are not burdened with the responsibilities of our social system; their sole interest is the pleasure of the listener (Havas 1973). A summary of coping strategies with MPA may therefore be: 'Playing is never difficult—it is either made easy or remains impossible' (Havas 1973).

## Appendix: Details of survey of MPA in three Swiss orchestras (2002)

Orchestras surveyed: Zürich Chamber Orchestra (ZKO) (full time professional musicians); Glarner Musikkollegium (GMK) (part-time professional musicians), and Orchester Gesellschaft Zürich (OGZ) (good amateurs).

The following abbreviations are used: FS, female, sometimes; FO, female, often; MS, male, sometimes; MO, male, often.

1. Did you ever experience anxiety while performing?
2. How much is stage fright disturbing you?
3. When does music performance anxiety begin?
4. Which symptoms do you experience?
5. What do you do to combat music performance anxiety?
6. Does it help?

| | | |
|---|---|---|
| ZKO | Women (6) | Never 0, sometimes 6, often 0 |
| | Men (4) | Never 0, sometimes 3, often 1 |
| GMK | Women (9) | Never 0, sometimes 8, often 1 |
| | Men (4) | Never 0, sometimes 3, often 1 |
| OGZ | Women (17) | Never 1, sometimes 11, often 5 |
| | Men (10) | Never 2, sometimes 6, often 2 |

Stage fright was experienced more often during solo performance (39: FS 21; FO 5; MS 10; MO 3) than when playing chamber music (21: FS 13; FO 4; MS 3; MO 1) or in an orchestral concert (11: FS 6; FO 3; MS 1; MO 1).

| | |
|---|---|
| Very much (7) | FS 3, FO 3, MS 1, MO 0 |
| Often (31) | FS 19, FO 2, MS 8, MO 2 |
| Not very much (4) | FS 0, FO 0, MS 3, MO 1 |

| | |
|---|---|
| Long before (6) | FS 4, FO 2, MS 0, MO 0 |
| Shortly before (18) | FS 8, FO 3, MS 6, MO 1 |
| Entering stage (11) | FS 5, FO 3, MS 2, MO 1 |
| Difficult passages (14) | FS 7, FO 1, MS 4, MO 2 |

| Symptom | Number | Distribution |
|---|---|---|
| Palpitation | 30 | FO 4, FS 18, MO 1, MS 7 |
| Intestinal problems | 24 | FO 5, FS 18, MO 1, MS 0 |
| Tremor | 22 | FO 3, FS 14, MO 2, MS 3 |
| Concentration difficulty | 15 | FO 3, FS 10, MS 2 |
| Dry mouth | 14 | FO 3, FS 9, MO 1, MS 1 |
| Wet hands | 13 | FO 0, FS 8, MO 1, MS 4 |
| Intonation | 11 | FO 0, FS 10, MO 0, MS 1 |
| Unreliable memory | 9 | FO 1, FS 5, MO 0, MS 3 |
| Increased sweating | 7 | FO 0, FS 4, MO 1, MS 2 |
| Irregular breathing | 4 | FO 0, FS 4, MO 0, MS 0 |

| | |
|---|---|
| Relaxation exercise (19) | FS 11, FO 3, MS 4, MO 1 |
| Autogenous training (8) | FS 6, FO 1, MS 1, MO 0 |
| Sport (5) | FS 3, FO 1, MS 1, MO 0 |
| Beta-blockers (7) | (only amateurs) |
| Tranquillizers (3) | (only amateurs) |

| | |
|---|---|
| Yes, a lot (26) | FS 14, FO 3, MS 7, MO 2 |
| A little (8) | FS 4, FO 3, MS 1, MO 0 |

## References

**Clark, D.B. and Agras, W.S.** (1991). The assessment and treatment of performance anxiety in musicians. Am. J. Psychiat, **148**, 598–605.

**Cox, W.J. and Kennardy, J.** (1993). Performance anxiety, social phobia and setting effects in instrumental music students. J. Anxiety Disord, **7**, 49–60.

**Damasio, A.** (1999). *The feeling of what happens. Body and emotion in the making of consciousness.* Harcourt Brace & Co., New York.

**Fishbein, M., Middlestadt, S., Ottati, V., Straus, S., and Ellis, A.** (1988). Medical problems among ICSOM musicians: overview of a national survey. *Med Prob Perform Artists* March, 1–8.

**Fredrikson, M. and Gunnarsson, R.** (1992). Psychobiology of stage fright: the effect of public performance on neuroendocrine, cardiovascular and subjective reactions. *Biol Psychol*, **33**, 51–61.

**Gratto, S.D.** (1998). The effectiveness of an audition anxiety workshop in reducing stress. *Med Prob Perform Artists*, **13**, 29–34.

**Havas, K.** (1973). *Stage fright—its causes and cures*. Bosworth, London.

**Krawehl, I. and Altenmüller, E.** (2000). Lampenfieber unter Musikstudenten: Häufigkeit, Ausprägung und 'heimliche Theorien'. *Musikphysiol Musikermed*, **7**, 173–178.

**Mantel, G.** (2003). *Mut zum Lampenfieber—Mentale Strategien für Musiker zur Bewältigung von Auftritts- und Prüfungsangst*. Schott Musik Verlag, Mainz.

**Smith, A.M., Maragos, A., and Van Dyke, A.** (2000). Psychology of the musician. In: R. Tubiana and P.C. Amadio (eds) *Medical problems of the instrumentalist musician*, pp. 135–170. Martin Dunitz, London.

**Steptoe, A. and Fidler, H.** (1987). Stage fright in orchestral musicians: a study of cognitive and behavioral strategies in performance anxiety. *Br J Psychol*, **78**, 243–249.

**Tarr Krüger, I.** (1993). *Lampenfieber. Ursachen, Wirkung, Therapie*. Kreuz Verlag, Stuttgart.

**Triplett, R.** (1983). *Stagefright. Letting it work for you*. Nelson-Hall, Inc., Chicago, IL.

**Van Kemenade, J.F.L.M., Van Son, M.J.M., and Van Heesch, N.C.A.** (1995). Performance anxiety among professional musicians in symphonic orchestras: a self-report study. *Psychol Rep*, **77**, 555–562.

**Wesner, R.B., Noyes, R., and Davis, L.T.** (1990). The occurrence of performance anxiety among musicians (University of Iowa School of Music). *J Affect Disord*, **18**, 177–185.

**Wine, J.D.** (1971). Test anxiety and the direction of attention. *Psychol Bull*, **76**, 92–104.

**Wilson, G.D.** (2002). *Psychology for performing artists*, 2nd edn. Whurr, London.

**Wilson, G.D. and Roland, D.** (2002). Performance anxiety. In: R. Parncutt and G.E. McPherson (eds) *The science and psychology of music performance*, pp. 47–61. Oxford University Press, Oxford.

# Index


accented strokes in drumming 129–32, 133, 134–5
adjacent superior pre-motor cortex 142
agraphia, musical 208
amusia, oral–expressive (vocal/motor) 208–10
analytic memory 88
anger, and music performance anxiety 312–13
anterior cerebellar hemisphere 163
anterior corpus callosum 158
anterior superior parietal cortex 142
anticipation tendency *see* mean negative asynchrony
anticipatory behaviour
  and memory 46–9
  in violinists 114–15
anti-phase coordination 58
anxiety
  and focal dystonia 251, 262, 271
  stage fright *see* music performance anxiety
aphasia
  oral–expressive amusia 210
  preserved vocal performance in 207–8, 210
  song memory 230–1, 232–3, 234–5, *Plate 14*
appoggiaturas 85
apraxia, instrumental 208
assimilation in finger spelling 83, 86
associative memory 32
attention 34
  in song 226–7
audience discernment, and historical increases in performance skills 9–10
audiovisual mirror neurons 183–4
auditions, music performance anxiety in 310–12
auditory attention 34
  in song 226–7
auditory cortex 163–4
  memory 32
    motor representations of piano performance 40–1
  primary 142, 176
auditory feedback
  mean negative asynchrony 56–7
  and memory 45, 46
auditory memory 88
auditory perception 26–8, 34, 223–4

Bach, Johann Sebastian
  complexity of works 15
  diverse skills 10
  students 8
bassoonists, breathing technique 16
Beethoven, Ludwig van
  complexity of works 15–16
  Hammerklavier sonata 16
  improvisation 6
  premieres of orchestral works 6
behavioural laterality measures of singing capabilities 211–12
Bernstein, Nicolai 93–4, 110
beta-blockers for music performance anxiety 314–15
Binet, A. 91
birds, song production in 206–7
body awareness techniques for focal dystonia 277
botulinum toxin (BT) injections 274, 275–6, 279, 300
bounce in drumming 128–9
bowing in violinists, motor control 109–10, 112–15, 120–1, *Plate 1*
  historical perspective 110–12
  movements and forces 118–20
  previous work 116
  principal movement types 116–17
Brahms, Johannes
  cello bowing movements 116
  left hand injury 13
brainstem inhibition in focal hand dystonia 284–5
brain structures of musicians 141
  children prior to training 144–6
  differences from non-musicians 141–4, 159–60, 161, 165, *Plates 2–4*
  frontal 'auditory' and 'auditory-motor' regions, functional significance of 147–9, *Plate 7*
  instrumentalist versus non-instrumentalist children 146, *Plates 5–6*
brain volume of musicians
  children prior to training 145, 146
  differences from non-musicians 143–4
  instrumentalist versus non-instrumentalist children 146
Braitenberg, V. 111, 116
brass players, focal dystonia 267, 269
Breithaupt, R. M. 92
Broca's area and its right-hemispheric homologue 174, 182, 183–4
  pianists 175
  precursor in monkeys 184
  singing 210
  string players 194, 195, 196–8, 199, *Plate 10*

cadence 84
cellists
  focal dystonia 269
  movement analysis 116
central sulcus (CS) 159, 160
  pianists 178
  string players 194, *Plate 10*
cerebellum 142, 143, 146, 159
  animal studies 162
  finger function, role of inhibition in motor control of 243, 245

cerebellum (*contd*)
and handedness 159
phase correction 61
singing 213, 214, 216, 217, *Plate 12*
string players 194, *Plate 10*
cerebral lesions, acquired focal 210
cerebral localization of oral–expressive amusia 208–10
motor or global aphasia 207–8
child prodigies, and historical increases in performance skills 16–18
Chopin, Frederic, complexity of piano works 15
chronocyclography 111
chunking 88
cingulate cortex 243, 245
circular breathing 16
Clementi, Muzio 10
co-articulation 47, 49, 79, 82–4, 115
and expression 85–6
in violinists 115
cognition 25–6, 35
attention and perception 34
cognitive control of musical performance 30–1
memory *see* memory
perception 26–8
sight-reading 28–9
visuospatial functions 34–5
cognitive-behavioural strategies for music performance anxiety 314
cognitive control of musical performance 30–1
cognitive hypothesis of mental practice 44
Coker, Jerry 173
competition among musicians 9–10
competitions between musicians 6, 9
complexity of pieces of music, and historical increases in performance skills 11–18
compulsive disorders 240
conductors
attention 34
musicians' synchronization with movements of 71
continuous breathing 16
corona radiata 161
corpus callosum 145, 146, 158, 161, 168
cortical inhibition in focal hand dystonia 284–5
cortical plasticity
maladaptive, in focal dystonia 297–9, 304–5
treatment strategies 300–4
in skill acquisition 293–7
corticothalamic circuit 61
Cortot, Alfred 7
Courtier, J. 91
Cramer, Johann 13
cultural context, historical increases in performance skills 8–9
cymbals 136
Czerny, Carl
complexity of piano works 13
teaching role 10

dance 55, 68
Darwin, Charles 205
degrees of freedom problem
anticipatory behaviour 47
cognitive control of musical performance 30–1
neurophysiology 165
design of instruments, improvements in 7, 8, 13–14, 16
didgeridoo players, breathing technique 16
disorientation, and music performance anxiety 313
dissimilation in finger spelling 83–4, 86
distractibility, and music performance anxiety 313
dominant pitch 84
dorsolateral pre-frontal cortex (DLPFC) 174
pianists 175
double bass players, focal dystonia 269
Dragonetti, Domenico 15
drummers
movement analysis 125, 136–7
diversity of instruments 125–6
empirical studies 130–6
pitch and timbre changes 126
playing techniques 127–30
short contact times 126
time constraint and need for compromises 126–7
synchronization 70–1
dystonia, inhibitory neuronal circuitry 240
*see also* focal dystonia

education *see* training
ELITE system for movement analysis 99–100
embouchure dystonia 267, 269
pathophysiology 289
treatment 273–9
emotional evaluation, and music performance anxiety 312–13
ensemble playing, synchronization 70–1
enslavement in finger movements 80
violinists 114, 120
ergonomic changes, focal dystonia treatment 274, 276–7
errors, performance
correction, sensorimotor synchronization
synchronization of complex movements with complex sequences 70
synchronization of complex movements with simple sequences 68
synchronization of simple movements with complex sequences 66–7
synchronization of simple movements with simple sequences 59–61
hand/finger movements 87–8
memory
anticipatory behaviour 47
interpretive influences 41–2
movement analysis in pianists 97
expression, musical
and motor control 84–8
movement analysis in pianists 103–4
percussionists 126
eye–hand span, sight-reading 29

family history of dystonia 266, 268, 271
feedback control 26
Field, John 13

finger function, role of inhibition in motor control of 239–40, 246
  implications for musician's brain 245–6
  neural correlates of motor inhibition 244–5, *Plates 15–16*
  neural correlates of movement retrieval 242–4
  paradigm 240–2
finger movements *see* hand/finger movements
finger spelling 80–1
  co-articulation 83–4
fingerstops in violinists, movement analysis 111–12, 113, 114, 115, 117
flutes, historical developments 18
flutists, focal dystonia 267
focal cerebral lesions, acquired 210
  cerebral localization of oral-expressive amusia 208–10
  motor or global aphasia 207–8
focal dystonia
  aetiology 293, 294
  demographics, epidemiology, and phenomenology 265–72, 280
  finger function, role of inhibition in motor control of 240, 246
  maladaptive plasticity 293–305
  neurophysiology 283–90
  pathophysiology 270–1, 284, 287–90
  in pianists 104–5
  risk factors 268–70
  Schumann 251–63
  treatment strategies and results 265, 272–80, 300–5
force measurements
  in pianists 100–4, 111
  in violinists 110, 111, 118–20
frontal 'auditory–motor' region 148–9
frontal 'auditory' region 147–9
frontal BA 9: 182
frontal cortex 148–9
  musical imagery 216
frontal gyrus 142, 147
  inferior 142, 147, *Plate 11*
  lesions, and oral–expressive amusia 208
  singing *Plate 13*
  string players 197, *Plate 11*
frontal lobe, inferior 146, 148–9
frontal operculum 163, 182
  singing 228, *Plate 13*
  string players 194, 195, 197, 199–200
frontal polar areas
  musical imagery 215
  string players 196, *Plate 11*
fronto-insular cortex, role in oral–expressive amusia 208–10
frontotemporal area of right cerebral hemisphere (rFTA) 178–81

gait patterns in stroke and Parkinson's disease patients 68
Gát, J. 94–5
gender factors
  focal dystonia 251, 261, 268, 271
  singing capabilities 211
Glennie, Evelyn 125
global aphasia, preserved vocal performance in 207–8
goal-directed behaviour 120–1
grey matter of musicians
  animal studies 161–2
  Broca's area 183
  children prior to training 144, 145, 146
  differences from non-musicians 142, 143, 159–60
  frontal 'auditory' and 'auditory–motor' regions, functional significance of 147
  instrumentalist versus non-instrumentalist children 146, *Plates 5–6*
  practice intensity 160–1
guitarists
  complexity increases 16
  focal dystonia 276
  *see also* plucked instrument musicians, focal dystonia; string players
gyrus occipitalis *Plate 11*
gyrus supramarginalis 196, *Plate 11*

hand dystonia *see* focal dystonia
handedness
  cerebellum, asymmetric volume of the 159
  focal dystonia 268, 269
  historical increases in performance skills 13
  and reversed pianos 157
hand/finger movements 79–80
  biomechanical and neural constraints 80–2
  co-articulation 82–4
  and expression 84–8
  memory
    anticipatory behaviour 46–9
    transfer of performance skills 42–3
  pianists *see* pianists, movement analysis
  violinists *see* violinists, fingering and bowing, motor control
harmonic processing in song 225, 226
harpists, complexity increases 16
Haydn, Joseph, complexity of works 12, 13, 16
Helmholtz, H. v
Heschl's gyrus lesions, and oral–expressive amusia 209
hippocampus 143
  animal studies 162
historical increases in performance skills 3–4, 18–20
  competition among musicians 9–10
  complexity increases 13–18
  cultural context 8–9
  empirical evidence 11–13
  reasons 6–8
  retrospective assessment of skills 4–6
  specialization and deliberate practice 10–11
Horowitz, Vladimir
  expression 85, 86
  practice 175
  specialization 10
Hummel, Johann Nepomuk 251, 252, 254–5, 256
hypermusia in aphasic syndromes 208

imagery abilities and memory 46
imagined performance
  cerebral substrates of musical imagery 215–16, *Plate 12*
  by pianists, brain activation during 175–6
  by string players, brain activation during 194–6, 197–8
immobilization therapy for focal dystonia 280, 301–2, 304
improvisation 6–7
individuation, lacking in finger movements (enslavement) 80
  violinists 114, 120
inferior frontal gyrus 142, 147
  string players *Plate 11*
inferior frontal lobe 146, 148–9
inferior parietal lobe 146
inferior temporal lobe 142
inhibition, role in motor control of finger function 239–40, 246
  implications for musician's brain 245–6
  neural correlates of motor inhibition 244–5, *Plates 15–16*
  neural correlates of movement retrieval 242–4
  paradigm 240–2
inhibitory circuits, excitability in focal hand dystonia 284–5
in-phase coordination 58
instruction *see* training
instrumental apraxia 208
instrument design, improvements in 7, 8, 13–14, 16
insula
  lesions, and oral–expressive amusia 209
  singing 213, 215, 216, 217, *Plate 13*
internal capsule 161
interpretive influences on performance memory 41–2
intracortical facilitation (ICF) 167
intracortical inhibition (IHI) 167
  focal hand dystonia 285
intra-parietal sulcus 147
intrasylvian cortex
  musical imagery 216
  singing 213, 214, 215, 216, 217
Iowa piano camera 94
IQ, effect of musical training on 35

jazz musicians, synchronization 70–1

keyboard players
  brain structures 143, *Plate 3*
  focal dystonia 267, 268, 269
    treatment 276
  improvisation 6
  movement analysis, historical overview 91
  specialization 10
keyboard-related parameters, analysis of 96–9
keyboards, complexity increases 16
kinesthetic memory 88
Klavier-Cembalo 91
Kraft, Nikolaus 16
Kreisler, Fritz 7
kymographs 92

language processing and music, comparison between 183
lateral inferior temporal lobe 142
lateralization
  focal dystonia 267–8, 269
  historical increases in performance skills 13
  singing 210, 211–15, 216, 217, 227–8, 229
    animal studies 206
leading tone 85
learning-based sensorimotor training (LBST) 302–5
left inferior frontal gyrus 142
linear phase correction model 59
lip-opening asymmetries, singing 212
Liszt, Franz
  improvisation 6
  Paganini Etude 91
  playing of Beethoven's Hammerklavier sonata 16
  violin technique transferred to piano 16
long-term memory 32
long-term working memory (LTWM) theory 33
lute players, complexity increases 16

marching 68
Matthay, T. 92
Mayer, Charles 254
McFerrin, Bobby 16
mean negative asynchrony (MNA)
  complex sequences 62, 64
  simplex sequences 56–7
melodic intonation therapy 208, 230
memory 28, 31–3, 39–40, 50
  anticipatory behaviour 46–9
  hand/finger movements 88
  interpretive influences 41–2
  mental and physical practice 43–6
  motor representations of piano performance 40–1
  sight-reading 29
  song 229–31
    aphasic patients 232–3, *Plate 14*
    neurologically intact participants 231–2
  transfer of performance skills 42–3
mental plans of movements 31
mental practice
  memory 43–6
  movement analysis in pianists 97
  music performance anxiety 315
  pianists 175–6
Merlin, Joseph 91
mesial Heschl's gyrus 142
metronomes
  synchronization of complex movements with 68–70
  synchronization of simple movements with 56–61
middle frontal gyrus 147
MIDI keyboards, movement analysis 96–9
  in musicians' medicine 100–1, 104–5
mirror neuron system 147, 148, 163, 183–4
  string players 199–200
mismatched negativity (MMN) 28
Moeller stroke 132
Moscheles, Ignaz 251, 252, 255
motor amusia 208–10

motor aphasia, preserved vocal performance in 207–8, 210
motor cortex
  focal hand dystonia 287
  memory 41
  primary 143, 159–60, 164, 166, 167, 174
    animal studies 162
    finger function, role of inhibition in motor control of 243, 245, *Plate 15*
    and maximum tapping frequency 154, 155
    pianists 175, 176
  singing 217, *Plate 12*
motor equivalence 121
motor feedback, and memory 45, 46
motor representations of piano performance 40–1
motor variance, sensorimotor synchronization 57
movement primitives 79
Mozart, Wolfgang Amadeus
  child-sized violin 18
  complexity of works 12–13, 15
  piano and violin skills 8
  Pushkin on v
  training 9
Mozart effect 234
  visuospatial functions 34
musical agraphia 208
musician's cramp/dystonia *see* focal dystonia
music performance anxiety (MPA)
  causes 310–12
  definition 309
  emotional evaluation and self-esteem 312–13
  prevalence 313–14, 316–17
  symptoms 309–10
  treatment strategies 314–15
music therapy 234–5
  melodic intonation therapy 208, 230

negative self-talk, and music performance anxiety 315
Nelson, W. L. 116
nerve entrapment, and focal dystonia 271
neuroplasticity
  maladaptive, in focal dystonia 297–9, 304–5
    treatment strategies 300–4
  in skill acquisition 293–7
neurovascular coupling 213
non-specific exercises for focal dystonia 274, 277–8

oboists, breathing technique 16
off-beat tapping, sensorimotor synchronization 64
Omega Sign (OS) *Plate 3*
on-beat tapping, sensorimotor synchronization 64
opercular cortex lesions, and oral–expressive amusia 209
oral–expressive amusia 208–10
Ortmann, O. 92–3, 94
overlearning 87
overuse injury
  and focal dystonia 271, 302
  in pianists 100

paced learning, and cortical plasticity 297
  learning-based sensorimotor training 303
Paganini, Niccolo, guitar technique transferred to violin 16
pain
  in focal dystonia 268, 271
  in pianists 100
parietal cortex 146, 147
  pianists 179
  posterior (PPC) 164, 174
    finger function, role of inhibition in motor control of 243, 245
    pianists 175, 176
    string players 194, *Plate 10*
parietal lobe, superior (SPL) 159, 163, 164
  finger function, role of inhibition in motor control of 243, 245
  pianists 177
Parish-Alvars, Elias 16
Parkinson's disease patients, gait improvements with music 68
pars opercularis *Plate 11*
Paumann, Conrad 16
pedagogical retraining in focal dystonia 274, 277–8, 279, 280, 300–5
perception 26–8, 34, 223–4
perceptual representation 32
perceptual span, sight-reading 29
perceptual variability, sensorimotor synchronization 57
percussionists
  movement analysis 136–7
    diversity of instruments 125–6
    empirical studies 130–6
    pitch and timbre changes 126
    playing techniques 127–30
    short contact times 126
    time constraint and need for compromises 126–7
  synchronization 71
perfectionism, and focal dystonia 251, 262, 271
performance conditions, and historical increases in skills 6
performance errors *see* errors, performance
period correction, sensorimotor synchronization
  complex movements with complex sequences 70
  simple movements with complex sequences 62, 66–7
  simple movements with simple sequences 59–61
personality factors
  focal dystonia 251, 262, 271
  music performance anxiety 313–14
phase correction, sensorimotor synchronization
  complex movements with complex sequences 70
  complex movements with simple sequences 68
  simple movements with complex sequences 62, 66–7
  simple movements with simple sequences 59, 60–1
phoneme monitoring in song 225
pianists
  brain activation when playing 173–4, 184
    distributed network 174–5
    EEG correlates of auditory sensorimotor integration during piano learning 177–81, *Plate 9*

pianists (*contd*)
imagined performance 175–6
maps, links, and plasticity 176–7
musical speech or musical language? 183–4
shared core network? 182–3
cognition, sight-reading 29
focal dystonia 267, 283
historical increases in performance skills
complexity increases 13–15, 16
empirical evidence 11–13
reasons 6
retrospective assessment of skills 5
specialization 10
memory 40, 88
anticipatory behaviour 47–9
interpretive influences 41–2
mental and physical practice 44, 45–6
motor representations 40–1
transfer of performance skills 42–3
motor representation 153–4, 169
anatomical peculiarities of brains 158–62
neurophysiological peculiarities in sensorimotor system 162–9, *Plate 8*
specific motor functions 154–8
movement analysis 91, 95, 96–9, 105–6
anticipatory behaviour 47–9
biomechanical and neural constraints 80
co-articulation 84
expression 85–6, 88
historical overview 91–5, 96, 111
keyboard-related parameters 96–9
in medicine 100–5
transfer of performance skills 42–3
upper extremities 99–100
and percussionists, similarity of challenges 136
Schumann 251–63
synchronization 68–9, 70
*see also* keyboard players
pianist's cramp *see* focal dystonia
pianos
historical developments 13–14
reversed 157–8
pitch changes, percussionists 126
pitch perception 27–8
singing 216
planum polare 228
planum temporale 142
plucked instrument musicians, focal dystonia 267, 268, 269
polyphonic playing, historical perspective 16
Popova, T. 93–4, 110–11
positive self-talk, and music performance anxiety 315
posterior corpus callosum 158
posterior middle frontal gyrus 147
posterior parietal cortex (PPC) 164, 174
finger function, role of inhibition in motor control of 243, 245
pianists 175
imagined performance 176
string players 194, *Plate 10*
practice
brain activation of pianists during imagined performance 175–6
cognitive control of musical performance 30
cortical plasticity in skill acquisition 294–7
focal dystonia 251, 262, 270–1, 283
pathophysiology 288, 289–90
hand/finger movements 87
historical increases in performance skills 8, 11
memory 43–6
anticipatory behaviour 47
motor representation in pianists and string players 153–4, 160–1
movement analysis in pianists 97
and music performance anxiety 312
*see also* rehearsals
pre-central gyrus 143, 159, *Plate 3*
lesions, and oral–expressive amusia 208–9
musical imagery 216
singing 213, 214, 216
pre-cuneus
finger function, role of inhibition in motor control of 243, 245
pianists 175
imagined performance 176
pre-frontal areas
dorsolateral pre-frontal cortex (DLPFC) 174
pianists 175
string players 199
pre-motor cortex (PMC) 148–9, 159–60, 164, 174, 182
finger function, role of inhibition in motor control of 243, 245
lesions, and oral–expressive amusia 209
pianists 175
imagined performance 176
string players 194, *Plate 10*
pre-supplementary motor area (pre-SMA) 165, 166
primary auditory cortex 142, 176
primary motor cortex 143, 159–60, 164, 166, 167, 174
animal studies 162
finger function, role of inhibition in motor control of 243, 245, *Plate 15*
and maximum tapping frequency 154, 155
pianists 175
imagined performance 176
maps, links, and plasticity 176
primary sensorimotor cortex 142
pianists 175
string players 194, 195, 196
primates, song production in non-human 207
processing levels in song 226–7
psychoneuromuscular hypothesis of mental practice 44
Pushkin, Alexander v

Rachmaninov, Sergei Vassilievich 10
rate limits of synchronization 58–9, 69
Ravel, Maurice 13
rebound in drumming 128–9, 130, 136
rehearsals 6–7
*see also* practice
repetitive strain injury 297
retraining in focal dystonia 274, 277–8, 279, 280, 300–5

retrieval structure
  long-term working memory theory 33
  skilled memory theory 32
retrospective assessment of performance skills 4–6
Rubinstein, Anton
  practice 175
  variability in performance 7

Salieri, Antonio
  Pushkin on v
  training 8
Schubert, Franz 116, 256
Schumann, Robert
  focal dystonia 251–63
  Toccata Opus 7: 251, 255, 258–9, 261, 262
Scrijabin, Aleksandr Nikolayevich 13
Seashore, C. E. 94
self-esteem, and music performance anxiety 312–13
semantic priming in song 225, 226
sensorimotor associations 26
sensorimotor cortex 142
  finger function, role of inhibition in motor control of 243
  and maximum tapping frequency 155
  neurophysiological peculiarities 162–9
  pianists 175, 179–81
  singing 216
  string players 194, 195, 196, 198–9, *Plates 10–11*
sensorimotor synchronization *see* synchronization
sensory input and motor output, interaction between
  in focal hand dystonia 287
  in writer's cramp 286–7
sensory–motor networks in singing and speaking, comparative approach 223–4, 233–5
  attention and levels of processing in song 226–7
  functional mapping of brain structures involved in song processing 227–9, *Plate 13*
  linguistic and musical dimensions of song, relationship between 225–6
  song cognition, reasons for studying 224–5
  song memory 229–33, *Plate 14*
sensory perception, integration, and mapping in focal hand dystonia 285–6
shame, and music performance anxiety 312
shock spectrum, percussion instruments 127
short-term memory 32
  brain activation in string players 199
  eye–hand span 29
sight-reading
  brain activation 183
    during piano playing 174, 177
  cognition 28–9
  historical increases in performance skills 6–7
sign language 80–1
  co-articulation 83–4
singing, cerebral correlates 205–6, 216–17
  acquired focal cerebral lesions, data from subjects with 207–10
  behavioural laterality measures of singing capabilities 211–12
  birds and non-human primates 206–7
  cerebral substrates of musical imagery 215–16
  experimental behavioural and electrophysiological data 210–13
  haemodynamic functional brain imaging of singing capabilities 213–15
  Wada test 210–11
  *see also* sensory–motor networks in singing and speaking, comparative approach
skilled memory theory 32–3
slow music, tapping to 64–8
snare drums 128–9
social phobia 311
soloists
  focal dystonia 271
  historical increases in performance skills 9
  music performance anxiety 313
somatosensory cortex
  finger function, role of inhibition in motor control of 243, 245
  focal dystonia 262, 270, 285–6, 297–8
  memory 40–1
  writer's cramp 286
spatial discrimination in focal hand dystonia 285
specialization, and historical increases in performance skills 7–8, 10–11
speech therapy 233
Spencer, Herbert 205
spinal inhibition in focal hand dystonia 284–5
splints for focal dystonia 277, 301–2
Stadler, Anton 15
stage fright *see* music performance anxiety
stave bars, and memory 33
striking velocities
  of drumming 132–3, 134, 135, 137
  of piano playing 137
string players
  brain activation 189–91, 199–200
    Broca's area and its right-hemispheric homologue 196–8
    electroencephalography 193–4
    functional MRI 192–3
    paradigms 191–2
    patterns 194–6, *Plates 10–11*
sensorimotor cortical network 198–9
  brain structures 143, *Plate 3*
  focal dystonia 267, 268, 269
    treatment 276, 278
  motor representation 153–4, 169
    anatomical peculiarities of brains 158–62
    neurophysiological peculiarities in sensorimotor system 162–9, *Plate 8*
    specific motor functions 154–8
stroke victims, gait improvements with music 68
strongly metrical (SM) beats 63
structural bars, and memory 33
structure of music, and sensorimotor synchronization 64–8
subcallosal gyrus 215
superior parietal lobe (SPL) 159, 163, 164
  finger function, role of inhibition in motor control of 243, 245
  pianists 177

superior temporal gyrus (STG) 158, 182
  pianists 178
  string players 196, 197, *Plate 11*
supplementary motor area (SMA) 162, 164, 165, 166, 174, 182
  finger function, role of inhibition in motor control of 243, 245
  singing *Plate 12*
  string players 194, 195, 196, 198–9, *Plates 10–11*
supramarginal gyrus 182
  memory 32
suspensions 85
symbolic hypothesis of mental practice 44
synchronization 55–6, 71
  between bowing and fingering in violinists 112, 115
  of complex movements with complex sequences 70–1
  of complex movements with simple sequences 68–70
  of simple movements with complex sequences 61–8
  of simple movements with simple sequences 56–61

tapping a beat
  synchronization with complex sequences 61–8
  synchronization with simple sequences 56–61
tapping frequency 154–7, 166
temporal discrimination in focal hand dystonia 285
temporal gyrus 158
  song *Plate 13*
  superior (STG) 158, 182
    pianists 178
    string players 196, 197, *Plate 11*
temporal lobe 142, 146
  musical imagery 216
  string players 197, 198
thalamus
  musical imagery 215, 216
string players *Plate 10*
Thalberg, Sigismund 15, 16
'third hand' technique 15, 16
three-dimensional movement analysis
  in pianists 99–100, 101
  in violinists 112–13
timbre changes, percussionists 126
timbre perception 27
timekeeper variance, sensorimotor synchronization 57
time-sharing interference paradigm, singing 211–12
timing
  drumming 125, 133–6
  inhibition, role in motor control of finger function 239–40
tonal music 84–5
tonic pitch 84–5
torticollis 287
Tourette syndrome 240
training
  and cognition 25
    attention 34
    IQ 35
    perception 27–8
  effects on brain structure 143–6, *Plate 4*
  hand/finger movements 86
  historical increases in performance skills 8, 9, 10, 13, 16–18
  of percussionists 136–7
transfer of performance skills/learning
  memory 42–3
  movement analysis in pianists 97
trauma, and focal dystonia 271
Trendelenburg, W. 110, 115, 118
trihexyphenidyl 273–5, 279
trombone players
  focal dystonia 267
  memory 45
typing, finger movements in 80, 81
  co-articulation 83

unilateral transient inactivation of cerebral hemispheres (Wada test) 210–11, 212–13
unspecific exercises for focal dystonia 274, 277–8
upper extremity movement analysis in pianists 99–100

variability
  of performance 7
  synchronization of simple movements with simple sequences 57–8
ventral pre-motor cortex 163
verbal memory 33
viola players
  focal dystonia 269
  solo function 9
violinists
  fingering and bowing, motor control 88, 109–10, 112–15, 120–1, *Plate 1*
    bowing movements 116–20
    historical perspective 110–12
  focal dystonia 269
    posture 267
  historical increases in performance skills
    complexity increases 16
    reasons 6, 7
  memory 88
    motor representations 41
  motor representation 153
    anatomical peculiarities of brains 159
    neurophysiological peculiarities in sensorimotor system 163, 165
    specific motor functions 154
violins, historical developments 7
violoncello players
  focal dystonia 269
  movement analysis 116
visual memory 88
visuospatial functions 34–5
Vivaldi, Antonio 15
vocal amusia 208–10

Wada test 210–11, 212–13
Wagner, C. 95

Wagner, Richard 16
weakly metrical (WM) beats 63
Wernicke's area 183, 197, 198
whip-cracking action, percussionists 127–8
white matter of musicians
  children prior to training 144, 145
  differences from non-musicians 143, 161
  instrumentalist versus non-instrumentalist children 146
Wieck, Clara 256, 257, 259
Wieck, Friedrich 251, 253–4, 256, 257
Wing–Kristofferson model of variability 57, 59, 68
Wittgenstein, Paul 13
woodwind players, focal dystonia 267, 268, 269
  treatment 276, 278
working memory 32
  brain activation in string players 199
  eye–hand span 29
writer's cramp 240, 246, 261
  family history 266
  immobilization therapy 302
  and musician's dystonia, differences between 283, 298, 304
  pathophysiology 270, 287, 289, 290
  sensory input and motor output, interaction of 286–7
  somatosensory cortex 286

The motor actions that can be witnessed as a virtuoso musician performs can be so fast, so accomplished, so precise, as to seem somehow superhuman. The musician has to produce the movements, monitor those they have already made and the subsequent result, co-ordinate their hands, fingers, eyes, and perhaps throat and diaphragm. These achievements are of course the product of hundreds, even thousands of hours of practice — playing scales and studies, time and time again. But those hours of practice by no means guarantee that great musicianship will result. This technical prowess has to be combined with a range of other, perhaps less tangible, qualities.

This book explores the secrets of musical virtuosity. It presents a comprehensive account of music and motor cognition, examining the neural basis of music making — our understanding of which is just starting to be enhanced by brain imaging. It considers the effect on our brains of prolonged music making. It explores the motor processes across a range of instruments (vocal, string, wind, percussion) and within different performance situations. It also considers what happens when things start to go wrong — why motor problems occur in so many professional musicians in later life, and the possible therapies for such problems.

Music is a topic of considerable interest within the brain sciences. With contributions from leading psychologists, neuroscientists, and neurologists, this book makes a unique contribution to our understanding of music and the brain.

## ALSO PUBLISHED BY OXFORD UNIVERSITY PRESS

**Musical Excellence:** *strategies and techniques to enhance performance*
Aaron Williamon

**The Cognitive Neuroscience of Music**
Isabelle Peretz and Robert Zatorre

**The Child as Musician:** *a handbook of musical development*
Gary McPherson

**Musical Communication**
Dorothy Miell, Raymond MacDonald, and David Hargreaves

**Motor Cognition**
Marc Jeannerod

*Front cover image* by Alexander Sitnikow in 'Alexander Puschkin: Mozart und Salieri. Szene aus dem Faust', reproduced courtesy of Grupello Verlag.

OXFORD
UNIVERSITY PRESS

**www.oup.com**

ISBN 0-19-929872-6

9 780199 298723

John Rink (ed.)



# La interpretación musical

Alianza Música

John Rink (ed.)

# La interpretación musical

Traducción
de Barbara Zitman

Alianza Editorial

Título original:
*Musical Performance. A Guide to Understanding*

Reservados todos los derechos. El contenido de esta obra está protegido por la Ley, que establece penas de prisión y/o multas, además de las correspondientes indemnizaciones por daños y perjuicios, para quienes reprodujeren, plagiaren, distribuyeren o comunicaren públicamente, en todo o en parte, una obra literaria, artística o científica, o su transformación, interpretación o ejecución artística fijada en cualquier tipo de soporte o comunicada a través de cualquier medio, sin la preceptiva autorización.

© Cambridge University Press, 2002
© de la traducción: Barbara Zitman, 2006
© Ed. cast. Alianza Editorial, S. A., Madrid, 2006
Calle Juan Ignacio Luca de Tena, 15; 28027 Madrid; teléf. 91 393 88 88
www.alianzaeditorial.es
ISBN: 84-206-6495-2
Depósito legal: M. 31.866-2006
Fotocomposición e impresión: EFCA, S. A.
Parque Industrial «Las Monjas»
28850 Torrejón de Ardoz (Madrid)
Printed in Spain

SI QUIERE RECIBIR INFORMACIÓN PERIÓDICA SOBRE LAS NOVEDADES DE ALIANZA EDITORIAL, ENVÍE UN CORREO ELECTRÓNICO A LA DIRECCIÓN:
alianzaeditorial@anaya.es

La interpretación musical

# *Índice*


Los autores ..... 9
Prefacio ..... 13

Primera parte: Conceptos y preconceptos

1. La interpretación a través de la historia, *por* COLIN LAWSON ..... 19
2. La interpretación histórica y el intérprete moderno, *por* PETER WALLS ..... 35
3. Análisis y (¿o?) interpretación, *por* JOHN RINK ..... 55
4. Comprender la psicología de la interpretación, *por* ERIC CLARKE ..... 81

Segunda parte: Aprender a interpretar

5. Sobre la enseñanza de la interpretación, *por* JANET RITTERMAN ..... 97
6. El desarrollo de la habilidad interpretativa, *por* JANE DAVIDSON ..... 111
7. Preparándose para interpretar, *por* STEFAN REID ..... 125
8. La memorización de la música, *por* AARON WILLIAMON ..... 137

Tercera parte: Hacer música

9. De la partitura al sonido, *por* PETER HILL ..... 155
10. El lenguaje del cuerpo durante la interpretación, *por* JANE DAVIDSON ..... 173
11. La interpretación en grupo, *por* ELAINE GOODMAN ..... 183
12. El miedo escénico, *por* ELIZABETH VALENTINE ..... 199

## Cuarta parte: Interpretar la interpretación

13. Escuchar la interpretación, *por* ERIC CLARKE ..... 217
14. El legado de las grabaciones, *por* PETER JOHNSON ..... 231
15. La crítica de la interpretación musical, *por* RAYMOND MONELLE ..... 249
16. Los intérpretes y la interpretación, *por* JONATHAN DUNSBY ..... 263

Índice analítico ..... 275

# *Los autores*

**Eric Clarke** es profesor de música de la Universidad de Sheffield y ha publicado numerosos trabajos sobre la psicología de la interpretación, la percepción y producción del ritmo, y el significado musical y la relación entre el lenguaje y la música. Es coeditor (junto con Nicholas Cook) del libro *Empirical Musicology: Aims, Methods, Prospects* y ha participado como improvisador en el Lapis String Quartet.

**Jane Davidson** es profesora adjunta de música de la Universidad de Sheffield. Ha publicado estudios sobre la expresión musical, los gestos, la práctica musical y los determinantes de la excelencia musical; es coeditora del libro *The Music Practitioner* y ha dirigido la revista *Psychology of Music.* Actúa con regularidad como solista vocal, ha dirigido y creado piezas de teatro musical y ha realizado la coreografía de diversas danzas y óperas.

**Jonathan Dunsby** es profesor de música de la Universidad de Reading. Ha escrito numerosas obras sobre la música de los últimos dos siglos, y entre sus libros se incluyen *Performing Music: Shared Concerns, Structural Ambiguity in Brahms, Schoenberg: Pierrot Lunaire* y (junto con Arnold Whittall) *Music Analysis in Theory and Practice.* Estudió piano con Fanny Waterman, ha ganado premios en diversos concursos internacionales y ha acompañado a muchos artistas distinguidos.

**Elaine Goodman** es profesora de música de la Universidad de Hull. Consumada violonchelista y pianista, toca con regularidad en diversas agrupaciones

de cámara y orquestas. Su tesis doctoral, presentada en el Royal Holloway College de la Universidad de Londres, trata sobre el ensayo y la interpretación en grupo, y sus campos de investigación incluyen la psicología de la interpretación, el análisis musical, la educación musical y los estudios sobre música popular.

**Peter Hill** es profesor de música de la Universidad de Sheffield. Entre las obras que ha publicado se encuentran *The Messiaen Companion* y un libro sobre *La consagración de la primavera* de Stravinsky. Como pianista ha grabado más de cien programas para la BBC y un galardonado ciclo de Messiaen. Sus discos recientes incluyen un recital de obras de Stravinsky y las obras completas para piano solo de Schoenberg, Berg y Webern.

**Peter Johnson** estudió en Cambridge y Oxford y ha publicado obras sobre Webern y diversos aspectos de la interpretación. Los objetivos de sus investigaciones recientes abarcan desde la filosofía de la interpretación hasta su análisis acústico, y ha trabajado ampliamente como director e intérprete de teclados. Es director de los departamentos de estudios de postgrado e investigación del Conservatorio de Birmingham.

**Colin Lawson** es conocido internacionalmente como intérprete de clarinete histórico y ha tocado con orquestas como The Hanover Band, el English Concert y la Orchestra of the Age of Enlightenment. Ha grabado como solista para muchas compañías discográficas y publicado numerosos escritos en la Cambridge University Press, incluyendo diversas ediciones y manuales. Es pro vicecanciller y rector del London College of Music & Media de la Thames Valley University.

**Raymond Monelle** ha ejercido la crítica musical durante treinta años y sus reseñas han aparecido en *The Independent, The Scotsman, Scotland on Sunday, Opera* y otras muchas publicaciones. Anteriormente era compositor, pianista y director. Su obra sobre teoría y semiótica musical es conocida internacionalmente y entre sus libros se incluyen *The Sense of Music,* Princeton University Press, 2000.

**Stefan Reid** es profesor de música de The Forest School de Londres. Estudió en la Universidad de Oxford, la Universidad de Surrey y el Royal Holloway College de la Universidad de Londres, en el que presentó su tesis doctoral sobre el arte de la práctica o estudio. Da conciertos regularmente como pianista y posee una distinguida reputación como director coral y orquestal.

**John Rink** es profesor de música del Royal Holloway College de la Universidad de Londres. Entre sus libros se incluyen *The Practice of Performance: Studies in Musical Interpretation, Chopin: The Piano Concertos* y (junto con Jim Samson) *Chopin Studies 2.* Es uno de los tres directores de la colección *The Complete Chopin —A New Critical Edition* y es coautor de un catálogo comentado de las primeras ediciones de Chopin junto con Christophe Grabowski. Como pianista, da conciertos y recitales-conferencia con regularidad.

**Janet Ritterman** ha sido directora del Royal College of Music desde 1993 y fue nombrada Dama de la Orden del Imperio Británico en 2002. Ha trabajado profesionalmente como pianista solista y acompañante, y se ha especializado en la educación y adiestramiento de intérpretes dotados. Sus temas de investigación incluyen el desarrollo del conservatorio, la pedagogía instrumental y la interpretación histórica. Ha publicado obras sobre compositores del siglo XIX y ha participado en numerosas conferencias nacionales e internacionales.

**Elizabeth Valentine** es profesora adjunta de psicología y directora del Psychology of Music Research Group del Royal Holloway College de la Universidad de Londres. Entre sus campos de investigación se incluye la psicología de la práctica y la interpretación musical, y ha escrito sobre temas como el efecto del canto sobre el estado de ánimo y la música de fondo como ayuda para la memoria autobiográfica en los casos de demencia. Es intérprete aficionada de teclados y flauta de pico.

**Peter Walls** es profesor de música de la Universidad Victoria de Wellington, Nueva Zelanda. Ha publicado numerosas obras sobre la práctica interpretativa, y su libro *Music in the English Courtly Masque 1604-1640* obtuvo el premio «Outstanding Academic Book» de la revista *Choice* y el «Derek Allen Prize» de la British Academy. Ha dado conciertos como violinista barroco y es director de orquesta y de ópera.

**Aaron Williamon** es investigador asociado y director del Centre for the Study of Music Performance del Royal College of Music. Su tesis doctoral, presentada en el Royal Holloway College de la Universidad de Londres, estudia cómo aprenden, ensayan y recuperan información los músicos expertos durante el estudio y la interpretación; su obra sobre la evaluación de las interpretaciones memorizadas recibió el premio Hickman de la SRPMME (Society for Research in Psychology of Music and Music Education). Otros objetivos de sus investigaciones son el miedo escénico y las técnicas de colaboración en el aprendizaje y el ensayo musical.

# *Prefacio*

La interpretación musical es una parte fundamental de la vida humana, pero hasta el más experimentado intérprete, maestro o erudito puede no apreciar lo que implica. Todos sabemos que una interpretación en público generalmente representa innumerables horas —incluso años— de aprendizaje y preparación, pero cómo se preparan las interpretaciones, sobre qué base y con qué resultados son aspectos menos conocidos. ¿Por qué algunas interpretaciones resultan «musicales»? ¿Deberíamos intentar respetar las intenciones del compositor? Y, de ser así, ¿cómo se pueden comprobar? ¿Cuál es la relación entre la partitura, la obra musical y las interpretaciones a las que éstas dan lugar? ¿Cómo se puede aumentar la eficacia de las prácticas y los ensayos, y cómo se puede superar el miedo escénico? Preguntas como éstas suelen rondar la mente de los intérpretes, por no mencionar la de sus maestros, pero hasta ahora ha sido difícil encontrar respuestas convincentes. Durante mucho tiempo, los músicos han tenido que recurrir a la tradición y la intuición para buscar las respuestas y, aunque aquéllas son indudablemente importantes, no siempre son suficientes. La creciente literatura académica sobre interpretación que se ha escrito en los últimos años ha ofrecido poco en términos de ayuda práctica: dirigida a lectores altamente especializados y escrita generalmente en un lenguaje un tanto impenetrable, ha tendido a descuidar los intereses de los intérpretes a pesar de que se necesitan obras claras e interesantes sobre asuntos como la práctica, la memorización, el miedo escénico, el análisis musical para la interpretación y las «responsabilidades» históricas del intérprete moderno. El hecho de que hoy en día sean cada vez más las universidades y conservatorios que ofrecen cursos en los que se fomenta la integra-

ción de la teoría y la práctica, en lugar de su tradicional separación, y que cada vez más los intérpretes profesionales se consideren a sí mismos tanto «hacedores» como «habladores» (términos de Joseph Kerman), hace que la ausencia de material adecuado sea aún más lamentable.

Los dieciséis ensayos de este libro tienen como objetivo desentrañar las complejidades de la interpretación y revelar los aspectos del aprender, ejecutar y responder a la música pertinentes a las interpretaciones en todos los niveles. Con un objetivo amplio y un lenguaje accesible, el libro está dirigido a los profesores, estudiantes y eruditos, así como a los amantes de la música, que deseen saber más sobre lo que implica la interpretación musical y no depender únicamente de la tradición y la intuición. Está dividido en cuatro partes, que corresponden a las etapas evolutivas de una interpretación: conceptos y preconceptos, aprender a interpretar, hacer música e interpretar la interpretación. Cada parte consta de cuatro capítulos, el primero de los cuales establece el ámbito de los otros tres. La Parte I comienza con una perspectiva general de la interpretación a través de la historia que sirve de introducción a la presentación de los conceptos básicos de la interpretación histórica, el análisis y la psicología de la interpretación. Los siguientes cuatro capítulos tratan sobre la historia de la enseñanza, el desarrollo de la habilidad interpretativa, la práctica o estudio y la memorización. La Parte III considera cómo se hace música partiendo de la partitura, así como los temas relacionados con la imagen personal, la interpretación en grupo y el miedo escénico. La última parte analiza el acto de escuchar, el legado de las grabaciones, la crítica musical y la opinión de los intérpretes sobre la interpretación.

Todos los autores de este libro son destacados expertos en el campo, y muchos son intérpretes de considerable experiencia y renombre. Ambos factores garantizan una obra vital y convincente, que al mismo tiempo plantea cuestionamientos provechosos. El libro pretende despertar un amplio interés, rehuyendo de las jergas y complejidades innecesarias y evitando concentrarse en un único instrumento o género interpretativo. En cambio, arroja una nueva luz sobre los principios y procedimientos que subyacen en la interpretación y su estudio, al menos dentro de la tradición clásica occidental, que constituye su principal centro de atención, aunque en el transcurso se hacen referencias a otras músicas y repertorios (incluyendo el jazz y el rock).

Existen varios hilos conductores comunes a lo largo del libro. El primero y principal es la importancia de la *experiencia* de la música a través de la interpretación, que transforma la partitura (en caso de haberla) en un evento musical único, reconociendo en todo momento que la partitura no es «la música» y que, aunque uno quiera, nunca se puede ser totalmente fiel a ella. (El libro también subraya que las grabaciones no reflejan «cómo suena la obra», a pesar de que pueden revelar mucho sobre las prácticas interpretativas

históricas y por supuesto sobre las visiones personales.) En varios capítulos se plantea, por consiguiente, la diferencia entre aprender «las notas» y aprender «la música» al desarrollar una visión de la obra —es decir, al construir un sentido del conjunto— que, como señala, es un primer paso vital. Varios autores tratan sobre el papel que desempeña el análisis en la interpretación y su aprendizaje, así como en la memorización y la lucha contra el miedo escénico. Se demuestran varias técnicas de análisis y se discute el papel de la intuición en contraposición a un conocimiento más profundo. De igual manera, al principio del libro se considera la respuesta del intérprete moderno al precedente histórico, particularmente los riesgos de la interpretación «históricamente *des*informada».

La naturaleza de la práctica o estudio —una actividad generalmente solitaria, totalmente opuesta a la interpretación como experiencia sumamente social— es otro tema común que permite considerar la relación entre las imágenes mentales y las representaciones físicas, la repetición en contraposición a la inspiración, la resolución de problemas y el proceso mediante el cual podemos «olvidarnos» de lo que hemos aprendido a medida que la música se hace más nuestra, para recordarlo «automáticamente» durante la interpretación. Varios autores analizan cómo memorizamos la música y ofrecen consejos prácticos para profesores e intérpretes por igual. Dejando atrás la fase preparatoria de la interpretación, los capítulos siguientes consideran los múltiples factores sociales que intervienen en la interpretación, como las relaciones entre los intérpretes y el público o entre los integrantes de un conjunto, así como los temas relacionados con la presentación, el papel que desempeña el cuerpo durante la interpretación y las semejanzas (y diferencias) entre la interpretación musical y la actuación. Dichas consideraciones están dirigidas a una amplia gama de intérpretes, incluyendo cantantes (por ejemplo, los capítulos 9, 11 y 14), pianistas (capítulos 2, 3, 7 y 9), ejecutantes de cuerdas (capítulos 8 y 14) y vientos (capítulo 14). La polémica que existe entre los intérpretes que intentan encontrar instrumentos históricamente «correctos» para tocar un repertorio determinado y los que desean «transcribir» música para tocarla en instrumentos modernos es abordada frontalmente en varios capítulos, al igual que la diferencia que existe entre las percepciones de los oyentes y los «hechos» que revelan los datos empíricos sobre las interpretaciones. Estas consideraciones forman parte de un debate más amplio sobre qué es lo que escuchan los oyentes y cómo lo escuchan, y sobre la manera en que los críticos y los intérpretes mismos han escrito sobre la interpretación —una actividad que para algunos ha sido tan apremiante como lo es para muchos intérpretes la necesidad de interpretar.

Como sugieren estos comentarios, el libro ofrece abundantes consejos prácticos, y la sección de «lecturas complementarias» al final de cada capítulo

ofrece una bibliografía concisa para aquellos que deseen estudiar más profundamente determinados temas. Esperamos sinceramente que este libro inspire a los lectores a investigar más sobre la interpretación, bien sea leyendo detenidamente otros textos generales sobre su estudio (por ejemplo, mi libro *The Practice of Performance: Studies in Musical Interpretation,* la obra de Jonathan Dunsby *Performing Music: Shared Concerns* o la de Colin Lawson y Robin Stowell *The Historical Performance of Music: An Introduction)*, bien ahondando en publicaciones más especializadas. Si realmente este libro ha despertado el interés por la interpretación y ha proporcionado respuestas para al menos algunas de las preguntas planteadas anteriormente, entonces habrá cumplido sin duda su propósito.

Quisiera agradecer a Sarah Smith por elaborar el índice y a Lucy Carolan por su cuidadosa edición de las copias; a Ruth Milsom y Mark Wells por preparar las ilustraciones de los capítulos 9 y 12; a mis alumnos de postgrado del Royal Holloway de la Universidad de Londres por su entusiasmo por el manuscrito, y a Penny Souster por su acostumbrada paciencia y sabios consejos.

JOHN RINK

# PRIMERA PARTE
## *Conceptos y preconceptos*

# 1 *La interpretación a través de la historia*

COLIN LAWSON

## *La interpretación en contexto*

En la mayoría de los conciertos clásicos de hoy en día esperamos que el público permanezca callado y absorto, pero ése es un fenómeno social bastante reciente, muy alejado de los eventos musicales anteriores a los comienzos del siglo XX. En el estreno de su Sinfonía «París», en julio de 1778, Mozart estaba encantado con un público respetuoso que, sin embargo, respondía activamente [1]:

> Yo estaba seguro de que iba a gustar un pasaje que se encuentra justo en medio del primer Allegro. El público se entusiasmó bastante, y hubo una tremenda ovación... Como había observado que aquí todos los primeros y últimos Allegros comienzan con todos los instrumentos a la vez, y generalmente al unísono, comencé el mío con dos violines solamente, tocando *piano* durante los primeros ocho compases, seguidos inmediatamente de un *forte*; el público, como yo esperaba, susurró «¡silencio!» durante el suave comienzo, y cuando escuchó el *forte*, comenzó a aplaudir inmediatamente [2].

Más de un siglo después, en 1893, tuvo lugar el concierto inaugural del Queen's Hall de Londres, representado en un cuadro que se encuentra actualmente en los archivos de la Royal Opera House del Covent Garden; el director y la orquesta aparecen en plena actividad; sin embargo, las conversaciones también fluyen libremente en las primeras filas del público [3]. Menos controlado fue el famoso y tumultuoso estreno de *La consagración de la primavera* de Stravinsky, en París en 1913, uno de los últimos ejemplos docu-

mentados de la respuesta activa del público dentro de la tradición concertística occidental.

Hoy en día, el público que responde libremente es más característico de los conciertos de música pop, en los que los intérpretes exhiben sus propias conductas cuidadosamente cultivadas. En las situaciones rituales que no requieren de música escrita hay menos diferenciación entre componer, ensayar e interpretar, y es muy probable que los oyentes y espectadores participen. Aunque la partitura constituye un registro y permite la difusión de la obra musical, también determina una relación más distante entre el compositor y el intérprete. Hasta bien entrado el siglo XIX, se daba por sentado que los intérpretes supieran improvisar, y en efecto lo hacían frecuentemente. Adaptar las partituras a las exigencias de la ocasión era algo rutinario, y también era bastante común insertar arias en una ópera que habían sido compuestas especialmente para los cantantes en cuestión. A principios del siglo XX, Stokowski, Elgar y otros no tenían por qué tener ningún reparo en hacer transcripciones orquestales de la música para teclado de Bach. Aunque la notación original del compositor se encontrara sobre el atril, los intérpretes adoptaban con frecuencia un enfoque libre y creativo. Joseph Joachim, un violinista amigo de Brahms, fue descrito como impredecible por un miembro de su cuarteto: «Tocar con él es sumamente difícil. Siempre con tempos diferentes, acentos diferentes» [4]. La abrumadora supremacía actual de la partitura, que exige fidelidad y exactitud a toda costa, no es de ningún modo característica de la historia de la interpretación en su conjunto. Sin embargo, la literatura musical suele dar la impresión de que el verdadero significado estético reside en la notación, y que la interpretación es, como mucho, una representación imperfecta y aproximada de la obra misma. Se dice que Brahms rechazó una invitación a una representación de *Don Giovanni* de Mozart, alegando que prefería quedarse en casa y leerla, una respuesta que seguramente habría horrorizado a Mozart, que era extremadamente práctico.

### *Una perspectiva mundial*

El estudio de las músicas del mundo pone en duda la mayoría de las hipótesis apoyadas por los intérpretes occidentales, proporcionando por lo tanto un provechoso sentido del contexto. En muchas culturas, la separación convencional entre el intérprete y el público nunca ha existido; mientras el instrumentista o cantante improvisa, los espectadores responden, bien sea dando golpecitos con la punta de los pies o los dedos de las manos, bien dando palmadas, bien cantando o bailando. Los músicos se encuentran tanto en las calles, mercados, ferias o tabernas como en lugares más formales, ya que la mú-

sica se ha mantenido estrechamente relacionada con los rituales como bodas, funerales o eventos del calendario agrícola, y trata sobre temas eternos como las heridas de amor. Los testimonios de la tradición oral se complementan a veces con un material básico más tangible; por ejemplo, en países como China, Corea, India y Japón, existen diversas tradiciones de teoría musical que datan de hasta 5.000 años atrás. Se han conservado grandes colecciones de música china escrita a partir del siglo XII, y muy poco después comenzó a gestarse el repertorio clásico turco.

La invención del fonógrafo, a finales del siglo XIX, contribuyó en gran medida a la investigación de las tradiciones orales y los sistemas musicales existentes. Otro hito fue el desarrollo de los métodos para medir intervalos menores que el semitono, con los que se ha dividido la octava de múltiples maneras desconocidas en Occidente. El análisis técnico de los sonidos musicales y las prácticas interpretativas se ha unido provechosamente al estudio antropológico, de modo que el trabajo de campo abarca generalmente tanto el proceso de creación como el de interpretación. Los instrumentos, melodías y ritmos de la India son desconocidos para los músicos occidentales, y también las ideas que originan su música están estrechamente relacionadas con conceptos filosóficos y religiosos que tienen poco que ver con los enfoques occidentales del tiempo, la materia o la realidad. En esas circunstancias, la música no se aprende de los libros: hay que tener un maestro que proporcione el secreto esencial y el conocimiento esotérico, y enseñe cómo se debe llevar la vida de músico. Desde el principio, el discípulo participa en el escenario observando e interactuando con su maestro. En la cultura hindú, los músicos pertenecen a una casta inferior, no sólo por su condición de comerciantes, sino porque su profesión implica romper tabúes de las clases altas, como tocar la piel de los tambores o tocar con los labios una flauta o una caña que puede haber sido tocada por otras personas. La complejidad rítmica de la música africana y su estrecha relación con la danza están igualmente alejadas de la experiencia occidental. Muchas lenguas africanas ni siquiera tienen una palabra que denote al músico o la música, abstracciones inútiles comparadas con sus términos concretos para significar cantante, bailarín o ejecutante de tambores. En el Extremo Oriente, las ceremonias religiosas y los rituales de la corte imperial de Japón están muy alejados filosóficamente de Occidente. En el elevado y estilizado arte dramático Noh se integran literatura, teatro y danza con una música en la que se combinan tambores con el sonido espiritual del shakuhachi.

El mundo sonoro de la música no occidental comprende muchos instrumentos de diseños y categorías poco conocidas. Un ejemplo antiguo es el qin chino, que se caracteriza por sus expresivos glissandi y etéreas armonías. El gamelán indonés incluye un conjunto de percusión afinada cuyos ejecutantes

participan en una disciplina espiritual vinculada a la danza, la poesía y el teatro cuyo objetivo es alcanzar un estado ideal de calmada ausencia emocional; naturalmente, no hay virtuosos ni solistas en esa tradición, básicamente oral, en la que no existen jerarquías. El sitar indio, con sus característicos glissandi, es un ejemplo de un instrumento que (como el gamelán) se ha dado a conocer en Occidente, especialmente a través de su interacción con la cultura popular.

### *Grecia y Roma*

Hace cuarenta mil años, el hombre hacía música; las pinturas paleolíticas descubiertas en la localidad francesa de Ariège muestran la imagen de un hombre con una máscara de animal rascando un arco musical ante un público formado por renos. Para la época clásica, la música despertaba el interés de los grandes filósofos, al tiempo que invitaba a una amplia participación ritual. Escritores como Platón y Aristóteles consideraban que la formación musical era de gran valor educativo. La relación entre música y poesía exigía respuestas profundas al intérprete y al oyente que hoy en día no se pueden recrear mediante una mera lectura de las fuentes que se han conservado. El canto coral era particularmente importante, con frecuencia incorporaba danzas y siempre estaba acompañado por profesionales que tocaban instrumentos de viento y de cuerdas. El ritual formaba parte esencial de la interpretación, ya que los coros cantaban en honor a los dioses o celebrando a hombres famosos o atletas victoriosos. Aunque es poca la música que ha llegado hasta nosotros, se han conservado tratados teóricos de Aristoxeno (siglo IV a.C.), Ptolomeo (siglo II a.C.) y otros. Podemos conjeturar que la música griega era fundamentalmente homofónica y que las voces cantaban juntas al unísono o a la octava. Los romanos también tenían música litúrgica y otras músicas de carácter público, música militar y canciones de trabajo, asimilando las influencias de los países que conquistaban. Los intérpretes de instrumentos de viento participaban en los rituales de sacrificio, en parte para alejar a los espíritus malignos e invocar a las deidades benignas; fueron representados con frecuencia en relieves y eran muy respetados. Las festividades en honor a los dioses se acompañaban de música procesional, que en el caso de Dioniso podía ser desenfrenada y orgiástica. En la música militar, los trompetistas ejecutaban señales establecidas y tocaban durante las marchas y ceremoniales; en combate, sus sonidos estaban ideados para alentar a las tropas y confundir al enemigo. Las canciones tradicionales proporcionaban un acompañamiento rítmico para ciertas actividades como remar, cosechar y tejer, y se agrupaban en géneros como nanas, canciones infantiles, canciones de cumpleaños y de

bodas. En el teatro cobraron importancia la mímica y la pantomima, un género en el que la música vocal e instrumental acompañaba a un mimo que representaba personajes mitológicos.

## *La iglesia*

Durante siglos, la iglesia ha sido un núcleo importante para el quehacer musical. Los romanos relacionaron la música con todo tipo de vicios e inmoralidades, por lo que las primeras autoridades cristianas se percataron de que podía tanto ennoblecer como degradar la moral del hombre. Una solución inicial al dilema fue ordenar que la música estuviera relacionada con palabras devotas y prohibir los instrumentos. La educación musical ha sido una función importante de la iglesia desde la fundación de la Schola Cantorum, atribuida tradicionalmente a Gregorio Magno (590-604). Allí los niños recibían una formación musical, así como una educación completa en las principales asignaturas académicas, participando igualmente en las festividades profanas y ferias. Esta prolongada tradición no se debe subestimar; cuando se fundó el primer conservatorio moderno en París, en 1795, después de la revolución, la educación musical francesa seguía estando en manos de casi cuatrocientas escuelas de música patrocinadas por la iglesia, en las que se enseñaba canto llano, contrapunto, algo de composición, un poco de francés, mucho latín y algo de aritmética. Aunque los métodos de enseñanza eran anticuados y se descuidaba la música instrumental, cada año unos cuatro mil alumnos iniciaban el seminario teológico o una vida como cantantes u organistas.

Las melodías eclesiásticas no fueron escritas hasta el siglo VI, y sólo en el siglo XII se ideó un sistema para indicar las alturas de los sonidos, aunque no su valor exacto. Una de las principales dificultades de recrear música medieval es que la improvisación y el acompañamiento instrumental no están representados en las notaciones que se han conservado. Los órganos pequeños, las cuerdas frotadas y pulsadas, los vientos y los instrumentos de percusión fueron introducidos en la iglesia, pero la plantilla era seleccionada rigurosamente según la ocasión y los recursos disponibles. Durante varios siglos, la complejidad de la música eclesiástica suscitó reiteradas críticas por parte de las autoridades, por ejemplo del papa Juan XXII en 1323. El Viernes Santo de 1555, el papa Marcelo II ordenó que en lo sucesivo el coro papal interpretara con voces correctamente moduladas y de manera que todo se pudiera escuchar y entender. De tiempo en tiempo, algunos compositores de música eclesiástica trascendieron completamente sus circunstancias profesionales, como es el caso de J. S. Bach. La Alemania luterana produjo muchos directores musicales que componían conscientemente para la mayor gloria de Dios,

satisfaciendo a sus superiores y edificando a la congregación. Bach vivió en una época en que los conceptos de genio y obra maestra aún no se habían formulado; trabajando de semana en semana con una pequeña orquesta de dieciocho a veinticuatro instrumentistas para acompañar a sus cantantes, sin duda le habría asombrado que las generaciones posteriores lo veneraran.

### *Música profana*

Mientras que la música eclesiástica era organizada y coordinada, la música profana se abordaba sin considerar en absoluto que su método u objetivo debía ponerse por escrito. El centro de gravedad de la música pasó de la iglesia a las cortes reales y aristocráticas sólo a comienzos de la época renacentista. Originalmente, la música formaba parte de la tarea de los animadores de la corte, que eran también juglares, acróbatas, bailarines, adiestradores de animales y cómicos. Estos ministriles formaban parte del personal de una casa aristocrática o bien de una compañía itinerante. Acompañados de instrumentos portátiles, no sólo entretenían sino también inmortalizaban gloriosas gestas reales en canciones, al tiempo que versaban sobre los temas habituales del amor, el duelo y la sátira. A medida que se desarrollaban las ciudades en la Edad Media, se formaron asociaciones de trompeteros o *Stadtpfeifer,* originalmente como guardias civiles. Éstas evolucionaron hasta convertirse en bandas locales que tocaban diversos instrumentos y formaban a los aprendices, que trabajaban como sirvientes hasta obtener las calificaciones necesarias. A menudo luchaban por mantener el monopolio de la música ceremonial y social en una ciudad determinada. Los remanentes de este sistema perduraron hasta entrado el siglo XIX, cuando en Leipzig los *Stadtmusiker* llegaron a ser miembros *ex officio* de la Orquesta de la Gewandhaus.

Las cortes renacentistas siguieron un modelo de desarrollo uniforme, con pequeños coros y un grupo de instrumentos para acompañar y hacer música instrumental. Crearon agrupaciones bien instruidas y flexibles para interpretar obras compuestas generalmente en los estilos más sorprendentes y avanzados de la época. El primer libro sobre instrumentos —*Musica getutscht*— fue publicado en Basilea en 1511 por Sebastian Virdung. En él se describen e ilustran mediante grabados el clavicordio, el virginal, el laúd, la viola, la dulcema, el arpa, la chirimía, la flauta, la corneta, la gaita, el trombón, la trompeta y varios instrumentos de percusión y órganos. La impresión de música comenzó a finales del siglo XV, comprometiendo al editor con la noción de que ciertas obras tendrían una larga existencia y creando la idea de un repertorio musical. Un aspecto importante de la clase media isabelina y jacobina fue el desarrollo del placer íntimo de la música de cámara vocal e instrumental.

## *La época barroca*

Durante el período barroco se apreciaba especialmente la espontaneidad en la interpretación. El auge de la música instrumental y la ópera en el siglo XVII produjo un nuevo tipo de músico que hacía especial hincapié en la técnica, demostrando habilidad para improvisar, ornamentar o crear el efecto deseado. Es significativo que las intérpretes femeninas fueron cada vez más aclamadas durante el barroco, no sólo en los escenarios operísticos sino también en conservatorios venecianos como el Ospedale della Pietà, que se hizo famoso por Vivaldi.

Los teatros de ópera fueron importantes por ser los primeros establecimientos independientes en los que se interpretaba música para públicos de diversas clases sociales que pagaban una entrada; el primero abrió sus puertas en Venecia en 1637. Eran relativamente pequeños, siendo el más grande de un tamaño similar al de una sala de cine pequeña de hoy en día. La nobleza alquilaba gran parte de los palcos por un año (a veces de por vida), y los asientos restantes estaban disponibles para los que adquirían una entrada única. El vocabulario sistemático de signos y gestos musicales que desarrollaron Monteverdi y sus contemporáneos llegó a representar las relaciones entre los personajes en escena, ya que a los cantantes se les exigió convertirse en actores. No podemos asegurar que siempre se usaran los mismos recursos para una misma ópera; de hecho, tanto las piezas como la plantilla podían variar en representaciones sucesivas. Entre los tipos de voz actualmente obsoletos, el más importante era el castrato, valorado por su fuerza y flexibilidad. Había muchos aspectos del buen estilo de canto que se escribían de una manera pero que, para ser más elegantes, se realizaban de otra bastante diferente. Desde sus comienzos, la ópera ha generado esa curiosa interacción entre consideraciones comerciales y artísticas que sigue existiendo hoy en día, siendo la interesante carrera de Haendel en Londres un excelente ejemplo del siglo XVIII.

Al igual que la ópera, la música de cámara era dirigida generalmente desde el clave, a menudo por el compositor; las instrucciones interpretativas detalladas eran por lo tanto innecesarias. Los ejecutantes seguían asumiendo un elevado nivel de responsabilidad en una época en la que gran parte del material musical estaba esbozado más que escrito en toda su extensión, especialmente las partes de *basso continuo.* En este aspecto se pueden establecer comparaciones significativas con la técnica del jazz, especialmente en el caso de las *notes inégales* —una sucesión de notas del mismo valor que se tocan desigualmente en el estilo francés según el tempo y el carácter de la música. No fue hasta alrededor del año 1700 cuando la música instrumental alcanzó el prestigio de la música vocal, pues hasta entonces había estado en gran parte confinada a las oberturas de ópera y el servicio del culto de la iglesia. Un fac-

tor importante a este respecto fue que la agrupación dirigida por Lully en la corte de Luis XIV había sido especialmente famosa por su disciplina y unanimidad, y estableció los instrumentos de cuerda como una base para la orquesta. Poco después surgió en Italia un nuevo tipo de escritura virtuosa para cuerdas, ejemplificada por el violinista-compositor romano Arcangelo Corelli.

### *El período clásico*

Durante la segunda mitad del siglo XVIII, la música llegó a difundirse regularmente a través de las ediciones publicadas. Incluso los textos fiables sólo ofrecen una idea parcial de lo que se interpretaba originalmente, ya que los instrumentistas y cantantes seguían aportando una parte importante. Sin embargo, los principales tratados de Quantz, Leopold Mozart y C. P. E. Bach abarcan muchos aspectos de la práctica contemporánea. En ellos se establece un paralelismo cercano entre la música y el lenguaje, y se valora especialmente la habilidad del intérprete para emocionar al público. Las condiciones sociales favorecieron el auge del concierto público, que al principio floreció en lugares pequeños, fomentando un estilo interpretativo claro e íntimo. En las reseñas, las descripciones del nivel de calidad de los conciertos son sumamente difíciles de interpretar. Un ejemplo famoso es la descripción que hace Charles Burney de la orquesta de Mannheim como un ejército de generales, tan capaz de planificar una batalla como de pelear en ella; luego matizó su entusiasmo con una crítica a la afinación de los vientos, cuya aspereza era un problema orquestal universal [5]. Basándose en razones que ahora consideraríamos mezquinas, una reseña notable de 1785-1786 de la orquesta de Lyon señala que el director no tenía inteligencia, ni tampoco un estilo interpretativo preciso, y que había ausencias no autorizadas entre sus colegas. La mayoría de los conciertos (a diferencia de la ópera) generalmente podía contar con un solo ensayo, y a veces no había ninguno. En los programas se valoraban especialmente la variedad y la novedad, y se combinaba música solista, de cámara y orquestal. Por ejemplo, el 25 de noviembre de 1781, el concierto inaugural de la Orquesta de la Gewandhaus de Leipzig incluía una sinfonía de Joseph Schmitt, un himno de Reichardt, el Concierto para violín de Berger, un cuarteto de un autor desconocido actualmente, una sinfonía de J. C. Bach, un aria de Sacchini y una sinfonía de E. W. Wolf.

En su obra *A General History of Music,* Burney describe la sed de música nueva que había en su siglo:

> Tan variable es el gusto en la Música, y tan pasajero el privilegio de cualquier estilo, que su historia es como la de un campo labrado: tal año produjo trigo, tal otro ce-

> bada, guisantes o tréboles: y tal año estuvo sin cultivar. Pero no queda ninguna de sus producciones, salvo quizá una pequeña parte de la cosecha del año pasado y las semillas y malas hierbas que ahora cubren su superficie [6].

Una excepción importante a esta regla general fue, por supuesto, la popularidad que tuvo Haendel en Inglaterra hasta mucho después de su muerte. Para finales del siglo, empresarios como Salomon, quien contrató a Haydn para sus conciertos en Londres después de su prolongado servicio en la corte de Esterházy, hicieron que el éxito fuese juzgado no sólo por la aprobación estética, sino también en términos comerciales. Mozart representaba un nuevo tipo de músico autónomo, presentándose astutamente al público vienés de los años 1780 con sus conciertos para piano en los que desempeñaba el papel tanto de compositor como de intérprete. También escribió papeles operísticos y obras solistas ajustadas específicamente a las habilidades de sus colegas intérpretes. Una generación después, Beethoven plantearía enormes retos a los ejecutantes de orquesta, que tenían que hacer frente a exigencias técnicas y estilísticas nuevas y difíciles en una época en que las condiciones para el ensayo y la interpretación eran desfavorables debido a factores sociales, políticos y también musicales.

Afortunadamente, se ha conservado un documento que retrata la formación musical austríaca del siglo XVIII en forma de un «Musick Plan», un sistema educativo proyectado en 1799 por Anton Stadler, clarinetista de Mozart, en respuesta a un encargo del conde húngaro Georg Festetics. Stadler recomendó una carrera de seis años en la que los alumnos estudiarían aspectos de teoría, interpretación y composición, incluyendo piano, órgano o bajo cifrado, violín e instrumentos de viento, para complementar sus estudios principales. Todos los alumnos de música estudiarían el arte del canto, cualquiera que fuese la calidad de sus voces. Haciendo hincapié en la importancia de una buena educación general, señaló que cualquiera que desee comprender la música ha de tener mucho mundo y saber de matemáticas, poesía, oratoria, arte y muchos idiomas [7].

## *El siglo XIX*

La era industrial trajo grandes cambios, especialmente las posibilidades de transporte que permitieron el florecimiento de las carreras de virtuosos e influyeron igualmente en la vida de los músicos de orquesta. A medida que las partituras se difundían más ampliamente, las indicaciones interpretativas tendían a ser aún más precisas. En la música alemana, la complejidad de los acompañamientos restringió gradualmente la libertad que tenían los cantan-

tes para ornamentar; un compositor advirtió que «las composiciones de Mozart, Haydn, Cherubini y Winter presentan menos adornos que las de Salieri, Cimarosa, Martin y Paisiello» [8]. El establecimiento de conservatorios en toda Europa estaba destinado a fomentar el virtuosismo técnico, especialmente porque es un aspecto de la interpretación que se puede evaluar sin dificultad. Naturalmente, el aficionado comenzó a desaparecer del mundo del concierto público. El piano mantuvo un puesto central tanto en el ambiente doméstico (había arreglos para piano a cuatro manos de casi toda la música orquestal) como en el escenario de concierto, donde Liszt era representativo de un nuevo tipo de virtuoso itinerante. Las innovaciones técnicas de Liszt permitieron una extraordinaria transcripción al piano de las expresiones orquestales, y su destreza virtuosística estaba inspirada en el violinista Niccolò Paganini, una de las figuras más hipnóticas del siglo XIX.

El desarrollo de las orquestas sinfónicas trajo consigo los conciertos públicos financiados por suscripciones y estableció el papel del director. Beethoven fue uno de los primeros músicos que dirigió un concierto en Viena sin tocar un instrumento, desde un atril separado; se dice que estaba especialmente interesado en transmitir la expresión de la música, y no tan sólo un pulso regular. El tamaño de las orquestas era muy variado; más avanzado el siglo, las experiencias de Brahms abarcaban desde su agrupación favorita de Meiningen, de cuarenta y ocho músicos, hasta la Gewandhaus de Leipzig, con una plantilla de noventa y ocho intérpretes. Para su época, el escuchar contemplativa e imparcialmente se había convertido en el objetivo mismo de la interpretación; las revistas y periódicos reflejan bastante bien cómo eran acogidos los conciertos, ya que la crítica profesional se estableció gracias a escritores como Eduard Hanslick y George Bernard Shaw. En Inglaterra, las sociedades corales tendían a relacionar la buena música con la religión, más que con la calidad musical, y el editor Vincent Novello les suministraba enormes cantidades de partituras vocales baratas; dichas asociaciones continúan brindando a los músicos aficionados la oportunidad, actualmente escasa, de aparecer en público. El movimiento de las bandas también formaba parte de una iniciativa de autoayuda y de superación moral y material.

La recuperación de la música antigua comenzó en serio, siendo ejemplos importantes la interpretación que hiciera Mendelssohn en 1829 de la *Pasión según San Mateo* de Bach y el repertorio coral dirigido por Brahms, que incluía obras de Morley, Schütz, Palestrina, Bach y Haendel. Mientras tanto, las ediciones de las obras completas de Bach, Haendel, Mozart y otros establecieron la noción del texto musical definitivo. Sin embargo, aún existía la creencia generalizada de que la música antigua tenía que ser actualizada en la interpretación; en la primera edición del *Grove's Dictionary,* el artículo de Ebenezer Prout sobre «acompañamientos adicionales» sostiene que la inter-

pretación literal de la música de Bach y Haendel no refleja las intenciones del compositor[9]. A la inversa, cada vez más músicos consideraban que los estilos interpretativos contemporáneos no se adecuaban necesariamente a la música de épocas anteriores. Músicos como Arnold Dolmetsch fueron pioneros de la concienciación histórica que tuvo gran aceptación durante el siglo XX. Sus actividades fueron oportunas, ya que la gran influencia de la dirección orquestal de Wagner había hecho que la música de las generaciones anteriores sonara más como la suya propia, reemplazando los elementos verbales de la interpretación por elementos pictóricos.

## *El siglo XX*

El desarrollo de las comunicaciones en el siglo XX permitió a los músicos degustar y asimilar interpretaciones en todo el mundo. Las posibilidades de grabación cambiaron fundamentalmente el carácter de la ejecución musical, y el estudio de grabación originó una nueva veneración por la precisión técnica que a su vez se introdujo en la sala de concierto. La industria de la grabación ha influido enormemente en la comercialización de diversos tipos de músico, tanto en el campo de la música popular como en el de la música clásica.

En el siglo XIX, la música popular aún coincidía parcialmente con el concierto, el teatro y la música doméstica. Después de la Revolución Industrial, el público iba cada vez más a los parques, jardines públicos, salones de baile y teatros de variedades en busca de música, mientras que las melodías más recientes también se escuchaban en los organillos de la calle. Gradualmente, la música seria fue alejada del alcance del público general, ya que la opereta y el teatro de variedades suplían la brecha de accesibilidad que habían creado Verdi y Wagner. Pronto llegaron de América los bailes populares y las marchas de Sousa, y después el ragtime y el foxtrot constituyeron un verdadero descubrimiento. Al principio, el jazz combinaba diversos elementos musicales de África, Europa y América, y encontró rápidamente un lugar en los barcos fluviales y furgonetas publicitarias, así como en los desfiles, funerales, salones de baile y cabarés. La improvisación era una característica esencial que combinaba habilidades individuales y colectivas y se valía de diversos estilos incluyendo la marcha, el rag, el himno y el blues. Los intérpretes de jazz dejaron de verse a sí mismos como meros animadores y empezaron a considerarse cada vez más los defensores profesionales de un arte sutil y complejo.

En el período de entreguerras, Paul Whiteman ofreció una imagen sumamente accesible del jazz, y las comedias americanas más importantes de Kern, Gershwin, Rodgers y Cole Porter alcanzaron una enorme popularidad.

Las barreras nacionales tendieron a desaparecer después de aproximadamente 1940, cuando la música podía ser trasladada y grabada con facilidad y, por lo tanto, formar parte de un complejo mundo comercial y financiero. Entre los estilos importantes se encontraba el swing, con su novedosa instrumentación, su sonido vibrante y su potente impulso rítmico; después, a mediados de la década de 1950, el rock and roll, con una diferencia esencial de sonido, instrumentación e intención expresiva; más tarde, el rock de los años 1960 y 1970, con una variedad y libertad formal y tonal estrechamente vinculadas a los sentimientos políticos y el radicalismo social. Sus textos abarcaban desde lo político hasta las experiencias intensamente personales, combinados con una creciente amplificación y colección de artilugios electrónicos. La importancia artística de la música popular seguía ocupando a la crítica seria, y en 1967 Kenneth Tynan afirmó que la publicación del álbum *Sgt. Pepper's Lonely Hearts Club Band* de los Beatles constituía «un momento decisivo en la historia de la civilización occidental» [10].

En las salas de concierto, los intérpretes debían adaptarse a situaciones nuevas y hacer frente a dificultades técnicas cada vez mayores. Las agrupaciones pequeñas de diversos tipos desafiaban la supremacía de la gran orquesta sinfónica. A principios del siglo XX, obras como *La consagración de la primavera* y *La historia del soldado* de Stravinsky señalaron una mayor complejidad en las exigencias rítmicas. En esa época, uno de los grandes logros educativos del suizo Émile Jacques-Dalcroze fue enseñar a los jóvenes a pensar y sentir diferentes patrones rítmicos simultáneamente. A la mayoría de los músicos de orquesta de 1910 les resultaba muy difícil ejecutar grupos de cinco notas en un compás de 3/4 o 4/4; en cambio, para los que se habían formado con el método Dalcroze eso era un juego de niños.

Además de las influencias del jazz y el ragtime, los cantantes e instrumentistas debían responder a los requerimientos de la composición serial, y a veces se les exigía aventurarse en terrenos aún más desconocidos, como los cuartos de tono. A la mayoría de los instrumentos de orquesta se les exigía abarcar ámbitos extendidos y técnicas nuevas como la producción de acordes en instrumentos de viento-madera. Precisar el límite que separa a los sonidos musicales de los ruidos se convirtió en el objetivo de innumerables experimentos, como las obras de Varèse y el *Ballet mécanique* de George Antheil, escrito en 1924 para diez pianos e instrumentos y máquinas eléctricas, incluyendo hélices de avión. Los recursos electrónicos desarrollados a lo largo del siglo XX amenazaron con usurpar del todo el papel del intérprete. En 1947, Stravinsky condenó la interpretación y exigió un enfoque estrictamente objetivo por parte del intérprete. Sin embargo, la notación musical se extendió incorporando elementos gráficos no convencionales para inspirar la improvisación, mientras que los principios aleatorios (o de probabilidad) también

constituían una fuente de experimentación. La indeterminación surge cuando al intérprete se le requiere elegir el orden del material o se le guía mediante algún proceso hacia un orden evidentemente aleatorio o fortuito. Influido por el budismo zen, John Cage ofreció tales opciones a sus intérpretes en su obra *Variations IV* (1963), «para cualquier número de intérpretes, cualquier sonido o combinaciones de sonidos producidos por cualquier medio, con o sin otras actividades».

En los años subsiguientes, los intérpretes se han beneficiado de las oportunidades que ofrecen la radio, la televisión, los CD y los CD-ROM; parece inevitable que éstos serán complementados por el internet, actualmente preparado para ofrecer nuevas posibilidades para la difusión de interpretaciones individuales. Mientras tanto, la gestión orquestal ha considerado los programas educativos como una manera de llegar a la comunidad y tender puentes entre la vida diaria y la sala de conciertos, reubicando al público como el centro de la actividad musical. Como señaló recientemente el director general de una orquesta londinense: «¿Quizá un siglo de rigurosa etiqueta musical ha reemplazado las emociones, el placer y la comunicación por el mundo académico, el orden social y el aislamiento?» [11].

## *Aproximaciones a la música de otras épocas*

La conciencia histórica estaba suficientemente desarrollada en 1952 como para que Hindemith afirmara que toda música ha de ser interpretada con los medios de producción que se utilizaban cuando el compositor la ofreció a sus contemporáneos. Pero también se percató de las limitaciones de esa aproximación: «Nuestra forma de ser no es idéntica a la de nuestros antepasados, y por lo tanto su música, aunque la recreemos con absoluta perfección técnica, nunca tendrá para nosotros exactamente el mismo significado que tenía para ellos» [12]. En respuesta a la aceptación generalizada de la interpretación histórica, algunos se han preguntado cómo se puede relacionar el conocimiento histórico con la pasión y la convicción personal de un intérprete. Paradójicamente, este aspecto siempre ha sido una importante fuente de inspiración para los compositores.

Richard Taruskin ha sostenido que, de hecho, la responsabilidad absoluta reduce la práctica interpretativa a una lotería, ya que el intérprete no puede ejercer ningún control sobre el estado de la evidencia. Además, la necesidad de satisfacer las intenciones de un compositor (suponiendo que ello fuera posible) demuestra una falta de valentía, por no decir una dependencia infantil [13]. El autor cita a Stravinsky como ejemplo positivo de un compositor que ha cambiado de opinión sobre diversos asuntos interpretativos a lo largo de

su carrera. Taruskin considera que la interpretación histórica se ha ganado su posición actual por su novedad más que por su antigüedad; también considera que es un fenómeno característico de nuestra época. Sin embargo, el estímulo de la historia puede avivar la imaginación musical, lo cual puede ser tan importante como intentar recuperar cada detalle. Recientemente, Margaret Bent señaló que «cada repertorio de música antigua tiene su propia gramática y su propio dialecto. Como ocurre con las lenguas "muertas", esas gramáticas son recuperables, permitiéndonos entender la literatura escrita y saber bastante sobre cómo se realizaban o pronunciaban, incluso sin que sepamos cómo sonaban exactamente» [14].

El hecho de que la articulación separada fuese la norma en la época de Mozart constituye una evidencia histórica tan fundamental que difícilmente podemos ignorarla. El conocimiento también puede complementar el gusto propio de un intérprete en cuanto a la interpretación de ligaduras cortas o la formulación adecuada de una disonancia y su resolución. Es necesario tener en cuenta que antes de 1900, el vibrato en las cuerdas se utilizaba como ornamento y no como un recurso extendido. En efecto, la conciencia histórica puede iluminar la interpretación de múltiples maneras. Recientemente ha habido una sana interacción entre las prácticas moderna e histórica, ya que los principios históricos han comenzado a influir gradualmente en la tendencia predominante del mundo musical. Mientras tanto, las afirmaciones de autenticidad o incluso exactitud histórica («el Beethoven más original que se haya grabado jamás») [15] se han atenuado cada vez más. Los estilos interpretativos apropiados para diferentes épocas serán seguramente uno de los mayores legados del siglo XX, y ciertamente se ha abierto una relación más profunda con la notación. Sin embargo, una obsesión con el pasado ¿no refleja también una terrible brecha entre los compositores modernos y su público? Resulta difícil no coincidir con el gran musicólogo Donald Jay Grout, quien ya en 1957 comentó que si un compositor de música antigua

> pudiera por algún milagro ser devuelto a la vida en el siglo XX para interrogarle sobre los métodos de interpretación de su época, su primera reacción sería indudablemente de asombro por nuestro interés acerca de semejantes asuntos. ¿Acaso no tenemos una tradición musical viva, y debemos intentar revivir una que está muerta? La pregunta puede resultar embarazosa. El arcaísmo musical puede ser síntoma de una civilización en decadencia [16].

Durante el siglo XX, el lugar de la música contemporánea fue, en efecto, sustituido por interminables repeticiones de un repertorio establecido y conocido que entonces invitaba a comparar detalladamente las pequeñas diferencias interpretativas. Los retos del nuevo milenio son, por lo tanto, bastante dis-

tintos de las experiencias descritas en este breve análisis de la interpretación a través de la historia.

## *Notas*

1. Véase Christopher Small, *Musicking,* Hanover, NH, y Londres, Wesleyan University Press, 1998, para un análisis estimulante de las relaciones entre el público y los intérpretes.
2. Carta dirigida el 3 de julio de 1778 en París a Leopold Mozart, Salzburgo; Wolfgang Amadeus Mozart, *The Letters of Mozart and His Family,* 3ª ed., Emily Anderson (trad. y ed.), Stanley Sadie y Fiona Smart (rev.), Londres, Macmillan, 1985, p. 558.
3. Reproducido en Michael Forsyth, *Buildings for Music,* Cambridge, Mass., MIT Press, 1985, p. 229.
4. Julius Levin en *Die Musik* (1926), citado por Joseph Szigeti, *Szigeti on the Violin,* Londres, Cassell, 1969, p. 176.
5. Charles Burney, *The Present State of Music in Germany, the Netherlands, and United Provinces* [1773], Nueva York, Broude Brothers, 1969, pp. 95-97.
6. Charles Burney, *A General History of Music,* 4 vols., Londres, Charles Burney, 1776-1789, vol. II, p. 380.
7. Véase Pamela Poulin, «A view of eighteenth-century musical life and training: Anton Stadler's "Musick Plan"», *Music & Letters* 71 (1990), pp. 215-224.
8. J. F. Schubert, *Neue Singe-schule,* Leipzig, Breitkopf & Härtel, 1804, p. 138.
9. Ebenezer Prout, «Additional accompaniments», en George Grove (ed.), *A Dictionary of Music and Musicians,* 4 vols., Londres, Macmillan, 1879-1889, vol. I, p. 30.
10. Citado por Allan F. Moore, *The Beatles: Sgt. Pepper's Lonely Hearts Club Band,* Cambridge, Cambridge University Press, 1997, p. 59.
11. Serge Dorny, «Vital Signs», *BBC Music Magazine* (octubre 1999), pp. 53-54.
12. Paul Hindemith, *A Composer's World,* Cambridge, Mass., Harvard University Press, 1952, pp. 170-171.
13. Richard Taruskin, *Text and Act: Essays on Music and Performance,* Nueva York y Oxford, Oxford University Press, 1995, p. 98.
14. Margaret Bent, «"Authentic" listening?», *Early Music* 25 (1997), p. 567.
15. Comentario sobre la grabación realizada por The Hanover Band de la Primera sinfonía y el Primer concierto para piano de Beethoven, Nimbus 5003 (1982).
16. Donald Jay Grout, «On historical authenticity in the performance of old music», en Jacob Maurice Coopersmith (ed.), *Essays on Music in Honor of Archibald Thompson Davison,* Cambridge, Mass., Harvard University Press, 1957, p. 346.

## *Lecturas complementarias*

Brown, Howard Mayer, y Stanley Sadie (eds.): *Performance Practice,* 2 vols., Londres, Macmillan, 1989.

Haskell, Harry: *The Early Music Revival: A History,* Londres, Thames and Hudson, 1988.

Kenyon, Nicholas (ed.): *Authenticity and Early Music,* Oxford y Nueva York, Oxford University Press, 1988.

Lawson, Colin, y Robin Stowell: *The Historical Performance of Music: An Introduction,* Cambridge, Cambridge University Press, 1999.

Myers, Helen (ed.): *Ethnomusicology: Historical and Regional Studies,* Londres, Macmillan, 1993.

Raynor, Henry: *A Social History of Music from the Middle Ages to Beethoven,* Nueva York, Taplinger, 1978.

— *Music and Society since 1815,* Nueva York, Taplinger, 1978.

# 2 *La interpretación histórica y el intérprete moderno*

PETER WALLS

## *«Interpretación»*

Este capítulo está dirigido a los «intérpretes modernos», independientemente de que toquen instrumentos históricos o modernos, o canten acompañados de cualquiera de ellos. Podríamos comenzar preguntándonos: ¿Qué es exactamente lo que deberíamos intentar al ejecutar una obra que ha sido compuesta antes de nuestra época? En realidad, la respuesta es esencialmente la misma para la música antigua que para una composición reciente. La mayoría de los intérpretes considerarán que hay que ser fiel a la obra, descubrir su contenido emocional e intentar respetar las intenciones del compositor. Todos valoramos la imaginación y la originalidad de los intérpretes, pero reconocemos que (normalmente) ellas están al servicio de la música que éstos ejecutan, ayudando a descubrir el carácter de la obra o a hacer perceptible su contenido emocional. Por lo general, no nos gusta mucho que la imaginación del intérprete distorsione o disfrace la música que está ejecutando.

Este proceso de realización sonora de una obra musical se conoce generalmente como «interpretación» —aunque algunos compositores han discrepado mucho en cuanto a su significado. Un caso famoso es el de Ravel, que afirmó: «No pido que mi música sea interpretada, sino solamente ejecutada», una afirmación de la que Stravinsky se hizo eco al escribir: «La música debe ser transmitida y no interpretada, ya que la interpretación revela la personalidad del intérprete en lugar de la del autor, y ¿quién puede garantizar que el ejecutante reflejará la visión del autor sin distorsionarla?» [1]. Se queja-

ban, claramente, de las interpretaciones voluntariosas en las que destacan los caprichos de los intérpretes que no se justifican de ninguna manera en la partitura.

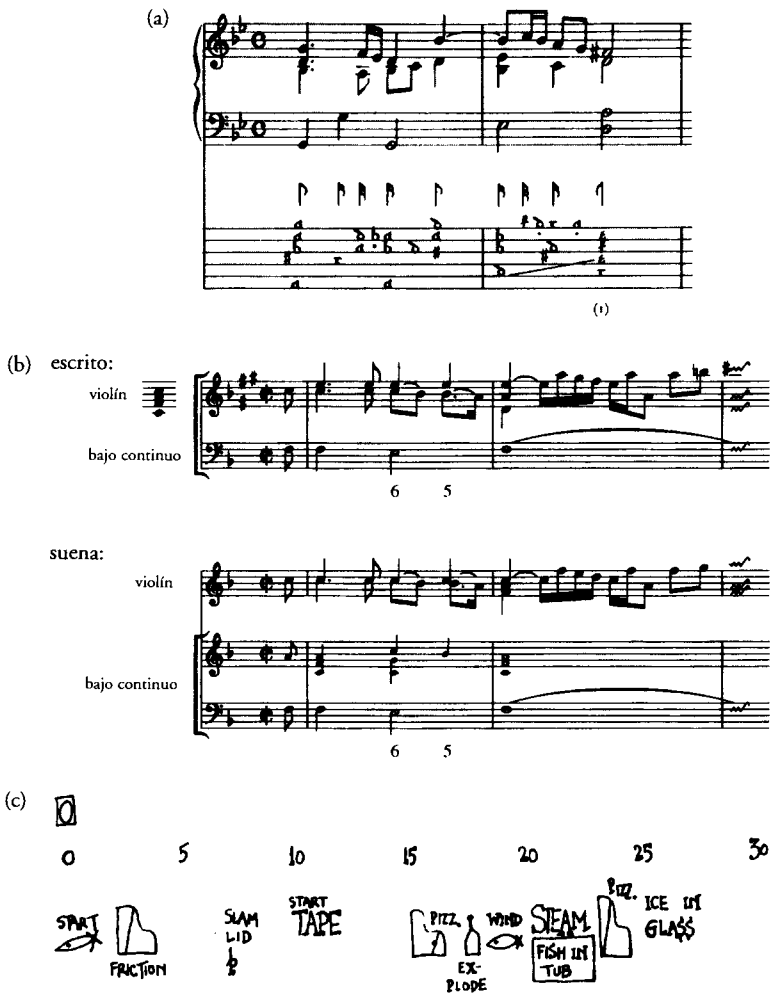
Sin embargo, la palabra «interpretación» debería ser adecuada para expresar lo que deseaban esos compositores. Una de las definiciones del término «interpretar» que ofrece el diccionario de la RAE es: «explicar o declarar el sentido de» una cosa [2]. En este sentido, se puede equiparar a la responsabilidad que tienen los intérpretes de lenguas extranjeras de no tergiversar el significado de lo que traducen.

Considerar la ambigüedad del término «interpretar» constituye una forma bastante útil de distinguir las interpretaciones que son fieles a la música que se está ejecutando de aquellas en las que se interponen otras cosas (como, por ejemplo, la habilidad para presentar algo de manera atractiva). De hecho, podríamos llamar «interpretación» al primer tipo de ejecución (vinculando el término a su significado esencial) y «apropiación» al segundo tipo (ya que, en cierto sentido, la obra musical se convierte en un medio que el intérprete utiliza para su objetivo personal). La interpretación, por lo tanto, puede ser sinónimo de intentar determinar y realizar las intenciones del compositor.

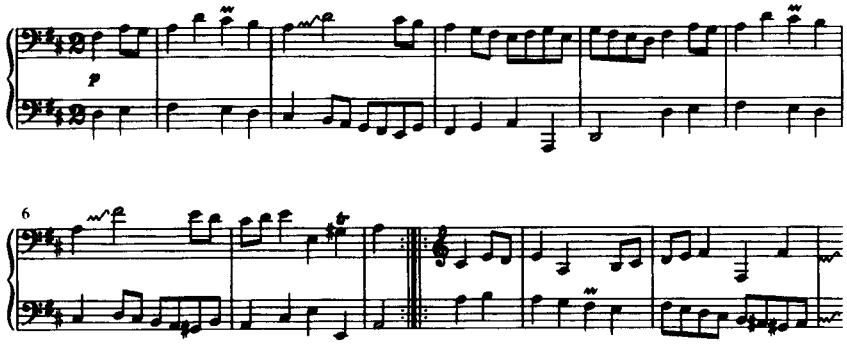
### *Lectura de la partitura*

Toda música escrita requiere de «interpretación». Las partituras musicales aguardan su realización en sonido y, como ocurre con las lenguas extranjeras, la capacidad de leer música es un requisito previo para llevarla a cabo. Esto es algo que todos damos por sentado hasta que nos topamos con una forma de notación que desconocemos, como por ejemplo la tablatura de laúd (Ejemplo 2.1a), la scordatura (Ejemplo 2.1b) o la notación gráfica (Ejemplo 2.1c). Sin embargo, debemos reconocer que, así como la pronunciación y los significados de las palabras cambian con el tiempo, la notación musical aparentemente estándar tiene significados sutilmente diferentes dependiendo de dónde y cuándo fue escrita. De hecho, siempre requiere de una contextualización como parte de este proceso de interpretación. En 1717, François Couperin escribió lo siguiente sobre el problema particular de la música de su propio entorno: «Existen, según mi punto de vista, ciertas deficiencias en la manera en que escribimos la música que están relacionadas con la manera en que escribimos nuestro lenguaje. Se trata de que escribimos algo diferente de lo que tocamos. Ésta es la razón por la que los extranjeros no tocan nuestra música tan bien como nosotros tocamos la suya» [3]. Se refería, específicamente, a que en una serie de corcheas que se mueven por movimiento conjunto, se

EJEMPLO 2.1 *(a) John Dowland, «Lachrimae», compases 1-2, con realización en el sistema superior (© Faber Music Ltd., 1974, 1978, 1981. Reproducido con permiso de Faber Music Ltd. de* The Collected Lute Music of John Dowland, *3ª ed., transcrito y editado por Diana Poulton y Basil Lam); (b) Heinrich Biber, «Sonata del Rosario» núm. 7 («La flagelación de Cristo»), compases 1-2; (c) John Cage,* Water Walk, *principio (Edición Peters núm. 6771; © Henmar Press Inc., Nueva York, 1961. Reproducido en nombre de los editores con la amable autorización de Peters Edition Limited, Londres).*



EJEMPLO 2.2 *J. S. Bach, Gavota II, «Overture nach Französischer Art» BWV 831, compases 1-10.*




debía alargar sutilmente la primera corchea de cada par *(notes inégales)*, aunque su valor escrito fuese el mismo en apariencia.

Muchos elementos, incluso en una pieza bastante simple de música para teclado, presuponen cierto sentido del contexto. El Ejemplo 2.2 muestra la segunda Gavota de la «Ouverture nach Französischer Art» (Obertura en el estilo francés) BWV 831 de J. S. Bach. Esta Obertura (que en este caso significa Suite) fue una de las pocas obras que Bach publicó durante su vida. Apareció en 1735 en la Parte II del *Clavierübung* (Estudio para teclado), cuya portada especifica que fue escrita para «clave con dos manuales». La indicación *piano* que aparece al principio del primer compás tiene un significado importante, ya que con ella el compositor indica que ha de tocarse en el manual superior, cuyos martinetes pulsan un solo juego de cuerdas afinado a 8′ y cuyo sonido corresponde a la nota escrita. En contraste, la Gavota I (que precede y sigue a la Gavota II, como en la relación entre el minueto y el trío) está marcada *forte,* indicando que se debe tocar en el manual inferior con sus dos juegos de cuerdas, probablemente enriquecido incluso con un tercer juego (a 4′) afinado a una octava por encima. Evidentemente, esta alternancia de manuales crea un contraste dinámico entre las dos gavotas pero, más que eso, produce un cambio significativo de timbre.

En la época, este instrumento habría estado «bien temperado». En otras palabras, no estaba afinado según el «temperamento igual», que constituye la manera habitual de resolver un problema fundamental en los instrumentos de teclado moderno, a saber, que el «círculo» de quintas acústicamente puras no regresa exactamente al punto de partida. En el siglo XVIII, la solución preferida consistía en una gama de temperamentos circulares en los que algunas

de las quintas (normalmente cuatro o cinco) son ligeramente menores que las acústicamente puras. Esto crea una especie de *chiaroscuro* (un término que empleó Tartini al recomendar uno de dichos temperamentos) [4] entre la agradable consonancia de los acordes más utilizados y el sonido más bien tenso de las armonías lejanas. Como consecuencia, cada tonalidad posee un carácter propio. Existen muchas instrucciones bastante detalladas del siglo XVIII sobre cómo establecer diversos temperamentos circulares. Determinar cuál era exactamente el que Bach prefería es cuestión de conjeturas, pero todos los temperamentos circulares comparten ciertos principios generales. Un método práctico bastante fiable para imaginar el efecto de estas afinaciones es considerar que cada nota del teclado está afinada adecuándose a su forma más simple, es decir, la que sea inferior según el orden de aparición de las alteraciones; de ahí que la nota inmediatamente por encima del Fa en el teclado sonará mejor cuando funcione como un Fa♯ que como un Sol♭, y así sucesivamente. Por lo tanto, el acorde de Re mayor —la tónica de la Gavota II— probablemente sonará más melodioso que el de Fa♯ (dominante de la Gavota I, que está en Si menor). En la Gavota II, un temperamento circular destacará sutilmente los alejamientos y acercamientos a Re mayor [5].

La indicación de que esta obra fue escrita para ser ejecutada en un clave de dos manuales puede considerarse como implícitamente más específica que lo que esos términos sugieren actualmente. No contamos con una información definitiva sobre los claves de Bach (aunque el inventario de su patrimonio señala que tenía cinco cuando murió). En 1719, Bach fue el encargado de comprar un instrumento de dos manuales de Michael Mietke para la corte de Köthen (de la que era Kapellmeister), y al parecer su constructor era muy respetado en Leipzig. Bach también trabajaba con el constructor de claves Zacharias Hildebrandt, radicado en Leipzig. Dos años antes de la publicación de la segunda parte del *Clavierübung,* un periódico de Leipzig había anunciado la reanudación de los conciertos del Collegium Musicum de Bach, que contarían con el atractivo de «un nuevo tipo de clave *[Clavicymbel]* que no se ha escuchado aquí anteriormente» [6].

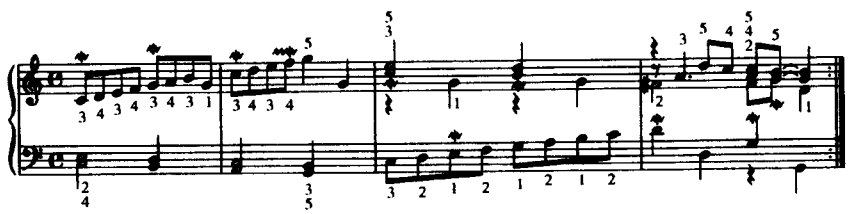
Independientemente de los modelos, el clave posee características que influyen en la técnica. Las teclas ligeramente más cortas hacen que la digitación pianística ortodoxa sea poco práctica (especialmente el paso del pulgar en los pasajes de escala). Según su hijo Carl Philipp Emanuel, Bach era innovador en sus métodos de digitación y empleaba el pulgar más de lo que había sido normal hasta entonces —aunque aún tenía poco que ver con la digitación pianística moderma. Además, los ejercicios de digitación que incluyó en el «Cuaderno» que recopiló para su hijo Wilhelm Friedemann demuestran que esperaba que practicara las digitaciones de los dedos emparejados que se utilizaban entonces habitualmente (véase el Ejemplo 2.3). Tales digitaciones

se prestan a la sutil prolongación de la primera corchea de un par, produciendo *notes inégales.*

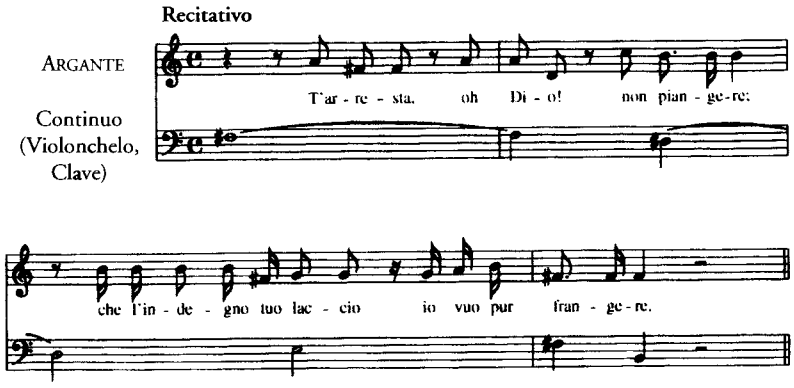
Tal vez las escalas de corcheas del Ejemplo 2.2 deban tratarse como *notes inégales,* ya que la Gavota II es explícitamente una pieza de estilo francés (y Bach era muy versado en la práctica interpretativa francesa). Los signos de ornamentación también son manifestaciones de un rasgo estilístico francés. En efecto, Bach incluyó una tabla de ornamentaciones en el «Cuaderno» de Wilhelm Friedemann que puede ser utilizado para descifrar los signos empleados en esa partitura. Por último, pero no por eso menos importante, es obvio que el hecho de que esta pieza sea una gavota (aunque evidentemente no danzada) influye directamente en el tempo y el carácter rítmico. Johann Joachim Quantz señaló que la gavota es un poco más lenta que una bourrée o un rigaudon, pero comparte el carácter alegre y bien articulado de esas danzas.

El Ejemplo 2.4, un breve recitativo que introduce un aria de Argante en la versión de 1731 de *Rinaldo* de Haendel, muestra un caso en el que es casi imposible leer la notación correctamente sin tener cierto conocimiento de algunos asuntos relativos a la práctica interpretativa. En primer lugar, cualquier cantante experimentado de la compañía de Haendel se habría dado cuenta de que hay que introducir apoyaturas allí donde las notas repetidas sobre un par de sílabas acentuada-no acentuada coinciden con los cambios de armonía y van seguidas de un silencio (en los llamados finales femeninos). Ésta es una manera de tratar debidamente la acentuación de las palabras y tiene una importancia sintáctica (más que decorativa). Dependiendo del gusto del cantante, en la frase «T'arresta» también se podría haber usado una apoyatura. El intérprete del clave continuo habría añadido una suspensión 4-3 en la dominante de la cadencia perfecta para evitar una disonancia desagradable con la apoyatura del cantante. Por otra parte, el violonchelista que toca junto con el clave habría sabido que las notas largas no se han de mantener con todo su valor escrito.

EJEMPLO 2.3 *J. S. Bach,* Klavierbüchlein für Wilhelm Friedemann Bach, *Applicatio BWV 994, compases 1-4.*



EJEMPLO 2.4 *Haendel,* Rinaldo *(versión de 1731), recitativo «T'arresta, oh Dio!».*




¿Cómo aprendían estos principios los intérpretes del siglo XVIII? Los aprendían de sus maestros, sencillamente al escuchar a sus colegas, de los tratados y de los ejemplos escritos. ¿Cómo los asimilamos hoy en día? De manera directa, sólo de las últimas dos fuentes (e indirectamente de dichas fuentes tal y como las interpretan nuestros maestros o intérpretes favoritos). Las apoyaturas de este tipo fueron descritas, por ejemplo, por Francesco Gasparini en su obra *L'armonico pratico al cimbalo* (Venecia, 1708)[7]. Un análisis aún más detallado puede encontrarse en la traducción profusamente comentada que hiciera Johann Friederich Agricola del tratado sobre canto de Pier Francesco Tosi[8]. Alessandro Scarlatti solía escribir exactamente las notas que quería que se cantasen en las cadencias como la que concluye el recitativo del Ejemplo 2.4. El aria que le sigue supone un conocimiento similar de las convenciones —particularmente las relacionadas con la ornamentación. Éstas también son analizadas por Tosi (y Agricola), pero existen algunos fascinantes ejemplos de arias operísticas ornamentadas por el propio Haendel[9].

¿Quién debería cantar esta música? En su producción de 1731, Haendel asignó el papel protagonista de Rinaldo al famoso castrato Francesco Bernardi, conocido como Senesino. Pero la parte de Argante (otro papel masculino) era cantada por la contralto Francesca Bertolli. Actualmente, los papeles masculinos de voces agudas tienden a ser asignados a falsetistas (contratenores). El contratenor operístico es un invento del siglo XX —una solución innecesaria al problema de la ausencia de castrati. Cuando Haendel se encontró en una situación similar (es decir, cuando no encontraba castrati que fueran suficientemente buenos), simplemente asignaba el papel a la mejor

cantante femenina de su compañía. Si queremos ser históricamente exactos, debemos seguir su ejemplo.

Resulta cómodo considerar que la notación es única o esencialmente preceptiva en relación a la altura y la duración de las notas. Sin embargo, los ejemplos de Bach y Handel que hemos analizado nos ayudan a identificar otros parámetros implícitos en la forma en que está escrita la música. Podríamos, por lo tanto, crear una lista de preguntas básicas con el propósito de alcanzar un conocimiento abundantemente contextualizado de la partitura musical:

— ¿Para qué tipo de voces o instrumentos fue escrita?
— ¿Cómo se deberán disponer (afinar, etc.) esos instrumentos?
— ¿Qué tipo de técnica se pretende y cuáles son las indicaciones de ésta en cuanto a la articulación, el fraseo, el timbre, etc.?
— ¿Ha supuesto el compositor la existencia de un conocimiento de las convenciones rítmicas o sintácticas no explicitadas en la partitura?
— ¿Qué tipo de ornamentación, si la hay, es la apropiada?
— ¿Existen recursos expresivos que hubiesen parecido usuales cuando la música fue compuesta, pero que ya no forman parte del vocabulario del músico experto?
— ¿Existen aspectos de la expresividad moderna que no son adecuados para este repertorio?

Todas estas preguntas están tan estrechamente relacionadas con el conocimiento de la notación y sus matices, que hasta dudo si considerarlas como aspectos del contexto. Sin embargo, a esta lista (la lista «A») se podrían añadir otras preguntas (la lista «B») que también pueden influir en la manera en que abordamos la ejecución de una obra. ¿Para qué ocasiones fue escrita? ¿Fue compuesta teniendo en mente un tipo de acústica en particular?, etcétera.

Dar con las respuestas «correctas» a todas estas preguntas nunca es fácil, y muchas veces es imposible; además, dichas respuestas no son sustitutos para la imaginación del intérprete. En el mejor de los casos, el campo en el que funciona la imaginación es uno en el que una serie de suposiciones adquiridas es reemplazada por una gama alternativa de posibilidades estilísticas. El proceso de reemplazo puede en sí mismo actuar como estímulo para la imaginación musical. Carl Dahlhaus sostuvo que lo que distingue a una historia de la música —y, por extensión, cualquier historia de las artes— de otros tipos de historia es que se ocupa no sólo de cosas del pasado (hechos, documentos), sino de un presente estético [10]. Podríamos hacer una afirmación similar sobre el papel de los intérpretes. El intérprete, cuya tarea es realizar esa

partitura para un público contemporáneo, está especialmente involucrado en ese acto de mediación entre un pasado histórico y un presente estético.

### *El instrumento «apropiado»*

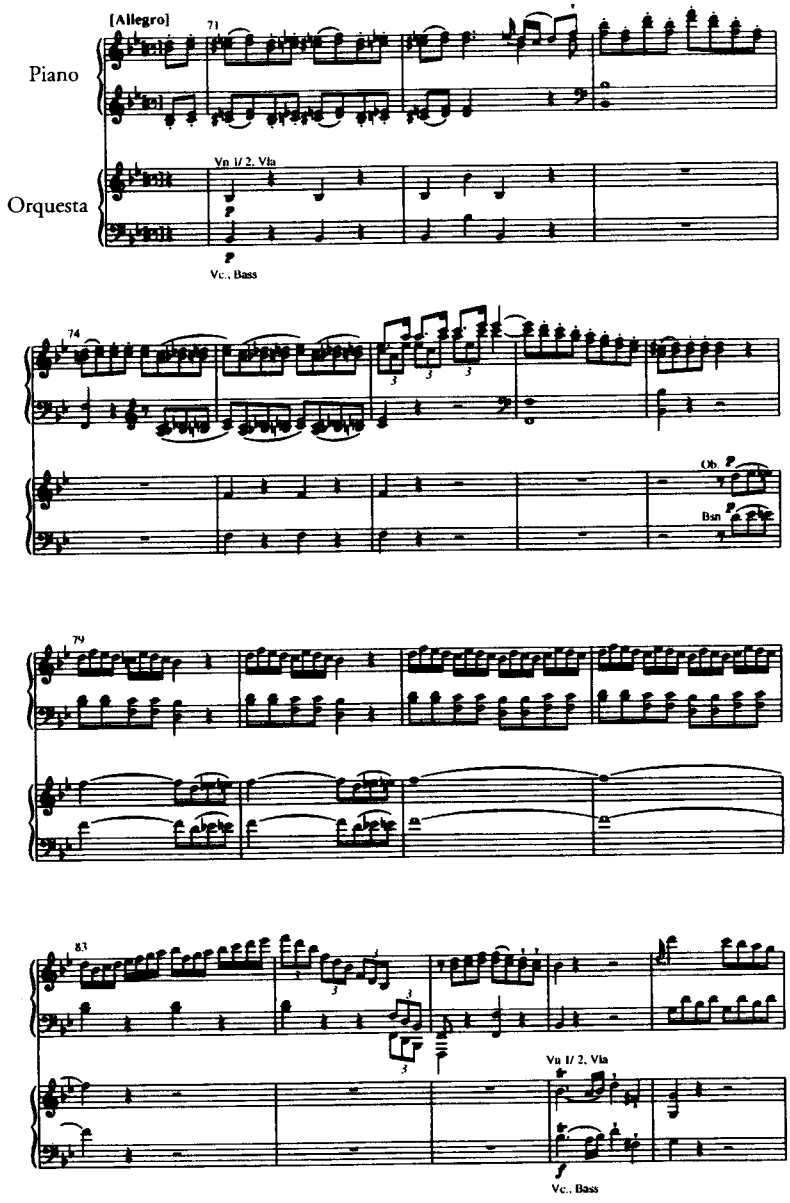
En este punto, quiero examinar el parámetro más obvio de los descritos anteriormente —la elección del instrumento— y considerar su importancia como parte esencial de un conocimiento riguroso de la música escrita antes de nuestra época. Esto a su vez nos conduce a otra pregunta: ¿Hasta qué punto los intérpretes deberían tener libertad para hacer caso omiso de (o seguir desconociendo) la información históricamente verificable sobre cómo sonaba originalmente la música que tocan?

El interés por los instrumentos históricos se vio impulsado por la convicción de que la mejor manera de tratar la música de épocas anteriores es recreando los timbres instrumentales previstos por el compositor (en la medida en que éstos se puedan establecer). Existe, claro está, un punto de vista alternativo, que encontró su expresión más elocuente en los años 1870, cuando Philipp Spitta afirmó que el pianoforte moderno «flotaba en la mente de Bach» [11]. Spitta, incapaz de apreciar la belleza de un buen clave alemán del siglo XVIII (y que quizá nunca pudo escuchar uno en condiciones), estaba convencido de que la concepción intelectual de Bach iba por delante de la tecnología instrumental —en otras palabras, que sus obras para teclado sólo fueron reconocidas cuando surgió un instrumento digno de ellas.

De hecho, es interesante especular sobre cuál habría sido la reacción, por ejemplo, de Mozart, ante un piano de mediados del siglo XIX, o incluso ante un Steinway moderno de nueve pies. Puede que hubiera pensado (a la manera de Spitta): «éste es el instrumento de mis sueños; es lo que realmente quería para mis conciertos», o en cambio, «si tan sólo hubiera conocido este maravilloso instrumento, habría compuesto alguna música para él». Por otra parte, puede que lo hubiera rechazado como un producto poco refinado y más bien vulgar de una sociedad industrializada. No podemos saber con certeza cuál de estas reacciones (en caso de que fuera una de ellas) es la que habría provocado el piano en Mozart —pero sí sabemos que le entusiasmaron los fortepianos que conoció a finales de la década de 1770. Este hecho en sí mismo debería ser más que suficiente para que cualquier intérprete que quiera tocar las obras de Mozart las tenga muy en cuenta.

Lo primero que uno nota cuando se ejecuta una obra de Mozart en un instrumento de los que él conoció y amó es que la imagen generalizada de su música como una expresión clásica de discreción y elegancia tiende a desvanecerse con bastante rapidez. Incluso una obra como el Concierto para piano

EJEMPLO 2.5 *Mozart, Concierto para piano en Si♭ mayor K. 450, primer movimiento, compases 71-78 (piano y reducción orquestal).*



en Sib mayor, K. 450, que parece evocar un mundo de *dramma giocoso* al estilo de *Figaro,* utiliza los recursos del piano al máximo. El Ejemplo 2.5 comienza con el punto en el que el piano expone una idea delineada anteriormente en los compases iniciales en un hermoso diálogo entre los vientos y las cuerdas.

Cuando la parte solista alcanza la cadencia en el compás 86 (donde las cuerdas realizan una transición de un compás a la relativa menor), se expande hasta abarcar el ámbito completo de cinco octavas ($Fa_1$ a fa''') del instrumento de Mozart. Además, el sonido transparente del instrumento conviene tanto a las dobles terceras, tan predominantes en este ejemplo, como a las ligeras semicorcheas.

El manuscrito autógrafo de la Sonata en Mib mayor Hob. XVI:49 de Haydn asigna la obra al «Forte-piano» —siendo la primera sonata que el compositor describe de esta manera. Además, en tres cartas que le escribió a Maria Anna von Genzinger durante la época en que completó la sonata, en junio de 1790, deja bastante claro que la había compuesto con un instrumento específico en mente:

> He escrito esta Sonata para Vuestra Alteza hace un año, y sólo el Adagio es bastante reciente; recomiendo que preste especial atención a este movimiento, ya que contiene muchas cosas que analizaré para Vuestra Alteza cuando llegue la ocasión; es bastante difícil pero está lleno de sentimiento. Es una pena, sin embargo, que Vuestra Alteza no tenga un fortepiano de Schantz [*sic*], ya que con él podría lograr un resultado doblemente eficaz [12].

En dos cartas subsiguientes explica y reafirma su preferencia, no sólo por el fortepiano, sino por uno de Schanz. En la primera, Haydn insiste en que esta sonata no funcionará muy bien en un clave:

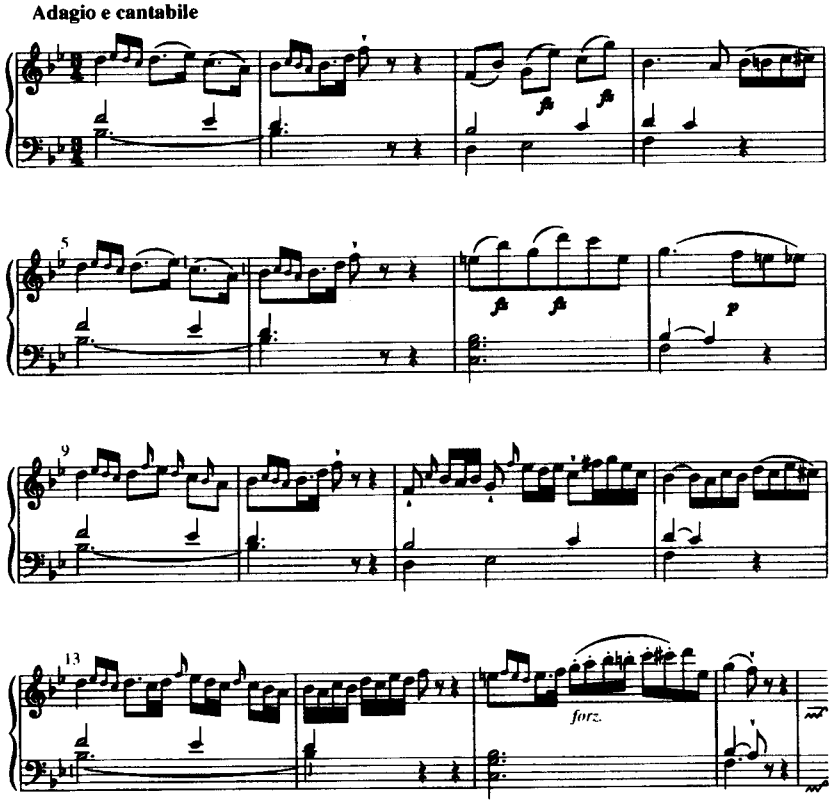
> Es una pena que Vuestra Alteza no posea un fortepiano de Schantz [*sic*], en el que todo se expresa mejor. He pensado que Vuestra Alteza podría traspasar su clave [*Flügl*], que aún está en buenas condiciones, a Fräulein Peperl, y comprar uno nuevo. Vuestras hermosas manos y su habilidad de ejecución se merecen eso y mucho más. Sé que he debido componer esta Sonata de acuerdo con las capacidades de su teclado [*Clavier*], pero me ha sido imposible porque ya no estoy acostumbrado a él... [13].

En la siguiente carta (fechada el 4 de julio de 1790) Haydn explica detalladamente por qué prefiere los instrumentos de Schanz a los de otros constructores vieneses, haciendo hincapié en la ligereza de su respuesta. Es evidente que destaca la importancia que la respuesta del instrumento tenía para él como

*compositor:* ya no estaba acostumbrado ni siquiera a un buen clave, y por ello se le hacía «imposible» escribir de una manera que conviniese a Maria Anna Genzinger. Como consecuencia, la mejor manera de *interpretar* esa obra era reproduciendo no sólo el sonido, sino también la respuesta del instrumento que tenía en mente.

Hay varios puntos en esta sonata en los que Haydn utiliza indicaciones que parecen presuponer la utilización de un fortepiano. Los movimientos primero y último tienen un solo *crescendo*, mientras que el Adagio e cantabile presenta varios *sforzando* y cambios dinámicos dentro de una frase (que no podían realizarse de manera convincente mediante un cambio de manual).

EJEMPLO 2.6 *Haydn, Sonata en Mi♭ mayor Hob. XVI:49, segundo movimiento, compases 1-16.*



Sin embargo, casi todas las características específicamente pianísticas son más sutiles —cuestiones del tipo de fraseo cantabile en el que destacaba el fortepiano (véase el Ejemplo 2.6). No obstante, para una persona suficientemente decidida, esta sonata podía tocarse en un clave.

Como hemos visto, Haydn le recomendó a la buena amiga para quien había escrito la Sonata en Mi♭ que adquiriera un instrumento más apropiado (en este caso, uno más moderno). Podríamos imaginarnos a Haydn adaptando su carta y su recomendación a nosotros. Al fin y al cabo, se justificaría igualmente que le señalara a los pianistas del siglo XXI que un instrumento ligero como los de Schanz le vendría mejor a la sonata que un Steinway, cuyo mecanismo es mucho más pesado. Y, como en el ejemplo de Mozart citado anteriormente, el sonido más transparente de un fortepiano vienés de finales del siglo XVIII es mucho más conveniente para las terceras que predominan en esta sonata. A. Peter Brown comenta:

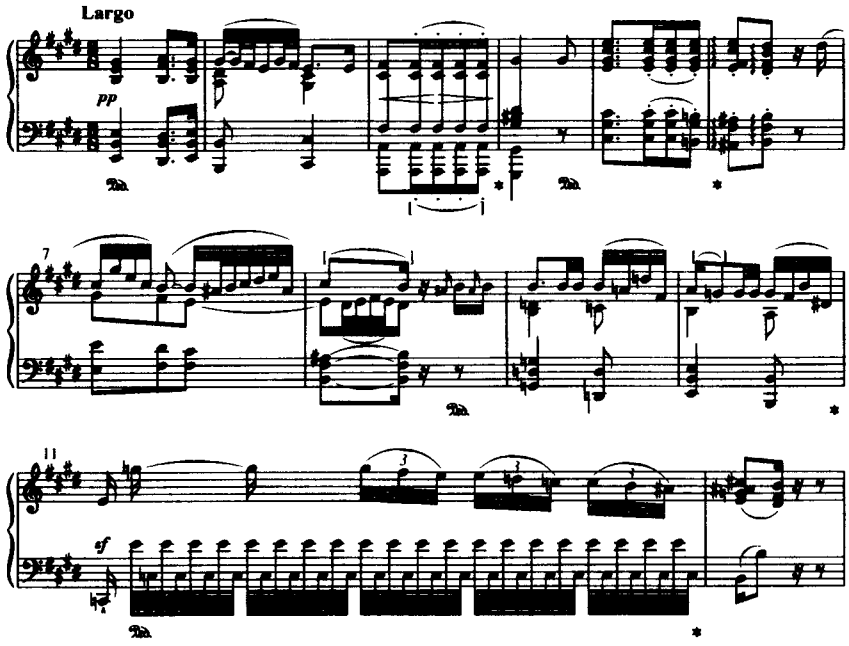
> las diferencias entre el clave y el *fortepiano* del siglo XVIII no son, ni con mucho, tan pronunciadas como la disparidad entre un Steinway o Bösendorfer moderno y el antiguo *fortepiano* Stein o Walter. Los instrumentos fabricados en Alemania y Austria durante el siglo XVIII tenían una articulación brillante y seca que poco tiene que ver estéticamente con los grandes pianos ingleses, de sonido grueso y oscuro, que son los verdaderos antecesores del instrumento moderno [14].

Podríamos señalar argumentos similares sobre el instrumento «correcto» en relación a las músicas posteriores. En el Tercer Concierto para piano, op. 37, de Beethoven, las indicaciones de pedal de los compases iniciales del movimiento lento (Ejemplo 2.7) son incomprensibles y hasta imposibles en un Steinway de concierto moderno, pero constituyen un recurso tímbrico evocativo en el instrumento para el que fue escrito [15]. Carl Czerny, alumno de Beethoven, dijo que el comienzo «debe sonar como una armonía sagrada, distante y celestial» [16].

El incremento del ámbito y del poder sonoro del piano durante los primeros treinta años del siglo XIX redefinió la relación entre el solista y la orquesta en los conciertos para piano. Schumann era muy consciente de ello. No recomendaba que los pianistas buscasen instrumentos antiguos para interpretar conciertos de Mozart o Beethoven, pero sí opinaba que los compositores debían descubrir una nueva manera de abordar el género. En un artículo de 1839 titulado «Das Clavier-Concert» (El concierto para piano) publicado en el *Neue Zeitschrift für Musik* (la revista que había fundado en 1832), escribió:

> Esta ruptura de la relación con la orquesta, como hemos visto, se estaba preparando hace tiempo. El arte pianístico moderno quiere retar a la [orquesta] sinfónica y so-

EJEMPLO 2.7 *Beethoven, Tercer concierto para piano op. 37, segundo movimiento, compases 1-12 (sólo la parte de piano).*




> bresalir mediante sus propios recursos; ello puede explicar la reciente extinción de los conciertos para piano... Sin duda sería una pérdida, en caso de que el concierto para piano y orquesta llegara a ser completamente obsoleto; pero, por otra parte, difícilmente podemos contradecir a los pianistas cuando afirman: «No necesitamos ninguna ayuda, nuestro instrumento puede bastarse por sí mismo». De manera que debemos esperar con confianza al genio que nos mostrará una forma nueva y brillante de combinar la orquesta y el piano, de modo que el autócrata al teclado pueda revelar la riqueza de su instrumento y su arte, mientras la orquesta, más que una mera espectadora, incrementa el conjunto artístico con sus numerosas capacidades expresivas[17].

A los que escuchan por primera vez un piano antiguo les resulta muy fácil advertir que es incapaz de hacer lo que un gran cola de nueve pies hace maravillosamente. Ver más allá de esto, sin embargo, proporciona grandes recompensas —escuchar su flexibilidad dinámica, sus gradaciones tonales y su brillante agilidad, que explican el entusiasmo que sintieron Mozart, Haydn y Beethoven por el instrumento.

Otros instrumentos pueden requerir inicialmente el mismo tipo de adaptación. En el caso del violín, vivimos la absurda paradoja de que, aunque Antonio Stradivari es reconocido popularmente como el mejor constructor de violines de todos los tiempos, en la práctica no confiamos en que supiera lo que estaba haciendo. Ninguno de los instrumentos de Stradivari que utilizan los virtuosos hoy en día conservan los ajustes que tenían cuando salieron del taller de Cremona. Sus barras armónicas han sido reforzadas, se les han insertado almas más gruesas, les han cambiado los puentes, se les ha recolocado el mango con otro ángulo, les han cambiado el tipo de cuerdas y se les han añadido barbadas. Por lo tanto, su sonido ha cambiado significativamente de diversas maneras que seguramente favorecen la ejecución de los conciertos del siglo XIX, pero perjudican a la música de la época de Stradivari.

## *El arte de lo posible*

Una vez que hemos aceptado que las intenciones de Haydn (o cualquier otro compositor) se realizarán probablemente de la mejor manera en los instrumentos para los que fueron concebidas sus obras, existe otra pregunta práctica y pertinente que debemos confrontar como músicos del siglo XXI: ¿hasta qué punto deberíamos esforzarnos por adaptar los instrumentos a los repertorios? Después de todo, Haydn recomienda no sólo un fortepiano vienés, sino los instrumentos de un constructor en particular. Sin embargo, para la época en que compuso su siguiente y maravillosamente exuberante Sonata en Mi♭ mayor Hob. XVI:52 en 1794, le habían cautivado los instrumentos que había conocido en Inglaterra de John Broadwood y sus contemporáneos, que eran más robustos, con mecanismos más pesados y de sonido más rico. Cuando Haydn regresó a Viena en 1795 después de su segunda visita a Inglaterra, trajo consigo un nuevo «Grand piano Forte» de Longman y Broderip (cuyo ámbito se extendía hasta el do′′′′). De modo que, ¿debemos suponer que las tardías sonatas «inglesas» le estaban vedadas a Maria Anna von Genzinger (muy poco después de haber convencido a su marido para que invirtiera en un Schanz) y debían ser evitadas por los fortepianistas que no tuviesen acceso a un piano de estilo inglés de finales del siglo XVIII? En realidad, las últimas tres sonatas de Haydn no sobrepasan el ámbito de los instrumentos vieneses (aunque lo explotan al máximo), de modo que quizá Haydn quería asegurarse de que esas obras, por más que hubiesen estado inspiradas en la sonoridad del piano inglés, siguieran siendo accesibles para sus compatriotas.

Beethoven proporciona una interesante analogía, ya que sus composiciones para piano abarcan el período de mayores cambios de la historia del ins-

trumento. Las sonatas tardías sencillamente no pueden tocarse en el instrumento para el que fueron concebidas sus primeras obras. Entonces, los pianistas que deseen interpretar todas las obras para piano solo de Beethoven ¿deberán adquirir media docena de instrumentos para hacerlo?

En casos como éste, se debe distinguir entre una solución ideal (a la que frecuentemente se puede acceder en los proyectos de grabación comerciales) y una en general más práctica. Resulta interesante reflexionar sobre los testimonios históricos del pragmatismo que, para la mayoría de los músicos, constituye la única alternativa a verse confinados a un segmento cada vez menor del repertorio para su instrumento. Todos los compositores analizados en este ensayo combinan unas preferencias instrumentales muy precisas con una disposición a reconocer las exigencias de las circunstancias interpretativas reales.

La historia de la creación y revisiones de numerosas obras de Bach demuestra su voluntad de adaptarse a las circunstancias, cambiar de idea o explorar otras alternativas. En la *Pasión según San Juan,* por ejemplo, los laúdes y las violas d'amore de «Erwäge, wie sein blutgefärbter Rücken» son reemplazados por el órgano y los violines con sordina en la tercera y cuarta versiones. Los siete conciertos para clave solo (BWV 1052-1058) son considerados transcripciones de obras que fueron compuestas originalmente para otros instrumentos. Sin duda, dos de ellos —el Concierto en Re mayor BWV 1054 y el Concierto en Sol menor BWV 1058— son arreglos de los conciertos para violín en Mi mayor BWV 1042 y en La menor BWV 1041, respectivamente.

Sabemos que a Mozart le encantaban los clarinetes. Después de visitar Mannheim en 1778, le escribió una carta a su padre en la que expresa un gran entusiasmo por la eficacia de los clarinetes en un grupo orquestal: «¡Ah, si tuviéramos además clarinetes! No puedes imaginar el glorioso efecto de una sinfonía con flautas, oboes y clarinetes» [18]. Mozart continuó, claro está, escribiendo para el instrumento de manera memorable. Pero estaba preparado para contemplar las sustituciones en circunstancias en las que, como le acababa de ocurrir en Salzburgo, no había clarinetes disponibles. En 1786, al enviar las partes del Concierto para piano en La mayor K. 488 que iban a ser utilizadas en la casa del príncipe von Fürstenberg, escribió:

> Hay dos clarinetes en el concierto en La mayor. En caso de que Su Alteza no disponga de ningún clarinetista en su corte, un copista competente puede transportar las partes a las tonalidades adecuadas, en cuyo caso la primera parte ha de ser tocada por un violín y la segunda por una viola [19].

Igualmente, Beethoven propuso el violín como una alternativa al clarinete en su Trío en Si♭ mayor op. 11 (1798), y, como ocurre con las de Haydn, las

portadas de muchas de sus sonatas para piano indican piano o clave como alternativas (aunque eso se debe más bien a las prioridades comerciales del editor que a una sincera flexibilidad por parte del compositor).

La evidente disposición de estos compositores a adaptarse a las circunstancias debería ser un incentivo para los intérpretes modernos. Al menos en ciertas ocasiones, ellos estaban preparados para contemplar sustituciones cuando el instrumento que tenían en mente para una obra en particular no estaba disponible. Sin embargo, esa actitud realista no es lo mismo que la indiferencia, y no puede utilizarse para justificar un incumplimiento total de las intenciones del compositor en cuanto a las alternativas instrumentales. Además, sería una locura, ya que para participar de la esencia de las obras del pasado es fundamental que nos acerquemos —literalmente— lo más posible a su mundo sonoro. (Por desgracia, los límites de este ensayo no nos permiten considerar la falacia de que para lograr algo parecido al efecto original habría que aumentar los decibelios y timbres para los oídos acostumbrados a los espectaculares conciertos actuales.)

## *Conclusión*

En este punto podemos regresar al concepto de «interpretación» tratado al principio del capítulo. Si es responsabilidad del intérprete realizar las intenciones del compositor, entonces el primer paso es, sin duda, intentar comprender plenamente la música. Algunos aspectos de esa comprensión no implicarán directamente (o al menos no conspicuamente) un sentido histórico. El análisis, obviamente, es una parte esencial de ese proceso. Pero son muchos los aspectos de la comprensión de la música (que en esencia es lo mismo que leerla correctamente) que requieren una perspectiva histórica.

Como he indicado antes, esa perspectiva puede verse simplemente como leer una partitura correctamente con el mayor conocimiento posible de sus matices. He definido esto como una lectura abundantemente contextualizada. Sin embargo, debemos refinar el concepto de contexto histórico, ya que al abordar prácticamente cualquier pieza musical tenemos que juzgar musicalmente cuáles son los aspectos históricos pertinentes a nuestra comprensión de la partitura y cuáles pueden ser desechados por irrelevantes o innecesarios. En otras palabras, debemos decidir cuáles son las preguntas que pertenecen a las listas «A» y «B», mencionadas anteriormente, y cuáles pueden relegarse a una «lista C», formada por características de las interpretaciones históricas que, por más interesantes que sean, nunca emularíamos [20]. El primer paso del juicio musical es decidir cuáles son los elementos históricos que no incorporaremos para preservar el aspecto estético de la música. Sin

embargo, el límite entre el «mero» contexto y los detalles que consideramos esenciales no es inalterable. Debemos permanecer abiertos a la posibilidad de descubrir la relevancia de detalles que parecen poco importantes.

El punto de vista sustentado en este ensayo se ha descrito frecuentemente como una búsqueda de la autenticidad histórica. Ese concepto, sin embargo, está rodeado de escollos. Como han señalado Peter Kivy y otros, es imposible en la práctica satisfacer las diversas facetas contenidas en una autenticidad absoluta —o, al menos, sería poco realista esperar una conformidad general con respecto a la autenticidad de un caso específico [21]. Así que, olvidémonos de la autenticidad, un término que ha creado más polémica de la que merece. Entre las diversas alternativas que se han sugerido, la «interpretación históricamente informada» es la que ha encontrado más aceptación. Pero también ésta ha sido cuestionada. Según Richard Taruskin, es inaceptable porque no supera el «test del antónimo», y resulta de poca ayuda para describir alternativas válidas [22]. No obstante, la objeción de Taruskin presupone que todo intérprete tiene la libertad de elegir entre adoptar o rechazar un enfoque que toma en cuenta lo que se puede demostrar de las intenciones del compositor. Esto sólo es cierto si aceptamos igualmente que el intérprete también puede elegir entre tocar las notas correctas o no. Por último, podríamos argumentar que es inaceptable tocar Bach en un instrumento apropiado para Brahms y sin tener en cuenta sus expectativas en cuanto a ciertos asuntos como la articulación y la ornamentación. Quizá deberíamos afrontar el hecho de que los intérpretes que pretenden tratar debidamente el aspecto estético de la música haciendo caso omiso de las implicaciones históricas de la partitura no merecen ser considerados como «diversamente capacitados» sino «históricamente desinformados».

## *Notas*

1. Respectivamente, Maurice Ravel, citado por Marguerite Long, *At the Piano with Maurice Ravel,* Londres, Dent, 1973, p. 16; Igor Stravinsky, *Autobiography,* Nueva York, Simon and Schuster, [1936]; Nueva York, Norton, 1962, p. 75. Long también cita el siguiente diálogo entre Ravel y Toscanini: Ravel —«Ése no es mi tempo»—, Toscanini —«Cuando toco su tempo, la pieza resulta ineficaz»—, Ravel —«Entonces, no la toque»—, *op. cit.*, p. 18.
2. Otras definiciones son: «Explicar, acertadamente o no, acciones, dichos o sucesos que pueden ser entendidos de diferentes modos» (RAE); «Atribuir cierto significado a una expresión o a otra cosa», «Particularmente, dar forma el artista, reflejando en ello su personalidad, a lo que copia, reproduce o ejecuta» (María Moliner: Diccionario de uso del español).
3. François Couperin, *L'Art de toucher le clavecin,* París, François Couperin, 1717, p. 39.
4. Giuseppe Tartini, *Trattato di musica secondo la vera scienza dell'armonia,* Padua, Seminario, 1754, p. 100, citado por Mark Lindley, «La "Pratica ben regolata" di Francescantonio Vallotti», *Rivista Italiana di Musicologia* 16 (1981), p. 72.

5. Para un punto de vista contrario, véase: Rudolf Rasch, «Does "well-tempered" mean "equal-tempered?"», en Peter Williams (ed.), *Bach, Handel, Scarlatti: Tercentenary Essays,* Cambridge, Cambridge University Press, 1985, pp. 293-310.
6. Citado por Hans T. David y Alfred Mendel (eds.), *The New Bach Reader,* revisado y ampliado por Christoph Wolff, Nueva York y Londres, Norton, 1998, p. 156. Véanse también la entrada «harpsichord» en Malcolm Boyd (ed.), *Oxford Composer Companions: J. S. Bach,* Oxford, Oxford University Press, 1999, pp. 207-208; John Koster, «The quest for Bach's *Clavier:* an historiographical interpretation», *Early Keyboard Journal* 14 (1996), pp. 65-84; y Sheridan Germann, «The Mietkes, the Margraves and Bach», en Peter Williams (ed.), *op. cit.,* pp. 119-148.
7. Traducido al inglés por Frank S. Stillings y Donald L. Burrows, *The Practical Harmonist at the Harpsichord,* New Haven y Londres, Yale University Press, 1963.
8. Johann Friedrich Agricola, *Anleitung zur Singekunst,* Berlín, Winter, 1757; para el uso de las apoyaturas en los recitativos, véase la traducción inglesa de Julianne C. Baird, *An Introduction to the Art of Singing by Johann Friedrich Agricola,* Cambridge, Cambridge University Press, 1995, pp. 174 ss.
9. Véanse George Frideric Haendel, *Three Ornamented Arias,* Winton Dean (ed.), Londres, Oxford University Press [c. 1976]; y Winton Dean, «Vocal embellishment in a Handel aria», en H. C. Robins Landon (ed.), *Studies in Eighteenth-Century Music,* Londres, Allen & Unwin, 1970, pp. 151-159.
10. Carl Dahlhaus, *Foundations of Music History,* J. B. Robinson (trad.), Cambridge, Cambridge University Press, 1983, pp. 4 ss.
11. Philipp Spitta, *Johann Sebastian Bach,* Clara Bell y J. A. Fuller-Maitland (trad.), 2 vols., Londres, Novello [1884-1885]; Nueva York, Dover, 1951, vol. II, p. 44.
12. Carta escrita en Esterháza el 20 de junio de 1790 a Maria Anna von Genzinger, Viena; H. C. Robbins Landon, *Haydn, Chronicle and Works,* vol. II: *Haydn at Esterháza 1766-1790,* Londres, Thames and Hudson, 1978, pp. 743-744.
13. Carta escrita en Esterháza el 27 de junio de 1790 a Maria Anna von Genzinger, Viena; *ibid.*, p. 745. He cambiado la traducción que hace Landon del término *Clavier* como «piano» por «teclado»; según A. Peter Brown (*Joseph Haydn's Keyboard Music: Sources and Style,* Bloomington, Indiana University Press, 1986, p. 146), «Haydn utiliza el término *Clavier* en un sentido genérico». Véase el texto original en Dénes Bartha y H. C. Robbins Landon (eds.), *Joseph Haydn: Gesammelte Briefe und Aufzeichnungen,* Kassel, Bärenreiter, 1965, p. 242.
14. A. Peter Brown, *op. cit.,* p. 171.
15. No es fácil determinar cuál era exactamente el instrumento que Beethoven tenía en mente en este caso. Dado que el último movimiento asciende hasta do'''' (en los compases 346-349), el candidato más probable es un Erard del tipo que adquirió en 1804 (después de la primera interpretación del op. 37 en abril de 1803). Sin embargo, ya había otros instrumentos de ese tipo en Viena.
16. Citado por Leon Plantinga, *Beethoven's Concertos,* Nueva York y Londres, Norton, 1999, pp. 143-144.
17. Citado por Leon Plantinga, *Schumann as Critic,* New Haven, Yale University Press, 1967, p. 204.
18. Carta escrita en Mannheim el 3 de diciembre de 1778 a Leopold Mozart, Salzburgo; Wolfgang Amadeus Mozart, *The Letters of Mozart and His Family,* 3ª ed., Emily Anderson (trad. y ed.), revisado por Stanley Sadie y Fiona Smart, Londres, Macmillan, 1985, p. 638.
19. Carta escrita en Viena el 30 de septiembre de 1786 a Sebastian Winter, Donaueschingen; *ibid.*, p. 900.

20. Sería absurdo, por ejemplo, considerar el hecho de que en 1778 Mozart interpretó sus Variaciones sobre un Minueto de J. C. Fischer K. 179 en medio de un frío glacial ante un público apático en París como estímulo para reproducir esas condiciones al interpretar esa obra hoy en día. Véase la carta que escribió Mozart el 1 de mayo de 1778 en París a su padre Leopoldo en Salzburgo, en *ibid.*, p. 531.
21. Véase Peter Kivy, *Authenticities: Philosophical Reflections on Musical Performance,* Ithaca y Londres, Cornell University Press, 1995.
22. Richard Taruskin, «The pastness of the present and the presence of the past», en Nicholas Kenyon (ed.), *Authenticity and Early Music,* Oxford y Nueva York, Oxford University Press, 1988, p. 140.

## *Lecturas complementarias*

Brown, Clive: *Classical and Romantic Performing Practice 1750-1900,* Oxford, Oxford University Press, 1999.

Brown, Howard Mayer, y Stanley Sadie (eds.): *Performance Practice,* 2 vols., Londres, Macmillan, 1989.

Cyr, Mary: *Performing Baroque Music,* Portland, Or, Amadeus, 1992.

Donington, Robert: *The interpretation of Early Music,* Londres, Faber [1963], 1989.

Dreyfus, Lawrence: «Early Music defended against its devotees: a theory of historical performance in the twentieth century», *Musical Quarterly* 49 (1983), pp. 297-322.

Haskell, Harry: *The Early Music Revival: A History,* Londres, Thames and Hudson, 1988.

Kenyon, Nicholas (ed.): *Authenticity and Early Music,* Oxford y Nueva York, Oxford University Press, 1988.

Kivy, Peter: *Authenticities: Philosophical Reflections on Musical Performance,* Ithaca y Londres, Cornell University Press, 1995.

Lawson, Colin, y Robin Stowell: *The Historical Performance of Music: An Introduction,* Cambridge, Cambridge University Press, 1999.

MacClintock, Carol (ed.): *Readings in the History of Music in Performance,* Bloomington y Londres, Indiana University Press, 1979.

# 3 *Análisis y (¿o?) interpretación*

JOHN RINK

La confusión y la controversia tienden a reinar cada vez que se utiliza el término «análisis» en relación con la interpretación musical. Algunos estudiosos consideran que el análisis «está implícito en lo que hace el intérprete», por más «intuitivo y asistemático» que sea [1], mientras que para otros, los intérpretes *deben* emplear el análisis teórico riguroso de los «elementos paramétricos» de la obra para poder sondear su «profundidad estética» [2]. Indudablemente, la interpretación musical requiere la toma de decisiones —conscientes o no— en cuanto a las funciones contextuales de determinadas características musicales y la manera de proyectarlas. Incluso el pasaje más sencillo —por ejemplo, una escala o una cadencia perfecta— estará formulado según la manera en que el intérprete considere que dicho pasaje encaja dentro de una pieza determinada y las prerrogativas expresivas que le conceda. Tales decisiones no son necesariamente intuitivas y asistemáticas: la mayoría de los intérpretes examinan meticulosamente cómo «funciona» la obra y cómo superar sus diversos retos conceptuales. Ese proceso es analítico en muchos sentidos, pero es necesario explicar en qué consiste.

El objetivo de este capítulo es examinar la dinámica que existe entre los pensamientos intuitivo y consciente que caracteriza potencialmente al proceso analítico en relación con la interpretación. Después de examinar algunos escritos en torno a este tema, describiré un método de análisis que puede beneficiar a los intérpretes en lugar de constreñirlos [3]. Dicho método será ilustrado en un estudio de caso del Nocturno en Do♯ menor op. 27 núm. 1 de Chopin.

## *Perspectivas sobre el análisis y la interpretación*

En un fascinante ensayo escrito en 1985, Janet Schmalfeldt le da voz a dos personalidades —la del analista y la del intérprete— para considerar la influencia mutua de sus respectivas interpretaciones de dos Bagatelles de Beethoven. Su diálogo es ingenioso pero cuestionable: el analista se expresa con mucha más autoridad que el intérprete, que parece subordinado a los mandatos teóricos de su contrario; además, la división en dos personalidades antagónicas apenas refleja cómo actúan la mayoría de los músicos. Sin embargo, la conclusión de Schmalfeldt parece acertada: «no existe una sola y única decisión interpretativa que pueda ser dictada por una observación analítica» [4]. Esta conclusión se hace eco del comentario que hizo Edward Cone en su clásica obra *Musical Form and Musical Performance,* «toda interpretación válida [...] representa, no una aproximación a un ideal, sino una decisión: ¿cuáles son las relaciones implícitas en la pieza que han de ser destacadas o explicitadas? [5].

Se han hecho comentarios similares en otros escritos, pero, como he señalado, también existe una escuela de pensamiento más doctrinaria que prácticamente exige a los intérpretes basar su interpretación en los resultados de un análisis riguroso. Uno de sus representantes es Eugene Narmour, quien en 1988 insistió en que «los intérpretes, como cocreadores [...] deben adquirir capacidad teórica y analítica [...] para saber cómo interpretar», y que se producen «muchas consecuencias negativas si las relaciones formales no son analizadas correctamente por el intérprete» [6]. Wallace Berry señaló igualmente que «la interpretación musical debe estar basada en un profundo análisis»; a diferencia de Schmalfeldt y Cone, Berry opina que «todo hallazgo analítico tiene una consecuencia en la interpretación» [7].

En 1989, Jonathan Dunsby cuestionó la legitimidad de tales preceptos, afirmando que «entender y tratar de explicar una estructura musical no constituye el mismo tipo de actividad que entender y comunicar la música. Estos polos de actividad se solapan indudablemente, pero no pueden superponerse completamente». Según Dunsby, un enfoque más deliberadamente analítico puede ayudar a los intérpretes a enfrentarse a los pasajes difíciles, pero sus enfoques generalmente pragmáticos excluyen el rigor metodológico asociado normalmente al análisis [8]. En una respuesta a su planteamiento, argumenté que «los intérpretes emplean continuamente un proceso de "análisis", sólo que es diferente del que se emplea en los análisis escritos. Este tipo de "análisis" no es un procedimiento independiente que se aplica al acto de interpretar», sino más bien «un componente del proceso interpretativo». Refiriéndolo como «análisis del intérprete» (es decir, «el estudio esmerado de la partitura, prestando especial atención a las funciones contextuales y la manera de proyectarlas»), destaqué que la «forma» musical es más importante que la estruc-

tura en la conceptualización que hace el intérprete de la música —una noción escurridiza pero esclarecedora cuya concepción es más temporal que la de la estructura. La temporalidad de la música es fundamental a este respecto, un factor que se ignora o subestima en algunos tipos de «análisis riguroso» [9], con consecuencias negativas cuando los resultados de dichos análisis se utilizan directamente en la interpretación. También propuse el término «intuición instruida», que reconoce la importancia de la intuición en el proceso interpretativo, y también reconoce que generalmente implica bastante conocimiento y experiencia —en otras palabras, que la intuición no necesariamente surge de la nada, ni tampoco es necesariamente caprichosa [10].

### *Objetivo y método*

A pesar de las muchas discrepancias entre los estudios publicados, el análisis relacionado con la interpretación puede dividirse en dos categorías principales:

(1) análisis previo a una interpretación determinada (que probablemente servirá de base para ésta),
(2) análisis de la interpretación misma.

El primero, bien sea riguroso o de carácter más pragmático, es potencialmente *preceptivo* con respecto a la interpretación, mientras que el segundo tipo de análisis es *descriptivo.* El análisis post facto no es mi objetivo principal en este capítulo, aunque hay que destacar su importancia. Los gráficos de tempo y dinámica de la Figura 3.1 (basados en interpretaciones grabadas y en directo, respectivamente) ofrecen un ejemplo [11]; además, los músicos normalmente analizan (es decir, descomponen y evalúan) sus interpretaciones posteriormente, intentando encontrar diferentes y mejores maneras de comprender la música para sus futuras ejecuciones en un proceso de constante evolución.

He delimitado en otra ocasión los problemas relacionados con algunas «prescripciones», y la conclusión es que los intérpretes quizá hacen bien oponiéndose a todo intento *sistemático* de correlacionar los resultados de los métodos analíticos rigurosos con la verdadera interpretación. Por ejemplo, las demostraciones analíticas de una unidad motívica pueden ser fascinantes sobre el papel, pero generalmente son más para ver que para escuchar; por otra parte, destacar obstinadamente cada aparición de un motivo importante en la interpretación podría producir resultados absurdos, aun cuando el conocimiento del desarrollo motívico de una pieza determinada pueda resultarle útil al intérprete (por ejemplo, para formular la música con respecto a los

timbres o la dinámica). Igualmente, aunque un análisis schenkeriano puede ilustrar elegantemente la complejidad jerárquica de una estructura tonal, sería cuestionable hacer que la interpretación se corresponda e intentar recrear el análisis en sonido, por más valioso que pueda ser el conocimiento de los procesos y relaciones implícitas en ese análisis para estructurar la interpretación.

No obstante, un estudio analítico riguroso puede ayudar a los intérpretes a resolver problemas conceptuales o técnicos (como he señalado antes), así como a memorizar y a combatir el miedo escénico. Los capítulos 8 y 12 tratan respectivamente sobre los dos últimos asuntos, mientras que en este capítulo se demostrará que el análisis es una herramienta para resolver problemas y se incluirán ejemplos de una concepción paralela del análisis con la cual planteo que los intérpretes frecuentemente entienden la música de un modo similar a los que realizan un «análisis riguroso» pero en términos diferentes —un paralelismo que ignoramos por nuestra cuenta y riesgo.

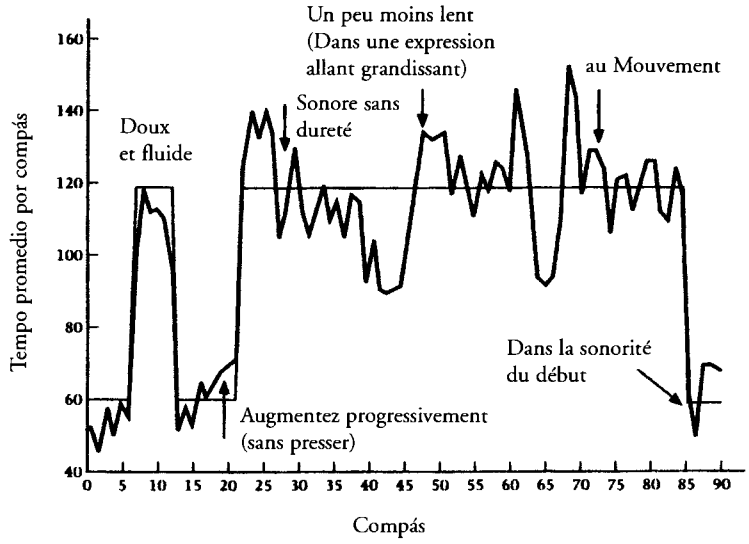



FIGURA 3.1 *(a) «Gráfico del tempo» de la interpretación de Claude Debussy, grabada en un cilindro de pianola (1913), de «La Cathédrale engloutie»; reproducido con autorización de Oxford University Press de José A. Bowen, «Finding the music in musicology: performance history and musical works», en Nicholas Cook y Mark Everist (eds.),* Rethinking Music, *Oxford, Oxford University Press, 1999, p. 440.*

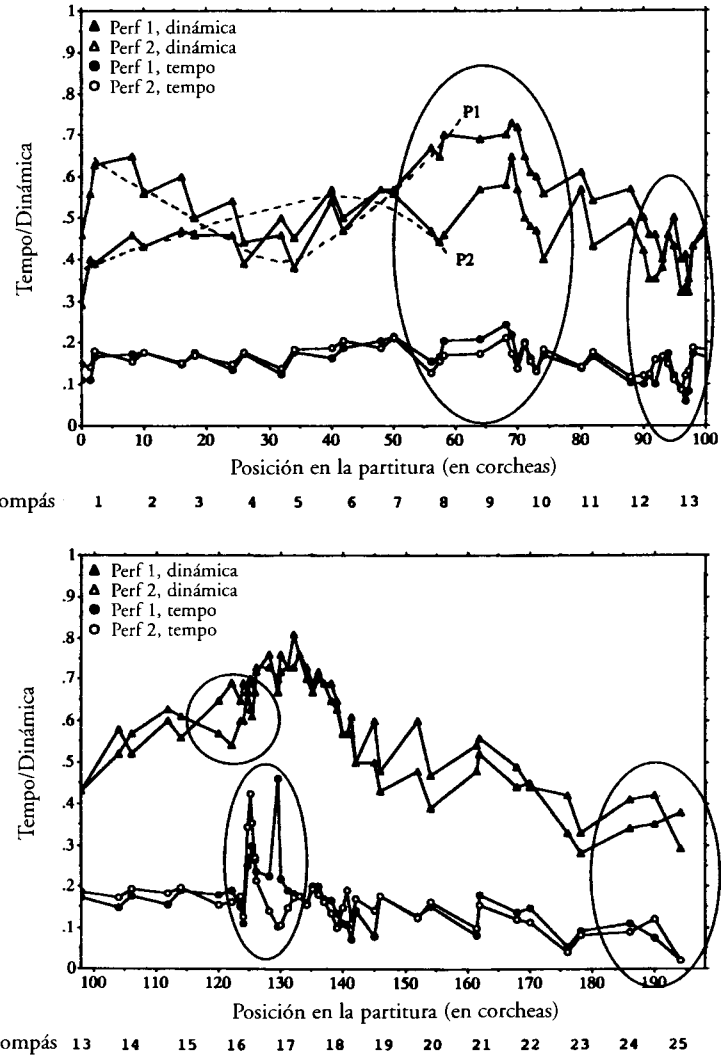



FIGURA 3.1 *(b) Gráficos de las fluctuaciones de dinámica y tempo en dos interpretaciones del Preludio en Mi menor op. 28 núm. 4 de Chopin (sólo la mano derecha; las unidades del eje vertical son arbitrarias pero parten de un cero real), reproducido de Eric Clarke, «Expression in performance generativity, perception and semiosis», en John Rink (ed.),* The Practice of Performance: Studies in Musical Interpretation, *Cambridge, Cambridge University Press, 1995, p. 32.*

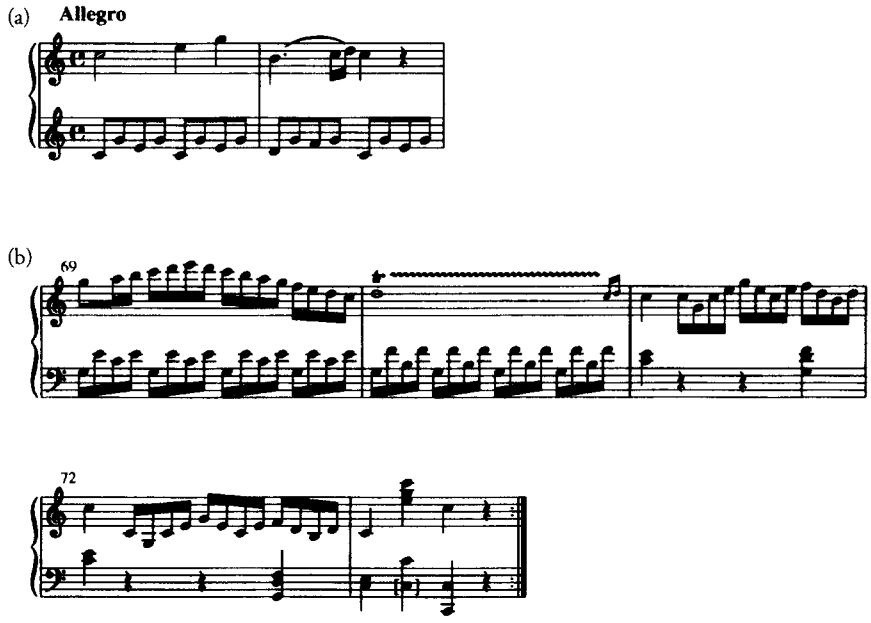
Para comprender ese paralelismo (un tema central de mi estudio de caso), es necesario que consideremos lo que el «análisis del intérprete» puede entrañar. La argumentación anterior se puede resumir en cinco principios:

(1) La temporalidad es inherente a la interpretación y por lo tanto es fundamental en el «análisis del intérprete».
(2) El objetivo principal de dicho análisis es descubrir la «forma» de la música, en oposición a la estructura, y la manera de proyectarla.
(3) La partitura no es «la música»; la música no está confinada a la partitura.
(4) Todo elemento analítico que repercuta en la interpretación será incorporado idealmente dentro de una elaboración más extensa en la que influyen consideraciones de estilo (en sentido amplio), género, tradición interpretativa, técnica, instrumento, etc., así como las capacidades artísticas del intérprete. En otras palabras, las decisiones que resultan del análisis no se deben priorizar sistemáticamente.
(5) La «intuición instruida» guía, o al menos influye, al proceso del «análisis del intérprete», aunque un enfoque analítico más deliberado también puede ser útil.

Hay que destacar que el «análisis del intérprete» se realiza principalmente durante la formulación y posterior reevaluación de una interpretación —es decir, mientras se está practicando, más que interpretando. Ello no excluye su posible influencia en la propia interpretación, ni que puedan surgir nuevos descubrimientos durante ésta. En general, sin embargo, el proceso analítico tiene lugar durante la fase (evolutiva) de diseño, y sus resultados son asimilados por el corpus de conocimientos que subyace en todo acto interpretativo.

Este tipo de análisis se puede ilustrar observando los primeros dos compases de la Sonata para piano en Do mayor K. 545 de Mozart (Ejemplo 3.1a). El pasaje consiste en un acompañamiento en forma de «bajo de Alberti» que apoya a una melodía en la mano derecha: una tríada arpegiada y un final cadencial ornamentado y de registro más limitado. La cadencia es débil: la séptima de dominante en la mano izquierda del compás 2 no está en posición fundamental, y el patrón rítmico de corcheas se mantiene uniforme. Además, la cadencia tiene lugar justo al comienzo de la pieza y carece de finalidad. Independientemente de que el intérprete conciba el pasaje en esos mismos términos, no hay duda de que estará al tanto de esas características y que ellas influirán correspondientemente en su interpretación. Por ejemplo, la mayoría de los pianistas destacarán mucho menos la cadencia en Do mayor del compás 2 que la cadencia que aparece al final del movimiento en los compases 69-71, mucho más definitiva y estructuralmente significativa

EJEMPLO 3.1 *Mozart. Sonata para piano en Do mayor K. 545, primer movimiento, (a) compases 1-2; (b) compases 69-73.*




(Ejemplo 3.1b) [12]. Tocar la cadencia del compás 2 con el mismo peso y la misma determinación sería difícil de justificar (¡aunque un Glenn Gould podría lograrlo!). En resumen, el conocimiento que tenga el pianista de la función contextual de cada cadencia determinará su formulación —el grado de dilatación del tempo (en caso de haberlo, en el compás 2), los niveles dinámicos de los elementos constitutivos en relación a ellos mismos y a la música en su conjunto, el tipo de articulación (pedales, pulsación, timbre), etc.

Se deberán tomar decisiones similares en cuanto a la melodía inicial del compás 1, que la mayoría de los pianistas articularán con un pequeño *crescendo*, destacando sutilmente, y quizá manteniendo, la primera corchea de cada grupo de cuatro en la mano izquierda. La tradición estilística y otros muchos factores influirán naturalmente en sus decisiones, pero en primer lugar el pianista debe reconocer explícitamente la tríada ascendente y arpegiada en la mano derecha y el acompañamiento en forma de «bajo de Alberti» en la izquierda —y cuando digo «explícitamente» no me refiero necesariamente a que deba hacerlo en esos términos. La cuestión es que si no se examinan de-

talladamente estos compases, ni se identifica la función de cada elemento dentro de su contexto inmediato como del contexto general, difícilmente se podrá concebir una interpretación viable.

El examen detallado y la identificación de las funciones —esencia de todo proceso analítico— se pueden efectuar a primera vista, como probablemente lo hará el músico experimentado, sobre la base de la «intuición instruida» mencionada anteriormente (es decir, sirviéndose de sus conocimientos sobre sintaxis, estructura melódica, patrones rítmicos, etc.). Sin embargo, la comprensión inmediata no excluye de ningún modo una reconsideración posterior y más deliberada de la manera en que funcionan diversos elementos en una pieza, y en esa fase es posible que se requiera de algo más que intuición.

Una de las maneras en que los músicos pueden satisfacer su «sed de conocimiento musical» [13] es aprendiendo técnicas más rigurosas, no tanto para aplicarlas posteriormente a la interpretación como para asimilar terminologías y conceptos que puedan mejorar su capacidad para explicarse a sí mismos y a los demás (alumnos, maestros, etc.), lo que está ocurriendo en la música. No es tanto al conocimiento a lo que temen algunos intérpretes como a los posibles *requerimientos* relacionados con él: comprensiblemente, a los intérpretes no les gusta que los eruditos les digan lo que deben hacer empleando un lenguaje dictatorial que coarta su libertad musical. No obstante, ser capaces de emplear un vocabulario más sofisticado, además del «arcano sistema de signos, gestos y gruñidos» [14] utilizado frecuentemente en los ensayos, y de comprender mejor las diversas maneras en que la música puede estar organizada, puede resultar provechoso para los músicos que se esfuerzan por alcanzar una intuición más instruida, un pensamiento consciente más profundo y una mayor capacidad de expresarse verbalmente.

Dichas técnicas incluyen las siguientes:

(1) identificar las divisiones formales y el esquema tonal básico,
(2) representar gráficamente el tempo,
(3) representar gráficamente la dinámica,
(4) analizar la línea melódica y los motivos/ideas que la componen,
(5) elaborar una reducción rítmica,
(6) reescribir la música.

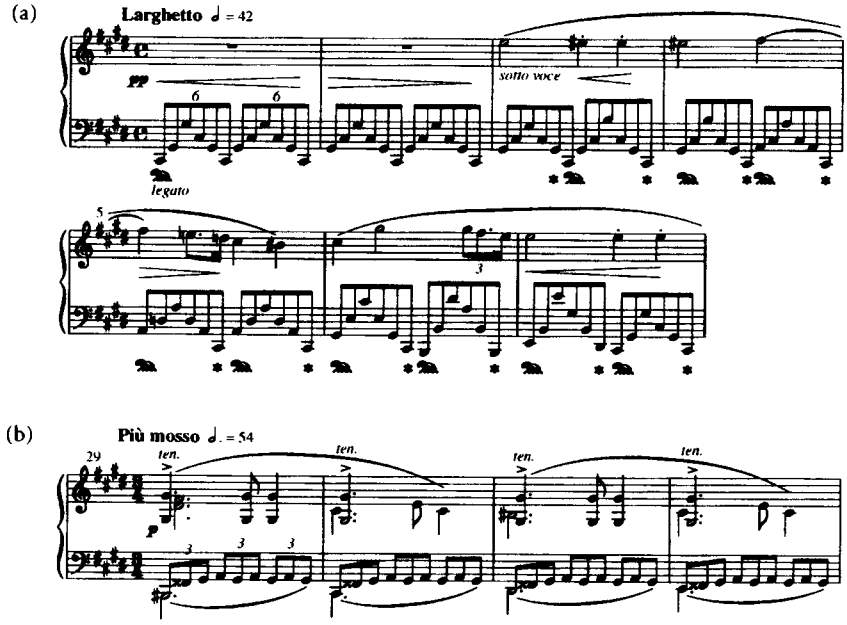
Describiré estas técnicas y el papel que desempeñan en la preparación del Nocturno op. 27 núm. 1 de Chopin, advirtiendo en primer lugar que su aplicación no está limitada a la música para piano solo. En principio, a cualquier músico le podrían resultar útiles estas técnicas, siempre y cuando no las considere como un objetivo en sí mismas, sino como un medio para mejorar su sentido del proceso musical. De nuevo, la temporalidad y la «forma» son

fundamentales para esta tarea analítica, como lo es igualmente reconocer la necesidad de una síntesis más extensa.

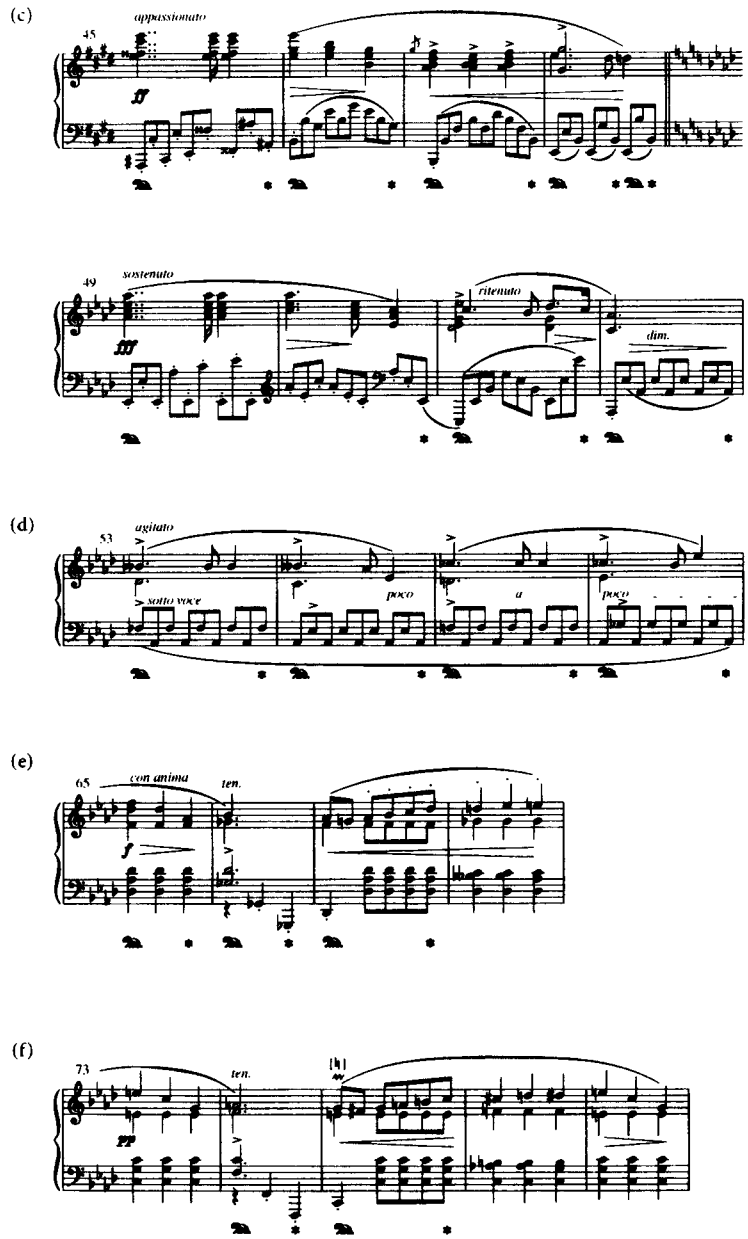
## *«Análisis de intérprete» del Nocturno en Do♯ menor de Chopin*

Chopin compuso los Nocturnos op. 27 en 1835. El comienzo del primero es tenue (Ejemplo 3.2a), pero a la larga la música comienza a tomar cuerpo hasta alcanzar un poderoso clímax, después del cual se retrae tan sólo para volver a crecer (Ejemplo 3.2b-d). Entonces entra una melodía triunfal (Ejemplo 3.2e), seguida de una retractación casi inmediata (Ejemplo 3.2f) y un repentino colapso, culminando con una cadencia catártica que conduce al regreso de la fútil melodía inicial (Ejemplo 3.2g). El ánimo se torna más esperanzador en la coda en Do♯ mayor (Ejemplo 3.2h), que anticipa la lírica tonalidad de Re♭ mayor del segundo nocturno [15].

EJEMPLO 3.2 *Chopin, Nocturno en Do♯ menor op. 27 núm. 1; (a) compases 1-7; (b) compases 29-32; (c) compases 45-52; (d) compases 53-56; (e) compases 65-68; (f) compases 73-77; (g) compases 78-84; (h) compases 94-101.*

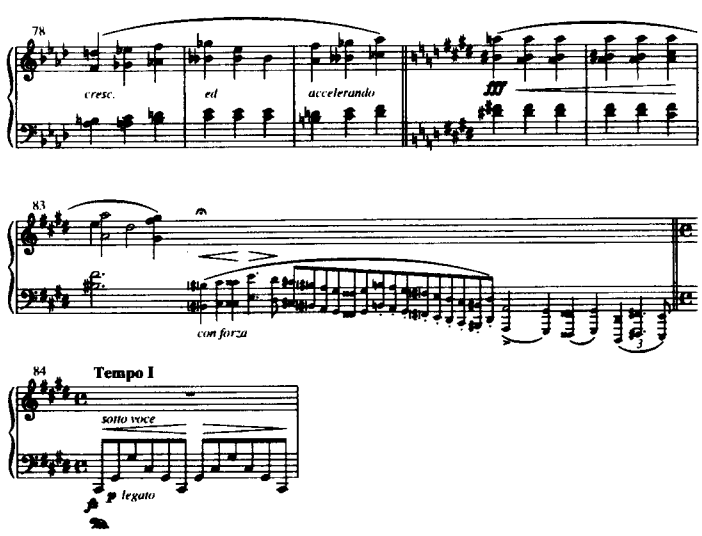


EJEMPLO 3.2 *(Continuación)*

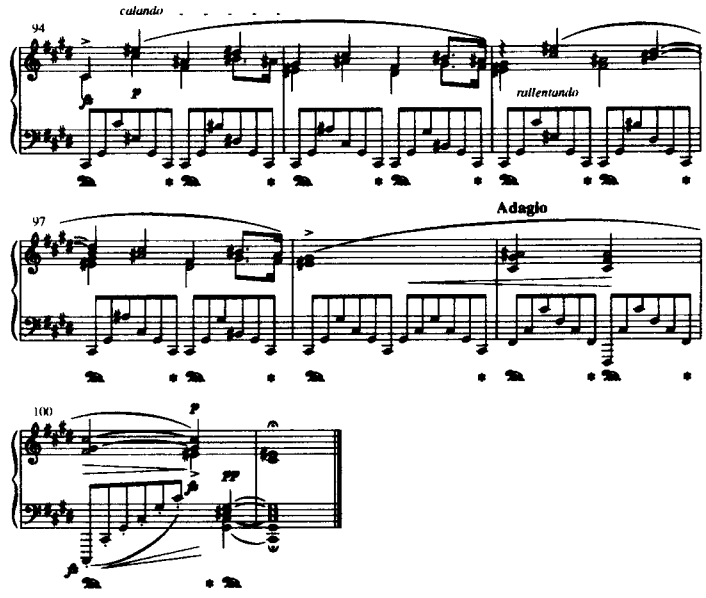


EJEMPLO 3.2 *(Continuación)*

(g)




(h)



El primer reto técnico que se le presenta al pianista es controlar el acompañamiento, *pianissimo* y de registro muy amplio, del principio y la repetición, así como la figuración más vehemente, también en la mano izquierda, que aumenta su intensidad y extensión a medida que se acerca el clímax. Entre los desafíos más sutiles se encuentran el mantener la energía de las secciones «tenues» y controlar el dramatismo de la sección central para poder realizar expansiones graduales y contracciones abruptas. Algunos de los siguientes recursos analíticos pueden contribuir a alcanzar esos objetivos.

### *Identificar las divisiones formales y el esquema tonal básico*

Una de las primeras tareas analíticas «deliberadas» que puede abordar el intérprete es determinar la forma de la música y su base tonal. La mayoría de los músicos saben lo que es una forma binaria o ternaria, la forma sonata, la forma rondó, etc., y puede ser útil analizar una obra en esos términos desde el principio, seguido quizá de una descripción más detallada y personal que destaque las principales secciones y subsecciones, el esquema tonal y otras características. Sin embargo, este tipo de diagrama no tiene más valor musical que un circuito eléctrico impreso, a no ser que uno comprenda sus símbolos arquitectónicos en términos diacrónicos, es decir, en términos de tiempo y proceso (sucesión de hechos en el tiempo). Esto es especialmente importante para el intérprete por las razones expuestas anteriormente.

La estructura del Nocturno de Chopin es relativamente sencilla, como demuestra la Figura 3.2. Está en forma ternaria, A-B-A′ más coda, con una división de la sección B en dos partes, definida por la melodía «con anima» en Re♭ mayor, el equivalente enarmónico de la tónica mayor (véase el Ejemplo 3.2e). La tonalidad de Re♭ mayor se ha alcanzado mediante una modulación desde Do♯ menor (sección A) pasando por Mi mayor a la dominante (escrita como La♭ mayor en el Ejemplo 3.2c), que actúa como un pivote importante en la sección B y la pieza en general. Después de una melodía nueva que aparece en el compás 65, y de la ruptura de los compases 78-80, regresa brevemente la dominante, que se mantiene hasta la repetición abreviada [16] de la sección A.

Una vez más, la manera en que los intérpretes conciben el Nocturno no debe corresponderse necesariamente con esta descripción «rigurosa». Los pianistas tienden en cambio a *sentir* las diversas áreas tonales como centros de gravedad que atraen o repelen el flujo de la música —áreas que serán subrayadas según su importancia, mereciendo menos énfasis (en la dinámica y la extensión temporal) la tonalidad transitoria de Mi mayor, por ejemplo, que la consecutiva dominante, que aparece en el clímax.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Compases | 1-28 | | | | 29-64 | 65-83 | | 84-94 | |
| Sección | A | | | | B | | | A′ | |
| Subsección | intro | $A_1$ | $A_2$ | 'coda' | $B_1$ | $B_2$ | cadenza | intro | $A_2'$ |
| Compás | 1 | 3 | 19 | 27 | 29 | 65 | 83 | 84 | 86 |
| Tonalidad | Do♯ menor | | | | ---Mi—La♭ | Re♭---Sol♯ | | Do♯ menor | |
| | i | | | | ---III–V | = I V | | i | |

FIGURA 3.2 *Chopin. Nocturno en Do♯ menor, op. 27 núm. 1: esquema formal y tonal.*

Igualmente, los pianistas no considerarán la forma A-B-A′ como una secuencia de bloques seccionados, sino como un desarrollo diacrónico —posiblemente como una interacción entre fases estables e inestables o estáticas y dinámicas— o como una exposición seguida de una sección contrastante y un regreso [17]. Aunque ninguna de estas formulaciones reproduce exactamente lo que ocurre en la pieza, al menos representan nuestra sensación básica de la música durante la interpretación cuando se expande y contrae, se mantiene fija o avanza, etc. Un *sentido* de la «forma como proceso» es lo que realmente importa al intérprete, aunque para lograrlo sea necesaria cierta experiencia práctica y no tan sólo una asimilación intuitiva. Los diagramas como el de la Figura 3.2 pueden ayudar en ese respecto a pesar de su aspecto rígido.

### *Representar gráficamente el tempo*

Aunque los gráficos de la Figura 3.1 proporcionan una valiosa imagen de la fluctuación del tempo durante la interpretación, es probable que ni los intérpretes ni los oyentes perciban tales fluctuaciones en un sentido literal: los «hechos» de la interpretación expresados por los datos sólo revelan una verdad parcial y un tanto engañosa [18]. Sin embargo, aunque lo anterior pueda repercutir en el análisis post facto de la interpretación, representar gráficamente la fluctuación del tempo puede resultar provechoso para los intérpretes como un análisis *previo a* la interpretación, es decir, mientras se prepara una interpretación. Una de las maneras de hacerlo es determinar las divisiones de tempo más generales de la obra y después delinear a mano el contorno de los tempos predominantes, teniendo en cuenta las pequeñas aceleraciones y ralentizaciones que ocurren durante una interpretación. Aunque este método a mano alzada carece del rigor de los diagramas de la Figura 3.1, puede

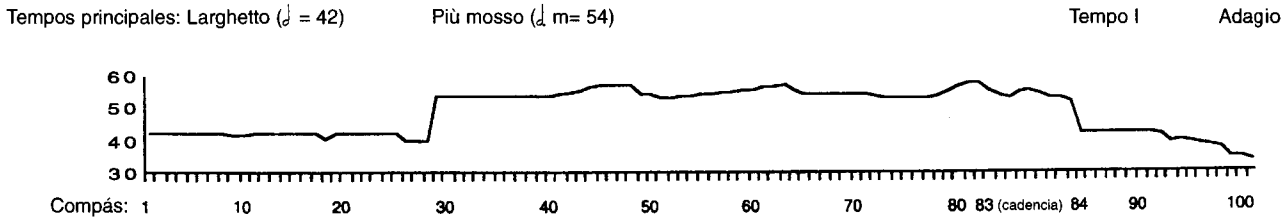



FIGURA 3.3 *Fluctuaciones de tempo en el Nocturno en Do♯ menor op. 27 núm. 1 de Chopin.*

representar más acertadamente la manera en que el intérprete «oye» la música, además de producir una imagen esclarecedora del proceso temporal.

Estos diagramas son mucho más fáciles de trazar cuando el compositor especifica las indicaciones metronómicas, y casualmente los Nocturnos op. 27 fueron las últimas obras en las que Chopin incluyó tales indicaciones. En el op. 27 núm. 1, la sección A está indicada «Larghetto 𝅗𝅥 = 42», y la sección B, «Più mosso 𝅗𝅥 = 54» (siendo la blanca y la blanca con puntillo, respectivamente, las unidades del pulso básico de cada sección). El tempo inicial regresa en la repetición, y los últimos tres compases se tocan Adagio (después de una ralentización gradual en los compases 93-96). Por lo tanto, la pieza puede verse como tres niveles de tempo que se suceden, seguidos del Adagio. Como en el caso anterior, los intérpretes no considerarán estos niveles como entidades, sino como una sucesión continua de pulsos predominantes que están sujetos a ajustes circunstanciales. He representado esa sucesión en la Figura 3.3 con un gráfico que simula los análisis post facto que hemos visto anteriormente, pero trazando el contorno del tempo de la música tal y como transcurre en el oído interior del intérprete. Además de ofrecer una imagen sincrónica, el gráfico también debe leerse de principio a fin, en otras palabras, como progresa el Nocturno. La interacción dialéctica entre el proceso diacrónico y el conjunto sincrónico es, de hecho, característica de la manera en que los intérpretes conciben la música en general —un punto importante al que regresaremos más adelante [19].

### *Representar gráficamente la dinámica*

Podemos mejorar nuestro conocimiento del proceso musical elaborando un gráfico de los niveles dinámicos de la obra en cuestión, de nuevo como en la Figura 3.1 pero con un objetivo potencialmente preceptivo. Para ello trazaremos un diagrama como el de la Figura 3.4, en el cual las principales indicaciones dinámicas aparecen en el eje vertical, mientras que el horizontal indicará la progresión de los compases según las indicaciones dinámicas señaladas por Chopin. Uno de los puntos débiles de este método es su enfoque limitado del contexto, que generalmente influye en el significado de una indicación dinámica. (Por ejemplo, una indicación *piano* puede tener un significado diferente según se encuadre en un *fortissimo* o en un pasaje *pianissimo*.) Sin embargo, aunque las indicaciones dinámicas se interpretan nominalmente, y aunque también carezca del rigor de la Figura 3.1b, el gráfico proporciona una excelente perspectiva general del terreno dinámico, así como la oportunidad de percibirlo a medida que avanza. Esa percepción se puede obtener imaginando que la primera parte de la pieza procede más o

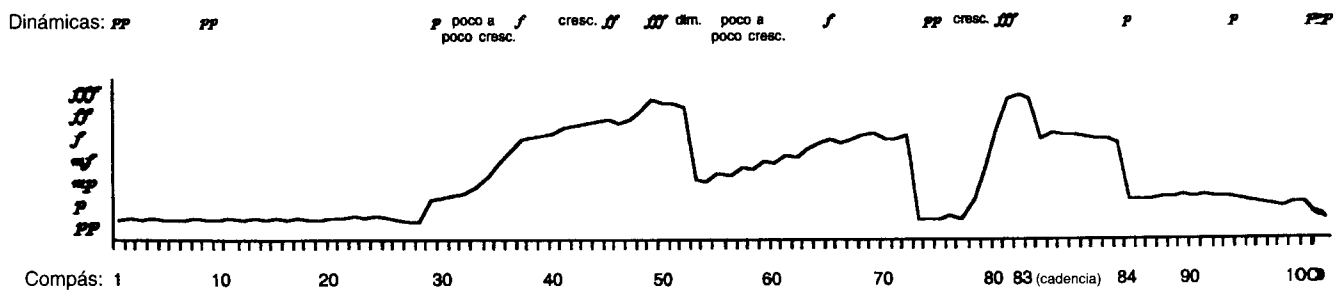



FIGURA 3.4 *Dinámicas del Nocturno en Do♯ menor op. 27 núm. 1 de Chopin.*

menos horizontalmente sobre un nivel básico, seguido de un rápido ascenso y un empinado descenso, después un ascenso más gradual hacia otro descenso abrupto y finalmente una tercera fase de ascenso y descenso que regresa más escabrosamente al nivel horizontal en el que comenzó la pieza. En mi opinión, Chopin vierte toda la energía potencial del Nocturno en la dinámica, o al menos representa ese flujo de energía en miniatura con las indicaciones dinámicas. En cualquier caso, el diagrama de las fluctuaciones dinámicas viene a ser una «curva de intensidad» de toda la pieza —es decir, «una representación gráfica de los altibajos de la música, de su "contorno" en el tiempo» [20]—, y si intentáramos plasmar la manera en que el intérprete siente el Nocturno, probablemente surgiría un diagrama como éste, cuya representación, una vez más, es importante tanto diacrónica como sincrónicamente.

### *Analizar la línea melódica y los motivos/ideas que la componen*

Aunque resulta difícil idear una metodología general para el análisis melódico, puede ser instructivo trazar contornos melódicos en un papel sin las figuras de las notas, sus plicas, cabezas y otros elementos que distraen nuestra vista de la subyacente trayectoria de los tonos. En general, como se ha sugerido antes, la representación de la música en el oído interior del intérprete se asemeja poco a su apariencia en la partitura: a medida que se desarrolla la interpretación, la notación se transforma en una serie de acciones físicas correspondientes a imágenes mentales de diversos tipos, incluyendo las auditivas (analizadas en el capítulo 4). Los intérpretes suelen concebir la melodía como una línea (continua o no) que cantan para sí mismos mientras hacen música, y un gráfico literal del contorno melódico, que abarque pasajes breves o extensos (que en este último caso probablemente representen sucesivas alturas), se puede parecer más a esa imagen auditiva que la notación original. La Figura 3.5 traza el contorno melódico de la cadencia del Nocturno, cuya ondulación y propiedades gestuales se hacen mucho más evidentes cuando se destaca de esta manera [21].

El conocimiento de los patrones importantes dentro de una melodía, así como de los más generales, también influirá en la manera en que los intérpretes proyectan la música, aunque su objetivo principal no es demostrar la unidad motívica. Dichos patrones pueden sugerir un timbre o una articulación característica, así como texturas, ritmos y otros elementos recurrentes que influyen en la forma de la música. En el Nocturno de Chopin, las secciones A y A′ —caracterizadas por un registro reducido y un movimiento melódico predominantemente conjunto y monódico— presentan una evocadora contramelodía en los compases 20-26 y 89-93 (en forma abreviada) que

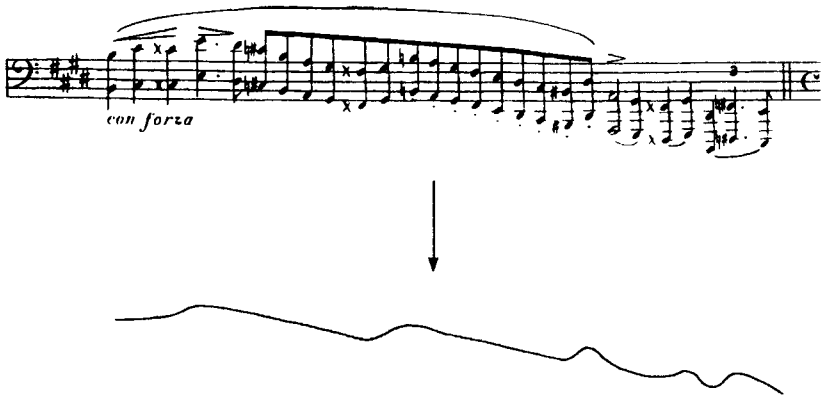


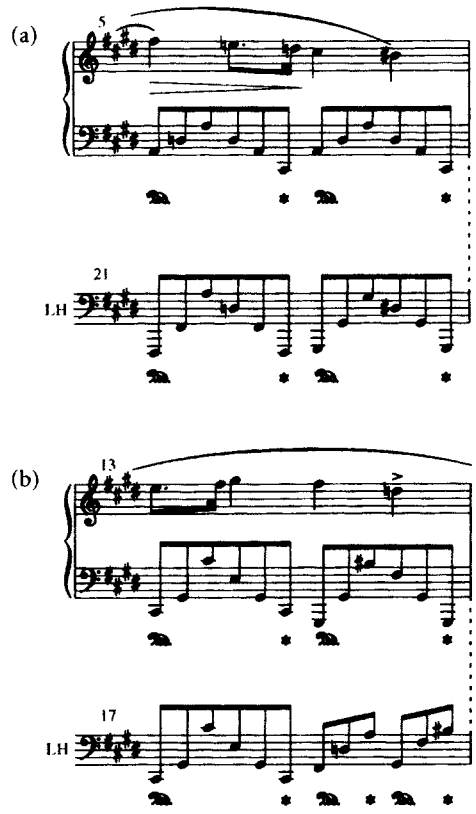

FIGURA 3.5 *Chopin. Nocturno en Do♯ menor op. 27 núm. 1: contorno melódico de la cadencia (partitura reproducida de la primera edición alemana: Breitkopf & Härtel, 1836).*

puede tocarse «nostálgicamente», es decir, con una pulsación pesada, aunque separada, y con una dinámica menos reticente que la tenue melodía a la que hace contrapunto, especialmente en los dos pasajes cadenciales (compases 25-26 y 92-93), el primero de los cuales se deja en suspenso, mientras que el segundo concluye el bloque principal de la pieza con su expresiva indicación «con duolo».

El poder expresivo de esas cadencias se debe en parte al sugerente Re♮ dentro del descenso melódico, que se repite a lo largo de A y A′ en dos disposiciones armónicas diferentes. La primera (Ejemplo 3.3a) emplea un Re mayor napolitano por encima de un insistente pedal de tónica, seguido de una cadencia. La segunda (Ejemplo 3.3b) contrapone la armonía de dominante al Re♮, que esta vez constituye un sustituto disonante del esperado Re♯ diatónico.

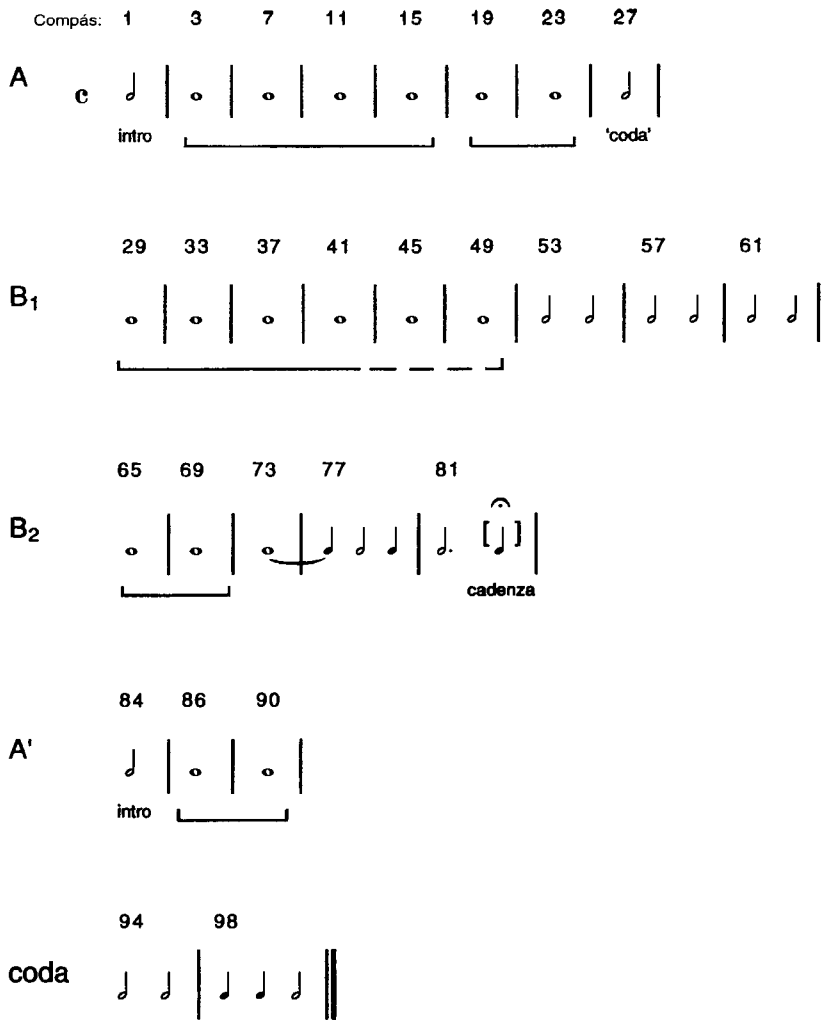
De nuevo, la terminología no es lo que importa: se trata de un técnica analítica pertinente para el intérprete —que en este caso implica extraer todas las apariciones del Re♮ de la partitura, yuxtaponerlas en papel para compararlas y después considerar sus respectivos roles dentro de la concepción de la interpretación. En mi opinión, los Re♮ proporcionan una clave para comprender las secciones A y A′; su proximidad a la tónica sugiere de alguna manera futilidad (particularmente cuando está armonizado con el acorde napolitano), mientras que la disonancia modal de los compases 13 y 17 parece arcaica y elegíaca, contrastando claramente con la belicosidad y el triunfo de la sección central.

EJEMPLO 3.3 *Chopin, Nocturno en Do♯ menor op. 27 núm. 1: Re♮ melódico en las secciones A y A′; (a) compases 5 y 21 (véanse también los compases 9-10, 25, 26-28, 88, 92 y 93); (b) compases 13 y 17.*




*Elaborar una reducción rítmica*

Cada sección del Nocturno posee un carácter rítmico particular, y si analizamos e identificamos la función de los principales elementos rítmicos descubriremos el papel que desempeñan en la generación, mantenimiento o suspensión del impulso rítmico. Por ejemplo, el invariable acompañamiento de corcheas de las secciones A y A′ y la coda produce la cualidad estática mencionada anteriormente; la figura con puntillo (♩. ♪ ♩) utilizada prácticamente




𝅗𝅥 = un compás en la notación

**c** indica un hipercompás de cuatro compases

FIGURA 3.6 *Chopin, Nocturno en Do♯ menor op. 27 núm. 1: reducción rítmica de la estructura de las frases.*

en toda la sección $B_1$ conduce la música hacia un clímax en dos fases; y las figuras de tres negras repetidas y de una negra y cuatro corcheas de la sección $B_2$ (𝅗𝅥 𝅗𝅥 𝅗𝅥 y ♩ ♫ respectivamente) ayudan a crear la sensación de triunfo pasajero.

Sin embargo, la reducción rítmica referida anteriormente no está relacionada con detalles rítmicos como ésos, sino con propiedades rítmicas en un nivel superior, a saber, la estructura de la frase. Los escritos sobre análisis describen diversas técnicas de reducción rítmica [22], aunque la más valiosa y accesible desde el punto de vista del intérprete podría ser representar cada compás (u otra unidad) como un valor rítmico proporcionalmente menor, como una negra, y combinar esos ritmos de acuerdo con las frases que se perciben en la música. El siguiente paso es discernir la forma dentro del diagrama resultante, es decir, las instancias de truncamiento, extensión, contracción, elisión, etc., en el nivel de la frase o superior. Esta técnica reductora funciona especialmente bien con Chopin, cuya predilección por las frases de cuatro compases es bien conocida, aunque la estructura de las frases de nuestro Nocturno no carece de impulso. En la Figura 3.6, cada redonda representa una frase de cuatro compases (o «hipercompás») [23], mientras que las blancas representan unidades de dos compases, y las negras, un compás. El diagrama revela que la sección A está formada por un patrón regular de frases de cuatro compases, exceptuando la introducción y la coda (de dos compases cada una), y, al igual que en la sección A', esa regularidad contribuye al carácter deliberadamente estático mencionado anteriormente. En la sección $B_1$, que comienza con una disposición de 4 + 2 destacando los compases culminantes 45-52, el ritmo de la frase se acelera en los compases 53-64 con seis unidades de dos compases en rápida sucesión. Ello forma una extensa anacrusa a la sección $B_2$, en la que el patrón de frases de cuatro compases, inicialmente equilibrado, se interrumpe en el compás 77, y las irregularidades resultantes en la estructura de la frase provocan el colapso en ese momento crítico. Aunque los intérpretes podrían intuir todas estas características, éstos comprenden un nivel estructural superior que puede resultar difícil incluso de sentir. Un diagrama como el de la Figura 3.6 ayuda a mostrar que en este Nocturno de Chopin la estructura de la frase y la forma de la música son dos caras de la misma moneda.

### *Reescribir la música*

La notación convencional no capta toda la complejidad de la música, como todo intérprete sabe. Las sutilezas rítmicas son especialmente difíciles de plasmar, al igual que los matices dinámicos, las pequeñas variaciones de altu-

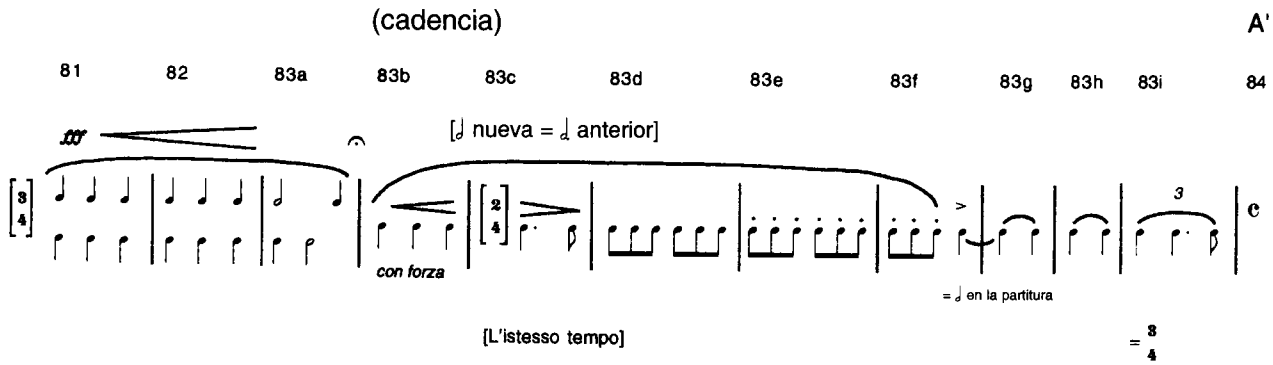



FIGURA 3.7 *Chopin, Nocturno en Do♯ menor op. 27 núm.1: reescritura de los compases 81-84 (con las dinámicas y ligaduras de Chopin).*

ra y la amplia gama de articulaciones que emplean los músicos. Reescribir la música puede reducir las insuficiencias de la notación original al arrojar luz sobre ciertas características ausentes o confusas en la partitura. Por ejemplo, podemos descubrir un esquema métrico alternativo, latente en el original, reescribiendo las partes relevantes de la música en el nuevo compás (o compases) y reagrupando las barras de compás cuando sea necesario para mover los tiempos fuertes y patrones de acentos que de otro modo estarían ocultos [24]. Ello no significa que la notación nueva capte totalmente «la música»: de hecho, la diferencia entre la versión original y la nueva es posiblemente lo más interesante para el intérprete, quien ha de reconocer que «la música» no se ajusta a ninguna de las dos, sino que existe, por así decirlo, en los intersticios entre ambas.

La reescritura también permite representar características como la ejecución rítmica a un nivel inferior, los estratos contrapuntísticos que pueden estar especialmente ocultos en la música para teclado, la distribución de las voces entre ambas manos o entre los miembros de una agrupación, etc.

La Figura 3.7 presenta una reescritura de los compases 81-84 del Nocturno de Chopin, elaborando particularmente las características métricas de la cadencia. Este diagrama surgió como respuesta a mi incertidumbre sobre cómo dar forma a la cadencia en una interpretación real: es producto de un análisis deliberado que emprendí porque la mera intuición era incapaz de resolver un problema muy difícil. El hecho de que el pasaje no esté dividido en compases en el original dificultó la ubicación de los acentos métricos, así como la organización rítmica interna del grupo de quince corcheas (véase Ejemplo 3.2g), factores que influyen en la dinámica, el *rubato* y la articulación y por lo tanto pueden extender o comprimir el contorno melódico de la Figura 3.5. En mi análisis, que desempeñó el referido papel de recurso para solucionar problemas, el compás 83b mantiene el 3/4 de los compases 81-83a, pero en el compás 83c se sugiere un 2/4, manteniendo constante la duración de los compases al equiparar la blanca con la blanca con puntillo (de modo que, retrospectivamente, el compás 83b parece un tresillo de negras). El grupo de quince corcheas se divide en los cuatro tresillos de los compases 83d-e, seguidos de un último tresillo en el compás 83f que impulsa la música hacia la octava sincopada sobre La (escrita como blanca en la partitura). A continuación, un ritmo más lento y regular conduce al tresillo de negras del compás 83i (el 3 que aparece por encima de la ligadura es de Chopin), que implícitamente reestablece el compás de 3/4 de los compases 81-83b al tiempo que prepara el regreso del ¢ (compás común) de la sección A′.

Resulta interesante considerar si Chopin tenía en mente este tipo de esquema organizativo cuando compuso el Nocturno. Indudablemente, sus ligaduras e indicaciones de dinámica y fraseo sugieren esa posibilidad, como

quizá también la estructura de las frases de ocho compases, que coincide con su práctica habitual. Sin embargo, si los pianistas «deben» o no tener en mente este tipo de esquema al interpretar el Nocturno es otro asunto. Crear un diagrama como el de la Figura 3.7 tampoco implica que haya que atenerse a él durante la interpretación; de hecho, yo toco esta música mucho más libremente de lo que sugiere esta nueva notación, por más reveladora que haya sido en un momento dado [25]. Los frutos de este tipo de ejercicios deben ser asimilados como cualquier otro conocimiento adquirido durante el proceso de elaboración de una interpretación —tal vez para que a partir de entonces los olvidemos y, sin embargo, los recordemos inconscientemente cada vez que la música se hace realidad.

## *Hacia una síntesis*

He subrayado a lo largo de este capítulo que los conocimientos adquiridos mediante el análisis —bien sea intuitivo o deliberado— no son más que un factor que influye en la concepción de la música por parte del intérprete. Tanto el intérprete como el público valorará el éxito de una interpretación, no tanto por su rigor analítico, su fidelidad histórica o incluso su exactitud técnica (al menos en ciertos círculos) como por la medida en que logra una «resonancia» al unir los elementos constitutivos en algo mayor que la suma de esas partes, en una síntesis musicalmente convincente y coherente. El análisis puede estar «implícito en lo que hace el intérprete», y también puede ser abordado explícitamente empleando estas u otras técnicas. Sin embargo, es importante no elevar el análisis por encima de la interpretación que produce, o utilizarla como una manera de subyugar y encadenar a los músicos. En cambio, debemos reconocer su utilidad al igual que sus limitaciones, con lo cual quiero decir sencillamente que «la música» trasciende el análisis y cualquier otro intento de comprenderla. Proyectar «la música» es lo más importante, y todo lo demás no son sino medios para ese fin.

## *Notas*

1. «La interpretación de una pieza musical es [...] la actualización de un acto analítico —aunque dicho acto haya sido intuitivo y asistemático— [...] el análisis está implícito en lo que hace el intérprete», Leonard Meyer, *Explaining Music,* Chicago, University of Chicago Press, 1973, p. 29.
2. «Los intérpretes no pueden sondear la profundidad estética de una gran obra sin examinar detalladamente sus elementos paramétricos.» Eugene Narmour, «On the relationship of analytical theory to performance and interpretation», en Eugene Narmour y Ruth A. Solie (eds.), *Explorations in Music, the Arts, and Ideas,* Stuyvesant, Pendragon, 1988, p. 340.

3. Cuando hablo de «intérpretes», me refiero principalmente a solistas, en oposición a los músicos de orquesta o agrupaciones, aunque las técnicas analíticas propuestas aquí pueden utilizarse para preparar interpretaciones de dos o más músicos. También pueden aplicarse a otros repertorios, a pesar de que este capítulo se centra en la música tonal de la tradición clásica occidental.
4. Janet Schmalfeldt, «On the relation of analysis to performance: Beethoven's Bagatelles Op. 126, núms. 2 and 5», *Journal of Music Theory* 29 (1985), p. 28; subrayado en el original.
5. Edward T. Cone, *Musical Form and Musical Performance,* Nueva York, Norton, 1968, p. 34.
6. Eugene Narmour, *op. cit.*, pp. 319 y 340; véase también el comentario de Narmour en la nota 2.
7. Respectivamente, *Form in Music,* 2ª ed., Englewood Cliffs, Prentice-Hall, 1986, p. 415, y *Musical Structure and Performance,* New Haven, Yale University Press, 1989, p. 44.
8. Jonathan Dunsby, «Guest editorial: performance and analysis of music», *Music Analysis* 8 (1989), pp. 7, 9 y ss. Cfr. la p. 271 de este libro, donde Dunsby comenta que «el análisis estructural en alguna de sus formas simplemente debe ser parte del trabajo del intérprete»; «el intérprete debe identificar la estructura, asimilarla y controlarla suficientemente bien en tiempo real para representarla e, idealmente, transmitirla con precisión».
9. Hay que reconocer que los términos «análisis riguroso» y «análisis del intérprete» son rudimentarios, aunque la distinción resultará útil. Entre las técnicas de «análisis riguroso» se incluyen el análisis schenkeriano, la teoría de conjuntos y el análisis paradigmático.
10. John Rink, reseña de Berry, *Musical Structure,* en *Music Analysis* 9 (1990), pp. 323, 324 y 328. Para un resumen más detallado de la bibliografía sobre análisis e interpretación, véase Nicholas Cook, «Analysing performance, and performing analysis», en Nicholas Cook y Mark Everist (eds.), *Rethinking Music,* Oxford, Oxford University Press, 1999, pp. 239-261.
11. Véanse también los capítulos 11 y 14 de este libro (especialmente las Figuras 11.1-11.3 y 14.1-14.3). Se utilizaron técnicas bastante diferentes para elaborar los diagramas de la Figura 3.1. Para crear un gráfico de las fluctuaciones de tempo en una grabación en CD, se puede utilizar un ordenador que tenga lector de CD-ROM, pulsando una tecla para marcar el comienzo de cada pulso o compás; después el ordenador une las marcas sucesivas con líneas, como en la Figura 3.1a. Otra alternativa son los programas informáticos como «Sforzando», que permiten trazar las fluctuaciones de tempo y dinámica en las interpretaciones grabadas. En cuanto a las interpretaciones pianísticas en vivo, la información digital obtenida mediante un Yamaha Disklavier define las propiedades expresivas (i.e., duración, dinámica y articulación) de cada nota, que puede representarse gráficamente como en la Figura 3.1b (véase también las figuras 11.2 y 11.3). Obtener dicha información digital de la voz y otros instrumentos es más difícil, lo cual explica en parte por qué predominan en estas publicaciones las investigaciones sobre música para piano.
12. La breve coda de los compases 71-73 refuerza esa cadencia.
13. Jonathan Dunsby, *Performing Music: Shared Concerns,* Oxford, Clarendon Press, 1995, p. 46.
14. Joseph Kerman, *Musicology,* Londres, Fontana, 1985, p. 196.
15. Recomiendo a los lectores que consulten una edición Urtext y escuchen la mayor cantidad de grabaciones posibles.
16. Tanto la sección $A_2$ como la $A_2$' carecen de un equivalente a los compases 11-18.
17. La nueva melodía de $B_2$ puede verse como otra exposición, y la coda como un regreso de al menos la tónica mayor de $B_2$, si no de su carácter triunfal.
18. Véase Bruno H. Repp, «Probing the cognitive representation of musical time: structural constraints on the perception of timing perturbations», *Cognition* 44 (1992), pp. 241-281; véase también el capítulo 13 de este libro.

19. Véase Edward T. Cone, *op. cit.*, pp. 88-98.
20. Una curva de intensidad está «determinada por todos los elementos activos (atmonía, melodía, ritmo, dinámica, etc.) actuando independiente, sincrónica o desfasadamente entre sí para crear los diversos niveles de energía y por tanto la forma total»; John Rink, «Translating musical meaning: the nineteenth-century performer as narrator», en Nicholas Cook y Mark Everist (eds.), *op. cit.*, p. 234.
21. Cfr. los gráficos en *ibid.*, pp. 231-232.
22. Por ejemplo, véanse Carl Schachter, *Unfoldings: Essays in Schenkerian Theory and Analysis,* Joseph N. Strauss (ed.), Nueva York y Oxford, Oxford University Press, 1999, y William Rothstein, *Phrase Rhythm in Tonal Music,* Nueva York, Schirmer Books, 1989.
23. Para un exposición detallada, véase William Rothstein, *op. cit.* Nótese el hipercompás C en la Figura 3.6, que indica que los hipercompases de cuatro compases son la norma.
24. Para ejemplos, véanse John Rink, «Playing in time: rhythm, metre and tempo in Brahms's *Fantasien* Op. 116», en John Rink (ed.), *The Practice of Performance: Studies in Musical Interpretation,* Cambridge, Cambridge University Press, 1995, pp. 274-275; y John Rink, «Opposition and integration in the piano music», en Michael Musgrave (ed.), *The Cambridge Companion to Brahms,* Cambridge, Cambridge University Press, 1999, p. 93.
25. Resulta instructivo comparar las interpretaciones de otros intérpretes (en vivo o grabadas). Por ejemplo, Alfred Cortot (EMI CDZ 7673652) formula la cadencia más o menos como en la Figura 3.7 —¡presumiblemente sin haber trazado primero un diagrama similar!

## *Lecturas complementarias*

Cone, Edward T.: *Musical Form and Musical Performance,* Nueva York, Norton, 1968.

Cook, Nicholas: «Analysing performance, and performing analysis», en Nicholas Cook y Mark Everist (eds.), *Rethinking Music,* Oxford, Oxford University Press, 1999, pp. 239-261.

Dunsby, Jonathan: «Guest editorial: performance and analysis of music», *Music Analysis* 8 (1989), pp. 5-20.

Lester, Joel: «Performance and analysis: interaction and interpretation», en John Rink (ed.), *The Practice of Performance: Studies in Musical Interpretation,* Cambridge, Cambridge University Press, 1995, pp. 197-216.

Rink, John: artículo sobre Wallace Berry, *Musical Structure and Performance,* en *Music Analysis* 9 (1990), pp. 319-339.

— «Playing in time: rhythm, metre and tempo in Brahms's *Fantasien* Op. 116», en John Rink (ed.), *The Practice of Performance,* cit., pp. 254-282.

Rothstein, William: «Analysis and the act of performance», en John Rink (ed.), *The Practice of Performance,* cit., pp. 217-240.

# 4 *Comprender la psicología de la interpretación*

ERIC CLARKE

## *Introducción: qué hacen los intérpretes*

La interpretación musical en su nivel más alto requiere una extraordinaria combinación de habilidades físicas y mentales. No es inusual para un pianista concertista tocar a velocidades de diez o más notas por segundo en ambas manos simultáneamente, distribuidas sobre el teclado en patrones espaciales complejos que cambian constantemente y con diversos patrones de ritmo, dinámica y articulación. Además, el intérprete debe conocer y comprender la estructura inmediata y general de la música, contar con una «estrategia» expresiva para realizarla y tener resistencia para enfrentarse a los requerimientos físicos y las tensiones psicológicas que plantean la interpretación en público. Habilidades como ésas no se desarrollan de la noche a la mañana, y para el momento en que los mejores intérpretes llegan a la edad de veintiún años generalmente han pasado más de 10.000 horas practicando su instrumento [1], bastante más que el tiempo dedicado a otros aspectos de la educación musical formal y a los componentes más informales de lo que puede llamarse «culturización musical». Incluso observaciones tan sencillas como éstas revelan que la interpretación musical representa un logro humano asombroso y es el resultado de una enorme inversión de tiempo y esfuerzo.

¿Qué es, entonces, lo que hacen los intérpretes? En un nivel, la respuesta a esta pregunta es obvia: producen realizaciones físicas de ideas musicales, independientemente de que dichas «ideas» hayan sido registradas por escrito, transmitidas oralmente (como en las culturas analfabetas) o inventadas en el momento (como en la improvisación libre). El requerimiento más básico es

que el intérprete debe producir las notas, ritmos, dinámicas, etc. (más o menos) correctos de una idea musical —en caso de que exista una referencia (escrita o conceptual) con respecto a la cual se puede medir la «fidelidad». Sin embargo, además de eso, se espera que los músicos den vida a la música, que vayan *más allá* de lo que proporciona explícitamente la notación o la transmisión oral —que sean «expresivos». Por lo tanto, no es de extrañar que una parte sustancial de la investigación sobre la psicología de la interpretación haya estado de alguna manera centrada en cuestiones de interpretación y expresión.

Este capítulo tratará sobre tres aspectos de la psicología de la interpretación: las habilidades físicas y mentales que intervienen en la interpretación, la expresión en la interpretación y el cuerpo durante la interpretación. Los tres temas serán considerados principalmente desde la perspectiva de la tradición concertística occidental, ya que hasta ahora se han hecho muy pocas investigaciones acerca de la psicología de otros tipos de interpretación —aunque posiblemente muchas de las habilidades y características que se pueden identificar y estudiar dentro de la tradición occidental también se emplean en otras tradiciones interpretativas, permitiendo al menos cierta extensión hacia otros contextos culturales. También existe una amplia literatura etnomusicológica sobre diversas tradiciones interpretativas, parte de la cual hace referencia directa a la teoría psicológica [2].

## *Habilidades interpretativas*

Como habilidad física, la interpretación musical ha despertado el interés de los psicólogos que estudian complejas habilidades relativas a los movimientos y el tempo. Los orígenes de ese interés se remontan a la obra de Carl Seashore [3], quien fue el primero en desarrollar técnicas para grabar las características detalladas del tempo y la dinámica de las interpretaciones. Seashore se dedicó principalmente a la precisión y reproductibilidad de la interpretación y a analizar la producción y control de sonidos en interpretaciones instrumentales, excluyendo los teclados. Henry Shaffer [4] y Caroline Palmer [5] desarrollaron esos objetivos de diversas maneras. Por ejemplo, Shaffer se centró en tres asuntos principales: el control del movimiento, los mecanismos del tempo en la interpretación y la coordinación e independencia de ambas manos en la interpretación pianística a solo y entre ambos ejecutantes en las interpretaciones a dúo. Su obra explora la manera en que la imagen mental de una pieza se traduce en los movimientos de la interpretación, y sostiene que el control del movimiento es bastante abstracto y conceptual hasta una fase relativamente tardía del proceso. Expone, por ejemplo, el caso de un intér-

prete que, leyendo a primera vista una fuga de Bach, no se fijó en un cambio de clave entre una página y la siguiente, de modo que produjo una serie de errores que, sin embargo, mantuvieron la armonía subyacente en el pasaje y no fueron simplemente las notas que habrían resultado de leer erróneamente la clave [6]. En cuanto el intérprete se dio cuenta de que algo iba mal, improvisó con un ritmo y una armonía apropiados mientras corregía el error. La presencia de información sobre la estructura musical dentro del programa mental que controla los movimientos es lo que permite una improvisación de emergencia como ésa.

Una de las maneras en que el control abstracto del movimiento ha de concretarse es en términos del tempo. El tempo es particularmente importante en la interpretación musical, ya que es un aspecto importantísimo de la estructura musical y también un medio de expresión. Shaffer y otros [7] han demostrado que los intérpretes expertos tienen un extraordinario control del tempo de la interpretación a niveles que abarcan desde una nota aislada hasta secciones o piezas enteras. Los intérpretes tienen la capacidad de juzgar y establecer el tempo absoluto de una interpretación con diversos grados de certeza y reproductibilidad, y se cree que algunos de ellos poseen un tempo equivalente al oído absoluto (es decir, la habilidad de establecer un tempo metronómicamente exacto sin necesidad de ninguna referencia externa). No existe una explicación generalmente aceptada del sentido del tempo, aunque se han sugerido varios tipos de «reloj interno» u oscilador interno [8]. Se ha propuesto que la formulación temporal a un nivel superior a la nota o el pulso es controlada por la representación (o imagen) mental que tiene el intérprete de la música. La estabilidad (o no) del tempo a un nivel superior puede por tanto atribuirse directamente a la estabilidad de la representación de la música que tiene el intérprete: los intérpretes que poseen una concepción clara y definida de una pieza musical, así como las habilidades técnicas necesarias, probablemente están mejor capacitados para tocarla de manera controlada y reproductible.

Una característica obvia de la interpretación de teclados es la coordinación entre ambas manos —y también el tema análogo de la coordinación entre los intérpretes de una agrupación. Shaffer demostró la considerable independencia entre ambas manos que pueden alcanzar los pianistas, bien sea porque lo exigen los polirritmos de la música, bien por razones expresivas [9]. También en este caso la independencia y coordinación de ambas manos se derivan de la estructura de la imagen mental. Es imposible lograr la independencia necesaria para tocar incluso un simple polirritmo flexiblemente si no se tienen imágenes o representaciones separadas de las partes de cada mano. Igualmente, la coordinación entre los dos pianistas de un dúo no depende (o tal vez no debería depender) de que uno de ellos siga rigurosamente al otro:

el análisis que hizo Shaffer de la interpretación de dos pianistas expertos tocando obras para cuatro manos de Beethoven indicó que implica un importante elemento de predicción por parte de ambos pianistas en cuanto al transcurso ulterior de la interpretación expresiva de cada uno [10]. Ello se puede explicar mediante la suposición de que los dos pianistas compartían una representación de la música y la utilizaban como punto de referencia a partir del cual podían anticipar mutuamente sus formulaciones expresivas. A pesar de la evidente importancia de escuchar y estar conscientes de la respiración, los movimientos corporales, los gestos faciales, etc., de cada uno como canales de comunicación, el papel de las respuestas auditivas y visuales en este proceso prácticamente no se ha investigado [11].

La lectura y la memorización son dos habilidades esenciales dentro de la tradición concertística occidental. A mediados de la década de 1970, Sloboda llevó a cabo un importante programa de investigación sobre la psicología de la lectura musical [12] para el que utilizó el «trayecto vista-mano» como indicador principal —una medida de la antelación con la que un músico procesa la notación durante la lectura. En sus experimentos, Sloboda midió dicha antelación pidiendo a los pianistas que siguieran tocando después de que se les retirara la partitura, repentinamente y sin previo aviso, utilizando cualquier imagen de la notación que perdurase en su mente. Sloboda descubrió que los lectores expertos tenían un «trayecto vista-mano» de unas siete notas, y que en los lectores más expertos esa medida se expandía y contraía coincidiendo con los límites de las frases de la música, lo que sugería que un factor importante en la lectura musical experta es la conciencia de la estructura musical. Los errores en la notación musical también demostraron la existencia de esa conciencia estructural. Sloboda demostró que los lectores expertos generalmente no se dan cuenta de los errores de imprenta aislados y tocan inconscientemente la nota «correcta». En efecto, en lecturas sucesivas de una pieza con errores de imprenta deliberados el número total de errores disminuía (como era de esperar), mientras que el número de «correcciones» inconscientes *aumentaba,* probablemente porque los lectores se familiarizaban con el material y hacían más suposiciones sobre su estructura. Cabe destacar que la lectura musical experta no es simplemente una habilidad visual: un estudio [13] ha demostrado que la capacidad de leer a primera vista está relacionada con la capacidad de hacer coincidir la notación con el sonido —en otras palabras, la habilidad de escuchar la música en el «oído interior» a partir de la notación.

Los estudios sistemáticos sobre el proceso de memorizar música para la interpretación a partir de la notación son sorprendentemente escasos [14], y se sabe relativamente poco acerca de los procesos que tienen lugar durante la memorización y la diversidad de estrategias que pueden utilizar los intérpre-

tes. La habilidad para memorizar música para la interpretación posee un valor cultural particular, aunque algunos podrían argumentar que hay tanto que decir en su contra como a su favor. En un estudio que evalúa el efecto de la memorización en las percepciones del público, Aaron Williamon halló ciertas pruebas de que mejora la comunicación musical [15], aunque sólo para los espectadores/oyentes con formación musical. Otro estudio [16] demostró que el enfoque visual es más eficaz (en cuanto al tiempo que se tarda en memorizar) que el auditivo, que, a su vez, es más eficaz que el enfoque cinestésico (la «memoria muscular»); también demostró que los que tienen más habilidad para memorizar son mejores leyendo a primera vista. Un estudio ulterior investigó el caso de un hombre autista (N. P.) que tenía una memoria musical auditiva excepcional [17] y demostró que a pesar de no haber visto nunca la partitura, este individuo era capaz de recordar y reproducir una pieza para piano de Grieg, de sesenta compases, prácticamente a la perfección después de haberla escuchado tan sólo cuatro veces, mientras que un pianista igualmente experimentado que participó como «factor constante» sólo pudo memorizar un fragmento de la pieza después de la misma cantidad de audiciones. Esta proeza evidentemente notable dependía de un conocimiento estilístico: cuando se realizó el mismo experimento con una pieza de Bartók, mucho más breve y estilísticamente desconocida por N. P., su memorización fue considerablemente peor, mientras que el pianista obtuvo en este caso un resultado muy superior. Los autores dedujeron que «la habilidad [de N. P.] tiene una base estructural», y que «necesita codificar el material en términos de estructuras y relaciones tonales, y su excepcional habilidad no puede actualmente traspasar los límites de esa estructura» [18].

### *La expresión en la interpretación*

La expresión, o «sentimiento», es tan fundamental en todo tipo de interpretaciones que, paradójicamente, es fácil de ignorar: perdemos la capacidad de escucharla. Un examen más detallado del fenómeno, sin embargo, revela un tema rico que ha originado muchas investigaciones. Seashore, uno de los pioneros de la investigación empírica sobre la interpretación durante los años 1930, definió la expresión de la siguiente manera: «la expresión artística del sentimiento en la música consiste en el alejamiento estético de lo que está regulado —el tono justo, la afinación real, la dinámica uniforme, el tempo metronómico, los ritmos rígidos, etc.» [19]. Debido a que su obra y la abrumadora mayoría de las investigaciones posteriores se han dedicado a la tradición escrita occidental, la frase «alejamiento estético de lo regulado» se ha interpretado frecuentemente como «alejamientos deliberados de las indicaciones de

la partitura». Pero ¿cómo se distinguen los alejamientos deliberados de los accidentes? ¿Cómo es la música cuando no hay una partitura que tratar? ¿Qué debe hacerse con las indicaciones expresivas que aparecen en la partitura (*accelerandos* y *ritardandos*, *crescendos* y *decrescendos*, etc.)? ¿Debemos considerar inexpresivos los correspondientes cambios de tempo y dinámica en la interpretación, simplemente porque vienen indicados en la partitura? Se han propuesto otras definiciones de la expresión que han intentado superar esos problemas manteniendo sin embargo el principio básico de que la expresión es un alejamiento de cierta norma [20].

Existen límites, claro está, en cuanto a los «alejamientos» esperados o permitidos convencionalmente. En la tradición concertística occidental, no se espera que el intérprete cambie las notas y ritmos de una pieza musical (aunque ésa es una postura relativamente reciente: en el siglo XIX, y más aún en siglos anteriores, los intérpretes solían adornar o hacer otro tipo de cambios en las composiciones que interpretaban), o que cambie el orden de los movimientos o secciones (excepto cuando esa opción está explícitamente permitida en la partitura, como ocurre en obras como *Klavierstück XI* de Stockhausen) o que interprete la música a una velocidad mucho mayor o menor que la indicada en la composición (aunque existe mucha música anterior a la época del metrónomo que carece de una indicación fija de tempo del compositor). Pero incluso dentro de esos límites, los intérpretes tienen muchísimas posibilidades de adoptar enfoques diferentes, lo cual plantea un problema interesante a la psicología de la música: determinar qué es lo que hace que una interpretación suene «humana» y musicalmente eficaz, y distinguir esa interpretación de una que suena exánime, inadmisiblemente afectada o caprichosa, o simplemente incompetente.

Una perspectiva general de la expresión es que el intérprete procura la transparencia entre la concepción y la acción, de modo que cada aspecto de su comprensión de la música encuentre una manifestación en la interpretación misma. Dicha perspectiva plantea al menos dos problemas. En primer lugar, es una visión idealizada de la interpretación, y la realidad se queda corta aunque sólo sea por razones prácticas: el intérprete puede tener más de una concepción de un evento o un pasaje de una pieza, pero en una interpretación determinada está obligado a decidirse por una de ellas; también puede ocurrir que el intérprete tenga una concepción de una característica a gran escala de la organización de la obra que sea imposible de proyectar de manera concreta dentro de la interpretación. En segundo lugar, y más esencialmente, el enfoque tiende a presentar la partitura como «la música», y la interpretación expresiva como cierto tipo de modificación de la misma. La partitura, claro está, no es la música, sino simplemente una de sus representaciones posibles (las otras son las grabaciones en CD, los vídeos y las des-

cripciones escritas), y puede considerarse como un anteproyecto para una interpretación —bastante incompleto, además— que sólo «tiene sentido» cuando se entiende dentro de un contexto cultural.

En los últimos tiempos, una cantidad considerable de investigaciones han procurado especificar los principios psicológicos que controlan la interpretación expresiva en la música [21]. Dichas investigaciones han utilizado una combinación de análisis detallado y modelos computerizados para examinar cómo los intérpretes utilizan expresivamente diversos parámetros musicales (tempo, dinámica, ataque, timbre, afinación, vibrato, etc.) en la interpretación. Algunos estudios experimentales han demostrado que la expresión puede ser sumamente estable en interpretaciones repetidas incluso después de varios años [22], que está presente incluso en las interpretaciones a primera vista [23] y que puede ser cambiada por el intérprete en el momento [24]. Estas observaciones han sido utilizadas para argumentar que la expresión no puede ser un patrón aprendido de tempo, dinámica y articulación que se recuerda y aplica a una pieza cada vez que se ejecuta, sino que debe surgir de la comprensión de la música por parte del intérprete en el transcurso de la interpretación.

En principio, todos los aspectos de la estructura musical contribuyen a especificar un perfil expresivo para una pieza, pero algunos estudiosos han demostrado que la estructura de las frases es particularmente relevante. Utilizando una simple regla, Neil Todd desarrolló un modelo que produce un patrón de tempo expresivo basándose en la jerarquía estructural de las frases [25]. Los perfiles de tempo resultantes son muy similares a los de las interpretaciones reales de instrumentistas profesionales. Igualmente, en un estudio sobre veintiocho interpretaciones de una pieza breve de Schumann, tomadas de grabaciones comerciales de muchos de los mejores pianistas del siglo XX, Bruno Repp encontró un elevado nivel de correspondencia entre los perfiles de tempo de las interpretaciones, todas las cuales estaban organizadas en torno a la estructura de la frase [26]. También demostró una gran diversidad en los niveles de expresión más superficiales, y sugirió que los intérpretes coinciden sustancialmente en cuanto a la forma general de una pieza musical y expresan su individualidad en los detalles más sutiles de la estructura y su ejecución expresiva. Repp realizó estudios ulteriores empleando una cantidad de muestras aún mayor (más de cien interpretaciones de la misma música) que han confirmado estas conclusiones [27].

La expresión puede entenderse como la consecuencia inevitable e irreprimible de la comprensión de la estructura musical; sin embargo, también es un intento consciente y deliberado del intérprete por hacer que sus interpretaciones sean perceptibles. Como prueba de la presencia inconsciente e inevitable de la expresión en la interpretación, algunos investigadores han demos-

trado que cuando se les pide a los intérpretes que toquen sin expresión, el nivel de variación de tempo y dinámica se reduce, pero no se elimina, conservando el mismo patrón general que se observa en condiciones normales [28]. Este aspecto de la expresión en la interpretación puede considerarse una consecuencia de la comprensión espontánea e inconsciente de los elementos básicos de la estructura musical por parte del intérprete: en términos sencillos, parece imposible *no* tocar una nota al final de una frase con características expresivas diferentes de las notas que se encuentran en medio de una frase. Sin embargo, es obvio que los intérpretes también dan forma a la expresión de sus interpretaciones, de manera consciente y deliberada, para lograr determinados resultados estilísticos y estructurales. Éste es uno de los propósitos de la práctica (además de superar los problemas puramente técnicos), e implica cambios en la comprensión que tiene el intérprete de la música y en la utilización equilibrada de diversos recursos expresivos. Todos estos procesos pueden destacar las propiedades expresivas implícitas en la música, o bien superponer un patrón sobre ellas; como ha señalado Shaffer entre otros, las pequeñas diferencias en la utilización y equilibrio de incluso unos pocos principios expresivos producirán caracterizaciones potencialmente distintas de una pieza [29]. Son relativamente pocos los estudios sistemáticos que se han realizado sobre los procesos psicológicos implicados en la práctica y el ensayo, así como en la memorización [30].

Aunque la expresión puede obedecer a ciertas «reglas», los intérpretes también poseen una habilidad (variable) para imitar un patrón expresivo que han escuchado —incluso cuando dicho patrón carece de una base estructural lógica [31]. En esas circunstancias, los intérpretes pueden crear un tipo de «imagen» auditiva de la interpretación que luego intentan copiar (más bien como un imitador); o pueden intentar recordar una descripción verbal de la interpretación (por ejemplo, «acelerar hacia el final de la primera frase, disminuir la velocidad en el medio de la segunda frase y después acelerar hacia el final»); o intentar concebir la interpretación como una imagen corpórea —una especie de coreografía mental que capta y representa la forma de la interpretación. Estas estrategias son más o menos parecidas a las que utilizan los alumnos para imitar a sus maestros o para calcar (quizá inconscientemente) aspectos de una representación grabada o en vivo que pueda haberles influido. Ellas nos recuerdan que la interpretación y la expresión, lejos de ser contribuciones singulares de ejecutantes aislados, se desarrollan dentro de una cultura a veces ensordecedoramente «ruidosa» de otras interpretaciones. Cualesquiera que sean la actitud y estrategia de los diversos intérpretes ante esa abundancia de influencias, está claro que una teoría de la interpretación presentada como una serie de reglas que relacionan la estructura con la expresión resulta demasiado abstracta e intelectual, y que la realidad es mucho

más práctica, tangible y, ciertamente, desordenada. En particular, el cuerpo no es tan sólo un «dispositivo de entrada y salida», y necesita que reconozcamos que tiene un papel más importante que el que sugiere un modelo basado en reglas.

### *El movimiento y el cuerpo durante la interpretación*

La relación de la expresión con el cuerpo humano y el papel de los movimientos corporales demuestran que la estructura no es la única determinante de la expresión. Una amplia gama de factores influirá en el resultado final, incluyendo las posibilidades del instrumento, la acústica del espacio, el tipo de público, el estado de ánimo y los objetivos del intérprete, las normas estilísticas y culturales, e incluso la ideología interpretativa que se adopte (por ejemplo, la interpretación «históricamente informada»). El movimiento y el cuerpo son particularmente importantes en esas complejas relaciones por varios motivos —el más obvio es que la música se produce con los movimientos del cuerpo y el instrumento, y por lo tanto está indeleblemente marcada por ellos. Los altibajos del movimiento y la tensión/relajación que experimentan los oyentes con la música se deben en parte a una identificación con el medio físico de la producción musical —tengamos o no una experiencia directa de los verdaderos instrumentos implicados. En un contexto más amplio, las investigaciones etnomusicológicas han demostrado que, en algunos casos, las estructuras características de un estilo musical determinado se remontan a factores físicos relacionados con el instrumento en el que se desarrolló dicha música. Por ejemplo, John Baily ha demostrado que la música para *dutar*, un instrumento de cuerda de Afganistán, evolucionó a partir del repertorio de otro instrumento de una región vecina, adquiriendo en el proceso características relacionadas con las cualidades físicas del *dutar*[32].

La estrecha relación que existe entre la música y el movimiento corporal ha sido estudiada durante mucho tiempo. En los años 1930, Alexander Truslit llevó a cabo una investigación experimental que demostró que al dar diferentes instrucciones (o imágenes) de movimiento a los intérpretes, éstos producían interpretaciones conmensurablemente diferentes[33]. Independientemente de lo anterior, otros investigadores han demostrado que los patrones espontáneos de tempo de los intérpretes concuerdan con la curva temporal del movimiento de los objetos en un campo gravitacional, lo cual sugiere que las interpretaciones que suenan «naturales» imitan el comportamiento de los objetos en movimiento en el mundo real[34]. Todd demostró que un modelo del tempo y la dinámica, basado en la velocidad y la fuerza de movimiento de los objetos bajo la influencia de la gravedad, puede explicar gran parte de

la expresión en las interpretaciones musicales espontáneas, confirmando así la idea de que la expresión en la interpretación musical es en ciertos aspectos un fenómeno físico y concreto sobre el cual se superponen elementos más elevados.

El movimiento en sí mismo forma parte de la interpretación entendida como evento audiovisual, y ha sido objeto de investigación y debate. Aunque ciertos movimientos son imprescindibles para producir música en un instrumento, la mayoría de los movimientos observados y estudiados en las investigaciones publicadas por Jane Davidson [35] están por encima de esa base puramente ergonómica —y por lo tanto pueden considerarse como expresivos en lugar de prácticos. Entre ellos se incluyen movimientos de balanceo de todo el cuerpo, así como gestos discretos que son, estrictamente hablando, «innecesarios» en la tarea básica de producir sonidos con el instrumento. En un estudio sobre un pianista profesional (analizado en el capítulo 10), Davidson concluyó que utilizaba algo parecido a un léxico de gestos expresivos en sus interpretaciones. Diferentes gestos parecían estar relacionados con funciones musicales específicas, y estaban tan unidos a su concepción de la música que se observaron incluso cuando produjo una interpretación imaginaria de la misma música sobre una mesa en lugar del teclado. El intérprete sólo estaba parcialmente consciente de los movimientos expresivos que hacía, y no había ninguna prueba que sugiriera que los movimientos habían sido desarrollados o ensayados deliberadamente. La impresión predominante (como se podría intuir) era que los movimientos formaban parte de su concepción y producción de la música, y eran generados durante la interpretación del mismo modo que se generan las características expresivas del sonido (tempo, dinámica, articulación, etc.) a partir de la idea de la música.

En este capítulo, el estudio de la interacción entre el cuerpo y la mente durante la interpretación refleja un dualismo cuerpo/mente persistente y más general. Cuando componemos, interpretamos, enseñamos, investigamos o escuchamos música, resulta muy fácil considerar que los límites impuestos por el cuerpo no son más que impedimentos para la mente todopoderosa y autosuficiente. La realidad es más diversa e integral, ya que la mente no está separada del cuerpo ni confinada al cerebro. Hasta el acto aparentemente pragmático de elegir una digitación está tan relacionado con pensar sobre lo que la música significa como con tocar las notas correctas en el momento adecuado [36]. La ejecución musical es una forma concreta de pensamiento musical, y el cuerpo forma parte del proceso de explorar la música y constituye igualmente un medio para su realización.

### *Conclusiones: mente, cuerpo y significado en la interpretación*

El enfoque psicológico de la interpretación ha tendido en ocasiones a separar radicalmente la estructura y la expresión. La estructura musical es sin duda un componente importante de aquello que motiva y da forma a la expresión, pero es sólo un elemento dentro de un amplio sistema de relaciones. Como he comentado antes, Shaffer ha propuesto que la expresión está relacionada con la caracterización de una pieza durante la interpretación, y que dos intérpretes con la misma imagen estructural de una pieza podrían producir ejecuciones distintas basadas en cómo caracterizan la música [37]. Esto plantea una pregunta sobre los límites de la noción de estructura en la música, y si la «caracterización» es viable como concepto. Más aun, sugiere que el término más importante —y en gran parte ausente en el análisis hasta ahora— es el término *significado,* y que, cuando un intérprete «caracteriza» una pieza en la interpretación, está construyendo un significado a través de la expresión. Una opción sería considerar que el intérprete utiliza su visión del significado de la música como «fuente generativa» de la expresión, pero como la expresión misma es parte del significado de la música, y no tan sólo un medio para el mismo, ello pone en duda al enfoque «generativo» en su conjunto —es decir, la idea de que la expresión es generada *desde* la estructura o el significado [38]. Avanzando otro paso por este camino, la investigación sobre la interpretación ha adoptado en gran parte un enfoque completamente individualista del intérprete y su mente. El contexto social de la interpretación (incluyendo a los cointérpretes, el público y la influencia de los maestros y tutores, así como las grabaciones e interpretaciones de otros, las actitudes sociales ante la interpretación y las «modas» interpretativas) es de vital importancia, pero hasta ahora es poco lo que se sabe de manera explícita. Basta considerar el ejemplo específico del miedo escénico para ver la magnitud que puede tener la influencia mutua de los componentes social e individual de la interpretación.

Plantear estas relaciones al final de un capítulo que ha intentado desglosar diferentes componentes de la interpretación y la mente puede parecer una revocación contradictoria de todo lo anterior. Pero no es así: aunque resulta productivo descomponer un fenómeno complejo para estudiarlo detallada y sistemáticamente, también es importante recordar que sus componentes no son independientes y que debemos intentar reintegrar y sintetizar. La interpretación musical es una construcción y articulación del significado musical en la que convergen las características cerebrales, corporales, sociales e históricas del intérprete, y si decidimos considerar esa convergencia como una expresión de la mente del intérprete, debemos recordar que la mente no está ni accionando el cuerpo ni confinada al cerebro.

## *Notas*

1. Véase Ralf Th. Krampe y K. Anders Ericsson, «Deliberate practice and elite musical performance», en John Rink (ed.), *The Practice of Performance: Studies in Musical Interpretation,* Cambridge, Cambridge University Press, 1995, pp. 84-102; véase también el capítulo 7 de este libro.
2. Por ejemplo, véanse John Baily, «Music structure and human movement», en Peter Howell, Ian Cross y Robert West (eds.), *Musical Structure and Cognition,* Londres, Academic Press, pp. 237-258; y Benjamin Brinner, *Knowing Music, Making Music,* Chicago, University of Chicago Press, 1995.
3. Carl E. Seashore, *Psychology of Music* [1938], Nueva York, Dover, 1967.
4. L. Henry Shaffer, «Performances of Chopin, Bach and Bartók: studies in motor programming», *Cognitive Psychology* 13 (1981), pp. 326-376; y «Timing in solo and duet piano performances», *Quarterly Journal of Experimental Psychology* 36A (1984), pp. 577-595.
5. Caroline Parker, «Music Performance», *Annual Review of Psychology* 48 (1997), pp. 115-138.
6. L. Henry Shaffer, «Performances of Chopin, Bach and Bartók: studies in motor programming», *cit.,* p. 365.
7. Por ejemplo, véase Manfred Clynes y Janice Walker, «Neurobiologic functions of rhythm, time and pulse in music», en Manfred Clynes (ed.), *Music, Mind, and Brain: The Neuropsychology of Music,* Nueva York, Plenum, 1982, pp. 171-216.
8. Por ejemplo, véanse L. Henry Shaffer, «Performances of Chopin, Bach and Bartók: studies in motor programming», *cit.;* y David Epstein, *Shaping Time: Music, the Brain, and Performance,* Nueva York, Schirmer Books, 1995.
9. L. Henry Shaffer, «Performances of Chopin, Bach and Bartók: studies in motor programming», *cit.;* y «Timing in solo and duet piano performances», *cit.*
10. L. Henry Shaffer, «Timing in solo and duet piano performances», *cit.*
11. Para un intento inicial, véase Lucy J. Appleton, W. Luke Windsor y Eric E. Clarke, «Cooperation in piano duet performance», en Alf Gabrielsson (ed.), *Proceedings of the Third Triennial European Society for the Cognitive Sciences of Music (ESCOM) Conference,* Uppsala, Uppsala University Press, 1997, pp. 471-474.
12. John A. Sloboda, «Experimental studies of music reading: a review», *Music Perception* 2 (1982), pp. 222-236.
13. Andrew J. Waters, Ellen Towsend y Geoffrey Underwood, «Expertise in musical sight reading: a study of pianists», *British Journal of Psychology* 89 (1998), pp. 123-149.
14. Véanse por ejemplo George R. Marek, «Toscanini's memory», en Ulric Neisser (ed.), *Memory Observed: Remembering in Natural Contexts,* San Francisco, W. H. Freeman, 1982, pp. 414-417; Grace Rubin-Rabson, «Studies in the psychology of memorizing piano music. VI: A comparison of two forms of mental rehearsal and keyboard overlearning», *Journal of Educational Psychology* 32 (1941), pp. 593-602 (véanse también los ensayos I-V de la serie); Michiko Nuki, «Memorization of piano music», *Psychologia* 27 (1984), pp. 157-163; Roger Chaffin, Gabriela Imreh y Mary Crawford, *Practicing Perfection: Memory and Piano Performance,* Mahwah, NJ, Erlbaum, 2002. Para una descripción de la investigación sobre la memorización en general, véase Alan D. Baddeley, *Human Memory: Theory and Practice,* Hove, Erlbaum, 1990.
15. Aaron Williamon, «The value of performing from memory», *Psychology of Music* 27 (1999), pp. 84-95; véase también el capítulo 8 de este libro.
16. Michiko Nuki, *op. cit.*, p. 160.

17. John A. Sloboda, Beate Hermelin y Neil O'Connor, «An exceptional musical memory», *Music Perception* 3 (1985), pp. 155-170.
18. *Ibid.*, p. 165.
19. Carl E. Seashore, *Psychology of Music,* cit., p. 9.
20. Por ejemplo, véase Peter Desain y Henkjan Honing, *Music, Mind and Machine,* Ámsterdam, Thesis Publishers, 1992, p. 175; véase también Bruno H. Repp, «The aesthetic quality of quantitatively average music performance: two preliminary experiments», *Music Perception* 14 (1997), pp. 419-444, que evalúa las cualidades estéticas de una interpretación «promedio» generada artificialmente.
21. Por ejemplo, véanse Eric F. Clarke, «Generative principles in music performance», en John A. Sloboda (ed.), *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition,* Oxford, Clarendon Press, 1988, pp. 1-26; Johan Sundberg, «Computer synthesis of music performance», en John A. Sloboda (ed.), *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition,* cit., pp. 52-69; y Neil P. McA. Todd, «A computational model of rubato», *Contemporary Music Review* 3 (1989), pp. 69-89.
22. Véase Manfred Clynes y Janice Walker, *op. cit.*
23. Véase L. Henry Shaffer, «Performances of Chopin, Bach and Bartók: studies in motor programming», *cit.*
24. Véase Eric F. Clarke, «Structure and expression in rhythnic performance», en Peter Howell, Ian Cross y Robert West (eds.), *Musical Structure and Cognition,* cit., pp. 209-236.
25. Neil P. McA. Todd, «A model of expressive timing in tonal music», *Music Perception* 3 (1985), pp. 33-58; véase también, del mismo autor, «A computational model of rubato», *cit.*
26. Bruno H. Repp, «Diversity and Commonality in music performance: an analysis of timing microstructure in Schumann's *Träumerei*», *Journal of the Acoustical Society of America* 92 (1992), pp. 2546-2568.
27. Bruno H. Repp, «A microcosm of musical expression: I. Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major», *Journal of the Accoustical Society of America* 104 (1998), pp. 1085-1100, y «A microcosm of musical expression: II. Quantitative analysis of pianists' dynamics in the initial measures of Chopin's Etude in E major», *Journal of the Accoustical Society of America* 105 (1999), pp. 1972-1988.
28. Véanse, por ejemplo, Carl E. Seashore, *Psychology of Music,* cit., y Caroline Palmer, «Mapping musical thought to musical performance», *Journal of Experimental Psychology: Human Perception and Performance* 15 (1989), pp. 331-346.
29. L. Henry Shaffer, «Musical performance as interpretation», *Psychology of Music* 23 (1995), pp. 17-38.
30. Véanse, por ejemplo, Linda M. Gruson, «Rehearsal skill and musical competence: does practice make perfect?», en John A. Sloboda (ed.), *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition,* cit., pp. 91-112; Susan Hallam, «Professional musicians' approaches to the learning and interpretation of music», *Psychology of Music* 23 (1995), pp. 111-128; Kacper Miklaszewski, «A case study of a pianist preparing a musical performance», *Psychology of Music* 17 (1989), pp. 95-109; Roger Chaffin, Gabriela Imreh y Mary Crawford, *Practicing Perfection: Memory and Piano Performance,* cit.; y el capítulo 7 de este libro.
31. Véase Eric F. Clarke, «Imitating and evaluating real and transformed musical performances», *Music Perception* 10 (1993), pp. 317-343.
32. John Baily, «Music structure and human movement», en Peter Howell, Ian Cross y Robert West (eds.), *Musical Structure and Cognition,* cit., pp. 242-256.

33. Alexander Truslit, *Gestaltung und Bewegung in der Musik,* Berlín-Lichterfelde, Chr. Friedrich Vieweg, 1938; para un análisis, véase Patrick Shove y Bruno H. Repp, «Musical motion and performance: theoretical and empirical perspectives», en John Rink (ed.), *The Practice of Performance: Studies in Musical Interpretation,* cit., pp. 55-83.
34. Neil P. McA. Todd, «The kinematics of musical expression», *Journal of the Acoustical Society of America* 97 (1995), pp. 1940-1949.
35. Jane W. Davidson, «Visual perception of performance manner in the movements of solo musicians», *Psychology of Music* 21 (1993), pp. 103-113, y «What does the visual information contained in music performances offer to the observer? Some preliminary thoughts», en Reinhart Steinberg (ed.), *Music and the Mind Machine: The Psychophysiology and Psychopathology of the Sense of Music,* Berlín, Springer Verlag, 1995, pp. 105-113; véase también el capítulo 10 de este libro.
36. Véase Eric F. Clarke, Richard Parncutt, Matti Raekallio y John A. Sloboda, «Talking fingers: an interview study of pianists' views on fingering», *Musicae Scientiae* 1 (1997), pp. 87-109.
37. Véase también Bruno H. Repp, «A microcosm of musical expression: I. Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major», *cit.*
38. Para otros enfoques relacionados, véanse John Rink, «Translating musical meaning: the nineteenth-century performer as narrator», y Nicholas Cook, «Analysing performance and performing analysis», en Nicholas Cook y Mark Everist (eds.), *Rethinking Music,* Oxford, Oxford University Press, 1999, pp. 217-238 y 239-261 respectivamente.

## *Lecturas complementarias*

Chaffin, Roger, Gabriela Imreh y Mary Crawford: *Practicing Perfection: Memory and Piano Performance,* Mahwah, NJ, Erlbaum, 2002.

Clarke, Eric, y Jane Davidson: «The body in performance», en Wyndham Thomas (ed.), *Composition, Performance, Reception: Studies in the Creative Process in Music,* Aldershot, Ashgate, 1998, pp. 74-92.

Davidson, Jane W.: «The social in musical performance», en David J. Hargreaves y Adrian C. North (eds.), *The Social Psychology of Music,* Oxford, Oxford University Press, 1997, pp. 209-228.

Dunsby, Jonathan: *Performing Music: Shared Concerns,* Oxford, Clarendon Press, 1995.

Gabrielsson, Alf: «The performance of music», en Diana Deutsch (ed.), *The Psychology of Music,* 2ª ed., San Diego y Londres, Academic Press, 1999, pp. 501-602.

Palmer, Caroline: «Music performance», *Annual Review of Psychology* 48 (1997), pp. 115-138.

Rink, John (ed.): *The Practice of Performance: Studies in Musical Interpretation,* Cambridge, Cambridge University Press, 1995.

Wilson, Glenn D.: «Performance anxiety», en David J. Hargreaves y Adrian C. North (eds.), *The Social Psychology of Music,* Oxford, Oxford University Press, 1997, pp. 229-248.

# SEGUNDA PARTE
# *Aprender a interpretar*

# 5 *Sobre la enseñanza de la interpretación*

JANET RITTERMAN

Toca con el alma, ¡no como un pájaro adiestrado!
C. P. E. BACH [1]

## *Introducción*

Este capítulo centra su atención en las interpretaciones en las que un músico o grupo de músicos «ejecuta música deliberadamente para un público», en la que se «realiza» una obra existente [2]. Su interés principal es la enseñanza de la interpretación instrumental solista de música culta occidental. Cronológicamente, abarca desde el siglo XVIII en adelante, sus límites geográficos son los de Europa occidental y destaca especialmente la enseñanza de aquellos que aspiran a un nivel profesional, más que de los aficionados o principiantes. Aunque frecuentemente se recomienda a los instrumentistas que aprendan de los cantantes, la enseñanza del canto no será abordada en este capítulo debido a la magnitud del tema y las numerosas ramificaciones particulares de la historia de la enseñanza vocal. En su lugar, los ejemplos citados se relacionan principalmente con la música para piano y violín. Debido a la popularidad de estos instrumentos y la importancia que se le da al dominio del repertorio solista y su presentación en concierto, ellos nos permitirán ver más claramente las tendencias de la enseñanza de la interpretación instrumental desde finales del siglo XVIII. En este capítulo, la frase «enseñanza de la interpretación» se referirá a la interpretación que implica una relación con un público, por lo que las tendencias educativas mencionadas anteriormente se refieren a la interpretación y comunicación de música en un concierto, más que a la técnica instrumental. Mientras que en la mayoría de las tradiciones de música culta occidental la enseñanza de la interpretación se ha centrado tanto en el último aspecto como en

el primero, el contexto de este capítulo será la enseñanza del estilo y la interpretación.

### *Los tratados musicales como guías para la interpretación*

El interés por los asuntos relacionados con la interpretación es una característica evidente de los tratados más importantes que comenzaron a proliferar durante la segunda mitad del siglo XVIII. Los tratados anteriores a esa época, que generalmente eran más breves, solían centrarse en las fases iniciales del aprendizaje, enfocando especialmente los rudimentos teóricos y las técnicas básicas. Aunque a menudo mencionaban incidentalmente la importancia de la ejecución expresiva y la opinión crítica, casi ninguno de ellos señala cómo se pueden demostrar o desarrollar esas cualidades, probablemente porque suponían que tales asuntos serían mejor explicados por el maestro y dependían del repertorio que estaba aprendiendo el alumno —frecuentemente composiciones del propio maestro.

A medida que aumentaba la cantidad de personas que querían aprender música, los tratados tendían a incluir comentarios más extensos sobre las cualidades más importantes de una interpretación eficaz. En su tratado de flauta de 1752, Quantz reiteró que el intérprete debe cultivar el buen gusto, tal y como habían aconsejado otros autores, y lo relacionó con la habilidad de identificarse con las diversas «emociones» de la música y ajustar correspondientemente el estilo de la ejecución, una opinión que apoyaron muchos autores posteriores:

> El intérprete de una pieza debe procurar implicarse en las emociones que ha de expresar. Y como en la mayoría de las piezas hay una alternancia constante de emociones, el intérprete debe saber discernir la naturaleza de la emoción que contiene cada idea, y hacer que su ejecución se ajuste a ella constantemente. Sólo así mostrará adecuadamente las intenciones del compositor y las ideas que tenía en mente cuando escribió la pieza» [3].

El tratado reconoce la necesidad de una enseñanza especializada para ayudar a los aspirantes a intérpretes a desarrollar ese criterio. Aunque Quantz desarrolló las derivaciones de sus consejos en el texto, consideró sin embargo que un buen maestro es esencial, como explica más o menos detalladamente en su introducción [4].

La importancia asignada en el tratado a las exigencias que plantea el arte interpretativo a los intérpretes en general —y no sólo a los flautistas— hizo que destacara entre la mayoría de las publicaciones anteriores. Sin embargo,

aunque abarca muchos temas, su enfoque es esencialmente retrospectivo, mientras que el *Versuch über die wahre Art das Clavier zu spielen* —un tratado de C. P. E. Bach, colega menor de Quantz en la corte de Federico el Grande en Postdam— tuvo una repercusión más directa en la enseñanza de la interpretación durante las décadas siguientes. Publicado en dos partes, la primera de las cuales apareció en 1753, el *Versuch* de Bach había sido reeditado tres veces para el año 1787 y no tardó en convertirse en la publicación más difundida, citada y admirada de su especie [5]. Fue el tratado que Beethoven le pidió al joven Carl Czerny para su primera clase de piano en Viena [6].

El tratado de Bach se centra principalmente en la enseñanza de la interpretación a solo de instrumentos de teclado. Como su autor señaló en cartas posteriores, iba dirigido a «aquellos para quienes la música es una meta», más que a los aficionados [7]. Por lo tanto, en su primer párrafo se subraya la importancia de la buena interpretación como una de las cualidades de las que depende el verdadero arte de tocar instrumentos de teclado [8]. Para Bach, como para Quantz, ello implicaba «la capacidad [...] de hacer que el oído sea consciente del verdadero contenido y sentimiento de una composición». Sin embargo, aunque al hacer hincapié en el «verdadero contenido» reconoce la necesidad de identificarse con las intenciones del compositor, la definición de Bach subraya especialmente el criterio del oído. Para Bach, la buena interpretación requiere que el oyente «escuche todas las notas y sus ornamentos ejecutados a tempo con la dinámica apropiada mediante una pulsación vinculada al verdadero contenido de la pieza» [9].

«Estudiar escuchando», una modalidad que Bach describió como «una especie de hurto legal», era esencial para los que querían convertirse en intérpretes consumados: era la única manera, según él, de que los asuntos «que no se pueden demostrar fácilmente, y mucho menos escribir», se hicieran evidentes [10]. Al igual que Quantz, a Bach le preocupaba la falta de una buena enseñanza de la interpretación. Lamentó el hecho de que a la mayoría de los alumnos se les exigía tocar las composiciones de su maestro y no se les enseñaban obras de compositores anteriores o de otros países, «con el pretexto de que eran obsoletas o demasiado difíciles» [11]. Sin embargo, y al igual que el tratado de Quantz, el *Versuch* no incluía consejos sobre el repertorio de otros compositores que fuese apropiado para el estudio.

## *Fuentes para la enseñanza de la interpretación*

La mayor parte de lo que C. P. E. Bach aprendió como intérprete lo asimiló de su padre y de un constante «estudiar escuchando», particularmente en la corte. En la mayoría de los países de la Europa occidental anterior a 1800, la

enseñanza musical de los que pretendían convertirse en músicos profesionales era generalmente de ese tipo —proporcionada dentro de la familia, la corte o la iglesia, o mediante métodos de aprendizaje. Las biografías de Haydn, Mozart y Beethoven revelan que desarrollaron sus habilidades interpretativas de esa manera, recibiendo una formación que solía hacer hincapié tanto en las habilidades de la composición como en las de la interpretación misma.

El adiestramiento más intensivo para la interpretación lo proporcionaban los conservatorios de Nápoles y Venecia —instituciones establecidas desde finales del siglo XVI por determinadas fundaciones benéficas para el cuidado de los huérfanos y otros necesitados. En algunos de ellos, la formación para la interpretación se convirtió en una actividad prominente, y adquirieron una reputación envidiable por su quehacer musical. Sin embargo, para las últimas décadas del siglo XVIII, la convulsión política había provocado el declive de la mayoría de esas instituciones. No obstante, de las mismas raíces comenzó a emerger un nuevo tipo de conservatorio. Estas instituciones nuevas fueron establecidas generalmente con el financiamiento de la ciudad o el Estado, en lugar de la iglesia o las organizaciones benéficas, y reflejaban la creciente importancia que se asignaba a la actividad musical pública. A través de los conciertos ofrecidos por los estudiantes o auspiciados por la institución, muchas llegaron a ser de vital importancia para el desarrollo del concierto público en sus respectivas ciudades.

Para el cambio de siglo, las transformaciones sociales, económicas y culturales que se extendieron por Europa ayudaron a asegurar que la enseñanza de la interpretación llegara a ser más generalizada, cada vez más solicitada por la clase media y menos influida por las tradiciones pedagógicas anteriores. En las grandes ciudades, la demanda de educación musical aumentaba vertiginosamente, como evidencian los anales, periódicos musicales y otras referencias especializadas de la época [12]. La enseñanza privada floreció, pero en general los nuevos conservatorios fueron los que ejercieron una mayor influencia en la sistematización gradual de la enseñanza de la interpretación a un nivel avanzado. Muchos de ellos incluyeron en su profesorado de instrumentistas a grandes intérpretes de la época, cuya reputación era conocida a lo largo de Europa occidental por sus carreras concertísticas. Su influencia contribuyó mucho a determinar lo que llegó a aceptarse como el repertorio instrumental estándar.

El primero de estos conservatorios modernos surgió en París después de la revolución. Fundado en 1795 con carácter nacional, y añadiendo escuelas más antiguas y especializadas, el Conservatorio de París fue establecido con responsabilidades conjuntas para la enseñanza y la interpretación: los alumnos eran adiestrados para interpretar en festivales nacionales y otros tipos de

interpretaciones públicas. Los alumnos, inicialmente 600 [13], eran seleccionados de todo el país y recibían enseñanza gratuita.

La influencia y el prestigio de la nueva institución se difundieron rápidamente, y para los años 1820 ciudades como Praga, Viena, Milán, Bruselas, Londres y La Haya habían seguido el ejemplo. La tendencia continuó, y con el tiempo se establecieron conservatorios en muchas ciudades de habla alemana, Rusia (San Petersburgo y Moscú), Gran Bretaña y América, algunos de los cuales utilizaron extensamente los enfoques de la enseñanza de la interpretación desarrollados en París. Muchos comenzaron con una gama de especialidades relativamente limitada que reflejaba las prioridades y posibilidades locales. En París, a pesar de la intención expresa de adiestrar a los alumnos «en todas las ramas del arte de la música» [14], la enseñanza instrumental en 1806 se centraba en la flauta, el oboe, el clarinete, el fagot, la trompa, el violín, el violonchelo y el piano. La introducción de otras especialidades fue gradual: por ejemplo, el arpa se introdujo en 1825, la trompeta y el trombón durante los años 1830.

La necesidad de establecer sistemas y modelos de enseñanza de la interpretación condujo a la publicación, en París, de una serie de métodos [15]. Estos métodos eran encargados especialmente por los profesores del conservatorio y estaban sujetos a un proceso de aprobación institucional formal. Aunque eran comparables a sus antecesores en cuanto a organización y contenido general, estos métodos superaban ampliamente el alcance de los tratados instrumentales más antiguos. La mayoría proporcionaba indicaciones extensas y detalladas sobre asuntos de técnica, estilo e interpretación, complementadas con numerosas ilustraciones musicales y material pedagógico, incluyendo composiciones de épocas anteriores y diversas escuelas nacionales, así como obras contemporáneas. Entre los métodos más influyentes que surgieron del Conservatorio de París durante la primera década de su existencia se encuentran los de cuerdas y piano. El método de piano, creado inicialmente por Adam y Lachnith en 1798 [16], proporciona una amplia colección de fragmentos, lo que muestra la amplitud del repertorio que se estudiaba. Principios similares dieron forma al *Méthode de violon* publicado en 1803 por Baillot, Rode y Kreutzer [17], y el *Méthode de violoncelle* de 1804, en el que Baillot colaboró con Baudiot, Levasseur y Catel.

París indicaba el camino, y los demás le seguían. El siglo XIX vio el nacimiento de un verdadero raudal de métodos, particularmente para los instrumentos más populares [18]. Muchos de ellos le debían bastante a los tratados antiguos, aunque no siempre lo reconocían. Algunos fueron escritos por maestros relacionados con otros conservatorios; otros por profesores-intérpretes establecidos que no estaban vinculados a ninguna institución, como Czerny, quien dio clases de piano en Viena durante unos cincuenta años, atrayendo a

estudiantes talentosos de muchas partes de Europa. Al igual que los tres volúmenes de la obra *Complete Theoretical and Practical Piano Forte School,* op. 500 de Czerny [19], los métodos creados por distinguidos intérpretes de la época —como los pianistas Hummel [20] y Moscheles [21] y el violinista Spohr [22], famosos en muchos de los principales centros musicales de Europa por sus giras como virtuosos— traspasaron rápidamente las fronteras nacionales. Reseñados concienzudamente en las principales revistas musicales de la época, los métodos más valorados fueron publicados en otros idiomas —generalmente en Viena, Londres y París—, ayudando así a desarrollar las perspectivas cada vez más cosmopolitas de la enseñanza de la interpretación a medida que avanzaba el siglo.

### *Las escuelas nacionales de interpretación*

A pesar de estas tendencias unificadoras, la existencia de escuelas nacionales de interpretación relacionadas con determinados instrumentos fue reconocida generalmente a lo largo del siglo XIX y bien entrado el XX. Algunos críticos utilizaban el término «nacional» históricamente. Czerny lo empleó en ese sentido, distinguiendo seis escuelas pianísticas relacionadas con determinados compositores: Clementi; Cramer y Dussek; Mozart; Beethoven; Hummel, Kalkbrenner y Moscheles; y una escuela incipiente, representada por tres artistas jóvenes de la época, Thalberg, Chopin y Liszt. En cualquiera de ellas, sin embargo, la música de cada compositor debía ser «ejecutada en el estilo en que la compuso»; Czerny rechazó de entrada la idea de que todas podían ser interpretadas en el mismo estilo [23].

Particularmente desde mediados del siglo XIX, otros utilizaron el concepto de escuela nacional para diferenciar las tendencias contemporáneas de interpretación y enseñanza. Las distinciones, un tema comentado con frecuencia en los relatos de maestros o alumnos, se consolidaron con la formalización de los enfoques pedagógicos de los principales conservatorios nacionales, siendo impulsadas en algunos casos por tendencias nacionalistas. Ello se puede observar en el nacimiento de las escuelas rusas de interpretación instrumental. Las referencias a las escuelas interpretativas rusas se hicieron comunes sólo a partir de la segunda mitad del siglo —después del establecimiento de los Conservatorios de San Petersburgo y Moscú, en los años 1860. Hasta entonces, la enseñanza instrumental a nivel avanzado en Rusia había sido proporcionada principalmente por maestros-intérpretes identificados con las tradiciones interpretativas europeas más importantes [24].

Las características de las escuelas interpretativas nacionales reflejan a menudo las contribuciones de compositores e intérpretes: el término se ha utili-

zado para identificar determinados enfoques de la expresión musical, la técnica de ejecución, la elección de repertorio e incluso la organización de los currículums. En cuanto a la técnica, durante gran parte del siglo XIX había indudablemente diferencias que distinguían a ciertas escuelas interpretativas de otras. Sin embargo, en los casos del piano y las cuerdas, dichas diferencias se desvanecieron gradualmente por los continuos desplazamientos de los principales maestros-intérpretes y por la creciente movilidad de los alumnos. Para el fin del siglo, se había producido una gran difusión: aunque algunos maestros seguían identificándose con determinados métodos técnicos y repertorios especializados, las similitudes en cuanto al enfoque estilístico e interpretativo se habían acentuado, y las diferencias eran más difíciles de clasificar.

La escuela violinística francesa refleja una tendencia similar. A principios del siglo XIX, dicha escuela, identificada con la enseñanza del Conservatorio de París, se había convertido en el principal foco de atención, en parte porque asimiló aspectos de la escuela italiana del siglo XVIII que Viotti había llevado a París en los años 1780 [25]. La continuidad de la enseñanza desempeñó un papel importante en el mantenimiento de la reputación de la escuela: Baillot, profesor del Conservatorio de París desde 1795 hasta 1842 [26], fue reemplazado por violinistas que habían estudiado con él y se mantuvieron en el cargo hasta principios de la década de 1890. La minuciosidad de la documentación de la escuela francesa y su continua reevaluación también fueron factores favorables. En 1834, Baillot escribió un método nuevo y exhaustivo, *L'Art du violon,* que proporcionaba una perspectiva general actualizada de la enseñanza de la escuela violinística francesa y el repertorio estudiado por los alumnos en París, que ayudó a asegurar que sus rasgos característicos siguieran influyendo en la ejecución de las cuerdas durante las décadas siguientes [27].

Para 1842, año en que murió Baillot, estas tradiciones estaban siendo influidas por nuevas tendencias: se estaba desarrollando una escuela franco-belga basada en el Conservatorio de Bruselas e influida por la técnica interpretativa de Paganini. Le siguieron otras fusiones de estilos. Durante la segunda mitad del siglo, debido a que los alumnos extranjeros eran admitidos más fácilmente en el Conservatorio de París [28], las características de la escuela interpretativa francesa se difundieron en círculos cada vez mayores, modificando y siendo modificadas por otras tradiciones violinísticas. Entre los que estudiaron en París y contribuyeron a esa tendencia se encontraban Marsick, Wieniawski y Kreisler. El arte de Kreisler, vienés de nacimiento, llegó a ser mundialmente famoso a través de sus interpretaciones y grabaciones. Wieniawski, polaco, dio clases en el recién fundado Conservatorio de San Petersburgo en los años 1860, y más tarde en Bruselas. Marsick, un violinista belga que después de estudiar en París fue alumno de Joachim en Berlín, regresó a París para dar clases en el conservatorio en los años 1890; entre sus alumnos

se encontraban Carl Flesch, de origen húngaro, y Georges Enescu, rumano; ambos habían estudiado en Viena. Para comienzos del siglo XX, eran relativamente pocos los maestros-intérpretes importantes que podían clasificarse como representativos de una determinada escuela nacional. Los estilos de ejecución, que en una época se relacionaban con un compositor o una nacionalidad, se identificaban cada vez más con el intérprete o el maestro que con una localidad geográfica.

### *Estilo e interpretación*

Aunque se consideraba que las diferencias nacionales existían, la mayoría de los comentarios sobre estilo e interpretación de los tratados del siglo XIX presentaban opiniones generalmente similares sobre las características esenciales de la interpretación expresiva y la importancia de la aportación del intérprete en dicha interpretación. Uno de los temas recurrentes era la habilidad del intérprete para imprimir su individualidad en la interpretación —un tema que fue planteado por Adam [29] y desarrollado por otros. Spohr, quien explicó la distinción entre lo que consideraba una mera *«expresión* o *estilo correcto»* y el *«estilo excelente»,* apoyó la opinión de Quantz en cuanto a la importancia de las intenciones del compositor. El *«estilo excelente»*, afirmó, no significaba simplemente que los intérpretes añadían sus propias ideas, sino que eran «capaces de recrear intelectualmente el tema, de modo que el oyente pueda descubrir y ser partícipe de las intenciones del compositor» [30].

Los críticos del siglo XIX prestaron especial atención a la importancia de las fluctuaciones o alteraciones de tempo como una de las variantes personales que podía introducir el intérprete. En la última sección de su método, Hummel ilustra ese aspecto con extensos fragmentos de su Concierto para piano en La menor op. 85, que muestran mediante indicaciones escritas en la partitura (e.g. «un tanto acelerado», «un poco más lento», «relajando un tanto el tempo») la medida en que se esperaba que hubiera modificaciones de tempo [31]. Dedicado exclusivamente a «ejecutar con expresión», el tercer volumen de la *Escuela* de Czerny incluye un capítulo sobre el buen gusto en el empleo de los *ritardandos* y *accelerandos.* Aunque comienza afirmando que toda composición «debe tocarse en el tempo prescrito por el compositor», los ejemplos musicales de Czerny apoyan su opinión de que «en casi toda línea de música hay ciertas notas y pasajes en los que es necesario un pequeño *ritardando* o *accelerando*, para embellecer la lectura y aumentar el interés [32].

Las variaciones elegantes de los tonos o ritmos de la partitura, por razones de elocuencia o para dar espontaneidad a la ejecución, también eran vistas como una parte importante del aporte del intérprete consumado. Las varia-

ciones podían incluir cambios en los detalles de un pasaje brillante o de los ritmos, la interpolación de pasajes al estilo de una cadencia o, en el caso de los pianistas, la ejecución arpegiada de acordes. Aunque advertían que para ello es necesario tener un criterio musical sólido y que el contexto determina lo que es adecuado, la mayoría de los métodos consideraban que la escasez era una falta tan grave como el exceso. Tanto Hummel como Kalkbrenner editaron conciertos de Mozart modernizando las figuraciones de acuerdo con el gusto contemporáneo. La escuela de Czerny, ampliada en publicaciones separadas sobre ornamentación y expresión «en el estilo moderno» [33], adoptó una postura similar. Existen muchas pruebas de que esa práctica fue perpetuada por maestros influyentes como Liszt y Leschetizky, quienes realizaban cambios textuales en las obras que ejecutaban y enseñaban. Liszt es conocido por haber «enriquecido» un pasaje brillante de uno de los conciertos de Weber que consideraba pobre [34], mientras que Leschetizky interpretaba las partituras con una «gran libertad y flexibilidad» [35].

Al mismo tiempo, había síntomas de que la actitud hacia la relación entre el intérprete y la obra estaba cambiando, y que dichos cambios estaban repercutiendo en la enseñanza de la interpretación. Clara Schumann ejerció una gran influencia en ese sentido: sus clases, que transcurrieron paralelas a su carrera de intérprete durante más de cincuenta años, influyeron a generaciones de intérpretes y maestros de Europa continental e Inglaterra [36]. Cuando tocó en Viena en 1856, el crítico Eduard Hanslick señaló que «la subordinación artística de su propia personalidad a las intenciones del compositor» era «un principio» para ella, y describió su ejecución como «una representación sumamente fiel de composiciones magníficas, pero no una manifestación de una personalidad magnífica». Otras características de su interpretación que Hanslick destacó como inusuales eran su claridad («Todo es preciso, claro, nítido como un dibujo a lápiz») y, con ciertas reservas sobre su conveniencia, su «estricta conformidad con el compás» [37].

En estos aspectos, el enfoque de Clara Schumann era marcadamente diferente del de Anton Rubinstein. Ampliamente admirado por «el fuego y la pasión de su ejecución y por su imaginación y espontaneidad» [38], Rubinstein, primer director del Conservatorio de San Petersburgo y personaje destacado de la vida musical europea durante la segunda mitad del siglo, adoptó un enfoque más liberal de las obras que interpretó. Según Josef Hofmann, posiblemente el más famoso de sus alumnos, Rubinstein aconsejaba: «Toca primero exactamente lo que está escrito. Si le has hecho plena justicia y aun así le quieres añadir o cambiar algo, pues hazlo» [39]. Pero para cuando murió Rubinstein, en 1894, tales opiniones, que ya eran cuestionadas desde varias tendencias, no eran aceptadas incuestionablemente. Hofmann, quien en 1926 se convertiría en el primer director del Curtis Institute de Filadelfia, opinaba

que el intérprete «siempre ha de sentirse convencido de que toca únicamente lo que está escrito», afirmando que la correcta comprensión de una obra —prerrequisito de una «interpretación verdadera»— dependía «únicamente de una lectura escrupulosamente exacta» [40]. El desarrollo, a mediados del siglo XIX, del concepto de ejecutante como intérprete, en lugar de cocreador, que condujo a algunos críticos a emplear metáforas casi religiosas para describir el acto de interpretar, fomentó este sentido de la interpretación como regla. Durante el siglo XX, el «respeto al texto» llegó a dominar cada vez más la enseñanza de la interpretación.

### *Los retos de la enseñanza de la interpretación*

El violinista Duncan Druce se encuentra entre los que lamentan la tendencia de los músicos modernos a producir interpretaciones admirablemente claras y precisas pero «relativamente sosas y carentes de personalidad». En su escrito sobre los enfoques históricos de la ejecución del violín, señala que ello «es parte de un declive general y gradual del papel creativo del intérprete» [41]. Aunque la enseñanza de los instrumentos que poseen un extenso repertorio solista decimonónico ha abrazado cada vez más la tendencia hacia un enfoque más histórico de la interpretación, es cierto que rara vez ha estado acompañada de una comprensión de la libertad creativa que implica dicha influencia. Muy frecuentemente ha tendido a fomentar la creencia de que la partitura ha de ser interpretada literalmente. La obligación de responder individualmente al espíritu de la obra —inequívocamente apoyada en épocas anteriores— ha sido suplantada con demasiada facilidad, en una era de menos confianza, por la creencia errónea de que el respeto al texto de la partitura ofrece seguridad frente al reto crítico. Las ediciones Urtext, junto con los sistemas y criterios de evaluación de las interpretaciones, han favorecido esta tendencia normalizadora, como también lo ha hecho la creciente autoridad que se le otorga a las grabaciones.

Comentando esa tendencia, Alfred Brendel, en un ensayo que escribió en 1966, rechazó categóricamente la noción de que una obra «hablará por sí misma siempre y cuando el intérprete no interponga su personalidad», recordando a los intérpretes que «las notas musicales sólo pueden sugerir, y las indicaciones expresivas sólo pueden complementar y confirmar lo que debemos, antes que nada, leer en la representación de la composición misma». Apoyando las opiniones que otros expresaron mucho antes, describió que el proceso requiere de «nuestras propias emociones, nuestro propio sentido, nuestro propio intelecto, nuestros propios oídos refinados» [42]. Sin embargo, aunque en muchas escuelas de interpretación se valora hipócritamente el pa-

pel del criterio y la intuición musical del alumno para darle forma a una interpretación, resulta sumamente fácil neutralizar su influencia en la práctica.

Aunque ha sido valioso en muchos aspectos, el aumento repentino del interés por la interpretación «auténtica» que tuvo lugar durante las décadas de 1980 y 1990 no ha hecho nada por resolver esta diferencia intrínseca. Mientras que los enfoques reconocidamente «auténticos» a menudo dicen tanto sobre el gusto contemporáneo como sobre los gustos y valores musicales de épocas remotas (como han argumentado Richard Taruskin y otros) [43], destacan sin embargo la necesidad de que los intérpretes y maestros sigan un camino en el que permanezcan fieles tanto a sí mismos como a la música que eligen interpretar. En una época de creciente diversidad de enfoques y actitudes, la enseñanza de la interpretación debe ayudar a los músicos a esclarecer sus convicciones personales como intérpretes; como afirma Jonathan Dunsby: «Nos guste o no, hoy en día no puede haber un intérprete inocente, si es que pudo haberlo alguna vez» [44]. Una de las consecuencias perdurables del movimiento de la interpretación histórica debería ser animar a los intérpretes a adoptar un enfoque menos coercitivo y más creativo de la música que interpretan, cualquiera que sea el repertorio. Como ha afirmado Nicholas Kenyon, «la tradición no es lo que era» [45]. En vista del alcance del repertorio que actualmente se considera como normativo para los intérpretes, el conocimiento bien fundado de las nociones que aportan las disciplinas como la estética, el análisis, la musicología (incluyendo la etnomusicología) y la psicología ya no puede ser una excepción. Para el intérprete en formación, su importancia es mayor que nunca.

Sin embargo, los intérpretes jóvenes necesitan que se les ayude a adquirir ese conocimiento y a no darle demasiado peso: no puede ser un sustituto del instinto musical, o llegar a ser tan importante que silencie la voz personal. Schnabel valoraba la enseñanza de Leschetizky porque funcionaba como «una corriente que activaba o liberaba toda la vitalidad latente de la naturaleza del alumno. Iba dirigida a la imaginación, al gusto y a la responsabilidad personal» [46]. Ésas son las cualidades que la buena enseñanza de la interpretación siempre ha fomentado. La coherencia estilística requiere ese enfoque integral, en el que las ideas del compositor se complementan con los conocimientos del intérprete —conocimiento de sí mismo como de la música que ejecuta. Sólo así los intérpretes podrán vestirse con su propia individualidad, como lo describió Baillot, e interpretar con convicción y coherencia, no como el «pájaro adiestrado» de C. P. E. Bach, sino utilizando con confianza «las inspiraciones del corazón y los estallidos de la imaginación» [47].

# Notas

1. Carl Philipp Emanuel Bach, *Essay on the True Art of Playing Keyboard Instruments (Versuch über die wahre Art das Clavier zu spielen* [1753, 1762]), William J. Mitchell (trad. y ed.), Londres, Eulenburg, 1949, p. 150.
2. Véase John A. Sloboda, *The Musical Mind: The Cognitive Psychology of Music,* Oxford, Oxford University Press, 1985, p. 67.
3. Johann Joachim Quantz, *On Playing the Flute* [1752], 2ª ed., Edward R. Reilly (trad.), Londres, Faber and Faber, 1985, pp. 124-125.
4. *Ibid.*, pp. 15-17.
5. Para comentarios, véase Carl Philipp Emanuel Bach, *Essay on the True Art of Playing Keyboard Instruments*, cit., pp. 1-23 passim.
6. Véase Carl Czerny, «Recollections from my life», *Musical Quarterly* 62 (1956), p. 307.
7. Carta publicada en *Hamburger unpartheiischer Correspondent,* 7 (1773); citado de Carl Philipp Emanuel Bach, *Essay on the True Art of Playing Keyboard Instruments,* cit., p. 8.
8. Las otras cualidades —respaldadas igualmente por autores anteriores— son la digitación correcta y los ornamentos apropiados; Carl Philipp Emanuel Bach, *Essay on the True Art of Playing Keyboard Instruments,* cit., p. 30.
9. *Ibid.*, p. 148.
10. *Ibid.*, p. 28.
11. *Ibid.*, pp. 30-31.
12. Véase por ejemplo César Gardeton, *Annales de la musique ou Almanach musical... [1819, 1820],* Ginebra, Minkoff, 1978, pp. 39-43, que enumera trece escuelas de música inauguradas en la época en París. En una carta que Mendelssohn escribió en París a Carl Friedrich Zelter, Berlín (el 15 de febrero de 1832), comenta que en la ciudad había 1.800 profesores de piano, pero que «aún no era suficiente» para satisfacer la demanda (Felix Mendelssohn, *Letters,* trad. y ed. G. Selden-Goth, Londres, Paul Elek, 1946, p. 190). En *Musique et musiciens* (París, Pagnerre, 1862), Oscar Comettant, un importante crítico parisino de la época, señala que el número de profesores de piano en París había ascendido a 20.000. Para comentarios sobre la situación en Inglaterra, véase Cyril Ehrlich, *The Musical Profession in England since the Eighteenth Century,* Oxford, Clarendon Press, 1985, pp. 76-120.
13. El número de alumnos variaba de año en año (véase Constant Pierre, *Le Conservatoire National de Musique et de Déclamation...,* París, Imprimerie nationale, 1900, p. 873. Para 1828 se habían inscrito 11.000 alumnos; de los 3.800 que se graduaron, 1.127 eran pianistas o acompañantes (véase François-Joseph Fétis, «Statistique musicale de la France. Premier Article», *Revue musicale* 6 (1828-1829), pp. 126-127.
14. Constant Pierre, *Le Conservatoire National de Musique et de Déclamation,* cit., p. 125.
15. Estos métodos, diez en total, eran para flauta, clarinete, fagot, trompa (dos), serpentón, violín, violonchelo y voz, y otro de armonía, solfeo y los rudimentos de la música.
16. El método de Louis Adam y Louis-Wenceslas Lachnith, *Méthode ou principe général du doigté pour le forté-piano,* París, Sieber, 1798, fue editado y publicado de nuevo por Adam en 1802 y 1804 con el título *Méthode de piano du Conservatoire adoptée pour servir à l'enseignement dans cet établissement,* París, Conservatoire Royal, 1802, 1804. La versión reeditada en 1805, *Méthode complète pour le piano,* París, Conservatoire impérial de musique, 1805, adquirió aún más popularidad y fue publicado en dos ediciones alemanas y dos italianas.
17. Este método reemplazó al de Jean-Baptiste Cartier, *L'Art du violon...* (París, Decombe, [1798]), que el autor había dedicado al Conservatorio recién establecido. La edición de 1803 se siguió utilizando para la enseñanza hasta que fue reemplazado por *L'Art du violon* de Pierre Baillot (París, Dépôt Central de la Musique [1834]).

18. Para detalles sobre los métodos para violín, véase Robin Stowell, *Violin Technique and Performance Practice in the Late Eighteenth and Early Nineteenth Centuries,* Cambridge, Cambridge University Press, 1985, pp. 5-10, 368-374. Para una lista de los principales métodos para violonchelo, véase Valerie Walden, *One Hundred Years of Violoncello: A History of Technique and Performance Practice, 1780-1840,* Cambridge, Cambridge University Press, 1998, pp. 300-301. Para tratados sobre teclado, véase Glyn Jenkins, «The legato touch and the "ordinary" manner of keyboard-playing from 1750 to 1850: some aspects of the early development of piano technique», tesis doctoral, Universidad de Cambridge, 1976.
19. Carl Czerny, *Complete Theoretical and Practical Piano Forte School,* op. 500 (versión inglesa de J. A. Hamilton, 3 vols., Londres, R. Cocks & Co. [1839]; publicado también en París y Viena).
20. Johann Nepomuk Hummel, *Ausführlich theoretisch-practische Ausweisung zum Piano-forte Spiel,* 3 vols., Viena, Haslinger, 1828.
21. François-Joseph Fétis e Ignaz Moscheles, *Méthode des méthodes de piano,* París, Schlesinger, 1837.
22. Louis Spohr, *Violinschule,* Viena, Haslinger [1832].
23. Carl Czerny, *Complete Theoretical and Practical Piano Forte School,* cit., III, pp. 99-100. Este comentario hace eco a la opinión de Louis Adam, *Méthode de piano du Conservatoire adoptée pour servir à l'enseignement dans cet établissement,* cit. (1802), p. 234.
24. La enseñanza de piano en San Petersburgo antes del establecimiento del Conservatorio, por ejemplo, estaba marcadamente influida por John Field (alumno de Clementi) y después por Adolphe Henselt (alumno de Hummel).
25. En 1792, debido a presiones revolucionarias, Viotti se mudó a Londres, donde fue una figura clave en la vida musical de la ciudad durante los siguientes cinco años.
26. La escuela de Baillot fue descrita por su alumno Eugène Sauzay, quien más tarde enseñó en el Conservatorio (de 1860 a 1892); véase Brigitte François-Sappey, «La Vie musicale à Paris à travers les *Memoires* d'Eugène Sauzay», *Revue de musicologie* 60 (1974), pp. 159-210.
27. En 1835 se publicaron dos ediciones alemanas; le siguieron otras ediciones en francés y en español. Está publicado en inglés con el título *The Art of the Violin,* traducido y editado por Louise Goldberg, Evanston, Ill., Northwestern University Press, 1991.
28. Al principio, los alumnos extranjeros sólo eran admitidos en el Conservatorio de París con una aprobación previa del ministerio.
29. «Los intérpretes, así como los compositores, han de tener su estilo personal.» Louis Adam, *Méthode de piano du Conservatoire adoptée pour servir à l'enseignement dans cet établissement,* cit. (1802), p. 233.
30. Louis Spohr, *Grand Violin School,* C. Rodolphus (trad.), Londres, Wessel, s.f., p. 179; subrayados del original.
31. Johann Nepomuk Hummel, *A Complete Theoretical and Practical Course of Instructions on the Art of Playing the Piano Forte* [1828], 3 vols., Londres, T. Boosey & Co., 1830, III, pp. 43-50.
32. Carl Czerny, *Complete Theoretical and Practical Piano Forte School,* cit., III, pp. 34-35. Para una descripción posterior, que abarca un repertorio más extenso, véase el capítulo 16, «The accelerando and rallentando» en Adolph Kullak, *The Aesthetics of Pianoforte-Playing,* Theodore Kullak (trad.), Hans Bischoff (ed.), 3ª ed., Nueva York, G. Schirmer, 1903, pp. 280-294.
33. La portada de la obra de Czerny *L'École des expressions,* Londres, R. Cocks & Co., s.f., señala que contiene melodías que incluyen «Ornamentos elegantes, en el estilo moderno, compuestas expresamente como modelos, para la ornamentación improvisada». Su obra *L'École des ornements,* Londres, R. Cocks & Co., s.f., incluye ejemplos similares.

34. Véase Mme Auguste Boissier, *Liszt Pédagogue,* París, Honoré Champion, 1927, p. 32.
35. Harold C. Schonberg, *The Great Pianists,* Londres, Victor Gollancz, 1964, p. 280. Para más ejemplos, véase William S. Newman, *The Sonata since Beethoven,* Chapel Hill, University of North Carolina Press, 1969, pp. 104-108.
36. Para una descripción de Clara Schumann como maestra, véase Nancy B. Reich, *Clara Schumann: The Artist and the Woman,* Ithaca y Londres, Cornell University Press, 1985, pp. 288-296.
37. Eduard Hanslick, *Vienna's Golden Years of Music, 1850-1900,* Henry Pleasants (trad. y ed.), Londres, Victor Gollancz, 1951, pp. 39-44 passim.
38. Amy Fay, *Music-Study in Germany,* Nueva York, Dover, 1965, p. 275.
39. Citado de Harold C. Schonberg, *The Great Pianists,* cit., pp. 355-356.
40. Josef Hofmann, *Piano Playing,* Nueva York, Dover, 1976, pp. 54-55.
41. Duncan Druce, «Historical approaches to violin playing», en John Paynter, Tim Howell, Richard Orton y Peter Seymour (eds.), *Companion to Contemporary Musical Thought,* 2 vols., Londres y Nueva York, Routledge, 1992, II, p. 1013.
42. Alfred Brendel, «Notes on a complete recording of Beethoven's piano works», *Musical Thoughts and Afterthoughts,* Londres, Robson, 1976, pp. 23 y 25.
43. Véanse por ejemplo Richard Taruskin, «The pastness of the present and the presence of the past», en Taruskin, *Text and Act: Essays on Music and Performance,* Nueva York y Oxford, Oxford University Press, 1995, pp. 90-154, y Robert Philip, *Early Recordings and Musical Style: Changing Tastes in Instrumental Performance, 1900-1950,* Cambridge, Cambridge University Press, 1992, pp. 239-240.
44. Jonathan Dunsby, *Performing Music: Shared Concerns,* Oxford, Clarendon Press, 1995, p. 41.
45. Nicholas Kenyon, «Tradition isn't what it used to be», conferencia dictada ante la Royal Phillarmonic Society, Londres, 24 de febrero de 2001.
46. Artur Schnabel, *My life and Music,* Gerrards Cross, Bucks., Colin Smythe, 1970, p. 125.
47. Pierre Baillot, *The Art of the Violin,* Louise Goldberg (trad. y ed.), Evanston, Ill., Northwestern University Press, 1991, pp. 13, 477.

## *Lecturas complementarias*

Dunsby, Jonathan: *Performing Music: Shared Concerns,* Oxford, Clarendon Press, 1995.

Godlovitch, Stan: *Musical Performance: A Phisolophical Study,* Londres y Nueva York, Routledge, 1998.

Howat, Roy: «What do we perform?», en John Rink (ed.), *The Practice of Performance: Studies in Musical Interpretation,* Cambridge, Cambridge University Press, 1995, pp. 3-20.

Kivy, Peter: «Live performances and dead composers: on the ethics of musical interpretation, en Peter Kivy, *The Fine Art of Repetition: Essays in the Philosophy of Music,* Cambridge, Cambridge University Press, 1993, pp. 95-116.

Sessions, Roger: «The performer», en Roger Sessions, *The Musical Experience of Composer, Performer, Listener* [1950], Princeton, Princeton University Press, 1971, pp. 68-86.

Taruskin, Richard: «On letting the music speak for itself», en Richard Taruskin, *Text and Act: Essays on Music and Performance,* Nueva York y Oxford, Oxford University Press, 1995, pp. 51-66.

# 6 *El desarrollo de la habilidad interpretativa*

JANE DAVIDSON

Este capítulo analiza los factores más importantes que intervienen en el desarrollo de las habilidades musicales. En primer lugar, examina el concepto de potencialidad musical; después presenta una investigación sobre las habilidades específicas requeridas para aprender a tocar un instrumento. Por último, examina cuál es la mejor manera de cultivar la habilidad de interpretar para un público. Se ha procurado un equilibrio entre la teoría y la práctica haciendo referencia tanto a las investigaciones publicadas sobre psicología y educación musical como a informaciones anecdóticas y al material que contienen los manuales educativos. Aunque se hacen algunas referencias al aprendizaje musical en otras culturas, este estudio se centra en el desarrollo de la habilidad interpretativa de los niños dentro de un marco cultural occidental.

## *Los roles relativos de la genética y el entorno en la configuración de la potencialidad musical*

Las aparentes diferencias de la capacidad musical entre la población en general han despertado un interés considerable. Gran parte del debate se ha centrado en los roles relativos de las influencias biológicas y ambientales en las habilidades humanas —el debate sobre «naturaleza versus crianza» [1]. Según los descubrimientos científicos, parece ser que los factores genéticos influyen de varias maneras en el desarrollo general:

- *Desarrollo físico por etapas*
  Un ejemplo es la necesidad de gatear antes de andar, o, en términos musicales, del desarrollo gradual de la destreza y coordinación manual y visual para poder sincronizar el arco y las cuerdas en la ejecución del violín.
- *Algunas ventajas físicas*
  Por ejemplo, las personas altas y delgadas suelen ser más aptas para los saltos de altura que las bajas y más bien robustas. En música, las personas de manos anchas tienen un mayor potencial para desarrollarse como pianistas que las de manos estrechas.
- *Algunas ventajas psíquicas*
  Generalmente, por ejemplo, algunas personas «sintonizan» más rápidamente con las tareas que implican la resolución de problemas que otras. Por lo tanto pueden ser capaces de identificar patrones musicales y realizar tareas de discriminación auditiva con mayor rapidez.

El componente biológico es sin duda importante, y los investigadores como Robert Plomin [2] sugieren que la identificación del gen que determina los logros musicales sólo es cuestión de tiempo. No obstante, si admitimos el concepto de la «distribución normal» [3], aparentemente la gran mayoría de la población —alrededor de un 95 por cien— está incluida dentro de la «norma» en cuanto a capacidad y estructura física y mental, sugiriendo que más o menos todos tenemos un potencial similar [4]. Ello se aplica igualmente al desarrollo de las habilidades musicales, en las que se exagera con frecuencia la importancia de las capacidades innatas. Sería de utilidad, por lo tanto, examinar los asuntos relacionados con la influencia del entorno en las habilidades musicales.

Para destacar el papel decisivo del entorno, consideremos el campo políticamente delicado de los estereotipos raciales y las diversas explicaciones del tipo «naturaleza versus crianza» que han sido propuestas para dar cuenta de las habilidades musicales. Los negros africanos, por ejemplo, han sido caracterizados a menudo por los blancos como una raza que lleva «el ritmo en la sangre», mientras que a los japoneses se les considera que poseen por naturaleza capacidades intelectuales elevadas y que despliegan un gran virtuosismo técnico en sus interpretaciones musicales. Aunque los factores genéticos pueden haber intervenido en cierta medida, el análisis de las respectivas culturas y prácticas sociales revela que esos estereotipos se desarrollaron a partir de las reglas socioculturales de participación y conducta. Por ejemplo, para la tribu Anang Ibibo de Nigeria, la música forma parte de la vida cotidiana. Todos los niños han de participar en la música desde su nacimiento. Cuando cumplen cinco años de edad, la gran mayoría de esos niños sabe tocar y cantar un enorme repertorio de canciones, bailes y patrones de percusión. La noción de

ser poco musical o carecer de musicalidad no existe. Además, aunque algunos destacan un poco más que otros, todos han de participar a un nivel elevado. El principio de la distribución normal funciona dentro de esa cultura: la música es para todos, y los logros interpretativos están distribuidos de manera bastante equitativa. En Japón, la disciplina que la sociedad impone al individuo es mucho más severa que la del mundo occidental. Por ello, los niños dedican muchas horas a los deberes de todo tipo, incluyendo la música. Como tiende a haber un enfoque más sistemático de la educación y por lo tanto unos patrones de conducta más disciplinados y estructurados, los niños japoneses adquieren un nivel más elevado de habilidad musical que sus equivalentes occidentales.

Tal vez una de las mejores maneras de considerar la importante influencia del entorno en la adquisición de habilidades musicales sea examinando el caso de un personaje que a primera vista no parece haber tenido apoyo alguno del entorno, pero que llegó a ser uno de los músicos más talentosos del mundo occidental: el trompetista de jazz Louis Armstrong. Nacido en condiciones de extrema pobreza, pasó mucho tiempo arreglándoselas solo en la calle (su madre era prostituta), y no recibió prácticamente ninguna educación formal. Cabe destacar que de niño ni siquiera tenía una trompeta. Sin embargo, si examinamos detalladamente la biografía de Armstrong [5], podemos identificar cinco factores ambientales sumamente importantes:

(1) *Estímulos musicales fortuitos pero frecuentes*
Armstrong creció en Nueva Orleans, rodeado constantemente de jazz.

(2) *Muchas oportunidades, durante un período prolongado, de explorar libremente el medio del jazz y de desarrollar habilidades de presentación de una interpretación*
Armstrong formó y cantó en un coro callejero cuando era niño, y tuvo oportunidades de participar en las actividades colectivas de jazz en las que se toleraban los errores y en las que podía elegir el nivel de riesgo y dificultad de su interpretación. Por lo tanto, pudo aprender informalmente las «reglas» de la armonía, la improvisación y la etiqueta del jazz.

(3) *La oportunidad temprana de experimentar estados emocionales o estéticos sumamente positivos en respuesta a la música*
Podemos suponer que Armstrong desarrolló un «amor» por el jazz basado en cierta experiencia emocional relacionada con la música.

(4) *La oportunidad de acumular muchas horas de práctica*
Armstrong pasaba todo el día con músicos de jazz, y podía pedir prestada una trompeta para tocar.

(5) *Varios factores externos que lo motivaron, como la presencia de un adulto clave*
Armstrong recibió una educación informal de los músicos de jazz con los que pasaba el día, y ese contacto desarrolló su conocimiento musical, sus habilidades interpretativas y su deseo de relacionarse con la música.

Los cinco factores que al parecer fueron decisivos para que Armstrong se interesara por la música y se relacionara con ella han sido estudiados sistemáticamente por psicólogos de la música, y hoy en día se consideran fundamentales para el desarrollo de las habilidades musicales [6]. A continuación, se examinarán en profundidad dichos factores y se harán algunas sugerencias sobre la mejor manera de fomentar las habilidades interpretativas.

### *Oportunidades tempranas para relacionarse con la música*

Existen cada vez más pruebas de que el entorno comienza a influir en el individuo antes de su nacimiento. Por ejemplo, el feto es capaz de oír música y voces varias semanas antes del nacimiento y de reconocer esa música después de nacer [7], lo que implica que ha tenido lugar algún tipo de aprendizaje. También se puede deducir que la experiencia fetal es en gran parte «musical» porque el feto siente los ritmos del cuerpo de su madre y oye las inflexiones —los contornos melódicos— de los sonidos externos, incluyendo el habla y la música. Es posible que la estimulación prenatal sea una de las experiencias ambientales más importantes, y que los niños estimulados de esa manera se beneficien en cuanto a su educación en general y su desarrollo musical en particular. El mundo de la música comercial da por sentada la posibilidad del aprendizaje prenatal, y ofrece a las mujeres embarazadas una variedad de programas musicales especiales que supuestamente tienen un efecto positivo en el niño que está por nacer.

Se ha demostrado que en la infancia el «lenguaje materno» [8] constituye la base para el desarrollo de las habilidades de alternancia recíproca [9]. La comunicación gestual también es de vital importancia, y el ritmo, tempo y dinámica de sus interacciones [10] son determinados por movimientos y prácticas lúdicas como dar palmadas. El «lenguaje materno» y los juegos gestuales se desarrollan para comunicar emociones y necesidades básicas. De esta manera se forma la relación entre la emoción, el gesto físico y el significado musical. Una característica musical expresiva como el *rubato* puede demostrar una relación con las primeras experiencias infantiles. Por ejemplo, las interacciones físicas y vocales se ralentizan cuando están a punto de finalizar, y en la inter-

pia musical [14] y la educación musical [15] exploran la música como un medio para calmar, aliviar y animar a las personas. Algunos de los resultados son sorprendentes: en un caso, una clase de niños difíciles, con «necesidades educativas especiales», mejoró su trabajo y su rendimiento al escuchar obras de Mozart como música de fondo durante sus clases de ciencias.

Se ha sostenido que la respuesta emocional ante un contenido musical constituye un desencadenante fundamental de la motivación musical intrínseca [16]. Un estudio [17] reveló que las personas comprometidas con la música durante toda su vida son mucho más proclives a rememorar las experiencias emocionales que tuvieron en respuesta a un contenido musical durante la primera infancia que aquellos que nunca aprendieron a tocar un instrumento musical más allá de la fase inicial. Por ejemplo, una mujer señaló que había escuchado por primera vez una música para clarinete clásico cuando tenía siete años; el sonido del instrumento le pareció «líquido» y «hermoso», y le hizo sentir un «hormigueo» por todo el cuerpo.

Es importante destacar que, además del contenido musical, el contexto musical constituye un estímulo muy importante de la motivación. Por ejemplo, estar con un amigo íntimo mientras se escucha una pieza determinada puede fomentar una relación futura con esa música, mientras que, al contrario, es posible que evitemos la música que hayamos escuchado en circunstancias desagradables.

### *La práctica para la perfección*

Una serie de estudios han demostrado que la manera más eficaz de adquirir habilidad es la práctica deliberada [18]. En el campo de la música, se ha encontrado una clara relación entre las horas acumuladas de práctica «formal» (escalas, ejercicios técnicos y repertorio) y los logros alcanzados. Por ejemplo, en un estudio sobre alumnos violinistas [19], K. Anders Ericsson et al. descubrieron que los mejores alumnos habían acumulado casi 10.000 horas de práctica formal para la edad de veintiún años, mientras que los alumnos cuyos logros eran mucho menores habían acumulado menos de la mitad. Además, un estudio de John A. Sloboda et al. [20] reveló que los alumnos más sobresalientes no sólo practicaban mucho más que sus compañeros menos exitosos, sino que además había una correlación directa entre las horas de práctica y la calidad interpretativa en los exámenes musicales. Los investigadores estudiaron cinco grupos de jóvenes que abarcaban desde los que habían abandonado la música después de dos años de aprendizaje hasta aquellos que estaban recibiendo una educación musical especializada con vistas a una carrera profesional. Descubrieron que los alumnos de educación especializada avanza-

que los individuos aprenden tareas si valoran la actividad o el resultado, o si anticipan el éxito. La valoración depende de varios tipos de motivación:

- *extrínseca* (cuando las tareas se llevan a cabo por la posibilidad de alguna recompensa externa, como aprobar un examen),
- *social* (el deseo de complacer a otros o encajar con ellos),
- *éxito* (para realzar el ego, ser mejor que otros),
- *intrínseca* (interés en la actividad misma, participar por puro placer personal).

Las investigaciones sobre cómo los niños se vinculan con la música sugieren que las tres fuentes de motivación externas —extrínseca, social y éxito— a menudo preceden y se desarrollan convirtiéndose en motivación intrínseca. Previamente en este capítulo, cuando se consideró la importancia de las respuestas emocionales, se demostró que la motivación intrínseca es el tipo de motivación que facilita un compromiso a largo plazo con la música. Hallam [26] ha escrito detalladamente sobre cómo ayudar a desarrollar la motivación intrínseca, y sugiere las siguientes técnicas de enseñanza:

(1) Los alumnos deben estar expuestos a muchos tipos de música de una manera libre, para abrir las posibilidades de respuesta emocional a la música.
(2) Los alumnos deben participar en las decisiones sobre el estilo y la dificultad de la música que van a aprender y decidir si quieren o no interpretar en público y examinarse.
(3) Para estimular a los alumnos, los profesores han de generar sorpresa, dudas, contradicción y debate.

En otras palabras, las tareas musicales deben ser importantes para los alumnos y han de tener un nivel de complejidad óptimo, es decir, que los enfrente a un reto, pero que sea practicable. Hallam desarrolló sus técnicas observando a muchos alumnos, pero también examinando las características biográficas que comparten diversos músicos afamados [27].

En cuanto a las motivaciones externas, parece ser que los apoyos sociales son los más comunes —por ejemplo, los amigos, la familia y los maestros. Ellos pueden utilizar refuerzos materiales como dulces, dinero, pequeños obsequios o convites, abrazos y sonrisas para animarlos, pero su presencia y participación son lo más importante para el progreso en el aprendizaje.

### *Apoyo social y aprendizaje*

Además de la estimulación temprana del interés musical mediante el «lenguaje materno» y el juego gestual, más tarde, cuando se ha elegido un instrumento, los padres y cuidadores pueden desempeñar un papel muy importante en el desarrollo del vínculo con la música. Una investigación psicológica a gran escala [28] demostró que es bastante común que los niños abandonen el aprendizaje de un instrumento musical por no recibir suficiente apoyo de los padres. Doscientos cincuenta y siete niños, de edades comprendidas entre los ocho y los dieciocho años, que recibían clases instrumentales y poseían diversos niveles de rendimiento fueron entrevistados en relación con el papel que desempeñan los padres y los profesores. El grupo más destacado era el que recibía el apoyo más frecuente y consecuente de sus padres, que a menudo asistían a las clases y tomaban nota de lo que debían practicar sus hijos en casa; después, de vuelta en casa, les apoyaban escuchando o acompañándoles durante las prácticas.

Por lo tanto, los padres deben establecer ciertas actitudes hacia la música. Los hermanos también pueden hacer que el entorno del aprendizaje sea más atractivo, así como las personas que el niño admira y quiere imitar. En el caso de Louis Armstrong, los músicos de jazz proporcionaban un ambiente de seguridad y apoyo en el que los elogios y el éxito desempeñaban un papel de refuerzo. En ese sentido, eran como los profesores del sistema educativo convencional.

### *El papel del maestro*

Las investigaciones [29] han demostrado que los alumnos destacados consideran que sus maestros son divertidos y amables, y también músicos competentes, mientras que los alumnos de menor éxito recuerdan a sus maestros como personas antipáticas e incompetentes. A medida que aumenta la edad de los alumnos, el grupo inferior en cuanto a logros mantiene la misma descripción, mientras que los más destacados comienzan a distinguir entre las cualidades profesionales y personales de sus maestros, considerando más pertinentes las profesionales. Los niños que abandonaron sus clases no establecieron esa distinción. Los que consideraban que la personalidad del profesor era agradable también percibían que era profesionalmente satisfactorio. Podemos observar cuán importantes son los aspectos personales; especialmente a edades tempranas, dichos aspectos pueden motivar al niño a tocar. Ello pone de relieve el clima emocional en el que se acumulan las experiencias musicales. Los niños que desarrollaron un nivel instrumental extraordinario aprendieron en una atmósfera positiva, agradable y despreocupada. Al contrario, el

ambiente de aprendizaje de los niños que se dieron de baja era negativo y angustioso.

Los investigadores también descubrieron que a partir de los once años de edad muchos niños no necesitan fuentes externas de motivación, ya que desarrollan su propio vínculo con el instrumento [30]. En efecto, un estudio sobre músicos adolescentes y sus actividades [31] reveló que su identidad estaba estrechamente vinculada con tocar un instrumento musical. Se ha encontrado que lo mismo es cierto en otros campos, como la música pop, en los que las recompensas musicales llegan a vincularse con la identidad propia [32].

De esta manera, el análisis de las investigaciones revela que la biografía de Armstrong muestra los factores ambientales más importantes relacionados con la adquisición de habilidades musicales. Pero ¿qué podemos decir sobre las características de habilidades mismas?

### *Las habilidades interpretativas*

Según Hallam [33], un primer paso importante en la motivación para el aprendizaje instrumental es elegir el instrumento adecuado. Ello implica elegir un instrumento que le guste al niño, tal vez adecuándose incluso a los estereotipos de género (por ejemplo, las niñas prefieren la flauta, y los niños la batería). Con el instrumento elegido, las prácticas debidamente apoyadas por los padres y un maestro afectuoso y simpático dirigiendo al niño, los siguientes elementos parecen formar la base de las habilidades a desarrollar:

(1) *Habilidades de estructura, notación y lectura*
Éstas se adquieren desarrollando una base de conocimientos sobre las reglas de la estructura musical, como por ejemplo la formación de las frases y el funcionamiento de la armonía; examinando el aspecto que tiene la música sobre el papel, y aprendiendo por lo tanto a descifrar el código de la notación musical; descubriendo cómo leer a primera vista para poder aprender y procesar más música en menos tiempo mediante una lectura más fluida; aprendiendo a memorizar para poder tocar la música sin tener que concentrarse completamente en la notación, y, a partir de todo ello, desarrollando mapas mentales sólidos que relacionen el conocimiento estructural con la ejecución musical.

(2) *Habilidades auditivas*
Las habilidades auditivas son particularmente necesarias para que el alumno pueda desarrollar una buena afinación y calidad de sonido en el instrumento. Adquirir la habilidad de emplear en la práctica el conocimiento sobre la estructura musical mediante la improvisación

es muy importante, como lo es también saber cómo suena la música sin tener que tocarla.

(3) *Habilidades técnicas y motrices*
Además de crear estructuras mentales para comprender la música, es necesario entrenar el cuerpo para que automatice el proceso de ejecución de modo que se pueda lograr fluidez y agilidad.

(4) *Habilidades expresivas*
Implican saber cómo emplear las reglas estructurales para crear una música en la que se percibe un contenido emocional y que es original del intérprete.

(5) *Habilidades de presentación*
Éstas incluyen aprender a presentar tanto la música como el cuerpo del intérprete con seguridad en el escenario, de modo que el intérprete, los cointérpretes y el público perciban una sensación de tranquilidad en torno al acto interpretativo y lo que está comunicando. También incluyen el dominio de las reglas de etiqueta en el escenario —por ejemplo, saludar con una reverencia después de la interpretación.

Existen muchos manuales para desarrollar las subhabilidades enumeradas anteriormente, y hoy en día se reconoce que con una actitud persistente y determinada se pueden desarrollar las tres primeras series de subhabilidades hasta niveles elevados [34]. Sin embargo, desde hace mucho tiempo se considera que las habilidades expresivas son cruciales a la hora de diferenciar a los intérpretes promedio de los excelsos, y que una especie de «don» misterioso e imposible de cultivar determina la medida en que una persona puede ser expresiva musicalmente. No obstante, la investigación [35] revela que los «recursos expresivos» concuerdan con una serie de características. En primer lugar, son *sistemáticos*. Por ejemplo, la ralentización siempre se utiliza en determinados momentos estructurales de una pieza, como por ejemplo los límites de las frases. En segundo lugar, dichos recursos mejoran la *comunicabilidad* de la característica estructural para el que lo percibe, de modo que, por ejemplo, el compás será más claro si se acentúa el primer tiempo. En tercer lugar, esos recursos se emplean de manera *estable,* de modo que un intérprete puede lograr repetidamente efectos expresivos similares en múltiples interpretaciones, a menudo durante períodos prolongados. En cuarto lugar, la *flexibilidad* es posible, y los intérpretes expertos pueden cambiar o exagerar el contorno de una línea expresiva. Por último, esos recursos expresivos se emplean de manera *automática,* y a menudo los intérpretes no son conscientes de lo que están haciendo. Partiendo de estas conclusiones, es evidente que las cinco características anteriores se pueden enseñar dada su estabilidad.

¿Las habilidades de presentación son un factor decisivo para distinguir a los buenos intérpretes de los grandes? Según los estudios publicados, aparentemente no dependen más del «talento» que las otras subhabilidades. Una investigación reciente [36] ha demostrado que las personas aprenden a armonizar entre sí y ajustarse a las convenciones sociales en el contexto de la interpretación, y existen programas de adiestramiento como Eurythmics de Dalcroze que destacan la importancia de las indicaciones no verbales y la interacción interpersonal en el desarrollo de las habilidades de presentación. Sin embargo, como señala perspicazmente Andrew Evans [37], es necesario que nos dediquemos mucho más, cultural y educativamente, al desarrollo de las habilidades de presentación. Por ejemplo, comenta ampliamente las terribles ansiedades, referidas por músicos de todos los niveles, relacionadas con una falta de seguridad en sí mismos en cuanto al contexto y la evaluación de la interpretación. Los estudios publicados sugieren que conocer diversos contextos interpretativos, tocar un repertorio que se conozca bien y tener muchas oportunidades para la interpretación constituyen maneras de superar posibles problemas con las habilidades de presentación.

Una de las mayores dificultades que plantea la forma en que se enseña y se concibe la interpretación como acto interpretativo en la cultura occidental es que tiene lugar en el aislamiento: practicar es una actividad solitaria, y las clases suelen ser individuales. Para desarrollar habilidades interpretativas es necesario luchar contra la tendencia cultural dominante. Los músicos clásicos y los maestros de música del mundo occidental podrían aprender mucho de los Anang Ibibo de Nigeria, aunque la naturaleza de su material musical sea muy diferente. Para los Anang, la música forma parte de la vida diaria: no asusta ni inquieta a nadie, ni es privilegio de unos pocos. La música es para todos, y todos participan como creadores, oyentes e intérpretes.

## *Notas*

1. Véase John A. Sloboda, Jane W. Davidson y Michael J. A. Howe, «Is everyone musical?», *The Psychologist* 7 (1994), pp. 349-354.
2. Robert Plomin, John C. DeFries y John C. Loehlin, «Genotype-environment interaction and correlation in the analysis of human behavior», *Psychological Bulletin* 84 (1977), pp. 309-322.
3. Véase Rita L. Atkinson, Richard C. Atkinson, Edward E. Smith, Daryl J. Bem y Ernest R. Hillgard, *Introduction to Psychology,* 10ª ed., Orlando, Fla., Harcourt Brace Jovanovich, 1990, pp. 779-780.
4. Los extremos superior e inferior de esa distribución presentan respectivamente la misma cantidad de aventajados o incapacitados en alguna forma, pero como la suma de ambos es de tan sólo el 5 por ciento, constituyen casos excepcionales.

5. James L. Collier, *Louis Armstrong: An American Genius,* Nueva York, Oxford University Press, 1983.
6. Véase John A. Sloboda, Jane W. Davidson, Michael J. A. Howe y Derek G. Moore, «The role of practice in the development of expert musical performance», *British Journal of Psychology* 87 (1996), pp. 287-309.
7. Véanse Peter G. Hepper, «An examination of foetal learning before and after birth», *Irish Journal of Psychology* 12 (1991), pp. 95-107, y Jean-Pierre Lecanuet, «Prenatal auditory experience», en Irène Deliège y John A. Sloboda (eds.), *Musical Beginnings: Origins and Development of Musical Competence,* Oxford, Oxford University Press, 1996, pp. 3-34.
8. Es decir, las vocalizaciones de tipo «gu-gu, ga ga» entre un adulto y un niño.
9. Véase Colwyn Trevarthen, «Musicality and the intrinsic motive pulse: evidence from human psychobiology and infant communication», *Musicae Scientiae,* número especial (1999-2000), pp. 155-215.
10. Véase Mechthild Papousek, «Intuitive parenting: a hidden source of musical simulation in infancy», en Irène Deliège y John A. Sloboda (eds.), *Musical Beginnings: Origins and Development of Musical Competence,* cit., pp. 88-114.
11. Véase Hanus y Mechthild Papousek, «Beginning of human musicality», en Reinhart Steinberg (ed.), *Music and the Mind Machine: The Psychophysiology and Psychopathology of the Sense of Music,* Berlín, Springer Verlag, 1995, pp. 27-34.
12. Michael J. A. Howe, Jane W. Davidson, Derek G. Moore y John A. Sloboda, «Are there early childhood signs of musical ability?», *Psychology of Music* 23 (1995), pp. 162-176.
13. Anthony Kemp, *The Musical Temperament,* Oxford, Oxford University Press, 1996, viii, pp. 234-255.
14. E.g. Gary Ansdell, *Music for Life: Aspects of Creative Music Therapy with Adult Clients,* Londres, Jessica Kingsley, 1995.
15. E.g. Anne Savan, «The effect of background music on learning», *Psychology of Music* 27 (1999), pp. 138-146.
16. John A. Sloboda, «Musical Structure and emotional response: some empirical findings», *Psychology of Music* 19 (1991), pp. 110-120.
17. John A. Sloboda, «Empirical studies of emotional response to music», en Mari R. Jones y Stephen Holleran (eds.), *Cognitive Bases of Musical Communication,* Washington, D.C., American Psychological Association, 1992, pp. 33-46.
18. Véase el capítulo 7 de este libro. Un caso examinado en los estudios de psicología que se han publicado es el de S. F., un hombre capaz de aumentar su capacidad normal de memorización de cinco a siete ítems para recordar inmediatamente a ochenta ítems después de 200 horas de práctica sistemática de la tarea. Para un análisis, véase William G. Chase y K. Anders Ericsson, «Skilled memory», en John R. Anderson (ed.), *Cognitive Skills and their Acquisition,* Hillsdale, NJ, Erlbaum, 1981, pp. 141-189.
19. K. Anders Ericsson, Ralf Th. Krampe y Clemens Tesch-Römer, «The role of deliberate practice in the acquisition of expert performance», *Psychological Review* 100 (1993), pp. 363-406.
20. John A. Sloboda, Jane W. Davidson, Michael J. A. Howe y Derek G. Moore, «The role of practice in the development of expert musical performance», *op. cit.*
21. Linda M. Gruson, «Rehearsal skill and musical competence: does practice make perfect?», en John A. Sloboda (ed.), *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition,* Oxford, Clarendon Press, 1988, pp. 91-112.
22. Kacper Miklaszewski, «A case study of a pianist preparing a musical performance», *Psychology of Music* 17 (1989), pp. 95-109.
23. Susan Hallam, «Professional musicians' orientations to practice: implications for teaching», *British Journal of Music Education* 12 (1995), pp. 3-19.

24. Véase Susan Hallam, *Instrumental Teaching: A Practical Guide to Better Teaching and Learning,* Oxford, Heinemann, 1998.
25. Véase Norman Feather (ed.), *Expectations and Actions,* Hillsdale, NL, Erlbaum, 1982.
26. Susan Hallam, *Instrumental Teaching: A Practical Guide to Better Teaching and Learning,* cit., pp. 104-106.
27. Algunos detalles pueden encontrarse, por ejemplo, en John A. Sloboda, Jane W. Davidson y Michael J. A. Howe, «Is everyone musical?», *cit.*
28. Jane W. Davidson, Michael J. A. Howe, Derek G. Moore y John A. Sloboda, «The role of parental influences in the development of musical ability», *British Journal of Developmental Psychology* 14 (1996), pp. 399-412.
29. Jane W. Davidson, Michael J. A. Howe, Derek G. Moore y John A. Sloboda, «The role of teachers in the development of musical ability», *Journal of Research in Music Education* 46 (1998), pp. 141-160.
30. Véase Jane W. Davidson, John A. Sloboda y Michael J. A. Howe, «The role of parents and teachers in rhe success and failure of instrumental learners», *Council for the Bulletin of Research in Music Education* 127 (invierno 1995-1996), pp. 40-44.
31. Jane W. Davidson, «Self and desire: a preliminary exploration of why students start and continue with music learning», *Research Studies in Music Education* 12 (1999), pp. 30-37.
32. Véase Daniel Cavicchi, *Tramps Like Us: Music and Meaning among Bruce Springsteen Fans,* Nueva York, Oxford University Press, 1998.
33. Susan Hallam, *Instrumental Teaching: A Practical Guide to Better Teaching and Learning,* cit., p. 17.
34. Véanse Desmond Sergeant, «Experimental investigation of absolute pitch», *Journal of Research in Music Education* 17 (1969), pp. 135-143, y John A. Sloboda, *The Musical Mind: The Cognitive Psychology of Music,* Oxford, Oxford University Press, 1985.
35. John A. Sloboda y Jane W. Davidson, «The young performing musician», en Irène Deliège y John A. Sloboda (eds.), *Musical Beginnings: Origins and Development of Musical Competence,* cit., pp. 171-190.
36. Jane W. Davidson, «The social in musical performance», en David J. Hargreaves y Adrian C. North (eds.), *The Social Psychology of Music,* Oxford, Oxford University Press, 1997, pp. 209-228.
37. Andrew Evans, *The Secrets of Musical Confidence,* Londres, Thorson, 1994, pp. 79-133.

## *Lecturas complementarias*

Deliège, Irène, y John Sloboda (eds.): *Musical Beginnings: Origins and Development of Musical Competence,* Oxford, Oxford University Press, 1996.

Evans, Andrew: *The Secrets of Musical Confidence,* Londres, Thorson, 1994.

Hallam, Susan: *Instrumental Teaching: A Practical Guide to Better Teaching and Learning,* Oxford, Heinemann, 1998.

Hargreaves, David J.: *The Developmental Psychology of Music,* Cambridge, Cambridge University Press, 1986.

Hargreaves, David J., y Adrian C. North (eds.): *The Social Psychology of Music,* Oxford, Oxford University Press, 1997.

# 7 *Preparándose para interpretar*

STEFAN REID

## *Introducción*

Aunque la interpretación musical es una experiencia sumamente pública, la preparación para una interpretación generalmente tiene lugar en privado, y muchas veces en un completo aislamiento. La mayoría de las interpretaciones sólo se realizan una vez, y casi nunca duran más de un par de horas como mucho, pero se preparan durante días, semanas o incluso años de intenso trabajo. El punto de partida de una interpretación de música occidental suele ser una partitura, que consiste en una serie de instrucciones con diversos grados de imprecisión que el intérprete debe traducir en sonidos. La imprecisión inherente a la notación musical occidental implica que la decodificación de la partitura requiere un considerable aporte y conocimiento por parte del intérprete. Por consiguiente, cada interpretación de una obra es diferente.

El término que más se utiliza para referir el proceso de prepararse para una interpretación es «practicar» (o «estudiar») —pero dicho término puede significar varias cosas. Mientras que los psicólogos definen la práctica como «adquirir habilidades mediante la repetición» [1], la práctica o estudio musical consiste en una variedad de actividades diferentes pero relacionadas entre sí que incluyen la memorización, el desarrollo de la destreza técnica y, finalmente, la planificación de una interpretación. Este capítulo se centra tanto en la cantidad como en la calidad de las prácticas que se realizan para preparar una interpretación. La primera sección examina cuánto necesitan practicar los músicos para desarrollar sus habilidades. La segunda sección analiza cómo practican los músicos eficazmente, enfocando específicamente el desa-

rrollo de la habilidad técnica, la planificación de las interpretaciones y la relación entre esos dos elementos. La última sección presenta un resumen que sugiere diversas maneras en las que los músicos pueden mejorar la eficiencia y eficacia de sus prácticas.

Examinando los temas anteriores, el capítulo hace referencia a dos principales fuentes de información sobre la práctica: los estudios pedagógicos sobre música desde el siglo XVIII hasta hoy, y los estudios psicológicos sobre la adquisición de habilidades en general y en el campo de la destreza musical. Aunque su naturaleza es en gran parte subjetiva y anecdótica, los escritos pedagógicos constituyen la base de la educación musical occidental. Su subjetividad se ha compensado mediante referencias a los escritos psicológicos que, utilizando métodos experimentales, procuran proporcionar un conocimiento más objetivo sobre cómo se adquieren habilidades y cómo practican los músicos.

## *¿Cuánto es suficiente?*

A la mayoría de los estudiantes de música, los maestros le han dicho en algún momento que «no han practicado suficiente». En sus referencias a la práctica, las obras pedagógicas escritas desde el siglo XVIII en adelante incluyen una gama de preceptos sobre cuánto se debe practicar para alcanzar el éxito en la música. Por ejemplo, el respetado profesor de canto Pier Francesco Tosi opinaba que «una hora al día no es suficiente, ni siquiera para la mente más dispuesta» [2]. En su *Klavierschule,* Daniel Gottlob Türk sugiere que «el que aprende a tocar el clavicordio por puro placer hará suficiente practicando dos horas al día» [3]. Johann Joachim Quantz, maestro de flauta de Federico el Grande de Prusia, propuso cuatro horas de práctica al día, dos por la mañana y dos por la tarde [4], mientras que el afamado pianista Johann Nepomuk Hummel opinaba que un máximo de tres horas al día era suficiente «para alcanzar la excelencia» [5]. La discrepancia entre estas sugerencias bien puede hacer que los músicos se pregunten: ¿cuánta práctica es necesaria?

Algunos investigadores que estudian la psicología de la música han procurado abordar esa cuestión. K. Anders Ericsson, Ralf Krampe y Clemens Tesch-Römer realizaron un estudio sobre la cantidad de «práctica deliberada» [6] que realizan tres grupos de violinistas estudiantes durante el transcurso de su educación musical. Su experimento se basó en cálculos aproximados del tiempo que cada alumno había practicado desde su infancia y en diarios de práctica que reflejaban sus actividades musicales actuales. Los alumnos fueron divididos en tres grupos según la evaluación que hicieron sus maestros de sus habilidades y posibles carreras en el futuro: «mejores violinistas»,

«buenos violinistas» y futuros «profesores de música». Los resultados indicaban que, para la época del experimento, los «mejores violinistas» habían acumulado unas 7.410 horas de práctica, los «buenos violinistas», 5.301 horas, y los posibles «profesores de música», 3.420 horas. Los investigadores dedujeron por tanto que el éxito y la habilidad musical, en lugar de ser producto del talento innato, son el resultado de la acumulación de muchas horas de práctica deliberada. En cuanto a la práctica regular, descubrieron que los «mejores» y «buenos» violinistas practicaban una media de 24,3 horas por semana, mientras que el promedio de los futuros profesores de música era de 9,3 horas por semana [7].

Por lo tanto, en respuesta a la pregunta «¿cuánta práctica es necesaria?», los descubrimientos de Ericsson et al. sugieren que la cantidad requerida depende del objetivo primordial del músico. Para alcanzar los más altos niveles de competencia musical, el intérprete ha de dedicar muchas más horas a la práctica que el músico que simplemente desea adquirir un nivel razonable de destreza.

Mientras esta investigación parece sugerir que la cantidad de práctica es la clave del éxito musical, el término «práctica deliberada» destaca la importancia de la calidad y la naturaleza de las actividades de la práctica. Este énfasis en la calidad se repite en todos los escritos sobre pedagogía musical. Hummel refiere la necesidad de un «estudio *atento*» [8], y en su ampliamente respetado *Versuch über die wahre Art das Clavier zu spielen,* C. P. E. Bach comenta que con la «práctica inteligente es fácil lograr lo que nunca se puede alcanzar mediante un excesivo esfuerzo muscular» [9]. Si las incontables horas de práctica acumuladas son un prerrequisito para el éxito musical, la eficacia de esas horas dependerá de la calidad y naturaleza de las actividades prácticas que se realicen. Para obtener los mayores beneficios de sus horas de práctica, los músicos deben desarrollar y perfeccionar sus técnicas de estudio.

## *¿Cómo practican los músicos?*

«Muy pocas personas obtienen con su práctica todos lo beneficios que esperan, porque, en general, practican mal» [10]. Eso comentó el famoso pianista Frédéric Kalkbrenner en su tratado de 1831, *Méthode pour apprendre le piano-forte.* Probablemente es cierto que la mayoría de los músicos podrían mejorar sus técnicas de estudio y con ello aumentar al máximo los resultados del tiempo que pasan estudiando. Como se dijo en la introducción, el músico que se prepara para una interpretación tiene dos objetivos fundamentales: formular la interpretación de una obra musical y desarrollar la destreza técnica necesaria para realizar esa interpretación. Aunque esos dos objetivos están

estrechamente relacionados —de hecho, para una buena interpretación son completamente interdependientes—, son habilidades diferentes que requieren métodos de trabajo diferentes. Eso se demuestra por el hecho de que, como señala el psicólogo John Sloboda [11], un intérprete puede tener una técnica magnífica pero carecer de perspicacia musical, y viceversa. El músico experto es capaz de equilibrar la habilidad técnica con la comprensión interpretativa.

### *El desarrollo de la destreza técnica*

En su obra *We Piano Teachers,* Victor Booth sugiere que durante la práctica «los actos conscientes o voluntarios se vuelven inconscientes o involuntarios» [12]. En otras palabras, una actividad que inicialmente requiere el pensamiento consciente se ensaya hasta que se vuelve automática. Los psicólogos Paul Fitts y Michael Posner han sugerido que la adquisición de una habilidad se realiza en tres fases:

(1) *Fase cognitiva,* una fase inicial que requiere la atención consciente.
(2) *Fase asociativa,* caracterizada por el perfeccionamiento de la actividad y la eliminación de errores.
(3) *Fase autónoma,* en la que la habilidad ya no requiere la atención consciente porque se ha vuelto automática [13].

El progreso de una fase a la siguiente es gradual a medida que la habilidad es transferida de la memoria a corto plazo a la memoria a largo plazo.

El método principal que utilizan los músicos para automatizar una actividad es la repetición de esa actividad. Al repetir una acción o serie de acciones, el intérprete la consolida en la memoria hasta que ya no requiere la atención consciente que se necesita en las fases iniciales del aprendizaje. Las investigaciones sobre la adquisición de habilidades en otros campos de especialización han demostrado que la memoria a corto plazo tiene una capacidad limitada, que se considera generalmente de unos siete «ítems». En el almacenamiento a corto plazo, esos «ítems» pueden ser de tamaños variables, pero se perderán a menos que se transfieran al almacenamiento a largo plazo mediante la repetición. Los constreñimientos de la memoria a corto plazo interfieren en la habilidad de una persona para realizar una tarea compleja como aprender una pieza musical. Para que la práctica repetitiva sea eficaz, es necesario seccionar la tarea compleja en tareas más pequeñas y sencillas, un procedimiento que los psicólogos denominan *chunking* (dividir en trozos grandes o bloques). En términos musicales, puede consistir en practicar sólo

una sección pequeña de la obra y después, cuando ya se domina, incorporarla dentro de un «bloque» más grande, y así sucesivamente. De esta manera, la pieza se fragmenta antes de reconstruirla gradualmente. Este método de trabajo fue recomendado por el pianista inglés James Ching, entre otros [14].

El tamaño de los bloques que se elija para la repetición dependerá de la complejidad de la obra y la habilidad del intérprete. Esto fue demostrado por Linda Gruson en un estudio sobre las prácticas de varios pianistas con diversos niveles de habilidad [15]. El estudio reveló que los pianistas más experimentados dividían la música en bloques más grandes que los novatos. Kacper Miklaszewski [16] ha analizado detalladamente el procedimiento de *chunking* en un estudio sobre un pianista ruso que preparaba la obra «Feux d'artifice» (*Preludios,* Libro II) de Debussy. Durante el proceso, el pianista trabajó la pieza sistemáticamente, estudiando primero bloques pequeños y después, progresivamente, bloques más grandes. Este tipo de enfoque del aprendizaje suele diferenciar el trabajo de los músicos experimentados y novatos. Los novatos generalmente carecen del enfoque sistemático de los intérpretes más experimentados que, como señala Robert Gerle, «separan los problemas y los resuelven uno por uno» [17].

En prácticas posteriores, el pianista ruso del estudio de Miklaszewski no sólo practicaba secciones más largas sino que además las interpretaba a un tempo más rápido. La práctica lenta es, naturalmente, otra manera en que los músicos pueden simplificar una tarea. Dale Reubart la recomienda en su obra *Anxiety and Musical Performance,* donde sugiere específicamente que la práctica lenta durante las primeras fases del aprendizaje puede ayudar a prevenir errores innecesarios, y que el lema de todo intérprete debería ser «hazlo bien la primera vez» [18].

La concentración, por tanto, es un elemento esencial de la práctica técnica, no sólo para asegurar la precisión sino también para mantener la eficiencia. Los músicos experimentados estudian sistemáticamente los problemas que presenta una obra musical, utilizando su concentración para diagnosticarlos y resolverlos. Estos objetivos cambiarán constantemente a medida que progrese el aprendizaje.

En resumen, el desarrollo de la destreza técnica necesaria para una interpretación requiere que la habilidad sea transferida de la memoria a corto plazo a la memoria a largo plazo. La manera más eficaz de realizar esa transferencia es mediante la repetición, que produce sus mejores resultados cuando el material musical se divide en bloques manejables de modo que el intérprete domine un «problema» a la vez. A medida que progresa la práctica, los bloques pueden aumentar de tamaño y el tempo también puede aumentar. La práctica repetitiva se debe supervisar constantemente para evaluar su eficacia y aumentar al máximo sus beneficios.

## *La formulación de las interpretaciones*

Cuando un actor prepara un nuevo papel, no sólo deberá memorizar un texto, sino también decidir cómo lo va a expresar. Las palabras en sí mismas ofrecen posibilidades, pero para realizar todo su significado han de considerarse dentro del contexto de una escena, un acto o toda la obra. En la actuación, el actor comunica al público su conocimiento e interpretación de la obra a través de la manera en que ha decidido expresar el texto. Se puede establecer un paralelismo con el músico que prepara una interpretación, que debe decidir cómo tocará cada nota si quiere transmitir correctamente su papel dentro del contexto inmediato y dentro de la interpretación más amplia de la pieza.

Las interpretaciones musicales se transmiten mediante los parámetros expresivos como el tempo, la dinámica, la articulación y el timbre, entre otros. Las partituras pueden contener una variedad de indicaciones expresivas para ayudar en las decisiones interpretativas del intérprete, pero, como señalé antes, la notación expresiva carece de precisión. El intérprete tendrá que decidir cuán fuerte va a tocar los *forte,* cuánto tiempo mantendrá un calderón o cuán agudo será su *staccato.*

Los intérpretes han de recorrer el difícil camino entre la necesidad de respetar la partitura, que representa las intenciones del compositor, y el deseo de poner en práctica su propia perspicacia creativa. Además, los ideales estéticos varían según la persona y las modas de la época. Por lo tanto, interpretar música es un proceso muy subjetivo y resistente a las recomendaciones normativas. Mientras que los textos pedagógicos y psicológicos contienen instrucciones específicas para el desarrollo de la destreza técnica, las recomendaciones interpretativas son, comprensiblemente, más diversas y menos detalladas.

Tanto los textos pedagógicos como los psicológicos sugieren que escuchar las interpretaciones de otros es la mejor manera de desarrollar la habilidad interpretativa. Tosi recomienda a los alumnos de canto que «escuchen cuanto puedan a los cantantes más famosos, y también a los mejores instrumentistas, porque de la atención que uno les presta al escucharles se obtienen más beneficios que de cualquier tipo de instrucción» [19]. Sloboda se hace eco de esa recomendación al explicar por qué escuchar buenas interpretaciones puede ser beneficioso:

> La pericia interpretativa requiere, en primer lugar, una gran capacidad de escuchar analíticamente para poder captar las diminutas variaciones de tempo e intensidad [i.e. dinámica] que hacen excelsa una interpretación y después *imitarlas.* Según mi propia experiencia, y la de muchos músicos, no existe una manera realmente satis-

> factoria de *describir* las variaciones expresivas de modo que uno pueda incorporarlas en su propia interpretación. Las técnicas expresivas son transmitidas de un músico a otro mediante la demostración. Ésa es la razón por la que los grandes músicos suelen estar muy interesados en escuchar las interpretaciones de otros grandes músicos [20].

Sin embargo, no basta simplemente con imitar la buena interpretación. El músico debe asimilar las características expresivas de la buena interpretación para enriquecer su propio vocabulario interpretativo. Por ejemplo, los músicos pueden emplear la audición como una manera de familiarizarse con un estilo musical desconocido, como observó Susan Hallam en sus encuestas a veintidós músicos profesionales [21].

A los músicos se les recomienda particularmente escuchar buen canto para desarrollar una interpretación melódica eficaz. El pianista húngaro György Sándor comentó que «podemos aprender mucho escuchando y observando a los buenos cantantes, que pueden respirar y frasear con mucha más libertad y espontaneidad que cualquier instrumentista» [22]. Doscientos años antes, C. P. E. Bach había propuesto igualmente que los instrumentistas cantasen melodías, y añadía que «esa forma de aprender es mucho más útil que leer extensos volúmenes o escuchar discursos eruditos» [23]. Al cantar una línea melódica, el intérprete la separa de la técnica, permitiendo así que surja un contorno expresivo que después se puede imitar en el instrumento.

El canto de una línea instrumental podría definirse como un procedimiento comparativo que ayuda a la interpretación. La melodía vocalizada actúa como una imagen mental que enlaza las notas de la línea melódica en una secuencia expresiva. Esa interpretación virtual actúa entonces como el punto de referencia para la verdadera interpretación durante la práctica. Otro procedimiento comparativo que utilizan los músicos a veces es la creación de un programa extramusical como por ejemplo un relato o una secuencia descriptiva o emocional. Al igual que el canto de una línea instrumental, dicho programa establece relaciones entre los diversos elementos de la obra musical y permite al intérprete realzar la importancia de los eventos musicales por separado dentro del contexto de la interpretación general.

Otro recurso mediante el cual los músicos formulan sus interpretaciones es el análisis, que a menudo adopta la forma de un «estudio considerado de la partitura» [24], más que de una técnica académica rigurosa. Algunos intérpretes analizan para obtener un conocimiento y comprensión detallados de la obra que luego utilizarán para sus decisiones interpretativas, mientras que otros utilizan el análisis como un medio para resolver problemas interpretativos específicos.

En las obras pedagógicas se pueden encontrar ejemplos de cómo analizan los músicos. Las sugerencias interpretativas de Quantz requieren que el intérprete realice algún tipo de análisis armónico para generar expresión; por ejemplo, afirma que «las disonancias deben tocarse más fuerte que las consonancias» [25]. Otro ejemplo de cómo pueden analizar los músicos es el de Fanny Waterman, profesora de piano y fundadora del Concurso Internacional de Piano de Leeds, que escribe: «Así como las oraciones contienen palabras clave, cada melodía tiene un clímax —aunque cada artista tendrá naturalmente su opinión sobre la ubicación de ese clímax. Los pianistas deben aprender a avanzar rítmicamente hacia el clímax, y no detenerse nunca, ni siquiera en un silencio o una pausa» [26]. Por lo tanto, para tocar «hacia el clímax», los músicos han de realizar algún tipo de análisis para decidir dónde se ubica ese clímax.

A pesar de la naturaleza subjetiva de la interpretación, existen recursos específicos mediante los cuales los músicos pueden desarrollar sus habilidades interpretativas y mejorar su habilidad para formular interpretaciones. Al interpretar una obra, el intérprete debe disponer sus diferentes elementos en una interpretación coherente; algunas técnicas útiles para mejorar ese proceso de perfeccionamiento son escuchar buenas interpretaciones, cantar, establecer analogías y analizar.

### *La relación entre la técnica y la interpretación*

Al preparar una obra para su interpretación, el músico debe negociar entre los procesos requeridos para formular la interpretación y los que se requieren para desarrollar las habilidades técnicas adecuadas. Wicinski investigó cómo organizan los músicos sus prácticas para poder satisfacer ambos requisitos [27]. Basó sus resultados en entrevistas realizadas a importantes pianistas rusos, quienes mostraron dos métodos de trabajo diferentes. Un grupo sugirió que sus prácticas podían dividirse en tres fases:

(1) Obtener una perspectiva general de la obra y desarrollar ideas interpretativas iniciales.
(2) Superar las dificultades técnicas de la obra.
(3) Combinar las dos primeras fases y perfeccionar la interpretación.

El otro grupo no estableció fases en sus prácticas y manifestó que simplemente estudiaban la pieza de principio a fin.

Obtener una perspectiva general de una obra antes de comenzar las prácticas ofrece dos ventajas básicas. En primer lugar, proporciona un medio para planificar las sesiones de práctica de modo que las dificultades técnicas se

puedan identificar en una primera fase, lo cual permite al intérprete idear la manera de superarlas. En segundo lugar, permite al intérprete formular ideas interpretativas iniciales antes de realizar las prácticas técnicas de modo que estas últimas estén fundamentadas en los requerimientos interpretativos. La perspectiva general se puede obtener con o sin el instrumento, dependiendo de las habilidades y los métodos de trabajo preferidos de cada músico. Sin embargo, en ambos casos requiere una participación mental consciente en el proceso de la práctica, fomentando por lo tanto la precisión y la eficiencia, además de destacar la importancia de la interpretación como el objetivo fundamental de la ejecución musical.

Las diversas maneras en que surgen las interpretaciones durante las prácticas difiere considerablemente de un músico a otro. En su estudio de encuestas, Hallam descubrió tres enfoques prácticos diferentes de la formulación de las interpretaciones durante las prácticas: un grupo de músicos empleaba el análisis, otro grupo utilizaba una mezcla de análisis e intuición y un tercer grupo trabajaba únicamente sobre una base intuitiva [28], es decir, no prestaban una atención consciente a la interpretación durante la práctica, sino que la dejaban surgir «automáticamente» como un producto de la práctica técnica. En lugar de reflejar cierto poder místico o habilidad natural, la intuición puede ser la respuesta interpretativa automatizada de un intérprete a determinado aspecto de una obra o partitura musical; John Rink lo denomina «intuición instruida» [29]. Los intérpretes expertos han acumulado muchos años de experiencia tocando y escuchando una amplia gama de músicas, lo que resulta en una habilidad «intuitiva» sumamente desarrollada. Los músicos pueden emplear conscientemente el análisis cuando falla la intuición o como un recurso para desarrollar las interpretaciones en general, en oposición a los detalles expresivos de la ejecución.

Dado que los métodos de trabajo de un intérprete evolucionan con los años de experiencia musical, no es de extrañar que los músicos exhiban una amplia variedad de estilos de trabajo, especialmente en cuanto a la formulación de interpretaciones. Sin embargo, una característica que comparten los métodos de práctica de los intérpretes más experimentados es el enfoque sistemático del aprendizaje musical, que se demuestra en una conciencia de los objetivos de la práctica y una habilidad para considerar estrategias musicales eficaces que les permitan alcanzar esos objetivos.

### *Resumen: eficiencia y eficacia en la práctica*

En la preparación de una interpretación musical, el músico aborda una partitura y después la traduce en sonidos mediante un instrumento o la voz. Para

ello debe desarrollar las habilidades técnicas apropiadas para tocar o cantar las notas y realizar su propia interpretación. Las prácticas técnicas e interpretativas requieren técnicas diferentes pero relacionadas entre sí.

A los músicos estudiantes y aficionados les resultará beneficioso estudiar los métodos de trabajo de los intérpretes experimentados y profesores y reconocer que la práctica más productiva es aquella en la que hay una participación mental consciente. Bien sea para desarrollar la técnica o para formular una interpretación, la práctica será más eficaz si el músico tiene objetivos claros y factibles y ha elegido los medios para realizarlos. El enfoque más común es simplificar la tarea dividiendo el material musical en bloques suficientemente pequeños como para asegurar un progreso.

En este sentido, la práctica se puede considerar una actividad mediante la cual el intérprete identifica los problemas y busca la manera de resolverlos. Los intérpretes deben supervisar sus propias acciones para evaluar la eficacia de sus técnicas de práctica y para evitar los errores. Por otra parte, pueden mejorar la eficiencia de sus prácticas planteándose continuamente las siguientes preguntas mientras practican:

(1) ¿Qué es lo que espero lograr?
(2) ¿Cuáles son los métodos que puedo emplear para lograr mi objetivo?
(3) ¿Han tenido éxito mis métodos?

## *Notas*

1. Véase Allen Newell y Paul Rosenbloom, «Mechanisms of skill acquisition and the law of practice», en John R. Anderson (ed.), *Cognitive Skills and their Acquisition,* Hillsdale, NJ, Erlbaum, 1981, p. 1; compárense las consideraciones sobre la práctica de los capítulos 9 y 10.
2. Pier Francesco Tosi, *Observations on the Florid Song,* John Galliard (trad.), Michael Pilkington (ed.), Londres, Stainer and Bell [1723/1742], 1987, p. 27.
3. Daniel Gottlob Türk, *School of Clavier Playing,* Raymond H. Haggh (trad.), Lincoln, Nebr., University of Nebraska Press [1789], 1982, p. 19.
4. Johann Joachim Quantz, *On Playing the Flute,* 2ª ed., Edward R. Reilly (trad.), Londres, Faber and Faber [1752], 1985, p. 118.
5. Johann Nepomuk Hummel, *A Complete Theoretical and Practical Course of Instructions on the Art of Playing the Piano Forte,* 3 vols., Londres, T. Boosey & Co. [1828], 1830, I, p. iii.
6. Definieron la «práctica deliberada» como una «actividad sumamente estructurada con el objetivo explícito de mejorar algún aspecto de la interpretación»; véase Ralf Th. Krampe y K. Anders Ericsson, «Deliberate practice and elite musical performance», en John Rink (ed.), *The Practice of Performance: Studies in Musical Interpretation,* Cambridge, Cambridge University Press, 1995, p. 86.
7. *Ibid.*, p. 93.
8. Johann Nepomuk Hummel, *A Complete Theoretical and Practical Course of Instructions on the Art of Playing the Piano Forte,* cit., I, p. iv.

9. Carl Philipp Emanuel Bach, *Essay on the True Art of Playing Keyboard Instruments [Versuch über die wahre Art das Clavier zu spielen],* William Mitchell (trad. y ed.), Londres, Eulenburg [1753, 1762], 1949, p. 101.
10. Frédéric Kalkbrenner, *Learning the Pianoforte,* Sabilla Novello (trad.), Londres, J. Alfred Novello, 1866, p. 19.
11. John A. Sloboda, *The Musical Mind: The Cognitive Psychology of Music,* Oxford, Oxford University Press, 1985, p. 90.
12. Victor Booth, *We Piano Teachers,* Londres, Hutchinson [1946], 1971, p. 77.
13. Paul M. Fitts y Michael I. Posner, *Human Performance,* Belmont, Calif., Brooks/Cole, 1967, pp. 11-15.
14. James Ching, *The Amateur Pianist's Companion,* Oxford, Hall, 1950, p. 102.
15. Linda M. Gruson, «Rehearsal skill and musical competence: does practice make perfect?», en John A. Sloboda (ed.), *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition,* Oxford, Clarendon Press, 1988, pp. 91-112.
16. Kacper Miklaszewski, «A case study of a pianist preparing a musical performance», *Psychology of Music* 17 (1989), pp. 95-109.
17. Robert Gerle, *The Art of Practising the Violin,* Londres, Stainer Books, 1987, p. 17; cfr. el capítulo 8 de este libro.
18. Dale Reubart, *Anxiety and Musical Performance: On Playing the Piano from Memory,* Nueva York, Da Capo Press, 1985, p. 79; véase también el capítulo 9 de este libro.
19. Pier Francesco Tosi, *Observations on the Florid Song,* cit., p. 37.
20. John A. Sloboda, *The Musical Mind: The Cognitive Psychology of Music,* cit., p. 88.
21. Susan Hallam, «Professional musicians' approaches to the learning and interpretation of music», *Psychology of Music* 23 (1995), pp. 111-128.
22. Gÿorgy Sándor, *On Piano Playing,* Nueva York, Collier Macmillan, 1981, p. 211.
23. Carl Philipp Emanuel Bach, *Essay on the True Art of Playing Keyboard Instruments [Versuch über die wahre Art das Clavier zu spielen],* cit., p. 152.
24. Véase John Rink, reseña de Wallace Berry, *Musical Structure and Performance,* en *Music Analysis* 9 (1990), p. 323; véase también el capítulo 3 de este libro.
25. Johann Joachim Quantz, *On Playing the Flute,* cit., 254.
26. Fanny Waterman, *On Piano Teaching and Performance,* Londres, Faber and Faber, 1983, p. 34; cfr. los comentarios de Rachmaninoff citados en la p. 271 de este libro.
27. A. A. Wicinski, «Psichologiceskii analiz processa raboty pianista-ispolnitiela nad muzykalnym proizviedieniem [A psychological analysis of how the pianist-performer works on a musical composition], *Izviestia Akademii Piedagogiceskich Nauk Vyp* 25 (1950), pp. 171-215, citado en Kacper Miklaszewski, «A case study of a pianist preparing a musical performance», *cit.,* p. 96.
28. Susan Hallam, «Professional musicians' approaches to the learning and interpretation of music», *cit.,* pp. 117-121.
29. John Rink, reseña de Wallace Berry, *Musical Structure and Performance,* cit., p. 328; véase también el capítulo 3 de este libro.

## *Lecturas complementarias*

Gruson, Linda M.: «Rehearsal skill and musical competence: does practice make perfect?», en John A. Sloboda (ed.), *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition,* Oxford, Clarendon Press, 1988, pp. 91-112.

Jørgensen, Harald, y Andreas C. Lehmann (eds.): *Does Practice Make Perfect? Current Theory and Research on Instrumental Music Practice,* Oslo, Norges Musikkhøgskole, 1997.

Krampe, Ralf Th., y K. Anders Ericsson: «Deliberate practice and elite musical performance», en John Rink (ed.), *The Practice of Performance: Studies in Musical Interpretation,* Cambridge, Cambridge University Press, 1995, pp. 84-102.

Miklaszewski, Kacper: «A case study of a pianist preparing a musical performance», *Psychology of Music* 17 (1989), pp. 95-109.

Reubart, Dale: *Anxiety and Musical Performance: On Playing the Piano from Memory,* Nueva York, Da Capo Press, 1985.

Sloboda, John A.: *The Musical Mind: The Cognitive Psychology of Music,* Oxford, Oxford University Press, 1985.

# 8 La memorización de la música

AARON WILLIAMON

Interpretar música de memoria es una actividad que puede alcanzar niveles de exigencia sumamente elevados. No sólo presenta la dificultad inicial de retener miles de notas y complejas estructuras musicales, sino también la tarea igualmente ardua de recordarlas y ejecutarlas en situaciones interpretativas de mucha tensión. Casi siempre, dichas exigencias han hecho que los intérpretes acumulen horas de práctica meramente repetitiva, intentando desarrollar múltiples maneras de recordar la música para que su interpretación pueda continuar pase lo que pase [1]. Tales estrategias, sin embargo, pueden ser ineficaces y no garantizar una memoria perfecta. Como consecuencia, los músicos, profesores e investigadores han intentado encontrar la respuesta a dos preguntas: ¿por qué los intérpretes memorizan la música? y ¿cuál es la manera más eficiente y eficaz de hacerlo? Las respuestas a estas preguntas se han obtenido tradicionalmente de un extenso corpus de evidencias incoherentes y anecdóticas. Este capítulo abordará ambas preguntas desde la perspectiva ofrecida por estudios recientes que han examinado más sistemáticamente la memoria musical.

## *¿Por qué los músicos interpretan de memoria?*

En 1828, Clara Schumann sentó un precedente para los intérpretes musicales al tocar en público sin partitura. Poco después, Franz Liszt, con su «predilección por el espectáculo, aprovechó la oportunidad para convertir este nuevo acontecimiento en un ritual emocionante» [2]. Una vez causó verdadero

furor lanzando no sólo sus guantes blancos a las primeras filas del público —un gesto por el que era conocido—, sino también su partitura. Semejantes efectos teatrales no eran bien recibidos por los críticos. De hecho, las interpretaciones de memoria fueron consideradas ostentosas y de mal gusto incluso hasta finales del siglo XIX. En su libro *With Your Own Two Hands,* Seymour Bernstein describe sucintamente cómo la percepción y aceptación de la interpretación memorizada han cambiado desde el siglo XIX:

> En 1861, sir Charles Hallé, el virtuoso británico, dio el primero de una serie de conciertos dedicados a todas las Sonatas de Beethoven. Tocó de memoria. Para el tercer concierto, se rindió ante los críticos londinenses y apareció en escena con las partituras. Sin embargo... siguió tocando de memoria. Aún muchos años después, en 1870, el dr. Hans von Bülow, cuyas proezas de memoria eran legendarias, se encontró con que los críticos londinenses se oponían tanto como entonces a las interpretaciones de memoria. En este caso, sin embargo, el *Daily Telegraph* censuró más al público que al mismo von Bülow: su memoria, y no su musicalidad, es la que parece haber causado la mayor impresión, reclamaron los críticos. No fue sino a finales del siglo XIX cuando tocar de memoria llegó a considerarse como una práctica seria y no mero sensacionalismo. Indudablemente, hoy en día es mucho más inusual que un solista toque con partitura que sin ella [3].

En efecto, tocar de memoria se ha convertido en una medida de la competencia profesional para solistas de todo tipo —pianistas, violinistas, cantantes y trompetistas por igual. Hacerlo, sin embargo, puede resultar arduo y provocar ansiedad. ¿Por qué entonces los intérpretes insisten en tocar de memoria? ¿Lo hacen sólo para impresionar y deslumbrar al público o existe una justificación musical más satisfactoria? Varios autores han apoyado la segunda razón. Edwin Hughes, por ejemplo, sugirió en 1951 que «tocar con un fajo de partituras» obstaculiza «la absoluta libertad de expresión y la relación psicológica más directa con el público» [4]. Muchos intérpretes han estado abiertamente de acuerdo con esta opinión, y han afirmado que tocar de memoria les permite crear y comunicar interpretaciones musicales nuevas y desechar la innecesaria muleta psicológica y musical de la página impresa. Sin embargo, no todos los intérpretes han tenido una opinión tan firme sobre la memorización; entre las excepciones notables se incluyen el famoso pianista francés Raoul Pugno, Béla Bartók y Dame Myra Hess, quienes tocaban en público con sus partituras.

Un estudio reciente ha investigado sistemáticamente el valor de la interpretación de memoria analizando cuatro preguntas:

(1) ¿Ofrece realmente la interpretación de memoria una «libertad de expresión absoluta»?

(2) ¿Proporcionan las interpretaciones memorizadas la «relación psicológica más directa con el público»?
(3) Si la interpretación de memoria ofrece una «libertad de expresión absoluta» y la «relación psicológica más directa con el público», ¿tienen más peso estas ventajas que las horas adicionales de práctica requeridas para lograr tales proezas?
(4) ¿Influye la formación musical en la habilidad para distinguir las interpretaciones memorizadas de las no memorizadas? [5]

Para responder a estas preguntas, se le pidió a un grupo de personas que viesen diversas interpretaciones grabadas en vídeo de una violonchelista tocando los Preludios de las Suites para Violonchelo núms. 1, 2 y 3 de Bach. Debían calificarlas según el conocimiento musical del intérprete, su capacidad comunicativa, su nivel técnico y la calidad general de la interpretación.

Las interpretaciones, sin embargo, eran de cinco tipos diferentes en cuanto a la memorización, la información visual recibida por las personas que vieron los vídeos y el tiempo dedicado a la preparación de las interpretaciones. Por otra parte, los espectadores fueron clasificados en músicos y no músicos con el fin de determinar si la formación musical influye en la calificación de las interpretaciones. Las características específicas de los cinco tipos de interpretaciones están detalladas en la Tabla 8.1, y la información visual que se le ofrecía al grupo de espectadores se muestra en la Figura 8.1.

Las calificaciones de las interpretaciones fueron examinadas y comparadas para investigar y obtener posibles respuestas a las cuatro preguntas expuestas anteriormente. ¿Ofrecen las interpretaciones memorizadas una «libertad de expresión absoluta»? Libertad, sí; absoluta, no. El estudio revela que los espectadores calificaron las interpretaciones memorizadas más favorablemente que las no memorizadas, independientemente de que un atril obstaculizara o no la imagen de la violonchelista.

Además, las personas con formación musical calificaron aún más favorablemente las interpretaciones memorizadas en cuanto a la capacidad comunicativa de la violonchelista, lo que indica que los músicos pueden estar mejor preparados para percibir los sutiles signos comunicativos que forman parte de las interpretaciones de memoria. Estas respuestas sugieren que, hasta cierto punto, los intérpretes pueden expresarse más libremente cuando tocan sin partitura. Sin embargo, para que esa libertad de expresión sea «absoluta», como propone Hughes, debería ser similarmente identificable por todos los oyentes, es decir, independiente de su formación musical.

¿Ofrecen las interpretaciones memorizadas la «relación psicológica más directa con el público»? Sí, especialmente para los integrantes del público con formación musical. Todos los espectadores calificaron más positivamente

TABLA 8.1 *Características de las interpretaciones, incluyendo una descripción de los cinco tipos de interpretación e indicaciones sobre la presencia o no de un atril y una partitura*

| *Tipo de interpretación* | *Descripción* | *Atril* | *Partitura* |
|---|---|---|---|
| 1 | La violonchelista tocó los Preludios cuando se sintió capaz de ofrecer una interpretación pública satisfactoria. Para estas interpretaciones iniciales, se le pidió que no memorizara las piezas. La imagen de la violonchelista estaba parcialmente tapada por un atril (véase Figura 8.1a). | √ | √ |
| 2 | Los Preludios fueron filmados de nuevo cuando la intérprete podía tocarlos de memoria (aproximadamente un mes después de las interpretaciones del tipo 1). Aunque estaba vacío, el atril que seguía delante de la violonchelista daba la impresión de que ésta estaba leyendo una partitura (véase Figura 8.1a). | √ | × |
| 3 | La interpretación del tipo 3 era la misma que la del tipo 2, excepto que se había retirado el atril (véase Figura 8.1b). | × | × |
| 4 | Se le pidió a la violonchelista que ensayara y tocara de nuevo los Preludios con las partituras. La imagen de la intérprete era similar a la de las interpretaciones de los tipos 1 y 2, ya que un atril la tapaba parcialmente (véase Figura 8.1a). | √ | √ |
| 5 | La interpretación del tipo 5 era igual que la del tipo 4, excepto que la cámara se colocó del lado derecho de la violonchelista, por lo que se excluye el atril de la imagen (véase Figura 8.1c). | × | √ |

las interpretaciones sin atril que aquellas en las que aparece el atril, independientemente de que fuesen memorizadas. Esta respuesta fue aún más categórica para los espectadores músicos, lo que sugiere que su formación musical les puede haber inculcado una inclinación a favor de las interpretaciones sin atril. Las investigaciones futuras podrían aclarar la naturaleza de esta predis-

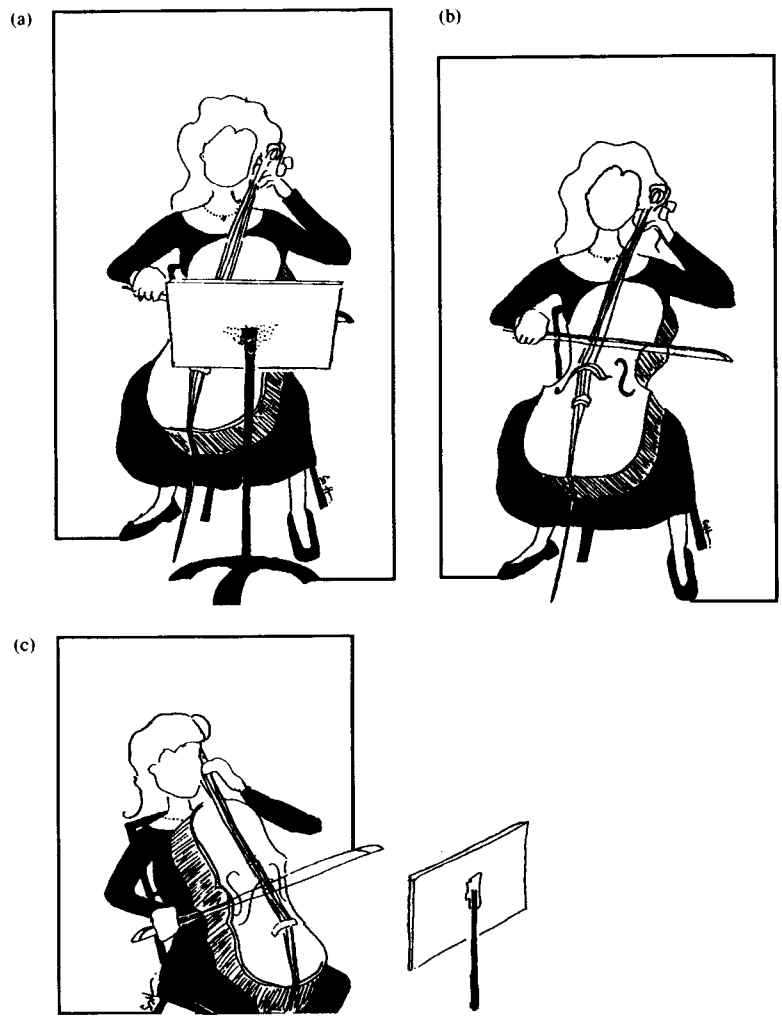



FIGURA 8.1 *La imagen de la violonchelista durante las interpretaciones de los tres Preludios de las Suites núms. 1, 2 y 3 de J. S. Bach. Los espectadores vieron las interpretaciones filmadas de cada Preludio. En algunas de ellas (las de los tipos 1, 2 y 4), la imagen de la violonchelista estaba parcialmente tapada por un atril (Figura 8.1a). En otras interpretaciones (las del tipo 3), la imagen de la violonchelista era como la de una interpretación memorizada, ya que no estaba obstaculizada por el atril (Figura 8.1b). Finalmente, en la interpretación del tipo 5, la intérprete fue filmada desde su lado derecho, excluyendo el atril de la imagen (Figura 8.1c).*

posición, por ejemplo, pidiéndole a un grupo de músicos que califiquen interpretaciones después de informarles correcta o incorrectamente de que dichas interpretaciones son memorizadas.

¿El tiempo de preparación adicional que requiere una interpretación de memoria vale la pena? Sí. Esta respuesta se obtuvo comparando las calificaciones de las interpretaciones no memorizadas anteriores al mes adicional de práctica (es decir, las interpretaciones del tipo 1) con las posteriores (tipos 4 y 5). Todos los espectadores —músicos y no músicos por igual— calificaron más favorablemente las interpretaciones posteriores a la práctica adicional, especialmente con respecto al conocimiento musical y la capacidad comunicativa. Es necesaria una investigación ulterior para saber si esta respuesta se aplica igualmente a otros intérpretes. La intérprete que participó en este estudio consideró que su capacidad técnica había sido una gran ventaja en su ejecución, y no es de extrañar que interpretara las piezas en cuanto satisfizo los requerimientos técnicos. Sin embargo, los intérpretes que no poseen el mismo nivel de competencia técnica pueden mejorar todos los aspectos interpretativos durante un mes adicional de práctica. A la inversa, los que poseen mayores habilidades musicales o comunicativas pueden no mostrar ningún progreso. Para identificar estas diferencias individuales se deberá investigar y comparar sistemáticamente la práctica que conduce a las interpretaciones tanto memorizadas como no memorizadas en estudios futuros.

Este estudio proporciona un examen inicial y empírico de la opinión de que las interpretaciones memorizadas son más expresivas y comunicativas que las no memorizadas. Sus resultados pueden aplicarse, aunque con discreción, tanto a la interpretación como a la pedagogía musical. Por ejemplo, se puede asegurar a los intérpretes que las interpretaciones de memoria ofrecen experiencias más intensas al público, bien sea mediante medios auditivos o visuales. Por consiguiente, los profesores deben preparar a sus alumnos para la exigente expectación del público —especialmente el que está formado por músicos— ofreciéndoles más oportunidades de prepararse para las interpretaciones de memoria. Además, una vez que las investigaciones ulteriores hayan identificado el papel específico que desempeña la información visual en la valoración por parte del público de las interpretaciones memorizadas, los intérpretes y profesores podrán obtener suficiente información para decidir si los atriles deberían excluirse completamente del escenario o si con unos atriles más pequeños y menos imponentes se puede ofrecer una experiencia igualmente satisfactoria para el público.

A pesar de las deducciones anteriores, este estudio presenta varias limitaciones que hay que destacar. En primer lugar, el intérprete, a pesar de su elevado nivel interpretativo, no era una solista eminente. Las ventajas de interpretar de memoria pueden ser aún mayores para los expertos. En segundo

lugar, la interpretación de memoria puede resultar más ventajosa en ciertas piezas musicales. Las obras elegidas para este estudio fueron seleccionadas por la coherencia en cuanto a compositor, estilo, dificultad y longitud. Examinar piezas que difieren en estos aspectos permitirá demostrar con mayor precisión el valor de la interpretación de memoria. Memorizar una composición atonal, por ejemplo, puede ser una tarea pesada y abrumadora si el intérprete se ve obligado a hacerlo nota por nota. Semejante procedimiento aumentaría inevitablemente las exigencia de los procesos mentales durante la interpretación y por lo tanto reduciría —o eliminaría del todo— las ventajas de interpretar de memoria. En tercer lugar, las interpretaciones grabadas en vídeo no ofrecen las condiciones óptimas de visualización a los espectadores. Algunos aspectos comunicativos (por ejemplo, los movimientos sutiles de los ojos, las expresiones faciales y la intensidad de la respiración) pueden disminuir o ser eliminados en el proceso de la filmación. Abordando las limitaciones mencionadas, las investigaciones futuras podrán revelar con precisión cómo el público evalúa y valora la interpretación, proporcionando información que ayude a los intérpretes a elegir o no la memorización.

## *¿Cómo interpretan de memoria?*

### *Métodos de memorización musical*

Edwin Hughes y Tobias Matthay fueron dos de los primeros músicos y pedagogos que escribieron ampliamente sobre cómo los músicos memorizan las composiciones [6]. Hughes, por ejemplo, destacó tres formas principales de memorización musical: auditiva, visual y cinestésica. La *memoria auditiva* permite escuchar composiciones en el «oído interior», anticipar eventos que se aproximan en la partitura y realizar evaluaciones simultáneas del progreso de una interpretación. La *memoria visual* consiste en imágenes del papel escrito y otras características de la interpretación que se retienen en la mente. Los pianistas y otros teclistas, por ejemplo, pueden recordar posiciones «de la mano y los dedos, el aspecto de los acordes sobre el teclado y el patrón que forman las diversas figuras cuando se tocan» [7]. La *memoria cinestésica* (es decir, la memoria muscular o táctil) permite a los intérpretes tocar notas automáticamente. En el caso de los pianistas, se adquiere con un intenso entrenamiento de los dedos, las muñecas y los brazos, y puede existir en dos formas: (1) posición y movimiento de nota a nota y (2) sentido de la resistencia de la tecla. Por otra parte, Hughes y Matthay insisten en que «no es posible una memorización realmente inteligente sin un conocimiento de la armonía y la forma musical, y, para las formas de composición más estrictamente polifó-

nicas, también del contrapunto y la fuga» [8]. Ni Hughes ni Matthay clasificaron este conocimiento como un método independiente para memorizar música. Más bien, simplemente insistieron en que las memorias auditiva, visual y cinestésica no pueden funcionar correctamente sin esos conocimientos. Sin embargo, otros intérpretes y pedagogos posteriores lo han referido explícitamente como *memoria conceptual* para destacar su importancia en el proceso de hacer que la «seguridad sea más segura» [9].

Indudablemente, las interpretaciones de memoria sólo pueden resultar de las complejas interacciones entre todos los métodos de memorización descritos anteriormente. Casi ningún músico puede ejecutar una composición nota por nota sin antes —consciente o inconscientemente— escuchar la música, observar las notas en la página impresa, ver y sentir los movimientos musculares que se producen al tocar y estudiar la partitura hasta cierto punto. Por lo tanto, Hughes y Matthay recomendaron que cada músico se centrase en el método o la combinación de métodos que le resultase más fácil y seguro.

### *Entrevistas con músicos*

Dos investigaciones recientes explican cómo los músicos utilizan estos métodos cuando interpretan. Susan Hallam entrevistó a veintidós músicos profesionales y cincuenta y cinco novatos [10]. Las entrevistas revelaron una diversidad individual considerable en el uso de las memorias auditiva, visual y cinestésica. Un músico, por ejemplo, declaró que «a veces me resulta muy útil la memoria visual, el aspecto de la página [...] Mi memoria visual no es nada buena, excepto para la música». Otro afirmó que «una vez que te encuentras ejecutando un pasaje rápido, o uno en el que la música fluye continuamente, el oído tiende a asumir el control» [11].

Hallam descubrió que todos los músicos profesionales adoptaban una especie de estrategia «analítica» consciente que implicaba desarrollar un conocimiento de la estructura de la música. Un director señaló: «no creo que el método que yo utilizo para aprenderme las partituras [...] sea de ninguna utilidad para un instrumentista, ya que lo hago enteramente por análisis, que para mí resulta infalible» [12]. También afirmaron emplear una codificación múltiple de la información (es decir, una combinación de memoria auditiva, visual y cinestésica con estrategias analíticas) para proporcionar una mayor seguridad durante la interpretación. Por ejemplo, uno de los músicos profesionales comentó:

> Necesito comprender lo que ocurre armónicamente, o si es una melodía con una secuencia definida. No tocaría simplemente con la esperanza de recordar la músi-

> ca [...] Hay piezas que he tocado tantas veces a lo largo de los años que simplemente sé cómo van y puedo tocarlas. Eso no es analítico, pero no me gustaría en absoluto salir a escena y tocarlas porque creo que es posible que la memoria automática me falle [13].

Independientemente de que los músicos fuesen profesionales o novatos, Hallam encontró que sus estrategias dependían frecuentemente de la naturaleza de la obra que iban a interpretar. Informó que algunos músicos se sentían seguros confiando en procesos automáticos para memorizar piezas breves, pero para las obras más largas y complejas, como un concierto, muchos adoptaban un enfoque más analítico. Un músico señaló que

> depende de lo que se trate. Si tiene melodías «fijas», puedo aprenderla en un día. Pero si la música es moderna, comienzo por el principio, la trabajo, y entonces es cuando interviene la práctica, porque me aseguro de que realmente la puedo tocar. Entonces la estudio una y otra vez [...] Y después creo una imagen mental [14].

En un estudio similar, Rita Aiello también examinó los métodos utilizados por músicos profesionales para memorizar música [15]. Entrevistó a siete pianistas concertistas para investigar cómo utilizan la memoria auditiva, visual y cinestésica y cómo enfocan la memorización de diversos tipos de música (por ejemplo, composiciones polifónicas y atonales). Además, examinó el papel que desempeñan la digitación, los recursos mnemotécnicos, la improvisación y el conocimiento de la estructura musical en la memorización. Descubrió varios temas en las respuestas de los participantes. En primer lugar, los músicos afirmaron que confiaban mucho más en la memoria auditiva que en la cinestésica, aunque todos coincidieron en que es importante conocer los componentes cinestésicos de una pieza (por ejemplo, la digitación). En segundo lugar, todos los pianistas destacaron que tener una idea clara de la estructura musical y analizar la partitura eran los métodos de memorización más importantes y seguros. En tercer lugar, cinco de los siete pianistas comentaron que memorizaban de una manera «holística» dividiendo la pieza en secciones breves para obtener una «recuerdo más sólido de toda la obra» [16]. En cuarto lugar, la capacidad de improvisar fue considerada una ventaja importante para memorizar música; incluso un pianista que no sabía improvisar la consideró como una herramienta valiosa. Basándose en los resultados de su investigación, Aiello concluyó que, en general, los maestros deberían ayudar a sus alumnos a memorizar música, fomentando en ellos el conocimiento detallado de la estructura musical y ayudándoles a analizar la partitura.

Las investigaciones de Hallam y Aiello, junto con las obras de Hughes y Matthay, aportan un conocimiento profundo sobre cómo los músicos inter-

pretan de memoria. En suma, estas investigaciones revelan que los músicos tienden a utilizar una gran variedad de métodos para memorizar música. Desafortunadamente, no existe ninguna prueba de que la memoria auditiva, visual o cinestésica, o una combinación de ellas, sea la más eficiente o proporcione la mayor seguridad en las interpretaciones memorizadas. La diversidad de opiniones en torno a la importancia de estos métodos —incluso entre los solistas— es muy sorprendente. Algunos músicos afirman que su manera de memorizar música es superior a otras. Indudablemente, cada método propio será uno de los mejores para esa persona, sencillamente porque lo ha elegido y desarrollado durante años de práctica.

Sin embargo, las obras de Hallam y Aiello —apoyadas por las afirmaciones anecdóticas de Hughes y Matthay— sugieren que los músicos más experimentados de sus investigaciones confiaban ampliamente en las estrategias analíticas a la hora de memorizar. Los profesores y analistas musicales han destacado desde hace tiempo la importancia de estudiar la estructura de las composiciones musicales. Nicholas Cook, por ejemplo, ha sostenido que

> la capacidad de dejar a un lado los detalles y observar las relaciones a gran escala pertinentes al contexto musical particular, que es lo que fomenta el análisis, forma parte esencial de la manera en que el músico percibe el sonido musical. Para el intérprete, es obvio que el análisis desempeña un papel en la memorización de partituras extensas, y, hasta cierto punto, en el discernimiento de las relaciones dinámicas y rítmicas a gran escala (aunque algunas de las afirmaciones de que el análisis schenkeriano, en particular, es indispensable para intérpretes y directores han sido indudablemente exageradas) [17].

Sin embargo, es necesaria una investigación ulterior para determinar si el estudio de la estructura de la música proporciona el medio más seguro para la memorización y si ha de ser la base para aprender las obras auditiva, visual y cinestésicamente. Las investigaciones sobre las habilidades excepcionales de memorización en otros tipos de actividades pueden aportar conocimientos valiosos sobre este asunto.

### *Los principios de la memoria experta*

Los investigadores han estudiado la memoria excepcional en diversos tipos de actividades, incluyendo el ajedrez, la física, la medicina, la escritura y el atletismo. A partir de sus trabajos, se han propuesto varias teorías para explicar cómo las personas son capaces de desarrollar y mantener las extraordinarias habilidades que demuestran con frecuencia los profesionales eminentes y

expertos en esos campos [18]. Antes de aplicar a la música cualquier teoría sobre la memoria experta, debemos determinar si la memoria musical es comparable a otros tipos de memoria experta. Aunque la naturaleza de la interpretación musical difiere fundamentalmente de la de otras actividades como el ajedrez o el atletismo, los músicos expertos comparten algunas de las mismas exigencias de memoria con otros especialistas. Por ejemplo, la memoria desempeña un papel importante en deportes como la gimnasia y el patinaje artístico, ya que los deportistas han de memorizar las secuencias de los movimientos que constituyen una actuación determinada y ejecutar dichos movimientos con referencia a unos modelos técnicos definidos. Debido a que las habilidades técnicas son valoradas de acuerdo con unas formas ideales, los ejecutantes deben hacer comparaciones entre las técnicas «reales» e «ideales» durante la ejecución. Las bases para estas comparaciones se obtienen de la memoria a largo plazo, adquirida a lo largo de muchos años de entrenamiento [19]. Los solistas, al igual que los expertos en los deportes mencionados, practican minuciosamente para poder recordar las series de acciones que constituyen una interpretación; se refieren a los modelos técnicos definidos que esperan el público y el jurado; y supervisan su ejecución a medida que transcurre para evaluar su calidad.

Habiendo introducido el tema de las comparaciones, podemos comenzar a aplicar ciertos aspectos de las teorías existentes a la memoria musical. En general, las investigaciones han revelado que las habilidades de memoria sobresalientes están apoyadas en una amplia base de conocimientos específicos de la actividad. En esa base de conocimientos, se recopila información continuamente y se almacena en grupos significativos (o «bloques»), que se asocian frecuentemente con órdenes o actos físicos específicos [20]. Entre los ejemplos básicos de esos bloques en la música se incluyen las escalas y los arpegios. Los músicos practican escalas y arpegios durante horas para poder reconocerlos y ejecutarlos con facilidad cuando leen a primera vista o memorizan una pieza.

Una vez que ha adquirido una extensa base de datos significativos, el individuo puede utilizarla para crear una representación mental de cualquier información compleja que ha de ser memorizada [21]. En términos sencillos, una representación mental es una especie de mapa interno que se puede utilizar con confianza para recordar un material específico. Como en cualquier mapa, se deben identificar en la práctica los hitos (o señales) importantes dentro de la información y después usarlos para guiar la recuperación, de modo que la interpretación resultante pueda seguir su curso [22].

En términos musicales, una composición puede representarse en una jerarquía que abarca desde el nivel más alto de comprensión global (la pieza en su conjunto) hasta el nivel inferior (la pieza como notas separadas). Además,

la representación puede estar basada en la memoria auditiva, visual o cinestésica, o bien en diversas estrategias analíticas, dependiendo de las preferencias del intérprete y de la pieza en cuestión. Sin embargo, los límites humanos en cuanto al procesamiento de información y la atención hacen que sea poco probable que se active una representación completa durante la interpretación, incluso cuando se trata de una pieza de longitud moderada. En cambio, sólo una parte de la representación se activa en un momento dado, y esa región activa se desplaza a medida que avanza la interpretación de la música [23]. Por ejemplo, en medio de una frase musical profundamente grabada en la memoria, el músico puede estar fundamentalmente centrado en las relaciones que existen dentro de la frase misma. En ese caso, sólo estarán activas las señales auditivas, visuales y cinestésicas del nivel inferior. A la inversa, cuando se encuentra en los límites de una frase, es posible que el intérprete necesite saber cómo las frases previa y subsiguiente están relacionadas entre sí y con la estructura general de la pieza. En un momento así, se activarían las señales del nivel superior que especifican las relaciones a gran escala —aunque pueden ser necesarias ciertas señales del nivel inferior para ejecutar la sucesión inmediata de eventos [24].

Obviamente, existe una enorme posibilidad de que surjan diferencias individuales en el contenido específico de las representaciones mentales de los músicos. Tales diferencias se muestran claramente en las entrevistas de Hallam y Aiello. Sin embargo, las investigaciones han revelado que una representación puede ser potencialmente segura y perdurable siempre y cuando se haya organizado y ensayado minuciosamente [25]. Uno de los métodos más eficaces de organizar la música memorizada —tal y como indican los músicos profesionales en los estudios de Hallam y Aiello— es incorporar estrategias analíticas a la práctica. De esta manera, los intérpretes pueden establecer una base sólida sobre la cual se aprenden auditiva, visual y cinestésicamente las composiciones.

Roger Chaffin y Gabriela Imreh han explicado más detalladamente la importancia de las estrategias analíticas en la práctica organizativa y la recuperación [26]. Observaron sistemáticamente la práctica de una pianista concertista —Imreh misma— para determinar cómo formaba una representación mental del tercer movimiento del *Concierto italiano* de Bach. Descubrieron que la estructura formal de la composición sirvió de herramienta para organizar la representación. De hecho, la pianista integró su conocimiento de la estructura formal en las primeras fases de práctica y después confió ampliamente en dicho conocimiento para recordar la música durante la interpretación. Además, afirmó que sus señales eran de tres tipos —«básicas», «interpretativas» y «expresivas»— y que las había asociado con la música nota por nota mediante una combinación de las memorias auditiva, visual y cinestésica. Las señales

«básicas» incluían digitaciones, dificultades técnicas y grupos de notas que formaban unidades de información identificables. Las señales «interpretativas» incluían fraseo, dinámica, tempo y pedales. Las «expresivas» eran partes de la partitura en las que la pianista intentaba producir cambios en la expresión o el carácter de la música.

No todos los intérpretes estarán tan familiarizados con la estructura formal de una composición como Imreh; tampoco clasificarán los componentes más memorables e importantes de una composición mediante señales básicas, interpretativas y expresivas. Sin embargo, este estudio confirma la importancia de las estrategias analíticas a la hora de establecer una representación mental segura de una composición, ya que proporcionan señales de observación para ser usadas a lo largo del proceso de memorización. Otros músicos pueden emplear las estrategias analíticas particulares que conocen y prefieren, aprendiendo y almacenando correspondientemente señales importantes. Cualesquiera que sean esas preferencias, las investigaciones realizadas sugieren que, para que las señales sean seguras y estables durante la interpretación, deben estar organizadas de manera significativa y ser utilizadas para guiar la práctica [27].

## *Conclusiones*

Las investigaciones recientes han arrojado luz sobre las maneras en que los músicos memorizan música. Los músicos deberían ser conscientes de que una característica sobresaliente del proceso de memorización es la creación y empleo de representaciones internas para recuperar información específica, y que se ha demostrado claramente que los profesionales utilizan para ello estrategias analíticas. Aunque sería inverosímil recomendarles a todos los músicos que realicen un análisis formal minucioso de cada composición que ejecuten, se les debería animar a que desarrollen sus propias estrategias analíticas y las incorporen en las primeras fases del aprendizaje. Ello les permitirá reconocer señales importantes dentro de una pieza en las que podrán confiar con seguridad, a las que podrán referirse a medida que progresa la interpretación o que les ayudarán a reanudar una interpretación que se ha interrumpido inesperadamente.

Uno de los temas que deberían abordar las investigaciones venideras es la medida en que ciertas estrategias analíticas son las más efectivas y eficaces. Podemos sugerir naturalmente que a mayor diversidad de codificaciones de la música, mayor será la cantidad de asociaciones y conexiones que se formen con esa información y, por lo tanto, el ejecutante tendrá más posibilidades de recordarla. En este sentido, aquellos intérpretes que se esfuerzan por apren-

derse las composiciones auditiva, visual y cinestésicamente y que practican utilizando estrategias analíticas o de otro tipo (por ejemplo, improvisación y recursos mnemotécnicos) deberían tener la memoria más segura y eficaz para la interpretación. Sin embargo, tales excesos de codificación y práctica pueden no ser la manera más eficaz de abordar la memorización. Idealmente, los intérpretes deberán procurar un equilibrio entre la eficacia y eficiencia si quieren responder satisfactoriamente a las elevadas y variadas exigencias de una carrera musical. Es posible que las investigaciones futuras revelen dónde se encuentra ese equilibrio.

## *Notas*

1. Véase Roger Chaffin y Gabriela Imreh, «'Pulling teeth and torture': musical memory and problem solving», *Thinking and Reasoning* 3 (1997), pp. 315-316.
2. Seymour Bernstein, *With Your Own Two Hands: Self-Discovery through Music,* Nueva York, Schirmer, 1981, p. 220.
3. *Ibid.*, p. 220.
4. Edwin Hughes, «Musical memory in piano playing and piano study», *Musical Quarterly* 1 (1915), p. 595.
5. Aaron Williamon, «The value of performing from memory», *Psychology of Music* 27 (1999), pp. 84-95.
6. Respectivamente, Edwin Hughes, *op. cit.*, y Tobias Matthay, *On Memorizing and Playing from Memory and On the Laws of Practice Generally,* Londres, Oxford University Press, 1926.
7. Edwin Hughes, *op. cit.*, p. 598.
8. *Ibid.*, p. 599.
9. *Ibid.*, p. 599; véase Gabriela Imreh y Roger Chaffin, «Understanding and developing musical memory: the views of a concert pianist and a cognitive psychologist», *The American Music Teacher* (diciembre 1996/enero 1997), pp. 20-24.
10. Susan Hallam, «The development of memorisation strategies in musicians», discurso presentado en el Séptimo Congreso sobre Psicología del Desarrollo (Cracovia, Polonia, 1995).
11. *Ibid.*, p. 8.
12. *Ibid.*, p. 6.
13. *Ibid.*, p. 5.
14. *Ibid.*, p. 7.
15. Rita Aiello, «Strategies for memorizing piano music: pedagogical implications», ponencia cartel presentada en la Eastern Division of the Music Educators National Conference (Nueva York, NY, 1999).
16. *Ibid.*, p. 1.
17. Nicholas Cook, *A Guide to Musical Analysis,* Londres, Dent, 1989, p. 232; véase también el capítulo 3 de este libro.
18. Véanse William G. Chase y Herbert A. Simon, «Perception in chess», *Cognitive Psychology* 4 (1973), pp. 55-81; K. Anders Ericsson y Walter Kintsch, «Long-term working memory», *Psychological Review* 102 (1995), pp. 211-245; Eric F. Clarke, «Generative principles in music performance», en John A. Sloboda (ed.), *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition,* Oxford, Clarendon Press, 1988, pp. 1-26; y Andreas C. Lehmann y K. Anders Ericsson, «Expert pianists's mental representation of me-

morized music», ponencia cartel presentada en la 36º Reunión Anual de la Psychonomic Society (Los Ángeles, 1995). Los primeros dos estudios utilizan el término «estructura de recuperación» para referirse a las representaciones mediante las cuales los expertos en diversos campos recuperan la información almacenada para la ejecución. Para los objetivos de este capítulo se ha utilizado el término más general de «representación mental» en concordancia con los últimos dos estudios, dirigidos a las representaciones que emplean los músicos en particular.

19. Véase Fran Allard y Janet L. Starkes, «Motor-skill experts in sports and other domains», en K. Anders Ericsson y Jacqui Smith (eds.), *Toward a General Theory of Expertise: Prospects and Limits,* Cambridge, Cambridge University Press, 1991, pp. 126-152.
20. Véanse William G. Chase y Herbert A. Simon, *op. cit.*; K. Anders Ericsson y Walter Kintsch, *op. cit.*
21. Véase K. Anders Ericsson y Walter Kintsch, *op. cit.;* véase también el capítulo 7 de este libro.
22. Véanse Endel Tulving y Z. Pearlstone, «Availability versus accessibility of information in memory for words», *Journal of Verbal Learning and Verbal Behavior* 5 (1966), pp. 381-391; Alan D. Baddeley, *Human Memory: Theory and Practice,* Boston, Allyn & Bacon, 1990; K. Anders Ericsson y Walter Kintsch, *op. cit.*
23. Véase Eric F. Clarke, *op. cit.*
24. Véase *ibid.*, p. 4.
25. Véase William G. Chase y K. Anders Ericsson, «Skilled memory», en John R. Anderson (ed.), *Cognitive Skills and Their Acquisition,* Hillsdale, NJ, Erlbaum, 1981, pp. 141-189.
26. Véanse Roger Chaffin y Gabriela Imreh, «'Pulling teeth and torture': musical memory and problem solving», cit.; y Roger Chaffin y Gabriela Imreh, «A comparison of practice and self-report as sources of information about the goals of expert practice», *Psychology of Music* 29 (2001), pp. 39-69; véase también Roger Chaffin, Gabriela Imreh y Mary Crawford, *Practicing Perfection: Memory and Piano Performance,* Mahwah, NJ, Erlbaum, 2002.
27. Véanse Aaron Williamon y Elizabeth Valentine, «The role of retrieval structures in memorizing music», *Cognitive Psychology* 44 (2002), pp. 1-32; y Aaron Williamon, Elizabeth Valentine y John Valentine, «Shifting the focus of attention between levels of musical structure», *European Journal of Cognitive Psychology* (de próxima aparición).

## *Lecturas complementarias*

Bernstein, Seymour: *With Your Own Two Hands: Self-Discovery through Music,* Nueva York, Schirmer, 1981.

Chaffin, Roger, Gabriela Imreh y Mary Crawford: *Practicing Perfection: Memory and Piano Performance,* Mahwah, NJ, Erlbaum, 2002.

Matthay, Tobias: *On Memorizing and Playing from Memory and On the Laws of Practice Generally,* Londres, Oxford University Press, 1926.

Williamon, Aaron: «The value of performing from memory», *Psychology of Music* 27 (1999), pp. 84-95.

# TERCERA PARTE
# *Hacer música*

# 9 *De la partitura al sonido*

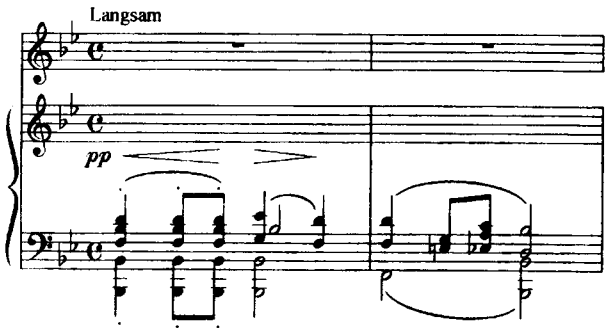
PETER HILL

Muchos intérpretes se refieren a las partituras con el término «la música». Eso no es correcto, por supuesto. Las partituras contienen información musical, en parte exacta y en parte aproximada, junto con indicaciones sobre cómo puede interpretarse esa información. Pero la *música* misma es algo imaginado, en primer lugar por el compositor, después en asociación con el intérprete, y finalmente transmitido en sonido.

El momento en que la música nace como sonido, junto con los correspondientes placeres y frustraciones, ha sido hermosamente descrito por el gran acompañante Gerald Moore. En el siguiente fragmento, Moore analiza microscópicamente los compases iniciales, o *Vorspiel,* del «Wandrers Nachtlied» de Schubert (véase el Ejemplo 9.1). La elección de una canción tan «fácil» es reveladora, porque destaca que la técnica no es simplemente una cuestión de reproducir la partitura con exactitud —en este caso las notas son suficientemente sencillas para ser leídas a primera vista— sino de dar vida a la partitura en sonido. Moore se deleita demostrando lo escurridizo que eso puede ser:

> Dinámicamente, este pequeño *Vorspiel* es todo *pianissimo,* pero dentro de los límites de ese *pianissimo* debe haber un ligero aumento y una subsiguiente reducción de intensidad. Es una curva que asciende y luego desciende; la más suave de las curvas, en la que cada acorde se enlaza con el siguiente. Tan restringida es su gama, tan estrecho el margen entre tu acorde más suave y tu acorde menos suave, que si sobrepasas en una fracción el límite superior de la curva, todo se vendrá abajo. Cada acorde, aunque esté relacionado y enlazado con su vecino, tiene un peso diferente, y la diferencia no es mayor que el peso de una pluma. Escuchas con autocrítica

EJEMPLO 9.1 *Schubert, «Wandrers Nachtlied» D. 768, compases 1-2.*




> mientras la practicas. Experimentas. Tocas cada acorde con una presión uniforme y ligera de modo que no haya aumento ni disminución de intensidad —todo *pianissimo.* Después intentas darle ese *crescendo* y *diminuendo* infinitesimales que se desean realmente para darle forma y sentido a la frase: pero está desproporcionado —te has sobrepasado—, así que empiezas de nuevo. Ahora encuentras que tus acordes suenan turbios, tu pedal es incorrecto, un acorde invade el terreno de otro en lugar de fundirse suavemente con él sin empañarlo todo. Intentas solucionarlo. Pero, a pesar del *pianissimo* que estás logrando, comienzas a darte cuenta de que tus acordes no tienen carácter —son pesados, y a toda la frase le falta vida. Entonces, muy delicadamente, pruebas dándole una fracción más de peso al dedo superior de tu mano derecha [...] Todo ello es la búsqueda más fascinante que se pueda imaginar. En tu búsqueda de luz y sombra te sientes tan feliz y abstraído como un pintor mezclando los colores en su paleta, y la satisfacción del intérprete cuando ha logrado producir esa ondulación perfecta, esa línea clara, cuando siente que por fin todo el diseño es proporcionado y hermoso, es inmensa. Pero esa satisfacción o autosatisfacción se experimenta muy pocas veces. Con la misma seguridad con la que te das cuenta de que *lo has logrado,* tu aguzada sensibilidad te recuerda lo escurridiza que es la recompensa y cuántas veces te has quedado sin ella [1].

Esta clase magistral de Moore es sumamente instructiva. Ya hemos visto cómo invierte con elegancia la visión normal de lo que hace que la música sea difícil. Moore argumenta que, precisamente porque hay tan pocas notas, cada una requiere un arte y un control muy refinados. Una nota inexacta en el finale de una sonata de Beethoven afectará poco o nada al resultado musical; pero en el «Wandrers Nachtlied» hasta el menor error de cálculo significa un desastre. Después expone la secuencia de procedimientos: la concepción —la imagen inicial de cómo debe sonar la frase— conduce a la experimentación en el piano, después la reflexión y la autocrítica, seguidas de más experi-

mentos. Un punto clave es que el trabajo progresa en un estado de serenidad: la búsqueda del ideal es su propia recompensa, por más escasas que sean las ocasiones en las que realmente se logra ese ideal. El equilibrio entre el extremado cuidado y la paciencia, por un lado, y el profundo deseo de hacer música es uno de los equilibrios más difíciles que ha de conseguir el intérprete. En un sentido debemos ser impacientes; pero si la impaciencia se derrama sobre la obra misma, las consecuencias son negativas, ya que tomaremos atajos y bajaremos el listón. Además es probable que incorporemos dudas y tensiones que reaparecerán perjudicialmente en la interpretación y que a largo plazo pueden ser difíciles o imposibles de erradicar.

Moore finaliza su descripción citando la desventaja de su método: «Se puede argumentar que trabajando y experimentando de manera repetitiva, como yo recomiendo, se puede perder el espíritu, la inspiración y la frescura de la música. En efecto, ése es el mayor peligro» [2]. Este problema fundamental, que Moore deja sin resolver, ocurre porque la perfección que él busca se construye mediante capas y capas de hábitos arraigados. Surge entonces la pregunta de cómo podemos aglutinar esos hábitos sin aglutinar al mismo tiempo nuestra concepción musical. Ejercitar repetidamente la interpretación, como reconoce Moore, embota la creatividad del intérprete. Además, debilita o elimina nuestra capacidad de adaptarnos tanto a las vicisitudes que forman parte esencial de una interpretación en vivo como a las ideas diferentes de los demás. Moore debe de haber sufrido tanto como cualquiera al trabajar con músicos encerrados en su propia interpretación, indiferentes o incapaces de responder al punto de vista del pianista.

Éste es, pues, el dilema central del músico ejecutante. ¿Cómo podemos practicar y practicar sin perder el entusiasmo o la receptividad que nos permite desarrollar nuestras ideas? ¿Cómo, de hecho, podemos *interpretar,* en el sentido de crear un acontecimiento único, con la agudeza que surge de la inspiración del momento? Evidentemente, si la repetición es el núcleo del problema, debemos encontrar una manera de practicar que reduzca su papel al mínimo, o al menos incorporar a nuestro método una manera de mitigar sus desventajas.

El primer paso es analizar la manera en que comenzamos el aprendizaje de una obra. La primera fase es la más importante: no sólo establecemos las pautas para el trabajo subsiguiente, sino que nuestras decisiones —tomadas rápida y quizá irreflexivamente— pueden llegar a dominar la manera en que tocamos la pieza durante el resto de nuestra vida. ¿Cuáles son las suposiciones que hacemos sobre el estilo de la pieza? ¿Cuál es el significado de la notación del compositor? ¿Cuál es nuestro papel como intérpretes en relación al compositor? Esta última pregunta es tan fundamental que requiere una digresión. En el ejemplo que hemos comentado —Moore y Schubert— la acti-

tud es típica del intérprete de mediados del siglo XX: respeta el texto y las intenciones del compositor, y ve la interpretación como el arte de leer entre líneas. Sin embargo, si nos trasladamos a la música escrita un siglo después de Schubert, nos encontramos con que el espacio «entre las líneas», por lo visto, ya está lleno de indicaciones, como en las obras de la Segunda Escuela Vienesa, que parecen especificar todos los matices. El aumento general de la información especificada por los compositores durante los últimos dos siglos sugiere que no confían en la capacidad de comprensión del intérprete o tienen un afán de controlar todo, posturas de las que Stravinsky fue portavoz en sus diatribas contra la «interpretación». En estas circunstancias, no es de extrañar que los intérpretes se hayan inclinado por la interpretación textual. Esto es especialmente cierto en cuanto a la música más reciente. En los ensayos, incluso de grupos de especialistas selectos, uno tiene la impresión de que la partitura *es* la música: se le da prioridad absoluta a ejecutar correctamente todas las indicaciones de la partitura, mientras que lo que la partitura pueda *expresar* se da automáticamente.

El problema de la interpretación literal es que puede ser completamente errónea. Un ejemplo clásico y bien documentado es el caso de las Variaciones para piano op. 27 de Webern. El estudio de la historia de la interpretación de esta obra revela que durante años los pianistas tomaron el aspecto austero de la partitura al pie de la letra: las interpretaciones eran rigurosamente impersonales, sin añadir nada a las escasas indicaciones expresivas. Con el tiempo, y gracias a los esfuerzos de los alumnos de Webern [3], se supo que aquello era lo opuesto a sus intenciones, y que lo que realmente tenía en mente era una constante interacción de *rubato* y luces y sombras. ¿Por qué, entonces, Webern no anotó esos matices? Quizá eran demasiado sutiles para poder especificarlos, o quizá dependían demasiado del instinto del pianista (incluyendo al mismo Webern) como para imponerlos. Cualquiera que fuese la razón, el error que cometieron los pianistas fue pasar por alto la diferencia del enfoque de la notación entre las obras tempranas y tardías de Webern.

Otro ejemplo es uno que proviene de mi experiencia personal. Cuando comencé a estudiar la música de Messiaen, las diferencias frecuentemente mínimas de sus ritmos hicieron que me esmerara por lograr una precisión extrema; pero cuando finalmente tuve la oportunidad de estudiar con él, ese aspecto de mi ejecución le pareció mecánico. Messiaen opinaba que el ritmo y el fraseo deben ser flexibles, y por más compleja que sea la notación, la música nunca ha de sonar «como un estudio». Para Messiaen, la labor del intérprete es inferir el significado y el «carácter» de lo que está escrito en la partitura —como en el «Nachtlied» de Schubert [4].

Suponiendo que se han considerado los asuntos más generales —el estilo, la notación y el papel del intérprete—, ¿cómo comenzamos con la partitura

en sí? Cuando les planteo esta pregunta a mis alumnos, recibo una variedad de respuestas. Una de las sugerencias es comenzar organizando la técnica: digitación, articulación, etc. El problema de ese enfoque es que trata a la partitura «como un estudio» y no como una pieza musical. Como la digitación y la articulación pueden cambiar, incluso radicalmente, la manera en que suena la música, no podemos establecerlas con antelación, sin conocer el objetivo al que habrán de servir. Otra escuela de pensamiento prefiere establecer un primer contacto con la partitura leyéndola a primera vista de principio a fin. Esto también tiene sus desventajas. Las correcciones, modificaciones y mejoras se realizarán por fragmentos, y, si no se tiene una visión fundamental del funcionamiento de la obra, el resultado será arbitrario. Los intérpretes adquieren hábitos con una facilidad asombrosa. Y por lo general esa facilidad se emplea mal, en lugar de aprovecharla. Si no tenemos una visión de cómo debe sonar la música, seremos propensos a acostumbrarnos inconscientemente a los defectos de nuestra ejecución —defectos que son inevitables en esta primera fase— y a formar malos hábitos que, como todos sabemos, son fáciles de adquirir pero muy difíciles de corregir.

Otra objeción que puede hacerse a ambos enfoques es que son retornos a la infancia: el alumno obediente que primero se aprende las notas y al que después, como recompensa, se le permite «ponerles la música». Por importantes que sean, si «las notas» van primero, el peligro es que «la música» seguirá a nuestra técnica, en lugar de guiarla, condicionada por nuestro nivel de habilidad y también, hay que decirlo, por nuestros defectos.

En respuesta a estas objeciones, mis alumnos frecuentemente sugieren escuchar una grabación, como un modelo de interpretación. Sin embargo, aunque es un paso en la dirección correcta —ya que da prioridad a la música—, es un atajo de perezosos, el equivalente musical a buscar la respuesta al final del libro. Si nos limitamos a copiar la manera en que otros tocan, estamos admitiendo que lo que nosotros podemos ofrecer es de segunda categoría. De nuevo, es un mal empleo de la primera fase: si no tenemos nuestras propias ideas, no podemos evitar que otros nos influyan excesivamente[5].

Entonces, ¿qué podemos hacer? Descartando las posibilidades mencionadas, la conclusión inevitable es que el primer paso es desarrollar nuestra propia visión de la obra: el ideal es aprovechar el precioso momento en el que la pieza nos es desconocida, cuando tenemos la mayor capacidad de aplicar una percepción fresca. Sin embargo, debemos hacerlo sin que nos influyan inconscientemente las inevitables limitaciones de nuestras habilidades. La solución es trabajar intensamente con la partitura, pero escuchándola en nuestra mente, *alejados del instrumento.* El beneficio musical es que podemos desarrollar nuestra comprensión de la obra sin estar limitados por las consideraciones técnicas, centrándonos completamente en los aspectos musicales que

se deben decidir antes de practicar en el instrumento. Aunque pueda ser obvio, debemos saber, al menos en el aspecto general, lo que estamos intentando hacer antes de buscar la manera de lograrlo.

A continuación hay dos ejemplos de cómo uno podría comenzar a pensar sobre una partitura. La canción «L'ombre des arbres» de Debussy pone música al poema de Verlaine en dos estrofas de cuatro versos. La música evoca el río brumoso, cuya superficie actúa como un espejo que separa la realidad de la ilusión, los objetos reales de sus reflejos:

> La sombra de los árboles en el río brumoso
> se muere como el humo,
> mientras que en el aire, entre las ramas reales
> las tórtolas se quejan.
>
> ¡Cuánto, oh viajero, este paisaje lívido
> reflejó tu propia palidez
> y qué tristes lloraban en los altos ramajes
> tus esperanzas ahogadas!

El motivo principal está formado por las dos armonías iniciales, que se mecen entre acordes de séptima separados por un tritono, sobre Do♯ y Sol♮ (véase el Ejemplo 9.2). Es quizá un equivalente musical de la imagen de Verlaine, en el cual el Mi♯ (común a ambos acordes) actúa como una especie de «horizonte», análogo a la superficie del agua, que más adelante se recalca (como Fa♮) en el compás 5 cuando la armonía cambia a una novena menor sobre Mi, dándole un giro amargo a «Meurt comme de la fumée». Hasta aquí la relación entre el poema y la música es clara. Pero en la segunda mitad, misteriosamente, cada cual toma su camino. Verlaine utiliza su imagen de la naturaleza como metáfora de la condición humana, en el «viajante» cuya palidez se asemeja a la del paisaje y cuyas mayores aspiraciones —como las ramas más altas— se pierden en las profundidades del agua.

Esta resignación no es apoyada por la parte de piano, que avanza vertiginosamente con elocuentes secuencias cromáticas y, mientras la voz se apaga («Tes espérances noyées»), prorrumpe en una versión enormemente magnificada de la armonía de $Sol^7$. Finalmente, y más misteriosamente aún, este atolladero armónico se resuelve tranquilamente (en una armonía de Do♯ mayor), mientras el «horizonte» (Mi♯) late con calma en la voz superior. De este modo, la desesperanza del poema es aparentemente contradicha por la serenidad y la aceptación del postludio.

Naturalmente, lo anterior es tan sólo un esbozo de cómo podemos comenzar a desarrollar nuestros pensamientos cuando se trata de comprender el poema, la música y la relación que existe entre ambos. La cuestión es que evidentemente la cantante y el pianista tienen mucho material que analizar —especialmente la paradoja de la segunda estrofa— antes de poder tomar las decisiones musicales específicas.

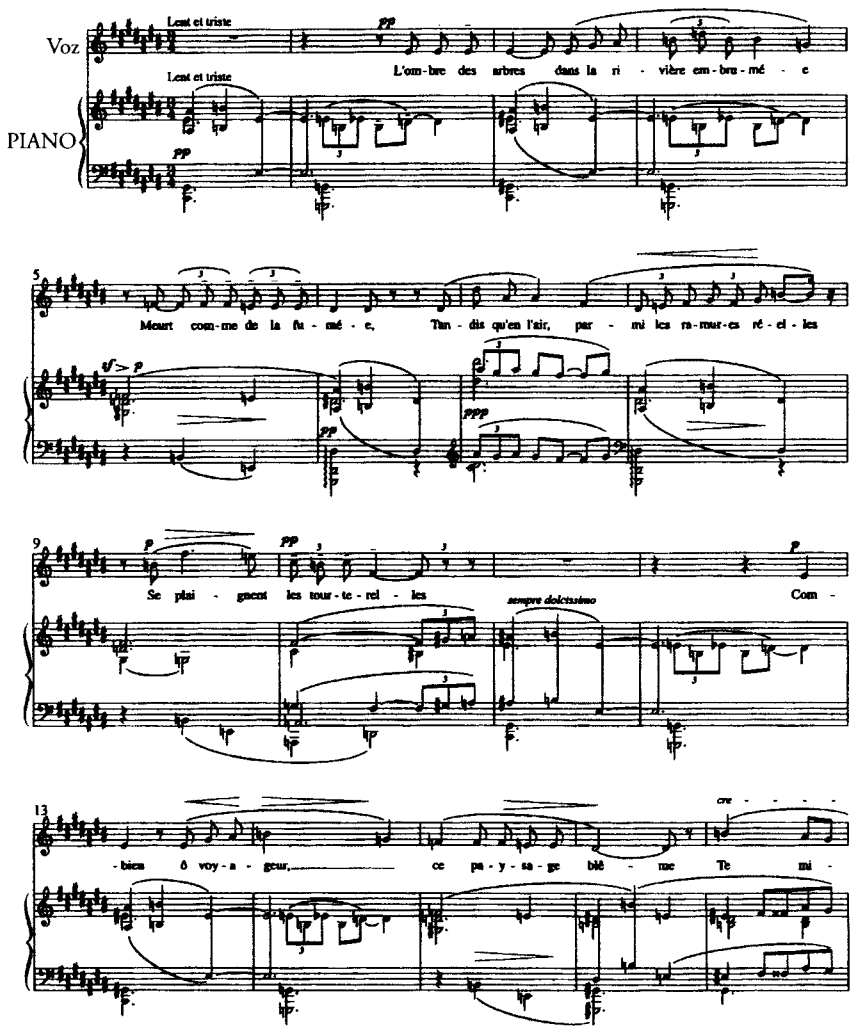
Pasar de la destilada concentración de Debussy y Verlaine a un gran movimiento de una de las últimas sonatas de Beethoven —el primer movimiento de la Sonata en Lab mayor op. 110— es adentrarse en otro mundo. Sin embargo, el proceso de pensar sobre la música comienza, de manera similar, considerando una curiosidad.

La sección de desarrollo (compases 40-55; Ejemplo 9.3) es asombrosamente diferente del carácter expansivo del resto del movimiento: es breve, curiosamente formal, y se limita a pasar el tema por una serie de frases secuenciales. Esta observación sugiere a primera vista un estilo interpretativo modesto y «objetivo», confirmado inmediatamente por la manera en que Beethoven diseña la recapitulación. La línea melódica prorrumpe en unos trinos (compás 55) —una variante apresurada del trino cadencial de la frase que abre el movimiento (compás 4; Ejemplo 9.4)— que introducen el regreso del tema principal en el compás 56, esta vez en una versión apasionada, impulsada por las fusas de la mano izquierda. El flujo uniforme del desarrollo se puede considerar, retrospectivamente, como un preludio a la transformación decisiva del movimiento. La magnitud de esa transformación puede verse volviendo a la primera página de la partitura (Ejemplo 9.4). Entonces podemos entender que el comienzo de la recapitulación es una compresión de lo que al inicio era una serie de gradaciones sutiles, desde la exposición del tema en el estilo de un himno, a través del acompañamiento en semicorcheas (desde el compás 5), hasta las fusas que comienzan en el compás 12. Dado que el ritmo armónico permanece constante, el efecto de esas gradaciones es el de una serie de variaciones, realzado por el hecho de que en cada caso el primer movimiento armónico es el mismo. Esta visión de la música sugiere que el pianista debe realizar los cambios de velocidad con suavidad y exactitud, sin la repentina aceleración de tempo que frecuentemente se le da al comienzo de las fusas. Esta línea de investigación se puede extender, teniendo en cuenta las diferencias entre la recapitulación y la exposición. Por ejemplo, hay una modulación a la tonalidad lejana de Mi mayor para la repetición de los arpegios de fusas (desde el compás 70; Ejemplo 9.5) y una reestabilización (en Lab mayor) en la extensa coda (compases 105 y ss.). Más significativo aún es el último acorde de la exposición, con su discanto en espirales ascendentes de semicorcheas (desde el compás 36; Ejemplo 9.6), que en la reexposición florecen en una nueva variante —suspendidas en los agudos del

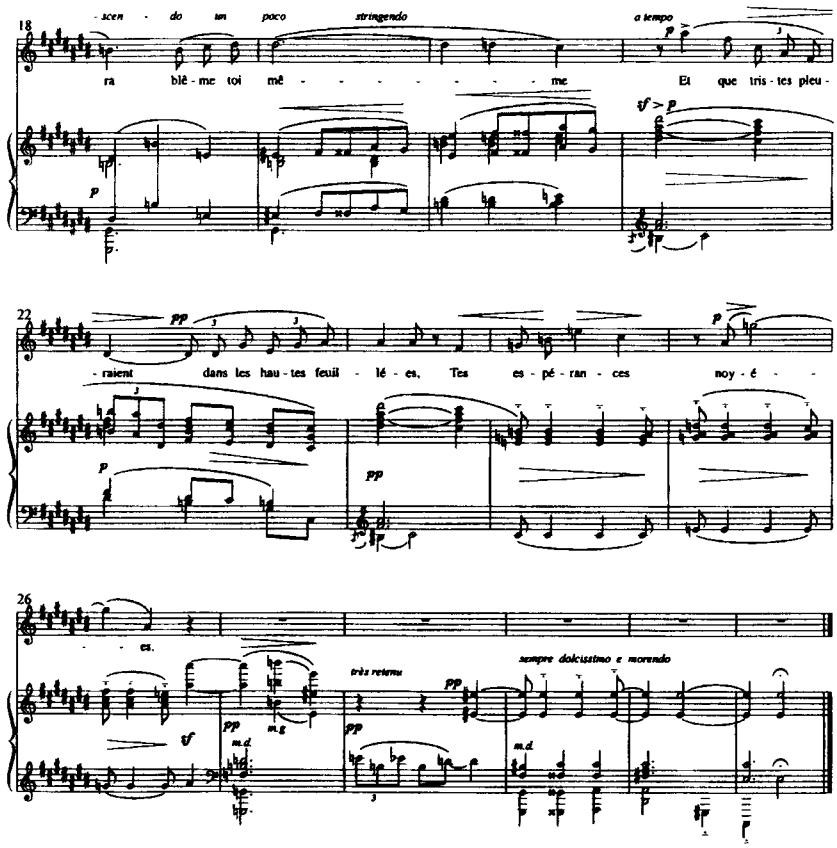
teclado— en un pasaje que parece ser el núcleo emocional del movimiento (compases 97-104; Ejemplo 9.7).

De esta manera, uno descubre gradualmente no sólo lo que la música está expresando, sino cómo cada frase o sección forma parte del conjunto, un

EJEMPLO 9.2 *Debussy, «L'ombre des arbres»,* Ariettes oubliées.



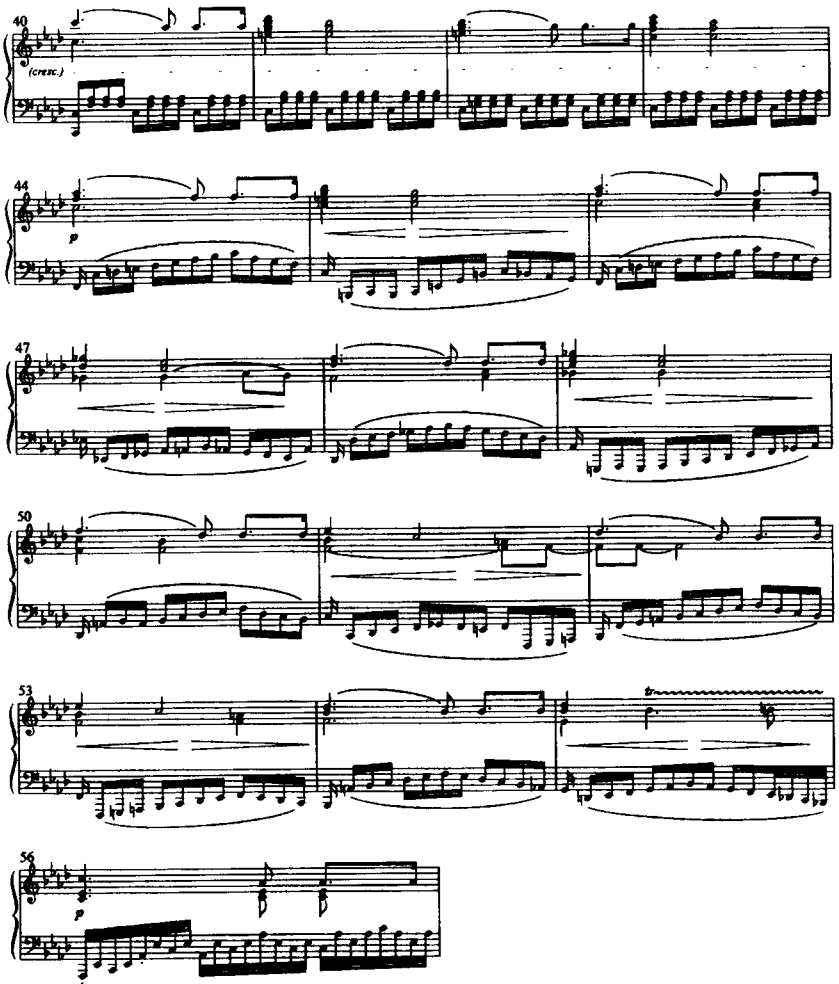
EJEMPLO 9.2 *(Continuación)*




sentido de la lógica dramática de Beethoven a medida que las piezas del puzle encajan en su sitio. La gran ventaja de trabajar alejados del instrumento se hace evidente, ya que uno puede experimentar y reflexionar a voluntad, desembarazado de los detalles técnicos que en esta fase representarían una distracción inoportuna.

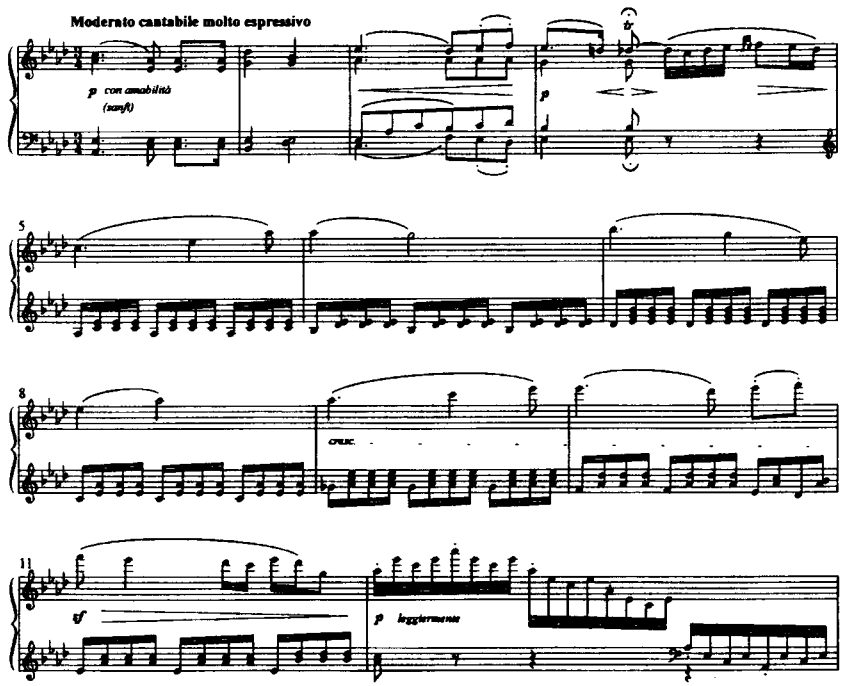
Todo este trabajo mental es similar a la preparación que realiza un director antes del primer ensayo con orquesta. Como el director, ahora podemos abordar el primer «ensayo» —la primera sesión de prácticas— con los fundamentos esenciales en su sitio: un esbozo de una concepción; un conocimiento del contexto, de qué es lo que ocurre en la música y por qué; y, lo más importante, la habilidad de escuchar la música en todas sus dimensiones. Incluso los problemas técnicos

EJEMPLO 9.3 *Beethoven, Sonata en La♭ mayor op. 110, primer movimiento, compases 40-56.*




pormenorizados se habrán anticipado y resuelto al menos parcialmente. Nótese también que habremos memorizado la pieza en gran parte, o quizá del todo: se trata de una memoria «activa», que surge del conocimiento de la lógica de la música en lugar de poceder del aprendizaje repetitivo de las notas. De esta manera uno espera abordar el dilema principal esbozado anteriormente, para poder *conocer* la música completamente sin quedar atrapados en una serie de hábitos.

EJEMPLO 9.4 *Beethoven, Sonata en Lab mayor op. 110, primer movimiento, compases 1-12.*




EJEMPLO 9.5 *Beethoven, Sonata en Lab mayor op. 110, primer movimiento, compases 70-72.*

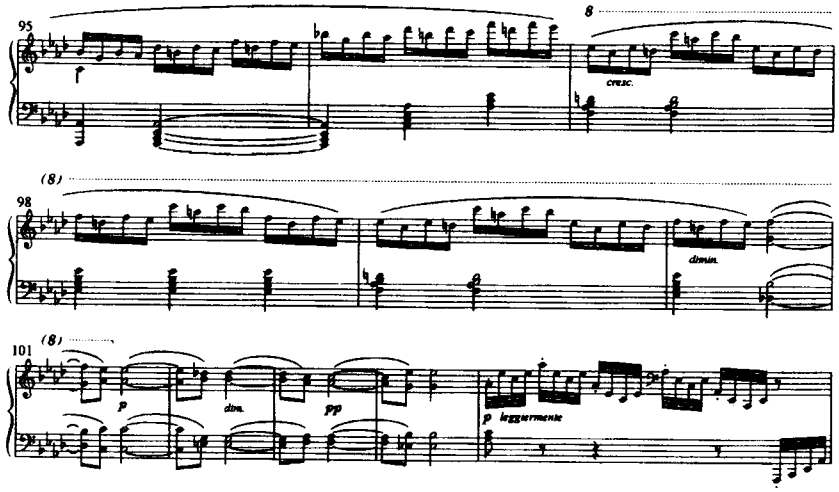



Sólo ahora estamos en condiciones de comenzar un experimento minucioso en el instrumento como el que describió Moore. Escuchando muy atentamente el sonido que producimos y comparándolo con nuestra concepción interna, podemos lograr que ambos coincidan. Sin embargo, siempre hay una delgada línea entre intentar que una idea funcione y estar dispuestos a revisar nuestro método.

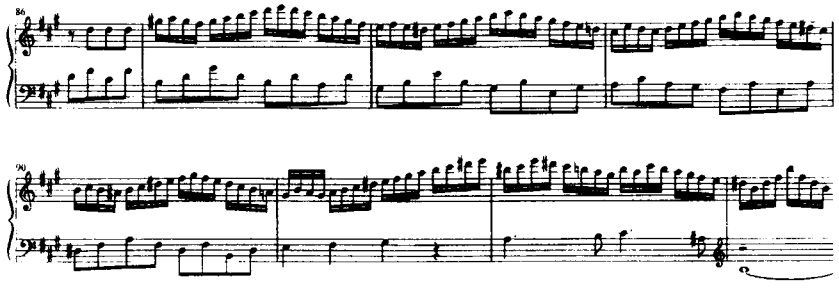
EJEMPLO 9.6 *Beethoven, Sonata en La♭ op. 110, primer movimiento, compases 36-37.*




EJEMPLO 9.7 *Beethoven, Sonata en La♭ op. 110, primer movimiento, compases 95-105.*




EJEMPLO 9.8 *Mozart, Concierto para piano en La mayor K. 488, primer movimiento, compases 86-93.*



En este punto llegamos a la segunda fase de resolución de nuestro problema. Con demasiada frecuencia, los intérpretes, y me incluyo, atacan los problemas técnicos con un enfoque provisional. Los errores deben corregirse, claro está, pero si nos limitamos a eso estaremos tratando el síntoma sin tratar la causa. Un ejemplo sencillo podría ser la afinación: antes de ajustar el dedo sobre la cuerda, ¿estamos seguros de poder *escuchar* (o cantar) el pasaje correctamente? De no ser así, «echarle la culpa» a los dedos es inútil.

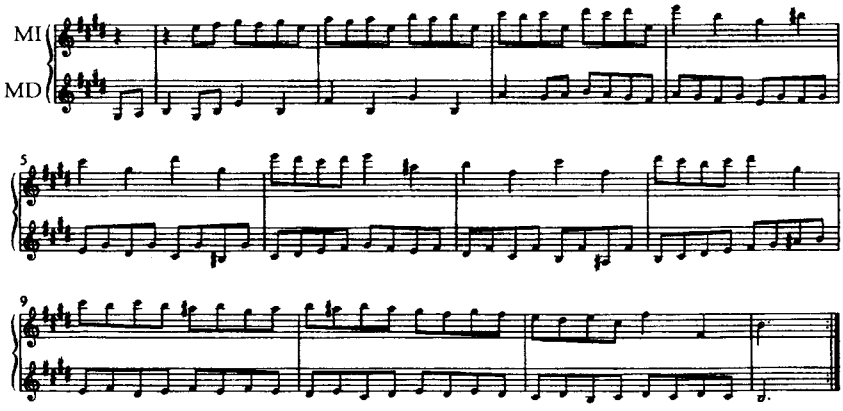
Esta verdad me fue inculcada, siendo un novato mayormente autodidacta de diecinueve años, en mi primera clase con el pianista Cyril Smith. Estaba tocando el primer movimiento del Concierto para piano en La mayor K. 488 de Mozart, y mis dedos se confundieron momentáneamente en la transición anterior al segundo tema (véase el Ejemplo 9.8). Inmediatamente, comencé a repetir el pasaje y corregir el desliz; sin embargo, Smith me detuvo y me dijo que practicara el pasaje tocando las semicorcheas de la mano derecha lentamente (sin la partitura) pero con un solo dedo. Esa extraña instrucción resultó inmediatamente reveladora. Al eliminar la memoria *física* (de mi digitación) me vi forzado a reconstruir el pasaje de oído, cosa que (vergonzosamente) no fui capaz de hacer. Smith me dio tiempo para que me aprendiera el pasaje auditivamente; cuando regresó, le toqué toda la sección como una «interpretación de concierto» —desinhibido y en el tempo justo— y milagrosamente claro y seguro. Por haber descubierto dónde se encontraba el error, en este caso en mi oído y no en mis dedos, pude corregirlo —permanentemente—, y destacaría que lo hice sin realizar ninguna práctica repetitiva convencional.

Se pueden abordar problemas más persistentes extendiendo este principio, de nuevo con el objetivo de atacar sus raíces. La clave para ello es la práctica «indirecta», un concepto conocido por los pianistas en particular gracias a las famosas ediciones de Alfred Cortot. Por ejemplo, Cortot añade a cada Estudio de Chopin unos ejercicios preliminares, de acuerdo con su precepto de que uno no debería practicar «el pasaje difícil, sino la dificultad que contiene, comenzando dicha práctica desde su misma raíz» [6]. Smith, que era un maestro interpretando los conciertos de Rachmaninoff y otros repertorios virtuosos, extendió aún más esa norma. Recomendaba un método feroz con el que, además de abordar la dificultad por etapas (como hacía Cortot), se exageraba el problema, extendiendo el nivel de habilidad más allá de lo requerido. Por poner un ejemplo sencillo: si el problema era de «geografía» del teclado —encontrar con seguridad las notas correctas—, la solución era aprender a tocar el pasaje con la misma facilidad, pero a ciegas, sin utilizar la vista para dirigir las manos. Para lograr uniformidad y brillantez en los pasajes virtuosos, proponía que se estudiasen las figuraciones en diferentes tonalidades, empleando la misma digitación para hacerlo aún más difícil —lo que de paso tiene la ventaja de evaluar nuestro conocimiento auditivo y armóni-

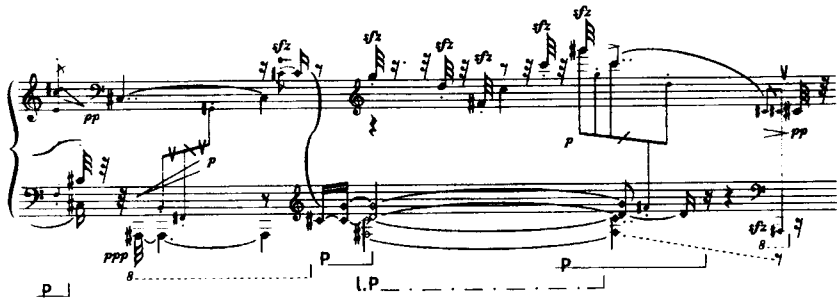
co de la obra. En estos días, he estado practicando en otra forma «indirecta» la Bourrée de la Suite Francesa núm. 6 en Mi mayor de Bach. Para mejorar el difícil equilibrio y coordinación de las dos voces, he estudiado la pieza con las manos cruzadas, transportando la parte de la mano izquierda dos octavas más arriba y la de la mano derecha una octava más abajo (véase el Ejemplo 9.9). Esto ayuda a contrarrestar la tendencia «diestra» que tenemos los pianistas, por la que escuchamos la voz superior a expensas de la mano izquierda. También resulta útil, especialmente en la música contrapuntística de este tipo, aprender a cantar la voz inferior tocando la superior. Además de aportar un beneficio inmediato, estas técnicas tienen otras ventajas. Una de ellas es que el acto de reaprender la música de diferentes maneras expone cualquier incertidumbre persistente en nuestro conocimiento básico —auditivo, intelectual, musical— de la música; por otra parte, saber que uno es capaz de afrontar un reto más allá de lo estrictamente necesario refuerza mucho nuestra seguridad. Por último, nótese que todas estas técnicas ahorran mucha práctica repetitiva de la música misma.

La experiencia de aprender música contemporánea me confirmó la importancia de estudiar mentalmente antes de comenzar a practicar. En las partituras que hemos analizado hasta ahora el intérprete puede aplicar mucha habilidad o «técnica» a la música: en efecto, como señala Moore, el hecho de que uno pueda leer a primera vista el «Wandrers Nachtlied» encubre lo

EJEMPLO 9.9 *Bach, Suite Francesa núm. 6 en Mi mayor BWV 817, Bourrée, compases 1-12 (mano izquierda transportada dos octavas más arriba; mano derecha transportada una octava más abajo).*



EJEMPLO 9.10 *Stockhausen,* Klavierstück VII, *p. 1, sistema 2;* © *1965 Universal Edition (Londres) Ltd., Londres.*



difícil que es realmente la pieza. Sin embargo, en el Ejemplo 9.10, del *Klavierstück VII* de Stockhausen, las técnicas de ejecución son tan inusuales que el intento de leerlo a primera vista es en gran parte inútil. Como muestra el fragmento, *Klavierstück VII* está basado en aislar y combinar las resonancias del piano. Estos sonidos «reflejados» se producen de dos maneras. El pianista puede pulsar una o varias teclas sin que lleguen a sonar (indicadas en la partitura mediante notas con forma de rombo) levantando así sus apagadores y hacer sonar los armónicos de las cuerdas correspondientes pulsando otras teclas con un agudo *staccato.* También puede prolongar los ataques como ecos, volviendo a liberar la cuerda, que aún vibra, mediante el pedal o una nueva pulsación insonora. Otra dificultad para el pianista es la eliminación gradual de las resonancias, soltando las teclas una por una.

Estudiar este tipo de música «experimental», en la que las técnicas interpretativas le deben poco o nada a las técnicas del repertorio tradicional, nos obliga a reconsiderar nuestro enfoque del aprendizaje y la práctica. Antes de intentar tocar *Klavierstück VII,* dediqué varias horas a hacer anotaciones en la partitura, asimilando y «probando» los detalles mentalmente, como una manera de comprender sus efectos musicales. La recompensa de una preparación esmerada es que, una vez al piano, podemos concentrarnos en el problema esencial, el equilibrio entre la intensidad del ataque y las resonancias. (Por cierto, ello varía notablemente según el piano y la acústica, y por lo tanto se debe planificar de nuevo para cada interpretación.) El proceso de recrear la partitura en sonido es fascinante, ya que uno explora diferencias de sonoridad que a menudo son extremadamente sutiles. Esta nueva conciencia del potencial del instrumento influye indudablemente en nuestra actitud hacia el sonido en la interpretación clásica. Igualmente, uno comienza a ver la «técnica» desde una perspectiva más creativa: tocar música contemporánea nos ani-

ma a buscar las soluciones que funcionen, aunque sean poco ortodoxas, en lugar de aplicar fórmulas habituales.

Sería un error deducir que el aprendizaje de una pieza es una sencilla progresión del trabajo mental a la práctica en el instrumento, del silencio al sonido. Las prácticas deben alternarse con períodos de descanso, durante los cuales la pieza se «olvida». También es posible que necesitemos períodos de reflexión: la práctica puede haber revelado deficiencias en nuestro conocimiento de la partitura. Esto ocurre especialmente cuando abordamos repertorios extremadamente difíciles. En las obras que explotan las posibilidades pianísticas hasta el límite, o incluso lo sobrepasan —como por ejemplo *Klavierstück X* de Stockhausen, *Evryali* de Xenakis y *La Fauvette des jardins* o el sexto movimiento fugado de *Vingt regards* de Messiaen—, se debe examinar muy cuidadosamente cada dificultad. La claridad mental absoluta es un prerrequisito, claro está. Asimismo, estudiar regularmente alejados del instrumento nos ayudará a refrescar y profundizar nuestra visión, y a recordar hacia dónde vamos. Es un contrapeso necesario a lo que pueden ser semanas de estudio esmerado, durante las cuales es fácil desanimarse o quedarse atascado en las laderas musicales.

En este ensayo he destacado el aspecto mental de la labor del intérprete, porque constituye la clave del problema central descrito anteriormente, y también porque es el aspecto que se descuida con mayor frecuencia. En el caso de los intérpretes en edad estudiantil, puede que sea simplemente una cuestión de ignorancia, ya que siguen utilizando los métodos de la infancia y dependiendo del profesor para tomar decisiones. El simple apremio de tiempo puede hacer que los intérpretes adultos omitan las primeras fases del aprendizaje y vayan directamente al instrumento, especialmente a una edad avanzada, cuando se deben dedicar más esfuerzos a los aspectos físicos de la técnica. En cuanto a la educación, uno de los factores es la naturaleza fragmentada del plan de estudios, en el que se enseñan las habilidades auditivas, la armonía y el análisis en compartimientos separados y pocas veces se explica su repercusión directa sobre la interpretación. Y, como profesor, sé lo fácil que es caer en la tentación de las soluciones rápidas, ajustando el sonido mismo en lugar de investigar su causa.

Sin embargo, la razón más común por la que se descuida el aspecto mental es un lógico recelo hacia lo que puede considerarse un enfoque demasiado intelectual. El argumento es que para los músicos es suficiente con ser «musicales». La labor del intérprete es dominar el reto técnico de una partitura; una vez logrado esto, nuestros instintos aportarán la interpretación. De hecho, un exceso de pensamiento puede impedir realmente el funcionamiento natural de esos instintos.

Debemos escuchar con atención lo que surge «naturalmente», por supuesto, pero la idea de que uno puede comprender la sonata de Beethoven

que hemos analizado, simplemente desde una perspectiva natural, parece peligrosamente engañosa. Además, el estudio mental, lejos de reprimir las reacciones instintivas ante la música, está ideado para sacar el mayor provecho de ellas. El problema, como he señalado antes, es la naturaleza de la práctica. Debemos practicar, claro está, pero también debemos reconocer que la práctica puede embotar nuestra inteligencia creativa y enredarnos en hábitos que, al acostumbrarnos a nuestros defectos, no nos permiten escucharnos como somos realmente. Una solución es aprender el máximo posible sobre la obra antes de llevarla al instrumento. Pero el objetivo principal del estudio mental es liberar nuestra musicalidad, asegurar que los objetivos musicales —y no los constreñimientos técnicos— sean lo primero. El ideal es que la música no esté guiada por lo que podamos (o no podamos) hacer, sino por lo que queremos y necesitamos hacer.

## *Notas*

1. Gerald Moore, *Am I too loud? Memoirs of an Accompanist,* Londres, Hamish Hamilton, 1962, pp. 197-198.
2. *Ibid.*, pp. 198-199.
3. Véase Peter Stadlen, «Webern's ideas on the work's interpretation», prefacio a las *Variationen für Klavier* Op. 27 de Anton Webern, Viena, Universal Edition, 1979, pp. v-vii; véanse también las notas del CD de Peter Hill, *Schoenberg, Berg, Webern Piano Music,* Naxos CD 8.553870, 1999, p. 4.
4. Véase Peter Hill (ed.), *The Messiaen Companion,* Londres, Faber and Faber, 1994, pp. 277-278.
5. No pretendo negar la enorme importancia de estudiar grabaciones; sin embargo, debe reservarse para una fase más tardía del estudio, cuando estamos más preparados para hacer una evaluación crítica independiente. Los capítulos 13 y 14 de este libro proporcionan un análisis más amplio de la importancia que pueden tener las grabaciones para los intérpretes.
6. Frédéric Chopin, *Douze Études Op. 10,* ed. Alfred Cortot, París, Éditions Maurice Senart, 1915, p. 5.

## *Lecturas complementarias*

Cone, Edward T.: *Musical Form and Musical Performance,* Nueva York, Norton, 1968.
Menuhin, Yehudi: *Unfinished Journey,* Londres, Macdonald and Janes, 1977.
Moore, Gerald: *Am I too loud? Memoirs of an Accompanist,* Londres, Hamish Hamilton, 1962.
Neuhaus, Heinrich: *The Art of Piano Playing,* K. A. Leibovitch (trad.), Londres, Barrie and Jenkins [1958], 1973.
Solti, Georg: *Solti on Solti,* Londres, Chatto and Windus, 1997.
Stein, Erwin: *Form and Performance,* Nueva York, Knopf, 1962.

# 10 *El lenguaje del cuerpo durante la interpretación*

JANE DAVIDSON

## *Introducción*

Este capítulo describe el importante papel que desempeña la «presentación» en las interpretaciones solistas dentro de la tradición clásica. Como punto de partida reconoce la enorme transición que existe entre el aislamiento de la práctica individual y las interacciones sociales que se producen en el escenario. Sostendrá que el trabajo que tiene lugar durante la práctica ha de prever el contexto social de la interpretación y la inevitable excitación fisiológica y psicológica que esa situación provocará; también considerará la necesidad de que la interpretación sea espontánea y al mismo tiempo consciente de la dinámica social de la experiencia interpretativa. El aspecto visual de la presentación es fundamental en este estudio. Es de vital importancia señalar las múltiples valoraciones que tienen lugar entre el intérprete y el público en función de la información transmitida por señales visuales como las normas en el vestir, los hábitos característicos y otros comportamientos no verbales, como las miradas y los gestos que se realizan mientras se interpreta [1].

## *El papel de la práctica*

Las investigaciones han revelado que la maestría musical implica un elevado nivel de perfeccionamiento de las actividades física y mental, hasta tal punto que la relación entre la información almacenada en la memoria, los esquemas y planes que organizan dicha información y los pensamientos y acciones aso-

ciados a ella se vuelve completamente automática y fluida [2]. La práctica es esencial para lograr esa fluidez. En las etapas iniciales del aprendizaje, los pensamientos y las acciones son torpes, deliberados y difíciles de coordinar, pero después de mucha práctica y experiencia el músico es capaz de tocar sin prestar una atención consciente a los pensamientos y acciones que realiza para producir la interpretación. Sin embargo, la interacción entre el automatismo, el pensamiento consciente y la acción se caracteriza por varios fenómenos curiosos. Por ejemplo, muchos intérpretes dirigen repentinamente su pensamiento consciente al proceso automático de tocar y entonces se percatan de que se encuentran en medio de una pieza, pero sin saber exactamente dónde. La excitación causada por la presencia de otras personas también puede interrumpir la atención y el equilibrio entre los sistemas mentales automático y consciente empleados al tocar. Una forma de desarrollar y aprender a coordinar el pensamiento y la acción automática y consciente es ensayar ante otras personas para intentar familiarizarse con esos efectos producidos por la excitación y ejercitar mecanismos conscientes para mantenerse concentrado y evitar ser distraído por el público.

La música tiene, por supuesto, un marcado contenido emocional: puede inspirar un sentimiento de alegría o contemplación, o hacer que un oyente responda con movimientos, ya sea lenta o enérgicamente. Estas posibilidades se deben examinar ampliamente durante la preparación de una pieza para poder consolidar una interpretación individualizada. Los manuales y textos sobre el aprendizaje de las habilidades musicales frecuentemente recomiendan a los estudiantes que encuentren una articulación óptima de sus ideas expresivas utilizando su cuerpo. Por ejemplo, un cantante puede emplear el gesto de levantar lentamente el brazo para ayudar a transmitir la idea de que está cantando más fuerte a lo largo de una frase musical [3]. Durante el siglo XIX, muchos tratados sobre interpretación mencionaban la interacción esencial entre la estructura musical, la expresión y el cuerpo. Por ejemplo, en su obra *L'Art du violon* de 1834, Baillot sugiere que diferentes tipos de movimiento producen y al mismo tiempo provienen de los efectos relacionados con diferentes tempos musicales. Señaló que un Adagio requiere una ejecución «más amplia» que la de un Allegro, en el que los trinos se deben «agitar» más rápidamente, mientras que la ligereza y la vivacidad han de mantenerse en un Presto y en los pasajes de «mayor desenfreno» [4]. Un siglo después, Alexander Truslit [5], un psicólogo alemán, descubrió que los intérpretes producen interpretaciones musicales conmensurablemente diferentes en respuesta a distintas indicaciones o nociones de movimiento. Además, la información sobre las características estructurales de la música ejecutada y las intenciones expresivas del intérprete es comunicada a los espectadores mediante movimientos corporales [6]. Por ejemplo, un grupo de espectadores a los que se les

mostró un vídeo en el que sólo se veían los movimientos del intérprete fue capaz de diferenciar claramente las interpretaciones de una misma pieza con tres intenciones expresivas diferentes —interpretar sin expresión, con expresión normal o con expresión exagerada.

Considerar el cuerpo como la fuente de la expresión musical implica que ésta constituye un medio para comunicar cualidades básicas de la naturaleza humana, cualidades que emergen del movimiento y que se traducen y abstraen en formas musicales. En efecto, la retórica que rodea a la música revela que la expresión musical utiliza mucho el cuerpo como metáfora —por ejemplo, «ha sido una interpretación conmovedora», o «su canto realmente me animó». Por otra parte, escritores como Ray Jackendoff[7] han sugerido que la capacidad de movimiento que le atribuimos frecuentemente a la música (e.g. «la música corría por sus venas») puede tener realmente un origen corporal (por ejemplo, activa nuestra fisiología del movimiento y equilibrio) o estar asociada a sonidos o imágenes que transmiten una sensación de movimiento corporal. De acuerdo con esta línea de pensamiento, la expresión musical está estrechamente relacionada con quiénes y qué somos. Roger Scruton advierte lo siguiente sobre lo que hacemos los oyentes:

> al escuchar ritmo escuchamos la música como algo activo; parece estar haciendo algo (a saber, bailando) que los sonidos no pueden hacer [...] al mismo tiempo, no creemos que eso esté ocurriendo en el ámbito del sonido: fundamentalmente, somos conscientes del movimiento como algo nuestro[8].

Por lo tanto, al preparar una interpretación, parece ser de vital importancia que examinemos el papel que desempeña el cuerpo en la producción de las características técnicas y expresivas.

## *Observando al intérprete*

Habiendo establecido que el cuerpo es vital en la generación de las cualidades técnicas y expresivas de una interpretación musical, es preciso analizar qué comprenden los movimientos corporales del intérprete. Un extenso estudio de caso de un pianista, realizado por Jane Davidson[9], intentó decodificar los movimientos utilizados en la interpretación y determinar qué información podrían transmitir dichos movimientos al público. En resumen, se descubrió que, para el pianista, una interpretación deliberadamente expresiva siempre implicaba un movimiento giratorio y de balanceo. Generalmente, estos movimientos cíclicos eran expresivos, pero en su transcurso había determinados momentos que eran más expresivos que otros. De hecho, parecía

haber una correlación entre las valoraciones de la expresividad por parte del público y el tipo específico de movimiento que empleaba el intérprete en ese momento. Se encontraron varios gestos identificables: diversos grados de elevación y descenso de brazos y manos, que abarcaban desde ampulosos movimientos giratorios hasta pequeñas rotaciones de las muñecas, un movimiento ondulante de los omoplatos y movimientos frontales y laterales de la cabeza. Se encontró que todos los movimientos, incluyendo el balanceo generalizado, provenían de la región de la cadera. Dado que el pianista estaba sentado, se teorizó que la cadera representaba el punto de apoyo de su centro de gravedad, actuando por lo tanto como pivote para todos los movimientos del torso. Este centro de gravedad parecía ser el punto central en el que se genera la expresión física.

Se buscaron correspondencias entre los gestos utilizados por el pianista y momentos específicos de la estructura musical, pero se descubrió que esos gestos eran empleados de una manera bastante flexible, de modo que en las interpretaciones que se repitieron con intenciones expresivas similares podía aparecer, por ejemplo, un movimiento de cabeza donde previamente se había producido un movimiento ondulante de los omoplatos. Sin embargo, también se descubrió que tales gestos siempre aparecían en los mismos puntos de la música, lo que sugería una estrecha relación entre la producción física de la expresión y su correlativo efecto sonoro expresivo. Por ejemplo, siempre se producía un gesto en el límite o en el clímax de una frase, aunque variaba de interpretación en interpretación.

Las investigaciones sobre la expresión física son escasas, pero se puede establecer un paralelismo entre los descubrimientos de este estudio de caso y el trabajo de James Cutting y sus colaboradores [10]. Este último examina la naturaleza de la expresión física observando la forma de andar de las personas; en este caso, la «expresión» era la identidad y el género de cada persona. Los estudios demostraron que cualquier articulación del cuerpo (por ejemplo, la cadera o el tobillo) proporciona una información igualmente expresiva en el movimiento cíclico del andar. Los investigadores demostraron que, cuando una persona anda, hay un punto en su cuerpo que funciona como una referencia para los movimientos de todas las demás partes. Al tocar el piano, no todo el cuerpo participa en la misma medida en la ejecución; por lo tanto, el punto de referencia para los movimientos expresivos en el piano estará inevitablemente relacionado con la posición sentada y el centro de gravedad, y, dados los hallazgos de Davidson, es muy probable que se encuentre en el movimiento de balanceo.

La descripción de Cutting proporciona claramente una base para comprender la importancia del movimiento de balanceo en las interpretaciones pianísticas. Además, el autor ha refinado su descripción [11] para explicar el ha-

llazgo de que distintas partes del cuerpo transmiten información igualmente expresiva, y propone que algunas zonas del cuerpo son indicadoras generales de expresión, mientras que otras partes proporcionan una información más específica. Esto explica por qué en la ejecución pianística encontramos una información expresiva continua (movimiento de balanceo) y una cantidad mucho mayor de información localizada (por ejemplo, un gesto específico de la mano). Es posible que los indicadores locales actúen cuando el potencial expresivo del cuerpo es mayor. De hecho, muchos de estos movimientos sumamente pronunciados tuvieron lugar en momentos en los que el cuerpo se encontraba «libre»: por ejemplo, una gran elevación del brazo ocurría frecuentemente cuando éste no debía cumplir ningún requerimiento técnico, como durante un acorde mantenido en una cadencia.

Como se ha sugerido, la información que transmiten los movimientos está totalmente al alcance del público, que la utiliza para evaluar y enjuiciar una interpretación. Por lo tanto, parecería que tales movimientos se deberían reconocer y fomentar, en lugar de desaprobarlos. De hecho, un estudio a pequeña escala [12] realizado con un grupo de pianistas reveló que no podían tocar con una expresión óptima si sus cuerpos se encontraban constreñidos de alguna manera durante el proceso de aprendizaje o de interpretación. Por consiguiente, encontrar la intención física y mental adecuada, y permitir que ésta se comunique libremente a través del cuerpo, parece ser esencial para la producción de una interpretación fluida y expresiva.

Naturalmente, los movimientos y gestos han de adecuarse al contexto social y cultural. Por ejemplo, los pianistas clásicos tocan sentados y se mueven relativamente poco, mientras que un teclista de rock puede tocar de pie, bailando sin moverse de su sitio y haciendo gestos ampulosos. Éstos son tan sólo dos ejemplos de las maneras específicas en las que los intérpretes aprenden a presentar sus interpretaciones.

### *Factores sociales*

La etiqueta social de la interpretación, incluyendo conductas como hacer reverencias y vestirse adecuadamente, es significativa para los intérpretes y el público [13]. En efecto, un estudio reciente sobre los procesos de evaluación musical [14] revela que los evaluadores de interpretaciones vocales consideran que el atractivo físico del intérprete es uno de los criterios más importantes en su evaluación de interpretaciones solistas. Curiosamente, los evaluadores consideran que para comunicar lo más íntimo del contenido musical y hacer frente a la naturaleza pública de la interpretación el intérprete ha de revelarse a sí mismo y presentar su personalidad de una manera especialmente intere-

sante o atractiva. De este modo, los gestos pomposos que hace Luciano Pavarotti con el brazo, agitando un pañuelo mientras canta, representan un acto público y exagerado de expresión emocional, mientras que, al mismo tiempo, los sutiles movimientos de sus ojos y su boca cuando canta una línea melancólica parecen expresar sus pensamientos y emociones más profundas, convirtiéndose en actos de intensa revelación emocional y personal.

Naturalmente, todas las conductas que implican movimientos se aprenden de otras personas, y en la interpretación musical se puede reconocer con frecuencia al profesor del alumno por la forma en que éste aborda el instrumento. Por ejemplo, hay una profesora de violín muy respetada que enseña a todos sus alumnos a doblar las rodillas de un modo particular para añadir peso y relajación a su ejecución. Dicha profesora es inmediatamente reconocible en los movimientos de cada uno de sus alumnos.

Martin Gellrich [15] ha investigado gestos superficiales aprendidos, como los que hace Pavarotti con el brazo, y la manera en que éstos interactúan con los movimientos expresivos esenciales e intrínsecos, como los gestos de los ojos y la boca. Gellrich sugiere que los movimientos y gestos aprendidos aportan una intención expresiva a la interpretación y pueden tener efectos tanto positivos como negativos en la producción de la interpretación. Pueden aportar una información al espectador que le ayudará a comprender la interpretación, ya que los gestos pueden esclarecer significados, aunque el movimiento en sí mismo sea «superfluo» en la producción del conjunto de la música. Puede haber un nivel «superficial» de movimiento que el intérprete añade a la interpretación. A la inversa, si estos gestos no son coherentes con las intenciones del intérprete, pueden provocarle tensiones físicas, al inhibir la fluidez y causar una incongruencia entre el gesto empleado y el objetivo de la interpretación.

Sverkar Runeson y Gunnar Frykholm [16] han descrito las consecuencias negativas de los gestos incongruentes en un trabajo que demuestra que las disposiciones mentales se hacen explícitas en el movimiento. Utilizando como experimento la simple tarea de levantar una caja, los investigadores pidieron a un grupo de observadores que dieran información sobre lo que podían ver cuando una persona levantaba una caja, y descubrieron que los observadores podían detectar no sólo el peso de la caja, sino también cuánto difería ese peso del que esperaba encontrar el que la levantaba. Más relevante aún es el hecho de que los observadores detectaron la información falsa que se les dio en algunos casos sobre el peso de la caja. Por lo tanto, la información engañosa y el peso verdadero de la caja se explicitaban conjuntamente.

Los gestos «superficiales» pueden contribuir de manera importante a la producción y percepción de una interpretación musical. En efecto, en el estudio de las interpretaciones del pianista mencionado anteriormente, otra in-

terpretación del hallazgo de que ciertos fragmentos de dos segundos eran mucho más expresivos que otros podría ser que en algunos momentos de la ejecución el pianista utilizaba gestos imitativos, y que esos movimientos acentúan el efecto expresivo del momento en cuestión. Por ejemplo, un amplio movimiento lateral de la cabeza puede tener una forma propia y característica y, sin embargo, formar parte del continuo movimiento de balanceo. Los estudios sobre los gestos en el lenguaje verbal [17] indican la existencia de repertorios gestuales asociados a significados específicos; por ejemplo, el dedo pulgar levantado es un gesto emblemático en la cultura occidental que significa «bien». Es posible que el pianista hubiera desarrollado gestos específicos para determinadas expresiones musicales —en otras palabras, un repertorio de movimientos gestuales.

### *Presentar una interpretación*

Es evidente que el público distingue todo tipo de informaciones sutiles sobre los intérpretes a partir de sus movimientos corporales y su aspecto general y, por supuesto, de los sonidos musicales que produce. El intérprete también percibe muchas señales diferentes del público: el nivel general de entusiasmo y concentración, así como la forma de vestir, de sentarse, etc. De modo que existe una comunicación recíproca. Pero al presentar una interpretación, el músico no está sencillamente presentándose en un espacio público y ofreciendo una versión bien ensayada de una obra musical, atribuyéndose cierto mérito por la excitación que produce la presencia del público. Ocurren muchas más cosas: el intérprete y el público intercambian información continuamente mediante señales visuales y auditivas. En tales circunstancias, el intérprete sigue el «fluir» del momento [18]. Los investigadores han observado que si el intérprete percibe las numerosas señales del contexto de la interpretación en vivo y las interpreta positivamente, puede alcanzar un estado de conciencia psicológica diferente que le permite concentrarse profundamente en su tarea y analizar pensamientos y sentimientos espontáneos de una manera creativa [19].

Así pues, en la interpretación en vivo se pueden tomar decisiones diferentes sobre la interpretación y la interacción con otros intérpretes y el público. Como se ha mencionado en la sección sobre aprendizaje, la excitación psicológica y física es consecuencia inevitable de las interacciones sociales [20]. En niveles de rendimiento óptimos, el ritmo cardíaco acelerado y el consecuente aumento de la provisión de oxígeno causados por la producción de adrenalina pueden conducir a la circunstancia intensificadora referida anteriormente como el «fluir». Sin embargo, una sobreexcitación puede provocar falta de

confianza y preocupación sobre asuntos como la capacidad de recordar, que será de vital importancia para el éxito o el fracaso. De hecho, un elevado nivel de adrenalina en combinación con esos pensamientos puede provocar una reacción en cadena de consecuencias fisiológicas nefastas para la interpretación: temblores que conducen a errores, confusión de pensamientos que produce fallos en la memoria, etc.

## *Resumen*

En conclusión, la interpretación musical es una actividad sumamente expresiva y abstraída que emana de una base corporal. Para el intérprete, es necesario un equilibrio cuidadoso entre el control y la entrega, entre obedecer ciertas reglas y desobedecer otras. La interpretación supone un conocimiento profundo de la música, del instrumento y del propio yo en la música. No es tarea fácil llegar a ser un intérprete, mucho menos un solista, por la enorme preparación que requiere, pero presentar una interpretación parece ser una oportunidad inusual para los intérpretes de compartir con otros quiénes y qué es lo que son exactamente en ese momento.

## *Notas*

1. Para más detalles sobre estos temas, véase Jane W. Davidson, «The social in musical performance», en David J. Hargreaves y Adrian C. North (eds.), *The Social Psychology of Music,* Oxford, Oxford University Press, 1997, pp. 209-228. El capítulo 11 del presente libro analiza la naturaleza de la comunicación visual y auditiva entre los intérpretes de una agrupación.
2. Para un estudio general de la habilidad, véase John A. Sloboda, *The Musical Mind: The Cognitive Psychology of Music,* Oxford, Oxford University Press, 1985, pp. 215-236; véase también el capítulo 7 de este libro.
3. Véanse, por ejemplo, las indicaciones dadas en Esther Salaman, *Unlocking Your Voice,* Londres, Victor Gollancz, 1990; y Eloise Ristad, *A Soprano on Her Head: Right-side Up Reflections on Life and Other Performances,* Moan, Ut., Real People Press, 1982.
4. Citado por Robin Stowell, *Violin Technique and Performance Practice in the Late Eighteenth and Early Nineteenth Centuries,* Cambridge, Cambridge University Press, 1985, p. 271.
5. Para una traducción y resumen, véase Bruno H. Repp, «Music as motion: a synopsis of Alexander Truslit's (1938) *Gestaltung und Bewegung in der Musik*», *Psychology of Music* 21 (1993), pp. 48-73.
6. Véanse Jane W. Davidson, «Visual perception of performance manner in the movements of solo musicians», *Psychology of Music* 21 (1993), pp. 103-113; y, de la misma autora, «What type of information is conveyed in the body movements of solo musician performers?», *Journal of Human Movement Studies* 6 (1994), pp. 279-301.

7. Ray Jackendoff, *Consciousness and the Computational Mind,* Cambridge, Mass., MIT Press, 1988, p. 131.
8. Roger Scruton, *The Aesthetic Understanding: Essays in the Philosophy of Art and Culture,* Londres, Methuen, 1983, p. 84.
9. Véanse Jane W. Davidson, «The perception of expressive movement in music performance» (tesis doctoral, City University, Londres, 1991); «What type of information is conveyed in the body movements a solo musician performers?», cit.; y, «Understanding the expressive movements of a solo pianist», *Deutsches Jahrbuch für Musikpsychologie* (de próxima aparición).
10. James E. Cutting y Lynne T. Kozlowski, «Recognising friends by their walk: gait perception without familiarity cues», *Bulletin of the Psychonomic Society* 9 (1977), pp. 353-356; y James E. Cutting, Dennis R. Proffitt y Lynne T. Kozlowski, «A biomechanical invariant for gait perception», *Journal of Experimental Psychology: Human Perception and Performance* 4 (1978), pp. 357-372.
11. Véase James. E. Cutting y Dennis R. Proffitt, «Gait perception as an example of how we may perceive events», en Robert D. Walk y Herbert L. Pick (eds.), *Intersensory Perception and Sensory Integration,* Nueva York, Plenum Press, 1981, pp. 37-59.
12. Jane W. Davidson y Jeremy C. Dawson, «The development of expression in body movement during learning in piano performance», en David Wessel (ed.), *Conference Proceedings of Music Perception and Cognition Conference,* Berkeley, University of California Press, 1995, p. 31.
13. Descrita detalladamente en Jane W. Davidson «The social in musical performance», cit.
14. Jane W. Davidson y Daniela C. C. Coimbra, «Investigating performance evaluation by assessors of singers in a music college setting», *Musicae Scientiae* 5 (2001), pp. 33-54.
15. Martin Gellrich, «Concentration and tension», *British Journal of Music Education* 8 (1991), pp. 167-179.
16. Sverkar Runeson y Gunnar Frykholm, «Kinematic specification of dynamics as an informational basis for person-and-action perception: expectations, gender, recognition and deceptive intention», *Journal of Experimental Psychology: General* 112 (1983), pp. 585-615.
17. Véanse la obra de Paul Elkman y Wallace V. Friesen, «The repertory of nonverbal behaviour: categories, origins, usage, and coding», *Semiotica* 1 (1969), pp. 49-204; y Andrew Ellis y Geoffrey Beattie, *The Psychology of Language and Communication,* Londres, Weidenfeld and Nicholson, 1986.
18. Véanse Mihalyi Csikszentmihalyi, *Finding Flow: The Psychology of Engagement with Everyday Life,* Nueva York, Basic Books, 1997.
19. Para más detalles, véase Jane W. Davidson, Stephanie E. Pitts y Jorge S. Correia, «Reconciling technical and expressive elements in young children's musical instrument learning», *Journal of Aesthetic Education* (de próxima aparición).
20. Véase un resumen detallado en Glenn D. Wilson, «Performance Anxiety», en David J. Hargreaves y Adrian C. North (eds.), *op. cit.,* pp. 229-248; véase también el capítulo 12 de este libro.

## *Lecturas complementarias*

Clarke, Eric, y Jane Davidson: «The body in performance», en Wyndham Thomas (ed.), *Composition, Performance, Reception: Studies in the Creative Process in Music,* Aldershot, Ashgate, 1998, pp. 74-92.

Evans, Andrew: *The Secrets of Musical Confidence,* Londres, Thorson, 1994.
Hargreaves, David J., y Adrian C. North (eds.): *The Social Psychology of Music,* Oxford, Oxford University Press, 1997.
Sloboda, John A.: *The Musical Mind: The Cognitive Psychology of Music,* Oxford, Oxford University Press, 1985.

# 11 *La interpretación en grupo*

ELAINE GOODMAN

La interpretación en grupo implica una interacción musical y social entre varios intérpretes. El término *ensemble* significa «juntos» en francés, y se utiliza para designar una variedad aparentemente infinita de agrupaciones musicales, desde un dúo hasta una orquesta sinfónica. Al mismo tiempo, y al igual que su equivalente español «conjunto», también refiere el grado de precisión con el que los músicos interpretan juntos: se dice que un grupo está «muy conjuntado» cuando es bueno, mientras que uno inferior puede estar «poco conjuntado».

Nuestra experiencia en interpretación en grupo puede mejorar entendiendo algunos de los procesos involucrados en el acto de hacer música juntos. Este capítulo considera cuatro aspectos en particular: la coordinación, la comunicación, el papel del individuo y los factores sociales. La mayoría de las observaciones se refieren a pequeños grupos de cámara dentro de la tradición occidental, pero de vez en cuando también se hará referencia a grupos más grandes.

## *Coordinación: mantener el tempo*

El requisito más importante de cualquier conjunto es que las partes individuales encajen entre sí. Es necesario, por tanto, que cada músico sea capaz de tocar a tempo con el resto del grupo; en efecto, la coordinación de un conjunto depende únicamente del tempo. Cualquier intérprete de grupo consumado sabe que contar es vital para mantener el tempo, al menos para saber

qué compás están tocando y cuándo entrar o parar. Sin embargo, hay otros asuntos que considerar en relación a la coordinación del tempo, y analizaremos tres de ellos en esta sección: el «reloj» del grupo, las habilidades para mantener el tempo y la ilusión de sincronía.

### *El «reloj» del grupo*

El punto de partida de un ensayo de un conjunto pequeño, sin director, es establecer el tempo general de una pieza, lo cual puede consistir en probar diferentes velocidades hasta acordar un tempo determinado. En cambio, el director de una agrupación grande indicará a los intérpretes la velocidad del pulso básico. En ambos casos, el pulso básico (i.e., el tempo general) proporciona un punto de referencia para mantener el tempo. En efecto, cada músico generará un pulso interno de acuerdo con el compás (que a veces es visible cuando marcan con el pie, aunque tal vez sería mejor que se limitaran a mover los dedos del pie, de modo que no se vea). El tempo general funciona por tanto como el reloj del grupo, ya que proporciona una fuente de coordinación y controla el pulso interno de cada músico. En otras palabras, los intérpretes de cámara poseen un «marcador de tiempo compartido» para el tempo general de la obra [1].

Mientras el pulso básico es una referencia importante para mantener el tempo, es posible que en determinadas circunstancias algunos músicos quieran subdividir o agruparlo en diferentes puntos. En esos casos, el compás constituye una estructura importante para la coordinación, ya que es una unidad de tiempo en la que podemos organizar individualmente el pulso básico.

### *Técnicas para mantener el tempo*

Tocar a tempo con otras personas requiere algo más que simplemente contar y seguir el tempo general, ya que, durante la interpretación, la ejecución de cada compás ha de estar cuidadosamente administrada. Ello implica dos técnicas básicas: la *anticipación* y la *reacción*. Los intérpretes de cámara realizan predicciones complejas que están estrechamente relacionadas con las reacciones adquiridas por retroalimentación o *feedback*: sobre la base de la nota anterior, ¿cuándo sonará la nota siguiente de un compañero? Michael Tree, integrante del Cuarteto Guarnieri, afirma que «cada momento de nuestra ejecución está condicionado por lo que acaba de ocurrir o lo que pensamos que ocurrirá a continuación» [2]. La anticipación de cada pulso y la reacción a

la producción de cada pulso están prácticamente definidas por la naturaleza de la interacción manifiesta entre los intérpretes.

Los intérpretes de cámara actúan recíprocamente de diversas maneras para mantener el tempo. Una de ellas puede compararse con el proceso de *cazar,* en el sentido de que un músico sigue a otro del mismo modo en que un cazador sigue la pista de su presa, anticipándose y reaccionando a sus movimientos. El proceso de «cazar», que también se puede considerar como seguir o rastrear, es unidireccional, ya que uno de los músicos responde al tempo de otro(s) [3].

El segundo tipo de interacción implica un ajuste mutuo, o *cooperación,* entre los intérpretes. La cooperación es un proceso dinámico (en contraposición al unidireccional), aunque también puede presentar elementos de «caza», ya que los intérpretes «intercambian roles [...] para permitir que uno u otro sea el guía» [4]. En efecto, la línea que separa a la cooperación de la caza es muy delgada, y ambas habilidades pueden verse afectadas por factores como una personalidad dominante o la naturaleza de la música (por ejemplo, es posible que el acompañante siga al intérprete que ejecuta la melodía).

Las interacciones musicales pueden planificarse hasta cierto punto en los ensayos: quién seguirá a quién durante un pasaje determinado, o quién será el guía en el pasaje subsiguiente. Este tipo de planificación consciente puede ser provechosa para la coordinación, pero no equivale a la «caza» y cooperación del momento que ocurren necesariamente a lo largo de la interpretación.

## *La ilusión de sincronía*

Cuando se mantiene el tempo con exactitud, se obtiene la coordinación de las notas que produce cada músico en un grupo, pero la realidad de tocar juntos no es tan sencilla. De hecho, ejecutar las notas exactamente al mismo tiempo sobrepasa los límites de la habilidad y la percepción humanas: siempre habrá pequeñas diferencias de tempo —es decir, asincronismo— entre las notas que deben ser ejecutadas simultáneamente [5]. La figura 11.1 ilustra el asincronismo que podría producirse en un trío de cuerda, mostrando que las notas del violín pueden sonar, por término medio, unas fracciones de segundo antes que las del violonchelo y la viola respectivamente. Por tanto, el arte de tocar juntos es crear la *ilusión* de estar perfectamente conjuntados [6].

Hay varios factores que tener en cuenta cuando se intenta lograr la coordinación en un conjunto, especialmente las diferencias físicas entre los instrumentos (incluyendo la voz humana). Los músicos deben ser conscientes del tiempo que necesita cada instrumento para hacer sonar una nota. Por

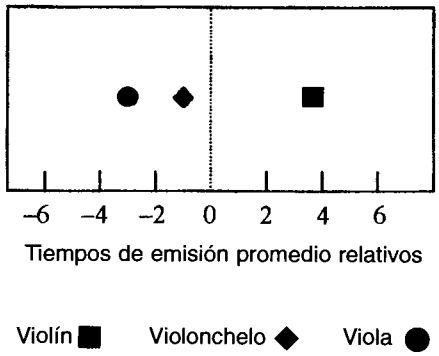



FIGURA 11.1 *Ejemplo hipotético de los tiempos de emisión promedio, en milisegundos, de un trío de cuerda (adaptado de Rudolf A. Rasch, «Timing and synchronisation in ensemble performance»); nótese que el violín es el primero, le sigue el violonchelo y por último la viola.*

ejemplo, un instrumento de viento suena antes que uno de cuerda porque su *tiempo de subida* —el tiempo que transcurre entre la ejecución de una nota y su percepción auditiva— es más breve.

El reto que deben afrontar los intérpretes de grupo es controlar la percepción de las notas de cada instrumento y cada músico, particularmente cuando se utilizan diversas combinaciones de instrumentos. Además pueden surgir problemas cuando los intérpretes están colocados a diferentes distancias del público: el director de orquesta, por ejemplo, debe considerar el tiempo que tarda el sonido que producen los músicos al fondo del escenario en ser percibido por el público.

La velocidad de la música también influye en la coordinación de un conjunto: las pruebas demuestran que es más difícil sincronizar las notas en los tempos lentos, ya que los músicos necesitan subdividir el pulso básico[7]. Además, los pasajes que contienen cambios de tempo y las entradas después de una pausa o silencio son más difíciles de coordinar. En estas últimas, la ausencia de sonido durante un silencio implica que los intérpretes dependen de la comunicación visual —así como del «reloj» del grupo— para ayudar a coordinar una entrada.

### *Comunicación: señales auditivas y visuales*

La información comunicada entre los intérpretes de un conjunto es principalmente auditiva y visual: constantemente se transmiten señales mediante

sonidos, contactos visuales, gestos y movimientos del cuerpo. Los intérpretes de cámara se observan y escuchan unos a otros no sólo para coordinar las acciones, sino también para comunicar ideas sobre la expresión o la interpretación de la música. Sin embargo, la comunicación auditiva (poder oírse unos a otros) es más importante que la visual (poder verse unos a otros). La explicación más sencilla es que oímos la música —no la vemos [8]. Las siguientes secciones se centrarán respectivamente en esos dos aspectos de la comunicación, considerándolos hasta la planificación de señales auditivas y visuales.

### *Comunicación auditiva*

Los intérpretes de cámara necesitan oírse unos a otros para lograr la coordinación, es decir, para asegurar que las partes individuales suenen juntas en los sitios correctos. En un nivel más importante, sin embargo, cada intérprete escucha continuamente los matices expresivos de sus compañeros, como las fluctuaciones de tempo, las gradaciones dinámicas y los cambios en la articulación, el timbre y la entonación. En efecto, la concentración de cada músico está repartida entre controlar su propio sonido y atender al sonido que producen los demás intérpretes del grupo. Por consiguiente, realizan pequeños ajustes, consciente o inconscientemente, en el equilibrio y la calidad del sonido del grupo a lo largo de la interpretación.

Independientemente del tamaño del grupo, todos los matices que emergen de las partes individuales influyen en la expresión de la música. Las dos técnicas mencionadas anteriormente —la anticipación y la reacción— entran en juego cuando los intérpretes intentan predecir y responder a las señales que transmiten sus compañeros. Al mismo tiempo, también contribuyen al sonido produciendo sus propias señales. Por ejemplo, si un intérprete de un grupo pequeño comienza un *crescendo*, los demás músicos pueden ajustar la intensidad de sus respectivos sonidos aumentando correspondientemente su nivel dinámico o tal vez disminuyéndolo para que el *crescendo* individual predomine en el sonido general. Además, los intérpretes pueden predecir las acciones físicas entre ellos basándose en un conocimiento auditivo; por ejemplo, si en un quinteto de viento la intensidad del tono de un oboísta comienza a decaer, los compañeros pueden acelerar el tempo hacia el final de la frase para evitar que el oboísta se quede sin aliento. De igual manera, el grupo puede reaccionar ante los cambios de sonido compensando sus acciones entre ellos: en el caso del oboísta, los compañeros pueden aumentar sus niveles dinámicos para mantener la intensidad del sonido general o disminuirlos para asegurar que la parte de oboe se proyecte por encima del grupo.

Algunos matices expresivos se pueden planificar en los ensayos, de modo que los intérpretes puedan al menos comprender (si no responder a) las intenciones de cada uno, aunque, paradójicamente, algunos intérpretes ven con recelo la planificación metódica de los matices expresivos [9]. Además, las investigaciones han demostrado que la práctica prolongada puede reducir la habilidad para controlar la expresión del sonido durante la interpretación, ya que las ideas se vuelven fijas y más difíciles de adaptar. Caroline Palmer afirma que «el control de las características expresivas de una pieza determinada será mayor si no se ha practicado prolongadamente [...] La práctica prolongada de una interpretación concreta puede hacer que los gestos expresivos sean más difíciles de reprimir [10].

Este asunto tiene consecuencias trascendentales para los intérpretes de grupo, porque sugiere que cada uno debe ser cauteloso en cuanto al tiempo que dedica a sus prácticas individuales para una interpretación en conjunto. Después de la práctica individual, al intérprete le puede resultar más difícil «reprimir» sus propias ideas expresivas para mezclarse con el sonido general. Esta observación contradice las creencias valoradas por algunos intérpretes de grupo, como por ejemplo Abram Loft, quien recomienda: «desarrolla cuanto puedas tu propia idea de la pieza y su mundo particular [...] después evalúa y modifica esa concepción desde la perspectiva de las opiniones expresadas por otros músicos» [11]. Al contrario, William Pleeth aconseja a sus lectores que practiquen «con todo el lienzo a la vista» y que desarrollen el sonido «teniendo en mente el concepto total» [12].

De igual manera, las estrategias desarrolladas en los ensayos de grupo pueden ser más difíciles de ajustar si los músicos no son capaces de desviarse de sus planes. Algunos intérpretes, como los pianistas acompañantes experimentados, muestran una extraordinaria flexibilidad en el trabajo colectivo y pueden (con bastante razón) rechazar la paradoja anterior: ellos son capaces de adaptar rápidamente y sin ningún problema sus propias intenciones expresivas para acomodarlas o combinarlas con otra parte después de haber practicado poco o mucho.

### *Comunicación visual*

Los intérpretes también pueden comunicar ideas entre ellos sobre la expresión utilizando el lenguaje corporal, es decir, por el medio visual. Jane Davidson advierte que «las expresiones físicas de carácter musical proporcionan un medio para comprender y compartir las intenciones musicales» [13]. Por ejemplo, los intérpretes de un cuarteto de cuerda no se moverán exageradamente durante un pasaje calmado y lento; mantendrán, en cambio, una postura fija o, como mucho, un suave balanceo del torso superior.

La coordinación del sonido se puede mejorar planificando señales visuales —por ejemplo, determinando quién da la entrada al comienzo de la pieza o quiénes establecerán un contacto visual después de una pausa prolongada. Como he sugerido antes, sin embargo, tales coreografías pueden menoscabar la espontaneidad de la interpretación. Curiosamente, los miembros del Cuarteto Guarnieri afirman que captan mutuamente sus movimientos físicos en un nivel inconsciente durante un período, de tal manera que apenas son conscientes de las señales visuales transmitidas durante la interpretación: «Cada uno de nosotros tiene un determinado lenguaje corporal cuando toca. Llegas a conocer los de tus colegas y reaccionas adecuadamente. Con los años, gran parte de ello llega a ser intuitivo» [14]. En ese sentido, nos recuerdan que los intérpretes de cámara no tienen necesariamente que mirarse directamente unos a otros, ya que la visión periférica siempre está funcionando.

Como he señalado antes, los intérpretes de grupo se fían mucho más de la comunicación auditiva (la información que sus compañeros expresan en sonido) que de la comunicación visual para hacer música juntos. Sin embargo, las investigaciones revelan que la retroalimentación visual «contribuye significativamente a la precisión y la libertad expresiva de las interpretaciones colectivas» [15]. Los directores, cuya forma de comunicación es puramente visual, pueden ayudar a mejorar la coordinación del tempo en la interpretación colectiva. Los intérpretes de orquesta observan constantemente la batuta del director para aprender a anticipar y responder. Para un director, sin embargo, las señales visuales son transmitidas sólo parcialmente por la batuta. El movimiento de la batuta de Simon Rattle, por ejemplo, está incorporado dentro de su movimiento corporal general —que incluye gestos aparentemente caprichosos, cambios de postura y diversas expresiones faciales— que transmite información sobre el talante y el carácter de la música. En efecto, el director comunica mucho más que un simple pulso, ya que los miembros de la orquesta pueden percibir señales visuales sobre la expresión a través del lenguaje corporal del director, de la misma manera en que los intérpretes de un cuarteto de cuerda pueden proyectar ideas interpretativas observando los movimientos físicos entre ellos.

## *El papel del individuo*

Un intérprete de una agrupación musical desempeña un papel similar al de un actor en una obra de teatro. El actor desarrolla su personaje por su cuenta, pero ese personaje cambia cuando entra en contacto con los demás personajes de la obra, especialmente porque es entonces cuando se establece el ritmo y la expresión de cada parte. Igual que los actores, los músicos de un

conjunto moldean y ajustan sus partes para encajarlas entre sí. Sin embargo, el actor siempre tiene presente a su propio personaje, de modo que la obra se define simultáneamente por cada actor como por la combinación de todos los actores. Lo mismo ocurre con la música de conjunto, ya que cada músico imprime su propio sello al grupo, pero también intenta fusionarse con él. Esta relación entre el individuo y el grupo será analizada a continuación.

En 1996, Mitch Waterman investigó las reacciones emocionales de los intérpretes. Pidió a los músicos de un dúo de piano y un dúo de piano y violonchelo que comentaran, después de una interpretación, sus sentimientos según los momentos en que consideraban que «la música despertaba algo» [16]. Waterman observó que cada intérprete tenía respuestas emocionales diferentes ante la música, por lo que concluyó que «los intérpretes no coinciden en sus opiniones sobre los momentos emotivos dentro de una interpretación conjunta» [17]. Esto sugiere que los individuos no necesariamente experimentan lo mismo cuando tocan juntos, ni comparten ideas sobre una interpretación, sino que mantienen identidades bastante independientes.

Existen muchos tópicos en torno a la noción de conjunto musical que menoscaban el papel del individuo: por ejemplo, la agrupación ideal se ha descrito como «un organismo musical integrado que refleja la perspicacia colectiva de todos sus miembros» [18]. En realidad, un conjunto es mucho más complejo, y ciertamente menos «integrado», de lo que sugiere esa descripción, ya que está relacionado tanto con los caracteres individuales como con la combinación de éstos. Entonces, ¿qué sucede con el individuo cuando toca en un grupo?

El siguiente análisis se centra en el caso de una pianista profesional que interpreta el tercer movimiento (Largo) de la Sonata para violonchelo en Sol menor, op. 65, de Chopin. En esta investigación comparé determinadas características expresivas —tempo y dinámica— de dos interpretaciones previas a un ensayo del dúo: la primera fue individual de la pianista, es decir, que interpretó su parte en un contexto solista, y la segunda fue con el violonchelista, o sea, en un contexto de grupo. La pianista preparó su parte sin realizar consultas con el violonchelista, de modo que la interpretación a dúo representó la primera ocasión en que los dos intérpretes se reunieron y tocaron la obra juntos. Por tanto, los datos obtenidos reflejan la manera en que cambió la realización «a solo» de la pianista cuando tocó con el violonchelista.

La figura 11.2 muestra, compás por compás, las fluctuaciones de tempo de ambas interpretaciones [19]. En primer lugar, es evidente que el tempo general de la interpretación a dúo fue más lento que el de la interpretación a solo de la pianista (lo cual se puede ver en el gráfico porque el perfil de tempo de la interpretación a dúo se encuentra por debajo del perfil de tempo de la interpretación a solo). Esto demuestra que la interpretación de la pianista se

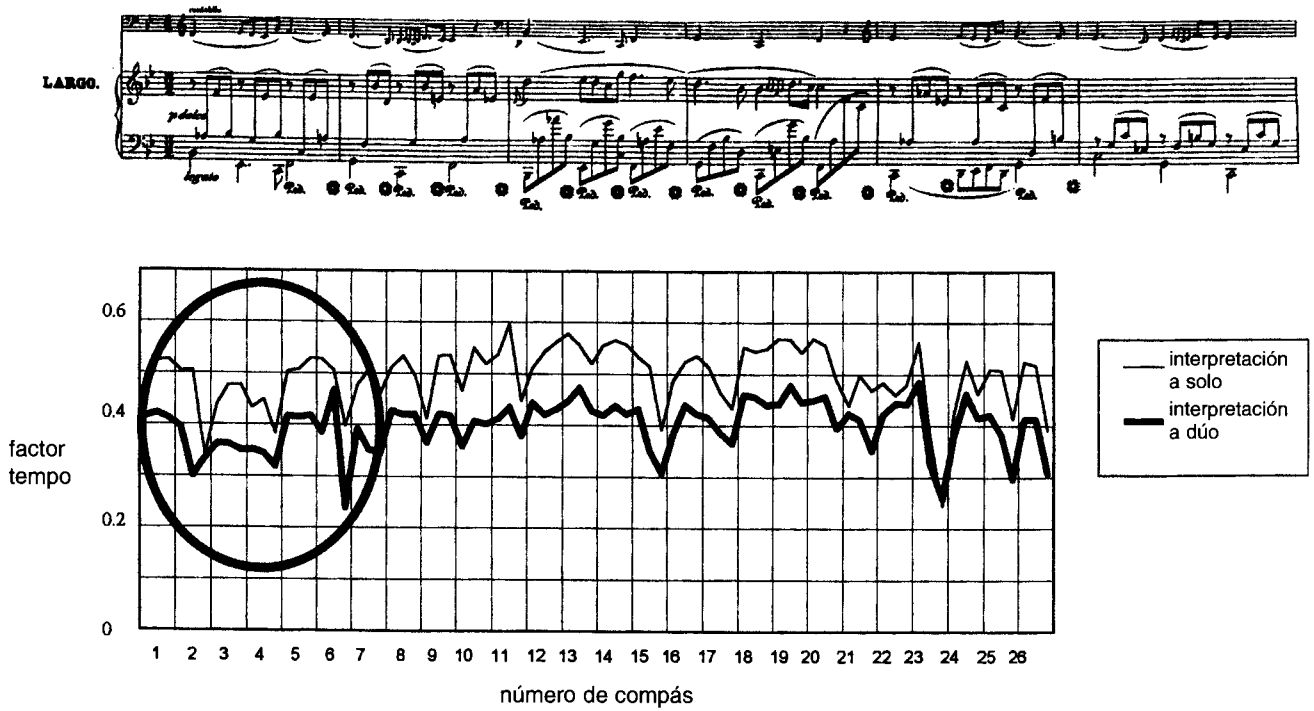



FIGURA 11.2 *Perfiles de las fluctuaciones de tempo de las interpretaciones a solo y a dúo de la Sonata para violonchelo op. 65 de Chopin, tercer movimiento, compases 1-26; se anexa un fragmento de la partitura (compases 1-6 de la primera edición alemana; Leipzig, Breitkopf & Härtel, 1846) para facilitar las referencias que se hacen en el análisis a los compases agrupados dentro del círculo.*

vio afectada inmediatamente al tocar con el violonchelista, ya que tuvo que negociar su tempo para ajustarlo al del violonchelista. Sin embargo, los puntos más altos y más bajos de ambos perfiles son muy similares, lo cual sugiere que parte de las tendencias «a solo» de la pianista fueron mantenidas en el dúo. Los niveles de *rubato*, sin embargo, fueron diferentes en las dos interpretaciones: en general, hubo más *rubato* en la interpretación a solo de la pianista que en la interpretación conjunta (compárense, por ejemplo, la medida de las fluctuaciones en los compases 7-11 y la proporción de *ritardando* al final de los compases 15 y 23). Parece ser, por tanto, que el *rubato* de la pianista disminuyó en la interpretación a dúo.

Los primeros seis compases merecen especial atención (véase el fragmento de la partitura y los compases del gráfico marcados con el círculo en la figura 11.2), ya que en ellos las fluctuaciones de tempo de la pianista difieren notablemente en las dos interpretaciones. (Esto refleja lo que ocurrió durante los primeros momentos de su interpretación conjunta.) En su interpretación a solo, la pianista redujo la velocidad al final del compás 2, mientras que en la interpretación a dúo el tempo disminuyó durante los compases 1 y 2, particularmente hacia la mitad del segundo compás (el punto en el que el violonchelista ejecuta un grupeto en la línea melódica). La ralentización en la interpretación a dúo alude a la manera en que la pianista adoptó gradualmente el tempo general del violonchelista, y acomodó la ornamentación de la línea melódica de éste. Además, la pianista cambió su realización de la segunda frase (compases 3-4) en la interpretación a dúo, ya que el tempo aumentó justo antes de la frase en los compases 2-3, pero después se mantuvo relativamente estable. En efecto, en cuanto la pianista tocó con el violonchelista, ajustó su interpretación para encajarla con la otra parte, por lo que las fluctuaciones de tempo fueron diferentes en los dos contextos.

La figura 11.3 muestra, compás por compás, las fluctuaciones dinámicas de la pianista en las mismas interpretaciones, a solo y a dúo [20].

Mientras que los puntos más altos y más bajos de los dos perfiles coinciden generalmente, es evidente que la pianista tendía a tocar un poco más suave en el dúo (compárense, por ejemplo, los niveles dinámicos de los compases 1-6 y 10-15). Por lo tanto, en este caso también se observa cierto constreñimiento en la interpretación a dúo de la pianista, aunque dentro de la misma gama dinámica general.

El pasaje de los compases 21-23 (véanse los compases dentro del círculo en la figura 11.3) es particularmente interesante porque la pianista tocó más fuerte en su interpretación con el violonchelista que en su interpretación a solo. Este pasaje constituye el clímax principal del movimiento y está marcado por un cambio de textura (como puede verse en el fragmento de la partitura): el movimiento en corcheas en la parte de piano cesa en el compás 21 y

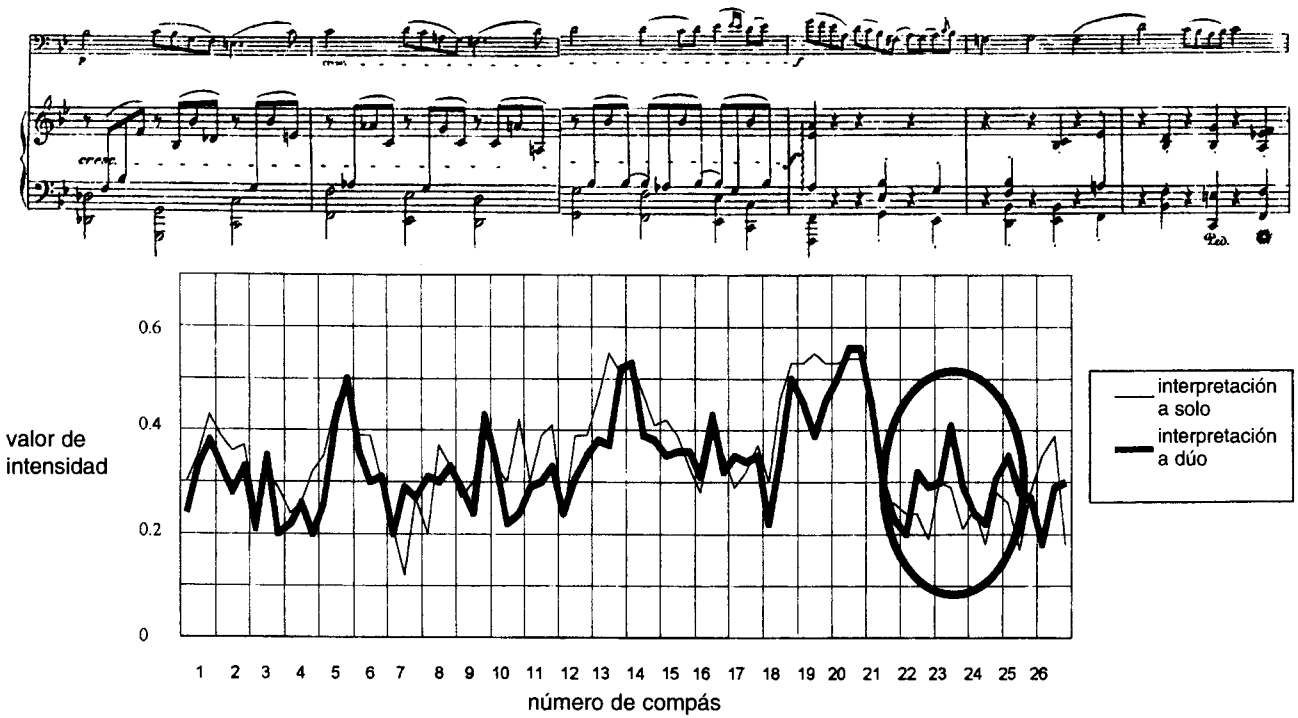



FIGURA 11.3 *Perfiles de las fluctuaciones dinámicas de las interpretaciones a solo y a dúo de la Sonata para violonchelo op. 65 de Chopin, tercer movimiento, compases 1-26; se anexa un fragmento de la partitura (compases 18-23 de la primera edición alemana; Leipzig, Breitkopf & Härtel, 1846) para facilitar las referencias que se hacen en el análisis a los compases agrupados dentro del círculo.*

es reemplazado por acordes en tiempos débiles por debajo de una declamación en estilo recitativo del violonchelo (la indicación dinámica es *forte* en ambas partes). Curiosamente, en su interpretación a solo, la pianista respondió al cambio de textura tocando relativamente suave en los compases 21-23, pasando por alto, aparentemente, la indicación *forte* de la partitura. En cambio, su dinámica en la interpretación a dúo aumentó en los compases 22-23. En dichos compases cambió rápidamente su enfoque anterior produciendo gradualmente un sonido más fuerte, dándole así un mayor apoyo al violonchelista (esto es un ejemplo de la importancia de la comunicación auditiva, ya que la pianista respondió indudablemente a las señales dinámicas del violonchelista). Curiosamente, la pianista comentó posteriormente en un ensayo que su parte estaba marcada *forte*.

La realización individual de la música puede verse afectada de diferentes maneras por la interacción con otros intérpretes: las características expresivas pueden moderarse o intensificarse (i.e. exagerarse) o simplemente mantenerse en un mismo nivel. Al mismo tiempo, cada intérprete puede mantener los patrones de expresión. En el ejemplo anterior, los aspectos expresivos de la interpretación a solo de la pianista fueron moderados en el contexto de la interpretación conjunta, y aunque algunas fluctuaciones de tempo o dinámica cambiaron, otras se mantuvieron iguales. Esto sugiere que los músicos no necesariamente renuncian a sus propias ideas cuando tocan en un grupo: el papel del individuo es el de negociar, o dar y recibir [21].

### *Factores sociales: permanecer juntos*

> Una agrupación musical es [...] un tipo inusual de grupo social cuyas interacciones implican un nivel de intimidad y sutileza que posiblemente no sea igualado por ningún otro tipo de grupo [22].

La interpretación en grupo es un trabajo colectivo: la mitad de la batalla de hacer música juntos (y de permanecer juntos como grupo) se libra en el terreno social. El asunto más importante es el liderazgo, ya que todo grupo necesita al menos un líder. Algunos son designados (como los directores de orquesta), mientras que otros surgen sin haber sido elegidos como tales (por ejemplo, un miembro dominante de un cuarteto de cuerda). Los factores que influyen en la elección del líder de un grupo incluyen las personalidades de los intérpretes, la música y los estereotipos sociales.

Una persona dominante puede guiar a un grupo dirigiendo el curso de un ensayo y dictando los aspectos interpretativos de una actuación. La manera en que está compuesta la música también puede repercutir en la relación

entre los intérpretes [23] (de modo que el compositor ejerce cierta influencia en las interacciones entre los músicos). Además, los intérpretes que son conscientes de los estereotipos sociales asociados a determinados grupos de instrumentistas (como el papel supuestamente inferior del segundo violín en un cuarteto, o del subordinado pianista «acompañante») pueden ser influidos por ellos, consciente o inconscientemente.

El estilo de liderazgo generado en un grupo determinará su ambiente social. Por una parte, los miembros de un cuarteto de cuerda pueden intentar trabajar juntos democráticamente, y así probablemente asegurarán que «los miembros del grupo realizarán gustosamente una parte razonable del trabajo, no sólo durante los ensayos y conciertos (cuando el líder está presente), sino también cuando practican sus partes en casa» [24]. Por otra parte, el director de orquesta puede esforzarse por lograr un enfoque menos dictatorial delegando gradualmente el control «por el bien del esfuerzo colectivo» [25].

Aunque el papel del líder es vital al determinar la dinámica de grupo dentro de un conjunto, la interacción entre los cointérpretes también es importante. Huelga decir que una retroalimentación positiva entre los músicos inspirará un elevado nivel de confianza y ayudará al grupo a tocar bien juntos. La interacción social entre los intérpretes es como un «proceso de desarrollo constante» [26], que puede realizarse de la manera más eficaz durante los ensayos, cuando los músicos negocian sus ideas, tratan los conflictos e intentan llegar a un acuerdo.

Un grupo florecerá mientras «cada individuo sienta que está contribuyendo a su máxima capacidad artística y al mismo tiempo colaborando con sus colegas para producir algo más hermoso de lo que se puede producir individualmente» [27]. En los deportes de equipo se hace hincapié en la «sinergia» (el potencial del grupo, que es mayor que la suma de los potenciales individuales) y la «confluencia» (la sensación de pertenecer a un equipo). Los músicos también pueden intentar generar el mismo espíritu cuando tocan juntos [28].

Por último, ¿por qué algunos grupos permanecen juntos durante más tiempo que otros? Aparte de las razones prácticas (como que un miembro del grupo se mude a otro sitio), los grupos exitosos intentan reconocer la influencia de los estereotipos sociales y las paradojas del liderazgo y la democracia, así como la necesidad de confrontación y compromiso [29]. Aunque es imposible predecir si los miembros de un grupo serán compatibles, y por tanto tampoco la perdurabilidad de éste, está claro que las relaciones entre ellos se desarrollarán y cambiarán continuamente. Después de todo, los intérpretes de cámara profesionales no sólo hacen música juntos, también pasan mucho tiempo juntos, por ejemplo, viajando. Las relaciones sociales y musicales se deben cultivar, y sin duda las incompatibilidades de carácter pueden surgir con el tiempo.

## *Conclusión*

En resumen, en este capítulo hemos analizado cuatro aspectos de la interpretación en grupo: la coordinación (mantener el tempo), la comunicación (auditiva y visual), el papel del individuo y los factores sociales. Ahora podemos percatarnos de las exigencias que plantea la interpretación en grupo y apreciar la verdadera proeza que implica dicha actividad. En efecto, el intérprete de un grupo exhibe diversas habilidades musicales y sociales. Para lograr la coordinación del tempo los cointérpretes anticipan las acciones de los demás y responden a ellas siguiendo a un colega o cooperando con él. Al mismo tiempo, comunican sus ideas expresivas mediante señales auditivas o visuales, de modo que los intérpretes se observan y escuchan unos a otros mientras transmiten y evalúan constantemente las señales musicales. El intérprete de grupo exhibe tendencias individuales, a solo, en la interpretación, y al mismo tiempo intenta fusionarse con el resto del grupo. Las relaciones sociales y musicales entre los cointérpretes están en constante desarrollo, de modo que cada grupo genera continuamente un nuevo espíritu, que tal vez sea la razón por la que la interpretación en grupo puede ser tan refrescante y excitante.

### *Notas*

1. Véase L. Henry Shaffer, «Timing in solo and duet piano performances», *Quarterly Journal of Experimental Psychology* 36A (1984), p. 592.
2. Citado por David Blum, *The Art of Quartet Playing: The Guarnieri Quartet in Conversation with David Blum,* Ithaca, Cornell University Press, 1986, p. 20.
3. Véase L. Henry Shaffer, «Timing in solo and duet piano performances», *cit.,* p. 587.
4. Lucy J. Appleton, W. Luke Windsor y Eric F. Clarke, «Cooperation in piano duet performance», en Alf Gabrielsson (ed.), *Proceedings of the Third Triennial European Society for the Cognitive Sciences of Music (ESCOM) Conference,* Uppsala, Uppsala University Press, 1997, pp. 471-474.
5. Véase Rudolf A. Rasch, «Timing and synchronisation in ensemble performance», en John A. Sloboda (ed.), *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition,* Oxford, Clarendon Press, 1988, pp. 70-90.
6. *Ibid.,* p. 71; véase también Jonathan Dunsby, *Performing Music: Shared Concerns,* Oxford, Clarendon Press, 1995, pp. 64-65.
7. Véase Rudolf A. Rasch, «Timing and synchronisation in ensemble performance», *cit.,* pp. 80-81.
8. Véanse Anthony Clayton, «Coordination between players in musical performance», tesis doctoral, Universidad de Edimburgo, 1985, y Elaine C. Goodman, «Analysing the ensemble in music rehearsal and performance: the nature and effects of interaction in cello-piano duos», tesis doctoral, Universidad de Londres, 2000.
9. Según Blum (*The Art of Quartet Playing: The Guarnieri Quartet in Conversation with David Blum,* cit., p. 16), «un exceso de *rubato* planificado puede coartar la verdadera libertad»; véase también el capítulo 9 de este libro.

10. Caroline Palmer, «On the assignment of structure in music performance», *Music Perception* 14 (1996), p. 54.
11. Abram Loft, *Ensemble! A Rehearsal Guide to Thirty Great Works of Chamber Music,* Portland, Ore., Amadeus Press, 1992, p. 17.
12. William Pleeth, *Cello,* compilado y editado por Nona Pyron, Londres, Kahn and Averill [1982], 1992, pp. 113 y 119.
13. Jane W. Davidson, «The social in musical performance», en David J. Hargreaves y Adrian C. North (eds.), *The Social Psychology of Music,* Oxford, Oxford University Press, 1997, p. 222.
14. John Dalley, citado por David Blum, *The Art of Quartet Playing: The Guarnieri Quartet in Conversation with David Blum,* cit., p. 14.
15. Lucy J. Appleton, W. Luke Windsor y Eric F. Clarke, «Cooperation in piano duet performance», *cit.,* p. 474; véase también Anthony Clayton, «Coordination between players in musical performance», cit., pp. 72-108.
16. Mitch Waterman, «Emotional responses to music: implicit and explicit effects in listeners and performers», *Psychology of Music* 24 (1996), p. 56.
17. *Ibid.*, p. 65; en cambio, el estudio reciente de Goodman sobre los ensayos de grupo reveló que a menudo los intérpretes intercambian opiniones sobre las ideas emocionales como una manera de reconciliar los puntos de vista conflictivos sobre la interpretación de la música (véase Elaine C. Goodman, «Analysing the ensemble in music rehearsal and performance: the nature and effects of interaction in cello-piano duos», cit., pp. 105-152).
18. Abram Loft, *Ensemble! A Rehearsal Guide to Thirty Great Works of Chamber Music,* cit., p. 18.
19. El «factor tempo» de la Figura 11.2 expresa la diferencia entre el valor de las notas ejecutadas por el intérprete en tiempo real y su valor metronómico «normal» (de modo que cuanto mayor es el factor tempo, mayor es la velocidad real de la interpretación). La pianista tocó en un Yamaha Disklavier, lo que permitió recuperar sus interpretaciones en formato MIDI. Los datos MIDI fueron subsiguientemente analizados utilizando programas informáticos avanzados como POCO.
20. El «valor de intensidad» en la Figura 11.3 refleja la dinámica o volumen de sonido de una nota en una escala arbitraria de 0 a 1, partiendo de un cero real. A mayor valor, más fuerte es la dinámica.
21. Para un análisis más amplio, véase Elaine C. Goodman, «Analysing the ensemble in music rehearsal and performance: the nature and effects of interaction in cello-piano duos», cit., pp. 165-225.
22. Vivienne M. Young y Andrew M. Colman, «Some psychological processes in string quartets», *Psychology of Music* 7 (1979), pp. 12-13.
23. Como afirma Loft (*Ensemble! A Rehearsal Guide to Thirty Great Works of Chamber Music,* cit., p. 18), «la prominente voz superior imprime un sello inevitable en el temperamento del grupo».
24. Vivienne M. Young y Andrew M. Colman, «Some psychological processes in string quartets», *cit.,* p. 16.
25. Yaakov Atik, «The conductor and the orchestra: interactive aspects of the leadership process», *Leadership and Organisation Development Journal* 13 (1994), p. 27.
26. John Dalley, citado por David Blum, *The Art of Quartet Playing: The Guarnieri Quartet in Conversation with David Blum,* cit., p. 7.
27. John Harvey-Jones, *All Together Now,* Londres, Heinemann, 1994, p. 137.
28. Véase John Syer, «Team building: the development of team spirit», en Stephen J. Bull (ed.), *Sports Psychology: A Self-Help Guide,* Marlborough, Crowood Press, 1991, pp. 123-143.

29. Véase J. K. Murningham y D. E. Conlon, «The dynamics of intense work groups: a study of British string quartets», *Administrative Science Quarterly* 36 (1991), pp. 165-186.

## *Lecturas complementarias*

Appleton, Lucy J., W. Luke Windsor y Eric F. Clarke: «Cooperation in piano duet performance», en Alf Gabrielsson (ed.), *Proceedings of the Third Triennial European Society for the Cognitive Sciences of Music (ESCOM) Conference,* Uppsala, Uppsala University Press, 1997, pp. 471-474.

Atik, Yaakov: «The conductor and the orchestra: interactive aspects of the leadership process», *Leadership and Organisation Development Journal* 13 (1994), pp. 22-28.

Blum, David: *The Art of Quartet Playing: The Guarnieri Quartet in Conversation with David Blum,* Ithaca, Cornell University Press, 1986.

Rasch, Rudolf A.: «Timing and synchronisation in ensemble performance», en John A. Sloboda (ed.), *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition,* Oxford, Clarendon Press, 1988, pp. 70-90.

Shaffer, L. Henry: «Timing in solo and duet piano performances», *Quarterly Journal of Experimental Psychology* 36A (1984), pp. 577-595.

# 12 *El miedo escénico*

ELIZABETH VALENTINE

## *El análisis del miedo escénico*

La ansiedad de actuación, conocida generalmente como «miedo escénico», es un problema muy antiguo, pero el interés por su naturaleza, sus causas y sus tratamientos se ha intensificado durante los últimos quince años, en los que han surgido un número creciente de clínicas, conferencias y revistas dedicadas a la medicina de las artes escénicas. La ansiedad de actuación musical se ha definido como «la persistente y angustiosa aprensión en torno a las habilidades interpretativas, y/o su deterioro real, en un contexto público, hasta un nivel injustificado dadas la aptitud, adiestramiento y nivel de preparación musicales del individuo» [1]. Entre los músicos famosos que lo han padecido se encuentran Maria Callas, Enrico Caruso, Pau Casals, Leopold Godowski, Vladimir Horowitz, Ignacy Paderewski y Sergei Rachmaninoff. Varias encuestas han revelado que el miedo escénico constituye un problema grave para un porcentaje sustancial de músicos. En un estudio a gran escala sobre orquestas estadounidenses [2], el miedo escénico fue el problema interpretativo que se refirió con mayor frecuencia, el 24 por ciento de los músicos afirmó que lo padecía, y el 15 por ciento consideró que era un problema grave. El miedo escénico disminuye con la edad y la experiencia [3], aunque no se sabe si es por los efectos beneficiosos de las frecuentes presentaciones en público o por la «supervivencia del más apto».

Los síntomas del miedo escénico son conocidos, y se clasifican en tres tipos: fisiológicos, conductuales y psíquicos. Los *síntomas fisiológicos* son el resultado de una sobreexcitación del sistema nervioso autónomo e incluyen un

aumento del ritmo cardíaco, palpitaciones, falta de aliento, hiperventilación, sequedad de la boca, sudoración, náuseas, diarrea y mareos. Esta reacción de «huir o pelear», que ayudaba a nuestros antepasados cazadores a huir de las fieras, es sumamente perjudicial para los músicos, que necesitan destreza y un control muscular minucioso. Los temblores en las extremidades y las manos sudorosas no ayudan al intérprete. Además, esa excitación autónoma puede estar asociada al temor como resultado de una experiencia pasada. El aumento de la excitación generalmente conduce a una disminución de la atención, que también puede ser perniciosa. Los *síntomas conductuales* del miedo escénico pueden manifestarse como síntomas de ansiedad, como temblores, rigidez muscular o inexpresividad, o deteriorar la interpretación misma. Los *síntomas psíquicos* son una sensación subjetiva de ansiedad y pensamientos negativos sobre la actuación. Más que miedo a la interpretación *per se*, es un miedo a la interpretación en público, con el riesgo de una evaluación negativa y la consecuente pérdida de autoestima. Ese miedo puede ser producto de una identificación demasiado rigurosa de la autoestima con la perfección interpretativa, incluso la creencia de que la autoestima depende del éxito. Los pensamientos negativos influyen perjudicialmente en la calidad de la interpretación. La preocupación conduce a una falta de concentración, lo que dispersa la atención y provoca un desperdicio de recursos valiosos, y probablemente también actúa como una señal que incrementa aún más la ansiedad.

Una variante del pensamiento negativo es el «catastrofismo», la exageración irracional de la posibilidad de un fracaso —por ejemplo, «estoy casi seguro de que voy a cometer un error terrible que arruinará todo». Esta variante tiende a estar relacionada con niveles elevados de ansiedad de actuación en profesionales, estudiantes y aficionados, mientras que la valoración realista —reconocer la inevitabilidad de los errores, la naturaleza generalmente tolerante del público y la necesidad de concentrarse en los aspectos técnicos e interpretativos de la música— está relacionada con niveles moderados de miedo escénico [4].

Otra variante es la estrategia del «autoimpedimento» [5], que consiste en que el intérprete promueve determinadas situaciones que servirán para excusar una mala interpretación, como, por ejemplo, no practicar, trasnochar la noche anterior, dañar el material de trabajo, etc. Estos actos pueden convertirse en profecías que se autorrealizan e incluyen el riesgo de sabotear realmente la interpretación. Su propósito es el de atribuir la mala interpretación a factores externos y reivindicar un mérito adicional en el caso de que se produzca una interpretación buena, que habrá ocurrido a pesar de las circunstancias adversas. Sin embargo, estas estrategias son esencialmente destructivas, ya que evaden la situación, lo cual suele empeorar las cosas. Pueden

surgir cuando existe una preocupación exagerada por la capacidad, quizá como condicionante del amor de los padres, o cuando el intérprete percibe que su capacidad es frágil, tal vez porque no puede mantener un nivel de interpretación alcanzado anteriormente.

Una descripción exhaustiva del miedo escénico ha de incluir, por lo tanto, factores fisiológicos (como el pulso y la tensión arterial), medidas de la conducta (ansiedad, calidad de la interpretación) y autoinformes (de los pensamientos y sentimientos). Las reacciones de estos tres sistemas no están necesariamente correlacionadas [6]. En particular, es común tener solamente los síntomas fisiológicos. La correlación es más probable en los estados de elevada ansiedad. Craske y Craig [7] descubrieron que los pianistas relativamente ansiosos mostraron ansiedad en los tres sistemas al interpretar en una situación de tensión ante un público, mientras que los pianistas relativamente no ansiosos sólo mostraron los síntomas fisiológicos. Ello también plantea la pregunta de si el tratamiento debe ajustarse a los síntomas, por ejemplo, la prescripción de fármacos para los síntomas fisiológicos y de tratamiento psicológico para los síntomas psíquicos. Las pruebas no son concluyentes en este sentido.

Varios estudios de encuestas han identificado diferentes componentes de la ansiedad de actuación musical, algunos de los cuales pueden ser positivos. Una encuesta realizada a unos 200 músicos, principalmente profesionales, identificó cuatro factores: nerviosismo/aprensión, seguridad/capacidad, cohibición/distracción, y un factor excitación/intensidad [8]. Un grupo cuantitativamente similar pero más diverso distinguió cinco componentes independientes: preocupación sobre la ansiedad y sus efectos; miedo a la evaluación, es decir, preocupación por las reacciones de los demás ante la interpretación; actitudes críticas en relación a las habilidades interpretativas; distracción y problemas de memoria, y el desarrollo deliberado de técnicas para hacer frente a la ansiedad. De todos ellos, el primero es el que está más estrechamente relacionado con la ansiedad que desmejora la actuación, pero todos excepto el último tienen consecuencias negativas en ella [9].

Es importante distinguir las ansiedades beneficiosas de las perjudiciales —o, para ser más precisos, entre las ansiedades reactiva, maladaptativa y adaptativa. La ansiedad reactiva, que resulta de una preparación inadecuada, es realista, y la mejor manera de tratarla es con el análisis musical y el ensayo. La ansiedad es considerada generalmente como perjudicial, pero todo intérprete sabe que cierto nivel de excitación es beneficioso para la actuación. Generalmente, la mejor ejecución es la que se realiza con un nivel moderado de excitación: si el nivel es demasiado bajo, la interpretación será floja y aburrida; si el demasiado elevado, el intérprete y la interpretación se pueden desmoronar por completo. Esto se puede representar como una U invertida (trazando la calidad de la interpretación como una función de excitación), y

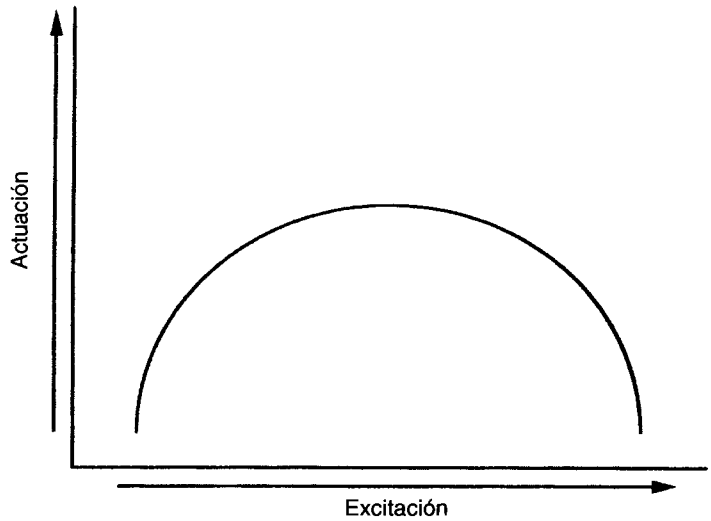



FIGURA 12.1 *La ley de Yerkes-Dodson, que muestra la relación hipotética entre la excitación y la calidad de la actuación.*

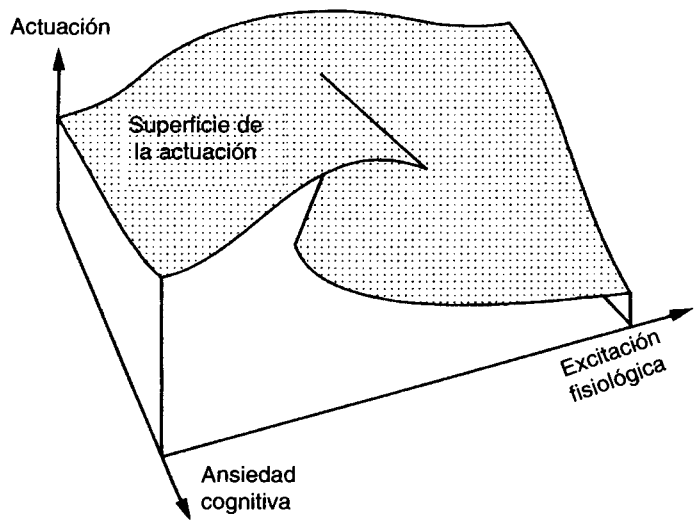



FIGURA 12.2 *Modelo de catástrofe de la relación entre la ansiedad y la actuación (© 1996 John Wiley & Sons Limited; reproducido con autorización de Lew Hardy, Graham Jones y Daniel Gould,* Understanding Psychological Preparation for Sport: Theory and Practice of Elite Performers, *Chichester, John Wiley & Sons Limited, 1996, p. 151).*

se conoce como la ley de Yerkes-Dodson (véase la Figura 12.1) [10]. Steptoe [11] confirmó la aplicación de este patrón a los alumnos de canto y cantantes profesionales, a los que pidió que evaluaran su tensión emocional y la calidad de su interpretación en diferentes situaciones. Ambos grupos consideraron que la calidad de la interpretación alcanzaba su punto máximo en el nivel intermedio de tensión emocional, después de lo cual descendía verticalmente. Para los alumnos, el contexto ideal era una clase; para los profesionales (afortunadamente), era una interpretación pública. Fazey y Hardy [12] sostienen que es necesario distinguir el componente psíquico o cognitivo de la ansiedad (el miedo al fracaso y sus consecuencias) del componente corporal o somático (las respuestas fisiológicas ante el estrés). Para ello se necesita un modelo tridimensional de la relación entre la ansiedad y la actuación (véase la Figura 12.2). Prosiguen señalando que cuando la ansiedad cognitiva es baja, la relación entre la tensión y la actuación sigue la ley de Yerkes-Dodson (representada en la parte posterior de la Figura 12.2), pero cuando es alta sigue el modelo de catástrofe: cuando aumenta el estrés, la actuación es propensa a un declive catastrófico, del cual es difícil que se recupere. La razón es que las preocupaciones y cavilaciones conducen a una espiral viciosa de pensamientos negativos, haciendo que el intérprete se «sobrepase» y se venga abajo.

Hardy y Parfitt [13] han proporcionado pruebas de este modelo en estudios sobre atletas. Wilson [14] ha sugerido que puede aplicarse igualmente a la interpretación musical.

### *Causas*

Existen tres factores que contribuyen a la ansiedad de actuación: la persona, la tarea y la situación. Wilson los ha representado en una ampliación tridimensional del modelo de Yerkes-Dodson (véase la Figura 12.3) [15]. Un alto nivel de uno de ellos puede ser compensado con un bajo nivel de otro. Por tanto, a una persona ansiosa por naturaleza le convendría elegir una pieza relativamente fácil o bien ensayada para una ocasión exigente como una audición; a la inversa, los intérpretes experimentados pueden necesitar situaciones más desafiantes que los novatos para dar lo mejor de sí [16].

Las personas difieren considerablemente en cuanto a las reacciones de sus sistemas autonómicos y su sensibilidad tanto a los estímulos internos como a las críticas de los demás. Existen pruebas de que los músicos son más ansiosos que la población general [17] y que los músicos de orquesta pueden ser más susceptibles a la ansiedad que otros artistas escénicos (como cantantes, bailarines y actores) [18]. La ansiedad de actuación musical está relacionada con otras formas de ansiedad, particularmente el neuroticismo y la fobia social [19].

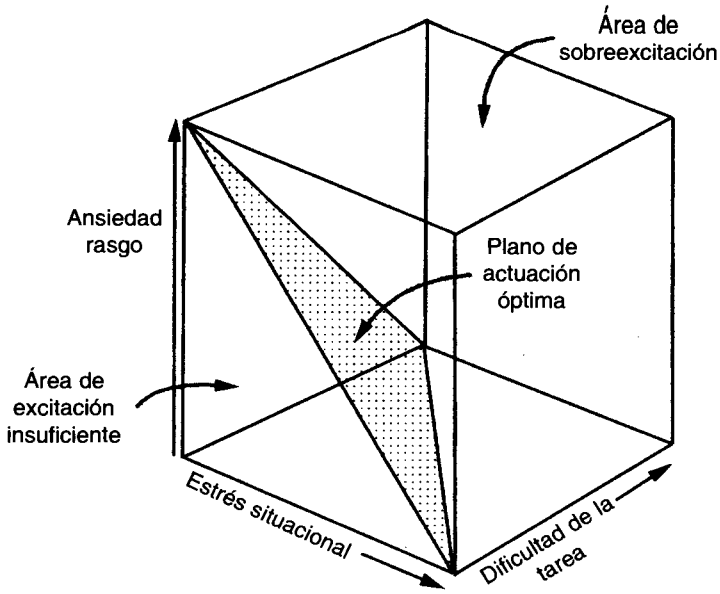



FIGURA 12.3 *Ampliación de Wilson (1994) del modelo de Yerkes-Dodson, ilustrando tres factores que contribuyen a la ansiedad de actuación (reproducido con autorización de Glenn Wilson y Jessica Kingsley; todos los derechos reservados).*

Cuanto más difícil sea la tarea, más ansiedad provocará al intérprete. A medida que aumentan los niveles de habilidad y dominio de la tarea, será necesaria una tarea correspondientemente más difícil para provocar el mismo nivel de ansiedad, como he señalado antes.

Varios estudios han demostrado la influencia del contexto. LeBlanc et al. [20] encontraron que la ansiedad referida por los músicos de banda de institutos de secundaria aumentaba significativamente en el transcurso de tres situaciones interpretativas: tocando solo en una sala de prácticas, tocando en una sala de prácticas en presencia de un investigador y en una sala de prácticas con cuatro investigadores y de quince a diecinueve colegas.

Leglar [21] descubrió que tocar de memoria ante un gran público generalmente produce el efecto más debilitante de la actuación y la máxima respuesta autonómica en los organistas. Abel y Larkin [22] observaron aumentos del pulso, la tensión arterial y la ansiedad autorreferida en un grupo de estudiantes de música que tocó ante un jurado. Curiosamente, los varones mostraron aumentos mayores en la tensión arterial, mientras que las mujeres mostraron

aumentos mayores en la sensación de ansiedad, revelando que existen diferencias de género y que las diferentes manifestaciones de la ansiedad a menudo no se corresponden. Brotons [23] encontró que era indiferente que las audiciones se realizaran abiertamente, es decir, que los intérpretes y el jurado pudieran verse los unos a los otros (como en la mayoría de las audiciones para los cursos de interpretación de las escuelas superiores), o que el intérprete se colocara detrás de una pantalla, de modo que tanto él como el jurado desconocieran sus respectivas identidades (como suelen realizarse las audiciones para un empleo).

Los tres factores que hemos mencionado —la persona, la tarea y la situación— interactúan entre sí: el efecto de uno de ellos depende del nivel de los otros. Un estudio realizado por Cox y Kenardy [24] demuestra la interacción entre las características de la personalidad y la situación. Los intérpretes que tienen fobias sociales se mostraron mucho más ansiosos en un contexto solista que los que no tienen esas fobias, mientras que había menos diferencia entre ellos en las interpretaciones en grupo o los ensayos. La interacción entre la tarea y la situación puede ilustrarse haciendo referencia a un estudio realizado por Hamman y Sobaje [25]. La mejoría de actuación que se observó en una condición de ansiedad elevada (en la que un jurado evaluaba la interpretación) con respecto a una condición de ansiedad reducida (en la que el intérprete se encontraba solo en una habitación con un magnetófono) fue mayor en los intérpretes más experimentados, es decir, aquellos con más años de estudio formal y un nivel de habilidad elevado. La ansiedad facilitó más la actuación para aquellos que poseían un mayor dominio de la tarea.

### *Tratamientos*

Brodsky y Sloboda [26] enumeran los siguientes métodos para afrontar el miedo escénico: las técnicas de Alexander y Feldenkreis, los ejercicios aeróbicos, el entrenamiento en el manejo de la ansiedad, la concentración, el entrenamiento autogénico, la desensibilización sistemática, desarrollar intereses y aficiones fuera de la música, exponerse a situaciones de interpretación, el ensayo mental, la retroalimentación de la tensión muscular y de la temperatura de los dedos, la terapia nutricional, las autoafirmaciones positivas, el rezo, el entrenamiento de la relajación, la autohipnosis, la terapia de inoculación del estrés, el ensayo sistemático y el yoga. Esta lista ofrece una amplia gama de posibilidades, por no decir más. Hanser [27] establece convenientemente una diferencia entre las técnicas «instrumentales», cuyo objetivo es cambiar el entorno del yo, y las medidas paliativas dirigidas a regular las emociones y respuestas ante el estrés.

El 48 por ciento de una muestra de músicos profesionales y aficionados, a quienes se les pidió que describieran métodos eficaces para hacer frente a la ansiedad de actuación, respondió que utilizaba al menos uno. De los 478 métodos sugeridos, dos tercios se centraban en las emociones, por ejemplo: «Intento sentarme tranquilamente y rezo para calmarme antes de cantar»; otro tercio apuntaba a los problemas, por ejemplo, «estar bien preparado —que no haya ninguna dificultad técnica en la música que no esté dominada». Los músicos que utilizaban predominantemente los métodos centrados en las emociones manifestaron sentir más seguridad y capacidad y menos timidez y distracción que los que empleaban predominantemente los métodos enfocados en los problemas. Por otra parte, el grupo que se centraba en los problemas recordó haber tenido menos pensamientos perturbadores durante su interpretación más reciente [28]. Esto confirma la idea de que diferentes métodos pueden tratar distintos aspectos de la ansiedad de actuación.

La siguiente evaluación de la eficacia de los tratamientos para la ansiedad de actuación comenzará con los más físicos y procederá hacia los más psicológicos. Naturalmente, muchas técnicas combinan elementos de ambos enfoques [29].

### *Técnicas físicas*

La *relajación,* inducida mediante la respiración profunda o mediante la técnica de relajación progresiva de Jacobson que tensa y relaja alternadamente los músculos, es la técnica más utilizada [30]. Un estudio demostró que con seis sesiones de entrenamiento de la relajación progresiva se puede reducir el ritmo cardíaco y la ansiedad autorreferida de los instrumentistas ansiosos [31].

El *biofeedback* (retroalimentación) se utiliza a veces para proporcionar información sobre la tensión y para ayudar a entrenar la relajación. Un grupo de violinistas fue entrenado para reducir la tensión excesiva de la mano izquierda, colocándoles electrodos en los músculos de la mano que controlan los movimientos del pulgar. Los electrodos estaban conectados a una máquina que emitía un sonido cuando la tensión superaba un nivel determinado. Dicho nivel fue disminuido sistemáticamente a medida que los intérpretes aprendían a relajarse [32]. Se han utilizado técnicas similares para reducir la tensión en los antebrazos de violinistas y clarinetistas, los músculos laríngeos de cantantes y los músculos faciales de ejecutantes de instrumentos de viento. Una desventaja del *biofeedback* es que los intérpretes no pueden llevar el aparato consigo en los conciertos.

El 22 por ciento de una muestra de músicos de orquesta londinenses afirmó que tomaba *alcohol,* y el 12 por ciento *sedantes,* para hacer frente al mie-

do escénico. Es probable que estas cifras subestimen el consumo real, aunque otros estudios [33] han reportado un consumo mínimo. El problema de estos sedantes generales es que disminuyen las funciones y el juicio, por lo que el ejecutante puede creer que su interpretación es mejor de lo que realmente es; además, pueden tener efectos secundarios negativos, creando adicción si se toman habitualmente en dosis elevadas y produciendo un deterioro general de la salud. Los sedantes pueden parecer beneficiosos a corto plazo, pero impiden el desarrollo de métodos más adecuados para combatir la ansiedad a largo plazo [34].

Sin embargo, los *betabloqueadores* (bloqueadores de los betaadrenoceptores), que actúan sobre el sistema nervioso periférico sin producir efectos pronunciados en el sistema nervioso central, han despertado mucho interés. Una encuesta estadounidense a gran escala reveló que más del 20 por ciento de los ejecutantes profesionales toman estos fármacos, en muchos casos sin supervisión médica [35]. Muchos músicos han recurrido al betabloqueador conocido como Inderal (propanolol). Una encuesta europea indicó que el 15 por ciento de los músicos de orquesta lo utilizan ocasionalmente. Varios estudios cuidadosamente controlados [36] han demostrado la eficacia de los beta-bloqueadores para aliviar los síntomas fisiológicos (aumento del pulso y la tensión arterial), los síntomas conductuales (temblores) y la ansiedad autorreferida, además de mejorar la calidad de la actuación. El beneficio más notable que se ha observado es la reducción de temblores en ejecutantes de cuerdas, lo que mejora la uniformidad del vibrato y los trinos rápidos [37]. Sin embargo, también se ha reportado un deterioro del control rítmico y una dinámica más monótona [38], sensación de inseguridad, fallos en la memoria y falta de concentración [39]. Es importante recalcar que estos medicamentos pueden producir efectos secundarios perjudiciales y por tanto sólo han de tomarse bajo supervisión médica. La opinión general es que sólo son adecuados para ser utilizados ocasionalmente.

### *Técnicas físicas y mentales*

Muchas de las técnicas que se pueden emplear para combatir la ansiedad de actuación *combinan la relajación física con la agudeza mental,* especialmente las que provienen de Oriente, como el yoga y el tai chi, o la concentración sin esfuerzo del «Inner Game» [40]. Una terapia que incluía sugestiones hipnóticas para la relajación, control de la respiración, imágenes visuales y sugestiones verbales que relacionaban las imágenes con un mayor control mental disminuyó la ansiedad de actuación en estudiantes de conservatorio con miedo escénico inmediatamente después del tratamiento y en un tratamiento

complementario después de seis meses, mientras que las sesiones de conversación no dieron resultado [41].

La *técnica de Alexander* es empleada por muchos músicos y otros profesionales para mantener un adecuado equilibrio entre distintas partes del cuerpo y reducir las tensiones innecesarias. Un estudio [42] demostró que era tan eficaz como los betabloqueadores, y preferible a éstos, para reducir la tensión arterial de los músicos de orquesta antes de la interpretación. En otro estudio [43], los estudiantes de música que habían seguido un curso de la técnica de Alexander mostraron progresos, en comparación con aquellos que no lo habían seguido, en cuanto a la calidad musical y técnica en general, las variaciones del ritmo cardíaco, la ansiedad autorreferida y la actitud positiva hacia la interpretación. Sin embargo, esos efectos positivos estaban limitados al contexto de un salón de clases, relativamente menos estresante, y no se extendieron a la situación de estrés elevado de un recital. Dichos efectos tampoco se relacionaron con progresos en cuanto a la «utilización» del cuerpo según expertos en la técnica de Alexander; por tanto, su ubicación sigue siendo dudosa. Sin embargo, todos los estudiantes que tomaron el curso manifestaron haber mejorado su conciencia de la tensión y su habilidad para relajarse, y la mayoría afirmó que lo recomendaría a otros.

### *Técnicas psicológicas*

La terapia de la conducta está basada en principios tomados de la psicología del aprendizaje. En la *desensibilización sistemática* se le enseña al individuo a mantener un estado calmado y relajado durante una exposición progresiva a un estímulo que produce temor, puesto que la relajación es incompatible con el temor. Muchos estudios han demostrado, mediante mediciones fisiológicas, de actuación y autorreferencia, que es un tratamiento eficaz para la ansiedad de actuación. En uno de ellos [44], la desensibilización sistemática combinada con prácticas en situaciones «en vivo» resultó más eficaz para reducir la ansiedad subjetiva que el análisis musical o la ausencia de entrenamiento; también redujo el ritmo cardíaco. En otro estudio [45], la desensibilización sistemática redujo en mayor medida los errores técnicos y respuestas verbales ansiosas que el entrenamiento de la improvisación y las habilidades cognitivas o la ausencia de tratamiento.

El pensamiento positivo es un método comprobado que se recomienda con frecuencia [46]. Las investigaciones han demostrado que su equivalente moderno, la *terapia cognitiva de la conducta,* es la técnica más eficaz para tratar la ansiedad de actuación y la menos utilizada por los músicos, que tienden a preferir métodos más físicos, como la relajación y los fármacos.

Resulta terapéuticamente beneficioso aceptar la ansiedad anticipatoria como un elemento natural de la actuación y utilizar la consiguiente tensión para activar nuestra preparación. La *terapia de inoculación del estrés*[47] se basa en la idea de aceptar y reevaluar los síntomas de la ansiedad, convirtiéndolos en valores positivos. El procedimiento consiste en imaginar una situación que produce ansiedad, seguida de una actividad preparatoria o una conducta eficaz para afrontarla.

También se puede lograr una mejoría con un adecuado *timing de la ansiedad* y preocupándose con antelación. Los estudios sobre paracaidistas han demostrado que los más experimentados controlan su ansiedad para que ocurra antes y no durante la actividad. Lo mismo se ha demostrado con un grupo de estudiantes de música a quienes se les pidió que clasificaran su ansiedad en varios momentos antes de una interpretación ante un jurado. Los intérpretes más experimentados mostraron una ansiedad máxima antes de la interpretación, mientras que los menos expertos alcanzaron el máximo nivel de ansiedad durante la interpretación[48].

En cuanto a la eficacia de algunas de estas técnicas psicológicas, Sweeney y Horan[49] encontraron que entrenando un control de la relajación mediante una palabra como «calma» o del análisis de las manifestaciones contraproducentes y su sustitución por otras más positivas —por ejemplo, desplazar el centro de atención de la ansiedad a la tarea— se redujo el ritmo cardíaco y la ansiedad autorreferida de un grupo de pianistas estudiantes de licenciatura en música. La primera técnica también mejoró la capacidad interpretativa, mientras que la segunda redujo las manifestaciones conductuales de la ansiedad, como por ejemplo los temblores. Un grupo de comparación que recibió adiestramiento en análisis musical no mostró ventajas frente a un grupo de «lista de espera» —es decir, pianistas que esperaban a recibir el entrenamiento.

En un estudio sobre pianistas con ansiedad de actuación debilitante, Kendrick et al.[50] compararon la eficacia relativa de la terapia cognitiva de la conducta (que incluía la autoinstrucción, la concentración y el reemplazo de los pensamientos negativos ajenos a la tarea con un autohabla positiva orientado a ésta) y el ensayo conductual (practicar la interpretación ante un público amable y comprensivo). Inmediatamente después del tratamiento no se observaron diferencias entre los tres grupos del estudio, pero cinco semanas después, los dos grupos que recibieron tratamiento mostraron progresos en comparación con un grupo de «lista de espera» en casi todos los parámetros de ansiedad de actuación (pensamientos negativos, signos visibles de ansiedad, eficacia esperada y calidad real de la ejecución). Sin embargo, ninguno de los tratamientos redujo la *sensación* de ansiedad. La terapia cognitiva de la conducta fue más eficaz que el ensayo conductual para reducir los signos vi-

sibles de ansiedad y aumentar la autoeficacia. Además, las mediciones del pensamiento positivo y la autoeficacia inmediatamente después del tratamiento predijeron una mejoría de los signos visibles de ansiedad y/o de la calidad de la ejecución en el seguimiento de la investigación.

### *Tratamientos combinados*

Frecuentemente se emplean tratamientos combinados que reúnen diversas técnicas. Dos ejemplos deberán bastar en este caso. Un programa de seis semanas de terapia de grupo, que consistía en una combinación de relajación muscular progresiva, terapia cognitiva (identificar pensamientos negativos y sustituirlos por otros positivos al tiempo que se cuestionan los pensamientos irracionales) y entrenamiento del *biofeedback* de la temperatura, demostró ser eficaz para reducir la ansiedad de actuación autoevaluada en doce pianistas competentes aquejados de ansiedad de actuación debilitante, en comparación con un grupo de lista de espera [51]. Clark y Agras [52] tuvieron éxito con el siguiente tratamiento combinado: identificar las autoverbalizaciones negativas y el pensamiento distorsionado, hacer deberes para modificarlos, proporcionar modelos en diapositivas y cintas magnetofónicas para hacer frente a la ansiedad, entrenar la relajación y dar ánimos para encarar las situaciones de interpretación que se temen. En conjunto, estos tratamientos redujeron significativamente la ansiedad subjetiva y mejoraron la confianza y la calidad de la interpretación musical en intérpretes que padecían fobia social. En cambio, el tratamiento con Buspirona (un fármaco que reduce la ansiedad) no fue eficaz. Estos estudios demuestran claramente que existen tratamientos eficaces para la ansiedad de actuación. Lo que no demuestran es cuáles son exactamente. Son necesarias muchas investigaciones minuciosas para aislar los componentes fundamentales y determinar las combinaciones óptimas.

### *Recomendaciones*

En base a lo expuesto anteriormente, se pueden dar los siguientes consejos a los lectores que sufren de ansiedad de actuación. En primer lugar, distinga cuidadosamente los diferentes tipos de ansiedad —reactiva, maladaptativa y adaptativa. Considere si su ansiedad es realmente perjudicial para su interpretación. Recuerde que cierto nivel de ansiedad es normal e incluso beneficiosa, ya que convierte una interpretación monótona en una dinámica y excitante. Si su ansiedad es excesiva, considere si existen medidas prácticas que pueda tomar para reducirla (teniendo presente el modelo tridimensional que

examinamos al principio), como elegir una pieza más fácil o estudiarla más concienzudamente. Fármacos como los betabloqueadores pueden considerarse en circunstancias excepcionales, pero sólo como una medida temporal y bajo supervisión médica: producen efectos secundarios indeseables y pueden ser adictivos. Tener confianza en sí mismo ha de ser su objetivo. Si tiende a ser ansioso y a reaccionar exageradamente en situaciones estresantes, considere una o varias de las siguientes medidas: (a) técnicas físicas de diversos tipos que incluyen la relajación —por ejemplo, la autosugestión o la técnica de Alexander—; (b) técnicas conductuales, como la desensibilización sistemática y el ensayo conductual; (c) el pensamiento positivo y el «autohabla», centrándose en la tarea misma en lugar de evaluar la interpretación, y en el placer propio y del público en lugar de adoptar una actitud crítica. También se pueden hacer combinaciones, como imaginar interpretaciones exitosas. Si, a pesar de estas precauciones, los costes siguen siendo mayores que los beneficios, entonces tal vez la interpretación pública no sea lo suyo.

## *Notas*

1. Paul G. Salmon, «A psychological perspective on music performance anxiety: a review of the literature», *Medical Problems of Performing Artists* 5 (1990), p. 3.
2. Martin Fishbein, Susan E. Middelstadt, Victor Ottati, Susan Strauss y Alan Ellis, «Medical problems among ICSOM musicians: overview of a national survey», *Medical Problems of Performing Artists* 3 (1988), pp. 1-8.
3. Véase *ibid.*; Mary L. Wolfe, «Correlates of adaptive and maladaptive musical performance anxiety», *Medical Problems of Performing Artists* 4 (1989), pp. 49-56; y Andrew Steptoe y Helen Fidler, «Stage fright in musicians: a study of cognitive and behavioural strategies in performance anxiety», *British Journal of Psychology* 78 (1987), pp. 241-249.
4. Véase Andrew Steptoe y Helen Fidler, «Stage fright in musicians: a study of cognitive and behavioural strategies in performance anxiety», *cit.*
5. Véase Edward E. Jones y Steven Berglas, «Control of attributions about self through self-handicapping strategies: the appeal of alcohol and the role of underachievement», *Personality and Social Psychology Bulletin* 4 (1978), pp. 200-206.
6. Véase Peter J. Lang, *The Application of Psychophysiological Methods to the Study of Psychotherapy and Behaviour Change: An Empirical Analysis,* Nueva York, Willey, 1971.
7. Michelle G. Craske y Kenneth D. Craig, «Music performance anxiety: the three-systems model and self-efficacy theory», *Behavior Research & Therapy* 22 (1984), pp. 267-280.
8. Véase Mary L. Wolfe, «Correlates of adaptive and maladaptive musical performance anxiety», *cit.*
9. Véase *ibid.*
10. Véase Robert M. Yerkes y J. D. Dodson, «The relation of strength of stimulus to rapidity of habit-formation», *Journal of Comparative Neurology and Psychology* 18 (1908), pp. 459-482.
11. Andrew Steptoe, «The relationship between tension and the quality of musical performance», *Journal of the International Society for the Study of Tension in Performance* 1 (1983), pp. 12-22.

12. John A. Fazey y Lew Hardy, *The Inverted-U Hypothesis: A Catastrophe for Sport Psychology*, Leeds, The National Coaching Foundation, 1988.
13. Lew Hardy y Gaynor Parfitt, «A catastrophe model of anxiety and performance», *British Journal of Psychology* 82 (1991), pp. 163-178.
14. Glen D. Wilson, *Psychology for Performing Artists: Butterflies and Bouquets*, Londres, Jessica Kingsley, 1994.
15. *Ibid.*, la «ansiedad de rasgo» de la Figura 12.3 se refiere a una característica estable (es decir, permanente y persistente) de la personalidad.
16. Véase Andrew Steptoe, «The relationship between tension and the quality of musical performance», *cit.*
17. Véase Anthony Kemp, «The personality structure of the musician: I. Identifying a profile of traits for the performer», *Psychology of Music* 9 (1981), pp. 3-14.
18. Véase Susan E. Marchant-Haycox y Glenn D. Wilson, «Personality and stress in performing artists», *Personality and Individual Differences* 13 (1992), pp. 1061-1068.
19. Véase Andrew Steptoe y Helen Fidler, «Stage fright in musicians: a study of cognitive and behavioural strategies in performance anxiety», *cit.* Para otro análisis de los factores personales en la ansiedad de actuación, véase Jolanta Ossetin, «Psychological aspects of performance anxiety. Part I: Personality characteristics», *Journal of the International Society for the Study of Tension in Performance* 5 (1988), pp. 26-30.
20. Albert LeBlanc, Young Chang Jin, Mary Obert y Carolyn Siivola, «Effect of audience on music performance anxiety», *Journal of Research in Music Education* 45 (1997), pp. 480-486.
21. Mary A. Leglar, «Measurement of indicators of anxiety levels under varying conditions of musical performance», tesis doctoral, Universidad de Indiana, 1978.
22. Jennifer L. Abel y Kevin T. Larkin, «Anticipation of performance among musicians: physiological arousal, confidence, and state anxiety», *Psychology of Music* 18 (1990), pp. 171-182.
23. Melissa Brotons, «Effect of performing conditions on music performance, anxiety, and performance quality», *Journal of Music Therapy* 31 (1994), pp. 63-81.
24. Wendy J. Cox y Justin Kenardy, «Performance anxiety, social phobia, and setting effects in instrumental music students», *Journal of Anxiety Disorders* 7 (1993), pp. 49-60.
25. Donald L. Hamman y Martha Sobaje, «Anxiety and college musicians: study of performance conditions and subject variables», *Psychology of Music* 11 (1983), pp. 37-50.
26. Warren Brodsky y John A. Sloboda, «Clinical trial of a music generated vibrotactile therapeutic environment for musicians: main effects and outcome differences between therapy subgroups», *Journal of Music Therapy* 34 (1997), pp. 2-32.
27. Suzanne B. Hanser, «Music therapy in stress reduction research», *Journal of Music Therapy* 22 (1985), pp. 193-206.
28. Véase Mary L. Wolfe, «Coping with musical performance anxiety: problem-focused and emotion-focused strategies», *Medical Problems of Performing Artists* 5 (1990), pp. 33-36.
29. Resulta muy problemático diseñar estudios para investigar la eficacia de los diferentes tratamientos. Los requisitos básicos son la asignación aleatoria de los participantes a los tratamientos y las evaluaciones «ciegas», i.e. que los evaluadores no deben saber cuál es el tratamiento que ha recibido el participante. Idealmente, el tratamiento en cuestión ha de compararse con una alternativa plausible (un «tratamiento de control»), de modo que se puedan descartar los efectos de la expectativa —es decir, la creencia de que un tratamiento tendrá efecto. En los estudios que incluyen la utilización de fármacos, se emplean placebos. Dado que los efectos suelen ser de corta duración, es necesario incluir tratamientos complementarios a largo plazo.

30. Mary L. Wolfe, «Coping with musical performance anxiety: problem-focused and emotion-focused strategies», *cit.*, y Andrew Steptoe, «Stress, coping and stage fright in professional musicians», *Psychology of Music* 17 (1989), pp. 3-11.
31. Véase Alan Grishman, «Musicians' performance anxiety: the effectiveness of modified progressive muscle relaxation in reducing physiological, cognitive, and behavioral symptoms of anxiety», tesis doctoral, Universidad de Pittsburg, 1989.
32. Véase William R. LeVine y Jeffrey K. Irvine, «In vivo EMG biofeedback to reduce left-hand tension for string players», *American String Teacher* 31 (1981), pp. 10-12.
33. Robert Wesner, Russell Noyes, Jr. y Thomas L. Davis, «The occurence of performance anxiety among musicians», *Journal of Affective Disorders* 18 (1990), pp. 177-185, y Geoffrey I. D. Wills y Gary Cooper, *Pressure Sensitive: Popular Musicians Under Stress,* Londres, Sage, 1988.
34. Véase Yvon-Jacques Lavallee, Yves Lamontagne, Gilbert Pinard, Lawrence Annable y Léon Tétreault, «Effects on EMG feedback, diazepam and their combination on chronic anxiety», *Journal of Psychosomatic Research* 21 (1977), pp. 65-71.
35. Martin Fishbein, Susan E. Middelstadt, Victor Ottati, Susan Strauss y Alan Ellis, «Medical problems among ICSOM musicians: overview of a national survey», *cit.*, pp. 1-8.
36. Para un informe, véase Paul M. Lehrer, Raymond C. Rosen, John B. Kostis y Daniel Greenfield, «Treating stage fright in musicians: the use of beta-blockers», *New Jersey Medicine* 84 (1987), pp. 27-34.
37. Véase Klaus A. Neftel, Rolf H. Adler, Louis Kappeli, Maria Ross, Martin Dolder y Hans E. Kaser, «Stage fright in musicians: a model illustrating the effect of beta blockers», *Psychosomatic Medicine* 44 (1982), pp. 461-469.
38. Véase James e Imogen Savage, «Beneficial effect of nadolol on anxiety-induced disturbances of musical performance in musicians: a comparison with diazepam and placebo», *American Heart Journal* 108 (1984), pp. 1150-1155.
39. Véanse Michael Nielsen, «A study of stress amongst professional musicians», en Christopher Stevens (ed.), *Medical and Physiological Aspects of the Alexander Technique,* Aalborg, Dinamarca, International School for the Alexander Technique, 1988; y C. O. Brantigan y N. Joseph, «The effect of beta blockade on stage fright», *Rocky Mountain Medical Journal* 76 (1979), pp. 227-232.
40. Véase Barry Green y Timothy C. Gallwey, *The Inner Game of Music,* Garden City, NY, Doubleday, 1986.
41. Véase Harry E. Stanton, «Reduction in performance anxiety in music students», *Australian Psychologist* 29 (1994), pp. 124-147.
42. Michael Nielsen, «A study of stress amongst professional musicians», *cit.*
43. Elizabeth R. Valentine, David F. P. Fitzgerald, Tessa L. Gordon, Jennifer A. Hudson y Elizabeth R. C. Symonds, «The effect of lessons in the Alexander technique on music performance in high and low stress situations», *Psychology of Music* 23 (1995), pp. 486-497.
44. Sylvia S. Appel, «Modifying solo performance anxiety in adult pianists», *Journal of Music Therapy* 13 (1976), pp. 1-16. Aunque es menos eficaz que la desensibilización sistemática para reducir la ansiedad subjetiva, el análisis musical sí redujo el número de errores de ejecución.
45. Barbara A. McCune, «Functional performance anxiety modification in adult pianists», tesis doctoral, Columbia University Teachers' College, 1982.
46. Véanse Robert Triplett, *Stage Fright: Letting it Work for You,* Chicago, Nelson-Hall, 1983; Eloise Ristad, *A Soprano on Her Head: Right-side Up Reflections on Life and Other Performances,* Moan, Ut., Real People Press, 1982; y Kato Havas, *Stage Fright: Its Causes and Cures, with Special Reference to Violin Playing,* Londres, Bosworth, 1973.

47. Véase Donald Meichenbaum, *Stress Inoculation Training,* Nueva York, Pergamon, 1985.
48. Véase Paul Salmon, R. Schrodt y J. Wright, «A temporal gradient of anxiety in a stressful performance context», *Medical Problems of Performing Artists* 4 (1989), pp. 77-80.
49. Gladys Acevedo Sweeney y John J. Horan, «Separate and combined effects of cue-controlled relaxation and cognitive restructuring in the treatment of musical performance anxiety», *Journal of Counseling Psychology* 29 (1982), pp. 486-497.
50. Margaret J. Kendrick, Kenneth D. Craig, David M. Lawson y Park O. Davidson, «Cognitive and behavioral therapy for musical-performance anxiety», *Journal of Consulting and Clinical Psychology* 50 (1982), pp. 353-362.
51. Véase Julie J. Nagel, David P. Himle y James D. Papsdorf, «Cognitive-behavioural treatment of musical performance anxiety», *Psychology of Music* 17 (1989), pp. 12-21.
52. Duncan B. Clark y W. Stewart Agras, «The assessment and treatment of performance anxiety in musicians», *American Journal of Psychiatry* 148 (1991), pp. 598-605.

## *Lecturas complementarias*

Lehrer, Paul M.: «A review of the approaches to the management of tension and stage fright in music performance», *Journal for Research in Music Education* 35 (1987), pp. 143-152.

Salmon, Paul G.: «A psychological perspective on music performance anxiety: a review of the literature», *Medical Problems of Performing Artists* 5 (1990), pp. 2-11.

Steptoe, Andrew: «Performance anxiety. Recent developments in its analysis and management», *Musical Times* 23 (1982), pp. 537-541.

Wilson, Glenn D.: «Performance anxiety», en David J. Heargraves y Adrian C. North (eds.), *The Social Psychology of Music,* Oxford, Oxford University Press, 1997, pp. 229-248.

— «Stage fright and optimal performance», en *Psychology for Performing Artists: Butterflies and Bouquets,* Londres, Jessica Kingsley, 1994, pp. 185-203.

# CUARTA PARTE
# *Interpretar la interpretación*

# 13 *Escuchar la interpretación*

ERIC CLARKE

## *Introducción*

Los oyentes tienen acceso hoy en día a una mayor cantidad y variedad de interpretaciones musicales que en cualquier otra época de la historia, y su rápida expansión, que comenzó con la llegada de la industria de la grabación a principios del siglo XX, no muestra señales de disminuir. Con la proliferación y diversificación de la interpretación a través de las grabaciones, ha llegado también un renovado interés por los ejecutantes como intérpretes y personalidades. Este capítulo trata sobre las diversas circunstancias en las que se escuchan las interpretaciones, sus respectivas repercusiones y los procesos perceptivos implicados en lo que los oyentes pueden escuchar y escuchan en la interpretación.

El título de este capítulo sugiere que es posible escuchar el componente interpretativo de lo que podríamos llamar «el sonido total de la música». ¿Se puede defender este punto de vista? ¿Hasta qué punto y de qué manera los oyentes son conscientes de la interpretación como un elemento que se puede separar de la música? De hecho, ¿se puede escuchar *otra cosa* que la interpretación? En el caso de la música que se basa en la partitura, y que conforma la tradición concertística occidental, podríamos justificar una distinción entre «la música» y «la interpretación», aunque sólo sea porque la partitura es una representación de la música, independiente de cualquier interpretación *particular.* Pero ¿qué ocurre con la enorme cantidad de oyentes que rara vez, si acaso, ven las partituras, o con la abrumadora mayoría de músicas que carecen de partituras? En resumidas cuentas, ¿acaso la gente escucha

la interpretación (en oposición a la música)? Y, de ser así, ¿qué es lo que escuchan?

Una prueba fehaciente de la importancia que tiene la interpretación para los oyentes es el enorme empaque que nuestra cultura le da a la interpretación: en la cultura musical contemporánea, los intérpretes son estrellas, como demuestran frecuentemente los diseños de las carátulas de los CDs [1] y como revelan igualmente sus honorarios y sus perfiles públicos. La industria de la grabación depende para su supervivencia del apetito casi insaciable de un público consumidor de grabaciones nuevas o reeditadas de un repertorio mayormente estático, y la industria fomenta a su vez la importancia de la interpretación por razones comerciales obvias. Esta prueba por sí sola plantea claramente una cuestión que requiere una respuesta, del mismo modo que el florecimiento de los estudios sobre la interpretación refleja el creciente interés académico por ella.

### *Circunstancias y actitudes*

Hasta aproximadamente 1910, sólo existía la interpretación en vivo: la música se escuchaba en presencia de los intérpretes, que generalmente estaban a la vista. En esas circunstancias sería difícil *no* percatarse de la interpretación (o al menos de los intérpretes) en el plano directo de la presencia física. Sin embargo, antes del nacimiento de la cultura concertística, hacia 1800, la interpretación y los intérpretes estaban profundamente integrados dentro de las funciones sociales de las que formaban parte explícita (la iglesia, la corte, las celebraciones profanas, el entretenimiento doméstico, las funciones militares). Esa integración social puede implicar que era difícil separar la interpretación musical de su función y sus circunstancias, y por lo tanto es poco probable que destacara como un objeto de contemplación o un tema para el análisis. Evidentemente, ése no era el caso: incluso san Agustín, por ejemplo, cuando escribe sobre la cuestión de la música por la música y la música al servicio de Dios, reconoce que una interpretación lograda le podía «cautivar», y que el placer estético podía superar a la razón religiosa [2].

Sin embargo, en comparación con la posición que ocupaba la música dentro de sus contextos sacro y profano, la cultura concertística trajo consigo una perspectiva enfocada en la interpretación y la colocó, literalmente, en el centro del escenario. Los intérpretes fueron separados de sus oyentes, y a menudo se les colocaba sobre una plataforma; sólo tocaban cuando había un público, cuyos miembros escuchaban con mucha atención y relativamente callados. Estas distribuciones físicas y sociales forman parte de una actitud históricamente específica hacia la interpretación y la audición, y todos esos

factores —que siguen caracterizando a las condiciones convencionales para las interpretaciones en vivo del repertorio concertístico [3]— se conjugan para presentar la música como algo autónomo (independiente, autosuficiente) que ocupa un ámbito especial. Por tanto, no es de extrañar que los grandes intérpretes —los «realizadores» de la música— sean considerados con el mismo respeto que han inspirado los eruditos o los genios.

Sin embargo, la interpretación en vivo ha sido eclipsada por la grabación (al menos en términos cuantitativos), y hoy día se escucha mucha más música grabada o transmitida por los medios de comunicación, lo cual repercute significativamente en las actitudes de audición. Cuando los discos y gramófonos comenzaron a estar al alcance de muchos, durante el primer cuarto del siglo XX, mantenían algo del carácter social de escuchar una interpretación en vivo. Pero el gramófono también permitía una solitaria «audición acusmática», en la que la fuente instrumental del sonido está oculta, que fomenta una concentración en el sonido mismo que anteriormente apenas era posible [4]. Aunque tal vez no representen el mayor refinamiento del sonido, las diversas versiones de reproductores portátiles con audífonos pueden considerarse como la culminación de la audición acusmática, permitiendo al oyente llevar consigo y controlar un mundo sonoro virtual localizado aparentemente dentro de su propia cabeza.

Los efectos que produce la grabación en la audición son paradójicos. Por un lado, la naturaleza acusmática de las grabaciones —la ausencia física del intérprete— implica que la interpretación misma se encuentra en un plano más secundario, más implícito, que en una interpretación en vivo. Por otra parte, el advenimiento de la grabación, junto con un repertorio más o menos fijo de música que se regraba constantemente, significa que la interpretación en sí misma se ha convertido en el centro de atención. Además, la historia de la grabación demuestra que persiste el dilema sobre cómo se deberían entender la grabaciones —como interpretaciones plasmadas o como creaciones de los estudios de grabación [5]. Algunas compañías especializadas en grabaciones «de una sola toma», como Nimbus, del Reino Unido, apoyan el punto de vista de que las grabaciones deberían ser como las interpretaciones, del mismo modo que muchos intérpretes afirman que lo que intentan hacer en un estudio es grabar una interpretación. Por el contrario, hay intérpretes —entre los cuales Glen Gould fue quizá el representante más famoso— que insisten en que una grabación difiere mucho de una interpretación, que la edición permite crear algo que no se puede lograr en vivo y que las grabaciones deben entenderse como una manifestación distinta de la música. Las grabación sumamente editada de múltiples tomas no es un «remiendo»: es una expresión de la música que se ha construido deliberada y cuidadosamente, que posee sus propias características y requiere cri-

terios de evaluación específicos. Los procesos que tienen lugar en un estudio están completamente entrelazados con lo que se considera convencionalmente como factores de interpretación, y el resultado es simplemente otro tipo de objeto/evento en comparación con una interpretación real. Esto se evidencia más aún en otros géneros de música: un gran porcentaje de música pop y electroacústica implica un tipo de producción de estudio que no admite la interpretación en vivo en el sentido habitual del término, y hay piezas del repertorio concertístico que tienen versiones diferentes para la grabación y la interpretación en vivo. Por ejemplo, la obra *New York Counterpoint* de Steve Reich puede ser interpretada en vivo por once clarinetistas o grabada por uno solo utilizando un multipistas[6].

En resumen, la grabación incorpora varias consideraciones (como las técnicas de grabación y posproducción, la edición, el enfoque interpretativo inducido o requerido por las circunstancias del estudio, la fidelidad y el carácter acústico del medio de grabación y la naturaleza física y acústica del soporte en el que se reproduce la grabación) que implican que debemos ser cuidadosos al equiparar la música grabada con la interpretación grabada. Algunas grabaciones pretenden claramente ser lo mismo que las interpretaciones —las más obvias son las grabaciones en vivo de conciertos y óperas. Pero incluso en esos casos, su carácter acusmático cambia significativamente la relación del oyente con la música, y el resultado tiene propiedades (entre ellas, que se puede repetir, interrumpir y transformar en el momento de la reproducción) que pueden cambiar profundamente la manera en que los oyentes utilizan y responden a esas «interpretaciones». El no reconocer que las grabaciones *no* son simplemente interpretaciones convenientemente empaquetadas, y que las prácticas y expectaciones auditivas relacionadas con las grabaciones también son bastante diferentes de las relacionadas con los conciertos, son quizá dos de las razones por las que muchas personas encuentran que la relación entre sus experiencias con las grabaciones y los conciertos es incómoda y quizá decepcionante. No debemos pretender escuchar ambas de la misma manera.

He sugerido que las grabaciones pueden destacar los factores interpretativos. ¿Existen circunstancias musicales particulares en la interpretación en vivo que induzcan a los oyentes a percatarse de la interpretación como tal? El fracaso y el éxito pueden hacer que la interpretación sea particularmente audible: cuando las cosas van francamente mal o el intérprete es irremediablemente incompetente, el abismo entre lo que la música requiere y lo que el intérprete es capaz de dar se hace demasiado evidente. Por el contrario, cuando el intérprete parece tener un control absoluto de una música técnicamente exigente, sus proezas técnicas y su «atleticismo» pueden ser claramente audibles, tal y como pretenden demostrar las piezas sobresalientes de los

conciertos de gala y concursos musicales. Así pues, en sus extremos, la interpretación es audible como incompetencia o virtuosismo.

Pero los extremos son inusuales, y lo que los oyentes escuchan en circunstancias más normales resulta menos obvio. Un factor importante puede ser el tipo de oyente en cuestión. Por una parte, están aquellos para quienes distinguir los niveles de excelencia interpretativa forma parte de su profesión y que han cultivado un oído agudo, aunque a veces muy particular, para la interpretación (por ejemplo, los jurados de los concursos y audiciones musicales). Por otra parte, hay oyentes que pueden desconocer completamente un determinado estilo musical y sus convenciones y tradiciones interpretativas. Los oyentes aprenden a reconocer y evaluar las características interpretativas de cualquier estilo durante un período prolongado de continuo aprendizaje perceptual, y generalmente de manera informal y no supervisada. La mayoría de esas habilidades auditivas se desarrollan mediante el simple contacto y un constante intercambio de opiniones y preferencias con otros oyentes, junto con la influencia de las reseñas de CDs y conciertos, los programas de radio en los que se evalúan diferentes grabaciones y la influencia más indirecta de las representaciones de interpretaciones y audiciones en películas, programas de televisión y textos. No hay una investigación sistemática sobre cómo se desarrollan estas habilidades auditivas usuales, cuáles son los factores e influencias básicas y hasta qué punto las evaluaciones de los profesionales son consecuentes o fiables. Sin embargo, los hechos confirman reiteradamente que los oyentes con diferentes habilidades e intereses escuchan y responden a las interpretaciones en formas que pueden ser radicalmente distintas.

Sin embargo, las opiniones estéticas también pueden desempeñar un papel importante, ya que los oyentes pueden desear o esperar diferentes *tipos* de relaciones entre el intérprete y el material —y para distintos tipos de música. Dentro de la tradición concertística occidental, podemos identificar dos extremos de una escala que influyen significativamente en la audición, definida por Wilfrid Mellers como la diferencia entre el ejecutante intermediario y el ejecutante intérprete [7]. Mellers describe a los «intermediarios» como ejecutantes que sirven como medio de expresión para las intenciones del compositor, y a los «intérpretes» como aquellos para quienes las intenciones del compositor (o, más concretamente, la partitura) son un punto de partida para su propia labor creativa/interpretativa. Aunque este asunto tiene una variable histórica, también se pueden establecer diferencias entre los diversos enfoques de una misma época: en los años 1940 y 1950, por ejemplo, los directores Toscanini y Furtwängler fueron calificados respectivamente de «intermediario» e «intérprete», siendo el primero venerado (o severamente criticado) por su «fidelidad a la partitura», y el segundo por su agudeza creativa y su idiosincrasia. Si existen discrepancias tan fundamentales sobre lo

que los intérpretes deberían *intentar* hacer, no es de extrañar que los oyentes puedan tener reacciones radicalmente divergentes: lo que a uno le resulta díscolo y caprichoso (basándose quizá en que «se aleja de la partitura») a otro le puede parecer comprometido e inspirado (por la «originalidad e idiosincrasia de su interpretación»).

En muchas tradiciones musicales fuera de la música culta occidental, la situación es aún más variable. No es que la distinción entre el intérprete y el material desaparezca en las tradiciones orales, pero la relación cambia. Por ejemplo, en el jazz tradicional, que posee un núcleo de melodías «clásicas» que son reinterpretadas continuamente, tiene poco sentido considerar si un intérprete es más o menos fiel al original, aunque sólo sea porque a menudo existe cierto grado de incertidumbre sobre lo que es (o era) realmente el «original» y un elevado nivel de ambigüedad incluso en los ejemplos más establecidos o definidos. Un «clásico» tiene una identidad que frecuentemente no es mucho más que una estructura armónica/métrica y una melodía, y representa una base sobre la cual se espera que los intérpretes exhiban sus habilidades improvisatorias. Algunos oyentes pueden tener un sentido de la fidelidad al «espíritu» del original y rechazar por consiguiente los tratamientos radicales de los intérpretes peculiares, pero existe claramente un terreno considerable para la libertad interpretativa del ejecutante —así como para la expectación de los oyentes, quienes cuentan con que efectivamente sea así. Los oyentes de jazz están más interesados en la interpretación y los intérpretes que en las piezas y repertorios.

Sin embargo, el interés principal de este capítulo es la tradición concertística occidental y qué es lo que los oyentes escuchan en la interpretación de esa música. Uno de los ideales de interpretación que se ha propuesto a menudo es que el oyente debería sentirse arrebatado por la fuerza de la interpretación sin percatarse del ejecutante como intermediario o como intérprete. Para algunos oyentes, por lo tanto, la audición gratificante y concentrada hace que, paradójicamente, la interpretación sea inaudible: consideran que la música les habla a ellos directamente. Peter Johnson escribió lo siguiente: «La interpretación extraordinaria de una buena pieza musical es, sugiero, una invitación a la audición trascendental, en el sentido de que, paradigmáticamente, evita llamar la atención hacia sí misma *como* una interpretación (bien sea por razones positivas o negativas)» [8]. Si el no percatarse de la interpretación es para algunos oyentes una medida del éxito de ésta, entonces el ideal del intérprete podría ser estar tan involucrado en la música, tan poseído por ella, que se sienta identificado con el compositor o con la obra. El punto de vista contrario sería una fascinación por la interpretación misma, con la manera en que los ejecutantes hacen lo que hacen y las ilimitadas maneras en que la música se realiza en la interpretación.

## *¿Qué especifican los sonidos de la interpretación?*

Aunque es posible que un oyente atento e interesado sea poco o nada consciente de las características específicas de una interpretación, ese oyente puede *llegar a ser* consciente de los atributos de ésta. Además, como la percepción implica no sólo enterarse conscientemente, sino también un proceso continuo e inconsciente de ajuste y reacción al entorno, los oyentes también serán influidos por las características de la interpretación, independientemente de que se den cuenta o no. Los sonidos de una interpretación tienen la capacidad de transmitir una gran cantidad de información al oyente, desde las características físicas relacionadas con el espacio en el que tiene lugar la interpretación y la naturaleza del instrumento hasta las propiedades menos tangibles, como la ideología interpretativa del ejecutante. Una forma de abordar esa gran diversidad es preguntarse qué especifican los sonidos de la interpretación, y al hacerlo prestar toda la atención a las características de los sonidos de la interpretación y su riqueza como fuente de información.

### *Espacio/instrumento/cuerpo*

Las propiedades más físicas y tangibles de las interpretaciones son los espacios en los que se realizan, los instrumentos que se utilizan y las interacciones físicas del ejecutante con ellos. En algunas músicas (por ejemplo, la música antifonal de Andrea Gabrieli, o la obra *Gruppen* de Karlheinz Stockhausen, en la que la música está distribuida espacialmente), la «geometría» del espacio es importante como componente de la estructura musical. Pero en gran parte de la música de la tradición concertística occidental, el espacio es importante no como elemento estructural, sino por la manera en que transmite el «aura» o carácter de la interpretación. Muchos oyentes prefieren escuchar una obra como la *Pasión según San Mateo* de Bach, por ejemplo, en un «espacio» (bien sea el espacio real de la interpretación o el espacio virtual de una grabación) amplio y resonante, por al menos dos razones. En primer lugar, una acústica amplia y resonante se corresponde con el tipo de espacio para el cual fue escrita esa música (una iglesia), y al transmitir ese espacio, la acústica también transmite los materiales (piedra y madera), los entornos y la ideología (el culto cristiano) que están relacionados con la interpretación de esa música. En segundo lugar, la música fue compuesta probablemente para explotar la acústica a la que estaba destinada, y por tanto funciona mejor con la reverberación amplia (y la consiguiente ola de sonido) que proveen dichos espacios. Por el contrario, la mayoría del repertorio de música de cámara fue compuesto para la interpretación doméstica, y la música puede verse afectada

si se toca en una acústica muy reverberante, no sólo porque las relaciones serán erróneas, sino también porque las texturas musicales se confunden.

Además de los espacios de interpretación, también influyen los instrumentos y los ejecutantes, cuyas ejecuciones en el instrumento se reflejan en sonidos. El sonido de una interpretación proyecta no sólo propiedades básicas como el instrumento que se está tocando, sino también la interacción física del ejecutante con él. La facilidad o el esfuerzo, la energía, la confianza y el nivel de control de una interpretación son transmitidos claramente en el sonido de una interpretación, y proporcionan al oyente una información sobre el cuerpo del intérprete y su relación con el instrumento.

Las actitudes convencionales hacia el repertorio concertístico estándar han tendido a subestimar el carácter físico de la interpretación, con respecto al cual los oyentes suelen ser ambiguos. Dentro de una tradición que admira la virtuosidad, los oyentes generalmente prefieren no percatarse de los esfuerzos que realiza el intérprete para mantener el control técnico de las piezas difíciles, pero sí quieren percibir su presencia física, implicados en mente *y* cuerpo en la «lucha heroica» por expresar y comunicar. Uno de los defectos que presentan las simulaciones de interpretaciones humanas con sistemas informáticos (como el sistema CHANT de simulación vocal desarrollado por el ICAM de París en los años 1980) es que carecen de toda sensación de compromiso y «esfuerzo» físicos.

### *Eventos y procesos*

Escuchar la interpretación tiene mucho en común con escuchar el habla, ya que el oyente debe seguir la pista de un flujo continuo de eventos en el tiempo y encontrarles un sentido. Aunque gran parte de la teoría de la música ha destacado la importancia de la estructura musical a gran escala [9], también existen pruebas de que los oyentes no son sensibles a esas características a gran escala de una manera directamente perceptual [10]. Jerrold Levinson [11] ha sostenido que los oyentes no construyen estructuras jerárquicas elaboradas y a gran escala, sino que se centran fundamentalmente en el desarrollo continuo de una pequeña porción de música «presente». Esa idea concuerda con los estudios psicológicos sobre la percepción del tiempo, que sugieren que en todo momento somos conscientes de un «trozo» relativamente breve de tiempo (unos 6-10 segundos), al que denominan presente perceptual [12]. Las estructuras más grandes son una función de la memoria, y Levinson afirma que la memoria del oyente para esas estructuras suele ser bastante imprecisa y no admite una «arquitectura» a gran escala. En otras palabras, el acto de escuchar es como una ola en la que los eventos presentes y contiguos producen

un efecto vívido e inmediato, que deja tras de sí una estela de recuerdos cuya historia de eventos e influencias acumuladas le da a la ola una forma determinada.

Al escuchar una interpretación, somos fundamentalmente conscientes de lo que ocurre en el *presente* de un flujo continuo. Sin embargo, ese flujo no es uniforme: en la interpretación de la mayoría de las músicas, hay determinados eventos que lo segmentan, proporcionando periódicamente la oportunidad de organizar los eventos de los últimos segundos y almacenarlos en la memoria. De la misma manera, cuando escuchamos a alguien hablando, nos percatamos en todo momento de los contenidos activos de la frase u oración en curso —las palabras dichas en unos pocos segundos. Cuando el discurso llega a un punto o un punto y coma, sintetizamos los contenidos de dicha frase u oración como la esencia de lo que el hablante acaba de decir —tal vez recordando literalmente la frase, pero principalmente en términos de lo que significa. Y al mismo tiempo que esa parcela de significado está siendo registrada, se abre una nueva ventana de percepción inmediata que permite al oyente proseguir su audición. En la música, esta segmentación del flujo de información se deriva de una mezcla de propiedades estructurales (notas largas, discontinuidades de alturas, funciones tonales, etc.) y características interpretativas (cambios de tempo, dinámica, articulación, etc.). Las diversas maneras en que los intérpretes dan forma a los eventos de la interpretación son analizadas en varios capítulos de este libro; la siguiente sección considera si los oyentes captan o no esas características interpretativas, cómo lo hacen y cómo las interpretan.

## *Expresión, emoción y estilo*

¿En qué medida los oyentes son conscientes de las microvariaciones de interpretación, y de qué manera les dan sentido? Varios estudios han demostrado que, bajo condiciones controladas, los oyentes pueden detectar cambios mínimos en el tempo, del orden de los 30 a 50 milisegundos [13], y otros estudios han revelado que perciben igualmente los cambios en la dinámica y la entonación [14]. Sobre la perceptibilidad de la articulación, el timbre y el vibrato se tiene menos información, en parte porque son características más complejas que la duración, altura o intensidad del sonido. Igualmente, se han realizado pocas investigaciones sobre la percepción de dichas características por parte de los oyentes en condiciones normales de audición.

Bruno Repp, sin embargo, ha demostrado que la capacidad que tienen los oyentes de detectar características detalladas de interpretación está marcadamente influida por el contexto musical. Cuando a un grupo de oyentes se

les hizo escuchar una breve sección de una sonata para piano de Beethoven (el inicio del Minueto de la Sonata op. 31 núm. 3) en la que se prolongó la duración de determinadas notas en varios puntos, éstos mostraron una variación perceptual considerable dependiendo de la ubicación de la prolongación [15]. En los puntos donde esperaban fluctuaciones expresivas en el tempo (características de los finales de frases), los oyentes no detectaron los cambios —como si por el hecho de ser previsibles dichos cambios resultaran imperceptibles. Por el contrario, cuando las fluctuaciones expresivas de tempo ocurrían en mitad de una frase, donde son menos probables, los oyentes fueron particularmente conscientes de su presencia —como si resultaran más evidentes por ser incongruentes.

Aunque sólo considera un aspecto de la interpretación, el estudio de Repp ilustra un principio general importante: que la capacidad de los oyentes de detectar las características detalladas de la interpretación depende mucho del contexto. Este principio puede influir tanto en la *conciencia* que tiene el oyente de las características de la interpretación como en su manera de interpretarlas —es decir, cómo funcionan esas características como signos [16]. Una ralentización del tempo en medio de una frase puede interpretarse como una intensificación expresiva, mientras que la misma ralentización al final de una frase puede entenderse como una disminución de la tensión, marcando la conclusión de un pasaje musical. Las diferentes interpretaciones de una «misma» característica interpretativa son producto del sistema de relaciones estructurales y expresivas en el que se inserta cada característica de la interpretación.

Las propiedades expresivas de la interpretación transmiten cualidades emocionales y características estructurales. Patrik Juslin ha demostrado que cuando a un intérprete se le pide que toque una pieza de tal modo que transmita un estado emocional específico, los oyentes pueden reconocer con seguridad qué estado está comunicando [17]. Ésta es una versión controlada (y un tanto artificial) de un hecho común, a saber, que los oyentes perciben emociones transmitidas por la interpretación y suelen atribuirlas a los intérpretes —independientemente de que las escuchen como verdaderos estados emocionales de los intérpretes o como emociones actuadas de una personalidad adoptada.

Más de cien años de interpretaciones grabadas, junto con el desarrollo de las prácticas interpretativas históricas, han puesto a nuestro alcance una fascinante diversidad de estilos e ideologías interpretativas. El estudio sistemático sobre las diversas maneras en que éstas se especifican en sonido apenas ha comenzado (aunque la obra de Robert Philip constituye un punto de partida) [18], pero es potencialmente fascinante. Para escuchar la expresión, la emoción, el estilo y la ideología de una interpretación, el oyente ha de identificar

características de la interpretación que destacan sobre un fondo implícito de neutralidad —una especie de «norma» teórica con respecto a la cual se distinguen la expresión, la emoción, el estilo y la ideología (como se explica en el capítulo 4). En otras palabras, la interpretación se escucha sobre el telón de fondo de las interpretaciones y grabaciones previas. Los continuos cambios históricos implican que las normas de una época se convierten en las idiosincrasias o anacronismos de otra, y ésa es la razón por la que las grabaciones antiguas pueden resultarles amaneradas o peculiares a los oyentes contemporáneos. De nuevo, la enorme proliferación de interpretaciones grabadas y transmitidas por la radio y la televisión durante el siglo XX trajo cambios profundos para los intérpretes y oyentes, quienes ahora tienen acceso a una cantidad y variedad inusitada de interpretaciones de un enorme repertorio musical. Algunos lo han considerado un proceso enriquecedor y provechoso, mientras que para otros, como Hans Keller, ha contribuido a la desastrosa erosión de las habilidades interpretativas y auditivas [19].

## *Conclusión*

Existe una opinión según la cual la interpretación exitosa y la audición concentrada son aquellas en las que los oyentes no nos percatamos de la interpretación —y, por consiguiente, este capítulo no tendría razón de ser. Sin embargo, es una opinión innecesariamente limitada: supone una estética auditiva particular que está lejos de ser compartida universalmente, se relaciona con circunstancias particulares de audición, depende de condiciones idealizadas y sólo se refiere a la percepción consciente. Como se ha demostrado en este capítulo, existen muchas circunstancias en las que los oyentes se percatan indudable, y a veces desafortunadamente, de la interpretación, y quizá los oyentes *siempre* responden a la interpretación, incluso cuando no se percatan de ella. Sin embargo, una característica de la música de concierto y la actitud con que se escucha, que resulta del concepto integral de la obra musical y su (separable) interpretación, es la arraigada incertidumbre o ambigüedad sobre qué es lo que uno está, o debería estar, escuchando. La «obra como concepto» sugiere que la interpretación debería pasar inadvertida, mientras que el auge del intérprete como artista proviene de un interés por la interpretación en sí, y al mismo tiempo fomenta dicho interés. Pero, independientemente de que los oyentes crean que están escuchando la interpretación o «la obra misma», no pueden negar que lo que oyen es una interpretación (o una grabación), y el segundo objetivo de este capítulo ha sido demostrar la cantidad y diversidad de información que se transmite a través de los sonidos de una interpretación. Por último, en una época en que se escucha mucha más

música grabada y emitida por los medios audiovisuales que interpretada en vivo, aún no hemos definido claramente qué es una grabación —una interpretación «plasmada» o un producto de un estudio de grabación—, con todas las consecuencias que ello implica en cuanto a nuestras respuestas y actitudes hacia las grabaciones. En una época en que se cuestiona cada vez más la prioridad que se le da a la partitura cuando se escribe sobre música, y en que las interpretaciones y grabaciones se están convirtiendo en objetos de estudio mucho más comunes, sabemos más sobre lo que se puede *medir* de las interpretaciones y grabaciones, pero aún sabemos muy poco sobre lo que los oyentes realmente *oyen* en ellas.

## *Notas*

1. Véase Nicholas Cook, «The domestic *Gesamtkunstwerk*, or record sleeves and reception», en Wyndham Thomas (ed.), *Composition, Performance, Reception: Studies in the Creative Process in Music*, Aldershot, Ashgate, 1998, pp. 105-117.
2. Oliver Strunk, *Source Readings in Music History: From Classical Antiquity through the Romantic Era*, Nueva York, Norton, 1950, p. 74.
3. Para una descripción fascinante de los rituales y circunstancias de la audición de conciertos en el siglo XX, véase Christopher Small, «Performance as ritual: sketch for an enquiry into the true nature of a symphony concert», en Avron Levine White (ed.), *Lost in Music: Culture, Style and the Musical Event*, Londres, Routledge and Kegan Paul, 1987, pp. 6-32.
4. Véase Denis Smalley, «The listening imagination: listening in the electroacoustic era», en John Paynter, Tim Howell, Richard Orton y Peter Seymour (eds.), *Companion to Contemporary Musical Thought*, 2 vols., Londres y Nueva York, Routledge, 1992, I, pp. 514-554.
5. Véase Michael Chanan, *Repeated Takes: A Short History of Recording and its Effects on Music*, Londres, Verso, 1995.
6. Un ejemplo de esto último es la grabación que realizó Evan Ziporyn para el sello Nonesuch (7559-79451-2).
7. Wilfrid Mellers, «Present and past: intermediaries and interpreters», en Paynter et al. (eds.), *Companion to Contemporary Musical Thought*, cit., II, pp. 920-930.
8. Peter Johnson, «Performance and the listening experience: Bach's "Erbarme dich"», en Nicholas Cook, Peter Johnson y Hans Zender (eds.), *Theory into Practice: Composition, Performance and the Listening Experience*, Leuven, Leuven University Press, 1999, p. 85.
9. E.g., Heinrich Schenker, *Free Composition (Der freie Satz)*, Ernst Oster (trad. y ed.), Nueva York, Longman, 1979; y Fred Lerdahl y Ray Jackendoff, *A Generative Theory of Tonal Music*, Cambridge, Mass., MIT Press, 1983.
10. Véase, por ejemplo, Nicholas Cook, «The perception of Large-scale tonal closure», *Music Perception* 5 (1987), pp. 197-205.
11. Jerrold Levinson, *Music in the Moment*, Ithaca y Londres, Cornell University Press, 1997.
12. Véanse John A. Michon, «The making of the present: a tutorial review», en John Requin (ed.), *Attention and Performance VII*, Hillsdale, NJ, Erlbaum, 1978, pp. 89-111; y Eric F. Clarke, «Levels of structure in the organization of musical time», *Contemporary Music Review* 2 (1987), pp. 211-238.
13. Véase Eric F. Clarke, «The perception of expressive timing in music», *Psychological Research* 51 (1989), pp. 2-9.

14. Para una perspectiva general, véase Brian Moore, *An Introduction to the Psychology of Hearing,* 4ª ed., Londres, Academic Press, 1997.
15. Bruno H. Repp, «Probing the cognitive representation of musical time: structural constraints on the perception of timing perturbations», *Cognition* 44 (1992), pp. 241-281.
16. Véase Eric F. Clarke, «Expression in performance generativity, perception and semiosis», en John Rink (ed.), *The Practice of Performance: Studies in Musical Interpretation,* Cambridge, Cambridge University Press, 1995, pp. 21-54.
17. Véase, por ejemplo, Patrik Juslin, «Emotional communication in music performance: a functionalist perspective and some data», *Music Perception* 14 (1997), pp. 383-418.
18. Robert Philip, *Early Recordings and Musical Style: Changing Tastes in Instrumental Performance, 1900-1950,* Cambridge, Cambridge University Press, 1992; véase también el capítulo 14 de este libro.
19. Hans Keller, «The gramophone record», en Robert Matthew-Walker (ed.), *The Keller Column: Essays by Hans Keller,* Londres, Alfred Lengnick, 1990, pp. 22-25.

## *Lecturas complementarias*

Chanan, Michael: *Repeated Takes: A Short History of Recording and its Effects on Music,* Londres, Verso, 1995.

Day, Timothy: *A Century of Recorded Music,* New Haven, Yale University Press, 2000.

Francès, Robert: *The Perception of Music,* W. Jay Dowling (trad.), Hillsdale, NJ, Erlbaum, 1988.

Gaver, William: «What in the world do we hear? An ecological approach to auditory event perception», *Ecological Psychology* 5 (1993), pp. 1-29.

Johnson, James H.: *Listening in Paris: A Cultural History,* Berkeley, University of California Press, 1995.

Storr, Anthony: *Music and the Mind,* Londres, Harper Collins, 1992.

# *14 El legado de las grabaciones*

PETER JOHNSON

La imagen visual que más ha perdurado en toda la historia de la grabación es, sin duda, la de Nipper el perro, escudriñando nostálgicamente el enorme altavoz de un viejo gramófono. Su título, claro está, es «His Master's Voice» («La voz de su amo»). En cierto modo, fue una elección curiosa para un sello discográfico, ya que habla de ausencia: el amo había fallecido años atrás. Las grabaciones musicales representan más bien una presencia, las voces de los intérpretes y el sonido de una música real.

Tal vez sea porque las grabaciones se han incorporado tanto a nuestra vida diaria que habitualmente las aceptamos en lugar de las interpretaciones en vivo. De alguna manera nos las arreglamos sin el contacto visual entre el intérprete y el oyente, o sin esa sensación intangible de la música tomando cuerpo ante nosotros. Aceptamos que el pasaje virtuoso siempre será ejecutado de la misma manera, y que al intérprete probablemente le supuso una docena de tomas alcanzar su nivel de perfección. Y, habitualmente, no consideramos que una grabación profesional sea menos auténtica que una interpretación en vivo: al contrario, la grabación reafirma su autoridad por encima del efímero concierto que pronto se olvida. Aún así, las grabaciones no han reemplazado a los conciertos en vivo: proporcionan un modo alternativo de producción musical, de gran valor para los intérpretes y su público, y para los estudiosos y compositores.

### *Voz y personaje*

Si las grabaciones difieren significativamente de las interpretaciones en vivo, ¿qué es lo que tienen en común? La «voz» interpretativa de todo gran intérprete, sea Håkan Hardenberger o Heinz Holliger, Yo Yo Ma o Yuri Bashmet, es inconfundible una vez que la hemos escuchado, bien sea en CD o en concierto. Esa voz interpretativa es lo que las grabaciones captan tan eficazmente, y lo que se reproduce cada vez que ponemos un disco o un CD. Una grabación es, literalmente, *de* el intérprete o el conjunto que interpreta la obra. En ese sentido, las compañías discográficas tienen razón al comercializar sus productos por intérpretes y no por obras [1].

Cada obra, sin embargo, tiene sus propias exigencias interpretativas y, al igual que el actor, el buen intérprete ajusta su voz correspondientemente. Me referiré a este segundo nivel de presencia como el «personaje» del intérprete [2]. En las interpretaciones en vivo vemos a los intérpretes entrar y salir de su papel: la figura modesta de Alfred Brendel se transforma en el pianista magistral en el momento en que sus dedos tocan las teclas, pero retorna cuando agradece tranquilamente los aplausos. Las grabaciones, al igual que las películas, no nos dejan ver esas transformaciones: nos muestran solamente al personaje. Lo que sí nos permiten es comparar a esos personajes: la *Madame Butterfly* de Callas es muy diferente de su *Tosca;* sin embargo, la voz es siempre la de Callas.

### *Las grabaciones como testimonios*

Muchas veces se ha preguntado si los artistas de antaño como Schnabel o Cortot realmente tocaban como sugieren sus grabaciones. En cierto modo, es una pregunta que no viene al caso, ya que las grabaciones son lo que son, y no lo que podrían haber sido. En las grabaciones posteriores a la guerra, el productor ejercía probablemente un control considerable sobre el sonido e incluso sobre la interpretación; sin embargo, el producto en sí mismo puede ser de gran interés musical e histórico. Las grabaciones pueden revelar muy claramente cómo cantaban o tocaban realmente los intérpretes desde comienzos del siglo XX. La cuestión es cómo las analizamos.

Hasta ahora ha habido muy pocos estudios musicológicos serios sobre las grabaciones, ya que la musicología del siglo XX ha tenido otras prioridades. Por lo tanto, este capítulo está dirigido principalmente al intérprete en activo: ¿qué puede esperar aprender el intérprete actual del legado de las grabaciones? Los tres problemas más inmediatos son la magnitud del archivo, la complejidad del arte interpretativo en sí mismo y la influencia que ejerce el

medio de grabación en lo que oímos. Podríamos limitar el campo concentrándonos en nuestro propio tipo de voz o instrumento, o en las múltiples grabaciones de una obra o un movimiento [3]. Sin embargo, este capítulo adoptará un enfoque temático, destacando algunas de las áreas más provechosas e ilustrando ciertas técnicas analíticas. Todos esos enfoques plantean cuestiones mucho más amplias —la interpretación y el significado— que consideraré brevemente en una sección de conclusión. Quisiera subrayar que este capítulo es una mera introducción, ya que el tema es muy amplio; en un terreno tan poco explorado, el viaje promete ser excitante, aunque a veces arriesgado.

## *Métodos de grabación*

El primer paso consiste en conocer y comprender las diversas tecnologías y cómo pueden haber influido en lo que oímos en una grabación [4]. Desde 1900 han existido cinco medios principales de grabación: el disco de laca grabado acústicamente (1900-1925); el disco de laca grabado eléctricamente (1925-1954); el LP de vinilo monofónico (1950-*ca.*1960); el LP de vinilo estereofónico (1958-*ca.*1985), y la grabación digital moderna que comenzó con el CD a principios de los años 1980. A los pianistas también les interesará el rollo o cilindro de piano, que tuvo una historia breve pero importante desde su invención en 1904 hasta su desaparición alrededor de 1930. Cabe señalar que el disco de laca de 78rpm existió durante más de cincuenta años, pero en su primera etapa, anterior a 1925, era completamente diferente de su versión posterior. El acontecimiento más importante de la historia de la grabación fue probablemente la introducción de la grabación eléctrica en 1925.

En líneas generales, cuanto más antigua es una grabación, mayor es la distorsión que presenta debido al procedimiento de grabación, y menor es la manipulación deliberada por parte del productor. Los primeros discos de 78 rpm no permitían ningún tipo de edición, ya que una vez que se hacía el «máster» o disco maestro de cinc, había que enviarlo a la planta de procesamiento y prensar un disco para poder escuchar la grabación. Pero los mecanismos de grabación y reproducción tenían sus propias limitaciones, incluyendo una respuesta irregular a lo largo del espectro grabado y una gama de frecuencias y duración extremadamente limitadas.

La introducción de la grabación eléctrica en 1925 tuvo varias consecuencias inmediatas. En primer lugar, la gama de frecuencias fue ampliada desde un límite superior de 3 kHz (véase la Figura 14.3, analizada más adelante) hasta aproximadamente 5 kHz, que era suficiente para la mayor parte de la información acústica contenida en las principales obras de música clásica. En

segundo lugar, permitió la grabación de agrupaciones grandes con un balance que se acercaba bastante al real, aunque durante muchos años los micrófonos impondrían sus propias características al sonido grabado. A medida que evolucionaba la grabación eléctrica, se comenzaron a explotar las posibilidades de balancear artificialmente la grabación, un proceso en el que destacó especialmente Leopold Stokowski, quien realizó todo tipo de experimentos con la colocación de la orquesta y el control del balance en busca de un sonido ideal [5].

Para los años 1930, se utilizaban varios platos de grabación para poder grabar una interpretación continua sin tener que parar a los músicos cada pocos minutos, y para 1940 el problema del tiempo se había resuelto finalmente empleando una cinta para los másters [6]. La longitud de la cinta era teóricamente ilimitada y se podía cortar y editar de maneras que eran imposibles con el disco de 78 rpm. A partir de esa época comenzó la era de la grabación manufacturada en estudio. Desde los años 1960, la edición y la grabación se han vuelto tan sofisticadas que actualmente hay poca garantía de que una grabación sea el resultado de una única interpretación, exceptuando las grabaciones realizadas «en vivo». John Culshaw, productor del sello discográfico Decca y responsable del primer ciclo completo de «El Anillo del Nibelungo» (1959-1965), aprovechó las posibilidades de la grabación estereofónica para crear interpretaciones que hubieran sido imposibles con cualquier otro medio. Reuniendo a las mejores voces y orquestas del mundo (por ejemplo, la Filarmónica de Viena para su versión de «El Anillo»), logró efectos auditivos que no se habrían podido obtener en el teatro [7]. El problema, claro está, es que las interpretaciones en vivo de «El Anillo» tienden a palidecer en comparación con la extraordinaria presencia de la grabación de Culshaw.

Las grabaciones posteriores a la guerra muestran varios ideales diferentes del sonido grabado. Durante una época de los años 1960 hubo predilección por una acústica «cerrada» y muy seca; después se puso de moda una reverberación molesta y artificial. En ambos casos, los niveles dinámicos y las cualidades tímbricas de la interpretación original eran manipulados casi con seguridad durante el proceso de grabación. Más recientemente, algunos productores han procurado un sonido natural, utilizando un solo par de micrófonos colocados a una distancia apropiada del conjunto. Sólo en este tipo de grabaciones se puede analizar sensatamente el balance de la interpretación en lugar del de la grabación.

Un problema importante que se nos presenta al estudiar cualquier grabación es que el sello puede nombrar en sus carátulas a intérpretes diferentes de los que realmente grabaron. Existe un rollo de piano realizado supuestamente por Ravel tocando sus «Miroirs» y otras piezas, del que hoy en día se sabe

que fue grabado por Robert Casadesus [8]. Un ejemplo más trivial, de principios de los años 1950, es el doblaje de los Do agudos de Kirsten Flagstad en una grabación de EMI del Segundo Acto de «Tristán e Isolda» con Wilhelm Fürtwangler. Esos Do agudos fueron cantados en realidad por Elizabeth Schwarzkopf[9].

### *Instrumentos y técnica*

Una de las características más interesantes de las grabaciones antiguas es el sonido de los instrumentos que se usaban entonces y la manera en que se tocaban. Por ejemplo, las flautas de pico, de sonido dulce y nasal, fueron usadas ampliamente por los intérpretes ingleses hasta finales de los años 1950, pero cayeron en desuso en las décadas siguientes. Actualmente están retornando a nuestras orquestas, sin duda debido al auge de las interpretaciones históricas, pero también, quizá, debido a la difusión de las remasterizaciones en CD de grabaciones realizadas en LP. El sonido característico de las maderas francesas se puede escuchar en innumerables grabaciones antiguas, como los maravillosos discos que grabó Roger Désornière de *Pelléas et Mélisande* de Debussy en París en 1941 [10]. Las sonoridades son muy diferentes de las de los instrumentos actuales, pero conmovedoramente hermosas; el oboe, asociado a Mélisande a lo largo de la ópera, es de una delicadeza particular que parece totalmente adecuada a ese papel. En las grabaciones anteriores a la guerra aún se pueden escuchar trombones de taladro estrecho y violines con cuerdas de tripa. Quizá algún día será habitual que los principales intérpretes elijan sus instrumentos de acuerdo con el carácter de la composición, como hacen hoy en día muchos trompetistas; o tal vez las orquestas vuelvan a encontrar una voz característica retornando a los instrumentos de fabricación local.

Las grabaciones antiguas de piano son tal vez las que resultan más chocantes para los gustos contemporáneos. La mayoría de los pianistas anteriores a aproximadamente 1925 parecen caprichosos por la libertad con que se toman la notación de los compositores; sin embargo, resulta fascinante escuchar, por ejemplo, las múltiples maneras en que se puede arpegiar un acorde [11], u observar la naturaleza improvisatoria de gran parte de su ejecución, destacando las voces intermedias o añadiendo cadencias no indicadas en la partitura [12]. Resulta fácil condenar tales prácticas, pero es mucho más interesante intentar comprenderlas e incluso emularlas. La regla moderna según la cual debemos tocar únicamente lo que está escrito en la partitura no viene al caso cuando la notación del compositor no pretendía ser tan restrictiva. Naturalmente, eso no quiere decir que el intérprete actual pueda hacer caso omiso de las indicaciones del compositor sin correr ningún riesgo.

### *Estudio de caso 1: tempo y duración*

El tempo plantea una serie de problemas especiales, particularmente al considerar las grabaciones antiguas. Se ha afirmado a menudo que, debido a la duración limitada del disco de 78rpm, no se pueden establecer conclusiones categóricas sobre el tempo original de las grabaciones anteriores a aproximadamente 1950. Sin embargo, hay muy pocas pruebas de que los intérpretes realmente aceleraran sus interpretaciones. ¿Qué es, entonces, lo que podemos deducir con certeza sobre el tempo partiendo del testimonio de las grabaciones?

En primer lugar, debemos diferenciar las grabaciones realizadas antes de aproximadamente 1922, cuyo límite de duración era más o menos de tres minutos, de los discos de 78rpm posteriores, que duraban bastante más de cuatro minutos. La mayoría de las grabaciones clásicas realizadas en la época acústica son de arias operísticas, y presentan las siguientes tendencias generales:

(1) Las piezas fueron reducidas sistemáticamente a dos minutos y medio, lo cual sugiere que los productores preferían sacrificar secciones del aria en lugar de apresurar a los intérpretes.
(2) Los tempos no son obviamente rápidos (salvo ciertos momentos muy expresivos en los que solían deleitarse menos que hoy en día), y las secciones de recitativo y parlando suelen tomarse en tempos que hoy consideramos naturales.
(3) Según los parámetros modernos, los *tenutos* de las cadencias principales parecen excesiva e innecesariamente exagerados, lo que indica que no existía una preocupación apremiante por el reloj.

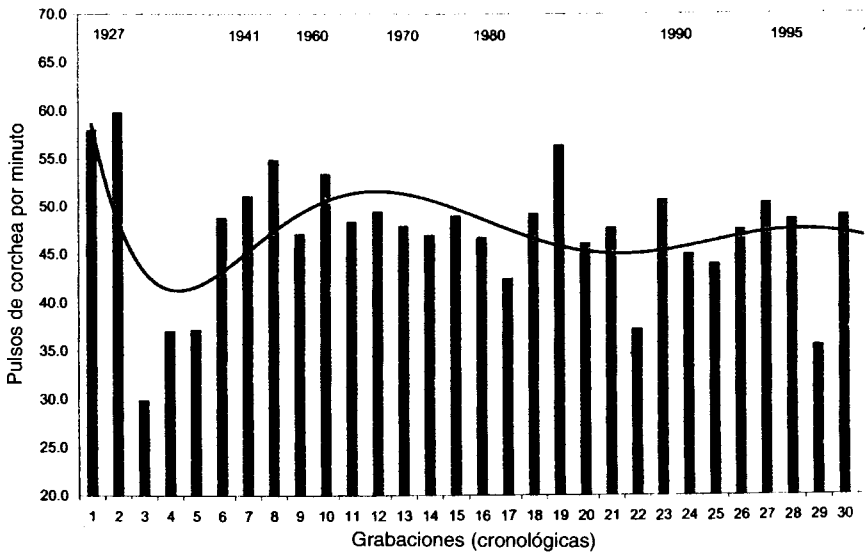
Cuando escuchamos las grabaciones de Caruso o del Otello original de Verdi, Francesco Tamagno, cantando una selección de la ópera en 1903, percibimos interpretaciones con un maravilloso tempo dramático, sin ningún indicio, al menos para el oído moderno, de constreñimiento ni premura [13].

Existen testimonios de músicos que mencionan la incomodidad de los estudios de grabación antiguos y las limitaciones de las «tomas» de tres minutos, pero generalmente datan de épocas muy posteriores y no invalidan los testimonios de las grabaciones en sí. En un análisis provechoso de esta cuestión, John Steane compara las duraciones de grabaciones anteriores a la guerra con interpretaciones en vivo realizadas en la misma época por los mismos artistas y descubre que las grabaciones solían durar menos [14]. Sin embargo, los asuntos de escena generalmente requieren más tiempo en una interpretación en vivo, y su argumento no es de ningún modo conclusivo.

Probablemente nunca sabremos si ciertos tempos rápidos se debieron a la limitada longitud de una cara del disco, pero sabemos que muchos no. Robert Phillip cita varios ejemplos en los que la longitud no pudo haber influido en el tempo [15]. A esos ejemplos yo añadiría el caso de dos grabaciones antiguas del movimiento lento del último cuarteto de Beethoven, op. 135, que son, respectivamente, casi la más rápida y sin duda la más lenta que se han grabado hasta ahora. Beethoven indicó el movimiento «Lento assai, cantante e tranquilo» (véase el Ejemplo 14.1, más adelante). El Cuarteto Flonzaley, que lo grabó en 1927, interpreta esta indicación como un tempo fluido de ♪ = 58, asumiendo, en efecto, un pulso de negra con puntillo. Con un contraste asombroso, el Cuarteto Busch comienza en un tempo de ♪ = 32. La grabación de 1934 del Cuarteto Busch ocupa tres caras en lugar de las dos habituales, pero la versión del Flonzaley ocupa tan sólo una cara y media, dejando el resto de la segunda en blanco. En otras palabras, tenían todo el tiempo necesario para extenderse en el movimiento y sin embargo se apresuran hacia la cadencia final [16].

José Bowen ha analizado minuciosamente el tempo de las grabaciones de una selección de obras sinfónicas, únicamente para concluir que el tempo ha cambiado muy poco a lo largo de ochenta años [17]. De hecho, el método de Bowen no muestra los cambios que pueden haber ocurrido durante períodos más breves. La Figura 14.1 muestra los tempos de los compases 4-6 de treinta y cinco grabaciones del Lento assai de Beethoven, y demuestra que aunque los enfoques del tempo han seguido ciertas tendencias en el transcurso de los años, difícilmente se pueden resumir en una simple fórmula [18]. Además, las grabaciones que rompen la tendencia son a veces las más interesantes.

Se pueden realizar estudios muy útiles sobre el tempo utilizando únicamente un instrumento tan sencillo como el cronómetro. El análisis acústico formal revela un mundo fascinante de *rubatos* y cambios de tempo, incluso en interpretaciones recientes que parecen mantener un pulso estricto. La Figura 14.2 representa gráficamente el tempo de cada corchea del Ejemplo 14.1 tal y como lo tocaron los Cuartetos Flonzaley y Lindsays [19]: cuanto más alta es la barra, más breve es el pulso (y, por tanto, más rápido el tempo). El gráfico muestra que el *subito piano,* implícito en los compases 7 y 8, es articulado por ambos conjuntos mediante una marcada prolongación del tercer tiempo, mientras que el Cuarteto Flonzaley acelera el cuarto tiempo del compás 9, un manierismo que hoy en día nos puede sonar descuidado o caprichoso; sin embargo, la sintaxis musical lo justifica, ya que el acorde de tónica en posición fundamental no llega a cerrar su disposición. Curiosamente, los músicos del Cuarteto Flonzaley articulan el cambio de velocidad del tema de Beethoven entre los compases 3-6 y 7-9 mucho más explícitamente que los integrantes del Lindsays, quienes mantienen una calma un tanto ominosa a lo largo de toda la sección.

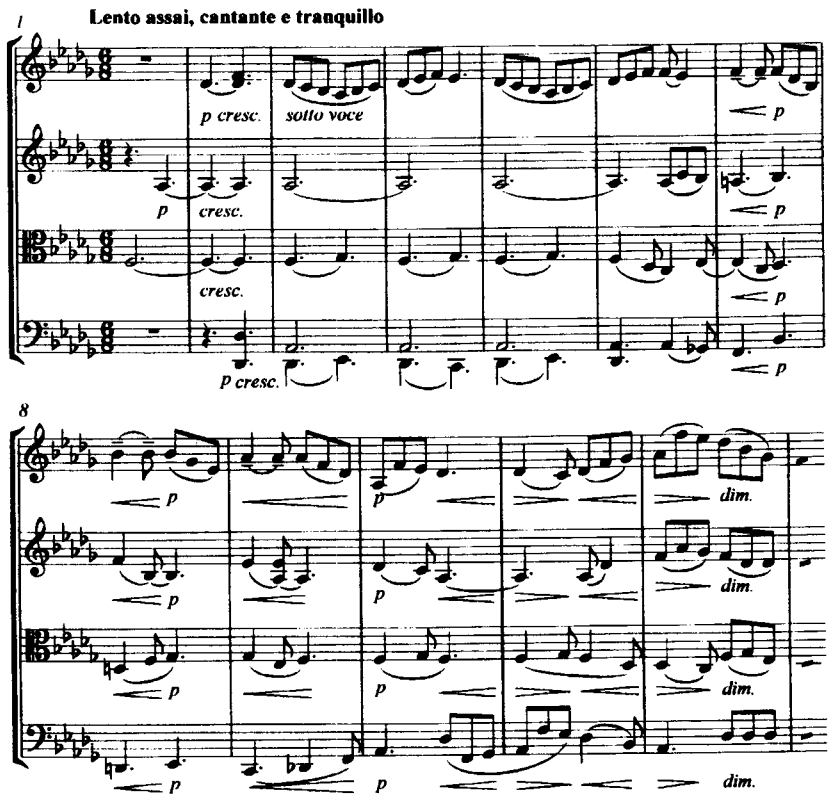



**Cuartetos**

1 Flonzaley (1927)
2 Léner (1927)
3 Busch (1934)
4 Toscanini (NBCO, 1938)
5 Budapest (1941)
6 Loewenguth (1947)
7 Hungarian (1935)
8 Hollywood (1957)
9 Fine Arts (*ca.* 1960)
10 Janáček (1963)
11 Italiano (1968)
12 Julliard (1969)
13 Amadeus (1970)
14 Végh (1973)
15 La Salle (1976)
16 Talich (1977)
17 Amadeus (1982)
18 Melos (1984)
19 Oxford (1985)
20 Vermeer (1985)
21 Guarneri (1987)
22 Lindsays (1987)
23 Emerson (1988)
24 Alban Berg (en vivo, 1989)
25 Hagen (1990)
26 Medici (1991)
27 New Budapest (1991)
28 Tokyo (1991)
29 Bernstein (VPO, 1992)
30 Cleveland (1995)
31 Julliard (1996)
32 Vanbrugh (1996)
33 Alexander (1997)
34 Leipziger (1998)
35 Mosaïques (en vivo, 1998)

FIGURA 14.1 *Tempos iniciales de treinta y cinco grabaciones del Cuarteto de cuerda en Fa mayor op. 135 de Beethoven, tercer movimiento (Lento assai), compases 4-6, en orden cronológico; la línea ondulada traza las tendencias aproximadas.*

EJEMPLO 14.1 *Beethoven, Cuarteto de cuerda en Fa mayor, op. 135, tercer movimiento (Lento assai), compases 1-12.*




Una pregunta que se ha planteado muy pocas veces es si las interpretaciones grabadas en LP o CD se han visto afectadas por cuestiones de tiempo: el escéptico puede suponer que los productores apresuran a sus intérpretes para que la música quepa en una cara del disco o en un CD. No tengo ninguna prueba de ello, pero, por ejemplo, la grabación de *Drumming* de Steve Reich muestra cómo una grabación en CD puede afectar directamente a la composición misma. La duración original de la obra es de unas tres horas; sin embargo, se redujo a 58 minutos para la grabación en CD realizada por la agrupación del propio compositor [20]. Aunque la decisión de grabarla en un solo CD haya sido motivada por razones económicas o artísticas, el resultado es una experiencia musical completamente diferente.

### *Estudio de caso 2: vibrato*

Muchos aspectos de la técnica de ejecución permanecen prácticamente inalterados por las técnicas de grabación, y uno de ellos es el vibrato [21]. Tanto en los instrumentos como en las voces, el vibrato se logra mediante una combinación de fluctuaciones periódicas de amplitud, frecuencia y timbre. Las grabaciones vocales más antiguas revelan que el estilo moderno de vibrato continuo ya era utilizado por los cantantes italianos durante el cambio de siglo, pero muchos cantantes noreuropeos y americanos siguieron empleando un vibrato intermitente o de muy poca amplitud hasta los años 1920. Sin embargo, a principios de 1900, ni siquiera los italianos empleaban las ondula-

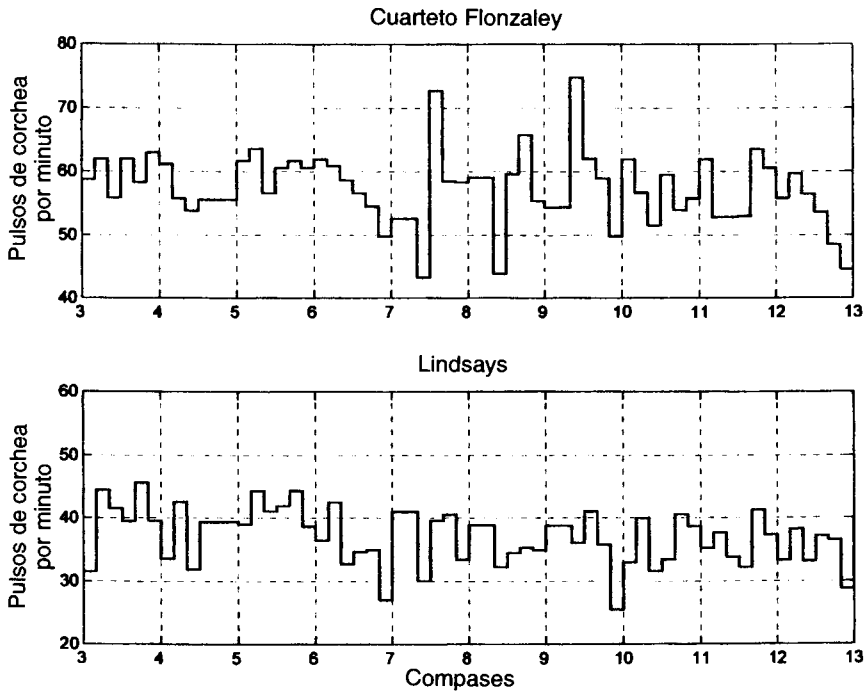
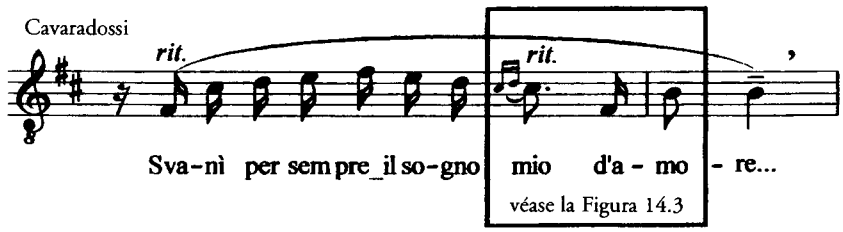



FIGURA 14.2 *Tempos del Cuarteto de cuerda en Fa mayor, op. 135 de Beethoven, tercer movimiento (Lento assai), compases 3-12, grabados por el Cuarteto Flonzaley (1927) y The Lindsays (1987), medidos en pulsos de corchea por minuto (i.e., valores metronómicos); cuanto más alta es la barra, más breve es el pulso.*

EJEMPLO 14.2 *Puccini, «E lucevan le stelle»,* Tosca, *Acto III, compases 20-21 (basado en la partitura vocal de Ricordi, p. 290).*




ciones amplias y lentas que se escuchan habitualmente en las salas de ópera hoy en día. La voz de Caruso se caracteriza por un vibrato rápido y continuo y por un empleo controlado pero notable del portamento, como muestra la Figura 14.3. Dicha Figura muestra el do♯[1], el fa♯ y el si del famoso pasaje de *Tosca* reproducido en el Ejemplo 14.2. Cada línea ondulada corresponde a un armónico de la nota cantada [22].

La Figura 14.3 muestra una enorme cantidad de detalles. Permite medir con mucha precisión la duración de cada nota, su timbre (los armónicos fuertes y ausentes), los modos de ataque y caída y el portamento. El emotivo declive del do♯[1] se puede ver claramente en el gráfico y se puede medir hasta aproximadamente una tercera mayor. La energía principal del do♯[1] está centrada en el segundo armónico, do♯[2] (*ca.* 550 Hz), mientras que la fundamental, do♯[1], está prácticamente ausente [23]. La Figura 14.3 también muestra la formante del cantante, es decir, el «zumbido» que se escucha en las voces masculinas bien adiestradas. Éste se encuentra una banda de frecuencia de aproximadamente 2.600 Hz (≈ re♯[4]), en la que las ondas largas y solapadas, causadas por el vibrato, garantizan que la formante forme una banda de sonido agudo casi continuo. Por último, contando los picos de las líneas onduladas y midiendo los puntos de comienzo y final, resulta fácil determinar que la frecuencia del vibrato es de aproximadamente 7,6 ciclos por segundo (cps), tal y como se indica entre las dos flechas [24].

Las primeras grabaciones de violinistas llegaron justo a tiempo para captar el final de una tradición interpretativa que cambiaría dramáticamente entre 1910 y 1930. Las únicas grabaciones de Joachim, realizadas en 1904, revelan un empleo selectivo y muy poco expresivo del vibrato, que no forma parte del tono básico [25]. En cuanto a los vientos de madera, las grabaciones demuestran que el vibrato no fue empleado antes de aproximadamente 1930, excepto por algunos flautistas de la tradición francesa. Léon Goossens

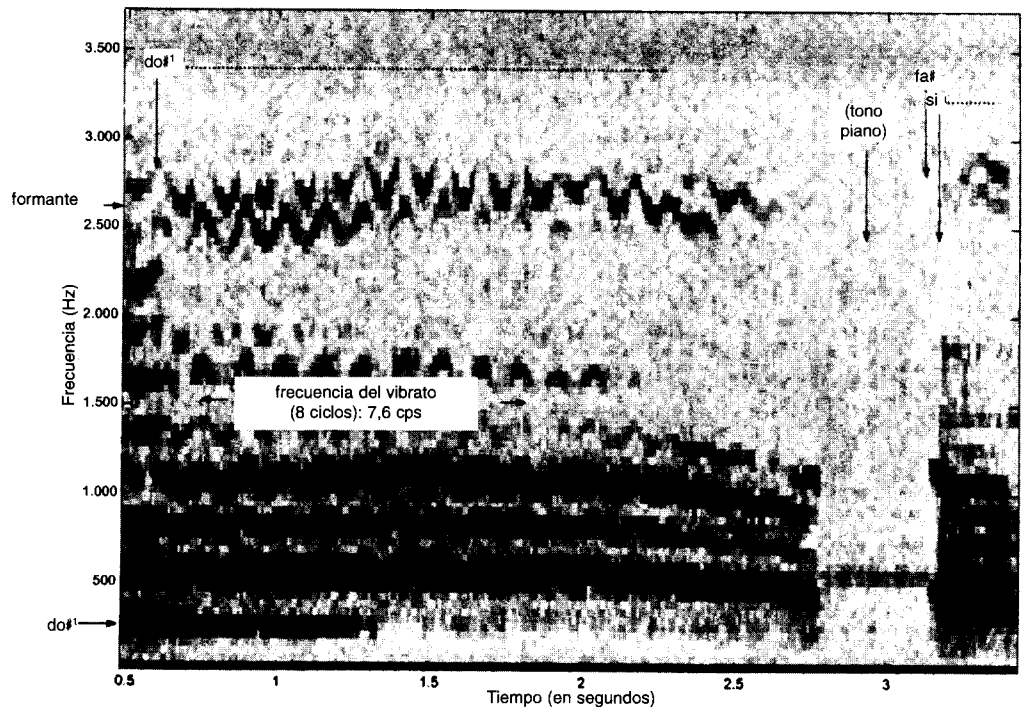



FIGURA 14.3 *Espectograma del pasaje marcado dentro del recuadro del Ejemplo 14.2. Las líneas onduladas muestran cada armónico del tono vocal; las líneas rectas, apenas visibles, representan los tonos piano. El tono se corta en el décimo armónico, ca. 2.770 Hz, que es mi♯⁴, tres octavas por encima del mi central.*

describió detalladamente sus intentos por introducir el vibrato durante los años 1920. Aunque muchos intérpretes lo insultaron y se burlaron de él, sus esfuerzos fueron recompensados por Thomas Beecham, quien lo eligió como primer oboe para la Orquesta Filarmónica de Londres en 1932 [26]. El primer clarinete de la misma orquesta, Reginald Kell, no tardó en seguir su ejemplo, y los efectos de esa revolución se pueden escuchar en varias grabaciones anteriores a la guerra. Sin embargo, algunas de la grabaciones orquestales más antiguas, como la notable versión de la Quinta Sinfonía de Beethoven que realizó Artur Nikisch al frente de la Filarmónica de Berlín en 1913, demuestran lo que puede lograr una buena sección de maderas sin utilizar el vibrato [27]. Muchas grabaciones, incluso de los años 1950, demuestran que se empleaba mucho menos vibrato del que sería habitual en las décadas siguientes.

El empleo del vibrato en otros instrumentos, como la trompeta o el fagot, es igualmente fascinante. De especial interés es la historia del vibrato en la trompa. Durante la primera mitad del siglo, las orquestas europeas orientales lo empleaban habitual y muy expresivamente. Las grabaciones de orquestas checas y soviéticas proporcionan ejemplos impresionantes para los oyentes curiosos, y para los ejecutantes de trompa muy valientes [28].

Aunque generalmente el proceso de grabación no altera el vibrato, el productor de la grabación puede intervenir indudablemente en el resultado final. Elizabeth Schwarzkopf señaló que su marido, Walter Legge (el más veterano de los productores de la posguerra, famoso por exigir los niveles más altos a sus intérpretes), no dudaba en quejarse cuando escuchaba el odiado «temblor» [29]. Legge produjo la mayoría de las grabaciones de EMI/HMV de los años 1950, incluyendo muchos LPs maravillosos de la Philarmonia, una orquesta que fundó específicamente para grabar.

## *La interpretación de las grabaciones*

Las grabaciones revelan claramente una gran cantidad de información sobre las diversas prácticas interpretativas del siglo XX, que en este capítulo apenas he tratado superficialmente. La pregunta básica es cómo se combinan los factores como la entonación, el vibrato, el timbre, el balance, la sincronización y el modo de ataque, la acentuación, la ornamentación [30] y, no menos importante, el control estructural, para producir la experiencia auditiva holística que ofrece la grabación misma. En este sentido, las mejores reseñas de grabaciones siguen siendo las de ciertos críticos, que han desarrollado un lenguaje sofisticado para abordar la experiencia auditiva. Mucha labor crítica se desperdicia en el juego pueril de descubrir la grabación «definitiva» o la «compra maestra». En efecto, comparar grabaciones es un método excelente para des-

cubrir y celebrar la maravillosa diversidad de interpretaciones y artistas del archivo de grabaciones. Y si el lenguaje de los críticos es demasiado impreciso para los musicólogos, éste se puede complementar y apoyar con el análisis acústico, como he intentado demostrar. Sin embargo, no hay música hasta que se interpretan los sonidos, y el lenguaje del crítico proporciona al menos una base para analizar por qué una grabación determinada es musicalmente tan conmovedora, emocionante, mágica o simplemente aburrida [31].

Existen muchas grabaciones en las que los compositores han participado en la producción o la interpretación. También se da el caso de compositores que han grabado la misma obra más de una vez, y esas grabaciones nos permiten hacer comparaciones muy interesantes. Las inevitables diferencias que existen entre ellas sugieren que una única interpretación o grabación «aprobada» por el compositor nunca debe considerarse como definitiva. Las dos grabaciones del Concierto para violín de Elgar dirigidas por el compositor constituyen un ejemplo. La primera, grabada en 1916, es una versión muy resumida que, aunque fue reorquestada por el mismo Elgar, apenas ofrece al oyente moderno una experiencia satisfactoria de la obra en sí misma. Sin embargo, la interpretación del violín es fascinante. La solista es Marie Hall, una violinista inglesa muy respetada a ambos lados del Atlántico, cuya grabación revela una sonoridad exquisita y una notable sensibilidad en el fraseo. Resulta interesante comparar su interpretación con la famosa grabación realizada por Yehudi Menuhin en 1932 [32]. La técnica de Menuhin es esencialmente moderna: el sonido «hilado» se debe a un vibrato compacto y continuo; la ejecución es audaz y virtuosística; y cada frase es expresada de una manera retórica que quizá delate la juventud de Menuhin; sin embargo, no es sustancialmente diferente del enfoque de un violinista moderno.

El enfoque más lírico de Hall quizá se acerca más al espíritu de la partitura de Elgar. Su ejecución está impregnada de ese patetismo y sensación de pérdida que parecen tan propios de la música de Elgar y de gran parte del arte y la literatura inglesas de principios del siglo XX. Y aunque Elgar manifestó estar encantado con la interpretación de Menuhin, no pudo haber imaginado una ejecución similar cuando compuso la obra, en 1910, ya que el estilo moderno aún no existía. Su ideal de ejecución violinística, e incluso de la obra misma, debe por tanto haber cambiado a medida que se desarrollaron los estilos interpretativos en los años 1920. Aún así, Hall demuestra que tales cambios no son necesariamente beneficiosos, y que la interpretación del concierto de Elgar no ha de ser necesariamente la demostración de poderío que suele ser hoy en día.

## *La influencia más amplia de las grabaciones*

La influencia de la grabación en nuestras prácticas musicales más amplias apenas está comenzando a ser reconocida [33]. ¿Se ha llegado a concebir la música como algo más abstracto, incorpóreo e ideal debido a la invisibilidad de sus intérpretes en las grabaciones?, ¿o es que los métodos de grabación simplemente han respondido a las ideologías actuales sobre lo que debe ser la música? Una cosa está clara: debemos rechazar la opción simplista de suponer que una determinada grabación representa «la obra misma» de manera autorizada. Las obras posteriores a la guerra son particularmente vulnerables en ese sentido. Cuando se conocen por una sola grabación y rara vez se tocan en vivo, resulta fácil suponer que las características, e incluso los errores, de la interpretación grabada representan lo que es la obra, que la grabación refleja «cómo suena la obra». Las grabaciones múltiples nos permiten explorar la obra como un objeto de muchas facetas, o incluso como algo no determinado materialmente hasta que es leído o interpretado [34].

Hindemith afirmó que toda interpretación es una corrupción de la obra, pero las grabaciones demuestran la importancia de la voz del intérprete como complemento de la del compositor. Es a través del personaje del intérprete —y quizá también del productor— como «la obra misma» cobra vida y adquiere significados musicales particulares. Cada grabación es una creación artística única, lograda mediante una síntesis de la composición, la interpretación y determinados métodos de grabación. No es de extrañar, pues, que la música grabada haya desempeñado un papel tan importante en nuestra cultura musical durante los últimos cien años. Con la llegada de nuevos formatos como el muestreo digital y la transmisión por internet, podemos esperar un futuro igualmente emocionante e influyente para la grabación de la música clásica en las décadas venideras.

## *Notas*

1. Véase Nicholas Cook, *Music: A Very Short Introduction,* Oxford, Oxford University Press, 1998, p. 13.
2. Véase Edward T. Cone, *The Composer's Voice,* Berkeley, University of California Press, 1974. Para la interpretación «como un tipo de actuación», véase Joel Lester, «Performance and analysis: interaction and interpretation», en John Rink (ed.), *The Practice of Performance: Studies in Musical Interpretation,* Cambridge, Cambridge University Press, 1995, p. 213; véanse también los capítulos 7 y 11 de este libro.
3. Para dos ejemplos recientes, véase Ronald Woodley, «Strategies of irony in Prokofiev's Violin Sonata in F minor Op. 80», en John Rink (ed.), *The Practice of Performance: Studies in Musical Interpretation,* cit., p. 170, y Peter Johnson, «Performance and the listening experience:

Bach's «Erbarme dich», en Nicholas Cook, Peter Johnson y Hans Zender (eds.), *Theory into Practice: Composition, Performance and the Listening Experience,* Leuven, Leuven University Press, 1999, pp. 55-101.

4. Para una historia general de la grabación, véase Michael Chanan, *Repeated Takes: A Short History of Recording and its Effects on Music,* Londres, Verso, 1995.
5. Véase Ronald Gelatt, *The Fabulous Phonograph: 1877-1977,* Londres, Collier Macmillan, 1977, p. 236.
6. Véase Joseph Horowitz, *Understanding Toscanini: A Social History of American Social Life,* Londres, Faber and Faber, 1987, para un excelente análisis crítico de las cintas de tres óperas completas grabadas en Salzburgo en 1937.
7. Para un análisis, véase John Culshaw, *Ring Resounding: The Recording of «Der Ring des Nibelungen»,* Londres, Secker & Warburg, 1967.
8. Véanse las notas editoriales de Roger Nichols en la edición Urtext de Peters de *Miroirs* (1991).
9. John Culshaw, *Ring Resounding: The Recording of «Der Ring des Nibelungen», op. cit.*, p. 55.
10. Remasterizado en el CD de EMI CHS 7 61038 2.
11. Para un análisis técnico de la asincronía, véase Carl E. Seashore, *Psychology of Music,* Nueva York, Dover [1938], 1967, pp. 248-253.
12. La colección *Great Pianists of the 20th Century* incluye grabaciones de los pianistas antiguos más importantes. Las grabaciones de Rachmaninoff están editadas en RCA 09026 61265.
13. Las grabaciones completas de Tamagno se encuentran en la colección Pearl Opal, CD 9846.
14. John Steane, *The Grand Tradition: Seventy Years of Singing on Record, 1900-1970,* 2ª ed., Londres, Duckworth, 1993, p. 10.
15. Robert Philip, *Early Recordings and Musical Style: Changing Tastes in Instrumental Performance, 1900-1950,* Cambridge, Cambridge University Press, 1992, pp. 35-36.
16. La grabación del Cuarteto Flonzaley se encuentra en el sello Biddulph LAB 089-90; y la del Cuarteto Busch, en EMI CHS 65306-2.
17. José A. Bowen, «Finding the music in musicology: performance history and musical works», en Nicholas Cook y Mark Everist (eds.), *Rethinking Music,* Oxford, Oxford University Press, 1999, p. 435.
18. La línea ondulada de la Figura 14.1 (un polinomio de sexto orden) traza las tendencias aproximadas de períodos más breves.
19. El nombre oficial de este cuarteto es «The Lindsays». Los tempos fueron tomados de ASV, CD DCS 403.
20. Nonesuch, CD 979 170-2.
21. Los términos «vibrato» y «tremolo» se han utilizado para referir aspectos de lo que yo llamo generalmente vibrato: véase Carl E. Seashore, *Psychology of Music,* cit., p. 33.
22. La Figura 14.3 fue creada a partir de la remasterización en CD del disco *Enrico Caruso, The Puccini Recordings (1902-1916),* Vocal Archives, VA 1216, pista 10, 1' 46".
23. Los tests psicoacústicos demuestran que no es necesario que la fundamental esté presente. Para un caso similar en un tono de violín, véase Carl E. Seashore, *Psychology of Music,* cit., p. 98. En su ensayo «Performance and the listening experience: Bach's *Erbarme dich,* cit., Peter Johnson ofrece un análisis más detallado de mis métodos.
24. Exceptuando a algunas soubrettes, son pocas las voces operísticas actuales que tienen un vibrato más rápido que 5 cps; las voces más maduras suelen oscilar a una frecuencia de 2 o 3 cps. Las conclusiones sobre el vibrato que ofrece John Steane en su excelente análisis crítico del canto del siglo XX se pueden verificar mediante los métodos que he ilustrado; por ejemplo, véase John Steane, *The Grand Tradition: Seventy Years of Singing on Record, 1900-1970,* cit., p. 135.

25. El primer volumen de la colección de Pearl *Great Virtuosi of the Golden Age* (GEMM CD 9101) incluye piezas grabadas por Joachim, Ysaÿe, Elman y Sarasate. Nacido en 1831, Joachim tenía setenta y dos años cuando realizó esa grabación.
26. Léon Goossens y Edwin Roxburgh, *Yehudi Menuhin Music Guides: Oboe,* 2ª ed., Londres, Macdonald, 1980, p. 87.
27. La grabación se encuentra remasterizada en DG 453 804-2GCB6. El segundo movimiento es particularmente interesante en este respecto.
28. Un ejemplo hermoso de la trompa orquestal checa es el movimiento lento del Concierto para violonchelo de Dvotak grabado en 1938 por Pau Casals y George Szell, remasterizado en EMI CDH 7 63498 2. Se dice que esta grabación era una favorita de Fred Gaisberg, principal productor de HMV desde las primeras grabaciones de Caruso de 1902 hasta su retiro en 1939.
29. Elizabeth Schwarzkopf, *On and Off the Record: A Memoir of Walter Legge,* Londres, Faber and Faber, 1982, p. 19.
30. Para un estudio de la ornamentación en la grabación, véase Will Crutchfield «Vocal ornamentation in Verdi: the phonographic evidence», *19th Century Music, 7* (1983), pp. 3-54.
31. Analizo esta cuestión en Peter Johnson, «Performance and the listening experience: Bach's *Erbarme dich,* cit., pp. 97 y ss.
32. La grabación de Hall se encuentra en Pearl GEMM CDS 9951/5; la de Menuhin, en EMI CDC 5 55221-2.
33. Dos libros en los que se comienza a lidiar con esta cuestión son Michael Chanan, *Repeated Takes: A Short History of Recording and its Effects on Music,* cit.; y Evan Eisenberg, *The Recording Angel: Music, Records and Culture from Aristotle to Frank Zappa,* Londres, Picador, 1987.
34. Véase Joel Lester, «Performance and analysis: interaction and interpretation», cit., p. 214.

## *Lecturas complementarias*

Chanan, Michael: *Repeated Takes: A Short History of Recording and its Effects on Music,* Londres, Verso, 1995.

Cook, Nicholas, Peter Johnson y Hans Zender (eds.): *Theory into Practice: Composition, Performance and the Listening Experience,* Leuven, Leuven University Press, 1999.

Day, Timothy: *A Century of Recorded Music,* New Haven, Yale University Press, 2000.

Eisenberg, Evan: *The Recording Angel: Music, Records and Culture from Aristotle to Frank Zappa,* Londres, Picador, 1987.

Horowitz, Joseph: *Understanding Toscanini: A Social History of American Social Life,* Londres, Faber and Faber, 1987.

Philip, Robert: *Early Recordings and Musical Style: Changing Tastes in Instrumental Performance, 1900-1950,* Cambridge, Cambridge University Press, 1992.

# 15 *La crítica de la interpretación musical*

RAYMOND MONELLE

Considerada como un tipo de escritura formal, la reseña de concierto es algo relativamente nuevo. Hasta hace poco se tenía un prejuicio contra ella; no estaba a la altura del interés de las personas sofisticadas. Uno podía escribir sobre teoría musical, sobre estilos, modas e influencias, sobre composiciones nuevas. Pero criticar conciertos e intérpretes era un tanto degradante y furtivo. Incluso un crítico distinguido como William J. Henderson afirmó: «Estamos confrontados con la demanda de artistas interpretativos. Cualquiera que ubique la crítica en un nivel alto querrá decir muy poco sobre ellos. Considerar al intérprete es el objetivo menos importante de la verdadera crítica» [1]. El gran especialista en música Ernest Newman comentó en una entrevista para la prensa, en 1931, que las interpretaciones repetidas de obras famosas dejaban a los críticos sin nada que hacer:

> Aunque proporcionan muchos placeres al aficionado común, los Promenade Concerts siguen ofreciendo muy pocos pretextos al crítico para dar a conocer sus opiniones, ya que los programas están compuestos en gran parte por obras muy famosas. La semana pasada hubo una noche Wagner, una noche Beethoven, una noche Brahms y una noche Tchaikowski, y como ninguno de estos compositores ha escrito mucho últimamente, no ha habido obras nuevas para que los críticos pudieran ejercer su infalible juicio [2].

Quizá debido a ese prejuicio, la historia de la crítica de la interpretación es bastante diferente de la historia de la crítica musical general. La mayoría de las historias de la crítica se centran en el juicio de composiciones, especialmente las nuevas, y las antologías estándar de críticos como Eduard Hanslick

y Bernard Shaw naturalmente otorgan más espacio a las obras maestras que eran nuevas en su día. Curiosamente, la descripción de ese elemento familiar de la vida musical, la reseña de concierto, aún estaba por escribirse.

Sin embargo, hoy en día ese tipo de reseña predomina en la crítica musical. ¿Cómo se ha producido ese cambio? Indudablemente, el abrupto declive de la interpretación de música contemporánea ha llevado a los críticos a concentrarse en los artistas intérpretes; en el siglo XVIII, mucha de la música que se escuchaba acababa de salir de la pluma del compositor. Además, a menudo los instrumentistas eran también compositores; interpretaban sus propias obras, y los críticos naturalmente se centraban en las obras y no en los intérpretes. Pero a medida que se consolidaba el canon de las obras clásicas —en gran parte gracias a los críticos musicales—, las múltiples interpretaciones de obras canónicas revelaban la medida en que los intérpretes contribuyen a la experiencia de la música. Al principio (y esto es especialmente evidente en los escritos de Shaw), se podían hacer comentarios, incluso desfavorables, sobre obras que ya habían sido aceptadas dentro del canon, pero en la época moderna, cuando la mayoría de las obras que se interpretan son canónicas, no se espera que los críticos alaben a Haendel, o refunfuñen sobre Brahms como lo hiciera Shaw. Se supone que han de centrarse en el intérprete.

Las primeras críticas musicales publicadas en revistas especializadas, como *Critica musica* (Hamburgo, 1722-1725) de Johann Mattheson o *New eröffnete musikalische Bibliothek* (Leipzig, 1736-1754) de Lorenz Mizler, contienen muy pocas evaluaciones interpretativas. A veces se encuentra alguna evaluación general del estilo de un intérprete, pero rara vez se publicaban reseñas de conciertos u óperas. Sin embargo, hacia finales del siglo XVIII los informes sobre las producciones operísticas podían incluir descripciones perspicaces de los cantantes; por ejemplo, la revista berlinesa *Litteratur-und Theater-Zeitung* de 1781 incluye muchos informes sobre óperas. Ese mismo año, la revista reseñó la visita a Berlín del bajo Ludwig Fischer, el primer Osmin de *Die Entführung aus dem Serail* (1782) de Mozart:

> Herr Fischer, cantante imperial y real, estuvo varios días en la ciudad; actuó dos veces como el mensajero en *Zémire et Azor* [de Grétry], y en una ocasión, entre los actos, cantó dos arias de la gran ópera *Günther von Schwarzburg* de Holzbauer. Los entendidos se maravillaron ante la pureza y extensión de su voz, y todo el público estaba encantado con su exitosa interpretación y la gran habilidad de su canto... [3].

Ésta es una crítica rudimentaria, que no demuestra ni exige un conocimiento sofisticado de la música o el canto.

La primera publicación periódica realmente influyente fue el semanal *Allgemeine musikalische Zeitung (AmZ)* [4] de Leipzig, que apareció por primera

vez en 1798. Su editor, Friedrich Rochlitz, se dio cuenta rápidamente de que los lectores querían algo más que los artículos y reseñas generales de obras nuevas que habían caracterizado a las publicaciones anteriores. Los primeros números contenían a veces «correspondencia» proveniente de los centros musicales alemanes, Berlín, Viena, Dresden, etc., a veces (aunque no siempre) con comentarios sobre conciertos. Para el año 1800, estas correspondencias se habían consolidado en noticias periódicas *(Nachtrichten)* en las que se describía la actualidad concertística y operística de las ciudades alemanas, muchas veces con comentarios detallados sobre los intérpretes, especialmente los cantantes. La siguiente reseña de un concierto que tuvo lugar en 1802 con la actuación del violonchelista Nikolaus Kraft, miembro del Cuarteto Schuppanzigh, demuestra un conocimiento de muchos intérpretes y del instrumento, así como un sólido criterio sobre lo que constituye una buena ejecución:

> Es un violonchelista muy notable; su sonido no es opaco y nasal como el que suele escucharse en este instrumento, sino potente, varonil, amplio (lleno, como dicen los músicos); su ejecución es rápida, segura, precisa, delicada y pura (aparte de algunos detalles en los registros extremos, que no fueron del todo exitosos); su manejo del arco es admirable... [5].

Ésta es una verdadera reseña de concierto, y no un artículo general sobre el estilo de un artista como los que habían sido comunes durante el siglo anterior.

La revista llegó a publicar algunas de las reseñas más detalladas y acertadas de todas las épocas, lo que exigía de sus lectores un conocimiento profesional de los instrumentos, las técnicas y los estilos, como demuestra una reseña del 28 de abril de 1830 sobre un gran violinista, Heinrich Wilhelm Ernst:

> Herr Ernst, un hombre joven todavía, posee todos los recursos que, combinados con la fuerza de su talento obviamente excepcional, pueden conducirle en pocos años al camino que le convertirá en uno de los más grandes artistas vivos, ya que tiene una facilidad inusual para los pasajes brillantes, los saltos, las dobles cuerdas, etc., además de un sonido potente y puro y un control diestro del arco; su ejecución de los Adagios tampoco es despreciable, aunque no del todo elogiable. Sin embargo, es de lamentar que Herr Ernst, sin saberlo, se encuentre en un camino artístico falso, ya que parece ser simplemente un imitador y quiere tocar «a la Paganini» [6].

Ésta no es una reseña para el lector común. La revista esperaba que los lectores tuvieran cierto conocimiento de la técnica instrumental. El verdadero arte del periodismo musical popular aún estaba por descubrirse.

La publicación de comentarios musicales en los diarios, a diferencia de las publicaciones semanales o mensuales, comenzó a principios del siglo XIX,

probablemente con las reseñas de J. C. F. Rellstab (1808-1813) en el *Vossische Zeitung* de Berlín. En Londres y París, la reseña diaria —foco principal de la crítica moderna— surgió a partir de la reseña teatral. De hecho, Julien-Louis Géoffroy (1743-1814), crítico teatral del *Journal des débats* parisino, se jactaba de su ignorancia musical, aunque reseñaba óperas con frecuencia.

Durante los primeros años del siglo XIX, *The Times* publicaba a veces reseñas de las funciones de los principales teatros londinenses, incluyendo óperas. Al principio, éstas eran escritas por un crítico de teatro, Barron Field, o por el director del periódico, T. M. Alsager. Más tarde, en 1843, contrataron a un verdadero crítico musical, Charles Lamb Kenney. Durante un tiempo, sin embargo, no se notó mucha diferencia; las reseñas, que aún aparecían sin firma, seguían haciendo comentarios sobre el público y los vestidos de las damas, y había muy poca crítica musical, aparte de las óperas. Cuando aparecía una reseña de concierto, los comentarios eran breves y generales —como en la siguiente descripción de un concierto del Festival de Birmingham, el 19 de septiembre de 1843, dirigido por William Knyvett, con la contralto Maria Hawes y la mujer de Knyvett, Deborah Travis: «La señora Knyvett cantó el aria "Pious orgies" [de *Judas Maccabaeus* de Haendel] con gran precisión y pureza de estilo. Su interpretación de la música de Haendel sigue siendo tan perfecta como siempre. El hermoso aria "Jehovah crown'd" le fue asignado a la señorita Hawes, quien lo cantó con fuerza pero con una energía tal vez demasiado elaborada en algunos pasajes» [7]. Aún no se consideraba que los lectores de *The Times* estuvieran preparados para comentarios musicales más detallados o sensibles.

En esa misma época tuvo lugar un acontecimiento en París que determinaría el futuro de la reseña de conciertos. En 1872, François-Joseph Fétis, bibliotecario del Conservatorio de París y, más tarde, director del Conservatorio de Bruselas, fundó el semanal *Revue musicale,* una revista seria y responsable que incluía reseñas de conciertos entre sus artículos generales sobre música. La mayor parte de cada número estaba escrita por el infatigable Fétis, quien se tomaba sus labores muy en serio, escribiendo en un estilo pesado e intransigente. El poeta Heinrich Heine, que también escribía críticas musicales para la *Allgemeine musikalische Zeitung,* comentó: «No conozco nada que inspire menos que una crítica del señor Fétis [...] escritas en una jerga particular y salpicadas de términos técnicos que la mayoría de los lectores educados desconocen» [8]. El estilo tedioso de Fétis se puede apreciar en una reseña de concierto del primer número de la revista (febrero 1827):

> El concierto se inició con el primer movimiento de la sinfonía de Haydn que comienza en Re menor seguido de un Allegro en Re mayor [núm. 104]. Una entrada falsa de los fagotes en la introducción estropeó el efecto. La primera parte del alle-

> gro estuvo bien ejecutada, pero, en la segunda, tan enérgica y vigorosa, la orquesta careció de brío y calidez. Varias fluctuaciones de tempo impidieron que los violines tocaran juntos, y a los bajos, por ser muy pocos, les faltó energía en el momento explosivo [¿el tutti del compás 32?].

En oposición a esta respetable publicación, el editor Maurice Schlesinger fundó la *Gazette musicale de Paris* en 1834. Esta rival tuvo tanto éxito que la *Revue musicale* no tardó en fusionarse con ella. Los escritores de la *Gazette,* entre quienes se incluían personajes importantes como Berlioz, Liszt, Wagner, Castil-Blaze y Joseph d'Ortigue, cultivaron un estilo nuevo e ingenioso derivado del *feuilleton* de crítica teatral, que reemplazó el lenguaje técnico con impresiones, burlas, referencias cultas y opiniones personales. La mayoría de los artículos ya aparecían firmados por sus autores. Era verdadero periodismo, pero sus mejores ejemplos estaban reservados a los artículos generales y reseñas de obras nuevas.

En la *Gazette,* las reseñas regulares de conciertos estaban escritas cada vez con mayor frecuencia por el dramaturgo, compositor y violinista Henri Blanchard (quien se convirtió en uno de los editores adjuntos en 1836), en lugar de las celebridades mencionadas anteriormente. La bravura de los famosos *feuilletons* de la *Gazette* no caló realmente en las reseñas elementales de Blanchard; sin embargo, Blanchard desarrolló un auténtico estilo de periodismo musical evocativo, dinámico y desprovisto de tecnicismos. Tal vez el arte de la reseña concertística moderna se desarrolló a partir de las reseñas de Blanchard y su sucesor Adolphe Botte. El 5 de diciembre de 1847, Blanchard describió un concierto en la Salle Hertz en el que se interpretó un cuarteto de Mozart —aparentemente el K. 589:

> La ejecución... fue perfecta. El señor Tilmant [primer violín] hace bien en no tomarse demasiado literalmente el término «música de cámara» que aparentemente sólo se aplica a la música íntima interpretada en un espacio pequeño. Su estilo amplio y orquestal confiere una especie de amplitud sinfónica a un cuarteto que algunos de nuestros violinistas tocan de manera bonita y tacaña. Para que suene apasionado y expresivo —signos característicos de la interpretación musical moderna—, conviene que en ciertas cadencias melódicas haya un tratamiento flexible de esa precisión y esa rigidez de tempo que al final le da una especie de aridez mecánica a la idea musical, despojándola a veces de la inspiración y la poesía que debe de haber animado al compositor cuando la escribió.

La *Gazette musicale* tenía muchos lectores, y desde el principio comenzó a influir en el mundo musical, y a crear o destruir carreras. Citando de nuevo a Heine: «Los artistas más afamados se postran sumisamente a los pies de Schlesinger, arrastrándose y moviendo sus colas, para tener una buena crítica

en su revista» [9]. Uno de los colaboradores era Ludwig Rellstab, hijo del crítico del *Vossische Zeitung*, que sabía bastante sobre las reacciones de los intérpretes ante sus críticas: el artículo que escribió en Berlín en 1826 sobre la cantante Henrietta Sontag le había dado problemas con las autoridades, y fue encarcelado durante tres meses en la fortaleza de Spandau, lo que lo convirtió en el primer mártir de los críticos musicales [10].

También en Londres el poder de la prensa musical se hizo evidente. En 1846, *The Times* contrató a su primer crítico musical profesional, J. W. Davison. Este escritor característico pronto comenzó a adquirir un enorme poder sobre la profesión musical [11]. En sus reseñas, da la impresión de ser un crítico indulgente y generoso, alabando incluso a aquellos artistas cuyos defectos también tenía la obligación de reseñar. De tal manera, su reseña de la representación de *Lucia di Lammermoor* en la Royal Italian Opera del Covent Garden en 1847, con Fanny Tacchinardi-Persiani y Lorenzo Salvi, combina característicamente un gran discernimiento y el deseo de no querer decir nada desfavorable sobre nadie:

> La ejecución extremadamente limpia y delicada de una profusión de ornamentos, la distinción de cada nota en los pasajes más rápidos, la variedad de colorido —la facilidad con la que hace todo— eran las mismas de siempre. El canto de Persiani ha sido comparado una y otra vez con un artista consumado tocando un instrumento delicado [...] En el segundo acto, se pudieron observar quizá algunos de los leves fallos de entonación que solía tener la cantante, por lo demás perfecta, pero el defecto solo fue transitorio [...]
>
> El debú de Salvi en el papel de Edgardo fue tal vez el evento más interesante de la noche, y su éxito debe de haber satisfecho las expectativas de sus admiradores más fervientes. Tiene una voz de tenor pura, de registro amplio, no destacable por su volumen, pero perfectamente uniforme en cada parte y poseedora de la mayor flexibilidad. La maneja con la habilidad de un artista consumado, introduciendo en ella las más hermosas variedades de luz y sombra, y haciéndola elocuente con una expresión dramática. Quizá su mayor habilidad como vocalista se demuestra en sus pasajes *piano*, con una voz completamente atenuada sin detrimento del valor de las notas [12].

El estilo y las virtudes de ambos cantantes están maravillosamente descritos, aunque quizá de manera poco fidedigna. En la época corría el rumor de que Davison aceptaba sobornos, y en esa ocasión el marido de Persiani, Giuseppe, que era el director del teatro, pudo haber influido en el juicio de Davison. En todo caso, cuando Francis Hueffer reemplazó a Davison en 1879, se decía que «tuvo que alquilar un carro de cuatro ruedas para devolverle a los distinguidos músicos los regalos que le llegaban a su puerta» [13]. Tratar de congratularse con los críticos es una actividad que siempre ha acosado al

mundo musical: aunque tal vez el puro soborno ha desaparecido de la escena, los críticos pueden ser capaces de aceptar invitaciones a fiestas y entradas gratuitas incluso hoy en día.

Los críticos pedantes y demasiado serios que seguían la tendencia de Fétis no tenían futuro en un mundo en el que todos leían sobre música. La crítica se había convertido en un género literario. Sin embargo, algunos grandes críticos seguían adoptando un tono solemne y altivo. Uno de ellos era Eduard Hanslick, profesor de historia y estética musical de la Universidad de Viena. Hanslick escribió reseñas de conciertos desde 1855, y en 1864 comenzó a colaborar en la *Neue freie Presse,* un diario que era leído no sólo en Austria sino también en Alemania. Los artículos de Hanslick solían aparecer en primera página. Hoy en día se le recuerda principalmente por su hostilidad hacia Wagner y por su folleto sobre estética de 1854, *Vom Musikalisch-Schönen* (Sobre lo musicalmente bello), pero en su día dominaba la crítica musical europea. Sus escritos combinaban un profundo conocimiento y un estilo serio pero fácil, basándose más en la descripción evocativa que en el análisis técnico. Escribía con claridad, precisión y elegancia, con un agudo discernimiento y elevados parámetros estéticos; pero siempre resulta agradable de leer. Cuando el pianista Rafael Joseffy dio un recital en 1874, Hanslick consideró que la esencia poética de la música se perdía en el deseo de lograr la perfección técnica:

> El tesoro poético [...] permaneció en general oculto. ¿Carece [el pianista] de la secreta resonancia del espíritu, o es que el puro placer del virtuosismo le constriñe? Tocó el *Traumeswirren* de Schumann, uno de los pequeños poemas más hermosos del compositor, como si fuera un *étude* cuyo objetivo es la máxima velocidad y el máximo nivel de *pianissimo* —casi a la manera de un reloj musical. Incluso el maravilloso movimiento central, ese núcleo del sueño, característicamente velado y nebuloso, estuvo demasiado expuesto al sol matutino. Ocurrió lo mismo con la maravillosa Novellette en Re, op. 21 núm. 4, de Schumann; hemos visto el salón de baile, pero no la conmovedora historia de amor que allí se representó [14].

Tal vez nos sorprenda que el destacado defensor del formalismo estético muestre semejante sensibilidad a la magia poética de la música.

Bastante diferente es el caso del contemporáneo menor de Hanslick, Bernard Shaw, quien escribió críticas musicales entre 1876 y 1894 para *The Hornet, The Star* (en el que firmaba como «Corno di Bassetto») y por último *The World,* una revista semanal. En Shaw se combinaban un verdadero genio para la crítica musical y una irresponsabilidad monstruosa. Era vulgar, insultante, divertidísimo, caprichoso e injusto, y el estándar de sus reseñas de conciertos era a veces lamentable. Le gustaba pretender ser un ignorante musical, pero en realidad su conocimiento del repertorio era extenso y sabía

mucho sobre música, especialmente sobre canto, ya que su madre era cantante. En lugar del aire de erudición que sustenta a los escritos de Hanslick, los textos de Shaw sugieren el *vox populi;* Shaw hace comentarios sobre el precio de los programas de mano, la avería de un vagón en las vías férreas o una visita al dentista, como si fuesen tan importantes como el concierto.

Sus prejuicios son conocidos. No le gustaba Brahms ni Bruch, elogiaba a Wagner, era cruel con Mendelssohn y admiraba a Hermann Goetz. Pero sus reseñas de conciertos a veces son exasperantes. Una y otra vez, Shaw escribía miles de palabras sobre un oratorio o una serie de conciertos sin decir nada más que la ejecución había sido execrable, o excelente, todo ello mezclado con muchos comentarios sobre lo anticuados que son los franceses o la plaga de violinistas femeninas. Siempre, invariablemente, expresaba una opinión personal. Critica un concierto de Joachim con el único objetivo de recomendarnos que no escuchemos el Concierto para violín de Bruch. Comenta el canto de Adelina Patti sólo para demostrar que Wagner, cantado correctamente, no cansa la voz.

El extraordinario ingenio y la brillantez de Shaw, por tanto, sólo son marginalmente pertinentes al estudio de la crítica musical. En realidad, Shaw podía escribir con agudeza y perspicacia sobre interpretación cuando quería. En el siguiente fragmento de una reseña de la Novena Sinfonía de Beethoven, dirigida por George Henschel, demuestra un profundo conocimiento y una enorme simpatía por el estilo de Beethoven, pero no puede evitar la ridiculización en sus últimas frases:

> En cuanto al señor Henschel y su interpretación de la Novena Sinfonía el pasado jueves, cuando digo que comprendió totalmente la naturaleza de la obra, y en ningún momento estuvo en peligro de cometer el tradicional y craso error de tratarla como un mero arabesco musical, quiero decir que el concierto fue un éxito [...] El vigor y la decisión con la que Henschel nos entregó el primer tema hicieron que nos tranquilizáramos rápidamente con respecto a su dominio de la situación, especialmente en esos dos últimos compases con los que Beethoven lo resuelve tan contundentemente [...] El señor Henschel, como el Maestro Contratista de Ibsen, y como todo buen director, tiene un troll dentro de sí; y a ese troll a veces le da por arrasar y obstaculizar, momentos en los que Henschel no sólo tolera e incluso se deleita al atacar brusca y descaradamente los pasajes imponentes, sino que también acelera a sus músicos [...] Pero Beethoven debe de haber sabido muy bien que ése es uno de los fallos comunes en las cualidades de un director; y para mí está claro que ese pavor a que cualquier urgencia o excitación vulgar estropeara la grandeza de su obra maestra sinfónica fue lo que le impulsó a indicar el tempo del primer movimiento no como un mero *allegro*, sino como «allegro, pero no demasiado, más bien majestuoso»...
>
> Más tarde, Henschel nos asombró a algunos con el tempo sumamente lento con el que tomó el *allegro assai,* en el que los bajos enuncian el tema de la oda a la

> alegría. Estamos tan acostumbrados a escucharlo a exactamente el doble de la velocidad demasiado rápida, como si las blancas y negras fueran corcheas y semicorcheas, y tratado como si fuese un allegro de Haydn en lugar de una melodía expresiva, que algunos de los oyentes más mayores se sintieron un poco indignados con Henschel por no haber tomado el camino erróneo habitual [...]
>
> [Durante el movimiento lento, una dama del coro se desmayó, dando lugar a] un despliegue de abanicos y frascos de sales, la conmoción entre los galantes tenores y bajos, al fondo del coro, y por último la desesperada recuperación de la paciente, que aplazó triunfantemente su caída desde lo alto de la empinada escalera al otro lado de la puerta. Durante todo aquello, la orquesta podría haber tocado «Pop Goes the Weasel» sin sentir más temor de que lo detectáramos del que habrían sentido si todos hubiésemos sido una clase de St. John Ambulance *...[15]

El sucesor de Shaw como principal crítico musical de Inglaterra fue Ernest Newman, pseudónimo de William Roberts, quien escribió para el *Manchester Guardian,* el *Birmingham Daily Post,* el *Observer* y finalmente el *Sunday Times* durante una larga carrera desde 1905 hasta 1958. Mientras que Shaw era un *enfant terrible,* Newman era humano y augusto. Incluso su pseudónimo era una expresión de sus creencias; se veía a sí mismo como un nuevo tipo de crítico, serio y juicioso. Sin embargo, las obras maestras de Newman eran sus ensayos eruditos sobre temas generales; sus reseñas de conciertos eran banales.

Mucho más personales e ingeniosas eran las críticas de Virgil Thomson en el *New York Herald Tribune.* Thomson era un compositor eminente, por supuesto; como crítico, era heredero de una distinguida tradición de crítica musical americana que incluía críticos magníficos como Henry T. Finck, quien escribió para *The Nation* y el *Evening Post* de Nueva York entre 1881 y 1924; el culto William J. Henderson [16], que escribía para el *New York Times* y el *Sun* entre 1887 y 1937; y el contemporáneo de Thomson, Olin Downes, cuyos artículos para el *New York Times* influyeron en la historia de la música (por ejemplo, ayudó a establecer a Sibelius dentro del repertorio canónico).

El enfoque de George Szell de la Quinta Sinfonía de Beethoven condujo a Thomson a examinar el contenido moral de dicha sinfonía:

> Un resultado moral estimulante es más importante que el peligro que entraña cualquier lectura errónea de la idea específica del compositor. Además, la pieza se recuperará de su actual servicio militar con la misma facilidad con que se libró de sus anteriores connotaciones metafísicas y políticas [...]
>
> Para llevar la sinfonía a la acción directa, el señor Szell se vio obligado a destacar la determinación del material masculino y franquear disimuladamente la im-

* Se refiere a las clases de primeros auxilios que imparte esa organización humanitaria, fundación de la Orden Hospitalaria de San Juan *(N. de la T.).*

> portancia de sus pasajes tiernos y gentiles. Hizo que las cuerdas tocaran fuerte y toscamente, con ese impacto feroz que producen tan admirablemente las cuerdas sinfónicas [...] Con la entrada de los trombones, al principio del último movimiento, las trompas se quedaron tan irremediablemente descalificadas en esa competencia de lanzamiento de peso como se habían quedado las maderas desde el principio. Todo el desequilibrio hizo que Beethoven sonara como un compositor muy autorizado y didáctico, que ciertamente era, pero también como un orquestador irremediablemente incompetente, que desde luego no era. Y precisamente esa ineficacia musical de los contrastes orquestales fue lo que demostró que, a pesar de la grandiosidad moral de la interpretación, se estaba quebrantando el espíritu de la obra, sea cual sea nuestro modo de concebirlo [17].

Tal combinación de crítica de la interpretación con una profunda percepción del contenido moral de la obra es realmente inusual.

En Alemania, la crítica permaneció en manos de eminentes músicos en activo y profesores de música. Su estilo no era ingenioso o elegante, sino digno, respetable y pesado. Heinrich Ehrlich, quien escribió para el *Berliner Tageblatt* desde 1878 hasta 1898, era un distinguido pianista. Su homólogo en Hamburgo, Max Loewengard, era profesor de teoría musical. El más distinguido de este grupo era el musicólogo Alfred Einstein, quien escribió críticas para el *Berliner Tageblatt* hasta 1933, además de revisar el catálogo de Köchel de las obras de Mozart y publicar un libro decisivo sobre el madrigal italiano.

En épocas más recientes, la crítica se ha vuelto menos solemne, menos *ex cathedra.* Hoy en día, los críticos con marcados intereses personales no son realmente aceptables; se espera que se centren en el intérprete y no en sus teorías favoritas sobre la música. Sin embargo, existen dos tendencias que han arruinado el arte del periodismo musical. En primer lugar, el espléndido espacio que se le asignaba a la crítica musical se ha reducido; sólo algunas publicaciones periódicas, como el *New Yorker,* continúan reservando millares de palabras —como las que Shaw o Hanslick podrían exigir— para la crítica musical. Los críticos modernos que compiten por el espacio con las reseñas de música pop, cine, radio y discos, y sobre todo con los anuncios publicitarios, frecuentemente deben expresar su opinión en 500 palabras o menos.

En segundo lugar, las ventajas en cuanto al transporte han tendido a centralizar la reseña musical en las ciudades más importantes. Si una orquesta da el mismo concierto en varias ciudades, o una compañía de ópera hace una gira en diversos teatros, un periódico nacional como *The Times* naturalmente lo reseñará una sola vez. Esto ha provocado el declive o traslado a la capital de importantes periódicos provinciales. Un proceso similar ha ocurrido en otros países, especialmente en Estados Unidos.

Pero, en todo caso, el estilo periodístico admitido ha cambiado. La escritura singular, digna y caballerosa de Ernest Newman ha dado paso a una es-

pecie de precisión lacónica, más lista que erudita, más bromista que ingeniosa y más popular que autorizada. En lugar de la majestuosidad de un extenso artículo de reseña, los mejores críticos buscan hoy en día la crítica memorable de una sola línea. Reseñando una interpretación del guitarrista Carlos Bonell de un canto mozárabe del siglo XI, la admirable escritora Hilary Finch comentó que la música «giraba en pequeños círculos pentatónicos, tañida justo por debajo de la boca de la guitarra, con una voz cobriza que parecía provenir de más allá del tiempo mismo» [18].

Quizá el cambio de mira más importante de la crítica moderna es el descubrimiento de la «interpretación». Mientras que el objetivo de los críticos anteriores era alabar o condenar, o al menos evaluar, la manera en que el intérprete realizaba las «intenciones» del compositor, los críticos contemporáneos han reconocido que el intérprete añade su personalidad y sus ideas a la obra. Por lo tanto, puede haber diferentes realizaciones de una misma pieza, y todas son válidas a su manera. En sus reseñas de una serie de conciertos de Mahler para *The Times,* Joan Chissell señaló que Otto Klemperer y Jascha Horenstein enfocaban de diferentes maneras la obra del compositor:

> [la Primera Sinfonía de Mahler] es el Mahler joven, y el señor Horenstein hizo que sonara joven. Incluyendo la repetición en el primer movimiento, la interpretación duró una hora, pero el tiempo voló porque los ritmos eran muy animados y las texturas muy traslúcidas. Mientras que en la séptima y última semana, el doctor Klemperer se mantuvo distante y objetivo, abandonando la mayoría de los detalles a su suerte, Horenstein hizo todo lo contrario; todo estaba fraseado y coloreado con precisión, y sin embargo no hubo el menor rastro de exageración emocional [19].

Chissell no sugiere que Horenstein, ni Klemperer, se hubieran acercado más al «verdadero» espíritu de Mahler; simplemente indica que sus personalidades eran diferentes. Esta actitud se ha visto fomentada por la disponibilidad de las grabaciones, que permiten una comparación objetiva de diferentes interpretaciones. Cuando Andrew Porter se encontraba redactando una reseña de la Novena Sinfonía de Mahler, salió a comprar un disco de la versión de Bruno Walter para compararla con la versión que acababa de escuchar, dirigida por Pierre Boulez. Se sorprendió al descubrir que los tempos eran muy similares a los de Walter, aunque la impresión era bastante diferente [20].

Finalmente los críticos reconocen diferentes estilos de interpretación de obras canónicas; ya no hay uno correcto y uno incorrecto. En agosto de 1999, un crítico del *Independent* describió a Alfred Brendel como un pianista «clásico», sugiriendo que hay otros tipos de pianista cuya interpretación puede ser igualmente buena:

> La música clásica, y la interpretación clásica, no son realmente una cuestión de lo que haces, sino más bien de lo que no haces. Alfred Brendel es, quizá, el pianista clásico más importante del mundo. Brendel no alarga brevemente cada nota importante. Tampoco hace un *crescendo* emotivo hacia la mitad de cada frase. No toca los pasajes rápidos demasiado rápido, ni los lentos demasiado lento. Y si la partitura indica que hay que repetir una sección, él la repite.
>
> El resultado es más narrativo que expresivo. Te encuentras más bien siguiendo una historia cautivadora que apasionándote. Todo está interpretado con el mayor refinamiento, y nunca toca un cliché, ni siquiera cuando el compositor lo ha escrito [22].

A veces los intérpretes y sus admiradores maldicen a los críticos por ser irresponsables, por dedicarse simplemente a llamar la atención, por tomar a la ligera el trabajo de los artistas serios. La mayoría de los que escriben sobre música repudiarían esas acusaciones. Los críticos son honestos y generalmente escriben con el corazón. Naturalmente, han de atraer el interés de sus lectores; la crítica musical no se puede escribir en el mismo estilo que los informes económicos o los comentarios sobre boxeo, aunque la mayoría de nosotros hemos aprendido algunas lecciones útiles de los cronistas deportivos [23]. Los que discrepan de sus opiniones también acusan a los críticos de ofrecer sólo la «opinión de una persona». Pero ése, claro está, es el único y verdadero objetivo de la crítica musical. Como bien sabía Shaw, nuestros peores enemigos seguirán comprando el periódico para leer nuestras críticas y maldecirlas de nuevo.

## *Notas*

1. *Musical Quarterly,* 1/1 (1915), p. 75.
2. *The Sunday Times,* 30 de agosto de 1931.
3. *Litteratur- und Theater-Zeitung,* 4/1 (1781), p. 52.
4. A. B. Marx fundó una revista en Berlín con el mismo nombre en 1824.
5. *AmZ,* 28 de abril de 1802, pp. 498-499.
6. *AmZ,* 28 de abril de 1830, p. 270. Ernst, quien fue elogiado más tarde por Berlioz, Joachim y Mendelssohn, había tocado en Stuttgart.
7. *The Times,* 20 de septiembre de 1843.
8. Citado de Max Graf, *Composer and Critic: 200 Years of Musical Criticism,* Londres, Chapman & Hall, 1947, p. 217.
9. *Ibid.*, p. 218.
10. *Ibid.*, pp. 173-174.
11. Para 1851 *The Times* había alcanzado una tirada de 40.000 ejemplares; véase *ibid.*, p. 260.
12. *The Times,* 14 de abril de 1847.
13. *The History of The Times. Part 2: The Tradition Established, 1841-1884,* Londres, Office of The Times, 1939, p. 445.

14. Citado de Henry Pleasants (trad. y ed.), *Eduard Hanslick, Music Criticism 1846-1899,* Harmondsworth, Penguin, 1950, pp. 123-124.
15. *The World,* 8 de marzo de 1893; citado de Bernard Shaw, *G. B. S. on Music,* con un prólogo de Alec Robertson, Harmondsworth, Penguin, 1962, pp. 118-120. Es posible que la reseña contenga una sutil indirecta a Hanslick, quien había descrito la música como un «arabesco».
16. La visión arrogante que tenía Henderson de la crítica de la interpretación ha sido citada al principio de este capítulo.
17. *New York Herald Tribune,* 18 de marzo de 1945; citado de Virgil Thomson, *The Art of Judging Music,* Nueva York, Greenwood Press, 1969, pp. 140-141.
18. *The Times,* 28 de diciembre de 1999.
19. *The Times,* 9 de octubre de 1968.
20. Andrew Porter, *Music of Three Seasons, 1974-1877,* Londres, Chatto & Windus, 1979, pp. 427-428; la reseña fue publicada en *The New Yorker,* 8 de noviembre de 1976.
21. *The Independent,* 30 de agosto de 1999.
22. Ernest Newman escribía sobre deporte y sobre música, y Neville Cardus era crítico de música y de cricket.

## *Lecturas complementarias*

Cone, Edward T.: «The pianist as critic», en John Rink (ed.), *The Practice of Performance: Studies in Musical Interpretation,* Cambridge, Cambridge University Press, 1995, pp. 241-253.

Haskell, Harry: *The Attentive Listener: Three Centuries of Music Criticism,* Londres, Faber, 1995.

Newman, Ernest: *A Musical Critic's Holiday,* Londres, Cassell, 1925.

Price, Kingsley (ed.): *On Criticizing Music: Five Philosophical Perspectives (The Alvin and Fanny Blaustein Thalheimer Lectures, 1978-1979),* Baltimore, Johns Hopkins University Press, 1981.

Thompson, Oscar: *Practical Musical Criticism,* Nueva York, Da Capo Press [1934], 1979.

# *16 Los intérpretes y la interpretación*

JONATHAN DUNSBY

## *Qué siente el intérprete*

«Puedo dar fe de que en mi larga vida jamás he escuchado a un cantante pronunciar esta feliz verdad: Hoy me siento bien.» Este comentario de una autoridad internacional del canto puede parecer divertido, pero también contiene varios indicios importantes acerca de lo que piensan los intérpretes sobre la interpretación. Existe, por ejemplo, una relación indudable entre la interpretación y el bienestar físico y anímico del intérprete. Hacer música tiene muchos efectos terapéuticos conocidos, pero siempre exige un esfuerzo físico, aunque también suele ser muy estimulante. La música occidental plantea muchas exigencias mentales, por ejemplo en cuanto a memoria y concentración. También requiere un control de las emociones, o al menos así parece, ya que nos sorprendería ver una lágrima deslizándose por la mejilla de un violinista durante un momento particularmente conmovedor de un concierto, a pesar de la importancia que tiene la expresión facial dentro del aspecto visual de la interpretación. Como ha señalado un estudioso al analizar la interpretación como una simulación: «mientras están interpretando, los intérpretes pueden tener la mente en cualquier cosa. Con el espíritu de entretenimiento profesional, un músico puede interpretar con sensibilidad y estar al mismo tiempo sumamente aburrido y anhelando un tranquilo trabajo en el sector inmobiliario» [1]. En segundo lugar, el contexto de la cita inicial de este capítulo es una argumentación sobre qué deberían sentir los cantantes, y su fuente representa toda una colección de escritos y otros testimonios de intérpretes a lo largo de los siglos que plantean que la interpretación tiene algún

tipo de «objetivo» y no es una actividad indiferente —una que, por así decir, es tan buena en un día malo como en uno bueno. Interpretar requiere un grado excepcionalmente alto de entrenamiento y habilidad, hasta un nivel «casi sobrehumano» según uno de los principales directores del siglo pasado [2]. En tercer lugar, la cita inicial no proviene de un pasado reciente, ni siquiera de los últimos dos siglos, sino de 1757 [3], lo que nos indica que puede transcurrir mucho tiempo sin que ocurran grandes cambios. En los años 1780 se escribía sobre la ejecución pianística de Mozart casi como si hubiera sido el ganador de un concurso internacional reciente [4], y la antigua mitología griega describía a Orfeo y sus poderes mágicos como intérprete de la lira casi de la misma manera en que la «realidad» de los medios de comunicación de principios del siglo XXI resucita la presencia casi divina de un Frank Sinatra o un Freddie Mercury —o, en el campo de la interpretación instrumental, un Jimi Hendrix a la guitarra o un Stephane Grappelli al violín.

Interpretar música es una actividad tan positiva, en la que participan tantos tipos de atención y reflejos humanos, que los intérpretes piensan que se trata en gran parte de lo que podríamos llamar «simplemente tocar» (y «simplemente cantar»). Pero indudablemente se necesita algo más para ir más allá de la mera ejecución, y es aquí donde se debe establecer una diferencia entre los intérpretes que se inclinan hacia el enfoque intuitivo y los que piensan que un apoyo cognitivo minucioso produce mejores resultados. Hans Keller —que no sólo era un notable crítico musical, sino también un excelente intérprete y maestro de algunos de los mejores intérpretes de cuerda occidentales de las últimas décadas del siglo pasado— escribió acerca del «intérprete irreflexivo que se enorgullece de "ser simplemente un intérprete", y que nunca ha pensado sino sólo sentido la música —quien, de hecho, cree que el pensamiento es enemigo del sentimiento, mientras que en la realidad musical no existe ningún sentimiento que no se pueda pensar claramente» [5]. En este fragmento nos está diciendo que la premisa «qué siente el intérprete» es inadecuada para el estudio y la práctica que estamos considerando en este libro. Más que nada, debemos recordar que, para el intérprete, el «sentimiento» musical es una mezcla de emoción e inteligencia, de respuesta y control, de empatía y dominio, de lo autónomo y lo propioceptivo (para utilizar términos fisiológicos más técnicos). Expresándolo de la manera más simple pero también más profunda, y en una forma que podamos comprender de alguna manera, el «sentimiento» es una mezcla del ser y el hacer.

Es evidente que la crítica de la interpretación, la crítica constructiva, suele ser extremadamente difícil. Escribir sobre *piezas* musicales, fenómenos tan inefables y volubles, se vuelve al menos doblemente difícil cuando se escribe sobre la *interpretación* de piezas musicales, sobre fenómenos que, por mucho que podamos grabarlos gracias a la tecnología, son como, por ejemplo, un

partido de fútbol —una vez terminado, ha terminado para siempre. Sin embargo, el deseo de comprender y grabar lo que hacemos forma parte de la naturaleza de los seres humanos. No hay una distancia inasible entre una pintura rupestre de hace 20.000 años que representa la caza de un animal salvaje —la desesperada y escabrosa inmediatez de lo que sucedió después nunca se podrá recuperar— y una reseña escrita hace cien años sobre una interpretación memorable de un aria que Puccini había compuesto entonces recientemente. Así como las prácticas de caza prehistóricas, esenciales para la supervivencia, fueron transmitidas y codificadas, algunos músicos han afrontado el hecho de que existe una *práctica* de la música occidental, se han preguntado cómo se interpreta y han tratado de expresarlo.

## *Qué piensa el intérprete sobre la música*

José Bowen ha escrito sobre la «*ilusión* de que existe un estilo interpretativo neutral (o natural)» [6]. Si realmente existe esa ilusión, y Bowen no está simplemente inventando una falacia para reafirmar su argumento, creo que es principalmente una ilusión del oyente; pero, aun así, quizá el término «ilusión» no sea el correcto. Por poner un caso concreto, aunque quizá nunca me acaben de gustar las grabaciones «auténticas» que hizo Eliot Gardiner de las sinfonías de Beethoven, comparadas con las interpretaciones de Herbert von Karajan con la Filarmónica de Berlín (la máxima referencia musical de los años 1960), nunca he pensado que el sello discográfico Decca estuviera atesorando algo «neutral» o «natural» en sus grabaciones posteriores a la Segunda Guerra Mundial, es decir, antes de la época de la «autenticidad». Todas las pruebas apuntan al hecho de que los intérpretes saben y siempre han sabido, aunque quizá no del todo, que su actividad tiene lugar dentro de un contexto cultural. Saben quién les enseñó y les influyó. Saben que, a medida que envejecen, están menos abiertos a las influencias en el sentido de un verdadero cambio en su forma de trabajar. Saben que su «mundo» interpretativo puede resultarles cada vez más ajeno a las nuevas generaciones, pero, a la inversa, en su vejez pueden desempeñar un papel importante en la preservación de un estilo interpretativo en vías de extinción.

Aunque Bowen representa el punto de vista de un estudioso realizando un examen a posteriori, y no tanto el de un intérprete que para un objetivo artístico necesita tener fe en algún punto de partida sin adentrarse demasiado en su credibilidad, tiene razón, sin embargo, cuando sugiere que los intérpretes siempre demuestran una actitud hacia su música. Planteemos tres preguntas que pueden considerarse representativas de las actitudes relativamente modernas que tienen los intérpretes hacia su música. ¿Por qué Glen Gould

tocaba Bach en el piano? ¿Qué pensaba Dietrich Fischer-Dieskau que hacía, aparte de «simplemente» cantar los lieder de Schubert con las notas correctas en el orden adecuado? Y ¿qué se supone que debe opinar un director sobre la partitura?

En el caso de Gould, debemos tener presente que este maestro pianista no sólo era, según la opinión unánime, uno de los más grandes intérpretes del siglo XX que se dedicó (casi patológicamente, según algunos) a la calidad del sonido grabado, en que los mejores elementos de diversas interpretaciones pueden juntarse en una interpretación trascendental que nunca, o casi nunca, puede esperarse en un concierto en vivo, sino que además era un filósofo musical que estaba convencido de que la música existe en un estado ideal que nunca puede realizarse completamente. Para él esto era obvio en un sentido práctico. En su ensayo «The Prospects of Recording» (1966) («Las perspectivas de la grabación sonora»), escribió: «determinar el valor de una obra de arte de acuerdo con la información que se tiene sobre ella es la forma más inaceptable de valoración estética» [7]. Según la visión de Gould, Bach no pretendió saber cómo se ha de escribir su música o hacerla realidad en sonido más allá de sus propias circunstancias. Basta decir que Bach era un genio de la invención musical, y que no deberíamos esperar que el Cantor del siglo XVIII fuera también un futurólogo infalible capaz de evocar una imagen del Steinway de gran cola de los siglos XIX y XX. Gould estaba completamente seguro de que Bach no podía anticipar el futuro (aunque se adentró en él con sus composiciones) más de lo que nosotros podemos recuperar el pasado (a pesar de nuestra necesidad de asimilarlo). Si Gould hubiera vivido los debates sobre autenticidad que se produjeron en el ámbito de la erudición musical y, a su debido tiempo, en el comercio musical a partir de 1982, probablemente habría atemperado algunas de las presunciones históricas y filosóficas que surgieron en torno a las importantes contribuciones hechas por teóricos [8].

En este contexto, por lo tanto, Gould no sólo creía que era correcto tocar Bach en el piano, sino que además lo consideraba un deber artístico. En un minucioso ensayo sobre el enfoque de Gould, Jean-Jacques Nattiez describe el punto más importante de este aspecto de la autenticidad:

> La tecnología debería liberar al intérprete de la tiranía de la cronología, permitiéndole elegir cómo interpretar, y el ejecutante tiene todo el derecho a la libertad de interpretación, incluso a tener más de una concepción de una misma obra. Para Gould, las partituras de Bach no son más que «una excusa para construir una variedad infinita de sistemas interpretativos pertinentes» [...] Lo mismo se aplica al oyente, a quien Gould anima a escuchar en un disco algo diferente de lo que él mismo habría predicho [...] y para quien espera un futuro en el que la tecnología le permita participar en la composición final de la grabación. Hoy en día, «juguetear

> con el sintonizador es, en su forma limitada, un acto interpretativo» [...] y el día de mañana, el oyente deberá poder elegir entre dieciséis tomas de un mismo pasaje [...] y juntar diferentes tipos de soluciones interpretativas[9].

Este texto presenta algunas ideas que hay que explicar. Como para Gould la música existe en un estado ideal —de hecho, una forma platónica—, para Nattiez no es una exageración referirse a la «tiranía» de la cronología. A modo de ilustración, propongo dos afirmaciones por lo visto similares:

(1) Bach murió en 1750 y no pudo haber leído estas palabras.
(2) Bach escribió música para el clave del siglo XVIII que nunca se podría interpretar correctamente en el piano en el siglo XX.

La primera afirmación es evidentemente cierta por sí misma de acuerdo con nuestro conocimiento de la realidad empírica. La segunda afirmación puede ser verdadera, pero indudablemente no es del mismo tipo de afirmación que la primera. Aunque su cualidad de verdadera o falsa dependa del término «correctamente», observamos sin embargo que existe una contingencia en la segunda afirmación que no se aplica a la primera. La interpretación, como Gould nos ha enseñado tan radicalmente, fluye desligada del tiempo (y del lugar y todas las demás contingencias del origen de la obra musical). Sin embargo, Gould no está afirmando que el intérprete pueda hacer cualquier cosa y salir triunfante. Lejos de eso, vincula concienzudamente el intérprete a la composición —y lo hace de diversas maneras en sus escritos—, en este caso, con la palabra «pertinente». Sin adentrarnos mucho más en este argumento, podemos observar que, en el nuevo mundo tecnológico de Gould, el oyente tiene una responsabilidad similar. No podemos ni imaginar lo que Gould habría opinado de la posibilidad, que ya existe, de escuchar una de sus grabaciones a una velocidad mucho mayor sin alterar la altura original (una técnica que aún estaba en pañales a principios de los años 1980). Quizá habría pensado que eso va más allá de lo «pertinente»; pero nunca lo sabremos, debido a la tiranía de la cronología.

Después de estas estimulantes ideas, tal vez sea reconfortante pasar a nuestro segundo tema, Dietrich Fischer-Dieskau, y su opinión sobre el canto. De entrada, el cantante está seguro de que interpretar es algo más que cantar las notas en el orden adecuado si se quiere transmitir el valor compositivo que se ha percibido: «si observamos [...] la trayectoria de la canción romántica alemana, vemos destacada, como punto de partida esencial para cualquier desarrollo futuro, la necesidad de expresar los sentimientos humanos, de comunicar, y de renunciar a la mera ejecución del sonido» [10]. Su cumplimiento de tal renuncia es indudable, como se puede escuchar en su

amplio legado de grabaciones, pero lo técnico nunca está muy lejos de su mente. Véase esta concienzuda reflexión sobre Wolf:

> las dificultades que planteaba Wolf tanto al intérprete como al oyente eran nuevas y múltiples. Exigía un nivel de *precisión técnica* que desconcertaba a sus contemporáneos. Antes que Wolf, nadie había requerido interpretaciones con una gama tan amplia de efectos auditivos. La forma declamatoria de la canción parecía insuperable; Wolf suprimió, sin reparos, el *bel canto* y la coloratura [11].

En un aspecto más general, Fischer-Dieskau opina que «el intérprete, como una especie de constructor de puentes para lo que está por venir [...] tiene la enorme responsabilidad de preservar, mediante una interpretación inmaculada, las creaciones existentes de los maestros» [12]. De nuevo, hay algo que debemos observar más allá de la evidente franqueza de la afirmación:

> Aunque alcanzamos un dominio perfecto del instrumento, la obra, su espíritu, amenaza con eludirnos. No es fácil para nosotros discernir los principios de organización que han proporcionado la sustancia espiritual a la obra. Sólo podemos esforzarnos por escuchar la música en nuestra mente con el mismo espíritu con que se le apareció a su creador [...] La música y la poesía poseen un ámbito común, del cual obtienen la inspiración y en el cual se desenvuelven: el paisaje del alma. Juntos, la obra y el intérprete, tienen el poder de dar forma intelectual a las sensaciones y emociones, de transmutarlas en un lenguaje que ningún otro arte puede expresar [13].

En esta declaración convergen muchas de las ideas expuestas anteriormente. No sólo apoya el argumento general, sino que en particular nos dice que el intérprete le da «forma intelectual a las sensaciones y emociones» que, como hemos indicado, parece ser un dato conocido de la música occidental. También respalda un tanto el compromiso de Gould con lo que los eruditos alemanes del siglo XX llamaron el *Werk-Konzept* (la obra musical como concepto) —en otras palabras, con la idea de que las piezas musicales existen en una especie de estado «ideal» gouldiano que los grandes intérpretes deben captar, como dijera Fischer-Dieskau, «con el mismo espíritu con que se le presentó a su creador». Fischer-Dieskau estaba mucho más comprometido con la idea de que existe un *Werk* históricamente pertinente que debe ser redescubierto, pero resulta instructivo observar las afinidades entre dos intérpretes trascendentales que, vistos desde un punto de vista limitado en su época, eran tan diferentes: por un lado, Gould, que abandonó sus actuaciones públicas y se dedicó a la interpretación «inauténtica», y por otro Fischer-Dieskau, que se deleitó con los aplausos del público por todo el mundo al servicio de ese santo grial, el «espíritu» con que una u otra canción «se presentó a su creador». Hay una semejanza subyacente entre estos genios del piano y la voz, cuyas

interpretaciones llenan extensos catálogos de la historia reciente de la música grabada. Fischer-Dieskau podría haber adoptado como un legado suyo lo que Jeremy Siepmann ha escrito sobre Gould:

> Planteadas como preguntas y no tanto como declaraciones, menos aún como respuestas, muchas de las interpretaciones más originales e indagadoras de Gould se encuentran entre las más reveladoras de todas las que han sido grabadas. Gould me ha impulsado una y otra vez a regresar a la partitura y aprender de ella de nuevo —no por emular a Gould, cuyos exagerados amaneramientos siguen siendo inimitables e indeseables, sino en busca de un contacto renovado con el espíritu y la infinita fascinación de la música misma. No se le puede pedir nada más a un músico. En sus mejores interpretaciones, y hasta un nivel que no ha sido superado por ningún otro intérprete, Gould transmite una sensación de enorme alegría y asombro, y una jubilosa conciencia del privilegio que tenemos de estar vivos [14].

Por otra parte, debido a que sus interpretaciones carecen de experimentos y parecen representar el estándar de la tendencia principal, Fischer-Dieskau transmitía al público y a los especialistas por igual una profunda impresión de «autenticidad» en el sentido común de la palabra, de reproducir las supuestas intenciones del compositor (supuestas, porque no podemos regresar a una época pasada, y, aun así, nunca sabemos realmente lo que otra persona está pensando). Otro notable artista, el acompañante Gerald Moore, escribió: «[Fischer-Dieskau] comprende lo que sentía el *compositor*, y sabe *desvelarlo y expresarlo* tan penetrantemente, que llega al corazón» [15]. Sin embargo, cuando no hay experimentación, existe necesariamente una tendencia al reciclaje en lugar de la reinvención de lo que los intérpretes piensan sobre el repertorio canónico, incluso una tendencia al «estancamiento cultural» [16].

Este tono amargo nos acompañará cuando consideremos lo que piensan los directores sobre la música, al menos en la opinión de nuestro principal informante, Gunter Schuller, quien tuvo una larga y exitosa carrera internacional como intérprete, director y compositor. Schuller afirma que «la idea de que un artista musical, un director, ha de servir a la música —y no que la música sirve al músico— es apoyada a veces en apariencia, pero raramente se pone en práctica» [17]. Prosigue señalando que los directores han sido con frecuencia menos competentes en ciertos aspectos de su arte, y además han adoptado una actitud equivocada ante la música. No deberían interpretarla en absoluto, sino que sus interpretaciones deberían ser «realizaciones» —y cita a Ravel como la autoridad en la que se apoya para exigirlo. Para mostrar un ejemplo concreto, podríamos estudiar su análisis del inicio de la Quinta Sinfonía de Beethoven, una descripción que abunda en detalles y extiende su crítica a la mayoría de las noventa grabaciones analizadas que, según el autor, acentúan erróneamente el segundo compás de cada frase de cuatro compases,

en lugar del primero [18]. Nadie culparía a Schuller por considerar asombroso, incluso irritante, que se siguiera haciendo caso omiso del argumento planteado por el teórico Heinrich Schenker en 1925, como lo hiciera el director Norman del Mar [19] casi setenta años después y como lo hicieron entretanto miles de interpretaciones. Éste es un caso particular de una cuestión general que Schuller confirma al estudiar grabaciones de ocho obras representativas del repertorio sinfónico:

> El público medio, e incluso muchos profesionales y «expertos» en música, tiene poca o ninguna idea de lo que contiene o indica realmente una partitura, por ejemplo, de Brahms. Como consecuencia, si la orquesta toca técnicamente bien [...] el público tiende a asumir (¿qué otra cosa podría hacer?) que lo que está escuchando corresponde a lo que el compositor escribió, especialmente si ha recibido alguna carga emocional de la interpretación. (Como he explicado muchas veces [...] la carga emocional que el público recibe de una determinada interpretación puede no coincidir con la que el compositor pretendía, puede ser bastante ajena a la música y puede incluso haber sido sustituida por el director.) Al público, al oyente medio, sólo le queda la capacidad de valorar una interpretación según la encuentre «emocionante» o «no emocionante» [...] ¿Cuál es la solución para estos problemas? [...] elevar el estándar de la industria, una mayor integridad y honestidad artística en el campo musical; un público y un gremio de críticos más selectivo, educado y culto —objetivos y sueños en los que nos hemos quedado muy atrás en los últimos tiempos [20].

Hay algo que podemos afirmar acerca de lo que opinaba este intérprete sobre la interpretación: ¡no estaba satisfecho! A ello hay que añadir que no se trataba en ningún modo de reclamar cierta «deshumanización» de una «realización». Schuller seguramente coincidiría con el pianista, notable escritor y conferencista Alfred Brendel en que «comprender las intenciones del compositor significa *traducirlos* a nuestra propia comprensión» [21].

### *Qué creen los intérpretes que deberían pensar sobre la música*

Las opiniones analizadas hasta ahora son necesariamente un tanto generalizadas, pero, aun así, el lector informado puede haber sospechado que un enfoque analítico de la música sería inevitable en esta perspectiva general del tema. Ciertamente, dos ejemplos emblemáticos nos permitirán plantear la posibilidad de que el enfoque analítico de la interpretación musical es absolutamente necesario y fiable. El primer ejemplo —quizá un tanto sorprendente para algunos, ya que proviene del intérprete archirromántico (al menos popularmente) Sergei Rachmaninoff— es un comentario sobre la importante configuración del clímax de cualquier pieza musical:

> Esta culminación, dependiendo de la pieza, puede encontrarse al final o en el medio y puede ser fuerte o suave; pero el intérprete ha de saber cómo abordarla con un *cálculo absoluto*, una precisión absoluta, porque si pasa inadvertida, toda la construcción se vendrá abajo, la pieza se tornará inarticulada e incoherente y no transmitirá al oyente lo que debe transmitir [22].

¿Qué quería decir Rachmaninoff exactamente con la expresión «cálculo absoluto»? Probablemente esté representado por un estado psicológico diferente para diferentes intérpretes (y es muy posible que esté relacionado con lo que los científicos llaman el equilibrio bipolar del cerebro humano, un asunto sumamente individual). Sin embargo, y regresando al *Werk-Konzept* mencionado anteriormente, Rachmaninoff afirma que cada pieza posee una estructura musical que existe independientemente de cualquier intérprete en particular. El intérprete debe identificar la estructura, asimilarla y controlarla suficientemente bien en tiempo real para representarla e, idealmente, transmitirla con precisión. Independientemente de la curiosa opinión de Rachmaninoff, que después de todo fue uno de los músicos más «inspirados» e inspiradores de principios del siglo XX, su compromiso con la estructura no debería resultar sorprendente proviniendo de un intérprete tan excelente que compuso obras a gran escala, extremadamente complejas y en muchos sentidos modernas [23]. El «cálculo», por lo tanto, no es un concepto sorprendente en este contexto, sino una aseveración del control que debe ejercer la estructura musical sobre el intérprete, y ello implica que el análisis estructural en alguna de sus formas simplemente debe ser parte del trabajo del intérprete, por más imperceptible que sea para el público y por más invisible que resulte en el producto final. En efecto, al público le puede resultar sorprendente que la interpretación en sí misma ocupe una parte más bien pequeña de la atención del intérprete. La verdadera interpretación es la punta de un iceberg formado por la práctica y el ensayo del intérprete, lo que de diversas e incontables maneras constituye el nivel «analítico» del hacer música, la etapa en que todo se ordena mental y físicamente para ese «cálculo» que se realizará en el escenario y tendrá una sola oportunidad de salir bien. En los demás casos, uno está «simplemente tocando» [24].

Los intérpretes, pues, ocupan una escala que abarca desde los que son levemente conscientes del nivel analítico (el intérprete literalmente «irreflexivo» de Keller, mencionado anteriormente, es tal vez raro, lo cual no quiere decir que haya que restarle importancia a la validez general de su argumento) hasta los que consideran el análisis como un asunto constante y consciente, como seguramente es el caso en los niveles más altos de la interpretación musical, y como siempre ha sido el caso en la música occidental. Los teóricos, sin embargo, están excepcionalmente preparados para ser conscientes del ni-

vel analítico y explicar con precisión lo que piensan sobre el asunto. La experiencia de Schenker como uno de los pianistas acompañantes más famosos de finales del siglo XIX y su convincente teoría sobre la estructura de obras maestras occidentales le permitieron asegurar que los intérpretes siempre deben reconocer la primacía del conocimiento sobre la práctica. En un ensayo de 1925 sobre el movimiento Largo de la Tercera Sonata para violín solo de Bach, Schenker escribió: «Una vez que se entiende completamente el contenido de una pieza, la interpretación no plantea problemas» [25]. Prosigue señalando que existe una correspondencia indudable entre, por ejemplo, la «matización dinámica primaria» de los eventos estructurales subyacentes de una pieza (los que articulan el «argumento», por así decir) y las «matizaciones propias» de la superficie musical: «Más allá de todas estas matizaciones entran en consideración otros matices, más delicados [...] Pero todos ellos han de integrarse en el esquema dinámico primario y las matizaciones propias de un orden estructural más elevado» [26]. Sin embargo, en otros puntos también nos recuerda que el «orden estructural más elevado» no es excusa para que el intérprete descuide los más pequeños detalles de una partitura, así como «el mapa de ruta del alpinista le evita tener que franquear todos los caminos, rocas y atolladeros». Cada detalle, dice, «debe significar lo mismo para el intérprete que para el compositor» [27].

### *Impresiones*

Los comentarios finales también deberían ofrecer perspectivas. Las voces del pasado, que con este capítulo se han introducido en los pensamientos futuros del lector, plantean unas exigencias claras. Piden compromiso y humildad, que deberían ser obvios. También plantean algunas cuestiones fundamentales, cuyas respuestas son necesarias para proseguir con la urgente labor de recrear la música del pasado en el futuro, pero que nunca llegarán a estar «resueltas» para siempre. Presumiblemente, nunca habrá una definición clara de lo que real y definitivamente «es» una pieza musical. Es probable que ni siquiera la pregunta en sí misma sea la más inteligente y provechosa. Lo que una pieza musical puede «ser» idealmente es una experiencia nouménica y memorable que transforma la conciencia del oyente. Eso es lo que ofrecían en su trabajo diario tanto Mozart y Murray Perahia, por citar a uno de los primeros virtuosos del piano junto a uno de los más recientes, como Schuller y Del Mar, dos directores recientes cuyas opiniones difícilmente podrían ser más contrarias [28]. Sin embargo, los intérpretes que nos hablan sobre la interpretación tienden a describir la punta del iceberg, como hemos visto, prefiriendo naturalmente concentrarse en el efímero objetivo, el producto, que en

el camino, el proceso sustancial mediante el cual llegan a ese objetivo. Si Orfeo realmente amansaba a las fieras con su lira —¿y quiénes somos nosotros para negarlo?—, debe de haber sabido exactamente lo que hacía como intérprete.

## *Notas*

1. Stan Godlovitch, *Musical Performance: A Philosophical Study,* Londres y Nueva York, Routledge, 1998, p. 127.
2. Hermann Scherchen, *Handbook of Conducting,* trad. M. D. Calvocoressi, Londres, Oxford University Press, 1933, p. 3.
3. Johann Friedrich Agricola, *Anleitung zur Singekunst,* Berlín, Winter, 1757; citado de Julianne C. Baird (ed. y trad.), *An Introduction to the Art of Singing by Johann Friedrich Agricola,* Cambridge, Cambridge University Press, 1995, p. 218. La obra de Agricola es una versión alemana con comentarios del libro *Opinioni de' cantori antichi e moderni* de Pier Francesco Tosi, escrito en 1723, que ya había aparecido en inglés en 1742 (Londres, Wilcox).
4. Véase por ejemplo el comentario que hace Michael Kelly sobre la técnica interpretativa de Mozart en Louis Biancolli (ed.), *The Mozart Handbook,* Nueva York, Grosset & Dunlap, 1962, p. 134.
5. Hans Keller, *The Great Haydn Quartets: Their Interpretation,* Londres, Dent, 1986, p. 15.
6. José A. Bowen, «Finding the music in musicology: performance history and musical works», en Nicholas Cook y Mark Everist (eds.), *Rethinking Music,* Oxford, Oxford University Press, 1999, p. 445; subrayado mío. Para más información sobre la idea de la «ilusión» en los estudios sobre interpretación, véase Jonathan Dunsby, *Performing Music: Shared Concerns,* Oxford, Clarendon Press, 1995.
7. Glenn Gould, «The prospects of recording», en Tim Page (ed.), *The Glenn Gould Reader,* Nueva York, Knopf, 1984, p. 341.
8. Una extraordinaria contribución histórica a este debate fue el ensayo de Richard Taruskin «The pastness of the present and the presence of the past», publicado por primera vez en Nicholas Kenyon (ed.), *Authenticity and Early Music,* Oxford, y Nueva York, Oxford University Press, 1988, pp. 137-207. La publicación de la obra de Peter Kivy, *Authenticities: Philisophical Reflections on Music Performance,* Ithaca y Londres, Cornell University Press, 1995, marcó un hito en la filosofía de la música. Para una notable psicobiografía de Glenn Gould, véase Peter Oswald, *Glenn Gould: The Ectasy and Tragedy of Genius,* Nueva York, Norton, 1997.
9. Jean-Jacques Nattiez, «Gould Hors-Temps», en *Le Combat de Chronos et d'Orphée: essais,* París, Bourgois, 1993, pp. 46-47.
10. Dietrich Fischer-Dieskau, *The Fischer-Dieskau Book of Lieder: The Texts of Over 750 Songs in German Chosen and Introduced,* Londres, Gollancz, 1976, p. 11.
11. *Ibid.*, p. 20.
12. *Ibid.*, p. 27.
13. *Ibid.*, pp. 27-28.
14. Jeremy Siepmann, «Glenn Gould and the interpreter's prerogative», *The Musical Times,* 131/1, enero 1990, p. 27.
15. Gerald Moore, *Am I too Loud? Memoirs of an Accompanist,* Londres, Hamish Hamilton, 1962, p. 178; subrayado mío.

16. Esta noción se analiza ampliamente en Leonard Meyer, *Music, the Arts, and Ideas,* Chicago, University of Chicago Press, 1967.
17. Gunther Schuller, *The Compleat Conductor,* Nueva York, Oxford University Press, 1997, p. 4.
18. *Ibid.*, pp. 109-124.
19. Norman Del Mar, *Conducting Beethoven,* Oxford, Oxford University Press, 1992.
20. Gunther Schuller, *op. cit.*, pp. 538-543.
21. Alfred Brendel, *Musical Thoughts and Afterthoughts,* Londres, Robson, 1976, p. 25; subrayado mío.
22. Geoffrey Norris, *Rachmaninoff,* Londres, Dent, 1993, p. 78; subrayado mío. Esta cita también es analizada brevemente en Jonathan Dunsby, *op. cit.*, pp. 93-94. Cfr. los comentarios de Fanny Waterman, citados en la página 132 de este libro.
23. Las Danzas sinfónicas y el Cuarto concierto para piano son ejemplos del modernismo compositivo de Rachmaninoff, no según los modelos de un Schoenberg o un Stravinsky, claro está; sin embargo, hoy en día es obvio que Rachmaninoff no era un compositor «conservador». Incluso Arnold Whittall, en su investigación sobre la historia esencial del modernismo del siglo XX, elogia la música de Rachmaninoff por ser «completamente convincente en su revitalización de las esencias de las tradiciones» (*Musical Composition in the Twentieth Century,* Oxford, Oxford University Press, 1999, p. 37).
24. Aunque este importante axioma quizá no necesite referencias, es una frase conocida que dijo de pasada el pianista y director Murray Perahia durante un ensayo en Londres con la Academy of St. Martin in the Fields, en abril del año 2000. En el capítulo 3 de este libro se amplían opiniones similares.
25. «The Largo of Bach's Sonata No. 3 for Solo Violin [BWV 1005], traducción de John Rothgeb, en Heinrich Schenker, *The Masterwork in Music: A Yearbook, vol. 1, 1925,* William Drabkin (ed.), Cambridge, Cambridge University Press, 1994, p. 37. El título genérico de los brillantes ensayos de Schenker es *Das Meisterwerk in der Musik.* Cfr. el análisis de la importancia de la preparación mental en el capítulo 9.
26. Heinrich Schenker, *op. cit.*, p. 38.
27. «Further consideration of the Urlinie: I», traducción de John Rothgeb, en *ibid.*, p. 109.
28. Dos de estos intérpretes también eran compositores —aunque evidentemente resulta raro decir que Mozart «también» era compositor.

## *Lecturas complementarias*

Brendel, Alfred: *Musical Thoughts and Afterthoughts,* Londres, Robson, 1976.

Dunsby, Jonathan: «Performance», en Stanley Sadie y John Tyrrell (eds.), *The New Grove Dictionary of Music and Musicians,* 2ª ed., 29 vols., Londres, Macmillan, 2000, vol. XIX, pp. 346-349.

— *Performing Music: Shared Concerns,* Oxford, Clarendon Press, 1995.

Moore, Gerald: *Am I too Loud? Memoirs of an Accompanist,* Londres, Hamish Hamilton, 1962.

Page, Tim (ed.): *The Glenn Gould Reader,* Nueva York, Knopf, 1984.

Schenker, Heinrich: *The Art of Performance,* Heribert Esser (ed.), Irene Schreier Scott (trad.), Nueva York y Oxford, Oxford University Press, 2000.

# *Índice analítico*


Abel J. L., 204, 212
Adam L., 101, 104, 108, 109
Adler, R. H., 213
Agras, W. S., 210, 214
Agricola, J. F., 41, 53, 273
Aiello, R., 145-146, 148, 150
d'Albret, J., 115,
Allard, F., 151
Alsager, T. M. 252,
Anderson, E., 33, 53
Anderson, J. R., 123, 134, 151
Annable, L., 231
Ansdell, G., 123
Antheil, G.
*Ballet mécanique,* 30
Appel, S. S., 213
Appleton, L. J., 92, 196, 197, 198
Aristóteles, 22, 247
Aristoxeno, 22
Armstrong, 113-114, 117, 119-120, 123
Atik, Y., 197
Atkinson, R. C., 122
Atkinson, R. L., 122
Agustín, san, 218

Bach, C. P. E., 26, 39, 97, 99, 107, 108, 127, 131, 135
Bach, J. S., 20, 23-24, 28, 29, 42, 43, 50, 52, 53, 83, 92, 266-267
*Clavierübung,* Parte II, 38, 39
*Concierto italiano,* 148
Concierto para clave en Re mayor BWV 1054, 50
Concierto para clave en Sol menor BWV 1058, 50
Concierto para violín en La menor BWV 1041, 30
Concierto para violín en Mi mayor BWV 1042, 30
*Klavierbüchlein für WilhelmFriedemann Bach,* Applicatio BWV 994, 39-40
«Overture nach Franzöischer Art» BWV 831, 38, 40
*Pasión según San Juan, La,* 50
*Pasión según San Mateo, La,* 28, 223; *«Erbarme dich»,* 228, 246
Sonata para violín solo núm. 3 en Do mayor BWV 1005, 272, 274
Suite francesa núm. 6 en Mi mayor BWV 817, 168
Bach, W. F., 39
Baddeley, A. D., 92, 151
Baillot, P., 101, 103, 107, 108, 174
Baily J., 89, 92, 93

Baird, J. C ., 53, 273
Bartha, D., 53
Bartók, B., 85, 92, 138
Bashmet, Y., 232
Baudiot, N., 101
Beatles, The
*Sgt. Pepper's Lonely Hearts Club Band*, 30, 33
Beattie, G., 181
Beecham, T., 243
Beethoven, L. van, 27, 28, 32, 33, 47, 48, 49, 50, 53, 84, 99, 100, 102, 110, 138, 156, 163, 164, 237, 249, 256, 257, 265, 274
Bagatelles op. 126 núms. 2 y 5, 56
Concierto para piano núm. 3 en Do menor op. 37, 48
Concierto para piano núm.1 en Do mayor op. 15, 33
Cuarteto de cuerda en Fa mayor op. 135, 237-240
Sinfonía núm. 1 en Do mayor op. 21, 33
Sinfonía núm. 5 en Do menor op. 67, 243, 257-258, 269
Sinfonía núm. 9 en Re menor op. 125, 256-257
Sonata para piano en Lab mayor op. 110, 161-166, 170-171
Sonata para piano en Mib mayor op.31 núm.3, 226
Trío en Sib mayor op. 11, 50
Bell, C., 53
Bem, D. J., 122
Bent, M., 32, 33
Berg, A., 171
Berglas, S., 211
Berlioz, H., 253, 260
Bernardi, F., 41
Bernstein., S., 138, 150, 151
Berry, W., 56, 79, 80, 135
Bertolli, F., 41
Biancolli, L., 273
Biber, H.
«Sonata del Rosario» núm. 7, 37
Bischoff, H., 109
Blanchard, H., 253
Blaze, F. H. J., 253
Blum, D., 196, 198
Boissier, Mme. A., 110
Bonell C., 259
Booth, V., 128, 135
Botte, A., 253
Boulez, P., 259
Bowen, J. A., 58, 237, 246, 265, 273
Boyd, M., 53
Brahms, J., 20, 28, 52, 80, 249, 250, 256, 270
*Fantasien* op. 116, 80
Brantigan, C. O., 213
Brantigan, T. A., 213
Brendel, A., 106-107, 110, 232, 259-260, 270, 274
Brinner, B., 92
Broadwood, J., 49
Brodsky, W., 205, 212
Brotons, M., 205, 212
Brown, A. P., 47, 53
Brown, C., 54
Brown, H. M., 33, 54
Bruch, M., 256
Concierto para violín, 256
Bull, S. J., 197
Bülow, H. von, 138
Burney, C., 26-27, 33
Burrows, D. L., 53

Cage, J.
*Variations* IV, 31
*Water Walk*, 37
Callas, M., 199, 233
Calvocoressi, M. D., 273
Cardus, N., 261
Cartier, J.-B., 108
Caruso, E., 199, 236, 241-242, 246, 247
Casadesus, R., 235
Casals, P., 199, 247
Castil-Blaze, *véase* Blaze, F. H. J.
Catel, C.-S., 101
Cavicchi, D., 124
Chaffin, R., 92, 94, 148, 151
Chanan, M., 228, 229, 246, 247
Chase, W. G., 123,150, 151, 151
Cherubini, L., 28
Ching, J.,129, 135
Chissell, J., 259
Chopin, F., 62, 69, 71, 75, 77, 92, 102
Estudio en Mi mayor op. 10 núm.3, 93
Estudios op. 10, 167, 171

Nocturno en Do♯ menor op. 27 núm. 1, 55, 62, 63-78
Nocturno en Re♭ mayor op. 27 núm. 2, 63, 69
Preludio en Mi menor op. 28 núm. 4, 59
*Sonata para violonchelo en Sol menor op. 65*, 190-194
Cimarosa, D., 28
Clark, D. B., 210, 214
Clarke, E. F., 59, 92, 93, 94, 150, 151, 181, 196, 198, 228
Clayton, A.,196
Clementi, M., 102, 109
Clynes, M., 92, 93
Coimbra, D. C. C ., 181,123
Collier, J. L., 123
Colman, A. M., 197
Comettant, O., 108
Cone, E. T., 56, 79, 80, 171, 245, 261
Conlon, D. E., 198
Cook, N., 58, 79, 80, 94, 146, 150, 228, 245, 246, 247
Cooper, G., 213
Corelli, A., 26
Correia, J. S., 181
Cortot, A., 80
Couperin, F., 36, 52
Cox, W. J., 205, 212
Craig, K. D., 201, 211, 214
Cramer, J. B., 102
Craske, M. G., 201, 211
Crawford, M., 92, 93, 151
Cross, I., 92, 93
Crutchfield, W., 247
Csikszentmihalyi, M., 181
Culshaw, J., 234, 246
Cutting, J. E., 176, 181
Cyr, M., 54
Czerny, C., 47, 99, 101, 102, 104-105, 108, 109

Dahlhaus, C., 42, 53
Dalcroze, *véase* Jacques-Dalcroze, É.
Dalley, J., 197
David, H.T., 53
Davidson, J. W., 90, 94, 122, 123, 124, 175, 176, 180, 181, 188, 197
Davidson, P. O., 214
Davis, T. L., 213
Davison, A. T., 33
Davison, J. W., 254-255
Dawson, J. C., 181
Day, T., 229, 247
Dean, W., 52
Debussy, C.
«Feux d'artifice», 129
«La Cathédrale engloutie», 58
«L'ombre ddes arbres», *Ariettes oubliées*, 160-163
*Pelléas et Mélisande*, 235
DeFries, J. C., 122
Del Mar, N., 270, 272, 274
Deliège, I., 123, 124
Desain, P., 93
Désormière, R., 235
Deutsch, D., 94
Dodson, J. D., 202, 203, 204, 211
Dolder, M., 213
Dolmetsch, A., 29
Donington, R., 54
Donizzeti, G.
*Luccia di Lammermoor*, 254
Dorny, S., 33
Dowland, J.
«Lachrimae», 37
Downes, O., 257
Drabkin, W., 274
Dreyfus, L., 54
Druce, D., 106, 110
Dunsby, J., 16, 56-57, 79, 80, 94, 107, 110, 196, 273, 274
Dussek, J. L., 102
Dvorák, A.
Concierto para violonchelo en Si menor op. 104, 247

Ehrlich, C., 108
Ehrlich, H., 258
Einstein, A., 258
Eisenberg, E., 247
Ekman, P., 181
Elgar, E., 20, 244
Concierto para violín en Si menor op. 61, 244
Ellis, Alan., 211
Ellis, Andrew, 181
Elman, M., 247
Enescu, G., 104

Enrique IV, rey de Francia, 115
Epstein, D., 92
Ericsson, K. A., 92, 116, 123, 126-127, 134, 136, 150, 151
Ernst, H. W., 251, 260
Evans, A., 122, 124, 182
Everist, M., 58, 79, 80, 94, 246, 272

Fay, A., 110
Fazey, J. A., 203, 212
Feather, N., 124
Federico el Grande, rey de Prusia, 99, 126
Festetics, conde G., 27
Fétis, F.-J., 108 , 109, 252, 255
Fidler, H., 211, 212
Field, B., 252
Field, J., 109
Finch, H., 259
Finck, H. T., 257
Fischer, L., 214
Fischer-Dieskau, D., 266, 267-269, 273
Fishbein, M., 211, 213
Fitts, P. M., 128, 135
Fitzgerald, D. F. P., 213
Flagstad, K., 235
Flesch, C., 104
Forsyth, M., 33
Francès, R., 229
François-Sappey, B., 109
Friesen, W. V ., 181
Frykholm, G., 178, 181
Fuller-Maitland, J. A., 53
Fürstenberg, príncipe von, 50
Furtwängler, W., 221, 235

Gabrieli, A., 223
Gabrielsson, A., 92, 94, 196, 198
Gaisberg, F., 247
Galliard, J., 134
Gallwey, T. C., 213
Gardeton, C., 108
Gardiner, J. E., 265
Gasparini, F., 41
Gaver, W., 229
Gelatt, R., 246
Gellrich, M., 178, 181
Genzinger, M. A. von, 45-46, 49, 53
Géoffroy, J.-L., 252
Gerle, R., 129, 135
Germann, S., 53
Gershwing, G., 29
Godlovitch, S., 110, 273
Godowsky, L., 199
Goetz, H., 256
Goldberg, L., 109
Goodman, E. C., 197
Goosens, L., 241, 243, 247
Gorton, T. L., 213
Gould, D., 202
Gould, G., 61, 219, 265-269, 273, 274
Graf, M., 260
Grappelli, S., 264
Green, B., 213
Greenfield, D., 213
Gregorio Magno, papa, 23
Grétry, A. E. M.
*Zémire et Azor*, 250
Grieg, E., 85
Grishman, A., 213
Grout, D. J., 32, 33
Grove, G., 28, 33
Gruson, L. M., 93, 117, 123, 129, 135,

Haggh, R. H., 134
Hall, M., 244, 247
Hallam, S., 93, 117, 118, 120, 123, 124, 131, 133, 135, 144-146, 148, 150
Hallé, C., 138
Hamann, D. L., 205, 212
Haendel, G. F., 25, 27, 28, 29, 40, 41, 53, 250
*Judas Maccabeus*, 252
*Rinaldo, K*, 40, 41
Hanser, S. B., 205, 212
Hanslick, E., 28, 105, 110, 249, 255, 256, 258, 261
Hardenberger, H., 232
Hardy, L., 202, 203, 212
Hargreaves, D. J., 94, 124, 180, 182, 197, 214
Harvey-Jones, J., 197
Haskell, H., 33, 54, 261
Havas, K., 213
Hawes, M., 252
Haydn, F. J., 26, 28, 45-47, 49, 50, 53, 100, 257, 273
Sonata en Mib mayor Hob. XVI:49, 45-47

Sonata en Mib mayor Hob. XVI:52, 49
Sinfonía núm. 104 en Re mayor, 252-253
Heine, H., 252, 253
Henderson, W. J., 249, 257, 261
Hendrix, J., 264
Henschel, G., 256-257
Henselt, A., 109
Hepper, P. G., 123
Hermelin, B., 93
Hess, M., 138
Hildebrandt, Z., 39
Hilgard, E. R., 122
Hill, P., 171
Himle, D. P., 214
Hindemith, P., 31, 33, 245
Hofmann, J., 105-106, 110
Holleran, S., 123
Holliger, H., 232
Holzbauer, I.
*Günther von Schwarzburg*, 250
Honing, H., 93
Horan, J. J., 209, 214
Horenstein, J., 259
Horowitz, J., 246, 247
Horowitz, V., 199
Howat, R., 110
Howe, M. J. A., 122, 123, 124
Howell, P., 92, 93
Howell, T., 110, 228
Hudson, J. A., 213
Hueffer, F., 254
Hughes, E., 138, 139, 143-144, 145-146, 150
Hummel, J. N., 102, 104, 105, 109, 126, 127, 134,
Concierto para piano en La menor op. 85, 104

Ibsen, H., 256
Imreh, G., 92, 93, 94, 148, 150, 151
Irvine, J. K., 213

Jackendoff, R., 175, 181, 228
James, I., 213
Jacques-Dalcroze, É., 30, 122
Jenkins, G., 109
Jin, Y. C., 212
Joachim, J., 20, 103, 241, 247, 256, 260
Johnson, J. H., 229
Johnson, P., 222, 228, 245, 246, 247
Jones, E. E., 211
Jones, G., 202
Jones, M. R., 123
Jørgensen, H., 136
Joseffy, R., 255
Joseph, N., 213
Juan XXII, papa, 23
Juslin, P., 226, 229

Kalkbrenner, E., 102, 105, 127, 135
Kappeli, L., 213
Karajan, H. Von, 265
Kaser, H. E., 213
Kell, R., 243
Keller, H., 227, 229, 264, 271, 273
Kelly, M., 273
Kemp, A., 115, 123, 212
Kenardy, J., 205, 212
Kendrick, M., 209, 214
Kenyon, N., 33, 54, 107, 110, 273
Kerman, J., 14, 79
Kern, J., 29
Kintsch, W., 150, 151
Kivy, P., 52, 54, 110, 273
Klemperer, O., 259
Knyvett, W., 252
Koster, J., 53
Kostis, J. B., 213
Kozlowsky, L. T., 181
Kraft, N., 251
Krampe, R. Th., 92, 123, 126, 134, 136
Kreisler, F., 103
Kreutzer, R., 101
Kullak, A., 109
Kullak, T., 109

Lachnith, L.-W., 101, 108
Lam, B., 37
Lamb, Kenney, C., 252
Lamontagne, Y., 213
Lang, P. J., 211
Larkin, K. T., 204, 212
Lavallee, Y.-J., 213
Lawson, C., 16, 34
Lawson, D., 214
LeBlanc, A., 204, 212
Lecanuet, J.-P., 122

Legge, W., 243, 247
Leglar, M. A., 204, 212
Lehmann, A. C., 136, 150
Lehrer, P. M., 213, 214
Leibovitch, K. A., 171
Lerdahl, F., 228
Leschetizky, T., 105, 107
Lester, J., 80, 245, 247
Levasseur, N., 101
Levin, J., 33
LeVine,W. R., 213
Levinson, J., 224, 228
Lindley, M., 53
Liszt, F., 28, 102, 105, 137, 253
Loehlin, J. C., 122
Loewengard, M., 258
Loft, A., 188, 197
Long, M., 52
Luis XIV, rey de Francia, 26
Lully, J.-B., 26

Ma, Y. Y., 232
MacClintock, C., 54
Mahler, G., 259
  Sinfonía núm. l, 259
  Sinfonía núm. 9, 259
Marcelo II, papa, 23
Marchant-Haycox, S. E., 212
Marek, G. R., 92
Marsick, M., 103
Martin, J.-B., 28
Marx, A. B., 260
Matthay, T., 143-144, 145, 146, 150, 151
Mattheson, J., 250
Matthew-Walker, R., 229
McCune, B. A., 213
Meichenbaum, D., 214
Mellers, W., 221, 228
Mendel, A., 53
Mendelssohn, F., 28, 108, 256, 260
Menuhin, Y., 171
Mercury, F., 264
Messiaen, O., 158, 171
  *La Fauvette des jardins,* 170
  *Vingt regards sur l'Enfant Jésus,* 170
Meyer, L., 78, 274
Michon, J. A., 228
Middelstadt, S. E., 211
Mietke, M., 39
Miklaszewski, K., 93, 117, 123, 129, 135, 136
Mitchell, W., 108, 135
Mizler, L., 250
Monteverdi, C., 25
Moore, A. F., 33
Moore, B., 229
Moore, D. G., 123, 124
Moore, G., 155-157, 165, 168, 171, 269, 273, 274
Morley, T., 28
Moscheles, I., 102, 109
Mozart, L., 26, 33, 50, 53
Mozart, W. A., 19, 27, 28, 32, 33, 43, 45, 47, 48, 50, 53, 54, 100, 102, 105, 116, 258, 264, 272, 273, 274
  Concierto para piano en La mayor K. 488, 50, 166, 167
  Concierto para piano en Si♭ mayor K. 450, 43-45
  Cuarteto de cuerda núm. 22 en Si♭ mayor K. 589, 253
  *Die Entführung aus dem Serail,* 250
  *Don Giovanni,* 20
  *Le Nozze di Figaro,* 45
  Sinfonía en Re mayor («París») K. 297, 19
  Sonata para piano en Do mayor K. 545, 60-62
  Variaciones sobre un minueto de J. C. Fischer K. 179, 54
Murningham, J. K., 198
Musgrave, M., 80
Myers, H., 34

Nagel, J. J., 214
Narmour, E., 56, 78
Nattiez, J.-J., 266-267, 273
Neftel, K.A., 213
Neisser, U., 92
Neuhaus, H., 171
Newell, A., 134
Newman, E., 249, 257, 258-259, 261
Newman, W.S., 110
Nichols, R., 246
Nielsen, M., 213
Nikisch, A., 243
Norris, G., 274
North, A. C., 94, 124, 180, 181, 197, 214
Novello, S., 135

Rothgeb, J., 274
Rothstein, W., 80
Roxburg, E., 247
Rubin-Rabson, G., 92
Rubinstein, A., 105
Runeson, S., 178, 181

Sacchini, A., 26
Sadie, S., 33, 53, 274
Salaman, E., 180
Salieri, A., 28
Salmon, P. G., 211, 214
Salomon, J. P., 27
Salvi, L., 254
Sándor, G., 131, 135
Sarasate, P., 247
Sausay, E., 109
Savage, I., 213
Savan, A., 123
Scarlatti, A., 41
Scarlatti, D., 53
Schachter, C., 80
Schenker, H., 58, 79, 146, 228, 270, 272, 274
Scherchen, H., 273
Schlesinger, M., 253
Schmalfeldt, J., 56, 79
Schmitt, J., 26
Schnabel, A., 107, 110, 232
Schoenberg, A., 171, 274
Schonberg, H. C., 110
Schreier Scott, I., 274
Schrodt, R., 214
Schubert, F., 157, 158
«Wandrers Nachtlied», 155-157, 158, 168
Schubert, J. F., 33
Schuller, G., 269-270, 273, 274
Schumann, C., 105, 110, 137
Schumann, R., 47, 54, 87
*Novellette* en Re mayor op. 21 núm. 4, 255
«Traumerei», 93
«Traumeswirren», 255
Schütz, H., 28
Schwarzkopf, E., 235, 243, 247
Scruton, R., 175, 181
Seashore, C. E., 82, 85, 92, 93, 246,
Seldden-Goth, G., 108
Senesino, *véase* Bernardi, F.
Sergeant, D., 124
Sessions, R., 110
Seymour, P., 110, 228
Shaffer, L. H., 82, 83, 88, 91, 92, 93, 196, 198
Shaw, G. B., 28, 250, 255-257, 258, 260, 261
Shove, P., 94
Sibelius, J., 257
Siepmann, J., 269, 273
Siivola, C., 212
Simon, H. A., 150, 151
Sinatra, F., 264
Sloboda, J. A., 84, 92, 93, 94, 108, 116, 117, 122, 123, 124, 128, 130-131, 135, 136, 150, 180, 182, 196, 198, 205, 212
Small, C., 33, 228
Smalley, D., 228
Smart, F., 33, 53
Smith, C., 167
Smith, E. E., 122
Smith, J., 151
Sobaje, M., 205, 212
Solie, R. A., 78
Solti, G., 171
Sontag, H., 254
Sousa, J. P., 29
Spitta, P., 43, 53
Spohr, L., 102, 104, 109,
Springsteen, B., 124
Stadlen, P., 171
Stadler, A., 27, 33
Stanton, H. E., 213
Starkes, J. L., 151
Steane, J., 236, 246
Stein, E., 171
Steinberg, R., 94, 123
Steptoe, A., 203, 211, 212, 214
Stevens, C., 213
Stillings, F. S., 53
Stockhausen, K.
*Gruppen,* 223
*Klavierstück VII,* 169
*Klavierstück X,* 170
*Klavierstück XI,* 86
Stokowski, L:, 20, 234
Storr, A., 229
Stowell, R., 16, 34, 54, 109, 180
Stradivari, A., 49

Strauss, J. N., 80
Strauss, S., 211
Stravinsky, I., 30, 31, 35, 52, 158, 274
*La consagración de la primavera,* 19, 30
*La historia del soldado,* 30
Strunk, O., 228
Sundberg, J., 93
Sweeney, G. A., 209, 214
Syer, J., 197
Symonds, E. R. C., 213
Szell, G., 247, 257
Szigeti, J., 33

Tachinardi-Persiani, F., 254
Tamagno, F., 236, 246
Tartini, G., 39, 52
Taruskin, R., 32, 33, 52, 54, 107, 110, 273
Tchaikovsky, P., 249
Tesch-Römer, C., 123, 126
Tétreault, L., 213
Thalberg, S., 102
Thomas, W., 94, 181, 228
Thompson, O., 261
Thompson Davison, A., 33
Thomson, V., 257-258, 261
Tilmant, T., 253
Todd, N. P. McA., 87, 89-90, 93, 94
Toscanini, A., 52, 92, 221, 246, 247
Tosi, P. F., 41, 126, 130, 134, 135, 273
Towsend, E., 92
Travis, E., 252
Tree, M., 184
Trevarthen, C., 123
Triplett, R., 213
Truslit, A., 89, 94, 174, 180
Tulving, E., 151
Türk, D. G., 126, 134
Tynan, K., 30
Tyrrell, J., 274

Underwood, G., 92

Valentine, E. R., 151, 213
Valentine, J., 151
Vallotti, F., 53
Varèse, E., 30
Verdi, G., 29, 236, 247
Verlaine, P., 160-161
Viotti, G. B., 103, 109
Virdung, S., 24
Vivaldi, A., 25

Wagner, R., 29, 249, 253, 255, 256
«El Anillo del Nibelungo», 234, 246
«Tristán e Isolda», 235
Walden, V., 109
Walk, R.D., 181
Walker, J., 92, 93
Walter, B., 259
Waterman, F., 132, 135, 274
Waterman, M., 190, 197
Waters, A. J., 92
Weber, C. M. von, 105
Webern, A., 158, 171
*Variationen für Klavier* op. 27, 158, 171
Wesner, R., 213
Wessel, D., 181
West, R., 92, 93
White, A. L., 228
Whiteman, P., 29
Whittall, A., 274
Wicinski, A. A., 132, 135
Wieniawski, H., 103
Williamon, A., 85, 92, 150, 151
Williams, P., 53
Wills, G. I. D., 213
Wilson, G. D., 94, 181, 203, 204, 212, 214
Windsor, W. L., 92, 196, 198
Winter, P., 28
Winter, S., 53
Wolf, E. W., 26
Wolf, H., 268
Wolfe, M. L., 211, 212, 213
Wolff, C., 53
Woodley, R., 245
Wright, R ., 214

Xenakis, I.
*Evryali,* 170

Yerkes, R. M., 202-203, 204, 211
Young, V. M., 197
Ysaÿe, E., 247

Zappa, F., 247
Zelter, C. F., 108
Zender, H., 228, 246, 247
Ziporyn, E., 228

3422102

PVP
21,00 €

Esta accesible guía para estudiantes, profesores e intérpretes de todos los niveles revela las complejidades de la interpretación musical y se centra en los aspectos fundamentales del aprendizaje, la ejecución y la manera en que escuchamos música y respondemos a ella. En primer lugar, un recorrido por la interpretación a través de la historia sirve de introducción a los conceptos analíticos y psicológicos básicos. A continuación, tras el análisis de la enseñanza, el desarrollo de la habilidad interpretativa, la práctica y la memorización, aborda la «traducción» de la partitura a sonidos, la imagen personal, la interpretación en grupo y el miedo escénico. Por último, trata el acto de escuchar, el legado de las grabaciones, la crítica musical y la opinión de los propios intérpretes sobre la interpretación. Evitando concentrarse en un único instrumento o género interpretativo, y rehuyendo las complejidades innecesarias, este libro ofrece abundantes consejos prácticos y arroja nueva luz sobre los principios y procedimientos que subyacen en la interpretación y su estudio.

John Rink es profesor de Música en el Royal Holloway, de la Universidad de Londres.

ISBN 84 206-6495-2

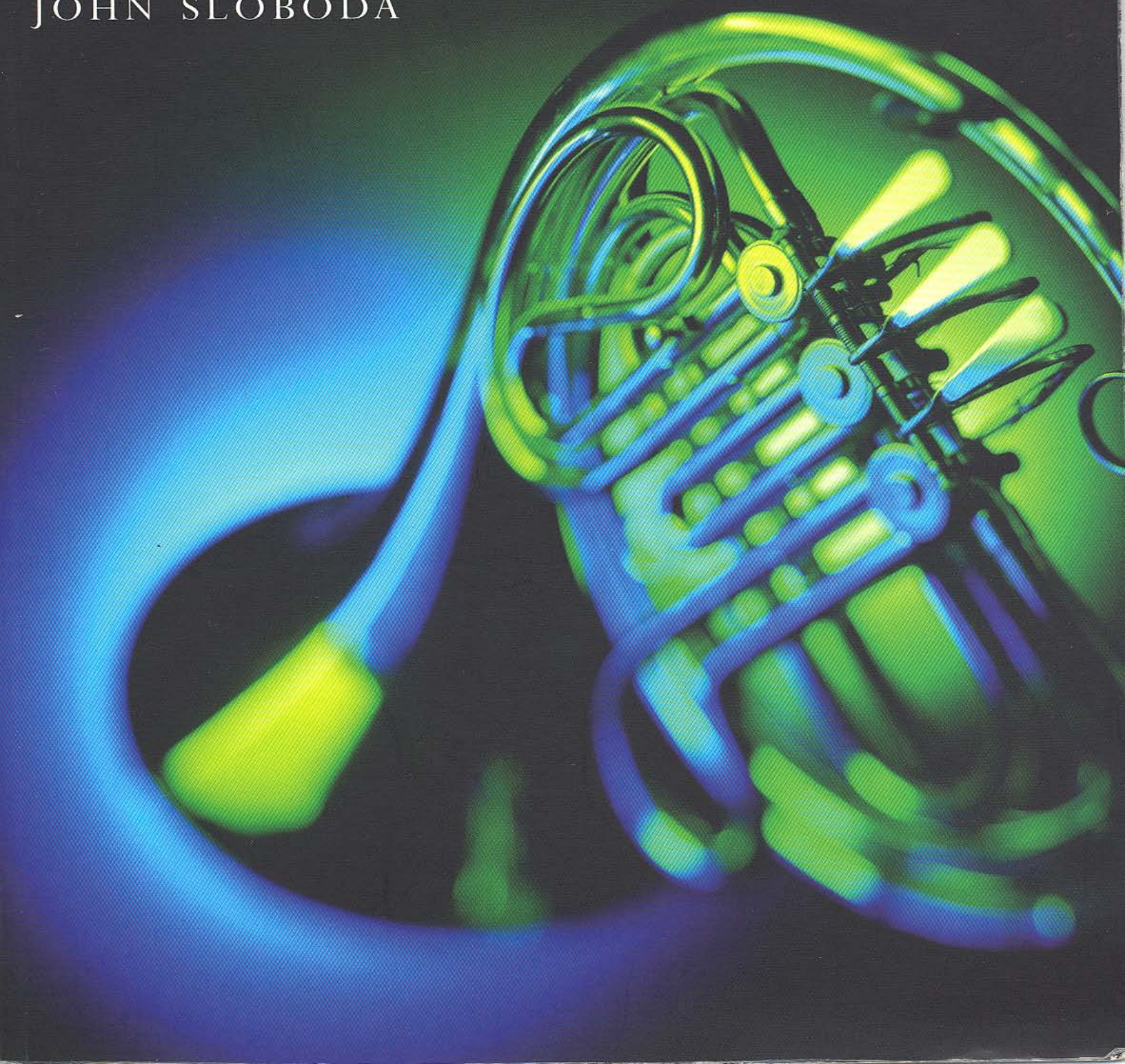



Alianza Editorial

# exploring the musical mind

COGNITION | EMOTION | ABILITY | FUNCTION

JOHN SLOBODA



# EXPLORING THE MUSICAL MIND

## cognition, emotion, ability, function

John Sloboda
*Professor of Psychology*
*Keele University*
*UK*

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research,
scholarship, and education by publishing worldwide in

Oxford New York
Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai
Nairobi São Paulo Shanghai Taipei Tokyo Toronto

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

© John Sloboda 2005

The moral rights of the author have been asserted

Database right Oxford University Press (maker)

First published 2005

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any
means, without the prior permission in writing of Oxford University
Press, or as expressly permitted by law, or under terms agreed with
the appropriate reprographics rights organization. Enquiries
concerning reproduction outside the scope of the above should be
sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this book in any other binding or cover
and you must impose this same condition on any acquirer

A catalogue record for this title is available from the British Library

Library of Congress Cataloging in Publication Data
(Data available)

ISBN 0 19 853012 9 (Hbk)
ISBN 0 19 853013 7 (Pbk)

10 9 8 7 6 5 4 3 2 1

Typeset by Cepha Imaging Private Ltd., Bangalore, India.

Printed in Great Britain

on acid-free paper by Biddles Ltd, King's Lynn

# PREFACE

## 1 Why this book

I first proposed this book to Oxford University Press for two reasons, one positive, one negative. The positive reason was that I wanted to bring together in one place a number of chapter-length pieces I had written about music over the years. These pieces all shared some of the following characteristics:

- They went beyond reporting of specific research studies to reflect on wider issues, of theory, methodology, or practical application.
- They were originally published in sources which have become hard to access (for example in books now out of print)
- They complemented, or went beyond, the issues and preoccupations of my only other 'single-author' book—*The Musical Mind*, first published in 1985.
- I had been told, by colleagues or readers, that they contained material of current value, which merited their reprinting in more accessible form.

This is a 'solo' book, in that almost all the contributions have me as sole author. However, I do see this book as in a very real sense 'belonging' to all those who have shared and supported my intellectual journey. What I have achieved would be impossible without their help. I have felt it important to fully acknowledge all who have contributed to my work, and I do this both within the acknowledgments section on pages xvii–xxiii, and also by including a complete bibliography of my academic publications (pp. 421–429), which explicitly records the contributions of each of my co-authors in the proper way.

The negative reason for this book was that, at the time I first made the proposal to Oxford University Press in early 2003, I sincerely believed that this book might be among my last published contributions to the field of music psychology. I felt that there might be more pressing calls on my time and intellectual energies for the foreseeable future.

In thinking about how to frame and introduce this set of readings, I have felt compelled to attempt to articulate, and possibly achieve some

reconciliation of, these positive and negative motivations. This is because the issues I have faced, as an individual scientist and writer, are issues which face many intellectual and professional people in one way or another. In exploring and debating these issues my goal is to assist anyone with an interest in music and the mind to situate their interest within the broadest possible context. These issues are, at root, about the responsibility of scientists, writers, and practitioners as citizens, and how we remain true to our wider responsibilities to society within our chosen specialist paths. Rather than have these broader concerns skew this preface out of shape, I have decided to include my reflections on these matters within a specially written chapter not published elsewhere, which appears as Chapter 23.

## 2 Structure, organization and context

The chapters are organized into four main Parts, which represent four distinct areas in which I have worked. Each of these areas would be generally accepted as core to an understanding of the musical mind.

**Part A (Cognitive processes)** starts with research on cognitive processes in music reading, and spreads out from reading to the activities with which reading normally interacts, namely, perception and memory of music, and music performance. The core concept binding together the contributions in this part is that of 'mental representations'—categorical and rule-based abstractions from the complex musical surface. The nature and functioning of these representations is the central concern of mainstream cognitive psychology. However, the last two contributions in this section demonstrate how concern with cognition (particularly semantics) leads inevitably to a consideration of the role of the emotions in music cognition.

Perception and cognition remains the core concern of English-language music psychology as evidenced by the preponderance of books and journal articles addressing this sub-area (see, for instance, Krumhansl, 1990, 1991; McAdams and Bigand, 1993; Deliège and Sloboda, 1997; Palmer, 1997; Deutsch, 1999). The chapters of this book are not intended to provide a comprehensive overview of the state of the field, rather they illustrate core concerns through issues (reading, memory recall, semantics) which have remained, by and large, rather under-studied, despite their intrinsic interest. Two current growth areas in music cognition are the study of brain mechanisms (e.g. Peretz and Zatorre, 2003) and modelling of music perception and cognition through computation and artificial intelligence (e.g. Tillman, Bharucha, and Bigand, 2000). Whilst acknowledging the vibrancy of these approaches, this book remains firmly within the core psychological tradition, where overt human behaviour is the central

form of data, and explanatory frameworks are developed in primarily intentional and functional terms (Sloboda and Juslin, 2001).

**Part B (Emotion and motivation)** examines how emotion impacts at every stage of the musical chain, from composer, through performer, to listener. It also addresses the question of how emotional responses to music develop through childhood, and how differences in emotional musical biographies can lead to gross differences in motivation for musical activity across the lifespan. A key concept of this Part is that of the 'peak experience', a high-intensity psycho-physiological response to music which has both immediate and long-lasting effects, and the attaining of which is a prime reason why many people engage with music. The research on emotion reflects an increasing engagement of the research community in the 1990s (see Juslin and Sloboda, 2001). It is now generally accepted that only a joint consideration of the 'cold' processes of cognition and the 'hot' processes of emotion and motivation allows proper understanding of skill development and ways in which music is used in society. Motivation has become a particularly key concept in research on music education and processes of learning (O'Neill and McPherson, 2002).

**Part C (Talent and skill development)** focuses on the psychological processes underlying the acquisition of different levels of musical skill, with a specific emphasis on performance skills within the classical conservatoire tradition. The question of whether there are innate differences in 'talent' looms large in this work, and the approach taken by my collaborators and myself has been to see how many individual differences could be accounted for by differences in learning experiences, and the motivational and social contexts in which that learning takes place (Howe, Davidson and Sloboda, 1998). A central concept of this Part is 'practice', and what its characteristics and support demands are (Ericsson, 1996).

The research on skill development is a contribution to the 'expertise' research field, pioneered by Herbert Simon, William Chase and others in the 1960s and 1970s (e.g. Chase and Simon, 1973), who provided a very clear intellectual grounding for much good contemporary work. Where our work differs from the earlier work is in its extensive use of biographical and interview methods to supplement more traditional experimental techniques. This approach has been further developed by some of the most influential contemporary studies of musical learning (e.g. Green, 2002; Chaffin, Imreh, and Crawford, 2002).

**Part D (Music in the real world)** reflects the way in which contemporary thinking about music has more fully embraced social and cultural considerations (e.g. Hargreaves and North 1997; Cook, 1998;

Small, 1998; De Nora, 2000; McDonald, Hargreaves and Miell, 2002). Musical skills (of both performing and listening) are highly dependent on the cultural and individual functions and values which music adopts within particular social and historical contexts. This section contains a range of contributions, mainly preliminary and to some extent speculative, exploring how these broader cultural considerations may impact on, and relativize, much of our scientific knowledge about musical activity. Uses of music in the home, in the street, in church, and in the concert hall, may have rather little in common with one another. A key concept for this Part is 'functionality'—what the music is *for*, in the minds of both producers and consumers.

This area of interest represents a natural intellectual progression from the concerns of the previous two Parts. Motivational structures and skill development systems are much more open to cultural influences than basic cognitive processes. However, such issues cannot be addressed solely with the intellectual resources of psychology. They open music behaviour up to the full range of the social sciences, including sociology, anthropology, and the political sciences. This implies a larger, and potentially less achievable, challenge than the previous two discontinuities, at least in relation to the constraints of an individual career within an individual discipline. These are issues that psychologists individually, and psychology in its current state of development, may simply be as yet unprepared to grasp productively.

*The Musical Mind* was deliberately open-ended. It tried not to achieve closure on any of the issues raised there, but aimed to express confidence that psychology possessed the resources to make further progress on important (and answerable) questions. The subsequent decades have fully vindicated that confidence. If a book covering the topic areas of *The Musical Mind* were to be written today, there would be much newer, and better, results to report.

*Exploring the Musical Mind* is similarly open-ended. Its trajectory leads from the relatively safe and comfortable 'normal science' of the cognitive paradigm (one of the most successful and stable paradigms of 20th-century psychology) through the 'unsafe' area of emotion (one of the most problematic and 'unparadigmatic' areas of modern psychology), to questions which cannot be solved, even in principle, using the intellectual resources of psychology alone. Although I think that a reader will find, within these pages, some better answers to important questions than were available 20 years ago, my more fervent hope is that the reader will be encouraged to embrace better, and more far-reaching questions, than were typically being asked in the field at the time of the publication of *The Musical Mind*.

I do not retain the same certainty that answers to these questions will be easily obtained, but I do retain the hope that, in another 20 years, the field of music psychology will be at least as different by then as it has become in the last quarter century.

## 3 The individual chapters

The decision to include a particular chapter in this volume was influenced by a number of factors. Paramount among them was that each chapter should be complete in itself and, in some sense, a 'good read'. By this I mean that each chapter is trying to make a point, carry through an argument, and engage a reader in an implicit dialogue.

My primary selection criterion for each chapter was that it should contain some substantial and unique material, and that each chapter should contain some material not previously included within any book that I have written or edited. For instance, although Chapters 1 and 2 both report research on the psychology of music reading, Chapter 1 is unique in its discussion of educational implications, and Chapter 2 is unique in its coverage of the ability of expert sight-readers to add appropriate non-notated expression to their sight-reading. A few chapters are completely unique, in that they contain, or refer to, no significant material elsewhere within the set. However, uniqueness should not be confused with importance. Where I have considered an idea, or a set of data, to be of great importance, I have tended to refer to it repeatedly. Indeed, overlaps provide a natural way for a reader with a specific set of interests to select the subset which most interests him or her, and to navigate from one chapter to another in a way that reflects his or her priorities rather than the imposed structure of the book.

To assist the reader to navigate the book with some sense of main and recurrent themes, I have provided a table at the end of this introduction which lists the main topic areas of each chapter, and an indication of the key chapters with which it has some overlap. As good musical themes bear repetition within a coherent musical composition, I also hope that repeated themes within the book allow key ideas to be set in different contexts, where different aspects of their importance may be highlighted.

The variation in style between these chapters is less than in a multi-author volume, although writings spanning a 30-year period cannot be totally stylistically consistent. The most strikingly dated aspect of earlier chapters is the use of the masculine where non-sexist conventions would now prescribe gender-neutral terminology. After some debate, I decided that history should not be rewritten, and that the original text should

stand, partly as a witness to the social changes that have taken place in my short professional career. I have, however, taken the opportunity to update references which were 'in press' at time of original citation, so that readers wishing to locate these sources could do so on the basis of accurate information.

## References

Chaffin, R., Imreh, G., and Crawford, M. (2002) *Practicing Perfection: memory and piano performance*. Mahwah, NJ: Lawrence Erlbaum.

Chase, W. G. and Simon, H. A. (1973) The mind's eye in chess. In W. G. Chase (Ed.) *Visual Information Processing*. New York: Academic Press.

Cook, N. (1998) *Music: a very short introduction*. Oxford: Oxford University Press.

Deliège, I. and Sloboda, J. A. (Eds.) (1997) *Perception and Cognition of Music*. Hove: Psychology Press.

De Nora, T. (2000) *Music In Everyday Life*. Cambridge: Cambridge University Press.

Deutsch, D. (Ed.) (1999) *The Psychology of Music, 2nd edn. New York: Academic Press.*

Ericsson, K. A. (Ed.) (1996) *The Road to Excellence: the acquisition of expert performance in the arts and sciences, sports and games*. Mahwah, NJ: Lawrence Erlbaum.

Green, L. (2002) *How Popular Musicians Learn: a way ahead for music education*. Aldershot: Ashgate.

Hargreaves, D, J. and North, A. C. (Eds.) (1997) *The Social Psychology of Music*. Oxford: Oxford University Press.

Howe, M. J. A., Davidson, J. W., and Sloboda, J. A. (1998) Innate talent: reality or myth? *Behavioural and Brain Sciences*, **21.3**, 399–442.

Juslin, P. N. and Sloboda, J. A. (Eds.) (2001) *Music and Emotion: theory and research*. Oxford: Oxford University Press.

Krumhansl, C. L. (1990) *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press.

Krumhansl, C. L. (1991) *Music psychology: tonal structures in perception and memory*. Annual Review of Psychology, 42, 277–303.

McAdams, S. and Bigand, E. (Eds.) (1993) *Thinking in Sound: the cognitive psychology of human audition*. Oxford: Oxford University Press.

McDonald, R., Hargreaves, D. J., and Miell, D. (Eds.) (2002) *Musical Identities*. Oxford: Oxford University Press.

O'Neill, S. A. and McPherson, G. E. (2002) In R. Parncutt and G. E. McPherson (Eds.) *The Science and Psychology of Music Performance: creative strategies for teaching and learning*. New York: Oxford University Press.

Palmer, C. (1997) Music performance. *Annual Review of Psychology,* **48**, 115–138.

Peretz, I. and Zatorre, R. (Eds.) (2003) *The Cognitive Neuroscience of Music.* New York: Oxford University Press.

Sloboda, J. A. (1985) *The Musical Mind: the cognitive psychology of music.* London: Oxford University Press.

Sloboda, J. A. and Juslin, P. N. (2001) Psychological perspectives on music and emotion. In P. N. Juslin and J. A. Sloboda (Eds.) *Music and Emotion: theory and research.* Oxford: Oxford University Press, pp. 71–104.

Small, C. (1998) *Musicking: The Meanings of Performing and Listening.* Hanover, NH: Wesleyan University Press.

Tillman, B., Bharucha, J. J., and Bigand, E. (2000) Implicit learning of tonality: a self-organising approach. *Psychological Review,* **107.4**, 885–913.

| Chapter | Key topics | Chapters sharing significant concerns |
|---|---|---|
| **PART A Cognitive processes** | | |
| 1 The psychology of music reading | Structural cues within notation<br>The role of musical knowledge in reading<br>Teaching music reading | 2 |
| 2 Experimental studies of music reading: a review | Review of early reading research<br>Experiments varying skill level and structure in text<br>Expressive sight-reading | 1, 13, 15 |
| 3 The uses of space in music notation | Comparison of text and music notations<br>Historical trends in notation development<br>Conceptual issues in the design of notation<br>Psychological issues in the design of notation | |
| 4 Immediate recall of melodies | Problems of transcription from sung performance<br>Analysis of errors in sung recall<br>Memory representations for melodies | 5 |
| 5 Cognition and real music: the psychology of music comes of age | Nature of scientific paradigms<br>Is music psychology paradigmatic?<br>The contribution of Lerdahl and Jackendoff | 4, 8, 14 |
| 6 Psychological structures in music: core research 1980–1990 | Citation as a means of assessing significance of research<br>The contribution of Krumhansl<br>Establishing a sense of tonal centre<br>Hierarchies in tonal representation | |
| 7 Book review of *Language, Music and Mind* | The relationship of philosophy of music to psychology of music<br>Nuance and ineffability in music<br>Different ways in which music might be said to have 'meaning' | 8, 20 |
| 8 Does music mean anything? | Is parsing sufficient for understanding in music?<br>Dynamic feelings as precursors of meaning<br>'Thrills' as proto-emotions | 5, 7 |

| Chapter | Key topics | Chapters sharing significant concerns |
|---|---|---|
| **PART B Emotion and motivation** | | |
| 9 Music as a language | Musical phonology, syntax, and semantics<br>Peak musical experiences in childhood<br>Motivation for long-term commitment to music | 11, 14 |
| 10 Music psychology and the composer | Different models for composer-psychologist dialogue<br>Relationship of composition to functions of music<br>Music as an aid to cognitive restructuring | 20, 21 |
| 11 Empirical studies of emotional response to music | Free verbal emotional responses to music<br>Verbal responses within forced categories<br>Extrinsic and intrinsic antecedents of musical emotion | 9, 12, 13 |
| 12 Emotional responses to music: a review | Accounting for variation between and within people<br>Methods for measuring emotional response<br>How musical events elicit emotion | 11 |
| 13 Musical performance and emotion: issues and developments | Different expressive roles of performers<br>Connection between music structure and expressive devices<br>Evidence for deliberate planning of expressive communication | 2, 11 |

*Continued*

| Chapter | Key topics | Chapters sharing significant concerns |
|---|---|---|
| **PART C Talent and skill development** | | |
| 14 Musical expertise | Social and cultural relativity of musical expertise<br>Musical skill acquisition in non-instructional settings<br>Precursors of musical expertise | 5, 9, 17 |
| 15 Musical ability | Cultural exposure as a primary source of basic ability<br>Expressive playing in sight-reading as a strong test of ability<br>How to develop technical and expressive skill together | 2, 16 |
| 16 The acquisition of music performance expertise | Is musical skill inherited?<br>The role of practice<br>Different learning mechanisms for technical and expressive skill | 15, 17 |
| 17 Are some children more gifted for music than others? | The role of early learning in skill development<br>Social support mechanisms for skill development<br>Motivational factors predicting success or dropout | 14, 16 |

| Chapter | Key topics | Chapters sharing significant concerns |
|---|---|---|
| **PART D Music in the real world** | | |
| 18 Everyday uses of music | Importance of situation and volition in determining effects of music<br>Free verbal descriptions of music's role in everyday life<br>Music in public places | 10, 19, 22 |
| 19 Music: where cognition and emotion meet | Mismatch between music's value to people and their level of musical skill<br>Tracking everyday uses of music in real time<br>Societal barriers to engagement with music | 18 |
| 20 Music and worship: a psychologist's perspective | Can music have assumed common effects within a defined situation?<br>Parallels between different ways of listening to music and different aspects of worship<br>Ineffability as a core concept in both music and worship | 7, 10 |
| 21 Emotion, functionality, and the everyday experience of music | Reasons why young adolescents give up instrumental engagement<br>Cultural trends associated with disengagement from formal music<br>Responses of the education system to cultural change | 10 |
| 22 The sound of music versus the essence of music | Cultural trends which deindividuate the musical experience<br>How science can be misused in the interests of cultural homogenization of the musical experience<br>Research which demonstrates the strong contribution of the individual to the nature of the musical experience | 18 |
| 23 Assessing music psychology research | The social benefits of music psychology research<br>Differentiating the above from the social benefits of music<br>Differing levels of engagement with social issues evident in research | |

John A. Sloboda
*Keele, April 2004*

# ACKNOWLEDGEMENTS

Anniversaries are no more than arbitrary round numbers, but powerful in human affairs nonetheless. The publication date for this book, 2004, marks exactly 30 years since my first academic publication in music psychology (Sloboda, 1974). It also marks 30 years of continuous membership of the faculty at Keele University, where, under the supportive championing of successive heads of Psychology (the late Ian Hunter 1974–1982, James Hartley 1982–1992, Angus Gellatly 1992–2000) music psychology has been encouraged and enabled to flourish and develop. This book is a tribute to them, and all those who have worked at Keele, or collaborated with members of the Keele music psychology group, to establish the knowledge and activity base on which the papers in this volume are built.

## Research collaborators

My main debt is to my co-authors and co-editors: teachers, colleagues, and students. Everything in this book shows their influence. They are listed in chronological order of date of first co-publication (given), with current or most recent institutional affiliation where known. Although a few of these 25 have not collaborated directly on music projects, their influence on the way I think about many things, including music, has been palpable.

| | | |
|---|---|---|
| Andrew Gregory | (1980) | University of Manchester |
| Judy Edworthy | (1981) | University of Plymouth |
| Don Rogers | (1983) | Keele University |
| Beate Hermelin | (1985) | Goldsmiths College, University of London |
| Neil O'Connor | (1985) | Deceased |
| David Parker | (1985) | Liverpool John Moores University |
| Angus Gellatly | (1989) | Open University |
| Michael Howe | (1990) | Deceased |
| Jeff Hopkins | (1992) | Keele University |
| John McLeod | (1993) | Abertay University |

| | | |
|---|---|---|
| Rita Aiello | (1994) | Julliard School of Music, New York |
| Warren Brodsky | (1994) | Ben Guiron University of the Negev, Be'er Sheva |
| Mitch Waterman | (1994) | University of Leeds |
| Jane Davidson | (1994) | University of Sheffield |
| Irene Deliège | (1995) | University of Liège |
| Derek Moore | (1995) | University of East London |
| Eric Clarke | (1997) | University of Sheffield |
| Richard Parncutt | (1997) | University of Graz |
| Matti Raekallio | (1997) | Sibelius Academy of Music, Helsinki |
| Susan O'Neill | (1997) | Keele University |
| Peter Desain | (1997) | University of Nijmegen |
| Stephen Newstead | (1997) | University of Plymouth |
| Patrik Juslin | (2001) | University of Uppsala |
| Andreas Lehmann | (2001) | University of Würzburg |
| Antonia Ivaldi | (2001) | Loughborough University |

I would wish to draw special attention to the two deceased members of the above list. Neil O'Connor had a far greater influence on me the bare record might suggest, which shows only one co-authored publication. From 1971 to 1974 I was jointly supervised in my PhD studies by Neil O'Connor and Beate Hermelin, in the MRC Unit for Developmental Psychology at University College London. For being in the right intellectual shape to be taken on by them I must also acknowledge my debt to several members of the Psychology Department at Oxford University, where I did my first degree. Of these, my greatest debt by far is to Patrick Rabbitt, under whose apprenticeship I learned something of what it meant to be a cognitive psychologist in the mould of Pat's great 'master' and father of post-war British experimental psychology, Donald Broadbent. It was with gentle promptings from O'Connor and Hermelin that I undertook my first experiment on music cognition in 1972. My original planned research topic was the development of language reading in children, and I might still be a language specialist rather than a music specialist if they had not seen that my extracurricular musical interests and my academic research interests begged to be integrated. Their act of faith (there being almost no serious music psychology research being undertaken in British psychology departments in those days) consistently reinforced and encouraged me to take creative and integrative risks. O'Connor and Hermelin were not specialists in music, indeed, they self-confessedly professed this. They were, however,

superb scientists, and inspired supervisors. In that sense, O'Connor and Hermelin are unacknowledged co-authors on almost everything I wrote, certainly in the first 15 years of my postdoctoral career. Their elegant and deceptively simple approach to the design and reporting of experiments (see, for instance O'Connor and Hermelin, 1978) influenced my own scientific practice enormously, and their commitment to charting and enabling the intellectual capacities of a marginalised and under-valued group, the learning disabled, was an inspiration which was to directly underpin the most important and productive research collaboration of my working life, that with the late Michael J. A. Howe.

The record shows 18 co-authored publications with Mike Howe over a ten-year period between 1990 and 2000. Many more of my publications (around another 20) refer to, and build on, the work I undertook with him and our principal collaborator Jane Davidson. Mike and I first met in 1988 and I was immediately swept up in his burning passion to understand and encourage excellence and its development across the range of human abilities. Like Hermelin and O'Connor, he wished to demystify both exceptional performance and the blocks which appear to limit people from achieving excellence. But unlike them, he was comfortable in ranging far beyond the confines of the factorial experiment to pursue his wide-ranging intellectual agendas, employing survey research, biographical and other text-based methods, which were marshalled into a range of outputs of which the academic journal article was only one (see, for instance, Howe, 1990). His willingness to paint on a very broad canvas, through books and book chapters, employing rhetoric and polemic (often motivated by strong and explicit liberal and humanitarian impulses), taught and encouraged me to be more bold, inclusive, and socially engaged, in my academic work.

## Research support

My second debt goes to those people, both students and colleagues, whose efforts, although not resulting in joint publications, have been directly influential or enabling for my own work. Former students in this category include Graca Mota, Laura Neilly, Vicki De Las Heras, and Jane Ginsborg. As with many university teachers, my research progress has depended on the support of those who have provided time and resources to do research, particularly in environments away from the home base. Key research advances have been made with the help of Henry Shaffer (University of Exeter), Maria Manturszewska and colleagues at the Chopin Academy of Music, Warsaw; Anders Ericsson (Florida State University); Kari Kurkela

(Sibelius Academy, Finland); Daniel Levitin (McGill University, Montreal). Research reported in this book has been financially supported by several special research awards from Keele University, the Economic and Social Research Council of Great Britain, the Nuffield Foundation, and the Leverhulme Trust.

My work has been hugely helped by the increasingly numerous and increasingly focused outlets for music psychology. In 1974 there was only one, very young, specialist journal (*Psychology of Music*) attached to the one scholarly body that then existed: the Society for Research in Psychology of Music and Music Education, now renamed SEMPRE—Society for Education, Psychology, and Music Research (*www.sempre.org.uk*). This UK-based organization offered one of my earliest and most supportive conference platforms, and published my very first research paper. The foresight of Arnold Bentley, Desmond Sergeant, Anthony Kemp, Rosamund Shuter-Dyson, and others mainly working at Reading University and the Roehampton Institute, in seeing the need in the early 1970s for reviving music psychology research cannot be overestimated. Reviving is the correct word, because although there had been leaders of the field, working on both sides of the Atlantic, their work left little trace in the years immediately after the Second World War when behaviourism was at its height, and even laypeople knew that psychology was mainly about rats running mazes. Herbert Wing (UK) and Carl E. Seashore (USA) were arguably the last in a line of key figures beginning a century earlier with Helmholtz and Stumpf (Germany). At the time I started my researches these key figures had all but been forgotten by the 'mainstream' of contemporary British psychology.

If SEMPRE provided the organizational support for my early endeavours, then two key individuals provided the direct intellectual nourishment: the late Christopher Longuet-Higgins in the UK, and Diana Deutsch in the USA. At the time I began my research into music, the work of these two were oases in a relative intellectual desert. I avidly read every word they wrote on music (such as Deutsch, 1972, and Longuet-Higgins, 1972). Their writings proved that forward-looking scientific work of the highest order, both methodologically and theoretically, could be carried out, borrowing from and extending the tools of experimental psychology and artificial intelligence. Both people were extraordinarily generous in promoting my work. If Christopher Longuet-Higgins ever tired of writing the numerous references and testimonials I requested of him, he never showed it. Diana Deutsch was responsible for giving me my first big publishing 'break', an invitation to contribute a chapter on music

performance to her seminal and hugely influential *The Psychology of Music* (Deutsch, 1982).

My second 'break', and the one which has undoubtedly created the longest working relationship of my professional career, came when a contemporary from Oxford, Adam Hodgkin, joined Oxford University Press as a junior editor, and was asked to scout among the people he knew for 'new talent' for the press. As early as 1978, he suggested to me that OUP might be interested in a book on the Psychology of Music. At the time I put the idea to one side, but it niggled at me, and around two years later I tentatively returned to the idea. The far-sighted Science and Medicine Editor of that time, Bruce Wilcock, not only accepted the idea, but signed me up as the fifth contributor to a new, prestigious, monograph series, the Oxford Psychology Series, whose series editors were Larry Weiskrantz, Endel Tulving and Donald Broadbent. *The Musical Mind* was the result, a book which is my single most widely cited publication, and which, to my great surprise, is still selling well today. Twenty-five years, three editors, and seven books later, I view my relationship to Oxford University Press as one of the most consistent and enabling ones of my publishing career. I wish to acknowledge fulsomely the tireless and consistently pro-active support of Bruce Wilcock's two successors in the Science and Medicine department at the Clarendon Press, Vanessa Whitting and Martin Baum. Under Martin's firm direction OUP has now built up an ever more impressive music psychology list, one which shows every sign of continual growth and renewal.

Diana Deutsch was to have further profound influences on my career path (as I am sure she has had on many others). In 1983 she founded the second English language journal devoted entirely to the psychology of music—*Music Perception*. I was invited to become a contributor and a consulting editor at a very early stage. *Music Perception* gave the discipline its first journal of fully international standing (to whose standards *Psychology of Music* and other more recent arrivals have needed to conform). These two journals, together with the more recent European *Musicae Scientiae*, have provided key outlets for the research community, and are all represented in the selection of items republished here. Each journal is also associated with a scholarly body. Diana Deutsch was the founding President of the North American Society for Music Perception and Cognition (SMPC). *Musicae Scientiae*, whose founding editor is Irene Deliège, is the journal of the European Society for the Cognitive Sciences of Music (ESCOM). ESCOM, and its journal, arose as a direct response to yet another initiative of Diana Deutsch. Together with Edward Carterette, Kengo Ohgushi, and others, in 1989 she founded the world conference series known as the

International Conference on Music Perception and Cognition (ICMPC). This was the first time that the music psychology confraternity was encouraged to organize in truly global fashion. Again, through Diana's good offices, I was invited to attend the First ICMPC in Kyoto, and discussions with co-attendees led to the realisation that (a) the conference needed to circulate around the globe in a planned way on successive occurrences, and (b) that each conference required strong local and regional organisation. ESCOM was formed initially to provide an organisational host for the Third ICMPC (Liège, Belgium, 1994), although its remit is now far wider. ICMPC, and its supporting organizations, have provided the means for researchers to interact and present their work informally, prior to publication. These conferences now routinely attract several hundred papers. Almost everything I and my colleagues and students have published in the last 15 years has first been shared with colleagues (and improved through the ensuing discussion) at a meeting sponsored by, or supported by, organizations such as ICMPC, ESCOM, or SEMPRE. The ICMPC web-site provides a convenient set of links to its supporting organisations and details of conferences, past and future (*www.icmpc.org*).

The life of a researcher does not proceed purely on academic support and resources. Support of family, teachers, and friends has been key. For encouraging me equally in both music and science, and expecting much of me, I am deeply indebted to my parents, Mary Sloboda and the late Mieczyslaw Sloboda. Equally important encouragement has also come from inspired teachers at all levels prior to university. In music, my experiences with my first piano teacher, the late Phyllis Leach, fitted everything that our later research discovered about successful first teachers of the musically able. The schools I attended, Pinner Park Primary School, and St Benedict's School Ealing, provided some outstanding teachers of science, culture, and music—with skills that I am only now coming to realize are rather rare. At many times, particularly times of discouragement and stress, friends and loved ones have been there to validate and re-invigorate. Now is not the time to list those still living—they know who they are. However, 2003 saw the death through illness of a particularly close friend and intellectual companion, Ruth Clayton. As a world-class academic in her own field of genetics, and as a passionate advocate of both the arts and radical social reform, she represented for me the best of what it means to be an intellectual. Her probing but validating engagement with my intellectual journey, and her constant attempts to help me clarify the personal and political meanings of that journey, have been a profound influence on the course of my life over the last 15 years.

I dedicate this book to the memory of Neil O'Connor (1917–1997), Ruth Clayton (1925–2003), and Mike Howe (1940–2002).

## References

Deutsch, D. (1972) Effect of repetition of standard and comparison tones on recognition memory for pitch. *Journal of Experimental Psychology*, **93.1**, 156–162.

Deutsch, D. (Ed.) (1982) *The Psychology of Music*. New York: Academic Press.

Longuet-Higgins, H. C. (1972) Making sense of music. *Proceedings of the Royal Institute of Great Britain*, **45**, 87–105.

Howe, M. J. A. (1990) *The Origins of Exceptional Abilities*. Oxford: Blackwell.

O'Connor, N. and Hermelin, B. (1978) *Seeing and Hearing and Space and Time*. London: Academic Press.

Sloboda, J. A. (1974) The eye-hand span: an approach to the study of sight-reading. *Psychology of Music*, **2**, 4–10.

# CONTENTS

Permissions xxvii

**A Cognitive processes**

1 The psychology of music reading *3*
2 Experimental studies of music reading: a review *27*
3 The uses of space in music notation *43*
4 Immediate recall of melodies *71*
5 Cognition and real music: the psychology of music comes of age *97*
6 Psychological structures in music: core research 1980–1990 *117*
7 Review of *Language, Music, and Mind* by Diana Raffman *153*
8 Does music mean anything? *163*

**B Emotion and motivation**

9 Music as a language *175*
10 Music psychology and the composer *191*
11 Empirical studies of emotional response to music *203*
12 Emotional response to music: a review *215*
13 Musical performance and emotion: issues and developments *225*

**C Talent and skill development**

14 Musical expertise *243*
15 Musical ability *265*
16 The acquisition of musical performance expertise: deconstructing the 'talent' account of individual differences in musical expressivity *275*
17 Are some children more gifted for music than others? *297*

**D Music in the real world**

18 Everyday uses of music listening: a preliminary study *319*
19 Music: where cognition and emotion meet *333*
20 Music and worship: a psychologist's perspective *345*
21 Emotion, functionality, and the everyday experience of music: where does music education fit? *361*
22 The 'sound of music' versus the 'essence of music': dilemmas for music–emotion researchers *375*
23 Assessing music psychology research: values, priorities, and outcomes *395*

Bibliography *421*

Index *431*

# PERMISSIONS TO REPRODUCE MATERIAL

| | Chapter | Acknowledgement requested by original publisher |
|---|---|---|
| 1 | The psychology of music reading | Reprinted by permission of Sage Publications Ltd from *Psychology of Music* (1978) Vol. 6, pp. 3–20. |
| 2 | Experimental studies of music reading: a review | © 1984 by The Regents of the University of California. Reprinted from *Music Perception*, Vol. 2.2, by permission of the University of California Press. |
| 3 | The uses of space in music notation | Reprinted from *Visible Language* (1981), Vol. 15, pp. 86–100 by permission of Prof. Sharon Poggenpohl, Editor/ Publisher, Institute of Design, Chicago, IL. |
| 4 | Immediate recall of melodies | Reprinted from P. Howell, I. Cross, and R. West (Eds.) *Musical Structure and Cognition*. London: Academic Press, pp. 143–167, copyright © 1985, with permission from Elsevier. |
| 5 | Cognition and real music: the psychology of music comes of age | Reprinted from *Psychologica Belgica*, Vol. 26(2), pp. 199–219, with permission from the Belgian Psychological Society. |
| 6 | Psychological structures in music: core research 1980–1990 | Reprinted from J Paynter, R. Orton, T. Seymour, and T. Howell (Eds.) (1992) *A Compendium of Contemporary Musical Thought*, London: Routledge, pp. 803–839, by permission of Thomson Publishing Services. |

| | | |
|---|---|---|
| 7 | Book review of *Language, Music and Mind* | Reprinted from *Mind and Language* (1993), Vol. 9.3, pp. 377–385, with permission from Blackwell Publishing Ltd. |
| 8 | Does music mean anything? | Reprinted from *Musicae Scientiae* (*http://musicweb.hmt-hannover.de/escom/english/MusicScE/MSstart.htm*) (1998), Vol. 2.1, pp. 21–32, by permission of the European Society for the Cognitive Sciences of Music. |
| 9 | Music as a language | Reprinted from F. Wilson and F. Roehmann (Eds.) *Music and Child Development: Proceedings of the 1987 Biology of Music Making Conference.* St. Louis, Missouri: MMB Music Inc. 1989, pp. 28–43. Permission sought. |
| 10 | Music psychology and the composer | Reprinted from S. Nielzen and O. Olsson (Eds.) *Structure and Perception of Electroacoustic Sound and Music.* Amsterdam: Elsevier. 1989, pp. 3–12. |
| 11 | Empirical studies of emotional response to music | Originally published in M. Riess-Jones and S. Holleran (Eds.) (1992) *Cognitive Bases of Musical Communication.* Washington: American Psychological Association, pp. 33–46. Copyright © 1992 by the American Psychological Association. Reprinted with permission. |
| 12 | Emotional responses to music: a review | K. Riederer and T. Lahti (Eds.) *Proceedings of the Nordic Acoustical Meeting (NAM96).* Helsinki: The Acoustical Society of Finland. 1996, pp. 385–392. Permission sought. |
| 13 | Musical performance and emotion: issues and developments | Reprinted from S.W. Yi (Ed.) *Music, Mind, and Science.* Seoul: Seoul National University Press. 1999, pp. 220–238. Permission sought. |
| 14 | Musical expertise | © Cambridge University Press. Reprinted with permission from K. A. Ericsson and J. Smith, Eds. (1991) *Towards a General Theory of Expertise: Prospects and Limits,* Cambridge University Press, pp. 153–171. |

| | | |
|---|---|---|
| 15 | Musical ability | © John Wiley and Sons Ltd. Reprinted with permission from G. Bock, and K. Ackrill, (Eds.), (1993) *The Origins and Development of High Ability*, Chapter 7, pp. 106–113. |
| 16 | The acquisition of music performance expertise | Reprinted from K. A. Ericsson (Ed.) *The Road to Excellence: the Acquisition of Expert Performance in the Arts and Sciences, Sport and Games*. Mahwah, NJ; Lawrence Erlbaum Associates. 1996, pp. 107–126. Permission sought. |
| 17 | Are some children more gifted for music than others? | First published in Italian as 'Dotu musicali e innatismo?'. In J. J. Nattiez (Ed.) *Enciclopedia della Musica: Vol II Il sapere musicale*. Torina: Giulio Einaudi Editore. 2002, pp. 509–529. Not previously published in English. Permission sought. |
| 18 | Everyday uses of music | Reprinted from S.W. Yi (Ed.) *Music, Mind, and Science*. Seoul: Seoul National University Press. 1999, pp. 354–369. Permission sought. |
| 19 | Music: where cognition and emotion meet | Reprinted from *The Psychologist*. **12.4,** 450–455. Permission sought from the British Psychological Society. |
| 20 | Music and worship: a psychologist's perspective | Reprinted from T. Hone, M. Savage, and J. Astley (Eds.) *Creative Chords: Studies in Music, Theology and Christian Formation*. Leominster: Gracewing. 2002, pp. 110–125. Permission sought. |
| 21 | Emotion, functionality, and the everyday experience of music | Originally published in *Music Education Research* (2001), Vol. 3 no. 2, pp. 243–253 (*http://www.tandf.co.uk/journals/carfax/14613808.html*) and reproduced here by permission of Taylor and Francis Ltd. |
| 22 | The sound of music versus the essence of music | Reprinted from *Musicae Scientiae* (*http://musicweb.hmt-hannover.de/escom/english/MusicScE/MSstart.htm*) (2002), Vol. 6.3 (special issue on music and emotion), pp. 235–253, by permission of the European Society for the Cognitive Sciences of Music |

PART A

# COGNITIVE PROCESSES

CHAPTER 1

# THE PSYCHOLOGY OF MUSIC READING

## 1.1 Introduction

The ability to read in one's native tongue is, in most cultures, an almost essential qualification for full membership of society. Accordingly, the attention devoted to the reading process by educationalists and psychologists has been immense. The ability to read music is, if not essential, an irreplaceable asset to anyone who indulges in musical activity. Yet the amount of attention devoted to music reading by teachers, educationalists and psychologists has, on the whole, been very small, It is not the intention here to provide an historical or sociological analysis of the neglect of this aspect of musical ability, nevertheless, before proceeding, two points need to be established: firstly that influential commentators on psychological aspects of music have, in fact, had very little to say about music reading; secondly that this neglect is unjustified in consideration of the importance of musical literacy for overall musical competence.

Seashore (1938) mentions sight-reading twice in his classic book *The Psychology of Music*. The first mention is in the context of 'Twelve rules for efficient learning in music' addressed to the aspiring pupil. He says:

> Build larger and larger units. At certain advanced stages we learn by wholes, but the best rule for learning in general is to learn one small specific thing at a time; then weave these larger units together, and so on, until the task is completed. In doing this you acquire the power to learn larger and larger units. Take the analogy of learning to read. The child first learns to see individual letters, to associate these with sounds, to weave the sounds into words, the words into phrases, the phrases into clauses. . . . As he learns to read, reading becomes easier because he reads in larger and larger units. This is exactly parallel to sight reading in music. . . .

This passage is misleading on at least two counts. Firstly it is not the case that the child learning to read prose always starts with letter–sound

associations and builds upwards. In a language like English there are so many irregular words that such a method would not suffice. It appears from the errors made by beginning readers that many of them operate at a higher level of abstraction, using the printed page as an aid to extracting the sense of the passage rather than individual letters (e.g. Weber, 1970). Secondly, the parallel to sight-reading in music cannot be 'exact' because there are no exact parallels to words, phrases, clauses, and sentences which make up the larger units of language. Until the aspiring reader knows what the larger units are—chords, bars, beats, or whatever, Seashore's parallel could not be applied.

Seashore's second mention comes in a section containing advice to the teacher. He rightly stresses the intimate connection between reading and aural training, but is unhelpful on the specifics of training methods. The instruction to 'Pass by natural stages from the mechanics of sight-reading to the singing and playing for pleasure and the preparation of repertoire' rather evades the issue. If he is saying—concentrate on the music and the reading will look after itself—then we must assume that he is rejecting the possibility that reading can be taught, and suggesting that it is something that 'just happens'.

Buck (1944) does not believe that reading ability 'just happens'. On the contrary, he proposes that hard work and virtuous application is required:

> If you are a slow reader, remember that anyone can read a piece at a bar a minute, and there is no other excuse than laziness for not acquiring speed. If the child struggling with the cat sentence were to lament that she would never be able to read rapidly 'like a grown-up', you know that she is talking nonsense, and that the speed at which you and I can read English is not due to cleverness or any special gift. And the same is true, with no qualification, about the reading of music. To confess that you are a bad reader is to confess laziness. . . .

Nonetheless, he might just as well have shared Seashore's view, for he really gives no concrete advice to the learner. Laziness can only be attributed to someone who knows what he should be doing but does not do it. Buck does not exactly say what one should do to acquire speed. It may be that poor readers work very industriously at their reading, but in an inappropriate way.

Shuter (1968), in a book specifically on musical ability, does not deal with reading as a substantive issue at all, although it is mentioned briefly in a couple of places as a component of musical skill. More recently, Critchley and Henson (1977) in the introduction to a collection of invited papers on

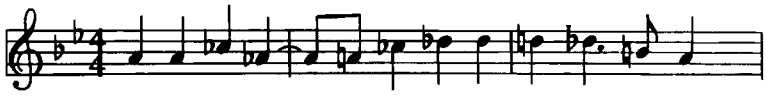
'Music and the Brain' explicitly acknowledge that a chapter on music reading is one of the major omissions of their collection. Finally, Davies (1978), in the first major text on psychology of music to be published since Shuter (1968), mentions reading only in passing. This is perhaps understandable, since the book is 'concerned almost wholly with music from the standpoint of the listener. . . . '. Experimental research into music reading has also been slight in quantity. Bean (1938), Weaver (1943), Van Neuys and Weaver (1943), Lannert and Ullman (1945), and Kwalwasser (1955) have studied various aspects of skilled sight-reading, mainly on the piano. Research carried out in an educational context has concentrated more on sight-singing, and ways of teaching it (e.g. Hillbrand, 1923; Hammer, 1963).

It is hardly necessary to state that a musician with sight-reading facility has an immense advantage over other musicians in nearly all walks of musical life. Many professional musicians could simply not perform their jobs without a high level of reading skill, and even among amateurs a good reader is more likely to be able to enter fully into the more rewarding and fulfilling aspects of musical life. It seems to be part of musical folk-lore that good sight-readers have poor musical memories, but the reason for this is probably that they do not need to develop efficient memorising strategies, rather than that they are incapable of so doing. Surely no one would wish to claim that musicians whose jobs demanded memorisation, such as opera singers, would be worse off for being good readers. After all, professional actors do not seem to suffer as a result of being able to read their scripts unaided. I wish to argue further, however, that reading facility is not simply a useful additional skill for a musician to have. It is, in a sense, necessary for full membership of the musical community.

There are basically two types of composer, the improviser and the writer. The improviser (or composer/performer) conceives a musical idea which he communicates to his listeners directly in performance. The writer communicates indirectly through a score which must be read, understood, and realised by a performer. It would not, I think, be unfair to say that the composer-writer predominates in our musical culture. Musical notation is, therefore the medium, or language, of his communication. It is not simply a set of instructions for performance (except, maybe for some forms of contemporary notation—see Karkoschka, 1972; Cole 1974)—it actually embodies aspects of the structure, or meaning, of the music which would not be present in a *physical* description of the sounds (i.e. details of frequency, duration, and amplitude of various notes).

Specifically, just knowing the pitch of a note does not tell one anything about its tonal significance; and just knowing the duration of a note does

not tell one anything about its rhythmic significance. The score makes explicit a musical relationship between notes by the use of key signatures, accidentals, and rhythmic notation, etc. Longuet-Higgins (1972) cites the following example, which would not instantly be recognised by most readers as a musically illiterate transcription of the opening few measures of the British National Anthem.



**Example 1.1**

Nonetheless, if executed by a performer, the auditory effect would be correct. All that the performer would miss would be a sense of comprehension, until he recognised the 'tune' through hearing what he played. An analogous situation in language can be obtained by asking someone to read out the following nonsensical French sentence: 'Pas de lieu Rhone que nous'. It often takes the reader some time to realise that the sounds he is producing can be interpreted sensibly as the English sentence 'Paddle your own canoe'.

When the composer of the National Anthem chose to write it in the way we are now accustomed to seeing it, it was because he wished to convey that the melody made sense conceived as being in triple time and containing the tonal sequence tonic–tonic–supertonic–, etc. This vital information is plain for any reader to see, without his having to execute a single note. When he does come to perform the notes, therefore, he already *knows* what the music means and can perform it in such a way as to make that meaning explicit. The poor reader cannot do this. He is in precisely the position of the prose reader trying to make sense of 'Pas de lieu Rhone que nous'. The meaning is not apparent to him on inspection so he has to listen to what he plays or sings and try to make sense of that. Whether he succeeds will, in part, depend upon his musicality, that is, his ability to hear what makes musical sense. But it will also depend upon something less tangible and predictable—his ability to switch between ways of looking at (or in this case hearing) a musical message.

A famous demonstration of fixation is that devised by the cartoonist W. E. Hill in 1915, entitled 'My Wife and My Mother in Law' (Figure 1.1). Many people can see only the old woman or only the young woman, even after instruction. One's first schema for viewing the picture can make it almost impossible to switch to the alternative perception. If a musician



**Figure 1.1** 'My Wife and my Mother-in-law'. A reversal figure created by the cartoonist W. E. Hill; first published in *Puck* in 1915. The young woman's chin is the old woman's nose.

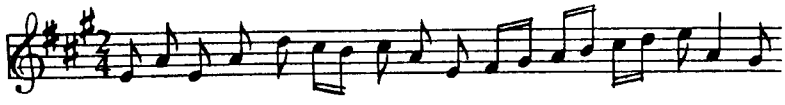
misperceives the musical meaning of a passage, he may therefore be unable to switch to the correct interpretation without difficulty.

There is, of course, an objection to this line of argument. Surely, one may argue, listeners in the concert hall experience no difficulty in perceiving the music as the composer intended. Why then, should a reader have any difficulty hearing the meaning once he has executed it? The answer to this objection lies, I believe, in a consideration of the difference of 'cognitive load' on a reader and a listener. Essentially, a listener has just one thing to do, that is, listen. A reader not only has to listen. He also has to decide what notes to play next with reference to the score, organising his motor behaviour, as well as trying to remember the aural impression created by the notes he has already played. He can often be in the worst position to appreciate the overall meaning behind what he is doing. Even so, listeners frequently do 'mis-hear' music. This is especially true, for instance, at the start of a piece, where misperceptions of rhythmic structure are quite common even among musically sophisticated listeners.

Longuet-Higgins (1972, 1976) and Steedman (1977) have studied this issue in the context of producing computer programmes which correctly transcribe

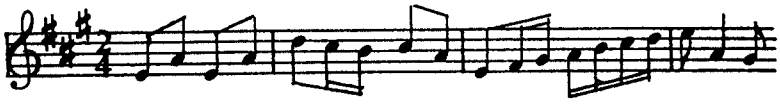
melodies, the programmes thus acting as 'intelligent' musical listeners. Using the themes of J. S. Bach's '48' fugues from the 'Well Tempered Clavier' performed in a 'dead-pan' fashion on an organ keyboard, they find that it is often impossible to correctly interpret the tonal or rhythmic structure of the theme till several measures have passed. Steedman (1977) develops some ingenious ideas as to how a listener might attempt to work out rhythmic structure in an ambiguous passage, which involve, among other things, looking out for repetitions of tonal sequences. There is a sense, though, in which this is placing rather too much responsibility on the listener. Very few performers will, in fact, give an entirely dead-pan performance of a theme. Even on an organ keyboard it is possible to indicate accent by variations in durations of notes. On most other instruments variations in amplitude, pitch, and timbre for individual notes are also possible. Thus, in all cases except those where one has good reason to believe the composer intended ambiguity it is part of the performer's job to present music to the listener in a way that creates no ambiguity, by whatever means are available to him.

Consider, for instance, the following sequence of notes:

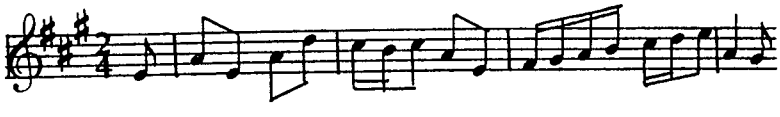


**Example 1.2**

When this is played in a 'dead-pan' fashion to listeners, some perceive the beat as coming on the first, third, fifth note, and so on; whilst others perceive the beat as coming on the second, fourth, sixth note, and so on. Both are valid interpretations which make reasonable musical sense. Of course, in a literate score, the two alternatives would be notated differently, as follows:

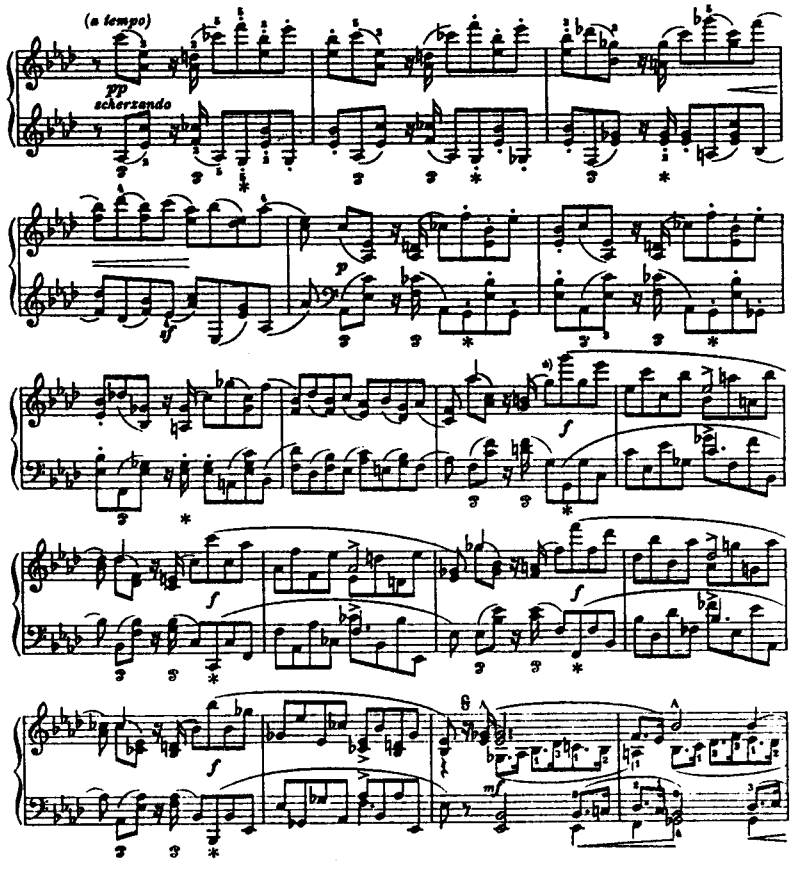


**Example 1.3**



**Example 1.4**

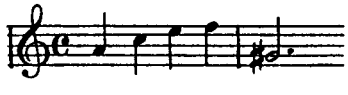
A good performer, confronted with one of these alternatives, would naturally play it in such a way as to make sure that listeners did not 'mis-hear' it as the alternative. In a reading situation, this would imply assigning to the passage, before playing it, one or other of the two alternative interpretations. This would then allow decisions about the duration, loudness, and quality of the various notes to be made. In this case, the cues which allow a reader to arrive at the correct interpretation are fairly straightforward. They include the placing of the bar-lines and the beaming of notes into pairs. There are, however, many examples where the assignation of a correct interpretation is less straightforward, and therefore demands greater literacy on the part of the performer. Take, for example, the second movement of the Schumann's *Phantasie* for piano, bars 141–157.




**Example 1.5** Bars 141–157 from second movement of *Phantasie* for piano by Schumann.

Here, the bar-lines and the beaming suggest a conventional 4/4 beat. If, however, a performer attempted to place the beats on the first, third, fifth and seventh quavers in each bar, musical nonsense would result. The 'beats' are in fact on the second, fourth, sixth, and eighth quavers. The phrase marks suggest this to some degree, but it takes considerable reading skill to see that the disposition of notes themselves demands this interpretation. In essence it entails the knowledge that in the upper register of the piano, all other things being equal, the uppermost notes will constitute the psychologically dominant material, especially if separated from lower notes by a significant interval (cf. Dowling, 1973; Deutsch, 1977). Why then, one may ask, did Schumann put the bar lines where he did? Presumably because he wanted to instil into the music a slight metrical 'unease' as though there were a suppressed 4/4 metre going on underneath all the time. This theory is substantiated by the fact that five times during the passage, the underlying metre 'breaks through' in the form of an accented 5th quaver. To capture all this in a performance demands a highly intelligent reading of the score. Few of us, maybe, would be capable of picking all this up at a first playing, even if we were technically proficient enough to play the passage at all, but Schumann is within his rights to demand that we should be able to understand his intentions from the score.

The claim that readers should be able to interpret a score before playing it should not be confused with the claim that they should be able to hear what the music should sound like 'in the mind's ear'. This second claim may be true of particular individuals, but one would wish to call it literacy only if mediated by an appreciation of the musical meaning. Let us take a linguistic analogy. Any competent reader of English should be able to imagine what the following would sound like if spoken: *Homo factus est.* However, only if the reader understood Latin would we wish to say that he had displayed full literacy in the task. Similarly, we could imagine a musician raised on a diet of atonal music who might somehow be able to imagine the pitches associated with the following sequence:



**Example 1.6**

This would, I believe, not be taken as evidence of musical literacy unless it was coupled with a knowledge that the sequence implied a movement from tonic to dominant harmony in the key of A minor. Furthermore

I believe that we should wish to ascribe literacy to a music reader who displayed evidence of this knowledge even if he was not imagining the actual sounds. There is mounting evidence now that language reading does not always involve a sight to sound translation, but can take place by a direct transition from sight to meaning (Allport, 1979; Bradshaw, 1975; Patterson and Marcel, 1977; Saifran and Mann, 1977). I wish later to review some evidence which suggests that competent music readers may use a similar process.

Finally, nothing which has been said above about the skills which might be employed by performers implies that performers should have full conscious awareness of these processes. On the contrary, they may be largely unaware of what is going on when they read. The notion that highly practised perceptual skills are executed without the involvement of consciousness receives general support in the psychological literature (e.g. LaBergc and Samuels, 1974; Schneider and Shiffrin, 1977).

## 1.2 Introspections and conversations

It is a common experience among keyboard players in particular that page turners do not turn early enough. This author has, on several occasions, agreed in advance with a page turner that a head nod would signal the turn, only to find in performance that the nod was ignored. 'I couldn't believe that you wanted me to turn so soon!' is a typical post-mortem remark. This experience alone is enough to tell the seasoned performer that he reads further ahead than does the average page-turner. A similar thing occurs in choirs where a good and a less good reader are sharing the same copy. if the poor reader is controlling the copy he will invariably turn later than is comfortable for the good reader. Nonetheless, when one asks musicians to estimate how far ahead they customarily read, they usually have only the vaguest of ideas.

Nearly all instrumental teachers have to spend some time with their pupils on sight-reading, if only to prepare them for examinations. If my experience is typical, this vagueness about one's own sight-reading extends to all facets of it and makes the teaching of sight-reading an almost impossible assignment. I literally do not know what to say to a pupil who cannot sight-read. I do not remember having to *do* anything in particular myself to achieve my adequacy in reading. On the contrary I have often used my reading facility to cover up for far more intractable technical difficulties, or for lack of practice. I suspect that most good readers acquired their skill early in life without too much trouble, with the unfortunate consequence that those who sight-read best may well be the worst

teachers of it. Teachers who in adult life have had to struggle to achieve a basic reading competence are probably in a far better position to impart useful hints to pupils. Nonetheless, it is of interest to discover what expert sight-readers have to say for themselves, and for this reason Wolf (1976) has performed a valuable service in interviewing four exceptional sight-readers of his acquaintance. Although only short extracts from the interviews are printed, some persistent themes emerge. One of these is the insistence that notes are seen in groups rather than singly. Nonetheless, some of the extracts suggest that introspection does not come easily. For instance, the following remark of Boris Goldovsky suggests that he is having to convince himself that a plausible statement actually finds corroboration in his own experience: 'When you read piano music, you see them (he staves) both simultaneously. You don't read them separately. The eye certainly covers both lines easily . . . I don't think I read the top line, then do the bass line. No, I'm quite sure that I see at least that much at once.' When Vladimir Sokoloff was asked what the principal problems involved in sight-reading were, his answer was: 'For me personally, there are none.' Although this is hardly useful for someone attempting to understand the sight-reading process, it is, I believe, an honest answer about a skill which, for him, is so highly overlearned that its execution requires almost no conscious effort.

Although, therefore, the comments of skilled sight-readers are of considerable interest as sources of hypotheses about the sight-reading process, it would be misleading to consider them descriptions of the processes involved. They are perhaps more like commentaries on their conscious experiences, or rationalisations of their own behaviour. Whilst such comments must be accounted for in a full theory of skilled music reading, they do not in themselves offer the means of achieving such a theory. This paper is motivated by the belief that only by direct observation of musicians engaged in reading tasks in the context of controlled experimental situations can one hope to achieve any substantial understanding of this complex skill. This belief has motivated much of the recent research on language reading (Williams, 1970), and the techniques and models developed there over the last 20 years have considerable relevance to the study of music reading.

## 1.3 Misreading

The phenomenon known as 'proof-readers' error' has been long familiar to cognitive psychologists (Vernon, 1931; Pillsbury, 1897; Huey, 1908). This occurs because of the tendency for incorrectly spelled words to be

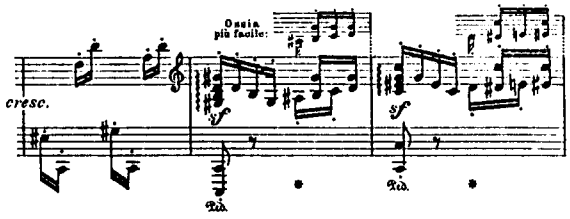
overlooked when the mis-spelling is slight. This is one of many observations which supports the view that reading involves 'top-down' or 'conceptually driven' processes, as well as 'bottom-up' or 'data-driven' processes (Runielhart and Siple, 1974; Smith, 1971; Kolers, 1970; Norman, 1976). In such processes a reader does not depend simply upon decoding of stimulus information to build up a mental representation of the text; he uses his prior knowledge and expectancy to supplement, or even replace stimulus information. Neisser (1967) enjoins us to characterise reading as 'externally guided thinking'. By looking at the kind of errors people make when reading, we can find out exactly what expectancies they have, and exactly what parts of the stimulus information they are using. For instance, it has been shown that readers are more likely to misperceive letters occurring in the middles of words than at the beginning or the end (Morton, 1964; Sloboda, 1976b). Thus, readers detect the mis-spelling in 'hwistle' much more readily than that in 'whitsle', suggesting that they identify words by identifying the exterior letters and then using the surrounding context and their knowledge of the language to pick the most appropriate word. Top-down processing is not confined to reading. It can be found in most types of perception: listening to speech (Marslen-Wilson and Welsh, 1978; Warren, 1970; Cole, 1973), looking at pictures (Palmer, 1975), perceiving chess boards (Chase and Simon, 1973).

The existence of 'proof-readers' error' in music has been demonstrated by the 'Goldovsky experiment' reported in Wolf (1976):

> a student whom Dr. Goldovsky describes as 'technically competent but a poor reader' prepared a Brahms *Capriccio* (Opus 76 no. 2) which she brought to her lesson. She began to play the piece through but when she arrived at the C sharp major chord on the first beat of the bar 42 measures from the end, she played a G natural instead of the G sharp which would normally occur in the C sharp major triad. Goldovsky told her to stop and correct her mistake. The student looked confused and said that she had played what was written. To Goldovsky's surprise, the girl had played the printed notes correctly—there was an apparent misprint in the music.

This misprint occurs in most standard editions of Brahms' piano music. It appears, therefore, that countless scores of musicians, let alone proof-readers and publishers, had never noticed this misprint. Goldovsky proceeded to test skilled readers by telling them that the piece contained a misprint which they were to find. He allowed them to play the piece as

many times as they liked and in any way they liked. None of them ever found the mistake. Only when he narrowed the field down to a single bar did most readers spot the mistake.




**Example 1.7** Bars 76–78 from *Capriccio*, op. 76, no. 2 by Brahms.

One may guess why this misprint was so difficult to spot. Firstly the chord occurred at the beginning of a bar. Skilled readers will be used to seeing C sharp major chords in other positions of a bar when the G sharp will have been signalled by an earlier note. Such chords may be visually identical to the 'misprint' in this case. Secondly the chord contains other accidentals. A quick glance at the chord may give the impression that one of them is actually attached to the G. Thirdly, and most importantly, we may suppose that most readers were expecting to find a C sharp major chord at that point. This expectancy, coupled with the fact that the printed chord was sufficiently similar to what was expected, led to the error. It is highly significant that the first person to notice the error was a poor reader. Perhaps one of the causes of poor reading is insufficient expectancy and the lack of top-down processing. Sloboda (1976b) studied misreading more systematically by introducing deliberate notational 'mistakes' into pieces of music which subjects were asked to read. Four excerpts from relatively little-known keyboard pieces by Marcello and Dussek were used. They were chosen for the unlikeliness of their having been encountered previously by the subjects used, and for their adherence to highly conventional harmonic and rhythmic devices. These excepts were copied out by hand, and had introduced into them a total of 72 copying errors distributed equally between the two staves and between the various positions within a musical phrase. Each error was produced by raising or lowering a note by one position on the stave, but transferring intact all accidentals, etc. The errors were chosen to be both playable and highly unlikely in the musical context as expressive of the composers' intentions.

These pieces were presented to a group of competent piano readers who were asked to sight-read each piece twice. Their instructions were to play exactly what was written with a view to seeing how much they could

improve on the second play-through. Most subjects suspected that there might be something odd about the passages, and, having spotted one misprint, were consciously on the lookout for others. In spite of this, no subject managed to execute all the altered notes correctly. Instead, many of them were played as they appeared in the original score. Most subjects showed genuine surprise when some of the altered notes were pointed out to them. They had not noticed them, playing what ought to have been there rather than what was there. Not all altered notes were equally easy to spot. Hardest to spot were those in the middle of right-hand phrases. This corresponds rather well with the pattern shown when reading words.

Another interesting finding was that, although subjects made fewer errors on their second attempts at the pieces, they made more errors on the altered notes, playing even more of them as they appeared in the original correct score. This suggests that, as subjects began to become familiar with the structure of the pieces they used this knowledge to 'skim' even more. In general, their knowledge allowed them to predict correctly what notes should be played, but in the case of the altered notes this knowledge would predict the wrong course of action. Thus, it seems that good sight-reading is based, at least in part, on the ability to decide on probable continuations within an idiom. For conventional tonal music this implies a knowledge (explicit or implicit) of harmonic and rhythmic rules, and an ability to translate this knowledge into the movements necessary to produce the appropriate sounds. This does not necessarily mean that a good sight-reader must be technically expert. Many proficient instrumental readers will often sing a passage that they just cannot get their fingers round, or provide a harmonically acceptable substitute. In these cases it is clear that they know what they should be playing. Another hallmark of the good reader is the ability to correct mistakes rapidly and efficiently. For instance, it is easy, on the piano, to find oneself playing a key too high or too low. The good reader will often correct himself after playing only a single wrong note, whilst the poor reader will often continue at the wrong pitch for several notes. This does not so much suggest that good readers have 'good ears' (because most people can detect wrong notes when listening to someone else playing) as that they have the ability to monitor their own performances. The lack of this skill in beginners is widespread, and affects technical and interpretative skill as well. This lack may be due to the fact that the mental effort required to process the visual input and produce a response is so great that there is little attentional capacity left for monitoring.

It would be interesting to know at what level good readers monitor their performance. One possibility is that they listen to what they produce.

Another is that they monitor kinaesthetic feedback, or even the 'motor commands' themselves, so that they know even before hearing a note that they have played it wrongly. In studies of other forms of skilled performance there is compelling evidence that performers cannot be simply monitoring their output, since there is often not the delay in making correction movements which would follow if this were the case. In other words, they make corrections too quickly. Discussions of this issue are given by Broadbent (1977), Poulton (1957), Young (1969).

## 1.4 Reading span

One good way of enhancing one's reading performance is to obtain preview (Shaffer, 1975; 1976). If one is taking in visual information well ahead of the time when performance is required, one has time to decode the information and provide a steady flow of output in spite of temporary hold-ups along the information processing route. It follows that preview requires some form of memory storage in which notes that have been seen, but not yet played, can be held on a short-term basis. The greater the capacity of this store, the greater is the opportunity for preview, and the greater is the opportunity for making reasonable predictions about subsequent notes. Given six notes to go on, success in predicting the next note is more likely than if there are only two notes. A standard method of estimating the amount of preview in normal reading, and thus by inference the capacity of short-term memory store, involves measuring the 'eye-voice span'. This technique, and a modification of it suitable for application to music reading, has already been described in this journal (Sloboda, 1974). It was shown that better readers had larger spans for single line melodies. The best reader had a span of 6.5 notes, whereas the poorest reader had a span of only 3.5 notes. So the common advice to 'look ahead more' seems to be enacted by proficient readers. But looking ahead is no use to the reader unless he is able to remember what he has seen, so we must ask how it is that the good reader can memorise more of a musical score than a poor reader.

One clue is supplied by the observation that span is not constant but appears to 'jump' from one phrase boundary to the next, suggesting that readers are treating the phrase as some kind of 'unit'. This appears to be not simply a matter of picking out a visual boundary (like using gaps between words in prose reading (Hochberg, 1970), but of using the structure of the music itself to relate notes within a phrase to one another. This was shown by an experiment (Sloboda, 1977) in which musicians were asked to read musical passages, of which some conformed to the conventional rules of melodic progression, and others did not. It was found that span

decreased for the unconventional melodies, as did the tendency for span to coincide with a phrase boundary (defined rhythmically). Thus we may conclude that one of the factors allowing good readers to hold larger 'chunks' of musical material in memory is their ability to organise the material into groups of notes having higher-order interrelationships. The effects of 'redundancy' on memory for linguistic material has been amply demonstrated in the psychological literature (Miller and Selfridge, 1950; Epstein, 1961; Marks and Miller, 1964). Thus, people read and remember less well items of the following sort—'of by method gravity is beer simplest the serving'—than items which conform to grammatical expectations, e.g. 'When eating spanners it is best to remove both feet'.

One may suppose that a lot of music is constructed according to rules which are similar to grammatical rules in that writing certain combinations of notes limits the range of notes which can follow them. For instance the sequence C–E–G–F–E–D suggests C, E or G as a probable continuation. B, on the other hand would be rather unlikely. Some writers (e.g. Pinkerton, 1956) have suggested that musicians learn a set of transitional probabilities between adjacent notes which makes certain intervals much more likely to be used than others. While this may be true, it is unlikely that it is the whole story. It is more likely that musicians learn that certain harmonic moves are more permissible than others within the tonal system (cf. Winograd, 1968; Sundberg and Lindblom, 1976; Longuet-Higgins, 1977), so that a melodic continuation is rendered acceptable if the implied harmonic progression allows it, regardless of the interval that it creates with the preceding note. Of course, musicologists devote much of their attention to the construction of musical grammars of considerable refinement (e.g. Forte, 1962; Cooper and Meyer, 1960). The question that concerns us here is how much of these grammars do musicians use (albeit automatically or unconsciously) to assist their reading? The word 'grammars' is deliberately put in the plural since it is clear that the rules of musical construction vary, if not from composer to composer, from period to period. It is likely that as we learn more about a particular composer, reading his music will become easier. Proficiency at reading Mozart only predicts proficiency in reading Messiaen insofar as their languages overlap.

While no experimental verification of this hypothesis exists, it is a common experience (among musicians who have to learn several works of a new composer writing in a consistent idiom) that it becomes easier to read the music as time passes. Chords which at first looked like meaningless jumbles of notes now start to 'lie under the fingers'. Rhythms which had to be pieced together note by note, now can be played 'as a pattern'. It is as if the reader is gradually learning new musical 'words' and entering

them into the musical equivalent of a 'lexicon'. The reader seems to be acquiring new 'distinctive features' in a musical pattern recognition task (cf. Gibson, 1969; Uhr and Vossler, 1963).

## 1.5 The nature of musical memory

If reading ability depends upon rapid memorisation of notes in advance of their performance it becomes of central importance to determine the form in which the material is held. Is the representation acoustic (a set of imagined sounds), articulatory (a set of specifications for muscular movement), or what? A similar question has held centre stage in the study of language perception for several years (reviewed by Neisser (1967), and more recently by Bradshaw (1975), Henderson (1977), Allport (1979), Gibson and Levin (1975). One of the ways in which experimentalists have tried to solve this problem is to attempt to occupy the hypothesised storage-system with an extraneous task and observe whether there is a decrement in performance on the primary task (Allport, Antonis and Reynolds, 1972; Baddeley and Hitch, 1974). If there is no decrement, then it is often plausible to argue that the primary task cannot be using that particular memory system.

Using a technique of this sort (Sloboda, 1976a) it was found that subjects could memorise lists of letters, or sequences of tones, at the same time as performing a task in which they were required to write down a briefly exposed (2 sec.) segment of musical notation, without any significant detriment to either task. Nonetheless, musically trained subjects were significantly better at the notation transcription task than a control group of untrained subjects. This superiority could not be due to greater writing facility since the two groups of subjects did not differ in a condition where exposure was very brief (20 msec.). Thus it appeared that the musicians' superiority arose because of more efficient storage of the material in some non-verbal, non-acoustic type of memory. It seemed as though the visual input was obtaining direct access to some abstract form of representation which was not aurally based. The precise nature of this representation is yet unknown, although visual contour seems to be an important aspect of the input for musicians (Sloboda, 1978). In a sense, the independence of the reading system from vision-to-sound translation is not surprising. Proficient readers are always looking at one thing while listening to something else, simply because they read ahead. If they are occupying acoustic mechanisms by listening to what they are playing, visual input could not be passed through the same system, and another route would have to be

developed. Of course, one cannot say that the sound system is *never* used in reading. In fact it probably is used when the reader is learning an unfamiliar idiom where no efficient procedures exist for determining the structure or 'meaning' of the music. It is probably the only way to sight-sing contemporary music efficiently. No tonal centre being present, each note has to be pitched as an interval from the preceding note, the pitch of the sound being matched to an internally generated image. Thus, the sound route must he available. The contention here is that it is not generally used by good sight-readers in conventional (tonally based) reading situations.

At the present time, it seems most plausible to suppose that the musical memory code is a representation of the visual input in abstract tonal and rhythmic space, consisting of a set of tonal centres, or harmonies, together with a more or less detailed specification of the structural-temporal relations between them. This representation may then be overlaid by visual specifications for frequently occurring sequences, such as scale passages. Thus the sequence CDEFGFEDC might be represented as 'C major (tonic+rising scale→dominant+falling scale→tonic)' in an hypothetical memory notation. Some such system would seem necessary to account for the mistakes made during attempts to memorise tunes (cf. Davies and Jennings, 1977). Such errors are generally tonally and rhythmically related to the stimulus, and are rarely haphazard. That such a memory system is directly involved in reading awaits experimental demonstration. A possible approach would be to look for interference between reading and a concurrent task which required tonal analysis.

## 1.6 Helping learners

The foregoing considerations lead to a number of suggestions for acquiring sight-reading proficiency, offered not because they have been tried and found to work, but rather in the hope that teachers might be encouraged to re-examine their teaching methods.

Firstly, the reader must build up the musical knowledge of form, style and 'language' to be able to make small-scale predictions about what is going to come next. This might be encouraged by making copies of scores in which, from time to time, one or two notes or chords have been left out. Students could attempt to play these scores, filling in with appropriate notes. Training in improvisation, providing continuations to given themes, keyboard harmony, and memory for melodies, should all help this ability.

Secondly, the reader must wean himself away from a direct association between a written note and a hand movement on the instrument. A good

reader should appreciate the music 'in his head' without playing it at all. Sight-singing should help this skill, checking on an instrument afterwards to see if one is right. Reading 'a bar at a time' will also help: that is, looking at a bar, putting the music away, and attempting to play the whole of that bar. The size of chunks can gradually be built up as proficiency is acquired. Following the score of a work whilst listening to it on a record or on the radio should also help, attempting to keep a note or two ahead of the sound point and anticipating what the next notes will sound like. This exercise would be best begun with something simple (slow movement of a sonata or quartet) before symphonic works are tackled. Reading the score of a familiar work imaging the sounds of the music, but without actually playing them may help, as will learning a keyboard instrument—keyboard music makes explicit the harmony which is only implicit in a solo line, thus providing many more cues for the beginner to guess what is coming next.

Thirdly, sight-reading performed note-by-note is unlikely to improve however frequently practised. The music must be understood before it is played, using the actual sound to check out one's hypothesis. The main aid to this is a fascination with music itself, a desire to find out what a piece sounds like, It is therefore important to sight-read music that is liked. For many children this may mean a diet of pop-music transcriptions. In fact, these often make excellent sight-reading material, since they use basic harmonic patterns a great deal, and familiar rhythmic and melodic patterns. In the early stages, the reader should be encouraged to try to associate the sound with its written counterpart. Occasional stopping may facilitate this in the beginner by providing opportunity for thought about what has been played, or judgement of correctness.

Fourthly, it would seem reasonable to seek to attempt to develop musical sensibility before embarking on training in reading. No one would consider teaching a normal child to read while he was at a very early stage of learning spoken language. Yet it seems the norm to start children off on reading at the very first instrumental lesson without establishing the level of musical awareness already present. Without some musical knowledge a beginner has no expectancies which can be used in reading.

Fifthly, the teacher might advantageously create situations in which the learner needs to be able to read in order to fulfil musical or social aspirations. Membership of choir or a chamber music group will provide motivation for development of reading. Placing expectations of pupil's reading performance at a deliberately high level is also likely to enhance development. If we seek to improve standards, we must not let the poor reader hold us to ransom in an ensemble situation.

## 1.7 Improving musical text

In all that has been said so far we have assumed a perfect, immutable score which it is the reader's job to understand and execute. Many of the difficulties involved in reading, however, may be attributable to faults in the score itself, and not the reader.

The amount of psychological research on the presentation of printed language is considerable (Hartley and Burnhill, 1976; Poulton, 1959; Hartley, 1978; Watts and Nisbet, 1974), but it is doubtful how much of this is relevant to music, where both the reading environment, and the nature of the symbols used, are very different. There has been little research on factors of presentation in musical text which affect the ease and speed of reading, although practically every musician has come across examples of presentation practices which have seriously upset his reading performance. One can name just a few of these: crowding of ledger lines; inconsistencies in horizontal separation of notes; staves too close together, so that it is difficult to tell which stave a dynamic mark applies to; decrescendo signs which look like accents; notes which are not accurately centred on a line or a space; page turns in awkward places; notes, stems, stave lines which are too thick or thin for comfortable reading; unhelpful word underlay, etc.

Some of these difficulties are caused by lack of thought in preparation of the score, and can be prevented. Other problems are less easy to solve. How big should a stave be? Are clef changes preferable to ledger lines? How wide should a page of music be, and how many notes to the inch should a line contain on average? What is the best shape for note-heads? How thick should stems be, and how high should they be? Many of these questions seem to have been decided on the basis of engraving tradition, and, where explicit stipulations are made (e.g. Weaner and Boelke, 1966), there is little evidence that alternatives have been tried out, or even seriously considered. Nonetheless, if one takes several editions of practically any classical work, the differences in presentation are usually immense. Some of the differences may have only aesthetic import, but there is reason to suppose that not all differences are of this sort.

In the absence of substantial research into these matters, is there any general principle which could guide those concerned with the preparation of music which other people must play? One guiding principle must be to adopt any practice which makes it easier for the reader to perceive the structure of the music. To take an example, until recently it was the custom to join quavers and semiquavers in vocal music by beams only if the notes were sung to the same syllable. When a bar contained many short notes it was often difficult to see which notes came on a strong beat,

and it was easy to overlook a single quaver amid many semiquavers. The modern practice of beaming the notes as in the manner of instrumental music clarifies the rhythmic structure of a vocal passage at no expense to other considerations (provided that the word underlay is set out intelligently). Another example concerns the use of accidentals. There are many examples of tonal music where, for several bars, a new key is established so remote from the original key that almost every note must be accompanied by an accidental of some sort. It would be perceptually and cognitively easier to change the key signature for the duration, since this would unambiguously signal the key change, so making the harmonic structure more apparent. The memory capacity of the average reader must also be taken into account. In a 12/8 piano bar containing 12 quaver chords with various accidentals, one should not expect the performer to be able to remember all through the bar that there was, say, an E sharp in the very first chord. If the last chord also contains an E sharp the accidental should be written again, even if the rules of notation do not strictly require it.

Many considerations of like kind are specific to a particular context, and demand knowledge of the particular work in question. For the serious student of music-presentation, there is no substitute for the detailed study of alternative editions of a work.

Those who produce instrumental parts which are to be played by professional orchestral musicians have, perhaps, the greatest need to consider psychological aspects of music reading. The amount of sight-reading required of professional instrumentalists (at least in Britain) is phenomenal. Slight blemishes in a part can waste valuable rehearsal seconds, and can contribute to faults in hastily rehearsed performances. A slight imperfection in a piano score which a solo performer has many months to study might not be important. In a part which may be seen by the performer only once, if at all, before the performance, such matters are of essential importance. It is, therefore, surprising to discover that the quality of many orchestral parts is very poor. Some sets of parts used by England's major symphony orchestras have to be seen to be believed.

Finally, a radical alternative to improving legibility of conventional notation is the devising of new notational systems (see Karkoschka, 1972). Some of these are necessary because the existing notation is incapable of specifying a composer's requirements, but others (e.g. Klavarscribo) have been proposed as improvements on conventional notation. These systems rarely offer advantages not available in conventional notation. They merely employ alternative symbols, which, their various proponents have claimed, assist reading in one way or another. None of these systems has gained wide adherence, but it would be presumptuous to suggest that they are all

inferior to conventional notation. However, only when a set of objective criteria for the psychological efficacy of musical symbols is developed shall we have a way of judging between the numerous alternative notational systems which contemporary music is generating.

In conclusion, I hope to have shown that the psychological study of music reading is an important, if hitherto neglected, area of the psychology of music which nevertheless has bearing on central theoretical and practical issues in music. It offers one important route to a better understanding of the nature of musical cognition itself.

## References

Allport, D. A. (1977) On knowing the meaning of words we are unable to report: the effects of visual masking. In S. Dornic (Ed.) *Attention and Performance VI.* Hillside, N.J.: Erlbaum Associates.

Allport, D. A. (1979) Word recognition in reading: a tutorial review. In H. Bouma, P. A. Kolers and M. Wrolstad (Eds.) *Processing of Visible Language* (proceedings of a conference held at the Institute for Perception Research, I.P.O., Eindhoven, Netherlands, September 1977).

Allport, D. A., Antonis, B., and Reynolds, P. (1972) On the division of attention: a disproof of the single channel hypothesis. *Quarterly Journal of Experimental Psychology*, **24**, 225–235.

Baddeley, A. D. and Hitch, G. (1974) Working memory. In Go Bower (Ed.) *Recent advances in learning and motivation VIII.* New York: Academic Press.

Bean, K. L. (1938) An experimental approach to the reading of music. *Psychological Monographs*, **50**, whole number 226.

Bradshaw, J. L. (1975) Three interrelated problems in reading: a review. *Memory and Cognition*, **3.2**, 123–134.

Broadbent, D. E. (1977) Levels, hierarchies, and the locus of control. *Quarterly Journal of Experimental Psychology*, **29**, 181–201.

Buck, P. C. (1944) *Psychology for Musicians.* London: Oxford University Press.

Chase, W. G. and Simon, H. A. (1973) The mind's eye in chess. In W. G. Chase (Ed.) *Visual Information Processing.* New York: Academic Press.

Cole, H. (1974) *Sounds and Signs: Aspects of Musical Notation.* London: Oxford University Press.

Cole, R. A. (1973) Listening for mispronunciations: a measure of what we hear during speech. *Perception and Psychophysics*, **11**, 153–156.

Cooper, G. W. and Meyer, L. B. (1960) *The Rhythmic Structure of Music.* Chicago: University of Chicago Press.

Critchley, M. and Henson, R. A. (1977) *Music and the Brain.* London: Heinemann.

Davies, J. B. (1978) *The Psychology of Music.* London: Hutchinson.

Davies, J. B. and Jennings, J. (1977) The reproduction of familiar melodies and the perception of tonal sequences. *Journal of Acoustical Society of America*, **61.2**, 534–541.

Deutsch, D. (1977) Memory and attention in music. In M. Critchley and R. A. Henson (Eds.) *Music and the Brain*. London: Heinemann.

Dowling, W. J. (1973) The perception of interleaved melodies. *Cognitive Psychology*, **5**, 322–337.

Epstein, W. (1961) The influence of syntactical structure on learning. *American Journal of Psychology*, **74**, 80–85.

Forte, A. (1962) *Tonal Harmony in Concept and Practice*. New York: Holt, Rinehart and Winston.

Gibson, E. J. (1969) *Principles of Perceptual Learning and Development*. New York: Appleton-Century-Crofts.

Gibson, E. J. and Levin, H. (1975) *The Psychology of Reading*. Cambridge, Massachusetts: MIT Press.

Hammer, H. (1963) An experimental study of the use of the teaching of melodic sight singing. *Journal of Research in Music Education*, **11**, 44–54.

Hartley, J. (1978) *Designing Instructional Text*. London: Kogan Page.

Hartley, J. and Burnhill, P. (1976) *Textbook Design: A Practical Guide*. UNESCO.

Henderson, L. (1977) Word recognition. In Sutherland, N. S. (Ed.) *Tutorial Essays in Experimental Psychology*, 1. Potomac, Md.: Erlbaum.

Hillbrand, E. K. (1923) *Hillbrand Sight-singing Test*. Yonkers, New York: World Book.

Hochberg, J. (1970) Components of literacy: speculations and exploratory research. In H. Levin and J. P. Williams (Eds.) *Basic Studies on Reading*. New York: Basic Books.

Huey, E. B. (1908) *The Psychology and Pedagogy of Reading*. Cambridge, Massachusetts: MIT Press, reprinted 1968.

Karkoschka, E. (1972) *Notation in New Music*. London: Universal Edition.

Kolers, P. A. (1970) Three stages of reading. In H. Levin and J. P. Williams (Eds.) *Basic Studies on Reading*. New York: Basic Books.

Kwalwasser, J. (1955) *Exploring the Musical Mind*. New York: Coleman-Ross Co. Inc.

LaBerge, D. and Samuel, S. J. (1974) Toward a theory of automatic information processing in reading. *Cognitive Psychology*, **6**, 293–323.

Lannert, V. and Ullman, M. (1945) Factors in the reading of piano music. *American Journal of Psychology*, **58**, 91–99.

Longuet-Higgins, H. C. (1972) Making sense of music. *Proceedings of the Royal Institute of Great Britain*, **45**, 87–105.

Longuet-Higgins, H. C. (1976) The perception of melodies. *Nature*, **263**, 646–655.

Longuet-Higgins, H. C. (1978) The perception of music. *Interdisciplinary Science Reviews*, **3**, 148–156.

Marks, L. E. and Miller, G. A. (1964) The role of semantic and syntactic constraints in the memorization of English sentences. *Journal of Verbal Learning and Verbal Behaviour*, **3**, 1–5.

Marslen-Wilson, W. D. and Welsh, A. (1978) Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, **10**, 29–63.

Miller, G. A. and Selfridge, J. A. (1950) Verbal context and the recall of meaningful material. *American Journal of Psychology*, **63**, 176–185.

Morton, J. (1964) The effects of context on the visual duration threshold for words. *British Journal of Psychology*, **55.2**, 165–180.

Neisser, U. (1967) *Cognitive Psychology*. New York: Appleton-Century-Crofts.

Norman, D. A. (1976) *Memory and Attention* (2nd edition). New York: Wiley.

Palmer, S. E. (1975) Visual perception and world knowledge. In D. A. Norman, D.E. Rumelhart and the L.N.R. Research Group, *Explorations in Cognition*. San Francisco: Freeman.

Patterson, K. E. and Marcel, A. J. (1977) Aphasia, dyslexia and the phonological coding of written words. *Quarterly Journal of Experimental Psychology*, **29**, 307–318.

Pillsbury, W. B. (1897) A study in apperception. *American Journal of Psychology*, **8**, 315–393.

Pinkerton, R. C. (1956) Information theory and melody. *Scientific American*, **194.2**, 77–86.

Poulton, E. C. (1957) On prediction in skilled movements. *Psychological Bulletin*, **54**, 467–478.

Poulton, E. C. (1959) Effects of printing types and formats on the comprehension of scientific journals. M.R.C. Applied Psychology Unit Report No. 346. Cambridge University, U.K.

Rumelhart, D. F. and Siple, P. (1974) The process of recognising tachistoscopically presented words. *Psychological Review*, **81**, 99–118.

Saifran, E. M. and Marin, O. S. M. (1977) Reading without phonology: evidence from asphasia. *Quarterly Journal of Experimental Psychology*, **29**, 515–525.

Schneider, W. and Shiffnin, R. M. (1977) Controlled and automatic human information processing. *Psychological Review*, **84.1**, 1–66.

Seashore, C. (1938) *The Psychology of Music*. New York: McGraw-Hill; reprinted by Dover Books: New York (1967).

Shaffer, L. H. (1976) Intention and performance. *Psychological Review*, **83.5**, 375–393.

Shaffer, L. H. (1975) Control processes in typing. *Quarterly Journal of Experimental Psychology*, **27**, 419–432.

Shuter, R. (1968) *The Psychology of Musical Ability*. London: Methuen.

Sloboda, J. A. (1974) The eye-hand span: an approach to the study of sight reading. *Psychology of Music*, **2**, 4–10.

Sloboda, J. A. (1976a) Visual perception of musical notation: registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology*, **28**, 1–16.

Sloboda, J. A. (1976b) The effect of item position on the likelihood of identification by inference in prose and music reading. *Canadian Journal of Psychology*, **30.4**, 228–238.

Sloboda, J. A. (1977) Phrase units as determinants of visual processing in music reading. *British Journal of Psychology*, **68**, 117–124.

Sloboda, J. A. (1978) Perception of contour in music reading. *Perception*, **7**, 323–331.

Smith, F. (1971) *Understanding Reading: A Psycholinguistic Analysis of Reading and Learning To Read*. New York: Holt, Rinehart and Winston.

Steedman, M. J. (1977) The perception of musical rhythm and metre. *Perception*, **6**, 555–570.

Sundberg, J. and Lindblom, B. (1976) Generative theories in language and music descriptions, *Cognition*, **4**, 99–122.

Uhr, L. and Vossler, C. (1963) A pattern recognition program that generates, evaluates, and adjusts its own operators. In E. A. Feigenbaum and J. Feldman (Eds.) *Computers and Thought*. New York: McGraw Hill.

Van Neuys, K. and Weaver, H. E. (1943) Memory span and visual pauses in reading rhythms and melodies. *Psychological Monographs*, **55**, 33–50.

Vernon, M. D. (1931) Characteristics of proof-reading. *British Journal of Psychology*, **21**, 368.

Warren, R. M. (1970) Perceptual restoration of missing speech sounds. Science, **167**, 392–393.

Watts, L. and Nisbet, J. (1974) *Legibility in Children's Books*. Windsor, Berks.: NFER Publishing Company.

Weaner, M. and Boelke, W. (1966) *Standard Music Engraving Practice*. New York: Music Publishers' Association of the U.S., Inc.

Weaver, H. E. (1943) A survey of visual processes in reading differently constructed musical selections. *Psychological Monographs*, **55**, 1–29.

Weber, R. M. (1970) First-graders' use of grammatical context in reading. In H. Levin and J. P. Williams (Eds.) *Basic Studies on Reading*. New York: Basic Books.

Williams, J. P. (1970) From basic research on reading to educational practice. In H. Levin and J. P. Williams (Eds.) *Basic Studies on Reading*. New York: Basic Books.

Winograd, T. (1968) Linguistics and the computer analysis of tonal harmony. *Journal of Music Theory*, **112**, 3–49.

Wolf, T. (1976) A cognitive model of musical sight reading. *Journal of Psycholinguistic Research*, **5**, 143–171.

Young, L. R. (1969) On adaptive manual control. *Ergonomics*, **12**, 635–674.

CHAPTER 2

# EXPERIMENTAL STUDIES OF MUSIC READING: A REVIEW

The psychological study of music reading is atypical of music perception studies in two ways. First, we can identify a clearly defined behavioral goal in most nonlaboratory music reading situations. This is not the case for many music listening behaviors. Second, the object of perception is visual, not auditory. The first of these features provides the investigator with some advantages; the second poses some problems. Both merit further comment.

## 2.1 Goals of music reading

The most common goal of music reading is the production of a coherent musical performance. The reader converts the visual input into a set of prescriptions for performance—he finds out which notes to play, in which sequence and combination they are to occur, and much else. His performance provides the investigator with a continuous measure of the effectiveness of his reading procedures, and it is possible to specify some of the conditions that must be satisfied if a person can be said to have read the music effectively. We will examine some of these conditions in more detail later.

In contrast, it is not easy to define well-formed goals in most music listening situations, and few of them result in a behavioral measure that varies from note to note as performance does in respect of a score. In our culture the most common goal of listening is some degree of aesthetic satisfaction or affective response. It is enough for many listeners that the experience was enjoyable, or evoked some emotion. When investigators take music into the laboratory they typically ask subjects to reproduce, or make comparisons between short musical passages. Although such activities can be found in some 'natural' musical contexts (such as the transmission of vocal melodies by oral imitation) they do not typify the behaviorally 'inert' listening activities that are more commonly associated with, say, hearing a new piece of orchestral music in the concert hall. The question of how we study and understand such processes has not really been addressed by mainstream contemporary research into music perception: and it is

clearly unsatisfactory, and unjustifiable, to suppose that the cognitive processes underlying the recognition and reproduction of simple melodies and chord sequences are the same as those underlying the activity of listening to a long contrapuntally organized composition. The problems of how to conceptualize the goals of listening, and how to measure the moment to moment cognitive processes that are deployed in service of those goals, urgently await solution. Luckily, in the study of music reading, these problems are, in outline, solved; at least for one central case. In this central case the goal is a performance that must satisfy certain conditions, and behavioral measures capable of shedding some light on the underlying cognitive processes can be derived from quantifiable aspects of the performance.

## 2.2 Music reading as a visual task

The advantages that are conferred on us by studying a perceptual activity with clear goals and a 'ready-made' behavioral measure are counterbalanced by the problems posed by the atypicality of the input medium. Principally, the onus is on the investigator to demonstrate that the activity of music reading is a species of *music* perception at all. The skeptic can argue that we are bound to hold a simpler hypothesis until data force us to reject it. This simple hypothesis is, in essence, that music reading is a visuo-motor task that does not engage any of the cognitive processes specific to and necessary to musical perception. On this hypothesis, visual signs are converted directly into prescriptions for action without any musical mediation. Thus, a particular symbol may be interpreted as a prescription for a particular movement pattern with respect to a particular instrument without any representation of the musical functions of the note (as, for instance, contributing to a particular harmonic or rhythmic structure). On this hypothesis, it is only *after* the notes have been produced, thereby allowing the performer to *hear* them, that any *musical* perception or cognition takes place. This hypothesis is falsifiable *only* by the demonstration that musical features of a score can be represented prior to performance and that these features can actually control performance in some systematic way.

At first glance, there appears to be an obvious falsifying counter-example to this null hypothesis. This is the case of people who claim to be able to read a score in complete silence, without mediation of instrument or voice. However, I choose not to make anything of this example here. There are several reasons for this. One is that there is no obvious way of assessing the claim. Another is that among the population of fluent music readers those who can read silently are statistically rare, and so, third they may be

employing processes in silent reading that are not normally employed in reading for performance. For these reasons, I prefer to base the case for accepting music reading as a true species of music perception on the evidence from tasks where each note in the visual stimulus must be given its own response in a performance or transcription mode.

## 2.3 Principal themes in early studies

The experimental literature on the perceptual and cognitive processes underlying music reading is not large. However, two important early studies exemplify two major areas of concern that have contemporary relevance; differences in reading processes *between* subjects as a function of expertise; and differences in reading processes *within* subjects as a function of the nature of the stimulus materials.

Bean (1938) presented pianists with briefly displayed musical extracts, which they were then required to perform. Here, as in all the studies reported in this article, the notation used was the standard 'orthochronic' notation of conventional Western music. (For a discussion of its principal characteristics see Read, 1974 and Sloboda, 1981). He found that 'span of apprehension,' as measured by the number of correctly played notes, was greater for the more experienced pianists. This important result mirrors and anticipates the findings that have emerged in many, if not all, domains of cognitive skill—experts have immediate recall for notation which is superior to that of novices. Unfortunately, Bean's study does not allow us to pinpoint the causes of this superiority as precisely as we would like. Does the superiority arise because experts have more rapid perceptual coding processes; or because they have better and more economical ways of storing what they perceive in memory; or because they have more efficient motor programs to organize the response? All these stages are plausible locations for the superiority. Subsequent research has helped clarify the situation, as the next section will outline.

The studies of Weaver and associates (Van Nuys and Weaver, 1943; Weaver, 1943) were concerned with the measurement of eye movements

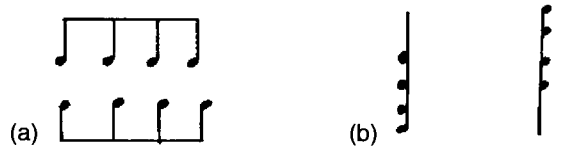



**Figure 2.1** Notational patterns typical of (a) contrapuntal music and (b) homophonic or chordal music.

as pianists played extended excerpts of classical music from score. One of the most striking findings from these studies is that the sequence of fixations made depends upon the type of music being played. Contrapuntal music elicits more horizontal than vertical fixation sequences. Chordal music elicits the reverse relationship. However, this result does not allow us to know whether *musical* considerations are responsible for this difference, or whether the visual appearance of the scores could account for the result. For instance, the eye could be 'captured' by groups of notes connected with beams of stems. Figure 2.1(a) shows an example of notation often associated with contrapuntal scores; notations such as the one in Figure 2.1(b) are more typical of chordal scores. Considerations of continuity and connectedness make it plausible that (a) would elicit a horizontal scan, and (b) a vertical scan, regardless of musical significance. Nonetheless, Weaver's work is a valuable and ecologically valid precursor of more recent studies that have attempted to manipulate the stimulus material in a more controlled way. It is a great shame that, despite the large amount of contemporary research on eye movements in language reading and scene scanning, there has been, to my knowledge, no further published research on eye movements in music reading since Weaver's work.

We can profitably combine the concerns of these two studies to generate a two-dimensional 'schema' for research into skills such as music reading (Figure 2.2). Such a schema may, ideally, be incorporated into the factorial design of a single experiment, where subjects of at least two levels of skills are examined on at least two conditions of stimulus presentation. This allows the examination of interaction effects, which are often as telling as

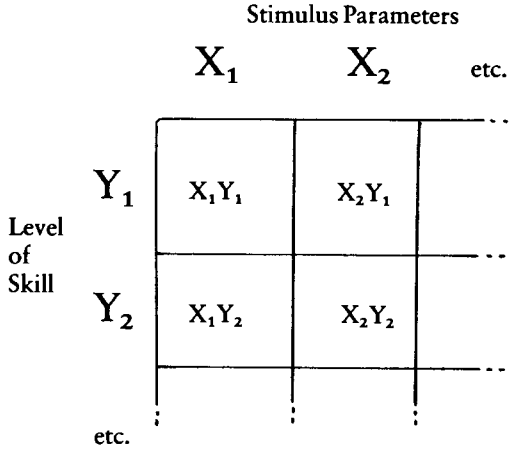



**Figure 2.2** Basic factorial schema for skills research.

the main effects of experience or stimulus material. For instance, experts may exhibit a difference in behavior toward two stimulus conditions not shown by the novices. This can allow us to pinpoint more exactly in what respect experts are different from novices.

## 2.4 Musicians compared to nonmusicians

We began our examination of more recent studies by taking a few examples of experiments that conform to the schema given in Figure 2.2. There is a small group of such experiments that set out to compare absolute novices with experienced readers on some aspect of music reading. The particular problem these studies have had to face is that of finding a way to measure the reading behavior of novices. Clearly one cannot expect absolute novices to sit down at a piano and play what they are shown. We need a response measure that does not put the nonmusician at a disadvantage, while retaining the requirement that a discrete response should be made to every note in the display. The solution adopted by Sloboda (1976a, 1978a) and Salis (1980) is to ask subjects to *copy* highly simplified excerpts of music notation (e.g. Figure 2.3) onto a bare stave after a brief viewing period. This parallels the 'reconstruction task' used by Chase and Simon (1973) in their now classic studies on chess skill. The copying task can be taught in a few minutes, and most musicians have not had extensive practice at this task.

In Sloboda's (1976a) experiments, subjects were presented with stimuli containing between one and six randomly chosen pitch symbols in the format exemplified by Figure 2.3. Under one condition subjects were able to observe the displays for 2 seconds; under another condition the stimuli were displayed for only 20 msec. Subjects were required to make a response for each note in the display, even if only guessing, and their responses were scored in terms of number of notes correctly copied. The two main effects were highly significant: musicians are better at this task than nonmusicians, and subjects perform best in the 2-second exposure condition. The interaction of these two effects was also significant: musicians performed much better than nonmusicians in the 2-second exposure condition, but they were not significantly better than nonmusicians at 20 msec, where both groups performed very poorly.

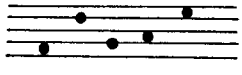


**Figure 2.3** Simplified notation for copying experiments.

These results tell us that musicians do not read better than nonmusicians simply because they have better response organization. Arguably, the copying response eliminates this particular advantage which the better musicians in Bean's study had. Thus, the advantage must relate to input coding or storage processes. The interaction also tells us that musicians are unlikely to have a highly speeded input scanning rate in comparison to nonmusicians. If they did, then we would expect superiority at the brief exposure in the number of notes correctly recorded. The results suggest, instead, that musicians have more efficient coding and storage mechanisms than nonmusicians, so that, with adequate viewing time, they are better able to retain the details of the display.

One obvious contender for the mechanism that gives musicians the advantage in this task is some specifically musical coding of the notes. Musicians might code the stimulus as a little tune, or as a series of pitch intervals. To test this hypothesis a further series of experiments was carried out (Sloboda, 1976a) using various attempts to interfere with musicians' supposed coding processes, so eliminating the advantage over nonmusicians. Thus, subjects carried out the task while listening to various kinds of music and speech; in one condition they even had to remember a short melody while doing the primary task. It was hypothesized that if musicians were using a specifically *musical* code in storing the visual display, then their performance should be selectively disrupted by concurrent musical interference. Disappointingly, the results gave no support for this hypothesis whatsoever. None of the interference conditions had the slightest detrimental effect on the performance of any subject; indeed, some subjects actually improved with interference.

Experiments using this copying paradigm have been carried out many times by students in Smith's laboratory, and it has been used as the basis of one of the laboratory workshops in the collection by Bennett, Hausfield, Reeve, and Smith (1981). Smith (personal communication) testifies that the principal results described above are robust and have been replicated many times by her students. It is particularly reassuring to have this confirmation, for the lack of interference effects is puzzling and counterintuitive. Halpern and Bower (1979, 1982) were equally puzzled by this finding, and their experiments examine the 'interference problem' further.

In the Sloboda (1976a) studies, the musical interference condition was as follows. The subject listened to a short tone sequence of a length preselected during training so as to ensure equal difficulty for all subjects. Immediately after the sequence ended the subject observed the visual display, and then made his attempt at reproduction. Finally he heard a second tone sequence which could either be identical to the first sequence

or differ from it by one note. He then had to judge whether the two tones were the same or different. So he was required to hold the tone sequence in memory while carrying out the reading task. Halpern and Bower reversed this design. In their experiment, the subject first observed the visual display, then carried out an interfering task, and finally attempted to reproduce the visual display. So the subject was required to hold the visual display in memory while carrying out the interfering task.

Halpern and Bower *did* obtain interference effects. Musicians performed worse with interference, while nonmusicians were not affected by the interference. Musicians were equally affected by two interference conditions, a visual and an auditory one. In the visual condition subjects classified a random selection of single notes into one of four sets, quarter or half-note, and high (above the middle line) or low (below the middle line). In the auditory condition subjects classified a similarly random collection of played notes into one of four sets, combining long–short in duration with high–low in pitch.

The most plausible explanation for the difference between the two studies is that musical interference disrupts retention and rehearsal processes, but does not disrupt coding and storage for immediate output. In other words, interference disrupts memorization for recall (in Halpern and Bower's case, after 15 seconds), but does not disrupt reading for immediate transcription. A possible reason for this is that musicians have overlearned mechanisms for immediate transcription gained precisely from years of reading experience, although they do not have well-developed mechanisms for retaining sections of score for recall many seconds later. This latter task is novel to most musicians, and performance on it, arguably, does not tell us much about reading mechanisms. In general it seems that automated and overlearned tasks are less susceptible to interference than novel tasks. What the Halpern and Bower study does suggest is that memorization from score does involve specifically *musical* mechanisms since only the musicians were disrupted by the two interference tasks, both of which could have engaged musical knowledge. However, it is possible that such knowledge is only mobilized *during* the 15-second retention period and does not contribute to normal reading.

A further study (Sloboda, 1978a) returned to another rather puzzling feature of the first experiments' results, the lack of any difference between musicians and nonmusicians at the 20-msec exposure. This was not a floor effect, since subjects were all performing significantly above chance. The result is puzzling because there is a body of literature that demonstrates that subjects are able to identify more of the information contained in a very briefly exposed display when the constituents of the display are

familiar than when they are not (Allport, 1968; Welsandt, Zupnick, and Meyer, 1973). Since musicians are much more familiar with music notation than nonmusicians, why do they not outperform the nonmusicians? One possibility is that they *do* outperform the nonmusicians, but that a more sensitive measure than number of notes correct is required to demonstrate that superiority. Accordingly, in the Sloboda (1978a) study, various methods of scoring were used, giving credit for nearness in pitch to the correct note and preservation of the up–down contour of the original. Sure enough, some scoring methods did yield the hoped for superiority. Most significantly, musicians were better than nonmusicians at retaining the correct up–down relationships between adjacent pairs of notes, even though not necessarily positioning either note of a pair on the right line or space. These results can be interpreted as showing that 20 msec is too short a time for accurate localization of notes on a stave, but that it *is* sufficient time for global attributes of contour and shape to be registered. This is consistent with our understanding of visual perception in other domains (e.g., Navon, 1977; Palmer, 1977) and does not necessarily implicate musical knowledge in the process. Familiarity with the visual characteristics of music notation could be a sufficient explanation of the musicians' superiority.

It is also noteworthy that, in contrast to studies of experts and novices in other domains like chess, musicians show a superiority over nonmusicians even when, as in the Sloboda (1976a, 1978a) studies, the stimulus notes are chosen at random and do not form musically coherent patterns (in that they do not represent common structures or sequences within the system of diatonic tonality). Why this should be is not at all clear. Maybe the complete absence in nonmusicians of any knowledge about the names or functions of the individual symbols puts them at a disadvantage in *any* situation involving music notation. In contrast, it is hard to find a novice chessplayer with similar levels of ignorance about chess notation. Nevertheless, it is still worth asking whether musicians gain an additional advantage over nonmusicians when the stimuli 'make musical sense.' Halpern and Bower carried out a second experiment that directly addresses this question. They used immediate recall without interference, thus making the task of more direct relevance to music reading, and found indeed that musicians performed better on 'good' than 'bad' melodies, whereas nonmusicians showed no difference between the melody types.

This result provides the strongest indication so far that musical knowledge is implicated in reading, even though the reconstruction task is not identical in its requirements to normal reading.

## 2.5 Good readers compared to poor readers

Nonmusicians are severely limited in the range of responses they can give to music notation. To make direct comparisons between musicians and nonmusicians we must grossly simplify both stimulus and response in comparison to the habitual fare of the performing musician. Accordingly, this line of investigation cannot take us very far. Fortunately, music reading skills are distributed very unevenly through the population of musicians. Some otherwise competent performers are very poor readers. Others excel at reading. The task on which differences emerge most strikingly is that of 'sight-reading,' where the performer is required to provide a coherent performance of a piece of music he or she has never seen before. Sight-reading is to be contrasted to the case where a performer uses a score to guide performance of an already learned piece. In the latter case, the performer is using the score (probably selectively) as an *aide-mémoire* rather than as the primary source of information about the music. One may often observe that a performer hardly looks at the music for pages at a time. Clearly, his primary source of information about the music is stored in his long-term memory. The score is used, from time to time, simply to act as a retrieval cue for the next section of music. Sight-reading is, in contrast, a 'cleaner' and 'simpler' case of reading. The performer must examine the great majority of, if not all, the notes, and the success of his performance is directly related to the knowledge gained from this examination, uncontaminated with long-term knowledge about the particular piece of music. Of course, long-term knowledge of a *general* kind may be implicated, as it is in prose reading, and the use of such knowledge is, as we will see, precisely what distinguishes the expert from the ordinary player. The remaining studies reviewed in this paper are all concerned with sight-reading. Other types of reading, important as they are, are not treated further.

In any continuous transcription task, such as copy typing, simultaneous interpretation, or music reading, there is a time lag between the perception of a symbol and its subsequent realization in output. When a particular symbol is being spoken or played, the eye or ear is already receiving input from a symbol further ahead in the input. A simple way of demonstrating the existence of this lag is to suddenly remove the stimulus while a reader is engaged with it. Typically, performance does not stop dead exactly simultaneously with the removal of the stimulus. The reader is able to produce a little more output. The average amount of extra output available is known, in prose reading, as the 'eye–voice span' or in typing or instrumental music reading the 'eye–hand span.' For normal prose the

eye–voice span (EVS) is about five or six words in adult readers. For simple tonal melodies the average instrumental eye–hand span (EHS) is about five or six notes. In a series of studies (Sloboda, 1974, 1977) it was found that EHS varied with both expertise and the nature of the material to be read. In one study (Sloboda, 1974), reading ability was estimated by counting the number of performance mistakes in a large corpus of sight-reading performances. People who made many mistakes had lower EHS (3–4 notes) than those who made few mistakes (6–7 notes). In addition (Sloboda, 1977), it was found that performers had greater EHS when reading tonally coherent music than when reading music that broke rules of tonal progression. In both these studies the experimental design was such that subjects were deprived of the score at various distances prior to a musical phrase boundary. It was found that there was a greater than chance likelihood of EHS coinciding with a phrase boundary. This effect interacted with reading ability. The best reader ended his performance at a phrase boundary on 72% of occasions, the worst reader on only 20% of occasions. Thus, for good readers, the EHS was not constant; it expanded and contracted to accommodate a phrase unit.

Here is incontrovertible evidence that musical knowledge is implicated in sight-reading. How, otherwise, would a reader know that a particular note marked a phrase boundary, prior to having heard the music? It is, of course, true that much printed music contains visual signs marking phrase boundaries other than the notes themselves, just as printed text contains punctuation marks. But in these experiments all such signs were eliminated from the score, and in one condition (Sloboda, 1977) even the long notes which often mark phrase endings were eliminated, so that the reader would have no clue to phrase subdivision other than the melodic, harmonic, and rhythmic logic of the note sequence itself. We conclude, then, that good sight-readers are particularly attuned to important superordinate structures within a score, structures that link notes together into musical groups. They organize their perception and performance in terms of discovering these higher order groupings, with consequential economy of coding.

It does not seem likely that these results could be explained purely in terms of response factors, such as the tendency for experts to improvise phrase endings that were not seen, or to stop short at a phrase close. On the first count, there are several plausible ways to improvise an ending to most phrases, but the data show almost no instances of erroneous yet plausible responses among the better readers. Their performances were almost error free. On the second count, readers were under strict instructions to play *all* the notes they saw and *only* the notes they saw. The subjects believed themselves to have complied with these instructions.

In sum, then, the EHS studies show that good readers 'look farther ahead' than poor readers and are more sensitive to the musical structure of what they read than poor readers. It is plausible to infer that the latter phenomenon is, in part, responsible for the former.

Turning away from EHS, we may ask whether good readers differ from poor readers in the *type* of sight-reading mistakes they make, in addition to their frequency. This question has been examined mainly in relation to 'proofreaders' error,' the tendency for readers to misperceive small inaccuracies in text as the 'correct' or 'intended' sequences. In prose, many readers fail to see small spelling or typing errors. Do musicians also fail to see 'misprints'?

Wolf (1976) reports the experience of a distinguished piano teacher and sight-reader, Boris Goldovsky, who discovered a misprint in a much used edition of a Brahms *Capriccio* only when a relatively poor pupil played the (musically impossible) printed note at a lesson. Goldovsky stopped the pupil, thinking her to have misread, but soon discovered that it was *he*, and countless other pupils and colleagues, who had misread, inferring a sharp sign in front of a note because in the musical context it *had* to be a G sharp, not, as printed, a G natural. So struck was he with his misperception, that he devised the 'Goldovsky experiment', which consisted in telling skilled readers that there was a misprint somewhere in the piece and asking them to find it. He allowed them to play the piece as many times as they liked and in any way that they liked. No musician ever found the mistake. Only when he told his subjects which bar (measure) the mistake was in did most of them spot it. (The piece is Brahms' op. 76, no.2, and the mistake occurs in bar 78, which is reprinted in Sloboda, 1978b).

Misprints of this nature are, thankfully, rare in printed music. To ascertain the generality of this effect, it is necessary to construct scores with 'deliberate' misprints. This was the purpose of an experiment reported in Sloboda (1976b). In this experiment subjects were presented with pieces of piano music that were littered with misprints. The pieces of music were in conventional classical style, composed by lesser known contemporaries of Mozart, and each misprint mislocated a note by one scale step on the stave in such a way that the resulting sequence was blatantly out of character, in virtue of containing an unprepared and unresolved discord. The subjects were all competent pianists and they were asked to give two performances of each piano piece at sight, being particularly careful to play exactly what was written. Although the overall level of error in performance was very low for all subjects (2.9% for first performance, 1.7% for second performance) the level of error on the misprints was high (38% for first performance, 41% for second performance). All the misplayings of these misprints conformed

to proof readers' error, that is, subjects substituted the original correct notes, although, of course, they had never seen or heard the original music.

It is noteworthy that, although overall errors decreased on a second playing, proofreading errors *increased*. This suggests that what subjects had learned from a first playing were *not* only individual notes, but also some things about the overall melodic and tonal structure in virtue of which the misprint became even less likely to be spotted, contradicting, as they did, the tonal coherence of the pieces. Here, then, is further evidence that musical knowledge is implicated in reading. The Goldovsky experiment suggests that it is skilled, rather than unskilled, readers who are most likely to be susceptible to proofreaders' error, presumably because they are more likely to be reading for superordinate structures.

## 2.6 Master readers compared to good readers

Studies on skill such as those of Chase and Simon (1973) have taught us that we must expect significant changes in the structure of skill beyond the point of mere proficiency. There are, in any skill domain, those practitioners who by virtue of their excellence and experience deserve the title 'master.' A factor correlating very strongly with mastery is quite simply the time spent practicing the skill in question. A master has probably spent upward of 10 000 hours engaged with his skill, whereas the merely proficient practitioner has probably only spent some couple of thousands of hours on the same skill. We may expect to find as many significant differences between the master and the merely proficient as we do between the proficient and the novice.

For sight-reading, such differences have been studied with respect to expressive aspects of piano performance (Sloboda, 1983). For a performance to be musically effective, it must do more than present a listener with the right notes at the right speed; the performer must modulate tempo and dynamic in such a way that he provides a coherent, and possibly novel, insight into the structure of the music. The level of structure I wish to concentrate on here is that specified by the meter of a piece of music. Most conventional notated music specifies a subdivision of the musical stream into bars (measures) each containing the same number of beats (pulses) of notionally equal duration. The first beat of each bar is marked for the highest prominence (stress) within that bar, and for a listener to understand the music he must be able to identify the metrical contour from the first few notes of a performance. Sometimes the notes themselves allow only one coherent metrical reading (Steedman, 1977) but often the music is metrically ambiguous until the performer adds some expressive content. The most

ambiguous sequence of all would be one made up of the same note repeated over and over again. Unless the performer modulates such a sequence in some way a listener may infer any meter he likes. Here, then, is one condition which, it can be argued, must be satisfied if a performer is to be said to have read the music effectively. Metrical information specified by the score should be specified in the performance too.

Shaffer (1980, 1981) has devised a means whereby all the objective features of a piano performance can be recorded on computer. If we deny a pianist the use of the pedals, then he has only three dimensions of variation available for expressive performance; the loudness of a note, onset timing, and offset timing. Shaffer's equipment allows one to record all these features by means of photocells attached to the piano mechanism. Using Shaffer's system it was possible to obtain full records of the expressive devices used by a group of pianists asked to sight-read a metrically ambiguous melody, notated twice, with different placement of metrical stress. Meter was signified in the score by the placement of bar lines and beams, and in no other way. The second notation was generated from the first by shifting the bar lines and beams one semiquaver (eighth-note) to the right in the score.

Subjects were asked to play each melody at sight five times in a row. The melodies were embedded in a longer series of test melodies; and no subject noticed that the two melodies contained the same notes. The instructions encouraged the performers to give each melody some musical character, to be reproduced as exactly as possible on the five consecutive performances, but no mention was made of meter. In a second experiment, audio recordings of the performances were played to listeners who had to try and identify which of the two notated forms each performance came from.

The subjects themselves ranged considerably in playing experience. The least experienced subject had been playing for about 11 years, and the most experienced subject, a professional pianist, had been playing for more than 40 years. The results showed many significant differences as a function of experience. First, the number of notes where the performance of the two melodies differed increased with experience. The masters made the two performances more distinct from one another. Second, the types of expressive devices used increased in number with experience. In particular, the less experienced players tended to concentrate on loudness variations at the expense of timing variations. The more experienced players used loudness variations too (stressed notes louder than unstressed notes) but supplemented them with timing variations (such as the slight delaying in the onset of the note *following* a stress). Third, the most experienced player was significantly better at communicating meter than the least experienced player. Listeners identified many more performances correctly.

Expressive aspects of piano performance result from the detailed planning of finger movements in advance of sound production. The pianist must know how loud he wants a particular note to sound, when he wants it to sound, and for how long, *before*, not after, he actually hears it. If it can be shown, as I believe we have shown here, that his allocation of expressive variation is governed by a fundamental aspect of musical structure such as meter, then it seems indisputable that knowledge of such musical structure is a direct consequence of the reading process, so proving music reading to be a real species of, and window onto, music perception.

## 2.7 Summary

This review of the small literature on music reading has been concerned to show that music reading is a genuine species of music perception, by demonstrating the various ways in which musical structure affects performance in readers of varying levels of expertise. Written recall of notes is improved when the sequence to be recalled is tonally coherent (Halpern and Bower, 1979, 1982); the eye–hand span is sensitive to musical phrase structure (Sloboda, 1974, 1977); errors in reading tend to be musically plausible (Sloboda, 1976b; Wolf, 1976); and expressive aspects of sight performance respond to structural aspects of the noted music (Sloboda, 1983). In general, the effects of musical structure are most pronounced in the more experienced readers, suggesting increasing mediation of representations couched in terms of formalisms which define relationships between groups of notes. Such mediating representations are being discovered to play an essential part in many music perception tasks (cf. Deutsch, 1982), and it is to be hoped that a fuller articulation of the nature of these representations will be instrumental in advancing our understanding of the whole range of skills (such as composition and improvisation) that can properly be called musical (cf. Sloboda 1982, 1985).

## References

Ailport, D. A. The rate of assimilation of visual information. *Psychonomic Science*, 1968, **12**, 231–232.

Bean, K. L. An experimental approach to the reading of music. *Psychological Monographs*, 1938, **50**, Whole number 226.

Bennett, A., Haustield, S., Reeve, R. A., and Smith, J. *Workshops in cognitive processes*. London: Routledge and Kegan Paul, 1981.

Chase, W. G. and Simon, H. A. The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press, 1973.

Deutsch, D. (Ed.) *The psychology of music.* New York: Academic Press, 1982.

Halpern, A. R. and Bower, G. H. Interference tasks and music reading. Unpublished manuscript, Stanford University, 1979.

Halpern, A. R. and Bower, G. I-I. Musical expertise and melodic structure in memory for musical flotation. *American Journal of Psychology,* 1982, **95**, 31–50.

Navon, D. Forest before trees: the precedence of global features in visual perception. *Cognitive Psychology,* 1977, **9**, 353–383.

Palmer, S. E. Hierarchical structure in perceptual representation. *Cognitive Psychology,* 1977, **9**, 441–474.

Read, G. *Music notation.* London: Gollancz, 1974.

Salis, D. L. Laterality effects with visual perception of musical chords and dot patterns. *Perception and Psychophysics,* 1980, **28**, 284–294.

Shaffer, L. H. Analysing piano performance: A study of concert pianists. In G. E. Stelmach and J. Requin (Eds.), *Tutorials in motor behaviour.* Amsterdam: North Holland, 1980.

Shaffer, L. H. Performance of Chopin, Bach and Bartok: Studies in motor programming. *Cognitive Psychology,* 1981, **13**, 326–376.

Sloboda, J. A. The eye–hand span: an approach to the study of sight reading. *Psychology of Music,* 1974, **2**, 4–10.

Sloboda, J. A. Visual perception of musical notation: registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology,* 1976a, **28**, 1–16.

Sloboda, J. A. The effect of item position on the likelihood of identification by inference in prose reading and music reading. *Canadian Journal of Psychology,* 1976b, **30**, 228–236.

Sloboda, J. A. Phrase units as determinants of visual processing in music reading. *British Journal of Psychology,* 1977, **68**, 117–124.

Sloboda, J. A. Perception of contour in music reading. *Perception,* 1978a, **7**, 323–331.

Sloboda, J. A. The psychology of music reading. *Psychology of Music,* 1978b, **6**, 3–20.

Sloboda, J. A. The uses of space in music notation. *Visible Language,* 1981, **15**, 86–110.

Sloboda, J. A. Music performance. In D. Deutsch (Ed.) *Psychology of music.* New York: Academic Press, 1982.

Sloboda, J. A. The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology,* 1983, **35A**, 377–396.

Sloboda, J. A. (1985) *The musical mind: An introduction to the cognitive psychology of music.* London: Oxford University Press.

Steedman, M. J. The perception of musical rhythm and metre. The perception of musical rhythm and metre

Van Nuys, K. and Weaver, H. E. Memory span and visual pauses in reading rhythms and melodies. *Psychological Monographs,* 1943, **55**, 33–50.

Weaver, H. F. A study of visual processes in reading differently constructed musical selections. *Psychological Monographs,* 1943, **55**, 1–30.

Welsandt, R. F. Jr., Zupnick, J. J., and Meyer, P. A. Age effects in backward masking (Crawford Paradigm). *Journal of Experimental Child Psychology,* 1973, **15**, 454–461.

Wolf, T. A. cognitive model of musical sight reading. *Journal of Psycholinguistic Research,* 1976, **5**, 143–171.

CHAPTER 3

# THE USES OF SPACE IN MUSIC NOTATION

## 3.1 Introduction

In this article I want to describe how space is used in music notation. Music notation is unfamiliar to many people and so I have attempted to make what I say intelligible to the non-musician by defining all the technical terms I use. Music notation (score) is very different from language notation (text) and so some prefatory comparative remarks may help to place this notation in a wider context.

Perhaps the most fundamental difference between score and text is that a score must be able to specify different events as occurring simultaneously whereas text portrays a single sequence of events. The problem of how to link up parallel streams of information has thus been fundamental to the development of score. No analogous problem exists for text. A second difference concerns use. Score readers are mainly concerned with producing a musical performance. Text readers are more concerned with understanding and remembering what they read. This difference makes issues of layout of foremost importance in a score. The music reader cannot afford to lose his place or experience ambiguity even for a second if he is to maintain the flow of performance. This demand has expressed itself in intense and prolific experimentation with different ways of arranging score material over the centuries. In contrast, the solitary text reader is able to pace his own reading to accommodate deficiencies in layout. Thus there has not been the same historical incentive to experiment with layout. Unhelpful layout may be a nuisance: but it hardly ever leads to total breakdown. A third difference concerns the levels at which spacing and layout become important. Score material presents the reader with diverse spatial information at a microscopic level. Any two-centimeter square of a modern score will contain a rich array of symbols at varying distances and directions from one another. In contrast a similarly sized portion of text is most likely to contain evenly spaced letters in evenly spaced rows. Interesting spatial differentiations occur, if at all, at the macroscopic level, and concern such matters as how paragraphs are sited. Another way of putting this difference

is to say that in a score there are complex spatial constraints which determine the positioning of each notational element with respect to its neighbors, whereas in text the determination of the position of one letter in relation to its neighbors is trivially simple. Thus, layout is an integral part of the music notation system. In language notation it is an optional extra.

There are, of course, many points one could make which would tend to blur the clear lines drawn up by these distinctions. There is not one use for scores any more than there is for texts. It would be fair, however, to claim that these differences do characterise broad and prevailing tendencies within the two systems.

What follows from these comparisons? Firstly, spatial considerations in music notation are more complex and multi-levelled (and arguably more interesting) than spatial considerations in text. Secondly, perhaps as a result of the complexity, there has been little, if any, controlled experimentation using different spatial arrangements of music. And so, thirdly, this article cannot draw on published psychological research. Rather its aim is to present an historical and conceptual perspective on some of the major motivations for, and problems with, using space as a notational tool. In this task I draw most heavily on the work of musicologists and music historians, in particular the monumental study of Apel (1953). Other useful source and background material has been provided by Abraham (1979), Cole (1974), Hyatt-King (1964), Karkoschka (1972) and Read (1974).

Western musical civilization of the last 1500 years has produced a bewildering multiplicity of notational systems devised in different contexts and for different purposes. Nonetheless, today we would recognize one particular system as central. This is, to adopt Read's (1974) terminology, the *orthochronic* system. Figure 3.1 shows a typical portion of this notation. It has been the major western notational system for over 400 years, holding its place against all competitors even in this [twentieth] century which has seen a mushrooming growth of new notations. As the discussion proceeds I shall touch on various aspects of orthochronic notation, looking both at their development and at notational systems which have used space in different ways to achieve similar ends. Also, since any notational system can only meet limited needs, it is necessary to compare orthochronic notation with systems whose aims are different. Differing aims place differing demands on the available space.

## 3.2 Directions for action

A most important fact about orthochronic notation is, apparently, negative. It is impossible to tell, just by examining a typical extract, which instrument

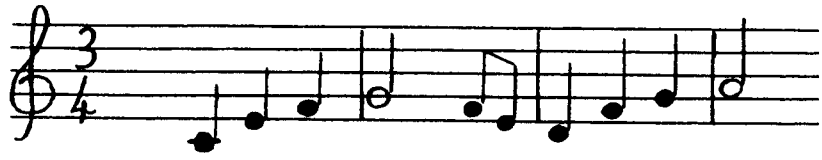


**Figure 3.1** A section of orthochronic notation showing the principal notational features.

the extract is intended for. The central features of the notation are abstract. They specify pitch and rhythmic relationships between notes and groups of notes, *not* what keys are to be pressed on an instrument nor precisely how the notes are to sound. They are intended to convey something of the musical structure to a reader. Psychological implications of this fact have been discussed by Sloboda (1978a, 1982). This characteristic probably explains why orthochronic notation has retained its central position. It is a lingua franca for all musicians, as understandable to a violinist as to a singer.

The corollary of this is that each instrumentalist must have additional practical knowledge before he can actually play the music. For instance, orthochronic notation may tell a flautist to play the note D, but it does not tell him which fingers to place over which holes to achieve the D. The wish to have a more direct notational aid to performance has sometimes led to the development of notations based upon a different principle—that of specifying actions on a particular instrument. Such notations achieve this end at the expense of universality. A notation devised for the violin would be meaningless to a trumpet player.

One large sub-group of such systems are the tablatures. Figure 3.2 is part of a French lute tablature of the sixteenth century. The horizontal lines represent the strings of the lute. The letters designate frets on the finger-board. In this example, the frets are marked by the letters B to H, with A denoting an open string. Tablatures use letters or numbers to specify at least part of

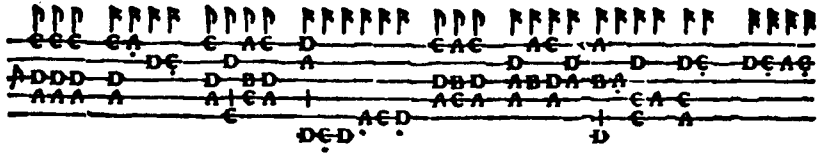


**Figure 3.2** 'Fortune a bien couru sur moi' from *Tres breve et familiere introduction pour entendre et apprendre par soy mesme a iouer toute chanson reduictes en la tabulature du Lutz.* Source: Apel (1953), p. 65.

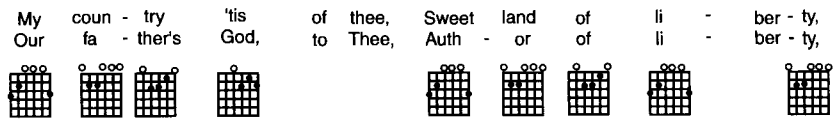



**Figure 3.3** Chord sequence for 'My Country, 'tis of thee' (first line) in modern guitar notation.

the actions required. Some modem successors have eliminated alphabetic symbols completely in favor of graphic representation. A simple example of this is modern guitar notation (Figure 3.3) in which strings and frets are represented on a two-dimensional matrix, with dots to show the finger positions.

A more complex example is a modern system for the piano called Klavarscribo. In this system vertical lines represent the black keys on the piano, the spaces between representing white notes. The score is read from top to bottom, and the small circles represent the keys to be struck (unfilled circles are white notes). Figures 3.4 and 3.5 show the same extract from F. Chopin's Sonata op. 58 written in orthochronic notation and Klavarscribo. To the eye there is little to suggest that it is the same music in both cases.

When we go on to consider the detailed requirements of a developed music notation system, I shall wish to return to these examples to argue that the price paid for instrumental specificity is, in general too high, since it restricts the number of dimensions remaining for representing other important aspects of the music. An alternative approach which has worked very well for many instruments is to incorporate additional symbols into

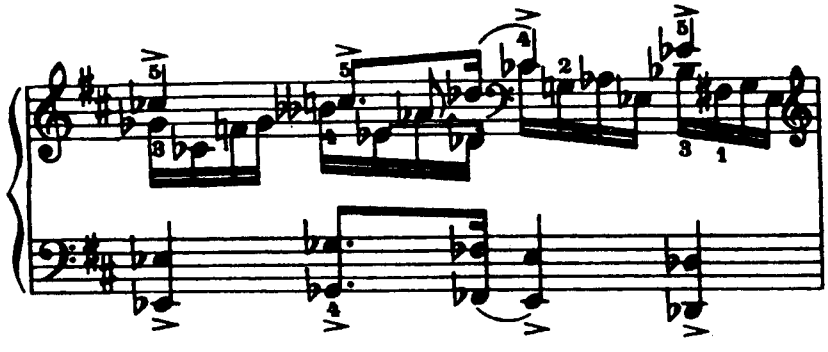


**Figure 3.4.** A bar from F. Chopin, Sonata op. 58, in orthochronic notation. Source: Karkoschka (1972). p. 12.

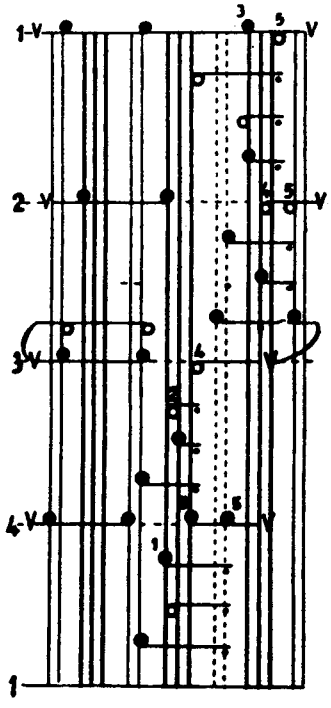


**Figure 3.5** The same music as Figure 4 notated in Klavarscribo. Source: Karkoschka (1972), p. 12.

the orthochronic system without altering its basic characteristics. Thus, for instance, many piano scores contain small numbers over some of the notes, which prescribe fingering. There are also many simple symbols which can be placed over or under notes to specify such things as the nature of attack, loudness, phrasing. The spatial constraints here are rather indefinite. If there are too many of these additional symbols the score takes on a cluttered appearance, and there is a danger of overloading the reader with detail. Compare Figures 3.6 and 3.7 which are different editions of the same music. Figure 3.6 is crowded with specific performance directions. Figure 3.7 uses such directions sparingly. In general, editors must take a middle line between giving the performer no help at all (trusting to his own musical competence) and giving him so much that some of it is bound to be ignored. The trend recently has been towards the former approach, at least in respect to classical music. It is not coincidental that Figure 3.6 predates Figure 3.7 (a contemporary edition) by some decades.

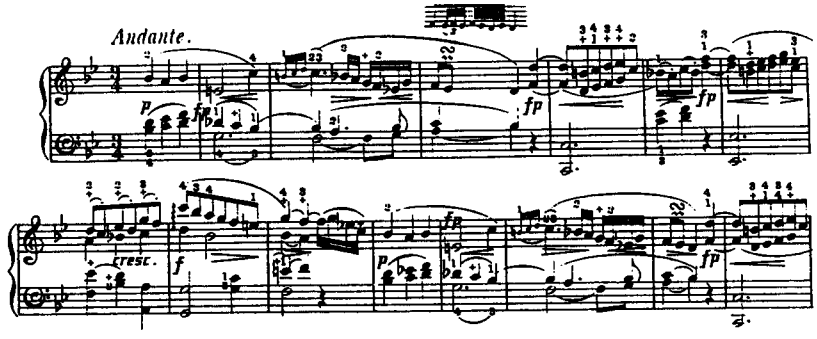



**Figure 3.6** Opening bars of 'Andante' from W. A. Mozart, Sonata in F, K. 533. (Ed. A. Zimmerman). London: Novello. Undated (c. 1900).

## 3.3 The staff

The most noticeable and characteristic feature of any segment of orthochronic notation is the grid of five horizontal lines on which the notes are placed. These lines are called staff-lines, with the five lines known, collectively, as a staff (plural staves). Unlike the tablatures, the function of the staff in orthochronic notation is to represent the pitch dimension. From earliest times pitch was seen as one of the most important aspects of music to notate. The earliest notations were alphabetic, dating at least as far back as pre-Christian Greece. In these systems each note of the scale was designated by a letter or other symbol. To notate a melody one simply wrote out the letters in the correct order. The first use of space to indicate pitch was the development of *neumes* in around the

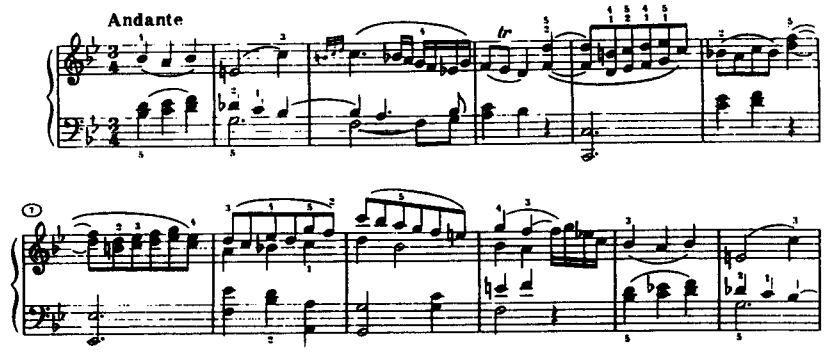



**Figure 3.7** The same music as Figure 3.6 (Ed. W. Lampe). München: Henle, 1955. Reproduced with the publisher's permission.

eighth century AD. These were used almost exclusively in the liturgy of the Christian Church to notate the plainchant melodies to which prayers were sung in religious communities. Figure 3.8 shows one of the earliest examples of neumes from a French manuscript c.871. The neumes are written above the Greek text and indicate the directions of pitch change. Thus ⌣ denoted a rise, ⌝ a fall, and ✓ a fall followed by a rise.

This system had no power to specify the extent of a pitch change, only its direction. Read (1974) comments 'Useful solely as a reminder to the singers, refreshing their memory of a general rise and fall originally learned by rote, the neumes provided only the over-all contour of the plainsong melodic line—not an exact map by which a novice might approach an unknown musical territory' (p. 7). This comment introduces a very important consideration, that of the context in which a notational system is used. Neumes did not do badly what orthochronic notation does better. They probably served the cultural and psychological requirements of their users better than modern notation would have done. They had an economy of symbolism which exactly served its purpose within monastic choirs, where novices would learn most of what they needed through repeated hearing of the same hymns. It was only, possibly, as chants elaborated and diversified that more precise mnemonic systems were necessary.

The next conceptual development was the use of vertical distance to indicate the extent of the pitch change. Figure 3.9 shows an early example from the tenth century which places neumes relative to a single horizontal line given a precise pitch (F below middle C in this example). The distance of a neume from this line gives an indication of its pitch distance from F.

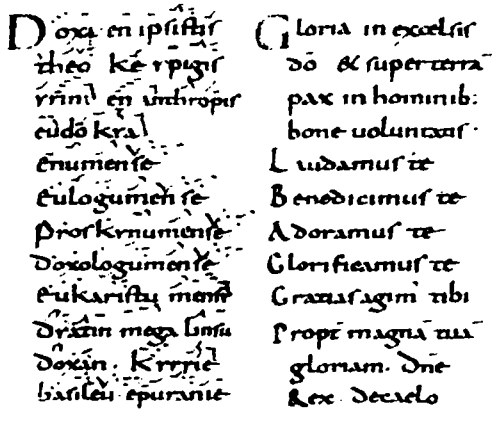


**Figure 3.8** 'Gloria' from a St. Amand manuscript c. 871. Paris: Bibliotheque Nationale, ms. lat. 2291. Reproduced with the Trustees' permission. Source: Abraham (1979), p. 62.

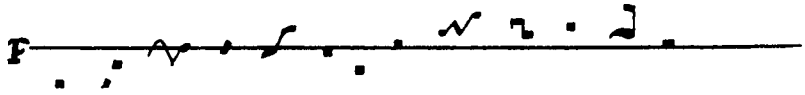


**Figure 3.9** Tenth century neume notation (author unknown). Source: Read (1974), p. 10.

The difficulty with this, from a modern musician's point of view, is that the y-axis is not calibrated. A reader cannot easily tell what range of pitches is covered, nor the exact distance of any particular neume from the baseline. Both these problems were solved over the next two centuries by the addition of further horizontal lines denoting other precise pitches. By the 12th century the number of lines had reached four. This number seems to have been supported by two principles. One was that lines were provided for alternate pitches in a scale. Thus, the bottom line might represent A, the space above it B, the next line C, the next space D, and on. This principle remains in operation today in orthochronic notation. It means that the pitch of a note need never be estimated by reference to its physical distance from another point on the paper. It can be identified precisely by counting up the staff lines. For experienced readers, even this is not necessary. The position of the note is directly known by a kind of spatial subitising. The look of each position on a staff becomes so distinctive that no conscious counting is required.

The second principle was that the number of staff lines was tailored to the pitch span of the melodies. Most melodies could be encompassed within nine successive pitches of a scale. This span (about an octave) would be the comfortable range of an average untrained male voice. Even in the 12th century most notated music was church music sung exclusively by monks. As time passed, however, it became more customary for instrumental music to be noted on a staff. Instruments have larger pitch spans than voices, and so more lines were needed to encompass the melodies that could be played on instruments. Six lines were fairly common and up to fifteen were known. Somehow things settled down and the five lines of modern orthochronic notation became the norm. Although it is impossible to know this with certainty, it seems probable that one of the reasons why the five-line staff survived is that it represents the best compromise between good span and readability. Figure 3.10 presents a ludicrously extreme example of the readability problem. This is part of a composition by E. Brown (1952). There are fifty staff lines, and it is practically impossible for a reader to keep any bearings at all. Luckily for the performer, Brown does not intend the lines to be taken seriously, but

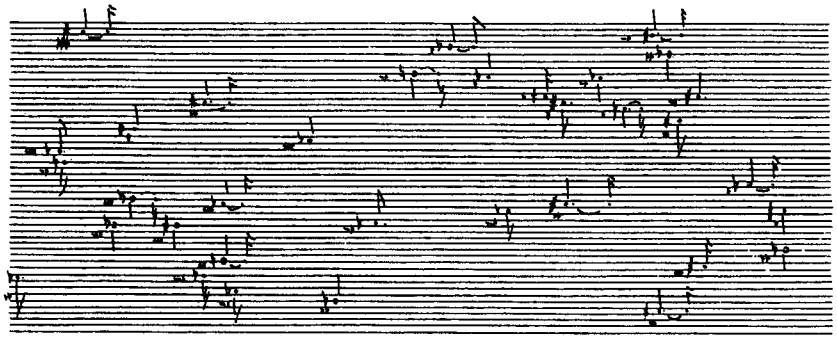


**Figure 3.10** 'November 1952' from Folio by E. Brown. New York: Associated Music Publishers Inc. Reproduced with the publisher's permission. Source: Karkoschka (1972), p. 90.

the moral is clear enough. There is a limit to the number of lines a staff can contain if a performer is to use it to rapidly identify the pitches of successive notes. This presents a problem within orthochronic notation when melodies span more than eleven notes of a scale, as frequently happens. The problem has been solved in three ways, all in common use today.

The first solution is to recalibrate the staff as the melody rises above or falls below the range of the staff. This is conventionally done by arbitrary symbols known as clef signs. The leftmost symbol in Figure 3.1 is the treble clef sign, and it indicates that the bottom line of the staff is to be read as the E above middle C. There are four or five such symbols in common use, and when one of them occurs in the course of a piece of music it instructs the reader to recalibrate everything from that point onwards until another clef sign is encountered.

For tunes which skip about a lot, clef signs are, however, rather unsatisfactory. If the clef must be changed every few notes the reader will spend as much time interpreting the clef symbols as the notes themselves. Another solution which is in common use is to temporarily extend the staff by ledger lines. The first note of Figure 3.1 is on a ledger line and the reader is asked to interpret it as if it were another staff line below the lowest one shown. In fact, the reader may imagine the staff is extending indefinitely both above and below its visible portion, but to avoid the problems associated with Figure 3.10 only that portion of it becomes visible which allows a single note to be identified. Thus, ledger lines extend only as far on either side of a note as to make them clearly visible. Most fluent readers become adept at rapidly interpreting up to five ledger lines. With more lines, problems of readability again develop. Figure 3.11 shows two

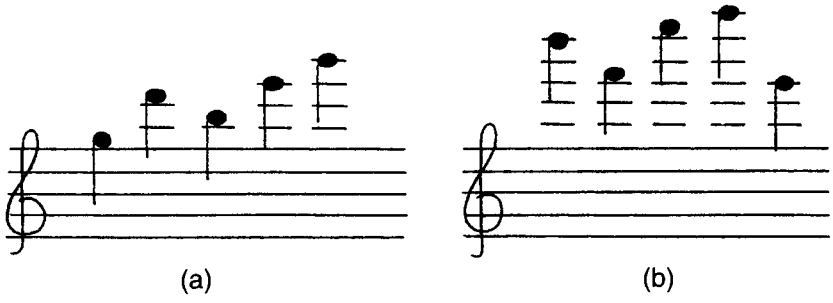



**Figure 3.11** Two examples of the use of ledger lines: (b) is more difficult to read than (a).

extracts using ledger lines; (a) would be easily read by most competent readers today; (b) would not be fluently read.

At this juncture a comment about the vertical spacing of staff lines and ledger lines is pertinent. Although not logically necessary, it has always been standard practise to make the distances between adjacent staff-lines equal to one another. This has meant that the vertical distance between notes is an accurate measure of the pitch distance between them.

Although this information is technically redundant, it does provide readers with an additional analog cue to pitch which arguably supplements the 'digital' information supplied by the lines themselves. There is, for instance, some evidence that music readers can be aware of the contour of a notated melody before they can precisely identify the constituent notes (Sloboda, 1978b). It is unclear how much use competent musicians make of this information in performing situations, but one informal observation supports its importance. It sometimes happens that staves are printed so close together that there is a difficulty about placing a note on, say, the fifth ledger line. If the spacing of the ledger lines were to preserve the spacing of the staff lines, they would come too close to, or overlap, the adjacent staff. In such a situation the ledger lines are squashed closer together. This, of course, upsets the analog representation of pitch, and many musicians complain bitterly that such notes are very difficult to read, especially if the music is unfamiliar. In the best publishing houses this crowding of ledger lines never occurs. The solution is, of course, to move the staves further apart, but with fewer staves to the page, production expenses increase. This probably accounts for the persistence of this troublesome practice. An allied point concerns the size of a staff. This has both psychological and economic implications. If staff lines are too close together the reader has problems of discrimination, especially if he must be at some distance from

the score when performing. Conversely, if the lines are too far apart then it may be impossible to see the stave clearly with a single fixation. Similarly, from the printer's point of view, closely spaced lines demand finer (and thus more expensive) printing. Widely spaced lines mean fewer staves per page and thus more pages. Engraving tradition seems to have settled on a between-staff-line distance of about 0.20 to 0.25 cm as suitable for most performing scores.

The third solution to the problem of increased melodic span we owe to the rise of polyphonic music in the middle ages. Polyphonic music is that where several streams of notes are performed simultaneously. This type of music can be vocal, where different singers take different melodies simultaneously. It can be for an instrumental group, where, for instance, different pipe instruments play different notes. Or it can be for a keyboard instrument played by one performer producing simultaneous sounds with different hands and fingers.

In all these cases, the notational problem is to indicate which person or hand takes which stream of notes. To put all the notes on to a single staff would be to overcrowd it hopelessly if more than a couple of parts were involved. So the solution arrived at fairly early was to give each part a separate staff. These staves were vertically aligned and joined in some way to indicate simultaneity. A typical example of this is given in Figure 3.4. This is part of a piano piece—the curved bracket at the left indicates that the two staves are to be played simultaneously, the upper by the right hand, the lower by the left hand. This arrangement is necessary when a single performer must read both staves simultaneously: the relevant portions must clearly be as close to one another as possible. It also provides an ideal vehicle for extending the span of pitches which can be notated. This is because the two staves can be calibrated differently by clef signs, as in Figure 3.4. Suitably calibrated, the piano score has over twice the pitch span of a single staff. In the notation of symphonic music this system can be extended indefinitely, twenty staves being a not uncommonly large number for a conductor's score. Of course, in this case many of the staves overlap closely in pitch calibration, or, indeed, are identical. Each staff is identified with the name of the appropriate instrument at the left of the page.

It remains to add that this score arrangement is not the only possible or desirable one for polyphonic music. When many performers are involved it often makes sense to arrange things so that each performer has primary access to his own part. Figure 3.12 shows in diagrammatic form the three major ways in which a polyphonic composition can be arranged in space. In this example three parts are shown, but the same principles can apply to any number. The full, dashed, and dotted lines represent the staves for the three

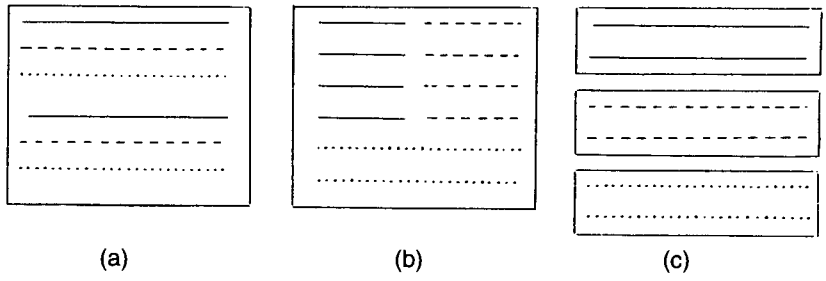



**Figure 3.12** Three arrangements for polyphonic scores.

separate parts or instruments; (a) is the score arrangement just described, and would be typified by any organ score (separate staves for right hand, left hand and feet); (b) is an arrangement whereby each part is confined to a circumscribed area on the page. Although the other parts would be visible to a performer they would not get in his way. This arrangement is typical for one piano, three players music and for much medieval vocal music where each part was to be taken by one singer. In (c) each part is physically separate and can be distributed around the room among the various performers. This is the usual format for modern orchestral players where their own part may be one of as many as fifty different parts going on simultaneously. Of these systems, (a) is undoubtedly the oldest. It allows the reader to keep a direct check on the synchrony of the parts. Examples (b) and (c) make visual checks progressively more difficult. The survival of (c) depends upon the development of cues to synchrony. One, outside the scope of this article, is the conductor, who provides an external reference for a group of performers through his gestures. Probably more crucial was the development of methods for notating cues to synchrony and timing within individual parts. It is to this second major dimension that we now turn.

## 3.4 Bars, timing and rhythm

We have seen how orthochronic notation has commandeered the vertical dimension for notating pitch. It is only a first approximation to say that the horizontal dimension represents time. Temporal aspects are certainly represented, but not in any strict analog sense. Distance from left to right is not a direct measure of time elapsed. To understand how time notation works it is necessary to return to its beginnings in history.

As we have seen, most of the early notations were designed for singing liturgical chant. Thus there were two parallel notations—one giving the

words in an orthography essentially that of today—the other, usually above the words, giving musical directions. In addition, the early music was all homophonic. No matter how many people sang, they all sang the same melody. It would not be unfair to say that up to about the thirteenth century the only timing information explicitly supplied by the musical notation was that of order; the order of the successive words and the pitches to which they were to be sung. If there were differences in the duration of the notes then these were possibly suggested by the flow of the words, or were part of an oral tradition of timing which had grown up in particular communities. Thus, while pitch notation continued to develop towards the modern staff system, the temporal aspect remained undeveloped until the thirteenth century. At that time notational distinctions between long and short notes began to be made.

For some reason, the strategy of using space to indicate time was not adopted. A compelling explanation for this is that scribes felt themselves constrained by the parallel vocal text. From the appearance of Figure 3.13 one could guess that the scribe had written out the words before embarking on the music. The spacing of the letters and the words is very regular; that of the notes is quite irregular. Clearly, the main consideration was to ensure that the right notes fell over the right syllables. This is what determined the spacing of notes, and so timing had to be indicated in some other way. In general, this was done by altering the shapes or appendages of the notes. In this example, from a thirteenth century French manuscript there are three types of note: squares with stems, squares, and diamonds, denoting long, short and shorter notes respectively. It is along this path that rhythmic notation continued to develop despite the fact that much notated music now came to be instrumental, and so without constraint of words.

The history of the subsequent development of rhythmic notation is of some complexity, but certain principles emerged which have remained with us to the present day. One was that note durations were conceived as simple multiples of each other. The other was that simple combinations of durations would often recur several times thus setting up a rhythmic grouping. A simple example of this, as apposite to the twentieth as to the thirteenth century, is given in orthochronic time notation in Figure 3.14(a). In this example the white (unfilled) notes have twice the duration of the black (filled) notes. If the duration of the black note is taken as one unit, then we can see a repeating pattern every three units. This is expressed today by saying that a triple metre is being used. When the rhythm is more elaborate, it is not so easy to determine what the repeating grouping should be just by examination of the note values. Sometimes a sequence

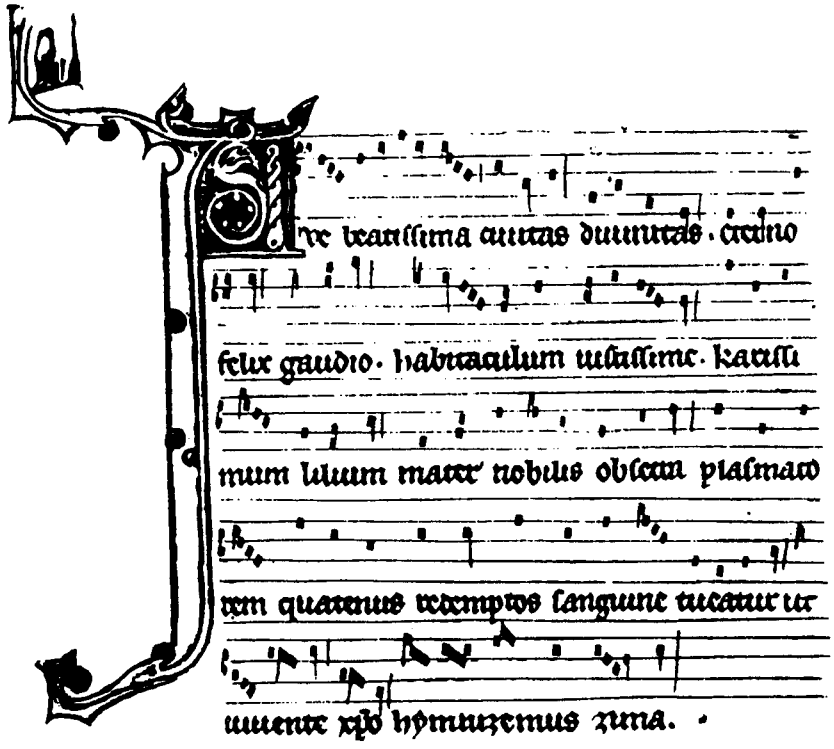


**Figure 3.13** 'Ave beatissima' from the Codex Montpellier (thirteenth century). Montpellier: Bibliotheque de l'Ecole de Medicine. MS. H 196, p. 93. Reproduced with the Trustees' permission. Source: Apel (1953), p. 291.

can be grouped in more than one way. In these cases, some notational means of specifying the grouping directly is required. It is required because rhythmic groupings have important performing implications. For instance, the first note in each rhythmic group is typically accented in some way (e.g. played louder). This is particularly important for communicating the rhythmic structure to a listener. A direct and visually compelling

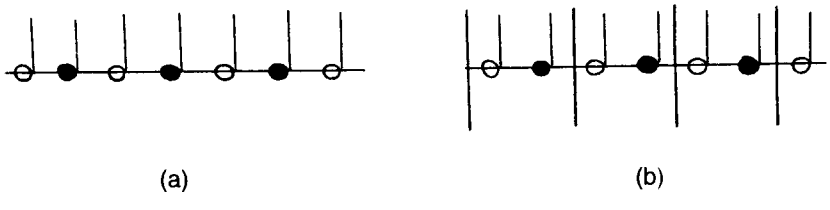



**Figure 3.14** Rhythmic notation.

cue to rhythm is particularly useful when the musician is reading unfamiliar music 'at sight'. Under these conditions he will not have time to work out the rhythmic structure from the note sequence itself. In more leisured study circumstances, or where the music is already partly familiar, direct cues are not so crucial. In modern times sight-reading is required of musicians in many circumstances. The same was not true 700 years ago. The possibility of effective sight-reading has rested upon the emergence of two notational devices.

The first of these is an indication, at the beginning of a sequence, of the number and type of notes in each metrical grouping. This is achieved by two numbers (see Figure 3.1). The upper number tells the reader whether the recurring grouping is 2, 3, 4 or whatever number of units. The bottom number specifies the unit. These two numbers together comprise the time signature of a piece, and in Figure 3.1 this signature indicates a triple meter, with the filled note (known as a quarter-note in America) as the unit.

The time signature, however, leaves a very important question unanswered, for it does not tell the reader which note is the first, accented, note in each group. One cannot simply assume that the first note printed is given the accent. Many tunes begin on an unaccented note. Furthermore, it would be quite difficult for a reader to keep track of the meter through a long piece without dropping a note. A recurrent cue is required. This is supplied, in orthochronic notation, by bar lines. These are vertical lines, running the extent of a staff, which precede the first note in each metrical unit. A bar is the space between two adjacent bar lines, and so each bar contains a complete metrical unit with its first, accented, note at its leftmost edge. The correct barring for Figure 3.14(a) is shown in Figure 3.14(b).

Regular metrical barring did not become widespread until the sixteenth century, but when it did if offered a number of cues to synchronization. The first is the use of bars as a tally in part performance (Figure 3.12(c)). If a performer knows the meter and the speed he can keep track of the music by counting the bars. In symphony orchestras a primary responsibility of a conductor is to maintain the beat, moving his arms in such a way as to clearly indicate the timing of the first note in each bar (the down-beat) and the regular progression of metrical units during each bar. Performers may then tally each downbeat with a bar on their score. In addition, many parts actually number the bars so that in rehearsal performers can agree to 'start from bar 50'.

The second cue to synchronization is available only in parallel score arrangements (Figure 3.4 and 3.12(a)). This is achieved by vertically aligning bar lines in the separate parts. In many instances the bar line will be ruled right down the whole system of parallel staves to make more salient the

points of synchrony. Figure 3.15 shows an early example from Italy (Naples, 1603), transcribed using modern conventions in Figure 3.16. Two features are of note. One is that when some parts contain many short notes the bars must be widened to accommodate them legibly. Thus bars are not of regular size. The trend in modern orthochronic notation has been to attempt to minimize difference in size between successive bars, although it is not possible to eliminate it entirely. In some music, a bar may contain so many notes that it requires the whole width of a page. To maintain this width for bars containing only one or two notes has little psychological advantage, and severe economic disadvantage. The other feature is that the distance between adjacent notes in parallel parts cannot be kept constant if the bars are to be aligned. The clearest example of this in Figure 3.15 comes in the third bar, where the top two parts have many closely spaced notes, whilst the bottom two parts contain only few notes. In this example, and in general up to at least the eighteenth century, no consistent method for spacing out the notes within a bar was adopted. Thus, although bar-lines provided conceptual points of synchrony for a reader, one could not drop a plumb-line at any other point within a bar and expect any synchrony between the notes encountered in the various parts. The time dimension from left to right was stretched and contracted quite arbitrarily within each bar. It is worth noting that in some early scores the bar lines were not even straight. What appears to have happened is that bar-lines became so useful to musicians that scribes often went back over old un-barred scores to put the bar-lines in. In these older scores there was often no attempt whatsoever to maintain vertical alignment of parts, and scribes were apparently too hard-pressed to write out the music afresh. So one finds examples like the splendid British keyboard

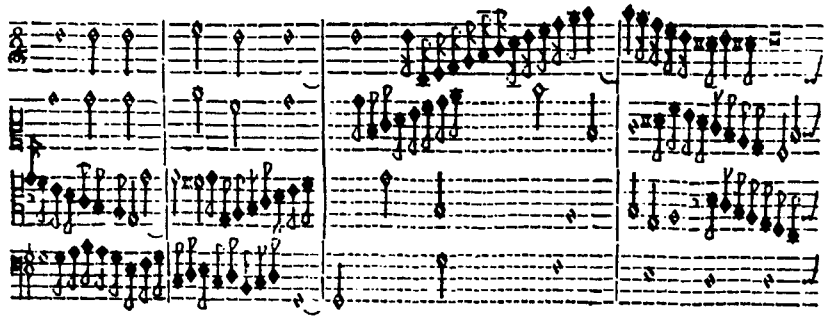


**Figure 3.15** From A. Mayone *Primo libro di diversi capricci per sonare* Naples, 1603, p. 70. Source: Apel (1953), p. 17.

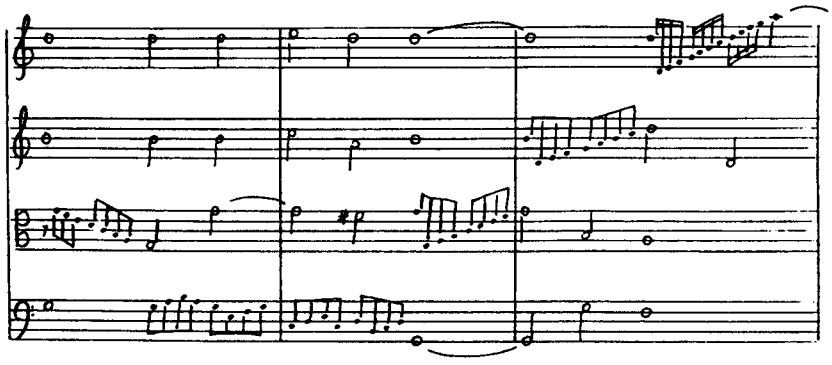


**Figure 3.16** Transcription of first three bars of Figure 3.15 using the conventions of modern orthochronic notation.

score (circa 1540) in Figure 3.17, where the clumsy barring indicates the extent of the visual splits which early keyboard players were required to perform. Notational clarity apart, a modern keyboard performer would find it impossible to sight-read such a score. I suspect that a sixteenth century player would have found it equally difficult, and it suggests that such a score would have had rather different purposes—perhaps archival, perhaps as an aid to a performer attempting to memorize a piece, or perhaps as a starting point for an improvisation.

It was not really until the nineteenth century that what we now see as a logical corollary of bar-lines, namely proportional spacing of notes, came to be widely accepted. It is now standard notational practice to make systematic use of the space within the bar. Two main principles summarize this practice:

Firstly, notes which begin simultaneously must be vertically aligned. This has the useful consequence for, say, a pianist, that all the notes he

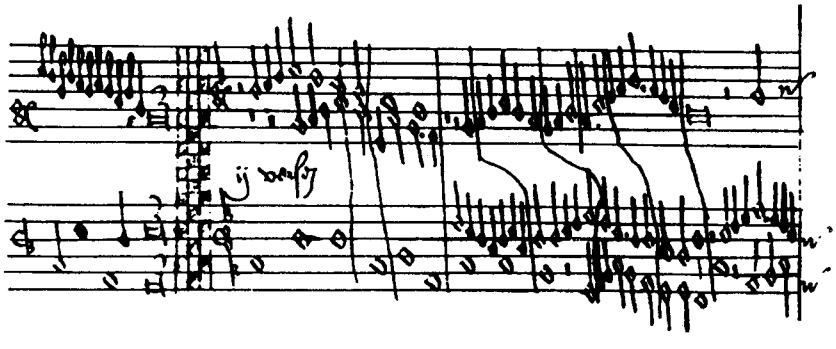


**Figure 3.17** From a keyboard manuscript. London, 1540. British Museum, Ms. Add. 29996, p. 117. Reproduced with the Trustees' permission. Source: Apel (1953), p. 11.

must play simultaneously can be seen from a single vertical slice of score. This supersedes an earlier, intermediate, principle of spacing whereby a note occupied the centre of the time-space allotted to it, rather than the left-hand boundary of that space. This meant that a keyboard player had to scan horizontally as well as vertically to find the notes which had to be played together. Figure 3.18 shows a small extract from a score of a Handel organ concerto; (a) is an early printed score (c. 1800?); (b) is a modern reprinting. The clearest example of the change can be seen at the beginning of the second bar. In (b) the three notes to be played simultaneously are precisely aligned; in (a) there is considerable divergence, particularly for the top note, which, because it lasts for the whole bar, is placed almost in the middle of the bar.

Secondly, the space between a note and the following note must be proportional to the time between their respective onsets. This has always been interpreted rather loosely. It is not necessary that the space following a particular note need be twice that of a note with half its duration, only that the space be discernibly larger to the reader. This principle, in fact, follows as a necessary consequence of the first principle applied to parallel staves.

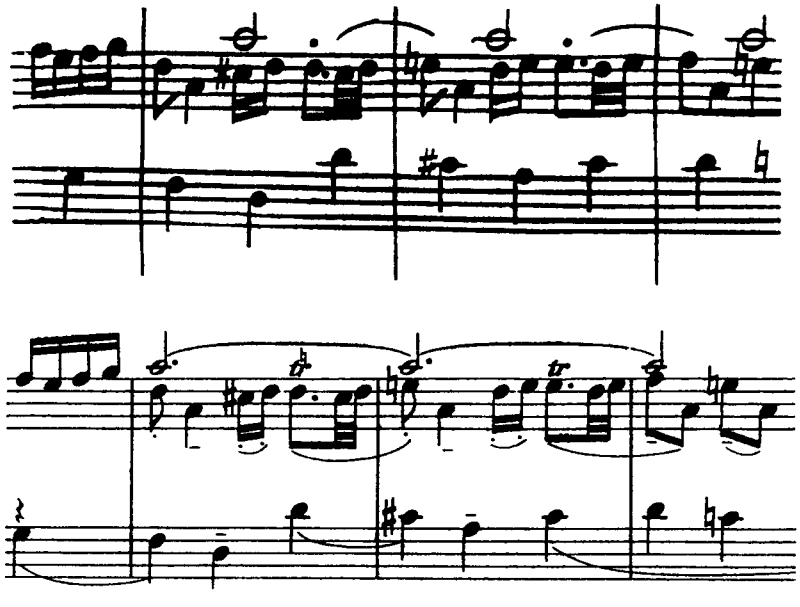


**Figure 3.18** (a) From G. F. Handel, Organ concerto Op. 4 No. 3. Publisher unknown, c. 1800. Source: Cole (1974), p. 58. (b) The same music as (a). Ed. K. Matthaei. Kassel: Barenreiter, 1956, p. 59. Reproduced with the publisher's permission.

It is, however, used nowadays, even in cases where the constraints of alignment do not strictly require it. Thus, as in the pitch dimension, analog spatial information about timing is available to the reader as a formally redundant cue. Again, what psychological evidence exists is consistent with the idea that, at least in some circumstances, readers make use of proportional spacing of notes in organizing their perception of the score (Sloboda, 1977). To illustrate the principles of proportional spacing in an extended orthochronic example, I offer as Figure 3.16 my own transcription of the example in Figure 3.15 (first three bars).

## 3.5 How well does music notation use space?

The preceding sections outlined the principal spatial characteristics of orthochronic notation and its forbears. In this final section I wish to turn to questions of greater generality: to look at music notation not from the point of view of musical history, but from that of the formal study of symbolic systems. This shift in perspective allows us to focus less on questions of development and more on questions of information content, symbolic structure and psychological effectiveness. Any musical notation system must code certain types of information about the music. It uses particular symbols in particular spatial arrangements to represent this information, and has a greater or lesser effectiveness in conveying this information to the reader in a manner consistent with his requirements. These issues can be examined without reference to the historical context of particular notational systems. What follows is an outline of the way in which these issues may be expanded and articulated to form a conceptual framework within which systematic scientific evaluation of alternative notations could be carried out. It is necessary, I believe, to say something about these three issues separately, although the overall goodness of a system will be the result of a combination of factors from the three levels.

(a) *Information content.* The information conveyed by different musical systems can vary in two major respects: how specific the information is, and what type of information is conveyed. The notion of specificity is best conveyed by two extreme examples. Figure 3.19 is the score of a contemporary composition in which the performer is asked to contemplate what he sees and play anything that the pattern suggests to him. No detailed rules are supplied for interpreting the various elements. Thus there is no sense in which particular elements of the display correspond to particular sounds or features of the music. At the other extreme are notations which almost totally specify each sound. Some modern electronic compositions achieve such specificity by giving the precise details of all the oscilators,

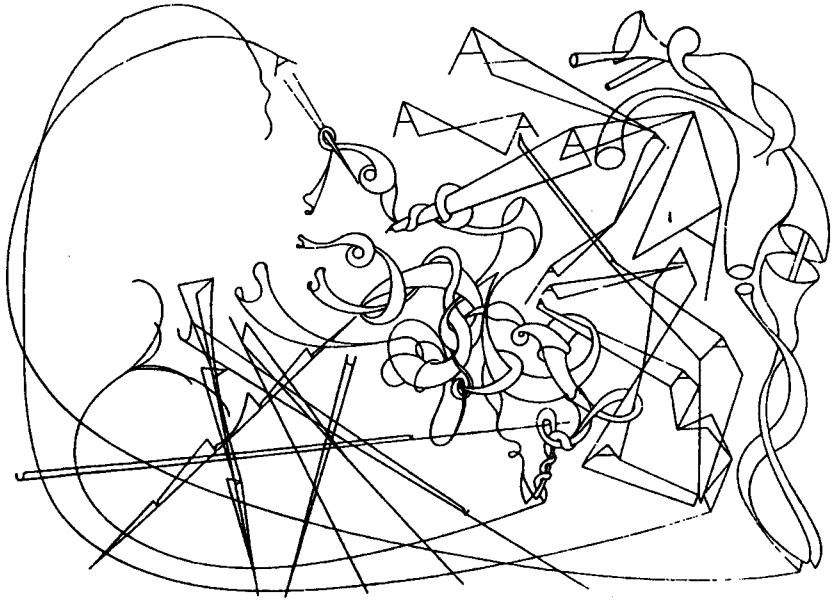


**Figure 3.19** Sheet 3 of A. Logothetis, *Cycloide*. Munich; Edition Modern, MusikVerlag Hans Werwerka, 19. Reproduced with the publisher's permission. Source: Karkoschka (1972), p. 129.

filters, etc. Orthochronic notation is specific in that it provides information about each note in a piece of music. Of course, even in such a case, specificity can vary according to how much information about each note is provided. Sometimes only one or two dimensions of the sound (e.g. pitch and duration) are notated. Sometimes many more dimensions of the sound are specified. The need for specificity poses certain formal problems for a notation system. One is how to represent the order of events. Another is how to correlate the elements in the various dimensions which refer to the same note. How do we show, for instance, that a given note has a pitch $x$, a duration $y$, and an intensity $z$?

At a given level of specificity there can be further variation in the *type* of information represented. One distinction of type was mentioned in Section 3.2, that between systems which notate actions (e.g. the tablatures) and those which notate aspects of the sound. If any consideration were to limit the number of dimensions which could be notated, then the different types of information would have to be traded of against one another. Increased specificity in one area would imply decreased specificity in another.

(b) *Symbolic structure.* By structure is meant both the nature of the symbols chosen and their spatial arrangement. Design of the structure becomes a formal problem when some degree of specificity is required.

There are a whole set of solutions to this problem which we could characterize as essentially non-spatial. As an example, we could assign each event a number (1 being the first event, 2 being the second, and so on) and then make certain statements about the various characteristics of each event (e.g. '2 has a pitch of 565 Hz', '7 has an intensity of 65 dB', etc.) These statements could be arranged on the page in an arbitrary order. One's immediate feeling that such a system would be pointless (i.e. ineffective) does not invalidate it as a perfectly consistent and correct way of symbolizing the specified dimensions of the music. Nonetheless, a notation system could utilize the space provided by the writing surface in a more systematic fashion. The most primitive use of space is to order the elements in some way. With music, the ordering is nearly always temporal. Information about the first note is spatially adjacent to information about the second note, and so on. This is, presumably, because temporal order is the dimension which *makes* music. In the example just given, one could achieve this characteristic by grouping all the statements about event 1 together at the top of a page, with all the statements about event 2 following immediately below, and so on. This could be structurally described as a list. The next advance on a list is a matrix. Here, the two spatial dimensions of a page are utilized systematically. Thus, for instance, the horizontal dimension might represent the temporal order of events, with the various characteristics of each event separated out vertically down the page. So, for example, reading across one row would give the pitches of successive notes, reading across another row would give their intensities, and reading down a column would give all the features of a single event. (I owe much of my conceptualization of structure and some of the terminology to Twyman (1979).)

A different use of the two dimensions might be called 'co-ordinate'. Here, both horizontal and vertical dimensions represent ordinal axes on which some 2-dimensional function may be plotted. Orthochronic notation can be seen, in part, as a graph of the pitch ($y$)–time ($x$) function, each note plotted as a point on the graph. Matrix and co-ordinate notations can be made very powerful by the use of two other supplementary devices, partition and clustering. In partitioning, one divides the page up into discrete areas where space may serve a different function. Thus, for instance, in Figure 3.15 the pitch dimension is not represented continuously in the vertical dimension. It only operates within each staff, recalibrated for each. The space between the staves has no pitch implications at all, and so other, pitchless information may be placed here. In some modern scores,

for instance, a rising and falling line between staves is given to indicate increases and decreases in speed. A most imaginative use of partitioning is shown in Figure 3.20. This is the second page of Busotti's *Siciliano*. Here, vertical space within staves is interpreted conventionally, but the slant of a staff indicates the degree of acceleration or deceleration of speed. Dotted vertical lines indicate points of synchrony.

Clustering refers to the practice of representing several dimensions of an event, not by separate symbols, but by different aspects of a composite symbol. Thus, a note's position could indicate its pitch, whilst its shape indicated its duration, and its color indicated its intensity. This is a useful economy of space, which allows one to compact several rows or columns of a matrix into one.

These strategies, in their detailed operation, would seem to account for most of what one could say about the possible structures of music notation

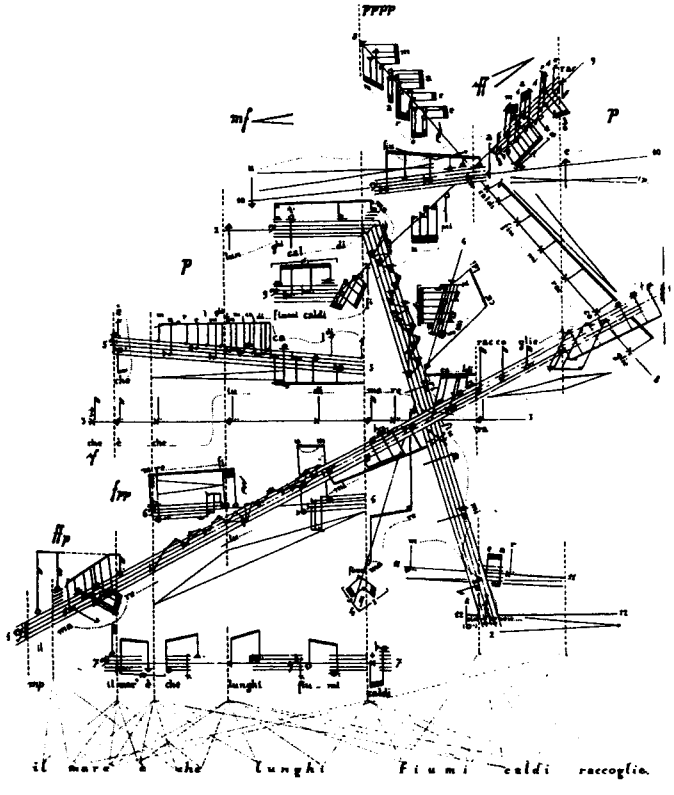



**Figure 3.20** Page 2 of S. Bussotti, *Siciliano*. Florence: Aldo Bruzzichelli, 1962. Reproduced with the publisher's permission. Source: Karkoschka (1972), p. 94.

systems. When a combination of all the strategies is available there would seem to be very few formal limits on the number of specific dimensions which could be represented by a symbolic system. If questions of how well space is used have significance at this level, it is with respect to the economy of space. A system which allows clear representation of twenty pieces of information has a claim to greater efficiency than one which allows representation of only ten pieces of information within the same area of space. But for a criterion of what constitutes a clear representation, we must turn to the final, psychological level of analysis.

(c) *Psychological effectiveness.* Music notations are for use by readers. Their effectiveness is the extent to which readers are able to retrieve the information about the music which they supply. Effectiveness is not simply a function of the symbolic structure of the score, nor of the information it contains, but of the conditions under which a reader is expected to retrieve the information. These conditions may include general factors of human psychological functioning; they certainly include factors specific to the particular reading situation. All I am able to do here is to enumerate a representative sample of issues pertaining to psychological effectiveness. Although these issues are all empirically determinable, almost none have received rigorous empirical treatment.

A fundamental issue I have already referred to is the extent to which a reader is expected to provide a coherent performance of the music at first sight. Clearly, the more the reader must do without prior knowledge, the more constraints are placed upon what is an acceptable notation. Some of these constraints are straightforward matters of legibility. Even within a single notational system clarity and consistency can vary greatly according to the care with which a score is laid out and the nature of the reproduction process. It is distressing to discover that many of the parts which players in major symphony orchestras must use (at least in Britain) are shockingly presented and reproduced. Figure 3.21 shows a not untypical example of the kind of thing an orchestral player must put up with. It has not been determined how much this upsets performance. Certainly the players themselves complain bitterly. A major difficulty is spatial uncertainty. This can occur when staff lines are badly printed, or too close together. In this case it is sometimes difficult to see which line a note is centered on. A similar problem occurs when noteheads are too large with respect to the staff lines. A different type of spatial uncertainty is caused when ancillary marks are not properly aligned with the notes, or when they are too far from the staff to allow easy estimation of their alignment.

These considerations lead to the wider issue of discriminability. When a notation system uses a large number of different symbols, two different

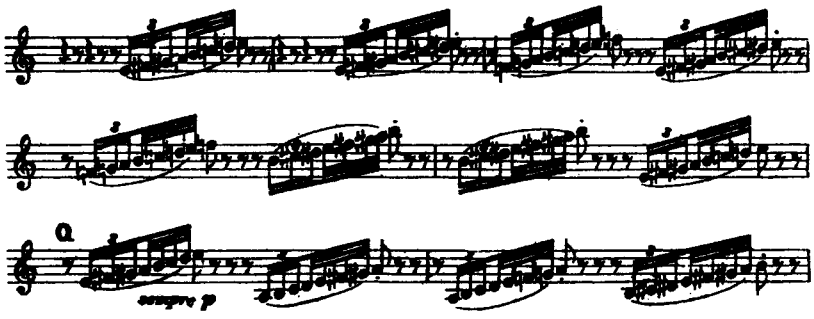


**Figure 3.21** Extract from an orchestral part used by a British symphony orchestra, date and publisher unknown.

symbols can often look very much alike. If a reader is expected to discriminate the symbols at speed, then he will arguably be helped by symbols which are as different as possible in appearance. On the other hand, if he is to perceive relations between different notes then he will also arguably be helped by symbols which preserve that relationship (e.g. symbols for notes close in pitch will be more like one another than symbols for notes not similar in pitch). The analog representations of pitch and time in orthochronic notation are based directly on this latter principle. Here, closeness in space maps directly on to closeness in pitch and time. When space is the notational dimension then it is possible to achieve increased discriminability simply by expanding the axes (i.e. make the staves larger and the notes farther apart). But to do this raises problems in a different area, that of the visual span and eye-movements.

If a reader is to use the score to organize immediate performance then he must take in all the information required for a particular action with as few fixations as possible. An average rate of performance for a moderately demanding piano piece might be ten notes per second (some of them executed simultaneously). It would be difficult to achieve more than four or five separate fixations in this time, and so the information required must necessarily be contained in a small area. Expanding the dimensions of the symbols would increase the number of fixations required to read the same number of symbols. Thus an effective notational system also has to be compact, and this poses problems when more than one dimension of the same event must be notated. The conventional matrix would be unhelpful because different aspects of one note could be spread right down a page. The device which arguably saves orthochronic notation is the clustering described in the previous sub-section. By using different visual aspects of the same symbol (position, shape) to indicate different sound aspects of

the same note (pitch, duration) one can increase the visual density of the information at little expense to clarity. In general, a system such as the staff system would seem to be able to support about five different dimensions of sound without difficulty. For instance, the two spatial dimensions could represent pitch (vertical) and order (horizontal). Shape of note is a third dimension which could be used to represent duration. Ancillary symbols above and below the note could represent intensity and phrasing. There is, however, clearly some limit to the number of dimensions which could be usefully added in this way. This shows itself in lute notation (Figure 3.2) where note shape is used to designate which fret to touch, at the expense of a means of signifying duration. Klavarscribo fits in an extra dimension only because the arrangement of keys on a piano is isomorphic with the pitch dimension. Thus, the same horizontal information specifies both pitch and which key should be pressed (Figure 3.5).

A final requirement of a system that is to be read at sight is universality. One can only become proficient at reading any system if one encounters a large body of music written in that notation. Many contemporary notation systems are ruled out on this count because they are designed by the composer with a particular composition in mind, and are used only for this composition. Thus, Brown's fifty-line staff (Figure 3.10) has never been used before or since his one composition. Many contemporary notations are designed with the expectation that a performer will devote considerable time to a study of the particular score and its notation before a performance is possible. It would be inconceivable to arrange a performance of Busotti's *Siciliano* (Figure 20) at sight. Neither he nor his performers would wish for that. Similarly no notational system devised prior to about 1500 would have been read at sight. Musical culture just did not require it.

Today, musicians accustomed to reading orthochronic notation at sight become very sensitive to slight changes in notational practice. One simple spatial example concerns the positioning of note-stems (the vertical lines attached to notes). The modern convention is that descending stems are attached to the left edge of a note-head, ascending stems to the right. Some earlier scores reversed this convention. Informationally and structurally this has absolutely no consequences at all, but it is psychologically disruptive. Such scores are more difficult to read at sight, and the subjective impression is of something quite wrong about them.

Even when reading at sight is not required, a score's primary function is to assist a musician in performance. In such cases the score may be used to elicit learned response patterns in the correct sequence. This means that the performer must be able to keep place in the score, even if he does not use all the information in it. When detailed prompting is required, he must be

able to find the appropriate information rapidly and effectively. Thus, many of the psychological considerations pertinent to reading at sight remain important in less extreme reading situations. Compactness and discriminability are still necessary. Consistency in positioning of information is also important. A reader must be able to know exactly where to go for particular types of information, even if he does not always need them. For instance, in a conventional orthochronic score, information about intensity is always to be found below the staff.

Two concluding remarks are in order. First, I hope to have shown that spatial factors in music notation are of some complexity and importance, and that there are several frameworks within which empirically determinable questions about efficiency of alternative arrangements may be asked. Second, I need to emphasize that nearly all my remarks about the psychological aspects of music reading are based on personal experience and accounts of other musicians rather than on rigorous investigation. Although it is to be hoped that serious empirical work will be undertaken. I should perhaps finish by firing one warning shot across the bows. Efficiency can only be meaningfully estimated where a reader is as familiar with the system in question as he is with the accustomed system with which the experimental system is being compared. Failure to find an immediate effect of some notational change does not imply that there will be no effect after the months or years of familiarization that has been received by the accustomed system. Any evaluation of notational reform is going to be a long-term process, and one which will require a considerable commitment from musicians who are required to learn and operate with new systems. Any short-term experimentation is almost bound to lead to the erroneous conclusion that orthochronic notation as it now stands is the best of all possible systems.

## References

Abraham, G. *The Concise Oxford History of Music.* London: Oxford University Press, 1979.

Apel, W. *The Notation of Polyphonic Music 900–1600.* Cambridge, Massachusetts: The Mediaeval Academy of America (fourth Edition), 1953.

Cole, H. *Sounds and Signs: Aspects of Musical Notation.* London: Oxford University Press, 1974.

Hyatt King, A. *Four Hundred Years of Music Printing.* London: Trustees of the British Museum, 1964.

Karkoschka, E. *Notation in New Music.* London: Universal Edition, 1972.

Read, G. *Music Notation.* London: Victor Gollancz, 1974.

Sloboda, J. A. Phrase units as determinants of visual processing in music reading. *British Journal of Psychology*, 1977, **68**, 117–124.

Sloboda, J. A. The psychology of music reading. *Psychology of Music*, 1978a, **6.2**, 3–20.

Sloboda, J. A. Perception of contour in music reading. *Perception*, 1978b, **7**, 323–331.

Sloboda, J. A. Music performance. In D. Deutsch (Ed.) *Psychology of Music*. New York: Academic Press, 1982.

Twyman, M. A schema for the study of graphic language. In P. A. Kolers, M. E. Wroistad, H. Bouma (Eds.) *Processing of Visible Language 1*. New York: Plenum Press, 1979.

CHAPTER 4

# IMMEDIATE RECALL OF MELODIES

## 4.1 Introduction

The primary purpose of this chapter is to present musical transcriptions of the attempts of eight adult subjects to recall part of a folk melody that was repeatedly presented to them. We also present the results of some analyses of these transcripts, which seem to point particularly clearly to the involvement of structural knowledge in musical memory.

### 4.1.1 Problems of transcription

The study of free recall in music has not commended itself to many researchers. This is understandable: For one, there is no generally accepted criterion for what counts as a reliable transcription. The conventional notation of Western musical culture forces the transcriber to translate what may be a fluctuating and richly variable signal into a limited set of pitch and time categories prescribed by the notation. Much information is inevitably lost in the process. This is, of course, true of linguistic transcription, too, but the problems are less acute because it is often possible to assume that the speaker intends to address the categories the notation embodies (e.g. he or she intends to utter words contained within a shared natural language). In music, it is less clear to what extent a singer *intends* a performance to be interpreted within the categories of tonal music. The problem arises largely because the intentions of a musical performer are not bound by *practical* constraints in the way that they usually are in language. Linguistic communication is embedded in an extralinguistic context, and it is often easy enough to discover when a speaker's intentions have been misinterpreted. Music is not a communication system in the way that language is. It is relatively self-contained and has no major consequences for domains of cognition beyond music.

How, then, can we know when a singer intends his or her performance to be interpreted by a listener within the categories supplied by conventional notation?

We think there are at least three factors that should be taken into consideration:

1. If a Western listener *hears* the bulk of a performance as fitting the pitch/time categories of tonal music, then it is likely that the performer intends it to be heard this way. In other words, if a performance *makes sense* tonally, then it is probable, though not completely certain, that such sense was intended by the performer, especially if the performance is fairly long. There are so many degrees of freedom for a musical performance that it would be coincidence indeed if a sequence generated non-tonally had consistent and strong tonal implications. Successful transcription is in itself some indication that an intention exists to address the categories embodied in the notation.

2. There is evidence that children immersed in a musical culture (such as our own) internalize the structures that are implicit in the bulk of music they hear (e.g. Gardner, 1981; Zenatti, 1969). We therefore, feel that most adults, conceive of music in terms of these enculturated structures. This is as true of performance as it is of perception.

3. Given positive indications on the preceding two counts, a performance is most likely to address intentionally the normal categories of the culture when its context encourages this. One such context involves the presentation of an unambiguously well-formed musical sequence for recall. We may argue that if a singer *hears* the music as tonal, then he or she will attempt to reproduce it as tonal.

These criteria raise particular problems for the transcription of the performances of children or of people from cultures whose musical structures are different from those of the transcriber. It is especially difficult to interpret transcriptions of songs by very young children (e.g. Dowling, 1982; Moog, 1976) or of songs from specialized or isolated subcultures (e.g. the Serbo-Croatian folk songs transcribed by Bartok and Lord, 1951). We are not concerned with resolving these difficulties here. Rather, we argue that our own data do not fall foul of these difficulties and are, therefore, optimal for transcription into conventional notation.

The foregoing does not entirely dispose of the problems associated with making transcriptions. One very important and well-documented fact about performance is that, even when the intention is unambiguously directed towards the categories of pitch and time enshrined in our notation, performance deviates significantly and systematically from the notated ideal. Conventional notation is not up to the task of recording the expressive

variation that is present in any musical performance. Furthermore, such variations are not reliably diagnosed as variations by listeners. This means that whereas variations may contribute to the sense of musical life or articulation and may even help a listener to assign the intended tonal or metrical structure (Sloboda, 1985), they are unlikely to be capable of reliable transcription by ear, even if a notation were available to represent them. They are reliably captured only by mechanical transcription that records the exact timing and pitch parameters of successive notes (e.g. Seashore, 1938; Gabrielsson, 1974; Shaffer, 1981).

If the primary concern of an empirical study is with such expressive variation, then, clearly, conventional transcription is of little use. If, however, the primary concern is with the basic pitch and timing structures on which expressive variations operate, then there is some merit in a transcription process that necessarily discards much of these variations. Only when a variation is so large as to cause some uncertainty about which category a note should be assigned to (e.g. crotchet or quaver, G or G#) does it need to find its way into a transcription. If the performance is, indeed, one based on the structures inherent in our culture, then the proportion of variations causing uncertainty should not be large. If it is, we can hardly say that the performance strikes the ear as unambiguous. Such a performance does not fulfil the first of our criteria for transcription. In our own data, subjects produced a total of 1069 notes. We judged 62 (5.8%) of these to be ambiguous in the time domain and 33 (3.1%) in the pitch domain. Of this small proportion, many timing deviations were apparently due to the taking of a breath that distorted an established metre; and several pitch deviations were associated with large pitch leaps, suggesting response programming problems rather than intrinsic deviations from tonality as an explanatory cause.

We are satisfied that our transcriptions do not misrepresent the performance in any way that is germane to the issues under discussion. We do not rule out, however, the possibility that we—and other potential transcribers—are subject to systematic perceptual distortions. For instance, the effects of categorical pitch and time perception will almost inevitably be to render transcriptions more coherent than they might be. One merit of transcription in conventional notation is, of course, that it is immediately comprehensible to literate musicians. More literal transcriptions, in terms of precise performance parameters, are invaluable for some purposes, but they must be interpreted into something approaching conventional notation before they can be understood as music. Formal procedures for translating between numerical performance parameters and conventional notation (e.g. Longuet-Higgins, 1976; Steedman, 1977)

have not yet solved the full range of problems presented by real auditory signals. We believe that, for the type of research reported here, there is, at present, no viable alternative to aural transcription by musically trained listeners.

### 4.1.2 The analysis of recall data

A different reason for the paucity of empirical work on musical recall is the lack of agreed upon and well-motivated methods of describing and analysing the content of a performance in relationship to an original model. It is clear that crude measures such as number of notes correct tell us little of interest about the nature of the recalls.

The recent study of recall processes in language provides some pointers to the type of analysis that might be appropriate (Bower, 1976; Kintsch, 1977). We know that verbal recall is rarely word-for-word correct, but it matches the original at a higher level of meaning and structure. We find substitution and inference, selective loss of information, and other distortions, but preservation of essential meaning and structure. We do not wish to be constrained by the details of recall analysis methods in language, but we are concerned to explore methods of musical analysis that provide information at an analogous level of abstraction. This makes certain conditions necessary. First, the music to be remembered must be long enough to contain within it a variety of the types of structural progressions and relationship that characterize the typical music of the culture; we have chosen the folk song or popular song melody as such a typical structure. Second, the music to be remembered must be short enough to allow for the possibility that subjects can encode it completely at some level of abstraction within a single hearing; pilot observations suggest that a melody containing about 30 notes fulfils these conditions. It is worth pointing out that most contemporary research on musical memory has used some form of recognition procedure (e.g. asking subjects to judge whether two melodies are the same or different) and has used sequences containing much fewer than 30 notes.

### 4.1.3 Recall versus recognition

Recall and recognition testing both have a long and well-established history within the study of memory. Both have strengths and weaknesses, and to a large extent, complement one another. Recognition studies are particularly amenable to factorial hypothesis-testing designs in which a high degree of stimulus control is exercised. By restricting subjects' responses to a few preordained categories, it is possible to obtain precise answers to preformulated questions. Yet, as many, including Newell (1973) and Allport (1975) have

noticed, such research strategy can be counterproductive. In many cases, the binary theoretical choices that this forces one to make are naive and ill-conceived. Successive refinements and elaborations can lead one away from what was important about the original question, so that the research, although technically sound, sheds little light on the most salient aspects of the phenomenon under study.

There are other reasons for being cautious about the overextensive use of the recognition paradigm in music research. In many instances, subjects are required to do things that are not familiar to them and that find no parallel in their normal musical behaviour. Experiments which, for instance, use the probe technique of extracting a small segment from a longer one for recognition are of doubtful ecological validity, as are those requiring recognition of any short, decontextualized musical fragments. The most prevalent type of natural recognition behaviour involves the identification of the name of a piece of music or the feeling that 'I have heard this before,' where 'this' can be as long as a whole movement or as short as a theme in context (Halpern, 1983; Pollard-Goit, 1983).

Recall studies provide more 'messy' data, since subjects' responses are not confined to preordained categories. Although hypothesis testing is still possible. the amount of behaviour left unaccounted for is much larger. This is the price that many researchers are prepared to pay for data that bear on integrated sets of complex processes and that reflect more directly the exercise of preexisting skill. Recall data provide many more opportunities for abrasive contact between theoretical preconceptions and reality. They are also, perhaps, the most appropriate kind of data to gather when theory is relatively unrefined. One is less likely to overlook some vital aspect of the behaviour in question when faced by a rich data base than when restricting oneself to the impoverished behaviour that is the typical outcome of recognition experiments.

A particular difficulty with the recall method as applied to music is that many subjects have no well-practised response mode, have no instrumental skill, and may be totally inept at vocal techniques such as singing, humming, or whistling. Such ineptitude naturally predisposes the researcher toward recognition testing. We wonder, however, whether psychologists of music have sometimes made too pessimistic assumptions about the ability of untrained subjects to retain material and make coherent responses. With a methodological and theoretical predisposition toward recognition studies, such assumptions may not have been as critically examined as they should have been. We have been pleasantly surprised by the quantity, quality, and coherence of the data we have obtained from committed but relatively untrained subjects.

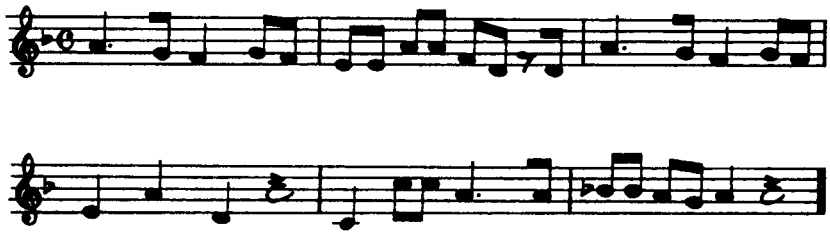
Our investigation was prompted by the wish to provide a data base of musical recalls. We began with no strong hypotheses about what we would find, other than the intuition that high-level structures would be implicated. Rather than present a rational reconstruction that embeds our data in theoretical preconceptions, we proceed roughly historically, describing first our method of data collection, then our results and some of the analyses performed, and finally, theoretical comments. This seems the most appropriate way to proceed in a study that has a strong component of natural history.

### 4.2 Method

Our subjects were eight students drawn from a volunteer pool in the Department of Psychology, City of Liverpool College of Higher Education. All were females between the ages of 19–22. Four of the subjects had not had any special musical training but enjoyed listening to music. These we will call the 'non-musicians'. The other four subjects were musically trained and were active performers of classical music.

We recorded piano performances of several sequences taken from folksong melodies. Each subject was asked to listen to each sequence six times in a row, providing an attempt at sung recall after each hearing (to 'la' or any other chosen syllable). Each recall was recorded for subsequent analysis. After preliminary examination of the recall data, it was decided to discard the results for all sequences other than the first one presented. This was because material from the first set intruded massively into subsequent recalls. Sometimes, the first recall of the second sequence was almost identical to the last recall of the first sequence, even though the presented melodies were quite different. It was surprising for us to obtain this high degree of intersong contamination, and the finding is intriguing in itself. It is also something that only a recall study could have shown. However, we decided to leave it aside in this context as an unnecessary complicating factor.

Accordingly, the data we present are from the six recalls of the first sequence given to the subjects alone. The sequence, given in Example 4.1, was played at a speed of about two crotchets per second. It comprises the first three phrases of the Russian folk song, ~Sailor,'' contained in a collection of Russian songs (O'Toole, 1974). This song was not known to any of our subjects. The sequence contains 29 notes, subdivided into three phrases (11 + 9 + 9). There are multiple cues to this phrase subdivision: (1) the rests after Notes 11 and 20, (2) the fact that each phrase occupies an identical metrical slot of two bars, and (3) the fact that Note 13 begins a repetition of the melodic and rhythmic pattern that starts the sequence (Note 12 is an added upbeat). In terms of traditional analysis, the sequence shows an $A_1A_2B$ form, where $A_1$ and $A_2$ are variants of the same material



**Example 4.1** The stimulus sequence used in the study.

and *B* introduces new material. In fact, the *B* phrase shares features with *A* as well as differs from it. Notes 25–28 repeat the rhythmic pattern of notes 6–9 at an equivalent point in the metrical structure (third half-bar). Harmonically, the *A* phrase implies a I–I–V–I movement in D minor, and the *B* phrase is best interpreted as V–I–V–I in F major. In the full song, *B* is followed by a repetition of *A*, to bring the verse to a close in D minor.

We believe that the melody is highly unambiguous in that most listeners within the Western tonal culture hear the structural groupings we have informally assigned, in strong preference to other possible groupings. We base this belief on our own musical intuitions, but also on the formalizations provided by Lerdahl and Jackendoff (1983), Longuet-Higgins (1976), and Steedman (1977). These formalizations embody heuristics for assigning metrical and harmonic structure to note sequences in a way that appears to match the intuitions of experienced listeners. As we understand them, such heuristics provide unanimous support for our parsing.

The appendix to this chapter provides complete transcriptions of the six recall attempts by each of the eight subjects. Subjects 1–4 are the non-musicians, subjects 5–8 are the musicians. The recalls are numbered in sequence so that, for instance, 3.4 denotes the fourth recall of Subject 3. The transcriptions use conventional notation, but with certain differences and additions:

1. All recalls are transcribed in D minor/F major even when, as in a few cases, subjects transposed their response to a new key (usually a semitone up or down). We may take such transpositions to be unintentional and of little interest in the present context, except to show that subjects are coding in terms of pitch relationships rather than absolute pitches.

2. We have not given metrical interpretations above the level of the crotchet beat. Thus, bar lines are omitted.

3. The symbols × and o denote approximations to notated duration and pitch, respectively. When we were in doubt about which pitch or time

category was nearest to the note sung, we assigned the note to the category that best preserved a consistent metrical and harmonic interpretation. Although this does make the transcriptions look more coherent than arguably they ought to be, the proportion of notes involved is small enough to make us feel that this is not a serious problem.

4. The curved slur mark encloses notes sung within a single breath.
5. The symbol v denotes a slight but discernible pause that bends rather than disrupts an established metre. It is usually associated with the taking of a breath.

## 4.3 Results of data analyses

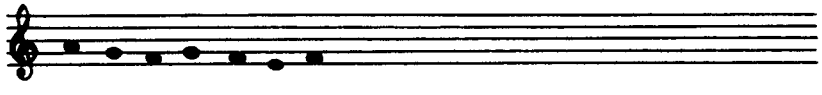
### 4.3.1 Melodic contour analysis

The first analysis is a fairly low-level one. It measures the degree to which the melodic patterns of the original are retained in the recalls. For this purpose, we derived a melodic contour for Example 4.1 that shows the pitch movements within the three phrases without regard for duration or contiguous repetition. This is shown in Example 4.2.

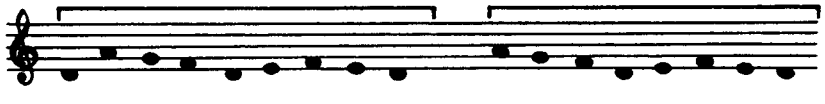


**Example 4.2** The melodic contour of Example 4.1.

We then obtained the melodic contour of each recall by the same method. For instance, the melodic contour of 1.1 is given in Example 4.3, and that of 1.6 in Example 4.4. When the subject sang one phrase or less, we looked across the entire contour in Example 4.2 for the best-fitting phrase. Thus, 1.6 was evaluated against the contour for the second phrase of Example 4.2, on account of the D–A movement at the beginning. When



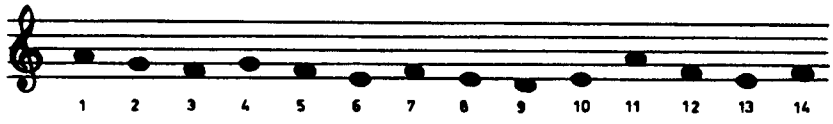
**Example 4.3** The melodic contour of Recall 1.1.



**Example 4.4** The melodic contour of Recall 1.6.

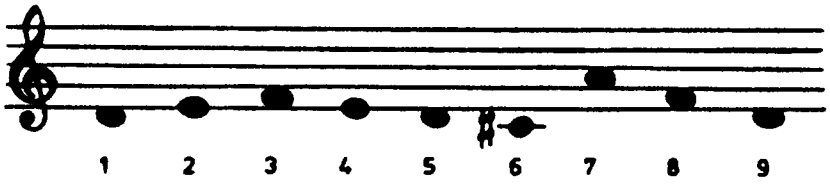
a subject provided two or more phrases, as in most cases, then the first phrase sung was matched against the first phrase of the original, and so on.

Within each phrase, a recall scored 1 for each of its notes that formed part of a melodic pattern (of two or more notes) found in the original. For example, Example 4.3 scores 6, because the first six notes match the first six notes of Example 4.2. Recall 1.2 provides a more difficult example. Its contour is given in Example 4.5. To score this recall, one has, in effect, to slide the original contour along the recall contour. Thus, the first six notes of the original are found at Notes 1–6 in Example 4.5. Notes 7–10 do not match anything in the original, but then Notes 11–12 match Notes 7–8 of the original. Therefore, this recall scores 8 on 14 for this measure.



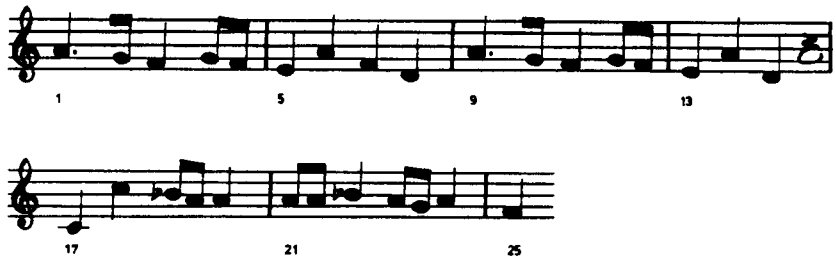

**Example 4.5** The melodic contour of Recall 1.2.

A final example of our scoring procedure is given for Recall 8.4. The melodic contour of its first phrase is given in Example 4.6. In this case, we find no match to the beginning of the original, so we move through the original contour until we find a match. Notes 7–9 of the recall match notes 7–9 of the original, so this phrase scores 3 on 9.



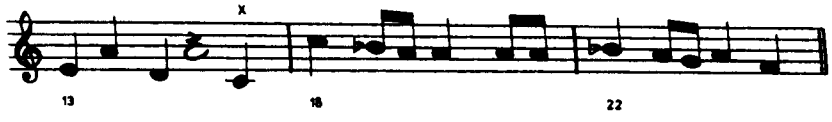

**Example 4.6** The melodic contour of Recall 8.4.

Table 4.1 presents summary data from this analysis. Column 1 gives, for each subject, the melodic contour score summed over the six recalls; Column 2 gives the total number of melody notes in the recall contours; and Column 3 expresses 1 as a percentage of 2. These percentage data show wide individual differences ranging from 0% for Subject 6 to 88% for Subject 5. Is there any particular reason for the exceptionally low scores of Subjects 4 and 6? Our view is that these two cases can be explained, but differently. Let us take Subject 6 first. If we examine a recall such as 6.5, we experience a strong intuition that, despite its low melody score, it is *very* similar to Phrases 1 and 2 of the original. We argue later that its perceptual similarity is based on a shared underlying metrical and harmonic structure; that this subject has constructed a new melody




**Example 4.8** Alternative metrical analysis of Recall 5.5.

matches the metrical feel of this sequence most exactly is given in Example 4.9. This makes Beat 17 into an additional upbeat, enlarging Bar 4 to contain five beats. This beat is, therefore, scored as anomalous, as is Beat 20, which is noticeably longer than the consistent metre would predict. This recall scores 23 on 25.




**Example 4.9** Alternative metrical analysis of Recall 5.5.

Table 4.3 presents summary data from this analysis for each subject. Not only is the total percentage of metrical consistency very high, it is also high for each individual subject. Whatever subjects are doing with the details of the melody, they are, on the whole, ensuring that their recalls use the metre of the original.

**Table 4.3** Metrical analysis

| Subject | Total beats played | Metre score (beats) | Score as percentage of total |
|---|---|---|---|
| 1 | 45 | 43 | 95 |
| 2 | 135 | 108 | 80 |
| 3 | 146 | 137 | 94 |
| 4 | 14/1 | 120 | 86 |
| 5 | 102 | 92 | 90 |
| 6 | 54 | 46 | 85 |
| 7 | 145 | 112 | 77 |
| 8 | 112/889 | 115/773 | 94/88 |

### 4.3.3 Rhythmic analysis

One possible reason for the high scores on the metrical analysis is that subjects have accurate memories for the rhythmic patterns of the original, if they correctly reproduce most of the rhythms, then it automatically follows that they will reproduce the metre. Accordingly, we assessed the degree of rhythmic reproduction by awarding each half-bar of a recall 1 point if it matched the rhythm at the same position in the original. Summary results of this analysis are shown in Table 4.4. Fewer than 50% of rhythms are exact copies of the original rhythmic sequence. It seems, therefore, that metrical consistency cannot be explained in terms of rhythmic reproduction. Rather, subjects are creating new rhythmic combinations within an underlying quadruple metre.

Table 4.5 presents the rhythmic reproduction analysis summed across subjects for each trial. Like the melodic reproduction scores, these scores do not seem to increase across trials. A two-way ANOVA on the percentage data factored by subject and trial shows a significant effect of subject ($F(7, 35) = 3.09$, $p < 0.02$) but not of trial ($F(5, 35 = 2.27$, $p > 0.05$). If subjects do improve across trials, their improvement is not due to more accurate rhythmic recall.

### 4.3.4 Breath analysis

We now turn to some scoring measures that try to take direct account of the phrase structure of the recalls. We first examine the way that subjects group notes by breathing. In this analysis, a score of 1 is given for each audible breath between notes or groups of notes. We partitioned the recalls into two unequally sized segments: (1) the last notes in Phrases 1 and 2, and (2) all other notes. Those in the first group were determined by examining the barring assigned to recalls in the metrical analysis. The last note of Phrase 1 was defined as the last note in Bar 2 whose duration (including

**Table 4.4** Rhythm analysis, by subject

| Subject | Total produced | Score | Percentage |
|---|---|---|---|
| 1 | 27 | 10 | 37 |
| 2 | 62 | 14 | 23 |
| 3 | 66 | 36 | 55 |
| 4 | 66 | 29 | 44 |
| 5 | 47 | 31 | 66 |
| 6 | 24 | 10 | 42 |
| 7 | 64 | 33 | 52 |
| 8 | 56/412 | 25/188 | 45/46 |

**Table 4.5** Rhythm analysis, by trial

| Trial | Total produced | Score | Percentage |
|---|---|---|---|
| 1 | 55 | 21 | 38 |
| 2 | 66 | 31 | 47 |
| 3 | 62 | 26 | 42 |
| 4 | 75 | 40 | 53 |
| 5 | 77 | 35 | 45 |
| 6 | 77 | 35 | 45 |

any subsequent rest) was a crotchet or longer. The last note in Phrase 2 was similarly defined for Bar 4. The only recall where this gave a result at odds with our intuitions was 7.5. Our procedure places the phrase boundary after the crotchet F (Note 10), whereas we would view it as occurring after the D quaver (Note 11). This uncertainty presents no problems in the present analysis since neither note is associated with a breath.

Table 4.6 presents summary data from this analysis, showing the percentage of breaths following notes in Categories (1) and (2) for each subject. Almost all phrase endings were followed by breaths, but less than 10% of other notes were followed by breaths. This is overwhelming evidence that subjects are choosing to breath at phrase ends rather than at other places.

### 4.3.5 Phrase structure analysis

Our next analysis attempts to quantify the degree to which subjects construct their phrases using the patterns of melodic and rhythmic imitation present in the original. That is, do they preserve the $A_1A_2B$ phrase structure? To answer this question, we need a measure of the degree to which Phrases 2 and 3 imitate Phrase I. If subjects are reproducing this aspect of the original,

**Table 4.6** Breath analysis

| Subject | Percentage of breaths at phrase endings | At other places |
|---|---|---|
| 1 | 100 | 0 |
| 2 | 91 | 10 |
| 3 | 92 | 1 |
| 4 | 100 | 7 |
| 5 | 100 | 0 |
| 6 | 100 | 6 |
| 7 | 92 | 10 |
| 8 | 90 | 5 |

**Table 4.7** Percentage match of phrase analysis

| Subject | Phrase 1 to Phrase 2 | Phrase I to Phrase 3 |
|---|---|---|
| 1 | 79 | – |
| 2 | 71 | 70 |
| 3 | 69 | 41 |
| 4 | 73 | 44 |
| 5 | 78 | 22 |
| 6 | 88 | – |
| 7 | 70 | 11 |
| 8 | 58 | 44 |
| Mean | 71 | 41 |
| Original | 63 | 13 |

then we predict that the similarity of Phrase 3 to Phrase I is significantly less than the similarity of Phrase 2 to Phrase I. Accordingly, we compared each beat in one phrase with the corresponding beat in the other phrase, awarding a score of 1 if the pitch(es) in the two beats were identical and a score of 1 if the timing pattern was the same. These scores, represented as a percentage of total possible score, are given in Table 4.7. The missing data occur because Subjects 1 and 6 did not provide a third phrase in any of their recalls.

All six subjects for whom comparison is possible show a greater match between Phrases 1 and 2 than between Phrases 1 and 3. Nonetheless, most still show a considerable degree of similarity between Phrases 1 and 3. Applying our scoring technique to the original melody yields the percentage values given at the bottom of Table 4.7. It is apparent that most subjects make Phrases 1 and 3 more similar than in the original. This suggests two things: (1) Subjects do reproduce the $A_1A_2B$ schema of the original, and (2) there is a tendency to make the recall more consistent than the original. Memory tends towards structural simplification.

### 4.3.6 Harmonic analysis

Our next analysis concerns the harmonic structure of the recalls and the degree to which they match the harmonic structure of the original. The original has a simple and unambiguous harmonic structure that is primarily conveyed by harmonic notes falling on the strong beats in each bar (Beats 1 and 3) with passing notes falling on weak beats (2 and 4). The harmonic implication of each two-beat segment is determined by its strong beat and is shown as follows:

// I I / V I// I I /V I // V I / $V^7$ I //

(D minor) (F major)

We scored the recalls by counting the number of half-bars that began with a note drawn from the chord of the harmonic sequence prescribed at that point for the original. For example, the first note of Bar I had to be D, F, or A for it to be scored positively. This scoring matched our intuitions about the harmonic structure of the recalls in all cases except the recalls of Subject 2, where it seemed to us that the E that began the fourth half-bar functioned as an appoggiatura to the following D and thus fulfilled a tonic function. This was therefore allowed as a correct harmony.

Table 4.8 shows the percentage of harmonic replication for each subject, broken down over phrases. We observe a great difference among subjects on this measure, ranging from 8% to 97%. The distribution of scores across phrases is, however, also highly variable, with Subjects 4 and 7 representing the two extremes of a distribution. The only previous analysis that showed such intersubject variability was the melodic contour score previously discussed. The possibility exists that the harmonic score might represent simple differences in melodic contour recall. One may argue that high melodic contour recall will necessarily lead to an improved harmonic score. If this is so, we might expect a high positive correlation between the two measures. Accordingly, we computed Spearman's rho and found a non-significant positive correlation of .429. Inspection of Tables 4.1 and 4.8 shows striking dissimilarities. For instance, Subject 6 scores lowest on melodic contour recall (zero) but scores highest on harmonic recall (93%). Her recalls show that she has constructed a new melody on the harmonic and rhythmic structure of the original. Although this subject demonstrates the most wholesale use of this strategy, it crops up in the other subjects, too. Good examples are found in Recalls 1.6, 3.1, and 8.4. Even subjects who show poor harmonic retention overall display isolated pockets of retention. Recall 2.3, for instance, shows a good harmonic match in Phrase 3, even though the previous two phrases show an odd amalgam of

**Table 4.8** Percentage of replication of harmony analysis

| Subject | Phrase 1 | Phrase 2 | Phrase 3 | Total |
|---|---|---|---|---|
| 1 | 63 | 100 | – | 70 |
| 2 | 38 | 39 | 36 | 38 |
| 3 | 83 | 100 | 13 | 66 |
| 4 | 0 | 0 | 25 | 8 |
| 5 | 92 | 75 | 82 | 84 |
| 6 | 91 | 100 | – | 93 |
| 7 | 79 | 83 | 0 | 56 |
| 8 | 100 | 100 | 90 | 97 |
| average | 68 | 67 | 37 | |

Phrases 1 and 3. Likewise, Subject 4 seems to have been totally dominated by Phrase 3 of the original (as remarked upon previously). If we allow her data to be matched against the harmonic structure of Phrase 3 in all cases, then her score rises to 82%.

### 4.3.7 Between-subjects analyses

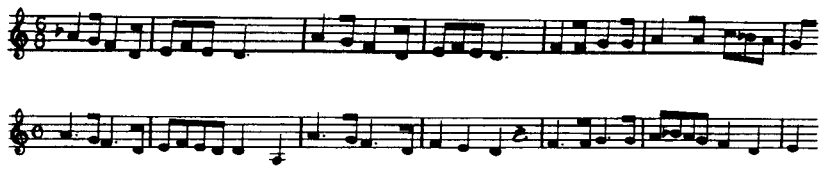
Our final analysis constitutes a sweep over all the previous analyses to ask whether any shows a clear difference between musicians and non-musicians. Table 4.9 presents the mean percentage score for non-musicians and musicians on each measure, and the result of two-way ANOVA carried out on the percentage scores factored by subject group and trial. The only analysis yielding any significant differences is that on the harmonic retention scores. Musicians as a group do significantly better on this measure than non-musicians.

There are doubtless many other analyses that these results could afford. For instance, it is possible to examine the degree to which crude contour (up–down movement) is retained even when precise melodic imitation is lost (e.g. the start of Recall 6.5 has the same contour as the start of the original—two downward steps). One may assess the degree to which general stylistic features of the original (such as characteristic pitch or timing intervals) are reproduced in the recalls. Some of these features may be perceived as particularly salient by some subjects and overdone in the recalls. For instance, Subject 2 overuses the octave leap from the beginning of Phrase 3 of the original. One might also try to ask whether themes and motives from a subject's wider musical knowledge are incorporated into the recalls. For instance, Recalls 3.1–3.4 suggest a contribution from this subject's knowledge of another tune that resembles the original. This is the Christmas carol, 'We Three Kings of Orient Are'. Example 4.10

**Table 4.9** Comparison between musicians and non-musicians

| Analysis | Non-musicians | Musicians | *F*-ratio | *p* |
|---|---|---|---|---|
| Melodic contour | 32 | 52 | .935 | NS |
| Metre | 89 | 87 | .441 | NS |
| Rhythm | 39 | 51 | 1.342 | NS |
| Breaths[a] | 94 | 94 | .175 | NS |
| Phrase match[b] | 14 | 43 | 1.131 | NS |
| Harmony | 45 | 82 | 18.718 | .0001 |

[a]Proportion of phrase endings followed by a breath (omitting Subjects 1 and 6 for missing data).
[b]Percentage match of Phrases 1 and 2 minus percentage match of Phrases 1 and 3. Trials 5 and 6 only. Subjects 1 and 6 omitted for missing data. See Table 4.7.



**Example 4.10** The carol, 'We Three Kings,' compared to Recall 3.1.

gives the first three phrases of this carol above Recall 3.1. The recall resembles the carol's melodic structure more than that of the experimental sequence, particularly at Bars 2 and 5. Other analyses will no doubt suggest themselves. Our data present a small but rich base of material against which to test future hypotheses, and we make no claim to have provided an exhaustive account of the data here. Nonetheless, we feel that the analyses performed so far allow us to draw some quite definite conclusions.

## 4.4 Conclusions

Let us summarize our main findings:

1. Recall of this simple melody is never note-for-note perfect, even in the best case.
2. The exact recalls provided by these subjects are highly related to the original in many respects.
3. The most fundamental feature that is preserved in this melody is the metrical structure. Almost all the recalls are interpretable as maintaining a quadruple metre. This suggests that metre is a primary structural frame for melodic comprehension and recall.
4. Most subjects articulate the metrical frame at a level higher than the bar. The stimulus sequence is coded into two-bar phrases, and the recalls show overwhelming evidence of being governed by this phrase subdivision. Subjects breathe more often between phrases than in other places and reproduce the melodic and rhythmic imitations between phrases that preserve an $A_1A_2B$ structure.
5. Within the metrical phrase structure, subjects do not reproduce the exact rhythms of the original. Rather, they substitute metrical equivalents in about half of the cases.
6. Subjects vary significantly in the degree to which their recalls match the harmony and melodic pattern of the original. This suggests that

listeners can attain a metrical representation independently of attaining a harmonic representation. Memory for harmonic structure seems to be related to musical expertise.

7. There is evidence that harmonic structure may be coded even when exact melodic structure is lost. On some occasions, this leads to a radically new melody, although, more commonly, subjects make small variations on the original melody that are harmonically and metrically consistent.

8. Musicians and non-musicians differ significantly on only one of our measures. This is the ability to retain the harmonic structure of the original.

9. Subjects do not show an improvement in performance over the six trials on any of our measures. Some recalls get longer, but they do not get any better.

What do these results tell us about memory for music? They point toward the notion that memorizing simple, well-formed tonal melodies involves building a mental model of the underlying structure in which not all of the surface detail is necessarily retained. Recall involves processes akin to improvisation, which fill in structurally marked slots according to general constraints about what is appropriate to the piece or genre. We have evidence that different levels of structure are available to people with differing amounts of musical expertise. Musicians code harmonic relationships that seem less accessible to non-musicians. in both groups, however, there is evidence of a great amount of common processing. Subjects seem to share the pool of basic melodic and rhythmic building blocks. Contiguous movements up and down scales and chords of the key account for much of the melodic content; and simple dactylic or equal interval rhythms for almost all of the timing content.

It is not our intention to make detailed theoretical proposals for the nature of the representational system implied by these results. However, we notice that our results fit quite well with the system proposed by Sundberg and Lindblom (1976) for generating Swedish nursery tunes (whose structures are similar to our own material). In their system, a hierarchical metrical tree is generated, which specifies binary subdivisions of a piece into phrases, bars, and beats. Harmonies and rhythms are assigned to terminal elements on the basis of their status within the tree structure, and pitches are chosen to satisfy the harmonies and rhythms specified. To account more precisely for our results, we need to modify the system to allow for the possibility of only *part* of a tree being generated, and we also need to construct a melodic generation process that samples some lexicon of possible melodic

patterns. Some individual differences could be accounted for by a weighting system that makes it more likely that a frequently encountered pattern will be chosen for any particular role. It would also be necessary to build in some real-time constraints. The speed at which a melody is presented clearly has an important bearing on the detail of representation achieved, although we are not in a position to make detailed predictions.

We see the development of theory in this area as working towards a formal system, embodied in a computer program, which takes melodies like ours as its input, and models the type of recall obtained as its output. Sundberg and Lindblom's (1976) system provides the most detailed sketch of a possible output generation system. Several theoretical approaches are relevant to the input procedures. At the level of the individual note, the notions of tonal space developed by, for instance, Sheperd (1982) and Krumhansl (1983) might be used to predict the kind of individual pitch error most likely to find its way into recalls. Our intuition is, however, that such simple relationships will often be overridden in extended melodic representation by factors relevant to melodic shape and imitation and to the preservation of metrical and harmonic structure. We look, therefore, to theories of music perception that generate temporally extended hierarchic structures by means of parsing procedures (in particular, Lerdahl and Jackedoff, 1983; Longuet-Higgins, 1976; and Steedman, 1977).

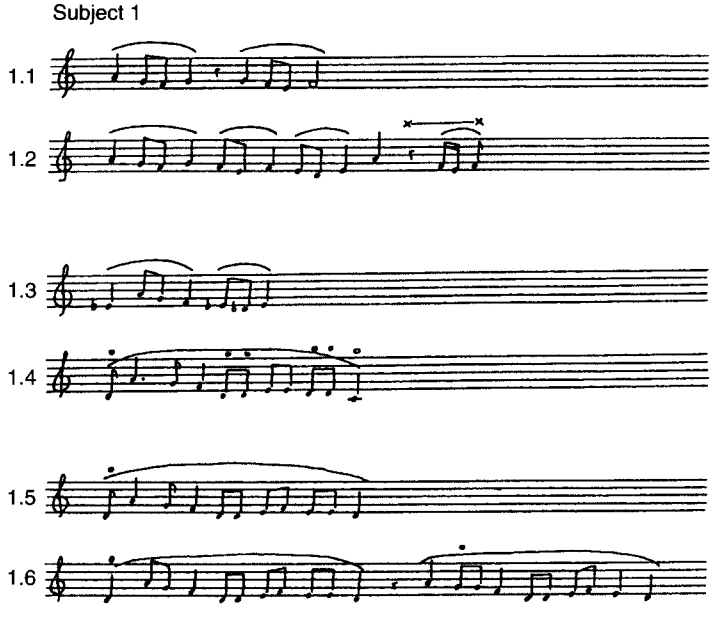
It may be argued that musical recall data cannot directly disconfirm these theories, because a particular unpredicted recall can always be accounted for by some additional factor ignored by the theory (e.g. the freedom exercised by an individual to preserve some linear or statistical feature of the music at the expense of harmony, metre, and hierarchical structure). It would, however, be disconcerting for such theories if a significant proportion of recall distortions could not be accounted for under their assumptions. Nonetheless, we may wish to see recall data not so much as detailed testing grounds for existing theories, but rather as pointers to aspects of memory and representation for which, as yet, no well-developed theory exists. There is no psychological theory of melodic or thematic identity; neither is there any detailed theoretical framework in terms of which we could articulate the nature of the intrusion errors that we noted.

Whatever view one takes about the relationship between recall data and theory testing, we believe our data to be a more concrete and appropriate explanandum for the development of theory than the oft-appealed-to 'intuitions of experienced listeners,' concerning which there can be legitimate and unresolvable dispute.

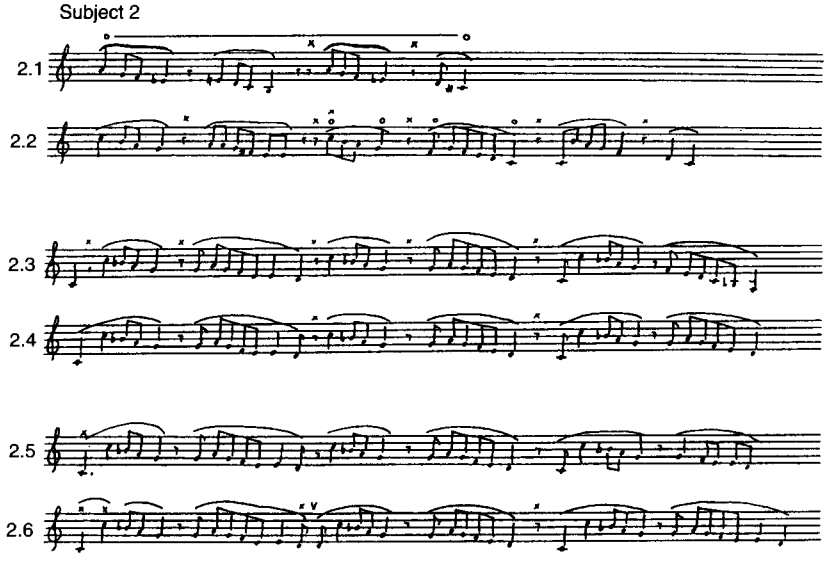
## Appendix

This appendix contains transcriptions of six recall attempts by each of eight subjects aurally presented with Figure 4.1. The following are the special symbols used; details of interpretation are given in the text.

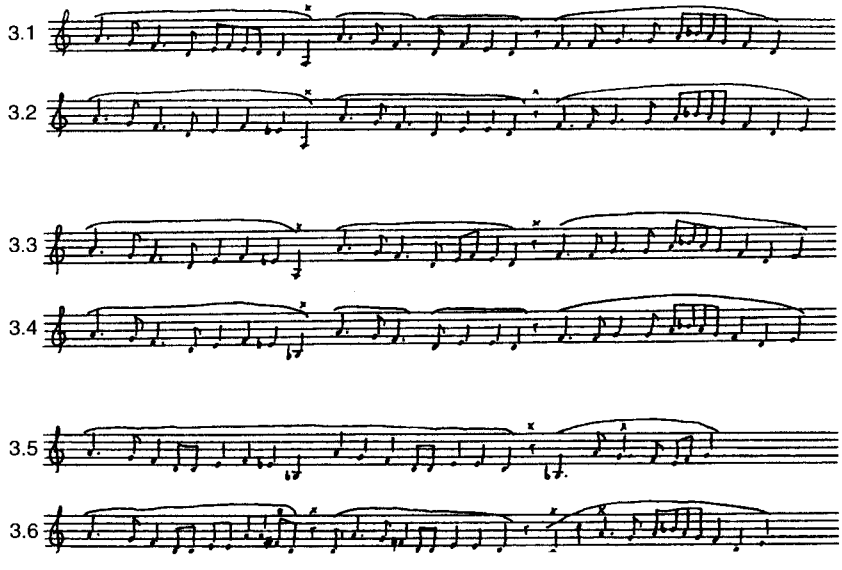
| | |
|---|---|
| slight pause | V |
| in one breath | ⌒ |
| approximate duration | X |
| approximate pitch | O |
| slowing of tempo | ~~~ |
| quickening of tempo | → |



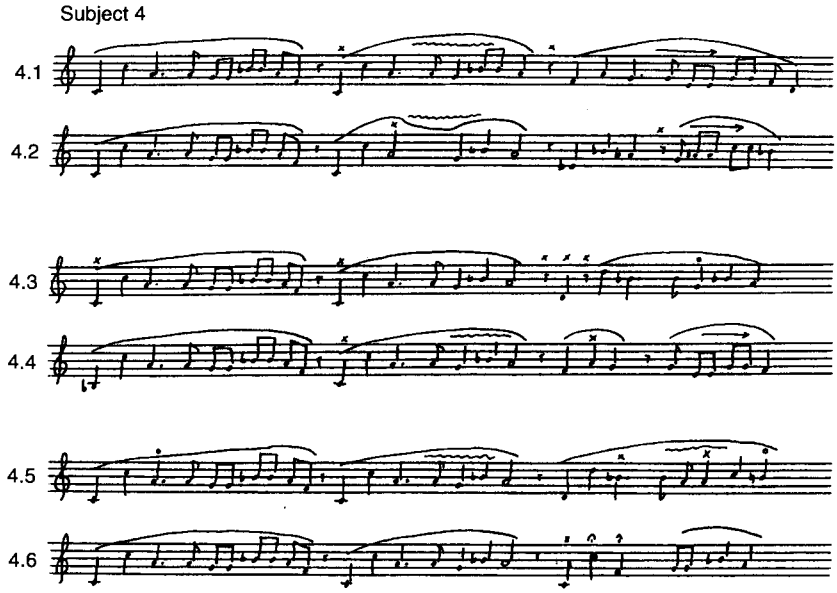
Subject 2



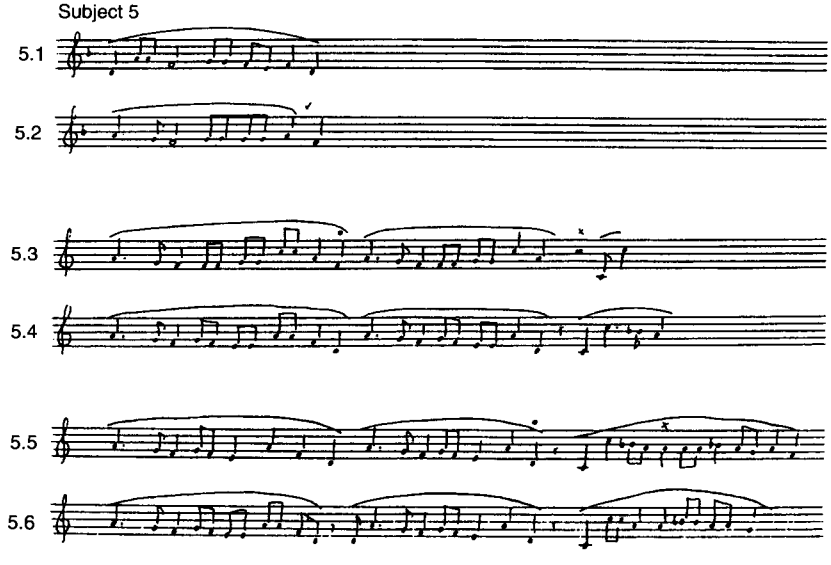

Subject 3

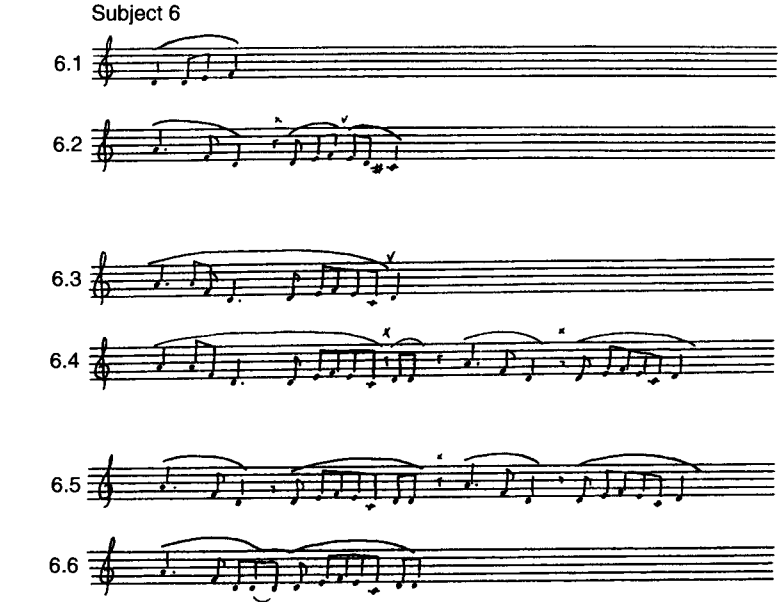


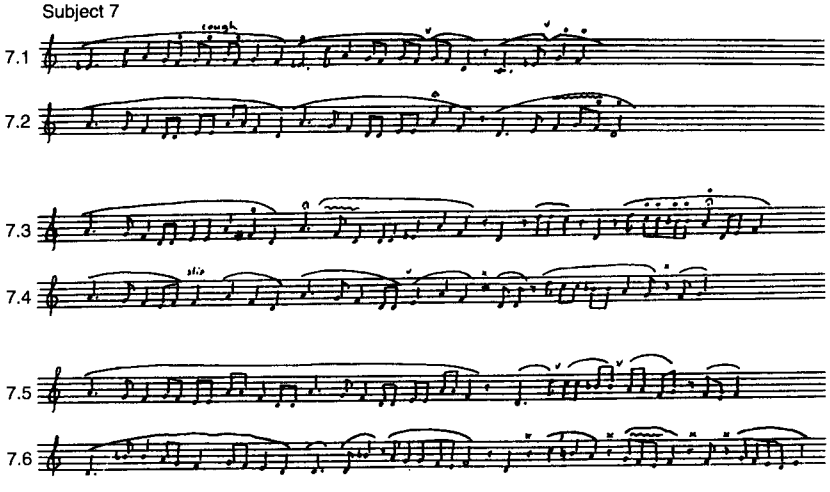
Subject 4

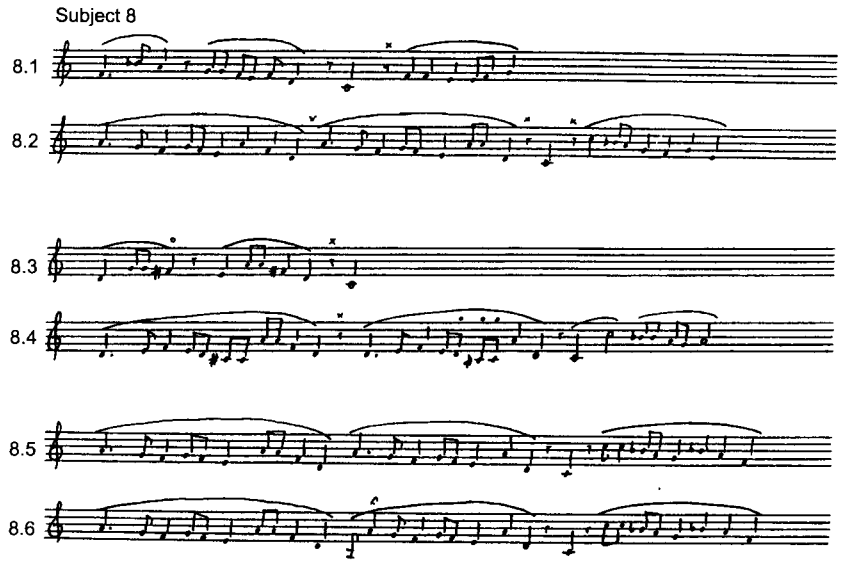



Subject 5




Subject 6
6.1
6.2
6.3
6.4
6.5
6.6


Subject 7
7.1
7.2
7.3
7.4
7.5
7.6




## References

Allport, D. A. The state of cognitive psychology. *Quarterly Journal of Experimental Psychology,* 1975, **27**, 141–152.

Bartok, B., and Lord, A. B. *Serbo-Croatian folk songs.* New York: Columbia University Press, 1951.

Bower, G. H. Experiments in story understanding and recall. *Quarterly Journal of Experimental Psychology,* 1976, **28**, 511–534.

Dowling, W. J. Melodic information processing and its development. In D. Deutsch (Ed.), *The psychology of music.* New York: Academic Press, 1982.

Gabrielsson, A. Performance of rhythm patterns. *Scandinavian Journal of Psychology.* 1974, **15**, 63–72.

Gardner, H. Do babies sing a universal song? *Psychology Today,* 1981, **14**(4), 18–27.

Halpern, A. Organization in memory for music. *Bulletin of the British Psychological Society,* 1983, **36**, A124.

Kintsch, W. On comprehending stories. In M. A. Just and P. A. Carpenter (Eds.), *Cognitive processes in comprehension.* Hillsdale, NJ: Erlbaum, 1977.

Krumhansl, C. L. Perceptual structures for tonal music. *Music Perception,* 1983, **1**, 24–58.

Lerdahl, F., and Jackendoff, R. *A generative theory of tonal music.* Cambridge, MA: MIT Press, 1983.

Longuet-Higgins, H. C. Perception of melodies. *Nature*, 1976, **263**, 646–653.

Moog, H. *The musical experience of the pre-school child* (C. Clarke, Trans.). London: Schott, 1976.

Newell, A. You can't play 20 questions with nature and win. In W. G. Chase (Ed.), *Visual information processing*. London: Academic Press, 1973.

O'Toole, L. M. *The Gateway Russian song book*. London: Collets, 1974.

Pollard-Gott, L. Emergence of thematic concepts in repeated listening to music. *Cognitive Psychology*, 1983, **15**, 66–94.

Seashore, C. E. *The psychology of music*. New York: McGraw-Hill, 1938.

Shaffer, L. H. Performance of Chopin, Bach, and Bartok: Studies in motor programming. *Cognitive Psychology*, 1981, **13**, 326–376.

Shepard, R. N. Structural representation of pitch. In D. Deutsch (Ed.), *The psychology of music*. New York: Academic Press, 1982.

Sloboda, J. A. *The musical mind: The cognitive psychology of music*. London: Oxford University Press, 1985.

Steedman, M. J. The perception of musical rhythm and metre. *Perception*, 1977, **6**, 555–570.

Sundberg, J., and Lindblom, B. Generative theories in language and music descriptions. *Cognition*, 1976, **4**, 99–122.

Zenatti, A. Le développement génetique de la perception musicale. *Monographs Français de Psychologie*. No. 17.

CHAPTER 5

# COGNITION AND REAL MUSIC: THE PSYCHOLOGY OF MUSIC COMES OF AGE

The past ten years have seen such qualitative and quantitative changes in the psychology of music that it is tempting to make a provocative claim. The claim is that until about ten or so years ago there was no scientific discipline which deserved the title 'psychology of music'. Books had been written with this name. Research into music had been carried out using the techniques of experimental psychology. But something essential was lacking. I would loosely adopt Kuhn's (1962) thinking and say that only in the last ten years has the basis for normal paradigmatic science in the psychology of music emerged.

I believe that there are five important characteristics of a healthy paradigm: an agreed set of central problems; agreed methods for working on these problems; agreed theoretical frameworks in which to discuss them; techniques and theories which are specific to the paradigm; research which is appropriate to the whole range of phenomena in the domain being studied.

Although parts of this recipe have been present in different research endeavours, the full set has only recently emerged at one time and within a research community who intercommunicate their ideas and results.

How and why such a discipline should emerge at this particular point in time is a question for sociologists and historians to pursue. Nonetheless, it is abundantly clear that public foci such as conferences and journals have played a crucial part in sustaining the paradigm. Journals such as *Music Perception* and regular conferences, (e.g. Clynes, 1982) have been, and continue to be vitally important to the health of the discipline.

Let us take each of the five characteristics of a paradigm in turn, and elaborate on them in the context of the psychology of music.

## 5.1 The agreed problem

There seems to be a growing consensus that a central problem for the psychology of music is to explain the structure and content of musical experience. This is not to deny the importance of studying the processing

of musical material, but this approach puts the emphasis onto the endpoint of processing—what the processing is supposed to achieve. This involves examining two major questions: a. What is the nature of musical knowledge or representation?; b. How does music have aesthetic and emotional effects?

Both questions are of central importance, but the majority of research has focused on the first of these two questions. I would judge that this definition of the problem excludes much that has been loosely called psychology of music. It, for instance, excludes many studies in music education where the prime aim is to test a teaching method, and where the prime research measure is some measurement of the technical improvement of the learner. I am not denying that music learners have musical experiences, nor that the nature of these experiences contributes to the nature of the learning. It does seem, however, that much research in this area either ignores, or makes simplistic assumptions about, the nature of the experience.

There are, however, research studies in other apparently more cognitively oriented programmes which I would also want to claim are somewhat peripheral to our major concerns. These are the studies which derive their data from musical situations, but which do not attempt to explain their musical significance. Such findings are certainly important and worthy of discovery, but they do need considerable interpretation before they contribute to the psychology of music. Research by Vos and Rasch (1982) that I much admire illustrates this well.

They have shown that perceived onset of a musical note is related to its rise time. The slower the rise, the later the onset appears to be. This has practical implications. For instance, instruments with slower rise times will have to play earlier than others if they are to sound simultaneous. But these findings generalize to any generated sounds whatsoever, musical or not. What might be the specifically musical significance of synchrony and asynchrony? McAdams (1982) has shown that perceived onset asynchrony encourages us to hear two sounds as separate, whilst synchrony encourages fusion. This now begins to relate to musical issues. Perhaps asynchrony is a way of emphasizing polyphony, whilst synchrony encourages perception of homophony. We may then begin to ask questions about what other features of a musical stimulus contribute to the experience of polyphony or homophony. And now we begin to move into the realm of core psychology of music. The questions which we would want to answer in this are: What effects on the musical experience do different types of synchrony patterns have? Why do musical performers choose particularly large or small asynchronies at particular points? Can we identify structural features of music which we would predict to be associated with certain asynchrony patterns? And so on.

One of the seminal publications in our young discipline was Longuet-Higgins' (1976) paper in *Nature*. Longuet-Higgins wrote a computer simulation which took a sequence of pitches and durations as input, and produced a sensibly notated version as output. His thesis was that a central part of experiencing any music as music is the assignation of sounds to positions in tonal and metrical space. This involves determining the key of the melody and its principal stresses. The heuristics in the program are embodiments of strong psychological hypotheses about an experienced listener's processing mechanisms. This research is central to the psychology of music not only because it directly addresses the central problems, but also because the analyses it provides actually accord well with 'the intuitions of experienced listeners', and thus have a strong claim to be an empirically verified account of music experience. The insights of Longuet-Higgins have found echoes at the core of several more recent theoretical enterprises (e.g. Balzano, 1980; Lerdahl and Jackendoff, 1983).

## 5.2 Agreed methods

Some of these methods have been around for a long time. Maybe no one is specific to psychology of music, but recent years have seen researchers exercising a greater freedom to use many different methods to complement one another. In particular, many researchers are eschewing the strait-jacket of traditional methodologies such as the same–different recognition experiment, and are exploring a whole range of methods:

1. Direct recording of performance parameters. This has seen a recent renaissance following the explosion of availability of computer technology for handling and analysing such data (Clarke, 1985; Gabrielsson, 1974; Michon, 1974; Rasch, 1979; Shaffer, 1981, 1984).

2. Computer synthesis of sound patterns to test perceptual hypotheses, sometimes coupled with prior analysis of music with method 1 (Gabrielsson, 1985; Sloboda, 1985a; Sundberg, Askenfelt and Fryden, 1983).

3. Measurement of memory performance. Recognition measures, usually of computer synthesized stimuli, are common (Cuddy, Cohen and Miller, 1979; Deutsch, 1982; Dowling, 1978; Edworthy, 1985; Tan, Aiello and Bever, 1981). Less common are studies of musical recall (Deutsch, 1980; Sloboda and Parker, 1985).

4. Studies of verbal behaviour connected with music, such as judgements about qualities of music, and protocols. These have been most

particularly connected with aesthetic aspects (Crozier, 1974; Smith and Cuddy, 1986), but also such processes as composition (Reitman, 1965; Sloboda, 1985b).

5. Studies of physiological correlates of musical experience. These include evoked potentials, heart rate and other autonomic measures, also studies of hemispheric differences (Davidson and Schwartz, 1977; Harrer and Harrer, 1977; Hirshkowitz, Earle and Paley, 1978, McKee et al., 1973).

6. Simulations of aspects of music perception (Longuet-Higgins, 1976), performance (Sundberg and Fryden, 1985), or generation (Sundberg and Lindblom, 1976; Ulrich, 1971).

A principal concern of all these methods is increasingly to measure or control the *microstructure* of the music or the behaviour, investigators are concerned to discover how the specific actions and events contribute to the global phenomenon. It is no longer adequate, for instance, to simply record the number of errors in some task. We wish to know exactly what errors were made and where.

## 5.3 Agreed theoretical frameworks

A convergence is detectible in the central assumptions made by researchers working in this discipline. Some of these are: Music is constructed according to grammar(s) (Sundberg and Lindblom, 1976); representation is often hierarchic (Deutsch and Ferowe, 1981; Martin, 1972); tonality, metre, and rhythm are psychologically real organizing principles (possibly instantiations of musical universals) (Lerdahi and Jackendoff, 1983); different processes (composition, performance, perception) access the same core representations (Sloboda, 1985b).

This by no means shows that theoretical agreement has been reached, but it does prescribe the kind of things that are theoretically interesting to researchers. In particular, it is arguable that those who hold the above assumptions will particularly want theory to account for: errors in performance, their nature and position; distribution of expressive parameters in performance; differential memorability of musical passages; differential perceived salience of musical events; judgements of well-formedness by listeners; cultural differences in music perception; processes of enculturation and learning, and skill-level differences; structural determinants of musical meaning.

## 5.4 Specificity of the activity

There has been a tendency among researchers to become interested in music because it is a good example of something else: e.g. a complex motor skill; a language-like phenomenon; a complex auditory phenomenon; a set-theoretic entity.

Of course, music is all these things, and it is important to bring in ideas from other areas to elucidate music; but it is possible that many of us, including myself, have been too eager to see music through the spectacles of our own extra-musical preoccupations, and underplay its uniqueness.

It is most precisely at the point where our studies characterize music as wholly unlike anything else that we may be most secure at establishing a real psychology of music. We may locate this uniqueness of music at the confluence of three aspects: 1. Music creates a non-referential world; 2. It does so through the psychological dimensions of tonality and metre; 3. Despite the lack of reference, music has deep emotional significance.

Rhythm and metre are, of course, concepts which have parallels in other domains, such as speech (Martin, 1972). Tonality, however, seems something quite unique to music. The study of scales, harmony, melody, and key seems to be what gives our subject its specificity and its heart. We have begun to unravel the rich and complex structure which makes the tonal system such a fruitful medium for building interesting forms (e.g. Longuet-Higgins, 1972; Balzano, 1980). We have done less to relate tonality to the emotions, but it is becoming much more respectable to pose this as a question for research effort (e.g.Minsky, 1982).

## 5.5 Appropriate research

1. Music psychology demonstrates its maturity by starting to address the range of activities properly described as musical. Although the majority of research relates to perception, increasing research effort is devoted to examining such activities as performance (Clarke, 1985), memory (Sloboda and Parker, 1985), composition (Lerdahi, 1988) and improvisation (Davidson, 1985; Pressing, 1988).

2. Music psychologists are becoming more concerned to use real music in their investigations instead of impoverished music-like stimuli. It is being realized that certain aspects of music will not be understood unless materials of some complexity are used. This does not entail giving up stimulus control; however, more researchers use relatively long sequences (two or more phrases) (Sloboda and Gregory, 1981;

Tan *et al.*, 1981). ensure that both rhythmic and tonal structures are present (DeLiege, 1985), and try to use 'composed' material rather than 'constructed' material where feasible (Pollard-Gott, 1983). This is part of the move towards greater 'ecological validity'.

3. Music psychology increasingly employs subjects in behaviours that are close analogues of common musical activities, such as memorizing for recall (Sloboda and Parker, 1985), sight-reading (Sloboda, 1984), extended listening (Pollard-Gott, 1983), improvising (Sagi, 1988).

4. Music psychology has come to address issues which are not simply of interest to psychologists, but to practitioners of the various arts and skills of music (performers, music theorists, etc.). We find increasing numbers of professional musicians undertaking psychological research, often in collaboration with psychologists (Rosner and Meyer, 1982; Sternberg, Knoll and Zukofsky, 1982). Collaboration of various kinds has been particularly encouraged by such bodies as the Institut de Recherche el de Coordination Acoustique/Musique (IRCAM) in Paris.

## 5.6 The coming of age

The title of this paper commits me, if I am serious about my analogy, to identify a precise historical moment which signified the coming of age of our discipline. I have no doubt at all about what this moment was. It was the publication in 1983 of Lerdahl and Jackendoff's *Generative Theory of Tonal Music*. This is for several reasons:

1. It must be one of the most comprehensively reviewed books in music/psychology (Cady, 1983; Child, 1984; Clarke, 1986; Longuet-Higgins, 1983; Peel and Slawson, 1984; Rosner, 1984). The avid interest with which it was awaited and then devoured by a very wide constituency signals the community of concern around the problems it addresses. It was recognized that the authors had articulated a core problem for music theory and psychology, and had attempted an ambitious answer.

2. It has the problem of the structure of musical knowledge as its main theme.

3. It is explicit and detailed. It offers rules for understanding the perceived importance of each note in an extended piece of music.

4. It is 'universal'. It can account for anything from a small fragment of music right up to a symphonic movement. It also attempts to be universal in another sense, in proposing that culturally specific rules are special cases of culture free rules which apply to all music.
5. It is ambitious. It tries to offer a complete theory which will cover most real music.
6. It is specific. Although many of the structures are borrowed from other domains, notably generative linguistics, the way they are applied to music is unique for a. the notion of several independent but interactive structural analyses of the same passage, b. the distinction between well-formedness rules and preference rules, c. the proposal of special rules to deal with such music-specific events as cadences.
7. It is empirically oriented. The authors claim that specific predictions of the theory may be tested.
8. It is written with the contribution of a highly trained working musician who is prepared to use his own musical intuitions and insights as the basis of theory.

These features do not necessarily imply that Lerdahl and Jackendoff have produced a theory without flaws. The longer we live with it the more apparent its shortcomings are. This is not the point. The mere existence of the theory has significantly accelerated the amount of thinking and research around the problems it addresses (Deliège, 1985; Narmour, 1983; Todd, 1985). In 'coming out' the authors have undoubtedly laid their heads on many chopping blocks, but had they not done so, fewer people would have been motivated to sharpen up the particular hatchets. Lerdahl and Jackendoff are to be congratulated for making themselves so explicit that criticism of them is relatively easy and fruitful. By seeing what they do not do, we should be able to make rapid progress in developing better theories. Any discipline which cannot produce a steady stream of theories of the clarity, complexity and elegance of Lerdahl and Jackendoff, is certainly not to he judged mature. We may have come of age, but we still have a long way to go.

It is particularly heartening to see that Lerdahl and Jackendoff have stimulated diverse type of endeavour. Music theorists have held it up against other theoretical and analytical perspectives (see Clarke, 1986). Scientists have started to conduct empirical studies aimed at testing certain of its predictions. For instance, work carried out in Shaffer's laboratory (Todd, 1985) shows that the pattern of timing deformations in expressive

playing can be predicted quite closely from patterns of prominence derived from a Lerdahl–Jackendoff analysis.

Although some studies have arisen specifically with reference to Lerdahl and Jackendoff's theory, many more have been carried out within the general climate that has been responsible for the intense interest in their work. I would like to illustrate this point with reference to two studies that I have been recently involved with (Sloboda and Parker, 1985; Sloboda, Hermelin and O'Connor, 1985). They both concern a 'real' musical task, that of memorizing composed music for subsequent recall. One study looks at 'normal' processes in relatively unselected subjects. The other study looks at a single case of 'exceptional' memory. In each study a principal concern is with the type of representation which must be postulated to account for memory distortions.

## 5.7 Recall of folk tunes

Although memory has been a major concern of the psychology of music, almost all the serious research has been conducted using a recognition method. Subjects are presented with two short music-like sequences and they are required to detect some difference.

David Parker and myself decided to explore a more direct approach, of playing a real tune to subjects and asking them to attempt to recall, by singing, what they could remember (see Sloboda and Parker, 1985 for further details). The study was devised as a preliminary step on the road to investigating how musicians go about memorizing long pieces of music, a topic we know almost nothing from the viewpoint of empirical science.

Secondly, it seemed to us that a memory recall would provide the most detailed, rich, and explicit clues to the nature of the structures and representations that people use when dealing with tonal music. In particular, an analysis of the type of error made should tell us a great deal about the structures being picked up.

Thirdly, extensive literature search failed to discover any similar analysis of a corpus of recalls.

The data were collected as follows. Short but unfamiliar folk melodies were played on the piano and tape recorded. Subjects were tested individually by listening to each melody and then trying to sing it back to 'la' or any other convenient syllable. Each melody was repeated six times in succession, and so six consecutive recalls were obtained.

Figure 5.1 shows the melody to which we have devoted our major attention. It is the first 12 bars of a Russian folk song called 'Sailor'. It contains 30 notes. None of our subjects had heard it before.

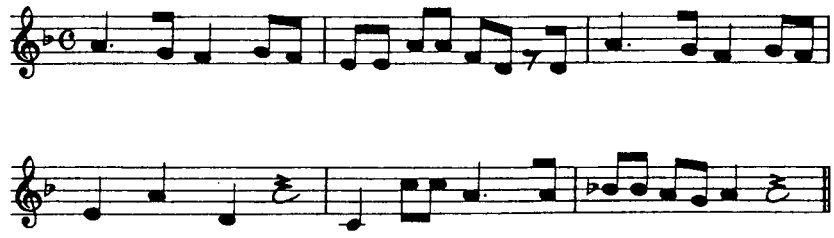


**Figure 5.1** Stimulus melody in Sloboda and Parker (1985) study.

Figure 5.2 shows transcriptions of the six recalls of a single subject on repeated presentation of the tune. This set of recalls displays features typical of the corpus as a whole. Firstly, the recalls are remarkably coherent and definite. This makes them easy to transcribe and analyse. We had fears of vague mumbled hummings which would be impossible to transcribe. Our fears were groundless. Second, not a single recall in our whole corpus is note perfect. Third, everyone who has heard these recalls agrees that the subjects display real memory for the music.

We have subjected the data to a number of analyses which have corroborated these intuitions. First we showed that subjects do not retain a large proportion of the actual melodic and rhythmic sequences present in

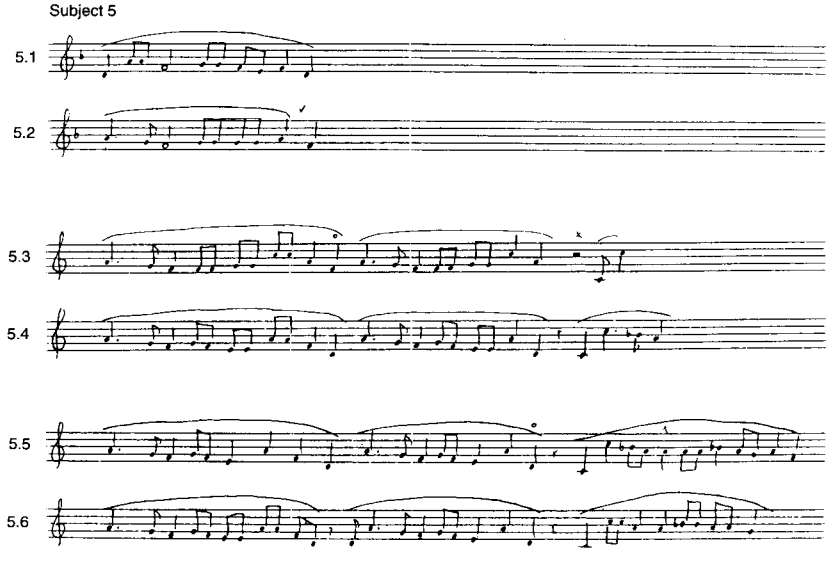



**Figure 5.2** Transcriptions of six consecutive recalls of a typical subject in Sloboda and Parker's (1985) study.

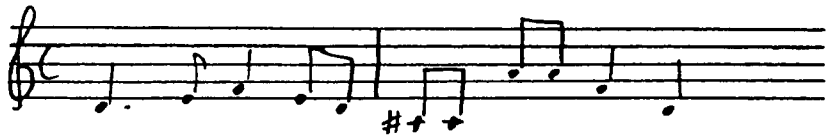


**Figure 5.3** Example of recall of phrase I in Sloboda and Parker's (1985) study.

the original. Even when we give considerable latitude in the positioning of material average scores are still low. Expressed as a percentage of total notes recalled, only 43% of the notes matched the melody given, and only 46% matched the rhythm given. One interpretation to place upon this would be that for about half the time subjects were improvising at random to sustain the melody. We have evidence, however, that their productions were constrained by high-level features of the original melody.

The first of these features is the metrical structure. We found that we could satisfactorily parse 88% of the recalls as in consistent 4-beat metre, like the original, by applying the kind of parsing rules formalized by Lerdahl and Jackendoff, and others. In other words, subjects were extracting the metre of the original as a frame for their recalls.

The second of these features is harmony. In this case we found quite wide subject differences. Half of our subjects were music students, the other half were just interested in music. In the case of the music students 81% of the material was harmonically appropriate. That is to say, whilst the material was melodically and rhythmically new, it used notes from the chords implied in the original. We scored the recalls by looking at the notes subjects chose to begin each half-bar, giving a score of 1 if the note came from the chord prescribed by the original at that point. Figure 5.3 gives an example of memory which is harmonically and metrically perfect but contains almost none of the actual notes of the original.

A third higher-order feature of these recalls concerns phrases. It should be clear that the original melody falls into three phrases of two bars each. The second phrase rather closely imitates the first, but the third phrase introduces new material. Subjects showed awareness of this structure in two ways, through breath positions, and through imitative relations between phrases:

1. *Breaths.* In each subject's recall we defined the end of phrase 1 as the last note in bar 2 whose duration was a crotchet or longer. The end of phrase 2 was similarly defined as the last long note in bar 4. We then compared the number of times subjects breathed at a phrase end to the number of times they breathed elsewhere. Almost every phrase

end was associated with a breath, and almost no other point was. Subjects made the phrase structure explicit through their breathing.

2. *Imitative structure.* We asked whether subjects preserved the imitative relationships between phrases 1 and 2, making the third phrase different. To do this we took each recall, and compared phrase 1, beat by beat, with phrase 2, awarding a score of 1 if the pitches were identical, and a score of I if the rhythms were identical. We found that phrases 1 and 2 had a 71% match, whereas phrases 1 and 3 had only a 41% match. So, as in the original, subjects preserved the closer relation of phrases 1 and 2.

How do we summarize our findings? They show that a trivially simple ecological technique coupled with appropriate data analyses yields good insights into structural aspects of musical memory. They suggest that memory of a tonal melody involves building up a model of the underlying structure in which not all of the surface details are necessarily retained. Recall involves a reconstruction process which fills in structural slots according to general constraints. The type of structure noticed depends to a certain extent on experience. In our sample only the musically trained show a majority retention of harmonic structure. All the subjects were able to access the metrical and phrase structure.

## 5.8 An exceptional musical memory

The second study I wish to report is of a much more dramatic and sensational kind. We saw in the previous experiment that ordinary people, even those with musical training, are not very good at retaining exact details of even a short melody. Yet musical folklore abounds with stories of people gifted with quite exceptionally accurate musical memories. The most famous story probably concerns Mozart's exploits with Allegri's *Miserere* (Anderson, 1966). However trustworthy we believe the observers of this feat of memory to be, the actual record of the attempt does not survive. It would be of considerable interest to examine a contemporary musician with memory powers equal to that claimed for Mozart.

A rigorous examination of such a memory would not only be of intrinsic fascination, it would also have theoretical significance. We could ask whether a prodigious musical memory is the result of some abnormal mode of representing music (such as a vivid 'eidetic' store) or whether it comes about through the same kind of structural assimilation as more mundane memory.

If the former, then we would expect it hardly to matter whether the material to be remembered was easily assimilable to well learned structures. If the latter, however, we would expect music containing familiar structures to be remembered better than music lacking some of these structures. We would also expect any errors that did occur to be predominantly structure preserving.

In 1984, an opportunity presented itself to examine a young man of allegedly remarkable memory powers for piano music (see Sloboda, Hermelin and O'Connor, 1985 for further details). He had been discovered during a large scale study of *idiots savants*. These are people of low intelligence, often mentally handicapped. who appear to have an isolated pocket of intellectual excellence. A small number of such people excel in music. Sometimes their excellence is only relative to their mental age; and the skill would not be unusual in a normal person with some musical training. This particular man, who we shall call NP, was particularly interesting because his alleged memory skill was far in excess of any normal accomplishment. The claim was that he could memorize an extended movement in two or three hearings.

NP was 22 years old at time of testing. He lives in a home for the autistic, and displays the typical range of autistic symptoms, including language deficiency, social withdrawal, gaze aversion, and obsessive behaviour patterns. His major obsession is music. He spends as much of every day as possible playing the piano or listening to music.

Prior to the experimental sessions, a number of piano pieces were prerecorded on tape. In the event, we were able to expose NP to two of them, neither of which, to our knowledge, he had learned before: 1. a lyric piece by the nineteenth-century Norwegian composer Edward Grieg (*Melodie* Opus 47 no 3): 2. a short atonal piece by the twentieth-century Hungarian composer Bela Bartok ('Whole Tone Scale' from *Mikrokosmos* Book 5).

Neither piece was technically virtuosic, and well within the technical capacity of the subject. In the trials, NP simply listened to the recordings. At no stage did he see anyone visually demonstrating the appropriate motor movements on the piano, nor did he have the chance to examine a score. NP heard each entire piece through, then heard it in shorter sections, as judged appropriate from the signals he gave, and his moment to moment capacity. After each presentation of the piece (or section) he was asked to try and play what we had heard. The whole session was recorded on tape, and all recall attempts were transcribed for analysis.

We found that NP learned the Grieg piece to almost note-perfect criterion in about 12 minutes, during which time he heard no more than four repetitions of any section of the piece. The piece is 66 bars long and contains

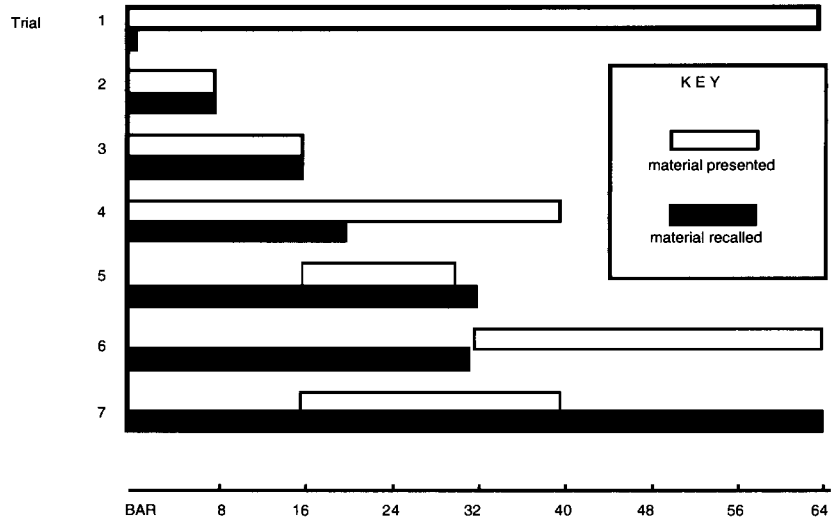



**Figure 5.4** Learning history for subject NP on Grieg Op. 47 no. 3 (from Sloboda, Hermelin and O'Connor, 1985).

310 notes (discounting adjacent repetitions of the same chord). Figure 5.4 shows a summary learning history. The upper white bar of each pair shows what NP heard on a given trial. The lower black bar shows what was recalled. Although recall was never 100% perfect, the deviations from score were very minor; an occasional omission or addition of a subsidiary note which hardly changed the character, and never the structure.

Table 5.1 shows some summary quantitative data. The top figure is the number of notes whose recall was attempted over all trials. The second figure is the number of wrong notes (sum of notes omitted which should have been played and notes played which should not have been). The third figure is the number of times that there was a stumble, pause, or backtrack. The fourth figure represents the wrong notes and stumbles as a percentage of notes attempted, to give a gross error rate.

**Table 5.1** Summary quantitative data of the exceptional musical memory study

| | **NP** | | **AS** | |
|---|---|---|---|---|
| | **Grieg** | **Bartok** | **Grieg** | **Bartok** |
| Notes attempted | 798 | 277 | 354 | 153 |
| Notes wrong | 54 | 137 | 265 | 12 |
| Stumbles, pauses | 11 | 37 | 19 | 10 |
| Total % error | 8 | 63 | 80 | 14 |

Table 5.1 includes data from a number of comparisons that help us to understand better the nature of NP's achievement with the Grieg. The second column shows data from NP's attempt to recall the Bartok piece. Although shorter than the Grieg, and with just as simple a rhythmic and formal structure, the reproduction attempts were hesitant, error-prone and unconfident. For instance, in four attempts to recall the first 51 notes of the piece, NP never played more than 36 correctly, and each time with frequent stumbles. Although many fewer notes were attempted overall, the error rate was a massive 63%. It appears that NP cannot function at a high level outside the framework supplied by conventional tonal harmony.

A second comparison summarized in Table 5.1 comes from the attempts of a professional pianist AS to memorize the same two pieces. AS was an approximate age and experience control for NP. Figure 5.5 shows the learning history for the Grieg piece in the same format as Figure 5.4 for NP. The hatched bars are recalls containing large numbers of errors.

This shows that good recall was obtained for the first eight bars alone. When further material was attempted, even the original level of performance on the first eight bars was lost. It was clear that AS was simply overwhelmed by the amount of material in the Grieg. Although some of the right hand melody was reproduced reasonably well, the accompanying

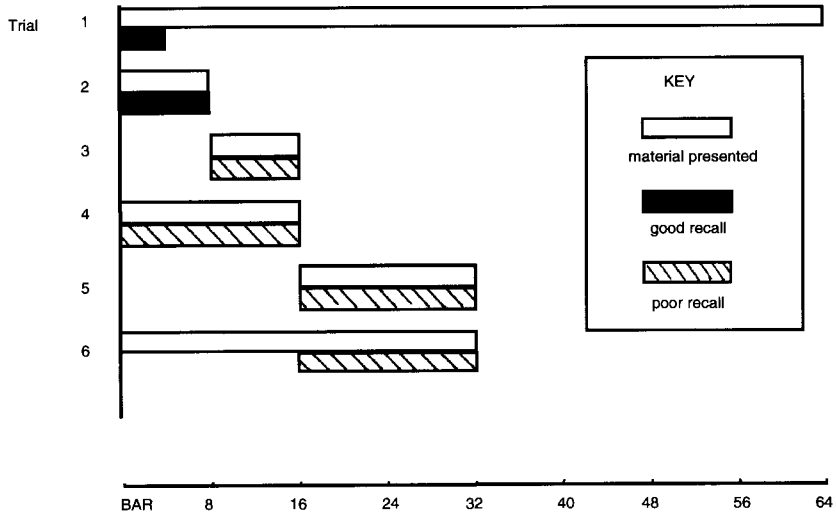



**Figure 5.5** Learning history for subject AS on Grieg Op. 47 no. 3 (from Sloboda, Hermelin and O'Connor, 1985).

chords bore no resemblance to the original for significant portions of the recalls.

The Bartok, in contrast, showed negligible error. This was partly due to the brevity of the section attempted (first eight bars only), but it was clear that the atonality did not disrupt AS to anything like the extent that it did for NP. Like many contemporary musicians, AS had considerable experience of this kind of music.

A third comparison was obtained by looking at the results of some simple tests of NP's verbal memory. This was found to be consistent with his general level of intellectual functioning. He had a forward digit span of 5 (as compared to an average of 7) and a backwards span of 4 (as compared to an average of 6). He was able to recall simple 7-word sentences haltingly but correctly, but could only recall 4 out of 7 words when word order was jumbled. There was no evidence of any exceptional memory capacity outside the realms of tonal music.

It is quite clear from these data that NP is not the possessor of some primitive but high-capacity 'echoic' or 'mimicking' memory, but has a narrow but intensely highly-developed window of 'normal' memory which operates, as in most known cases of expertise, by representing material in terms of familiar higher-order structures. This ability, like other exceptional skills, has been developed through many thousands of hours of practise over many years, in conjunction with an intense degree of absorption with the material concerned, and an unusually high degree of motivation to succeed at this particular task (cf. Chase and Ericsson, 1981).

In sum, these two studies illustrate how naturalistic musical situations may be exploited to provide insights into the nature of musical representation. They accord well with the theoretical perspectives formalized by Lerdahl and Jackendoff, in that they suggest overlapping structured hierarchic representations of music (rhythmic, metrical, tonal, etc.), where higher levels have greater 'durability' than lower levels.

## 5.9 Conclusions

The type and range of endeavours summarized in this paper convinces me that our discipline of psychology of music has indeed come of age, and that we can happily take risks. We no longer need to convince a sceptical psychological world that we have rigour, scientific respectability and relevance. Rather, we should boldly extend the frontiers of our subject by tackling some of the outstanding difficult issues head-on.

Lerdahl and Jackendoff have set new standards for comprehensiveness and boldness in theory. Let us hope that they stimulate not only criticism

Harrer, G. and Harrer, H. (1977) Music, emotion and autonomic function. In M. Critchley and R. A. Henson (Eds.). *Music and the brain. Studies in the neurology of music.* London: Heinemann.

Hirshkowitz, M., Earle, J., and Paley, B. (1978) EEG Alpha asymmetry in musicians and non-musicians: A study of hemisphere specialization. *Neuropsychologia*, **16**, 125–128.

Kuhn, T. S. (1962) *The structure of scientific revolutions.* Chicago: University of Chicago Press.

Lerdahl, F. and Jackendoff, R. (1983) *A generative theory of tonal music.* Cambridge, MA: MIT Press.

Lerdahl, F. (1988) Compositional systems and psychological reality. In J. Sloboda (Ed.). *Generative processes in music: The psychology of composition, improvisation and performance.* London: Oxford University Press.

Longuet-Higgins, H. C. (1972) Making sense of music. *Proceedings of the Royal Institute of Great Britain*, **45**, 87–105.

Longuet-Higgins, H. C. (1976) Perception of melodies. *Nature*, **263**, 646–653.

Longuet-Higgins, H. C. (1983) All in theory—the analysis of music. *Nature*, **304**, 93.

Martin, J. G. (1972) Rhythmic (hierarchical) versus serial structure in speech and other behaviour. *Psychological Review*, **79**, 487–509.

McAdams, S. (1982) Spectral fusion and the creation of auditory images. In M. Clynes (Ed.). *Music, mind and brain: The neuropsychology of music.* New York: Plenum.

McKee, G., Humphrey, B., and McAdam, D. W. (1973) Scaled lateralisation of alpha activity during linguistic and musical tasks. *Psychophysiology*, **10**, 441–443.

Michon, J. A. (1974) Programs and 'programs' for sequential patterns in motor behaviour. *Brain Research*, **71**, 413–424.

Minsky, M. (1982) Music, mind and meaning. In M. Clynes (Ed.). *Music, mind and brain: The neuropsychology of music.* New York: Plenum.

Narmour, E. (1983) Some major theoretical problems concerning the concept of hierarchy in the analysis of tonal music. *Music Perception*, **1.2**, 129–199.

Peel, J. and Slawson, W. (1984) Review of *A generative theory of tonal music. Journal of Music Theory*, **28.2**, 271–294.

Pollard-Gott, L. (1983) Emergence of thematic concepts in repeated listening to. music. *Cognitive Psychology*, **15**, 66–94.

Pressing, J. (1988) Improvisation: methods and models. In J. Sloboda (Ed.). Generative processes in music: *The psychology of composition, improvisation and performance.* London: Oxford University Press.

Rasch, R. A. (1979) Synchronization in performed ensemble music. *Acustica*, **43**, 121–131.

Reitman, W. R. (1965) *Cognition and thought.* New York: Wiley.

Rosner, B. (1984) Review of *A generative theory of tonal music. Music Perception,* **2.2**, 275–290.

Rosner, B. and Meyer, L. B. (1982) Melodic processes and the perception of music. In D. Deutsch (Ed.). *The psychology of music.* New York: Academic Press.

Sagi, M. and Vitanyi, I. (1988) Experimental research into musical generative ability. In J. Sloboda (Ed.). *Generative processes in music: The psychology of composition improvisation and performance.* London: Oxford University Press.

Shaffer, L. H. (1981) Performance of Chopin, Bach, and Bartok: Studies in motor programming. *Cognitive Psychology,* **13**, 326–376.

Shaffer, L. H. (1984) Timing in solo and duet piano performances. *Quarterly Journal of Experimental Psychology,* **36A**, 577–595.

Sloboda, J. A. (1984) Experimental studies of music reading: A review. *Music Perception,* **2.2**, 222–236.

Sloboda, J. A. (1985a). Expressive skill in two pianists: Style and effectiveness in music performance. *Canadian Journal of Psychology,* **39**, 273–293.

Sloboda, J. A. (1985b). *The musical mind: The cognitive psychology of music.* London: Oxford University Press.

Sloboda, J. A. and Gregory, A. H. (1981) The psychological reality of musical segments. *Canadian Journal of Psychology,* **34**, 274–280.

Sloboda, J. A.., Hermelin, B., and O'Connor, N. (1985) An exceptional musical memory. *Music Perception,* **3.2**, 155–170.

Sloboda, J. A. and Parker, D. H.H. (1985) Immediate recall of melodies. In P. Howell, I. Cross and R. West (Eds). *Musical structure and cognition.* New York: Academic Press.

Smith, K. C. and Cuddy, L. L. (1986) The pleasingness of melodic sequences: Contrasting effects of repetition and rule familiarity. *Psychology of Music,* **14.1**, 17–32.

Sternberg, S., Knoll, R. L. and Zukofsky, P. (1982) Timing by skilled musicians. In D. Deutsch (Ed.). *The psychology of music.* New York: Academic Press.

Sundberg, J., Askenfelt, A., and Fryden, L. (1983) Music performance: A synthesis by rule approach. *Computer Music Journal,* **7**, 37–43.

Sundberg, J., and Fryden, L. (1976) Teaching a computer to play melodies musically. In *Analytica. Studies in honor of Ingmar Begtsson.* Publication no 47 issued by the Royal Swedish Academy of Music, 67–76.

Tan, N., Aiello, R., and Bever, T. G. (1981) Harmonic structure as a determinant of melodic organization. *Memory and Cognition,* **9**, 533–539.

Todd, N. (1985) A model of expressive timing in tonal music. *Music Perception,* **3.1**, 33–58.

Ulrich. J. W. (1977) The analysis and synthesis of jazz by computer. *Proceedings of the 5th International Joint Conference on Artificial Intelligence,* pp. 865–872.

Vos, J., and Rasch, R. (1982) The perceptual onset of musical tones. In M. Clynes (Ed.). *Music, mind and brain: The neuropsychology of music.* New York: Plenum.

CHAPTER 6

# PSYCHOLOGICAL STRUCTURES IN MUSIC: CORE RESEARCH 1980–1990

## 6.1 Selecting the core topics

The 1980s saw a huge increase in the quantity and quality of psychologically oriented investigations into music. A full account of recent advances would fill several books. A brief chapter like this must make drastic selections if it is to be more than an annotated bibliography. It seems important to make the selection on explicit and rational bases, rather than on personal whim, and so I devote the first section of this chapter to outlining my method for selecting what to include.

In 1983 I completed a major review of the field of psychology of music as I saw it at the time. This was published as a book (Sloboda 1985a). In this book I paid as much attention to underdeveloped areas of study that I believed deserved attention (such a composition and performance) as I did to well-studied areas (such as the perception of brief musical fragments). My intention for this essay is to complement the earlier book by concentrating on work published since 1983, and by focusing on the central themes that emerge from this work. This means that if a topic is poorly represented in the central literature, I will not deal with it here.

It is always difficult to define what counts as central work, but my task has been made easier by the inauguration in 1983 of the journal *Music Perception* under the inspired and dedicated editorship of Diana Deutsch. This journal has rapidly established itself as the flagship of the psychology of music confraternity. Its bias is experimental and theoretical, seeking to understand the basic perceptual and cognitive processes that underlie musical behaviour; and it is the only psychological journal that attracts significant contributions from music theorists as well as psychologists. Because of its unrivalled pre-eminence, it seems entirely sensible to confine my review to the issues raised within the pages of that journal (volumes 1–5, 1983–8).

Even this, however, is too wide a brief, since the first five volumes contain over 100 contributions. To determine which papers could be classed as relating to core issues, I undertook a citation analysis of all the papers in the first five volumes. Table 6.1 shows the ten most cited publications

**Table 6.1** The ten most frequently cited publications in Volumes 1–5 of *Music Perception*

| Author | Date | Title and source | Number of citations |
|---|---|---|---|
| Deutsch, D. and Feroe, J. | 1981 | The internal representation of pitch sequences in tonal music', *Psychological Review*, **88**, 503–22. | 17 |
| Lerdahl, F. and Jackendoff, R. | 1983 | *A Generative Theory of Tonal Music*, Cambridge, MA: MIT Press. | 17 |
| Dowling, W. J. | 1978 | Scale and contour: two components of a theory of memory for melodies. *Psychological Review*, **85**, 341–354. | 15 |
| Krumhansl, C. L. | 1979 | The psychological representation of musical pitch in a musical context. *Cognitive Psychology*, **11**, 346–374. | 15 |
| Schenker, H. | 1979 | *Free Composition*, trans. E. Oster, New York: Longman. Originally published as *Die freie Satz* (1935). | 12 |
| Helmholtz, H. | 1954 | *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, trans. A. J. Ellis, New York: Dover. Originally published 1863. | 12 |
| Krumhansl, C, L. and Kessler, E. | 1982 | Tracing the dynamic changes in perceived tonal organisation in a spatial representation of musical keys. *Psychological Review*, **89**, 334–368. | 12 |
| Shepard, R. N. | 1982 | Structural representations of musical pitch. In: D. Deutsch (ed.) *Psychology of Music*. New York: Academic Press. | 11 |
| Krumhansl, C L. and Shepard, R. N. | 1979 | Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, **5**, 579–94. | 10 |
| Meyer, L. B. | 1973 | *Explaining Music*, Berkeley: University of California Press. | 9 |

summed across all contributions to the journal. These ten key publications were then used as a 'sieve' to select papers from the journal to form the basis of this review. The criterion adopted for inclusion was that the paper should cite at least three of these ten key publications. Table 6.2 gives a list of the nineteen papers selected in this way. They are listed in citation

**Table 6.2** Nineteen core articles from Volumes 1–5 of *Music Perception*

| Author(s) | Volume | Number of items from Table 6.1 cited by author |
|---|---|---|
| Rosner, B. S. and Meyer, L. B. | 4.1 | 8 |
| Lerdahl, F. | 5.3 | 8 |
| Krumhansl, C. L. | 1.1 | 7 |
| Brown, H. | 5.3 | 7 |
| Deutsch, D. and Boulanger, R. C. | 2.1 | 6 |
| Roberts, L. A. and Shaw, M. L. | 2.1 | 6 |
| Bharucha, Jj. | 5.1 | 6 |
| Kessler, E. J., Hansen, C., and Shepard, R. N. | 2.2 | 5 |
| Hantz, E. | 2.2 | 5 |
| Cuddy, L. | 2.3 | 5 |
| Lerdahi, F. and Jackendoff, R. | 1.2 | 4 |
| Butler, D. and Brown, H. | 2.1 | 4 |
| Swain, J. P. | 4.1 | 4 |
| Krumhansl, C. L., Sandell, G. J., and Sergeant D. C. | 5.1 | 4 |
| Cross, I., West, R., and Howell, P. | 2.3 | 3 |
| Todd, N. | 3.1 | 3 |
| Stoffer, T. H. | 3.2 | 3 |
| Deliège, I. | 4.4 | 3 |
| Cook, N. | 5.2 | 3 |

order, so that the paper citing the largest number of the ten key references appears at the top.

## 6.2 The psychological approach to music

The fundamental characteristic that distinguishes the psychologist of music from the music theorist or analyst is the former's concern with empirical measurement of musical behaviour or response. A secondary concern is with 'generalizability' (both across listeners and across pieces). Psychologists are concerned to discover commonalities of response that hold across members of a population. Let us take a hypothetical example of what this would mean in practice.

Theorist X, writing about musical passage Y, states that a certain sequence leads to a sense of resolution at point Z. This statement is primarily based on the theorist's own informed intuitions. He or she has listened to the passage in question, has compared it to other passages, and has applied to it formal descriptions that may or may not be drawn from historically established

modes of musical discourse. It is up to other readers to agree or disagree with the analysis offered.

Psychologist A, wishing to verify this statement, would begin by asking whether the notion of 'resolution' can be operationalized. Is the meaning of the term clear enough that a task can be devised that would allow judgements of 'resolution' to be measured in an objective and unambiguous way? A typical psychological operationalization of such a concept might be embedded in an instruction such as this: 'please listen to each of the following short musical sequences and rate, on a scale of 1 to 10, the degree to which you perceive each sequence as coming to an acceptable close (i.e. the degree to which the sequence sounds 'finished'). Award the mark of 1 if you hear the sequence as not at all closed, and 10 if you hear it as perfectly closed.'

Let us suppose, for convenience, that passage Y contains twelve chords, and that the theorist proposes the eighth chord in the sequence as point Z (the point of maximum resolution). A possible tactic for the psychologist would be to construct a set of sequences made up of between one and twelve chords, each starting on chord 1 of Y but ending on a different chord of the sequence. A psychological verification of X's statement would be obtained if a random sample of an appropriate population rated the sequence ending on chord 8 as significantly more closed than other sequences. Typically, sample size would be twenty to forty individuals, drawn from student populations with general musical interests but little formal theory or analysis training.

Problems of operationalizing musical responses are not trivial ones. These problems explain, to a significant degree, why psychologists have made so little progress in exploring the higher levels of musical response, including the emotional and aesthetic aspects of the response. Even if all such problems were solved, however, it would not make psychological work easy. There are just too many statements in the music literature for psychologists to verify each one of them. The work required to verify the single sentence uttered by Theorist X concerning passage Y could easily keep a psychological researcher busy for weeks, if not months. At the end of the research, all that would have been accomplished would be a verification of X's acumen with respect to that passage.

The selection of research topics must be guided by further considerations, and these considerations are ones that psychologists and music theorists probably share. Both professions are centrally concerned with explaining responses to music. So if X's views on Y are to be accorded any professional weight they should contain reference to some general principles by virtue of which the specific judgement is made (i.e. Z is the point of closure

because it has formal properties K, L, M, etc.). It is proposed that closure is generally signalled by such properties within a musical genre.

Given such potential explanations the psychologist is now in a much better position to carry out useful research. Musical passages may be collected (or constructed) in which the presence or absence of the various formal features K, L, M, etc., is systematically manipulated, and listeners' responses to these passages measured in some way. This activity will commend itself to psychologists, particularly when the features so defined are common to a whole range of music, rather than a small sub-set of pieces or genres. For this reason, the central topics in psychological research have been concerned with aspects of musical structure that are common to the widest body of music for which there exists a formal literature of theory and analysis: Western tonal music, whether art, folk, or popular.

This concern is reflected in the citation orientation of contemporary psychological research (as shown in Table 6.1). More psychologists take Schenker or Meyer, or Lerdahl and Jackendoff as their analytic reference point than, say, Babbitt (who rates sixty-fifth in the citation ranking, with only three out of one hundred authors referring to him). This does not simply reflect the musical conservatism (or populism) of psychologists (although this may be a contributing factor), but a genuine scientific concern to explain historically and culturally pervasive aspects of music before looking at forms that inhabit a more prescribed and closed domain.

The other thing that is abundantly clear from Table 6.1 is the predominance of pitch (over rhythm, metre, timbre and form) as a dimension for psychological investigations. Only one of the top ten citations (Lerdahl and Jackendoff) deals with rhythm, metre, or large-scale form in any substantial way.

## 6.3 Krumhansl's research: a benchmark

Table 6.1 shows that the work of Carol Krumhansl and associates has achieved an unrivalled pre-eminence in the field, with three out of the top ten citations going to work from her laboratory. This work provided a solid benchmark for the whole field of psychology of music for the 1980s. It is not necessarily the most adventurous or far-reaching research, but it represents a secure platform of scientific achievement which has undoubtedly raised the general standards and expectations of the field to new levels, and against which any new development must be compared.

Krumhansl's work (ably summarized in Krumhansl 1983) is essentially an experimental verification of traditional theoretical notions about tonal relationships (between notes, chords and keys). It is common in many

formal and informal theories of tonality to refer to the keys of C major and F# major as 'distant', whilst C major and G major are 'close'. Similarly, the third and fifth degrees of the major scale are held to be musically closer to the tonic than the second degree (which is actually closer in pitch). Do these, and other theoretically accepted statements, have good grounding in the actual responses of theoretically naive listeners?

Krumhansl's method for investigating such questions is to use experimentally constructed 'minimal' stimuli which are usually computer generated. This means that all aspects other than the ones under investigation can be strictly controlled (duration, loudness, attack, etc.). Each stimulus consists of two parts: (a) a tonal context, which could be a scale, a tonic chord, or a perfect cadence in the relevant key, (b) a test stimulus, which could be a note, a chord or a pair of items. When the test stimulus is a single item, listeners are asked to rate the degree to which it fits with or completes the context. In the case of pairs of items, listeners may be asked to rate the similarity or closeness of the two items to one another in the context. In all cases, such ratings are made on a numerical scale.

### 6.3.1 Rating of notes in a tonal context

A specific example of this approach is a study by Krumhansl and Kessler (1982) which asked adult listeners to rate each note of the chromatic scale for how well it fitted, in a musical sense, with the key-defining context. The listeners had a moderate to high level of musical experience, but no explicit instruction in music theory. Figure 6.1 shows the mean results for the major and minor key contexts (averaged over all keys as if in C). The results are clear and unequivocal. Notes of the tonic triad were given the highest rating, followed by notes of the diatonic scale, with chromatic notes being given the lowest rating. This exactly confirms traditional music theory.

Thinkers about music have long been concerned to offer explanations for why certain pitch relationships (such as the octave and the fifth) seem particularly privileged. The appearance of Helmholtz high in the citation rankings signals psychologists' recognition of his seminal contribution in this area. He pointed to the fact that many pitch combinations that we judge most consonant or harmonically central have a strong coincidence of harmonic overtones. So, for instance, the first (and usually strongest) non-octave harmonic of a natural tone is the fifth degree of the scale, the next is the major third, and so on.

Although overtone consonance is a probable factor in the historical development of musical scales, it has a number of shortcomings as a psychological explanation. For one thing, consonance rankings do not match Krumhansl's rankings in a minor scale context, as Figure 6.1 shows. In particular, the

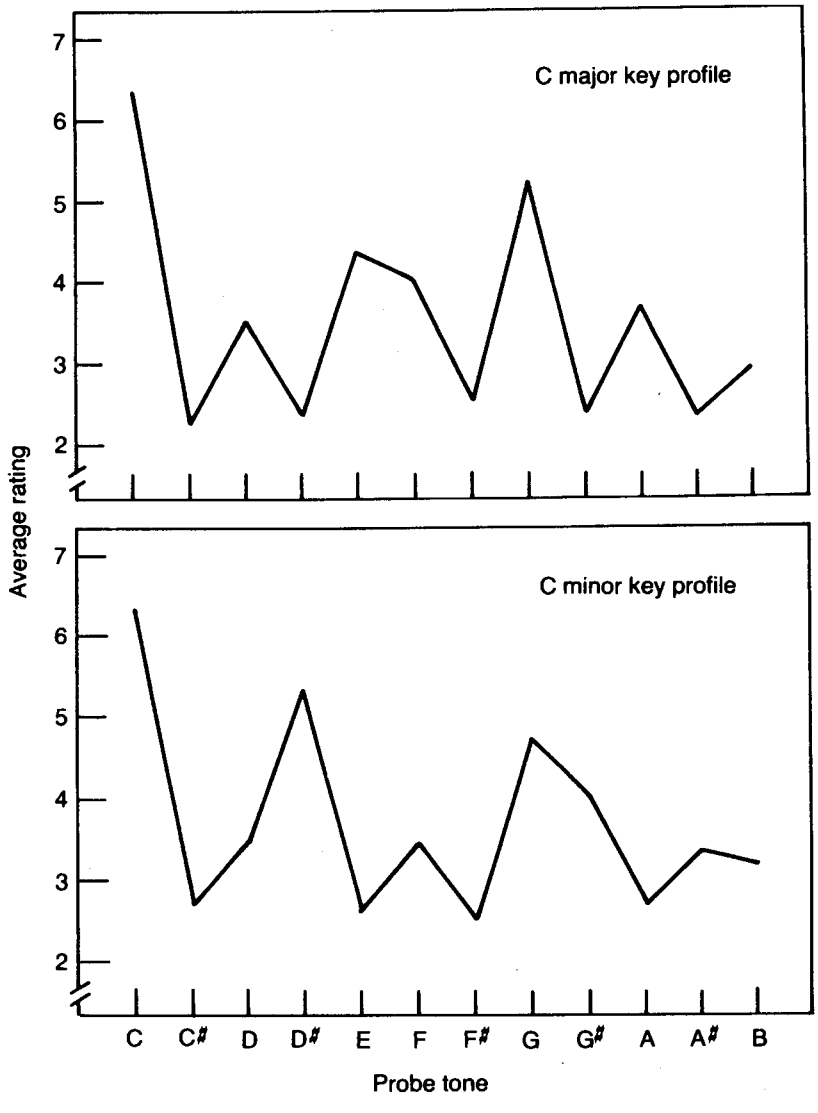



**Figure 6.1** The major and minor key profiles obtained by Krumhansl and Kessler. The profile for the major key (upper graph) is the average rating given each of the twelve notes of the chromatic scale following a tonic triad chord or a cadence in a major key. The minor key profile (lower graph) is averaged over the minor tonic triad and minor cadences. The profiles are shown with respect to C major and minor, respectively. Source: Krumhansl and Kessler 1982. (© 1982 the American Psychological Association, reprinted by permission)

minor third is a less consonant interval than the major third, yet the minor third degree of the scale is judged more fitting than the major third in the minor context.

Another sign of the inadequacy of the consonance explanation is the existence of individual differences between listeners based on their degree of musical experience. For instance, Krumhansl and Shepard (1979) tested adults with negligible musical instruction on a pitch-rating task, and found that these subjects based their ratings primarily on pitch height rather than harmonic relatedness. Even more dramatically, Kessler, Hansen and Shepard have undertaken cross-cultural work using similar tasks, comparing the responses of Balinese adults who have rarely heard Western music to the responses of those familiar with Western music (Kessler *et al.* 1984*).

Although there were some similarities in response across groups, the results supported the notion that listeners are more likely to rank notes in terms of tonal relationships when the stimuli are based on scales with which they are familiar, in comparison with those with which they are not familiar.

A third source of evidence against consonance-based explanations is recent work by Krumhansl, Sandell and Sergeant, in which the probe-tone technique was applied to twelve-tone serial music (Krumhansl *et al.* 1987*). In a complex series of experiments, the authors show that listeners familiar with twelve-tone music made judgements that could be predicted from formal constraints of this idiom. For instance, they were more likely to judge a note as fitting the context of a preceding partial tone-row if it had *not* occurred in the context.

These are simply three examples from many which show that important perceptual responses to music are learned as a result of exposure to music (see Hargreaves 1986; Sloboda 1985a). Responses based purely on physical consonance should, arguably, be little affected by experience.

A more promising line of psychological explanation is based on the proposal that musical works of a genre contain within them statistical and sequential regularities which allow listeners to build up relatedness profiles for pitches and chords. For instance, note counts on melodic lines in a corpus of non-modulating tonal music (see Youngblood 1958; Knopoff and Hutchinson 1983; Krumhansl 1990) show that the tonic, third and fifth degrees of the scale are the most frequently repeated notes, followed by the diatonic set, with chromatic notes occurring least often. The correlation between these counts and the ratings in Krumhansl's own experiments is high (over 0.85). When counts include all notes in a piece (not just the melodic lines) and accord notes a weighting according to their duration (e.g. Hughes 1977, also cited by Krumhansl 1990), then the correlation becomes almost perfect (0.969).

However, the demonstration that such regularities exist in the music is not a demonstration that the listener can necessarily pick them up during normal music listening, nor does it lead to a proposal for a mechanism for doing so. A logical next step in this research would be the construction of a computer program that embodies a proposed learning mechanism, is then exposed to a large body of music in a particular idiom, such as Western tonal music or Balinese music, and ends up responding to Krumhansl-like tasks as appropriate human listeners respond. A principal problem for such a mechanism would be how to store the information in such a way that data from pieces in different keys could be combined. If the combination were to be done on grounds of absolute pitch, then the pre-eminence of particular pitches would probably be destroyed. If the combination were to be done relative to the tonic of each piece, then a mechanism is required for determining what is the tonic. If that mechanism itself depended on the outcome of a note frequency count, then it is not clear that circularity could be avoided. The issue of tonic determination is examined in more detail below (see section 6.4).

### 6.3.2 Rating of chord pairs in a key context

A different set of music theory statements evaluated by Krumhansl concerns the relationship of chords and keys. Within a key, traditional music theory proposes that the chords of IV and V are particularly closely related to I (because of their closeness in the cycle of fifths, and their shared notes). An experiment by Bharucha and Krumhansl (1983) verified such statements by asking listeners to rate how closely related a pair of chords sounded when preceded by an orienting tonal context. The context could be either C major or F# major, and the pair of chords could also be either chords based on the C major or F# major scale. In order to avoid complications raised by inversions, voice leading, or pitch height (see Deutsch and Boulanger 1984), the chords were computer generated according to a technique first developed by Shepard (1964). Each pitch of the chord was duplicated at each audible octave, with a tailing-off in amplitude towards each edge of the audible range. This effectively destroys pitch height information, leaving 'pure' harmony as the only discriminable factor.

Within the experiment, listeners were exposed to every possible combination of pairs of chords within each key, and preceded by each key context. This allowed the application of two particularly powerful statistical techniques to the data: multi-dimensional scaling (Kruskal 1964) and hierarchical cluster analysis (Johnson 1967). These provide alternative representations of the psychological closeness of each item of a set to each other item. Figure 6.2 shows the basic results of such analyses applied to the chords of a single key within that key context. The left-hand box

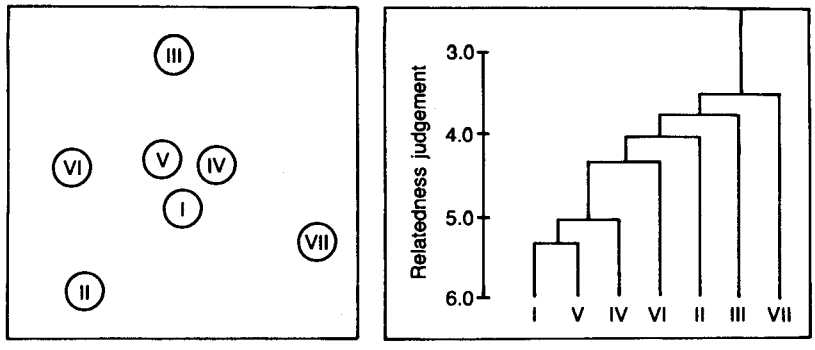



**Figure 6.2** The multidimensional scaling solution (left panel) and the hierarchical clustering solution (right panel) of the relatedness ratings of chords in the basic set of harmonies of a key (I–VII). Source: Krumhansl *et al.* 1982b. (© 1982 the American Psychological Association, reprinted by permission)

shows a two-dimensional scaling solution where relatedness is given by spatial proximity. The right-hand box shows a cluster analysis, where relatedness is given by the level at which the branches for two chords meet. The lower the level, the higher the relatedness. Both analyses yield very similar pictures, confirming the high relatedness of I, IV and V within a key context.

One of the most important demonstrations of Bharucha and Krumhansl's study (1983) is that the perception of such relatedness is fragile and dependent on an appropriate key context. When chord pairs in C major were preceded by a C minor context, relationships like that in Figure 6.2 were discovered. When, however, the same chord pairs were preceded by an F# major context, their perceived relatedness dropped dramatically, and the privileged relation of I, IV and V disappeared. In informal musical language, the sudden key shift caused listeners to 'lose their tonal bearings' and be unable to interpret the new chords in a consistent tonal framework. The extent of this disruptive effect turns out to be directly related to key distance as measured round the cycle of fifths. A similar experiment by Krumhansl, Bharucha and Castellano, where chords in C major were preceded by a G major context, showed almost no disruption of the pattern obtained in Figure 6.2 (Krumhansl *et al.* 1982a). Listeners are able to retain tonal orientation provided that key shifts are small steps round the cycle of fifths. In such small steps, chords and notes are shared between the two keys, and these may perform some pivotal function (see also Lerdahl 1988*).

The key-distance effect is relatively robust, and can be found using several different tasks. An example of this is a study by Roberts and Shaw (1984*). They asked their subjects to judge whether or not two chords were of the

same type (major, minor, diminished or augmented). Judgement was easiest when the two chords were from keys close to one another on the circle of fifths, although this effect was only apparent for highly trained musicians.

Similar effects were also found in a series of experiments reported by Bharucha (1987*) where, for instance, judgements about the intonation of a probe chord were fastest when preceded by a chord context from a relatively close key. Bharucha uses such data to construct a connectionist model of harmony, where notes, chords and keys are conceived of as nodes in a hypothetical 'neural' network. A web of connections between nodes means that when one node is activated, activation spreads through the network to nearby nodes. Such activation forms the basis of tonal expectation, and the model accounts neatly for data from a number of different sources. It is, however, a model for the hypothesized nodes and connections of an experienced listener. A satisfactory account of how such a network is developed from a primitive starting point has yet to be constructed. Also, as Bharucha is quick to point out, the model does not yet incorporate non-harmonic influences (e.g. voice leading) or activation patterns caused by the recollection of specific pieces of music. What is particularly important about Bharucha's work is the application of connectionist models to music. Connectionism is having a profound effect on theorizing in many other domains of psychology, and even though intense debate still rages about the ultimate usefulness of the approach, it is good to know that music is as good a testing ground for its assumptions as any other domain of cognitive functioning.

### 6.3.3 Limitations of Krumhansl's approach

The experimental approach outlined above has many merits from the scientific point of view. It provides rigorous quantified answers to questions of deep theoretical interest, and shows that the precision of a basically psychophysical approach need not be limited to single isolated stimuli, but can be used to study musically interesting phenomena such as the effects of key context.

None the less, like all single approaches, it has difficulties and shortcomings. A major problem is the unusualness and unspecificity of the judgements required from listeners. Most normal musical listening does not require overt judgements of the similarity or relatedness of simple events, and it is possible that such judgements may be dissociated from deeper and more automatic modes of musical processing. In addition, the instructions allow listeners a great deal of freedom to choose what dimensions of the sound they will include in their judgements of relatedness. This freedom may be exercised in arbitrary, or highly task-dependent ways. This consideration becomes particularly important when different groups

of listeners (by age or culture) are examined. Differences in response may be due to differences in understanding of what is being asked, rather than in the organization of musical perception. This second problem is to a certain extent reduced as stable and consistent patterns emerge from the application of this technique across a wide variety of people on a wide variety of different tasks.

A different kind of objection has recently been raised by Butler (1989). He suggests that the ratings of a single tone in a tonal context might be an artefact of the specific context stimuli used in Krumhansl's experiments. The note profile of Figure 6.1 closely resembles a profile derived from the relative frequency of each note in the immediately preceding contexts. It is, therefore, possible that Krumhansl's results derive from short-term activation of specific notes rather than from an enduring internalized tonal schema. Butler's criticisms clearly merit examination and empirical substantiation. Whether or not the criticisms turn out to be well-founded they point to the dangers of relying too heavily on any one methodology as a basis for strong theoretical claims. The best validation of any theory is the discovery of converging evidence from different methods of enquiry. In this case, such evidence does exist, particularly from the study of musical memory tasks.

There are a large number of studies that have used a memory paradigm to investigate structural factors in music perception. Most of these studies have used recognition memory, in which listeners are required to say whether two musical sequences are the same or different. Generally these studies show that different sequences are most likely to be judged the same when the substituted item is close to the original along one of the metrics defined by Krumhansl's data. So, for instance, listeners detect a diatonic substitution (for example, E for G in C major) in a melody less well than a chromatic substitution (for example, F# for G in C major) (Cuddy *et al.* 1979). Similar results have been obtained for chord sequences (Bharucha and Krumhansl 1983).

A final shortcoming of Krumhansl's approach which relates particularly to the multidimensional scaling and cluster analysis techniques is that one has to test every possible pair of items in a set. This makes for a combinatorial explosion as the set size increases, and makes it very difficult to run experiments on large sets or on long musical passages. The issue of length also places limits on what can be included in recognition memory experiments as well. An alternative, but less well used technique is recall or reproduction, where listeners hear a musical passage and are asked to sing or play it back. Where it has been tried (e.g. Sloboda and Parker 1985), it has yielded data that are congruent with the recognition studies, but that display a variety of

richer features that cannot be captured by the ecologically sparse response categories of the more typical experiment. In addition, such tasks are more musically natural, resembling, as they do, the way in which much music is actually learned.

There has, in fact, been a valuable interplay between those who start from the basis that strict experimental controls are paramount and those who start from the assumption that musical realism is paramount. On the first count, it is notable how Krumhansl's most recent work has moved towards greater musical realism (e.g. Palmer and Krumhansl 1987). That particular study used Bach fugue subjects as its stimulus materials and subjected them to ingenious experimental manipulations. On the second count, it is apparent that work based on realistic situations (such as performance) has become increasingly sophisticated in terms of the controls and theoretical models applied to the situation. This progression can be seen by comparing a relatively early study such as that by Gabrielsson (1974), which simply measures timing variations in free performance and attempts to induce some underlying regularities, with studies by Clarke (1988) and Sloboda (1983, 1985b) which manipulate performance materials in controlled ways in order to test specific theoretical ideas. This marrying of scientific control with musical realism is one of the factors which has led me to suggest (Sloboda 1986) that the psychology of music as a scientific discipline has now 'come of age'. Clearly, however, the tensions between these conflicting requirements are still strong, and will continue to act as a spur to new developments.

## 6.4 How do listeners establish a tonal centre?

The work of Krumhansl and others has shown that, with a suitable tonal context, listeners rate musical stimuli with respect to that context in accordance with many of the traditional tenets of music theory. Several investigators have sought to examine a further question: what is it that determines whether a context will or will not establish a strong sense of key or tonal centre? This is an immensely important issue for all tonal music because arguably listeners must establish the key of a piece within the first few bars (if not notes) if they are to code and store the musical material appropriately.

Hantz (1984*) raises this issue from the perspective of a music theorist, and suggests that some psychologists' views of what constitutes a tonal sequence are too naive. In particular, it is incorrect to suppose that a pitch sequence made up of notes from a single diatonic set will necessarily have strong tonal implications. The converse is also true: sequences with chromatic notes may have very strong tonal implications.

Butler and Brown (1984*) have suggested that one important factor facilitating identification of tonal centres is the detection of rare pitch intervals within the diatonic set. The rarest interval is, of course, the tritone (e.g. C–F#). Any given tritone can be found only in two major keys (in the case of C–F#, the keys are G major and C# major). Any added third note will specify a unique key (since the scales concerned share no other notes than the particular tritone). In contrast, a common interval like a perfect fourth (e.g. C–F#) can be found in five keys (C, F, B<flat>, E<flat>, A<flat> and D<flat>), and most additional single notes will not eliminate the ambiguity completely. Butler and Brown verified this hypothesis in an experiment where musicians were asked to listen to a three-note sequence and specify a tonic for the sequence. Where the sequence included a tritone, listeners consistently agreed on the correct tonic (which was not necessarily present in the stimulus sequence). Where the sequence did not contain a tritone, listeners did not agree, even when one of the three notes present was a possible candidate for tonic.

The order in which notes appear is also important, a finding confirmed and extended by Brown (1988*). She took short extracts from classical compositions which contained a variety of pitch sets, ranging from those containing too few pitches to specify uniquely a key ('unspecific'), through those uniquely specifying a single key ('specific'), to those containing pitch sets that could not be contained within a single key ('incongruent'). Examples of each type of extract are given in Figures 6.3, 6.4 and 6.5. In one part of the experiment, subjects heard the untransformed extracts and were asked to identify the tonic. Not surprisingly, there was greatest agreement in the case of the specific extracts.

Brown then constructed tone sequences—three for each extract—which contained one (and only one) of each of the pitch classes contained in each extract (see Figures 6.3–6.5). The A sequences were designed strongly to evoke the tonal centre of the corresponding music excerpt; B sequences were designed to evoke some tonal centre other than that of the corresponding excerpt. In both cases this was done by making the rarest interval from the diatonic set of the chosen key prominent in the sequence and supporting it with other notes that could be interpreted in the same key. This can be seen clearly in Figure 6.5, where the A sequence makes the B<flat>—E tritone prominent, while the B sequence makes the G—C# tritone prominent. The C sequences were designed to be ambiguous by obscuring and de-emphasizing rare intervals. The figures beneath the notes show the percentage of listeners choosing each note as tonic. In each case, the results showed strong effects of order on the choice of tonic. In the case of the 'specific'

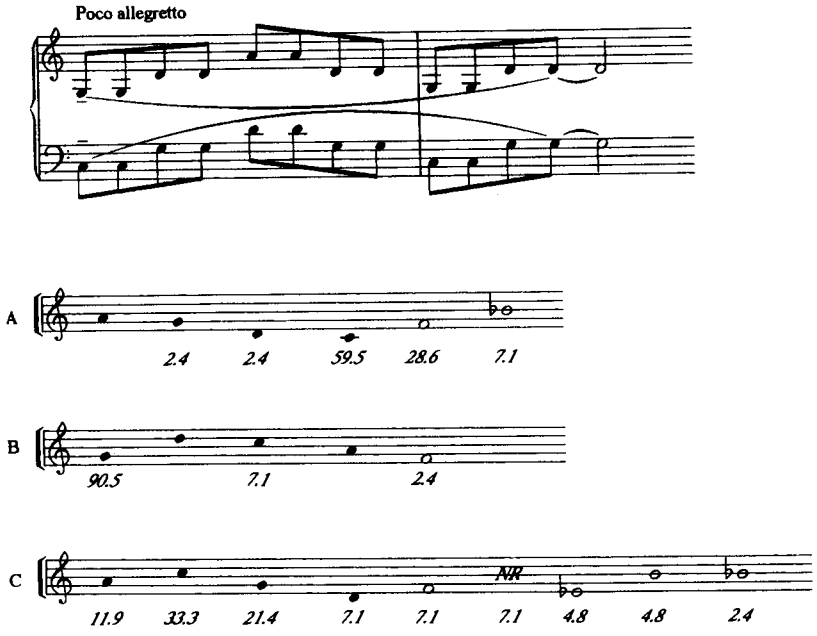



**Figure 6.3** Sample of Brown's stimuli with 'Unspecific' tonal content (Bartok: 85 pieces without octaves for children, 'Romance', mm. 8–12). Response percentages are shown for pitch strings (A) ordered to promote C major, local tonic of musical excerpt) (B) ordered to promote key of G major, and (C) ordered to promote tonal ambiguity. Source: Brown 1988. (© 1988 the Regents of the University of California, reprinted by permission)

extracts such as Figure 6.4, Brown's A sequences actually provoked more unanimity amongst listeners than did the original extract.

The work of Brown and Butler is a salutary reminder of the dangerous gap that can develop between theoretical formalisms and what listeners actually do. Listeners operate in real time and are subject to memory and other processing limitations. Many influential theories of musical structure are time—independent and assume no processing limitations. At a theoretical level, a pitch sequence may be tonally unambiguous. As a listener processes it from left to right, it may appear locally ambiguous, or even unambiguous in a different key. Such local perceptions may (and arguably often do) colour the perception of the whole passage.

Related work reported by Cross, West and Howell confirms this general point in a rather different way by demonstrating that the effects obtained also depend crucially on the task required of listeners (Cross *et al.* 1985*). In one of their experiments, listeners were asked to detect 'wrong notes'

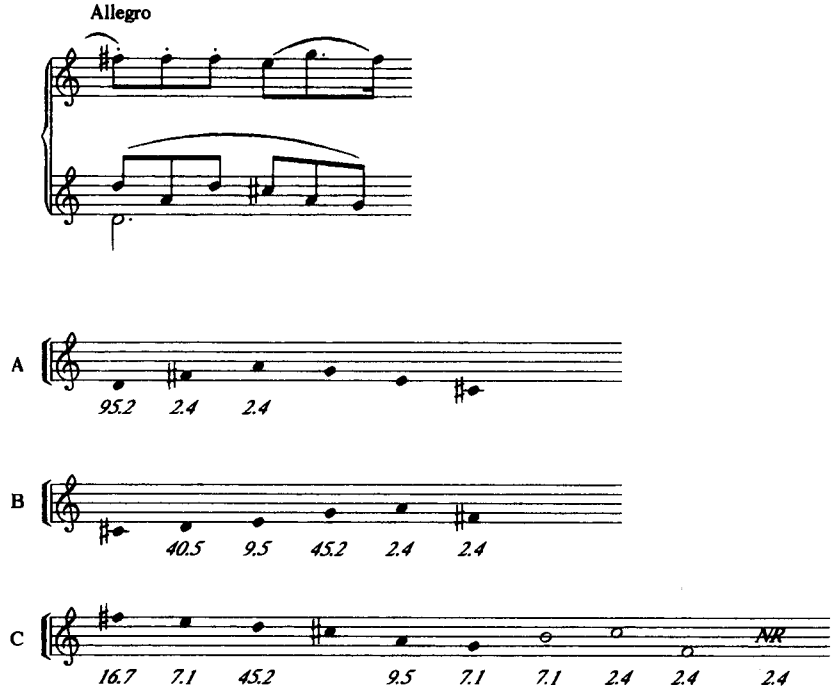



**Figure 6.4** Sample of Brown's stimuli with 'Specific' tonal content (Schubert: Sonata, D.664 (III), m. 49). Response percentages are shown for pitch strings (A) ordered to promote D major, local tonic of musical excerpt, (B) ordered to promote key of G major, and (C) ordered to promote tonal ambiguity. Source: Brown 1988. (© 1988 the Regents of the University of California, reprinted by permission)

in sequences. Wrong notes were notes that could not be contained in the scale(s) implied by the previous notes of the sequence. They found that wrong-note detection was better after sequences such as C–F–G than after sequences containing the tritone, such as B–F–G. On first examination this seems to contradict Brown and Butler. Surely the tritone B–F creates the strongest key sense, making the wrong note 'stand out' most strongly. Cross *et al.* argue convincingly that the better detection of a wrong note in the first sequence is a function of its uniqueness. There is only one note, F#, that is incompatible with any major scale that contains C–F–G. There are five possible wrong notes with respect to the B–F–G sequence (C#, D#, F#, G# and A#). So an F# in the first sequence sounds more wrong than an F# in the second sequence.

Salutary as the work of Brown and Butler is, it also has some puzzling aspects. A straightforward application of their findings suggests that tonal music compositions should contain the fa–te tritone (to use sol-fa terms) in

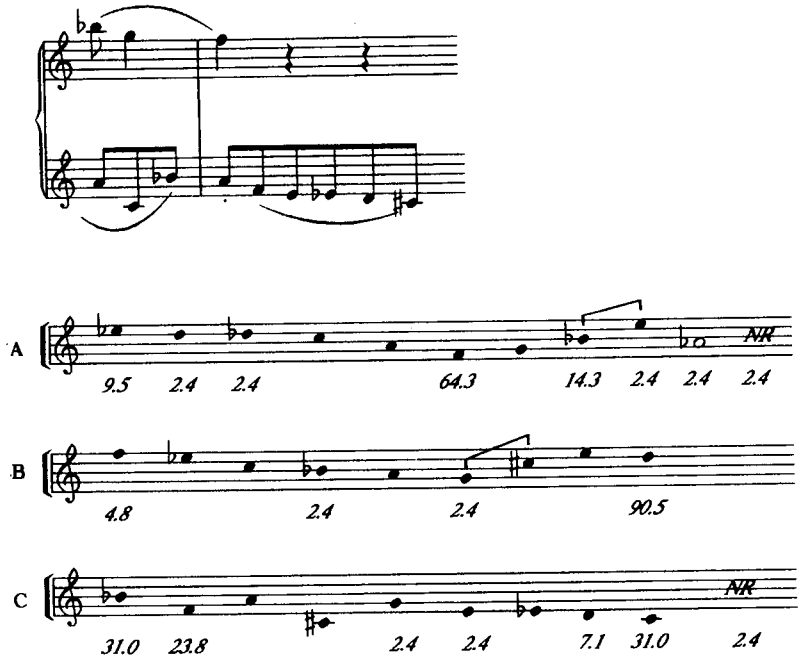



**Figure 6.5** Sample of Brown's stimuli with 'Incongruent' tonal content (Mozart: Sonata, K.570 (I), m. 49). Response percentages are shown for pitch strings (A) ordered to promote F major, local tonic of musical excerpt, (B) ordered to promote key of D major, and (C) ordered to promote tonal ambiguity. Source: Brown 1988. (© 1988 the Regents of the University of California, reprinted by permission)

the first few notes if listeners are to identify the key most efficiently. The reality is that almost no tonal pieces start in this way, yet listeners are still able to identify the key extremely well. Figure 6.6 shows the opening bars of Mozart's piano sonata K.576 in D major. Although I do not have experimental data, I would be very surprised if a large majority of listeners did not identify the tonic as D, even though the tritone is absent. Although it is theoretically possible that the piece is in G major or A major, it is psychologically impossible for me to hear it that way. It has to be in D.

A number of investigators have attempted to tackle the issue of key identification from a purely theoretical perspective by constructing algorithms

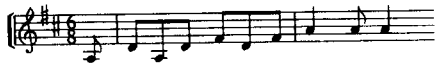


**Figure 6.6** Mozart: Sonata, K.576 (1), bars 1–2.

that take the early notes of a piece as input, and produce candidate keys as output (e.g. Longuet-Higgins and Steedman 1971; Krumhansl 1990). These algorithms work from left to right through a sequence, and their output is therefore sensitive to pitch order. The Longuet-Higgins and Steedman algorithm takes single-line melodies as input and discards all information except pitch-class order. As it moves from left to right through the melody, it successively eliminates keys within which the notes heard so far cannot be contained. When this process fails to come up with a unique key (or eliminates all keys) there is a fall-back procedure that looks at the very first note of the sequence to see if it is the tonic of one of the candidate keys remaining (or eliminated last). If not, the first note is assumed to be the dominant of the correct key. This algorithm assigns the correct key to all forty-eight of the Bach fugue subjects from the *Well-Tempered Clavier*. It is apparent straight away that this algorithm will be greatly helped by encountering a tritone early in the sequence, thus supporting Brown and Butler's data. It would, however, fail to reach a conclusion concerning Figure 6.6 above.

Krumhansl's algorithm combines pitch-class information with durational information so as to compute a running duration total for each note of the chromatic set. This yields a duration profile which is correlated with the prominence profile for each major and minor key obtained in the Krumhansl and Kessler (1982) study (see Figure 6.1 above). The algorithm does not yield a unique key, but assigns to each of the twenty-four keys a relative likelihood rating, which can be updated from note to note, as the piece progresses. It is clear from Figure 6.1 that this algorithm will be helped most by early emphasis on the primary harmonic notes of the scale, particularly tonic and dominant. By note 2 of Figure 6.6, the correlations to D major and D minor should be well above all other values, with note 5 (F#) pushing D major further ahead. The system also handles the tritone reasonably well, not because a tritone produces high correlations to the two candidate keys, but because it produces even lower correlations for all other keys. If we suppose that it is not the absolute level of correlation that counts psychologically, but relative differences in correlation across the twenty-four key 'monitors', with keys that 'stand out' from the others being considered as candidates, then Krumhansl's algorithm appears to be a much more satisfactory one than that of Longuet-Higgins and Steedman.

Neither algorithm, however, is entirely plausible as a psychological mechanism. The Longuet-Higgins and Steedman algorithm places too great a load on short-term memory, since the system must hold the first note of the piece together with a list of keys eliminated (or not eliminated).

The memory of the first note will be particularly important when the sequence does not establish an unambiguous tonal centre early on. But, as Deutsch (1970) and others have demonstrated, memory for pitch is most fragile precisely in that situation, and is unlikely to survive more than a handful of intervening notes.

Similarly, the memory of the Krumhansl algorithm is probably too good. If it accumulates all information about pitch frequency from the beginning of a piece, then it will accumulate great tonal inertia, and will not hear a modulation as establishing a new key until the new notes have been heard as often as the ones left behind. However, this problem could probably be quite easily 'fixed' by building in a time-decay function to the pitch-duration counters. My intuition is that the best decay function would be exponential, with a rapid early decay, but a lingering residual. This would allow for reasonably rapid establishment of new keys after modulations, but would allow for an advantage of returning to the 'home' key after a tonal excursion.

A more fundamental shortcoming of all such algorithms is that they operate entirely without respect to other structural cues, such as rhythm and metre, dynamics, articulation and timbre. It is evident that these factors interact in the psychological determination of key. For instance a note placed on a metrical accent or played at a greater intensity than its neighbours is more likely to be considered as a candidate for tonic than an unaccented passing note.

## 6.5 Hierarchies as psychological structures for the representation of music

One of the reasons why music appears interesting and comprehensible is that it contains various kinds of structure that the human mind is capable of apprehending. It has often been claimed that making sense of music *is* the process of discovering and representing its structure. An important class of such structures is the class of hierarchical structures. Rather than attempt a definition of hierarchies, I will give examples of two major types of hierarchic structure which I will call *reductional* hierarchies and *grouping* hierarchies. These are not exclusive categories. A given hierarchy can be of both types, but it is difficult to imagine a hierarchical structure that is of neither type.

A good example of a reductional hierarchy is that proposed by Deutsch and Feroe (1981—the most cited reference in *Music Perception*). In Lerdahl's adaptation of their model (Lerdahl 1988*) the chromatic set of pitches is represented as a five-level hierarchy (see Figure 6.7). Each level describes

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| level 1: | C | | | | | | | G | | | | | (C) |
| level 2: | C | | | | E | | | G | | | | | (C) |
| level 3: | C | | D | | E | F | | G | | A | | B | (C) |
| level 4: | C | Db | D | Eb | E | F | F# | G | Ab | A | Bb | B | (C) |
| level 5: | C | | | | | | | | | | | | |

**Figure 6.7** The pitch space orientated towards the tonic chord of C major as postulated by Lerdahl. Source: Lerdahl 1988. (© 1988 the Regents of the University of California, reprinted by permission)

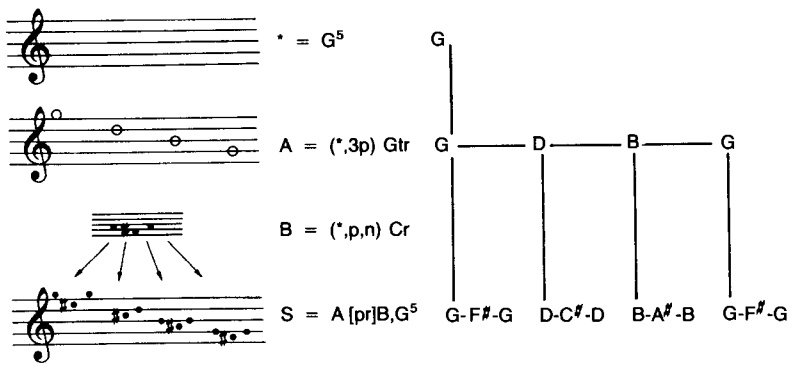
an 'alphabet' (i.e. an ordered functional pitch set). Level 1 contains notes related to one another at the octave; level 2 adds the open fifth space; level 3 adds the notes of a diatonic triad; level 4 adds the remaining notes of the diatonic set (in this example the major diatonic scale), and level 5 adds the remaining chromatic notes.

What makes this a reductional hierarchy is the fact that as one progresses from lower to higher levels, elements are dropped from the representation. Only the most 'important' or 'salient' elements survive to higher levels. In this model, the alphabets are defined culturally, and so may be different for different groups of people or types of music.

A major psychological prediction of the Deutsch and Feroe model is that sequences that can be represented in terms of stepwise movements up or down any (or several) of these alphabets will be easier to represent (and thus remember) than other types of sequence. Such movements tend to produce repeating patterns of a type very common in classical tonal music, and experimental evidence (e.g. Deutsch 1980) supports the prediction that such sequences are easy to remember. Figure 6.8 gives one such sequence, and shows how it can be represented as a chromatic turn (one step down and one step up at level 5) repeated at four successive descending positions at level 3.

Lerdahl (1988) also points out that the model accounts for some of the implicative characteristics of melody suggested by Meyer (1973). A stepwise movement at any level of the hierarchy seems to create an implication for further movement at that level. The implication is strongest when the initial movement is on to a note that is not represented at the next highest level. For instance, in C major, a move from C to D strongly suggests a movement to E, because D is not contained in level 3, but E is. The implication for further movement to F is not so strong, because E serves as a structural resolution to the sequence.

An example of a grouping hierarchy is provided by Lerdahl and Jackendoff (1983, 1984). In their theory, which will be discussed more fully below, it is claimed that an experienced listener parses a musical sequence into non-overlapping groups of events, which are given such standard musical names




Key:

| | |
|---|---|
| * | reference note; |
| p | downward movement of one step within an alphabet; |
| n | upward movement of one step within an alphabet; |
| Gtr | alphabet of the G major triad; |
| Cr | alphabet of the choromatic scales; |
| A[pr]B | the application of pattern B to each element of pattern A. |

**Figure 6.8** Formal representation of a hierarchically organized tone sequence after Deutsch and Feroe. Source: Sloboda 1985a. (© 1985 Oxford University Press, reprinted by permission)

as *motifs*, *phrases*, and *sections*. A specific set of grouping preference rules is provided, which is claimed to capture intuitions about the musical features that incline a listener to place notes together in the same group, or to place a group boundary between them. An example of a grouping hierarchy produced by this theory is given in Figure 6.9.

The grouping mechanism produces a hierarchy because the groups formed at a basic level are then grouped into 'groups of groups', and so on, until the final level is reached where the whole piece is a single group. Unlike a reductional hierarchy, the grouping hierarchy includes all elements at every level. What changes is their degree of relatedness. At the top level, every element is related to every other element in the same way by being a member of the same group (the piece). At lower levels an element is most strongly related to other elements in its own group, progressively less related to elements in more distant groups.

We will consider the psychological evidence for this grouping hierarchy below. But first, we should note the contrast between the two hierarchies just described. The reductional hierarchy of Deutsch and Feroe (1981) is not a grouping hierarchy. For instance the relationship of C# to its neighbours C and D is not specified in the hierarchy. It forms a group with neither of them. Lerdahl and Jackendoff's grouping hierarchy is not reductional.

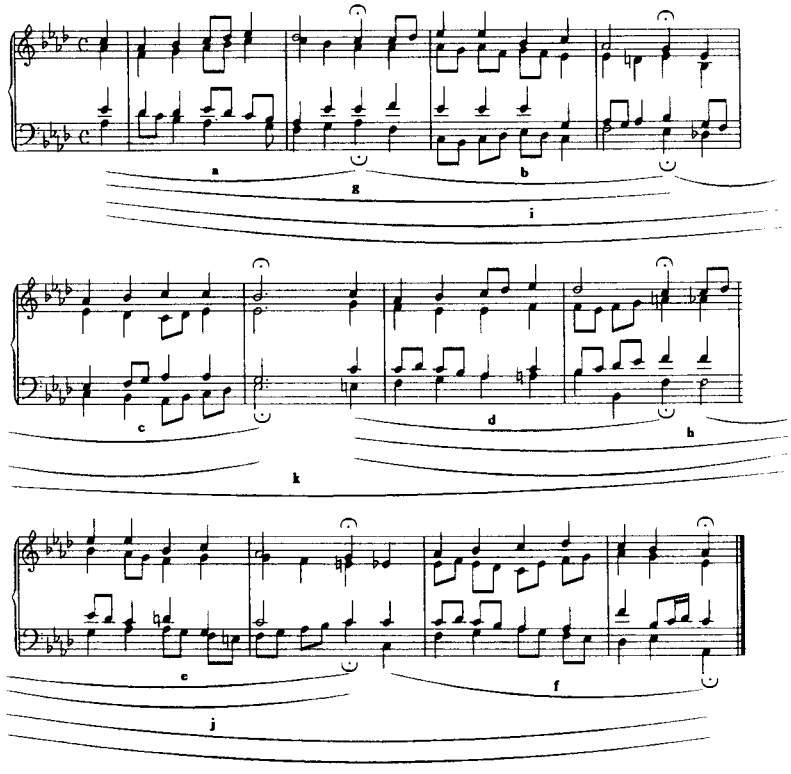



**Figure 6.9** The grouping structure of Bach: 'Ich bin's, ich soilte bussen' from *St Matthew Passion*. Source: Lerdahl and Jackendoff 1984. (© 1984 the Regents of the University of California, reprinted by permission)

It does not nominate a sub-set of elements from a lower level to be given prominence at a higher level. Some hierarchies, however, can be both reductional and grouping. In these cases, groups are formed, but a single item from each group is nominated as the head member of the group, and goes forward to form an element at the next highest level. An example of such a hierarchy is the *time-span reduction* of Lerdahl and Jackendoff (1983, 1984).

Hierarchies are very appealing to human beings. The vast majority of social structures depend upon them, and many of our theories of how the world is and works are couched in hierarchic terms. Many children enjoy writing their address as, for instance, 35 Any Street, Newtown, Overshire, England, Europe, The World, The Solar System, The Milky Way, The Universe. In doing so they are building a grouping hierarchy of physical

locations. The command chain of any organization usually represents a reductional hierarchy.

Because hierarchies are so appealing they bring with them a danger. People tend to impose hierarchical explanations to an extent that may be inappropriate. In particular, what works well at a local level may not necessarily translate to higher levels, and vice versa. There is a satisfying elegance to complete hierarchic descriptions of entire pieces of music (such as those supplied by Schenker (1979) and Lerdahl and Jackendoff (1983)), but many observers, including psychologists, have pointed out that the ability to describe something in hierarchic terms does not necessarily mean that listeners represent what they hear according to such descriptions (see, for instance, Rosner 1984).

### 6.5.1 The psychological status of Lerdahl and Jackendoff's theory

Lerdahl and Jackendoff's generative theory of tonal music (Lerdahl and Jackendoff 1983)—henceforth GTTM—is the most ambitious attempt to date at a complete theory of what a listener knows when he or she knows a well-formed piece of tonal music. In the words of Lerdahl and Jackendoff (1984) the theory

> takes the form of explicit rules that assign, or 'generate', heard structures from musical surfaces. By 'musical surface' we mean, broadly, the physical signal of a piece when it is played. By 'heard structure' we mean all the structure a listener unconsciously infers when he listens to and understands a piece, above and beyond the data of the physical signal.

GTTM has a number of self-confessed limitations. It represents the supposed understanding of an experienced listener who knows a piece well. It does not attempt to construct a model of how such a representation is built up in real time and over repeated hearings. As justification for this limitation the authors appeal to Marr (1981) and others, who argue that one cannot sensibly investigate mechanism until one knows the endpoint that the mechanism is meant to achieve. A second limitation is its failure to deal with non-hierarchic aspects of music, or polyphonic counterpoint. GTTM is unique in a number of ways. It goes beyond the reductional approach of Schenker in specifying explicit rules for moving from lower levels of the hierarchy to higher ones. It also captures the intuition that musical awareness is multidemensional. GTTM proposes not one hierarchical reduction but four, which proceed independently, although they

interact at higher levels. In addition to *grouping structure* mentioned above, there is *metrical structure*, *time-span reduction*, and *prolongational reduction*. Metrical structure is the hierarchy of repeating patterns of strong and weak beats. GTTM proposes that metrical structure is only extracted up to intermediate levels, spanning a few seconds. This accords with the intuition that we do not hear large-scale timing regularities as metre. Time-span reduction creates a tree-like hierarchy for the whole piece. At each level, the most prominent element (defined metrically or harmonically) is chosen as the head element to be represented at the next higher level. Such reduction often ends with the Schenkerian *Ursatz* (I–V–I) at the higher level. Prolongational reduction is analogous to time-span reduction, but attempts to capture the flux of harmonic tension and resolution within a piece.

GTTM is also unique in proposing that most of the rules determining the formation of these hierarachies are *preference* rules. They do not specify mandatory decisions, but incline the decision in one way or another. Grouping rules may conflict with one another. In such situations, the music may be heard as ambiguous. Lerdahl and Jackendoff arrange the preference rules in a rough weighting order by specifying that some preferences are 'strong' and others are 'weak', but they do not attempt to quantify these rules by assigning numbers to the weights. They do, however, present the preference rules as empirically testable, and explicitly invite experimental verification.

This challenge has been taken up by a number of researchers, notably Deliège (1987*) and Todd (1985*). Deliêge focused on the local Grouping Preference Rules (see Figure 6.10) by asking musician and non-musician subjects to listen to brief musical excerpts and mark the positions of group boundaries. In a first experiment using recorded performances of classical instrumental music, Deliège found that the majority of musicians' choices (77%) were in accordance with GTTM rules. Choices not in accordance with the rules were of several types. For instance, some subjects did not group according to the length rule (Figure 6.10, Rule 6, right-hand example) but included the first long note in the same group as the shorter notes preceding it, suggesting that, in real time, such sequences are likely to be interpreted under proximity rules (Figure 6.10, Rule 2). Other results suggested the need for new rules. For instance, some subjects placed a boundary at the point of change in direction of melodic contour or of underlying harmony.

In a second experiment using constructed nine-note sequences, Deliège introduced features that would lead to conflicting segmentations on two different grouping rules. An example of such a sequence is given in Figure 6.11. Grouping by dynamic difference would produce a segmentation after note 3, but grouping by timbre would produce a segmentation after note 4.

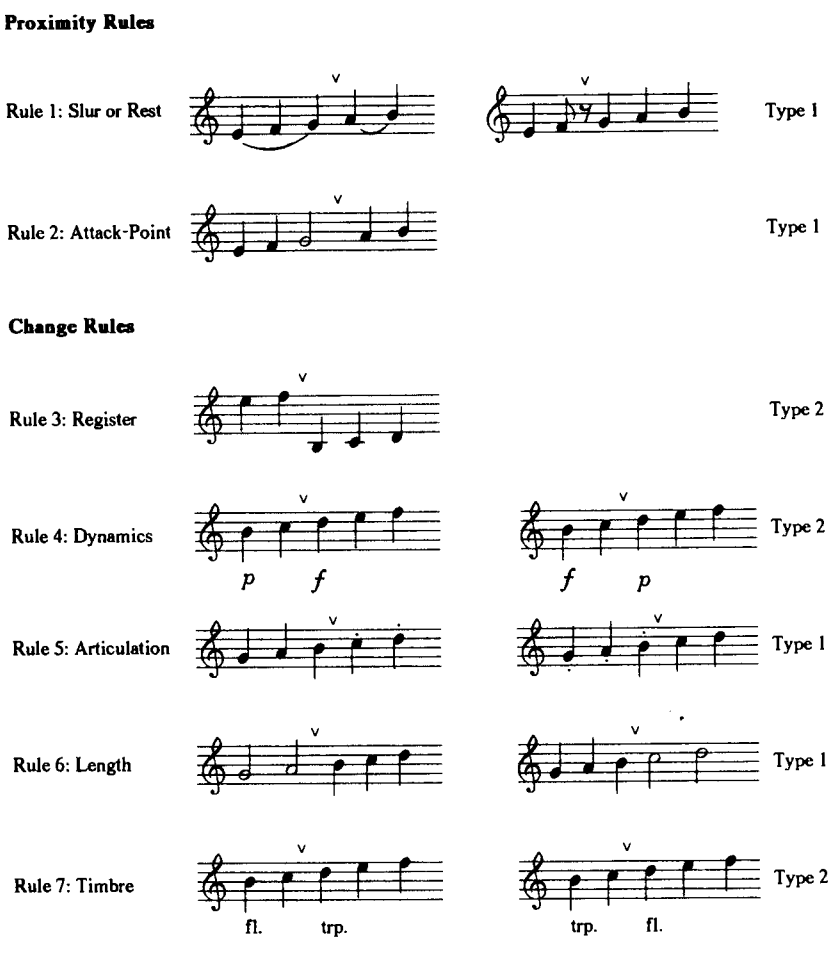



**Figure 6.10** The grouping preference rules of Lerdahi and Jackendoff. Source: Deliège 1987. (© 1987 the Regents of the University of California, reprinted by permission)

Subjects were allowed to choose only one grouping boundary in each sequence. The results were not simple, but were in general accordance with a higher-level grouping rule of GTTM (Symmetry) which says: 'Prefer grouping analyses that most closely approach the ideal subdivision of groups into two parts of equal length'. Subjects in Deliège's experiment showed a tendency to choose the grouping that most nearly divided the sequence into two equal halves. A general problem affecting the interpretation of Deliège's results is that the grouping preference expressed may depend on the relative magnitude of conflicting musical cues (for example,

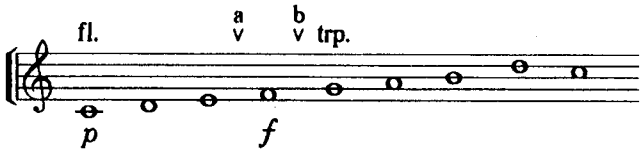



**Figure 6.11** Example of conflict between two rules: (a) segmentation for rule 4 (change in dynamics); (b) segmentation for rule 7 (change in timbre). Source: Deliège 1987. (© 1987 the Regents of the University of California, reprinted by permission)

a large dynamic contrast may 'win' over a small timbral change, but the preference may be reversed if the small timbral contrast is increased).

Todd's work (1985) stems from the now well-established position that expressive variation in musical performance is systematic, and performs the funtion of reinforcing a structural interpretation of the music. For instance, my own work (Sloboda 1983, 1985b) has shown how the local deviations in timing and loudness of pianists assists listeners to extract the intended metre of a musical passage. Shaffer (1981) has shown how such deviations tend to be stable over repeated performances by the same musician, not just at a local level, but also at more global levels, as indicated through the use of rubato or dynamic progressions. Todd focused on the expressive slowing at phrase and segment endings that is a characteristic of the performance of much Classical and Romantic music. His basic argument is that the degree of such slowing should represent the relative importance of that particular boundary in the time-span reduction of GTTM. Boundaries of major units (represented at high levels in the hierarchy) should receive greater amounts of slowing than lower-level boundaries. Todd incorporated these suppositions into a mathematical model, and Figure 6.12 shows the predictions of the model for a sixteen-bar phrase with the time-span reduction given at the top of the figure. The height of the line on the lower graph is an indication of the degree of deviation from tempo. Data from performances of the theme of Mozart's A Major Sonata (K.331) fit the predictions of the model extremely well. Since the time-span reduction of GTTM depends in part on identifying points of greater or lesser harmonic stability, an unambiguous reduction structure is less likely for music where the tonality is not so well defined. This is reflected in data from performances of Chopin's Etude No. 3 from *Trois Nouvelles Etudes*, where slowings are sometimes determined by local harmonic tensions and relaxations (prolongational structure) rather than the time-span reduction. This reflects the tonal instability of the central section of this piece. On the whole, however, the fit to GTTM is remarkably impressive. Todd's work is particularly encouraging because it shows that performance data can be

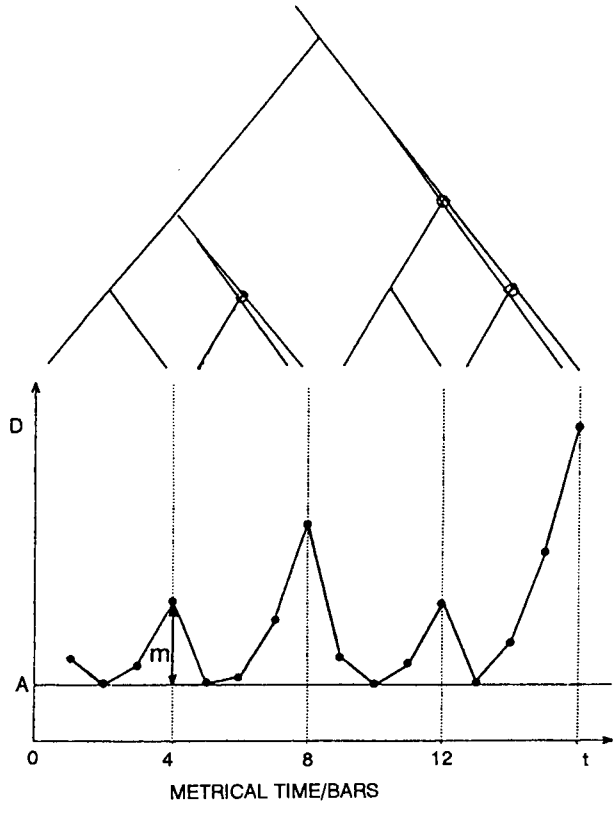



**Figure 6.12** A hypothetical performance duration structure generated by Todd with the given time-span reduction above. Source: Todd 1985. (© 1985 the Regents of the University of California, reprinted by permission)

applied to the larger-scale aspects of GTTM. It is rather difficult to see what kind of listening tests could provide equally compelling evidence of large-scale awareness. Indeed, Rosner (1984) suggests that such levels are available only to musicians who can study the score of a piece, and are therefore represented in the minds of analysts, but not necessarily other listeners.

The amount of empirical work stimulated by GTTM is not yet large. One unexplored area is in respect of Lerdahl and Jackendoff's claim that many of the rules of GTTM are universal (i.e. they embody general facts about human perceptual and cognitive propensities and capacities, not just contingent facts about certain genres). It follows that music of all cultures and periods should be susceptible to analysis by modified versions of GTTM, and that cross-cultural data should support such analysis.

### 6.5.2 Other empirical investigations into structural perceptions

Although there have been few direct empirical tests of GTTM, a number of investigators have taken more general and less precise music-theoretic proposals as the basis for psychological investigation. For instance, Cook (1987*) investigated whether listeners were capable of perceiving large-scale tonal closure within a movement. To this end, he 'adjusted' a number of classical piano pieces so that they ended in a different key to that in which they started. University music students listened to both versions and were asked to rate them on a number of dimensions, including coherence and completion. For pieces lasting less than one minute, students rated the 'correct' version higher, but on longer pieces, there were no consistent differences between pieces. Cook argues that such data are damaging to many theories that attempt to explain the appeal of classical tonal music in terms of the tonal journey which starts and ends in the same key.

Rosner and Meyer (1986*) looked at the influence of melodic form (AA' vs. AA'B) on the perception of musical structure that has generally been held to have perceptual salience. In one of Rosner and Meyer's experiments (Experiment 5) melodies also differed in what the authors call 'melodic process'. This idea, originated by Meyer, 'specifies the principal motions by which a melody leaves its main starting point and finally achieves closure'. For instance a *gap fill* process 'has a large early skip, usually upward, of a fourth or more, followed by a more or less stepwise return to a point of closure'. In contrast an *Adeste Fideles* process, like the Christmas carol after which it is named,

> always involves two characteristic skips. The first spans a fourth and the next one a fifth. Both usually occur in the first half of the melody; the second skip leads to upward motion to the third or fourth of the scale, followed by downward resolution often to the tonic. (Meyer 1973: 19)

Examples of these two types of melody are given in Figure 6.13.

Rosner and Meyer asked listeners to rate all possible pairs of melodies within their experimental set for similarity. This allowed a multidimensional scaling to be produced (of a similar sort used by Krumhansl—see p. 126 above). A difficulty with using multidimensional scaling in this context is that it is not possible to compare all possible pairs in the universe under consideration: the universe is too ill-defined. And so there is a certain amount of 'ad-hoccery' in the choice of stimuli to include. This is in contrast to the studies of Krumhansl, which tend to sample every element of a given universe (for example, the chromatic set, the diatonic chord set).

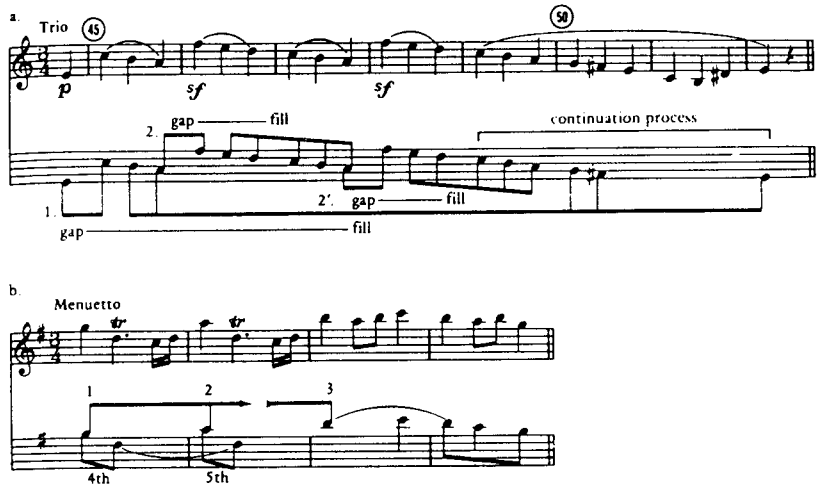



**Figure 6.13** (a) Melody from Beethoven's Piano Sonata in A, op. 2 no. 2 (III) bars 45–52 with gap-fill processes shown in lower staff. (b) Melody from Handel's Flute Sonata in G, Op. 1 no 5 (V) bars 1–4 with Adeste Fideles process indicated on lower staff. Source: Rosner and Meyer 1986. (© 1986 the Regents of the University of California, reprinted by permission)

However, the broad results of this analysis showed that listeners grouped melodies in terms of melodic process but not in terms of form. In all their experiments, however, there were 'rogue' melodies that seemed to be classified on the basis of prominent local events. Rosner and Meyer argue that although analysts typically conduct their analyses in terms of large-scale features of the music, listeners are actually more likely to construct a representation that is primarily based on local features, although they do not rule out the possibility of achieving higher-order representations.

Implicit in the work of Cook and of Rosner and Meyer is the notion that there are processing limits on human cognitive capacity that limit what can be extracted from music, at least on early hearings (Swain 1986 makes this point more explicitly). Their findings do not show that people cannot hear out large-scale structure. They simply show that they do not usually do so in early listening *given the experimental tasks required of them*. This last caveat is important, because competence is sometimes tied quite firmly to particular tasks, and does not show itself in other tasks. With longer exposure and/or different tasks, it is still quite possible that we may find evidence of large-scale psychological representations. At the very least, the existence of unusual individuals who can reproduce large musical structures from

memory after one or two hearings (see Sloboda *et al.* 1985) is a finding to put on the other side of the balance sheet.

The search for appropriate measures of structural awareness has led some researchers down what seem at first to be rather curious avenues. For instance, Stoffer (1985*) asked listeners to detect the location of clicks superimposed on musical sequences. This follows work in language (Fodor and Bever 1965) and music (Gregory 1978; Sloboda and Gregory 1980) showing that clicks are more accurately located when they occur at phrase boundaries. Indeed clicks near to phrase boundaries seem perceptually to migrate towards the boundary. The explanation for this phenomenon is that listeners try to organize their perception so that the extraneous event (the click) falls between the two structural units (the phrases). Stoffer's particular contribution was to extend the technique so that it could be used with non-musicians (the earlier studies relied on the ability to read a score to mark the click location). In his second experiment listeners were simply required to press a button as soon as they heard a click. The prediction was that reaction times would be faster at phrase boundaries since, in some sense, listeners were 'expecting' the click to fall there. This prediction was confirmed for both musicians and non-musicians, although the effects were stronger for musicians. The studies also show that 'simple' explanations of the skill are not sufficient. For instance, one possibility is that listeners simply divide melodies into two equal halves. Stoffer showed that when a phrase boundary occurred after one-third of a passage, the click judgements moved accordingly. Another possibility is that listeners identify phrase boundaries on the basis of some gross physical cue like a long note or a gap. Stoffer ruled out this explanation by making all notes of equal length and intensity. Phrase awareness must have been dependent on melodic or harmonic factors.

What is particularly interesting about this technique is that subjects are not required to make any judgement about the music at all, yet their 'subconscious' analyses reflect structural awareness in the absence of any formal knowledge about phrase structure. In this respect the technique resembles reproduction, but it does not depend (as reproduction does) on the practical musical accomplishment of the experimental subject. We need more 'indirect' measures of this sort to probe structural awareness in ordinary untrained listeners. Evidence from cognitive and perceptual psychology suggests that this 'back door' approach can often yield a richer picture of what is known than a direct judgement would suggest. A dramatic example of this is the work on subliminal perception. A stimulus of which a subject can give no conscious report can none the less measurably affect performance on some other task (Dixon 1981).

## 6.6 Concluding remarks

In this chapter I have attempted to give a fairly detailed picture of some central preoccupations of psychologists of music as represented through a citation-based selection of papers from the journal *Music Perception*. By any account these preoccupations relate to core issues in music. Tonal music is claimed by many analysts to gain its power and appeal from the complex structures that tonality can support, both on a small and a large scale. Psychologically based investigations have demonstrated the psychological reality of some of these structures, particularly those operating on a small scale. The investigations have been more equivocal in relation to larger structures. This may be because such structures reside only in the heads of specialists who analyse and write about music; but it may also be because we have not yet developed measuring techniques capable of capturing large-scale awareness.

Although the topics of this chapter are central, they do not by any means exhaust the psychology of music. The wide range of perceptual and cognitive topics studied by music psychologists is well illustrated in books such as Deutsch (1982), Howell *et al.* (1985), Dowling and Harwood (1986), Sloboda (1985a, 1988). These topics range from the nature of absolute pitch (Ward and Burns 1982), through the measurement of ensemble playing in chamber music (Rasch 1988) to the connections between musical structure and emotional response (Dowling and Harwood 1986). There is also a large body of research on developments and educational aspects of musical skill (e.g. Shuter-Dyson and Gabriel 1981; Hargreaves 1986). Most of the research reported in this chapter relates to adults familiar with a musical idiom. The story of the acquisition of this familiarity over the first years of life is a fascinating one which must be brought together into an overall account of the psychology of musical perception and performance.

Although I began the chapter with an undertaking to confine myself to contemporary achievements rather than hopes for the future, I cannot resist ending without choosing one area that I believe should see a substantial growth throughout the 1990s. This is the study of psychological functions of music. The studies outlined in this chapter show that human beings go to a great deal of cognitive effort to build up complex structured representations of the music they hear. The studies give very few clues, however, as to why people do this, what psychological purposes are fulfilled. Many volumes have, of course, been filled by commentators on music setting forward their own philosophies and hunches. What we have lacked is any really convincing empirical work on, for instance, the links between musical structure and emotional response. Work has started in

one or two places, and my hope is that future work will be able to provide a sketch of an answer to the question that many of us cannot even answer articulately for ourselves: why do we devote so much of our lives to these organized sounds we call music?

*Note:** denotes a core article listed in Table 6.2.

## References

Bharucha, J. (1987) Music cognition and perceptual facilitation: a connectionist framework. *Music Perception*, **5**, 1–30.

Bharucha, J. and Krumhansl, C. L. (1983) The representation of harmonic structure in music: hierarchies of stability as a function of context. *Cognition*, **13**, 63–102.

Brown, H. (1988) The interplay of set content and temporal context in a functional theory of tonality perception. *Music Perception*, **5**, 219–50.

Butler, D. (1989) Describing the perception of tonality in music: a critique of the tonal hierarchy theory, and a proposal for a theory of intervallic rivalry. *Music Perception*, **6**(3), 219–42.

Butler, D. and Brown, H. (1984) Tonal structure versus function: studies of the recognition of harmonic motion. *Music Perception*, **2**, 6–24.

Clarke, E. F. (1988) Generative principles in music performance, in J.A. Sloboda (ed.) *Generative Processes in Music: The Psychology of Performance, Improvisation and Composition*. London: Oxford University Press.

Cook, N. (1987) The perception of large-scale tonal closure. *Music Perception*, **5**, 197–206.

Cross, I., West, R., and Howell, P. (1985) Pitch relations and the formation of scalar structure. *Music Perception*, **2**, 329–44.

Cuddy, L. L., Cohen, A. J., and Miller, J. (1979) Melody recognition: the experimental application of musical rules. *Canadian Journal of Psychology*, **33**, 148–57.

Deliège, I. (1987) Grouping conditions in listening to music: an approach to Lerdahl and Jackendoff's grouping preference rules. *Music Perception*, **4**, 325–60.

Deutsch, D. (1970) Tones and numbers: specificity of interference in short-term memory. *Science*, **168**, 1604–5.

Deutsch, D. (1980) The processing of structured and unstructured tonal sequences. *Perception and Psychophysics*, **28**, 38 1–9.

Deutsch, D. (ed.) (1982) *The Psychology of Music*. New York: Academic Press.

Deutsch, D. and Boulanger, R. C. (1984) Octave equivalence and the immediate recall of pitch sequences. *Music Perception*, **2**, 40–51.

Deutsch, D. and Feroe, J. (1981) The internal representation of pitch sequences in tonal music. *Psychological Review*, **88**, 503–22.

Dixon, N. F. (1981) *Preconscious Processing*. Chichester: Wiley.

Dowling, W. J. (1978) Scale and contour: two components of a theory of memory for melodies. *Psychological Review*, **85**, 341–54.

Dowling, W. J. and Harwood., D.L. (1986) *Music Cognition*. New York: Academic Press.

Fodor, J. A. and Bever, T. G. (1965) The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behaviour*, **4**, 414–20.

Gabrielsson, A. (1974) Performance of rhythm patterns. *Scandinavian Journal of Psychology*, **15**, 63–72.

Gregory, A. H. (1978) Perception of clicks in music. *Perception and Psychophysics*, **24**, 171–4.

Hantz, E. (1984) Studies in musical cognition: comments from a music theorist. *Music Perception*, **2**, 245–64.

Hargreaves., D. J. (1986) *The Developmental Psychology of Music*. London: Cambridge University Press.

Helmholtz, H. von (1954) *On the Sensations of Tone as a Physiological Basis for the Theory of Music* (1863). trans. A. J. Ellis, New York: Dover.

Howell, P., Cross, I., and West, R. (eds) (1985) *Musical Structure and Cognition*. New York: Academic Press.

Hughes., M. (1977) A quantitative analysis, in M. Yeston (ed.). *Readings in Schenker Analysis and Other Approaches*. New Haven: Yale University Press.

Johnson, S. C. (1967) Hierarchical clustering schemes. *Psychometrika*, **32**, 241–54.

Kessler, E. J., Hansen, C. and Shepard, R.N. (1984) Tonal schemata in the perception of music in Bali and in the West. *Music Perception*, **2**, 131–65.

Knopoff, L. and Hutchinson, W. (1983) Entropy as a measure of style: the influence of sample length. *Journal of Music Theory*, **27**, 75–97.

Krumhansl, C. L. (1979) The psychological representation of musical pitch in a musical context. *Cognitive Psychology*, **11**, 346–74.

Krumhansl, C. L. (1983) Perceptual structures for tonal music. *Music Perception*, **1**, 28–62.

Krumhansl, C. L. (1990) *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press.

Krumhansl, C. L. and Kessler, E. (1982) Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, **89**, 334–68.

Krumhansl, C. L. and Shepard, R. N. (1979) Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology. Human Perception and Performance*, **5**, 579–94.

Krumhansl, C. L., Bharucha, J. J., and Castellano, M. A. (1982a) Key distance effects on perceived harmonic structure in music. *Perception and Psychophysics*, **32**, 96–108.

Krumhansl, C. L., Bharucha, J. J., and Kessler, E. J. (1982b) Perceived harmonic structure of chords in three related musical keys. *Journal of Experimental Psychology: Human Perception and Performance*, **8**, 24–36.

Krumhansl, C. L., Sandell, G. J., and Sergeant, D. C. (1987) The perception of tone hierarchies and mirror forms in twelve-tone serial music. *Music Perception*, **5**, 31–78.

Kruskal, J. B. (1964) Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, **29**, 28–42.

Lerdahl, F. (1988) Tonal pitch space. *Music Perception*, **5**, 315–50.

Lerdahl, F. and Jackendoff, R. (1983) *A Generative Theory of Tonal Music*. Cambridge, Mass.: MIT Press.

Lerdahl, F. and Jackendoff, R. (1984) An overview of hierarchical structure in music. *Music Perception*, **1**, 229–52.

Longuet-Higgins, H. C. and Steedman, M. J. (1971) On interpreting Bach. *Machine Intelligence*, **6**, 221–41.

Marr, D. (1981) *Vision*. San Francisco: Freeman.

Meyer, L. B. (1973) *Explaining Music*. Berkeley: University of California Press.

Palmer, C. and Krumhansl, C. L. (1987) Independent temporal and pitch structures in perception of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance*, **13**, 116–26.

Rasch, R. A. (1988) Timing and synchronization in ensemble performance, in J. A. Sloboda (ed.) *Generative Processes in Music: the Psychology of Performance, Improvisation, and Composition*. London: Oxford University Press.

Roberts, L. A. and Shaw, M. L. (1984) Perceived structure of musical triads. *Music Perception*, **2**, 95–124.

Rosner, B. S. (1984) Review of *A Generative Theory of Tonal Music* by Fred Lerdahl and Ray Jackendoff. *Music Perception*, **2**, 275–90.

Rosner, B. S. and Meyer, L. B. (1986) The perceptual roles of melodic process, contour, and form. *Music Perception*, **4**, 1–40.

Schenker, H. (1979) *Free Composition*, trans. E. Oster. New York: Longman, originally published as *Die freie Satz* (1935).

Shaffer., L. H. (1981) Performance of Chopin, Bach, and Bartok: studies in motor programming. *Cognitive Psychology*, **13**, 326–76.

Shepard, R. N. (1964) Circularity in judgements of relative pitch. *Journal of the Acoustical Society of America*, **36**, 2346–53.

Shepard, R. N. (1982) Structural representations of musical pitch. in D. Deutsch (ed.) *Psychology of Music*. New York: Academic Press.

Shuter-Dyson, R. and Gabriel, C. (1981) *The Psychology of Musical Ability*, 2nd edn. London: Methuen.

Sloboda, J. A. (1983) The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology*, **A35**, 377–96.

Sloboda, J. A. (1985a) *The Musical Mind. The Cognitive Psychology of Music.* London: Oxford University Press.

Sloboda, J. A. (1985b) Expressive skill in two pianists: style and effectiveness in music performance. *Canadian Journal of Psychology*, **34**, 274–80.

Sloboda, J. A. (1986) Cognition and real music: the psychology of music comes of age. *Psychologica Belgica*, **26**, 199–219.

Sloboda, J. A. (ed.) (1988) *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition.* London: Oxford University Press.

Sloboda, J. A. and Gregory, A. H. (1980) The psychological reality of musical segments. *Canadian Journal of Psychology*, **34**, 274–80.

Sloboda, J. A. and Parker, D. H. H. (1985) Immediate memory for melodies, in P. Howell, I. Cross, and R. West (eds) *Musical Structure and Cognition.* London: Academic Press.

Sloboda, J. A., Hermelin, B., and O'Connor, N. (1985) An exceptional memory. *Music Perception*, **3**(2), 155–70.

Stoffer, T. H. (1985) Representation of phrase structure in the perception of music. *Music Perception*, **3**, 191–220.

Swain, J. P. (1986) The need for limits in hierarchical theories of music. *Music Perception*, **4**, 121–48.

Todd, N. (1985) A model of expressive timing in tonal music. *Music Perception*, **3**, 33–58.

Ward, W. D. and Burns, E. M. (1982) Absolute pitch. in D. Deutsch (ed.) *Psychology of Music.* New York: Academic Press.

Youngblood, J. E. (1958) Style and information. *Journal of Music Theory*, **2**, 24–35.

CHAPTER 7

# REVIEW OF *LANGUAGE, MUSIC, AND MIND* BY DIANA RAFFMAN

Are there things about musical experience that we know but cannot say? This is the central question that this important book unpacks in the light of philosophical and psychological analysis. The question throws into stark opposition two quite distinct philosophical positions. On one side is the conviction that much knowledge, particularly aesthetic knowledge is ineffable. This position is well expressed by the following remark of John Dewey:

> If all meanings could be adequately expressed by words, the arts of painting and music would not exist. There are values and meanings that can be expressed only by immediately visible and . . . audible qualities, and to ask what they mean in the sense of something that can be put into words is to deny their distinctive existence.

On the other side is the view, championed most vociferously by Daniel C. Dennett, that all conscious knowledge is propositional in form, capable of expression in verbal form if only we attended closely enough and possessed the necessary vocabulary. This 'bracingly Spartan' view of consciousness leads Dennett to assert that:

> There are no colours, images, sounds, gestalts, mental acts, feeling tones or other Proustian objets trouvés to delight the inner eye: only featureless—even wordless—conditional-intentions-to-say-that-p for us to be intimately acquainted with.

Raffman is broadly in sympathy with Dennett's minimalist agenda. In particular she shares with him the rejection of phenomenological 'raw feels', but believes that music provides a good example of something that challenges his strict propositionalism. The basic thrust of her argument is that musical knowledge involves a level of representation that is necessarily non-propositional. There is, in other words, a type of musical knowledge that is not, even in principle, expressible in words.

Raffman's book cogently and meticulously maps out the territory that provides the intellectual underpinnings of her position, from which she is able to define and articulate three different types of musical ineffability. These are not only philosophically exact, but psychologically plausible, and are used to elucidate a number of secondary, but equally important issues, particularly the nature of musical meaning, and the possible role of emotion as a vehicle for musical meaning.

The enterprise is deeply imbued with the 'cognitivism' that characterizes most of the influential work in music psychology of the last two decades. Drawing on the general computational approach of Fodor and others, as exemplified in music psychology through the seminal theorizing of Lerdahl and Jackendoff, Raffman's premise is that 'conscious musical perception or experience results from the unconscious computation of a series of increasingly abstract mental representations of an acoustic signal.' (p. 4) In order to understand her proposals for musical ineffability, it is necessary to sketch an outline of Raffman's characterization of these representations.

At the lowest, or shallowest, level is a representation of the sound in terms of the finest pitches, durations, timbres and loudnesses that we can discriminate (although Raffman, for the purposes of establishing her arguments talks mainly about pitch). Such a representation is inferred from outputs of auditory transducers of physical stimulation. Such outputs are (in the case of pitch) a function of acoustic frequency. This level is called by Raffman, the N-level, or *nuance* level, of representation, because it is a non-structural representation, which captures as much as possible of the richness of the original acoustic stimulus. In language perception, an analogous level of representation might be the level that carries information about the voice of the speaker, accent, intonation and modulation, speed, etc.

At the next level is a representation of the music as a collection of pitch and duration classes within a symbolic system, which Raffman calls the mental score because of its close isomorphism with an actual conventional score. The mental score represents music in terms of such description as can be cashed by trained listeners in verbal expressions such as C#, F, tonic, crotchet, quaver, etc., which Raffman designates collectively as chromatic-pitch-events. This level is the lowest or shallowest at which the representation can be said to be structural or grammatical, in that it is composed of a finite set of tokens or elements out of which all legitimate pieces of western tonal music can be constructed. The linguistic analogy might be phonemic representation.

Chromatic-pitch events are inferred from the N-level representation via mental schemata, built from experience, which represent certain 'institutionalized' relationships obtaining among pitches, intervals, chords, etc.

These schemata are generally designated as 'tonal schemata' within the psychological literature, and their psychological reality has been attested by a large number of experimental studies. They assign n-level events to differentiated classes of event such as the twelve steps of the chromatic scale, chromatic intervals and chord types. With this level of representation, a listener should, at least in principle, be able to take down a musical excerpt from dictation even if, in the sense to be articulated below, the music is not understood.

The third level of representation allows a listener to represent a musical work not simply as a collection of successive notes with tonal and rhythmic/metrical 'names', but as an articulated hierarchical structure in which certain events have greater or lesser importance along a number of dimensions. Raffman designates this level the M-grammatical level (musical grammar). Lerdahl and Jackendoff's *Generative Grammar of Tonal Music* is the most comprehensively worked out proposal for the kinds of structural description that underlie the musical intuitions of experienced listeners of tonal music. The details of their grammar need not concern us here. It has many features analogous to the generative linguistic grammars proposed by Chomsky to underlie syntactic competence, and like Chomsky's grammar, relies for its operation on certain supposedly universal constraints on human cognitive functioning, including in the case of music, gestalt principles of perceptual grouping. Raffman's interest is more in the function of this level of representation (p. 19):

> You, the experienced listener, have unconscious knowledge of certain rules for musical analysis. As you hear an incoming musical signal, you unconsciously represent it and analyse it according to those rules: that is to say, you assign it a structural description. *Ex hypothesi* it is in virtue of assigning a structural description that you have the conscious experience you do—that you hear or, as we often say, *feel* music as you do. For example, it's in virtue of assigning such an analysis that you feel the tonic as being the most stable pitch in a scale, or an accent as being relatively strong in its metrical context, or a harmonic progression as being 'tense' or 'relaxed.' As Lerdahl and Jackendoff will put it, having the right sort of musical structure in your head is what *understanding* the music consists in.

Raffman argues, I believe correctly, that all three levels of representation are potentially available to consciousness. She reviews, with clarity and fairness, the 'categorical perception' data, and concludes that subcategorical information from the n-level cannot be suppressed as fully as it sometimes

is in language perception. We hear both that a note is C (chromatic pitch event) and that it is mistuned (n-level event), and that it elicits, in this particular context, a sense of relaxation or repose (M-grammar).

Although Raffman outlines the above distinctions with admirable clarity, in doing so she has largely reflected a mainstream psychological 'received view', which in itself is not something to get hugely excited about, although it is admirable and unusual for a philosopher to be so clear and up-to-date about both psychological and musical thinking. Where Raffman makes her distinctive and original contribution is in tackling the following key question: in what sense might any of these levels be supposed to involve *ineffable* or unsayable knowledge?

As remarked earlier, Raffman articulates three different types of musical ineffability. The first she calls *structural ineffability*, and it relates to limitations in the ability to verbalize outcomes of representing music according to the M-grammar. She provides some examples of experiences falling into this category, which will be recognizable to many experienced musicians (pp. 31–2):

> Having easily reported the pitches, rhythms and time-signature of a heard piece, along perhaps with some of its local grouping, time-span, and prolongational structures, a listener (perhaps even a highly expert one) finds himself at a loss: 'I am feeling that E-natural in a certain distinctive way, but I can't say just how. I know it's the leading tone, and that it's preparing the return to the tonic, and I know it's a weak prolongation of the E-natural in the previous bar . . . but somehow I feel there is more going on.' Or a performer feels himself compelled to play a certain passage in a certain way (e.g., to slow down, or to get louder, or to increase the vibrato speed)—consciously feels, knows, that it must be played thus, and yet cannot say why . . . What I want to suggest is that unconscious structural representations – in particular, relatively global levels of representation – are 'making themselves felt' in our conscious experience, yet their contents elude our verbal grasp.

The proposed cause of this type of ineffability is simple and intuitively plausible. At more global levels of M-grammar assignation, there will be much more variability of structural assignment across listeners and across hearings (due to training and listening history and performance variation) than there will be at local levels. Thus, the psychological conditions for generation of descriptive language are adverse. This does not mean, however, that verbal descriptions of these ineffable intuitions are impossible. A lot will depend on the experience of the listener, and on the efforts made to

infer an appropriate description. Sometimes, as Raffman suggests (p. 32) 'the structural ineffability dissolves, or anyway seems to dissolve, with explicit music-theoretic analysis: one applies pencil to score and discovers an element of the musical architecture that seems, intuitively, to account for the character of the experience.'

Raffman has pinpointed a real phenomenon that deserves more serious psychological study than it has received. However, she claims that structural ineffability is relatively uninteresting from a philosophical point of view. Examples of structural ineffability certainly do not meet the stringent criterion of being necessarily non-propositional: they are only contingently so for a particular person at a particular time.

The second type of musical ineffability is characterized by Raffman as *feeling ineffability*. This comes about because of the essential sensory-perceptual or experiential nature of music. Knowing a piece of music involves, along with other things, knowing what it sounds like. 'A person deaf from birth cannot know a piece of music.' (p. 40) This does not mean that a trained musician has to hear a particular new piece of music to know it. It is possible to compute what it sounds like (at least to some degree of specification) by examining a score, provided that one has a rich enough base of prior musical experience to fall back on. But without this sensory underpinning there is no possibility of full knowledge.

This type of ineffability is by no means unique to music. All sensory experiences, from sunsets to 'the taste of last night's chicken curry' have the same ineffability. They require 'actual occurent sense-perception of the relevant stimuli at some point' (p. 4). Yet, according to Raffman, there seems to be a special quality to the feeling ineffability of musical knowledge that commands our particular attention in the way that last night's curry does not. Her thought is that it is the grammatical structure of music that (mis)leads us to expect something effable as an end result of 'following' a piece of music. 'Music's grammatical structure may mislead us into semantic temptation' (p. 41) in a unique way not brought about when we observe a sunset, or even a painting. This comes about because of music's formal similarity to language.

These considerations allow a lengthy but extremely useful digression into the possible nature of musical meaning. Raffman argues that there is a sense in which to experience a class of peculiarly musical feelings (feelings of beat strength, metrical stress, prolongational tension, and so forth) is to understand the music. These feelings serve something of the same function as linguistic meanings do with respect to structural representations of language. For instance, experienced listeners by and large cannot help but recover these feelings when listening to music, and there is a sense in

which these feelings can be determined as 'right' or 'wrong'. 'If a person fails to feel the C-natural in a C-major scale as the most stable pitch in the collection . . . then we will be justified in saying that he has misunderstood the music, that he ought to listen more carefully—or, perhaps, that he does not yet 'know the language.' (p. 54) Also, such musical feelings are arguably the basis of musical communication in that 'the performer's mission is to sculpt the nuances of pitch, duration, loudness, and the rest in such a way that his listeners feel the music in the same way he does'. (p. 55)

There is, however, in Raffman's mind a key distinction between linguistic semantics and the 'quasi-semantics' of music described here. 'Whereas the relation of a linguistic string to its meaning is a more or less conventional one, the relation of a musical string to the relevant feelings is nonconventional: we are presumably just *wired* in such a way as to have these feelings upon tokening those mental representations. Not surprisingly, the tie between music and feelings is considerably tighter than the tie between a sentence and its meaning.' (p. 55)

From this perspective, Raffman sheds interesting new light on the proposal that the meaning of music is somehow tied up with the emotional responses it evokes. Basically she claims that emotions are not the basis for musical meaning because, contingently, musicians do not take emotional responses as normative. That is, they are not open to dispute. I may be surprised if you find the last movement of the Jupiter Symphony morose and life-negating, but I cannot say you are wrong. On the other hand, musicians can and do argue about the correct way to feel a particular note in terms of its stability or prominence (p. 59):

> Such disputes may prove unresolvable . . . as when listeners 'hear the music differently' owing to disparities in their unconscious structural descriptions. But such disputes, far from undermining the normativity of the judgments at issue, are possible only in its presence. . . . Musicians argue about phrasings and dynamics and resolutions. They do not argue about the emotions they feel or otherwise ascribe to music.

Raffman also brings to our attention an ambiguity in the use of the term 'meaning', which sometimes refers to content (as in "Bachelor" means unmarried man') and value (as in 'your support means a lot to me'). 'No doubt music means something to us, in the sense of being important or valuable. . . . Perhaps the valuing of musical works for their emotive properties, in concert with their apparent possession of grammatical structure, fosters

the mistaking of those emotive properties for musical *contents*.' (p. 60) It seems to me that this simple but profound distinction probably renders obsolete libraries full of conceptually confused writing on musical meaning!

Insofar as this second type of ineffability attaches to knowledge of musical structure (in principle verbalizable), it also fails to meet the stringent criterion of necessary non-propositionality. For this, we will need to show that a further ineffability also attaches to other levels of musical knowledge incapable of full verbalization. According to Raffman, there is such a level: it is the N-level, and so her third and final type of ineffability is *nuance ineffability*.

This type of ineffability comes about because we do not have categorizations or schemas as fine-grained as the stimuli we can discriminate. For this to lead to the strong conclusion of necessary ineffability, Raffman needs to establish that it is impossible that we could acquire such schemata. Her arguments are basically psychological. First she argues that 'it is hard to see what point there would be to a schema whose "grain" was as fine as that of perception. Indeed the C-interval schemas presumably serve to reduce the information load imposed precisely by the nuance representations of the N-level'. In other words, human processing limitations make total effability of the N-level a psychological impossibility. Second she argues that the just noticeable distinction (JND) for perceptual properties varies with stimulus and observer characteristics (including alertness, health, etc.) Such inconsistencies make it hard to see how long-term mental representations of the required sort could be established. Thirdly, she argues that the perception of, for instance, out-of-tune intervals as out-of-tune (rather than as in-tune instances of microtonal n-intervals) is necessary for us to establish structural awareness at all. If every nuance were recognized and categorized 'we would not be hearing the music as tonal: tonal music just is music whose melodies and harmonies soar out from the background welter'. (p. 86)

I have to say that I find the first and third of these arguments rather less convincing than the second. The first seems to be an unverified assumption. We simply do not know how fine-grained a schema could be developed with adequate experience. Clearly such a schema would not produce in a person possessing it an awareness of music as tonal, in the sense we now understand. Probably new types of music would need to be written to train and exploit this type of schema, but I don't find it inconceivable that, for instance, a microtonal schema or a timbral schema could form the basis of coherent and comprehensible musical structures.

The second argument seems much more powerful. If there is inherent instability in our perceptual discriminations at or around the JND, then the conditions for acquiring schemata of an appropriate grain can never

be established. There will, therefore, always be a residual ineffable component in any conceivable representation of music. As Raffman poetically describes it, musical experience will always be characterized by 'an evanescent corona shimmering around the structural frame of the piece'. (p. 96)

But a final problem remains. If we don't have schemata for nuances, then 'there is an important sense in which we do not know what N-pitch or N-interval we are hearing at any given moment. If that is so, then what is it that we know but cannot say? In other words, what do we know ineffably?' (p. 87) Raffman's solution is to propose that (p. 88):

> you hear (n-events) as nuances, and therein know that they are nuances (as opposed to, say, C-pitches); it's just that, for lack of interval schemas sufficiently fine-grained to permit type identifications, you don't know *which* nuances they are. Hence although in hearing the nuances you know how they sound, there will be nothing you can say, no report you can make, that could, by itself serve to reproduce that knowledge in the mind-brain of another listener. On the contrary, your only hope of getting him to know what you know is to ostend the signal: he must hear it for himself.

I am fully in agreement with Raffman's conclusion that the only way to communicate nuance is by demonstration. This is why so much of music performance teaching is a matter of demonstration: at the end of the day, a teacher has to show what she wants through performance and hope that the pupil can pick it up—'play it like this'. But I think that Raffman is overly pessimistic in believing that musicians do not, or cannot, know which nuances they are experiencing. Indeed, if there is no possibility of identifying and remembering nuances at some level, then there is no possibility of learning by imitation. Recent research has shown that musical performers can retain and reproduce substantial portions of the nuances of heard performances (for instance, rubato and other timing deformations below the level of the smallest note-length). Therefore, they must know, at least in broad outline, which nuances they are experiencing. It is simply that this knowledge informs performance programs more directly and explicitly than it informs verbal consciousness.

Again, if listeners really were incapable of knowing which nuances they were hearing, there would be little possibility of identifying particular performers or performances. But it is common knowledge that connoisseurs can tell Gilels from Schnabel, or even two different Gilels performances from one another. So, although a listener may not be able to identify a

nuance in the language of tonality (e.g. 'a slightly prolonged up-beat'), he may be able to identify it as symptomatic of a style (e.g. 'late Rubenstein'). Thus, he not only knows that he heard a particular nuance, but can identify that nuance as part of one performer's expressive repertoire, rather than that of another performer.

What does all this have to say about the original Dewey–Dennett opposition? It seems to me that Raffman's challenge to Dennett's strong propositionalism is upheld. The JND argument clinches for me that there will always be a residual ineffable 'corona'. Where I think more work needs to be done is in a deconstruction of the n-level, which is more complex than Raffman would have us believe. Even taking the pitch domain alone (and there are many interesting considerations in the time, timbre, and loudness domains which Raffman does not address), then it seems to me that there are certain aspects of nuance knowledge that can be, as it were, gathered up into a kind of 'proto-effability' which may be parasitic on structural representations of music. These are probably the grosser nuances which can be used to 'mark' the c-level representations. So, for instance, we don't just know that this note is an F#—we also know that it is significantly sharpened, lengthened, and softened, relative to its neighbours. We don't necessarily have available the vocabulary to capture all of this, but our ability to imitate or recognize style shows that, in some sense of the term, the knowledge is, indeed, propositional, in the sense that it is symbolic and computable, and can be brought into conscious awareness, even if it is not always there.

The performance research of recent years suggests that the degree of proto-effability of particular nuances is heavily dependent on the cognitive economy with which they can be represented. Nuances that have the effect of heightening the impact of legitimate structural representations are much better assimilated than those that have an arbitrary relationship to structural factors. This is what I mean by saying that the representations are parasitic on structural representations. Another way of saying this is that it is the existence of global maps of the musical territory that allows us to get a fix on some of the local detail. There would be no way of capturing nuance 'from the bottom up'.

There are many aspects of this excellent book to which I have not done justice, in particular a detailed analysis of the impact of Raffman's thinking on our understanding of Nelson Goodman's seminal work on the nature of aesthetic understanding. Whether or not Raffman's conclusions are ultimately defensible does not really matter. What she has done is set a new agenda for the interdisciplinary study of music—at the interfaces of philosophy, psychology, and music analysis—and has set impeccably high

standards in the posing and critical analysis of fundamental questions. There is a strong tendency in the study of music perception for fragmentation of aims. Psychologists pursue goals that are unrelated to major questions in music or philosophy, and vice versa. This leads to mutual accusations of triviality or irrelevance. Our discipline needs particularly strong work at the disciplinary interfaces to ensure that it continues to be 'progressive' rather than 'degenerative', in Lakatosian terminology. Raffman's book provides the proof that such work can still be done. It deserves to be read by anyone who has a serious interest in the study of music cognition.

## References

Raffman, D. (1993) *Language, Music, and Mind.* Cambridge, MA: MIT Press.

CHAPTER 8

# DOES MUSIC MEAN ANYTHING?

## 8.1 Introduction

In 1985 I gave a paper to the Belgian Psychological Society entitled 'Cognition and real music: the psychology of music comes of age' (Sloboda, 1986). It was an optimistic paper, focusing on what I then saw as the best and most promising signs of our young and growing discipline. At that point in time, when we were still basking in the afterglow of Lerdahl and Jackendofi's ground-breaking *Generative Theory of Tonal Music* (1983), it seemed that music psychology had, at last, reached paradigmatic status, in Kuhnian terminology. We were no longer a rag-bag of eccentrics operating at the unfashionable fringes of psychology, but a proper sub-discipline with a programme! I proposed that there were five important characteristics of a healthy scientific paradigm: an agreed set of central problems; agreed methods for working on those problems; agreed theoretical frameworks in which to discuss them; techniques and theories specific to the paradigm; research which is appropriate to the whole range of phenomena in the domain being studied.

In 1985 I believed I had identified at least one central problem or task for music psychology: that was 'to explain the structure and content of musical experience', which could be further elucidated by two more specific questions:

1 What is the nature of musical knowledge or representation?

2 How does music have aesthetic and emotional effects?

Both of these questions still seem to me to be fundamental to music psychology and are absolutely central to the enterprise of determining if, and in what ways, music might be said to mean something. If the mental representation of music is not simply a copy (however reduced the 'grain') of the original acoustic signal, and if this representation is in some respects shared by composer/performer and listener, then such a representation might be held to comprise the 'meaning' of the music. This would be analogous to

saying that the meaning of a sentence is to be found, not in its acoustic or visual characteristics, but in the underlying proposition which speaker intends and listener recovers.

But many commentators have resisted the apparent linguistic analogy, on the grounds that whatever the nature of mental representations of music, they do not and cannot carry referential meaning of the sort carried by sentences. I believe this resistance to be correct. Music is unlike a natural language in many ways, and because it does not have to serve practical communicative functions, there can be much looser coupling between the representations of different individuals. An alternative view of 'music as a language of the emotions' has received expressions of varying degrees of cogency and clarity. According to this view, the meaning of music is somehow tied up with the emotional states it evokes or reminds of.

The interfaces of philosophy and psychology are natural environments to attempt to clarify these two questions, and I will sketch the broad parameters of a 'received' direction in which answers to the two questions above have been framed within the discipline of psychology, illustrating some of these directions from my own research. In doing so, I am very aware that psychology of music has, to a large extent, progressed independently of developments in musicology, and that some of the positions articulated here may appear musicologically naive or off-centre. However, a short paper of this sort cannot go about rectifying the divergences of a generation of scholars, and makes no attempt to do so.

## 8.2 The nature of musical representation

The book by Lerdahl and Jackendoff (1983) remains the paradigmatic expression of the 'received' view on representation of tonal music (which has been taken by most psychologists as the paradigm example of a musical language). It is grounded in the following assumptions:

- **music is constructed according to grammar(s)**—i.e. sets of explicit or implicit rules or constraints determining the legality or 'well-formedness' of musical sequences
- **representation is often hierarchic**—i.e. grammars determine elements within a musical sequence which control or subsume neighbouring elements, and this process is recursive so that lead elements at one level become subsumed at the next higher level (metre is probably the simplest example of this)
- **tonality, metre, and rhythm are psychologically real organising principles (possibly instantiations of musical universals)**—i.e. these

music-theoretic descriptions correspond to the natural tendencies of perceivers to 'parse' acoustic input

- **different processes (e.g. composition, performance, perception) access the same core representations,** i.e. there is a unified representational substrate which all explicit activities feed on, even if there are additional representational modes specific to the activity (e.g. motor representations of performance).

Lerdahl and Jackendoff's main achievement was to deliver a set of specific operations which, if applied to a piece of tonal music, would yield a complete description of that music, in which each event was marked as being more or less prominent in one of several hierarchical systems. Such descriptions yielded strong empirical predictions about perception and memory (e.g. listeners would recall prominent events more readily than subsidiary ones, would hear prominent events as more important or accented, would hear unstable events as subsumed by more stable ones, etc.).

A strong claim that could be made on the basis of this type of programme is that to have understood a piece of music is to have parsed it according to the principles embodied in some representational theory which is shared within a community of composers, performers and listeners. Although there may be areas for legitimate dispute (indeed the interest of music analysis is often exactly in assessing plausible alternative parsings), there are some key characteristics of pieces which could not be represented in a radically different way which had any real claim to be a meaningful representation of that piece. If anyone were to propose that the most prominent notes in the opening theme of Mozart's Symphony no 40 (E<flat> D D E<flat> D D E<flat> D D B<flat>) were (in descending order) the 2nd, 5th and 8th notes we would be strongly inclined to say that he or she didn't understand the music. Most of us would agree 3, 6, and 9 in ascending order. Without going into detail, much perceptual and cognitive research in this area can be seen as showing that there are, indeed, large areas of agreement between individuals about such matters. These agreements often transcend different levels of training, and sometimes transcend cultures.

If the strong claim were to be true, the search for musical meaning must stop at the level of structural description. Everything else would be, as Stravinsky and others have claimed, irrelevant noise.

## 8.3 Shortcomings of the representational approach

Any musician presented with the above type of account of the meaning of music is likely to complain that, however real the processes and

representations described above are, they miss out the core of what makes a musical experience musical. A structural description is merely the skeleton and framework on which flesh and blood must be put. To have constructed a structural representation is only a small part of musical understanding. One could characterise this response as proto-Searlian. Everything described above could be put inside Searle's Chinese Room (Searle, 1980).

I once had the privilege to study a young autistic man, Noel Patterson, who was a rather clear example of a person who could do everything implied by the strong claim but who failed to demonstrate 'musical understanding' in some very important sense (Sloboda, Hermelin, and O'Connor, 1985). He was able to memorize heard tonal (but not atonal) piano music rather quickly, and reproduce it with a high degree of accuracy. His recall errors were 'structural', in that omissions tended to be subsidiary elements (as characterized by such theories as that of Lerdahl and Jackendoff) and substitutions were structurally appropriate. But after a very short time his performances became stripped of all expression (even that contained in the original from which his performance was derived) so that there was nothing of interest in his performance for an audience. He played totally mechanically. From a musician's point of view it is almost as if Noel Patterson had 'missed the point' of bothering to play the music at all.

## 8.4 Feelings as the dynamic aspect of music

What is this 'point' that both the strong claim and Noel Patterson's performance leave out? It is that the structures underlying music animate an experience which is *dynamic*. The experience involves sensations of tension and resolution, anticipation, growth and decay. One account of this would hold that such experiences *depend* upon having assigned a particular structural analysis to a sequence. We feel the music as increasing in tension towards a particular point *in virtue* of its being assigned a structural importance, but the structural description does not in itself *incorporate* that feeling. The structural description is a precondition for the feeling. Another account would have the experience of dynamism as an essential precondition of a listener recovering the structural description, given that there is no proof that an analysis such as that offered by Lerdahl and Jackendoff can be carried out algorithmically on the basis of information presented in a score. We may need the dynamic feelings to bootstrap the parsing process. Lehrdahl and Jackendoff, whilst acknowledging that there is a relationship between some of their structural descriptors and dynamic experiences (e.g. of tension or resolution) did not seek to establish a direction of psychological causality. Either way, it is not at all clear that the feelings we

are speaking of here are identical with those feelings that we characterize as emotions or moods in other contexts. Emotions and moods are valenced (positive or negative) feelings where one's own orientation to the object of the emotion is a key component (like–dislike, approve–disapprove, happy–unhappy, etc.; cf. Frijda, 1986). We may indeed feel such emotions and moods to music (and these feelings are the stuff of our aesthetic and other responses to music, determining the importance or personal significance of our experience—a different kind of meaning, as Raffman (1993) has so clearly pointed out), but the dynamic feelings or sensations of tension are essentially unvalenced. Given a particular structural understanding of the music feelings of tension or resolution are just a correlate of that understanding (whether I like them or not).

So, a musical performance could be one in which the performer recognizes the dynamic implications of the structure and enhances the acoustic signal in such a way as to boost or highlight those implications. Returning, for instance, to the initial theme of Mozart 40, one 'musical' performance response to this would be to make each of 3, 6 and 9 a little louder and a little more prolonged than the previous element. How does a performer know that this is a good or appropriate way to highlight the dynamic characteristics of the passage? And how does a performer choose appropriate levels of loudening and slowing? A number of theorists are converging on the notion that this knowledge is a consequence not only of the structural characteristics of a given piece of music but also of a whole range of other signals that impinge upon humans and can be mapped, by processes of analogy, onto the musical process. In other words, it is the whole process of being a biological human inhabiting a physical and social world that provides the fuel and the impetus for a dynamic understanding of music. Putative non-human sentient beings would have a hard job understanding music in the way we do, because their set of mapping analogies would be so different.

The analogical mapping we are speaking of is entirely different from the iconic or associative process whereby musicians can imitate real-life sounds, such as birdsong. That is a rather uninteresting mimicking of surface characteristics. What is implied in the above proposal is somewhat deeper and more abstract. Imagine, for instance, that I come to a swinging door in a corridor. I push it, and it doesn't yield. I push again, a little harder. Still it doesn't quite yield, but I sense that a third push will do it. I push for a third time, harder still, and the door suddenly swings open, propelling me through to the next corridor. This is just one of many concrete examples of a dynamic process which involves repeated application of increasing force and a sudden release of whatever was holding the obstacle in place.

One reading of the opening theme of Mozart 40 (and, of course, there are many others) could assign it to this dynamic shape, representing something that has got stuck at a particular level, here represented by the small semitone movements from D to E<flat>. As the force is repeated, the obstacle is overcome and the music is 'freed' to fly up to the B<flat>. The increasing force is well represented by louder and longer sounds because this would be the exact consequence of reproducing this sequence of actions on the bow of a violin, for instance. The actual bodily process of interacting with an instrument in a real physical environment makes this analogy directly available to the performing musician. Incidentally, this is probably why computer music often sounds 'inhuman'. Its creators are not bound by the constraints that bind musicians interacting with real instruments in real time using real human bodies!

If this line of argument is accepted, then it follows that a major task for those concerned with explaining how we understand music is to analyse and describe the whole range of analogies of the above sort that are available to musicians. The notion that the instantiation of music practice in actual bodily movements may profoundly affect its structure and our conceptions of it, has been forcefully championed by Baily (1985), Blacking (1995) and Cuniming (1997) among others. Suggestions by a number of researchers (e.g. Clynes, 1977; Sundberg and Verillo, 1980; Todd, 1995) indicate the type of analogies which might be available to even the most unsophisticated listener. The results presented by Watt and Ash (1998), in which music is reliably ascribed characteristics such as male–female, good–evil, may come about through the integration of clusters of appropriate basic analogies into 'virtual persons'. But this is not the only, or even a necessary, consequence of dynamic awareness of music. The type of ascriptions that listeners are prepared to make on the basis of musical experience will undoubtedly be determined by a range of cultural and contextual factors.

## 8.5 Physical manifestations of emotion as core unvalenced reactions to music

It is commonplace knowledge, although rather little researched, that music evokes in people physical concomitants of emotion. Music 'moves people to tears' with great reliability and regularity. Although I wouldn't have used precisely this terminology at the time, it struck me some years ago that such concomitants were rather attractive candidates for one major class of 'unvalenced dynamic sensations'. My own subjective experience of such sensations is often of noticing them first, interpreting them (at an

aesthetic level) second. For instance, there comes a tightening or pricking sensation at the back of the throat which has an increasing intensity over a number of seconds. I then notice sensations behind my eyes. As sensations increase, I notice my eyes watering. A point of maximum tension is reached, then the feeling ebbs and subsides. Whether I like or value this feeling depends upon a very large range of considerations. Do I think the music is profound (or maudlin)? Am I alone or with others? Am I performing this music, and will indulging these sensations help or inhibit the performance, etc.? All of these later evaluative processes are part of the aesthetic experience. The raw 'tears' seem much closer to the form of the music itself. It is also interesting to note that it is sometimes quite difficult to ascribe a particular emotion to such an experience. Often it seems neither sad nor happy in a straightforward way.

Many musicians report having experienced these raw manifestations rather early in life (Sloboda (1989) found them beginning at around age 7–8), and musicians often claim that these experiences had a profound effect on their motivation for future involvement with music (similar findings are reported in a series of studies on 'strong emotions to music' by Gabrielsson and Lindstrom, 1993). Over 90% of respondents in a survey claimed to have cried to music in the last 5 years (Sloboda, 1991).

To substantiate the hunch that these feelings are tied to musical structure, it is necessary to plot the exact time-course of the feelings and see if there is any systematic link to structure. Waterman (1996) played a variety of listeners the same musical extracts (in individually testing sessions) and asked them to indicate by pressing a button whenever they 'felt something'. He found significant agreement between listeners in location of button presses. There was also within-listener agreement when the same pieces were played one year later. This suggests that these reactions are stable and locatable within specific portions of the musical structure.

To widen the range of musical materials included, Sloboda (1991) surveyed 75 individuals and asked them to identify specific pieces of music that had reliably elicited physical reactions such as tears, shivers down the spine and gooseflesh, or racing heart. All examples supplied were of western tonal music, but ranged from Bach to contemporary pop and jazz. Musical analysis of those segments which participants could identify precisely (within a measure or phrase) yielded an interesting and unexpected correlation. Tears were associated with melodic or harmonic sequences, appoggiaturas and suspensions, and downward harmonic movement through the circle of fifths. Shivers were associated with enharmonic changes and other harmonic, textural, or dynamic discontinuities. Racing heart was

associated with syncopation, and other forms of accentual anticipation. These effects were insensitive to style, and examples of all these effective structures were found in both classical and pop music. This is another indication that these responses are 'pre-aesthetic'.

While a detailed explanation for the specific feeling-structure linkage has not yet emerged, it is clear that the features of music associated with these feelings are features of the sort that would be marked as more or less prominent within a Lerdahl and Jackendoff type of analysis, yielding appropriate responses of tension-resolution (e.g. appoggiatura) or violation of expectancy (e.g. enharmonic change, syncopation). This is a vindication for the structurally-based explanations of emotional response to music pioneered by L. B. Meyer (1956, and his followers, e.g. Narmour, 1977).

I do not wish to claim that such sensations as described above constitute a level of meaning for music over and above the structural. They are, however, intense symptoms of the operation of such a level. This level can, no doubt, be investigated using other techniques (for instance monitoring of brain and physiological response, or continuous manual response of the kind used by Madsen, *et al.* 1993). Such techniques all share the feature that they measure essentially non-propositional or 'hot' cognitions, rather than the 'cold' cognitions of the propositional level. It is also important to re-emphasize that these sensations or feelings are not specific emotions, although they may easily give rise to specific emotions if appropriate contexts or associations are to hand. In this respect I believe that the absolutist position on emotion (as, for instance, argued by Meyer, 1956) may be more robust than some of us might have thought!

## 8.6 Summary

Mapping from language to music might be an overplayed tactic, but the above analyses suggest that there is a way of doing it which doesn't do too much injustice to either form. Insofar as linguists still allow a division between syntax and semantics (and I am not fully conversant with the latest thinking), it seems to be that one could see the structural description of the music as a form of syntax, and the dynamic sensations of flux, tension, expectations fulfilled or violated as the semantics. This is because the structural description does not refer to anything outside the music, while the dynamic aspects do so refer, if only by analogy. Dynamic awareness of music involves reading the music as an embodiment of something else, and my proposal for that something else is, broadly, the physical world in motion, including that very special sub-class of moving objects, the living organism. But this analogical linking up is not at the level of specific

mimicry of the auditory signals of that moving world, but at various deeper levels concerned with the underlying forms of these motions as initiated, experienced, and mediated by the conscious human agent. Specifying what is included in this, and finding a principled reason for excluding what should be excluded is a task hardly yet begun.

## References

Baily, J. (1985) Music structure and human movement. In P. Howell, I. Cross, and R. West (Eds.) *Musical Structure and Cognition*. London: Academic Press.

Blacking, J. (1995) *Music, Culture and Experience*. Chicago: University of Chicago Press.

Clynes, M. (1977) *Sentics: the Touch of Emotions*. New York: Doubleday.

Cumming, N, (1997) The subjectivities of 'Erbarme Dich'. *Music Analysis*, **16.1**, 5–44.

Frijda, N. (1986) *The Emotions*. Cambridge UK: Cambridge University Press.

Gabrielsson, A. and Lindstrom, S. (1993) On strong experiences of music. *Jahrbuch der Deutschen Gesellschaft für Musikpsychologie*, **10**, 114–125.

Lerdahl, F. and Jackendoff, R. (1983) *A Generative Theory of Tonal Music*. Cambridge: MIT Press.

Madsen, C. K., Byrnes, S. R., Capperella-Sheldon, D. A., and Brittin, R. V. (1993). Aesthetic responses to music: musicians versus non-musicians. *Journal of Music Therapy*, **30**, 174–191.

Meyer, L. B. (1956) *Emotion and Meaning in Music*. Chicago: University of Chicago Press.

Narmour, E. (1977) *Beyond Schenkerism The Need for Alternatives in Music Analysis*. Chicago: University of Chicago Press.

Ortony, A, Clore, G. L., and Collins, A. (1988) *The Cognitive Structure of the Emotions*. New York: Cambridge University Press.

Raffman, D. (1993) *Language, Music and Mind*. Cambridge, MA: MIT Press.

Searle, J. R. (1980) Minds, brains, and programs. *The Behavioural and Brain Sciences*, **3**, 417–457.

Sloboda, J. A. (1986) Cognitive psychology and real music: the psychology of music comes of age. *Psychologica Belgica*, **26.2**.

Sloboda, J. A. (1989) Music as a language. In F. Wilson and F. Roehmann (Eds.) *Music and Child Development*. St Louis, Missouri: MMB Press Inc.

Sloboda, J. A. (1991) Music structure and emotional response: some empirical findings. *Psychology of Music*, **19.1**, 110–120.

Sloboda, J. A., Hermelin, B., and O'Connor, N. (1985) An exceptional musical memory. *Music Perception*, **3**, 155–170.

Sundberg, J. and Verillo, V. (1980) On the anatomy of the ritard: a study of timing in music. *Journal of the Acoustical Society of America*, **68**, 772–779.

Todd, N. P. McA. (1995) The kinematics of musical expression. *Journal of the Acoustical Society of America*, **97**, 1940–1949.

Waterman, M. (1996) Emotional responses to music: implicit and explicit effects in listeners and performers. *Psychology of Music*, **24.1**, 53–67.

Watt, R. and Ash, R. L. (1998) A psychological investigation of meaning in music. *Musicae Scientiae*, **2.1**, 33–53.

PART B

# EMOTION AND MOTIVATION

CHAPTER 9

# MUSIC AS A LANGUAGE

## 9.1 Introduction

I guess that all of us are here because, in one way or another, we love music. It is that love which has led us to become teachers, performers, researchers, concerned to deepen our awareness and effectiveness in ensuring that music and musicians are held in high esteem by society and given society's best resources.

But as well as bringing love with us, most of us also probably carry a deep puzzlement; despite our deep familiarity with it, music still remains at many levels a mystery to us. We don't really understand what music is, how it comes to have such a profound effect on us, why it moves us, fascinates us, brings us back to it again and again.

To many of us, the mystery is part of the love. As in any love affair, it is the sense of never ending discovery which gives impetus to the love. Once we feel that we have found everything out about the other, the love dries up. It is, I am sure, for this sort of reason that many of us have a deep aversion to the scientific analysis of music. It seems that if we could reduce music to a set of explicit categories or formulae, if we could derive an all embracing theory, then music would lose its mystery and power.

I very much respect this intuitive aversion. I think scientists labour under a constant temptation to replace what they are studying with theories about what they are studying, and then talk about their theories as if they were the total reality of the situation. To overcome this tendency to narrow the sights, I believe that every scientist who studies music has a duty to keep his or her love of music alive, and be constantly aware of what escapes the net of particular theories.

On the other hand, I believe that science has much to offer musicians. One thing that it does is attempt to make theoretical assumptions explicit, and thus open to discussion and test. Everybody, whether scientist or not, is operating on theories and assumptions. Whatever else scientists may or may not do, they can offer people new ways of looking at old issues.

And sometimes these new ways are more profitable. They can also offer hard data to replace speculations.

Let me give one example of what I mean. Recently I attended a conference on what in Britain we call 'ear training' or 'aural training.' There was a lot of talk among the teachers there about 'developing a good ear' for music, and some rather sterile controversy about whether some people had better ears than others. As a psychological scientist, I was able to remind them that the ear as such is not the problem. Most people's ears function excellently, and there is nothing anyone can do to enhance their functioning. Everyone's ear is constantly sending to the brain highly sophisticated, fine grain information about all sounds received. It is what the brain does with it that determines musical differences between people.

As a result of a large amount of experimentation and theorizing in cognitive psychology, we are now able to say rather precisely what needs to happen to produce the behaviour that musicians would associate with 'a good ear.' First, the relevant dimensions of the sound have to be *attended* to, then the listener requires some method of *coding* or *categorizing* the individual sounds. Thirdly, the listener must be able to hold the various sounds together in some *structure* or *pattern*. Finally, the listener must be able to translate what has been received into some form of *response*. Many unfortunate people have been labelled 'tone deaf' because they can't sing in tune. The analysis just outlined shows that the problem could be entirely one of response; the other stages may have been completed perfectly.

I believe that the four stage theory of musical awareness is more useful to practitioners than an unarticulated theory about 'the ear.' It allows us to ask at what stage a problem might be occurring, and it gives us better ideas about what we might do to overcome the difficulty. All these stages are known to be susceptible to profound alteration through learning and experience. The theory does not solve all a practitioner's problems. The business of teaching and learning is still largely an art; but I contend that the availability of such ways of thinking will enhance a practitioner's functioning, by providing better and more varied tools to tackle particular problems.

It is in the same spirit that I offer some observations on music as a language. I do not claim that thinking of music as a language is a completely satisfactory way of accounting for everything about it. Indeed, some of the most important things about music seem to arise from the ways it is different from a language. But it offers a way of thinking about music which is, I believe, useful and illuminating.

Like music, the primary medium of language is sound. We can capture and measure the various physical dimensions of speech sounds through the many sophisticated measuring instruments we have. In themselves,

however, these sounds are just sounds, with various physical characteristics such as pitch, amplitude, and timbre. They are not language. What makes them language is what human brains do with them. And it seems that what human brains do is to attempt to map these sounds onto internal structures. When some kind of match has been made with these structures, then language can be said to have come into existence. These internal structures seem to be divided into these types: *phonological, syntactic,* and *semantic.* Generally, some contact has to be made at each level for speech to be understood or created.

## 9.2 Phonology

Phonology describes the way in which the brain parcels up continuously varying sounds into discrete and separable units. These units constitute the basic building blocks of a language. It seems clear from studies on language that phonology is learned. Different cultures place phonological boundaries in particular places. What are taken as two separate speech sounds by users of one language are heard as the same by users of another.

There are some dramatic scientific demonstrations of the way that phonological mechanisms transform the raw sound. One of these is the phenomenon of *categorical perception.* We can take two naturally occurring speech sounds, such as 'da' and 'ba,' and artificially synthesize them. We discover that what distinguishes them physically is the frequency of the initial burst of sound. When it is high we hear 'da'; when it is low we hear 'ba.' What happens if we synthesize a series of sounds which move from 'da' to 'ba' through small steps of frequency change? How are they heard? The common sense view would be something like this. The sound definitely starts as 'da.' Gradually it becomes more and more ambiguous until a point is reached where it sounds like neither, or an amalgam of the two. Then it gradually becomes more and more like 'ba'.

Experimental evidence (e.g., Liberman *et al.,* 1961) shows that this common sense view is wrong. What actually happens is that we hear a series of apparently indistinguishable 'da's', and then, quite suddenly, the perception flips, and the experience is of 'ba'. Our brains categorize intermediate sounds, and make it very difficult for us to notice any difference between two sounds falling within the same category. Then, as some internal boundary is crossed, the perception quite suddenly changes, as the sound is assimilated to a new category.

What, then, of music? The best evidence we have (Locke and Kellar, 1973) is that similar categorization processes go on for most of us. That is, we tend not to perceive every slight change in sound dimensions, but

assimilate to broad categories. One example of this is the musical scale. If we take a major triad, and change the third note of the chord in several equal small steps from major to minor, then many people experience an effect analogous to the language experiment. There seems to be a rather sudden shift from major to minor somewhere in the middle of the sequence.

We can find analogous effects in the time domain. If, for instance, we ask people to reproduce rhythmic patterns, there is a tendency for the reproductions to be inexact but categorical. That is, if the ratio of two notes in the original is roughly 2 : 1 then the reproduction will tend to be roughly 2 : 1. We have internalized categories which tell us that durations tend to be whole number multiples of one another. We find it hard to hear truly ambiguous rhythms as ambiguous. We try to fit them to one or other definite category.

When comparing to language, we need to make two important caveats. One is that the degree of categorization tends to depend on musical experience. In most of the published studies 'musicians' (however defined) tend to show stronger categorization effects than 'non-musicians'.

The second caveat is that even when we do categorize, we are still able to be aware of the differences within a category. So, for instance, a musician will be able to say 'that C# is flat.' This shows both categorization (an identification of the note as in the category C#, and the awareness that the note differs from an ideal central pitch. Sometimes we are simply aware that two notes within the same category differ in some way, without being able to consciously identify the difference. So one finds musicians who say 'I know that the note is badly pitched, but I can't tell whether it is sharp or flat' (Siegel and Siegel, 1977). The same phenomenon also explains why we are able to find two performances of the same notes different. We judge performance A to be livelier, clearer, more satisfying, than performance B, but cannot precisely identify what it is that makes them different. In general, these differences are due to minute expressive deviations within categories.

## 9.3 Syntax

The second level of structuring in language is syntax. This concerns the ordering of the basic phonological building blocks. In all languages, and in music, it is not enough to notice which sound categories (be they phonemes or notes) are present; their combination and ordering contains the crucial information. Most languages (and many musical styles) seem, at least in principle, describable by a grammar. This is a set of rules which is capable of generating (or recognizing) all sequences which are acceptable, and fails to generate (or rejects) sequences which are not acceptable.

There are many types of evidence that people handle language through syntactic structures. One is that it is much easier to remember word sequences that conform to grammatical order than the same sequences that have their order scrambled. This is so, even when the sentences make little or no sense. For instance, it is easier to remember *Colourless green ideas sleep furiously* than *sleep green furiously ideas colourless.*

In music, it is easier to remember sequences which conform to conventional rules of tonality than those that don't. This is increasingly so as one becomes more experienced with a particular musical culture. For instance, there are studies which suggest that the memory advantage for tonal sequences does not develop until about age seven in the average child (Zenatti, 1969).

In my own research (Sloboda, 1985) I have been interested in the growing ability of children to reject what we would judge to be illegal musical sequences. At age five, most children I tested seemed unable or unwilling to reject gross chordal dissonance as wrong. By the age of nine, they were overtly laughing or screwing up their faces at the 'wrong' chords, and scoring at an adult level. In another test, each chord was tonally coherent in itself, but the ordering of the chords could be either tonally conventional (e.g. ending with a cadence) or scrambled (e.g. ending without resolution). On this test children did not achieve adult levels of performance until the age of 11. I was particularly interested to find that these abilities seemed to develop regardless of formal music tuition. Children who had regular lessons did not develop these abilities earlier than other children. It seems that mere exposure to the standard musical culture is enough for children to build grammatical structures.

## 9.4 Semantics

I take semantics to be the systematic set of processes whereby the symbols of a language are able to be mapped onto, or represent, objects, states of affairs, and events that are not part of that language. To be a full language user it is not enough to recognize a sequence as grammatical; one has to understand it. This means at least two things: identifying what the individual words refer to, and identifying what kind of event or proposition is being described.

I see semantics as the driving force behind language acquisition. Children don't learn the meaning of their first words until they are already familiar with the objects and states that the words represent. Similarly they cannot learn to express relationships between objects (such as wants, movements) until they know and experience such relationships.

Their map of the world and their understanding of it undoubtedly develops as a result of becoming language users, but the whole process would never start unless they first had something there before language.

There was once a belief, generated through some of Chomsky's ideas, that children might have a 'language acquisition device' which would build up competence through mere exposure to language. Most experts would now reject this idea, emphasizing instead the necessity of a child's embeddedness in the social and physical worlds as the major stimulus to language development.

The study of semantic aspects of language development has proved rather fruitful. This is because it is relatively easy to determine a child's understanding through practical behaviour. We know that the child understands the concept 'under' if she is able to carry out the request to 'put the green block under the red one.' We all know the domain (or one important domain) onto which language has to be mapped. It is the public world of social and physical action that we all inhabit.

When we come to music, we have a difficulty. I think we would find it very difficult to agree on a set of criteria for demonstrating that a person had 'understood' some music. The domain, if there is one, onto which music maps is not wholly (and some may argue not even partly) public. Although there is general agreement that it has in some way to do with emotion, feeling, or affect, there are scant hard data to constrain our theorizing.

What little research has been done has tended to focus on the ability of listeners to verbally identify the general character of different pieces of music. Hevner's (1936) pioneering work still stands as the major contribution in this area. She showed that within our culture, adults tend to generally agree which adjectives best describe the character of a musical excerpt, that is, whether it is happy or sad, lively or solemn, restful or agitated. Gardner (1973) has more recently shown that this ability develops slowly through childhood and into adolescence.

The ability to describe and talk about music is clearly an important aspect of the awareness of musical meaning, but it is not primary. Anyone who has had any significant musical experience will know that the words to describe it are not always at hand. This will be particularly so for children. We should not take what children say about music to fully reflect their experience of it.

Another thing which limits many studies is their mode of data collection. Subjects are played several brief extracts in an experimental context. In contrast, real musical experiences are embedded in rich social and personal context. If we wish to understand more about how musical meanings are acquired, we need to turn to the everyday real-life acts of musical involvement that make up a child's musical life.

With substantial resources, I guess one would want to observe directly many children over many years. With fewer resources, I have had to adopt a different approach, but one which is nonetheless turning out to be rich and promising.

## 9.5 The autobiographical memory study

I have been asking people to write down the details of any autobiographical memories they could access which involved music in any way, from the first 10 years of their life. To date I have collected 113 usable memories from about 70 adults, ranging from professional musicians to people with little or no current involvement in music.

Despite the apparent indirectness of this approach, it allows us to answer questions not addressed by earlier research. First, we get the responses of people to the naturally occurring musical events of their lives, rather than to contrived experimental situations where a response is forced. Secondly, we can get some indication of the effect of such events on subsequent musical development. Thirdly, adults often have the words to describe, in retrospect, an experience that as a child he or she would have been inarticulate about.

It seems to me that for there to be any kind of consensus in a culture about musical meaning there would have to be fairly general kinds of experiences with music shared by large numbers of children in a culture. Furthermore, these experiences should be expected to generate regularly the same kinds of affective responses. I would not, for example, expect crucial experiences to occur very often within formal musical training. Such training is only given to a minority of children and could not account for wider cultural patterns.

I am aware of two great shortcomings of my study. One is that few people are able to access specific memories before the age of three or four. If the crucial experiences occurred in the first three years of life then my study will not capture them. The second shortcoming is that memory is selective, biased and inaccurate. I cannot guarantee that this sample of memories would necessarily give the same picture as a sample of observations on the subjects of children.

It is however, possible to argue that a memory study actually highlights the kind of event we are looking for. This is because memory for specific events very often seems linked to a high emotional tone. For instance, almost everyone of an appropriate age remembers exactly what they were doing when they heard the news of John F. Kennedy's death. If the meaning of music is connected to affect, then specific memories related to the

acquisition of this meaning should be high in affect, so more likely to be remembered. Memory should act as a sort of filter or concentrator.

Although my volunteers were asked to write freely, I offered them 10 questions as cues:

1 How old were you?

2 Where did the experience take place?

3 Of what event did the music form a part?

4 Who were you with?

5 Can you identify the piece of music, or say anything about it.

6 If you can identify it, have you experienced it recently, and do you know it well?

7 What significance or meaning did the experience have for you at the time?

8 Did the experience influence your subsequent behaviour or attitudes in any way?

9 How often, if at all, does this memory come to you?

10 Any other comment.

After subjects had finished the recall task, they were asked to provide information on:

1 Current level of involvement with music.

2 Amount of formal music tuition received prior to the age of ten.

3 The degree to which experiences such as that reported determined current level of involvement with music.

4 The number of other memories available which were not written down.

On examination of the results, a coding scheme was devised so that quantitative statistical analysis could be carried out. It very quickly became clear that events could have two distinct types of significance. Firstly, the music itself could have some effect on the person. This I have called the *internal* significance of the event. Secondly, the context in which the music was taking place could have some effect. This I have called the *external* significance.

In the initial classification, I coded each memory on each of these two dimensions according to whether the significance was positive, neutral, or negative. Table 9.1 shows the frequency breakdown for all combinations of internal and external significance. Several points are of interest. First, there are almost no cases of negative internal significance. People rarely remembered events where the music itself was disliked or aroused negative affect. This is not simply a reluctance to recall negative experiences or some form of repression, since plenty of memories with negative external significance were recalled.

Secondly, the number of events having neutral classification on both dimensions was rather small (17% of the total). Thirdly, there are almost no cases of positive internal significance where the external significance is negative. One way of interpreting this is to say that you can't enjoy the music when you are not enjoying the circumstances in which it takes place.

I would like to give a few examples of specific memories illustrating the main categories of response. Here is one which shows positive internal significance but neutral external significance.

*Subject 42: age twenty; regular private performer; tuition from seven*

> I was seven years old, and sitting in morning assembly in school. The music formed part of the assembly service. I was with my friends Karen, Amelia, Jenny, Allan. The music was a clarinet duet, classical, probably by Mozart. I was astounded at the beauty of the sound. It was liquid, resonant, vibrant. It seemed to send tingles through me. I felt as if it were a significant moment. Listening to this music led to me learning to play first the recorder and then to achieve my ambition of playing the clarinet. Playing the clarinet has altered my life; going on a paper round and saving up to buy my own clarinet; meeting friends in the county band. . . . Whenever I hear clarinets being played I remember the impact of this first experience.

**Table 9.1** Rated significance of 113 memories

| Internal | External Positive | Neutral | Negative |
|---|---|---|---|
| Positive | 16 | 25 | 3 |
| Neutral | 25 | 19 | 24 |
| Negative | 0 | 1 | 0 |

I would like to give one further example in this category to make a specific point:

*Subject 49: age nineteen; regular private performer; tuition from six*

> I was about ten years old, at school with the teacher and the other kids. The music was a very sad song about the loneliness of the people which we learned in a music lesson. The music made me very sad. I haven't heard it recently and don't know it well. I loved this piece so much that I bought some music instruments and played them every night. I used to think of this song very often, but then I forgot it almost totally until this experiment.

What this memory shows is that sadness in music can be a positive and sought-after experience, even for children. The next example shows positive external significance, but neutral internal significance.

*Subject 23: age twenty; occasional private performer; tuition from seventeen*

> I was about ten years old. I was singing in the school choir for the school carol concert. We were in the school hall. The carols were some more unusual ones. I can't remember them now. I remember being pleased at missing history lessons, and also because it was a really cheerful occasion. This memory returns occasionally, but I don't think the event had any particular influence on me.

The final example shows negative external significance with neutral internal significance. It is sadly by no means unique in this set of memories.

*Subject 31: age fifty; occasional listener; tuition from seven*

> I was five years old. I was at the local Infant School, and a music lesson was being taken by the Infant Head Teacher, Miss Linkler. She played a piece of music. I have no idea what it was called. I don't think I was ever told the name of the piece. The class had to listen to the music, and then beat out the time. Suddenly the teacher pounced on me, screaming that I was beating 3/4 time to a 4/4 piece of music. She then produced a battledore and gave me six smacks on the back of my thighs. At the time I ran home, and my mother had the greatest difficulty in making me go to school

> after that. As far as the music goes, I have always considered that I am not musical, and for many years I refused to sing or do anything connected with music. . . .

Since most volunteers provided a wealth of information about the social context of their memories, it has been possible to isolate factors which are associated with positive internal significance. The most significant of these are:

- *Place:* When in a concert hail, church or at home, over 60% of the memories had positive significance, as opposed to only 29% of events at school.
- *Event:* 59% of events where the subject was listening to music were positively significant but only 17% of events where the subject was performing were positively significant.
- *With:* When with family and friends, 50% of memories were positive; when with teachers only 27% were positive.

All this suggests that the occasions when significant meaning was transmitted were mainly informal, relaxed occasions when the person was not being evaluated, and was in the company of loved ones. A stepwise discriminant analysis revealed that the event was the only one of these variables to enter into the discriminant function. This suggests that listening is the crucial factor. It just so happens in our culture that opportunities to listen to music are provided more often away from school and teachers.

Positive internal significance is also predictive of indices relating to subsequent behaviour:

- *Status:* 53% of people currently claiming to be regular performers of music produced a memory of positive internal significance; whereas only 26% of those not regularly performing did so.
- *Influence:* Positive memories were recalled by 68% of those who claimed that the memory had an enhancing influence on their subsequent musical involvement. Such memories were only recalled by 14% of those who claimed a neutral or negative subsequent effect.

It seems, therefore, that children experiencing events with positive internal significance were more likely to pursue a high level of involvement with music in later life. Interestingly, this was not true of external significance. Positive external significance did not go with a higher level of subsequent involvement in music. People will not pursue musical activity just because they have experienced positive events around music. The positive experience has to come from the music itself.

Finally, it turns out that the amount of formal music tuition received before the age of ten relates weakly to experience of positive internal significance. If, however, we compare the age of the positive experience to the age of starting lessons, we find that the experience often precedes the start of lessons. Of the 36 events for which full information was available, only 9 (25%) occurred after the beginning of formal music tuition. It looks as though such experiences are spurs to a child seeking lessons, rather than the lessons providing a basis for such experiences.

On the sample I have so far it is hard to conclude a great deal about specific types of music eliciting these experiences. Sometimes people could not name the piece in question, and when they did, they covered a vast range of classical, folk, and popular pieces. The one thing which has emerged so far is that memories of nursery rhymes, hymns, and carols (of which there were a significant number) almost never carried positive internal significance. My explanation for this is that it seemed to be usually the *first* experience of a piece which led to the most profound effects. Many of the rhymes and hymns will have been repeated many times in the child's life, and the first experience of them may well have been before the age of concrete memories.

A similar analysis was carried out to see what factors seemed to be responsible for *external* significance. The picture here is simple. Negative significance is most often associated with performing situations (56% of cases), almost never with listening (4% of cases). It often occurs in a formal educational setting with teacher and/or other children present and almost never with parents or friends. Experiences with positive significance are spread roughly across all categories of occurrence.

The main negative experiences are of nervousness, embarrassment, humiliation, and criticism (either feared or actual) associated with the performing situation. These negative events almost invariably (85% of cases) had a negative influence on the subsequent involvement of these subjects with music, inhibiting either performing or listening involvement. These results will not be surprising to anyone familiar with the research literature on motivation, but they do remind us of the immense care that teachers need to take when asking children to perform. They confirm the reports of a recent British study which roundly condemned music competitions for their damaging effects on the development of young talented musicians.

To conclude the brief treatment of these results, I would like to turn back to the data on internal significance and look more closely at the content of these experiences. What kinds of things did respondents say about the musical experiences? I found that people's statements were of two kinds; statements about the *feelings* that the music evoked in them (e.g. 'The music

made a deep impression on me,' and *judgements* about the characteristics of the music itself (e.g. 'the music was very lively'). I tabulated such statements according to similarity of meaning, and also according to the age at which the music was experienced. A statistically significant age progression was found on both the feeling and the judgment dimensions.

Adults who reported no particular feeling associated with the music had an average age of 4.8 years at the time of the memory. An *enjoyment* cluster (including words such as *love, like, enjoy, excited, elated, happy*) came from children with an average age of 6.2. Then followed a cluster which I have categorized as *wonder.* (It includes such words as *enthralled, incredulous, astounded, overwhelmed, awe-struck,* etc.) The average age associated with such feelings was 8.1. Finally a small cluster of sad feelings (e.g. melancholy, sad, apprehensive) was experienced by children with an average age of 8.7 years.

A similar progression was observed in *judgements,* although fewer memories contained judgement elements overall. The average age for events containing no judgment element was 6.7. A cluster of memories used what I have called *emotional neutral* descriptive words (e.g. *fast, loud, simple,* and the anodyne *nice*). The average age of such events was 4.8. The remaining memories all occurred at an average age of around eight years. They were mainly *emotional/sensual* (e.g. *beautiful, romantic, liquid, funny*) but with a few memories mentioning particular imagery or qualities of the performance itself.

Although the number of memories is as yet rather small for definite conclusions, a suggestive pattern has begun to emerge. At around four years of age, children begin to remember music for its general level of activity (*fast, loud, simple*). Such characteristics are capable of evoking memories of general excitement (*happy, enjoy, elated*) by around six years of age. By the eighth birthday, children are remembering music for its sensual and expressive characteristics (*liquid, funny, beautiful*), and by the age of nine are regularly remembering feelings of wonder or sadness as a result.

It looks as if the memories I have gathered reveal an evolutionary process in the acquisition of musical meaning. Music has meaning at many different levels. Excitement at the activity, movement, and energy of music seems to precede wonder at the sensuous and expressive qualities of music. The age of seven seems to signal the progression to the new awareness. It is useful to look at what other changes in the awareness of music as a language are taking place at this age. Evidence mentioned earlier suggests that the age of seven is when the grasp of tonal syntax becomes particularly apparent. It is at this age that a memory advantage of tonal music emerges, as well as the ability to reject non-tonal chords. I suspect that the syntactic and semantic developments are not unrelated.

This study has been exploratory. It has shown a number of things:

1 The autobiographical memory technique is a rich and workable one.

2 Significant musical experiences are most likely to occur in particular contexts (relaxed, informal listening).

3 Such experiences have long-lasting effects on musical behaviour and involvement.

4 Negative environmental factors generally preclude the possibility of the music itself acquiring positive significance, and can inhibit the level of subsequent involvement with music.

5 Positive environmental factors do not in themselves lead to an increase in involvement with music.

6 The general nature of the experiences does not seem to depend upon the amount of prior formal tuition. Most memories with high internal significance precede formal tuition.

7 Changes occur between the ages of six and eight in the nature of the emotional and judgmental responses to music. They move from responses related to energy and excitement towards ones of wonder and beauty.

We have a long way to go before we can explain precisely what features of music are capable of eliciting these early significant responses, but we now know they are common, deeply influential, and we know something about the situations likely to encourage them. I think this is crucial information for any attempt to build an understanding of musical semantics.

## References

Hevner, K. (1936) Experimental studies of the elements of expression in music. *American Journal of Psychology*, **48**, 246–68.

Gardner, H. (1973) Children's sensitivity to musical styles. *Merrill-Palmer Quarterly of Behavioural Development*, **19**, 67–77.

Liberman, A. M., Harris, K. S., Kinney, J. A., and Lane, H. (1961) The discrimination of the relative onset time of the components of certain speech and non-speech pattern. *Journal of Experimental Psychology*, **61**, 379–88.

Locke, S. and Kellar, L. (1973) Categorical perception in a non-linguistic mode. *Cortex*, **9**, 355–69.

Siegel, J. and Siegel, V. (1977) Categorical perception of tonal intervals: musicians can't tell sharp from flat. *Perception and Psychophysics*, **21**, 399–407.

Sloboda, J. A. *The Musical Mind: The Cognitive Psychology of Music*. Oxford University Press, London, 1985.

Zenatti, A. (1960) Le développement génétique de la perception musicale. *Monographs Français Psychologique*, **17**.

CHAPTER 10

# MUSIC PSYCHOLOGY AND THE COMPOSER

## 10.1 Introduction

According to one participant at a recent conference where composers and psychologists were both present, the level of mutual interest and communication was not high. It seemed that 'whenever a psychologist spoke the composers left the room, and whenever a composer spoke the psychologists left the room'! This is not a unique phenomenon for psychology. Psychology lays great claims to being relevant in every sphere of human activity, but psychologists are not always able to convince practitioners that their researchers deserve attention. Some of this may be attributable to resistances in the practitioners, but much is also to be laid at the door of the psychologist who, typically, is not good at interpreting the implications of his or her research to practitioners, or indeed directing the research towards questions which are of direct relevance to practitioners.

This paper attempts to outline some of the issues which may prevent psychologists and composers achieving a profitable meeting of minds, examines some aspects of a framework for mutual activity, and, through this framework, attempts to point to the type of psychological findings which might be truly relevant for composers.

## 10.2 Is music psychology credible to composers?

One sobering fact for psychologists is that composers have managed very well without them for hundreds of years. Masters of composition, from Machaut through Mozart to Messiaen, have achieved profoundly significant and lasting art without the slightest help from professional psychology. Indeed modern psychological research itself would support the view that the factors that help make a good composer are (among other things) long hours of practice at composing (from as early an age as possible), belief in self, perseverence, unwillingness to accept the first solution as best, exposure to the products of other composers, the existence of mentors and supportive but searching critics, etc. Academic knowledge, be it of music theory,

philosophy, history or whatever has not had any proven effect on compositional ability. Why expect psychology to be any different?

One factor which may propel contemporary composers to search outside their own discipline for enlightenment is the breakdown of cultural consensus and traditions concerning the most basic parameters of music, its form, its instruments, its functions. Art music has broken free of longstanding constraints imposed not only by culture, but also by technology. Anything is now possible, and faced with almost infinite possibilities composers must generate their own constraints and discipline. They may rightly ask, 'is there anything that psychologists can tell us about the limits of human perception, cognition, or aesthetic and emotional response which can tell us that one avenue of exploration is more profitable than another'. In terms of theories of problem solving, composers must somehow 'reduce the search space'.

If some brave composer were to turn to the literature on psychology of music, there would be several immediate barriers to the discovery of useful Information. For instance, the vast majority of research is concerned with the way in which people perceive and handle conventional tonal music. Also, the research is mainly concerned with microstructural factors occurring within fragments of a few seconds of length. The question of large scale architecture is almost unstudied. Overarching all this is the fact that none of the research literature is *addressed* to the composer, but normally to other psychologists (who are, rightly, concerned with statistical correctness, methodology, and the development of detailed models of single aspects of a complex phenomenon).

It is not to be expected that psychology will provide composers with any ready made answers to questions which may be unasked, or even unformulated. Some mutual work is required. Composers must be able to formulate questions in precise and unambiguous ways. In their turn, psychologists must be prepared to accept composers' preoccupations and priorities without immediately rejecting them as 'unscientific' or 'too difficult'. These are not easy or trivial preconditions. What is required is a common framework.

## 10.3 What model of music is appropriate for composer–psychologist dialogue?

Some (maybe all) composers obtain satisfaction from their compositional work. The mere act of self-exploration or self-expression, the intense manipulation of compositional materials to achieve a personally meaningful end, can be quite enough motivation for the composer. Some would go further to say that such satisfaction is the only thing that can guide a composer.

Given this view, then a piece of music may be seen as an *expression* of the craft, effort, and personality of the composer. If we take such a view, then the only and ultimate arbiter of the work, its value and its purposes is the composer himself. The work's value is simply its value to the composer. Within such a framework a psychologist's role becomes rather limited—explaining how compositional activity can provide satisfaction to the composer, and, maybe, help the composer to obtain more satisfaction. This starting point, whilst logically feasible, is socially and ethically flawed. What makes it worthwhile investing effort and resources in composers (or any other creative artists) is that their work has value for others.

The exact opposite of this approach treats music entirely in terms of its *effects* on listeners. It is almost a *pharmaceutical* mode, ascribing to the music a mandatory power of bringing about certain perceptual, cognitive, or emotional responses in a listener. The question then becomes to discover what aspects of music have what effects and under what conditions. Such an approach is tempting for a psychological or physiological scientist, who can then treat music as a causal agent in the manner of the physical sciences. It does, indeed, underly much perceptual research on music in more or less explicit ways. It, however, reduces the composer to a technician manipulating the responses of a passive potential audience. This view is equally flawed, and has its uglier outcomes in the work of such agencies as the Muzak corporation, where social and behavioural control is the explicit purpose of their activities.

Neither of the above models provides a wholesome or productive meeting ground for composers and psychologists. The first makes it hard for the psychologist to make an authentic contribution, the second makes it equally hard for the composer. What alternatives do we have?

A model which many, including myself, have tried to develop in recent years is that of music as a *language*. This has worked particularly well in developing notions of music as structure (syntax), with the listener conceived of as decoding or analysing an acoustic surface into categorial elements which are then placed in hierarchically organized groups. This analysis has proved interesting and useful, at least for tonal music (see Sloboda, 1985; Lerdahl and Jackendoff, 1983). It is an advance on the pharmaceutical model because it frees the listener from the passive role of someone on whom music has 'effects', and allows that s/he may or may not learn the underlying rules of a particular genre of music: thus the existence of cultural and developmental differences.

However, the model fails in many places, particularly in the arena of semantics. It assumes that music has a 'meaning', intended by the composer and received by the listener. Success, on this model, is achieved if the listener

reconstructs the composer's underlying intention, so as to 'understand what the composer was trying to communicate'. In doing so it possibly ascribes an unrealistically tight coupling between the composer's activity and that of the listener.

I wish to propose a different model as a potentially useful meeting ground for psychologists and composers, and this is the analogy of music as *architecture or artefact.*

## 10.4 Music as architecture

We will conceive of a composer as an architect of a structure (a musical composition). Beyond him we have a range of 'middle men' who perform the function of making the achitect's plan into reality. In building work, they are the surveyors, builders, manufacturers. In music they are the publishers, performers, and instrument makers. The customer of all this work is the 'user', in music's case, often, although not always exclusively, the listener.

The steps in the creation of an artefact such as building are usually, (a) determine the function(s) the building may perform, (b) design a structure that can serve that function, (c) choose materials which will allow the structure to be made. (b) and (c) often interact, in the sense that the structure can be to a certain extent determined by the available materials.

Psychologists and composers can only begin to fruitfully interact if the composer will agree to be explicit about intended functions of this work. A composer who refuses to do this is in effect reverting to the expressive model, claiming a right to do what he likes without having to specify any outcome for others. A psychologist can be of no use to an architect if the architect does not care what use people make of his building. Some really useful dialogues have taken place between psychologists and real architects. The psychologist's contribution has sometimes been to say 'you may have intended that people would use your building in such-and-such a way. Well, I have to tell you that, because of various properties of your building, and various characteristics of the people who inhabit them, the buildings are not serving their intended function'. On some occasions the psychologists may be able to suggest relatively minor changes which make the buildings usable; in other cases it may turn out that the building is so unworkable that it has to be demolished (as happened in the case of many high-rise apartment developments of the 1960s).

But there is another side to the issue of function. An architect can never entirely predetermine all the uses that people will make of his building. Users to a certain extent choose what use to make of a building. A building

may have been designed to live in, but someone decides to run a business from it. A church may have been designed for worship, but later users turn it into a museum or concert hall. The best the architect can do is provide *affordances* (salient features) which offer opportunities for certain kinds of use. There are always degrees of freedom for a user. This is true of music too.

Psychologists may have a role to play in this respect in providing knowledge of people's ability to pick up certain kinds of affordance, and the likelihood of their doing so given their motivations, background, and preoccupations.

A building is usually a site of multiple affordances. Different people use it in different ways. Much sterile debate about 'the' meaning of a piece of music can be avoided by looking at the music as an artefact which can be used in different ways. Each listener chooses what use to make of it in accordance with the perceived affordance and his or her own agendas. A composer may well take satisfaction in designing music which can be used in many ways, including unplanned for ways, but this does not negate the requirement to design the music with some uses in mind.

## 10.5 Psychologists as reactionaries or realists?

The typical psychologist of music will be less concerned with grand designs than components. If we translate into architectural terms, he is more likely to be studying staircases than elevations. An imaginary psychologist of staircases might tell a story something like this. 'Staircases have a structure, Involving recursive incremental platforms. People exposed to this structure are able to learn to move up and down it. A skilled staircase user will adjust leg and body movements quite finely to the features of the steps—s/he is therefore highly sensitive to the structure. People build up certain kinds of expectancies. You can show this by devising a staircase which violates recursive rules (e.g. has steps of different unpredictable sizes—see Figure 10.1A below). People take longer to get up such steps and tend to trip over. People are also sensitive to higher-order recursive structures (e.g. landing placements). if landings are placed at regular intervals (Figure 10.1B) people show awareness of them (by, for instance, taking pauses more often at these places than elsewhere). Some psychologists go even further and claim that there is an optimum step size for staircases which is determined by characteristics of the human body, and that this configuration is likely to remain an invariant characteristic of all usable staircases into the indefinite future. This proposal is resisted strongly by some avant-garde architects who claim that different step sizes will provide new challenges and possibilities to users (Figure 10.1C).

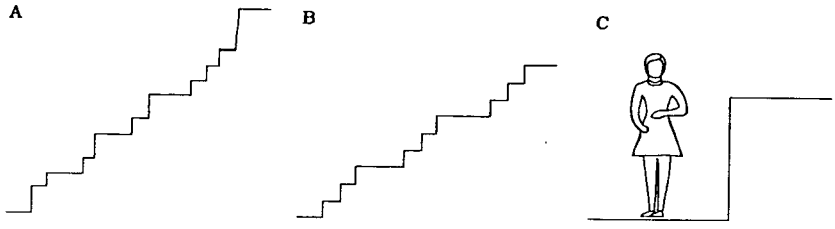



**Figure 10.1** Steps of different sizes.

This somewhat 'tongue in cheek' parody of psychological studies of, say scales and tonality, nevertheless has a number of serious points. In order to make any contribution, the psychologist has to analyse the situation with respect to function (at some level). In the absence of knowledge of any explicit statement from the architect he may take a rather low level function, such as 'going up or down a staircase'. Considerations of why a person might not want to go up or down a staircase are left to one side. In music, psychologists tend to work at a functional level of 'perceiving similarities or differences' or 'being able to remember' or 'forming perceptual groups', showing what factors in music promote or hinder such functions. The question, 'why should the music need to be remembered?' is not often asked. Some composers may not intend that the details of a composition be remembered by a listener, but even where they do, we need to know why. What higher functions is memory in music serving? On these grounds it would seem essential for psychologists to be enquiring of composers 'What are you hoping that your music will allow people to do with it?' and of listeners 'Why are you listening to this music, what are you using it for?' Few psychologists I know of are doing this. Ethnomusicologilsts such as Merriam (1964) have come closest to systematic treatment of this area, although Merriam's distinction between functions and uses of music may not be a particularly useful one for psychological analysis. In this paper, I am using the two words interchangeably.

Another point of the parody is to show that psychologists are more able to assess human performance than human competence. They look at what people do. This is realistic. If you want to change people you need to know where they are starting from. It may also be reactionary, in that there is a strong temptation to use what people are currently doing to draw strong inferences about what they are able to do. Thus you will find psychologists of music claiming that the tonal system is the best possible system of pitch organization, the human voice is the best possible instrument, etc. I take a middle view. Change is possible, but it proceeds slowly and incrementally,

each step being only a small change from the preceding one. So called cultural revolutions are, in fact, surface perturbations which conceal underlying continuities. Change through evolution seems to characterize almost every historical phenomenon from the differentiation of species to the growth of ideas. Psychologists may be able to remind composers that people need preparation before they can handle radical change.

## 10.6 The functions of music

Given the above argument, it follows that the functions of music will be relatively stable over time. New functions will arise, and old ones will drop out, but at a slow rate. Functions will be largely determined by cultural and psychological requirements of listeners. Composers may be able to invent radically new forms and structures: they are hardly capable of inventing new functions. Of course, we are here talking of generalized functions. Each specific composition (building) has unique features which may be related (or may not be related) to the precise requirements of its location, intended users, etc. This is why there is always new work for composers (and architects). Thus, at a given cultural point, functions of music are not arbitrary, but maybe (in principle) enumerated by observation. This is an empirical process in which a composer has no particularly privileged position.

Functions, or goals, are also capable of indefinite hierarchical embedding. Disagreements about ultimate functions can legitimately arise. For one composer, it may be a satisfactory goal to have some listeners like his composition. Others will argue that a person only likes something if it is fulfilling some further function, and that we cannot build for liking without understanding what that further function is. By analogy, it may be argued incoherent to say that a person likes his house without being able to specify in what way it fulfils useful functions for him. I take the latter view, and see as non-explanatory any account of music which posits as ultimate function the achievement of affective or aesthetic states (such as liking, pleasure, or the experience of some emotion). Affective states are concomitants or symptoms of processes of behavioural or cognitive adaptation and change, and have no value in themselves.

This may be made clearer by an example. It is very likely that music has served mnemonic functions in many cultures. That is, the music serves as a means of structuring the recall and organization of some other behaviour. The use of music to help recall of words is a paradigm example of this. Songs aid retention of culturally important material, particularly in non-literate or semi-literate societies. The ability of the music to guide or remind or help the user anticipate, will lead to a sense of satisfaction or pleasure in the

music. A composer will need to know what features promote music's mnemonic functions if he is to compose apppropriate music for this purpose. A person's aesthetic or emotional reaction to the music in respect of this function will be entirely based on the degree to which the function is satisfactorily fulfilled. My strong claim is that all such reactions are contingent on their fulfilling some psychological function. There is no such thing as 'just' liking a piece of music.

Of course, many functions of music may be grasped by composers and listeners only intuitively and implicitly. It is not expected that a full formal description will always be forthcoming. A listener may gain intense value from a musical experience without necessarily being able to articulate and explain the basis of this value. In other cases, the listener may be clear enough about the outcome (e.g. 'It renewed my sense of hope') but be unable to articulate the mechanism by which this outcome was achieved. Psychology in general (and cognitive psychology in particular) has much experience in formalizing and uncovering those processes which seem to be hard to consciously report.

There are two main reasons for self-report limitations. One is that the processes are truly subconscious, rapid and automatic. Examples of this are mechanisms of reading and perceptual analysis. They can be uncovered by relating task performance to structural factors in controlled stimuli. Another reason is that the listener quickly forgets the precise sequence of mental events. This is true particularly of complex time-extended activities such as problem-solving, day-dreaming, and, arguably, listening to music (see Ericsson and Simon 1980). A few music psychologists, including myself, are beginning to explore ways of allowing people to report on their experiences during extended music listening, either retrospectively or concurrently. I am looking at listener's ability to report either retrospectively or concurrently on moments of high emotional 'charge' during pieces of music. This is being linked to structural analysis of these moments. Preliminary work is confirming the insights of Meyer (1956) and others, that patterned or motivated tension and resolution seem to promote significant psychological outcomes for listeners. Indeed, an initial analysis of a small data-base of musical 'moments' which provoke strong physiological reactions shows strong and systematic relationships between the type of reaction (tears, shivers, or racing heart) and specific structural features of these moments. Full details of this research will be reported in due course.

One important line of psychologically informed theorizing is being undertaken by Lerdahl (1988), who is trying to articulate the defining characteristics of music that is capable of being grasped as structure (and

thus as capable of supporting motivated tension and resolution). He arrives at a number of general postulates which would include tonal music but exclude serial music such as Boulez' *Le Marteau sans Maitre*. The postulates would not, however, exclude all novel compositional forms and devices, and he himself is exploring the possibility of using *timbre* in ways which are consistent with his postulates (Lerdahl, 1988).

Another equally important line of thinking is the attempt by some psychologists (e.g. Dowling and Harwood, 1986; Gayer and Mandler, 1987; Ortony, Clore and Collins, 1988) to articulate general theories about the cognitive antecedents and consequences of emotion. This work promises the opportunity of establishing formal links or analogies between musical events which promote emotion and non-musical life events which have similar effects. It may be easier to see the functions of the emotions in 'external' life events than in the 'internal' world of music.

## 10.7 Emotions in music serve non-trivial personal purposes for listeners

Taking the assumptions of modern cognitive science that many emotions serve rational purposes rather than being embarrassing lapses from functionality, it is possible to make a number of assumptions about emotional functions in music.

(a) *Each emotion experienced in response to music is matched by an identical emotional response to a (potential) analogous non-musical situation.* This statement is asserting that there are no emotional response which are unique to music. It is also asserting that a particular non-musical analogue is not necessarily experienced frequently (or at all) by a given individual. For instance, music may be accompanied by a sense of relief from anxiety that is seldom experienced in the extra-musical life of a particular individual.

(b) *The experience of emotion arises when substantial cognitive re-appraisal is required, and accompanies such re-appraisal.* This statement asserts that emotions function with respect to a cognitive system which is monitoring situations for changes of various sorts. They signal the detection of particular classes of change and are, in some yet to be fully understood ways, implicated in mobilizing further cognitive resources to handle these changes.

(c) *The need for re-appraisal is precipitated by the violation of current goals, plans, or assumptions. The situation will continue to be emotion bearing until such*

*re-appraisal is complete.* This statement defines important types of change that are likely to be emotion-bearing. Such violations can be either negative (the death of a loved one) or positive (winning a lottery).

(d) *Such re-appraisal is promoted by any context which suggests the possibility of an alternative or a resolution to the precipitating situation.* It is a common observation that a small kindness can often restore the morale of someone who is discouraged or disillusioned. It is not that the kindness itself solves the major problems, but it reminds the recipient of 'another world', and can motivate a renewed effort to tackle the problem. In a similar way, many prisoners of war (or conscience) have testified to the immense value of tiny scraps of Information or encouragement from the outside world in reviving hope and energy. Some formal therapies are based on a systematic application of this phenomenon.

(e) *Such contexts are typically less available than complete re-appraisal would require. People, therefore, accumulate a backlog of unfinished business.* A colloquial way of expressing this statement is that life events leaves most people with emotional scars. If, for instance, a person continues to feel hurt or angry about some long past event, this signals that the individual has not yet been able to deal successfully with all the consequences and implications of that event.

(f) *People will seek and use appropriate subsequent context to continue the unfinished re-appraisals. Such context will contain (a) some reminder or analogue of the initial emotion-provoking disturbance and (b) a demonstration of the resolution of that disturbance.* A clear example of such a context would be the 'story with a happy ending'. A huge number of the world's most loved and retold stories are based on the theme of an individual or group struggling against adversity to win through to a goal.

(g) *Certain musical structures meet these criteria. Cultural or pan-cultural commonalities of experience tend to make such structures widely effective, rather than individualistically so.* The discovery and specification of such musical structures is a task for the future. This statement suggests that few emotional responses to music are 'pure'. In the sense of being detached from an individual's life experiences. However, it relates life experiences to emotional response in a completely different way to the 'associationistic' model, where a person re-experiences the same emotion that he or she happens to have felt at the time of an earlier hearing of the music. Such associations certainly occur and are used by composers, but many, perhaps most of these responses, are completely individualistic and cannot be systematized. It suggests, rather,

that people will look to different pieces of music at different times of life and in different situations because they contain material relevant to particular types of re-appraisal.

(h) *Engagement with music can thus be seen as Important and purposeful work. Conceptions of music as recreation or diversion are short sighted and incomplete.* To make explicit connections between music and non-musical ends is not in any way to devalue the composer's work. For one thing, music may supply contexts which are rare or unavailable in other situations. Secondly, as stated before, music can serve multiple purposes, and no one purpose exhausts music's significance. Music is, therefore, not a substitute for or 'just like' something else.

## 10.8 Conclusion

I have suggested that a fruitful point of contact for musicians and psychologists is the architectural analogy, which focuses on articulating the functions of a composition, and then specifying the characteristics of the affordances which may allow these function to be fulfilled. Some functions may necessitate the possibility of the listener making discriminations and comparisons, of classifying events along dimensions (of relative tension or prominence), and of recalling earlier events. Psychologists are in a good position to formalize for composers the properties of sound structures which encourage or do not encourage such activities, and also to help assess the degree to which listener actually respond in intended ways. It is, however, up to psychologists to prove that their work is relevant and powerful enough to replace or supplement the time-honoured strategy of composers whereby they use themselves and their own responses as the means of assessing the adequacy of what they are doing. This may require a more intimate and personal relationship between individual composers and psychologist than has up to now been achieved, and conferences such as this may be an important catalyst to the development of such relationships.

## References

Dowling, W. J. and Harwood, D. L. (1986) *Music Cognition*. Academic Press, New York.

Ericsson, K. A. and Simon, H. A. (1980) Verbal reports as data. *Psychological Review*, **87**, 215–251.

Gayer, W. W. and Mandler, G. (1987) Play it again Sam: on liking music. *Cognition and Emotion*, **1**, 259–282.

Lerdahl, F. and Jackendoff, R. (1983) *A Generative Theory of Tonal Music.* MIT Press, Cambridge, Mass.

Lerdahl, F. (1988) In: Sloboda J (ed) *Generative Processes in Music. The psychology of performance, composition, and improvisation. Cognitive constraints on compositional systems.* Oxford University Press, London.

Merriam, A. P. (1964) *The Anthropology of Music.* Northwestern University Press, Chicago.

Meyer, L. B. (1956) *Emotion and Meaning in Music.* University of Chicago Press, Chicago.

Ortony, A., Clore, G. L., and Collins, A. (1988) *The Cognitive Structure of the Emotions.* Cambridge University Press, New York.

Sloboda, J. A. (1985) *The Musical Mind: The Cognitive Psychology of Music.* Oxford University Press, London.

CHAPTER 11

# EMPIRICAL STUDIES OF EMOTIONAL RESPONSE TO MUSIC

There is a general consensus that music is capable of arousing deep and significant emotion in those who interact with it. In a study of autobiographical memories of musical events (Sloboda, 1989a), one subject recalled the following event from the age of 7:

> I was sitting in morning assembly at school. The music formed part of the assembly service. . . . The music was a clarinet duet, classical, probably by Mozart. I was astounded at the beauty of the sound. It was liquid, resonant, vibrant. It seemed to send tingles through me. I felt as if it was a significant moment. Listening to this music led to me learning to play first the recorder and then to achieve my ambition of playing the clarinet. . . . Whenever I hear clarinets being played I remember the impact of this *first* experience. (p. 37)

This study asked a wide variety of adults to recall any memories from the first 10 years of their lives that involved music in any way. There was no attempt by the experimenter to suggest particular types of memory. Forty-four out of 113 memories so produced (39%) shared an important characteristic with the aforementioned example: a valued emotional experience that was derived from an awareness of some aspect of the musical sound itself. Although there were other memories in the sample where the emotional tone of the experience (positive or negative) derived from the context in which the music took place, the study showed that sustained involvement with music was more likely in those subjects who could recall a musical event of the positive type.

Because such experiences are widespread and seem to be important motivators for engagement with music, the study of these experiences deserves a more central position in the psychology of music than it currently enjoys. The purpose of this chapter is to explore how empirical science may study this type of phenomenon and to summarize some of what is known about it.

## 11.1 The cognitive content of the emotional experience

### 11.1.1 Free verbal response—adult experience

In a previously unpublished study, I asked 67 regular listeners to music to describe in their own words the nature of their most valued emotional experiences of music. Although every account was different in detail, some common themes emerged. The most commonly mentioned concept was that of music as *change agent (n = 41)*. Some common comments were:

- 'Music relaxes me when I am tense and anxious'
- 'One feels understood and comforted in pain, sorrow, and bewilderment'
- 'Involvement in music detaches me from emotional preoccupations'
- 'Through hearing emotions in someone else's music, it is possible to feel that emotions are shared and not your burden alone'
- 'Music motivates and inspires me to be a better person (e.g. more agreeable and loving).'

Common to these examples is the characterization of music as offering an alternative perspective on a person's situation, allowing him or her to construe things differently.

A second cluster of responses focused on the notion of music as promoting the *intensification* or *release* of existing emotions *(n = 34)*. These responses included:

- 'Music releases emotions (e.g. sadness) that would otherwise be bottled up'
- 'Music helps me discover what I am actually feeling'
- 'Music reconnects me to myself when my emotions are ignored or suppressed'
- 'Music makes me feel more alive, more myself'
- 'Music can provide a trigger for the outlet of my emotions concerning memories of pleasurable or painful experiences in my past.'

Common to these examples is the notion that music does not create or change emotion; rather it allows a person access to the experience of emotions that are somehow already 'on the agenda' for that person, but not fully apprehended or dealt with. Gabrielsson (1989) reports a similar finding in a study where he asked 149 people to describe the 'strongest (most intense, deep-going) experiences' they ever had to music. A significant number of

respondents reported 'feelings that this music deals with myself; it reflects or clarifies my feelings and situations.'

### 11.1.2 Free verbal response—childhood experience

Gardner (1973) has demonstrated an age progression in the kind of adjectives children spontaneously use to describe the character of music they hear. Children aged 6 to 14 were asked to say whether or not two musical extracts came from the same composition and were then asked to verbally justify their choices. The 6-year-old children tended to base their justifications on such simple dimensions as fast/slow or loud/soft. Eight-year-olds used metaphors drawn from outside music, such as 'peppy,' 'dull,' 'churchy,' and so forth. Neither this study nor any other known to the author has directly asked children to describe their emotional experiences to music.

The previously mentioned retrospective study of Sloboda (1989a) provides some indirect evidence of an age progression in the nature of emotional response to music. For each memory, adults were asked to describe the nature and significance of the event and also to recall the age at which the experience had occurred. Many subjects spontaneously used emotion words or concepts to describe their experiences. The emotion words were divided into three broad categories: an *enjoyment* category (including love, like, enjoy, excited, elated, and happy), a *wonder/surprise* category (including enthralled, incredulous, astounded, overwhelmed, and awestruck), and a *sadness* category (including melancholy, sad, and apprehensive). There was a significant difference in the mean age associated with the recalled experiences relating to these categories. For enjoyment experiences the mean age was 6.2 years, for wonder/surprise experiences the mean age was 8.1 years, and for sad experiences 8.7 years.

This study also provided independent corroboration of Gardner's (1973) results. Descriptions of the music in terms of simple characteristics (e.g. fast, loud) were associated with experiences at a mean age of 4.8. Descriptions in terms of metaphor and character (e.g. liquid, funny, romantic) came at a mean age of 8.0. Gardner's results may have been partly due to the expanding vocabulary of the children he tested rather than a fundamental change in experience. However, the Sloboda data—coming as they do from adults with sophisticated vocabularies—are less plausibly explained in a similar way. It appears increasingly likely that there is a fundamental change in the experience of emotion to music and in its cognitive construal between the ages of 6 and 8. At the heart of this change is a heightened response to aspects of music which strike children as surprising or unexpected. This may be related to a widely reported increase in sensitivity to aspects of tonality at about this age (Dowling, 1982;

Imberty, 1969; Thackray, 1976; Zenatti, 1969). The more one has understood and internalized the implicit 'grammar' or 'rule system' of a musical language, the greater the possibility for surprise at novelty.

### 11.1.3 Forced verbal response—adult experience

The great majority of studies on emotional characteristics of music have asked subjects to ascribe such characteristics to pieces of music by choosing an adjective from a checklist (e.g. Hevner, 1936; Scherer & Oshinsky, 1977; Wedin, 1972). These studies show that, by and large, adults in a given musical culture agree on the broad characterization of a musical passage. The relevance of these studies to emotional *experience*, however, is not proven. It is possible to make character judgments on the basis of conventional characteristics without experiencing any emotion whatsoever.

A study by Waterman (personal communication, March 1990) approached the issue more directly by asking 76 college students to indicate on a checklist of emotions which ones they have actually experienced to music. The checklist consisted of 25 emotion words or phrases, each representing a major emotion type in the theory of Ortony, Clore, and Collins (1988). This theory proposes that experienced emotion arises through cognitive appraisal of situations construed as either *events, actions of agents,* or *objects.* Until such construal takes place, no emotion is possible. The most undifferentiated emotion arises from simply placing a positive or negative valence on a situation so appraised. The basic event-related emotion pair is PLEASED–DISPLEASED. The corresponding pair for actions of agents is APPROVE–DISAPPROVE, and that for objects is LIKE–DISLIKE.

Within the theory, other emotions are produced by appraising the situation to finer levels of differentiation. For instance, within the event branch, an event can be appraised as relevant to oneself or relevant to someone else. It can be further appraised as a present event or one in the future. Each level of differentiation is achieved by the consideration of an additional variable: The more variables added, the more specific is the emotion and the more specific the eliciting situation. A relatively unspecific event-based emotion is that of being DISPLEASED at the occurrence of an event judged UNDESIRABLE. A typical word for such emotion would be SADNESS. A more specific event-based emotion is that of being DISPLEASED at the nonoccurrence of a event whose PROSPECT was judged as DESIRABLE for ONESELF. A typical word for such an emotion would be DISAPPOINTMENT.

Waterman found that the emotions reported most often did not fall under any one appraisal category (events, actions of agents, or objects). People frequently experienced emotions relevant to appraisals of music as event (for instance, 72% of the subjects reported experiencing HOPE in

response to music), action of agent (53% experienced ANGER), and object (92% experienced LIKING). However, there was a significant negative correlation between the proportion of subjects experiencing an emotion and its specificity in the theory, in terms of the number of cognitive variables requiring appraisal. Less specific emotions such as JOY (71%) were experienced more often than more specific emotions such as RESIGNATION (33%). The full list of emotion labels used in the study is given in Table 11.1, together with the percentage of respondents experiencing that emotion and the number of local appraisal variables relevant to that emotion in Ortony *et al.*'s theory.

**Table 11.1** Proportion of subjects who have felt each emotion to music, and the number of local variables influencing the intensity of each emotion, according to Ortony, Clore, and Collins (1988). Data are from Waterman (personal communication, March 1990)

| Emotion | % | Number of local variables influencing emotional intensity | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 |
| Sadness | 96.2 | ■ | | | |
| Joy | 93.4 | ■ | | | |
| Liking | 92.1 | ■ | ■ | | |
| Appreciation | 86.8 | ■ | ■ | | |
| Dislike | 81.6 | ■ | ■ | | |
| Satisfaction | 78.9 | ■ | ■ | ■ | ■ |
| Suspense | 77.6 | ■ | ■ | ■ | |
| Hope | 72.4 | ■ | ■ | | |
| Anger | 52.6 | ■ | ■ | ■ | |
| Hopelessness | 52.6 | ■ | ■ | ■ | |
| Disappointment | 52.6 | ■ | ■ | ■ | ■ |
| Pride | 51.3 | ■ | ■ | ■ | |
| Relief | 50.0 | ■ | ■ | ■ | ■ |
| Sympathy | 48.7 | ■ | ■ | ■ | ■ |
| Happiness for | 46.1 | ■ | ■ | ■ | ■ |
| Resignation | 43.3 | ■ | ■ | ■ | |
| Fear | 42.1 | ■ | ■ | | |
| Remorse | 39.5 | ■ | ■ | ■ | ■ |
| Gratitude | 32.9 | ■ | ■ | ■ | |
| Resentment | 30.3 | ■ | ■ | ■ | ■ |
| Self-satisfaction | 30.3 | ■ | ■ | ■ | ■ |
| Fears confirmed | 23.7 | ■ | ■ | ■ | ■ |
| Shame | 17.1 | ■ | ■ | ■ | |
| Reproach | 15.8 | ■ | ■ | ■ | |
| Gloating | 13.2 | ■ | ■ | ■ | ■ |

Waterman's data demonstrate that most adults have experienced a range of emotions to music and that frequency of response is less a function of the type of cognitive appraisal underlying the emotion than of its cognitive specificity. There are, however, major problems with any empirical approach which limits subjects' responses to a set of experimenter-determined categories. A more direct way of eliciting data would be to ask subjects to specify a nonmusical situation in which they have experienced the same emotion as during a particular episode of music listening. It may then be possible to map formal characteristics of the nonmusical situation directly onto formal characteristics of musical events without any need to employ the vocabulary of emotional terms at all. This approach is currently being explored.

## 11.2 The antecedents of emotional experience

### 11.2.1 Factors external to the music

It is clear that the same event, be it musical or otherwise, does not always result in the same emotional experience. For example, I can listen to the same recording on two different occasions and be moved to tears on one of them, while remaining completely detached on the other.

There are several possible reasons for this inherent variability. One is that the same piece of music may be appraised on different criteria. A musical passage appraised as event may incline a listener to an event-based emotion such as sadness. The same piece appraised as the action of an agent may incline a listener towards an emotion such as gratitude.

Another possibility is that one's prevailing mood determines the extent to which a piece of music will lead to strong emotional experience of particular sorts. It may be hard to experience grief or sadness in response to music when one's prevailing mood is cheerfulness, even though one is carrying out the relevant cognitive appraisals of the music, which in other circumstances would lead to sadness.

A third possibility is that there may be some emotional or mood states which simply preclude any relevant appraisal-based emotion to the music from occurring at all. One of the most striking features of the Sloboda (1989a) study was the almost complete absence of differentiated responses to the music itself when the context was being appraised as negative. The most typical negative context was a situation of threat, anxiety, or humiliation, brought about by being placed in a situation where one's performance was being assessed. Such situations included singing or playing in front of an audience or a teacher. They often included fear of verbal or physical punishment (and sometimes its realization). For many subjects these contexts

had long-lasting effects, leading to 'refusing to have anything to do with music' or 'considering myself completely unmusical.'

The evidence of the study is that the likelihood of experiencing positively valued emotional responses from music is a direct function of the degree to which the subject feels relaxed and unthreatened. If someone is feeling frightened, anxious, or under threat, this usually means that there are factors in the environment which are dominating the appraisal system, probably because they impinge on significant personal goals (e.g. preservation of self-esteem). Under these circumstances the music itself will not be likely to yield appraisal-based emotions. This may be because the appraisals are not undertaken at all (the attention is on the threat-provoking context) or because the appraisals are in some way 'blocked' from access to the system which produces emotional experience.

### 11.2.2 Factors intrinsic to the music

Although it is well established that people respond emotionally to music, little is known about precisely what it is in the music that they are responding to. A recent study (Sloboda, 1991) attempted to establish something about the specific types of musical events that are associated with physically based concomitants of emotional response to music. Physical reactions such as crying, shivering, a racing heart, and so forth avoid some of the problems associated with verbal emotion labels that were discussed previously. It is very difficult to be mistaken about whether you cried or not to a piece of music. Such reactions are stereotyped, memorable, distinct from one another, and shared by all humans regardless of culture and vocabulary. They are arguably more closely connected to the experience of emotion than verbalizations which may be infected with rationalizations.

In a study by Sloboda (1991), subjects were asked to specify particular pieces of music to which they could recall having experienced any of a list of 12 physical manifestations commonly associated with emotion. Having identified such pieces, they were then asked to specify the location within the music that provoked these reactions. They were encouraged to do this, where possible, with reference to a score or a recording, but since this was a postal questionnaire study, no direct check on method was made. Surprisingly, and in contradistinction to Meyer's (1956) misgivings, a significant proportion (about one third) of the people were able to locate their reaction within a theme or smaller unit. Those specific segments for which a published score was available (i.e. the classical citations) were classified according to the musical features they contained, and according to the emotional reaction they provoked.

**Table 11.2** Music-structural features associated with physical-emotional responses

| Feature | | Number of musical passages provoking a response | | |
|---|---|---|---|---|
| | | TEARS | SHIVERS | HEART |
| 1 | Harmony descending cycle of fifths to tonic | 6 | 0 | 0 |
| 2 | Melodic appogiaturas | 18 | 9 | 0 |
| 3 | Melodic or harmonic sequence | 12 | 4 | 1 |
| 4 | Enharmonic change | 4 | 6 | 0 |
| 5 | Harmonic or melodic acceleration to cadence | 4 | 1 | 2 |
| 6 | Delay of final cadence | 3 | 1 | 0 |
| 7 | New or unprepared harmony | 3 | 12 | 1 |
| 8 | Sudden dynamic or textural change | 5 | 9 | 3 |
| 9 | Repeated syncopation | 1 | 1 | 3 |
| 10 | Prominent event earlier than prepared for | 1 | 4 | 3 |
| Total number of musical passages | | 20 | 21 | 5 |

From *Music Structure and Emotional Response: some Empirical Findings*, by J. Sloboda, *Psychology of Music*, 1992. Copyright 1991 by Society for Research in Psychology of Music and Music Education. Adapted by permission.

Table 11.2 shows the main results of this analysis. Twenty musical passages provoked the cluster of responses labeled TEARS (i.e. crying, lump in the throat). The majority of these passages contained melodic appoggiaturas and melodic or harmonic sequences. Twenty-one passages provoked the cluster of responses labeled shivers (i.e. goose pimples, shivers down the spine). The majority of these contained a new or unprepared harmony. Only 5 passages provoked HEART reactions (i.e. racing heart and pit-of-stomach sensations). These were associated with repeated syncopations and prominent events occurring earlier than prepared for.

A more recent unpublished analysis of the popular and jazz citations (by transcription from audio tracks) has confirmed this basic picture. The same features tend to be operative, although the picture is somewhat obscured due to the fact that combined physical reactions are more common in these passages than in classical music. That is to say, subjects frequently reported TEARS, Shivers, and HEART reactions simultaneously to the same segment of music.

Figure 11.1 shows an example of an excerpt that provoked tears. It is the opening passage from Albinoni's *Adagio for Strings*. Note that the first 7-note melodic phrase contains 3 consecutive appoggiaturas, C–B<flat>, A–G, G–F#. The phrase is then repeated in sequence form (exact repetition at one higher scale step). This is followed by a second phrase which is repeated twice more in sequential fashion. What seems to characterize this, and other TEARS-provoking passages, are the successions of harmonic

tensions or dissonances which are created and resolved within a structure which has a high degree of repetition and implication-realization (see Narmour, 1977). In general terms, an implication is set up when a musical segment contains within it some parameter, such as a pitch movement. The implication is realized if the parameter is repeated. For example, in Figure 11.1 the initial scalar movement D–C sets up implications for further downward scalar motions, which are realized. A sequence realizes implications that a theme, once stated, will be repeated. If the first repetition is at a certain pitch distance from the first statement, then the implication is that a second repetition will repeat this pitch jump. Psychologically, passages with strong implication-realization are to some extent predictable. In cases like that of Figure 11.1 it is possible that the predictable but repeated appoggiaturas, which 'tease' the sense of consonance, are responsible for the particular form that the emotional response takes.

Figure 11.2 shows an example similar to an excerpt that provoked SHIVERS (copyright restrictions preclude citing the actual example). It is characterized by a sudden key shift from E to C# in the context of a rising sequential pattern based on E, then F#, and then G#. In terms of music analytic theories of expectancy/implication (see, for instance, Meyer, 1956, 1973; Narmour, 1977; Schmuckler, 1989), the sequential progression E–F# sets up an implication for a progression G# (which is fulfilled). At the same time the harmonic progression (successive chords within E-major tonal space) sets up an implication for further chords within this tonal space (which is violated). This dual characteristic of an event being both expected (at one level or according to one criterion) and unexpected (at another level or on another criterion) seems to be shared by several of the SHIVERS-provoking passages. For example, an enharmonic change fulfills one

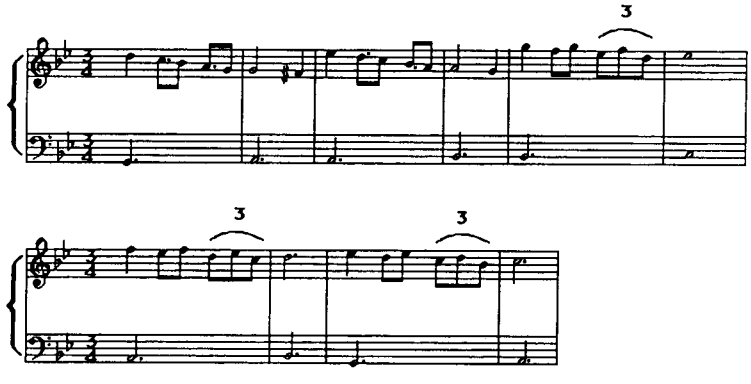



**Figure 11.1** Opening theme from Albinoni's *Adagio for strings*.

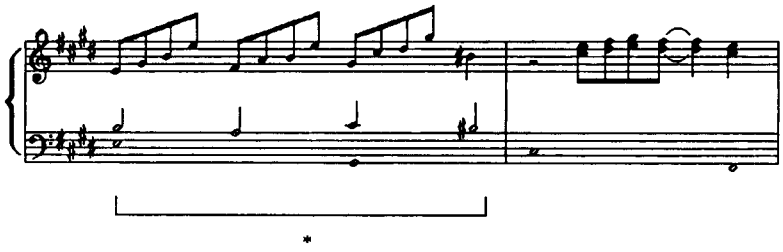


**Figure 11.2** An example similar to an excerpt that provoked shivers.

expectancy (melody note stays the same) while violating another (harmonic function of melody note changes). In other cases, the response seems to be associated with a significant change in some characteristic of the music at or near a major structural boundary (e.g. new phrase or verse in a 32-bar song). If there is some psychological dissociation or independence between the various mental analyses that may be carried out on music, then it could well be that some emotional responses come about when there is a mismatch between the output of two processing units or 'modules.' The boundary may be expected according to the computations of one 'module' (e.g. a rhythmic/metrical analyzer assigning events a hierarchical position within a phrase-structure tree), but the specific parameter change can still 'surprise' another 'module' (attuned to, say, texture or dynamic).

Figure 11.3 shows a passage provoking heart reactions. This is bar 191 from the last movement of Beethoven's *Piano Concerto No. 4* in G major, op. 58. Here the phrase structure of the whole movement is built on multiples of even numbers of bars (2, 4, 8, etc.). The piano solo starting at bar 184 reinforces this. A new phrase starts at 188, and there is an implication that the next phrase will commence at bar 192. Instead it arrives at 191 with a sudden increase in dynamic. Each phrase beginning functions as an accent within the phrase structure. Here, then, is a case of an expected accent arriving earlier than it 'should.' This is also the principal feature of syncopation, another provoker of heart reactions.

In each case, it is clear that expectancies and violations of expectancy are playing a major part in the promotion of emotional reactions to the music. These results are an empirical vindication for the music-theoretic tradition whose main inspiration is L. B. Meyer. However, what the data demonstrate is that the nature of the emotional experience depends on the particular character and pattern of an interlocking set of implications and their realization or nonrealization. A major task for the future is therefore the specification of the precise way in which emotions map onto different kinds of musical events. A theory providing at least the differentiation of Ortony *et al.*'s will be necessary.

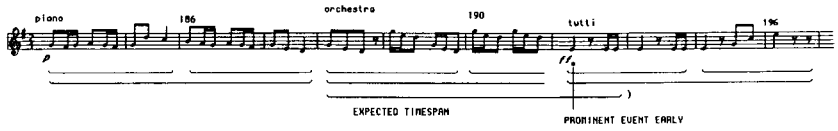



**Figure 11.3** Main features surrounding bar 191, last movement, in Beethoven's *Piano Concerto No. 4* in G major, op. 58. From *Eulenberg score no. 705.* Copyright by Ernest Eulenberg, Ltd. Adapted by permission.

## 11.3 Conclusions

The empirical study of emotional responses to music is in its infancy, despite several recent attempts to construct a theoretical framework for considering emotions in music (e.g. Dowling and Harwood, 1986; Sloboda, 1989b). The subject poses serious methodological and theoretical problems. Since there is no generally accepted theory of the emotions and how they interact with cognition, I believe that open-ended empirical investigations with a strong element of natural history continue to be the most profitable way of exploring this area at this time. Music psychology is already a successful interdisciplinary study, with major advances at the junctions of music theory and cognitive science. A satisfactory incorporation of the study of emotion into this work will require methods and theoretical approaches drawn from a wider range of subdisciplines. I hope this chapter has given at least a hint that this is a promising area ripe for further development.

## References

Dowling, W. J. (1982) Melodic information processing and its development. In D. Deutsch (Ed.), *The psychology of music.* New York: Academic Press.

Dowling, W. J. and Harwood, D. L. (1986) *Music cognition.* New York: Academic Press.

Gabrielsson, A. (1989) Intense emotional experiences of music. In *Proceedings of the First International Conference on Music Perception and Cognition* (pp. 371–376). Kyoto, Japan: The Japanese Society of Music Perception and Cognition.

Gardner, H. (1973) Children's sensitivity to musical styles. *Merrill-Palmer Quarterly of Behavioural Development,* **19**, 67–77.

Hevner, K. (1936) Experimental studies of the elements of expression in music. *American Journal of Psychology,* **48**, 246–268.

Imberty, M. (1969) *L'acquisition des structures tonales chez l'enfant* [The acquisition of tonal structures in the child]. Paris: Klinksiek.

Meyer, L. B. (1956) *Emotion and meaning in music.* Chicago: University of Chicago Press.

Meyer, L. B. (1973) *Explaining music.* Berkeley: University of California Press.

Narmour, E. (1977) *Beyond Schenkerism: The need for alternatives in music analysis.* Chicago: University of Chicago Press.

Ortony, A., Clore, G. L., and Collins, A. (1988) *The cognitive structure of the emotions.* New York: Cambridge University Press.

Scherer, K. R. and Oshinsky, J. S. (1977) Cue utilization in emotional attribution from auditory stimuli. *Motivation and Emotion,* **1**, 331–346.

Schmuckler, M. A. (1989) Expectation in music: Investigation of melodic and harmonic processes. *Music Perception,* **7**, 109–149.

Sloboda, J. A. (1989a) Music as a language. In F. Wilson and F. Roehrnann (Eds.), *Music and child development.* St. Louis, MO: MMB, Inc.

Sloboda, J. A. (1989b) Music psychology and the composer. In S. Nielzen and O. Olsson (Eds.) *Structure and perception of electroacoustic sound and music.* Amsterdam: Elsevier.

Sloboda, J. A. (1991) Music structure and emotional response: Some empirical findings. *Psychology of Music,* **19**, 110–120.

Thackray, R. (1976) Measurement of perception of tonality. *Psychology of Music,* **4**, 32–127.

Wedin, L. (1972) A multidimensional study of perceptual–emotional qualities in music. *Scandinavian Journal of Psychology,* **13**, 1–17.

Zenatti, A. J. (1969) *Le développement génétique de la perception musicale* [The genetic development of musical perception] (Monographie Françaises de psychologie No. 17). Paris: Centre Nationale des Recherches Scientifiques.

CHAPTER 12

# EMOTIONAL RESPONSE TO MUSIC: A REVIEW

## 12.1 Introduction

The reason that many people engage with music, as performers or listeners, is that it has power to evoke or enhance valued emotional states. Studies supporting this general assertion are reviewed in Section 12.2. In the light of this one might have expected the study of emotion to be central to the psychology of music. This has not been the case. There are conceptual and methodological reasons for this. Conceptual issues are discussed in Section 12.3, of which the primary issue is the inherent variability of emotional response, from listener to listener, and within the same listener on different occasions. Section 12.4 addresses some of the methodological issues, arising primarily from difficulties of measurement.

Notwithstanding the difficulties inherent in research on this topic, some progress has been made, and Section 12.5 outlines some major findings, including some evidence that a major subset of emotional responses are cued by confirmations and violations of expectancy within the musical structure, in line with the predictions arising from the theoretical proposals first articulated by L. B. Meyer (1956).

## 12.2 The importance of emotions to music participants

### 12.2.1 Music listeners

Music has the power to evoke strong and significant emotions in a wide cross-section of the general population (Gabrielsson and Lindstrom, 1993). For instance, in a survey of the crying behaviour 331 North American adults (Frey, 1985) it was discovered that 63 (8%) of the 800 crying episodes reported were triggered by music.

When 67 regular listeners of music were asked to describe the nature of their most valued responses to music (Sloboda, 1992), two common themes emerged. The first theme was of music as change agent, in which it helped the individual to move from a less desirable to a more desirable psychological state. This was evident in statements such as 'music relaxes me when

I am tense and anxious', 'music motivates and inspires me to be a better person (e.g. more agreeable and loving)'. The second theme was of music as promoting the intensification or release of existing emotions. This was indicated by statements such as 'Music helps me discover what I am actually feeling', and 'Music reconnects me to myself when my emotions are ignored or suppressed through sheer busyness'.

### 12.2.2 Music performers

In an ongoing project on communication of emotion in piano performance, ten professional pianists were asked to what extent their own emotional reactions to the music they performed were important in helping them determine their own interpretation. All of them judged their own responses to be important. One pianist said 'I choose repertoire that I can put my whole emotional energy into. I couldn't play without it'. Another judged the effectiveness of a performance by taping it and listening to it 'as audience' to see if it had the intended emotional effect.

Strong emotional responses to music appear to underlie the decisions of many individuals to commit to careers in music. A study of childhood musical experiences (Sloboda, 1989a) showed that individuals reporting intense positive experiences to music were more likely to continue to engage with music over the lifespan. In a therapeutic study (Brodsky, 1995) it was discovered that professional musicians may even use their emotional involvement with music as some compensation for perceived deficiencies in human relationships. For instance, one violinist said 'when I am playing the violin I have a deeper relationship and experience than any human could offer'.

### 12.2.3 Psychological functionality of music

The above findings suggest that music serves a major psychological function in many people's lives, closely connected to the explicit purposes of many forms of psychotherapy. Musical engagement can thus become a form of self-administered therapy (Sloboda, 1989b). Conceptions of music as 'entertainment' or 'diversion' grossly underestimate the seriousness of many people's engagement with music.

## 12.3 Conceptual issues

### 12.3.1 Definitional problems

The psychological literature on emotion contains a number of unresolved problems. For instance, which psychological states are to be counted as

emotions? One proposal would include only positively or negatively valenced states ('anger', 'joy', etc.) and exclude unvalenced or neutral states (such as 'alertness' or 'surprise'). Another would make a distinction between emotions (acute reactions to time-limited events) and moods (chronic prevailing affective states engendered by the internal functioning of the organism). Yet again, researchers disagree about whether there are some emotions which are more basic than others (Frijda, 1986; Ortony and Turner, 1990). Research with music cannot resolve these problems. Indeed it may seem that music renders all these distinctions harder to make. Emotions to non-musical events can often be clarified by their behavioural consequences—e.g. if a person runs away then he or she is likely to be feeling fear. The behavioural consequences of emotional response to music are likely to be less tangible and less immediate. Emotions, moods, and other affective states may thus be less distinct from one another than in non-musical settings. In particular, it remains an open question as to whether emotional responses ascribed to specific time-limited musical events are independent of and clearly delineated from 'moods' evoked by whole pieces.

### 12.3.2 Between-person variation

A particularly acute problem presented by the study of emotion is its variability between individuals, and across time within individuals. There are some indications that cultural differences account for much of the between-person variance. Developmental studies (e.g. Gardner, 1973) confirm that knowledge of the conventions and structures of a musical culture are acquired over time, and influence the kind of emotional responses that are possible. Concepts of music as some form of universal 'language of the emotions' are probably without foundation (although one study—Pinchot Kastner and Crowder (1990)—has shown that US children as young as 3 years old can reliably map the major–minor distinction onto the happy–sad dimension).

A further source of between-person variability comes from the biographical associations from particular pieces of music. Musical stimuli, like olfactory ones, appear to have a particular power to evoke memories of earlier situations in which the same music was experienced, particularly if these situations had emotional significance. Although such effects have not received systematic study, their ubiquity and power is without question. In these circumstances, emotional response is dominated by the idiosyncratic emotional tone of the previous context, and, unless such contexts are widely shared in a culture (e.g. organ music connotes weddings and funerals) the responses they engender will be idiosyncratic.

### 12.3.3 Within-person variation

There is a considerable literature on liking and preference which suggests that familiarity with a piece of music affects these variables according to an inverted-U function (e.g. Hargreaves, 1984). When a complex piece is totally unfamiliar response will be low. As it becomes known, liking increases. When overfamiliar it becomes boring and liking decreases again.

On the other hand, many musicians report being emotionally moved by music which is very familiar to them. This could be a function of the very high complexity of the particular pieces for which this is true. It might take an inconceivably large number of exposures to move from the left to the right of the U-curve. Another intriguing suggestion (Jackendoff, 1992) is that some aspects of music processing are modular in the sense developed by Fodor (1983), and, as a consequence, denied full access to conscious memory. 'However well one knows a piece, then, expectation, suspense, satisfaction, and surprise can still occur within the processor. In essence, the processor is always hearing the piece for the first time—and that is why affect remains intact.'

What many discussions of emotional response to music seem to overlook is the dependency of this response on factors which have little to do with the music itself, and much to do with participants' own decisions and attitudes. Listeners can engage greater or lesser degrees of attention and concentration on a piece of music. They can choose different aspects of the music as a focus for their emotional attention. For instance, I can focus on a music event as structure, and experience an emotion commensurate with the interplay of tension and release, form and dynamic within the music. I can focus on the same event as the action of an agent or group of agents, and feel emotions such as admiration, in respect of the skill of the performer, for instance. Again, I can focus on the event as sound, and feel emotions relevant to the purity or sonority of a particular timbre. I can focus on the music as cultural symbol, and feel emotions relevant to group membership, such as patriotism. Because music is multidimensional and multifunctional, there is no one way to listen to a piece of music, and no one emotion appropriate to it. A recent study probing the variety of emotional responses to a single piece of music identified 13 discrete categories of antecedents of such responses, each of which could trigger the whole range of available emotions, but in response to different attributes of the situation (Waterman, 1996).

## 12.4 Methodological issues

### 12.4.1 Measurement context

Emotions by their nature are immediate and evanescent: they do not survive long after the triggering event. Measurement in relation to music would

therefore seem to demand taking measurements during music listening or performing itself. Some studies have done this (Madsen *et al.* 1993; Waterman, 1996) but, where measurement requires the conscious deliberation of participants in the choice of a response, such experiments are open to the criticism that the focus of emotional response will be on the task rather than the music, so that the task destroys the very thing it is trying to measure. This may possibly be overcome by very simple or well-practised response modes.

Other studies have attempted to preserve the integrity of the musical experience by collecting retrospective data. Such data can be immediate, as when participants are asked to report on their experiences directly a musical passage has finished (Hevner, 1936; Gregory and Varney, 1996), or they can be delayed, as when people are asked to report on typical responses to pieces they hear quite often (Sloboda, 1991). This kind of study is open to the criticism that memory loss and other kinds of bias can be introduced. To address such criticisms requires careful choice of the type of data collected.

### 12.4.2 Measurement content

One possible form of data is verbal commentary describing the actual emotion felt or perceived. To avoid the problems and ambiguities associated with free verbal description, the majority of researchers have asked participants to choose their responses from predetermined word-lists or bipolar dimensions (e.g. Hevner, 1936; Gregory and Varney, 1996). One problem with such a method is that it is ambiguous as to whether the respondent has actually felt the emotion concerned, or simply recognized the music as possessing that character.

Three other alternatives have been investigated. One is the use of a non-verbal response device, such as a pad to squeeze (Clynes, 1977; de Vries, 1991; Waterman, 1996), or a pointer to move (Madsen *et al.* 1993). In this way, intensity of some experienced dimension is indicated by the participant. It is still not entirely clear whether this response is representing feeling or judgement. A second alternative is the use of direct physiological measures (Collins, 1988). Unfortunately, very little of such work has been published, and it seems quite difficult to establish clear correlations between the nature and intensity of physiological response and the nature and intensity of an emotional experience. A third alternative is the use of self-report on physiological manifestations of strong emotion, such as crying or 'shivers down the spine' (Goldstein, 1980; Sloboda, 1991). These universal, stereotypical, and highly memorable occurrences seem the closest we can get to tapping the emotional experience itself.

## 12.5 How are musical events and emotions linked?

### 12.5.1 Gross characteristics

A significant number of studies have collected listener responses across a wide variety of contrasting musical stimuli, predominantly Western and classical (Rigg, 1964; Wedin, 1972; Gabrielsson and Juslin, 1996). The response have usually been verbal judgements of character using predetermined lists or bipolar scales. These studies yield some unsurprising correlations between certain gross characteristics and certain emotions (see Table 12.1). Performers have been shown to make use of similar characteristics under their control to change the emotional character of their performance (Gabrielsson and Juslin, 1996).

### 12.5.2 Internal structural characteristics

Within a prevailing mood or emotional character, determined by quite general characteristics of a musical passage, individual events seem to carry or heighten the specific emotional character or intensity of a piece of music. The music theorist Deryck Cooke (1959) argued that certain tonal intervals, and melodic sequences constructed from these intervals, carried very specific emotional content. He based his proposal on analysis of a very large body of Western vocal music, showing that particular tonal sequences repeatedly accompany words depicting particular types of emotion. Unfortunately, only one small-scale psychological study (Gabriel, 1978) has ever attempted to verify the theory with experimental data. Although the study failed to find the predicted effects, this could have been because of methodological weaknesses in the study (Sloboda, 1985).

A different theoretical approach to the issue of emotion was introduced by the music theorist Leonard B. Meyer (1956). This was based on the psychological insight that emotions in general are normally experienced due to violations of expectancy (surprises) of some kind. In music, such expectations are built up through the patterning of melodic, rhythmic, harmonic and other structural elements. Sometimes violations of expectancy are quite crude (e.g. sudden changes of key or dynamic). On other occasions they are

**Table 12.1** Examples of emotion–music links (after Gabrielsson and Juslin, 1996)

| Emotion | Musical characteristics |
|---|---|
| Serious, solemn | Slow, low-pitched, regular rhythms, low dissonance |
| Sad | Slow, low-pitched, minor mode, high dissonance |
| Happy | Fast, high-pitched, major mode, low dissonance |
| Exciting | Fast, loud, high dissonance |

**Table 12.2** Examples of thrill-structure links (after Sloboda, 1991)

| Thrill type | Structural characteristics |
|---|---|
| Tears | Appogiatura or suspension melodic or harmonic sequences harmonic movement descending through circle of 5ths |
| Shivers | Enharmonic change or sudden key-shift sudden dynamic or textural change |
| Heart | Repeated syncopation prominent event occurring earlier than expected |

more subtle, and not necessarily consciously noticed by the listener. Meyer was somewhat pessimistic about the possibility of empirically verifying his proposals, for many of the reasons outlined in Sections 12.3 and 12.4 above.

I have argued (Sloboda, 1991) that the retrospective elicitation of information about the nature and musical location of 'thrills' provides a good chance of generating data relevant to isolating structural determinants of emotional response. In a study where music-lovers were asked to identify known pieces of music that reliably provoked 'thrills', a significant minority were able to identify the thrill-producing segment to within a single phrase. Three main types of thrill were reported: tears (crying, lump in throat); shivers (goose flesh, skin sensations); and heart (racing heart, sinking stomach and other visceral sensations). Table 12.2 shows the main musical features associated with each of these types of thrill.

A recent study has shown that thrills may be elicited during *in vivo* experiments with large numbers of listeners in the same room, and that these responses tend to occur at the same points in the music, although no formal analysis of musical structure at these points was reported (Panksepp, 1995).

The discovered structure–thrill link is somewhat enigmatic. Although each structural device relies for its effect on creation and violation of expectancy, our understanding of these responses in the context of non-musical stimuli is not advanced enough for intelligent analogising in the realm of music. I finish by raising a couple of questions which must be answered if we are to know that we are making real scientific progress in this area:

1. What characteristics of the musical situation make it possible for participants to experience most emotions (including crying) as positive and desirable?
2. What characteristics of musical performances (acoustic, expressive, timbral, social) make different renditions of the same musical work so different in emotional power?

3. What mechanisms allow someone to be emotionally moved by imagining themselves to be hearing a performance of a known piece of music, in the absence of any actual acoustic stimulation?

## References

Brodsky, W. (1995) Career Stress and Performance Anxiety in Professional Orchestra Musicians: a Study of Individual Differences and their Impact on Therapeutic Outcomes. PhD Thesis, University of Keele, Staffordshire, UK.

Clynes, M. (1977) *Sentics: the Touch of Emotions.* New York, Doubleday.

Collins, S. C. (1988) Subjective and autonomic responses to western classical music. Unpublished PhD Thesis, University of Manchester, UK.

Cooke, D. (1959) *The Language of Music.* Oxford, Oxford University Press.

de Vries, B. (1991) Assessment of the affective response to music with Clyne's Sentograph. *Psychology of Music,* **19**(1), 46–64.

Fodor, J. A. (1983) *Modularity and the Mind.* Cambridge, MA, MIT Press.

Frey, W. H. (1985) *Crying: The Mystery of Tears.* Minneapolis, Minnesota, Winston Press.

Frijda, N. (1986) *The Emotions.* Cambridge, UK, Cambridge University Press.

Gabriel, C. (1978) An experimental study of Deryck Cooke's theory of music and meaning. *Psychology of Music,* **6**, 13–20.

Gabrielsson, A. and Lindstrom, S. (1993) On strong experiences of music. *Musikpsychologie, Jarbuch der Deutschen Gesellschaft fur Musikpsychologie,* **10**, 114–125.

Gabrielsson, A. and Juslin, P. (1996) Emotional expression in music performance. *Psychology of Music,* **24**(1), 68–91.

Gardner, H. (1973) Children's sensitivity to musical styles. *Merrill-Palmer Quarterly of Behavioral Development,* **19**, 67–77.

Goldstein, A. (1980) Thrills in response to music and other stimuli. *Physiological Psychology,* **8**, 126–129.

Gregory, A. H. and Varney, N. (1996) Cross-cultural comparisons in the affective response to music. *Psychology of Music,* **24**(1), 47–52.

Hargreaves, D. J. (1984) The effects of repetition on liking for music. *Journal of Research in Music Education,* **32**, 35–47.

Hevner, K. (1936) Experimental studies of the elements of expression in music. *American Journal of Psychology,* **48**, 248–268.

Jackendoff, R. (1992) Musical processing and musical affect. In: M. R. Jones and S. Holleran (Eds.) *Cognitive Bases of Musical Communication.* Washington DC, American Psychological Association, pp. 51–68.

Madsen, C. K., Byrnes, S. R., Capperella-Sheldon, D. A., and Brittin, R. V. (1993) Aesthetic responses to music: musicians versus non-musicians. *Journal of Music Therapy*, **30**, 174–191.

Meyer, L. B. (1956) *Emotion and Meaning in Music.* Chicago, University of Chicago Press.

Ortony, A. and Turner, T. J. (1990) What's basic about basic emotions? *Psychological Review*, **97**, 315–331.

Panksepp, J. (1995) The emotional sources of 'chills' induced by music. *Music Perception*, **13**, 171–208.

Pinchot Kastner, M. and Crowder, R. G. (1990) Perception of the major/minor distinction: IV. emotional connotations in young children. *Music Perception*, **8**(2), 189–202.

Rigg, M. G. (1964) The mood effects of music: a comparison of data from earlier investigations. *Journal of Psychology*, **58**, 427–438.

Sloboda, J. A. (1985) *The Musical Mind: the Cognitive Psychology of Music.* Oxford, Oxford University Press.

Sloboda, J. A. (1989a) Music as a language. In F. Wilson and F. Roehmann (Eds.) *Music and Child Development*, St Louis, Missouri, MMB Inc., pp. 8–43.

Sloboda, J. A. (1989b) Music psychology and the composer. In: S. Nielzen and O. Olsson (Eds.) *Structure and Perception of Electroacoustic Sound and Music.* Amsterdam. Elsevier, pp. 3–12.

Sloboda, J. A. (1991) Music structure and emotional response: some empirical findings. *Psychology of Music*, **19**, 110–120.

Sloboda, J. A. (1992) Empirical studies of emotional response to music. In: M. R. Jones and S. Holleran (Eds.) *Cognitive Bases of Musical Communication.* Washington DC, American Psychological Association, pp. 3–50.

Waterman, M. (1996) Emotional responses to music: implicit and explicit effects in listeners and performers. *Psychology of Music*, **24**(1), 53–67.

Wedin, L. (1972) A multi-dimensional study of perceptual-emotional qualities in music. *Scandinavian Journal of Psychology*, **13**, 1–17.

CHAPTER 13

# MUSICAL PERFORMANCE AND EMOTION: ISSUES AND DEVELOPMENTS

## 13.1 Emotional response to music is dynamic

There now exists substantial scientific confirmation for two commonplace observations about the experience of listening to music. The first of these observations is that the intensity of emotional response to a piece of music often rises and falls as the music unfolds. Musical discourse, both formal and informal, talks of climaxes and points of repose, tension and relaxation. In other words there are peaks, where intense emotions (or other affective sensations) are prone to be experienced, and troughs, where the intensity is weak. These peaks and troughs occur at similar moments in the music across different people within a culture, and are repeated in individual hearings separated by as much as one year (Waterman, 1996). Peaks appear to be associated with particular structural features of the music (Sloboda, 1991), and often are accompanied by physical manifestations of emotion such as weeping, pilo-erection, heart-rate changes, etc.

## 13.2 Performers contribute to emotional power

The second observation is that the nature of the specific performance contributes significantly to the quality of the affective experience. Again, musical discourse abounds with comparisons between different interpretations of the same composition. Music critics earn their living by describing the differences in their reactions to different performances, and attempting to convince readers that their characterizations of one performance as 'thrilling' and another as 'dull' are apt and perceptive. Two converging lines of evidence are pertinent to this phenomenon. On the one hand, listeners within a culture show agreement when asked to characterize complete professional performances along bipolar dimensions such as 'powerful–weak', 'masculine–feminine', 'expressive–inexpressive', and the values of these dimensions correlate to objective features of the performance such as the timing profile of individual notes or beats (e.g. Repp, 1990). On the other hand, when performers are asked to play the same music

with different expressive character (happy, sad, angry, expressive, deadpan), their performances show objective differences in expressive microstructure, and listeners are able to assign the intended character to each performance with greater than chance accuracy (Gabrielsson and Juslin, 1996; Kendall and Carterette, 1990). These findings show that, at least in some respects, judgments about performance qualities have some inter-rater reliability, and are not entirely a matter of idiosyncratic taste or whim.

So, there is absolutely no doubt that listeners experience greater or lesser levels of emotional intensity as a result of what composers and performers do. The focus of this paper is on *how* these effects are brought about.

## 13.3 Intensity matters more than content

I need to make a number of clarifications or qualifying remarks. Firstly, in this paper I am not very much interested in *what* emotion is present, detected, or felt. My interest is in the mechanisms whereby the intensity of any emotion may be brought to a peak. Indeed, I do not even want to make strong claims for music's capacity to embody, express, communicate or induce specific emotions (such as happiness, anger, disappointment, etc.). Roughly speaking I take the position that there is a range of characteristics which an emotion may contain. The more of these characteristics are present in a psychological state the more 'emotion-like' it becomes, but there is no clear cut-off between emotions, feelings, states of arousal, etc. Our interactions with music certainly take us into the territory or domain where to talk of emotions being present is legitimate. In everyday contexts I think emotions can be identified and discussed with precision because they generally involve an event, often enacted by human agents, accompanied by biologically pre-programmed gestures and facial expressions, which has social and personal consequences.

Our identification of, for instance, anger, is aided by the fact that A has acted towards B in a way that has violated B's dignity, rights, or prospects; B is showing signs of emotional arousal, is enacting certain bodily movements and facial gestures, and is verbally expressing that she is aware that A's act has damaged her in some ways. If any one of the cues is missing, it is still possible to make an emotional attribution. The more cues are missing, the more tenuous the identification becomes. Music (at least the 'pure' instrumental music which has dominated high art for the last 200 years) is of its nature impoverished in cues. Music does not directly describe or depict agents with purposes who enact behaviours with social consequences. In this respect it is unlike other emotion-bearing art forms, such as drama, literature, even dance. It does, however, suggest or evoke

varying degrees of energy, tension, and arousal. It can suggest or resemble certain types of human gestures and actions. Given shared experience and understandings, it is easy to see how a group of people might be able to extrapolate specific emotions from such cues. It is equally easy to see how there is room for ambiguity and imprecision. Very often we feel that there is an emotion present, we know it is of one general type rather than of another, but we cannot quite tie it down. In such a state of ambiguity and cue-impoverishment we may well expect the profound and semi-mystical experiences that music seems to engender. Our own subconscious desires, memories, and preoccupations rise to flesh out the emotional contours that the music suggests. The so-called 'power' of music may very well be in its emotional cue-impoverishment. It is a kind of emotional Rorschach blob.

## 13.4 Music blurs the boundary between feeling and judgment

My second qualification is that I am not very concerned with the distinction between recognition of emotion (as somehow inhering in the music) and the experience of emotion within the listener. These are, of course, rather distinct psychological states. Knowing that someone else is angry is not the same as being angry oneself. But again, I think the distinction has much more force in the everyday world of human interaction than it does in the case of music. If you are angry, and I detect your anger, then my own emotional reaction will be a function of my own role in the social drama that is being enacted. If you are angry with *me*, I may well feel scared, or defensive. If you are angry with someone else, then my own feelings may be determined by my relationship to both you and the person you are angry with, as well as my value system. I may feel anger too in the case where I see the injustice as serious, and value your welfare. I may, on the other hand, feel pity or even amusement, if I feel that the event about which you are angry is trivial. The precise logic of the interacting set of beliefs, actions, intentions, and interpretations will almost totally determine my own emotions. In music, most elements of this logic are uncoupled. If I detect some piece of music as containing elements appropriate to anger, then there is a sense in which I am pretty free to ascribe that anger where I will. I may ascribe it to the composer, or to the performer, to myself, to God. None of these ascriptions need bring me up against any of the immediate constraints or contradictions provided by a real-life event in which my own interests and relationships figure. A possible hypothesis is that people are likely to ascribe the emotion they detect within the music to *themselves* at those points where their general level of

music-entrained arousal is highest, and other sources of specific emotion are not obtruding or interfering at the same moment.

## 13.5 Intrinsic versus extrinsic sources of emotion

Why do we need to consider other sources of emotion? Consider the case of a performance of a piece I know well which contains obvious technical shortcomings. I might well become irritated with the performance. But it is not the music as music which is causing the irritation but the actions of the performer as social agent *reflected in* the performance, not arising from the specifically musical intentions of either composer or performer. For instance, at a professional concert I may resent having my time and money wasted by someone who appears not to respect his audience enough to prepare well. Given the existence of this emotion, I may not have the capacity or inclination to ascribe any of my emotion to the musical content, even if my level of emotional arousal is rising or falling as a consequence of the specifically musical contour. I may just feel myself as more or less irritated!

There is some suggestive evidence that negative non-musical emotions (including anger, fear, humiliation, anxiety) are particularly inconsistent with experiencing music-entrained emotion in oneself (Sloboda, 1990). In general, positive emotions (such as love, being proud of, admiration, etc.) seem to make listeners more able to 'own' or 'identify with' the emotional peaks and troughs in the music. It is as if these states both enhance general emotional arousal, and at the same time lower the threshold for self-ascription of the emotions discovered in the music.

## 13.6 Emotional peaks ascribable to compositional devices

The contributions made by the research community to addressing this issue have to date been inductive in nature. That is so say, we have measured the contour of emotional responsiveness to pre-composed pieces, chosen either at random or by the participants, and have then tried to discover whether the peaks have anything in common.

A variety of methodologies have been tried. One methodology focuses on the time-limited physical manifestations of emotion such as the feeling of tears welling to one's eyes, pilo-erection, and other phenomena which have variously been called thrills, even skin-orgasms. These have been monitored either retrospectively ('tell me which pieces give you reliable thrills', Sloboda, 1991) or concurrently (Panksepp, 1995). Other researchers have

used more internal criteria ('press the button when you feel something' (Waterman, 1996). Yet others have used some kind of continuous measure, such as finger-pressure on a pressure-sensitive pad (DeVries, 1991; Madsen *et al.*, 1993, Krumhansl, 1996).

We probably have too little data yet to make definitive conclusions, but I don't know of clear proof that these methods produce hugely differing results. What the data confirm in a general way is that the emotional system is sensitive to change. We don't feel much emotion to steady state. One possible exception to this is music of such generally high levels of sound energy that it is neurologically arousing in its own right (i.e. very loud or very fast music). There is clear evidence that pieces of music which vary in overall energy tend to have emotional peaks at points of high energy, wherever else they may have peaks.

Another possible exception is provided by musical events whose sound characteristics resemble biologically programmed emotional triggers (such as the sound of a female voice).

## 13.7 Are there mechanical rules for composing emotional peaks?

In 1991, I published a list of 10 music-structural characteristics that appeared repeatedly in passages which Western listeners cited as reliably provoking 'thrills'. Despite a considerable increase in the amount and quality of work on musical emotion in the intervening years, no one has proposed specific modifications or refinements to this list.

I have conjectured that what all these characteristics have in common is that they cause perturbations of one sort or another in expectations, brought about either in the light of the structure of that specific piece, and/or (more commonly) in the light of general features of the body of tonal music as a whole. For instance, an appogiatura creates a local dissonance that violates a basic expectation that melody notes will be taken from the actual or implied harmony at that point. This dissonance is then resolved in the direction of the expected note. A syncopation provides an accent (point of high energy) earlier than the established and conventional metrical structure would predict. And so on.

The problem with both my list and the conjecture which links its members is that this hugely overdetermines the location of emotional peaks. First, it is trivially easy to find instances of these devices scattered all through each of the pieces that provided the specific examples for my list. If the mere presence of such a device were enough for an emotional peak, then there would hardly be a measure of music which did not

contain one. The conjecture is also too inclusive. Almost any change in any parameter could be construed as violating some expectancy or other. On this conjecture, 'Twinkle twinkle, little star' (*Ah vous dirai-je, Maman*) should elicit an emotional peak as it moves from the first two repeated notes to the second two. The expectation is that two repeated notes will be followed by a third.

There are two rather obvious moves that one could take to bring theory into a better relation with reality. First, one could postulate that the degree of unexpectedness matters: the more unexpected the more emotional. Second, one could postulate that the density of devices matters: so that the greater number of these devices present within a measure the greater likelihood of that measure being experienced as a peak. These are plausible ideas, but they would need a lot of work before they could be operationalized in empirical studies to test them. It is hard to see how one might provide appropriate data without an automated procedure for detecting the relevant devices in a large corpus of music. Writing the relevant software would be in itself a major exercise in artificial intelligence since it is hard to see how many of these devices could be captured with simple algorithms.

My intuition is that doing the above work might capture some of the variance in emotionality within a piece but by no means all. At this point I will simply make two further conjectures which seem to resonate with musicians that I respect.

## 13.8 Structural junctions and asynchrony of levels

Almost all descriptions of music (whether professional or informal) characterize it as made up of sections of differing lengths, differing levels of hierarchical embeddedness, and differing levels of importance within the overall plan of the piece. In sonata form, for instance, the start of the recapitulation of the first theme is a very significant and pivotal moment. We could conceive the listener as attempting to obtain some form of structural map of the piece from the cues available. These cues would by and large be boundary cues, signifying the end (or the impending end) of one section and the start of the next. Compositional devices which perturbed expectations at such boundary or junction points should be far more emotionally arousing than similar devices at other points, simply because they are much more crucial to helping or hindering the process of establishing the structural map.

Secondly, emotionally and aesthetically satisfying works of music have a sense of unity and rightness about them. Not any violation of expectancy

will cause musically relevant emotion. One could simply stop a piece before the final cadence. That would be surprising, but not musical. Because music is multidimensional and multi-levelled, a device which breaks expectancies on one dimension or one level can confirm expectancies on another. Our aesthetic reaction to such a device is 'how unexpected, but how beautiful, and how right'. A simple example of what I mean is given by the opening melodic phrase of the Albinoni *Adagio for strings* (D C B<flat> A G G F#). On one level of description there is a series of appogiaturas creating and then resolving a dissonance through each successive note. On another level of description the whole sequence conforms to a smooth descending scalar motion, each next note expected in the light of what went before as a single scale step lower than its predecessor. This is a very simple example, but much of the art of professional music analysis is to uncover more complex multiple relationships within works of interest. And this is a task which must be left to analysts. Psychologists have no credible alternative route at this stage of development of the field.

## 13.9 The problem of experimental control

There is a fundamental scientific problem that besets this research area. Any real piece of music has multiple characteristics at any point. Where we gain our data of emotional response from such pieces, we can never be sure that we are able to fully separate the potentially contributory dimensions of the sound, even if we were able to list them all. Suitable controls may not exist within the body of composed music (for instance where two dimensions of sound are significantly correlated with one another). Yet we cannot hope to easily elicit valid judgments of emotionality from listeners with experimenter-designed atomistic fragments designed to fit some factorial experimental design. Early naive attempts to do this (e.g. Gabriel, 1978) were a complete failure for reasons which are now pretty obvious. Emotions in music are sustained by the rich set of cultural, historical, structural, and situational cues which a properly embedded musical event can elicit. These cues provide a set of enabling conditions for the ups and downs of moment-by-moment emotion. Take these broader cues away and the whole emotional enterprise collapses.

## 13.10 The performer's contribution: faithfulness to the composers' intentions?

What it is that a performer or interpreter of a notated composition adds to the notation in order to create a compelling experience for a listener?

One suggestion concerning the function of interpretative variation is that it is intended to reinforce and make more apparent to a listener a structural feature which the performer has reason to believe is intrinsically important. Very often, at least in the classical 'notated' canon, the score (or rather, an culturally accepted way of reading the score) determines some of the important structural features, through metrical and phrase marks, notated accents and dynamic marks. So, for instance, pianists have been shown to perform melodies admitting of two valid metrical interpretations in such a way as to communicate the actual notated metre to listeners. This explanation of performance is consistent with the performance tradition which talks of a performance's purpose as to be 'faithful to the composer's intentions'. In this tradition the key interpretative process is the detailed study of the score, accompanied (if the composer is no longer alive) by consultation of historical commentaries, and other markers of 'authenticity'. This perspective allows a position from which judgments of value may be made. One performance is more compelling than another because it emphasizes, rather than obscures, a key structural phenomenon that prior score analysis reveals.

## 13.11 Faithfulness to convention

There are, however, two distinct types of structural phenomenon within most music, which may engage a listener. The first might be called the *conventional*. This includes that set of phenomena which characterize the music of a whole period or genre. For instance, it is normal for Western melodies to be constructed out of bars or measures of equal duration, linked together into superordinate units, phrases, which normally contain $2n$ measures, where $n$ is an integer not greater than 4. Compositional practices, such as rhythmic repetition, longer than average note durations, etc. are deployed to mark significant structural units and the boundaries between them. There also exist a whole set of conventional performance practices which emphasize phrase structure (e.g. slowing at phrase ends and speeding in phrase middles). Listeners within a culture are so accustomed to these conventions that they find it hard to explicitly notice small performance perturbations in line with the conventions, as Repp (1998) has shown.

## 13.12 Faithfulness to the dialectical structure

The second kind of structural phenomenon might be called *dialectical*, in that listeners' conventional expectations are challenged or thwarted in

some way. These expectations have sometimes been called *schematic* because they derive from schemas that have been built up through extensive exposure to a wide range of music, rather than the *veridical* representation of any single musical object. A basic schema derives from the diatonic scale, and the harmonic relations between its members. Schematic perception 'expects' two consecutive members of a scale to be followed by the next consecutive member in ascending or descending order. A skip or change of melodic direction is thus something that challenges structural expectations at a very basic level. Such dialectical features of compositions are exactly those that we have already found evidence stimulate more intense emotional experiences.

This distinction between conventional and dialectical structure means that performers can 'emphasize' structure in two quite different ways. By emphasizing the conventional structure, the performer can 'feed' schematic expectations. By emphasizing the dialectical structure, the performer can heighten the experience of the unexpected. A simple example of this might be to slightly delay the onset of an enharmonic change. The surprise of the unexpected harmony is heightened by its occurring at a slightly unexpected time. Being 'faithful to the score' does not, therefore, mean giving the listener no surprises. The surprises are, however, precisely those which the performer believes the composer intended the listener to experience.

## 13.13 Performance as expression of individuality

Alongside, and in some ways in contradistinction to the notion of performer as faithful carrier of the composer's message, is the notion of performer as creator in his or her own right. This tradition of performance lays on the performer the duty to come up with something new, an interpretation which allows us to see the composition in a different light, to make manifest aspects of the music which are normally latent. The commercial realities of performance often press performers in this direction. Why should someone pay money to listen to a performance which is not distinctively different from the one they heard the other day (whether in live performance, or on a recording)? Repp (1997) has provided conclusive documentation of the reality of this pressure towards individuality. In a series of analyses of solo piano performances, some from commercially available professional recordings and some from competent amateurs, he was able to show that amateurs conform more closely to an average interpretation than do professionals. Some professionals are, however, very close to the average, so we can rule out the notion that the average performance is in some

sense the result of lack of skill or appropriate musicality. We could see such professionals as operating on 'faithfulness to the score' as their interpretative philosophy. On the other hand, the most extreme performances, as measured on a variety of dimensions, including speed, exaggeration, etc. are almost always those of 'world class' professionals. Typically, many of these performers do things which are not obviously prescribed by the score. Yet, in order to be received as a legitimate performance of the work in question, their interpretative decision must in some way be able to be seen as 'fitting' the music. Not every exaggeration will be acceptable.

## 13.14 Virtuosity as a boundary case

A rather interesting example of exaggeration which might be emotionally compelling in and of itself, is one where the virtuosity of the performance is, in itself, a source of emotion. I can recall numerous occasions of high emotion brought about by the way that a performer, or group of performers, executes a particularly difficult passage with exceptional accuracy or panache. For instance, we could consider a performance of Chopin's Prelude Op 26 no 16 in B flat minor that was given by Krystian Zimmerman at the finals of the 9th International Chopin Competition in 1975 (from *Frederyk Chopin: International Chopin Piano Competitions Vol. 4*. Polskie Nagrania, 1991, CK 1015). At the time he was 19 years of age. This, and other performances, won him the prize.

The virtuosity in this performance is the predominant aspect of it—it is the thing that anyone who knows anything about piano playing cannot ignore! As I listen to this I can detect a variety of conflicting strands to my emotion. One of these is a kind of stylistic distaste—a sense that Chopin ought not in general to be played this way! Another, much stronger emotion is a complicated and multi-levelled identification with the Polish character (I am half-Polish myself). At a basic level there is just shared pride in the identity, particularly its musical history and passions (this is, after all, a Polish performer, winning a Polish competition with the work of Poland's greatest composer). One level further in there is a sense that this performance reflects something of the Polish character itself—strong, impetuous, passionate, even intemperate. Within its historical context the performance could be further interpreted as an expression of the outraged defiance of an oppressed people at the height of the dead hand of Moscow's cold-war rule of that country, a defiance which was shortly to take shape in the Solidarity movement. Everything I have said so far informs and infects my emotional state for the whole 54 seconds of the performance, from its shattering first few chords to the abrupt final cadence. Then there

are further feelings which come and go at particular points in the music. For me, the sections in parallel octaves have incredible emotional power. This is because I know, as a pianist, that the coordination required to bring these off effectively is of an order of magnitude greater than that required to accompany a right-hand line with left-hand chords. This is especially so in the lower registers of the piano where the acoustic interference between notes is so much greater, and therefore requires added precision. When you think he cannot possibly play any faster, there is an accelerando towards the end. This together with a section where the performer thumps the left-hand chords to the point where they become ugly and overpowering, makes me feel that the performer is almost trying to pound the piano and the piece into total destruction. I sometimes feel pummelled and pounded myself as if I was witnessing or being part of an assault taking place. According to one's reading of the social situation, the assault can take on a symbolic significance. It could represent a perception of how the communist regime was undermining Poland's national heritage; demonstrating through the semi-destruction of Chopin what was being done to Poland as a whole. It could simply be an arrogant young pianist assaulting the prejudices of a rigid music establishment.

## 13.15 Art versus spectacle

The question for this presentation is whether the fact that more or less virtuosity can be detected at various points, apparently leading to more or less emotion, is in itself a case for saying that such emotions are ascribable to the music. My sketch of the Zimmerman performance is designed to suggest that there is no simple or universal answer to that question. In certain cultural and artistic contexts, virtuosity becomes a salient aesthetic variable which is at the intentional service of the artist. But it will be the specifics of the piece, the historical and cultural context of both performer and listener, and the precise variations in virtuosity, which will determine whether or not we put virtuosity on one side or other of the divide we are discussing. In this case I am inclined to ascribe the emotion brought about by virtuosic accelerando as a music-relevant emotion—being a part of the headlong rush towards the untimely and brutal demise of the piece. On the other hand, the particular thrill I receive from the bass parallel octaves is a consequence of my pianistic experience with modern grand pianos (which might be unlike the piano used by either Zimmerman or Chopin) and sits pretty much outside any musical considerations. In the former case virtuosity is at the service of structure, and contributes to artistry. In the latter case virtuosity floats free of structure, and contributes only to

spectacle (which is a legitimate source of emotion but has nothing to do with music as such, even though some music may, as it happens, be spectacular).

I have indulged in this rather personal analysis primarily to show just how complex and multi-layered are the elements of emotion to any piece of music. Even something as apparently straightforward as virtuosity is actually a deeply socially embedded judgment, sensitive to one's beliefs about the context, the level of skill displayed, and its aesthetic appropriateness. Of course psychology must take isolated fragments of this picture in order to make rigorous scientific progress, but we have constantly to relate our discrete findings back to the full richness of emotional and aesthetic experience if we are to retain ecological generality.

## 13.16 Mapping relationships between performance characteristics and listener response

We (Sloboda *et al.* 1997) have described a preliminary attempt to characterize, for a single short piece of romantic piano music, the nature of the relationship between performer decisions and listener's judgments of emotionality.

In order to systematically explore relationships between moment-to-moment variations in a musical sequence and corresponding variations in listener response, a means is required of assessing listener response concurrently with the musical stimulation. Our understanding of the relationship between physiological measures and emotional impact is not yet very detailed. However, several investigators have successfully used a method whereby listeners operate a continuously variable manual control during music listening in order to indicate the degree of some quality of the music as it unfolds over time (e.g. Krumhansl, 1996). The functions so obtained appear to be reliable, both within and between individuals.

The main purpose of our study was to provide continuous ratings of emotionality for a range of performances of the same piece of music, in order to determine (1) that there is enough agreement between different listeners for reliable differences between the emotionality functions of different performances to be established, and (2) whether there are reliable correlates of emotionality within performances: whether arising from conventional or dialectic structures, or performer's creative departures from such structures. Results showed that listeners showed significant agreement about which performances (and which portions of them) were most emotionally powerful.

Our means of isolating emotionally interesting segments of individual performances was as follows:

- 27 listeners heard each performance, and a mean emotionality trace for that performance was generated on the basis of the average of all 27 ratings at each point (*individual trace*).
- A grand mean trace of all 20 performances was derived by averaging the mean traces for the individual performances (*average trace*).
- Three independent raters identified locations where the individual traces substantially diverged in slope from the average trace over at least an entire measure of the music (i.e. one trace went up where the other trace stayed level or went down). 24 such locations were discovered.
- We took the 10 largest divergences, and undertook analysis of the MIDI files generated by performances at these points. Rather than create an artificial average performance, we used the individual performance that yielded an emotionality trace that was closest to the average, and compared different individual performances to that.
- We examined the MIDI files for up to two measures prior to the start of the emotionality divergence, on the grounds that emotional responses are often integrative and require time to 'come on line'.
- We looked at loudness, inter-onset-interval, and left–right onset asynchronies, and compared their values in the key segments both to what that performer did immediately before, and also to what the emotionally average performer did at that point.

In 9 out of 10 cases we found clear evidence of a change or contrast in at least one of these parameters that could unambiguously be linked to the change in emotionality.

In cases where a rise in emotionality was involved, the performance changes tended to be within one measure either side of a phrase boundary, offering some support for the notion that points of structural importance are key locations for emotion-enhancing performance variations. Within these, some clear examples of dialectic devices were evident—for instance, performers using conventional performance markers for phrase boundaries ahead of the phrase boundary.

Detailed analysis of the MIDI outputs of the performances demonstrated that high-emotion performances mostly demonstrated 'dialectical' rather than 'conventional' expressive gestures, going against genre-specific expectations for locations of timing or dynamic extremes. The question of whether the particular dialectical devices were faithful to the composer's intentions is a difficult one. It may be an unanswerable one for conceptual

as much as historical reasons (i.e. the composer may not have a specific performance in mind, and may recognize in a performance a legitimate aspect of his work which he had not explicitly designed to be emphasized). In this particular piece, all we can say is that several of the emotion-bearing performance devices were not related to obvious dialectical aspects of the composed structure, but creative additions to this dialectic by the performer.

## 13.17 Do performers intend to communicate emotion?

There is much talk of music as being a medium for emotional communication. Although we can undoubtedly feel emotion in music, can we really say that there is emotional communication going on? For communication to happen, we would seem to require that the communicator intended something, and there is some evidence that the intended message was received by the audience. For the communicator to know that communication has been successful, he or she requires some feedback, direct or indirect, from the audience.

I have nothing at this point to say on how composers' emotional intentions might be realized or studied, but it is rather easier to make some preliminary remarks about performers. After all, they are here to speak. Composers are often unavailable.

As a follow-up to Sloboda *et al.*'s (1997) study, we interviewed each pianist about their communicative intentions in the specific performances we recorded. We found that pianists agreed in broad terms about what emotion the specific piece designated. This agreement could be a complex function of (1) their own unmediated response to hearing a person perform it, who could be themselves, (2) their knowledge of the discourse surrounding this piece in the musical community (texts, opinions of teachers, etc.). Their comments, however, focused around the notion of the music 'speaking directly' to them, either through their own interaction with the score in rehearsal, or, in some cases, through influential public performances by other pianists, which showed them something important about the piece.

Although pianists agreed about the overall character of the work, they disagreed, sometimes quite extremely, concerning the emotional microstructure of the piece. Does the second phrase have a different emotional character to the first, and if so what is it? Here pianists had quite different things to say. They were quite ambiguous about the extent to which they sought or valued audience response. Many of them made quite strong statements rejecting the idea that what an audience feels is of any concern to them at all. A common tactic for performers is to use their own

emotional reactions to their performance as the yardstick of effective performance. Since this creates a closed loop which excludes the audience altogether, one might question whether the concept of emotional *communication* is valid at all in many music performance contexts.

What allows us as scientists to break out of that loop is the measuring of listener emotional response against explicit performance intentions (even though performers are habitually unable or unwilling to do that). In our study, we asked each performer to give an account of the major interpretative decisions made, immediately after their performance. Even though there was no pressure on them to be systematic or exhaustive, such that the average number of expressive intentions mentioned was only 5, six of the 10 locations identified above as leading to major shifts in emotionality were explicitly referred to by the performers.

## 13.18 Summary

It is a commonly held folk belief that the emotional power and freshness of music performance is in its intuitive spontaneity, which cannot be captured in words or specific moments consciously planned by performers and consciously identified by listeners. The work which I have outlined suggests the trail from performance intention to listener response may be more transparent than the folk belief would hold. Musicians study musical texts, experiment with different interpretative devices, and make conscious decisions to incorporate specific devices (based on performance parameter contrasts), which can be described and accounted for at some level of verbal specificity. These devices drive listeners' responses, which, in the cases we examined, lead to divergences in rated emotionality at, or immediately after the location of the devices. None of this devalues the performer's art, because it is the precise nature and location of these contrasts, chosen with an expert knowledge of piece and genre, neither overstated nor so tiny as to be imperceptible, which determines whether the listener's emotions remain tied to and focused on the unfolding musical structure, or whether they jump out of the frame because the performance is too boring, too exaggerated, or in some other way inappropriate to the work and the context. We may know something about what makes a performance emotional. We know almost nothing about what makes it count as a contribution towards art!

## References

DeVries, B. (1991) Assessment of the affective response to music with Clynes' Sentograph. *Psychology of Music*, **19.1**, 46–64.

Gabriel, C. (1978) An experimental study of Deryck Cooke's theory of music and meaning. *Psychology of Music*, **6**, 13–20.

Gabrielsson, A. and Juslin, P. (1996) Emotional expression in music performance. *Psychology of Music*, **24.1**, 68–91.

Kendall, R. A. and Carterette, E. C. (1990) The communication of musical expression. *Music Perception*, **8**, 129–164.

Krumhansl, C. L. (1996) A perceptual analysis of Mozart's Piano Sonata K. 282: segmentation, tension, and musical ideas. *Music Perception*, **13.3**, 401–432.

Madsen, C. K., Byrnes, S. R., Capperella-Sheldon, D. A., and Brittin, R. V. (1993) Aesthetic responses to music: musicians versus non-musicians. *Journal of Music Therapy*, **30**, 174–191.

Panksepp, J. (1995) The emotional sources of 'chills' induced by music. *Music Perception*, **13.2**, 171–208.

Repp, B. (1990) Patterns of expressive timing in performances of a Beethoven Minuet by nineteen famous pianists. *Journal of the Acoustical Society of America*, **88**, 622–641.

Repp, B. (1997) The aesthetic quality of a quantitatively average music performance: two preliminary experiments. *Music Perception*, **14.4**, 419–444.

Repp, B. (1998) The detectability of local deviations from a typical expressive timing patters. *Music Perception*, **15.3**, 265–290.

Sloboda, J. A. (1990) Music as a language. In F. Wilson and F. Roehmann (Eds.), *Music and child development*. St. Louis, Miss.: MMB Music Inc.

Sloboda, J. A. (1991) Music structure and emotional response: some empirical findings. *Psychology of Music*, **19**, 110–120.

Sloboda, J. A., Lehmann, A. C., and Parncutt, R. (1997) Perceiving intended emotion in concert-standard performances of Chopin's Prelude No. 4 in e-minor. In A. Gabrielsson (Ed.) *Proceedings of the 3rd ESCOM Conference*. Uppsala: Department of Psychology.

Waterman, M. (1996) Emotional responses to music: implicit and explicit effects in listeners and performers. *Psychology of Music*, **24.1**, 53–67.

PART C

# TALENT AND SKILL DEVELOPMENT

CHAPTER 14

# MUSICAL EXPERTISE

This chapter treats six connected issues having to do with musical expertise. Section 14.1 examines the difficulties associated with characterizing expertise in a way that offers a genuine foothold for cognitive psychology, and I suggest that expertise may not, in fact, be 'special' in any cognitively interesting sense. Section 14.2 goes on to review some experimental studies of music, which suggest that most members of a culture possess tacit musical expertise, expressed in their ability to use high-level structural information in carrying out a variety of perceptual tasks. This expertise seems to be acquired through casual exposure to the musical forms and activities of the culture. Section 14.3 provides two detailed examples of exceptional musical expertise (a musical savant and a jazz musician) that apparently developed in the absence of formal instruction, suggesting that normal and 'exceptional' expertise may be parts of a single continuum. The evidence presented in section 14.4 suggests that a major difference between musical expertise and many other forms of expertise is that musical expertise requires an apprehension of a structure–emotion mapping. Without this, the ability to perform with 'expression' cannot be acquired. Section 14.5 outlines some evidence to suggest that these structure–emotion links become firmly established during middle childhood, under certain conditions, and that these conditions are predictive of future development of musical expertise. Finally, section 14.6 reviews some research efforts that are attempts to clarify the precise nature of the structure–emotion link and are showing that definite types of structures seem to mediate distinct emotions.

## 14.1 What is expertise?

In beginning to think about how a psychologist who deals with music could contribute in a specific way to a volume on expertise, it became clear to me that most of the recently published work on musical competence has made little attempt to define or characterize musical expertise. What we

have, instead, is a varied collection of empirical studies on single aspects of what some musicians do. The topics of such studies range from pitch memory (Ward and Burns, 1982), through synchronization in performance (Rasch, 1988), to planning a composition (Davidson and Welsh, 1988), and it is not immediately clear that such accomplishments have anything in common other than the fact that they are different aspects of handling the organized sounds our various societies label as music.

That observation led back to a logically prior question: Is there anything that all examples of expertise *in general* should or might have in common? More precisely, is there anything about the *internal* psychological structures of certain accomplishments that marks them out as examples of expertise? It is important to remember that when someone is declared an expert, that is a social act that may or may not correspond to an intrinsic characteristic of the person so designated.

One possible definition of an expert is 'someone who performs a task significantly better (by some specified criterion) than the majority of people'. According to this definition, Chase and Ericsson's (1981) digit memorizer SF is an expert. If, however, digit-span recall became a popular hobby, then he might well be overtaken by sufficient numbers of people so that he would cease to be considered an expert. Such a relativistic attribution of expertise clearly would preclude the possibility of any *cognitive* account of expertise, because the cognitive apparatus that earned SF expert status would remain precisely the same after SF was no longer labeled an expert. It does, however, seem to me that exactly such a relativistic conception underlies much common talk of expertise, and to a certain extent determines the agendas of 'expertise' research.

For cognitive psychology to have an authentic foothold, we have to find a characterization of expertise that will allow any number of people (up to and including all) to be expert in a particular area. For instance, many would, I think, agree that the vast majority of people are expert speakers of their native languages. I shall later suggest that the majority of our population possess particular types of musical expertise. A possible definition with this outcome might relate to the reliable attainment of specific goals within a specific domain. So, for instance, one is an expert diner if one can get a wide variety of foodstuffs from plate to mouth without spilling anything.

An apparent problem with this definition, however, is that there is no lower limit to the simplicity or specificity of the task to which one can apply it. For instance, this definition would allow each of us to be expert at pronouncing his or her own name or at folding his or her arms. It may seem that we need more than goal attainment to attribute expertise. For instance, one may want to say that an expert is someone who can make

an appropriate response to a situation that contains a degree of unpredictability. So the expert bridge player is one who can work out the play most likely to win with a hand that the player has never seen before; the expert doctor is one who can provide an appropriate diagnosis when faced with a configuration of symptoms never before encountered. In this way we might be able to carve out precisely the set of activities in which the contributors to this book have been interested.

On further examination, however, it is not as easy to apply this distinction as it might first appear. Pronouncing one's own name can also be seen as an act requiring the handling of unpredictability. It is an act that is occasioned by cues (external or internal) that can vary. One must be able to retrieve and execute the required motor program regardless of the immediate mental context. The complexity of these apparently simple acts is soon revealed when one attempts to construct machines that can do the same tasks, as the discipline of artificial intelligence has amply documented (e.g. visual recognition [Marr, 1982]).

It is difficult for me to escape the conclusion that we should abandon the idea that expertise is something special and rare (from a cognitive or biological point of view) and move toward the view that the human organism is in its essence expert. The neonatal brain is already an expert system. 'Becoming expert' in socially defined ways is the process of connecting 'intrinsic' expertise to the outside world so that it becomes manifest in particular types of behaviors in particular types of situations. I believe that Fodor (1975), from another point of view, was articulating a similar proposal: To broadly paraphrase Fodor, 'You can't learn anything you don't already know'.

To look at expertise in this way may require reversal of some of our perspectives on familiar situations. For instance, when considering Chase and Ericsson's (1981) study of SF, it is easy to allow one's focus of attention to fall on the two hundred hours of practice that moved him from average to the world's best, implicitly equating the acquisition of the expertise with the work that went on in the practice period under observation. The perspective to which I am increasingly drawn suggests that we focus our attention instead on what SF brought to the experimental situation. SF's intimate knowledge of running times was, from this perspective, the principal manifestation of expertise that 'bootstrapped' the digit-span task, and it seems to me that the most interesting psychological considerations are how and why that knowledge came to be applied to the task in hand when it did. What determined that it would be applied after about fifteen hours of practice rather than instantaneously or not at all? A plausible answer to that question may well be 'chance' (e.g. a particular sequence of numbers that strongly reminded SF of a well-known running time).

In other words, the broad answer to the question of how SF became expert at the digit-span task is that he was able to increment his expertise by approximately 0.01 percent in a situation in which he was already expert at a number of things, including running times, that supplied the other 99.99 percent of what was needed. And each of those preceding areas of expertise was likewise resting on other forms of expertise in the same relationship in a constant, unbroken sequence back to birth and beyond. What made SF 'exceptional' in conventional terms was no more than a unique set of life experiences. In the sections that follow, I pursue some implications of this way of looking at expertise as applied to music.

## 14.2 Acquiring musical skill

One of the principal reasons for studying expertise is practical. Given that it would be socially desirable for certain manifestations of expertise to be more widespread than they are, we want to know what we can do to assist people to acquire them. The issue becomes acute in relation to formal education, where the general perception is that we set up environments that are supposed to encourage expertise, but that many individuals still do not achieve levels that we know to be possible (whether it be learning a foreign language, a musical instrument, or physics skill). We want to be able to tell teachers that there are principled things that they can do to increase the frequency of those 0.01 percent increments in learning.

Music is no exception to this, and music teachers are continually inquiring of psychologists how psychological insights can inform their work. It is their perception that musical expertise is taught and acquired with great difficulty. They speak of 'tone-deaf' children (usually children unable to sing in tune); they speak of the difficulty of teaching sight-reading, of teaching rhythm, of teaching good intonation on a string instrument, and so on.

My early research on the skill of sight-reading has been summarized elsewhere (Sloboda, 1984). That research was carried out under the influence of the previously published work of Chase and Simon (1973) on chess perception. Their research showed that, like playing chess, reading of music depended on an ability to pick up various sorts of patterns in the stimulus. For instance, good sight-readers were found to be much more prone than poor sight-readers to a sort of 'proofreader's error' (Sloboda, 1976a) whereby notational mistakes out of character with the genre were automatically corrected back to what the genre would have predicted. Their ability to use music structure to 'chunk' notes could account for their superior short-term memory for notation (Halpern and Bower, 1982; Sloboda, 1976b).

Encouraging as it was to find results for music that so clearly paralleled Chase's findings, I became progressively more disheartened as I talked about those results to groups of teachers. The question they all asked was of what prescriptions I would draw from my results for the teaching of sight-reading, and after some hand waving I really had to admit that there were no prescriptions that I could draw at that time. I did not know how one could teach children to 'see' structures.

Since then I have come to realize that in order to 'see' musically significant structures, one first must be able to 'hear' those structures, and I have learned from reading some excellent recent research that the process of coming to 'hear' musical structure is a process that occurs quite naturally for the majority of children as a function of normal enculturation. For instance, Zenatti (1969) showed that children at age 7 show a distinct memory advantage for sequences conforming to rules of normal tonal progression, as compared with atonal sequences. This advantage is not shared by children of age 5. Similar results were obtained from studying children's songs (Dowling, 1982, 1988; Gardner, Davidson, and McKernon, 1981). There is a definite age progression from tonal inconsistency and instability toward conformity to the norms of the tonal culture.

An experiment I conducted earlier (Sloboda, 1985a) showed that the progressing of the ability to discriminate between 'legal' and 'illegal' sequences did not seem to depend on children's receiving any sort of formal music instruction. Almost no children at age 5 made meaningful discriminations, whereas almost all 11-year-olds made discriminations in accordance with those of adults (and music harmony textbooks). The children who were receiving formal music lessons did not fare better than other children.

Although many experiments with adults have shown cognitive differences between musicians and nonmusicians, some studies have shown little difference. For instance, Deliège and El Ahmahdi (1990) showed that musicians and nonmusicians were remarkably similar in the segmentations they suggested for an atonal piece. That may have been partly attributable to the relative unfamiliarity of the genre to both groups. More strikingly, Bigand (1990) showed that nonmusicians had an ability similar to that of musicians to classify superficially different conventional tonal melodies into groups containing underlying structural similarities. Studies of memory recall for melodies (Sloboda and Parker, 1985) have shown that musicians and nonmusicians have similar abilities to preserve higher-order structure at the expense of note-to-note detail.

The research literature, therefore, leads to the conclusion that human beings pick up quite high-level implicit (or tacit) knowledge about some major structural features of the music of their culture. They gradually

improve their ability to do this over the first ten or more years of life and preserve this ability into adulthood. We may presume that this is achieved through informal engagement in the everyday musical activities that abound in almost all human cultures (e.g. nursery rhymes, hymns, dances, popular songs, playground games). In our own culture these forms are, of course, massively reinforced through the broadcast media.

In this way, almost every member of a culture is a musical expert, but the expertise is usually hidden and tacit. It may not exhibit itself in abilities to sing or play. It is, however, manifested in a variety of perceptual and memory tasks. Nearly all of us can identify some kinds of 'wrong notes' when we hear them, even though we cannot always say why the notes are 'wrong'.

Tacit expertise depends, in part, on being in a culture in which one is exposed to products in the specified domain without the necessity for active engagement. This allows the dissociation between receptive expertise and productive expertise. Such a dissociation would not normally occur in chess, or bridge, or physics, because the only way one normally gets exposure to the relevant structures is by *doing* the activity.

It is not the purpose of this essay to give an account of the various developments in understanding what it is that people know when they know about music structure. Suffice it to say that it seems necessary to postulate mechanisms for representing music that are multidimensional and hierarchical. This means that music can be characterized by points of greater or lesser prominence or distance from one another and that various dimensions may be in synchrony or in opposition. This gives rise to complex patterns of tension and resolution at different hierarchical levels. Some of the most influential characterizations of musical representation have been offered by Lerdahl and Jackendoff (1983), Krumhansl (1990), and Meyer (1973).

More pertinent for our current purposes is the observation that at least some of these structures seem capable of being represented in a connectionist network (Bharucha, 1987). A connectionist model of the brain shows one way in which it might be possible for knowledge of complex structures to be built up simply as a result of frequent exposure to relevant examples. Such an activity seems to be an essential requirement of any mechanism that acquires expertise from environments that are not engineered to be instructional (i.e. most environments).

## 14.3 Acquisition of musical expertise in noninstructional settings

Musical expertise, in the foregoing sense, is possessed by the majority of untutored members of any culture. This is not, however, what most people

mean when they refer to musical expertise; they mean overt skills of performance or composition. Surely these cannot be acquired other than through formal instruction. It is certain that such skills are acquired mainly through instruction, at least in our culture, but there is some evidence that such instruction is not necessary. Several cases of overt expertise have apparently arisen without any formal tuition or intervention by other experts. An examination of these cases is particularly important if we are to isolate the general conditions for the acquisition of expertise.

### 14.3.1 Musical prodigies and savants

There have been several documented cases of children who showed exceptional precocity at various musical skills. Some of them, such as Mozart, went on to become exceptional adults. Others did not sustain their exceptionality into adult life (see Bamberger, 1986, for a cognitive account of adolescent 'burnout' among musical prodigies). One of the fullest accounts of a child musical prodigy was given by Revesz (1925), who made an intensive study of the young Hungarian prodigy Erwin Nyherigazy (EN). Although EN had a great deal of formal tuition and support from professional musicians from an early age, he soon surpassed his teachers in his ability to commit tonal piano music to memory on one or two exposures.

There is another group of prodigies who, by and large, do not receive formal instruction: the so-called idiots savants (see Treffert, 1988, for a recent review). The savant is a person of generally low IQ, usually male, and often autistic, who has developed a skill in one defined area to a level quite exceptional compared with the general population. Although such cases have been reported in the literature for many years, the reports have mostly been only anecdotal and impressionistic contributions to the psychiatric literature. Only in the past decade have systematic investigations of musical savants been reported in the cognitive literature (e.g. Miller, 1987).

One of these studies concerned the autistic savant NP (Sloboda, Hermelin, and O'Connor, 1985). At the time of detailed study, NP was in his early twenties, and we were able to document his ability to recall a tonal piano movement almost perfectly twelve minutes after first hearing it. Two features of the study were particularly noteworthy: (1) His ability did not extend to a simple atonal piece, and (2) the few errors in his recall of the tonal piece were largely in conformity with the rules of the genre. We concluded that NP's recall ability was predicated on his ability to code and store tonal music in terms of its structural features. In that respect, NP's ability was every bit as 'intelligent' as the memory performance of chess masters. Other studies of musical savants (Hermelin, O'Connor, and

Lee, 1987; Miller, 1987; Treffert, 1988) have confirmed the importance of structural knowledge in supporting their skills.

Because NP was still relatively young when studied, it was possible to talk to people who knew him at different points in his life and observed his ability develop. It seems that NP's early life was one of considerable cultural deprivation. As far as we know, he had few, if any, opportunities to interact with musical instruments and was not encouraged to sing or to engage with music. His precocity was first noticed at about the age of 6 years, when he spontaneously reproduced at the piano a song that a staff member at his day-care center had just played. From the point on, he was given many opportunities and encouragements to interact with music and musical instruments, although nothing approaching 'instruction' was ever possible with this profoundly nonverbal individual. Even now his 'lessons' consist of a pianist playing pieces that NP then reproduces. A tape recording of his accomplishments at the age of 8 years shows memory and performance skills that were impressive for an autistic child, though by no means as polished and outstanding as his current performances.

How did NP's skill compare with 'normal' skill at the various stages of his life? At age 6 or 7, it was not clear that his memorization abilities were abnormally good. Most untutored children of that age are capable of memorizing short songs, and many can succeed in picking them out on a piano by a process of trial and error. What distinguished NP at that age was his ability to map his internal knowledge of songs directly and without error onto the piano keyboard and to choose appropriate fingering patterns. His performances of tonal music have always been characterized by an absence of hesitation or experimentation, no doubt assisted by his possession of absolute pitch. We have no information that would help us to explain how NP acquired his knowledge without having had any known opportunity to practice before the age of 6.

For the period of his early twenties, the comparison with normals showed a somewhat different pattern. His technical accomplishments were then not unusual. Many reasonably proficient pianists can choose appropriate fingerings for musical passages immediately and automatically. What made NP

be a high degree of intrinsic motivation for engagement with a single activity sustained over many years. Such motivation usually has a strong obsessional component, in that given freedom, the savant will spend all available time on the activity, without ever tiring of it.

The second factor is an environment that provides frequent opportunities for the practice of the skill in question. In the case of a musical savant, this may include the provision of regular access to instruments, broadcast media, and musical events. It is possible to suppose that whatever level of cultural deprivation NP suffered during his earliest years, he at least would have been exposed to music through the broadcast media.

The third factor is, of course, the exceptional amount of time spent in cognitive engagement with the materials and activities relevant to the skill in question (practice). It is difficult to estimate the amount of time NP spent thinking about music when not playing or listening to it, but obvious external involvement probably amounted to four to five hours per day.

The fourth factor, therefore, is the availability of the time and opportunity to 'indulge' the obsession. It may be because fewer societal demands are made on people with low IQs that they are 'allowed', even encouraged, to devote their attentions in this way.

The fifth factor is the complete absence of negative external reinforcement related to attainment or lack of it. There is, therefore, little possibility of a savant's developing self-doubt, fear of failure, or any of the other blocks that inhibit and sometimes prevent normal or exceptional accomplishment.

### 14.3.2 Jazz musicians

It is probable that many of the world's musical cultures, particularly the informal, nonliterate 'folk' cultures, have been breeding grounds for expertise. Some anthropological work (e.g. Blacking, 1976) suggests that this is true of indigenous Third World cultures. The jazz culture of New Orleans in the early part of this century may not have been greatly different from those other cultures in many respects. Its advantage for us is that jazz rapidly spread from New Orleans to become part of mass culture and contributed an entirely new facet to the face of Western culture. Its leaders became cult heroes, and jazz itself became a subject for intensive academic scrutiny. For these reasons, we have far more detailed biographical information about jazz musicians than about the musicians from all of the world's other nonliterate cultures put together.

It appears that most of the early jazz players were self-taught. Among the self-taught players who became international names were Bix Beiderbecke, Roy Eldridge, and Louis Armstrong. Collier's (1983) study of Armstrong is particularly detailed, and it allows us to look at Armstrong's

musical development in some detail as a 'prototype' of untutored expertise.

Armstrong spent most of his early years in a neighborhood known as 'Black Storeyville', an area designated for black prostitution. One of the features of that neighborhood was the continual live music, performed by dance bands and 'tonk' bands, which often would play on the street to attract custom. Having little knowledge of the world outside, Armstrong had little more than pimps and musicians as male role models. His father had abandoned his mother before he was born. His childhood was one of extreme poverty and deprivation, and from the age of 7 years he had to work, steal, and hustle to make money for his mother and himself. At the age of 8 or 9 years he formed a vocal quartet with some other boys in order to pick up pennies on street corners. The group lasted two or three years and probably practiced and performed in public two or three times per week. That provided several hundred hours of improvised part singing, which as Collier observed, 'would have constituted a substantial course in ear training—far more than most conservatory instrumentalists get today'.

At the age of 13 or 14 years, Armstrong was involved in an incident with a gun and was, as a result, sent to the Colored Waif's Home (known as the Jones Home). There the boys were taught reading, writing, and arithmetic, with gardening as a sideline. The home had a band that played once a week around the city. After six months in the home, Armstrong was allowed to join the band, first playing tambourine, then drums, then alto horn. It is clear from contemporary accounts that many of the bands playing in the streets of New Orleans were fairly informal groups with an 'anything goes' attitude. It was quite easy for a novice to join in the general noise, just playing the notes he knew, and his mistakes and split notes would pass without comment. Armstrong quickly learned how to get sounds out of the horn, and his vocal experience made it easy for him to work out appropriate parts to the songs the band played. His talent was noticed, and he was promoted to bugle player. He gradually improved to become the band's leader, but he left the home and the band after two years, at age 16. Nothing he experienced in the home would merit the term 'formal teaching'.

Armstrong found casual work driving a coal cart, which occupied his days, but during the evenings he began playing jazz in the blues bands of the tonks. He did not at that stage own a cornet, and so it was impossible for him to practice. He simply went around to the various bands asking cornetists to let him sit in for a few numbers. Blues music provided a good vehicle for gaining jazz expertise. Blues songs featured slow tempos in two or three of the easiest keys. The set melodies were of the simplest sort; in

many cases there was no set melody at all, and the cornetist would string phrases together from a small repertoire of stock figures.

At age 17, Armstrong acquired his first cornet and began to practice and work regularly at one of the tonks. The work paid little, and so he kept his coal job during the day. At some point in that period Armstrong met Joe Oliver, acknowledged as the best cornetist in New Orleans. Armstrong began hanging around the places where Oliver played, running errands, carrying his case, and eventually sitting in for him. Oliver became Armstrong's sponsor and to some extent his teacher. According to Collier, however, Oliver did not influence Armstrong's style and probably did little more than show Armstrong some new tunes and possibly a few alternative fingerings.

By age 19, Armstrong was finding employment on local riverboat excursions. Then, for three summers running, he made long trips, playing every day. For the first time in his life music had become his predominant activity. The band played seven nights per week, doing fourteen numbers and encores each night. They rehearsed two afternoons per week, and the repertoire changed every two weeks. It was only after joining the riverboats that Armstrong learned how to read music and had to acquire the discipline of playing what was written rather than what he felt like playing. When he left the riverboats at age 23, he was an established professional musician.

If Armstrong's early life was a prototype for untutored acquisition of expertise, which of its features might we highlight for future corroboration? One obvious feature was the casual immersion in a rich musical environment with many opportunities to listen and observe. A second feature was the early systematic exploration of a performance medium (in his case, voice). Third, as far as we can judge, his early experiences allowed a great deal of freedom to explore and experiment without negative consequences. A fourth feature was a lack of distinction between 'practice' and 'performance'. The learning took place on the job. A fifth feature was an enduring motivation to engage in music—in Armstrong's case, a complex mix of internal and external motivations, but arguably with internal motivations dominating. A sixth feature was a graded series of opportunities and challenges available or sought out as the expertise developed.

In many ways, this list of features fits the case of a savant such as NP. The principal differences in the two examples cited here relate to motivation and challenge. NP's motivation did not have a significant external component, and partly for that reason it is not clear that his challenges either arose or were grasped with the same frequency as those of Armstrong. It is easy to imagine NP remaining on a performance plateau. Armstrong went on growing and changing throughout his life.

What these case studies show is that high levels of expertise are achievable without instruction. This does not, of course, mean that instruction is useless. By providing a structured progression of information and challenges for a learner, geared precisely to the learner's capacities at a given time, a teacher may be able to accelerate a learner's progress. Not every person has the opportunity to extract the relevant experiences from the 'natural' environment that Armstrong had. A formal instructional environment can engineer the conditions for such extraction. The danger of all such environments is that goals and standards are imposed on the learner, rather than being chosen. The consequence can be to inhibit intrinsic motivation and originality (Amabile, 1983). If external constraints are extreme, it may even be that the ability to enjoy music will be destroyed.

In this connection, one other difference between NP and Armstrong has not been brought out thus far. One of the most striking aspects of NP's musical life was its lack of affect. All pieces in his repertoire were played in a 'wooden', unexpressive manner. Although his immediate reproduction showed some of the expressive features of the model, within twenty-four hours all expressive variation was 'washed out', leaving a rigid metronomical husk. It was as if NP had no means of understanding (and thus relating to the structure of) the small variations in timing, loudness, and timbre that are the lifeblood of musical performances. From the earliest recording we have of Armstrong's music, in contrast, we find a richly expressive, flexible performance that bends tone and time in ways that have strong impact on many listeners. Armstrong is not hailed as the king of jazz for his technique, impressive as it was. There are others who match or surpass him in technique. He is revered for the life he could breathe into the simplest material.

NP was one of a rather small number of people who appear to gain complete satisfaction from relating to music as pure structure or syntax. What brings the vast majority of us to music, and keeps us with it, is something additional: its power to mediate a vast range of emotionally toned states, ranging from the subtle to the overwhelming. Because modern systematic studies of music have approached it with the tools of cognitive science and linguistics, the emotional aspect of music has been virtually overlooked, and naive readers of modern research studies might be forgiven for thinking that music is simply another kind of complex structure to be apprehended, like chess or physics.

I know that those who are expert in chess or physics say that there is beauty and emotion in those activities too, but there is a sense in which such things are not central to the skill. One can write a perfectly effective computer program for chess that will not need any information about how

particular chess positions or games will affect the emotions of certain human players. I think there is a strong case for saying that a computer could never adequately simulate Louis Armstrong without some implementation of a theory of the emotions.

## 14.4 Expression and emotion as foundational aspects of musical expertise

Those approaching music with the prejudices and preoccupations of experimental psychology have been wary of examining the emotional aspect, for methodological and conceptual reasons. Rather than examine these reasons in detail, I should like to point to some recent investigations that seem to have 'opened doors' into this area.

The advent of the microcomputer and microtechnology has, for the first time, made possible easy and accurate transfer of detailed performance information into computers for sophisticated analysis. The past decade has seen a number of studies (Clarke, 1985; Gabrielsson, 1983; Shaffer, 1981; Sloboda, 1983; Sundberg, 1988; Todd, 1985) that have measured minute expressive variations in performance loudness and timing. These studies have shown several things: (1) A given player can consistently repeat given variations on successive performances; (2) these perturbations are not random but, rather, are intentional, and performers can alter them to a greater or lesser extent at will; (3) many of these perturbations are rule-governed and relate to the formal structure of the music in systematic ways.

My own studies (Sloboda, 1983, 1985b), for instance, have shown that timing deformations are organized around the strong metrical beats of tonal melodies in a way that makes the metrical structure clearer for listeners than it is when such deformations are not present. Although we do not yet have the evidence, this line of research suggests that all effective expression may be systematic and rule-governed in this way, helping to highlight musical structures in a way that makes their emotion-bearing content more manifest to listeners.

The other line of contemporary thinking that converges with the experimental work on expression is the music-theory work of such writers as Leonard Meyer (Meyer, 1956, 1973) and Fred Lerdahl (1988a, 1988b; Lerdahl and Jackendoff, 1983). Meyer has convincingly argued that emotion in music arises out of the complex, often subliminal web of expectations and violations of expectations that musical structures unfold over time (Narmour, 1977). Lerdahl (1988b) takes this a step farther by suggesting that only structures that have certain formal properties (such as discreteness and hierarchical organization) can be directly detected by listeners (Balzano,

1980; Shepard, 1982). Only such structures will be effective in creating the types of tensions and resolutions that can support the emotional activities and responses peculiar to music. Lerdahl has particularly enraged certain sections of the avant-garde music community by claiming that traditional tonal music satisfies his criteria, whereas such forms as serial music do not. This could be used as an explanation of why tonality has been able to resist all attempts to oust it from center stage in music and why many avant-garde genres have but limited appeal. The general thrust of all this thinking about music gets independent support from cognitive theorists (e.g. Ortony, Clore, and Collins, 1988) who characterize the cognitive substrate of all emotion in terms of the violations of various classes of expectations.

These strands of work lead toward the following set of working hypotheses about the vast central bulk of the world's music:

1. One major function of music is to suggest or mediate a range of emotional responses.
2. Common musical structures have particular perceptible properties that support the patterns of expectation underlying such emotions.
3. Expression in musical performance has the effect of making these structural features more prominent, and thus of heightening the emotional response.

## 14.5 The roots of musical expertise

At the beginning of this chapter, I asked whether all aspects of musical expertise have anything in common. By a rather circuitous route I now come to a proposed answer, which is that they involve apprehension and use of the structure–emotion link. At whatever level, and for whatever activity, what makes the behavior *musically*, as opposed to technically or perceptually, expert is its manifestation of this link. I take it as axiomatic that emotions do not have to be learned (although they may be refined and differentiated through experience). They are part of the 'expert system' with which we are born. So what must be learned is how to apprehend those features of musical structures that can be mapped onto and therefore evoke our existing emotions.

Hevner's (1936) pioneering work showed that adult members of a culture generally agree on the emotional characterization of a passage of music, in that they tend to select similar adjectives to describe it (e.g. majestic, gloomy, playful). Gardner (1973) has shown that this ability develops through childhood, with younger children able to use only rather crude

descriptions (such as 'loud' or 'jumpy'). It is, of course, possible that particular kinds of music have come to acquire conventional meanings by routes that do not involve the listener's own emotions. Laboratory studies of people's abilities to *describe* music do not show how these abilities were acquired.

Direct observational studies of children's emotional responses to music have been rare. Moog's (1976) studies showed that preverbal infants could demonstrate quite strong expressions of delight or fear on hearing music. The available evidence suggests that tone quality is the aspect of music that elicits the strongest early reactions. Smooth, treble-register sounds seem to elicit the strongest reactions of attention and pleasure. Most children below the age of 5 years seem not to be particularly interested in unpitched rhythms and seem not to differentiate emotionally between music played in conventional harmony and that played dissonantly.

As children grow older, it is less easy to record emotional responses by direct observation. Socialization leads to significant suppression of direct emotional expression. An alternative approach that I have been pursuing (Sloboda, 1989) is to ask adults to recall musical experiences from the first ten years of life. The literature on autobiographical memory (Brown and Kulik, 1977; Rubin and Kozin, 1984) suggests that experiences connected with significant emotion may be particularly retrievable. The method also has the advantage of tapping musical experience in a range of naturalistic contexts, rather than in restricted experimental contexts. In addition to asking these adults for information about childhood events and their contexts, I also ask them if those experiences had any particular significance for them. Information about the involvement of music in their lives, including formal music tuition, is also collected.

The findings from these studies indicate that most subjects seem to be capable of producing at least one memory. Some people readily recalled as many as ten different events. No event was recalled from an age earlier than 3 years, but from 4 to 10 years the age spread was fairly even. Analysis of the words used by adults to describe the character of their experiences (both of the music itself and of their reaction to it) showed an interesting age progression. Memories from around age 5 tended to characterize music in rather neutral descriptive terms (e.g. 'fast', 'loud', 'simple'), and the responses to it in terms of general positive enjoyment (e.g. 'love', 'like', 'enjoy', 'excited', 'happy'). Looking back to age 8, subjects characterized music in terms of its affective or sensual characteristics (e.g. 'beautiful', 'liquid', 'funny'), and the responses to it were recalled in terms of wonder or surprise (e.g. 'enthralled', 'incredulous', 'astounded', 'overwhelmed', 'awe-struck'). Finally, harking back to around age 9, some

memories contained strong feelings of sadness (e.g. 'melancholy', 'sad', 'apprehensive').

It is of particular significance that the ability to respond to music in terms of wonder arises at about the age when children can be shown to distinguish reliably between tonal and atonal music. This strongly suggests that the particular violations of expectations that mediate some of the more 'advanced' emotional responses to music require the ability to represent music in terms of the structural categories of tonal music. It is also significant that the progression of responsivity seems to owe nothing to explicit formal instruction. The majority of the experiences reported *preceded* the onset of formal musical training, and in several cases such an experience spurred the child to seek instruction. Learning the structure–emotion link seems to proceed in the absence of formal instruction.

Some of the memories reported clearly had the status of what some people call 'peak experiences'—unusual and deeply rewarding experiences of a complex emotional/intellectual character. The research showed that people who have had such peak experiences were more likely than others to pursue involvement with music for the rest of their life. The experiences provided a strong source of internal motivation to engage with music in a systematic way (arguably in part to increase the likelihood of replicating the experiences). Educators wishing to raise the general level of musical skill might well be advised to consider how they can help increase the frequency of such experiences in the population, because it is clear that not every child has them.

The memory research provided some interesting clues on this latter point as well. It was discovered that almost none of those peak experiences had occurred in situations of external constraint or anxiety. The most likely environment for a peak experience was at home, on one's own or with friends and family, and while listening to music. The least promising environment was at school, with teachers, while performing. The individual stories graphically revealed the kinds of anxieties and humiliations many children were made to suffer in relation to music by insensitive adults or through insensitive educational practices. These acted as strong disincentives to further engagement with music and seemed to block the possibility of making links between emotions and the intrinsic characteristics of music.

A similar lesson emerges from a recent study of leading American concert pianists by Sosniak (1989). None of those in her sample showed exceptional promise as a child, but in every case their early lessons were associated with fun and exploration, rather than with practical achievement. It seems that, at least for the crucial early stages of musical development, there is no special strategy we should recommend to educators, other than to stop worrying

about particular apparent skill deficiencies and concentrate on not getting in the way of children's enjoyment and exploration of music. In such contexts, children become natural experts who spontaneously seek what they require to bring their expertise to bear on particular practical accomplishments.

## 14.6 Musical structure and emotion

The final question I wish to raise in this chapter concerns the precise nature of the structure–emotion link: What structures elicit what emotions, and why? Although musicologists have long debated this point (e.g. Cooke, 1959; Meyer, 1956), there have been remarkably few attempts to collect empirical data on it. A few physiological studies (e.g. Goldstein, 1980; Nakamura, 1984) have shown that reliable changes in such indices as heart rate and skin conductance can be shown as people listen to specific pieces of music. But such studies generally have not involved subjecting the music itself to detailed structural analysis. A particular characteristic of emotional responses to music is that they often change in nature and intensity over the duration of a piece and are linked to specific events (rather than being a general 'wash' of a particular mood). In this respect, they are similar in nature to emotional responses to drama or fiction. To my knowledge, no published studies provide data on the specific points in musical compositions at which intense or peak emotional experiences take place. One problem is that it is difficult to get intersubjective agreement on how to characterize these experiences. Some of my own recent research entails an attempt to circumvent this problem by asking people to report (retrospectively, at this stage) on the locations in musical compositions at which they reliably experience direct physical manifestations of emotion (e.g. tears, shivers). A significant minority of subjects have been willing and able to do this and have provided a corpus of some 165 'moments' of reliable emotional response. Full details of this study are reported in Sloboda (1991). An analysis of the subset comprising classical instrumental excerpts has revealed three clusters of structural features associated with three different types of responses. These are summarized in Table 14.1.

This pattern requires confirmation with other types of music and also by direct observation in experimental situations. If confirmed, it will show that many of the emotional responses to music require that the listener, at some level, represent high-level structure. For instance, one cannot define 'melodic appoggiatura' apart from a description of music in terms of strong and weak beats within a metrical structure and of discord and resolution within a tonal framework. This is one reason we find it difficult to respond

**Table 14.1** Emotion and musical structure

| Emotional response | Associated structural features |
|---|---|
| Tears or lump in throat | Melodic appoggiatura |
| | Melodic or harmonic sequence |
| | Harmonic or melodic acceleration to cadence |
| Shivers down spine or goose pimples | Enharmonic change |
| | Delay of final cadence |
| | New or unprepared harmony |
| | Sudden dynamic or textural change |
| | Descending circle of 5ths in harmony |
| Racing heart and 'pit of stomach' sensations | Harmonic or melodic acceleration |
| | Sudden dynamic or textural change |
| | Repeated syncopation |
| | Prominent event arriving earlier than expected |

emotionally to the music of other cultures as do the members of those cultures. We have not yet assimilated the means of representing their musical structures that would allow the appropriate structure–emotion links to be activated.

We have many interesting and important questions to explore, such as why these particular structures mediate these particular emotions in the way that they do. Research, however, has begun to clarify a major strand in musical expertise that distinguishes it starkly from the other forms of expertise represented in this volume. It suggests that the central conditions for acquisition of musical expertise are as follows:

1 Existence in a musical culture of forms that have perceptible structures of certain kinds (as specified by Lerdahl and others)

2 Frequent informal exposure to examples of these forms over a lifetime

3 Existence of a normal range of human emotional responses

4 Opportunity to experience these emotions mediated through perceived musical structures, which in itself requires

5 Opportunity to experience music in contexts free of externally imposed constraints or negative reinforcements

If we can ensure these conditions, then the problems associated with bringing individuals to levels of achievement we would currently regard as exceptional may turn out to be trivial.

## References


Amabile, T. M. (1983) *The social psychology of creativity.* New York: Springer-Verlag.

Balzano, G. J. (1980) The group-theoretic description of twelvefold and microtonal pitch systems. *Computer Music Journal,* **4**, 66–84.

Bamberger, J. (1986) Cognitive issues in the development of musically gifted children. In R. J. Stemberg and J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 388– 416). Cambridge University Press.

Bharucha, J. J. (1987) Music cognition and perceptual facilitation: A connectionist framework. *Music Perception,* **5**, 1–30.

Bigand, E. (1990) Abstraction of two forms of underlying structure in a tonal melody. *Psychology of Music,* **19**, 45–59.

Blacking, J. (1976) *How musical is man?* London: Faber.

Brown, R. and Kulik, J. (1977) Flashbulb memories. *Cognition,* **5**, 73–99.

Chase, W. G. and Ericsson, K. A. (1981) Skilled memory. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 141–189). Hillsdale, NJ: Erlbaum.

Chase, W. G. and Simon, H. A. (1973) The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215–281). New York: Academic Press.

Clarke, E. F. (1985) Structure and expression in rhythmic performance. In P. Howell, I. Cross, and R. West (Eds.), *Musical structure and cognition* (pp. 209–236). London: Academic Press.

Collier, J. L. (1983) *Louis Armstrong: An American genius.* New York: Oxford University Press.

Cooke, D. (1959) *The language of music.* London: Oxford University Press.

Davidson, L. and Welsh, P. (1988) From collections to structure: The developmental path of tonal thinking. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation and composition* (pp. 260–285). London: Oxford University Press.

Deliège, I. and El Ahmahdi, A. (1990) Mechanisms of cue extraction in musical groupings: A study of perception, on *Sequenza VI* for viola solo by Luciano Berio. *Psychology of Music,* **19**, 18–44.

Dowling, W. J. (1982) Melodic information processing and its development. In D. Deutsch (Ed.), *The psychology of music* (pp. 413–430). New York: Academic Press.

Dowling, W. J. (1988) Tonal structure and children's early learning of music. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation and composition* (pp. 113–128). London: Oxford University Press.

Foder, J. A. (1975) *The language of thought.* Hassocks, Sussex: Harvester Press.

Gabrielsson, A. (1988) Timing in music performance and its relation to music experience. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation and composition* (pp. 27–51). London: Oxford University Press.

Gardner, H. (1973) Children's sensitivity to musical styles. *Merrill-Palmer Quarterly of Behavioral Development*, **19**, 67–77.

Gardner, H., Davidson, L., and McKernon, P. (1981) The acquisition of song: A developmental approach. In *Documentary report of the Ann Arbor Symposium*. Music Educators' National Conference, Reston, VA.

Goldstein, A. (1980) Thrills in response to music and other stimuli. *Physiological Psychology*, **8**, 126–129.

Halpern, A. R. and Bower, G. H. (1982) Musical expertise and melodic structure in memory for musical notation. *American Journal of Psychology*, **95**, 31–50.

Hermelin, B., O'Connor, N., and Lee, S. (1987) Musical inventiveness of five idiots-savants. *Psychological Medicine*, **17**, 79–90.

Hevner, K. (1936) Experimental studies of the elements of expression in music. *American Journal of Psychology*, **48**, 246–268.

Krumhansl, C. (1990) *Tonal structures and music cognition*. New York: Oxford University Press.

Lerdahl, F. (1988a). Tonal pitch space. *Music Perception*, **5**, 315–350.

Lerdahl, F. (1988b). Cognitive constraints on compositional systems. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation and composition* (pp. 231–259). London: Oxford University Press.

Lerdahl, F. and Jackendoff, R. (1983) *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Marr, D. A. (1982) *Vision*. San Francisco: Freeman.

Meyer, L. B. (1956) *Emotion and meaning in music*. Chicago: University of Chicago Press.

Meyer, L. B. (1973) *Explaining music*. Berkeley: University of California Press.

Miller, L. K. (1987) Sensitivity to tonal structure in a developmentally disabled musical savant. *Psychology of Music*, **15**, 76–89.

Moog, H. (1976) *The musical experience of the preschool child* (C. Clarke, Trans.). London: Schott.

Nakamura, H. (1984) Effects of musical emotionality upon GSR and respiration rate: The relationship between verbal reports and physiological responses. *Japanese Journal of Psychology*, **55**, 47–50.

Narmour, E. (1977) *Beyond Schenkerism: The need for alternatives in music analysis*. Chicago: University of Chicago Press.

Ortony, A., Clore, G. L., and Collins, A. (1988) *The cognitive structure of the emotions*. Cambridge: Cambridge University Press.

Rasch, R. A. (1988) Timing and synchronization in ensemble performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation and composition* (pp. 70–90). London: Oxford University Press.

Revesz, G. (1925) *The psychology of a musical prodigy*. London: Kegan Paul, Trench, and Trubner.

Rubin, D. C. and Kozin, M. (1984) Vivid memories. *Cognition*, **16**, 81–95.

Shaffer, L. H. (1981) Performance of Chopin, Bach, and Bartok: Studies in motor programming. *Cognitive Psychology*, **13**, 326–376.

Shepard, R. N. (1982) Structural representations of musical pitch. In D. Deutsch (Ed.), *The psychology of music* (pp. 344–390). New York: Academic Press.

Sloboda. J. A. (1976a) The effect of item position on the likelihood of identification by inference in prose reading and music reading. *Canadian Journal of Psychology*, **30**, 228–236.

Sloboda, J. A. (1976b) Phrase units as determinants of visual processing in music reading. *British Journal of Psychology*, **68**, 117–124.

Sloboda, J. A. (1983) The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology*, **A35**, 377–396.

Sloboda, J. A. (1984) Experimental studies of music reading: A review. *Music Perception*, **2**, 222–236.

Sloboda, J. A. (1985a) *The musical mind: The cognitive psychology of music.* London: Oxford University Press.

Sloboda, J. A. (1985b) Expressive skill in two pianists: Style and effectiveness in music performance. *Canadian Journal of Psychology*, **39**, 273–293.

Sloboda, J. A. (1989) Music as a language. In F. Wilson and F. Roehmann (Eds.), *Music and child development: Proceedings of the 1987 Biology of Music Making Conference* (pp. 28–43). St. Louis: MMB Music.

Sloboda, J. A. (1991) Music structure and emotional response: Some empirical findings. *Psychology of Music*, **19**(2).

Sloboda, J. A., Hermelin, B., and O'Connor, N. (1985) An exceptional musical memory. *Music Perception*, **3**, 155–170.

Sloboda, J. A. and Parker, D. H. H. (1985) Immediate recall of melodies. In P. Howell, I. Cross, and R. West (Eds.), *Musical structure and cognition* (pp. 143– 167). London: Academic Press.

Sosniak, L. (1989) From tyro to virtuoso: A long-term commitment to learning. In F. Wilson and E Roehmann (Eds.), *Music and child development: Proceedings of the 1987 Biology of Music Making Conference* (pp. 274–290). St. Louis: MMB Music.

Sundberg, J. (1988) Computer synthesis of musical performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation and composition* (pp. 52–59). London: Oxford University Press.

Todd, N. (1985) A model of expressive timing in tonal music. *Music Perception*, **3**, 33–58.

Treffert, D. A. (1988) The idiot savant: A review of the syndrome. *American Journal of Psychiatry*, **145**, 563–572.

Ward, W. D. and Burns, E. M. (1982) Absolute pitch. In D. Deutsch (Ed.), *The psychology of music* (pp. 431–452). New York: Academic Press.

Zenatti, A. (1969) *Le développement génétique de la perception musicale.* Monographies Français Psychologiques No. 17.

CHAPTER 15

# MUSICAL ABILITY

## 15.1 What is musical ability?

The use of the term 'musical ability' may already seem to presuppose too much. Such a term suggests that there is some common factor, or set of factors, underlying all accomplishments in the sphere of music. How does this square with the fact that there are singers who cannot read music, pianists who cannot sing in tune, performers who cannot compose, and music critics who can neither play an instrument nor compose? Do all such people possess some common attribute in virtue of which they can be said to be musically able? Those who have applied the concepts and methods of contemporary cognitive psychology to music would answer 'yes' to this question. They would say that musical ability is a particular sort of acquired cognitive expertise, entailing at its core the ability to *make sense* of musical sequences, through the mental operations that are performed on sounds (whether real or imagined). The term 'make sense' is somewhat analogical. Musical idioms are not languages, and do not have referential meaning in the way that languages such as English do. They do, however, have complex multi-levelled structural features which resemble syntax or grammar (Lerdahl and Jackendoff 1983).

How do we know if and when a person is 'making sense' of some music? There are a number of ways in which psychologists have operationalized this ability. First, people who are 'making sense' of music should be able to remember music which conforms to a cultural 'language' better than music which does not. There are many demonstrations that this is so, even in fairly young children (e.g. Deutsch 1980, Zenatti 1969). This exactly parallels findings in other domains such as chess. Chess players recall meaningful (i.e. game-plausible) boards better than random boards (Chase and Simon 1973). Second, people who make sense of music tend to make plausible substitutions when asked to recall music they have just heard (Sloboda and Parker 1985, Sloboda *et al.* 1985, Oura and Hatano 1988). This is similar to the well-established finding that people rarely remember verbal information word-for-word, but reconstruct in their own words

something that *means* the same as what they heard. This kind of finding suggests that people generally store something more abstract than the actual words or notes.

A third criterion for making sense of music is the ability to correctly judge whether or not given sequences are acceptable according to cultural rules. There is strong evidence that most children can reject blatant violations, such as discords, by the age of seven, and more subtle ones, such as unfinished cadences, by the age of ten (Sloboda 1985). A final criterion is the ability to correctly identify the consensual mood or emotion of a musical passage. Again, we have rather strong evidence that the broad parameters of this ability are in place for most children by the age of five (Pinchot-Kastner and Crowder 1990).

Although people vary quite widely in the level of sophistication to which they have developed their ability to make sense of music, the available evidence points to the conclusion that the vast majority of the population have acquired a common receptive musical ability, clearly evident through experimental demonstration, by the end of the first decade of life, *regardless of accomplishment in any particular sphere of musical performance, and regardless of having been in receipt of any formal musical education or training.*

## 15.2 How is musical ability acquired?

What features of music and the human mental apparatus make this widespread ability possible? There is broad agreement between music theorists and psychologists (e.g. Bregman 1990, Deutsch 1992, Huron 1991, Lerdahl 1988, Meyer 1956, Narmour 1990) that the most prevalent musical idioms have structural and mathematical properties that make them easily analysable by universal pre-cultural mechanisms of auditory perceptual grouping. Connectionist models of learning applied to music (e.g. Bharucha 1987, Gjerdingen 1990) demonstrate one way in which complex mental representations might be built up from such simple groupings on the basis of repeated exposure to a variety of musical examples sharing similar structures. There is also the beginnings of an understanding of how the confirmation and violation of expectations built up within such a system may lead to the experience of music-induced emotion and mood (Jackendoff 1991, Narmour 1991, Sloboda 1991). We may, therefore, make the strong conjecture that exposure through enculturation to certain types of music can be a sufficient condition for the development of musical ability. What remains to be explained is why individuals develop at different rates and to different levels.

Cognitive psychologists who study expertise in general have come to a simple but far-reaching general solution to the problem of individual differences, which may be summed up in the old proverb 'practice makes perfect'. More precisely, level of expertise seems to be a monotonic function of the duration of relevant cognitive activity. Most computer-based models of expert systems, including connectionist models, function according to this principle, because the relevant learning takes place through the processing of a large number of examples. The more examples processed, the more elaborate are the knowledge structures. If exposure to music is at least *one* of the preconditions for relevant cognitive activity, it is clear that different developmental trajectories may start very early indeed. For example, pregnant mothers who sing may provide their fetus with a particularly rich early database. Lest it be thought that the neonatal brain does not have the capacity to analyse and store music, I should point to recent evidence that infants exposed to particular pieces of music before birth show distinct preferences for those same pieces after birth (Hepper 1991).

Perhaps, however, the most common way by which young people come to increase their degree of cognitive involvement with music is by starting to perform it, through singing, or by learning a musical instrument. Such involvement forces cognitive engagement in a way that mere exposure may not. For most observers, it is the ability of people to perform well which constitutes the evidence on which we judge their musical ability.

We need, however, to be quite clear that there are several distinct abilities involved in performing a musical instrument, and only some of them are 'musical' in the sense I have been discussing. The other abilities are what I would call 'technical'. For example, a pianist may be able to play a difficult passage very quickly, very evenly, and very accurately. This will be possible because of very highly developed control systems for execution of hand and finger movements. The development of such systems to high levels will have required much practice, and they are manifestations of expertise. It is, however, possible that a person could learn to play a piece of music 'technically' without having 'understood' it musically at all. Such a technically perfect performance could equally well be generated by mechanical means.

## 15.3 Expressive performance

It is well established that notationally perfect performances of music are experienced by most listeners as dull, mechanical and uninteresting. What makes any performance musically interesting are the slight fluctuations in duration, loudness, pitch and timbre which together constitute expressive performance. Such fluctuations are what distinguishes one performance

from another, and their importance is what keeps professional performers in business. Were it not for this, we could get computers to generate once-and-for-all perfect performances of musical works, and close most of our symphony halls and conservatories!

Recent research has made it increasingly clear that the existence of expressive performance is the best evidence we can obtain that musicians *understand* the music they are playing. This is because expressive variations are by no means random or idiosyncratic. A wide range of studies (e.g. Clarke 1988, Gabrielsson 1988, Shaffer 1981, Sloboda 1983) has shown that microvariations are highly systematic, both within the same performer and across different performers. Many of such variations are designed to make important structural features of the music more prominent to the listener. Their systematicity therefore depends upon the performer having understood what *are* the important structures. The best test that this understanding is deeply rooted in a performer is to ask that performer to sight-read some music he or she has never seen before. If a performer can apply appropriate expression in such a situation, then his or her understanding must be fully internalized. Evidence of such ability does indeed exist (Sloboda 1983), and it is often so deeply internalized that performers are not fully aware of the details of their own expressive repertoires; they have become intuitive and semi-automatic. This should not surprise us, because it is another well-documented characteristic of all cognitive expertise. Many people mistakenly assume that intuitive behaviour must be innate. This is a major fallacy. Any well-practised habit eventually becomes automatic.

It may seem at first sight that the statement that expression is systematic conflicts with the existence of interesting differences between performers. The contradiction is, however, more apparent than real. Performers differ from one another not so much in the nature of expressive devices used, but in their distribution and intensity. So, for example, one performer may achieve by subtle dynamic fluctuations what another achieves by temporal fluctuations. Such differences between performers, or between the same performer at different times, can be characterized as differences in expressive 'style' (Sloboda 1985).

## 15.4 Developing expressive skills alongside technical skills

Because high level musical performance has this dual technical–expressive aspect, the conditions for the development of expert performers must allow development on both fronts. This is quite difficult to achieve.

There are many music professionals who are technically adept but lack a high level of expressive skill. There are also many non-professionals who have a high level of expressive insight, but cannot realize their expressive intentions in music of any technical complexity. The master performer is that relatively rare person who has managed to develop both technical and expressive skills to a high level in tandem. How is it done?

We now have evidence from a number of biographically based research studies (Ericsson *et al.* 1993, Sloboda and Howe 1991, Sosniak 1985) that confirms that music is no exception to the general rule that 'practice makes perfect'. The study of Ericsson *et al.* shows that student violinists rated as excellent by their professors have, by the age of 21, accumulated on average around twice as many hours of practice over the lifespan (10 000 hours) as more average players (5000 hours). Sosniak, in a study of 24 top pianists in the USA, showed that none of them showed particular signs of exceptional ability at the outset of training. Exceptionality was something which developed gradually through the early years of formal training. Indeed, the notion that very early achievement is the normal precursor of adult excellence finds very little support in the documented research literature. The child prodigy may be the exception rather than the norm.

It is not very easy for a young person to accumulate 10 000 hours of practice between starting an instrument (typically around the age of six or seven at the earliest) and the age of 21. To get some idea of the workload that figure implies, one can calculate that the accumulation of 10 000 hours of practice would require two hours every day of the year for 14 years. In reality, the typical 'expert' practice pattern begins with 15–30 minutes per day, and increases to 4–6 hours per day by the age of 21. Practice activities are not always inherently rewarding to young children, and we (Sloboda and Howe 1991) found that most children in a sample of students at a specialist music school had not always enjoyed practice as young children. They often required parental support, even direct supervision, to accomplish regular practice.

Indeed, the quality of the child's relationship with significant adults, both parents and teachers, appeared crucial in predicting long-term involvement with musical performance. The parents in our sample (Sloboda and Howe 1991) were characterized by a very high level of involvement in their children's musical life. In addition, students tended to remember their first teachers, not so much for their technical adeptness as for the fact that they made lessons fun. They communicated both their love of music and their liking for their pupil. Personal warmth seemed an important characteristic. Critical, confrontational or achievement-oriented approaches seemed not to be successful, at least in the early stages of learning.

We (Sloboda and Howe 1991) also found some surprising differences between the best and the average students in our sample. It turned out that the best students had done *less* formal practice in their early years than had the average students, and that their parents were less accomplished musically than the parents of the average children. The rich qualitative data obtained from interviews with these students strongly suggests a social–motivational explanation for these findings. Children whose parents were not musically accomplished tended to receive a high level of praise and admiration from their families for objectively rather modest achievements. This, in its turn, encouraged the children to develop a strong sense of themselves as special, as the 'musician' in the family. Musical parents tended to have higher standards, which were reflected in a more rigid, achievement-oriented, family environment, with much less admiration for modest achievements.

There is some indication that the best students did not spend less time on musical activity than their less able peers, but that they simply spent less time on formal task-oriented practice and more time on free exploration of the musical medium (improvisation, and other activities often perjoratively described by both children and parents as 'messing about'). I can certainly remember from my own childhood that I would despatch my lesson tasks in 10 minutes or so, and then spend hours 'tinkering' on my own projects, with the forbearance of my long-suffering parents, who none the less were often driven to enquire 'what *is* that strange noise you are making: are you sure it is part of your practice?'

These initially unexpected results now make a lot of sense to me. One may suppose that, by and large, formal task-oriented practice encourages the development of technical rather than expressive skills, whereas exploratory and improvisatory activities encourage the individual's expressive development. Successful musicians are those who have been able to achieve a proper balance between these two types of activity.

There is a considerable body of evidence, much of it anecdotal, but some experimental (e.g. Sloboda 1989), that there are two types of motivation to engage with music. One motivation is what one might call 'intrinsic'. It develops from intense pleasurable experiences with music (of a sensual, aesthetic or emotional kind) which lead to a deep personal commitment to music. The other motivation is what one might call 'extrinsic', and is concerned with achievement. Here, the focus is not so much on the music itself as on achieving certain goals (e.g. approval of parents, identification with role models, winning competitions). Clearly, any one individual will have a mixture of the two types of motivation. There is some evidence, however, that a too early emphasis on achievement can inhibit intrinsic

motivation. In simple terms, children become so concerned about what others may be thinking of their performance that they have little attention left to allow the potential of the music to engage their aesthetic and emotional sensibilities deeply. All music becomes a source of anxiety.

Research on people's experiences with music in school (Sloboda 1989) shows that many remember their earliest encounters with music as ones of anxiety or humiliation which left them with the sense that they were not musically able. Those people who survived such environments were often those who had already experienced music as intensely pleasurable in some other context.

It seems that our society—particularly our system of formal education—is set up to produce a large number of musical 'walking wounded'. What may be needed to remedy that is the rehabilitation of the notion that, in the sphere of music if nowhere else, deep emotional experiences and sensual enjoyment are as important and valid as hard work and technical achievement. One of the reasons why music seems to be seen by many children as an essentially female activity may be the strong male conditioning against emotional display, which is fully present by the age of six or seven in most cultures. One of the unexpected findings of our study was the huge sense of relief and safety felt by some boys when they were able to get to a specialist music school where music was at last seen as a quite normal activity for boys (Howe and Sloboda 1991).

Finally, there is considerable evidence that children are aware of the associations of different types of music with social group identity. Schools often present children with the notion that classical music is superior to the popular and folk idioms with which they more naturally identify. Children can reject classical music as too highbrow and also become intimidated by the apparent requirement to step outside what seems comfortable and familiar before they can produce 'acceptable' music. Most of us are aware of the child who presents as incompetent in a singing class, but can be found the same day in the school bus joining in an informal sing-along session with competence and enthusiasm. There are strong social pressures from peers to conform to the norm and hide a potential liking for music which is 'different' (Finnas 1987). This means that teachers need to introduce music through, rather than in opposition to, the cultural musical norms of their students.

## References

Bharucha, J. J. (1987) Music cognition and perceptual facilitation: a connectionist framework. *Music Perception*, **5**, 1–30.

Bregman, A. S. (1990) *Auditory scene analysis: the perceptual organisation of sound.* MIT Press, Cambridge, M. A.

Chase, W. G. and Simon H. A. (1973) Perception in chess. *Cognitive Psychology,* **4**, 55–81.

Clarke, E. F. (1988) Generative principles in music performance. In: Sloboda J. A. (ed) *Generative processes in music: the psychology of performance, improvisation, and composition.* Oxford University Press, Oxford, pp. 1–26.

Deutsch, D. (1980) The processing of structured and unstructured tonal sequences. *Perception and Psychophysics,* **23**, 215–218.

Deutsch, D. (1982) Grouping mechanisms in music. In: Deutsch, D. (ed) *The psychology of music.* Academic Press, New York, pp. 99–134.

Ericsson, K. A., Krampe, R. Th., and Tesch-Römer, C (1993) The role of deliberate practice in the acquisition of expert performance. *Psychological Review,* **100**, 363–406.

Finnas, L. (1987) Do young people misjudge each other's musical tastes? *Psychology of Music,* **15**, 152–166.

Gabrielsson, A. (1988) Timing in music performance and its relations to music experience. In: Sloboda J. A. (ed) *Generative processes in music: the psychology of performance, improvisation, and composition.* Oxford University Press, Oxford, pp. 27–51.

Gjerdingen, R. O. (1990) Categorization of musical patterns by self-organizing neuronlike networks. *Music Perception,* **7**, 339–369.

Hepper, P. G. (1991) An examination of foetal learning before and after birth. *Irish Journal of Psychology,* **12**, 95–107.

Howe, M. J. A. and Sloboda, J. A. (1991) Problems experienced by young musicians as a result of the failure of other children to value musical accomplishments. *Gifted Education,* **8**, 102–111.

Huron, D. (1991) The avoidance of part-crossing in polyphonic music: perceptual evidence and musical practice. *Music Perception,* **9**, 93–104.

Jackendoff, R (1991) Musical parsing and musical affect. *Music Perception,* **9**, 199–230.

Lerdahl, F (1988) Cognitive constraints on compositional systems. In: Sloboda J. A. (Ed.) *Generative processes in music: the psychology of performance, improvisation, and composition.* Oxford University Press, Oxford, pp. 231–259.

Lerdahl, F. and Jackendoff, R. (1983) *A generative theory of tonal music.* MIT Press, Cambridge, MA.

Meyer, L. B. (1956) *Emotion and meaning in music.* University of Chicago Press, Chicago, IL.

Narmour, E. (1990) *The analysis and cognition of basic melodic structures: the implication– realisation model.* University of Chicago Press, Chicago, IL.

Narmour, E. (1991) The top-down and bottom-up systems of musical implication: building on Meyer's theory of emotional syntax. *Music Perception,* **9**, 1–26.

Oura, Y. and Hatano, G (1988) Memory for melodies among subjects differing in age and experience in music. *Psychology of Music,* **16**, 91–109.

Pinchot-Kastner, M. and Crowder, R. G. (1990) Perception of the major/minor distinction: IV. Emotional connotations in young children. *Music Perception,* **8**, 189–202.

Shaffer, L. H. (1981) Performances of Chopin, Bach, and Bartok: studies in motor programming. *Cognitive Psychology,* **13**, 326–376.

Sloboda, J. A. (1983) The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology,* **35**, 377–396.

Sloboda, J. A. (1985) Expressive skill in two pianists: style and effectiveness in music performance. *Canadian Journal of Psychology,* **39**, 273–293.

Sloboda, J. A. (1989) Music as a language. In: Wilson, F., Roehmann F (eds) *Music and child development.* MMB Music, Saint Louis, MO, pp. 28–43.

Sloboda, J. A. (1991) Music structure and emotional response: some empirical findings. *Psychology of Music,* **19**, 110–120.

Sloboda, J. A. and Howe, M. J. A. (1991) Biographical precursors of musical excellence: an interview study. *Psychology of Music,* **19**, 3–21.

Sloboda, J. A. and Parker, D. H. H. (1985) Immediate recall of melodies. In: Howell, F., Cross, I., West R (eds) *Musical structure and cognition.* Academic Press, London, pp. 143–167.

Sloboda, J. A., Hermelin, B., and O'Connor N. (1985) An exceptional musical memory. *Music Perception,* **3**, 155–170.

Sosniak, L. A. (1985) Learning to be a concert pianist. In: Bloom B. S. (ed) *Developing talent in young people.* Ballantine Books, New York, pp. 143–167.

Zenatti, A. (1969) *Le développement génétique de la perception musicale.* Monographies Français Psychologique No. 17.

CHAPTER 16

# THE ACQUISITION OF MUSICAL PERFORMANCE EXPERTISE: DECONSTRUCTING THE 'TALENT' ACCOUNT OF INDIVIDUAL DIFFERENCES IN MUSICAL EXPRESSIVITY

## 16.1 The curious case of 'musical talent'

Considering how widespread a phenomenon music is within our society and the fact that listening to music is an important part of so many people's daily lives, it is something of an enigma that so few people develop any significant level of music performance ability. Not only do many people never attempt to learn an instrument, but the majority of those who undertake some form of structured learning activity abandon their efforts within a few years.

In countries such as the United States and the United Kingdom, almost every child in school receives classroom music instruction from an early age. Yet the general level of musical achievement in the school-age population is surely well below that of many other skills addressed by the school curriculum.

An apparently almost irresistible popular line of explanation for this general lack of musical accomplishment in the population is the invocation of the presence or absence of 'musical talent'. Sloboda, Davidson, and Howe (1994) proposed the existence of a folk psychology of talent, which postulates substantial innately determined differences between individuals in their capacity for musical accomplishment. According to the folk psychology account, few people become expert musical performers because few people have the necessary talent. This folk psychology is evident in the structures and rhetoric of the music conservatoire (see, e.g., Kingsbury, 1988), and also in the beliefs of people involved in schools. For instance, O'Neill (1994) demonstrated that schoolchildren as young as 8 years old are already more likely to believe that music ability cannot be improved by effort than

sporting skill. A recent survey by Davis (1994) shows that more than 75% of a sample comprised mainly of educational professionals believed that composing, singing, and playing concert instruments required a special gift or natural talent. Other activities, such as playing chess, performing surgery, and writing nonfiction, were seen as requiring talent by less than 40% of the sample, as, curiously, was orchestral conducting!

Although not wishing (or being in a position) to deny that inherited differences between individuals may play some role in determining to what extent musical skills are acquired, one major purpose of this chapter is to marshal evidence and arguments for an alternative view to the prevalent folk psychology. This alternative view holds the capacity for musical accomplishment of one sort or another to be a species-defining characteristic (along with the capacity to learn a language, develop motor skills, etc.). As with these other skills, musical development may be impeded by inherited deficiencies, but such deficiencies only cause major problems in a relatively small minority of cases. Most humans are 'primed' to become musicians. Therefore, for an account of why they fail to develop significant levels of performance skill, we need to look elsewhere than lack of supposed talent.

In reviewing the state of this debate, I summarize some material and go over arguments already published elsewhere. The specific issue that I want to address in greater detail in this chapter is the distinction between technical and expressive aspects of musical performance. Those who are prepared to concede that talent might not be the best explanation of technical development are much more reluctant to concede on the issue of expressivity. Expressive musicianship is widely considered to be intuitive, spontaneous, unteachable to those who do not display it, and a prime manifestation of talent. If we can rescue expressivity from the folk psychology of talent, then the rest of musical expertise comes with it rather easily.

## 16.2 What can we know about inheritance of musical ability?

The usual source of evidence on heritability of psychological characteristics is, of course, the literature on twin studies. The methodological and theoretical traps of this literature are well known, particularly in the context of IQ studies, and it is outside my concerns or expertise to rehearse the arguments here. There are two points to make with respect to music. First, there exists no generally accepted psychometric measure of musical ability that has the validity and reliability of standard measures of intelligence. There are no widely accepted 'MQ' tests. Second, even if all the difficulties with twin studies are set aside, such studies seem to conclude that differences in

musical ability are considerably less dependent on inherited factors than are differences in IQ (Coon and Carey, 1989).

Plomin and Thompson (1993) recently offered the hope of a means of transcending the inherent difficulties of twin studies by a new technique of *allelic association,* which allows direct comparison of genetic material between large numbers of unrelated individuals varying on some known psychological property such as IQ. This technique uses analysis of genetic material found in human blood cells. Hints that genetic correlates of high IQ have been established by this group have, however, yet to be substantiated in print, and the technique is too new to have been properly evaluated, let alone applied to music or other skill domains.

A different source of evidence concerning genetic foundations of musicality comes from direct research on the capacities of human infants pre- and postterm. Research on prenatal capacities (see Hepper, 1991; Lecanuet, 1995) and postnatal musical cognition (see Papousek, 1995; Trehub, 1990) suggests that the vast majority of human infants have rather similar, and surprisingly sophisticated, means of handling musical stimuli. These universal capacities lead to an astonishing flowering (from the sixth month on) of spontaneous premusical generativity. All but the most severely learning-disabled children display a rich vocabulary of pitch- and rhythm-modulated utterances. As in the case of language, these utterances can be shown to pass through a number of stages that can well be described in terms of different systems of underlying grammatical rules that succeed one another in an orderly way (Imberty, 1995).

Although individual differences in infant functioning are remarkably difficult to research, for methodological and practical reasons (see Colombo, 1993), it would seem that developmental research has shown that the average human infant displays musical capacities as early, if not earlier, than linguistic capacities, and in a far more developed and overt form than any capacities that might be supposed to underlie other skills such as mathematical, artistic, or athletic skills. Nothing in this literature prepares us for the dismal musical outcomes observed in our culture in late childhood and adulthood.

It would be impossible for me to address the issues of musical expertise without recurrent reference to the most recent large-scale study in which I have been involved, with colleagues Michael Howe, Jane Davidson, and Derek Moore (Davidson, Sloboda, and Howe, 1994; Howe, Davidson, Moore, and Sloboda, 1995; Sloboda and Davidson, 1995; Sloboda, Davidson, Howe, and Moore, 1996; Sloboda, Davidson, Moore, and Howe, 1994). This study, named the Leverhulme Project in honor of its funding source, the Leverhulme Trust, obtained a wide variety of information,

mostly biographical, on 257 young people in England. They were selected to represent a range of current performance achievement on musical instruments, from the outstandingly able through to the modal 'tried an instrument but gave it up'. Individuals at different levels of accomplishment were roughly matched for age, gender, instrument, and socioeconomic class. We established five major ability groups, whose rank ordering was validated by differences in objective performance measures (musical examinations set by a national UK examining body). Group 1 comprised 119 students of a specialist music school, where entry is by competitive audition. This group represents the highest ability level. Graduates of the school regularly win national performance competitions, and typically go on to conservatory study and careers as professional performers. Groups 2, 3, and 4 comprised current players of musical instruments, classified in terms of their levels of aspiration for excellence in musical performance. The members of Group 2 had failed the entry audition to the specialist school. The members of Group 3 had inquired about but had failed to apply for entry to the specialist school. The members of Group 4 were current 'players for pleasure' in an ordinary state school, undistinguished in its general level of musical education. Group 5 comprised pupils of the same school as Group 4, who had begun to learn an instrument, but had given up study after a period of 6 months or more. Each participant was interviewed, and answers cross-checked for reliability in individual interviews with a parent of each participant.

It is very important for later steps in the argument to note that all these young people were learning musical instruments in what might be described as the 'classical conservatoire culture', where emphasis is on reproduction of musical artworks within the formal classical tradition, represented by composers such as Mozart, Beethoven, Rachmaninov, and Stravinsky. The characteristics of this tradition are (a) concern with accurate and faithful reproduction of a printed score, rather than with improvisation or composition; (b) the existence of a central repertoire of extreme technical difficulty; (c) definitions of mastery in terms of ability to perform items from a rather small common core set of compositions within a culture; and (d) explicit or implicit competitive events in which performers are compared with one another by expert judges on their ability to perform identical or closely similar pieces, such judgments forming an important element in decisions about progression and reward within the culture. Other musical cultures exist, such as jazz, pop, or folk cultures, which have very different characteristics. It should not be supposed that findings from studies of conservatoire culture musicians can be straightforwardly applied to other forms of musical activity.

It has often been supposed that individual differences in musical accomplishment are prefigured by individual differences in early signs of musicality. Building on accounts of one or two legendary 'geniuses' such as Mozart, and other accounts of child prodigies (e.g. Revesz, 1925/1970), commentators have often assumed that individuals displaying unusually high ability in adolescence or later must have always been unusual. We attempted to test this claim in the Leverhulme Project by asking detailed and specific questions of the parents of all participants concerning the age of occurrence of particular early manifestations of musical ability, such as the first age at which the child could reproduce by singing a recognizable tune from the culture.

The results were clear (Howe *et al.*, 1995). Similarities between the five groups far outweighed the differences, which were nonexistent on many measures, and when present did not always favor the highest ability group. The difference most favorable to the 'talent' account was a 6-month average advantage for Group 1 in singing a recognizable tune (18 months of age, as opposed to 24 months for the other groups). However, this group showed earlier parent-initiated musical activity, too, and so the result may be a consequence of differential parental behavior. We have considerable evidence from other parts of our study that the parents of the highest achieving group showed unique behavioral characteristics. Although there are obvious interpretational difficulties with retrospective studies, this study lends little credence to the notion of large individual differences in early musicality that prefigure later achievement. The results are entirely consistent with the data obtained by direct study of very young children. Most children show a variety of overt musical behaviors at an early age, and few stray far from established norms.

## 16.3 The role of practice

Ericsson, Krampe, and Tesch-Römer (1993) provided compelling evidence for a conclusion of some generality with respect to acquisition of expertise. Their conclusion is that level of expertise is a direct function of the amount of effortful formal practice of that skill undertaken by an individual. Their own work on student and professional instrumental players has shown that the highest achieving individuals consistently undertook around twice as much daily practice as moderate achievers, over long periods of childhood, adolescence, and early adulthood. Similar findings have been obtained in other domains such as chess and sport.

The Ericsson position offers a strong challenge to the folk psychology of talent, which, in the case of music and many other skills, assumes that

high-achieving individuals acquire their outstanding skill with the same or even less practice than others.

The Leverhulme Project has provided data that fully confirm and strengthen the practice hypothesis (Sloboda *et al.*, 1994, 1996). All participants were asked to provide estimates of average daily formal practice for each year of life since beginning to learn an instrument, with formal practice defined as activities directly related to tasks set by the instrumental teacher in the previous lesson. A subset of the sample also kept a detailed record of practice activities over a 42-week period.

The two main subcomponents of such practice were named as scales and pieces. *Scales* is a shorthand designation for work on scales, arpeggios, and other technical exercises without musical interest or merit in their own right, but designed to provide exhaustive and repetitive opportunities for mastery of technical aspects of playing. *Pieces* designates work on items from the musical repertoire toward the end goal of polished performance.

The relationship between practice and achievement level was strongly confirmed, and extended across the entire sample. Members of Group 1 were, even by the age of 12, practicing an astonishing 800% more than the members of Group 5 (2 hours a day, as compared to 15 minutes a day). The intermediate groups fell between these extremes in exact order of achievement. The practice–achievement relationship does not, therefore, break down at low levels of achievement. We found no individuals in our sample who practiced as much as 2 hours per day yet failed to achieve high levels of skill! Such cases would surely be predicted in considerable numbers were the folk psychology view correct.

A chance feature of the UK instrumental education system allowed us to be even more definitive about the practice–achievement relationship than Ericsson *et al.* (1993). The vast majority of children learning instruments in the UK are entered by their teachers into a national system of graded instrumental examinations, which are set and examined by a nationally constituted panel of trained assessors, whose purpose is to provide equivalent measurements across the country. There is a preliminary grade, and eight main grades, which, as a very rough guide, might each represent a year's additional work on an instrument. The pedagogic practices of many teachers are totally determined by the requirements of these grade examination syllabi.

Every child in our sample had entered for these examinations, and we were able to ascertain at what age each grade level had been reached. This allowed, for each group, the calculation of the average number of hours of formal practice required to achieve each grade. This amount proved to be not significantly different between groups. The relationship between grade

level and practice for the whole sample is shown in Figure 16.1. This indicates, for instance, that, regardless of ultimate achievement level, it takes an average of 1200 hours of formal practice in total to achieve Grade 5 standard (or 300 hours to get there from Grade 4). The reason that Group 1 members achieve grade exam successes earlier than other groups seems entirely due to accumulation of the requisite amount of practice more quickly. There is absolutely no evidence of a 'fast track' for high achievers. Indeed, there is a nonsignificant trend for high achievers to practice more than low achievers to reach a particular grade.

Of course, as one might expect, there is a fairly high degree of within-group variance in amounts of practice undertaken to achieve a given grade.

One can suppose differences in instrument, practice strategy, teaching efficiency, concentration, and other factors to account for this (see Lehmann, 1997 for an elaboration of this argument). There might even be some contribution from inherited differences. The important thing for this discussion is that this variability is present at all levels of achievement, and therefore is not constitutive of achievement, as the folk psychology account would demand.

Several studies of expertise have been criticized for not including a 'dropout' group. Group 5 of this study comprises such a dropout group. The fact that this group shows exactly the same relationship between practice

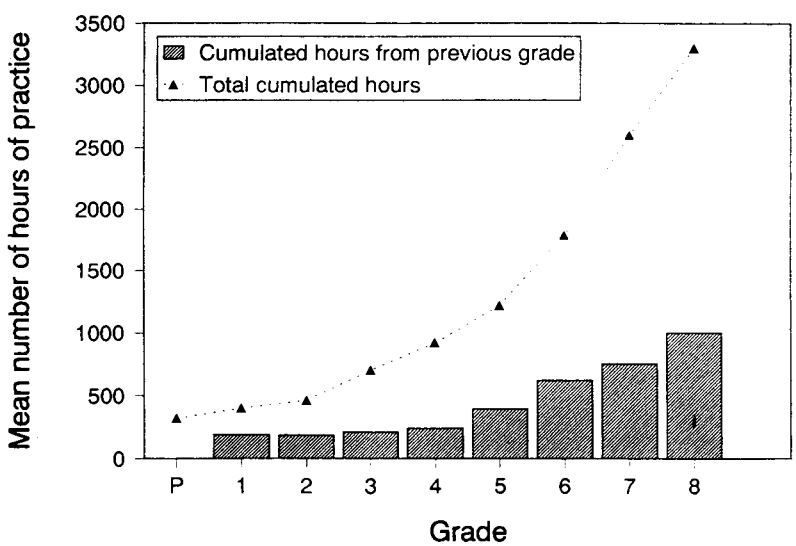



**Figure 16.1** Relationship of cumulated hours of practice to grade achievement for young instrumentalists. Data are means for 257 instrumentalists in five ability groups. There are no significant differences between groups at any grade level.

and achievement as any other group suggests that putative differences between persisters and dropouts (in terms of 'talent', self-beliefs, and motivation) influence only the duration of practice activities, and not the effectiveness of those activities once undertaken, in improving skill.

The Leverhulme study also provided some data that helped explain why the groups differed so much in the amount of daily practice they were capable of (Davidson, Howe, Moore, and Sloboda, 1996; Davidson *et al.*, 1994). At the outset of instrumental study, groups were broadly equivalent on a crude measure of practice motivation. All reported having periods when motivation to practice was low, and when they probably would not have practiced at all without parental intervention. They displayed little of the 'rage to master' reported in the case of some precocious young artists. Two things marked the higher achieving children. First, the nature and level of early parental involvement was significantly different. Parents of Group 1 students were more likely to attend instrumental lessons with their children, obtain detailed feedback and instructions from teachers, and actively supervise daily practice on a moment-to-moment basis, often at some considerable cost to their own schedule. For instance, one family with three high-achieving siblings adopted a daily routine in which the father would supervise the individual practice of each child in 30-minute intervals from 6.00 a.m. until 7.30 a.m. prior to the family dispersing to school and work. Parents of low-achieving children were less likely to have meaningful contact with the teacher, and were likely to confine their domestic interventions to telling children to 'go and do your practice', without any direct involvement in it. In sum, therefore, it seems that abnormally high levels of early practice are sustained by abnormal levels of social and cognitive support, mainly from parents.

The second feature of the higher achieving children was a gradual age-related move toward self-motivation for practice as adolescence progressed. These individuals appeared to be finding their own motivations for practice, although the study did not systematically probe what these were. Arguably, those who make faster progress get to work on more interesting and complex music, obtain greater mastery, perceive themselves as successful, and identify with teachers or other high-level performers as role models. All of these factors have been experimentally demonstrated in other domains to have strong effects on motivation (e.g. Amabile, 1983; Dweck, 1986), and we found examples of all of these factors in the informal comments of the interviewees.

The diary study showed, however, that even among the highest achieving groups, there were large fluctuations in daily practice duration as a function of external circumstances. These individuals practiced most before

important concerts, and least during school vacation. Despite these fluctuations, however, Group 1 still demonstrated the most week-to-week stability in practice duration, and confirmed Ericsson *et al.*'s (1993) finding that high-achieving individuals tend to do more of their practice in the morning. In our sample this effect was particularly strong for scales practice. High-achieving young people tend to concentrate the more grueling repetitive technical practice at the time of day when they are physically and cognitively at their peak.

## 16.4 Technical versus expressive aspects of music performance

Advanced music performance requires high levels of technical skill. This is not only because much core repertoire requires extremely rapid movements, but also because very small differences in the positioning or timing of movements can make huge differences in the perceptual qualities of the sound. For instance, accurate tuning in violin playing can be disrupted by positioning deviations of fingers on the string of less than 1 millimeter. The technical difficulty of classical instrumental performance is recognized by the long-standing tradition of manuals, tutors, and exercises designed solely to address technical problems through highly systematic and repetitive practice. The sequences that are used for such purposes often have little or no aesthetic interest. From the very outset of instrumental learning, most teachers encourage their pupils to devote the first part of any practice period to technical exercises, before going on to work at repertoire. The teaching of technique is commonly based on some set of theoretical assumptions about optimal posture, muscle tension, transitions between different positions, and so on. These assumptions have general components that resemble those found in the pedagogy of other skills. For instance, many musical theories make general assumptions about the desirability of economies in both effort and movement, which resemble assumptions behind sporting or athletic technique. Most of the theories, however, devote the bulk of their efforts toward the peculiar problems of the interface between the human body and the particular instrument.

Because the shape, size, and sound-producing mechanisms of musical instruments are so different from one another, there is very little in common between, say, violin technique and piano technique. Perhaps even more important, musical instruments require, by and large, bodily movements of a type not common in other domains of human activity. These almost always involve small but precisely timed and located finger movements that show high degrees of independence from one another. A very fundamental issue

in nearly all instrumental performance is the move toward complete independence of fingers and hands one from another. The nonmusical skill with the most similar set of requirements is typing, but even there, the level of specificity of finger movement speed and timing is of a much lower order. By and large it does not matter when a key is depressed, so long as the sequence of keys is in the correct order. Furthermore, typing rarely involves the simultaneous depression of two or more fingers. Skills such as piano playing require this routinely.

For all these reasons, we can predict that the acquisition of high levels of technical competence on an instrument will take a long time and will require much practice specifically tailored to the requirements of each instrument. People have, in effect, to learn a whole new set of motor programs for each different instrument. There are few previously learned skills that the learner can build on. We may suppose that if there is any transfer of training between different instruments it reflects either their physical similarity (e.g. piano and harpsichord, violin and viola) or other nonmotor skills (e.g. the ability to read music) that both instruments require.

It has long been recognized by musicians, and more recently confirmed by psychological research (Clarke, 1988; Gabrielsson, 1988; Shaffer, 1981; Sloboda, 1983) that mere technical prowess does not make an effective musical performer. Playing the notes as written with speed and fluency is the starting point for artistic excellence, not the finish.

High-level performers add value to the score by a range of expressive devices that affect the microstructure of timing, loudness, timbre, and other elements, resulting in notes that are equal in the notation being performed unequally in some respect. We know a number of things for certain about expressive performance:

1 The deviations from exactitude are intentional. That is, they are not unintentional by-products of technical deficiencies, or products of random noise in the motor system. We know this because a performer can (a) reproduce the same deviations on occasions that may be widely separated in time (Shaffer, 1984), and (b) produce a different expressive profile in response to changes in task demand (Palmer, 1989; Sloboda, 1983).

2 The deviations are, in part, applied without conscious awareness. Performers asked to produce expressionless 'deadpan' performances are normally incapable of doing this. Reduced features of expressive performances are still present (e.g. Gabrielsson, 1974, 1988). Even when

performers know they are being deliberately expressive, they do not have full conscious awareness of every detail of their performance.

3 The deviations are systematic. That is to say, a given performer applies the same type of deviation at analogous points in a musical structure, and performers within a culture tend to use the same types of expressive devices.

4 The deviations are constrained by the structure of the music. This means that expression takes account of structure, and in some situations is intended to reinforce the structure by making it more manifest or salient to a listener. Musicians can have rational debates about the relative merits of different ways of playing the same piece, because it is not the case that 'anything goes' in expression.

5 The deviations are unique in some respects to a given performer. At first glance this seems to contradict the third point already given. However, even if two performers use exactly the same set of expressive devices, they can still sound quite different from one another because of their differences in frequency of use of the various devices, the distribution of magnitudes of deviations, and other 'free' parameters. Hence, each performer has a distinct performing style that is recognizable to connoisseurs (Sloboda, 1985).

6 The deviations are detectable by listeners, and assist their processes of forming representations of the music. For instance, where the music is metrically ambiguous, expert performance disambiguates the meter (Sloboda, 1983).

7 Expressive skill develops over long timespans. Commentators on internationally known performers often trace significant changes in expressive character of recorded performances over a lifespan. It is generally held that 'mature' interpretations are unlikely to be achieved within the first 10 years of performing an instrument. Where the issue has been studied experimentally, it has been found that performers with several decades of performing experience show distinct superiority in the consistency and effectiveness of expression, as compared to performers with only one decade of performing experience (Sloboda, 1983).

Expressive performance, unlike technical performance, is not systematically taught or acquired through the use of manuals or sets of exercises. We can observe two processes at work in the acquisition of expression. First is the use of expressive models. A teacher will play something and

ask the student to imitate the expression. In later stages of development, performers avidly seek out the performances of other musicians for purposes of analysis and comparison. The second process is the encouragement of the development of an individual's own expressive style. An individual is invited to generate alternative expressive solutions and exercise some independent judgment in deciding among them.

## 16.5 Deconstructing the talent account of expression

When conversing with music professionals, particularly those who teach, a very common experience recounted by these professionals is the existence of young musicians who seem to be 'primed' to make unusually rapid progress in the acquisition of expressive performance when compared with the norm. This unusual rapidity can be observed in both imitation and generativity. So, on the one hand, there appear to be individuals who are capable of imitating expressive models with unusual accuracy or rapidity; and on the other, there appear to be individuals who spontaneously generate appropriate expressive variations in the absence of an immediate model. To a great extent, it is such observations that support the notion of a small group of talented individuals, separated from the norm.

### 16.5.1 Imitative expression

For the purposes of the argument, assume that these observations by teachers are reliable, accurate, and free from the kinds of biases that can so easily infect judgments of this sort. Let us first of all take the case of imitation. There is, of course, a limiting condition that must pertain if the listener is to detect an intended imitation of expression; that is, that the performance as a whole must be technically more stable than the parameters of the expressive device being detected. To take a concrete example, let us suppose that the device to be imitated is an expressive lengthening of 25% over and above the notated duration of a given note in a phrase. In this case, in order for the lengthening to be detected the average variability in note lengths over the passage as a whole must be no greater than, say, 10% (psychophysical considerations would determine the exact ratio required). If technical stability is not achieved, then we have the possibility that an individual is able to form a representation of a heard expressive device, generate a performance intention in which this device is present, and thus intended for performance, but fail to translate the intention into an adequate motor program. This has the consequence that some differences between individuals in supposed observed expressive ability are actually

technical differences. Technical difficulties can be caused by many things, one of them, of course, being lack of practice. However, there are other factors at play in any individual instance. For instance, an overanxious child might well suffer technical disruption more easily than a confident child. A child who knows the piece being played, or knows pieces like it, might be able to form a representation of the notes it contains more rapidly and efficiently. Many of these factors will lead to fairly stable and reliable differences between individuals in their ability to 'deliver' expressive imitation, at least in the short term, but they have nothing to do with differences in expressive ability, *per se*. It may not be possible for an individual informal observer to distinguish these cases of production deficiency from true differences in expressivity. Only controlled formal quasi-experimental observation may have the necessary discriminatory power.

Let us assume, however, that there is a residue of cases that do not fall foul of these problems. These cases would be individuals of equal technical capacity, in general, and in respect to the piece presented in this situation, who nevertheless show significant differences in capacity to imitate the expressive deviations being demonstrated. What account might we give of these differences that did not rely on talent?

My proposal is based on how individuals might solve the problem of storing the large amount of analog information that appears to be involved in remembering expressive details. There is a great deal of evidence that musical memory, like any other kind of memory, requires categorical structural representation that reduces the information load to manageable proportions. Most tonal music solves the problem by allowing that all musical inputs are reducible to a canonical form in which, for instance, all pitches are subsumed into one of a small number of categories (5–12 per octave according to specific genre), and all durations are likewise subsumed (usually equal subdivisions of a primary beat, 2, 4, 8, 3, and 6 being the most common). To have represented a piece of music tonally means to have extracted these categories from a much richer surface. Expressive deviations are contained in that part of the input that the categorization process must ignore, so representing it means recovering some of it before it is lost. This recovery cannot be a simple list of all the analog information present, because humans do not have the capacity to store such information, which would resemble the output from a performance to a MIDI file, containing the precise duration, loudness, and timbral qualities of every successive note.

We need to translate the expressive information into something more abstract before it can be stored. This abstract representation has two requirements: (a) it must link otherwise apparently arbitrary changes in timing, loudness, and other characteristics of successive notes one to

another through some function or formula that can then be run in reverse during performance to regenerate something approximating the original pattern of deviations; and (b) it must mark the starting and end positions of a particular function or formula in the canonical representation of pitch time categories within the tonal representation of the piece. To put this in more everyday musical language, the expressive device must have a shape that has a definite location in the musical structure. We store the shape and its location rather than the analog data. Change the shape or change the location of the shape and you have a different expressive outcome.

My contention is that what bootstraps the process of representing expressive devices in music is the existence of extra-musical functions and formulas that act as ready-made templates onto which musical expression can be mapped. The individual who presents as 'talented' with respect to expressive imitation is the individual who has (a) acquired an appropriate repertoire of extra-musical templates, and (b) has made the appropriate connections between domains that allows them to be applied 'by analogy' to music.

These templates arise from a number of domains, the most plausible being those of bodily and physical motion, gesture, speech and vocal intonation, and expressions of emotion. Such linkage has been suggested by a number of studies (e.g. Clynes, 1983, 1986; Repp, 1992, 1993; Todd, 1992). There are a number of factors that can account for individual differences in the facility of this linkage:

1 Differences in nonmusical expressivity. Most theories of emotional expression (see Frijda, 1986, for instance) seem to accept that the fundamental forms of such expression are universal, innate, and species defining. What determines individual differences are cultural and social factors, concerned with repression of spontaneous expression and replacement with stylized variants. Some of these factors may be rather specific (e.g. a family environment in which emotional expression is discouraged). Others may be more general. For instance, Sloboda (1991a) provided evidence for the assertion that many children experience performance-related anxiety in music education settings, which overrides any possibility of attending to and representing the expressive characteristics of the music itself. O'Neill and Boulton (1996) and others (Abeles and Porter, 1978; Delzell and Lepplar, 1992) have demonstrated how gender-role stereotypes play a major role in determining what kind of musical engagement is thought to be appropriate by boys and girls. These stereotypes will impact on the kind of expressive or emotional awareness that children will be prepared to bring to music.

2 Differences in the ability to identify and retain the musical location of features requiring expressive treatment. Expressive gestures in music typically have a focus, even if they are spread over a number of notes. So, for instance, a particular note may be expressively marked by a increase in volume prior to it, and a decrease in volume after it. How is someone to remember where the focus of an expressive device is? In the light of the characteristics of expert expressive performance described earlier, it is probable that events are marked for expressive treatment in virtue of their structural characteristics. Therefore, a particular note is the focus for expressive treatment because it is harmonically, rhythmically, and thematically important. The ability to locate an expressive device will therefore be affected by the extent to which those structures are encoded in the person's representation. This will be affected by the degree of experience with that particular genre of music.

3 Differences in the degree of awareness of, and elaboration of, the cross-domain analogy. It is relatively well accepted that analogies need to be noticed before they can be incorporated in cognitive strategies, and that the initial noticing of the analogy may happen quite suddenly. For instance, in Chase and Ericsson's (1981) seminal study on Falloon's digit span, they report that on Day 4 of trials he had reached an asymptote of performance, whereas on Day 5, after discovery of the 'running times' analogy, he was able to dramatically increase his span. We may speculate that what caused the analogy to be noticed was some number sequence that triggered a particularly strong association to a running time that was highly salient for this subject. For another runner, it might have been a quite different number that triggered the analogy, and this may have happened sooner, later, or never. In the case of music, we can suppose that the expressive analogy is bootstrapped by the experience of an expressive gesture that evokes a strong association to a salient nonmusical gesture for that individual.

4 Differences in the degree of monitoring of one's own performance for intended expressive outcomes. There is quite clear evidence available that so-called tone deafness is nothing of the sort. Individuals identified as tone-deaf are unable to match a heard pitch by singing. This is not, however, due to any deficit in pitch perception or memory, which is just as acute as the average correct singer. Rather, it is due to an inability to monitor one's own vocal output and make appropriate adjustments. By providing enhanced feedback, for instance, computer-generated visual representation of one's own sung pitch as compared to a standard, it is possible to teach individuals to adjust their vocal pitch, and, eventually,

to do so without the visual feedback (Welch, Howard, and Rush, 1989). This is essentially the same technique that is used in biofeedback to teach people to control their own heart rate or skin conductance. In relation to expressive performance, it is commonplace for teachers to observe that their pupils do not monitor the expressive outcomes of their own performances. Rather, they seem to be monitoring their expressive intentions, and take the intention for the deed. Such a situation occurs when, for instance, a teacher asks a pupil to copy a dynamic shape (getting louder then getting softer), the pupil does something, the teacher says she cannot hear any difference, but the pupil believes that he did play louder and then softer. Such scenarios can easily lead to incorrect attributions ('the child is inexpressive'), which may not encourage the teacher to seek ways of helping the pupil monitor their own performance more appropriately.

5 Differences in degree of emotional response to the expressive device. There is considerable suggestive evidence (see Sloboda, 1991a) that the ability of individuals to experience music-generated emotional response is a function of the degree to which their emotional response is not captured by factors extrinsic to the music. Thus, where the context of music is a source of threat or anxiety, the emotional response to the situation is determined by sell-related emotions (self-esteem, self-preservation, group conformity) in a way that seems to block the kind of relaxed attentiveness that is a precondition for strong aesthetically based responses to music. Such emotional blocks act as overarching factors that reduce the likelihood of noticing and internalizing the analogy between musical and nonmusical expression. In some individuals, the blocks seem to have become chronic, and independent of actual external context. Music itself triggers anxiety, by association (e.g. 'Whenever I hear music I remember the humiliation of being told by my teacher that I was singing out of tune').

There are no doubt other factors that could be adduced to account for individual differences. The demonstration of their existence is a task for future empirical research, but each of them is a specific causal hypothesis in which considerations of supposed talent or its absence are really of not much help in moving toward a better understanding of individual differences.

### 16.5.2 Generative expression

Almost all that one can say about imitation can also be said about individual differences in expressive generation, but there are some additional

sources of variance. First, we need to be clear that there are two different types of generativity One (Type 1) is closely related to imitation in that the sequence of notes is prescribed—and an expressive contour must be fitted to this pre-existent sequence. The other (Type 2) is rather different, and relates to improvisatory mode, where the notes are chosen by the performer along with their expression. This latter kind of expressivity almost never occurs in the conservatoire culture.

Type 1 generativity requires the recognition that a certain musical sequence admits of expressive treatment in the absence of the immediate presentation of a specific expressive model for that sequence. A clear example of generative expression is in the case where an individual sight-reads from an unedited score, or where an individual imitates an inexpressive model, or incorporates different expression from that offered. It is highly likely that another kind of analogical thinking is needed here. The logic would run something like this: The present sequence is similar in musical structure or form to a sequence experienced in another context; in that other context such-and-such an expressive tactic seemed to work well; therefore let us try it here. Another way of expressing this logic within musical discourse is to say that there are expressive conventions within a musical culture, and that these conventions may be learned more or less thoroughly. Nothing in this requires that a performer be able to consciously articulate the known convention. It may be entirely embodied within a non-verbal performance 'habit' or 'production rule', to use a more theoretically fashionable term.

Individual differences in expressive generativity may, in this case, be linked to individual differences in the opportunities that an individual has had to learn these conventions. This will be partly a matter of exposure to appropriately expressive performances of others, but also of factors such as the five adduced earlier with respect to imitative performance.

Type 2 generativity has a rather different etiology. The developmental literature shows that expressive improvisation is the norm in infants and young children. In the vocal play of young children, expressive gesture is almost the guiding force of vocalizations. Infants sigh, chortle, coo, and make other directly recognizable emotional signals, then they embellish these with pitch and temporal patterning. They also sing while engaged in physical activity, and so naturally coordinate their vocalizations with their bodily movements. What then seems to happen in midchildhood is that the formal characteristics of music separate out from the expressive. Children become very concerned with imitating recognizable tunes of their culture, and at the same time, are learning the control of emotional expression as part of socialization. By the early school years, at least in our own

culture, improvisation (except where cherished and nurtured by perceptive adults) has all but dried up. Group singing and group instrumental playing, at the early stages of education, encourage this process of suppression of individual expression in favor of subordination to the group. Regrafting of expression within the elaborated formal musical structures of midchildhood represents a formidable challenge for many children. It requires going directly against the mainstream of cultural separation and repression. I believe that children require very clear enabling conditions for that to happen. We learn from music therapists that emotionally repressed individuals can discover strong and intuitive emotional expressivity within themselves when offered a music performance medium in which they come realize there are no preordained standards of correctness. Typical music therapy, at least within the UK, will not offer models that are part of the standard conservatoire culture, but might, for instance, offer an individual a large array of percussion instruments, and encourage significant bodily and sonic exploration. In such a context, for instance, anger maybe gesturally expressed in a very direct way through the force with which an instrument is struck. Children may need an opportunity of a similar sort to be expressive. It is extremely significant to me that so many teachers and parents tell stories of children who appear 'unmusical' (or at least unremarkable) in standard music lessons, but then are found engaging in highly expressive and creative forms of musical behavior outside the formal classroom culture, be it involvement in playground games involving singing, or in the formation of peer-organized groups, using popular forms such as pop, jazz, or folk. This seems to me to suggest key contexts in which the individual senses 'permission' to be expressive. For many young people, the conservatoire culture does not provide such permission, and we may need to discover more precisely how those few that do flourish expressively within that culture are protected from the normal inhibitory processes. They are the ones who can, to some extent at any rate, feed on the wellsprings of their own primal musical expressive generativity.

## 16.6 Concluding remarks

I realize that many of the comments I have made in this chapter are somewhat speculative. I believe it is appropriate in a conference of this sort to go out on a limb, and push a particular line of argument for all it is worth. I do not wish any of the arguments I put forward to be taken as conclusive, but rather suggestive. In many cases the evidence for particular kinds of processes needs to be gathered. In other cases, weaknesses in the argument need to be exposed. I hope that other participants will be merciless in this

respect. We need to find the holes in each other's positions in the pursuit of scientific clarity.

What I have hoped to do is sketch out the framework of a position in which talent is not so much disproved as dissolved into a whole set of complex interacting factors and causes, each of which has its own logic and determining conditions. It is the intricacy and complexity of musical growth and development that makes for fascination and scientific discovery, not some ultimately arid dispute between oversimplistic extremes.

To summarize the steps in my argument:

1 Music seems to be biologically constitutive of early human functioning.

2 Music education in Western cultures produces a dismal yield of achievement.

3 Heritability estimates, where available, are low.

4 Technical expertise within the conservatoire culture requires practice levels far in excess of cultural norms because of its unique properties with respect to particular instruments and the specific requirements of that culture to master a technically demanding canon.

5 Such practice is sustained by external motivators in the early years, but by increasing development of internal sources through development.

6 Expressive expertise has rationality and develops through practice.

7 Significant individual differences in expression are noticed by teachers at early stages in instrumental learning.

8 Unlike technique, expression has characteristics that are similar to extra-musical activities (bodily and emotional gestures).

9 This creates opportunities for learning by analogy.

10 A whole range of plausible factors can influence the ease of uptake of this analogy.

11 The articulation and investigation of these factors constitutes a progressive agenda for the scientific study of musical skill.

## References

Abeles, H. F. and Porter, S. Y. (1978) Sex-stereotyping of musical instruments. *Journal of Research in Music Education*, **26**, 65–75.

Amabile, T. M. (1983) *The social psychology of creativity*. New York: Springer Verlag.

Chase, W. G. and Ericsson, K. A. (1981) Skilled memory. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 141–189). Hillsdale, NJ: Lawrence Erlbaum Associates.

Clarke, E. F. (1988) Generative principles in music performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 1–26). London: Oxford University Press.

Clynes, M. (1983) Expressive microstructure in music, linked to living qualities. In J. Sundberg (Ed.), *Studies of music performance* (pp. 76–181). Stockholm: Royal Swedish Academy of Music.

Clynes, M. (1986) Music beyond the score. *Communication and Cognition*, **19**, 169–194.

Colombo, J. (1993) *Infant cognition: Predicting later intellectual functioning.* Newbury Park, CA: Sage.

Coon, H. and Carey, G. (1989) Genetic and environmental determinants of musical ability in twins. *Behavior Genetics*, **19**, 183–193.

Davidson, J. W., Sloboda, J. A., and Howe, M. J. A. (1994) The role of family and teachers in the success and failure of music learners. *Proceedings of the Third International Conf erence for Music Perception and Cognition* (pp. 359–360). University of Liège, Belgium: URPM.

Davidson, J. W., Howe, M. J. A., Moore, D. G., and Sloboda, J. A. (1996) The role of parental influences in the development of musical performance. *British Journal of Developmental Psychology*, **14**, 399–412.

Davis, M. (1994) Folk music psychology. *The Psychologist*, **7**(12), 537.

Deizell, J. K. and Lepplar, D. A. (1992) Gender association of musical instruments and preferences of fourth-grade students for selected instruments. *Journal of Research in Music Education*, **40**, 93–103.

Dweck, C. S. (1986) Motivational processes affecting learning. *American Psychologist*, **41**(10), 1040–1048.

Ericsson, K. A., Krampe, R., and Tesch-Romer, C. (1993) The rule of deliberate practice in the acquisition of expert performance. *Psychological Review*, **100**, 363–406.

Frijda, N. H. (1986) *The emotions.* Cambridge, UK: Cambridge University Press.

Gabrielsson, A. (1974) Performance of rhythm patterns. *Scandinavian Journal of Psychology*, **15**, 63–72.

Gabrielsson, A. (1988) Tuning in music performance and its relation to musical experience. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 27–51). London: Oxford University Press.

Hepper, P. G. (1991) An examination of fetal learning before and after birth. *Irish journal of Psychology*, **12**, 95–107.

Howe, M. J. A., Davidson, J. W., Moore, D. G., and Sloboda, J. A. (1995) Are there early childhood signs of musical ability? *Psychology of Music*, **23**(2), 162–176.

Imberty, M. (1995) Linguistic and musical development in pre-school and school age children. In I. Deliège andJ. A. Sloboda (Eds.), *Musical beginnings: The origins and development of musical competence* (pp. 191–213). London: Oxford University Press.

Kingsbury, H. (1988) *Music, talent, and performance: A conservatory cultural system.* Philadelphia: Temple University Press.

Lecanuet, J. P. (1995) Prenatal auditory experience. In I. Deliège and J. A. Sloboda (Eds.), *Musical beginnings: The origins and development of musical competence* (pp. 5–54). London: Oxford University Press.

Lehmann, A. C. (1997) Acquisition of expertise in music: Efficiency of deliberate practice as a mediating variable in accounting for sub-expert performance. In I. Deliège and J. A. Sloboda (Eds.), *Perception and cognition of music.* Mahwah, NJ; Lawrence Erlbaum Associates.

O'Neill, S. (1994) Factors influencing children's motivation and achievement during the first year of instrumental music tuition. Paper presented at the Third International Conference on Music Perception and Cognition, Univerity of Liège, Belgium, July 1994.

O'Neill, S. and Boulton, M. (1996) Boys' and girls' preferences for musical instruments: A function of gender? *Psychology of Music,* **24.2**, 171–183.

Palmer, C. (1989) Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance,* **15**, 331–346.

Papousek, H. (1995) Musicality and infancy research. In I. Deliège and J. A. Sloboda (Eds.), *Musical beginnings: The origins and development of musical competence* (pp. 37–55). London: Oxford University Press.

Plomin, R. and Thompson, L. A. (1993) Genetics and high cognitive ability in G. Bock and K. Ackrill, (Eds.), *The origins of high ability: Proceedings of Ciba Symposium 178* (pp. 67–84). London: Wiley.

Repp, B. H. (1992) A constraint on the expressive timing of a melodic gesture: Evidence from performance and aesthetic judgment. *Music Perception,* **10**, 221–242.

Repp, B. H. (1993) Music as motion: A synopsis of Alexander Truslit's (1938) *Gestaltung und Bewegung in der Musik. Psychology of Music,* **21**, 48–72.

Revesz, C. (1970) *The psychology of a musical prodigy.* Freeport, NY: Books for Libraries Press. (Original work published 1925).

Shaffer, L. H. (1981) Performance of Chopin, Bach, and Bartok: Studies in motor programming. *Cognitive Psychology,* **13**, 326–376.

Shaffer, L. H. (1984) Timing in solo and duet piano performances. *Quarterly journal of Experimental Psychology,* **36A**, 577–595.

Sloboda, J. A. (1983) The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology,* **35A**, 377–396.

Sloboda, J. A. (1985) Expressive skill in two pianists: Metrical communication in real and simulated performances. *Canadian Journal of Psychology,* **39**(2), 273–293.

Sloboda, J. A. (1991a) Music as a language. In F. Wilson and F. Roehmann (Eds.), *Music and child development* (pp. 28–43). St. Louis, MO: MMB Music.

Sloboda, J. A. (1991b) Musical expertise. In K. A. Ericsson and J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 153–171). New York: Cambridge University Press.

Sloboda, J. A. and Davidson, J. W. (1995) The young performing musician. In I. Deliège and J. A. Sloboda (Eds.), *The origins and development of musical competence* (pp. 171–190). London: Oxford University Press.

Sloboda, J., Davidson, J., and Howe, M. J. A. (1994) Is everyone musical? *The Psychologist*, **7**(8), 349–354.

Sloboda, J. A., Davidson, J. W., Howe, M. J. A., and Moore, D. C. (1996) The role of practice in the development of performing musicians. *British Journal of Psychology*, **87**, 287–309.

Sloboda, J. A., Davidson, J. W., Moore, D. C., and Howe, M. J. A. (1994) Formal practice as a predictor of success or failure in instrumental learning. In *Proceedings of the Third International Conference for Music Perception and Cognition* (pp. 125–126). University of Liège, Belgium: URFM.

Todd, N. P.M. (1992) The dynamics of dynamics: A model of musical expression. *Journal of the Acoustical Society of America*, **91**, 3540–3550.

Trehub, S. E. (1990) The perception of musical patterns by human infants: The provision of similar patterns by their parents. In M. A. Berkeley and W. C. Stebbins (Eds.), *Comparative perception: Vol. 1. Basic mechanisms* (pp. 429–59). New York: Wiley.

Welch, C. F., Howard, D. M., and Rush, C. (1989) Real-time visual feedback in the development of vocal pitch accuracy in singing. *Psychology of Music*, **17**, 146–157.

CHAPTER 17

# ARE SOME CHILDREN MORE GIFTED FOR MUSIC THAN OTHERS?

## 17.1 Is 'talent' the best account of individual differences?

Imagine a common situation, in which a teacher gives a simple musical task to two 5-year-old children (for instance to reproduce a heard melody by singing it). Child A listens to the melody once, and then delivers a lively and accurate reproduction of it. Child B listens to the melody several times, but can only deliver a hesitant and inaccurate response. This pattern is repeated in a range of different tests. Clearly, the level of musical accomplishment of these two children is very different. How are we to explain such a difference?

It is very common for observers to invoke the notion of 'talent' or 'gift' at this point. Child A is said to musically gifted, whereas child B is not. Child A may be supposed to have a 'natural aptitude' for music, and 'inborn ability' to handle sounds. This is usually connected to the notion that such capacities are fixed and immutable. These kinds of lay explanations can then lead to prescriptions for action. For instance it may be concluded that child A will be able to profit from music instruction whereas child B will not.

It is really important to grasp that there is absolutely nothing within the performances of these two children which logically invokes explanations in term of 'talent', even if neither child has engaged in formal musical activities up to that point. The observed outcome is entirely consistent with a different explanation, in which we assume that the two children are equally 'talented', that is, they have exactly equal innate potential or aptitude for musical activity. The differences in their performances may be entirely due to differences in the knowledge-handling structures and strategies for dealing with music that they have acquired prior to testing.

Evidence for inheritance of differences in specific intellectual and mental characteristics is, in fact, very hard to find. There are clearly some specific genetic disorders which result in gross deficiencies in mental operation

over a wide range of tasks. In some cases, it has been possible to identify a single gene or a small number of genes which are responsible for the disorder (e.g. Huntingdon's disease). Researchers have not yet discovered any single genes which appear to accompany exceptionally high levels of ability. It is generally believed that the neurological underpinning for complex psychological abilities is multi-genetic, that is, the result of the action of many different genes. Although some researchers (e.g. Plomin and Thompson, 1993) are confident that new genetic techniques will eventually isolate individual genes associated with high IQ, no such genes have been discovered to date.

On the other hand there is a large and growing body of evidence from a number of sources (see Storfer, 1990 for a comprehensive review) that differences in early childhood experience can have a profound effect on later cognitive functioning. These early experiences can begin before birth, because the fetal hearing system achieves some degree of maturity during the second trimester. Indeed, research by Hepper (1991) shows that specific pieces of music, played to prenatal infants through loudspeakers placed on the mother's stomach, were then subsequently recognized by the infants when played to them after birth (as evidenced by changes in their attentiveness to these pieces as compared with novel pieces). Such learning can, of course, take place naturally. Sounds at above 80 dB sound pressure level are normally audible within the uterus. If pregnant mothers sing or play instruments, then their unborn children may, therefore, be able to hear them. Loud music in the immediate environment of the mother may also be heard clearly.

Most learning experiences, however, occur after birth. In our own studies of the early lives of young musicians (Sloboda and Howe, 1991; Howe and Sloboda, 1991a; Howe *et al.*, 1995) we have found that many, but not all, of the parents sang to their children (particularly at sleep time) every day from birth. Many also engaged in song games, encouraging children to dance and sing to music, and other kinds of informal home activities. Because these are seen by many parents as 'ordinary' activities, their importance as learning opportunities can be seriously underestimated (Papousek, 1982).

Although research on the long-term effects of early musical stimulation has not been carried out, there is good evidence for language stimulation (Whitehurst *et al.*, 1988; Fowler, 1990). In one study, parents were asked to engage with their one- and two-year-old children in simple structured interactive language games for a few minutes each day over a period of a few months. These games included such things as naming objects and then asking the child to repeat the name. Many parents do these activities as a matter of course, and are surprised to learn that other parents almost

never do these things regularly with their children. The study found that parents who did these games with their children secured substantial and long-lasting gains in the language competence of their children.

The conclusion from these studies is that very young children can receive extremely different levels of exposure to and engagement with music simply as a result of the informal musical activities of their immediate family members. By the time they start school, these differences can lead to wide disparities in ability to do a variety of musical tasks.

So far the discussion has treated young children as passive recipients of differing amounts of musical stimulation. But, of course, children are active participants in their own learning. As a result of pleasurable associations, or experiences of mastery (e.g. Renninger and Wozniak, 1985), some children will begin to seek out musical stimulation and activity in preference to other kinds of activity. They will ask for music games to be begun or continued. They will engage in spontaneous singing or dancing activities. They will ask for certain songs or records to be played over and over again. In this way, a child becomes to be seen by those around him/her as particularly sensitive to and involved in music, and therefore be considered to have a 'music oriented' personality.

Although some people like to think that a particular preference for musical sounds might be inborn in some infants, I think this is unlikely. An orientation towards sound is likely to be a universal feature of human beings when one considers its importance in early life (see, for instance, Fassbender, 1993). It seems to me much more likely that early preferences develop as a result of positive or negative experiences which incline the individual to engage or disengage from certain types of activity. These factors are essentially emotional or motivational, and I will return to discuss them in more detail in a later section. It is particularly important to realize that these positive or negative experiences can be private and unnoticed (by others). Parents may be convinced that they treated two siblings in exactly the same way, but because they cannot control every aspect of a child's experience, there is no guarantee that the two children will develop equal levels of musical skill. In fact, intra-family dynamics are often such that differences between siblings within a family can often be very marked indeed (Storfer, 1990). There is no greater fallacy than assuming that children in the same family have had the same experiences.

To summarize the argument so far, I hope I have demonstrated that it is generally impossible to conclude, from observing two children differing in musical behaviour, that they differ in musical talent, if by talent one means an inherited or inborn difference in capacity. This is true no matter how early one observes the behaviour, or what kind of behaviour it is one

observes. All 'tests' of musical 'ability' or 'aptitude' measure actual performance, not potential. It is, of course, true that well-designed tests are somewhat predictive of future achievement. But this is because those children that are already ahead are likely to stay ahead if all other things remain equal. If Child A knows more than Child B at time 1, and both learn equal amounts between time 1 and time 2, then Child A will still know more than Child B at time 2.

This is just one reason why the ascription of differences in talent tends to become a 'self-fulfilling' prophecy. Another reason concerns the effect that attributing 'talent' to an individual has on the beliefs and behaviours of that individual and those that interact with him or her. Dweck and colleagues (e.g. Dweck, 1986) have shown that children who believe intellectual capacity to be fixed are less resistant to failure on academic tasks, and will cease from problem-solving effort when faced with failure (the typical cognition is 'I can't do it' or 'I'm no good at maths', etc.). In contrast, children who believe that ability can be improved through learning and effort persist with tasks even when faced by initial failure. There is evidence (Vispoel and Austin, 1993) that children's beliefs about musical ability have a similar effect on their willingness to persist in solving difficult classroom musical tasks.

A detailed investigation of motivational factors affecting musical learning has been undertaken recently by O'Neill (O'Neill, 1996, 1997, O'Neill and Sloboda, 1997). In this study, 50 children were tested before starting their first instrument lesson, and then again after one year. A variety of tests of musical and intellectual ability were administered prior to instruction. There was also a test of 'resistance to failure'. In this, children were given a non-musical task at which they first succeeded, then failed due to experimenter manipulation. On a third attempt at the task, some children renewed their efforts and improved their performance, others stopped making an effort and their performance dropped. The first type of child was described as having 'mastery orientation', while the second type had 'helpless orientation'. These motivational orientations did very well at predicting instrumental progress after one year. In contrast, the music and intellectual ability measures did not predict progress at all.

It is also well established (Rosenthal and Jacobson, 1968; Brophy and Good, 1970) that teachers' beliefs about their students abilities affect their behaviour towards those children. When a teacher believes a child has high potential, the teacher is more likely to set the child challenging tasks, and encourage him/her to succeed. When a teacher believes a child has low potential, the teacher is more likely to set the child non-challenging tasks, and give less encouragement for high expectations. Several studies

have shown that one can improve a child's academic performance simply by giving the teacher apparently substantiated (but in fact bogus) information about the child's high academic potential.

Not only, therefore, is the ascription of differences in musical talent usually without empirical foundation or logical justification, it can actually reinforce and increase the very achievement differences upon which the original ascription was made. That being so, the most consistent and educationally sound conclusion to adhere to at the present time is that if there is such a thing as musical talent (interpreted as an inborn capacity for and predisposition towards musical activity), then it exists to equal measure in the vast majority of the population, and in no way accounts for the very wide range of adult musical accomplishment that exists in the populations of industrialized societies such as ours.

This conclusion is, I have to admit, not accepted by all researchers in the field. Debate continues to flourish, and readers interested in assessing alternative positions are referred to the sets of published commentaries on two target articles which elaborate the broad position taken here (Sloboda *et al.*, 1994; Howe *et al.*, 1998). A forceful defence of 'talent' may be found in Winner (1996).

With our conclusion in mind we can turn to a question which I believe to be much more profitable. That is, what are the psychological mechanisms that underlie musical achievements? Such psychological mechanisms taken together constitute the 'ability' of a person. I am much happier with the term 'ability' than I am with the term 'talent'. To label someone as able does not seem to me to involve any assumptions about the origins of that ability. When we understand something of the nature of musical ability we will be in a much better position to explain, predict, and bring it about (Sloboda, 1996).

## 17.2 The nature of musical ability

Although the term 'musical ability' is somewhat less laden with interpretation than the term 'talent', it may still presuppose too much. It suggests, for instance, that there is some common factor, or set of factors, underlying all accomplishments in the sphere of music. How does this square with the fact is that there are singers who cannot read music, pianists who cannot sing in tune, performers who cannot compose, and music critics who can neither play an instrument nor compose? Do all such people possess some common attribute in virtue of which they can be said to be musically able?

Much contemporary research on music supports the notion that there is indeed such a common attribute, which is the ability to *make sense* of

musical sequences, through the mental operations that are performed on sounds (whether real or imagined). The term 'make sense' is something of an analogy. Musical idioms are not languages in the sense that they do not refer to states of affairs in the world in the way that languages like English do. They do, however, share with languages structures which resemble syntax or grammar (Lerdahl and Jackendoff, 1983). Musical ability entails the ability to detect and use these structures in the mental manipulation of music. Another term that is sometimes used for the process of making sense of music is 'audiation'.

How do we know if and when a person is 'making sense' of some music? There are a number of ways in which psychologists have operationalized this ability. First, people who are 'making sense' of music should be able to remember music which conforms to a cultural 'grammar' (e.g. diatonic tonality) better than music which does not. There are many demonstrations that this is so, and that fairly young children already show such an advantage (e.g. Zenatti, 1969; Deutsch, 1980). Similar things have been found for other skills such as chess. For instance, chess players recall meaningful (i.e. game plausible) boards better than random boards (Chase and Simon, 1973). Second, people who 'make sense' of music tend to make grammatically plausible substitutions when asked to recall music they have just heard (Sloboda and Parker, 1985; Sloboda *et al.*, 1985; Oura and Hatano, 1988). This is similar to the well-established finding that people rarely remember verbal information word-for-word, but reconstruct in their own words something that *means* the same as what they heard. This kind of finding suggests that people generally store something more abstract than the actual words or notes.

A third criterion for 'making sense' of music, is the ability to correctly judge whether or not given sequences are acceptable according to cultural rules. There is strong evidence that most children can reject blatant violations, such as discords, by the age of seven, and more subtle ones, such as unfinished cadences, by the age of ten (Sloboda, 1985). A final criterion is the ability to correctly identify the consensual mood or emotion of a musical passage. Again, we have rather strong evidence that the broad parameters of this ability are in place for most children by the age of five (Pinchot-Kastner and Crowder, 1990; Terwogt and Van Grinsven, 1991).

Although people vary quite widely in the level of sophistication to which they have developed their ability to make sense of music, the available evidence points to the conclusion that the vast majority of the population have acquired a common ability to make sense of music, clearly evident through experimental demonstrations such as the ones described above, by the age of 10, regardless of accomplishment in any particular sphere of

musical performance, and regardless of having been in receipt of any formal musical education or training.

This is quite an important conclusion to bear in mind, because most people think of specific performance skills, like being able to sing or play the piano, as being the principal signs of ability. These are the kinds of skills that tend to be most often addressed, noticed, and assessed in formal education. My point is that these specific visible performance skills usually rest on previously acquired but 'hidden' receptive skills. If we return to the two hypothetical children discussed in the opening section of this chapter, we can see that it is quite possible that both children are equally 'able' (same level of ability to 'make sense' of the music) but perform differently because child A has had specific singing-related experiences which child B has not had. In this case, appropriate singing experience should help child B catch up with child A. On the other hand, child B may do less well than child A because she has not acquired the 'ability' (cognitive procedures) necessary to make sense of the music. In this case, specific singing experience may not allow child B to catch up with child A, because the deficit is at a more fundamental level. I believe that it may sometimes be this latter case that persuades teachers and parents that there are differences in 'talent' between two children. The argument of this section is that these differences should more properly be seen as differences of ability, susceptible to reduction or removal by appropriate learning. It is simply that the kind of learning required in the two cases is different.

All of the preceding argument does not mean, of course, that no genetic factors are ever important in the growth of musical ability and achievement. Operatic singers require certain vocal tract characteristics which affect resonance and quality. Much piano repertoire demands a large hand span. Sometimes children who have good voices may be labelled as 'gifted' even when their musical ability is not high. This may lead to the kinds of life experiences which eventually produce high levels of musical ability. But having a good voice is not, in itself, any indication of musical ability, any more than possessing a Stradivarius will of itself turn an indifferent violinist into a good one.

## 17.3 The acquisition of musical ability

What features of music and the human mental apparatus make it possible for most people to acquire a basic level of musical receptive ability without formal training? There is broad agreement between music theorists and psychologists (e.g. Meyer, 1956; Deutsch, 1982; Lerdahl, 1988; Bregman, 1990; Narmour, 1990; Huron, 1991) that the most prevalent musical

idioms have structural and mathematical properties that are easily picked up by basic perceptual mechanisms that are built in to the 'hard-wiring' of the brain. These properties include repetition and grouping according to pitch or duration. Contemporary theories of learning applied to music (e.g. the 'connectionist' models of Bharucha, 1987 and Gjerdingen, 1990) have demonstrated how quite complex mental representations may be built up from such basic processes as a result simply of repeated exposure to a variety of musical examples sharing similar structures. These representations will include knowledge of what we would describe as scale or tonality, such that pairs of sequences like C–D–E–F–G and F–G–A–B–C will be detected by the system as 'the same' (i.e. both scale sequences in the key of C-major) even though their intervallic and physical characteristics are different. The sophistication of the patterns that the system can match in some way will be a direct function of the number and range of musical examples to which it has been exposed. In human terms this could come about by precisely the differing levels of childhood experiences discussed in Section 17.1. The more sophisticated the system, the better 'sense' it makes of a wider range of musical material. Through learning of this sort, a basic and universal system is transformed into a culturally tailored system that reflects and responds best to the music in the surrounding environment of the individual. This allows us to understand why it is that people find it hard to make initial sense of music from other cultures. For instance, most Westerners find it hard to reproduce melodies of Indian or Arabic origin, which are built on different scale systems.

We also have the beginnings of an understanding of how the confirmation and violation of expectations built up within such a system may lead to the experience of emotion and mood to music. Many emotions to music seem to be responses to various kinds of surprise within the harmonic or rhythmic structure (Meyer, 1956; Jackendoff, 1991; Narmour, 1991; Sloboda, 1991a).

## 17.4 Performance ability

The foregoing discussion should have helped to make it clear that there are several distinct abilities involved in performing a musical instrument, and only some of them are 'musical' in the sense we have been discussing. The other abilities are what I would call 'technical'. For instance, a pianist may be able to play a difficult passage very quickly, very evenly, and very accurately. This will be possible because of very highly developed control systems for execution of hand and finger movements. Such systems will have required much practice to develop to high levels, and are manifestations of expertise. It is, however, possible that a person could learn to play a piece

of music 'technically' without having 'understood it' musically at all. Such a 'technically perfect' performance is one that could equally well be generated by mechanical means.

It is well-established that notationally 'perfect' performances of music are experienced by most listeners as dull, mechanical and uninteresting. What makes any performance musically interesting are the slight fluctuations in duration, loudness, pitch, and timbre which together constitute *expressive performance* (see Chapter 9). Such fluctuations are what distinguish one performance from another, and their importance is what keeps professional performers in business. Were it not for this, we could get computers to generate once-and-for-all perfect performances of musical works, and close most of our symphony halls and conservatories!

Recent research has made it increasingly clear that the existence of expressive performance is the best evidence we can obtain that a musician *understands* the music she is playing. This is because expressive variations are by no means random or idiosyncratic. A wide range of studies (e.g. Shaffer, 1981; Sloboda, 1983; Clarke, 1988; Gabrielsson, 1988) has shown that microvariations are highly systematic, both within the same performer and across different performers. Many of such variations are designed to make important structural features of the music more prominent to the listener. Their systematicity therefore depends upon the performer having understood what *are* the important structures. The most convincing evidence that this understanding is deeply rooted in a performer may well be to ask that performer to sight-read some music she has never seen before. If she can apply appropriate expression in such a situation, then her understanding must be fully internalized. In contrast, in a rehearsed performance there always exists the remote possibility that expressive devices have been learned 'by rote' without any real understanding. Evidence of expressively appropriate sight-reading does indeed exist (Sloboda, 1983), and it is very often so deeply internalized that the performer is not fully aware of the details of her own expressive repertoire. It has become intuitive and semi-automatic. This should not surprise us. It is another well-documented characteristic of all cognitive expertise. Many people mistakenly assume that intuitive behaviour must be innate. This is a major fallacy. Any well-practised habit eventually becomes automatic.

It may seem at first sight that the statement that expression is systematic conflicts with the existence of interesting differences between performers. The conflict is, however, more apparent than real. Performers differ from one another, not so much in the nature of expressive devices used, but in their distribution and intensity. So, for instance, one performer may achieve by subtle dynamic fluctuations what another one achieves by

temporal fluctuations. Such differences between performers, or between the same performer at different times, can be characterized as differences in expressive 'style' (Sloboda, 1985).

Because high-level musical performance has this dual technical–expressive aspect, the conditions for the development of expert performers must allow development on both fronts. This is quite difficult to achieve. There are many music professionals who are technically adept but lack a high level of expressive skill. There are also many non-professionals who have a high level of expressive insight, but cannot realize their expressive intentions in music of any technical complexity. The master-performer is that relatively rare person who has managed to develop both technical and expressive skills to a high level in tandem. How is it done?

## 17.5 Making the expert performer

We now have evidence from a number of biographically based research studies (Sosniak, 1985; Sloboda and Howe, 1991; Ericsson *et al.*, 1993; Sloboda *et al.*, 1996) that confirm that music is no exception to the general rule 'practice makes perfect'. The study of Ericsson *et al.* shows that student violinists rated as excellent by their professors have, by the age of 21, accumulated on average around twice as many hours of practice over the lifespan (10 000 hours) as more average players (5000 hours). These data were obtained from a study where violinists were asked to estimate retrospectively the average number of hours per week of practice for each year of their life. For the current year these data were confirmed by practice diaries kept by all participants.

Sosniak, in a study of 24 top US pianists, showed that none of them showed particular signs of exceptional ability at the outset of training. Exceptionality was something which developed gradually through the early years of formal training. Indeed, the notion that very early achievement is the normal precursor of adult excellence finds very little support in the documented research literature. The 'child prodigy' may be the exception rather than the norm.

Sloboda *et al.* (1996) extended the generality of the above findings by obtaining practice data from various categories of young musician, from the highly able to the least able. Their data showed that rate of progress (as measured through objective examination scores) was entirely a function of the number of hours of practice undertaken. A common belief held by teachers and parents is that 'talented' children can achieve a given level of performance after less practice than 'untalented' children. The data from this study firmly refute this idea. Able musicians achieve more because

they practise more. For instance, at age 12, the children who were most advanced were practising on average for 120 minutes per day. The children who were least advanced at the same age were achieving average daily practice of only 15 minutes. This represents an 800% difference between the best and worst players. No wonder the able performers did so well!

It is not very easy for a young person to accumulate 10 000 hours of practice between starting an instrument (typically around the age of 6 or 7 at the earliest) and the age of 21. To give some idea of the workload that implies, it would require 2 hours every day of the year for 14 years to accumulate 10 000 hours of practice. In reality, the typical 'expert' practice pattern begins with 15–30 minutes per day, and increases to 4–6 hours per day by the age of 21. Practice activities are not always inherently rewarding to young children, and even the most able young musicians often required parental support, even direct supervision, to accomplish regular practice (Sloboda and Howe, 1991; Davidson *et al.*, 1996).

Indeed, the quality of the child's relationship with significant adults, both parents and teachers, appears crucial in predicting long-term involvement with musical performance. The parents or high achievers are characterized by a very high level of involvement in their children's musical life. In addition, students tend to remember their first teachers, not so much for their technical adeptness, as for the fact that they made lessons fun. They communicated both their love of music and their liking for their pupil. Personal warmth seems an important characteristic. Critical, confrontational, or achievement-oriented approaches, seems not to be successful, at least in the early stages of learning (Davidson *et al.*, 1998).

Of course, sheer amount of practice is a rather crude measure from a psychological point of view. The nature of learning will depend on what is done during practice as well as how much of it is done. Surprisingly, there is little reliable scientific information on how musicians practise (what exists is well summarized by Hallam, 1997). One clear finding is that experienced performers pay more attention to structural matters when practising. Clearly, since expressive performance is closely related to the detailed structure of a composition, practice that shows awareness of structural factors is more likely to lead to the development of an appropriate expressive repertoire.

## 17.6 Emotion and motivation

There is a considerable body of evidence that there are two types of motivation to engage with music. One motivation is what one might call 'intrinsic'. It develops from intense pleasurable experiences with music (of a sensual, aesthetic, or emotional kind) which lead to a deep personal commitment

to music. The other motivation is what one might call 'extrinsic', and is concerned with achievement. Here, the focus is not so much on the music itself as on achieving certain goals (e.g. approval of parents, identification with role models, winning competitions). Clearly any one individual will have a mixture of both types of motivation. There is some evidence, however, that too early an emphasis on achievement can inhibit intrinsic motivation. In simple terms, children become so concerned about what others may be thinking of their performance, that they have little attention left for the potential of the music to engage their aesthetic and emotional sensibilities deeply. All music becomes a source of anxiety.

This conclusion is supported by a study of autobiographical memories (Sloboda, 1990). In this study, musician and non-musician adults were asked to recall events from the first 10 years of life that had any connection at all with music. They were given a number of questions to stimulate recall, such as where the event took place, what event the music was part of, who they were with, and what significance the experience had for them. The 113 'stories' obtained in this way could be classified along two dimensions. One related to the 'internal' significance of the event, that significance which was attributed to the music itself. The other related to the 'external' significance, attributable to the context. In each case the value of the dimension could be positive, neutral, or negative.

Many of the musicians reported deeply felt and intensely positive early experiences to the 'internal' aspect of musical events, which seemed to lift them outside the normal state of awareness. For instance, one young woman reminisced as follows:

> I was seven years old, and sitting in morning assembly in school. The music formed part of the assembly service. I was with my friends Karen, Amelia, Jenny, Allan. The music was a clarinet duet, classical, probably by Mozart. I was astounded at the beauty of the sound. It was liquid, resonant, vibrant. It seemed to send tingles through me. I felt as if it was a significant moment. Listening to this music led to me learning to play first the recorder and then to achieve my ambition of playing the clarinet. Playing the clarinet has altered my life; going on a paper round and saving up to buy my own clarinet; meeting friends in the county band. . . . Whenever I hear clarinets being played I remember the impact of this first experience.

Others, more often the non-musicians, recalled events where the music itself was not remembered as significant in itself, but rather its context, which was often one of anxiety, humiliation, or embarrassment.

Being made to perform in front of other, being criticized, being laughed at, were common experiences. In almost every mixed audience where I report these results, a large proportion of the non-musicians identify just such experiences which they believe turned them away from music and persuaded them that they were not 'musical'. In almost none of these cases was the person able to experience the music as a source of positive significance.

Frequency analyses of the subcomponents of the stories showed three factors which were statistically associated with positive internal experiences:

- the event occurred at home, in church, or at a concert hall, rather than at school
- the event occurred while the child was listening rather than performing
- the child was on his/her own, with family or friends, rather than with a teacher.

In each case, these conditions seem to be connected with lack of expectation of or assessment of the child, but rather a relaxed, non-threatening environment where nothing is being asked of the child. Perhaps such an environment is necessary for music to work its strongest effects on individuals.

These deeply positive 'peak' childhood experiences seem important for the development of musical ability for two reasons. First, these experiences are so pleasurable that children often increase their engagement with music in the hope of repeating them. Second, these experiences seem to be intimately connected to the person's understanding of the musical structures which are crucial to performance expression. Sloboda (1991a) has shown that adult music listeners identify many of their moments of most intense emotional response with quite specific musical structures such as appogiaturas or enharmonic changes. These structures are ones which manipulate listener's expectancies in some way. Unless one has experienced the 'delicious' surprise of an enharmonic change through listening, it is hard to see how one could effectively add the appropriate performance expression to such a change to heighten its effect for listeners (such expression might arguably be a slight slowing which delays and emphasizes the onset of the unexpected chord). Experience must precede performance. Children who are focusing emotional attention on the performance and other extrinsic factors rather than on the music itself may not be able to build the structure–emotion links that are the key to spontaneous expressive playing.

It seems that our society, particularly, our system of formal education, is set up to produce anxiety and self-doubt rather than love of music. What may be needed to remedy that is the rehabilitation of the notion

that, in the sphere of music if nowhere else, deep emotional experiences and sensual enjoyment are as important and valid as hard work and technical achievement. One of the reasons why music seems to be seen by many children as an essentially female activity may be the strong male conditioning against emotional display, which is fully present by the age of 6 or 7 in most male cultures. One of the unexpected findings of Sloboda and Howe's study was the huge sense of relief and safety felt particularly by some boys, when they were able to get to a specialist music school where music was at last seen as a quite normal activity for boys, and where they could escape the mockery and misunderstanding of their peers (Howe and Sloboda, 1991b).

If school experiences are such strong inhibitors of musicality for many children, this may go a long way to explain why so few children ever reach high levels of performance ability. Children who succeed may well be the children who had lucky early experiences, supportive rather than undermining musical relationships, and encouragement of musical autonomy and feeling. It is salutary to remember that many excellent musicians received little or no formal training at all (see for example Collier's (1983) detailed description of Louis Armstrong's acquisition of jazz skill, summarized and commented on by Sloboda, 1991b), and that many young people develop considerable expertise as popular musicians with little or no help from their classically trained teachers. There also exist many cultures, especially traditional non-Western cultures, where the spread of musical accomplishment in the population is much more even than it is in modern industrialized cultures. In such cultures, participation in musical activity is often more rooted in the whole everyday life and work of the community, rather than being conceived of as something separate or special (e.g. Blacking, 1976).

Finally, there is considerable evidence that children are very aware of the associations of different types of music with social group identity. Very often schools present children with the notion that classical music is superior to the popular and folk idioms with which they more naturally identify, and with which they have the greatest experience. Children can both reject classical music as 'too highbrow' and also become intimidated by the apparent requirement to step outside what seems comfortable and familiar before they can produce 'acceptable' music. Most teachers are aware of the child who presents as incompetent in singing class, but can be found the same day in the school bus joining in an informal 'sing-along' session with competence and enthusiasm. There are strong social pressures from peers to conform to the norm and hide a potential liking for music which is 'different' (Finnas, 1987). This means that teachers need to introduce

music through, rather than in opposition to, the cultural musical norms of their students.

## 17.7 From able child to able adult

Most of what we have discussed to this point relates to the development of musical ability through childhood into adolescence. Many promising adolescents never, however, reach the highest levels of musical excellence. Some of these are 'happy' drop-outs, individuals who have many other interests and skills and are not prepared to make the level of commitment required for a professional music career. They are content to remain as proficient amateurs. Others drop out as part of the process of developing self-knowledge and autonomy. They realize that their parents' investment in their musical development has always been considerably greater than their own and that they have never really acquired their own individual commitment to music. A third group (described in some detail by Bamberger, 1986) are those who experience a difficult transition from the unreflexive and intuitive modes of musical thinking characteristic of children to the more analytical and self-critical modes of adult thinking where what once seemed easy now appears to 'fall apart' and become problematic.

Among those who continue to higher levels of training, the most successful musicians appear to be those with the highest levels of commitment and self-discipline. The best violinists in Ericsson *et al.*'s (1993) study were those who practised regularly every morning at the same time, spent more time in daytime sleeping or 'napping', particularly in the afternoon (arguably needed to recover from the effortful activity of practice) and spent least time on leisure activities.

A study by Manturszewska (1990) suggests that a further precondition of transition to musical maturity is the existence of a deep and significant 'master–student' relationship with a respected older professional, very often the student's instrumental teacher. Even at this relatively late stage the relationship is concerned with more than technical and performance development. It is a relationship of personal involvement which often leads to lifelong friendship. The existence of a older and more well-connected 'mentor' has been observed in many other fields of creative and professional endeavour as a frequent accompaniment to the highest levels of achievement. The development of such relationships seems to depend, in part, on the stability of the student's relationship with his or her own parents (Butler, 1993). Where parents have been over-critical and under-supportive, students tend to be fearful of their teachers and project onto them the role

of 'punishing parent' in a way which undermines the possibility of healthy 'mentor' relationships developing.

## 17.8 Conclusions

I hope that this chapter has provided the basis of possible answers to the kinds of question that lay people often pose to psychologists. Individual differences, even within families, are easily accounted for by individual differences in experience, leading to different motivations and learning outcomes. No family can provide the 'same' environment for two siblings. Indeed the mere fact that one sibling takes up music may incline the other sibling not to.

It is impossible to accurately predict the future musical achievement of a given child from its behaviour at age 5. Much depends on the circumstances in which music is introduced and how parents and teachers handle the musical relationship. Although there may be a statistical correlation between early and later achievement in the population as a whole, this should not be taken as a deterministic causal relationship.

What we can say with some certainty is that there are a set of circumstances which will increase the chances of attainment of high levels of excellence. These are, early and frequent casual exposure to the music of a culture, opportunities to engage with this music in a relaxed enjoyable environment without performance expectations, early positive encouragement from parents and teachers to engage in musical activity but respect for the child's own autonomy and explorations, a balance between technical and expressive activities, an increasing commitment to high levels of time investment in practice, and the eventual commitment to excellence in the context of a supportive and long-lasting 'master–student' relationship.

This picture contradicts three deeply ingrained cultural myths. The first myth is that musical achievement depends upon the pre-existence of a rare inherited quality 'talent'. I have argued instead that it builds on a basic aptitude which is shared by almost all human beings. The second myth is that musical excellence grows from within by solitary effort. The studies summarized here emphasize the social nature of musical development, in which the achievement of one is the result of effort by many. The third myth is that work and pleasure are separate. The evidence I have outlined here suggests that unless one gains and retains deep pleasure in music throughout the years of learning, the pinnacles of expressive performance will always remain out of reach.

## References

Bamberger, J. (1986) Cognitive issues in the development of musically gifted children. In R. J. Sternberg and J. E. Davidson (Eds.) *Conceptions of Giftedness*. Cambridge: Cambridge University Press.

Bharucha, J. J. (1987) Music cognition and perceptual facilitation: a connectionist framework. *Music Perception*, **5**, 1–30.

Blacking, J. (1976) *How Musical is Man?* London: Faber.

Bregman, A. S. (1990) *Auditory Scene Analysis: the perceptual organisation of sound*. Cambridge, MA.: MIT Press.

Brophy, J. E. and Good, T. L. (1970) Teachers, communication of differential expectations for children's classroom performance: some behavioural data. *Journal of Educational Psychology*, **61**, 365–74.

Butler, C. (1993) The effects of psychological stress on the success and failure of music conservatoire students. MPhil Thesis: University of Keele, England.

Chase, W. G. and Simon, H. A. (1973) Perception in chess. *Cognitive Psychology*, **4**, 55–81.

Clarke, E. F. (1988) Generative principles in music performance. In J. A. Sloboda (Ed.) *Generative Processes in Music: the psychology of performance, improvisation, and composition*. London: Oxford University Press.

Collier, J. L. (1983) *Louis Armstrong: An American genius*. New York: Oxford University Press.

Davidson, J. W., Howe, M. J. A., Moore, D. G., and Sloboda, J. A. (1996) The role of parental influences in the development of musical ability. *British Journal of Developmental Psychology*, **14**, 399–412.

Davidson, J. W., Sloboda, J. A., Moore, D. G., and Howe, M. J. A. (1998) Characteristics of music teachers and the progress of young instrumentalists. *Journal of Research in Music Education*, **46.1**, 141–160.

Deutsch, D. (1980) The processing of structured and unstructured tonal sequences. *Perception and Psychophysics*, **23**, 215–218.

Deustch, D. (1982) Grouping mechanisms in music. In D. Deutsch (Ed.) *The Psychology of Music*. New York: Academic Press.

Dweck, C. S. (1986) Motivational processes affecting learning. *American Psychologist*, **41**, 1040–1048.

Ericsson, K. A., Krampe, R. T., and Tesch-Romer, C. (1993) The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, **100.3**, 363–406.

Fassbender, C. (1993) *Auditory Grouping and Segregation Processes in Infancy*. Norderstedt: Kaste Verlags.

Finnas, L. (1987) Do young people misjudge each other's musical tastes? *Psychology of Music*, **15**, 152–166.

Fowler, W. (1990) Early stimulation and the development of verbal talents. In M. J. A. Howe (Ed.) *Encouraging the Development of Exceptional Skills and Talents.* Leicester: BPS Books.

Gabrielsson, A. (1988) Timing in music performance and its relations to music experience. In J. A. Sloboda (Ed.) *Generative Processes in Music: the psychology of performance, improvisation, and composition.* London: Oxford University Press.

Gjerdingen, R. O. (1990) Categorization of musical patterns by self-organizing neuronlike networks. *Music Perception,* **7**, 339–369.

Hallam, S. (1997) What do we know about practising? Towards a model synthesising the research literature. In H. Jorgensen and A. C. Lehmann (Eds.). *Does Practice Make Perfect? Current theory and research on instrumental and music practice.* Oslo: Norges Musikkhogskole, pp. 53–70.

Hepper, P. G. (1991) An examination of fetal learning before and after birth. *Irish Journal of Psychology,* **12**, 95–107.

Howe, M. J. A. and Sloboda, J. A. (1991a) Young musicians' accounts of significant influences in their early lives: 1. the family and the musical background. *British Journal of Music Education,* **8**, 39–52.

Howe, M. J. A. and Sloboda, J.A. (1991b) Problems experienced by young musicians as a result of the failure of other children to value musical accomplishments. *Gifted Education,* **8**, 102–111.

Howe, M. J., A., Davidson, J. W., Moore, D. G., and Sloboda, J. A. (1995) Are there early childhood signs of musical ability? *Psychology of Music,* **23**, 162–176.

Howe, M. J. A., Davidson, J. W., and Sloboda, J. A. (1998) Innate talent: reality or myth? *Behavioural and Brain Sciences,* **21.3**, 399–442.

Huron, D. (1991) The avoidance of part-crossing in polyphonic music: perceptual evidence and musical practice. *Music Perception,* **9**, 93–104.

Jackendoff, R. (1991) Musical parsing and musical affect. *Music Perception,* **9**, 199–230.

Lerdahl, F. (1988) Cognitive constraints on compositional systems. In J. A. Sloboda (Ed.) *Generative Processes in Music: the psychology of performance, improvisation, and composition.* London: Oxford University Press.

Lerdahl, F. and Jackendoff, R. (1983) *A Generative Theory of Tonal Music.* Cambridge, MA: MIT Press.

Manturszewska, M. (1990) A biographical study of the life-span development of professional musicians. *Psychology of Music,* **18**, 112–139.

Meyer, L. B. (1956) *Emotion and Meaning in Music.* Chicago: University of Chicago Press.

Narmour, E. (1990) *The Analysis and Cognition of Basic Melodic Structures: the implication-realisation model.* Chicago: University of Chicago Press.

Narmour, E. (1991) The top-down and bottom-up systems of musical implication: building on Meyer's theory of emotional syntax. *Music Perception,* **9**, 1–26.

Oura, Y. and Hatano, G. (1988) Memory for melodies among subjects differing in age and experience in music. *Psychology of Music*, **16**, 91–109.

O'Neill, S. A. (1996) Factors influencing children's motivation and achievement during the first year of instrumental music tuition. Unpublished PhD Thesis, Keele University, UK.

O'Neill, S. A. (1997) The role of practice in children's early musical performance achievement. In H. Jorgensen and A. C. Lehmann (Eds.). *Does Practice Make Perfect? Current theory and research on instrumental and music practice.* Oslo: Norges Musikkhogskole, pp. 53–70.

O'Neill, S. A. and Sloboda, J. A. (1997) The effects of failure on children's ability to perform a musical test. *Psychology of Music*, **35.1**, 18–34.

Papousek, M. (1982) The 'mother tongue method' of music education: psychobiological roots in oroverbal parent-infant communication. In J. Dobbs (Ed.). *International Music Education.* ISME Yearbook 1982.

Pinchot-Kastner, M. and Crowder, R.G. (1990) Perception of the major/minor distinction: IV. Emotional connotations in young children. *Music Perception*, **8**, 189–202.

Plomin, R. and Thompson, L. A. (1993) Genetics and high cognitive ability. In K. Ackrill (Ed.) *The Origins and Development of High Ability: Proceedings of Ciba Symposium 178.* London: Wiley.

Renninger, K. A. and Wozniak, R. N. (1985) Effect of interest on attentional shift, recognition and recall in young children. *Developmental Psychology*, **21**, 624–632.

Rosenthal, R. and Jacobsen, L. (1968) *Pygmalion in the Classroom: teacher expectation and pupil's intellectual development.* New York: Holt, Rinehart and Winston.

Shaffer, L. H. (1981) Performances of Chopin, Bach, and Bartok: studies in motor programming. *Cognitive Psychology*, **13**, 326–376.

Sloboda, J. A. (1983) The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology*, **35A**, 377–396.

Sloboda, J. A. (1985) Expressive skill in two pianists: style and effectiveness in music performance. *Canadian Journal of Psychology*, **39**, 273–293.

Sloboda, J. A. (1990) Music as a language. In F. Wilson and F. Roehmann (Eds.) *Music and child development.* St. Louis, MI: MMB Inc.

Sloboda, J. A. (1991a) Music structure and emotional response: some empirical findings. *Psychology of Music*, **19**, 110–120.

Sloboda, J. A. (1991b) Musical expertise. In K. A. Ericsson and J. Smith (Eds.) *Toward a General Theory of Expertise: prospects and limits.* New York: Cambridge University Press.

Sloboda, J. A. (1996) The acquisition of musical performance expertise: deconstructing the 'talent' account of individual differences in musical expressivity. In K. A. Ericsson (Ed.) *The Road to Excellence: the acquisition of expert performance*

*in the arts and sciences, sports and games.* Mahwah, NJ: Lawrence Erlbaum Associates, pp. 107–126.

Sloboda, J. A. and Howe, M. J. A. (1991) Biographical precursors of musical excellence: an interview study. *Psychology of Music,* **19**, 3–21.

Sloboda, J. A. and Parker, D. H. H. (1985) Immediate recall of melodies. In Howell, P., Cross, I., and West R (eds) *Musical Structure and Cognition.* London: Academic Press.

Sloboda, J. A., Davidson, J. W., and Howe, M. J. A. (1994) Is everyone musical? *The Psychologist,* **7.7**, 349–354.

Sloboda, J. A., Hermelin, B., and O'Connor, N. (1985) An exceptional musical memory. *Music Perception,* **3**, 155–170.

Sloboda, J. A., Davidson, J. W., Howe, M. J. A., and Moore, D. G. (1996) The role of practice in the development of expert musical performance. *British Journal of Psychology.* **87**, 287–309.

Sosniak, L. A. (1985) Learning to be a concert pianist. In B. S. Bloom (Ed.) *Developing Talent in Young People.* New York: Ballantine.

Storfer, M. D. (1990) *Intelligence and Giftedness: the contributions of heredity and early environment.* San Francisco: Jossey Bass.

Terwogt, M. M. and Van Grinsven, F. (1991) Musical expression of moodstates. *Psychology of Music,* **19**, 99–109.

Vispoel, W. P. and Austin, J. R. (1993) Constructive response to failure in music: the role of attribution feedback and classroom goal structure. *British Journal of Educational Psychology,* **63**, 110–129.

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., DeBaryshe, B. D., Valdez-Menchaca, M. C., and Caulfield, M. (1988) Accelerating language development through picture book learning. *Developmental Psychology,* **24**, 552–559.

Winner, E. (1996) *Gifted Children: myths and realities.* New York: Basic Books.

Zenatti, A. (1969) Le developpement genetique de la perception musicale. *Monographs Français Psychologique,* **17**.

PART D

# MUSIC IN THE REAL WORLD

CHAPTER 18

# EVERYDAY USES OF MUSIC LISTENING: A PRELIMINARY STUDY

## 18.1 Introduction and research background

There exists a particular attitude to music listening which implicitly dominates much psychological theorizing and experimentation in the study of individual response to music. This is what the author has caricatured (Sloboda, 1989) as the *pharmaceutical* model. Listeners are construed as passive recipients of musical stimuli which have the psychological effect they do because of the way the human brain is constructed on the one hand, and the way the music is structured on the other. Although almost all psychologists would reject the crudest versions of this model (for instance, by stressing effects of experience and learning on listener response) dominant research practices nonetheless perpetuate key assumptions of this model. For instance, most studies of music listening present listeners with music which is not of their choosing in an environment (the laboratory) which is controlled and constructed by the experimenter.

The pharmaceutical approach may be one of the reasons for the particularly slow progress made in the scientific understanding of emotional and aesthetic response to music. Laboratory studies which show, for instance, that most listeners within a given culture will agree which adjectives, from a large checklist, best describe the emotional character (or their aesthetic evaluation) of a particular piece or performance, may have been used to draw over-strong inferences about how musical stimuli mediate affective and evaluative responses (Gabrielsson and Juslin, 1996; Hevner, 1936; Scherer and Oshinsky, 1977). It might be that, in some other listening context, quite different forms of consistency are obtained, or that consistency breaks down altogether.

This paper takes a diametrically opposite position to the pharmaceutical one, and emphasizes the role of the listener as an active agent, who makes choices about what music to listen to, where and when to listen to it, and what to listen for, according to needs, goals, and purposes. Music listening

is, thus, intensely situational. One cannot begin to give a full account of the cognitive, affective, or aesthetic response of a listener, without paying detailed attention to the reasons why the person is listening to that particular music at that particular time (indeed, we also need a much fuller characterization of different *types* of listening that might be going on in different people in different contexts). Some recent psychological research has begun to acknowledge the importance of situation to musical response. For example, the work of North and Hargreaves (1996, 1997) shows how people's liking of and preference for different types of music is influenced by the activity which the music is accompanying (e.g. aerobics versus meditation). One could, however, go well beyond this, and surmise that the situation can determine, in a much more fundamental way, what it is that the listener hears in or extracts from the music.

Since the seminal work of Bourdieu (1984) it is generally accepted within the field of popular culture studies that consumers of cultural products appropriate these products to support their social identities, lifestyles, and personal purposes. These appropriations can be quite different from those predicted or planned for by producers and retailers (as is evident, for instance, in the account of the Sony Walkman's appropriation given by du Gay *et al.* 1997). In the light of this, it is somewhat surprising that theories of appropriation are typically supported by rather thin data-sets, relying on journalism or anecdote at worst, or a small number of in-depth case studies at best.

As far as we are aware, the 'functional terrain' of music listening situations has not yet been systematically mapped within contemporary Western culture. It is a curious paradox that we probably know (through the work of anthropologists and ethnomusicologists) more about the different purposes for which music is used within certain non-Western societies than we do about how it is used in western consumer societies (see, for example, Gregory, 1997). Only music performance seems to have attracted any significant and comprehensive work (Finnegan, 1989).

Nonetheless, some strands of psychological work provide some pointers regarding possible functions of music listening for particular people and situations. For instance, both Gabrielsson and Lindstrom (1996) and Sloboda (1992) have collected data on strong experiences to music, and have established a range of mood-altering functions that could be generally classified as therapeutic. Behne (1997) has shown that adolescents reporting high frequency of personal problems also report higher frequency of listening to music as some kind of escape (a classification described as 'sentimental'). Music is also used to reinforce social identities, a phenomenon particularly commented on in adolescence (Zillman and Gan, 1997).

The present study provides preliminary groundwork for a planned large-scale investigation into everyday uses of music.

## 18.2 Pilot tests

Some suggestive data has been obtained in two unpublished student projects supervised by the author (Neilly, 1995, De Las Heras, 1997). Neilly (1995) asked 20 undergraduate students to keep a diary of each time they exposed themselves to music by their own choice over a period of three days. She found that 60% of the sample used music to wake up/get up in the morning, 50% used music while they studied, 40% used music while they got ready to go out in the evening, and 30% used it at night to relax/go to sleep. Although it offers some suggestive pointers, this small sample of undergraduate students obviously cannot be taken to be representative of the population.

De Las Heras (1997) asked 82 music listeners to complete a comprehensive questionnaire about their use of records, tapes, CDs and radio in the home. The study included 45 Likert-scale items probing reasons and contexts for music use, which had been generated from a range of sources including intensive interviews. These items were factor analysed, and the six factors accounting for the greatest amount of the variance (50% in total) are listed in Table 18.1, with example items. The factor names are those suggested by De Las Heras on examination of the items contained within each factor, and should not be taken as definitive.

These factors and their associated items illustrate the way in which music listening behaviours are ranged along a particularly important dimension, that of control (related to choice and autonomy). For instance,

**Table 18.1** Principal factors influencing music use (De Las Heras, 1997)

| Factor name | % of variance | Example of item |
|---|---|---|
| Energy use | 15 | I use music to sing along to while I am tidying up the house |
| Transcendent | 12 | This music connects me with a higher self |
| Memory | 7 | I want to reminisce about times associated with earlier listening |
| Analytical | 6 | I like experiencing how the different parts of the music are put together |
| Radio | 5 | I like the radio because it is unpredictable: you don't know what will follow |
| Compensation | 5 | I play this type of music to get rid of negative moods |

the sample item shown in Table 18.1 for the 'energy use' factor requires that the music being listened to is known by the listener (otherwise how can she sing along to it?), and has certain speed and rhythm characteristics that are compatible with the bodily movement of housework. This means that the listener must make very definite choices about the precise track to accompany such activity. Similar considerations come into play in the 'memory' factor, for obvious reasons. on the other hand the 'radio' factor suggests a desire to reduce control and choice. It is the surprise of someone else's choice which is valued, although even here the listener must exercise some choice (for instance, in choosing one station rather than another). It is also interesting to note that 'analytical' hearing, which corresponds most closely to the type of structure-based listening which is assumed to be necessary for serious 'high-art' appreciation, is only one of a wide range of listening modes employed by this group of self-defined music-lovers. Although this sample of listeners was more broadly based than Neilly's, they were still self-selected as regular music listeners with substantial CD collections. Their music-listening behaviours could arguably be both more specialized and more differentiated than an unselected sample of the population.

## 18.3 The Mass-Observation project

Since 1981, Sussex University has involved around 500 British people in an ongoing recording of their everyday lives (Sheridan, 1998). Participants are recruited through media advertisements, and agree to respond by post to a set of questions mailed about three times each year. Correspondents are encouraged to write in a free and open-ended fashion, answering those questions which most engage them. All the writers are volunteers and therefore self-selected. Women outnumber men by 3 to 1, and there is a preponderance of people over 40. Table 18.2 shows the list of general topics on which panel members were asked to write in 1995–1998. Respondents are encouraged to write to their Mass-Observation contact in personal letter, like a friend. All material is confidential, and cannot be removed from the archive. The archive has thus collected a uniquely broad picture of the views and experiences of individual citizens, offering many untapped opportunities for comparisons across different domains of activity.

The Autumn 1997 mailing focused on music. Although the overall style and presentation of questions was determined by Mass-Observation staff, the present author played a major role in determining the music questions and their areas of investigation. The overarching question was 'Please could

**Table 18.2** Topics of Mass-Observation mailings in 1995–1998

| | |
|---|---|
| 1995 | Daily meals |
| | The countryside |
| | TV soap operas |
| | Shopping |
| | Images of where you live |
| | Mothers and literacy |
| 1996 | The National Lottery |
| | Beef: BSE and the 'mad cow' debate |
| | The supernatural |
| | Using the telephone |
| | Unpaid work |
| | The next general election |
| 1997 | You and the National Health Service |
| | Doing a job |
| | Being overweight |
| | The death of Princess Diana |
| | Music |
| | Dancing |
| 1998 | The garden and gardening |
| | Having an affair |

you tell us all about you and music'. All other questions were optional cues to focus correspondents on areas they might like to talk about. This paper looks in particular detail at two areas, which represent the extremes of the control/autonomy dimension identified by De Las Heras (1997): (1) personal use as chosen by the participant; and (1) non-chosen music experienced by the participant in public places. The key cues were (1) 'Do you use music in different ways? My mother used to play fast Greek music to get her going with the housework—do you have habits like this? Are they linked to particular times, places, activities or moods? For instance, you might use music in different ways at home, outdoors, or at work; in company or on your own; while you exercise, cook, study, make love, travel, or sleep; to cheer you up or calm you down'; and (2) 'Do you enjoy music in pubs? Restaurants and cafes? Supermarkets? Shops? Streets? Do you ever dislike music in public places?'. Other areas of questioning on music, not analysed in this paper, were: going to concerts; money spent on music; nature and location of collection and music playing devices in the home.

The archive holds data on the age, gender, occupation, marital status, and home address of all participants, which allowed selection of stratified sub-sets for analysis, and for the examination of possible effects of these variables on music listening. No attempt was made to examine effects of

occupation, but possible effects of all other variables were tested for. There were no significant effects of marital status or geographical location (north, midlands, south) within the UK, so no further mention is made of these variables.

The Music Directive was mailed to 460 panel members (137 male, 323 female), and replies were received from 249 panel members. This represents a 54% response rate, which is typical for this panel. Of the respondents, 69 were male (50% of the men in the panel), and 181 were female (56% of the women in the panel). For detailed analysis, it was decided to select at random an equal number (15) of men and women in each of 3 age-bands, under 40, 40–60, and over 60. This proved possible for women, providing 45 female responses in total. Unfortunately there were only three male respondents under 40, and only 13 between 40 and 60, so the male sample for analysis was 31 in total (3+13+15).

## 18.4 Chosen personal uses of music

The selected sample of respondents' comments were analysed to discover recurrent functional and contextual themes in the responses. Table 18.3 shows the categories of response obtained. Explicit functions referred to by correspondents match quite closely the types of categories found in earlier studies. The use of music as a cue to reminiscence is the single most frequent use reported.

The activities which music accompanies are predominantly domestic and solitary. Indeed some respondents made a point of qualifying their use of music (e.g. 'the car is the only place where I can listen to it loud enough without annoying other people'). Several respondents made explicit links between activities/contexts and their psychological functions (e.g. on arrival home from work 'music lifts the stress of work: it has an immediate healing effect').

There are age and gender effects on the distribution of reported activities and functions. The mean number of reported activities in the home that are accompanied by self-chosen music decreases with age (2.4 for the youngest group, 0.6 for the oldest). However, it is not clear at this stage whether this result is artefactual (it may be that older people write less). For activities outside the home there is a similar age effect, but also a significant gender effect. Men report 0.7 activities on average, whereas women report only 0.3 activities.

In terms of functions, there are no significant age or gender effects associated with mention of the memory-association functions of music, but there is a significant tendency for women to mention mood altering or

**Table 18.3** Percentage of Mass-Observation Respondents reporting various functions and activities chosen for music

| | | |
|---|---|---|
| *Functions* | | |
| Reminder of valued past event | 50 | Memory |
| Spiritual experience | 6 | Transcendent |
| Evokes visual images | 2 | Sensorial |
| Tingles/goose pimples/shivers | 10 | |
| Source of pleasure/enjoyment | 6 | |
| To put in a good mood | 16 | Mood change |
| Moves to tears/catharsis/release | 14 | |
| Excites | 2 | |
| Motivates | 2 | |
| Source of comfort/healing | 4 | |
| Calms/soothes/relaxes/relieves stress | 8 | |
| Mood enhancement | 8 | Mood enhancement |
| To match current mood | 6 | |
| *Activities* | | |
| To wake up to | 6 | |
| While having a bath | 4 | |
| While exercising | 4 | |
| To sing along to | 6 | |
| To work to (desk work) | 14 | |
| To work to (housework) | 22 | |
| On arrival home from work | 2 | |
| While having a meal | 12 | |
| Background while socializing | 4 | |
| To accompany sexual/romantic events | 4 | |
| Whilst reading | 6 | |
| In bed/to get to sleep | 14 | |
| While driving/running/cycling | 22 | |
| While on public transport (Walkman) | 4 | |

enhancing functions more frequently than men (mean of 1.0 for women, as compared to 0.2 for men).

Women in general offer more articulate and detailed descriptions of their use of music for mood-related functions than do men (e.g. 'When I'm down I listen to this and go down as far as I can, then I cry, I cry deep from inside. I wallow in self-pity and purge all the gloom from my body. Then I dry my eyes, and wash my face, do my hair, put on fresh makeup, and rejoin the world.').

In sum, people tend to mention between one and four activities in which they use music of their own choice, and between one and two functions associated with that use. In this context of free verbal response,

people are thus more able to retrieve activities than explain their functions. It is an aim of planned research to discover if more intensive and direct research methods will be able to uncover explicit or implicit functions associated with every deliberate act of music listening.

## 18.5 Music in public places

The main locations in which respondents reported experiencing unchosen music were pubs, shops (predominantly supermarkets), eating places, and streets (including covered shopping malls). There were a small number of mentions of other sources (such as on-hold telephone music, lifts, and music overheard from a neighbour's domestic premises).

Each mention of music in a public place was given a score of +1 if it indicated a liking for music in that context, and −1 if it indicated a dislike. Mean ratings for pubs and restaurants were both 0.0 (equal frequency of like and dislike). The mean rating for shops was −0.3 (more people dislike than like), and for streets was 0.1 (more people like than dislike). The mean rating of music for the separate categories of pubs, restaurants, shops, or the street did not differ according to age or gender, however if all instances of music in public places were combined, there was a significant age effect, such that the greatest dislike was expressed by the 40–60 year olds. There was also a significant gender effect, such that men were more likely than women to express dislike of music in public places.

The reasons people gave for liking or disliking music in public places were categorized into four major groups: (1) noise (e.g. intrudes, unable to hear, unable to concentrate); (2) the type of music or the quality of the performance; (3) whether the music was live or prerecorded; (4) whether the music matched or was appropriate for the atmosphere/context of its location. Again, a reason contributing to liking was given a score of +1 and a reason contributing to disliking was given a score of −1. There was no significant association of reason type to location, age, or gender, except in the case of noise, where there was an interaction of gender and age, such that around half of men at all ages cite noise-related reasons, whereas only women between 40 and 60 mention noise to the same extent. Older and younger women tend not to give this reason. Overall, noise was generally a contributory factor to disliking music in public places, whereas the other reasons were equally often positive as negative.

Table 18.4 provides representative quotes regarding each of the four main categories, giving in each case a predominantly positive, a mixed, and a predominantly negative comment.

**Table 18.4** Selected quotations from Mass-Observation participants regarding music in public places

| | |
|---|---|
| *Pubs* | 'I love music in pubs restaurants and cafes but NOT TOO LOUD because that stops me being able to talk to people, unless I've gone to a pub to specifically hear a band. I like jazz in restaurants or piano music, in pubs I like folk or the rock music of my liking.' |
| | 'I do not enjoy music in pubs. It is such a shame that the days of going to the pub for a drink and conversation has gone. The noise racket is awful.' |
| *Restaurants* | 'Music in pubs etc.—I like a bit of music and noise when I eat out. I hate silence in public places where you can hear every sound. It makes me feel very nervous.' |
| | 'I find music played in public places can often be intrusive. As a background, it can be beneficial, helping to break up the silence which can be inhibiting to conversation. All too often, though, the music is played at a level which prevents audible conversation and so has the effect of reducing conversation. And if you have the misfortune to be seated under the loudspeaker in a restaurant, it can put you off your food too.' |
| | 'I would not patronize a restaurant which piped music – not even "Egmont." Food and wine demand undivided attention.' |
| *Shops* | 'Slow music in shops and supermarkets is soothing, fast music makes you buy more. In any case its a distraction, but quite enjoyable to hear chart hits I wouldn't otherwise hear (so THAT'S what Oasis sounds like!).' |
| | 'I have discovered Waterstones Bookshop in Newcastle plays good music: the last time I was there Beethoven's "Egmont" Overture was playing, and the trouble is that I pay more attention to the music than to finding the book I want.' |
| | 'Most shops have it much too loud and I hate it. In fact I've gone out of shops because the noise stops me thinking and have gone into shops without music playing purposely.' |
| *Streets* | 'Both my husband and I have performed on the streets. My husband is still a regular busker. This to me is musical entertainment at its purest form. The joy of busking is its spontaneity. Your audience is free to come and go as it wishes, to pay or not pay, to listen or not listen. There is a beautiful freedom about busking that I love, and I hope we never lose street entertainment.' |
| | 'The music in the streets is acceptable. Quite often the musicians are quite good and anyway the noise is dispersed.' |
| | 'I also dislike the din that some cars make when they pass by, infesting the streets with their thumping noises from within. How they can drive properly with such a din in their cars, God knows. It concerns me too that they wouldn't be able to hear the sound of an ambulance |

*continued*

**Table 18.4,** *cont'd* Selected quotations from Mass-Observation participants regarding music in public places

| | |
|---|---|
| | or police car with such a noise going on inside their cars. It seems to me, rather like the fastest drivers, it is usually young men between the age of 17 and 25 who are the main culprits. They usually like to have their driving-seat window open, elbow leaning out of the car and looking macho.' |
| *General* | 'I dislike music in public places like shops, pubs, etc. Much as I like listening to music, I like it to be the music I have chosen, and when I choose to listen.' |
| | 'I positively dislike music in public places (pubs, restaurants, supermarkets, shops, streets), mainly because it's usually the sort of music that I wouldn't choose to listen to anyway, and I resent having it inflicted on me.' |
| | 'Another thing I hate is music or radio in a car—I feel a real captive audience then and don't like it.' |
| | 'I've got to listen to bad music as much as I have got to listen to good. As such I object to it. I didn't choose it, nor did anyone else—in the absence of getting immediate consensus from people of what they would like to hear—lets have none of it.' |
| | 'But just as I don't want to intrude on others, I wish others wouldn't intrude on me.' |
| | 'I heard that a mall in America solved the problem of rowdy teenagers hanging around all day by playing Mantovani all day long. This was such an assault on the teenagers' sense of 'cool' that they found other places to go. Must be powerful stuff.' |

Much of the data suggests an ambivalence towards music in public places. Music in pubs and restaurants can be welcome if it is seen as enhancing the atmosphere, but it easily becomes disliked if the style is not to the respondent's taste, or it is too loud. Much music makes conversation difficult, and can distract people from their intentions. The area of most unequivocal liking appears to be live street-music provided by buskers. There is a strong hint that lack of control is at the root of much dislike of music in public. Where the music is pre-recorded, often emerging from hidden speakers, the sense of lack of control is heightened. Live music, with its possibility of interacting with the performer, offers a greater sense of control and participation. It would be consistent with the centrality of the control/autonomy dimension that groups most used to being able to exercise control and choice (men between 40 and 60, at the height of their earning power and associated social status) would have the most negative reactions to music in public, as these data show. The final section of Table 18.4 (General) highlights the overarching presence of control as a theme in respondents' comments.

A final concern of this paper was to explore possible relationships between chosen personal uses of music and attitudes to music in public places. Table 18.5 shows the outcomes of selected correlations on the quantitative indices reported above. Respondents who reported more self-chosen activities to music outside the home reported greater levels of liking of music in pubs.

Respondents who reported more mood enhancing or altering functions for self-chosen music reported greater levels of liking of street music, and were more likely to cite its live nature as a reason for liking it. These relationships, demonstrated in free-response data, offer a promising indication that patterns of music use in high-control contexts have implications for attitudes towards music in low-control contexts. One would expect more substantial relationships to emerge in research designs where such relationships were explicitly probed.

## 18.6 Discussion and conclusions

The data presented here confirm that listening to music (perhaps unlike performing it) is used by most people as a means of enhancing everyday activities that are mainly solitary and individualistic. The valued outcomes are mainly self-referring, internal, even solipsistic. Music seems to be a ready source of conflict between people as much as it is something which draws them together. This contrasts rather markedly with the participatory and social cohesion functions identified by ethnomusicologists as dominant forces in many non-western societies, where public music is the medium

**Table 18.5** Correlations between chosen activities/functions and attitudes to music in public places

| | Activities | Functions | | |
|---|---|---|---|---|
| | Home | Outside Home | Memory | Mood |
| *Dislike* | | | | |
| Pub | –0.13 | –0.25* | –0.05 | –0.21 |
| Restaurant | 0.01 | 0.03 | –0.08 | –0.14 |
| Shop | –0.11 | –0.03 | –0.02 | –0.08 |
| Street | –0.04 | –0.03 | –0.15 | –0.28* |
| *Reason* | | | | |
| Noise | 0.11 | 0.08 | 0.01 | –0.08 |
| Type | 0.05 | 0.06 | 0.00 | –0.06 |
| Live | 0.05 | –0.03 | –0.14 | –0.23* |
| Atmosphere | 0.01 | –0.12 | –0.12 | –0.08 |

*$p < 0.05$.

through which people share experience and relate to one another and the social group as a whole (e.g. Blacking, 1976). Indeed, one sees in the responses an explicit contrast between the private and personal work (often of considerable importance) which music allows an individual to accomplish, and the rest of their life which is enacted in the public arena (e.g. 'When I'm down I listen to this and go down as far as I can, then I cry, I cry deep from inside. I wallow in self-pity and purge all the gloom from my body. Then I dry my eyes, and wash my face, do my hair, put on fresh makeup, and rejoin the world.').

If music is a powerful source of reliving important memories then, unless such memories are shared, imposed music can alienate rather than unite. Maybe music can only perform the social cohesion functions noticed in smaller societies where the body of music doesn't change too much over generations and involves everyone (true folk music). Although music continues to play an important part in public life, the data from this study reinforce the caution that is needed in drawing strong inferences about emotional and other forms of significance of music from studies which present subjects with music not of their choosing in public or semi-public contexts.

Future research cannot, of course, rely on the free reports of a self-selected, and therefore unrepresentative sample of the population. A range of more focused methodologies, tracking the everyday use of music as it happens, and probing far more explicitly into its perceived functions and psychological outcomes, is the necessary next step in making progress in this area.

## References

Behne, K. E. (1997) The development of 'Musikerleben' in adolescence: how and why young people listen to music. In I. Deliège and J. A. Sloboda (Eds.) *Perception and Cognition of Music.* Hove: Psychology Press.

Blacking, J. (1976) *How Musical is Man?* London: Faber.

Bourdieu, P. (1984) *Distinction* (tr. R. Nice). London: Routledge.

De Las Heras, V. (1997) What does music collecting add to our knowledge of the functions and uses of music? Unpublished MSc dissertation, Department of Psychology, Keele University.

du Gay, P., Hall, S., Janes, L., Mackay, H., and Negus, K. (1997) *Doing Cultural Studies: the Story of the Sony Walkman.* London: Sage.

Finnegan, R. (1989) *The hidden musicians: music-making in an English town.* Cambridge: Cambridge University Press.

Gabrielsson, A. and Juslin, P. (1996) Emotional expression in music performance: between the performer's intention and the listener's experience. *Psychology of Music,* **24**, 68–91.

Gabrielsson, A. and Lindstrom, S. (1996) Can strong experiences of music have therapeutic implications? In R. Steinberg (Ed.) *Music and the mind machine: the psychophysiology and psychopathology of the sense of music.* Berlin: Springer.

Gregory, A. H. (1997) The roles of music in society. In D. J. Hargreaves and A. C. North (Eds.) *The social psychology of music.* Oxford: Oxford University Press.

Hevner, K. (1936) Experimental studies of the elements of expression in music. *American Journal of Psychology*, **48**, 248–268.

Neilly, L. (1995) The uses of music in people's everyday lives. Unpublished undergraduate dissertation, Department of Psychology, Keele University.

North, A. C. and Hargreaves, D. J. (1996) Responses to music in aerobic exercise and yogic relaxation classes. *British Journal of Psychology*, **87**, 535–547.

North, A. C. and Hargreaves, D. J. (1997) Experimental aesthetics and everyday music listening. In D. J. Hargreaves and A. C. North (Eds.) *The social psychology of music.* Oxford: Oxford University Press.

Scherer, K. R. and Oshinsky, J. S. (1977) Cue utilisation in emotion attribution from auditory stimuli. *Motivation and Emotion*, **1.4**, 331–346.

Sheridan, D. (1998) Mass-Observation revived: the Thatcher years and after. In D. Sheridan, D. Bloome and B. Street (Eds.) *Writing ourselves: literacy practices and the Mass-Observation project.* Cresskill, NJ: Hampton Press.

Sloboda, J. A. (1989) Music psychology and the composer. In S. Nielzen and O. Olsson (Eds.) *Structure and perception of electroacoustic sound and music.* Amsterdam: Elsevier.

Sloboda, J. A. (1992) Empirical studies of emotional response to music. In M. Riess Jones and S. Holleran (Eds.) *Cognitive bases of musical communication.* Washington, DC: American Psychological Association.

Zillman, D. and Gan, S. (1997) Musical taste in adolescence. In D. J. Hargreaves and A. C. North (Eds.) *The social psychology of music.* Oxford: Oxford University Press.

CHAPTER 19

# MUSIC: WHERE COGNITION AND EMOTION MEET

Music presents a puzzle. On the one hand, people love music and devote much time and effort to putting themselves in the way of it. On the other hand the levels of musical skill achieved by the vast majority of people in contemporary Western society are surprisingly low. On the face of it, music has all the characteristics which would lead one to predict that many people should be highly skilled at it. A long tradition of research into skill and its acquisition (cf. Ericsson and Smith, 1991; Ericssson, 1996) suggests that *structure* plus *motivation* plus *practice* leads to *skill.*

In reviewing these four aspects I will first provide evidence that most music has the kind of structure which is easily learned and understood by the human mind. Second, I will review evidence relating to motivation, and the very high value which many individuals place on their engagement with music. Third, I will review evidence of the intimate link between level and nature of practice activities and achievement in music. Fourth, and more speculatively, I will outline some potential inhibitors to musical achievement, in an attempt to formulate hypotheses about why, given such an apparently propitious set of circumstances, population outcomes in the area of music achievement are so dire. These inhibitors, whilst having individual psychological consequences, are rooted in wider social and cultural phenomena.

## 19.1 Naturally detected structures

The parallels between music and language are very great. Just as it has been shown that very young infants have already picked up certain structural regularities of spoken language, so similar research shows they know something about the structure of music. Sloboda (1985) has showed that children as young as 7 can consciously choose between well-formed and ill-formed musical sequences. Asked which of two teddies played the 'right' tune, and which 'made a mistake', children demonstrated knowledge of

both simultaneous and sequential constraints that exist within Western tonal music.

Structural sensitivity has also been demonstrated by untrained adults in a wide variety of tasks. For instance, Sloboda and Parker (1985) asked psychology students to repeatedly listen to a short folk melody and attempt to reproduce it after each trial by singing it. Although almost no participant gave note-perfect recall, recalls showed structural sensitivity. For instance, recalls tended to preserve the metre and phrase structure of the original.

These and many other studies have established that music is represented in the mind in terms of the structures and regularities that it contains. When such structures cannot be detected, it is impossible for people to process and store the information efficiently. The music of a culture has a familiar syntax, and when this is missing, as for example in some forms of atonal music, processing is severely disrupted. Music is just like chess and the many other cognitive skills that psychology has studied. If a person cannot discover the structure, they cannot become an expert. The research literature on music shows that almost everyone in our culture has 'found the structure of music', through exposure to music irrespective of formal training. Therefore a prime condition for the acquisition of expertise has been fulfilled by the vast majority of our population.

## 19.2 How and why do people value music?

The proportion of people who can gain material advantage from being skilled at music must be small. Why is music attractive to people? Why are they motivated to listen to it, perform it, create it? Does it serve any useful purpose?

Many types of research have demonstrated that music can and does have important psychological benefits. Not only does music engagement seem inherently pleasurable, but it is often used for essentially therapeutic purposes. Strong and valued emotions seem to be at the core of music engagement (Sloboda, 1992).

There are a number of laboratory and clinical demonstrations of the mood altering powers of music, but more recently an interest has developed in real life uses of music, outside the laboratory or the clinic. For instance, Sloboda (1990) obtained 113 accounts of musical life events from 70 adults of varying musical background and experience. The question asked was 'do you have memories of any specific incident from the first 10 years of life which involved music in any way'. Many of the memories so elicited were deeply memorable, and often motivated lifelong enthusiasm for, and involvement with, music. The contexts in which these highly valued

**Example 19.1** Excerpt from the response of a participant in the Sloboda (1990) study

'I was sitting in morning assembly at school. The music formed part of the assembly service. The music was a clarinet duet, probably by Mozart. I was astounded at the beauty of the sound. It was liquid, resonant, vibrant. It seemed to send tingles through me. I felt as if it was a significant moment. Listening to this music led me to learning to play first the recorder and then to achieve my ambition of playing the clarinet . . .'

experiences occurred were somewhat restricted, and this provides a very important clue for framing an answer to the puzzle with which this paper opened. Hardly any of these experiences occurred during music lessons or in the presence of a teacher. They occurred in time 'off task', alone or with friends, at home or at school (See Example 19.1)

Currently, a research team at Keele is gathering data on music use rather closer to the coal face (Sloboda *et al.*, 2001). Adults without musical training have been carrying pagers around with them for an entire week. The pager calls them at random intervals throughout the day. Every time they are paged they are asked to stop what they are doing as soon as they can safely do so, and complete a brief questionnaire in a booklet of identical sheets that they are also asked to keep with them at all times. They are asked to focus on any music that was occurring as the pager sounded, or, if no music was occurring, to recall the most recent occurrence of music since the previous paging.

What we have found so far is that music is experienced in an incredibly wide variety of circumstances. Hardly any of these have music as the main focus. The music is accompanying some other activity (such as washing up, socializing, exercising, working, travelling). Two sets of data are particularly pertinent to the current discussion.

First, participants were asked to estimate their mood on a number of bipolar scales both before and after the music. These scales factored into three major mood dimensions which we have labelled, in order of contribution to the variance, positivity, present-mindedness, and arousal (see Table 19.1). In general music increased emotional state towards greater positivity (e.g. more happy), greater arousal (e.g. more alert), and greater present-mindedness (e.g. less bored). But what seemed of particular significance to us is that mood change was greatest when participants exercised choice over the music they were hearing. Music maximally enhances well-being when participants exercise some degree of autonomy and self determination in the type of music they hear (see Table 19.2).

The second result of interest is the dissociation of emotional factors in some music-listening cases. While it is true overall that music makes people more positive, more aroused and more present-minded, a more

**Table 19.1** Factor analysis of bipolar mood scales in pager study

| | |
|---|---|
| *Positivity (36%)* | |
| Distressed | Comforted |
| Sad | Happy |
| Irritable | Generous |
| Insecure | Secure |
| Tense | Relaxed |
| *Present-mindedness (14%)* | |
| Bored | Interested |
| Detached | Involved |
| Lonely | Connected |
| Nostalgic | In-the-present |
| *Arousal (12%)* | |
| Drowsy | Alert |
| Tired | Energetic |

differentiated pattern was evident in about 12% of the episodes. In some episodes, for example, positivity increased, while the other two dimensions decreased. In others, arousal increased, while nothing else did, and so on. Each of these patterns was associated with a particular set of circumstances and psychological purposes. People deliberately use music for different intended outcomes. Some examples of these episodes are given in Table 19.3.

From results such as these, a quite detailed picture is beginning to build up of *what* music does for people. But a much more intriguing scientific puzzle is to explain *how* music can 'mean so much' to people.

## 19.3 How are emotions mediated through music?

A range of research studies suggests that emotions are mediated through music in at least three quite distinct ways, *episodic associations, iconic associations,* and *structural expectancies.*

**Table 19.2** Change in mood as a function of degree of musical choice in pager study

| | Degree of choice | | |
|---|---|---|---|
| | **Low** | **Medium** | **High** |
| Change in: | | | |
| Positivity | 0.8 | 0.3 | 2.3 |
| Present | 0.6 | 3.0 | 3.2 |
| Arousal | 0.8 | 1.2 | 1.8 |
| Average change | 0.8 | 1.5 | 2.4 |

**Table 19.3** Examples of mood change dissociation in pager study

| | |
|---|---|
| Mood change pattern | Positivity up, present-mindedness down, arousal down |
| Everyday descriptor | 'Chilling out' |
| Example | At home, relaxing with 6 friends and acquaintances. Wanted to do it. Ambient music on CD. Little choice over music. 'The music was very tranquil and relaxing. I was very, very tired.' |
| Mood change pattern | Positivity down, present-mindedness down, arousal up |
| Everyday descriptor | 'nostalgic wallowing' |
| Example | At home, alone, washing up. Rock on radio. High choice over music. 'Favourite song I had not heard for some time. It brought back certain memories.' |

Episodic associations are the type of effects which can be explained by what John Booth Davies (1978) has memorably described as the 'Darling They're Playing Our Tune' theory of emotion. Music can provide a powerful reminder of earlier events or periods in our lives, and the significant people or places that figured in them, particularly when these life events were strongly emotional. Strong and real as these associations are, they are not particularly interesting from a theoretical point of view, because they are entirely driven by idiosyncratic autobiographical contingencies.

Iconic associations are brought about by physical characteristics of the music that mimic or resemble the sound effects that could be created by non-musical events. Crude examples would be the mimicking of birdsong or of natural phenomena such as thunderstorms. More interesting are the ways in which music can suggest a particular kind of emotional character, by creating sounds typical of that emotion. Watt and Ash (1998), for instance, have recently developed the analogy, previously developed for visual materials by Michotte, of musical objects as 'virtual persons'.

Research at Keele has been more concerned with the third source of emotional engagement, that seems to come about by tracking the unfolding structures of a piece of music, and reacting to the confirmations and violations of expectancy that are created within those structures. Waterman (1996) showed that there are particular 'hot spots' in many pieces of music where people are prone to experience emotions particularly intensely. These can be tracked by asking people to report on 'thrills' (Sloboda, 1991). Thrills are reliable physical concomitants of emotional response, which include shivers down the spine, tears, or a lump in the throat.

It has been found that music hot spots usually involve particular structural events which tease structural expectancies, by repeatedly creating and resolving tensions, or by manipulating timing parameters that cause expected

events to happen earlier or later than expected. Emotional response to music is thus an integral outcome of the intuitive structural analysis that goes on while listening. This is an important finding for a long-running debate in musicology. Some musicologists have tried to argue that 'pure' music listening means stripping away all mundane associations from music and hearing it as pure sound. Even if that were possible and desirable, our findings suggest that music would still be an intensely emotional experience. It is interesting and pertinent to note that in many traditional academic music contexts the official discourse is such that emotions don't get a look in. Everything is to do with form, content, history, and analysis. Perhaps this is the price that musicians felt they had to pay to have music accepted as a 'proper' discipline. This could be one more clue to the puzzle that motivates this lecture. You can't get fully inside music without becoming emotionally involved, yet such emotional involvement is exactly what many of the traditional institutions of music education have tended to inhibit and discount.

In sum, there is good evidence that people are strongly motivated to engage with music because of the valued psychological outcomes. These outcomes can be traced in part to the powerful emotions engendered when we listen to music, emotions which are enhanced by the structural expectations that we acquire within a musical genre or culture.

## 19.4 Acquiring musical expertise

Some recent research has been concerned with the processes by which a transition is made from the basic level of receptive competence, which is shared by all listeners within a culture, to the status of expert. There are, of course, many types of expertise, but much recent research has focused on a particular and relatively widespread form of expertise, the ability to perform notated music on a traditional acoustic instrument, such as piano or violin. This is a form of expertise which has been particularly widely promoted within the educational system. Almost every school has a recorder or wind band, or an orchestra, and most qualified music teachers, whether private or public, work within this tradition. Much of the research in this area is concerned to show the intimate link between achievement and focused deliberate cognitive engagement of the sort which is generally called practice.

Sloboda *et al.* (1985) made an extensive single-case study of an autistic musical savant, who was reported to be able to memorize long piano pieces just by hearing them over a few times. His quite phenomenal ability turned out not only to be specific to music (his verbal memory span was subnormal)

but also specific to tonal music. Given a piece of mildly atonal music to memorize, his ability fell apart. We found biographical documentation which suggested that his ability to memorize grew gradually over 15 years of obsessive and repetitive work at this single task, to which he devoted almost all his waking hours.

Sloboda *et al.* (1996) probed systematically into the musical histories of 250 young people learning musical instruments. The data showed a strong relationship between objectively measured level of skill and the amount of practice undertaken, in any give year, and also cumulatively over the life span. The highest achieving group were undertaking on average 800% more daily practice than the lowest achieving group at the age of 12. Because differences in practice time were evident at considerably younger ages, we found similarly large disparities in total hours of accumulated practice time by early adolescence (see Table 19.4).

More recently, O'Neill (1997) tracked beginning instrumental learners over the first year of music lessons. This allowed her to administer a battery of measures before any exposure to the instrument at all. These included measures of IQ, standardized musical aptitude tests, and a motivational measure of resistance to failure (an adapted version of the Wisconsin Card Sorting Task). This is a particularly pertinent measure for early musical learning, where failure to be able to reproduce the fluent musical products they hear all around them is a constant experience of beginners. After 9 months all children were given the same piece to learn for two weeks, and their performance on this piece was video-recorded for evaluation by expert music examiners. There were large disparities between the children in the standard of these performances. Neither starting IQ nor musical aptitude predicted outcomes, but both the pre-instruction motivational variable and amount of practice on the piece were strong predictors.

There is a considerable amount yet to be discovered about what types of practice lead to better and faster learning outcomes. It is naive to suppose that an hour of one type of engagement is as good as any other, and some

**Table 19.4** Key practice indicators from Sloboda *et al.* (1996)

| | | | | |
|---|---|---|---|---|
| *Average minutes of daily music instrument practice at age 12 as a function of skill level* | | | | |
| Lowest achievement | | | Highest achievement | |
| 1 | 2 | 3 | 4 | 5 |
| 15 | 30 | 60 | 60 | 120 |
| *Average hours of accumulated practice at age 13 as a function of skill level* | | | | |
| Lowest achievement | | | Highest achievement | |
| 1 | 2 | 3 | 4 | 5 |
| 450 | 800 | 1400 | 1400 | 2500 |

preliminary evidence on efficiency of different practice strategies is now being collected. But none of this negates the main conclusion of this body of work, which is that in order to attain high levels of performance expertise within the classical performance tradition, large amounts of deliberate practice are required.

## 19.5 Where have all the musicians gone?

This more speculative section attempts to describe and explain some of the barriers to the achievement of musical expertise that appear to exist in our society.

The first barrier to achievement is a significant reduction (in comparison to many earlier periods) of the societal scaffolding which allowed people to progressively occupy intermediate rungs on a ladder of skill progression. There are decreasing numbers of widespread social institutions where moderate levels of performance skill are encouraged and celebrated. In an earlier age, one could expect to hear and join in music sung and played in the home. There would be sing-songs at the local pub, or at village festivals, where all and sundry could join it, at their own level. Playing along on a tin whistle or a violin would be tolerated, even encouraged. At a slightly more formal level, people might receive a structured learning environment within a church choir or a brass band. Here, there would be a level of discipline and correction of blatant errors, together with a regular cycle of rehearsals and concerts. These institutions could allow a gradual progression in skill and accomplishment, so that someone might move from the back desk to a soloist position. From this backdrop, the few might indeed move on to professional activity. But significant opportunities were there for many to reach some significant intermediate level of achievement. In times of mass conscription, even the armed forces provided significant opportunities for regular formal and informal music making.

The present society has seen the rungs in this ladder of progression gradually rot and fall out. The decline in church attendance and in the cohesion of local communities, has caused these institutions to wither, and they are not being replaced. Thus there is a widening gap between everyday contexts in which people operate as novices and those in which they can come to operate as experts. In general, what Simon Frith (1996) has called 'the academy' has come to appropriate the gateways which provide cultural scaffolding for moving up the ladder of achievement. The academy is constituted by 'the music departments, conservatories, and the whole panoply of formal arrangements and practices in which classical music in its various forms is taught and handed down the generations'.

It is significant that in our large-scale study of young people's musical achievement (Davidson *et al.*, 1996) the family environment of high achievers was quite unusual. Parents devoted great amounts of time and energy to the practical support of their children's learning. They often hired private tutors, they supervised practice sessions on a daily basis, they encouraged, cajoled, and sometimes even fought with their children when the children felt like watching TV rather than practicing. They have fought for their children's access to the academy.

I suggest that such unusual family environments are necessary these days to replace the scaffolding that local communities no longer provide. Most families are neither able to nor see the point in making these arrangements for their children. Arguably, the ubiquity of recorded music also inhibits practical music-making in the home and elsewhere.

The second barrier to achievement is the increasing framing of official discourse about music performance in terms of concepts of talent, achievement and success, rather than in terms of community, fulfilment, or transcendence. The National Curriculum for music is only the most recent manifestation of this, where attainment targets are more salient than any notion of why it might be interesting or personally relevant to achieve these targets. This reflects the very ambivalent attitude towards art that typifies many modern consumer societies. Art is often seen as of no value except as a commodity to be purchased by consumers for 'entertainment' in exchange for hard cash. Therefore musical expertise is only valued to the extent that it can earn money for the purveyors of entertainment, be they popular music recording companies or opera houses. What then matters most is to be better, more skilful, more innovative, more 'professional' than ones' peers. Hard work is taken for granted, but on top of that, only those with that special indefinable extra quality 'talent' are, at the end of the day, going to be able to command the attention which will earn their sponsors the kinds of profits that they seek. And so, the impossibly polished outputs of musical superstars are rubbed in the faces (or more precisely, the ears) of young people through constant media exposure. Young music learners are pitted against each other, in exams, competitions, festivals, with the aim of weeding out all but the 'really talented'. Even at the highest levels of training, in the conservatories and music colleges, where everyone is way above the average level of achievement, to come second in a competition is seen by many as having failed. In this context it is unsurprising that young people are discouraged from participation in an activity where there are so few winners and so many losers.

The notion that music could be engaged in purely for personal fulfilment, for the building up of community and friendship, for the sheer joy

of making beautiful sounds together, is a strange, and almost reprehensible concept in many people's minds. Music is the poor relation in many schools—what has it got to do, after all, with the real business of equipping people to contribute to wealth creation? The message is projected that if you haven't got talent, you should stop wasting your time messing about with music, and concentrate on your maths or business studies.

The third and final barrier I want to postulate is the barrier of elitism or high art. The academy in most of its manifestations promotes the classical performance tradition as the paradigm and paragon of what music 'really' is, and what it is to be a musician. Limited nods in the direction of jazz and an assortment of exotica labelled as 'world musics' hardly modify the message that the academy sends to most people. The traditions and forms of the academy are, despite what some apologists claim, inaccessible to most people. Their inaccessibility is of two sorts. First, they do not reflect to most people the values and identities which they bring to music. Music of the academy is seen to be about class privilege and maintenance of a cultural *status quo* in which an elite minority dictate to the majority what constitutes good music. Second, the core exemplars of the forms of the academy (such as concertos and symphonies) demand such a level of individual and corporate proficiency and resource to execute that they have almost no points of contact with the levels of music-making that still survive in our culture (e.g. the karaoke bar). If elite music has neither resonances with one's own cultural identity, nor appears to be potentially learnable, then it hardly encourages mass participation.

It is, of course, encouraging when symphony orchestras go into schools, or when sympathetic school music teachers take an interest in students' love of pop. But if the academy, whether represented in schools or symphony orchestras, carries within its very structures and discourses the seeds of the problems I have identified, then the solution will not lie in minimal reforms within the academy but in the creation of new cultural and folk institutions. We need living and socially relevant forms to replace the church choir and the village brass band. Recent pleas by major figures in classical music for the government to reverse cuts in school instrumental provision may have come too late, if the social institutions which support music-making outside the academy are no longer there.

I have no idea what these social institutions might be, what they might build on, nor how they can be encouraged. Psychology cannot provide the answer to such questions. All it can do is indicate some of the conditions which must be met if these institutions are to enable the flourishing of individual development. The evidence reviewed above indicates to me

that performance potential could be unlocked in millions of people if we could recreate social institutions which focused on musical enjoyment, and personal and communal fulfilment, rather than on the need to be best, or to meet the taxing performance requirements of a professional elite.

## References

Davidson, J. W., Howe, M. J. A., Moore, D. G., and Sloboda, J. A. (1996) The role of parental influences in the development of musical ability. *British Journal of Developmental Psychology*, **14**, 399–412.

Davies, J. B. (1978) *The psychology of music.* London: Hutchinson.

Ericsson, K. A. (Ed.) (1996) *The road to excellence: the acquisition of expert performance in the arts and sciences, sports and games.* Mahwah, NJ, Lawrence Erlbaum Associates.

Ericsson, K. A. and Smith, J (Eds.) (1991) *Towards a general theory of expertise: prospects and limits.* New York: Cambridge University Press.

Frith, S. (1996) *Performing rites: on the value of popular music.* Oxford: Oxford University Press.

O'Neill, S. A. (1997) The role of practice in early musical performance achievement. In H. Jorgensen and A. C. Lehmann (Eds.) *Does practice make perfect? current theory and research on instrumental music practice.* Oslo: Norges Musikhogskole, pp. 53–70.

Sloboda, J. A. (1985) *The musical mind: the cognitive psychology of music.* Oxford: Oxford University Press.

Sloboda, J. A. (1990) Music as a language. In F. Wilson and F. Roehmann (Eds.) *Music and child development.* St Louis, Miss: MMB Inc.

Sloboda, J. A. (1991) Music structure and emotional response: some empirical findings. *Psychology of Music*, **19**, 110–120.

Sloboda, J. A. (1992) Empirical studies of emotional response to music. In M. Riess-Jones and S. Holleran (Eds.) *Cognitive bases of musical communication.* Washington, DC: American Psychological Association.

Sloboda, J. A. and Parker, D. H. H. (1985) Immediate recall of melodies. In P. Howell, I. Cross and R. West (Eds.) *Musical structure and cognition.* London: Academic Press.

Sloboda, J. A., Hermelin, B., and O'Connor, N. (1985) An exceptional musical memory. *Music Perception*, **3**, 155–170.

Sloboda, J. A., Davidson, J. W., Howe, M. J. A., and Moore, D. G. (1996) The role of practice in the development of expert musical performance. *British Journal of Psychology*, **87**, 287–309.

Sloboda, J. A., O'Neill, S. A., and Ivaldi, A. (2001) Functions of music in everyday life: an exploratory study using the Experience Sampling Methodology. *Musicae Scientiae*, **5**(1), 9–32.

Waterman, M. (1996) Emotional responses to music: implicit and explicit effects in listeners. *Psychology of Music*, **24**, 52–67.

Watt, R. and Ash, R. (1998) A psychological investigation of meaning in music. *Musicae Scientiae*, **2**, 33–54.

CHAPTER 20

# MUSIC AND WORSHIP: A PSYCHOLOGIST'S PERSPECTIVE

## 20.1 Introduction

The role of music in worship is not a topic that has greatly exercised psychologists. Both the psychology of music and the psychology of religion are very much 'fringe' topics in contemporary psychology. The number of psychologists interested in both must be tiny. Accordingly, the most a paper like this can hope to do is ask some of the right questions, and speculate about possible directions in which answers may be sought.

It seems that the particular sequence of questions that a psychologist might hope to elucidate would be something like the following:

1 What goes on in people's minds when they are engaging with music? Or, more formally, what are the mental processes specific to the domain of music?

2 What goes on in people's minds when they are engaging in worship? What are the mental processes specific to the domain of worship?

3 In what way do the mental processes involved in music and in worship overlap? In other words, are there particular aspects of worshiping which are enhanced, reinforced, or afforded, by acts of 'musicking'.

4 How can we account for, or explain, these overlapping mental processes in terms of our broader understanding of human mental functioning?

These are, of course, weighty questions, requiring the input of musicians, theologians, liturgists, philosophers, as well as psychologists. This paper can hardly do more that make some preliminary comments in the direction of these questions.

In what follows I have tried to take a phenomenological approach to worship. That is to say, I concentrate on what I understand people to be doing or feeling in worship, and try to describe this in theologically neutral language.

Where it has seemed necessary, for ease of expression, to use terms like 'God', I am neither claiming to be theologically precise nor attempting to use any particular set of dogmatic assumptions, whether Christian or otherwise, to support the arguments.

## 20.2 The basis of human response to music: free or constrained?

Can we choose what mental processes to engage during some musical activity? If we cannot, then there is the possibility of a rather simple approach to the psychology of music. Each type of music will have a particular mental effect on a person, just as a particular drug will have a reliable effect on human physiology. In an earlier article I have called this model of musical response the *pharmaceutical* one (Sloboda, 1989).

It should not take too much reflection to reject the pharmaceutical model as hopelessly inadequate. When confronted with the same musical situation I have many options that I can exercise. Here are just a few:

1 I can pay no attention to the music at all, and pursue my own thoughts, memories, and fantasies.

2 I can derive personal associations to the music. When have I heard this music before? How was I feeling at the time I experienced this music? What personal events, feelings, reactions am I reminded of by this music?

3 I can exercise an analytic or critical faculty with respect to some aspect of the music. Is the soprano soloist singing in tune? What was that interesting harmony? Could this be Mozart? Is this a better performance of this piece than the one I heard last week?

4 I can engage with the music in a non-analytic, contemplative mode, in which, while being attentive to it, I am not focusing on any particular event, or making specific judgements. I am 'letting the music wash over me'.

5 I can engage with the music by attending to the relationship between myself and others also engaged in the music. Am I following the conductor? Does my voice blend with that of my neighbour? Is my neighbour experiencing what I am experiencing?

This list, which is not necessarily complete, shows that we have choices about the mental processes we engage in a musical context. Is there,

in that case, anything we can usefully say about such processes that has any generality, or is it simply a case of 'you do your thing and I'll do mine'?

I believe we can go beyond total individualism by use of the notion of *affordances*. The term 'affordance' was introduced to psychology by the perceptual psychologist James J. Gibson (1996). In Gibson's eyes, human response can often be explained, not so much by reference to inner states, but by reference to characteristics of the environment which constrain and direct that response. So, for instance, a chair *affords* sitting. It has the necessary physical characteristics to allow someone to sit on it. If well designed it almost, by its very nature, *invites* you to sit on it. No one has to explain to you that it is for sitting. Even a very small child with no understanding of the concept of a chair would still naturally end up sitting on it in the course of physical exploration. This is because the object fits to a set of human characteristics. Chairs have to be a particular way because humans are the way they are.

By affording some kinds of activities, environments discourage or even preclude others. It is rarely that the same object will afford both sitting and eating, for instance. On the other hand, most objects have multiple affordances. The same object that affords sitting can also afford throwing (you could pick up a light chair and throw it at an assailant).

Designers of everyday objects are beginning to understand how they can make life simpler by building in the right sorts of affordances (Norman, 1988). No one should ever have to struggle with a door, not sure whether to push or pull it, not sure whether the handle goes up or down. Appropriately designed handles will afford only a single movement, the correct one. A correctly designed concourse will naturally draw people in the intended direction.

When objects are used in social contexts, culture can shape the normative affordances of those objects, and these can change over time. Objects (e.g. Roman vessels) which were once used for cooking are now used as collectors' items, or objects of historical research. But there is a cultural inertia. Uses of objects change relatively slowly.

And so it is with music. We can choose what to do with music, but for most of us, the choices are limited by culture, background, social context, and the way the music itself has been designed (in Gibsonian terminology, the affordances it possesses). A key context for music has always been worship. It must be the case that (at least some) music affords worship. Our task is to demonstrate how.

Rather than attempt a formal definition of worship, I shall assume for the moment a shared intuitive understanding of the concept, and sweep back over the five 'ways of responding' to music that I outlined above.

## 20.3 Music as unattended background

It is tempting to say that if one is not attending to music then it cannot be said to be affording anything, let alone worship. But there are different ways of not attending to music. On the one hand, one can simply ignore the music and get on with what one was doing as if it wasn't there. This tends to be my response to subdued canned music in supermarkets, restaurants, and other public places. I can also be at work, deeply involved with a piece of writing, and be almost totally unaware of concurrent recorded music. On the other hand, in some contexts the mere presence of the music may limit my responses and direct them in certain ways. For instance, music in a worship setting may define a period of time in which participants neither speak nor move about. So, although I may not be listening to the anthem, I am, perforce, sitting still, in the presence of others, with the memories of particular acts or words that have preceded the music. Worship may certainly happen in such contexts. But if I am truly ignoring the music, then the most that the music can be contributing to my worship is to signal a beginning and end to the time period within which particular sorts of mental activity may take place. Such signals can be provided in many ways, and are not in any sense special to music.

## 20.4 Music as a source of personal associations

There has now accumulated a significant body of data which confirms that music is a particularly powerful evoker of other times and other places. It is no accident that almost every single guest who has ever appeared on *Desert Island Discs* has been able to use music to appropriately illustrate his or her life story. The majority of Western adults seem to remember particularly well the popular tunes of their adolescent years, and retain a strong emotional response to them. In one of the earliest contemporary texts on the psychology of music (Davies, 1978), this aspect of music psychology was wittily characterized as the 'Darling, They're Playing Our Tune' theory of musical response.

There is considerable psychological evidence to show that we remember highly emotionally charged events in a unique way. Brown and Kulik (1977) coined the term 'flashbulb memory' to characterize this type of memory, which seems to be characterized by a vivid literal memory of all the incidental details surrounding the emotion-provoking incident. In the earliest study on flashbulb memories individuals were asked whether they could remember exactly what they were doing when they first heard the news that President J. F. Kennedy had been assassinated. Most people

recalled exactly where they were, what time it was, who they were with, what was said. If music was playing at the time, then the music would also be remembered.

The situation is complicated for music, however, in that the past musical event may, in itself, have been a source of strong emotion. I will discuss the nature of such strong emotion and the evidence for it in a later section. Sufficient for now to say that the association of a present musical experience may be self-referring; the present experience retrieves the thoughts and feelings of an earlier experience of that music, rather than of the other events which accompanied the music. Such experiences are often positive, sometimes profoundly so.

The associationistic function of music may be tight or loose. A tight association is the association generated by a specific piece of music (e.g. Bach's *Prelude and Fugue in G* for organ). A loose association is to a type or genre of music (e.g. baroque organ music). Many people operate with loose associations which determine their willingness to engage with music in certain circumstances. Individuals whose strong personal associations with popular music are to situations of alienation and degeneracy are unlikely to be spurred to worship by such music. Such associations can be self-reinforcing. Because many people in our culture associate organ music with worship settings (as a result of experiencing many previous worship settings involving organ music) but do not so associate pop music, organ music can in itself appear 'sacred' and pop music 'profane', leading to cultural practices which reinforce the role of some types of music in worship, and exclude others. Appearances generated through such associations alone are, however, arbitrary and culturally contingent. The issue of whether some kinds of music have intrinsic features which better afford worship than others is a separate question, and one to which we will return.

The consequences of musical associations for worship are, I think, rather clear. Music has a particular power to remind us of important past events in our life, and the profound feelings which accompanied them. This can lead to celebration, joyful remembrance, or the revisiting of pain and suffering. It can lift us out of our current preoccupations and re-focus our attention on the wider landscape of our lives. By bringing together in our consciousness events which may be distant in time, it can help us to get a better sense of the cradle-to-grave continuity and unity of ourselves as persons.

Music is by no means the only source of powerful reminders of times past. But because of its inherent capacity to evoke strong emotion, it is likely to be a particularly powerful route to personal associations.

## 20.5 Music as a vehicle for the exercise of judgement

The preceding section glossed over a crucial problem, the problem of representation and equivalence. How does a person recognize a piece of music experienced in the present as an instance of something they have heard before in the past? It can rarely be the case that the physical stimulus is identical in all respects. Only repeated hearings of the same recording in exactly the same acoustic conditions would guarantee that. What is it, then, that we use to make judgements of equivalence?

This is not a problem unique to music. It relates to any experience that we classify and categorize. How do we identify different instances of the handwritten letter 'a' as such, given that they may be written in different handwritings, orientations, sizes, etc.? Simplifying 30 years of intense scientific work in a few sentences is not easy, but the essence of mainstream theorizing about the matter is that humans make progress by forming representations of objects which are less like 'photographs' and more like 'lists'. In other words, a letter 'a' might be represented as

- a bounded circular space
- a vertically oriented line attached to the right edge of that space.

In the process of representing objects in this way, some detailed information from the original is lost, but a more useful level of underlying structure is gained, because the representation is more abstract than the stimulus from which it is derived. Several physically different objects can be captured under the same description in a way that identifies their common function.

Psychologists are not claiming that individuals always have conscious verbal access to the nature of their representations. In many cases, the existence of the representations can only be inferred on the basis of how people actually classify objects, i.e. which objects are perceived by people as similar or identical to one another.

In music, for instance, we know that, by and large, people do not represent music in terms of its absolute pitches or frequencies. Unless a person has an acutely developed sense of absolute pitch, he or she will judge as 'identical' the sequences C–D–E–F–G and D–E–F#–G–A, so long as they are not played consecutively. This judgement will be made by anyone with sufficient exposure to tonal music. It does not require formal training. Formal training will simply allow someone to verbalize what it is about these two sequences that make them sound the same. They are both ascending major scales starting on their respective tonics and ending on the fifth degree.

To summarize a very large amount of research, it appears that in listening to music, humans are normally representing what they hear, at least in

part, as a (non-verbal, intuitive) structural description of that music (in terms of pitch classes, tonality, metre, tension and resolution, continuities and discontinuities, repetitions and transformations etc). What it means to recognize a piece of music is to discover a match between the structural description of the current experience and a previously stored structural description.

In a very important sense then, the mere fact of being able to recognize a piece of music means that (intuitive and subconscious) judgements were exercised on both previous and current hearing. These were the judgements that assigned *this* event to *this* pitch class rather than the neighbouring one, and so on. Of course, there are different levels of judgement. It does not require detailed note-by-note judgement to identify two pieces of music as 'organ music' or 'pop music'. It requires much detailed judgement to recognize two events as performances of the same piece, and even more judgement to reliably differentiate the two performances.

## 20.6 Ineffability and judgement

What the preceding discussion highlights for us is the fact that there is an great deal of 'knowing' of one sort or another going on in engagement in music, which is not necessarily translated into 'saying'. I have a sense that this gap between knowing and saying is one that offers particular affordances for worship. At the heart of much worship is the sense of being in the presence of that which is beyond capture by human concepts. In approaching the object of worship we are approaching that which is at the limits of our apprehension. And yet, neither the object of worship nor the activity of worship is alien to us. In worship, there is fulfilment. We discover in the object of our worship intimate knowledge and understanding but cannot adequately say what it is we know or understand. The adjective which describes those things which we can know but not say is 'ineffable'. Human experience of God has always possessed an ineffable core.

One recent project in the philosophy of music which seems to have crucial implications for our understanding of the role of music in worship is the attempt to clarify in what ways, if at all, musical knowledge is ineffable. This project has been undertaken by Diana Raffman in her recent book *Language, Music, and Mind* (1993), and what I have to say here draws very deeply on her thinking.

Raffman identifies three distinct ways in which individuals may know more than they can say about a musical experience. She has a particular concern to determine whether any of these types of ineffability are logically necessary for all listeners at all times, but this concern need not be ours. If ineffability can be experienced at all, then it may be relevant to worship.

### 20.6.1 Structural ineffability

The first type of ineffability identified by Raffman is *structural ineffability*, and it relates to limitations in the ability to verbalize high-level structural features which affect the way that music is represented. She provides some examples of experiences falling into this category which will be recognizable to many experienced musicians:

> Having easily reported the pitches, rhythms and times-signature of a heard piece, along perhaps with some of its local grouping, time-span, and prolongational structures, a listener (perhaps even a highly expert one) finds himself at a loss: 'I am feeling that E-natural in a certain distinctive way, but I can't say just how. I know it's the leading tone, and that it's preparing the return to the tonic, and I know it's a weak prolongation of the E-natural in the previous bar . . . but somehow I feel there is more going on'. Or a performer feels himself compelled to play a certain passage in a certain way (e.g. to slow down, or to get louder, or to increase the vibrato speed)—consciously feels, knows, that it must be played thus, and yet cannot say why . . . What I want to suggest is that unconscious structural representations—in particular, relatively global levels of representation—are 'making themselves felt' in our conscious experience, yet their contents elude our verbal grasp. (pp. 31–32)

Another way of experiencing this kind of ineffability is the sense of 'rightness' about a particular musical event. We feel that 'it had to be like that', even though we have no conscious access to explicit reasons for the judgement. It seems to me that this maps rather straightforwardly onto an aspect of worship which concerns our apprehension of 'the larger design of things'. We can marvel or wonder at the rightness and beauty of aspects of creation at the same time as being unable to do more than catch hints of the underlying nature of the phenomenon. There is a kind of inexhaustibility to this level of ineffability, which spurs scientific and other intellectual endeavours. Our efforts to understand can be likened to peeling an infinitely large onion from the inside. There is always another higher level of explanation beyond the one that has been discovered. The permanent gap between intuitive contemplation and intellectual understanding is a central feature of the worshipping experience.

### 20.6.2 Feeling ineffability

The second type of musical ineffability is characterized by Raffman as *feeling ineffability*. This comes about because of the essential sensory-perceptual

or experiential nature of music. Knowing a piece of music involves, along with other things, *knowing what it sounds like*. 'A person deaf from birth cannot know a piece of music' (p. 40). This does not mean that a trained musician has to hear a particular new piece of music to know it. It is possible to compute what it sounds like (at least to some degree of specification) by examining a score, provided that one has a rich enough base of prior musical experience to fall back on. But without this sensory underpinning there is no possibility of full knowledge.

This type of ineffability is by no means unique to music. All sensory experiences, from sunsets to 'the taste of last night's chicken curry' have the same ineffability. They require 'actual occurrent sense-perception of the relevant stimuli at some point' (p. 4). Yet, according to Raffman, there seems to be a special quality to the feeling ineffability of musical knowledge that commands our particular attention in the way that last night's curry does not. Her thought is that it is the grammatical structure of music that (mis)leads us to expect something effable as an end result of 'following' a piece of music. 'Music's grammatical structure may mislead us into semantic temptation' (p. 41) in a unique way not brought about when we observe a sunset, or even a painting. This comes about because of music's formal similarity to language.

I'll try and put this another way. Music's formal structure and its similarities to speech (sequential, unfolding over time) suggests to us that it is saying something to us, or pointing our attention to something. What it is pointing to, among other things, is the actual sensual experience itself, which by definition cannot be fully described in words. Unlike a sunset, therefore, music *announces itself to us*. This power of music to draw our attention to our own sensations and encourage us to consider them as significant seems also to me to have an obvious purpose in worship. Many traditions of worship encourage the development of an attentiveness and a readiness to be 'spoken to'. Yet, few traditions encourage worshippers to expect that the 'word of God' will always come directly through language, spoken or imagined. Rather, the message may come through images, experiences, sensations, memories. And so worshippers are encouraged to be still and pay attention to everything. Music may smooth our path towards such attentiveness by in some way simulating a situation in which we sense we are being spoken to, thus sharpening our attentiveness accordingly. Since the music is not literally saying anything (in the strict linguistic sense), our attentiveness is available for other purposes. We are made ready to 'hear God's voice'.

### 20.6.3 Nuance ineffability

Raffman proposes that the third and final type of musical ineffability is *nuance ineffability*. This type of ineffability comes about because we do not

have verbal categorizations or schemas as fine-grained as the stimuli we can discriminate. For instance, we may be able to characterize a pitch as 'middle C'. If we are particularly acute, we may be able to label it as 'a slightly sharp middle C', but none of us would be able to operate at the level of cycles per second. Although most of us could detect that there was some difference between two pitches of 250 Hz and 255 Hz respectively, we would not be able to name the difference at that level.

The reasons for nuance ineffability are inherently psychological. Raffman cites a number of such reasons, of which I think two are compelling. First, as Raffman argues, 'it is hard to see what point there would be to a schema whose 'grain' was as fine as perception'. The fact that we describe musical stimuli in broad categories (such as C, D, etc.) is arguably a result of our need reduce the intolerable information load that would be placed on our cognitive system by representing every nuance. In other words, human processing limitations make total effability nuances a psychological impossibility. Second she argues that the just-noticeable-difference (JND) for perceptual properties varies with stimulus and observer characteristics (including alertness, health, etc.).

If these arguments are correct, there will always be a residual ineffable component in any conceivable representation of music. As Raffman poetically describes it, musical experience will always be characterized by 'an evanescent corona shimmering around the structural frame of the piece' (p. 96).

This ineffability is a kind of opposite to structural ineffability. Structural ineffability occurs because, however powerful our mental telescopes. We can never quite get a fix on the biggest picture. Nuance ineffability occurs because our mental microscopes can never reach the smallest level of magnification. I forget who said that there is a complete universe in a drop of water, but such sentiments describe exactly the inexhaustible detail within a musical performance. Such detail is, of course, present in any sensory experience, but again, as was the case with feeling ineffability, we can suppose that the quasi-linguistic nature of music points us towards the 'shimmering corona' and encourages us to endow it with significance. In worship there are contemplative states in which tiny, usually ignored, details of experience are perceived in great clarity, as if for the first time. Musical experience may encourage such states.

To conclude this rather lengthy analysis of musical ineffability, we can see that there are strong arguments for supposing that much musical experience challenges us to go to the limits of our ability to pass easy judgements. We are constantly getting tripped up over one or another form of ineffability. It doesn't seem to me that the act of making well-practiced

verbal judgements (about music or anything else) has any particularly special relevance for worship. Where worship seems to be afforded a particularly strong foothold is precisely at the boundaries of what can be said. Music brings us particularly effectively into an awareness of the ineffable, and thus into a core attribute of worship.

## 20.7 Music as a vehicle for non-judgemental contemplation

The preceding section has suggested what I now need to say explicitly. There may be a sense in which, while listening to music, we cannot help but make judgements about it. These are not so much verbal judgements, as those subconscious and automatic acts of assigning notes to the basic categories of representation. There is much research literature that shows that when we are very familiar with particular structures and codes, our cognitive system performs these routine analyses whether we like it or not. Therefore, when I refer to non-judgemental contemplation, I really mean an attitude of mind in which a listener does not engage in specific verbally based activities leading to verbal judgements (e.g. 'what key is this in?', 'is the trumpet playing in tune?', etc.).

In a restricted sense, therefore, the pharmaceutical model rejected earlier may have some merit. Within a given musical culture, and among individuals who have had comparable prior musical experience, there may be some automatic and subconscious mental processes which are indeed determined primarily by the nature of the particular piece of music being heard. However, if these processes are predominantly inaccessible to consciousness, what is it that a listener may be aware of that derive directly from these processes but are neither associations, nor verbal judgements? One major category of such objects of awareness is the category of affective experiences: feelings, moods, and emotions. It seems clear that there are at least some musically generated affective states which are not simply triggered through association to previously experienced instances of the same or similar music. Music has some inherent characteristics which promote affective responses.

Two lines of evidence (more extensively reported elsewhere, Sloboda 1991) support this assertion. First, music is reported by a wide range of individuals as evoking, often on first hearing, emotions and feelings of an unprecedented intensity and unique character. If someone has feelings she never had before to music she never heard before, then specific associations cannot be playing a major role in the generation of that state. Second, the nature of the affective response to music is sometimes determined by very

precise structural characteristics of the music, such as a particular harmonic structure, and seems to be shared by members of a culture. In general these structures are rather abstract, and can be realized in a wide range of specific sequences. It is not the case that there is any simple one-to-one mapping between a key, melody, interval, or chord, and a specific emotion.

There are two main structural determinants of affective response to music. The first, which we could call 'global', relates to overall characteristics of a piece or a section. For instance, smooth, quiet, slow music tends to evoke moods of repose or resignation. Fast, jagged, load music tends to evoke agitated moods, whether of happiness or anger. There seems often to be some rather straightforward analogy being drawn to the characteristics of a human being's behaviour when in such moods. So, for instance, a sad person tends to behave slowly and quietly.

The second determinant, which we could call 'local' relates to the moment to moment changes in affect which are experienced as the music unfolds. So, although we may experience a whole piece as 'sad', we find that sadness reaches a peak of intensity at a particular moment. It appears that many of these moments are determined by the precise confirmations and violations of expectations that are built up in the course of listening. Of course, two people listening to the same piece of music may not always experience the same emotion to such an event, any more than the same person may do so on different hearings. But what the available data do strongly suggest is that if these moment-specific emotions are felt at all, then they are felt at the same points in the music by all listeners. Variations in emotional experience are not caused so much by idiosyncratic ways of representing the music, as by contextual factors (e.g. the prevailing degree of emotional arousal in the individual, the importance of the event of which the music forms a part, etc.) and by the individuals own construal of the emotional response (e.g. one individual may find it wholly therapeutic to be moved to tears by a musical passage, another individual might find it excruciatingly embarrassing).

An example of a musical feature that appears to provoke emotional responses associated with crying (e.g. tears, lump in the throat) is the repeated suspension or appogiatura. Many composers and commentators have explicitly recognized this characteristic as indicative of pathos, and this does not seem to be an arbitrary feature to serve the same effect. The emotional response has something to do with the repeated harmonic tension and relaxation that such structures produce within a tonal context.

It is particularly relevant to the discussion of the role of genre in worship to note that these structural features are not the province of narrow genres.

One can find exactly the same use of appogiaturas in baroque orchestral music as in 1980s pop ballads, yielding exactly the same kind of emotional effect. What can differ, but not in a simple way that is mapped onto genres, is the complexity, subtlety or 'disguisedness', of the way in which a basic structure is deployed by the composer. It is such subtlety which distracts us from noticing too readily that the same underlying structure is present in two pieces of music that affect us in similar ways, and adds a sense of wonder or newness to what might otherwise seem manipulative and overly sentimental.

History has a refining effect. By and large, what survives from ages past is the interesting and unusual rather than the trite. Contemporary music, on the other hand, contains much which is of little value, precisely because history has not yet been able to weed out that which people do not value. I am convinced that many of the value judgements about genres suitable for worship are based on average 'triteness' ratings on the pieces of each genre known to an individual. On these grounds, the average pop track is bound to be triter than the average sixteenth-century anthem. So, also, I have to say, is the average piece of music especially written for church use in the twentieth century.

Here, then, is where individual differences will 'get in the way' of the contemplative response. Many people will resist the emotional effect of some music, simply because the effect is too blatant or obvious. Such resistance is not unique to music. I have heard people complain that they find themselves resisting being moved to tears by the crude manipulative dramas of soap operas. Several musicians I know complain that they find tears coming to their eyes during the performance of 'Land of Hope and Glory' at the Last Night of the Proms, even though they claim to loath the music, and the shallow sentimentality of the occasion. They almost feel ashamed of their response and try to suppress it.

For true contemplation, it is necessary to mentally instruct oneself to forget where the music comes from, to forget one's associations to it, to forget whether one likes it or not, and simply let it work its effect. There is a strong element in worship of quieting one's own inner voices to 'let God speak'. This is often expressed as a subjugation of one's own will and inclinations to God's. Precisely because music does have such strong personal associations, but also *of its own nature* evokes emotional responses which may be quite at variance with one's personal predispositions towards it, there is a clear sense in which the contemplation of music embodies the particular challenge to 'self' that marks out contemplative worship.

## 20.8 Music as a vehicle for experiencing personal relationships

Engaging with music can bring the participant into particular kinds of relationship with others. I have already alluded in the previous section to the fact that music has a tendency to coordinate at least the shape of the rise and fall of emotional response to music in a body of people. We are all likely to feel most strongly at the same point, even if the precise colour of our feelings differ from one to another.

Such coordination becomes even more focused when we are engaged together in making music. In music-making we are contributing to a larger whole, so that our small individual contribution becomes more significant. This happens in two ways. First, and more simply, our effort is amplified (as when a large body of people are all singing the same tune). Secondly, and more profoundly, new things are achieved through coordination of different activities. New aspects of a melody emerge when it is put in counterpoint with another one, because added to the individual characteristics of each melody on is now the harmonic and rhythmic structure that is created by the relationship of the two melodies.

Music is thus a very good analogy for certain key aspects of worship. In almost every religious tradition there is the notion of a worshipping community, that in some way the coordinated worshipping of a body of people united by common activity is somehow more powerful and effective than individual prayer. Worship reflects human solidarity. And secondly, worship is expressive of relationships between one individual and another. Any relationship is more than the sum of the two individuals. Something new and valuable is realized in the meeting of persons; and in many traditions, including the Christian one, some of the profoundest theological concepts are attempts to grasp just what the newness consists in.

There is, of course, a major problem about communality. A sense of solidarity is not, in and of itself, a guarantee of goodness. Music, and other forms of communal activity can inspire people to unite for evil as well as good ends. Sometimes, the 'godly' act is to refuse to be caught up in communality, and stand alone. It seems to me, therefore, that music can only be used for effective communal worship where there is already an existing community, bound by common understandings and commitments, where trust and 'good works' underpin relationships. In this context music in worship deepens and strengthens mutual commitments. Music cannot, in and of itself, create community.

## 20.9 Conclusions

I hope that this paper has demonstrated that there are several characteristics of human response to music which are particularly effective affordances for worship. Many of the mental processes encouraged by music are also present in worship. What I have not done is to show, in any very systematic way, that activities other than music do not have the same auspicious set of affordances, thus demonstrating music's unique role. We really need a comprehensive psychological analysis of all human activities that might impinge on worship, comparing and contrasting the affordances of each. But that is a task for a different time and place.

## References

Brown, R. and Kulik, J. (1977) Flashbulb memories' *Cognition*, **5**, 73–99.

Gibson, J. J. (1966) *The senses considered as perceptual systems.* Boston: Houghton Mifflin.

Norman, D. A. (1988) *The psychology of everyday things.* New York: Basic Books.

Raffman, D. (1993) *Language, music, and mind.* Cambridge, MA: MIT Press.

Sloboda, J. A. (1989) Music psychology and the composer. In S. Nielzen and O. Olsson (Eds.) *Structure and perception of electroacoustic sound and music.* Amsterdam: Elsevier.

Sloboda, J. A. (1991) Empirical studies of emotional response to music. In M. Riess-Jones (Ed.) *Cognitive bases of musical communication.* Washington, DC: American Psychological Association, pp. 33–46.

CHAPTER 21

# EMOTION, FUNCTIONALITY, AND THE EVERYDAY EXPERIENCE OF MUSIC: WHERE DOES MUSIC EDUCATION FIT?

## 21.1 Introduction

This paper is intended to raise broad issues about the nature and purpose of general music education. Although it is grounded in an empirically observable phenomenon (the decline of engagement of young people in traditional forms of music activity), it goes far beyond scientifically validated data in an attempt to locate the phenomenon we observe within a broad historical and cultural context. I am, however, not a social historian, and I am painfully aware that I may be doing badly what others unknown to me have done much better. I hope readers will take this as an opening contribution to a debate which not enough people in music education research appear to be participating in, rather than a polished and finalized set of positions.

## 21.2 Why do children drop out of instrumental music?

A phenomenon which has been reported with concern by observers of UK music education is the decline of engagement by young people with those forms of musical activity which have been traditionally encouraged and supported within the school system (particularly the playing of traditional acoustic musical instruments). This decline may be mirrored in other countries, but the focus of this paper is on data from the UK, and its implications for music education in the UK.

The transition from primary to secondary education appears to be a point of particular vulnerability. Many children appear to 'leave instrumental music behind' as they make this transition. This has recently been confirmed by data from a large-scale study of young people's participation in music

(cf. Ryan *et al.*, 2000). In this longitudinal study we have been able to collect data from 684 school-aged children (aged 11–12) at three points in time: (a) in the last half of the final year at primary school; (b) in the first half of the first year at secondary school; and (c) in the second half of the first year at secondary school.

Children were asked the same question at each point: 'do you play a musical instrument?'. Among these 684 children, 420 answered 'yes' at point (a). By point (c) only 229 of these were still playing; 191 children (45% of the players) had given up playing.

Although detailed quantitative analysis of the correlates of this massive drop-out is still under way, I'd like to present here two illustrative vignettes, drawn from intensive interviews with two of the children concerned. A sample of 76 children were interviewed at point (a) and again at point (c). The vignettes presented here are from two children who showed a very high degree of instrumental engagement at point (a) but who had given up playing by point (c), one year later. Their names have been changed.

Mary was receiving weekly group violin lessons at school when we first interviewed her. She had learned the violin for two years, but had not taken any Grade exams. She reported playing the violin for upwards of 21 hours per week (3 hours per day).

Lucy was receiving weekly individual piano lessons at her teacher's home when we first interviewed her. She had been learning the piano for 3 years, and had passed Grade 3. She reported playing the piano for between 11 and 15 hours per week (2 hours per day).

Mary and Lucy's weekly investment in their playing puts them both in the top 5% in terms of instrumental time investment within the playing cohort. These levels of investment equal or exceed the age-average (2 hours per day) for the highest achieving players in the study of practising time carried out by Sloboda *et al.* (1996). Mary and Lucy could, therefore, have been expected to be on a trajectory similar to those reported for those children in Sloboda *et al.*'s study who succeeded in competitive entry to a nationally recruiting specialist music school.

Three significant features characterize the accounts given by Mary and Lucy at point (a) while identifying as players.

1 *'Fun' is a key motivating concept, both for individual playing and participation in lessons.* When asked 'What does being involved in music mean to you?', Mary responded: 'It gives me a chance to practice because I like playing my violin, and most children in my class play instruments, so I just thought it would be fun.'. In response to the question: 'What do

you like about playing the violin?' she said, 'I just like practising it, and I like playing it.' To a similar question, Lucy answered, 'Well its just fun really. I just enjoy doing it. It's a hobby, something to do.'

2 *While playing, achievement is valued (e.g. grade success, being able to play a difficult piece, playing in front of people).* Mary expanded her liking of playing the violin by saying 'sometimes the music we have to play is quite hard but it seems easy after a while'. When asked, 'What do you like about playing in front of people?' Lucy replied 'Um, its just that you can show off your talent because you only do that to like your family and music teacher really'. Later she explained: 'I enjoy playing the piano. You can get different grades for it and things, so like all your hard work to pay off when you get to pass an exam or something.'

3 *Parents and instrumental teachers are key players in the discourse concerning support.* In response to the question: 'What did your parents think when you first started playing the violin?' Mary said: 'Me dad was quite interested and happy for me, me mum, she doesn't like the violin because its screechy, she says I'm happy for you, and she's glad that I've got an instrument because I've been on about it all like for a long time, she says I'm happy that you've got the instrument the violin because you've always wanted it . . . '. In response to the same question, Lucy said 'they were really pleased and they said that I was making good progress as well.'

Of her violin teacher, Mary said; 'She's really kind. I did have one teacher and I didn't like her much because if we forgot our violins she used to go tell secretary and then we were in real trouble. But that wasn't all the reason. She didn't give us the chance to pick what music we wanted to play.' Lucy described her piano teacher thus: 'He's always like cheery, and he doesn't shout at you or anything if you get it wrong, he buys books for you to play out of and he just helps you work out notes and things.'

When asked if they thought they would still be playing their instruments in secondary school they both said yes.

Five new features characterize the discourse of these two interviewees at point (c):

1 *Playing an instrument is seen as 'boring'.* When asked 'Can you tell me why you decided to give up the violin?' Mary said 'Cos I didn't like it.' 'What didn't you like about it? 'I don't know. When I first started

playing it, I used to really like instruments but then I just went off it. I just found it boring . . . '. In response to the question, 'Why did you give up the piano?, Lucy said 'Well I'd been playing since year 2, and I'd got up to grade 3, and I just got bored with doing exams and that.'

2 *Previously valued achievements are discounted.* Mary, who at point (a) had enjoyed practising, now 'hated' it. When asked how her teacher reacted to her giving up she said: 'I think she was a bit upset really cos I had been doing quite well just before I quit, and she was like a bit upset that I was, she said I was making good progress, and if I'd carried on a bit longer I would have got some more grades.' 'And how do you feel about having given up now?' I feel quite glad actually. When Lucy was asked 'Do you ever regret having given up?' she replied: 'No . . . . It was.. the exams and practising all the different ones I didn't like doing.'

3 *Other activities are valued more highly than music (e.g. academic lessons, homework, seeing friends).* Mary said 'I hated practising, and I had other things to do, like homework and going outside and playing with my friends, and its just like that took over.' In response to a question about whether she might take up an instrument the following year, Lucy said no, because 'the first year at high school, when you do instruments you miss your lessons, and the you have to catch up, so you worry about having to catch up on your lessons.'

4 *Dropping out of music is associated with discourses of autonomy and self-determination.* When asked whether she missed playing, Mary said no 'because I always used to play at five o'clock till six o'clock, an hour every night. My mum used to push me about doing my practice. I just missed it for a bit, but then I got used to it, doing things I really wanted to do.' Lucy got explicit support from her mother for autonomous choice to stop—'I told my mum first, and she said she wouldn't make me do anything I didn't want to do. So she told my (private) teacher for me.'

5 *Future engagement in music is not ruled out but is conceptualized in terms of instruments and social networks that were marginal to (and possibly conflicting with) previous involvement.* When asked 'If you were to take up an instrument again, is there anyone who you think would be particularly helpful or supportive to you?' Mary nominated her grandmother. In the earlier interview Mary had told us that she used to play the piano, which she actually preferred to the violin (although found it more difficult). Her grandmother played the piano, and Mary sometimes went to her grandmother's house to play the piano there, which used to be in Mary's house. When Lucy was asked whether there was

any instrument she might consider taking up in the future she nominated the saxophone because 'my friend plays cornet, and I go to their jazz concerts, and I like the sound of the saxophone.'

## 21.3 The sociocultural environment for young people

In gaining a clearer understanding of how a well-embedded daily activity can collapse like a house of cards within a 12-month period, I'd like to propose four conjectures which link these vignettes to wider social and cultural trends.

*Conjecture 1.* Instrumental playing is a 'hobby'. It is in the nature of hobbies to 'grow out of them'. Therefore an understanding of the 'take' that a lot of children have on musical activities requires a broader understanding of the psychology and sociology of hobbies.

*Conjecture 2.* Instrumental playing 'comes from' the culture of the primary school, where musical activity is just 'part of the furniture'. It's underpinnings are undermined through the transition to secondary school because:

- instruments are 'less important' than 'academic' subjects
- the peer groups through which instrumental activities are reinforced have been disrupted
- secondary music teachers are (for structural and cultural reasons) in a less good position than their primary colleagues to maintain a shared set of values and expectations within which positive engagement can proceed.

*Conjecture 3.* Instrumental playing is associated with parental control and a focus of everyday life on the home. The desired focus of the preadolescent life shifts in the direction of increasing assertions of autonomy, identification with peer groups, and life outside the home.

*Conjecture 4.* For many young people instrumental playing has no inherent 'purpose' which relates to their activities and goals, even those which might involve music in other ways.

For these conjectures to be validated, evidence from a number of sources will be relevant. Some of this evidence exists already. One of the most important sources of evidence comes from detailed studies on the musical lives of individuals in different subcultures within UK society. This shows how people appropriate music for their specific personal and social uses in ways which are often unique, and which have little to do with the forms and activities of music which are found within educational establishments. Two major examples of this approach are provided by Green (2001) in relation to the lives of popular musicians, and De Nora (2000)

in relation to the personal lives of adult non-musician women. Another example is given by Sloboda *et al.* (2001) who obtained a one-week 'diary' of musical engagement from adult non-musicians using an 'experience sampling methodology' (originated by Czikszentmihalyi and Lefevre, 1989). Participants carried pagers with them at all times for a week, and made a structured entry into a 'diary' every time the pager sounded. They were required to report on concurrent activity and the nature, purpose, and effect any music happening in the context of that activity. The study showed that the vast majority of music takes place in the context of (and subsidiary to) non-musical activity. The most common types of activity are what we call 'maintenance' activities, those essential routines of everyday life which everyone has to perform many times each day (e.g. washing, cleaning, cooking, travelling, etc.). Usually these activities are enacted alone, and music seems to be a resource through which the individual addresses issues of personal autonomy and identity. Because of this, the accounts of participants from a range of studies suggest that music is often a source of conflict where different individuals occupying the same space have different relationships to, and needs from, the music. The same music that affirms and reinforces one person's psychological agendas can simultaneously threaten another's.

Sloboda *et al.* (2001) also investigated change in emotional state as a result of the music experienced. In general, participants experienced greater and more positive change when the music was chosen by them.

Unfortunately, all the studies of this sort known to us have been undertaken with adult participants. The details of the intimate hour-by-hour musical lives of children in contemporary society are almost unknown to us. We really need to know much more about what children autonomously use music for in their everyday lives. In particular, we need to know what are the 'natural' varieties of performance that are given meaning within their solitary, family, and social settings. Then we can begin to understand better how formal instrumental playing maps (or fails to map) onto these natural categories.

## 21.4 The goals and attitudes of classroom music teachers

A second source of evidence relevant to an understanding of why children drop out of formal music relates to the nature of what it is they are running from. An important new source of information on this matter comes from a recent survey of 750 heads of music in UK secondary schools catering for

ages 12–16 (York, 2001). Seventy-eight per cent of respondents had degrees based in classical music, with 50% being either classical pianists, organists, or singers. The most commonly used public sources of support and resources for their teaching were BBC Radio 3, Classic FM, and *BBC Music Magazine*. The most commonly used specialist sources of information were the Qualifications and Curriculum Authority (55%) and the local music service (50%). Use of agencies such as British Phonographic Institute, the Musicians Union, or the Performing Rights Society was almost nonexistent.

Respondents were asked if they knew 20 pieces of music from various genres and periods , and could name the most associated composer or performer. Although over 95% could correctly answer for classical titles such as *The Four Seasons* (Vivaldi) or the 'Ode to Joy' (Beethoven), only half answered correctly for 'Wannabe' (Spice Girls), and only tiny numbers were correct for 'Kind of Blue' (Miles Davies—15%), and 'Rockafeller Skank' (Fatboy Slim—9%).

York concludes from these and other data that:

> School music culture tends to be introverted and avoids looking for models of current practice from the art of music rather than relying on the received knowledge of music education. . . . Many teachers seem to be engaged in using pop and rock to extend pupils' musical interests into other styles. . . . but the vast majority of teachers are doing this with little or no training or professional experience in pop or rock.

Although York's analysis focuses on the genres used by music teachers, there is implicit in his data the assumption that this is about more than genres. It is about different approaches to music, including ways of thinking about it and responding to it. One cannot just insert a popular genre into the set of classroom practices that have been developed to deal with classical music.

Musical subcultures are defined by much more than the style of music they use (it is also how used, in what contexts, for what purposes, assuming what type of inter-personal relationships, accruing what meanings). Labels such as 'pop' and 'rock' are coarse-grained and uninformative. Take, for instance, the case of 'techno'.

Techno is neither 'performed', 'composed', or 'appraised' within parameters that fit neatly with UK National Curriculum formulations. The music is 'constructed' in real time out of computer-manipulated elements at the disposal of a 'DJ'. Its primary function is to support communal (but individualistic) dancing designed to induce certain altered states of awareness.

Dancers may only experience the intended effects after several continuous hours of engagement. A short extract experienced in a classroom setting provides an incomplete, even misleading, basis for appraisal. The basis for valid appraisal exists only for someone attending a techno club and dancing to the music. No classroom teacher could hope to adequately address issues relating to techno with their students without specific understanding of, and exposure to, that sub-genre and its role for its habitual users. The same point can be made for almost any other sub-genre.

## 21.5 Determining the place of music in education

Music education in schools cannot function effectively without an implicit agreement between stakeholders (e.g. teachers, student, parents, government, etc.) about what it is for. The 'meaning of music' is a constantly shifting function of the discourses of these diverse groups, which may coalesce around a 'dominant ideology' which gains enough inter-group consensus to generate a stable educational agenda. I would argue that such a stable agenda existed in mid twentieth-century music education, but its underpinning consensus collapsed as a result of major cultural shifts, most evident from the 1960s onwards.

At risk of caricature, but emboldened by the analyses of writers such as Cook (1998) and Small (1998), the dominant paradigm of the early to mid twentieth century might be described as follows.

Classical artworks (as epitomized by Bach or Beethoven) represent the pinnacle of musical value. Deeper appreciation and understanding of such artworks is the most important (and universally applicable) aim of music education. Such appreciation and understanding is most accessible to those who have the technical and theoretical skills to perform these works. All performance-based education is oriented towards enabling a significant minority (or even a majority) to acquire the necessary skills to be able to perform at least some works from the canon. The most profound end result of a musical education trajectory is the development of the ability to add to the canon of masterworks through compositional activity. This is necessarily confined to the few. Music is necessarily best taught by people trained in the understanding and performance of the classical canon.

A revised version of this characterization emerged in the 1960s–1980s, primarily articulated through the liberal educational establishment, justifying some broadening of the syllabus but still very much framed by the pre-existing dominant paradigm. It ran something like the following.

Lesser forms of musical activity (e.g. world music, pop, jazz, etc.) have their legitimate place within the syllabus, and may be necessary 'stopping points' for many people. A wide range of musical activities and genres permit the development of valuable skills, including technical skills relating to the perception and production of organized sound, the ability to subordinate the personal to the goals of a group, self-discipline, etc. Music can be a useful vehicle for interdisciplinary education, relating it to its cultural, historical, and scientific context. *But,* whatever broadening of the syllabus is contemplated, music education must remain controlled by those who have been through a full classical training themselves, since this remains the pinnacle of the musical pyramid, to which all, in the end, aspires.

Our present dilemmas may be an indication of the unsustainability of even this liberalized view of music education. It may be no longer possible to muster stakeholder consensus around any version of the educational enterprise which prioritizes the classical canon. Strong cultural forces have been at work which account for the collapse of this consensus. I identify seven such forces. There may, of course, be more.

## 21.6 Key cultural trends

### 21.6.1 Multiculturalism

An important consequence of the multiculturalism that increasingly has characterized the United Kingdom since the Second World War is the fact that European history and culture no longer defines British history and culture. The classical canon is the product of European history and culture, and its defence therefore increasingly becomes identified with a xenophobic 'fortress' mentality which places it on the margins of the cultural spectrum.

### 21.6.2 Youth culture

In developed countries, prosperity gives young people unprecedented freedoms and spending power. This allows them to set cultural agendas rather than accept agendas laid down by others. The effects of this are being felt throughout secondary education, not just in music, but it is arguable that stakeholder consensus is still holding together (just) in disciplines where the educational agendas relate more directly to occupationally relevant skills. Parents and children will ally themselves with school management where the earning potential of their child is at stake. They are less likely to do this when neither school management nor government is able to articulate a shared vision of the value of music which is stronger that the child's own lived experience.

### 21.6.3 Electronic communication

Unprecedented choice, and miniaturization of delivery technology, gives much greater individual autonomy in musical experience than has been possible hitherto. Young people can easily and cheaply create their own musical 'worlds'. Institutions such as schools no longer comprise a privileged route of access.

### 21.6.4 Feminism

This movement has had many profound effects on culture. One important effect is the progressive replacement of the hierarchic by the democratic as the paradigm of cultural organization. In the context of music, this has rendered the paradigm forms of the classical canon increasingly problematic. The symphony concert represents the pinnacle of such forms. Everything from the nature of the building to the social organization of symphony orchestras points to the subjugation of the entire cultural project to the will of the (overwhelmingly predominantly male) 'maestro' who is himself subjugated to the will of the (male) composer. Feminism foregrounds alternative modes of organization where leadership is more shared and fluid, and where the musicians are as likely to have generated their own music as they are to be reproducing music of others. Although the pop group is not without major contradictions as an expression of feminism (not least because the membership of such groups is still overwhelmingly male, and because the music industry attempts to be far more controlling of pop musicians than classical patrons ever had the power to be) it is hard to imagine major industrial nations according symphony orchestras the explicit or implicit role as 'cultural ambassador' that is now accorded to a country's major popular artists. When the London Symphony Orchestra visits the USA no one notices or cares. When the Spice Girls visit, this is front page news.

### 21.6.5 Secularism

Christianity (along with all forms of organized religion) has suffered a massive decline as a social force in the last 40 years (Brown, 2001). People look within themselves and their personal networks rather than outside for sources of personal meaning and fulfilment. Personal emotional work, rather than confirmation of the social/divine order, becomes a key focus of musical activity. Key musical organizations to which almost all children used to be exposed (the church choir) are no longer experienced by most children. I have argued elsewhere that churches have probably been responsible for keeping alive a 'ladder of opportunity' for classical music performers in every town and village for the last century. This ladder now

has many missing rungs (Sloboda, 1999, Chapter 19, this volume), and schools have neither the resources nor the social influence to bridge the gaps.

### 21.6.6 Niche cultures

In the age of easy electronic communications affiliations between individuals are increasingly based on 'special interests' rather than shared social or civic obligations (and the geographical proximity which this presupposes). Young people can be swapping information (and music clips) with people on the other side of the world. Schools, which could easily relate to the civic (e.g. village brass bands, the church, the Town Hall) are cut out of the loop of shifting grass-roots affiliations. The decline of the 'civic' has been most prominently documented by Putnam (2000).

### 21.6.7 Postmodernism

All the above are sometimes considered as key manifestations of a postmodern society, characterized by a free, even anarchic cultural 'market', in which the conditions for one segment to acquire cultural dominance do not exist. There are significant national differences in all these factors. For instance, the cultural homogeneity of small-town America, with the far stronger grip of religion, and relatively weak incursions of feminism, can probably still sustain a music education system which meets more dominant cultural needs.

The consequence of cultural fragmentation is that we music educators no longer occupy a privileged vantage point. We represent a small (and increasingly marginal) subset of these subcultures that coexist in the population. Furthermore it is an assumption of postmodernism that the conditions for the re-establishment of a new 'dominant cultural ideology' do not exist. We cannot hope for any easy return to the stability we once enjoyed. A national curriculum for music was probably introduced at the very moment in history when its sustainability had never been less certain. The classical conservatoire culture is vibrant and valuable. It just won't do for the majority of our children. In this context it is important to defend teachers, who are not to blame for the 'failure' of music education. The present rapid rate of cultural change is leaving many public institutions 'high and dry' in one way or another.

## 21.7 What should viable and engaging UK music education look like?

On the basis of the above it is at least conceivable to hypothesize that classroom music, as currently conceptualized and organized, is an inappropriate

vehicles for mass music education in twenty-first century Britain. Hints of the parameters of a more effective music education environment may well be found within the somewhat anarchic mixed economy of out-of-school music provision in this country. Such anarchy may be crucial breeding-ground for the celebration of personal autonomy and cultural differentiation that is a prerequisite for focused and goal-directed musical engagement in a postmodern society. This is not simply an anarchy of musical styles and genres, or even of technologies, but an anarchy of social relationships, where boundaries between the teacher and student role are creatively redrawn.

Variety may well be the key concept. I indicate eight areas where it may be possible to significantly increase the variety available to our young people:

1. *Varied providers.* There should be no monopolies of provision but a healthy alliance of national, regional, and local interest groups.
2. *Varied funding.* There should be a mixture of state, voluntary, and commercial funding, reflecting the different interests that have a legitimate stake in music education.
3. *Varied locations.* Music education should be delivered throughout the community, in schools, music centres, community centres, shops, homes, cultural centres (e.g. libraries, museums).
4. *Varied roles for educators.* To support a wider range of musical cultures and activities, those supporting young people need to adopt a wider range of roles than has traditionally been conceived. These include teacher, animateur, coach, mentor, impresario, fundraiser, programmer, composer, arranger, and studio manager.
5. *Varied trajectories.* Young people need a wide range of 'entry and exit' points from musical engagement. These should vary from long-term syllabi for some, but should include a far wider variety of short-term projects for others, who may dip in and out according to the needs and parameters of the project, rather than be subject to the tyranny of the school term or year.
6. *Varied activities.* The range of activities available to young people should mirror more closely the types of musical activities available in the subcultures they value. This will involve a wider menu of visits, workshops, concerts, talks, programme planning, and assistance in activities such as DJing. It is probable that the DJ is the most common 'deliverer' of music within the community. How many young people get meaningful help in developing the role of DJ?

7. *Varied accreditation of achievement.* A range of certificates of participation/competence is needed to accredit and document participation. Existing agencies need to broaden the range of qualifications on offer, and new agencies should develop them. There is nothing wrong with graded music exams or GCSEs. We just need more choices for young people who value accreditation.

8. *Varied routes to training competence (beyond GRSM, BMus, PGCE).* If people are to be suitably trained to support young people's music-making, then a far wider range of training and continuing professional development opportunities must be on offer than then traditional 'teacher education' model. Indeed, until this happens, it is hard to see how major change will occur.

It is quite difficult to predict what the final role of schools might be in a wider more inclusive view of music education. However, school music is the one aspect of the provision which we can guarantee all children to receive. Perhaps its most useful role is to provide a core anchor point where diverse experiences may be reflected upon, integrated, and coordinated.

## References

Brown, C. (2001) *The death of Christian Britain.* London: Routledge.

Cook, N. (1998) *Music: a very short introduction.* Oxford: Oxford University Press.

Czikszentmihalyi, M. and Lefevre, J. (1989) Optimal experience in work and leisure. *Journal of Personality and Social Psychology,* **56**, 815–822.

De Nora, T. (2000) *Music in everyday life.* Cambridge: Cambridge University Press.

Green, L. (2001) *How popular musicians learn: a way ahead for music education.* London: Ashgate.

Putnam, R. (2001) *Bowling alone: the collapse and revival of American community.* New York: Simon and Schuster.

Ryan, K., Boulton, M., O'Neill, S. A., and Sloboda, J. A. (2000) Perceived social support and children's participation in music. In C. Woods, G. Luck, R. Brochard, F. Seddon, and J. Sloboda (Eds.) *Science, music, and society: Proceedings of the 6th International Conference on Music Perception and Cognition.* Newcastle, Staffordshire: Keele University.

Sloboda, J. A. (1999) Music: where cognition and emotion meet. *The Psychologist,* **12.4**, 450–455.

Sloboda, J. A., Davidson, J. W., Howe, M. J. A., and Moore, D. G. (1996) The role of practice in the development of performing musicians. *British Journal of Psychology,* **87**, 287–309.

Sloboda, J. A., O'Neill, S. A., and Ivaldi, A. (2001) Functions of music in everyday life: an exploratory study using the Experience Sampling Methodology. *Musicae Scientiae,* **5**(1), 9–32.

Small, C. (1998) *Musicking: the meanings of performance and listening.* Hanover NH: Wesleyan University Press.

York, N. (2001) *Valuing school music: A report on school music.* London: University of Westminster and Rockschool Ltd.

CHAPTER 22

# THE 'SOUND OF MUSIC' VERSUS THE 'ESSENCE OF MUSIC': DILEMMAS FOR MUSIC–EMOTION RESEARCHERS

## COMMENTARY ON *CURRENT TRENDS IN THE STUDY OF MUSIC AND EMOTION*, A SPECIAL ISSUE *OF MUSICAE SCIENTIAE* (2001–2) *EDITED BY PATRIK N. JUSLIN AND MARCEL R. ZENTNER*

### 22.1 Introduction

I was a musician long before I was a psychologist. Specific performances of individual compositions spoke to me as no human being I knew had spoken, certainly by the age of 5. These, mainly recorded, performances became friends, comforters, inspirers, pointers to possibilities beyond the norms of my life. I knew these pieces as a lover knows a beloved. I knew every crevice, every corner, every quirk which made this piece individual, even odd, but identifiable as distinct and itself, different from every other piece, even though sharing so much in common. Within many pieces, I had favourite parts, and parts that I got bored with, or struggled to understand. And after a while my infatuation with a particular piece waned, and I sought new experiences with different pieces. Were emotions involved? Obviously! But specific emotions were embedded within something much more complex and multifaceted. Being 'into' these pieces of music was about learning who I was, what I might hope for, and something about what love and commitment could be for me. In other words, my involvement with these pieces of music was about the deepest and most ultimate aspects of being: self-knowledge, spirituality, transcendence, and growth. They were not only my friends, they were my guides.

Because I heard some of these pieces on the radio (used for example as interludes between spoken items), without any textual or background information, I did not always know their names. But I could often conjure up a complete auditory 'image', just as a lover might visualize the behaviour and personality of someone watched from afar, but to whom one had never been introduced. Name and genre knowledge came much later: sometimes many years later.

I was astonished and surprised to make the discovery, probably some seven or eight years after my first exposure, that a number of short pieces, which I had heard separately at different times, actually were part of a single entity – the Tchaikovsky *Nutcracker Suite*. It was a similar surprise to finding out that several quite different people one knew in different contexts all turned out to be brothers and sisters. How could these characters, all so different, so individual in such distinct and quirky ways, known and loved in such different times and places, have come from the same source?

Why have I made this small excursion into the autobiographical? It is because reading some of the contributions to this special issue raised in me an echo of this childish surprise. For instance the *Nutcracker Suite* appears as one of the items in Table 2 of Vastfjall's masterly analysis of the music mood induction literature. As I surveyed this list, I wondered how it was that this kaleidoscopic collection of deeply loved friends could be 'reduced' by Parrott (1982) to one of several dozen other named items that induced 'positively valenced affect' when used in a mood induction experiment, with the implication that it didn't much matter which of these several pieces were being heard at the time, or who was hearing them, or how well they knew the pieces, which bits of them were liked more or less, or indeed that there might be subtle differences between the induced feelings which were far richer than any psychometric scale could ever detect. I felt the same jolt as I might if overhearing a local bigot write off individual members of a family that I knew well as 'all the same' because they were members of a particular ethnic minority.

Is it not understandable that many people (particularly those most intimately involved with music) feel a sense of uneasiness, possibly verging on outrage in the most extreme cases, that psychological research may be failing to address the most fundamental and essential aspects of what it is to be humanly involved in music? Loved pieces of music are being reduced to a set of 'effects' such as might similarly be realized by the colour of paint on a wall, or the administration of caffeine.

For many years I held this unease at arm's length. To do this I had two main defences: the scientific reification of the variable, and a rejection of elitism (or at least, my own simplistic takes on these hugely complex issues).

Reification of the variable is an intellectual viewpoint which asserts that the only things one can really understand are the things one can control and measure (in the instrumental ways available within experimental or quasi-experimental settings). While control over variables is completely necessary for very many aspects of scientific activity, including hypothesis testing, it is in fact an error to believe that one can do good science if one limits one's thinking and observations to that which can be expressed in terms of those variables. On the contrary, without experiencing the fullest complexity of a phenomenon, one cannot know what it is that the variables within one's theory are supposed to be explaining. The danger for scientists is that we 'filter out' those aspects of the experience that are not encompassed within the dimensions that we use to describe and talk about them, and so what we work on becomes an almost unrecognizable parody of the phenomenon we are supposedly investigating. In the case of 'musical mood induction procedures' it is hard to recognize much that would be considered music listening by a person such as the child I once was. Such activity might indeed be taking place in some of the people being studied, but the level of description and measurement is often such as to completely exclude that experience from the discourse.

I am not here talking principally about ecological validity, although this is a linked concept. Situations are ecologically valid if they resemble actual life experiences in significant ways. People listen to recorded tracks that they may not have heard before in everyday life whilst performing a range of cognitive tasks, so there is *prima facie* ecological validity when people are asked to listen to such tracks in a laboratory setting. My concern is more about which aspects of that listening situation are being measured, and whether they can lead to real knowledge unless they are continually being held up against the known fullness of human musical experience.

My second defence was based on a rejection of elitism. For many years, it appeared to me (and others) that the experimental psychology of music could be seen as contributing to the liberation of the music-listening public from the grip of notions promoted by the academic music establishment that derived from a ninteenth-century bourgeois conception of art-music as a small set of masterworks that formed the canon, and whose greatness transcended culture and history. Some of our work, and parallel work undertaken by ethnomusicologists, and the increasing number of proponents of the 'new musicology', was engaged in a process of 'deprivileging' these masterworks and their cultural niches (i.e. the classical concert-hall *habitus*, to use the term coined by Bourdieu, 1977, whose sad death was announced as I write this in January 2002; see also Small, 1998). This process had, and continues to have, worthy aims (excellently outlined and

critiqued by Cook, 1998). It is a part of a programme of democratization of taste whereby any type of music experienced in any kind of context is as worthy of interest and study as any other. Thus, chart music experienced through a Walkman in the subway is no more or less worthy of study than Beethoven experienced in the concert hall. Indeed, one might even argue that chart music in the subway deserves more study, since this is a predominant cultural experience, whereas concert-hall music is a specialized niche. Science has therefore allied itself with grassroots sensibilities and present-day realities over and against the outmoded cultural imperialism rooted in a yearning for an idealized past.

It may not be coincidental that the intellectual and cultural home of this grand democratization project is, and has always been, North America. In popular culture it is hard to argue that any figures have been of greater global significance in this project than Elvis Presley and Andy Warhol. In academia, the pivotal figures have also been American (e.g. Susan McClary, 1991). Music sciences (particularly psychological sciences) have been profoundly influenced by the democratizing principles of behaviourism, of whom the most profound twentieth-century icon remains the American psychologist B. F. Skinner (who was at his apogee just as Presley and Warhol were emerging to an unsuspecting world). Behaviourism allowed a major escape from 'reputation and pedigree' in describing and accounting for human behaviour. It told us to ignore a person's heritage, rank, wealth, or status, and just concentrate on what he or she actually did (described in language which was similarly 'value free' and which could equally well be—and was—applied to non-human animals). The 'cognitive revolution' of the 1960s never abandoned the fundamental tenets of behaviourism, it just showed how behaviourist ideals could be applied to the measurement of the mental (including emotions and subjective reports) as well as the physical.

North American journals retain the highest scientific prestige within most areas of psychology, and most of the English-speaking research community adopts the conventions and research agendas that such prestigious outlets set. The contributions to this volume predominantly reflect 'Anglo-Saxon' intellectual values derived from the behaviourist programme (I say 'Anglo-Saxon' rather than North American, since these values have a second historic axis in Great Britain, and neighbouring countries of north-west Europe).

What first gave me cause to explicitly question the Anglo-Saxon programme as applied to music was my exposure to the ideas of the British sociologist of music, Simon Frith (1996). Frith made the observation that the rejection of elitism did not necessarily involve (as many had assumed)

the rejection of connoisseurship. People who listened extensively to the products of popular rather than 'high-art' musical culture could have a knowledge of that music which was just as articulated, subtle, and differentiated as the commentaries of classical music theorists. In other words, just because the materials and contexts of much popular music were mundane and everyday, it did not follow that the mental contents of such listening must be similarly 'stripped down' and mundane. What was true of my relationship to the *Nutcracker Suite* at age 5 could be true of any person's relationship to any music of any genre.

If this is the case, then the unease of those who question the Anglo-Saxon scientific agenda cannot be written off as an elitist backlash. Rather it suggests that 'cultural democratization' contains within it elements of a new cultural imperialism, an imperialism which seeks to downplay those elements of the musical experience which are individual, personal, complex, subtle, unreplicable; in favour of those elements which are communal, public, simple, plain, and replicable.

Although such imperialism may mesh well with strands of Anglo-Saxon science, it surely cannot relate in any simple way to the needs of science. Far more likely are economic motivations, deriving from the commodification of musical goods and services, and the need to predict and control musical consumption in the interests of profit. If significant psychological effects of music can be linked to a few key variables, along which musical products can be measured and classified, then the possibilities for mass marketing and mass consumption are hugely increased. If people can be encouraged, in subtle ways, to believe that these key variables are the 'main story' about music, then they will be much more susceptible to indiscriminate consumption of musical products which fit these descriptors.

The most clear, and bizarre, example of this process in recent years is the commercialization of the 'Mozart effect' (Rauscher *et al.*, 1995), whereby a small series of exploratory and rather inconclusive experiments was seized on by commercial interests bent on persuading anxious parents that they could improve their child's chance of economic success (through higher school grades) if they played Mozart to their child. Forget that Mozart wrote several hundred distinct works, which have received several thousand different recorded interpretations. The consequence of the way that music has been commodified is that a smart profiteer may take any set of Mozart performances (one guesses those for which the copyright and reproduction rights costs are the least), put them together on a compilation CD, and then reach thousands of eager purchasers. Whether families may benefit from the use of such recordings is not the issue here. The issue is that experiential horizons may be insidiously narrowed. When a piece of Mozart is appropriated

instrumentally as a kind of auditory vitamin pill, it becomes less likely to be encountered as a unique and highly personal cultural communication which has the power to speak at many levels, teach, inspire, become deeply known.

I would claim that the Mozart effect is one (although probably among the oddest) of a set of commercially inspired projects whose outcomes are gradually but inexorably 'thinning out' the musical habitus of the industrialized consumer, in the interests of production efficiency and profit. Instead of drawing on a multifaceted cultural discourse where Mozart's music is understood as a highly differentiated collection with a rich history of usage and assimilation, an entire repertoire can be 'reinvented' as 'the music that makes you smart', thus cutting it away from 250 years of cherishing, and re-presenting it as the vapid and culturally monochrome product of a advertising copywriter's pen.

## 22.2 Orientation and purpose

My reasons for opening this commentary with remarks which might seem overly speculative and culturally barbed are twofold. First, scientific activity can never escape from the cultural and economic forces which surround its practitioners. It befits us to examine our own practices and the degree to which our assumptions of objectivity are in fact deeply conditioned by our cultural outlook. Second, this commentary is being written at a unique juncture in history, following a series of events (11 September 2001) which many (including myself) believe will change the world more profoundly than any other set of events in living memory. These events have revealed that an ever deepening understanding of, and rational reaction to, the cultural and economic programme of the world's richest nations, is the most important task facing all intellectuals, whatever their discipline. That programme must neither be hailed as self-evidently good, nor self-evidently evil, but like all other cultures that have existed in the world, a complex mixture of the functional and the dysfunctional. The way we do our science can reinforce either of these trends, and, for many people at this moment in history, it is just not an option to 'carry on as normal' without evaluating the effect of the programmes we participate in.

My main question, therefore, to the contributions in this special issue, is to what extent the research reported points to, explicates, and encourages the understanding of diversity and complexity in musical experience. For diversity to be authentic, rather than stochastic, such diversity requires relating to the autonomous experience of the individual listener as agent, meeting the musician as autonomous agent, and generating an interaction

through the music which reflects and draws upon the motives, beliefs, aspirations, and goals of the players.

In posing this specific question of these contributions I am using them as an 'opportunity sample' of cutting-edge research in music and emotion which is somewhat representative of the field as a whole. My comments are thus intended to be interpreted in relation to 'the state of the discipline', rather than as specific critiques of these contributions which, having been through vigorous and rigorous peer review, should be assumed to represent a sample of the best quality work that is going on within the discipline. My question is more, what values do these contributions promote or undermine (whether explicitly or implicitly).

In doing this I am not for a moment suggesting that individual studies which explicitly address the issues I have raised are more scientifically worthwhile than those which don't, or that all scientific studies should be of a particular nature.

The papers fall into three main groups. The first group contains the contributions of Gabrielsson (2002), London (2002), and Vastfjall (2002). These are summative reviews which highlight in different ways the continuing attempt of the research community to achieve greater clarity on whether and how real emotions and moods may be induced by exposure to music, and how this relates to other ways of 'knowing' emotion-based aspects of music (e.g. through judgements of emotional 'character' or association). A key question I address through the contributions of this group as 'Does music really make us feel things; and if so, how?'

The second group contains the contributions of Scherer *et al.* (2002), and Schubert (2002). These present new data on listeners' emotional experiences, and are indicative of the ways in which music-emotion research has gone beyond the summative Likert-type rating scale as a means of tapping listener response. A key question for the papers in this group is 'How do we gain the most accurate picture of what is going on emotionally when people listen to music?'

The third and final group contains the contributions of Juslin *et al.* (2002), and Trehub and Nakata (2002). These also present new data, but explore the link between performer-controlled variations and listener response. Although the performer–audience situation is very different in these two papers (expert instrumental performers to adult listeners in Juslin *et al*; parents to babies in Trehub and Nakata) they share a concern with mapping communicative options available to performers. This group allows us to address the question 'How can performers affect the emotional experience of listeners?'

I will deal with these issues in turn, and then conclude with some summative remarks.

## 22.3 Does music really make us feel things, and why?

The celebrated English conductor Sir Thomas Beecham (a noted wit) is reported to have once said: 'The English don't like music: but they do love the sound it makes'. London expresses this quite profound distinction when he says:

> The acoustic qualities (of music) have expressive correlates and may trigger emotional responses, and of course one cannot have music without sound. However, musical expression is more than this: it requires the attention to the music *qua* music, rather than as mere sounds.

This distinction has encouraged several scholars (among whom London identifies Kivy) to see to what extent 'garden variety' emotion induction can be loaded onto music as sound, leaving music *qua* music free to engender feelings by a different route. Kivy (1999), for instance, wishes to claim that it is the aesthetic qualities of music which engage music-specific emotion—'Sad music emotionally moves me, qua *sad* music, by its musically beautiful sadness'.

My reading of this is that *recognition* of emotion (e.g. sadness) is a function of identifying characteristics of the sound that could equally well be instantiated in non-musical sounds (e.g. speech), or by picking up specific cultural associations of the sound (e.g. organ music is 'churchy'). In some contexts, this may lead to the listener *feeling* that emotion (or some other emotion). However, that entire process may proceed without any music-specific processes being invoked. I may cry when I hear sad music for exactly the same reason as I cry when I hear a person expressing grief through their vocalizations. Therapeutic and other behavioural effects of music may be mediated at this level alone; in which case it is more appropriate to say that the effects rely on the shadow cast by music on to the soundscape, but that they are not effects of the music *qua* music.

Musically relevant emotion requires, according to my reading of Kivy and others, an aesthetic response to the individuality of that work, not what it has in common with other musical works, far less with non-musical stimuli. Kivy identifies one possible perceived quality that can drive this process, 'beauty'. I am sure this is the right kind of quality to support his arguments, but wonder if we need (I) to unpack this rather imprecise

concept, and (2) characterize a wider range of ways in which music can be experienced and individuated. Beauty can, for instance, be related to both biologically and culturally induced factors, as well as nature and degree of prior exposure to related stimuli. Familiarity may be another key determinant of aesthetic judgement (as developed in the Berlynean approach, for instance). However, more work is probably needed to explain precisely what specific kinds of familiarity and unfamiliarity yield strong aesthetic attraction. If one uses analogies from human relationships (e.g. falling in love), they suggest a complex interplay between familiarity (the tendency of many people to fall in love with someone that resembles a parent) and novelty (we tend not to fall in love with people we know well unless we discover ways of seeing them in a 'new light'). The notion of 'revelation' seems pertinent—which may be characterized as the joint and simultaneous experience of something as wholly new and unexpected together with the experience of its absolute 'rightness and inevitablity' (see Sloboda 1989, 1999, 2000, for further exploration of these issues).

In order to engage these kinds of conceptual domains, it seems that a decision of some sort (whether conscious or subconscious, whether influenced by learned or innate factors) is required of the listener. This is the decision to treat the musical stimulus as intentional and personal (i.e. emanting from, or representing, an autonomous agent with communicative intentions). It doesn't much matter whether the stimulus actually so emanates (for instance many people experience natural phenomena, flowers, mountains and sunsets, in similar ways, and ascribe these to the intention of a creator). It is that decision that allows us to respond to music in the way that Kivy and others are pointing to.

The elaboration and development of these kinds of concepts seem not to be a priority of the mainstream experimental literature on musical emotion. The mere fact that music is presented to a listener is very often taken as a grounds to assume that music-*qua*-music listening is going on, and so the emotional effects are ascribed to 'music'. This is a gross oversimplification.

It was therefore a great relief to read, in Vastjfall's review of the music mood induction literature, that many researchers recognize that 'the music by itself will not automatically induce the desired mood state, and that (listeners) should try really hard to get into the mood, using whatever means they find most effective'. Furthermore 'within the same experiment, some participants report to be highly affected by the manipulation, whereas others remain unaffected'. Allowing participants choice over the music used in mood induction appears to increase the reliability and intensity of the outcome. All of these findings suggest that the apparent 'power' of music

rests on the full engagement of the listener, and is not a 'pharmaceutical' property of the sound stimulus. Commonalities of response between listeners arise only when their intentions, beliefs, cultural background, and experience can also be commonalized. Cultural homogenization may be the prerequisite for 'induction laws' to be derived by science. Vastjfall's review suggests that, despite the strong cultural pressures for homegenization of taste, individuals resist, and continue to respond to music in ways which individuate and differentiate them. They will not accept *Swan of Tuonela* in place of 'Nothing compares to you', whatever similarities at the level of 'sound characteristics' these two items may have.

The data reviewed by Gabrielsson (2002) shows how difficult it is to show commonality of response to music, either at the level of basic physiology or at the level of aesthetic response. Studies which correlate physiological responses to basic emotion categories (e.g. Nyklicek *et al.* 1997) rarely yield a straightforward outcome. The clearest outcome of this study was a correlation between respiratory variables and arousal level of the music. When musical sound parameters are arousing, people tend, on the whole, to breathe faster and deeper. What is missing from many of these studies is an appropriate non-music control (e.g. the sound of machinery or traffic). If the arousing properties of the musical stimulus are the same as those of non-musical stimuli, then we learn nothing much from these studies about the specific ways in which music-*qua*-music engages the emotions and the corresponding physiology.

One of Gabrielsson's most important contributions to the research literature is his series of studies on 'strong experiences to music'. These studies show that musical experiences sometimes allow people to enter into what Maslow (1976) has described as 'peak experiences'. Although these experiences have an emotional component, they are hugely more complex than the 'garden variety' basic emotions of happiness, sadness, etc. in ways that are not fully explored. There is little evidence that these experiences are 'determined' by musical parameters in any clear way. There is no 'set' of musical works which are guaranteed to provide these experiences. The accounts given by Gabrielsson's respondents show that strong experiences arise in situations which often contain complex biographical and non-musical factors which interact with the musical experience in unique ways. According to Maslow (1976), music and sex are the most reliable ways of engendering peak experiences. Understanding why this is requires a great deal more systematic and painstaking collection of idiographic data than has been undertaken up till now, both with musical and non-musical experiences. There is also a need for imaginative theorizing if the work is to progress beyond the stage of anecdote and natural history.

I am aware that any attempts to separate varieties of listening to music-*qua*-music from other modes of engagement with it are open to the very claims of elitism (and prescriptiveness) that I have been at pains to distance myself from. One might argue that if some people listen to music without those mental processes which are being proposed 'special to music-*qua*-music' then they legitimize these modes of listening as 'real music listening' by the mere fact that they exist. I admit that the issues here are sensitive. What need re-emphasizing is that there is much evidence to suggest that individuated and personally engaging involvement with music appears to be a primary human impulse, which is demonstrated in young children (further discussion of this point may be found in Section 22.5 below), in people with no formal musical training at all, and in many of the world's musical cultures. 'Taking music personally' seems to be a fundamental human response. It is at least possible that depersonalizing and deinvidid-ualizing music in the particular ways noted in commercially oriented cultures requires various types of psychological narrowing or 'reining in'. It is precisely such 'narrowing' tactics that cultural elites have traditionally employed. When the narrowing I have focused on is apparently presented as part of 'popular' culture, then those who resist it can be branded as elitist. This, however, I would argue, is a misreading of the situation. The resistance to the narrowing of the musical habitus is not driven by middle-aged white male Beethoven lovers. It is as likely to come from rock lovers, techno lovers, jazz or folk lovers; and as likely to be felt by someone with no musical training at all as someone with a conservatoire diploma. All these forms of music engagement are coming to have common cause as marginal cultural forms, under threat because their economic yield is insufficient. If cherishing relatively powerless minority values under threat from powerful global interests is elitism, then I gladly sign up to being an elitist!

## 22.4 What is going on emotionally when people listen to music

Gabrielsson's 'strong experience' studies are limited by the fact that they are retrospective, and allow participants to select one experience from a life-time's experiences. Such experiences must, however, be grounded in (and perhaps experienced as special in contrast to) everyday musical experience. Various studies (e.g. Sloboda and O'Neill, 2001) suggest that the average person experiences music perhaps 3–4 times every day. People are, thus, very experienced and accomplished 'handlers' of musical experience, and emotions are bound to be shaped by and responsive to these experiences.

The two papers in this section represent two different responses to the need to get closer to the actual musical experience. Scherer *et al.* undertook a retrospective study, but within a narrowed time frame ('the last time when listening to.. music produced a strong emotional state in you'), whereas Schubert gathered 'on-line' continuous responses during the music experience itself. Scherer *et al.* were unable to find features of the music or of the situation which strongly predicted emotional response. Out of a list of proposed determinants, a catch-all category of 'other factors' was the most frequently cited determinant. Scherer *et al.* conclude that 'since these determinants vary enormously over participants one can assume that whether an emotional reaction occurs or not may be determined, to a very large extent, by factors that are very specific for each individual'. Furthermore, Scherer *et al.* found that, in a free response situation, where participants wrote down whatever words they chose to describe their musical experience, the basic emotions, such as happy, sad, angry, etc., occurred rather rarely. Most often participants described their experience in more complex ways, in which 'being moved' was the most frequent component. From these findings they conclude that 'it is all the more surprising that a large portion of current research on the emotional effects of music seems to focus on such a small number of such "basic" emotions'.

Schubert's studies have brought 'continuous response' methodologies to new levels of methodological sophistication. These methodologies attempt to reflect the fact that emotional and other evaluative responses are not constant over the entire duration of a piece of music, but change dynamically in some kind of relation to what is going on in the music at the time. Although some studies (e.g. Krumhansl, 1997) have attempted to measure continuous response physiologically, most studies have used the manual response of the listener, who continuously adjusts a scale cursor according to the relative level of some experienced quality (such as tension, intensity, and specific emotions).

Schubert's contribution to this volume is a critical analysis of the significance that we can place on the reported fact that inter-rater correlations (between different people listening to the same piece of music) are high. Such reports are often used to support claims that different people have similar emotional reactions to the same piece of music. Schubert shows that a significant amount of this covariance can, in fact, be artefactual, due to serial correlation effects. He proposes statistical techniques for minimizing these artefacts in future studies, and concludes that we should treat inter-rater correlations of less than 0.8 with suspicion, even if statistically significant.

The implication of Schubert's work is that listeners are probably more different from one another in their response curves than a statistically naive approach to curve similarity would suggest. One possible reason for this is that the 'cognitive' load on a listener who is continuously responding is just too high. It may be that the respondent cannot meaningfully bring the cursor to bear on musical experience at every moment, and so there are 'dead patches' where the cursor simply remains at the last deliberately chosen location, or is placed at a 'resting-point' until the next event where a deliberate linkage can be made. Both these strategies could produce spurious inter-rater correlations.

A possible alternative to the continuous response methodology is a discrete sampling methodology where the listener makes a discrete response every time a particular experience occurs, or when the intensity of the experience moves over some threshold. This is the approach adopted by Waterman (1996) who asked respondents to press a button every time they 'felt something'. A typical respondent might press 3–4 times per minute, and Waterman found clear positive statistical relationships between different listeners in the places they chose to press, and structural commonalities between the musical locations of these presses. However, retrospective accounts of the reasons for these choices yielded wide individual differences. It may be concluded from data like this that there may be aspects of musical response (such as surprise or arousal) which can be predicted from measurable aspects of the musical stimulus, but that the experiences generated by those effects will be a complex function of the way in which the person concerned is engaging with the task.

## 22.5 How can performers affect the emotional effects of listeners?

Trehub and Nakata summarize intriguing evidence to suggest that infants have very clear preferences for particular types of adult vocalizations, and that these characteristics are shared by infant-directed speech and song.

In general, these characteristics (including high pitch and slow rhythm) are recognized (by adult listeners) as communicating positive emotional engagement (high pitch = joy, slow speed = tenderness). These preferences are accompanied by behavioural and physiological outcomes, which suggest that these vocalizations have an arousal-regulating function. Infants are actually comforted, engaged, and calmed by such vocalizations.

Caregivers' infant-directed vocalization patterns are also characterized by the repetitive use of rather few rhythmic and melodic contours. Trehub and

Nakata hypothesize that this repetition assists the infant in both recognition (of the caregiver) and learning (of speech and communication). However, they suggest that each caregiver's repertoire of vocal patterns is somewhat individualistic at the level of individual melodic and rhythmic elements. There is no such thing as a universal caregiver's song. This means that, from the very earliest times, infants are encouraged to identify specific musical patterns or elements which characterize and individuate the various 'significant others' in their lives.

In this context it may not be surprising that music listeners of all ages are prone to attempt to relate to music as if a significant other was present in the music and attempting to communicate something. In doing so, we are simply extending and enriching our earliest memories of the context in which we first heard and emotionally responded to communication from caregivers.

The data summarized by Trehub and Nakata suggest that adults (and older children) provide distinctive vocalizations to infants not because they are infants *per se*, but because infants elicit in most of us warm, caring, and affectionate responses. It is the affective nature of our response which determines the expressive content of our speech and song, rather than the age of the person we are speaking or singing to. This is confirmed by the fact that infants are equally happy listening to adult-directed speech and song if that speech or song is expressive of love or comfort. Both types of communication share similar acoustic qualities.

In sum, it appears that humans in the role of infant caregivers (and this can extend as young as preschool siblings) have an intuitive grasp of the ways in which their infant-directed vocalizations may be productive in the lives and development of infants. Shared characteristics of the signal (relating to overall pitch, speed, repetition, etc.) underpin potentially universal arousal and mood outcomes, whilst individually idiosyncratic characteristics assist with person recognition and differentiation.

Juslin *et al.* are also concerned with a decomposition of musical performance parameters into different sources with different potential outcomes on listeners. Their concern is with trained adult performers delivering short but complete musical compositions for adult listeners to hear. Their computational approach assumes (and formally characterizes) separate sources of variability between performers, makes predictions concerning differential effects of these sources on listener judgements, and then tests these by selectively manipulating presence or absence of the different sources in synthesized performances.

Their data demonstrate that the presence or absence of different sources of variation do, indeed, have systematic (if rather complex) effects on listener

judgements of the peformance as 'clear', 'gestural', 'human', 'musical', 'sad', and 'expressive'. The specific emotion of sadness was best communicated by the presence of performance parameters (labelled by the authors as 'emotional expression') relating to overall characteristics of tempo, sound level, articulation, tone attacks, etc. These are exactly the 'statistical' characteristics which Juslin and others (Juslin, 2001; Juslin and Laukka, 2003) have demonstrated link emotionally communicative music and speech (and which therefore are, in one sense, not special to music). This can be taken as a rather elegant demonstration that the communication of emotion does not require the processing (by either performer or listener) as 'music *qua* music'. For that to happen (as indicated by, for instance, a listener judgement that the performance is 'musical') further sources of variation must be inserted (those that Juslin *et al.* characterize as 'generative' and 'motional'). Note that emotional expression remains important. Taking this out reduces perceived musicality; but to influence judgements of musicality it must act through and with other sources of variability, including those generative sources which are intimately bound up with the structural (e.g. syntactic and segmental) characteristics of the music.

This study demonstrates important listener effects by computationally manipulating the presence or absence of formally defined sources of variance. This does not in itself demonstrate that human performers manipulate these parameters independently, let alone under intentional control. Evidence does exist from a variety of sources that musicians can significantly alter their performances according to emotional and musical intentions (e.g. Davidson, 1993). The most convincing analyses of what this achieves (for listeners) has been provided by Juslin and colleagues in respect of the 'emotional expression' component of the GERM model. Substantial further work will be required to tie down the communicative parameters of other components (either of this model or yet to be formally characterized).

However, even this preliminary work allows a very clear conclusion. Performance effects of music on listener experience are multicomponential, with some components directly addressing basic or 'garden' emotional parameters, and other components (in combination or separately) affecting a broader range of judgements and reactions that listeners may legitimately make to music. Indeed, given the huge diversity between individual performers and performances that Juslin and others have documented, it would be somewhat strange if listener responses were not characterizable by an equivalent diversity.

## 22.6 Concluding remarks

My question, 'to what extent the research reported in this issue points to, explicates, and encourages an understanding of diversity and complexity in musical experience' has a clear answer. Although based on 'behaviourist/cognitivist' foundations, the present studies actually do point to, and explicate, this diversity in a range of ways.

What goes on when a listener engages with music is above all a function of how that listener construes the musical event. In particular, it matters if and how the listener believes that there is an intentional agent behind the music. From earliest infancy, we are set up to 'read' music-like stimuli as emanating from a human agent who has intentions and feelings towards us. In the earliest times, our concerns are probably limited. Is this a person I recognize and know, and do they mean well towards me? As we grow and develop our conceptual apparatus for construing music as indicative of agency grows and becomes more complex in ways which empirical science has perhaps not addressed as fully as it might.

As we grow, we come to realize that particular communications are called music, and others, with some shared characteristics, are not. Within the domain of music we come to make finer distinctions (distinctions of genre, function, quality, culture). We come to appreciate that much music is crafted (often with much sweat, pain, and commitment) by specific individuals to whom this matters more than almost anything else they do. We learn that certain pieces of music have been preserved lovingly by many generations, and that entire and costly cultural institutions have been set up to ensure that these works are available to all citizens and remain alive in the minds, hearts, and fingers of performers. We learn that some pieces of music are loved fanatically by some groups of individuals, and loathed passionately by others. We learn that some pieces of music make their progenitors rich and famous. We learn that some societies ascribe to music such power for potential harm that they strictly regulate it, or even ban it. All of this frames and potentiates our reactions, before we even switch on the radio, select a track, or walk into a church, concert hall, or disco. Our knowledge about music and its cultural position reaches down to the most mundane act of reception (e.g. half-hearing piped music in a shopping mall) and transforms it (see for instance, De Nora (2000).

We also come to know that, within the cultural market place, there are competing ideologies that use music as a battleground. Disputes rage about what is proper music, music of value, music worth teaching in schools, music worth subsiding with taxpayers' money. The role of academia in

general, and science in particular, is to be aware of all this, to document, and explain. Science, however, is also a cultural product, and its adherents are no more immune to their social and cultural background than anyone else. The science of music, with its particular origins, can be presumed to be under constant pressure to align with values of the dominant cultures which support and fund it. These cultures have predominantly economic imperatives which underpin science's aims to classify and de-individuate in favour if isolating broad trends and key marker variables.

These imperatives can lead to the deconstruction of the richness of a cultural object like a piece of music into a set of 'effects' whose impact on health, economic success, consumer behaviour, etc. can be measured, predicted, and controlled.

Fortunately or unfortunately (depending on one's perspective) the papers in this volume have demonstrated that the process of deconstruction is a hugely difficult task, and that any apparently firm statement regarding the emotional effects of music have to be hedged with so many caveats and qualifications as to significantly hinder the prospects for large-scale commercial exploitation of these findings. Individual variations, both between different pieces of supposedly 'similar' music, and between supposedly 'similar' people remain large, and mainly unexplained. Even mood induction often seems to require those submitting to the induction to 'try hard' to get into the required mood!

What we can say, with increasing scientific certainty, is that there are certain characteristics of the 'sound of music', which are shared by several non-musical stimuli (particularly speech, and gesture) which allow simple emotional messages to be 'read off' the sound surface, and which allow certain basic effects of arousal modulation to occur. That being so, much music may be used by some agents as something else than (and something less than) music. They throw away the kernel and market the husk.

There are moral and scientific reasons to adopt a critical stance to this programme. The moral reasons relate to the purposes and intentions of those producing the music. It is a signal disrespect to those who conceived of their music as addressing ultimate issues of life, death, love, identity, war and peace, to appropriate it purely for analgesia or profit. If music is to be used for such stripped-down purposes, let people who have these purposes make their own stripped-down products rather than diminish the work of others by recycling it. The scientific reasons relate to our duties to represent, not ignore, the huge variety of musical objects, activities, and contexts, which this world actually contains. Underlying laws and patterns may exist, but we are very far from having adequately sampled the universe

of these aspects. Even within those that we have sampled, the sheer richness individuality of each music experience must be among the most salient facts that we have to come to terms with. In doing music science we may indeed be pushing at the limits of what science can tackle or achieve.

I could be accused of coming dangerously close to implying that the present authors are catering to commercial interests simply because they attempt to deconstruct music and emotion into more manageable parts. This is certainly not my belief. As scientists we should, of course, be wary of letting others misuse musical emotion findings, but there is nothing inherently wrong with attempts to simplify a phenomenon to make it manageable to study empirically. The only people who can afford to take in everything at once may be those that rely entirely on armchair speculation. Any attempt to delineate various subcomponents represents a laudable step towards a more theoretically informed approach that can help to organize disparate findings. It is a basic principle of any theorizing and modelling that one must simplify or 'deconstruct' the essentials. In science there is no serious alternative. My comments should not, therefore, be seen as favouring a romantic notion of art as an inscrutable mystery, too complex to be ever dealt with properly by science. Rather, our attempts to deconstruct need to be constantly held up against the richness of everyday (and peak) musical experience to ensure that it is the full experience we are attempting to explain, and not some conveniently simplified portion of it.

## References

Bourdieu, P. (1977) *Outline of a theory of practice* (Trans R. Nice). Cambridge, UK: Cambridge University Press.

Cook, N. (1998) *Music: a very short introduction.* Oxford: Oxford University Press.

Davidson, J. W. (1993) Visual perception of performance manner in the movements of solo musicians. *Psychology of Music,* **21.2**, 103–113.

De Nora, T. (2000) *Music in everyday life.* Cambridge: Cambridge University Press.

Frith, S. (1996) *Performing rites: on the value of popular music.* Oxford: Oxford University Press.

Gabrielsson, A. (2002) Perceived emotions and felt emotions: same or different? *Musicae Scientiae, Special Issue 2001–2,* 123–148.

Kivy, P. (1999) Feeling the musical emotions. *British Journal of Aesthetics,* **39**, 1–13.

Juslin, P. N. (2001) Communicating emotion in music performance: a review and empirical framework. In P. N. Juslin and J. A. Sloboda (Eds.) *Music and emotion: theory and research.* Oxford: Oxford University Press. pp: 309–337.

Juslin, P. N. and Laukka, P. (2003) Communication of emotion in vocal expression and musical performance. Different channels, same code? *Psychological Bulletin*, **129.5**, 770–814.

Juslin, P. N., Friberg, A., and Bresin, R. (2002) Toward a computational model of expression in music performfnace: the GERM model. *Musicae Scientiae, Special Issue 2001–2*, 63–122.

Krumhansl, C. L. (1997) An exploratory study of musical emotions and physiology. *Canadian Journal of Experimental Psychology*, **51**, 336–352.

London, J. (2002) Some theories of emotion in music and their implications for research in music psychology. *Musicae Scientiae, Special Issue 2001–2*, 23–36.

Maslow, A. H. (1976) *The farther reaches of human nature*. New York: Penguin Books.

McClary, S. (1991) *Feminine endings: music, gender and sexuality*. Minnesota, MN: University of Minnesota Press.

Nyklicek, I., Thayer, J. F., and Van Doornen, L. J. P. (1997) Cardiorespiratory differentiation of musically-induced emotions. *Journal of Psychophysiology*, **11**, 304–321.

Parrott, A. C. (1982) Effect of paintings and music, both alone and in combination, on emotional judgements. *Perceptual and Motor Skills*, **54**, 635–641.

Rauscher, F. H., Shaw, G. L., and Ky, K. N. (1995) Listening to Mozart enhances spatial-temporal reasoning: towards a neurpsychological basis. *Neuroscience Letters*, **185**, 44–47.

Schubert, E. (2002) Correlation analysis of continuous emotional response to music: correcting for the effects of serial correlation. *Musicae Scientiae, Special Issue 2001–2*, 213–236.

Scherer, K. R., Zentner, M. R., and Schact, A. (2002) Emotional states generated by music: an exploratory study. *Musicae Scientiae, Special Issue 2001–2*, 149–172.

Sloboda, J. A. (1989) Music psychology and the composer. In S. Nielzen and O. Olsson (Eds.) *Structure and perception of electroacoustic sound and music*. Amsterdam: Elsevier. pp. 3–12.

Sloboda, J. A. (1999) Music performance and emotion: issues and developments. In S. W. Yi (Ed.) *Music, mind, and science*. Seoul: Seoul National University Press, pp. 220–238.

Sloboda, J. A. (2000) Music and worship: a psychologist's perspective. In T. Hone, M. Savage, and J. Astley (Eds.) *Creative chords: studies in music, theology and Christian formation*. Leominster: Gracewing, pp. 110–125.

Sloboda, J. A., and O'Neill, S. A. (2001) Emotions in everyday listening to music. In P. N. Juslin and J. A. Sloboda (Eds.) *Music and emotion: theory and research*. Oxford: Oxford University Press.

Small, C. (1998) *Musicking: The meanings of performance and listening*. Hanover, NH: Wesleyan University Press.

Trehub, S. E., and Nakata, T. (2002) Emotion and music in infancy. *Musicae Scientiae, Special Issue 2001–2*, 37–62.

Vastfjall, D. (2002) A review of the musical mood induction procedure. *Musicae Scientiae, Special Issue 2001–2*, 173–212.

Waterman, M. (1996) Emotional responses: implicit and explicit effects in listeners and performers. *Psychology of Music*, **24**, 53–67.

CHAPTER 23

# ASSESSING MUSIC PSYCHOLOGY RESEARCH: VALUES, PRIORITIES, AND OUTCOMES

## 23.1 Introduction

What criteria can we use to evaluate research in music psychology? Do these criteria provide any guide for researchers, or prospective researchers, wishing to decide what to research, or whether to do research at all?

These are the questions which motivate this chapter, and, in exploring them it will be necessary to address broader questions, to do with the social responsibilities of scientists, academics, and educated citizens. How do we decide where to devote our energies? Can we make such decisions in rational ways, which optimize both our own personal fulfilment, and our need to earn a living, but also address the needs of society? These questions are hard to answer, but I believe we all have a responsibility to think about them to the best of our abilities.

I write this chapter in the main because these issues have been facing me personally and relentlessly in recent years. Although I am concerned to offer a systematic and informed treatment of these issues, I will, at some points, need to be more personal than is customary in academic discourse. It seems to me that in this particular area of debate the personal is appropriate, and maybe even necessary. Attitudes towards values and priorities are so intimately connected to a person's social context and background that removing them from the discourse can hide key assumptions that motivate the argument, and thus obscure the basis of the argument as rooted in the realities of lived experience.

Many intellectual and moral issues which 'come to a head' at a particular point of time do so because they have been 'building up steam' for a long time. Although dilemmas about the value of research came to dominate my concerns in the last half decade, they have been there in a less urgent form for considerably longer.

A question which I have posed repeatedly to anyone who will listen is: suppose all the music psychology in the world had never been written,

and was expunged from the collective memory of the world, as if it had never existed, how would music and musicians be disadvantaged? Would composers compose less good music, would performers cease to perform so well, would those who enjoy listening to it enjoy it any less richly?

In all my years of asking this question, I have to admit to never having received a fully convincing answer that was truly specific and substantial. Some respondents have pointed to the intrinsic interest of music psychology research, and the fact that many musicians and music-lovers seem to find it interesting to learn something about music psychology. However, interest or curiosity is not a very substantial outcome. I can be interested in the personal lives of media celebrities. This does not of itself make their lives useful, to them or to me.

What is even more alarming than the absence of convincing answers is the sense that to many who have been posed it, this just doesn't seem an important enough question to have any sort of ready or worked-out answer. This appears to be, in part, because there appears to be little or no well-developed literature where such questions are framed and answers attempted, certainly not in music psychology. Social benefit is not a consideration in the training of many psychology researchers (perhaps even most). Ethical considerations tend to be confined narrowly to the treatment of experimental volunteers (for instance, British Psychological Society, 1993).

The lack of attention to social benefit may also be a consequence of employment culture. Schmidt (2000) has suggested that professional employment, such as that offered to most academic researchers, can often subvert and distort the idealism and social responsibility of professionals, subordinating their work to the narrow political and economic goals of their employing organizations. Organizational culture (including a tendency to burden professionals with cripplingly high workloads) can inhibit professionals from seeking to make the meaningful difference to society that may have first inspired them to join their chosen profession. Few university researchers are explicitly encouraged to put social benefit at the top of their work agendas.

## 23.2 The social and personal benefits of music

One area in which there has now been substantial research investment is that of the effects or benefits of music engagement on non-musical aspects of health, well-being, and psychological functioning.

The main aim of much of this research is to document the nature of the benefits that are experienced through music engagement. In the most notorious case (the 'Mozart effect'; see for example Hetland, 2002; Rauscher, 2002) it has been discovered that exposing people to certain

kinds of background music improves their performance on some cognitive tasks, even though they many not be paying focal attention to the music. Other more complex studies have demonstrated that involvement in music education programmes improves intellectual and social outcomes for children across a range of measures.

Some of the evaluated uses of music are genuinely new, as for instance the use of recorded music in the pre-operative stage of surgical procedures, as a means of enhancing anaesthetic effects (see, for instance, Nilsson *et al.*, 2003). Other research focuses on pre-existing types of musical interventions, and examines effects in a systematic way. For instance there are now a number of studies (Costa-Giomi, 1999; Schellenberg, 2004) which show clear effects of traditional instrumental instruction on intelligence.

It is of considerable significance that this research is being undertaken against a backdrop of cutbacks and retrenchment in school music provision in a number of countries. The research is therefore part of a wider 'advocacy' movement (Hallam, 2000; ISME, 2003) to save music education from further cuts in public spending and to argue for a restoration, or even an increase, in the availability of music education within populations.

However, there are a number of issues to address before we can unambiguously propose that such research has demonstrable social benefit.

### 23.2.1 Demonstrating benefits versus producing benefits

The first issue relates to a potential confusion between 'research which demonstrates benefits' and 'research which produces benefits'. A very clear example of research which demonstrates benefits is given by Schellenberg (2004). He tracked 144 schoolchildren who were randomly assigned to four different year-long educational enrichment programmes, keyboard lessons, voice lessons, drama lessons, and no lessons. By comparing WISC-R IQ measures before and after the intervention, he demonstrated improvements in IQ for the two music groups which were significantly greater than for the other two groups.

In this research, Schellenberg did not devise or employ innovative or unusual music education programmes. These were 'off the peg' programmes, within well-established pedagogic traditions, that have already benefited thousands of children. His contribution was to undertake research in the tradition of the 'random controlled trial' model perfected within medical research, to prove a causal relationship between existing well-used musical interventions and a socially beneficial and possibly non-obvious outcome (increased IQ).

Schellenberg's research is good and original science. That is not in doubt. However, the research itself has not, at this point, yielded any new

social benefits. It has simply documented existing benefits. For the research itself to have benefits, it would, for instance, have to be influential on the decisions and resource-allocation activities of providers and consumers of music education programmes, so that the number of children participating in them significantly increased over a period of time.

### 23.2.2 Research as documentation versus research as advocacy

This observation moves the argument to a second set of issues. These are concerned with the process of turning research findings into socially beneficial outcomes, and the role of the researcher in such processes. Let us examine some of the potential difficulties that arise. Paramount among these is the unpredictability of the relationship between the publication of a research finding and the realization of its social benefits. It could be plausibly argued that the primary responsibility of a scientist is to publish findings at the earliest opportunity. However, changes in policy and other social arrangements which might enshrine the results of research in new outcomes might take years, or even decades, to bring about. They might never be realized. For instance, research which has showed the viability and environmental benefits of motor traction engines that do not use fossil fuel has not led to the large-scale development of such engines. This in, in part, because the development costs of the necessary infrastructure are so substantial. For instance, the BMW corporation has developed a prototype 'Hydrogen Car', but there are only currently a handful of filling stations in the world where such cars could be refuelled (Dunlop, 2001). There is always a development cost to the implementation of any new research finding. Resources (of time, money, advocacy) must be committed up front before the predicted benefits will start to flow. Researchers, as researchers, can document these processes, but in order to play a role in the social implementation of their research, they must step out of the researcher role into the advocacy role, to assist in mobilizing the necessary resources.

The question which these considerations lead to is this. How can a scientist plan, conduct, and report research in a fashion which does not compromise principles of objectivity, but which maximizes the likelihood of mobilizing the necessary advocacy to realize whatever potential social benefits accrue to the research?

One model, which is increasingly available to researchers in the more technological sciences, is the business model. Such a model is, indeed, at the heart of policy in many corporations and institutions that employ researchers. Agencies exist to assist researchers to protect their intellectual property and to exploit it in conjunction with entrepreneurs, manufacturers, and those in the business of selling and advertising. Researchers are

increasingly advised by those who employ them that if they suspect that a line of work on which they are engaged might have commercial spin-offs, these potential spin-offs are identified at the earliest opportunity, so that the research can be 'shadowed' and 'championed' by those more expert than the researcher in advocacy and development. This model is generally most appropriate when the outcome of the research is a tangible product, such as a new technology, a new invention, or, in psychology, a new psychometric test. It is generally pursued with most vigour when institutional managers detect that there is a high potential for short-term or medium-term financial rewards.

Although such a model may be relevant to some music researchers, operating at the more technological and 'product-oriented' end of our discipline (for instance in research which has direct implications for sound-engineering or sound-reproduction) it appears largely irrelevant for the greater bulk of research in music psychology.

A search of two major research databases (CAIRSS and PSYCINFO) using the search terms social+benefits+music, and conducted on 25 November 2003, yielded 14 articles in total for the period 1983–2003. These papers are listed in Table 23.1.

Although this is not necessarily a comprehensive list of all research on social benefits (see for instance Hallam, 2000), it can be seen as a representative sample of the extent and topic spread of research studies in which the social benefits of music are directly and explicitly addressed by the researcher. The absolute number of studies is low: as would be expected from informal observations about the low explicit priority given to social benefits within the research community. The articles all fall within four clearly defined topic areas:

- care of the elderly (5 items)
- child development (3 items)
- music in education (2 items)
- therapy (4 items).

Most of the studies describe complex behavioural applications involving music. They examine 'ways of behaving' with challenging, disabled, or needy subgroups, that are suggested as having subtle 'quality of life' benefits. Few of these subgroups are likely to form powerful self-advocacy organizations, and few are clearly defined consumer groups that would attract the obvious interest of the business community. The tools offered within these research studies are largely addressed to 'carers' whether professional or voluntary. Several of the carer groups identified are small or relatively low-status groups (e.g. music therapists, elderly-care nurses). There is unlikely to be

**Table 23.1** Research on social benefits of music

| Title | Author(s) | Citation details | Topic |
|---|---|---|---|
| Music and quality of life in older adults | D. D. Coffmann | *Psychomusicology*, **18.1**, 2002, 76–88 | Care of the elderly |
| Neuroscience applications in marital and family therapy | E. A. Tootle and L. Sperry | *Family Journal: Counselling and Therapy for Couples and Families*, **11.2**, 2003, 185–190 | Therapy |
| The magic within the music: exploring the use of music in psychotherapy with adolescents | C. D. Diraimondo | *Dissertation Abstracts International: Section B: the Science and Engineering*, **63(5-B)**, 2002, 2578 | Therapy |
| The contribution of music to positive aging: a review | T. Hays, R. Bright, and V. Minchello | *Journal of Aging and Identity*, **7(3)**, 2002, 165–175 | Care of the elderly |
| First-time mothers' use of music and movement with their young infants: the impact of a teaching program | W. Vlismas and J. Bowes | *Early Child Development and Care*, **159**, 1999, 43–51 | Child development |
| *Talent Development IV: Proceedings from the 1998 Henry B and Jocelyn Wallace National Research Symposium on Talent Development* | N. Colangelo and S. G. Assouline (Eds.) | 2001 (publisher not given) | Music in education |
| Popular music styles in Nordoff–Robbins clinical improvisation | K. S. Aigen | *Music Therapy Perspectives*, **19(1)**, 2001, 31–44 | Therapy |

| | | | |
|---|---|---|---|
| The use of music therapy to address the suffering in advanced cancer pain | L. Magill | *Journal of Palliative Care*, **17(3)**, 2001, 167–172 | Therapy |
| Music therapy: enhancing communication between family caregivers and their loved ones with dementia | C. I. Gardner | *Dissertation Abstracts International*, **60(60-A)**, 1999, 1953. | Care of the elderly |
| Long-term benefits of music intervention in the newborn intensive care unit: a pilot study | J. M. Standley | *Journal of the International Association of Music for the Handicapped*, **6.1**, 1990–1, 12–22 | Child development |
| Play as improvisation: the benefits of music for developmentally delayed young children's social play | A. S. Gunsberg | *Early Child Development and Care*, **66**, 1991, 85–91 | Child development |
| The unique power of music therapy benefits Alzheimer's patients | S. Smith | *Activities, Adaptation and Ageing*, **14.4**, 1990, 59–63 | Care of the elderly |
| It's cool because we like to sing: junior high school boys' experience of choral music as an elective | M. A. Kennedy | *Research Studies in Music Education*, **18**, 2002, 24–34 | Music in education |
| Music and movement for the elderly | M. Morris | *Nursing Times*, **82.8**, 1986, 44–45 | Care of the elderly |

any potential for 'patenting' or 'commercial exploitation' of the research findings reported. Therefore, the 'business model' for advocacy is likely to be inappropriate, without considerable modification.

The constituencies ready to undertake advocacy in such areas would seem to those whose principal motivations are humanitarian, ethical, and compassionate rather than commercial.

### 23.2.3 Commissioned policy-relevant research

A research model which is increasingly used by government and larger corporations in the commissioning or supporting of research which aims directly to address key policy questions. Quite small organizations can also commission developmental or evaluative research on their activities. Two projects I have been recently involved in illustrate different applications of this model.

Between 1998 and 2001 a research project on Young People's Participation in Music took place, with funding and support from a major grant awarding agency (the Economic and Social Research Council of Great Britain (ESRC), and *an Advisory Committee drawn from* key agencies and organizations concerned with music participation in young people (see *http://www.keele.ac.uk/depts/ps/ESRC/esrcmenu.htm*). Some results from this project are outlined in Sloboda (2001—reprinted as Chapter 21 of this volume) and Ryan *et al.* (2000).The social issue which drew this constituency together was the experienced decline or dropout of participation in active music-making (particularly the playing of musical instruments) by British young people in their early adolescence. A better understanding was sought of the factors which influence long-term commitment to instrumental performance, in order to develop policy and practice in the area of music education. The advisory board worked with the researchers to develop and improve the research protocol (involving questionnaires and interviews to young people, their parents, and their teachers). Following the completion of the preliminary data analysis, the research team worked with members of the advisory board to produce and publish a set of 'Practitioner and Policy Implications' which were agreed to emerge from the research findings and which could be addressed to teachers and educational policy-makers (reproduced in the Appendix to this chapter). These have been widely distributed, and have fed forward into other discussions at national level about music education policy.

This level of engagement with policy-makers is, to our best knowledge, unusual within the music psychology confraternity. It has ensured that some of the ideas contained within the research have influenced subsequent policy debates, at least within the UK. The project team have also

taken as much trouble to disseminate these findings to groups of practitioners around the world as to more traditional academic audiences, undertaking what could be described as direct advocacy work alongside the more traditional 'objective' reporting of research results.

However, even with such an explicit focus on policy and practice it is not easy to prove that the research has contributed (or will contribute) to tangible social benefits. Such benefits would only accrue some time after significant increases in the number of agencies and individuals operating in accordance with these implications, and would probably only be measurable in relevant outcomes some years later. Even if such increases could be quantified, attributing them directly to this research project would be highly problematic. Many practitioners and policymakers may have operated on the basis of such principles without any knowledge of our specific research findings or recommendations. In general, we would have to fall back on a general faith that spreading sound ideas into relevant communities is likely to have more positive social benefits than not doing so.

The second project that illustrates a 'commissioning' approach, was a focused piece of evaluative research within a single educational establishment (Sloboda *et al.*, 2003). This research was designed to address a commonly observed problem among high-level performers, which is the unreliability of achieving optimal levels of expressive performance in repeated performances of the same piece of music. Even highly rehearsed performances can fail to 'come alive' for reasons which are often unclear to the performers themselves. A pedagogic approach to this problem, entitled 'Feeling Sound' had been developed by one of the authors (Gayford), and piloted in workshops with both amateur and professional musicians. This involved teaching participants a set of attentional tools for focusing more effectively and consistently on expressive (in addition to technical) parameters of the performance.

The researchers worked with institutional managers of a tertiary-level college of music to raise funds for the research, and recruit participants from among its students. Key pedagogic leaders within the institution assisted in student recruitment. The study itself was based on the random allocation clinical trial method, where volunteers, matched for gender and year on the course, were randomly assigned to either an experimental or a control group, Students in the experimental group received the 'Feeling Sound' training, students in the control group were placed on a waiting list for potential enrolment in a future course. Level of expressive performance of a test piece was assessed twice for all participants, once before the intervention, and once after it. A significant advantage was found for the experimental group as compared to the control group. In some cases, the

intervention led to radical changes both in the way that participants reported thinking and feeling about music, and also in the quality of their playing as observed by their instrumental teachers.

The measured effects of the course on performance-quality ratings were modest, and several participants observed that, even though the course had lasted for an entire academic year (9 months), they felt they were only at the beginning of incorporating the implications of what they had learned into their playing. Several participants predicted that the substantial gains of participating in the programme were still to be realized, some way into the future.

From a 'social benefits' perspective, the danger of this study (which is, in some ways quite typical of many intervention studies within the behavioural sciences) is that its benefits may not spread beyond the small group of individuals participating in the study. Funding problems have meant that the intended repetition of the course for a second cohort of students within the college has not taken place. Also, only one person (the progenitor of the method) is really qualified to teach such a course. For a method like this to achieve its full benefit, more teachers need to be trained and more institutions need to hire teachers to work with their performers. Creating the conditions in which such outcomes are made more possible takes one well beyond the realm of research *per se*, and into the arena of advocacy, organizational development, and fundraising. Few researchers and practitioners have the necessary skills to bring this about. As a result, many such initiatives fade away when the specific trial is completed and specific personnel move on, leaving little trace in the community as a whole.

The approach to these problems currently being developed within the 'Feeling Sound' team, is the creation and dissemination of a set of resources which can be used and modified by practitioners within their own contexts without the need for the direct involvement of the method's progenitor. Time will tell whether this is a productive approach.

### 23.2.4 Contextualizing benefits: ensuring that one person's benefit is not another person's loss

There is an ambiguity to much research carried out in a 'first world nations' context. This ambiguity has been noted in respect of other conditions which have sometimes been labelled 'diseases of affluence'. For instance, a great deal of medical research is aimed at problems caused by the over-consumption of the high-fat foodstuffs cheaply available in affluent nations (and increasingly in developing countries; see, for instance, James, 1998). Obesity is an outcome of the huge disparities in distribution of the world's resources. Medical research into obesity, and programmes for the obese,

have been criticized by some as addressing the symptoms of a social problem rather than its underlying cause. Obesity would cease to be a major social problem if there was a more just distribution of the world's resources, so that overconsumption was a financial impossibility for most people. Medical research on obesity may be seen as accepting and bolstering the social *status quo* by putting more resources in the way of people who already have too much.

Several influential organizations and foundations have publicly argued that the priority for health researchers (and those bodies which fund healthcare) should now be an exploration of the ways in which the cheap and effective advances in healthcare of the last century (now taken for granted in every developed country) can be extended to be available to every person in the world. A few low-cost and high-outcome remedies available to every person on this planet would be far more cost-effective over a shorter timescale, in terms of lives saved, than further development of high-cost procedures for low-frequency conditions (as was argued, for instance, by Bill Gates at the 2002 World Economic Forum).

A similar argument could be proposed for music. The availability of music education within the developed nations (including North American, Europe, and Australasia) is far greater than elsewhere in the world. Even within the developed nations, the level and extent of provision within the countries where most research is undertaken (the USA, Canada, the UK) exceeds that of others (such as Greece or Portugal). When put alongside the music education opportunities available to a Rwandan child, or an Indian child, what is available in the West is truly munificent. These facts would incline some to argue that any advocacy might be better deployed at providing proper schooling for children in these developing and war-torn countries, worrying further about advanced nations only when the rest of the world had caught up.

It could be further argued that the apparent shortcomings of music education in the West are also a result of the peculiarities of the affluent consumer society. Opportunities for music may have to be created in schools within advanced nations because such activity has almost ceased to exist within the home and everyday society. The fragmentation of sociability brought about by an individualistic and media-dominated lifestyle, where families and individuals conduct their personal lives 'behind closed doors' and 'in front of the TV' rather than in the street and in the fields, means that the natural social contexts in which group music-making has flourished for tens of thousands of years, without the intervention of educators or psychologists, have withered away (see Sloboda, 1998 (this volume Chapter 19); Sloboda, 2001 (this volume Chapter 21).

It might be argued that more of the full benefits of musical engagement are still felt by those societies where day-to-day sociability and the pooling of physical labour is still a necessity. These societies tend to be among the less affluent of the planet. One could argue that if the societies in which music psychologists flourished became less over-affluent, music would naturally take root again in people's lives. Increasing school music is trying to patch the symptoms without addressing the real problem, which is the dysfunctionality of the society in which citizens of developed countries live.

From this kind of perspective, it is no longer enough to identify potential benefits of research. One also needs to ask whether the benefits outweigh the costs. The training of a researcher is an expensive business—it takes many years and involves the resources of many institutions. Specific research projects can also easily cost thousands, or hundreds of thousands of pounds or dollars, much of which is often devoted to salary payment. Developing and implementing research findings can require considerably more resource.

Even more problematical is identifying the 'what if' costs. What if the money and resource that had been devoted to training and supporting music psychology research had instead been put into research on a cure for cancer, or had been invested in a regional development programme? How can we assess potential benefits of hypotheticals?

At a personal level I have constantly been struggling with the issue of justifying what I do for a living, in a world where there appear to be so many more pressing issues requiring the resources of collective human intelligence. In the context of the state of the world, the question that poses itself to me, repeatedly, and in increasingly stark form is:

> How can I allow myself the luxury of researching music psychology, when half the world's population lives below the poverty line, when millions of children die of malnutrition and avoidable disease every year, when the activities and consumption patterns of the rich world are destroying the ability of the planet to sustain human life into the long term future, and when violence between people grows increasingly dangerous as the means to inflict violence become every more lethal and large-scale.

These interlinked problems are recognized, at least in part, by almost all leaders on the world stage. Xavier Solana, EU High Representative for Commmon Foreign and Security Policy, recently wrote:

> Almost 3 billion people, half the world's population, live on less than 2 Euros a day. 45 million continue to die every year out of hunger and malnutrition. Sub-Sahara Africa is poorer now than

> it was 10 years ago. In many cases, the failure of economic growth has been linked to political problems and violent conflict. In some parts of the world, notably Sub-Saharan Africa, a cycle of insecurity has come into being. Since 1990, almost 4 million people have died in wars, 90% of them civilians. Over 18 million people worldwide have left their homes or their countries as a result of conflict. (Solana, 2003)

A secondary question is:

> If I can allow myself the privilege of conducting music psychology research, how is that privilege best and most responsibly exercised so as to be more than 'harmless' (which I believe almost all music psychology research to be), but actually contributing in a direct way to the greater human good?

## 23.3 The personal and the political in music psychology research

Most human decisions can be accounted for, in large part, by an analysis of social and historical context in which people live, and the ways in which events and social forces shape them as individuals. Researchers are no different. The choices made by researchers reflect their social contexts. There are no context-free tools for making completely rational decisions about how to expend one's energies. It would be dishonest, and incomplete, to present the ideas of this chapter as flowing from a disinterested and leisurely rational analysis. These considerations repeatedly and disruptively forced themselves upon me during a period of my life when societal events were generating enormous emotional charge, in my immediate circles and also in society at large. I was forced to rethink my priorities as a result of the huge waves of anger, sorrow, despair, guilt, compassion, mingled with hope and determination, that have swept through the air around me these past 5 years. To understand these processes an excursion into politics and biography is required.

### 23.3.1 Responding to world events—a personal account

In the 1980s I was active in the anti-nuclear movement. This was an international movement to encourage the two major nuclear powers of the world, the USA and the USSR, to make real and lasting mutual steps to dismantle their nuclear arsenals, whose use could signal the end of human civilization. We believed (and I still believe) that the only way this would

happen would be for one side to make a unilateral step, in the hope of encouraging the other side to do likewise. This would lead to further steps yielding a progressive mutual reduction in tension and danger. The reality we faced was that each side refused to move until the other side made the first move. This created a stalemate which meant that, far from disarming, both the USA and the USSR continued building up their arsenals until the point at which the USSR suddenly and catastrophically broke apart.

There was a peculiar kind of security in the cold-war world. The balance of power seemed to provide a crazy stability, based on the maintenance of a very low likelihood of a very terrible nuclear cataclysm. Within this global system, the anti-nuclear movement rarely caught the imagination of large swathes of the population, certainly not in the West, and it was perceived as marginal, ineffective, and 'communist', at least in the UK where I have always lived and worked. It never occurred to me for a moment that involvement in this movement could or should be central to my life. It was something I did weekends and evenings when other people played golf or did their garden. It had little or no direct impact on my research work or my career. It was part of my 'charitable work'.

The break-up of the Soviet Union, far from ushering in world peace, led to a world which was awash with weaponry of all sorts, and where social fragmentation along ethnic and other lines was leading to greater instability within regions and within states. It was almost as if the USSR–USA stalemate of the 1950s to the 1990s had kept the lid on a whole set of simmering social and political problems which began to boil over as soon as the lid was lifted. The certainties of the post-war period were swept aside, and although there were some hugely exciting and positive developments, such as the liberation of the South African people from Apartheid, and the re-democratization of Eastern Europe, they were soon overshadowed by new lethal threats.

The new threats were ones which in the main did not emanate in any simple way from the actions of nation states, nor were they confined to specific countries or regions of the world. 'Globalization' became an increasing reality, spreading both its benefits and its dangers into every corner of the world. The AIDS epidemic was perhaps one of the first phenomena to make it real to people that globalization was real, and potentially death-dealing. Increasingly the threat to the world's environment brought about by human consumption and pollution, spurred and moulded by the increasingly powerful multinational corporations, came to be recognized as real. The end of human life as we know it became something which was not a low probability outcome of nuclear cataclysm, but a certain consequence of continuing doing exactly what we were doing (for example,

through global warming and the depletion of the ozone layer). Many of the environmental challenges the world faces come about through the increasing disparity in wealth between the few percent of the world's richest people, and the majority of the world's poor. The profligate consumption of oil and other commodities by the people of the advanced nations, particularly the USA, destabilizes not only the environment, but the very means to life of those in the poorest parts of the world. Deprivation fuels anger and violence, and the 1990s saw a significant increase in the number of regions of the world where ethnic and tribal violence flared up in conflict of one variety or another.

Although Western leaders and opinion formers liked in general to characterize these new civil wars as evidence of 'ancient ethnic hatreds' which were bound to resurface, and proof of the 'uncivilized' nature of these peoples, and their needs to be brought 'democracy' and 'control', a more critical (and I believe accurate) reading recognizes that whenever people with little power and resource are deprived or oppressed by a greater power, it is tragically easy for the oppressed to turn against each other, rather than uniting with one another to challenge the oppressor. People fight one another for control of those small resources allowed to them by the higher powers, and in doing so, leave those powers freedom to extend their power even further.

My growing acceptance of this perspective on the world's problems was helped along by exposure to systems and people who were able to clearly articulate a world view which has been labelled as radical (although it is seeming much more mainstream these days). I have been particularly influenced by the writings of Noam Chomsky (2003) and George Monbiot (2003). I also gained much from 15 years of involvement with a radical self-help movement called Re-Evaluation Counselling, founded by the American activist Harvey Jackins (1963). A friendship network containing many strong radical feminists also assisted my convictions to mature. For most of the 1990s these perspectives trickled into my research work in rather unobtrusive ways. They were partly responsible for my increasing interest in social and motivational factors underlying the acquisition of musical skills, and the search for wider cultural explanations for the unequal distribution of 'talent', analogous to the explanations being developed in radical politics to explain the unequal distribution of wealth within societies. Still, however, I was able to see a very clear distinction between my career, as a music psychologist, and the rest of my life in which critical social perspectives took other forms.

The events that began to unravel my carefully constructed compartmentalization of my life began on 24 March 1999, when, without the

authority of either Parliament or the United Nations, my country, along with 18 other members of NATO, formed an alliance to start the first European air war since 1945. During this 78-day war, they dropped 23 000 bombs on Serbia and Kosovo, more than were dropped in Britain during the entire Second World War. Over 500 civilians were killed in the air-raids, thousands more were injured. Depleted uranium and cluster bombs were used in large numbers. These proscribed weapons will continue to maim and irradiate for generations. This was no feared hypothetical nuclear exchange. This was actual, unprovoked, old-fashioned, conventional invasion of one sovereign state by others, the kind of war which nuclear weapons were supposed to have rendered obsolete, particularly in Europe, the presumed 'theatre' for nuclear war.

The immediate and total outrage I felt towards my government was, luckily, shared by many. We disbelieved the claims of Tony Blair and Bill Clinton that there was no alternative. We remain convinced that events were manipulated from the White House to ensure that no matter what concessions Milosevic made, war would happen. We were appalled that war could be so readily accepted as an answer to an admittedly real problem by so many intelligent people, when the entire Soviet empire had crumbled only a few years earlier without a single bomb being dropped or army mobilized. We were shocked that people could so easily be convinced that replacing one ethnic cleansing (of Albanians by Serbs) with another (of Serbs, Jews, Roma and all other non-Albanians by Albanians) could be accepted as any kind of 'victory' for the people of the region. We were horrified that after the end of the war, a punitive sanctions regime was imposed on Serbia by NATO and the EU, which denied hospitals the most basic medicines and supplies, harming the very people that this war was supposed to help. My own sense of outrage and injustice was so complete that I joined a delegation to Serbia to help document at first hand the ruinous effect that the war and sanctions had had on the civilian population of Serbia, particularly the children. I found myself unable to sleep at nights, and often unable to concentrate on my work, which seemed trivial, irrelevant, and self-indulgent. I was able to focus some of my energies by researching and writing about the worst failings of the West in the Kosovo crisis (Sloboda, 2001).

For the first time in my life I realized, fully and on every emotional level, that the direction that the leaders of the West were taking the world, with money taken from my taxes, was diametrically opposite to the direction in which I wished them to go. These were not people who shared my world view, with a few wrinkles separating us. These were people whose notion of the world, and where it should be going, had diminishingly few

points of contact at all with mine, or with those of the people I considered to be my 'reference group'.

I became convinced that the 'new interventionism' foisted on the world by the USA and Britain, was part of an immensely dangerous wedge, which would augur instability and tribulation on a massive scale. I was in absolute accord with the words of the British journalist Marc Almond (1999) who said: 'Even those who wholeheartedly endorse the humanitarian justification for the Nato action must be aware that, when the first cruise missile smashed through the thin skin of international law, the world's door was opened on to a void.'

This void was, I feared, to be filled by those whose argued that if the world's most powerful and most heavily armed states were not going to play by international rules from now on, then anything was fair game for those with less power but growing hatred and frustration at the privations and humiliations imposed on them by the 'global world order'. A self-destructive cycle of local instability, terrorism, reprisals, more terrorism, and even greater reprisals seemed to me inevitable. The victims of all of this would be the ordinary people, particularly the world's poor, in whose name this violence would be enacted, while actually making life worse for them.

My worst fears were realized on 11 September 2001. If March 1999 lit a fuse within me, 9/11 was detonation. Within hours of the fateful event, my reading of the words coming from the White House convinced me that the US administration was pulling out of its back pocket its long-harboured plan to engineer total military, economic, and cultural control over the Middle East and Asia, in order to bring about the 'Project for the New American Century'. Within days of 9/11 I knew that the main focus of my life had to change from music psychologist to activist for global peace and justice, and I soon shared that commitment with some of my closest friends. If I could not achieve such a change, I felt I would fall apart, morally, mentally and spiritually.

At that time I made the personal commitment that I would undertake no significant writing project in music psychology, or on any other topic, without conscious and deliberate linkage to issues of global justice, through which exploration, articulation and promotion of alternatives to prevailing societal norms could be addressed. My first, tentative, enactment of that commitment is to be found in Sloboda (2002, Chapter 22 of this volume). A second piece entitled 'The War, What War? Psychology in Denial' (Sloboda, 2003), offers a more trenchant and forthright critique of psychology's general failure to respond in any substantive way to the main historical trends outlined above.

At the time that I first made my life commitment to peace and justice, I had no idea how or when I would manage to combine this with the need to earn a living and support my family. However, I made the assumption that I might have to abandon music psychology completely, and so began putting my academic house in order. The book of which this chapter is part was conceived as one part of that tidying up, a coda to a 30-year phase of my life which was about to come to an end.

More than two years on from 9/11, a lot of bad things have happened in the world, and my responses to those bad things have developed in a range of directions that I certainly could not have predicted. A few good things have happened too, one of which is the phenomenal resurgence of a global anti-war movement on a scale which no one could have predicted before the Iraq crisis. This is part of a wider grass-roots mobilization which includes environmental campaigners and 'anti-globalization' social activists who challenge the damage done to peoples by the policies of the World Bank and other instruments of neo-liberal capitalism. This movement brought together a wealth of creative energies and initiatives which I became swept up in. Out of all this a potential solution to my personal dilemmas has been found in the form of a part-time job with a peace and disarmament non-governmental organization, which will ensure that, taken together with my evenings and weekends, the majority of my time will, indeed, be directly devoted to the causes to which I am committed. This leaves a proportion of each week still available to devote to academic pursuits, if I so choose. Where music psychology fits within this is something I shall no doubt be exercising myself about for some years to come.

### 23.3.2 A typology of levels of social engagement

It is only in the context of my personal journey that I am able to attempt some characterization of what different levels of socially responsible academic engagement might look like. It arises from the issues and dilemmas I have had to face in my own career. However, in making this characterization I am hoping to provide some kind of template against which other researchers could assess their own position, and become more explicit about the role of considerations about social benefit in their own work, rather than leaving such considerations implicit or unaddressed. I am sure that the template offered is crude and insufficiently nuanced, but its simplicity will perhaps be more productive in allowing a variety of elaborations and modifications which might make more sense within the very diverse set of personal and cultural experiences that characterize a research community.

The template postulates four successive levels of social engagement, which involve increasingly prominent roles for considerations impacting on academic activity 'from the outside'.

*Level 1—Sensitivity to historic academic norms*

Within academia there are generally accepted norms which, if adopted and followed in both letter and spirit, are usually held to contribute to the general social good. These include the belief that all knowledge is public knowledge, and the consequent imperative on academic researchers to communicate their knowledge through public channels. Academic freedom includes the freedom to promote and support unpopular conclusions, regardless of social or political pressures to do otherwise. Academics are enjoined to be truthful, and to provide evidential support to their communications, by, for instance the full citation of sources, and the availability of their raw data. They are enjoined to subject their work to informed critique through peer review prior to publication. Scientific researchers are required to consider the ethical impact of their research, specifically in terms of avoiding potential harm to those upon whom they experiment. Researchers who conduct themselves in an exemplary fashion within their chosen area of research can be said to be exercising *Level 1 social engagement.*

Researchers who depart from these norms, whether through carelessness and unprofessionalism, or for personal political or financial gain, cannot easily claim to be contributing to the social good in any way.

*Level 2—Sensitivity to applicability*

Researchers may choose to follow their intellectual interests in a range of directions, but retain a sensitivity to the possibility that some research might turn out to have practical or policy implications that could yield specific social benefits. If a researcher, on realizing potential applicability, takes actions designed to explore or promote that applicability, then they can be said to have moved to *Level 2 of social engagement.*

The range of appropriate actions is considerable. They may, for instance, undertake additional research to develop the applicability of their findings. They may seek to present the results of their research to practitioners, or discuss their findings with policy makers. On this model applicability is an occasional spin-off of a programme whose motivations are varied, but should include Level 1 social engagement.

*Level 3—Focus on applicability*

At this level, researchers make choices about what to research which are influenced by the prior assessment of potential applicability. If the applicability assessment is not favourable, the research does not get done.

This model is often the enforced model for researchers working for companies or government agencies. They are hired to research a particular problem, and their continued employment depends on the outcomes of their research being applicable. However some researchers in universities, who are free to choose the topics they research, may still choose a focus on applicability. This choice has, in many Western academic circles, at least until recently, been seen as likely to produce second-rate science. Pure 'blue-skies' research, has, it is often argued, more often encouraged the intellectual creativity which has led to theoretical breakthroughs, which then, often considerably later, begin to have practical applications. It can also be argued, in a similar vein, that it is not always possible to make an accurate assessment of the applicability of research in advance of conducting the research. Finally, it is often argued that applicable research is often that which is subservient to the less than admirable motives of funders and commissioners, which may often be flawed from an ethical point of view.

*Level 4—Focus on values*

Applicable research can be, and often is, socially unbeneficial. Research could be conducted, for instance, on how to use music to encourage people to buy and consume more alcohol. The research is clearly applicable, but hardly beneficial (other than to breweries and wine shops). In some contexts, researchers may well find that most research which they are being asked to carry out conflicts with their core values. In other contexts, resources which are being devoted to research may, while not directly conflicting with core values, be considered by the researcher, be more properly devoted to more urgent, or larger, social needs. In such cases the researcher may be drawn to seek other activities which express their values more directly. It cannot always be assumed that such considerations may be expressed within the specialist sub-discipline in which a researcher has gained experience. They may require changing research fields, or they may require stopping doing research altogether.

Perhaps the most frequent rebalancing of life priorities experienced by Western researchers is the 'career break' normally taken to raise a family. Considerable numbers of people, mainly but not exclusively women, possess value systems which put the needs of their children above the needs of

their profession. Not only is this accepted in many societies, but the right to do this without detriment to later career is enshrined in employment law. But the issues facing parents actually face everyone all the time. What are my priorities at this moment? Why am I doing what I do? How much of my life's resources (of time and energy) should I devote to any one activity, and why? What should I do for pay, and what should I do without pay—in what proportion? Of course there are constraints. Not everyone is free to decide these things afresh at every life juncture, but it can be argued that the maintenance of personal integrity and fulfilment requires that people assess the fit between their values and their actual activities on a reasonably regular basis, and make adjustments to the nature and balance of their activities accordingly. In these adjustments, very little can be accepted as an unchanging given. Expertise and experience in a given field is only one of a number of considerations that must be taken into account.

## 23.4 Concluding remarks

In relation to my own trajectory through music psychology, as depicted in the four roughly chronological sections in this book, I would want to claim that Part A (cognitive processes) describes work which is primarily operating at Level 1 (driven by sensitivity to academic norms). Part B (Emotion and motivation) is focused at Level 2 (sensitivity to applicability), Part C (Talent and skill development) shifts to Level 3 (focus on applicability), and Part D (music in the real world) increasingly represents Level 4 (focus on values). I am well aware that a typology that fits so relatively neatly onto my own intellectual trajectory might lack generality, and I am only too ready to have these ideas modified, or even swept aside, by robust contact with other models.

What I do hope, however, is that the debate about values will continue within the music psychology community, and become more central to our published work than it has been hitherto, at least within the concerns of those parts of the community with which I have had most contact.

## References

Almond, M. (1999) What have we started? The Independent, 6 June. *http://www.independent.co.uk/sindy/stories/C0606907.html*

British Psychological Society (1993) *Code of Conduct, Ethical Principles and Guidelines.* Leicester. *http://www.bps.org.uk/documents/Code.pdf*

Chomsky, N. (2003) *Hegemony or survival: America's quest for global dominance.* London: Hamish Hamilton.

Costa-Giomi, E. (1999) The effects of three years of piano instruction on children's cognitive development. *Journal of Research in Music Education*, **47**, 198–212.

Dunlop, N. (2001) BMW's paradigm shift: the Hydrogen Car. *http://www.edmunds.com/news/innovations/articles/46906/article.html* 20 July.

Hallam, S. (2000) *The power of music.* Performing Rights Society, London. *http://www.thepowerofmusic.co.uk/*

Hetland, L. (2002) Listening to music enhances spatial-temporal reasoning: evidence for the 'Mozart effect'. *Journal of Aesthetic Education*, **34**(3/4), 105–148.

ISME (2003) Advocacy web pages. International Society for Music Education. *http://www.isme.org/article/archive/25*

Jackins, H. (1963) *The human side of human beings. Seattle:* Rational Island Publishers.

James, P. Y. (1998) Diet-related diseases shift global burden. *Global Health and Environment Monitor*, **6.2**, 1–2. *www.ceche.org/monitor*

Monbiot, G. (2003) *The age of consent.* London: Flamingo.

Nilsson, U., Rawal, N., and Unosson, M. (2003) A comparison of intra-operative or postoperative exposure to music—a controlled trial of the effects on postoperative *Anaesthesia*, **58.7**, 699.

Rauscher, F. (2002) Mozart and the mind: factual and fictional effects of musical enrichment. In J. Aronson (Ed.) *Improving academic achievement: Impact of psychological factors on education.* (pp. 267–278). San Diego: Academic Press.

Ryan, K. J., O'Neill, S. A., Boulton, M.J. and Sloboda, J. A. (2000) Perceived social support and children's participation in music. In: *Proceedings of the Sixth International Conference on Music Perception and Cognition.* Staffordshire: Keele University.

Schellenberg, G. (2004) Music lessons enhance IQ. *Psychological Science*, **15.8**, 511–514.

Solana, J. (2003) *A secure Europe in a better world.* European Council, June 2003. *http://ue.eu.int/pressdata/EN/reports/76255.pdf*

Schmidt, J. (2000) *Disciplined minds: a critical look at salaried professionals and the soul-battering system that shapes their lives.* Lanham, Maryland: Rowman and Littlefield.

Sloboda, J. A. (1998) Music: where cognition and emotion meet. *The Psychologist*, **12.4**, 450–455.

Sloboda, J. A. (2001) So much expended for so little good. In: K. Dreszov, R. Waller and B. Gokay (Eds.) *Kosovo: the politics of delusion.* London: Frank Cass.

Sloboda, J. A. (2002) The 'sound of music' versus the 'essence of music': dilemmas for music-emotion researchers (commentary). *Special issue on Music and Emotion, Musicae Scientiae,* **6.3**, 235–253.

Sloboda, J. A. (2003) The war, what war? Psychology in denial. *DECP Debate.* British Psychological Society, Issue 107, 6–11.

Sloboda, J. A., Gayford. C. and Minnassian, C. (2003) Assisting advanced musicians to enhance their expressivity—an intervention study. In: *Proceedings of the 5th triennial conference of the European Society for the Cognitive Sciences of Music (ESCOM).* Hannover University of Music and Drama.

## APPENDIX: YOUNG PEOPLE AND MUSIC PARTICIPATION PROJECT*

*S. A. O'NEILL, M. BOULTON, K. RYAN, AND J. A. SLOBODA*

### Practitioner and policy implications

1 *Provide structured opportunities for young people which give them increasing choice, control and responsibility for organizing and developing their own musical involvement.* Findings from numerous parts of the study point to the importance young people place on choosing their own musical instruments, music, and musical activities. This is particularly important following the transition to secondary school where many young people feel that compared to primary school, they are given less opportunity to take responsibility for various aspects of musical involvement and decision-making. One can sense the frustration that this might cause as young people reach a point in their development where they seek to establish a greater sense of independence and control over the choices they make. Young people who continue to play instruments following the transition to secondary school are more self-directed and autonomous in the musical activities they are involved in compared to those who give up. They are also more likely to report having a classroom music teacher or instrumental music teacher who communicates a belief in their potential to do well and enables them to choose music they like and want to learn to play.

2 *Provide young people with instruments that they associate with valued role models.* There is a mismatch between the instruments children want to

*Funded by the Economic and Social Research Council, February 2002. See *http://www.keele.ac.uk/depts/ps/ESRC/Practitionerimp.doc*

learn to play and the instruments they actually play, particularly at school. This mismatch is especially prominent among boys. The majority of children in both Year 6 and Year 7 reported playing the recorder. This certainly affords opportunities since it is a small, light, inexpensive and relatively durable instrument that most children can learn to play and therefore gain experience of active music-making. However, children do not associate playing the recorder with their musical role models in the adult world. As such, they view the recorder as 'not a real instrument' or a 'child's instrument' which is limited in its ability to express the music young people are most interested in playing. Those who are most likely to continue learning to play an instrument report valuing the instrument they play and identifying positively with adult role models who play a similar instrument.

3 *Start learning to play an instrument younger, and at least one year before making the transition to secondary school.* The findings indicate that those children who are more likely to continue playing after making the transition to secondary school had already been playing an instrument for at least a year while they were at primary school. Starting an instrument younger, and having played an instrument for longer, appears to act as a protective factor in terms of continued interest and commitment toward the instrument at secondary school.

4 *Provide opportunities and encourage young people to play in musical groups outside of school hours.* Young people who have more close associates who play instruments in terms of family members and friends, and who engage in group musical activities outside school hours, are more likely to continue playing instruments following the transition to secondary school.

5 *Help to establish within young people positive beliefs and values about musical involvement.* Young people who continue playing instruments following the transition to secondary school have greater confidence in their own ability, and find playing more important and enjoyable than those who give up. They also don't mind a challenge and believe hard work will lead to improvements compared to those who give up.

6 *Provide young people with structured goals and appropriate challenges.* Children who have been involved in performance groups and have taken a music examination are more likely to continue playing instruments after making the transition to secondary school. However, there was no difference between those who continue and those who give up in terms of the highest grade achieved or the mark awarded

on exams. What appears to be important is not so much that children attain the highest levels on music exams, but rather that they gain a sense of having a structured goal to work towards that is viewed as a challenge or opportunity to improve their skills and not just as an opportunity to display their competence. In other words, many teachers involve their students in music examinations, but what appears to be important is that these examinations are not viewed as merely opportunities to display competence (i.e. achieve high grades). Rather, it appears more important for music examinations to be viewed by children as a challenge that will motivate them to learn more and make more progress.

7 *Young people are helped if they believe their parents are supportive of their involvement in musical activities.* The most important differential source of support comes from parents. Children who stay involved in learning to play an instrument believe that their parents are supportive of this activity far more than children who give up. Siblings' and friends' direct support is far less important, although it helps if friends believe in music as an activity for all. The support offered by teachers appears to make less of an impact, although it helps if the child feels that music is valued at the school, that the music teacher and instrumental music teacher are able to communicate a belief in the child's potential to do well at music and teach music the child values and wants to learn to play.

# BIBLIOGRAPHY OF JOHN A. SLOBODA 1974–2003

This is not a complete bibliography of the period, but contains all psychology publications that have appeared in normal academic outlets (e.g. peer-reviewed journals and books) as well as a representative sample of items concerning music that have appeared in outlets used by music professionals (e.g. educational and practitioner journals, and some serious general-circulation periodicals, such as the *BBC Music Magazine*). Submitted conference proceedings and abstracts are excluded.

The list is ordered chronologically by date of publication, and, within each year, by alphabetical order of co-author surname, where appropriate. Items are numbered in sequence for convenience of reference.

Items that are reprinted in this collection are indicated by an asterisk. Permissions to reproduce are listed at the beginning of this book (pp. xxvii–xxix).

1 Sloboda, J. A. (1974) The eye-hand span: an approach to the study of sight-reading. *Psychology of Music*, **2**, 4–10.

2 Sloboda, J. A. (1976) The effect of item position on the likelihood of identification by inference in prose and music reading. *Canadian Journal of Psychology*, **30**, 228–238.

3 Sloboda, J. A. (1976) Visual perception of musical notation: registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology*, **28**, 1–16.

4 Sloboda, J. A. (1976) Decision times for word and letter search: a holistic word identification model examined. *Journal of Verbal Learning and Verbal Behaviour*, **15**, 93–101.

5 Sloboda, J. A. (1977) Phrase units as determinants of visual processing in music reading. *British Journal of Psychology*, **68**, 117–124.

6 Sloboda, J. A. (1978) Perception of contour in music reading. *Perception*, **7**, 323–331.

7 Sloboda, J. A. (1978) The locus of the word-priority effect in a target detection task. *Memory and Cognition*, **5**, 371–376.

8*Sloboda, J. A. (1978) The psychology of music reading. *Psychology of Music*, **6**, 3–20.

9*Sloboda, J. A. The uses of space in music notation. (1980) *Visible Language*, **15**, 86–110.

10 Sloboda, J. A. (1980) Visual imagery and individual differences in spelling. In: U. Frith (Ed.) *Cognitive Processes in Spelling*. London: Academic Press.

11 Sloboda, J. A. and Gregory, A. H. (1980) The psychological reality of musical segments. *Canadian Journal of Psychology*, **34**, 274–280.

12 Sloboda, J. A. and Edworthy, J. (1981) Attending to two melodies at once: the effect of key relatedness. *Psychology of Music*, **9**, 39–43.

13 Sloboda, J. A. (1982) Music performance. In: D. Deutsch (Ed.) *Psychology of Music*. New York: Academic Press.

14 Rogers D. R. and Sloboda, J. A. (Eds.) (1983) *Acquisition of Symbolic Skills* (Proceedings of a NATO Human Factors Programme Conference, Keele, July 1982). New York: Plenum Press.

15 Sloboda, J. A. (1983) Neil O'Connor and the MRC Developmental Psychology Unit. *British Journal of Developmental Psychology*, **1**, 405–408.

16 Sloboda, J. A. (1983) The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology*, **37A**, 377–396.

17 Sloboda, J. A. (1984) Critical review of J.R. Anderson, *Cognitive Psychology* (1981) *Quarterly Journal of Experimental Psychology*, **33**, 317–322.

18 Sloboda, J. A. (1984) Critical review of Wing and Seashore tests of musical ability. In: H. Goldstein and P. Levy (Eds.) *Tests in Education: a book of critical reviews*. London: Academic Press.

19*Sloboda, J. A. (1984) Experimental studies of music reading: a review. *Music Perception*, **2.2**, 222–236.

20 Sloboda, J. A., Hermelin, B., and O'Connor, N. (1985) An exceptional musical memory. *Music Perception*, **3.2**, 155–170.

21 Sloboda, J. A. (1985) Expressive skill in two pianists: style and effectiveness in music performance. *Canadian Journal of Psychology*, **39**, 273–293.

22 Sloboda, J. A. (1985) Infant Perception. In: A. Branthwaite and D. Rogers (Eds.) *Children Growing Up*. Milton Keynes: Open University Press, pp. 61–81.

23 Sloboda, J. A. (1985) Perception and knowledge. In A. Branthwaite and D. Rogers (Eds) *Children Growing Up*. Milton Keynes: Open University Press.

24 Sloboda, J. A. (1985) *The Musical Mind: the cognitive psychology of music*. Oxford: Oxford University Press (reprinted with a new preface in 1999).

25 Sloboda, J. A. and Parker, D. H. H. (1985) Immediate recall of melodies. In: P. Howell, I. Cross, and R. West (Eds.) *Musical Structure and Cognition*. London: Academic Press.

26 Sloboda, J. A. (1986) Achieving our aims in music education research. *Psychology of Music*, **14.2**, 144–145.

27 Sloboda, J. A. (1986) Acquiring skill. In: A. Gellatly (Ed) *The Skillful Mind*. Milton Keynes: Open University Press.

28*Sloboda, J. A. (1986) Cognitive psychology and real music: the psychology of music comes of age. *Psychologica Belgica,* **26.2**, 199–219.

29 Sloboda, J. A. (1986) Computers and cognition. In: A. Gellatly (Ed) *The Skillful Mind.* Milton Keynes: Open University Press.

30 Sloboda, J. A. (1986) Reading: a case study of cognitive skill. In: A. Gellatly (Ed) *The Skillful Mind.* Milton Keynes: Open University Press.

31 Sloboda, J. A. (1986) What is skill? In A. Gellatly (Ed) *The Skillful Mind.* Milton Keynes: Open University Press.

32 Sloboda J. A. and Rogers D. R. (Eds). (1987) *Cognitive Processes in Mathematics.* (Selected and edited papers from the first International Keele Cognition Seminar, March 1985). Oxford: Oxford University Press.

33 Sloboda, J. A. (Ed.) (1988) *Generative Processes in Music: the psychology of composition, performance, and improvisation.* London: Oxford University Press.

34 Sloboda, J. A. (1988) *L'esprit musicien: la psychologie cognitive de la musique.* Brussels: Pierre Mardaga. (French Language Translation of The Musical Mind).

35 Sloboda, J. A. (1988) *La mente musicale: psicologia cognitivista della musica.* Bologna: Il Mulino, 1988. (Italian translation of *The Musical Mind*).

36 Sloboda, J. A. (1988) Musique et memoire: le point de vue du psychologie. (Music and memory: the psychologist's view). *Inharmoniques,* **4**, 106–119.

37 Gellatly, A., Rogers, D. R., and Sloboda, J. A. (Eds) (1989) *Cognition and Social Worlds.* (selected and edited papers from the second International Keele Cognition Seminar, March 1987). London: Oxford University Press.

38*Sloboda, J. A. (1989) Music as a language. In: F. Wilson and F. Roehmann (Eds.) *Music and Child Development: Proceedings of the 1987 Biology of Music Making Conference.* St. Louis, Missouri: MMB.

39 Sloboda, J. A. (1989) Music as a skill. In: S. Nielzen and O. Olsson (Eds.). *Structure and Perception of Electroacoustic Sound and Music.* Amsterdam: Elsevier.

40 Sloboda, J. A. (1989) Music, composers and scientists: a relationship of equals. In: S. Nielzen and O. Olsson(Eds.) *Structure and Perception of Electroacoustic Sound and Music.* Amsterdam: Elsevier.

41*Sloboda, J. A. (1989) Music psychology and the composer. In: S. Nielzen and O. Olsson (Eds.) *Structure and Perception of Electroacoustic Sound and Music.* Amsterdam: Elsevier.

42 Sloboda, J. A. (1989) What do psychologists have to say about nuclear war? In. J. Hartley and A. Branthwaite (Eds.) *The Applied Psychologist.* Milton Keynes: Open University Press.

43 Sloboda, J. A. (1990) Combating examination stress among university students: an institutional survey. *British Journal of Guidance and Counselling,* **18**, 124–136.

44 Sloboda, J. A. (1990) Musical excellence: how does it develop? In M. J. A. Howe (Ed.) *Encouraging the Development of Exceptional Abilities and Talents.* London: Wiley.

45 Sloboda, J. A. (1990) Ongakuteki na Yushusa wa Dono Yoni Hattatsu Suruka? (How does musical excellence develop?) Translated by Fumiko Fujita. *Journal of Research in Music Education (Japan)*, **64**, 163–168.

46 Sloboda, J. A. (1990) Que puede ensenar la psicologia de la musica an los musicos? *Musica Y Educacion (Spain)*, **3.2**, 335–355. (translated by Ubaldo Perez).

47 Sloboda, J. A. and Howe, M. J. A. (1990) Reasons for choosing specialist music schooling: an interview study. *Journal of the Music Masters and Mistresses Association*, No 25, November, 17–20.

48 Howe, M. J. A. and Sloboda, J. A. (1991) Early signs of talents and special interests in the lives of young musicians. *European Journal of High Ability*, **2**, 102–111.

49 Howe, M. J. A. and Sloboda, J. A. (1991) Helping children to become highly competent adults: what can parents do? In: J. Radford (Ed), *Talent, Teaching and Achievement*. London: Jessica Kingsley.

50 Howe, M. J. A. and Sloboda, J. A. (1991) Young musicians' accounts of significant influences in their early lives: 1. The family and the musical background. *British Journal of Music Education*, **8**, 39–52.

51 Howe, M. J. A. and Sloboda, J. A. (1991) Young musicians' accounts of significant influences in their early lives: 2. Teachers, practising and performance. *British Journal of Music Education*, **8**, 53–63.

52* Sloboda, J. A. (1991) Empirical studies of emotional response to music. In: M. Riess-Jones (Ed) *Cognitive Bases of Musical Communication*. Washington, DC: American Psychological Association, pp. 33–46.

53 Sloboda, J. A. (1991) Music structure and emotional response: some empirical findings. *Psychology of Music*, **19**, 110–120.

54* Sloboda, J. A. Musical Expertise. (1991) In K. A. Ericsson and J. Smith (Eds) *The Study of Expertise: prospects and limits*. Cambridge: Cambridge University Press.

55 Sloboda, J. A. and Howe, M. J. A. (1991) Biographical precursors of musical excellence: an interview study. *Psychology of Music*, **19**, 3–21.

56 Howe, M. J. A. and Sloboda, J. A. (1992) Problems experienced by talented young musicians as a result of the failure of other children to value musical accomplishments. *Gifted Education*, **8.1**, 16–18.

57 Sloboda, J. A. (1992) Etudes empiriques de la reponse emotionnelle a la musique. *Analyse Musicale*, **26**, 14–21.

58* Sloboda, J. A. (1992) Psychological structures in music: core research 1980–1990. In: Paynter, R. Orton, T. Seymour, and T. Howell (Eds.) *A Compendium of Contemporary Musical Thought*. London: Routledge, pp. 803–839.

59 Sloboda, J. A. and Hopkins, J. (1992) Promoting mental health in the workplace: the Staff Care concept. In: D. R. Trent (Ed) *Promotion of Mental Health*, Vol. I. Aldershot: Avebury, pp. 297–304.

60 Sloboda, J. A. and Howe, M. J. A. (1992) Transitions in the early musical careers of able young musicians: choosing instruments and teachers. *Journal of Research in Music Education*, **40**, 283–294.

61 Sloboda, J. A. (1993) Begabung und Hochbegabung. In: H. Bruhn, R. Oerter and H. Rosing (eds) *Musikpsychologie: ein Handbuch*. Reinbek: Rowholt, pp. 565–578.

62*Sloboda, J. A. (1993) Musical ability. In: K. Ackrill and G. Bock (Eds). *The Origins and Development of High Ability* (Ciba Symposium 178). Chichester: Wiley, pp. 106–118.

63 Sloboda, J. A. (1993) Weighing of the talents. *Nature*, **362**, 115–116.

64 Sloboda, J. A., Hopkins, J. S., Turner, A., Rogers, D.R., and McLeod, J. (1993) An evaluated staff counselling programme in a public sector organisation. *Employee Counselling Today*, **5.5**, 4–12.

65 Aiello, R. with Sloboda, J. A. (1994) *Musical Perceptions*. New York: Oxford University Press.

66 Brodsky, W., Sloboda, J. A., and Waterman, M. (1994) An exploratory investigation into auditory style as a correlate and predictor of music performance anxiety. *Medical Problems of Performing Artists*, **9.4**, 101–112.

67*Sloboda, J. A. (1994) Book review of *Language, Music, and Mind* by Diana Raffman. *Mind and Language*, **9.3**, 377–385.

68 Sloboda, J. A. (1994) Book review of *Music and the Mind*, by Anthony Storr. *Music and Letters*, **75**, 122–124.

69 Sloboda, J. A., Davidson, J., and Howe, M. J. A. (1994) Is everyone musical? *The Psychologist*, **7.7**, 349–354.

70 Sloboda, J. A. (1994) Music performance: expression and the development of excellence. In: R. Aiello with J. Sloboda (eds) *Musical Perceptions*. New York: Oxford University Press, pp. 152–172.

71 Sloboda, J. A., Davidson, J., and Howe, M. J. A. (1994) Musicians: experts not geniuses. *The Psychologist*, **7.7**, 363–365.

72 Sloboda, J. A. (1994) What makes a musician? *European String Teachers Association Review*, **2**, 31–34.

73 Deliège, I. and Sloboda, J. A. (Eds) (1995) *Naissance et developpement du sens musical*. Paris: Presses Universitaires de France.

74 Deliège, I., and Sloboda, J. A. (Eds.) (1995) *The Origins and Development of Musical Competence*. London: Oxford University Press/Paris: Presses Universitaires de France.

75 Howe, M. J. A., Davidson, J. W., Moore, D. G., and Sloboda, J. A. (1995) Are there early childhood signs of musical ability? *Psychology of Music*, **23**, 162–176.

76 Rogers, D. R., McLeod, J, and Sloboda, J. A. (1995) Counsellor and client perceptions of the effectiveness of time-limited counselling in an occupational scheme. *Counselling Psychology Quarterly*, **8.3**, 221–231.

77 Sloboda, J. A. (1995) Are we all musicians? Evidence from the biographies of young instrumentalists. *The Physiological Society Magazine,* **19**, 32–33.

78 Sloboda, J. A. (1995) Psychology of music today: the need for applicable psychology. In: M. Manturszewska, K. Miklaszewski, and A. Bialkowski (eds.) *Psychology of Music Today*. Warsaw: Frederick Chopin Academy of Music, pp. 19–24.

79 Sloboda, J. A. (1995) Talent, motivatie, oefening en succes (Talent, achievement, motivation and success). In: F. Evers, M. Jansma, P. Mak and B. Devries (Eds.) *Musiekpsychology [Psychology of Music]*. Assen: Van Gorcum, pp. 27–40.

80 Sloboda, J. A. and Davidson, J. W. (1995) L'interprete en herbe. In: I. Deliège and J. A. Sloboda (eds.), *Naissance et developpement du sens musical.* Paris: Presses Universitaires de France, pp. 199–222.

81 Sloboda, J. A. and Davidson, J. W. (1995) The young performing musician. In: I. Deliège and J. A. Sloboda (Eds.) *The Origins and Development of Musical Competence.* London: Oxford University Press/Paris: Presses Universitaires de France.

82 Sloboda, J. A., and Howe, M. J. A. (1995) Biographical precursors of musical excellence: an interview study. In: M. Manturszewska, K. Miklaszewski, and A. Bialkowski (eds.) *Psychology of Music Today*. Warsaw: Frederick Chopin Academy of Music. pp. 338–357.

83 Davidson, J. W., Howe, M. J. A., Moore, D. G., and Sloboda, J. A. (1996) The role of parental influences in the development of musical ability. *British Journal of Developmental Psychology,* **14**, 399–412.

84 Howe, M. J. A. and Sloboda, J. A. (1996) Early signs of talents and special interests in the lives of young musicians. In: A. J. Cropley and D. Dehn (Eds.) *Fostering the Growth of High Ability: European perspectives.* Norwood, NJ: Ablex Publishing Corporation, pp. 233–246.

85*Sloboda, J. A. (1996) Emotional responses to music: a review. Invited keynote paper. In: K. Riederer and T. Lahti (Eds.). *Proceedings of the Nordic Acoustical Meeting (NAM96)*. Helsinki: The Acoustical Society of Finland, pp. 385–392.

86*Sloboda, J. A. (1996) The acquisition of musical performance expertise: deconstructing the 'talent' account of individual differences in musical expressivity. In: K. A. Ericsson (Ed.) *The Road to Excellence: the acquisition of expert performance in the arts and sciences, sport and games.* Mahwal, NJ: Lawrence Erlbaum Associates, pp. 107–126.

87 Sloboda, J. A., Davidson, J. W., Howe, M. J. A., and Moore, D. G. (1996) The role of practice in the development of expert musical performance. *British Journal of Psychology.* **87**, 287–309.

88 Brodsky, W. and Sloboda, J. A. (1997) Clinical trial of a music generated vibrotactile therapeutic environment for musicians: main effects and

outcome differences between therapy subgroups. *Journal of Music Therapy*, **24**, 2–33.

89 Clarke, E. F., Parncutt, R., Raekallio, M., and Sloboda, J. A. (1997) Talking fingers: an interview study of pianists' views on fingering. *Musicae Scientiae*, **1.1**, 87–108.

90 Davidson, J. W., Howe, M. J. A., and Sloboda, J. A. (1997) Environmental factors in the development of musical performance over the lifespan. In: Hargreaves, D. J. and North, A. C. (Eds.) *The Social Psychology of Music*. Oxford: Oxford University Press, pp. 188–206.

91 Deliège, I. and Sloboda, J. A. (Eds.) (1997) *Perception and Cognition of Music*. Hove: Psychology Press.

92 O'Neill, S. A. and Sloboda, J. A. (1997) The effects of failure on children's ability to perform a musical test. *Psychology of Music*, **35.1**, 18–34.

93 Parncutt, R.P., Desain, P., Clarke, E. F., Raekallio, M., and Sloboda, J. A. (1997) An ergonomic model of keyboard fingering for melodic fragments. *Music Perception*, **14.4**, 341–382.

94 Sloboda, J. A. and Newstead, S. E. (1997) Guidelines for PhD examiners: an evaluation of impact. *The Psychologist*, **10.9**, 407–410.

95 Davidson, J. W., Sloboda, J. A., Moore, D. G., and Howe, M. J. A. (1998) Characteristics of music teachers and the progress of young instrumentalists. *Journal of Research in Music Education*, **46.1**, 141–160.

96 Howe, M. J. A., Davidson, J. W., and Sloboda, J. A. (1998) Innate talent: reality or myth? *Behavioural and Brain Sciences*, **21.3**, 399–442.

97 Sloboda, J. A. (1998) Brain waves to the heart. *BBC Music Magazine*, November, pp. 31–33.

98* Sloboda, J. A. (1998) Does music mean anything? *Musicae Scientiae*, **2.1**, 21–32.

99 Sloboda, J. A. (1998) Peer counselling as a resource in the teaching of psychology. In: Radford, J., Van Laar, D., and Rose, D. (Eds.) *Innovations in Teaching Psychology*. Edgbaston: SEDA, pp. 110–120.

100 Sloboda, J. A., Clarke, E. F., Parncutt, R., and Raekallio, M. (1998) Determinants of finger choice in piano sight-reading. *Journal of Experimental Psychology: Human Perception and Performance*, **24.1**, 185–203.

101 Parncutt, R., Sloboda, J. A., and Clarke, E. F. (1999) Interdependence of right and left hands in sight-read, written, and rehearsed fingerings of parallel melodic piano music. *Australian Journal of Psychology*, Special Issue on 'Music as a Brain and Behavioural System', **51**(3), 204–210.

102 Sloboda, J. A. (1999) *Cognition, Emotion, and Performance: three lectures on the psychology of music*. Warsaw: Chopin Academy of Music.

103* Sloboda, J. A. (1999) Everyday uses of music: a preliminary study. In: S. W. Yi (Ed.) *Music, Mind, and Science*. Seoul: Seoul National University Press, pp. 354–369.

104 Sloboda, J.A. (1999) (contributions to) Music and the mind. In: M. Oliver (Ed.) *Settling the Score: a journey through the music of the 20th century*. London: Faber and Faber, pp. 211–235.

105*Sloboda, J. A. (1999) Music performance and emotion: issues and developments. In: S. W. Yi (Ed.) *Music, Mind, and Science*. Seoul: Seoul National University Press, pp. 220–238.

106*Sloboda, J.A. (1999) Music: where cognition and emotion meet. *The Psychologist*, **12.4**, 450–455.

107 Sloboda, J. A. (1999) Preface to the 1999 reprinting of *The Musical Mind: the cognitive psychology of music*. Oxford: Oxford University Press, pp. ix–xx.

108 Sloboda, J. A. and Howe, M. J. A. (1999) Musical talent and individual differences: a reply to Gagne. *Psychology of Music*, **27.1**, 52–54.

109 Sloboda, J. A. (2000) Individual differences in music performance. *Trends in Cognitive Science*, **4.10**, 397–403.

110*Sloboda, J. A. (2000) Music and worship: a psychologist's perspective. In: T. Hone, M. Savage, and J. Astley (Eds.) *Creative Chords: studies in music, theology and Christian formation*. Leominster: Gracewing, pp. 110–125.

111 Sloboda, J. A. (2000) Why are musical enjoyment and musical achievement at such different levels in contemporary society? In W. Jankowski (Ed.) *Czlowiek – Muzyka – Psychologia: ksiazka dedykowana Professor Marii Manturzewskiej [Education – Music – Psychology: in honour of Professor Maria Manturszewska]*. Warsaw: Frederik Chopin Academy of Music.

112 Juslin, P. N. and Sloboda, J. A. (Eds.) (2001) *Music and Emotion: theory and research*. Oxford: Oxford University Press, pp. 487.

113*Sloboda, J. A. (2001) Emotion, functionality, and the everyday experience of music: where does music education fit? *Music Education Research*, **3**(2), 243–253.

114 Sloboda, J. A. (2001) Entries on 'Affect', 'Memory'. In: S. Sadie (Ed.) *The New Grove Dictionary of Music and Musicians*. London: Macmillan.

115 Sloboda, J. A. and Juslin, P. N. (2001) Psychological perspectives on music and emotion. In: Juslin, P. N, and Sloboda, J. A. (Eds.) *Music and Emotion: theory and research*. Oxford: Oxford University Press, pp. 71–104.

116 Sloboda, J. A. and Lehmann, A.C. (2001) Performance correlates of perceived emotionality in different interpretations of a Chopin piano prelude. *Music Perception*, **19.1**, 87–120.

117 Sloboda, J. A. and O'Neill, S. A. (2001) Emotions in everyday listening to music. In: Juslin, P. N. and Sloboda, J. A. (Eds.) *Music and Emotion: theory and research*. Oxford: Oxford University Press, pp. 415–430.

118 Sloboda, J. A., O'Neill, S. A., and Ivaldi, A. (2001) Functions of music in everyday life: an exploratory study using the Experience Sampling Methodology. *Musicae Scientiae*, **5**(1), 9–32.

119*Sloboda, J. A. (2002) Dotu musicali e innatismo. In: J. J. Nattiez (Ed.) *Enciclopedia della Musica: Vol II Il sapere musicale.* Torina: Giulio Einaudi Editore, pp. 509–529.

120*Sloboda, J. A. (2002) The 'sound of music' versus the 'essence of music': dilemmas for music-emotion researchers (commentary). Special issue on Music and Emotion. *Musicae Scientiae,* **6.3**, 235–253.

# INDEX


11 September 2001 attacks 380, 411
ability see musical ability
academic norms 413
acquisition
  of musical ability 266-7, 303-4
  of skill 246-8
aesthetic attraction to music 382-3
*affordances* 195, 347
age, individual differences 324-5
algorithms for key identification 134-5
ambiguity 8
applicability of research 413-14
  *see also* benefits of music
appogiaturas 210, 221, 260, 356-7
appropriation theories 320
architecture, music as 194-5
Armstrong, Louis 251-5
art, attitude towards 341
associationistic function *see* memory
asynchrony 98
'audiation' 302
autistic subject (NP) 108-11, 166, 249-51, 254
autobiographical memory study 181-8, 257-8

background, music as unattended 348
bar lines 57-60
beaming notes 21-2
behaviourism 378
beliefs, teacher's 300-1
benefits of music
  aims of research 396-7
  and the 'business model' 399
  commissioned policy-relevant research 402-4
  demonstrating benefits vs. producing benefits 397-8
  documenting vs. advocacy 398-402
  IQ measures 397-8
  justifying research 406
  'Mozart effect' 379-80, 396-7
  musical interventions 397
boys, gender stereotyping of 271, 288, 310
business model 398-9
busking 328

categorization of sounds 177-8
change
  mood-altering effect of music 335, 384
  music as agent of 204
  *see also* cognitive re-appraisal
children
  autobiographical memory study 181-8, 257-8
  development of emotional response to music 205-6
  development of skill by 247-8, 256-9
  development to able adult 311-12
  dropping instrumental tuition 362-5, 402
  expressive performance by 291-2
  importance of non-threatening environment for 258-9, 308-11
  individual differences in ability 277-9, 297-301, 303
  infant musicality 277, 291, 298-9, 387-8
  relationships with significant adults 269-71, 363-4, 388
  sibling differences 299
  *see also* music education
chord relationships 125-7
church *see* worship
clef signs 51
closure, tonal 144
cognitive re-appraisal 199-201
commercialization of music 379-80, 391-2
commissioned policy-relevant research 402-4
commodification of music 379-80, 391-2
commonalities of response 384, 386-7
competitions, music 186, 341

composers
  as architects 194-5
  credibility of music psychology to 191-2
  dialogue with psychologists 192-4
  intentions of 231-2
computer music 168
conflict, over choice of music 366
connectionist models 127, 248, 266, 267, 304
consonance, overtone 122-4
contemplative response to music 357
continuous response methodology 386
contour
  melodic contour analysis 78-81
  visual 18
control
  and mood change 335, 384
  and music listening 328-9
cross-cultural research
  individual differences 217
  non-Western music 304
  pitch relationships 124
cue-impoverishment of music 226-7
cultural democratization 378-80
cultural position of music 390-1
cultural trends 369-72

'Darling, They're Playing Our Tune' theory 348
democratization, cultural 378-80
demonstration, teaching as 160
Deutsch, Diana 117
discrete sampling 387

ecological validity 101-2, 129, 131, 377
education *see* music education
electronic communication 370
elitism 341-2, 377-9, 385
emotion
  aesthetic attraction to music 382-3
  between-person variation 217
  children's development 205-6
  cognitive content of 204-8
  commonalities of response 384, 386-7
  cue-impoverishment of music 226-7
  definitional problems 216-17
  dynamic intensity of 225
  episodic associations 336-7
  expectancy violations 211-12, 229-31, 232-3, 337
  experimental control problems 231
  expertise and 254-6
  external factors 208-9, 228
  feelings vs. judgment 227-8
  forced verbal response 206-8
  free verbal response 204-6, 386
  function of music and 199-201
  heart reactions 212, 221, 260
  iconic associations 337
  intensification of 204-5
  intensity over content 226-7
  intrinsic musical factors 209-13, 228
  link to musical events 220-2
  and meaning (Raffman's view) 158
  measuring 218-19
  'peak experiences' 258, 309, 384-5
  performance-response relationship 236-9
  performer's 216, 225-6
  performer's intentions and 238-9
  performer's interpretation and 233-4
  physical manifestations of 168-70
  release of 204
  self-ascription of positive 228
  'shivers' response 210-12, 221, 260
  and structures 259-60
  'tears' response 210, 221, 260, 356-7
  violation of expectations 211-12, 229-31, 232-3, 337
  virtuosity as source of 234-6
  within-person variation 218
  *see also* autobiographical memory study; mood
enjoyment, importance of in learning 258-9, 308-11, 341-2, 363
ensemble playing, reading ability and 20
episodic associations 336-7
examinations, music 280-1, 341, 373
expectation
  misreading and 14-15, 37-8
  of teacher 300-1
  violation of 211-12, 229-31, 232-3, 337
expertise
  acquisition in noninstructional settings 248-55
  acquisition of skill 246-8, 258-9
  children's development of 247-8, 256-9
  definition of 243-6
  and emotion 254-6
  jazz musicians 251-5
  practice patterns and 269, 270, 279-83, 306-7, 339, 365
  prodigies 249-51
  savants 108, 249-51
  tacit 248
  *see also* musical ability
expressive performance
  by children 291-2
  devices used 284-5, 388-90
  'Feeling Sound' training 403-4
  generative 290-2
  imitative 286-90

individual differences in 217
individuality of 233-4
intuitive 267-8, 305
slowing 142-3
technical skills and 268-71, 283-6
understanding and 305-6
unreliability of achieving 403
when sight-reading 38-40
*see also* emotion
extreme performances 234
extrinsic motivation 270-1, 307-11
'eye-hand' span (EHS) 35-7

failure, resistance to 300, 339
family environment 270, 340-1, 363
feeling ineffability 157-9, 352-3
feelings
dynamism of 166-8, 225
vs. judgment 227-8
*see also* emotion
feminism 370
'fixation' by poor readers 6-11
'flashbulb memory 348-9
folk tunes *see* melody recall research
Frith, Simon 379
functions of music 197-9

gender
individual differences 324-5, 328
stereotypes 271, 288, 310
generative expression 290-2
generative theory of tonal music (GTTM) 139-43
*Generative Theory of Tonal Music* (Lerdahl and Jackendoff) 102-4
genetics 276-9, 297-8
giftedness *see* talent
Goldovsky experiment 13-14, 37-8
'good ear', a 176
grammar *see* language; structures
grouping hierarchies 136-9
*see also* generative theory of tonal music
GTTM (generative theory of tonal music) 139-43

heart reactions 212, 221, 260
'helpless orientation' 300
heritability of musical ability 276-9, 297-8
hierarchical representation 111, 135-46, 164-5
hierarchical tree theory 89-90
'hobby', instrument playing as 365

iconic associations 337
*idiots savants* 108, 249-51
imitative expression 286-90
improvisation 291-2
individual differences
age 324-5
in children's ability 277-9, 297-301, 303
in contemplative response 357
cross-cultural 217
in expressive performance 288-90
gender 324-5, 328
siblings 299
individuality, expression of 233-4
ineffability of music 153-4, 156-61, 351-5
infant musicality 277, 291, 298-9, 387-8
inheritance of musical ability 276-9, 297-8
intentions
composer's 231-2
performer's 238-9
'reading' into music 388, 390
interventions, musical 397
intrinsic
motivation 270, 307
musical factors 209-13, 228
intuitive performance 267-8, 305
IQ measures 397-8

jazz musicians 251-5
judgment, versus feelings 227-8

key, establishing 129-35
key-distance effect 126-7
Klavarscribo system 46-7, 67
Kosovo crisis 410
Krumhansl, Carol (work of) 121-9

language, music as a
autobiographical memory study 181-8, 257-8
cultural 'grammar' of music 302
phonology 177-8

language, music as a *(Continued)*
semantics 179-81
syntax 178-9
*Language, Music and Mind* (Raffman) (book review) 153-62, 351
learning
of pitch relationships 124-5
*see also* music education
ledger lines 51-2
Leverhulme Project 277-9, 280, 282
listening
age differences 324-5, 326
control and 328-9
gender differences 324-5, 326, 328
listener as active agent 319-20
Mass-Observation project 322-9
personal uses of music 324-6
'pharmaceutical' approach to 319, 346, 384
pilot studies into 321-2
in public places 326-9
uses of music 324-9
literacy, musical 10-11
Longuet-Higgins, seminal work of 99

Mass-Observation project 322-9
'mastery orientation' 300
meaning of music
emotion and (Raffman's view) 158
emotive properties mistaken for content 158-9
and music education 368-9
*see also* language
'melodic process' 144
melody recall research
breath analysis 83-4, 106-7
data analysis 74, 105-6
harmonic analysis 85-7, 106
hierarchical tree theory 89-90
intersong contamination 76
melodic contour analysis 78-81
method 76-8, 104-5
metrical analysis 81-2, 106
musicians vs. nonmusicians 87-8
phrase structure analysis 84-5, 106-7
recall vs. recognition 74-6
rhythmic analysis 83
transcription problems 71-4
untrained subjects in 75
memory
autobiographical memory study 181-8, 257-8
for carols 186
exceptional recall ability (NP) 107-11, 166, 249-51, 254
expressive performance and 287-90
'flashbulb' 348-9
for hymns 186
mnemonic functions of music 197-8, 217, 348-9
nature of musical 18-19
negative musical 184-5, 186, 271, 308-11
for nursery rhymes 186
reading span 16-18, 29
*see also* melody recall research
men's listening 324-5
'mentor' 311-12
methodology
agreed methods in psychology of music 99-100
continuous response 386
discrete sampling 387
forced verbal response 206-8
free verbal response 204-6, 386
metre
recall of 81-2
sight-reading of 38-40
metrical structure (in GTTM) 140
misreading 14-15, 37-8
modular processing 218
monitoring own performance 15-16
mood
-altering effects of music 334-8
*see also* emotion
motivation
autistic player's (NP) 253-4
extrinsic 270-1, 307-11
and 'helpless orientation' 300
intrinsic 270, 307
Louis Armstrong's 253-4
and 'mastery orientation' 300
resistance to failure 300, 339
'Mozart effect' 379-80, 396-7
multiculturism 369
musical ability
acquisition of 266-7, 303-4
definition 265-6
individual differences in children's 277-9, 297-301, 303
inheritance of 276-9, 297-8
nature of 301-3
teacher's belief about student's 300-1
'technical' 304-6
*see also* expertise
musical subcultures 367-8
music education
accreditation of achievement 280-1, 341, 373
cultural trends and 369-72
cutbacks in 397
dropping instrumental tuition 362-5, 402
funding for 372
and meaning of music 368-9
parental support and 270, 340-1, 363

practitioner and policy implications 417-19
providers of 372
purpose of 368-9
sociocultural environment and 365-7
'teacher education' 373
transition from primary to secondary 362, 365
variety as key concept in 372-3
*see also* music teachers
music notation
bar lines 57-60
clef signs 51
clustering in 64, 66-7
improving 21-3
information conveyed by 5-6
Klavarscribo system 46-7, 67
layout compared with text 43-4
ledger lines 51-2
legibility of 65-6
level of detail in 46-8
need for 67
orthochronic *see* orthochronic notational system
psychological effectiveness of 65-8
size of 52-3
specificity of information 61-2
symbolic structure of 63-5
tablatures 45-6
time notation 54-61
time signatures 57
transcriptions and 72-4
type of information 62
vertical alignment of notes 57-60
*see also* orthochronic notational system
*Music Perception* journal 117-19
music reading
cognitive load of 7
compared to prose reading 3-4
conscious awareness during 11
copying paradigm 31-4
early studies of 29-31
and ensemble playing 20
experimental research into 5
eye movement studies 29-30
'fixation' by poor readers 6-11
goals of 27-8
improving learning of 19-20
'interference problem' 32-3
lack of insight 11-12
and musical memory 5
musical text and 21-3
musicians vs. nonmusicians 31-4
neglected area of study 3-5
silent reading 28-9
span 16-18, 29
visual task hypothesis 28-9
*see also* sight-reading
music teachers
beliefs about student's ability 300-1
expectations from psychology of music 246, 247
knowledge of musical subcultures 367-8
not blamed for 'failure' of music education 372
training of 373
*see also* music education; teaching

National Curriculum for music 341, 371
negative musical memories 184-5, 186, 271, 308-11
neumes 48-50
niche cultures 371
nonmusicians
implicit knowledge of 247
melody recall by 87-8
and music reading 31-4
notation *see* music notation; orthochronic notational system
notational systems, new 22-3
nuance ineffability 159-61, 353-4
nursery rhymes 186

orchestral parts, set quality 22
orthochronic notational system
2-D function of 63
abstract nature of 44-5
clustering in 64, 66-7
multiple staves 53-4
spacing of 66
specificity of 62
staff in 48-54
time notation 54-61
universality of 67
overtone consonance 122-4

page-turning 11
paradigmatic status *see* 'psychology of music'
'peak experiences' 258, 309, 384-5
performance
technical ability 304-6
*see also* expressive performance
performers
children 186
emotion of 216, 225-6
expression of individuality of 233-4
extreme performances by 234

performers *(Continued)*
faithful to composer's intentions 231-2
virtuosity 234-6
personal associations 197-8, 217, 348-9
personal relationships, music as vehicle for 358
personal uses of music 324-6, 366
*see also* benefits of music
'pharmaceutical' approach to listening 319, 346, 384
phonology 177-8
physiological responses 100, 384
pitch
predominance of 121
rare intervals 130
relationships 122-5
plainchant melodies 49
policy-relevant research 402-4
political issues 407-15
polyphonic music 53
pop music 357, 367
postmodernism 371-2
practice 269, 270, 279-83, 306-7, 339, 365
prodigies 249-51
prolongational reduction (in GTTM) 140
'proof-readers' error' 12-13, 37-8
psychology of music
agreed methods 99-100
agreed problem 97-9
agreed theoretical frameworks 100
appropriate research 101-2
author's current position and 412
comes of age 102-4
composers and 191-2
and 'democratization of taste' 378
empirical measurement and 119-21
engagement with policy-makers 402-3
and musical experience 97-9
personal and political in 407-15
questions asked by 195-7
'reduction' of music in 376
reification of the variable 376-7
researcher's personal values 395-6, 414-15
and uniqueness of music 101
'value' of 395-6, 406-7
Western tonal music dominates 121
public places, music in 326-9

reading *see* music reading; sight-reading
re-appraisal *see* cognitive re-appraisal
recall *see* melody recall research; memory
reductional hierarchies 135-6, 137-9
Re-Evaluation Counselling 409
reification of the variable 376-7
relationships
children's with significant adults 269-71, 363-4, 388
chords 125-7
with 'mentor' 311-12
performance-responses 236-9
personal 358
pitch 122-5
religion *see* secularism; worship
representation
hierarchical 111, 135-46, 164-5
nature of musical 164-6, 350-1
Raffman's levels of 154-6
research
academic norms 413
applicability of 413-14
evaluation criteria 395-6
'Feeling Sound' training 403-4
focus on values 414-15
Krumhansl's 121-9
limitations of 127-9
personal values of researcher 395-6, 414-15
sensitivity to applicability 413
social engagement levels 412-15
*see also* benefits of music; melody recall research
resistance
to emotional effects of music 357
to failure 300, 339
rhythm, recall of 83
rhythmic notation 54-61
rise times 98

savants 108, 249-51
school
cutbacks in provision in 397
negative musical memories of 184-5, 186, 271, 308-11
*see also* music education; music teachers; teaching
scientific analysis of music
aversion to 175
benefit to musicians 175-6
and dominant cultural values 391
*see also* psychology of music
score *see* music notation
secularism 371
self-administered therapy, music as 216
semantics 179-81
'shivers' response 210-12, 221, 260
sibling differences 299
sight-reading
expressive variation in 38-40
'eye-hand' span (EHS) 35-7

good readers compared to poor readers 35-7
master readers compared to good readers 38-40
metrical reading 38-40
misreading 12-16, 37-8
notational convention and 67-8
*see also* music reading
sight-singing 20
skill *see* expertise
slowing, expressive 142-3
social factors
the academy 342
cultural trends 369-72
elitism 341-2, 377-9, 385
family environment 270, 340-1, 363
knowledge about cultural position of music 390-1
music education and 365-7
purpose of music in society 341
reduction in social opportunities for music 340-1, 405
social engagement in research 412-15
*see also* benefits of music
span
'eye-hand' (EHS) 35-7
reading 16-18, 29
staff-lines 48-54
stereotypes, gender 271, 288, 310
structural ineffability 156-7, 352
structures
conventional 232
dialectical 232-3
and emotion 259-60
hierarchical representation 111, 135-46, 164-5
internal musical characteristics and emotion 220-1, 356
naturally detected 333
*see also* representation
subcultures, musical 367-8
synchrony 98
syntax 178-9

tablatures 45-6
talent 275-6, 286, 297-301
teachers *see* music teachers
teaching
demonstration as 160
*see also* children; music education; music teachers; school
'tears' response 210, 221, 260, 356-7
technical skills
and expressive performance 268-71, 283-6
and performance ability 304-6
*see also* practice
'thrills' response 210-12, 221, 260, 356-7
time signatures 57
time-span structure (in GTTM) 140
tonal centre, establishing 129-35
tonal closure 144
tone-deafness 289
top-down processing 13-14

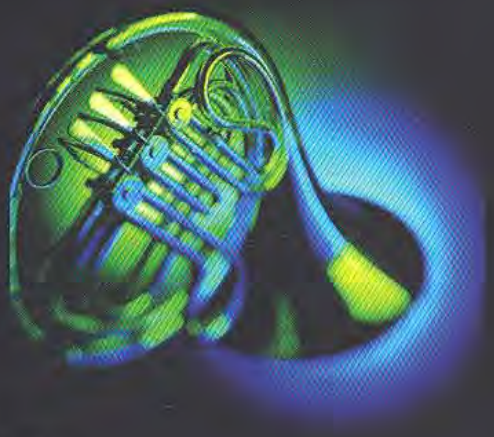
uniqueness of music 101

violations of expectancy 211-12, 229-31, 232-3, 337
virtuosity 234-6
visual contour 18

Western tonal music dominance 121
women's listening 324-5
worship
associations and 349
communality of 358
contemplative response to music 357
music's ineffability and 352, 353, 354-5
music as a source of personal associations 348-9
music as unattended background 348
psychology of music and 345-6

Young People's Participation in Music project 402
youth culture 370

Zimmerman, Krystian 234-5



In the 20 years since the publication of John Sloboda's landmark volume *The musical mind*, music psychology has developed as a vibrant area of research, exerting influence on areas as diverse as music education and cognitive neuroscience. This new book brings together 24 selected essays and reviews written by an internationally acclaimed authority on music and the mind. Chapters are grouped into four main areas of study: cognitive processes (including music reading, memory, and performance); emotion and motivation; talent and skill development; and music in the real world (including functions of music in everyday life and culture). The book ends with a newly written chapter on music psychology and social benefits. The book brings together a range of influential writings to provide a compendious overview of a subject that has come to maturity during the author's career, a career that has significantly contributed to the development of the field.

## BOOKS BY THE SAME AUTHOR

**Music and emotion: theory and research**
P. N. Juslin and J. A. Sloboda (OUP, 2001)

**Musical beginnings: origins and development of musical competence**
J. A. Sloboda and I. Deliege (OUP, 1995)

**Generative processes in music: the psychology of composition, performance, and improvisation**
J. A. Sloboda (OUP, 1988)

**The musical mind: the cognitive psychology of music**
J. A. Sloboda (OUP, 1985)

## REVIEWS FROM OTHER BOOKS BY THE SAME AUTHOR

*"the clarity of Sloboda's writing and his numerous suggestions for further research will make his book essential reading for anyone, student or researcher, interested in how minds and music interact."* Nature

*"this work is infused with informed observations of high interest to psychologists, musicians and teachers of music...Sloboda's research has been exemplary"*
Allan Shields in *Leonardo*, Journal of the International Society for the Arts, Sciences and Technology

*"Sloboda's book deserves loud applause, for its content, for its frankness in dealing with speculative issues largely ignored by other workers, and especially because it brings with it a new kind of approach."*
New Scientist

*"one of the most stimulating books it has been my lot to read"* Classical Music

*"as a vastly influential sourcebook ... even entirely non-academic music lovers should find plenty to interest them in Sloboda's lucid surveys of research into the ways performers read, master and memorise music, the way it is improvised and composed, how we hear and grasp it, what part it plays in personal development and, more broadly, in society and culture as a whole."* Bayan Northcott, BBC Music

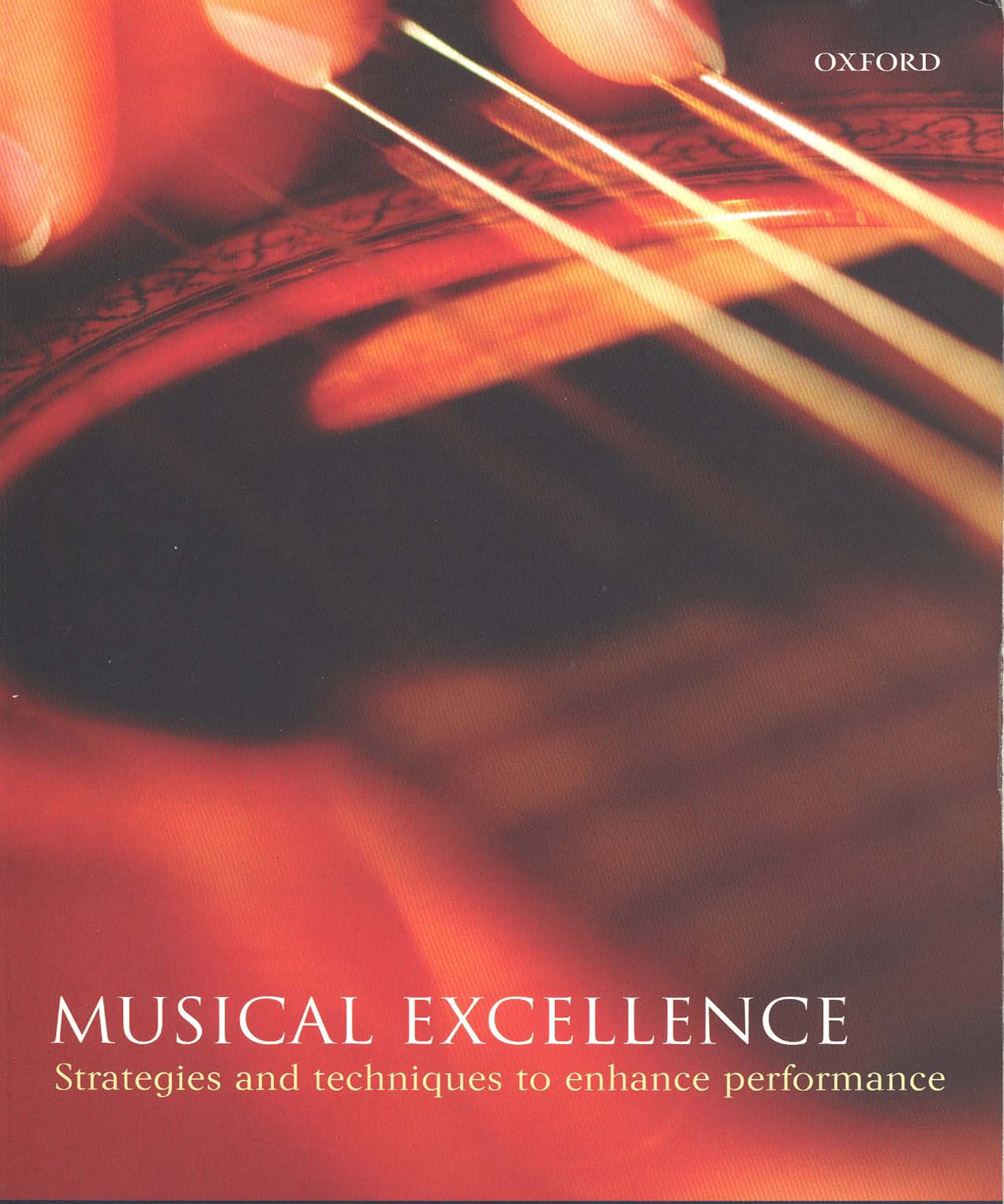
*Cover image © Getty Images*

OXFORD
UNIVERSITY PRESS

www.oup.com

EL ARGONAUTA
la librería de la música
0-19-853013-7
SLOBODA, JOHN
EXPLORING THE MUSICAL MIND.
P.V.P. (euros): 40.50

ISBN 0-19-853013-7
9 780198 530138



OXFORD

# MUSICAL EXCELLENCE

## Strategies and techniques to enhance performance

AARON WILLIAMON

# MUSICAL EXCELLENCE

## Strategies and techniques to enhance performance

Edited by

AARON WILLIAMON

*Royal College of Music, London*

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Taipei Toronto Shanghai

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan South Korea Poland Portugal
Singapore Switzerland Thailand Turkey Ukraine Vietnam

Published in the United States
by Oxford University Press Inc., New York

© Oxford University Press 2004

The moral rights of the author have been asserted

Database right Oxford University Press (maker)

First published 2004
Reprinted 2004 (with corrections)

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above.

You must not circulate this book in any other binding or cover
and you must impose this same condition on any acquirer.

A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
(Data available)

ISBN 0 19 8525346 (Hbk)
ISBN 0 19 8525354 (Pbk)

10 9 8 7 6 5 4 3 2

Typeset by Newgen Imaging Systems (P) Ltd., Chennai, India
Printed in Great Britain
on acid-free paper by
Biddles Ltd., King's Lynn, Norfolk

# ACKNOWLEDGMENTS

I should like to thank the contributors to this book, whose attention to detail, general erudition, and insight into musical performance have been both professionally rewarding and personally inspiring. I am also grateful to the Royal College of Music (RCM) and to Oxford University Press for their support at every stage of this venture.

My own involvement in research to enhance musical performance was sparked by the multidisciplinary project "Zoning In: Motivating the Musical Mind." The project was based on a partnership between the RCM and The Leverhulme Trust, and its aim was to enable musicians to improve their performance skills and manage the high levels of stress that often accompany performance situations. From 1999 to 2002, over 150 students at the College worked with a team of scientists and musicians to learn complementary mental and physical training routines drawn from four areas: (1) physical fitness, (2) Alexander technique, (3) neurofeedback, and (4) mental skills training (each of which is detailed in Chapters 9–12 of this volume). Since the inception of the project, there has been a burgeoning interest in applied music research among RCM faculty and students, and research outcomes have now been embedded into the curriculum of the College. For this, I am most grateful to my esteemed collaborators on the project, whose commitment to enhancing musical performance has been steadfast: Christopher Connolly, John Gruzelier, Tobias Egner, Brian Hawkins, Judith Kleinman, Janet Ritterman, Yonty Solomon, Adrian Taylor, Sam Thompson, Elizabeth Valentine, and David Wasley.

Finally, I wish to acknowledge the discerning advice of several individuals who have, in numerous of ways, made this book possible: Paul Banks, Jennifer Barnes, Martin Baum, Roger Chaffin, Eulalie Charland, Steve Cothran, Bob Crenshaw, Hubert Eiholzer, Nick Foster, Scott Gwara, Patrik Juslin, Andreas Lehmann, Tânia Lisboa, Karin Marson, Gary McPherson, Janet Mills, Javier Negrin, Karen Rayner, John Rink, Karin Rosenkranz, Charles Wiffen, and David Wright.

Aaron Williamon
London, 2004

# FOREWORD

Performers strive for excellence. Whatever the genre or specialism, at the highest level, expressive impact supported by flawless execution is the goal to which professional musicians aspire, knowing that only by this means will they be able to communicate effectively with their audiences and derive personal fulfilment from performing.

But the road is far from easy, as many professional performers can testify. In his monograph on performance (with its evocative subtitle, *Revealing the Orpheus Within*) Anthony Rooley, lutenist and director of The Consort of Musicke, writes of the journey that he has traveled as a performer. He contrasts the experience of performance when it feels as though nothing can prevent the "Orpheus within" from being revealed, with earlier occasions when he found himself hampered by aspects of his mental or physical preparation. Rooley's account of a situation in which he had "never felt less like a performer" will have familiar resonances for many: there can be few musicians who have not felt at some point in their career that their performances have fallen below the standards that they believe they can achieve.

It is to assist musicians in developing personal strategies that will enable them to "feel like performers" and to perform to the best of their ability that this book, *Musical Excellence*, has been produced. Bringing together the findings of applied research on performance enhancement, based in many cases on insights and intuitions contributed by experienced performers themselves, this volume offers guidance to musicians on types of performance preparation that have been shown to help in developing and maintaining musical excellence. Many of these are techniques that can also assist in reducing the performance-related anxiety that can make the experience less rewarding than it should be.

Admittedly, there are performers for whom such uncertainties seem not to hamper artistic development. Their performances are distinctive, often from an early age, and they appear to achieve and maintain high standards in a range of performance situations, with the minimum of self-questioning. Perhaps it is because we all know performers like this that there has been some reluctance among musicians to consider scientifically grounded ways—commonly explored in other performance fields—of enabling a wider range of performers to achieve their potential artistically. It is perhaps born of the conviction that the qualities that mark out the exceptional performer—to whose artistry and prowess listeners instinctively respond—are innate, intuitive, and highly personal, and that any attempt to analyze these is unlikely to reveal strategies that lend themselves to replication or adaptation by others apparently less gifted.

This book does not discount the significance of this viewpoint but complements it by sharing research findings on performance enhancement, findings which suggest that performers who achieve greater control over their mental and physical states can often surpass expectations of their potential as artists. The areas singled out for

attention in this book focus on matters that concern many performers throughout their careers—matters such as practice, memorizing, improvisation, and sight-reading. None of the strategies are offered as "quick fixes," but they are presented to encourage musicians to select and explore for themselves approaches to fulfilling their potential as performers and maintaining this momentum in their professional lives.

It is through sustained positive experience in individual practice, in rehearsal, and in performance itself that excellence is fostered. As this book shows, applied music research has the potential to contribute to each of these areas. I hope, therefore, that those individuals and organizations involved in the education, training, and employment of skilled performers will take account of the findings that research into performance enhancement is making available, will ensure that the information is shared and used, and will do whatever they can to foster further investigation. The music profession as a whole can only benefit if performers acquire, during their training, strategies for physical and psychological preparation which complement and support their musical preparation and which they can continue to apply and refine as part of their ongoing career development.

In the introduction to her autobiographical journal, *Full Circle*, which recounts her final year on the operatic stage, Janet Baker speaks of help she received that enabled her "to understand and participate in the joy of performing"—help that focused on the need to prepare "musically, physically and psychologically for a performance and then stand aside to allow the music to speak for itself." The relief that this gave from the fear that had dogged her for much of her professional life she describes as a miracle. *Musical Excellence* does not claim to achieve miracles, but it does offer insights into the kinds of preparation that may help performers retain the joy of performing while achieving the standards that are demanded of professional musicians today. If you are a performer, I hope that you will feel encouraged to explore some of these ideas for yourself.

Janet Ritterman
Royal College of Music
London, 2004

## References

Baker, J. (1982). *Full Circle.* London: MacRae Books.

Rooley, A. (1990). *Performance: Revealing the Orpheus Within.* Shaftesbury, UK: Element Books Limited.

# CONTENTS

Contributors *page xi*

## Part I Prospects and limits

1 A guide to enhancing musical performance *page 3*
*Aaron Williamon*

2 General perspectives on achieving musical excellence *page 19*
*Roger Chaffin and Anthony F. Lemieux*

3 Managing the physical demands of musical performance *page 41*
*Christopher B. Wynn Parry*

4 Measuring performance enhancement in music *page 61*
*Gary E. McPherson and Emery Schubert*

## Part II Practice strategies

5 Strategies for individual practice *page 85*
*Harald Jørgensen*

6 Strategies for ensemble practice *page 105*
*Jane W. Davidson and Elaine C. King*

7 Strategies for memorizing music *page 123*
*Jane Ginsborg*

8 Strategies for sight-reading and improvising music *page 143*
*Sam Thompson and Andreas C. Lehmann*

## Part III Techniques and interventions

9 Physical fitness *page 163*
*Adrian H. Taylor and David Wasley*

10 Alexander technique *page 179*
*Elizabeth Valentine*

11 Physiological self-regulation: Biofeedback and neurofeedback *page 197*
*John H. Gruzelier and Tobias Egner*

12 Mental skills training *page 221*
*Christopher Connolly and Aaron Williamon*

13 Feedback learning of musical expressivity *page 247*
*Patrik N. Juslin, Anders Friberg, Erwin Schoonderwaldt, and Jessika Karlsson*

14 Drugs and musical performance *page 271*
*Robert West*

Epilogue: A note on future directions for enhancing musical performance *page 291*
*Aaron Williamon*

Index *page 293*

# CONTRIBUTORS

**Roger Chaffin** is professor of psychology at the University of Connecticut. His recent work on piano performance reflects a longstanding interest in how knowledge is represented in the mind, which previously led him to study the representation of word meaning in memory, how new words are learned, analogical problem solving, what people know of their own memory abilities, and the effects of gender on thinking. His work has appeared in journals such as *Psychological Science*; *Journal of Experimental Psychology: Learning, Memory, and Cognition*; *Cognitive Science*; and *Memory and Cognition*. His books are *Practicing Perfection: Memory and Piano Performance* (Erlbaum, 2002), *Cognitive and Psychometric Analysis of Analogical Problem Solving* (Springer-Verlag, 1991), and *Memory in Historical Perspective: The Literature before Ebbinghaus* (Springer-Verlag, 1988). He is an amateur flute player who has performed in public exactly once.

**Christopher Connolly** is co-founder of SyCon: The Sporting Bodymind Group, a London-based consultancy that has pioneered applied sports psychology in mental training since 1979. He trained with Moshe Feldenkrais in the Israeli scientist's revolutionary movement education system, which he then taught throughout Europe before establishing the first professional training program in this method in the UK. Since, Connolly has worked as coach and mental trainer with professional, national, and Olympic teams in Britain, Europe, and the United States, as well as with several diverse organizations within the business community. He has co-authored numerous books with John Syer, including *Think to Win* (Simon and Schuster, 1991), *How Teamwork Works* (McGraw-Hill, 1996), and *Sporting Body, Sporting Mind* (Simon and Schuster, 1998). In 1999, he helped launch "Zoning In: Motivating the Musical Mind," a performance enhancement research project based at the Royal College of Music, London.

**Jane W. Davidson** is reader in music at the University of Sheffield. She has a background in psychology, music, drama, and contemporary dance, and teaches areas of the curriculum related to the social psychology of music and performance, including masters courses in psychology of music and music theater studies. Her extensive publication record reflects a strong interest in bodily expression, performance, communication, and also the therapeutic and developmental aspects of artistic creativity and performance. Davidson performs as a singer/dancer, but her main artistic output is in the direction of operas. She has worked with companies such as Complicite and Drama per Musica and forthcoming projects include the medieval music drama *Ludus Danielis* with Andrew Lawrence-King.

**Tobias Egner** is a research fellow in the fMRI Research Center at Columbia University. He studied psychology at Goldsmiths College, University of London, and undertook his PhD research at Imperial College London, involving a collaboration with the Royal

College of Music, London. He has worked extensively in the field of physiological self-regulation of the electroencephalogram (EEG) and its relation to event-related brain potentials and attention, as well as music performance measures. Other research interests encompass the investigation of central executive processes in normal and altered states of consciousness (e.g. hypnosis), particularly by integrating various neuroimaging techniques, such as functional magnetic resonance imaging and EEG measurements. His personal musical development has been limited to a short-lived career as a singer with a punk rock band.

**Anders Friberg** is a researcher in the Department of Speech, Music, and Hearing at the Royal Institute of Technology in Stockholm. He has been working mainly with the synthesis and analysis of music performance, leading to a patented rule system translating the score to a performance, modeling common principles used by musicians concerning phrasing, metrical patterns, intonation, etc. Recently, Friberg has been focusing on automatic extraction of music parameters from audio and its relation to the emotional/motional qualities of performance, as well as recognition of emotional qualities in dancing. This has resulted in the development of an interactive collaborative game, in which the players interact using singing and dancing. He teaches a course on musical communication and music technology and is an active pianist with a degree from Berklee College of Music, playing mostly salsa and jazz with groups in the Stockholm area.

**Jane Ginsborg** is a lecturer in psychology in the School of Health and Human Sciences at Leeds Metropolitan University. She is also an associate lecturer with the Open University. She has a BA in music from the University of York, an advanced diploma in singing from the Guildhall School of Music and Drama, and a BA in psychology from the Open University; she gained her PhD at Keele University. Her research and teaching interests include music psychology (particularly in relation to practice and performance), as well as cognitive, developmental, and lifespan psychology. In 2002, Ginsborg won the British Voice Association's Van Lawrence Award for her research on song memorization. In her first career as a professional singer, she performed and recorded a wide range of music as a recitalist, chamber musician, and soloist with a variety of ensembles and choirs, specializing in twentieth century and contemporary music. She continues to sing and enjoys giving concerts whenever the opportunity arises.

**John H. Gruzelier** is professor of psychology in the Faculty of Medicine at Imperial College London. His main research interests concern the relation between psychology and brain function, as well as the mind–body connection in health and illness. He has published over 250 scientific articles in journals and books on topics such as schizophrenia and psychosis-proneness, hypnosis, stress, immune enhancement, self-regulation of brain rhythms, and individual differences. Gruzelier has also been co-editor of the *International Journal of Psychophysiology* since its inception in 1984 and editor of *Contemporary Hypnosis* since 2001. His research in music began with the project "Zoning In: Motivating the Musical Mind" in conjunction with the Royal College of Music, London, and has focused on self-regulation of brain electrical activity and stress reduction with a view to strengthening musical performance.

**Harald Jørgensen** is professor of education at the Norwegian Academy of Music in Oslo. Within music education and psychology of music, he has published books on

theories of musicality, the experience of music, and the history of school music education in Norway, as well as journal articles (in both Norwegian and international journals) on the same topics. He has also published several reports and presented papers on time perspectives and planning in practice (mostly related to practice among students in higher music education) and has co-edited *Does Practice Make Perfect?* (Norwegian Academy of Music, 1997) and *Musical Imagery* (Swets and Zeitlinger, 2001). In addition, he has been engaged in research in schools in sparsely populated areas and has published books and articles about general research methodology. Jørgensen has served as rector of the Academy from both 1983–1989 and 2002 onwards and was head of research there from 1995 to 2002. He studied flute at the Oslo Conservatory of Music in the 1960s.

**Patrik N. Juslin** is associate professor of psychology at Uppsala University, where he teaches courses on emotion, perception, and music psychology and is the director of the interdisciplinary research project "Feedback-learning of Musical Expressivity." He is a member of the International Society for Research on Emotions and, in 1996, received the Young Researcher Award by the European Society for the Cognitive Sciences of Music. He has published numerous articles in journals that include *Psychological Bulletin*; *Emotion*; *Journal of Experimental Psychology: Human Perception and Performance*; and *Music Perception*. Juslin also edited the book *Music and Emotion: Theory and Research* for Oxford University Press (2001) together with John Sloboda. Research interests include music performance, music and emotion, music education, and nonverbal communication. Alongside his work as a researcher, Juslin has worked professionally as a guitar player and toured internationally with blues/jazz bands.

**Jessika Karlsson** is a PhD candidate in the Department of Psychology at Uppsala University. She holds a masters degree in cognitive science (human–computer interaction) from Linköping University and a candidate degree in psychology (specializing in music psychology) from Uppsala University. In addition, she has previously worked as research assistant in developmental psychology with Professor Claes von Hofsten at Uppsala University. Her doctoral work is part of the research project "Feedback-learning of Musical Expressivity," and it combines her experience in human–computer interaction with investigations of new methods of teaching expressivity in music (i.e. computer-based feedback learning). Her musical background encompasses, among other things, being a guitarist and singer in several punk rock bands.

**Elaine C. King** is a lecturer in music at the University of Hull. She completed her doctoral work on ensemble rehearsal and performance in 2000 at Royal Holloway, University of London. Her research interests include music analysis, psychology of music, performance studies, and popular music, and her publications include "Ensemble performance" in John Rink's edited book *Musical Performance: A Guide to Understanding* (Cambridge University Press, 2002). King intends to continue developing her research activity on aspects of performance, specifically on gestures and breathing in piano and ensemble playing. She is an active cellist, pianist, and conductor, performing regularly in regional chamber and orchestral ensembles.

**Andreas C. Lehmann** is professor of systematic musicology at the Würzburg Academy of Music in Germany. After a degree in music education and doctoral work in musicology at the Hanover University of Music and Drama, he spent several years as a postdoctoral

fellow in the Department of Psychology at the Florida State University. His areas of teaching relate mainly to the psychology and sociology of music, and he also teaches music education. His research interests are in skill acquisition, practice, and performance. Lehmann was awarded a Young Researcher Merit Award from the European Society for the Cognitive Sciences of Music and is associate editor of the society's journal *Musicæ Scientiæ*. Among other instruments, he plays the piano and likes listening to any type of music, including non-Western genres.

**Anthony F. Lemieux** is currently a PhD candidate in social psychology and a research fellow in the Center for Health/HIV Intervention and Prevention at the University of Connecticut. His main research area is the development and testing of music-based HIV prevention interventions for urban adolescents in a program of research funded by the US National Institute of Mental Health. Lemieux has an extensive background in music engineering and production, as well as in his primary instrument, guitar. He has performed widely and recorded five CDs with several of his bands.

**Gary E. McPherson** studied music education at the Sydney Conservatorium of Music before completing a master of music education at Indiana University and a doctorate of philosophy at the University of Sydney. He is an associate professor of music education at the University of New South Wales and from 2004–2006 will serve as President of the International Society for Music Education. His published research addresses visual, aural, and creative aspects of musical performance, and he has conducted two major longitudinal studies with young developing musicians. As an experienced adjudicator of solo and ensemble performances and examiner for the Australian Music Examinations Board, he has a special interest in the measurement and evaluation of musical performance. McPherson is currently on the editorial boards of all flagship journals in music education, as well as being editor for *Research Studies in Music Education*. As a trumpeter, he has performed with several of Australia's leading ensembles.

**Erwin Schoonderwaldt** is currently a PhD candidate in the Department of Speech, Music, and Hearing at the Royal Institute of Technology (KTH) in Stockholm. He has a degree in physics, and in 1999, he became involved in research on vibrato production in cello playing in a collaborative project of the music, mind, machine group at the University of Nijmegen and the faculty of Human Movement Sciences at Vrije Universiteit Amsterdam. In 2001, he traveled to KTH as a young researcher in the EU-funded MOSART IHP network, where he developed a rule-based model for violin vibrato. He is also involved in two projects for developing software for music educational purposes: "Feedback-learning of Musical Expressivity," in cooperation with the Department of Psychology at Uppsala University, and "Interactive Music Tuition System," a project within the information society technologies program of the European Commission. He also plays the violin.

**Emery Schubert** holds the Vice-Chancellor's postdoctoral fellowship in the School of Music and Music Education at the University of New South Wales. His formal qualifications are in music, electrical engineering, psychology, and education. His music research has included studies in perception, statistics, software for recording responses, continuous measurement, and expressive devices used by performers. These studies

have been published in *Musicæ Scientiæ, Psychology of Music,* several book chapters, and other journals. Much of his research is connected with the fundamental questions of measurement and assessment of performances and of musical structure. The impetus for addressing these kinds of questions stems in part from his practical involvement in administration (including organization and adjudication) at school band festivals, eisteddfods, and concerts. As an active musician, he plays the French horn and has conducted and taught numerous orchestras and bands.

**Adrian H. Taylor** is reader in exercise and health psychology in the School of Sport and Health Sciences at the University of Exeter. As a fellow of the British Association of Sport and Exercise Sciences, he has worked with various government agencies on policies and reviews of evidence for the effects of exercise on health. His research has focused on facilitating more physically active lifestyles among older people, smokers, drivers, and musicians, all with an aim to enhance psychological well-being and as a therapy for stress management and self-enhancement. Taylor is associate editor for the international journal *Psychology of Sport and Exercise* and has published in a variety of leading exercise and health journals, including *Cochrane Reviews, Addiction, Health Psychology, Epidemiology and Community Health,* and the *Journal of Sport Science.* His involvement in the research project "Zoning In: Motivating the Musical Mind" at the Royal College of Music, London, has enabled him to critically extend his interests to the potentially pressured context of a musical performance.

**Sam Thompson** is a researcher in the Centre for the Study of Music Performance at the Royal College of Music, London. He read music at the University of Cambridge and subsequently completed a masters degree in psychology of music at the University of Sheffield. His research interests include the theory and practice of musical performance evaluation (the major focus of his doctoral work, based at Royal Holloway, University of London) and the breakdown of musical abilities in performance situations, as well as various philosophical issues in music and psychology. Intermittently active as a composer and arranger, Thompson has re-scored operas for professional touring productions and had his own works performed in venues including London's Wigmore Hall. He is also a clarinetist and, recently, uilleann piper.

**Elizabeth Valentine** is professor of psychology at Royal Holloway, University of London, where she is director of the Psychology of Music Research Group. She is also an advisor to the Centre for the Study of Music Performance at the Royal College of Music, London. She is a member of the European Society for the Cognitive Sciences of Music as well as the Society for Education, Music, and Psychology Research. Valentine has published widely on theoretical and experimental psychology. Her publications in the psychology of music include articles on performance anxiety, Alexander technique, musical responsiveness in relation to blocked capacity for intimacy, the effect of singing on mood, relations between reading and musical ability, cognitive structures in music practice for memorized performance, and background music as an aid to autobiographical recall in dementia. She is an amateur singer and woodwind player.

**David Wasley** is a lecturer in exercise and health psychology at the University of Wales Institute, Cardiff. He completed a masters degree in sport and exercise behavior at

Southern Illinois University in 1999 and is currently working toward his PhD. He has published a number of papers on the psychophysiology of exercise and has over a decade of experience working at both practical and professional levels within the health and exercise industry.

**Robert West** is professor of health psychology in the Cancer Research UK Health Behaviour Unit at University College London. He specializes in substance dependence and particularly tobacco use. He is deputy editor-in-chief of the journal *Addiction* and co-author of national guidelines that have led to the development of comprehensive treatment services to aid smoking cessation in the UK's National Health Service. West also has an interest in the psychology of music and has co-edited texts on music comprehension and music therapy. He enjoys all types of music but particularly jazz, rock, and blues. He plays guitar and occasionally gigs with his sons Jamie and Daniel.

**Aaron Williamon** is the research fellow in psychology of music at the Royal College of Music, London, where he heads the Centre for the Study of Music Performance. He also holds a research fellowship in the Faculty of Medicine at Imperial College London. His research focuses on music cognition, expert performance, and (in particular) applied psychological and health-related initiatives that enable musicians to perform at their peak. He has led a number of research projects across these areas, including "Zoning In: Motivating the Musical Mind." In addition, Williamon is interested in how audiences perceive and evaluate musical performances and, in 1998, was awarded the Hickman Prize by the Society for Education, Music, and Psychology Research for his work on this topic. He has performed as a trumpeter in chamber and symphony orchestras, brass bands, and brass quintets in both Europe and North America.

**Christopher B. Wynn Parry** is a senior consultant physician and trustee to the British Association for Performing Arts Medicine, a charity that provides free advisory service to performing artists. Previously, he has held posts as a consultant adviser in rheumatology and rehabilitation to the Royal Air Force and as the director of rehabilitation at the Royal National Orthopaedic Hospital and at King Edward VII Hospital, Midhurst. He was co-founder and first president of the International Rehabilitation Medicine Association and is a past president of the British Society for Surgery of the Hand. His book *Rehabilitation of the Hand* (Butterworth, 1981) is now in its fourth edition, and he is co-editor of *The Musician's Hand* (Martin Dunitz, 1998) with Ian Winspur. His professional interests in the medical and psychological problems of performing artists is informed by both his past performing experience as a bass trombonist and his current involvement as a tenor.

PART I

# PROSPECTS AND LIMITS

CHAPTER 1

# A GUIDE TO ENHANCING MUSICAL PERFORMANCE

AARON WILLIAMON

Musicians routinely encounter an elaborate array of mental and physical demands during practice and performance, having to process and execute complex musical information with novel artistic insight, technical facility, and a keen awareness of audiences' expectations. One of the most exciting challenges facing music researchers is to develop ways of assisting performers to meet those demands efficiently and effectively. For various methodological and technological reasons, progress has been disappointingly slow—at least compared with other performance domains such as sports, where applied research has long informed the training of practitioners. Increasingly, however, researchers are forging cross-disciplinary collaborations and generating innovative methods for investigating how exceptional musical performances can be produced.

These collaborations serve as the inspiration for this volume, which brings together their emergent findings for the first time in a single collection. It does so in three parts. Part I details the prospects and limits of performance enhancement research and sets out ground rules for achieving musical excellence. What roles do innate talent, environmental influences, and sheer hard work play in attaining eminence (Chapter 2)? How can musicians best manage the physical demands of a profession that is intrinsically arduous, throughout a career that can literally span a lifetime (Chapter 3)? How can performers, teachers, and researchers effectively assess and reflect on performance enhancement for themselves, their colleagues, and their students (Chapter 4)?

Part II presents strategies for increasing the effectiveness and efficiency of practice. These are examined generally for the individual (Chapter 5) and ensembles (Chapter 6), and specifically for the tasks of memorizing (Chapter 7), sight-reading, and improvising music (Chapter 8). Musicians inevitably spend vast amounts of time and energy acquiring and refining their skills, but are there particular rehearsal strategies that they can employ to produce better performance results or to achieve the same results more quickly? What implication does existing knowledge of human information processing and physical functioning have for musical learning and practice?

Part III introduces readers to pioneering techniques and interventions from fields including exercise science, psychophysiology, sports psychology, cognitive science, and medicine that have been used to enhance performance in music. The selection of techniques is limited to those that have been evaluated scientifically, the goal of this being to provide an objective critique of the rationale and practical efficacy of each. They include research on physical fitness (Chapter 9), Alexander technique (Chapter 10),

biofeedback and neurofeedback (Chapter 11), mental skills training (Chapter 12), musical expressivity and communication (Chapter 13), and the effects of drugs on performance (Chapter 14). While some of these interventions may already be well known to musicians, many have only recently been applied and tested in the musical domain.

This first chapter provides a concise introduction to each part of the book, previewing the chapters contained therein and placing them in the broader context of performance enhancement research. It then considers the extent to which the physical and psychological skills required of today's performers are amenable to improvement through scientifically tested, systematic training procedures. Finally, it discusses how one may begin to embark on a personal course of enhancement using the pragmatic advice offered in this volume. The chapters in Parts II and III describe particular strategies and techniques and their effects on musical performance; this chapter takes a contrasting approach, exploring how enhancement techniques may be selectively applied to yield specific outcomes.

## 1.1 Introduction to Parts I, II, and III

### 1.1.1 Part I: Prospects and limits

Throughout this book, recommendations are drawn from findings of empirical research so that the practitioner can reflect on how (or whether) to implement certain training approaches in search of his or her own musical fulfillment. Part I lays down explicitly, in four chapters, some of the fundamental principles of performance that musicians should consider before moving on to the practical advice offered in Parts II and III.

Chapter 2, "General perspectives on achieving musical excellence," considers the source of exceptional performance abilities. The debate as to whether these abilities arise from innate gifts or environmental influences has long been fuelled by anecdote and personal conviction. Chapter 2 gives a balanced appraisal of the scientific debate and, in doing so, elucidates salient characteristics of cognition (or mental activity) common to all expert musicians. By removing the ambiguity and mystique that often surround high-level musical achievement, performers should be in a better position to know how best to direct their efforts and those of their students.

Inevitably, achieving high performance standards is an extremely effortful endeavor. It is not uncommon for musicians to experience physical challenges that result from that effort. Chapter 3, therefore, offers musicians guidance for "Managing the physical demands of musical performance." Based on decades of medical experience, it provides an account of common physical problems that befall performers, an analysis of the source of those problems (e.g. specific instruments and practicing habits), and frank advice for how they can best be avoided. Although performance success very often hinges upon mental agility and acuity, musical ideas are ultimately expressed through physical skill. To do this effectively—and to maintain this skill over the course of a long career—the preventive measures presented in Chapter 3 are vital.

For musicians to enhance their performance, it is instructive to understand how performances are perceived and evaluated by others. Chapter 4, "Measuring performance enhancement in music," sets out some of the musical and psychological factors that

influence the assessment of performance quality. It distinguishes among those factors that are under the performer's control, those that are determined by listeners, and those typically beyond the command of both. Of course, the extent to which a performer will let audience opinion influence his or her musical decisions is highly individual-specific; nevertheless, greater knowledge of how to communicate those decisions and the impact they will have on the audience can be distinctly advantageous.

Musicians can follow any number of possible routes in order to perform at the peak of their ability. Regardless of the exact path, the principles established in Part I provide indispensable starting points.

### 1.1.2 Part II: Practice strategies

The direction and quality of one's practice are integral to performance enhancement, whether for the fulfillment of short-term goals (e.g. performing well in an upcoming concert) or the realization of long-term ambitions. Part II explores the characteristics of quality practice for a variety of performance circumstances and relates them to core psychological mechanisms upon which they rely.

Chapter 5, "Strategies for individual practice," sets the tone for Part II by reviewing research on strategies for planning practice, executing rehearsal plans, evaluating the success of that execution, and (perhaps most importantly) reflecting on one's ability to exploit strategies for optimal benefit. The overt recognition of the processes that underlie performance skills, as promoted in Chapter 5, can be fundamental to performance enhancement and can lead to a clearer understanding of which strategies actually work and why. Indeed, self-reflection and self-evaluation are prerequisites for making the most of any strategy and technique discussed in this volume or encountered through any other source.

The ability to work with others to produce cohesive, well-coordinated performances is a hallmark of musical expertise. In such situations, verbal, aural, and visual cues are often established and shared between coperformers. Chapter 6, "Strategies for ensemble practice," weighs up the importance of these cues in driving successful rehearsals and performances. It examines a variety of ensemble types (from cello-piano duos to symphony orchestras) and the functions of members within those ensembles.

Chapter 7 explores "Strategies for memorizing music." The memorization of music is commonplace within virtually all performance traditions—although the extent to which precise recall is compared against established standards varies immensely between genres. A substantial body of evidence has been accumulated within the wider psychological literature that informs the process of committing music to memory, as well as its subsequent retrieval. Chapter 7 reviews that evidence alongside findings from interview, observational, and experimental studies specifically with musicians.

Finally, Chapter 8 focuses on "Strategies for sight-reading and improvising music." The abilities to read fluently at sight and improvise convincingly hold a curious and somewhat paradoxical status. On one hand, they are highly prized and seen as indicators of great musical ability. On the other, they are often regarded as natural gifts that one either possesses or does not. Chapter 8 considers how these two skills share a common etiology, both in their use of similar mental processes (e.g. pattern recognition) and their generative, "online" nature. In addition, it explores how these two tasks can

be systematically enhanced through structured training and how the augmentation of generative musical abilities as such could actually provide new insight for improving other musical skills.

Within the research reviews provided throughout Part II, specific recommendations for enhancing practice efficacy are bulleted in close proximity to the research from which they are drawn. In this way, their derivation is intended to be transparent for the reader. In some cases, the practice-strategies presented may be the same as those that musicians would employ automatically or as a matter of course. This is not surprising, since much of the research in this area has arisen directly from practitioners' intuitions; in fact, it is encouraging when the findings of empirical research and common practice correspond. However, not all musicians have (or follow) the same intuitions. By highlighting what research has shown, the chapters in Part II aim to supply a greater number of musicians with *explicit* access to verified strategies. Moreover, by organizing these strategies according to the psychological mechanisms to which they relate, it is hoped that practitioners will find new and useful ways of thinking about them.

### 1.1.3 Part III: Techniques and interventions

The performance enhancement methods presented in Part III represent the first collection of extramusical, scientifically validated techniques for enhancing musicians' achievement. They originate from a wide range of disciplines and, naturally, impact on performance in different ways. Each chapter provides a review of the theory behind the presented findings so that readers will understand why these interventions were first applied and are relevant to music. The techniques and interventions are ordered from the more physical (Chapters 9–10) to the psychological (Chapters 11–13) to the pharmacological (Chapter 14); however, they *all* address issues of both mental and physical significance for the musician.

It is well documented that keeping physically fit can be an effective strategy for managing stress. This includes the anxiety experienced in conjunction with specific events and the more general anxiety that can accumulate from working in a demanding, high-profile profession. Chapter 9, "Physical fitness," explores how musicians can integrate physical activity into their lifestyles in ways that can optimize their musical performance. In particular, both regular and single sessions of aerobic-type exercise have been shown to enable musicians to control physical and psychological responses to stress.

Chapter 10, "Alexander technique," provides a critique of this popular method of kinesthetic re-education, which is aimed at eliminating unnecessary tension and developing efficient use of the body. Since its development by F. M. Alexander in the first half of the twentieth century, an increasing number of studies have provided scientific evidence for both its anatomical and physiological benefits. They have shown positive effects of Alexander technique training on breathing, heart rate and blood pressure, posture and bodily use, quality of musical performance, and mental attitude.

Chapter 11, "Physiological self-regulation: Biofeedback and neurofeedback," details techniques that have been designed to empower individuals to control physiological processes through mental activity. The general rationale behind biofeedback is that physiological responses (such as heart rate, blood pressure, and in the case of *neuro*feedback, the electrical activity of the brain) can be recorded and fed back to an individual

so that he or she can eventually gain voluntary control over those responses. Such physiological self-regulation holds great potential for enhancing musical performance. Notably, biofeedback of electric muscle activity has been used successfully to reduce muscular tension in certain types of instrumentalists. Also, neurofeedback training of "slow wave" activity has been shown to facilitate remarkable improvements in musicians' performance skills.

Being able to achieve conscious control over one's mental and physical states can prompt successful performance. One route to such control is through the acquisition and refinement of mental skills aimed at relaxing and reducing tension, realizing one's artistic possibilities, and optimizing the time spent working toward goals. Chapter 12, "Mental skills training," surveys fundamental skills that are common to elite performance across disciplines and describes how they can be propagated through select relaxation and visualization techniques.

Chapter 13, "Feedback learning of musical expressivity," explores the very nature of musical expression and evaluates a novel method for teaching expressivity based on recent advances in musical science, psychology, technology, and acoustics. The empirically based approach presented here purports to grant musicians greater control over the expression of their own musical ideas and, importantly, the tools for conveying those ideas to listeners. The chapter also assesses the potential for implementing this novel method within music education contexts.

Finally, Chapter 14, "Drugs and musical performance," focuses on the multifarious effects of "lifestyle," "prescription," and "illicit" drugs on performance and general well-being. Regardless of their possible benefits or hazards, drugs do factor into many musicians' pre-performance routines and lives generally. A number of studies have examined the scope of such drugs as alcohol, sedatives, and beta-blockers in controlling symptoms of anxiety and improving performance quality. Here, a forthright and comprehensive appraisal is offered of these substances' impact on the demands of performance.

Collectively, the chapters in Part III stand as clear evidence that applied research can facilitate musicians' strive for performance excellence. The strength of this evidence lies in the fact that results have been obtained in real-world performing contexts and in relation to outcome variables that musicians themselves deem important (e.g. performance quality, anxiety reduction, and respiratory efficiency). In some cases, the extent to which the practitioner can directly implement a given intervention for his or her own benefit will vary, depending on access to technology and expert assistance. Nevertheless, these research initiatives represent a positive step toward establishing suitable educational and professional environments where artists are supported in their acquisition of secure, holistic strategies for maintaining and enhancing their performance skills.

## 1.2 The potential for enhancement

### 1.2.1 Physical and psychological enhancement

Musical performance, at its best, is indicative of the upper limits of human physical and mental achievement. The training required to reach such lofty heights can be substantial,

and invariably, the acquisition and refinement of requisite musical skills develop through the implementation of high-quality practice and the accrual of high-quality training and experience (see Chapter 2 for an extended discussion). Although there is no ideal performance standard that pervades all music, either within or across genres, the prospects for using specifically targeted training methods (even if adopted from domains other than music) to enhance the physical and mental skills needed for performance seem good. The chapters of this book set out to highlight ways for performers to make the most of their existing practice, training, and experience and to give them additional tools for acquiring and developing new skills.

In relation to long-term engagement in music, it is possible for musicians to experience both anatomical and broader psychological changes that permit the effective implementation of their abilities. As a consequence of repeated exposure to the physical demands of extensive practice, musicians are susceptible to changes in musculature, bone structure, circulation, and respiration to the same degree as expert performers in other domains (see Allard & Starkes, 1991; Ericsson *et al.*, 1993; Ericsson, 1996; Williamon & Valentine, 2000). For example, it will come as no great surprise that focused training can instill greater precision, speed, and stamina in specific muscle groups that contribute to music making; after all, countless generations of musicians have practiced scales, études, and the like for precisely that purpose.

Furthermore, extended training can also instigate performance-related refinements in the body's nervous system. Since the nervous system lies at the center of all skilled endeavor, its most fundamental characteristics are set out here as a precursor to discussion in subsequent chapters. As a matter of routine, the nervous system controls and monitors our basic support systems (e.g. respiration and circulation); directs movements; maintains balance and posture; receives, records, and interprets information from the outside world; and enables the generation of new ideas, communication, and emotional responses. It is a set of interrelated divisions that, at the most global level, is partitioned into the *central* and *peripheral* nervous systems (see Figure 1.1). The central nervous system (CNS) is composed of the brain and spinal cord and is connected to the rest of the body through nerve cells of the peripheral nervous system (PNS). The PNS is divided into the *somatic* and *autonomic* systems. The somatic system contains sensory neurons that carry impulses to the CNS and motor neurons that carry impulses from the CNS to the skeletal muscles. It allows for the production of voluntary movements such as walking, running, or throwing, as well as involuntary movements such as the adjustment of balance and posture. The autonomic system is responsible for the functioning of vital internal organs and glands (e.g. the heart, lungs, and stomach). Neurons within this system link these organs to the CNS and allow the brain to regulate life-supporting processes without the need for voluntary or conscious involvement. The autonomic system is divided further still into the *sympathetic* and *parasympathetic* systems. The former enables "fight or flight" responses that prepare the body physically for action during periods of high excitement or stress (e.g. adrenalin is secreted, glucose is released, heart rate accelerates, the bronchi dilate, and salivary flow decreases). The latter operates to conserve resources during periods of calm (i.e. heart rate is slowed, the bronchi constrict, and salivary flow is stimulated; for further discussion, see Rosenzweig & Leiman, 1989).

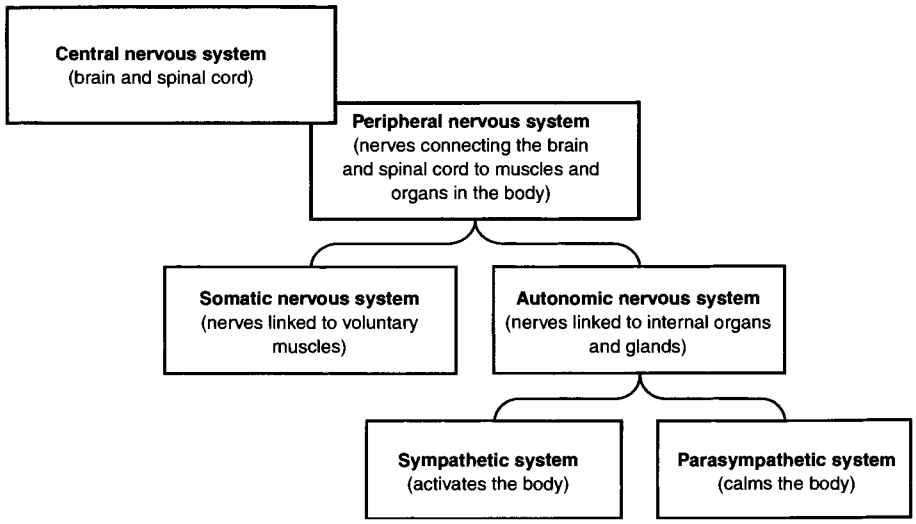



Figure 1.1 Major divisions of the human nervous system.

For musicians, the range and extent of refinements that are possible at all levels of the nervous system are extraordinary, from the higher-order functioning of the CNS to the primitive "fight or flight" responses of the autonomic system (see Halpern & Bower, 1982; Salthouse, 1991; Ericsson *et al.*, 1993; Fyk, 1995; Gabrielsson, 1999, 2003; Altenmüller & Gruhn, 2002; Williamon & Valentine, 2002). Evidence for the advancement of mental and physical skills within music has been so great over time that some authors have claimed that the violin virtuoso Paganini would indeed cut a sorry figure if placed upon the modern concert stage. Regardless of the extent to which this may be true, the wide availability of recordings and global marketing and distribution of music has meant that today's performers are increasingly held to exacting standards and required to meet a host of diverse expectations. Throughout this book, the chapters detail the existing research on and benefits of given strategies and techniques for reaching such high standards. In some cases, the professed benefits overlap between chapters; therefore, Table 1.1 offers a summary of the areas of physical, psychological, and musical improvement that emerge.

### 1.2.2 Managing and overcoming obstacles

The emphasis of this book is decidedly on *enhancing* performance rather than simply surviving the demands of it. One area, however, that has received considerable attention from researchers and practitioners alike is the management of performance anxiety. "Stage fright" has been documented as a problem for many performers, with one study indicating that 24% of orchestral musicians are regularly affected by it and 15% finding it a severe limitation to their performance (Fishbein *et al.*, 1988). Likewise, distinguished and veteran soloists such as Maria Callas, Enrico Caruso, Pablo Casals,

Table 1.1 A summary of the (a) physical, (b) psychological, and (c) general musical skills addressed in *Musical Excellence*. As musical performance ultimately relies on the overarching coordination of multifaceted skills, the delineation of physical, psychological, and musical areas for enhancement provided here is approximate and is intended only as a general guide

| Area for enchancement | Chapter | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| *(a) Physical skills* | | | | | | | | | | | | | |
| General physical fitness; health | | × | | | | | | × | | | | | × |
| Relaxation; tension reduction | | × | | | | | | × | × | × | × | | × |
| Musculoskeletal function | | × | | | | | | | × | × | | | × |
| Cardiovascular function | | | | | | | | × | × | | | | |
| Respiratory function | | | | | | | | | × | | | | |
| Movement generation; coordination | | × | × | | × | | × | | × | | | | × |
| *(b) Psychological skills* | | | | | | | | | | | | | |
| General psychological well-being | | | | | | | | × | × | | | | × |
| General stress reduction | | | | | | | | × | × | × | × | | × |
| Management of stage fright | | | | | | | | × | × | × | × | | × |
| Concentration; attention | × | | | | | | × | × | × | × | × | | |
| Memory | | | | | | × | | | | × | | | |
| Sensory awareness; perception | | | | × | × | × | × | | × | | × | × | |
| Motivation; attribution | × | | | | | | | | | | | | |
| Self-efficacy; confidence | × | | × | | | | | | | × | × | | |
| *(c) General musical skills* | | | | | | | | | | | | | |
| General practice efficacy | × | × | | × | × | | | | | | × | × | |
| Goal setting | × | | | × | | | | | | | × | | |
| Time management | × | | × | | | | | | | | × | | |
| Problem identification | × | | | × | | | × | | | | | × | |
| Self-evaluation of skill | × | | × | × | | | | | | | | × | |
| Evaluation of others' skill | | | × | | | | | | | | | | |
| Memorizing music | | | | | | × | × | | | | × | | |
| Sight-reading | | | | | | | × | | | | | | |
| Improvisation | | | | | | | × | | | | | | |
| Analytical music skills | | | | | | × | × | | | | | | |
| Expressive range | | | × | | | | | | × | × | × | × | |
| Expression of emotions | | | | | | | | | | × | | × | |
| Teaching expressivity | | | | | | | | | | | | × | |
| Communication within the ensemble | | | × | | × | | | | | | | | |
| Communication with the audience | | | × | | | | | | | × | × | × | |

Vladimir Horowitz, and Ignacy Paderewski have been noted as sufferers (Valentine, 2002; see also Schonberg, 1963).

To make the most of any technique for managing stage fright, it is often constructive first to understand its symptoms and causes. Salmon (1990) defines performance anxiety as "the experience of persisting, distressful apprehension about and/or actual impairment of performance skills in a public context, to a degree unwarranted given the individual's musical aptitude, training, and level of preparation" (p. 3). This apprehension can manifest itself in three general types of interrelated symptoms. *Physiological* symptoms are natural, sympathetic responses of autonomic nervous system. They include increased heart rate and palpitations, hyperventilation, dry mouth,

sweating, nausea, diarrhea, and dizziness. Valentine (2002) describes these symptoms as follows:

> This fight-flight response, which assisted our hunter-gatherer forebears in fleeing large animals, is highly detrimental to musicians requiring dexterity and fine muscular control over their instruments. Trembling limbs and slippery fingers are likely to hinder rather than help the performer. In addition, this autonomic arousal may have become associated with fear as a result of past experience. Increased arousal generally leads to a narrowing of the focus of attention, which may also be deleterious (p. 168).

*Behavioral* symptoms include shaking, trembling, fidgeting, stiffness, and dead-pan expression. Unfortunately, these symptoms can exude clear signals to others that the performer is nervous (thus potentially undermining his or her effort to communicate effectively with the audience) or actually impair performance itself (e.g. through the shaking of a violinist's bow arm). *Cognitive* (or mental) symptoms consist of subjective feelings of anxiety and negative thoughts about performing. These are often associated with the overidentification of self-esteem and self-worth with performance success. As a result, such worry can lead "to poor concentration, diverting attention and wasting valuable resources, possibly also acting as a cue to increase anxiety further" (Valentine, 2002, p. 169).

Wilson (2002) has identified three underlying causes to these symptoms. First, *trait anxiety* is a person's disposition to become anxious in response to situations of social stress. It may stem from an overreactive autonomic nervous system or low self-esteem and is generally related to neuroticism and social phobias (Steptoe & Fidler, 1987; Cox & Kenardy, 1993). Second, the degree of task mastery can impact greatly on one's subjective feelings of stress. Anxiety is likely to be lower when the person is well prepared or finds the task simple, than when he or she is under-rehearsed or when the task is complex. Adequate preparation and realistic appraisal of task requirements, therefore, are essential for eliminating this source of worry. Finally, situational stress is a pervasive determinant of stage fright (Leglar, 1978; Abel & Larkin, 1990; Brotons, 1994; LeBlanc *et al.*, 1997), and factors such as the size and make-up of the audience and whether a performance is from memory are classic examples.

Trait anxiety, task difficulty, and situational stress interact in multifarious ways to contribute to one's perception of threat in social settings (see Hamann & Sobaje, 1983; Cox & Kenardy, 1993), which in turn can trigger basic survival instincts. Unfortunately, as Wilson and Roland (2002) point out, "running from or attacking the audience is seldom appropriate" (p. 47). Beck and Emery (1985) posit that the perception of threat is usually linked to an overestimation of the severity of the feared event, an underestimation of coping resources (i.e. what can be done about it), or an underestimation of rescue factors (i.e. what other people can do to help). Therefore, it is reasonable to conclude that musicians can profitably employ techniques to reduce their fear of performance. Steptoe (1989) lists the following strategies used by professional musicians to reduce anxiety (some of whom reported more than one strategy): 38% deep breathing, 28% distraction, 23% muscle relaxation, 22% alcohol, and 12% sedatives. In a survey of almost 200 musicians, Wolfe (1989) reports a total of 478 strategies used, of which approximately two-thirds were "emotion-focused" (e.g. "I try to sit quietly and pray for

calmness before I sing") and about one third "problem-focused" (e.g. "Much like an athlete I stay in shape and have a regular warm-up routine, so that I can rely on my body to do what it has to do to operate my instrument" (Wolfe, 1990, p. 34).

Despite the debilitating nature of physiological, behavioral, and cognitive symptoms of anxiety, performers should be aware that arousal can, in fact, *benefit* performance. The Yerkes–Dodson law (Yerkes & Dodson, 1908) states that performance is best when arousal is at moderate levels—that is, when the performer is neither too relaxed nor too anxious. Unusually low arousal can result in dull or lifeless performances, whereas very high arousal can lead to physical and psychological impairment of ability. This long established relationship is commonly referred to as the *inverted U function*, as shown in Figure 1.2 (see Steptoe, 1983, for confirmation of the Yerkes–Dodson law within music). Undeniably, many musicians do claim that the thrill or "buzz" of the performance situation fosters spontaneity and encourages new musical insight. Some researchers, however, caution that the relationship between arousal and performance quality is often not so straightforward. Fazey and Hardy (1988) make a useful distinction between the effects of physiological and cognitive components on performance quality. They argue that when cognitive anxiety is low (i.e. when there is little fear of failure and its consequences), the Yerkes–Dodson law holds; however, when cognitive anxiety is high, performance quality is susceptible to a catastrophic drop, from which immediate recovery is difficult to achieve (see Figure 1.3). The basic idea, based on data from performance in sports, is that excessive apprehension and rumination can lead to a vicious spiral of negative thoughts and, hence, can cause performance to collapse.

Several chapters of this book offer advice on reducing unnecessary physiological, behavioral, and cognitive symptoms of performance anxiety (commonly measured and labeled as *state anxiety*, i.e. a person's level of anxiety at a given moment), as well as how those symptoms interact with a musician's general disposition for becoming anxious (i.e. *trait anxiety*; Spielberger *et al.*, 1983). Table 1.1 offers a synopsis of where this advice appears in this volume (for comprehensive critiques of methods specifically designed to manage anxiety, see reviews by Brodsky & Sloboda, 1997; Steptoe, 2001; Valentine, 2002; Wilson, 2002; Wilson & Roland, 2002).

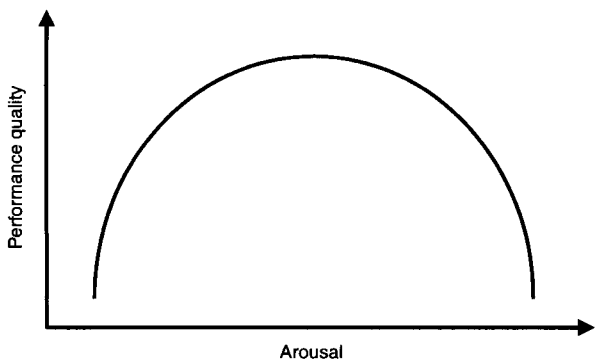



**Figure 1.2** The Yerkes–Dodson inverted-U curve, depicting the relationship between arousal level and performance quality.

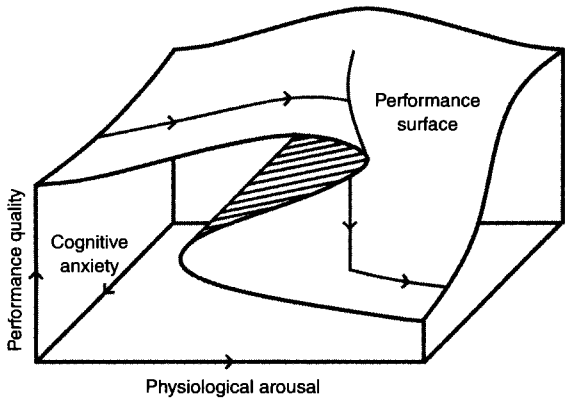



**Figure 1.3** The catastrophe model of the relationship between anxiety and performance. (From Fazey & Hardy, 1988. © Sports Coach UK, used by permission.)

## 1.3 Selecting a personal program for enhancement

Performance enhancement in music can take various forms, operating at different levels and on different timescales. The practical recommendations offered here are ultimately aimed at skilled performers (i.e. those approaching entry into higher music education and above), in that certain performance enhancement interventions may not, for instance, benefit novice musicians who have yet to acquire elementary performance proficiency. Even so, the contributors have prepared their chapters with educational implications in mind. Thus, it is hoped that this volume will serve as a stimulus for in-depth student–teacher discussions of a range of possibilities for exploiting physical and mental skills in practice and performance.

But how exactly might performers or teachers identify the most efficacious route for their own personal enhancement or that of their students? A helpful first step is to carry out a realistic assessment of individual strengths and weaknesses in skill. The number of differences that can exist between musicians in this respect, even those at the highest levels of performance, is tremendous. This is to be expected, given that individual differences pervade so many aspects of human functioning, including such well-documented factors as personality (Eysenck, 1981), attention (Hunt, 1980), reasoning (Sternberg, 1977), and intelligence (Sternberg, 1985; for discussions on the evolutionary roots of individual differences, see Lancaster, 1975; Baer & McEachron, 1982). Also, individual differences in artistic insight are commonly seen as positive characteristics and are features by which audiences of all types make their decisions about which concerts to attend and what records to buy. Hence, a candid evaluation of personal skills can open up a clear path for the musician to accentuate strengths and eliminate weaknesses, as well as to gain greater control over shaping and developing a unique musical persona. Classical musicians typically spend years working privately with teachers who, with expert knowledge, assist them in reflecting on how best to improve their skills. For professional players (who inevitably have less time to pursue further private study) or for other performers from nonclassical traditions where

extended tuition is not the norm, the ability to evaluate oneself and to make astute decisions about personal improvement is essential for artistic longevity. The development of this ability can be bolstered by learning about constituent physical and psychological skills that support musical performance (as detailed throughout this book) and by generally working to expand one's *metacognitive knowledge* (i.e. knowledge about learning itself, or simply put, thinking about thinking; see Nielsen, 1999; Barry & Hallam, 2002; and Chapter 5 of this volume).

After identifying strengths and weaknesses, careful attention should then be given to the selection of appropriate methods for achieving desired outcomes. Table 1.1 is provided as an aid for this in relation to the outcomes presented here. Although the empirical evidence for these has been obtained through controlled scientific investigations (which suggests that the benefits are potentially generalizable to the entire population of musicians), not every strategy or technique will be suitable or have the exact same impact for every musician. This could be determined by the performer's individual preferences for how to practice and learn, open-mindedness toward using a particular strategy or technique, and access to adequate training resources and facilities. Although a performance enhancement intervention may have scientific validation, this does not mean that it will necessarily work for a person who is unwilling to engage it.

Granted, it may be difficult to decide whether an enhancement method is entirely appropriate until hands-on experience can be obtained. As a natural next step, therefore, performers should use trial-and-error wherever possible in order to grasp how particular interventions can be used in specific performance situations and for more long-term gains. Furthermore, some techniques may have immediate and dramatic effects on one's physical and mental state; consequently, performers are advised to experiment with interventions conservatively, at a pace where the effects will be easily predictable and in situations where there is flexibility for experimentation (e.g. before rehearsals or leading up to low-pressure performances).

Finally, it is important to stress that this book does not supply a series of quick-fix solutions to musical problems. Rather, it provides a broad review of techniques that have been shown to optimize performance through careful and deliberate training. In many cases, the full potential of a performance enhancement method can only be realized through patience, concerted effort, and attention to subtle detail. Thus, performers should be aware of the timescale on which an intervention has been shown to deliver benefits and be willing to persevere for as long as is necessary. Indeed, readers should note that the benefits displayed throughout this book have been documented over different time periods and measured in diverse ways. Chapters 9–12, for instance, all stem from related investigations carried out by one research team and, to some extent, afford comparisons between techniques. Figure 11.4 shows the extent to which conservatory students receiving instruction in exercise, mental skills, three types of neurofeedback, and Alexander technique improved their scores on performance examinations from before to after training (which spanned approximately 16 weeks). In this case, the students receiving "alpha/theta" neurofeedback training improved their performances far more than those in any other group. At first glance, one could easily conclude that this type of neurofeedback is in everyway superior to the other interventions; however, it should be noted that the 16-week period between the two performances

was ideally suited to the timescale needed to administer the designated 10 alpha/theta neurofeedback sessions. The other techniques, conversely, may feasibly require a longer period of involvement and may also present different challenges for the musician to overcome (e.g. exploiting mental skills training to enhance performance is not easy and can take years of deliberate use, and some argue that Alexander technique should be viewed more as a way of life than as a temporary intervention). In addition, it is worth noting that not every technique is directly concerned with enhancing quality judgments of performance. As is shown in Chapter 9, being physically fit can offer a range of physical and psychological advantages for the musician without directly changing his or her ability to give higher-quality performances. In time, this benefit may come as either a direct or an indirect consequence of exercising, but this will depend largely on how the individual integrates it into his or her own routine for performance preparation.

## 1.4 Conclusions

The basic precept of this book is that the findings of applied music research can be used to inform the training and teaching of performers. A range of strategies and techniques is presented so that practitioners can select an enhancement approach that they feel is best for themselves or for those who rely on their informed counsel. This is not to say, however, that the methods presented in this book are exhaustive of the full range of holistic body–mind approaches for enhancing musical performance. Ostensibly, there could be countless additional interventions capable of sharpening physical and mental skills but which have yet to be validated through systematic research or which are not currently supported by adequate theoretical frameworks. Certainly, musicians have been enhancing their performance for centuries without the aid of applied research—the Tchaikovsky Violin Concerto, for instance, once considered unplayable, is now one of the standard works in the instrument's repertory.

So why should performers and teachers concern themselves with the findings of research? By knowing that the benefits of a given approach have been rigorously tested and empirically demonstrated, practitioners are offered a certain level of quality assurance, with greater transparency as to how the technique works, an unbiased assessment of how it can improve musical skill, and a realistic appraisal of how quickly such improvement can be expected. In addition, for those musicians who have developed a unique pedagogical and/or performance enhancement method themselves, the scientific investigation of its effects can provide an excellent means for demonstrating reliability and validity.

It seems that much of the future for exploiting applied research to enhance musical performance lies in the direct collaboration of performers, teachers, and researchers, all working toward an agenda of addressing musically relevant issues through methods that are scientifically sound. One of the limitations with the extant research in this field is that, for the most part, it has been conducted within the confines of the Western classical tradition; hence, further research is needed to establish the breadth and applicability of existing findings across the whole of music. Where appropriate, the chapters in this volume examine the relevance of research results for nonclassical performance

genres and consider possible benefits for all types of musicians: instrumentalists, vocalists, and conductors. Nevertheless, given that each chapter is underpinned by physical and psychological principles relevant to *all* of human performance, it is reasonable to argue that the recommendations offered here are pertinent to any performance tradition that demands extensive dedication and resilience, unique artistic vision, and effective communication.

## Acknowledgments

I am grateful to Elizabeth Valentine, Roger Chaffin, Tobias Egner, and Sam Thompson for very helpful comments and suggestions on an earlier version of this chapter.

## Further information and reading

Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.), *The Psychology of Music* (2nd ed., pp. 501–602). London: Academic Press.

Parncutt, R., & McPherson, G. E. (Eds.). (2002). *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning.* Oxford: Oxford University Press.

Rink, J. (Ed.). (2002). *Musical Performance: A Guide to Understanding.* Cambridge: Cambridge University Press.

Wilson, G. D. (2002). *Psychology for Performing Artists: Butterflies and Bouquets* (2nd ed.). London: Whurr.

## References

Abel, J. L., & Larkin, K. T. (1990). Anticipation of performance among musicians: Physiological arousal, confidence, and state anxiety. *Psychology of Music, 18,* 171–182.

Allard, F., & Starkes, J. L. (1991). Motor-skill experts in sports and other domains. In K. A. Ericsson & J. Smith (Eds.), *Toward a General Theory of Expertise: Prospects and Limits* (pp. 126–152). Cambridge: Cambridge University Press.

Altenmüller, E., & Gruhn, W. (2002). Brain mechanisms. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 63–81). Oxford: Oxford University Press.

Baer, D., & McEachron, D. L. (1982). A review of selected sociobiological principles: Application to hominid evolution I. The development of group social structure. *Journal of Social and Biological Structures, 5,* 69–90.

Barry, N. H., & Hallam, S. (2002). Practice. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 151–165). Oxford: Oxford University Press.

Beck, A., & Emery, G. (1985). *Anxiety Disorders and Phobias: A Cognitive Perspective.* New York: Basic Books.

Brodsky, W., & Sloboda, J. A. (1997). Clinical trial of a music generated vibrotactile therapeutic environment for musicians: Main effects and outcome differences between therapy subgroups. *Journal of Music Therapy, 34,* 2–32.

Brotons, M. (1994). Effect of performing conditions on music performance, anxiety, and performance quality. *Journal of Music Therapy, 31*, 63–81.

Cox, W. J., & Kenardy, J. (1993). Performance anxiety, social phobia, and setting effects in instrumental music students. *Journal of Anxiety Disorders, 7*, 49–60.

Ericsson, K. A. (Ed.). (1996). *The Road to Excellence: The Acquisition of Expert Performance in the Arts and Sciences, Sports, and Games.* Mahwah, NJ: Erlbaum.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363–406.

Eysenck, H. J. (1981). *A Model for Personality.* Berlin: Springer-Verlag.

Fazey, J. A., & Hardy, L. (1988). *The Inverted-U Hypothesis: A Catastrophe for Sport Psychology* (British Association for Sports Sciences Monograph No. 1). Leeds: The National Coaching Foundation.

Fishbein, M., Middelstadt, S. E., Ottati, V., Strauss, S., & Ellis, A. (1988). Medical problems among ICSOM musicians: Overview of a national survey. *Medical Problems of Performing Artists, 3*, 1–8.

Fyk, J. (1995). *Melodic Intonation, Psychoacoustics, and the Violin.* Zielona Góra, Poland: Organon.

Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.), *The Psychology of Music* (2nd ed., pp. 501–602). London: Academic Press.

Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology of Music, 31*, 221–272.

Halpern, A. R., & Bower, G. H. (1982). Musical expertise and melodic structure in memory for musical notation. *American Journal of Psychology, 95*, 31–50.

Hamann, D. L., & Sobaje, M. (1983). Anxiety and college musicians: Study of performance conditions and subject variables. *Psychology of Music, 11*, 37–50.

Hunt, E. (1980). Intelligence as an information processing concept. *British Journal of Psychology, 71*, 449–474.

Lancaster, J. B. (1975). *Primitive Behavior and the Emergence of Human Culture.* New York: Holt, Rinehart, and Winston.

LeBlanc, A., Jin, Y. C., Obert, M., & Siivola, C. (1997). Effect of audience on music performance anxiety. *Journal of Research in Music Education, 45*, 480–506.

Leglar, M. A. (1978). *Measurement of Indicators of Anxiety Levels Under Varying Conditions of Musical Performance.* Unpublished doctoral dissertation, Indiana University, Bloomington, IN.

Nielsen, S. G. (1999). Learning strategies in instrumental music practice. *British Journal of Music Education, 16*, 275–291.

Rosenzweig, M. R., & Leiman, A. L. (1989). *Physiological Psychology* (2nd ed.). New York: Random House.

Salmon, P. G. (1990). A psychological perspective on music performance anxiety: A review of the literature. *Medical Problems of Performing Artists, 5*, 2–11.

Salthouse, T. A. (1991). Expertise as the circumvention of human processing limitations. In K. A. Ericsson & J. Smith (Eds.), *Toward a General Theory of Expertise: Prospects and Limits* (pp. 286–300). Cambridge: Cambridge University Press.

Schonberg, H. C. (1963). *The Great Pianists.* New York: Simon and Schuster.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory (Form Y1)*. Palo Alto, CA: Consulting Psychologists Press.

Steptoe, A. (1983). The relationship between tension and the quality of musical performance. *Journal of the International Society for the Study of Tension in Performance, 1*, 12–22.

Steptoe, A. (1989). Stress, coping and stage fright in professional musicians. *Psychology of Music, 17*, 3–11.

Steptoe, A. (2001). Negative emotions in music making: The problem of performance anxiety. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research* (pp. 291–307). Oxford: Oxford University Press.

Steptoe, A., & Fidler, H. (1987). Stage fright in musicians: A study of cognitive and behavioural strategies in performance anxiety. *British Journal of Psychology, 78*, 241–249.

Sternberg, R. J. (1977). *Intelligence, Information Processing, and Analytical Reasoning.* Mahwah, NJ: Erlbaum.

Sternberg, R. J. (1985). *Beyond IQ: A Triarchic Theory of Human Intelligence.* Cambridge: Cambridge University Press.

Valentine, E. R. (2002). The fear of performance. In J. Rink (Ed.), *Musical Performance: A Guide to Understanding* (pp. 168–182). Cambridge: Cambridge University Press.

Williamon, A., & Valentine, E. (2000). Quantity and quality of musical practice as predictors of performance quality. *British Journal of Psychology, 91*, 353–376.

Williamon, A., & Valentine, E. (2002). The role of retrieval structures in memorizing music. *Cognitive Psychology, 44*, 1–32.

Wilson, G. D. (2002). *Psychology for Performing Artists: Butterflies and Bouquets* (2nd ed.). London: Whurr.

Wilson, G. D., & Roland, D. (2002). Performance anxiety. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 47–61). Oxford: Oxford University Press.

Wolfe, M. L. (1989). Correlates of adaptive and maladaptive musical performance anxiety. *Medical Problems of Performing Artists, 4*, 49–56.

Wolfe, M. L. (1990). Coping with musical performance anxiety: Problem-focused and emotion-focused strategies. *Medical Problems of Performing Artists, 5*, 33–36.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative and Neurological Psychology, 18*, 459–482.

CHAPTER 2

# GENERAL PERSPECTIVES ON ACHIEVING MUSICAL EXCELLENCE

ROGER CHAFFIN AND ANTHONY F. LEMIEUX

Are exceptional musicians born or made? The everyday explanation for outstanding accomplishment—whether it be a virtuoso performance by a concert soloist, the triple axel jump of an ice skater, or a kayaker shooting a class 6 rapid—is to attribute it to an innate gift or talent. The idea of "talent" reflects the feeling of awe that such displays of virtuosity induce in us: the sense that these abilities lie outside the range of normal mortals. It has been one of the goals of psychology, since its earliest days, to understand the nature of outstanding talent. Galton framed the issues in his book *Hereditary Genius* (1869/1979), in which he attributed outstanding achievement to natural, inherited gifts, combined with a capacity for hard work and the ability to focus one's skills on a particular goal. In support of this analysis, he offered the genealogies of eminent men to show that eminence runs in families.

While there are clear limitations in his evidence—which overlooks the roles of inherited wealth, social position, and gender—Galton's views are still widely espoused today (Ericsson & Charness, 1994). Talent is regularly invoked whenever outstanding achievement needs to be explained, and often seems to be the only plausible explanation for the dramatic differences commonly observed in musical accomplishment. The idea that extraordinary abilities are an inborn gift of nature flatters the talented, while relieving the rest of us of the expectation that our own efforts should be measured on the same scale. A great deal of scientific effort has been directed at pinning down the genetically based traits and abilities of which talent is composed. Many likely candidates have been identified (see Coon & Carey, 1989). Success in predicting outstanding achievement from such components has, however, been relatively meager (Winner, 1996a; Simonton, 1999, 2001). Good evidence of a genetic basis for intellectual ability is still lacking. This may be the result of overly simplistic assumptions about how traits combine to produce extraordinary talent (Simonton, 1999), and more sophisticated approaches may eventually succeed in demonstrating the biological basis for talent so widely taken for granted. Meanwhile, the jury is still out.

More progress has been made in understanding the other two components of Galton's analysis: a capacity for hard work and the ability to concentrate on a particular goal. These provide the main focus of the discussion here. In every field that has been examined, those who attain eminence do so only after prolonged hard work over

a period of years. This is as true of music as of any other field and suggests that, whatever the role of heredity, the aspiring performer must be willing to work for success. This chapter, therefore, begins by examining the amount of work involved in achieving musical excellence. As in the joke about the visitor to New York who asked, "How do I get to Carnegie Hall?", the answer is "Practice, practice, practice."

Accumulating hours of practice, however, is not enough. High levels of accomplishment also require that practice time be well spent. Given the vast amounts of practice required to reach the highest levels in any field, even small improvements in effectiveness may yield very large differences in achievement. The second section of this chapter describes the evidence for this claim, setting the stage for the more specific, practical suggestions about how to achieve specific musical goals that are described in Chapters 5–8 of this book. The third section of the chapter identifies five general characteristics of effective practice—concentration, goal setting, self-evaluation, strategy selection, and "the big picture"—that constitute, in Galton's terms, the ability to focus and hone one's skills. These are the characteristics that allow the kinds of strategies described in Chapters 5–8 to be implemented effectively.

But where does the drive to excel come from? Is the temperament needed to devote oneself intensively to music born or made? The evidence for genetically based traits of temperament is stronger than for talent, but as with talent, there is clear evidence that environment also plays an important role. This chapter concludes, therefore, by considering the social psychological antecedents of the motive to succeed. It is intended that, by laying out the general precursors to musical excellence, performers and teachers will be better able to recognize and assess their own skills (as well as those of their students and colleagues) and make the most effective use of strategies for enhancing performance.

## 2.1 Quantity of practice: The 10-year rule

There is now a compelling body of evidence for the "10-year rule": a minimum of 10 years of dedicated work and practice are required to become an expert in any field. The 10-year minimum has been documented in every field of human endeavor that has been examined—from those that are largely physical in nature (like running), to those that are predominantly mental (like chess), to those requiring a combination of both (like musical performance). Indeed, this rule holds for musicians, novelists, poets, mathematicians, chess players, tennis players, swimmers, long distance runners, livestock judges, radiologists, and doctors, with the only possible exceptions being in the visual arts (Ericsson *et al.*, 1993; Winner, 1996b). And 10 years is a minimum, not the norm; for most, the path to eminence is much longer. And this is just the beginning. At least in music, the skills acquired during the long years of training must be continuously maintained and developed through practice (Krampe & Ericsson, 1996; Krampe, 1997).

While the idea that practice is integral to success is not likely to surprise anyone, the amount of training involved is striking. It is estimated that more than 10 000 hours of practice is required before a performer is ready to begin a professional career (Ericsson *et al.*, 1993). The young pianists in a study by Sosniak (1985) started their careers as concert soloists after an average of 17 years of training. For composers, the period of

preparation is even longer: 20 years from first exposure to music to first notable composition for the 76 major composers whose careers were reviewed by Hayes (1981). After a lifetime of practice, the experienced pianists in Krampe's (1997) study had put in 60 000 hours of practice. Such prolonged training has profound effects on physical and mental characteristics of the sort that have been generally thought of as determined by heredity, such as the proportion of fast and slow twitch muscle fibers and the area of cortex devoted to particular motor and sensory functions (Ericsson *et al.*, 1993; Altenmüller & Gruhn, 2002; Chapter 3 of this volume reviews the effects of prolonged training on health and offers suggestions for how they may be managed over the course of one's career).

To those familiar with the field of music, apparent counterexamples to the 10-year rule spring readily to mind. Music has provided its share of the geniuses and prodigies whose histories appear to make the case for inborn talent. Closer examination, however, suggests that these cases support rather than demolish the 10-year rule. Even with the best of intentions, early achievements tend to be exaggerated, and given the market value of child prodigies, deliberate misrepresentation is not uncommon. For these reasons, the early achievements of prodigies tend to be obscured by myth and distortion (Hayes, 1981; Howe, 1990, 1996; Ericsson & Charness, 1994; Sloboda, 1996; Howe *et al.*, 1998). Among composers, the most striking cases of early achievement (e.g. Alban Berg, Liszt, Mozart, and Shostakovich) all turn out to have put in the requisite 10 years before producing their first masterwork, shaving off just a year in the case of Berg and Shostakovich. These extraordinary cases "prove" the 10-year rule by demonstrating how well it marks the outer limit for even the most extreme examples of early accomplishment. In this company, the paragon of childhood genius, Mozart, was a late developer, not producing his first masterwork until he had been composing for 13 years, at the tender age of 17 (Schonberg, 1970, cited in Hayes, 1981, p. 211).

Close examination of that other class of apparent exceptions to the 10-year rule, musical savants, points to the same conclusion. Savants are people with a special ability that contrasts with their generally low level of functioning in other areas. Often their abilities appear to emerge very suddenly and without the opportunity for practice. The cases that have been studied, however, suggest that savants learn their skills in the same way as other people and that it is the recognition of their skills that occurs suddenly (Sloboda *et al.*, 1985; Ericsson & Faivre, 1988; Howe, 1996). Studies of musical savants and geniuses, thus, both point to the same conclusion: there are no clear exceptions to the 10-year rule in the field of music. Even the most exceptional talents put in 9 or 10 years of prolonged, dedicated work to develop the skills that allowed them to make their mark. For most people, a great deal more time and work is needed.

In summary, the development and maintenance of musical skills require a tremendous amount of practice. To reach the finals of international piano competitions, for instance, young pianists must practice regularly from childhood, increasingly dedicating their lives to music. Estimates of the amount of practice required for high achievement are remarkably uniform: about 2500 hours by age 13, 6500 by age 17, and approaching 10 000 by age 21 (Ericsson *et al.*, 1993; Sloboda *et al.*, 1996). Those who practice less generally achieve less. This is true for students, and it is true for professionals, even after a lifetime of performing. Once a person has put in the 10 000 or more hours

of practice needed to acquire the skills of a concert soloist, additional practice is needed to maintain those skills. The relationship between amount of practice and level of skill continues to hold, even among professional pianists (Krampe & Ericsson, 1996; Krampe, 1997).

This is not to say that achievement is solely determined by the amount of practice. It is almost certainly not the case that anyone who puts in the required number of hours would achieve the same high level of skill. More accomplished musicians may simply be more motivated to engage in musical activities and, as a result, practice more; in turn, their abilities may make practice itself more rewarding. Regardless, the relationship between practice and achievement suggests that practice is an important, indeed essential, part of the road to high achievement.

## 2.2 Quality of practice

Simply putting in hours of practice, however, is not enough. Practice time must be managed effectively. The same amount of practice can produce very different levels of achievement in different people. In a study by Sloboda *et al.* (1996), there were considerable differences in the amount of reported practice *within* each of four groups of music students, with the groups representing different levels of musical achievement. Among the "elite" group (i.e. students enrolled in a selective, specialist high school for music), there was "a small handful of outliers who [did] vastly greater amounts of practice than anyone else" (p. 301), and there was a handful of students in all groups who managed on very little practice, less than 20% of the group average. In preparing for Associated Board examinations (a system of standardized music exams), there were students in each group who passed with one fifth as much practice as the other students, and there were others who did four times more practice. Similarly, there were large differences in the amount of practice reported by the pianists in Sosniak's (1985) study. In their early years, some spent "every free minute" at the piano, while others practiced as little as possible—although with practice enforced by parents, this was still a substantial amount (p. 34). These differences narrowed as the pianists became more serious about their practice in their middle years, but still ranged from 2 to 4 hours per day. Williamon and Valentine (2000, 2002) also found a wide range in the amount of practice student pianists required to learn a new piece, with some students taking two or three times as much time as others. Moreover, these differences in quantity of practice were not related to the quality of the final performance; rather, quality of performance was predicted by how practice was organized. As learning progressed, practice was increasingly organized in terms of the structure of the music (as identified by each pianist), and those who started doing this earlier gave the best performances at the end of the learning process. Musical achievement in this case was predicted by the ability to discern the musical shape of the piece and thus to form an "artistic image" of how it should be performed (Neuhaus, 1973), rather than by the amount of practice.

Similar differences in the effectiveness of practice also occur in beginning students. O'Neill (1997) found that, during their first year of music lessons, children who exhibited a "helpless" orientation to practice put in twice as much time to reach the same level of achievement as those exhibiting a "mastery" orientation. McPherson and

Zimmerman (2002) found that achievement during the first 9 months of music lessons was determined less by the amount of practice than by commitment to the instrument. Regardless of how much practice they did, students who felt they would be playing their instrument throughout their schooling made more progress than those who felt they would play for only a few years.

What is the source of these enormous differences in the amount of time needed to reach similar levels of accomplishment? One possibility is that the differences are due to raw, native talent. It may be that those who practice less are simply more talented and need to put in less work to achieve the same end. As noted above, however, psychologists have so far been unsuccessful in providing the evidence needed to substantiate this explanation. The contribution to achievement made by native talent has yet to be determined. Native talent is, in any case, not something that anyone can do anything about. Meanwhile, it is clear that enormous amounts of practice are needed, even for the most talented, and it seems very likely that at least some of the large differences in musical achievement that can be observed at every level of training are due to differences in how effectively practice time is spent.

Preparing a musical performance is a complex task and the necessary skills develop over many years. Musicians learn to practice more effectively as their skills develop. The pianist Misha Dichter noted, "I hate to think of the time I wasted as a student. Now it's just so easy to see certain shortcuts that would have saved thousands of hours" (Noyle, 1987, p. 57). Psychological studies of practice in fields other than music indicate that Dichter was probably right. Effective practice depends on finding strategies that work (Seashore, 1939; Chase & Ericsson, 1982; Ericsson & Faivre, 1988; Ericsson *et al.*, 1993). Without an effective strategy, practice does not lead to improvement (Chase & Ericsson, 1981). Within the field of music, there is ample evidence that use of more effective practice strategies results in faster and better learning. Details of this evidence with regard to individual and ensemble practice and in terms of achieving specific musical goals—such as memorizing, sight-reading, and improvising—are provided in Chapters 5–8 of this volume.

The characteristics of effective practice have been the subject of investigation in a variety of domains for over a century (e.g. Bryan & Harter, 1899; see reviews by Ericsson & Lehmann, 1996; Ericsson *et al.*, 1993; Ericsson, 1996, 1997). Effective practice is not simply a matter of going through the motions. The repeated exercise of a skill, even for professional purposes, does not necessarily lead to improvement. In most fields, skill development stops at a stable plateau when performance reaches the level required to get the job done, and further improvement occurs only when there is some new incentive, like a pay raise or opportunity for promotion, to motivate the hard work required (Bryan & Harter, 1899). Ericsson and colleagues have characterized the work needed to produce improvement as "deliberate" practice. Whether the skill involved is cognitive (such as solving algebra problems), physical (such as typing), or perceptual (such as wine-tasting), improvement requires setting goals that are attainable from the current skill level and which lead to the development of effective strategies (Chase & Ericsson, 1981). Progress must be constantly monitored and new routes to improvement continually explored. For those with the necessary skills, self-evaluation and self-directed exploration may be sufficient, but for students and for the less skilled,

the help of a teacher or coach is essential to provide the necessary feedback and to suggest productive strategies for overcoming problems. Although it is the everyday grind of practice that often seems most salient, effective practice is anything but routine. Rather, it is a matter of continuous, creative problem solving, self-evaluation, and striving.

## 2.3 Fundamental characteristics of musical excellence

In this section, five fundamental characteristics of effective practice are outlined. They are based on the literature on expertise and skill learning reviewed in the previous section and are general characteristics of effective practice, not limited to music. Evidence that these characteristics of effective practice also apply to the field of music can be found in the writings of noted piano pedagogues (e.g. Gieseking & Leimer, 1932/1972; Neuhaus, 1973; Sandor, 1981), in published interviews with eminent pianists (see Chaffin *et al.*, 2002, for a summary), and in interview studies that have investigated how accomplished musicians practice and perform (Wicinski, 1950, cited in Miklaszewski, 1989; Hallam, 1995a, 1995b, 1997a, 1997b; Aiello, 2001). In order to provide performers and teachers with useful examples, the discussion here has been restricted to characteristics that can be drawn from the limited number of empirical studies of the practice of expert musicians.

### 2.3.1 Concentration

The ability to concentrate fully on the task at hand is probably the most important characteristic of effective practice for musicians (Auer, 1921; Hallam, 1998). Heinrich Neuhaus, the eminent Russian pianist and teacher notes "The greater the . . . concentration, the better the result" (1973, p. 4). The eminent pianists whose comments about practice were collected by Chaffin *et al.* (2002) appear to agree. Misha Dichter warns,

> In practicing, never daydream. Never use the piano as a vehicle for simply moving the fingers and passing time. If you have only one moment when you're not aware of what you're doing musically or technically (and usually both), you're wasting your time . . . (Noyle, 1987, p. 59).

Once a piece is learned, it is easy to fall into mindless practice because the piece can be played automatically, without attention. As David Bar-Illan notes, "One can all too easily play music without actually listening to it" (Dubal, 1997, pp. 40–41). Playing without full attention is dangerous, Leon Fleischer warns, because "under the stress of public performance [motor memory is] the first thing that goes, if you are nervous" (Noyle, 1987, p. 95). Emile Sauer sums up the situation as follows:

> One hour of concentrated practice with the mind fresh and the body rested is better than four hours of dissipated practice with the mind stale and the body tired. With a fatigued intellect the fingers simply dawdle over the keys and nothing is accomplished (Cooke, 1913/1999, p. 238).

What does this kind of concentration look like? To answer this question, Chaffin *et al.* (2002) observed the practice of a concert pianist as she learned the Presto of J. S. Bach's Italian Concerto. The practice sessions, which were videotaped, give a predominant impression of continuous, urgent activity as the pianist played without

pause, stopping and starting continuously at top speed. When asked about the frantic pace of her work the pianist reported, "When I start on a new piece, I have such an appetite to take hold of it and make it mine" (p. 255).

In an attempt to quantify this "appetite to take hold," Chaffin *et al.* compared the *practice rate* (the number of beats actually played per minute) with the *mean target tempo* (number of beats per minute dictated by the tempo at which the pianist was trying to play). The expectation was that the two values would be nearly identical, showing that the pianist played almost continuously throughout each practice session. The reality proved very different, and much more interesting. The practice rate was less than a fourth of the target tempo; the pianist played only a quarter the number of notes expected, based on the tempo at which she was nominally playing. This might suggest that the pianist had spent three quarters of her practice time daydreaming, except that the videotapes showed otherwise. Rather, the discrepancy between the two measures reflected time spent in micro-pauses and momentary decreases in tempo, each no more than a few seconds, which occurred continually throughout every practice session. During these pauses and hesitations the pianist was thinking: evaluating what she had just done, planning what to do next, mentally previewing the upcoming passage, allowing time for what Matthay (1926, p. 5) calls "pre-listening."

The pianist was indeed engaged in nonstop practice; the impression from the videotape of continuous, unrelenting effort was not mistaken. But nonstop practice does not mean nonstop playing. Effective practice must be guided by thought; and thinking takes time. The measurements showed that only a quarter of the time was spent actually playing (moving the fingers over the keys) and that the remaining three quarters of the time was spent in thinking. These measurements suggest that a central characteristic of effective practice is the mental effort and concentration that is involved and that, at least for accomplished musicians, the level of effort may be reflected in the ratio of rate of practice to target tempo (pp. 130–135).

The need to maintain full attention is the reason that practice is most effective if done in short sessions of an hour or less, separated by breaks for recuperation (Auer, 1921; Seashore, 1938/1967; Rubin-Rabson, 1940a). As the pianist Rudolf Firkusny noted, "Concentration is very tiring . . . Sometimes one or two hours of concentrated practicing is much more tiring than playing seven or eight hours" (Noyle, 1987, p. 81). Bella Davidovich reports,

> I find that a one-hour period is where I achieve the utmost in terms of concentration. I work very intensively for one hour and then take a ten, fifteen, or twenty minute break during which I will occupy myself with something completely different, whether it's to eat or something else. This method works out well so that I can continue for eight hours, in one-hour periods (Noyle, 1987, p. 43).

The same is true for the practice of all kinds of skills, not just music (see Ericsson *et al.*, 1993, for a review) and the ability to practice effectively for even this much time develops only with experience and training (Sosniak, 1985). For maximum effectiveness, the length of practice sessions must be tailored both to the task at hand and to the energy required. In Chaffin *et al.*'s (2002) account of the concert pianist learning the Presto of the Italian Concerto, the length of practice sessions decreased steadily from

an hour at the beginning to half an hour towards the end of the learning process nearly a year later. Throughout, the pianist monitored her energy level, making comments like, "I definitely feel like I am running out of steam" (p. 163) and stopped when she became too tired, even when this conflicted with her goals for the practice session (e.g. "I'll stop now and take a break and come back to it later," p. 163). Ericsson *et al.* (1993) noted that, in a study of student violinists, the most accomplished student performers took more naps than less accomplished performers and students preparing for careers as music teachers. The implication is that musicians should take care to be well rested when they practice in order to practice effectively.

The need to practice attentively is also responsible for the common preference among experienced musicians to practice in the morning and to use the morning hours for their most demanding tasks. For many people, mental capacities are at their peak during the morning (May *et al.*, 1993; Hasher *et al.*, 2002; Yoon *et al.*, 2000) and this appears to be the case for many of the pianists surveyed by Chaffin *et al.* (2002). Bella Davidovich reports that, "The best hours for practice are in the morning when a person's mind is fresh and the ears are fresh" (Noyle, 1987, p. 43). Emil Sauer said, "I find in my own daily practice that it is best for me to practice two hours in the morning and then two hours later in the day" (Cooke, 1913/1999, p. 238). Janina Fialkowska confirms May *et al.*'s (1993) finding that morning practice becomes more important later in life:

> For me, the best time to learn something new is in the morning. I, absolutely, at my advanced age cannot learn a new piece, I cannot memorize, in the afternoon after lunch. I don't know why, I cannot . . . (Noyle, 1987, pp. 67–68).

More systematic evidence comes from Sloboda *et al.*'s (1996) study of student musicians, in which the most accomplished students reported doing a higher proportion of their scales in the morning while less accomplished students practiced scales more in the evening. The more accomplished students were apparently devoting their most effective practice time to the type of practice that, for most people, is least rewarding and therefore harder to concentrate on (see Chapters 11 and 12 of this volume for suggestions of specific techniques for enhancing concentration).

### 2.3.2 Goal setting

A second general characteristic of effective practice is that it involves setting and meeting specific goals. Neuhaus (1973) describes an incident in which Sviatislav Richter reported that he had repeated one difficult passage of 10 bars over 100 times and recommends,

> You have to put the kettle on the stove and not take it off until it boils. . . . Mastering the art of working, of learning compositions is characterized by an unwavering determination and an ability not to waste time. The greater the part . . . played by willpower (going straight for the goal) . . . the better the result. . . . (pp. 3–4).

This is why students are generally urged to work on a piece initially in small segments (Barry & Hallam, 2002). Limiting the number of problems to be dealt with makes it

possible to focus attention on a small number of problems and solve them, mastering the passage at once instead of returning to it time and time again. By not playing through passages before they have been mastered, the musician also avoids developing bad motor habits that will later have to be laboriously unlearned.

Expert musicians appear to follow Neuhaus's advice, organizing practice into periods of *work* in which one passage of a few bars is repeated many times, separated by *runs* in which progress is evaluated and new problems are identified for work (Gruson, 1988; Miklaszewski, 1989; Chaffin *et al.*, 2002; Williamon *et al.*, 2002). Chaffin *et al.* (2002) describe an example of the effectiveness of the strategy in which the pianist they were studying worked through one eight-bar section for the first time, repeating different short segments for a total of more than 150 repetitions. In a second practice session, the same passage received another 50 repetitions and, from then on, needed no further work; it had been mastered in two focused encounters. This pattern of work interspersed by runs continued throughout the 10 months of preparation. Even when the pianist was able to perform fluently and was polishing the piece for performance, intensive work on short passages continued to account for about 5% of the music played in each session. As playing became more fluent, runs increased in length, but work did not. Instead, work segments became slightly shorter as the pianist was able to focus more narrowly on individual problems. This pattern of activity is indicative of practice that is goal-directed with problems being identified and eliminated in every session, as Neuhaus recommends (strategies for setting goals and action planning are discussed in Chapters 5 and 12 of this volume).

### 2.3.3 Self-evaluation

To know when a goal has been accomplished the performer needs feedback about success and failure. Ultimately, this must come from within. While teachers play an essential role in setting standards and developing discrimination in their pupils, they cannot be present during every practice session and cannot provide the continuous, moment-to-moment feedback that determines the microstructure of practice with its constant starts, stops, and repetitions. It has already been noted that the opportunity for feedback is a necessary condition for the improvement of any kind of skill (Ericsson *et al.*, 1993). Here, the suggestion is more specific—that effective practice in the field of music requires the feedback provided by self-evaluation and that without it, progress is slower. For example, McPherson and McCormick (1999) found that students who practiced more also made more evaluative judgments about the success and failure of their efforts. Similarly, the majority of the comments made by the pianist observed by Chaffin *et al.* (2002) during practice were evaluative in nature. General evaluative remarks were the most frequent topic, accounting for nearly 40% of all remarks, and most of the comments classified under other headings also contained an evaluative component, suggesting that self-evaluation was part of every decision and action. It may also be important that most of these evaluative remarks, whether positive or negative, were delivered with an air of dispassionate detachment—for example, "That needs work too. It's coming along" (p. 159). Chapter 4 of this volume offers specific suggestions for how self-evaluation can be used to enhance performance.

### 2.3.4 Strategies

The ability to meet goals depends on being able to come up with an effective practice strategy to meet the needs of the moment. Effective practice depends on a wide range of strategies that can be flexibly deployed. Hallam (1995a, 1995b) interviewed 22 professional musicians selected for their reputation for technical and musical excellence, asking about their practice habits and how they would go about learning a new piece. The musicians reported a wide range of strategies that they used flexibly to address every aspect of their task: learning new repertory, maintaining skills, preparing for performance, and managing the physical and emotional demands of their challenging careers. Direct observations of the practice of experienced musicians confirms these reports. Chaffin *et al.* (2002) describe strategies dealing with everything from choice of fingering, what bars to use as starting places, and the spacing of practice sessions, to the management of frustration, memorization techniques, and deployment of attention during performance (see also Miklaszewski, 1989, 1995; Hallam, 1994; Chaffin & Imreh, 1997, 2001, 2002; Lehmann & Ericsson, 1998; Nielsen, 1999, 2001; Ginsborg, 2002; and Chapters 5–8 of this book for recommended strategies for individual and ensemble rehearsal).

### 2.3.5 The big picture

A fifth characteristic of effective practice is the ability to keep in mind the larger musical picture, the "artistic image" of a piece (Neuhaus, 1973), at the same time as attending to details of technique and interpretation. The performer needs to have an idea of the overall expressive shape of the piece while making initial decisions about technique. If these decisions are made without the big picture in mind, many of them may have to be changed later, which greatly increases the learning time. The ability to switch attention from detail to big picture and back increases with experience (Chaffin *et al.*, 2002, p. 90; Williamon *et al.*, 2002). Most of the experienced performers interviewed by Hallam (1995a, 1995b) and the eminent Russian pianists interviewed by Wicinski (1950, cited in Miklaszewski, 1989) reported beginning work on a new piece by first looking at the big picture and thinking about how it should be performed. Similarly, the pianist studied by Chaffin *et al.* (2002) began by identifying the overall shape of a new piece and locating the main difficulties for performance, anticipating from the outset decisions that would not be made until later in the learning process (see Chaffin *et al.*, 2003, for details). The ability to anticipate in this way is characteristic of expert problem solving in many fields. For example, chess experts are able to play "lightning chess," intuitively making snap decisions that anticipate later developments in the game (Chase, 1983; Glaser & Chi, 1988; Gobet & Simon, 1996). Similarly, when accomplished performers acquaint themselves with a new work there is an "instantaneous and subconscious process of 'work at the artistic image' " (Neuhaus, 1973, p. 17).

One effect of paying attention to the larger musical picture is the use of the formal structure of a piece to organize practice, something recommended by many pedagogues (e.g. Matthay, 1926; Sandor, 1981; Shockley, 1986). Experienced musicians spontaneously do this from the outset, starting practice segments at the beginnings of sections when first learning a new piece (Miklaszewski, 1989, 1995; Nielsen, 1999, 2001; Chaffin & Imreh, 2002; Chaffin *et al.*, 2002). Later in the learning process, the use

of section boundaries as starting places is abandoned, at least for a time, in order to ensure fluent transitions between sections (see Chaffin *et al.*, 2002). For less experienced musicians, in contrast, use of the structure to organize practice begins later in the learning process and increases across sessions. In Williamon and Valentine's (2002) study, the students did not initially have a good grasp of the structure and only began using it to organize practice later in the learning process as they began to appreciate its musical significance. Earlier use of the structure was associated with a more musical final performance. Students who were quicker to grasp the musical shape of the piece gave better performances.

### 2.3.6 Caveat

The five characteristics identified above have been noted as important by distinguished teachers, reported of their own practice by eminent musicians, and identified in empirical observations of the practice of expert or experienced musicians. The list is necessarily tentative since it remains to be shown, for any of these characteristics, that they are more common among professionals than among students. For this, further studies comparing the practice of musicians at different levels of training are needed. Eventually, experimental studies will be needed to determine which characteristics can be usefully induced in student practice by instruction (Hallam, 1997a, 1997b), and for this, the pioneering studies of Rubin-Rabson (1937, 1939, 1940a, 1940b, 1941a, 1941b, 1941c) still provide the model.

## 2.4 The social psychological antecedents of musical excellence

Concentration, setting clear goals, evaluating progress, using strategies flexibly, and looking for the big picture are complex skills. Developing them requires unusual motivation. Where does that motivation come from? It is likely that motivation is one of several necessary components of musical talent, and it seems to be the one most strongly constrained by biology (Ericsson *et al.*, 1993). Heredity plays a large role in determining other aspects of temperament (Kagan *et al.*, 1992, 1994), and it may be that the same is true of the motivation to master musical skills. As with the case for abilities, it is still uncertain as to the extent to which genetics contribute to the motivation to play music. Meanwhile, there is ample evidence of the powerful contribution of the social environment to both the motivation to succeed and the development of musical excellence.

Social factors involved in the development of musical excellence are summarized in Figure 2.1, which provides a conceptual model integrating the motivational and social cognitive principles discussed in the remainder of the chapter. Those who develop the deep commitment needed to master a musical instrument (1) play for their own satisfaction, rather than to please someone else or for other external reasons, (2) attribute their achievements (and failures) to their own efforts, and (3) feel that they have the capacity to improve.

### 2.4.1 Sources of motivation

Motivation determines what people like and what they do. Motivation to engage in an activity such as instrumental practice can come either from inside a person (intrinsic,

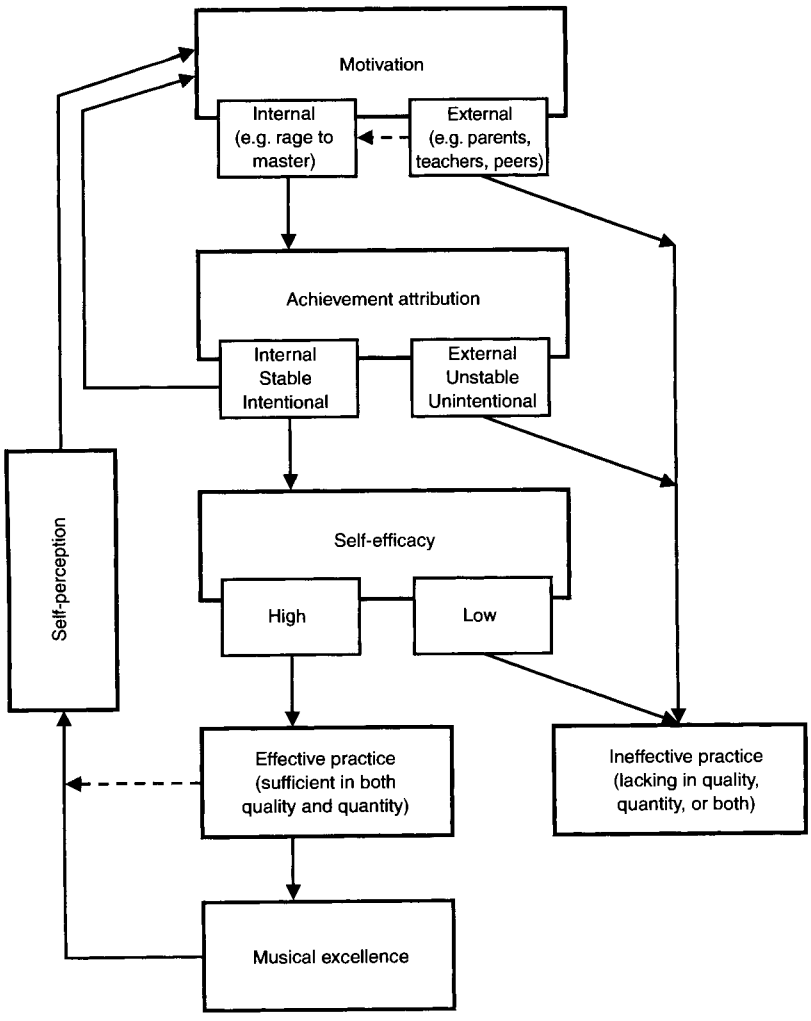



Figure 2.1 A conceptual model of the social antecedents of musical excellence.

e.g. am I doing this because I want to?) or from someone or something else (extrinsic, e.g. am I doing this because someone else—parents, teachers, or peers—wants me to?), or some combination of the two (Deci & Ryan, 1985; Ryan & Deci, 2000; see also Hallam, 1997a). Intrinsic motivation appears to be essential for the development of effective practice strategies (see Maehr *et al.*, 2002 and O'Neill & McPherson, 2002, for reviews). Studies of parental involvement in the practice of young instrumental students show that children's initial motivation to practice comes largely from their parents (Pitts *et al.*, 2000a, 2000b; McPherson & Zimmerman, 2002). Internalization of the motivation to practice represents the development of intrinsic motivation as students begin to undertake practice of their own volition. As intrinsic motivation develops,

students need fewer parental reminders and are more likely to engage in goal-directed and self-regulated practice. Parents who continue to supervise or compel practice can ultimately hinder the development of the sustainable, intrinsic motivation that is essential if students are to engage in the kinds of effective practice described in the previous section (O'Neill, 1997; Pitts *et al.*, 2000a, 2000b; McPherson & Renwick, 2001; McPherson & Zimmerman, 2002).

The motivation to undertake the intense effort required to practice effectively has been identified in gifted children as the "rage to master." Winner (1996a, 1996b) describes it as a powerful motive to excel in a given behavioral or intellectual arena that emerges when children work on appropriately challenging, self-assigned tasks that they are highly motivated to accomplish. The rage to master is an unusually intense form of intrinsic motivation, and Winner's descriptions of its development provide good examples of the transition from external to internal motivation. Children need to have at least some modicum of success at a task before they develop their own internal motivation to perform it. Some initial positive experience is needed for the child to decide that he or she is "good" at a task. Children come to recognize their own talents, as well as their likes and dislikes, from observation of their own successes and failures (e.g. Bem, 1972).

Although Winner introduced the construct of rage to master to describe the motivation of gifted children, it can be usefully extended to adults. In fact, the term provides an excellent description of the kind of concentrated, goal-oriented absorption identified in the previous section as a hallmark of effective practice (Chaffin *et al.*, 2002). In adults, this state of intense creative engagement with a task has been characterized as "flow" (Csikszentmihalyi & Csikszentmihalyi, 1988; Csikszentmihalyi, 1990; Csikszentmihalyi *et al.*, 1993; O'Neill & McPherson, 2002). The trance-like state often described by musicians as occurring during their best performances is another manifestation of flow. The groundwork for these transcendent experiences during performance is laid during practice when many hours must be spent in a similar state of absorption in the task of mastering the music. When the same intense, focused attention is finally transferred from the details of practice to the expressive effects of a fluent performance, the result is the very special, transcendent experience so often reported by performers (Chaffin & Imreh, 2002).

### 2.4.2 Attribution

People constantly try to understand themselves and other people by identifying the causes of their behavior, a process that social psychologists refer to as "attribution" (Heider, 1958; Jones & Davis, 1965; Kelley, 1971; Dweck & Goetz, 1978). The attributions people make in explaining their own achievements have important effects on their subsequent motivation (Weiner *et al.*, 1972, 1976; Weiner, 1974, 1985, 1986).

Weiner *et al.* (1972) identified three elements in the attribution process that affect expectations for future performance and, thus, motivation. The first element is whether the causal factor has an internal or external locus (e.g. is my success due to my own efforts or abilities?). The second element is intentionality (e.g. did I try?). The third element is whether the causal factor is stable or unstable over time and across situations (e.g. can I expect to succeed again next time?). To the extent that a student

feels herself to be in control of her successes or failures, engages intentionally in practice activities, and perceives the causal factors associated with her successes and failures as having stability across time and situation, she assumes a greater level of responsibility for her performance. She also assumes responsibility for her own improvement and for problem solving. On the other hand, if a student feels that her success or failure is due to someone or something else (external attributions), feels that her efforts were not important to the overall outcome (lack of intentionality), and that the results could just as well be otherwise next time (lack of stability over time or situation), she is unlikely to assume personal responsibility for addressing and remedying problems with her practice and their effects on her performances.

### 2.4.3 Self-efficacy

Thus far, the importance of motivation and the role of attribution have been addressed. However, a person's intense desire to excel and sense of personal responsibility must also be accompanied by feelings of competence, capability, and personal agency (e.g. Bandura, 1977, 1982, 1986, 1997). Studies by McPherson and McCormick (1999) and Pitts *et al.* (2000a) provide evidence of the importance of feelings of self-efficacy in music students. So, while students must have both the motivation to engage in effective practice and a sense of personal agency, these factors alone may not be sufficient to sustain effective practice. It is also necessary that students feel that they either have or are capable of developing the requisite skills.

As the student makes more internal attributions for his successes and failures, self-efficacy becomes increasingly important. For example, if a student makes internal attributions and feels a sense of personal responsibility for his achievement (or lack thereof), but does not have a sense of self-efficacy, he could become discouraged and ultimately stop practicing. However, if internal attributions for successes and failures are accompanied by a sense of self-efficacy, the student not only feels personally responsible for his progress, but also feels confident that he will indeed have the skills and ability necessary for further improvement.

## 2.5 Conclusions

Some of the variability in musical achievement can be attributed to the amount of time spent in practice. Tens of thousands of hours of practice are needed to reach and sustain a professional level of playing in the Western classical music tradition. The idea that very gifted individuals can achieve the highest levels of accomplishment in this field with minimal effort is a myth. In music, as in most fields of human endeavor, even the most exceptional prodigies put in at least 10 years of dedicated effort and practice in order to develop their skills to the point that they are capable of attaining eminence.

Another important factor determining the level of musical accomplishment is the effectiveness of those thousands of hours of practice time. Because preparation time is so long, and because so many hours must be spent in practice, even small differences in efficiency may accumulate over years, resulting in very large differences in accomplishment.

To be effective, practice must be guided by efficient strategies and continuous monitoring of their effectiveness. It seems indisputable that some of the differences between individuals in musical accomplishment are due to the effectiveness of their practice. Five characteristics of effective practice are:

- concentration
- setting and meeting specific goals
- constant self-evaluation
- flexible use of strategies
- seeing the big picture

This list of characteristics is offered in the hope that it will stimulate research on the characteristics of effective music practice. In addition, identifying the characteristics of effective practice is worthwhile because it can help to understand the basic psychological processes involved in skill acquisition and because of its pedagogical value. Knowing that practicing in a particular way systematically produces better results should be helpful to both students and teachers in their efforts toward excellence. Concentration, goal setting, self-evaluation, use of strategies, and attention to the big picture can all be seen as manifestations of an intense desire to produce music, of the "rage to master." Such motivation must be nurtured; it cannot be produced strictly by instruction. There are, however, important situational and social factors involved in the development of the motivation toward mastery, and some of the most important have been enumerated here. There is still much to learn about the long and complex process of developing musical skill. The process is, however, well worth understanding, since it is both a universal of human culture and a microcosm of cognitive development.

## Acknowledgments

We would like to thank Mary Crawford for helpful comments on an earlier version of this chapter.

## Further information and reading

Chaffin, R., Imreh, G., & Crawford, M. (2002). *Practicing Perfection: Memory and Piano Performance.* Mahwah, NJ: Erlbaum.

Ericsson, K. A. (Ed.). (1996). *The Road to Excellence: The Acquisition of Expert Performance in the Arts and Sciences, Sports, and Games.* Mahwah, NJ: Erlbaum.

Howe, M. J. A. (1990). *The Origins of Exceptional Abilities.* Oxford: Blackwell.

## References

Aiello, R. (2001). Playing the piano by heart. *Annals of the New York Academy of Sciences, 930,* 389–393.

Altenmüller, E., & Gruhn, W. (2002). Brain mechanisms. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 63–81). Oxford: Oxford University Press.

Auer, L. (1921). *Violin Playing as I Teach It.* New York: Stokes.

Bandura, A. (1977). Self-efficacy: Towards a unifying theory of behavioral change. *Psychological Review, 84,* 191–295.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37,* 122–147.

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control.* New York: Freeman.

Barry, N. H., & Hallam, S. (2002). Practice. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 151–165). Oxford: Oxford University Press.

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 6, pp. 1–62). London: Academic Press.

Bryan, W. L., & Harter, N. (1899). Studies on the telegraphic language: The acquisition of a hierarchy of habits. *Psychological Review, 6,* 345–375.

Chaffin, R., & Imreh, G. (1997). "Pulling teeth and torture": musical memory and problem solving. *Thinking and Reasoning, 3,* 315–336.

Chaffin, R., & Imreh, G. (2001). A comparison of practice and self-report as sources of information about the goals of expert practice. *Psychology of Music, 29,* 39–69.

Chaffin, R., & Imreh, G. (2002). Practicing perfection: Piano performance as expert memory. *Psychological Science, 13,* 342–349.

Chaffin, R., Imreh, G., & Crawford, M. (2002). *Practicing Perfection: Memory and Piano Performance.* Mahwah, NJ: Erlbaum.

Chaffin, R., Imreh, G., Lemieux, A., & Chen, C. (2003). "Seeing the big picture": Piano practice as expert problem solving. *Music Perception, 20,* 465–490.

Chase, W. G. (1983). Spatial representation in taxi drivers. In D. R. Rogers & J. A. Sloboda (Eds.), *Acquisition of Symbolic Skills* (pp. 391–405). New York: Plenum.

Chase, W. G., & Ericsson, K. A. (1981). Skilled memory. In J. R. Anderson (Ed.), *Cognitive Skills and Their Acquisition* (pp. 141–189). Hillsdale, NJ: Erlbaum.

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 16, pp. 1–58). London: Academic Press.

Cooke, J. F. (1913/1999). *Great Pianists on Piano Playing: Godowsky, Hofmann, Lhévinne, Paderewski and 24 Other Legendary Performers.* New York: Dover. (Original work published in 1913; expanded edition published in 1917).

Coon, H., & Carey, G. (1989). Genetic and environmental determinants of musical ability in twins. *Behavior Genetics, 19,* 183–193.

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience.* New York: Harper and Row.

Csikszentmihalyi, M., & Csikszentmihalyi, I. S. (Eds.). (1988). *Optimal Experience: Psychological Studies of Flow in Consciousness.* Cambridge: Cambridge University Press.

Csikszentmihalyi, M., Rathunde, K. R., Whalen, S., & Wong, M. (1993). *Talented Teenagers: The Roots of Success and Failure.* Cambridge: Cambridge University Press.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior.* New York: Plenum.

Dubal, D. (1997). *Reflections from the Keyboard: The World of the Concert Pianist* (2nd ed.). New York: Schirmer.

Dweck, C. S., & Goetz, T. E. (1978). Attributions and learned helplessness. In J. H. Harvey & W. Ickes & R. F. Kidd (Eds.), *New Directions in Attribution Research* (Vol. 2). Hillsdale, NJ: Erlbaum.

Ericsson, K. A. (Ed.). (1996). *The Road to Excellence: The Acquisition of Expert Performance in the Arts and Sciences, Sports, and Games.* Mahwah, NJ: Erlbaum.

Ericsson, K. A. (1997). Deliberate practice and the acquisition of expert performance: An overview. In H. Jørgensen & A. C. Lehmann (Eds.), *Does Practice Make Perfect? Current Theory and Research on Instrumental Music Practice* (pp. 9–51). Oslo: Norwegian Academy of Music.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49,* 725–747.

Ericsson, K. A., & Faivre, I. A. (1988). What's exceptional about exceptional abilities? In L. K. Obler & D. Fein (Eds.), *The Exceptional Brain: Neuropsychology of Talent and Special Abilities* (pp. 436–473). London: Guildford.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363–406.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology, 47,* 273–305.

Galton, F. (1869/1979). *Hereditary Genius.* London: Julian Friedman. (Original work published in 1869).

Gieseking, W., & Leimer, K. (1932/1972). *Piano Technique.* New York: Dover. (Original work published in 1932).

Ginsborg, J. (2002). Classical singers learning and memorising a new song: An observational study. *Psychology of Music, 30,* 58–101.

Glaser, R., & Chi, M. T. H. (1988). Overview. In M. T. H. Chi, R. Glaser & M. J. Farr (Eds.), *The Nature of Expertise* (pp. xv–xxviii). Hillsdale, NJ: Erlbaum.

Gobet, F., & Simon, H. A. (1996). The roles of recognition processes and look ahead search in time-constrained expert problem solving: Evidence from grand-master-level chess. *Psychological Science, 7,* 52–55.

Gruson, L. M. (1988). Rehearsal skill and musical competence: Does practice make perfect? In J. A. Sloboda (Ed.), *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition* (pp. 91–112). Oxford: Clarendon Press.

Hallam, S. (1994). Novice musicians' approaches to practice and performance: Learning new music. *Newsletter of the European Society for the Cognitive Sciences of Music, 6,* 2–10.

Hallam, S. (1995a). Professional musicians' approaches to the learning and interpretation of music. *Psychology of Music, 23,* 111–128.

Hallam, S. (1995b). Professional musicians' orientations to practice: Implications for teaching. *British Journal of Educational Psychology, 12,* 3–19.

Hallam, S. (1997a). Approaches to instrumental music practice of experts and novices: Implications for education. In H. Jørgensen & A. C. Lehmann (Eds.), *Does Practice Make*

*Perfect? Current Theory and Research on Instrumental Music Practice* (pp. 89–107). Oslo: Norwegian Academy of Music.

Hallam, S. (1997b). What do we know about music practising? Towards a model synthesising the research literature. In H. Jørgensen & A. C. Lehmann (Eds.), *Does Practice Make Perfect? Current Theory and Research on Instrumental Music Practice* (pp. 179–231). Oslo: Norwegian Academy of Music.

Hallam, S. (1998). The predictors of achievement and dropout in instrumental tuition. *Psychology of Music, 26*, 116–132.

Hasher, L., Chung, C., May, C. P., & Foong, N. (2002). Age, time of testing, and proactive interference. *Canadian Journal of Experimental Psychology, 56*, 200–207.

Hayes, J. R. (1981). *The Complete Problem Solver.* Philadelphia: Franklin Institute Press.

Heider, F. (1958). *The Psychology of Interpersonal Relations.* New York: Wiley.

Howe, M. J. A. (1990). *The Origins of Exceptional Abilities.* Oxford: Blackwell.

Howe, M. J. A. (1996). The childhoods and early lives of geniuses: Combining psychological and biographical evidence. In K. A. Ericsson (Ed.), *The Road to Excellence: The Acquisition of Expert Performance in the Arts and Sciences, Sports, and Games* (pp. 255–270). Mahwah, NJ: Erlbaum.

Howe, M. J. A., Davidson, J. W., & Sloboda, J. A. (1998). Innate talents: Reality or myth? *Behavioral and Brain Sciences, 21*, 399–442.

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 16, pp. 219–266). London: Academic Press.

Kagan, J., Arcus, D., Snidman, N., Feng, W. Y., Hendler, J., & Greene, S. (1994). Reactivity in infants: A cross-national comparison. *Developmental Psychology, 30*, 342–345.

Kagan, J., Snidman, N., & Arcus, D. M. (1992). Initial reactions to unfamiliarity. *Current Directions in Psychological Science, 1*, 171–174.

Kelley, H. H. (1971). *Attribution in Social Interaction.* Morristown, NJ: General Learning Press.

Krampe, R. T. (1997). Age-related changes in practice activities and their relation to musical performance skills. In H. Jørgensen & A. C. Lehmann (Eds.), *Does Practice Make Perfect? Current Theory and Research on Instrumental Music Practice* (pp. 165–178). Oslo: Norwegian Academy of Music.

Krampe, R. T., & Ericsson, K. A. (1996). Maintaining excellence: Deliberate practice and elite performance in young and older pianists. *Journal of Experimental Psychology: General, 125*, 331–359.

Lehmann, A. C., & Ericsson, K. A. (1998). Preparation of a public piano performance: The relation between practice and performance. *Musicæ Scientiæ, 2*, 67–94.

Maehr, M. L., Pintrich, P. R., & Linnenbrink, E. A. (2002). Motivation and achievement. In R. Colwell & C. Richardson (Eds.), *The New Handbook of Research on Music Teaching and Learning* (pp. 348–372). Oxford: Oxford University Press.

Matthay, T. (1926). *On Memorizing and Playing from Memory and on the Laws of Practice Generally.* Oxford: Oxford University Press.

May, C. P., Hasher, L., & Stoltzfus, E. R. (1993). Optimal time of day and the magnitude of age differences in memory. *Psychological Science, 4*, 326–330.

McPherson, G. E., & McCormick, J. (1999). Motivational and self-regulated learning components of musical practice. *Bulletin of the Council for Research in Music Education, 141*, 98–102.

McPherson, G. E., & Renwick, J. M. (2001). A longitudinal study of self-regulation in children's musical practice. *Music Education Research, 3*, 169–186.

McPherson, G. E., & Zimmerman, B. J. (2002). Self-regulation of musical learning: A social cognitive perspective. In R. Colwell & C. Richardson (Eds.), *The New Handbook of Research on Music Teaching and Learning* (pp. 327–347). Oxford: Oxford University Press.

Miklaszewski, K. (1989). A case study of a pianist preparing a musical performance. *Psychology of Music, 17*, 95–109.

Miklaszewski, K. (1995). Individual differences in preparing a musical composition for public performance. In M. Manturzewska, K. Miklaszewski & A. Bialkowski (Eds.), *Psychology of Music Today: Proceedings of the International Seminar of Researchers and Lecturers in the Psychology of Music* (pp. 138–147). Warsaw: Fryderyk Chopin Academy of Music.

Neuhaus, H. (1973). *The Art of Piano Playing*. New York: Praeger.

Nielsen, S. G. (1999). Learning strategies in instrumental music practice. *British Journal of Music Education, 16*, 275–291.

Nielsen, S. G. (2001). Self-regulating learning strategies in instrumental music practice. *Music Education Research, 3*, 155–167.

Noyle, L. J. (1987). *Pianists on Playing: Interviews with Twelve Concert Pianists*. London: Scarecrow Press.

O'Neill, S. A. (1997). The role of practice in children's early musical performance achievement. In H. Jørgensen & A. C. Lehmann (Eds.), *Does Practice Make Perfect? Current Theory and Research on Instrumental Music Practice* (pp. 53–70). Oslo: Norwegian Academy of Music.

O'Neill, S. A., & McPherson, G. E. (2002). Motivation. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 31–46). Oxford: Oxford University Press.

Pitts, S. E., Davidson, J. W., & McPherson, G. E. (2000a). Models of success and failure in instrumental learning: Case studies of young players in the first 20 months of learning. *Bulletin of the Council for Research in Music Education, 146*, 51–69.

Pitts, S. E., Davidson, J. W., & McPherson, G. E. (2000b). Developing effective practice strategies: Case studies of three young instrumentalists. *Music Education Research, 2*, 45–56.

Rubin-Rabson, G. (1937). *The Influence of Analytical Pre-Study in Memorizing Piano Music*. New York: Archives of Psychology.

Rubin-Rabson, G. (1939). Studies in the psychology of memorizing piano music: I. A comparison of the unilateral and the coordinated approaches. *Journal of Educational Psychology, 30*, 321–345.

Rubin-Rabson, G. (1940a). Studies in the psychology of memorizing piano music: II. A comparison of massed and distributed practice. *Journal of Educational Psychology, 31*, 270–284.

Rubin-Rabson, G. (1940b). Studies in the psychology of memorizing piano music: III. A comparison of the whole and part approach. *Journal of Educational Psychology, 31*, 460–476.

Rubin-Rabson, G. (1941a). Studies in the psychology of memorizing piano music: IV. The effect of incentive. *Journal of Educational Psychology, 32*, 45–54.

Rubin-Rabson, G. (1941b). Studies in the psychology of memorizing piano music: V. A comparison of pre-study periods of varied length. *Journal of Educational Psychology, 32*, 101–112.

Rubin-Rabson, G. (1941c). Studies in the psychology of memorizing piano music: VI. A comparison of two forms of mental rehearsal and keyboard overlearning. *Journal of Educational Psychology, 32*, 593–602.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*, 68–78.

Sandor, G. (1981). *On Piano Playing: Motion, Sound, Expression.* New York: Schirmer.

Seashore, C. E. (1938/1967). *Psychology of Music.* New York: Dover. (Original work published in 1938.)

Seashore, R. H. (1939). Work methods: An often neglected factor underlying individual differences. *Psychological Review, 46*, 123–141.

Shockley, R. (1986). A new approach to memorization. *Clavier, July–August*, 20–23.

Simonton, D. K. (1999). Talent and its development: An emergenic and epigenetic model. *Psychological Review, 106*, 435–457.

Simonton, D. K. (2001). Talent development as a multidimensional, multiplicative, and dynamic process. *Current Directions in Psychological Science, 10*, 39–43.

Sloboda, J. A. (1996). The acquisition of musical performance expertise: Deconstructing the "talent" account of individual differences in musical expressivity. In K. A. Ericsson (Ed.), *The Road to Excellence: The Acquisition of Expert Performance in the Arts and Sciences, Sports, and Games* (pp. 107–126). Mahwah, NJ: Erlbaum.

Sloboda, J. A., Hermelin, B., & O'Connor, N. (1985). An exceptional musical memory. *Music Perception, 3*, 155–170.

Sloboda, J. A., Davidson, J. W., Howe, M. J. A., & Moore, D. G. (1996). The role of practice in the development of performing musicians. *British Journal of Psychology, 87*, 287–309.

Sosniak, L. A. (1985). Learning to be a concert pianist. In B. S. Bloom (Ed.), *Developing Talent in Young People* (pp. 19–67). New York: Ballantine Books.

Weiner, B. (1974). An attributional interpretation of expectancy-value theory. In B. Weiner (Ed.), *Cognitive Views of Human Motivation* (pp. 51–70). London: Academic Press.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92*, 548–573.

Weiner, B. (1986). *An Attributional Theory of Motivation and Emotion.* New York: Springer-Verlag.

Weiner, B., Russell, D., & Lerman, D. (1976). The cognition-emotion process in achievement-related contexts. *Journal of Personality and Social Psychology, 20*, 82–92.

Weiner, B., Frieze, I., Kukla, A., Reed, L., Rest, S., & Rosenbaum, R. M. (1972). Perceiving the causes of success and failure. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins & B. Weiner (Eds.), *Attribution: Perceiving the Causes of Behavior* (pp. 95–120). Morristown, NJ: General Learning Press.

Williamon, A., & Valentine, E. (2000). Quantity and quality of musical practice as predictors of performance quality. *British Journal of Psychology, 91*, 353–376.

Williamon, A., & Valentine, E. (2002). The role of retrieval structures in memorizing music. *Cognitive Psychology, 44*, 1–32.

Williamon, A., Valentine, E., & Valentine, J. (2002). Shifting the focus of attention between levels of musical structure. *European Journal of Cognitive Psychology, 14*, 493–520.

Winner, E. (1996a). *Gifted Children: Myths and Realities.* New York: Basic Books.

Winner, E. (1996b). The rage to master: The decisive role of talent in the visual arts. In K. A. Ericsson (Ed.), *The Road to Excellence: The Acquisition of Expert Performance in the Arts and Sciences, Sports, and Games* (pp. 271–301). Mahwah, NJ: Erlbaum.

Yoon, C., May, C. P., & Hasher, L. (2000). Aging, circadian arousal patterns, and cognition. In D. C. Park & N. Schwarz (Eds.), *Cognitive Aging: A Primer* (pp. 151–171). Hove, UK: Psychology Press.

CHAPTER 3

# MANAGING THE PHYSICAL DEMANDS OF MUSICAL PERFORMANCE

CHRISTOPHER B. WYNN PARRY

In order to achieve the highest levels of performance, a musician needs to be physically, emotionally, and mentally fit, for what is a hugely demanding profession. Sadly, the education and training of performers very often fail to incorporate advice to students on care of the body and prevention of injury. Until quite recently, in fact, it seems to have been accepted that pain is an inevitable accompaniment of a life in music. The pernicious exhortation "no pain, no gain" has been accepted without challenge. However, provided that technique is sound, body fit, mind and spirit calm, the musician can spend a lifetime in music, with only occasional aches and pains, such as affect most people.

Until 1986, there was little understanding of the scope and nature of the medical problems confronting musicians. There were sporadic reports in the literature that musicians suffered cramps, fatigue, and neuritic symptoms (e.g. Singer, 1932), but there were no indications of their incidence and effect on a playing career. It took the publicity of an article in the *New York Times* on Gary Graffman's "dystonia" to reveal the true magnitude of the predicament. Graffman (1986) records that, following the appearance of this article, he was inundated by calls from musicians recounting a wide variety of physical ills.

The medical profession has responded by setting up specialist clinics and instigating surveys into the scope of the problem. From this, one point has emerged as certain: many musicians suffer seriously, yet unnecessarily, for their art. Clearly, the act of practicing and performing intensely over several years can bear profound effects on human anatomy, including musculature, bone structure, circulation, and the respiratory system. Adopting a healthy approach to engaging in music, especially early on in one's career, can have substantial impacts on achieving and maintaining peak performance and, therefore, on one's ability to exploit the practice strategies and performance enhancement interventions included in this book.

The underlying theme of this chapter is that the majority of medical problems facing the musician are preventable. Relatively few medical investigations have been carried out on the physical demands of performance, partly because the music profession is quite naturally secretive about ailments. For example, the incidence of damage to joints or muscles from a professional lifetime of repetitive movements is not known—nor will

it be known without further longitudinal research. Nonetheless, this chapter reviews the existing scientific literature on music and physical health and offers a clinical perspective on some of the more common medical conditions that affect musicians—namely, musculoskeletal structural disorders and focal dystonia. Moreover, it explores some of the more general, nonstructural problems that typically arise in conjunction with certain types of music making (playing the piano, violin, cello, guitar, flute, and conducting) and provides advice on how to prevent such problems before they become career threatening (the equally pertinent and closely related issue of psychological well-being and health is discussed in detail in Chapter 9).

## 3.1 Surveys

The first large-scale surveys on the physical and mental demands of performance have revealed some disturbing facts. Of the 2212 musicians in a survey by Fishbein *et al.* (1988), 76% reported at least one serious medical condition that affected their playing. Brandfonbrener (1990) surveyed 22 000 music teachers, of whom 19% reported performance-related medical problems. Caldron *et al.* (1986) found that, of 250 musicians studied, 57% suffered musculoskeletal symptoms, and 37% had to seek medical help. Newmark and Lederman (1987) studied a group of 79 amateur musicians who normally practiced for an hour a day; when on an intensive course playing seven hours per day, 72% reported musculoskeletal symptoms. A consensus among these and many other reports ascribe performance-related symptoms to a variety of causes, the most common being excess practice, change of instrument, difficult repertory, and stress of a demanding lifestyle. Among students, these include overpractice, change of teacher, stressful conditions, and examinations.

In 1997, the Fédération Internationale des Musiciens (FIM) surveyed 57 orchestras worldwide (James, 2000). This was the most comprehensive survey to date and enquired into a host of factors that might affect the performer—physical, psychological, and pedagogical. Throughout the world, the results were consistent: 56% had suffered pain when playing within the last year and 34% experienced pain more than once a week. The most common sites of pain were the neck, shoulder, and back. In 19%, pain was severe enough to stop performance. To the question "do you think your training college or academy gave you sufficient help in preparing for the stresses and strains of being an orchestral musician," the answer in 83% was "no." It is, thus, clear that musicians can suffer for their art and that their education should be sufficiently geared toward identifying and preventing these problems.

## 3.2 Specific medical diagnoses

The author has run specialist clinics for musicians over the last 40 years and in the last 12 years as a founding member of the British Association for Performing Arts Medicine (BAPAM). A detailed analysis has been carried out on 1046 musicians personally seen at these clinics (see Figure 3.1). Half (just over 48%) had a clear-cut pathology, in which a specific diagnosis could be made. In the remaining 52%, few

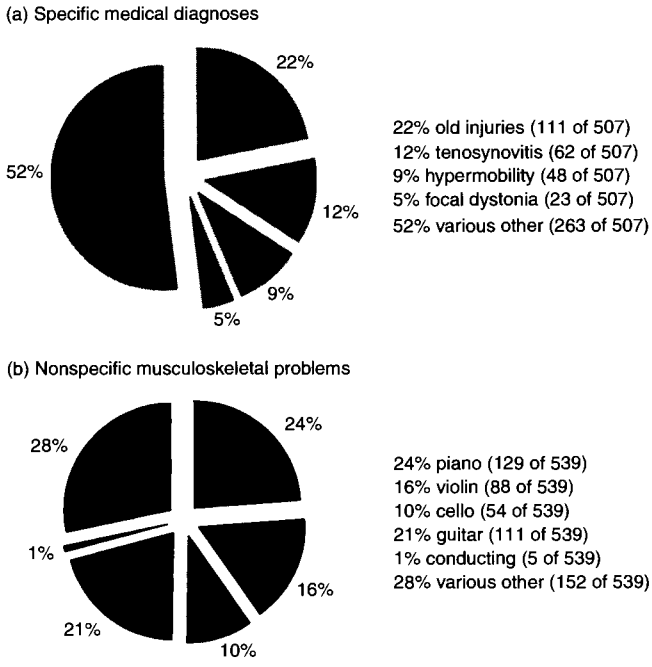



**Figure 3.1** Musicians seen at the British Association for Performing Arts Medicine (BAPAM). A total of 1046, including (a) 507 with specific medical diagnoses and (b) 539 with nonstructural problems resulting, generally, from poor posture, tense neck and shoulder muscles, inappropriate practice regimes, lack of fitness, and stress.

(if any) physical signs could be found. Symptoms were often vague and general, and the musicians themselves attributed these to "doing it wrong" (e.g. poor posture, bad practice technique, lack of fitness, and stress). The fact that such a high proportion of patients were suffering from nonstructural, performance-related problems is a major challenge to the profession, for many of these problems are preventable (see "Non-structural musculoskeletal problems" below for further discussion).

### 3.2.1 Structural disorders

The structural disorders observed resulted from three broad sources: old injuries, tenosynovitis, and hypermobility. The most common of these was the result of an old injury (22%), which had left the musician with residual stiffness of joints, weak muscles, and a lack of confidence that was impairing performance. The moral from this is clear: full rehabilitation to 100% capacity is vital after injury.

Tenosynovitis (or inflammation of tendons) was relatively rare (12%). The label "tenosynovitis" is often too liberally used to describe any ache or pain in a specific area. True tenosynovitis is associated with redness, swelling, crepitus, and severe pain or

restricted movement. This is rare in musicians and usually follows extramusical activities, such as household chores and gardening. The exception is De Quervain syndrome, which affects the long abductor and extensors of the thumb and can occur in percussionists as a result of repetitive, high-pressure activity and in cellists who grip the neck of the cello too tightly.

Just over 9% of the musicians with structural disorders were hypermobile. The term hypermobility refers to the increased range of movement beyond the accepted norm, often described in lay terms as "double jointed." The joints most commonly affected are the fingers, thumb, elbow, shoulder, spine, and knee. To be classified as hypermobile, for example, the knuckle should be passively hyper-extendable to more than 90°; the thumb can be opposed to the volar (front) aspect of the forearm; the elbow and knee can be extended 10° beyond 180°; the hands can be placed flat on the floor without bending the knees; the hand can be placed behind the back as high as the seventh cervical vertebra. There is undoubtedly a genetic element to hypermobility, and its expression is at its height during youth, peaking at 12–15 years and tending to decline slowly thereafter.

It seems that there is a higher incidence of hypermobility among musicians than in the population at large. An increased range of movement can, in fact, be of benefit to the musician (e.g. Liszt's hands could together span three octaves—i.e. about 20 inches or 50 cm). Larsson *et al.* (1987) studied 660 musicians and concluded that hypermobility of the thumb, wrist, and fingers can be an asset in playing instruments such as the flute, violin, and piano for passages requiring repetitive motions, but that it can be a liability in situations where joints are required to be used as stabilizers (e.g. knees and spine for timpanists and others who stand to play).

In general, hypermobile musicians are more liable to strains and sprains of the affected joints. There is evidence to suggest that hypermobile joints are less sensitive to relaying information about position and rate of change of joint movement (i.e. proprioception; see Hall *et al.*, 1995). This can result in greater pressure being exerted by fingers on keys or strings, with the possibility of chronic strain. In addition, hypermobile individuals can suffer generalized rheumatic symptoms. The "full house" of multiple joints being affected is rare; it is, in fact, more common to see only one or two joints affected. At BAPAM, every patient is examined for hypermobility. If the muscles of hypermobile joints are found to be weak, either from repeated or previous injury or from general poor muscle tone, an exercise program to restore full power and improve proprioception is normally recommended. Often a musician's "nonspecific pains" can relate to this, and even if there is a clear cause of symptoms unrelated to hypermobility, guidance for general care of the joints is always given. Advice from a therapist on joint protection and general lifestyle is strongly recommended in those with persistent symptoms.

### 3.2.2 Focal dystonia

Focal dystonia (literally meaning, localized abnormal muscle tone) refers to the incoordination of movement—often with cramps—that can affect fingers or embouchure. It is perhaps the most devastating medical condition that can affect the musician, and of the 507 patients with a clear-cut pathology seen by the author, 5% were diagnosed

with dystonia. The problem usually arises in mid-career, although it may affect musicians sooner, particularly if they began playing in early childhood. Several reports have suggested that dystonia can follow trauma or a particularly stressful situation, such as significant increase in work or the tackling of particularly difficult repertory. Brandfonbrener (1995) reviewed the cases of 58 musicians with dystonia and found that 10 had suddenly increased their playing/practice, 8 had recently undergone an extensive change in technique or a change in instrument, and 7 had engaged in intensive playing/practice after a long repose. In 19, however, no antecedent factor could be implicated. Gary Graffman (1986) described the condition as it affected him:

> To begin with, I had no pain. No numbness, either; nor any such symptom that one could possibly explain to a doctor. For all normal purposes—opening jars, tying shoelaces, manipulating chopsticks, and even at the piano 99.9% of the time, my hand operated as always. It was only in certain extended positions—playing a series of octaves, for example—that my trouble surfaced. I was able to play the first octave normally, but striking each subsequent one caused my fourth finger to draw in more and more, dragging the fifth along with it as the hand contracted, and, of course, hitting wrong notes. This behavior was in no way affected by fatigue, warming up, or "good" or "bad" days. Either I could play a certain pattern, or I could not. Every time it was exactly the same. An octave span of eight notes on the keyboard is about seven inches. A child can deal with that. But I, at fifty, was no longer capable of doing so (p. 3).

The condition is a form of "occupational cramp," and the characteristic sudden loss of coordination has been observed in conjunction with a number of motor skills, including telegraphy, typing, watchmaking, tailoring, snooker playing, and golf (the "yips"). The condition is usually highly task-specific and depends on sensory input. The musician notices the problem only when playing his or her instrument, not when playing other instruments or in any of the usual activities of daily living, such as using a computer. The amount of practice seems to be an important factor in the development of dystonic symptoms. Soloists are much more often affected than orchestral musicians. Furthermore, orchestral musicians can cope by resting or avoiding difficult passages and, thus, abort its development; soloists have nowhere to hide and often must keep playing with ever-increasing symptoms. The condition is often made worse by practicing passages that trigger the dystonic movements.

There even seem to be specific patterns of symptoms for different instruments. Wind players, for instance, may find that they split notes or blur a run (e.g. in clarinetists, it often results in an inability to extend the middle finger but can also interfere with the little finger flexing and the index finger extending). Jazz drummers may not be able to control their foot on the pedal. In guitarists, it is commonly the third or middle finger that is affected (see Hayes, 1987). In piano playing, as in the case of Gary Graffman, the ring and little fingers flex into the palm. The task specificity might be so distinct that symptoms occur only when playing particular passages, or playing scales upward and not downward, or vice versa. Therefore, in some cases, it is not possible to track the appearance of symptoms to one single finger rather than to a complex movement pattern. Another interesting phenomenon is that, in most pianists, the dystonic movement pattern appears only when playing on an actual piano keyboard, rather than when "playing" on the wooden cover of the piano or on plastic keys. This clearly highlights

that focal dystonia is not simply a motor disorder but a disorder of sensorimotor integration, which requires transforming incoming sensory input from an ongoing movement into an adequate motor output, as is necessary for the guidance of skilled motor acts.

The salient feature for diagnosis is that the condition is quite painless, unlike other local disorders that can interfere with playing. The course is variable, but it may be months or years before the musician finally realizes that the situation is unacceptable. The clinician can carry out tests for any local or central neurological disorder that might have precipitated or caused the dystonia, but in most cases, full examination reveals normal functioning—all the tendons, joints, and nerves being in full working order.

It is still not fully understood as to how musicians at the top of their profession (whose technique is seemingly faultless) can be so affected, but in recent years, the knowledge about the underlying complex pathophysiology has increased alongside the development of research. Byl *et al.* (1996) described a primate model for focal hand dystonia. They trained primates to perform a gripping task 300 times a day for several weeks. The primates developed incoordination and cramps similar to symptoms of focal hand dystonia. Topographical mapping of the primates' sensory cortex revealed that the receptive fields of single fingers were remarkably enlarged and overlapping. Their conclusion was that this degradation of sensory representations of single fingers might have caused the observed motor disturbance. Elbert *et al.* (1998), in measuring cortical finger representations in the somatosensory cortex with magnetencephalography (MEG), described a smaller distance (i.e. fusion) between the representations of the digits in the somatosensory cortex for the affected hand of dystonic musicians than for the hands of nonmusicians. They suggested that use-dependent susceptibility to digital representation fusion in the cortex may be involved in the etiology of focal dystonia.

Similar to these findings in the somatosensory cortex, enlargements and blurring of motor cortical representations of single hand muscles have been described in focal hand dystonia patients (Byrnes *et al.*, 1998). These changes in the motor cortical representation may be secondary to altered afferent inputs from both clinically affected and unaffected muscles, which underlines the strong link between the alterations in the sensory and motor system forming a kind of "dysfunctional network" in focal hand dystonia. Part of this network is formed by the basal ganglia, deep-lying structures in the forebrain that are involved in motor control in terms of filtering and focusing sensory input to respective pre-motor and motor cortical areas. "Blurred" sensory input certainly leads to imprecise output, which in turn can lead to an unfocused input to motor cortical areas.

In the FIM survey (James, 2000), one of the questions asked was "do you ever suffer from clumsiness or incoordination of your fingers." Surprisingly, 39.5% answered "yes," and 21% admitted to loss of lip seal or coordination. This highlights that incipient dystonia may be more common than is realized, but on the other hand, it shows that the symptoms are often very unspecific in the beginning. Focal hand dystonia should really only be diagnosed after ruling out any incoordination or painless playing problems that may result from other sources or simply from bad playing habits.

Treatment of this condition remains a difficult task. Improvement of the symptoms can be attained in most cases, but it is understandable that for many patients this is not

enough. Complete recovery is rare and some restrictions may remain with respect to particular pieces, or to the speed and power required in certain motor movements.

Drugs can help in some cases, but of course, possible side effects have to be considered. Botulinum toxin (Botox), which temporarily paralyses the nerve endings in muscles, can relieve symptoms for a while, but when the drug wears off, they reappear. It can also cause muscle weakness; therefore, injections of botulinum toxin should only be administered by an experienced neurophysiologist in order to avoid any unwanted widespread weakening effect. There have not, as of yet, been long-term follow up studies, so it is not entirely clear whether the repeated application of this drug has negative effects itself. Its use in cases with musicians has now been restricted to those where nothing else has worked.

Recently, a very promising therapeutic approach has been developed that is based on the assumption that use-dependent re-organization of cortical networks (as described above) contribute to the fundamental abnormalities seen in focal dystonia. It has used a method of sensory motor retuning that involves the immobilization of different unaffected fingers and the training of sequential finger movements (including the affected one). Through this procedure, a re-differentiation of the finger representation in the somatosensory cortex has been achieved, which has also been accompanied by an amelioration of dystonic symptoms (Candia *et al.*, 2003).

Rehabilitation is slow and relapses not uncommon; therefore, the musician must show great patience. Naturally, dystonic patients can be deeply distressed and tense and may benefit from a general muscular relaxation program, such as Feldenkrais technique (which teaches whole body control over a progressive period) or basic yoga. At the same time, any postural or musculoskeletal dysfunction must be tackled. All authorities agree that the earlier the patient is treated, the more likely a cure is possible. Musicians who present themselves to physicians in late stages of the disorder, with inability to play even quite simple runs without mistakes, rarely improve. This is a difficult area, and much research still is needed to identify the causes of dystonia and to determine an adequate course for rehabilitation and management. It is imperative, however, that affected players seek expert advice as early as possible—indeed, as soon as it is clear that a painless incoordination of hand or embouchure is affecting performance.

## 3.3 Nonstructural musculoskeletal problems

### 3.3.1 General remarks

This section deals with the 52% of musicians seen at BAPAM whose musculoskeletal symptoms were adjudged to be due to technical faults (i.e. improper technique, poor posture, misuse of the body, or inappropriate practicing procedures). These musicians spanned all instruments and styles of music, but the vast majority were pianists (129), violinists (88), cellists (54), and guitarists (111). Common faults and symptoms have been analyzed and studied according to instrument, and a selection of these are discussed below.

Every instrument presents problems in terms of physical interface, some being notoriously user-unfriendly (e.g. the violin, viola, and flute). It is particularly important

that performers are aware of these problems early on in their careers and take steps to prevent injury. Generally speaking, one can alter either the player's approach or the instrument itself. The latter is occasionally possible, but it is the player in virtually all cases who has to make allowances for the instrument. Unlike other performers, the singer *is* the instrument and needs especial care to keep body and mind in first-class order. To that end, regular sessions with a voice coach and, if necessary, joint discussions with a sympathetic laryngologist should be part of a singer's life (despite the common association between body size and the tradition of monumental Wagnerian opera, singers should not be overweight, as obesity reduces vital capacity and physical endurance and is associated with reflux laryngitis).

Common to all the musicians seen at BAPAM were the following faults: poor posture, tense neck and shoulder muscles, inappropriate practice regimes, lack of fitness, and stress.

*Posture*

Postural deficits are common, but as Dommerholt *et al.* (1998) point out, musicians are usually not aware of their own misalignments. When shown their playing position on video, performers are often stunned to see how inappropriate their posture is. The relationship of head to neck and neck to trunk is all-important, as F. M. Alexander discovered when analysing the neck and shoulder symptoms associated with his own loss of function as an actor (Rosenthal, 1987; see also Chapter 10 of this volume). In addition, pelvic and spinal alignment is equally important for trunk and pelvic stability and is essential—whether standing or sitting to play—for muscle control, as well as awareness of body position and rate at which movement occurs.

The most common fault is the forward head posture. Haughie *et al.* (1995) showed a statistically significant relationship between the degree of forward head posture and pain. A typical example in music is the need for string players sitting at the front desks in orchestras to look up and to the side to see the conductor. Similarly, musicians may need to crane forward—and, thus, be susceptible to neck pain—when reading from poorly photocopied scores or if their sight is not properly corrected. A hunched position at the keyboard and forward stoop with the cello are common and also highly likely to cause chronic pain. Furthermore, excessive lordosis and kyphosis (i.e. rounded shoulders and excessive arching of the lower back, respectively) inevitably lead to pain.

Singing teachers have long known of the vital importance of good balanced posture, which is at the heart of comfortable and effective singing. As Davies and Jahn (1998) point out, singing involves vocal athletics, and a fit body and mind are essential. Indeed, performers and teachers of all types could benefit tremendously from a similarly healthy approach to postural awareness. It is encouraging to see that Alexander technique has been incorporated so widely into performance education (see Chapter 10), but more could be done in terms of bringing posture and bodily awareness to the forefront of musical training—be it through the Alexander technique or other methods. Doing so could dramatically decrease the number of *unnecessary* strains on the musician's body. Given the amount of time that musicians spend (and are expected to spend) practicing—as documented in the preceding chapter—it is vital that those hours are spent ergonomically correctly.

*Practice, tension, and fitness*

The human body is not designed to remain in one position for hours at a time; prolonged muscle tension causes fiber shortening and restriction of circulation, which results in inadequate dispersal of muscle metabolites and chronic pain. In response to this, the musician may unknowingly adjust the hold of the instrument, causing stresses elsewhere. To relieve stress in the neck and shoulder, for example, string players frequently alter the grip of the left hand and cause wrist strain; they then complain of wrist pain when the source of the mischief is actually higher up.

Rigid advice on how many hours to practice per day cannot be given, as much depends on the individual capacity for prolonged concentration and level of fitness (see Chapters 9, 11, and 12 for suggestions on how to enhance concentration and fitness). Learning new repertory in a hurry may mean putting in an excessive amount of practice (say, over 5 hours per day), but painful symptoms and damage can be minimized by regular breaks and short spells of physical activity, involving stretching the contracted muscles and freeing up stiffened joints. John Williams, a professional guitarist whose father was a distinguished musician and taught him care of the body from an early age, never practices for more than 20 minutes at a time. He stretches, walks around, carries out deep breathing, and then resumes play. He has said that it takes him 4 hours to do 2.5 hours of good practice, which includes warming up, cooling down, breaks, and taking notes of his performance (Williams, 1998).

Insistence on many hours of practice a day by some music schools and conservatories has led to a culture of work intensity (a recent brochure of one academy has stated that violin pupils are expected to practice 6 hours per day). An excessive workload on tired muscles—compounded by lack of instruction in correct posture, practice regime, and care of the body in general—is the most common cause of injuries with students. While 52% of all musicians seen at BAPAM were suffering because of technical faults, the figure was 70% in students. Warrington *et al.* (2002) found that nonspecific causes of pain (clearly resulting from technical faults) were present in 4% of patients over 40 years old, 19% in the 25–40 age range, and 48% in those under 25.

There are several general recommendations that clinicians can make in this area. Firstly, Spaulding (1988) points out that a child's early teachers and family should emphasize prevention of injury. After the mid-twenties, preventive techniques can be difficult to learn and perhaps even more difficult to incorporate into everyday musical life. There are many aspects to effective prevention—correct technique, good posture, general attitude, and emotional set, as well as a sensible perspective on appropriate quantity of practice. Young students should be regularly encouraged to think about these issues so that the process of self-reflection becomes automatic. This is a serious responsibility for music educators and one that is far too frequently neglected, either from exigencies of time or ignorance. Redmond and Tiermond (2001) carried out a survey of music teachers and found that less than half had received any education on preventive measures. In a pilot study, Spahn *et al.* (2001) found that musculoskeletal symptoms were significantly reduced in students who received a weekly 2-hour instruction on anatomy and physiology relevant to music making, along with specific exercises for the particular instrument.

Secondly, all athletes warm up before and cool down after practice and performance. Coaches have known for years that injuries are common if muscles are not slowly exercised to maximum output. A surprising number of musicians admit that they neglect these precautions. The body dislikes sudden exertion, particularly in an unheated environment. Zaza and Farewell (1997) showed that warming up before and taking breaks during practice can protect musicians from performance-related musculoskeletal disorders.

Thirdly, the intensity of practice advisable can vary with different degrees of fitness. In general, the more physically fit one is—with strength, stamina, flexibility, and good cardiovascular reserve—the more one can play without harm. It has been suggested that the prolonged periods of practice required for musical performance at the highest levels is all that is necessary for fitness. This may be true to a limited extent, in that no specific extra exercises are required for strength and dexterity in the wrists or hands other than playing and practice (except in the case of weakness from old injuries or hypermobility). However, other physical parameters are required, including strong postural muscles to maintain neck and back position (particularly for keyboard and string players). It is equally important to develop stamina, the ability to sustain repeated stress to the musculoskeletal system over a period of time. This is achieved by a program of high-repetition, low-load exercises (conversely, increased strength is achieved through high-load, low-repetition exercises). In all cases, it is wise for musicians to consult an expert physiotherapist or trainer for advice on a particular regime suited to their starting fitness level, physique, and specific instrumental requirements. Also, it is as much the demands of long-distance traveling, adjusting to new countries, irregular meals, change in climate, and anxiety about arrangements that cause fatigue as the actual physical demands of music performance. An extensive discussion of physical fitness, psychological well-being, and stress management is provided in Chapter 9 of this volume.

Fourthly, it has been recognized of late that much good practice can be performed in the mind. Mental rehearsal has been shown to be highly effective (Freymuth, 1993; Stanton, 1994), and electrophysiological studies have revealed that the same neural pathways are excited when musical imaging occurs as when the instrument is actually played. Serious professional musicians are increasingly making mental rehearsal a significant part of their practice, and sports psychologists have shown how such techniques can be used to improve subsequent performance. Chapter 12 reviews and assesses specific mental skills training techniques.

Finally, Zaza and Farewell (1997) insist that other repetitive tasks, such as computing and knitting, should be avoided or limited by the musician. It is not uncommon to find that symptoms develop in musicians who spend 2–3 hours on a computer, on top of their music making.

### 3.3.2 Individual instruments

*Piano*

Of the 129 pianists whose symptoms were adjudged to be due to technical faults, 58 felt pain in the wrists and hands and 26 in the neck and shoulders. Common sources of pain were stiff wrists, a hunched posture, excessive pressure on the keys, and overpracticing

without adequate breaks and thorough warm up. There can be a number of antecedent factors that lead to symptoms from technical faults, including individual differences in basic anatomy and in how pianists engage physically with the instrument.

In terms of anatomy, Charles Rosen (1999), in an article in the *New York Review of Books*, surveyed a number of great pianists and found that size and dimensions of the hands varied considerably. Richter, for instance, could play intervals as large as a twelfth, but Hofmann, one of Rosen's heroes, could only span an octave. Denison (2000) has argued that manufacturers should make keyboards smaller, so as to accommodate those with small hands. It seems that practicing parallel octaves is a typical cause of trouble (Rosen cites Schubert's *Erlkonig* as a particularly vicious example). Sakai (1997) analyzed the physical problems of 40 professional pianists and found that octaves and chords were the most common cause of pain. Pianists would therefore be wise to monitor (and possibly limit) the amount of time given to repetitive practice in this area. Rosen notes that Steinway built a piano in which the ivories were slightly narrower, so that he could manage a ninth. Thus, it is possible that, in time, the instrument may conform to the player's anatomy, but for the moment, pianists with small hands must vary technique in order to master certain physical challenges.

Similarly, problems can emerge as a result of anomalous tendon linkages. Richard Beauchamp, Head of Piano at St. Mary's Music School (Edinburgh), has made a special study of these (see Beauchamp, 1996, 1997, 1999, 2001, 2003). The most common is the absence of the sublimus flexor tendon in the little finger (i.e. the tendon that allows independent bending at the middle joint of the finger). This can be determined by holding the hand flat on a table—palm upwards, with the index, middle, and ring fingers held by the other hand—and trying to bend the little finger at the middle joint without bending the end joint. An inability to do this can present problems, for example, when attempting to use fingerings such as those indicated in bars 415–416 of Figure 3.2. A number of other anomalous linkages can be problematic, including the presence of a connecting tendon between the profundus tendons used in bending the end joints of the little, ring, and middle fingers. These can be closely joined as far up as the wrist, which can interfere with achieving independence of the third, fourth, and fifth fingers. Beauchamp (2003) has recommended several practical solutions—ranging from alternative fingerings to rearrangement of passages between the hands—to avoid

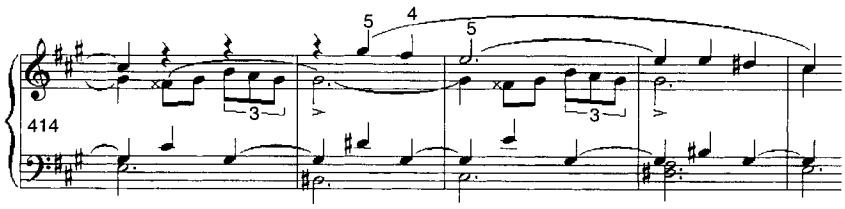



**Figure 3.2** Chopin's Scherzo No. 2 in B-flat minor, Op. 31. Rather than using the indicated fingering, Beauchamp (2003) recommends two possible solutions: (1) playing the *E* with the third finger and rapidly changing to the fifth or (2) not actually trying to join the F# to the *E* physically, as the sustaining pedal will almost certainly be down at that point.

strain caused by these anomalies (see also Winspur & Wynn Parry, 1998, for further information on the musician's hand).

In terms of differences in how individuals engage physically with the piano, Rosen (1999) notes that Horowitz played with his fingers flat, Iturbi held his wrist below the level of the keyboard, Gould used a very low piano stool, and Richter preferred a very high stool. It seems that many approaches have been proposed as to the ideal curvature of the hand when playing. Matthay (1903/1950) stressed the importance of a curved hand, using wrist and forearm rotation, and moving the body rather than the wrist for major stretches (certainly, from a medical viewpoint, playing with stiff wrists and too flat hands can cause problems). Bejjani *et al.* (1989) compared three types of hand—flat, arched, and flexed—using video and sound recording to study the finger, wrist, elbow, and shoulder joints. The flexed-finger approach involved more ulnar deviation (i.e. rotating the wrist in the direction of the little finger), which is a potent cause of injury. Until a long-term study is mounted with state-of-the-art electronic recording of the muscles used in playing (comparing different techniques over many years and the incidence of musculoskeletal symptoms that result from each), pianists can only be guided by traditional teaching methods that obey the laws of human ergonomics. Grindea (2001) strongly emphasizes the importance of a relaxed wrist and the position of the piano—upright posture, head level, shoulders relaxed, and wrists in line with the fifth finger and elbow. Brown (2000) has reviewed her own experience from master-classes over many years and found a direct correlation between injury and tense or awkward hand and finger positions. She counsels against a down-flexed wrist, and keeping the elbows too close to the body.

In general, the pianist must think of the upper limb as a single connected unit and make full use of the shoulders for power—not the hand. Excessive lifting of the fingers off the keys is dangerous. Many players attempt to overcome the natural weakness of extension of the ring finger (which has less independence in this movement than the other fingers) by resistance exercise, which can be a major cause of pain. It is better to use the excursion of the wrist "dropping" on to the key.

*Violin and viola*

Among the 88 violinists seen at BAPAM, symptoms were equally divided between the neck and shoulder and the forearm and hand/wrist. Common faults were poor posture, overpractice, unsatisfactory chin/shoulder rest, and too tight pressure on the bow—and in students, a change of teacher and re-jigging of technique.

Filling the gap between chin and chest is a major problem. A common fault among violinists, typically due to an ill-fitting shoulder rest, is to hunch up the left shoulder, which can cause excess tension in the neck and shoulder muscles. Using electrical recordings, Levy *et al.* (1992) analyzed the activity in the shoulder and neck muscles in 15 professional violinists with and without shoulder rests. There was a highly significant reduction in muscle activity with the rest. Also, violinists should ideally not be too tall because additional height makes the gap between chin and chest difficult to fill comfortably. Of course, one option for players with long or unusually shaped necks is to obtain a custom-built chin/shoulder rest. Weilerstein and Neal (2000) strongly recommend all players not to have the instrument always under the chin and never to

practice for more than 5 hours per day. They also note that injuries can arise from stubborn tuning pegs.

Fischer (1998) has written extensively on technique and musculoskeletal problems. He points to a multitude of interrelated factors, many subtle and invisible, that can set off a chain reaction of tension. Some important factors are speed of dropping the finger in relation to tempo, speed of bow related to pressure and distance from the bridge, and different speeds and widths in vibrato. Fischer points out that problems arise when players point the violin too far forward or use too much clockwise rotation of the left hand or finger pressure. Drop release, not drop press, avoids locking the hand, which is a common cause of pain, as is too strong a grip on the bow with counter pressure of the thumb. Raising the fingers too high off the string in fast passages triples the energy required. Also, postural errors are common, with the hip joint collapsing forward, pulling down from the diaphragm causing the upper back to bend forward and the shoulders to pull up. Fischer concludes that, in almost every example of a player suffering from debilitating tension, there are simple technical/physical causes that can quickly be rectified. "It is really never too late, and players of any age and stage can continue to polish and refine their overall technique, and move ever closer to truly effortless playing" (p. 22).

Similar problems are seen in viola players, and with the instrument being heavier, it can be more disabling. Violas vary surprisingly in size and weight, and symptoms may arise after a change of instrument. Donald Ehrich of the San Francisco Symphony Orchestra had severe problems with his bowing arm. The Rivinus viola provided a deeper cut away on the bowing side and angled the fingerboard towards the right hand, thus "untwisting" his left arm with excellent results (Markison, 1998). Gelberman *et al.* (1981) have cut away the body on the right side of the instrument to allow reduction in wrist flexion in higher positions, particularly useful for musicians with incipient median nerve compression (i.e. carpal tunnel syndrome).

### *Cello*

With the 54 cellists, symptoms were equally divided between shoulder/neck and wrist/hand. Causes were excessive playing, poor posture, weak shoulder and arm muscles, and poor bow hold. Extended thumb with excess pressure in the bow hand was a common cause of pain and loss of sonorous tone, particularly in students.

It seems that cellists suffer some degree of backache as part of their life as a result of the playing position. Most take great trouble to ensure satisfactory seating with back cushions and front blocks, and many employ back exercises to avoid chronic pain. The cellist Bernard Gregor-Smith (1998) advises that the torso should incline towards the cello, counteracting the tendency for the weight of the cello to push back on the torso. One should lean forward from the base of the spine—not halfway up the back, which is a common cause of back pain.

Joan Dixon, a professional cellist whose masterclasses were legendary, noted that symptoms frequently arise when the cellist's area of motion is too restricted. If movements are free and full, the posture balanced, and there is no tension in back or neck, playing should be pain free.

Cellists would be wise to adopt a program of exercises to maintain strength in the spinal extensor muscles and abdominal muscles, in addition to the fitness and warm up procedures recommended for all musicians (see Chapter 9).

*Guitar*

A relatively large number of guitarists (111) seen at BAPAM had technical faults. This was principally because most of the guitarists were self-taught and had no formal instruction in care of the body or posture while playing. Pain was felt predominantly in the fingers, wrist, and hand (twice as common as in the neck and shoulders), and common faults were poor posture, tense neck/shoulders leading to painful adjustments of the wrist, poor practice regime, change of teacher, and many hours spent on the computer (as most of the guitarists had jobs outside of music).

In the classical guitar, the wrist and thumb tends to be hyperflexed, and resting the inner arm near the elbow against the instrument can irritate the ulnar nerve where it runs superfically at the inner elbow. This results in pins and needles in the fourth and fifth fingers and weakness of the small hand muscles. As John Williams (1998) has pointed out, unlike piano and violin, little large-scale movement is required of the guitarist. Indeed, apart from movement of the fingers of the left hand along the strings and (to some extent) the fingers of the right hand up and down the strings, no natural movements of the upper limb are necessitated by the physical production of music.

The inevitable slight twisting of the torso consequent on the left foot being on a stool is likely to lead to neck and back pain, and it is essential that players move their bodies and stretch frequently during practice. Kember (1995a, 1995b, 1996a, 1996b, 1996c, 1996d, 1996e), a senior physiotherapist, has made a special study of guitarists' problems and, in a series of articles in *Classical Guitar*, has summarized good practice to avoid injury: (1) the choice of guitar should relate to body shape (a broad, wide palm with long fingers needs a wider neck than with stout fingers); (2) posture must allow for no rotation of the pelvis, shoulders must be level, and head central (a slumped position with a twisted pelvis is all too common); (3) if the guitar is held too high, there will be pain in the neck and shoulder; (4) the seat should not slope too far forward (a wedge shaped cushion of firm foam on a poorly shaped chair can avoid back strain); and (5) warm-up should include uncurling of the spine from a flexed position, general postural alignment, breathing exercises, and attention to balance of the shoulders (too early movement of the shoulder blades may follow imbalance and cause stress in the trapezii).

Adequately broad straps are vital for rock guitarists who stand with a heavy instrument suspended from the neck. These musicians, more than any, must pay strict attention to posture and build up powerful neck and shoulder muscles. Commonly, they experience forearm and wrist pain and a painful neck and back as a result of balancing too large a guitar in a lifted playing posture. This tempts the player to suspend the guitar too low to enable the hand to get on to the strings. Perrin (1997) has written at length on problems of student guitarists, illustrating good and bad hand, wrist, and trunk positions.

*Flute*

It is difficult to imagine a more diabolically designed instrument than the flute, which demands contorted positions of both arms. The left forearm has to be hyperextended,

and the left wrist deviated radially. Hyperextending the left index finger provides a sort of shelf and, together with the radial deviator of the wrist, can compress the posterior interosseous nerve, causing pain and weakness in the fingers. Pain is not uncommon in the right thumb, particularly if excessive tension is used. Norris (2000) advocates an orthosis (the "Stedirest"), which uses a nonslip cork-lined metal grip to handle the instrument. It distributes the pressure over a wider area of the thumb and directs the weight of the rods through the base of support preventing the flute from rolling inwards.

Katherine Butler, a hand therapist and professional flute player, finds that muscle pain in the forearm is not uncommon among musicians she has treated. This can require re-education of technique, for which a finger block over the flute keys can be of great value. Also, she finds that neck pain can result from inadequate power in the shoulder and neck muscles with inability to maintain the playing position for extended periods, particularly when the arms are held too high.

Failure to take adequate breaks in practice is particularly likely to lead to musculoskeletal problems in the upper limb in flute players, owing to the "strained" position that players need to adopt. In cases where musculoskeletal symptoms persist, the Markison angled flute has an 88° bend that allows untwisting of both arms. If the player has very small muscles of the hand and wrist, extensions can be added to keys and a program of hand exercises instituted.

*Conductors*

Conductors are notoriously long lived and seem relatively free of musculoskeletal problems (Simons, 1986). This is probably due to the regular movements they make of arms and trunk. Simons studied 153 choral conductors, few of whom reported any significant musculoskeletal problems, other than backache.

At BAPAM, five international-level conductors have been seen who developed disabling damage to the tendon mechanisms of the shoulder. Repeated friction of these tendons against the end of the collarbone can cause severe inflammation and, ultimately, rupture. This results from the restricted arc of movement when the shoulder is held at about 70–90°. The shoulder is put under particular stress when conducting opera and chorus, when the arm has to be lifted up well above 90° for the baton to be seen. Physiotherapy and the judicious use of a steroid injection (a maximum of three in one year) usually solve this problem; however, there have been conductors who have submitted to overfrequent use of steroid injections and have subsequently ruptured tendons, necessitating major surgery and prolonged rehabilitation.

### 3.3.3 General advice to prevent injury

Musculoskeletal symptoms and injuries can be greatly reduced on all instruments and within all performance genres by the following precautions:

- Warming-up: this does not mean playing scales and specific exercises for the instrument; rather, it means stretching the arms above the head, stretching the neck in all directions, arm circling, elbow rotation, wrist shaking, knee bends, trunk rotation, and deep breathing. Even 5 minutes can be beneficial (see Chapter 9).
- Cooling down in a similar manner.

- Good posture throughout playing (see Chapter 10).
- Regular breaks every 20–30 minutes, with arm stretching, walking about, and deep breathing.
- Practicing a relaxation technique, using it before performance and auditions in order to control the body and not let the body control one's performance (see Chapter 12).
- Maintaining a good diet.
- Taking time to relax and pursue general cultural and sporting activities.

## 3.4 Conclusions

Life as a musician is fraught with physical, emotional, and mental demands. When injury or illness strikes, the show has to go on, often to the detriment of the player's future health. In the FIM survey (James, 2000), respondents were asked to list the top 10 stressors in orchestra life. The order and nature were virtually universal throughout the world:

- a conductor who saps confidence
- an incompetent conductor
- experiencing problems with the instrument
- playing an orchestral solo
- illegible music
- disorganized rehearsal time
- an incompatible desk partner
- having medical problems that affect work
- making a mistake when performing
- inadequate financial reward.

Clearly, musicians lead a demanding life. In Britain, for example, performers must hasten between rehearsals, concerts, recording sessions, and teaching (with many musicians having to seek work outside of music to make ends meet). In addition, they need to practice, and on tour, schedules are inordinately demanding with practically no respite on return. A top London orchestra recently toured the USA for 6 weeks, with almost nightly concerts, and was back in the concert hall the day after returning to London. In Britain, musicians are underpaid, overworked, and undervalued, yet Danziger (1997), who interviewed over 50 members of a leading orchestra about their hugely demanding lifestyle, reported that only 3 would not enter a career in music if given their time again.

It is difficult to say whether the situation will improve. Increasingly, more demands are being made on musicians, both to cram in more concerts and to take their art into the community. This is entirely laudable but has obvious implications for health. BAPAM advocates that musicians take every opportunity to improve their physical and

mental condition. There is a need for a continuing dialogue with administrators to minimize stress and to ensure proper travel arrangements, satisfactory seating and lighting, more attention to the hazards of noise, regular time off (tours are *not* holidays), and a better career structure.

The whole question of damage to hearing from playing music has been opened up with the publication of *A Sound Ear* by the Association of British Orchestras (Wright Reid, 2001). There is no question that hearing suffers from playing and is not only related to the proximity of loud instruments (e.g. woodwind in front of brass) but may be due to sensitivity of the player to the high frequencies of his or her own instrument. The right answer has yet to be made clear, but it is likely to be a matter of compromise, with use of protective shields, planning repertory to avoid too many loud pieces in one concert, and improved earplugs. It is generally agreed that something has to be done to minimize this potent source of harm to orchestral players.

The implications of this discussion on physical health for students are obvious. There must be an overhaul of the curricula of music schools and conservatories. Students should be given thorough instruction on care of the body—one lecture in 4 years is inadequate. This must include posture, general and specific fitness of their musculoskeletal system, good breathing techniques, satisfactory compromise with problems of the instrumental interface, sensible practice regimes (including warming up and cooling down), coping strategies for performance anxiety and stage fright, the need for appropriate sport and recreation, and relaxation techniques. There is ample evidence to show that it is possible for students to perfect their techniques and live a fulfilling social and cultural life. When asked by one of his best pupils how he could improve his playing, Brahms reportedly replied "practice an hour less a day, and read a book."

## Further information and reading

The British Association for Performing Arts Medicine offers free guidance at its London headquarters and its Manchester regional clinic (see *www.bapam.org.uk* or write to BAPAM, Fourth Floor, Totara Park House, 34–36 Grays Inn Road, London WC1X, UK, for further information). Details of similar such clinics around the world can be obtained through the websites of the following organizations: the Performing Arts Medicine Association (*www.artsmed.org*), the International Foundation for Performing Arts Medicine (*www.ifpam.org*), Médecine des Arts (*www.arts-medicine.com*), and the International Association for Dance, Medicine, and Science (*www.iadms.org*). For further reading on music and health, see:

Brandfonbrener, A. G., & Kjelland, J. M. (2002). Music medicine. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 83–96). Oxford: Oxford University Press.

Grindea, C. (Ed.). (1995). *Tensions in the Performance of Music: A Symposium* (6th ed.). London: Kahn and Averill.

Tubiana, R., & Amadio, P. C. (Eds.). (2000). *Medical Problems of the Instrumentalist Musician.* London: Martin Dunitz.

Winspur, I., & Wynn Parry, C. B. (Eds.). (1998). *The Musician's Hand: A Clinical Guide.* London: Martin Dunitz.

## References

Bejjani, F. J., Ferrara, L., Xu, N., Tomaino, C. M., Pavlidis, L., Wu, J., & Dommerholt, J. (1989). Comparison of three piano techniques as an implementation of a proposed experimental design. *Medical Problems of Performing Artists, 4*, 109–113.

Beauchamp, R. (1996). *Curved Fingers and Tension?* Retrieved August 28, 2003, from the Music and Health website: *www.musicandhealth.co.uk/tension.html.*

Beauchamp, R. (1997). Curved fingers and tension? *Classical Piano Magazine, March/April*, 27.

Beauchamp, R. (1999). *Our Anatomical Differences.* Retrieved August 28, 2003, from the Music and Health website: *www.musicandhealth.co.uk/differences.html.*

Beauchamp, R. (2001). *The Ergonomics of Piano Playing.* Retrieved August 28, 2003, from the Music and Health website: *www.musicandhealth.co.uk/ergonomics.htm.*

Beauchamp, R. (2003). *Examples of Passages Which May Cause Problems due to Tendon Linkages or Absences.* Retrieved August 28, 2003, from the Music and Health website: *www.musicand health.co.uk/linkages.htm.*

Brandfonbrener, A. G. (1990). The epidemiology and prevention of hand and wrist injuries in performing artists. *Hand Clinics, 6*, 365–378.

Brandfonbrener, A. G. (1995). Musicians with focal dystonia: A report of 58 cases seen during a ten-year period at a performing arts clinic. *Medical Problems of Performing Artists, 10*, 115–120.

Brown, S. (2000). Promoting a healthy keyboard technique. In R. Tubiana & P. C. Amadio (Eds.), *Medical Problems of the Instrumentalist Musician* (pp. 559–571). London: Martin Dunitz.

Byl, N. N., Merzenich, M. M., & Jenkins, W. M. (1996). A primate genesis model of focal dystonia and repetitive strain injury: I. Learning-induced dedifferentiation of the representation of the hand in the primary somatosensory cortex in adult monkeys. *Neurology, 47*, 508–520.

Byrnes, M. L., Thickbroom, G. W., Wilson, S. A., Sacco, P., Shipman, J. M., Stell, R., & Mastaglia, F. L. (1998). The corticomotor representation of upper limb muscles in writer's cramp and changes following botulinum toxin injection. *Brain, 121*, 977–988.

Caldron, P. H., Calabrese, L. H., Clough, J. D., Lederman, R. J., Williams, G., & Leatherman, J. (1986). A survey of musculoskeletal problems encountered in high-level musicians. *Medical Problems of Performing Artists, 1*, 136–139.

Candia, V., Wienbruch, C., Elbert, T., Rockstroh, B., & Ray, W. (2003). Effective behavioral treatment of focal hand dystonia in musicians alters somatosensory cortical organization. *Proceedings of the National Academy of Science, 100*, 7942–7946.

Danziger, D. (1997). *The Orchestra: The Lives Behind the Music.* London: Harper Collins.

Davies, D. G., & Jahn, A. F. (1998). *Care of the Professional Voice: A Management Guide For Singers, Actors, and Professional Voice Users.* Oxford: Oxford University Press.

Denison, C. (2000). Performer's perspective: Hand size versus the standard piano keyboard. *Medical Problems of Performing Artists, 15*, 111–114.

Dommerholt, J., Norris, R. N., & Shaheen, M. (1998). Therapeutic management of the instrumental musician. In R. T. Sataloff, A. G. Brandfonbrener, & R. J. Lederman (Eds.), *Performing Arts Medicine* (pp. 277–290). San Diego: Singular Publishing.

Elbert, T., Candia, V., Altenmuller, E., Rau, H., Sterr, A., Rockstroh, B., Pantev, C., & Taub, E. (1998). Alteration of digital representations in somatosensory cortex in focal hand dystonia. *Neuroreport, 9*, 3571–3575.

Fischer, S. (1998). The interface: Violin. In I. Winspur & C. B. Wynn Parry (Eds.), *The Musician's Hand: A Clinical Guide* (pp. 16–22). London: Martin Dunitz.

Fishbein, M., Middelstadt, S. E., Ottati, V., Strauss, S., & Ellis, A. (1988). Medical problems among ICSOM musicians: Overview of a national survey. *Medical Problems of Performing Artists, 3*, 1–8.

Freymuth, M. (1993). Mental practice for musicians: Theory and applications. *Medical Problems of Performing Artists, 8*, 141–143.

Gelberman, R. H., Hergenroeder, P. T., Hargens, A. R., Lundborg, G. N., & Akeson, W. H. (1981). The carpal tunnel syndrome: A study of carpal canal pressures. *Journal of Bone & Joint Surgery, 63*, 380–383.

Graffman, G. (1986). Doctor can you lend an ear? *Medical Problems of Performing Artists, 1*, 3–4.

Gregor-Smith, B. (1998). The interface: Cello. In I. Winspur & C. B. Wynn Parry (Eds.), *The Musician's Hand: A Clinical Guide* (pp. 24–25). London: Martin Dunitz.

Grindea, C. (2001). *Healthy Piano Technique.* London: Richard Schauer.

Hall, M. G., Ferrell, W. R., Sturrock, R. D., Hamblen, D. L., & Baxendale, R. H. (1995). The effect of the hypermobility syndrome on knee joint proprioception. *Rheumatology, 34*, 121–125.

Haughie, L. J., Fiebert, I. M., & Roach, K. E. (1995). Relationship of forward head posture and cervical backward bending to neck pain. *Journal of Manual and Manipulative Therapy, 3*, 91–97.

Hayes, B. (1987). Painless hand problems of string pluckers. *Medical Problems of Performing Artists, 2*, 39–40.

James, I. M. (2000). Survey of orchestras. In R. Tubiana & P. C. Amadio (Eds.), *Medical Problems of the Instrumentalist Musician* (pp. 195–201). London: Martin Dunitz.

Kember, J. (1995a). You and your guitar: 1. *Classical Guitar, November*, 24–25.

Kember, J. (1995b). You and your guitar: 2. *Classical Guitar, December*, 16–18.

Kember, J. (1996a). You and your guitar: 3. *Classical Guitar, January*, 16–18.

Kember, J. (1996b). You and your guitar: 4. *Classical Guitar, February*, 26–27.

Kember, J. (1996c). You and your guitar: 5. *Classical Guitar, March*, 31–32.

Kember, J. (1996d). You and your guitar: 6. *Classical Guitar, April*, 26–27.

Kember, J. (1996e). You and your guitar: 7. *Classical Guitar, August*, 26–27.

Larsson, L.-G., Baum, J., & Mudholkar, G. S. (1987). Hypermobility: features and differential incidence between the sexes. *Arthritis & Rheumatism, 30*, 1426–1430.

Levy, C. E., Lee, W. A., Brandfonbrener, A. G., Press, J., & Levy, A. E. (1992). Electromyographic analysis of muscular activity in the upper extremity generated by supporting a violin with and without a shoulder rest. *Medical Problems of Performing Artists, 7*, 103–109.

Markison, R. E. (1998). Adjustment of the musical interface. In I. Winspur & C. B. Wynn Parry (Eds.), *The Musician's Hand: A Clinical Guide* (pp. 149–159). London: Martin Dunitz.

Matthay, T. (1903/1950). *The Act of Touch in All its Diversity: An Analysis and Synthesis of Pianoforte Tone-Production.* London: Bosworth and Co. (Original work published in 1903).

Newmark, J., & Lederman, R. J. (1987). Practice doesn't necessarily make perfect: Incidence of overuse syndrome in amateur instrumentalists. *Medical Problems of Performing Artists, 2*, 93–97.

Norris, R. (2000). Applied ergonomics. In R. Tubiana & P. C. Amadio (Eds.), *Medical Problems of the Instrumentalist Musician* (pp. 595–613). London: Martin Dunitz.

Perrin, P. (1997). Réflexion autour du schéma postural des guitaristes. *Médecine des Arts, 22,* 13–19.

Redmond, M., & Tiermond, A. M. (2001). Knowledge and practice of piano teachers in preventable playing-related injuries in high school students. *Medical Problems of Performing Artists, 16,* 32–38.

Rosen, C. (1999). Playing the piano. *New York Review of Books,* October 21, pp. 49–52.

Rosenthal, E. (1987). The Alexander technique. What it is and how it works: Work with three musicians. *Medical Problems of Performing Artists, 2,* 53–57.

Sakai, N. (1992). Hand pain related to keyboard techniques in pianists. *Medical Problems of Performing Artists, 7,* 63–65.

Simons, H. (1986). Health and the choral conductor. *Medical Problems of Performing Artists, 1,* 56–57.

Singer, K. (1932). *Diseases of the Musical Profession: A Systematic Presentation of Their Causes, Symptoms, and Methods of Treatment* (W. Lakond, Trans.). New York: Greenberg.

Spahn, C., Hildebrandt, H., & Seidendanz, K. (2001). Effectivess of a prophylactic course to prevent playing-related health problems of music students. *Medical Problems of Performing Artists, 16,* 24–31.

Spaulding, C. (1988). Before pathology: Prevention for performing artists. *Medical Problems of Performing Artists, 3,* 135–139.

Stanton, H. E. (1994). Reduction of performance anxiety in music students. *Australian Psychologist, 29,* 124–127.

Warrington, J., Winspur, I., & Steinwede, D. (2002). Upper extremity problems in musicians related to age. *Medical Problems of Performing Artists, 17,* 131–134.

Weilerstein, D., & Neal, C. (2000). Violin technique. In R. Tubiana & P. C. Amadio (Eds.), *Medical Problems of the Instrumentalist Musician* (pp. 573–594). London: Martin Dunitz.

Williams, J. (1998). The interface: Guitar. In I. Winspur & C. B. Wynn Parry (Eds.), *The Musician's Hand: A Clinical Guide* (pp. 22–24). London: Martin Dunitz.

Winspur, I., & Wynn Parry, C. B. (Eds.). (1998). *The Musician's Hand: A Clinical Guide.* London: Martin Dunitz.

Wright Reid, A. (2001). *A Sound Ear.* London: Association of British Orchestras.

Zaza, C., & Farewell, V. T. (1997). Musicians' playing-related musculoskeletal disorders: An examination of risk factors. *American Journal of Industrial Medicine, 32,* 292–300.

CHAPTER 4

# MEASURING PERFORMANCE ENHANCEMENT IN MUSIC

GARY E. McPHERSON AND EMERY SCHUBERT

Musical performance incorporates many facets—hours of laborious practice, mental preparation, and physical well-being—and for many, it is the performance itself that is the ultimate goal of all this preparation. However, once the performance is complete (and sometimes even before), listeners will cast judgment upon it. The assessment may be in the form of a rapturous ovation or a cool murmur of unconvincing applause, a grade or report that provides an evaluation of how well the musician has satisfied certain criteria, or alternatively, a critic writing polemical or glowing reports in the next day's paper. Regardless of the assessment process, it is an integral part of the performance.

The first part of this chapter attempts to deconstruct some of the (often unwritten) assumptions of assessment, so that readers are aware of the flaws and limitations that are inherent in any form of evaluation made by humans. This leads to the central question addressed in the final part of the chapter, which details what performers can do, in addition to performing at their peak, to enhance how their performance is received and evaluated by an audience. While most of this book focuses on how musicians can improve specific aspects of their performance, the current chapter examines performance from a different perspective, by exploring how musicians might enhance their own performance *assessment.*

## 4.1 The process of assessing musical performances

Typical forms of music assessment seek either to determine the ranking of a performance in comparison with others (*norm referenced*), or how well a performance satisfies predetermined examination criteria (*criterion based*). Norm-referenced procedures are common in competitions, eisteddfods, and music festivals, where the intention is to rank the musicians or ensembles from most to least accomplished, or to determine a winner and runners up. Criterion-based methods are more common in institutions that seek to determine how much progress or learning has been achieved and whether the musician or ensemble can demonstrate mastery of certain predetermined skills. The work of a music critic in some ways bridges both purposes in that the critique will often compare the musician to other performances that have taken place in the same venue or city, in addition to commenting on how well the performance satisfies generally agreed principles about how the work should be and was performed.

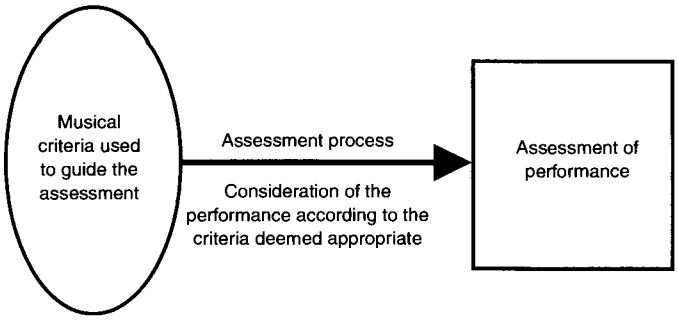



**Figure 4.1** Assumed assessment process.

In both the norm referenced and criterion based forms of assessment, it is typical for judges to focus their assessment on a series of *musical* parameters to guide their evaluation—be they the personal criteria that result from their previous work as evaluators and musicians, or the type of criteria that have been identified by authorities assigned with the task of defining the most appropriate parameters to guide an evaluation. Figure 4.1 depicts this process.

This perspective has several limitations because any music evaluation will be influenced by a number of additional factors that impact on the quality of the evaluation. Thus, a more complete understanding of the music assessment process only comes with an expanded view that includes—along with the musical criteria—extramusical and nonmusical parameters in addition to errors in measurement that influence all forms of evaluation. So, while adjudicators may *think* they are evaluating the "musical value" of a performance, they are also necessarily including other factors in their judgment. Figure 4.2 depicts this broader perspective. In this figure, the three areas indicated by dashed lines on the left side of the diagram refer to factors that may be beyond the assessor's control. The lines flowing from left to right indicate that the adjudication is not simply an assessment of musical value. Instead, it is masked by other factors that can, for ease of interpretation here, be categorized as measurement errors, as well as other extramusical and nonmusical factors. Extramusical factors overlap musical value because it is a fuzzy component, with some aspects of it lying outside the boundary of the assessor's aim of recording musical value. Each component of this model is explicated in the four sections which follow.

## 4.2 Musical value

Much assessment literature tacitly assumes that it is possible to access the true value of a behavior, a view that is also evident when inspecting the criteria traditionally used to assess musical performances. Undoubtedly, aiming to enhance the value of a musical performance is an essential, though not singular, goal of achieving successful assessment, but musical evaluation in all practical terms is far more complicated.

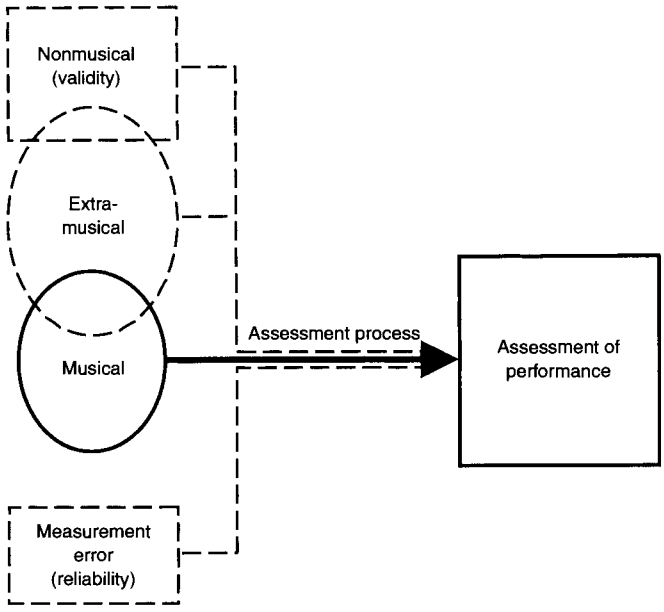



Figure 4.2 Expanded model of music assessment process.

The published literature on the criteria used to assess performances suggests that there are at least four types of competencies that are typically used by music institutions, from which appropriate performance assessment criteria are devised (see, for example, Hollis, 2001a, 2001b; Hunter & Russ, 1996; Hunter, 1999; Palatine Report, 2000; Stanley *et al.*, 2002). These can be summarized as follows:

- Technique

  *Physiological*

  —breathing

  —posture

  —relaxation—tension

  —balance

  —coordination

  *Physical*

  —sound: production, projection, and control of instrument/voice and consistency, clarity, and focus of tone across all registers and dynamic levels

  —range

  —intonation

  —physical control (e.g. stamina, endurance)

—bodily coordination

*Instrumental*

—ensemble coordination, balance, and cohesion

—accuracy, assuredness, and facility of rhythm, pitch, articulations, dynamics, timing, as well as the degree to which errors undermine and detract from the overall quality of the performance

—pacing of performance

—sensitivity to intonation, both individual and ensemble

- Interpretation
  - —authenticity: understanding of the style/genre and established performance practice (e.g. use of a reliable edition)
  - —accuracy: based on faithful reading and/or memorization of the score, and realization and exploration of the composer's intentions
  - —musical coherence: perceptive choice of tempo, phrase shaping, dynamic shadings, sense of line, and understanding of overall structure
- Expression
  - —understanding of the emotional character of the work
  - —projection of mood and character of the work
  - —communication of the structural high points and turning points of the work
  - —sensitivity to the relationship between parts within a texture
  - —appropriate use of tone and color, light and shade, and/or drama
- Communication
  - —among the members of the ensemble (e.g. listening and leadership)
  - —confidence, as demonstrated in performances that are both convincing and purposeful
  - —interest, in terms of the degree to which the performer holds the audience's attention, maintains a sense of direction, creates a sense of occasion, and ends the work convincingly
  - —projection of expressive, interpretative, and structural features of the composition performed

The above assessment criteria contain a mixture of categories that could be classed as either skill or artistry. A commonly discussed difficulty in assessment is the issue of marking interpretation. If it is subjective, then any assessment of interpretation can be no more than unreliable speculation. However, over the past decade, music psychologists have begun to clarify the microstructural changes that affect expressive playing, and there now exists a substantial set of "rules" of expression that performers can incorporate into their playing that help them convey their expressive intentions to their audiences (see Chapter 13 of this volume). Another difficulty is the inevitable

overlap between constituent elements of technique, interpretation, expression, and communication. For example, as an evaluative dimension, the "projection of mood and character of the work," classified above as expressive, could be seen to overlap with the "projection of expressive, stylistic, and structural features of the composition performed," as in the communication category.

## 4.3 Measurement error

The process of assessing musical performances is often based on several implicit and often flawed assumptions. One is that the musical value of a performance can be assessed accurately and reliably. A second is that experienced listeners are able to make consistently accurate judgments. This leads to a third assumption, that expert judges possess the ability to make finer discriminations than average listeners due to their more refined abilities to determine which of the components of a performance were effective and which were not. This situation is analogous with using a microscope to examine a chemical sample. Most people have access to a basic microscope, but the adjudicator has an especially powerful one and can, therefore, see more (see Thompson & Williamon, 2003, for further discussion of these assumptions). However, in all fields of human pursuit where fine esthetic discriminations are made, we can never know everything about the thing we wish to measure. Hence, all measurement is subject to some kind of coloration by error, which leads to the serious implication that it is impossible to judge with perfect accuracy and, even worse, fairness.

The most common kind of error found in assessment is measurement error. In a statistical sense, measurement errors do not produce a systematic bias, but consist of more or less random fluctuation that cannot be easily controlled. Measurement error has important implications in music assessment; it means that a judge cannot be expected to behave in an idealized, machine-like way because there is always likely to be some kind of unbiased, random fluctuation. For example, Fiske (1978, 1983) and Manturzewska (cited in Davidson & Coimbra, 2001) both asked evaluators to judge performances, some of which were the same, and found very low correlations between scores. For this reason, statistical measurement error corresponds to what psychometric researchers call *reliability*, or the consistency and accuracy of a measurement. In this chapter, little attention is given to the manipulation of measurement error for the purpose of enhancing performance assessment because, by definition, measurement error cannot be manipulated to favor one performance over another in a systematic manner. Measurement error could be minimized, but this would lead to fairer assessment across all performances, regardless of the strategies used.

## 4.4 Extramusical factors

Extramusical factors are an unclearly defined, fuzzy set that may belong to true musical value or to nonmusical aspects of performance assessment. The location of these factors are largely subjective or dependent on circumstances. Some selected extramusical factors are reviewed below. Although this is not a definitive list, the labeling of factors as such is useful in parsing the assessment process into smaller, more manageable

units. This area is complex and controversial, however, because the many items that could be grouped in this category are considered by some as essential parts of musical value and by others as something that only produces messiness in the evaluation, as it introduces biases that are not related to the essential formal properties of a musical performance. In some cases, the issues mentioned may seem more relevant to measurement error. The decision to include such items under extramusical factors was guided by the criterion that the performer might be able to use some knowledge about the factor to systematically enhance their performance assessment.

While there are clearly marked boundaries between such aspects as interpretation and technique, which are part of the performance, it is often less clear whether stage presence, individuality, and flair of the performer should have an influence on evaluation. It is these underlying factors that are included in the definition of extramusical used here. The existing literature suggests that categories related to the performer, the context in which the performance takes place, and the evaluator help explain the types of extramusical factors that can potentially impact on the evaluation of a musical performance. The following sections are, therefore, divided into three sources of extramusical assessment enhancement:

- those that the player can directly control (including less obvious issues such as self-efficacy and cognitive mediation)
- those that depend on the playing context
- those that require research about the adjudicator.

Understanding these can be useful in knowing how to maximize one's performance assessment.

### 4.4.1 Performer-related aspects

The following factors are those over which the performer can exercise direct control to enhance their performance assessment. The internal thoughts and mechanisms related to a musician's preparation for the performance are examined first, followed by an explanation of research on expressive variations from the norm, attractiveness, flair, and movement.

#### *Self-efficacy*

How individuals control and monitor their own cognition before, during, and after engaging in a challenging activity provides a powerful insight into how well they are able to perform. One of the most important cognitive processes is self-efficacy, which refers to the perception or judgments a person holds for being able to perform a certain task successfully (Bandura, 1997; see also Chapter 2 of this volume). As an example, opera singers will "differ in their perceived efficacy to fulfill the vocal, emotive, and theatrical aspects of their artistic craft and to fuse them into dramatic performances" (Bandura, 1997, p. 36), because their self-beliefs are linked to distinct realms of functioning that influence not only how they perform, but their thought processes, motivation, and affective and physiological states as well.

Self-efficacy has been shown to be the strongest predictor of a musician's performance in assessment situations, such as a graded music examination (McCormick & McPherson, 2003). The extensive literature on this topic in academic areas demonstrates that students who display slightly higher perceptions of their own ability to accomplish a task than is justified by their actual ability, are more likely to choose more challenging tasks, exert more effort, persist longer, and be less likely to experience debilitating anxiety (Bandura, 1986; Pajares, 1996; Zimmerman, 2000). In contrast, even accurate self-perceptions can lower optimism, which then affects effort, persistence, and perseverance (Pajares, 1996). All of this reinforces the importance of training the mind, as well as the body, because how musicians feel about their own ability and level of performance has a powerful effect on how they project themselves to their audience and, as a result, how their audience reacts to their performance.

*Expressive variations*

Expressive variation in performance is an extramusical factor directly related to interpretation, and here it is referred to as deviations from the expressive norm. Contrary to common belief, it is possible to measure expression empirically (see Chapter 13), and also to establish an expressive norm that relates to the typical amount of deviation from a musical score that is considered necessary for expressive playing. Layered upon this norm are deviations from it to obtain greater expression, and this deviation from the norm is classified here as extramusical. The music expression literature has identified some of these deviations (e.g. Repp, 1997; Woody, 2000), but there is controversy as to how and when to apply them. For example, Repp (1997, 1999) reports that student performers tend to exhibit fewer deviations from the expressive norm for a given piece than famous performers: "expert pianists' performances are much more diverse in timing than student performances, which tend to stay close to the norm" (Repp, 1999, p. 1983). In line with Woody's (2000) report regarding the locations at which musically appropriate deviations are noticeable, Schubert (2002) argues that there exists a quantifiable boundary of acceptable deviation in expression, where deviations too far beyond these hypothetical limits produce dull, awkward, inappropriate, or even comical-sounding effects. This boundary theory suggests that students can maximize the impact of their performances by deliberately varying expressive content from the norm, even though culturally specific performance conventions, expectations, and standards may dictate that they need to stay within well-defined boundaries.

*Attractiveness and flair*

Social psychologists have known for some time that a person's physical appearance affects the impression they project to others (Myers, 1999), and so it is of no surprise that performers' physical *projections* of themselves influence how their audiences perceive them and their performance. For example, audiences who are asked to provide written comments on a performance will often start by referring to the attractiveness of the performer (Wapnick *et al.*, 1997; Davidson & Coimbra, 2001; Stanley *et al.*, 2002). This seems especially true for singers, with studies showing that maintaining eye contact and smiling at appropriate times in the performance influence how they

are rated by judges (Bermingham, 2000; Wapnick *et al.*, 1997, 1998, 2000). But it is important to note that this dimension acts along a continuum, such that the stage presence required to bring off a performance of a solemn piece will be quite different from that required to perform a piece that is joyful and lively. Expert performers know that it is all a matter of degree and appropriateness to the repertory being performed and the type of audience that is listening.

*Movement*

Some estimates suggest that vision accounts for more than 75% of all information learned (Gladney, 1997; Long, 1997), and as the dominant sense, it can inhibit the processing of information from other modalities (Posner *et al.*, 1976; Smyth, 1984). In terms of the visual component of a musical performance, physical movements and gestures provide important expressive information about a musician's intentions and, thereby, help an audience to judge the interpretation and "musicality" of a performance (Davidson, 1993, 1994, 1995, 2001). This is not surprising, given how closely physical movements are linked to and driven by the mental representations of the music being performed (Lehmann & Davidson, 2002).

The types of visual cues that influence an audience include the actual quantity of the performer's movements, as well as specific gestures that are an integral part of a performer's way of expressing specific musical intentions (Davidson, 2002). A key point is that the context is all-important; better performers are able to achieve a "balance in the articulation of musical expression, coordination, and virtuosic display within the social convention" (Lehmann & Davidson, 2002, p. 554). In this way, communication is a two-way process in which performers:

> are just as likely to pick up information from the audience as audiences are likely to pick up information from the performer. Not only do general behaviors such as applause provide the performer with feedback about their performance, but moment by moment, smiles, frowns, laughs, coughs, etc. can all have an effect on the performer (Davidson, 1997, p. 224).

### 4.4.2 Context related aspects

While not directly under the control of the performer, there are several aspects of the performer's environment that can be used to enhance one's assessment. The communication of players within an ensemble, the acoustics, social elements, and the audience response are included in this subcategory as examples of extramusical factors.

*Communication within the ensemble*

An extension to the processes involved in attending to and concentrating on the music being performed is the way the performer communicates with the other musicians. Again, evidence suggests that these social interactions are extremely important. The way performers maintain eye contact and use nonverbal information such as gestures and body movement to converse "musically" helps their coperformers and their audience discern finely grained cues about their musical and expressive intentions (Davidson, 1993, 2001; Williamon & Davidson, 2002). This is especially evident for nonmusicians, who are more reliant on visual rather than aural cues to discern the extent to which a musician is playing with or without expression (Davidson, 2001).

Highly competent accompanying musicians affect an audience's overall impression of a performance, not only because they are more likely to provide an accurate reading of their own part and interact at a higher level with the soloist, but because the presence of superior coperformers acts to heighten or "lift" a musician to perform closer to his or her peak. This type of enhancement has been reported by psychologists for the past few decades: from Zajonc (1965) who proposed that the presence of others raises a person's general arousal level and can affect task performance, to investigations by Seta (1982), who explains that feedback from a superior co-actor can alter an individual's performance because it serves as a cue for competition and as a source of motivation to achieve at higher levels (see also Bryant & Driskell, 1997).

Some of the more illuminating research on this topic has been undertaken by Williamon and Davidson (2002) who, in one of their studies, explored social and nonverbal communication between two expert pianists who prepared and then presented a recital of duet music. Results show that the pianists used a rich variety of visually based information—such as eye-contact, facial expressions, physical gestures, and swaying movements—to communicate specific musical ideas while they performed. These forms of visually based feedback help performers coordinate their playing (Clayton, 1985) and are, therefore, an integral part of effective interaction between musicians (Yarbrough, 1975). They enable musicians to converse musically (see Chapter 6 for further discussion of ensemble communication).

*Acoustics*

There is a significant literature on how room acoustics affect sound quality (e.g. Naylor, 1988; Gade, 1989a, 1989b; Meyer, 1993) but virtually no mention of it in performance assessment literature. It is self-evident that room acoustics may serve one instrument over another, and perhaps even one performance technique over another. Moreover, the size, shape, and design of a room all have a bearing on the sound an instrument can produce, in addition to the position at which the assessor sits (Ando, 1988). Wolfe (2003), an expert in acoustics, makes suggestions of which many musicians are aware, such as playing softer in an auditorium with little background noise and playing slower with more separation between the notes in a larger, reverberant auditorium. While a performer may not necessarily be interested in becoming an expert in acoustics, there are some commonsense practices that may benefit the well prepared performer:

- attending other performances in the concert venue, especially those performances featuring the same instrument or ensemble combination and taking notes on tone quality and how sitting in different parts of the auditorium affect the sound
- practicing in the venue beforehand if possible, remembering that the acoustics of an auditorium are dramatically altered when there is a shift of people and equipment
- being prepared to compensate and adjust for any weaknesses caused by the acoustic during the performance (see Ando, 1988).

Working on this extramusical aspect of performance can potentially enhance the assessment one receives under a musical category such as tonal production, and as such, the contribution made to tone quality by the acoustic environment is an important consideration.

*Social factors*

The context in which a performance will take place can be thought of as representing yet another facet of extramusical influences. First, the *purpose of the assessment*—a music competition, concert, festival, end of semester recital/examination, audition, or public concert—can influence the way an evaluator or music critic will listen to, and therefore, evaluate a performance (McPherson & Thompson, 1998). Second, the *type of performance*—whether the musician is performing from memory, with music, by ear, or by improvising—can affect not only the criteria relevant to how the performance should be assessed but also how judgments will be formed. Some instruments are considered harder to play than others, or may have different technical skills and repertory, all of which can potentially impact on the assessment process.

Clearly, the way a performance will be perceived by an audience depends on a number of complex social factors (Davidson, 1997). Social etiquette in performances of all types of music operates via a two-way process between what the audience expects from the performer, and what the performer expects from the audience. So, actively integrating the cultural values—such as dress codes and stage manner—is an important means of maximizing one's performance assessment. For example, in art music, performances from memory have been shown to allow players to perform more freely and expressively, as well as to communicate their expressive intentions more effectively to their audience (Aiello & Williamon, 2002; see also Chapter 7 of this volume).

*Audience support*

Another interesting factor in assessment is the audience support for the performer. This may appear to be a nonmusical factor, but audience response can enhance the perception of a performance, especially in situations where greater audience support creates a sense of excitement that encourages us to believe that a performance is better than its true musical value. It is for this reason that the factor of audience support is positioned in the extramusical category. While no studies have been cited in the empirical music assessment literature, crowd noise has been shown to have a significant effect on judges' decisions in various sports, where distinct advantages have been found in favor of the home team. Nevill and Holder (1999) argue that referees make decisions that are influenced by the audience. So, the more "noise" and auditory support (clapping, cheering, etc.) the audience gives a performance, the more positive the adjudication may be. As with many extramusical factors, adjudicators may require training and experience to minimize the unfair influence this may have, but as mentioned, it is conceivable that audience support may enhance a performance in this extramusical way.

### 4.4.3 Evaluator characteristics

The ways in which individuals draw a conclusion based on available evidence can be quite different, depending on a number of personal factors. For this reason, additional extramusical factors relevant to the evaluator (or listener) need to be recognized. Below, a selection of evaluator-related aspects of assessment are discussed that can be understood and controlled by obtaining reliable information about the characteristics of the assessor(s). If it is possible and ethical, collecting such information may help the

performer to enhance his or her performance. Several factors seem to be relevant here, which follow the path from knowledge about the performer before the performance, to first impressions of the performance, to the mood of the adjudicator at the time of the assessment, and finally, to the adjudicator's overall familiarity with and preferences for the literature being performed.

*Memory influences*

Memory-influenced biases occur when individuals are primed with information or have expectations about someone that subsequently color their expectations about how that person will react or perform (Fyock & Stangor, 1994). In music, Radocy (1976) provides clear evidence that listeners' impressions are influenced by what they expect to hear from a performer and that prior knowledge or expectations can exert a major influence on their assessment of a performance. He assigned undergraduate students to one of five conditions where bias was manipulated by introducing false information about performers before the students made their evaluations. Students in the moderate bias conditions were given misleading information, such as being told that the performer was a "former symphony player" or "young graduate assistant." Students in the strong bias conditions were given misleading information about the player and reasons why the performances by the professional performer were supposedly preferred by prior listeners. Results showed overall biases in the undergraduate students' evaluations, in that they preferred the performances by musicians they thought were professionals or more experienced. Interestingly, however, Radocy reports that some instruments (e.g. piano) were more susceptible to these biases than others (e.g. trumpet). Unfortunately, there is very little research on how susceptible highly experienced adjudicators are to such memory priming effects.

*First impressions*

First impressions are often more lasting than we might think. In music, evidence shows that rankings of performers are not markedly different when evaluators hear entire performances as compared to only the first sections of the performances (Vasil, 1973) and that expert musicians are able to form a relatively quick global assessment very early in a performance (Stanley *et al.*, 2002). For example, in a study at the Sydney Conservatorium of Music, one member of the staff stated: "I look at them and I say 'Distinction, high Credit'. I have bands in my own mind and then the number is immaterial—to me the number is way more negotiable than the actual range" (Stanley *et al.*, p. 47). This examiner then moved to a process of justifying his immediate reaction, by stating

> This is why I feel this way. . . . I may find out that all that appealed to me was the person looked smart or behaved in a good way and then I would say: "Well the presentation was so spectacular that I forgot to listen to the playing, but now when I come to think of it the playing was not all that really good" (p. 47).

Studies in social cognition provide an explanation for this fairly typical, though complex, response. According to this line of research, strong first impressions we have of a person persevere and color our subsequent judgments about that person. Initial opinions affect how we will draw in (i.e. adapt) or resist drawing in (i.e. isolate) subsequent

information to form an extant impression (Ybarra, 2001). Initial *positive* impressions tend to exert a different influence on subsequent judgments than initial *negative* impressions. For example, if a performer creates a good initial impression at the start of his or her program but fails to sustain this level of performance in subsequent parts of the performance, a judge will be more capable of integrating new information (in this case, a poorer level of performance) into his or her final judgment. In contrast, an initial *negative* impression appears to exert a more stable, long-term influence, in that a judge will be less inclined to include new information (e.g. better playing) as the basis of changing his or her overall impression. This asymmetry emerges because people find it easier to infer dispositional characteristics from negative behavior rather than positive behavior. Negative information about a person is more likely to result in a more stable evaluation over time than positive information (Carston, 1980), because negative first impressions make it harder for a person to *isolate* their initial impressions from information that occurs subsequently.

*Mood of the assessor*

The mood of the assessor probably has some effect on his or her adjudication, and if it does, there exists ample literature demonstrating how different kinds of music can evoke different moods that vary in degrees of valence (happiness and sadness) and arousal (activity and sleepiness; Gorn *et al.*, 2001), as well as how a listener's mood can be altered by listening to a particular piece of music (Västfjäll, 2001–2002). An important part of this dimension for performers to consider is the principal of habituation, which suggests that listeners will become bored when hearing a series of pieces that do not vary sufficiently in interpretative or expressive intent (Schubert, 1996). It is, therefore, of paramount importance for musicians to strike a balance between the *intensity* and *variety* of emotions expressed in the repertory they are performing, in order to optimize the pleasure their listeners derive from it, thereby enhancing their own assessment.

*Familiarity and preference*

Judges often find it more difficult to assess a work that they have never heard before than a work that is highly familiar to them. Flores and Ginsburgh (1996) studied whether judges are more severe in adjudications of works heard for the first time by analysing results over a number of decades from the Queen Elisabeth music competition in Belgium. Their study provides convincing evidence that judges tend to be more severe for the first few performances of a new commissioned work than for later performances after they have had a chance to form an impression of how difficult the work is to perform and are, consequently, more forgiving of problems in the performance. In this way, memories of literature being performed—whether formed immediately before the performance such as in a music competition with a newly commissioned work or long-term as a result of previous exposure—have the potential to exert a major influence on a judge's assessment of the performer's efforts.

From another perspective, this can be explained by applying theories from the literature on preference, of which the most well known is the inverted-U curve that relates preference to massed exposure. Massed exposure is equivalent to a listener hearing the same piece many times in sequence (see Martindale, 1988; North & Hargreaves, 2000).

The inverted-U theory suggests that preference for the new work will start to rise with familiarity and then decay as the listener becomes bored from overexposure. In this way, a preference peak will occur after the first performance, which confirms the type of bias reported in the study by Flores and Ginsburgh (1996).

*Halo effect*

The halo effect arises from a tendency of a judge to be unduly influenced by a single factor, such as a particular strength or weakness in a person's performance, their physical appearance, or another aspect of their behavior (Anastasi & Urbina, 1997). This often occurs in situations where a person's performance is particularly high on a factor that an assessor personally feels is important, which then unduly influences them to inflate their assessment on other attributes of the person's performance. The phenomenon has received much attention in areas such as marketing (e.g. Leuthesser *et al.*, 1995), gymnastics (Ansorge & Scheer, 1988), and management (e.g. Wirtz & Bateson, 1995), but little is understood about its influence and how it might be controlled in music performance.

The most obvious type of halo effect in music is where a judge will draw on previous knowledge of a performer as being a very good or very poor musician, and incorrectly or unjustly inflate or decrease their assessment rating, even though a particular performance may be better or worse than would normally be achieved. There is evidence that knowing the performer can lead to "in group" and "out group" biases, such that assessors give more generous marks to performers they know (Manturzewska, 1970, cited in Davidson & Coimbra, 2001). For example, Hunter and Russ (1996) found that unrealistic marks were sometimes awarded for musicians based on reputation, particularly in situations "when a performer has a high reputation and there is a buzz of expectancy when they enter the hall" (p. 71; see also Hunter, 1999). In such instances, this type of bias can occur at different levels, from knowing the performer personally to having heard the performer play on previous occasions. A professional relationship of this sort does not necessarily mean a bias will occur on all occasions—only that the likelihood increases.

In other areas, techniques are used to reduce the effect (Bagozzi, 1996), and there is evidence that training might be essential to alert the adjudicator of this subconscious bias. Given that the halo effect can influence many aspects of a performance and that the aim in adjudications should be to minimize this, it is perhaps better to view it as a measurement error rather than an extramusical factor. It is treated as an extramusical factor here as it could conceivably be used to enhance one's own performance assessment. For example, a musician who has a noticeable strong point (especially one that an adjudicator is known to appreciate), should consider how to maintain and exploit this aspect of performance, and perhaps develop it through the repertory he or she performs.

## 4.5 Nonmusical factors

Nonmusical factors are defined as those related to *validity*—that is, whether evaluators are actually assessing what they think they are assessing. Because nonmusical factors produce unfair biases, it is important that educators, adjudicators, and researchers work

toward understanding them. However, from the performer's perspective, apart from being aware of their existence, there is little they can, or should have to do to manipulate these factors in order to enhance their performance assessment. Only two kinds of nonmusical factors are discussed here, but several issues raised in the extramusical section could also be listed as nonmusical, such as attractiveness and audience support.

### 4.5.1 Stereotyping

Stereotyping based on gender or race has the potential to impact on assessment in subtle ways (McCrary, 1993; Green, 1997; O'Neill, 1997; Morrison, 1998; Bermingham, 2000; Dibben, 2002). This was shown most convincingly in a study by Elliott (1995), who used videotaped performances of four trumpet players and four flute players. Each instrument had a male/female and black/white performer. Elliott's master videotape of the eight performances controlled the soundtracks for each of the four trumpet and flute performances so that the musicians pretended to coordinate their playing over the same sound recording for each instrument. Undergraduate and graduate music specialists who rated the performances were unaware that the soundtrack for each instrument was identical, but still consistently rated the performances according to both dimensions. Black performers scored significantly lower than their white peers, with the female trumpeters scoring lower than female flautists. Other studies have shown that people sharing the race of the judge tend to be awarded higher marks (Green, 1997), and that the higher percentage of women now playing in professional orchestras is a direct result of the introduction of "blind" auditions where the auditioning panel sits behind a screen and cannot see the performer (Goldin & Rouse, 2000).

### 4.5.2 Order of performance

Order is known to affect judgments. In wine tasting, there is a bias that favors the first sample tasted (Filipello, 1956), and in gymnastics, the fourth position within a team tends to receive the more advantageous adjudication (Scheer & Ansorge, 1975; Plessner, 1999). In international music competitions, order biases have also been observed. For example, Flores and Ginsburgh (1996) report that rankings of performers in the Queen Elisabeth competition were significantly related to the order in which they performed. Performers near the beginning had a statistically lower chance of being ranked highly in comparison with those who performed near the final day, who tended to be ranked in the top category by the adjudicators. For international competitions, Flores and Ginsburgh suggest that repeated hearings of the same piece may nurture a greater appreciation, which in turn casts a more positive light on the performance itself (see also "Familiarity and preference of repertory" above). Another explanation is that adjudicators tend to start with "higher expectations and more strict rules, and then progressively adapt them to the reality of the actual performance" (Flores & Ginsburgh, p. 102; see also McPherson & Thompson, 1998).

## 4.6 A model with practical implications: The Johari window

So far, a variety of issues have been presented that can or do affect performance assessment. Each issue varies in relevance to different performers, depending on their age,

expertise, mood, self-efficacy, and so on. Furthermore, several adjudicator factors, such as experience, training, and mood will determine how much these issues influence assessment. This final section draws together the various elements of the assessment model as a means of helping musicians understand which aspects of assessment most concern them and, consequently, which aspects can be improved to maximize their chances of enhancing their own assessment.

The "Johari window" is the framework for this new model. The Johari window is a model of awareness of behavior and motivation (Luft, 1969), used primarily in psychodynamic therapy, but it has received a wide variety of applications and continues to attract interest (e.g. Afolabi, 1993; Sole, 1997). The model proposes that, when interacting with others, "awareness" can be divided into four areas:

- *public area*: an individual will be aware of some behaviors and motivations that are also noticeable to others.
- *blind area*: some behaviors and motivations will be inaccessible to the individual but accessible to others (hence, this is like a blind spot for the individual).
- *secret area*: the individual will hide certain motivations and behaviors, and therefore this is the secret part of the model.
- *hidden area*: there is a part of behavior and motivation of which neither the individual, nor others, are aware.

The model is represented as a window with four quadrants. In this application of the Johari window, "others" are the person (or persons) who makes the assessment, such as an adjudicator or an audience. Consequently, it is suggested that performers organize their personal windows according to what they are aware of themselves, and to what they believe the adjudicator to be aware of that is relevant to their performance. An example of some of the issues discussed above could be sorted into a typical, but hypothetical, individual's Johari window (see Figure 4.3).

From the example shown, it becomes clear how the general principles of our model apply to individuals. The leitmotif of the present review is the questioning of the assumption that musical value can be measured. From the Johari window perspective, this translates to musical value being in the top left quadrant, where this is known by both the performer and the assessor. However, it has been argued that the assessor may not be able to access the musical value of a performance, and so only certain aspects may reveal themselves in this quadrant. Measurement error, on the other hand, falls into the bottom right quadrant where assessment factors are hidden from both assessor and performer. This hidden region appears to be a general location where *reliability* of assessment is threatened. The secret and blind areas appear to mark regions where assessment *validity* is threatened. Various other factors in performance can be sorted into the four panes. No two individuals will have identical windows. For the hypothetical relationship shown, several factors have been added that are not mentioned in this chapter because literature could not be cited that specifically describes how they can be used to enhance a player's assessment. For instance, preparation and effort have been added, speculatively, to the secret pane. As an exercise, performers should attempt to produce their own personal Johari window and sort the various factors discussed in this chapter, and elsewhere in the book, into the appropriate panes.

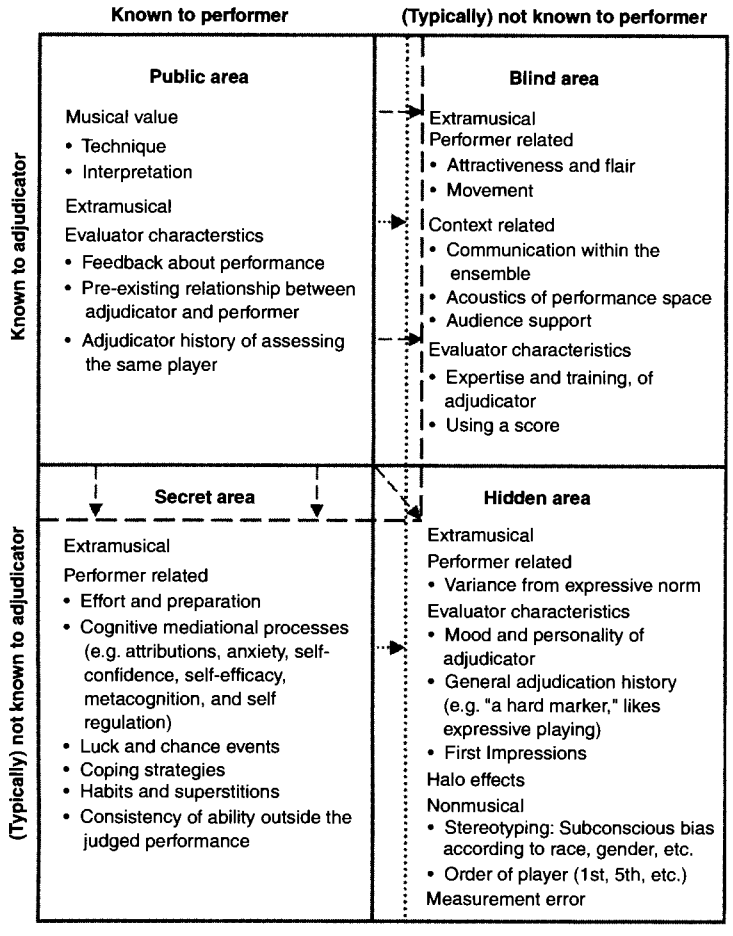



**Figure 4.3** Example of the Johari window applied to an imaginary individual and adjudicator awareness relationship. The dotted line and arrows indicate the direction in which the individual should aim the quadrant shapes—that is, to increase the left side and reduce the right side. The dashed line and arrows indicate the traditional direction in which the Johari window should be changed—that is, to make the public quadrant larger and all other quadrants smaller.

Apart from organizing assessment and performance related issues, the window has an important application. In general, performers should strive to reduce the area of the right quadrants of the window, and widen the left quadrants. That is, individuals will be more likely to enhance their performance if they know as much about their performance-controlling factors as possible. For example, if the adjudicator is not aware of the halo effect, and the performer becomes aware of how a particular extramusical factor might enhance his or her performance (of which he or she was not previously aware), then an item from the right side of the window will shift to the bottom left pane, in effect expanding the left side of the window: the aim of the model from the

performer's perspective. Furthermore, suppose the hypothetical performer plays with insufficient expressive movement and is penalized as a result. If the adjudicator, teacher, or colleague alerts the performer that this is a performance problem, the player can then work on developing this extramusical factor and hence shift it into the top left (public) quadrant. The window also explains how an adjudicator may vary the criterion for individuality and expression by deducting marks for an overly expressive, though appropriate, performance. In this case, the performer's blind window-pane has been enlarged, because the adjudicator knows that the criterion has been changed but the performer does not.

The window also helps to allow recognition that some things might never be moved from the right to the left, particularly some items in the bottom right (hidden) quadrant, because no one has access to these factors. The main example is measurement error. It may be possible, however, for the performer to develop some insights into the bottom right corner of the window and to be able to shift these into one of the left panes. For instance, it is vital that stereotyping issues are brought to the adjudicator's awareness so that the impact of such unfair biases can be minimized or eliminated. Importantly, researchers and educators need to provide ways of enabling adjudicators to shift the bottom two panes up to the top two, so that the adjudicator's reliability and validity can improve via appropriate assessment recording techniques.

## 4.7 Conclusions

The present chapter has presented a framework to explain the complexities of the assessment process that are not commonly understood or considered by musicians. In the sampling of literature discussed, it is evident that extramusical factors of a performance are a largely untapped source for performance enhancement because performers, adjudicators, and in many cases even music researchers are unaware of the major impact these could and do have upon the assessment of a performance. This does not suggest that performers do not need to work on the traditional elements associated with musical value. The chapters that follow provide methods for developing these features, but for players of a very high standard, this approach may give them an edge. By the same token, apart from measurement error, it is not implied that all adjudicators are prone to making the various kinds of errors that allow performers to use some less traditional means of enhancing their assessment. However, it does need to be acknowledged that an "evaluation of a performer does not mean anything until we know how reliable the judge was who evaluated that performance" (Fiske, 1994, p. 76). Like learning how to play proficiently, learning how to provide reliable and accurate assessments of musical performance is an acquired skill that requires reflection and practice for any degree of success. While more research is needed, it is acknowledged that there will be some adjudicators who have a very deep understanding of the process of adjudication. The wider adjudicating community could be well served by discovering and applying the knowledge such individuals possess.

Understanding the process of assessment is a key to enhancing one's performance beyond the traditional mantra of "practice makes perfect." The expanded theory outlined here, therefore, presents a perspective that crystallizes many of the problems that

have previously been hidden beneath the general guise of assessment. While many issues await investigation by researchers to ensure that fair, unbiased, and reliable assessment regimes are provided, performers can maximize their reception by applying the theory to their own personal Johari window; as research develops, so will the nature and shape of the window. The main point, however, is that performers should organize their window in such a way as to enhance their assessment, and therefore further enhance their performance. Such reflective practices have considerable potential in the area of performance enhancement.

## Acknowledgments

We thank James Renwick and two anonymous referees who provided very helpful suggestions and insightful comments on an earlier version of this chapter, plus Aaron Williamon, whose attention to detail and critical mind helped at all stages of the process.

## Further information and reading

Hollis, E. (2001a). *The Guildhall School's* clear *Performance Assessment System: How* clear *Works.* London: Guildhall School of Music and Drama Publications.

Hollis, E. (2001b). *The Guildhall School's* clear *Performance Assessment System: Marking Schemes for the Assessment Categories.* London: Guildhall School of Music and Drama Publications.

Luft, J. (1969). *Group Processes: An Introduction to Group Dynamics.* Bethesda, MD: National Press Books.

McPherson, G. E., & Thompson, W. F. (1998). Assessing music performances: Issues and influences. *Research Studies in Music Education, 10,* 12–24.

Palatine Report (2000). *Palatine Learning and Teaching Support Network.* Retrieved August 28, 2003, from Palatine website: *www.lancs.ac.uk/users/palatine/report8nov.htm.*

## References

Afolabi, M. (1993). Application of JOHARI Communication Awareness Model to special libraries management. *Library Management, 14,* 24–27.

Aiello, R., & Williamon, A. (2002). Memory. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 167–181). Oxford: Oxford University Press.

Anastasi, A., & Urbina, S. (1997). *Psychological Testing* (7th ed.). Englewood Cliffs, NJ: Prentice Hall.

Ando, Y. (1988). *Architectural Acoustics: Blending Sound Sources, Sound Fields, and Listeners.* New York: Springer.

Ansorge, C. J., & Scheer, J. K. (1988). International bias detected in judging gymnastic competition at the 1984 Olympic Games. *Research Quarterly for Exercise and Sport, 59,* 103–107.

Bagozzi, R. P. (1996). The role of arousal in the creation and control of the halo effect in attitude models. *Psychology & Marketing, 13,* 235–264.

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control.* New York: Freeman.

Bermingham, G. A. (2000). Effects of performers' external characteristics on performance evaluations. *Update: Applications of Research in Music Education, 18,* 3–7.

Bryant, B. M., & Driskell, J. E. (1997). Presence of others and arousal: An integration. *Group Dynamics: Theory, Research, and Practice, 1,* 52–64.

Carston, D. E. (1980). The recall and use of traits and events in social inference processes. *Journal of Experimental Social Psychology, 16,* 303–328.

Clayton, A. M. H. (1985). *Coordination between Players in Musical Performance.* Unpublished doctoral dissertation, University of Edinburgh, Edinburgh, UK.

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music, 21,* 103–113.

Davidson, J. W. (1994). Which areas of a pianist's body convey information about expressive intention to an audience? *Journal of Human Movement Studies, 26,* 279–301.

Davidson, J. W. (1995). What does the visual information contained in music performances offer the observer? Some preliminary thoughts. In R. Steinberg (Ed.), *The Music Machine: Psychophysiology and Psychopathology of the Sense of Music* (pp. 105–113). Berlin: Springer-Verlag.

Davidson, J. W. (1997). The social in musical performance. In D. J. Hargreaves & A. C. North (Eds.), *The Social Psychology of Music* (pp. 209–228). Oxford: Oxford University Press.

Davidson, J. W. (2001). The role of the body in the production and perception of solo vocal performance: A case study of Annie Lennox. *Musicæ Scientiæ, 5,* 235–256.

Davidson, J. W. (2002). Understanding the expressive movements of a solo pianist. *Deutsches Jahrbuch für Musikpsychölogie, 16,* 9–31.

Davidson, J. W., & Coimbra, D. D. C. (2001). Investigating performance evaluation by assessors of singers in a music college setting. *Musicæ Scientiæ, 5,* 33–53.

Dibben, N. (2002). Gender identity and music. In R. R. MacDonald, D. Hargreaves, & D. Miell (Eds.), *Musical Identities* (pp. 117–133). Oxford: Oxford University Press.

Elliott, C. A. (1995). Race and gender as factors in judgments of musical performance. *Bulletin of the Council for Research in Music Education, 127,* 50–56.

Filipello, F. (1956). Factors in the analysis of mass panel wine-preference data. *Food Technology, July,* 321–326.

Fiske, H. E. (1978). *The Effect of a Training Procedure in Music Performance Evaluation on Judge Reliability.* Toronto, Canada: Ontario Educational Research Council.

Fiske, H. E. (1983). Judging musical performances: Method or madness? *Update: Applications of Research in Music Education, 1,* 7–10.

Fiske, H. E. (1994). Evaluation of vocal performances: Experimental research evidence. In G. Welch & T. Murao (Eds.), *Onchi and Singing Development* (pp. 74–103). London: David Fulton Publishers.

Flores, R. G., & Ginsburgh, V. A. (1996). The Queen Elisabeth musical competition: How fair is the final ranking? *The Statistician, 45,* 97–104.

Fyock, J., & Stangor, C. (1994). The role of memory biases in stereotype maintenance. *British Journal of Social Psychology, 33,* 331–343.

Gade, A. C. (1989a). Investigations of musicians' room acoustic conditions in concert halls: I. Methods and laboratory experiments. *Acustica, 69*, 193–203.

Gade, A. C. (1989b). Investigations of musicians' room acoustic conditions in concert halls: II. Field experiments and synthesis of results. *Acustica, 69*, 249–262.

Gladney, G. A. (1997). Impact of visuals in selecting NNA General Excellence winners. *Newspaper Research Journal, 18*, 157–168.

Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Association, 90*, 715–741.

Gorn, G., Pham, M. T., & Sin, L. Y. (2001). When arousal influences evaluation and valence does not (and vice versa). *Journal of Consumer Psychology, 11*, 43–55.

Green, L. (1997). *Music, Gender, Education.* Cambridge: Cambridge University Press.

Hollis, E. (2001a). *The Guildhall School's* clear *Performance Assessment System: How* clear *Works.* London: Guildhall School of Music and Drama Publications.

Hollis, E. (2001b). *The Guildhall School's* clear *Performance Assessment System: Marking Schemes for the Assessment Categories.* London: Guildhall School of Music and Drama Publications.

Hunter, D. (1999). Developing peer-learning programmes in music: Group presentations and peer assessment. *British Journal of Music Education, 16*, 51–63.

Hunter, D., & Russ, M. (1996). Peer assessment in performance studies. *British Journal of Music Education, 13*, 67–78.

Lehmann, A. C., & Davidson, J. W. (2002). Taking an acquired skills perspective on music performance. In R. Colwell & C. Richardson (Eds.), *The New Handbook of Research on Music Teaching and Learning* (pp. 542–560). Oxford: Oxford University Press.

Leuthesser, L., Kohli, C. S., & Harich, K. R. (1995). Brand equity: The halo effect measure. *European Journal of Marketing: Bradford, 29*, 57–66.

Long, K. (1997). *Visual-Aids and Learning.* Retrieved August 28, 2003, from the University of Portsmouth, Department of Mechanical and Manufacturing Engineering website: *www.mech.port.ac.uk/av/AVALearn.htm.*

Luft, J. (1969). *Group Processes: An Introduction to Group Dynamics.* Bethesda, MD: National Press Books.

Martindale, C. (1988). Cognition, psychobiology, and aesthetics. In H. F. Farley & R. W. Neperud (Eds.), *The Foundations of Aesthetics, Art, and Education* (pp. 7–42). New York: Praeger.

McCormick, J., & McPherson, G. E. (2003). The role of self-efficacy in a musical performance examination: An exploratory structural equation analysis. *Psychology of Music, 31*, 37–51.

McCrary, J. (1993). Effects of listeners' and performers' race on music preferences. *Journal of Research in Music Education, 41*, 200–211.

McPherson, G. E., & Thompson, W. F. (1998). Assessing music performances: Issues and influences. *Research Studies in Music Education, 10*, 12–24.

Meyer, J. (1993). The sound of the orchestra. *Journal of the Audio Engineering Society, 41*, 203–213.

Morrison, S. J. (1998). A comparison of preference responses of white and African-American students to musical versus musical/visual stimuli. *Journal of Research in Music Education, 46*, 208–222.

Myers, G. (1999). *Social Psychology* (6th ed.). New York: McGraw-Hill.

Naylor, G. (1988). Modulation transfer and ensemble music performance. *Acustica, 65*, 127–137.

Nevill, A. M., & Holder, R. L. (1999). Home advantage in sport: An overview of studies on the advantage of playing at home. *Sports Medicine, 28*, 221–236.

North, A. C., & Hargreaves, D. J. (2000). Musical preferences during and after relaxation and exercise. *American Journal of Psychology, 113*, 43–67.

O'Neill, S. A. (1997). Gender and music. In D. J. Hargreaves & A. C. North (Eds.), *The Social Psychology of Music* (pp. 46–60). Oxford: Oxford University Press.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*, 543–578.

Palatine Report (2000). *Palatine Learning and Teaching Support Network.* Retrieved August 28, 2003, from Palatine website: *www.lancs.ac.uk/users/palatine/report8nov.htm.*

Plessner, H. (1999). Expectation biases in gymnastics judging. *Journal of Sport & Exercise Psychology, 21*, 131–144.

Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review, 83*, 157–171.

Radocy, R. E. (1976). Effects of authority figure biases on changing judgments of musical events. *Journal of Research in Music Education, 24*, 119–128.

Repp, B. H. (1997). Expressive timing in a Debussy Prelude: A comparison of student and expert pianists. *Musicæ Scientiæ, 1*, 257–268.

Repp, B. H. (1999). A microcosm of musical expression: II. Quantitative analysis of pianists' dynamics in the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America, 105*, 1972–1988.

Scheer, J. K., & Ansorge, C. J. (1975). Effects of naturally induced judges' expectations on the ratings of physical performances. *Research Quarterly, 46*, 463–470.

Schubert, E. (1996). Enjoyment of negative emotions in music: An associative network explanation. *Psychology of Music, 24*, 18–28.

Schubert, E. (2002). Continuous response methodology applied to expressive performance. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the Seventh International Conference on Music Perception and Cognition* (pp. 83–86). Adelaide, Australia: Causal Productions.

Seta, J. J. (1982). The impact of comparison processes on coactors' task performance. *Journal of Personality and Social Psychology, 42*, 281–291.

Smyth, M. M. (1984). Perception and action. In M. M. Smyth & A. M. Wing (Eds.), *The Psychology of Human Movement* (pp. 119–152). London: Academic Press.

Sole, D. (1997). Johari's window for generating questions. *Journal of Adolescent & Adult Literacy, 40*, 481–483.

Stanley, M., Brooker, R., & Gilbert, R. (2002). Examiner perceptions of using criteria in music performance assessment. *Research Studies in Music Education, 18*, 43–52.

Thompson, S., & Williamon, A. (2003). Evaluating evaluation: Musical performance assessment as a research tool. *Music Perception, 21*, 21–41.

Vasil, T. (1973). *The Effects of Systematically Varying Selected Factors on Music Performing Adjudication.* Unpublished doctoral dissertation, University of Connecticut, Storrs, CT.

Västfjäll, D. (2001–2002). A review of the musical mood induction procedure. *Musicæ Scientiæ, Special Issue*, 173–211.

Wapnick, J., Darrow, A. A., Kovacs, J., & Dalrymple, L. (1997). Effects of physical attractiveness on evaluation of vocal performance. *Journal of Research in Music Education, 45*, 470–479.

Wapnick, J., Mazza, J. K., & Darrow, A. A. (1998). Effects of performer attractiveness, stage behavior, and dress on violin performance evaluation. *Journal of Research in Music Education, 46*, 510–521.

Wapnick, J., Mazza, J. K., & Darrow, A. A. (2000). Effects of performer attractiveness, stage behavior, and dress on evaluation of children's piano performances. *Journal of Research in Music Education, 48*, 323–335.

Williamon, A., & Davidson, J. W. (2002). Exploring co-performer communication. *Musicæ Scientiæ, 6*, 53–72.

Wirtz, J., & Bateson, J. E. G. (1995). An experimental investigation of halo effects in satisfaction measures of service attributes. *International Journal of Service Industry Management, 6*, 84–102.

Wolfe, J. (2003). From idea to acoustics and back again: The creation and analysis of information in music. Paper presented at the Seventh Western Pacific Regional Acoustics Conference, Melbourne, Australia.

Woody, R. H. (2000). Learning expressivity in music: An exploratory study. *Research Studies in Music Education, 14*, 14–23.

Yarbrough, C. (1975). Effect of magnitude of conductor behaviour on students in mixed choruses. *Journal of Research in Music Education, 23*, 134–146.

Ybarra, O. (2001). When first impressions don't last: The role of isolation and adaptation processes in the revision of evaluative impressions. *Social Cognition, 19*, 491–520.

Zajonc, R. B. (1965). Social facilitation. *Science, 149*, 269–274.

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25*, 82–91.

PART II

# PRACTICE STRATEGIES

CHAPTER 5

# STRATEGIES FOR INDIVIDUAL PRACTICE

HARALD JØRGENSEN

For the most part, individual practicing is a solitary activity. The performer is alone, with his or her instrument, and must rely on personal skill to achieve progress throughout the practice session. What musicians strive for, and what this chapter is about, is achieving *effective* individual practice. Hallam (1998) describes effective practice as "that which achieves the desired end product, in as short a time as possible, without interfering with long term goals" (p. 142).

By selecting and implementing appropriate practice *strategies*, a performer can acquire necessary tools to accomplish this. For the sake of clarity, practice strategies can be defined as thoughts and behaviors that musicians engage in during practice that are intended to influence their motivational or affective state, or the way in which they select, organize, integrate, and rehearse new knowledge and skills (adapted from Weinstein & Mayer, 1986). For instance, when a musician makes plans for a practice session, this is a "thought" strategy, and when a musician gradually increases the tempo of a performance, this is a "behavior" strategy. Strategies are (usually) consciously applied but may become automatic with repetition.

Most strategies are neither functional nor dysfunctional in and of themselves. Practicing a piece of music by playing it over and over again from beginning to end without interruption may be beneficial for some, but ineffective for others. Indeed, advanced students and professional musicians often hold idiosyncratic views as to what constitutes effective practice. Such differences in approach are to be expected, as performers work to accentuate their individual strengths and eliminate their weaknesses. In spite of this diversity, however, there are common strategies that musicians can employ to produce better performance results or to achieve the same results more quickly.

A number of music teachers have suggested that it is convenient for students to view practicing as a means of "self-teaching", where in the absence of the teacher, students must act as the teacher's deputy, assigning themselves definite tasks and supervising their own work (Galamian, 1964). Based on this view, one can emphasize that effective practice (both within a single session and across longer periods of time) should include three self-teaching phases:

- planning and preparation of practice
- execution of practice
- observation and evaluation of practice.

An alternative view, offered by educational psychologists, is the metaphor of "self-regulated learning" (i.e. how students acquire the tools necessary to take control of their own learning and, thereby, learn effectively). In this approach, the practitioner should engage in three self-regulating phases:

- *forethought*: the thought processes and personal beliefs that precede efforts to engage in a task
- *performance/volitional control*: processes that occur during learning that affect concentration and performance
- *self-reflection*: the learner's reaction and subsequent response to the experience (see McPherson & Zimmerman, 2002; Zimmerman, 1998).

The three phases identified in this approach can easily be mapped onto those of self-teaching described above. Moreover, both approaches view the respective phases of practice as interwoven, or cyclical.

This chapter presents and classifies practice strategies in terms of the self-teaching phases of planning, execution, and observation and evaluation, as these terms tend to be more commonly used by performers and teachers (see Figure 5.1 for a model of strategies based on self-teaching). Nevertheless, within a larger frame of reference, theory

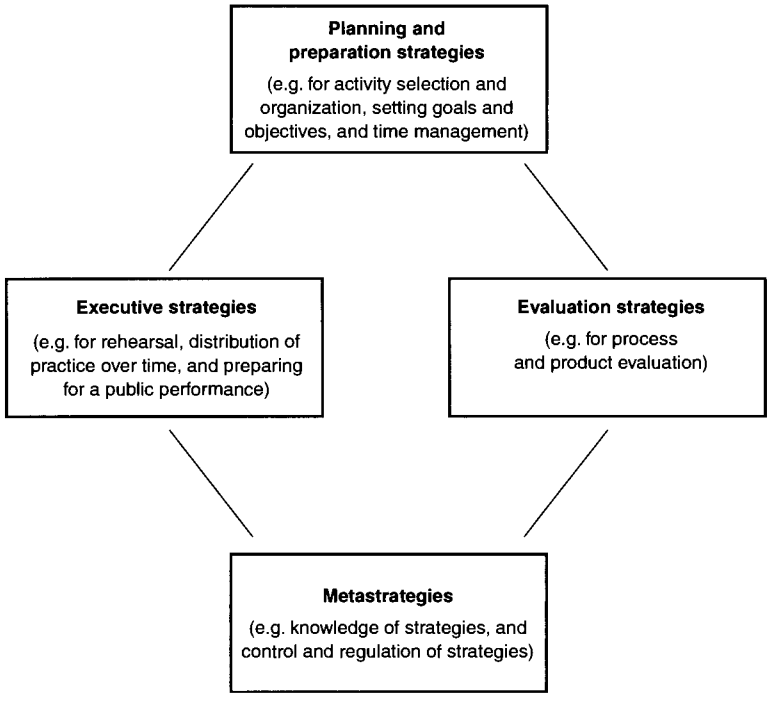



**Figure 5.1** A model of strategies for individual practice.

and research on self-regulation of learning form an important basis for this chapter. Directing the practitioner's attention to three phases to be employed within practice is not, however, sufficient for optimal use of individual strategies. Every practitioner—from the student to the professional musician—must have a thorough knowledge of his or her *repertory* of strategies and must be able to control, regulate, and exploit this repertory. Therefore, the issue of metastrategies (i.e. knowing about and controlling one's strategies) is addressed toward the end of the chapter.

## 5.1 Individual practice: A historical perspective

Since roughly 1700, an abundance of books have been written about instrumental methods or approaches, each containing material for individual practice (e.g. Quantz, 1752/1985; Tosi, 1723/1987; Türk, 1789/1982). Most of these works are based on the personal experience and opinions of the authors, and their views are often contradictory.

Empirical research on individual practice has a much shorter history. The Hungarian pianist, Sandor Kovacs, published the first empirical study in 1916. He was puzzled by the problems his students had in performing difficult pieces from memory. His conclusion was that musicians should engage in mental practice, especially when initially learning a new piece. During the subsequent 20 years, only three research studies were published. Two of these (Brown, 1928; Eberly, 1921) examined the issue of the "whole" versus "part" approaches to practicing (see Section 5.3.1), while the third dealt with the coordination of hands in piano playing (Brown, 1933). Then, in 1937, Grace Rubin-Rabson, an American psychologist and educationist, published the first of a series of 10 research studies on piano practicing. The last of her studies appeared in 1947. After that, there was a period of silence for almost 30 years.

Since 1975, the number of published studies on individual practice strategies has gradually increased with each decade, and approximately two thirds of the published works from 1916 to present are dated 1990 or later. All in all, this research has explored the efficacy of a wide range of practice strategies, as well as the psychological and pedagogical principals that underpin learning and teaching in music. Although some conflicting results have emerged within the body of empirical work since 1916, the strategies offered below are drawn from a selection of the most salient and robust findings.

## 5.2 Planning and preparation strategies

### 5.2.1 Strategies for activity selection and organization

The quality of a practice session is characterized and defined by the specific activities in which the practitioner engages. Some activities are *not* related to learning, such as daydreaming or avoidance behaviors that include taking a long time to set up a music stand or perform maintenance on an instrument (see Pitts *et al.*, 2000a; McPherson & Renwick, 2001). It almost goes without saying that these strategies are to be avoided.

Practice activities *with* learning objectives can be divided into "playing practice" and "nonplaying practice" (e.g. studying the score, writing in the score or in a notebook,

silent fingering, etc.). Research suggests that playing activities tend to dominate the practice session, and within those, musicians most commonly engage in repertory work, followed by practice on technical issues (Geringer & Kostka, 1984; Hallam, 1995; Sloboda *et al.*, 1996; Jørgensen, 1998; McPherson & Renwick, 2001). From this, the following suggestions emerge:

- Reduce the nonlearning time in practice sessions by enhancing concentration and motivation (see Chapters 11 and 12).
- Balance "playing practice" with "nonplaying practice" in a single session or over a period of time. Focused, nonplaying practice will give more time for mental rehearsal and reflection and prevent overuse of muscles (as discussed in Chapter 3). In the long run, this is not a waste of time, but in fact, quite the opposite (see Section 5.3.1).

There appears to be a wide range of strategies used by practitioners to select and organize practice activities, from those who do as they "usually do" to those who change practice activities from session to session (Jørgensen, 1998).

- If you have a specific routine, try to introduce new elements into your practice sessions or practice schedule. For instance, you can reserve one session a day for "unexpected demands" (e.g. a sight-reading task) or use a whole session for technical work.
- If you vary practice content and sequence of activities from session to session, reflect on the benefits of a more regular regime that touches on all the immediately relevant aspects of performance before you turn to something else.

Warm-up exercises are used by most musicians and are included in most sessions. These exercises are so closely related to the instrument's peculiarities and physical demands (as well as to the individual's muscular and technical proficiency) that it is difficult to offer specific recommendations. However:

- Ask yourself: why is this particular warm-up exercise necessary for me in this session? Are other exercises better suited for this particular practice situation? For my short-term needs? For my long-term needs?

These same questions can be applied to practice on exercises and etudes that are intended to prepare particular technical and musical aspects of performance. Of course, such exercises can benefit the general advancement toward performance excellence, but exercises should be well focused and carefully selected if they are to address specific challenges within a given composition (Harvey *et al.*, 1987; Pacey, 1993; Pierce, 1992).

Finally, research has shown that, within practice sessions, musicians normally begin with technical exercises and then move on to repertory work (Duke *et al.*, 1997; Jørgensen, 1998). Within this repertory work, however, there is evidence to suggest that many students practice only new pieces and do not have a strategy for systematically reviewing "old" pieces (Jørgensen, 1998).

- Include deliberate work on formerly rehearsed music in your practice schedule to keep your core repertory active.

### 5.2.2 Strategies for setting goals and objectives

Goals are necessary to achieve learning and improvement (as shown in Chapters 2 and 12 of this volume). However, in a study of conservatory students, Jørgensen (1998) found that many practice sessions were started without efforts to plan and establish goals. Nevertheless, there seems to be a tendency to focus on two types of goals for the practice of a specific piece of music: goals for technical performance quality and goals for expressive performance quality.

The expressive intentions for a piece should be regarded as part of a *performance plan*, which is comprised of more or less conscious musical ideas or communicative messages to be developed and incorporated into performance. Some musicians let their expressive ideas guide their technical work. Others develop their performance with no preconceived plan, allowing it to evolve as practice progresses. Still others use a combination of a preconceived plan and intuitive processes in working toward a performance (Hallam, 1997a). Another approach is to develop a performance plan after mastering most of the technical challenges in a piece (Nielsen, 2001). There is no research indicating that one approach is more efficient than the others independent of circumstances. This is certainly one of the strategies where effectiveness depends on individual-specific preferences for learning and particular characteristics of the performer, in addition to the nature of the music. However, one general recommendation is:

- At the beginning of the session, formulate some of your overall intentions for the practice session (whatever these may be) as clearly and precisely as possible. What do you specifically want to develop? To master?

Goals come and go, and they will change during a long rehearsal process. Individuals who plan what to do—and know why they do something and what they wish to accomplish—can make their practice sessions more efficient. Practitioners must not refrain from formulating specific plans just because they know that unexpected events and frustrating moments will inevitably pop up and may alter any planning efforts. Goals and related issues are dealt with further in several sections of this chapter, as well as in Chapter 12.

### 5.2.3 Time management strategies

Time management is integral to effective practice. For novice students, the two most frequently asked questions are "how much time do I have to practice" and "how often do I have to practice." Both of these questions concern the management of practice quantity, and the general answer to all practitioners is that there is a positive relationship between quantity of practice and overall musical achievement (Jørgensen, 2002).

As expertise develops, both the length and frequency of practice sessions tend to increase (Sloboda *et al.*, 1996). During the formative years of childhood and youth, increasing practice quantity is typically more the result of longer sessions than of more frequent sessions or more practice days (Hallam, 2001a). Also, Hallam noted, not surprisingly, that 95% of the novices and advanced students in her study reported an increase in practice during the weeks preceding examinations. Another event that seems to generate more practice activity is an upcoming or recently concluded private

lesson (Lehmann & Ericsson, 1998), while school holidays usually result in a decrease in practice activity (Sloboda *et al.*, 1996).

The notion of what constitutes a "normal" frequency and volume of practice for students across the performance programs in higher music education seems to be fairly consistent. First of all, students are expected to practice regularly—that is, every day or most of the days in a week. This means that they have to manage the distribution of practice sessions throughout the week and during a practice day within the context of a busy study schedule. Several management strategies have been observed in relation to this. Some students practice at the same time every day (Duke *et al.*, 1997), and some integrate practice into a daily or weekly plan (Jørgensen, 1997b). Most students, however, try to fit in practice sessions between other activities without any preconceived plan. There may be many sound reasons for this, but some crucial questions to ask are:

- Do you give practice sessions the highest priority among your study activities throughout a study day?
- Do you say "no" to other activities that interfere with your planned sessions, or do you look at practice sessions as something that can easily be subordinated to other demands, cancelled, or postponed?

The answers to these questions can offer an important indicator of practicing behavior and may help trigger a new way of thinking about study habits.

Is there an optimal time of day for practicing? Certainly, common sense tells us that we need to be awake and alert. Thus, one strategy is to practice those pieces that are considered to be most difficult in the morning (Lehmann & Ericsson, 1998). Ericsson *et al.* (1993) found that the students at the highest skill level in one conservatory practiced in the morning, took naps in the afternoon, and then put in more practice in the evening. Although this may be a strategy that can be adopted by only a select few, its general relevance is clear:

- Attack what you regard as your day's most challenging practice task at a time when you feel most alert and ready for serious work.

There is no universally applicable answer to the question of how much to practice. For students in higher education, however, an unwritten norm for the amount of individual practice seems to average around 20–25 hours per week. There are pronounced variations in both directions from this average, primarily depending on the nature of the instrument. In general, keyboard players have a higher quantity of individual practice than string players, who practice more than wind players (Jørgensen, 1997a). There are also differences between specific instruments within these groups. Violinists, for instance, tend to practice more than double bass players, and flute players practice more than tuba players. These differences are partly related to the restrictions posed by the instrument's physical and technical demands and partly on the aspirations and motivation of the performer, which is why such large differences can be observed in the amount of individual practice among students who play the same instrument. While one trumpet player may practice 10 hours in a week, another may practice for 20 or more. Similarly, a violin student practicing 20 hours in a week may have a friend who practices for 35 or 40 hours.

Looking at these differences, and keeping in mind that there is a positive relationship between the amount of practice and rise in general proficiency, it is reasonable to conclude that different levels of musical accomplishment may be demonstrated between those who practice less versus those who practice more during the 4–6-year duration of higher music education. Suppose there are two students of the same instrument, one of whom has practiced 1500 hours in 4 years of study, while the other has practiced in excess of 4500 hours. Assuming that they started their studies at the same general level of proficiency and engaged in the same quality of practice, it is highly probable that they will exhibit a clear difference in proficiency after those 4 years.

In order to manage the amount of time spent in practice successfully, you should observe and reflect upon the relationship between the time you invest in individual practice, and the progress you witness in your playing. Do you invest too little time to make progress? Do you invest just enough time to progress slowly and manage your short-term demands and goals?

Are there strategies that a practitioner can use to help increase practice volume? Individuals who are not sure of how much time they actually invest in individual practice should start by systematically observing and registering their individual practice time.

- If you realize that you ought to practice more, you can formulate a contract (strictly for yourself) of how much time you intend to invest in individual practice for the next week, month, etc. Then, it is up to you to comply with the contract.
- The term "invest" has been consciously used in relation to practice time as a potentially helpful metaphor. View the practice you do over days, weeks, and years as an investment in a "practice amount account," and plan your investment accordingly.

Quantity, however, must not be the only concern of the practitioner. In their study of three young wind students in their first year of practice, Pitts *et al.* (2000a) observed that both the mother of one of the students and her practicing daughter were exclusively concerned with quantity, with the mother insisting on "more" without addressing the quality of the playing and practice. In this respect, there is reason to wonder how many students have only one time management strategy: the dysfunctional "filling time" strategy. Effective practice is not a question of "filling time," but of filling the optimal amount of time with the optimal quality of practice, which means employing carefully selected strategies.

It is intended that, by using appropriate practice strategies to enhance *practice* quality, *performance* quality will, in turn, be enhanced. It is important to note that focusing on practice quality (initially through adequate planning and preparation) will most probably reduce the need for quantity.

## 5.3 Executive strategies

### 5.3.1 Rehearsal strategies

*Mental versus playing rehearsal strategies*

Mental rehearsal is a type of nonplaying practice (see Section 5.2.1) that has been championed by a number of pedagogues and psychologists. It is usually defined as the

cognitive or imaginary rehearsal of a physical skill without overt muscular movement and is often presented as an alternative or a supplement to the playing approach. However, it is important to remember that there is virtually no playing without cognitive activity, which means that the playing approach is inevitably a combination of mental and physical effort.

No specific mental strategy can be identified as the "right" one. There are a myriad of different techniques that can be applied; particular suggestions are offered in Chapter 12 of this volume. The basic purpose of mental strategies is to establish and activate visual, aural, and kinesthetic images of the music for use in both practice and performance.

- Regard the score as a wellspring of information for cognitive activity. Respond to it as a source of information about signs, keys, structures, sounds, movements, touch, etc.

In order to exploit all the possible cues inherent in a score, performers should activate their analysis and ear training skills, as well as their ability to develop images of movements and touch, their knowledge of style and history of music, and their powers of reflection and memorization. There are marked differences between musicians in the breadth and quality of images that they receive and process from the score (Grøndahl, 1987), and only the individual's personal effort can expand and differentiate the repertory of mental strategies.

The relative merits of mental strategies, as compared with playing strategies, have been addressed in a number of studies (see Jones, 1990; Kopiez, 1990; Ross, 1985; Rubin-Rabson, 1937), with some conflicting results and recommendations. Trying to determine which of these strategies function best in a practice situation is not, however, the most important issue because mental and playing strategies are not inseparable. The most important point to consider is how to combine these two approaches in the most constructive fashion during practice.

- Always reflect on how you can combine mental strategies with playing strategies to ensure that they support and enhance each other. For instance, when is it convenient to bring in mental strategies: before you start playing, used as an interruption in playing practice, or after playing as a mental reminder? How do the style and complexity of the music (including the length and familiarity of the piece) influence your choice of strategy?

*Strategies to master the whole piece versus smaller parts*

Every piece of music can be broken down into smaller sections or parts. How can this be applied in practice? Three main approaches have emerged in recent research. One strategy is to play through the entire piece several times, not stopping for special treatment of parts of the piece. A second strategy is to concentrate only on parts of the music before attempting to master the whole; going through each part several times before going on to the next is a characteristic of this approach. The third strategy combines the whole and the part approaches. One could, for instance, play through the music and stop to practice larger parts *en route* (Hallam, 1997b; Miklaszewski, 1989); alternatively, one could start with the whole to gain sufficient knowledge of the piece

(either by studying the entire score or by playing through the piece as well as possible) in order to select and relate parts to larger sections or to the whole piece (Chaffin & Imreh, 1997; Nielsen, 1997, 1999b; Williamon & Valentine, 2002; Williamon *et al.*, 2002).

Although the first strategy is more common among novices than professionals, each of the three strategies has its own merits, depending on the circumstances that define the practice situation. These include the time available for the rehearsal of the piece, the style and complexity of the music, the individual's familiarity with the music, the goal of the practice, and the individual's skill at evaluating and correcting performances on repeated rehearsals.

- Make use of all three major strategies to master both the whole and its parts; do not stick to a single strategy regardless of circumstances.

How is a "part" to be selected? The formal structure of the piece has been shown in some studies to be a good starting point. Even so, there may be technical difficulties within a formal part that forces the practitioner to subdivide it further (Chaffin & Imreh, 1997; Miklaszewski, 1989; Miklaszewski & Sawicki, 1992). Parts may also be selected because of their visual layout and harmonic progressions (Williamon & Valentine, 2002). One study of organ students found that some chose parts that had a musical or technical content that was either new or vaguely related to other sections, rather than sections that were identical (Nielsen, 1999b).

- You are not bound to partition a piece into formal parts in accordance with the composer's view; however, the parts you select ought to be musically meaningful to you.
- Adopt a procedure that combines overview and care for the "whole," with special treatment of difficult or challenging parts that are connected to the whole.

The parts on which one focuses are likely to become longer as practice progresses, and this process has been linked to the production of higher quality musical, communicative, and technical performances (Williamon & Valentine, 2000, 2002). Nevertheless, it is not unusual for attention to be focused on small parts and details throughout the duration of practice on a piece, even in sessions just before a performance (Williamon *et al.*, 2002). Further consideration of the part approach is offered in Chapter 7 in relation to memorizing music.

*Strategies for difficult and challenging parts*

An issue that is closely tied to the above involves practicing difficult or challenging parts. Clearly, "difficulty" depends on each musician's skill and experience, as well as instrument-specific challenges (e.g. double stops, triple tonguing, etc.). One strategy is to use general exercises and etudes that simplify the problem and gradually bring one toward a solution that is applicable to the specific piece. When this happens, it is called positive "transfer of learning" (Nielsen, 1999b).

Another strategy is to practice the difficult part within the given piece repeatedly until it is mastered. For example, choral singers in one study relied on this approach when practicing specific intervals, rather than using exercises to train them to master

these intervals (Harvey *et al.*, 1987). Research on this issue is scant, but the crucial question for the practitioner is:

- Are these exercises really bringing you toward mastery of the specific problem? If you use exercises to master a difficult passage, make sure there is positive transfer between the exercises and the passage itself, with the conditions for learning and performance being as similar as possible (Hallam, 1997a).

*Strategies concerned with tempo of performance*

Research has shown that there are three primary strategies used when a practitioner encounters a composition that he or she is unable to play immediately at the required tempo. The first, starting slowly and then gradually increasing tempo, is the most frequently recommended (see Barry, 1992). Interestingly, however, slower performances may contain more errors than those that are faster (Drake & Palmer, 2000). This is not necessarily an argument against the strategy, as the piece will no doubt grow more familiar and contain fewer errors as practice progresses. A more important objection, though, is that starting out using arm, finger, or mouth movements that are far from the physical and muscular demands of the final movements can cause the initial learning not only to be a waste of time, but also counterproductive in developing the "correct" muscular responses. The reason for this is that qualitatively different movements activate different muscles (Winold *et al.*, 1994). Another strategy is to alternate, if possible, between a slow tempo and the faster performance tempo (Donald, 1997). This may bring the piece up to the performance tempo more quickly, since it entails regular tests of one's ability to play at that tempo. A third strategy is to practice at the performance tempo from start to finish.

The general rule of thumb here is that a tempo strategy should be chosen only after careful consideration of what it can accomplish. There is one issue, however, that is often overlooked by students who restrict their use of alternative tempo strategies. This involves the ability to identify and classify errors that are and are not acceptable during practice. It seems that most students are willing to accept rhythmic and dynamic deviations from the "target" performance, as they are extremely anxious to play the "right notes" (and to play them in tune). The issue of errors will be dealt with further in Section 5.4.2; as for now:

- Decide the types and number of performance errors you are willing to accept to bring the tempo up to speed as quickly as possible, and choose a tempo strategy in accordance with this decision.

Practicing tempo variations can also be useful for trying out different expressive renderings of the music (see Oura & Hatano, 2001).

### 5.3.2 Strategies for the distribution of practice over time

Imagine a situation in which there is one week to rehearse a new piece and that this rehearsal, in total, will take approximately 10–12 hours. In such a situation, one could spread practice over the entire week (the "distributed" approach) or cluster it together in just a few sessions (the "massed" approach). There are, of course, many available

options for deciding on the length and distribution of practice sessions, but one must consider what is going on in the period between sessions. Does this comprise rest, practice on other pieces, or activities unrelated to practicing? In light of an almost complete lack of research in this area and the complexity of the issue, the best recommendation may still be from Fleischman and Parker (1962):

- Use distributed practice when the practice sessions are long enough to bring about improvement and when the time between sessions is long enough to overcome fatigue, but not so long that forgetting occurs.

### 5.3.3 Strategies for preparing for a public performance

There are several issues that influence public performance that could benefit from close attention during practice. Performance is often accompanied by worries regarding memory lapses, anxiety, and stage fright, and these issues are addressed in detail elsewhere in this book (see Chapters 7, 9, 10, 12, and 14). In addition, it is important to consider the audience's perspective during practice. For example, experienced performers have reported taking the audience's perspective into account when developing their expressive ideas (Oura & Hatano, 2001) and working deliberately on the communication of musical ideas in the final stages of practice on a piece (Williamon *et al.*, 2003). Many have also reported imagining how the audience will perceive the performance in a range of different acoustical environments (Edlund, 2000). These considerations lead to questions such as:

- Who is your audience and what type of concert venue are you going to perform in? When you practice, imagine how the audience will perceive and react to the soundscape you create.

In performance, many students seem preoccupied solely with the music and forget to prepare their stage behavior adequately. Understanding the role of body movement in performance, for instance, may help harmonize musical and bodily messages to the audience (Davidson, 1993; see also Chapter 4 of this volume for a discussion of factors that influence audiences' evaluations of performance).

- Observe your body movements and ask whether they are supporting your musical intentions? Are they counterproductive? Viewing videos of previous concerts, combined with written self-evaluations of your stage behavior, is recommended for improvement (Daniel, 2001).

## 5.4 Evaluation strategies

### 5.4.1 The necessity of evaluation

Knowledge of results, or informative feedback, is essential to the learning process. This means that a practitioner should explicitly diagnose his or her strengths and weaknesses in prescribing solutions to problems. Practicing requires constant vigilance and attention to all kinds of feedback, as well as a repertory of knowledge and skills to remedy problems. This is an area in which many practitioners could improve. For instance,

the end of a practice session is an important time for evaluation. However, only 21% of conservatory students in one study regularly evaluated and made subsequent plans after practice (Jørgensen, 1998). Thus, sessions may be ended rather haphazardly, with no reflection on progress and how to continue practicing.

- Do you use evaluation as a regular strategy in your efforts to improve?

### 5.4.2 Strategies for process evaluation

*Aural and visual models*

In process evaluation, evaluative thoughts and behaviors are interwoven with ongoing practice. One approach to this type of evaluation is for the individual to have a model performance in mind during practice (see also Chapters 7 and 12). Accomplished performers are expected to develop their own models as mental images, while less experienced performers may need to use others' performances as models (e.g. from a recording of a piece). With unfamiliar and complex music, research suggests that experts, too, can benefit from studying others' performances as models (see Lisboa *et al.*, in press, for discussion).

- Develop a performance plan or use a recording to guide your evaluation of your performance.

Viewing videos of your own previous performances is also recommended, especially if this is combined with written self-evaluations (Daniel, 2001). This has the additional advantage of enhancing student independence in assessing performance, since students usually rely on comments from teachers to improve.

*Strategies to detect and correct errors*

Detecting and correcting errors is a major task during practice. The musician must have a notion of what errors are and the ability to recognize them when they occur, including errors in pitch, rhythmic accuracy, dynamics, intonation, steadiness of pulse, and tonality (see Drake & Palmer, 2000; Palmer & van de Sande, 1995). Some of these are closely related to the specific demands of the instrument, but the question each practitioner must ask is:

- Do you have a preoccupation with some errors and ignore others? Develop a broad knowledge of error types. Write out the most important types of errors related to your instrument or your own most recurring errors, and keep a watchful eye on them.

The German cellist, Gerhard Mantel (1984), recommends that the practitioner direct his or her attention to different aspects of the music studied in a systematic and planned manner. This will aid concentration and also aid in detecting and correcting errors. In many cases, errors are made but go unobserved because the performer has not studied or grasped essential features of the score.

- Use a thorough study of the score to identify important aspects of the music and thus prevent unnecessary errors.

How can errors be corrected? One strategy is to ignore errors and play on, hoping that they will disappear on replaying. Another is for the musician to stop and correct errors whenever they occur. Both approaches have certain advantages, as long as the practitioner remains aware of major pitfalls of each. Regarding the first of these, errors once made tend to be practiced in, instead of removed (Hallam, 1997a). As for the second strategy, "stopping" may also be learned, prompting small pauses at the location of errors in subsequent performances. A third approach is probably the most generally useful:

- Practice whole sections followed by specific attention to errors and error-correction within these sections.

*Self-guiding strategies*

Self-guidance involves covertly or overtly describing how to proceed, giving comments on progress, and reminding oneself of concentration lapses and motivational aspects of the situation (Nielsen, 2001). There are warning strategies ("Oh, here comes that tricky bit, I have to play it slower"), comments as strategies ("No, not so fast!"), supportive verbalizations as strategies ("One, two, three, one, two, three"), and reminder strategies ("Remember the melody in the left hand"). Guidance strategies also involve singing along with playing.

- Regardless of which approach is used, it is important to give constructive messages to yourself when practicing.

## 5.5 Metastrategies

### 5.5.1 Knowledge about strategies

The importance of having a repertory of strategies, ready for use, has been mentioned several times, and by reading this chapter, it is intended that the practitioner will learn more about the strategies that best suit his or her needs. Such knowledge about strategies is commonly referred to within psychology as "metacognitive knowledge." Borkowski and Turner (1990) have proposed three different areas of metacognitive knowledge.

- *Specific knowledge* is comprised of the strategies that the individual actually knows.
- *Relational knowledge* covers the knowledge of how different strategies may contribute toward satisfying different tasks and goals.
- *General knowledge* has to do with an understanding of how much effort is required when employing practice strategies; for instance, concentrated practice on problem areas in a piece is typically regarded by students as requiring greater effort than playing familiar repertory (Lehmann, 2002).

A distinction may also be made between *knowledge about personal variables* and *task knowledge* (Flavell, 1987). The first offers insight into individual learning processes and the individual's personal thoughts about these processes. The second is knowledge about how tasks differ from each other (e.g. in complexity) and how an individual's goals influence the ways in which tasks will be addressed (e.g. different strategies for sight-reading and memorizing music).

Several studies have observed that practitioners have a restricted knowledge of specific strategies and that there are considerable differences between beginners, novices, and experts (e.g., Hallam, 2001b; Pitts *et al.*, 2000a, 2000b; Renwick & McPherson, 2002; Sloboda *et al.*, 1996), as well as individual differences among musicians at the same level of accomplishment (Nielsen, 1997, 1999a, 1999b, 2001). The same applies to relational and general knowledge and personal and task knowledge. Developing a repertory of strategies and acquiring knowledge about these, as well as one's own cognitive functioning, is certainly one of the most important objectives for any practitioner and should be emphasized by teachers in students' lessons and rehearsal time.

- Observe your practice behavior, and try to register the strategies you use. Do you have strategies for all issues covered in this chapter? Doing this, you will probably find that you have a broader range of strategies than expected.

### 5.5.2 Control and regulation of strategies

A second metacognitive area is the control and regulation of strategies. Control of strategies is, first of all, dependent on knowledge and classification of strategies as described above. Based on this knowledge, one can then proceed with *checking, evaluation,* and *prediction* (Kluwe, 1987).

- What strategies are being used? (*knowledge and classification*).
- How well is your practice proceeding? (*checking and evaluation*).
- What type of result will occur if . . . ? (*prediction*).

The regulation of strategies involves deliberate *effort, task* selection, *speed,* and *intensity* (see Nielsen, 1999b). Questions such as the following may be helpful in regulating strategy use in specific practice situations:

- How much effort are you willing to invest in this problem, and consequently, what strategies will you use? (*effort*).
- Where will you start? (*task selection*).
- How many strategies are you willing to use to solve this problem? (*speed*).
- How long will you persist with a strategy if it does not give results? (*intensity*).

## 5.6 Conclusions

This chapter has addressed a number of strategies for individual practice. The development from beginner and novice to higher levels of accomplishment and expertise is accompanied by a broadening of one's repertory of strategies and skills in using them. At all levels of accomplishment, however, research indicates that there may be a range of individual differences in the extent to which strategies are appropriate, applicable, and effective. This is a reminder that the person, task, and context must always be kept in mind when reflecting on strategy use. There is, as Hallam (1995) observes, no simple correspondence between a particular practice strategy and success as a performer.

Developing a thorough knowledge of strategies and strategy use is not necessarily straightforward. Many instrumental teachers are of the opinion that they teach their

students how to practice. When Barry and McArthur (1994) asked teachers if they included specific instructions on how to practice in their lessons, 84% answered "always" or "almost always." However, what the student perceives and remembers may not correspond to what the teacher thinks he or she is communicating. For three consecutive years, Jørgensen (2000) asked students entering a music conservatory how much their former instrumental teachers had emphasized practice. Approximately 40% reported that their teachers had put "very little" or "no" emphasis on practice behavior. This discrepancy, therefore, presents a challenge to all teachers to include instruction on practice strategies in their lessons in a way that directly captures the interest of the student, and then to follow that up with appropriate learning and performance assignments.

Barry (1992) has demonstrated that strategies can, indeed, be successfully taught. She found, for example, that wind instrument students, when instructed, were able to examine the music visually, to "finger through" the music silently, and to tap the rhythm of an entire étude before playing it. These strategies were generally not observed among students who were simply asked to "practice in any way you think best." Similarly, Kenny (1992) reported positive results from a study in which beginning band students were instructed in strategies such as problem identification, practice planning, goal selection, monitoring, and evaluation. This study lasted for only 18 weeks, and there was no difference in technical proficiency and sight-reading ability between the students that were instructed in practice strategies and those who were not. However, the instructed students expressed stronger favorable opinions about practicing, were more likely to engage in practice planning and problem identification, were better able to select appropriate performance goals, and were able to formulate goals that were more cognitively complex than their control group counterparts. These results remind us that:

- Practice must be practiced! Observe your practice sessions regularly, and concentrate your observation on one or two specific features at a time. For example, try to register how much time you spend on specific tasks during a practice session, what types of errors you produce, how you include repetition of previously rehearsed repertory into your practice schedule, etc. The number and range of issues are limitless. The main point is that practicing must be taken seriously, as the *professional* musician is dependent on a *professional* level of practicing.

At the beginning of this chapter, it is stated that books about instrumental methods and approaches are primarily based on personal experience and on the opinions of the respective author. Thus, the views put forth are often contradictory at a very profound level. What is called for now is for teachers and practitioners to employ their personal experience, the experience of others, as well as research findings, in order to reflect on and experiment with practice, while remaining open to both traditional and nontraditional strategies.

## Further information and reading

Bernstein, S. (1981). *With Your Own Two Hands: Self-Discovery Through Music.* New York: Schirmer. [See Chapter 1 "Why do your practice?" and Chapter 2 "Why don't you practice?" for a discussion of the relationship between practicing and personal self-integration.]

Hallam, S. (1997). What do we know about music practising? Towards a model synthesising the research literature. In H. Jørgensen & A. C. Lehmann (Eds.), *Does Practice Make Perfect? Current Theory and Research on Instrumental Music Practice* (pp. 179–231). Oslo: Norwegian Academy of Music. [Gives an overview of research up to 1997.]

Lehmann, A. C. (1997). The acquisition of expertise in music: Efficiency of deliberate practice as a moderating variable in accounting for sub-expert performance. In I. Deliège & J. Sloboda (Eds.), *Perception and Cognition of Music* (pp. 161–187). Hove, UK: Psychology Press. [Relates a discussion of deliberate practice to research on expertise and development of practice throughout the life-span.]

McPherson, G. E., & Zimmerman, B. J. (2002). Self-regulation of musical learning: A social cognitive perspective. In R. Colwell & C. Richardson (Eds.), *The New Handbook of Research on Music Teaching and Learning* (pp. 327–347). Oxford: Oxford University Press. [Contains a thorough review of the self-regulation phases of forethought, performance/volitional control, and self-reflection.]

## References

Barry, N. H. (1992). The effects of practice strategies, individual differences in cognitive style, and gender upon technical accuracy and musicality of student instrumental performance. *Psychology of Music, 20*, 112–123.

Barry, N. H., & McArthur, V. (1994). Teaching practice strategies in the music studio: A survey of applied music teachers. *Psychology of Music, 22*, 44–55.

Brown, R. W. (1928). A comparison of the "whole," "part" and "combination" methods of learning piano music. *Journal of Experimental Psychology, 11*, 235–247.

Brown, R. W. (1933). The relation between two methods of learning piano music. *Journal of Experimental Psychology, 16*, 435–441.

Borkowski, J. G., & Turner, L. A. (1990). Transsituational characteristics of metacognition. In W. Schneider & F. E. Weinert (Eds.), *Interactions Among Aptitudes, Strategies, and Knowledge in Cognitive Performance* (pp. 159–176). New York: Springer.

Chaffin, R., & Imreh, G. (1997). "Pulling teeth and torture": musical memory and problem solving. *Thinking and Reasoning, 3*, 315–336.

Daniel, R. (2001). Self-assessment in performance. *British Journal of Music Education, 18*, 215–226.

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music, 21*, 103–113.

Donald, L. S. (1997). *The Organization of Rehearsal Tempos and Efficiency of Motor Skill Acquisition in Piano Performance.* Unpublished doctoral dissertation, University of Texas, Austin, TX.

Drake, C., & Palmer, C. (2000). Skill acquisition in music performance: Relations between planning and temporal control. *Cognition, 74*, 1–32.

Duke, R. A., Flowers, P. J., & Wolfe, D. E. (1997). Children who study piano with excellent teachers in the United States. *Bulletin of the Council for Research in Music Education, 132*, 51–84.

Eberly, L. E. (1921). *Part versus Whole Method in Memorizing Music.* Unpublished masters dissertation, Columbia University, New York.

Edlund, B. (2000). Listening to oneself at a distance. In C. Woods, G. Luck, R. Brochard, F. Seddon, & J. A. Sloboda (Eds.), *Proceedings of the Sixth International Conference on Music Perception and Cognition.* Keele, UK: Keele University.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363–406.

Flavell, J. H. (1987). Speculations about the nature and development of metacognition. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, Motivation, and Understanding* (pp. 21–29). Hillsdale, NJ: Erlbaum.

Fleischman, E. A., & Parker, J. F. (1962). Factors in the retention and relearning of perceptual-motor skill. *Journal of Experimental Psychology, 64,* 215–226.

Galamian, I. (1964). *Principles of Violin Playing and Teaching.* London: Faber & Faber.

Geringer, J. M., & Kostka, M. J. (1984). An analysis of practice room behavior of college music students. *Contributions to Music Education, 11,* 24–27.

Grøndahl, D. (1987). *Thinking Processes and Structures Used by Professional Pianists in Keyboard Learning.* Unpublished masters dissertation, Norwegian Academy of Music, Oslo.

Hallam, S. (1995). Professional musicians' approaches to the learning and interpretation of music. *Psychology of Music, 23,* 111–128.

Hallam, S. (1997a). What do we know about practising? Towards a model synthesising the research literature. In H. Jørgensen & A. C. Lehmann (Eds.), *Does Practice Make Perfect? Current Theory and Research on Instrumental Music Practice* (pp. 179–231). Oslo: Norwegian Academy of Music.

Hallam, S. (1997b). Approaches to instrumental music practice of experts and novices: Implications for education. In H. Jørgensen & A. C. Lehmann (Eds.), *Does Practice Make Perfect? Current Theory and Research on Instrumental Music Practice* (pp. 89–107). Oslo: Norwegian Academy of Music.

Hallam, S. (1998). *Instrumental Teaching: A Practical Guide to Better Teaching and Learning.* Oxford, UK: Heinemann.

Hallam, S. (2001a). The development of expertise in young musicians: Strategy use, knowledge acquisition and individual diversity. *Music Education Research, 3,* 7–23.

Hallam, S. (2001b). The development of metacognition in musicians: Implications for education. *British Journal of Music Education, 18,* 27–39.

Harvey, N., Garwood, J., & Palencia, M. (1987). Vocal matching of pitch intervals: Learning and transfer effects. *Psychology of Music, 15,* 90–106.

Jones, A. R. (1990). *The Role of Analytical Prestudy in the Memorization and Retention of Piano Music with Subjects of Varied Aural/Kinesthetic Ability.* Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign, IL.

Jørgensen, H. (1997a). Time for practising? Higher level music students' use of time for instrumental practising. In H. Jørgensen & A. C. Lehmann (Eds.), *Does Practice Make Perfect? Current Theory and Research on Instrumental Music Practice* (pp. 123–139). Oslo: Norwegian Academy of Music.

Jørgensen, H. (1997b). Higher instrumental students' planning of practice. In A. Gabrielsson (Ed.), *Proceedings of the Third Triennial ESCOM Conference* (pp. 171–176). Uppsala, Sweden: University of Uppsala.

Jørgensen, H. (1998). *Planlegges øving? [Is Practice Planned?].* Oslo: Norwegian Academy of Music.

Jørgensen, H. (2000). Student learning in higher instrumental education: Who is responsible? *British Journal of Music Education, 17,* 67–77.

Jørgensen, H. (2002). Instrumental performance expertise and amount of practice among instrumental students in a conservatoire. *Music Education Research, 4*, 105–119.

Kenny, W. E. (1992). *The Effect of Metacognitive Strategy Instruction on the Performance Proficiency and Attitude Toward Practice of Beginning Band Students.* Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign, IL.

Kluwe, R. H. (1987). Executive decisions and regulation of problem solving behavior. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, Motivation, and Understanding* (pp. 31–64). Mahwah, NJ: Erlbaum.

Kopiez, R. (1990). *Der Einfluss kognitiver Strukturen auf das Erlernen eines Musikstücks am Instrument.* Frankfurt am Main, Germany: Peter Lang.

Kovacs, S. (1916). Untersuchungen über das musikalische Gedächtnis. *Zeitschrift für angewandte Psychologie, 11*, 113–135.

Lehmann, A. C., & Ericsson, K. A. (1998). Preparation of a public piano performance: The relation between practice and performance. *Musicæ Scientiæ, 2*, 67–94.

Lehmann, A. C. (2002). Effort and enjoyment in deliberate practice: A research note. In I. M. Hanken, S. G. Nielsen, & M. Nerland (Eds.), *Research in and for Higher Music Education: Festschrift for Harald Jørgensen* (pp. 55–67). Oslo: Norwegian Academy of Music.

Lisboa, T., Williamon, A., Zicari, M., & Eiholzer, H. (in press). Mastery through imitation. *Musicæ Scientiæ.*

Mantel, G. (1984). Üben schwerer Stellen nach dem Prinzip der rotierenden Aufmerksamkeit. *Üben und Musizieren, 3*, 147–150.

McPherson, G. E., & Renwick, J. M. (2001). A longitudinal study of self-regulation in children's musical practice. *Music Education Research, 3*, 169–186.

McPherson, G. E., & Zimmerman, B. J. (2002). Self-regulation of musical learning: A social cognitive perspective. In R. Colwell & C. Richardson (Eds.), *The New Handbook of Research on Music Teaching and Learning* (pp. 327–347). Oxford, UK: Oxford University Press.

Miklaszewski, K. (1989). A case study of a pianist preparing a musical performance. *Psychology of Music, 17*, 95–109.

Miklaszewski, K., & Sawicki, L. (1992). Segmentation of music introduced by practicing pianists preparing compositions for public performance. In R. Dalmonte & M. Baroni (Eds.), *Secundo Convegno Europeo di Analisi Musicale* [Proceedings of the Second European Conference on Musical Analysis] (pp. 113–121). Trento, Italy: University of Trento Press.

Nielsen, S. G. (1997). Self-regulation of learning strategies during practice: A case study of a church organ student preparing a musical work for performance. In H. Jørgensen & A. C. Lehmann (Eds.), *Does Practice Make Perfect? Current Theory and Research on Instrumental Music Practice* (pp. 109–122). Oslo: Norwegian Academy of Music.

Nielsen, S. G. (1999a). Learning strategies in instrumental music practice. *British Journal of Music Education, 16*, 275–291.

Nielsen, S. G. (1999b). Regulation of learning strategies during practice: A case study of a single church organ student preparing a particular work for a concert performance. *Psychology of Music, 27*, 218–229.

Nielsen, S. G. (2001). Self-regulating learning strategies in instrumental music practice. *Music Education Research, 3*, 155–167.

Oura, Y., & Hatano, G. (2001). The constitution of general and specific mental models of other people. *Human Development, 44*, 144–159.

Pacey, F. (1993). Schema theory and the effect of variable practice in string teaching. *British Journal of Music Education, 10,* 91–102.

Palmer, C., & van de Sande, C. (1995). Range of planning in music performance. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 947–962.

Pierce, M. A. (1992). The effects of learning procedure, tempo, and performance condition on transfer of rhythm skills in instrumental music. *Journal of Research in Music Education, 40,* 295–315.

Pitts, S. E., Davidson, J. W., & McPherson, G. E. (2000a). Developing effective practice strategies: Case studies of three young instrumentalists. *Music Education Research, 2,* 45–56.

Pitts, S. E., Davidson, J. W., & McPherson, G. E. (2000b). Models of success and failure in instrumental learning: Case studies of young players in the first 20 months of learning. *Bulletin of the Council for Research in Music Education, 146,* 51–69.

Quantz, J. J. (1752/1985). *On Playing the Flute* (E. R. Reilly, Trans. 2nd ed.). London: Faber and Faber. (Original work published in 1752).

Renwick, J. M., & McPherson, G. E. (2002). Interest and choice: Student-selected repertoire and its effect on practicing behaviour. *British Journal of Music Education, 19,* 173–188.

Ross, S. L. (1985). The effectiveness of mental practice in improving the performance of college trombonists. *Journal of Research in Music Education, 33,* 221–230.

Rubin-Rabson, G. (1937). *The Influence of Analytical Pre-Study in Memorizing Piano Music.* New York: Archives of Psychology.

Rubin-Rabson, G. (1947). Studies in the psychology of memorizing music. VIII: The inhibitory influence of the same and of different degrees of learning. *Journal of Musicology, 5,* 13–28.

Sloboda, J. A., Davidson, J. W., Howe, M. J. A., & Moore, D. G. (1996). The role of practice in the development of performing musicians. *British Journal of Psychology, 87,* 287–309.

Tosi, P. F. (1723/1987). *Observations on the Florid Song* (J. Galliard, Trans.). London: Stainer and Bell. (Original work published in 1723).

Türk, D. G. (1789/1982). *School of Clavier Playing* (R. H. Haggh, Trans.). Lincoln, NE: University of Nebraska Press. (Original work published in 1789).

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock (Ed.), *Handbook of Research on Teaching* (3rd ed., pp. 315–327). New York: Macmillan.

Williamon, A., & Valentine, E. (2000). Quantity and quality of musical practice as predictors of performance quality. *British Journal of Psychology, 91,* 353–376.

Williamon, A., & Valentine, E. (2002). The role of retrieval structures in memorizing music. *Cognitive Psychology, 44,* 1–32.

Williamon, A., Valentine, E., & Valentine, J. (2002). Shifting the focus of attention between levels of musical structure. *European Journal of Cognitive Psychology, 14,* 493–520.

Williamon, A., Lehmann, A. C., & McClure, K. (2003). Studying practice quantitatively. In R. Kopiez, A. C. Lehmann, I. Wolther, & C. Wolf (Eds.), *Proceedings of the Fifth Triennial ESCOM Conference* (pp. 182–185). Hanover, Germany: Hanover University of Music and Drama.

Winold, H., Thelen, E., & Ulrich, B. D. (1994). Coordination and control in the bow arm movement of highly skilled cellists. *Ecological Psychology, 6,* 1–31.

Zimmerman, B. J. (1998). Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-Regulated Learning: From Teaching to Self-Reflective Practice.* London: Guilford Press.

CHAPTER 6

# STRATEGIES FOR ENSEMBLE PRACTICE

JANE W. DAVIDSON AND ELAINE C. KING

Within the music profession, ensemble rehearsals are commonly geared toward public performance, most typically with the focus of attention being on the achievement of musical fluency and group coordination. Ultimately, this relies on high-quality practice for motor, cognitive, and social skill development, but it is well documented that musical knowledge, even among professionals, varies in degree, with some players being for example "better" at sight-reading or manipulating the expressive features of the music. As a result, the absolute time involved in rehearsing specific pieces can differ enormously. Indeed, there are clearly no prescriptive rules for how much practice is to be done, and in this sense, the advice offered in Chapter 5 (that the person, task, and context must be considered when deciding on appropriate practice strategies) applies equally well to the ensemble situation. Nevertheless, within ensembles, it is vital that each person not only exploit his or her own individual skills, but think beyond them for the sake of the group's musical and interpersonal cohesion—at the very least, to an extent that will convince audiences. What is certain is that, for the music to be negotiated between the players and then communicated effectively to the audience, two levels of knowledge must be integrated and articulated.

The first is a general musical and social knowledge that provides performance rules and regulations based on historical, social, and cultural factors (e.g. scale systems, harmonic and melodic expressive rules, and sociocultural interaction and performance practices). Whilst this knowledge is fairly genre specific (with art music, jazz, pop, folk, etc. all emerging out of different musicohistorical traditions, and so, different knowledge bases), it has a sense of being *permanent knowledge* for both the producer and perceiver of the music. It forms, therefore, a stable base on which a musical performance is built and received. Although such knowledge is a requisite for all types of performance, be it alone or in a group, the social psychological aspects of it are of paramount importance to ensemble practice and performance.

The second is the specific moment-by-moment information that must be processed and responded to in an ongoing manner—for instance, accommodating a coperformer's sudden change of tempo or coping with a memory slip. Of course, the ability to deal with moment-by-moment novelty depends on familiarity with other similar situations and, hence, relies on knowledge of a permanent sort. Thus, although each performance is a unique confluence of new interactive elements in the moment-by-moment process, familiarity with such experiences can help prepare for likely scenarios

and outcomes. Consider, for instance, any musical tradition in which improvisation takes place (e.g. jazz or Indian classical music), where the musicians learn a repertory of musical materials and stock-cultural practices about the length and style of improvised solos. When the music is being created afresh, it emerges from a context and depends on shared knowledge, so that coperformers and audience alike can experience the unfurling of musical ideas (this is akin to the long-term working memory theory of Ericsson & Kintsch, 1995; see also Chapter 7).

Working on the basis that musicians and audiences use and constantly develop the two forms of knowledge described above, the current chapter explores ways in which ensemble rehearsals operate, and how they may be facilitated for public performance. Largely because of existing research, the central focus is on classical music ensembles, ranging from piano duos through to the orchestra. Of course, all ensembles depend on the successful interaction of individual members; therefore, the chapter begins with a discussion of the dynamics of group membership. This is followed by an analysis of the rehearsal situation itself (both the general structure of rehearsals and approaches to individual pieces), as well as the functions of verbal and nonverbal communication. To close, two case studies of ensemble practice, with two very different types of group membership, are presented so as to elucidate the specific practical recommendations arising throughout the chapter.

## 6.1 Group dynamics

It is obvious that in ensembles the players must have a common connection to the *music*, and research suggests that this is the primary source of cohesion for the group (Davidson & Good, 2002). In small ensembles, cohesiveness can be supplemented by factors such as the potential for equal musical input and the fact that all players use essentially comparable instrumental techniques (e.g. in string quartets or flute choirs; Davidson, 1997). Even when the ensembles are comprised of quite different instruments, as in the case of a wind quintet where the French horn is juxtaposed alongside a flute, the complementarity of the instrumental timbres is often regarded by the group members as a cohesive element too (Ford & Davidson, 2003). Thus, similarity and complementarity can contribute to ensemble cohesion from a compositional/instrumental point of view. However, is the same true in terms of the social, interpersonal relationships?

Social psychological research suggests that, in order for any group to function, the members need to have a personal sense of "affiliation" (Douglas, 1993). That is, there needs to be a connection or association between individuals and to the group as a whole. The existing music research literature suggests that this affiliation is highly dependent on an individual's sense of self within the ensemble. Young and Colman (1979), who investigated string quartets, point out that group closeness achieved through trust and respect of each individual's boundaries is absolutely vital in order for quartets to continue functioning, with leaders and followers conforming to their interactive roles, which are mirrored in their musical roles.

Previous research, however, indicates that different types of ensembles function in unique ways. Work on piano duos shows that personal friendships seem critical to

the functioning of the rehearsal and performance situation (Blank & Davidson, 2003), with male–female or female–female dyads working far more successfully than male–male partnerships. They argue that this is because men in Western culture in general are not encouraged to be in working relationships with close friends, and so find the intimacy of the physical proximity of the piano duo difficult.

At the other extreme, Atik (1994) has shown that the orchestra, which itself is comprised of smaller sections, elicits very different group interactions. Atik discovered that, in this context, the players' sense of self has to be sublimated to a certain extent in order to facilitate the sense of group affiliation. Since between 50 and 100 people make up an orchestra, with section leaders and a conductor, it is evident that complex and subtle interactions are required in order for the ensemble to function successfully. An individual occupying a relatively low rank within the hierarchical structure may be asserting an inappropriately prominent voice to make his or her personal views common knowledge within the context of a whole orchestra rehearsal; here, the role of section leader—as a filter to the conductor—is the mechanism through which the player can express views.

- The prerequisite for effective ensemble rehearsal is that the operational principles of the ensemble are established, understood, and complied with. This depends, of course, on the nature of the ensemble itself, as well as the task.

In small ensembles, both conflict and compromise have been commonly found (see Young & Colman, 1979; Murninghan & Conlon, 1991). That is, when one person disagrees, the best solutions have been shown to come about through compromise, rather than one individual taking a dictatorial approach. Generally, in groups of four individuals, there is the potential for a three-to-one conflict, where one player can literally be overruled by the others; worse still, a two-to-two deadlock can occur. So, ironically, despite their prevalence as music groups, quartets can be problematic ensembles in terms of social dynamics and functioning. Intriguingly, although a democratic approach is generally favored, it has been found that the second violinist—above all other players—is the one who needs to feel respected and encouraged; if he or she is overwhelmed by the opinions of the others, especially that of the first violinist, the group has been shown simply to break down. It has been argued that this is because the two violinists need similar skill levels and play the same instrument, but owing to the nature of the part-writing, the first violinist is often perceived by others to be the superior player (Davidson & Good, 2002).

- In rehearsals, it is important that every voice is heard, or at least for every individual participant to feel that he or she can contribute as desired.

In a slightly larger ensemble—the wind quintet, for instance—members appear to be more relaxed about their roles. Perhaps this is because the ratio of disagreement is rarely four-to-one; any disagreement is more likely to feel like a group split, rather than one against the group. However, as mentioned earlier, the complementarity of the timbres and techniques of the instruments within a wind ensemble may well ensure that each person feels his or her own sense of belonging to the group.

## 6.2 The rehearsal situation

### 6.2.1 General structure of rehearsals

The existing research literature on the structure of a rehearsal has shown that approaches are widely varied (see Davidson, 1997). However, those of a professional standard share several overriding traits (see Goodman, 2000). Firstly, there is a rough plan of the material to be covered across the span of rehearsals in preparation for performance. Secondly, there is always room in the schedule for extra practice, accounting for unexpected circumstances, such as a player being unable to attend a rehearsal. Of course, very often when people come together, they are surprised that some pieces take longer to rehearse than others. Blank and Davidson (2003) discussed with a top piano duo the unexpected difficulties they faced when preparing Bartók's Sonata for Two Pianos and Percussion. One player intriguingly commented:

> We could not believe how difficult it was not only to play the notes, but to coordinate that second bar of the opening movement. We spent most of our rehearsal time on that . . . but having said that, the act of concentrating on that singular difficulty seemed to facilitate the preparation of the rest of the piece. We, if you like, faced the most difficult challenge, and that put everything else into context (personal communication; taken from the interview transcript but not quoted by Blank & Davidson, 2003).

This comment is revealing, for it shows that it is often the need to find an appropriate manner for approaching the rehearsal process that takes time. The singular difficulty of the Bartók focused everyone's attention, and thus they had all developed a similar level of concentration and awareness of the difficulties involved in the piece. The comment has a flip side, which is that often ensembles do not achieve such a coordinated manner of working, with the consequence that rehearsal time may end up unfocused and unproductive. In Davidson and Good's (2002) string quartet research, there was a difficult social dynamic in the rehearsals in which one member had an obsession with tuning, while the others were more interested in working on entrances and exits of the individual lines of the pieces being rehearsed. Although the group adjusted to this individual particularity, the researchers observing the rehearsals sensed that the player who continually tuned during discussions created a distraction to the main musical activity.

- A key rehearsal aim is for you to establish shared musical goals within the ensemble as soon as possible (be they of a technical or expressive nature). In relation to the literature on group dynamics, it is also necessary for you to find a means of coordination that leaves room for individual needs and styles.
- Leave space in your practicing schedule for unexpected difficulties, such as a piece being more challenging than expected or unanticipated absences of ensemble members.

In addition to the overall context of a series of rehearsals, the content of single rehearsals is critically important. Specific strategies for effective single rehearsals have been suggested. Goodman (2000) found that a warm-up routine can provide a helpful focus for groups. Vikram Seth's (1999) novel *An Equal Music*, in which the members of

a professional string quartet are described warming up, seems to capture much of the sentiment found in Goodman's data:

> Every rehearsal of the Maggiore Quartet begins with a very plain, very slow three-octave scale on all four instruments in unison: sometimes major, as in our name, sometimes minor, depending on the key of the first piece we are to play. No matter how fraught our lives have been over the last couple of days, no matter how abrasive our disputes about people or politics, or how visceral our differences about what we are to play and how we are to play it, it reminds us that we are, when it comes to it, one... No one appears to lead. Even the first upbeat is merely breathed by [the first violinist], not indicated by any movement... When I play this I release myself into the spirit of the quartet. I become the music of the scale. I mute my will, I free my self (p. 12).

In addition to warming up together, Cox (1989) has stressed the role of the pace of rehearsals. In particular, there needs to be periods where players are required to concentrate intensely, such as in a very fast passage, which is then set against less intense work, such as running through for an overview of a section or the entire piece. This is what Blum (1986) considers to be the "balance" of key components of the rehearsal (e.g. sight-reading versus working on familiar pieces or parts therein). Allied to this is the fact that, according to most of the existing studies, all players need to be equally involved in the rehearsal. Durrant (1994), for instance, found that in conducting situations, the conductor could easily lose control if players and/or singers showed signs of boredom after having been neglected. Thus, sustaining the interest of all participants seems to be critically important. In summary:

- For the small ensemble, warming up together can bring focus and unity to the rehearsal.
- Balance the key components of practice, in order to achieve both technical progress and musical fluency.
- Make sure that each member of the ensemble is, in some way, engaged in the rehearsal musically and mentally, so as to prevent boredom or a sense of exclusion.

For the professional and advanced student, a concert appears to be an obvious key focal point for rehearsals. Indeed, Blank and Davidson (2003), Goodman (2000), and Murninghan and Conlon (1991) all found that the concrete goal of a performance on a specific date aided focus and effectiveness. However, this is where professional and amateur players differ. Blank and Davidson (2003) found that for the amateur player, it was the rehearsal itself that was the pleasurable activity and not the planning for a specific goal. Within the rehearsal time, there was an emphasis on "fun" and "chat," with the session often lasting several hours and being broken by coffee breaks. The selection of repertory covered was also less goal-oriented and included familiar pieces being played through to give a sense of "reward," alongside some time spent tackling new and difficult pieces for a smaller proportion of time. This amateur approach has been well-reported by Booth (1999) in a personal account of the joys of music-making for fun. For him, the social aspect is clearly as important as the musical. The question that each practitioner should ask when thinking about

a rehearsal approach is:

- Are your rehearsals strictly for the pleasure of participation and for entertainment, or are they in preparation for professional-level performance? Depending on the answer to this question, the focus and structure of rehearsal may be very different; however, achieving a balance between pleasure and work is important for all ensembles, and healthy music making generally.

### 6.2.2 Approaches to individual pieces

The particular demands of individual pieces often mean that effective practice strategies can be as many and varied as performers themselves. Goodman (2000, 2002), Gruson (1988), Miklaszewski (1989), and Williamon *et al.* (2002) all discovered that there can be many differences in approach to a piece. Some groups only play through from beginning to end, whilst others tend to stop and start to focus on "tricky" passages. Hallam (1995, 1997) has examined these differences in strategy in detail and, although she does not favor one type of strategy above another, she comments that they are indicative of individual differences in cognitive style: either a "holist" or "serialist," respectively. Still, looking at the different approaches found to exist, ensembles can work in the following manner to achieve thorough and accurate learning:

- Identify and work within and up to structural boundaries in the music. This immediately highlights the form of the piece, and so can strengthen your ensemble's shared understanding of it (Chaffin *et al.*, 2002; Miklaszewski, 1989; Williamon & Valentine, 2002).
- Balance the focus in rehearsal between long and short segments of the piece. Doing this within a piece can help sustain concentration and involvement in the rehearsal (Goodman, 2000) and has been shown to be a salient characteristic of effective learning (Williamon *et al.*, 2002).
- Integrate "run-throughs" during the rehearsal process, for it is important to put expressive ideas and/or technical problems into the context of the whole. This can be important for reinforcing memory and consolidating common tempi, dynamics, and other expressive elements (see Chaffin and Imreh, 1997; Chaffin *et al.*, 2002).
- Consider working both sequentially and nonsequentially through an individual movement or composition. These strategies can help you place portions of the music in their chronological context and also allow common material, or sections of similar levels of difficulty, to be explored. Goodman (2000) found that different ensembles tended to use one or both of these strategies, but the former was particularly used to generate new thoughts about sub-sections of a work.

Other strategies that could be explored include:

- When time is short, have an "economical rehearsal approach"—that is, only work on passages that need attention, so as not to waste time going over parts that are sufficiently grasped.

- For chamber groups: you can save much time and effort if all players have access to a copy of the full score and even take time to go through the score together. Typically in chamber music, individual parts are handed out, so unless each person in the ensemble can be acquainted with the others' parts through repeated playing, it is often difficult for one instrument to know what is going on around them. This strategy has been found to be useful and a good shortcut to lengthy rehearsal processes (Goodman, 2000).
- For conductors: you need to think about potential areas of difficulty in the piece to ensure technical security and to confirm that your interpretation is robust (to know what you want from the players in each passage and for the player to understand what is required; Durrant, 1994).

From the existing literature, it is evident that despite individual differences, a clear rehearsal method can be prescribed to ensure positive outcomes. Nonetheless, these strategies are offered with some caution, since all players have to display similar levels of commitment to the rehearsal and have the same goals—whether it be for social amateur entertainment or for professional engagement.

## 6.3 Verbal and nonverbal communication

### 6.3.1 Between a conductor and ensemble

There can be various types of "talk" used in rehearsals. Directions such as "find letter 'A' in your score" or "let's try that again but at a slower tempo" are the most common types of exchanges. Typically, when there is a conductor, he or she leads the speech. Time and again, however, ensemble research indicates that too much talking, either from the conductor or in the form of questions from an individual player, disrupts the flow of the music in rehearsals (see Yarbrough, 1975; Durrant, 1994; Weeks, 1996; Price & Byo, 2002). That is, if too much questioning is permitted from one individual, it is likely to lead to interference as others may become irritated or distracted.

Obviously, talk cannot be avoided. Instructional talk, however, seems acceptable, whereas discussion is often less welcomed (although there needs to be space for this). In the larger ensemble context, discussion should be carefully managed, for it is easy for the talking to turn into confrontation. Davidson (1997) gives an example of this through an incident with Leonard Bernstein and José Carreras when making a television documentary on the CD recording of Bernstein's *West Side Story*. Carreras was singing the role of Toni and had mislearned or could not sing accurately a part of one of his songs. Beginning relatively patiently, Bernstein, who was conducting, pointed out the tenor's error. Bernstein stopped when the tenor repeatedly made the same mistake, but after several corrections in front of the whole orchestra, Bernstein became extremely vexed and asked the tenor why he could not do this simple thing. The tenor, furious and shamed by the situation, swore profusely and stormed out of the rehearsal.

- It is evident that it is best for the conductor not to allow a situation to escalate into confrontation, otherwise there can be dire consequences. By contrast, of course, the odd word of praise, especially to an individual within an entire group setting, has

been found to be extremely beneficial. However, even an individual comment needs to be moderated, otherwise other group members could feel undervalued.

Verbal instructions can be necessary, but are best minimized. Talk with any sort of social dynamic context needs to be carefully monitored, as the stakes for all parties can be high (see Davidson, 1997, for a further exploration of this issue). In the light of the comments above, it is not surprising to learn that, although both verbal and nonverbal indicators to others can be employed in rehearsals, conductors are at their most efficient when using nonverbal cues (Durrant, 1994). The conductor, of course, typically controls the ensemble through the use of gestures, and the rehearsal period is the time for the players to learn the meaning behind those gestures.

The extensive literature on general nonverbal communication has demonstrated that we typically use four main categories of gesture in our interactions:

- *adaptors*: movements of self-stimulation (e.g. head-scratching and touching the finger to the bridge of the nose)
- *regulators*: movements that allow for entrances and exits to speech or movement (e.g. using a hand movement to encourage someone to make a point during a discussion, or using an open arm gesture to indicate to someone to walk ahead and pass through the door first)
- *illustrators*: self-explanatory gestures of emphasis with direct musical or speech translations (e.g. making a cradling arm movement when talking about a baby)
- *emblems*: culturally defined gestures (e.g. the "V for Victory" sign used by Churchill to the British troops during the Second World War; see Ekman & Freisen, 1969, for further discussion).

In performance where a conductor is involved, regulatory gestures and illustrators are obviously required in order to indicate when and how something should happen. Durrant (1994) has found that some conductors use emblems extensively. These, however, are often very elaborate codes, and part of the rehearsal period must then be concerned with performers learning and quickly decoding what these emblems signify. The emblems are usually complex signs intended to substitute for spoken technical and expressive instruction (e.g. using the hand at the side of the head in a rotational gesture to indicate the notion of creating "space" for the singing voice to sound, and accompanying this with a sharp downward pointed gesture to indicate the vocal attack on the sound). Although there is no formal research to date, Davidson (1997) has hypothesized that the most successful conductors on the international professional circuit are those who not only keep to nonverbal interaction, but who are also efficient in their use of illustrators and regulators and keep emblems to a minimum, given the time involved in learning their meanings. In choral music, it is more usual for a conductor to work with the same ensemble for long periods, so a common understanding of emblems can be built up over time.

◆ An overly zealous, gesticulating conductor can become a distraction (Davidson, 1997). Scrutinize your movements and consider whether these can be easily and quickly deciphered by others.

- Besides being able to give instructions using gestures, you need to read cues from the ensemble. Be aware of possible signals (in practice and performance) that are used by musicians in your ensemble, and determine which are consistent across the entire group and which are individual-specific. In doing so, you will be able to work toward establishing an effecting three-way communication between yourself, the ensemble, and the audience.

### 6.3.2 Between coperformers

Coperformers need to be able to decipher what is going on both within and between musical parts in order to achieve a fluent performance. Davidson and Good (2002) discovered that there are two major categories of interaction between string quartet members: (1) those of musical content and its coordination (e.g. achieving entrances and exits) and (2) those of a more personal nature (e.g. the first violinist making a head-nodding gesture of approval to the second violinist after a particularly well-played passage). Davidson and Good reported a power struggle going on in a rehearsal based principally on sexual politics (one flirtatious male among three compliant, admiring females), and it was noted that the flirtation worked around the use of nonverbal gestures. The male teased the first violinist about her lack of clarity, or even the sudden jerkiness of her regulatory gesture to say "start the upbeat, now." Ironically, the two researchers did not find her cue remotely ambiguous, but the example demonstrates how multifunctional the role of nonverbal communication can be on both a musical and social level in the rehearsal.

Williamon and Davidson (2002) found that in the practice of two professional pianists forming a duo, the quantity of nonverbal interaction increased significantly from first meeting through to a performance. Here, the most used movement was a slight head-and-eye indication for regulation. Also and somewhat intriguingly, through the course of the rehearsals, the movement styles of the two individuals coalesced. The player who moved the most moderated and reduced the quality of his movements somewhat, while the more conservative player produced more movement. It was as if they began to move as one, rather than two individuals. According to interviews with the pianists, this was for the sake of musical coherence. In fact, these physical expressions provided a means of understanding and sharing musical intentions, as they were most coordinated at key points in the music's structure (as identified by the pianists); for an audience, this level of coordination can give an additional access point to the music's meaning (Davidson, 1993). The musicians' movements also seemed to have a social dimension as well. Evidently, they sensed that they were performing the pieces "together," as one "unit." Davidson and Good (2002) also observed this with a string quartet playing Britten's *Rhapsody*. They found it much easier to achieve the canonical musical entries if they were all ebbing and flowing in exactly the same manner.

- Movement can help to bring about musical coherence and social unity. Consider the movements made within your ensemble. Are these readily recognized and shared among the group? Are they timed to correspond with certain points in the music? Will audiences consciously notice and be able to interpret them, or are the movements purposefully too subtle for this?

Goodman (2002) and Murninghan and Conlon (1991) found that nonverbal communication during playing often solved problems where talk had failed. By simply playing and sorting out coordination "on the hoof," there was often no need to discuss a particular point at all. Alternatively, when a verbal disagreement about musical interpretation emerged, playing the piece through in several different ways would often lead to a decision without verbal discussion (or at least nothing more than exchanging a few syllables).

Finally, it is important to note that nonverbal cues can be discussed in rehearsal too. Blank and Davidson (2003) interviewed one duo partnership who said that, early on in their career together, they would discuss where they were going to give such cues. As time went by and familiarity with their musical repertory grew, the cues often did not occur as they had done originally, and new ones emerged.

- It is important for you to think about how your ensemble is to use talk and nonverbal gestures in rehearsal. Strategies based on the regulator and illustrator can help with musical coordination, but eye contact and smiles (found in the rehearsals recorded by Goodman, 2000, and Williamon & Davidson, 2002) can be extremely important in helping you provide clarity in a message and express the degree of satisfaction between ensemble members.

## 6.4 Case studies

The case studies presented here are on two cello–piano duos. Both duos were asked to rehearse and perform (at the end of their rehearsal) the second movement (Minuet and Trio) of Brahms's Sonata for Piano and Cello, Op. 38. The players in both ensembles were unfamiliar with one another before the investigation and were matched together according to age, playing standard, and experience. None of the musicians had played this particular movement for at least 5 years. Neil (cello) and Matthew (piano) were grouped together as professionals, and Mary (cello) and Mikako (piano) were grouped as advanced students (i.e. they had each recently completed undergraduate studies at a conservatory). One month before the rehearsal, they were sent copies of the movement and were asked to learn the relevant part in preparation for the rehearsal. For the purposes of analysis, the rehearsals were videotaped and transcribed by documenting all verbal and pertinent nonverbal communication. Three aspects of the duo rehearsals are addressed below: rehearsal method, communication, and social interaction (the complete method and set of analyses are reported in Goodman, 2000).

### 6.4.1 Rehearsal method

In the first instance, the duos began the rehearsal by playing through the movement in its entirety without stopping. Thereafter, their practice proceeded in quite different ways, although in both cases this involved breaking down the music into long and short segments in order to address specific issues. Interestingly, the performers did not discuss *how* they would tackle the piece, or propose a specific agenda for the rehearsal, but simply got on with the task of playing the movement. On the one hand, Mary and Mikako rehearsed the piece in a sequential fashion by working progressively from

beginning to end. On the other, Neil and Matthew maintained a nonsequential agenda, in so far as they isolated certain parts of the score throughout the rehearsal with no particular concern for the movement's chronological progression. They used their initial run-through of the movement as a benchmark or baseline for discussion, and they modified, clarified, and reshaped aspects of their playing accordingly. These different styles of rehearsing the music are illustrated in Figure 6.1.

In view of the contrasting styles of approach undertaken by the two ensembles, it is useful to speculate which of the two is the most effective. First, it is possible to suggest that a nonsequential method of practice is more economical in terms of time than a sequential approach, for Neil and Matthew's rehearsal lasted 30 minutes while Mary and Mikako's rehearsal was strikingly longer at 80 minutes. Arguably, however, the total duration of the rehearsal is not an indication of its quality, but rather a reflection on how the performers have been trained to practice and how they decided to apply their training in the context of the new partnership established in this study (cf. Krampe & Ericsson, 1995). It is plausible to suggest, however, that the professional musicians, Neil and Matthew, were used to working under time constraints and, therefore, adopted a "quick-fit solution" approach to the rehearsal, while Mary and Mikako wanted to take the time to ensure that all corners of the music were worked out in rehearsal.

Second, the sequential agenda could be regarded as more demanding or thorough because it allowed for a step-by-step approach to the movement. Mary and Mikako seemed to explore details of the piece and inspire each other to consider ideas about the piece *en route*. In effect, the sequential method allowed for the generation and confirmation of ideas in the rehearsal. By contrast, in studying select areas of the movement in relation to their initial run-through, Neil and Matthew adopted a more retrospective or reflective style. To this end, the nonsequential agenda could be regarded as more "risky," in the sense that the performers did not address every part of the movement in rehearsal and, hence, left a certain amount of passagework to the whim of the performance. As professional musicians, however, Neil and Matthew obviously felt comfortable with their preparation, and they had the confidence in one another to be able to perform the movement at the end of the rehearsal.

There is inevitably a degree of overlap between these two rehearsal methods, and it is possible that they could be combined quite successfully. Indeed, other data from Goodman's (2000) research showed that one of the duos rehearsing the Prologue from Debussy's Cello Sonata adopted a nonsequential agenda and then switched to a sequential approach midway through the rehearsal. Nevertheless, in the cases discussed here, the ensembles achieved the goal of performing the movement but adopted alternative methods of preparation, both of which can be regarded as appropriate in their own right.

### 6.4.2 Rehearsal communication

The performers communicated using verbal and nonverbal discourse, the latter including singing, humming, counting aloud, gesturing, and clicking the pulse. There was a relatively high amount of talking, as the musicians made many of their ideas explicit through verbal discourse either before or after playing segments of the

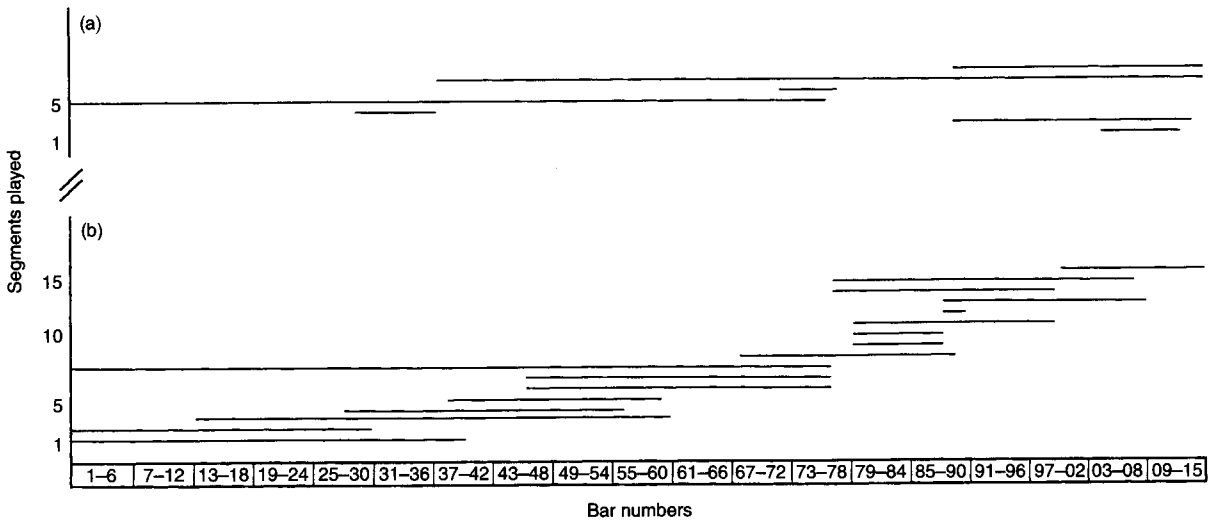



**Figure 6.1** Sample transcriptions of (a) the non-sequential practice of Neil and Matthew and (b) the sequential practice of Mary and Mikako in their respective rehearsals of the second movement of Brahms's Sonata for Piano and Cello, Op. 38. Horizontal lines (read from bottom to top) represent segments of the music played, which correspond to the bar numbers listed along the x-axis (graphical method adapted from Chaffin & Imreh, 1997).

music (above all, the relatively high degree of talking seems to reflect the musicians' unfamiliarity with one another). Nonetheless, some issues were resolved "tacitly" while playing. For example, Neil and Matthew realized that they were "not together" at the end of the Trio, but the problem of coordination ironed itself out in a subsequent run-through of the section, so there was no further discussion about the point.

Of particular interest, however, are the issues that the performers addressed verbally, and the manner in which they negotiated their ideas. The musicians in both ensembles divided their rehearsal time relatively equally between the Minuet and the Trio. There were several common issues of consideration: the dynamics in the Minuet (particularly across the climactic build up in bars 50–58), the rubato in the Trio, and the tempo at the end of the Minuet. The performers discussed these issues in different ways, thus highlighting the potential variety and complexity of musical interaction manifest in chamber rehearsal. For example, the realization of the *crescendo poco a poco* in bar 50 of the Minuet caused varying problems for them (see Figure 6.2). Matthew suggested that the music should be played more softly at bar 50 in order to enhance the subsequent build-up: "start that bit quietly, then we can make more of the crescendo." Indeed, Neil confessed that he found the cello part difficult to interpret at that point ("I've got a complex about that bit") and admitted that he originally conceived the start of the crescendo a few bars earlier (i.e. bar 48). Neil readily conformed to Matthew's suggestion to "play quieter" at bar 50. In this case, negotiation was achieved easily on the basis of adherence to the dynamic indication in the score, and Matthew's practical direction.

Mikako made a similar request to Mary in their rehearsal of the same passage: "I know you have a crescendo, but can you do it . . . in a softer way?" In contrast to Neil and Matthew's relatively quick resolution of the point, Mary and Mikako discussed the

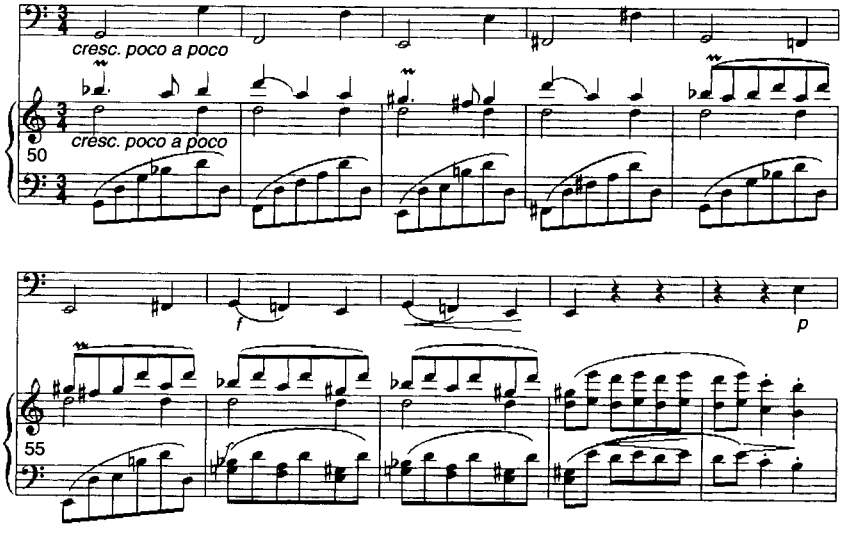



**Figure 6.2** Bars 50–59 of the Minuet from the second movement of Brahms's Sonata for Piano and Cello, Op. 38.

issue at length. Mikako mentioned the difficulty of exposing the melody in her right hand because of its high register. For this reason, she suggested that the cellist should delay the crescendo until bar 54 and exemplified her point by talking and playing through the passage. Essentially, she wished to use dynamic contrast to highlight the melodic repetition in the piano part: bars 50–51 are echoed by bars 52–53, which are quieter, and from bar 54, the crescendo begins with the cello. By contrast, Mary conceived a "broad flow forwards. . . not crescendoing too early, and doing it really with a very fine gradient." To this end, she claimed that Mikako's dynamic fluctuations contradicted the marking in the score, "which is. . . just to crescendo really gradually". In order to negotiate their conflicting points of view, Mary and Mikako played the passage in several different ways, hence using trial-and-error. In addition, they gave further attention to performance indications in the score, discussed the music's process (i.e. the concept of build-up across the passage), and reconsidered the nature of the melodic line in the piano part. There was a certain degree of compromise by both performers in order to reach a workable solution: Mikako conformed to Mary's notion of a broad build-up, while Mary agreed to play quietly at the beginning of bar 57 to exaggerate the hairpin dynamic marking in her part. Similarly, Mikako accepted Mary's conception of the arrival point at bar 58 as "louder than *forte*," and they both agreed that the dynamic level at bar 50 should be reduced.

The above example reflects the potential scope of communication between performers about a single issue, and the fact that there are alternative ways in which musical negotiation can arise. In this case, the two ensembles used the following strategies to negotiate their ideas:

- awareness of score indications
- personal judgment based on trial-and-error
- analytical reasoning
- consideration of the music's form as process.

Other types of negotiation are reported by Goodman (2000) in the duo rehearsals, including practical reasoning and reconciliation of emotional insights. In both ensembles, the most effective strategy for resolving the issue of dynamics in the Minuet was to follow the indications in the score. Still, the other strategies usefully supported the discussion of this issue (particularly for Mary and Mikako), which suggests that a wide range of strategies should be explored to reach a workable solution when members of an ensemble are in disagreement.

### 6.4.3 Social interaction

From an analysis of the performers' discourses, it is evident that the majority of social interaction in both ensembles was task-related, notably through "giving opinions" about the music. This can be regarded as quite characteristic of small-group behavior in general (see Bales, 1950, 1999). At the same time, a high level of agreement existed in the rehearsals, although phatic acknowledgements mainly consisted of "yeah." In fact, the amount of agreement expressed in the performers' discourses was tremendous (as

compared with that witnessed in other, nonmusical small group scenarios; see Bales, 1999), which indicates that the performers gave mainly positive socioemotional responses. Once again, the musicians' unfamiliarity with one another is a likely explanation for this point, as they possibly felt more inclined to express positive reactions in order to establish a "friendly" group dynamic (undoubtedly, the relationship would develop in time, and it is possible that negative reactions would be expressed more frequently in ensuing rehearsals).

It is apparent that Neil dominated much of the task-related discussion in his rehearsal, for he gave more suggestions, opinions, and orientation than Matthew. Indeed, Neil initiated a majority of the points or areas of discussion (e.g. he mentioned "What I'd like to do first is actually the very end of the Trio"). By contrast, Matthew asked for more orientation, which perhaps reflects his desire, as an accompanist, to ascertain as much information from the cellist as possible. In general, however, Neil and Matthew generated a very relaxed and positive atmosphere in their rehearsal, not least because there were no signs of negative tension or antagonism; they simply got on with the rehearsal.

The interaction between Mary and Mikako appeared to be equally positive on the surface, although the relationship between them soured as the rehearsal progressed. There were several moments of disagreement, which created negative tension. Mikako offered more suggestions, opinions, and orientation in the rehearsal, and Mary produced more utterances of agreement and disagreement. In fact, Mikako appeared to lead the discussion by virtue of the fact that she did most of the talking and contributed fully to the realization of musical ideas. Mary, however, showed signs of discontent with Mikako's persistent verbal explanation of points and her constant desire to stop and start, for Mary clearly preferred to concentrate on letting the music "just play itself." There was a certain amount of conflict here between "musical" personalities (one wanting to analyze the musical detail intently and the other wanting simply to let the music happen), as well as differences in perception of each other's standard. The social dynamic was invariably strained.

Nevertheless, Mary and Mikako continued to offer supportive opinions to one another throughout the rehearsal, even if they were used as a tactic to soften critical suggestions or disagreement. For instance, Mary remarked, "[the dynamic] makes a really nice contrast, but can we try it just once where we follow exactly what it says in the copy. . . ?" and "I know that you're really playing out your tune and I think that's really good, but I think it's a little bit too much. . ." Likewise, Mikako commented, "I thought that was good, but if you could come in less. . ." These utterances reflect the performers' sensitivity toward each other, for they seemed wary of giving comments without offering solidarity first.

The performers in both of these ensembles thus maintained working relationships by sustaining a "friendly," positive group dynamic. This was primarily achieved by ensuring that the discussion was focused on the task at hand and that opinions were offered sensitively. However, the rehearsals were led by different instrumentalists, Neil (cellist) and Mikako (pianist), which indicates that the stereotypical hierarchy between a soloist and an accompanist does not necessarily dictate the social relationship between performers in rehearsal.

## 6.5 Conclusions

The case studies above demonstrate that there is no single "best" method or strategy for rehearsing a particular piece of music—rather, that individual ensembles will necessarily find their own "best" ways of rehearsing. Here, it has been revealed that two ensembles can effectively rehearse the same piece in contrasting ways in order to achieve the same goal, whilst adopting alternative approaches and using different strategies to negotiate musical ideas. At the same time, the rehearsal must be underpinned by a strong social framework, with musicians working hard to achieve a positive socioemotional ambience through sensitive and friendly behavior.

Undoubtedly, musical interaction in rehearsal is most effective and pleasurable when musicians "click" with one another at both musical and social levels, but when this is not possible, can an ensemble still be successful? Research data reveal that a social framework can be sustained in different ways, yet still enable a group to achieve its goal. In the case studies, one partnership "clicked" quite comfortably, and the musicians simply proceeded with the task at hand. The other duo experienced conflict, resulting in a much more tense rehearsal environment. Nonetheless, both ensembles achieved their goal of successfully performing Brahms's Minuet and Trio.

The crucial advice for performers—be they in small or large ensembles—is to develop a greater awareness of the social psychological principles that govern group interaction and cohesion. At the same time, players should reflect over the effectiveness of their rehearsal methods and ways of communicating. Such knowledge will invariably offer new perspectives in approaching and subsequently optimizing ensemble practice and performance.

## Further information and reading

Blum, D. (1986). *The Art of Quartet Playing: The Guarneri String Quartet in Conversation with David Blum.* New York: Cornell University Press.

Cox, J. (1989). Rehearsal organisational structures used by successful high school choral directors. *Journal of Research in Music Education, 37,* 201–218.

Goodman, E. (2002). Ensemble performance. In J. Rink (Ed.), *Musical Performance: A Guide to Understanding* (pp. 153–167). Cambridge: Cambridge University Press.

Murninghan, J. K., & Conlon, D. E. (1991). The dynamics of intense work groups: A study of British string quartets. *Administrative Science Quarterly, 36,* 165–186.

Price, H. E., & Byo, J. L. (2002). Rehearsing and conducting. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 335–351). Oxford: Oxford University Press.

Williamon, A., & Davidson, J. W. (2002). Exploring co-performer communication. *Musicæ Scientiæ, 6,* 53–72.

## References

Atik, Y. (1994). The conductor and the orchestra: Interactive aspects of the leadership process. *Leadership and Organisation Development Journal, 13,* 22–28.

Bales, R. F. (1950). A set of categories for the analysis of small group interaction. *American Sociological Review, 15*, 257–263.

Bales, R. F. (1999). *Social Interaction Systems: Theory and Measurement.* London: Transaction Publishers.

Blank, M., & Davidson, J. W. (2003) A consideration of the effect of gender on co-performer communication in piano duos. Manuscript submitted for publication.

Blum, D. (1986). *The Art of Quartet Playing: The Guarneri String Quartet in Conversation with David Blum.* New York: Cornell University Press.

Booth, W. (1999). *For the Love of It: Amateuring and Its Rivals.* Chicago: University of Chicago Press.

Chaffin, R., & Imreh, G. (1997). "Pulling teeth and torture": Musical memory and problem solving. *Thinking and Reasoning, 3*, 315–336.

Chaffin, R., Imreh, G., & Crawford, M. (2002). *Practicing Perfection: Memory and Piano Performance.* Mahwah, NJ: Erlbaum.

Cox, J. (1989). Rehearsal organisational structures used by successful high school choral directors. *Journal of Research in Music Education, 37*, 201–218.

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music, 21*, 103–113.

Davidson, J. W. (1997). The social in musical performance. In D. J. Hargreaves & A. C. North (Eds.), *The Social Psychology of Music* (pp. 209–228). Oxford: Oxford University Press.

Davidson, J. W., & Good, J. M. M. (2002). Social and musical co-ordination between members of a string quartet: An exploratory study. *Psychology of Music, 30*, 186–201.

Douglas, T. (1993). *A Theory of Groupwork Practice.* New York: Macmillan.

Durrant, C. (1994). Towards an effective communication: A case for structured teaching of conducting. *British Journal of Music Education, 11*, 56–76.

Ekman, P., & Friesen, W. V. (1969). The repertoire of non-verbal behaviour: Categories, origins, usage and coding. *Semiotica, 1*, 49–98.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211–245.

Ford, L., & Davidson, J. W. (2003). An investigation of members' roles in wind quintets. *Psychology of Music, 31*, 53–74.

Goodman, E. (2000). *Analysing the Ensemble in Music Rehearsal and Performance: The Nature and Effects of Interaction in Cello-Piano Duos.* Unpublished doctoral dissertation, University of London.

Goodman, E. (2002). Ensemble performance. In J. Rink (Ed.), *Musical Performance: A Guide to Understanding* (pp. 153–167). Cambridge: Cambridge University Press.

Gruson, L. M. (1988). Rehearsal skill and musical competence: Does practice make perfect? In J. A. Sloboda (Ed.), *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition* (pp. 91–112). Oxford: Clarendon Press.

Hallam, S. (1995). Professional musicians' approaches to the learning and interpretation of music. *Psychology of Music, 23*, 111–128.

Hallam, S. (1997). The development of memorisation strategies in musicians. *British Journal of Music Education, 14*, 87–97.

Krampe, R. T., & Ericsson, K. A. (1995). Deliberate practice and elite musical performance. In J. Rink (Ed.), *The Practice of Performance: Studies in Musical Interpretation* (pp. 84–102). Cambridge: Cambridge University Press.

Miklaszewski, K. (1989). A case study of a pianist preparing a musical performance. *Psychology of Music, 17*, 95–109.

Murninghan, J. K., & Conlon, D. E. (1991). The dynamics of intense work groups: A study of British string quartets. *Administrative Science Quarterly, 36*, 165–186.

Price, H. E., & Byo, J. L. (2002). Rehearsing and conducting. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 335–351). Oxford: Oxford University Press.

Seth, V. (1999). *An Equal Music*. London: Phoenix.

Weeks, P. (1996). A rehearsal of a Beethoven passage: An analysis of correction talk. *Research on Language and Social Interaction, 29*, 247–290.

Williamon, A., & Davidson, J. W. (2002). Exploring co-performer communication. *Musicæ Scientiæ, 6*, 53–72.

Williamon, A., & Valentine, E. (2002). The role of retrieval structures in memorizing music. *Cognitive Psychology, 44*, 1–32.

Williamon, A., Valentine, E., & Valentine, J. (2002). Shifting the focus of attention between levels of musical structure. *European Journal of Cognitive Psychology, 14*, 493–520.

Yarbrough, C. (1975). Effect of magnitude of conductor behaviour on students in mixed choruses. *Journal of Research in Music Education, 23*, 134–146.

Young, V. M., & Colman, A. M. (1979). Some psychological processes in string quartets. *Psychology of Music, 7*, 12–16.

CHAPTER 7

# STRATEGIES FOR MEMORIZING MUSIC

JANE GINSBORG

It is common in many kinds of music for artists to perform from memory. Within certain musical genres, in fact, social convention and audience expectation are such that performing with a notated score would be inconceivable. Rarely does one see a folk, rock, or pop musician reading from the printed page during performance; within the Western art music tradition (where audiences have come to expect almost note-perfect renditions of well-known pieces), memorized performances are customary for concert soloists. Why is this the case? Clearly, there are a number of practical advantages to playing or singing without a score. These include not having to turn pages, being able to monitor the physical aspects of a performance (e.g. looking at the hands or at other ensemble members), and in the case of concert pianists, avoiding the need to share the platform with a page-turner.

More importantly, perhaps, performing from memory is often seen to have the effect of enhancing musicality and musical communication. It is commonly argued that the very act of memorizing can guarantee a more thorough knowledge of and intimate connection with the music (see Plunkett Greene, 1912; Hughes, 1915; Matthay, 1926). In addition, memorization can enable the use of direct eye contact with an audience (and gaze, in other ways, for expressive purposes) that is more convincing than when referring to the score. Those who "possess" the music in this way often convey the impression that they are spontaneously and sincerely communicating from the heart, and indeed, contemporary evidence suggests that musicians who achieve this are likely to find their audiences more responsive (see Davidson, 1993, 1994; Williamon, 1999). Moreover, when performers receive and react to visual feedback from the audience, a performance can become truly interactive, involving genuine communication between all involved.

While performing from memory may have benefits for the performer and audience alike, the task of memorizing music and giving a memorized performance can be arduous. Memory failure—however momentary—can have catastrophic consequences for a performer's self-confidence. The musician does not, in fact, need to have experienced such failure to be frightened of forgetting, and this may well contribute to a more generalized performance anxiety. Furthermore, the fear of memory lapses is as common among experienced professional musicians as with novices—perhaps more so, as they feel they have more to lose. Artur Rubinstein, for example, "abandoned the concert stage mainly on account of the fact that he could no longer depend on his memory in public performance" (Hughes, 1915, p. 603).

This chapter reviews a range of memorization strategies that are available to musicians. An analysis of their efficacy is based on a critique of existing observational and empirical research, most of which stems from studies carried out within the context of Western classical music. Nevertheless, the benefits of each are presented so that they may be applied equally well to performance in other musical styles. The chapter is in three main sections. It begins by setting out basic principles of human memory, including a discussion of the psychological structures thought to be involved in memorization and some recommendations for improving memory generally. The second section explores musical memory specifically and offers a selection of strategies commonly used by musicians and tested through systematic research. These include rote memorization, methods for memorizing visual and auditory information, methods for developing and exploiting "conceptual" memory, and memorization for words and music. Finally, the chapter closes with a discussion of how these strategies can be used in conjunction with one another and how individual differences in learning styles may impact on one's capacity to memorize.

## 7.1 Memory: Structure and general recommendations

### 7.1.1 Structure

Before considering strategies that musicians can use to exploit their memory effectively and efficiently, it is important to place musical memory into a broader theoretical context. Memory can be conceptualized in a variety of ways. In Atkinson and Shiffrin's (1968) modal model, for example, memory consists of three stores, as depicted in Figure 7.1. The first is the *sensory store*, which has sub-stores for the different kinds of environmental information we receive via our senses: visual, auditory, tactile, and so on. Much of this incoming information is not attended to and, consequently, remains only briefly in memory. Information that does capture attention, however, is transferred into *short-term memory* (subsequent researchers have proposed that short-term memory has a *working memory* component that is used to manipulate different kinds of information in conscious awareness; see Baddeley & Hitch, 1974). Memory storage at this level plays a vital role in basic, everyday functioning (e.g. remembering phone numbers and keeping track of an ongoing conversation). In terms of musical performance, this type of memory is integral to tasks such as sight-reading and improvisation, where new information must be retained and related to past, present, and predicted future events (see Chapter 8 for further details on sight-reading and improvisation, Chapter 11 for research on working memory enhancement, and Chapter 14 for sources of memory impairment).

Once in short-term memory, information can be rehearsed and elaborated in such a way that it is stored alongside or assimilated into existing knowledge. Unlike short-term memory, which has a capacity limited to around seven "chunks" of information (Miller, 1956), *long-term memory* seems to be of unlimited capacity and duration. It holds different kinds of information, including procedural knowledge ("knowing how" to do something), semantic knowledge (facts, i.e. "knowing that"), and episodic memories (details of events or episodes in one's life). These memories are organized

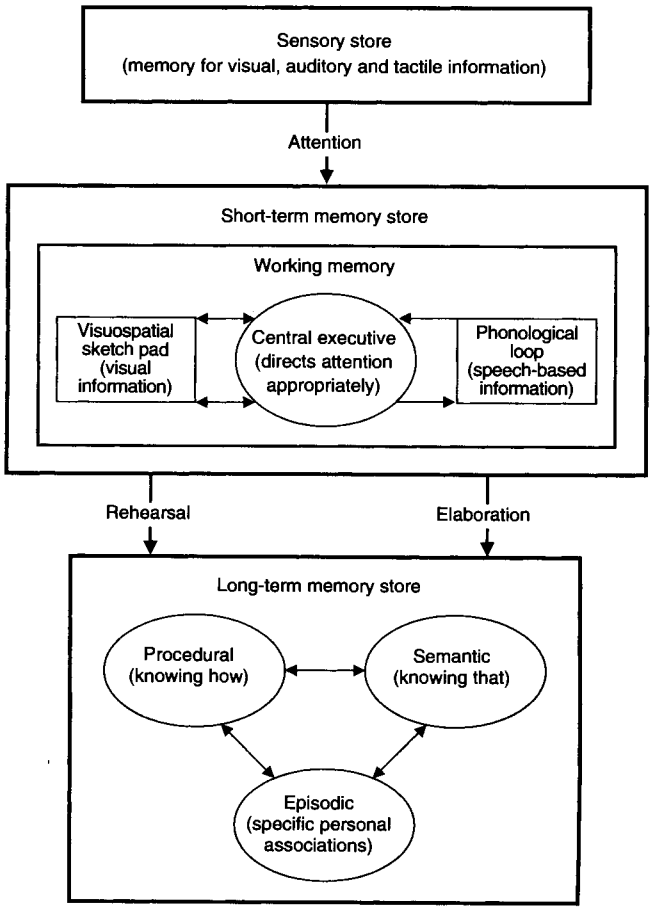



**Figure 7.1** A graphical representation of the components of memory, consisting of (1) the sensory store, which temporarily holds environmental information received via the senses; (2) short-term memory, which contains information that captures attention and enables the manipulation of that information through working memory; and (3) long-term memory, which stores information that has been rehearsed and elaborated in such a way that it becomes integrated into one's knowledge. The last of these is particularly vital to the memorization of music and houses procedural and semantic knowledge ("knowing how" and "knowing that," respectively), as well as episodic memories about events in one's life (diagram adapted from Baddeley, 1990; Eysenck & Keane, 2000).

hierarchically, enabling us to make associations both within and between different categories of information. Procedural knowledge underlies the performer's ability to make music, in terms of coordinating his or her movements to produce sequences of notes, such as scales and arpeggios, on an instrument. However, the knowledge that a particular sequence of notes represents certain patterns (a scale of D major, say, or a favorite melody) depends on semantic memory. Meanwhile, the specific associations

the melody has for the individual (e.g. where it was first heard, events going on at the same time as first hearing, others who were present) are held in episodic memory.

Of course, one of the primary objectives of storing information in long-term memory is to be able to retrieve that information accurately and on demand. Explicit retrieval of stored information takes place through the processes of recognition and recall. Recognition involves matching new, incoming information with information already stored in long-term memory. Recall involves actively (re-) producing that information. Both processes can occur spontaneously (the same cue can trigger recognition or recall), but recall makes greater demands on our ability to remember. The focus of this chapter is on the rehearsal and elaboration strategies that contribute to deliberate memorization in music: recall from long-term memory.

### 7.1.2 General recommendations

While all people possess the same memory structures and use the same basic cognitive mechanisms, some people seem to have "better" memories than others. To a limited extent, this may be due to differences in "natural" general memory ability. However, Wilding and Valentine (1997) conclude from their research with exceptional memorizers that most superior memory performance can be attributed to the use of highly effective strategies for learning (or "encoding") and, independently, for retaining information in memory.

There are many accounts of professional musicians exhibiting extraordinary feats of memory (Revesz, 1925/1999; Marcus, 1979; Marek, 1982; Charness *et al.*, 1988). The notion of individual differences in the ability to memorize music, even among those who achieve such feats, is supported by research with musicians memorizing music in the laboratory (Lehmann & Ericsson, 1995; Ginsborg, 2000) and under more naturalistic circumstances in the studio (Ginsborg, 2002). Meanwhile, everyone who plays or sings, no matter what their level of natural ability, can benefit from learning and practicing effective strategies for memorizing music.

Before considering the strategies that have been found to be particularly useful for memorizing music, there are general recommendations for encoding information and storing it in long-term memory for easy retrieval that can aid the musician. These can be categorized in four ways: (1) improving memory in its most general sense; (2) staving off or remediating age-related memory deterioration; (3) enhancing study skills; and (4) using mnemonics by associating meaningless information with the material to be remembered. Firstly, recommendations for improving memory in its most general sense include understanding how memory works and having the motivation to try to remember (Arden, 2002). Secondly, those who fear age-related memory deterioration are advised to reduce stress, eat a balanced diet, take vitamin supplements, avoid alcohol and drugs, keep an active mind, and do "brain exercises" (Small, 2003; see also Chapters 3, 9, 10, 12, and 14 of this volume). Thirdly, students can improve their learning, and specifically their performance in examinations, by using memory strategies such as attending to the material to be remembered (e.g. in lectures) before organizing, interpreting, and understanding it, and also by visualizing, reciting, and reviewing it (Pauk, 1994). Different techniques focus on enhancing different aspects of memory. For example, "mind-mapping" (Buzan, 1989) involves organizing

and making appropriate links between chunks of information. The end-product (the "mind-map") provides an elaborated set of notes; more importantly, the process of devising an idiosyncratic visual representation serves to reinforce memory. Finally, mnemonics include "peg-words," such as rhyming words and numbers (one = bun, two = shoe); the method of *loci*, dating from classical times, whereby objects or concepts are associated with familiar locations such as landmarks on a regular journey, or different rooms in one's home; and linking unrelated items by weaving them into a story (Arden, 2002).

Many of these recommendations, such as those relating to health and well-being, are supported by continuing research into the functioning of the brain (Small, 2003). Understanding how memory works may seem an ambitious aspiration, yet it is well known that metacognition, or "knowing about knowing," is essential to learning (see Chapter 5 for specific recommendations for individual practice). When students are advised to attend to and later review what has been said in lectures, this is based on the notion of transferring information from the sensory store into short-term memory and thence into long-term memory. Interpreting new information in order to understand it, and linking it with information already stored in long-term memory is a highly effective way of extending the storage capacity of short-term memory by increasing the size of information chunks. This was shown dramatically in a study carried out by Ericsson *et al.* (1980), in which one of the authors, a keen amateur runner, learned to encode up to 82 digits by thinking of them as combinations of running times (e.g. 3594 was stored as 3 minutes 59.4 seconds for a mile run, which was meaningfully classified as "fast"). Visualization and recitation are forms of elaboration and rehearsal, drawing on visual, kinesthetic, and auditory processes; reviewing involves reorganization of the material to be remembered in long-term memory. Similarly, mind-mapping combines the creative use of visual and associative memory. As for the mnemonics outlined above, they all rely on forming associations between items, using visual and/or auditory memory, in order to create new meanings and links between the new information to be remembered and existing information stored in long-term memory. Still, how do these recommendations relate to memorizing music?

- You are well advised to take care of your general health for sharp and long-lasting memory performance.
- One of the main aims of this chapter is to help you find out what is currently known about how memory works and how musicians go about memorizing music so that you can learn to use effective memorizing strategies yourself. These are based on similar principles to those that underpin most study skills.
- However, while study skills usually relate to the learning and memorization of semantic knowledge, you need procedural knowledge as well to sing or play. This is gained through practice involving some inevitable repetition, in order to "automatize" performance.
- As for mnemonics, they may not at first glance seem relevant to music memorization. Yet, they can be useful for encoding sequences of musical events within a piece, or the words of songs that do not seem to make sense otherwise.

## 7.2 Musical memory

One way of describing a memory for a specific piece of music—whether acquired through repeated hearings or deliberate memorization—is a "mental representation," stored in long-term memory, on which the musician can draw when performing. Lehmann (1997) suggests a structure for expert musicians' mental representations with three facets: (1) the desired performance goal, (2) the knowledge of how to produce that goal, and (3) the actual performance. The desired performance goal represents an ideal performance: a rehearsed, highly detailed interpretation of the music. This interpretation can be conceptualized at different levels, forming a hierarchy (Williamon *et al.*, 2002). At the most global level, the music is represented mentally as a whole piece, from beginning to end; at the most local level, it is represented by the notes forming each musical sequence. Of course, there are numerous possible hierarchical levels in between. In a classical sonata movement, these might be the exposition, development, and recapitulation; in a song they might be introduction, verse, chorus, and so on. The desired performance goal is achieved (to a greater or lesser extent) via the second and third facets of the mental representation suggested by Lehmann—that is, the production aspects and the actual performance. The representation of production aspects is constructed from repeated rehearsals of musical sequences that become fully automated but which must continually be monitored. The representation of the actual performance—the interpretation of the music—also involves monitoring in the context of the performance situation and feedback from the audience. In terms of the different kinds of long-term memory outlined in Section 7.1.1, musicians' semantic and episodic memories for the music they perform inform the desired performance goal. Meanwhile, actually playing, singing, or conducting music requires procedural memory.

Different kinds of memorizing strategies can contribute to the formation of mental representations at their different levels and enable attention to be shifted, during practice, from one level to another (Williamon *et al.*, 2002). Whether the information coming into the sensory store is tactile, auditory, and/or visual (e.g. the feel of the fingers on the keys of the instrument, the sound of the music, and the look of the musical score, respectively), sensory information is essential for mental representations of music at the local level. It is also essential for producing and monitoring the desired performance goal. Sequences of musical events are usually memorized—and automatized—via rote repetition using "kinesthetic" memory. However, they could not be learned in the first place without the use of "visual" memory of the notated music and of the way the music is created, and "aural" memory of sounds that are imagined, heard, or produced by the performer for the first time. Depending on the learning preferences of the musicians, music can thus be rehearsed and elaborated kinesthetically, visually, and aurally for representation in long-term memory.

However, in order to represent music mentally at global and local levels and to move between the different levels of representation during practice, analytic or conceptual strategies are required. These strategies have already been introduced in relation to study skills; they enable us to understand and organize the material (in this case, the music) to be remembered. Unlike kinesthetic, visual, and aural representations,

conceptual memory for music can be thought of as a form of semantic knowledge (as well as procedural), which facilitates performance.

So how do musicians go about memorizing using these different strategies? Some memorize, apparently, almost without conscious awareness. When asked how they memorize they cannot—or will not—say. Many musicians, on the other hand, have shared their experiences of memorizing, in published books and interviews, passing on the techniques they have found useful to their students. The list that follows is based on these, as well as on a growing body of research in psychology of music. Of course, no single strategy is guaranteed for all musicians and individual differences in learning styles are discussed in Section 7.3.

### 7.2.1 Memorizing by rote

Rote memorization is perhaps the most common form of rehearsal, particularly for novice musicians attempting to commit a piece of music to memory. Essentially, it uses kinesthetic memory (otherwise known as tactile, motor, finger or muscular memory) and consists of repetition of a bar, phrase, or page until it can be played, automatically, by "feel." For example, many pianists report memorizing with separate hands and consistently using the same fingerings so as to develop their kinesthetic memory for a piece of music (Chaffin *et al.*, 2002). In a study of singers memorizing words and music separately and together, however, the latter performed with greater accuracy and confidence, since the words and music acted as cues for each other (Ginsborg, 2000).

- If you are a keyboard player, you could memorize by rote with your hands separately as well as together; for singers, try to use it to link the words and music together.

Expert musicians use rote memorization to "overlearn" their performances, to ensure that they will be as secure as possible, but they almost always use other memorizing strategies in addition to rote memorization. This is because kinesthetic memory is vulnerable to interference. A stray thought, a cough from the audience, hearing something in the music, as it were, for the first time can "throw" the performer.

- You should use rote memorization in order to free your mind to monitor your performance, to focus on the meaning you wish to convey, and to communicate with your fellow musicians and/or your audience.

In an interview study carried out by Hallam (1997) with novice musicians and professional orchestral musicians, the majority reported memorizing—initially—without conscious awareness, as they learned their parts. When they identified passages they could not play accurately from memory, they chunked them into smaller sections and repeated each section until it was secure. Then, they grouped the small sections into longer sections, and repeated these longer sections until they too were secure. In this way, the whole piece would eventually be successfully memorized. This technique was useful for short pieces; however, for longer, more complex works, half of the professional musicians in Hallam's sample reported using what she refers to as "cognitive analysis" (see Section 7.2.5) to construct a framework into which the memorized sections could be placed.

### 7.2.2 Memorizing visual information

Most classical musicians begin learning and practicing new pieces from a notated source, and this can form an important foundation for memory. Indeed, many musicians report "knowing where they are on the page" when they play or sing notated music from memory—a very useful tool for preventing or recovering from memory lapses. As early as 1800, long before it was customary for classical musicians to perform from memory, Anfossi (cited in Monahan, 1978) recommended that singers should memorize at least the turn of the page.

- It is not a good idea to rely simply on the visual memory you acquire in the course of learning the music. This is similar to memorizing without conscious awareness; like rote memorization, it can be an unreliable technique on its own. Rather, use memory for visual information deliberately; you could mark the score with a pencil or colored pen.

Although there is anecdotal evidence of musicians' (particularly conductors') memory for musical scores (see Marcus, 1979), empirical research on memory purely for visual information is scant. Nuki (1984) asked piano and composition students first to sight-read and then to memorize a novel piece of music. When they had memorized it, they performed it from memory. They were then asked which of four strategies they had used to memorize: kinesthetic, visual, aural, or a mixture of the three. The students who reported using a visual strategy were significantly quicker to memorize than those who used one of the other strategies or a mixture of the three. However, by "visual" strategy, Nuki infers the use of the "inner ear" (see Section 7.2.3), and indeed, the quickest memorizers in her sample were the students of piano and composition who were also good sight-readers and had gained high grades in solfège at their universities.

- Deliberate memorization of visual information, whether signals from another musician or the score itself, is linked to conceptual memory (see Section 7.2.5). You should carry out some form of analysis of the music in order to decide which information will serve as the most reliable cues and then mark those cues in the score. Each time you rehearse the piece thereafter, you are reinforcing (sometimes unknowingly) your memory of those cues.

Visual memories can also consist of visions of the position of the instrument and of the body. Where musicians play in groups they often rely on each other's visual, as well as aural, cues.

- If you are playing or singing with an ensemble, visual memory may be triggered by something as basic as a conductor's down-beat or a particular movement by a section leader. These are not just trivial remembrances, but rather they can play a major role in the unfolding of a performance. For maximal efficacy in practice, particularly useful cues as such should be explicitly acknowledged to ensure their consistent delivery.

### 7.2.3 Memorizing by ear

Many musicians—particularly musicians who are not part of the Western classical tradition—do not rely on musical notation to learn or memorize, but rather memorize

through listening and imitating what they hear. For some, this is laborious and time-consuming; for others, comparatively easy. Macek (1987) reports the case of a jazz musician who developed what he called a "photographic" ear (analogous to eidetic memory for visually-presented stimuli; Haber & Haber, 1988), which enabled him to recall and reproduce complex music in detail. Aural memory is as hard as visual memory to disentangle from the formation of mental representations and conceptual memory. Two studies in particular illustrate this. First, Sloboda *et al.* (1985) studied the exceptional verbatim memory for music of an autistic savant and showed that his memory was reliable only when he listened to music in a familiar genre. Second, Sloboda and Parker (1985) asked musicians and nonmusicians to listen to short folk melodies and then sing them as accurately as possible from memory. On the whole, the versions produced were inaccurate and/or incomplete, but they preserved the metrical and—particularly in the case of the musicians—harmonic structure of the originals. This suggests that we construct memory for specific pieces of music using what we already know about and have experienced with music; every new piece of music that we hear contributes to this knowledge.

- While some musicians rely on the aural memories they acquire in the course of learning music, this is rarely reliable without conceptual memory to serve as a framework for monitoring "where you are" in performance and "where you are going," and to provide you with a repertory of alternative musical sequences with which to improvise, should you have a memory lapse.
- Once you have a conceptual framework, using aural memory as a basis for rote repetition, either playing or singing aloud, or in your imagination, is an excellent strategy.

### 7.2.4 Translating visual information into imaginary sound

An important tool for musicians who do use musical notation is the ability to imagine the way a performance would sound, or how it would feel to play or sing, from the score. Gieseking and Leimer (1932/1972) call this "visualization" and recommend that pianists should memorize, referring to the printed score, using their "inner ear" *before* beginning to practice at the keyboard. Gordon (1993) refers to this as "audiation": being able to hear and understand music that is no longer, or may indeed never have been, audible. Equally, it is possible for musicians who do not use musical notation to hear a piece of music and imagine the way they, or someone else, would perform it. The research by Nuki outlined in Section 7.2.2 highlights the role of the ability to read music, translating visual information into imaginary sound and thus creating a mental representation for music.

- The difference between reading music and playing or singing aloud, and reading music and conjuring up imaginary sound (and "feel") is like the difference between reading a piece of text aloud, as beginners do, and reading silently. If you are a musician who must learn and memorize notated music, then learning to hear it in your mind as you read it—so that you can memorize away from your instrument—is one of the most valuable skills you can develop.

### 7.2.5 Developing and exploiting conceptual memory

The ability to form mental representations of music at the global and local levels, particularly for extended compositions, depends on "conceptual" memory: the musician's existing semantic knowledge—held in long-term memory—of the structures that underlie the music. These structures may vary according to musical genre, and they include chord progressions in jazz, rhythmic patterns, *ragas* in Indian music, and the conventions of tonality in Western classical music.

Lehmann and Ericsson (1995) tested the hypothesis that the ability to construct and draw on a mental representation facilitates performance from memory that is both stable (i.e. the musician can reliably "keep the music the same") and flexible (i.e. variations can be made in performance, depending on the requirements of the situation). They asked advanced students of piano to memorize two short pieces of music, providing "think-aloud" verbal reports as they did so. The pianists memorized the pieces—accompaniments to an instrumental "pacing" track—in the laboratory by carrying out a number of pairs of trials, playing each piece once with the musical score and once without until they could perform it from memory without errors twice, consecutively: the fewer trials they needed, the faster they had succeeded in memorizing. Then they carried out a series of tasks designed to test their ability to "manipulate" the pieces from memory. The results of the study showed that the pianists who were quickest to memorize made the fewest errors when transposing from memory. Thus, rather than learning musical sequences by rote and reproducing them using kinesthetic memory, it is clear that efficient memorization must involve the ability to construct and draw on a detailed mental representation of the music, allowing the performer both to reproduce it reliably and to manipulate it in performance from memory.

Ginsborg (2000) carried out a similar study to investigate the nature of mental representations for the words and music of songs with 20 singers. Each singer memorized two songs, in two separate sessions, paced by a recorded accompaniment, using the Lehmann and Ericsson (1995) memorizing paradigm. Once the songs were memorized, the singers carried out a number of tasks. Transposition—the task the quickest memorizers in Lehmann and Ericsson's study performed most accurately—is relatively simple for most singers, so instead they were asked to retrieve the words and melody of the song independently, to "deconstruct" the words and melody, to show an awareness of the compositional structure of the song, to use visual memory to retrieve information about different aspects of the musical score, and to respond to cues taken from the words and music of the song.

There was no correlation between speed of memorization and performance on some of the more straightforward tasks, such as performing the song accurately with accompaniment and at the same tempo as it was originally memorized. This would suggest a role for the automatization of performance, through the rote repetition of text and melody together (see also Section 7.2.6). On the other hand, speed of memorization was significantly correlated with performance on the seven most difficult tasks: speaking the words of the whole song at speed; retrieving fragments of the text and fragments of the music, with the words and pitches in reverse order; responding to cues, both forward and reverse; singing the phrases of the song in reverse order; and singing the pitches of the melody of the song only, without rhythm, to the regular beat

of a metronome. Such tasks could not have been performed successfully in the absence of a well-formed, multicode mental representation of the song including conceptual, kinesthetic, visual, and aural information. As in Lehmann and Ericsson's study, the faster the singers acquired these mental representations, the better they performed the tasks. Mental representations of music must, therefore, underlie the ability to produce performances from memory that are both stable and flexible.

- Once you have memorized a piece for the first time, you should test your memory in order to simulate different possibilities for recovering from lapses. You could try starting mid-section instead of at a structural boundary, for example, or by practicing jumping from one section to a previous section and to later sections. When you have done this, you will know you have created a mental representation on which you can draw to produce a stable, yet flexible, performance.

Research has focused on two main ways of developing and exploiting mental representations using conceptual memory: analysis and the use of structural boundaries to organize practice. These are discussed and recommendations are offered in turn.

*Analysis*

It is often suggested that musicians should analyze the music they are to memorize (e.g. Hughes, 1915; Matthay, 1926), partly so as to be able to "chunk" or to divide the music into shorter sections; once each one is memorized they can be "grouped" together. Many researchers have investigated the efficacy of analysis as a memorization strategy. For example, half the professional musicians—though none of the novice musicians—who took part in Hallam's (1997) study reported using an analytic approach to develop a framework into which they could fit the material they had memorized by rote. According to Hallam, this included noting where the music changed from one key to another, particular harmonic progressions, the length of rests, and difficult exit points.

Rubin-Rabson (1937), on the other hand, compared four groups of piano students' memorization of short pieces. One group memorized at the keyboard. Two groups analyzed the pieces before memorizing them, one independently, the other with Rubin-Rabson's guidance. The fourth group listened to a recording of the pieces and then memorized in one of the three ways already detailed. Three weeks later, the students who had either analyzed the pieces themselves or with guidance were the quickest to relearn the pieces and to perform them accurately from memory.

Ross (1964) replicated this study with the participation of competent wind players randomly divided into experimental and control groups. Following a pre-test to establish baseline memorizing ability, the experimental group undertook a weekly hour-long session with a teacher, for 6 weeks, in which they analyzed and memorized a series of training pieces, in preparation for memorizing a test piece at the end of the 6-week period. Again, the experimental group accurately memorized the test piece significantly faster than the control group.

The most detailed account of the use of analysis to develop conceptual memory for music is provided by Chaffin and Imreh who have been collaborating for almost a decade on a project exploring the cognitive processes involved in learning and memorizing

music (see Chaffin *et al.*, 2002). The purpose of their first study was to produce a systematic record of the activities of an experienced professional concert pianist (i.e. Imreh) while memorizing the third movement of Bach's Italian Concerto (Chaffin & Imreh, 1994). Imreh undertook 57 practice sessions over the course of 10 months, 34 of which were recorded on videotape and 11 on audiotape. She talked aloud about her decision-making as she worked; she was interviewed and, once the piece was successfully memorized, she commented on her videotaped practice and annotated the score to show how she organized the music conceptually. One of the findings of this study was that she identified the compositional structure of the piece and used it not only to guide her practice, through chunking and grouping, but also to ensure accuracy when she played from memory. She practiced and memorized acknowledged sections in the music separately and together. When she came to play the whole piece from memory, she would deliberately recall the particular features of each section in order to remind herself which section she was actually playing and which came next (Chaffin & Imreh, 1997).

- The use of conceptual memory is the crucial overarching strategy that no musician can do without.
- Whatever kind of music you play or sing, you need to identify how it fits together (i.e. its structure), whether this is at a highly detailed level such as that described above or at a much simpler level—remembering, for example, that you must sing two verses interspersed with three choruses, the last time in a higher key.
- How you go about identifying structure and using it to help you memorize will depend on the nature of your musical training as well as the nature of the music. Some musicians rely on an understanding of underlying harmony; others rely on memory for melodic sequences; others rely on identifying and being able to reproduce rhythmic patterns. In Western classical music, all of these are useful ways of conceptualizing music, separately and together.

*Structural boundaries*

Williamon and Valentine (2002) explored the development of the ability to identify structure, to use it as a tool for organizing practice, and to employ it successfully for retrieval in performance. They asked 22 pianists to learn and memorize a piece by J. S. Bach appropriate to their level of expertise. All of the pianists' practice sessions were recorded on audiotape. The final performances that they gave from memory, as part of a recital, were recorded on videotape and assessed by experienced adjudicators. The pianists made verbal and written comments on their practice, annotated their musical scores and were interviewed following their final performances about the practice and memorizing process. The results of the study showed that pianists with higher levels of skill were more likely than those at lower levels to practice segments of the piece beginning and ending on bars at the structural boundaries (these were defined according to the structure of the piece as understood and described by the individual pianist; only three pianists' descriptions conformed to the "textbook" formal structure of the piece). The likelihood that a pianist would use structural boundaries to guide his or her choice of the part of the piece to be practiced increased over the course of practice sessions.

The earlier the pianist began to do so, however, the better his or her final performance was likely to be judged.

Research has shown that singers, too, use structural boundaries to guide their practice during the course of memorization. Ginsborg (2002) carried out an observational study in which 13 singers learned and memorized the same song over the course of up to six 15-minute practice sessions. The sessions, which included concurrent verbal commentaries made by the singers, were taped and transcribed. Finally, an assessment of the singers' speed and accuracy of memorization was made in order to determine which of the various memorizing strategies used by the singers were more and less effective. Fast, accurate memorizers were no more likely than slower, less accurate memorizers to use structural boundaries to guide their practice. Furthermore, the use of structural boundaries did not change significantly between the beginning and end of the practice sessions. However, it was used widely: more than four fifths of "attempts", the segments of the song that were practiced, began on one of its eight phrases. Nearly half of the attempts began at the start of a phrase and finished at the end of the same or a subsequent phrase. The majority of attempts were on the first phrase, followed closely by the third phrase (i.e. the beginning and middle of the first verse); a much smaller proportion of attempts was made on the fifth and seventh phrases (the beginning and middle of the second verse). An unexpectedly high proportion of attempts on the final phrase was attributable to the fact that participants found it "technically" difficult to sing (it had two large-interval "leaps") and also to memorize.

- You should try to divide the music into sections in a way that makes sense to you, so as to memorize the sections one at a time before re-combining them into longer sections and fitting them into your conceptual framework for the whole piece.
- The more you organize your practice around sections identified in the music (especially in the early stages of learning a piece), the more likely you are to remember this basic architecture in performance.

### 7.2.6 Memorizing strategies for singers: Combining words and music

Like instrumental musicians, singers use kinesthetic, visual, aural, and conceptual memory. Unlike instrumental musicians, however, they must also memorize words. It should be remembered that music and words are not necessarily equal partners: sometimes the voice is used instrumentally (e.g. scat singing in jazz or choral music by Bach) and sometimes the words are more important than the melody (e.g. operatic recitative, *Sprechstimme*). The relative importance of words and music in a particular musical genre will determine the memorizing strategies used, and whether the words and music are memorized separately or together. In an observational study, Ginsborg (2002) identified a variety of "modes" of attempt used in singers' practice sessions. These were: focusing on the words alone (speaking the words, while reading from the score and from memory); focusing on the music alone (playing the melody, playing the accompaniment, vocalizing the melody, and counting—the last of these while reading from the score and from memory); and singing the words and music together, while reading from the score and from memory. Fast, accurate memorizers used more modes of attempt than the slower, less accurate memorizers. They were also more likely to

count aloud, but most importantly, they began memorizing earlier and sang more of the song from memory during the course of their practice sessions.

Next, the higher-level strategies of the "best" memorizer (the first of the 13 participants in the sample to sing the whole song entirely accurately from memory) and the "worst" memorizer (the one who took longest to memorize and made the most errors when she sang from memory) were compared, with reference to their verbal commentaries as well as their practice data. The "best" memorizer sang the words and music of the song together rather than separately. She started memorizing early and tested her memory throughout the practice sessions. She worked on a variety of lengths of practice unit. She made plans and implemented them, monitored and corrected her errors, and explicitly evaluated her practice. Her strategies were varied and complex; furthermore, her verbal commentary was detailed and self-guiding. In contrast, the "worst" memorizer implemented plans, monitored errors, and evaluated her practice to a much lesser extent. She preferred to sing the music only, started to memorize comparatively late and consistently went back to the beginning of the song and repeated it in its entirety, rather than dividing it into smaller sections.

Ginsborg (2000) also investigated which of two strategies is more effective for memorizing a song: memorizing the words and melody separately or memorizing them together. Sixty singers took part in this study. Each singer was categorized in two ways: by their experience of memorizing and by their level of musical expertise. Half of them were solo singers with considerable experience of memorizing and performing from memory, while the other half were choral singers with little or no experience of memorizing songs. Half of them had undertaken musical training to advanced levels, while the other half had lower levels of musical training.

They were randomly divided into three groups and asked to memorize a short, unaccompanied song in a single 20-minute practice session. The first group memorized the words first, then the melody, and finally both together; the second group memorized the melody first, then the words, and finally both together; the third group memorized the words and melody simultaneously throughout the course of the practice session. When the singers had memorized the song, they were asked to sing it from memory twice, their two performances separated by a 10-minute interview. Their second performances from memory were transcribed, coded, and analyzed. Errors and hesitations were noted and counted. Generally, the singers who memorized the words and melody of the song together gave more accurate and more confident performances from memory. They made fewer word-only errors, and fewer simultaneous word-and-music errors. They also made fewer hesitations. So, as shown in the comparison of the strategies used by the "best" and "worst" memorizer in the observational study, memorizing words and melody together appears to be a more effective strategy than memorizing them separately.

A comparison of the performances of singers with high and low levels of memorizing experience, and with high and low levels of musical expertise, was also instructive. Those singers with considerable experience of memorizing songs were no more accurate or confident in their performances from memory than those with no experience of memorizing songs. In other words, experience is no guarantee of success, at least in

terms of memorizing. On the other hand, there were observable differences between those singers who had advanced levels of musical training and those who did not. The singers with high levels of musical expertise made fewer music-only errors, and fewer simultaneous word-and-music errors, than the nonexpert singers. Furthermore, while—as one might predict—they were generally more accurate in their performances from memory, this was *only* the case when they had memorized the words and melody together.

- Almost every memorizing strategy recommended in this chapter is as applicable to singers as it is to instrumental musicians. The only difference between them is that singers must memorize words as well as music. Think of the words as an element of the music rather than the other way round.
- You might think of the words in terms of the way they look on the page (visual memory), the way they sound (aural memory), or the way it feels to produce them (kinesthetic memory). If you are learning a melody using visualization/audiation, mouth the words to yourself. Once they are securely memorized, it is helpful, of course, to think about what they mean.

## 7.3 Conclusions

The aim of this chapter has been to report on memorizing strategies that have been shown to be effective for all musicians, both instrumentalists and singers. To reiterate: music is stored using multiple representations in long-term memory. Effective memorization strategies capitalize on this.

- Analyze the music you are to memorize in order to understand it and organize it, using conceptual memory as an overarching framework.
- Chunk the music and practice it in small sections, from one structural boundary to the next, increasing the size of chunk as you become more familiar with the music.
- Use rote, kinesthetic memory in conjunction with either visual or aural memory, or both.

For the most part, discussion has focused on memorization of a specific piece of music; however, some final, more general suggestions can be made:

- First, it has been shown in a number of domains that "little and often" is a better strategy for memorizing than trying to memorize in just one or two extended sessions. Sometimes you find that your memory for a piece of music has improved between practice sessions without conscious effort; alternatively, you can use the time between practice sessions for recalling the music you have memorized and identifying sections that may need more work.
- Second, there are strategies that can be used in the final stages of preparing to give a performance from memory. One is *state-dependence*: learning to control your frame of mind so that you can conjure up the same sense of exhilaration, for instance, when you are practicing as when you are actually performing. The other is *context-dependent learning*: practicing in the room or on the stage where the performance is

to take place. These strategies, like so many others, can be useful whatever kind of music you are playing or singing, and whether you are a soloist or in an ensemble.

In addition, we have to remind ourselves that different tasks may require different strategies. Furthermore, we also need to remember that different individuals have different capacities for memorizing easily and successfully. While most research investigating memorization strategies has focused on the average numbers of trials required to memorize and re-learn pieces of music, these mask a wide variety of memorization speeds. For example, Lehmann and Ericsson (1995) found that their slowest memorizer needed 6–7 times as many trials to memorize as their fastest memorizer; using the same paradigm, the singers in Ginsborg's (2000) study required from 2 to 17 trials to memorize a song with a word-text, and from 3 to as many as 49 trials to memorize the song with a digit-text. Similarly, in Ginsborg (2002), in which 13 singers learned and memorized under more (although not completely) naturalistic circumstances, the fastest memorizer produced 3232 beats before her first accurate performance from memory (equivalent to just under 20 renditions of the song) while the slowest produced 5287 beats (equivalent to 32 renditions).

These findings may well reflect the extent to which the musicians who took part in these studies found that the strategies they were encouraged to use resembled or interfered with their usual, or preferred, strategies. From interview and experimental data, it is clear that musicians not only have different memorizing abilities, but also a wide variety of views on memory and preferences for using different approaches to memorization (e.g. kinesthetic, visual, and aural). These preferences have been characterized by some educationists as learning modalities or learning styles, which can be identified through the use of checklists such as the Swassing–Barbe Modality Index (SBMI; Barbe & Swassing, 1979) and the Learning Styles Index (LSI; Price & Dunn, 1997; Dunn *et al.*, 2003). Although learning style theories have been criticized on the grounds that they lack a psychological foundation (Sternberg & Grigorenko, 2001; Renzulli & Dai, 2001), some music teachers advocate teaching their students according to their particular learning strengths (see Garcia, 2002).

While the impact of learning styles on expert musicians' memorizing abilities and strategies has not yet been explored, research has been carried out using the SBMI and LSI with children: (1) to identify "gifted" musicians and investigate the extent to which their learning styles differ from those of the general population (Kreitner, 1981); (2) to find out if memory for rhythm patterns can be enhanced by presenting them in children's preferred modalities (Persellin & Pierce, 1988); (3) to examine the relationships between the perceptual elements of learning style, musical aptitude, and attitudes to music learning (Faulkner, 1994); and (4) to examine the influence of children's preferred modalities on the way they listen to music. The findings of these studies indicate that children's preferred modalities may influence some aspects of basic music listening and learning. However, there is only anecdotal evidence, as yet, to suggest that these persist into adulthood. If individual musicians do possess particular learning strengths then they must also have complementary learning weaknesses. Future research might well seek to determine the nature of the training whereby such weaknesses in memorization can be systematically improved.

## Further information and reading

Chaffin, R., Imreh, G., & Crawford, M. (2002). *Practicing Perfection: Memory and Piano Performance.* Mahwah, NJ: Erlbaum.

Gieseking, W., & Leimer, K. (1932/1972). *Piano Technique.* New York: Dover. (Original work published in 1932.)

Lehmann, A. C. (1997). Acquired mental representations in music performance: Anecdotal and preliminary empirical evidence. In H. Jørgensen & A. C. Lehmann (Eds.), *Does Practice Make Perfect? Current Theory and Research on Instrumental Music Practice* (pp. 141–163). Oslo: Norwegian Academy of Music.

Wilding, J., & Valentine, E. (1997). *Superior Memory.* Hove, UK: Psychology Press.

Williamon, A. (2002). Memorising music. In J. Rink (Ed.), *Musical Performance: A Guide to Understanding* (pp. 113–126). Cambridge: Cambridge University Press.

## References

Arden, J. B. (2002). *Improving Your Memory for Dummies.* New York: Wiley.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 2, pp. 89–195). London: Academic Press.

Baddeley, A. D. (1990). *Human Memory: Theory and Practice.* Boston: Allyn and Bacon.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 8, pp. 47–90). London: Academic Press.

Barbe, W. B., & Swassing, R. H. (1979). *Teaching through Modality Strengths: Concepts and Practices.* Columbus, OH: Zaner-Bloser.

Buzan, T. (1989). *Use Your Head.* London: BBC Consumer Publishing.

Chaffin, R., & Imreh, G. (1994). *Memorizing for Performance: A Case Study of Expert Memory.* Paper presented at the Third Practical Aspects of Memory Conference, University of Maryland.

Chaffin, R., & Imreh, G. (1997). "Pulling teeth and torture": musical memory and problem solving. *Thinking and Reasoning, 3,* 315–336.

Chaffin, R., Imreh, G., & Crawford, M. (2002). *Practicing Perfection: Memory and Piano Performance.* Mahwah, NJ: Erlbaum.

Charness, N., Clifton, J., & MacDonald, L. (1988). Case study of a musical "mono-savant": A cognitive-psychological focus. In L. K. Obler & D. Fein (Eds.), *The Exceptional Brain: Neuropsychology of Talent and Special Abilities* (pp. 277–293). London: Guilford Press.

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music, 21,* 103–113.

Davidson, J. W. (1994). Which areas of a pianist's body convey information about expressive intention to an audience? *Journal of Human Movement Studies, 26,* 279–301.

Dunn, R., Dunn, K., & Price, G. E. (2003). *Learning Style Inventory (LSI) Grades 5–12.* Lawrence, KS: Price Systems.

Ericsson, K. A., Chase, W. G., & Falloon, S. (1980). Acquisition of memory skill. *Science, 208,* 1181–1182.

Eysenck, M. W., & Keane, M. T. (2000). *Cognitive Psychology: A Student's Handbook* (4th ed.). Hove, UK: Psychology Press.

Faulkner, D. L. (1994). *An Investigation of Modality Preferences, Musical Aptitude, and Attitude Toward Music at the Third-Grade Level.* Unpublished doctoral dissertation, University of Mississippi, Oxford, MS.

Garcia, S. (2002). Learning styles and piano teaching. *Piano Pedagogy Forum, 5(1).* Retrieved August 28, 2003, from *www.music.sc.edu/ea/keyboard/PPF/5.1/5.1.PPFpp.html.*

Gieseking, W., & Leimer, K. (1932/1972). *Piano Technique.* New York: Dover. (Original work published in 1932.)

Ginsborg, J. (2000). Off by heart: Expert singers' memorisation strategies and recall for the words and music of songs. In C. Woods, G. Luck, R. Brochard, F. Seddon, & J. A. Sloboda (Eds.), *Proceedings of the Sixth International Conference on Music Perception and Cognition.* Keele, UK: Keele University.

Ginsborg, J. (2002). Classical singers learning and memorising a new song: An observational study. *Psychology of Music, 30,* 58–101.

Gordon, E. E. (1993). *Learning Sequences in Music: Skills, Contents, and Patterns: A Music Learning Theory.* Chicago: GIA Publications.

Haber, R. N., & Haber, L. R. (1988). The chacteristics of eidetic imagery. In D. Fein & L. K. Obler (Eds.), *The Exceptional Brain: Neuropsychology of Talent and Special Abilities* (pp. 218–241). London: Guilford Press.

Hallam, S. (1997). The development of memorisation strategies in musicians. *British Journal of Music Education, 14,* 87–97.

Hughes, E. (1915). Musical memory in piano playing and piano study. *Musical Quarterly, 1,* 592–603.

Kreitner, K. R. (1981). *Modality Strengths and Learning Styles of Musically Talented High School Students.* Unpublished masters dissertation, Ohio State University, Columbus, OH.

Lehmann, A. C. (1997). Acquired mental representations in music performance: Anecdotal and preliminary empirical evidence. In H. Jørgensen & A. C. Lehmann (Eds.), *Does Practice Make Perfect? Current Theory and Research on Instrumental Music Practice* (pp. 141–163). Oslo: Norwegian Academy of Music.

Lehmann, A. C., & Ericsson, K. A. (1995). *Expert pianists' mental representation of memorized music.* Poster presented at the 36th Annual Meeting of the Psychonomic Society, Los Angeles.

Macek, K. (1987). The photographic ear. *Piano Quarterly, 35*(137), 46–48.

Marek, G. R. (1982). Toscanini's memory. In U. Neisser (Ed.), *Memory Observed: Remembering in Natural Contexts* (pp. 508–511). San Francisco: Freeman.

Marcus, A. (1979). *Great Pianists Speak.* Neptune City, NJ: Paganiniana Publications.

Matthay, T. (1926). *On Memorizing and Playing from Memory and on the Laws of Practice Generally.* Oxford: Oxford University Press.

Monahan, B. J. (1978). *The Art of Singing: A Compendium of Thoughts on Singing Published between 1877 and 1927.* London: Scarecrow Press.

Miller, G. A. (1956). The magic number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–93.

Nuki, M. (1984). Memorization of piano music. *Psychologia, 27,* 157–163.

Pauk, W. (1994). *How to Study in College* (5th ed.). Boston: Houghton-Mifflin.

Persellin, D. C., & Pierce, C. (1988). Association of preference for modality to learning of rhythm patterns in music. *Perceptual and Motor Skills, 67*, 825–826.

Plunkett Greene, H. (1912). *Interpretation in Song.* London: Macmillan.

Price, G. E., & Dunn, R. (1997). *Learning Style Inventory (LSI) Manual.* Lawrence, KS: Price Systems.

Renzulli, J. S., & Dai, D. Y. (2001). Abilities, interests, and styles as aptitudes for learning: A person-situation interaction perspective. In R. J. Sternberg & L.-F. Zhang (Eds.), *Perspectives on Thinking, Learning and Cognitive Styles* (pp. 23–46). Mahwah, NJ: Erlbaum.

Revesz, G. (1925/1999). *The Psychology of a Musical Prodigy.* London: Routledge. (Original work published in 1925).

Ross, E. (1964). Improving facility in music memorization. *Journal of Research in Music Education, 12*, 269–278.

Rubin-Rabson, G. (1937). *The Influence of Analytical Pre-Study in Memorizing Piano Music.* New York: Archives of Psychology.

Sloboda, J. A., & Parker, D. H. (1985). Immediate recall of memories. In P. Howell, I. Cross, & R. West (Eds.), *Musical Structure and Cognition* (pp. 143–167). London: Academic Press.

Sloboda, J. A., Hermelin, B., & O'Connor, N. (1985). An exceptional musical memory. *Music Perception, 3*, 155–170.

Small, G. (2003). *The Memory Bible: Ten Commandments for Keeping your Brain Young.* New York: Hyperion Books.

Sternberg, R. J., & Grigorenko, E. L. (2001). A capsule history of theory and research on styles. In R. J. Sternberg & L.-F. Zhang (Eds.), *Perspectives on Thinking, Learning and Cognitive Styles* (pp. 1–21). Mahwah, NJ: Erlbaum.

Wilding, J., & Valentine, E. (1997). *Superior Memory.* Hove, UK: Psychology Press.

Williamon, A. (1999). The value of performing from memory. *Psychology of Music, 27*, 84–95.

Williamon, A., & Valentine, E. (2002). The role of retrieval structures in memorizing music. *Cognitive Psychology, 44*, 1–32.

Williamon, A., Valentine, E., & Valentine, J. (2002). Shifting the focus of attention between levels of musical structure. *European Journal of Cognitive Psychology, 14*, 493–520.

CHAPTER 8

# STRATEGIES FOR SIGHT-READING AND IMPROVISING MUSIC

SAM THOMPSON AND ANDREAS C. LEHMANN

Why discuss sight-reading and improvising together in one chapter? Seeing the title, many musicians would be forgiven for wondering just what could possibly connect the two. This confusion is partially due to the fact that the two skills are normally situated in different performance practices and stylistic areas. Sight-reading has to do with the performance of notated music, mostly from the Western classical repertory, while improvisation is associated with oral traditions, chiefly jazz. Another, arguably no less important factor is that the "creative" aspect of improvisation is usually given the most emphasis in discussions. Improvisation is seen to be an art that has the potential for individuality and self-expression. Sight-reading, by contrast, is regarded as a largely mechanical task, undoubtedly worthy but essentially unimaginative. However, from a psychological perspective the two skills take place under similar constraints, both involving the performance of musical material without overt preparation. The aim of this chapter is to show how taking this kind of analytical approach can be useful in thinking about ways to enhance ability in both sight-reading and improvising. In making the comparison, it is hoped to shed light on a perplexing question: how can one perform at a level usually encountered in a rehearsed performance, *without* preparing explicitly for the particular performance?

Psychomotor skills (which can be defined, loosely, as deliberate physical movements of the body; see also Chapter 14) are often described as being "open" or "closed." A rehearsed musical performance can be said to be a closed skill in the sense that the same movement has to be performed in an essentially unchanging environment—to draw a sporting analogy, this is something akin to the task facing a competitive swimmer. Improvisation and sight-reading, on the other hand, are open skills in that they require the performer to *adapt* constantly to a changing environment—something more like the task that confronts a soccer player. While swimmers can hone their technique, endurance, and strength, and develop strategies to cope with the psychological demands of the situation, there is no uncertainty about the activity itself—successful performance in a race is a matter of executing a particular, well-practiced motor sequence as fluently as possible. Soccer players never know ahead of the game exactly what combination of motor sequences they will be required to execute, and so must have the ability to adapt their particular technical skills quickly to the situation at hand.

In sight-reading, the musician is confronted with unfamiliar music and has to play through it at first sight, imitating—as convincingly as possible—a rehearsed performance with regard to tempo and expression. Improvisation is even more adventurous in that the performer has to chart unknown territory, sometimes in the company of other musicians who are also improvising. However, unlike sight-reading, which covers a clearly delineated range of activities, improvisation is a broad category ranging from free jazz (with virtually no points of reference) to highly patterned improvisation in classical music, not to mention the many complex types of improvisation outside Western classical music traditions (see Nettl & Russell, 1998).

There are several questions that arise from the difference between an extensively rehearsed and memorized performance, and a sight-read or improvised performance:

- What are the psychological mechanisms that underpin the ability to perform music without rehearsal?
- How are such abilities acquired?
- How can performance in sight-reading and improvising be enhanced?

These questions guide the present survey of the two skills, as they are investigated from two perspectives: (1) the underlying psychological processes and (2) practical strategies for skill acquisition and enhancement. Throughout the chapter, a number of hints are given to the practitioner; some of these are proposals for actual activities, while others offer wider suggestions for increasing knowledge and understanding of the task at hand.

Two general observations should be made at this point. Firstly, research on sight-reading has tended to focus almost exclusively on music of the Western classical tradition. Similarly, the greater part of research in improvisation has focused on American jazz styles. In the case of sight-reading, this is less of a deficit than it might appear since, as discussed below, the classical tradition is more or less unique in identifying and prizing the skill at all (or, indeed, in having a notation system capable of supporting it). With improvisation, the focus on jazz has tended to obscure the many other musical practices that involve a strong improvisational element. Despite these limitations, the intention in this chapter is to make suggestions that are broadly applicable across styles. The second observation is that in both domains the great majority of research has been conducted using keyboard instruments. This is little more than a function of the available technology; clearly it is not just keyboard players who sight-read and improvise. Again, therefore, the suggestions made are intended to have wide, cross-instrument relevance.

## 8.1 Comparing sight-reading and improvisation

For the purposes of this chapter, it is interesting to note that there is a historical link between improvisation and sight-reading. Sight-reading became more common at roughly the same point that improvisation, as a concert spectacle, began to lose popularity. Although sight-reading was always important in ensemble situations, be it in church choirs or various orchestras associated with the nobility, solo performances

were generally given by composers who would perform their own works and then improvise in various ways (there are many reports, for example, of J. S. Bach's renowned ability to improvise a fugue on themes suggested by the audience). Toward the middle of the eighteenth century, however, emerged the virtuoso—a solo performer who was not (primarily, or at all) a composer but was instead heralded for his or her skilled and insightful performances of music by others. In some cases, feats of difficult sight-reading replaced improvisation as the performer's favorite "party trick." These days, of course, classical musicians rarely sight-read in public at all, and certainly do not advertise the fact if they are doing so. Sections 8.1.1 and 8.1.2 consider the two skills separately, in terms of the psychological processes that underpin them.

### 8.1.1 Task analysis: Sight-reading

Any performance of music in which the player reads from a written score could be construed as "sight-reading" in a literal sense. However, in its most common usage the term refers to the practice of playing a piece of music directly from the score on first encounter or after brief rehearsal (see Lehmann & McArthur, 2002, and Lehmann, in press, for reviews of the sight-reading literature). Defined thus, sight-reading is a normal part of the musical experience of classically trained musicians, especially pianists who work as accompanists. Ability at sight-reading involves the capacity to play the music accurately and fluently (i.e. without pauses or breaks in the musical flow) at an acceptable tempo and with adequate musical expression.

Examination systems such as the graded syllabi of the Associated Board of the Royal Schools of Music have traditionally included a statutory test of sight-reading at all levels. Rigorous assessments of sight-reading ability at the keyboard (such as reading an orchestral score at sight or sight-reading vocal scores in multiple clefs) are commonly found in higher education music courses and are usually compulsory for all students, irrespective of whether or not keyboard is their favored instrument. It is not hard to see how the ability to sight-read competently is likely to be useful in a number of common musical environments; professional musicians, for example, particularly those whose work involves recording sessions, frequently find themselves required to sight-read a new score.

Because of the dominance of notation in the Western classical tradition, sight-reading remains an almost exclusively Western classical skill, although music notation is also used in other musical styles and cultures for didactic purposes. Music in the classical tradition is learned and passed-down through notated sources, whereas many other musics are preserved aurally or, increasingly, in the form of sound recordings. While it is the case that, for example, rock music can be described in Western notation (with limitations), rock musicians do not typically learn through notation or perform from it, and so often have little need to sight-read even if the music is available to them in a notated form.

The most immediate psychological observation about sight-reading is that it is an "online" activity, which is to say it requires that a sequence of movements be produced in response to a succession of visual stimuli presented in real-time. The speed of stimulus presentation is a function of the chosen tempo and the relative density of musical events. Although any musician could read anything given enough time and a slow

enough tempo, an effective minimum speed for a particular piece is generally dictated by stylistic convention. Moreover, to maintain fluency it is not possible to pause without interrupting the musical stream—and this is one important way in which musical sight-reading differs from superficially similar skills such as typing from copy or reading aloud. The sequence of events is thus, (1) perceiving notation, (2) processing it, and finally (3) executing the resulting motor program.

*Perceiving notation*

Perceiving notation is an intricate process involving low-level perception routines as well as higher-level cognitive functioning. The lower-level routines are most likely acquired in the same way that we learn to read, namely by slowly deciphering individual symbols at first and gradually increasing fluency (although, it is worth noting that sight-reading music and reading text involve different neurological processes; see Sargent *et al.*, 1992.) It is only after these initial routines have been developed that sight-reading—in the sense considered here—really begins.

Our eyes do not function like a photo camera, but rather more like a flashlight searching around in the dark. Gaze is directed at the page of printed material and only a small circular portion in the middle of our field of vision is in focus at any one time. This area is called the fovea, and around it lays a blurry circle called the parafovea (see Raynor & Pollatsek, 1989, for a description in the context of reading). What we experience as a continuous and coherent picture is actually the result of piecemeal information gathering by our visual system combined with a good deal of cognitive processing to fit it all together. The individual gazes, of which there are roughly five per second, and the connecting intermediate trajectories constitute what is known generically as occulomotor behavior. In sight-reading, this means that the score has to be assembled by jumping to different places on the page trying to find the relevant information.

A number of studies have investigated the movement of the eyes during sight-reading (see Lehmann, in press, for a review). From this research, it has become clearly established that unskilled readers differ markedly from experienced readers with regard to their looking behavior. In particular, better readers tend to look further ahead of the point where they are currently playing and do not fixate on every note as beginners tend to. There are many variables that could account for these differences, some physiological, but since looking behavior is not under conscious control, it is probably safe to assume that the effects of training and expertise play a major role.

- Perception is an interactive process of gathering information from the environment and responding to it. As such, "practicing" looking by, for example, rolling your eyes with the head still will do little to change the behavior of the occulomotor system during sight-reading. You can train the system to be more efficient, but only by performing realistic, interactive tasks.
- How far do you look ahead while sight-reading? A number of studies have shown that accomplished sight-readers consistently read around six or seven notes ahead (in a single-line melody), while novices read only two or three notes ahead. Try closing your eyes at random when reading and carrying on as far as you can—how many notes did you manage to play successfully?

- Think carefully about phrase boundaries and bar lines. Sloboda (1984) has suggested that the perceptual span of better sight-readers changes adaptively with the music. If the distance to the end of a phrase is shorter than six or seven notes, then the perceptual span decreases. In other words, skilled sight-readers are not just reading the music ahead of them note by note, but *interpreting* aspects of the musical structure first and changing their perception routines accordingly. Do you know where the end of the phrase is before you arrive at it?
- Novice sight-readers often find the ends of lines problematic, due to having to make large eye movements across the page. If you find that you frequently lose fluency while moving down a line, spend time practicing these moments particularly.

*Cognitive processing of visual information*

Research established early on that, in sight-reading tasks where a brief musical stimulus is presented, better and more experienced sight-readers remember longer sequences than less skilled players (Bean, 1938). This effect has been replicated several times since (e.g. W. B. Thompson, 1985). Integrating these results with the findings from research in eye movement implies that experts make eye movements that allow for a more efficient encoding of note sequences. This in turn leads to better memory for presented note sequences and a longer eye–hand span, since a longer lasting memory will facilitate encoding of a motor program.

These findings point toward a theoretical framework known as *long-term working memory* (Ericsson & Kintsch, 1995), which was developed as a general model of memory expertise in domains of skill such as chess playing and has more recently been applied to music (e.g. Williamon & Valentine, 2002). The long-term working memory concept proposes that during performance, experts are able to access the contents of their long-term memory (analogous to a computer hard drive) with an ease that is typically possible only for the contents of working memory (analogous to RAM in computers; see Chapter 7). In brief, experts acquire a kind of high-speed connection to long-term memory through extensive training, where previously only slower transmission rates existed. However, this high-speed connection is restricted to the domain in which the expert works (e.g. baroque music for a baroque specialist). This facilitating process has been shown to be a direct function of extensive training.

Related to memory is pattern recognition. Combinations of notes that occur frequently within a style are, over time, stored in memory as discrete entities rather than sets of individual events. They can then be easily retrieved from memory and used in novel contexts. An expert musician may see a certain familiar scale and immediately recognize what to play without further looking or thinking (see Waters *et al.*, 1997). More than this, though, different musical styles have their own characteristic patterns of notes and rhythms that recur and that are instantly recognizable from the page. This essentially means that highly familiar stimuli "stick out" more and attract attention immediately without the performer having to do much. Expert sight-readers can identify familiar patterns (e.g. a series of diminished seventh arpeggios) extremely quickly and with minimal cues.

Experienced sight-readers do not just take in larger chunks of visual information than those less experienced, they also generate more accurate predictions about what

may be coming up next. This process encompasses mechanisms of inference, anticipation, and guessing. Evidence for these mechanisms comes from at least three different sources. Firstly, sight-readers can be misled into making so-called *proof-reader's errors* (Sloboda, 1976), whereby they play notes that are not really there becuase they identify a familiar pattern without looking at the individual notes. (Did you spot a misspelling in the previous sentence? If not, that is a proofreader's error.) Secondly, experts are relatively more distracted by unexpected information (Waters *et al.*, 1997), suggesting that they rely on the context (e.g. harmonies) more in planning ahead. Finally, good sight-readers perform better on tasks where they are required to fill in blank spaces in a score with an appropriate note (Lehmann & Ericsson, 1996).

- The ability to recognize and infer patterns is based on a profound knowledge of the style in question. Listen often to the styles you have to sight-read, so that your anticipations become informed. With people in whom this is well entrenched, it is often interpreted as "intuition."
- Sometimes a scale or familiar pattern may contain unusual notes. Here, errors may occur in seemingly "easy" passages. If you are experienced, do not let yourself be completely guided by what you expect. Often you are correct, but sometimes you may not be. Beginners should try the opposite, namely to let the musical context help them anticipate what is coming up. Nobody can see everything on the page, and sometimes a "best guess" is all there is time for.

*Generation of movements*

While sight-reading does require the online generation of movements, clearly much of the material actually played will not be totally novel to the player. As musicians become familiar with a musical style they acquire a large body of knowledge in the form of common patterns, allowing them to draw on a storehouse of previously established motor sequences; for example, scale and arpeggio patterns are easy to play at sight because most musicians already know how to play them. However, sight-reading is more than simply a chaining together of predetermined motor sequences—as argued above, it is an open skill. It seems that some sort of plan must be formed *before* the action is actually executed.

In fact, performance on a musical instrument is an example of a task that implies the so-called *motor program* account of action control (Lashley, 1951). In this account, complex movements are specified by cognitive representations of individual actions, arranged hierarchically into programs. These programs are assembled before the actual execution phase and are best thought of as cognitive "maps" of the actions that are required.

Debate in motor programming research from the musical perspective has centered largely on how the timing of movement sequences is controlled. After all, one of the distinctive things about musical performance as opposed to other, outwardly similar skills is that the timing of actions is paramount (in typing, for example, the relative timing of keystrokes is irrelevant). By contrast, playing the right notes of a piece in the right order is not enough to give a musical result. But how are motor sequences timed? There are several possibilities, of which the most popular is the "internal timekeeper"

model, which proposes that individual components of the motor program are triggered at the correct moment by comparison with an imagined beat (see Palmer, 1997).

- When practicing sight-reading, choose a tempo at which you think you will be able to play the most difficult parts, and go no faster.

The precise mechanisms by which complex motor programs are constructed and converted into movements remain to be fully explicated. For the purposes of this chapter, however, it is enough to note that while musicians' knowledge of a musical style is primarily acquired through playing, factors unrelated to the musical material itself may impact on the motor program requirements. On woodwind and brass instruments, there is generally only one way of playing a particular note, and thus there are no alternatives open to the performer regarding which finger to use or where to position themselves in relation to the instrument. By contrast, keyboard and string instruments have a significantly greater freedom. This means that motor programs on these instruments must take account of ergonomic considerations. Sloboda *et al.* (1998) have provided evidence that expert sight-readers tend to select the most efficient, ergonomically viable fingering when sight-reading, enabling them to play faster and more fluently. For pianists, then, fingering choice may indeed be a determining factor of sight-reading expertise. In some cases, of course, there may be several alternative ways to play a given passage, all of them feasible.

- If applicable to your instrument, practice common patterns using as many different fingerings as you can.

### 8.1.2 Task analysis: Improvisation

Improvisation is most commonly discussed in the context of jazz, although many other musical styles incorporate a greater or lesser degree of improvisation (e.g. rock guitar solos and Indian art music; see Pressing, 1984). As with sight-reading, there is no one single concept of improvisation, but rather a spectrum of activities that involve improvisational elements. However, the range of cultures in whose music improvisation plays a part is significantly wider than sight-reading, since non-notated musics are the norm in our world rather than the exception (as we sometimes tend to assume).

Despite the popular cliché, improvisation is anything but "making it up as you go along." Genuinely free improvisation, in which the music has absolutely no externally observable point of reference (or *referent*; Pressing, 1984) is rare and has never become especially popular (the jazz saxophonist Ornette Coleman being perhaps the best-known exponent). Instead, improvised music is typically based around some predetermined structure or form: the chord sequence in jazz, the tune in Celtic folk music, the *maqam* in Islamic and the *raga* in Indian musical traditions, and so on. In each of these styles, some aspect of the music is left undetermined or unrealized—often the lead melodic line—and the performer is expected to improvise an acceptable solution. The acceptability of the improvisation lies in its relation to stylistic conventions and, consequently, appreciation of the improvisation often requires that the listener have an understanding of the stylistic boundaries within which the performer is working.

Good improvising is thus, in one sense, the production of novelty within identifiable constraints.

- Have you ever tried improvising? Because improvisation is now so strongly associated with jazz, many classical musicians never even attempt it; however, it is quite possible to improvise in any style. Many of the more obviously improvisational elements in classical music have become less common these days (e.g. improvised cadenzas), but this does not mean that they cannot be revived, even if only for use in practice rather than performance.
- Good improvising in any musical genre relies on a thorough knowledge and understanding of the style. The improvisation is successful because of the way it relates to established stylistic norms. It is, therefore, essential for performers to be well-versed with the style in which they wish to improvise.

*Pattern chaining versus more complex models*

Judging from method books on jazz improvisation, and from hearing amateurs play, one may think that improvising consists simply of chaining together pre-learned patterns. However, the process seems to be more complex than it appears at first sight. Johnson-Laird (1991) was the first to develop a cognitive theory that contradicts the "motif" notion of improvisation. This model assumes the operation of a generative process, somewhat akin to that proposed for speech production. More recently, Johnson-Laird (2002) has developed a model that entails the procedural production of possible patterns according to certain rules. For experts, this process is more or less automatic. The resulting patterns are then subjected to conscious, deliberate evaluation that allows the musician to pick the most successful of the produced patterns. This procedure can account for situational aspects, such as the mood of the audience ("they seemed to like that phrase; I'll do it again"), broader aesthetic considerations ("with these cats, I have to be a lot more sophisticated"), or the response to music just played by another musician in the group. Taking on Johnson-Laird's view means having to rethink the idea of improvisation as simply pulling patterns from a storehouse.

- Try to be aware of your thought processes while improvising—what kind of things go through your mind? Do you make conscious decisions about what to play? While some experienced improvisers report achieving so-called "flow" states, in which they are not aware of exerting deliberate control, it seems likely that conscious decision-making is a normal part of skilled improvisation.

*Previous knowledge in a secondary function*

While successful improvisation is clearly not just a matter of pattern-chaining, it certainly does require a large amount of previous knowledge. This is simply because of lack of time at the moment of production. Having to produce everything during performance would be impossible; therefore, the musician needs to buy time. Here, knowledge can be viewed as having the primary function of serving as a building block during skill acquisition, when production rules and patterns have to be learned. Later, however, this application may serve simply as a subordinate process to the more

important goal of making good music. This can be done by either retrieving patterns from memory or by generating adequate patterns using entrenched algorithms (experienced public speakers such as politicians seem to do this quite well). The remaining cognitive resources can then be used to address artistic issues.

- Do you have your routines sufficiently well practiced that you could do something else at the same time? Perhaps test yourself by conversing with someone while playing.

*Motor programming*

As in sight-reading, the improvising musician must be able to generate and execute motor programs "online." However, while the number of variables open to manipulation by the performer seems to be greater in the case of improvised music than sight-reading, their solutions are inevitably expected to be style-specific. Therefore, successful improvising, like sight-reading, relies on the player having access to a large stock of style-specific knowledge, in the form of common patterns and sequences, from which to compile a suitable motor program.

Improvisers tend to program what Schramowski (1973, p. 239) calls "movement images;" they anticipate entire movement sequences that hinge on particularly salient notes. Improvisers are also subject to constraints determined by the physical characteristics of the instrument, which can be obvious (e.g. pianists cannot "bend" notes) or rather more subtle. Furthermore, individual performers will have anatomically or biographically determined typical movement patterns that factor into the equation. A pianist with large hands will be able to play chords and patterns that would not be available to others. The great jazz guitarist Django Reinhard suffered severe burns to his left hand and was forced to adopt a novel technique that used only two fingers on the fretboard; jazz guitarists have subsequently found his style of playing very difficult to imitate.

- Observe yourself in a mirror as you improvise; are there distinctive patterns of movement that occur in your playing? Try using particular types of movement as the basis for your improvisation and see what effect this has.
- What are the characteristics of your instrument? Spend some time identifying the patterns that "sit under the fingers," and then try improvising without using any of them. Another idea is to attempt to copy improvisations from other instrument. If you are a saxophone player, try playing improvised lines by, say, pianists or trumpet players.
- What are your own physical characteristics? Are there capabilities that you have which you do not exploit, or perhaps that you overexploit?

### 8.1.3 Two sides of the same coin?

From the analyses developed above, it should be clear that sight-reading and improvising have a good deal more in common than might have been initially predicted, and certainly more than their popular descriptions would suggest. To recap: both sight-reading and improvising require the "online" creation and execution of motor program sequences in response to a stimulus. This essentially breaks down into two component

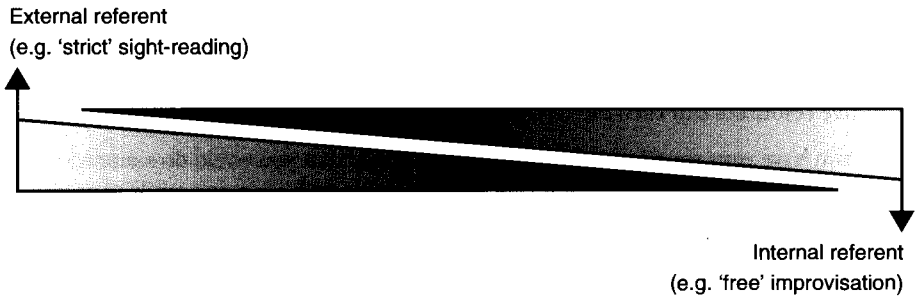



**Figure 8.1** Schematic diagram illustrating the difference between sight-reading and improvising in terms of the internal/external nature of their referents. Most real musical tasks fall somewhere between the two extremes.

skills: (1) *planning*, the ability to encode and process information from some stimuli quickly and fluently, and (2) *execution*, the ability to translate those plans into controlled, accurate motor movements.

In this analysis, the key difference between improvisation and sight-reading seems to lie in the nature and extent of the referents that specify what is to be played. Specifically, referents could be characterized as falling between the two extremes of *external* and *internal*. Completely external stimuli would be those with sufficient referents to specify the desired outcome in its entirety, without any freedom available to the player (e.g. in a "strict" sight-reading task the score details every aspect of the music to be performed). Conversely, wholly internal stimuli would be those for which no referent exists other than the performer's own intention.

This relationship can be encapsulated schematically, as illustrated in Figure 8.1. Both extremes are hypothetical. Even the most complex modern notation is not sufficiently specific as to detail every aspect of the musical outcome. As for "free" improvisation, it is certainly debatable whether this is even possible in principle, given the highly acculturated experience of most musicians). Nonetheless, typical sight-reading tasks lie toward the "external" end of the scale; similarly, the most commonly discussed improvisation practices lie toward the "internal" end. Taking this perspective on sight-reading and improvising implies at least one provocative hypothesis, namely that a player who is skilled at sight-reading within a given musical style should feasibly be able to improvise within that style after some practice (see McPherson, 1995, for supporting empirical evidence).

## 8.2 Acquiring and enhancing sight-reading and improvising skills

Research has been slow in answering the practitioner's most pressing question, namely how sight-reading and improvising skills are best learned and improved. The reason may be that common practices seem to suffice in producing satisfactory results on the whole and that, consequently, no one is eager to invest time in researching new ways to improve the acquisition process. Such differences as exist between individuals' sight-reading and improvising abilities are often ascribed to variations in innate disposition.

This somewhat apathetic view is regrettable because it is clear from the relatively little research that does exist that both sight-reading and improvising are amenable to enhancement through structured training (see e.g. Sudnow, 1978; Kornicke, 1995; Lehmann & Ericsson, 1996; Kenny & Gellrich, 2002; Souter, 2002; Hoffmann & Lehmann, 2003). Unfortunately, these findings have barely filtered into music education systems.

### 8.2.1 Skill acquisition and development

Improving sight-reading is commonly presented as a matter of trial and error, the methodology being simply to practice slowly and gradually build up to the desired speed and complexity. If improvisation is deliberately taught at all, the emphasis is usually on the need to listen to skilled improvisers and copy them; a "mastery through imitation" approach rather similar to the manner in which painting skills were traditionally acquired (see Lisboa *et al.*, in press). In both cases, the mode of learning—and thus the knowledge gained—tends toward being *implicit* rather than *explicit* (i.e. when the individual has conscious access to, and is able to articulate, constituent parts of the skill). While much is known about the differences between good and bad sight-readers, the acquisition of ability in improvisation is still largely uncharted territory, which is probably due to the difficult methodological problems that beset researchers trying to study it.

It is likely that all musical skills (playing by ear, sight-reading, improvisation, playing rehearsed music) improve with time, if the student keeps playing the instrument and rehearses new pieces of music. This prediction is based on the assumption that some transfer occurs from rehearsed music to the other skills. This is hardly surprising, given that all the skills share so many basic requirements. No study has explicitly tested the claim, although McPherson's work (1994, 1995) suggests moderate to large correlations between ability in all these skills. However, despite this transfer it seems likely that significant progress can be made in performance of the individual skills when they are trained separately.

Individual differences among piano sight-readers have been related to varying amounts of sight-reading experience (Kornicke, 1995) and to more rigorous measures of sight-reading training (Lehmann & Ericsson, 1993, 1996). Better sight-readers on the piano tend to have accumulated many hours as accompanists and have a larger repertory of pieces they can use for accompanying. It might seem that one follows from the other, but the effects of the two factors (experience accompanying and repertory size) have been shown to be statistically independent. Most pianists start accompanying about 3 years after beginning piano lessons and work with instrumentalists and singers that they can technically accommodate at the piano. Gradually, their accompanying becomes more demanding. As this cycle continues, sight-reading skills improve, as do general pianist skills.

Not a great deal is known from a scientific standpoint about how to acquire improvisation skills. Hargreaves *et al.* (1991) found that novices tended to be much more detail-oriented than experts, who had a plan before they started playing. Also interesting is Sudnow's (1978) account, which documents how the tedious and effortful process of improvisation becomes quasi-automatic, and how eventually the performer

(in this case Sudnow himself) merely watches his hands do surprising things while concentrating on aspects that may or may not bear relation to them (see Kenny & Gellrich, 2002). As discussed above, successful improvising in jazz is highly contingent upon a large knowledge base of style-specific patterns. At this point, however, it is not really known how learning to use these patterns ultimately results in innovative and fluent performance, and not merely a stringing together of overlearned patterns (see Johnson-Laird, 2002). Kenny and Gellrich (2002) call for *deliberate practice* and *transcendence*, whereby the first is a process of amassing a large knowledge base through systematic practice routines, and the second is "understood as a heightened state of consciousness that moves beyond the confines of (thereby often jettisoning) the accumulated knowledge base itself" (p. 124). What is not clear is precisely how to get from one stage to the other.

The relative paucity of applied pedagogical research in either area makes it difficult to recommend specific practice strategies with confidence. The following suggestions, then, are intended as ideas to stimulate practice rather than tried-and-tested methods for improving.

*Acquiring the knowledge base*

- This has been noted above, but is so important it merits repetition—*know the style of music in which you wish to play.* Successful performance on both skills depends absolutely and utterly on having a wide knowledge of the style in question.
- You can acquire knowledge of a style without having to play it. Both listening to recordings and looking at scores are excellent ways of becoming acquainted with a range of music. Try to ask yourself questions about what you hear. How is it put together? What are its characteristic features and patterns?
- Devise your own exercises. While there are many books of technical exercises available, it can be more beneficial to create your own. This requires not only playing the exercise(s) until you achieve a high level of competence, but also analyzing the music to extract the most relevant aspects from it. This, again, is a good way to expand the knowledge base.
- Sing, hum, or whistle. A common frustration of learning to improvise or sight-read is having a strong idea of what you wish to play, but being unable physically to find the notes. Ultimately, of course, perceptual and cognitive skills will have to meet with motor skills in order to produce a successful improvisation or sight-read performance. However, research suggests that the two areas can develop independently, especially at early stages. Try sight-singing, or singing an improvised melody line.

*Things to consider when practicing*

- McPherson (1994) has suggested that better sight-readers are able to assimilate more information about the score *before* beginning to play than those who are less skilled. In other words, they scan the music and quickly notice features such as time and key signature, phrase structure, passages likely to be problematic, and so on. Try scanning a new piece for a few seconds and then writing down what you remember about it. How much information did you gather on first glance?

- The illusion of a well thought-out interpretation in sight-reading comes from applying rules of musical expression, and arguably these "performance rules" can be deliberately learned in much the same way that we learn the more fundamental parameters of music like the meaning of pitch and rhythm symbols (see Chapter 13 of this volume). Think about what some of these rules could be by analyzing a rehearsed performance. For example, what usually happens to the tempo at major cadence points in classical music?
- Be aware of what your referents are. When improvising, think carefully about what requirements your improvisation needs to meet. Do you need to arrive and finish on a particular beat or highlight a particular chord change? If you are sight-reading, it can be useful to reflect on the information that is *not* provided—what alternatives are open to you in your interpretation?

*Some general thoughts*

- Do not be put off if things are difficult at first. Initially, sight-reading and improvising take a great deal of cognitive effort, but by practicing, you are gradually making automatic many of the processes that at first seem to require a lot of thinking.
- Both improvising and sight-reading require a certain degree of nerve and (at least at the beginning) self-confidence. Because of this, both are also good ways of learning to take artistic risks, making you rely on your own abilities and technical facilities. The confidence to take risks at the right moment can be learned, like anything else.
- There is little evidence to suggest that "talent" has anything much to do with proficiency in sight-reading or improvising. Rather, it is a case of diligent and inventive practice.

### 8.2.2 Identifying problems

After the initial stages of improvement, it is common for skills to reach a plateau where the same problems seem to recur, and it is difficult to make further progress. If this happens, it can be useful to try to understand the nature and causes of these problems; examining errors provides a means for players to make explicit the implicit knowledge they possess. By identifying the types of error to which they are prone, players can become aware of weaknesses in their technical ability and develop strategies that focus on these specific areas.

General theories of failure in human performance (e.g. Reason, 1990) posit that errors can arise from one of two sources. Either the action itself can be incorrectly performed (this is often labeled an *execution failure* or a *slip*), or the action can prove to have been insufficiently specified to achieve its intended outcome, correct execution notwithstanding (this is a *planning failure* or *mistake*). This distinction accords well with the analyses of sight-reading and improvising given above. Note also that in both cases the *intention* of the person performing the action is critical to determining whether an error has taken place.

The types of errors that can be made in sight-reading may be usefully thought of as arising from either misreading (where the error occurs during the processing stage) or misexecution (where the error occurs during the execution stage). In terms of the

distinction between mistakes and slips, a wrong note caused by misreading seems to be a mistake in that it reflects the correct execution of a flawed plan. By contrast, the case of a correct note being misstruck (e.g. on a piano keyboard) is a slip, as it reflects not a failure of planning but of execution (see Thompson *et al.*, 2002, for further discussion of this distinction).

This system can provide a useful method of self-diagnosis. By making careful note of the types of error that they commit most frequently, players may be able to identify the particular aspects of sight-reading that are causing problems and in most need of work. This in turn could prompt the development of appropriate exercises focusing on that particular aspect, perhaps along the lines suggested earlier in this chapter.

- There is more to errors than just wrong notes. Record yourself sight-reading and then listen to the mistakes you made. Try dividing them into categories; for example, incorrect rhythm, incorrect note, note misstruck, and so on. Then, in each case, try to decide whether the error occurred due to misreading (e.g. the number of leger lines), forgetting (e.g. an accidental from earlier in the bar), or failed execution (e.g. missing the key). Does a pattern emerge?

The issue of making errors in improvised music is at once the same and rather more complex than in sight-reading. While there are "no wrong notes in jazz, just possibilities," conceptually speaking there is not much difference between errors in sight-reading and errors in improvising. It is perfectly possible to claim that an error has occurred in an improvised performance when a player either fails to execute a passage in the way he or she intended, or else deliberately plays a passage in such a way that he or she later regards to be unsuccessful (e.g. due to an incorrect anticipation of the harmonic progression). The difficulty arises in trying to identify precisely what the intention was in the first place, and at what stage in the process a failure occurred that led to it being unsatisfactorily realized. Unlike in the sight-reading situation, there may be little or no outward clue as to what was actually intended, and this is something about which players themselves often find it difficult or even impossible to be explicit. Some educated guesswork may help but necessarily relies on *a priori* knowledge of the player's personal style and musical habits. To those only familiar with "traditional" jazz, the modal improvisations of Miles Davis over otherwise functional tonal harmony might sound like a succession of wrong notes played with haphazard rhythm. The kind of self-analysis recommended for improving sight-reading may be a profitable exercise for those concerned with enhancing their improvisational skills; however, individual instances of error may be hard to identify.

- There need be no contradiction between improvising with originality and having conscious awareness of what you want to play. Much colloquial discussion of improvising speaks as if it is a mystical process by which notes come "from within"—this is just a function of conscious and deliberate processes becoming automated over time. Do not be afraid of thinking about what you play.
- As with sight-reading, using recordings is often the best way to analyze your own playing. When listening to yourself improvise, highlight the areas of the improvisation that you

are unhappy with and try to work out precisely what you played, and what you were intending to play.

- Sometimes when improvising you do not make errors as such, but feel that what you are playing is unoriginal or otherwise unsatisfying. If this happens, try imposing artificial constraints on your playing (for example, only allowing yourself to play within a certain range, or deliberately avoiding certain notes). This can force you to take your improvisation in different directions by breaking your so-called "mental set."

## 8.3 Conclusions

The most crucial challenge in the mastery of either sight-reading or improvisation is to maximize the use of the knowledge base, both by becoming as familiar as possible with the musical style in question and by mastering the motor sequences that are required to execute the patterns that typically occur. The more stylistically relevant patterns are "under the fingers," the more one's limited cognitive resources will be free for use in artistic decision-making or for dealing with unexpected performance problems. A dividing factor between sight-reading and improvising is that the ultimate goal of the improviser is to become a uniquely recognizable artist with a distinct personal voice, whereas the goal of sight-reading is to create the perfect illusion of a rehearsed performance despite the lack of rehearsal. A common factor is the necessary ability to adapt skills to a changing environment.

It seems likely that many musicians would benefit from spending more time engaged in both sight-reading and improvising, and within a range of musical styles. They are not mysterious and extraordinary feats of skill, out on their own in terms of requirements. On the contrary, they are heavily reliant on the same knowledge and psychomotor skills that underpin successful performance, memorization, and (to an extent) ensemble playing. Certainly, they are both useful abilities for the musician to possess, and in some cases they are essential ones. More than this, however, they are skills that—if properly and diligently practiced—can develop intuition and bring new levels of musical awareness. Even for the musician who does not need, or dare, to engage in sight-reading or improvising outside the practice room, the potential benefits to their overall musical ability are clear. One could regard them as enhancement strategies in themselves.

## Further information and reading

Kenny, B. J., & Gellrich, M. (2002). Improvisation. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 117–134). Oxford: Oxford University Press.

Hill, P. (2002). From score to sound. In J. Rink (Ed.), *Musical Performance: A Guide to Understanding* (pp. 129–143). Cambridge: Cambridge University Press.

Lehmann, A. C., & McArthur, V. (2002). Sight-reading. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 135–150). Oxford: Oxford University Press.

Rayner, K., & Pollatsek, A. (1989). *The Psychology of Reading.* Hillsdale, NJ: Erlbaum.

## References

Bean, K. L. (1938). An experimental approach to the reading of music. *Psychological Monographs, 50,* 1–80.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102,* 211–245.

Hargreaves, D. J., Cork, C. A., & Setton, T. (1991). Cognitive strategies in jazz improvisation: An exploratory study. *Canadian Journal of Research in Music Education, 33,* 47–54.

Hoffmann, A., & Lehmann, A. C. (2003). Anfänger und Profis bei der Improvisation: Unterschiede als Wegweiser für Übung und Unterricht [Experts and novices in improvisation: Differences as a guide for practice and teaching]. *Üben & Musizieren, 20,* 35–42.

Johnson-Laird, P. N. (1991). Jazz improvisation: A theory at the computational level. In P. Howell & R. West & I. Cross (Eds.), *Representing Musical Structure* (pp. 291–325). London: Academic Press.

Johnson-Laird, P. N. (2002). How jazz musicians improvise. *Music Perception, 19,* 415–442.

Kenny, B. J., & Gellrich, M. (2002). Improvisation. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 117–134). Oxford: Oxford University Press.

Kornicke, L. E. (1995). An exploratory study of individual difference variables in piano sight-reading achievement. *Quarterly Journal of Music Teaching and Learning, 6,* 56–79.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral Mechanisms in Behavior: The Hixon Symposium* (pp. 112–136). New York: Wiley.

Lehmann, A. C. (in press). Vomblattspiel und Notenlesen [Sight-reading and note reading]. In H. Stoffer & R. Oerter (Eds.), *Enzyklopädie der Psychologie*: Vol. D/VII/1. *Allgemeine Musikpsychologie.* Göttingen, Germany: Hogrefe.

Lehmann, A. C., & Ericsson, K. A. (1993). Sight-reading ability of expert pianists in the context of piano accompanying. *Psychomusicology, 12,* 182–195.

Lehmann, A. C., & Ericsson, K. A. (1996). Structure and acquisition of expert accompanying and sight-reading performance. *Psychomusicology, 15,* 1–29.

Lehmann, A. C., & McArthur, V. (2002). Sight-reading. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 135–150). Oxford: Oxford University Press.

Lisboa, T., Williamon, A., Zicari, M., & Eiholzer, H. (in press). Mastery through imitation. *Musicæ Scientiæ.*

McPherson, G. E. (1994). Factors and abilities influencing sightreading skill in music. *Journal of Research in Music Education, 42,* 217–231.

McPherson, G. E. (1995). The assessment of musical performance: Development and validation of five new measures. *Psychology of Music, 23,* 142–161.

Nettl, B., & Russell, M. (Eds.). (1998). *In the Course of Performance: Studies in the World of Musical Improvisation.* Chicago: University of Chicago Press.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48,* 115–138.

Pressing, J. (1984). Cognitive processes in improvisation. In W. R. Crozier & A. J. Chapman (Eds.), *Cognitive Processes in the Perception of Art* (pp. 345–363). Amsterdam: Elsevier.

Rayner, K., & Pollatsek, A. (1989). *The Psychology of Reading.* Hillsdale, NJ: Erlbaum.

Reason, J. (1990). *Human Error.* Cambridge: Cambridge University Press.

Sargent, J., Zuck, E., Terriah, S., & MacDonald, B. (1992). Distributed neural network underlying musical sight-reading and keyboard performance. *Science, 257,* 106–109.

Schramowski, H. (1973). Schaffenspsychologische Untersuchungen zur instrumentalen Improvisation [Psychology of creative behavior regarding instrumental improvisation]. *Beiträge zur Musikwissenschaft, 15,* 235–251.

Sloboda, J. A. (1976). The effect of item position on the likelihood of identification by inference in prose reading and music reading. *Canadian Journal of Psychology, 30,* 228–236.

Sloboda, J. A. (1984). Experimental studies of music reading: A review. *Music Perception, 2,* 222–236.

Sloboda, J. A., Clarke, E. F., Parncutt, R., & Raekallio, M. (1998). Determinants of fingering choice in piano sight-reading. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 185–203.

Souter, T. (2002). Manipulating working memory to improve sight-reading skills. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the Seventh International Conference on Music Perception and Cognition* (pp. 639–642). Adelaide, Australia: Causal Productions.

Sudnow, D. (1978). *Ways of the Hand: The Organisation of Improvised Conduct.* London: Routledge and Kegan Paul.

Thompson, S., Valentine, E., & Williamon, A. (2002). Error patterns in piano sight-reading. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the Seventh International Conference on Music Perception and Cognition* (pp. 643–646). Adelaide, Australia: Causal Productions.

Thompson, W. B. (1985). *Sources of Individual Differences in Music Sight-Reading Skill.* Unpublished doctoral dissertation, University of Missouri, Columbia, MO.

Waters, A. J., Underwood, G., & Findlay, J. M. (1997). Studying expertise in music reading: Use of a pattern-matching paradigm. *Perception and Psychophysics, 59,* 477–488.

Williamon, A., & Valentine, E. (2002). The role of retrieval structures in memorizing music. *Cognitive Psychology, 44,* 1–32.

PART III

# TECHNIQUES AND INTERVENTIONS

CHAPTER 9

# PHYSICAL FITNESS

ADRIAN H. TAYLOR AND DAVID WASLEY

The link between a healthy body and healthy mind is not a new concept for optimizing performance and general functioning in life. Recent acknowledgment of this has come from violinist Maxim Vengerov in a television documentary, who noted: "I try to work out regularly. Keeping fit as a musician is important" (NVC Arts, 1998). Indeed, physical *inactivity* is not something for which the human body was designed; however, in modern life, it is all too easy to become enticed by technological and social change that can remove opportunities to be physically active. As a result, sedentary (or "hypokinetic") related health problems seem to be dramatically on the rise (Sallis & Owen, 1999; US DHHS, 1996).

Across the broad spectrum of educational and professional music performance contexts, general physical fitness is an area that receives surprisingly little formal emphasis. Admittedly, the connections between physical exercise, personal well-being, and performance are easy to overlook. Until recently, in fact, this link has been largely ignored by researchers, with only scant evidence in the current literature on how physical and psychological well-being can interact for the betterment of the musician.

This chapter is broadly about these links and is written for performers and anyone else responsible for enhancing the well-being and performance of musicians. Its aim is to introduce applications of physical activity to music and to explore how practitioners can integrate exercise into their lifestyle in ways that can optimize their skills. It considers how different aspects of physical activity—including frequency, intensity, duration, and type—impact on physiological and psychological responses to performance, focusing primarily on the well-documented health benefits of moderate intensity, aerobic-type exercise (see Table 9.1 for typical examples of vigorous, moderate, and low-intensity activities). However, because of the absence of literature directly applied to musicians, much of the initial evidence and rationale behind the music-specific benefits are drawn from other fields. Nevertheless, research findings from recent studies with musicians are presented in which the effects of both regular and single sessions of exercise (just before performance) have been examined. The results confirm Vengerov's observation and support the notion that exercise can give the musician valuable insight and tools for managing the demands of being a performing artist.

## 9.1 Background

Exercise scientists and epidemiologists have, in the past 50 years, systematically accumulated a considerable body of knowledge about the benefits of physical activity, to

Table 9.1 Examples of vigorous, moderate, and light exercise, sports, and general physical activity (adapted from Blair *et al.*, 1985)

| Vigorous | Moderate | Light |
|---|---|---|
| *Exercise and sports* (if out of breath or sweaty) | *Exercise and sports* (if not out of breath or sweaty) | *Exercise and sports* (if not out of breath or sweaty) |
| Aerobics | Aerobics | Bowls |
| Cycling | Cycling | Darts |
| Racquetball/squash | Golf | Fishing |
| Running | Swimming | Golf |
| Soccer | Soccer | Pool/snooker |
| Swimming | Table Tennis | Table tennis |
| Tennis | Tennis | Long walks (average pace) |
| | Long walks (brisk or fast pace) | |
| *General physical activity* | *General physical activity* | *General physical activity* |
| Occupations that involve frequent climbing, lifting, or carrying heavy loads | Occupations that are active but not vigorous | Occupations that are not entirely sedentary |
| | Heavy "do it yourself" activities (e.g. mixing cement) | Light "do it yourself" activities (e.g. decorating) |
| | Heavy gardening (e.g. digging) | Light gardening (e.g. weeding) |
| | Heavy housework (e.g. spring cleaning) | Social dancing |
| | Social dancing | |

the point where a consensus now exists on its impact on many facets of health and well-being. Expected health gains from specific increases in physical activity intensity, duration, frequency, and type have partly been established for:

- cardiovascular disease and associated risk factors, such as obesity, hypertension, and type 2 diabetes
- cancer (Batty & Thune, 2000)
- musculoskeletal problems (Videman *et al.*, 1995; Vuori, 2001)
- osteoporosis (Bonaiuti *et al.*, 2003).
- Dimensions of psychological well-being and mental health, such as reduced anxiety and depression, and enhanced cognitive functioning and self-perceptions (Steptoe & Cox, 1988; Moses *et al.*, 1989; Kubitz *et al.*, 1996; Morgan, 1997; Biddle *et al.*, 2000; Lawlor & Hopker, 2001).

Being overly sedentary can, for instance, triple the risk of cancer and diabetes. At the other end of the continuum of physical activity, it is equally important to consider how physical inactivity impacts on our emotions, mood, and affect; inactivity may lead to negative moods (e.g. restlessness, irritability, fatigue, and depression) and a negative affective state (low activation and high tension or "bottled-up energy"). Guidelines for the amount of physical activity that people should do—musicians included—to prevent health problems associated with an inactive lifestyle have been produced according to the above evidence base (see Pate *et al.*, 1995). The internationally recommended minimum amount for cardiovascular health is to accumulate at least 30 minutes of moderate intensity exercise (e.g. brisk walking) on five or more days per week

(Murphy *et al.*, 2002). A debate still exists about whether it is the number of sessions or total energy expenditure that is more important.

What is evident, however, is that lasting changes take time and are rarely as rapid as people want. Take weight management, for example. It is now accepted that calorie restriction alone is insufficient to bring about long-term, positive weight alterations and body change. Achieving such change requires both increased physical activity as well as a balanced diet, roughly according to the following relationships:

- Weight gain occurs when "energy in" is more than "energy out".
- Weight loss occurs when "energy in" is less than "energy out".
- Weight stability is achieved when "energy in" equals "energy out".

Here, "energy in" is the food and drink one consumes; "energy out" is the physical activity and daily living energy (i.e. metabolic rate) of the individual. Importantly, in comparison to diet restriction, physical activity positively stimulates muscular development in a number of ways to deal with the demands placed on the body. Firstly, a change in the balance between fat and lean muscle can be realized. Muscle is denser than fat and takes up less space for an equal amount of weight; thus, individuals who become physically active over a prolonged period experience positive changes in body shape (e.g. tone and firmness, as well as possible reductions in overall weight). Secondly, more muscle is beneficial because it increases metabolism, which in turn impacts on the energy relationships mentioned above. Specifically, a person with more muscle mass is able to eat more than an equivalently weighted person and can still maintain weight stability.

It is worth noting that the general evidence for the health-related benefits of increased physical activity is based on statistical relationships and provides no guarantee for any one individual. Genetics and other environmental factors may limit or refine the predictive power of these statistical relationships. For example, it is possible to find a healthy older person—or a top performer, for that matter—who has never been particularly active. It is therefore common for an individual's vicarious experience to determine levels of motivation for changing lifestyle, rather than relying on health information based on statistics and probability. Nevertheless, the fields of preventive and public health medicine are now beginning to employ strategies for changing the motivations of populations and individuals for adopting a healthy lifestyle.

But how exactly can physical activity offer a positive impact on musical performance? As shown above, there are a number of benefits that can be derived from being physically fit, and certainly, any of these may have a direct or indirect effect on an individual's performance ability. This chapter focuses on what is commonly described as stress or performance anxiety (see also Chapter 1). Individual performers will prepare for a performance in different ways, with each method consciously or subconsciously designed to facilitate the "right" mental and physical state for optimum performance. Performance stress can arise from cognitive appraisal and an imbalance between the perceived demands of the situation and the perceived ability of the performer (Lazarus & Folkman, 1984). In addition, other events outside of the musical setting may cause stress, which can then bear on one's performance state. The challenge of achieving optimum performance is to replicate one's own ideal performance state consistently.

To do this, it is necessary to be calm—yet energized—and confident, and not to be distracted by negative emotions and thoughts. It is typically recommended, therefore, that a pre-performance routine (i.e. the hours just before a performance) be designed to provide consistency in building one's mental and physical state for performance (also see Chapter 12).

Research on the effects of physical fitness on responses to stress has generally focused on two broad types of exercise programs: (1) *chronic exercise*, which is regular exercise participation, and (2) *acute exercise*, which is a single session of exercise (lasting at least 20 minutes) that can be embedded into a pre-performance routine. Firstly, the "cross-stressor hypothesis" suggests that fitter people will show less cardiovascular response to physical and psychological stressors (Sothmann *et al.*, 1996). Chronic exercise improves fitness and, thereby, reduces cardiovascular (e.g. heart rate and blood pressure) and neuroendocrine (e.g. cortisol) response to both physical and psychological stressors. Awareness of reduced physiological arousal may, consequently, result in lower perceived physical symptoms (somatic anxiety) and tension. Secondly, within a pre-performance routine, acute exercise has been shown to impact positively on arousal, anxiety, and other dimensions of affect and mood. Exercise physiologists have long known about post-exercise hypotension (i.e. reduced blood pressure) that can last for several hours after moderate intensity aerobic exercise. Given that some musicians and other performing artists turn to beta-blockers, alcohol, and tranquilizers to cope with the stress from performing (Steptoe & Fidler, 1987; Steptoe, 1989; also see Chapter 14 of this volume), it seems reasonable to ask whether a behavioral approach might be equally effective. From the existing research literature on physical activity and well-being, several key questions emerge in relation to music:

- Is physical activity particularly common among musicians?
- Can physical activity impact on the psychological well-being of musicians (which may, in turn, impact on musical performance), in what are clearly occupationally demanding circumstances?
- Do fitter musicians respond differently to the psychological stress of a musical performance?
- Does engaging in regular exercise change how musicians physiologically and psychologically respond to a performance?
- Can a single session of exercise (within a pre-performance routine) influence physiological and psychological responses to a musical performance?

A top musical performer, like an athlete, should be a finely tuned individual. To date, however, more is known about tuning an athlete than a musician. Nevertheless, the following sections consider the questions raised above, and evidence is presented that weighs up the musical relevance of chronic and acute exercise.

## 9.2 Chronic exercise

*Fitness* is usually used to describe an ability to perform a particular task. Being fit as a musician will imply having the necessary endurance, strength, and coordination to

function effectively on a specific musical instrument. In exercise physiology, *endurance fitness* is defined by assessing physiological response on an exercise ergometer (i.e. a machine, such as a treadmill or stationary cycle, devised to precisely control exercise intensity). By noting the change in response (e.g. in heart rate) to increments in exercise load/intensity, a measure of endurance fitness can be derived. Some people naturally (due to genetics) have lower cardiovascular responses to physical and psychological stress and at rest. Comparing fit and less fit people on different indicators of psychological well-being is, therefore, limited. Some naturally "fit" people may do little or no exercise and still be assessed as fitter than more active people on standard fitness tests. However, many studies have examined the effects of chronic exercise on psychological well-being, compared with inactive control groups. In one useful review, Biddle *et al.* (2000) conclude that chronic exercise is associated with increases in positive mood and affect (e.g. vigor and activation) and reductions in negative mood (e.g. tension, fatigue, and depression). Mutrie (2000; see also Lawlor & Hopker, 2001) concludes that chronic exercise reduces depression, while Fox (1997, 2000) reviews considerable evidence that shows improved physical self-perceptions and physical self-concept. Boutcher (2000) reports that exercise training improves cognitive functioning (i.e. speed and capacity to process information), and Taylor (2000) has confirmed previous evidence (Petruzzello *et al.*, 1991) that supports a positive effect of chronic exercise on trait anxiety (i.e. a disposition to be anxious). Other research has even begun to show beneficial effects of low to moderate intensity exercise (e.g. brisk walking), lasting only 10 minutes (e.g. Thayer, 1987; Ekkekakis *et al.*, 2000).

One of the complexities of this work is that there is undoubtedly personal meaning attached to exercise. For example, the thought of having to join a group of other exercisers may bring back memories of physical education classes and elicit fear of being breathless. Displaying our bodies in exercise settings may also bring out what is called social physique anxiety. On the other hand, exercise performed in supportive groups or with friends, and which provides a sense of mastery over time (e.g. becoming fitter or improving a movement skill), can facilitate anxiety reduction. The challenge for exercise scientists has been to identify the psychobiological versus psychosocial outcomes of different exercise intensities, durations, and frequencies. Initially, it was believed that the dose of exercise had to be sufficiently high to improve cardiovascular fitness before psychological benefits would accrue. It is now known that this is not necessarily the case. Reductions in trait anxiety can result from regular exercise at low to moderate intensity without observation of an improvement in cardiovascular fitness as determined from a fitness test. However, this finding may be due to the fact that fitter and more active people have tended to be involved in these studies and that there is less opportunity for them to improve anyway. Inactive and less fit people who are generally anxious are likely to improve fitness with chronic exercise of sufficient duration—at least 4 weeks—while concurrent improvements in anxiety will also be evident.

If chronic exercise (particularly of the aerobic type) is to be recommended for reducing one's disposition to be anxious and help manage stress, then some insight needs to be provided as to how this could possibly work. If the same effects could be observed from a more passive form of exercise (e.g. yoga, tai chi, or Alexander technique) with fewer environmental constraints, then it is likely that, disregarding the

other physical benefits that accrue from this form of exercise, many would choose the easiest option. However, research has shown that low-intensity exercise does not have the same beneficial effects on health as moderate-intensity exercise, even when carried out for the same duration and frequency. To date, there is a much greater body of knowledge about moderate and vigorous aerobic exercise than less intensive (e.g. yoga, tai chi, and Alexander technique) or resistance exercise (e.g. circuit and weight training), but recent valuable work has begun to examine more closely forms of low-intensity activity that are sustainable, accessible, and popular (e.g. tai chi, particularly among older people).

One possible explanation for the anxiety reducing effects of chronic exercise is distraction. Regular mental focus on something that prevents rumination about ongoing worries and stress can certainly be useful. On the other hand, an exercise that demands only limited information processing may very well not have the same benefits; for example, walking alone at a low intensity can still allow for rumination. More quickly paced exercise forces the mind to think about the physical symptoms of the activity itself. Similarly, exercise with others (at a pace that still enables conversation) or in an environment that distracts or demands attention is also likely to be beneficial. Either way, exercise needs to be performed regularly to develop optimal attention control.

Another explanation is that exercise involves improvements in self-confidence and self-esteem. People may label themselves as "not the sporty kind" or "not very good at exercise." A hierarchical model of physical self-perceptions (Fox, 1997) would suggest that the more negative labels one has, the lower will be the person's self-perception across different domains (e.g. academic, musical, social, and physical). According to this theory, if there is only one domain in which a person feels good about him- or herself (e.g. music performance), then that person's self-esteem will be relatively fragile. Also, self-esteem and neuroticism (i.e. a disposition to be anxious) are negatively related—that is, people with low self-esteem tend to be more anxious. Exercise provides an excellent medium in which to experience mastery under the right conditions. First, performance can be easily assessed and monitored. At exercise facilities, trained staff can help with this, but it is equally easy to monitor improvement alone, anywhere. Assuming there are no health reasons or contraindications to exercise, after a warm up and stretch, a brisk walk or jog can be timed between two fixed points. Regular exercise will lead to improvements in this time or reductions in any perceived negative physical symptoms such a breathlessness or fatigue. More complex monitoring may involve taking a pulse by hand or with a heart rate monitor after the exercise, keeping the time and pace constant across assessments. Following 6–8 weeks of regular exercise, there is likely to be a reduction in heart rate of at least 5–10 beats per minute (bpm), depending on initial levels of fitness. Second, the improvement curve is likely to be particularly great for those who are least active and least fit. The feedback from self-monitoring provides a rapid sense of achievement after only a few sessions and weeks. As a word of caution, the experience of mastery is dependent on "goal orientation." If goals are set that involve comparison with others, then outcomes are often less positive. There are always people who may be better at a given task, so aiming to be better than others can be self-defeating. In contrast, if goals are centered on personal achievement (with self-referenced targets), then this can lead to a sense of mastery with greater improvements

in affect and motivation to persist. As a further word of caution, particularly for people who have been sedentary, it is better not to seek rapid gains but to build up the exercise dose slowly.

A final explanation for how exercise reduces anxiety involves various physiological adaptations (for reviews, see Morgan, 1997; Biddle & Mutrie, 2001). Regular exercise participation may change some physiological systems that impact on the subjective appraisal of stress and, hence, on musical performance. Some researchers have focused on how the central nervous system adapts to regular exercise. There is evidence that chronic exercise reduces muscle tension (determined from self-report and electromyography). Stress hormones such as catecholamine (adrenaline and noradrenaline) and corticosteroids (e.g. cortisol) also appear to be reduced over time following chronic exercise. In terms of cardiovascular adaptations, resting heart rate and blood pressure are also reduced. It would also be easy to speculate that improved lung function could aid certain types of musicians (i.e. woodwind or brass players) through enhanced pulmonary or lung function, which often results from aerobic exercise training. Interestingly though, there is little or no literature on the specific physiological benefits of chronic exercise on particular performance specialties.

There have been at least two studies that have considered stress reactivity to a musical performance as a stressor (Abel & Larkin, 1990; Valentine *et al.*, 1995). The music stressor did elicit an increase in heart rate and systolic blood pressure, as might be expected. However, the research did not consider the role of cardiovascular fitness or chronic exercise as a moderator of these effects. Anxious people, smokers, less active people, and those with a family history of hypertension tend to be more responsive to mental and physical challenges. In other words, these people show greater increases (from rest) in cardiovascular indices (e.g. heart rate and blood pressure) in anticipation of and during psychosocial stressors. Some research suggests that following chronic exercise, there are small reductions in cardiovascular reactivity (Taylor, 2000). The importance for musicians is that, as a result of chronic exercise, if smaller changes in physiological symptoms of arousal are perceived, this may lead to lower perceptions of activation and enhanced feelings of calm.

One of the proposed ways in which regular exercise impacts on psychological outcomes may be simply from repeated benefits from single sessions of exercise. In other words, should someone stop exercising, the benefits would be quickly lost. There are mixed views about this; however, a considerable research literature exists on the effects of a single session of exercise on emotions, mood, and affect. The conclusions drawn from this literature are considered and the potential implications for musicians are explored in the next section.

## 9.3 Acute exercise

Extensive literature surveys (Petruzzello *et al.*, 1991; Taylor, 2000) have concluded that acute exercise can reduce self-reported state anxiety and physiological indices of stress (e.g. blood pressure). Research suggests that the physiological effects of vigorous exercise may last up to 24 hours—but not beyond 2 days—while the psychological impact seems to be more transitory. Stronger anxiety-reducing effects have been observed

when the exercise is of moderate intensity, as compared with low and vigorous intensities. From a practical perspective, it is important that exercise should not be thrown blindly into a musician's pre-performance routine without rehearsal and refinement. The time between exercise and performance should depend on fitness and personal preference, gained from experience and trial-and-error before music rehearsals rather than performances. Given the range of emotions and moods attached to a musical performance, and indeed to specific musical works, it is impossible to give a blanket exercise prescription for eliciting optimal psychological states. The use of acute exercise requires individual exploration and experimentation.

There is evidence for an impact on basic affect (founded on reasonably well understood psychophysiological responses to exercise), but exactly how acute exercise impacts on specific emotions and mood is less well understood. A variety of mechanisms have been proposed. Distraction and mastery, as identified above, may play a part in gaining benefits from a single session of exercise before performance. Additional explanations include the release of endorphins (endogenous opioids), increase in core body temperature, and rhythmic contraction and relaxation of muscles to reduce tension, but no clear conclusions have been agreed (Morgan, 1997). Recent evidence also points to changes in brain activation associated with positive affect (Petruzzello *et al.*, 2001). It is likely that a number of psychological or physiological mechanisms are simultaneously involved or have greater or lesser impact, depending on variations in exercise intensity and duration.

Taylor (2000) has argued that a single session of exercise can reduce reactivity to a psychosocial stressor. In other words, the stress response (e.g. in terms of blood pressure change) is dampened or attenuated following exercise. It is likely that the effects will be greater for those with parental hypertension and could well be greater when the stressor has been appraised as particularly threatening, rather than just a challenge. One main criticism of this work is that it has typically employed laboratory-based stressors (or mental challenges) rather than naturalistic stressors, thereby raising issues about applicability to musical performance stress.

In order to address the questions raised in Section 9.1 earlier, the authors conducted two studies within the context of a large-scale research initiative, "Zoning In: Motivating the Musical Mind," aimed at optimizing the performance of music conservatory students (see also Chapters 10, 11, and 12 of this volume). In particular, the effects of regular and single sessions of exercise were explored to determine possible benefits for musicians on psychological and physiological indices of stress (Wasley & Taylor, 2002).

## 9.4 Studies with musicians

### 9.4.1 Chronic exercise and music performance

Initially, a musical performance was staged to simulate a natural stressor. The conservatory students were asked to prepare two compositions, each lasting 3–5 minutes. Assessors were in the audience and the performances were videotaped. On a separate day, students also underwent a sub-maximal fitness test and a laboratory-based stressor

(i.e. a color–word Stroop test). Before and during both stressors, a variety of psychological and physiological indicators of stress were taken—heart rate, blood pressure, and state anxiety. Students were then assigned to one of two groups: the first (12 students) did 16 weeks of exercise training involving a weekly 1-hour group session at a local gym and an expectation to do at least 3–5 days of 30 minutes of moderate intensity exercise. The other group (9 students) had similar contact time in Alexander technique lessons. Finally, the performance, laboratory stressor, and fitness test were repeated, under identical conditions, 16 weeks later. The exercise group completed on average 13 contact sessions, and increased the amount of physical activity they did to an average of 180 minutes per week.

The results showed that, as expected, responses to the fitness test improved only for the exercise group, confirming sufficient involvement in the program among most of the students. The ability to utilize oxygen improved significantly, with no change in the Alexander group. Heart rate at rest and during the actual performance was also reduced by about 7 bpm among the exercise group (though not significantly so), with no change for the Alexander group (see Table 9.2). Interestingly, some individual heart rates were as high as 140 bpm during the actual performance, indicating a high level of both physical and emotional engagement in the performance. The exercise (but not the Alexander) group significantly reduced diastolic blood pressure after training, both at rest and before the performance.

Exercise training, therefore, appears to be valuable in reducing the physiological symptoms of occupational stress; however, the fact that there was no reduction in heart rate before the musical performance suggests that the anticipation of a stressor is not influenced by exercise training. It may be that being more fit does little to help cope with a passive stressor such as anticipating a musical performance, but when an active response to a stressor is needed (e.g. as in an actual performance), then the data suggest that being more fit results in a downregulation of the sympathetic nervous system (i.e. attenuation of arousal response).

In terms of general anxiety levels, both the exercise and Alexander groups showed a similar trend of reduced trait anxiety from before to after training; although, without a passive control group, it is not possible to attribute changes to the interventions. Also, the exercise group found the performance situation to be as anxiety eliciting as the Alexander group, but their transient state anxiety levels after performance were marginally lower,

**Table 9.2** Mean heart rate (standard deviation in brackets) under three conditions (i.e. rest, before performance, and during performance), before and after involvement in a 16-week exercise or Alexander technique training program

| Regime | Resting heart rate (bpm) | | Heart rate before performance (bpm) | | Heart rate during performance (bpm) | |
|---|---|---|---|---|---|---|
| | Before | After | Before | After | Before | After |
| Exercise | 72.04 (7.51) | 67.83 (11.36) | 81.23 (12.41) | 84.33 (14.91) | 120.17 (20.74) | 113.25 (20.79) |
| AT | 75.72 (10.44) | 74.11 (7.36) | 82.44 (6.69) | 85.67 (11.68) | 117.33 (20.53) | 123.44 (20.98) |

suggesting that chronic exercise may have expedited recovery after this stressful event (Wasley & Taylor, 2002). Further work is needed to consider how regular exercise may influence both psychological and physiological post-performance recovery.

### 9.4.2 Acute exercise and music performance

In a second study with different students, the simulated musical performance described above was repeated, and a financial incentive (in the form of a prize to one winner) was used to increase the perceived pressure/competitiveness associated with the performance. Thirteen musicians performed on two occasions, and each was asked to select one of their two audiotaped performances to be sent to independent assessors to determine the winner. The students were also given a familiarization session, in which they were introduced to the experimental procedure; this was deemed necessary in order to reduce any concerns the students may have had about the monitoring of physiological indices or about the study in general. Before one performance, students undertook 20 minutes of moderately intense acute exercise, followed by 20 minutes of rest; they were then given a warm-up period on their instruments. Before the other performance, students sat quietly without playing for an equivalent period of time, before also engaging in their normal pre-performance activities. To control for order effects, half the students were assigned to the exercise condition first.

Students then listened to and selected one of their two performances to be put through to the competition. Even though the students had not rehearsed or practiced using exercise within a pre-performance routine, 9 of the 13 students selected their performance after acute exercise as the better of the two. Following the acute exercise, state anxiety was significantly lower than for the nonexercise event. Diastolic (but not systolic) blood pressure after exercise was also slightly lower before the performance, as compared with the resting condition. Mean changes in heart rate from baseline to during the performance was 18 and 23 bpm, following exercise and no exercise, respectively (a statistically significant difference).

### 9.4.3 Subjective responses from students

In addition to the data presented above, the students in both of these studies were interviewed about their subjective responses to the chronic and acute exercise interventions. As a sample, two are provided below verbatim. They provide valuable insight into some benefits that had not been expected, but stimulate further thoughts about the potential implications of physical exercise for enhancing musical performance.

*A female oboist*: I have always been fairly active but hadn't thought about exercising before a performance. Before a performance, I sometimes resorted to drinking alcohol or taking tranquilizers. The exercise program taught me how to build aerobic exercise into a pre-performance routine, and I no longer resort to alcohol or tablets. Now, I walk briskly for 25–30 minutes, ending about 90 minutes before performing. I then go through the rest of my routine, which involves rehearsal, imagery, and relaxation.

*A male pianist*: Previously, I couldn't eat before a performance and would go into it feeling hungry and empty. Now, I always do at least 20 minutes of exercise ending at least 2 hours before a performance, then shower and feel relaxed enough to eat a meal. After that, I go into my normal routine leading up to the performance. I wouldn't want to perform without exercising.

## 9.5 Guidelines for the musician

When the students in the above studies were first interviewed, their self-reported physical activity levels were below UK national recommendations for most students. A few were quite active, but most were keen to participate in the research in order to learn more about the benefits of exercise. No student had systematically incorporated exercise into his or her pre-performance routine (as described above). This section offers recommendations for physical activity and exercise for the musician, both from a general fitness perspective and for pre-performance. It is important to bear in mind that the following guidelines are general and that there is no one-size-fits-all approach. The following guidelines are provided for those who are not at present actively involved in exercise and who have, as is typical among musicians, a busy schedule.

*General guidelines for regular exercise*

- Incorporating more physical activity into your daily living contributes to your overall health, but exercise, like other activities, has a risk attached to it. The level of risk depends on you (the exerciser), as well as the chosen activity and intensity. There are ways to appraise and reduce this risk, and this normally involves exercising within your ability until you are sure how hard you can push yourself. Also, you should ask yourself a number of questions which, if the answer is "yes," you may want to take necessary care. For example, do you suffer from asthma or another form of breathing difficultly? If yes, you should keep your medication with you just in case. Do you have an injury that may be negatively affected by the exercise? If yes, consider other forms of exercise that do not aggravate your condition. If you are in any doubt, visit your physician to discuss your intended program of exercise. This being said, do not let this put you off from planning and starting to become more physically active, as the benefits very often far outweigh the risks.
- The type of exercise in which you decide to engage can be a highly individual choice and may require trial and error. Broadly speaking though, all forms of aerobic activity (e.g. walking briskly, jogging, swimming, and cycling) offer similar health benefits. Thus, regardless of changes in the year and the advantages and challenges to maintaining physical activity that this brings, it is important to know that the same health outcomes can be achieved through several routes. But obviously, the more you enjoy your choice, the more likely you are to maintain your activity.
- You may want to consider exercising with a friend to give the event a social aspect. This has been suggested as one option for enhancing motivation and making exercise more fun. Alternatively, joining a gym can be highly motivating, provide a safe exercise environment, and give structure to your exercise regime; however, fitness centers can be somewhat expensive, inconvenient, and daunting to the new exerciser. Again, it is a matter of your preference as to which approach you take. You can generally achieve the same health and fitness results whether in a gym or not.
- Above all, it is important to keep your chosen exercise intensity in mind. The general program recommended in Table 9.3 starts with one session per week lasting between 20–30 minutes and focuses on moderate intensity activities, as this appears to afford

Table 9.3 General guidelines for engaging in regular (or chronic) exercise for aerobic fitness benefits (assuming a sedentary starting point)

| Weeks | Number of sessions per week | Duration (minutes) | Intensity |
|---|---|---|---|
| 1–4 | 1 | 20–30 | Moderate |
| 4–8 | 2 | 20–40 | All moderate |
| 8–16 | 2 | 20–60 | All moderate |
| 16 onwards | 3 | 20–60 | Two moderate (with one session more vigorous) |

*Note*: For general health benefits, the aim should be to accumulate *at least* 30 minutes of moderate or vigorous activity on five or more days per week. This can be achieved through multiple short bouts of physical activity lasting only 10 minutes.

the best health outcomes. In moderate intensity exercise, you need only feel a little sweaty and mildly breathless, while still able to maintain a conversation. After 3–4 weeks, it is recommended that you increase to two sessions per week at the same intensity.

- This program requires two considerations, however. Firstly, you may find it useful to accumulate the suggested amount of exercise over two or more shorter sessions that add up to your goal. Secondly, the exercise program should progress from a generally easy start with incremental additions of time, frequency, and intensity to avoid injury and unnecessary, unpleasant post-exercise experience. It is relevant to remember that the adage "no pain, no gain" does *not* apply here. Yes, you must push yourself somewhat—but not beyond your limits. It has been observed that many people start exercise programs at unsuitable intensities and frequencies, which can then lead to injury, negative experiences, and ultimately dropping out of exercise. A steady approach will deliver a healthier and more enjoyable experience.
- There are numerous books on exercise, from simple texts on walking to highly technical guidelines set by the American College of Sports Medicine (2001). In addition to general guidelines, as noted above, you may want to explore how best to increase physical activity around musical practicing (such as during breaks) and performance.

*Pre-performance guidelines*

- As a first step, try monitoring how your physical activity patterns change before a performance. It is strongly recommended that the introduction of any moderate intensity exercise before a performance be initially rehearsed and any response closely monitored before attempting to use it within an actual pre-performance routine. This will enable you to work out what is best for you.
- As a general guide, your exercise should take place more than an hour before the performance, as this appears to allow for adequate recovery; however, the intensity at which you work will also impact on the amount of time required to recover.

- Try exercises that are suitable for the environment you are in. For example, go for a brisk, moderate intensity walk for 20 minutes or longer; also, use stairs that may be available in order to get to the desired intensity.
- An exercise session might be incorporated into getting to the performance location, but being late and rushing to a performance is highly discouraged, as this is likely to have a deleterious effect on the necessary calm and focus required for performing. Rather, acute exercise should be planned deliberately with time incorporated for recovery at the venue. Also, remember to rehydrate during or after exercise with sufficient fluid.

## 9.6 Conclusions

In terms of practical experience, research confirms that there are individual differences in how musicians respond to chronic and acute exercise. A process of trial and error to find the right intensity and duration is essential, with an adequate period after exercise for recovery. Clearly, exercise can play a particularly important role if the performer has a predisposition or learned response pattern to be overly anxious. Some musicians may fear that exercising will leave them too tired to practice and perform, but the often-cited notion that exercise leaves one fatigued is largely a mistake. Moderate-intensity exercise is advocated here, which is not likely to have this effect. In fact, there is growing evidence that moderate intensity exercise has revitalizing or energizing capabilities, in addition to reducing anxiety (Ekkekakis *et al.*, 2000). This chapter has focused on aerobic exercise, which has known health and stress management properties. Other exercise involving stretching and resistance training may also be integrated into an exercise program.

## Acknowledgments

We wish to acknowledge the support of The Leverhulme Trust and the Royal College of Music, London, for the work reported in this chapter. Also, we would like to thank the David Lloyd Club (South Kensington, London) for generously providing exercise facilities, as well as the participating students for their enthusiastic support and involvement.

## Further information and reading

Relevant material on the benefits of exercise, suggestions for exercise routines, and information on how to find a personal fitness instructor is available through the following websites: American College of Sports Medicine (*www.acsm.org*), The Exercise Register (*www.exerciseregister.com*), Fit Launch (*www.fitlaunch.com*), and Just Walk (*www.justwalk.com*; track your exercise and diet online). Further particulars of the research results presented in this chapter are provided at the project website for "Zoning In: Motivating the Musical Mind": *www.zoningin.rcm.ac.uk*. For additional reading, see:

American College of Sports Medicine. (2001). *ACSM's Resource Manual for Guidelines for Exercise Testing and Prescription*. London: Lippincott, Williams, and Wilkins.

Biddle, S. J., Fox, K., & Boutcher, S. H. (Eds.). (2000). *Physical Activity and Psychological Well-Being*. London: Routledge.

Franks, B. D., & Howley, E. T. (1998). *Fitness Leader's Handbook.* Champaign, IL: Human Kinetics.

Karageorghis, C. I., & Terry, P. C. (1997). The psychophysical effects of music in sport and exercise: A review. *Journal of Sport Behavior, 20,* 54–68.

McPherson, B. D. (Symposium Chair). (2001). *Medicine and Science in Sports and Exercise, 33* (June Supplement), S345–641.

Pate, R. R., Pratt, M., Blair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C., Bucher, D., Ettinger, W., Heath, G. W., King, A. C., Kinska, A., Leon, A. S., Marcus, B. H., Morris, J., Paffenbarger, R. S., Patrick, K., Pollock, M. L., Rippe, J. M., & Sallis, J. W. (1995). Physical activity and public health: A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *Journal of the American Medical Association, 273,* 402–407.

Weinberg, R. S., & Gould, D. (1999). *Foundations of Sport and Exercise Psychology* (2nd ed.). Champaign, IL: Human Kinetics.

## References

American College of Sports Medicine (2001). *ACSM's Resource Manual for Guidelines for Exercise Testing and Prescription.* London: Lippincott, Williams, and Wilkins.

Abel, J. L., & Larkin, K. T. (1990). Anticipation of performance among musicians: Physiological arousal, confidence, and state anxiety. *Psychology of Music, 18,* 171–182.

Batty, D., & Thune, I. (2000). Does physical activity prevent cancer? Evidence suggests protection against colon cancer and probably breast cancer. *British Medical Journal, 321,* 1424–1425.

Biddle, S. J., & Mutrie, N. (2001). *Psychology of Physical Activity: Determinants, Well-Being, and Interventions.* London: Routledge.

Biddle, S. J., Fox, K., & Boutcher, S. H. (Eds.). (2000). *Physical Activity and Psychological Well-Being.* London: Routledge.

Blair, S. N., Haskell, W. L., Ho, P., Paffenbarger, R. S., Vranizan, K. M., Farquhar, J. W., & Wood, P. D. (1985). Assessment of habitual physical activity by a seven-day recall of physical in a community survey and controlled experiments. *American Journal of Epidemiology, 122,* 794–804.

Bonaiuti, D., Shea, B., Iovine, R., Negrini, S., Robinson, V., Kemper, H. C., Wells, G., Tugwell, P., & Cranney, A. (2003). Exercise for preventing and treating osteoporosis in postmenopausal women (Cochrane Review). *The Cochrane Library, Issue 4.* Oxford: Update Software Ltd.

Boutcher, S. H. (2000). Cognitive performance, fitness, and ageing. In S. J. Biddle, K. Fox, & S. H. Boutcher (Eds.), *Physical Activity and Psychological Well-Being* (pp. 118–129). London: Routledge.

Ekkekakis, P., Hall, E. E., Van Landuyt, L. M., & Petruzzello, S. J. (2000). Walking in (affective) circles: Can short walks enhance affect? *Journal of Behavioral Medicine, 23,* 245–275.

Fox, K. R. (1997). The physical self and processes in self-esteem development. In K. Fox (Ed.), *The Physical Self: From Motivation to Well-Being* (pp. 111–140). Champaign, IL: Human Kinetics.

Fox, K. R. (2000). The effects of exercise on self-perceptions and self-esteem. In S. J. Biddle, K. Fox, & S. H. Boutcher (Eds.), *Physical Activity and Psychological Well-Being* (pp. 88–117). London: Routledge.

Kubitz, K. A., Landers, D. M., Petruzzello, S. J., & Han, M. (1996). The effects of acute and chronic exercise on sleep: A meta-analytic review. *Sports Medicine, 21,* 277–291.

Lawlor, D. A., & Hopker, S. W. (2001). The effectiveness of exercise as an intervention in the management of depression: Systematic review and meta-regression analysis of randomised controlled trials. *British Medical Journal, 322,* 1–8.

Lazarus, R. S., & Folkman, S. (1984). *Stress, Appraisal, and Coping.* New York: Springer.

Morgan, W. P. (1997). *Physical Activity and Mental Health.* Washington, DC: Hemisphere.

Moses, J., Steptoe, A., Mathews, A., & Edwards, S. (1989). The effects of exercise training on mental well-being in the normal population: A controlled trial. *Journal of Psychosomatic Research, 33,* 47–61.

Murphy, M., Nevill, A., Neville, C., Biddle, S. J., & Hardman, A. (2002). Accumulating brisk walking for fitness, cardiovascular risk, and psychological health. *Medicine and Science in Sports and Exercise, 34,* 1468–1474.

Mutrie, N. (2000). The relationship between physical activity and clinically defined depression. In S. J. Biddle, K. Fox, & S. H. Boutcher (Eds.), *Physical Activity and Psychological Well-Being* (pp. 46–652). London: Routledge.

NVC Arts (producer) (1998). *Maxim Vengerov Masterclass* [television broadcast], December 29. London: Channel 4.

Pate, R. R., Pratt, M., Blair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C., Bucher, D., Ettinger, W., Heath, G. W., King, A. C., Kinska, A., Leon, A. S., Marcus, B. H., Morris, J., Paffenbarger, R. S., Patrick, K., Pollock, M. L., Rippe, J. M., & Sallis, J. W. (1995). Physical activity and public health: A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *Journal of the American Medical Association, 273,* 402–407.

Petruzzello, S. J., Hall, E. E., & Ekkekakis, P. (2001). Regional brain activation as a biological marker of affective responsivity to acute exercise: Influence of fitness. *Psychophysiology, 38,* 99–106.

Petruzzello, S. J., Landers, D. M., Hatfield, B. D., Kubitz, K. A., & Salazar, W. (1991). A meta-analysis on the anxiety-reducing effects of acute and chronic exercise: Outcomes and mechanisms. *Sports Medicine, 11,* 143–182.

Sallis, J. F., & Owen, N. (1999). *Physical Activity and Behavioral Medicine.* London: Sage.

Sothmann, M. S., Buckworth, J., Claytor, R. P., Cox, R. H., White-Welkley, J. E., & Dishman, R. K. (1996). Exercise training and the cross-stressor adaptation hypothesis. *Exercise and Sport Sciences Reviews, 24,* 267–287.

Steptoe, A. (1989). Stress, coping and stage fright in professional musicians. *Psychology of Music, 17,* 3–11.

Steptoe, A., & Cox, S. (1988). Acute effects of aerobic exercise on mood. *Health Psychology, 7,* 329–340.

Steptoe, A., & Fidler, H. (1987). Stage fright in musicians: A study of cognitive and behavioural strategies in performance anxiety. *British Journal of Psychology, 78,* 241–249.

Taylor, A. H. (2000). Physical activity, stress and anxiety: A review. In S. J. Biddle & K. Fox & S. H. Boutcher (Eds.), *Physical Activity and Psychological Well-being* (pp. 10–45). London: Routledge.

Thayer, R. E. (1987). Energy, tiredness, and tension effects of a sugar snack versus moderate exercise. *Journal of Personality and Social Psychology, 52,* 119–125.

US DHSS (1996). *Physical Activity and Health: A Report of the Surgeon General.* Pittsburgh, PA: Department of Health and Human Services, Centers for Disease Control and Prevention, and National Center for Chronic Disease Prevention and Health Promotion.

Valentine, E. R., Fitzgerald, D. F. P., Gorton, T. L., Hudson, J. A., & Symonds, E. R. C. (1995). The effect of lessons in the Alexander Technique on music performance in high and low stress situations. *Psychology of Music, 23*, 129–141.

Videman, T., Sarna, S., Battie, M. C., Koskinen, S., Gill, K., Paananen, H., & Gibbons, L. (1995). The long-term effects of physical loading and exercise lifestyles on back-related symptoms, disability, and spinal pathology among men. *Spine, 20*, 699–709.

Vuori, I. M. (2001). Dose-response of physical activity and low back pain, osteoarthritis, and osteoporosis. *Medicine and Science in Sports and Exercise, 33* (June Supplement), S551–S586.

Wasley, D., & Taylor, A. (2002). The effect of physical activity and fitness on psycho-physiological responses to a musical performance and laboratory stressor. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the Seventh International Conference on Music Perception and Cognition* (pp. 93–96). Adelaide, Australia: Causal Productions.

CHAPTER 10

# ALEXANDER TECHNIQUE

ELIZABETH VALENTINE

The Alexander technique (AT) is a method of kinesthetic re-education developed by F. M. Alexander, who was born in Tasmania in 1869, came to live and work in England, and died in London in 1955. His career as an actor became jeopardized by periodic attacks of loss of voice, particularly when under stress. On finding that rest did not cure his problem and that the medical profession was unable to help him, he turned to his own resources, examining his manner of bodily use by observing himself in a mirror. He noticed that whenever he was on the point of reciting he would pull his head back and down, tighten his neck muscles, and suck in air through his mouth. Eventually, he managed to train himself to inhibit this maladaptive strategy of anticipatory tensing.

Alexander (1932) considered that the key to efficient control of bodily movement was the relation of the head to the neck and back, the "primary control" as he called it, a position of mechanical advantage. His view was that in general people develop habitual patterns of muscles, perhaps owing to the failure of instinctive control of the body to keep pace with civilization in the evolution to upright posture. Typically, the head is pulled back and down, as in the startle reaction, when performing simple acts like sitting down. He believed this was particularly insidious because kinesthetic feedback is unreliable and misleading. Alexander's aim was to re-educate people kinesthetically in what he termed "good use," a coordinated and balanced distribution of muscle tension throughout the body, resulting in an attentive state attainable under stress. Habitual, automatic reactions are inhibited and replaced by consciously controlled, directed actions. An important element is maintaining a state of openness with regard to the eventual act to be performed: thus, at the critical moment, a decision may be made to inhibit the intended action, to perform an alternative action, or to continue with the intended one.

Alexander preferred the term "posture" not to be used: Dart's (1947) "poise" gives a much better idea of what is intended. The aim of the technique is to condition the association of a new postural model to verbal instructions such as "direct the head forward and upward" or "lengthen and widen the back," first given overtly by the teacher but subsequently covertly by the student. Movements in accordance with the technique are characterized by economy of effort, a subjective sense of lightness, and a balanced and appropriate distribution of tension (not an absence of tension, but no more than is necessary). Such optimal functioning constitutes good use and deviations from it are termed "misuse." Similarities to the oriental martial arts and the concept of "flow" (Green & Gallwey, 1986; Csikszentmihalyi, 1992; Goleman, 1998) have not gone unnoticed (Gehman, 2002).

In this chapter, the principles of the technique are described and its application to music performance discussed. Although often regarded as on the fringe of science, an increasing number of studies have provided scientific evidence for both its anatomical and physiological underpinnings, and its beneficial effects on behavior and experience. Anatomical and physiological studies of AT in the general population are reviewed, followed by a consideration of studies on its effects in musicians, firstly physiological, and then behavioral and experiential. Although the evidence is limited and not always consistent, a number of studies have shown beneficial effects of training in AT on breathing, heart rate and blood pressure, posture and bodily use, quality of musical performance, and mental attitude. Finally, a list of resources is supplied where further information can be obtained both on the research that has been conducted and on how to find out more about the technique and to locate a teacher. Lessons are typically individually based, and training in AT is commonly available in music conservatories and university music departments.

## 10.1 Application of AT to music performance

Although considered something that can be of use to everyone, AT has been applied clinically in areas of pain management (Fisher, 1988; Bajwa *et al.*, 2001), Parkinson's disease (Stallibrass, 1997), and for the mentally and physically disabled (Maitland *et al.*, 1996). It has been popular in equitation and generally in the performing arts, especially ballet (e.g. Richmond, 1994; Bluethenthal, 1996). Given the fine neuromuscular coordination, and physical and mental strains and stresses to which performing musicians are subject, it is not surprising that AT has been popular among them. The technique has been endorsed by such eminent and diverse musicians as Adrian Boult, Julian Bream, Colin Davis, James Galway, Paul McCartney, Yehudi Menuhin, Sting, and Barry Tuckwell. It is taught at most, if not all, leading international music conservatories and has been applied to *voice* (Lewis, 1980; Duarte, 1981; Lloyd, 1986; Englehart, 1989; Rohmert *et al.*, 1990; Hudson, 2002); *wind instruments* (Bosch, 1997; Holm, 1997; Bosch & Hinch, 1999); *strings* (Tursi, 1959; Richter, 1974; Hamilton, 1986; Conable, 1993; McCullogh, 1996; Stein, 1999); and *keyboard* (Ben-Or, 1987; Kaplan, 1994).

Although group lessons are occasionally given, lessons are typically and preferentially one-to-one and hands-on. The teacher will guide the student's movements by light touch and verbal instructions. Attention is usually focused initially on familiar, daily actions—in which the student is taught to connect thought to movement—and is later extended and applied to specific areas of skilled performance, such as playing a musical instrument. The student may lie on the floor (or for teaching purposes, on a long wide table) in semi-supine position (i.e. on the back with the knees bent up pointing to the ceiling), thus allowing for a natural lengthening of the spine. Another common feature of a lesson is the movement from sitting to standing and vice versa. Such common everyday activities enable the teacher to demonstrate the importance of keeping the head and neck free, and the back lengthening at all times. Other activities include walking, bending, reaching, lifting, and breathing, which is of course crucially important to music performance. In each case, the teacher will help students become more aware of themselves, where they are, and how they move. She or he might draw

their attention to the way they put on their shoes or jacket, pick up a bag, or open the door at the end of a lesson. Particularly in the case of a musician, the teacher will help the student become aware of maladaptive, habitual reactions that are often associated with underlying patterns of tension and anxiety (e.g. singers skewing their head to one side on high notes or string players arching their back) and encourage ease and freedom of movement. Of great importance to the musician is the proper support of the arms, hands, and fingers from the lower back. The student may also be given homework, which can include keeping a diary of "use" and reading relevant texts.

There is evidence that musicians both need and use therapeutic techniques to aid them in their work. Several studies have indicated that levels of anxiety are higher among musicians than in the general population. Watson and Valentine (1987) administered a number of personality inventories and a questionnaire about the use of complementary medical techniques to 60 members of the London Sinfonietta. Both sexes scored significantly above the norms for general or *trait* anxiety (see Chapter 1)—for instance, the sample mean on Spielberger's Trait Anxiety Inventory (Spielberger *et al.*, 1983) was 41.95 for men, compared with the norm of 34.89 for the general population, and 44.25 for women, compared with the norm of 34.79. This is consistent with the work of Kemp (1981), who also found that professional performing musicians scored significantly above population norms for anxiety, that music students were more anxious than nonmusic students, and that "talented" musicians were more anxious than "nontalented." In Watson and Valentine's sample, 53% reported using anxiety-reduction techniques of some kind. AT was easily the most frequently used technique, with 43% currently taking or having taken lessons. This suggests that AT is used by at least 20% of the general population of musicians (this allows for possible sampling bias; only 50% of the sample returned usable data).

## 10.2 AT and science

The work of eminent scientists such as George Coghill, Rudolph Magnus, and Charles Sherrington was invoked in support of Alexander's ideas, though in some cases now considered an unjustified extrapolation. Many famous people were numbered among Alexander's admirers, including John Dewey (McCormack, 1958), Aldous Huxley, George Bernard Shaw, and Nikolaas Tinbergen, who devoted much of his Nobel prizewinner's lecture to a discussion of Alexander's work. The theory, however, has remained on the fringe of science and medicine. Among psychologists, Alexander's admirers have included Harold Schlosberg and Tony Buzan.

Alexander developed his method of posture and movement re-education about the same time (i.e. soon after the turn of the twentieth century) as Sherrington's theories on the nature of proprioception were becoming known. His methods of releasing tension patterns in the neck coincided with the discovery of massive numbers of muscle spindles in the cervical musculature (Batson, 1996, in reviewing recent work relevant to the peripheral neuroanatomic basis of AT, remarks on the prescience of Alexander's conclusions, in the light of current knowledge in neuroscience). Alexander does not mention this work, though he may have known about it. Like Freud, he seems not to have felt the need to validate his ideas scientifically, perhaps considering that he had

created a practical, experiential means of body learning, accessible through a method of "hands-on" instruction. Hence, while belief in the technique remains strong and practitioners are fulsome in its praise and testify to its benefits in their personal statements, rigorous objective data have been hard to come by (this is all the more surprising in view of the large number of musicians who use the technique and claim to benefit from it). This may be partly because of the ineffable nature of its effects, inherently difficult to describe verbally or quantify, and partly because it is subject to effects of expectation and suggestibility, or both of these.

In 1937, 19 physicians wrote to the *British Medical Journal* to urge that "as soon as possible, steps be taken for an investigation of Alexander's work" (Bouce-Porter *et al.*, 1937, p. 1137). Dewey undertook to get support for a scientific investigation from the Rockefeller Foundation, but Alexander set up too many requirements and it fell through. In 1949, a grateful student gave $500 (US) for research which was used for a pilot electromyographic study at Tufts University (see Section 10.3 below). Then Dewey died, and not long afterwards Alexander himself. As late as 1978, Wilfred Barlow wrote: "I appreciate only too well that the statistical background which doctors and scientists will eventually need is not very much in evidence" (p. 11).

There is a substantial literature on AT (for bibliographies, see Priest, 1992; Evans, 2000), not only in English, but also in French, German, Spanish, and Italian. Articles are scattered among professional journals, scientific periodicals, theses, and dissertations. Nevertheless, with very few notable exceptions, it is still the case that the majority of studies are anecdotal, lacking in adequate experimental controls, on small samples, or simply case studies. While considerable progress has been made on pharmacological and cognitive-behavioral therapeutic interventions for the treatment of performance anxiety (see Chapters 12 and 14 of this volume), there has been very little systematic investigation of the efficacy of AT on behavioral and experiential manifestations.

## 10.3 Anatomical and physiological studies on the general population

Most of the attempts to provide objective evidence of the effect of AT have focused on anatomical and/or physiological measurements. Some of the earliest research on AT was carried out by Wilfred Barlow on army cadets (see Barlow, 1998, for a review). He suspended a tape measure from the back of their heads, enabling measurement of the distance to the bottom of their necks. The vast majority showed a shortening of this distance (head retraction) by an inch or more as they sat down or stood up, which they were unable to prevent and of which they were unaware (see Figure 10.1). This confirms Alexander's view that kinesthetic sensations are unreliable.

Frank Pierce Jones (Jones *et al.*, 1963; see Jones, 1976, for a review), working at Tufts University in the 1950s and 1960s, measured muscle activity and movement patterns for guided and unguided movements using electromyography (i.e. electrical recording of muscle activity), X-ray and multiple-image photography, and a strain-gauge force platform (which measures directional forces). The X-ray photographs showed that employing the Alexander principle resulted in increased length of the sternomastoid muscles on the side of the neck, increased width of the discs between the neck vertebrae, and a forward movement of the center of gravity of the head. The multiple-image photographs showed that the movements were quicker and more direct. This work has

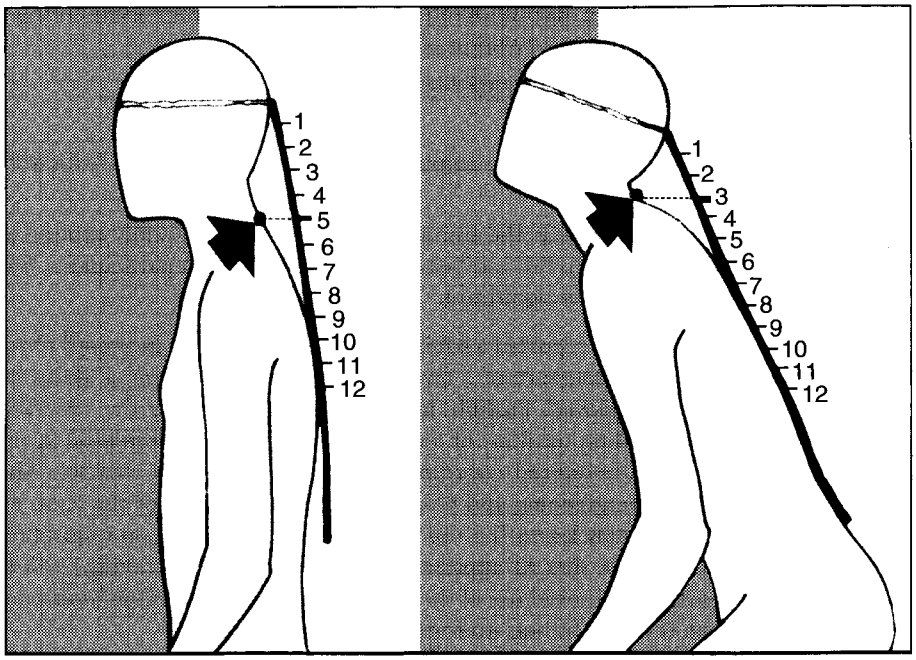



**Figure 10.1** Pulling back the head when moving from a standing position (left) to sitting down (right). (From Barlow, 1998. © Orion Publishing group, used by permission.)

been criticized by Stevens (1987) on the grounds of artifacts due to touching the participant in the guided movements, slow sampling rates of the photography which induce various kinds of error, and outdated anatomical and physiological knowledge.

Stevens *et al.* (1989) repeated Jones's work, using more sophisticated electronic apparatus. Their results confirmed those of Jones in showing that guided movements, in which habitual postural adjustments of the head and neck are inhibited, require less force, show less muscle activity, and are quicker and smoother than habitual movements. In their study, guided sit-to-stand movements showed a decrease in head movements, ground reaction forces, and electromyographic activity in the trapezius, sternomastoid, and erector spinae muscles (located in the shoulders, neck, and back respectively).

Austin and his colleagues have investigated the effects of AT instruction on respiratory function in healthy adult volunteers. In an exploratory study, Austin and Pullin (1984) studied six normal volunteers using standard spirometric (i.e. lung capacity) and mouth pressure tests before and after a single lesson in the technique, but no differences were found. They then studied eight normal adult volunteers before and after a course of 20 weekly private lessons. The results showed significant increases in the following measures of respiratory function:

- forced vital capacity (i.e. the volume of air expelled by a forced maximal expiration from a position of full inspiration)

- peak expiratory flow (i.e. the maximum rate of airflow that can be achieved during a sudden forced expiration from a position of full inspiration)
- forced expiratory flow at 25% of forced vital capacity
- maximum voluntary ventilation
- maximal inspiratory pressure (i.e. the amount of air that can be inspired from normal expiration)
- maximal expiratory pressure (i.e. the amount of air that can be forcibly expired at the end of normal expiration); this increased for seven of the eight participants, but the change was not statistically significant.

Six of the eight participants reported a subjective sense of distinctly increased physical well-being and that performance of a favorite physical activity, such as skiing or playing the piano, had improved. Possible mechanisms suggested for the observed increases are decreased resting tensions of the respiratory muscles, decreased intra-abdominal pressures, and enhanced coordination of the respiratory muscles. The authors speculate that as lengthening and poise develop, compression of the torso by habitual muscular actions may lessen (i.e. the "normal" corseting effect on respiration may decrease). They conclude that AT appears to offer promise for improving respiratory function, by teaching improved use of the musculoskeletal system. One limitation of the study, however, is that there was no comparison control group.

In a study that did compare an experimental and a control group, Austin and Ausubel (1992) obtained further evidence that AT musculoskeletal education may enhance respiratory muscular function in normal adults. Following a course of 20 private AT lessons at weekly intervals, the experimental group (consisting of 10 participants) showed significant improvements on a number of spirometric measures:

- peak expiratory flow
- maximal voluntary ventilation
- maximal inspiratory pressure
- maximal expiratory pressure.

The control group (also of 10 participants), matched for age, gender, height, and weight, given no lessons, did not show any significant changes. Again, suggested mechanisms for the observed effects are increased length and decreased resting tension of muscles in the torso (which, in turn, may increase their strength), increased thoracic compliance (i.e. a measure of lung and chest recoil, the inverse of stiffness), and/or enhanced coordination.

## 10.4 Studies with musicians

### 10.4.1 Physiological studies

Doyle (1984) applied physiological measurements to musicians, undertaking photographic and electromyographic analysis of the head–neck relationship during violinists' initiation of the act of playing. He claimed that the demonstrated association

between neck tonus and the attitudinal response in the player supports Alexander and discusses implications for the pedagogy of violin playing.

Nielsen (1988) assigned members of the Aarhus Symphony Orchestra to an exercise group (who ran 7 kilometers 3 times a week for 8 weeks), an Alexander group (who received 20 private lessons in the technique during the same period), and a beta-blocker group (who took 40 mg of propranolol 1.5 hours before the test concert). Measures of heart rate and blood pressure were taken just before the first rehearsal, the final rehearsal, and the concert on two separate occasions (i.e. before and after the intervention). Exercise led to a significant reduction in heart rate. Lessons in AT and beta-blockers led to significant reductions in systolic blood pressure and significantly reduced increases in systolic blood pressure from the final rehearsal to the concert.

It appears, therefore, that AT can be almost as effective as beta-blockers in preventing inappropriate responses to stress (the effect was somewhat stronger in the latter than in the former). Nielsen suggests that this may be by resolving the demands of the voluntary and reflex components of behavior in a more conscious way. The musicians reported that participation in the study had resulted in benefits for the orchestra as a whole, predominantly social but also musical. Of 13 members of the exercise group, 12 members of the Alexander group, and 7 members of the beta-blocker group, 10, 8, and 3 respectively reported that the intervention had led to lowered unwanted stress; 11, 10, and 2 respectively that they would like to continue with the training; and 11, 8, and 2 respectively that they would recommend their treatment to others. Exercise was reported to increase general well-being, and Alexander training to result in improved breathing and surplus energy with which to solve daily problems. Beta-blockers, however, were reported to produce indifference to the job, deterioration in concentration, feelings of being isolated, and hammering in the fingers and toes. As one participant said: "If you become totally de-stressed and cold after taking a beta-blocker, the result is that what you play becomes cold and uninteresting" (see Chapter 14 for further discussion).

### 10.4.2 Behavioral studies

Early work on the behavioral effects of AT was undertaken by Barlow (1956), who compared 44 students at the Central School of Speech (CSS; where posture was considered important), who were taught by conventional methods (verbal instructions, manual adjustment, and exercises), with 20 students from the Royal College of Music, who were sent to him for lessons in AT. The students were photographed by an independent investigator but scored by Barlow for postural faults, including retracted head (see Figure 10.2), dropped shoulders, rotated pelvis, lordosis, and internally rotated knees. He sanguinely remarks, "A given observer soon achieves a uniform pattern of scoring. My average scoring for the first and second year group only varied by 1%, which indicates that the method is accurate" (p. 671). The students were categorized as follows:

- excellent (0–3 faults)
- slight defects (4–5 faults)
- moderately severe defects (6–9 faults)

- severe defects (10–14 faults)
- very severe defects (over 16 faults).

Barlow claimed that the RCM group showed considerable improvement and the CSS group deterioration, but no statistical analysis was conducted. Applying Kendall's tau to the results shows that there is indeed statistically significant evidence of improvement in the RCM group but that the deterioration in the CSS group is not statistically significant. However, the difference between the two groups is highly significant. This is consistent with Barlow's conclusion that "postural re-education will be ineffective unless a new 'body schema' is taught, by associating the postural model with an improved postural awareness," but more evidence would be needed to demonstrate that this is indeed the crucial element in Alexander training, not to mention the need for the assessments to be made blind by an independent observer.

In addition, Barlow (1955; see also 1978) reported a study of 50 opera students from a London voice and drama college, who had on average 37 lessons in AT. However, the

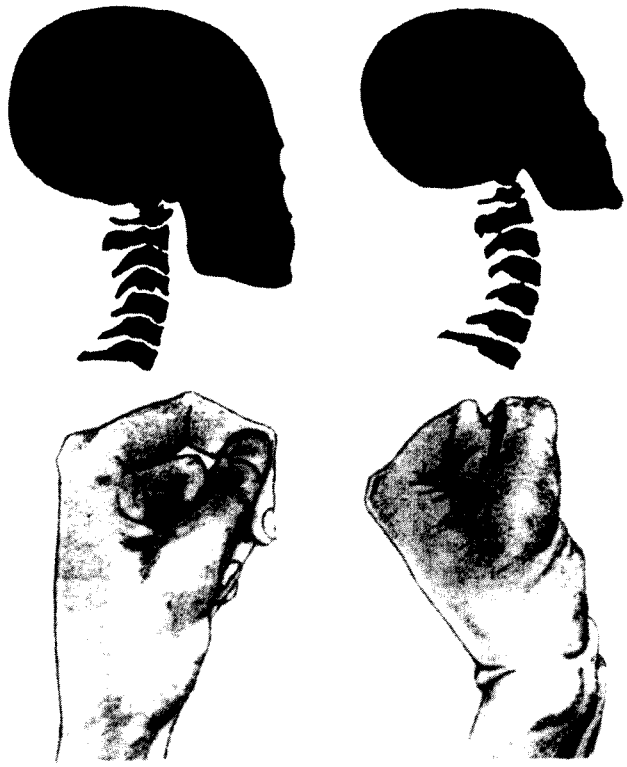


**Figure 10.2** Skull and neck vertebrae, and "parallel" picture of hand and wrist. Cervical lordosis corrected by movement similar to wrist movement. (From Barlow, 1998. © Orion Publishing group, used by permission.)

study was unsystematic (he admits that "a statistician might like to see it all planned out a little more formally", p. 298), and the results are incompletely reported. Twenty of the students were subjected to Sheldon somatotype photography (while standing nude on a turntable!). Pictures for a selected five of these before and after re-education appear in the published report. Anthropocentric measurements were taken, and it is reported that "all the subjects with one or two exceptions show small but definite increases in standing and sitting height and shoulder width" (p. 300). A report was provided by two professors at the college, which included the following generalizations:

- In each case, there was marked physical improvement, which was usually reflected vocally and dramatically.
- In all cases, students since re-education were easier to teach.
- Each student reacted in a different characteristic way. The time it took to get results varied greatly between one student and another. The utilization of the approach depended largely on the student.
- Of eight students who entered a national singing competition, in which there were over 100 competitors, six reached the semifinal, which is quite out of proportion to what would be expected.

They concluded that "In our opinion, this approach is the best means we have yet encountered for solving the artist's problem of communication and should form the basis of training" (p. 300). However, there are a number of obvious criticisms of this study. Participants were not randomly assigned to the treatment, the sample included problem cases as well as "normals," there was no real control group, assessments were not free from observer bias, and the data are incompletely reported.

Jones (1972) made recordings of a singer when in the habitual mode and when the balance of the head was altered by the experimenter in accordance with the Alexander principle (however, note Stevens's criticism mentioned above). The singer reported that she sang more easily and with greater resonance and better breath control under the experimental conditions. The difference in the quality of singing was confirmed by other musicians who listened to the tape. A spectrogram confirmed increased richness of overtones and the virtual disappearance of breathing sounds during the experimental session.

Dennis (1987) randomly assigned 13 wind players to 20 lessons in AT or to a control group. Music performance was video-recorded and various measures of respiratory function, using standard spirometry and maximal static mouth pressures, were taken before and after training. For music performance, measures were taken of posture and movement during nonplaying, posture and movement during playing, breath control, and overall performance quality. The only significant difference between groups favored the control group, who showed a significantly greater improvement in maximal voluntary ventilation following training. No significant differences were obtained between the groups on the five other measures of respiratory function (forced vital capacity; forced expiratory volume at one second; peak expiratory flow; expiratory flow rates at 25%, 50%, and 75% of forced vital capacity; maximal inspiratory and expiratory pressures) or on any of the measures of musical performance (overall

performance, posture and movement during playing and nonplaying, and breath control). It is important to note that, as wind players, these musicians may already have been performing at ceiling on respiratory function. As is often the case, subjective reports by the musicians suggested a range of beneficial effects of the lessons.

Using a design that avoided some of the methodological weaknesses of earlier studies, Valentine *et al.* (1995) aimed to investigate the experiential and behavioral effects of lessons in AT on music performance. Specifically, participants were randomly allocated to treatment and control conditions, and judges were blind to participants' condition assignment (i.e. they did not know which students had received lessons). Music students taking classes in performance were randomly assigned to an experimental group, which was offered a course of 15 lessons in AT, and to a control group. They included singers, violinists, cellists, pianists, organists, flautists, clarinetists, an oboist, and a trombonist. Physiological, behavioral, and self-report measures were taken on four occasions: in both a high-stress situation (an audition or a recital, subject to evaluation with serious consequences) and a low-stress situation (class performances, still subject to appraisal by the class teacher and peers), before and after treatment.

Measures of height and peak expiratory flow were included to test frequent claims that training in AT increases the realization of potential height and improves breath control. Heart rate mean and variance were chosen as measures of arousal under stress, as well as self-report questionnaires on mental state (i.e. a mood checklist and the Musical Performance Anxiety Self-Statement Scale, which measures a positive attitude to performance and task-focused attention; Craske *et al.*, 1988). Performances were video-recorded and subsequently rated by expert judges blind to participants' condition assignment, for degree of misuse by two Alexander teachers (other than those who gave the lessons) who specialize in teaching musicians, and for music performance by two music faculty members who specialize in teaching performance. Judgments of misuse (i.e. deviations from optimal performance in accordance with the principles of the Alexander technique) were made on a 7-point rating scale, drawn up and agreed by the two judges, where 1–2 indicated low levels, a generally good (above average) standard of use and coordination; 3–5 indicated more significant misuse likely to impair performance; and 5–7 indicated serious misuse. Judgments were made purely on the basis of visual information, with the sound turned down.

Positive effects of AT were obtained for physiological, performance, and self-report measures. The experimental group showed improvement relative to the control on the following measures:

- overall musical and technical quality
- heart rate variance
- self-rated anxiety
- positive attitude to performance.

Thus, quality of musical performance and positive attitude to performance increased, and heart rate variance and self-rated anxiety declined, as a result of training in AT. However, with the exception of heart rate variance, these effects were restricted to performance in the low stress class situations (it is worth noting, however, that the degree

of stress is relative: mean heart rate was 102.19 beats per minute in the high-stress situations compared with 90.83 in the low-stress situations). There were no significant changes in height or peak flow as a result of the intervention.

There was good agreement between the two judges on overall musical quality and on technical quality (this is of interest in view of the skepticism concerning the validity of such "subjective" judgments). On the other hand, inter-rater agreement for judgments of misuse by the Alexander experts was very low. Qualitative comments made by one of the judges indicated that four participants in the experimental group, but none in the control group, were considered to have improved, whereas one participant in the experimental group and two in the control group were considered not to have shown improvement. There were no significant effects of training in AT on misuse as judged on the basis of video-recordings of behavior. This may indicate that it is difficult to assess misuse on the basis of purely visual information (which in addition may be limited in perspective).

The musicians' average ratings of general benefit derived from the course of lessons was 3.86 and for musical benefit was 4.77 (both on a scale from 1, not at all, to 7, very beneficial). Responses ranged from very little general effect to "marvelous, fantastic, terrific." All participants mentioned increased awareness of tension and improvement in the ability to relax. Some participants mentioned help with physical effects such as breathing or posture—for example, "helps poise," "induces a feeling of lifting up, being elongated" (two participants said they felt taller after lessons), "increases awareness of the relation of particular movements to overall body use," "eliminates bad habits," "avoids a bull-in-a-china-shop approach," "muscles don't become tense and bones are where they should be." Others emphasized mental attitudes towards playing (e.g. "teaches you how to go about setting yourself up to play," "decreases nervousness and uptightness in relation to performance," "releases tension, helps you relax"). Eight out of twelve participants said they would have more lessons if they could. Seven said they would definitely recommend lessons to someone else, three that they probably would, and two that they might possibly, depending on whether the person had tension problems and the abilities of the teacher. No one said they would definitely not recommend the technique.

In sum, these results suggest that AT may have beneficial effects on the quality of performance, the mental state of the performer, and may help to modulate increased variability of heart rate under stress. However, attention should be drawn to two cautionary qualifications. First, the sample size was small and the participants were predominantly female. Second, with the exception of heart rate variance, the effects of the experimental treatment were restricted to the low stress class situations and did not extend to the high stress recital situation. It is possible that 15 lessons were insufficient to develop the level of skill required to apply the technique in a state of high arousal, or that an enhanced ability to deal with performance anxiety is not the main benefit to be derived from the technique. Far more dramatic effects have been obtained with cognitive-behavioral methods (see Kendrick *et al.*, 1982; Sweeney & Horan, 1982; Clark & Agras, 1991; Chapter 12 of this volume). It is possible that demand characteristics may have contributed to the results, since it was considered necessary to inform the participants of the purpose of the study in order to enlist their cooperation, given the

small size of the available sample. While demand characteristics may have affected the self-report measures, this is less plausible as an explanation for the more objective, physiological measures and independent behavioral ratings.

In a recent study (Valentine & Williamon, 2003), conservatory students were randomly assigned to various training groups: three groups received different kinds of neurofeedback and one group received AT lessons (students were also assigned to other interventions as part of the project "Zoning In: Motivating the Musical Mind"; see also Chapters 9, 11, and 12). The musicians in the neurofeedback training groups received 10 15-minute sessions over a period of 6–8 weeks. Those in the AT group received 12 30-minute sessions of one-to-one training on a weekly basis. Music performances before and after training were video-recorded. These were then randomly ordered and assessed by expert judges, external to the conservatory and blind to students' group membership, on perceived instrumental competence, musical understanding, and communication (each of which had three subscales) in addition to overall quality. Only one of the groups receiving neurofeedback (that receiving "alpha/theta" neurofeedback) showed significant changes on these rating scales as a result of training (for further discussion of the neurofeedback results, see Gruzelier *et al.*, 2002, and Chapter 11 of this volume). Following training, all four groups showed significant reductions in self-rated *state* anxiety (see Spielberger *et al.*, 1983, and Chapter 1) before performance. Video-recordings for the alpha/theta neurofeedback and the AT groups were rated for quality of bodily use by an experienced musician and AT practitioner, with extensive experience of teaching AT to musicians. She devised her own rating scales as follows (ratings were made from 1, very poor, to 7, excellent, purely on the basis of visual information with the sound turned down):

- head–neck–back relationship
- upper limb/back
- hips/balance
- direction of knees
- face and eyes
- breathing
- fingers
- thought direction
- inhibition
- overall impression/poise.

The AT group showed a significant improvement relative to the neurofeedback group on seven of the ten scales of "use": head–neck–back relationship, upper limb/back, face and eyes, fingers, thought direction, inhibition, and overall impression/poise (change scores, i.e. the difference between scores before and after training, were all significantly correlated with the exception of those among head–neck–back relationship, face and eyes, and breathing, suggesting that these are relatively independent of the rest and of each other). This is the first demonstration of the validity of

judgments of AT use based on video-recorded performance and suggests not only that training in AT can enhance use but also that this can be demonstrated objectively. There were also some interesting differences among instruments on several of the scales: hips/balance, fingers, and thought direction. The greatest improvements were shown by singers.

Student feedback indicated a high degree of satisfaction and perceived benefit from the training: "I'd love it if the lessons were longer—half an hour goes very quickly"; "I really enjoyed and found useful the practical sessions—when I played piano or oboe and [the teacher] worked specifically with my needs"; "I found everything really interesting and incredibly helpful"; "I found the work we did with my instrument during lessons particularly valuable"; "Slowly making the Alexander technique a part of my life is a really fulfilling experience. I think it will always be part of my life from now on."

## 10.5 Conclusions

Although results have not always been consistent, there is now sufficient scientific research confirming that training in AT can:

- improve respiratory function
- minimize increases in heart rate and blood pressure under stress
- improve bodily use
- enhance the quality of musical performance.

Practitioners report improved mental attitudes, and increased awareness of ability to minimize tension. Results are best when general training in the technique can be integrated with work specifically directed at music performance in situ. For details of useful contacts and how to obtain further information, see "Further information and reading" below.

## Acknowledgments

I should like to thank Judith Kleinman and Elisabeth Waterhouse for their help, and an anonymous reviewer for useful comments on a draft of this chapter. The research reported at the end of this chapter was supported by the Research Strategy Fund of Royal Holloway, University of London, and by the Royal College of Music, London.

## Further information and reading

Additional information (including how to locate a teacher) can be obtained through the following professional associations: The Society of Teachers of the Alexander Technique (see *www.stat.org.uk* or write to STAT, First floor, Linton House, 39–51 Highgate Road, London NW5 1RS, UK), The American Society for the Alexander Technique (see *www.alexandertech.org* or write to AmSAT, 30 North Maple, P.O. Box 60008, Florence, MA 01062, USA), and Alexander Technique International (see *www.ati-net.com*). Further details of the project "Zoning In:

Motivating the Musical Mind" are available at *www.zoningin.rcm.ac.uk.* For additional reading on research conducted on the Alexander technique and what to expect in a lesson, see:

## Books and journal articles

Alcantara, P. D. (1997). *Indirect Procedures: A Musician's Guide to the Alexander Technique.* Oxford: Clarendon Press.

Brennan, R. (1999). *The Alexander Technique Manual: A Step-by-Step Guide to Improve Breathing, Posture and Well-Being.* London: Little, Brown and Company.

Conable, B. (2000). *What Every Musician Needs to Know About the Body: The Practical Application of Body Mapping & the Alexander Technique to Making Music.* Portland, OR: Andover.

Gelb, M. (1981). *Body Learning.* London: Aurum.

Mackie, V., & Armstrong, J. (2002). *Just Play Naturally.* Aarhus, Denmark: Novis/Duende.

Rosenthal, E. (1987). The Alexander technique. What it is and how it works: Work with three musicians. *Medical Problems of Performing Artists, 2,* 53–57.

## Articles in music periodicals

Anon (1989). The Alexander technique: From pupil to teacher. *Music Teacher, 68*(4), 19–20.

Babits, L., & Mayers, H. (1988). The path to productive practicing: An introduction to the Alexander technique. *American Music Teacher, 38,* 24–26 and 60.

Benham, B., Melborn, M., & Sanders, J. (1996). Three musicians talk about the Alexander Technique. *Journal of the International Double Reed Society, 24,* 121–127.

Bondurant, N. (1992). Movement, music, and the Alexander technique. *The Double Reed, 15*(1), 59–61.

Farkas, A. (1995). Coach's notebook . . . Exploring the Alexander technique and opera. *Opera Journal, 28*(3), 16–32.

Naylor, J. (1992). Getting better: Applying Alexander technique to instrumental study. *Music Teacher, 71*(7), 30–31.

Sanders, J. (2002). Freedom to breathe. *Strad, 113,* 631–635.

Stein, J. (1996). To schlep or not to schlep. *Music Teacher, 75*(5), 8–13.

## References

Alexander, F. M. (1932). *The Use of Self.* New York: Dutton.

Austin, J. H. M., & Ausubel, P. (1992). Enhanced respiratory muscular function in normal adults after lessons in proprioceptive musculoskeletal education without exercises. *Chest, 102,* 486–490.

Austin, J. H. M., & Pullin, G. S. (1984). Improved respiratory function after lessons in the Alexander technique of musculoskeletal education. *American Review of Respiratory Disease, 129,* A275.

Bajwa, Z. H., Gupta, W., Warfield, C. A., & Steinman, T. I. (2001). Pain management in polycystic kidney disease. *Kidney International, 60,* 1631–1644.

Barlow, W. (1955). Anxiety and muscle tension. In D. F. O'Neill (Ed.), *Modern Trends in Psychosomatic Medicine* (pp. 285–309). London: Butterworth.

Barlow, W. (1956). Postural deformity. *Proceedings of the Royal Society of Medicine, 49,* 670–674.

Barlow, W. (1978). Research at the Royal College of Music. In W. Barlow (Ed.), *More Talk of Alexander* (pp. 190–195). London: Victor Gollancz.

Barlow, W. (1998). *The Alexander Principle: How to Use Your Body Without Stress.* London: Victor Gollancz.

Batson, G. (1996). Conscious use of the human body in movement: The peripheral neuroanatomic basis of the Alexander technique. *Medical Problems of Performing Artists, 11,* 3–11.

Ben-Or, N. (1987). The Alexander technique and performance. In C. Grindea (Ed.), *Tensions in the Performance of Music* (revised ed., pp. 84–95). London: Kahn and Averill.

Bluethenthal, A. (1996). Before you leap: The Alexander technique, an alternative approach to dance. *Dance Theatre Journal, 13,* 44–46.

Bosch, A. (1997). *The Use of the Alexander Technique in the Improvement of Flute Tone.* Unpublished masters dissertation, University of Pretoria, Pretoria, South Africa.

Bosch, A., & Hinch, J. (1999). The application of the Alexander technique for flute teaching: Two case studies. *British Journal of Music Education, 16,* 245–251.

Clark, D. B., & Agras, W. S. (1991). The assessment and treatment of performance anxiety in musicians. *American Journal of Psychiatry, 148,* 598–605.

Conable, W. (1993). The Alexander technique. *Journal of the Violin Society of America, 13,* 125–135.

Craske, M. G., Craig, K. D., & Kendrick, M. J. (1988). The musical performance anxiety self-statement scale. In M. Hersen & A. S. Bellack (Eds.), *Dictionary of Behavioral Assessment Techniques.* Oxford: Pergamon.

Csikszentmihalyi, M. (1992). *Flow: The Psychology of Happiness.* London: Rider.

Dart, R. (1947). The attainment of poise. *South African Journal, 21,* 764–791.

Dennis, R. J. (1987). *Music Performance and Respiratory Function in Wind Instrumentalists: Effects of the Alexander Technique on Musculoskeletal Education.* Unpublished doctoral dissertation, Columbia University, New York.

Doyle, G. (1984). *The Task of the Violinist: Skill, Stress and the Alexander Technique.* Unpublished doctoral dissertation, Lancaster University, Lancaster, UK.

Duarte, F. (1981). The principles of the Alexander technique applied to singing: The significance of the "preparatory set." *Journal of Research in Singing, 5,* 3–21.

Englehart, R. J. (1989). *An Electromyographic Study of Preparatory Set in Singing as Influenced by the Alexander Technique.* Unpublished doctoral dissertation, Ohio State University, Columbus, OH.

Evans, N. (2000). Alexander technique: An introduction and bibliography of recommended resources for the musician. *Brio: Journal of the United Kingdom Branch of the International Association of Music Libraries, Archives and Documentation Centres, 37,* 9–12.

Fisher, K. (1988). Early experiences of a multidisciplinary pain management programme. *Holistic Medicine, 3,* 47–56.

Gehman, S. (2002). *The Alexander Technique and Tai Chi Chuan?* Retrieved June 24, 2003, from *www.drizzle.com/~stacyg/ATforTaiChi.html.*

Goleman, D. (1998). *Working with Emotional Intelligence.* London: Bloomsbury.

Green, B., & Gallwey, W. T. (1986). *The Inner Game of Music.* New York: Doubleday.

Gruzelier, J. H., Egner, T., Valentine, E., & Williamon, A. (2002). Comparing learned EEG self-regulation and the Alexander Technique as a means of enhancing musical performance. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the*

*Seventh International Conference on Music Perception and Cognition* (pp. 89–92). Adelaide, Australia: Causal Productions.

Hamilton, B. (1986). *The Alexander Technique: A Practical Application to Upper String Playing.* Unpublished doctoral dissertation, Yale University, New Haven, CT.

Holm, C. P. (1997). *Corrective to Breathing Hindrances in Flute Performance, With Emphasis on the Alexander Technique.* Unpublished doctoral dissertation, Southern Baptist Theological Seminary, Louisville, KY.

Hudson, B. (2002). The effects of the Alexander technique on the respiratory system of the singer/actor. I: F. M. Alexander and concepts of his technique that affect respiration in singer/actors. *Journal of Singing, 59,* 9–17.

Jones, F. P. (1972). Voice production as a function of head balance in singers. *Journal of Psychology, 82,* 209–215.

Jones, F. P. (1976). *Body Awareness in Action: A Study of the Alexander Technique.* New York: Schocken.

Jones, F. P., Hanson, J. A., Miller, J. K., & Bossom, J. (1963). Quantitative analysis of abnormal movement: The sit and stand pattern. *American Journal of Physiological Medicine, 42,* 208–218.

Kaplan, I. (1994). *The Experience of Pianists Who Have Studied the Alexander Technique: Six Case Studies.* Unpublished doctoral dissertation, New York University.

Kemp, A. (1981). The personality structure of the musician: I. Identifying a profile of traits for the performer. *Psychology of Music, 9,* 3–14.

Kendrick, M. J., Craig, K. D., Lawson, D. M., & Davidson, P. O. (1982). Cognitive and behavioral therapy for musical-performance anxiety. *Journal of Consulting and Clinical Psychology, 50,* 353–362.

Lewis, P. P. (1980). *The Alexander Technique: Its Relevance for Singers and Teachers of Singing.* Unpublished doctoral dissertation, Carnegie Mellon University, Pittsburgh, PA.

Lloyd, G. (1986). *The Application of the Alexander Technique to the Teaching and Performing of Singing: A Case Study Approach.* Unpublished masters dissertation, University of Stellenbosch, Stellenbosch, South Africa.

Maitland, S., Horner, R., & Burton, M. (1996). An exploration of the application of the Alexander technique for people with learning disabilities. *British Journal of Learning Disabilites, 24,* 70–76.

McCormack, E. D. (1958). *Frederick Matthias Alexander and John Dewey: A Neglected Influence.* Unpublished doctoral dissertation, University of Toronto.

McCullogh, C. P. (1996). *The Alexander Technique and the Pedagogy of Paul Rolland.* Unpublished doctoral dissertation, Arizona State University, Tempe, AZ.

Nielsen, M. (1988). A study of stress amongst professional musicians. In C. Stevens (Ed.), *The F. M. Alexander Technique: Medical and Psychological Aspects* (pp. 14–16). Aalborg, Denmark: International School for the Alexander Technique.

Priest, J. (1992). *From Stage-Fright to Seat-Height: An Annotated Bibliography on the Alexander Technique and Music 1907–1992.* North Grosvenordale, CT: The author.

Richmond, P. (1994). The Alexander technique and dance training. *Impulse: The International Journal of Dance Science, Medicine, and Education, 2,* 24–38.

Richter, E. (1974). *The Application of the Alexander Technique to Cello Playing.* Unpublished doctoral dissertation, Florida State University, Tallahassee, FL.

Rohmert, F., Rehders, H., & Rohmert, G. (1990). Alexandertechnik: Auswirkungen der Korperaufrichtung auf das Klangspektrum der menschlinchen Stimme. In *Beitrage zum 1* (pp. 81–90). Kolloquium Praktische Musikphysiologie Koln: Schmidt.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory (Form Y1).* Palo Alto, CA: Consulting Psychologists Press.

Stallibrass, C. (1997). An evaluation of the Alexander Technique for the management of disability in Parkinson's disease: A preliminary study. *Clinical Rehabilitation, 11,* 8–12.

Stein, C. J. (1999). The Alexander technique: Its basic principles applied to the teaching and performing of stringed instruments. *American String Teacher, 49,* 72–77.

Stevens, C. H. (1987). *The Alexander Technique.* London: MacDonald.

Stevens, C. H., Bojsen-Møller, F., & Soames, R. (1989). Influence of initial posture on the sit to stand movement. *European Journal of Applied Physiology, 58,* 687–692.

Sweeney, G. A., & Horan, J. J. (1982). Separate and combined effects of cue-controlled relaxation and cognitive restructuring in the treatment of musical performance anxiety. *Journal of Counseling Psychology, 29,* 486–497.

Tursi, F. (1959). The problem of excessive psycho-physical tension in string performance. *American String Teacher, 9,* 6–8.

Valentine, E. R., & Williamon, A. (2003). Alexander technique and music performance: Evidence for improved "use". In R. Kopiez, A. C. Lehmann, I. Wolther, & C. Wolf (Eds.), *Proceedings of the Fifth Triennial ESCOM Conference* (pp. 145–147). Hanover, Germany: Hanover University of Music and Drama.

Valentine, E. R., Fitzgerald, D. F. P., Gorton, T. L., Hudson, J. A., & Symonds, E. R. C. (1995). The effect of lessons in the Alexander Technique on music performance in high and low stress situations. *Psychology of Music, 23,* 129–141.

Watson, P., & Valentine, E. (1987). The practice of complementary medicine and anxiety levels in a population of musicians. *Journal of the International Society for the Study of Tension in Performance, 4,* 26–30.

CHAPTER 11

# PHYSIOLOGICAL SELF-REGULATION: BIOFEEDBACK AND NEUROFEEDBACK

JOHN H. GRUZELIER AND TOBIAS EGNER

A high degree of control over both mental and emotional processes is of particular importance to performing musicians. Musicians must attain high concentration levels when performing—without getting overly tense or excited. At the same time they must achieve a degree of mental relaxation, flexibility, and assurance in order to express their own creativity and individuality and to communicate this to an audience. EEG biofeedback, or "neurofeedback," is a technique that teaches the individual how to control the rhythmic electrical activity of the brain. Brain rhythms differ in frequency (or speed) and in amplitude (or voltage). The amount of activity at certain frequencies reflects information processing, such as focused concentration or day-dreaming, and aspects of arousal, such as being tense, awake, relaxed, or sleepy. Neurofeedback makes individuals aware of these processes by *feeding back* a representation of their own electrical brain activity and allowing them to change it. Since particular frequencies are closely associated with, for example, states of high concentration or deep relaxation, musicians can use this feedback to guide themselves into brain states that are beneficial for performing at their peak. This chapter provides a survey of the theoretical and clinical underpinnings of neurofeedback and reviews recent studies that have systematically examined this training and its application to music. First, however, it places neurofeedback in the broader context of general biofeedback training.

The real impetus for this review has been the surprising efficacy of neurofeedback in recent studies with music conservatory students. After neurofeedback training, students have shown substantive increases in ratings by experts on the overall quality of musical performance and on artistic aspects of performance such as "interpretative imagination," "musicality," "stylistic accuracy," and "emotional commitment and conviction." These were controlled studies and the results were replicated across two years of research (see Gruzelier *et al.*, 2002; Egner & Gruzelier, 2003). At the same time, neurofeedback training has produced improvements in cognitive and neurophysiological measures of attention and in semantic working memory in either the same musicians or in concurrent studies with medical students (Egner & Gruzelier, 2001, 2004a; Vernon *et al.*, 2003). These results contribute to growing evidence of performance enhancement

due to neurofeedback in a diverse range of applications, and to what is anticipated to be a renaissance in the field of biofeedback training, both for optimizing performance in healthy individuals and for clinical applications.

## 11.1 Biofeedback

"Physiological self-regulation" describes voluntary control over functions of the central nervous system (CNS) and the peripheral nervous system (PNS), the latter comprising the somatic nervous system (SNS) and the autonomic nervous system (ANS; see Chapter 1 for an outline of the divisions of the human nervous system). Such control is acquired by means of "instrumental learning," which refers to the adjustment of behavior in response to the behavior's perceived consequences. These perceptions can then either act as positive or negative reinforcement (making the repetition of that behavior more likely) or as punishment (making the occurrence of the behavior less likely in the future). For any such trial-and-error learning to occur, the presence of feedback of a given behavior's consequences is essential. For instance, in order to learn to play the piano at an elementary level, it is crucial to receive tactile feedback from the fingers' interplay with the keyboard and auditory feedback in the shape of the sounds that are produced by this interaction.

Instrumental learning of psychomotor skills such as playing the piano, which are mediated by interaction between the CNS (the brain and spinal cord) and the SNS (motorneurons that innervate the musculoskeletal system), has never been subject to great controversy. On the other hand, the learning of voluntary control over ANS functions, such as heart rate and blood pressure, has historically been assumed to be impossible. However, the careful study of instrumental learning of ANS responses in animals has revealed that, provided accurate feedback linked to reinforcement, animals can achieve learned control over such autonomic measures as blood pressure and galvanic skin response (which reflects sweat gland activity; for a review, see Kimmel, 1974) and that control over these peripheral responses can be achieved in the absence of voluntary muscular changes (Miller & DiCara, 1967). These findings have provided an important foundation for the development of clinical applications in humans that make use of the individual's capacity for learned physiological self-regulation through a process called *biofeedback* training.

The biofeedback rationale holds that, if one can record a physiological response and *feed back* information about ongoing changes in this measure to the person from whom it is recorded, this feedback loop can enable instrumental learning of control over that physiological response. The progressive stages that comprise the successful therapeutical biofeedback training process have been identified by Stoyva and Budzynski (1993) as follows:

- acquiring awareness of the maladaptive response (e.g. high muscular tension)
- learning to associate certain changes in mental and bodily states with changes in the maladaptive response through the biofeedback process, leading to learned self-regulation
- the transference of the learned skill into everyday life.

Over the last three decades, this simple realization has spawned a great variety of effective clinical applications. For instance, biofeedback of blood pressure levels and electrodermal activity (i.e. skin conductance) has been employed in the treatment of hypertension (e.g. Benson *et al.*, 1971; Patel, 1973). Likewise, thermal biofeedback (e.g. finger temperature) has been used in the treatment of Raynaud's disease (e.g. Freedman *et al.*, 1983) and tension headache (see Blanchard & Andrasik, 1987). Probably the most widely used biofeedback modality is electromyography (EMG), which measures electric muscle activity. EMG biofeedback has been employed in the treatment of such diverse ailments as tension headaches (e.g. Reinking & Hutchins, 1981) and asthma (e.g. Kotses *et al.*, 1991), as well as for neuromuscular re-education after stroke (e.g. Basmajian *et al.*, 1975). Given that the most commonly reported medical problems in musicians are related to muscle and tendon injuries (Lockwood, 1989), EMG biofeedback constitutes a particularly promising treatment modality for performers.

Typically, however, EMG and other biofeedback regimes are employed as complementary therapeutic tools for general relaxation training in the treatment of many anxiety and stress-related conditions. It is likely that, in addition to primary gains associated with control over the specific physiological measure that represents the symptomatology of a given clinical problem, all of the mentioned biofeedback treatments lead to appreciable nonspecific relaxation effects. In fact, all biofeedback training protocols that address ANS functions appear to be geared at shifting the organism's ANS from an activated, sympathetic state (i.e. the "fight or flight" system) to a more relaxed, parasympathetic tone. Additional nonspecific factors that may contribute to biofeedback efficacy are an enhancement of general awareness of one's bodily state, as well as feelings of self-efficacy and control stemming from perceived success at the biofeedback task. From this portrayal of specific and nonspecific gains that can be derived from biofeedback training, it would appear quite obvious that the application of these techniques to issues related specifically to musicians and music performance could be a fruitful endeavor.

## 11.2 Biofeedback and music performance

A small number of research studies have explored benefits of biofeedback training within music (mostly specific to muscular control) and supplied promising results. For example, Morasky *et al.* (1981) studied the use of EMG biofeedback for reducing excess left-arm extensor muscle tension in string players. They reported that the performers learned to reduce muscular tension during playing and that these reductions were successfully transferred to their playing in the absence of feedback. Morasky *et al.* (1983) performed a similar study involving clarinet players, where forearm extensor EMG feedback was again found to facilitate significantly reduced tension that generalized beyond the feedback condition itself. In order to allay fears that a more relaxed muscle tone would negatively affect musical performance, they also showed that performance speed in playing trills and scales, which would have been most affected by this particular muscle relaxation, remained at comparable levels to pre-training performances.

Biofeedback regimes have also been employed to help musicians with specific career-threatening clinical symptoms. For instance, Levee *et al.* (1976) reported a case

study of a woodwind player suffering from tics due to high tension in throat and facial muscles. Specific EMG biofeedback, targeting the muscle groups in question, led to successful symptom reduction and improved (self-reported) musical performance ability. Similarly, LeVine (1983) documented dramatic improvements in a professional violinist suffering from functional palsy (also known as "craft palsy," of which writer's cramp and focal dystonia are examples) after she participated in a mixed course of behavioral therapy and thermal biofeedback.

Further intriguing examples of EMG biofeedback as a pedagogical tool in music include its application for removing unwanted left-hand tension in violin and viola players by feeding back muscle activity recorded from the adductor pollicis muscle. Self-rated improvements in muscle tension were evident even at 5-month follow-up interviews (LeVine & Irvine, 1984). Montes *et al.* (1993) found that EMG feedback training of activity of the abductor pollicis brevis muscle in pianists facilitated peak amplitude and relaxation rate values in this muscle group during thumb attacks in trills, a pattern of muscular activity that is seen in advanced pianists but not in beginners. From this brief survey of biofeedback applications in general and some peer-reviewed publications of its specific use in relation to music performance, it can be summarized that biofeedback training in principle could be employed to address profession-related clinical symptoms, general clinical and nonclinical levels of tension and anxiety, and as an instrument-specific pedagogical aid. It has to be stressed, however, that the use of objective performance measures for determining the benefits of such interventions has been almost entirely absent from this literature.

## 11.3 Neurofeedback

A different feedback modality to those discussed above is one that taps directly into the self-regulation of brain activity: biofeedback of the electroencephalogram (EEG) or neurofeedback. The EEG is commonly recorded via noninvasive electrodes attached to the scalp with paste or gel. These sensors allow for the measurement of minute electric potentials (i.e. brain rhythms or "brainwaves") emanating from the cortical neuronal populations just beneath the scalp electrode. This brainwave activity can be analyzed, by means of filtering or mathematical transformation, in terms of the amount of activity present in different "frequency bands" (i.e. at different speeds of brainwaves).

The body of knowledge concerned with mechanisms underlying cerebral rhythmic activity is elaborate, but nevertheless far from complete. In a general sense, rhythmic activity recorded at the scalp is the sum of electrical field potentials generated by cortical neurons in the vicinity of the electrode site, with one scalp electrode possibly integrating activity from neurons across 10 $cm^2$ of cortical surface (Nunez, 1995a). As random fluctuations in electrical potentials cancel each other out, any electrical activity detectable at the scalp necessitates approximately synchronous activity of a large number of neurons (for further discussion of the principles underlying fundamental characteristics and measurement of the human EEG, see Rosenweig & Leiman, 1989; Nunez, 1995b; Gazzaniga, 2000).

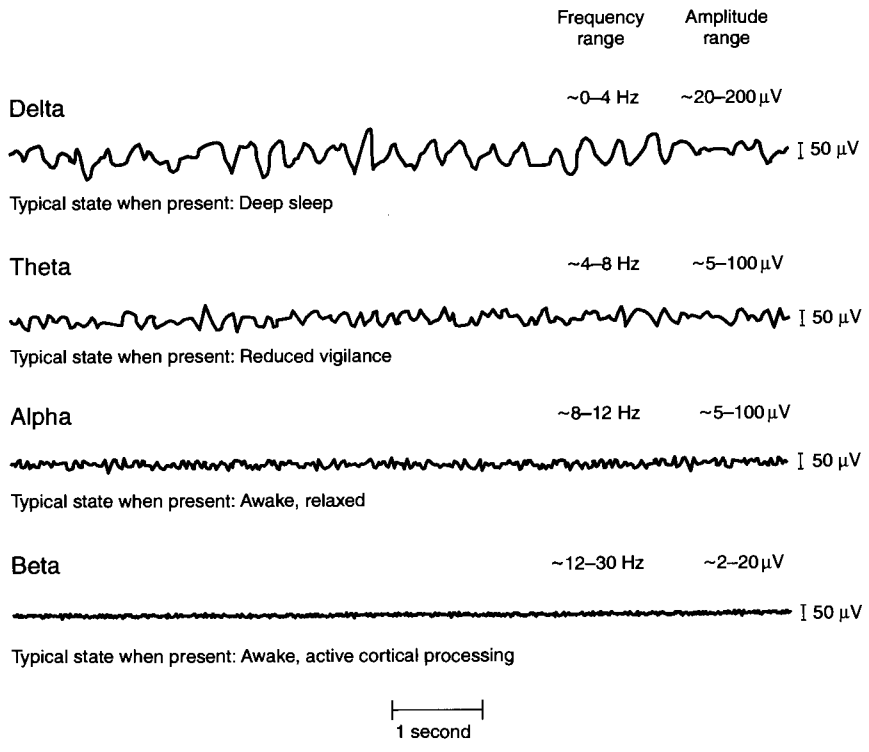



**Figure 11.1** Spectral components and characteristics of human brain rhythms.

The feasibility of modifying aspects of the EEG by means of instrumental learning has been demonstrated in animals by supplying a food reward for the production of a particular frequency component of the EEG. For example, it has been shown that cats (Wyrwicka & Sterman, 1968) as well as rhesus monkeys (Sterman *et al.*, 1978) readily learn to enhance brainwaves at specific frequencies in their EEG. In humans, intentional modulation of brain activity is usually achieved by displaying an approximately real-time representation of some EEG parameter to the trainee participant (thus providing biofeedback), paired with reinforcement to facilitate successful instrumental learning of the desired response.

The idea of employing EEG frequency-band activity as a feedback criterion in biofeedback training partly stems from the close association observed between the speed of EEG frequencies and the arousal state of the organism. For example, very slow brainwaves in the so-called "delta" range (~0–4 Hz) are primarily found in the human EEG during deep sleep. Slightly faster "theta" waves (~4–8 Hz) on the other hand are often associated with drowsiness and early sleep stages, while the adjacent "alpha" frequency (~8–12 Hz) is characteristic of a relaxed, waking state. Faster frequencies in the

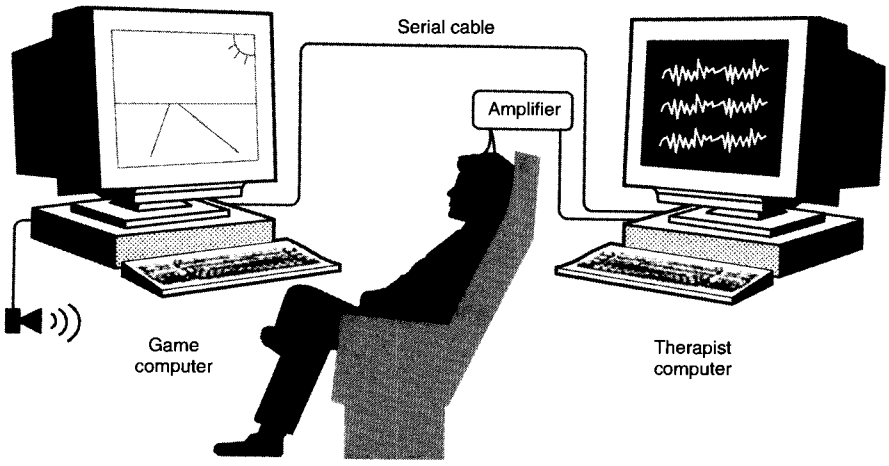



**Figure 11.2** The neurofeedback loop. Scalp electrodes are attached, with reference and ground electrodes attached to the earlobes. The EEG signal is filtered into different frequency bands and displayed to the experimenter or therapist. For the training of high frequencies (e.g. beta and SMR), the frequency components may be translated into a graphic computer-game display, visible to the participant. For lower frequencies (e.g. alpha and theta), the participant listens in a reclining position with eyes closed to sounds through headphones (e.g. waves on the shore, a babbling brook, a Thai gong). (From Neurocybernetic, 1998. © EEG Spectrum International, Inc., used by permission.)

"beta" (~12–30 Hz) and "gamma" range (above 30 Hz) are associated with more aroused, active cortical processing during mental operations in the alert brain (see Figure 11.1).

It is important to stress that this association between EEG rhythms and arousal/activational state of the organism is but one of many functional correlates of EEG activity and, as such, constitutes only a convenient simplification. For instance, in the context of carrying out different tasks, any one particular brain rhythm may reflect many diverse functional states of neural communication, and may be generated through different processes by various anatomical structures. Furthermore, many aspects of EEG generation and functional significance are very much under active research and, as of yet, not entirely understood.

In the typical neurofeedback set-up, EEG activity is recorded from a single scalp electrode, and referenced and grounded to electrodes attached to the earlobes. The recorded signal is then filtered in order to extract activity in the frequency bands of interest, and ongoing changes in activity in these bands are relayed in real time to the trainee participant, usually in the form of audio-visual feedback similar to a computer game (see Figure 11.2). The trainee gets rewarded (e.g. by winning points) for producing the desired changes in brainwave patterns, and over the course of a number of training sessions (which typically last between 15–30 minutes each) learns to self-regulate the EEG activity.

## 11.4 Neurofeedback and attention

A line of pioneering neurofeedback research was set in motion by Sterman's experiments on cats in the 1960s (for a review, see Sterman, 1996). Sterman and associates noted that during learned suppression of a previously conditioned response (pressing a bar for food), a particular brain rhythm emerged over the cats' sensorimotor cortex (Roth *et al.*, 1967; Wyrwicka & Sterman, 1968; Sterman & Wyrwicka, 1967; Sterman *et al.*, 1969). This rhythm was characterized by a frequency range of 12–20 Hz, with a spectral peak in the area of 12–14 Hz, and has since been referred to as the "sensorimotor rhythm" (SMR; Roth *et al.*, 1967). The researchers decided to study this distinct rhythm directly, attempting to teach the cats to produce SMR through instrumental learning, by making a food reward contingent on the occurrence of SMR "bursts" (Wyrwicka & Sterman, 1968; Sterman *et al.*, 1969). Cats learned this feat of EEG self-regulation with apparent ease, and the behavior associated with SMR production was one of behavioral stillness, with SMR bursts regularly preceded by a drop in muscle tone.

In a serendipitous twist, Sterman's laboratory was also commissioned to establish dose-response functions of a highly epileptogenic rocket fuel. When using the cats that had previously taken part in SMR conditioning, these cats were found to display elevated epileptic seizure thresholds compared to untrained cats. This research was successfully extrapolated to humans, where it was repeatedly documented that epileptic motor seizure incidence could be lowered significantly by SMR feedback training (Sterman & Friar, 1972; Sterman *et al.*, 1974; Sterman & MacDonald, 1978; Lantz & Sterman, 1988; for a review, see Sterman, 2000).

In acknowledgement of the apparent quieting effect of SMR training on the excitability of the sensorimotor system, Lubar and coworkers applied a protocol of SMR enhancement to the treatment of attention deficit hyperactivity disorder (ADHD; Lubar & Shouse, 1976). These researchers reported that enhancement of SMR with concurrent suppression of slow-wave theta activity (4–8 Hz) resulted in attentional improvements primarily facilitated by reduced motor hyperactivity (Lubar & Shouse, 1976; Shouse & Lubar, 1979). Subsequently, the training of SMR has often been complemented or substituted by another training protocol that combines suppression of theta activity with increments in higher beta components, such as the beta-1 band (15–18 Hz). These types of "beta" protocols have been conceptualized as targeting improvements in attentiveness (e.g. Lubar & Lubar, 1984).

The application of beta/SMR protocols to attentional disorders has since evolved into probably the most widely employed application within the field of neurofeedback. However, only rather recently have appropriately controlled studies begun to supply a proper scientific basis for the training's efficacy. For instance, Rossiter and LaVaque (1995) and Fuchs *et al.* (2003) have reported that beta/SMR neurofeedback has led to significant improvements in laboratory attention tests, as well as observational ratings of behavior, in children with ADHD; these improvements were at comparable levels to those seen with stimulant medication. Furthermore, Monastra *et al.* (2002) have shown that an extensive course of beta band training in addition to standard pharmacological treatment can lead to lasting benefits even after medication has been suspended. These studies have provided evidence for neurofeedback's

potential for enhancing attentional function in clinical groups, and offer the promise of possible applications to improving attentional abilities in healthy people (see Section 11.6).

## 11.5 Neurofeedback and relaxation

In what is believed to be the first ever application of neurofeedback training, Kamiya (1962) found that participants who were made aware of alpha-frequency bursts (8–12 Hz) in their EEG recorded from the occipital scalp regions (i.e. at the back of the head) could eventually identify alpha in the absence of feedback. They also appeared to be able to increase voluntarily the incidence of alpha rhythms, and they reported the subjective experience of the "alpha state" as being relaxing and peaceful. This initiated the pursuit of neurofeedback research that aimed to use EEG frequency characteristics as criteria for allowing the brain to be guided into different functional and activational states. Following this line of inquiry, a number of studies reported alpha density enhancement accompanied by a reduction in anxiety (and/or physiological arousal) by means of alpha feedback training (Kamiya, 1969; Brown, 1970; Budzynski & Stoyva, 1972; Hardt & Kamiya, 1978). This endeavor, however, did not always result in reliable replication of the subjective and physiological states originally associated with the "alpha state." Participants did not always succeed in actually enhancing alpha activity above pre-feedback levels (Lynch *et al.*, 1974), and in some studies, the behavioral and phenomenological effects could be attributed to social variables such as positive expectancies (Pressner & Savitzky, 1977) and perceived success at the biofeedback task (Plotkin & Rice, 1981).

More recently, an alpha/theta neurofeedback protocol that has been conceptualized as an "EEG-based relaxation therapy" (Peniston & Kulkosky, 1999, p. 158) has emerged as a complementary therapeutic tool in the treatment of alcoholism (Peniston & Kulkosky, 1989, 1990; Saxby & Peniston, 1995), as well as post-traumatic stress disorder (PTSD; Peniston & Kulkosky, 1991). The protocol is geared at facilitating a rise in levels of theta (4–8 Hz) over alpha (8–12 Hz) activity in a state of eyes-closed relaxation. Normally, upon closure of the eyes and onset of relaxation, the EEG displays high-amplitude rhythmic alpha activity. When the organism deactivates further, alpha activity slowly subsides and slower theta (and delta) activity gradually becomes predominant (see Broughton & Hasan, 1995; De Gennaro *et al.*, 2001). The point in time when theta activity supersedes alpha activity (the so-called "theta/alpha crossover") is in this context commonly associated with loss of consciousness and the onset of early sleep stages (i.e. sleep stage I). By teaching trainees to raise theta over alpha activity while not falling asleep, the alpha/theta protocol aims to teach people to enter a state of deep relaxation and deactivation consciously, apparently resembling a meditative state that would normally be unconscious.

Initial development of this kind of protocol (e.g. Budzynski & Stoyva, 1972) was based on the proposed association between alpha activity and subjective phenomenology of relaxation, as well as the implication of elevated theta amplitudes in certain types of meditative (Anand *et al.*, 1961) and "reverie" states (Green *et al.*, 1970). The

application of alpha neurofeedback to alcoholism was originally justified only in terms of the high levels of anxiety reported among alcoholics (Passini *et al.*, 1977). Subsequent research (Gabrielli *et al.*, 1982) has furthermore confirmed the negative correlations between alcoholism and alpha activity first reported by Funderburk (1949), as well as increased alpha band levels in alcoholics after alcohol consumption (Pollock *et al.*, 1983).

Peniston and Kulkosky have developed a treatment protocol that customarily combines alpha/theta neurofeedback with various other relaxation-inducing techniques, such as thermal (finger temperature) biofeedback, respiratory and autogenic instructions, and systematic desensitization procedures (Peniston & Kulkosky, 1999). The application of this combined treatment package, as an additional intervention to conventional therapy, has resulted in spectacular improvements in relapse rates in alcoholics (Peniston & Kulkosky, 1989; Saxby & Peniston, 1995). Dramatic symptom reductions in veterans with combat-related PTSD have also been reported from the same program of intervention (Peniston & Kulkosky, 1991). However, given the heterogeneous nature of their treatment package, it has been impossible to assess unambiguously the specific importance of the contribution of alpha/theta neurofeedback to the therapeutic process, and it has been suggested that it may be no different from alternative relaxation techniques (Taub & Rosenfeld, 1994; Graap & Freides, 1998). Even the basic feasibility of gaining voluntary control over this specific EEG signature has long remained unsubstantiated. However, in a recent study, Egner *et al.* (2002) have demonstrated that accurate alpha/theta feedback did successfully facilitate learned control over these frequency components and led to significantly elevated theta-to-alpha ratios in comparison with relaxation under noncontingent feedback conditions.

Regardless of the ambivalent status of its scientific evaluation, the alpha/theta neurofeedback protocol has gained considerable popularity with biofeedback practitioners in application to a number of tension- and anxiety-related clinical conditions and, as such, could provide a potential tool for addressing similar issues commonly found in musicians.

## 11.6 Neurofeedback and performance enhancement in musicians

It is clear that peak performance in music requires tremendous control over the brain processes underlying shifts of attentional and activational states. Working on the basis that such control is of central importance to the musician, one could argue that a program of experimental application of the clinical neurofeedback paradigms described previously could be applied to music with professionally significant consequences. In recent research based at a music conservatory, the authors have conducted two studies to examine the strength of this line of reasoning. In particular, the studies were carried out as part of the project "Zoning In: Motivating the Musical Mind" (see also Chapters 9, 10, and 12), with the intention of establishing the impact of these neurofeedback training paradigms on laboratory-based behavioral and neurophysiological measures of attention and on the quality of music performance.

### 11.6.1 Enhancing attention

In the first study, a group of music students were trained on both an attention-targeting beta1/SMR neurofeedback protocol (i.e. 10 15-minute sessions of both beta-1 and SMR training) and a relaxation-targeting alpha/theta neurofeedback protocol (also 10 15-minute sessions). The assessment of attention performance was carried out employing a computerized test displaying two classes of stimuli: (1) "targets," which require the participant to respond as quickly and accurately as possible (by pushing a response switch) and (2) "nontargets," which require the participant to refrain from responding. Thus, two types of errors can be incurred on such a task: (1) errors of *omission* (i.e. failing to respond to a target stimulus) and (2) errors of *commission* (i.e. erroneously responding to a nontarget stimulus). Respectively, these errors are held to reflect inattentiveness and impulsiveness. A further attention measure, derived from signal detection theory (Green & Swets, 1966), is termed "perceptual sensitivity" or "d prime" (d′) and takes into account both of these error types by expressing a ratio of hit rate to false alarm rate.

From the emergent data, it was established that 10 training sessions of both beta-1 and SMR neurofeedback led to a significant reduction in errors (Egner & Gruzelier, 2001). More specifically, students showed a significant reduction in commission errors as compared with measures taken before training. When exploring the link between the process of learned EEG self-regulation and this reduction in impulsive mistakes, it was found that the relative success at enhancing the SMR was highly positively correlated with reduced commission errors, meaning that participants who did well on the SMR feedback task tended to decrease the amount of impulsive mistakes after training. These findings support the notion that learned SMR enhancement is associated with improved response inhibition and constitute the first evidence for cognitive performance enhancement through neurofeedback in healthy volunteers (Egner & Gruzelier, 2001).

It was furthermore established that SMR neurofeedback training can also improve the incidence of omission errors (inattentive mistakes) in healthy subjects (Egner & Gruzelier, 2004a), and recent findings even suggest that memory function can be enhanced by similar means (Vernon *et al.*, 2003). In addition, the attention-enhancing potential of beta-1 neurofeedback has been corroborated by electrocortical performance measures related to selective attention processes. Specifically, beta-1 neurofeedback training is associated with increments in the P300 event-related brain potential (Egner & Gruzelier, 2001; Egner & Gruzelier, 2004a), which has been conceptualized as representing activity in neuronal sources responsible for updating relevant stimulus environment information in working memory (Donchin & Coles, 1988; see Chapter 7 of this volume for a short discussion of working memory). While these data offer support for the general feasibility of employing neurofeedback for improving cognitive performance in healthy individuals, they might be of no meaningful consequence at all for the enhancement of such a complex blend of requirements as live musical performance.

### 11.6.2 Achieving peak performance

The variable of relevance to the issue of achieving peak performance is, of course, the measurement of music performance quality itself. While some form of quantitative evaluation of music performance quality has long formed a routine part of the

assessment policies within music schools and conservatories, the issue of reliable measurement of such seemingly subjective concepts as a performer's musical interpretation is, not surprisingly, an issue of some debate. The use of performance measures as dependent variables in music psychology research has gained popularity in recent years (e.g. Juslin & Laukka, 2000; Williamon & Valentine, 2000, 2002), and the actual process of performance evaluation is under active investigation (Thompson & Williamon, 2003). Attempts at quantifying quality judgments of musical performance in order to achieve reliable assessments have typically involved having expert evaluators complete structured evaluation forms that elicit judgments along a range of performance aspects (e.g. "technical ability" versus "artistic expression"; see Chapter 4 for a discussion of performance assessment schemes).

The purpose of the music performance measurements in the neurofeedback research was to assess performance variables of high ecological validity and pedagogical relevance; therefore, it seemed appropriate to employ an evaluation scheme by which students are assessed customarily. For this reason, the marking scheme of the Associated Boards of the Royal Schools of Music (Harvey, 1994) was adapted to take into account segmented aspects of performance quality (including "instrumental competence," "musicality," and "communicative ability"), as well as overall performance quality (see Table 11.1 for a listing of the four main assessment categories and the adapted sub-scales).

While examination assessments within a conservatory are normally carried out on live performances by teachers both internal and external to the institution, in the context of a scientific investigation, this type of evaluation procedure would seem to introduce potential sources of bias. Foremost, the assessors would be aware of the order of pre- and post-training performances, which might induce expectations of improved performances. In order to exert maximum control over such biases, the performances

**Table 11.1** Music performance evaluation scales and Pearson product-moment correlation coefficients between change scores in music performance evaluation and alpha/theta learning (from Egner & Gruzelier, 2003). Major evaluation categories are in italics, with their associated sub-scales following below

| | Alpha/theta learning |
|---|---|
| *Overall quality* | $r = 0.47$, $p = 0.038$ |
| *Instrumental competence (i.e. technical ability)* | $r = 0.50$, $p = 0.029$ |
| Level of technical security | $r = 0.39$, $p = 0.086$ |
| Rhythmic accuracy | $r = 0.65$, $p = 0.003$ |
| Tonal quality and spectrum | $r = 0.39$, $p = 0.140$ |
| *Musicality* | $r = 0.54$, $p = 0.017$ |
| Stylistic accuracy | $r = 0.58$, $p = 0.007$ |
| Interpretative imagination | $r = 0.48$, $p = 0.037$ |
| Expressive range | $r = 0.53$, $p = 0.016$ |
| *Communicative ability* | $r = 0.55$, $p = 0.013$ |
| Deportment | $r = 0.45$, $p = 0.052$ |
| Communication of emotional commitment and conviction | $r = 0.51$, $p = 0.021$ |
| Ability to cope with situational stress | $r = 0.44$, $p = 0.052$ |

here were videotaped, randomized, and then evaluated only by experienced assessors external to the institution and blind as to the order of performances and to experimental group membership of the students. In order to assess whether any changes in music performance ratings were related to pre-performance anxiety, the Spielberger's State Anxiety Inventory (Spielberger *et al.*, 1983) was administered to each participant just before performance.

As mentioned earlier, a group of music students took part in a mixed course of beta-1/SMR and alpha/theta training in the first study (Gruzelier *et al.*, 2002; Egner & Gruzelier, 2003). Prior and subsequent to the training program, the students were assessed on the quality of their performances. The impact of participation in this mixed course of beta-1/SMR and alpha/theta neurofeedback was assessed in comparison with a no-training control group, as well as with a randomly selected subgroup of neurofeedback participants who engaged in additional interventions. These additional training components consisted of a physical exercise regime (see Chapter 9) and a course of "mental skills training" composed of interventions derived from sports psychology (see Chapter 12). These comparison groups were conceived in order to allow for the assessment of neurofeedback's impact on performance as a stand-alone intervention versus a subcomponent of a more elaborate training program, while controlling for the presumed natural progression in performance quality resulting from the standard conservatory training (through the no-training control group).

The blind rating scores of music performance quality before and after training obtained from two expert judges were averaged for each rating scale for each student. It was found that improvements in performance were evident in the neurofeedback-only group, but neither in the neurofeedback group engaging in additional interventions nor in the no-training control group (see Figure 11.3). The neurofeedback group improved most markedly on ratings of overall quality of performance and of their "communication of emotional commitment and conviction."

Interestingly, an alpha/theta training learning index, reflecting increasing ease and depth of relaxation across the training process, correlated highly positively with music

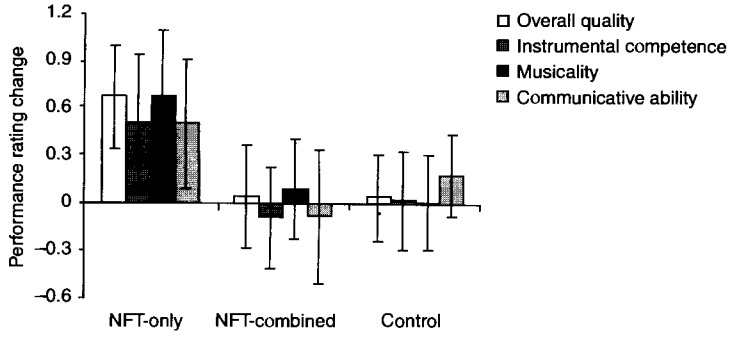



**Figure 11.3** Mean music performance rating change scores (i.e. scores from the post-training performance minus those of the pre-training performance, on a 10-point rating scale) on the main evaluation categories for neurofeedback-only (NFT-only), neurofeedback and additional interventions (NFT-combined), and a no-training control group. The whiskers indicate mean ± standard error.

performance improvements. As can be seen in Table 11.1, alpha/theta learning was significantly associated (i.e. $p < 0.05$) with trends of improvement in 9 out of 13 rating criteria, including the main four. Neither the SMR nor the beta-1 protocols were related to improvements in music performance. Furthermore, differential improvement rates between the experimental groups in this study were not related to pre-performance state anxiety, as generally decreased anxiety levels between the first and second performances did not differ between groups.

These findings supply evidence for a potential benefit of neurofeedback on a highly ecologically valid music performance measure. The fact that music performance quality changes were not related to SMR and beta-1 learning suggests that improvements were not mediated by attention-related variables, whereas the fact that alpha/theta learning correlated highly with changes in virtually all music evaluation categories would seem to point to a single, pervasive factor mediating these correlations. Given the nature of the alpha/theta protocol and these relationships, a prime candidate for effect mediation would appear to be pre-performance anxiety, but the data did not support this assumption.

In order to clarify the seemingly strong association between alpha/theta neurofeedback and performance enhancement, a quasi-replication study was devised (Gruzelier *et al.*, 2002; Egner & Gruzelier, 2003), where participants were randomly allocated to one of the following: an alpha/theta, SMR, or beta-1 neurofeedback training group, a physical exercise program, or a mental skills training program. As before, music performances were assessed before and after training. In addition, a further comparison group was integrated into the study consisting of students involved in a course of Alexander technique training, as this technique is considered an established tool for improving performance in music conservatories worldwide. This program engaged participants in a comparable amount of one-to-one interaction as the neurofeedback intervention (see Chapter 10 for further discussion of the Alexander technique).

Analysis of music performance ratings from three expert judges blind to the experimental conditions revealed that the alpha/theta group displayed significant improvements, while neither the beta-1 nor the SMR group exhibited any post-training performance changes. Similarly, students from the Alexander technique, physical exercise, and mental skills training groups showed no post-training changes. In the alpha/theta group, evaluation scores for "musicality," "stylistic accuracy," "interpretative imagination," and "overall quality" were all significantly improved (see Figure 11.4). These increments represent average alpha/theta group improvements between 13.5% and 17%, with a mean improvement rate of 12% across all evaluation scales. Individual participants displayed improvements of over 50% on some evaluation criteria. As in the first study, all groups reported significantly less pre-performance anxiety before the post-training performance, with no differences between groups.

In the previous study, alpha/theta learning correlated with musical improvements across all evaluation categories. Here, the protocol's performance enhancing effects proved to be replicable particularly with respect to parameters on the "musicality" evaluation category (see Table 11.1). These data could be interpreted as indicating that alpha/theta training led especially to improvements on attributes of artistic expression (as opposed to technical skills), which in turn improved overall performance. These

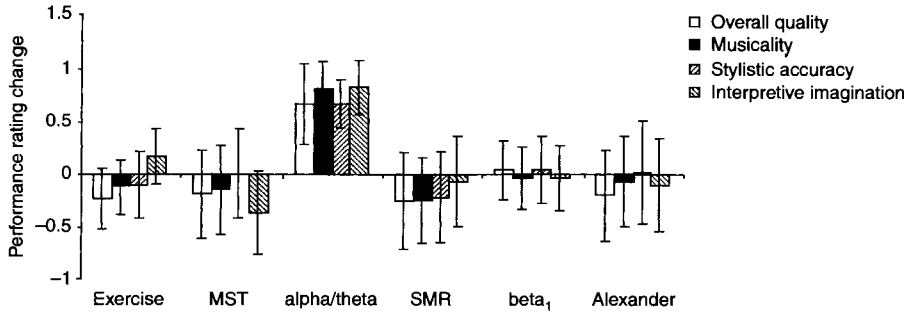



**Figure 11.4** Mean music performance rating change scores (i.e. scores from the pre-training performance minus those of the post-training performance, on a scale of 1–10) for the exercise, mental skills training (MST), alpha/theta, SMR, $beta_1$, and Alexander technique groups. The whiskers indicate mean ± standard error on a 10-point rating scale of musical evaluation criteria (from Egner & Gruzelier, 2003).

results suggest that alpha/theta training appears reliably to enhance artistic aspects of musical performance skills, independently of training on additional neurofeedback protocols and that these effects are superior to the other interventions in this respect. These data confirm a significantly beneficial effect of alpha/theta training on a highly ecologically valid and pedagogically relevant performance measure. Indeed, the improvements were equivalent to two academic grades within the conservatory's assessment system. These effects cannot be accounted for by invoking practice, motivational, or generic neurofeedback factors. An explanation based on generic relaxation can also be discounted on the grounds that, once again, the alpha/theta training was not associated with a greater decrease in pre-performance anxiety than that seen in other groups. Additionally, a mental skills training group, which engaged in extensive relaxation training, showed no detectable performance improvements.

The size of the performance improvements implies great potential for the implementation of this application in music performance contexts. But how does alpha/theta training achieve these remarkable effects? In theoretical terms, the data do not appear to fit with a conceptualization of training efficacy primarily based on relaxation. The need to look beyond a mere relaxation model is further emphasized by a study showing that significantly different theta-to-alpha ratios between real and false alpha/theta feedback conditions were not reflected in differences in reports of subjective relaxation (Egner *et al.*, 2002).

When inspecting EEG changes within the sessions across the training process, an increase in the production of theta amplitudes has been noted (Egner & Gruzelier, 2004b). EEG theta activity has been implicated in a number of mental and affective states ostensibly unrelated to relaxational processes. Furthermore, the clinical efficacy of alpha/theta training in patients suffering trauma and chronic drug addiction, extending in recent evidence to young offenders with crack and cocaine misuse, implies a more powerful impact than can be achieved with relaxation and anxiety reduction alone. Theta is generated throughout the cortex in hypnogogic states and following training in meditation and hypnosis (Vaitl *et al.*, 2003); its production can

often outlast the experience itself (Williams & Gruzelier, 2001). These states are typically associated with pleasure, and in the case of meditation, theta production has been positively associated with ratings of bliss (Aftanas & Golocheikine, 2001). Accordingly, theta appears associated with enhanced feelings of well-being as well as with relaxation; in the words of one participating music student: "it lets the mind breathe" (Gruzelier *et al.*, 2002).

Relevant too is the growing evidence of the role of theta in a range of memory control processes, enabling a fluent retrieval of memories into conscious awareness and working memory (Klimesch, 1996; Burgess & Gruzelier, 1997; Sarnthein *et al.*, 1998; Klimesch *et al.*, 2001). It is self-evident that memory control processes are exercised to the hilt in musical performance, when both virtuoso technical accomplishment and artistry are demanded.

Intriguingly, a bold theory of the relation between theta, dreams, and survival by Winson (2002) chimes with the findings presented here. Winson has long observed that across species theta from the hippocampus (a structure situated deep in the brain's limbic system) is generated both in situations that challenge survival and in dream sleep. Consistent with the evidence of the importance of theta for memory control processes, he posits that theta assists in making available to memory information that is necessary for survival. This follows his observations of survival behavior in animals, which (though species-specific) is united by the occurrence of theta. Theta also occurs in all species during dream sleep, which serves the purpose of facilitating access to memory control processes of information regarding events that threaten survival in order to deal more effectively with any recurrence of threat. The process of alpha/theta training, in fact, involves the transient entering and re-entering of dream-like states. Let us consider that the typical music student's career trajectory has taken 10 years or so of singular dedication to the practice and performance of music at the highest level. Now his or her musical ability and career ambition is under the challenge of a stressful performance to be judged by teachers and other experts. Accordingly, the exercise of theta-related processes through alpha/theta training may well facilitate success, for the successful student orchestrates a dazzling technical and artistic display that is dependent on memory control processes. This involves a complex coordination of sensory-motor, postural, and respiratory functions, together with the expression and modulation of emotion and the expression of creative and artistic ability (as is inherent in the concept of musicality).

A further line of interpretation of these results stems from considerations of the post-training effects of alpha/theta neurofeedback on the EEG amplitudes in different frequency bands across the whole scalp. From measuring whole-head multi-channel EEG activity during rest, there is evidence for alpha/theta training being related to a reduction in high beta-frequency activity (15–30 Hz) over frontal brain regions (Egner *et al.*, 2003). The training's effect of reducing frontal beta band activity could quite readily be interpreted as reflecting decreased agitation and negative affect, based on associations between frontal beta activity and depression (Pollock & Schneider, 1990) and post-traumatic stress (Begic *et al.*, 2001), and with stress responses to brief painful stimulation in nonclinical groups (Chen *et al.*, 1989; Chen, 2001). Based on these data, a theoretical account of alpha/theta neurofeedback efficacy has to integrate its apparent

role in reducing vigilant arousal and its positive effects on artistic expression in musical performance.

Irrespective of the precise mechanisms underlying the training effects, the findings presented here have important implications with respect to both the psychological and physical well-being of musicians. The significant enhancement of performance skills may in the long run alleviate excessive worry about performing, the most commonly cited impediment to musicians' successful careers (Fishbein *et al.*, 1988; Steptoe, 1989). Furthermore, the enhancement of artistic aspects of performance warrant application to the performing arts in general.

Although a theoretical link between functional correlates of theta activity (as well as alpha/theta training effects on frontal beta activity) and improved artistic expression in music performance can easily be envisaged, no effect mediation has yet been established. However, the remarkable neurofeedback training effects demonstrated in these studies should serve as a motivation for empirically addressing the exact mechanisms underlying the training in the future. Firstly, further clarification of the precise nature of the alpha/theta effects is required for advancing the theoretical framework. With regards to subjective phenomenology, a wide scope of self-report measures assessing the experiences associated with the training process could prove elucidating for pinpointing psychological mediators of the training's efficacy. Furthermore, the impact of alpha/theta training on arousal states within training sessions, as well as at post-training, could be assessed by peripheral physiological measures, such as the skin conductance response. The monitoring of whole-scalp EEG changes within each training session would also be of great interest, above all in order to elucidate the origin of the theta activity generated during the training and the way in which the training may affect frontal beta band activity.

Of more immediate practical concern to the musician and educator are the questions of who is most likely to benefit from the training, how to optimize the nature and duration of the training, and how long the effects may be expected to last. These issues are at the present moment unresolved but under active investigation. Elaborate psychometric testing could possibly allow one to determine personality trait predictors of likely responsiveness to alpha/theta neurofeedback and successful performance enhancement. It is also conceivable that particular instrument groups are more likely to benefit from this intervention than others.

The relevance of applying this technique to music performers hinges decidedly on its potential to evoke long-term effects. The research to date does not permit any inferences regarding this important aspect, as all post-training music performance measures were taken within a time-span of maximally four weeks after the last training session. In the future, studies that involve regular follow-up assessments over a longer interval will have to be conducted in order to determine whether the costs of neurofeedback training (both in terms of time and money) as a performance enhancement tool are justified by long-term returns.

## 11.7 Conclusions

The field of learned physiological self-regulation through means of biofeedback holds promising potential for enhancing musical performance (see Table 11.2 for a summary

Table 11.2 A summary of the clinical and music performance benefits of biofeedback and neurofeedback

| | Clinical benefits | Music performance benefits |
|---|---|---|
| *Biofeedback* | | |
| Blood pressure | Treatment of hypertension (Benson *et al.*, 1971; Patel, 1973) | |
| Electrodermal | Treatment of hypertension (Benson *et al.*, 1971; Patel, 1973) | |
| Thermal | Treatment of Raynaud's disease (Freedman *et al.*, 1983)<br>Treatment of tension headaches (Blanchard & Andrasik, 1987) | Treatment of functional palsy in a professional violinist (LeVine, 1983) |
| EMG | General relaxation training<br>Treatment of tension headaches (Reinking & Hutchins, 1981)<br>Treatment of asthma (Kotses *et al.*, 1991)<br>Neuromuscular re-education after stroke (Basmajian *et al.*, 1975) | Reduction of left arm tension in string players (Morasky *et al.*, 1981; LeVine & Irvine, 1984)<br>Reduction of forearm tension in clarinet players (Morasky *et al.*, 1983)<br>Reduction of tension in facial muscles of a woodwind player (Levee *et al.*, 1976)<br>Facilitation of peak amplitude and relaxation rate values in pianists' trills (Montes *et al.*, 1993) |
| *Neurofeedback* | | |
| SMR | Treatment of epilepsy (Sterman & MacDonald, 1978; Lantz & Sterman, 1988)<br>Treatment of attention deficit hyperactivity disorder (Rossiter & LaVaque, 1995; Fuchs *et al.*, 2003) | Improved attention performance in music students (Egner & Gruzelier, 2001, 2004a) |
| Beta | Treatment of attention-deficit hyperactivity disorder (Rossiter & LaVaque, 1995; Monastra *et al.*, 2002; Fuchs *et al.*, 2003) | Enhanced attention-related P300 event-related brain potential in music students (Egner & Gruzelier, 2001, 2004a) |
| Alpha | Reduction of anxiety in anxiety-prone students (Hardt & Kamiya, 1978) and in alcoholics (Passini *et al.*, 1977) | |
| Alpha/theta | Treatment of alcoholism (Peniston & Kulkosky, 1989, 1990) and posttraumatic stress disorder (Peniston & Kulkosky, 1991) | Improved music performance in music students (Egner & Gruzelier, 2003) |

of emergent findings). Biofeedback of peripheral nervous system measures, notably EMG activity, has successfully been employed as an educational aid and as a therapeutic tool in reducing muscular tension in string and wind players. More recently, an exciting new possibility of improving music performance has arisen through the use of EEG biofeedback of alpha and theta frequencies. Research has shown that alpha/theta neurofeedback can facilitate replicable improvements in music performance skills in conservatory students. These effects were not related to learning of a relaxation technique per se, nor to reduced pre-performance anxiety levels. In order to understand fully the way in which the alpha/theta protocol produces its effects, the training process

will have to be monitored closely by assessing both cerebral and peripheral electrophysiological responses and by elucidating in detail the resultant phenomenology.

## Acknowledgments

This research was supported by grants from The Leverhulme Trust, the Research Strategy Fund of Royal Holloway, University of London, and the Royal College of Music, London.

## Further information and reading

Additional information on physiological self-regulation, biofeedback, and neurofeedback can be obtained from the websites of the following societies: the Association for Applied Psychophysiology and Biofeedback (*www.aapb.org*), the Society for Neuronal Regulation (*www.snr-jnt.org*), and the European chapter of the Society for Neuronal Regulation (*www.e-isnr.org*). Further particulars of the research results presented in this chapter are provided at the project website for "Zoning In: Motivating the Musical Mind": *www.zoningin.rcm.ac.uk.* For related reading, see:

Evans, J. R., & Abarbanel, A. (Eds.) (1999). *Introduction to Quantitative EEG and Neurofeedback.* London: Academic Press.

Freeman, W. J. (1999). *How Brains Make Up Their Minds.* London: Weidenfeld & Nicolson.

Gazzaniga, M. S. (Ed.) (2000). *Cognitive Neuroscience: A Reader.* Oxford: Blackwell.

Hatch, J. P., Fisher, J. G., & Rugh, J. D. (Eds.) (1987). *Biofeedback: Studies in Clinical Efficacy.* New York: Plenum.

Nunez, P. L. (Ed.) (1995b). *Neocortical Dynamics and Human EEG Rhythms.* Oxford: Oxford University Press.

Shapiro, D., Barber, T. X., DiCara, L. V., Kamiya, J., Miller, N. B., & Stoyva, J. M. (Eds.) (1972). *Biofeedback and Self-Control.* Chicago: Aldine.

## References

Aftanas, L. I., & Golocheikine, S. A. (2001). Human anterior and frontal midline theta and lower alpha reflect emotionally positive state and internalized attention: High-resolution EEG investigation of meditation. *Neuroscience Letters, 310,* 57–60.

Anand, B. K., China, G. S., & Singh, B. (1961). Some aspects of electroencephalographic studies in yogis. *Electroencephalography and Clinical Neurophysiology, 13,* 452–456.

Basmajian, J. V., Kukulka, C. J., Narajan, M. G., & Takebe, K. (1975). Biofeedback treatment of foot drop after stroke compared with standard rehabilitation technique: Effects on voluntary control and strength. *Archives of Physical Medicine and Rehabilitation, 56,* 231–236.

Begic, D., Hotujac, L., & Jokic-Begic, N. (2001). Electroencephalographic comparison of veterans with combat-related post-traumatic stress disorder and healthy subjects. *International Journal of Psychophysiology, 40,* 167–172.

Benson, H., Shapiro, D., Tursky, B., & Schwarz, G. E. (1971). Decreased systolic blood pressure through operant conditioning techniques in patients with essential hypertension. *Science, 173,* 740–741.

Blanchard, E. B., & Andrasik, F. (1987). Biofeedback treatment of vascular headache. In J. P. Hatch, J. G. Fisher, & J. D. Rugh (Eds.), *Biofeedback: Studies in Clinical Efficacy* (pp. 1–79). New York: Plenum.

Broughton, R., & Hasan, J. (1995). Quantitative topographic electroencephalographic mapping during drowsiness and sleep onset. *Journal of Clinical Neurophysiology, 12*, 372–386.

Brown, B. B. (1970). Recognition of aspects of consciousness through association with EEG alpha activity represented by a light signal. *Psychophysiology, 6*, 442–452.

Budzynski, T. H., & Stoyva, J. M. (1972). Biofeedback techniques in behavior therapy. In D. Shapiro, T. X. Barber, L. V. DiCara, J. Kamiya, N. B. Miller, & J. M. Stoyva (Eds.), *Biofeedback and Self-Control* (pp. 437–459). Chicago: Aldine.

Burgess, A. P., & Gruzelier, J. H. (1997). Short duration synchronization of human theta rhythm during recognition memory. *NeuroReport, 8*, 1039–1042.

Chen, A. C. N. (2001). New perspectives in EEG/MEG brain mapping and PET/fMRI neuroimaging of human pain. *International Journal of Psychophysiology, 42*, 147–159.

Chen, A. C. N., Dworkin, S. F., Haug, J., & Gehrig, J. (1989). Topographic brain measures of human pain and pain responsivity. *Pain, 37*, 129–141.

De Gennaro, L., Ferrara, M., & Bertini, M. (2001). The boundary between wakefulness and sleep: Quantitative electroencephalographic changes during the sleep onset period. *Neuroscience Letters, 107*, 1–11.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11*, 357–374.

Egner, T., & Gruzelier, J. H. (2001). Learned self-regulation of EEG frequency components affects attention and event-related brain potentials in humans. *NeuroReport, 12*, 4155–4159.

Egner, T., & Gruzelier, J. H. (2003). The effects of neurofeedback training on the spectral topography of the healthy electroencephalogram. Manuscript submitted for publication.

Egner, T., & Gruzelier, J. H. (2004a). EEG Biofeedback of low beta band components: Frequency-specific effects on variables of attention and event-related brain potentials. *Clinical Neurophysiology, 115*, 131–139.

Egner, T., & Gruzelier, J. H. (2004b). The temporal dynamics of electroencephalographic responses to alpha/theta neurofeedback training in healthy subjects. *Journal of Neurotherapy, 8*, 43–57.

Egner, T., Strawson, E., & Gruzelier, J. H. (2002). EEG signature and phenomenology of alpha/theta neurofeedback training versus mock feedback. *Applied Psychophysiology and Biofeedback, 27*, 261–270.

Egner, T., Zech, T. F., & Gruzelier, J. H. (2003). *The effects of neurofeedback training on the spectral topography of the healthy electroencephalogram.* Manuscript submitted for publication.

Fishbein, M., Middelstadt, S. E., Ottati, V., Strauss, S., & Ellis, A. (1988). Medical problems among ICSOM musicians: Overview of a national survey. *Medical Problems of Performing Artists, 3*, 1–8.

Freedman, R. R., Ianni, P., & Wenig, P. (1983). Behavioural treatment of Raynaud's disease. *Journal of Consulting and Clinical Psychology, 51*, 539–549.

Fuchs, T., Birbaumer, N., Lutzenberger, W., Gruzelier, J. H., & Kaiser, J. (2003). Neurofeedback treatment for attention-deficit/hyperactivity disorder in children: A comparison with methylphenidate. *Applied Psychophysiology and Biofeedback, 28*, 1–12.

Funderburk, W. H. (1949). Electroencephalographic studies in chronic alcoholics. *Electroencephalography and Clinical Neurophysiology, 1*, 369–370.

Gabrielli, W. F., Mednick, S. A., Vlovka, J., Pollack, V. E., Schulsinger, F., & Itil, T. M. (1982). Electroencephalograms in children of alcoholic fathers. *Psychophysiology, 19*, 404–407.

Gazzaniga, M. S. (Ed.). (2000). *Cognitive Neuroscience: A Reader.* Oxford: Blackwell.

Graap, K., & Freides, D. (1998). Regarding the database for the Peniston alpha-theta EEG biofeedback protocol. *Applied Psychophysiology and Biofeedback, 23*, 265–272.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics.* New York: Wiley.

Green, E. E., Green, A. M., & Walters, E. D. (1970). Voluntary control of internal states: Psychological and physiological. *Journal of Transpersonal Psychology, 2*, 1–26.

Gruzelier, J. H., Egner, T., Valentine, E., & Williamon, A. (2002). Comparing learned EEG self-regulation and the Alexander Technique as a means of enhancing musical performance. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the Seventh International Conference on Music Perception and Cognition* (pp. 89–92). Adelaide, Australia: Causal Productions.

Hardt, J. V., & Kamiya, J. (1978). Anxiety change through electroencephalographic alpha feedback seen only in high alpha subjects. *Science, 201*, 79–81.

Harvey, J. (1994). *These Music Exams.* London: Associated Board of the Royal Schools of Music.

Juslin, P. N., & Laukka, P. (2000). Improving emotional communication in music performance through cognitive feedback. *Musicæ Scientiæ, 4*, 151–183.

Kamiya, J. (1962). *Conditioned Discrimination of the EEG Alpha Rhythm in Humans.* Paper presented at the Western Psychological Association, San Francisco, CA.

Kamiya, J. (1969). Operant control of the EEG alpha rhythm and some of its reported effects on consciousness. In C. T. Tart (Ed.), *Altered States of Consciousness* (pp. 519–529). New York: Wiley.

Kimmel, H. D. (1974). Instrumental conditioning of autonomically mediated response in human beings. *American Psychologist, 29*, 325–335.

Klimesch, W. (1996). Memory processes, brain oscillations and EEG synchronisation. *International Journal of Psychophysiology, 24*, 61–100.

Klimesch, W., Doppelmayr, M., Yonelinas, A., Kroll, N. E., Lazzara, M., Rohm, D., & Gruber, W. (2001). Theta synchronisation during episodic retrieval: Neural correlates of conscious awareness. *Cognitive Brain Research, 12*, 33–38.

Kotses, H., Harver, A., Segreto, J., Glaus, K. D., Creer, T. L., & Young, G. A. (1991). Long-term effects of biofeedback induced facial relaxation on measures of asthma severity in children. *Biofeedback and Self-Regulation, 16*, 1–21.

Lantz, D., & Sterman, M. B. (1988). Neuropsychological assessment of subjects with uncontrolled epilepsy: Effects of EEG biofeedback training. *Epilepsia, 29*, 163–171.

Levee, J. R., Cohen, M. J., & Rickles, W. H. (1976). Electromyographic biofeedback for relief of tension in the facial and throat muscles of a woodwind musician. *Biofeedback and Self-Regulation, 1*, 113–130.

LeVine, W. R. (1983). Behavioural and biofeedback therapy for a functionally impaired musician: A case report. *Biofeedback and Self-Regulation, 8*, 101–107.

LeVine, W. R., & Irvine, J. K. (1984). In vivo EMG biofeedback in violin and viola pedagogy. *Biofeedback and Self-Regulation, 9*, 161–168.

Lockwood, A. H. (1989). Medical problems of musicians. *New England Journal of Medicine, 321*, 51–53.

Lubar, J. F., & Shouse, M. N. (1976). EEG and behavioural changes in a hyperkinetic child concurrent with training of the sensorimotor rhythm (SMR): A preliminary report. *Biofeedback and Self-Regulation, 3*, 293–306.

Lubar, J. O., & Lubar, J. F. (1984). Electroencephalographic biofeedback of SMR and beta for treatment of attention deficit disorders in a clinical setting. *Biofeedback and Self-Regulation, 9*, 1–23.

Lynch, J. J., Paskewitz, D. A., & Orne, M. T. (1974). Some factors in the feedback control of human alpha rhythm. *Psychosomatic Medicine, 36*, 399–410.

Miller, N. E., & DiCara, L. (1967). Instrumental learning of heart rate changes in curarized rats: Shaping and specificity to discriminative stimulus. *Journal of Comparative and Physiological Psychology, 63*, 12–19.

Monastra, V. J., Monastra, D. M., & George, S. (2002). The effects of stimulant therapy, EEG biofeedback, and parenting style on the primary symptoms of attention-deficit/hyperactivity disorder. *Applied Psychophysiology and Biofeedback, 27*, 231–249.

Montes, R., Bedmar, M., & Martin, M. S. (1993). EMG biofeedback of the abductor pollicis brevis in piano performance. *Biofeedback and Self-Regulation, 18*, 67–77.

Morasky, R. L., Reynolds, C., & Clarke, G. (1981). Using biofeedback to reduce left arm extensor EMG of string players during musical performance. *Biofeedback and Self-Regulation, 6*, 565–572.

Morasky, R. L., Reynolds, C., & Sowell, L. E. (1983). Generalization of lowered EMG levels during musical performance following biofeedback training. *Biofeedback and Self-Regulation, 8*, 207–216.

Neurocybernetics (1998). *Biofeedback Program User's Manual* (Version 2.3). Canoga Park, CA: EEG Spectrum International, Inc.

Nunez, P. L. (1995a). Mind, brain, and electroencephalography. In P. L. Nunez (Ed.), *Neocortical Dynamics and Human EEG Rhythms* (pp. 133–194). Oxford: Oxford University Press.

Nunez, P. L. (Ed.). (1995b). *Neocortical Dynamics and Human EEG Rhythms.* Oxford: Oxford University Press.

Passini, F. T., Watson, C. B., Dehnel, L., Herder, J., & Watkins, B. (1977). Alpha wave biofeedback therapy in alcoholics. *Journal of Clinical Psychology, 33*, 292–299.

Patel, C. H. (1973). Yoga and biofeedback in the management of hypertension. *Lancet, 2*, 1053–1055.

Peniston, E. G., & Kulkosky, P. J. (1989). Alpha-theta brainwave training and beta endorphin levels in alcoholics. *Alcoholism: Clinical and Experimental Results, 13*, 271–279.

Peniston, E. G., & Kulkosky, P. J. (1990). Alcoholic personality and alpha-theta brainwave training. *Medical Psychotherapy, 3*, 37–55.

Peniston, E. G., & Kulkosky, P. J. (1991). Alpha-theta brainwave neurofeedback for Vietnam veterans with combat-related post-traumatic stress disorder. *Medical Psychotherapy, 4*, 47–60.

Peniston, E. G., & Kulkosky, P. J. (1999). Neurofeedback in the treatment of addictive disorders. In J. R. Evans & A. Abarbanel (Eds.), *Introduction to Quantitative EEG and Neurofeedback* (pp. 157–179). London: Academic Press.

Plotkin, W. B., & Rice, K. M. (1981). Biofeedback as a placebo: Anxiety reduction facilitated by training in either suppression or enhancement of alpha brainwaves. *Journal of Consulting and Clinical Psychology, 49*, 590–596.

Pollock, V. E., & Schneider, L. (1990). Quantitative, waking EEG research on depression. *Biological Psychiatry, 27*, 757–780.

Pollock, V. E., Volavka, J., Goodwin, D. W., Mednick, S. A., Gabrielli, W. F., Knop, J., & Schulsinger, F. (1983). The EEG after alcohol in men at risk for alcoholism. *Archives of General Psychiatry, 40*, 857–864.

Pressner, J. A., & Savitsky, J. C. (1977). Effect of contingent and noncontingent feedback and subject expectancies on electroencephalogram biofeedback training. *Journal of Consulting and Clinical Psychology, 45*, 713–714.

Reinking, R. H., & Hutchins, D. (1981). Follow-up to: "Tension headaches: What form of therapy is most effective?". *Biofeedback and Self-Regulation, 6*, 57–62.

Rossiter, T. R., & LaVaque, T. J. (1995). A comparison of EEG biofeedback and psychostimulants in treating attention deficit hyperactivity disorders. *Journal of Neurotherapy, 1*, 48–59.

Roth, S. R., Sterman, M. B., & Clemente, C. C. (1967). Comparison of EEG correlates of reinforcement, internal inhibition, and sleep. *Electroencephalography and Clinical Neurophysiology, 23*, 509–520.

Sarnthein, J., Petsche, H., Rappelsberger, P., Shaw, G. L., & von Stein, A. (1998). Synchronization between prefrontal and posterior association cortex during human working memory. *Proceedings of the National Academy of Sciences of the United States of America, 95*, 7092–7096.

Saxby, E., & Peniston, E. G. (1995). Alpha-theta brainwave neurofeedback training: An effective treatment for male and female alcoholics with depressive symptoms. *Journal of Clinical Psychology, 51*, 685–693.

Shouse, M. N., & Lubar, J. F. (1979). Operant conditioning of EEG rhythms and ritalin in the treatment of hyperkinesis. *Biofeedback and Self-Regulation, 4*, 299–312.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory (Form Y1)*. Palo Alto, CA: Consulting Psychologists Press.

Steptoe, A. (1989). Stress, coping and stage fright in professional musicians. *Psychology of Music, 17*, 3–11.

Sterman, M. B. (1996). Physiological origins and functional correlates of EEG rhythmic activities: Implications for self-regulation. *Biofeedback and Self-Regulation, 21*, 3–33.

Sterman, M. B. (2000). Basic concepts and clinical findings in the treatment of seizure disorders with EEG operant conditioning. *Clinical Electroencephalography, 31*, 45–55.

Sterman, M. B., & Friar, L. (1972). Suppression of seizures in an epileptic following sensorimotor EEG feedback training. *Electroencephalography and Clinical Neurophysiology, 33*, 89–95.

Sterman, M. B., & MacDonald, L. R. (1978). Effects of central cortical EEG feedback training on incidence of poorly controlled seizures. *Epilepsia, 19*, 207–222.

Sterman, M. B., & Wyrwicka, W. (1967). EEG correlates of sleep: Evidence for separate forebrain substrates. *Brain Research, 6*, 143–163.

Sterman, M. B., Wyrwicka, W., & Roth, S. R. (1969). Electrophysiological correlates and neural substrates of alimentary behavior in the cat. *Annals of the New York Academy of Sciences, 157*, 723–739.

Sterman, M. B., MacDonald, L. R., & Stone, R. K. (1974). Biofeedback training of the sensorimotor electroencephalogram rhythm in man: Effects on epilepsy. *Epilepsia, 15*, 395–416.

Sterman, M. B., Goodman, S. J., & Kovalesky, R. A. (1978). Effects of sensorimotor EEG feedback training on seizure susceptibility in the rhesus monkey. *Experimental Neurology, 62*, 735–747.

Stoyva, J. M., & Budzynski, T. H. (1993). Biofeedback methods in the treatment of anxiety and stress disorders. In P. M. Lehrer & R. L. Woolfolk (Eds.), *Principles and Practice of Stress Management.* London: Guilford Press.

Taub, E., & Rosenfeld, J. P. (1994). Is alpha/theta training the effective component of the alpha/theta therapy package for the treatment of alcoholism? *Biofeedback, 22,* 12–14.

Thompson, S., & Williamon, A. (2003). Evaluating evaluation: Musical performance assessment as a research tool. *Music Perception, 21,* 21–41.

Vaitl, D., Birbaumer, N., Gruzelier, J., Jamieson, G., Kotchoubey, B., Kübler, A., Lehmann, D., Miltner, W. H. R., Ott, U., Pütz, P., Sammer, G., Strauch, I., Strehl, U., Wackermann, J., Weiss, T. (2003). *Psychobiology of altered states of consciousness.* Manuscript submitted for publication.

Vernon, D., Egner, T., Cooper, N., Compton, T., Neilands, C., Sheri, A., & Gruzelier, J. H. (2003). The effect of training distinct neurofeedback protocols on aspects of cognitive performance. *International Journal of Psychophysiology, 47,* 75–85.

Williamon, A., & Valentine, E. (2000). Quantity and quality of musical practice as predictors of performance quality. *British Journal of Psychology, 91,* 353–376.

Williamon, A., & Valentine, E. (2002). The role of retrieval structures in memorizing music. *Cognitive Psychology, 44,* 1–32.

Williams, J. D., & Gruzelier, J. H. (2001). Differentiation of hypnosis and relaxation by analysis of narrow band theta and alpha frequencies. *International Journal of Clinical and Experimental Hypnosis, 49,* 185–286.

Winson, J. (2002). The meaning of dreams. *The Hidden Mind: Scientific American,* special edition, 54–61.

Wyrwicka, W., & Sterman, M. B. (1968). Instrumental conditioning of sensorimotor cortex EEG spindles in the waking cat. *Physiology and Behavior, 3,* 703–707.

CHAPTER 12

# MENTAL SKILLS TRAINING

CHRISTOPHER CONNOLLY AND AARON WILLIAMON

For decades, elite athletes have enhanced their performance by exploiting advancements in applied psychology and science. An important constituent of the training received by these performers has focused on the interplay between mental and physical skills. Indeed, any witness to elite sporting achievement can appreciate that physical exactitude and excellence go hand in hand with mental agility, determination, focus, and an unimpeded vision of the performance outcome. Thus, the potency and worth of pursuing holistic body–mind training approaches for developing both physical and mental skills have been taken as self-evident by sports trainers and coaches.

Clearly, the same argument can be made for musical performance. A concert soloist performing Dvořák's Cello Concerto, for example, will need to be in command of a number of psychological features, not least among these are the ability to fuse his or her musical ideas with those of the conductor and other ensemble members, to communicate effectively that combined vision to the audience, and to manage any excess anxiety that may accompany the event. In music, as with sports, it is very often the case that the skills needed to master such demands are accrued through years of hard work, training, experience, and self-reflection. This chapter aims to offer performers a range of musically validated techniques for honing their mental skills in order to perform at the peak of their personal abilities. The chapter emerges from work with conservatory students as part of the multidisciplinary project "Zoning In: Motivating the Musical Mind," which examined the efficacy of several interventions for enhancing musical performance (see also Chapters 9, 10, and 11). Considering the wide applicability of mental skills training (also referred to as "cognitive-behavioral" training) to other performance domains, it was hypothesized that particular techniques within the cognitive-behavioral repertory would be of use to highly skilled musicians. The objective of the mental skills program, therefore, was to design and pilot a curriculum of training derived from applied sports psychology. Within the program, a total of 58 students took part in both group and individual training sessions.

This chapter lays the foundation for mental skills training in music, first, through introductions to relaxation and mental rehearsal techniques. It then builds upon this by examining how mental skills can be used to prepare for and meet the demands of specific performance situations and to set and achieve long-term performance goals. Exercises found to be most effective by the participating conservatory students are provided as a guide for how mental skills can be practiced, and readers are strongly encouraged to make note of how the techniques can be employed toward their own musical and personal development. It is important to stress that mental skills should

be learned and developed in the same way as any musical skill—that is, they should be practiced regularly. People who get the most from this training find ways of integrating techniques into their everyday lives, alongside and in ways complementary to physical engagement in the domain.

## 12.1 Relaxation

"Arousal" is the degree of activation in the body, and as is described in Chapter 1, level of arousal can have an impact on performance preparation and quality. A relaxed state is characterized by an absence of unnecessary activity and tension; it is a period of stillness, in which the need for activity or any sense of deficiency is subjugated or at least interrupted for a period. In terms musical performance, relaxation techniques can be used (1) to manage levels of *over*arousal that can interfere with pre-, during-, and post-performance functioning and (2) to control general life and occupational stress that can detract from physical and psychological well-being.

Firstly, a number of studies in music have produced notable results when using cognitive-behavioral interventions to reduce the "stage fright" typically associated with physical, mental, and emotional arousal (Kendrick *et al.*, 1982; Sweeney & Horan, 1982; Nagel *et al.*, 1989; Clark & Agras, 1991). Although the degree of arousal appropriate for performance will vary between individuals and with the particular performance occasion, relaxation training can help instill an awareness of bodily and mental states. This awareness can then be used, for example, to reduce muscle tension that interferes with the coordination of physical skills or to quiet distracting, rational "chattering" in the mind. In this sense, it augments focus on the task at hand. Of course, degrees of relaxation and tension exist on a continuum of arousal, but when used as part of a pre-performance routine, relaxation techniques can and should be paired with mental alertness. Achieving a relaxed, lethargic state will be of no benefit to the performer, particularly as part of an overall performance cycle; therefore, relaxation techniques that are to be implemented before a performance should be practiced and tested well in advance of any crucial performance.

Secondly, relaxation techniques also offer broad, lifestyle benefits, which may then impact on practice and performance activities. For instance, they can regenerate the body, mind, and emotions through the induction of a positive state of being; they can be used to enhance learning and recall; and they can improve sleep and energy conservation (see Woolfolk *et al.*, 1976). In fact, research suggests that skills are best learned in a relaxed state, and by interspersing periods of learning with relaxation sessions, more can be absorbed without deterioration of learning (Benson, 1975).

Relaxation techniques tend to be either physical (e.g. the reduction of tension in muscles) or mental (e.g. the visualization of a natural, quiet scene). Both can be used to reach either a deep, sustained level of relaxation or a momentary, less profound state; the depth to which one relaxes will depend on the time and focus devoted to the exercise. Practice of deeper methods of relaxation is best carried out after, rather than before, a musical activity (e.g. after an intensive practice session) and more generally as part of an overall daily routine (e.g. before an evening meal and going to sleep, when they can be particularly beneficial for combating fatigue and excess worry). Momentary relaxation can be used in relation to musical performance in a number of ways: before

warming up on the instrument (to determine a sense of physical, mental, and emotional state), in gaps during practice and performance (to reduce unnecessary tension and refocus thoughts on the task), as part of warming down (to reduce the build up of tension and re-establish a balanced physical state), and before practicing any form of mental rehearsal (see Section 12.2).

**Exercise 12.1:** Introduction to relaxation

The following steps can be used as a guide to relaxation. As with all exercises presented in this chapter, you may want to have someone read these points aloud to you at first or tape record them, so that you are then free to experience the effects without having to remember each step. With time and practice, the steps will come naturally.

- Sit down. Put your feet flat on the floor, with your hands placed easily on your lap. Close your eyes.
- First, pay attention to the sounds outside of the room; then, within the room; then, notice your own thoughts.
- Pay attention to the contact you make with the chair and floor, and notice your breathing.
- Take a deep breath in for four counts; hold it for four, and let it out slowly to the count of either four or eight.
- As you continue to breathe, easily pay attention to the muscles around your face: your forehead, the muscles around your eyes, cheeks, mouth, jaw, and neck. As you exhale, imagine all tension draining away through your shoulders and on through your arms, wrists, hands, and fingers.
- Notice how your chest relaxes, as well as your back, and your stomach, waist, and pelvis.
- Pay attention to your thighs, and notice any tension you may find in them. On your next outbreath, imagine all of the tension draining away through your knees, calves, ankles, feet, and into the floor, leaving you calm, relaxed, and alert.
- Remain in this state for a few minutes before taking three deep breaths, stretching, and opening your eyes.

How we breathe has a considerable impact on our state of physical, mental, and emotional arousal. Deep-centered breathing from the abdomen can be relaxing and energizing. Stopping to take a few slow, steady breaths can help to center focus on the here-and-now and to check mental state. The following exercise can be used to release tension just before the start of a performance or in short gaps between successive performance events.

**Exercise 12.2:** Breathing through the body

For this exercise, you will be breathing in a similar way as in the exercise above, relaxing three areas of your body with each exhale. It is worth noting that one physiological symptom of performance anxiety is hyperventilation, which is an abnormally increased pulmonary ventilation. Therefore, in high-anxious situations, it is advisable to clear the lungs of air *before* taking in deep breaths.

**Exercise 12.2** (*Continued*)

- Start breathing from the center of your body (just behind your navel). Draw in air slowly through your nose for a count of four. As you inhale this long, slow breath, allow the air to fill your belly out, like a balloon filling with air.
- Hold your breath for a count of four.
- Now, slowly exhale for a count of eight, and as you do, release any tension from your forehead, jaw, and neck through your shoulders and arms and on through to your fingers.
- Inhale for four and hold for four in the same way.
- Slowly exhale for a count of eight, releasing any tension from your chest and stomach and relaxing all of your vertebrae down your spine.
- Inhale for four and hold for four in the same way.
- And now, slowly exhale for a count of eight, releasing any tension from your hips and thighs through your lower leg and ankle, all the way to your toes.

Reflections on relaxation

- When you were relaxing, what physical sensations did you notice? What did you notice about your thoughts?
- What did you feel during the relaxation, and what do you feel now?
- Where and when might you practice this technique? Why?

## 12.2 Mental rehearsal

### 12.2.1 Introduction to mental rehearsal

Mental rehearsal is the cognitive or imaginary rehearsal of a physical skill without overt muscular movement. The basic idea is that the senses—predominantly aural, visual, and kinesthetic for the musician—should be used to create or recreate an experience that is similar to a given physical event. Everyone has the capacity to rehearse mentally, and in making the most of this, musicians would be well served when developing their mental rehearsal (as recommended in Chapter 5) to activate their analysis and ear training skills, their abilities to notice and attend to movements and touch, their knowledge of style and history of music, and their powers of reflection and memorization.

Some inconsistency has emerged within the music research literature regarding the extent to which mental rehearsal can positively impact performance skills (e.g. Rubin-Rabson, 1937, 1941; LaBerge, 1981; Ross, 1985; Coffman, 1990; Jones, 1990; Kopiez, 1990a, 1990b, 1991). This is not surprising, however, given the wide disparity in the methods employed in the research, including core differences in the groups of musicians recruited to take part (i.e. their level of skill and instrument specialization), the procedure implemented (i.e. the amount of time given for mental rehearsal, the selection and length of the pieces to be rehearsed, and the purpose of the rehearsal itself), and the outcome measures of the performance skill. What is clear from the

extant literature is that musicians themselves vary considerably in their use of mental rehearsal (see Grøndahl, 1987). It seems also that the points of most practical significance are that (1) informed physical practice at the highest levels of musicianship can hardly take place without some sort of cognitive or mental activity and (2) only through committed, personal effort can the musician expand, differentiate, and fully exploit his or her repertory of mental strategies.

With regard to the former point, evidence suggests that mental and physical practice are inextricably linked on a psychophysiological level. Hale (cited in Freymuth, 1993) found that internal imagery produced the same muscle activity that would have been used in the actual movements (see further evidence provided by Stippich *et al.*, 2002). In addition, Bird and Wilson (1988) studied electroencephalographic (EEG) and electromyographic (EMG) data obtained from student conductors and their teacher during imagery. The results showed that, despite large individual differences, the teacher and the more skilled students produced more repeatable EEG patterns than the less skilled students. In addition, the teacher displayed EMG patterns during mental rehearsal that were similar to those of the actual performance. Given that conductors are by necessity forced to engage in much mental rehearsal, these findings suggest that, with practice, mental rehearsal can indeed be an effective supplement to physical rehearsal. As such, it can then be integrated into the musician's practice routine so as to avoid the musculoskeletal problems that may arise from overpracticing (see Chapter 3 for further discussion).

Findings presented by Rubin-Rabson (1937, 1941), Ross (1985), and Coffman (1990) all provide support for the notion that combining physical and mental practice can be favorable for musical learning. Gabrielsson (1999) concludes, however, that "it seems the less advanced the person is on the instrument and the more difficult the music is, the more important is the motor practice" (p. 507). But what exactly can musicians gain from integrating mental rehearsal into their regular practicing routines? Interview data with the conservatory students mentioned above suggest that mental rehearsal can be used as follows:

- To improve learning and memory

  A horn player: "I had to play pieces from memory for my final [recital]. I found that really difficult because I've never done anything from memory before. I spent an awful lot of time playing the music through my head, without it in front of me, playing it through, seeing what it felt like."

- To make practice more efficient

  A clarinetist: "My fundamental practicing has been a lot more focused. I achieve things in a shorter amount of time . . . because I make sure I am doing it properly."

- To overcome technical difficulties and develop skills

  A pianist: "Mental rehearsal helps with body/mind awareness, which helps correct technical difficulties in pieces; rather than go over and over, I try and fix it the first time."

- To heighten sensory awareness

  A cellist: "I am more aware of my sound, more aware of what I am doing, more aware of what I can do. . . . I experiment . . . and have clarity of thought for executing something difficult."

- To gain more interest in the music itself

  An oboist: "Mental practice made me aware of other things I wanted to check out about the piece. . . . When you start thinking about playing the music rather than just doing it, it helps."

- To refocus attention during performance

  A singer: "I use quick images and color when I notice my attention drift during a performance."

- To enhance general confidence and resilience on stage

  A violinist: "I think it's the self-confidence. The confidence in my own performance. . . . You have rehearsals before a concert, and often that's just to remind yourself how it all goes and how it works with other instruments. . . . That's what I'm basically doing, but I'm just cutting out the actual, physical side of it—doing it mentally—which seems to work just as well."

- To achieve greater control over negative emotions

  A horn player: "By imagining myself performing, I think it takes away the fear of the unknown—unknown places, unknown circumstances."

- To establish a greater connection and presence with an audience

  A flute player: "I'm closer to the music, I suppose, than I have ever been before and hopefully, therefore, the audience is too. They're closer to what they've come to experience. . . . I link images with the music, and this helps me to project my feelings and the meaning of the music."

- To achieve peak experience.

  A violinist: "I had the first real performance after last year's course of mental skills on Thursday of last week. There were 800 people in the audience, and I was being sponsored to play. So, it was quite important. It was lovely because I went on, and the pieces were internalized from when I learned them last summer, so it was a question of bringing them out. For the first time, I forgot I was playing—not forgot, although I was communicating with the audience, I detached myself from them and concentrated on the music and let the music bridge, rather than myself playing a violin and looking at the audience. It was the music. I looked around while I was playing, and I felt very easy. . . . The high afterwards was great."

Plainly, the above benefits may not apply universally to all musicians. In fact, specific benefits are likely to vary depending, for instance, on the individual's current technical skill, personal preferences for learning and self-reflection, the conditions under which a particular piece is to be performed, and most importantly, how and the extent to which mental rehearsal techniques are practiced. Ultimately, it is not a question of adopting either a mental or physical approach, but rather how to make the most of both approaches. The two simply are not mutually exclusive at the highest levels of performance.

Research on motor skill acquisition has identified conditions and situations in which mental rehearsal is optimally effective. For example, the person should have prior experience in executing the task (or one similar to it). Also, mental rehearsal can be particularly effective during the early stages of learning, when novel insights and ideas about the task can be formulated, and during the later stages to reinforce cultivated performance strategies. Furthermore, the performer should endeavor to imagine

responses in muscles that would actually perform the movement (Weinberg, 1982). In addition, years of hands-on experience in the field of sports psychology have led to several guiding principles for mental rehearsal (see Loehr, 1987; Martens, 1987; Syer & Connolly, 1991, 1998; Butler, 1996), which include the following:

- Practice regularly, especially in the morning (May *et al.*, 1993; Yoon *et al.*, 2000; Hasher *et al.*, 2002).
- It is better to carry out short, regular mental rehearsal sessions (e.g. of 5–10 minutes) than long, infrequent sessions (Weinberg, 1982).
- Start with relaxation, so that clear signals can be communicated between mind and body. You need to be relaxed and alert at the same time.
- Mentally rehearse specific skills or qualities you are working on in your technical training, above and close to your current level of performance.
- Be positive; move toward what you want to focus on, not on aspects that do not contribute directly to improving your performance. Stay in control of what you rehearse.
- Use all of your senses, so that you believe that you are actually in the situation executing the skill. Keep working to improve the clarity and vividness of mental images; include emotions and feelings.
- Notice how you visualize. Do you observe yourself performing as if on a screen (externally), or do you see, hear, and feel from within your body (internally)? Both strategies can be useful (see Section 12.3.2 below). In general, when you are correcting skills, it is helpful to observe externally (from the outside) until the skill is corrected, and then to experience it from an internal perspective, noticing the physical sensations and feelings associated with the skill.

In short, to use mental rehearsal successfully for performance enhancement, relaxation should come first, focus should be intently on the task at hand, the experience should be realistic, and it should be practiced regularly.

**Exercise 12.3:** Developing imagery skills

Mental rehearsal for performance enhancement is not just visual. Rather, it involves using all of the senses. This introductory exercise is intended to develop such imagery.

- Before visualizing, take a moment to relax. Let yourself imagine, one after another, the following sensory experiences:
- Now *see*: A sunset over the ocean, white clouds racing over the sky, your musical instrument, the face of a friend, a building that you like, a rose as it opens and blooms. . . .
- Now *hear*: The sound of a rainstorm on a tin roof, church bells ringing in the distance, the applause of the audience at the end of a concert, the sound of wind in the trees, your favorite piece of music, the voice of a famous singer. . . .

**Exercise 12.3** (*Continued*)

- Now *feel*: The sun on your back on a hot day, jumping into a cold bath or pool, the grasp of a firm handshake, walking barefoot on grass, your fingers moving when you play your instrument. . . .
- Now *smell*: Your favorite perfume, a newly mowed lawn, burnt toast, a new musical instrument. . . .
- And now *taste*: A lemon as you bite into it, a cool refreshing drink, your favorite food. . . .
- Let the scene fade, slowly stretch, and open your eyes.

**Exercise 12.4:** Developing control of your mental rehearsal

It is important that you have control over your experience during mental rehearsal. The following exercise can be helpful in developing this.

- Sit down in a place where you will not be disturbed. Uncross your legs and arms. Close your eyes, and relax from your head downward.
- In your imagination, walk into a practice room. Notice what the room looks like.
- See your instrument (or any object generally related to your performance, such as a music stand or a page of a score). Notice details in what you see.
- Walk around it. Notice what you can see and hear as you do this.
- Look at the instrument from above, and then from below.
- Now, imagine your instrument growing and make it larger. Then, see it getting smaller, first back to its normal size and then smaller still.
- Now, change its color. What color is it?
- Take a breath, and go back to the beginning. Walk into the room again.
- See your instrument again. What details do you notice now?
- Reach toward your instrument and touch it. Notice what it feels like, its texture. Either pick it up or put yourself into a position ready to play it.
- Notice how you make contact with the instrument. Take time to "feel right," and stay with this feeling for a moment.
- Begin to play and listen. What can you hear?
- Stop playing. Breathe.
- Begin to play again and listen.
- Stop playing. Breathe. Do this a few more times, but when your attention begins to waver or you begin to lose control of your imagery, finish the practice.
- Stop playing, put the instrument down, and leave the room.
- Let the scene fade, and slowly bring your attention back to your present environment.

### Exercise 12.5: Skill development mental rehearsal

Skill development mental rehearsal complements the physical practice of a skill. Adding this process to your weekly training can increase confidence and speed the learning process. You may also develop a greater awareness of what you are doing. Choose a specific skill or piece of music you want to practice. You may want to choose a piece that you need to memorize, are having technical difficulties with, or want to practice communicating to an audience.

- Before visualizing, take a moment to relax as you have practiced previously. Notice how the more you train yourself to relax, the easier it becomes.
- See yourself in a place where you can practice or perform the skill you want to improve. Create as real and as genuine a scene as possible, incorporating information from all your senses.
- Now, see yourself approaching your instrument, and when you have reached it, prepare to play. Find the place that feels right, as you did before. Notice what you do to prepare yourself.
- Begin to play. Notice what you can see and hear as you observe yourself. How do you use your arms and head? How is your body positioned? How do you use your legs and feet? Notice as much detail as possible.
- Take a deep breath, and now watch the same scene from somewhere else—a different viewpoint, closer or further away. Again, pay attention to all the details. Listen to the music, and take note of what you can see.
- Rewind the image to the beginning, and this time step into the image of yourself in the scene. You are now inside your body looking through your own eyes. As you play, notice what you can see and hear, and notice physical sensations and feelings. Notice what parts of your body seem most alive. Notice where you are relaxed and where you are tense. Be aware of all the sensations and feelings you identify with this practice. Notice what your whole body feels like when you are playing.
- Rewind and repeat this one more time, this time focusing on different parts of your body or different aspects of your technique.
- Stop the practice. Take a breath, move away from your instrument, and leave the room.
- Let the scene fade, and slowly bring your attention back to your present environment.

**Reflections on mental rehearsal**

- What did you discover about using all your senses?
- Which sense was the easiest to evoke? Which was the most difficult? What happened when you explored your physical sensations and feelings?
- What did you notice about your ability to control your visualization?
- What did you see and hear yourself doing when you observed yourself playing?
- What did you notice when you looked from a different viewpoint?
- What skills do you want to improve using mental rehearsal?

### 12.2.2 Evocative symbols and images

Images, sounds, symbols, and words can evoke certain feelings and physiological changes in us. Indeed, music teachers have long used metaphors to suggest new musical ideas and to encourage students to express music in particular ways. In one study of 135 music students, 81% at one point had been taught to express music through this method, and 42% ranked it as most effective, compared with "felt emotion" (preferred by 39%) and "aural modeling" (preferred by 25%; Lindström *et al.*, 2003). The aim of mental skills training in relation to evocative symbols and images is to develop within performers the ability to exploit symbols and images for themselves and, subsequently, to make the mental rehearsal in which they engage all the richer and more detailed. To do so, it is helpful to be aware of images, stories, and sounds that may naturally emerge when practicing and performing. As such, one's intuition, imagination, and emotion are stimulated.

**Exercise 12.6:** Developing qualities in performance

For this exercise, you will be playing through a piece of music in your mind.

- Start by using a relaxation exercise to prepare yourself for mental rehearsal.
- Imagine that you are about to perform a piece of music. Look around you, and notice what you can see and hear.
- When you are ready, begin playing. Pay attention to your physical sensations as you play. Play everything perfectly. Notice the sound you are creating and how you feel when you hear the music played in this way.
- When you are fully immersed in the music and the experience, take a deep breath. As you exhale, notice what image the music evokes in you. Notice what the image, metaphor, symbol, or object is. Stay with the image, and let it continue to get clearer in your mind.
- Now, continue to play the passage while maintaining the image in your mind.
- As you finish the piece, notice what you are feeling.
- Gradually let the scene fade, and when you are ready, open your eyes.

Over the coming week, imagine this image or symbol, and in your imagination, play the piece with this quality. As soon as you are able, evoke this same quality by imagining the image or symbol before playing the piece physically.

**Reflections on developing qualities in performance**

- What image or symbol was evoked when you expressed the music?
- How will you use this symbol or image in practice and performance?

**Exercise 12.7:** Acting "as if"

Acting "as if" is like acting a part, performing as if you were someone or something else that represents a quality you want to emphasize in your performance. The someone or something will be very specific to you.

- Before starting the mental practice, identify someone who expresses a performance quality that you would like to emulate.
- Use a relaxation exercise that you find works best for you.
- Observe this person in a specific situation where he or she expresses this quality. Step into the situation yourself and imagine being this person in this situation. Notice physical sensations, and notice what it feels like to be the person.
- Keep this image in your mind, and imagine playing a piece of music you want to perform. Notice what you see, hear, and feel subjectively when you play like this person with the qualities that they possess.
- And as you finish the piece, let the scene fade. When you are ready, bring your attention back to your present environment.

Over the coming week, take the best qualities you discover about yourself or the "model performer" in this exercise, and play the piece with these qualities in your imagination. Before you physically practice the piece, picture this image and allow it to evoke the quality you want to express in the music. In practice, you may want to try playing the piece as if you are this person. Many people find this exercise can be very effective. Some, however, prefer to use only themselves as a model. This can be done equally well, and you may want to model yourself on one of your own "peak performances" when you played at your very best.

**Reflections on acting "as if"**

- What did you notice when you acted the part of the other person?
- What quality of performance was evoked?
- What did you notice about your performance when you played the music with this quality?

## 12.3 Using and developing mental skills for specific performance situations

### 12.3.1 The pre-performance routine

Partington (1995) found that highly skilled musicians give a central role to physical and mental strategies that enable them to feel confident and ready to perform. During the one or two days before a performance, he found that the performers deeply immersed themselves in musical, physical, and mental preparation aimed at allowing them to perform at their best. They developed preparation routines through personal experience. On the day of the performance, they displayed a wide range of individual differences in preparation. Technical and physical readying included staying connected with the music through easy practicing, not changing interpretation, and ensuring adequate sleep and nutrition. To prepare themselves mentally, some strove to feel excited and ready to perform, yet not too anxious. To achieve this, some simply "took it easy;" others used relaxation techniques or positive imagery; and some engaged in constructive internal dialogue and reasoning. The final stages of preparation (i.e. the few hours before performance) were also highly individualized. To avoid anxiety associated with waiting, some arrived only 15 minutes before playing. Others arrived 1–2 hours beforehand and worked through a systematic physical and mental warm-up that

sometimes consisted of reviewing the score through audiation, engaging in an internal dialogue to achieve an appropriate orientation to play, or completing a predetermined warm-up on their instrument. In the minutes just before going on stage, some of the musicians attempted to create a mood in themselves appropriate for their intended expression of the music, while others struggled to isolate themselves or remain relaxed through easy playing.

Despite the clear individual differences emerging from Partington's account, it was clear that each player made a systematic, patient effort to nurture and maintain a state optimal for concert performance. Two key factors seemed to apply in facilitating this process: (1) a long-term commitment to high-quality physical, technical, and artistic preparation and (2) the development of an individualized, flexible pre-performance routine that includes physical activity, nutrition, and rest, as well as warming up using mental, emotional, technical, and musical strategies.

**Exercise 12.8:** Optimizing performance preparation and review

The following will help you optimize your performance preparation over time and will encourage you to incorporate mental and emotional strategies, as well as physical.

- Think back to your last important performance, and write down answers to the questions that follow:
- How did you prepare physically, emotionally, and mentally?
- What worked well?
- What, if anything, do you want to change?
- What else could you have done?

Once you have answered these questions, think back to how you *reviewed* the success of that performance and write down answers to the following questions:

- How did you review your performance physically, emotionally, and mentally?
- What worked well?
- What, if anything, do you want to change?
- What else could you have done?

**Exercise 12.9:** Developing consistency in preparation

Consistency in preparation leads to consistency in performance. Segment your preparation on concert day into four sections: (1) the morning of the performance day, (2) the afternoon of the performance day, (3) arrival at the performance venue, (4) 10 minutes before performing. Within each section, consider what you are doing to prepare technically, musically, mentally, and emotionally. It is vital that you be specific.

In addition, when warming up, you may find it is useful to pay attention and tune into five different factors that can affect your performance:

- The place where you are to perform (identify supportive factors and distractions in the environment; make yourself at home).
- Yourself (what are your physical sensations, thoughts, and feelings?).
- The people you are with, especially others with whom you are going to perform.
- Your technical and artistic intentions.
- The way in which you will meet and convey these intentions.

It is best to start with a simple routine and use trial and error to try out variations on this in situations where you are able to experiment. Begin to integrate the results of your experimentation into your preparation for actual performances when you feel ready. Importantly, ensure that there is sufficient flexibility within your routine so that you can adjust it to changing external circumstances (adapted from Emmons & Thomas, 1998).

### 12.3.2 Concentration

It is common to find musicians asking themselves why they have difficulty focusing in high-pressure performance situations. Being able to concentrate and "shift" one's focus of attention is a salient component of musical expertise (Williamon *et al.*, 2002), which is typically acquired over years of practice and training (see Chapter 2 for an extensive discussion). Still, to make the most of this training in intense performing environments, it can be helpful for performers to understand some mechanics of concentration.

Concentration can differ and vary in duration and intensity. High concentration is an unwavering awareness of a specific subject to the momentary exclusion of other subjects. For musical performance, such concentration should ideally take the form of a relaxed state of being alert. Moreover, the performer should be mentally agile enough to change focus rapidly and smoothly in conjunction with the flow of relevant factors and external events. In work with athletes, Nideffer (1976) developed a model that is helpful in becoming aware of how to direct attention before, during, and after a performance. The model recognizes that concentration can be broken down into two key dimensions: direction (internal or external) and width (broad or narrow). These dimensions, in turn, form four quadrants of concentration (see Figure 12.1).

In terms of concentration training, performers need to maintain awareness of themselves and what is happening around them, allowing only the most relevant factors to come into momentary focus at the expense of all else. Practicing concentration and shifting the focus of attention can help performers become aware of where attention is actually drawn and why it is there, and then to break outdated patterns of undesirable thought or unproductive behaviors. Performers can then become more closely familiar with parts of their performance that are within their control, and can interrupt or distance themselves from distractions or worries that stem from sources beyond their control.

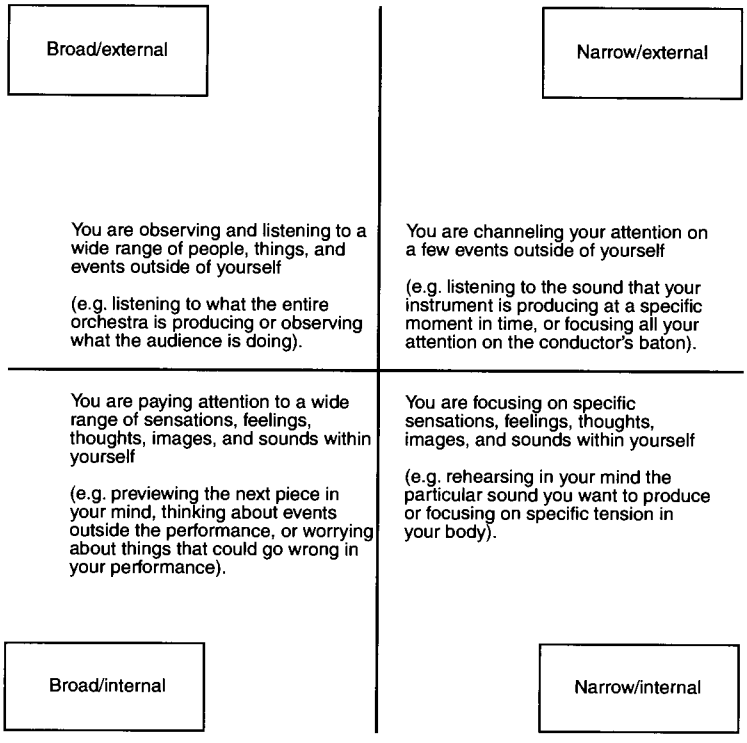



**Figure 12.1** Four quadrants of concentration, composed of the direction dimension (internal or external) and the width dimension (broad or narrow; adapted from Nideffer, 1976).

In terms of when performers should concentrate, an obvious answer is throughout the entire performance event, but many performances offer periods of more and less activity. These can demand continual change in the direction, span, and intensity of concentration. Discovering where best to focus attention during periods of inactivity requires practice. Generally speaking, changing focus between external and internal and between broad and narrow gives the performer a way to review and refocus. When the focus of attention is on distractions that are unhelpful or irrelevant for performance, it is important to redirect attention to task-relevant cues. The ability to do this can be enhanced by identifying occasions or moments when distraction habitually occurs. Once counterproductive patterns in attentional focus have been identified, mental skills training can then be directed toward managing or (when possible) replacing negative habits with more appropriate ones. It is important to mention that when seeking to change behaviors, a negative habit needs to be replaced with a more productive habit: simply stopping a habit without substituting another for it can result in the negative habit eventually reasserting its presence.

### Exercise 12.10: Identifying drifts in attention

- Review your last few performances and notice when and where your attention drifted. Attentional drifts very often result when you are focusing on factors that are unhelpful to your performance. Concentrate on what you need to do, not what to avoid.
- Are you distracted by factors in the environment?
- Are you distracted by past or future experiences?
- Are you distracted by task-irrelevant thoughts?
- What patterns do you notice?

### Exercise 12.11: Concentration and focused breathing

- Breathe from the center of your body.
- Hold your breath for a count of 4.
- Then, for a count of 4, exhale through your mouth, and allow the air to flow out of your belly.
- Repeat for 10 more times. Count the number of breaths you take. See how many times you can count up to 10 without distraction.

Review of focused breathing

- How many times were you able to count to 10 without being distracted?
- Where did your attention drift?
- What helped you maintain your focus on the here and now?

### Exercise 12.12: Shuttling

This is an exercise designed to help you avoid getting stuck in the "middle zone," a term that describes situations in which you are physically doing one thing (e.g. performing a movement of a piece) while mentally doing another (e.g. thinking about how you will be performing a technically difficult passage in the next movement). This preliminary "drill," which should be practiced with a partner, will help you develop your ability to control attention and learn to shift your attention intentionally.

- Close your eyes, and tune in to some sensation, feeling, thought, sound, or image from inside and say "now I am aware of. . . ." For example: "now I am aware of my breathing," or "now I am aware of a pain in my leg," or "now I am feeling silly," etc.
- Then, open your eyes, and say "now I am aware of . . .", adding something from outside, such as "now I am aware of the sunlight" or "now I am aware of your eyes."
- Repeat the process, shuttling your attention inside and out; first an inside statement, and then an outside statement for a few minutes.

**Exercise 12.12** (*Continued*)

- Let your partner do the same.
- Now, try the same exercise again with your eyes open all the time.

**Reflections on shuttling**

- Did you find it easier to shift your attention in or out?
- How aware were you of your internal experience; your external experience?

**Exercise 12.13:** Zooming in and out

When you are anxious, your attention narrows, and your ability to shift attention lessens—that is, your ability to shift attention from either broad to narrow or internal to external is hampered. This can have a negative effect on your performance, particularly if your ability to pick out and pay attention to salient performance cues is reduced. The purpose of this exercise is to practice shifting your attention across all four quadrants shown in Figure 12.1.

Nervous zoom

- Imagine a time when you were very nervous or anxious. Recall the whole experience; notice what you can see, hear, and feel. Step into the experience, only as much as you feel comfortable. Stay with this feeling.
- Pay attention to what is happening inside of you. First, focus on a specific feeling, sensation, thought, image, or sound. Now broaden your internal lens. Keep broadening and narrowing your internal lens, and notice what you become aware of each time.
- Now, shift your attention outside. Zoom your external lens in and out.
- Shuttle your internal and external lenses from broad to narrow and vice versa a few more times.
- Stop, stand up, and shake out the feeling of nervousness.

Confident zoom

- Imagine a time when you were confident and assured. Step into the whole experience, noticing what you see, hear, and feel. Stay with this feeling.
- Pay attention to what is happening inside of you. First, focus on a specific feeling, sensation, thought, image, or sound. Now broaden your internal lens. Keep broadening and narrowing your internal lens, and notice what you become aware of each time.
- Now, shift your attention outside. Zoom your external lens in and out.
- Shuttle your internal and external lenses from broad to narrow and vice versa a few more times. In each, try to achieve a greater flexibility in shifting your attention.

**Reflections on zooming in and out**

- During the nervous zoom, where did your attention switch to most easily? What impact does switching from broad internal to broad external have on your state?
- During the confident zoom, what did you notice about your ability to switch your attention?

**Exercise 12.14:** The present moment technique

This exercise can help you to focus on the here and now—the present moment. It is a simple exercise that takes 30 seconds and can help you refocus *during* a performance.

- For 10 seconds, focus all of your attention externally, and notice the visual or auditory detail of what you are looking at or listening to.
- For 10 seconds, focus all of your attention on your breathing and total body awareness.
- For the last 10 seconds, imagine playing your next phrase perfectly. Breathe deeply, and then play it.

There are several additional strategies that can be adopted that may have a positive effect on enhancing concentration. These include (1) building and practicing a stronger pattern (e.g. adopting the five-point warm-up procedure described in Exercise 12.9), (2) strengthening the attraction of the object or action on which attention must be placed, (3) deliberately paying attention to the distraction (i.e. treating it as if it has a life of its own; discovering its "needs" and making a contract to deal with it later), (4) "making friends" with the distraction (i.e. practicing in an environment where distractions are simulated and/or finding a way to view the distraction positively), and (5) removing the focus from emotional distraction to physical or mental patterns (e.g. if the distraction is emotionally charged anxiety and tension, a physical technique such as Exercise 12.2 could be used to relax). It is important to stress that the selection of the "right" strategy(s) for each musician will inevitably depend on that individual's experimentation and preferences. Ultimately, intense and consistent concentration is most likely to be possible when the performer is well prepared artistically, when he or she has developed clear intentions for the performance.

### 12.3.3 Ideal performance states and simulation

What a musician focuses on during a performance seems to be one of the keys to perceived inspiration (Partington, 1995). Experienced performers concentrate on a range of features when making music, including making a personal statement, conveying beauty, and establishing a rapport with the audience (see Chaffin *et al.*, 2002). They are often aware that the consequences of concentrating on playing each note with technical perfection are generally disappointing and can lead to high anxiety about making an error, disappointment following the error, and a generally cool response from the audience to an accurate (yet unmoving) rendition of a piece. Furthermore, they seem to focus on doing, not on *how* to do what they are doing. The following are characteristics associated with individuals in peak performance states (for further discussion, see Csikszentmihalyi & Csikszentmihalyi, 1988; Csikszentmihalyi, 1990; Syer & Connolly, 1991, 1998):

- energized, yet relaxed (a subtle balance of quiet intensity)
- confident (an expectation of success; lapses in performance do not undermine self-belief)

- focused (absorbed in the moment; focused solely on the task at hand)
- effortless and automatic (mind and body working in unison; little or no conscious thought involved; instincts and intuition "just let the performance happen")
- fun (incomparable enjoyment)
- in control (full control of the situation, no matter what).

The purpose of this section is to assist musicians in attaining peak performance states. The exercises covered previously provide the foundation for this. Relaxation training helps generate a relaxed state, which is vital to peak performance; mental rehearsal helps to fine tune performance and develop skills; a consistent routine before both practice and performance allows for adequate mental, physical, and emotional preparation and warm-up. The exercises below can help the performer use mental rehearsal to re-experience peak performance states both mentally and in actual performance.

**Exercise 12.15:** Top performance mental rehearsal

- First, relax yourself to ensure that your mind, body, and emotions are quiet.
- Imagine yourself walking into a big room, and in front of you, you see a chair and a large screen. Sit down, and watch a film of your best ever performance. Notice what you are doing to prepare yourself, and when you begin playing, notice what the music sounds like. Notice how you are expressing yourself, the positioning of your body, and your facial expression. As you come to the end of your performance, notice what you look and sound like.
- Having just finished watching yourself, what is your response? What are you feeling?
- Rewind the film back to the beginning, and this time, choose another camera angle—perhaps from above or behind—and then play the film through again. What can you see and hear? What is different when you look from this angle? Complete the performance.
- Now, see yourself about to perform again, and this time, imagine yourself getting up from your chair and stepping into the film. You are now about to play from the inside, looking through your own eyes.
- Prepare yourself and play. Give your best performance. Notice where and what you are paying attention to, notice what the music sounds like, and notice how you are expressing yourself. In particular, pay attention to the physical sensations when you play. What are the key aspects of your performance? Complete the performance and notice what you are feeling now.
- Give the performance from inside for a second time. Focus on the key aspects of your performance and hold onto the feelings that are specific to when you play your best. Complete the performance and return to your seat, where you can see yourself on the screen again.
- Watch yourself one more time, performing at your very best, and notice your response. As the performance finishes, notice what you are feeling and bring this feeling back with you to your present environment.

**Reflections on top performance mental rehearsal**

- What did you notice when you observed and heard your performance from the outside?
- What focus and feeling did you associate with your best performance?
- What are you going to do before future performances to evoke this state in yourself?

### Exercise 12.16: Simulating performance

Before you do this visualization, it helps if you have seen the venue in which you are going to be performing.

- Notice the details of the room and what the room sounds like.
- Relax your mind and body in the usual way and now recall the memory of a moment in your musical career when you felt you were in the "right place at the right time"—just one moment, the first that occurs to you. This is a moment when your performance was seemingly flawless, when you seemed to know in advance what was going to happen.
- Notice what you are doing, where you are, and who you are with. Above all, notice your physical sensations and what it feels like to be performing in this way. Play through this moment once or twice more, focusing on both your physical and mental sensations.
- Hold these feelings and imagine you have arrived at the venue for an upcoming performance. Find a place in the venue where you can momentarily store your equipment and feel at home. Walk into the room where you will be playing and notice what you see and hear as you walk around. Perhaps go to back of the room and notice what you can see from the audience's point of view. Now walk onto the stage and look around and listen carefully to the room. Make yourself at home on stage.
- Imagine that you have now completed your warm-up routine, and when you are ready, step out to perform. As you do so, bring the experience and feeling of being in the right place at the right time with you. Notice how you feel as you step out onto the stage and ready yourself. Notice the audience that you are communicating with and the place you are in now. Focus on yourself, and notice what you are feeling. Now play.
- Notice what you can see and hear as you play, and pay attention to your physical sensations and feelings. Play your performance perfectly. For a moment, float out of your body and notice what you can see and hear from the audience's perspective. Notice how you are responding as a member of the audience. Now, float back into yourself and complete your performance. As you finish, notice what and how you feel, and then bring yourself back to your present environment.

**Reflections on simulated performance**

- What did you feel when you were in the right place at the right time?
- When and where will you use this visualization again?

## 12.4 Using mental skills for long-term learning and preparation for performance

As detailed throughout the rest of this book, there are several qualities that performers should incorporate into their long-term performance preparation. These include maintaining high-quality individual and group practice (e.g. the adoption of a systematic practice approach, selecting an appropriate environment for practice with minimal internal and external distractions, and using a warm-up routine that prepares the body and mind to play), technical preparation and learning (e.g. the removal of extraneous and unwanted physical movements and tension), and artistic preparation and expression (e.g. seeking to make a personal statement in interpreting and expressing music and working to communicate that statement to the audience).

Mental skills training can also be used to help performers realize their long-term goals. Essential ingredients to any long-term preparation plan include the ability to (1) assess performance quality, (2) set goals and plan, (3) take action toward meeting those goals, and (4) monitor progress. In terms of the first, it is important to acknowledge one's current performance ability in order to conceive of how performance can be improved (see Chapter 4 for a discussion on how to increase the power of one's "microscope" for assessing performance quality). Second, setting challenging and specific goals can lead to better performances than setting goals merely to "do my best" or not setting goals at all. Goals provide direction, maintain motivation, increase effort, direct attention, and improve learning (Reitman, 1965; Simon, 1973; Wessels, 1982; see also Chapters 2 and 5 of this volume). Third, this action plan must be implemented. Finally, it is important to review set goals regularly in order to update and adjust the plan accordingly. The following exercise can assist long-term preparation; it will take about an hour, and since it is potentially an important step in planning career choices, it should be done with undivided attention.

**Exercise 12.17:** Choosing and prioritizing goals; deciding on actions and scheduling

**Choosing goals**

- To do this, you will need five sheets of paper.
- On the first sheet, write "LIFETIME GOALS." Now, write as many such goals as you can for 5 minutes. Write down anything that comes to your mind.
- Take the second sheet and write "1 TO 3 YEAR GOALS," and do the same.
- Take the third sheet and write "3 TO 12 MONTH GOALS," and do the same.
- Take the forth sheet and write "1 MONTH GOALS," and do the same.

**Prioritizing goals**

- Review each piece of paper in turn and decide on the goals that are most important to you (e.g. those that you are most excited about).

- Choose the three goals that are most important to you and that have a sense of urgency about them. Write "A" beside them.
- Write "B" next to three of the remaining important goals. These goals either do not have the same urgency or are not as important as the "A" goals.
- Now, write "C" next to the remaining goals. There is neither the same importance nor urgency for these, but at some point in the future, you may want to review their relative importance.

Deciding on actions and scheduling

- On the fifth sheet, draw up a table with the following headings in the left-hand column: (1) Specific actions for 1 MONTH "A" goals, (2) Specific actions for 1 MONTH "B" goals, (3) Specific actions for 3 TO 12 MONTH "A" goals, (4) Specific actions for 1 TO 3 YEAR "A" goals, and (5) Specific actions for LIFETIME "A" goals. In the right-hand column, beside each of these five headings, write "when."
- Look at your 1 MONTH goals. For each "A" goal, choose three actions you will take in the coming week. For each "B" goal, choose one action you will take in the coming week. Schedule each item with a date to be completed.
- Look at your 3 TO 12 MONTH goals. For each "A" goal, choose one action you will take in the coming week. Schedule each item with a date to be completed.
- Look at your 1 TO 3 YEAR goals. For each "A" goal, choose one action you will take in the next month. Schedule each item with a date to be completed.
- You now have a number of specific actions to do over the next month. How are you going to remember to pay attention to these actions?
- After 2 weeks, review your progress and decide to what extent you have achieved your 1 MONTH goals, and then schedule further actions.
- Also, look at your LIFETIME goals. For each "A" goal, choose one action you will take in the next year. Schedule each item with a date to be completed and review these actions monthly.

This exercise should be reviewed and revised on a fortnightly or monthly basis. Once a year, you may want to conduct the whole exercise again, putting aside your previous goals setting exercise. After you have completed the exercise again, you may then contrast your goals of the previous year with the newly developed ones to see how you have progressed in your career path and your priorities.

## 12.6 Conclusions

Mental skills are integral to performance success. The ability to guide oneself consistently into a calm, focused, flexible, goal-oriented state of mind is so central to performance that renowned pedagogues have long championed its use. Karl Leimer, for example, teacher of the famous pianist Walter Gieseking, advocated mental training for all of his students, most notably in relation to memorization but also for refining technique and developing the "inner ear" (see Gieseking & Leimer, 1932/1972). As for the

general benefits of mental training, Leimer remarks:

> To practice the piano five, six, or seven hours daily is generally done without concentration and is at the same time injurious to the health. The mental study demanded by the method here outlined is naturally very strenuous, and pupils afflicted with mental inertness greatly dislike it. But it is the only way in which really good and astonishing results can be obtained (p. 48).

Leimer continually stressed the importance of practicing mental skills but remarked that "when pupils assert that they make use of mental study, it is usually only talk" (p. 49). The purpose of this chapter has been to provide musicians with some representative mental skills exercises that they can apply within the framework of their own performance schedules and their own identified goals. Performers are encouraged to experiment with the exercises presented, adapt them where necessary, and reflect on how their mental and physical skills can interact to enhance their performance. To promote the development of metacognitive skills (as discussed in Chapter 5; also see Nielsen, 1999; Barry & Hallam, 2002), reflective guiding questions have been provided along with the exercises presented here.

Despite the fact that mental skills training is largely specific to the individual, one factor that generalizes across all potential users is that mental skills must be practiced regularly and should be viewed as a long-term commitment. By adopting and sticking with a personalized mental skills regime (embedded within a larger program of high-quality physical practice), the benefits can manifest themselves in ways that are both personally and professionally meaningful.

## Acknowledgments

We would like to thank David Buswell and Mark Mercer for their assistance on earlier drafts of this chapter and The Leverhulme Trust and the Royal College of Music, London, for their support of our efforts to develop a mental skills training curriculum for music students.

## Further information and reading

Further details of mental skills training and sports psychology can be obtained from the websites of the following organizations: the Exercise and Sports Psychology Division of the American Psychological Association (*www.psyc.unt.edu/apadiv47*), the Association for the Advancement of Applied Sports Psychology (*www.aaasponline.org*), and the European Federation of Sport Psychology (*www.psychology.lu.se/Fepsac*). Additional information on the material presented in this chapter is provided at the project website for "Zoning In: Motivating the Musical Mind": *www.zoningin.rcm.ac.uk*. For further reading on mental skills, see:

Emmons, S., & Thomas, A. (1998). *Power Performance for Singers: Transcending the Barriers.* Oxford: Oxford University Press.

Godøy, R. I., & Jørgensen, H. (Eds.). (2001). *Musical Imagery.* Lisse, The Netherlands: Swets and Zeitlinger.

Martens, R. (1987). *Coaches Guide to Sports Psychology.* Champaign, IL: Human Kinetics.

Nideffer, R. M. (1976). *The Inner Athlete.* New York: Thomas Crowell.

Partington, J. T. (1995). *Making Music.* Montreal: Carleton University Press.

Syer, J., & Connolly, C. (1991). *Think to Win.* London: Simon and Schuster.

Syer, J., & Connolly, C. (1998). *Sporting Body, Sporting Mind* (2nd ed.). London: Simon and Schuster.

## References

Barry, N. H., & Hallam, S. (2002). Practice. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 151–165). Oxford: Oxford University Press.

Benson, H. (1975). *The Relaxation Response.* New York: Morrow.

Bird, E. I., & Wilson, V. E. (1988). The effects of physical practice upon psychophysiological response during mental rehearsal of novice conductors. *Journal of Mental Imagery, 12,* 51–64.

Butler, R. J. (1996). *Sports Psychology In Action.* Oxford: Butterworth-Heinemann.

Chaffin, R., Imreh, G., & Crawford, M. (2002). *Practicing Perfection: Memory and Piano Performance.* Mahwah, NJ: Erlbaum.

Clark, D. B., & Agras, W. S. (1991). The assessment and treatment of performance anxiety in musicians. *American Journal of Psychiatry, 148,* 598–605.

Coffman, D. D. (1990). Effects of mental practice, physical practice, and knowledge of results on piano performance. *Journal of Research in Music Education, 38,* 187–196.

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience.* New York: Harper and Row.

Csikszentmihalyi, M., & Csikszentmihalyi, I. S. (Eds.). (1988). *Optimal Experience: Psychological Studies of Flow in Consciousness.* Cambridge: Cambridge University Press.

Emmons, S., & Thomas, A. (1998). *Power Performance for Singers: Transcending the Barriers.* Oxford: Oxford University Press.

Freymuth, M. (1993). Mental practice for musicians: Theory and applications. *Medical Problems of Performing Artists, 8,* 141–143.

Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.), *The Psychology of Music* (2nd ed., pp. 501–602). London: Academic Press.

Gieseking, W., & Leimer, K. (1932/1972). *Piano Technique.* New York: Dover. (Original work published in 1932.)

Grøndahl, D. (1987). *Thinking Processes and Structures Used by Professional Pianists in Keyboard Learning.* Unpublished masters dissertation, Norwegian Academy of Music, Oslo.

Hasher, L., Chung, C., May, C. P., & Foong, N. (2002). Age, time of testing, and proactive interference. *Canadian Journal of Experimental Psychology, 56,* 200–207.

Jones, A. R. (1990). *The Role of Analytical Prestudy in the Memorization and Retention of Piano Music with Subjects of Varied Aural/Kinesthetic Ability.* Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign, IL.

Kendrick, M. J., Craig, K. D., Lawson, D. M., & Davidson, P. O. (1982). Cognitive and behavioral therapy for musical-performance anxiety. *Journal of Consulting and Clinical Psychology, 50,* 353–362.

Kopiez, R. (1990a). *Der Einfluss kognitiver Strukturen auf das Erlernen eines Musikstücks am Instrument.* Frankfurt am Main, Germany: Peter Lang.

Kopiez, R. (1990b). Der Einfluss grafischer vs. verbal-analytischer kognitiver Strukturierung mentalen Erlernen eines Musikstücks. *Jahrbuch der Deutschen Gesellschaft für Musikpsychologie, 7*, 147–155.

Kopiez, R. (1991). Structural aids to the cognitive practice of music: Graphic or verbal analysis? *Psychologica Belgica, 31*, 163–171.

LaBerge, D. (1981). Response to Robert G. Sidnell. In *Documentary Report of the Ann Arbor Symposium. Applications of Psychology to the Teaching and Learning of Music* (pp. 39–41). Reston, VA: Music Educators National Conference.

Lindström, E., Juslin, P. N., Bresin, R., & Williamon, A. (2003). "Expressivity comes from within your soul": A questionnaire study of music students' perspectives on expressivity. *Research Studies in Music Education, 20*, 23–47.

Loehr, J. E. (1987). *Mental Toughness Training For Sports.* Harrisonburg, VA: R. R. Donnelley and Sons.

Martens, R. (1987). *Coaches Guide to Sports Psychology.* Champaign, IL: Human Kinetics.

May, C. P., Hasher, L., & Stoltzfus, E. R. (1993). Optimal time of day and the magnitude of age differences in memory. *Psychological Science, 4*, 326–330.

Nagel, J., Himle, D., & Papsdorf, J. (1989). Cognitive-behavioural treatment of musical performance anxiety. *Psychology of Music, 17*, 12–21.

Nideffer, R. M. (1976). *The Inner Athlete.* New York: Thomas Crowell.

Nielsen, S. G. (1999). Learning strategies in instrumental music practice. *British Journal of Music Education, 16*, 275–291.

Partington, J. T. (1995). *Making Music.* Montreal: Carleton University Press.

Reitman, W. R. (1965). *Cognition and Thought: An Information Processing Approach.* New York: Wiley.

Ross, S. L. (1985). The effectiveness of mental practice in improving the performance of college trombonists. *Journal of Research in Music Education, 33*, 221–230.

Rubin-Rabson, G. (1937). *The Influence of Analytical Pre-Study in Memorizing Piano Music.* New York: Archives of Psychology.

Rubin-Rabson, G. (1941). Studies in the psychology of memorizing piano music: VI. A comparison of two forms of mental rehearsal and keyboard overlearning. *Journal of Educational Psychology, 32*, 593–602.

Simon, H. A. (1973). The structure of ill-structured problems. *Artificial Intelligence, 4*, 181–202.

Stippich, C., Ochmann, H., & Sartor, K. (2002). Somatotopic mapping of the human primary sensorimotor cortex during motor imagery and motor execution by functional magnetic resonance imaging. *Neuroscience Letters, 331*, 50–54.

Syer, J., & Connolly, C. (1991). *Think to Win.* London: Simon and Schuster.

Syer, J., & Connolly, C. (1998). *Sporting Body, Sporting Mind* (2nd ed.). London: Simon and Schuster.

Sweeney, G. A., & Horan, J. J. (1982). Separate and combined effects of cue-controlled relaxation and cognitive restructuring in the treatment of musical performance anxiety. *Journal of Counseling Psychology, 29*, 486–497.

Weinberg, R. S. (1982). The relationship between mental preparation strategies and motor performance: A review and critique. *Quest, 33*, 195–213.

Wessels, M. G. (1982). *Cognitive Psychology.* New York: Harper and Row.

Williamon, A., Valentine, E., & Valentine, J. (2002). Shifting the focus of attention between levels of musical structure. *European Journal of Cognitive Psychology, 14,* 493–520.

Woolfolk, R. L., Carr-Kaffashan, K., & Lehrer, P. M. (1976). Meditation training as a treatment for insomnia. *Behavior Therapy, 7,* 359–365.

Yoon, C., May, C. P., & Hasher, L. (2000). Aging, circadian arousal patterns, and cognition. In D. C. Park & N. Schwarz (Eds.), *Cognitive Aging: A Primer* (pp. 151–171). Hove, UK: Psychology Press.

CHAPTER 13

# FEEDBACK LEARNING OF MUSICAL EXPRESSIVITY

PATRIK N. JUSLIN, ANDERS FRIBERG,
ERWIN SCHOONDERWALDT, AND JESSIKA KARLSSON

One of the recurrent themes in treatises on music is that music is expressive (Davies, 1994). Although this notion seems generally accepted among scholars, the concept of expression itself has remained elusive. This is reflected to some extent in both research and education. In research on music performance, many different approaches to expression have been adopted (for reviews, see Gabrielsson, 1999, 2003), but few attempts have been made to integrate different approaches into a coherent framework (Juslin *et al.*, 2002). Within music education, studies have indicated that the explicit teaching of expressive skills has largely been neglected (Tait, 1992; Persson, 1993; Rostwall & West, 2001) and that the teaching that is available seems to be based on folk theory and tradition, rather than on empirically validated knowledge (Persson, 1996). Why is this so?

First, the nature of musical expressivity does not lend itself easily to formalized description. For instance, much knowledge about expressivity is tacit and, therefore, difficult to express in words (Hoffren, 1964). Second, researchers have not been able to provide music teachers with theories that can guide the development of novel strategies for teaching expressive skills (Juslin & Persson, 2002). Expression in music performance is usually studied in isolation from any concerns with teaching of expressivity (Palmer, 1997), and studies of teaching strategies rarely investigate the nature of expressivity itself (Marchand, 1975). Yet, these aspects should not be separated because the development of new and perhaps more effective methods of teaching expressivity may benefit from a more complete understanding of the nature of expressivity.

Fortunately, research carried out during the last decade or so makes it possible to develop more theoretically informed teaching strategies aimed at expressive skills. The purpose of this chapter is to consider how theory and research on expressivity may inform day-to-day practice in the music profession. First, a critical discussion of traditional views on expressivity is provided, and some of the myths that surround the concept of expressivity are dispelled. Then, a revised view of expressivity based on modern research is proposed. Finally, a new and empirically based approach to learning expressivity, termed *cognitive feedback* (CFB), is described and evaluated. The bulk of this chapter is devoted to expression of emotion, but in the process, other important issues regarding expressivity are considered. The focus is on psychological research rather than philosophical inquiry (for a review of the latter, see Davies, 1994).

## 13.1 The nature of expressivity

One necessary requirement for the development of a good teaching strategy is to have a theory of the skill that is desired. A natural point of departure is to look at how music performers and researchers have traditionally conceptualized expressivity and then to propose a revised view based on current research on expressivity. It should be noted from the outset that the discussion is restricted to Western music, especially classical and popular music from the eighteenth century to the present day.

### 13.1.1 Common myths about expressivity

Views on expression and emotion have changed over time (Ratner, 1980; Cook & Dibben, 2001) as have, without doubt, the expressive features of performance themselves. Certain views have, however, been remarkably consistent across time and musical styles. Below, five myths about expressivity are considered, which may have had a negative impact on the teaching of expressivity in music education. Some of these myths stem from historical treatises and folk theories, and certainly not all students and teachers today subscribe to them. However, the discussion below highlights key issues that are helpful in understanding the context of the novel teaching strategy presented later.

*Expressivity is a completely subjective entity that cannot be studied objectively*

Much has been written about expressivity by philosophers, musicologists, and musicians—often with the implication that there is something *mysterious* about expressivity. Different authors have defined expressivity in strikingly different ways. This has led to the belief that expressivity is a completely "subjective" quality that cannot—or at least should not—be described in scientific terms (Hoffren, 1964; cf. Howard, 1989). Musicians are often unable, or unwilling, to define expressivity or to probe its underlying mechanisms: "great interpretation certainly is a mystery" (Dubal, 1985, p. 62); "the interpretation should speak for itself" (Menuhin, 1996, p. 329); "expressiveness should not be demystified, this could make the expression itself disappear" (teacher cited in Laukka, in press); "the most inspired performances are always those that are inexplicable from a logical point of view" (Ruth Laredo, cited in Dubal, 1985, p. 245).

Does this mean that it is impossible to study expressivity objectively? Not necessarily; acoustic correlates of expressivity can readily be obtained and manipulated in music performances, and listeners' judgments of expressivity can be systematically and reliably related to such acoustic correlates (see Section 13.1.2 below). Note that one common teaching strategy aimed at developing expressive skills, modeling, actually presumes that there are objective correlates of expressivity: the teacher's performance provides an acoustic model of what is desired from the student and the student is required to learn simply by imitating the teacher (Dickey, 1992). The problem is that the student is required to pick up relevant features of the model, and that it may be difficult for a student to know precisely what to listen for and how to represent it in terms of specific skills (Lehmann, 1997; see Woody, 1999, for data supporting this view).

*You must feel the emotion in order to convey it to your listeners*

"A musician cannot move others unless he too is moved," C. P. E. Bach (1778/1985) argued in his book on the true art of playing keyboard instruments (p. 152). This idea is

common even among today's students (Lindström *et al.*, 2003) and teachers (Laukka, in press). Focusing on felt emotions may help a performer to naturally translate emotions into appropriate sounds, but felt emotion does not guarantee that the emotion will be communicated to listeners—neither is it necessary to feel an emotion in order to communicate it successfully. A listener has no direct access to the performer's emotion. What reaches the listener is the sound properties of the music. Emotion is communicated only to the extent that the sound properties contain information about emotion.[1] As noted by Sloboda (1996), students rarely monitor the expressive outcome of their own performances. Instead, they monitor their own intention and "take the intention for the deed" (p. 121). Indeed, recent questionnaire data indicate that 50% of students at conservatories listen to recordings of their own performances "seldom" or "never" (Lindström *et al.*, 2003). This is unfortunate, since listening to the performance may reveal just how far from the intended expression the actual sounds are. Some professional musicians do listen to their own performance: "It can be very instructive. When I hear it back, I think 'did I really play it *that* fast?' or 'Why didn't I give that section more poetry or feeling?' This never fails to amaze me" (Murray Perahia, cited in Dubal, 1985, p. 260).

Undoubtedly, focusing on felt emotions may help a performer to activate existing schemata for emotional expression, and a performer can use emotional memories as a means to self-induce moods (Stanislavski, 1937/1988; Persson, 1993). This is not an all-encompassing strategy, however, because a performer must be able to accomplish a convincing expression, regardless of what he or she feels (which could be nervousness, exhilaration, or intense concentration, depending on the specific circumstances). This is all the more true considering that a performer may need to move swiftly between different expressions within the same piece of music. Emotional engagement is not necessarily bad, but it cannot substitute for teaching that addresses the actual playing in specific and informative ways (see also Tait, 1992).

*Explicit understanding is not beneficial to learning expressivity*

Treatises on art often focus on intuition: "The artist activates . . . aesthetic experience through an intuitive combination of sensory stimuli" (Eibl-Eibesfeldt, 1989, p. 685). This is true also of musicians: "ultimately, the paramount role is that of intuition. For me the determining factor in creativity, in bringing a work to life, is that of musical instinct" (Casals, 1970, p. 97). It seems to be generally agreed that musicians are usually not aware of the details of how their expressive intentions are realized in performance (Sloboda, 1996), although there are individual differences among performers in this regard. To the extent that expressive features are used *tacitly*, this presents a problem for the teaching of expressivity, which relies to a large extent on verbal instruction (Tait, 1992). Moreover, because expert performers do not consciously think about how to apply expressive features in performance, one might wrongly conclude that students do not benefit from thinking about how to apply such features, even in early learning stages.

The notion that learning of expressive skills is "best left untouched" by conscious thoughts reflects a misunderstanding that pervades commonsense teaching based on tradition and folklore. It is important to note that willfully applied, goal-directed strategies normally undergo automation as a result of practice. For instance, although a performer may initially have to apply expressive features in a conscious manner, soon

the associations between features and emotions become internalized by that performer and no longer in need of conscious control. Recent studies have shown that explicit instruction is beneficial to learning expressivity. Woody (1999), for example, noted that "the most effective approach for expressive performance involves conscious identification and implementation of specific expressive features" (p. 339; see Chapter 5 of this volume for a discussion of the benefits of metastrategies for individual learning and practice).

*Emotions expressed in music are very different from everyday emotions*

The most popular idea among musicians and others about what music expresses is emotion (e.g. Budd, 1985; Davies, 1994; Juslin & Sloboda, 2001; Lindström *et al.*, 2003; Laukka, in press). However, several authors have argued that the emotions expressed in music are very different from the emotions expressed in everyday life (e.g. Lippman, 1953; Swanwick, 1985). It is often suggested that music expresses emotions so subtle and complex that they cannot be described in words—that is, they are *ineffable* (Raffman, 1993). There are certain aspects of music experience that are difficult to describe in words, but this does not necessarily apply to emotions (note that most Western languages have several hundreds of emotion words that can be used to describe emotional nuances). Perhaps, listeners confuse the subtlety and variety of the music itself with the emotions actually expressed? The notion that emotions expressed in music are ineffable is not helpful in music education. How could teachers say anything about the ineffable?

An arguably more fruitful approach is to regard emotions expressed in music as similar to those expressed outside of music (Davies, 1994; Juslin, 2001). The acoustic factors used to express emotions in performance (e.g. tempo, sound level, articulation, timbre) gain much of their power from the nonverbal communication of emotions. For instance, there is evidence that performers' expression of emotions in music largely reflects the nonverbal aspects of speech (see Section 13.2). Hence, if one defends the notion that musical emotions are different from everyday emotions, music students are effectively robbed of their most powerful sources of expressive form.

One of the most common strategies for teaching expressive skills (and apparently the most popular among music students; see Lindström *et al.*, 2003) is the use of metaphors, which presumes that emotions expressed in music are similar to everyday emotions. Metaphors are useful because emotions experienced in a nonmusical context can help shape musically relevant emotions. Musical emotions touch deeply not because they are so different from everyday emotions, but because they are so similar. However, the problem with metaphors is that they depend on the performer's personal experience with words and images. For that reason, metaphors may be ambiguous as directions to a performer (e.g. Persson, 1996; but see Barten, 1998, for a positive view). Still, the fact that music students find extramusical sources, such as life situations, useful in developing expression (see Woody, 2000; Lindström *et al.*, 2003) suggests that emotions expressed in music have much in common with emotions in everyday life.

*Expressive skills cannot be learned*

The four myths about musical expressivity described thus far will tend to reinforce the fifth—namely, that expressive skills cannot be learned. If expressivity is completely

subjective and passive and has nothing to do with an explicit understanding or emotions as they are known, it will obviously be difficult to teach expressivity to students. Expertise in music performance is often viewed as the synthesis of *technical* and *expressive* skills (Sloboda, 1996; Gabrielsson, 1999). However, technical aspects of playing are often regarded as learnable skills, whereas expressive aspects are regarded as instinctive (Boyd & George-Warren, 1992, pp. 104–105). Some teachers conceive of expression as "something that cannot be taught," a view that is shared by some students: "there is no technique to perform expressively. You have to use your soul" (student cited in Woody, 2000, p. 21).

As discussed by Sloboda (1996), one consequence of this view is that musical expressivity is wrongly believed to reflect only musical talent that is beyond learning and development. True enough, expressive skills sometimes reflect the emotional sensitivity of the musician in regard to musical sounds (as suggested by the personality concept of *auditory style*; see Brodsky *et al.*, 1994). But this does not mean that it is impossible to develop expressive skills through training. This is really an empirical question: studies that have addressed this issue have demonstrated that expressive skills can be improved by instruction and training (Marchand, 1975; Johnson, 1998; Woody, 1999; Juslin & Laukka, 2000; Sloboda *et al.*, 2003; see also Section 13.2.4 below).

Perhaps, the critical issue is *how* expressivity is being taught to students. Essential to the learning process is the provision of informative feedback. Technically, *feedback* has been defined as "the process by which an environment returns to individuals a portion of the information in their response output necessary to compare their present strategy with a representation of an ideal strategy" (Balzer *et al.*, 1989, p. 412). Consideration of this definition suggests that the traditional teaching strategies mentioned earlier (i.e. modeling and metaphors) rarely provide informative feedback, for they do not provide the performer with a comparison of his or her current performance strategy with an optimal strategy. Could empirical research help to address this problem?

### 13.1.2 Perspectives from empirical research

Experimental studies of music performance have been conducted for at least 100 years (Gabrielsson, 1999). Researchers have investigated performance expression mainly by measuring various acoustic parameters in the performance and relating these measures to a musical notation. In this particular context, expression usually refers to the "systematic variations" or "deviations" in timing, dynamics, timbre, and pitch that form the so-called *microstructure* of a performance, and differentiate it from another performance of the same music (Palmer, 1997). Sometimes, attempts have been made to relate such acoustic properties to the perceived expression of the performance, as evident from listener judgments.

Much like the individual performer who first learns to play a musical instrument, studies of expressivity in music performance have progressed from a consideration of basic principles of expression (e.g. how does the performer convey the phrase structure?) toward more subtle and individualistic aspects (e.g. what makes a performance truly special?). However, it is fair to say that researchers are still struggling with the problem of how to account for the microstructure of a typical performance of music. Contributing to the difficulty are the vast number of factors that, in principle, may

**Table 13.1** Examples of factors that may influence expressivity in music performance (based on Juslin, 2003)

| | Examples of factors |
|---|---|
| *Piece-related* | The musical composition itself |
| | Notational variants of the piece |
| | Consultations with composer or composer's written comments |
| | Musical style/genre |
| *Instrument-related* | Acoustic parameters available |
| | Instrument-specific aspects of timbre, pitch, etc. |
| | Technical difficulties of the instrument |
| *Performer-related* | Performer's structural interpretation |
| | Performer's expressive intention with regard to the mood of the piece |
| | Performer's emotion-expressive style |
| | Performer's technical skill |
| | Performer's motor precision |
| | Performer's mood while playing |
| | Performer's body language |
| | Performer's interaction with co-performers |
| | Performer's perception of/interaction with audience |
| *Listener-related* | Listener's musical preferences |
| | Listener's musical expertise |
| | Listener's personality |
| | Listener's current mood |
| | Listener's state of attention |
| *Context-related* | Acoustics |
| | Sound technology (e.g. recording technique, stereo equipment) |
| | Listening context (e.g. recording, concert) |
| | Other individuals present |
| | Visual performance conditions |
| | Larger cultural and historic setting |
| | Whether the performance is formally evaluated |

influence expressivity in music performance (see Table 13.1). Music performance is a complex process affected by numerous factors related to the music: the instrument, the performer, the listener, and the social context. As a result of the practical limitations associated with experimental studies, researchers have been forced to bracket most of the factors shown in Table 13.1 and to focus instead on the core principles of music performance (discussed later in this section) that transcend a specific time and place (in certain styles of music in the Western world, at least).

Still, some consistent themes have emerged from this field of research, one of which is that performance expression is best conceived of as a multidimensional phenomenon, which may be decomposed into different subcomponents. That is, there are different kinds of expressivity that arguably make different contributions to the aesthetic impact of a given performance. Expressive features may serve different and sometimes opposing functions. The exact nature of these components of expressivity is still being discussed by researchers. However, drawing on previous research, Juslin (2003) argues

that expressivity derives from five main sources (collectively referred to as the GERMS model):

1. *Generative rules (G)* that mark the structure in a musical manner (Clarke, 1988). By means of variations in such parameters as timing, dynamics, and articulation, a performer is able to communicate group boundaries (Gabrielsson, 1987), metrical accents (Sloboda, 1983), and harmonic structure (Palmer, 1996). Generative features may increase the emotional impact of the music, but do so mainly by enhancing the expression that is inherent in the structure of the piece. Hence, part of the expression reflects the *structure* of the piece performed—as filtered through the performer's interpretation.
2. *Emotional expression (E)* that serves to convey emotions to listeners (Juslin, 1997a). By manipulating overall features of the performance, such as tempo, timbre, and loudness, a performer is able to play the same structure with different emotional characters. Often the performer will try to support the emotional character suggested by piece-specific features (e.g. melody, harmony), although by contradicting the expression of the piece to some degree, the performer may convey "conflicting," "mixed," or "complex" emotions.
3. *Random fluctuations (R)* that reflect human limitations in motor precision (Gilden, 2001). It has been revealed in several studies that even when expert performers attempt to play perfectly even time intervals (e.g. in a tapping task), there are still small and involuntary fluctuations in timing in their performance. Practice instills consistency in speed, precision, and fluency—among other things—but it seems that the small random variations contribute to the "living" character of music; the slight unpredictability that makes every performance absolutely unique (Juslin *et al.*, 2002).
4. *Motion principles (M)* that hold that tempo changes should follow natural patterns of human movement (e.g. gesture). A pleasing performance is one in which the expressive microstructure satisfies basic constraints of animate motion (Shove & Repp, 1995). For instance, Friberg and Sundberg (1999) showed that final ritards of music performances that are preferred by listeners tend to follow a tempo curve similar to that of runners' decelerations.[2]
5. *Stylistic unexpectedness (S)*, which reflects a performer's deliberate attempt to deviate from stylistic expectations concerning performance conventions in order to add tension and unpredictability to the performance (e.g. Meyer, 1956, p. 206). This aspect is probably the least researched so far, but it may be critical to developing a truly original interpretation.

All of these components reflect psychophysical relationships among acoustic characteristics of the performance (specific patterns of information) and psychological characteristics of the human listener (e.g. categorical perception). In reality, all five components occur together in complex interactions, but for certain purposes (e.g. research, teaching), it may be useful to consider them separately. The reasons for this are that the components have different origins, display different characteristics, and are processed partly by different regions of the brain (Juslin, 2003). Juslin *et al.* (2002) made an attempt to simulate the first four of the GERMS components in

synthesized performances of music, and a listening test indicated that all four components contributed to the perceived expressivity of a performance, although in different degrees. Statistically, the emotion component (E) had the largest impact on listeners' judgments.

What are the implications of the GERMS model for music education? One is that, at certain stages, the different aspects of expressivity may need to be taught separately because they have different characteristics. According to the GERMS model, a performance should (1) convey the generative structure of the music, (2) express emotions, (3) exhibit motor precision, (4) be suggestive of human motion and gesture, and (5) deviate from stylistic expectations in aesthetically pleasing ways. The conveying of structure (G) is probably the aspect that has received most attention in research and education, perhaps because it is the most tangible aspect: "structure can be seen, heard, demonstrated" (Epstein, 1995, p. 126). Certainly, there is no substitute for a thorough understanding of the musical structure in shaping a musical interpretation. This aspect of expressivity may be guided by structural analysis, by consulting the composer or his or her writings, and also by exploring visual graphs of performance parameters alongside the notation (e.g. Riley-Butler, 2001). The motion component (M) may be trained by using techniques outlined by Davidson and Correia (2002), such as asking the student to conduct the expression in the playing of the teacher and developing nonverbal narratives of physical gesture for individual phrases. The stylistic unexpectedness component (S) requires extensive knowledge about musical styles and performance conventions, as well as an ability to conceive of the performance structure in terms of a narrative of musical expectations that can be experimented with in creative ways. The aspect of expressivity in most need of a formalized teaching strategy is emotional expression (E).

*Expression of emotions: Summary of findings*

Research on emotional expression, as manifested specifically in *performance features* (e.g. timing, timbre) as opposed to features of the composition (e.g. melody, harmony), have matured in the last decade (Juslin, 2001). Most of these studies have used a procedure in which musicians were asked to play various pieces of music in order to express specific emotions (e.g. tenderness). The performances were recorded and used in listening tests to see whether listeners could (accurately) recognize the intended expression.[3] In addition, many studies have analyzed the acoustic features of the performances to investigate *how* each emotion was expressed. Below, the key findings from this research are summarized (for more extensive treatment, see Juslin & Laukka, 2003).

1. Professional performers are able to communicate basic emotions (happiness, sadness, anger, fear, tenderness) to listeners with accuracy about as high as in facial and vocal expression of emotions (Juslin & Laukka, 2001, 2003). Amateur performers communicate emotions less effectively and tend to apply the expressive features inconsistently (Juslin & Laukka, 2000; Rohwer, 2001).
2. The communicative process appears to operate on a fairly broad level of emotion categories, whereas finer distinctions within these categories are difficult to convey (Juslin & Lindström, 2003; Juslin, 1997c), at least without additional context (lyrics, program notes, visual cues).

Table 13.2 Summary of acoustic features used to express 12 "basic" and "complex" emotions by professional performers (adapted from Juslin & Lindström, 2003)

| | Acoustic feature | | | | | | |
|---|---|---|---|---|---|---|---|
| | Tempo | Sound level | Articulation | Timbre* | Attack* | Vibrato extent* | Vibrato rate* |
| Anger | ↑ | ↑ | = | ↑ | ↑ | ↑ | = |
| Contentment | ↓ | = | ↑ | = | ↓ | ↓ | ↓ |
| Curiosity | = | = | ↓ | = | ↓ | ↑ | = |
| Disgust | = | ↑ | ↓ | = | ↑ | = | ↑ |
| Fear | = | ↓ | ↓ | ↓ | ↑ | = | ↑ |
| Happiness | ↑ | ↑ | ↓ | = | ↑ | ↓ | = |
| Jealousy | ↑ | ↑ | = | ↑ | = | ↑ | ↓ |
| Love | ↓ | = | ↑ | = | ↓ | = | ↓ |
| Pride | = | ↑ | = | ↑ | ↑ | = | ↑ |
| Sadness | ↓ | ↓ | ↑ | ↓ | ↓ | ↓ | ↓ |
| Shame | ↓ | ↓ | ↓ | ↓ | = | = | ↑ |
| Tenderness | ↓ | ↓ | ↑ | ↓ | ↓ | ↓ | = |

*Note*: ↑ indicates a high value, ↓ indicates a low value, = indicates a medium value. A high value indicates (respectively for each cue) fast (versus slow) tempo, high (versus low) sound level, legato (versus staccato) articulation, much (versus little) high-frequency energy in the spectrum, fast (versus slow) attack velocity, large (versus small) vibrato extent, and fast (versus slow) vibrato rate ($N = 216$, except * $N = 144$).

3. The timbre of specific musical instruments may affect the effectiveness of the communicative process to some extent (e.g. Behrens & Green, 1993).
4. Many studies have attempted to describe the "code" (the acoustic means) that performers use to convey each emotion. This has shown that performers manipulate variables such as tempo, sound level, articulation, timbre, timing, attack, decay, intonation, portamento, and vibrato to express different emotions (Juslin, 2001). Both mean levels and patterns of variability in these variables may influence the perceived expression (Juslin, 1997b; Juslin & Madison, 1999). Table 13.2, for example, summarizes, in a simplified fashion, patterns of acoustic features used by professional performers to express 12 "basic" and "complex" emotions (on guitar, piano, and saxophone). In reality, the emotional expression will often change rapidly within a piece, or even within a single note. Many different musical styles have been studied (classical music, folk music, rock, opera, jazz), but it is not yet known whether the obtained relationships hold for other cultures as well.
5. Whether researchers' descriptions of acoustic features used to express various emotions are valid can be confirmed by creating synthesized performances of music on the basis of these findings. Synthesized performances with appropriate acoustic settings can express emotions as reliably as human performers (Juslin, 1997b; see also Bresin & Friberg, 2000).
6. An extensive review of 145 studies (101 speech and 44 music studies) strongly indicates that patterns of acoustic features used to convey each emotion largely derives from the nonverbal features of emotional speech (Juslin & Laukka, 2003), as suggested by Spencer (1857).

7. The expression created by overall features such as tempo can be enhanced by accents on specific notes in the melodic structure (e.g. those with much melodic tension; Lindström, 2003).
8. The communication of emotion in music is generally successful despite individual differences in the use of acoustic features among both performers and listeners (e.g. Juslin, 2000), and despite the fact that different musical instruments provide different features at the musician's disposal.

## 13.2 A novel approach: Feedback-learning of musical expressivity

In a project called "Feedback-learning of Musical Expressivity" (Feel-ME), the authors aim to develop new methods for teaching expressive skills based on recent advances in musical science, psychology, technology, and acoustics. *Cognitive feedback* (CFB), an empirically based approach to learning expressivity, has been developed and implemented in user-friendly computer software. A number of computer programs aimed at enhancing performance skills have been created, as computers have become both less expensive and more easy to use (Webster, 2002). None of these programs, however, seems to focus on emotional expression.

### 13.2.1 Criteria for a useful teaching strategy

One requirement for teaching emotional expression is that it improves the communication of emotion as effectively as possible. However, the ultimate goal should be to provide performers with the tools they need to develop their own personal expression (originality is often emphasized as an important criterion for an artistic work). To achieve a creative interpretation of a piece of music, the performer may wish the emotional expression to be clear or ambiguous, stable or variable, specific or general. But only detailed knowledge about the relationships among expressive features and their perceptual effects will help the performer to achieve the desired effects on listeners reliably. This knowledge may serve as the core from which interpretations are developed, with knowledge about the particular musical instrument, piece, style, and composer. Given these aims, what are the specific criteria for a useful teaching strategy?

- First, it is essential that the teaching strategy is well-suited to the nature of the communicative process as described in research. Such communication involves a number of acoustic features that are used largely tacitly by performers and listeners. A useful teaching strategy should take these characteristics into account, and help to render transparent the communicative process.
- Second, the teaching strategy should include the three elements required for *deliberate practice* (Ericsson *et al.*, 1993): (1) a well-defined task, (2) informative feedback, and (3) opportunities for repetition and correction of errors.
- Third, the teaching strategy should allow the performer to compare his or her playing with an "optimal" (or reference) model, in accordance with the definition of feedback discussed earlier.
- Fourth, the teaching strategy should relate sound properties of the performance to experiential concepts (e.g. emotion) relevant in interpretation. Previous strategies

focus either on acoustic aspects (e.g. modeling) or experiential aspects (e.g. metaphors), but a useful strategy should preferably resolve this dualism by describing the relationships between the two.

- Finally, the efficacy of the teaching strategy should have been empirically demonstrated.

Below, a teaching strategy is presented that meets these criteria. First, however, a general model is outlined that is suited to describing the communicative process.

### 13.2.2 The lens model

Communication of emotion requires that there is both a performer's intention to express an emotion and recognition of this same emotion by a listener.[4] The performer may, for instance, wish to highlight an emotional quality of a piece of music (e.g. to bring out love and joy in Beethoven's Appassionata; Dubal, 1985, p. 107). Emotional communication requires a controlled application of dynamic contrasts in various expressive features: fast–slow, loud–soft, staccato–legato, bright–dark, and so on. These features have some peculiar characteristics that are essential to understand in order to devise an efficient teaching strategy.

One way of capturing the characteristics of the communicative process is to conceptualize it in terms of a variant of Brunswik's (1956) *lens model*. The lens model was originally intended as a model of visual perception but was later used mainly in studies of human judgment (Cooksey, 1996). The lens model (see Figure 13.1) can be used to illustrate how musicians express emotions by means of a set of *cues* (e.g. tempo, sound level, timbre) that are uncertain, but partly redundant (the cues covary to some extent).

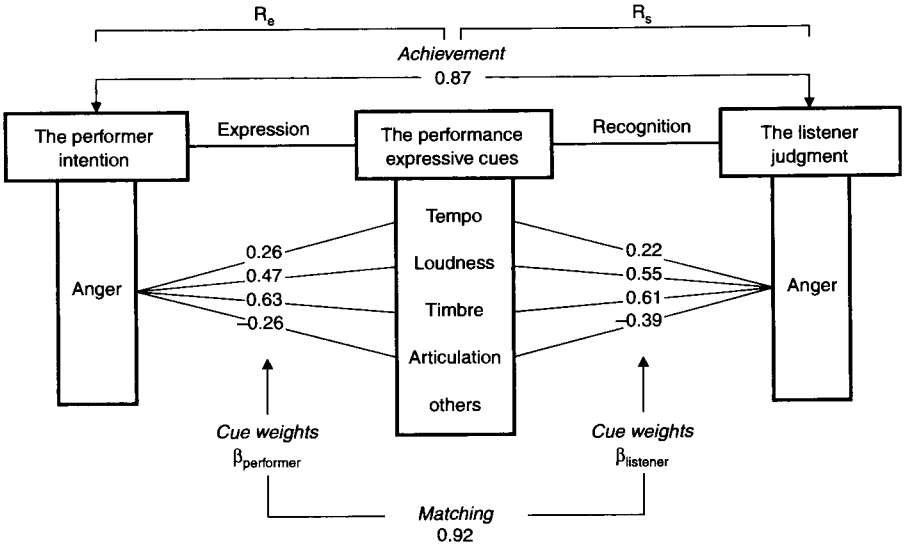



Figure 13.1 A Brunswikian lens model of communication of emotion in music performance. The cue weights should be interpreted as follows: positive (as opposed to negative) signs indicate, respectively for each cue, fast (versus slow) mean tempo, high (versus low) sound level, sharp (versus soft) timbre, and legato (versus staccato) articulation (for details regarding methodology, see Juslin, 2000).

The expressed emotions are recognized by listeners, who use the same cues to recognize the expression. The cues are uncertain because they are not perfectly reliable indicators of the intended expression. Thus, for instance, fast tempo is not perfectly correlated with expression of happiness, because fast tempo is also used in expression of anger. None of the cues are completely reliable when used in isolation, but by integrating the values of various cues, performers and listeners can achieve reliable expression and recognition, respectively. This is not simply a matter of pattern matching (as suggested by Table 13.2), since cues contribute in a mainly additive fashion to listener judgments (Juslin, 1997b). This can explain how the communication can be successful on different musical instruments, which provide different cues. Redundancy between cues partly reflects how sounds are produced on instruments (e.g. a harder string attack produces a tone that is both louder and sharper in timbre). There are several indices in the lens model that are key in understanding the communicative process (for a description of how each index is measured, see Juslin, 2000):

- *Achievement* ($r_a$) refers to the relationship between the performer's expressive intention and the listener's judgment. It is a measure of how well the performer succeeds in communicating a particular emotion to listeners.
- *Cue weight* ($\beta_1$, $\beta_2$, $\beta_3$ . . .) refers to the strength of the relationship between an individual cue, on the one hand, and expressive intentions and listeners' judgments, on the other (indexed by correlation statistics). Cue weights indicate how individual cues are used by performers and listeners, respectively (e.g. that the performer uses fast tempo in expression of anger).
- *Matching* (G) refers to the degree of similarity between the performer's and the listeners' utilization of acoustic cues, respectively. For effective communication to occur, the performer's use of cues (i.e. their cue weights) must be reasonably matched to listeners' use of cues.
- *Consistency* ($R_e$ and $R_s$) refers to the degree of consistency with which the performer and listener, respectively, are able to utilize the various cues. Other things equal, the communication will be more effective if the cues are used in a consistent manner.

The important point is that the upper limit of communication accuracy is set by the degree of matching, performer consistency, and listener consistency. If the emotional communication is *not* successful, this may be because (1) performer and listeners use the cues differently (poor matching), (2) the performer uses the cues inconsistently, and (3) the listeners use the cues inconsistently. By analyzing these indices separately, one can interpret the success of communication in a particular situation. This information is needed in order to be able to improve the communicative process (Juslin, 2000; see also Hursch *et al.*, 1964).

The lens model has several important implications. First, perfect accuracy of communication cannot be expected. If the acoustic cues are uncertain, it means that the communicative process is uncertain too. Second, to understand the success of the communication in a specific situation, both expression and recognition of emotions must be described in terms of the same concepts. The model implies that the extent to which the emotional expression of a piece of music is recognized depends equally on the

sender and the receiver. Third, because the cues are partly redundant, more than one cue utilization strategy could lead to the same level of accuracy (e.g. Juslin, 2000). Since there is no pressure toward uniformity in utilization of cues, performers can communicate successfully with listeners, without having to compromise their unique playing styles. The key is to profit from the cue redundancy through *vicarious functioning* of cues (Brunswik, 1956)—one cue may substitute for another to some extent. Of course, there are clear limits to this tolerance, and the use of vicarious functioning requires a certain kind of knowledge (detailed in Section 13.2.3). Further, this flexibility comes with a price: redundancy limits the information capacity of the communication because the same, rather than unique, information is conveyed by many cues. This explains why performers find it hard to communicate anything other than broad categories of emotion (Juslin & Lindström, 2003).

### 13.2.3 Cognitive feedback

The lens model offers a useful tool for improving the communication of emotions in music. Recall that emotional communication in music performance involves a large number of acoustic cues that are used by both performers and listeners. Both expression and recognition of emotion are made by integrating these cues. Such integration requires knowledge about the *relationships* between performers, cues, and listeners. The notion of CFB is to allow the performer to compare a model of his or her playing to an "optimal" model of playing. This is obviously a feature of feedback as such; however, it is important to understand how this is accomplished in the present method. Several of the performer's manipulations of acoustic cues are audible to listeners in general and to music teachers in particular. Nevertheless, it is extremely difficult for a human perceiver to infer the statistical relationships that exist among expressive intentions, acoustic cues, and listener impressions (see Figure 13.1), let alone the relationships among the cues themselves. It is well-known from extensive research on human judgment that judges are usually unable to verbalize the basis of their own judgments, especially in situations that involve many uncertain cues (Cooksey, 1996). CFB benefits from a statistical method (i.e. regression analysis) that makes it possible to describe the complex relationships with a precision that human perceivers would find difficult to achieve (e.g. Juslin, 2000). The authors recently implemented CFB in computer software. In the following summary of the basic procedure of CFB (Box 13.1), the plain text explains what the user (i.e. the musician) is doing in each phase, and the text in *italics* explains what the software is doing in each phase. This is to highlight that, although the principles by which CFB is generated are complicated, the actual use of the software is relatively straightforward.

It should be noted that using this software may comprise two different aspects of learning expressivity: (1) a "normative" aspect that involves trying to maximize communication accuracy through effective use of cues (the three phases described in Box 13.1 may be repeated as many times as judged necessary, e.g. until a certain level of matching is accomplished), and (2) an "aesthetic explorative" aspect that means that the performer may simply experiment with different ideas for interpretation, and check how listeners would be likely to perceive the resulting expression. The latter aspect

## Box 13.1 CFB: Basic procedure

- In the first phase, the performer records a set of different performances of the same melody with the goal of expressing particular emotions (e.g. happiness, sadness, anger, tenderness) that are selected at the outset. The performer should record several versions of each emotional expression to obtain a representative sample of performances. The recording of each version can be repeated until the performer is satisfied with the result. When the performer has recorded the performances, he or she requests feedback from the software.[5]
- *Performances are recorded and acoustic cues are automatically analyzed by the software. The recording is first segmented into tone boundaries from an analysis of both sound level and pitch. Potential tone onsets and offsets are detected by identifying areas of similar pitch and substantial dips in the sound level. An example of the segmentation is presented in Figure 13.2. For each detected tone, the following cues are computed: pitch (in semitones), sound level (dB, upper quartile of sound level within onset–offset), instantaneous tempo (notes per second), articulation (percentage of pause duration), attack velocity (dB/s), spectral content (dB, difference between high and low spectral content, i.e. a correlate of perceived timbre), vibrato rate (Hz), and extent of vibrato (semitones). Then, statistical analysis is used to model the correlations between the performer's expressive intentions and the acoustic cues. This produces indices of consistency and cue weights for the performer (explained in Section 13.2.2). The performer's model is also mathematically related to a stored model of*

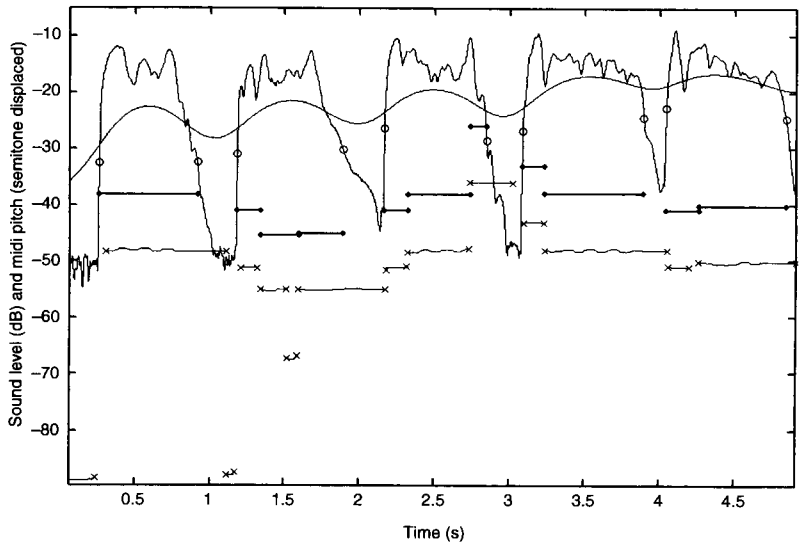



**Figure 13.2** Automatic extraction of musical notes. This is an example of the onset and offset detection applied to a short phrase played on the violin. The upper two lines show the sound level and the "smoothed" sound level. Circles indicate tone onset and offset as indicated by sound level detection. The lines with crosses at the bottom are the tones suggested by the pitch analysis. The resulting tones from the combination of pitch and sound level analysis are the thick lines in the middle. (For clarity, the pitch graphs are displaced by −110 and −120 semitones, respectively.)

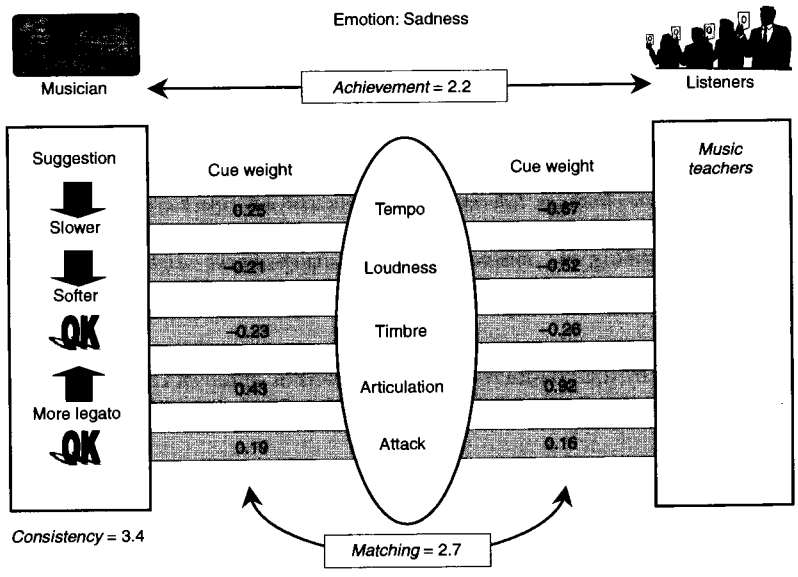



**Figure 13.3** The graphical interface for cognitive feedback used in the CFB software.

*listeners' judgments of emotion in music performances. This listener model was obtained from listening tests in which listeners were asked to judge the emotional expressions of a wide range of performances with varying emotional expression. Statistical analysis was used to model the correlations between listeners' judgments and acoustic cues, thus producing a general model that may be used to simulate new listener judgments. From this, the program obtains indices of achievement, matching, consistency, and cue weights. This information is finally transformed to a graphical interface, according to set criteria (for further details, see Juslin* et al., *2003).*

- In the second phase, the performer receives CFB from the software: this consists of a visual and numerical description of the performer's use of cues, the listeners' use of cues, the degree of matching between performer's and listeners' cue weights, the consistency of the performer's use of cues, and the achievement. All this is shown in a graphic interface that resembles the lens model (see Figure 13.3). This makes it possible to compare directly how performers and listeners use the same cues in the performances. Instances of poor matching are highlighted by arrows that signal that a change in utilization of a cue in a specific direction is recommended. The recommendation is also expressed verbally (e.g. "softer"). If the performer is using cues in an inconsistent manner, this will also be apparent. Quantitative indices of the lens model—achievement, matching, consistency—are converted to "scores" (1–5). From this point, the performer should try to revise his or her use of cues in accordance with provided feedback (e.g. to use more legato articulation in expression of sadness).
- In the final phase, the user repeats the first task once again (i.e. recording of performances with the same emotional expressions). The aim is to see whether the performer has improved his or her accuracy by changing the use of cues in the ways recommended by the feedback. Thus, it can easily be examined which cues are used effectively and

**Box 13.1** (*Continued*)

which cues need continued attention. The performer may also request statistics on the progress over repeated feedback cycles.

- *The computer software again records and analyzes the acoustics of the performances, and uses simulated listener judgments to obtain updated lens model indices, which may be compared with previous findings. Information from successive feedback cycles are stored in the computer software and can be retrieved at will from a database to plot learning curves for achievement, matching, and consistency, as well as changes in cue weights.*

is further enhanced by the possibility to select models of different types of listeners (e.g. music teachers, expert performers, or laymen) for the simulation of listeners' judgments.

### 13.2.4 Research on cognitive feedback

A pilot study evaluated the efficacy of the CFB procedure (although using manual acoustic measurements), and indicated that the method is highly effective (Juslin & Laukka, 2000). Eight pop/rock guitarists (aged 16–29) at an intermediate level of skill went through one cycle of CFB (the three phases described in Box 13.1) playing three different pieces of music. The results indicated that CFB yielded about a 50% increase in accuracy after a single feedback session, as shown by listening tests. Moreover, a questionnaire suggested that the performers reacted positively to the CFB and believed that it had improved their skills. In contrast, a no-training contrast group did not improve its accuracy.

After these promising findings, the authors began developing computer software that could provide automatic feedback to musicians. Research is currently underway to evaluate the usability of the software. *Usability* is the quality that arises from a well-functioning interplay between (1) a task, (2) a user who is solving the task, and (3) a system that supports the user in solving the task. Four features of usability are regarded as important in the evaluation (Briggs, 1987): *efficacy* (does the software improve the communication?), *user satisfaction* (how is the software experienced by the user?), *user understanding* (has the user adequate knowledge regarding the software?), and *training cost* (how quickly can the user learn to use the software?). In one study (Juslin *et al.*, 2003), 36 semi-professional jazz/rock guitarists (aged 21–49) were randomly assigned to one of three conditions:

1. *CFB provided by the computer software*: After a brief interactive exploration of the software supervised by the experimenter, the participant was instructed to go through one cycle of CFB as described above in Section 13.2.3. The melody used was "When the Saints" and the emotions to be expressed were happiness, sadness, anger, and fear. Measures of achievement were obtained before (pre) and after (post) feedback. The participant also completed a questionnaire regarding usability.
2. *Teacher feedback*: The participant carried out the same task as in condition 1, except that the feedback was provided by a teacher. The teacher was allowed to use any kind of verbal instruction (e.g. metaphors, technical directions) to help the participant improve the emotional expression.

3. *No feedback (contrast group)*: The participant received no feedback, but simply performed the musical material (same as in conditions 1 and 2) twice (pre- and posttest) with a break in between.

Figure 13.4 presents the main findings from this study in terms of measures of achievement (pre and post) for each of the three conditions. The largest improvement

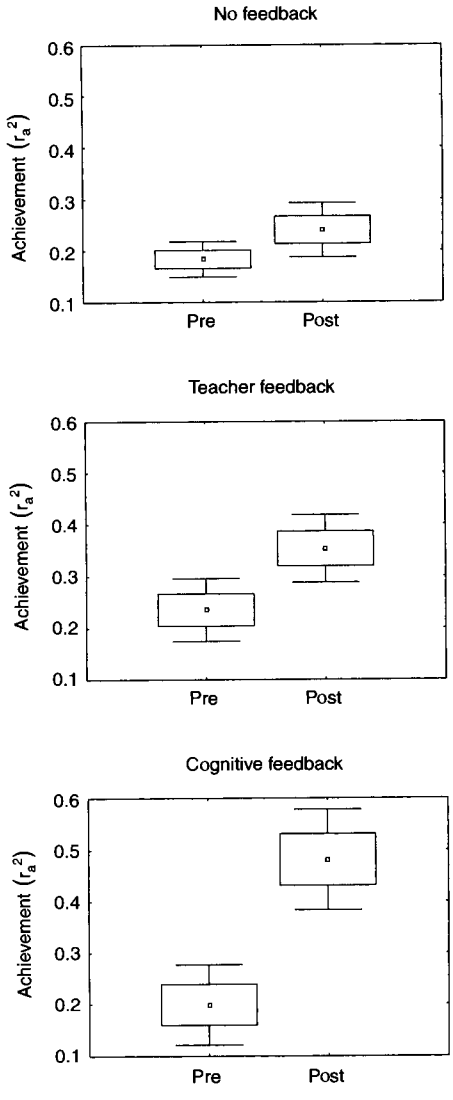



**Figure 13.4** Results from a study that investigated the efficacy of cognitive feedback from the CFB software in comparison with teacher feedback and a no-feedback contrast condition (from Juslin *et al.*, 2003). Markers indicate the mean level of achievement (squared), boxes indicate the standard error of the estimate, and whiskers indicate 95% confidence intervals around the mean.

in achievement occurred for the CFB group, followed by the teacher group, and the no-feedback group. In sum, CFB seems to be highly effective in improving performers' communication of emotions.

### 13.2.5 Generality of the CFB approach

The present approach offers a certain level of generality, since the basic procedure of CFB (recording, analysis, simulation of listener judgments, feedback) can, in principle, be used with any style of music. What is required to adapt the CFB software to a specific style is (1) that all acoustic parameters that are relevant to the style are featured in the acoustic analysis and (2) that the listener model used to simulate listener judgments is based on listening tests in which musical examples, emotion terms, and listeners are appropriate for the musical style of interest. Could the CFB software work with music performance based in a different cultural context? Again, there is nothing in the basic procedure of CFB as such that precludes its use with regard to music from other cultures. In the current version of the software, the acoustic cue extraction and the listener models were based on characteristics of Western music (e.g. classical music, pop music, jazz), but both of these parts of the software could, in principle, be adapted to other cultural settings.

However, there are also some limitations of the CFB approach to learning expressivity. For example, in its current form, the software can only analyze cues from monophonic performances of music (melody). Thus, the software is mainly suitable for single line, solo instruments, such as the violin, flute, guitar, saxophone, and voice. Still, researchers are currently attempting to develop cue extraction for polyphonic performances as well. Another limitation of the CFB software is that it is restricted to fairly brief extracts of music; the software analyzes cues in terms of average measures across each recorded performance, and these measures are not meaningful for longer pieces where the expression may change substantially. One solution to this problem could be to practice the various sections of a longer piece in short segments that are appropriate for the CFB software.

There are also some conceptual problems. The issue of developing expressive performance raises the crucial question of what constitutes an "optimal" performance. Arguably, this issue spans many different artistic aspects, including originality, recognition, arousal, beauty, emotion, balance, and personal expression. In this chapter, the focus has been on only one of these aspects—namely, emotional expression. In the very specific context of the CFB software, it is easy to define what constitutes an optimal performance: an optimal performance is one that communicates the intended emotion reliably to listeners by incorporating cues in accordance with how listeners use the same cues in their judgments of expression. As this is only one aspect of expressivity, performers obviously have to develop other aspects as well, using other means (see Section 13.1.2).

It is not clear whether music performers at all skill levels would benefit from using the CFB software. Obviously, the student must be able to manipulate various acoustic cues independently in order to benefit from the detailed feedback. Hence, the software may be too demanding for novices. On the other hand, expert musicians may already have perfected many of the features addressed by the software. Thus, the software may be most useful for performers at an intermediate skill level.

### 13.2.6 Implications for music education

The use of computer feedback brings certain advantages: it (1) can provide critical feedback but in a nonthreatening environment, (2) is easily available, (3) provides possibilities for flexible and individually-based learning, and (4) explicitly describes cues that are typically embedded in tacit knowledge. There could be drawbacks as well. First, the usefulness of computer feedback depends on the availability of computers, and not all institutions or individuals may have access to these. Fortunately, some recent estimates show that the availability of computers in music educational contexts is steadily increasing (Webster, 2002). Second, one may fear that use of CFB could lead to standardized performances of music. It must be emphasized that the decision about how to interpret the music is left to the performer. The CFB software only serves to help performers achieve intended interpretations more reliably, whatever those may be, by giving performers a better understanding of cue–emotion relationships. Third, computers lack a human touch that may be highly valued by the student. However, the teacher could play a supporting role also when using computer-assisted teaching strategies, especially in shaping aesthetic judgment and achieving balance among different aspects of expressivity. One final problem with computer-based instruction (CBI) is that students seem to have a negative attitude toward using computers in music teaching (Lindström *et al.*, 2003). Such skepticism may be unwarranted; studies that have compared CBI with traditional approaches suggest that CBI can be just as effective as teacher-based instruction (Taylor, 1981; Webster, 2002).

## 13.3 Conclusions

A number of conclusions can be drawn from the evidence outlined in this chapter. First, expressivity is an empirically tractable problem: acoustic correlates of expressivity have been identified and simulated (Section 13.1.2) and important aspects of the communicative process can now be modeled in a systematic way (Sections 13.2.2–13.2.4). This research has various limitations but shows that it is time to discard the view that expressivity is a mystery, impossible to describe in scientific terms. Second, music teaching should incorporate up-to-date theories and findings about expressivity to replace the common myths about musical expressivity described in Section 13.1.1. Third, although traditional teaching strategies (e.g. the use of metaphors) may be useful, teaching of expressivity should at some stage, at least, address the actual sound properties of the performance (Woody, 1999) and should help the student to relate sound properties to experiential properties. Finally, CFB as provided in user-friendly software offers an efficient method of teaching emotional expressivity to students that could complement traditional teaching strategies.

Findings from a questionnaire study indicate that the majority of music students would like to practice more on expressivity than is currently the case, and that most of them are willing to try out new teaching strategies (Lindström *et al.*, 2003). This chapter has attempted to show that new strategies for teaching expressivity are indeed within reach. Adopting these strategies in music education will require openness among students and teachers. Some of the traditional views on expressivity may need to be revised in the light of empirical findings; however, all aspects of musical expressivity may never be fully

explained. Be that as it may, it must not prevent practitioners from making an effort to improve teaching and learning of musical expressivity on the basis of empirical research.

## Acknowledgments

The writing of this chapter was supported by the Bank of Sweden Tercentenary Foundation. We are grateful to Aaron Williamon, Peter Johnson, Chris Gayford, Caroline Minassian, and two anonymous reviewers for helpful comments on an earlier version of this chapter.

## Notes

1. An important exception to this is when the *visual* impression of a performer may influence the listeners' experience of the music (see Chapter 4). Listeners may be much more likely to engage emotionally with a performance if they believe the performer is also engaged, as suggested by facial expressions and body language. However, such features may be either feigned or genuine expressions, just like the sounds produced by the performer.
2. A subset of movement patterns in performance expressivity are those involuntary patterns that reflect instrument-specific constraints in combination with anatomical constraints of the human body. These patterns should be distinguished from a performer's intentional attempt to render a performance with a certain motion character (Juslin *et al.*, 2002).
3. In this chapter, the focus is on *expression* and *recognition* of emotion, rather than on the more labile process of *induction* of emotion, which is influenced by a number of factors outside the performer's control (see Chapters 4, 16, 18, and 19 in Juslin & Sloboda, 2001).
4. Certainly, performers may not always strive to achieve communication. Sometimes, the artist may prefer to create complex and ambiguous scenarios for the listener or simply try to discover the composer's intention. However, questionnaire research indicates that musicians commonly attempt to express specific emotions in their performances (Lindström *et al.*, 2003).
5. It should be apparent that this procedure requires that the performer momentarily sets aside his or her concerns with the composer's intentions regarding the piece for the sake of the practice. As it happens, performers usually find it interesting and thought-provoking to attempt many different interpretations of the same piece of music.

## Further information and reading

Musicians, teachers, researchers, and others interested in receiving more information about the CFB software are encouraged to contact the authors. For further information, see the "Feedback-learning of Musical Expressivity" website: *www.psyk.uu.se/hemsidor/musicpsy.*

## References

Bach, C. P. E. (1778/1985). *Essay on the True Art of Playing Keyboard Instruments* (W. J. Mitchell, Trans.). London: Eulenburg Books. (Original work published in 1778).

Balzer, W. K., Doherty, M. E., & O'Connor, R. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin, 106*, 410–433.

Barten, S. S. (1998). Speaking of music: The use of motor-affective metaphors in music instruction. *Journal of Aesthetic Education, 32,* 89–97.

Behrens, G. A., & Green, S. B. (1993). The ability to identify emotional content of solo improvisations performed vocally and on three different instruments. *Psychology of Music, 21,* 20–33.

Boyd, J., & George-Warren, H. (1992). *Musicians in Tune: Seventy-Five Contemporary Musicians Discuss the Creative Process.* New York: Fireside.

Bresin, R., & Friberg, A. (2000). Emotional coloring of computer-controlled music performance. *Computer Music Journal, 24,* 44–62.

Briggs, P. (1987). Usability assessment for the office: Methodological choices and their implications. In M. Frese, E. Ulich, & W. Dizda (Eds.), *Human Computer Interaction in the Workplace* (pp. 381–401). Amsterdam: Elsevier.

Brodsky, W., Sloboda, J. A., & Waterman, M. G. (1994). An exploratory investigation into auditory style as a correlate and predictor of music performance anxiety. *Medical Problems of Performing Artists, 9,* 101–112.

Brunswik, E. (1956). *Perception and the Representative Design of Experiments.* Berkeley, CA: University of California Press.

Budd, M. (1985). *Music and the Emotions: The Philosophical Theories.* London: Routledge.

Casals, P. (1970). *Joys and Sorrows.* London: MacDonald.

Clarke, E. F. (1988). Generative principles in music performance. In J. A. Sloboda (Ed.), *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition* (pp. 1–26). Oxford: Clarendon Press.

Cook, N., & Dibben, N. (2001). Musicological approaches to emotion. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research* (pp. 45–70). Oxford: Oxford University Press.

Cooksey, R. W. (1996). *Judgment Analysis.* London: Academic Press.

Davidson, J. W., & Correia, J. S. (2002). Body movement. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 237–250). Oxford: Oxford University Press.

Davies, S. (1994). *Musical Meaning and Expression.* Ithaca, NY: Cornell University Press.

Dickey, M. R. (1992). A review of research on modeling in music teaching and learning. *Bulletin of the Council for Research in Music Education, 113,* 27–40.

Dubal, D. (1985). *The World of the Concert Pianist: Conversations with 35 Internationally Celebrated Pianists.* London: Victor Gollancz.

Eibl-Eibesfeldt, I. (1989). *Human Ethology.* New York: Aldine.

Epstein, D. (1995). A curious moment in Schumann's Fourth Symphony: Structure as the fusion of affect and intuition. In J. Rink (Ed.), *The Practice of Performance: Studies in Musical Interpretation* (pp. 126–149). Cambridge: Cambridge University Press.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363–406.

Friberg, A., & Sundberg, J. (1999). Does music performance allude to locomotion? A model of final ritardandi derived from measurements of stopping runners. *Journal of the Acoustical Society of America, 105,* 1469–1484.

Gabrielsson, A. (1987). Once again: The theme from Mozart's piano sonata in A major. A comparison of five performances. In A. Gabrielsson (Ed.), *Action and Perception in Rhythm and Music* (pp. 81–103). Stockholm: The Royal Swedish Academy of Music.

Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.), *The Psychology of Music* (2nd ed., pp. 501–602). London: Academic Press.

Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology of Music, 31*, 221–272.

Gilden, D. L. (2001). Cognitive emissions of 1/f noise. *Psychological Review, 108*, 33–56.

Hoffren, J. (1964). A test of musical expression. *Council for Research in Music Education, 2*, 32–35.

Howard, V. A. (1989). Expression as hands-on construction. In H. Gardner & D. Perkins (Eds.), *Art, Mind, & Education* (pp. 133–141). Chicago: University of Illinois Press.

Hursch, C. J., Hammond, K. R., & Hursch, J. L. (1964). Some methodological considerations in multiple-cue probability studies. *Psychological Review, 71*, 42–60.

Johnson, C. M. (1998). Effect of instruction in appropriate rubato usage on the onset timings and perceived musicianship of musical performances. *Journal of Research in Music Education, 46*, 436–445.

Juslin, P. N. (1997a). Emotional communication in music performance: A functionalist perspective and some data. *Music Perception, 14*, 383–418.

Juslin, P. N. (1997b). Perceived emotional expression in synthesized performances of a short melody: Capturing the listener's judgment policy. *Musicæ Scientiæ, 1*, 225–256.

Juslin, P. N. (1997c). Can results from studies of perceived expression in musical performances be generalized across response formats? *Psychomusicology, 16*, 77–101.

Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1797–1813.

Juslin, P. N. (2001). Communicating emotion in music performance: A review and a theoretical framework. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research* (pp. 309–337). Oxford: Oxford University Press.

Juslin, P. N. (2003). Five facets of musical expression: A psychologist's perspective on music performance. *Psychology of Music, 31*, 273–302.

Juslin, P. N., & Laukka, P. (2000). Improving emotional communication in music performance through cognitive feedback. *Musicæ Scientiæ, 4*, 151–183.

Juslin, P. N., & Laukka, P. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion, 1*, 381–412.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129*, 770–814.

Juslin, P. N., & Lindström, E. (2003). *Musical expression of emotions: Modeling composed and performed features.* Manuscript submitted for publication.

Juslin, P. N., & Madison, G. (1999). The role of timing patterns in recognition of emotional expression from musical performance. *Music Perception, 17*, 197–221.

Juslin, P. N., & Persson, R. S. (2002). Emotional communication. In R. Parncutt & G. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 219–236). Oxford: Oxford University Press.

Juslin, P. N., & Sloboda, J. A. (Eds.). (2001). *Music and Emotion: Theory and Research.* Oxford: Oxford University Press.

Juslin, P. N., Friberg, A., & Bresin, R. (2002). Toward a computational model of expression in music performance: The GERM model. *Musicæ Scientiæ, Special Issue 2001–2002,* 63–122.

Juslin, P. N., Friberg, A., Karlsson, J., Lindström, E., & Schoonderwaldt, E. (2003). *Once more with a feeling: Feedback-learning of expressivity in music performance.* Manuscript submitted for publication.

Laukka, P. (2004). Instrumental music teachers' views on expressivity: A report from music conservations. *Music Education Research, 6,* 45–56.

Lehmann, A. C. (1997). Acquired mental representations in music performance: Anecdotal and preliminary empirical evidence. In H. Jørgensen & A. C. Lehmann (Eds.), *Does Practice Make Perfect? Current Theory and Research on Instrumental Music Practice* (pp. 141–163). Oslo: Norwegian State Academy of Music.

Lindström, E. (2003). The contribution of immanent and performed accents to emotional expression in short tone sequences. *Journal of New Music Research, 32,* 269–280.

Lindström, E., Juslin, P. N., Bresin, R., & Williamon, A. (2003). "Expressivity comes from within your soul": A questionnaire study of music students' perspectives on expressivity. *Research Studies in Music Education, 20,* 23–47.

Lippman, E. A. (1953). Symbolism in music. *Musical Quarterly, 39,* 554–575.

Marchand, D. J. (1975). A study of two approaches to developing expressive performance. *Journal of Research in Music Education, 23,* 14–22.

Menuhin, Y. (1996). *Unfinished Journey.* London: Methuen.

Meyer, L. B. (1956). *Emotion and Meaning in Music.* Chicago: University of Chicago Press.

Palmer, C. (1996). Anatomy of a performance: Sources of musical expression. *Music Perception, 13,* 433–453.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48,* 115–138.

Persson, R. S. (1993). *The Subjectivity of Musical Performance: An Exploratory Music-Psychological Real World Enquiry into the Determinants and Education of Musical Reality.* Unpublished doctoral dissertation, University of Huddersfield, Huddersfield, UK.

Persson, R. S. (1996). Concert musicians as teachers: On good intentions falling short. In A. J. Cropley & D. Dehn (Eds.), *Fostering the Growth of High Ability: European Perspectives* (pp. 303–320). Norwood, NJ: Ablex.

Raffman, D. (1993). *Language, Music, and Mind.* Cambridge, MA: MIT Press.

Ratner, L. G. (1980). *Classic Music: Expression, Form, and Style.* New York: Schirmer.

Riley-Butler, K. (2001). *Teaching Expressivity through Feedback and Replication.* Paper presented at the Meeting of the Society for Music Perception and Cognition, Kingston, Canada.

Rohwer, D. (2001). Instrumental music students' cognitive and performance understanding of musical expression. *Journal of Band Research, 37,* 17–28.

Rostwall, A.-L., & West, T. (2001). *Interaktion och Kunskapsutveckling* [Interaction and Learning. *A Study of Music Instrument Teaching*]. Unpublished doctoral dissertation, KMH Förlaget, Stockholm.

Shove, P., & Repp, B. H. (1995). Musical motion and performance: Theoretical and empirical perspectives. In J. Rink (Ed.), *The Practice of Performance: Studies in Musical Interpretation* (pp. 55–83). Cambridge: Cambridge University Press.

Sloboda, J. A. (1983). The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology, 35A*, 377–396.

Sloboda, J. A. (1996). The acquisition of musical performance expertise: Deconstructing the "talent" account of individual differences in musical expressivity. In K. A. Ericsson (Ed.), *The Road to Excellence: The Acquisition of Expert Performance in the Arts and Sciences, Sports, and Games* (pp. 107–126). Mahwah, NJ: Erlbaum.

Sloboda, J. A., Minassian, C., & Gayford, C. (2003). Assisting advanced musicians to enhance their expressivity—An intervention study. In R. Kopiez, A. C. Lehmann, I. Wolther, & C. Wolf (Eds.), *Proceedings of the Fifth Triennial ESCOM Conference* (p. 92). Hanover, Germany: Hanover University of Music and Drama.

Spencer, H. (1857). The origin and function of music. *Fraser's Magazine, 56*, 396–408.

Stanislavski, C. (1937/1988). *An Actor Prepares.* London: Methuen. (Original work published in 1937).

Swanwick, K. (1985). *A Basis for Music Education.* Windsor, UK: NFER-Nelson.

Tait, M. (1992). Teaching strategies and styles. In R. Cowell (Ed.), *Handbook of Research on Music Teaching and Learning* (pp. 525–534). New York: Schirmer.

Taylor, J. A. (1981). *Introduction to Computers and Computer-Based Instruction in Music.* Tallahassee, FL: Center for Music Research, Florida State University.

Webster, P. R. (2002). Computer-based technology and music teaching and learning. In R. Colwell & C. Richardson (Eds.), *The New Handbook of Research on Music Teaching and Learning* (pp. 416–439). Oxford: Oxford University Press.

Woody, R. H. (1999). The relationship between explicit planning and expressive performance of dynamic variations in an aural modeling task. *Journal of Research in Music Education, 47*, 331–342.

Woody, R. H. (2000). Learning expressivity in music: An exploratory study. *Research Studies in Music Education, 14*, 14–23.

CHAPTER 14

# DRUGS AND MUSICAL PERFORMANCE

ROBERT WEST

Musical performance can be among the most difficult and demanding of human activities and generally occurs in situations when the personal cost of failure or even suboptimal performance is very high. It is not surprising, therefore, that individuals who earn a living by performing often turn to artificial aids to help them cope either emotionally or physically (Wesner *et al.*, 1990). Use of "psychotropic" drugs (i.e. drugs that affect the mind and behavior) has been linked to music certainly for centuries and probably since both were discovered by our ancestors. The question is: can they genuinely help the performer? If the answer to this is yes, we must also ask who might benefit and is there a price to pay? This chapter provides a wide-ranging review of psychotropic drug effects and their implications for musical performance. There is actually very little direct evidence linking the two, but research on drug effects more generally can provide useful clues and practical guidance for the performing musician.

## 14.1 Musical performance as a psychomotor task

Before scrutinizing drug effects, it is worth briefly considering musical performance as a "psychomotor task;" this should help to set these effects in context. Psychomotor performance refers to the execution of tasks requiring physical actions to achieve a predetermined goal (Lemon, 1999; see also Chapter 8 of this volume). At its simplest, it can involve pressing a button as quickly as possible in response to a tone or a light, and it includes tasks as varied as typing and driving a car. Musical performance is a particular kind of psychomotor task which is self-initiated (i.e. the musician starts when he or she is ready) and largely self-contained (i.e. other than group improvisation, it involves executing an internal plan rather than responding to a changing environment). Figure 14.1 shows a simplified overview of this kind of task. Schlaug (2001) provides a very interesting review of the development of musical skill and how this relates to changes in brain function.

Psychomotor performance involves a wide range of component activities or processes, each of which may in principle be affected in different ways by different drugs. Musical information needs to be stored, recalled, or derived from a score or extemporized. Sensory information needs to be monitored and used to initiate and control action sequences within certain parameters. Most obviously, there is a need for fine control of one's muscles so that notes are played or sung as intended with the

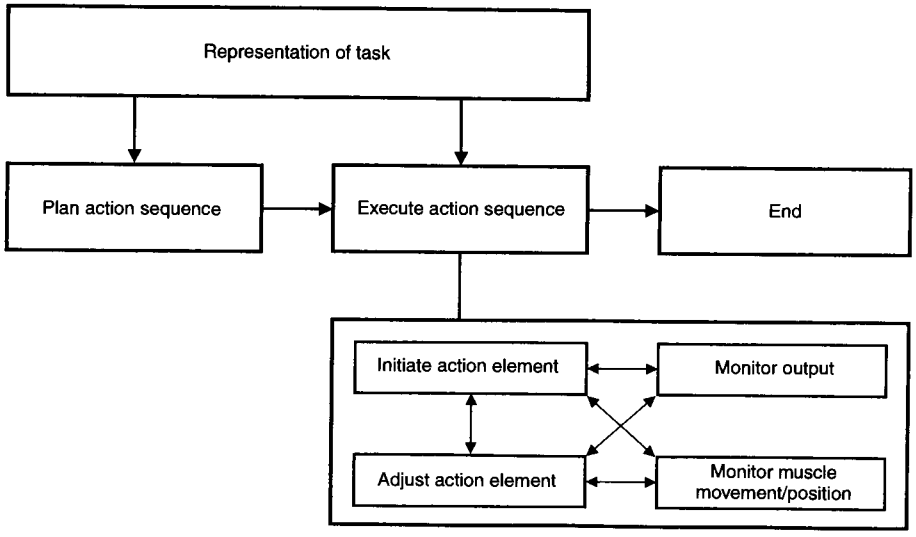



**Figure 14.1** Elements of psychomotor performance for self-initiated, self-contained tasks.

required quality and intonation. This requires very precise and often very rapid movements. All of this takes place in a context where the performer knows he or she is being subject to evaluation, possibly by hundreds, thousands, or even millions of people. Performers also put great pressure on themselves to achieve what they consider to be perfection.

### 14.1.1 The effects of anxiety on psychomotor performance

It is well recognized that a little anxiety can improve performance but that a high level can be anything from mildly unpleasant to disastrous (Fredrikson & Gunnarsson, 1992). Individuals differ in their propensity to experience anxiety; this manifests itself first in childhood and does not change much over the years (Chorpita & Barlow, 1998). However, individuals can be more or less successful in finding ways of coping with it or masking or suppressing it. Simple task elements involving rapid responses can benefit from even quite high levels of anxiety but more complex task elements requiring problem solving or fine muscle control can be badly affected (see Table 14.1). Anxiety can also affect the way information is processed. Anecdotally, there are reports that it tends to expand one's sense of time—pauses can seem very long.

All of this can give some guide to the way in which drugs might affect performance, but in practice, much depends on the dose, the dosing methods, and contextual and individual factors. Indeed, the only way to know for sure whether a drug will have a particular effect is to test it in the context in which it will be used in people who are likely to use it. Unfortunately, there is very little research of this kind in the literature. However, current evidence is provided below and conclusions are drawn where possible.

Table 14.1 Aspects of musical performance, and their relationship with anxiety and level of arousal

| Activity | Effect of high anxiety | Optimum arousal level |
|---|---|---|
| Recall of musical information | Interferes | Low |
| Perception of time | Expands | Moderate |
| Coordination of movement | Interferes | Moderate |
| Fine control of muscles | Interferes (e.g. tremor) | Moderate |
| Monitoring of muscle activity | Not known | Not known |
| Monitoring of sound output | Not known | Not known |
| Monitoring of external sounds | Not known | Not known |
| Speed of movement | Increased | High |

## 14.2 Types of drug

Drugs that affect the mind or performance can be classified in many ways. One system for classifying drugs is according to their effects on mental state. Drugs can be thought of as "stimulant" if they increase alertness, wakefulness, or agitation. They are "sedative" if they have the opposite effects. Drugs that relieve anxiety are said to have "anxiolytic" properties, just as drugs that help relieve depressed moods are known as "antidepressants." Some drugs, independent of stimulant or sedative properties, induce a feeling of well-being; these have "euphoriant" actions. Some drugs can induce distorted perceptions; these are "hallucinogens."

Drugs are also classified according to their effects on physiological systems. For example, drugs that mimic the body's naturally occurring neurotransmitters (i.e. chemicals released that transmit a signal between nerve cells and, thereby, enable our nervous system to encode and process information and function) are labeled by the neurotransmitter they mimic: "cholinergic" drugs mimic a neurotransmitter called acetylcholine; "dopaminergic" drugs mimic dopamine; "adrenergic" drugs mimic adrenaline, and so on.

This section starts with the most commonly used licit, nonprescription drugs (alcohol, caffeine, and nicotine), then considers prescription drugs (antidepressants, beta-blockers, and tranquilizers), and finally drugs that are used illicitly and are subject to abuse (amphetamines, cocaine, ecstasy, hallucinogens, and opiates). Of course, there are many other drugs that could be considered, but the ones reviewed here are probably the most significant for performers.

### 14.2.1 Lifestyle drugs

*Alcohol*

Alcohol is very much a part of many musicians' lives. This may be partly because of lifestyle or possibly the personality characteristics of people who become musicians. However, there are numerous anecdotal reports of musicians who consciously use alcohol to help them cope with the stress of performing. There is little doubt that, acutely, alcohol can "dull the senses;" it can also ease feelings of anxiety. Unfortunately, the intoxicating and sedating effects of alcohol impair performance on most tasks

(Ellinwood *et al.*, 1981; Horne & Gibbons, 1991; Grant *et al.*, 2000; Farquhar *et al.*, 2002). Alcohol also increases impulsiveness (Nagoshi *et al.*, 1991). Intoxicated individuals are often not aware of the level of impairment because alcohol appears to disrupt this perception as well (Lubin, 1977; Ridderinkhof *et al.*, 2002). Moreover, consuming alcohol in large quantities over a period of time leads to a chronic increase in anxiety and neuropsychological problems (Eckardt *et al.*, 1980; Capitani *et al.*, 1983; Schaeffer *et al.*, 1989; Beatty *et al.*, 2000; Di Sclafani *et al.*, 2002).

In this way, musicians who start using alcohol to calm their nerves before performing can become trapped in a downward spiral, in which they come more and more to rely on alcohol and without it become ever more anxious. This is compounded, of course, by the fact that alcohol adversely affects technique. Coordination, memory, and muscle control are impaired in a dose–response manner (Burns & Moskowitz, 1980; Chait & Perry, 1994). For performers who experience severe anxiety about performing, there may be a difficult line to tread between the need to control the nerves enough to perform and the effect of alcohol on that performance.

Of equal concern is that excessive drinking, particularly bingeing (i.e. drinking large quantities on a single occasion), is damaging to physical health. Consistently drinking above 21 units per week by a man or 14 units by a woman increases risk of liver damage, heart disease, and other organ damage (a unit is 330 ml of ordinary strength beer, a medium sized glass of wine, or a single measure of spirit). In addition, excessive drinking is linked with an increased risk of accidental injury or death and involvement in violence both as a victim and a perpetrator. These latter effects are probably linked to the effect of alcohol in removing inhibitions (de Wit *et al.*, 2000). This loss of inhibition is a double-edged sword because many drinkers, particularly if they are socially anxious, feel that they are funnier and more attractive when they drink (and to be fair, some are).

In fact, moderate drinking probably has beneficial psychological and physical effects. It seems to reduce risk of heart disease, possibly by reducing the amount of low-density lipoprotein (a fat that damages the arteries) in the blood. Psychologically, it is linked with greater social integration and feelings of adjustment. Some have argued that alcohol can increase creativity. This has not been studied much, but an interesting study of well-known writers and musicians suggests that for most of them alcohol had the opposite effect (Ludwig, 1990).

Alcohol dependence is quite common, affecting 5–10% of the adult population in the USA and UK, and more in countries such as Russia. It is characterized by compulsive use, strong cravings, and drinking to relieve withdrawal symptoms. These symptoms include anxiety and often physical symptoms such as tremor and even on occasions seizure. It has a strong genetic component (Nestler, 2000). If one or more parent is or has been alcohol dependent, the offspring are significantly more likely to develop problems of the same kind. It is very difficult for a musician to perform to an acceptable standard if he or she is severely alcohol dependent, and eventually, something has to give.

Treatment for alcohol dependence involves an initial period of "detoxification" in which abstinence from alcohol is made more comfortable with the use of drugs such as benzodiazepines (Edwards *et al.*, 2003). The detox can be undertaken with the patient

as either an outpatient or an inpatient, depending on the person's life circumstances and the severity of dependence. After this, the patient is usually given an extended period of counseling, often supplemented by the drug Campral (acamprosate) which can help reduce the cravings. While many alcoholics respond well to this treatment protocol, relapse is very common, and for a large proportion of alcoholics (but not all), their condition is best conceived of as a chronic condition that can be improved for quite long periods but is ever-present. The goal of the sufferer, as well as of various treatment agencies, is to keep it at bay as often and for as long as possible. It has been suggested that it is possible for people who are alcohol dependent to resume a life of "controlled drinking" rather than total abstinence; while there are clearly instances where this has occurred, it is rare.

Where all this leaves the musician is debatable. The majority can gain the benefits of alcohol while retaining control over their drinking; however, onset of dependence is insidious. The alcohol "abuser" is typically in no position to detect the downward spiral, perhaps because alcohol is affecting his or her judgment. Given the strong genetic contribution to alcohol dependence, perhaps an individual could use his or her own family history as a guide to assessing how far he or she can go with alcohol consumption before running into difficulties. Another approach is to monitor one's drinking—to ensure that one goes at least a few days in any given month without drinking and to ensure that one does not exceed the maximum recommended level over a period of time. Musicians who feel the need for something to relax them should consider other, less harmful approaches, including exercise (see Chapter 9), counseling, and prescription medications (see Section 14.2.2 below).

*Caffeine*

Caffeine is a mild central nervous system stimulant (Battig, 1991; Donovan & DeVane, 2001; Mandel, 2002). It is a member of the group of chemically related alkaloids known as methylxanthines. In moderate doses (i.e. less than 400 mg per day), which can be obtained from no more than a few cups of coffee, it can improve vigilance and endurance and reduce fatigue (Tarnopolsky, 1994; Spriet, 1995; Graham, 2001; Smith, 2002). A typical cup of "real" coffee contains about 100 mg of caffeine, and soft drinks such as colas contain up to 50 mg. At these doses, it has not been found to have adverse health effects (Nawrot *et al.*, 2003). In fact, there is some evidence that it can have a modest effect in opening the airways and easing breathing difficulties. However, in higher doses, it can increase feelings of anxiety and has been reported to impair fine muscle control. It can also have adverse effects on the cardiovascular system and calcium balance. In women, there is some evidence that in high doses it may affect fertility and fetal development. Thus, while it may be invaluable in keeping the musician awake and alert, it is not necessarily harmless.

Heavy caffeine use over a long period can lead to withdrawal symptoms during abstinence (Griffiths & Chausmer, 2000; Dews *et al.*, 2002). A common symptom is headache, although drowsiness and other symptoms are also reported. It has been reported that a proportion of regular caffeine users develop cravings and show signs of compulsive use, but there is little evidence of this from controlled studies. Given that

caffeine is relatively benign, the pressure on users to abstain is not great and so it is not clear what proportion of users would find it impossible to abstain if they had to.

For the musician, it makes sense to consider caffeine in moderate doses as a potentially useful tool when alertness needs to be maintained or to mitigate the effect of alcohol (Liguori & Robinson, 2001; Drake *et al.*, 2003), but it would probably be unwise to maintain a high dose for long periods or, obviously, to use the drug if it appears to be impairing ability to execute fine movements.

*Nicotine*

Nicotine is very rapidly absorbed from cigarette smoke (West & Shiffman, 2003). It is also obtained from other tobacco products such as cigars and pipes. A smoker will typically absorb 10–20 mg of nicotine each day—about 1 mg from each cigarette. Nicotine is also absorbed, albeit more slowly, from smokeless tobacco products such as nasal or oral snuff and chewing tobacco. There is also a wide range of pure nicotine delivery systems available such as nicotine chewing-gum, inhaler, skin patch, lozenge, and nasal spray. These are marketed as aids to stopping smoking but, in principle, can also be used instead of smoking; pure nicotine does not cause cancer, so they are much less harmful than cigarettes.

Nicotine affects the central and peripheral nervous system by partially mimicking the effects of a naturally occurring neurotransmitter called acetylcholine (RCP, 2000). It has complex effects on the body, and there are many misunderstandings about what these are. Its most noticeable effects are essentially "stimulant;" causing an increase in heart rate and blood pressure, sweating, a slight "buzz," and feelings of agitation. Nicotine also causes feelings of nausea in individuals who are not used to it or who take more than they are used to.

The hands of smokers typically shake more than those of nonsmokers. When smokers stop, the tremor reduces within a few hours as nicotine leaves the body. Of course, nervousness at performing in public also typically causes tremor. It may be that the effect of nicotine would aggravate this problem, but that has not been tested.

There is some research to show that nicotine can increase the speed of simple movement tasks such as tapping on a keyboard (West, 1993). It is plausible that this would help with passages that require fast movement. It is also possible that nicotine withdrawal would cause problems with this, but no scientific studies have been conducted on this.

Musicians who smoke might be concerned that they would experience an increase in anxiety when they stop. Contrary to what many people believe, smoking does not typically calm the nerves (West, 1993). The idea that it calms the nerves probably arises out of the fact that, when abstaining, smokers experience unpleasant withdrawal symptoms, such as irritability and restlessness, and smoking makes these go away. Strangely, although anxiety itself is listed as a nicotine withdrawal symptom in the American Psychiatric Association diagnostic manual, many studies actually find a decrease in anxiety within a week of stopping smoking and an increase if the smoker relapses (West & Hajek, 1997). One thing is clear; musicians who stop smoking can expect to experience less chronic stress than if they continue, quite apart obviously from the health benefits which accrue.

One potentially important nicotine withdrawal symptom is impaired ability to concentrate (RCP, 2000). Obviously this could be a problem for musicians, but it can be ameliorated by using one of the pure nicotine delivery systems that are available.

Smokers who fail to stop smoking experience diseases of old age about 12 years earlier than nonsmokers and die an average of 8 years earlier. Lung cancer and heart disease are well known as two of the causes of death. Less well-known is "chronic obstructive pulmonary disease," in which there is a progressive deterioration in ability to breathe. Any smoker who is beginning to experience breathlessness or a chronic cough should stop immediately, as they may well be on a steep decline in lung function. The good news is that stopping halts the accelerated decline straight away. For people who play wind instruments, the importance of maintaining lung function is obvious.

Musicians who perform in clubs and pubs can be exposed to quite high levels of smoke from other people's cigarettes. Although the effects of such exposure are considerably less than from active smoking, there is an increase in risk from lung cancer and heart attack. Passive smoking can also trigger asthma attacks and other respiratory problems. This has led to legislation to protect nonsmokers from exposure to tobacco smoke in some countries, and in the future, it is likely that smoking restrictions will extend further, affecting clubs and music venues.

In summary, it is unlikely that tobacco use significantly improves musical performance, but in addicted users, there may be impairment during periods of abstinence. Pure nicotine products may mitigate these effects. Musicians, like any other smokers, would be well advised for the sake of their mental and physical health to try their best to stop and keep trying until they succeed. They should use proven treatment approaches, including nicotine replacement therapies (gum, patch, etc.), bupropion (Zyban), and behavioral support. In the UK, one should simply be able to go to a general practitioner to get an immediate referral to the specialist stop smoking treatment service, which is paid for by the national health service.

### 14.2.2 Prescription drugs

*Antidepressants*

Like anyone else, musicians can suffer from depression. In fact, depression as a clinical syndrome is much more common than most people realize. About 1 in 7 people will suffer at least one bout in their lives, and many are plagued by repeated occurrences. One of the common symptoms is an inability to motivate oneself to do anything, in many cases even to get out of bed. This is accompanied by feelings that one is worthless as a person, sometimes irritability, and often suicidal thoughts; one is overwhelmed by a sense of hopelessness. Obviously, this is not good news for the performing musician, who has to get up on stage and ply his or her craft.

There are several different classes of antidepressant on the market. One of these is the *tricyclics*; the name derives from their chemical structure which includes three hydrocarbon rings. Commonly used examples are imipramine (Tofranil) and nortriptyline (Pamelor, Aventyl). They can help about 20–30% of users to recover within a couple of months who would not have done otherwise (surprising as it may seem, about a quarter to a third of depressive episodes resolve within about 3 months by

themselves). They can also be effective against anxiety. However, they are not without side-effects. The most common are dry mouth, a slight tremor, fast heartbeat, constipation, sleepiness, and weight gain. Depending on the instrument one plays, these could in principle affect performance, and this needs to be taken into account by the prescribing doctor.

The second and probably most popular class of antidepressants are called *selective serotonin re-uptake inhibitors* (SSRIs). Probably the best known of these is fluoxetine (Prozac), but there are many others, such as paroxetine (Seroxat) and fluvoxamine (Luvox). These do not seem to be more effective in treating depression than the older tricyclic antidepressants, but users generally find the side-effects more tolerable. In the first couple of weeks of taking these, some users feel sick and some more anxious. They can also interfere with sexual functioning. However, these symptoms, if they occur, generally resolve after a few weeks. Paroxetine and other drugs in this class have recently come under scrutiny because of concerns that they might cause anxiety disturbances or even suicide attempts in some users.

The third main class of antidepressants in common use is the *monoamine oxidase inhibitors* (MAO inhibitors). These block the breakdown of the neurotransmitters dopamine, noradrenaline, and serotonin, making them more readily available in the central nervous system. These drugs can be quite effective, and there is little evidence that they impair psychomotor performance (Tiller, 1990).

In general, antidepressants clearly have a potential role for musicians who are suffering from clinical depression or anxiety, and the effect on performance is likely to be modest (Hindmarch, 1992; Linnoila *et al.*, 1993). However, there is no particular reason for a musician to use an antidepressant as a performance-enhancing medication or to control anxiety specifically related to performance.

### *Beta-blockers*

Beta-blockers are drugs that block the action of a naturally occurring neurotransmitter, noradrenaline, in the autonomic nervous system (see Chapter 1). These drugs are available only on prescription and are not suitable for people who have low blood pressure, bronchial asthma, or some other conditions. They were developed to control high blood pressure. Apart from reducing blood pressure, they reduce other physical symptoms of overarousal and anxiety, including tremor (Brantigan *et al.*, 1978, 1979, 1982). For this reason, many musicians regularly take beta-blockers to control "stage fright" (Neftel *et al.*, 1982). In fact, one survey found that about 20% of orchestral musicians in the USA took beta-blockers for performance anxiety. One can take the pill an hour or so before an important performance, and the effect lasts several hours. There is reasonably strong evidence that they are indeed quite effective at this and can improve performance as a result (Lidén & Gottfriess, 1974; James *et al.*, 1977, 1978, 1983; Pearson & Simpson, 1978; James, 1984; James & Savage, 1984; Gates *et al.*, 1985). However, these drugs do not deal with the emotional experience of anxiety and should not be considered a panacea by any means.

There are quite a few different brands of beta-blocker which differ slightly in their action. Drugs such as metoprolol (Lopresor) and atenolol (Tenormin) are "beta-1 selective". This means they block the beta-1 form of the noradrenaline receptor, which

is found primarily in the heart, but not the "beta-2" receptors, which are found in the lungs. As a result, they are less likely to cause bronchoconstriction and breathing difficulties. Propranolol (Inderal) is a beta-1 and beta-2 antagonist. A dose of 40 mg (less than that typically used to control blood pressure) is recommended to control "situational anxiety".

Like any drug, beta-blockers have side effects, although these are generally mild. In someone who has normal blood pressure, they can cause dizziness from low blood pressure, particularly when suddenly getting up ("postural hypotension"). They can also cause tiredness or reduce physical endurance (Kaiser *et al.*, 1981; Bengtsson, 1984; Cupido *et al.*, 1994) and may have a modest effect on speed of reactions (Broadhurst, 1980). Beta-blockers are not known to be addictive.

In summary, views vary about the appropriateness of using beta-blockers to ease symptoms of anxiety and help performance. Taken infrequently (e.g. before important auditions) and under the supervision of a doctor, they may benefit many musicians. However, it is also important to consider other methods of controlling anxiety at its source. This can include exercise, self-help guides, and cognitive-behavioral therapy (Clark & Agras, 1991; see Chapters 9 and 12 of this volume).

*Tranquilizers*

The most commonly used tranquilizers today are benzodiazepines such as diazepam (Valium) and chlordiazepoxide (Librium). These are used to control anxiety (anxiolytics) or aid sleep (hypnotics). They work by enhancing the inhibitory actions of a naturally occurring neurotransmitter called GABA. In this respect, they are quite similar to alcohol.

There are many different drugs in this class, some of which are short-acting and some longer-acting. They differ in the extent to which they cause sedation (de Wit *et al.*, 1993). In general, their effects on ability to perform complex tasks and tasks requiring vigilance are as one would expect from sedative drugs: they impair it (Erwin *et al.*, 1986; Kunsman *et al.*, 1992; Simpson & Rush, 2002). Impairment is made worse when alcohol is taken as well (Lister & File, 1983). It is possible that, in individuals with very high levels of anxiety, they could lead to a net improvement by combating the deterioration caused by that anxiety, although this has not been adequately tested.

Benzodiazepines are also used illicitly as recreational drugs for their subjective effects (Evans *et al.*, 1990). Concerns have also been expressed that users can become dependent on them, and it is true that there is often a rebound increase in anxiety and sleeplessness when people come off them. However, it is not clear how much this is just a re-emergence of a problem that the drugs have been suppressing. In that event, it seems that there is a decision to be made as to whether the benefits of continuing with long term use of the drugs outweigh the costs.

Because of their sedative properties, benzodiazepines are probably not very suitable for performance anxiety. However, they can have a role to play in treating clinical anxiety disorders or dealing with anxiety associated with an acute traumatic event such as a bereavement.

Newer tranquilizers have been developed that seem to be suffer less from problems of performance impairment and sedation (Mattila *et al.*, 1982). Most notable among these is buspirone (Goldberg, 1984). If a drug is needed to treat clinical anxiety, it is

certainly worth considering one of these modern anxiolytics. Antidepressants (see above) are also commonly used nowadays in the treatment of anxiety disorders. These have a range of side effects but little effect on psychomotor performance.

### 14.2.3 Illicit drugs

Performers appear more likely than the general population to take illicit drugs, particularly those musicians who are involved in jazz, rock, blues, and other modern idioms. In many cases, the usage is recreational, but in others, the motivation may be related to musical performance *per se*. The drugs discussed in this section (listed alphabetically) are classified as illicit because they are frequently used illicitly (i.e. their possession or use without authority constitutes a criminal offence). However, this is for presentational purposes only; in many cases there are also licit uses, and some drugs considered in previous sections can also be used illicitly.

*Amphetamines*

Amphetamines are stimulants that, like caffeine, increase feelings of agitation and arousal (Koelega, 1993). They work by stimulating release of the neurotransmitters dopamine and noradrenaline in the central nervous system and also blocking their re-uptake back into the nerve terminals, making them more available. They are chemically similar to the drug ephedrine, which was once used as a decongestant and diet pill. The three main types in use are amphetamine sulfate (Benzedrine), dexamfetamine (Dexedrine), and methamphetamine (Methedrine). They can improve endurance and performance on tasks that are physically demanding, and there is some evidence that they can mitigate the effect of alcohol on psychomotor performance (Perez-Reyes *et al.*, 1992). Unlike caffeine they can also produce a strong feeling of well-being and increase self-confidence. These are attractive attributes for musicians, who have a demanding schedule or feel that they need to maintain a high energy level in their performance. Amphetamines can help with breathing disorders because they dilate the bronchioles. They have also been used as weight loss pills. Unfortunately, sessions of amphetamine use are often followed by a "crash" or period of feeling depressed and tired. This can act as a stimulus to take more of the drug.

Amphetamines can be ingested in a variety of ways but probably most commonly as pills. "Ice" is recrystallized methamphetamine hydrochloride and is a potent stimulant. It generally takes the form of clear crystallized chunks that can be smoked. Amphetamines can also be sniffed and even injected.

Unfortunately, amphetamines can also create feelings of paranoia and stimulate aggressive behavior (Schiorring, 1981; de Wit *et al.*, 2000). Some users become addicted to them, although the pattern of addiction is generally not the same as for nicotine and alcohol, where there is regular use throughout the day. Amphetamine use more commonly involves bingeing on large quantities on specific occasions, often at weekends. Heavy amphetamine use can lead to psychosis and to serious physical effects including cardiac arrest.

Although it may be tempting to use amphetamines to keep up one's energy levels, maintain an attractive weight, and to feel good, for many users the harm outweighs the benefits, and musicians should treat these drugs with caution.

*Cannabis*

Cannabis is the most widely used illicit drug in the world, and of course, it is a staple among entertainers and musicians of all kinds. About 30% of adults in the UK have tried cannabis, and about 7% are current users. The average of lifetime use in Europe is around 20%, with about 7% being current users; figures for other English-speaking countries are very similar. The active ingredient in cannabis is tetrahydrocannabinol or THC. This produces a range of effects that include a sense of well-being, a feeling of relaxation, and enhanced sociability (Heishman *et al.*, 1997). It also results in increased appetite, difficulty in concentrating, a distorted sense of time, and altered perceptions, and can in high doses lead to hallucinations (Perez-Reyes *et al.*, 1988; Chait & Perry, 1994; Liguori *et al.*, 2002). Physical effects include increased heart rate and decreased muscle tone. The psychomotor impairment caused by cannabis is modest (Pickworth *et al.*, 1997) but exacerbated when consumed with alcohol (Perez-Reyes *et al.*, 1988; Liguori *et al.*, 2002).

There is a heated debate about the harmfulness of this drug. Physically, the main harm would be expected to arise from smoking it, particularly with tobacco. The extent of the risk has yet to be quantified, but it is likely that there is at least some because inhaling smoke particles on a regular basis in itself damages the respiratory system. On the psychological side, the potential for addiction appears to be similar to that for alcohol, with approximately 1 in 10 users developing a clinical dependence syndrome involving compulsive use. There is some evidence that high levels of chronic use can lead to intellectual impairment, although the effects are generally quite subtle (Schwartz *et al.*, 1989). Perhaps of greater concern is the likelihood that use precipitates psychiatric disturbance in susceptible individuals. Users are more likely to develop psychotic symptoms and to attempt suicide than nonusers, even once other predisposing factors have been taken into account.

Advising musicians (particularly hip-hop, pop, jazz, rock musicians, and suchlike) to avoid taking cannabis is probably akin to advising a dog not to bark. However, it is important to be aware that its effects may detract from performance, and it is not without risks.

*Cocaine*

Cocaine is very similar to amphetamines, with possibly a greater potential to induce feelings of well-being. It works in a somewhat similar manner by blocking re-uptake of the neurotransmitters dopamine and noradrenaline into the nerve terminals, thereby increasing their potency. It is also easy to take in rapidly acting forms either through sniffing into the nose (cocaine hydrochloride powder) or through inhaling the "freebase" form produced by heating cocaine hydrochloride and baking soda ("crack"). Subjectively, the main difference from amphetamines is probably that there is less agitation.

Unfortunately, perhaps more than amphetamines, cocaine can be addictive and can come to take over the life of the user. There is some evidence of neurological problems with chronic use (Di Sclafani *et al.*, 2002). Chronic use can also damage the parts of the brain that are involved in feeling pleasure, so that normal sources of pleasure lose

their attractiveness. As with amphetamines, a bout of cocaine use is often followed by a "crash" or period of depression.

Also as with amphetamines, the physical effects of cocaine use include constricted peripheral blood vessels, dilated pupils, and increased temperature, heart rate, and blood pressure. Even in first-time users, cocaine can cause cardiac arrest. It can also interact with alcohol to cause liver damage (Hoyumpa, 1984). The duration of cocaine's psychological effects depends on the route of administration. The faster the absorption, the more intense the "high" but the shorter the duration of action. The high from sniffing may last 15–30 minutes compared with perhaps 5–10 minutes from smoking. As with amphetamines, high doses of cocaine and/or prolonged use can trigger paranoia. Chronic cocaine snorting can result in ulceration of the mucous membrane of the nose and can damage the nasal septum enough to destroy it.

It is possible to recover from problem use of cocaine. As yet there is little scientific evidence that specific treatments improve the chances of recovery, but cognitive-behavioral therapy and contingency management (where rewards and punishments are used to maintain abstinence, checked by urine analysis) have shown some limited effect. A number of medications are also being tried as treatments for cocaine dependence, but at present, there is no clear evidence that they are of lasting benefit.

### *Ecstasy*

Ecstasy (MDMA) is a stimulant quite similar chemically to amphetamines. The main effect on the brain is to increase the release of the neurotransmitters serotonin (5-hydroxytryptamine) and dopamine. The drug remains in the system and is active for several hours. Physiological effects are similar to other stimulants; it can increase alertness and physical energy. In addition, it raises body temperature, and this has been blamed for a number of deaths among users. Users report that taking ecstasy makes them feel more sociable, with heightened sensory awareness. There is slight "depersonalization", an altered sense of time and greater feeling of warmth toward others.

Deaths from ecstasy have achieved prominence in the media, although they are rare considering the number of users. Concerns have been raised about the long-term damage that taking ecstasy might inflict. There is some evidence that ecstasy can cause damage to the nerves that normally produce serotonin, and it has been speculated that this may cause chronic depression in heavy users. Chronic use has also been found to be associated with memory deficits and impairment on tasks involving complex information processing (Simon & Mattick, 2002).

For the musician, the benefits of the increased energy and greater sense of sociability and well-being produced by ecstasy have to be weighed against the significant risk of longer term mental health problems and memory impairment.

### *Hallucinogens*

Hallucinogens (or "psychedelic drugs") are a class of drugs characterized by the sensory distortions that they induce. Best known among these is lysergic acid diethylamide (LSD). These drugs were popular among rock and pop musicians in the 1960s because at that time there was considerable interest in altered states of consciousness

and expanding one's experiences. Recently they have become popular again with young people. The dangers in using them lie in the distorted perception of reality they create during acute intoxication, which can lead to self-harm. In the longer term, there is a risk of a chronic syndrome of unpredictable recurrence of the drug effects ("flashbacks") and possibly longer-term mood disturbance (Strassman, 1984).

Acute intoxication with these drugs makes controlled performance extremely difficult. Some have argued that it might assist with more free-form extemporization, though scientific evidence for this is lacking.

*Opiates*

Opiates are a class of drug made from the opium poppy. They include opium itself, morphine, and diacetylmorphine (diamorphine or heroin). Methadone is a synthetic opiate which is longer acting. These drugs have powerful pain-relieving properties because they mimic the actions of naturally occurring "endogenous" opiates in the brain. They can also produce a powerful feeling of euphoria or well-being, particularly heroin and opium (Hill & Zacny, 2000). Of course, they are also highly addictive.

Heroin is the most commonly used illicit form of the drug in countries such as the UK and USA. It can be sniffed, smoked, or injected. In all cases it carries a risk of death from overdose due to respiratory suppression—the user stops breathing. A large proportion of overdoses occur not because the drug is unusually pure or because it has been taken in larger quantities than normal, but because it is taken with alcohol or benzodiazepines which compound the effect. Overdose is the most common cause of death in heroin users, but death from HIV, hepatitis and other infections, suicide, and homicide are also quite common. A heroin addict is about 30 times more likely to die in a given year than the average adult of the same age.

After a few weeks of regular use the body adapts, and users experience withdrawal symptoms when they abstain. Sometimes overdoses occur when a user has abstained for a while (e.g. when in prison) and then resumed taking the drug without compensating for the loss of tolerance. In a chronic opiate user, abstinence will bring on withdrawal symptoms within a few hours. They peak at about 2–4 days and are usually gone within a week or two. These include gastrointestinal upset, "goose bumps," shivering, yawning, sweating, and muscle cramps. Some have likened them to a bout of flu. They are not life-threatening. Many people think that it is the need to avoid these withdrawal symptoms that keeps people addicted to heroin. Certainly, heroin users report fear of withdrawal symptoms as a major factor keeping them seeking out the drug; however, there is far more to heroin dependence than this. Chronic heroin users continue to get a rush or high from the drug and will typically relapse back to the drug even after they have gotten over the withdrawal symptoms.

There are many famous examples of musicians—particularly in rock, blues, jazz, and pop—who have become addicted to heroin. As with drugs such as cocaine and alcohol, how far this reflects their temperament or their lifestyle is difficult to determine; neither is it clear whether opiates have any beneficial effect in terms of creativity or musical expression. They are probably not as damaging as alcohol to the execution of a performance. Chronic illicit opiate and methadone use has been linked with some intellectual and psychomotor impairment (Mason, 1977; Darke *et al.*, 2000;

Specka *et al.*, 2000; Mintzer & Stitzer, 2002), but patients on long-term opiate treatment for pain control do not appear to be impaired (Fishbain *et al.*, 2003).

However, the control that these drugs come to exert over the individual's life can be so profound that indirectly they make it impossible to continue a musical career. While the situation varies considerably, many heroin addicts are often "busy people." Their life is dominated by acquiring the means to obtain the drug, obtaining it, using it, and then starting the cycle over again. This leaves little time or energy to pursue work and relationships.

Treatment for heroin addiction typically involves trying either to stabilize the patient on prescribed methadone to enable him or her to lead as normal a life as possible or to assist the patient to come off opiates altogether. In the latter case, the regimen consists of medication, such as methadone or buprenorphine, to control withdrawal symptoms, accompanied by counseling to help the patient adapt to life without drugs. The medication is then gradually withdrawn. In practice, this is rarely effective, and it is more common these days (at least in the UK) to try to manage the heroin addict's condition by maintaining him or her on a long-term methadone or buprenorphine program designed to minimize the need to use street heroin. Some treatment centers will provide patients with heroin on prescription, although this is still controversial and whether it confers any benefit over methadone is not clear.

In fact, most heroin addicts also take other drugs such as cocaine, cannabis, benzodiazepines, and amphetamines, and almost all smoke cigarettes. This raises the question of how far to go in treating dependence on these other drugs. In general, drug treatment programs tend to focus on heroin and stimulant use and not worry too much about cannabis, alcohol, or smoking.

For musicians, the direct effect of opiates on performance is probably less of a problem than the more general effect on lifestyle and motivation, coupled obviously with a substantial increase in risk of early death.

## 14.3 Conclusions

Drugs are an important part of all our lives. Society's views on drugs are inconsistent across time and across drugs and have more to do with the way it regards the groups who use the drugs than the effects of the drugs themselves. As performers, musicians have particular traits and needs that seem to make them more likely to use a variety of drugs, either recreationally or to help them cope. The need to maintain a high energy level, performance anxiety, and boredom when traveling all probably contribute to self-medication by musicians. Acutely, drugs such as alcohol, cocaine, and amphetamines no doubt do, on occasion, meet a need on the part of the musician. Unfortunately for many, the cost in terms of dependence and the adverse effects of intoxication is high and sometimes disastrous.

Before closing this chapter with some general observations, it is important to note that how far an individual experiences problems with a drug is related to his or her psychological makeup and life circumstances. Use of a drug such as heroin or alcohol may result in a miserable short life, or the user may be able to retain control and reap the benefits while experiencing few of the adverse effects. Heroin addicts often turn out to

have other psychiatric problems, which would be there even without the addiction, and so do alcoholics. It would be tempting, therefore, to imagine that one could inspect one's own personality and lifestyle and make an informed decision about drug taking. In practice, however, this turns out not to be a very successful strategy. Most of us simply do not have the self-insight to achieve it. The "sensible" advice to performing musicians, therefore, is:

- Drink alcohol in moderation, but avoid intoxication while performing.
- By all means, drink coffee or use caffeine if you need to maintain alertness, but watch out for its effects on your ability to execute fine movements.
- Do not smoke tobacco; it will probably bring on diseases of old age and premature death.
- If you suffer badly from performance anxiety, consider getting a prescription for beta-blockers from your doctor, but also consider exercise and psychological treatments (see Chapters 9 and 12 of this volume).
- Unless you suffer from a clinical anxiety disorder or depression, avoid tranquilizers and antidepressants; even if you do suffer clinical anxiety or depression, consider psychological treatments rather than pills.
- If you use cannabis, it is safer not to smoke it; be aware that it is not a safe drug and that it can be addictive.
- It is best to avoid stimulants such as amphetamines, cocaine, and ecstasy; the acute euphoriant effects may carry a price that you would prefer not to pay in terms of long-term damage to the brain.
- Probably steer clear of hallucinogens; the "expansion of your mind" may carry a price of short- and long-term effects you would want to do without.
- Avoid heroin; it is among the most addictive of all drugs, and there is a good chance it will ruin your life.

Of course, this advice is easier to give than to follow.

## Acknowledgments

I would like to thank Andy McEwen, Andrew Preston, Mike Gossop, and Aaron Williamon for helpful comments on this chapter.

## Further information and reading

For additional information on drugs and addiction, see the websites of the Society for the Study of Addiction (established since 1884) at *www.addiction-ssa.org* or Addiction Search at *www.addictionsearch.com*. Recommended books for further reading on this topic include:

Gossop, M. (2000). *Living with Drugs*. London: Ashgate Publishing.

Heather, N., & Robertson, I. (1997). *Problem Drinking*. Oxford: Oxford University Press.

Shapiro, H. (2003). *Waiting for the Man: The Story of Drugs and Popular Music.* London: Helter Skelter Publishing.

West, R., & Shiffman, S. (2003). *Smoking Cessation.* Oxford: Healthcare Press.

## References

Battig, K. (1991). Coffee, cardiovascular, and behavioral effects current research trends. *Reviews on Environmental Health, 9,* 53–84.

Beatty, W. W., Tivis, R., Stott, H. D., Nixon, S. J., & Parsons, O. A. (2000). Neuropsychological deficits in sober alcoholics: Influences of chronicity and recent alcohol consumption. *Alcoholism: Clinical and Experimental Research, 24,* 149–154.

Bengtsson, C. (1984). Impairment of physical performance after treatment with beta blockers and alpha blockers. *British Medical Journal, 288,* 671–672.

Brantigan, C. O., Brantigan, T. A., & Joseph, N. (1979). The effect of beta blockade on stage fright. *Rocky Mountain Medical Journal, 76,* 227–233.

Brantigan, C. O., Brantigan, T. A., & Joseph, N. (1982). Effect of beta blockade and beta stimulation on stage fright. *American Journal of Medicine, 72,* 88–94.

Brantigan, T. A., Brantigan, C. O., & Joseph, N. (1978). Beta-blockade and musical performance. *Lancet, 2,* 896.

Broadhurst, A. D. (1980). Comparison of effect on psychomotor performance of single doses of propranolol and acebutolol. *Current Medical Research and Opinion, 7,* 33–37.

Burns, M., & Moskowitz, H. (1980). Effects of diphenhydramine and alcohol on skills performance. *European Journal of Clinical Pharmacology, 17,* 259–266.

Capitani, E., Della Pria, M., Doro, G., & Spinnler, H. (1983). Is memory impairment greater than cognitive impairment in moderate chronic alcoholics? *Italian Journal of Neurological Sciences, 4,* 443–449.

Chait, L. D., & Perry, J. L. (1994). Acute and residual effects of alcohol and marijuana, alone and in combination, on mood and performance. *Psychopharmacology, 115,* 340–349.

Chorpita, B. F., & Barlow, D. H. (1998). The development of anxiety: The role of control in the early environment. *Psychological Bulletin, 124,* 3–21.

Clark, D. B., & Agras, W. S. (1991). The assessment and treatment of performance anxiety in musicians. *American Journal of Psychiatry, 148,* 598–605.

Cupido, C. M., Hicks, A. L., McKelvie, R. S., Sale, D. G., & McComas, A. J. (1994). Effect of selective and nonselective beta-blockade on skeletal muscle excitability and fatiguability. *Journal of Applied Physiology, 76,* 2461–2466.

Darke, S., Sims, J., McDonald, S., & Wickes, W. (2000). Cognitive impairment among methadone maintenance patients. *Addiction, 95,* 687–695.

de Wit, H., Dudish, S., & Ambre, J. (1993). Subjective and behavioral effects of diazepam depend on its rate of onset. *Psychopharmacology, 112,* 324–330.

de Wit, H., Crean, J., & Richards, J. B. (2000). Effects of d-amphetamine and ethanol on a measure of behavioral inhibition in humans. *Behavioral Neuroscience, 114,* 830–837.

Dews, P. B., O'Brien, C. P., & Bergman, J. (2002). Caffeine: Behavioral effects of withdrawal and related issues. *Food and Chemistry Toxicology, 40,* 1257–1261.

Di Sclafani, V., Tolou-Shams, M., Price, L. J., & Fein, G. (2002). Neuropsychological performance of individuals dependent on crack-cocaine, or crack-cocaine and alcohol, at 6 weeks and 6 months of abstinence. *Drug and Alcohol Dependence, 66*, 161–171.

Donovan, J. L., & DeVane, C. L. (2001). A primer on caffeine pharmacology and its drug interactions in clinical psychopharmacology. *Psychopharmacology Bulletin, 35*, 30–48.

Drake, C. L., Roehrs, T., Turner, L., Scofield, H. M., & Roth, T. (2003). Caffeine reversal of ethanol effects on the multiple sleep latency test, memory, and psychomotor performance. *Neuropsychopharmacology, 28*, 371–378.

Eckardt, M. J., Ryback, R. S., & Pautler, C. P. (1980). Neuropsychological deficits in alcoholic men in their mid thirties. *American Journal of Psychiatry, 137*, 932–936.

Edwards, G., Marshall, E. J., & Cook, C. C. H. (2003). *The Treatment of Drinking Problems: A Guide for the Helping Professions* (4th ed.). Cambridge: Cambridge University Press.

Ellinwood, E. H., Linnoila, M., Easler, M. E., & Molter, D. (1981). Onset of peak impairment after diazepam and after alcohol. *Clinical Pharmacology and Therapeutics, 30*, 534–538.

Erwin, C. W., Linnoila, M., Hartwell, J., Erwin, A., & Guthrie, S. (1986). Effects of buspirone and diazepam, alone and in combination with alcohol, on skilled performance and evoked potentials. *Journal of Clinical Psychopharmacology, 6*, 199–209.

Evans, S. M., Funderburk, F. R., & Griffiths, R. R. (1990). Zolpidem and triazolam in humans: Behavioral and subjective effects and abuse liability. *Journal Pharmacology and Experimental Therepeutics, 255*, 1246–1255.

Farquhar, K., Lambert, K., Drummond, G. B., Tiplady, B., & Wright, P. (2002). Effect of ethanol on psychomotor performance and on risk taking behavior. *Journal of Psychopharmacology, 16*, 379–384.

Fishbain, D. A., Cutler, R. B., Rosomoff, H. L., & Rosomoff, R. S. (2003). Are opioid-dependent/tolerant patients impaired in driving-related skills? A structured evidence-based review. *Journal of Pain and Symptom Management, 25*, 559–577.

Fredrikson, M., & Gunnarsson, R. (1992). Psychobiology of stage fright: The effect of public performance on neuroendocrine, cardiovascular, and subjective reactions. *Biological Psychology, 33*, 51–61.

Gates, G. A., Saegert, J., Wilson, N., Johnson, L., Shepherd, A., & Hearne, E. M. (1985). Effect of beta blockade on singing performance. *Annals of Otology, Rhinology, and Laryngology, 94*, 570–574.

Goldberg, H. L. (1984). Buspirone hydrochloride: A unique new anxiolytic agent. Pharmacokinetics, clinical pharmacology, abuse potential, and clinical efficacy. *Pharmacotherapy, 4*, 315–324.

Graham, T. E. (2001). Caffeine, coffee, and ephedrine: Impact on exercise performance and metabolism. *Canadian Journal of Applied Physiology, 26* (Supplement), S103–119.

Grant, S. A., Millar, K., & Kenny G. N. (2000). Blood alcohol concentration and psychomotor effects. *British Journal of Anaesthesia, 85*, 401–406.

Griffiths, R. R., & Chausmer, A. L. (2000). Caffeine as a model drug of dependence: Recent developments in understanding caffeine withdrawal, the caffeine dependence syndrome, and caffeine negative reinforcement. *Japanese Journal of Psychopharmacology, 20*, 223–231.

Heishman, S. J., Arasteh, K., & Stitzer, M. L. (1997). Comparative effects of alcohol and marijuana on mood, memory, and performance. *Pharmacology, Biochemistry, and Behavior, 58*, 93–101.

Hill, J. L., & Zacny, J. P. (2000). Comparing the subjective, psychomotor, and physiological effects of intravenous hydromorphone and morphine in healthy volunteers. *Psychopharmacology, 152*, 31–39.

Hindmarch, I. (1992). A review of the psychomotor effects of paroxetine. *International Clinical Psychopharmacology, 6* (Supplement 4), 65–67.

Horne, J. A., & Gibbons, H. (1991). Effects on vigilance performance and sleepiness of alcohol given in the early afternoon ("post lunch") vs. early evening. *Ergonomics, 34*, 67–77.

Hoyumpa, A. M. (1984). Alcohol interactions with benzodiazepines and cocaine. *Advances in Alcohol & Substance Abuse, 3*, 21–34.

James, I. M. (1984). Practical aspects of the use of beta-blockers in anxiety states: Situational anxiety. *Postgraduate Medical Journal, 60* (Supplement 2), 19–25.

James, I., & Savage, I. (1984). Beneficial effect of nadolol on anxiety-induced disturbances of musical performance in musicians: A comparison with diazepam and placebo. *American Heart Journal, 108*, 1150–1155.

James, I. M., Griffith, D. N., Pearson, R. M., & Newbury, P. (1977). Effect of oxprenolol on stage-fright in musicians. *Lancet, 2*, 952–954.

James, I. M., Pearson, R. M., Griffith, D. N. W., Newbury, P., & Taylor, S. H. (1978). The effect of beta-adrenoceptor blockade on the somatic manifestations of anxiety. A study of stage fright. *Journal of Psychosomatic Research, 22*, 327–337.

James, I. M., Burgoyne, W., & Savage, I. T. (1983). Effect of pindolol on stress related disturbances of musical performance: Preliminary communication. *Journal of the Royal Society of Medicine, 76*, 194–196.

Kaiser, P., Rossner, S., & Karlsson, J. (1981). Effects of beta-adrenergic blockade on endurance and short-time performance in respect to individual muscle fiber composition. *International Journal of Sports Medicine, 2*, 37–42.

Koelega, H. S. (1993). Stimulant drugs and vigilance performance: A review. *Psychopharmacology, 111*, 1–16.

Kunsman, G. W., Manno, J. E., Przekop, M. A., Manno, B. R., & Kunsman, C. M. (1992). The effects of temazepam and ethanol on human psychomotor performance. *European Journal of Clinical Pharmacology, 43*, 603–611.

Lemon, R. N. (1999). Neural control of dexterity: What has been achieved? *Experimental Brain Research, 128*, 6–12.

Lidén, S., & Gottfries, C.-G. (1974). Beta-blocking agents in the treatment of catecholamine-induced symptoms in musicians. *Lancet, 11*, 529.

Liguori, A., & Robinson, J. H. (2001). Caffeine antagonism of alcohol-induced driving impairment. *Drug and Alcohol Dependence, 63*, 123–129.

Liguori, A., Gatto, C. P., & Jarrett, D. B. (2002). Separate and combined effects of marijuana and alcohol on mood, equilibrium, and simulated driving. *Psychopharmacology, 163*, 399–405.

Linnoila, M., Stapleton, J. M., George, D. T., Lane, E., & Eckardt, M. J. (1993). Effects of fluvoxamine, alone and in combination with ethanol, on psychomotor and cognitive performance and on autonomic nervous system reactivity in healthy volunteers. *Journal of Clinical Psychopharmacology, 13*, 175–180.

Lister, R. G., & File, S. E. (1983). Performance impairment and increased anxiety resulting from the combination of alcohol and lorazepam. *Journal of Clinical Psychopharmacology, 3*, 66–71.

Lubin, R. A. (1977). Influences of alcohol upon performance and performance awareness. *Perceptual and Motor Skills, 45*, 303–310.

Ludwig, A. M. (1990). Alcohol input and creative output. *British Journal of Addiction, 85*, 953–963.

Mandel, H. G. (2002). Update on caffeine consumption, disposition, and action. *Food and Chemical Toxicology, 40*, 1231–1234.

Mason, M. F. (1977). Drug impairment reviews: Opiates, minor tranquilizers. *NIDA Research Monograph, Series 11*, 44–60.

Mattila, M. J., Aranko, K., & Seppala, T. (1982). Acute effects of buspirone and alcohol on psychomotor skills. *Journal of Clinical Psychiatry, 43*, 56–61.

Mintzer, M. Z., & Stitzer, M. L. (2002). Cognitive impairment in methadone maintenance patients. *Drug and Alcohol Dependence, 67*, 41–51.

Nagoshi, C. T., Wilson, J. R., & Rodriguez, L. A. (1991). Impulsivity, sensation seeking, and behavioral and emotional responses to alcohol. *Alcoholism: Clinical and Experimental Research, 15*, 661–667.

Nawrot, P., Jordan, S., Eastwood, J., Rotstein, J., Hugenholtz, A., & Feeley, M. (2003). Effects of caffeine on human health. *Food Additives and Contaminants, 20*, 1–30.

Neftel, K. A., Adler, R. H., Kappeli, L., Rossi, M., Dolder, M., Kaser, H. E., Bruggesser, H. H., & Vorkauf, H. (1982). Stage fright in musicians: A model illustrating the effect of beta blockers. *Psychosomatic Medicine, 44*, 461–469.

Nestler, E. J. (2000). Genes and addiction. *Nature Genetics, 26*, 277–281.

Pearson, R. M., & Simpson, A. F. (1978). Effect of oxprenolol in stage fright in musicians. *Transactions of the Medical Society of London, 95*, 46–53.

Perez-Reyes, M., Hicks, R. E., Bumberry, J., Jeffcoat, A. R., & Cook, C. E. (1988). Interaction between marijuana and ethanol: Effects on psychomotor performance. *Alcoholism: Clinical and Experimental Research, 12*, 268–276.

Perez-Reyes, M., White, W. R., McDonald, S. A., & Hicks, R. E. (1992). Interaction between ethanol and dextroamphetamine: Effects on psychomotor performance. *Alcoholism: Clinical and Experimental Research, 16*, 75–81.

Pickworth, W. B., Rohrer, M. S., & Fant, R. V. (1997). Effects of abused drugs on psychomotor performance. *Experimental and Clinical Psychopharmacology, 5*, 235–241.

Ridderinkhof, K. R., de Vlugt, Y., Bramlage, A., Spaan, M., Elton, M., Snel, J., & Band, G. P. (2002). Alcohol consumption impairs detection of performance errors in mediofrontal cortex. *Science, 298*, 2209–2211.

RCP (2000). *Nicotine Addiction in Britain.* London: Royal College of Physicians.

Schaeffer, K. W., Parsons, O. A., & Errico, A. L. (1989). Performance deficits on tests of problem solving in alcoholics: Cognitive or motivational impairment? *Journal of Substance Abuse, 1*, 381–392.

Schiorring, E. (1981). Psychopathology induced by "speed drugs". *Pharmacology, Biochemistry, and Behavior, 14* (Supplement 1), 109–122.

Schlaug, G. (2001). The brain of musicians: A model for functional and structural adaptation. *Annals of the New York Academy of Sciences, 930*, 281–299.

Schwartz, R. H., Gruenewald, P. J., Klitzner, M., & Fedio, P. (1989). Short-term memory impairment in cannabis-dependent adolescents. *American Journal of Diseases of Children, 143*, 1214–1219.

Simon, N. G., & Mattick, R. P. (2002). The impact of regular ecstasy use on memory function. *Addiction, 97,* 1523–1529.

Simpson, C. A., & Rush, C. R. (2002). Acute performance-impairing and subject-rated effects of triazolam and temazepam, alone and in combination with ethanol, in humans. *Journal of Psychopharmacology, 16,* 23–34.

Smith, A. (2002). Effects of caffeine on human behavior. *Food and Chemical Toxicology, 40,* 1243–1255.

Specka, M., Finkbeiner, T., Lodemann, E., Leifert, K., Kluwig, J., & Gastpar, M. (2000). Cognitive-motor performance of methadone-maintained patients. *European Addiction Research, 6,* 8–19.

Spriet, L. L. (1995). Caffeine and performance. *International Journal of Sport Nutrition, 5* (Supplement), S84–S99.

Strassman, R. J. (1984). Adverse reactions to psychedelic drugs: A review of the literature. *Journal of Nervous and Mental Disease, 172,* 577–595.

Tarnopolsky, M. A. (1994). Caffeine and endurance performance. *Sports Medicine, 18,* 109–125.

Tiller, J. W. (1990). Antidepressants, alcohol, and psychomotor performance. *Acta Psychiatrica Scandinavica, Supplement 360,* 13–17.

Wesner, R., Noyes, R., & Davis, T. L. (1990). The occurrence of performance anxiety among musicians. *Journal of Affective Disorders, 18,* 177–185.

West, R. (1993). Beneficial effects of nicotine: Fact or fiction? *Addiction, 88,* 589–590.

West, R., & Hajek, P. (1997). What happens to anxiety levels on giving up smoking? *American Journal of Psychiatry, 154,* 1589–1592.

West, R., & Shiffman, S. (2003). *Smoking Cessation.* Oxford: Healthcare Press.

# EPILOGUE
## A note on future directions for enhancing musical performance

Research aimed at enhancing musical performance demands ingenuity and perseverance, as overcoming the methodological obstacles inherent to this area is no trivial undertaking. The studies reported throughout this book provide models for how such obstacles can be surmounted in future work. There are, nonetheless, a number of fundamental issues that need to be addressed in subsequent efforts to enhance performance. In my own reading of the existing research, four areas seem particularly important for further development:

- the refinement of systems used to measure performance enhancement
- the exploration of individual differences and responses to performance enhancement strategies and techniques
- the investigation of the long-term impact of interventions
- the study of how distinct approaches to enhancing performance can be combined.

First, those engaged in this research must not only be clear about which aspects of performance are responsive to enhancement through a given intervention, but also about how that enhancement can be measured reliably and consistently. As shown throughout this book, there is a multitude of ways to measure enhancement, but to date there is insufficient knowledge, in many cases, of the extent to which these methods are comparable and how they interrelate. Grassroots work aimed at establishing "industry standard" scales, systems, and levels of accepted impact could offer significant research and pedagogical benefits and could also save valuable time and resources for future generations of researchers. (This, of course, would not rule out the possibility of devising entirely novel systems of measurement, which seems necessary for investigations of unique performance characteristics, but it would enable benchmarks to be established against which new measurement systems could be compared and contrasted.) A natural next step would be to accumulate large stores of performance-related data—psychological, physical, and musical—across a range of reasonably controlled and monitored conditions. These databases could then be shared and added to internationally, in order to build up a more informative picture of the demands of performance and the ways in which musicians across the world respond to those demands.

Second, it is a common pattern within psychological research to work toward an understanding of principles and mechanisms that govern the functioning of all humans. From these endeavors, one can learn a great deal about the full extent of human behavior and achievement. In performance enhancement research, however,

there seems to be tremendous potential for identifying how individual musicians differ with respect to their skills and how they respond idiosyncratically to different performance contexts and interventions. Understanding how individual variability arises within the confines of generalizable psychological mechanisms would certainly benefit any research initiative aimed at enhancing performance. In fact, one goal toward which research in this area should be geared is the identification of intervention-specific predictor variables, enabling the immediate assessment of whether a performer could indeed benefit from an intervention. Aside from attributes of the person (e.g. personality traits and musical abilities), these predictor variables should also be oriented around the musician's performance genre and his or her instrumental specialization.

Third, efforts should be made to determine the extent to which performance enhancement benefits will last over the long-term and how musicians should engage with an intervention over the course of their careers. This information could have an immediate impact on decisions to apply an intervention in educational and professional environments. For instance, a particularly effective technique may be of little practical value to an educational institution if the yield on investment (in terms of time, personnel, and expense) is comparatively short-lived and unfavorable. Thus, it is important where possible to pursue research agendas that are longitudinal in nature so that long-term benefits (or side-effects) can be documented.

Finally, much research (particularly in relation to the techniques discussed in Part III of this book) has been aimed at determining the effects of only one type of intervention on performance. This is understandable, as it is impossible to know the precise source of a given benefit if the musician is exposed to an elaborate cocktail of strategies. Still, for those approaches for which singular benefits have been identified, much could be gained by pairing them with complementary physical and psychological interventions. A performer cannot achieve excellence only by isolating and refining a limited range of skills; likewise, it seems reasonable to argue that the full potential for performance enhancement research lies in exploring how musicians' physical and psychological abilities can be enhanced in unison.

Other such methodological issues have been raised throughout this book (see, for example, Chapters 1, 11, and 13), and still others will invariably arise as the discipline expands and develops. It is vital, in my view, that systematic attempts be made to address such issues rigorously in future research. In this way, a greater theoretical and practical understanding of musical performance—as well as the educational and professional contexts in which it occurs—can be achieved, and research to enhance performance can begin to offer even richer benefits for practitioners.

Aaron Williamon

# INDEX


Aarhus Symphony Orchestra 185
acamprosate 275
achievement, lens model 258
acoustics 69
acting 'as if' 230–1
activity selection strategies 87–8
acute exercise 166, 169–70, 172
adaptors 112
adrenergic drugs 273
aesthetic explorative aspect 259, 262
affect 164, 167
alcohol 166, 204–5, 273–5, 279, 281–5
Alexander technique 6, 15, 48, 167–8, 171, 179–91, 209–10
  anatomical studies 182–4
  application to music performance 180–1
  behavioral studies 185–91
  good use 179
  physiological studies 182–5
  and science 181–2
alpha neurofeedback 201, 204–5, 213
alpha/theta neurofeedback 14–15, 190, 204–13*passim*
American College of Sports Medicine 174
American Psychiatric Association 276
amphetamine sulfate 280
amphetamines 280–2, 284–5
antidepressants 273, 277–8, 280, 285
anxiety 272–3
  state 12, 190, 208
  trait 11–12, 181
  *see also* performance anxiety
anxiolytics 273, 279, 280
arousal 222–3
artistic image 22, 28–9
assessment process 61–2
assessor *see* evaluator characteristics
Associated Board of the Royal Schools of Music 145, 207
atenolol 278
attention 203–4
  deficit hyperactivity disorder 203
  identification of drifts in 235
  refocusing 226
  *see also* concentration
attractiveness 67–8
attribution 31–2
audience
  connection and presence with 226
  support 70
auditions 74
auditory style 251
aural
  memory 130–1, 135, 137–8
  modeling 96, 230
autonomic nervous system 8–9, 198–9
Aventyl 277

behavioral
  strategies 85
  studies 185–91
  symptoms of anxiety 11
Benzedrine 280
benzodiazepines 274, 279, 283–4
beta-1 band 201–3, 206, 208–10
beta-1 selective drugs 278–9
beta activity 211–13
beta-blockers 166, 185, 278–9, 285
bias 71, 73–74
biofeedback *see under* physiological self-regulation
blind area 75–6
'blind' auditions 74
blood pressure 213
Botulinum toxin (Botox) 47
brass instruments 149
  *see also* French horn; trombone; trumpet; tuba
breaks 56
breathing 223–4, 235
British Association for Performing Arts Medicine 42–4, 47–9, 52, 54–6
buprenorphine 284
bupropion 277
buspirone 279

caffeine 275–6, 280, 285
Campral 275
cancer 164, 277
cannabis 281, 284–5
catastrophe model 12–13
cello 96, 114–19, 188, 221, 225
  physical demands 43–4, 47–8, 53–4
central nervous system 8–9, 198

Central School of Speech 185–6
characteristics 24–9
checking 98
chlordiazepoxide 279
chronic exercise 166–72
chronic obstructive pulmonary disease 277
clarinet 45, 188
classification 98
cocaine 210, 281–5
cognitive analysis 129
cognitive feedback 247, 256, 259–62, 265
  basic procedure 260–1
  computer software 256, 259, 262–5
  generality 264
  research 262–4
  *see also* feedback learning of musical expressivity
cognitive processing of visual information 147–8
cognitive–behavioral training *see* mental skills training
communication 64, 111–18, 207–8
  coperformers 113–14
  rehearsal 115–18
  within ensemble 68–9
concentration 24–6, 233–7
  *see also* attention
conceptual memory 129, 133–5
conducting 43, 55, 130, 225
  and ensemble communication 111–13
confidence 236–7
consistency, lens model 258
context–dependent learning 137–8
control of strategies 98
cooling down 55
correction strategies 96–7
'crack' 210, 281
criterion-based procedures 61–2
cross-stressor hypothesis 166
cue weight, lens model 258

d prime 206
De Quervain syndrome 44
deliberate practice 23, 154, 256
delta waves 201
depression 211
deviations 251
dexamfetamine 280
Dexedrine 280
diabetes 164
diacetylmorphine 283
diamorphine 283
diazepam 279
diet 56
difficult/challenging parts strategies 93–4
direction dimension, concentration 233–4
distraction 168, 170
distribution of practice over time strategies 94–5
dopaminergic drugs 273
double bass 90
drugs and musical performance 7, 210, 271–85
  illicit drugs 280–4
  lifestyle drugs 273–7
  performance as psychomotor task 271–3
  prescription drugs 277–80
drums 45
dystonia, focal 44–7

ecstasy 282, 285
efficacy, cognitive feedback 262
electrodermal biofeedback 213
electroencephalogram (EEG) 197, 200–6, 210–13, 225
electromyography (EMG) 199, 200, 213, 225
emblems 112
emotion/emotional 72, 164, 248–50
  commitment and conviction 197
  expression 253–4
  felt 230
  focused strategies 11–12
  intensity 72
  negative 226
  variety 72
endurance fitness 167
enhancement of performance 3–16, 205–12
  measurement 61–78
    assessment process 61–2
    error 65
    extramusical factors 65–73
    Johari window 74–7
    musical value 62–5
    nonmusical factors 73–4
  obstacles, management and overcoming of 9–13
  peak performance achievement 206–12
  personal program selection 13–15
  physical and psychological enhancement 7–9
  practice strategies 5–6
  prospects and limits 4–5
  techniques and interventions 6–7
ensemble practice strategies 5, 105–20
  case studies 114–19
  group dynamics 106–7
  rehearsal situation 108–11
  verbal and nonverbal communication 111–14
  *see also* practice
episodic memory 126, 128
errors of commission 206
  detection 96–7
  of omission 206
eurphoriant actions 273

evaluation 98
  strategies 86, 95–7
evaluator characteristics 70–3
execution 152
  failure 155
executive strategies 86, 91–5
exercise 154, 185, 210
  *see also* physical fitness
explicit understanding 249–50
expression 64, 67, 251
external locus 31
extramusical factors 65–73
  context related aspects 68–70
  evaluator characteristics 70–3
  performer-related aspects 66–8
extrinsic motivation 29–30

familiarity 72–3
Féderation Internationale des Musiciens 42, 46, 56
feedback learning of musical expressivity 7, 247–66
  cognitive feedback 259–62
  cognitive feedback, generality of 264
  cognitive feedback research 262–4
  common myths 248–51
  criteria for useful teaching strategy 256–7
  implications for music education 265
  lens model 257–9
  perspectives from empirical research 251–6
  *see also* cognitive feedback
Feldenkrais technique 47
felt emotion 230
first impressions 71–2
fitness 49–50
  endurance 167
  *see also* exercise; physical fitness
flair 67–8
flow 31
fluoxetine 278
flute 47, 54–5, 74, 90, 106, 188, 226, 264
fluvoxamine 278
focal dystonia 44–7
focused characteristic 238
forced expiratory flow 184, 187
forced vital capacity 183, 187
forethought 86
French horn 106, 226
functional palsy (craft palsy) 200

generative rules 253
GERMS model 253–4
goals
  choosing 240
  orientation 168
  prioritizing 240–1
  setting 26–7, 89
good use and Alexander technique 179
group dynamics 106–7
guitar 43, 45, 47, 49, 54, 151, 255, 264

hallucinogens 273, 282–3, 285
halo effect 73, 76
heart disease 277
height 188–9
'helpless' orientation 22
heroin 283–5
hidden area 75–6
humming 154
hypermobility 43–4, 50
hypnotics 279

'ice' 280
ideal performance states 237–9
illicit drugs 280–4
illustrators 112
imagery skills development 227–8
images 230–1
imipramine 277
implicit knowledge 153
improvising *see* sightreading and improvising
'in group' bias 73
Inderal 279
individual pieces, approaches to 110–11
individual practice strategies 5, 85–99
  evaluation strategies 95–7
  executive strategies 91–5
  historical perspective 87
  metastrategies 97–8
  planning and preparation strategies 87–91
  *see also* practice
injury prevention 55–6
instrumental
  competence (technical ability) 207–8
  learning 198
  technique 64
instrument-related factors 252
intensity 98
  of emotion 72
intention 155
intentionality 31–2
internal
  locus 31
  stimuli 152
  timekeeper model 148–9
interpretation 64
interpretive imagination 197, 209–10
interventions 6–7

intrinsic motivation 29–31
inverted U function 12, 72–3

Johari window 74–7

keyboard 48, 50, 90, 129, 144–5, 149, 180
kinesthetic memory 128–30, 135, 137–8
knowledge 98
  about personal variables 97
  about strategies 97–8
  base acquisition 154
  general 97
  implicit 153
  metacognitive 14
  previous in secondary function 150–1
  relational 97
  specific 97
  task 97

learning
  context-dependent 137–8
  improvement 225
  instrumental 198
  long-term 240–1
  transfer of 93
Learning Styles Index 138
lens model 257–9
Librium 279
lifestyle drugs 273–7
limits 4–5
listener-related factors 252
*loci* 127
long-term learning 240–1
long-term memory 124–8, 132, 147
Lopresor 278
lung cancer 277
Luvox 278
lysergic acid diethylamine (LSD) 282

massed approach 94
mastery 22, 153, 170
matching, lens model 258
maximal expiratory pressure 184, 187
maximal inspiratory pressure 184, 187
maximum voluntary ventilation 184
MDMA *see* ecstasy
mean target tempo 25
medical diagnoses, specific 42–7
memorizing music strategies 5, 123–38
  aural memory 130–1
  conceptual memory 132–5
  recommendations 126–7
  rote 129
  structure 124–6
  visual information 130–1
  words and music combining 135–7
memory
  aural 130–1, 135, 137–8
  conceptual 129, 133–5
  episodic 126, 128
  improvement 225
  influences 71
  kinesthetic 128–30, 135, 137–8
  long-term 124–8, 132, 147
  semantic 125, 128
  short-term 124–5, 127
  visual 128, 130–1, 135, 137–8
  working 124
  *see also* memorizing music strategies
mental rehearsal 224–31
  acting 'as if' 230–1
  developing control of 228
  evocative symbols and images 230–1
  imagery skills development 227–8
  qualities in performance 230
  and skill development 229
mental skills training 7, 221–42
  long–term learning and performance preparation 240–1
  mental rehearsal 224–31
  relaxation 222–4
  specific performance situations 231–9
mental versus playing rehearsal strategies 91–2
metacognitive knowledge 14
metaphors 250, 257, 262, 265
metastrategies 86, 97–8
methadone 283–4
methamphetamine 280
  hydrochloride 280
Methedrine 280
metoprolol 278
mind-mapping 126–7
mistakes 155–6
  misexecution 155–6
  misreading 155–6
misuse and Alexander technique 179
mnemonics 127
modeling 257
moment-by-moment information 105–6
monoamine oxidase inhibitors 278
mood 164, 167, 188
  of the assessor 72
morphine 283
motion principles 253–4
motivation 29–31
motor programming 148, 151

movement 68
  generation 178–9
  images 151
musical expressivity *see* feedback learning of musical expressivity
Musical Performance Anxiety Self-Statement Scale 188
musical value 62–5
musicality 68, 197, 207–11

negative
  emotions 226
  impresssions 72
nervous zoom 236
neurofeedback *see under* physiological self-regulation
nicotine 276–7, 285
  replacement therapy 277
nonmusical factors 73–4
nonplaying practice 87–8
nonstructural musculoskeletal problems 47–56
  cello 53–4
  conductors 55
  flute 54–5
  guitar 54
  injury prevention 55–6
  piano 50–2
  posture 48
  practice, tension and fitness 49–50
  violin and viola 52–3
nonverbal communication 111–14
normative aspect 259
norm-referenced procedures 61–2
nortriptyline 277
notation perception 146–7

oboe 172, 188, 226
obstacles, management and overcoming of 9–13
old injuries 43, 50
opiates 283–4
opium 283
optimal (reference) model 256
order bias 74
order of performance 74
organ 188
organization strategies 87–8
'out group' bias 73
overall quality 207–10

Pamelor 277
parasympathetic nervous system 8–9
paroxetine 278
passive smoking 277
pattern chaining models 150
pattern recognition 147
peak
  experience 226
  expiratory flow 184, 187–9
  performance achievement 206–12
perceptual sensitivity 206
percussion 44
  *see also* drums
performance
  anxiety 9–10, 165, 182, 222, 278
  features 254
  order 74
  plan 89
  preparation 95, 232, 240–1
  qualities 230
  rules 155
  volitional control 86
performer-related aspects 66–8, 252
peripheral nervous system 8, 198
permanent knowledge 105
personal program selection 13–15
personal variables 97
physical demands 4, 41–57
  medical diagnoses, specific 42–7
  nonstructural musculoskeletal problems 47–56
  surveys 42
physical enhancement 7–9
physical fitness 6, 15, 163–75
  acute exercise 169–70, 172
  background 153–6
  chronic exercise 166–72
  guidelines 173–5
  subjective responses 172
physical technique 63–4
physiological
  adaptations 169–72
  self–regulation 6–7, 197–214
    biofeedback 198–9
    and music performance 199–200
    neurofeedback 200–2
    and attention 203–4; and performance enhancement 205–12; and relaxation 204–5
  symptoms of anxiety 10–11
  technique 63
  variables 146
piano 21–2, 24–5, 27, 87, 172, 242, 255
  Alexander technique 184, 188
  ensemble practice 107, 113–19
  memorizing music 123, 130–5
  performance enhancement 67, 69
  physical demands 43, 45, 47, 50–2
  physiological self-regulation 198, 200
  sight-reading and improvising 145, 149, 153, 156

piece-related factors 252
planning 152
  failure 155
  strategies 86–91
playing practice 87–8
poise 179
positive impresssions 72
post-traumatic stress disorder 204–5, 211
posture 48, 56, 179
practice 49–50, 154–5
  deliberate 23, 154, 256
  efficiency 225
  quality 22–4, 91
  quantity 20–2, 89, 91
  rate 25
  strategies 5–6
*see also* individual practice strategies, ensemble practice strategies
pre-performance
  guidelines 174–5
  routine 231–3
preference 72–3
preparation
  consistency 232–3
  strategies 86–91
prescription drugs 277–80
present moment technique 237
previous knowledge in secondary function 150–1
primary control 179
problem
  focused strategies 12
  identification 155–7
process evaluation strategies 96–7
proof-reader's errors 148
propranolol 185, 279
prospects 4–5
Prozac 278
psychedelic drugs 282–3
psychological
  enhancement 7–9
  impact 169–71
  processes 144–5
psychomotor
  skills 143
  task, musical performance as 271–3
psychotropic drugs 271
public area 75–6
purpose of assessment 70

quality
  in performance 230
  of practice 22–4, 91
quantity of practice 20–2, 91
Queen Elisabeth music competition (Belgium) 72, 74

'rage to master' 31
random fluctuations 253
regulation of strategies 98
regulators 112
rehearsal
  communication 115–18
  method 114–15
  situation 108–11
  strategies 91–4
relational knowledge 97
relaxation 56, 204–5, 222–4
reliability 65
  of assessment 75
repertory of strategies 87
rote 129
Royal College of Music 185–6
runs 27
Russia 274

savants 21, 131
saxophone 149, 255, 264
scheduling 241
secondary function, previous knowledge in 150–1
secret area 75–6
sedatives 273
selective serotonin re–uptake inhibitors 278
self-confidence 168
self-efficacy 32, 66–7
self-esteem 168
self-evaluation 27
self-guiding strategies 97
self-reflection 86
self-regulated learning 86–7
self-teaching 85–6
semantic memory 125, 128
sensorimotor rhythm (SMR) 202–3, 206, 208–10, 213
sensory
  awareness 225
  store 124–5
Seroxat 278
short-term memory 124–5, 127
shuttling 235–6
sightreading and improvising 5–6, 143–57
  comparisons 144–52
    common aspects 151–2
    task analysis: improvisation 149–51
    task analysis: sight-reading 145–9
  skills acquisition and enhancement 152–7
simulation 237–9

singing *see* voice
situational stress 11
skills
  acquisition and enhancement 152–7
  development 225, 229
slips 155
smoking *see* nicotine
social
  factors 70
  interaction 118–19
  physique anxiety 167
  psychological antecedents 29–32
somatic nervous system 8, 9, 198
specific knowledge 97
specific performance situations 231–9
  concentration 233–7
  ideal performance states and simulation 237–9
  pre-performance routine 231–3
speed 98
Spielberger's State Anxiety Inventory 208
Spielberger's Trait Anxiety Inventory 181
stability 31–2
stage fright *see* performance anxiety
state anxiety 12, 190, 208
state-dependence 137
stereotyping 74
stimulants 273, 276
stress 11, 165
stringed instruments 48–50, 90, 149, 180–1, 213
  ensemble practice 106, 108–9, 113
  *see also* cello; double bass; guitar; viola; violin
structural disorders 43–4
structure 124–6, 253–4
style of music 154
stylistic accuracy 197, 209–10
stylistic unexpectedness 253–4
subjective responses 172
Swassing–Barbe Modality Index 138
Sydney Conservatorium of Music 71
symbols 230–1
sympathetic nervous system 8, 9
systematic variations 251

tai chi 167–8
task
  analysis: sight-reading 145–9
  difficulty 11
  knowledge 97
  selection 98
teacher feedback 262
teaching strategy 256–7
technical
  difficulties 225
  directions 262
  skills 251
technique 6–7, 63–4
tempo of performance strategies 94
ten-year rule 20–2
Tenormin 278
tenosynovitis 43–4
tension 49–50
thermal biofeedback 213
theta waves 201, 211–12
  *see also* alpha/theta neurofeedback
'thought' strategies 85
time management strategies 89–91
tobacco *see* nicotine
Tofranil 277
top performance mental rehearsal 238–9
training cost 262
trait anxiety 11–12, 181
tranquilizers 166, 279–80, 285
transfer of learning 93
tricyclics 277–8
trombone 188
trumpet 74, 90
tuba 90
type of performance 70

usability, cognitive feedback 262
user
  satisfaction, cognitive feedback 262
  understanding, cognitive feedback 262

validity 73
  of assessment 75
Valium 279
variety of emotion 72
verbal communication 111–14
vicarious functioning 259
viola 47, 52–3, 200
violin 26, 43, 47, 49, 52–3, 90, 200
  Alexander technique 184–5, 188
  ensemble practice 107, 113
  feedback learning 260, 264
visual information 130
  cognitive processing 147–8
  translation into imaginary sound 131
visual memory 128, 130–1, 135, 137–8
visual models 96
voice 48, 66–7, 111, 154, 226, 264
  Alexander technique 180–1, 187–8
  memorizing music 129–37

warming-up 55, 88
weight management 165
whistling 154
whole piece versus small part stategies 92–3
width dimension, concentration 233–4
wind instruments 45, 213, 277
  Alexander technique 180, 187
  ensemble practice 107
  individual practice 90–1
  memorizing music 133
  physiological self-regulation 200
  sight-reading and improvising 149
  *see also* clarinet; flute; oboe; saxophone
words and music combining 135–7
working memory 124

Yerkes–Dodson law 12
yoga 47, 167–8

zooming in and out 236
Zyban 277

*Musical Excellence* offers performers, teachers, and researchers new perspectives and practical guidance for enhancing performance and managing the stress that typically accompanies performance situations. It draws together, for the first time in a single collection, the findings of pioneering initiatives from across the arts and sciences. Specific recommendations are provided alongside comprehensive reviews of existing theory and research, enabling the practitioner to place the strategies and techniques within the broader context of human performance and encouraging novel ways of conceptualizing music making and teaching.

**Part I**, Prospects and Limits, sets out ground rules for achieving musical excellence. What roles do innate talent, environmental influences, and sheer hard work play in attaining eminence? How can musicians best manage the physical demands of a profession that is intrinsically arduous, throughout a career that can literally span a lifetime? How can performers, teachers, and researchers effectively assess and reflect on performance enhancement for themselves, their colleagues, and their students?

**Part II**, Practice Strategies, presents approaches for increasing the effectiveness and efficiency of practice. These are examined generally for the individual and ensembles and specifically for the tasks of memorizing, sight-reading, and improvising music. Musicians spend vast amounts of time and energy acquiring and refining their skills, but are there particular rehearsal strategies that they can employ to produce better performance results or to achieve the same results more quickly? What implication does existing knowledge of human information processing and physical functioning have for musical learning and practice?



**Part III**, Techniques and Interventions, introduces scientifically validated methods for enhancing musical achievement, ordered from the more physical to the psychological to the pharmacological; however, they all address issues of both mental and physical significance for the musician. Collectively, they stand as clear evidence that applied, cross-disciplinary research can facilitate musicians' striving towards performance excellence.

Throughout, the book highlights ways for musicians to make the most of their existing practice, training, and experience and gives them additional tools for acquiring and developing new skills. Each chapter is underpinned by physical and psychological principles relevant to all performance traditions that demand dedication and resilience, unique artistic vision, and effective communication.

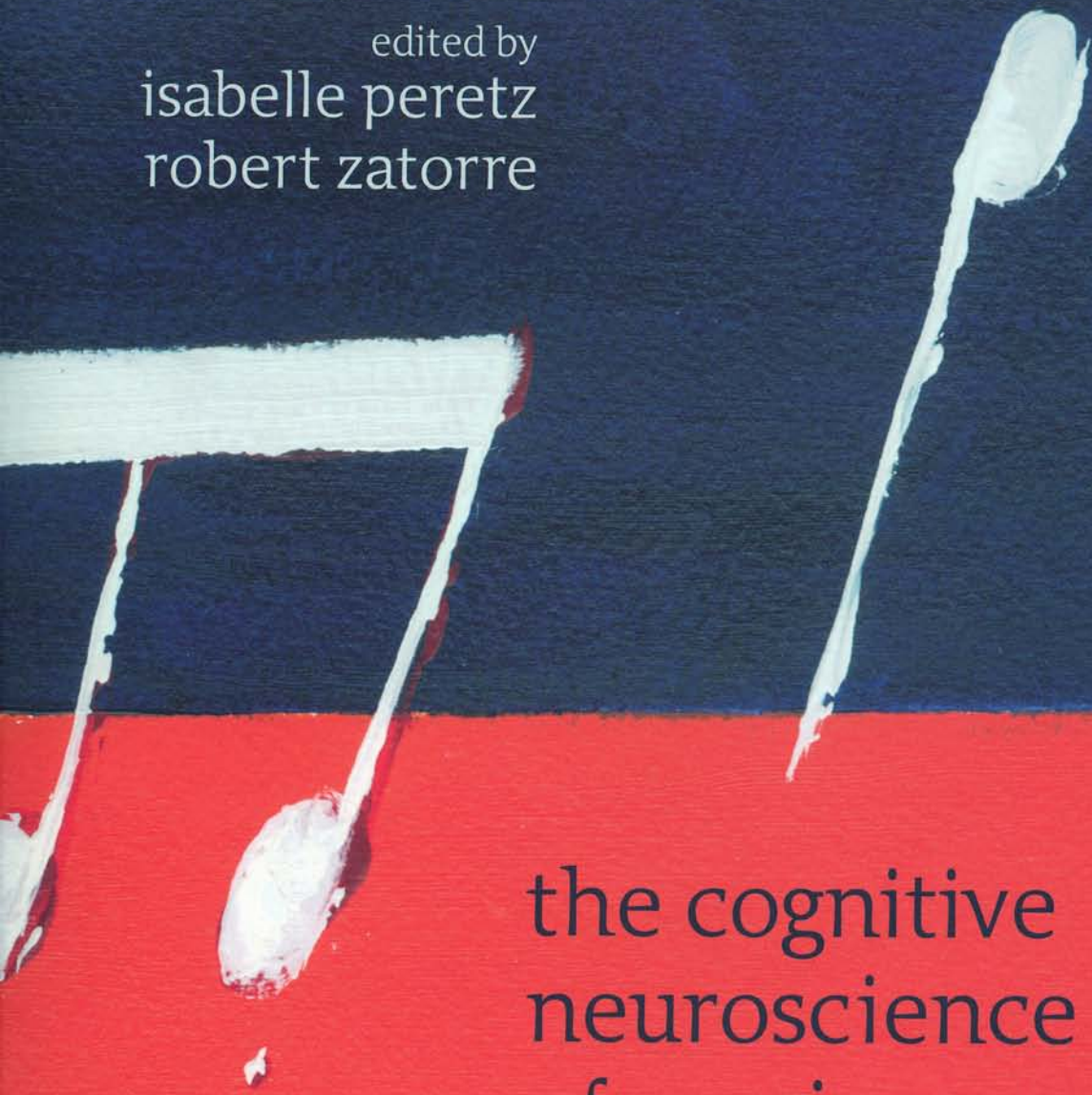
*Cover image:* © Corbis

EL ARGONAUTA
la librería de la música
0-1985-2535-4
WILLIAMON, AARON:
MUSICAL EXCELLENCE
P.V.P. (euros): 40.50

OXFORD
UNIVERSITY PRESS
www.oup.com

ISBN 0-19-852535-4
9 780198 525356

edited by
isabelle peretz
robert zatorre



# the cognitive neuroscience of music

OXFORD

# THE COGNITIVE NEUROSCIENCE OF MUSIC

Edited by

ISABELLE PERETZ

*Département de Psychologie,*
*Université de Montréal, C.P. 6128,*
*Succ. Centre-Ville, Montréal,*
*Québec, H3C 3J7, Canada*

and

ROBERT J. ZATORRE

*Montreal Neurological Institute,*
*McGill University, Montreal,*
*Quebec, H3A 2B4, Canada*

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford ox2 6dp

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Taipei Toronto Shanghai

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan South Korea Poland Portugal
Singapore Switzerland Thailand Turkey Ukraine Vietnam

Published in the United States
by Oxford University Press Inc., New York

© The New York Academy of Sciences, Chapters 1–7, 9–20, and 22–8, and Oxford University Press, Chapters 8 and 21. Most of the materials in this book originally appeared in *The Biological Foundations of Music,* published as Volume 930 of the Annals of the New York Academy of Sciences, June 2001 (ISBN 1-57331-306-8). This book is an expanded version of the original Annals volume.

The moral rights of the author have been asserted

Database right Oxford University Press (maker)

First published 2003
Reprinted 2004 (twice)

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above.

You must not circulate this book in any other binding or cover
and you must impose this same condition on any acquirer.

A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
(Data available)

ISBN 0 19 852519 2 (Hbk)

0 19 852520 6 (Pbk)

10 9 8 7 6 5 4 3

Typeset by Newgen Imaging Systems (P) Ltd., Chennai, India
Printed in Great Britain
on acid-free paper by Biddles Ltd., King's Lynn, Norfolk

# PREFACE

Over the past decade there has been an explosion in research activities on music perception and performance, and their correlates in the human brain. This sudden increase in scientific work on music has been motivated in part by the idea that music offers a unique opportunity to better understand the organization of the human brain. The other major motivation for exploring the neural substrates of musical activities is that they may shed light on the functional origin and biological value of music. Like language, music exists in all human societies. Like language, music is a complex, rule-governed activity, and appears to be associated with a specific brain architecture. Moreover, sensitivity to musical structure develops early in life, without conscious effort, in the large majority of the population. Music also appears to be specific to humans, although some investigators have begun to examine its possible evolutionary origins in other species'.[1]

Unlike most other high-level functions of the human brain—and unlike language—only a minority of individuals become proficient performing musicians through explicit tutoring. This particularity in the distribution of acquired skills confers to music a privileged role in the study of brain plasticity. Another distinction from language is that a large variety of sensory-motor systems may be studied because of the many different ways of producing music. These variable modes of auditory expression enable interesting comparisons across systems.

Given both its similarities with language and its divergence from it, we believe that the relationship between music and the brain is of central importance for the domain of cognitive neuroscience. The fact that musical activities have generally been considered as an exquisite product of human culture, and as such are often assumed to be merely a cultural artifact, should not be taken as an impediment to achieving a scientific understanding of its underlying basis. In fact, we would argue that the cultural overlay associated with music confers upon it a key role for understanding the biology of human cognitive functions. Indeed, the ubiquity of music, its developmental features, and its brain substrates raise the question of the nature and the extent of its biological foundations, and how these interact with culture.

In this context, it is remarkable to note that the study of music as a major brain function has been relatively neglected. Neuropsychological questions related to musical abilities have been of occasional interest to neurologists and psychologists since the last century (e.g. Ref. 2 for pioneering work), but systematic, sustained investigations have been rare until recently. Several developments—both theoretical and technological—have created a profound change in the way in which music studies are perceived, thereby enabling us to present this volume as a beginning to the scientific study of the neurobiology of music. First, one should mention the development of cognitive psychology during the latter half of the twentieth century; cognitive psychologists were among the first to recognize the

value of music as a means of studying perception, memory, attention, and performance (see Refs 2–5 for the three benchmark textbooks). Second, neuropsychology, which had always been based on knowledge from neurophysiology and neuroanatomy, began also to adopt experimental and cognitive paradigms, permitting advances in understanding lesion effects on musical functions. Third, developmental psychologists exploited new techniques that allowed them to probe the mind of even neonates. Finally, and most dramatically perhaps with the arrival on the scene of neuroimaging techniques such as fMRI and MEG, the stage was set for the explosion to which we alluded above, and which forms the core of the present volume.

The objective of the conference, from which most of the present texts are derived, was to demonstrate the dynamism and richness of the discipline. The conference was held in May 2000 in New York, and was initiated and sponsored by the *New York Academy of Science.* Major scientists from six different countries, working in a variety of interrelated disciplines and pursuing sustained research activities on music were invited to present their work. Most responded enthusiastically and addressed issues crucial to a better understanding of the neural substrates underlying musical functions. Their presentations first appeared in the *Annals of the New York Academy of Sciences,* volume 930, under the title 'The Biological Foundations of Music' in 2001. Given the success of this issue, Oxford University Press invited us to edit an updated version of it in their collection of books. We eagerly accepted and invited a few other major players in the field who originally presented a poster to publish a full chapter in the present volume. Therefore, you will find here an updated version of the 25 original chapters as well as three new chapters.

The contributions in the present volume will reveal how much progress has been made in the field over the last decade. Besides providing research overviews, many texts also delineate significant avenues for future research. Indeed, it is clear that this is a very young field still, and that much remains to be done. We hope that the material presented in this volume, coupled with an increase of interest for the biological foundations of music in the scientific and general community, represent the initial spark for building a solid and stimulating cognitive neuroscience of music.

I.P. and R.Z.

February 2003

## References

1. **Wallin, N., B. Merker and S. Brown** (ed.) (2000) *The Origins of Music.* Cambridge, MA: MIT Press.
2. **Bouillaud, J.** (1865) Sur la faculté du langage articulé. *Bulletin de l'Académie de Médecine* 30, 752–68.
3. **Deutsch, D.** (ed.) (1982) *The Psychology of Music.* New York: Academic Press.
4. **Dowling, W. and D. Harwood** (1986) *Music Cognition.* Series in cognition and perception. New York. Academic Press.
5. **Sloboda, J.** (1985) *The Musical Mind: The Cognitive Psychology of Music.* London: Oxford University Press.

# CONTENTS

List of contributors *page xi*

## Part I The origins of music

1 Musical predispositions in infancy: an update *page 3*
*Sandra E. Trehub*

2 The quest for universals in temporal processing in music *page 21*
*Carolyn Drake and Daisy Bertrand*

3 Mechanisms of musical memory in infancy *page 32*
*Jenny R. Saffran*

4 Music, cognition, culture, and evolution *page 42*
*Ian Cross*

5 Is music an evolutionary adaptation? *page 57*
*David Huron*

## Part II The musical mind

6 The roots of musical variation in perceptual similarity and invariance *page 79*
*Stephen McAdams and Daniel Matzkin*

7 Tonal cognition *page 95*
*Carol L. Krumhansl and Petri Toiviainen*

8 Learning and perceiving musical structures: further insights from artificial neural networks *page 109*
*Barbara Tillmann, Jamshed J. Bharucha, and Emmanuel Bigand*

## Part III The neurons of music

9 Neurobiology of harmony perception *page 127*
*Mark Jude Tramo, Peter A. Cariani, Bertrund Delgutte, and Louis D. Braida*

10 Intracerebral evoked potentials in pitch perception reveal a functional asymmetry of human auditory cortex *page 152*
*Catherine Liégeois-Chauvel, Kimberly Giraud, Jean-Michel Badier, Patrick Marquis, and Patrick Chauvel*

11 The neural processing of complex sounds *page 168*
*Timothy D. Griffiths*

**Part IV Musical brain substrates**

12 Music and the neurologist: a historical perspective *page 181*
*John C. M. Brust*

13 Brain specialization for music: new evidence from congenital amusia *page 192*
*Isabelle Peretz*

14 Cerebral substrates for musical temporal processes *page 204*
*Séverine Samson and Nathalie Ehrlé*

15 Cerebral substrates of musical imagery *page 217*
*Andrea R. Halpern*

16 Neural specializations for tonal processing *page 231*
*Robert J. Zatorre*

17 Exploring the functional neuroanatomy of music performance, perception, and comprehension *page 247*
*Lawrence M. Parsons*

18 Comparison between language and music *page 269*
*Mireille Besson and Daniele Schön*

19 Musical sound processing: EEG and MEG evidence *page 294*
*Mari Tervaniemi*

20 Processing emotions induced by music *page 310*
*L. J. Trainor and L. A. Schmidt*

21 A new approach to the cognitive neuroscience of melody *page 325*
*Aniruddh D. Patel*

22 How many music centres are in the brain? *page 346*
*Eckart O. Altenmüller*

**Part V Musical expertise/brain plasticity**

23 Functional organization and plasticity of auditory cortex *page 357*
*Josef P. Rauschecker*

24 The brain of musicians *page 366*
*Gottfried Schlaug*

25 Representational cortex in musicians *page 382*
*C. Pantev, A. Engelien, V. Candia, and T. Elbert*

26 The brain that makes music and is changed by it *page 396*
*Alvaro Pascual-Leone*

## Part VI Relation of music to other cognitive domains

27 The sounds of poetry viewed as music *page 413*
*Fred Lerdahl*

28 Does exposure to music have beneficial side effects? *page 430*
*E. Glenn Schellenberg*

# LIST OF CONTRIBUTORS

**Eckart O. Altenmüller** Institut für Musikphysiologie und Musiker-Medizin der Hochschule für Musik und Theater Hannover, 30175 Hannover, Germany

**Daisy Bertrand** Laboratoire de Psychologie Expérimentale, CNRS UMR 8581, Université René Descartes, Institut de Psychologie, Centre Universitaire de Boulongne, 92774 Boulogne-Billancourt Cedex, France

**Mireille Besson** Language and Music Group, Institut de Neurosciences Physiologiques et Cognitives, CNRS, 31 ch. Joseph Aiguer, 13402 Marseille, Cedex 20, France

**Jamshed J. Bharucha** Language and Music Group, Institut de Neurosciences Physiologiques et Cognitives, CNRS, USA

**Emmanuel Bigand** Université de Bourgogne, LEAD-CNRS, Bd Gabriel, F-21000 Dijon, France

**John C.M. Brust** Department of Neurology, Harlem Hospital Center and Columbia University College of Physicians and Surgeons, New York, New York 10037, USA.

**V. Candia** Department of Psychology, University of Konstanz, Konstanz, Germany

**Ian Cross** Faculty of Music, University of Cambridge, West Road, Cambridge CB3 9DP, UK

**Carolyn Drake** Laboratoire de Psychologie Expérimentale, CNRS UMR 8581, Université René Descartes, Institut de Psychologie, Centre Universitaire de Boulogne, 92774 Boulogne-Billancourt Cedex, France

**Nathalie Ehrlé** Service de Neurologie, Hôpital Maison-Blanche, 45 Rue Cognacq-jay, 51092 Reims Cedex, France

**T. Elbert** Department of Psychology, University of Konstanz, Konstanz, Germany

**A. Engelien** Institute for Experimental Audiology, University of Münster, Münster, Germany

**Timothy D. Griffiths** Departments of Neurology and Physiological Sciences, Newcastle University Medical School, Newcastle-upon-Tyne NE2 4HH, United Kingdom

**Andrea R. Halpern** Psychology Department, Bucknell University, Lewisburg, Pennsylvania 17837, USA

**David Huron** School of Music, Ohio State University, Columbus, Ohio 43210, USA

**Carol L. Krumhansl** Department of Psychology, Cornell University, Ithaca, New York 14853, USA

**Fred Lerdahl** Department of Music, 602 Dodge Hall, University of Columbia, 2960 Broadway, New York NY 10027 6902, USA

Catherine Liégeois-Chauvel INSERM EPI 9926, Laboratoire de Neurophysiologie et Neuropsychologie, Faculté de Medecine, 13005 Marseille, France

**Daniel Matzkin** IRCAM-CNRS, 1 place Igor Stravinsky, F-75004 Paris, France

Stephen McAdams IRCAM-CNRS, 1 place Igor Stravinsky, F-75004 Paris, France

Christo **Pantev** Language and Music Group, Institut de Neurosciences Physiologiques et Cognitives, CNRS, Germany

**Lawrence M. Parsons** University of Texas Health Science Center at San Antonio, San Antonio, Texas 78284-6240, USA

**Alvaro Pascual-Leone** Behavioral Neurology Unit, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02215, USA

**Aniruddh D. Patel** The Neurosciences Institute, 10640 John Jay Hopkins Drive, San Diego, CA 92121, USA

**Isabelle Peretz** Département de Psychologie, Université de Montréal, C.P. 6128, succ. Centre-Ville, Montréal (Qué), H3C 3J7, Canada

**Josef P. Rauschecker** Department of Physiology and Biophysics, Georgetown Institute for Cognitive and Computational Sciences, Georgetown University School of Medicine, Washington, DC 20057-1460, USA

Jenny **R. Saffran** Department of Psychology, University of Wisconsin-Madison, Madison, WI 53706, USA

**Séverine Samson** Laboratoire URECA, UFR de Psychologie, Universite Charles de Gaulle Lille-III, BP 149-59653 Villeneuve d'Ascq Cedex, France

E. **Glenn Schellenberg** Department of Psychology, University of Toronto at Mississauga, Mississauga, ON L5L 1C6, Canada

**Gottfried Schlaug** Beth Israel Deaconess Medical Center and Harvard Medical School, 330 Brookline Ave., Boston, MA 02215, USA

**Louis Schmidt** Department of Psychology, McMaster University, Hamilton, ON L8S 4B2, Canada

**Daniele Schön** Language and Music Group, Institut de Neurosciences Physiologiques et Cognitives, CNRS, 31 ch. Joseph Aiguer, 13402 Marseille, Cedex 20, France

**Mari Tervaniemi** Cognitive Brain Research Unit, Department of Psychology, P.O. Box 13, FIN-00014 University of Helsinki, Finland

**Barbara Tillmann** Université Claude Bernard Lyon 1, Laboratoire Neurosciences et Systèmes Sensoriels, CNRS-UMR 5020, 50 Avenue Tony Garnier, F-69366 Lyon Cedex 07, France

**Petri Toiviainen** Department of Music, University of Jyväskylä, Jyväskylä, Finland

**Laurel J. Trainor** Department of Psychology, McMaster University, Hamilton, ON, L8S 4B2 Canada

**Mark Jude Tramo** Department of Neurology, Harvard Medical School and Massachusetts General Hospital, Boston, Massachusetts 02114-2696, USA

**Sandra E. Trehub** Department of Psychology, University of Toronto at Mississauga, Mississauga, ON L5L 1C6, Canada

**Robert J. Zatorre** Montreal Neurological Institute, McGill University, Montreal, Quebec, H3A 2B4, Canada

PART I

# THE ORIGINS OF MUSIC

CHAPTER 1

# MUSICAL PREDISPOSITIONS IN INFANCY: AN UPDATE

SANDRA E. TREHUB

For contributors to this volume on the cognitive neuroscience of music, the consideration of musical predispositions seems reasonable. By contrast, the larger scientific community is highly skeptical about links between music and biology. For example, Pinker,[1] who popularized the notion of a *language instinct*, dismisses music as 'auditory cheesecake'—a pleasant but superfluous confection. 'As far as biological cause and effect are concerned, music is useless. It shows no signs of design for attaining a goal such as long life, grandchildren, or accurate perception and prediction of the world. Compared with language, vision, social reasoning, and physical know-how, music could vanish from our species and the rest of our lifestyle would be virtually unchanged' (p. 528).[2]

Pinker's[2] thinking in this regard is flawed. He makes no distinction between musical *competence* and *performance*, or between underlying knowledge and demonstrable skill, as he does for language. He notes that while *all* normal children speak and comprehend language without explicit instruction, many adults are unmusical, as evidenced by their inability to sing in tune. Although developmental disorders of language occur in the absence of other disabilities,[3] developmental disorders of music are rare[4] even though music perception requires finer pitch resolution than does speech perception.

Pinker[2] emphasizes the training and practice required for mastering a musical instrument while de-emphasizing the daily language (speaking and listening) practice that all children experience. Caregivers, peers, and others function as language mentors, modelling target behaviours in age-appropriate ways. Similarly, informal musical mentoring is much more common than formal musical training in much of the world. Admittedly, individuals without certain kinds of mentoring or training may be limited in their explicit musical knowledge, but their implicit knowledge is similar, in many respects, to that of trained individuals.[5] Pinker also contrasts the wide variations in musical ability (i.e. tuneless singing to expert performance) with the presumed uniformity in language ability (i.e. the basic ability to speak and listen). In so doing, he applies a magnifying lens to music, which exaggerates individual differences, and a minifying lens to language, which reduces or eliminates individual differences. There is no evidence, however, of anything other than a normal distribution of language and musical abilities. Finally, Pinker's vision of life without music betrays considerable ethnocentrism—music as material for concerts, dance halls, and movie soundtracks rather than something entwined in the fabric of life. It ignores the historical and cross-cultural importance of music in ritual ceremonies,[6–9] work,[10,11] and child care[12] as well as the inextricable links between music and movement.[13]

Miller,[14] by contrast, considers music to exemplify many of the classic criteria for a complex human adaptation: (1) No culture in any period of recorded history has been without music (i.e. universality). (2) The development of musical abilities is orderly.[15] (3) The ability is widespread in the sense that almost every normal adult can appreciate music and carry a tune (i.e. basic perceptual and performing skills). (4) Adults can recognize thousands of melodies, implying specialized memory. (5) Special-purpose cortical mechanisms are involved.[16] (6) There are analogues in the signals of other species such as songbirds,[17] gibbons,[18] and whales,[19] raising the possibility of convergent evolution. (7) Finally, music can evoke strong emotions,[20–23] which implies receptive as well as productive adaptations.

If music is a complex biological adaptation rather than a by-product of other evolutionary processes, it must have conferred survival or reproductive benefits through the cumulative effect, over generations, of natural selection or sexual selection.[14,24,25] The costs of music production—energy expenditure—must be balanced by benefits. Because no survival benefits are obvious, Miller[14] argues for the evolution of music on the basis of its reproductive benefits. In other words, he makes a case for music as sexual courtship. He argues that although music is typically performed *in* groups, it is not *for* groups (but see Ref. 26; Huron, Chapter 4, this volume). In the context of the tribal group, dancing (i.e. vigorous movement to music) could have provided a means for young women to assess the strength, endurance, and motor coordination of potential mates. The relative stability of ancestral social groups would have allowed women to scrutinize prospective partners over an extended period. Miller[14] acknowledges, however, that aesthetic preferences would have been equally effective as a basis for sexual selection.

Sexual (or natural) selection of this nature would lead to the replication of genes related to receptive and productive musical potential. In fact, genetic correlates of musical pitch processing have been identified.[27] Musical processing predispositions or biases might be evident in infancy, well before they have obvious utility. As Cross[28] argues, it may be possible to observe the impact of evolutionary processes in infancy before any musical culture has left its indelible impression.

## Uncovering musical predispositions

Accessing the music processing skills of preverbal listeners is no easy matter. Nevertheless, researchers have made considerable headway by means of conditioning procedures,[29] habituation procedures,[30,31] and preference procedures.[32,33] The most common conditioning procedure involves the presentation of a repeating melody or tone sequence from a laterally displaced loudspeaker while infants watch a puppet show. Periodically, the repeating pattern changes in some way, which prompts infants to turn towards the loudspeaker. Correct responses (i.e. turns) immediately following the change are rewarded by the presentation of a colourful mechanical toy (e.g. dancing bear, acrobatic monkey). Failure to turn at such times and turns at other times have no consequences. Such response-contingent rewards motivate attentive listening for 10–15 min, or 20–30 test trials. Aside from the bears, monkeys, puppets, and other paraphernalia that facilitate communication with nonverbal participants, the task is similar to adult same–different tasks. Infants turn

to the loudspeaker when they consider the comparison pattern to be different from the standard pattern (i.e. response of *different*). They continue watching the puppet show (i.e. no turning response) when they fail to notice the change or if they consider the change unimportant (i.e. response of *same*). Significantly more responses (turns) to trials with changes than to those without changes indicate that infants can detect the target change. Successful detection of such changes also sheds light on infants' encoding of the standard melody. For example, infants' detection of a pitch change in the comparison melody implies that they noticed the specific pitches or pitch patterning of the original melody.

### Relational processing of pitch and duration

Infants' perception of frequency,[34,35] timing,[36,37] and timbre[38,39] is finer than that required for musical purposes. Moreover, perceptual grouping principles that are relevant to music[40,41] are operative in infancy. Like adults, infants group isochronous tone sequences on the basis of similarities in pitch, loudness, and timbre.[42,43]

Infants go well beyond these minimal criteria for music perception. They recognize a melody in transposition, that is, when its pitch level is shifted upward or downward so long as the relations between tones are preserved.[30,44] An analogous situation prevails with respect to timing. Infants recognize a tone sequence when the tempo is altered so long as the relative durations remain unchanged.[45] In other words, they can focus on relational attributes of a melody, transforming a potentially unmanageable feat—encoding the exact pitches and durations—into a manageable task. Saffran[33,46] contends that infants preferentially attend to absolute rather than relative pitch cues in music, but support for her claim is lacking.[47] In principle, infants could retain absolute information along with relational information. In fact, infants and adults differentiate instrumental performances heard previously from subtly different performances of the same melody.[48] On balance, however, relational information seems to dominate melodic processing in infancy as in adulthood.[49,50]

One relational feature, melodic contour, is especially potent. For example, infants notice contour similarities and differences when the standard and comparison melodies are separated by a full 15 s[30] or when distractor tones are inserted between standard and comparison sequences.[29] Indeed, there are indications that pitch contour is the most salient musical feature for infant listeners.[50,51] Pitch contour may also be the most salient feature of mothers' speech to prelinguistic infants.[52–55] In the case of adult listeners, contour processing seems to be fundamental and relatively impervious to musical experience.[15] During contour-processing tasks, for example, the amplitude and latency of event-related potentials do not differ for musicians and nonmusicians, in contrast to interval-processing tasks, where experience-dependent differences are apparent.[56]

### Interval processing

For adults, the extraction of pitch information is not limited to melodic contour, except for novel, unconventional melodies.[57,58] When adults hear unfamiliar but conventional melodies (i.e. those that conform to familiar musical conventions), they remember more detailed pitch information than they would otherwise.[59–61] In the case of highly familiar

melodies, adults are thought to encode and retain precise pitch distances between tones (i.e. intervals) but not the absolute pitches.[62] There are indications, however, that even untrained adults retain the pitch level of familiar music, as reflected in their reproduction of familiar pop songs within two semitones of the original recordings.[63] More compelling evidence of the encoding of absolute pitch cues comes from adults' ability to distinguish the original pitch level of television soundtracks from versions that are pitch shifted by as little as a semitone.[64]

Despite infants' contour processing bias,[29] they can extract detailed interval information from patterns that elicit enhanced pitch processing in adults. For example, infants detect interval changes in brief melodies that conform to the conventions of their culture-to-be, but they fail to do so for melodies that violate those conventions.[65–68] It is highly unlikely that 6-month-old infants have acquired such musical knowledge. Indeed, the available evidence indicates that 6- to 10-month-old infants lack implicit knowledge of Western musical conventions.[61,69,70] The mistaken impression that they are sensitive to these conventions arises from commonalities in musical structure across cultures.[49,50]

In studies of infant melody perception, the presence of a prominent seven-semitone interval, or *perfect fifth*, has been associated with success in interval discrimination, and its absence with failure.[65–68,71] The *perfect fifth* is not alone in enhancing infants' perception of pitch relations.[72] Rather, infants, children, and adults detect interval changes more easily in the context of small-integer ratios—the octave (2 : 1), *perfect fifth* (3 : 2), and *perfect fourth* (4 : 3)—than in the context of large-integer ratios such as the *tritone* (45 : 32 ratio).[72–74] For example, Schellenberg and Trehub[72] had 6-month-old infants listen to a repeating tone sequence consisting of alternating tones separated by a *perfect fifth* (seven semitones), *perfect fourth* (five semitones), or *tritone* (six semitones). Their task was to respond (i.e. by turning) whenever a semitone change occurred. As can be seen in Figure 1.1, they succeeded when the standard pattern was consonant (i.e. *perfect fourth* or *fifth*) but not when it was dissonant (i.e. *tritone*). The implication is that *perfect fifths* and *fourths* are

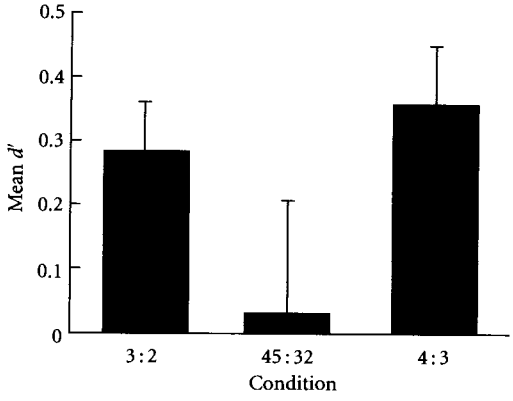



**Figure 1.1** Infant discrimination ($d'$) scores as a function of the frequency ratios of tones in melodic (sequential) intervals. Figure reprinted from Schellenberg and Trehub[72] with permission. Error bars indicate standard errors.

inherently easier to encode than are tritones. No doubt, the inherent ease of processing consonant intervals has contributed to their prominence in most musical systems.[50,75,76] Historical and cross-cultural records document the inequality of intervals, with numerous references to the beauty of octaves, *perfect fifths*, and *perfect fourths*[77,78] and the ugliness of the *tritone*—'the devil in music' in medieval times.[79]

Not only do consonant intervals promote ease of processing, they also influence infant attention and affect. For example, infants are more attentive and exhibit more positive affect when listening to consonant music than to music with many dissonant intervals.[80–82] In principle, these preferences could arise from the predominance of consonant intervals in the infant's environment (e.g. speech, music, environmental sounds). In the absence of such environmental input, as is the case for newborn infants whose deaf parents communicate by means of sign language, a similar pattern of preferences is evident.[83]

### Processing of scale structure

Scales across cultures are diverse, but they exhibit some common features. For example, the octave is typically divided into five to seven discrete pitches, a design feature that is often attributed to cognitive constraints.[84] Another similarity is the ubiquitous *perfect fifth*, which is associated with enhanced processing in listeners of all ages.[72,74,85] Scales, for the most part, have unequal steps[86]—for example, one- and two-semitone steps in diatonic scales—a feature that is thought to confer processing advantages.[87,88]

To evaluate potential processing advantages of equal steps in scales, Trehub *et al.*[70] presented 9-month-old infants and adults with one of three ascending–descending scales: (1) the major scale, which is highly familiar to adults and potentially familiar to infants, (2) an artificial analogue of the major scale that was created by dividing the octave into 11 equal units, and selecting a seven-tone subset (featuring 1- and 2-unit steps), and (3) an artificial equal-step scale that was created by dividing the octave into seven equal steps.[89,90] These scales are depicted in Figure 1.2. Infants and adults were required to detect a mistuned tone (0.75- and 0.5-semitone change for infants and adults, respectively) in the comparison pattern. Adults' performance was highly accurate on the familiar major scale and equally poor on the two unfamiliar scales. By contrast, scales with unequal steps, major or artificial, enhanced infant performance relative to their performance on the equal-step scale (see Figure 1.3). Thus, unequal-step scales, which are universal or near-universal, facilitate musical pitch processing in infancy.

Infants' performance cannot be attributed to familiarity. It is possible, although highly unlikely, that infants heard adults or siblings practising the major scale, but they could not have heard the invented unequal-step scale that generated similar performance. By extension, other instances of enhanced processing—small-integer ratios, for example—are unlikely to stem from exposure. Adults' performance is informative in revealing how culture-specific experience, such as long-term exposure to the major scale, can override inherent processing biases (e.g. advantages of unequal-step scales in general).

There are other examples of culture-specific exposure attenuating or eliminating initial processing biases. For example, pitch processing in infancy is enhanced when the standard and comparison patterns are presented in related keys—specifically, keys standing in a 3 : 2 ratio—rather than unrelated keys.[66] For adults, such enhancement is limited to melodies that

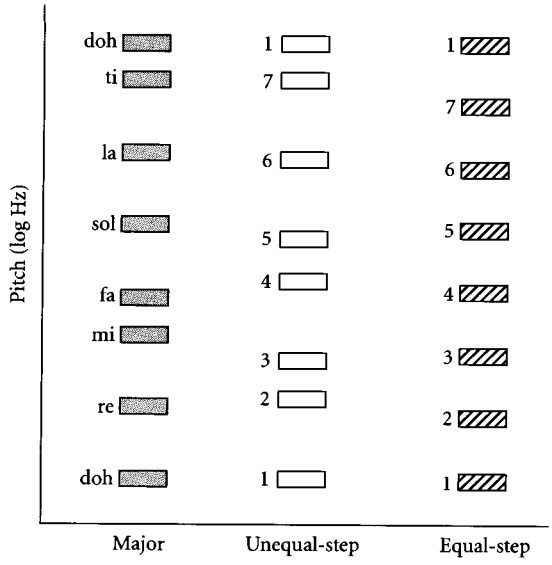



**Figure 1.2** Schematic depiction of the major, artificial unequal-step, and equal-step scales. Figure reprinted from Trehub *et al.*[70] with permission.

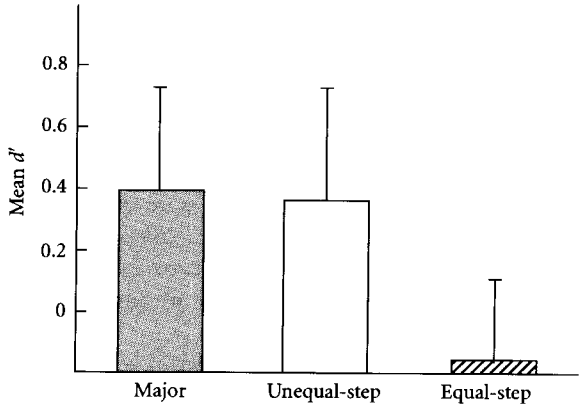



**Figure 1.3** Infant discrimination ($d'$) scores on the ascending-descending major scale, artificial, unequal-step scale, and equal-step scale. Error bars indicate standard errors. Figure reprinted from Trehub *et al.*[70] with permission.

are structured in a conventional manner.[66] Similarly, melodic redundancy enhances infants' retention of conventional and unconventional melodies, but comparable redundancy enhances adult performance for conventional melodies only.[61] In general, musics of the world capitalize on human processing predispositions (e.g. the preference for consonance and

unequal scale steps), which means that musical experience tends to intensify the initial biases or predispositions.

### Temporal processing

Similar adult–infant parallels are observable in the temporal domain (see Drake and Bertrand, Chapter 2, this volume). For example, infants differentiate tone sequences with identical pitches but contrasting rhythmic arrangements.[31,45,91] They also impose rhythmic groupings on isochronous tone sequences on the basis of pitch or timbre similarities.[42,43] The temporal organization of a sequence also has implications for infants' processing of pitch and timing details.[92] For example, melodic processing is enhanced when tone durations are related by binary ratios,[93,94] as is the case for children and adults.[95–97] Temporal processing predispositions (Drake and Bertand, Chapter 2, this volume) may account for the fact that adults' preferred rhythmic arrangements of melodies generate better infant performance than do non-preferred arrangements.[94] They may also account for superior infant performance on patterns with regular rather than syncopated rhythms.[93]

### Hemispheric specialization

Well before infants understand language, they show a right-ear (i.e. left hemisphere) advantage for speech and a left-ear (i.e. right hemisphere) advantage for music.[98,99] Moreover, 8-month-old infants exhibit a left-ear superiority for contour processing and a right-ear superiority for interval processing.[100] Such laterality effects parallel those reported for adults[101,102] (but see Zatorre, Chapter 16, this volume).

## Maternal music

The relevance of these processing predispositions to the daily lives of infants is unclear. Typical laboratory studies use musical stimuli (e.g. sequences of pure tones) that differ dramatically from the music that infants experience. Caregivers across cultures produce distinctive, solo performances of songs for infant listeners.[12,103–105] They also speak to infants in a sing-song manner that incorporates a number of musical features.[52,106,107] For example, mothers' speech differs from women's usual speech (e.g. pitch level, pitch contours, tempo, rhythm) in ways that reflect its greater emotional expressiveness.[52,55,108] Infants prefer this maternal speaking style to the usual adult style.[109–111]

When caregivers sing to infants, they generally use a special-purpose repertoire consisting of lullabies and play songs.[12,104] To the naïve adult listener, the tunes of unfamiliar lullabies, including those from foreign musical cultures, sound much like those of familiar lullabies. In fact, adults successfully identify unfamiliar lullabies from pairs of novel lullabies and nonlullabies matched on culture of origin and tempo.[112] Adults also perceive foreign lullabies as simpler and more repetitive than nonlullabies.[113]

Among the performance alterations that distinguish singing to infants from other singing are raised pitch level, decreased tempo, and an emotive voice quality.[114–116] In addition, mothers' song renditions are tailored to the age and needs of the listener. For example, maternal singing is approximately one-semitone higher in pitch level for infant

than for preschool audiences, and the articulation of lyrics is more slurred for infants than for preschoolers.[117] Expressive features of the 'maternal' style are also evident when fathers sing to their infant offspring[116,118,119] and when preschoolers sing to their infant siblings.[120]

There are parallels between maternal speech and singing to infants,[12,51,104] but there are differences as well. In general, women's singing is higher pitched than their speech, but maternal speech to infants is higher pitched than maternal singing.[121] This situation results from a three- to four-semitone increase in pitch level when mothers speak to infants,[122] in contrast to an increase of approximately one-semitone when they sing to infants.[114,116]

Despite considerable variation across mothers, there is dramatic intra-individual consistency. When mothers sing the same song to infants on different occasions separated by a week or more, their pitch level and tempo (see Table 1.1) are virtually identical.[121] Variations in mothers' pitch and tempo are smaller than those observed in adults' repeated renditions of pop or folk songs.[63,123,124] By contrast, mothers' repeated utterances to infants exhibit large variations in pitch and tempo (see Table 1.1) but relatively stable rhythms. If singing to infants affects maternal mood or state, it could contribute to the stability of pitch and tempo over extended periods—a possible consequence of mood-dependent memory.[125] Frequent singing to infants may also implicate motor memory. In any event, mothers' sung performances seem to become ritualized, which could promote infants' recognition of mother's voice and the songs she sings.

### Responsiveness to maternal music

How do infants respond to these maternal singing rituals? In the newborn period and beyond, infants listen significantly longer to audio recordings of infant-directed singing than to typical informal singing.[126,127] They also exhibit greater attention to higher- than to lower-pitched versions of the same song.[128] These preferences indicate the potency of the maternal style of singing, even in the absence of the familiar voice and expressive gestures that typically accompany maternal performances.

Table 1.1 Absolute pitch and tempo differences for mothers' infant-directed speech and singing across weeks

| Mother | Pitch (semitones) | | Tempo (beats per minute) | |
|---|---|---|---|---|
| | Speech | Singing | Speech | Singing |
| 1 | 4.0 | 1.0 | 16 | 0 |
| 2 | 6.0 | 0.0 | 32 | 4 |
| 3 | 5.0 | 1.0 | 6 | 0 |
| 4 | 3.0 | 2.0 | 32 | 6 |
| 5 | 4.0 | 0.0 | 88 | 0 |
| 6 | 4.5 | 0.0 | 0 | 0 |
| 7 | 5.5 | 0.0 | 4 | 12 |
| 8 | 16.0 | 0.0 | 16 | 0 |
| 9 | 6.5 | 3.0 | 18 | 18 |
| 10 | 0.0 | 0.0 | 8 | 0 |
| 11 | 0.0 | 1.0 | 31 | 0 |

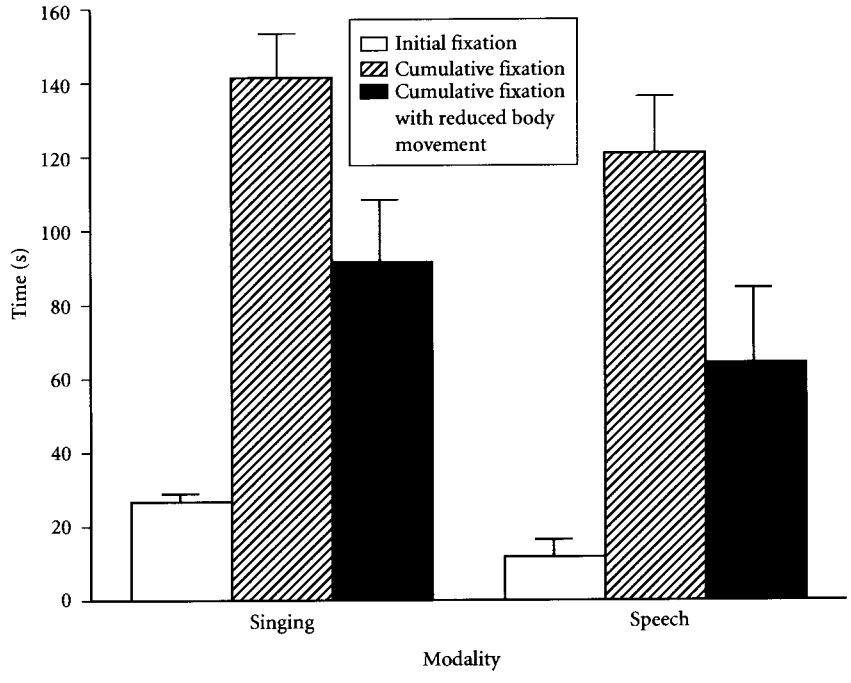



Figure 1.4 Initial fixation, cumulative fixation, and cumulative fixation with minimal body movement as a function of maternal speech and singing. Error bars indicate standard errors.

When 6-month-old infants view (and hear) videotaped speech and singing by their own mother, they exhibit longer visual fixations and greater reductions in body movement during singing than during speaking episodes (see Figure 1.4).[129,130] These sung performances have a hypnotic effect on infants, who remain glued to their mother's image. Maternal speaking is also engaging but not nearly as engaging as maternal singing.

There has been much discussion of maternal speech and singing as modulators of infant arousal,[12,52,108] but the available research has focused largely on attention rather than arousal. To explore the arousal consequences of maternal speech and singing, Shenfield, Trehub, and Nakata[131] took saliva samples from 6-month-old infants before the onset of maternal speech and singing (on different days) as well as 20 and 25 min later. As can be seen in Figure 1.5, cortisol levels decreased systematically across the three sampling periods of the singing condition. There was a similar decrease in cortisol levels 20 min after speaking onset, but this decrease was not sustained at the 25-minute period (see Figure 1.5). Maternal music as a means of optimizing infant mood or arousal parallels adolescents' and adults' use of music for self-regulation.[132,133] It appears that some forms of music, maternal singing in particular, modulate arousal level in healthy, nondistressed infants, as they do for distressed newborns in intensive care units.[134] There are indications, moreover, that maternal singing is a more effective regulator of arousal than is maternal speech.

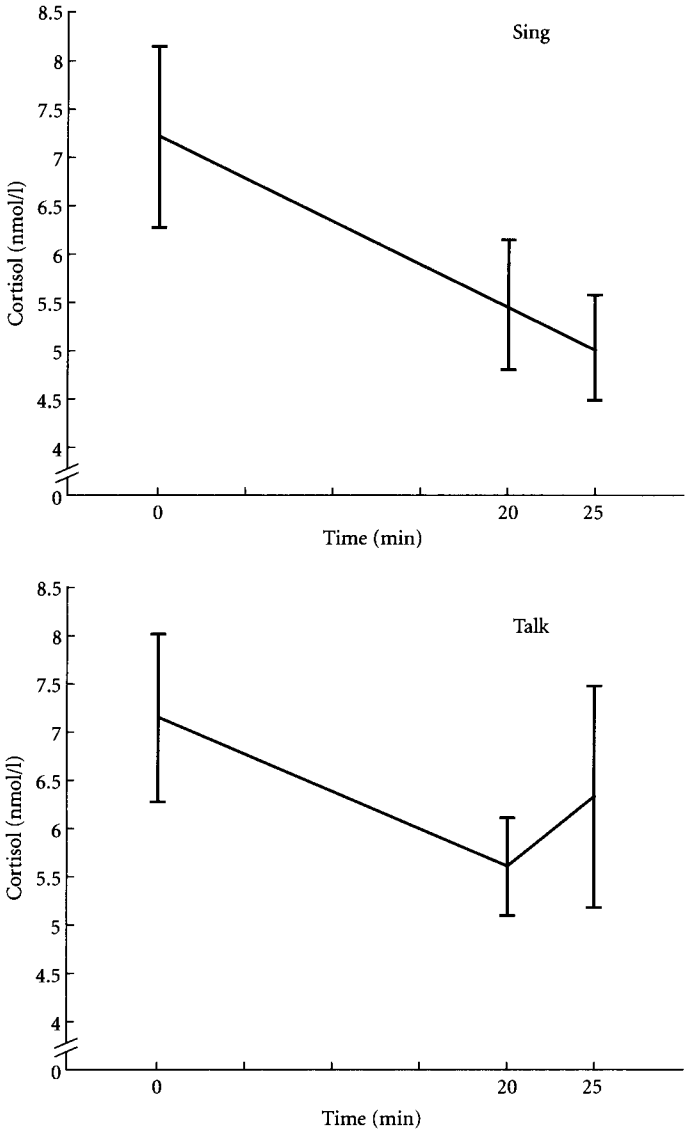



**Figure 1.5** Baseline cortisol levels and cortisol levels at 20 and 25 min after the onset of singing (upper panel) and speaking (lower panel). Error bars indicate standard errors.

## Biological significance of singing to infants

The ubiquity of maternal singing and the impact of such singing on infant attention and arousal raise the possibility that maternal singing could have enhanced infant survival in difficult (e.g. ancestral) conditions. To the extent that maternal singing optimizes infant

mood, it could contribute to infant growth and development by facilitating feeding, sleeping, and even learning. Children's extended period of helplessness creates intense selection pressures for parental commitment and for infant behaviours to reward such commitment. Falling asleep to lullabies or entering trance-like states to performances of other songs might be suitable rewards for maternal effort. In general, favourable consequences of maternal singing on infant arousal, whether through cry reduction, sleep induction, or positive affect, would contribute to infant well-being while sustaining such maternal behavior. Presumably, the healthy and contented offspring of singing mothers would be more likely to pass on their genes than would the offspring of non-singing mothers.

Maternal singing may also be self-soothing, providing a safe outlet for negative as well as positive feelings.[12] There are numerous examples of lullabies across cultures that combine laments on life with sleep-inducing singing.[12,135] Maternal singing is likely to strengthen the emotional ties between mother and infant[106,119] just as singing in other contexts reduces the psychological distance between singer and listener.[136,137] Indeed, maternal singing may set the stage for the subsequent role of music in group bonding.[26,138] In any case, singing to infants seems to enhance maternal as well as infant well-being.

If women's expressive singing captures the attention and hearts of infant listeners, it may have been equally captivating to young adult males in the distant past. In some avian species, females sing to attract mates, and they lead duets with their male partner.[17,139] There is speculation, moreover, that such duetting fosters long-term pair bonding.[140] These notions are inconsistent with Miller's[14] view of music as male sexual courtship. To evaluate his hypothesis, Miller compared the recorded output of prominent jazz, rock, and classical musicians. Male musicians produced 10 times as much music as female musicians, their output reaching its peak near the age of peak mating effort. This led Miller[14] to conclude that music continues to function as a courtship display to attract females. To date, however, there is no evidence of sexual dimorphism in musical ability or achievement that cannot be accounted for by sociocultural factors.

Tooby and Cosmides[141] offer a radically different evolutionary account of the universal attraction to the imaginative arts (e.g. fiction, visual art, music). Unlike human activities that have fitness-enhancing consequences in the external world (e.g. sex) or on the body (e.g. eating), aesthetic activities promote fitness-enhancing activities in the human brain. According to Tooby and Cosmides,[140] readying the brain for its mature role is one of the most critical and most exacting developmental tasks. They contend that artistic or imaginative actions (or action tendencies) play a key role in mental development. In line with their view, infants' attraction to music may reflect motivational guidance systems designed to develop and fine-tune the brain mechanisms that subserve cognitive and social competence. Part of what may make music so important in this regard is its lack of immediate consequences (Cross, this volume)—the very factor that leads others[2] to dispute its biological basis.

## Conclusion

It is clear that infants do not begin life with a blank musical slate. Instead, they are predisposed to attend to the melodic contour and rhythmic patterning of sound sequences, whether music or speech. They are tuned to consonant patterns, melodic as well as

harmonic, and to metric rhythms. Surely these predispositions are consistent with a biological basis for music, specifically, for the core abilities that underlie adult musical skill in all cultures. As Cross[28] notes, evolution acts on the mind by shaping infant predispositions, and culture shapes the expression of those predispositions. In effect, music is part of our nature as well as being part of our culture.

Mothers cater to infants' musical inclinations by singing regularly in the course of caregiving and by adapting their singing style in ways that are congenial to infant listeners. These ritualized vocal interactions may reflect caregivers' predisposition to share affect and forge emotional ties by means of temporal synchrony.[142] Perhaps it is not surprising that infants are predisposed to attend to and appreciate the species-typical vocalizations of their primary caregiver. Is it surprising that they are also predisposed to attend to specific structural features of music? Only if music is viewed in a narrow sense, as fully developed musical systems of particular cultures. Viewed broadly, music embraces what all musical systems have in common. In that sense, infants begin life as musical beings, being responsive to the primitives or universals that are the foundation of music everywhere.

## Acknowledgements

The preparation of this paper was assisted by grants from the Natural Sciences and Engineering Research Council and the Social Sciences and Humanities Research Council of Canada.

## References

1. **Pinker, S.** (1994) *The Language Instinct: How the Mind Creates Language.* New York: William Morrow.
2. **Pinker, S.** (1997) *How the Mind Works.* New York: Norton, p. 528.
3. **Fletcher, P.** (1999) Specific language impairment. In M. Barrett (ed.) *The Development of Language.* Hove, East Sussex, UK: Psychology Press, pp. 349–71.
4. **Peretz, I., J. Ayotte, R. Zatorre, J. Mehler, P. Ahad, V. Penhune,** and **B. Jutras** (2002) Congenital amusia: a disorder of fine-grained pitch discrimination. *Neuron* 33, 185–91.
5. **Smith, J. D., D. G. Kemler Nelson, L. A. Grohskopf,** and **T. Appleton** (1994) What child is this? What interval was that? Familiar tunes and music perception in novice listeners. *Cognition* 52, 23–54.
6. **Basso, E.** (1985) *A Musical View of the Universe: Kalapalo Rhythmic and Ritual Performance.* Philadelphia: University of Pennsylvania Press.
7. **Dissanayake, E.** (1992) *Homo Aestheticus: Where Art Comes From and Why.* New York: Free Press.
8. **Merriam, A. P.** (1964) *The Anthropology of Music.* Evanston, Ill: Northwestern University Press.
9. **Sarno, L.** (1993) *Song from the Forest: My Life among the Ba-Benjelle Pygmies.* Boston: Houghton Mifflin.
10. **Deng, F. M.** (1973) *The Dinka and their Songs.* London: Oxford University Press.
11. **Firth, R.** (1961) *Elements of Social Organization.* London: Watts.
12. **Trehub, S. E.** and **L. J. Trainor** (1998) Singing to infants: lullabies and play songs. *Advances in Infancy Res.* 12, 43–77.

13. **Blacking, J.** (1995) *Music, Culture, and Experience.* London: University of Chicago Press.
14. **Miller, G.** (2000) Evolution of human music through sexual selection. In N. L. Wallin, B. Merker, and S. Brown (eds) *The Origins of Music.* Cambridge, MA: MIT Press, pp. 329–60.
15. **Dowling, W. J.** (1999) The development of music perception and cognition. In D. Deutsch (ed.) *The Psychology of Music.* San Diego, CA: Academic Press, pp. 603–25.
16. **Peretz, I. and J. Morais** (1993) Specificity for music. In F. Boller and J. Grafman (eds) *Handbook of Neuropsychology,* vol. 8. Amsterdam: Elsevier, pp. 373–90.
17. **Slater, J. B.** (2000) Birdsong repertoires: their origins and use. In N. L. Wallin, B. Merker, and S. Brown (eds) *The Origins of Music.* Cambridge, MA: MIT Press, pp. 49–63.
18. **Geissmann, T.** (2000) Gibbon songs and human music from an evolutionary perspective. In N. L. Wallin, B. Merker, and S. Brown (eds) *The Origins of Music.* Cambridge, MA: MIT Press, pp. 103–23.
19. **Payne, K.** (2000) The progressively changing songs of humpback whales: a window on the creative process in a wild animal. In N. L. Wallin, B. Merker, and S. Brown (eds) *The Origins of Music.* Cambridge, MA: MIT Press, pp. 135–50.
20. **Blood, A. J. and R. J. Zatorre** (2001) Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc. Nat. Acad. Sci.* 98, 11818–23.
21. **Gabrielsson, A.** (2001) Emotions in strong experiences with music. In P. N. Juslin and F. A. Sloboda (eds) *Music and Emotion: Theory and Research.* Oxford: Oxford University Press, pp. 431–49.
22. **Peretz, I.** (2001) Listen to the brain: a biological perspective on musical emotions. In P. N. Juslin and J. A. Sloboda (eds) *Music and Emotion: Theory and Research.* Oxford: Oxford University Press, pp. 105–34.
23. **Sloboda, J. A.** (1992) Empirical studies of emotional response to music. In M. R. Jones and S. Holleran (eds) *Cognitive Bases of Musical Communication.* Washington, DC: American Psychological Association, pp. 33–50.
24. **Dawkins, R.** (1996) *Climbing Mount Improbable.* London: Penguin Books.
25. **Williams, G. C.** (1966) *The Selfish Gene. Adaptation and Natural Selection.* Princeton, NJ: Princeton University Press.
26. **Brown, S.** (2000) Evolutionary models of music: from sexual selection to group selection. In F. Tonneau and N. S. Thompson (eds) *Perspectives in Ethology, vol.* 13: *Evolution, Culture, and Behavior.* New York: Kluwer Academic, pp. 231–81.
27. **Drayna, D., A. Manichaikul, M. de Lange, H. Sneider,** and T. Spector (2001) Genetic correlates of musical pitch recognition in humans. *Science* 291, 1969–72.
28. **Cross, I.** (2001) Music, mind, and evolution. *Psychol. Music* 29, 95–102.
29. **Trehub, S. E., D. Bull,** and L. A. Thorpe (1984) Infants' perception of melodies: the role of melodic contour. *Child Dev.* 55, 821–30.
30. **Chang, H. W. and S. E. Trehub** (1977) Auditory processing of relational information by young infants. *J. Exp. Child Psychol.* 24, 324–31.
31. **Chang, H. W. and S. E. Trehub** (1977) Infants' perception of temporal grouping in auditory patterns. *Child Dev.* 48, 1666–70.
32. **Krumhansl, C. L. and P. W. Jusczyk** (1990) Infant's perception of phrase structure in music. *Psychol. Sci.* 1, 70–3.

33. **Saffran, J. R.** and **G. J. Griepentrog** (2001) Absolute pitch in infant auditory learning: evidence for developmental reorganization. *Developmental Psychol.* 37, 74–85.

34. **Olsho, L. W., C. Schoon, R. Sakai, R. Turpin,** and **V. C. Sperduto** (1982) Preliminary data on frequency discrimination in infancy. *J. Acoust. Soc. Am.* 71, 509–11.

35. **Werner, L. A.** (1992) Interpreting developmental psychoacoustics. In L. A. Werner and E. W. Rubel (eds) Developmental Psychoacoustics. Washington, DC: American Psychological Association, pp. 47–88.

36. **Trehub, S. E., B. A. Schneider,** and **J. L. Henderson** (1995) Gap detection in infants, children, and adults. *J. Acoust. Soc. Am.* 98, 2532–41.

37. **Werner, L. A, G. C. Marean, C. F. Halpin, N. B. Spetner,** and **J. M. Gillenwater** (1992) Infant auditory temporal acuity: gap detection. *Child Dev.* 63, 260–72.

38. **Clarkson, M. G.** and **R. K. Clifton** (1995) Infants' pitch perception: inharmonic tonal complexes. *J. Acoust. Soc. Am.* 98, 1372–9.

39. **Trehub, S. E., M. Endman,** and **L. A. Thorpe** (1990) Infants' perception of timbre: Classification of complex tones by spectral structure. *J. Exp. Child Psychol.* 49, 300–13.

40. **Bregman, A. S.** (1990) *Auditory Scene Analysis.* Cambridge, MA: MIT Press.

41. **Deutsch, D.** (1999) Grouping mechanisms in music. In D. Deutsch (ed.) *The Psychology of Music.* San Diego, CA: Academic Press, pp. 299–348.

42. **Thorpe, L. A.** and **S. E. Trehub** (1989) Duration illusion and auditory grouping in infancy. *Developmental Psychol.* 25, 122–7.

43. **Thorpe, L. A., S. E. Trehub, B. A. Morrongiello,** and **D. Bull** (1988) Perceptual grouping by infants and preschool children. *Developmental Psychol.* 24, 484–91.

44. **Trehub, S. E., L. A. Thorpe,** and **B. A. Morrongiello** (1987) Organizational processes in infants' perception of auditory patterns. *Child Dev.* 58, 741–9.

45. **Trehub, S. E.** and **L. A. Thorpe** (1989) Infants' perception of rhythm. Categorization of auditory sequences by temporal structure. *Canadian J. Psychol.* 43, 217–29

46. **Saffran, J. R.** (2003) Absolute pitch in infancy and adulthood: the role of tonal structure. *Developmental Science* 6, 37–45.

47. **Trehub, S. E.** (2003) Absolute and relative pitch processing in tone learning tasks. *Developmental Science* 6, 46–7.

48. **Palmer, C., M. K. Jungers,** and **P. W. Jusczyk** (2001) Episodic memory for musical prosody. *J. Mem. Lang.* 45, 526–45.

49. **Trehub, S. E.** (2000) Human processing predispositions and musical universals. In N. L. Wallin, B. Merker, and S. Brown (eds) *The Origins of Music.* Cambridge, MA: MIT Press, pp. 427–48.

50. **Trehub, S., E. G. Schellenberg,** and **D. Hill** (1997) The origins of music perception and cognition: a developmental perspective. In I. Deliège and J. Sloboda (eds) *Perception and Cognition of Music.* Hove, UK: Psychology Press, pp.103–28.

51. **Trehub, S. E., L. J. Trainor,** and **A. M. Unyk** (1993) Music and speech processing in the first year of life. *Advances Child Dev. Behav.* 24, 1–35.

52. **Fernald, A.** (1991) Prosody in speech to children: prelinguistic and linguistic functions. *Annals Child Dev.* 8, 43–80.

53. **Fernald, A.** and **P. K. Kuhl** (1987) Acoustic determinants of infant preference for motherese. *Infant Behav. Dev.* 10, 279–93.

54. **Lewis, M. M.** (1951) *Infant Speech.* London: Routledge and Kegan Paul.

55. **Papoušek, M.** (1992) Early ontogeny of vocal communication in parent–infant interactions. In H. Papoušek, U. Jürgens, and M. Papoušek (eds) *Nonverbal Vocal Communication: Comparative and Developmental Approaches.* Cambridge: Cambridge University Press, pp. 230–61.

56. **Trainor, L. J., R. N. Desjardins, and C. Rockel** (1999) A comparison of contour and interval processing in musicians and nonmusicians using event-related potentials. *Australian J. Psychol.* 51, 147–53.

57. **Bartlett, J. C. and W. J. Dowling** (1980) Recognition of transposed melodies: a key-distance effect in developmental perspective. *J. Exp. Psychol.: Hum. Percept. Perform.* 6, 501–15.

58. **Dowling, W. J.** (1978) Scale and contour: two components of a theory of memory for melodies. *Psychological Rev.* 85, 341–54.

59. **Cuddy, L. L., A. J. Cohen, and D. J. K. Mewhort** (1981) Perception of structure in short melodic sequences. *J. Exp. Psychol.: Hum. Percept. Perform.* 7, 869–83.

60. **Krumhansl, C. L., J. J. Bharucha and E. J. Kessler** (1982) Perceived harmonic structure of chords in three related keys. *J. Exp. Psychol.: Hum. Percept. Perform.* 8, 24–36.

61. **Schellenberg, E. G. and S. E. Trehub** (1999) Culture-general and culture-specific factors in the discrimination of melodies. *J. Exp. Child Psychol.* 74, 107–27.

62. **Dowling, W. J. and D. Fujitani** (1971) Contour, interval, and pitch recognition in memory for melodies. *J. Acoust. Soc. Am.* 49, 524–31.

63. **Levitin, D. J.** (1994) Absolute memory for musical pitch: evidence from the production of learned melodies. *Percept. Psychophys.* 56, 414–23.

64. **Schellenberg, E. G. and S. E. Trehub** (2003). Good pitch memory is widespread. *Psychol. Sci.* 14, 262–6.

65. **Cohen, A. J., L. A. Thorpe, and S. E. Trehub** (1987) Infants' perception of musical relations in short transposed tone sequences. *Canadian J. Psychol.* 41, 33–47.

66. **Trainor, L. J. and S. E. Trehub** (1993) Musical context effects in infants and adults: key distance. *J. Exp Psychol.: Hum. Percept. Perform.* 19, 615–26.

67. **Trainor, L. J. and S. E. Trehub** (1993) What mediates infants' and adults' superior processing of the major over the augmented triad? *Music Percept.* 11, 185–96.

68. **Trehub, S. E., L. A. Thorpe, and L. J. Trainor** (1990) Infants' perception of *good* and *bad* melodies. *Psychomusicology* 9, 5–19.

69. **Trainor, L. J. and S. E. Trehub** (1992) A comparison of infants' and adults' sensitivity to Western musical structure. *J. Exp. Psychol.: Hum. Percept. Perform.* 18, 394–402.

70. **Trehub, S. E., E. G. Schellenberg, and S. B. Kamenetsky** (1999) Infants' and adults' perception of scale structure. *J. Exp. Psychol.: Hum. Percept. Perform.* 25, 965–75.

71. **Lynch, M. P., R. E. Eilers, D. K. Oller, and R. C. Urbano** (1990) Innateness, experience, and music perception. *Psychol. Sci.* 1, 272–6.

72. **Schellenberg, E. G. and S. E. Trehub** (1996) Natural musical intervals: evidence from infant listeners. *Psychol. Sci.* 7, 272–7.

73. **Schellenberg, E. G. and S. E. Trehub** (1994) Frequency ratios and the discrimination of pure tone sequences. *Percept. Psychophys.* 56, 472–8.

74. **Schellenberg, E. G. and S. E. Trehub** (1996) Children's discrimination of melodic intervals. *Developmental Psychol.* 32, 1039–50.

75. **Meyer, L. B.** (1956) *Emotion and Meaning in Music.* Chicago: University of Chicago Press.

76. **Sachs, C.** (1943) *The Rise of Music in the Ancient World: East and West.* New York: Norton.

77. **Bower, C.** (1980) Boethius, Anicius, Manlius Severinus. In S. Sadie (ed.) *The New Grove Dictionary of Music and Musicians*, vol. 2. London: Macmillan, p. 844.
78. **Hill, C. C.** (1986) Consonance and dissonance. In D. M. Randel (ed.) *The New Harvard Dictionary of Music*. Cambridge, MA: Belknap Press, pp. 197–9.
79. **Kennedy, M.** (1994) Song. *The Oxford Dictionary of Music*, 2nd edn. Oxford: Oxford University Press.
80. **Trainor, L. J.** and **B. M. Heinmiller** (1998) Infants prefer to listen to consonance over dissonance. *Infant Behav. Dev.* 21, 77–88.
81. **Zentner, M. R.** and **J. Kagan** (1996) Perception of music by infants. *Nature* 383, 29.
82. **Zentner, M. R.** and **J. Kagan** (1998) Infants' perception of consonance and dissonance in music. *Infant Behav. Dev.* 21, 483–92.
83. **Masataka, N.** (1999) Personal communication.
84. ***Burns*, E. M.** (1999) Intervals, scales, and tuning. In D. Deutch (ed.) *The Psychology of Music*, 2nd edn. San Diego, CA: Academic Press, pp. 215–64.
85. **Schellenberg, E. G.** and **S. E. Trehub** (1994) Frequency ratios and the perception of tone patterns. *Psychonomic Bulletin and Rev.* 1, 191–201.
86. **Sloboda, J. A.** (1985) *The Musical Mind: The Cognitive Psychology of Music*. Oxford: Clarendon Press.
87. **Butler, D.** (1989) Describing the perception of tonality in music: a critique of the tonal hierarchy theory and proposal for a theory of intervallic rivalry. *Music Percept.* 6, 219–42.
88. **Shepard, R. N.** (1982) Geometrical approximations to the structure of musical pitch. *Psychological Rev.* 89, 305–33.
89. **Jordan, D. S.** and **R. N. Shepard** (1987) Tonal schemas: evidence obtained by probing distorted scales. *Percept. Psychophys.* 41, 489–504.
90. **Shepard, R. N.** and **D. C. Jordan** (1984) Auditory illusions demonstrating that tones are assimilated to an internalized scale. *Science* 226, 1333–4.
91. **Demany, L., B. McKenzie,** and **E. Vurpillot** (1977) Rhythm perception in early infancy. *Nature* 266, 718–19.
92. **Trainor, L. J.** and **B. Adams** (2000) Infants' and adults' use of duration and intensity cues in the segmentation of tone patterns. *Percept. Psychophys.* 62, 333–40.
93. **Bergeson, T. R.** (2002) *Perspectives on Music and Music Listening in Infancy*. Unpublished doctoral dissertation, University of Toronto.
94. **Trehub S. E., D. S. Hill,** and **S. B. Kamenetsky** (1997) *Infants' Perception of Melodies with "Good" or "Bad" Rhythms*. Presented at the biennial meeting of the Society for Research in Child Development. Washington, DC.
95. **Drake, C.** (1993) Reproduction of musical rhythms by children. *Percept. Psychophys.* 53, 25–33.
96. **Drake, C.** (1998) Psychological processes involved in the temporal organization of complex auditory sequences: universal and acquired processes. *Music Percept.* 16, 11–26.
97. **Parncutt, R.** (1994) A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Percept.* 11, 409–64.
98. **Bertoncini, J., J. Morais, R. Bijeljac-Babic, S. McAdams, I. Peretz,** and **J. Mehler** (1989) Dichotic perception and laterality in neonates. *Brain Lang.* 37, 591–605.
99. **Best, C. T., H. Hoffman,** and **B. B. Glanville** (1982) Development of infant ear asymmetries for speech and music. *Percept. Psychophys.* 31, 75–85.

100. **Balaban, M. T., L. M. Anderson, and A. B. Wisniewski** (1998) Lateral asymmetries in infant melody perception. *Developmental Psychol.* 34, 39–48.

101. **Peretz, I.** (1987) Shifting ear differences in melody comparison through transposition. *Cortex* 23, 317–23.

102. **Peretz, I. and M. Babaï** (1992) The role of contour and intervals in the recognition of melody parts: Evidence from cerebral asymmetries in musicians. *Neuropsychologia*, 30, 277–92.

103. **Brakeley, T. C.** (1950) Lullaby. In M. Leach and J. Fried (eds), *Standard Dictionary of Folklore, Mythology, and Legend.* New York: Funk and Wagnalls, pp. 653–4.

104. **Trehub, S. E. and E. G. Schellenberg** (1995) Music: its relevance to infants. *Annals Child Dev.* 11, 1–24.

105. **Tucker, N.** (1984) Lullabies. *History Today* 34, 40–6.

106. **Dissanayake, E.** (2000) Antecedents of the temporal arts in early mother–infant interaction. In N. L. Wallin, B. Merker, and S. Brown (eds) *The Origins of Music.* Cambridge, MA: MIT Press, pp. 389–410.

107. **Fernald, A.** (1992) Meaningful melodies in mothers' speech to infants. In H. Papoušek, U. Jurgens, and M. Papoušek (eds) *Nonverbal Vocal Communication: Comparative and Developmental Approaches.* Cambridge, UK: Cambridge University Press, pp. 262–82.

108. **Trainor, L. J., C. M. Austin, and R. N. Desjardins** (2000) Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychol. Sci.* 11, 188–95.

109. **Cooper, R. P. and R. N. Aslin** (1990) Preference for infant-directed speech in the first month after birth. *Child Dev.* 61, 1584–95.

110. **Fernald, A.** (1985) Four-month-old infants prefer to listen to motherese. *Infant Behav. Dev.* 8, 181–95.

111. **Werker, J. F. and P. J. McLeod** (1989) Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Canadian J. Psychol.* 43, 230–46.

112. **Trehub, S. E., A. M. Unyk, and L. J. Trainor** (1993) Adults identify infant-directed music across cultures. *Infant Behav. Dev.* 16, 193–211.

113. **Unyk, A. M., S. E. Trehub, L. J. Trainor, and E. G. Schellenberg** (1992) Lullabies and simplicity: a cross-cultural perspective. *Psychol. Music* 20, 15–28.

114. **Trainor, L. J., E. D. Clark, A. Huntley and B. Adams** (1997) The acoustic basis of infant preferences for infant-directed singing. *Infant Behav. Dev.* 20, 383–96.

115. **Trehub, S. E., A. M. Unyk, and L. J. Trainor** (1993) Maternal singing in cross-cultural perspective. *Infant Behav. Dev.* 16, 285–95.

116. **Trehub, S. E., A. M. Unyk, S. B. Kamenetsky, D. S. Hill, L. J. Trainor, J. L. Henderson and M. Saraza** (1997) Mothers' and fathers' singing to infants. *Developmental Psychol.* 33, 500–7.

117. **Bergeson, T. R. and S. E. Trehub** (1999) Mothers' singing to infants and preschool children. *Infant Behav. Dev.* 22, 51–64.

118. **O'Neill, C., L. J. Trainor, and S. E. Trehub** (2001) Infants' responsiveness to fathers' singing. *Music Percept.* 18, 409–25.

119. **Trehub, S. E., D. S. Hill, and S. B. Kamenetsky** (1997) Parents' sung performances for infants. *Canadian J. Exp. Psychol.* 51, 385–96.

120. **Trehub, S. E., A. M. Unyk, and J. L. Henderson** (1994) Children's songs to infant siblings: parallels with speech. *J. Child Lang.* 21, 735–44.

121. **Bergeson, T. R.** and **S. E. Trehub** (2002) Absolute pitch and tempo in mothers' songs to infants. *Psychol. Sci.* 13, 71–4.

122. **Fernald, A.** and **T. Simon** (1984) Expanded intonation contours in mothers' speech to newborns. *Developmental Psychol.* 20, 104–13.

123. **Halpern, A. R.** (1989) Memory for the absolute pitch of familiar songs. *Memory and Cognition* 17, 572–81.

124. **Levitin, D. J.** and **P. R. Cook** (1996) Memory for musical tempo: additional evidence that auditory memory is absolute. *Percept. Psychophys.* 58, 927–35.

125. **Eich, E.** and **D. Macaulay** (2000) Are real moods required to reveal mood-congruent and mood-dependent memory? *Psychol. Sci.* 11, 244–8.

126. **Masataka, N.** (1999) Preference for infant-directed singing in 2-day-old hearing infants of deaf parents. *Developmental Psychol.* 35, 1001–5.

127. **Trainor, L. J.** (1996) Infant preferences for infant-directed versus noninfant-directed playsongs and lullabies. *Infant Behav. Dev.* 19, 83–92.

128. **Trainor, L. J.** and **C. A. Zacharias** (1998) Infants prefer higher-pitched singing. *Infant Behav. Dev.* 21, 799–805.

129. **Nakata, T.** and **S. E. Trehub** (2000) *Maternal Speech and Singing to Infants.* Presented at the Society for Music Perception and Cognition. Toronto, ON.

130. **Trehub, S. E.** and **T. Nakata** (2001–2002) Emotion and music in infancy. *Musicae Scientiae,* Special Issue, 37–61.

131. **Shenfield, T., S. E. Trehub,** and **T. Nakata** (2002) *Salivary Cortisol Responses to Maternal Speech and Singing.* Presented at the International Conference on Infant Studies, Toronto, ON.

132. **Sloboda, J. A.** and **S. A. O'Neill** (2001) Emotions in everyday listening to music. In P. N. Juslin and J. A. Sloboda (eds) *Music and Emotion: Theory and Research. Series in Affective Science.* London, UK: Oxford University Press, pp. 415–29.

133. **Zillman, D.** (1988) Mood management: using entertainment to full advantage. In L. Donohew, H. E. Syhper, and E. T. Higgens (eds) *Communication, Social Cognition, and Affect.* Hillsdale, NJ: Erlbaum, pp. 147–71.

134. **Standley, J. M.** and **R. S. Moore** (1995) Therapeutic effects of music and mother's voice on premature infants. *Pediatric Nursing* 21, 509–12.

135. **Masuyama, E. E.** (1989) Desire and discontent in Japanese lullabies. *Western Folklore* 48, 144–8.

136. **Lomax, A.** (1968) *Folk Song Style and Culture.* Washington, DC: American Association for the Advancement of Science.

137. **Pantaleoni, H.** (1985) *On the Nature of Music.* Oneonta, NY: Welkin Books.

138. **Kogan, N.** (1997) Reflections on aesthetics and evolution. *Critical Rev.* 11, 193–210.

139. **Levin, R. N.** (1996) Song behaviour and reproductive strategies in a duetting wren, Thryothorus Nigricapillus: II playback experiments. *Animal Behavior* 52, 1107–17.

140. **Hooker, T.** and **B. I. Hooker** (1969) Duetting. In R. A. Hinde (ed.) *Bird Vocalizations.* Cambridge: Cambridge University Press, pp. 185–205.

141. **Tooby, J.** and **L. Cosmides** (2001) Does beauty build adapted minds? Toward an evolutionary theory of aesthetics, fiction and the arts. *SubStance* 30, 6–27.

142. **Dissanayake, E.** (2001) Becoming homo aestheticus: sources of aesthetic imagination in mother-infant interactions. *SubStance* 30, 85–103.

CHAPTER 2

# THE QUEST FOR UNIVERSALS IN TEMPORAL PROCESSING IN MUSIC

CAROLYN DRAKE AND DAISY BERTRAND

> One would ultimately hope to specify these cognitive principles or "universals" that underlie all musical listening, regardless of musical style or acculturation. To what extent is it learned, and to what extent is it due to an innate musical capacity or general cognitive capacity?
>
> Lehrdal & Jackendoff[1]

## Abstract

Music perception and performance rely heavily on temporal processing: for instance, each event must be situated in time in relation to surrounding events, and events must be grouped together in order to overcome memory constraints. The temporal structure of music varies considerably from one culture to another, and so it has often been supposed that the specific implementation of perceptual and cognitive temporal processes will differ as a function of an individual's cultural exposure and experience. In this paper we examine the alternative position that some temporal processes may be universal, in the sense that they function in a similar manner irrespective of an individual's cultural exposure and experience. We first review rhythm perception and production studies carried out with adult musicians, adult nonmusicians, children, and infants in order to identify temporal processes that appear to function in a similar fashion irrespective of age, acculturation, and musical training. This review leads to the identification of five temporal processes that we submit as candidates for the status of 'temporal universals'. For each process, we select the simplest and most representative experimental paradigm that has been used to date. This leads to a research proposal for future intercultural studies that could test the universal nature of these processes.

**Keywords:** Temporal processing; Rhythm perception; Segmentation; Grouping; Regularity; Duration ratios

## Towards intercultural research in temporal processing

We plead guilty to the charge of cultural egocentricity. In previous research, we have proceeded as if the musical environment the world over is identical to that in France, Belgium, England, and the United States (the countries in which we have carried out our experiments). The musical environment in these countries is relatively homogenous in the sense that it is dominated by a Western tonal tradition (although we acknowledge significant differences

even between these environments). From a rhythmic point of view, both our 'classical' and 'popular' musics are dominated by relatively simple rhythmic structures, organized around a regular beat, with binary or ternary multiplications and subdivisions of this beat.[1,2] By focusing on the auditory world around us, we have identified some of the psychological rhythmic processes that appear to function in our particular instance when listening to or producing 'our' type of music. Following the results of this developmental prospective, we have even made the claim that some of these processes function the same way whatever the musical environment in which a listener is immersed.[3] Despite this previous highly focused position, we are fully aware that our assumptions concerning the generalizability of these processes to other cultures are only that—assumptions. This paper is an attempt to make amends for our previous cultural egocentricity by suggesting how our psychological developmental prospective could be adapted to an intercultural one, into which existing knowledge from the field of ethnomusicology and future cross-cultural studies could be incorporated.

Our previous research has focused on how people hear musical and nonmusical rhythms in order to identify the underlying psychological processes that make the perception of music such a fulfilling and satisfying activity. Music is typical of all forms of sustained activity over time, in that it can be successfully perceived or performed only if the individual events from which it is composed are perceptually integrated into larger units spread over time. Indeed, music has been defined as the art of organizing events in time. As such, it provides an ideal opportunity to investigate the perceptual and cognitive temporal processes that make such activity possible.

In the past we have adopted the experimental principle of comparing performances on both perceptual and motor rhythmic tasks by people varying in levels of rhythmic skill, be it by age (as children get older there is a gradual increase in their exposure to, and experience with, the rhythmic structures around them) or musical training (the musical training common to our culture is particularly characterized by the development of explicit knowledge about musical structure). Such an experimental principle allows us to tease apart the processes that appear to be 'innate' or 'hard-wired' (functioning at birth, determined by genes, independent of environmental influence, and experience) and those that develop with maturation, acculturation (learning by immersion in the auditory world around us), or explicit training. The principle has been that if young infants, children, and nonmusician and musician adults display similar functioning modes on a particular task, then we conclude that this process may be 'innate' or at least 'functional' at an early age. Alternatively, if differences are observed between these populations, we then conclude that this type of functioning is acquired, either through acculturation or explicit learning.

We have, in the past avoided the word *innate* due to strong theoretical and philosophical connotations, referring rather to processes that may be universal—that is, that function in the same way in everyone. However, as has been pointed out to us, this wording is ambiguous. If we claim that a process is universal, we must demonstrate that it occurs the world over, irrespective of the cultural environment in which the individual lives and grew up. The present paper proposes such a research project. Whereas we have the know-how about both the psychological processes involved and appropriate paradigms to demonstrate the functioning of these processes, we are quite ignorant of the enormously exciting field of ethnomusicology, and our attempts at contact have so far been limited. Such a project cannot be

accomplished on our own, or even with just one or two colleagues, but requires numerous sites the world over with researchers collaborating within a network.

## Temporal processing is limited by memory space and processing time

If a computer programmer wanted to make a computer reproduce a musical rhythm, the program would probably record the precise duration (in milliseconds or computer clicks) of each interval in the sequence, put them in a lookup table, and then recall these values to produce the rhythm. The result would be that the reproduced rhythmic structure would be identical to that of the model. Is this a good model for the functioning of the human perceptual system? Probably not. Whereas human beings are able to reproduce musical rhythms so that they sound satisfactorily similar to the model, previous research suggests that our perceptual system does not function in the same way as the computer. The main difference in the way computers and humans function concerns memory limitations: computer memory is usually not a problem, while human memory is severely limited. As psychologists, our task is thus to describe the way in which our perceptual system analyses incoming temporal information and how it overcomes the main enemies: memory space and processing time.

Sound events usually do not occur in isolation, but rather are surrounded by other events, with each sound embedded within a sequence. Each event takes its existence from its relation with these surrounding events, rather than from its own specific characteristics. The task of situating each event in relation to surrounding events is quite simple when the sequence is short in duration and when it contains a limited number of events (i.e. three or four intervals). However, as the sequence becomes longer we very quickly run into problems of memory space and processing time. Imagine the number of intervals that would need to be stored and accessed when listening to a Beethoven sonata!

The idea is that all events would be stored in a memory buffer lasting several seconds (psychological present, probably corresponding to echoic or working memory), allowing the system to extract all relevant information (interval duration). One can imagine a temporal window gliding gradually through time, with new events arriving at one end and old events disappearing at the other due to decay. Thus only events occurring within a span of a few seconds would be accessible for processing at any one time, and only events occurring within this limited time window can be situated in relation to each other by the coding of relevant relational information. However, when the sequence becomes more complicated (longer and/or more events), the number of events that must be processed quickly goes beyond the limits of this buffer. Consequently, simple concatenation models that propose the maintenance in memory of the characteristics of each event are not able to account for the perception of long sequences because of problems of memory overload.

In this chapter, we present a set of five temporal processes that overcome these constraints, at least partially. These processes have been selected because previous research in the field of rhythmic perception suggests that they function in a similar fashion in all the populations examined so far. They are therefore candidates for the status of 'temporal universal'. This is not an exhaustive list, and other candidates could certainly be added as the

project advances. As mentioned above, this list emerges from previous developmental and comparative studies. In each case we select key paradigms that we have used successfully in the past to demonstrate the lack of difference between populations differing in age and musical experience. An essential characteristic of developmental research is that the tasks used must be conceptually simple; must not require reading, writing, or any other specific knowledge such as musical notation; must be short (usually not more than 20 min of experimenting time); and must be technically transportable (into schools). As such, the paradigms we have used for children should be easily adaptable for intercultural research, which faces similar constraints. Taken together, this set of experimental paradigms could be used as a basis for answering the question of the true 'universal' nature of these psychological processes.

## Candidate 1: segmentation and grouping

***We tend to group into perceptual units events that have similar physical characteristics or that occur close in time.***

One way to overcome processing limitations and to allow events to be processed together is to group the events into small perceptual units. These units result from a comparison process that compares incoming events with events that are already present in memory. If a new event is similar to those that are already present, it will be assimilated. If the new event differs too much (by its acoustical and/or temporal characteristics), the sequence will be segmented. This segmentation leads to the closure of one unit and the opening of the next. Elements grouped together will be processed together within a single perceptual unit and thus can be situated in relation to each other.

Several studies support these ideas. Listeners usually segment sequences as a function of the surface characteristics (timbre, pitch, intensity, event duration, pauses, etc.), following the principles laid down by the Gestalt psychologists: a change in any sound parameter leads to the perception of a break in the sequence and thus to the creation of groups separated by the changes (see Ref. 5 for a summary). For instance, the occurrence of a longer temporal gap or a major change in pitch leads to the segmentation of the sequence at that point, with the termination of one perceptual unit and the beginning of the next.

### Paradigm 1: online segmentation

Many variations on a simple segmentation paradigm have been used. Usually participants listen to a musical excerpt and are asked to indicate, by pressing a button or drawing a line on a musical score, whenever they hear a 'break' in the sequence, so that events that belong together go together. The sequences are constructed in such a way as to establish whether each type of segmentation is used, as well as to indicate the relative importance of each segmentation principle.

### Arguments in favour of universal status

*Comparisons across musical skill levels.* Grouping appears unaffected by musical training, as similar segmentation principles are observed for adult musicians and nonmusicians, although musicians are more systematic in their responses.[6–8]

*Infants.* Grouping also appears to function from an early age. For instance, studies using a gap detection paradigm suggest that six- to eight-month-old infants segment in a similar fashion to adults, at least for timbre and pitch, although a much greater pitch change was necessary for infants compared with adults.[9,10]

*Comparisons across ages.* Findings are even less clear in children. Among the few previous studies on rhythmic development, we often find suggestions that this sort of grouping process is functioning in relatively young (five- to seven-year-old) children. However, these studies employ extremely indirect methodologies, and the conclusions are tentative.[10–15] With this criticism in mind, a major research project has been undertaken to examine the principles governing segmentation in 4- to 12-year-old musician and nonmusician children.[16] The children's task was to listen to short musical sequences composed of nine tones and to indicate, during a second listening, the point at which they would break the sequence into two (group boundary). Each sequence contained two changes in either pitch, intensity, tone duration, or pause duration. The question was whether or not their perceived boundaries were induced by the physical changes in the sequence, as observed previously in adults. The results provide a clear answer to this question: segmentations were well above chance level for all four segmentation principles, even for the youngest children (four years). On average, more than three-quarters of the segmentations corresponded to the physical change; this effect was equally strong for the four indices (pitch, intensity, tone duration, and pause duration). There was a slight improvement with age and musical experience, probably due to improved task-related skills.

*Comparisons across cultures.* The online segmentation paradigm can easily be adapted for intercultural research, as the stimuli are nonculture-specific, and the task does not require any specific learned skill. If these basic processes of segmentation and grouping are unaffected by experience with a particular type of sequence, then the same types of segmentation should be observed in all cultures.

## Candidate 2: predisposition towards regularity

***Processing is better for regular than irregular sequences. We tend to hear as regular sequences that are not really regular.***

As mentioned above, rather than coding the precise duration of each interval, our perceptual system compares each newly arriving interval with preceding ones. If the new interval is similar in duration to preceding intervals (within an acceptable temporal window, the 'tolerance window'), it will be categorized as 'same'; if it is significantly longer or shorter than the preceding intervals (beyond the tolerance window), it will be categorized as 'different'. There may be an additional coding of 'longer' or 'shorter'. Thus, our system may code two or three categories of durations (same/different, or same/longer/shorter); but note that this is a relative, rather than absolute, coding system.

One consequence of this type of processing is that if a sequence is irregular (each interval has a different duration) but all the intervals remain within the tolerance window, then we will perceive this sequence as the succession of 'same' intervals and so perceive a regular sequence. Such a tolerance in our perceptual system is quite understandable when we examine the temporal microstructure of performed music: local lengthenings and shortenings of

more than 10 per cent are quite common[17,18] and are not necessarily picked out by listeners as irregularities.[19]

### Paradigm 2: tempo discrimination for regular and irregular sequences

This predisposition towards regularity has been investigated using a tempo discrimination paradigm. Listeners hear two sequences that differ slightly in tempo, and they must say which is the fastest. The degree of regularity of the sequences is varied from completely regular (isochronous) to extremely irregular (based on the standard deviation of the interval durations).[20] Results show that up to a certain degree of irregularity, irregular sequences are processed as well as the regular sequences: they are 'assimilated' towards regularity. However, above a certain degree of irregularity, discrimination performance drops considerably. This cut-off point corresponds to the tolerance window mentioned above.

### Arguments in favour of universal status

*Comparisons across musical skill levels.* The tolerance window appears to function in a similar fashion for both musicians and nonmusicians, with slightly irregular sequences being assimilated to regular sequences, although temporal thresholds are lower in musicians than nonmusicians.[20]

*Comparison across ages.* To our knowledge, nothing is known about children's ability to process irregular sequences. We do know, however, that from the age of four years, children are able to detect a small change in tempo of an isochronous sequence.[21] Another way of demonstrating the importance of temporal regularity in processing sequences involves rhythm reproduction tasks. For instance, when five- and seven-year-old children reproduced both regular and irregular rhythms, performance was much better for the regular ones.[22]

*Infants.* We also know little about infants' ability to process irregular sequences. We do know, however, that the capacity to detect a small change in tempo of an isochronous sequence is already functional at two months in infants.[23] They are able to habituate to a particular tempo, and there is a reaction to novelty if the tempo changes.

*Comparison across cultures.* This tempo discrimination task for regular and irregular sequences is easily adaptable to people from other cultures. If this process is universal, we should observe the same low tempo discrimination thresholds for regular and slightly irregular sequences, with considerably higher thresholds for very irregular sequences.

## Candidate 3: active search for regularity

***We spontaneously search for temporal regularities and organize events around this perceived regularity.***

Coding events in terms of temporal regularity is thus an economical processing principle, and it has implications. If an incoming sequence can be coded in such a fashion, the needed processing resources are reduced, thus making it easier to process such a sequence. Indeed, we can say that the perceptual system exploits this predisposition, by actively 'looking for' temporal regularities in all types of sequences. We therefore suggest that when listening to a piece of music, we are predisposed to finding a regular pulse, that which is

emphasised by tapping our foot in time with the music (*tactus* in musical terms). Once this underlying pulse has been identified, it is used as an organizational framework around which other events are situated.

### Paradigm 3: synchronization with musical sequences

A simple way of investigating this process is to ask people to listen to a musical sequence, and then to tap in time with the music in a regular fashion at the rate that they think 'goes best' with the music. Synchronization is considered to be accomplished if successive taps coincide with tones within the music (at a particular hierarchical level) within a 10 per cent window. This is quite a strict criteria when you take into consideration the considerable temporal variations observed in performed music. In order to accomplish this task, listeners must abstract an underlying regularity, even if the performance variations tend to mask 'pure' temporal regularity.

### Arguments in favour of universal status

*Comparison across musical skill levels.* Synchronization success rates for both musicians and nonmusicians are very high (more than 90 per cent) when they are asked to tap in synchrony with music, even if the music contains many temporal and other performance microvariations.[24]

*Comparison across ages.* All 20 four-year-olds (as well as the older children aged 6–10 years) examined in our study were able to successfully synchronize with the beginning of Ravel's *Boléro*.[21]

*Infants.* Infants are able to adapt their spontaneous sucking rate to the rate of an auditory sequence,[25] at least under certain circumstances.

*Comparison across cultures.* In order to adapt this paradigm to intercultural research, musical excerpts must be selected to include both music that respects the temporal structure with which the participants are familiar (their own musical idiom) and music that does not.

## Candidate 4: temporal zone of optimal processing

***We process information best if it arrives at an intermediate rate.***

A fourth processing principle concerns the rate of temporal sequences. People spontaneously 'listen for' important events occurring at equally spaced moments in time, and the rate at which they 'search for' important information is specific to each individual.[26,27] Thus, the search for temporal regularities described above occurs at a particular rate. A zone of optimal processing has been demonstrated with numerous paradigms and types of sequences. The results are concordant: sensitivity to change is highest if events occur about every 600 ms, with a range stretching between about 300 and 800 ms interonset interval (IOI).

### Paradigm 4: irregularity detection in complex sequences

How can we demonstrate that people focus spontaneously on events occurring at intermediate rates? A first method is simply to compare tempo discrimination thresholds of different

rates,[20] as mentioned above. However, a more ecologically valid paradigm is currently being developed that would be ideal for the present requirements. When participants listen to complex sequences composed of two cooccurring isochronous subsequences, they should focus on the subsequence closest in rate to the optimal processing zone (300–800 ms IOI). We therefore introduce a temporal irregularity into one of the subsequences, which listeners should be able to detect only if they are focusing on that particular subsequence. The overall rate of the complex sequence is varied from very fast to very slow. Our results indicate that, as expected, detection was better for the slowest subsequence when the complex sequence was fast, and better for the fastest subsequence when the complex sequence was slow.[28] This promising paradigm has been applied only to nonmusician adults.

### Arguments in favour of universal status

*Comparison across musical skill levels.* Although musicians demonstrate a wider range of optimal tempi than nonmusicians, this optimal zone is still centred around the same value of 600 ms IOI.[20]

*Comparison across ages.* Children between the ages of 4 and 10 years also demonstrate the same zone of optimal tempi, although the range increases with age.[21]

*Infants.* Two-month-old infants demonstrate a reaction to novelty only for sequences at 600 ms IOI, demonstrating the same optimal zone from a very young age.[23]

*Comparison across cultures.* This paradigm can be easily adapted to people of other cultures.

## Candidate 5: predisposition towards simple duration ratios

***We tend to hear a time interval as twice as long or short as previous intervals.***

A fifth principle concerns the perceptual status of the 'longer' and 'shorter' durations: longer time intervals tend to be perceived and produced twice as long as the 'same' intervals, and shorter intervals tend to be perceived and produced twice as short as the same intervals.[29–31] The implication is that a categorization process is involved here, with clear-cut passages from one category to another.[30,32] Such an organization principle results in the dominance of binary, rather than ternary or more complex ratios between the three categories of intervals.

### Paradigm 5: rhythm reproduction

One consequence of this process is that people are better able to reproduce rhythms containing only 1 : 2 ratios than rhythms containing 1 : 3 ratios[33] or even more complicated ratios.[34] Also, when people reproduce complex musical rhythms, the duration of some intervals undergoes a distortion towards one of the categories—that is, the rhythm is simplified so that the produced intervals respect a 1 : 2 ratio.[29]

### Arguments in favour of universal status

*Comparison across musical skill levels.* The same pattern of perceptual and motor distortions are observed in musicians and nonmusicians.[34]

*Comparison across ages.* From the age of 5 years, children are able to reproduce short rhythms based on binary (1:2), but not ternary (1:3) ratios.[33] When five- and seven-year-old children reproduce complex rhythms, their reproductions undergo simplification towards simple ratios.[22] Their incorrect reproductions contain only two durations, in a 2:1 ratio.[22]

*Infants.* To our knowledge, no data exists concerning the functioning of this duration categorization process in infants.

*Comparison across cultures.* A rhythm reproduction task can be adapted for intercultural research. Participants can be asked to reproduce rhythms varying in the number of different durations they contain. Reproductions should be best when they contain only two different durations in a ratio of 1:2. Ratios of 1:3 should be harder, and more complex ratios (which are perceived as irregular—see Candidate 2) even harder. The reproductions of complex rhythms should demonstrate simplification towards simple ratios.

## Conclusion

We have thus proposed a series of five experimental paradigms, designed and tested to demonstrate the functioning of five temporal processes. If our hypotheses are correct concerning the universal nature of these processes, then whatever the culture or origin of the people being tested, we should observe the same results. This list of potential candidates for the status of temporal universals is probably far from complete, but it provides a starting point. Further suggestions are welcome.

## References

1. **Lerdahl, F.** and **R. Jackendoff** (1983) *A Generative Theory of Tonal Music.* Cambridge, MA: MIT Press.
2. **Cooper, G.** and **L. G. Meyer** (1960) *The Rhythmic Structure of Music.* Chicago: University of Chicago Press.
3. **Yeston, M.** (1976) *The Stratification of Musical Rhythm.* New Haven, CT: Yale University Press.
4. **Drake, C.** (1998) Psychological processes involved in the temporal organization of complex auditory sequences: universal and acquired processes. *Music Percept.* 16, 11–26.
5. **Handel, S.** (1989) *Listening. An Introduction to the Perception of Auditory Events.* Cambridge, MA: MIT Press.
6. **Deliège, I.** (1987) Grouping conditions in listening to music: an approach to Lerdahl and Jackendoff's grouping preference rules. *Music Percept.* 4, 325–60.
7. **Fitzgibbons, P. J., A. Pollatsek,** and **I. B. Thomas** (1974) Detection of temporal gaps within and between perceptual tonal groups. *Percept. Psychophys.* 16, 522–8.
8. **Peretz, I.** and **J. Morais** (1989) Music and modularity. *Contemp. Music Rev.* 4, 279–93.
9. **Krumhansl, C. L.** and **P. W. Jusczyk** (1990) Infants perception of phrase structure in music. *Psychol. Sci.* 1, 70–3.

10. Thorpe, L. A. and S. E. Trehub (1989) Duration illusion and auditory grouping in infancy. *Dev. Psychol.* 25, 122–7.
11. Drake, C. (1993) Influence of age and experience on timing and intensity variations in the reproduction of short musical rhythms. *Psychol. Belg.* 33, 217–28.
12. Drake, C. (1993) Perceptual and performed accents in musical sequences. *Bull. Psychon. Soc.* 31, 107–10.
13. Drake, C., J. Dowling, and C. Palmer (1991) Accent structures in the reproduction of simple tunes by children and adult pianists. *Music Percept.* 8, 313–32.
14. Gérard, C. and C. Drake (1990) The inability of young children to reproduce intensity differences in musical rhythms. *Percept. Psychophys.* 48, 91–101.
15. Bamberger, J. (1980) Cognitive structuring in the apprehension and description of simple rhythms. *Arch. Psychol.* 48, 177–99.
16. Bertrand, D. (1999) Groupement rythmique et representation mentale de mélodies chez l'enfant. Ph.D. Thesis. Belgium: Liège University.
17. Penel, A. and C. Drake (1997) Perceptual and cognitive sources of timing variations in music performance: a psychological segmentation model. *Psychol. Res.* 61, 12–32.
18. Repp, B. H. (1992) Diversity and commonality in music performance: an analysis of timing microstructure in Schumann's Träumerei. *J. Acoust. Soc. Am.* 92, 2546–68.
19. Repp, B. H. (1992) Probing the cognitive representaion of musical time: structural constraints on the perception of timing perturbations. *Cognition* 44, 241–81.
20. Drake, C. and M. C. Botte (1993) Tempo sensitivity in auditory sequences: evidence for a multiple-look model. *Percept. Psychophys.* 54, 277–86.
21. Drake, C., M. R. Jones, and C. Baruch (2000) The development of rhythmic attending in auditory sequences: attunement, reference period, focal attending. *Cognition* 77, 251–88.
22. Drake, C. and C. Gérard (1989) A psychological pulse train: how young children use this cognitive framework to structure simple rhythms. *Psychol. Res.* 51, 16–22.
23. Baruch, C. and C. Drake (1997) Tempo discrimination in infants. *Infant Behav. Dev.* 20, 573–7.
24. Drake, C., A. Penel, and E. Bigand (2000) Tapping in time with mechanically and expressively performed music. *Music Percept.* 18(1), 1–24.
25. Pouthas, V. (1995) The development of the perception of time and temporal regulation of action in infants and children. In I. Deliège and J. A. Sloboda (Eds) *Musical Beginnings: The Origins and Development of Musical Competence.* New York: Oxford University Press, pp. 115–41.
26. Jones, M. R. (1976) Time, our last dimension: toward a new theory of perception, attention, and memory. *Psychol. Rev.* 83, 323–55.
27. Jones, M. R. and M. Boltz (1989) Dynamic attending and responses to time. *Psychol. Rev.* 96, 459–91.
28. Brochard, R. (1997) Role of attention in the perceptual organisation of complex auditory sequences. Ph.D. Thesis, University of Paris.
29. Fraisse, P. (1956) Les structures rythmiques. Publications Universitaires de Louvain. Louvain.
30. Clarke, E. F. (1987) Categorical rhythm perception: an ecological perspective. In A. Gabrielsson (ed.) *Action and Perception in Rhythm and Musica.* Stockholm: Royal Swedish Academy of Music, pp. 19–34.
31. Parncutt, R. (1994) A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Percept.* 11, 409–64.

32. **Schulze, H. H.** (1989) Categorical perception of rhythmic patterns. *Psychol. Res.* 51, 10–15.

33. **Drake,** C. (1993) Reproduction of musical rhythms by children, adult musicians, and adult nonmusicians. *Percept. Psychophys.* 53, 25–33.

34. **Sternberg, S., R. L. Knoll,** and **P. Zukofsky** (1982) Timing by skilled musicians. In **D. Deutsch** (ed.) *The Psychology of Music.* New York: Academic Press, pp. 182–237.

CHAPTER 3

# MECHANISMS OF MUSICAL MEMORY IN INFANCY

JENNY R. SAFFRAN

## Abstract

How do infants learn about their auditory world, and what do they remember? In this chapter, we review recent findings suggesting that as previously observed in the language domain, infants possess powerful memory abilities that allow them to represent myriad aspects of their musical environments. These memories provide a corpus of musical experiences from which infants can begin to acquire the structures that characterize their native musical systems.

## Introduction

From the perspective of an infant just beginning to make sense of the world, music is an extremely complex and attractive stimulus. Not only is music itself of great interest to the youngest listeners, but it is also the case that the musical aspects of language—the pitch and rhythmic patterns that constitute native language prosodic structure—are the most likely to capture infants' attention, and are the first components of language that infants learn (e.g. Refs 1–3). Moreover, infant-directed music may be more speech-like, with more communicative value, than adult-directed music, suggesting that the input to infant listeners may be particularly similar in these two domains (e.g. Ref. 4 and Chapter 1). Linguistic and musical input both present a vast array of information to listeners—including many levels of structure that are simultaneously available—that must be acquired by infants on the road to becoming native listeners. How does this process unfold, and what are the mechanisms that allow infants to learn from their musical experiences?

While we will focus primarily on music in this chapter, we will use the literature on infant speech and language perception to suggest avenues of overlap and difference, and as a source of potentially fruitful areas for future study. Before doing so, however, it is necessary to step back and consider the hypothesized relationships between these two systems. In particular, the connection between linguistic and musical processes in infancy is currently unknown. Language and music are clearly not isomorphic in the adult brain, and are dissociable in many important ways (for recent reviews, see Chapters 12 and 13). Despite these differences in the mature state, it is possible that some of the same mechanisms are engaged in learning in both domains. It is important to note that any observed modularity in the adult brain does not necessarily mean that the learning mechanisms which initially supported learning during infancy are themselves modular. While there is some evidence that some specifically musical processes are subject to congenital impairments,[8] it is also the case that some learning mechanisms have been shown to

operate equivalently for both speech-like and music-like stimuli in both infancy and adulthood.[9] Similarly, related patterns of adult brain electrical activity occur in response to some aspects of music and language, but not to others (for review, see Chapter 18). The level of plasticity evident for both types of auditory learning, particularly during childhood, suggests that at least some of the underlying processes for music and language may overlap developmentally.

Consider a specific example of the possible relationship between representations of auditory information in memory with respect to speech and music. The relevance of pitch for music perception is obvious; what is less obvious but also important is the role played by pitch in language learning. The most easily observed use of linguistic pitch lies in the melodic contours which signal sentence meaning, such as the rising contour that is indicative of question prosody. There are also many languages which use pitch and/or pitch contour contrastively with respect to word meaning; these tonal languages, such as Mandarin, Thai, and Vietnamese, are acquired readily by infants and young children. For this learning to occur, all infants—regardless of their native language-to-be—must possess the capacity to attend to pitch cues as a possible component of lexical items. At the same time, many aspects of pitch in linguistic stimuli must be ignored by infants engaged in learning their native language. As in music perception, the absolute pitches with which words are spoken are not informative as to their meaning. Even deeper relationships between the use of pitch in speech and music processing have been hypothesized; for example, processing speech may lead the auditory system to acquire the specific pitch relationships between harmonics in speech sounds, which become the intervals in music.[11] It is thus likely that the learning mechanisms that subserve the acquisition both speech and music are related to one another in some fashion, despite the eventual deep differences in the knowledge acquired; however, the nature of this relationship is currently unknown.

## Infant music perception and possible relationships with speech perception

The task of learning about how music works appears to be partially solved for the infant quite early in life, either via biological predispositions or through extremely rapid learning. Infants show early preferences for consonant musical intervals such as octaves and perfect fifths, are highly attuned to rhythm, can represent both relational and absolute pitch information under different circumstances, and are sensitive to cues correlated with musical phrase boundaries (for a recent review, see Chapter 1). These early emerging capacities provide support for the infant's initial representations of musical structure and events.

Despite this early head-start, there remains a tremendous amount of information for infants to learn. Just as in language, where learners must acquire knowledge of the systematic structure in their native language (e.g. phonological and grammatical generalizations) as well as specific knowledge about individual experiences (e.g. individual phonemes, lexical items, and voices), infants must learn the structure of the musical system present in their environment (e.g. tonality, conventional structures) as well as specific pieces of music. Infants must therefore possess abilities which allow them to represent musical experiences, and to learn from those experiences.

Of particular interest is the question of which aspects of music are represented in infants' memories. Any given musical experience contains information which is relevant for discerning the deep structure of the piece, including its melodic structure, harmonic structure, and rhythm. There is also myriad information which is less important musically, such as the specific key of the piece, the exact tempo of the performance, and the instruments upon which it is played. Despite the seeming irrelevance of these highly specific aspects of particular musical experiences, adult listeners' representations of highly familiar pieces include information about absolute pitch,[12] absolute tempo,[13] and timbre.[14] It appears that at least for adult listeners, 'memory representations for complex auditory stimuli contain information about the absolute properties of the stimuli, in addition to more meaningful information abstracted from the relations between stimulus components'[14] (p. 646). This model of auditory memory also captures what is currently known about adult representations of linguistic auditory events, which include seemingly irrelevant specific information such as talker-specific cues as well as linguistically meaningful information; for example, adult listeners appear to include speaker voice and rate (though not amplitude) in their representations of spoken words.[15]

Which aspects of auditory experiences are represented by infant listeners in memory? Research focused on infant speech perception suggests that infants as young as 7.5 months of age can reliably remember new words after just a few exposures embedded in fluent speech, and that these representations are specific enough that infants exposed to 'cup' do not incorrectly recognize 'tup' (e.g. Ref. 16). Moreover, infants at this age can remember words heard multiple times in stories over multi-week delays.[17] Even younger infants can recognize their names and other words that are heard frequently in their environments (e.g. Refs 18 and 19). The literature also provides some indications that like adults, infants represent multiple levels of auditory stimuli, including both those that are linguistically relevant and those that are not. For example, 7.5-month-old infants include talker-specific cues in their representations of spoken words.[20] However, infants are able to ignore talker-specific properties under other circumstances—in particular, infants readily exhibit vowel normalization, categorizing individual exemplars according to vowel identity despite differences in speaker sex (e.g. Refs 21 and 22). Infants thus appear to process linguistic auditory events at multiple levels of detail simultaneously.

These findings from the domain of speech suggest potential avenues of study for infant musical learning and memory. Is infant memory for musical experiences as powerful and as detailed as memory for language? It is possible that infants' representations of music are even more specific than their representations of speech; mothers' songs to infants are remarkably consistent over time with respect to their pitch and tempo, providing a highly reliable set of musical cues to infant listeners, whereas their infant-directed-speech is less consistent in pitch and tempo.[23]

To learn about music, infants must acquire knowledge both about musical structure in general and about specific pieces of music. The literature suggests that the former process—learning one's native musical system—is quite extended developmentally. Indeed, while there are some indications that infants acquire rudimentary knowledge about how tonal systems work during their first year of life (e.g. Ref. 24), most of this knowledge is not fully acquired until childhood (e.g. Refs 25–27). For example, Schellenberg and Trehub[28]

demonstrated that the degree to which novel melodies adhere to Western tonal conventions affected the performance of 5-year-old children, but not 9-month-old infants. These findings diverge from the literature on infant speech perception, which suggests that infants have acquired the bulk of their native language's sound structure by the time they are a year of age (for extensive review, see Ref. 29). It is possible that the acquisition of Western tonal structure is more akin to the acquisition of syntax and morphology than to the acquisition of phonology, as the grammatical aspects of language are acquired slightly later (primarily in the second and third postnatal year). Alternatively, the fact that infants presumably receive more exposure to speech than to music (although this has never been quantified) may account for the relative rapidity with which knowledge of the native language emerges relative to knowledge of the native musical system.

Infants' memory capacity for specific musical experiences—individual pieces of music present in the infant's environment—has received far less attention in the literature. In the remainder of this chapter, we will review recent evidence suggesting that infants possess the capacity to remember individual pieces of music over multiweek delays. These accumulated memory representations presumably provide a corpus of musical knowledge from which infants can learn about the structure of their native musical system, analogous to the specific linguistic experiences from which infants begin to glean phonological and grammatical generalizations. As described above with reference to adult musical memory and infant linguistic memory, infants appear to attend to many levels of musical structure which are manifest in their memory representations, suggesting interesting parallels in memory both across age (infants vs adults) and domain (music vs language).

## Long-term memory for music in infancy

To what extent do young infants maintain specific musical experiences in long-term memory? Recent findings in the language domain suggest that 8-month-old infants can remember words from stories over time delays of two weeks.[17] Because music, along with language, is among the most complex auditory learning tasks facing infants, it is possible that infant memory for musical stimuli is equally powerful. Little research, however, has addressed the question of whether infants remember particular musical pieces heard during passive listening experiences.

To ask whether infants possess long-term memory for musical materials, we adapted the experimental procedures used by Jusczyk and Hohne[17] to assess the maintenance of linguistic materials over multiweek delays.[30] We exposed a group of 7-month-old infants to a 10 min. recording of two Mozart piano sonata movements: KV 281 in B-flat major, Andante, and KV 282 in E-flat major, Adagio. Each infant listened to the music once a day, at home, for 14 consecutive days. Following a two week retention interval, during which the infants did not hear these pieces, we assessed the infants' long-term memory for the Mozart sonatas in the laboratory. To do so, we measured listening preferences using the head-turn preference procedure (e.g. Ref. 31), a standard computer-automated method used in studies of infant auditory perception.

We hypothesized that if infants remembered the music with which they were familiarized at home, they would show recognition by listening longer to passages of familiar music

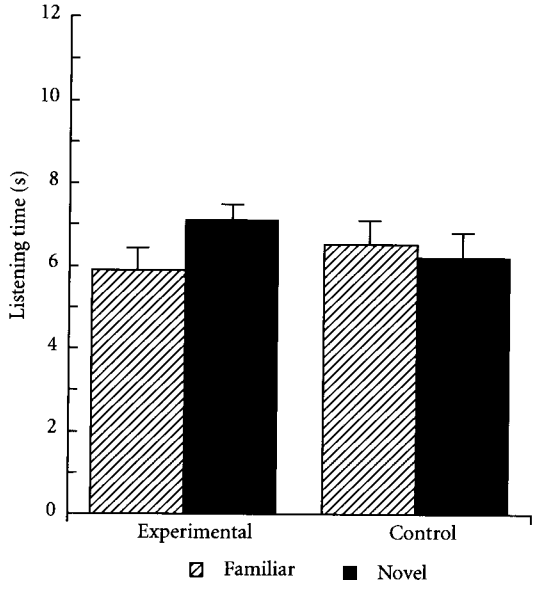



**Figure 3.1** Experiment 1: listening times for familiar and novel test passages for infants from the experimental (prior exposure) and control (no prior exposure) groups.[30]

than to novel passages (as observed in Ref. 17 research on infant linguistic long-term memory). The test consisted of two familiar and two novel passages, all 20 s in duration. The familiar test passages were drawn from the middle of each of the two familiarized sonata movements. The novel test stimuli were drawn from two other Mozart sonatas (taken from the same CD recording and also composed during late 1774-early 1775: KV 280 in F-major, Adagio, and KV 283 in G-major, Andante). To ensure that arbitrary properties of the test stimuli could not account for the infants' performance, we also tested a control group of 8-month-old infants who received no prior exposure to any of these pieces.

Mean listening times for the familiar and novel musical passages were calculated for each infant (see Figure 3.1). Consistent with the hypothesis that infants maintain musical experiences in long-term memory, there was a significant difference in listening times for the familiar and novel passages for infants in the experimental group. However, the direction of the preference was contrary to our hypothesis; the infants preferred the novel passages over the familiar passages. To ensure that this pattern of results was not due to biased test materials, we also examined the performance of the infants in the control group. These infants did not show a significant preference for either type of test passage, suggesting that the preferences which emerged for the experimental group reflected long-term memory for the familiar passages.

Why did the infants prefer the novel passages in the first experiment? This direction of preference is the opposite of the results from Jusczyk and Hohne's linguistic study,[17] in which infants preferred familiar stimuli after a two week delay. There are (at least) two possible explanations: (1) Mastery: If infants mastered the familiarized music, they might now

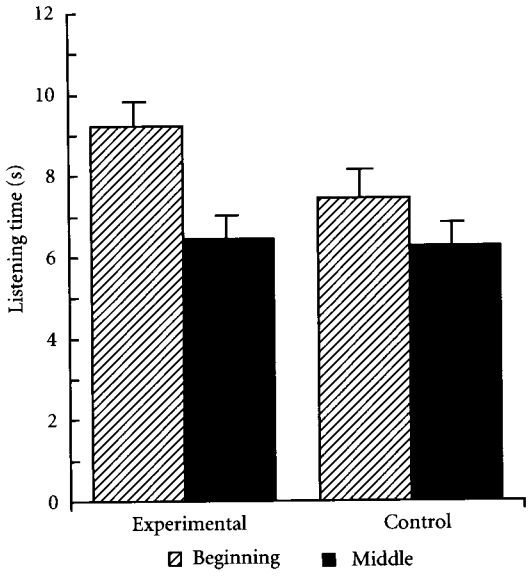



**Figure 3.2** Experiment 2: listening times for passages from the beginnings and middles of the familiar music for infants from the experimental (prior exposure) and control (no prior exposure) groups.[30]

prefer to hear something new, despite the two week retention period during which they did not hear the familiarized music. (2) Musical context: The familiar test passages were taken from the middles of the familiar sonata movements. Thus, when they were played during testing, they were removed from the musical context within which they were originally learned. Because infants typically prefer to listen to stimuli that are maximally natural, in both the linguistic and musical domains (e.g. Refs 32 and 33), passages removed from context may have been dispreferred.

The second experiment was designed to test these two hypotheses. A second group of 7-month-old infants was familiarized with the same two sonata movements, following procedures identical to Experiment 1. Following a two week retention interval, the infants were tested in the laboratory on two different types of *familiar* passages: passages taken from the middles of the sonatas (which were the familiar passages used in Experiment 1), and the beginning passages from the two familiarized sonatas. If the results of Experiment 1 reflected mastery as evidenced by relative disinterest in the familiar music, then the infants should now prefer the *middle* passages, as the beginnings are presumably more memorable than the middle passages. However, on the musical context hypothesis, infants should prefer the familiar passages taken from the *beginnings* of the sonatas. Even in the context of the full sonata movements, the beginning passages were always preceded by silence. Thus, unlike the passages taken from the middles of the sonatas, the beginning test passages were not removed from their musical context when played in isolation.

As in Experiment 1, there was a significant difference in listening times (see Figure 3.2). The direction of the preference was consistent with the musical context hypothesis: the

infants preferred the passages from the beginnings rather than the middles of the familiarized sonatas. Again, we tested a separate control group to ensure that biased test passages could not account for the experimental group's performance; these infants did not show a significant preference for either type of test passage. Along with Experiment 1, the results of Experiment 2 strongly suggest that infants do remember the music that they hear; the existence of a preference for passages drawn from the beginnings, rather than the middles, of the familiar music indicates that the infants remembered the pieces last heard at home two weeks before. Moreover, the infants preferred music that was not taken out of context. The control infants showed no such bias; even out of context, these passages sounded coherent to listeners who were not familiar with the pieces from which they were drawn.

These results suggest that infants are sophisticated musical listeners. Their representations of music in long-term memory are not undifferentiated sequences of notes. Instead, passages are linked together to form coherent musical events. Other recent studies investigating infant long-term memory for music similarly suggest that infants' representations are quite nuanced. For example, infants can represent far more complex pieces of music (such as Ravel) in long-term memory.[34] Moreover, the content of infants' memories includes some extremely specific aspects of musical performances. For example, 10-month-olds represent acoustic patterns specific to the particular performances with which they were previously familiarized.[35] Six-month-old infants remember the specific tempo and timbre of music with which they are familiarized, failing to recognize pieces when they are played at new tempos or with new timbres.[36] It thus appears that infant representations are extremely specific, without the flexibility to generalize to include changes in tempo or timbre. This level of specificity must change with age, as either a function of experience or development, else listeners would not be able to recognize familiar music played on different instruments or at different rates. Indeed, preschool children can recognize familiar tunes across many different types of transformations, suggesting that children are in the process of developing more flexible musical representations.[37]

Interestingly, infants can recognize previously familiarized melodies when they are played in a new key,[36] suggesting that for highly familiar pieces, infants represent the relative pitches of melodies. Such interval-level information is the appropriate level of specificity for generalizing memories from one experience to another, and is the preferential level of processing for most adult listeners. These results reinforce the suggestion that task demands determine the level of processing of pitch cues in musical materials: infants represent pitch contour and/or relative pitches more readily given familiar pieces, or given familiarization with repeating brief tunes (for review, see Ref. 38), but appear to rely on absolute pitches given unsegmented stretches of previously unfamiliar tone sequences.[24,39] Infants have access to multiple levels of pitch representations, much like adults (e.g. Refs 40 and 12); the structure of the task and the familiarity of the stimulus domain may influence which levels of representation are prioritized. For example, when absolute pitch cues are rendered uninformative in a task which previously elicited the use of absolute pitch information, infants shift to rely on relative pitch cues (Saffran, Reeck, Niebuhr, and Wilson, in prep.). These results mirror data from an avian species—starlings—who can switch from relying on absolute pitch cues to using relative pitch cues when necessitated by the structure of the task.[41]

## Conclusion

We thus see a surprising level of similarity in infant memory representations for music and for language. In both cases, infants readily represent the input at multiple levels of analysis, from highly specific surface-level cues to the structural information that eventually conveys meaning. Moreover, these processes of learning and memory proceed without instruction or any reinforcement other than the pleasure of listening and the instinct to learn. These findings lend further support to an emerging picture of infants as remarkably adept at implicitly learning and remembering the structured information that characterizes the environment in which they develop. From the infant's perspective, music and language may not be nearly as different as they are for the adult listener, at least when considering what is to be gleaned from a listening experience. Just how distinct these processes are remains to be seen. Might the eventual differences in how the adult brain represents music and language be due to predetermined differences in the brain, or to differences in the structure of the input drawn from these two domains, or both? Regardless, it is evident that even the youngest listeners are already accomplished musicians in their own right.

## Acknowledgements

The preparation of this chapter was made possible by a grant from the International Foundation for Music Research. Thanks to the many students and research assistants who have collaborated on the conduct of this research, and to Erik Thiessen for helpful comments on a previous draft.

## References

1. **Cooper, R. P.** and **R. N. Aslin** (1990) Preference for infant-directed speech in the first month after birth. *Child Dev.* 61, 1584–95.
2. **Fernald, A.** (1989) Intonation and communicative intent in mothers' speech to infants: is the melody the message? *Child Dev.* 60, 1497–510.
3. **Mehler, J., P. W. Jusczyk, G. Lambertz, N. Halsted, J. Bertoncini,** and **C. Amiel-Tison** (1988) A precursor of language acquisition in young infants. *Cognit.* 29, 144–78.
4. **Trehub, S. E.** and **L. J. Trainor** (1998) Singing to infants: lullabies and play songs. In C. Rovee-Collier, L. Lipsitt, and H. Hayne (eds) *Advances in Infancy Research.* Stamford, CT: Ablex, pp. 43–77.
5. **Ayotte, J., I. Peretz,** and **K. Hyde** (2002) Congenital amusia: a group study of adults afflicted with a music-specific disorder. *Brain,* 125, 238–51.
6. **Saffran, J. R., E. K. Johnson, R. N. Aslin,** and **E. L. Newport** (1999) Statistical learning of tone sequences by human infants and adults. *Cognit.* 70, 27–52.
7. **Terhardt, E.** (1974) Pitch, consonance, and harmony. *J. Acoust. Soc. Am.* 55, 1061–9.
8. **Levitin, D.** (1994) Absolute memory for musical pitch: evidence from the production of learned melodies. *Percept. Psychophy.* 56, 414–23.
9. **Levitin, D.** and **P. R. Cook** (1996) Memory for musical tempo: additional evidence that auditory memory is absolute. *Percept. Psychophy.* 58, 927–35.

10. Schellenberg, E. G., P. Iverson, and M. C. McKinnon (1999) Name that tune: identifying popular recordings from brief excerpts. *Psychonomic Bulletin Rev.* 6, 641–6.
11. Bradlow, A. R., L. C. Nygaard, and D. B. Pisoni (1999) Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Percept. Psychophy.* 61, 206–19.
12. Jusczyk, P. W. and R. N. Aslin (1995) Infants' detection of the sound patterns of words in fluent speech. *Cognit. Psychol.* 29, 1–23.
13. Jusczyk, P. W. and E. A. Hohne (1997) Infants' memory for spoken words. *Sci.* 277, 1984–86.
14. Mandel D. R., P. W. Jusczyk, and D. B. Pisoni (1995) Infants' recognition of the sound patterns of their own names. *Psychological Sci.* 6, 314–17.
15. Tincoff, R. and P. W. Jusczyk (1999) Some beginnings of word comprehension in 6-month-olds. *Psychological Sci.* 10, 172–5.
16. Houston, D. M. and P. W. Jusczyk (2000). The role of talker-specific information in word segmentation by infants. *J. Exp. Psychol.: Hum. Percept. Perform.* 26, 1570–82.
17. Kuhl, P. K. (1979) Speech perception in early infancy: perceptual constancy for spectrally dissimilar vowel categories. *J. Acoust. Soc. Am.* 66, 1668–79.
18. Kuhl, P. K. (1983) Perception of auditory equivalence classes for speech in early infancy. *Infant Behav. Dev.* 6, 263–85.
19. Bergeson, T. R. and S. E. Trehub (2002) Absolute pitch and tempo in mothers' songs to infants. *Psychological Sci.* 13, 72–5.
20. Saffran, J. R. (2003) Absolute pitch in infancy and adulthood: the role of tonal structure. *Dev. Sci.* 6, 37–49.
21. Krumhansl, C. L. and F. C. Keil (1982) Acquisition of the hierarchy of tonal functions in music. *Mem. Cognit.* 10, 243–51.
22. Lynch, M. P. and R. E. Eilers (1992) A study of perceptual development for musical tuning. *Percept. Psychophy.* 52, 599–608.
23. Trainor, L. J. and S. E. Trehub (1992) A comparison of infants' and adults' sensitivity to Western musical structure. *J. Exp. Psychol.: Hum. Percept. Perform.* 18, 394–402.
24. Schellenberg, E. G. and S. E. Trehub (1999) Culture-general and culture-specific factors in the discrimination of melodies. *J. Exp. Child Psychol.* 74, 107–27.
25. Jusczyk, P. W. (1997) *The Discovery of Spoken Language.* Cambridge: MIT Press.
26. Saffran, J. R., M. M. Loman, and R. R. W. Robertson (2000) Infant memory for musical experiences. *Cognit.* 77, 15–23.
27. Kemler Nelson, D. G., P. W. Jusczyk, D. R. Mandel, J. Myers, A. Turk, and L. A. Gerken (1995) The head-turn preference procedure for testing auditory perception. *Infant Behav. Dev.* 18, 111–16.
28. Krumhansl, C. L. and P. W. Jusczyk (1990) Infants' perception of phrase structure in music. *Psychological Sci.* 1, 70–3.
29. Hirsh-Pasek, K., D. G. Kemler Nelson, P. W. Jusczyk, K. Wright Cassidy, B. Druss, and L. Kennedy (1987) Clauses are perceptual units for young infants. *Cognit.* 26, 269–86.
30. Ilari, B. and L. Polka (2002) Memory for music in infancy: the role of style and complexity. Poster presented at the International Conference on Infant Studies, Toronto, April, 2002.
31. Palmer, C., M. K. Jungers, and P. W. Jusczyk (2001) Episodic memory for musical prosody. *J. Mem. Language* 45, 536–45.

32. **Trainor, L. J., L. Wu, C. D. Tsang,** and **J. Plantinga** (2002) Long-term memory for music in infancy. Poster presented at the International Conference on Infant Studies, Toronto, April, 2002.
33. **Schellenberg, E. G.** and **S. E. Trehub** (2000) Children's memory for familiar music. Paper presented at the meeting of the Psychonomic Society, New Orleans.
34. **Trehub, S. E., E. G. Schellenberg,** and **D. Hill** (1997) The origins of music perception and cognition: a developmental perspective. In I. Deliège and J. Sloboda (eds) *Perception and Cognition of Music.* East Sussex, UK: Psychology Press, pp. 103–28.
35. **Saffran, J. R.** and **G. J. Griepentrog** (2001) Absolute pitch in infant auditory learning: evidence for developmental reorganization. *Dev. Psychol.* 37, 74–85.
36. **Halpern, A. R.** (1989) Memory for the absolute pitch of familiar songs. *Mem. Cognit.* 17, 572–81.
37. **MacDougall-Shackleton, S. A.** and **S. H. Hulse** (1996) Concurrent absolute and relative pitch processing by European starlings. *J. Comparative Psychol.* 110, 139–46.
38. **Jusczyk, P. W.** and **C. L. Krumhansl** (1993) Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *J. Exp. Psychol.: Hum. Percept. Perfor.* 19, 627–40.
39. **Schellenberg, E. G.** and **S. E. Trehub** (1996) Natural musical intervals: evidence from infant listeners. *Psychological Sci.* 7, 272–7.

CHAPTER 4

# MUSIC, COGNITION, CULTURE, AND EVOLUTION

IAN CROSS

## Abstract

We seem able to define the biological foundations for our musicality within a clear and unitary framework, yet music itself does not appear so clearly definable. Music is different things and does different things in different cultures; the bundles of elements and functions which are music for any given culture may overlap minimally with those of another culture, even for those cultures where 'music' constitutes a discrete and identifiable category of human activity in its own right. The dynamics of culture, of music as cultural praxis, are neither necessarily reducible, nor easily relatable, to the dynamics of our biologies. Yet music appears to be a universal human competence. Recent evolutionary theory, however, affords a means for exploring things biological and cultural within a framework within in which they are at least commensurable. The adoption of this perspective shifts the focus of the search for the foundations of music away from the mature and particular expression of music within a specific culture or situation and on to the human capacity for musicality. This paper will survey recent research that examines that capacity and its evolutionary origins in the light of a definition of music that embraces music's multifariousness. It will be suggested that 'music', like speech, is a product of both our biologies and our social interactions: that 'music' is a necessary and integral dimension of human development: and that 'music' may have played a central role in the evolution of the modern human mind.

We can express our understanding of biology within a framework that enables us to relate it, if not reduce it, to our understanding of the world in physical and material terms. Biological and physical understandings of the world are *commensurable*, in at least one of the senses that Lakoff[1] (p. 322) proposes. An understanding of ourselves as biological beings appears to be an understanding of 'natural kinds'.[a] But is music a 'natural kind', comprehensible within the generalized framework that is science?

[a] This is not to endorse the idea that there are 'natural kinds', that science provides an account of the essences of things in the world. The term 'natural kind' is used here simply as a concise way of referring to the objects of scientific discourse. As Rose[2] (p. 42) points out, even a concept as seemingly 'natural' and unambiguous as a 'protein' is susceptible to multiple and differing levels of definition that are dependent on 'the purposes for which we need to make the definition'. There is an incontestably societal dimension to the make-up of what is taken to constitute science at any time. But the notion that scientific procedures and understandings are simply varieties of social practice definable by their particular vocabularies[3] or by their poverty and abstraction[4] is insufficient to account for the instrumentality of those procedures and for the commensurability of the understandings that they afford.

Many argue that music is not a natural kind. Indeed, following a conventional dictionary definition of music—'The art of combining sounds of voices or instruments to achieve beauty of form and expression of emotion'—it would be difficult to do so. The consensual view from within the humanities appears to be that music is cultural rather than natural; music is viewed as constituted of practices, concepts, and perceptions that are grounded in particular social interactions and constructions. Molino[5] (p. 169), in questioning the status of music as a natural kind, proposes that 'Nothing guarantees that all the forms of human music contain a nucleus of common properties that would be invariant since the origination of music'.

As Geertz[6] (p. 5) put it, in promoting a semiotic and interpretive approach to culture, 'man is an animal suspended in webs of significance he himself has spun', and within Geertz's humano-centric web of culture there is little room for the 'natural'. For Treitler[7] (p. 203), 'Meaning in music is a function of the engagement of codes or orders by the note complexes of which the musical event is comprised', and musical phenomena are thus 'intelligible only in the light of an interpretation which intuits the purpose or intention that they embody'. Tomlinson, inciting musicologists to embrace Geertz's concept of culture, makes explicit the idea that scientific generalization is incompatible with musicological method; he asserts[8] (p. 352) that the essence of cultural—and hence musicological—explanation is 'not to codify abstract regularities but to make thick description possible, not to generalize across cases but to generalize within them'. Indeed, Abbate[9] has suggested (p. xv) that 'There is nothing immanent in a musical work (beyond the material reality of its written and sonic traces) and our perceptions of forms, configurations, meanings, gestures, and symbols are always mediated by verbal formulas, as on a broader scale by ideology and culture'. And Garnett[10] proposes that 'there is... no extra-cultural locus from which to observe music, nor extra-cultural meaning to observe'. 'Music' is seen as the expression of discrete, contingent, and socially conditioned factors in respect of which a generalizable—and hence scientific—account is neither relevant nor possible.

Such an approach to understanding music appears justified in view of the heterogeneity of forms that music can take. What 'nucleus of common properties' other, perhaps, than the very concept of the musical work (see Ref. 11) underlies such diverse products of western culture (or, as Slobin[12] has put it, musical microcultures) as (i) a performance of a Machaut motet; (ii) the autograph score of a Beethoven string quartet; (iii) the concept of Alvin Lucier's 'I am sitting in a room...'; (iv) a live broadcast of Brian Ferneyhough's 'Transit'; (v) the grooves in the vinyl that constitute the recording of the Holy Modal Rounders' track 'Half a mind to have a mind'; (vi) and the samples that make up part of the mastering materials of a Dr Dre CD? And indeed this last example undermines the very concept of the musical work itself. It might be suggested that 'the art of combining sounds of voices or instruments...' provides such a common property (though the issue of 'the achievement of beauty...' seems moot). But if we look beyond what Tomlinson[8] refers to as a 'presentist'[b] view of our own culture, that common property evanesces.

[b] Tomlinson[8] (p. 358): 'The presentist view of art works as transcendent entities fully comprehensible without reference to the conditions of their creation sacrifices Geertz's expansion of human discourse for a solipsistic and ultimately narcissistic aestheticism'.

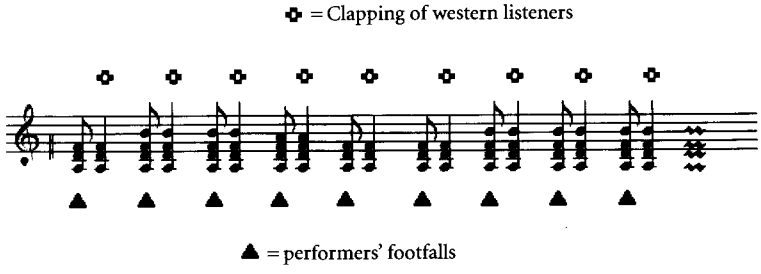
Indeed, when we look to other cultures, the notion of music *per se* is called into question. For example, Malm[13] (p. 5) questions whether or not the sound of the bull roarer used in some Arnhem Land aboriginal ceremonies constitutes 'music', as 'its sounds in secret rituals are not considered as independent sonic events but rather are thought to be the sounds of the supernatural itself'. And a culture may be constituted such that it does not distinguish a discrete category of practices that map onto those that would comprise music from a western perspective. Gourlay[14] notes that some cultures employ terms far more inclusive than the western notion of music; for the Igbo of Nigeria, *nkwa* denotes 'singing, playing instruments, and dancing'. Thus the anthropologist approaching Igbo culture with a view to examining its 'music' is confronted by a dilemma; as Gourlay[14] puts it (p. 35) 'By forcing the Igbo concept into the Procrustean bed of western conceptualization, she is in fact surrendering to the dominance of western ideas—or at least to the dominance of the English language. How different things would be if the Igbo tongue had achieved the same 'universality' as English! We should then have been seeking for universals in *nkwa*, and regarding the whole process of western "serious" music as an aberration because it excluded dance'. If the very concept of 'music' is so variable and inextricable from its cultural context how can we expect to seek, far less find, its biological foundations?

Perhaps some space for the 'natural' in conceptions of music might be found through recourse to cognitive anthropology rather than to anthropology *per se*, using the notion of mind to connect culture with biology. As D'Andrade[15] notes, the tacit notion of culture even within the anthropological consensus is that it is *in the mind.* He suggests (p. 146) that 'Since the 1950s most anthropologists have defined culture as a purely *mental* phenomenon'. Hence if culture is a mental phenomenon 'The structures that exist in the physical world as objects or events... are all thought to be... more or less a reflection of these mental cultural structures'. And given the success of cognitive science in instating mind in the material, it seems reasonable to expect that notions of mind might furnish terms that could connect the discourses of anthropology and of the natural sciences.

But cognitive anthropology, or from another perspective, cultural psychology (see Ref. 16) offers little comfort. For there is little consensus about the depth to which culture permeates the grain of our experience. The question of whether we should conceive of some psychic unity upon which culture forms an overlay, or a psychic diversity that is principally constructed by culture, remains unresolved. D'Andrade[15], for example, leaves space for the notion of a primitive and perceptual psychic unity when he suggests (p. 217) that 'culture seems to have its greatest effect... on semantic memory and complex reasoning'. But he undermines that notion by suggesting (p. 184) that 'In general, naive perception can be influenced by cultural schemas'. A practical musical example might uphold the second suggestion by showing that culture indeed appears to determine the grain of our unreflective perceptions.

One factor that appears to apply to almost all the world's musics is that there is a level of temporal organization that is regular and periodic, sometimes called the tactus. It is taken to correspond to the regular points in the music where one would tap one's foot or clap along. In listening to a piece in a familiar idiom, all listeners are capable of tapping along without thinking (even if it is frowned upon in the concert hall). And even when encountering previously unheard music from an unknown culture, a listener can still 'keep a beat'. Most western

listeners confronted by this particular piece from Northern Potosí in Bolivia (recorded by my friend and colleague Henry Stobart, an ethnomusicologist specializing in Andean music) are likely to clap or tap along with the regularly spaced longer and louder notes:




An example of a recording of this piece can be heard on the Internet by accessing http://www.mus.cam.ac.uk/~cross/BJE/ and selecting Example III(a). At the end of the recording something occurs that gives the game away; for the first time the regular footfalls that accompany—that are part of—the performance can be heard. And they coincide not with the longer and louder notes, but with the short, sometimes almost inaudible notes that alternate with the long notes. For the people of the culture from which this music comes, these appear to be the appropriate places to tap the feet—that is, the point at which the tactus occurs. For western listeners, while one can *learn* to clap or tap at those points in the piece, it does not appear to *feel*—and I speak from long experience—as though one is tapping on the beat. It always feels as though one is tapping on an offbeat.

While certain features of this music's organization can be accounted for by invoking the operation of perceptual processes underpinning the experience of time that appear genuinely universal, the fact that longer, more intense, notes do not mark out the tactus cannot be explained in this way. In fact it seems likely that prosodic features of the language of Northern Potosí, Quechua, relate to the way in which tactus is organized, projected, and experienced in that culture's music (see Ref. 17).

That even such an apparently unreflective act as regularly tapping the foot in time to a piece of music is so susceptible to cultural differentiation appears to suggest either that tapping one's foot in time to music has a semantic component (if D'Andrade's proposal that culture impacts cognition primarily at the conceptual level is accepted) or that human cognitive capacities are so grounded in culture that any elementary commonalities are over-ridden, and that minds are only susceptible to explanation in terms specific to the particular cultures in which those specific minds are rooted—a return to the position of Geertz. In other words, culture is in the bones and science has no place in its understanding.

We appear to have reached an impasse; it seems that music is cultural, variable and particular, and not susceptible to explanation in general and scientific terms. Yet there are those who argue that music is, nevertheless, a human universal. Blacking[18] (p. 224) states that 'every known human society has what trained musicologists would recognize as "music", while Merriam[19] (p. 227) bluntly asserts that music 'is a universal behavior'. How can these claims be squared with music's cultural particularity?

As a first step we must enquire what Blacking and Merriam mean by 'music' in this universal manifestation. For both, music is not just sound. Indeed, the musical example

given above is not 'the music'; it was a recording of the sound of a musical activity in a particular cultural context. To experience 'the music', you might have to undergo what has been called[20] the 'Total Turing Test' of lifelong immersion in the culture; at the least you would have get the feel of the instrument and of the movements involved. The 'music' would involve embodied action as much as disembodied sound. Even in our own culture it is only in the last hundred years, with music becoming an increasingly commodified aural consumable, that the self-evident ties between musical sound and human movement have been rendered obscure. For John Blacking, the claim that music is a human universal explicitly involves acknowledgement of the embodied nature of music, the *indivisibility* of movement and sound in characterizing music across times and cultures; he claims[18] (p. 241) that '"Music" as a human capability is a cognitive, and hence affective, activity of the body'. For the greater number of cultures in the world, and for the greater part of the historical existence of our culture, 'music' appears to have involved and to involve *movement* just as much as sound.

But music in its universal guise not only involves sound and movement, but also it involves multiplicity of reference and meaning; for Blacking and Merriam music is intrinsically polysemic. For example, music can function as a means of communication with the dead for the Kaluli of Papua New Guinea, binding birds, souls, places, and people at a time of transformation; or music can function in the restructuring of social relations, as in the *domba* initiation of the Venda.[21] In each of these two very different ceremonies, music is central, its meaning rarely if ever explicit but its fugitive significances essential. Blacking[18] notes (p. 237) 'the "same" sound patterns... can... have different meanings within the same society because of different social contexts'. And as Merriam[19] (p. 221) suggests, a defining characteristic of the musical utterance is its property of being 'unrepudiable in form but repudiable as to content'. In other words, music has the capacity to *lack* consensual reference; it can be *about* something, but its *aboutness* can vary from context to context and even within context.

Not only may music's significance vary according to social context, but the significances of a singular musical activity can vary from individual to individual. We know this from our experience of music in our own cultures and we can see it in others; for example, in the *gisalo* ceremonies of the Kaluli, some performers seek to dominate and direct while the performances of others appear to emerge from performer-audience interaction[22] and in any particular performance some participants weep while others do not.[23]

And finally, music in general has a further, peculiarly negative, feature; it appears to have no immediate and evident efficacy. Music neither ploughs, sows, weaves nor feeds; in itself, if it can be considered to exist outwith its context of use, it does not seem to be capable of being a material cause of anything other than a transient hedonic encounter. It is inefficacious.

From these considerations of the 'universal' characteristics of music we can return to the original question of whether music can be construed as a natural kind. We now appear to have a basis for proposing an operational definition of music that might afford the commonalities that would allow an instatement of the natural in the musical. It seems that a generalizable definition of music would refer to music's two roots in sound and movement, to music's heterogeneity of meaning, to its grounding in social interaction yet personalized

significance, and to its inefficacy. Putting these four premises together yields the following operational definition:

> Musics can be defined as those temporally patterned human activities, individual and social, that involve the production and perception of sound and have no evident and immediate efficacy or fixed consensual reference.

When applied to the mature expressions of music in particular cultures, this definition does no more than provide a conceptual umbrella for otherwise potentially heterogeneous ethnographies; in that context it is descriptive rather than explanatory. It constitutes a general proposition that applies to the universal presence of something like 'music' in all human cultures. However, in claiming that music is universal Blacking goes further than this; he suggests that music is ubiquitous not only across human societies but across all members of those societies. As he states[18] (p. 236) '... the almost universal distribution of musical competence in African societies suggested that musical ability [is] a general characteristic of the human species rather than a rare talent'. This suggestion squares with recent research into the precursors of musical ability in a western context, notwithstanding those such as Barrow[24] (p. 194) who asserts that 'musical ability is... limited in its distribution'. While it is self-evidently true that the production of 'music' in contemporary western society is in the hands of a specialized class of performers and composers (and lawyers), musical ability cannot be defined solely in terms of productive competence; (almost) every member even of our own, highly specialized, society is capable of listening to and hence of *understanding* music. Indeed, recent research[25] can be interpreted as suggesting that musical productive abilities in a western context, rather than being rare capacities that are evidence of some inborn 'talent', are better explained in terms of the effects of motivation and of practice. In other words, music is not just universal across cultures; it appears that everyone has the capacity to be musical, though this capacity is likely to be realized to different degrees and in different ways in different cultural and social environments.

It is within this broader notion of music as a universal human attribute that the operational definition of music given above might be informative, as applied to the human *capacity* for musicality; if borne in mind in the exploration of the propensities for, and the functionalities of, music for infants and children, it might yield an understanding of the commonalities that appear necessary in order to relate music to our biologies. And as Sandra Trehub notes in this volume, investigations of infant and childhood musical capacities do appear to reveal cross-cultural invariants.

Sandra Trehub and her collaborators have shown that even young infants possess the capacities to perceive significant structural and affective features of musical sounds.[26] The real-world context in which such capacities are most evident is the typically affect-laden interaction of the infant with a care-giver (and it is notable that, irrespective of culture, even adults appear sensitive to features of infant directed song[27].) Other researchers such as the Papouseks and, more recently Colwyn Trevarthen and his collaborators have focused closely on the musicality of such interactions.

The Papouseks have been particularly concerned with infant vocal capacities and interactions. Hanus Papousek[28] (p. 43) has noted that 'musical elements participate in the process of communicative development very early', suggesting that 'they pave the way to

linguistic capacities earlier than phonetic elements'. He sees (pp. 46–47) infant and early childhood musical behaviours as forms of play involving higher level integrative processes that act to nurture 'exploratory competence' (a notion that seems to rely on the idea of musical signification as transposable); these exploratory competences entail the participation of emotions and constitute precursors of artistic or scientific competencies.

Mechthild Papousek focuses on the musicality of infant-caregiver interactions, stressing the indivisibility in these of music and movement and the fact that they appear to involve patterns of infant *and* caregiver behaviour that are singularly invariant across cultures. She notes[29] (p. 100) that 'parents' multimodal stimulation is tailored to infants' early competence for perceiving information through different senses as coordinated wholes', and that 'regular synchronization of vocal and kinaesthetic patterns provides the infant with multimodal sensory information including tactile, kinaesthetic, and visual information'.

The work of Trevarthen and his colleagues has centred on these temporal characteristics of infant-caregiver interaction. Trevarthen[30] states that, from birth, central to our neuronal anatomy is a 'body-imaging core system' that comes to act so as to integrate attention, learning, and self-regulating physiology with actions of expression and execution; this he terms the Intrinsic Motive Formation (IMF). In operation, the IMF incorporates periodic timing mechanisms that give rise to a 'hierarchy of motor rhythms'; these, governing movement and binding affect in rhythmic time, he calls the Intrinsic Motive Pulse. For Trevarthen[30] (p. 160) 'Musicality...is the aurally appreciated expression of the IMF with the Intrinsic Motive Pulse as its agent'. From these premises, Trevarthen develops a conceptual framework to explore the expression and development of communication—of intersubjectivity, in his terms—through empirical observations and analyses of infant-caregiver interaction.

The rhythmicity of caregiver-infant interaction, in terms of the capacity of the infant to follow and respond in kind to temporal regularities in vocalization and in movement, and in time to initiate temporally regular sets of vocalizations and movements, is seen here as central to the development of human significative and communicative capacities; its embodied nature enables the sharing of patterned time with others and facilitates harmonicity of affective state and interaction. For Trevarthen, that rhythmicity is also a manifestation of a fundamental musical competence. As he frames it[30] (p. 194), 'Musicality is part of a natural drive in human socio-cultural learning which begins in infancy'.

There is, thus, an increasing amount of evidence that musicality is in our birthright; the capacity for music is an integral component of the infant mind. However, the notion of music as innate that emerges from the research just cited does not sit easily with current general theories of the infant mind. While these theories are increasingly nativist in suggesting that the infant mind, rather than being domain general, is endowed with either modular or domain-specific competences, they tend to account for the existence of these competences on the basis of their adaptive value in evolution and most theorists see no adaptive role for human musicality in evolution. The present consensus (see, e.g. Ref. 31) suggests that the infant mind is primed for the rapid emergence of competence in (at the least) interpreting social relations, physical and mechanical interactions, and the behaviour of biological systems. This view is supported by a substantial quantity of empirical research, in particular by recent work[32] that supports the cross-cultural generality of some of these domains.

There have been some suggestions that musicality might constitute one of these 'native domains'. Gelman and Brenneman[33] propose that a domain-specific competence in music is

evidenced by the results of Trehub and others, suggesting that sensitivities to harmonicity of tonal relations and to melodic contoural formations constitute evidence of a music-specific competence. However, their conclusions are rather undermined by their focus on perceptual capacities of which the existence might equally well be accounted for by their utility in other cognitive domains; the tendency to link sounds that are perceived as harmonically related, and to differentiate between sequences of sounds that differ in their contour, appears more likely to derive from a general capacity for auditory scene analysis[34] than to testify to the existence of an early and specifically musical competence. In other words, the suite of perceptual capacities that Gelman and Brenneman identify as making up the domain of musical competence might be epiphenomenal; each capacity might be more securely considered as being proper to other domains that are more self-evidently and immediately functional than is music.

Indeed when music has been viewed from an evolutionary perspective it has often been viewed as contingent, at best exaptive, a view most clearly exemplified by Steven Pinker[35] and endorsed by others such as Barrow[24] and Sperber.[36] For Pinker, music is famously 'auditory cheesecake'; while music in his view is bound to the domains of language, auditory scene analysis, habitat selection, emotion, and motor control, it does no more than exploit the capacities that have evolved to subserve each of these areas. Music is thus 'exaptive', an evolutionary by-product of the emergence of other capacities that have direct adaptive value. Barrow[24] similarly suggests that human musicality has had no role in our survival as a species, suggesting that it derives from an 'optimal instinctive sensitivity for certain sound patterns' that itself arose because it proved adaptive. Sperber[36] goes furthest in condemning music as an evolutionary 'parasite', though he explicitly disavows serious intent in formulating that view. Nevertheless, he does suggest that music is a human activity that arose to exploit parasitically the operation of a cognitive capacity to 'process complex sound patterns discriminable by pitch variation and rhythm' that was originally functional in primitive human communication but that fell into disuse with the emergence of the modern vocal tract and the finer shades of differentiation in sound pattern that it afforded. For Pinker, Sperber and Barrow, music exists simply because of the pleasure that it affords; its basis is purely hedonic, and, as Pinker puts it 'Compared with language, vision, social reasoning, and physical know-how, music could vanish from our species and the rest of our lifestyle would be virtually unchanged'.

These three views appear to constitute the beginnings of a consensus that would relegate music to the status of evolutionary footnote and would seem to vitiate the idea that its biological foundations deserve any attention. However, all three theories suffer from an attribute that disqualifies their conclusions from serious consideration, that of ethnocentricity. Theirs is a 'culture-lite' view of music. They take no account in their conclusions of the *indivisibility* of movement and sound in music,[c] focusing on only one dimension of music as defined above, that of music's inefficacy in any domain other than the individually hedonistic. Despite lip-service paid to the notion that music might take other forms in other cultures, music appears in these theories largely as disembodied sound oriented

[c] Although Pinker explicitly addresses the link between music and movement he treats this simply in terms of music 'tapping in' to systems of motor control.

towards individual hedonism, a notion quite untenable before the advent of recording technology. Indeed, over the last 100 years, recording technology together with the reification of intellectual property and the globalization of its law has sanctioned the subsumption of music into the capitalist economy as a tradable and consumable commodity. It might well be that Pinker's view of music is an accurate reflection of what music is now for some within western culture, but that culture-specific 'music' is scarcely representative of the complex and embodied set of activities and interpretations that are evident in most nonwestern 'musics'. To put it another way, what music is for some at present is not what music is for others, was for our predecessors or could be for our children.

Even when music has been viewed as adaptive in human evolution, the problem of ethnocentricity can remain. Miller,[37] in promoting the notion that 'Machiavellian intelligence' played and still plays a significant role as an agent in processes of sexual selection, suggests that musicality constitutes a marker for possession of such intelligence; musical performance constitutes a display of protean behaviours and functions so as to advertise to prospective mates the possession of the 'protean' capacity to be 'unpredictable', a capacity that he suggests is of value in social interaction. The putative link between music and sex certainly motivates many adolescents in our society to engage in 'musical' behaviours, but most will realise only too quickly that mere presence onstage is no guarantee of successful subsequent sexual interaction. If it could be demonstrated from a comprehensive cross-cultural survey that music's primary function is as a vehicle for the display of 'protean behaviours', it would be reasonable to infer that this was music's *raison d'être*. However, the available evidence does not sustain this view; music is and has been employed for many different ends by different societies, and in most the role of music in courtship is positively subsidiary to its value in activities of healing, praying, mourning, or instructing. Miller's view of music seems as bound to the peculiarities of current western practice as does that of Pinker.

For Pinker, Barrow, Sperber, and Miller, the effects of music are at the level of the individual, whether in terms of affording hedonic experience or exhibiting protean attributes. It's notable that much of the research into infant musicality (particularly that of Trevarthen) suggests a different locus for music's functionality (if any), that of human interaction. Several recent theories of music as adaptive in human evolution have located its functionality at the level of the group, including the writings of Kogan and of Brown. Kogan[38] (p. 197) notes that current evolutionary theory suggests that 'natural selection operates not only within groups but also between them'. He follows McNeill's[39] notion of 'muscular bonding' in proposing that the communal experience of affect elicited by moving together rhythmically in music and dance could have enhanced cooperative survival strategies for early humans, for example, in hunting or in inter-group conflict. This efficacy of rhythmic synchronicity in promoting group identity can be related back to the 'time-sharing' capacities exhibited in infant-caregiver interaction, though here it seems limited to its impact on affect, ignoring any broader functionality.

For Brown,[40] the adaptive features of music for the group go beyond those on which Kogan relies. Brown adduces the notion of music as reinforcing 'groupishness', which he defines as a 'suite of traits that favour the formation of coalitions, promote cooperative behaviour towards group members and create the potential for hostility towards those outside the group'. Music supports these traits through the opportunities that it offers for the

formation and maintenance of *group identity*, for the conduct of *collective thinking* (as in the transmission of group history and planning for action), for *group synchronization*—the sharing of time—between members of a group, and for *group catharsis*, the collective expression and experience of emotion. Ultimately, Brown sees music as having become instantiated in human cultures through its role as 'ritual's reward system'; music, for him, is a type of 'modulatory system acting at the group level to convey the reinforcement value of these activities... for survival'. And if Brown is correct in his portrayal of music's role in promoting 'groupishness', music is likely to have been a major contributor to what Smith and Szathmáry[41] hold to have constituted one of the major transitions in evolution: the very emergence of human culture.

Dissanayake[42] see the mature expression of music in human culture as intimately linked to the characteristics of mother–infant interaction. She views music (p. 390) as 'multimodal or multimedia activity of temporally patterned movements' that has 'the capacity to coordinate the emotions of participants and thus promote conjoinment'. She suggests that features of the musicality of mother–infant interaction might lay the foundations for a 'grammar of the emotions' that can be expressed in mature musical (and other artistic) activities. For the developing child, the musical characteristics of mother-infant interaction are of critical importance in the acquisition of capacities for 'social regulation and emotional bonding'; these characteristics also provide the elements in the 'musical play' of later childhood that will equip the adult with the predisposition and capacity to engage in the structured interactions of ceremony and ritual as well as in specifically musical behaviours. However, other significant and functional roles have been proposed for music in individual development and in the development of capacities for social interaction; music can be both a consequence free means of exploring social interaction and a 'play space' for rehearsing processes that may be necessary to achieve cognitive flexibility.[43]

Music is consequence free in that it is not *directly* functional; it is nonefficacious. It is specifically suited to testing out aspects of social interaction by virtue of both its noneffi-caciousness and its polysemic nature, its multiple potential meanings. For each child in a group ostensibly involved in a cooperative musical activity, that musical activity can mean something different yet the singularity of the collective musical activity is not threatened by the existence of multiple simultaneous and potentially conflicting meanings. Music provides for a child a medium for the gestation of a capacity for social interaction, a risk-free space for the exploration of social behaviour that can sustain otherwise potentially risky action and transaction.

Just as one can posit a role for music in the socialization of the child, one can also postulate a role for music in the development of the child's individual cognitive capacities that is quite distinct from its efficacy in the child's acculturation. Again, this role is motivated by the intrinsically polysemic nature of music, the fact that its significances can modulate from situation to situation and can even be simultaneously multiple. If music is *about* anything, it exhibits a deictic intentionality, a 'transposable aboutness'. And it is conceivable that music's 'transposable aboutness' is exploited in infancy and childhood as a means of forming connections and interrelations between different domains of infant and childhood competence such as the social, biological, and mechanical. To give a crude

example; the arc of a ball thrown through the air, the prosodic contour of a comforting utterance, the trajectory of a swallow as it hawks an insect, the pendular ballistics of a limb swung in purposive movement, might, for a child, each underlie the significances of a single musical phrase or proto-musical behaviour on different occasions. Indeed, these heterogeneous incidents may be bound together simultaneously in the significance of that phrase or behaviour, the music thus exhibiting what I have called elsewhere a 'floating intentionality'. The 'floating intentionality' of the music can provide for the child a space within which she can explore the possible bindings of these multidomain representations. Hence one and the same musical activity might, at one and the same time, be about the trajectory of a body in space, the dynamic emergence or signification of an affective state, the achievement of a goal and the unfolding of an embodied perspective. All these 'aboutnesses' exist not in respect of objects but events, ongoing structures in time, and music or proto-musical behaviours afford the opportunity to explore the cross-domain mappings that the representation of temporal sequences of object states as events makes available.

From this perspective one can advance a second definition of music, one that rests on the idea that what 'music' is for any given culture may vary immensely but will derive from the same general human propensities:

> Musics are cultural particularisations of the human capacity to form multiply-intentional representations through integrating information across different functional domains of temporally extended or sequenced human experience and behaviour, generally expressed in sound.

In this view, music, or proto-musical behaviours, subserve a *metaphorical* domain or perhaps more appropriately, underpin a metaphorical stance, acting to create and to maintain the cognitive flexibility that marks off humans from all other species. And it could be that the emergence of proto-musical behaviours and their cultural realization as music (and, for the matter, dance) might themselves have been crucial in precipitating the emergence of the cognitive flexibility that marks the appearance of *Homo sapiens*. For if Smith and Szathmáry[41] are correct in maintaining that human culture constitutes one of the major transitions in evolution, and Mithen[44] is correct in claiming that the appearance of *Homo sapiens sapiens*, ourselves, is marked by the emergence of a flexible cross-domain cognitive capacity, then music is uniquely fitted to have played a significant role in facilitating the acquisition and maintenance of the skill of being a member of a culture—of interacting socially with others—as well as providing a vehicle for integrating our domain-specific competences so as to endow us with the multipurpose and adaptive cognitive capacities that make us human.

Of course, what music, or more appropriately proto music, is for infants and children and what it might have been in evolution, is not necessarily what music is for a mature culture or society. Culture shapes and particularizes proto-musical behaviours and propensities into specific forms for specific functions, and those, as noted at the outset of this paper, can be so divergent that they do not appear to be mutually reducible—they do not appear to exhibit a 'nucleus of common properties' (after Molino[5]). However, it is noteworthy that for most cultures, music—and here one might almost prefer to use the Igbo term *nkwa*, as

it seems to capture the interlinking of sound and action that characterizes music for most cultures—the functions that music fulfils, the contexts in which it appears most efficacious, often lie in the realm of ritual and of psychic healing. That is, music often functions in individual and group encounters with the numinous and in the modulation of affective state.

The affective functionality of music can certainly be referred to its embodiment in action and to the contexts within which proto-musical activities occur for infants and children,[42] as well as to a broad range of different circumstances in specific cultures.[45] But it seems that music's ubiquity and efficacy in encounters with the numinous are best accounted for by reference to proto-music's polysemy, its 'floating intentionality'. This property of proto-musical activities may facilitate the mature use of music in those cultural contexts that deal with what Sperber has called 'relevant mysteries'. Sperber applies this term to situations where beliefs or mental representations arise which are contradictory but are each separately related to (and hence relevant in respect of) other mental representations and beliefs. When simultaneously foregrounded by actions or circumstances, these contradictory beliefs then become 'mysteries' and 'achieve relevance because of their paradoxical character—that is because of the rich background of everyday empirical knowledge from which they systematically depart'[36] (p. 72). Within the framework of Sperber's theory, religious ideas are distinguished from everyday beliefs by their paradoxicality and their relevance, by their broad applicability and their ambiguity; and the view of music's functionality outlined above would suggest that music is also distinguished by just such a broad applicability and ambiguity. By virtue of these attributes music may thus be particularly appropriate as a means of amplifying, exemplifying or reinforcing in the course of ongoing experience just these attributes of belief that are interpretable as religious; music's indeterminacy may suit it for use as a means of pursuing and perhaps even parsing the numinous.

But the factors that endow music with its efficacy for individual cognitive development and socialization in infancy and childhood cannot by themselves determine the multiplicitous forms and functions that music takes and that music serves in mature cultural contexts. The meaning of a musical activity for a mature individual will necessarily depend at any given moment on that person's own history and narratives, and on the situational significances that culture's 'shared system of meanings' confer on that activity. The polysemic potential that characterises proto-musical activity is likely to underpin the social functionality of music and to contribute to, but not determine, music's meaning. The functionalities and functions of music or proto-musical behaviours for the individual, whether in their own cognitive development or in their socialization, must be set in the context of the functionalities and functions of music as a cultural phenomenon. Music, like language, cannot be wholly private; it is a property of communities, not individuals. And these different levels at which music may be efficacious must be integrated in any understanding of its foundations. Music's very existence is best evidenced in interaction. If music is of importance in human development, evolution and life, then an attempt to render commensurable our understanding of music as interaction with our understanding of music's biological foundations is crucial in coming to terms with what Henry Plotkin[46] (p. 222) calls 'the most complicated thing in the universe—the collective of human brains and their psychological processes that make up human culture'.

To return to the beginning; what are the implications for this view of music as something more than patterned sound, for an understanding of its biological foundations? We would expect its neurophysiology to be complex, reaching beyond the auditory pathways to the limbic system and to centres of motor behaviour. We would expect that the cultural context of music—the forces that shape music for any given culture—should condition its neurophysiological correlates. And we might expect music and language to share many, but not all, neurophysiological correlates.[47] While music and language might meet somewhere near poetry, music can never attain the unambiguous referentiality of language (which Deacon[48] holds to be language's primary defining characteristic), nor language the absolute ambiguity of music.

But at the limit, while music may be in our biologies, our culture is in our music. If the roots of human musicality are to be found in infancy, particularly in infant-caregiver interaction, its potency might be tied to the support provided by society for those interactions. In an intriguing study Maya Gratier[49] investigated the coherence of interactions between caregivers and infants in three different contexts: French mothers in France, Southern Indian mothers in South India, and Southern Indian mothers who were recent immigrants to France. She found a difference between the coherence of the immigrant mother-infant interactions and those of the other two culturally embedded groups; interaction between the immigrant mothers and their infants was significantly less coherent than was interaction in the other two groups. She suggests that the cultural dislocation of the immigrant mothers had impacted directly on their capacities to interact 'musically' with their infants. Something as individual and putatively innate as the capacity of a mother to interact coherently in time with her child seems to be dependent on the mother's rootedness in her cultural environment. In other words, if music is in our birthright, its inheritance appears to be a fragile gift that rests on the humaneness and sympathy of the culture that surrounds us.

## References


1. **Lakoff, G.** (1987) *Women, Fire and Dangerous Things.* Chicago: University of Chicago Press.
2. **Rose, S.** (1996) *Lifelines: Biology, Freedom, Determinism.* London: Allen Lane.
3. **Rorty, R.** (2000) Being that can be understood is language. *London Rev. Books* 126, 23–5.
4. **Feyerabend, P.** (1981) *Problems of Empiricism*, vol. 2. Cambridge: Cambridge University Press.
5. **Molino, J.** (2000) Towards an evolutionary theory of music and language. In N. Wallin, B. Merker, and S. Brown (eds) *The Origins of Music.* Cambridge, MA: MIT Press, pp. 165–76.
6. **Geertz, C.** (1973) *The Interpretation of Cultures: Selected Essays.* New York: Basic Books.
7. **Treitler, L.** (1980) History, criticism and Beethoven's ninth symphony. *Nineteenth Century Music* 3, 193–210.
8. **Tomlinson, G.** (1984) The web of culture. *Nineteenth Century Music* 7, 350–62.
9. **Abbate, C.** (1991) *Unsung Voices: Opera and Musical Narrative in the Nineteenth Century.* Oxford: Princeton University Press.
10. **Garnett, L.** (1998) Musical meaning revisited: thoughts on an 'epic' critical musicology. *Critical Musicology.* http://www.leeds.ac.uk/music/Info/CMJ/Articles/(1998)/01/01.html.
11. **Goehr, L.** (1992) *The Imaginary Museum of Musical Works.* Oxford: Clarendon Press.

12. Slobin, M. (1993) *Subcultural Sounds: Micromusics of the West.* Hanover: Weslyan University Press.
13. Malm, M. (1977) *Music Cultures of the Pacific, the Near East and Asia.* Englewood Cliffs, NJ: Prentice-Hall.
14. Gourlay, K. A. (1984) The non-universality of music and the universality of non-music. *World Music* 26, 25–36.
15. D'Andrade, R. (1996) *The Development of Cognitive Anthropology.* Cambridge: Cambridge University Press.
16. Cole, M. (1996) *Cultural Psychology.* London: Belknap Press of Harvard University Press.
17. Stobart, H. F. and I. Cross (2000) The Andean Anacrusis? rhythmic structure and perception in Easter songs of Northern Potosí, Bolivia. *Br. J. Ethnomusicol.* 9(2), 63–94.
18. Blacking, J. (1995) *Music, Culture and Experience.* London: University of Chicago Press.
19. Merriam, A. P. (1964) *The Anthropology of Music.* Chicago: Northwestern University Press.
20. Harnad, S. (1992) There is only one mind/body problem. *Intl. J. of Psychol.* 27, 521.
21. Blacking, J. (1976) *How Musical is Man?* London: Faber.
22. Scheifflin, E. L. (1993) Performance and the cultural construction of reality: a New Guinea example. In R. Rosaldo, S. Lavie, and K. Narayan (eds) *Creativity/Anthropology.* London: Cornell University Press, pp. 270–295.
23. Feld, S. (1982) *Sound and Sentiment: Birds, Weeping, Poetics and Song in Kaluli Expression.* Philadelphia: University of Pennsylvania Press.
24. Barrow, J. D. (1995) *The Artful Universe.* Oxford: Clarendon Press.
25. Sloboda, J. A., J. W. Davidson, M. J. A. Howe, and D. G. Moore (1996) The role of practice in the development of performing musicians. *Br. J. Psychol.* 87, 287–309.
26. Trehub, S. E., G. Schellenberg, and D. Hill (1997) The origins of music perception and cognition: a developmental perspective. In I. Deliège and J. Sloboda (eds) *Perception and Cognition of Music.* Hove: The Psychology Press, pp. 103–28.
27. Trehub, S. E., A. M. Unyk, and L. J. Trainor (1993) Adults identify infant-directed music across cultures. *Infant Behavior Development* 16, 193–211.
28. Papousek, H. (1996) Musicality in infancy research: biological and cultural origins of early musicality. In I. Deliège and J. Sloboda (eds) *Musical Beginnings.* Oxford: Oxford University Press, pp. 37–55.
29. Papousek, M. (1996) Intuitive parenting: a hidden source of musical stimulation in infancy. In I. Deliège and J. Sloboda (eds) *Musical Beginnings.* Oxford: Oxford University Press, pp. 88–112.
30. Trevarthen, C. (1999) Musicalty and the intrinsic motive pulse: evidence from human psychobiology and infant communication. *Musicae Scientiae* (Special Issue), 155–215.
31. Spelke, E. (1999) Infant cognition. In R. A. Wilson and F. C. Keil (eds) The MIT Encyclopedia of Cognitive Sciences. Cambridge, MA: MIT Press, pp. 402–4.
32. Medin, D. and S. Atran (1999) *Folk Biologies.* Cambridge, MA: MIT Press.
33. Gelman, R. and K. Brenneman (1994) First principles can support both universal and culture-specific learning about number and music. In L. A. Hirschfeld and S. A. Gelman (eds) *Mapping the Mind: Domain Specificity in Cognition and Culture.* Cambridge: Cambridge University Press, p. 396.
34. Bregman, A. (1990) *Auditory Scene Analysis.* Cambridge, MA: MIT Press.
35. Pinker, S. (1997) *How the Mind Works.* London: Allen Lane.

36. **Sperber, D.** (1996) *Explaining Culture.* Oxford: Blackwell.

37. **Miller, G. F.** (1997) Protean primates: the evolution of adaptive unpredictability in competition and courtship. In A. Whiten and R. W. Byrne (eds) *Machiavellian Intelligence II: Extensions and Evaluations.* Cambridge: Cambridge University Press, pp. 312–40.

38. **Kogan, N.** (1997) Reflections on aesthetics and evolution. *Critical Rev.* 11, 193–210.

39. **McNeill, W. H.** (1995) *Keeping Together in Time.* London: Harvard University Press.

40. **Brown, S.** (2000) Evolutionary models of music: from sexual selection to group selection. In F. Tonneau and N. S. Thompson (eds) *Perspectives in Ethology. 13: Behavior, Evolution and Culture.* New York: Plenum Publishers, pp. 231–81.

41. **Smith, J. M.** and E. **Szathmáry** (1995) *The Major Transitions in Evolution.* Oxford: Oxford University Press.

42. **Dissanayake, E.** (2000) Antecedents of the temporal arts in early mother-infant interactions. In N. Wallin, B. Merker, and S. Brown (eds) *The Origins of Music.* Cambridge, MA: MIT Press, pp. 389–407.

43. **Cross, I.** (1999) Is music the most important thing we ever did ? Music, development and evolution. In Suk Won Yi (ed.) *Music, Mind and Science.* Seoul: Seoul National University Press, pp. 10–39.

44. **Mithen, S.** (1996) *The Prehistory of the Mind.* London: Thames and Hudson.

45. **Sloboda, J. A.** and **P. Juslin** (eds) (2001) *Music and Emotion: Theory and Research.* Oxford: Oxford University Press.

46. **Plotkin, H.** (1997) *Evolution in Mind.* London: Allen Lane.

47. **Peretz, I., S. Belleville,** and **S. Fontaine** (1997) Dissociation between music and language following cerebral hemorrhage – another instance of amusia without aphasia. *Canadian J. Exp. Psychol.* 51, 354–68.

48. **Deacon, T.** (1996) *The Symbolic Species: The Co-evolution of Language and the Human Brain.* London: Allen Lane.

49. **Gratier, M.** (1999) Expressions of belonging: the effect of acculturation on the rhythm and harmony of mother-infant interaction. *Musicae Scientiae,* (Special Issue), 93–122.

CHAPTER 5

# IS MUSIC AN EVOLUTIONARY ADAPTATION?

DAVID HURON

## Abstract

In contemplating the function and origin of music, a number of scholars have considered whether music might be an evolutionary adaptation. This article reviews the basic arguments related to evolutionary claims for music. Although evolutionary theories about music remain wholly speculative, musical behaviours satisfy a number of basic conditions, which suggest that there is indeed merit in pursuing possible evolutionary accounts.

**Keywords:** Evolutionary theories of music; Music industry; Evolutionary origin of language; Music and social bonding; Oxytocin; Mood regulation

Addressing the question of music's origins has a long history and is patently speculative. Although several cultures have provided colourful stories describing how people acquired the capacity for music, most contemporary scholars have focused on possible psychological, social, and cultural beginnings. In this chapter, I propose to offer a social account of music's origins that is explicitly linked to one of the most successful theories yet devised: the theory of evolution by natural selection.

Evolution is often thought of in purely physiological rather than psychological terms.[1–5] It is not simply that evolution has shaped immune systems, digestive tracts, and knee caps. Evolution has also shaped our attitudes, dispositions, emotions, perceptions, and cognitive functions. Some of our deepest convictions can be traced to plausible evolutionary origins: we love life, we fear death, and we nurture our children because these dispositions better ensure the propagation of our genes.

The theory of evolution by natural selection is a distal theory. It is not a theory that explains specific behaviours, such as why you chose to cook ravioli for dinner last night, or why you parked in a particular parking spot this morning. Similarly, if music is an evolutionary adaptation, this will not allow us to account for particular musical acts, such as Mendelssohn's writing of the *Scottish Symphony*. Evolution proceeds by selecting traits that are adaptive to an organism's environment. For example, evolution did not 'originate' or 'create' the phenomenon of altruism. Instead, given a certain environment, natural selection

favoured individuals who exhibited certain altruistic traits. Evolution does not dictate our behaviour: it selects which behaviours are likely to be passed on to subsequent generations—and it selects only those behaviours that have a genetic component. So in discussing possible evolutionary origins for musical behaviours, the question is not, What caused people to make music? but rather, How might music-making behaviours have escaped the hatchet of natural selection? or more precisely, What advantage is conferred on those individuals who exhibit musical behaviours over those who do not?

## Does music have survival value?

Many knowledgeable people have concluded that music has no survival value. Indeed, a number of esthetic philosophers have argued that an essential, defining characteristic of the arts is that they serve no practical function.[6] Accordingly, any music that is created for biological (or even economic) reasons cannot be considered art. Even among evolutionary psychologists such as Steven Pinker, it has been common to suppose that music is not adaptive.[7]

Many linguists—Pinker included—believe that language is likely an evolutionary adaptation.[8] However, the evidence in support of language as an adaptation is not notably stronger than comparable evidence for music. Where pertinent evidence is available, music exhibits the essential properties of all adaptations. My goal in this chapter is not to attempt to prove that music is adaptive. (Like others, I am not at all convinced that music has evolutionary origins.) Rather, my goal here is to convince you that the question of music's origins remains open and warrants further investigation (see also several articles in Ref. 9).[10]

Of course there are a number of dangers attending evolutionary speculation.[11,12] Popper (1935/1959) pointed out that no scientist has yet formulated the theory of evolution by natural selection in such a way that a set of observations could, in principle, be used to falsify it.[13] Gould and Lewontin[14] have noted that evolutionary reasoning is plagued by *post hoc* reasoning. Evolutionary theory has been used to defend all sorts of nefarious ideologies, from racism to sexism. Philosophers note that evolutionary arguments often lead to the naturalist fallacy where what is is confused with what ought to be. In the case of music, there is an undistinguished history of polemical writing where certain kinds of music have been condemned for being 'unnatural'. Finally, by focusing on biological issues, one can leave the false impression that the effects of culture on music are minimal.

If music is an evolutionary adaptation, then it is likely to have a complex genesis. Any musical adaptation is likely to be built on several other adaptations that might be described as premusical or protomusical. Moreover, the nebulous rubric *music* may represent several adaptations, and these adaptations may involve complex coevolutionary patterns with culture (see Ref. 15). In biological matters, things are rarely straightforward.

Given these possible dangers, why bother attempting to formulate an evolutionary theory of music? Isn't it premature? First, as noted above, my goal here is not to convince you that music is adaptive; my goal is only to convince you that this is a worthwhile question. Understanding the possible origins of music might help inform us about some of the reasons we tend to respond in certain ways. Second, in the spirit of Popper, I will aim to tell an evolutionary story that is able to generate testable hypotheses. Like other evolutionary

accounts, my own theory will draw on existing knowledge, and so be *post hoc* in character. As long as this account remains *post hoc*, Gould and Lewontin's criticisms raise justified and paramount difficulties. However, it is my hope that the theory can be developed to the point where testable hypotheses might be derived.

Before entertaining some possible evolutionary views of music's origins, let us first consider two pertinent complicating points of views. One view is that music is a form of nonadaptive pleasure seeking (NAPS). A second view is that music is an evolutionary vestige.

## NAPS theory of music

Most pleasurable activities, such as eating and sex, have clear links to survival. Such activities ultimately stimulate brain mechanisms that are specifically evolved to reward and encourage adaptive behaviours. Note that once brain mechanisms are in place that permit the experience of pleasure, it may be possible to stimulate those mechanisms in ways that do not confer a survival advantage. We can call these behaviours NAPS. An example of NAPS behaviour is found in the human taste for sugars and fats. In premodern times, sugars and fats were rare in human diets, but highly nutritious in the amounts available. There are good reasons why human tastes would evolve to reward the ingestion of foods with high fat or sugar content. However, centuries of human ingenuity have succeeded in generating a modern diet that contains unnaturally high levels of fats and sugars—levels so high as to cause health problems such as diabetes and heart disease. Although such tastes originally conferred an increased chance of survival, in the modern environment, these behaviours have become less adaptive.

Another example of NAPS behaviour is found in drug use, such as heroin or cocaine. These drugs can directly activate the brain's pleasure centers, simply by injecting or imbibing a substance. Although the channel for pleasure exists for good evolutionary reasons, it may be possible to exploit the channel without any concomitant survival-enhancing result.

As in the case of drugs, it is possible that musical behaviours are forms of nonadaptive pleasure seeking. That is, music itself may not enhance human survival; music may merely exploit one or more existing pleasure channels that evolved to reinforce some other adaptive behaviour(s). We might call this view the 'NAPS theory of music'.

One way to determine whether some pleasure-seeking behaviour is adaptive or nonadaptive is to consider how long the behaviour has been around. In the long span of evolutionary history, nonadaptive pleasure-seeking behaviours tend to be short-lived. For example, heroin users tend to neglect their health and are known to have high mortality rates. Furthermore, heroin users make poor parents; they tend to neglect their offspring. Poor health and neglect of offspring are infallible ways of reducing the probability that one's genes will be present in a future gene pool. After many generations, natural selection will tend to militate against heroin use. Those individuals who are not disposed (for whatever reason) to use heroin, are much more likely to procreate and so pass along their aversion to the use of such drugs, provided that the aversive behaviour is somehow linked to a gene or genes.

The use of alcohol already suggests how NAPS behaviours can transform a gene pool. Although no gene has been identified, either for alcohol susceptibility or for alcohol tolerance, the responses of different human populations to alcohol show a suggestive

pattern. Large quantities of alcohol became possible only with the advent of agriculture. Modern European and Asian descendents of early agrarian cultures (such as originated in Mesopotamia) manage to deal with alcohol better than descendents of traditional hunter-gatherer societies, such as indigenous peoples in the Americas and in the arctic regions of Europe. Of course there are certain to be nongenetic factors influencing alcohol tolerance and abuse. However, alcohol researchers suspect that genetic factors are at work. Those people who have descended from traditional agricultural societies have a clear statistical advantage in dealing with the nonadaptive consequences of alcohol, and this would be expected if alcohol had been prevalent in these societies for thousands of years.

If music itself has no survival value (and merely exploits an existing pleasure channel), then any disposition towards musical behaviours would tend to worsen one's survival. Spending inordinate amounts of resources (such as time and money) on music would be expected to place music lovers at an evolutionary disadvantage. In other words, if the NAPS theory of music is true, then we would expect music appreciation to be correlated with marginal existence: as in the case of alcohol, music lovers would be disproportionally more likely to end up on 'skid row'.

In addition, if music is nonadaptive, then it must be the case that music is historically recent; otherwise music lovers would have become extinct some time ago. As we will see, the archaeological evidence indicates that music is very old—much older than agriculture—and this great antiquity is inconsistent with music originating as a nonadaptive pleasure-seeking behaviour. In short, there is little evidence that musical behaviours have been selected against. All of these suggest that there is little support for the NAPS theory of music.

## Music as an evolutionary vestige

Another possibility is that, although music at one time did indeed confer some survival values, it is now merely vestigial. Like the human appendix, at one time this 'organ' may have contributed directly to human survival, but now it is largely irrelevant—a piece of evolutionary litter. If this view is true, then we would have to ask What advantage did music once confer? and How have things changed so that music is no longer adaptive?

## Measuring the adaptive value of music

The adaptive value of some function is often evident in the individual survival costs arising from that function. For example, the larynx of newborn infants is anatomically arranged so that breathing and swallowing can happen at the same time. When the larynx enlarges, our physiological capacity for speech is purchased at the price of the danger of choking. In effect, one measure of the evolutionary advantage of speech is the mortality rate due to choking.

Similarly, an estimate of the evolutionary advantage conferred by music is to measure the amount of time people spend in musical behaviours. In the Atlas mountains of Morocco, full-time Jujuka mountain musicians are supported by the local villagers. That is, there is an entire caste of people whose principal productive activity is music making. A ready index of the importance of music in such a society may be the ratio of the number of musicians to the number of farmers and herders.

## Some evolutionary theories of music

Of the various proposals concerning a possible evolutionary origin for music, eight broad theories can be identified:

*Mate selection.* In the same way that some animals find colourful or ostentatious mates attractive, music making may have arisen as a courtship behaviour. For example, the ability to sing well might imply that the individual is in good health.

*Social cohesion.* Music might create or maintain social cohesion. It may contribute to group solidarity, promote altruism, and so increase the effectiveness of collective actions such as defending against a predator or attacking a rival clan.

*Group effort.* More narrowly, music might contribute to the coordination of group work, such as pulling a heavy object.

*Perceptual development.* Listening to music might provide a sort of 'exercise' for hearing. Music might somehow teach people to be more perceptive.

*Motor skill development.* Singing and other music-making activities might provide (or have provided) opportunities for refining motor skills. For example, singing might have been a necessary precursor to the development of speech.

*Conflict reduction.* In comparison with speech, music might reduce interpersonal conflict. Campfire talk may well lead to arguments and possible fights. Campfire singing might provide a safer social activity.

*Safe time passing.* In the same way that sleep can keep an animal out of harm's way, music might provide a benign form of time passing. Evolutionary biologists have noted, for example, that the amount of sleep an animal requires is proportional to the effectiveness of food gathering. Efficient hunters (such as lions) spend a great deal of time sleeping, whereas inefficient feeders (such as grazing animals) sleep relatively little. Sleep is thought to help keep an animal out of trouble. A lion is more apt to injure itself if it is engaged in unnecessary activities. As early humans became more effective at gathering food, music might have arisen as a harmless pastime. (Note, e.g. that humans sleep more than other primates.)

*Transgenerational communication.* Given the ubiquity of folk ballads and epics, music might have originated as a useful mnemonic conveyance for useful information. Music might have provided a comparatively good channel of communication over long periods of time.

### Sexual selection

Before continuing, we should take a moment to discuss a variant of the mate selection theory. Charles Darwin identified a form of natural selection known as sexual selection. The classic example of sexual selection is the peacock's tail. The function of the tail is not to promote the survival of the peacock; rather, the function is to promote the survival of the peacock's genes. Sexual selection arises once a particular genetic preference is established by the opposite sex—in this case, the preference of the peahen for flashy tails. Even if one peahen is not particularly impressed by Las Vegas-style tails, it remains to the female's benefit to mate with the most colourful male if her offspring are more likely to be desired by other females who are fond of colourful tails.

Darwin himself suggested that music might have arisen due to sexual selection in mating calls.[16] Like the peacock's tail, the preferences of hominid women could create an

escalating competition for ever more elaborate and beautiful melodies. Miller[17] has suggested that sexual selection accounts for why musical interests appear to peak in adolescence. However, the members of the all-male Vienna Philharmonic notwithstanding, there is nothing to indicate that one sex is more musical than the other, and so there is no evidence of the dimorphism commonly symptomatic of sexual selection. Women may be impressed by men who serenade them outside their balcony windows, but unlike female songbirds, female humans are perfectly capable of serenading men. Currently, there is no other known example of sexual selection that does not exhibit high sexual dimorphism—although one should note that biology is full of surprises.

## Types of evidence

In presenting a case for the evolutionary origins of music, we can consider four types of evidence:

*Genetic evidence.* The best evidence of an evolutionary origin would be the identification of genes whose expression leads to the behaviour in question. Unfortunately, it is rare for scientists to be able to link particular behaviours to specific genes. Although behaviour-linked genes have been discovered in other animals (such as fruit flies), no behaviour-linked gene has yet been conclusively established in humans. As in so many other areas, music has attracted a kind of folklore related to heritability. In some cultures, it is common for people to assume or believe that musical talent is partly inherited. More recently, work at the University of California, San Francisco by Baharloo *et al.*[18] appears to suggest a genetic component for absolute pitch.

*Neurological evidence.* The existence of specialized brain structures is neither a sufficient nor a necessary condition for music to be an evolutionary adaptation. Nevertheless, if stable anatomical brain structures exist for music, then this is consistent with music arising from innate development rather than being due solely to a generalized learning.

*Ethological evidence.* Are musical behaviours consistent with survival and the propagation of genes? In order for music to be an evolutionary adaptation, music-related behaviours must somehow increase the likelihood that the musical person's genes will be propagated.

*Archaeological evidence.* Since complex evolutionary adaptations arise over many thousands of generations, we must ask how widespread music is in biological history? If music originated in the past few thousand years, then it is highly unlikely to be an evolutionary adaptation. Evolution does not work that fast.

As noted, there is currently no evidence that links music to any gene. Let us consider the other areas of evidence in more detail.

## Archaeological evidence

Let us begin by considering some of the archaeological facts. The archaeological record shows a continuous record of music making in human settlements. Wherever you find evidence of human settlement, you find evidence of musical activities.

In 1995, paleontologist Ivan Turk discovered a bone flute while excavating an ancient burial mound in Divje Babe, Slovenia.[19,20] This flute has been determined to be between

43,000 and 82,000 years old, using electron spin dating. The instrument was fashioned from the femur of the now-extinct European bear.

Of course, finding this flute does not mean we have found the earliest musical instrument; this is just the earliest found instrument. As musical instruments go, flutes are rather complicated devices. If we look at contemporary hunter-gatherer societies, the most common instruments are rattles, shakers, and drums. For example, prior to the arrival of Europeans, by far the most common instruments in native American cultures were rattles and drums. The same pattern of preferred instruments is evident in African and Polynesian cultures. If we assume that rattles and drums typically predated the use of flutes, then the ancient music makers of Slovenia might well have been creating instrumental music somewhat earlier than 100,000 years ago.

What sort of music making, however, might have existed prior to the fashioning of musical instruments? It is not unreasonable to assume that singing preceded the making of musical instruments by some length of time. If we suppose that singing predated instrument making by 50 per cent of the intervening time, then music making might have existed 150,000 years ago—roughly twice the age of the older estimate for the Divje Babe flute. Even this figure might be a conservative estimate, and the actual origin of music might be twice as old, say around 250,000 years ago.

On the other hand, the Divje Babe flute might truly be an early specimen, and singing might have developed about the same time. Using the most recent estimate for the Divje Babe flute would therefore place the origins of music making about 50,000 years ago.

In summary, the archaeological record implies that music making likely originated between 50,000 years ago and a quarter of a million years ago. Although Wurlitzer organs, American Bandstand, and MTV are relatively recent phenomena, music making, in general, is really quite old. The evidence pointing to the great antiquity of music satisfies the most basic requirement for any evolutionary argument. Evolution proceeds at a very slow pace, so nearly all adaptations must be extremely old. Music making satisfies this condition.

Incidentally, the antiquity of music raises problems for those who would wish to use evolutionary arguments in esthetic debates by claiming that one music is 'more natural' than another. Whatever the origins of music, the vast majority of people have long ceased to live in Pleistocene conditions. In deciding whether Twisted Sister is better or worse than the Grateful Dead, appealing to the sounds of neolithic caves is unlikely to help.

## Anthropological evidence

Turning to contemporary anthropology, we can ask, What does the plethora of existing human cultures tell us about music? Without taking time to review the evidence, there is one overwhelming conclusion from the modern anthropological record. There is no human culture known in modern times that did not, or does not, engage in recognizably musical activities.

Not only is music making very old, it is ubiquitous; it is found wherever humans are found.[21] Moreover, I neglected earlier to mention one important fact about the bone flute at Divje Babe: the flute was found in a Neanderthal burial site. The Divje Babe flute isn't

even a human artifact. In short, it may be that music making is not just ubiquitous among *Homo sapiens*; music making may possibly be characteristic of the entire genus *Homo*.

The evidence pointing to the ubiquity of music satisfies another important basic requirement for any evolutionary argument. Relatively few adaptations are not found throughout the entire population of the affected species. For example, if eyelashes confer an evolutionary advantage, then just about everyone should have eyelashes. There are some exceptions to this principle, some of which are very important. For example, humans divide into female and male versions, so there are some genes that are not shared by everyone. Another more subtle example is the gene that codes for sickle-cell—a gene that protects against malaria, but that can also cause anemia. In general, however, adaptive genes are typically ubiquitous throughout a gene pool.

## Ethological evidence

When studying a particular animal, ethologists often begin by making an inventory of observed behaviours. What does the animal do, and how often does it do it? Activities that require a great deal of time and large expenditures of energy are understandably considered important. Ethologists assume that behaviours are likely to be optimized. Even behaviours that seem unimportant (such as infant play or sleeping) often have a serious or critical purpose.

Primates, for example, spend an extraordinary amount of time grooming each other. Ethologists feel obliged to formulate theories that account for the various proportions of resources dedicated by an animal to different activities.

Let us apply the ethological approach to the behaviours we call musical. For the purposes of illustration, we will consider two case descriptions. The first case is that of the Mekranoti Indians of the Brazilian Amazon, and the second case is that of contemporary society in the United States.

### The Mekranoti Indians

The Mekranoti Indians are primarily hunter-gatherers who live in the Amazon rain forest of Brazil. In Mekranoti culture, singing plays a prominent role in daily life. For several months of the year, every morning and evening the women lay banana leaves on the ground where they sit and sing for between one and two hours. The men sing early every morning, starting typically around 4:30, but sometimes as early as 1:30 AM. The men sing for roughly two hours each day, and often they will also sing for a half hour or so before sunset.

When singing, the Mekranoti men hold their arms in a sort of cradling position and swing their arms vigorously. The men endeavour to sing in their deepest bass voices and heavily accent the first beats of a pervasive quadruple meter with glottal stops that make their stomachs convulse in rhythm. Anthropologist Dennis Werner[22] describes their singing as a 'masculine roar'. When gathering in the middle of the night, the men are obviously sleepy, and some men will linger in their lean-tos well after the singing has started. These malingerers are often taunted with shouted insults.

Werner reports that 'Hounding the men still in their lean-tos [is] one of the favourite diversions of the singers. "Get out of bed! The Kreen Akrore Indians have already attacked

and you're still sleeping", they [shout] as loudly as they [can] . . . . Sometimes the harassment [is] personal as the singers [yell] out insults at specific men who rarely [show] up'.[23]

What is extraordinary about the Mekranoti singing is the amount of time involved—roughly two hours per day. (Remember, this is a subsistence society.) For the evolutionary ethologist, the important question arising from the Mekranoti Indians is why music making would attract so much of the tribe's resources. We will return to this question later.

### Modern United States of America

By way of comparison, consider now the prevalence of music in a modern industrialized society like the United States. For the ethologist looking at modern human behaviours, a crude though ready index of the amount of resources we dedicate to a particular activity can be found by measuring economic activity.

There is a widespread misconception that the foremost export sector in the US economy is 'high technology'. In fact, the preeminent export sector in the US economy is entertainment. Of the various component areas—films, sports, television, toys, and games—it is music that ranks foremost.

How big is the music industry? The music industry is comparable in size to the pharmaceutical industry. People spend more money on music than on prescription drugs. We purchase recordings, go to concerts, buy sheet music, take our children to music lessons, listen to commercial radio, watch film accompanied by music, and encounter Muzak in the local shopping mall. The most active concert venues in the world are freeways: a major preoccupation for millions of drivers is listening to music.

Of course financial measures are crude indicators of behavioural significance. The ethological point is simple. In both a hunter-gatherer society and a modern industrial society, we find humans dedicating a notable proportion of the available resources to music making and listening. Music may not be more important than sex; but it is arguably more expensive, and it is certainly more time consuming.

In order to put these behaviours in perspective, suppose you were a Martian anthropologist visiting Earth. There are many aspects of human behaviour that would have recognizable value. You would see people engaged in growing and preparing food, in raising and educating children, people involved in transportation, health, and governance; but even if Martian anthropologists had ears, I suspect they would be stumped by music.

If you are still not convinced that music attracts a peculiarly excessive proportion of human resources, consider another comparison. Think of how important food is to human well-being, of how tasty and enjoyable food is and can be. Now how many universities have departments of cuisine or nutrition, or departments of food sciences, or even departments of home economics? Now consider how many universities have departments of music. Why would music figure more prominently than food? To a visiting tourist from Mars, music sticks out; it is a remarkable and bizarre activity that earthlings do.

Of course, we must be careful in drawing any conclusions about adaptations based on observations of modern behaviours. If music making is an adaptive behaviour, then it must have arisen long ago in the environment of evolutionary adaptedness—namely, the Pleistocene period, when the vast majority of human evolution occurred.

## Ethology and evolution

Just because an animal spends a lot of time on certain activities does not mean that the activity represents an evolutionary adaptation. Ethologists must connect the behaviour to an explicit evolutionary account. That is, there must exist a plausible explanation of how the behaviour would be adaptive.

Before considering such a theory for music, let us examine a nonmusical example—an example that has a richer theoretical literature about its origins. Specifically, we will consider some of the evolutionary arguments that have been advanced to account for the origins of language.

## On the evolutionary origin of language

As in the case of music, views concerning the origins of language are necessarily speculative. Nevertheless, we can learn a great deal by considering some of the theories that have been advanced concerning its origin. Until recently, the principal view of language was that it facilitated complex collaborative activities such as coordinating actions during hunting. This account seems unlikely, first, because talking is a bad idea when tracking prey, and second, because men as a group display inferior language skills compared with women.

A number of anthropological psychologists have suggested that language (and even music) evolved as surrogates for social bonding.

### The grooming and gossip theory of language origins

The most empirically grounded of the recent theories of language origins is what might be called the 'grooming and gossip hypothesis'. Its principal advocate is Robin Dunbar.[24] The theory proposes the following logic.

Animals often live in groups for mutual protection against predators. In general, larger groups are more effective in detecting and warding off predators than smaller groups, but there are costs associated with maintaining a large group. One cost is that feeding must be much more intensive in a given area and so a larger group must travel greater distances in search of food. A second cost is that as group size increases, threats are more likely to arise from internal conflict within the group rather than from external predators. That is, there is a point where group size effectively minimizes predation, but at the cost of threats from members of the group itself. Nowhere is this more evident than in primates. As a consequence of internal threats, animals within the group begin to form alliances with one another. These alliances reduce the likelihood of conflict due to the threat of group retaliation.

In primates, the principal means by which alliances are formed and bonds maintained is through grooming. Grooming accounts for between 10 and 20 per cent of an individual's daytime activities.

There is good evidence to suggest that the principal purpose of grooming is to form alliances between individuals. First, grooming partners are much more likely to come to the defence of one another when threatened by another member of the group. Even more important evidence comes from relating the amount of time spent grooming to the size of the group. Different primate species have different typical group sizes. Gorillas, macaques,

chimpanzees, and bonobos, among other groups, tend to form groups that have different average sizes. Primatologists have measured the different amounts of time each species engages in grooming.[25]

A major discovery has been that there is a consistent relationship between group size and the amount of time spent grooming. As the group size increases, the average grooming time also increases. This is an unusual finding: there is no reason to suppose that animals in larger groups tend to get dirtier than animals in smaller groups, so the increase in grooming is unlikely to be related to cleanliness. Primatologists widely agree that the increase in grooming time for larger groups arises from the need to form more extensive networks of alliances. In a large group, an individual fares better by having a wider circle of friends, and the way to build primate friendships is through mutual grooming.

In the case of humans, the common 'group size' has been estimated at roughly 150 people. This is approximately the size of most rural villages in the world. This means that human groups are especially large when compared with other primates. As Dunbar has pointed out, 'If modern humans tried to use grooming as the sole means of reinforcing their social bonds, as other primates do, then the equation for monkeys and apes suggests we would have to devote around 40 per cent of our day in mutual mauling'.[26]

Dunbar has suggested that language evolved as an alternative to physical grooming. In effect, physical grooming was replaced by 'vocal grooming', whose purpose remains the formation and maintenance of friendships or alliances. Such vocal grooming has a distinct advantage over physical grooming: we can talk to several people simultaneously. This increases the number of people we can bond with at the same time.

Note, however, that even language has significant limitations for multiple concurrent social interaction. Dunbar[27] has noted that 'there appears to be a decisive upper limit of about four on the number of individuals who can be involved in a conversation'. When a fifth or sixth person joins a conversation there is a marked tendency for the group to subdivide into two or more concurrent conversations. It is only in hierarchical situations (such as in a formal lecture) where a single conversation can be maintained in a larger group.

All of this suggests that language is most useful in close interpersonal interactions, such as grooming, gossiping, courting, and conspiring. Note, however, that there are other activities that are of value to members of a social group that involve the entire group (or at least large segments) rather than groups of twos or threes. Chief among these group activities is defence. When under threat, uniform group action is indeed a mighty force, much more powerful than smaller groups of twos and threes.

## Music and social bonding

At this point, we might speculate how music might fit into this account. Let us assume, for the moment, that the hypothesis that language evolved as a surrogate for physical grooming is true, and that language thereby allowed humans to live in larger groups with their attendant complex social relations. We could certainly conceive of a similar function for music. In some ways, music provides several advantages over language. Singing is much louder than speaking, so singing may facilitate group interactions involving more than the four individuals posited as the upper limit for conversation.

This view of the possible origins for music was essentially proposed by Juan Roederer in 1984: '... the role of music in superstitious or sexual rites, religion, ideological proselytism, and military arousal clearly demonstrates the value of music as a means of establishing behavioural coherency in masses of people. In the distant past this could indeed have had an important survival value, as an increasingly complex human environment demanded coherent, collective actions on the part of groups of human society'.[28]

In light of later work by primatologists such as Dunbar, there appears to be merit in Roederer's hypothesis. Music might have originated as an adaptation for social bonding—more particularly, as a way of synchronizing the mood of many individuals in a larger group. That is, music helps to prepare the group to act in unison.

Perhaps a helpful image is to imagine the cackling of geese prior to their taking off. How is it that individual geese manage to synchronize their actions so that the entire flock takes flight more or less simultaneously? For anyone who has watched geese take off, there is a clear increase in the volume of cackling: more and more geese start honking. The general hubbub of honking geese is apt to raise the arousal levels of all geese in the vicinity. This increased arousal (which includes increased heart rate) would prepare the geese for a significant collective expenditure of energy.

## Music and social bonding—further evidence

It is this theory of music and social bonding that I believe holds the greatest promise as a plausible evolutionary origin for music. For the remainder of this article, I would like to review further phenomena that provide support for this hypothesis. The evidence is going to come from the following five sources: various mental disorders imply a strong link between sociability and musicality; child development implies a social role for music; brain structures related to music are linked to social and interpersonal functions; the most popular musical works imply social functioning; music modifies hormone production in groups of people.

### Complementary disorders: Williams syndrome and Asperger autism

Consider two mental disorders: Williams syndrome and Asperger-type autism. The principal feature of Williams syndrome is mental retardation. Williams syndrome is unique in that sufferers display three additional characteristics. One characteristic is high verbal abilities. Individuals suffering from Williams syndrome take a great interest in words. Their speech is fluent and peppered with a remarkably sophisticated vocabulary. In fact, when first encountering someone with Williams syndrome, the language fluency tends to mask the mental handicap.

In addition to high verbal abilities, Williams syndrome individuals also exhibit high sociability. They are gregarious and sociable. Coupled with the high verbal abilities, this makes Williams syndrome children a delight to work with. Finally, Williams syndrome children exhibit high musicality.

Daniel Levitin and Ursula Bellugi[29] have described the musical activities of Williams syndrome children at a summer camp in New York State. The children are remarkable. The entire camp is alive with music. Although the children have been shown to have no greater

musical aptitude than normals, they clearly relish both the musical activities and the social environment in which their musical enthusiasms can flower.

Now consider the case of Asperger-type autism. Autism is a general diagnosis that is applied to individuals who exhibit a diverse range of symptoms. For example, some autistics are hyper-musical whereas others exhibit complete amusia. Autism is characterized by a strong aversion to social interaction. Although most autism is associated with reduced mental functioning, mental retardation is not always evident. There are autistic individuals with normal and above-average intelligence as well. Autism is related to an emotional deficit—notably the failure to develop the so-called secondary or social emotions, including shame, pride, guilt, love, and empathy. For normal children, these secondary emotions typically appear by the age of about four.

Temple Grandin is a high functioning Asperger-type autistic who has become well known through her writings about her own condition. Concerning love, Grandin talks about her confusion in high school when reading Shakespeare's *Romeo and Juliet.* 'I never figured out what it was all about', said Grandin. In a trip through the Rocky Mountains with Oliver Sacks, Grandin remarked, 'The mountains are pretty, ... but they don't give me a special feeling, the feeling you seem to enjoy'. 'You get such joy out of the sunset', she said. 'I wish I did, too. I know it's beautiful, but I don't "get" it'.[30] Grandin's experience of music is similar. Although Grandin has perfect pitch and what she describes as a tenacious and accurate auditory memory, she finds music leaves her cold. She finds the sounds 'pretty', but in general, she just doesn't get it.[31] All the fuss about music leaves her mystified.

Grandin's own explanation is that not all of the 'emotional circuits' are connected. Sacks interprets the phenomenon as follows: 'An autistic person can have violent passions, intensely charged fixations and fascinations, or, like Temple [Grandin], an almost overwhelming tenderness and concern in certain areas. In autism, it is not affect in general that is faulty but affect in relation to complex human experiences, social ones predominantly, but perhaps allied ones—esthetic, poetic, symbolic, etc. No one, indeed, brings this out more clearly than Temple herself... She feels that there is something mechanical about her mind, and she often compares it to a computer ... She feels that there are usually genetic determinants in autism; she suspects that her own father, who was remote, pedantic, and socially inept, had Asperger's—or, at least, autistic traits—and that such traits occur with significant frequency in the parents and grandparents of autistic children'.[32]

The contrast between Asperger-type autism and Williams syndrome is striking. On the one hand we have a group of people whose symptoms include high sociability linked with high musicality. On the other hand we have a group of people whose symptoms include low sociability often linked with low musicality. Together, these mental conditions are consistent with a relationship between sociability and musicality—and this link is the principal assumption of a group-oriented evolutionary account.

## Music and social function

Consider the following question: What is the most successful piece of music in modern history? Of course the answer to this question depends on how we define success—and this is far from clear, as esthetic philosophers have shown. Nevertheless, I want to use a

straightforward criterion: let us assume that the most successful musical work is the one that is most performed and most heard. Using this criterion, you might be surprised by the answer. The most successful musical work was composed by Mildred and Patti Hill in 1893, and revised in the 1930s.[33] The piece in question is, of course, 'Happy Birthday'. 'Happy Birthday' has been translated into innumerable languages and is performed on the order of a million times a day. It remained under copyright protection until the middle of the twentieth century. For many people, the singing of 'Happy Birthday' is the only time they sing in public. For other people, the singing of 'Happy Birthday' constitutes the only time they sing.

In some ways, 'Happy Birthday' is the quintessential feminist work. Its composers remain unknown and uncelebrated; the work was created by the collaboration of two women rather than as an egotistic expression of one man. It is a thoroughly domestic work: 'Happy Birthday' is performed in the kitchen or lunch room rather than in the concert hall. No other musical work has evoked so much spontaneous music making. The work is domestic, amateur, and relationally oriented. Despite its extraordinary success, it remains undervalued as a musical creation.

'Happy Birthday' plays a role in our evolutionary story because I suspect that for the vast majority of human history, music making was of this ilk. In Western culture, it is surely the camp songs sung by Girl Scouts or the songs sung by British soccer fans that come closest to what might be imagined in Pleistocene *Homo sapiens*. In all these cases, the music serves an obvious social role and in helping to define a sense of identity and common purpose.

In light of our evolutionary hypothesis, let us return and reconsider the singing of the Mekranoti Indians. Recall some of the characteristic features—especially the singing done by the men: the men's singing is done late at night and in the early morning, and their singing is associated with a high degree of machismo. Like most native societies, the greatest danger facing the Mekranoti Indians is the possibility of being attacked by another human group. The best strategic time to attack is in the very early morning while people are asleep. Recall the insult shouted at men who continued to sleep in their lean-tos: 'Get out of bed! The Kreen Akrore Indians have already attacked and you're still sleeping'. The implication is obvious. It appears that the nightly singing by the men constitutes a defensive vigil. The singing maintains arousal levels and keeps the men awake.

Of course music making is also associated with stirring a war party. North American Indians famously sang and danced prior to initiating an attack on another tribe. One might suppose that engaging in an activity that publicly announces a hostile intention would be counterproductive. War dances might possibly warn an enemy of an impending attack. However, the music making seems to serve a more important role: that of raising arousal and synchronizing individual moods to serve the larger goal of the group.

## Social bonding and hormones

Apart from arousing individuals, music can also pacify. Experimental work by Fukui has shown that listening to music can reduce testosterone levels. Fukui himself was quick to point out the possible social and evolutionary significance of this finding. In human social groups, lower levels of testosterone are likely to result in less aggression, less conflict, less sexual confrontation or sexual competition, and consequently more group cohesiveness.

A problem with Fukui's experiment is that he did not manipulate the type of music heard by his listeners. Listeners simply listened to their favourite music. Depending on his sample of listeners, we might expect whole genres of music were not represented. We might suppose, for example, that heavy metal, hard rock, or thrash music might well have increased rather than decreased testosterone levels. Further research is necessary to document the specific hormonal changes associated with different types of musical experiences. However, Fukui's work at least shows that music can have marked effects on hormone levels—specifically, hormones that relate especially strongly to sociability.

## Oxytocin and the biology of social bonding

An important question to ask is how precisely music might bring about social bonding. Neurophysiologist Walter Freeman[34] has proposed a pertinent theory related to the hormone oxytocin.

Oxytocin is most commonly associated with the 'let-down' response in new mothers—that is, the response that enables the flow of breast milk following childbirth. The presence of oxytocin also has dramatic effects on the brain. For example, when a ewe gives birth to a lamb, the olfactory bulb in the ewe's brain is bathed in oxytocin. Following the birth of the new lamb, a ewe will imprint on the smell of the new lamb, but will subsequently fail to recognize the smell of her former offspring. The result is that the ewe will suckle only the newborn lamb.

Neurophysiological research has shown that oxytocin acts as a sort of 'eraser' that wipes away previous memories and simultaneously facilitates the storage of new memories. When linked with significant life events, oxytocin is the cement that binds new memories. The amnesic properties of oxytocin are evident in all kinds of learning episodes. However, their strongest effects occur during major limbic activations, such as those resulting from trauma or from ecstasy. Pavlov discovered this phenomenon when serious spring flooding affected his lab and nearly drowned his caged dogs. Following their rescue it was discovered that the dogs had to be retrained from scratch.[35]

In his book, *Societies of Brains*, Freeman chronicles a number of circumstances where oxytocin release occurs and the effects of these releases on neural organization. As we have noted, oxytocin releases are associated with trauma and ecstasy. In addition to childbirth, oxytocin is released in males and females following sexual orgasm. Freeman also suggests that oxytocin is released during trance and while listening to music.

In many cases, the presence of oxytocin is correlated with human and animal bonding circumstances. For example, in the case of sexual orgasm, oxytocin may significantly facilitate pair bonding in the same way that oxytocin following childbirth facilitates mother–child bonding. Freeman's suggestion that music causes oxytocin to be released has important repercussions for instances of peer-group bonding and social identity. If Freeman is correct, there would be good neurophysiological reasons for lovers to enjoy music while courting, for union members to sing while on the picket line, for religious groups to engage in collective music making, for colleges to promote alma mater songs, and for warriors to sing and dance prior to fighting.

## Mood regulation

Thayer and his colleagues have carried out a number of studies concerning how people regulate their moods. One study attempted to determine what people do to try to get out of a bad mood. Of 29 categories of activities, the foremost activity was calling or talking to a friend. The second most frequently reported activity was trying to think positive thoughts—to give oneself a sort of 'pep talk'. The third most frequently reported activity—ahead of a wide variety of behaviours—was listening to music. Forty-seven per cent of respondents reported that they used music to temper or eliminate a bad mood.[36]

Thayer *et al.*[37] carried out a similar study to determine what people do to raise their alertness or energy level. Listening to music was reported by 41 per cent of respondents, following activities such as sleeping, taking a shower, getting some fresh air, and drinking coffee. Finally, in a third study investigating what people do to reduce nervousness, tension, or anxiety, listening to music ranked third at 53 per cent, following after only calling or talking to someone, and trying to calm down by thinking about a situation.

There are two points to highlight from these studies. The first is that the foremost category of behaviour for mood regulation is being with or conversing with a friend. That is to say, our first tendency is to seek mood regulation through social interaction. Moods are contagious, and we rely to some extent on each other to modulate, reinforce, or temper our moods. Although we know that moods are highly influenced by the individual's physiological state—notably through food, exercise, or rest—behaviours such as eating, exercise, and rest are less frequently used for mood regulation than music.

The second point to highlight is the obvious point that music appears to figure prominently as a method for mood regulation. Although in contemporary society music tends to be experienced in a personalized or individualized listening context, we already know that this context is historically unprecedented. Most music making in hunter-gatherer societies occurs in a social or group context. Until the invention of the phonograph, the vast majority of music in Western culture was also experienced in social or group contexts. In short, music is not out of place in the list of socialized behaviours used for mood regulation.

## Conclusion

By way of conclusion, first let me reiterate that I do not think the evidence in support of music as an evolutionary adaptation is strong. The purpose of this article has been to show that there are no obvious or fatal impediments that rule out a possible evolutionary origin.

We might summarize the basic evidence as follows:

1. Complex evolutionary adaptations arise only over many millennia. Accordingly, in order for a behaviour to be adaptive, it must be very old. As we have seen, music making does indeed conform to the criterion of great antiquity.
2. Behavioural specializations are often expected to be associated with specific anatomical or functional brain structures. Lesions and other neurological assaults can leave an individual with impaired musical functioning. There are double-dissociations between

various amusias and virtually every other kind of functional mental loss. This does not prove that music is not acquired by general learning, but the neurological evidence is at least consistent with the possibility that there are specialized music-related brain structures.

3. In order for a behaviour to be adaptive, the behaviour itself must enhance the propagation of the individual's genes. As we have seen, musical behaviours are consistent with mood modification and group mood synchronization—and these synchronous states are at times clearly associated with situations where group efforts are adaptive—such as in the case of defence against other human groups. In addition, high musical involvement is not associated with dereliction or poor survival (such as the case for alcohol); this raises problems for the view that music is a form of nonadaptive pleasure seeking.

The evidence we have for mood regulation and synchronization is suggestive:

1. We have noted contrasting disorders in Williams syndrome and Asperger-type autism. In one case, we see a group of individuals who are highly sociable and also highly musical. In the other case, we see some individuals who display extremely low sociability and also low musical understanding or affinity.
2. Although we did not review this literature, the emergence of the secondary or socialized emotions in child development is strongly associated with musical empathy, understanding, and sophistication. The pertinent research on child development implies a social role for music.
3. We noted that the most popular musical works often imply some sort of social function. 'Happy Birthday' is only one example. Group identity is often expressed through, for example, folk songs, Girl Scouts' camp songs, sports, and war dances.
4. Although we did not review the literature, it is also known that the emergence of musical tastes relates to postpubescent socializing and group identity.
5. Finally, we discussed how music modifies hormone production in groups of people.

As noted at the beginning of this essay, there is a long history of abuse of genetic claims serving ulterior and often nefarious motives. Even if we assume that musicality has some adaptive function, the repercussions for modern music making and modern musical enjoyment are likely to be minimal. Music is now deeply embedded in a cultural/historical context where human musical memories span centuries, and the fashion cycle is a significant engine of change. Music is now part of a Lamarckian system where acquired characteristics are transmitted in Dawkinsean 'meme pool' rather than in Mendelian 'gene pool'.[38] Like language, the details of musical culture and tastes are largely a product of enculturation.

Nevertheless, it remains worthwhile to attempt to understand where music comes from and why it has achieved such a ubiquitous presence in human lives. Evolutionary theorizing about music may well remain in the realm of *Just-So* stories. There is always the possibility, however, of a testable hypothesis emerging, and if so, we will all wait with interest to see the results.

## Acknowledgements

I would like to extend my thanks to Dr Kristin Precoda for drawing my attention to the work of Werner regarding the Mekranoti Indians, and to Dr David Wessel for drawing my attention to Freeman's work concerning oxytocin. This lecture was originally presented at the Department of Music, University of California, Santa Barbara, 6 March 1998. A shortened version was presented at the Society for Music Perception and Cognition Conference, Evanston, Illinois, 16 August 1999.

## References

1. Barkow, J., L. Cosmides, and J. Tooby (eds) (1992) *The Adapted Mind: Evolutionary Psychology and the Generation of Culture.* Oxford: Oxford University Press.
2. Tooby, J. and L. Cosmides (1992) The psychological foundations of culture. In J. Barkow, L. Cosmides, and J. Tooby (eds). *The Adapted Mind: Evolutionary Psychology and the Generation of Culture.* Oxford: Oxford University Press, pp. 19–136.
3. Baron-Cohen, S. (ed.) (1997) *The Maladapted Mind: Classic Readings in Evolutionary Psychopathology.* East Sussex, UK: Psychology Press.
4. Shepard, R. N. (1992) The perceptual organization of colors: an adaptation to regularities of the terrestrial world? In J. Barkow, L. Cosmides, and J. Tooby, (eds) *The Adapted Mind: Evolutionary Psychology and the Generation of Culture.* Oxford: Oxford University Press, pp. 495–532.
5. Wright, R. (1994) *The Moral Animal: The New Science of Evolutionary Psychology.* New York: Vintage Books.
6. Dissanayake, E. (1988) *What Is Art For?* Seattle: University of Washington Press.
7. Pinker, S. (1997) *How the Mind Works.* New York: W.W. Norton.
8. Pinker, S. (1994) *The Language Instinct.* New York: Morrow.
9. Wallin, N. L., B. Merker, and S. Brown (2000) *The Origins of Music.* Cambridge, MA: MIT Press.
10. Williams, L. (1980) *The Dancing Chimpanzee: A Study of the Origin of Music in Relation to the Vocalising and Rhythmic Action of Apes.* Revised edition London: Allison and Busby.
11. Lewontin, R. C. (1991) *Biology as Ideology; The Doctrine of DNA.* Concord, Ontario: House of Anansi Press.
12. Symons, D. (1992) On the use and misuse of Darwinism in the study of human behaviour. In J. Barkow, L. Cosmides, and J. Tooby (eds) *The Adapted Mind: Evolutionary Psychology and the Generation of Culture.* Oxford: Oxford University Press, pp. 137–59.
13. Popper, K. (1935/1959) *Logik der Forschung.* Vienna, 1935. Translated as The Logic of Scientific Discovery. New York: Basic Books.
14. Gould, S. J. and R. C. Lewontin (1979) The spandrels of San Marco and the Panglossian program: a critique of the adaptationist programme. *Proc. R. Soc. Lond.* 250, 281–8.
15. Durham, W. (1991) *Coevolution: Genes, Culture, and Human Diversity.* Stanford, CA: Stanford University Press.
16. Darwin, C. (1872) *The Expression of Emotion in Man and Animals.* London: Murray.
17. Miller, G. (2000) Evolution of human music through sexual selection. In N. L. Wallin, B. Merker, and S. Brown (eds) *The Origins of Music.* Cambridge, MA: MIT Press, pp. 329–60.

18. **Baharloo, S., P. A. Johnston, S. K. Service, *et al.*** (1998) Absolute pitch: an approach for identification of genetic and nongenetic components. *Am. J. Hum. Genet.* 62, 224–31.
19. **Anon** (1997) Neanderthal notes: did ancient humans play modern scales? *Sci. Am.* 277, 28–30.
20. **Turk, I.** (ed.) (1997) *Mousterian "Bone Flute" and Other Finds from Divje Babe I Cave Site in Slovenia.* Ljubljana: Zalozba ZRC.
21. **Brown, D.** (1991) *Human Universals.* New York: McGraw-Hill.
22. **Werner, D.** (1984) *Amazon Journey; An Anthropologist's Year among Brazil's Mekranoti Indians.* New York: Simon and Schuster.
23. **Werner, D.** (1984) *Amazon Journey; An Anthropologist's Year among Brazil's Mekranoti Indians.* New York: Simon and Schuster, pp. 245–47.
24. **Dunbar, R.** (1997) *Grooming, Gossip and the Evolution of Language.* New York: Faber & Faber.
25. **Aiello, L. and R. I. M. Dunbar** (1993) Neocortex size, group size and the evolution of language. *Curr. Anthropol.* 34, 184–93.
26. **Dunbar, R.** (1997) *Grooming, Gossip and the Evolution of Language.* New York: Faber & Faber, p. 78.
27. **Dunbar, R.** (1997) *Grooming, Gossip and the Evolution of Language.* New York: Faber & Faber, p. 121.
28. **Roederer, J.** (1984) The search for a survival value of music. *Music Percept.* 1, 350–6.
29. **Levitin, D. L. and U. Bellugi** (1997) *Musical abilities in individuals with Williams Syndrome.* Paper presented at the 1997 Society for Music Perception and Cognition. Cambridge, MA: Massachusetts Institute of Technology.
30. **Sacks, O.** (1996) *An Anthropologist on Mars.* New York: Random House, p. 124.
31. **Sacks, O.** (1996) *An Anthropologist on Mars.* New York: Random House, p. 122.
32. **Sacks, O.** (1996) *An Anthropologist on Mars.* New York: Random House, p. 123.
33. **Fuld, J. J.** (1995) *The Book of World-famous Music; Classical, Popular and Folk,* 4th edn. New York: Dover.
34. **Freeman, W. J.** (1995) *Societies of Brains: A Study in the Neuroscience of Love and Hate.* Hillsdale, NJ: Lawrence Erlbaum Associates.
35. **Pavlov, I. P.** (1955) *Selected Works.* Translated by S. Belsky, edited by J. Gibbons. Moscow: Foreign Languages Publishing House.
36. **Thayer, R. E.** (1996) *The Origins of Everyday Moods.* New York: Oxford University Press.
37. **Thayer, R. E., J. R. Newman, and T. M. McLain** (1994) The self-regulation of mood: strategies for changing a bad mood, raising energy, and reducing tension. *J. Pers. Soc. Psychol.* 67, 910–25.
38. **Dawkins, R.** (1976) *The Selfish Gene.* Oxford: Oxford University Press.

PART II

# THE MUSICAL MIND

CHAPTER 6

# THE ROOTS OF MUSICAL VARIATION IN PERCEPTUAL SIMILARITY AND INVARIANCE

STEPHEN McADAMS AND DANIEL MATZKIN

## Abstract

Perceptual similarity underlies a number of important psychological properties of musical materials including perceptual invariance under transformation, categorization, recognition, and the sense of familiarity. Mental processes involved in the perception of musical similarity may be an integral part of the functional logic of music composition and thus underly important aspects of musical experience. How much and in what ways can musical materials be varied and still be considered as perceptually related or as belonging to the same category? The notions of musical material, musical variation, perceptual similarity and invariance, and form-bearing dimensions are considered in this light. Recent work on similarity perception has demonstrated that the transformation space for a given musical material is limited by several factors ranging from degree of match of the values of auditory attributes of the events composing the sequences, to their relations at various levels of abstraction, and to the degree that the transformation respects the grammar of the musical system within which the material was composed. These notions and results are considered in the light of future directions of research, particularly concerning the role of similarity and invariance in the understanding of musical form during listening.

**Keywords:** Similarity in music; Invariance in music; Musical variation; Perception of music

## Introduction

Perceptual similarity underlies a number of important psychological properties of musical materials including perceptual invariance under transformation, categorization, recognition, and the sense of familiarity. Mental processes involved in the perception of musical similarity may be an integral part of the functional logic of music composition and may thus underly important aspects of musical experience. We are interested in the perception of musical materials and of their musical transformations, that is, 'themes' and 'variations' but in a much larger sense than is usually attached to this specific form in Western tonal music.

A longer term goal of this approach is to understand the role of the perception of similarity and invariance, as well as that of the change produced by variations of the

original materials in the experience of musical form. We hesitate to say 'perception' of musical form since most musical forms are too long to be 'perceived' as such. By understanding the mental processing underlying similarity perception, it may be possible to explain in part how listeners form hierarchical and associative relations among features of a musical piece belonging to a given musical system and thereby explain in some musical situations how they manage or fail to establish these relations. How much and in what ways can musical materials be varied and still be considered as perceptually related or as belonging to the same category? Can there be a repertoire of possible transformational rules that the cognitive system can decipher, allowing listeners to perceive a relation between a variation and a theme and perhaps even the nature of the transformation process itself?

We will first of all consider in a general way the notions of musical material, musical variation, perceptual similarity, and perceptual invariance, and then what these all suggest for the development of the notion of form-bearing dimensions in music as originally proposed in McAdams.[1] Subsequently, preliminary experiments on musical similarity that we have conducted[2,3] will be summarized. In this work, pitch and duration variations are studied for both tonal/metric and nontonal/nonmetric musical systems. A number of theoretical issues will finally be developed in the hopes of stimulating and orienting further work in this area, notably concerning: (1) the nature of the representation of musical materials and transformation processes, (2) modularity and dimensional interactions within the realm of music cognition, (3) parallelism and associative structures in music, and (4) the notion of musical development and musical process.

## Musical materials

We take the notion of musical material in a very broad way. Materials may be more conventionally considered as simple figures or themes such as the opening 5-note figure of Anton Webern's *Sechs Stücke für großes Orchester*, op. 6 (1909) (only slightly more complex than Beethoven's famous ta-ta-ta-tum, although not as explicitly developed) or slightly more lengthy themes such as a Bach fugue subject. They may also be more fully developed musical ideas such as the thematic materials used by Roger Reynolds in pieces such as *Archipelago* for chamber orchestra and computer-processed sound (1982–3) or *The Behavior of Mirrors* for guitar (1986). He calls these materials core elements.[4] They may vary in duration anywhere from several seconds to several tens of seconds and any given piece will have several of them, each designed to have distinctive characteristics in terms of pitch materials, durational proportions, textures, and gestural movements. Musical materials can also be complex textures: the dense blocks of sound used emblematically by Krzysztof Penderecki in *Threnody for the Victims of Hiroshima* for 52 strings (1960), the architectonic gestures often used by Iannis Xenakis in his orchestral works such as *Pithoprakta* (1956), or the rich micromelodic tapestries in early orchestral and choral works by György Ligeti such as *Atmosphères* (1961) and *Lux Æterna* (1966). And finally, we should not leave out the realm of electroacoustic music in which spectral and temporal sound structures can be imagined from scratch and become the basis for compositional development or in which recorded sounds can become musical materials in their own right as the composer delves into

and transforms the inner structure of the sound itself. Notable examples of the latter class are the deep church bell tone and the boy soprano's voice that are the two concrete sound sources from which the computer-generated piece *Mortuos Plango, Vivos Voco* (1980) was composed by Jonathan Harvey or the interpolations between unlikely partners such as a horse's whinny and a baby's cry in *Vox 5* (1986) by Trevor Wishart.

## Musical variation

Given the variety of types of musical materials, there is obviously an infinity of ways that even simple ones can be varied in order to generate new material to fill the needs of musical discourse, all the while maintaining (most often) some kind of perceptual link to the original material so that the listener not only recognizes the relation (implicitly or explicitly), but also senses the kind of variation that has been applied. This trajectory of change is one of the perceptual components of musical development, and our intuition is that it contributes to the sense of larger-scale movement and cohesion in a work. In the simplest (operational) terms, musical variation or transformation may be taken to be simply the change along one or more musical dimensions of a given musical material.

In most musics, the principal dimensions targeted for transformation of materials are pitch and duration, although space and timbre have also become the focus of musical structuring in the twentieth century. A complete catalogue of transformation devices is of course impossible, so only a few clearly definable types will be mentioned below. The simplest kind of (non)variation is the repetition of the same material at a different point in time. In this case there is absolute identity in a physical sense and the perception of similarity and the recognition of the return of the material depend only on the listener and the intervening materials.

Another version would keep the pitch and rhythmic values constant and change the timbre through orchestration. A blatantly clear example of this approach is the piece *Boléro* for symphonic orchestra (1928) by Maurice Ravel in which two themes are alternated AABB four times and then AB at the end. However each time, the A and B melodies and their accompaniment are orchestrated differently. The similarity and sense of repetition are strong and it is the pattern of change in orchestration, dynamics, and articulation that create the expressive trajectory of the work. Similarly, in *The Angel of Death* for solo piano, chamber orchestra and computer-processed sound (2001) by Roger Reynolds, the five thematic elements are composed both for solo piano and for the 16-instrument chamber orchestra.[5] The piece is organized in two halves, in both of which the central 'core identities' of the five themes appear. However, if one core appears in the piano in one half, it returns in nearly identical form as concerns pitch and rhythm in the other half, but in the orchestra. Other core elements have the reverse relationship. Bigand *et al.*[6] have shown that for these contemporary musical materials musicians are more perturbed in a recognition task by the instrumentation change than are nonmusicians, although both groups perform in equivalent manner for tonal materials. Further, for materials from the piece by Reynolds, the similarity relations among excerpts are affected by changes in orchestration for some materials and are not affected for others,[7] demonstrating the complexity of perception and memory processes implicated in orchestration practice.

A more complex approach would be to change the timbre on each note, a true *Klangfarbenmelodie* as proposed by Schoenberg[8] and then have different timbral compositions for each new return of the theme. The danger in this case, as will be discussed below is that the melody may become fragmented perceptually if the timbre change is too drastic, bringing about the assignment of different subsets of notes to different auditory streams and thus compromising the psychological coherence of the melody as a whole.[9]

Transposition of pitch and change of tempo are the simplest kinds of transformations and testify to the fact that the representation of relations between events (exact intervals or number of scale steps related to a given scale) are prominent in musical memory. The perception of similarity of transposed melodies is indicative of the influence of internalized musical systems on recognition and similarity perception. An exactly transposed melody (one that maintains the same pitch intervals between successive notes) will in many cases be perceived as a change in key. A melody that is transposed, keeping the same number of scale steps between successive notes, but with pitches that remain within the current key is often perceived as more similar as suggested by recognition studies, although this ability depends on musical training, musical context, and task instructions.[10] At any rate, what seems to be the important aspect of the perceptual representation underlying melodic and rhythmic similarity here is the relations among notes, that is, the relative distance in terms of intervals or scale-steps. Note that the slight interval changes generally preserve the contour (pattern of ups and downs), a component of melody representation that has been shown to be particularly cogent for the similarity of unfamiliar melodies as measured by false recognitions.[11]

Some studies have shown that the perception of timbral intervals is also possible.[12,13] This work was based on the notion of timbre space in which the perceptual relations among sounds equated for pitch, duration, and loudness are represented mathematically as a multidimensional space. Timbral intervals are defined as oriented vectors in the space and timbral contours could be defined by the relative patterns of up and down along the various dimensions. A transformation of a timbre melody such as transposition would simply be the translation of the timbre vectors in the space, keeping their lengths and orientations constant. Although the precise reproduction of timbral intervals is problematic with instrumental music, such musical structures are quite possible with synthesized sounds.

A variant on the simple translation operation is the inversion operation in which the sign of the intervals and contour directions are changed. This technique is used in mirror canon writing and is also quite prominent in serial music in the twentieth century.

Another class of transformations involves temporal rearrangement of the musical material. For example, playing a melody backward is called retrogradation and is often found in music of the Common Practice, romantic, and contemporary periods of Western music. Mirror and cancrizans canons use this technique and it also became a prominent operation in serial music. A particularly intriguing compositional problem-solving task involves creating a melody that can be played forwards and backwards (the cancrizans or 'crab' canon). Often the material that is retrograded may not reproduce the exact interval structure for reasons related to the musical syntax. In such cases, certain prominent intervals or the inverted melodic contour may preserve the sense of similiarity. Take, for example, the opening five-note figure of Anton Webern's *Sechs Stücke*, op. 6 mentioned above in which

the flute plays C#4-E4-F4-G4 followed by a D5 on the muted trumpet. Subsequently the sustained trumpet tone becomes the first note of the retrograded pattern D5 (trumpet), Bb4-Eb4-Db4 (flute), B3 (French horn). Although the intervals are slightly different, the pitch contour (small skip, two steps, large skip) is inverted, as is a timbral contour, moving from the low register flute to a bright muted trumpet in one direction and the flute to a lower darker French horn in the other. The mirror image is quite convincing perceptually. Various other techniques of recombination have been used as well, often involving the choice of particularly salient figures from the theme and their subsequent development, a practice quite prominent in Beethoven piano sonatas, for example.

All of the preceding examples are based on rearrangements of symbolically notated events. However, in electroacoustic music, where the material is a continuous sound waveform, the operations need not be limited to predefined events. Granular synthesis, for example, takes bits of recorded sound from different parts of the original sample, applies a smoothing envelope to them and then recombines them in ingenious ways to create rich sound textures.[14] The larger the grains, the more the original sound quality 'comes through' in the recombination. Roger Reynolds has used a technique involving segmentation and rearrangement in several pieces, applied either to sound samples (as in *Archipelago*, 1982–3) or to note lists representing symbolic events (as in *Variation* for piano, 1988).[4] He has developed two classes of algorithms (SPLITZ and SPIRLZ) that transform the original samples within a space of variation constrained by the variables of the algorithm. The first one segments the sound sequence into units of different durations and then plays the odd ones in one order and the even ones in reverse order. The spacing and timing of the segmented unfolding can be controlled, as can the amplitude envelopes applied to the segments (usually different for even and odd ones). This gives a kind of alternation between the front and back ends of the sound sequence that move towards the centre and then back to the extremities. The technique is obvious when applied to musical gestures with a clear trajectory, but can give stunning results in other cases even though the listener may not be able to follow explicitly the algorithm's logic. As for granular synthesis, the transformed sequences may have a more or less strong similarity to the original material depending on the size and ordering of the segments. When applied within a piece to several different themes, it may be that the transformation device itself takes on an identity across materials that is stronger than the relations between the original and transformed materials in some cases. The perception of these kinds of materials is currently under study in our laboratory in collaboration with the composer, particularly as concerns interactions in memory between original and transformed materials.[15]

Staying within the realm of electroacoustic music, there are a plethora of digital techniques now that allow an extremely fine-grained analysis of a sound's spectral and temporal structure, and then either a direct modification of certain of these parameters, or even a modelling of their behaviour and a subsequent resynthesis with changes in the model parameters. This kind of sonic elaboration is a class of transformation that remains quite close to the timbre and sound source identity of the materials to be transformed. Compelling examples of this approach can be found in Trevor Wishart's *Vox 5* (1986), Jean-Claude Risset's *Sud* (1985), and Jonathan Harvey's *Ritual Melodies* (1990). In particular, in the Harvey piece, models of various kinds of non-Western instruments and vocal styles are used to

create interpolations between the sound universes of each, including not only the sound qualities themselves, but also the styles of playing (forms of vibrato, tremolo, portamento, etc.).[16] Harvey found that without the expressive performance gestures, many of the musical phrases for a given simulated instrument or voice were unconvincing, and the zone in which the heard sound source was between the two originals was musically untenable.

By far the most widely used type of variation involves change in the surface level structure while some invariant is retained at a higher hierarchical level. This kind of technique is quite common in theme-and-variations forms in music of the Common Practice period. The notions of elaboration and reduction have been formalized to a large extent by Lerdahl and Jackendoff,[17] although their applicability to contemporary music as proposed by Lerdahl[18] has been questioned. Indeed Dibben[19] suggests that since structurally important events that 'stand for' sets of subordinate events in tonal music, do not have the same kind of representation in atonal music, it may be necessary to consider more associational structures to explain formal coherence in this latter kind of music as will be discussed at the end of this chapter. Evidence for similarity among musical materials related by a common reduced structure in tonal/metric music is clear, for example,[20–22] and will not be belaboured here. However, notions of elaboration also exist in contemporary music, although it is not clear to what extent commonalities of mental processing exist between the different musical systems.

Some of the psychological questions that arise from this brief and extremely schematic consideration of the possibilities of musical variation may include the following: How far can one transform an original musical material before it becomes completely new, and perceptually unrelated, material, that is, what are the bounds of perceptual similarity? What does the answer to this question tell us about the representations of musical materials and the nature of the psychologically realistic processes of transformation that can operate on them?

## Similarity and invariance

At the crux of the issue of musical variation are the notions of perceptual similarity and invariance, provided of course that the varied material appears to have at least some relation to the original material. What we actually mean by 'similarity' is not that easy to pin down operationally. Let us just assume that it is some degree of perceptual match between the properties and features of two materials. The experimental question is thus to understand what properties contribute to perceived similarity and how the degree of match between these (to-be-determined) properties affects its strength. We will speculate on the nature of various kinds of properties and levels of abstraction that might contribute to musical similarity as judged directly by listeners.

Properties that contribute to similarity may be of various levels of abstraction, moving from the most concrete (specific values of attributes of individual events in the sequence, such as melodic and rhythmic configurations) to the most abstract (higher-level reductions of the hierarchical event structure, if such exists for the material being heard, or harmonic contours that provide a kind of skeleton for a tonal work).

Concerning surface values, the surface content of a material and of its transformation (specific pitches, durations, loudnesses, timbres, etc.) may have more or less events in

common along the different auditory dimensions. Also the temporal order of these events may be the same or may be transformed to a greater or lesser extent. Exact repetition would of course have the highest similarity and completely different values would have no similarity. However, under some situations the exact same number of events with specific attributes in completely different orders (a random rearrangement of the events, e.g.), may still be experienced as similar due to identical statistical distributions of material properties. Krumhansl[23] has shown, for example, that listeners are sensitive to the combinations of pitch and duration in Olivier Messiaen's *Modes de Valeurs et d'Intensités* for piano (1950). After having heard half of the piece, they were asked to rate new material according to the degree with which it corresponded to that already heard. The specific melodic and rhythmic structures were different, but the tight melodic-rhythmic correspondence was either identical to that of the first half (it was indeed drawn from the second half) or this correspondence was perturbed to a greater or lesser extent in a series of systematic modifications of the score materials. Listeners gave the highest ratings to the repetition of (already heard) excerpts from the first half, followed by excerpts from the second half and then the various modified versions that changed the correspondences between chroma, register, and duration.

As mentioned above, however, musical material can also be represented in terms of relations among attributes in a more or less precise way: exact intervals, relations within a musical system such as a scale, contours, and so on. The perception of similarity would thus be based on a match between relations abstracted from the material attributes rather than directly on the values of the attributes themselves. A melodic contour in Béla Bartók's *Music for Strings, Percussion and Celesta* (1936), for example, might undergo an expansive transformation in which all of the intervals are increased, and yet the whole stays within a certain pitch system and the pitch contour is preserved. Since the rhythmic material is relatively similar between the versions, listeners may hear a strong link between this variation and the original.

Finally, as also mentioned previously, research[20,21] has shown that materials that are quite different in surface structure (affecting specific surface values and the relations between adjacent events) may still be perceived as similar if they share an underlying structure such as the reduced structure proposed by Lerdahl and Jackendoff[17] and Schenker[24] before them. In these cases, it is necessary to postulate that the representation of the event structure of the material is organized to some extent in a hierarchical fashion and that it is some level of this hierarchy that is compared between versions. The experimental question here concerns the depth of the hierarchical representation that can be used to derive a sense of similarity between materials that differ drastically in surface structure. It would seem that it is this kind of structural abstraction that would be one of the main bases of the musical invariance underlying theme and variations forms. Bigand,[20] for example, has demonstrated that listeners group together musical excerpts that have the same underlying reduced structure although they may vary considerably at the surface level.

From these various considerations, we propose that there are (at least) three types of similarity that should be taken into consideration, all the while understanding that they are overlapping and not mutually exclusive categories. The first is relatively abstract and concerns similarity of the statistical distribution of surface values and their derivatives as well as of surface relations of first and higher orders, including perhaps even transitional

probabilities from one value or relation to another (cf. Ref. 25 for implicit learning of transitional probabilities for pitch and Ref. 26 for those for timbre). This kind of similarity would seem to be involved in the perception of similar textures, where it is not the exact values that count but the probability of occurrence and of transition that define the texture. The second kind concerns figural similarity, specific patterns of attributes (perhaps associated intimately across auditory dimensions), that give perceptual landmarks within an extended material that allow its recognition when a transformed version still contains the same identifying pattern, or something very similar to it. This notion of figural similarity is akin to the notion of imprint proposed by Deliège.[27] The third kind of similarity is structural and relies on the abstraction of structural invariants related to the event hierarchy, often defined by underlying harmonic and metric templates that are maintained constant in cases of elaborative or reductive transformations of thematic materials.

## Form-bearing dimensions

A framework concerning the psychological constraints on form-bearing dimensions in music has been developed.[1] Among these psychological constraints, some seem particularly cogent for the present concerns. Form-bearing dimensions are those along which invariant relations can be configured. For them to be useful in musical discourse these relations must be easily discriminable. Further in evaluating the degree of similarity between thematic materials and their transformations, time has necessarily passed between the two, often several tens of seconds in real music. The configurations must thus be memorizable to some degree of precision either in their absolute form or in some abstracted relative form.

In real music, several dimensions may be varied for a given material. For example, transposing a melody in pitch that is played by the same instrument maintains the same rhythmic structure, and while globally the timbre may change slightly, it remains roughly constant. One psychological constraint on form-bearing dimensions that merits extensive consideration is the resistance of perceived invariance on one dimension to independent variation on other dimensions. In other words, to what extent are the dimensions processed independently or do they interact? This raises the modularity question that has been addressed in both theoretical and experimental terms.[28–30] Both the psychological and the musical issues are extremely complex here since it becomes clear that one may have varying degrees of identity and difference on different dimensions, and the question would be to know how strong the sense of similarity was under certain conditions, or even whether similarity could be maintained. The perception of such transformations would depend on several factors. At the level of the stimulus, it would depend on the way the composer used variation across dimensions and on the types of interactions between dimensions, which may indeed by asymmetric. For example, in Olivier Messiaen's *Mode de Valeurs et d'Intensités* for piano (1950), many dimensions are used: pitch, dynamics, articulation, duration. However, as mentioned above, there is a coupling between the values of one and those of the others. Krumhansl[23] has shown that listeners pick up on these interdimensional correlations and can judge that new material fits or does not fit with the rules developed by the composer. In integral serialism on the other hand, the composer may have series of pitches, dynamics, articulations, and durations, each series having a different

number of elements, and they thus become uncorrelated with one another as in Pierre Boulez' *Structures* for two pianos (1952). It would be interesting to repeat Krumhansl's experiment on such material. However, she and her colleagues[31] have shown that many listeners do hear different transformations of row forms (retrograde, inversion, and retrograde inversion) as being related to the original prime series in twelve-tone rows drawn from Schoenberg's *Wind Quintet* (1924) and *Fourth String Quartet* (1936).

A psychological factor that needs to be considered is the resistance of perceived invariance/similarity to perturbation by intervening material. We often presume as score gazers that recognition of the return of a theme is unaffected by new intervening material. However, work in auditory short-term memory has shown that relations of attribute similarity between the to-be-remembered information and the intervening material can degrade recognition and discrimination performance. This kind of auditory memory does seem to be somewhat modular however in that intervening material that is different on the criterial dimension but similar on other dimensions does not necessarily degrade performance, at least when the attributes of single tones are to be remembered.[11,31–40] At a more complex level and for longer term effects in the memory of musical materials varying along several auditory dimensions, it is unclear at present what kinds of contextual effects may exist that would render some intervening material more or less perturbing of the memory trace of the original material and thus hinder its recognition. If such were the case then the 'success' of a given theme in contributing to a sense of coherence in a musical form would certainly be compromised.

Given that materials can vary on different dimensions and that there have been strong tendencies in many cultures to focus primarily on pitch and duration, another psychological question that comes to the fore is the extent to which listeners can learn to focus on changes along a dimension not usually used in one's past experience as a structuring dimension in musical discourse, such as timbre or space, for example. That listeners do not have experience in listening for such structures would probably show up in psychological testing, but we could not conclude from such results that the use of these dimensions was impossible or doomed to failure from the outset. And so the question of perceptual learning will need to be addressed within the framework of contemporary musical experience.

Finally, one of the crucial problems of the perception of musical transformations in relation to their original material concerns perceptual coherence. Variations that affect stream formation or change the availability of properties of the material in working memory may indeed render what is perceived unsimilar to the original material since relations among perceptual attributes would appear to be computed on events belonging to the same auditory stream.[9,41]

## Experiments on musical similarity

Having drawn with broad strokes a rough and largely intuitive account of what musical similarity might be and what it's role in music might be, let us examine a few studies that have specifically addressed this question, focusing in particular on recent work from our laboratory that takes a few small steps towards the answer. Several experiments have studied the implication of musical training, musical memory, pitch hierarchy, and musical

context in the perception of tonal and nontonal transformations and their similarity ratings.[11,31,37–40] For example, similarity perception is not a symmetrical phenomenon: the similarity perceived between a melody A and a melody B does not automatically predict the similarity perceived between B and A.[37] Finally, although participants show an ability to recognize the type of transformations applied to a nontonal material presented isochronously, this ability tends to decline when transformations are multidimensional, that is when both pitch and rhythm dimensions vary simultaneously.[11,31,38] All these facts however do not explore or explain the computational basis of melodic similarity perception. In fact, is it plausible to imagine that melodic similarity perception could rely on general mental processes that are common to all listeners, independently from the melody's morphology (i.e. number of pitches, rhythmic figures, metronomical speed, etc.) and the listener's musical aptitude?

In recent work, we asked the following questions: What is the nature of musical transformation space and how is it constrained by the type of musical material and musical training (Ref. 2, exp. 1)? To what extent do pitch and duration components of musical materials contribute independently to global similarity judgements and discrimination focussed on each (Ref. 2, exp. 2)?

## Musical transformation space: limits of similarity

Three tonal melodies were used as reference material (Ref. 2, exp. 1). Following Serafine *et al.*[21] a lure melody was constructed for each one such that the surface characteristics were very similar, while the underlying reduced structure was different. For each melody, 16 transformations were composed that differed from the original in terms of pitch content, rhythmic content or both. The transformations could be either elaborations or reductions of the original material, reduction being taken in the sense of a simplified abstract representation of a musical structure that keeps structural information such as important notes in the tonal and metric hierarchies.[17] Further, the pitch and rhythmic transformations either respected tonal/metric syntax (reasonable variations within the scale and meter) or departed from it (unlikely chromatic and unmetrical variations). Corresponding structural transformations were applied to the lure melodies as well. The melodies were verified with respect to the desired properties by a practising composer (Joshua Fineberg) working at IRCAM.

Musician and nonmusician listeners heard one of the three reference melodies followed by a test melody. The test melodies could be the reference or its corresponding lure or any of their transformations. The listeners made a similarity rating on a scale of 1–9 with the high end corresponding to maximum similarity.

The experiment was designed to test several hypotheses:

1. Musicians should be able to hear similarity to greater degrees of transformation if the transformations respect the syntactical rules.
2. If listeners are sensitive to commonalities at certain levels of hierarchical reduction, transformations that respect the reduction should be more similar than those that violate it (including the lure transformations).

3. Drastic transformations in both pitch and rhythm should render the variation as dissimilar to the reference as unrelated material, represented by the lure transformations.
4. The same types of transformation should have similar effects on the three reference melodies, although local morphological differences may prevail due to figural properties of the melodies.

The data revealed no systematic differences between the two populations of listeners. This last finding is concordant with previous studies that have specifically examined effects of musical training on musically realistic tasks.[42–46] Transformations that respected the reduced structure were judged significantly more similar to the reference than those that did not. The dissimilarity was particularly strong for chromatic transformations and for combined chromatic and nonmetric transformations, many of which were judged as dissimilar from the original as were the lure melody and its transformations. This latter result, when combined with the fact that the large majority of lure transformations were judged quite dissimilar to the original, indicates that some transformations of the original were perceived as completely unrelated. And finally, the three reference melodies were different concerning the perceived effects of the transformation types, demonstrating the importance of local surface features in the perception of melodic similarity and the fact that a transformational algorithm must necessarily interact complexly with the to-be-transformed material. However, rank correlations between reference melodies were very high indicating similar qualitative effects of the set of transformations applied to all three melodies.

## Modularity of pitch and rhythm in perceived similarity

The answer to one question that cannot be easily gleaned from the results of the first study is the degree to which pitch and rhythmic materials interact in the perception of similarity. Another experiment (Ref. 2, exp. 2) was designed to address this problem more explictly. We therefore introduced a simple computational model based on a modular hypothesis that could represent the basis of the mental processing underlying melodic similarity perception. If the computational model turns out to be reliable for tonal melodies, can it also predict similarity perception of nontonal melodies and thus indicate general processing mechanisms independently of the musical system? One may wonder whether a modular approach is relevant or not to the study of music perception in general and of similarity perception in particular. Examples taken from vision, language, audition, and neuropsychology lead us to consider the modular hypothesis in order to clarify the specificities of similarity perception. Interestingly, aside from modular considerations in music theory,[30] the only consequential empirical research done within an explicitly modular perspective comes from the neuropsychological work of Peretz and collaborators.[29,47,48]

Based on previous findings, which argue in favour of independent pitch and duration processing structures,[40,49] the main hypothesis posited here is that a similarity judgement made between a reference melody and a transformation of that melody can be understood as an additive integration of two similarity coefficients computed by the two hypothetical pitch and duration modules. While such an additive approach is certainly a simplistic way

to consider the relations between the two modules, such a general principle should be applicable to any musical system without regard to the nature of the underlying musical grammar, the tonal and twelve-tone systems in our case.

To test this independence, a paradigm similar to that of Garner and Morton[50] was used in which discriminations of changes on one dimension were made in the presence or absence of changes along another dimension, and vice versa, by different groups of listeners. A third group was required to make both discriminations. We coupled this discrimination task with a similarity judgement task. This paradigm was used in order to investigate whether the listeners could recognize a pitch or duration pattern independently from the values taken by the other dimension. Note that since the judgement is focused on a single dimension, at times in the face of variation on the other dimension, perfect discrimination performance across all conditions indicates an ability to ignore the other dimension, implying perceptual independence.

Four reference melodies were composed: one major, one minor, and two twelve-tone melodies. In addition to the original pitch or rhythm patterns, three variations of each dimension were created. These variations were designed to be progressively less similar to the reference pattern. The pitch and rhythm patterns were crossed giving 15 transformations of each reference. The melodies were all composed to achieve the desired properties by a practising composer (Jacopo Baboni-Schilingi) working at IRCAM. The transformations could thus differ only in pitch, only in rhythm or in both. Listeners were divided into three groups: pitch discrimination, rhythm discrimination or both. They first made discrimination judgements on the pitch and/or temporal dimension and then judged globally on a 10-point scale the strength of similarity for each type of transformation in relation to the reference melody.

The discrimination results globally corroborate the modularity hypothesis, although complete perceptual independence was not observed. In particular, the pitch and duration modules do not exhibit the same processing properties since duration processing seems to be less permeable to pitch information than pitch is to duration (at least for tonal melodies). This is reflected when comparing participants' discrimination performance on the two musical dimensions for the two musical systems. For all melodies, duration configurations were better discriminated than pitch configurations and in particular identical pitch patterns were judged as being progressively different when the rhythm pattern was indeed more and more removed from the reference, even though the pitch pattern was identical. By deduction this finding is coherent with existing empirical data in melodic similarity perception where it has been shown that the duration dimension was the major one used by participants for establishing similarity judgements.[51] The pitch permeability to duration configuration was limited to tonal melodies suggesting a functional interaction between the two in that musical system which is decoupled in the twelve-tone system.

Variations among listeners' similarity judgements showed that tonal and nontonal transformations were globally perceived in different ways. While the tonal transformations were hierarchically distinguished, this was not the case for the perception of the nontonal transformations where the pitch and duration modalities were not hierarchically differentiated. Further, whereas the perception of tonal transformations appeared to be perceived in a systematic way over repetitions and for the two reference melodies, this was not the case for

the nontonal transformations where the range of variation of listeners' judgements was found to be much greater. Consequently, it seems reasonable to infer that similarity perception of tonal and nontonal melodies depended at least partly on different mental processing mechanisms.

Under the perceptual independence hypothesis, we claimed that similarity perception could be approximated by an elementary additive model. The model decomposes the listeners' similarity ratings into two numerical values: basic pitch and duration values (averaged across listeners) for each level of transformation of that dimension without change on the other dimension. The mean of the basic values for pitch and duration gives the model prediction. For three of the four reference melodies (including both tonal melodies and the first nontonal melody), listeners' mean similarity judgements for the nine two-dimensional transformations were highly correlated with model predictions. So although some differences between the tonal and nontonal melodies exist, they also seem to have certain commonalities as concerns processing in terms of rhythmic and pitch similarity. There are nonetheless deviations of the data from model predictions for each transformation, and these will need to be examined in detail to determine their cause, particularly for the second nontonal melody. There are perhaps salient surface or structural features not taken into account by this simplistic model that played a role in the perceived similarity.

## Conclusions/questions

From the data, it is clear that the space of possible variation of thematic material that still maintains a link of perceptual similarity to the original is limited. Although conditions of direct comparison are different from those in real music where intervening material or competing material would often be present, a number of factors can be targeted in provisory fashion for more contextually realistic work in the future. A transformation that maintains a majority of exact values or relations among values in the correct temporal order will be more strongly similar than one that changes all values and all relations. Although this has not been explicitly tested, it may be that maintaining the same distribution of values or relations could still elicit a sense of similarity if compared with material that had dramatically different distribution of attribute values. Further relations of similarity can exist at more abstract levels of a hierarchical reduction, indicating something about the hierarchical nature of the mental representation of musical material in both tonal and nontonal music. From examining the similarity judgements on elaborative and reductive transformations, it would seem that these kinds of transformations retain similarity to the extent that the transformations do not violate the 'grammaticality' of the musical system within which the material has been conceived, suggesting a close link between comprehension of the musical system and the categorization processes that one may hypothesize to operate in the recognition of variations of an original theme.

Another issue of major import for the notion of musical similarity and invariance is the notion of perceptual independence between patterns on different musical dimensions. While there seems to be some degree of independence between pitch and duration structures at a global level, there are also asymmetrical interactions in which duration changes

more strongly affect the perceived similarity of identical pitch patterns than is the case for the reverse. More refined experimentation will be necessary to probe the nature of the interaction between the dimensions that gives rise to such asymmetries. Also future research should clearly study more thoroughly different musical systems, including those of other cultures, in an attempt to distinguish the more general cognitive processes from those that are specific to a particular musical grammar.

And finally future research should examine the role that the perception of musical similarity plays in the apprehension or experience of musical form. The process of similarity perception necessarily involves comparison across time of different manifestations of musical material and thus memory processes. One might imagine that the recognition of similarity could thus establish some kind of associative link between different points in the piece and thus contribute in some way to a sense of formal cohesion over longer time spans. Further, one not only recognizes the link between two manifestations but can also appreciate the difference between the two, which perhaps contributes to a perception of the kind of change or the trajectory of change that has taken place. This trajectory may contribute in itself to a sense of musical development over longer time spans and also reinforce the integration of materials at different points in time into a greater whole. There are numerous theoretical and methodological obstacles to be overcome in the design of appropriate experiments to test these contributions in ongoing music listening.

## Acknowledgements

This work has benefitted from fruitful discussions with several composers (Jacopo Baboni-Schilingi, Benjamin Carson, Joshua Fineberg, and Roger Reynolds), a musicologist (François Madurell), and two fellow psychologists (Emmanuel Bigand and Sandrine Vieillard), to whom we express our gratitude.

## References

1. **McAdams, S.** (1989) Psychological constraints on form-bearing dimensions in music. *Contemporary Music Rev.* 4(1), 181–98.
2. **Matzkin, D.** (2001) Perception de similarité de mélodies tonales et non tonales: Etude pluridisciplinaire. Unpublished PhD thesis, Ecole des Hautes Etudes en Sciences Sociales, Paris.
3. **Matzkin, D.** and S. McAdams (2000) Similarity perception of variations of tonal and twelve-tone melodies. In *Proceedings of the 6th International Conference on Music Perception and Cognition.* Keele.
4. **Reynolds, R.** (2002) *Form and Method: Composing Music* (The Rothschild Essays). New York: Routledge.
5. **Reynolds, R.** (2002) Compositional strategies in The Angel of Death for piano, chamber orchestra and computer-processed sound. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, and J. Renwick (eds) *Proceedings of the 7th International Conference on Music Perception and Cognition.* Sydney, Adelaide: Causal Productions.
6. **Bigand, E., S. Vieillard, F. Madurell, S. McAdams,** and **B. Poulin** (2002) Effects of instrumentation on the memorization of musical materials. In C. Stevens, D. Burnham, G. McPherson, E. Schubert,

and J. Renwick (eds) *Proceedings of the 7th International Conference on Music Perception and Cognition.* Sydney, Adelaide: Causal Productions.

7. **Vieillard, S., S. McAdams, O. Houix,** and **R. Reynolds** (2002) Perceptual and cognitive criteria used in the categorization of thematic excerpts from a contemporary musical piece. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, and J. Renwick (eds) *Proceedings of the 7th International Conference on Music Perception and Cognition.* Sydney, Adelaide: Causal Productions.
8. **Schoenberg, A.** (1978) *Theory of Harmony.* Berkeley, CA: University of California Press.
9. **McAdams, S.** and **A. S. Bregman** (1979) Hearing musical streams. *Computer Music J.* 3(4), 26–43.
10. **Dowling, W. J.** (1986) Context effects on melody recognition: scale-step versus interval representations. *Music Percept.* 3, 281–96.
11. **Dowling, W. J.** (1972) Recognition of melodic transformations: inversion, retrograde, and retrograde inversion. *Percep. and Psychophysics* 12, 417–21.
12. **Ehresman, D.** and **D. L. Wessel** (1978) Perception of timbral analogies. IRCAM Report no. 13, IRCAM, Paris.
13. **McAdams, S.** and **J. C. Cunibile** (1992) Perception of timbral analogies. *Philosophical Transactions of the Royal Society, London, Series B* 336, 383–9.
14. **Roads, C.** (1985) Granular synthesis of sound. In C. Roads and J. Strawn (eds) *Foundations of Computer Music.* Cambridge, MA: MIT Press, pp. 145–59.
15. **Vieillard, S.,** E. **Bigand, F. Madurell,** S. **McAdams,** and R. **Reynolds** (2002) Can listening to excerpts of original versions of contemporary musical materials facilitate recognition of their transformed versions? In C. Stevens, D. Burnham, G. McPherson, E. Schubert, and J. Renwick (eds) *Proceedings of the 7th International Conference on Music Perception and Cognition.* Sydney, Adelaide: Causal Productions.
16. **Vandenheede, J.** (1992) Jonathan Harvey's Ritual Melodies. *Interface* 21(2), 149–83.
17. **Lerdahl, F.** and **R. Jackendoff** (1983) *The Generative Theory of Tonal Music.* Cambridge, MA: MIT Press.
18. **Lerdahl, F.** (1989) Atonal prolongational structure. *Contemporary Music Rev.* 4, 65–87.
19. **Dibben, N.** (1994) The cognitive reality of hiearchical structure in tonal and atonal music. *Music Percept.* 12, 1–25.
20. **Bigand, E.** (1990) Abstraction of two forms of underlying structures in a tonal melody. *Psychol. of Music,* 18, 45–60.
21. **Serafine, M. L., N. Glassman,** and **C. Overbeek** (1989) The cognitive reality of hierarchic structure in music. *Music Percept.,* 6, 397–430.
22. **Stoffer, T. S.** (1985) Representation of phrase structure in the perception of music. *Music Percept.* 3, 191–220.
23. **Krumhansl, C. L.** (1991) Memory for musical surface. *Memory and Cognit.* 19, 401–11.
24. **Schenker, H.** (1935/1979) *Free Composition.* New York: Longman.
25. **Saffran, J. R., E. K. Johnson, R. N. Aslin,** and **E. L. Newport** (1999) Statistical learning of tone sequences by human infants and adults. *Cognition,* 70, 27–52.
26. **Bigand, E., P. Perruchet,** and **M. Boyer** (1998) Implicit learning of artificial grammar of musical timbres. *Current Psychol. of Cognit.* 17, 577–600.
27. **Deliège, I.** (1989) A perceptual approach to contemporary musical forms. *Contemporary Music Rev.* 4, 213–30.
28. **Fodor, J.** (1983) *The Modularity of Mind.* Cambridge, Mass: MIT Press.

29. **Peretz, I.** and **J. Morais** (1989) Music and modularity. *Contemporary Music Rev.* 2(1), 279–94.
30. **Temperley, D.** (1995) Motivic perception and modularity. *Music Percept.* 13, 141–70.
31. **Krumhansl, C. L., G. J. Sandell,** and **D. C. Sergeant** (1987) The perception of tone hierarchies and mirror forms in twelve tone serial music. *Music Percept.* 5, 31–78.
32. **Deutsch, D.** (1970) Tones and numbers: specificity of interference in immediate memory. *Science* 168, 1604–5.
33. **Deutsch, D.** (1972) Mapping of interactions in the pitch memory store. *Science* 175, 1020–2.
34. **Semal, C.** and **L. Demany** (1991) Dissociation of pitch from timbre in auditory short-term memory. *J. Acoust. Soc. of Am.* 89, 2404–10.
35. **Semal, C.** and **L. Demany** (1993) Further evidence for an autonomous processing of pitch in auditory short-term memory. *J. Acoust. Soc. Am.* 94, 1315–22.
36. **Starr, G.** and **M. Pitt** (1997) Interference effects in short-term memory for timbre. *J. Acoust. Soc. Am.* 102, 486–94.
37. **Bartlett, J. C.** and **W. J. Dowling** (1988) Scale structure and similarity of melodies. *Music Percept.* 5, 286–313.
38. **Francès, R.** (1972) *La perception de la musique.* Paris: Vrin.
39. **Halpern, A. R., J. C. Bartlett,** and **W. J. Dowling** (1998) Perception of mode, rhythm, and contour in unfamiliar melodies. Effect of age and experience. *Music Percept.* 15, 335–55.
40. **Pitt, M. A.** and **C. B. Monahan** (1987) The perceived similarity of auditory polyrhythms. *Percept. Psychophys.* 41, 534–46.
41. **Bregman, A. S.** (1990) *Auditory Scene Analysis: The Perceptual Organization of Sound.* Cambridge, MA: MIT Press.
42. **Bigand, E.** (1993) The influence of implicit harmony, rhythm, an musical training on the abstraction of tension-relaxation schemas in tonal musical phrases. *Contemporary Music Rev.* 9, 123–37.
43. **Bigand, E.** (1997) Perceiving musical stability: the effect of tonal structure, rhythm and musical expertise. *J. Exp. Psychol., Hum. Percep. Perform.* 23, 808–18.
44. **Bigand, E.** and **M. Pineau** (1996) Context effects on melody recognition: a dynamic interpretation. *Curr. Psychol. of Cognit.* 15, 121–34.
45. **Bigand, E., B. Poulin, B. Tillmann, F. Madurell,** and **D. D'Adamo.** Sensory versus cognitive components in harmonic priming. *J. Exp. Psychol.: Hum. Percept. Perform.*, in press.
46. **Bigand, E., B. Tillmann, B. Poulin,** D. D'Adamo, and **F. Madurell** (2001) The effect of harmonic context on phoneme monitoring in vocal music. *Cognition*, 81, 11–20.
47. **Peretz, I.** (1993) Auditory agnosia: a functional analysis. In S. McAdams and E. Bigand (eds) *Thinking in Sound: The Cognitive Psychology of Human Audition.* Oxford: Oxford University Press, pp. 199–230.
48. **Peretz, I.** and **R. Kolinsky** (1993) Boundaries of separability between melody and rhythm in music discrimination: a neuropsychological perspective. *Quarterly J. Exp. Psychol.* 46A, 301–25.
49. **Palmer, C.** and **C. L. Krumhansl** (1987) Independent temporal and pitch structures in the determination of musical phrases. *J. Exp. Psychol: Hum. Percept. Perform.* 13, 116–26.
50. **Garner, W. R.** and **J. Morton** (1969) Perceptual independence: definition, models, and experimental paradigms. *Psychol. Bulletin* 72, 233–59.
51. **Monahan, C. B.** and **E. C. Carterette** (1985) Pitch and duration as determinants of musical space. *Music Percept.* 1, 1–32.

CHAPTER 7

# TONAL COGNITION

CAROL L. KRUMHANSL AND PETRI TOIVIAINEN

## Abstract

This chapter presents a self-organizing map (SOM) neural network model of tonality based on experimentally quantified tonal hierarchies. A toroidal representation of key distances is recovered in which keys are located near their neighbours on the circle of fifths, and both parallel and relative major/minor key pairs are proximal. The map is used to represent dynamic changes in the sense of key as cues to key become more or less clear and modulations occur. Two models, one using tone distributions and the other using tone transitions, are proposed for key-finding. The tone transition model takes both pitch and temporal distance between tones into account. Both models produce results highly comparable to those of musically trained listeners, who performed a probe tone task for ten nine-chord sequences. A distributed mapping of tonality is used to visualize activation patterns that change over time. The location and spread of this activation pattern is similar for experimental results and the key-finding model.

Keywords: Music; Tonality; Cognition; Probe tone

## Introduction

Tonality induction refers to the process through which the listener develops a sense of the key of a piece of music. The concept of tonality is central to Western music, but eludes definition. From the point of view of musical structure, tonality is related to a cluster of features, including musical scale (usually major or minor), chords, the conventional use of sequences of chords in cadences, and the tendencies for certain tones and chords to suggest or be 'resolved' to others. From the point of view of experimental research on music cognition, tonality has implications for establishing hierarchies of tones and chords, and for inducing certain expectations in listeners about how melodic and harmonic sequences will continue. One method for studying the perception of tonality is the probe tone method, which quantifies the tonal hierarchy. When applied to unambiguous key-defining contexts, it provides a standard for determining key strengths when more ambiguous and complex musical materials are presented. In addition to experimental studies, considerable effort has been spent developing computational models. This effort has produced various symbolic and neural network models, including a number that take musical input and return a key identification, sometimes called key-finding models.

## Probe tone methodology

An experimental method introduced to study tonality is sometimes referred to as the probe tone method.[1] It is best illustrated with a concrete example. Suppose you hear the tones of the ascending C major scale: C D E F G A B. There is a strong expectation that the next tone will be the tonic, C, first, because it is the next logical tone in the series and, second, because it is the tonic of the key. In the experiment, the incomplete scale context was followed by the tone C (the probe tone), and listeners were asked to judge how well it completed the scale on a numerical scale (1 = very bad, 7 = very good). As expected, the C received the maximal rating. Other probe tones, however, also received fairly high ratings, and they were not necessarily those that are close to the tonic C in pitch. For example, the most musically trained listeners also gave high ratings to the dominant, G, and the mediant, E. In general, the tones of the scale received higher ratings than the nonscale tones, C♯ D♯ F♯ G♯ A♯. This suggested that it is possible to get quantitative judgements of the degree to which different tones are perceived as stable, final tones in tonal contexts.

A subsequent study[2] used this method with a variety of musical contexts at the beginning of the trials. They were chosen because they are clear indicators of a key. They included the scale, the tonic triad chord, and chord cadences in both major and minor keys. These were followed by all possible probe tones in the chromatic scale, which musically trained listeners were instructed to judge in terms of how well they fit with the preceding context in a musical sense. Different major keys were used, as were different minor keys. The results for contexts of the same mode were similar when transposed to a common tonic. Also, the results were similar independent of which particular type of context was used. Consequently, the data were averaged over these factors. We call the resulting values the K-K profiles, which can be expressed as vectors. The vector for major keys is: K-K major profile = ⟨6.35, 2.23, 3.48, 2.33, 4.38, 4.09, 2.52, 5.19, 2.39, 3.66, 2.29, 2.88⟩. The vector for minor keys is: K-K minor profile = ⟨6.33, 2.68, 3.52, 5.38, 2.60, 3.53, 2.54, 4.75, 3.98, 2.69, 3.34, 3.17⟩.

We can generate K-K profiles for 12 major keys and 12 minor keys from these. If we adopt the convention that the first entry in the vector corresponds to the tone C, the second to C♯/D♭, the third to D, and so on, then the vector for C major is: ⟨6.35, 2.23, 3.48, 2.33, 4.38, 4.09, 2.52, 5.19, 2.39, 3.66, 2.29, 2.88⟩, the vector for Db major is: ⟨2.88, 6.35, 2.23, 3.48, 2.33, 4.38, 4.09, 2.52, 5.19, 2.39, 3.66, 2.29⟩, and so on. The vectors for the different keys result from shifting the entries the appropriate number of places to the tonic of the key.

## Tracing the developing and changing sense of key

The probe tone method was then used to study how the sense of key develops and changes over time.[2] Ten nine-chord sequences were constructed, some of which contained modulations between keys. Musically trained listeners did the probe tone task after the first chord, then after the first two chords, then after the first three chords, and continued until the full sequence was heard. This meant that 12 (probe tones) × 9 (chord positions) × 10 (sequences) = 1080 judgements were made by each listener. Each of the 90 sets of probe tone ratings was compared with the ratings made for the unambiguous key-defining contexts. That is, each set of probe tone ratings was correlated with the K-K profiles for the 24 major and minor keys. For some of the sets of probe tone ratings (some probe positions in

some of the chord sequences), a high correlation was found, indicating a strong sense of key. For other sets of probe tone ratings, no key was highly correlated, which was interpreted as an ambiguous or sense of key.

As should be obvious from the above, the probe tone task requires an intensive empirical effort to trace how the sense of key develops and changes, even for short sequences. In addition, the sequence needs to be interrupted, and the judgement is made after the sequence has been interrupted. For these reasons, the judgements may not faithfully mirror the experience of music in time. Hence, we[3] have tested an alternative form of the probe tone methodology. In this method, which we call the concurrent probe tone task, the probe tone is presented continuously while the music is played. The complete passage is sounded together with a probe tone. Then the passage is sounded again, this time with another probe tone. This process is continued until all probe tones have been sounded. Preliminary results suggest this methodology produces interpretable results, at least for musically trained participants. Our focus here, however, is on how the sense of key can be represented, whether the input to the representation is from a probe tone task or from a model of key-finding as described later.

## A geometric map of key distances from the tonal hierarchies

The K-K profiles generated a highly regular and interpretable geometric representation of musical keys.[2] The basic assumption underlying this approach was that two keys are closely related to each other if they have similar tonal hierarchies. That is, keys were assumed to be closely related if tones that are stable in one key are also relatively stable in the other key. To measure the similarity of the profiles, a product–moment correlation was used. It was computed for all possible pairs of major and minor keys, giving a $24 \times 24$ matrix of correlations showing how similar the tonal hierarchy of each key was to every other key. To give some examples, C major correlated relatively strongly with A minor (0.651), with G major and F major (both 0.591), and with C minor (0.511). C minor correlated relatively strongly with E♭ major (0.651), with C major (0.511), with A♭ major (0.536), and less strongly with F minor and G minor (both 0.339).

A technique called multidimensional scaling was then used to create a geometric representation of the key similarities. The algorithm locates 24 points (corresponding to the 24 major and minor keys) in a spatial representation to best represent their similarities. It searches for an arrangement such that points that are close correspond to keys with similar K-K profiles (as measured by the correlations). In particular, non-metric multidimensional scaling seeks a solution such that distances between points are (inversely) related by a monotonic function to the correlations. A measure (called 'stress') measures the amount of deviation from the best-fitting monotonic function. The algorithm can search for a solution in any specified number of dimensions. In this case, a good fit to the data was found in four dimensions.

The four-dimensional solution located the 24 keys on the surface of a torus (generated by one circle in dimensions 1 and 2, and another circle in dimensions 3 and 4). Because of this, any key can be specified by two values: its angle on the first circle and its angle on the second circle. Thus, the result can be depicted in two dimensions as a rectangle where it is understood that the left edge is connected to the right edge, and the bottom edge is connected to the top edge. The locations of the 24 keys were interpretable in terms of music theory. There was one circle of fifths for major keys (...F#/G♭, D♭, A♭, E♭, B♭, F, C,

G, D, A, E, B, F♯/G♭ ...) and one circle of fifths for minor keys (...f♯, c♯, g♯, d♯/e♭, b♭, f, c, g, d, a, e, b, f♯ ...). These wrap diagonally around the torus such that each major key is located near both its relative minor (e.g. C major and A minor) and its parallel minor (e.g. C major and C minor). (See Figures 7.1 and 7.2.)

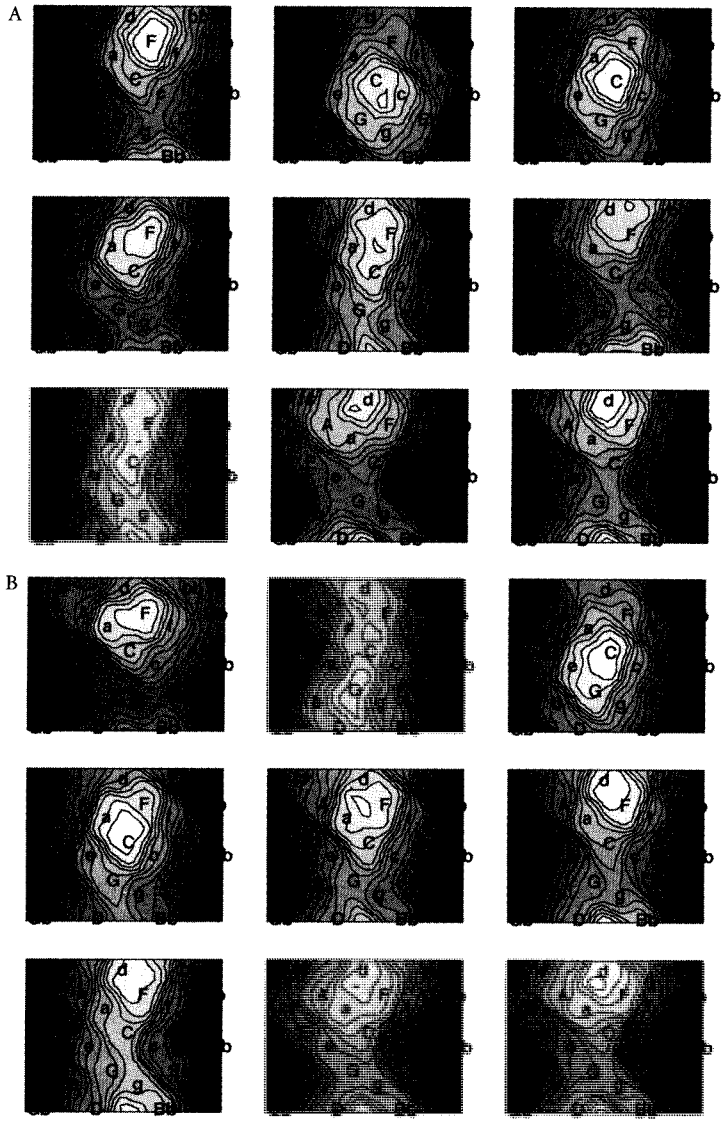



**Figure 7.1** (A) Distributed mapping of tonality on the SOM for probe tone judgements for the nine-chord sequence modulating from C major to D minor. The *top row* shows the results for chords 1, 2, and 3 (*left to right*), the *second row* for 4, 5, and 6 (*left to right*), and the *bottom row* for chords 7, 8, and 9 (*left to right*). (B) Distributed mapping of model 1 for the same sequence.

G, D, A, E, B, F♯/G♭ ...) and one circle of fifths for minor keys (...f♯, c♯, g♯, d♯/e♭, b♭, f, c, g, d, a, e, b, f♯ ...). These wrap diagonally around the torus such that each major key is located near both its relative minor (e.g. C major and A minor) and its parallel minor (e.g. C major and C minor). (See Figures 7.1 and 7.2.)

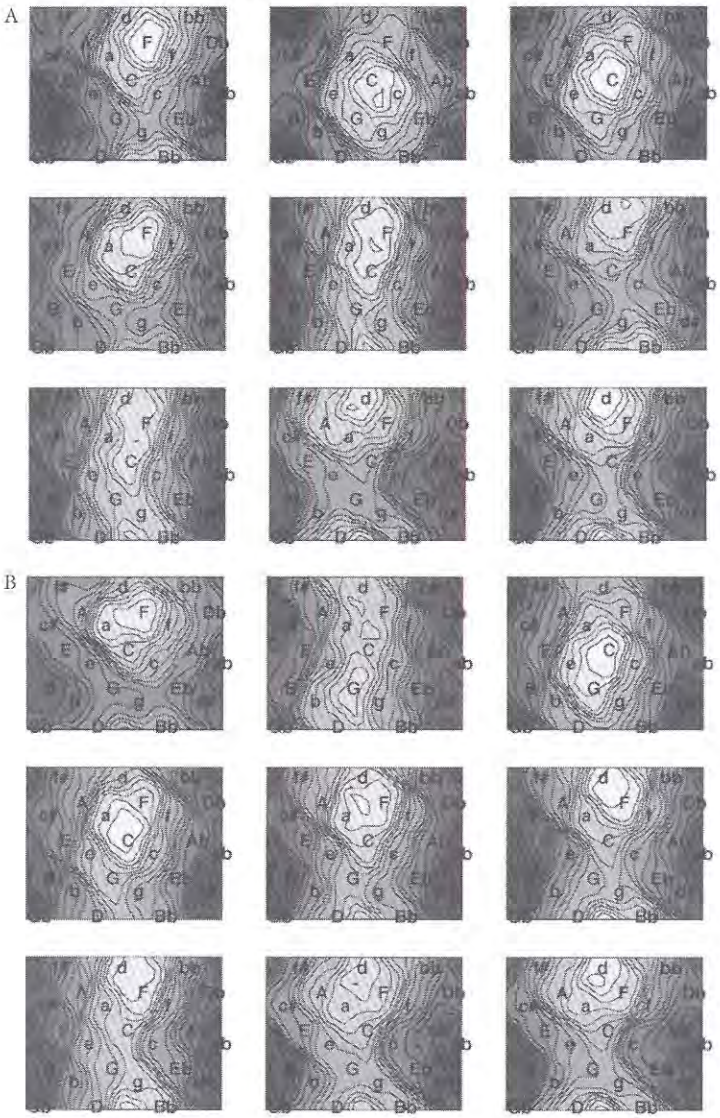



Figure 7.1 (A) Distributed mapping of tonality on the SOM for probe tone judgements for the nine-chord sequence modulating from C major to D minor. The *top row* shows the results for chords 1, 2, and 3 (*left to right*), the *second row* for 4, 5, and 6 (*left to right*), and the *bottom row* for chords 7, 8, and 9 (*left to right*). (B) Distributed mapping of model 1 for the same sequence.

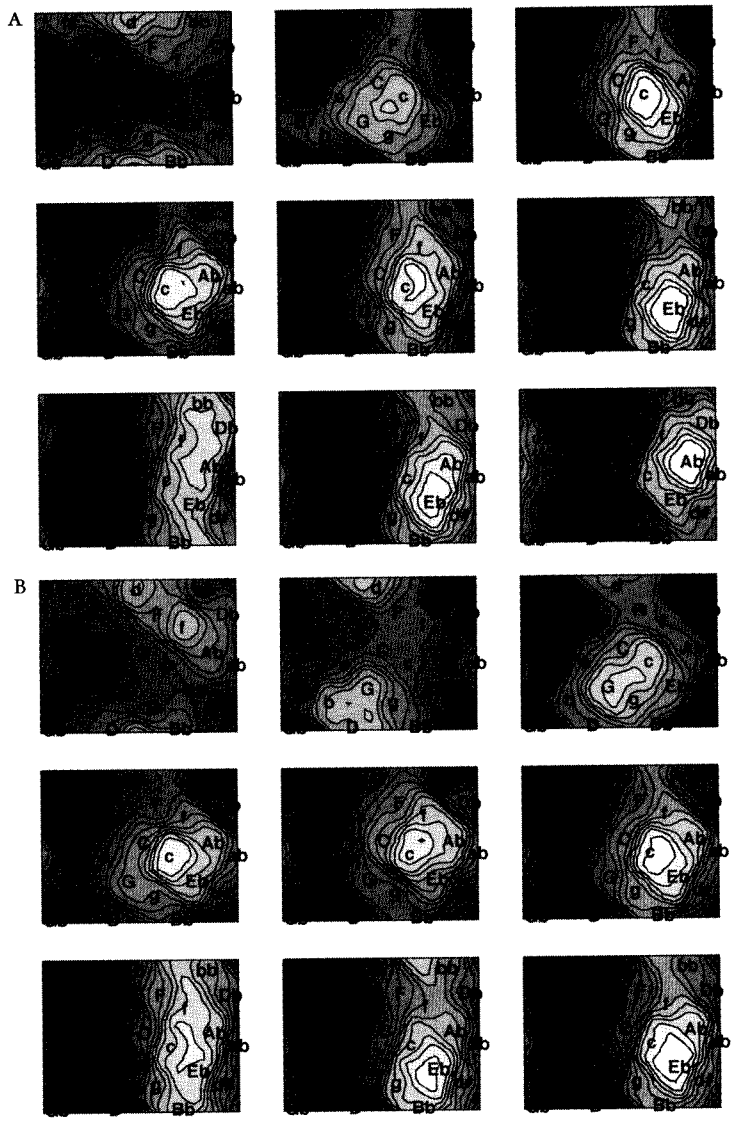



Figure 7.2 (A) Distributed mapping of tonality on the SOM for probe tone judgements for the nine-chord sequence modulating from C minor to Ab minor. Distributing mapping of model 1 for the same sequence.

## Representing the sense of key on the torus

The continuous spatial medium containing the 24 major and minor keys affords representing the sense of key in a graphical form,[2] using a technique called multidimensional unfolding. It is a method that is closely related to multidimensional scaling.

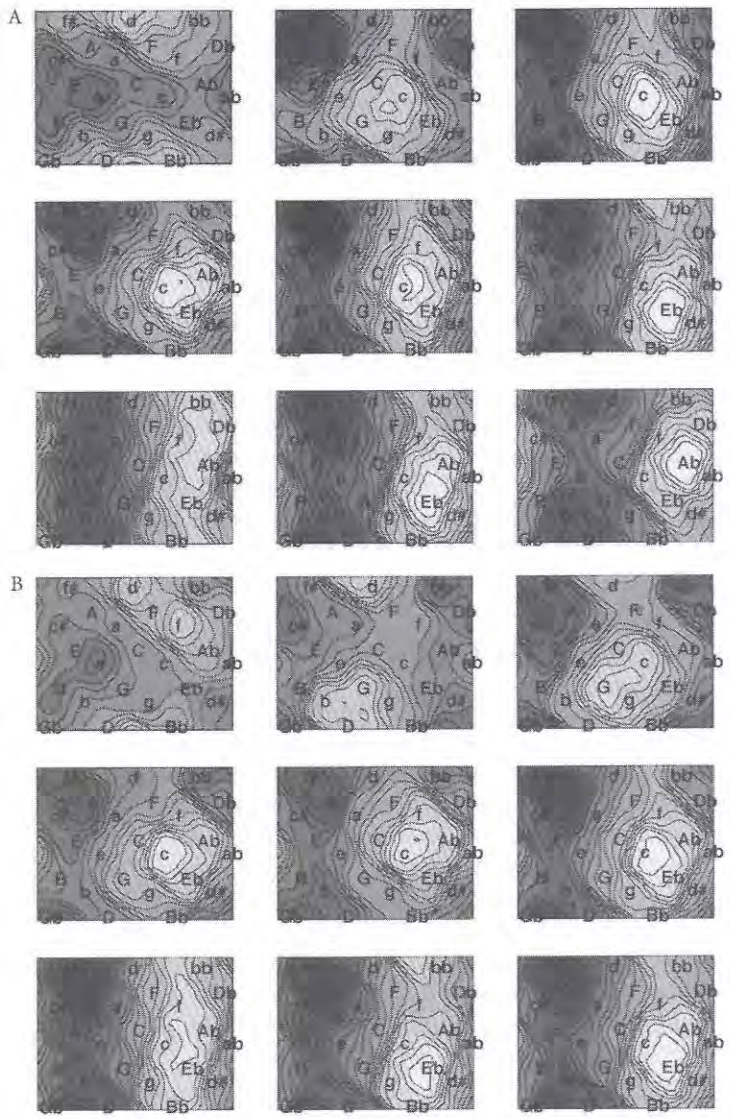



Figure 7.2 (A) Distributed mapping of tonality on the SOM for probe tone judgements for the nine-chord sequence modulating from C minor to Ab minor. Distributing mapping of model 1 for the same sequence.

## Representing the sense of key on the torus

The continuous spatial medium containing the 24 major and minor keys affords representing the sense of key in a graphical form,[2] using a technique called multidimensional unfolding. It is a method that is closely related to multidimensional scaling.

Multidimensional unfolding begins with a multidimensional scaling solution, in this case the torus representation of the 24 major and minor keys. This solution is considered fixed. The algorithm then finds a point in the multidimensional scaling solution to best represent the sense of key at each point in time. Let $P_1$ be the probe tone ratings after the first chord in a sequence; it is a 12-dimensional vector of ratings for each tone of the chromatic scale. This vector was correlated with each of the 24 K-K vectors, giving a 24-dimensional vector of correlations. The unfolding algorithm finds a point to best represent these correlations. Suppose $P_1$ correlates highly with the K-K profile for F major and fairly highly with the K-K profile for D minor. Then the unfolding algorithm will produce a point near these keys and far from the keys with low correlations. Then the vector of correlations was computed for $P_2$, giving a second point. This process continues until the end of the sequence.

In this manner, each of the ten nine-chord sequences[2] generated a series of nine points on the torus representation of keys. For nonmodulating sequences, the points remained in the neighbourhood of the intended key. For the modulating sequences, the first points were near the initial intended key, then shifted to the region of the second intended key. Modulations to closely related keys appeared to be assimilated more rapidly than those to distantly related keys. That is, the points shifted to the region of the new key earlier in sequences containing close modulations than in sequences containing distant modulations.

## Measurement assumptions of the multidimensional scaling and unfolding methods

The above methods make a number of assumptions about measurement, only some of which will be noted here. The torus representation is based on the assumption that correlations between the K-K profiles are appropriate measures of interkey distance. It further assumes that these distances can be represented in a relatively low-dimensional space (four dimensions). This latter assumption was supported by the low stress value (high goodness-of-fit value) of the multidimensional scaling solution. It was further supported by a subsidiary Fourier analysis of the K-K major and minor profiles, which found two relatively strong harmonics.[4] In fact, plotting the phases of the two Fourier components for the 24 key profiles was virtually identical to the multidimensional scaling solution. This supports the torus representation, which consists of two orthogonal circular components. Nonetheless, it would seem desirable to see whether an alternative method with completely different assumptions recovers the same toroidal representation of key distances.

The unfolding method also adopts correlation as a measure of distances from keys, this time using the ratings for each probe position in the chord sequences and the K-K vectors for the 24 major and minor keys. The unfolding technique finds the best-fitting point in the four-dimensional space containing the torus. It does not provide a way of representing cases in which no key is strongly heard, because it cannot generate points outside the space containing the torus. Thus, an important limitation of the unfolding method is that it does not provide a representation of the strength of the key or keys heard at each point in time. For this reason, we sought a method that is able to represent both the region of the key or keys that are heard, together with their strengths.

In addition, the unfolding approach assumes that spatial locations between the 24 points for major and minor keys are meaningful. An intermediate position would result from a blend of tonal hierarchies of nearby keys. However, other sets of probe tone ratings might also map to the same position. Thus, the identification of points between keys is not necessarily unique. This motivated an alternative model which explicitly specifies the meaning of positions between keys.

## The self-organizing map (SOM) of keys

The self-organizing map (SOM)[5] is an artificial neural network that simulates the formation of ordered feature maps. The SOM consists of a two-dimensional grid of units, each of which is associated with a reference vector. Through repeated exposure to a set of input vectors, the SOM settles into a configuration in which the reference vectors approximate the set of input vectors according to some similarity measure; the most commonly used similarity measures are the Euclidean distance and the direction cosine. The direction cosine between an input vector $\mathbf{x}$ and a reference vector $\mathbf{m}$ is defined by

$$\cos\theta = \frac{\sum_i x_i m_i}{\sqrt{\sum_i x_i^2}\sqrt{\sum_i m_i^2}} = \frac{\mathbf{x}\cdot\mathbf{m}}{\|\mathbf{x}\|\|\mathbf{m}\|}. \tag{1}$$

Another important feature of the SOM is that its configuration is organized in the sense that neighbouring units have similar reference vectors. For a trained SOM, a mapping from the input space onto the two-dimensional grid of units can be defined by associating any given input vector with the unit whose reference vector is most similar to it. Because of the organization of the reference vectors, this mapping is smooth in the sense that similar vectors are mapped onto adjacent regions. Conceptually, the mapping can be thought of as a projection onto a nonlinear surface determined by the reference vectors.

We trained the SOM with the 24 K-K profiles. The SOM was specified in advance to have a toroidal configuration, that is, the left and the right edges of the map were connected to each other as were the top and the bottom edges. Euclidean distance and direction cosine, when used as similarity measures in training the SOM, yielded identical maps. (The resulting map is displayed in Figures 7.1 and 7.2.) The map shows the units with reference vectors that correspond to the K-K profiles. The SOM configuration is highly similar to the multidimensional scaling solution[2] and the Fourier-analysis-based projection[4] obtained with the same set of vectors. Unlike those maps, however, all locations in the map are explicitly associated with a reference vector so that they are uniquely identified.

## Representing the sense of key on the SOM

A distributed mapping of tonality can be defined by associating each unit with an activation value. For each unit, this activation value depends on the similarity between the input vector and the reference vector of the unit. Specifically, the units whose reference vectors

are highly similar to the input vector have a high activation, and vice versa. The activation value of each unit can be calculated, for instance, using the direction cosine of Equation 1. Dynamically changing data from either probe tone experiments or key-finding models can be visualized as an activation pattern that changes over time. The location and spread of this activation pattern provides information about the perceived key and its strength. More specifically, a focused activation pattern implies a strong sense of key and vice versa.

## Key-finding models

A variety of key-finding algorithms have been proposed. The objective of these models is to assign a key to an input sample of music (such as the fugue subjects of J. S. Bach's *Well-Tempered Clavier*). Symbolic models, reviewed elsewhere,[4] have taken into consideration a number of factors in assigning key, including scales that contain the tones of the sample, and the presence of such cues to key as the tonic-fifth, tonic-third, and tonic—leading-tone intervals, characteristic tone sequences, and cadences. Some effort has been made to take phrasing, melodic accent, and rhythm into account. More recently proposed symbolic models[6,7] have made advances in both computational and music-analytic sophistication.

Concurrently, neural network models have provided an alternative subsymbolic approach.[8,9] In these, the input sample typically gives rise to activation levels of units associated with different keys. Thus, these models return graded measures of key strength. The problem of representing these in a way that takes into account the distances between keys, however, remains unsolved. In addition, little has been done to compare the output of key-finding algorithms to perceptual judgements, which may change over time as the cues to key become more or less clear and as the music may modulate to other keys. For the most part, the model's output has been compared with the composer's key signature.

Both these issues were addressed by a key-finding algorithm developed using the K-K profiles.[4] The input to this algorithm was a 12-dimensional vector specifying the total durations of the twelve chromatic scale tones in the musical selection to which the key was to be assigned. One application was to Bach's *C Minor Prelude* from Book II of the *Well-Tempered Clavier*. It was treated on a measure-by-measure basis. Each input vector was projected onto the toroidal map of keys using the phases of the two strong Fourier components. Two musical experts gave quantitative judgements of key strength in each measure of the piece, which were also projected onto the map. The key-finding algorithm followed the same pattern of modulations found by the experts, suggesting the approach lends itself to tracing dynamic changes in key as modulations occur. One obvious limitation of this approach is that it ignores the order of the tones in the input sample. This potentially important cue for key led to the development of the tone transition model described below.

## Tone transitions and key-finding

The order in which tones are played may provide additional information that is useful for key-finding. This is supported by studies on both tone transition probabilities[10,11] and perceived stability of tone pairs in tonal contexts.[4,13] In samples of compositions by Bach,

Beethoven, and Webern, only a small fraction of all the possible tone transitions were actually used (the fractions were 23, 16, and 24 per cent, respectively).[10] Furthermore, in a sample of 20 songs by Schubert, Mendelssohn, and Schumann, there was an asymmetry in the transition frequencies in the sense that certain tone transitions were used more often than the same tones in the reverse temporal order.[11] For instance, the transition B-C was used 93 times, whereas the transition C-B was used only 66 times. A similar asymmetry was found in studies on perceived stability of tone pairs in a tonal context.[4,13] After the presentation of a tonal context, tone pairs that ended with a tone that was high in the tonal hierarchy were given higher ratings than the reverse temporal orders. For instance, in the context of C major, the ratings for the transitions B-C and C-B were 6.42 and 3.67, respectively.

Determining tone transitions in a piece of polyphonic music is not a trivial task, especially if one aims at a representation that corresponds to perceptual reality. Even in a monophonic piece, the transitions can be ambiguous in the sense that their perceived strengths may depend on the tempo. Consider, for example, the tone sequence C4-G3-D4-G3-E4, where all the tones have equal durations. When played slowly, this sequence is heard as a succession of tones oscillating up and down in pitch. With increasing tempi, however, the subsequence C4-D4-E4 becomes increasingly prominent. This is because these tones segregate into one stream due to the temporal and pitch proximity of its members, separate from G3-G3. With polyphonic music, the ambiguity of tone transitions becomes even more obvious. Consider, for instance, the sequence consisting of a C major chord followed by a D major chord, where the tones of each chord are played simultaneously. In principle, this passage contains nine different tone transitions. Some of these transitions are, however, perceived as stronger than others. For instance, the transition G–A is, due to pitch proximity, perceived as stronger than the transition G–D.

It seems, thus, that the analysis of tone transitions in polyphonic music should take into account principles of auditory stream segregation.[14] Furthermore, it may be necessary to code the presence of transitions on a continuous instead of a discrete scale. In other words, each transition should be associated with a strength value instead of just coding whether that particular transition is present or not. Below, a dynamic system that embraces these principles is described. In regard to the evaluation of transition strength, the system bears a resemblance to a proposed model applying the concept of apparent motion to music.[15]

## Tone transition model

Let the piece of music under examination be represented as a sequence of tones, where each tone is associated with pitch, onset time, and duration. The main idea of the model is the following: given any tone in the sequence, there is a transition from that tone to all the tones following that particular tone. The strength of each transition depends on three factors: pitch proximity, temporal proximity, and the duration of the tones. More specifically, a transition between two tones has the highest strength when the tones are proximal in both pitch and time and have long durations. These three factors are included in the following dynamic model.

### Representation of input

The pitches of the chromatic scale are numbered consecutively. The onset times of tones having pitch $k$ are denoted by $t_{ki}^{n}$, $i = 1, \ldots, n_k$, and the offset times by, $t_{ki}^{f}$, $i = 1, \ldots, n_k$, where $n_k$ is the total number of times the $k$th pitch occurs.

### Pitch vector $\mathrm{p}(t) = (p_k(t))_k$

Each component of the pitch vector has nonzero value whenever a tone with the respective pitch is sounding. It has the value of 1 at each onset at the respective pitch, decays exponentially after that, and is set to zero at the tone offset. The time evolution of $\mathbf{p}$ is governed by the equation

$$\dot{p}_k = -p_k \tau_p + \sum_{i=1}^{n_k} \delta\left(t - t_{ki}^{n}\right) - p_k \sum_{i=1}^{n_k} \delta\left(t - t_{ki}^{f}\right), \tag{2}$$

where $\dot{p}_k$ denotes the time derivative of $p_k$ and $\delta(\cdot)$ the Dirac delta function (unit impulse function). The time constant $\tau_p$ has the value of $\tau_p = 0.5$ s. With this value, the integral of $p_k$ saturates at about 1 s after tone onset, thus approximating the durational accent as a function of tone duration.[16]

### Pitch memory vector $\mathrm{m}(t) = (m_k(t))_k$

The pitch memory vector provides a measure of both the perceived durational accent and the recency of notes played at each pitch. In other words, a high value of $m_k$ indicates that a tone with pitch $k$ and a long duration has been played recently. The dynamics of $\mathbf{m}$ are governed by the equation

$$\dot{\mathbf{m}} = \mathbf{p} - \mathbf{m}/\tau_m, \tag{3}$$

The time constant $\tau_m$ determines the dependence of transition strength on the temporal distance between the tones. In the simulations, the value of $\tau_m = 3$ s has been used, corresponding to typical estimates of the length of the auditory sensory input.[17–19]

### Transition strength matrix $\mathrm{S}(t) = (s_{kl}(t))_{kl}$

The transition strength matrix provides a measure of the instantaneous strengths of transitions between all pitch pairs. More specifically, a high value of $s_{kl}$ indicates that a long tone with pitch $k$ has been played recently and a tone with pitch $l$ is currently sounding. The temporal evolution of $\mathbf{S}$ is governed by the equation

$$s_{kl} = m_k \frac{1 + \mathrm{sgn}(p_i - p_k)}{2} p_l \mathrm{e}^{-(k-l)^2/\alpha^2}, \tag{4}$$

In this equation, the nonlinear term $(1 + \mathrm{sgn}(p_l - p_k))/2$ is used for distinguishing between simultaneously and sequentially sounding pitches. This term is nonzero only when $p_l > p_k$, that is, when the most recent onset of pitch $l$ has occurred more recently than that of pitch $k$.

The term $e^{-(k-l)}2/\alpha 2$ weights the transitions according to the interval size. For the parameter $\alpha$, the value $\alpha = 6$ has been used. With this value a perfect fifth gets a weight of about 0.37 times the weight of a minor second.

Dynamic tone transition matrix $\mathbf{N}(t) = (n_{kl}(t))_{kl}$

The dynamic tone transition matrix is obtained by temporal integration of the transition strength matrix. At a given point of time, it provides a measure of the strength and recency of each possible tone transition. The time evolution of N is governed by the equation

$$\dot{\mathbf{N}} = \mathbf{S} - \mathbf{N}/\tau_N, \tag{5}$$

where the time constant $\tau_N$ is equal to $\tau_m$, that is, $\tau_N = 3$ s.

To examine the role of tone transitions in key-finding, we developed two key-finding models. Model 1 is based on pitch class distributions. Model 2 is based on tone transition distributions. Below, a brief description of the models is given.

## Key-finding model 1

Model 1 is based on pitch class distributions only. Like the earlier algorithm based on the K-K profiles,[4] it does not take tone transitions into account. However, it has a dynamic character in that both the pitch vector and the pitch memory vector depend on time. It uses a pitch class vector $\mathbf{p}_c(t)$, which is similar to the pitch vector $\mathbf{p}(t)$ used in the dynamic tone transition matrix, except that it ignores octave information. Consequently, the vector has 12 components that represent the pitch classes. The pitch class memory vector $\mathbf{m}_c(t)$ is obtained by temporal integration of the pitch class vector according to the equation

$$\dot{\mathbf{m}}_c = \mathbf{p}_c - \mathbf{m}_c/\tau_d, \tag{6}$$

is correlated with the K-K profiles for each key. Alternatively, or in addition, the vectors $\mathbf{m}_c(t)$ can be projected onto the toroidal key representation using activation values as described earlier. Both approaches will be taken here.

## Key-finding model 2

Model 2 is based on tone transitions. Using the dynamic transition matrix N, it calculates the octave-equivalent transition matrix $\mathbf{N}' = (n'_{ij})_{ij}$ according to

$$n_{ij}'(t) = \sum_{\substack{i = p \bmod 12 \\ j = p \bmod 12}} n_{pq}(t). \tag{7}$$

In other words, transitions whose first and second tones have identical pitch classes are considered equivalent, and their strengths are added. Consequently, the melodic direction of the tone transition is not taken into account. To obtain estimates for the key, the pitch

class transition matrix is correlated with the matrices representing the perceived stability of two-tone transitions for each key.[4]

## Application of models 1 and 2 to the chord sequence data

The probe tone data for the ten chord sequences described earlier[2] were compared with the results of models 1 and 2 applied to those sequences. An interchord onset time of one second was used. The perceptual judgements were correlated with the K-K profiles, and these were compared with the key correlations of the two models. Model 1 results correlated highly ($r(2158) = 0.87$, $p < 0.0001$) with the experimental data. Model 2 also correlated highly ($r(1918) = 0.86$, $p < 0.0001$). (Model 2 produces an output only after the first two chords.) Model 2 contributed additional precision when combined with model 1, as shown by a multiple regression predicting the experimental data from both models ($R(2, 1917) = 0.89$, $p < 0.0001$, with both models contributing significantly at $p < 0.0001$; the standard coefficient was 0.50 for model 1 and 0.43 for model 2). Thus, the two models generally matched well the experimental results, and modelled slightly different aspects of the listeners' responses.

Figures 7.1 and 7.2 show the distributed mapping of tonality on the SOM for the listeners (Figures 7.1A and 7.2A) and model 1 (Figures 7.1B and 7.2B) for two illustrative chord sequences. (Because of issues about how best to visualize the results of model 2, we show only model 1 here.) In these representations, a single contour outlines correlations 0.70 or above, double contours outline correlations 0.80 or above, triple contours outline correlations 0.90 or above, and quadruple contours outline correlations 0.95 or above.

The sequence depicted in Figure 7.1 consists of the chords F G C F d B♭ e° A d, containing a relatively distant modulation from C major to D minor. After the first chord, F, appeared, the listeners and the model had a clear focus on the key of F major in which the chord is I. When the second chord, G, was sounded, listeners apparently interpreted this as a IV–V progression in C major, resulting in a focus near that key. The model, however, did not find this focus until the tonic triad, C, in the third position. For the fourth and fifth chords, F and d, the focus of both listeners and model remained in the region of C major but was shifted somewhat towards F major and D minor. The sixth chord, B♭, which is contained in D minor but not C major, shifted the focus farther towards D minor. The seventh chord, e° , greatly weakened the sense of key for both listeners and model; a diffuse elongated region of weak activation was found. With the last two chords, A and d, the final focus on D minor (in which these are V–I) was arrived at.

The sequence depicted in Figure 7.2 consists of the chords d° G c A♭ F D♭ b♭ E♭ A♭, containing a modulation from C minor to A♭ major, a modulation to what is considered a relatively close key. Both listeners and model yielded an extremely weak activation pattern for the first two chords, d° and G, although the focus for listeners after the second chord was weakly near C minor. The third chord, c, clarified this focus near C minor for both listeners and model, where it remained for the fourth and fifth chords, A♭; and f. The sixth chord, E♭, which is the tonic triad in the relative major of C minor (E♭ major) produced a strong focus on E♭ major for listeners and a shift in that direction for the model. The seventh chord, B♭, diffused this focus for both listeners and model before the

new key of A♭ major clarified with the E♭ and A♭ chords in the last two positions (which are V–I in the A♭ major key).

## Conclusion

Experimental studies suggest that the sense of tonality undergoes dynamic and subtle changes when a listener hears music. The sense of key develops and strengthens as certain cues appear, then may weaken or shift to a new key as subsequent events are sounded. An important step in understanding this process is a suitable means of representing such changes. Towards this end, we have developed a spatial representation based on psychological data. The distributed map of the SOM provided a visually accessible representation of these subtle dynamic changes. The key-finding models described here suggested that listeners' sense of key can be modelled quite well using tone distributions and tone transitions. Two models were formulated to incorporate various psychological phenomena, including durational accent as a function of tone duration, the duration of sensory memory, temporal-order asymmetries, and pitch streaming as a function of the distances between tones in pitch and time. The models accounted well for the results of an experiment with relatively simple and short chord sequences. Further tests of such models with more complex, extended, and musically realistic materials may point to additional factors, such as rhythm, meter, phrasing, and form, that may also influence the sense of tonality.

## References


1. **Krumhansl, C. L. and R. N. Shepard** (1979) Quantification of the hierarchy of tonal functions within a diatonic context. *J. Exp. Psychol. Hum. Percept. Perform.* 5, 579–94.
2. **Krumhansl, C. L. and E. J. Kessler** (1982) Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychol. Rev.* 89, 334–68.
3. **Krumhansl, C. L. and P. Toiviainen** (2000) Dynamics of tonality induction, a new method and a new model. In C. Woods, G. Luck, R. Brochard, F. Seddon, and J.A. Sloboda (eds) *Proceedings of the Sixth International Conference on Music Perception and Cognition.* Keele, UK: Keele University, pp. 1504–13.
4. **Krumhansl, C. L.** (1990) *Cognitive Foundations of Musical Pitch.* New York: Oxford University Press.
5. **Kohonen, T.** (1997) *Self-organizing Maps.* Berlin: Springer-Verlag.
6. **Vos, P. G. and E. W. Van Geenen** (1996) A parallel-processing key-finding model. *Music Percept.* 14, 185–223.
7. **Vos, P. G. and M. Leman** (Eds) (2000) Tonality induction. Special issue of *Music Percept.*
8. **Leman, M.** (1995) A model of retroactive tone-center perception. *Music Percept.* 12, 439–71.
9. **Bharucha, J. J.** (1999) Neural nets, temporal composites, and tonality. In D. Deutsch (ed.) *The Psychology of Music,* 2nd edn. San Diego, CA: Academic Press, pp. 413–40.
10. **Fucks, W.** (1962) Mathematical analysis of the formal structure of music. *I R E Trans. Information Theory* 8, 225–8.

11. **Youngblood, J. E.** (1958) Style as information. *J. Music Theory* 2, 24–5.
12. **Knopoff, L.** and **W. Hutchinson** (1978) An index of melodic activity. *Interface* 7, 205–29.
13. **Krumhansl, C. L.** (1979) The psychological representation of musical pitch in a tonal context. *Cognit. Psychol.* 11, 346–74.
14. **Bregman, A. S.** (1990) *Auditory Scene Analysis.* Cambridge, MA: M.I.T. Press.
15. **Gjerdingen, R. O.** (1994) Apparent motion in music? *Music Percept.* 11, 335–70.
16. **Parncutt, R.** (1994) A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Percept.* 11, 409–64.
17. **Darwin, C. J., M. T. Turvey,** and **R. G. Crowder** (1972) An auditory analogue of the sperling partial report procedure, evidence for brief auditory storage. *Cognit. Psychol.* 3, 255–67.
18. **Fraisse, P.** (1982) Rhythm and tempo. In D. Deutsch (ed.) *The Psychology of Music,* 1st edn. San Diego, CA: Academic Press, pp. 149–80.
19. **Treisman, A. M.** (1964) Verbal cues, language, and meaning in selective attention. *Am. J. Psychol.* 77, 206–19.

CHAPTER 8

# LEARNING AND PERCEIVING MUSICAL STRUCTURES: FURTHER INSIGHTS FROM ARTIFICIAL NEURAL NETWORKS

BARBARA TILLMANN, JAMSHED J. BHARUCHA, AND EMMANUEL BIGAND

## Abstract

Artificial neural networks expand our understanding of the acquisition and representation of knowledge of our environment and the influence of this knowledge on perception. In the present chapter, we review applications of artificial neural networks to music cognition, notably to the learning and perceiving of musical structures. We present a hierarchical self-organizing model that learns basic regularities of the Western tonal system by mere exposure and simulates tonal acculturation. After learning, the model simulates a variety of experiments dealing with the processing of tone, chord, and key relationships. It provides a parsimonious account of these data sets by postulating activation as the unifying mechanism underlying various cognitive tasks. The modelling of music cognition presented in this chapter is restricted to behavioural data. Nevertheless, the computational processes are based on fundamental neural constraints. Future developments of artificial networks simulating neuropsychological cases and establishing direct links to neural correlates will contribute to enhance our understanding of mechanisms of music perception.

**Keywords:** Music cognition; Connectionist modeling; Self-organization; Tonal acculturation

## Introduction

Artificial neural networks are rooted in the biological structure of the neural system: they consist of a set of artificial neurons (units) that are linked via synaptic connections of different strengths. A unit does not represent one neuron, but simulates a population of neurons. The goal of artificial networks is not to describe neural anatomy and physiology, but to be founded on neural principles in order to simulate different levels of perceptual and cognitive processing. Recent developments in artificial networks establish correspondences to neural mechanisms by relating components of a model to brain structures,

creating artificial lesions or suggesting biologically plausible networks (i.e. neurobiological models with autonomous and adaptive behaviour).

The present chapter reviews some artificial neural networks that have been proposed in auditory and music perception. Neural nets are helpful in understanding how we learn musical patterns by mere exposure, how these patterns might be represented and how this knowledge arising from acculturation influences perception. To date, applications of networks for music cognition have been restricted to behavioural data. However, future developments in combination with neuropsychological cases and brain imaging methods should make it possible to establish links to brain structures and neural circuitry for the investigation of music perception. These developments are in progress in other domains (e.g. language perception, visual perception, memory). In these domains, artificial neural networks are used not only to simulate cognitive processes and behavioural data, but to postulate correspondences to cerebral structures, their interactions and the effects of brain damage. For behavioural data, a variety of networks have been proposed in connectionist psycholinguistics for language acquisition, sentence processing, reading, and production (cf. Ref. 1 for a recent review). Well-known networks developed by McClelland and colleagues on visual word and speech perception[2–5] simulate a variety of behavioural data and generate numerous hypotheses for further behavioural experiments. The creation of 'artificial lesions' in neural net models simulating behaviour of healthy participants have enabled the simulation of neuropsychological cases, and have provided further insights into the normal functioning of the brain, possible representations of knowledge and the modularity of subprocesses (see Ref. 6 for a review). Simulations in computational neuropsychology have been proposed by eliminating network units or by weakening connections in order to model clinical cases of aphasia,[7–9] dyslexia[10,11] or semantic impairments.[12,13] For example, comparing the behaviour of damaged connectionist models to double-dissociations observed in the behaviour of brain-injured patients suggests that behavioural deficits arise from impairments of two types of lexical information (semantic and phonological) and explains past tense formation of English verbs without rules or symbolic mechanisms.[14,15]

In what follows, we briefly review some basics of neural net modelling and summarize the regularities underlying Western tonal music. Next, we present our recent work with Self-Organizing Maps (SOMs) as an example of neural networks providing further insights to music cognition. We then suggest some possible, though still speculative, future directions in which neural nets can help us further to investigate the processing of music, notably by simulating neuropsychological cases and establishing links to neural structures and circuitry.

## Some basics of artificial neural networks

A principal advantage of artificial neural networks is their capacity to adapt in such a way that representations, categorizations or associations between events can be learned. Connectionist models have the characteristic that rules governing the materials are not explicit, but emerge from the simultaneous satisfaction of multiple constraints represented

by the connections; these constraints can be learned by repeated exposure. This learning process is either supervised by an external teaching exemplar or unsupervised via passive exposure. In the present review, we focus on unsupervised learning algorithms that extract statistical regularities and encode events that often occur together. The competitive learning algorithm[16] is an unsupervised learning algorithm that provides the basis for learning in self-organizing maps[17,18] and ART (adaptive resonance theory) networks.[19,20] Unsupervised learning algorithms seem to be well suited to model music cognition, as acculturation to musical structures presumably occurs without supervision in listeners.

In an artificial neural network, a set of units is linked via synaptic connections of different strengths. The units are generally arranged into layers, with an input layer coding the incoming information. When a stimulus is presented to the model, the input units are activated. This activation is sent via the connections to units in other layers, and the strength of the transmitted activation is controlled by the strengths of the connections (i.e. weights). In parallel with biological networks, the learning process is defined as a modification of the connection weights. For the learning process, a set of training stimuli is presented repeatedly to the network. The network adapts its connections in such a way that it becomes sensitive to the underlying correlational structure between events of the training set. Before learning starts, all connections are generally set to random values, a state that represents the 'ignorance' of the model. Over the course of learning, the neural net units gradually become sensitive to different input events or categories. With the competitive learning algorithm, the specialization takes place by competition among the units.[16] When an input is presented to the network, the unit that is best able to represent it wins the competition and is allowed to learn the representation of this input even better. The unit's response will be subsequently stronger for this same input pattern and weaker for others. In a similar way, other units learn to respond selectively to other input patterns. The competitive learning algorithm is generalized in SOMs that contain a spatial layout of the units. In SOMs, the winning unit and its neighbour units learn. This learning process leads to topographic mappings between input data and neural net units on the map: For two similar input patterns, map units that respond maximally are located near each other. This mirrors principles of cortical information processing, such as spatial ordering in sensory processing areas (i.e. somatosensory, vision, audition). In the primary visual cortex, the orientation of stimuli to which cells respond best changes in an orderly fashion with nearby cells responding best to similar orientations.[21] In the auditory system, tonotopic organization is found at almost all major stages of processing (i.e. inner ear, auditory nerve, cochlear nucleus, auditory cortex), and the auditory cortex displays a tonotopic organization in which nearby cells respond best to similar frequencies.[22–25]

According to Kohonen,[26] the SOM algorithm parallels the adaptive processes underlying plasticity in biological networks: the theoretical definition of neighbourhood results in learning that is similar in consequence to lateral connections modelled by chemical processes. The principles of self-organization are general across domains, but input representations are domain specific. The SOMs have been applied to a variety of domains, including visual perception,[27] semantic processing[28] or auditory perception.[29,30] Mercado *et al.*[30] used an SOM to simulate experience-dependent reorganization of representation

maps in auditory cortex. The network architecture was constrained by properties observed in auditory cortex and which include the modulating influences of the basal forebrain. The simulations allowed the authors to formulate hypotheses on neuromodulatory effects in auditory cortical plasticity and to suggest future simulations including other neural regions (e.g. hippocampus) in order to model effects on learning and memory. Recently, we used the SOM algorithm to simulate an example of cognitive plasticity, notably the cognitive capacity to extract underlying regularities and to become sensitive to musical structures and regularities via implicit and unsupervised learning processes.[31]

## An overview of regularities underlying Western tonal music

Western tonal music is an example of a highly structured system in our environment. It contains a set of regularities (i.e. regularities of cooccurrence, frequency of occurrence, and psychoacoustic regularities) based on a limited number of elements. The Western tonal system that contains a three-level organization is based on a set of 12 tones (C-C#-D-D#-E-F-F#-G-G#-A-A#-B). The 12 tones are combined into subsets of seven tones, the scales. For each tone of a scale, chords (e.g. major or minor) are constructed by adding two notes—creating a second order of musical units. The tones in a scale and their associated chords define a third level of organization in the tonal system: the keys. Keys sharing numerous tones and chords are harmonically related, and the strength of harmonic relationship depends on the number of shared events. The seven tones that belong to the underlying scale are in-key tones, the remaining five tones are out-of-key tones. Similarly for chords, in-key chords are differentiated from out-of-key chords. Hierarchies of functional importance exist among tones and chords belonging to a given key: For example, chords built on the first, fifth, and fourth scale degrees (referred to as tonic, dominant, and subdominant, respectively) have more important functions than chords built on other degrees. From a psychological point of view, the hierarchically important events of a key act as stable cognitive reference points[32] to which others are anchored.[33] As the tonal system is based on a restricted set of events, the same event can define an in-key or an out-of-key event or can take different levels of functional importance (i.e. tonal stability)—always depending on the established key.

The three levels of musical units occur with strong regularities of cooccurrence. Tones and chords belonging to the same key are more likely to cooccur in a musical piece than tones and chords belonging to different keys. Changes between keys are more likely to occur between closely related keys than between less related ones. Also within-key hierarchies are strongly correlated with the frequency of occurrence of tones and chords in Western musical pieces. Tones and chords that have more important music-theoretic functions in a given key are used with higher frequency (and longer duration).[34–36]

Despite the complexity of the system, Western tonal listeners become sensitive to the musical structures by mere exposure to musical pieces obeying this system of regularities.[35–37] As the number of opportunities to listen to musical pieces is so great in everyday life, this implicit knowledge is plausibly learned by the same implicit processes as those investigated in the laboratory.[38–41] Behavioural and neurophysiological research has provided evidence that even nonmusician listeners are sensitive to the context dependency of musical events. The

implicit knowledge embodies the functions of tones and chords in a key,[35,42–44] the relations between different keys,[45–48] and the change in tonal functions depending on the context.[49–53] The influence of this internalized representation has been reported for musical memory,[54–57] musical expectancies,[52,58–61] and with electrophysiological and brain imaging methods.[62–69]

## Learning musical structures

The Western tonal system embodies strong regularities to which listeners become sensitive via mere exposure to musical pieces.[35–37] We proposed to simulate implicit learning processes of tonal regularities with SOMs and used the trained network to simulate experimental data on music perception via activation spreading through the learned representation.[31]

In these simulations, a hierarchical network with two SOMs was defined (i.e. a multi-layer hierarchical self-organizing map HSOM[70]). The input layer consisted of 12 units coding the 12 pitch classes. Each unit of this layer was connected to a second layer (a self-organizing map) that in turn was connected to a third layer (a second self-organizing map). Before learning started, the units of two layers were fully interconnected with connection weights set to random values. Four learning simulations that differed in training material and richness of input coding were realized. The first SOM was trained with a set of chords presented separately. The second SOM was trained with either groups of chords that belonged to the same key or short chord sequences. The short chord sequences represented more realistic harmonic material since they respected chord transitions and frequencies of occurrence in Western tonal music. In addition, each chord was weighted according to recency in order to mimic memory decay in short-term memory. All training material was coded by either a sparse input pattern (i.e. a unit is activated (set to '1') when a note is present in the chord, and set to '0' otherwise) or a richer input pattern that coded a chord as the weighted sum of subharmonics of the present pitch classes (as proposed in Refs 71 and 72).

Independent of training material and input coding, the network became sensitive to regularities underlying the musical material (i.e. chords, chord sets, chord sequences). The units of the first SOM became specialized in the detection of chords and the units of the second SOM in the detection of keys. Both layers showed a topological organization of the specialized units. In the chord layer, units representing chords that share tones (or subharmonics) were located close to each other on the map, but chords not sharing tones were not represented by neighbouring units. In the key layer, the units specialized in the detection of keys were organized in a circle: keys sharing numerous chords and tones were represented close to each other on the map and the distance between keys increased with decreasing number of shared events. The organization of key units reflects the music theoretic organization of the cycle of fifths: the more the keys are harmonically related, the closer they are on the cycle (and on the network map). After training with the rich input coding, also the chord layer showed a more global organization reflecting harmonic relationships: chords from one side of the circle of fifths are represented on one half of the map, the other side of the circle on the other half of the map.

After learning, the connection weights were not at random values any longer, but reflected the regularities of cooccurrence. With the sparse input coding, a chord unit was linked to three tone units, tones that are components of this chord, and a key unit was

linked to six chord units, chords that belong to the key. In addition, for the simulations using chord sequences, the strength of the connections reflected the frequency of occurrence of chords in the sequences, and the tonic chord in a given key had stronger links than the other chords. For the simulations with chord sets, all connections between chord and key units had the same strength. In sum, specialized representational units were formed for combinations of musical events (tones, chords) that occur with great regularity. The self-organization leads to a hierarchical encoding in which tones occurring together are represented by chord units, and similarly, chords occurring together are represented by key units. Interestingly, the learned connections mirrored the pattern of connections that had been hardwired in a previously proposed music perception network (MUSACT[73]).

The trained HSOM networks were used as feedforward and reverberation systems. In a feedforward system, activation consisted of bottom-up information only. Consequently, activation of chord units reflected the number of component tones shared with the input: the more tones are shared, the stronger the unit is activated. In a reverberation system, phasic activation spreads between the units of the three layers until an equilibrium is reached.[73] After reverberation, the activation pattern reflects top-down influences of learned, schematic structures, and respects tonal relations: the activation of major chord units, for example, decreased monotonically with increasing distance on the circle of fifths. After reverberation, top-down processes overwrote influences of coding richness and the imposed activation pattern was analogous for sparse and rich input coding. Finally, activation patterns after reverberation for the four trained HSOM networks showed strong correlations with activation patterns of MUSACT ($0.984 < r < 0.999$). This outcome suggests that MUSACT's hard-wired structure can emerge from self-organization.

The hierarchical self-organizing map thus manages to learn Western pitch regularities. The input layer of the network was based on pitch classes, the 12 tones of the chromatic scale. In Western tonal music, these pitch classes are repeated over octaves in different pitch heights. The phenomenon of octave equivalence (i.e. tones separated by an octave are perceived as similar in pitch) permits the presentation of the tonal system as based on 12 events. The abstract pitch-class coding in the network was based on previous simulations of octave category learning. An SOM learned octave-equivalent pitch classes via mere exposure to spectral representations of tones.[74] In addition, it has also been shown that neural nets are able to learn the extraction of pitch height from frequency,[75–77] and to transform a spectral representation of an acoustic source into a spatial distribution of pitch strengths.[75] These models represent the application of neural nets to the learning of low-level processes in auditory perception (pitch extraction, octave equivalence). The HSOM model with its three organizational layers of tones, chords, and keys can be conceived of as subsequent to these simulated phases of auditory preprocessing.

In the music perception domain, other neural net models have been proposed to simulate the learning of higher-level organizations. In contrast to the HSOM network presented above, these models focused on either one or two organizational layers in music perception, as for example, models of chord classification[78] or melodic sequence learning.[79–81] Griffith[82] and Leman[72,83] used SOMs to simulate the learning of key representations based on either melodies,[82] chords[72] or recordings of real musical pieces which are preprocessed by an auditory module.[83] After training, the specialized units of the SOMs showed a topological

organization of the detected key centres that reflects the harmonic distances between keys (as on the cycle of fifths) with the distances on the map. Leman's work provided evidence that higher order units of Western music (i.e. tonal centres) can be learned via passive exposure to a rich acoustic input. These models focused on the learning of tonal centres, but did not account for relations between tones, chords, and keys. Gjerdingen[84] proposed to simulate a more complex aspect of musical learning: the perception of musical style. An ART network[19,20] was trained to categorize and memorize stylistic features of Mozart pieces and the learned feature patterns were compared to music theoretic concepts.

## Perceiving musical structures

In order to be compelling, a cognitive model of Western music should not only simulate the internalization of Western pitch regularities via mere exposure, but should also simulate the behaviour of listeners after having adapted to Western tonal music. Numerous connectionist models in music have not been developed to reflect cognitive processes and their predictions are not directly tested with behavioural data observed with human listeners. In Leman's work,[72] for example, musical pieces were presented to the network and tonal centre activation was compared with music theory analyses. In Ref. 83, a context followed by a chord was presented to the network and changes in activation patterns were correlated with probe-chord judgements of human listeners.[36] The activation changes simulated indirectly the differences in subjective judgements, but the network was not able to generate predictions for the chords themselves. A different approach to compare performance of network and human listeners has been proposed by Stevens[85] and Stevens and Latimer.[86,87] A network was trained to distinguish musical excerpts (standard) from modified excerpts. Results of simulations were compared with performance of human listeners and the influence of musical expertise was modeled by length of training. However, the model was explicitly restricted to a recognition task without the goal to simulate other tasks or a more general representation of tonal knowledge. Krumhansl *et al.*[80] compared performance of human listeners and SOM networks for the specific case of melodic expectation in Finnish spiritual folk hymns. Expertise was simulated by contrasting two SOM networks trained with either hymns or different finish folk songs. Activation patterns of the hymn SOM correlated more strongly with melodic continuation judgements of human expert listeners and suggested that this SOM became sensitive to tone distributions and tone transitions underlying the Finnish hymns.

In Tillmann *et al.*,[31] the HSOM network was tested for its capacity to simulate a variety of empirical data about perceived relations between and among tones, chords, and keys in Western tonal music. Simulations were run with the experimental material used in the empirical studies. The underlying rationale of the simulations was to test activation (spreading through a representation of tonal knowledge) as a mechanism unifying a range of cognitive tasks. The experimental material was presented to the model and the activation levels of network units were interpreted as levels of tonal stability.[a] The more a unit (i.e. a chord

[a] For event sequences, activation due to each event is accumulated and weighted according to recency.[73] The total activation of a unit is thus the sum of the stimulus activation, the phasic activation accumulated during reverberation and the decayed activation due to previous events.

unit, a tone unit) is activated, the more stable the musical event is in the corresponding context. For the experimental tasks, it was hypothesized that the level of stability affects performance (e.g. a more strongly activated, stable event is more expected or judged to be more similar to a preceding event).

For perceived relations between chords, the model succeeded in simulating data obtained with similarity judgements,[49,53] recognition memory,[49] harmonic priming[52,59,88,89] and electrophysiological measures.[69] Harmonic priming and ERP data provided evidence for the development of harmonic expectations in a prime context that then influence the processing of a target chord: a target is processed more slowly and less accurately and evokes a larger late-positivity component LPC, when it is unrelated to the context than when it is related. When the experimental material is presented to the network, activation levels of chord units mirrored patterns of processing in the priming task (e.g. with higher activation for chord units representing facilitated targets), and activation changes in chord units after the target mirrored the amplitude of the LPC (e.g. with stronger activation changes for distant key targets). Both human listeners and network are sensitive to changes in chord stability caused by key contexts. Further behavioural experiments have shown that human listeners perceive keys underlying a given context, detect modulation (i.e. temporary changes in key) occurring in a musical excerpt and have implicit knowledge of distances between keys.[46,47,90] A further set of HSOM simulations showed that activation levels of key units and activation changes over time mirror listeners' behavioural data. For example, the changes of activation patterns in key units were more important when the excerpt modulated to distantly related keys.[46] As activation accumulates in the network over time, the model also tracks the key changes in a sequence and simulates a dynamic aspect of tonality sensation.

The HSOM network also simulated behavioural data on the perceived relations between tones, even if it was trained with chords only and not with melodies. Notably, the activation levels of tone units simulate the stability profiles obtained in probe-tone ratings,[90] the patterns of similarity ratings[32] and memory for melodies.[56] For example, when human listeners rate the similarity of a tone pair presented after a tonal context, the ratings reflect the differences in tonal functions and show a perceived asymmetry with higher similarity for a pair ending on a stable tone. The activation levels and activation changes of the tone units mirrored these data sets.

Overall, the simulations showed that activation in the trained self-organizing network mirrored data of human participants in tonal perception experiments. This outcome suggests that the level of activation in tone, chord, and key units is a single unifying concept for human performance in different tasks.

## General discussion

The HSOM network was presented as an application of artificial neural networks to further our understanding of learning and perceiving music. Based on self-organization (i.e. a general unsupervised learning algorithm), the structure of the network adapted to tonal regularities through repeated exposure to musical material. The network combines three levels of music perception which can be placed on the top of networks simulating lower perceptual processing steps (pitch extraction, octave equivalence). After training, the network

simulates a large set of experimental data on perceived relations between tones, chords, and keys. These simulations proposed activation as a unifying mechanism for different cognitive tasks. A context activates the network and the activation state then influences the perception (i.e. expectation, perceived similarity, memory). The HSOM network provided a model not only for the learning of implicit knowledge about the tonal system, but also for a low-dimensional and parsimonious representation of tonal knowledge. For example, the network proposes a parsimonious account of the contextual dependency of musical functions of an event and the changes associated with different keys. One single event has not multiple instances of representation, but its corresponding function in a given key emerges from activation spreading in the system. As a consequence, events representing prototypes and anchor points in a given key do not have to be stored separately, but emerge from the activation levels. A further emerging property of the network relates to the identification of a key; the underlying key does not need to be inferred employing separate processing steps, but emerges from the activation pattern in the key units.

The network also offers a framework for generating new predictions for further behavioural studies on music perception. These studies and associated simulations could investigate the perception of tones, chords, and keys. The experimental material will be presented to the network and on the basis of the activation levels precise hypothesis can be derived for the performance of human listeners. For example, the studies could be related to key identification (e.g. number of notes necessary to establish a key, disturbing effect of an unrelated event), to key modulation (e.g. how long the trace of a key remains) or to the eventual link between activation decay and musical short-term memory span.

In the following, we point out two limitations of the HSOM network and discuss future developments of the model. A first limitation is the restriction to pitch structures only, and the second one is that the network is more abstract than brain structures and neural circuitry.

The simulations have focused on how regularities of the pitch dimension may be internalized through passive learning processes. Although pitch is the most obvious form-bearing dimension of Western tonal music, temporal regularities also contribute to listeners' perceptual experience. Temporal regularities include the sensation of meter (sensation of a regular succession of strong and weak beats superimposed over an isochronous pulse) and patterns of intervals creating rhythms perceived against the metrical background. Two frameworks have been proposed to account for the respective contributions of regularities of pitch and time in the perception of musical events.[91] A single-component model[92] suggests that temporal and harmonic accents are not processed independently, but are integrated and together guide the attention of the listener during the unfolding musical piece.[51,54,93,94] Based on neuropsychological cases (with double-dissociations of amelodia without arhythmia and vice versa), and experimental data,[91,95–98] Peretz and Kolinsky[91] proposed a two-component model suggesting independence between temporal and nontemporal information processing which are integrated only at a later stage. Two types of neural net architectures can be used to mimic these two theoretical approaches and to account for learning and perception of regularities in pitch and time. A single-component neural net may learn the two regularities conjointly without a supplementary integrative step, by adapting input codings of metrical information as proposed by Berger and Gang[99,100] or by Stevens and Wiles.[101] A two-component network may learn separately regularities of pitch

(as in the HSOM[31]) and time (as proposed in Ref. 102), followed by a later integration step in order to simulate the development of expectations. The network modelling shall allow further investigation of how pitch and time are combined—a question that arises for learning, perception, and memorization (cf. Ref. 103 for visual stimuli). In addition, networks integrating both pitch and time create the possibility to study neuropsychological cases. The goal would be to simulate neuropsychological cases with amelodia without arythmia (and vice versa) by creating artificial lesions to the networks of normal perception (see Ref. 6 for applications in other domains).

Even if the HSOM model is conceived on an abstract level and not in relation to brain structures or neural circuitry, it can be a source of inspiration to generate hypotheses about the neural circuitry underlying music perception. For example, the self-organizing algorithm conforms to principles of cortical information processing, such as spatial ordering in sensory cortex or tonotopic organization in auditory cortex. The outcome of the simulations, together with aspects of cortical organization, leads to the question whether there exist higher order maps such as a tonotopic organization of key centres.

In the domain of music perception, simulations are restricted to learning and perception in combination with behavioural data. Future developments of networks, however, should also have the goal of suggesting simulations that make the bridge to actual neural structures and circuitry in the brain. In other domains, connectionist models in computational neurosciences are created on the bases of existing knowledge about cerebral structures and neural circuitry.[104,105] In visual perception, Otto *et al.*[105] investigated invariant object recognition (same object despite changes in size or location). The creation of the network was based on a set of neurophysiological constraints known about the ventro-lateral and dorsal-parietal pathways associated with the processing of object identification ('what') and spatial localization ('where'). Based on the simulations' outcome, the authors derived new predictions about interactions between the neural pathways and motivated neurophysiological, biological and behavioural experiments. Recently, the hypothesis of two types of pathways ('what', 'where') has been proposed for auditory perception.[106,107] Once more details are known about their functioning, simulations following the line proposed by Otto *et al.*[105] will enhance the understanding of their interaction and stimulate further explorations.

## References

1. **Christiansen, M. H.** and **N. Chater** (2001) *Connectionist Psycholinguistics.* Westoirt, CT: Ablex.
2. **McClelland, J. L.** and **D. E. Rumelhart** (1981) An interactive activation model of context effects in letter perception: part 1. An account of basic findings. *Psychol. Rev.* 86, 287–330.
3. **Rumelhart, D. E.** and **J. L. McClelland** (1982) An interactive activation model of context effects in letter perception. part 2. *Psychol. Rev.* 89, 60–94.
4. **McClelland, J. L.** and **J. L. Elman** (1986) The TRACE model of speech perception. *Cognit. Psychol.* 18, 1–86.
5. **Seidenberg, M. S.** and **J. L. McClelland** (1989) A distributed, developmental model of word recognition and naming. *Psychol. Rev.* 96, 523–68.

6. **Cohen, G., R. A. Johnston,** and **K. Plunkett** (2000) *Exploring Cognition: Damaged Brains and Neural Nets: Readings in Cognitive Neuropsychology and Connnectionist Modelling.* Hove, East Sussex, UK: Psychology Press.
7. **Dell, G. S.** (1986) A spreading-activation theory of retrieval in sentence production. *Psychol. Rev.* 93, 283–321.
8. **Dell, G. S., M. F. Schwartz, N. Martin, E. M. Saffran,** and **D. A. Gagnon** (1997) Lexical access in aphasic and nonaphasic speakers. *Psychol. Rev.* 104, 801–38.
9. **French, R. M.** (1997) Selective memory loss in aphasics: an insight from pseudo-recurrent connectionist networks. In J. Bullinaria *et al.* (eds) *Connectionist Representations.* London: Springer, pp. 183–95.
10. **Shallice, T.** and **D. Plaut** (1992) From connectionism to neuropsychological syndroms. In J. Alegria *et al.* (eds) *Analytic Approaches to Human Cognition.* Amsterdam: Elsevier Science, North Holland, pp. 239–58.
11. **Harley, T. A.** (1993) Connectionist approaches to language disorders. *Aphasiology* 7, 221–49.
12. **Farah, M. J.** and **J. L. McClelland** (1991) A computational model of semantic memory impairment: modality specificity and emergent category specificity. *J. Exp. Psychol. Gen.* 120, 339–57.
13. **Devlin, J. T., L. M. Gonnerman, E. S. Andersen,** and **M. S. Seidenberg** (1998) Category-specific semantic deficits in focal and widespread brain damage: a computational account. *J. Cognitive Neurosci.* 10, 77–94.
14. **Joanisse, M. F.** and **M. S. Seidenberg** (1999) Impairments in verb morphology after brain injury: a connectionist model. *Proc. Nat. Acad. Sci.* 96, 7592–7.
15. **Lavric, A., D. Pizzagalli, S. Forstmeier,** and **G. Rippon** (2001) Mapping dissociations in verb morphology. *Trends in Cognitive Sci.* 5, 301–8.
16. **Rumelhart, D. E.** and **D. Zipser** (1985) Feature discovery by competitive learning. *Cognitive Sci.* 9, 75–112.
17. **von der Malsberg, C.** (1973) Self-organizing of orientation sensitive cells in the striate cortex. *Kybernetic* 14, 85–100.
18. **Kohonen, T.** (1995) *Self-Organizing Maps.* Berlin: Springer.
19. **Grossberg, S.** (1970) Some networks that can learn, remember and reproduce any number of complicated space-time patterns. *Studies Appl. Math.* 49, 135–66.
20. **Grossberg, S.** (1976) Adaptive pattern classification and universal recoding: I. parallel development and coding of neural feature detectors. *Biological Cybernetics* 23, 121–34.
21. **Hubel, D. H.** and **T. N. Wiesel** (1962) Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *J. Physiol.* 160, 106–56.
22. **Brugge, J. F.** and **R. A. Reale** (1985) Auditory cortex. In A. Peters and E. G Jones (eds) *Cerebral Cortex: Association and Auditory Cortices.* New York, NY: Plenum Press.
23. **Wessinger, C. M., M. H. Buonocore, C. L. Kussmaul,** and **G. R. Mangun** (1997) Tonotopy in human auditory cortex examined with functional magnetic resonance imaging. *Human Brain Map.* 5, 18–25.
24. **Howard, M. A. I., I. O. Volkov, P. J Abbas, H. Damasio, M. C. Ollendieck,** and **M. A. Granner** (1996) A chronic microelectrode investigation of the tonotopic organization of human auditory cortex. *Brain Res.* 724, 260–4.
25. **Bilecen, D., K. Scheffler, N. Schmid, K. Tschopp,** and **J. Seelig,** (1998) Tonotopic organization of the human auditory cortex as detected by BOLD-FMRI. *Hearing Res.* 126, 19–27.

26. **Kohonen, T.** (1993) Physiological interpretation of the self-organizing map algorithm. *Neural Net.* 6, 895–905.
27. **Schyns, P. G.** (1991) A modular neural network of concept acquisition. *Cognitive Sci.* 15, 461–508.
28. **Ritter, H.** and **T. Kohonen** (1989) Self-organizing semantic maps. *Biological Cybernetics* 61, 241–54.
29. **Kohonen, T.** (1988) The 'neural' phonetic typewriter. *Computer IEEE* 21, 11–22.
30. **Mercado, E. I., C. E. Myers,** and **M. A. Gluck** (2001) A computational model of mechanisms controlling experience-dependent reorganization of representational maps in auditory cortex. *Cognitive, Affective & Behavioral Neurosci.* 1, 37–55.
31. **Tillmann, B., J. J. Bharucha,** and **E. Bigand** (2000) Implicit learning of tonality: a self-organizing approach. *Psychol. Rev.* 107 (4), 885–913.
32. **Krumhansl, C. L.** (1979) The psychological representation of musical pitch in a tonal context. *Cognit. Psychol.* 11, 346–74.
33. **Bharucha, J. J.** (1984) Anchoring effects in music: the resolution of dissonance. *Cognit. Psychol.* 16, 485–518.
34. **Budge, H.** (1943) *A Study of Chord Frequencies. Teacher College.* New York: Columbia University.
35. **Francès, R.** (1958) *La perception de la musique.* Paris: Vrin.
36. **Krumhansl, C. L.** (1990) *Cognitive Foundations of Musical Pitch.* Oxford, England: Oxford University Press.
37. **Dowling, W. J.** and **D. L. Harwood** (1986) *Music Cognition.* Orlando, Florida: Academic Press.
38. **Altmann, G.T., Z. Dienes,** and **A. Goode** (1995) Modality independence of implicitly learned grammatical knowledge. *JEP-LMC* 21, 899–912.
39. **Dienes, Z., D. Broadbent,** and **D. C. Berry** (1991) Implicit and explicit knowledge bases in artificial grammar learning. *JEP-LMC* 17, 875–87.
40. **Reber, A. S.** (1967) Implicit learning of artificial grammars. *J. Verbal Learning and Verbal Behavior* 6, 855–63.
41. **Reber, A. S.** (1989) Implicit learning and tacit knowledge. *J. Exp. Psychol. Gen.* 118, 219–35.
42. **Cuddy, L. L.** and **B. Badertscher** (1987) Recovery of tonal hierarchy: some comparisons across age and levels of musical expertise. *P&P* 41, 609–20.
43. **Hébert, S., I. Peretz,** and **L. Gagnon** (1995) Perceiving the tonal ending of tune excerpts: the roles of pre-existing representation and musical expertise. *Can. J. Exp. Psychol.* 49, 193–209.
44. **Tillmann, B., E. Bigand,** and **F. Madurell** (1998) Influence of global and local structures on the solution of musical puzzles. *Psychol. Res.* 61, 157–74.
45. **Bartlett, J. C.** and **W. J. Dowling** (1980) The recognition of transposed melodies: a key-distance effect in developmental perspective. *J. Exp. Psychol. Hum. Percept. Perform.* 6, 501–15.
46. **Cuddy, L. L.** and **W. F. Thompson** (1992) Asymmetry of perceived key movement in chorale sequences: converging evidence from a probe-tone analysis. *Psychol. Res.* 54, 51–9.
47. **Cuddy, L. L.** and **W. F. Thompson** (1992) Perceived key movement in four-voice harmony and single voices. *MP* 9, 427–38.
48. **Thompson, W. F.** and **L. L. Cuddy** (1989) Sensitivity to key change in chorale sequences: a comparison of single voices and four-voice harmony. *MP* 7, 151–68.
49. **Bharucha, J. J.** and **C. L. Krumhansl** (1983) The representation of harmonic structure in music: hierarchies of stability as a function of context. *Cognit.* 13, 63–102.

50. **Bigand, E.** (1993) The influence of implicit harmony, rhythm and musical training on the abstraction of 'tension-relaxation schemes' in a tonal musical phrase. *Contemporary Music Rev.* 9, 128–39.

51. **Bigand, E.** (1997) Perceiving musical stability: the effect of tonal structure, rhythm and musical expertise. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 808–22.

52. **Bigand, E. and M. Pineau** (1997) Global context effects on musical expectancy. *P&P* 59, 1098–107.

53. **Krumhansl, C. L., J. J. Bharucha, and M. Castellano** (1982) Key distance effects on perceived harmonic structure in music. *P&P* 32, 96–108.

54. **Bigand, E. and M. Pineau** (1996) Context effects on melody recognition: a dynamic interpretation. *Current Psychol. of Cognit.* 15, 121–34.

55. **Cuddy, L. L., A. J., Cohen, and D. J. Mewhort** (1981) Perception of structure in short melodic sequences. *J. Exp. Psychol. Hum. Percept. Perform.* 7, 869–83.

56. **Dowling, W. J.** (1978) Scale and contour: two components of a theory of memory for melodies. *Psychol. Rev.* 85, 341–54.

57. **Dowling, W. J.** (1991) Tonal strength and melody recognition after long and short delays. *P&P* 50, 305–13.

58. **Bharucha, J. J. and K. Stoeckig** (1986) Reaction time and musical expectancy: priming of chords. *J. Exp. Psychol. Hum. Percept. Perform.* 12, 403–10.

59. **Bharucha, J. J. and K. Stoeckig** (1987) Priming of chords: spreading activation or overlapping frequency spectra? *P&P* 41, 519–24.

60. **Bigand, E., F. Madurell, B. Tillmann, and M. Pineau** (1999) Effect of global structure and temporal organization on chord processing. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 184–97.

61. **Cuddy, L. L. and C. A. Lunney** (1995) Expectancies generated by melodic intervals: perceptual judgements of melodic continuity. *P&P* 57, 451–62.

62. **Besson, M. and F. Faieta** (1995) An Event-Related Potential (ERP) study of musical expectancy: comparison of musicians and nonmusicians. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 1278–96.

63. **Besson, M., F. Faita, and J. Requin** (1994) Brain waves associated with musical incongruities differ for musicians and nonmusicians. *Neurosci. Lett.* 168, 101–5.

64. **Besson, M. and F. Macar** (1987) An Event-Related Potential analysis of incongruity in music and other non-linguistic context. *Psychophysiology* 24 (1), 14–25.

65. **Koelsch, S., T. Gunter, and A. D. Friederici** (2000) Brain indices of music processing: 'nonmusicians' are musical. *J. Cogni. Neurosci.* 12 (3), 520–41.

66. **Maess, B., S. Koelsch, T. Gunter, and A. D. Friederici** (2001) 'Musical syntax' is processed in the Broca's area: an MEG-study. *Nature Neurosci.* 4, 540–5.

67. **Regnault, P., E. Bigand, and M. Besson** (2001) Event-related brain potentials show top-down and bottom-up modulations of musical expectations. *J. Cognit. Neurosci.* 13, 241–55.

68. **Tillmann, B., P. Janata, and J. J. Bharucha** (2003) Inferior frontal cortex activation in musical priming. *Cognit. Brain Res.* 16(2), 145–61.

69. **Patel, A. D., E. Gibson, J. Ratner, M. Besson, and P. J. Holcomb** (1998) Processing syntactic relations in language and music: an event-related potential study. *J. Cognit. Neurosci.* 10, 717–33.

70. **Lampinen, J. and E. Oja** (1992) Clustering properties of hierarchical self-organizing maps. *J. Mathematical Imaging and Vision* 2, 261–72.

71. **Parncutt, R.** (1988) Revision of Terhardt's psychoacoustical model of the roots of a musical chord. *MP* 6, 65–94.

72. **Leman, M.** (1995) *Music and Schema Theory.* Berlin: Springer.

73. **Bharucha, J. J.** (1987) Music cognition and perceptual facilitation: a connectionist framework. *MP* 5 (1), 1–30.

74. **Bharucha, J. J.** and **W. E. Mencl** (1996) Two Issues in auditory cognition: self-organization of octave categories and pitch-invariant pattern recognition. *Psychol. Sci.* 7, 142–9.

75. **Cohen, M. A., S. Grossberg,** and **L. L. Wyse** (1995) A spectral network model of pitch perception. *J. Acoust. Soc. of Am.* 98, 862–79.

76. **Sano, H.** and **B. K. Jenkins** (1991) A neural network model for pitch perception. In N. Todd and G. Loy (eds) *Music and Connectionism.* Cambridge, MA: MIT Press, pp. 42–9.

77. **Taylor, I.** and **M. Greenhough** (1994) Modelling pitch perception with adaptive resonance theory artificial neural networks. *Connection Sci.* 6, 135–54.

78. **Laden, B.** and **D.H. Keefe** (1991) The representation of pitch in a neural net model of chord classification. In (P. Todd and G. Loy (eds) *Music and Connectionism.* Cambridge, MA: MIT Press, pp. 64–83.

79. **Bharucha, J. J.** and **N. Todd** (1989) Modelling the perception of tonal structures with neural nets. *Computer Music J.* 13, 44–53.

80. **Krumhansl, C. L., J. Louhivuori, P. Toivianinen, T. Jarvinen,** and **T. Eerola** (1999) Melodic expectation in Finnish spiritual folk hymns: converging evidence of statistical, behavioral and computational analyses. *MP* 17, 151–95.

81. **Page, M. A.** (1994) Modelling the perception of musical sequences with self-organizing neural networks. *Connection Sci.* 6, 223–46.

82. **Griffith, N.** (1994) Development of tonal centres and abstract pitch as categorizations of pitch use. *Connection Sci.* 6, 155–75.

83. **Leman, M.** and **F. Carreras** (1998) Schema and Gestalt: testing the Hypothesis of Psychoneural Isomorphism by computer simulation. In M. Leman (ed.) *Music, Gestalt, Computing.* Berlin: Springer, pp. 144–68.

84. **Gjerdingen, R. O.** (1990) Categorization of musical patterns by self-organizing neuronlike networks. *MP* 8, 339–70.

85. **Stevens, C.** (1992) *Derivation and Investigation of Features Mediating Musical Pattern Recognition.* Sydney: University of Sydney.

86. **Stevens, C.** and **C. Latimer** (1992) A comparison of connectionist models of music recognition and human performance. *Minds and Machines* 2, 379–400.

87. **Stevens, C.** and **C. Latimer** (1997) Music Recognition: an illustrative application of a connectionist model. *Psychol. of Music* 25, 161–85.

88. **Tekman, H. G.** and **J. J. Bharucha** (1998) Implicit knowledge versus psychoacoustic similarity in priming of chords. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 252–60.

89. **Tillmann, B., E. Bigand,** and **M. Pineau** (1998) Effects of global and local contexts on harmonic expectancy. *MP* 16 (1), 99–117.

90. **Krumhansl, C. L., J. J. Bharucha,** and **E. J. Kessler** (1982) Perceived harmonic structures of chords in three related keys. *J. Exp. Psychol. Hum. Percept. Perform.* 8, 24–36.

91. **Peretz, I.** and **R. Kolinsky** (1993) Boundaries of separability between melody and rhythm in music discrimination: a neuropsychological perspective. *Q. J. Exp. Psychol.* 46A, 301–27.

92. **Jones, M. R.** and **M. Boltz** (1989) Dynamic attending and responses to time. *Psychol. Rev.* 96, 459–91.

93. **Boltz, M.** (1991) Some structural determinants of melody recall. *Memory and Cognition* 19, 239–51.

94. **Boltz, M.** (1993) The generation of temporal and melodic expectancies during musical listening. *P&P* 53, 585–600.

95. **Palmer, C.** and **C. L. Krumhansl** (1987) Independent temporal and pitch structures in determination of musical phrases. *J. Exp. Psychol. Hum. Percept. Perform.* 13, 116–26.

96. **Palmer, C.** and **C. L. Krumhansl** (1987) Pitch and temporal contributions to musical phrase perception: effects of harmony, performance timing, and familiarity. *P&P* 41, 505–18.

97. **Peretz, I.** (1990) Processing of local and global musical information in unilateral brain-damaged patients. *Brain* 113, 1185–205.

98. **Peretz, I.** and **J. Morais** (1989) La musique et la modularité. In S. McAdams and I. Deliege (eds) *La musique et les sciences cognitives.* Bruxelles: P. Mardaga, pp. 393–414.

99. **Berger, J.** and **D. Gang** (1997) A neural network model of metric perception and cognition in the audition of functional tonal music. In *Int. Computer Music Conference.* 1997 September, Thessaloniki, Greece.

100. **Berger, J.** and **D. Gang** (1998) A computational model of meter cognition during the audition of functional tonal music. In *Int. Computer Music Conference.* 1998 October, Ann Arbor, Michigan.

101. **Stevens, C.** and **J. Wiles** (1994) Tonal Music as a componential code: learning temporal relationships between and within pitch and timing components. In J. D. Cowan *et al.* (eds) *Advances in Neural Information Processing Systems,* vol. 6. San Francisco, CA: Morgan Kaufmann, pp. 1085–92.

102. **Large, E. W.** and **J. F. Kolen** (1994) Resonance and the perception of musical meter. *Connection Sci.* 6, 177–208.

103. **Mayr, U.** (1996) Spatial attention and implicit sequence learning: evidence for independent learning of spatial and nonspatial sequences. *JEP-LMC* 22, 350–64.

104. **McClelland, J. L., B. L. McNaughton,** and **R. C. O'Reilly** (1995) Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* 102, 419–57.

105. **Otto, I., P. Grandguillaume, L. Bouthkhil, Y. Burnod,** and **E. Guigon** (1992) Direct and indirect cooperation between temporal and parietal networks for invariant visual recognition. *Cognit. Neurosci.* 4, 35–57.

106. **Rauschecker, J. P.** and **B. Tian** (2000) Mechanisms and streams for processing of 'what' and 'where' in auditory cortex. *Proc. Nat. Acad. Sci.* 24, 11800–6.

107. **Belin, P.** and **R. J. Zatorre** (2000) 'What', 'where' and 'how' in auditory cortex. *Nature Neurosci.* 3, 965–6.

PART III

# THE NEURONS OF MUSIC

CHAPTER 9

# NEUROBIOLOGY OF HARMONY PERCEPTION

MARK JUDE TRAMO, PETER A. CARIANI, BERTRUND DELGUTTE, AND LOUIS D. BRAIDA

## Abstract

Basic principles of the theory of harmony reflect physiological and anatomical properties of the auditory nervous system and related cognitive systems. This hypothesis is motivated by observations from several different disciplines, including ethnomusicology, developmental psychology, and animal behaviour. Over the past several years, we and our colleagues have been investigating the vertical dimension of harmony from the perspective of neurobiology using physiological, psychoacoustic, and neurological methods. Properties of the auditory system that govern harmony perception include (1) the capacity of peripheral auditory neurons to encode temporal regularities in acoustic fine structure and (2) the differential tuning of many neurons throughout the auditory system to a narrow range of frequencies in the audible spectrum. Biologically determined limits on these properties constrain the range of notes used in music throughout the world and the way notes are combined to form intervals and chords in popular Western music. When a harmonic interval is played, neurons throughout the auditory system that are sensitive to one or more frequencies (partials) contained in the interval respond by firing action potentials. For consonant intervals, the fine timing of auditory nerve fibre responses contains strong representations of harmonically related pitches implied by the interval (e.g. Rameau's fundamental bass) in addition to the pitches of notes actually present in the interval. Moreover, all or most of the partials can be resolved by finely tuned neurons throughout the auditory system. By contrast, dissonant intervals evoke auditory nerve fibre activity that does not contain strong representations of constituent notes or related bass notes. Furthermore, many partials are too close together to be resolved. Consequently, they interfere with one another, cause coarse fluctuations in the firing of peripheral and central auditory neurons, and give rise to perception of roughness and dissonance. The effects of auditory cortex lesions on the perception of consonance, pitch, and roughness, combined with a critical reappraisal of published psychoacoustic data on the relationship between consonance and roughness, lead us to conclude that consonance is first and foremost a function of the pitch relationships among notes. Harmony in the vertical dimension is a positive phenomenon, not just a negative phenomenon that depends on the absence of roughness—a view currently held by many psychologists, musicologists, and physiologists.

**Keywords**: Consonance; Dissonance; Harmony, musical; Intervals, musical; Perception of harmony; Psychoacoustics of harmony

## Introduction

Why do some combinations of simultaneous tones sound more harmonious than others? Pythagoras's curiosity about the nature of harmony inspired some of the earliest experiments

relating mathematics and physics to perceptual phenomena.[1] Approaching the problem from the perspective of neurobiology, we ask: Are there physiological and anatomical properties of the auditory system and related cognitive systems that determine the degree to which simultaneous notes sound harmonious?

We restrict our consideration of harmony to basic tenets articulated by Piston,[2] among others. Harmony has a vertical dimension and a horizontal dimension. The vertical dimension encompasses the relationships among simultaneous notes. By convention, *note* refers to a pitch in the musical scale, and *harmonic interval* refers to two notes sounded simultaneously (Figure 9.1A–D). When a note is played on a musical instrument, its pitch corresponds to the fundamental frequency ($F_0$) of the complex tone generated by the instrument. Some synthesizers and other types of equipment are capable of generating pure tones, in which case the pitch of the note corresponds to the frequency of the pure tone. Harmonic intervals are a type of *dyad.* Three or more notes played simultaneously make up a *chord.* Chords with three notes are called *triads.* The time window over which acoustic information is integrated in the vertical dimension spans about a hundredth of a second to a few seconds (e.g. sixteenth notes to tied whole notes at a tempo of 120 beats per min). The horizontal dimension encompasses successive tones (melodic intervals and melodic progressions) and successive harmonic intervals and chords (harmonic progressions). Certain intervals and chords are treated as consonant (e.g. fifths, major triads) and others as dissonant (e.g. minor seconds, diminished triads). Acknowledging that different psychologists have attached different perceptual attributes and meanings to the terms *consonance* and *dissonance,* we nonetheless find considerable agreement among music texts and dictionaries that consonant means harmonious, agreeable, and stable, and that dissonant means disagreeable, unpleasant, and in need of resolution.[2–5] Psychoacoustic experiments bear out semantic overlap among the terms *consonant, pleasant, beautiful,* and *euphonious.*[6,7]

These basic concepts apply to a wide range of musical styles enjoyed by people throughout much of the industrialized world: contemporary pop and theater (including rock, rhythm and blues, country, and Latin-American), European music from the Baroque, Classical, and Romantic eras (1600–1900), children's songs, and many forms of ritualistic music (e.g. church songs, processionals, anthems, and holiday music). The overlap in their harmonic structure incorporates commonalities in musical phonology and syntax.[8] In our view, the widespread popularity of Western pop taps into (1) universal competence in auditory functions needed to extract the pitch of a note and to analyse the harmonic relationships among different pitches, and (2) universal competence in cognitive functions that parse acoustic information and associate perceptual attributes with emotion and meaning. Experimental results suggest that similar perceptual attributes can be associated with similar emotions and social contexts across different cultures.[9,10] Moreover, listeners from different cultures often use similar cognitive schemata to structure the processing of pitch-sequences.[11,12]

Much has been written in the psychology literature about the terms *harmony, consonance,* and *dissonance.* At present, many psychologists and musicologists subscribe to Terhardt's[13] two-component model of *musical consonance,* which subsumes all these terms. One component is *sensory consonance,* the absence of annoying features, such as roughness, in both musical and nonmusical sounds. The other component, *harmony,* is based on music-specific principles that govern pitch relationships in melodic and harmonic progressions. Terhardt

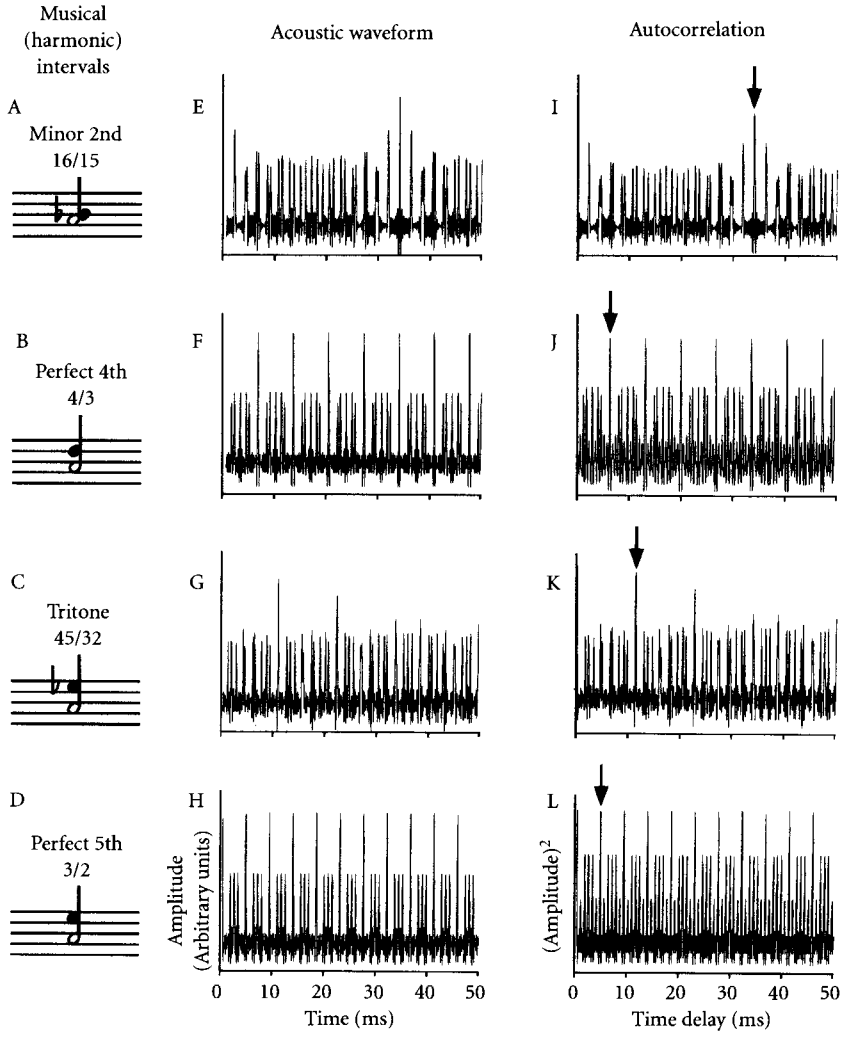



Figure 9.1 Acoustic representations of musical (harmonic) intervals in the time domain. *Left column* (A–D): Musical interval stimuli depicted in standard notation (G clef) below the name of each interval (e.g. minor second) and the $F_0$ ratio of its notes (16/15). *Middle column* (E–H): Acoustic waveforms of each interval. *Right column* (I–L): Autocorrelations of the acoustic waveforms. *Arrows* indicate the peaks corresponding to the period of each interval's 'missing' $F_0$.

asserts that 'sensory consonance... dominates the evaluation of single isolated chords... whereas harmony does not enter into the subject's response' (p. 282).[13]

With respect to terminology, we adhere in this discussion to the simple distinction between the vertical and horizontal dimensions of harmony. We restrict use of the terms

*consonance* and *dissonance* to the vertical dimension and keep the term *harmony* supraordinate to them. We make no assumptions about the level of auditory processing (e.g. sensory, peripheral) where the perceptual attribute of consonance takes shape. We consider it likely that a listener's implicit (or explicit) knowledge about harmony in the horizontal dimension[14] bears on harmony perception in the vertical dimension.

In this paper, we present neurophysiological, neurological, and psychoacoustic evidence to support our contentions that (1) pitch relationships among tones in the vertical dimension influence consonance perception and (2) consonance cannot be explained solely by the absence of roughness. First, we review terminology and basic psychoacoustics pertinent to our subsequent discussion of experimental results. Second, we demonstrate that the harmonic relationships of tones in musical intervals are represented in the temporal discharge patterns of auditory nerve fibres. Third, we critically reevaluate the psychoacoustic literature concerning the consonance of isolated intervals and chords, paying particular attention to (1) the relationships among interval width, roughness detection thresholds, and consonance ratings; and (2) the predictions of roughness-based computational models about relative consonance as a function of spectral energy distribution. Finally, we discuss evidence that impairments in consonance perception following auditory cortex lesions are more likely to result from deficits in pitch perception than to deficits in roughness perception. This evidence highlights the dependence of so-called low-level perceptual processing on the integrity of the auditory cortex, the highest station in the auditory nervous system (Figure 9.2).

## Psychoacoustics and neurophysiology of harmony

For authoritative reviews of the psychoacoustics of harmony, we refer the reader to Krumhansl,[14] Parncutt,[18] and Deutsch.[19] Here, only basic concepts and terminology pertinent to our subsequent discussion of psychoacoustic and neurophysiological experiments are covered.

Let us consider a modern restatement of Pythagoras's observation: The degree to which two simultaneous notes (a harmonic interval) sound consonant is determined by the simplicity of the ratio $x:y$, where $x$ is the $F_0$ associated with one tone and $y$ the $F_0$ of the other, lower tone. In musical terms, $y$ is the root of the interval. $x$ and $y$ can take on any value between about 25 Hz and about 5 kHz. This upper limit coincides with (1) the upper $F_0$ of notes on a piccolo (~4500 Hz); (2) the upper $F_0$ for which octave similarity can be reliably judged;[20,21] and (3) the upper $F_0$ of strong phase locking by auditory nerve fibres—that is, the highest frequency at which neurons can fire in time with amplitude fluctuations in the acoustic waveform.[22,23] This convergence of facts from music, psychoacoustics, and physiology suggests that limitations in the phase-locking capacity of neurons in the auditory periphery constrain the range of note $F_0$s that are used in music and the way they are combined in the vertical dimension of harmony. Other authors have discussed the relationships among the temporal discharge patterns of auditory nerve fibres, fundamental pitch perception, octave equivalence, and the consonance of intervals formed by simple integer ratios.[24–29]

By convention, notes in popular Western music are tuned to the scale of equal temperament, which chunks the $F_0$ continuum into octaves (i.e. doublings of $F_0$) and each octave

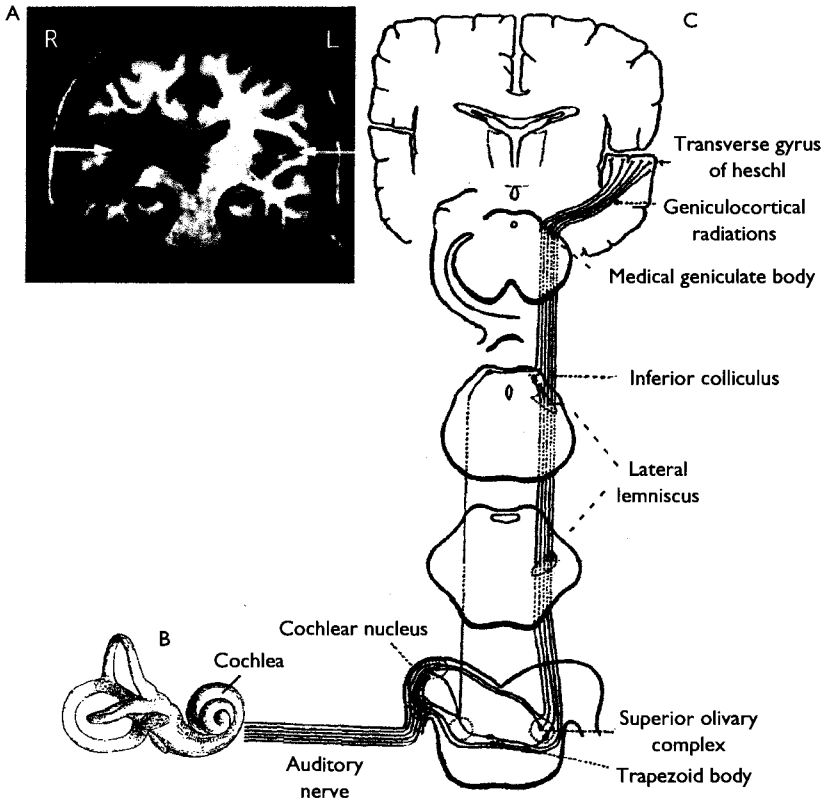



**Figure 9.2** The auditory system. (A) Magnetic resonance image (coronal section) through the transverse gyri of Heschl and the superior temporal gyri of case MHS, who suffered bilateral infarction nine years earlier. *White arrows* point to the region of the transverse gyrus of Heschl in the right (R) and left hemispheres (L). The results of psychoacoustic experiments performed by MHS are shown in Figure 9.7. (B) The inner ear and the auditory nerve, the obligatory pathway from the cochlea to the cochlear nuclei in the lower brain stem. (C) Schematic of the central auditory pathway showing the main relay stations and projection patterns. Adapted from Tramo *et al.*[15] (A), Helmholtz[16] (B), and Davis[17] (C).

into 12 discrete, equal, logarithmic steps. Each step within the octave is called a semitone, and the $F_0$s of adjacent semitones differ by a factor of $2^{1/12}$, or about 6 per cent. The chromatic scale is made up of all 12 tones in the octave, whereas the major and minor (diatonic) scales are made up of partially overlapping sets of seven tones in the octave. Harmonic intervals and melodic intervals are named according to the scale relationship of the upper note to the lower note. Thus when the fifth note in the major or minor scale sits atop the first note on the scale, the interval is called a *fifth* (Figure 9.1D). The note named $A_4$ is assigned an $F_0$ value of 440 Hz. The letter name of each note corresponds to one of the 12 notes in the octave; the number of each note indicates the octave the note is in, with increments at each occurrence of C along the scale (e.g. ... $A_4$—$A\#_4$—$B_4$—$C_5$ ...). The

frequency range over which pure tones are audible to humans extends from approximately $E_0$ (20.6 Hz) to $E_{10}$ (21.1 kHz). For the purposes of this discussion, we will set $y$, the $F_0$ of the root of a harmonic interval, equal to 440 Hz ($A_4$).

In music theory, the interval formed by notes that are an octave apart (e.g. $A_4$ and $A_5$) is the most consonant interval, followed by the fifth ($A_4$ and $E_5$), fourth ($A_4$ and $D_5$), major third ($A_4$ and $C\#_5$), and minor third ($A_4$ and $C_5$). In the scale of just intonation, these intervals correspond to $x:y$ ratios of 2:1, 3:2, 4:3, 5:4, and 6:5, respectively, consistent with Pythagoras's claim that the simplicity of the integer ratio correlates with perceived consonance. Combinations of some other notes on the scale between $A_4$ and $A_5$ have more complex $x:y$ ratios and sound dissonant. For example, the minor second and the tritone (also known as the augmented fourth, which, in equal-tempered tuning, is equivalent to the diminished fifth) have $x:y$ ratios of 16:15 and 45:32 (or approximately 7:5), respectively. The dependence of consonance on the simplicity of $F_0$ ratios tolerates small deviations from perfect integer relationships. For example, because the scale of equal temperament is based on equal logarithmic steps within each octave, a major third in this scale has a ratio of 5.04:4, not 5:4. This deviation amounts to 0.8 per cent. Even highly practised listeners participating in psychoacoustic experiments under ideal listening conditions cannot reliably detect a mistuned lower harmonic embedded within a harmonic complex tone if the deviation is less than 0.9 per cent (harmonics 1–12 at 60 dB SPL [sound pressure level] and isophase, $F_0 = 200$ Hz, duration $\leq 410$ ms).[30] Moreover, conservatory students who excel at interval identification cannot reliably judge whether a mistuned major third composed of two harmonic complex tones has been stretched or compressed away from a perfect 5:4 ratio if the deviation is less than 1.2 per cent (each tone with harmonics 1–20 in isophase, first harmonic at 80 dB SPL, higher harmonics at a 6 dB decrease per octave, $F_0$ between 260 and 525 Hz, and duration = 1000 ms).[31]

All experimental studies that have used stimuli consisting of single, isolated, harmonic intervals formed by two complex tones (as would be the case if the intervals were sung or played on guitar or piano) show that listeners consistently perceive the fifth and fourth as more consonant than the minor second and tritone.[6,32–35] This convergence of results across study populations from different countries (United States, Germany, Japan), generations (1909–69), and musical backgrounds, combined with results obtained in infants from the United States[36] and European starlings,[37] motivates the hypothesis that common, basic auditory mechanisms underlie perceptual categorization of harmonic intervals as consonant or dissonant. However, there is disagreement about the nature of the underlying neural mechanisms, and few physiological experiments have systematically analysed responses to harmonic intervals at any level of the auditory nervous system.[25,27,38–41] Still, a large body of data is available on the responses of neurons to other types of complex tones in the auditory nerve,[42] cochlear nucleus,[43] inferior colliculus,[44,45] medial geniculate nucleus,[46] and auditory cortex[47–49] (Figure 9.2; only a few of the many available papers are cited here; for review see Ref. 50).

## Neural coding of pitch relationships as a physiological basis of harmony

We synthesized simultaneous complex tones forming four musical intervals: the minor second ($F_0$ ratio = 16/15), perfect fourth (4:3), tritone (45:32), and perfect fifth (3:2,

Figure 9.1). Each of the two complex tones in the interval contained the first six harmonics with equal amplitude (60 dB SPL re: 20 μPa) and equal phase (cosine, Figure 9.3). Each interval had a duration of 200 ms (a bit shorter than an eighth note at a tempo of 120 beats per minute), including 5-ms rise and fall times. These stimuli are acoustically similar to the inputs into the computational models used by Plomp and Levelt,[51] Kameoka and Kuriyagawa,[35] and Hutchinson and Knopoff[52] to predict the consonance of complex-tone intervals on the basis of psychoacoustic data on puretone intervals.

Figure 9.1 illustrates two time-domain representations of our stimuli: the acoustic waveform, which plots sound pressure amplitude as a function of time (Figure 9.1E–H); and the autocorrelation of the waveform (Figure 9.1I–L). In the acoustic waveform of the

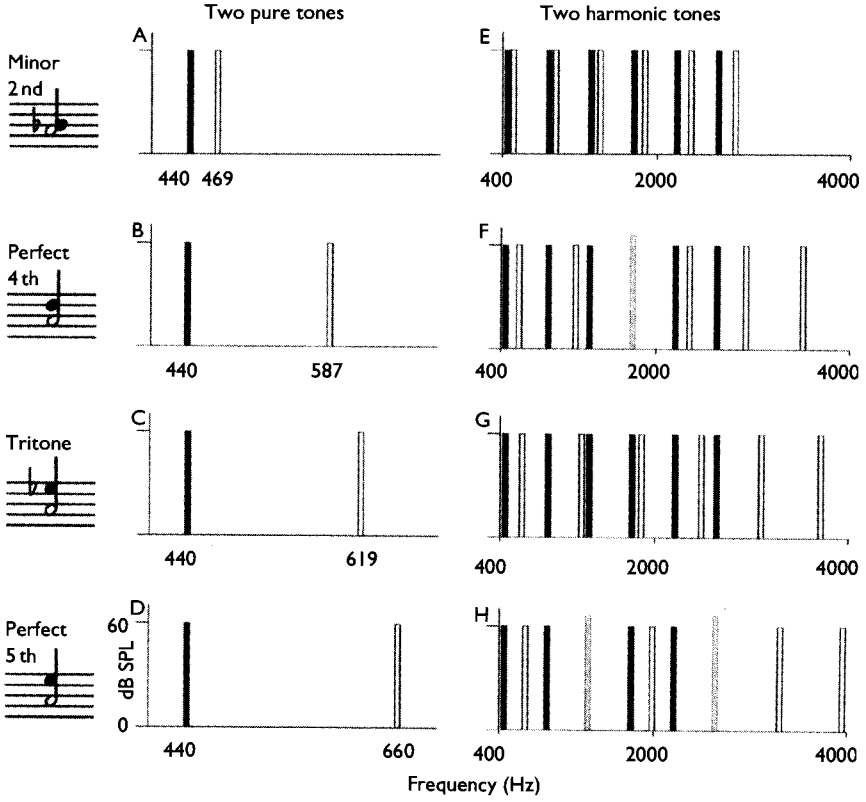



Figure 9.3 Line amplitude spectra of the four musical intervals shown in Figure 9.1. *Black* indicates the root and its harmonics, *white* the interval and its harmonics, and *gray* the frequencies at which harmonics of the root and interval overlap. (A–D) A minor second, perfect fourth, tritone, and perfect fifth composed of two pure tones. (E–H) The same intervals composed of two harmonic complex tones containing the first six harmonics at equal amplitude. These are the spectral domain representations corresponding to the time domain representations in Figure 9.1E–L.

most consonant interval, the perfect fifth, we see a regular pattern of major and minor peaks (Figure 9.1H). The pattern with one major peak and three minor peaks repeats every 4.55 ms ($1/x = 220$ Hz). This periodicity corresponds to the missing $F_0$ of a harmonic series containing energy at the second harmonic (440 Hz, $A_4$) and third harmonic (660 Hz, $E_5$), the $F_0$s of the notes actually present in the stimulus. Rameau's concept of the 'basse fondametale' (fundamental bass) in his *Treatise on Harmony*[53] is related to the missing $F_0$ of a harmonic interval.

Autocorrelation functions provide another representation of temporal regularities and irregularities embedded in acoustic waveforms (Figure 9.1I–L). Autocorrelation functions are computed by multiplying the waveform with a delayed copy of itself and integrating over time. A large value at a given delay indicates the presence of a dominant periodicity in the waveform whose period equals the delay. Like pitch percepts, but unlike acoustic waveforms, autocorrelation functions are stable despite changes in the relative phases of frequency components. In the autocorrelation functions plotted in Figure 9.1I–L, the periodicity at the upper limit of the $x$ axis (50 ms) corresponds to 20 Hz, the lowest audible frequency.

In the autocorrelation function of the perfect fifth (Figure 9.1L), the first major peak again corresponds to the missing $F_0$, $A_3$ (220 Hz). The second major peak occurs at 9.09 ms, which corresponds to $A_2$ (110 Hz), the bass note an octave below. In fact, all the major peaks up to 50 ms correspond to the fundamental bass and its subharmonics (undertones) at $A_2$, $D_2$, $A_1$, and on down to $A_0$, the lowest note on the piano ($F_0 = 27.5$ Hz).

In between the major peaks is a set of three, evenly spaced, minor peaks. The first of these minor peaks occurs at 1.51 ms, which corresponds to $E_5$ (660 Hz), the upper note of the interval. The second minor peak occurs at 2.27 ms, which corresponds to $A_4$ (440 Hz), the root of the interval. The third minor peak occurs at 3.03 ms, which corresponds to $E_4$ (330 Hz), the octave below the fifth and the fifth above the fundamental bass at $A_3$.

Temporal regularities are also seen in the waveform and autocorrelation of the perfect fourth, the other consonant interval in our stimulus set (Figure 9.1F and J). Here the major peaks are at 6.82 ms ($D_3$, $F_0 = 146.7$ Hz) and 13.6 ms ($D_2$, $F_0 = 73.3$ Hz). Thus, in addition to a representation of the fourth, there is a representation of its inversion as a fifth with the implied root at $D_3$ and the fundamental bass at $D_2$. The autocorrelation function of the fourth is a bit more complicated than that of the fifth, as there are two more peaks between each pair of major peaks. In the first set of minor peaks, the following notes are represented: $A_4$ (the root), $D_4$ (the interval), $A_3$ (the octave below the root), and $G_3$ (the fifth below $D_4$, and the fourth of an interval rooted at $D_3$). Thus, we find representations of notes that function as fourths and fifths in the major and (all) minor scales of A and D.

In summary, the temporal fine structure of the perfect fifth and fourth contains representations of the two notes constituting the interval, plus harmonically related bass notes that are implied by the interval. In music, these bass notes support the deep structure of harmony. Parncutt[18] demonstrated experimentally that listeners associate major triads with pitches that are harmonically related to note $F_0$s, including the fundamental bass, plus the pitches of note $F_0$s actually in the stimulus. These pitches cannot be accounted for simply on the basis of combination tones (for review, see Ref. 54). Houtsma and Goldstein[55] showed that musicians can use missing $F_0$ pitches to identify melodic intervals

(major and minor seconds and thirds), even when two upper harmonics are presented separately (dichotically) to each ear.

For the dissonant intervals in our stimulus set, the minor second and tritone, we find no such temporal regularity in the acoustic waveform and autocorrelation function. For the minor second (Figure 9.1I), the largest peak in the autocorrelation function occurs at 34.1 ms, which corresponds to a frequency of 29.3 Hz, and it decays rapidly into the background. This periodicity lies outside the range associated with strong periodicity pitch percepts (for review, see Ref. 56). In addition, we find multiple peaks between zero and the maximum peak. The largest of these is the first peak at 2.20 ms, which corresponds to mean of the note $F_0$s, 455 Hz. This pitch does not correspond to any of the notes in any scale that has $A_4$ in it. In short, there is no strong representation of any pitch below the note pitches in the interval, and the dominant pitch is off the scale. Both of these factors contribute to the dissonance of the minor second.

Likewise, the autocorrelation function of the tritone (Figure 9.1K) does not show a simple, regular pattern of peaks. The largest peak occurs at 11.4 ms, which corresponds to $F_0 = 88$ Hz, and it, too, decays into the background. This periodicity corresponds to a near coincidence between the fifth subharmonic of the root (440 Hz divided by 5) and the seventh subharmonic of the tritone (618.7 Hz divided by 7). It also lies close to $F_2$, which is related to the fundamental bass of an F dominant-seventh chord in its first inversion. Thus the autocorrelation function of the tritone implies a chord that is, in music theory, less consonant than the chords implied by the autocorrelation functions of the fifth and fourth.

To summarize, for the consonant intervals (the fifth and fourth), the pattern of major and minor peaks in the autocorrelation is perfectly periodic, with a period related to the fundamental bass. This pattern is obtained because these stimuli have a unique, clearly defined fundamental period. By contrast, for dissonant intervals (the minor second and tritone), no true periodicity is seen in the autocorrelation function. While some peaks occasionally stand out at specific delays, indicating a pseudoperiod, either there are no consistent peaks at multiples of this pseudoperiod, or the amplitudes of these peaks decay rapidly with increasing multiples of the pseudoperiod.

These observations suggest that the consonance of harmonic intervals reflects regularities in their temporal fine structure in the range of tenths to tens of milliseconds. Do neurons in the auditory system represent this information using a time code? Galileo,[57] who wrote about consonance while he was under house arrest for his work on the solar system, may have been the first to postulate that temporal coding in the auditory periphery was the physiological basis for consonance:

> Agreeable consonances are pairs of tones which strike the ear with a certain regularity; this regularity consists in the fact that the pulses delivered by the two tones, in the same interval of time, shall be commensurable in number, so as not to keep the ear drum in perpetual torment, bending in two different directions in order to yield to the ever-discordant impulses... The unpleasant sensation produced by [dissonances] arises, I think, from the discordant vibrations of two different tones which strike the ear out of time. Especially harsh is the dissonance between notes whose frequencies are incommensurable;... this yields a dissonance similar to the augmented fourth or diminished fifth [*tritono o semidiapente*]". (Ref. 57, 1638, pp. 103–4.)

To investigate the neural coding of consonance in the auditory periphery, we analysed the responses of over 100 cat auditory nerve fibres to the minor second, perfect fourth, tritone, and perfect fifth. Auditory nerve fibres are the central axons of spiral ganglion cells that synapse on cochlear nucleus neurons in the brain stem (Figure 9.2 for review, see Ref. 58). In humans, each auditory nerve contains about 30,000 auditory nerve fibres. Spiral ganglion cells also have peripheral axons that synapse on sensory receptors in the cochlea—the inner hair cells that ride atop the basilar membrane. Virtually all information about sound is transmitted from the ear to the brain via trains of action potentials fired by auditory nerve fibres.

When a minor second or some other interval is sounded, an auditory nerve fibre will increase the number of action potentials it fires only if it is sensitive to the frequencies present in the interval (Figure 9.4A). The time between consecutive action potentials in the train is called an *interspike interval* (ISI), and a plot of the number of times each ISI occurs in the spike train is called an *ISI histogram* (Figure 9.4B–E). We measured all the ISIs between all possible pairs of spikes (Figure 9.4A, $ISI_1$, $ISI_2$... $ISI_N$) with a precision of approximately one microsecond. The corresponding plot is called an *all-order ISI histogram*, which is equivalent to the autocorrelation of the spike train. The spike train of each fibre in the auditory nerve can be analysed in this way, and the resultant ISI histograms can be combined to show the ISI distribution in the entire population of auditory nerve fibres. Single-unit physiology experiments[59–61] and computational models[62] have shown that the first among the major peaks in the all-order ISI histogram computed from the entire auditory nerve fibre ensemble (the population ISI distribution) matches the fundamental period of complex tones and thus their periodicity pitch. This is essentially the time-domain equivalent of Terhardt's[63] spectrally based subharmonic sieve for virtual pitch extraction.

Figure 9.4B–E illustrates the population ISI distributions embedded in the spike trains fired by over 50 auditory nerve fibres in response to the minor second, fourth, tritone, and fifth. In the response to the fifth (Figure 9.3E), we see major peaks corresponding to the fundamental bass ($A_3$, 4.55 ms) and its subharmonics, just as we did in the acoustic waveform (Figure 9.1H) and the autocorrelation of the waveform (Figure 9.1L). Indeed, for all four intervals, the autocorrelation histograms of neural responses (Figure 9.4B–E) mirror the fine structure of acoustic information in the time domain (Figure 9.1).

Thus the peaks in the population ISI distribution evoked by consonant intervals reflect the pitches of each note, the fundamental bass, and other harmonically related pitches in the bass register. By contrast, the dissonant intervals (the minor second and tritone) are associated with population ISI distributions that are irregular. These contain little or no representation of pitches corresponding to notes in the interval, the fundamental bass, and related bass notes.

To obtain a physiological measure of the strength of the fundamental pitch of each interval relative to other pitches, we measured the number of intervals under the peak in the all-order ISI distribution corresponding to the missing $F_0$ (arrows in Figure 9.3B–E; bin width = 300 μs). We then divided that value by the value of $y$ in each $x$ bin from $x = 0$–50 ms. We found a high correlation ($r = 0.96$) between our physiological measure of fundamental pitch strength and previous psychoacoustic measures of the 'clearness' of musical intervals composed of two complex tones.[35]

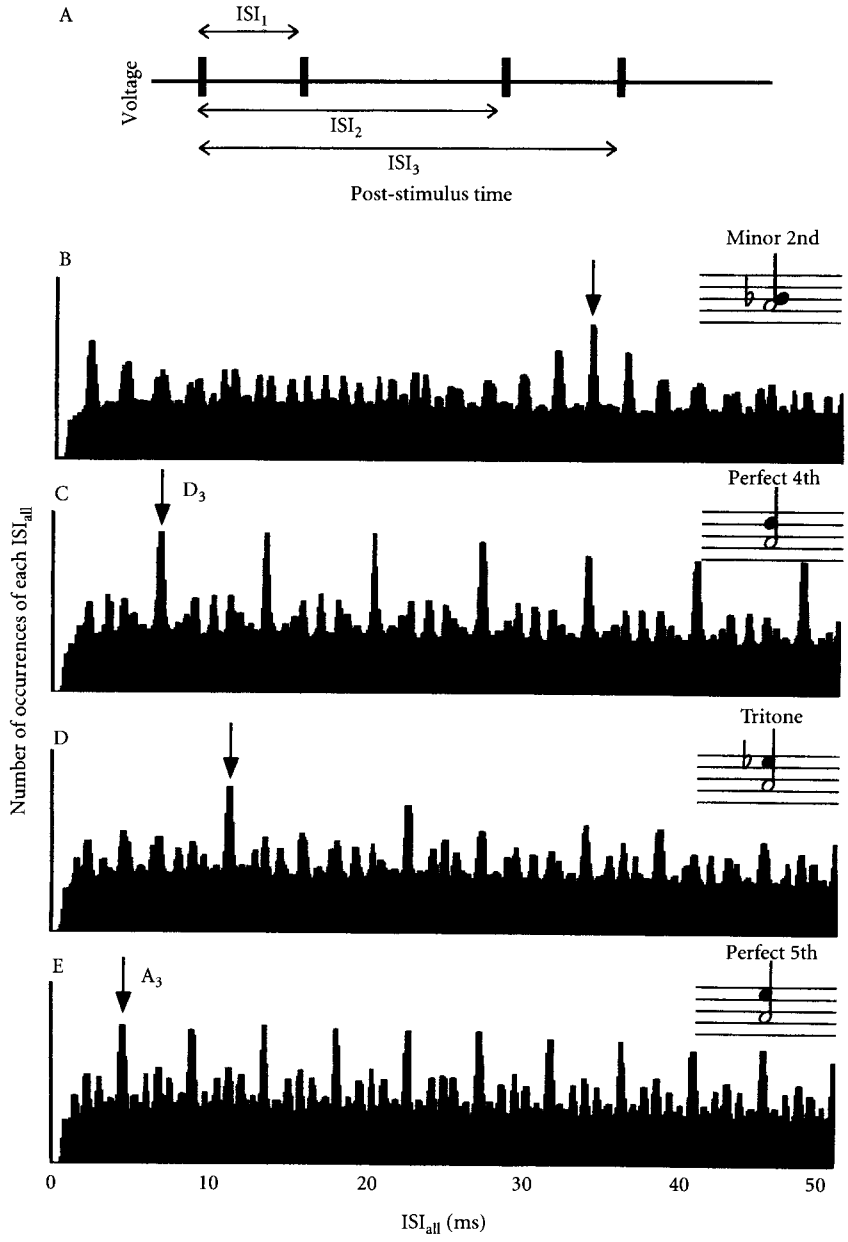



**Figure 9.4** Interspike interval (ISI) distributions embedded in the responses of axons throughout the auditory nerve during stimulation with four musical intervals. (A) Schematic of a train of action potentials (or 'spikes', *vertical lines*) fired by an auditory nerve fibre when a musical interval is played. *Double arrows* demarcate some of the ISIs between pairs of spikes in the spike train. $ISI_1$ refers to the first-order ISI, $ISI_2$ and $ISI_3$ to higher-order ISIs. All possible ISIs are in 'all-order' ISI histograms. (B–E) All-order ISI histograms showing ISIs embedded in the responses of 50 auditory nerve fibers to musical intervals composed of two complex tones (Figure 9.1E–H and Figure 9.3E–H). *Arrows* mark peaks in the ISI pattern that correspond to the missing $F_0$ of the interval. If the peak corresponds to a note on the scale, the name of the note is given.

In summary, the all-order ISI distribution embedded in auditory nerve fibre firing patterns contains representations of the pitch relationships among note $F_0$s that influence the perception of musical intervals as consonant or dissonant. The neural coding mechanisms that provide representations of these pitch relationships form part of the neurobiological foundation for the theory of harmony in its vertical dimension.

## Neural coding of roughness as the physiological basis for harmony perception

Whereas pitch-based accounts treat consonance as a positive perceptual phenomenon associated with the presence of highly structured temporal information, roughness-based accounts treat consonance as a negative phenomenon associated with the absence of annoying perceptual attributes. Terhardt's notion that the consonance of isolated intervals and chords depends on the absence of roughness[13] echoes one of the main points in Helmholtz's[16,64] monumental work, *On the Sensations of Tone as a Physiological Basis for the Theory of Music*:

> As long as several simple tones of a sufficiently different pitch enter the ear together, the sensation due to each remains undisturbed in the ear, probably because entirely different bundles of [auditory] nerve fibers are affected. But tones of the same, or of nearly the same pitch, which therefore affect the same nerve fibers, do not produce a sensation which is the sum of the two they would have separately excited, but new and peculiar phenomena arise which we term interference... and beats[16] (p. 160)... Rapidly beating tones are jarring and rough... the sensible impression is also unpleasant[16] (p. 168). Consonance is a continuous, dissonance an intermittent tone sensation. The nature of dissonance is simply based on very fast beats. These are rough and annoying to the auditory nerve.[64] (Helmholtz 1863,[64] 1885.[16])

Because frequency selectivity throughout the auditory nervous system is finite, simultaneous pure tones that are separated by small frequency differences ($\Delta F$s), such as a minor second (Figure 9.3A), cannot be separated or 'filtered out' from one another. Consequently, their waveforms are effectively summed, and the pitch of the tone combination matches their mean frequency.[65] The envelope of the summed waveform contains periodic amplitude fluctuations whose frequency equals $\Delta F$ (Figure 9.5, top). If these envelope fluctuations fall in the range of 20–200 Hz (the precise values depend on the frequencies of the two tones, Figure 9.6A), interruptions in continuous tone sensation are perceptible. These interruptions make the tone combination sound 'rough', analogous to the interruptions one feels on the fingertips when touching coarse sandpaper. At smaller frequency differences, and thus slower amplitude modulations, one perceives a single, continuous tone that is slowly fluctuating in loudness, or 'beating'. Auditory nerve fibres[38,39] (Figure 9.5, *bottom*), inferior colliculus neurons,[40,41] and populations of primary auditory cortex neurons[48] can fire in synchrony with amplitude fluctuations in the $\Delta F$ range associated with perception of roughness and beats.

The concept of *critical bandwidth* refers to the limits of $\Delta F$ over which frequency selectivity operates in the auditory system. Critical bandwidth has been estimated psychoacoustically in several ways that have yielded somewhat different results depending on the method (for

reviews, see Refs 56 and 66). One estimate of critical bandwidth is based on the $\Delta F$ above which roughness disappears.

When musical intervals are composed of two complex tones (Figure 9.3E–H), the partials may interfere with one another and produce amplitude fluctuations at the corresponding $\Delta F$. There is more interference between adjacent partials in the minor second and tritone (Figure 9.3E and G) than in the fourth and fifth (Figure 9.4F and H). Several computational models of consonance[16,35,51,52] assume that (1) the roughness generated by all the partials in the interval are added together (presumably by a central processor in the auditory brain stem or cortex), and (2) this total roughness determines the degree to which the interval is perceived as consonant.

Figure 9.6A shows Plomp and Steeneken's data on the relationships among the $\Delta F$ between two pure tones, the frequency of the lower tone (or root), and just-noticeable

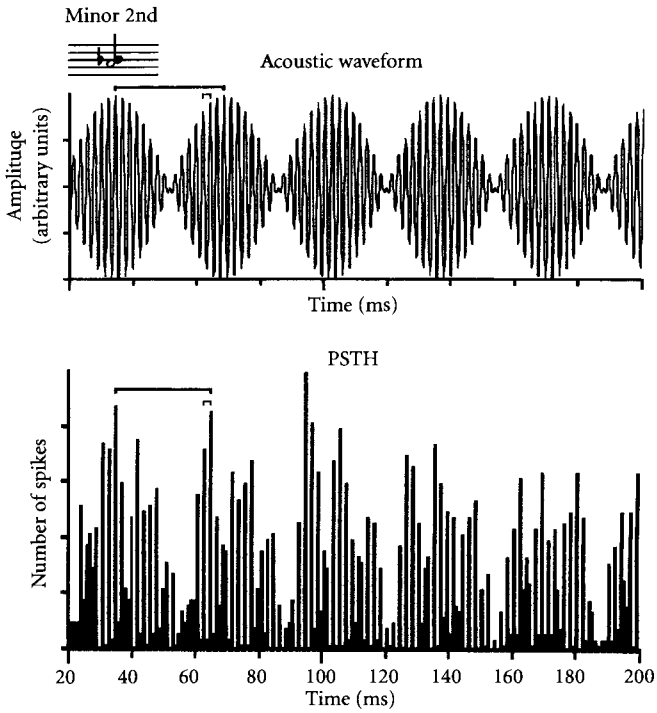



**Figure 9.5** (*Top*) Acoustic waveform of a minor second composed of two pure tones with the root at $A_4$. *Thick bars* show the period of envelope fluctuations that render the minor second rough ($P = 1/\Delta F = 34.1$ ms). *Thin bars* show the period of fluctuations under the envelope that corresponds to the mean frequency of the tones and the pitch of the interval ($P = 2.20$ ms). (*Bottom*) Poststimulus time histogram (PSTH) showing the number of spikes fired by a single auditory nerve fibre during the steady state portion of its response to the minor second. Note that the global and local fluctuations in firing rate mirror those seen in the acoustic waveform of the minor second. This fibre was sensitive to frequencies at both the root and the interval at 60 dB SPL. Bin width = 1 ms. Number of stimulus repetitions = 100.

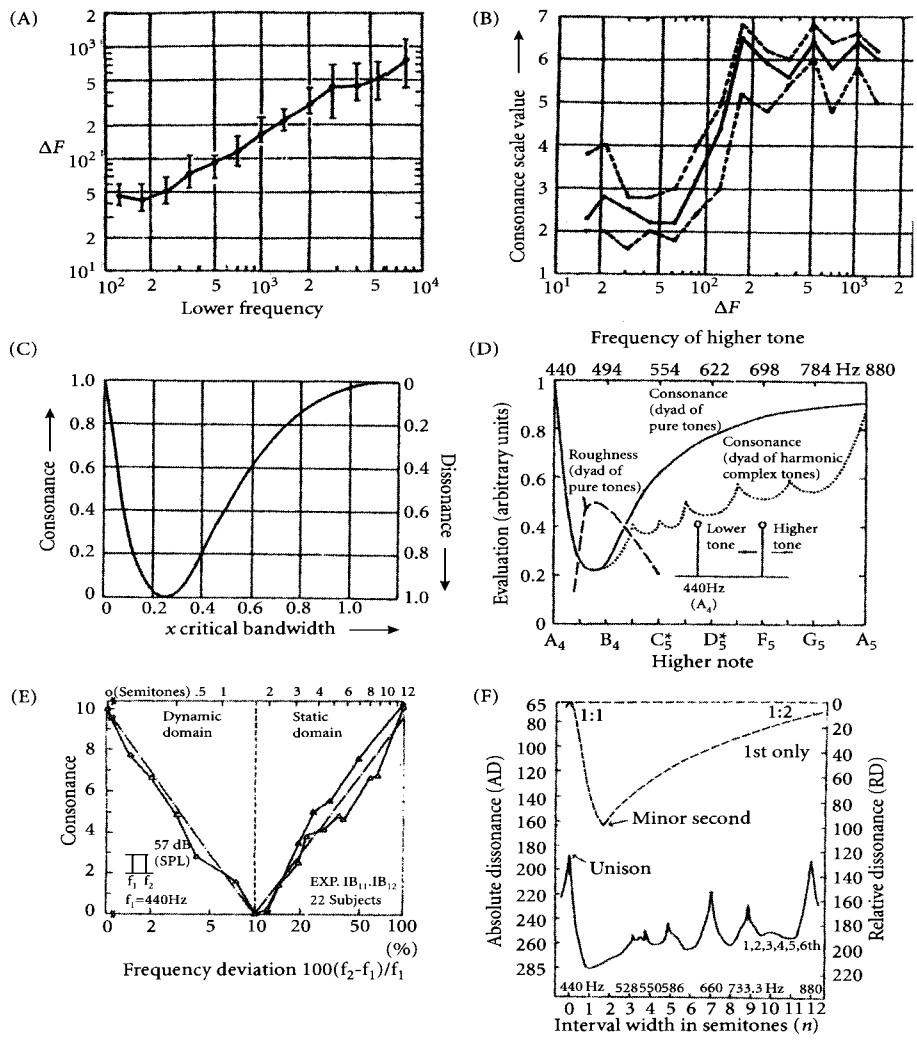



**Figure 9.6** (A) Just-noticeable roughness (*line*) as a function of the frequency difference between two pure tones ($\Delta F$) and the lower frequency of the tones. Tones were presented monaurally at 60 phons. The *bars* show the interquartile range of 20 subjects (musical background not given). (Adapted from Ref. 67 Figure 9.1, p. 883.) (B) Plomp and Levelt's data on consonance ratings as a function of $\Delta F$ for simultaneous pure tones with a mean frequency of 1000 Hz. Tones were presented in free field at 65 dB SPL. The *solid line* shows the mean consonance ratings of 10 subjects (musical background not given). *Dashed lines* show the interquartile range. (Adapted from Ref. 51 Figure 9.6, p. 554.) (C) Plomp and Levelt's idealized plot of the relationship between consonance and critical bandwidth. The *y* axis is in units of consonance (*left*) and dissonance (*right*). (From Plomp and Levelt[51] [Figure 9.10, p. 556]. Reproduced by permission.) (D) Terhardt's idealized plot showing the relationship of consonance and roughness to interval width (root at $A_4$, 440 Hz). The *solid line* shows consonance vs $\Delta F$ when the interval is composed of two pure tones. The *dotted line* shows consonance vs $\Delta F_0$ when the interval is composed of two complex tones, each containing several lower harmonics. The *dashed line* shows the roughness of two pure tones as a function of $\Delta F$. (From Terhardt[13] [Figure 9.1, p. 281]. Reproduced by permission.) (E) Kameoka and Kuriyagawa's data on consonance as a function of $\Delta F$ for simultaneous pure tones with the root at $A_4$ and intensities of 57 dB SPL. The *dashed line* is the mean performance of 22 audio engineers who performed the task twice (*solid lines* with *black* and *white triangles*). (From Kameoka and Kuriyagawa[72] [Figure 9.1, p. 1452]. Reproduced by permission.) (F) Kameoka and Kuriyzgawa's plot of dissonance as a function of interval width. The *dashed line* at the *top* shows the function when the interval is composed of two pure tones. The *solid line* at the *bottom* gives the calculated dissonance of an interval composed of two complex tones, each containing the first six harmonics at 57 dB SPL. (Adapted from Kameoka and Kuriyagawa,[35] Figure 9.8, p. 1465.)

roughness (line).[67] The data at $x = 500$ Hz would thus apply to the case of a harmonic interval composed of two pure tones with the root at $B_4$ ($x = 494$ Hz). The fifth of $B_4$, $F\sharp_5$, has a frequency of 741 Hz, so the $\Delta F$ between the root and the fifth is 247 Hz. According to Figure 9.6A, at $x = 500$ Hz, roughness disappears for $\Delta F$ values above 90–125 Hz. Therefore, the fifth should not be associated with roughness. Experimental studies agree that an isolated fifth composed of pure tones in this frequency register sounds 'consonant'[7,51] or 'pleasant'.[6,33] The pure-tone fifth is thus associated with the absence of roughness and with strong pitches associated with temporal regularities in its acoustic waveform and autocorrelation (similar to those described in the preceding section for a fifth composed of two harmonic complex tones; Figure 9.1H and L). The same set of observations applies to a perfect fourth composed of two pure tones.

Now consider the case of a minor second composed of two pure tones with the root at $B_4$ (494 Hz). Here, $\Delta F$ (between $B_4$ and $C_5$) is 33 Hz. This falls well within the range of noticeable roughness (Figure 9.6A). In fact, it lies near the $\Delta F$ associated with maximal roughness[67–69] (not shown). Experimental studies agree that a pure-tone minor second (and other tone combinations close to it) sounds 'dissonant'[7,51] or 'unpleasant'.[6,33]

The case of a tritone composed of two pure tones provides an interesting test of the roughness hypothesis. The $\Delta F$ between a tritone at $F_5$ and a root at $B_4$ is 201 Hz, well above the $\Delta F$ for just-noticeable roughness. Thus the fourth, tritone, and fifth are all above the roughness range. Does that mean they all have the same consonance?

Figure 9.6B shows Plomp and Levelt's data on consonance ratings as a function of $\Delta F$ for two pure tones whose mean frequency is 1000 Hz.[51] Because the $y$ axis is an ordinal scale, not an interval or ratio scale, it is inappropriate to assume that equal distances reflect equal differences in consonance. It follows that $4 > y > 4$ is not to be taken as the categorical boundary for dissonance and consonance, respectively. In addition, because Plomp and Levelt intentionally avoided using standard intervals like the fourth and tritone (they were concerned that interval recognition would influence consonance ratings), it is difficult to estimate where on the curve these intervals would fall. These caveats aside, it is clear that all pure-tone combinations with $\Delta F$s above approximately 150 Hz are consonant. This would apply to the fourth, tritone, and fifth with their roots in the vicinity of $A_5$. Tone combinations with $\Delta F_0 = 20$–80 Hz are dissonant; this would apply to a minor second with the root near $A_5$. Superficially, it would appear that we have a convergence between the disappearance of roughness at $\Delta F_0 = 150$–250 Hz (Figure 9.6A, interquartile range for a lower frequency of 1 kHz) and a steep increase in consonance ratings at $\Delta F_0 > 80$ Hz (at and above a minor third, Figure 9.6B). However, the $\Delta F$ associated with the highest mean consonance rating ($\Delta F \sim 180$ Hz) is within the range of noticeable roughness for many of Plomp and Steeneken's subjects.[67]

Beyond about $\Delta F = 180$ Hz, mean consonance ratings vary by only one rating point or less, but they are not perfectly flat (Figure 9.6B). One can discern alternating peaks and valleys out to about $\Delta F = 1200$ Hz. We estimated where the minor second, perfect fourth, tritone, and perfect fifth might fall on the interpolated lines drawn by Plomp and Levelt[51] (Figure 9.6B), and we estimated the frequency ratios and intervals that correspond to the peaks and valleys beyond $\Delta F = 1000$ Hz. The first peak is near 6:5, which would correspond to a minor third. The second peak falls close to 5:3, which would correspond to a

major sixth (or inverted minor third). The ratio 5:3 could also be thought of as the fifth and third harmonics of a harmonic series corresponding to the third and first notes of a major triad in its second inversion. The third peak is at or close to a ratio of 3:1, which corresponds to the interval of a twelfth, that is, a root and a fifth in the octave above. High consonance ratings for the twelfth are also found in Plomp and Levelt's data for two pure tones with a mean frequency of 500 Hz[51] (not shown). The first valley between the first two peaks is near the tritone, and the second valley appears to be a mistuned octave, with a frequency ratio near 2.03:1.

Figure 9.6C shows Plomp and Levelt's idealized plot of the relationship between consonance and critical bandwidth[51] (the latter is defined here by loudness summation[70]). Note that the curve reaches an asymptote near the end of the $x$ axis, at about one critical bandwidth. Thus a critical band account of consonance as the absence of roughness *cannot* apply to pure-tone intervals that are wider than a minor third or so.

Yet Terhardt's[13] idealized plot (Figure 9.6D) of the relationships among roughness, consonance, and pure-tone $\Delta F$ shows a monotonic increase in perceived consonance all the way out to the octave, well beyond the $\Delta F$ at which (1) roughness disappears (Figure 9.6A and D), (2) consonance ratings plateau (Figure 9.6B), and (3) loudness summation and masking effects are observed (Figure 9.6C; for review see Ref. 71). The representation of the psychoacoustic literature summarized in Figure 9.6D appears to draw upon Kameoka and Kuriyagawa's data showing increases in consonance well beyond the $\Delta F$ associated with disappearance of roughness[72] (Figure 9.6E). Although it is generally accepted that Kameoka and Kuriyagawa's work supports the idea that consonance is a function of roughness and critical bandwidth, comparisons of Figure 9.6A, C, and E reveal that their data actually argue against it, at least for musical interval widths greater than the $\Delta F$ for just-noticeable roughness.

The disagreement may arise from two sources. First, Kameoka and Kuriyagawa's Japanese audio engineers were instructed to judge tones for *sunda* (which they translate as 'clearness' in English) and *nigotta* ('turbidity').[72] Consequently, these listeners may have been rating different perceptual attributes than Plomp and Levelt's Dutch subjects, who were instructed to judge how 'consonant' [or *mooi* ('beautiful') or *welluidend* ('euphonious')] the intervals sounded.[51] Second, Kameoka and Kuriyagawa used an incomplete paired comparison paradigm—incomplete because only three or four adjacent intervals were paired for comparisons of relative consonance,[72] a much more restricted $\Delta F$ range than the one Plomp and colleagues used in their one-interval consonance rating paradigm.[7,51] When Kameoka and Kuriyagawa tried the method of magnitude estimation, presumably using all possible pairings, the task turned out to be 'rather difficult' for 'naive subjects'[72] (p. 1453), and they dropped it in favour of incomplete pairings. Comparing only adjacent intervals may have biased subjects to focus on differences they would not have otherwise attended to if all intervals had been paired with one another. The pattern of results suggests that pitch height, rather than absence of roughness, influenced consonance judgments beyond a minor third or so. Since the authors used these data to calculate the consonance of musical intervals formed by two complex tones (Figure 9.6F), they may have confounded roughness and pitch height in their predictions.

We reviewed previous studies that used isolated minor seconds, fourths, tritones, and fifths composed of two pure (or nearly pure) tones as experimental stimuli. Kaestner,[6] who

used a *Stimmgabelklangen* to generate tones that were 'poor in over-tones', found that subjects judged the fourth to be slightly more 'pleasant' than the tritone. Malmberg,[32] who used tuning forks, found a more marked preference for the fourth over the tritone for judgements of 'blending', 'purity', and 'smoothness'. Pratt,[73] who used a Stern variator that may have produced weak overtones, found that the fourth was judged to be more 'pleasant', 'smoother', and more 'unitary' than the tritone. Brues,[74] using a Stern variator that produced weak energy at the first overtone, found the fourth, tritone, and fifth were similar with respect to 'unitariness'. Guthrie and Morrill[75] used a Stern variator that produced 'very faint traces of the third partial' and reported that the fourth was judged to be more 'pleasant' than the tritone and of equal 'consonance'. Guernsey,[33] who used tuning forks, reported that nonmusicians, amateur musicians, and professional musicians found the fourth more 'pleasant' and 'smooth' than the tritone. Schellenberg and Trehub's recent experiments with nine-month-old babies are also relevant here.[76] When the upper pure tone of a repeating harmonic interval was flattened by one-fourth of a semitone, infants could detect the change if the interval was a fourth but not if it was a tritone. Their findings indicate that fourths provided a more stable background against which changes in tuning could be detected. All in all, these results indicate that the fourth, even when it is composed of two pure tones, is often perceived as more con-sonant than the tritone.

Another challenge for roughness-based accounts of consonance arises when we compare the consonance of pure-tone intervals and the consonance of complex-tone intervals. In Figure 9.6D, Terhardt[13] plots consonance ratings for pure-tone and complex-tone intervals on the same scale. In fact, Kameoka and Kuriyagawa's psychoacoustic data and calculations put them on different scales[35] (Figure 9.6F). Likewise, Plomp and Levelt use a dissonance scale from one to zero for their pure-tone data (Figure 9.6C) and a dissonance scale from six to zero for their complex-tone calculations (not shown).[51] Kameoka and Kuriyagawa's calculations predict that a minor second composed of two pure tones is more consonant than the unison of two complex tones with the first six harmonics at isoamplitude[35] (Figure 9.6F). Intuitively, this notion is untenable; however, direct comparisons between pure-tone intervals and complex-tone intervals have not been reported in the literature. We synthesized a pure-tone minor second and the unison of two complex tones (with the acoustic parameters specified by Kameoka and Kuriyagawa[35]) and asked several of our students to judge which of these two stimuli sounded more 'consonant'. These and Huron's (personal communication) informal observations raise the possibility that a combination of pitch height effects and loudness (shrillness), rather than or in addition to roughness, accounts for Kameoka and Kuriyagaw's predictions. For example, in the case of unison at $A_4$ (Figure 9.6F), spectral energy extends all the way up to 2200 Hz (fifth harmonic) and 2640 Hz (sixth harmonic), so there are high-frequency components that are greater in sensation level than the note $F_0$s. At the same time, $\Delta F$ (440 Hz) is higher than the highest $\Delta F$ associated with just-noticeable roughness when the root is at 2000 Hz (Figure 9.6A, interquartile range for $\Delta F \sim$ 250–400 Hz[67]).

In summary, the neural coding mechanisms that provide representations of roughness form part of the neurobiological foundation for the theory of harmony in its vertical dimension. However, our reappraisal of the psychoacoustic literature leads us to conclude that the dependence of consonance on the absence of roughness is overstated. We believe

pitch relationships, as well as roughness, influence the perception of intervals and chords as consonant or dissonant in the vertical dimension.

## Effects of auditory cortex lesions

Another approach to assessing the relative contributions of pitch and roughness to consonance perception might be to determine whether impairments in consonance perception caused by brain lesions are associated with deficits in one, the other, or both.

Consonance perception has been reported to be severely impaired following bilateral lesions of the auditory cortex.[15] In an experiment employing a one-interval, two-alternative, forced-choice paradigm, two types of stimuli were presented: a major triad, and a triad whose fifth was flattened by a fraction of a semitone. In each trial of the experiment, a young stroke patient, MHS, was asked if a single, isolated chord sounded 'in tune' or 'out of tune'. His response accuracy was 56 per cent, better than chance ($p < 0.05$), but more than two standard deviations below the mean of 13 normal controls (Figure 9.7A). Magnetic resonance imaging revealed that MHS's infarcts involved the primary auditory cortex in both hemispheres, all or almost all of the auditory association cortex in the right hemisphere, and about 20 per cent of the posterior auditory association cortex in the left hemisphere (Figure 9.2A). His pure-tone audiograms were within normal limits. Speech perception was impaired.

We subsequently compared MHS's performance on in-tune trials vs out-of-tune trials. If roughness perception were impaired, then MHS might make more errors in in-tune trials than out-of-tune trials. If frequency selectivity were coarsened and pitch perception impaired, then MHS might make more errors on in-tune trials. Consistent with the latter possibility, we found a marked response bias for out-of-tune judgments (Figure 9.7B). Two possible interpretations follow: (1) MHS was having difficulty extracting the pitches of chord frequency components and analysing their harmonic relationships; and/or (2) he heard more roughness in the chords than normals because his effective critical bandwidths were wider.

To assess whether MHS was having difficulty with frequency discrimination, we examined his performance on the Pitch Discrimination subtest of the Seashore Measures of Musical Talents.[77,78] This test uses the method of constant stimuli to measure one's ability to judge whether the second of two pure tones is higher or lower in pitch than the first tone. The $\Delta F$ between the tones gets smaller over successive blocks of trials. The tones were centred at 500 Hz, 600 ms in duration, and 600 ms apart, with an intensity of 35–40 dB above sensation level. Overall, MHS scored in the 15th percentile. His error pattern was again revealing. He performed poorly in the last third of the test, where the $\Delta F$s between the tones were smallest. We subsequently measured pure-tone frequency difference thresholds for pitch discrimination using an adaptive procedure and a two-interval, two-alternative, forced-choice paradigm.[79] Whereas normal controls and patient controls had frequency difference thresholds corresponding to Weber fractions ($\Delta F$/mean frequency) of around 1 per cent, MHS's Weber fractions were over 10 per cent. In short, his ability to judge the direction of a pitch change was markedly impaired. A similar deficit has since been reported in patients with surgical lesions of the right primary auditory cortex and right anterior auditory association cortex.[80] We also found that perception of the missing $F_0$ of

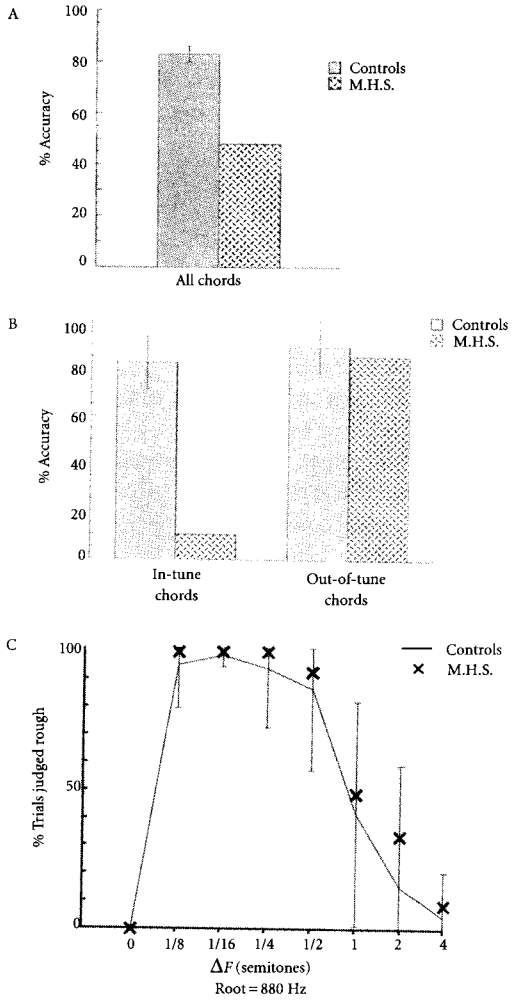



Figure 9.7 Perception of consonance (A and B) and roughness (C) in a patient, MHS, with bilateral lesions of auditory cortex (Figure 9.2A). (A) Data from Tramo *et al.*[15] showing response accuracy on a task that required MHS to determine whether a major triad was 'in tune' or 'out of tune'. (B) MHS's error rate as a function of stimulus condition. (C) Percentage of trials in which two simultaneous pure tones were judged to be fluctuating and rough (as opposed to steady and smooth) as a function of the $\Delta F\sharp$ between the tones. The mean (*line*) and standard deviation (*bars*) show the combined performance of 10 normal subjects and eight neurological patients who did not have bilateral lesions of auditory cortex. Three of the patients had chronic unilateral lesions of superior temporal cortex (left or right). and one had a lesion of the right inferior colliculus.

harmonic complex tones was impaired,[81] again consistent with observations in most right temporal lobectomy patients with partial or complete excisions involving the primary auditory cortex.[82] Taken together, these findings are consistent with the hypothesis that MHS's impaired consonance perception was related to deficits in pitch processing.

Was this increase in $\Delta F$ thresholds for pure-tone pitch discrimination associated with higher $\Delta F$ values for disappearance of roughness? We assessed MHS's ability to judge two simultaneous pure tones as 'steady and smooth' vs 'fluctuating and rough' using a one-interval, two-alternative, forced-choice paradigm and the method of constant stimuli. The lower tone was fixed at either 220 Hz ($A_3$) or 880 Hz ($A_5$), and the upper tone was above the root by a variable number of semitones: 0, 1/16, 1/8, 1/4, 1/2, 1 (a minor second), 2 (a major second), or 4 (a major third). Figure 9.7C shows the results when the root was at 880 Hz. When the tones were between 1/16 to 1/2 semitone apart, MHS, like controls, judged the combination as rough on $>80$ per cent of trials. When the tones were zero, two, or four semitones apart, MHS and controls judged the combination to be rough on less than 20 per cent of trials. At one and two semitones apart, MHS's performance fell near the mean of controls, but there is too much variability in the normal data to meaningfully assess MHS's performance. Still, these observations mitigate the possibility that consonance perception was impaired because he heard more roughness in chords than normals.

In summary, MHS's bias to hear major triads as mistuned appears to be associated with impairments in pitch perception but not roughness perception. Consistent with our physiological data and review of the psychoacoustic literature, this pattern of lesion effects indicates that pitch relationships influence harmony perception in the vertical dimension.

## Conclusions

Basic physiological and anatomical properties of auditory and cognitive systems determine why some combinations of simultaneous tones sound more harmonious than others. Distinctive acoustic features of consonant and dissonant intervals are translated into distinctive patterns of neural activity. A faithful representation of temporal regularities in the acoustic structure of consonant intervals exists in the population interspike interval (ISI) distribution of auditory nerve fibres. The most common ISIs in the distribution correspond not only to the pitches of note $F_0$s actually present in the consonant intervals, but also to the pitches of harmonically related notes in the bass register, such as the fundamental bass. By contrast, for dissonant intervals, the most common ISIs in the distribution do not correspond to one of the note $F_0$s, nor do they correspond to harmonically related notes. The relative strength of the missing $F_0$ in the population ISI distribution predicts the relative consonance of the minor second, perfect fourth, tritone, and fifth. Limits on the temporal precision and frequency selectivity of neurons throughout the auditory system constrain the range of note $F_0$s we can hear as strong pitches and how they are combined into intervals and chords. Implicit knowledge about the hierarchical relationships of pitches in a given tonal system is likely to exert cognitive influences on the degree to which intervals and chords sound consonant or dissonant, even when they are heard in isolation.

Representations of roughness exist in temporal patterns of neural activity at several levels of the auditory system. For the minor second, fourth, tritone, and fifth, the amount of 20- to 200-Hz temporal fluctuations in the firing patterns of auditory nerve fibres inversely correlates with perceived consonance. These representations of roughness are multiplexed with pitch representations in the spike trains of auditory nerve fibres. These two neural time codes operate over different time regimes. The fine timing of action potential

firing with precision in the submillisecond range carries information about fundamental pitch. Periodic fluctuations in discharge rate with precision in the range of milliseconds to tens of milliseconds carry information about roughness.

Bilateral lesions of primary auditory cortex and auditory association cortex can lead to severe impairments in consonance perception, with a bias to judge well-tuned chords as out of tune. In our patient, MHS, impaired consonance perception was associated with severely impaired pitch perception, but roughness perception appeared to be normal or near normal.

We interpret our findings and the results of previous psychoacoustic experiments as evidence in favour of the hypothesis that harmony in the vertical dimension, like harmony in the horizontal dimension, is principally a function of the pitch relationships among tones, with roughness playing a secondary role. In light of these observations, and in view of the likelihood that cognitive representations of pitch hierarchies influence harmony perception in the vertical dimension, we urge that the terms *sensory consonance* and *sensory dissonance* be reconsidered.

## Acknowledgements

This work was supported by NIH DC03382 (M.J. Tramo), DC03054 (P.A. Cariani), DC02258 (B. Delgutte), and DC00117 (L.D. Braida); and the McDonnell-Pew Program in Cognitive Neuroscience (M.J. Tramo). We gratefully acknowledge the instructive comments of David Cohen and Kay Shelemay of the Harvard Music Department; the technical support of Leslie Liberman of EPL, and David Lum, Andrew Grant, and Danielle Lemay of RLE; the graphics and manuscript assistance of Adda Kridler, Janelle Mallett, and Monique James; and, especially, our collaborators in the experimental studies cited here: Jamshed Bharucha, Martin McKinney, Michael Gazzaniga, Frank Musiek, Robert Zatorre, Carla Kovacs, Gaurav Shah, Eiling Yee. Special thanks go to Mark Smith and his family.

## References

1. **Cohen, M. R.** and **I. E. Drabkin** (1948) *A Source Book in Greek Science.* New York: McGraw-Hill Book Company, Inc.
2. **Piston, W.** (1941/1987) *Harmony,* 5th edn. Revised and Expanded by Mark DeVoto. New York: W. W. Norton & Company, Inc.
3. **Apel, Willi,** (ed.) (1972) *Harvard Dictionary of Music,* 2nd edn. Cambridge, MA: Belknap Press of Harvard University Press.
4. **Randel, D. M.** (1986) *The New Harvard Dictionary of Music,* 3rd edn. Cambridge, MA: Belknap Press of Harvard University Press.
5. **Tyrrell, J.** and **S. Sadie** (eds) (2001) *The New Grove Dictionary of Music and Musicians,* 2nd edn. New York: Grove's Dictionaries, Inc.
6. **Kaestner, G.** (1909) Untersuchungen uber den Gefuhlseindruck unanalysieter Zweiklange. *Psychol. Studien.* 4, 473–504.
7. **Van de Geer, J. P., W. J. M. Levelt,** and **R. Plomp** (1962) The connotation of musical consonance. *Acta Psychologia* 20, 308–19.

8. **Bernstein, L.** (1976) *The Unanswered Question.* Cambridge, MA: Harvard University Press.
9. **Gundlach, R. H.** (1935) Factors determining the characterization of musical phrases. *Am. J. Psychol.* 47, 624–43.
10. **Balkwill, L.** and **W. F. Thompson** (1999) A cross-cultural investigation of the perception of emotion in music: psycho physical and cultural cues. *Music Percept.* 17, 43–64.
11. **Castellano, M. A., J. J. Bharucha,** and **C. L. Krumhansl** (1984) Tonal hierarchies in the music of north India. *J. Exp. Psychol. Gen.* 113, 394–412.
12. **Kessler, E. J., C. Hansen,** and **R. N. Shepard** (1984) Tonal schemata in the perception of music in Bali and in the west. *Music Percept.* 2, 131–65.
13. **Terhardt, E.** (1984) The concept of musical consonance: a link between music and psychoacoustics. *Music Percept.* 1, 276–95.
14. **Krumhansl, C. L.** (1990) *Cognitive Foundations of Musical Pitch.* New York: Oxford University Press.
15. **Tramo, M. J., J. J. Bharucha,** and **F. E. Musiek** (1990) Music perception and cognition following bilateral lesions of auditory cortex. *J. Cognit. Neurosci.* 2, 195–212.
16. **Helmholtz, H.** (1877/1954) *On the Sensations of Tone as a Physiological Basis for the Theory of Music.* Translated by A.J. Ellis, 1885. New York: Dover Publications, Inc.
17. **Stevens, S. S.** and **H. Davis** (1938/1983) *Hearing: Its Psychology and Physiology.* New York: American Institute of Physics, Inc.
18. **Parncutt, R.** (1989) Harmony: *A Psychoacoustical Approach.* Berlin: Springer-Verlag.
19. **Deutsch, D.** (ed.) (1999) *The Psychology of Music,* 2nd edn. San Diego, CA: Academic Press.
20. **Demany, L.** and **C. Semal** (1990) Harmonic and melodic octave templates. *J. Acoust. Soc. Am.* 88, 2126–35.
21. **Demany, L., C. Semal,** and **R. P. Carlyon** (1991) On the perceptual limits of octave harmony and their origin. *J. Acoust. Soc. Am.* 90, 3019–27.
22. **Kiang, N. Y. -S., T. Watanabe, E. C. Thomas,** and **L. F. Clark** (1965) *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve.* Cambridge, MA: MIT Press.
23. **Rose, J. E., J. F. Brugge, D. J. Anderson,** and **J. E. Hind** (1967) Phase-locked response to low frequency tones in single auditory nerve fibers of the squirrel monkey. *J. Neuro-physiol.* 30, 769–93.
24. **Boomslitter, P.** and **W. Creel** (1961) The long pattern hypothesis in harmony and hearing. *J. Music Theory* 5, 2–31.
25. **Brugge J. F., D. J. Anderson, J. E. Hind,** and **J. E. Rose** (1969) Time structure of discharges in single auditory nerve fibers of the squirrel monkey in response to complex periodic sounds. *J. Neurophysiol.* 32, 386–401.
26. **Moore, B. C. J.** (1980) Neural interspike intervals and pitch. *Audiology* 19, 363–5.
27. **Rose, J. E.** (1980) Neural correlates of some psychoacoustic experiences. In D. McFadden (ed.) *Neural Mechanisms in Behavior.* New York: Springer-Verlag. pp. 1–37.
28. **Ohgushi, K.** (1983) The origin of tonality and a possible explanation of the octave enlargement phenomenon. *J. Acoust. Soc. Am.* 73, 1695–7.
29. **Patterson, R. D.** (1986) Spiral detection of periodicity and the spiral form of musical scales. *Psychol. Music* 14, 44–61.
30. **Moore, B. C. J., R. W. Peters,** and **B. R. Glasberg** (1985) Thresholds for the detection of inharmonicity in complex tones. *J. Acoust. Soc. Am.* 77, 1861–7.

31. **Vos, J.** (1982) The perception of pure and mistuned musical fifths and major thirds: thresholds for discrimination, beats, and identification. *Percept. and Psychophys.* 32, 297–313.

32. **Malmberg, C. F.** (1918) The perception of consonance and dissonance. *Psychol. Monogr.* 25, 93–133.

33. **Guernsey, M.** (1928) The role of consonance and dissonance in music. *Am. J. Psychol.* 40, 173–204.

34. **Butler, J. W.** and **P. G. Daston** (1968) Musical consonance as musical preference: a cross-cultural study. *J. Gen. Psychol.* 79, 129–42.

35. **Kameoka, A.** and **M. Kuriyagawa** (1969) Consonance theory part II: consonance of complex tones and its calculation method. *J. Acoust. Soc. Am.* 45, 1460–9.

36. **Zentner, M. R.** and **J. Kagan** (1998) Infants' perception of consonance and dissonance in music. *Infant Behav. and Dev.* 21, 483–92.

37. **Hulse, S., D. J. Bernard,** and **R. F. Braaten** (1995) Auditory discrimination of chord-based spectral structures by European starlings (Sturnus vulgaris). *J. Exp. Psychol. Gen.* 124, 409–23.

38. **Tramo, M. J., P. Cariani,** and **B. Delgutte** (1992) Representation of tonal consonance and dissonance in the temporal firing patterns of auditory nerve fibers. *Soc. Neurosci. Abstr.* 18, 382.

39. **Tramo, M. J., M. C. McKinney, P. A. Cariani,** and **B. Delgutte** (2000) Physiology of tonal consonance and dissonance. *Assoc. Res. Otolaryngol. Abstr.* 23, 275–6.

40. **McKinney, M. F., M. J. Tramo,** and **B. Delgutte** (2001) Neural correlates of musical dissonance in the inferior colliculus. In **A. J. M.** Houtsma *et al.* (eds) *Physiological and Psychophysical Bases of Auditory Function.* Maastricht, The Netherlands: Shaker Publishing, 71–7.

41. **McKinney, M. F., M. J. Tramo,** and **B. Delgutte** (2001) Neural correlates of the dissonance of musical intervals in the inferior colliculus. *Assoc. Res. Otolaryngol. Abstr.* 24, 54–55.

42. **Javel, E.** (1980) Coding of AM tones in the chinchilla auditory nerve: implications for the pitch of complex tones. *J. Acoust. Soc. Am.* 68, 133–146.

43. **Rhode, W. S.** (1995) Interspike intervals as a correlate of periodicity pitch in cat cochlear nucleus. *J. Acoust. Soc. Am.* 95, 2414–29.

44. **Schreiner, C. E.** and **G. Langner** (1988) Coding of temporal patterns in the central auditory nervous system. In *Auditory Function Neurobiological Bases of Hearing.* G. M. Edelman *et al.* (eds) New York: John Wiley & Sons pp. 337–61.

45. **Krishna, B. S.** and **M. N. Semple** (2000) Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. *J. Neurophysiol.* 84, 255–73.

46. **Preuss, A.** and **P. Muller-Preuss** (1990) Processing of amplitude modulated sounds in the medial geniculate body of squirrel monkeys. *Exp. Brain Res.* 79, 207–11.

47. **Bieser, A.** and **P. Muller-Preuss** (1996) Auditory responsive cortex in the squirrel monkey: neural responses to amplitude-modulated sounds. *Exp. Brain Res.* 108, 273–84.

48. **Fishman, Y. I., D. H. Reser, J. C. Arezzo,** and **M. Steinschneider** (2000) Complex tone processing in primary auditory cortex of the awake monkey. I. Neural ensemble correlates of roughness. *J. Acoust. Soc. Am.* 108, 235–46.

49. **Fishman, Y. I., D. H. Reser, J. C. Arezzo,** and **M. Steinschneider** (2000) Complex tone processing in primary auditory cortex of the awake monkey. II. Pitch versus critical band representation. *J. Acoust. Soc. Am.* 108, 247–62.

50. **Ehret, G.** and **R. Romand** (eds) (1997) *The Central Auditory System.* New York: Oxford University Press.

51. **Plomp, R.** and **W. J. M. Levelt** (1965) Tonal consonance and critical bandwidth. *J. Acoust. Soc. Am.* 38, 548–60.

52. **Hutchinson, W.** and **L. Knopoff** (1978) The acoustic component of Western consonance. *Interface* 7, 1–29.
53. **Rameau, J.-P.** (1722/1971) Treatise on Harmony. Translated by P. Gossett. New York: Dover Publications, Inc.
54. **Wightman, F. L.** and **D. M. Green** (1974) The perception of pitch. *Am. Sci.* 62, 208–15.
55. **Houtsma, A. J. M.** and **J. L. Goldstein** (1971) The central origin of the pitch of complex tones: evidence from musical interval recognition. *J. Acoust. Soc. Am.* 51, 520–9.
56. **Moore, B. C. J.** (1997) *An Introduction to the Psychology of Hearing.* San Diego, CA: Academic Press.
57. **Galileo Galilei** (1638/1954) *Dialogues Concerning Two New Sciences.* Translated by H. Crew and A. de Salvio, 1914. New York: Dover Publications.
58. **Pickles, J. O.** (1988) *An Introduction to the Physiology of Hearing.* London: Academic Press, Harcourt Brace Jovanovich, Publishers.
59. **Cariani, P. A.** and **B. Delgutte** (1996) Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *J. Neurophysiol.* 76, 1698–1716.
60. **Cariani, P. A.** and **B. Delgutte** (1996) Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase-invariance, pitch circularity, and the dominance region for pitch. *J. Neurophysiol.* 76, 1717–34.
61. **Cariani, P.** (1999) Temporal coding of periodicity pitch in the auditory system: an overview. *Neural Plast.* 6, 147–72.
62. **Meddis, R.** and **M. J. Hewitt** (1991) Virtual pitch and phase sensitivity of a computer model of the auditory periphery. II. Phase sensitivity. *J. Acoust. Soc. Am.* 89, 2883–94.
63. **Terhardt, E.** (1974) Pitch, consonance, and harmony. *J. Acoust. Soc. Am.* 55, 1061–9.
64. **Helmholtz, H. L. F.** (1863/1913) *Die Lehre von der Tonempfindungen als physiolo-gische Grundlage fur die Theorie der Musik,* 6th edn. Brauschweig: F. Vieweg, Quote translated by E. Terhardt,[13] pp. 283–4.
65. **Dai, H.** (1993) On the pitch of two-tone complexes. *J. Acoust. Soc. Am.* 94, 730–4.
66. **Greenwood, D. D.** (1991) Critical bandwidth and consonance in relation to cochlear frequency-position coordinates. *Hear. Res.* 54, 164–208.
67. **Plomp, R.** and **H. J. M. Steeneken** (1968) Interference between two simple tones. *J. Acoust. Soc. Am.* 43, 883–4.
68. **Mayer, A. M.** (1874) Researches in acoustics. *Am. J. Sci. Arts, 3rd Ser.* 8, 241–55.
69. **Cross, C. R.** and **H. M. Goodwin** (1891) Some considerations regarding Helmholtz's theory of consonance. *Am. J. Sci.* 58, 1–12.
70. **Zwicker, E., G. Flottorp,** and **S. S. Stevens** (1957) Critical band width in loudness summation. *J. Acoust. Soc. Am.* 29, 548–57.
71. **Yost, W. A.** (2000) Fundamentals of Hearing. San Diego, CA: Academic Press.
72. **Kameoka, A.** and **M. Kuriyagawa** (1969) Consonance theory part I: consonance of dyads. *J. Acoust. Soc. Am.* 45, 1451–9.
73. **Pratt, C. C.** (1921) Some qualitative aspects of bitonal complexes. *Am. J. Psychol.* 32, 490–515.
74. **Brues, A. M.** (1927) The fusion of non-musical intervals. *Am. J. Psychol.* 38, 624–38.
75. **Guthrie, E. R.** and **H. Morrell** (1928) The fusion of non-musical intervals. *Am. J. Psychol.* 40, 624–625.

76. **Schellenberg, E. G.** and **S. E. Trehub** (1996) Natural musical intervals: evidence from infant listeners. *Psychol. Sci.* 7, 272–7.

77. **Tramo, M. J.** (1990) Impaired perception of relative pure tone pitch following bilateral lesions of auditory cortex in man. *Soc. Neurosci. Abstr.* 16, 580.

78. **Seashore, C. E., D. Lewis,** and **J. G. Saetveit** (1960) *Seashore Measures of Musical Talents.* New York: The Psychological Corporation.

79. **Tramo, M. J., A. Grant,** and **L. D. Braida** (1994) Psychophysical measurements of frequency difference limens for relative pitch discrimination reveal a deficit following bilateral lesions of auditory cortex. *Soc. Neurosci. Abstr.* 20, 325.

80. **Johnsrude, I. S., V. B. Penhune,** and **R. J. Zatorre** (2000) Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain* 123, 155–63.

81. **Bharucha, J. J., M. J. Tramo,** and **R. J. Zatorre** (1993) Abstraction of the missing fundamental following bilateral lesions of auditory cortex. *Soc. Neurosci. Abstr.* 19, 1687.

82. **Zatorre, R. J.** (1988) Pitch perception of complex tones and human temporal-lobe function. *J. Acoust. Soc. Am.* 84, 566–72.

CHAPTER 10

# INTRACEREBRAL EVOKED POTENTIALS IN PITCH PERCEPTION REVEAL A FUNCTIONAL ASYMMETRY OF HUMAN AUDITORY CORTEX

CATHERINE LIÉGEOIS-CHAUVEL, KIMBERLY GIRAUD, JEAN-MICHEL BADIER, PATRICK MARQUIS, AND PATRICK CHAUVEL

## Abstract

One acoustic feature that plays an important role in pitch perception is frequency. Studies on the processing of frequency in the human and animal brain have shown that the auditory cortex is tonotopically organized: low frequencies are represented laterally whereas high frequencies are represented medially. To date, the study of the functional organization of the human auditory cortex in the processing of frequency has been limited to the use of either scalp-recorded auditory evoked potentials (AEPs), which have relatively poor spatial resolving power, or functional imagery techniques, which have poor temporal resolving power. The present study uses intracerebrally recorded AEPs to explore this topic in the primary and secondary auditory cortices of both hemispheres of the human brain. Recordings were carried out in 45 adult patients with drug-resistant partial seizures. In the right hemisphere, clear spectrally organized tonotopic maps were observed with distinct separations between different frequency-processing regions. AEPs for high frequencies were recorded medially, whereas AEPs for low frequencies were recorded laterally. In the left hemisphere, however, this tonotopic organization was less evident, with different regions involved in the processing of a range of frequencies. The hemisphere-related difference in the processing of tonal frequency is discussed in relation to pitch perception.

**Keywords:** Pitch perception; Auditory cortex; Neural processing

The ability to process pitch is of major importance to music and language perception. Although the acoustic cues necessary to the perception of pitch are mainly of a temporal nature, human subjects describe pitch as having a spatial dimension (i.e. musical scale). Traditionally, the literature on the processing of pitch in the human brain has been

dominated by two main theories: one that postulates an exclusively tonotopic coding and another that is in favour of an exclusively 'temporal' coding.[1–3] Recently, however, Langner[4,5] has suggested that temporal information, coded at the subcortical level as spikes synchronized to the periodicity of the acoustic signal, is subsequently converted to spatial information coded in the form of tonotopic maps in the auditory cortex. This hypothesis is consistent with the finding that the temporal resolution of the auditory nerve[6] and inferior colliculus[7,8] may be higher than at the cortical level, where the temporal resolution of auditory information is possible only at frequencies below 100 Hz for the majority of neurons.[9,10] At the cortical level, pitch appears to be coded according to a frequency spectrum (i.e. tonotopically).

Experimental studies using single-unit recording techniques indicate that for most auditory neurons, a 'best frequency' (BF) can be determined at which a low-intensity auditory stimulus evokes the greatest electrophysiological response in a given region. The auditory cortex is said to be tonotopically organized because an ordered change in BFs occurs as one moves along the cortex. The auditory cortex is organized according to a frequency gradient, with low frequencies represented laterally and high frequencies represented medially.[11–18]

The tonotopic organization of the auditory cortex is best understood in nonhuman animals. The cytoarchitecture of the auditory cortex is not always uniform,[14] however, rendering the mapping of different tonotopic regions difficult. Merzenich and Brugge[12] have found that the auditory cortex in the superior temporal plane of the macaque can be subdivided into at least six distinct fields on the basis of cytoarchitectonic characteristics and BFs. Likewise, Reale and Imig[16] observed a tonotopic organization in six different auditory areas of the cat, and a similar tonotopic organization has been described in the albino rat.[18]

In humans, the use of noninvasive techniques such as magnetoencephalography (MEG) and positron emission tomography (PET) has shed some light on the functional organization of the auditory cortex with respect to frequency. In their PET study, Lauter *et al.*[19] observed that changes in regional cerebral blood flow induced by a high-frequency tone (4 kHz) were located deeper and more posteriorly in the temporal lobe than those induced by a low-frequency tone (500 Hz). In a similar vein, MEG studies on the tonotopic organization of N1 generators have shown that their estimated depth can vary as a function of frequency: higher frequencies involve more medial regions of the Heschl's gyrus (HGs), whereas lower frequencies involve more lateral areas.[20–24]

In a previous study,[25] using depth electrodes implanted in the HG of human subjects with medically intractable epilepsy, we identified four separable components of middle-latency auditory evoked potentials (MLAEPs), which peaked at approximately 30, 50, 60, and 80 ms. The sources of these components were definable, although some overlap was observed. In a recent MEG study, Pantev *et al.*[23] explored the sources of different middle-latency components as well as their distribution. In all subjects and for all the frequencies studied, only two components were consistently observed: the 30-ms 'Pam' component and the N1m. Also, a tonotopic organization was observed in the primary auditory cortex that mirrored that observed in the secondary cortex. The localization of Pam's source moved more superficially with higher frequencies, while the localization of the N1m's source moved progressively deeper with higher frequencies. The authors, however, did not consistently observe the

50-ms component, which is thought to be generated in the lateral part of the primary auditory cortex in close proximity to the neighboring secondary areas.

The present paper explores the tonotopic organization of the human auditory cortex using intracerebrally recorded evoked potentials studied as a function of anatomical recording site. The sensitivity of a neuronal population to a given frequency was determined from fluctuations in AEP amplitude between different recording sites in the primary auditory cortex and surrounding secondary areas like the planum temporale. Indeed, previous data suggest that the auditory cortex is tonotopically organized for frequency and that even several different tonotopic maps may be described for different cortical regions. But is this tonotopic organization the same in both left and right cerebral hemispheres? A substantial body of data suggests that the perception and processing of pitch is specific to a single hemisphere,[26] and hemisphere-related differences in the functional organization of the auditory areas underlying such processing are not inconceivable. The main objective of this study was to investigate the tonotopic organization of the human auditory cortex in both cerebral hemispheres.

## Patients and methods

Intracerebral AEPs were recorded from 45 adult patients (aged 20–56 years) with medically intractable partial epileptic seizures of right and left temporal lobe origin (in 31 and 14 patients, respectively). These patients were undergoing stereo-electroencephalographic (SEEG) evaluation for the surgical relief of their seizures (for a description of this procedure, see Refs 27 and 28). Recordings were carried out in either the left or the right HG and right planum temporale, during which time subjects laid comfortably on a bed and remained vigilant and cooperative. All patients had previously given informed consent to participate in the study.

The anatomical localization of each implanted electrode was based on a stereotaxic procedure described elsewhere.[29–31] The three-dimensional coordinates of each electrode lead were defined and measured with respect to the temporal branch of the middle cerebral artery and allowed us to distinguish the primary from the secondary auditory areas in the different patients.[25,32]

The visualization of electrode tracks using stereotaxic MRI represents a means of confirming the anatomical localization of implanted sites and provides additional information on the anatomy of the HG in each subject (e.g. the presence of one or two gyri[33]). Such information can often be helpful in the interpretation of AEP data. MRI sections were cut along three spatial planes using bicommissural coordinates. Figure 10.1 shows the position of one electrode implanted orthogonally in the lateral part of the left HG.

### Auditory stimulation

Auditory stimuli were tone bursts (with a 0.3-ms rise and decay time and a total duration of 30 ms) of different frequencies (ranging from 250 Hz to 4 kHz). These stimuli were generated using an SM 700 NICOLET auditory stimulator and presented at a rate of 0.7/s via TDH 39 earphones to the ear contralateral to the implanted site. Stimuli were presented in random order at an intensity of 70 dB SL (sensation level) for each subject.

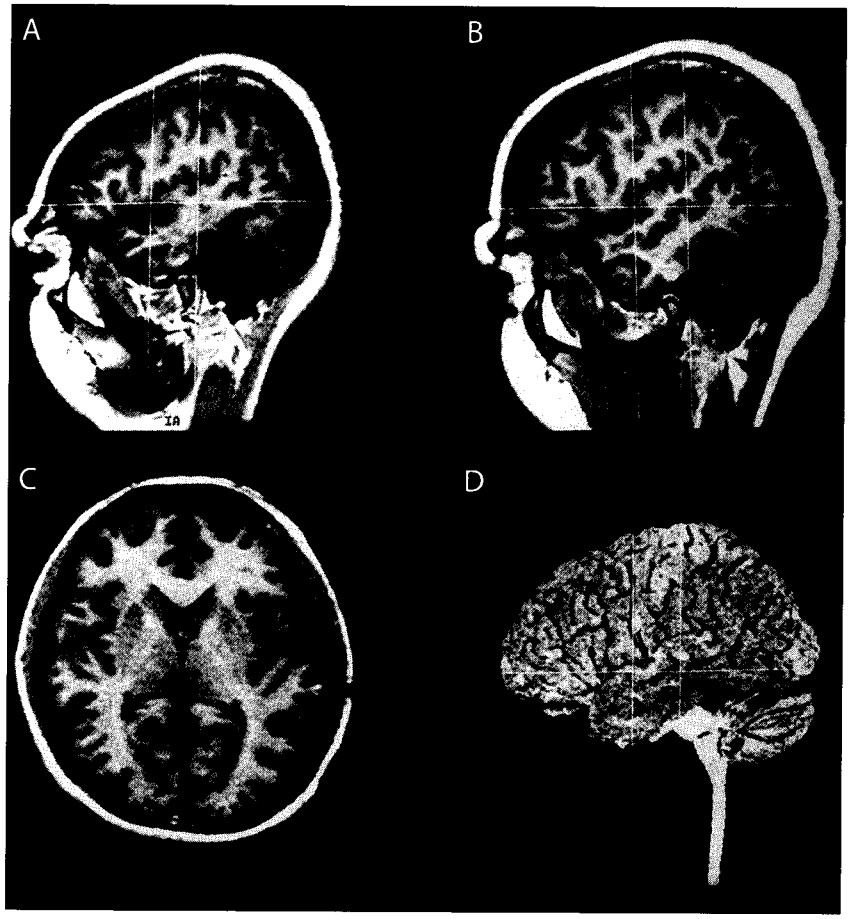



Figure 10.1 Sagittal MRI sections show the trajectory of the Sylvian fissure in the (A) left and (B) right temporal lobes. Note here the distinct shape of the end of the lateral sulcus (Sylvian fissure) on both sides (*white arrows*) and the longer length of the left planum temporale in this right-handed patient. The *three axes* of the bicommissural reference system (AC-PC, VAC, VPC) have been drawn. (C) The verticofrontal section shows the left Heschl's gyrus, and the *black arrow* indicates the trace of the electrode recorded from the auditory cortex. (D) On the surface view, the *arrow* shows the impact point of the same electrode.

## Depth AEP recording and averaging

Intracerebral AEPs were recorded from all leads of an electrode (using an extra-cephalic reference looped around the neck) and were simultaneously amplified and bandpass filtered 1 Hz–1.5 kHz. Recording epochs lasted 350 ms with a prestimulus time of 35 ms, 10 per cent of the epoch. Two trials of 100 samples were administered to each subject. Variations (i.e. polarity inversions and/or changes in amplitude in AEPs) recorded from

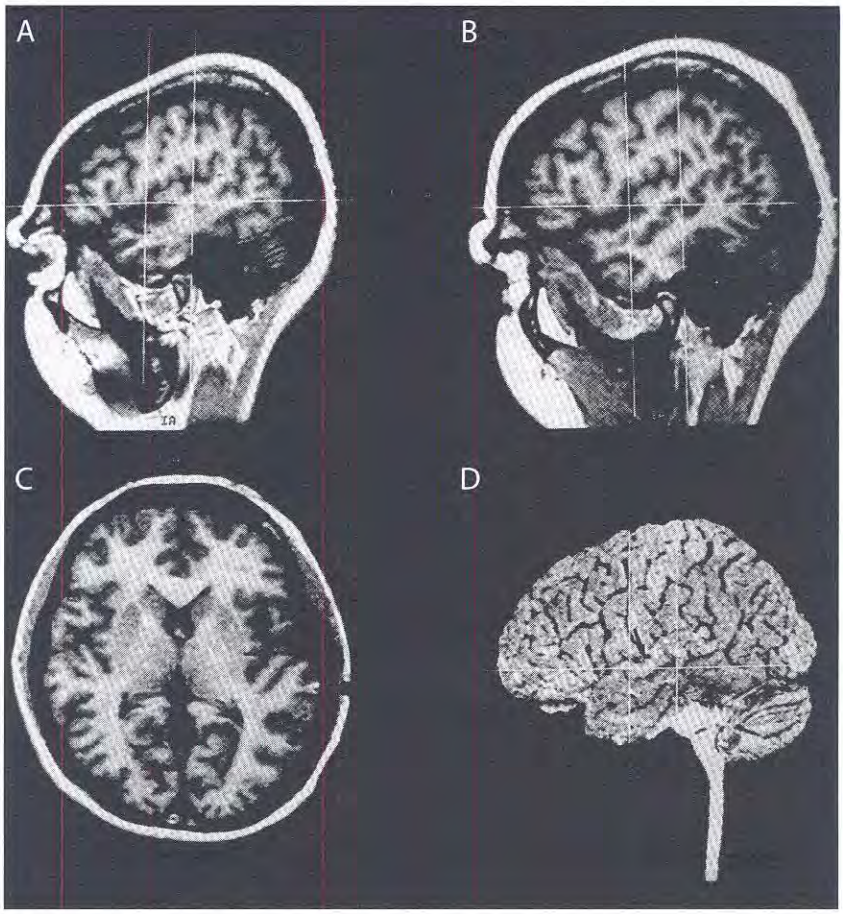



**Figure 10.1** Sagittal MRI sections show the trajectory of the Sylvian fissure in the (A) left and (B) right temporal lobes. Note here the distinct shape of the end of the lateral sulcus (Sylvian fissure) on both sides (*white arrows*) and the longer length of the left planum temporale in this right-handed patient. The *three axes* of the bicommissural reference system (AC-PC, VAC, VPC) have been drawn. (C) The verticofrontal section shows the left Heschl's gyrus, and the *black arrow* indicates the trace of the electrode recorded from the auditory cortex. (D) On the surface view, the *arrow* shows the impact point of the same electrode.

### Depth AEP recording and averaging

Intracerebral AEPs were recorded from all leads of an electrode (using an extra-cephalic reference looped around the neck) and were simultaneously amplified and bandpass filtered 1 Hz–1.5 kHz. Recording epochs lasted 350 ms with a prestimulus time of 35 ms, 10 per cent of the epoch. Two trials of 100 samples were administered to each subject. Variations (i.e. polarity inversions and/or changes in amplitude in AEPs) recorded from

different leads were studied for each patient. Tonotopic maps were based on a comparison of the results obtained from patients with different anatomical recording sites.

The acquisition system was a Nicolet Pathfinder II with 16 acquisition channels and an 8-bit-resolution digital converter. Negative potentials were displayed as up and positive as down. Recordings of brainstem evoked potentials carried out before SEEG recordings confirmed that all the patients studied had normal afferent pathway conduction.

### Temporospectral mapping

Data were stored on the hard drive of the computer and transferred via an asynchronous line to an IBM PC/AT. Off-line processing consisted of the construction of temporospectral maps, using a technique adapted from Rémond,[34] to depth recordings, particularly appropriate to the depth electrode geometry.[35,36]

## Results

Primary and secondary areas were identified on the basis of component latency (for a description, see Ref. 25). The selectivity of an auditory area's response (and hence its BF was determined from frequency-related fluctuations in a component's amplitude and/or polarity inversions.

### Recordings from the primary auditory cortex

**Right hemisphere** Figure 10.2 shows AEPs recorded from 10 contiguous leads of an electrode implanted in the right auditory cortex of one patient. The first four leads recorded activity from the primary auditory cortex; the other six recorded activity from secondary areas. Tone bursts at different frequencies were presented to the patient's left ear. A triphasic N30/P50/N80 complex was recorded from the first four leads for which a polarity inversion was observed between leads 4 and 5, as well as a shift in latency between leads 6 and 10. AEPs recorded from lead 4 were highly sensitive to tonal frequency. For tones higher than 1 kHz, the amplitude of the 50- and 80-ms components progressively decreased until an isopotential response (note the polarity inversion between leads 3 and 5) was observed for tones at 4 kHz, suggesting that the cortical region at lead 4 has a high sensitivity for frequencies in the 4 kHz range. AEPs did not differ in terms of tonal frequency on any of the other leads.

It should be noted, of course, that the amplitude changes observed were never an all-or-none phenomenon, but were observed as increasing or decreasing from lead to lead in progressive steps. Also, unlike the 50- and 80-ms components, the amplitude of the 30-ms component did not appear to be frequency dependent.

Figure 10.3 is a temporospectral representation of AEPs recorded for different tonal frequencies from a second patient. In this figure, the amplitudes of the same three components (30, 50, and 80 ms) as for the first patient are shown for the three most responsive leads (2, 3, and 6 for different tonal frequencies). In each case, frequency is represented along the $x$-axis and latency along the $y$-axis. On lead 2, amplitude was maximal for tone bursts with the highest frequencies (3–4 kHz), although frequency-dependent fluctuations

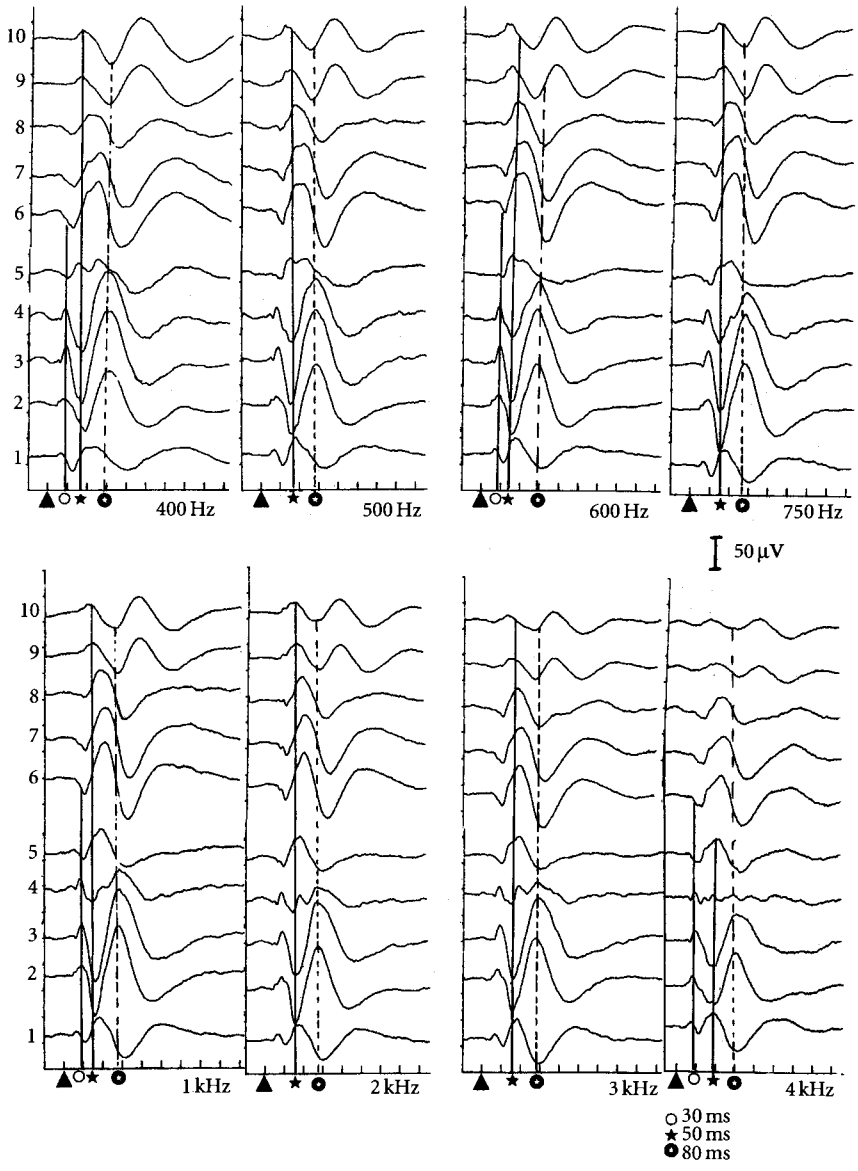



**Figure 10.2** Recordings from the medial part of the primary auditory cortex in right HG according to different tonal frequencies. Each response corresponds to a monopolar recording of each lead of the electrode that is referenced to an extracephalic site. The leads are numbered from 1 to 10, from the medial to lateral part of the auditory cortex. Note the amplitude variations of the 50-ms component (indicated by the *star*) according to frequency (see comments in the text). The *arrow* indicates the onset of the stimulus. Recording time ends at 350 ms.

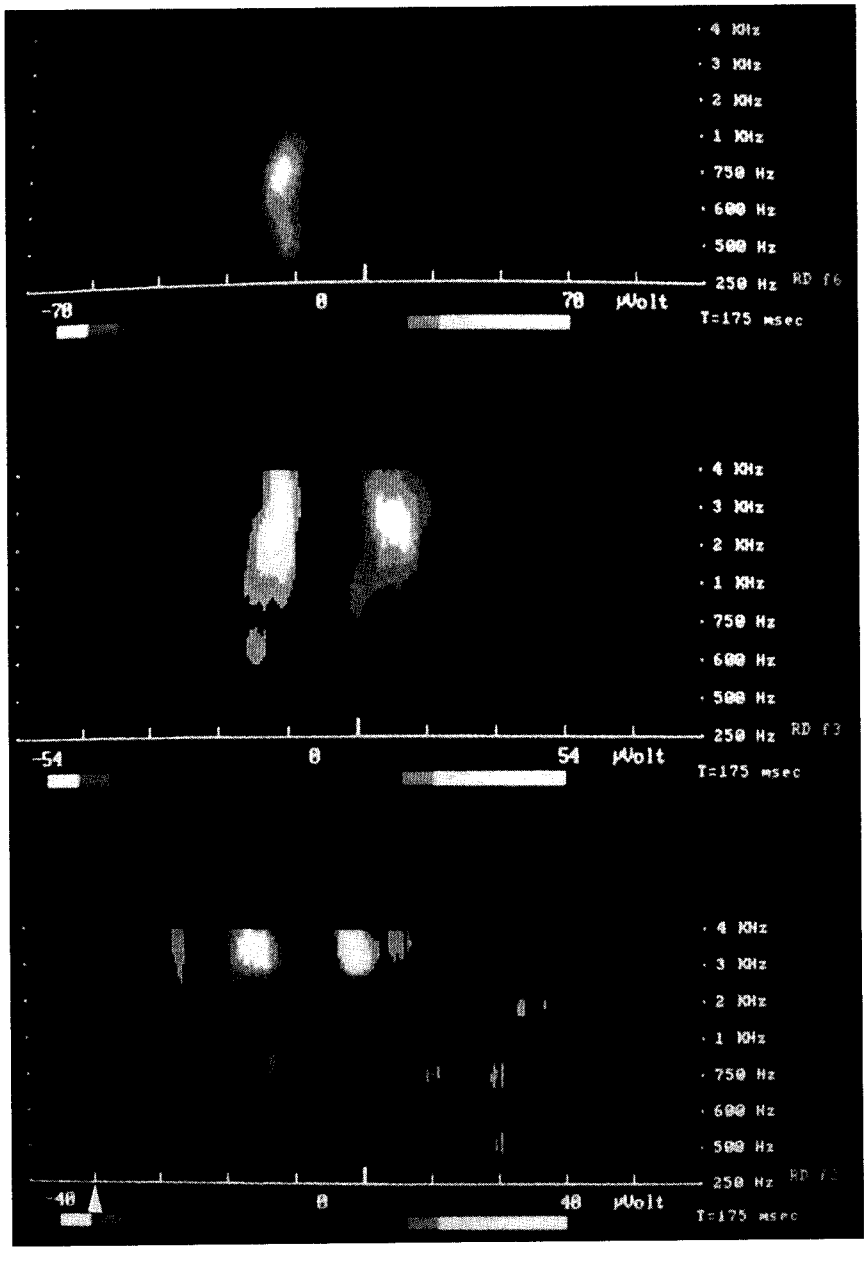



**Figure 10.3** Temporospectral maps of amplitude variations of potentials recorded from the right primary auditory cortex (leads 2, 3, and 6). Time, related in abscissae, is reduced to 175 ms, full scale. On the ordinate axis, each *yellow dot* represents the frequency. Amplitude is represented by colour. A cubic spline interpolation is used in order to obtain a continuous potential line between the different frequencies. Each spatial /reconstruction gives a continuous map both in time and in frequency. *Dark colours* are values close to the neutral potential, and *bright colours* are used for the activity—*blue* represents negativity and *red* represents positivity. The *white arrow* indicates the beginning of the sound. (For comments see text.) (See Plate 1 in colour section.)




**Figure 10.3** Temporospectral maps of amplitude variations of potentials recorded from the right primary auditory cortex (leads 2, 3, and 6). Time, related in abscissae, is reduced to 175 ms, full scale. On the ordinate axis, each *yellow dot* represents the frequency. Amplitude is represented by colour. A cubic spline interpolation is used in order to obtain a continuous potential line between the different frequencies. Each spatial /reconstruction gives a continuous map both in time and in frequency. *Dark colours* are values close to the neutral potential, and *bright colours* are used for the activity—*blue* represents negativity and *red* represents positivity. The *white arrow* indicates the beginning of the sound. (For comments see text.) (See Plate 1 in colour section.)

in amplitude were greater for the N50 and P80 components than for the N30. This was observed to an even greater extent on lead 3, where the amplitude of the N30 did not vary significantly according to tonal frequency, in spite of the highly significant fluctuations in P50 and N780 amplitude. On lead 3, the P50 and N80 were less sharply tuned and their amplitude was maximal at frequencies of 2–3 kHz. Lead 6 recorded neuronal activity in a region between the lateral part of the primary cortex and the medial part of the secondary cortex and yielded responses characterized, this time, by an almost complete absence of the 30-ms component. The N50 was the only component for which amplitude varied with tonal frequency. Its amplitude was maximal at 750 Hz.

**Left hemisphere** The left-hand side of Figure 10.4 shows AEPs recorded from the left primary auditory cortex (leads 4 and 5) for different tonal frequencies. As was the case in the right primary cortex, a triphasic N30/P45/N100 complex was observed. P45 and N100 amplitude recorded from the left hemisphere, however, did not vary as a function of

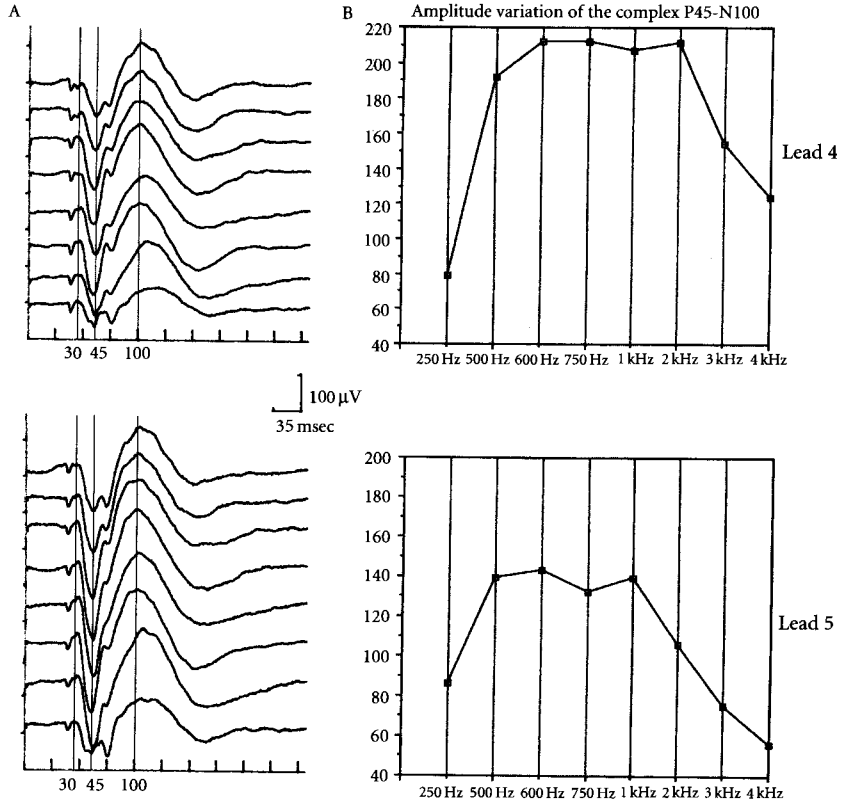



**Figure 10.4** (A) Evoked responses recorded from the medial part of the primary auditory cortex (leads 4 and 5) in the left HG. (B) Graphs representing amplitude variations for the P45/N100 complex according to different tonal frequencies (ordinate axis: absolute amplitude expressed in µV; abscissa axis: tonal frequency). The frequencies vary from 250 Hz (*first line*) to 4 kHz (*last line*) (see comments in the text).

frequency, as shown on the right in the same figure. For these components, amplitude was maximal at a range of frequencies between 600 Hz and 2 kHz.

### Recordings from the secondary auditory cortex in the left and right hemispheres

The secondary auditory areas studied were confined to an area anterior and lateral to the primary cortex, as previously identified in physiological[25,32] and morphological studies.[37,38] In this secondary area, frequency-dependent fluctuations in amplitude were observed for the 80-ms component. Figure 10.5 shows the localization of electrodes implanted in the right (left-hand side of figure) and left (right-hand side of figure) secondary auditory areas in 11 patients (six right; five left). The sagittal views (top) show the localization of each electrode, while the coronal views (below sagittal views) show the localization of those electrode leads yielding the greatest frequency-dependent fluctuations in amplitude. In the right hemisphere, a mediolateral as well as an anteroposterior distribution of BFs was observed for the 80-ms component. Neuronal populations sensitive to high frequencies (graphs A and E) 80-ms poststimulus onset were localized in the posterolateral part of right HG, while those sensitive to low frequencies (graph B) were localized in the anterolateral part of right HG. For the 80-ms component, high frequencies were also represented ventrally in an area adjacent to the lateral part of the primary cortex in the middle of right HG, overlapping the area in which low frequencies were represented for the 50-ms component.

The profile of amplitude curves for the right hemisphere was noticeably different from that for the left hemisphere. In the dorsolateral part of the left HG, the generators of the 80-ms component were most responsive over a range of high frequencies (1–4 kHz; graph B) or to two isolated sets of frequencies (500 Hz and 1–2 kHz; graph C); in the anterolateral part of left HG, this component was sensitive either to a range of low-to-medium frequencies (500 Hz–1 kHz) or to two isolated sets of frequencies (400 Hz and 1 kHz; graph D).

Figure 10.6 summarizes the main results. In the right primary cortex, the 30-ms component was not clearly tonotopic. A first tonotopic map is observed for the 50-ms component. A strict organization displaying high frequencies medially and low frequencies laterally is observed. A second map is observed for the 80- or 100-ms components generated in the secondary cortex. Neurons responsive to low frequencies were lateral to those responding to higher frequencies, but they were less finely tuned (broken line). Another map was observed for the 100-ms component generated in the planum temporale. High frequencies seemed to be more represented. On the left hemisphere, in the primary as well as in the secondary auditory cortex, the neurons were most sensitive to a range of frequencies (low or high). Nevertheless, the same tonotopic organization of the frequencies seems to be respected—high frequencies medially and low frequencies laterally represented.

## Discussion

The present study suggests that the auditory cortex is composed of frequency-dependent tonotopic maps and that these maps are more complex than previously reported in surface EP studies.[20,21,23,39] Although the previously reported mediolateral tonotopic organization was also observed in this study, it was not robust for all AEP components recorded and was hemisphere specific.

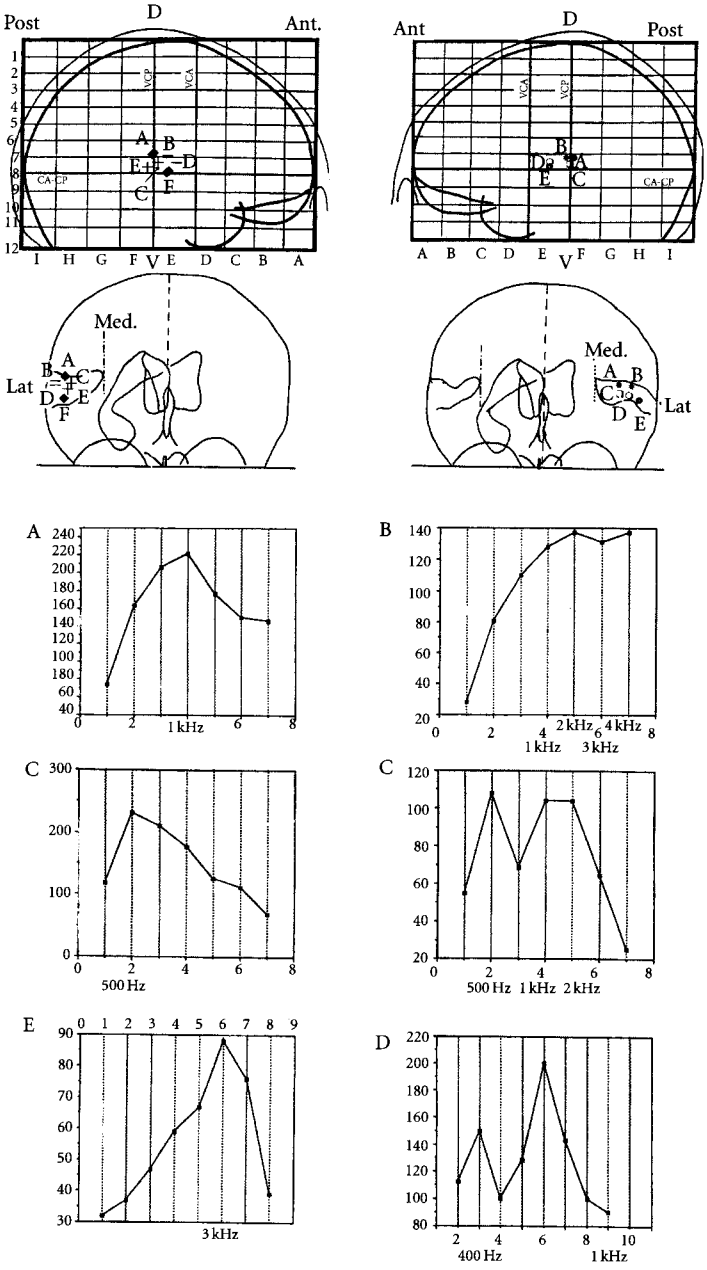



**Figure 10.5** Mapping of amplitude variations for the 80-ms component (ordinate axis: absolute amplitude expressed in μV; abscissa axis: frequency) recorded from different loci of the lateral part of the right and left HG referenced to Talairach's proportional standard grid on the lateral view. On the frontal view, the depth of the most responsive lead of each patient is reported in the Heschl's gyrus. The frequencies vary from 250 Hz *(first line)* to 4 kHz *(last line)*. The *peak of the curve* corresponds to the CF (see comments in the text).

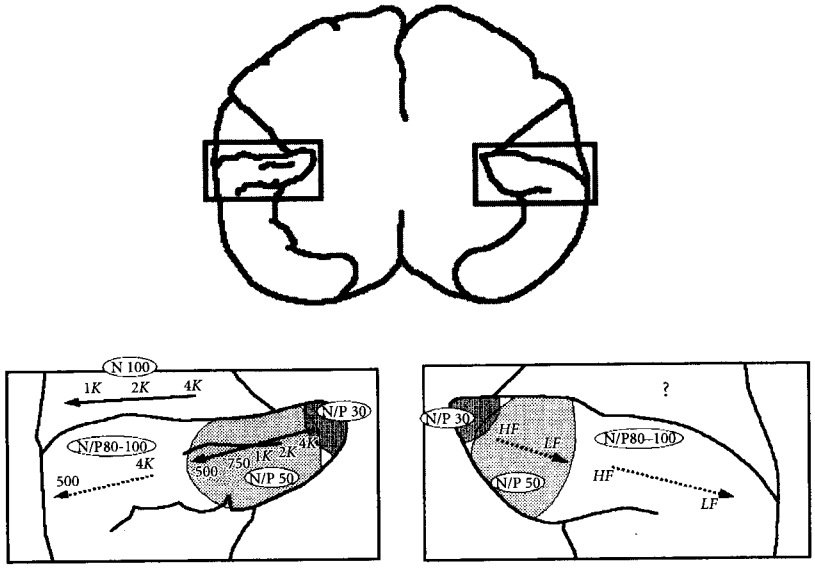



Figure 10.6 Schematic representation of the tonotopic maps recorded from the right and left auditory areas (see comments in the text).

In the primary auditory cortex, no frequency-dependent changes in amplitude or latency were observed for the earliest components (N13/P17) in either hemisphere. For the N30, only slight variations in amplitude were seen in response to tonal frequency, and most often its amplitude covaried with that of the 50-ms component. The amplitude of the 50-ms component was maximal for high frequencies (4 kHz) in medial regions of the primary cortex and for low frequencies (400–500 Hz) in lateral regions. Both mediolateral and anteroposterior organizations observed were marked by a progressive decrease in BFs toward the boundaries of the primary cortex.

The present data suggest the existence of an orderly, tonotopic representation of frequencies in the human primary auditory cortex, similar to that described in nonhuman animals.[13,40] In humans, BFs vary from 4 kHz medially to 400 Hz laterally for the 50-ms component. This range roughly corresponds to the frequency range useful for speech (this latter range actually begins near 250 Hz—a frequency at which a clear evoked response was difficult to obtain in this study). Studies exploring the auditory cortex of nonhuman primates have shed a great deal of light on the functional organization of the human auditory cortex, because its organization in primates appears to be similar to that in humans. In primates, BFs are represented in an orderly pattern in the primary auditory cortex and range from below 100 Hz to about 32 kHz, nearly the full frequency hearing range for primates. The lowest BFs are represented rostrally and laterally, whereas the highest BFs are found caudally and medially. Such an organization is thought to cover almost the entire auditory cortex, with all frequencies (except those at the extreme ends of the frequency hearing range) being represented.[2,16,41,42] The BFs observed in the present study were highly stable

from patient to patient, in spite of the intersubject variability in localization and orientation of the primary auditory cortex.

In a recent study, Pantev *et al.*[23] observed a tonotopic organization of the human auditory cortex for the Pam component (corresponding to the 30-ms intracerebral AEP component); unlike the results of the present study, however, higher frequencies were represented more laterally than were lower frequencies. This discrepancy may be explained by the fact that, in the present study, only slight frequency-dependent fluctuations in amplitude were observed for the 30-ms component, compared to those observed for the P/N 50. Furthermore, the P1m (corresponding to the P/N 50 in this study), generated in the primary auditory cortex, was not systematically discernible in Pantev *et al.*'s study and has been shown in several MEG[44–46] and EEG studies to have highly variable scalp topography. With intracerebrally recorded AEPs, on the other hand, the P/N 50 is robust and can reach amplitudes of up to 150 μV.[25] It is also probable that two different overlapping tonotopic maps exist, each corresponding to either the 30- or the 50-ms component, and that different investigation techniques have differential capacities to explore these maps.

Frequency-dependent fluctuations in amplitude were also observed for the 80-ms component, which is generated in secondary auditory areas anterior and lateral to the primary cortex.[25] In the right hemisphere, a mediolateral as well as an anteroposterior tonotopic organization was seen, with high frequencies represented posteromedially and low frequencies represented anterolaterally. High frequencies were also represented ventrally in an area adjacent to the lateral part of the primary cortex, which was also the area in which low frequencies were represented for the 50-ms component. While the distribution of BFs was roughly similar for the 80-ms component in both hemispheres, this distribution was not as well defined in the left hemisphere, where neuronal populations were often found to respond to more than one frequency or over a broad range of frequencies. Neither Merzenich *et al.*[13] nor Reale and Imig,[16] however, found a clear tonotopic organization of the secondary cortex (AII) in their studies with nonhuman animals: several groups of neurons had multiple BFs or responded over a broad range of frequencies. Evoked responses from this area had longer latencies and were less well tuned than those from the primary cortex. [16,47] In their study, Merzenich *et al.*[14] reported that neurons in these secondary areas often responded over a broad range of frequencies and that the assignment of BFs was difficult, with two maxima often being observed. It is possible that a second harmonic distortion produced this 'double range' because the pairs of maxima in this study were often harmonically related.[48]

With respect to area 22, in the present study no tonotopic organization was observed in either hemisphere. This finding is consistent with data from animal studies using single-unit recordings[49] showing that neurons in nonprimary auditory areas in the macaque are not selectively responsive to pure tones, but respond only to frequency-centred sound bursts containing many frequencies.

One of the most significant findings in this study is the difference in the response selectivity of neurons in the right and left cerebral hemispheres. Neurons in the right auditory cortex were more sharply tuned to frequency than neurons in the homologous region of the left hemisphere. In a previous study, we showed that left auditory neurons were specifically sensitive to the temporal features of auditory information.[50]

As discussed earlier in this paper, the perceptual identity of a sound depends on both a sound's temporal structure as well as its spectral content. For example, periodic modulations of a sound's amplitude evoke a pitch percept. Taken together, the results from the numerous studies on the functional organization of the human auditory cortex suggest a different functional hemispheric asymmetry in the processing of spectral and temporal information. Frequency (i.e. spectral information) may be processed in HG of the right hemisphere, while temporal information may be processed in the left hemisphere. Not all data, however, is consistent with this idea. Using MEG, Langner *et al.*[51] found that periodicity pitch and frequency are arranged orthogonally in the human auditory cortex, with no differences observed according to hemisphere.

The asymmetry observed in the tonotopic organization of the right and left auditory cortices could underlie hemispheric specialization and could account for a fair amount of behavioural data obtained in studies on human auditory function. Robin *et al.*[52] found that left temporoparietal lesions impair the perception of temporal but not spectral auditory information; lesions in homologous regions of the right hemisphere were found to have the opposite effect. The lesions referred to in their study affected only the associative auditory areas, but such lesions could also disconnect the primary cortex from the rest of the brain.[53,54] Taken together with these findings, the present study's results suggest that the frequency analysis of sounds corresponds to the neural sharpening for neurons tuned to frequency. This mechanism could be conferred on the right auditory cortex, which plays an important role in the processing of pitch.[26,55] Auditory neurons in the left hemisphere, on the other hand, have been shown to be more selective for temporal pattern[50] than for spectral pattern. Such data are consistent with the results from studies with language-disordered subjects who have left-hemispheric lesions, suggesting that many of these patients are also impaired in auditory temporal perception.[56]

In conclusion, these results suggest that the processing of spectral information is carried out primarily in the right hemisphere, and the processing of the temporal aspects of auditory stimuli in the left auditory cortex.[50] Future research in this direction could not only provide insight into the cortical processes underlying speech and music perception, but may also help us better understand the psychoacoustic sequelae of cortical lesions.

## Acknowledgements

The authors are grateful to Drs B. Devaux (Ste. Anne Hospital, Paris), A. Musolino (Clinica Neurochirurgica, Messina, Italy), and J.M. Scarabin (Pontchaillou Hospital, Rennes) for their collaboration.

## References

1. **Whitfield, I. C.** (1970) Central nervous processing in relation to spatio-temporal discrimination of auditory patterns. In R. Plomp and Smoorenburg (eds) *Frequency Analysis and Periodicity Detection in Hearing.* Leiden: Sijthoff, pp. 136–52.
2. **Shamma, S. A.** (1985) Speech processing in the auditory system II: lateral inhibition and the processing of speech evoked activity in the auditory nerve. *J. Acoust. Soc. Am.* 78, 1622–32.

3. **Shamma, S. A.** (1986) Encoding the acoustic spectrum in the spatio-temporal responses of the auditory nerve. In B. C. J. Moore and R. D. Patterson (eds) *Auditory Frequency Selectivity.* New York: Plenum, pp. 289–96.
4. **Langner, G.** (1997) Neural processing and representation of periodicity pitch. *Acta Otolaryngol.* (Stockh.) 532, 68–76.
5. **Langner, G.** (1998) Neuronal periodicity coding and pitch effects. In P. W. Poon and J. F. Brugge (eds) *Central Auditory Processing and Neural Modeling.* New York: Plenum, pp. 31–41.
6. **Palmer, A. R.** (1982) Encoding of rapid amplitude fluctuations by cochlear nerve fibers in the guinea pig. *Acta Otolaryngol.* (Stockh.) 236, 197–202.
7. **Schreiner, C. E. and G. Langner** (1984) Representation of periodicity information in the inferior colliculus of the cat. *Abstr. Soc. Neurosci.* 10, 395.
8. **Langner, G. and C. E. Schreiner** (1988) Periodicity coding in the inferior colliculus in the cat. I: neuronal mechanisms. *J. Neurophysiol.* 60, 1799–822.
9. **Goldstein, M. H., F. De Ribaupierre, and G. H. Yeni-Komoshian** (1971) Cortical coding of periodicity pitch. In M. B. Sach (ed.) *Physiology of the Auditory System.* Baltimore, MD: National Education Consultants, pp. 299–306.
10. **Ribaupierre, F. De, M. H. Goldstein, and G. H. Yeni-Komoshian** (1972) Cortical coding of repetitive acoustic pulses. *Brain Res.* 48, 205–25.
11. **Goldstein, M. H., M. Abeles, M. R. Daly, and J. McIntosh** (1970) Functional architecture in cat primary auditory cortex: tonotopic organization. *J. Neurophysiol.* 33, 188–97.
12. **Merzenich, M. M. and J. F. Brugge** (1973) Representation of the cochlear partition on the superior temporal plane of the macaque monkey. *Brain Res.* 50, 275–96.
13. **Merzenich, M. M., P. L. Knight, and G. L. Roth** (1975) Representation of cochlea within primary auditory cortex in the cat. *J. Neurophysiol.* 38, 231–49.
14. **Merzenich, M. M., J. H. Kaas, and G. L. Roth** (1976) Auditory cortex in the grey squirrel: tonotopic organization and architectonic fields. *J. Comp. Neurol.* 166, 387–402.
15. **Phillips, D. P. and D. R. F. Irvin** (1981) Responses of single neurons in physiologically defined primary auditory cortex (A1) of the cat: frequency tuning and responses to intensity. *J. Neurophysiol.* 45, 16–34.
16. **Reale, R. A. and T. J. Imig** (1980) Tonotopic organization in auditory cortex of the cat. *J. Comp. Neurol.* 291, 192–265.
17. **Phillips, D. P. and S. S. Orman** (1984) Responses of single neurons in posterior field of cat auditory cortex to tonal stimulation. *J. Neurophysiol.* 51, 147–63.
18. **Sally, S. L. and J. K. Kelly** (1988) Organization of auditory cortex in the albino rat: sound frequency. *J. Neurophysiol.* 59, 1627–38.
19. **Lauter, J. L., P. Herscovitch, *et al.*** (1975) Tonotopic organization in human auditory cortex revealed by positron emission tomography. *Hear. Res.* 20, 199–205.
20. **Romani, G. L., S. J. Williamson, and L. Kaufman** (1982) Tonotopic organization of the human auditory cortex. *Science* 216, 1339–40.
21. **Pantev, C., M. Hoke, K. Lehnertz, *et al.*** (1988) Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. *EEG Clin. Neurophysiol.* 69 160–70.
22. **Pantev, C., M. Hoke, K. Lehnertz, *et al.*** (1990) Identification of sources of brain neuronal activity with high spatiotemporal resolution through combination of neuromagnetic source localization (NMSL) and magnetic resonance imaging (MRI). *EEG Clin. Neurophysiol.* 75, 173–84.

23. **Pantev, C., O. Bertrand, C. Eulitz, *et al.*** (1995) Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *EEG Clin. Neurophysiol.* 94, 26–41.

24. **Cansino, S., S. J. Williamson,** and **D. Karron** (1994) Tonotopic organization of human auditory association cortex. *Brain Res.* 663, 38–50.

25. **Liegeois-Chauvel, C., A. Musolino, J. M. Badier, P. Marquis,** and **P. Chauvel** (1994) Evoked potentials recorded from the auditory cortex in human: evaluation and topography of the middle latency components. *EEG Clin. Neurophysiol.* 92, 204–14.

26. **Zatorre, R. J., A. C. Evans, *et al.*** (1992) Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256, 846–9.

27. **Bancaud, J., J. Talairach, *et al.*** (1965) *La stéréoélectroencephalographie dans l'epilepsie: informations neurophysiopathologiques apportées par l'investigation fonctionnelle stéréotaxique.* Paris: Masson.

28. **Bancaud, J., F. Brunet-Bourgin, P. Chauvel, *et al.*** (1994) Anatomical origin of *déjà vu* and vivid 'memories' in human temporal lobe epilepsy. *Brain* 117, 71–90.

29. **Szikla, G., G. Bouvier, T. Hori,** and **V. Petrov** (1977) Angiography of the human brain cortex. In *Atlas of Vascular Patterns and Stereotactic Cortical Localization.* Berlin: Springer Verlag.

30. **Musolino, A., J. Talairach, P. Tournoux,** and **O. Missir** (1988) Comparative study between stereotactic angiography and magnetic resonance imaging (MRI) data on spatial organization of sulci and convolutions in man. Boll. *Lega Italiana contro Epilessia* 62/63, 51–6.

31. **Musolino, A., P. Tournoux, O. Missir,** and **J. Talairach** (1990) Methodology of *in vivo* anatomical study and stereo-electroencephalographic exploration in brain surgery for epilepsy. *J. Neuroradiol.* 17, 67–102.

32. **Liegeois-Chauvel, C., A. Musolino,** and **P. Chauvel** (1991) Localization of primary auditory area in man. *Brain* 114, 139–53.

33. **Steinmetz, H., J. Rademacher, *et al.*** (1989) Cerebral asymmetry: Mr planimetry of the human planum temporale. *J. Comput. Assist. Tomogr.* 13, 996–1005.

34. **Remond, A.** (1961) Integrated and topological analysis of the EEG *Clin. Neurophysiol.* 20 (Suppl.), 64–7.

35. **Chauvel, P., P. Buser, J. M. Badier, *et al.*** (1987) La 'zone épileptogène' chez l'homme: représentation des évènements intercritiques par cartes spatio-temporelles. *Rev. Neurol.* (Paris) 143(5), 443–50.

36. **Badier, J. M.** (1991) Etude de la localisation des sources cérébrales d'activité paroxystique par cartographie. Thèse de Doctorat. Université de technologie de Compiègne. Bertrand *et al.*

37. **Braak, H.** (1978) *Architectonics of the Human Telencephalic Cortex.* Berlin: Springer Verlag.

38. **Galaburda, A. M.** and **F. Sanides** (1980) Cytoarchitectonic organization of the human auditory cortex. *J. Comp. Neurol.* 190, 597–610.

39. **Bertrand, O., F. Perrin,** and **J. Pernier** (1971) Evidence for a tonotopic organization of the auditory cortex observed with auditory evoked potentials. *Acta Otolaryngol.* (Stockh.) 491 (Suppl.), 116–23.

40. **Schreiner, C. E.** and **J. R. Mendelson** (1990) Functional topography of cat primary auditory cortex: distribution of integrated excitation. *J. Neurophysiol.* 64, 1442–59.

41. **Imig, T. J., M. A. Ruggero, L. M. Kitzes *et al.*** (1977) Organization of auditory cortex in the owl monkey. *J. Comp. Neurol.* 171, 111–28.

42. **Aitkin, L. M., M. M. Merzenich, D. R. F. Irvine, *et al.*** (1986) Frequency representation in auditory cortex of the common marmoset. *J. Comp. Neurol.* 252, 175–85.

43. **Imig, T. J.** and **R. A. Reale** (1980) Patterns of cortico-cortical connections related to tonotopic maps in cat auditory cortex. *J. Comp. Neurol.* 192, 293–332.

44. **Luethke, L. E., L. A. Krubitzer** and **J. H. Kaas** (1988) Cortical connections of electro-physiologically and architectonically defined subdivisions of auditory cortex in squirrels. *J. Comp. Neurol.* 268, 181–203.

45. **Reite, M., P. Teale, J. Zimmerman, *et al.*** (1988) Source location of a 50 Msec latency auditory evoked field component. *EEG Clin. Neurophysiol.* 70, 490–98.

46. **Makela, J. P., M. Hamalainen, R. Hari,** and **L. McEvoy** (1994) Whole-head mapping of middle-latency auditory evoked magnetic field. *EEG Clin. Neurophysiol.* 92, 414–22.

47. **Schreiner, C. E.** and **M. S. Cynader** (1984) Basic functional organization of second auditory cortical field (AII) of the cat. *J. Neurophysiol.* 51, 1284–305.

48. **Goldstein, M. H., J. L. Hall,** and **B. O. Butterfield** (1968) Single unit activity in the primary auditory cortex of unanesthetized cats. *J. Acoust. Soc. Am.* 43, 444–55.

49. **Rauschecker, J. P., B. Tian, *et al.*** (1997) Serial and parallel processing in rhesus monkey auditory cortex. *Science* 268, 111–14.

50. **Liegeois-Chauvel, C., J. De Graaf, V. Laguitton, *et al.*** (1999) Specialization of left auditory cortex for speech perception in man depends on temporal coding. *Cereb. Cortex* 9, 484–96.

51. **Langner, G., M. Sams, *et al.*** (1997) Frequency and periodicity are represented in orthogonal maps in the human auditory cortex evidence from magnetoencephalography. *J. Comp. Physiol.* 181 (6), 665–76.

52. **Robin, D. A., D. Tranel,** and **H. Damasio** (1990) Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. *Brain Lang.* 39, 539–55.

53. **Pandya, D. N.** and **H. G. J. M. Kuypers** (1969) Cortico-cortical connections in the rhesus monkey. *Brain Res.* 13, 13–36.

54. **Pandya, D. N.** and **B. Seltzer** (1982) Association areas of the cerebral cortex. *Trends Neurosci.* 5, 386–90.

55. **Zatorre, R. J.** (1988) Pitch perception of complex tones and human temporal-lobe function. *J. Acoust. Soc. Am.* 84, 566–72.

56. **Tallal, P.** and **P. Newcombe** (1978) Impairment of auditory perception and language comprehension in dysphasia. *Brain Lang.* 5, 137–52.

CHAPTER 11

# THE NEURAL PROCESSING OF COMPLEX SOUNDS

TIMOTHY D. GRIFFITHS

## Abstract

This chapter considers the temporal processing of complex sounds relevant to musical analysis. Functional imaging studies, using positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and magnetoencephalography (MEG), and the psychophysical assessment of patients with lesions allow two different approaches to this. Functional imaging allows the determination of structures *normally involved* in temporal analysis, while patient studies allow inference about the *necessary* structures for temporal analysis. Both approaches suggest a hierarchal organization in the brain corresponding to the processing of music. The features of individual notes are analyzed in the pathway up to and including the auditory cortices, while higher-order patterns formed by those features are analyzed by distributed networks in the temporal lobe and frontal lobes distinct from the auditory cortices.

**Keywords:** Neural processing; Complex sounds; Musical analysis; Functional imaging

## Introduction

This chapter might have been called 'The Neural Processing of Complex Sound *Features*'. *All* sound is processed by mechanisms for the analysis of simple acoustic features (intensity, frequency, onset), complex acoustic features (such as patterns of these simple features as a function of time), and semantic features (learned association of sound patterns and meanings). Here, I concentrate on the analysis of complex features likely to be important in musical analysis, below the level of semantic and affective processing. Such analysis is represented in the middle box in Figure 11.1.

For the analysis of music, analysis of temporal features is likely to be particularly important, although I would certainly not dismiss the role of spectral or spatial analysis. Temporal analysis can be considered at different levels that I will call fine temporal structure and higher-order temporal structure. Such a categorical distinction works well for music as a sound with a highly segmented structure. Fine temporal structure corresponds to temporal structure at the level of milliseconds or tens of milliseconds; this is a 'window' for the processing of the temporal regularity within individual notes that is relevant to complex pitch perception (see Refs 1 and 2 for discussion of mechanisms of pitch perception). Higher-order temporal structure corresponds to temporal structure at the level of patterns of the pitch, onset time, or duration of notes. This corresponds to a 'window' at

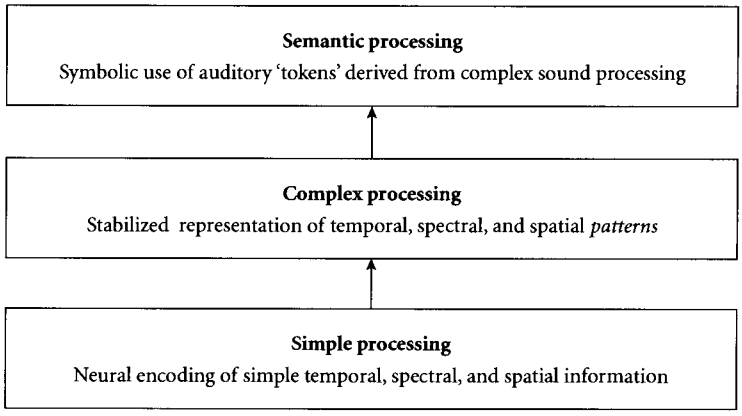



**Figure 11.1** Modular representation of complex sound perception relevant to music.

the level of hundreds of milliseconds or seconds. Higher-order structure might itself be subdivided into 'local patterns' (e.g. single intervals) and 'global patterns' (e.g. contour) as used by Peretz and others.[3,4] For present purposes, higher-order structure simply means any structure formed by patterns of the features in individual notes. This categorical distinction of temporal structure is convenient when considering processing relevant to music. This is not the only way of considering the temporal processing of sound. In particular, I would point out approaches based on the modulation transfer function. Here, responses to modulation of the sound amplitude or frequency are considered as *continuous* functions of modulation rate. This approach has been used in human psychophysical work,[5–8] evoked-potential work,[9,10] and functional imaging work.[11,12]

In the human auditory system much work on mechanisms of temporal processing, particularly in the case of music, has focused on the auditory cortex and cortical networks involving the auditory cortices. However, the ascending auditory pathway (Figure 11.2) affords an extensive mechanism for the processing of complex signals before the cortex is reached. Subcortical mechanisms for the processing of temporally complex sound in the form of continuous modulation have been investigated in animals[13–15] and in man.[11] Little work has been carried out on the processing of segmented sound closer to music in the ascending pathway, especially in man. Because of the extensive cross-connections in the auditory pathway, bilateral lesions of the ascending pathway are needed to produce temporal deficits. Such lesions are rarely compatible with life, but do occur, for example, in multiple sclerosis.[16] Imaging of human auditory brain stem processing is technically limited in functional imaging techniques, including PET, fMRI, and MEG. I point this out primarily to declare an open mind about the existence of ascending processing systems for sound sequences such as music. The concentration in this chapter on cortical mechanisms simply reflects a lack of data, rather than any a priori limitation on such processing in the ascending pathway. Another area of future interest is cortical 'molding' of auditory processing in the ascending pathway that has been demonstrated in other species such as the bat.[17]

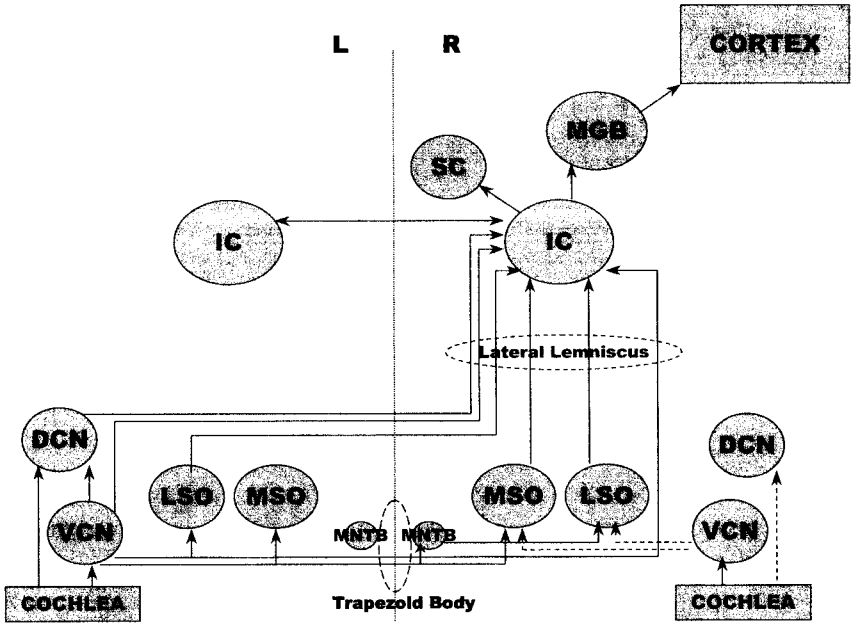



Figure 11.2 Simplified representation of ascending auditory pathway to the auditory cortex. This is based on animal and human studies (e.g. Refs 18 and 34) and demonstrates extensive processing of auditory information before the cortex is reached. For clarity, only the convergence of the input from both ears at the right-sided brain stem auditory nucleii is shown. A dorsal pathway, via the dorsal cochlear nucleus (DCN) and inferior colliculus (IC) is shown, which is involved in animals in the analysis of spectrally complex waveforms. The diagram also shows a ventral pathway, via the ventral cochlear nucleus (VCN), trapezoid body, and medial superior olive (MSO), thought to be important in animals for the analysis of spatial information. LSO = lateral superior olive, MNTB = medial nucleus of the trapezoid body, MGB = medial geniculate body, SC = superior colliculus.

The human auditory system[18] also contains a rich descending system that could mold behaviourally significant sounds such as music.

This chapter considers temporal processing in humans relevant to musical processing. I will consider functional imaging (using PET, fMRI, and MEG) and the psychophysical assessment of patients with lesions. Functional imaging allows the determination of structures *normally involved* in temporal analysis, and patient studies allow inference about the *necessary* structures for temporal analysis. For discussion of the extension of this approach above the level of acoustic pattern analysis, I refer the reader to Chapter 16 by Zatorre in this volume.

## Lesion studies

### Fine temporal structure

There have been a number of studies suggesting that the detection of sounds of different frequencies (as measured in the routine pure tone audiogram) depends on an intact

ascending auditory pathway to at least one of the auditory cortices. The rare condition of cortical deafness occurs following lesions to both superior temporal lobes (see Ref. 19 for review). The perception of pitch, as opposed to the detection of a pure tone at a given frequency, has been investigated using complex stimuli where the pitch is not immediately determined by the spectral characteristics of the sound. Zatorre[20] used a stimulus where the pitch is determined by the 'missing fundamental' of the spectral representation. The stimulus in this experiment, considered in the frequency domain, is a harmonic series where the lower harmonics including the fundamental stimulus are removed. The pitch of this stimulus is determined by the missing fundamental, against the idea that for a pitch to be heard there needs to be a representation of that pitch in the stimulus spectrum. The determination of pitch by any frequency domain model requires a form of template fitting;[21] the spectrum still contains harmonics that are *spaced* by an amount equal to the fundamental frequency. Alternatively, the pitch of this stimulus might be explained by temporal domain models; the waveform of this stimulus has a periodicity equal to the reciprocal of the fundamental, both before and after removal of the lower harmonics. Using the missing fundamental stimulus Zatorre investigated the perception of this complex pitch in subjects with well-characterized lesions affecting the superior temporal lobes. Subjects with right lesions involving the primary auditory cortex had deficits in perception of this pitch, but not subjects with right superior temporal lobe lesions sparing the primary auditory cortex, or subjects with lesions of the left primary auditory cortex. This suggests a particular involvement of the right primary auditory cortex in complex pitch perception.

Patterson, Yost, and others have developed stimuli containing regular intervals where the pitch properties are most parsimoniously explained by the fine temporal rather than the spectral characteristics of the sound.[22,23] Of course, both characteristics are interchangeable at the stimulus level, but these stimuli are manipulated in such a way that spectral template models are difficult to apply because of the filtering properties of the auditory system. Iterated rippled noise (IRN) is one such type of regular interval sound produced by taking bandpass noise and adding it to itself in a delay-and-add cycle. With increasing delay-and-add cycles (or gain in each cycle) the stimulus becomes more regular, as shown by the increasing first peak in the autocorrelation spectrum. In parallel with this temporal change, a pitch emerges from the noise with a signal-to-noise ratio that parallels the degree of temporal regularity of the auditory signal. In contrast, the spectrum of the auditory signal retains its bandpass properties. It is therefore easier to infer in the case of this signal (than for the missing fundamental) that changes in perceived properties are due to the fine temporal structure rather than the spectral structure in the auditory signal (see Ref. 2 for more detailed discussion). Using this stimulus I investigated one of Peretz's patients, I.R., with apperceptive amusia following bilateral temporal lobe lesions[24] (see Chapter 13 by Peretz, this volume, for further discussion of amusia). This patient has normal appreciation of pitch but is unable to perceive musical sequences. She has an intact pathway up to and including the right primary auditory cortex, but effective 'isolation' of that cortex from other areas of cortex that are lesioned. Testing this patient with IRN showed that she had a normal threshold for the detection of this stimulus presented to both ears. Testing each ear separately with contralateral masking noise showed that her normal performance was entirely accounted for by the performance at the left ear, the predominant input to the intact right auditory cortex.

Taken together these two studies are consistent with the pathway up to and including the primary auditory cortex being a necessary and sufficient substrate for the perception of complex pitch. I have discussed these studies under the heading fine temporal structure; both studies can be interpreted in terms of a need for primary cortex (particularly on the right) to perceive fine temporal structure, but the missing fundamental work is equally consistent with a spectral template model. These data would be consistent with activity of the human primary cortex being a neural correlate of the conscious perception of complex pitch.[25] Such processing is important for the perception of the pitch of the individual notes produced by most musical instruments.

### Higher-order temporal structure

Lesion studies can also allow inference about the detection of higher-order temporal patterns of segmented sounds, including music. A number of studies have looked at very simple tests of segmented sound processing in patients with auditory agnosias including amusia (see Ref. 19 for review). The tests used include gap detection and click-fusion threshold. Many studies have shown deficits, and been interpreted in terms of a temporal processing deficit, but conventional gap detection paradigms are confounded by the presence of spectral artifact in the stimulus (this is overcome in more refined paradigms such as those developed by Phillips[26]). Many of the patients in these studies had large bilateral lesions involving the superior temporal cortices, consistent with a role for the secondary and association areas of auditory cortex in this region being needed for the perception of this type of temporal pattern.

A deficiency in the evaluation of patients with auditory agnosia is the lack of a systematic battery of tests of sound sequence processing above the level of gap detection but below the level of actual musical sequences. Such a battery would have the potential to demonstrate 'generic' sound sequence mechanisms that might be relevant to a number of different sound processing disorders without higher-level confounds. A crude test of sequence discrimination in a patient, H.V., with amusia[27] suggested a loss of the perception of *rapid* but not slow sequences. This paralleled the phenomenology in a patient who had lost the ability to recognize and enjoy music with a high tempo. That patient had a unilateral lesion affecting the posterior superior temporal lobe and inferior parietal lobe, and the deficit in his case was less profound than in other cases reported following bilateral lesions.

Auditory agnosia can also be studied by a parallel approach based on the modulation transfer function as discussed earlier. In terms of the temporal 'window', the level of higher-order structure for patterns of notes (hundreds of milliseconds or seconds) parallels the level of low modulation rates (of tens of Hz or less). The patient with mild amusia discussed above (H.V.[27]) had a mild deficit in the detection of sinusoidal frequency modulation (FM) and amplitude modulation (AM) at a rate of 2 Hz. The dissociation between this and his profound deficit in the discrimination of rapid sequences of segmented sound was striking, and cautions against any universal approach in these patients based on modulation characterization. In contrast, the patient I.R., with a more marked amusia due to bilateral lesions, had a striking deficit in the detection of both AM and FM, particularly at low modulation rates (her threshold for 2-Hz sinusoidal AM was 11 times normal). Apart from demonstrating auditory deficits linked to the phenomenology, her case allows the inference

that a single intact pathway to the right auditory cortex is not a sufficient substrate for the normal perception of modulation. The presence of normal processing of IRN by I.R. provides an important control to allow the exclusion of other possible causes of the deficit. The IRN task demanded similar attention and working memory for simple acoustic features using a similar two-alternative forced-choice paradigm. The presence of such dissociated deficits is critical to these studies,[28] and studies showing a general deficit in performance in a range of tasks do not allow such interpretation in terms of distinct perceptual processes.

In terms of the psychophysical examination of higher-order temporal processing in subjects with lesions, these studies suggest that the deficits are produced by temporal lobe lesions that involve superior temporal lobe areas beyond the primary auditory cortex. Additionally, subject I.R. has a large right frontal lesion implicating this area in the perception of sound sequences and low-rate modulation. More studies of such patients using detailed modern structural imaging and systematic psychoacoustic testing will characterize the necessary structures for higher-order temporal processing. Such studies will never be supplanted by functional imaging, from which *sufficient* processing mechanisms cannot be immediately inferred.

## Functional imaging

### Fine temporal structure

The missing fundamental stimulus has been used in MEG experiments[29,30] where a mapping of perceived pitch has been demonstrated. For reasons discussed above, this map cannot be interpreted unambiguously in terms of the temporal structure of the stimulus, as template models could also explain the perceived pitch. But the mapping of perceived complex pitch to auditory cortex, derived by whatever means, is convergent with the lesion data discussed above. In the study of Langner, distinct representations of the frequency and temporal domains were suggested, with different orientations in the superior temporal planum. This would allow extensive processing in auditory cortex of spectrotemporal properties such as pitch as well as spectral (timbre) and temporal (attack) instrument characteristics. These are all features of individual notes.

IRN has also been used in functional imaging experiments, with the aim of identifying brain areas where the brain activity varies as a function of the fine temporal regularity of the stimulus.[2] In an initial study with PET in collaboration with Patterson, such a relationship was demonstrated in both auditory cortices. On the right, a region of activation was demonstrated in the medial part of Heschl's gyrus, containing the primary auditory cortex. On the left, the activation was situated more laterally and may correspond to secondary auditory cortex; these asymmetries are of interest in view of the lesion data suggesting a particular involvement of the right primary auditory cortex in the processing of temporal regularity. The activity shown in the PET study was argued to correspond to neural activity occurring after temporal regularity is converted into a more stable code in the ascending pathway, a process called temporal integration. The study is critically dependent on this assumption, but not on the exact mechanism by which such temporal regularity is stabilized, the simplest possibility being based on autocorrelation as mentioned above.

Models for this temporal integration mechanism[31] allow it to occur at an early stage in the auditory pathway, perhaps as early as the cochlear nucleus, and ongoing work is being carried out with fMRI to further delimit it to brain stem structures.

The work using IRN therefore also supports the existence of a representation of temporal regularity in the auditory cortex, with closer control of the auditory spectrum than in the case of the missing fundamental.

### Higher-order temporal structure

In the IRN experiment,[2] higher-order structure was also investigated. The individual notes with pitches determined by temporal regularity were presented as two pitch patterns: as a regular pitch 'staircase' and as novel diatonic melodies. We hypothesized that areas involved in the analysis of sound sequences would show greater activation as a function of the pitch strength of the individual notes for the melody condition than for the staircase condition. Sound sequence analysis represents an emergent property of the detection of the pitch of individual sounds, and areas involved in sequence analysis will show a more marked dependence on pitch strength for more complex pitch sequences like the melody condition. This was demonstrated in four areas, distinct from the primary auditory cortex, in both anterior temporal lobes near the poles and in the posterior superior temporal lobes. These areas are shown in Figure 11.3. A scheme is hypothesized, consistent with the data, where the temporal properties of the individual notes are extracted in the region of the primary auditory cortex before higher-order temporal properties of the stimulus sequence are analysed in the distinct network shown in Figure 11.3.

A subsequent study used PET to demonstrate areas involved in the processing of simple sequences that are not musical.[32] Here, simple six-element atonal sequences were presented to nonmusical subjects in order to demonstrate areas involved in the analysis of both pitch sequences and sequences of notes of different duration. Pairs of sequences were presented to subjects where either the pitch sequence, the duration sequence, neither, or both were varied between the two. During any scan, subjects were required to detect either changes in pitch sequence or duration sequence. The results showed, again, a network of activation distinct from the primary auditory cortex (in addition to activation in the primary auditory cortex in the principal comparison between the sound conditions and rest). Bilateral activation was demonstrated in the posterior superior temporal planum in the region of the planum temporale, and in both frontal opercula, with greater spatial extent of the clusters of activity demonstrated on the right. Highly significant activation was also demonstrated in both lobes of the cerebellum. The activation during the pitch sequence and duration sequence tasks was strikingly similar, and direct comparison of the two conditions showed no significant differences in activation. This would be consistent with a common mechanism for the two tasks at some level, consistent with some work in the musical domain.

These studies using 'submusical' stimuli, therefore, suggest the existence of networks for processing of higher-order structure in both temporal lobes that are distinct from the primary auditory cortices. The planum temporale is consistently activated in a number of studies, with less consistency in activation anterior to the primary auditory cortex. Activation anterior to the auditory cortex occurs in the anterior temporal lobe in the IRN

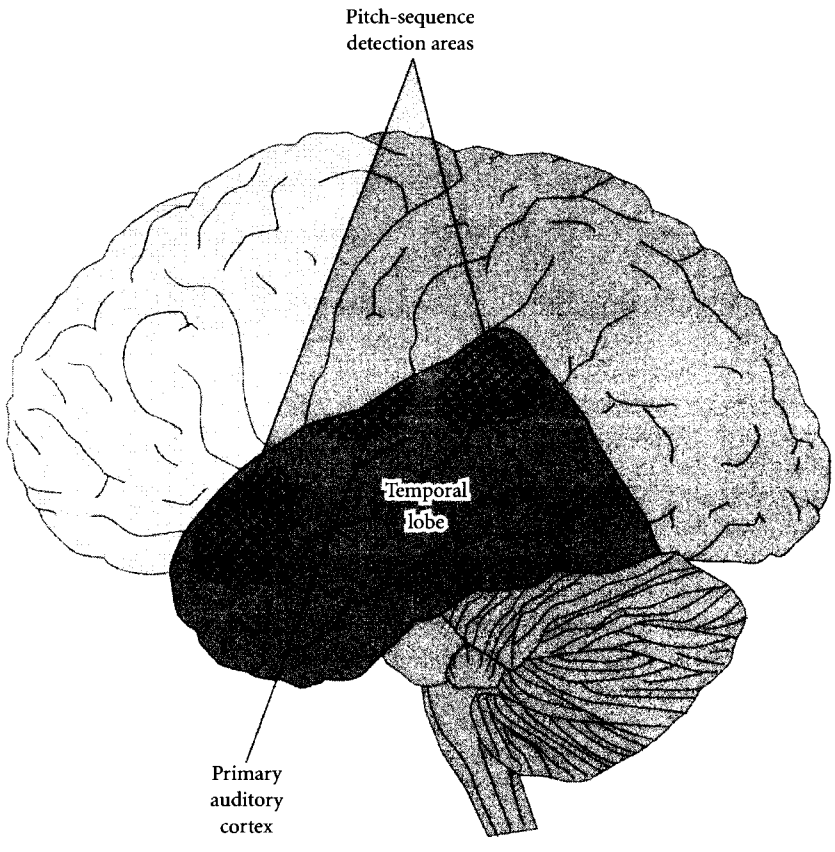



**Figure 11.3** Hypothesized basis for hierarchal processing of temporal structure in the temporal cortices suggested by the experiment using iterated rippled noise. The temporal structure within individual notes is processed in the pathway up to and including the primary auditory cortex, while pitch sequences are processed in the distinct areas shown. (Artist: © 1998, Bob Crimi. Reprinted with permission.)

study and frontal opercula in a number of other studies. The patterns of activation shown are similar to those demonstrated by studies using actual musical stimuli[33] (see also Chapter 16 by Zatorre, this volume), but show less striking lateralization. This suggests the possibility that the imposition of higher levels of cognitive processing in more musical studies contributes to the lateralization during musical processing.

## Conclusion

Both studies of humans with brain lesions and functional imaging provide convergent evidence for the existence of a neural substrate for the processing of sound sequences that is hierarchal in organization. The pathway up to and including the primary auditory cortices

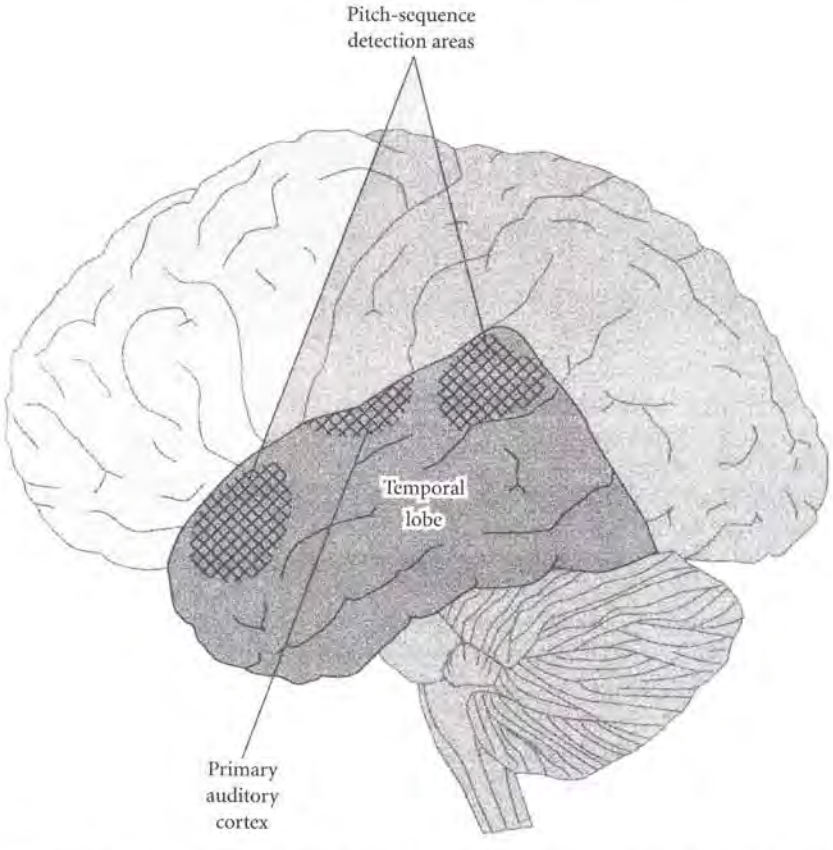



**Figure 11.3** Hypothesized basis for hierarchal processing of temporal structure in the temporal cortices suggested by the experiment using iterated rippled noise. The temporal structure within individual notes is processed in the pathway up to and including the primary auditory cortex, while pitch sequences are processed in the distinct areas shown. (Artist: © 1998, Bob Crimi. Reprinted with permission.)

study and frontal opercula in a number of other studies. The patterns of activation shown are similar to those demonstrated by studies using actual musical stimuli[33] (see also Chapter 16 by Zatorre, this volume), but show less striking lateralization. This suggests the possibility that the imposition of higher levels of cognitive processing in more musical studies contributes to the lateralization during musical processing.

## Conclusion

Both studies of humans with brain lesions and functional imaging provide convergent evidence for the existence of a neural substrate for the processing of sound sequences that is hierarchal in organization. The pathway up to and including the primary auditory cortices

may provide a sufficient mechanism for the processing of spectrotemporal features of individual sounds. Subsequent processing of higher-order temporal patterns formed by those individual sounds depends on processing in bilateral temperofrontal networks. These networks provide a first stage in the neural processing of music. Outstanding issues relate to the relationship between the elements of these networks and the perceptual attributes of the stimulus, the type of subject, and the type of task, to name but three.

## References


1. Hartmann, W. M. (1997) *Signals, Sound, and Sensation.* New York: American Institute of Physics Press.
2. Griffiths, T. D., C. Buechel, R. S. J. Frackowiak, *et al.* (1998) Analysis of temporal structure in sound by the human brain. *Nat. Neurosci.* 1, 421–7.
3. Peretz, I. (1990) Processing of local and global musical information by unilateral brain-damaged patients. *Brain* 113, 1185–205.
4. Liegeois-Chauvel, C., I. Peretz, M. Babai, *et al.* (1998) Contribution of different cortical areas in the temporal lobes to music processing. *Brain* 121, 1853–67.
5. Zwicker, E. (1952) Die Grenzen der Horbarkeit der Ampiltudenmodulation und der Frequenzmodulation eines Tones. *Acustica Akust. Beih.* 3, 125–33.
6. Kay, R. (1982) Hearing of modulation in sounds. *Physiol. Rev.* 62, 894–975.
7. Moore, B. C. J. and A. Sek (1995) Effects of carrier frequency, modulation rate, and modulation waveform on the detection of modulation and the discrimination of modulation type (amplitude modulation versus frequency modulation). *J. Acoust. Soc. Am.* 97, 2468–78.
8. Wojczak, M. and N. F. Viemeister (1999) *Adaptation Produced by Amplitude Modulation. Abstr. 59.* St. Petersburg, FL: Association for Research in Otolaryngology.
9. Rees, A., G. G. R. Green, and R. H. Kay (1986) Steady-state evoked responses to sinusoidally amplitude-modulated sounds recorded in man. *Hear. Res.* 23, 123–33.
10. Stefanatos, G. A., G. G. R. Green, and G. G. Ratcliff (1989) Neurophysiological evidence of auditory channel anomalies in developmental dysphasia. *Arch. Neurol.* 46, 871–5.
11. Harms, M. P., J. R. Melcher, and R. Weisskoff (1998) Time courses of fMRI signals in the inferior colliculus, medial geniculate body, and auditory cortex show different dependencies on noise burst rate [abstract]. *Neuroimage* 7, S365.
12. Giraud, A. L., C. Lorenzi, J. Wable, *et al.* (1999) Temporal envelope representation in the human auditory cortex [abstract]. *Neuroimage* 9, S787.
13. Moller, A. R. (1974) Response of units in the cochlear nucleus to sinusoidally amplitude modulated tones. *Exp. Neurol.* 45, 104–17.
14. Rees, A. and A. R. Moller (1983) Responses of neurons in the inferior colliculus of the rat to AM and FM tones. *Hear. Res.* 10, 301–30.
15. Langner, G. and C. E. Schreiner (1988) Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *J. Neurophysiol.* 60, 1799–822.
16. Quine, D. B., D. Regan, K. I. Beverly, *et al.* (1984) Patients with multiple sclerosis experience hearing loss for shifts of tone frequency. *Arch. Neurol.* 41, 506–7.
17. Yan, W. and N. Suga (1998) Corticofugal modulation of the midbrain frequency map in the bat auditory system. *Nat. Neurosci.* 1, 54–85.

18. **Moore, J. K.** (1994) The human brainstem auditory pathway. In R. K. Jackier and D. E. Brackmann (eds) *Neurotology*. St. Louis, MO: Mosby, pp. 1–17.
19. **Griffiths, T. D., A. Rees,** and **G. G. R. Green** (1999) Disorders of human complex sound processing. *Neurocase* 5, 365–78.
20. **Zatorre, R.** (1988) Pitch perception of complex tones and human cerebral lobe function. *J. Acoust. Soc. Am.* 84, 566–72.
21. **Cohen, M. A., S. Grossberg,** and **L. Wyse** (1995) A spectral network model of pitch perception. *J. Acoust. Soc. Am.* 98, 862–79.
22. **Patterson, R. D., S. Handel, W. A. Yost, *et al.*** (1996) The relative strength of the tone and the noise components in iterated rippled noise. *J. Acoust. Soc. Am.* 100, 3286–94.
23. **Yost, W. A., R. Patterson,** and **S. Sheft** (1996) A time domain description for the pitch strength of iterated rippled noise. *J. Acoust. Soc. Am.* 99, 1066–78.
24. **Griffiths, T. D., V. Penhune, I. Peretz, *et al.*** (2000) Frontal processing and auditory perception. *NeuroReport* 11, 919–22.
25. **Frith, C., R. Perry,** and **E. Lumer** (1999) The neural correlates of conscious experience: an experimental framework. *Trends Cognit. Sci.* 3, 105–14.
26. **Phillips, D. P., T. L. Taylor, S. E. Hall, *et al.*** (1997) Detection of silent intervals between noises activating different perceptual channels: some properties of 'central' auditory gap detection. *J. Acoust. Soc. Am.* 101, 3694–705.
27. **Griffiths, T. D., A. Rees, C. Witton, *et al.*** (1997) Spatial and temporal auditory processing deficits following right hemisphere infarction: a psychophysical study. *Brain* 120, 785–94.
28. **Ellis, A.** and **A. W. Young** (1998) *Human Cognitive Neuropsychology*. Hove, Sussex, UK: Lawrence Erlbaum Associates.
29. **Pantev, C., M. Hoke, B. Lutkenhoner, *et al.*** (1989) Tonotopic organisation of the auditory cortex: pitch versus frequency representation. *Science* 242, 486–8.
30. **Langner, G., M. Sams, P. Heil, *et al.*** (1997) Frequency and periodicity are represented in orthogonal maps in the human auditory cortex: evidence from magnetoencephalography. *J. Comp. Physiol.* 181, 665–76.
31. **Patterson, R. D., M. H. Allerhand,** and **C. Giguerre** (1995) Time-domain modeling of peripheral auditory processing: a modular architecture and a software platform. *J. Acoust. Soc. Am.* 98, 1890–4.
32. **Griffiths, T. D., I. Johnsrude, J. L. Dean, *et al.*** (1999) A common neural substrate for the analysis of pitch and duration pattern in segmented sound? *Neuroreport* 18, 3825–30.
33. **Zatorre, R. J., A. C. Evans,** and **E. Meyer** (1994) Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 14, 1908–19.
34. **Irvine, D. R. F.** (1986) *The Auditory Brainstem. A Review of the Structure and Function of the Auditory Brainstem Processing Mechanisms*. Berlin: Springer-Verlag.

PART IV

# MUSICAL BRAIN SUBSTRATES

CHAPTER 12

# MUSIC AND THE NEUROLOGIST: A HISTORICAL PERSPECTIVE

JOHN C. M. BRUST

## Abstract

Neurological disorders affecting musical function can produce either positive or negative symptoms. Positive phenomena include musicogenic epilepsy (seizures triggered by music), musical partial seizures (hallucinated music as the expression of the seizure), musical release hallucinations (nonepileptic musical hallucinations, usually associated with impaired hearing), and synesthesia (hallucinated colors triggered by musical tones). Negative phenomena comprise the amusias, which can be receptive, expressive, or both, and can selectively involve particular components of musical processing, including pitch, interval, contour, rhythm, meter, timbre, and emotional response. Amusia is often accompanied by aphasia, but each can occur in the absence of the other. Neurological disorders provide evidence that musical processing is multimodal and widely distributed in both cerebral hemispheres.

**Keywords:** Musicogenic seizures; Musical partial seizures; Release hallucinations; Amusia; Aphasia

## Introduction

In nearly all right-handed people the left cerebral hemisphere is 'dominant' for propositional or analytic processing (exemplified by language), whereas the right hemisphere is more responsible for appositional or gestalt processing (exemplified by spatial manipulation and aspects of emotional expression or response). How the brain processes music has intrigued neurologists and psychologists for over a century. A straightforward answer would not be expected, for the components of music—for example, pitch, timbre, duration, loudness, and rhythm—are likely processed through separate circuits. Music can be both linear (e.g. melody) and nonlinear (e.g. chords), and different kinds of music make different intellectual and emotional demands on different kinds of listeners or performers.

The biological survival value of music remains an enigma. Darwin believed that musical 'calling systems' evolved into speech,[1] whereas Herbert Spencer believed that music evolved as a stylized form of speech.[2] Steven Pinker notes, 'Of [all] mental faculties... music... shows

the clearest signs of not being [adaptive]';[3] yet in all cultures most people enjoy music. On the other hand, it takes severe mental retardation to prevent language acquisition, whereas musical aptitude ranges from tone deafness to Mozartian genius independently of other measures of intelligence.[4] Estimates of absolute pitch in the general population range from 1 in 1500 to 1 in 10,000.[5]

Clues as to how the brain processes music come from clinical case reports of patients with deranged musical function, from psychological testing of either normal or brain-damaged subjects, and, more recently, from studies using functional imaging (positron emission tomography [PET] and functional magnetic resonance imaging [fMRI]). Neurological symptoms can be positive or negative. Positive phenomena include musicogenic epilepsy, musical partial seizures, musical release hallucinations, and synesthesia. Negative phenomena include the amusias.

## Positive phenomena

Musicogenic seizures are triggered by hearing music. Attacks typically begin with subjective distress, agitation, tachycardia, and tachypnea and may or may not generalize to grand mal. Victims are often musically talented but do not have to be. The stimulus can be quite specific—for example, classical vs popular music, or a particular piece. Reported triggers include the second movement of Beethoven's Symphony No. 5 and a sequence of church bells with a frequency range of less than one octave.[6,7] In one case seizures occurred only if the patient actively listened to the music. In another, seizures could be precipitated by background music of which the patient was barely aware. Seizures have been triggered by the voices of particular singers.[8] A two-year-old boy's seizures were precipitated by his own singing.[9] In another case seizures were triggered by playing 'Now Thank We All Our God' on the organ but not by hearing it, singing it, silently reading it, or silently playing the notes on the keyboard.[10] Lesions causing musicogenic epilepsy are usually in the temporal lobe, right more often than left.[11]

Musical partial seizures consist of hallucinated music, which can be of any type but tends to be stereotypic for a particular epileptic subject. Such patients are less likely than those with musicogenic epilepsy to be musically talented. In a reported series of 666 temporal lobe epileptics, 16 per cent had auditory hallucinations, the great majority of which were unformed (e.g. banging, ticking).[12] Formed hallucinations could be verbal, musical, or both. In a more detailed description of hallucinatory seizures precipitated by cortical electrical stimulation, subjective phenomena included a male chorus and orchestra singing 'White Christmas', a radio jingle, organ music, and an orchestra playing in a music hall, which was simultaneously visually hallucinated.[13] Seizures most predictably followed stimulation of the superior temporal gyrus, right more often than left. Interestingly, spoken words also followed either left or right hemispheric stimulation, and there was no topographical separation between areas precipitating voices and those precipitating music.

Visual or auditory hallucinations associated with impaired vision or deafness ('Charles Bonnet syndrome', release hallucinations) can consist of music; they most often affect elderly people with presbycusis. Neither dementia nor psychiatric disease need be present,

and the hallucinations are often pleasurable rather than frightening or threatening. Some subjects sing along with the music, and some report that they can switch to another tune by consciously thinking about it. In some cases the pleasurable aspects of the music deteriorate over time, becoming distorted or too loud, or degenerating into tinnitus.[14,15] The pathophysiological basis of such hallucinations is uncertain; plausibly, reduced sensory input results in disinhibition of perception-bearing circuits, and perceptual traces are 'released'.

Synesthesia is an involuntary perception produced by stimulation of another sense. Thus, sounds produce the perception of colors. Geometric patters can occur, but more elaborate formed hallucinations do not. In some cases particular pitches produce the same colour; for example, C produced red, D green, and E blue in a boy with poor vision and musical talent (but not absolute pitch).[16] In addition to deafferentation, synesthesia is associated with hallucinogenic drug use (e.g. LSD) and, in one case, with a brain stem neoplasm. Many cases are idiopathic and noted in early childhood; some are familial. Women and left-handers are overrepresented, as are people who have difficulties with spatial relations.[17] Alexander Scriabin and Nicholai Rimsky-Korsakoff are believed to have had coloured-hearing synesthesia. The disorder, if it can be called that, is an example of unusually strong cross-modal association, but the physiological basis is otherwise unclear.

## Negative phenomena: amusia and aphasia

Amusia is an acquired impairment of musical processing. As with aphasia—a comparable impairment of language—the problem can be expressive, receptive, or both, and the disorder cannot be explained by damage either to the articulatory apparatus or to primary receptor mechanisms. In other words, tongue paralysis is not expressive amusia, and deafness is not receptive amusia. Amusia and aphasia can occur together or each can occur in the absence of the other.

Since the nineteenth century, clinicians investigating the amusias have tended to adopt either a localizationist/anatomical or a holistic/psychological approach. Among the localizers, Henschen classified amusia as either motor or sensory and then identified 'centres' in the left cerebral hemisphere: singing resided in the pars triangularis of the third frontal convolution, musical reception in the temporal pole, note reading near the angular gyrus, and instrument playing at the foot of the second frontal convolution (with a separate centre for violinists).[18,19] Similarly, Kleist described motor and sensory amusia, each affecting either tone or melody. His anatomical localization included separate centers for singing and for whistling.[20]

Opposed to such rigid diagramming was Feuchtwanger, who considered music too complex to localize or even to restrict to one cerebral hemisphere.[21] Ustvedt noted the heterogeneity of amusic patients and the lack, in most reports, of premorbid baseline information or of standardized tests.[22] A review of such reports reveals that heterogeneity is indeed the rule and that the presence or absence of aphasia does not predict the type of amusia.

The first report of aphasia and relatively preserved musical ability was in 1745; the patient, who had right hemiparesis and speech limited to the word 'yes', could sing hymns if someone else sang along.[23] Similar cases include two children with nonfluent aphasia,

one of whom could sing with words and the other without words;[24] a right-handed man with largely expressive aphasia following left hemispherectomy who could sing familiar songs with some errors in articulation but with little hesitation;[25] and a woman with transcortical aphasia (her verbal expression was limited to a few recurrent phrases, such as 'Hi, Daddy', but she had compulsively preserved speech repetition with echolalia) who could sing both old and newly taught songs.[26] Of 24 patients with Broca's aphasia and severely restricted speech output, 21 could sing (6 excellently), some sang only without words, some sang with paraphasic errors, and some needed help to get started.[27] A musically naïve patient with transcortical aphasia had hyperprosodic speech, whistled much of the time (including responses to questions), and listened to music much more often than before his illness.[28]

A professional pianist with Wernicke's aphasia, including alexia and agraphia for language, continued to play publicly; he could play unfamiliar melodies sung to him; reproduce heard chords; recognize dynamics, instruments, and intentional errors; and read a score nearly correctly.[29] A music professor with aphasia that included impaired speech comprehension and severe agraphia continued to write music.[30] The composer Shebalin developed Wernicke's aphasia with alexia and agraphia yet continued composing pieces 'which other musicians considered to be up to standard', including songs, quartets, choruses, and a symphony.[31] A conductor with global aphasia and ideomotor apraxia continued to conduct his orchestra.[32] A blind organist and composer who had written and read music in Braille developed Wernicke's aphasia and could no longer read or write language in Braille; yet he continued to read and write music in Braille.[33]

Cases of amusia plus aphasia are similarly diverse. A patient with anomia heard music as dissonant and voices, including his own, as too high pitched; he could neither sing nor reproduce rhythmic patterns.[34] Following surgery for a right frontal glioma, an accordionist had left hemiparesis and mild aphasia. Thereafter he sang 'wrong' and could no longer play his instrument (although he held it properly); he could correctly sing notes sung to him, but hearing them played on his accordion, he sang them a fourth too high or too low, and he could not imitate rhythm unless it was bound to a melody.[35] A violinist with absolute pitch had a stroke with right hemiparesis and Wernicke's aphasia, after which he had difficulty recognizing intervals, pitches, and rhythmic patterns; imitating heard rhythm; singing a melody previously unknown to him; and writing music spontaneously or to dictation.[36]

A man with probable Broca's aphasia was unable to identify popular songs or to sing melodies, but he could sing single tones and sequences without rhythm; rhythm recognition and production were completely lost whether presented visually, aurally, or tactily.[37] A musician with primary progressive aphasia, unable any longer to read music, could improvise but did so repetitively and without completion; he could not write or copy music or reproduce rhythm, but he could reproduce fragments of melodies with a sense of meter.[38] A woman with computerized tomographic evidence of right perisylvian atrophy developed aprosodic speech and could not recognize the prosody (e.g. sad, glad, mad) of others; she could no longer sing even single notes or identify melodies, although she could tell which of two notes was higher, and she could read and write music.[39]

Maurice Ravel had progressive aphasia with alexia, agraphia, and ideomotor apraxia; 'musical thinking' was said to be intact—he recognized melodies and errors of rhythm and

pitch and continued to listen to music critically—but he could no longer write or dictate music and thus could not compose the pieces he heard 'in his head'.[40,41] Following excision of a meningioma beneath the left temporal lobe, a young voice student had transcortical aphasia, alexia, and agraphia; she also had musical alexia and agraphia, yet she could still identify pitch, interval, melody, harmony, timbre, rhythm, and meter, and in a singing recital that included works in five languages, there were no apparent musical errors (although there may have been occasional paraphasic errors in the foreign languages).[42]

Cases of amusia without language disturbance are also heterogeneous; not surprisingly, in right-handers the lesions most often involve the right cerebral hemisphere. Reported symptoms include difficulty recognizing sounds as musical,[43] loss of a sense of rhythm,[44] hearing musical sounds 'out of tune',[45] and hearing both voices and music as monotonal.[46] Following a right temporoparietal infarct a man could no longer play the organ, yet other musical capabilities such as singing and imitating rhythm were intact ('expressive instrumental amusia').[47] Following a right temporal lobe infarct a man lost appreciation of timbre, including nonmusical sounds and people's voices; music sounded 'low and dull' with a kind of echo overlapping the sounds, and although he could recognize a guitar, a trumpet, or a drum, he could not recognize a violin, a piano, or an organ.[48]

A composer with right frontoparietotemporal infarction lost his emotional response to music and could no longer compose meaningfully; he could still compose serial music, '... but I have never been awed by the musical content of this style of composition'.[49] A man with a right temporoparietal vascular malformation had 'difficulty understanding the nuances of words and the inflections of sentences' and in understanding music; the relationship between a soloist and an accompanist became blurred, and there was a loss of aesthetic pleasure, yet there was no auditory agnosia, and he correctly recognized pitch, melody, timbre, harmony, and rhythm.[50]

A piano teacher with bilateral posterior cerebral atrophy of uncertain cause developed object-use apraxia and could no longer write or play the piano with two hands; her playing and singing were melodically disorganized and discontinuous, with poor rhythm and meter, and she read music note for note, with loss of 'global processing of pitch sequences'.[38] Following surgical clipping of bilateral middle cerebral artery aneurysms, a woman could no longer recognize or hum well-known tunes, and although she could tell if piano tones were the same or different, she could not tell if melodies were the same or different, whether melodic contour was preserved or violated, or whether a melody ended on the tonic or a nontonic; rhythm recognition and production were intact, however.[51]

## Studies of brain-damaged populations

Studies involving groups of patients with either left or right cerebral hemispheric disruption have also been inconsistent. Subjects receiving right carotid artery amobarbital injection could not sing properly, with pitch more distorted than rhythm, but they recognized their errors and identified songs sung to them.[52] Patients who had undergone either left or right temporal lobectomy were studied using the Seashore Test of Musical Abilities, which measures discrimination of pitch, loudness, rhythm, duration, timbre, and tonal

memory; right but not left lobectomy resulted in impaired recognition of timbre and tonal memory and a lesser degree of impairment involving pitch and rhythm.[53] A different testing instrument, the Musical Aptitude Profile, which involves tonal imagery (melody and harmony), rhythm (tempo and meter), and 'musical sensitivity' (phrasing, balance, and style) was used to study subjects with no musical training who had undergone either left or right temporal lobectomy; right-sided but not left-sided surgery resulted in impairment involving rhythm and sensitivity, not tonal imagery.[54]

Patients with either left or right temporal lobectomy were presented with spoken words, tunes without words, and words and tunes combined as an unfamiliar song. Left lobectomy resulted in impaired recognition of words alone, and right lobectomy resulted in impaired recognition of tunes alone, but either left or right lobectomy resulted in impaired recognition of melody combined with words.[55]

Several studies have compared local and global aspects of musical information processing in brain-damaged subjects. (For melody, pitch intervals are local, whereas melodic contour is global; for temporal patterns, rhythm is local, whereas meter is global.) In one study of patients with either right or left cerebral stroke, right hemispheric damage produced difficulty recognizing changes in either interval or contour, whereas left hemispheric damage produced difficulty only with intervals; either right or left hemispheric damage produced difficulty with rhythm but not with meter.[56] By contrast, although a similar study of patients following right or left cerebral corticectomy found comparable abnormalities involving interval and contour, either right or left hemispheric damage produced difficulty with meter, not rhythm.[57] In another study of patients with cerebral strokes, of eight patients with right-sided damage, five had impaired musical processing; of 12 patients with left-sided damage, nine had impaired musical processing. Right hemispheric damage caused abnormal processing of both contour and interval and of both meter and rhythm, whereas left hemispheric damage caused much more variable abnormalities: contour and interval could be affected alone or in combination, and meter and rhythm could be affected alone or in combination.[58]

Normal subjects recognize pitch even when the fundamental is missing if overtones are retained. This ability is lost following right temporal lobe damage that includes Heschl's gyrus.[59] Another study assessed the ability to recognize timbre derived either from spectral envelopes or from onset-offset times, and it was found that right temporal lobectomy produced greater impairment in either ability than did left temporal lobectomy.[60]

Following right but not left temporal lobectomy, patients had difficulty reproducing auditory but not visual rhythm; lesions producing such impairment encroached on the primary auditory cortex.[61]

In normal subjects the harmonic context established by a prime chord influences judgments of target chord intonation; in patients who had undergone corpus callostomy, this normal interaction was preserved in the right cerebral hemisphere but not the left.[62]

## Studies of normal subjects

Paralleling clinical case reports and studies of brain-damaged subjects have been studies of musical processing in normal people. Space does not allow a detailed review of such

investigations, the findings of which have often been as inconsistent as those of clinical reports. Early studies used the technique of dichotic listening, in which the two ears simultaneously receive different auditory information. In one such investigation musically naïve subjects appeared to have left ear advantage—that is, right cerebral hemispheric dominance—listening to melodies, whereas musically sophisticated subjects had right ear advantage; the findings were considered to reflect right hemispheric gestalt and emotional processing on the part of naïve listeners, in contrast to left hemispheric detailed analysis on the part of the musicians.[63] Other investigators, however, found no differences between musically sophisticated and musically naïve subjects,[64,65] and in some reports differences were found that seemed to depend on inherent musical aptitude,[66] age,[67] sex,[68] or right-ear advantage for higher pitch.[69,70] In one study the right hemisphere was used more for melody recognition when the tones had a complex timbre with rich overtones, again perhaps reflecting a right hemispheric advantage in dealing with simultaneous or gestalt, rather than linear or analytic, information.[71] It should be noted, however, that many investigators consider dichotic listening a highly unreliable technique for defining hemispheric functional 'dominance'.[72]

In normal subjects, functional imaging—positron emission tomography (PET) and functional magnetic resonance imaging (fMRI)—allows different components of musical processing to be assessed in real time. Such studies demonstrated that pitch recognition selectively activates the right prefrontal cortex, whereas active retention of pitch requires interaction of right frontal and right temporal cortices.[73,74] In another study, musical sight-reading, unlike reading words, activated the left occipitoparietal cortex, perhaps reflecting the dependence of musical reading on spatial location, in contrast to the dependence of word reading on feature detection.[75] In another study subjects listened to the same piece but were instructed to attend to different aspects of the music—for example, melody recognition, pitch, rhythm, or how many instruments were playing; even though the auditory information did not change with repeated trials, brain activation patterns were sharply altered by such changes in selective attention.[76]

As with brain-damaged patients, studies of normal subjects involve particular aspects of musical experience—auditory (pitch, harmony, timbre, interval, contour), visual (notational score, which includes real words; notes designated by letter but represented spatially, with duration indicated by appearance; purely musical symbols, such as rests, fermata, clef, crescendo, staccato, and repeat; and display of both simultaneous and sequential sounds), motor (performance, dance); memory; and emotion. Musical experience, however, is more than the sum of these parts. The multimodal components of musical processing cannot be sharply localized to one part of the brain or even to one cerebral hemisphere. The psychological whole that emerges from their interaction is even more widely distributed.

## References

1. **Darwin, C.** (1871) *The Descent of Man, and Selection in Relation to Sex.* New York: Appleton.
2. **Spencer, H.** (1857/1951) The origin and function of music. In *Literary Style and Music.* New York: Philosophical Library.

3. **Pinker, S.** (1997) *How the Mind Works.* New York: W. W. Norton & Co.
4. **Gardner, H.** (1983) Musical Intelligence. In *Frames of Mind. The Theory of Multiple Intelligences.* New York: Basic Books, Inc. pp. 99–127.
5. **Profita, J.** and **T. G. Bidder** (1988) Perfect pitch. *Am. J. Med. Genet.* 29, 763–71.
6. **Critchley, M.** (1937) Musicogenic epilepsy. *Brain,* 60, 13–27.
7. **Poskanzer, C., E. Brawn,** and **H. Miller** (1962) Musicogenic epilepsy caused by only a discrete frequency band of church bells. *Brain* 85, 77–92.
8. **Brien, S. E.** and **T. J. Murray** (1984) Musicogenic epilepsy. *Can. Med. Assoc. J.* 131, 1255–8.
9. **Herskowitz, J., N. P. Rosman,** and **N. Geschwind** (1984) Seizures induced by singing and recitation. A unique form of reflex epilepsy in childhood. *Arch. Neurol.* 41, 1102–3.
10. **Wieser, H. G., H. Hungerbühler, A. M. Siegel,** and **A. Buck** (1997) Musicogenic epilepsy: review of the literature and case report with ictal single photon emission computed tomography. *Epilepsia* 38, 200–7.
11. **Sutherling, W. W., L. M. Hershman, J. Q. Miller, *et al.*** (1980) Seizures induced by playing music. *Neurol.* 30, 1001–4.
12. **Currie, S., K. W. G. Heathfield, R. A. Henson,** and **D. F. Scott** (1971) Clinical course and prognosis of temporal lobe epilepsy. A survey of 666 patients. *Brain* 94, 173–90.
13. **Penfield, W.** and **P. Perot** (1963) The brain's record of auditory and visual experience. *Brain* 86, 595–696.
14. **Ross, E. D., P. B. Jossman, B. Bell, *et al.*** (1975) Musical hallucinations in deafness. *J. Am. Med. Assoc.* 231, 620–1.
15. **Hammeke, T. A., M. P. McQuillen,** and **B. A. Cohen** (1983) Musical hallucinations associated with acquired deafness. *J. Neurol. Neurosurg. Psychiatry* 46, 570–2.
16. **Rizzo, M. and P. J. Eslinger** (1989) Colored hearing synesthesia: an investigation of neural factors. *Neurol.* 39, 781–4.
17. **Cytowic, R. E.** (1989) Synesthesia and mapping of subjective sensory dimensions. *Neurol.* 39, 849–50.
18. **Henschen, S. E.** (1925) Clinical and anatomical contributions on brain pathology. *Arch. Neurol. Psychiatry* 13, 226–49.
19. **Henschen, S. E.** (1926) On the function of the right hemisphere of the brain in relation to the left in speech, music, and calculation. *Brain* 49, 110–23.
20. **Kleist, K.** (1928). Gehirnpathologie und localisatorische Ergebnisse über Horstorungen, Gerauschtaubheiten und Amusien. *Monatsschr. Psychiatr. Neurol.* 68, 853–60.
21. **Feuchtwanger, E.** (1930) Amusie. Studien zur pathologischen Psychologie der akustischen Wahrnehmung und Vorstellung undihrer Strukturgebiete, besonders in Musik und Sprache. Berlin: Julius Springer.
22. **Ustvedt, H. J.** (1937) Ueber die unersuchung der musikalischen funktionen bei patienten mit gehirnleiden, besonders bei patienten mit aphasie. *Acta Med. Scand.* 86 (Suppl.), 1–737.
23. **Dalin, O.** (1745) Berättelse om en dumbe, som kan siumga. *Kung Sven Vetensk Acad Handlinger* 6, 114–15. Cited in Benton, A. K. and R. J. Joynt. (1960) Early descriptions of aphasia. *Arch. Neurol.* 3, 205–22.
24. **Jackson, J. H.** (1871) Singing by speechless (aphasic) children. *Lancet* 2, 430–1.
25. **Smith, A.** (1966) Speech and other functions after left (dominant) hemispherectomy. *J. Neurol. Neurosurg. Psychiatry* 29, 467–71.

26. Geschwind, N., F. A. Quadfasel, and J. M. Segarra (1968) Isolation of the speech area. *Neuropsychologia* 6, 327–40.
27. Yamadori, A., Y. Osumi, S. Masuhara, and M. Okubo (1977) Preservation of singing in Broca's aphasia. *J. Neurol. Neurosurg. Psychiatry* 40, 221–4.
28. Jacome, D. E. (1984) Aphasia with elation, hypermusia, musicophilia, and compulsive whistling. *J. Neurol. Neurosurg. Psychiatry* 47, 308–10.
29. Assal, G. (1973) Aphasie de Wernicke sans amusie chez un pianiste. *Rev. Neurol.* 129, 251–5.
30. Assal, G. and J. Buffet (1983) Agraphie et conservation de l'écriture musicale chez un professeur de piano bilingue. *Rev. Neurol.* 139, 569–74.
31. Luria, A. R., L. S. Tsvetkova, and D. S. Futer (1965) Aphasia in a composer. *J. Neurol. Sci.* 2, 288–92.
32. Basso, A. and E. Capitani (1985) Spared musical abilities in a conductor with global aphasia and ideomotor apraxia. *J. Neurol. Neurosurg. Psychiatry* 48, 407–12.
33. Signoret, J. L., P. Van Eeckhout, M. Poncet, *et al.* (1987) Aphasie sans amusie chez un organiste aveugle. *Rev. Neurol.* 143, 172–81.
34. Pötzl, O. and H. Uiberall (1937) Zur Patholgie der Amusie. *Wien Klin. Wochenschr.* 50, 770–5.
35. Botez, M. and N. Wertheim (1959) Expressive aphasia and amusia following right frontal lesion in a right-handed man. *Brain* 82, 186–202.
36. Wertheim, N. and M. Botez (1961) Receptive amusia: a clinical analysis. *Brain* 84, 19–30.
37. Mavlov, L. (1980) Amusia due to rhythm agnosia in a musician with left hemispheric damage: a non-auditory supramodal defect. *Cortex* 16, 331–8.
38. Polk, M. and A. Kertesz (1993) Music and language in degenerative disease of the brain. *Brain Cogn.* 22, 98–117.
39. Confavreux, C., B. Croisile, P. Garassus, *et al.* (1992) Progressive amusia and aprosody. *Arch. Neurol.* 49, 971–6.
40. Alajouanine, T. (1948) Aphasia and artistic realization. *Brain* 71, 229–41.
41. Henson, R. A. (1988) Maurice Ravel's illness: a tragedy of lost creativity. *Br. Med. J.* 296, 1585–8.
42. Brust, J. C. M. (1980) Music and language. Musical alexia and agraphia. *Brain* 103, 367–92.
43. Pittrich, H. (1956) Sensorische Amusie mit Paramusie nach rechtsseitger Temporalverletzung. Cited by Spreen, O., A. L. Benton and R. W. Fincham (1965) Auditory agnosia without aphasia. *Arch. Neurol.* 13, 84–92.
44. Pötzl, O. (1939) Zur Pathologie der Amusie. *Z. Neurol. Psychiatr.* 165, 187–95.
45. Jellineck, A. (1956) Amusia. On the phenomenology and investigation of central disorders of the musical functions. *Folia Phoniatr.* 8, 124–49.
46. Sidtis, J. J. and B. T. Volpe (1988) Selective loss of complex-pitch or speech discrimination after unilateral lesion. *Brain Lang.* 34, 235–45.
47. McFarland, H. R. and D. Fortin (1982) Amusia due to right temporoparietal infarct. *Arch. Neurol.* 39, 725–7.
48. Mazzucchi, A., C. Marchini, R. Budai, *et al.* (1982) A case of receptive amusia with prominent timbre perception defect. *J. Neurol. Neurosurg. Psychiatry* 45, 644–7.
49. Judd, T. L., A. Arslanian, L. Davidson, *et al.* (1979) A right hemisphere stroke in a composer. *Presented at International Neuropsychological Society.* February 2. New York.
50. Mazzoni, M., P. Moretti, L. Pardossi, *et al.* (1993) A case of musical imperception. *J. Neurol. Neurosurg. Psychiatry* 56, 322.

51. **Peretz, I.** and **R. Kolinsky** (1993) Boundaries of separability between melody and rhythm in music discrimination: a neuropsychological perspective. *Q. J. Exp. Psychol.* 46A, 301–25.

52. **Gordon, H. W.** and **J. E. Bogen** (1970) Hemispheric lateralization of singing after intracarotid sodium amylobarbitone. *J. Neurol. Neurosurg. Psychiatry* 37, 737–8.

53. **Milner, B.** (1962) Laterality effects in audition. In V. B. Mountcastle (ed.) *Interhemispheric Relations and Cerebral Dominance.* Baltimore: John Hopkins University Press. pp. 177–95.

54. **Kester, D. B., A. J. Saykin, M. R. Sperling, *et al.*** (1991) Acute effect of anterior temporal lobectomy on musical processing. *Neuropsychologia* 29, 703–8.

55. **Samson, S.** and **R. J. Zatorre** (1991) Recognition memory for text and melody of songs after unilateral temporal lobe lesions: evidence for dual encoding. *Q. J. Exp. Psychol: Learn Mem. Cogn.* 17, 793–804.

56. **Peretz, I.** (1990) Processing of local and global musical information by unilateral brain-damaged patients. *Brain* 113, 1185–205.

57. **Liégois-Chauvel, C., I. Peretz, M. Babai, *et al.*** (1998) Contribution of different cortical areas in the temporal lobes to music processing. *Brain* 121, 1853–67.

58. **Schuppert, M., T. F. Münte, B. M. Wieringa, *et al.*** (2000) Receptive amusia: evidence for cross-hemispheric neural networks underlying music processing strategies. *Brain* 123, 546–59.

59. **Zatorre, R. J.** (1988) Pitch perception of complex tones and human temporal lobe function. *J. Acoust. Soc. Am.* 84, 566–72.

60. **Samson, S.** and **R. J. Zatorre** (1994) Contribution of the right temporal lobe to musical timbre discrimination. *Neuropsychologia* 32, 231–40.

61. **Penhune, V. B., R. J. Zatorre,** and **W. H. Feindel** (1999) The role of auditory cortex in retention of rhythmic patterns as studied in patients with temporal lobe removals including Heschl's gyrus. *Neuropsychologia* 37, 315–31.

62. **Tramo, M. J.** and **J. J. Bharucha** (1991) Musical priming by the right hemisphere post-callosectomy. *Neuropsychologia* 29, 313–25.

63. **Bever, T. G.** and **R. J. Chiarello** (1974) Cerebral dominance in musicians and non-musicians. *Science* 185, 537–9.

64. **Peretz, I.** and **J. Morais** (1980) Modes of processing melodies and ear asymmetry in non-musicians. *Neuropsychologia* 18, 477–89.

65. **Zatorre, R. J.** (1979) Recognition of dichotic melodies by musicians and non-musicians. *Neuropsychologia* 17, 607–17.

66. **Gaede, S. E., O. A. Parsons,** and **J. H. Bertera** (1978) Hemispheric differences in music perception: aptitude vs. experience. *Neuropsychologia* 16, 369–73.

67. **Wagner, M. T.** and **R. Hannon** (1981) Hemispheric asymmetries in faculty and student musicians and non-musicians during melody recognition tasks. *Brain Lang.* 13, 379–88.

68. **Shannon, B.** (1984) Asymmetries in musical aesthetic judgements. *Cortex* 20, 567–73.

69. **Deutsch, D.** (1974) An auditory illusion. *Nature* 251, 307–9.

70. **Deutsch, D.** (1992) Paradoxes of musical pitch. *Sci. Am.* 267, 88–95.

71. **Sidtis, J. J.** (1980) On the nature of cortical function underlying right hemisphere auditory perception. *Neuropsychologia* 18, 321–30.

72. **Sergent, J.** (1993) Mapping the musician brain. *Hum. Brain Map.* 1, 20–38.

73. **Zatorre, R. J., A. C. Evans, E. Meyer, *et al.*** (1992) Lateralization of phonetic and speech discrimination in speech processing. *Science* 256, 846–9.

74. **Zatorre, R. J., A. C. Evans,** and **E. Meyer** (1994) Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 14, 1908–19.

75. **Sergent, J., E. Zuck, S. Terriah, *et al.*** (1992) Distributed neural network underlying musical sight-reading and keyboard performance. *Science* 257, 106–9.

76. **Platel, H., C. Price, J.-C. Baron, *et al.*** (1997) The structural components of music perception. A functional anatomic study. *Brain* 120, 229–43.

CHAPTER 13

# BRAIN SPECIALIZATION FOR MUSIC: NEW EVIDENCE FROM CONGENITAL AMUSIA

ISABELLE PERETZ

## Abstract

Brain specialization for music refers to the possibility that the human brain is equipped with neural networks that are dedicated to the processing of music. Finding support for the existence of such music-specific networks suggests that music may have biological roots. Conversely, the discovery that music may have systematic associations with other cognitive domains or variable brain organization across individuals would support the view that music is a cultural product. Currently, the evidence favours the biological perspective. There are numerous behavioural indications that music-specific networks are isolable in the brain. These neuropsychological observations are briefly reviewed here with special emphasis on a new condition, that of congenital amusia (also commonly referred to as tone-deafness).

The notion that music might have biological foundations has only recently gained legitimacy. Over the past 30 years, music has mostly been studied as a cultural product. Musicologists were analysing each musical system in the context of its specific culture. Neuroscientists and psychologists were viewing music as a convenient window to the general functioning of the human brain. However, neuropsychological observations have consistently and recurrently suggested that music might well be distinct from other cognitive functions, in being subserved by specialized neural networks. As such, music might be viewed as pertaining more to biology than to culture. The goal of the present paper is to review the neuropsychological evidence that supports the biological perspective.

## Specialized neural networks for music processing

If music is biologically determined, then music is expected to have functional and neuroanatomical specialization. That is, music is expected to be subserved by neural networks that are dedicated to its processing, in being unresponsive or inadequate for dealing with nonmusical input. Presently, support for the existence of such specialized neural networks is compelling. Most evidence in this regard derives from the functional examination of individuals whose brain condition is disturbed in highly selective aspects. The brain anomaly may either impair or spare musical abilities exclusively. It can be revealed in three types of conditions: (1) Acquired disorders; (2) Congenital disorders; and (3) Brain stimulation.

### Acquired disorders

Acquired disorders refer to sequelae of a brain accident. Such injuries most frequently are subsequent to a cerebro-vascular accident at the adult age. In such cases, disorders are expected to reveal normal brain organization due to the fact that the disorders are constrained by the organization of the undamaged system, which was fully stabilized prior to the accident. Thus, the study of brain-damaged adults provides a unique opportunity to uncover the functional organization of the normal brain by a procedure akin to reverse engineering.

This research strategy is facilitated by the recurrent observation that brain damage does not affect cognition in its entirety, but rather, in particular aspects. Cognitive disorders can be highly selective. The selectivity of the disorder can take spectacular forms such as in brain-damaged composers who may lose their language and yet remain able to maintain their musical activities in their prior professional level.

> The most famous case is probably that of Shebalin, the Russian composer who, following successive vascular accidents occurring in his left hemisphere, suffered from severe disturbances of his language abilities. He remained aphasic for the rest of his life; he could neither understand nor speak intelligibly. Nevertheless, he continued to compose, notably completing his Fifth symphony, which Shostakovitch considered to be one of his most brilliant and innovative works.[1] Therefore, Shebalin displayed severe language deficits yet retained his musical skills to a remarkable degree.

This dissociation between language and music cannot simply be explained by the fact that these professional musicians were 'abnormally musical' from the start. Indeed, the reverse dissociation can also be observed, even in ordinary listeners. That is, persons devoid of any special talent, linguistic or musical, can experience spectacular losses of musical abilities, like losing the ability to recognize one's national anthem, without accompanying language difficulties.

> I have been fortunate to study in detail three such cases.[2,3] One of them, I.R., was a restaurant manager when she sustained bilateral brain damage as a consequence of successive surgeries in both sides of the brain for the clipping of ruptured aneurysms. Ten years post-onset, I.R. still suffers from severe and irreversible deficits in music perception and memory as a consequence of her brain damage.[3] Prior to her brain surgeries, music had great value to I.R. She was raised in a musically-inclined family since her only brother is a professional musician. Fortunately, I.R. did not lose her language skills. She understands speech perfectly and remains verbally fluent, being able to express herself quite effectively (see her poem recently published in French[3]).

Such musical disorders—called amusia in neurological terms—are not rare and are dissociable from disorders of language—that is, aphasia, as reviewed by Brust (Chapter 12, this volume). This recurrent finding suggests that most processing components that underlie language and music are not shared and are neuroanatomically separable.

Although dissociation between music and language has been reported in various spheres of activities, it remains that the two domains are rarely compared in analogous contexts. Hence, significant association between language and music might have escaped attention. It is important to examine music and language in similar tasks. One such simple task that

the auditory system performs constantly and without effort, in both music and language domains, is sound pattern recognition. Hence, recognition of tunes and lyrics provide a sound basis for cross-domain comparisons.

In comparing recognition of spoken lyrics to the recognition of instrumental tunes, it has been possible to show that these auditory functions are not performed by a single auditory system but by separate ones, each being specialized for its particular domain.[4,5] For instance, auditory recognition of music appears to recruit mechanisms that are not implicated in speech recognition or in environmental sound recognition. Indeed, one can find brain-damaged cases whose unique symptom is the loss of the ability to recognize and memorize music. The patients retain the ability to recognize and understand speech as well as to identify common environmental sounds normally.[2,3,6–8] The deficit can be remarkably selective. For example, C.N. was unable to recognize hummed melodies coming from familiar songs above chance. Yet, she could perfectly recognize the lyrics accompanying the melodies that she failed to recognize.[7] Moreover, C.N. was able to recognize the voice of speakers[2] and the intonation of speech.[9] The existence of such a specific problem with music alongside normal functioning of other kinds of auditory abilities, including speech comprehension, suggests damage to processing components that are not only specific to the musical domain but also essential to the normal process of music recognition. We refer to this condition, which is a particular form of amusia, as *music agnosia.*[7]

The reverse condition, that corresponds to a selective sparing of music recognition, has been reported (see Table 13.1) but not yet well documented. Nevertheless, in the available cases,[10–12] the lesion spared music processing relative to both speech comprehension and environmental sound recognition that were both severely disturbed. Such cases suggest isolated sparing of music recognition abilities, hence complementing the music-specific deficits described above. Thus, the current evidence, summarized in Table 13.1, is indicative of the

Table 13.1 Case reports of selective impairment and of sparing in the recognition of music

| Reports | Domain | | | | Lesions |
|---|---|---|---|---|---|
| | Music | Speech (not prosody) | Environmental sounds | Voices | |
| Peretz *et al.*,[2] C.N. and G.L. | − | + | + | + | Bilateral temporal lobes |
| | − | +* | + | + | Bilateral temporal lobes |
| Peretz *et al.*,[3] I.R. | − | + | + | + | Bilateral temporal lobes and right frontal lobe |
| Griffith *et al.*[6] | − | + | + | | Posterior right temporal lobe |
| Piccirili *et al.*[8] | − | + | + | + | Left superior temporal gyrus# |
| Laignel-Lavastine *et al.*[10] | + | − | − | | Right temporal lobe |
| Godefroy *et al.*[11]* | + | − | − | | Right posterior hemisphere |
| Mendez[12] | + | − | − | | Right temporal lobe# |

+ = normal recognition; − = impaired recognition.
* after or during recovery; # left-handed.

presence of specialized brain circuits for music recognition. These circuits are damaged in the cases of music agnosia and spared in the few cases who suffer from verbal agnosia coupled with agnosia for environmental sounds.

### Congenital disorders

Further neuropsychological evidence that is indicative of brain specialization for music comes from congenital disorders. These disorders refer to unexpected failures or achievements in musical abilities in comparison to the general level of intellectual and socio-emotional functioning. These deficiencies are termed congenital since their presence can be detected very early in development. One such well-known condition corresponds to the 'music-savant syndrome' that can often be observed in autistic individuals. The etiology of autism is not yet known; however, its incidence is relatively high, with one to two cases out of every 1000 births (about the same rate as the Down syndrome). Autism is currently viewed as deriving from some brain anomaly because of its frequent association with other brain defect (e.g. epilepsy), its genetic transmission and its atypical cerebral functioning as measured by brain imaging studies.[13] More interestingly from our perspective, 1–10 per cent of autistic individuals might be qualified as musicians.[14] In effect, autistic subjects are generally more apt in the area of music than in other domains, such as language.[15] Several even become *musical savants*, a term which refers to the observation of high achievements in musical activities in individuals who are otherwise socially and mentally handicapped.

> A well-known case (described in more details by L. Miller[14]) is that of "Blind Tom". Blind Tom was a young blind slave who gave piano concerts at the White House and all around the world. While his language repertoire consisted of less than 100 words, his musical repertoire contained more than 5000 musical pieces. This "music-savant" was sold in 1850 during a slavery sale in Georgia. Blind Tom was sold along with his mother to Colonel Bethune. Until age 5, he did not say a word and manifested no other sign of intelligence than his remarkable interest for the musical performance given by the colonel's daughters. At age 4, he was playing Mozart sonatas, which he had heard. At age 6, he was able to improvise, and at age 7, he gave his first recital. In 1862, despite the fact that he did not know how to read music, he was able to play without errors 14 pages of an original composition that he had heard just once. Blind Tom gave recitals until age 53 when following the colonel's death, he had to end his career.

The mirror image of this condition corresponds to individuals who are musically inept, despite normal exposure to music, normal intelligence, and social adaptation. Such individuals are sometimes called *tone-deaf* (see Ref. 16 for the first report).

> Che Guevara was known to be "tune-deaf".[17] He was well aware of his handicap as the following anecdote illustrates. At a party, by prior arrangement, Alberto, his best friend, was required to give a poke to the Che every time a tango was played. At some point during the party, the orchestra played an agitated Brazilian shora that had been Alberto's favorite. Alberto wished to share his enthusiasm with the Che. But the Che, with his eyes on a woman across the room, believed Alberto's nudge to be a tango signal and took to the floor, dancing a slow and passionate tango with everyone else jiggling to the shora. Realizing something was wrong, Che Guevara came over to ask Alberto for advice, who was too convulsed with laughter to be able to explain.

We refer to these rare individuals as *congenital amusics*. The term reflects better the likelihood that there are multiple forms of music developmental disorders, as there are various patterns of acquired amusia resulting from brain accident (see Chapter 12, this volume, for an illustration of the heterogeneity of acquired amusias).

Unlike other developmental disorders such as *dysphasia* and *dyslexia, congenital amusia* has not received much scientific attention. One obvious reason for this neglect is that attention is generally directed at learning disabilities that affect language because of their wide educational implications. Moreover, music educators are reluctant to suspect the presence of congenital disorders because such a diagnosis may mean discontinuation of musical studies.[18] The other reasons are less pragmatic. As mentioned at the outset, many scientists conceive music as the product of a general-purpose brain organization. In that context, amusia is not a developmental disorder that is expected to occur in isolation. Rather, amusia is expected to result from intellectual and/or socioemotional dysfunctioning. However, as *Music-savants* suggest, music proficiency does not seem to depend on the normal development of the cognitive and affective system. Musical proficiency can be achieved while sociocognitive functioning is globally deficient. Therefore, one can expect to observe the reverse condition. That is, we should be able to find congenital amusics in whom musical competence does not match the level of achievement reached in sociocognitive spheres.

We actively searched for such amusic cases by means of various advertisements over the last five years. The early discovery of a textbook case (called Monica, reported in Ref. 19, and represented in Figure 13.1) who closely matched the two case descriptions available in the literature (i.e. Grant-Allen's case[16] and Geshwind's case[20]) has greatly contributed to the advancement of the study. Since then, the presence of *amusia* has been confirmed in 21 adults and studied in detail in 11 of them.[21] All are self-declared 'musically impaired'. However, self-declaration does not suffice. Nonmusicians are prone to complain about their musical deficiencies, in general. Therefore, we selected only subjects who exhibited clear-cut performance deficits on our screening musical battery.[22–24] Furthermore, in order to exclude extraneous causalities, we carefully selected participants who had no psychiatry or neurological history and who possessed a solid level of education. To ensure adequate stimulation and to exclude lack of interest or of motivation, only volunteers who had experienced unsuccessful attempts at learning music during childhood were considered.

Up to now, we have found 21 individuals who exhibited a pattern of performance that unambiguously indicated the presence of a receptive musical disorder. Their results are represented in Figure 13.1 with two tests that allow comparisons of results on the melodic and temporal dimension because the tests are very close in structure. One test serves to evaluate the use of melodic cues (one altered note that modifies either pitch directions, key or pitch intervals) and the other to assess the use of rhythmic cues (one altered note that modifies rhythmic grouping by a change in duration values) in the discrimination of two successive short musical sequences as 'same or different'. The same set of novel but conventional musical phrases served in both types of tests.

Examination of the data in Figure 13.1 is informative in several aspects. First, most controls are confined to the right top quarter and do no overlap with amusics' performance. This confirms the usefulness of our screening battery to detect the presence of a music receptive disorder. Second, and more importantly, all amusic participants score below the normal range in the discrimination of musical stimuli that differ on the melodic dimension

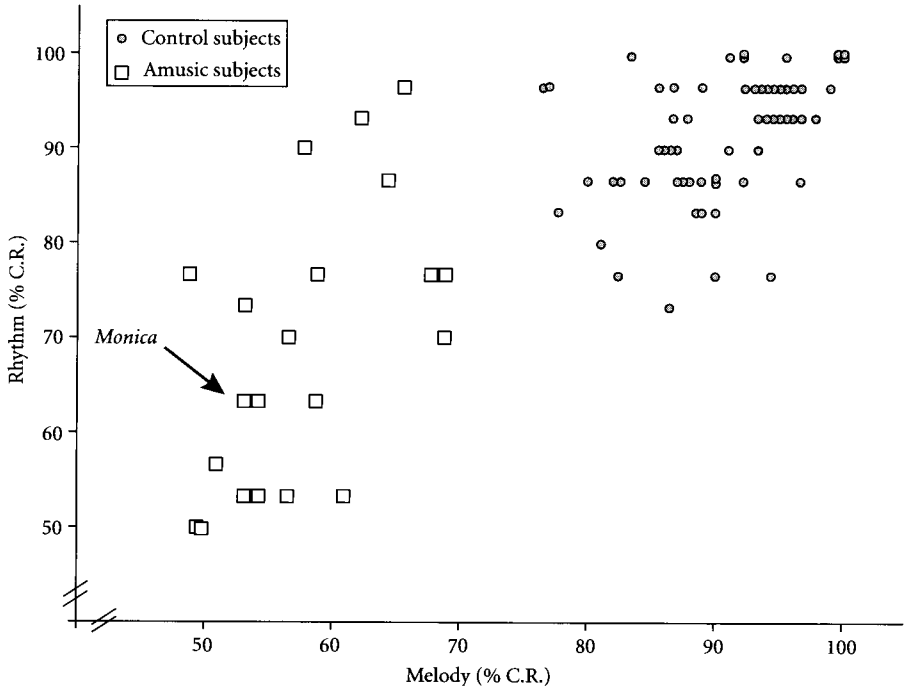



**Figure 13.1** Percentage of correct responses obtained by 21 congenital amusics relative to the normal distribution (constituted by 60 nonmusicians, age range: 14–74 years; education range: 7–20 years) in two series of tests of the montreal battery of evaluation of amusia[17–19] assessing melodic change discrimination (abscissa) and rhythmic change discrimination (ordinate axis). Chance level is at 50 per cent of correct responses.

while a majority of them succeed to discriminate the same stimuli when these differ in temporal structure. None of the amusic subjects is experiencing an isolated deficit on the rhythmic dimension. This difficulty in detecting pitch-related changes extends to dissonance for which amusics show little sensitivity.[21] The defect also extends to the detection of an anomalous pitch inserted in an otherwise conventional melody.[21] We propose that one likely origin for congenital amusia is related to a deficiency in musical pitch recognition.[19,21]

The deficit seems limited to music. Amusic subjects retain the ability to process nonmusical material as well as their matched controls. In one set of tests originally designed for music agnosic patients,[7] amusic participants, and their matched controls are presented with standard memory recognition tasks that only differ by the domain from which the test items are taken. In the music memory test, subjects are presented with 20 instrumental tunes (taken from familiar songs) to memorize. The melodies are then represented among 20 unstudied (but equally familiar) melodies that are randomly mixed. The subject is requested to indicate which melodies were heard in the study phase. For comparison purposes, in the lyrics and the environmental sound tests, subjects are given similar opportunities to learn and recognize 20 spoken lyrics (taken from the same familiar songs) and 20 environmental sounds (e.g. a barking dog), respectively. The three tests are performed in different sessions.

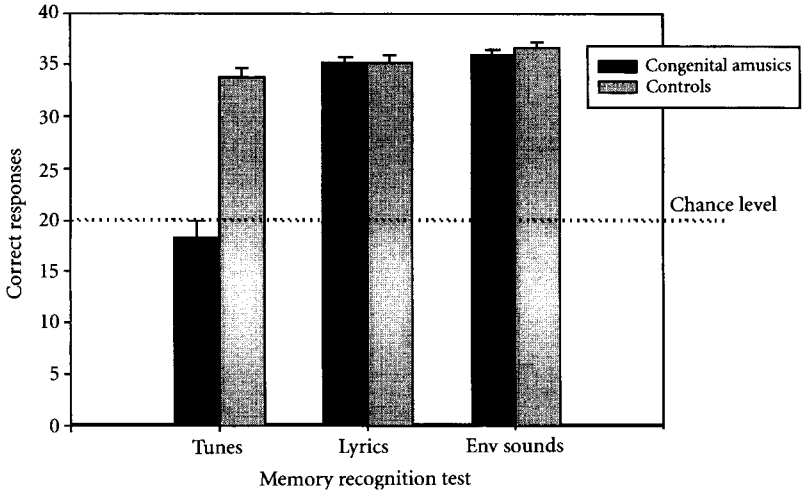



**Figure** 13.2 Number of correct responses obtained by 12 amusics and 20 matched controls in the memory recognition of studied melodies, lyrics, and environmental sounds. Error bars represent standard error from the mean. Chance level is at 20 correct responses.

As can be seen in Figure 13.2, amusics' performance is at chance in the music test whereas their scores are fairly high and lie within normal limits for nonmusical material, be they song lyrics or environmental sounds. The results clearly show the music specificity of the disorder.

In conclusion, congenital amusia appears as a new class of learning disabilities that are domain specific. Affected individuals appear to be born without the essential wiring elements for developing a normally functioning system for music, while achieving a high degree of proficiency in their professional and social life. This impairment seems to result, at least in some cases, from a more elemental problem that makes the individual unable to hear the pitch-relevant variations in music. Although it seems unlikely that a single defective mechanism, such as a pitch defect, is responsible for the proper operation of all processing components that subserve musical functions, it may be the case that a single essential defective component brings the development of the musical system to a halt. In this perspective, fine-grained pitch perception might be an essential component around which the musical system develops in a normal brain.[19]

The existence of congenital amusia coupled with the music-savant syndrome strongly suggests the presence of early pressures for the normal development of neural networks that are dedicated to music. Identification of these neural bases as well as of their associated heritability should provide crucial elements in the debate surrounding the existence of biological foundations for musicality.

### Brain stimulation

Investigations with epileptic patients provides the third source of evidence that speaks for the existence of neural networks that are dedicated to music. Indeed, in a few individuals,

music will be the exclusive trigger of the pathological firing of neurons conductive to the epileptic crisis. This form of epilepsy is called *musicogenic epilepsy*[25] (see also Brust, Chapter 12, this volume) and suggests that the epileptogenic tissue lies in a neural region that is tied to music processing.

> The British neurologist Macdonald Critchley[26] describes one regular visit of an epileptic patient who reports that hearing music exclusively provoked her attacks. The patient specified that music of the "popular" type had not effect, only "classical" music did, while confessing she had no particular preference for this kind of music. She was subsequently admitted to the hospital and, despite her discomfort, was presented with various kinds of music. The only record of classical music possessed by the experimenters, which contained "la valse des fleurs" de Tchaikovsy, immediately provoked a seizure with generalized convulsive movements, frothing of the lips and cyanosis. (Ref. 26, pp. 344–5).

During musicogenic epileptic seizures, abnormalities in electrical activity (recorded from the scalp) are generally observed at the temporal lobes, with a slight bias towards the right one (see Ref. 25 for a recent review). Thus, some processing component that is exclusively related to music must be located in those regions.

The direct application of electric stimulation to the brain of epileptic patients yields to a similar conclusion. *In vivo* electric stimulation of particular areas of the auditory associative cortex of awake patients may produce highly vivid *musical hallucinations.*[27] These provoked hallucinations suggest that the stimulation be applied to circuits that contain memories of musical experiences.

> Penfield & Perot report stimulating a particular region of the first right temporal circumvolution. The patient then says: "I hear music". The experimenters then repeat the stimulation, without telling the patient, who immediately says, "I hear the music again. It is like the radio". When asked what tune she was hearing, she said she did not know but it was familiar. The stimulation is again repeated and the patient shouts: "I hear it". The electrode was kept in place and the patient is asked to describe what she hears. The patient hummed the tune quite distinctly. The song comes out so clearly that one of the nurses recognizes "Rolling along together". The patient agreed that this sounded like the words in the song. (Ref. 27, case #5, p. 620)

The left, and slightly more often, right temporal regions are prone to evoke such musical experiences. The fact that musical memories can be exclusively elicited in individuals with no musical training underscores once again brain specialization for music.

In summary, the patient-based approach converges on one precise point: neuronal networks that are situated in or close to the superior temporal gyrus participate in music perception and memory in a decisive and exclusive manner. Presently, there is little evidence for music specificity coming from the study of normal brains (but see Chapters 15 and 18, this volume, for recent data).

## Localization of the music-specific networks

Precise localization of the music-specific networks is essential when biological determinism is at issue. Finding brain specialization for music is not sufficient. Brain specialization for

music may just result from the recruitment of a free neural space in the infant's brain. Music could modify that space to adjust it to its processing needs and hence be associated to neural specialization. This type of specialization does not require prewiring. It may occur as a response to early cultural pressures. In that case, a variable localization and distribution of the musical networks can be expected across individuals. Depending on the moment, quality, and quantity of exposure, various brain spaces might be mobilized. Thus, if music is a brain 'squatter', localization should vary capriciously across members of the same culture.

In contrast, biological functions are expected to be prewired. A prewired organization is expected to exhibit consistency in localization. For example, the primary auditory areas (located in the Heschl's gyri) are systematically buried in the sylvian fissure; this holds for all humans. Similarly, if brain specialization for music is pre-wired, then the music-specific networks are expected to have a relatively *fixed* arrangement. That is, brain implementation of music networks should be similar in the vast majority of humans, nonmusicians, and musicians alike. Moreover, this organization is not expected to vary as a function of the musical culture considered. Musical functions are expected to be similarly implemented in the brain of an isolated Pacific Islander, a Chinese opera singer, and a Western fan of rap music. This brain localization hypothesis can be easily tested with the new brain imagery techniques.

However, the demonstration of a similar brain organization for music in all humans remains elusive. Localization of the brain substrates underlying music has been an enduring problem for more than a century. In my view, the only consensus that has been reached today about the cerebral organization underlying music concerns pitch contour processing. The vast majority of studies point to the superior temporal gyrus and frontal regions on the right side of the brain as the responsible areas for processing pitch contour information.[28] However, it remains to be determined if this mechanism is music specific, since the intonation patterns of speech seem to recruit similarly located, if not identical, brain circuitries.[9,29]

Similarly, assuming comparable brain organization in musicians and nonmusicians may sound controversial, since there is strong evidence that musical training has a sizable effect on cortical morphology (see Chapter 24, this volume) and activity (see Chapters 25 and 26, this volume). It remains to determine to what extent these changes due to musical training are merely quantitative or rather represent qualitative modifications. In other words, it will be important to qualify the effect of training on brain organization. Presently, music training does not seem to distribute musical modules differently in the neural space of the expert compared to the ordinary listener, as a specialization by default would predict. Yet, the relevant evidence is still scarce.

Clearly, what is needed at the present stage is a grid that allows specification of the processing mechanisms that are essential for music appreciation. Once these essential ingredients have been identified, their respective localization may be tracked down in the brain of musicians and nonmusicians of different musical cultures. The research agenda involved is dense and will only be briefly sketched in section 'What is the content of the music-specific neural networks?'.

## What is the content of the music-specific neural networks?

The music-specific neural networks should correspond to a common core of musical abilities that is acquired by all normally developing individuals of the same culture. This common core should also be universal, in forming the essence of the musical competence acquired by members of all cultures. This universal competence can hopefully be reduced to a few essential processing components that represent the germ of brain specialization for music. In this perspective, there is no need for all musical abilities to have initial specialization. Brain specialization for a few mechanisms that are essential to the normal development of musical skills should suffice.

I am proposing that the two anchorage points of brain specialization for music are the encoding of pitch along musical scales and the ascribing of a regular beat to incoming events. The notion that a special device exists for tonal encoding of pitch has been developed in previous papers[30] and will thus not be elaborated further. Similarly, the notion that regularity might be fundamental to music appreciation is slowly emerging[31] (e.g. Chapter 2, this volume), although its specificity to music is rarely addressed.

Universality of musical scales and of pulse regularity is another issue that has received very little attention from ethnomusicologists. As mentioned earlier, ethnomusicologists cautiously avoid generalization across musical cultures. Yet, the plausibility of considering pitch scales and regularity as music universals has increased in recent years.[32]

In contrast, developmental psychologists have made significant progress in identifying plausible musical universals. Infants' studies have largely confirmed the presence of precocious sensitivity to musical scales and to temporal synchronicity in auditory processing. For example, 6- to 9-month-old infants exhibit enhanced discrimination in a consonant interval context compared to a dissonant one;[33] they also exhibit learning preferences for musical scales.[34] In most musical cultures, musical scales make use of unequal pitch steps. Infants already show a sensitivity bias towards musical scales, since they have been shown to be better at detecting a small pitch change in an unequal-step scale than in an equal-step scale (see Chapter 1, this volume). On the time dimension, infants prefer music that is subject to an isochronous temporal pulse. For instance, like adults, 4-month-old infants are biased towards perceiving regularity; they exhibit sensitivity to slight disruptions of temporal isochrony[31] (Chapter 2, this volume). All of these aspects of auditory pattern processing suggest the presence of innate learning preferences.

*In conclusion*, appreciation of music fits well with the product of a specialized cortical arrangement that is present and functional early in human development. Hence, music does not seem to be a game for the mind, for the neurons or for the senses. Music seems to serve needs that are so important to humans that their brain has dedicated some neural space to its processing. It remains to demonstrate that these music-specific networks are fulfilling needs that are not optional but have adaptive value.

## Acknowledgements

This chapter is based on studies supported by grants from the Canadian Institute of Health Research and the Canadian Natural Science and Engineering Research Council.

## References


1. **Luria, A., L. Tsvetkova, and J. Futer** (1965) Aphasia in a composer. *J. Neurolog. Sci.* 2, 288–92.
2. **Peretz, I., R. Kolinsky, M. Tramo, L. Labrecque, C. Hublet, and G. Demeurisse** (1994) Functional dissociations following bilateral lesions of auditory cortex. *Brain* 117, 1283–301.
3. **Peretz, I., S. Belleville, and S. Fontaine** (1997) Dissociations entre musique et langage après atteinte cérébrale: un nouveau cas d'amusie sans aphasie. *Revue Canadienne de Psychologie Expérimentale* 51, 354–67.
4. **Peretz, I.** (1993) Auditory agnosia: a functional analysis. In S. McAdams and E. Bigand (eds) *Thinking in Sound. The Cognitive Psychology of Human Audition.* New York, NY: Oxford University Press, pp. 199–230.
5. **Polster, M. and S. Rose** (1998) Disorders of auditory processing: evidence for modularity in audition. *Cortex* 34, 47–65.
6. **Griffiths, T., A. Rees, C. Witton, P. Cross, R. Shakir, and G. Green** (1997) Spatial and temporal auditory processing deficits following right hemisphere infarction: a psychophysical study. *Brain* 120, 785–94.
7. **Peretz, I.** (1996) Can we lose memories for music? The case of music agnosia in a non-musician. *J. Cognitive Neurosci.* 8, 481–96.
8. **Piccirilli, M., T. Sciarma, and S. Luzzi** (2000) Modularity of music: evidence from a case of pure amusia. *J. Neurol. Neurosurgery Psychiatry* 69, 541–5.
9. **Patel, A. D., I. Peretz, M. Tramo, and R. Labrecque** (1998) Processing prosodic and musical patterns: a neuropsychological investigation. *Brain Lang.* 61, 123–44.
10. **Godefroy, O., D. Leys, A. Furby, J. De Reuck, C. Daems, P. Rondepierre, B. Dabachy, J.-F. Deleume, and A. Desaulty** (1995) Psychoacoustical deficits related to bilateral subcortical hemorrhages: a case with apperceptive auditory agnosia. *Cortex* 31, 149–59.
11. **Laignel-Lavastine, M. and T. Alajouanine** (1921) Un cas d'agnosie auditive. *Revue Neurologique 37,* 194–8.
12. **Mendez, M.** (2001) Generalized auditory agnosia with spared music recognition in a left-hander. Analysis of a case with a right temporal stroke. *Cortex* 37, 139–50.
13. **Bailey, A., W. Phillips, and M. Rutter** (1996) Autism: towards an integration of clinical, genetic, neuropsychological, and neurobiological perspectives. *J. Child Psychol. Psychiatry* 37, 89–126.
14. **Miller, L.** (1989) *Musical Savants. Exceptional Skill in the Mentally Retarded.* Hillsdale, NJ: Erlbaum.
15. **Heaton, P., B. Hermelin, and L. Pring** (1998) Autism and pitch processing: a precursor for savant musical ability? *Music Percept.* 15, 291–305.
16. **Grant-Allen** (1878) Note-deafness. *Mind* 10, 157–67.
17. **Taibo II, P. I.** (1996) *Ernesto Guevara, también conocido como el Che.* Buenos Aires: Planeta.
18. **Kazez, D.** (1985) The myth of tone deafness. *Music Educators J.* 71, 46–7.
19. **Peretz, I., J. Ayotte, R. Zatorre, J. Mehler, P. Ahad, V. Penhune, and B. Jutras** (2002) Congenital amusia: a disorder of fine-grained pitch discrimination. *Neuron* 33 (13), 185–91.
20. **Geschwind, N.** (1984) The brain of a learning-disabled individual. *Annals of Dyslexia* 34, 319–27.
21. **Ayotte, I., I. Peretz, and K. Hyde** (2002) Congenital amusia: a group study of adults afflicted with a music-specific disorder. *Brain* 125, 1–14.

22. **Peretz, I.** (1990) Processing of local and global musical information by unilateral brain-damaged patients. *Brain* 11, 1185–205.
23. **Liégeois-Chauvel, C., I. Peretz, M. Babaï, V. Laguitton,** and **P. Chauvel** (1998) Contribution of different cortical areas in the temporal lobes to music processing. *Brain* 121, 1853–67.
24. **Ayotte, J., I. Peretz, I. Rousseau, C. Bard,** and **M. Bojanowski** (2000) Patterns of music agnosia associated with middle cerebral artery infarcts. *Brain* 123, 1926–38.
25. **Wieser, H. G., H. Hungerbühler, A. Siegel,** and **A. Buck** (1997) Musicogenic epilepsy: review of the literature and case report with ictal single photon emission computed tomography. *Epilepsia* 38, 200–7.
26. **Critchley, M.** (1977) Musicogenic epilepsy. In M. Critchley and M. Henson (eds) *Music and the Brain.* London: W. Heinemann, pp. 344–53.
27. **Penfield, W.** and **P. Perot** (1963) The brain's record of auditory and visual experience. *Brain* 86, 595–696.
28. **Peretz, I.** (2000) Music perception and recognition. In B. Rapp (ed.) *The Handbook of Cognitive Neuropsychology.* Hove: Psychology Press, pp.519–40.
29. **Zatorre, R. J., A. C. Evans, E. Meyer,** and **A. Gjedde** (1992) Lateralization of phonetic and pitch processing in speech perception. *Sci.* 256, 846–9.
30. **Peretz, I.** and **J. Morais** (1989) Music and modularity. *Contemporary Music Rev.*, 4, 277–91.
31. **Drake, C.** (1998) Psychological processes involved in the temporal organization of complex auditory sequences: universal and acquired processes. *Music Percept.* 16, 11–26.
32. **Arom, S.** (2000) Prolegomena to a biomusicology. In N. Wallin, B. Merker, and S. Brown (eds) *The Origins of Music.* Cambridge: MIT press, pp. 27–9.
33. **Schellenberg, E. G.** and **S. Trehub** (1996) Natural musical intervals: evidence from infants listeners. *Psycholog. Sci.* 7, 272–7.
34. **Trehub, S., G. Schellenberg,** and **S. Kamenetsky** (1999) Infants' and adults' perception of scale structure. *J. Exp. Psychol.: Hum. Percept. Perform.* 25, 965–75.

CHAPTER 14

# CEREBRAL SUBSTRATES FOR MUSICAL TEMPORAL PROCESSES

SÉVERINE SAMSON AND NATHALIE EHRLÉ

## Abstract

Music as well as language consists of a succession of auditory events in time, which require elaborate temporal processing. Although several lines of evidence suggest that the left dominant hemisphere is predominantly involved in the processing of rapid temporal changes of speech, very little is known about the cerebral substrates underlying such auditory temporal processes in music. To investigate this issue, we examined epileptic patients with either left (LTL) or right (RTL) temporal-lobe lesions as well as normal control subjects (NC) in two different tasks involving the processing of time-related (temporal) information. By manipulating inter-onset interval (IOI) in a psychophysical task, as well as in a task of detection of rhythmic changes in real tunes, we studied the processing of temporal microvariations in music. The first task assessed anisochrony (or irregularity) discrimination of sequential information according to different presentation rates (between 80 and 1000 ms IOI). For all subjects, an effect of tempo was obtained, thresholds were lower for the 80 ms IOI than for longer IOIs. Furthermore, there was a specific impairment of rapid anisochrony discrimination (80 ms IOI) for LTL patients as compared to RTL and NC subjects, but no deficit was observed for longer IOI. These findings suggest the specialization of LTL structures in processing rapid sequential auditory information. The second task involved the detection of inter-onset interval increments in familiar monodic tunes. Performance was measured for two increments (easy vs difficult to detect according to cognitive expectation) to assess the effect of cognitive expectation using a forced-choice paradigm (changed vs unchanged melody). The results showed that LTL patients but not RTL were impaired as compared to NC subjects in the increment detection. However, all groups showed differences between the two levels of difficulty, suggesting that top-down processing remains functional. These findings suggest that LTL structures are predominantly involved in perceiving time-related perturbations in familiar tunes as well as in isochronous sequences, extending to the musical domain findings previously reported in speech.

## Introduction

Music as well as language consists of a succession of auditory events in time, which require elaborate temporal processing. In the present study, we will consider the temporal coding of musical information which plays an important role in performing and perceiving rhythm. This time-related processing concerns a large range of time frames that may implicate different forms of information processing. In the neuropsychological domain, simple

auditory sequences and musical patterns have been used to investigate the cerebral structures underlying musical temporal processes. Studies reported in the literature will be reviewed first. Then, experimental findings that we obtained in two different studies will be reported to clarify the role of the left (LTL) as opposed to the right temporal lobe (RTL) structures in processing subtle temporal variations within a range of 10–100 ms.

### Temporal processing in simple auditory sequences

Several studies have explored time-related processing by using simple auditory sequences. In this domain, several paradigms classically used in experimental psychology have been adapted to neuropsychology. In a seminal paper, Efron[1] investigated the ability of brain-damaged patients to judge the temporal order of two-tones of different frequencies separated by a silent interval. The results showed that aphasic patients with left-hemisphere lesions required longer intervals (between 140 and 400 ms) to discriminate temporal order than nonaphasic patients with right-hemisphere lesion or normal subjects (75 ms interval). It was therefore suggested that the left hemisphere structures generally thought to be involved in language processing may also contribute to the temporal analysis of fast auditory sequential information. Subsequently, Tallal and Newcombe[2] provided convergent evidence by demonstrating that selective damage to the left but not to the right hemisphere disrupted the ability to process two-tones separated by a short interval (300 ms or less between the tones). Importantly, neither left- nor right-hemisphere damaged subjects were affected when longer intervals were used. This finding has been reproduced with similar paradigms and generalized to other tasks involving gap detection or perception of simulaneity and succession.[3–7] One case report of an amusic patient (HV) seems to contradict previous results since a deficit in temporal order processing of rapid patterns was observed in presence of a right-hemisphere lesion.[8] However, the authors noted that the cortical lesion of this patient was associated with an underlying bilateral white matter lesion which may have caused bilateral cortical deafferentation. Such a cerebral dysfunction, that presumably involved the left hemisphere, might have been responsible for the disruption of rapid temporal processing explaining therefore the apparent contradiction between the results.

However, inconsistent results were reported in studies investigating the perception of duration. Indeed, results of different studies have shown that impairments in this perceptual ability have been observed in patients with unilateral lesions implicating the right or the left hemisphere,[9] in a case of auditory agnosia following a bilateral cerebral dysfunction[10] as well as in patients with RTL lesions.[11] These findings suggest that the perception of continuous signals, as opposed to discrete events, depends on different processing involving distinct neural substrate.

Except for the perception of duration, the results previously reviewed suggest that left-hemisphere structures are predominantly involved in the processing of rapid sequential information. Recent electrophysiological results indicate that this function can be linked to auditory cortices of the LTL. By recording intracerebral evoked potentials to syllables in the right and the left human auditory cortices, Liégeois-Chauvel and her collaborators[12] demonstrated a specialization of the left auditory cortex for speech perception that depends on rapid temporal coding (within a few tens of milliseconds). If the left temporal cortex is dominant for fine grained time-related processing of language, it seems

plausible to hypothesize that it would also be important for such temporal processing in music.

### Temporal processing in musical sequences

Parallel to the previously reported studies using simple auditory sequences, several studies investigated temporal processes in musical sequences. It has been suggested that temporal processing, of which musical rhythm would be an example, is best performed by the functions of the dominant left hemisphere. However, the evidence reported in the literature provide little support for this hypothesis. Few studies investigating the perception of dichotically presented stimuli in normal listeners usually report a right ear advantage which is supposed to reflect left-hemisphere predominance for the perception of temporally complex nonspeech stimuli,[13–16] but this perceptual asymmetry has not always been obtained.[17] Similarly, studies carried out in brain damaged subjects have not systematically documented a deficit in rhythmic tasks in the presence of a left-hemisphere lesion. Although evidence supporting left-hemisphere involvement in rhythm have been reported in a few studies carried out in unilateral brain damaged subjects,[18,19] other studies have demonstrated the contribution of right-hemisphere structures and/or a bilateral cerebral involvement in rhythm discrimination.[20–25] Two studies have even failed to report deficits in rhythm discrimination after unilateral temporal lobe resection.[11,26] However, the use of musical stimuli for which familiarity and rhythmical complexity are extremely variable make the comparison between studies difficult. Based on these seemingly contradictory results obtained in lesion studies, it seems impossible to conclude that left-hemisphere structures are predominantly involved in musical rhythm.

One functional neuroimaging study investigated the temporal processing of musical patterns[27] but interpretation of the results in this study is made difficult because of methodological factors. Using Positron Emission Tomography in normal subjects, an increase of cerebral blood flow was obtained in left inferior Broca's area and in left insula when rhythmic and pitch judgement conditions were compared. Although this finding was interpreted by the authors as evidence supporting left-hemisphere superiority for temporal processing, it seems impossible to attribute these foci of activation to a purely temporal processing. Indeed, many other nontemporal aspects related to the task demands and to the stimuli, that differ between the two compared conditions, might have contributed to these metabolic changes.

In cognitive psychology, it has been proposed that the perception of rhythmic grouping, consisting of a sequential organization of relative durations of tones and silent intervals, should be differentiated from the perception of metre, referring to the underlying perceived beat marking off equal duration units.[28–31] Neuropsychological studies have demonstrated a dissociation between these two components involved in subjective organization of temporal patterns,[23,26,32] but this dissociation has not been systematically observed.[20] Moreover, no indication in favour of a lateralized deficit in metrical processing has been reported.

Based on this review, we can assume that the left-dominant hemisphere and more specifically the temporal lobe plays an important role in the processing of simple auditory sequences requiring rapid timing variations ranging from 10 to 100 ms. However, this left-hemisphere superiority has not been consistently demonstrated in studies using musical sequences. Although precise timing information is not usually reported in such studies, the

description of the stimuli indicates that the manipulated temporal information refers to slower changes resulting, for example, from the permutation of two notes. As previously emphasized, left-hemisphere damage affected the processing of fast but not slow temporal information explaining therefore why left cerebral lesions do not systematically interfere with time-related judgement involved in musical sequences (which usually concerned longer temporal information superior to 200 ms[33]).

To try to clarify the role of the LTL in fast temporal coding in music, we examined the consequences of unilateral temporal lobe dysfunction in processing rapid sequential information. In this chapter, we present evidence supporting the hypothesis that LTL structures are predominantly involved in auditory temporal processes within the 10–100 ms time frame. For this purpose, inter-onset interval (IOI) was manipulated in a psychophysical task of anisochrony discrimination by using simple auditory sequences, as well as in a detection task involving rhythmic changes in real tunes.

## Perception of inter-onset increment in simple auditory sequences: anisochrony discrimination

As frequently suggested, an important component of the rhythmic subjective structure is its underlying isochrony (i.e. a regular beat), which corresponds to musical metre (or tempo). In its simplest form, the metre can be expressed as a series of regular sounds. In the present study, we designed an experiment based on anisochrony perception which refers to the ability to differentiate a regular from an irregular sequence of sounds. The goal of this experiment was to test the effect of tempo on anisochrony (or irregularity) discrimination in patients with unilateral temporal lobe dysfunction, using an adaptive procedure. The rate of presentation or tempo (defined by the IOI separating the sounds in the sequence) was systematically manipulated to compare anisochrony discrimination using different tempos.

The anisochrony paradigm is very adequate for looking at temporal processing. As indicated by the results of experiments carried out in normal subjects,[34] anisochrony discrimination remains unaffected by duration or number of tones but it is influenced by tempo or presentation rate. By using sequences presented at various tempos ranging from 80 ms (fast) to 1000 ms (slow) IOI, it allows to compare perception of fast to slow sequential information without changing the task demands. Results obtained in normal listeners also showed that the discrimination thresholds for the anisochronous sequences were proportional to the size of the IOI for tempos between 300 and 1400 ms, whereas sensitivity deteriorated for IOIs of 80 ms as compared to slower tempos. This finding corroborates the assumption of different processes for temporal stimuli inferior or equal to 200–300 ms compared to longer ones.[35–37]

Furthermore, the paradigm that we designed uses an adaptive procedure which offers the opportunity to determine precise and fine individual thresholds which represent sensitive measures of perceptual abilities. In particular, the use of such a method in neuropsychology improves the chance of documenting subtle perceptual difficulties in brain-injured patients that could not be detected by classical paradigms involving fixed levels of stimulation common to all subjects.




**Figure 14.1** Example of the two sequences presented at one trial in the anisochrony discrimination task.

Two successive sequences of five tones with the same frequency, intensity, duration and attack, and decay time were presented on each trial (one regular and one irregular). The subject had to decide if the irregular sequence was in the first or second position (see Figure 14.1). Five conditions with an IOI of 80, 300, 500, 800, or 1000 ms were prepared in order to test the effect of tempo on anisochrony discrimination. The anisochrony was introduced by delaying either the second or the fourth sound of the isochronous sequence. To obtain reliable measures of anisochrony discrimination, a psychophysical procedure developed by Levitt[38] was used to determine the minimum temporal shift necessary for each individual to discriminate the regular from the irregular sequence. The size of the first shift corresponded to 10 per cent of the base interval (i.e. 50 ms for an IOI of 500 ms). The threshold of the shift detection was determined by an adaptive procedure in which the subject's response on one trial determined the size of the shift on the next trial. The perceptual threshold was then expressed as a percentage of the base IOI to allow comparison with other differential thresholds reported in the literature. Further details about the methods are provided in another paper.[39]

Eighteen patients with medically intractable epilepsy, candidates for surgical treatment at La Salpêtrière hospital (Paris) as well as a group of normal control subjects ($n = 6$), participated in this study. They all presented medial temporal lobe epilepsy associated with lateralized hippocampal sclerosis as identified by magnetic resonance imaging (MRI). None of them suffered from language disturbances and language function was lateralized in the left-hemisphere for all subjects. These patients were divided into two groups: those with right hippocampal atrophy ($n = 8$) and those with left hippocampal atrophy ($n = 10$).

According to the literature, we hypothesized that the thresholds of patients with LTL lesions would be significantly higher than the thresholds of patients with RTL lesions or normal control subjects for tempos with IOIs at or below 300 ms. In contrast, no specific deficit was predicted for slower tempos with IOIs greater than 300 ms. Finally, we predicted that for all subjects the thresholds obtained for the fastest tempo (80 ms) would be higher than those obtained for slower tempos ($\geq$300 ms) as indicated by previous results in normal listeners.

In keeping with our predictions, the results displayed in Figure 14.2 clearly showed that anisochrony discrimination thresholds for the rapid tempo (80 ms IOI) were significantly higher for the patients with LTL dysfunction (mean threshold: 27.5 per cent) than for the patients with RTL dysfunction (17.7 per cent) and for the normal control subjects (16.4 per cent). However, there were no differences between the groups for the slower tempos (from 300 to 1000 ms IOI). This finding is compatible with results of aphasic patients with left-hemisphere lesion[1,2,4] and provides strong evidence suggesting that fast auditory sequential information processing depends specifically on the integrity of LTL structures.

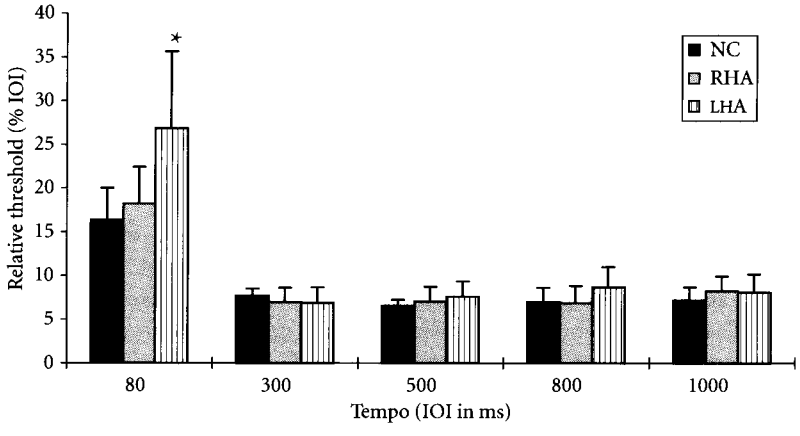



**Figure 14.2** Relative thresholds of the three groups of subjects (LHA = left hippocampal atrophy, RHA = right hippocampal atrophy and NC = Normal controls) for Inter Onset Intervals (IOI) of 80, 300, 500, 800, and 1000 msec. The bars represent the standard error of the mean.

In addition, an effect of tempo was systematically obtained for all the subjects. The thresholds for very rapid sequences (80 ms IOI = 21.43 per cent) were always higher than the thresholds obtained with slower tempos ranging from 300 to 1000 ms IOI (7.6 per cent on average), for which the different thresholds remained proportional to IOI. This finding confirmed results of a previous study carried out in healthy subjects[34] and is in agreement with the assumption that brief auditory information is processed differently from longer information.[35–37]

In a previous paper,[39] we proposed that the decreased sensitivity observed for the fastest tempo could be related to an auditory memory saturation. According to echoic memory studies (Ref. 40; see Ref. 41 for a review), acoustic properties of sounds necessitate a 200–300 ms delay following the presentation of a tone to be adequately retained. This time delay was obviously shorter for the fastest tempo (80 ms IOI). Indeed, the interval between the two sequences to be discriminated was twice the base interval, in order to maintain isochrony across the trial. Thus, the inter-sequence interval for the fastest tempo (80 ms IOI) was 160 ms whereas a minimum of 600 ms separated the sequences for the other tempos. Therefore, a saturation of echoic memory could have been responsible for the observed difference between the thresholds of the fasted tempo and the slower tempos, because the inter-sequence interval would be too short to allow consolidation of the first sequence. If such an explanation is validated in the future, the deficit documented in patients with LTL dysfunction in fast sequential auditory perception could be interpreted as an echoic memory dysfunction. Although an eventual relation between a temporal processing dysfunction and an auditory memory deficit in dysphasic children has already been suggested by Tallal and collaborators,[42] the possible link between an echoic memory dysfunction and an anisochrony discrimination deficit of fast auditory information in patients with LTL lesion remains to be clarified.

It is also possible that the decreased temporal sensitivity for the rapid tempo is related to a difficulty to determine if the irregular sequence was the first or the second one suggesting deficit in the temporal order processing. Although this judgement was not affected when slower tempos were presented, it becomes more difficult to make such a judgement in rapid sequences (subjects sometimes having trouble to separate the sequences *per se*). This additional constraint might have penalized more strongly patients with LTL dysfunction than the other subjects.

In previous studies, rapid temporal processing deficits were usually reported in language-impaired subjects (Ref. 43 for review). Results of the present investigation extend this finding to patients presenting a limited LTL lesion without associated massive language disorders, indicating that time-related disturbances are not necessarily associated with language disturbances. However, the present results differ from the data obtained in language-impaired subjects since in this latter case, subjects presented a deficit with a slower tempo (300 ms IOI) while the subjects with LTL dysfunction did not present any difficulties in this condition. It seems that the severity of the sequential auditory deficit observed in patients with or without verbal deficit varies. Although these divergent results could be explained by methodological differences between the studies, it is also possible that the severity of the temporal deficits could be correlated to the severity of verbal deficits. The patients tested in the present study might be situated at one end of the continuum, showing limited temporal processing deficits and relatively preserved verbal abilities, while dysphasic or dyslexic subjects could be located at the other end, with more severe temporal deficits and more severe language deficits. Conversely, it is also possible that rapid sequential temporal processing can be disturbed without affecting language functions, which would suggest that they depend on distinct neural substrates. Future research will be necessary to determine whether the same underlying mechanism is responsible for the deficit observed in epileptic and dysphasic patients.

## Perception of inter-onset increment in musical sequences: temporal microstructure of familiar tunes

The goal of this study was to extend to the musical domain the role of the LTL in processing small timing variations. As emphasized in the introduction, it seems that the perception of small timing differences within a musical sequence has not been investigated in neuropsychology. To explore this issue, we designed a task involving the detection of inter-onset interval increments in familiar monodic tunes improving therefore the ecological validity of the task.

Several descriptive studies showed that performers considerably modify the temporal structure as noted in the score when interpreting a piece of music (for review, see Ref. 44), a piece being never played exactly as it is written in the score. However, music played in this way seems perfect to the listeners, suggesting that we are expecting these temporal changes. Such temporal microvariations occur at specific locations and are supposed to reflect the subjective temporal structure of music. It has been demonstrated that consequently, a listener is expecting these musical microvariations in perceiving musical excerpts.[45] In an experimental study using classical pieces of music (microstructural expectations), Repp

showed that increment detection accuracy was correlated with the temporal microstructure profiles of expert performances suggesting that temporal increments of an interval which is usually lengthened are more difficult to detect than temporal increments of intervals for which no microvariations were introduced. In the cognitive literature, two hypothesis, which are not mutually exclusive, have been proposed to explain variations in the accuracy detection. A top-down hypothesis suggests that listeners' expectations reflect expressive performance of temporal microvariations[45] whereas a bottom-up hypothesis indicates that some expectations may be due to psychoacoustical characteristics of the stimulus.[44]

Based on these observations, we designed an experiment involving the detection of inter-onset interval increments introduced in musical sequences to test perception of these temporal microvariations in patients with unilateral temporal lobe lesions. The musical excerpts used were very familiar in order to generate high expectations and to prevent bias due to each listeners' musical background.

In the experimental task, half of the trials consisted of the presentation of a score (played exactly as it is written in the score) version of a familiar tune whereas the other half consisted of a modified version of the tune in which one inter-onset interval was increased by 25 per cent, producing an increment of 10–140 ms (mean = 85 ms; standard deviation = 62 ms) depending on the size (or duration) of the modified interval within the tune. The subject's task was to decide whether the presented trial corresponded to the mechanical or to the modified version of the excerpt using a two-alternative forced choice paradigm. Two types of inter-onset interval were introduced. One type corresponds to expected increments and the other one corresponds to unexpected increments. The level of expectation was determined by results of previous studies detailed elsewhere.[46] Basically, the recording of pianists' performances were compared to a score (or computerized) version of each selected tune to allow the analysis of temporal microstructure. The results of this experiment allowed the identification of one interval that was systematically lengthened in pianists' recordings as compared to metrical version, and another interval that was systematically preserved. Inter-onset interval increments were thus introduced at these specific locations and it was hypothesized that temporal increments located at 'lengthened intervals' would be more difficult to detect than temporal increments located at 'preserved intervals' since listeners would be expecting temporal increments in the first but not in the second condition. The ability to distinguish these two types of inter-onset increments was subsequently tested in a perceptual task confirming therefore the relevance of this temporal manipulation in normal nonmusician listeners. Based on these results, we designed an experimental task to evaluate the ability to detect subtle temporal changes by taking into account implicit knowledge of the rhythmical structure underlying musical listening. By manipulating temporal increments and cognitive expectations, especially in regard to the temporal (or rhythmic) dimension, it was possible to assess bottom-up as well as top-down processing involved in musical rhythm perception.

Twenty-two patients who had undergone a right ($n = 11$) or a left ($n = 11$) temporal lobe resection for the relief of medically intractable epilepsy as well as 14 normal control subjects were tested in this experiment. None of the patients presented language disorders or suffered from extra temporal lesions. Language was lateralized on the left side in all

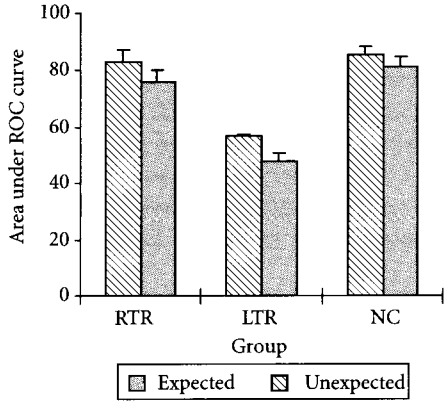



Figure 14.3 Detection accuracy derived from the area under the receiver operating curve (ROC) of the three groups of subjects (LTR = left temporal resection, RTR = right temporal resection and NC = Normal controls) for expected and unexpected inter onset increments. The bars represent the standard error of the mean.

subjects. As opposed to the previous study, the patients were tested post-operatively. The temporal resection includes the medial temporal lobe structures (hippocampal and surrounding cortex) and in some patients, the excision involves the temporal pole as well.

Considering the predominant contribution of LTL structures in processing rapid sequential information, we expected a deficit in detecting microtemporal variations in familiar tunes in patients with LTL lesion. In this study, the task of detection of temporal increments required no tonal judgment, for which the RTL is usually needed.[47–50] We therefore postulated adequate performance for patients with RTL lesion. Moreover, an effect of level of expectation of inter-onset interval was predicted for normal subjects and for patients with RTL lesion. If this effect reflects a perceptually driven processing, it should disappear in patients with LTL lesion, otherwise it might be also present in this patient group.

The results of this study are displayed in Figure 14.3 which represents detection accuracy of the three groups of subjects for expected and unexpected inter-onset increments. An area under the receiver operating curve (ROC) derived from the number of hits and false alarms was computed for each subject. As predicted, a significant deficit in detecting inter-onset increments was obtained in patients with LTL lesion as compared to subjects with RTL lesion and normal subjects, these two latter groups being not different. This deficit characterizing patients with LTL lesion cannot be attributed to difficulty in identifying the tunes since all the subjects were able to adequately recognize the melodies. The results of this study suggest that the ability to detect an inter-onset increment introduced in very well known tunes depends on the integrity of LTL structures, extending to the musical domain the previously reported role of the LTL in rapid sequential processing.

Moreover, the results showed that it is more difficult to detect expected than unexpected temporal increments for all the subjects indicating therefore that top-down processing remains functional for the three groups of subjects. Moreover, it means that expectations of nonmusician listeners were identical to those of pianists, extending to the simple monodic familiar tunes findings reported with polyphonic classical stock pieces.[45]

Although patients with LTL lesion presented a clear deficit in perceiving temporal microvariations, they seem to be able to generate temporal expectations to the same extent as healthy subjects.

To summarize, the results of this study illustrate a dissociation between a selective deficit in fine grained temporal sensitivity whereas temporal expectations seem to be relatively spared in the presence of LTL lesions, suggesting that top-down (or descending) processing may function independently from bottom-up (or ascending) processing. In addition, these data indicate that top-down processing involved in the temporal perception of musical information can still be preserved despite disrupted bottom-up processing.

## Conclusion

The experiments reported in this paper were designed to test the hypothesized role of LTL structures to musical temporal processing. We focused our interest on the perception of brief sequential information ranging from few tens to hundreds of milliseconds by comparing results obtained with very simple auditory sequences and real musical excerpts. In keeping with our predictions, the data reported in this paper suggest that LTL structures are predominantly involved in perceiving inter-onset interval increments in familiar tunes as well as in isochronous sequences, extending to the musical domain findings previously reported in speech.[12,43]

The patients included in both studies had lesions located within the medial temporal lobe structures and more specifically in the hippocampus. It remains difficult to know if the deficit could be attributed to the hippocampal lesion or to a more global dysfunction of the temporal neocortical structures. Evidence from a functional cerebral imaging study (Positron Emission Tomography with 18 Fluorodeoxyglucose, FDG-PET) carried out in a similar population of epileptic subjects indicates that the temporal lobe dysfunction associated with unilateral hippocampal sclerosis is not restricted to the hippocampus but extends to include the anterior part of the temporal neocortex corresponding to the temporal pole.[51] Although the structural lesion is apparently limited to the hippocampal formation, it does not mean that the connections between the medial temporal lobe structures and the neocortex or the surrounding and external temporal cortex *per se* remain functional. The present data does not allow to differentiate the role of the hippocampus from the contribution of the adjacent cortical areas in auditory sequential information processing. However, it seems relevant to suggest that abnormal metabolism reported in these surrounding cortical structures reflects a temporal lobe dysfunction or an interruption of the hippocampo-neocortical loop that may be responsible for the deficits obtained in our patients with LTL dysfunction.

In the present studies, the time-related deficit following LTL lesion was reproduced in two different tasks involving purely temporal stimuli as well as multidimensional musical sequences in the context of an adaptive or a fixed level paradigm, respectively. Despite numerous methodological differences, the results of these two studies provide strong support for the selective sensitivity of LTL structures in processing rapid time-related events. Indeed, the two methodologies require the detection of subtle temporal variations which was made possible by the use of regular sequences or very familiar tunes producing strong

temporal expectations. It is therefore tempting to consider that a single mechanism, depending on the processing of very brief sequential events, is responsible for impairments obtained in both tasks.

However, it is too early to assume that such impairments can be explained by a unique deficit. We already suggested that deficit in anisochrony discrimination of rapid sequences can also result from the saturation of sensory memory or from an inability to judge the temporal order. Such a hypothesis should be tested in the future to elucidate the nature of the deficit characterizing subjects with LTL dysfunction.

## Acknowledgements

We are also indebted to Xavier Gangand for programming the experiments. This research was supported by a grant from Contrat d'objectif Région Nord—Pas de Calais: 'Motricité et Cognition' (France) and by GIS 'Dysfonctionnement de la Cognition'.

## References

1. **Efron, R.** (1963) Temporal perception, aphasia and déjà vu. *Brain*, 86, 403–24.
2. **Tallal, P.** and **F. Newcombe** (1978) Impairment of auditory perception and language comprehension in dysphasia. *Brain and Language* 5, 13–24.
3. **Lackner, J. R.** and **H. L. Teuber** (1973) Alterations in auditory fusion thresholds after cerebral injury in man. *Neuropsychologia* 11, 409–15.
4. **Mills, L.** and **G. B. Rollman** (1980) Hemispheric asymmetry for auditory perception of temporal order. *Neuropsychologia* 18, 41–7.
5. **Robin, D. A., D. Tranel,** and **H. Damasio** (1990) Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. *Brain and Language*, 39, 539–55.
6. **Sherwin, L.** and **R. Efron** (1980) Temporal ordering deficits following anterior temporal lobectomy. *Brain and Language* 11, 195–203.
7. **Swisher, L.** and **I. J. Hirsh** (1972) Brain damage and the ordering of two temporally successive stimuli. *Neuropsychologia* 10, 137–52.
8. **Griffiths, T. D., *et al.*** (1997) Spatial and temporal auditory processing deficits following right-hemisphere infarction. *Brain*, 120, 785–94.
9. **Musiek, F. E., J. A. Baran,** and **M. L. Pinheiro** (1990) Duration pattern recognition in normal subjects and patients with cerebral and cochlear lesions. *Audiol.* 29, 304–13.
10. **Albert, M. L., *et al.*** (1972) A case study of auditory agnosia: linguistic and non-linguistic processing. *Cortex* 8, 427–43.
11. **Milner, B.** (1962) Laterality effects in audition. In V. Mountcastle (ed.) *Interhemispheric Relations and Cerebral Dominance*, Baltimore: John Hopkins University Press, pp. 173–201.
12. **Liégeois-Chauvel, C., *et al.*** (1999) Specialization of left auditory cortex for speech perception in man depends on temporal coding. *Cereb Cortex* 9, 484–96.
13. **Gordon, H.** (1978) Left-hemisphere dominance for rhythmic elements in dichotically-presented melodies. *Cortex*, 14, 58–70.
14. **Halperin, Y., I. Nachshon,** and **A. Carmon** (1973) Shift of ear superiority in dichotic listening to temporally patterned nonverbal stimuli. *J. Acousti. Soc. Am.* 53, 46–50.

15. Papçun, G., *et al.* (1974) Is the left-hemisphere specialized for speech, language and/or something else? *J. Acoust. Soc. Am.* 55, 319–27.
16. Robinson, G. M. and D. J. Solomon (1974) Rhythm is processed by the speech hemisphere. *J. Exp. Psychol.* 102, 508–11.
17. Peretz, I. and J. Morais (1980) Modes of processing melodies and ear asymmetry in non-musicians. *Neuropsychologia* 18, 447–89.
18. Mavlov, L. (1980) Amusia due to rhythm agnosia in a musician with left-hemisphere damage: a non-auditory supramodal defect. *Cortex* 16, 331–8.
19. Prior, D. A., G. Kinsella, and J. Giese (1990) Assessment of musical processing in brain-damaged patients: implications for laterality of music. *J. Clinical Exp. Neuropsychol.* 12, 301–12.
20. Kester, D., *et al.* (1991) Acute effect of anterior temporal lobectomy on musical processing. *Neuropsychologia* 29, 703–8.
21. Michel, F., F. Peronnet, and B. Schott (1980) A case of cortical deafness: clinical and electrophysiological data. *Brain and Language* 10, 367–77.
22. Penhune, V. B., R. J. Zatorre, and W. H. Feindel (1999) The role of auditory cortex in the retention of rhythmic patterns as studied in patients with temporal lobe removals including Heschl's gyrus. *Neuropsychologia* 37, 315–31.
23. Peretz, I. (1990) Processing of local and global musical information by unilateral brain-damaged patients. *Brain* 113, 1185–205.
24. Reitan, R. M. and D. Wolfson (1989) The Seashore rhythm test and brain functions. *The Clinical Neuropsychologist* 3, 70–8.
25. Shapiro, B. E., M. Grossman, and H. Gardner (1981) Selective musical processing deficits in brain damaged populations. *Neuropsychologia* 19, 161–9.
26. Liégeois-Chauvel, C., *et al.* (1998) Contribution of different cortical areas in the temporal lobes to music processing. *Brain* 121, 1853–67.
27. Platel, H., *et al.* (1997) The structural components of music perception. A functional anatomical study. *Brain* 120, 229–43.
28. Dowling, W. J. and D. Harwood (1986) *Music Cognition.* New York, NY: Academic Press.
29. Povel, D. J. (1981) Internal representation of simple temporal patterns. *J. Exp. Psychol: Hum. Percept. and Perform.* 7, 3–18.
30. Povel, D. J. and P. Essens (1985) Perception of temporal patterns. *Music percept.*, 2, 411–40.
31. Repp, B. H. (1995) Detectability of duration and intensity increments in melody tones: partial connection between music perception and performance. *Percept. Psychophy.* 57, 1217–32.
32. Polk, M. and A. Kertesz (1993) Music and language in degenerative disease of the brain. *Brain and Cognit.* 22, 98–117.
33. Fraisse, P. (1982) Rhythm and tempo. In D. Deutsch (ed.) *The Psychology of Music.* New York: Academic Press, pp. 149–78.
34. Ehrlé, N. and S. Samson (1996) Détection de l'irrégularité temporelle: influence de la durée et du tempo. In F. Anceaux and J. M. Coquery (eds) *Sciences Cognitives, Individus et Société. Actes du 6ème Colloque de l'Association pour la Recherche Cognitive.* Villeneuve d'Ascq, pp. 209–14.
35. Michon, J. A. (1964) Studies on subjective duration: I. Differential sensitivity in the perception of repeated temporal intervals. *Acta Psychol.* 22, 441–50.
36. Nakajima, Y., G. ten Hoopen, and R. Van der Wilk (1991) A new illusion of time perception. *Music Percept.* 8, 431–48.

37. **ten Hoopen, G., *et al.*** (1995) Auditory isochrony: time shrinking and temporal patterns. *Percept.* 24, 577–93.

38. **Levitt, H.** (1971) Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* 49, 467–77.

39. **Ehrlé, N., S. Samson,** and **M. Baulac** (2001) Processing of rapid auditory information in epileptic patients with left temporal lobe damage without dysphasia. *Neuropsychologia* 39, 525–31.

40. **Massaro, D. W.** (1972) Preperceptual images, processing time and perceptual units in auditory perception. *Psycholog. Rev.* 79, 124–45.

41. **Cowan, N.** (1984) On short and long auditory stores. *Psycholog. Bull.* 96, 341–70.

42. **Tallal, P., R. L. Sainburg,** and **T. Jernigan** (1991) The neuropathology of developmental dysphasia: behavioral, morphological, and physiological evidence for a pervasive temporal processing disorder. *Reading and Writing: an Interdisciplinary Journal* 3, 363–77.

43. **Tallal, P., S. Miller,** and **R. H. Fitch** (1993) Neurobiological basis of speech: a case for the preeminence of temporal processing. *Annals of the New York Academy of Sciences* 682, 27–47.

44. **Penel, A.** and **C. Drake** (1998) Sources of timing variations in music performance: a psychological segmentation model. *Psycholog. Res.* 61, 12–32.

45. **Repp, B. H.** (1992) Probing the cognitive representation of musical time: structural constraints on the perception of timing perturbations. *Cognit.* 44, 241–81.

46. **Ehrlé, N.** (1998) Traitement temporel de l'information auditive et lobe temporal. Ph.D. Thesis. France: University of Reims.

47. **Samson, S.** and **R. J. Zatorre** (1988) Melodic and harmonic discrimination following unilateral cerebral excision. *Brain and Cognit.* 7, 348–60.

48. **Johnsrude, I., B. B. Penhune,** and **R. J. Zatorre** (2000) Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain* 123, 155–63.

49. **Zatorre, R. J.** (1988) Pitch perception of complex tones and human temporal-lobe function. *J. Acoust. Soc. Am.* 84, 566–72.

50. **Zatorre, R. J.** and **S. Samson** (1991) Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain* 114, 2403–17.

51. **Semah, F. *et al.*** (1995) Is interictal temporal hypometabolism related to mesial temporal sclerosis? A positron emission tomography/magnetic resonance imaging confrontation. *Epilepsia* 36, 447–56.

CHAPTER 15

# CEREBRAL SUBSTRATES OF MUSICAL IMAGERY

ANDREA R. HALPERN

## Abstract

Musical imagery refers to the experience of 'replaying' music by imagining it inside the head. Whereas visual imagery has been extensively studied, few people have investigated imagery in the auditory domain. This chapter reviews a program of research that has tried to characterize auditory imagery for music using both behavioural and cognitive neuroscientific tools. I begin by describing some of my behavioural studies of the mental analogues of musical tempo, pitch, and temporal extent. I then describe four studies using three techniques that examine the correspondence of brain involvement in actually perceiving vs imagining familiar music. These involve one lesion study with epilepsy surgery patients, two positron emission tomography (PET) studies, and one study using transcranial magnetic stimulation (TMS). The studies converge on the importance of the right temporal neocortex and other right-hemisphere structures in the processing of both perceived and imagined nonverbal music. Perceiving and imagining songs that have words also involve structures in the left hemisphere. The supplementary motor area (SMA) is activated during musical imagery; it may mediate rehearsal that involves motor programs, such as imagined humming. Future studies are suggested that would involve imagery of sounds that cannot be produced by the vocal tract to clarify the role of the SMA in auditory imagery.

**Keywords:** Musical imagery; Behavioural studies; Cognitive neuroscience

## Introduction

Many people experience the sounds of music in two distinct but related manners. Listening to live or recorded music is, of course, the way we commonly think of enjoying music. However, many people also report that they can reexperience music by imagining it in their heads. This can be pleasurable or not, depending on the circumstances, but in either case, the experience appears to be a vivid one, even among people untrained in music. In fact, I am often asked how to 'stop' a tune from obsessively intruding into everyday thoughts. Highly trained musicians report that they can use auditory imagery to help them in their everyday tasks, such as 'hearing' music as they read musical notation.

In recent years, I have studied the characteristics of this auditory imagery experience, initially from a behavioural perspective and more recently from a cognitive neuroscientific perspective. I have been most interested in the auditory imagery experiences of untrained or moderately trained musicians, although many interesting questions derive from considering

experts as well. My paradigms and experimental logic have been derived partly from the more extensive literature in visual imagery; other approaches have been created anew as the need arose. In studying any mental imagery, the challenge is to externalize what is essentially an internal experience to examine what it means to have, in the case of musical imagery, a 'tune inside the head'.

## Behavioural approach

In the 1980s, I carried out an extensive series of experiments with young adult college students to explore how auditory imagery of music may be characterized. I explored the mental representation of *tempo*[1,2] and *pitch*[3] in familiar songs. To study representation of tempo, I asked participants to set their preferred tempo for a familiar tune being played on a computer. The program allowed continuous adjustments of tempo until the tune sounded right to the listeners. I then asked them to imagine the same tunes and set a metronome to the imagined tempo. Tempo settings were highly correlated in the perception and imagination tasks. I also showed that people can manipulate the tempo of an imagined song, within limits, to make it 'sound' very fast or very slow. Finally, I showed that musicians showed impressive consistency when asked to tap out the tempo of an imagined song on one day and again two to five days later (average standard deviation = 1.4 metronome settings over four attempts), although nonmusicians were somewhat less consistent.

In the domain of pitch, I asked people to imagine familiar songs and hum the starting note corresponding to their auditory image of the song. In a second task, I asked people to think of the starting note of a familiar song and select that note from a piano keyboard (all the tunes were popular folk or children's songs, unlikely to have been heard with any particular starting note in the past). Once again, consistency of pitch production or selection was impressive for musically unselected participants, even over a delay of several days (average standard deviation = 1.25 semitones for production, two semitones for selection over four attempts). Finally, I asked people to rate how similar a played note was to the opening note of a tune they had imagined and produced a few days earlier. The person's own preferred note was rated quite highly, as were starting notes a major third lower, minor third higher, and a perfect fifth higher than the preferred notes (these are all musically coherent intervals). However, subjects rated notes only one semitone higher or lower than their preferred note as dissimilar to their imagined pitch, showing a fairly acute sense of pitch representation.

These behavioural tasks seem to show a veridical representation of characteristics particular to music. However, I also explored the *extension in time* that is characteristic of almost all auditory stimuli and thus ought to be captured in auditory imagery. I explored this temporal aspect in the subsequent cognitive neuroscientific studies of auditory imagery for music.

To begin, I modified a paradigm introduced in visual imagery by Kosslyn *et al.*[4] In that study, they asked people to learn a map of an imaginary island. Pairs of features on the island were presented, and subjects had to 'mentally scan' between them. Latencies to do so were highly correlated with actual distance between the features on the map, suggesting that the mental representation was preserving an analogue of space.

To extend this paradigm to musical imagery,[5] I selected a number of familiar songs where unique lyrics fell on specified beats of the tune. For instance, the first line of the American national anthem, 'The Star Spangled Banner,' is 'Oh, say can you see by the dawn's early light?' In the most relevant study in that series, I presented the title of the song, followed by one lyric from the first line of the song, followed by the second lyric. The task was mental pitch comparison: Was lyric 2 higher or lower in pitch than lyric 1? No singing or humming was allowed. The lyrics were either close together ('Oh' and 'can') or far apart in the actual tune ('Oh' and 'dawn's'). I found that this was a difficult task for my musically unselected subjects; nevertheless, reaction times increased nearly linearly as the separation in beats between the two lyrics increased. Subjects reported using auditory imagery to accomplish the task, even though they were not instructed to do so. The consistency of reaction time data with this report strongly suggested to me that this task had captured the extension in time of auditory imagery for music.

## Cognitive neuroscientific approach

The behavioural studies cited above have in common the logic that if responding to an imagined stimulus resembles responding to a perceived stimulus, we may conclude that imagery is a particularly vivid and veridical form of mental representation. However, this comparison between imagery and perception may be strengthened by examining the similarities in neural underpinnings of the two processes. This approach is complementary to the behavioural approach in that at least to some extent behavioural responses might be influenced by external influences such as demand characteristics or experimenter expectancies.[6] However, it is unlikely that people can influence their own brain structures or activities. We may thus look to similarities in the brain loci involved in auditory imagery and perception to gain a better perspective on the processing of similarities and differences in the two types of tasks. An argument for this approach is well articulated by Farah[7] for the visual domain.

The strongest hypothesis is of course that the brain areas would be identical in auditory imagery and perception. This may serve as a guiding null hypothesis, but we would not in reality expect this amount of overlap; people other than those hallucinating can tell the difference between imagining and hearing a song. However, the extent of overlap may tell us how similarly the brain processes hearing and imagining hearing. Brain areas uniquely active in imaging tasks can by extension inform us as to the additional or alternative processing demands imposed in imagery by having to, in effect, create as well as perceive the stimulus.

As noted above, the approach has been reasonably successful in the visual domain. For instance, Kosslyn *et al.*[8] used positron emission tomography (PET) to measure brain activity during parallel perceptual and imagined visual tasks. They found quite a few areas activated in common, even to the extent that varying the size of the presented object and varying the size of an imagined object activated similar brain areas in similar ways. Few people have looked at auditory imagery using parallel perceptual and imagery tasks. However, the neural structures responsible for some aspects of musical perception are well defined. My partner for most of the studies described in the following sections has been

Robert Zatorre. When we began our first study, he and Samson, among others, had already established that a number of musical tasks are impaired after lesions in the right temporal neocortex.[9–12] Thus our initial hypothesis was that the same region would be involved in musical imagery tasks that resembled musical perceptual tasks.

What follows are brief descriptions of a series of studies using three different cognitive neuroscience techniques to investigate the cerebral substrates of musical imagery. The first is a lesion study, which can give information about the necessary involvement of some brain areas in an activity. The next two studies use PET paradigms to investigate brain areas that are active in auditory imagery for verbal and nonverbal songs, respectively. Finally, I describe a study using transcranial magnetic stimulation (TMS) that returns to the logic of the lesion studies, in that a brain area in normal people is disrupted for a brief time and the ensuing decrements in performance are measured.

## Lesion study

My first study with Zatorre[13] examined the effect of right temporal lobectomy on performance of a mental pitch comparison task that I described earlier.[5] Participants were patients having undergone surgical excision of the anterior portion of the right or left temporal lobe (excluding the primary auditory cortex) for relief of intractable epilepsy. We tested the patients either two weeks after surgery (approximately two-thirds of the patients), or at follow-up medical appointments a year or more after surgery. A control group consisted of age- and education-matched neurologically normal individuals. Preoperative testing insured that the patients had typical language representation. All participants were familiar with the songs we used, and they all passed a brief test of pitch discrimination ability.

The imagery task was essentially the one described earlier. Participants saw a title of a song, for instance, 'Jingle Bells', followed by the first line of the song, with two words in capital letters, such as 'Dashing through the SNOW, in a one-horse open SLEIGH'. They decided if SLEIGH was higher or lower in pitch than SNOW, and pressed a button to answer. The parallel perception task, which was always presented first, was the same except that the song, sung with lyrics, was actually presented to participants from a digitized sound file while they made their judgement. Accuracy and reaction times were recorded. The reaction time pattern replicated the pattern I had shown earlier[5] of increasing latency with increasing distance in beats between the lyrics, although accuracy turned out to be the measure of interest here.

Accuracy results are shown in Figure 15.1. It is clear that the imagery task was more difficult than the perception task, as expected. It is also clear that the right temporal lobectomy group was impaired relative to controls on both the imagery and perception tasks, to an equal extent, whereas the left temporal lobectomy group was impaired on neither.

This pattern was consistent with our hypothesis that the right temporal lobe is an important mediator of musical imagery, as it has been shown to be for musical perception. We considered which particular aspect of the task was most likely to have been subserved by the right temporal lobe. All participants passed a simple tone discrimination task, so a

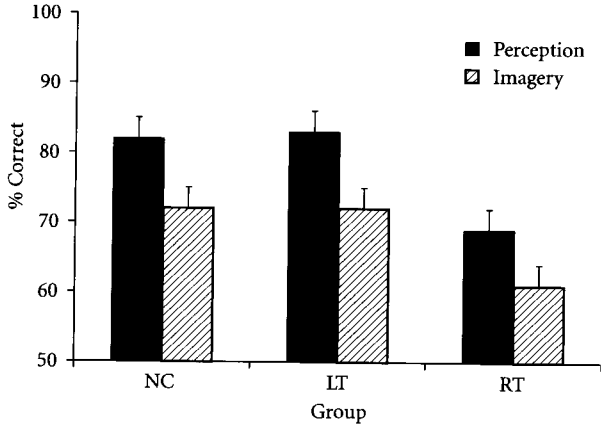



Figure 15.1 Mean per cent correct, plus standard error, for normal control subjects (NC), patients with left temporal lobe excision (LT), and patients with right temporal lobe excision (RT) in perception and imagery tasks in the lesion study.

deficit in pitch comparison itself was unlikely to be the cause of the group's impairment. Furthermore, this group was impaired in the perception task, when the tune did not need to be generated from memory, making memory retrieval unlikely to be the right temporal lobe's crucial contribution. Thus we concluded that the right temporal deficit is likely to be primarily due to a difficulty in remembering the (imagined) pitches to be compared.

## PET studies

Lesion studies allow one to draw conclusions about the areas *crucial* to perform certain tasks. They do not, however, allow any conclusions about what brain areas are *active* in the tasks, whether or not they are crucial. Lesion studies also are confined to the particular excisions made available by accidents of nature or deliberate surgery. If we want to study the somewhat broader question of what areas throughout the brain are active during imagery and perception tasks, then functional brain imaging technology is a useful addition to our investigations. Another advantage of brain imaging techniques is that neurologically normal participants can be studied, in contrast to the neurologically abnormal people studied, by definition, in lesion work. In our next study, Zatorre and I[14] used PET to observe the changes in cerebral blood flow (CBF) as participants performed musical imagery tasks similar to the ones already described.

In this approach, tasks are set up in a series so that more complex tasks include elements of the simpler tasks presented to subjects. Subtracting the activation observed in the simpler task from that observed in the more complex task allows one to isolate activation unique to the critical components of the more complex task. For instance, in almost any cognitive task, the brain activation due to the simple registration of auditory or visual stimuli, and the

motor activation involved in pressing buttons, are not of major interest. Therefore, control tasks are often presented that involve simply seeing or hearing a stimulus and pressing a button in response. Subtracting activation due to these lower-level components gives a better indication of the mental work involved in the tasks of more interest.

In our study, we presented the imagery and perception tasks that we used previously,[14] except we presented single words instead of complete first lines from the songs. A third task was meant to control for such simpler tasks as visual perception and button pressing. This visual baseline presented pairs of words used in the imagery and perception task, but each pair member came from a different song. For each pair presented on the screen, participants had to judge which one was longer in length and press a button for their choice. Because the pairs were scrambled, we assumed listeners would not be reminded of songs; thus auditory imagery should not be activated. But over the course of the baseline condition, all words used in the main task would be seen. All songs were very familiar to the subjects, who were 12 healthy right-handed young adults. Most had some musical training but none were serious musicians.

All participants underwent a magnetic resonance imaging (MRI) scan in order to allow us to later localize CBF activity to the appropriate anatomy for each person. For the test session, the three tasks were explained and there was a short practice session. After being placed in the scanner, the subjects received an injection of radioactively labelled water ($H_2{}^{15}O$) required to index the CBF activity, and performed the visual baseline task. This was then repeated for the perception task and then the imagery task, in that order.

As expected, people were more accurate on the perception task than the imagery task, and once again reaction times in both imagery and perception increased as a function of distance between the lyrics in the real tunes. The analysis then proceeded by subtracting the activation shown in the baseline task from the perception task and also the baseline task from the imagery task. A graphic representation of the results is shown in panel I of Figure 15.2. Here we see in the upper part of panel I that the primary auditory area (located in superior temporal gyrus, or STG) is quite active when listening to sounds, as one would expect. More interesting is the fact that several areas of the STG adjacent to the primary auditory cortex (secondary auditory cortex) are also active when people were just imagining the sounds (lower part of panel I). Panel II shows several other areas of correspondence between the imagery and perception tasks: several areas in the frontal lobe were active in both tasks, as was one area in the parietal lobe. Panel III shows activity in the supplementary motor area (SMA, involved in motor planning) in both tasks, stronger in imagery than perception. When we subtracted the activity in the perception task from that in the imagery task, only four brain areas were unique to imagery, two of which are pictured in panel IV. These two areas, the thalamus and inferior frontopolar areas, are known to be involved in memory functions, and may be associated with the extensive memory demands attached to the imagery task.

The results of this study supported several of our earlier ideas but also raised new questions. Consistent with our lesion study,[13] several areas of the temporal lobe classified as secondary auditory cortex were activated when people were carrying out mental pitch comparisons, in the absence of any overt auditory stimulation. Thus we concluded that the right STG is both *active* during such tasks (shown by the PET study), as well as *necessary*

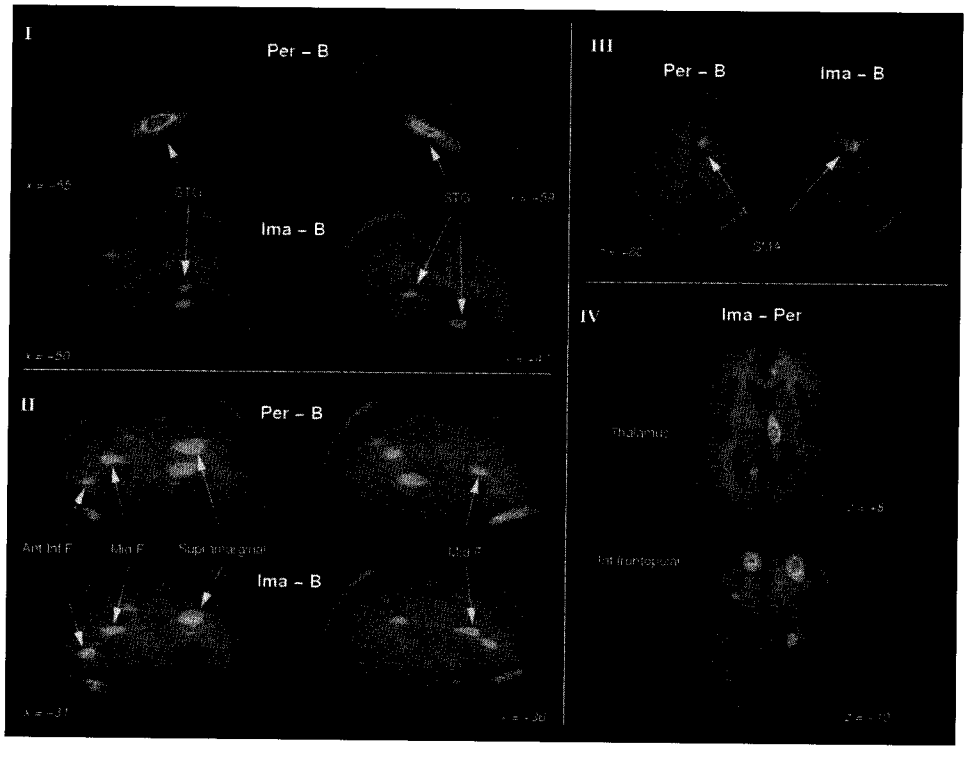



**Figure 15.2** Brain areas showing activation in the perception minus baseline, imagery minus baseline, and imagery minus perception conditions of the verbal tune PET study. (See Plate 2 in colour section.)

for carrying out the task successfully (shown by the lesion study). We were also not surprised by the areas of activation in the frontal lobes. Activation in frontal areas similar to the ones seen here was found by Zatorre *et al.*[15] in a PET study in which participants had to judge relative pitches of notes in novel (perceived) tunes, and in a study where the pitches of spoken syllables were compared.[16]

Two results were more unexpected. First, contrary to the literature cited earlier that consistently showed right-sided asymmetries in brain activation during music processing,[15] and deleterious consequences of right-sided lesions for music processing,[9–13] here almost all the activated areas were bilateral, except for right-sided activation in the thalamus. A second unexpected finding was the strong activation of the SMA during imagery even though subjects were not actually producing any motor movements. Both these findings might be related to the fact that the stimulus songs all had words, which participants had to access during the tasks. The left frontal regions have been implicated in various overt verbal tasks,[17,18] so that the bilateral activation seen here might be reflecting the dual nature of the stimulus to be processed (words plus music). Activation of the SMA has been observed when subjects are asked to generate internal speech[19] and to overtly vocalize

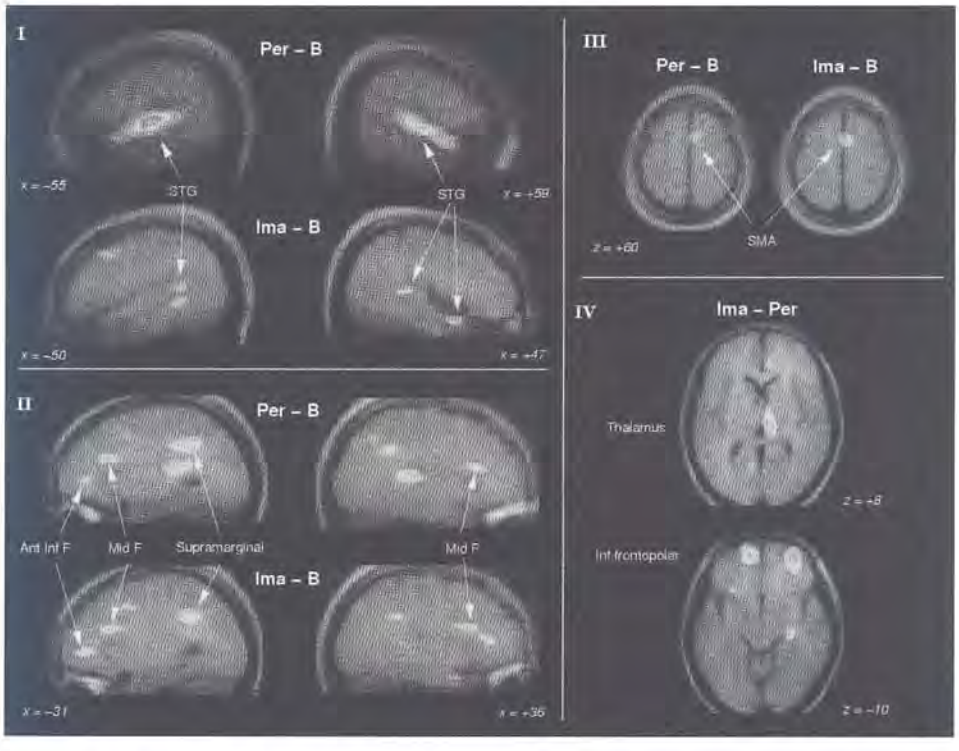



**Figure 15.2** Brain areas showing activation in the perception minus baseline, imagery minus baseline, and imagery minus perception conditions of the verbal tune PET study. (See Plate 2 in colour section.)

for carrying out the task successfully (shown by the lesion study). We were also not surprised by the areas of activation in the frontal lobes. Activation in frontal areas similar to the ones seen here was found by Zatorre *et al.*[15] in a PET study in which participants had to judge relative pitches of notes in novel (perceived) tunes, and in a study where the pitches of spoken syllables were compared.[16]

Two results were more unexpected. First, contrary to the literature cited earlier that consistently showed right-sided asymmetries in brain activation during music processing,[15] and deleterious consequences of right-sided lesions for music processing,[9–13] here almost all the activated areas were bilateral, except for right-sided activation in the thalamus. A second unexpected finding was the strong activation of the SMA during imagery even though subjects were not actually producing any motor movements. Both these findings might be related to the fact that the stimulus songs all had words, which participants had to access during the tasks. The left frontal regions have been implicated in various overt verbal tasks,[17,18] so that the bilateral activation seen here might be reflecting the dual nature of the stimulus to be processed (words plus music). Activation of the SMA has been observed when subjects are asked to generate internal speech[19] and to overtly vocalize

music.[20] Thus the SMA activation observed here in our imagery task may reflect a subvocal rehearsal process of either words or music to support performance on an otherwise difficult task.

In light of these issues, our next PET study[21] used only nonverbal tunes. We were interested to see whether removing words from our task would lead to more right-sided activation than we saw earlier, consistent with various findings in the music perception literature. We also wanted to see if SMA would be active even when potential rehearsal devices would not involve words. Finally, we wanted to try a different type of imagery task to see if our results would generalize over paradigms.

To this end, we developed a stimulus pool of tunes that were familiar but did not have lyrics. These included movie and television themes, classical excerpts, and miscellaneous tunes such as the Westminster (Big Ben) chimes. In our main task (cue/imagery) we played the first few notes of a theme as a cue, and asked participants to imagine the theme to the end of its first phrase (this task and the materials had been presented to people in advance of scanning to familiarize them with task parameters). To have a behavioural index of auditory imagery, the played excerpts differed in length. If subjects were carrying out instructions as we intended, latency to press the button should increase from our shortest (2.2 s on average) to longest excerpts (6.2 s on average).

We also had some control tasks. For these, we took the first few notes of each real tune, and scrambled the note order so that the cue did not elicit a memory of any real tune. In the control task for simple listening and button pressing we presented these 'fake' cues and simply had people press a button after each one (control). A second control task involved imagery but no retrieval from long-term memory. In this task, we presented the fake cue and asked for people to simply reimagine it immediately after presentation (control/imagery). The subtraction of interest for current purposes is the cue/imagery minus the control task.

As previously, PET scanning was undertaken in conjunction with an MRI to provide anatomical localization of CBF activation for each person. Eight healthy, right-handed young adults participated, who had from 3 to 16 years of musical training. The conditions were presented in order of control, cue/imagery, and control/imagery.

As predicted, the average time to press a button indicating imaging of the tune was complete in the cue/imagery condition varied proportionally to the length of the tune. Thus we are confident that subjects were following our imagery instructions. The results of the cue/imagery minus control subtraction are shown in Figure 15.3. As in our previous study,[14] we found activation in the secondary auditory cortex (marked STG), although this time the activation was in the right but not left temporal lobe. We also found activation in several regions of the frontal lobe (inf F), most of which were more prominent on the right than the left side. Finally we once again found strong activation in the SMA.

We thus confirmed several findings from our study with verbal tunes:[14] areas normally concerned with processing of auditory information are recruited even when the auditory information is internally generated. This occurred even with a different behavioural task and different songs than had been used earlier. We also confirmed the activity of the SMA in our task. The fact that this area was active even though no verbal rehearsal could logically have been taking place suggests that the SMA is involved with some kind of subvocal

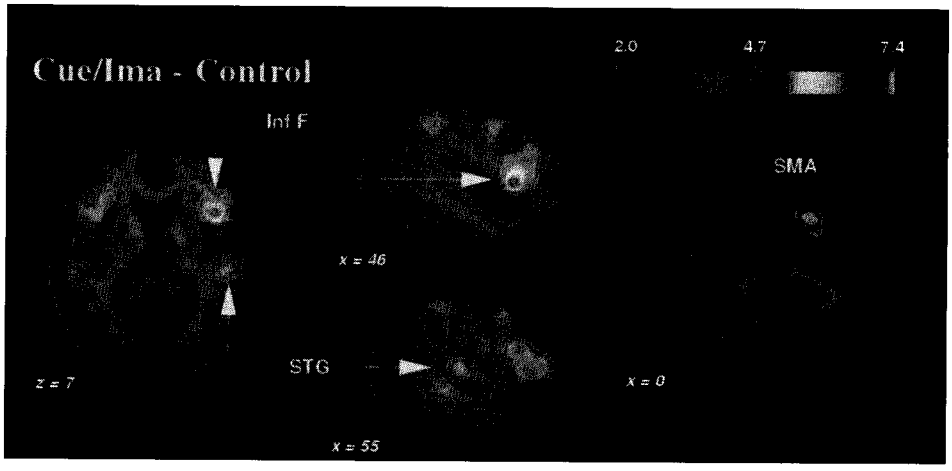



Figure 15.3 Brain areas showing activation in the cue/imagery minus control subtraction in the nonverbal tune PET study. (See Plate 3 in colour section.)

humming strategy during the imagery generation. The frontal areas activated here are also approximately the same as some of the frontal areas seen in our previous task. Because the current task did not require extensive involvement of working memory (mental pitch comparison was not required, merely internal scanning of the tune), it is likely that the areas in common in our two studies reflect retrieval from musical semantic memory rather than working memory.

The major new contribution here was that once words were removed from the stimulus and the task, we found prominent right-sided asymmetry in the areas active in the imagery task. The temporal lobe activation is consistent with our lesion study[13] that showed similar decrements from right-sided temporal lobectomy in both imagery and perception tasks. Thus we conclude that this area is both active in and necessary to the support of musical imagery tasks. The second right-sided asymmetry involved a region in the frontal lobes. This area, on the left, has been implicated in retrieval from verbal semantic memory,[22] suggesting that retrieval from semantic memory may be lateralized depending on the type of material. Another right-sided asymmetry was found in activation of the thalamus, a subcortical area involved in memory among other functions. In our earlier PET study,[14] that activation was above our statistical threshold; here it was just below the threshold and thus was not pictured with our main results. However, we can take this as at least suggestive evidence that a circuit involving temporal, frontal, and subcortical areas on the right are important in imagined music processing.

## Transcranial magnetic stimulation

The final study in the series I am reviewing is currently unpublished, and was conducted in partnership with Alvaro Pascual-Leone, Fumiko Maeda, and Gottfried Schlaug. TMS is

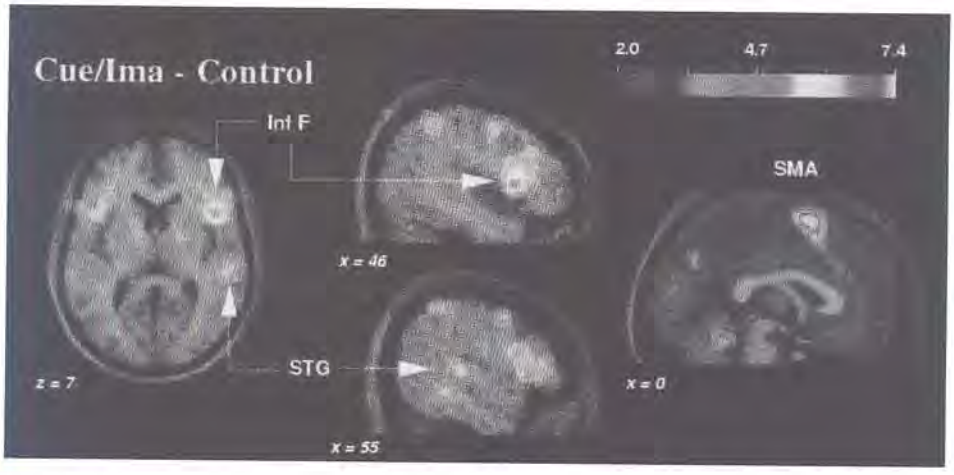



Figure 15.3 Brain areas showing activation in the cue/imagery minus control subtraction in the nonverbal tune PET study. (See Plate 3 in colour section.)

humming strategy during the imagery generation. The frontal areas activated here are also approximately the same as some of the frontal areas seen in our previous task. Because the current task did not require extensive involvement of working memory (mental pitch comparison was not required, merely internal scanning of the tune), it is likely that the areas in common in our two studies reflect retrieval from musical semantic memory rather than working memory.

The major new contribution here was that once words were removed from the stimulus and the task, we found prominent right-sided asymmetry in the areas active in the imagery task. The temporal lobe activation is consistent with our lesion study[13] that showed similar decrements from right-sided temporal lobectomy in both imagery and perception tasks. Thus we conclude that this area is both active in and necessary to the support of musical imagery tasks. The second right-sided asymmetry involved a region in the frontal lobes. This area, on the left, has been implicated in retrieval from verbal semantic memory,[22] suggesting that retrieval from semantic memory may be lateralized depending on the type of material. Another right-sided asymmetry was found in activation of the thalamus, a subcortical area involved in memory among other functions. In our earlier PET study,[14] that activation was above our statistical threshold; here it was just below the threshold and thus was not pictured with our main results. However, we can take this as at least suggestive evidence that a circuit involving temporal, frontal, and subcortical areas on the right are important in imagined music processing.

## Transcranial magnetic stimulation

The final study in the series I am reviewing is currently unpublished, and was conducted in partnership with Alvaro Pascual-Leone, Fumiko Maeda, and Gottfried Schlaug. TMS is

a noninvasive method of affecting underlying brain tissue by application of a magnetic pulse to the outside of the skull. The pulse is discharged via a coil that can be placed on the skull overlying the brain area of interest, using external anatomical landmarks, or in our case an MRI of each subject, to locate the areas of interest. A high-frequency series of pulses (10–20 Hz) has been shown to excite the underlying brain tissue, and low-frequency (1-Hz) sequences have been shown to inhibit the underlying tissue.[23] The inhibition is temporary, and the technique is safe for most people, although some discomfort may be experienced by the tapping sound the coil makes. Precautions are taken to exclude people with a history of seizure, as well as people who would be excluded from any study involving MRI, such as anyone with metal implants in the skull. The excitatory aspect of TMS is of interest to clinicians and researchers in the treatment of mental illness such as depression.[24] The inhibitory function of low-frequency TMS is of interest to researchers in that a temporary 'lesion' can be created in otherwise normal people, and subsequent effects on behaviour can be studied.

TMS has already been used to investigate visual imagery. Kosslyn, Pascual-Leone, and colleagues[25] showed that an application of 10 min of 1-Hz TMS to the visual cortex impaired later performance of both a visual perception and equivalent visual imagery task. They took this as evidence, in accord with lesion logic, that the primary visual cortex is necessary for performance of the visual imagery task, strengthening the argument that visual perception and imagery are mediated by common structures.

In our TMS study, we selected three brain areas that had been shown to be active during auditory imagery in my previous work, plus one control area. We also wanted to look at auditory imagery for both verbal and nonverbal tunes, as some different brain areas were shown to be active for these two types of tasks. To this end, we needed an auditory imagery task that would be suitable for both types of tunes. We modified a task used by Smith,[26] in which participants are given the title of a familiar tune and then asked about a pitch relationship between two notes indexed by ordinal position. In our case, we asked participants to judge whether the second note of a given tune was higher or lower than its first note. Verbal and nonverbal tunes could both be tested, and we intermixed the two types in the test.

The general outline of a session was to administer the auditory imagery task as a pretest. We then determined each person's motor threshold, or the lowest amount of energy that, when applied to the skull overlying the hand area of the motor cortex, just made the person's finger twitch. Each person received TMS at 90 per cent of that value. Administration of 10 min of 1-Hz TMS to a designated brain area ensued, followed immediately by another auditory imagery task as a posttest (using different tunes for pretest and posttest). When more than one brain area was investigated in a single session, 30 min of rest intervened between testing different brain areas, to allow the effects of the TMS to dissipate entirely.

Participants were nine right-handed adults with average age of 34, all of whom had some musical background (four were active musicians). Each person underwent an MRI scan before the TMS session. This allowed us to locate, using anatomical measurements, the areas of interest to us. The control area we selected was primary visual cortex (coil placed at the back of the skull), which had not been implicated in auditory imagery tasks before. The three experimental areas were the SMA (top of the skull a little more than midway towards the back), and the left and the right auditory cortex (above and slightly behind the

top of the ear in most people). Because of the anatomy of the auditory cortex, we could not localize our coil placement exactly on secondary auditory cortex. We located the coil instead as near as possible on the primary auditory cortex, with the assumption that the TMS activation would also likely affect the secondary cortex that is immediately adjacent to that area. Our prediction was that after 10 min of low-frequency stimulation, performance overall (in accuracy and/or time) would decrease for the experimental but not for the control areas. We made a further prediction that both left and right auditory cortex stimulation would impair processing of verbal tunes, but that the right auditory cortex stimulation would only impair nonverbal tunes.

Brain areas were tested in different orders for each person. However, as early results seemed to indicate that stimulation to the SMA was showing no effect, three volunteers did not receive SMA stimulation, in an effort to reduce subject time and discomfort. Thus all results except those for SMA were based on nine subjects; those for SMA were based on six subjects.

Accuracy was quite high in the task, so results concentrated on reaction times. Figure 15.4 shows the results for each brain area for the verbal and the nonverbal tasks, both before and after TMS stimulation. The nonverbal tunes are a little less familiar to people, and thus they were responded to on average more slowly than the verbal tunes. Inspection of the four panels shows that in accord with the prediction, the TMS made no difference when applied to the visual cortex. But, surprisingly, TMS made no significant difference when applied to SMA and left auditory cortex. The lower right panel does show a predicted effect: TMS applied to right auditory cortex slowed down responses to the imagery task, but only for the nonverbal tunes.

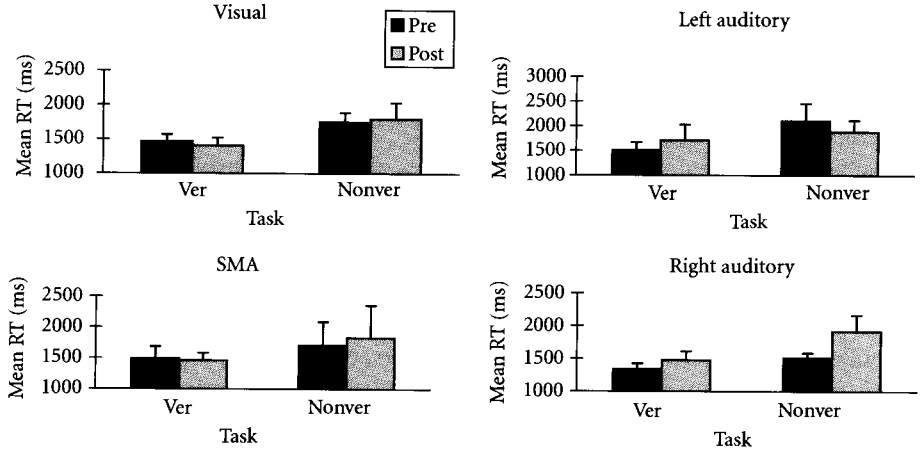



Figure 15.4 Mean reaction time before and after TMS for the verbal (Ver) and nonverbal (Nonver) imagery tasks, after application to the visual cortex, supplementary motor area (SMA), left auditory cortex, and right auditory cortex.

Stimulation to the two auditory areas were the most comparable in terms of subjective experience: the coil clicks were rather loud in the ear, and some jaw muscles were activated by the stimulation. Thus we compared those two areas statistically in an analysis of variance, using the three factors of tune (verbal, nonverbal), area (left, right auditory) and test (pre-, post-); all factors were within subjects. The main effect of tune was significant, $F(1, 8) = 6.24$, $p < 0.05$, but the crucial three-way interaction of tune, area, and test just failed to reach significance, $F(1, 8) = 3.26$, $p = 0.10$. Because of the strong prediction about these areas, we made a specific comparison of the pre- and posttests after right auditory stimulation. The comparison was not significant for verbal tunes, but was for nonverbal tunes, Newman–Keul's $t = 4.173$, $p = 0.05$.

As far as I know, this study was the first use of TMS to investigate auditory areas. As mentioned above, the coil had to be localized over primary auditory cortex, which in my previous studies was not activated in auditory imagery tasks. Thus we were relying on the somewhat weaker collateral activation to the secondary auditory cortex to show inhibition in the areas of more interest to us. Therefore, although the results were not as statistically strong as we would like, the increase in latency to nonverbal tunes after nonverbal tunes only is consistent with the pattern of data shown in our PET study with nonverbal tunes.[21] The isolation of this effect to non-verbal tunes is particularly notable, because the verbal tunes were tested in the same session with the same subjects, and did not elicit the inhibition.

The lack of effect after SMA stimulation, after strong activation of this area shown in two PET studies, is more puzzling. Null effects are always hard to interpret, but two possibilities suggest themselves. One is methodological: the neural fibres subserving SMA bend as they move ventrally from the skull surface, and it may be that our particular configuration of coil type and its placement simply did not activate the area sufficiently. On the other hand, it is possible that although SMA provides useful rehearsal mechanisms during auditory imagery tasks, it may not be an area that *must* be involved in such tasks. The task used here required retrieval of the tune once a title was given, but only the first two notes had to be retrieved. Perhaps SMA is more useful when auditory imagery tasks that extend more in time are required, such as 'scanning' between two notes several beats apart[14] or 'playing' several seconds of a familiar tune inside one's head.[21] Pilot versions of the TMS task, in fact, asked subjects to compare the first and third notes, which might involve SMA to a greater extent; this could be tried in the future (although accuracy rates would be lower, making interpretation of reaction times more problematic).

## Conclusions and directions

The studies just reviewed seem to converge on the general idea that parts of the cortex specialized for processing actual sound are also recruited to process imagined sound. Furthermore, the particular structures processing imagined music bear some similarity to those processing heard music. Specifically, the right temporal lobe has been shown to be active in, and necessary for, adequate performance on both perceived and imagined pitch comparison tasks derived from previously known music. This structure may be especially important for processing music that does not have words, when the left temporal structures cannot be recruited for verbal processing.

The role of the SMA remains unclear, in several respects. Because of our failure to disrupt imagery processing with TMS inhibition, we are not sure if a fully functioning SMA is necessary for good performance on musical imagery tasks. The strong activation of SMA shown by PET in both verbal and nonverbal auditory imagery tasks suggests that 'internal humming' may at least be helpful to maintain an auditory image over time.

This latter point can be subjected to test by devising auditory imagery tasks in which internal humming would not be helpful or even possible to support the task. One domain to investigate would be imagery for environmental sounds. Intons-Peterson[27] has shown that people can generate images of everything from volcanoes erupting to wind chimes tinkling, and make mental comparisons on subjective loudness of these sounds. As people cannot possibly produce most of these sounds, it would be useful to find a task that shows activation in secondary auditory areas (which we can, by now, use as a tentative correlate of the subjective state of auditory imagery) and see if SMA is also activated.

An alternative domain to explore might be that of timbre. Crowder[28] demonstrated in several behavioural paradigms that people asked to imagine timbres of different instruments can do so. He gave people a sine wave of a particular pitch and asked them to imagine it in an instrument timbre. A second tone was then presented in an actual timbre at the same or a different pitch and subjects had to say if the pitch was the same or different as the first note. People were slower to confirm that two pitches were the same if the imagined and perceived timbres did not match. As people cannot produce the sounds of guitars and clarinets, SMA support should not be necessary if it is providing a motor rehearsal program. Timbre imagery tasks would also have the advantage of providing us a lateralization hypothesis, as Samson and Zatorre[11] showed the importance of the right temporal lobe in timbre tasks. Thus brain imaging studies using either PET or functional magnetic resonance imaging in these domains should help us clarify the role of the SMA in auditory imagery tasks.

## Acknowledgements

I thank the National Science Foundation and the McDonnell-Pew Program in Cognitive Neuroscience for support of this research.

## References

1. **Halpern, A. R.** (1988) Perceived and imagined tempos of familiar songs. *Music Percept.* 6, 193–202.
2. **Halpern, A. R.** (1992) Musical aspects of auditory imagery. In D. Reisberg (ed.) *Auditory Imagery.* Hillsdale, NJ: Lawrence Erlbaum, pp. 1–27.
3. **Halpern, A. R.** (1989) Memory for the absolute pitch of familiar songs. *Mem. Cognit.* 17, 572–81.
4. **Kosslyn, S. M., T. M. Ball,** and **B. J. Reiser** (1978) Visual images preserve metric spatial information: evidence from studies of image scanning. *J. Exp. Psychol. Hum. Percept. Perform.* 4, 47–60.
5. **Halpern, A. R.** (1988) Mental scanning in auditory imagery for tunes. *J. Exp. Psychol. Learn. Mem. Cognit.* 14, 434–43.
6. **Intons-Peterson, M. J.** (1983) Imagery paradigms: how vulnerable are they to experimenters' expectations? *J. Exp. Psychol. Learn. Mem. Cognit.* 10, 699–715.

7. **Farah, M. J.** (1988) Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychol. Rev.* 95, 307–17.
8. **Kosslyn, S. M., N. M. Alpert, W. L. Thompson *et al.*** (1993) Visual mental imagery activates topographically organized visual cortex: PET investigations. *J. Cognit. Neurosci.* 5, 263–87.
9. **Zatorre, R. J.** (1985) Discrimination and recognition of tonal melodies after unilateral cerebral excisions. *Neuropsychologia* 23, 31–41.
10. **Zatorre, R. J.** (1988) Pitch perception of complex tones and human temporal-lobe function. *J. Acoust. Soc. Am.* 84, 566–72.
11. **Samson, S. and R. J. Zatorre.** (1994) Contribution of the right temporal lobe to musical timbre discrimination. *Neuropsychologia* 32, 231–40.
12. **Zatorre, R. J. and S. Samson.** (1991) Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain* 114, 2403–17.
13. **Zatorre, R. J. and A. R. Halpern.** (1993) Effect of unilateral temporal-lobe excision on perception and imagery of songs. *Neuropsychologia* 31, 221–32.
14. **Zatorre, R. J., A. R. Halpern, D. W. Perry *et al.*** (1996) Hearing in the mind's ear: a PET investigation of musical imagery and perception. *J. Cognit. Neurosci.* 8, 29–46.
15. **Zatorre, R. J., A. C. Evans, and E. Meyer** (1994) Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 14, 1908–19.
16. **Zatorre, R. J., A. C. Evans, E. Meyer, and A. Gjedde** (1992) Lateralization of phonetic and pitch processing in speech perception. *Science* 256, 846–9.
17. **Smith, E. E., J. Jonides, and R. A. Koeppe** (1996) Dissociating verbal and spatial working memory using PET. *Cereb. Cortex* 6, 11–20.
18. **Gabrieli, J. D. E., J. E. Desmond, J. B. Demb *et al.*** (1996) Functional magnetic resonance imaging of semantic memory processes in the frontal lobes. *Psychol. Sci.* 7, 278–83.
19. **Paulesu, E., C. D. Frith, and R. S. J. Frackowiak** (1993) The neural correlates of the verbal component of short-term memory. *Nature* 362, 342–5.
20. **Perry, D., M. Petrides, B. Alivasatos *et al.*** (1993) Functional activation of human frontal cortex during tonal working memory tasks. *Soc. Neurosci. Abst.* 19, 843.
21. **Halpern, A. R. and R. J. Zatorre** (1999) When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. *Cereb. Cortex* 9, 697–704.
22. **Nyberg, L., R. Cabeza, and E. Tulving** (1996) PET studies of encoding and retrieval. *Psychol. Bull. Rev.* 3, 135–48.
23. **Walsh, V. and M. Rushworth** (1999) A primer of magnetic stimulation as a tool for neuropsychology. *Neuropsychologia* 37, 125–35.
24. **Pridmore, S. and R. Belmaker** (1999) Transcranial magnetic stimulation in the treatment of psychiatric disorders. *Psychiatry Clin. Neurosci.* 53, 541–8.
25. **Kosslyn, S. M., A. Pascual-Leone, O. Felician *et al.*** (1999) The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science* 284, 167–9.
26. **Smith, J. D., M. Wilson, and D. Reisberg** (1995) The role of subvocalization in auditory imagery. *Neuropsychologia* 33, 1433–54.
27. **Intons-Peterson, M. J.** (1980) The role of loudness in auditory imagery. *Mem. Cognit.* 8, 385–93.
28. **Crowder, R. G.** (1989) Imagery for musical timbre. *J. Exp. Psychol. Hum. Percept. Perform.* 15, 472–8.

CHAPTER 16

# NEURAL SPECIALIZATIONS FOR TONAL PROCESSING

ROBERT J. ZATORRE

## Abstract

The processing of pitch, a central aspect of music perception, is neurally dissociable from other perceptual functions. Studies using behavioural-lesion techniques as well as brain imaging methods demonstrate that tonal processing recruits mechanisms in areas of the right auditory cortex. Specifically, the right primary auditory area appears to be crucial for fine-grained representation of pitch information. Processing of pitch patterns, such as occurs in melodies, requires higher-order cortical areas, and interactions with the frontal cortex. The latter are likely related to tonal working memory functions that are necessary for the on-line maintenance and encoding of tonal patterns. One hypothesis that may explain why right-hemisphere auditory cortices seem to be so important to tonal processing is that left auditory regions are better suited for rapidly changing broad-band stimuli, such as speech, whereas the right auditory cortex may be specialized for slower narrow-band stimuli, such as tonal patterns. Evidence favouring this hypothesis was obtained in a functional imaging study in which spectral and temporal parameters were varied independently. The hypothesis also receives support from structural studies of the auditory cortex, which indicate that spectral and temporal processing may depend on interhemispheric differences in grey/white matter distribution and other anatomical features.

**Keywords:** Hemispheric functional specialization; Pitch processing; Pitch patterns

How does the nervous system analyze and make sense of information coming from the environment? This central question in contemporary neuroscience has been addressed from many perspectives in the last two decades. In the auditory domain, a great deal of effort has gone into understanding the way in which the human brain processes speech sounds. Given the importance of speech and spoken language to human communication and behaviour, this is not surprising. It is interesting, however, to consider other aspects of auditory information processing that may help to round out the picture. Increasingly, music is being recognized as an important component of human activity that may help us to gain insight into the functional organization of the human brain. It may be argued that the perception, encoding, and reproduction of musical sounds requires neural mechanisms that are at least as complex as those for speech; indeed, it seems likely that speech and music must engage the most cognitively demanding aspects of auditory processing. It may further

be argued that these two classes of stimuli are uniquely human, making it particularly relevant to study them in the context of understanding those aspects of brain organization that are most relevant to human cognition and behaviour. Finally, an overarching theme of chapters presented in this volume is that music can and should be viewed as a biological phenomenon, not merely a cultural artifact; from this perspective, therefore, it is not only meaningful but even imperative to seek out the neural processes that may allow human beings to engage in the diverse array of behaviours we term music.

All of these considerations set the stage for the studies from my laboratory and from others that will be described in this chapter. The recent development of brain imaging techniques, coupled with more traditional behavioural approaches, has allowed for rapid developments in the field. A great deal of closely related information on music and brain function is presented in many of the other chapters in this volume, particularly in those by Besson, Griffiths, Halpern, Pantev, Parsons, Peretz, Samson, Schlaug, and Tervaniemi. The studies to be discussed in the present chapter focus specifically on aspects of tonal processing and their neural substrates. The underlying theme of the research approach is that by studying these processes we will achieve a more complete understanding not only of the auditory nervous system, but of the human brain in general. The idea that neural systems may display functional specializations is generally well accepted in cognitive neuroscience. The work described in this chapter therefore also has a goal of identifying these specializations in order to determine the extent to which musical processes depend on dedicated neural circuitry or may form part of shared neural architecture. In this context, another important feature of brain organization, hemispheric functional specialization, is also crucial, and will be important in understanding the pattern that emerges.

In order to make sense of a complex phenomenon such as music, it is necessary, at least as a starting point, to focus on a specific aspect; in the present set of studies, I have chosen to explore the processing of pitch, which is to say pitch information as used for musically relevant processes. This choice is motivated by the fact that pitch appears to be a central aspect of all music. Although other aspects of music, such as temporal organization for rhythm, for example, may play an equally important role, it is difficult to conceive of a musical system of any type that does not involve the patterning of pitches. Moreover, pitch processing is amenable to study not only because the physical parameters of sounds associated with pitch are relatively well understood and easily manipulated, but also because it affords us the opportunity to analyse different levels of processing. Thus, in what follows I will consider first of all studies examining relatively low-level aspects of pitch processing, such as pitch discrimination; second I will discuss higher-order aspects, including pitch patterns. Finally, I will present some anatomical findings that may be relevant to understanding the nature of the functional specializations that the other studies have uncovered.

## Basic aspects of pitch processing

To begin, let us consider a relatively simple aspect of a musically relevant function: the ability to distinguish sounds on the basis of their pitch. The neural basis for the encoding of pitch has been studied for some time. Indeed, much neurophysiological work both in the periphery and in the central nervous system has examined the processing of frequency information

(see Chapter 9, this volume). One of the salient features of the auditory nervous system, in fact, is that a tonotopic organization exists from the earliest level of the periphery, at the basilar membrane, to many fields within the auditory cortex. This topographic organization along the frequency axis points to the importance of pitch information in auditory processing generally. However, does the cortex carry out some essential function in the representation of pitch information? Several classic studies in experimental animals with bilateral destruction of auditory cortical areas have suggested that simple pitch discrimination remains unaffected, indicating that earlier levels of processing, in the midbrain or thalamus, for example, may be sufficient.[1–3] However, the nature of the pitch processing task appears to be critical in answering this question.

In a recent study from my lab, Johnsrude *et al.*[4] explored the specificity of pitch discrimination using a behavioural-lesion technique. We tested patients who had undergone surgical excision in the auditory cortex in the right or left temporal lobe; the patients' lesions were classified according to magnetic resonance imaging (MRI) as extending or not into Heschl's gyrus (HG), which contains the primary auditory cortex (see Ref. 5 for description of a lesion quantification procedure, and Figures 16.1 and 16.2). A simple psychophysical 'staircase' discrimination task was used, with pairs of pure tones to determine the threshold or minimal frequency separation needed to achieve a certain level of performance. Two separate tasks were administered. In the first, the two tones were either identical or different in pitch, and the subject answered accordingly. In the second task, the

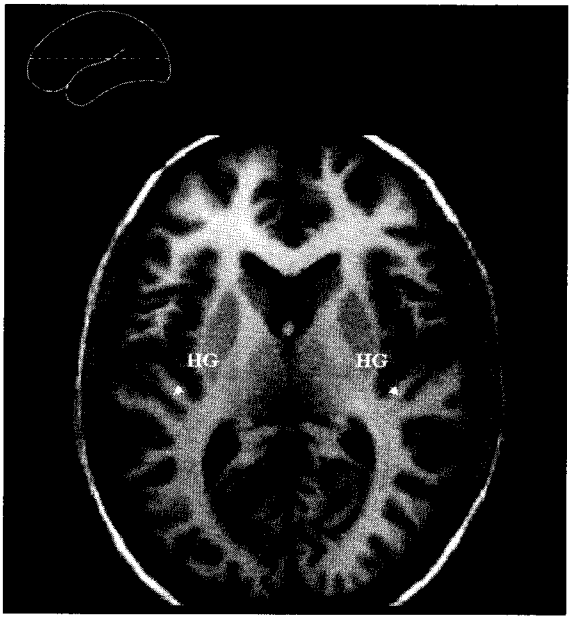



Figure 16.1 Magnetic resonance image of a normal human brain showing a horizontal section through the region of Heschl's gyrus, marked HG. The orientation and level of the horizontal section is indicated by the *dashed line* in the *inset*.

(see Chapter 9, this volume). One of the salient features of the auditory nervous system, in fact, is that a tonotopic organization exists from the earliest level of the periphery, at the basilar membrane, to many fields within the auditory cortex. This topographic organization along the frequency axis points to the importance of pitch information in auditory processing generally. However, does the cortex carry out some essential function in the representation of pitch information? Several classic studies in experimental animals with bilateral destruction of auditory cortical areas have suggested that simple pitch discrimination remains unaffected, indicating that earlier levels of processing, in the midbrain or thalamus, for example, may be sufficient.[1–3] However, the nature of the pitch processing task appears to be critical in answering this question.

In a recent study from my lab, Johnsrude *et al.*[4] explored the specificity of pitch discrimination using a behavioural-lesion technique. We tested patients who had undergone surgical excision in the auditory cortex in the right or left temporal lobe; the patients' lesions were classified according to magnetic resonance imaging (MRI) as extending or not into Heschl's gyrus (HG), which contains the primary auditory cortex (see Ref. 5 for description of a lesion quantification procedure, and Figures 16.1 and 16.2). A simple psychophysical 'staircase' discrimination task was used, with pairs of pure tones to determine the threshold or minimal frequency separation needed to achieve a certain level of performance. Two separate tasks were administered. In the first, the two tones were either identical or different in pitch, and the subject answered accordingly. In the second task, the

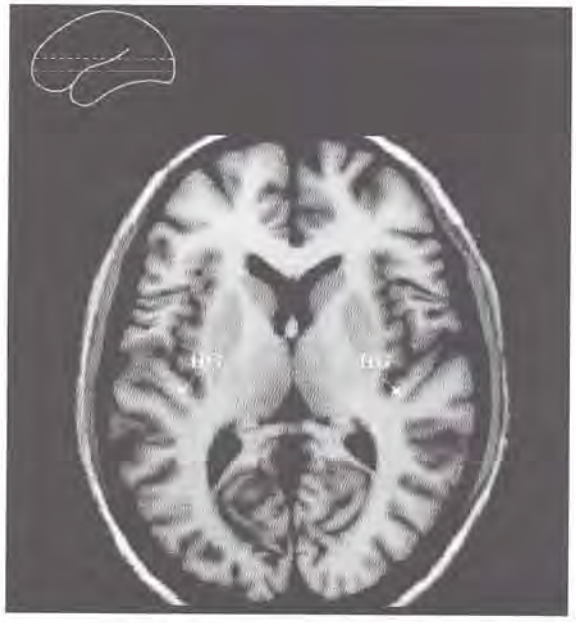


**Figure 16.1** Magnetic resonance image of a normal human brain showing a horizontal section through the region of Heschl's gyrus, marked HG. The orientation and level of the horizontal section is indicated by the *dashed line* in the *inset*.

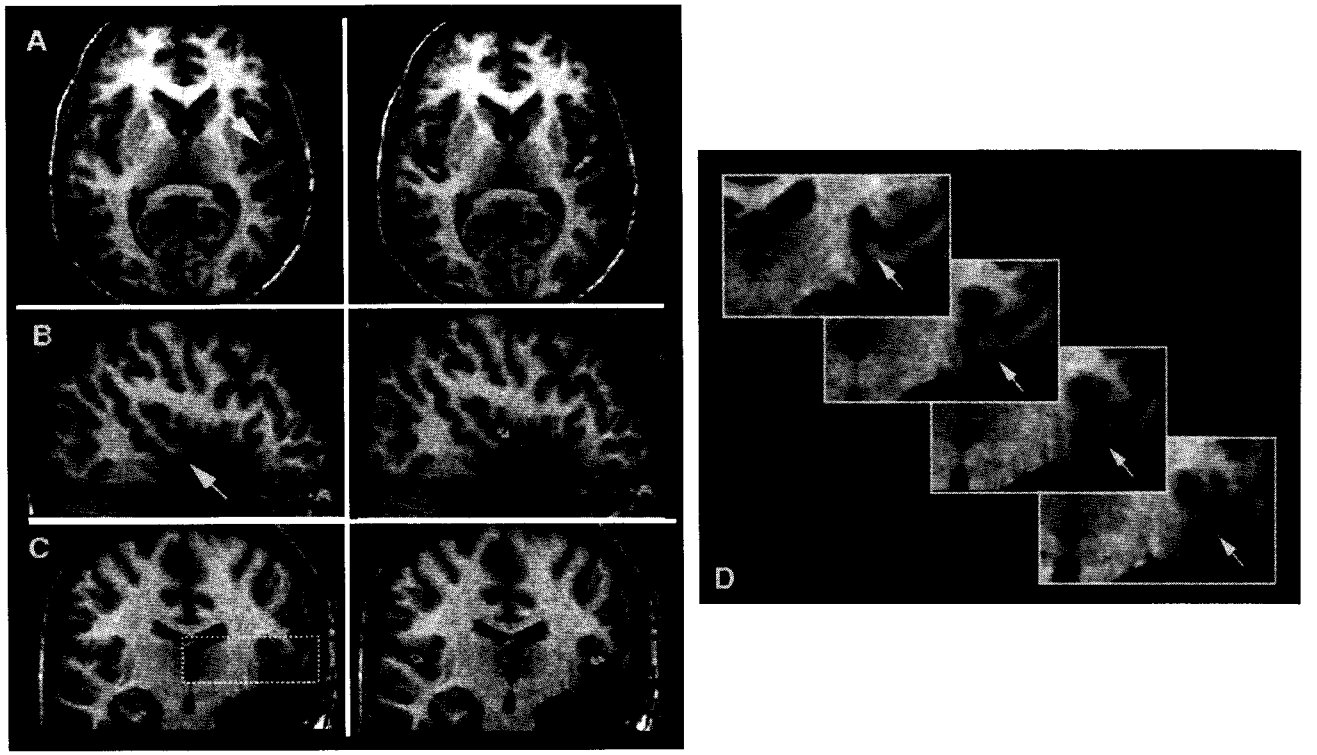



**Figure 16.2** MRI scan of a patient with a surgical excision in the right HG; illustrated are planes of section oriented horizontally (**A**), sagittally (**B**), and coronally (**C**). The left side of panels **A**, **B**, and C show the patient's scan alone, with an *arrow* indicating the region of excision/undercutting. The images on the *right side* of panels **A**, **B**, and C show the patient's scan coregistered with an anatomical probabilistic map of HG derived from normal individuals (Ref. 44). The *crosshairs* indicate the same position in the standardized space as the *arrow*. Note the correspondence between the position of HG as determined from the map and the patient's partially excised HG region. The *yellow box* in panel C indicates the region of the removal pictured in close-up in panel D, which illustrates the transition from intact, to undercut, to fully excised tissue (coronal sections taken at 3-mm intervals, posterior to anterior). *Arrows* again correspond to the *crosshairs* in the other panels and indicate the location of the HG region. (See Plate 4A–D in colour section.)

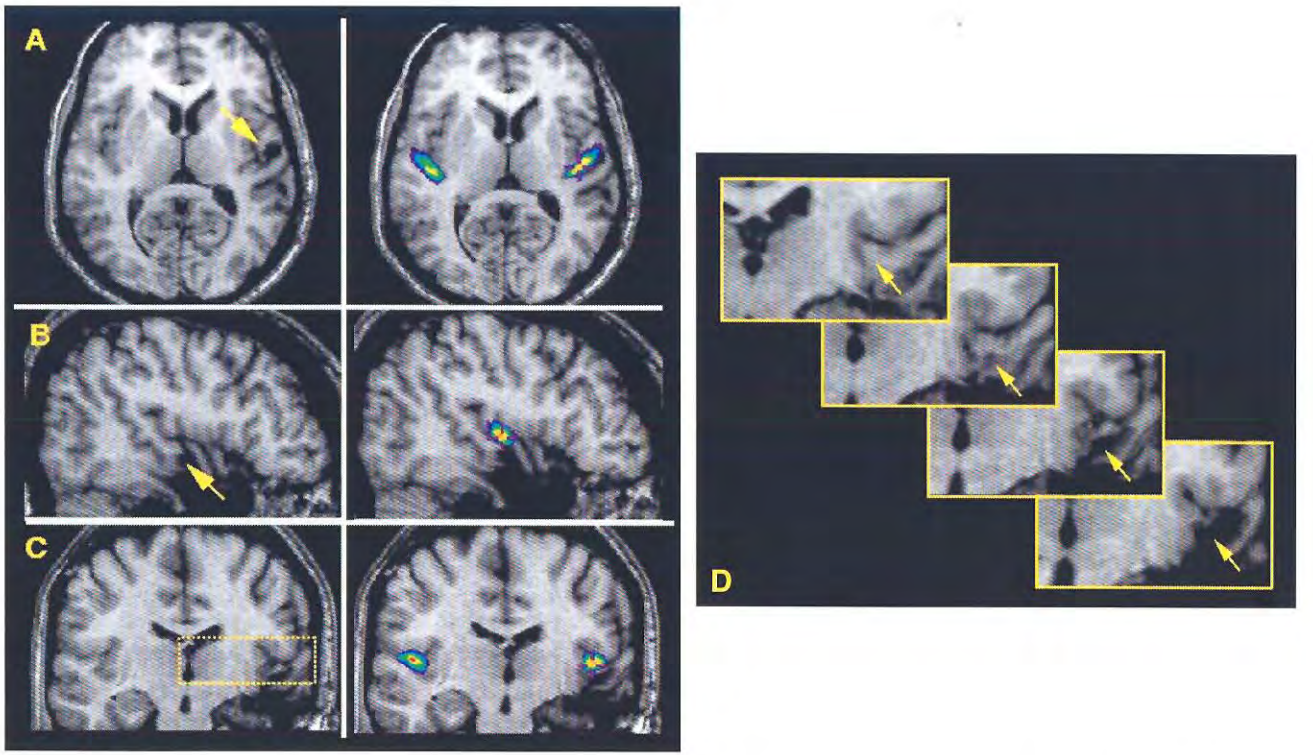



**Plate 4** MRI scan of a patient with a surgical excision in the right HG; illustrated are planes of section oriented horizontally (A), sagittally (B), and coronally (C). The left side of panels A, B, and C show the patient's scan alone, with an *arrow* indicating the region of excision/undercutting. The images on the *right side* of panels A, B, and C show the patient's scan coregistered with an anatomical probabilistic map of HG derived from normal individuals (Ref. 44). The *crosshairs* indicate the same position in the standardized space as the *arrow*. Note the correspondence between the position of HG as determined from the map and the patient's partially excised HG region. The *yellow box* in panel C indicates the region of the removal pictured in close-up in panel D, which illustrates the transition from intact, to undercut, to fully excised tissue (coronal sections taken at 3-mm intervals, posterior to anterior). *Arrows* again correspond to the *crosshairs* in the other panels and indicate the location of the HG region.

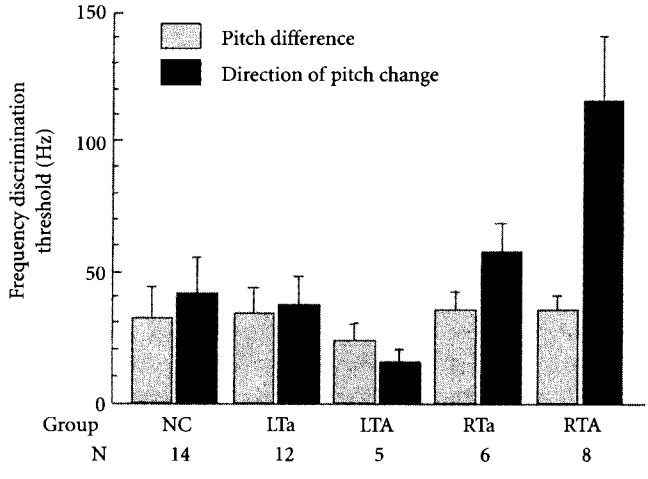



**Figure 16.3** Mean frequency difference discrimination thresholds observed in two pitch discrimination tasks in patients with right or left (RT or LT) temporal-lobe excision, extending into the region of HG (A) or sparing this area (a). The simple discrimination thresholds were not different across the groups. Thresholds in the direction-of-pitch-change task were significantly higher in the RTA group than in any of the other groups. (Reproduced with permission from Ref. 4).

two tones always differed, and the subject's task was to indicate if the second tone was higher or lower than the first. The tones themselves were identical across the two tasks. The idea here is that whereas the simple pitch discrimination task merely requires that a difference be noted, the second, pitch-direction task requires that the tones be discriminated and ordered along some underlying pitch scale.

The findings (Figure 16.3) indicated several interesting dissociations. First, the patients with lesions within the left temporal lobe were quite unimpaired as compared to control participants on either task, regardless of the extent of excision with HG. Second, the patients with damage to the right temporal lobe but excluding HG also had normal thresholds. The group of subjects with right HG damage, however, showed a fourfold increase in their pitch thresholds, but only for the pitch-direction task; they were normal on the simple pitch discrimination task. The findings are thus quite specific in indicating that only a certain restricted lesion in the right temporal lobe has an effect, and then only on a particular task. Thus, one may conclude that the right primary auditory area plays a special role not simply in discriminating one pitch from another, but in some aspect of organizing the sounds according to their pitch. This aspect of pitch organization, it is argued, is essential for any musical use of pitch information. Interestingly, all but one of the patients in question were able to perform the task itself; it was only the threshold that was abnormally high. This aspect of the result suggests that a much coarser level of pitch organization still exists in these patients, which may be mediated by left auditory cortex, a point to which we shall return below.

These results are consistent with those of a much earlier study[6] in which patients with right temporal lobe excisions with encroachment onto Heschl's gyri had significant difficulties with a missing-fundamental pitch discrimination task but not with the identical

task when the fundamental frequency was present. Patients with more anterior right temporal lobe damage, or those with similar damage on the left were unimpaired. Both studies thus point to a specificity for basic aspects of pitch processing, indicating a special role for areas in or adjacent to the right primary auditory cortex in these functions. Several other studies have added weight to this idea, although the degree of anatomical specificity has not always been possible; nonetheless, many authors have observed that the right temporal cortex plays a particularly important role for processing of pitch and of spectral pattern information.[7–11]

More recently, functional imaging studies have begun to add important new evidence to our emerging view of the functional organization of the human brain. Among the many advantages of techniques, such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), is that they provide a noninvasive means of obtaining anatomically accurate information about the cerebral areas that change their neural activity as a function of stimulus or task parameters. Thus, these techniques provide critical complementary evidence to the data derived from more traditional behavioural studies. However, whereas activation studies can identify the sites of neural activity associated with a particular function, they cannot determine the degree to which necessary computations are carried out. Conversely, only lesion techniques or stimulation techniques (such as transcranial magnetic stimulation, see, Chapter 26, this volume) can provide information about essential areas, but they are constrained by various other limitations. Clearly, whenever converging evidence can be obtained from a variety of methods, the conclusions that can be drawn will be correspondingly stronger.

Several recent studies have examined pitch processing using functional imaging techniques. One relevant piece of evidence in favour of a special role for the right auditory cortex in pitch processing comes from PET studies of pitch production. Perry *et al.*[12] compared a condition in which subjects repeatedly produced a single vocal tone of constant pitch to a control in which they heard synthetic tones at the same rate. Increased cerebral blood flow (CBF) was noted in several motor areas, but within the superior temporal gyrus (STG), only right-sided activations were noted in both primary and secondary auditory cortices. The authors interpret this finding as a reflection of the specialization of the right auditory cortex for pitch perception. They argue that accurate perception of the pitch of one's own voice is required during singing in order to use feedback to maintain the desired pitch. Additional findings in favour of this conclusion come from regression analyses performed on these data, in which the accuracy of the subjects' singing was measured, and the degree of pitch excursion was covaried against CBF in the whole brain volume.[13] Strikingly, a region of covariation was identified in the right HG, such that increased CBF was observed with increasing pitch deviation. This result confirms the special role played by this region in pitch processing and suggests that a feedback mechanism may indeed be operative to maintain consistency of vocalized pitch.

## Processing of pitch patterns

The initial phase of pitch analysis set the stage for processing of pitch patterns. The relationships between individual tones form the basis for the most fundamental aspect of musical

processing: the encoding and recognition of melodies. Melodic patterns, in turn, may be investigated at many different levels of analysis, from the basic interval relationships between tones to the semantic networks that are important for long-term representations of familiar tunes, and the output mechanisms necessary for singing or playing them. Here, we shall concentrate on the way in which auditory cortical systems are engaged in perception of melodies. One critical consideration in this respect is that all melodic perception (indeed, all auditory perception) involves working memory mechanisms. Since sounds unfold over time, in order to compute the relationships between successive events, an on-line retention system is evidently necessary. In the case of even a brief melody, this would mean that relationships between tones have to be computed and maintained over periods ranging from seconds to minutes in order to achieve a relatively stable and coherent internal representation.

Many studies of melodic perception with brain-lesioned individuals have demonstrated that damage to the right STG often leads to behavioural deficits in discrimination tasks,[7,14–16] although most of these studies have also noted that damage to the equivalent structures on the left may also lead to some degree of impairment. Indeed, Peretz *et al.*[17] (see also Chapter 13, this volume) have noted that patients with bilateral lesions that involve portions of the STG and frontal cortical structures demonstrate much more severe disturbances in melodic tasks than most patients with unilateral damage. In fact, in each of the cases that Peretz and her colleagues have studied, the patients did not appear to show the specific musical deficit until after a second lesion had occurred, implying that unilateral damage had had at most only a mild effect not noted by the patients or their families. This point is important in that it suggests that the putative hemispheric specializations that are one theme of this chapter are likely to be relative rather than absolute.

The relative importance of right-hemisphere STG mechanisms has been demonstrated specifically for tonal working memory in two studies from our laboratory. In one,[18] we adapted a paradigm first developed by Deutsch[19] in which a target and a comparison tone are to be compared for pitch; the tones are either separated by a brief silent interval or by a series of interfering random tones. Deutsch had shown that memory for tones was relatively specific because it was not disrupted by other sounds, but only by other tones (see also Ref. 20). When this tonal working-memory task was administered to temporal-lobectomy patients, performance was significantly worse following damage to the right than to the left temporal cortex for the interference task. On the other hand, unlike the studies of basic pitch processing discussed above, there was no difference in degree of behavioural deficit as a function of encroachment onto HG. Thus, areas of auditory cortex outside of the primary zone are most likely to be involved in this function. Finally, it was also noted that right frontal cortical excisions produced an impairment similar to that seen in the right temporal patients, a point that is taken up below.

To follow up on the question of tonal working memory function, Zatorre *et al.*[21] used PET to test 12 volunteer nonmusicians with unfamiliar melodies and noise bursts that had been constructed so as to approximate the acoustic characteristics of the melodies. Four conditions were run; subjects were asked to (1) listen to the noise bursts; (2) listen to the melodies without explicit instruction; (3) listen to the same melodies and compare the pitch of the first and second notes (two-note condition); and (4) listen to the melodies and compare the pitch of the first and last notes (first/last condition).

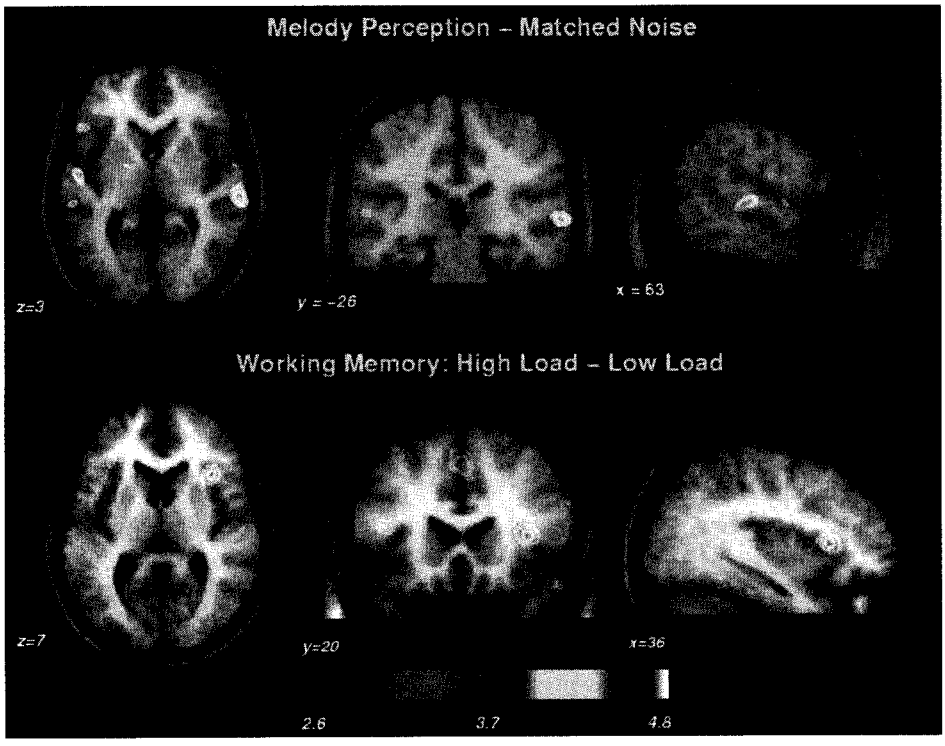



**Figure 16.4** Selected PET CBF activation sites associated with processing of tonal melodies. The figures show averaged PET subtraction images superimposed on a corresponding section of an average anatomical MRI. (*Top*) Horizontal, coronal, and sagittal views through a focus located in the right superior temporal gyrus representing a significant blood flow increase while subjects listened to a series of short, unfamiliar tonal melodies, as compared with a baseline condition in which they listened to noise bursts that were acoustically equated to the melodies. Also visible in the horizontal section is a region of significant activity in the left temporal cortex. (*Bottom*) Horizontal, coronal, and sagittal views through a focus in the right frontal/opercular region showing activity associated with making judgements about the pitch of the first and last notes of the tonal sequences as compared to making judgements about the first two notes. Data reanalysed from Zatorre *et al.*[21] (See Plate 5 in colour section.)

The comparison of listening to melodies minus noise permits examination of the cerebral regions specifically active during processing of novel melodies. The principal result indicated a significant CBF increase in the right superior temporal gyrus, anterior to the primary auditory cortex; a weaker CBF increase was also observed in the left STG. A reanalysis of these data with a smaller reconstruction filter and improved image registration[22] indicates that the right STG activation is in the ventrolateral aspect of the gyrus (Figure 16.4). These findings point to differential activation of primary vs secondary auditory areas within the STG, depending on the stimulus features. Caution must still be exercised in interpreting the results, for the noise bursts clearly differ from the melodies on a number of dimensions (e.g. periodicity, frequency modulation, and spectral composition). It remains to be established, therefore, which specific features of the melodies may lead to the observed pattern of

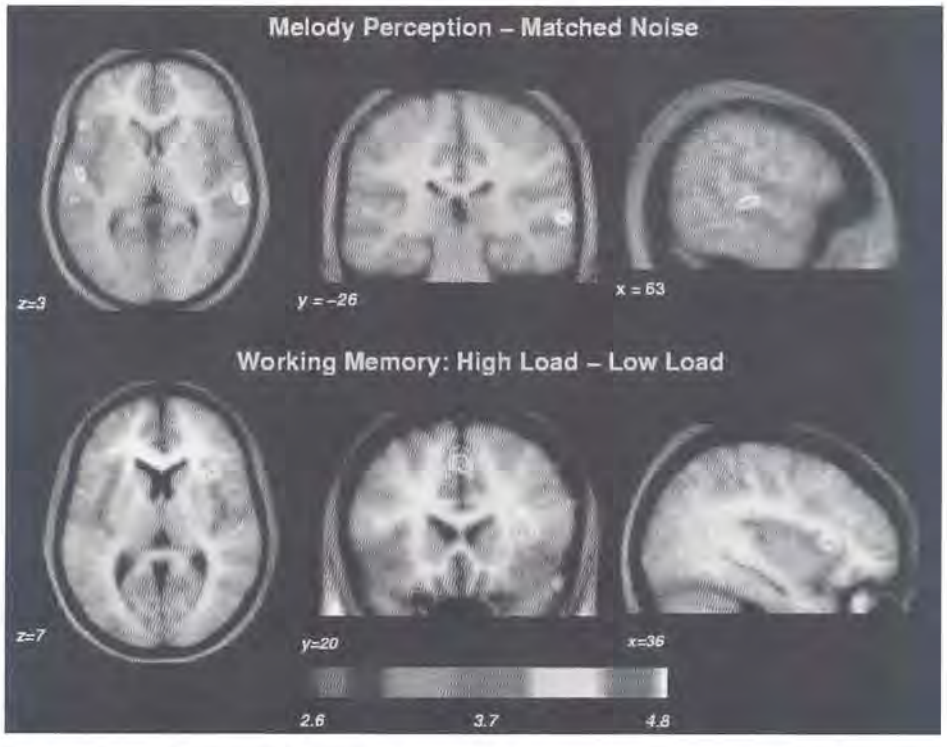



**Figure 16.4** Selected PET CBF activation sites associated with processing of tonal melodies. The figures show averaged PET subtraction images superimposed on a corresponding section of an average anatomical MRI. (*Top*) Horizontal, coronal, and sagittal views through a focus located in the right superior temporal gyrus representing a significant blood flow increase while subjects listened to a series of short, unfamiliar tonal melodies, as compared with a baseline condition in which they listened to noise bursts that were acoustically equated to the melodies. Also visible in the horizontal section is a region of significant activity in the left temporal cortex. (*Bottom*) Horizontal, coronal, and sagittal views through a focus in the right frontal/opercular region showing activity associated with making judgements about the pitch of the first and last notes of the tonal sequences as compared to making judgements about the first two notes. Data reanalysed from Zatorre *et al.*[21] (See Plate 5 in colour section.)

The comparison of listening to melodies minus noise permits examination of the cerebral regions specifically active during processing of novel melodies. The principal result indicated a significant CBF increase in the right superior temporal gyrus, anterior to the primary auditory cortex; a weaker CBF increase was also observed in the left STG. A reanalysis of these data with a smaller reconstruction filter and improved image registration[22] indicates that the right STG activation is in the ventrolateral aspect of the gyrus (Figure 16.4). These findings point to differential activation of primary vs secondary auditory areas within the STG, depending on the stimulus features. Caution must still be exercised in interpreting the results, for the noise bursts clearly differ from the melodies on a number of dimensions (e.g. periodicity, frequency modulation, and spectral composition). It remains to be established, therefore, which specific features of the melodies may lead to the observed pattern of

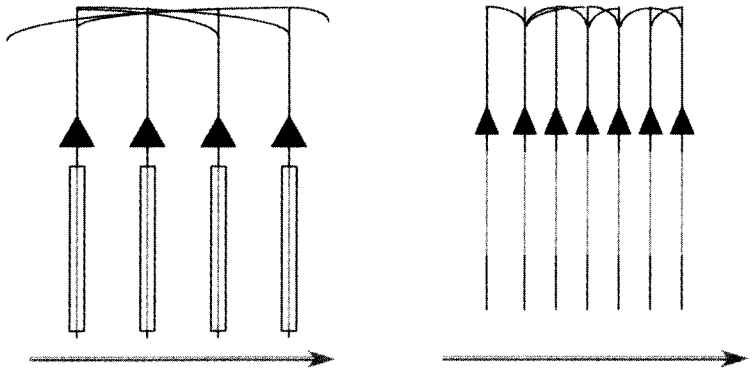


Figure 16.5 Cartoon depiction of putative differences in anatomical structure between left and right auditory cortical areas. *Triangles* and *lines* represent individual neurons with differing degrees of myelination (thicker on the left). Neurons on the left are larger than those on the right, have wider spacing across tonotopically organized cortex, and are more widely interconnected. These structural features would lead to relatively fast conduction but poor spectral resolution on the left, and to poorer temporal resolution but greater sensitivity to fine spectral differences on the right.

activation. The finding of an activation within the right STG while listening to melodies fits in well with the lesion data, however, and likely reflects the specialization of neuronal networks within the right secondary auditory cortices for perceptual analysis of tonal information.

Zatorre *et al.*[21] also examined working memory for pitch by subtracting the passive condition from the two-note condition, which resulted in significant activation within the right frontal lobe (both inferiorly and superiorly). The first/last minus passive melodies subtraction yielded a large number of cortical and subcortical activation sites in both hemispheres, including a focus in the right inferior frontal cortex identical to that observed in the two-note condition. A comparison not reported in the original study pits the first/last condition against the two-note condition. The advantage of this contrast is that both involve pitch comparisons, the main difference being the degree of working memory load. The interesting finding here (Figure 16.4) is once again an area of increased blood flow in the right inferior frontal region, though not in the same location as in the other subtractions. The pattern of results from these conditions therefore implicates frontal-lobe mechanisms in effecting pitch comparisons, with a particularly important contribution from right-frontal regions. It should be recalled, however, that in the first/last minus passive melody comparison, a large number of separate foci of CBF change were seen over a wide swath of cortical and subcortical territory; this finding perhaps reflects the complexity and increased cognitive demands of the task, which was also manifested in increased error rate and slower reaction times.

One may speculate that the numerous frontal lobe sites observed might be associated with successful performance of distinct aspects of the task. The findings of Perry *et al.*[23] are directly relevant to this issue; they observed CBF increases with a rightward asymmetry in dorsolateral frontal areas while subjects were engaged in an auditory tonal working memory task that required them to monitor a sequence of tones. This finding suggests a functional dissociation between the cognitive act of monitoring information as opposed to holding sensory

information in working memory. Perry *et al.*[23] suggest a hierarchical view, compatible with other evidence,[24] such that right inferior lateral frontal areas are important for maintenance of tonal information, whereas dorsolateral frontal areas are required for higher-level functions such as monitoring the contents of working memory. This conclusion would be in accord with the lesion study described earlier,[18] as well as with PET studies[25] that also indicated increased activity in the right frontal cortex during pitch judgements on speech syllables.

Several other imaging studies over the past few years have supported the idea that the right secondary auditory cortex may be particularly important in processing tonal patterns (but see Chapter 10, this volume, for evidence that the story is more complicated).The results of Démonet *et al.*[26,27] are relevant, since tone judgement tasks were used, in which subjects were required to detect pitch changes of tones within a sequence, and the difference in activity pattern between the tonal task and a language task was examined. The findings were consistent with data presented above, implicating the right auditory cortical areas in tonal processes, since foci were observed in posterior right STG and middle temporal areas. More recently, Binder *et al.*[28] reported whole-brain fMRI data using an active tone judgement task. This task, which required subjects to discriminate two pitches during a series of tones (which ranged from 3 to 7 tones in length) and then to make a judgement pertaining to the entire series of tones, emphasized not only pitch perception but also maintenance of simple pitch information in working memory. When this task was compared to an active task involving semantic categorization of words, stronger activation associated with the tone task was observed in the posterior STG bilaterally, the supramarginal region bilaterally (but much more on the right), the right posterior middle temporal gyrus, premotor cortex bilaterally (also much more on the right), and right superior parietal cortex.[28] These right temporoparietal foci are in agreement with the right posterior STG, middle temporal, and inferior parietal foci reported by Démonet in their tone-speech comparisons.[26,27]

A further source of relevant evidence comes from studies of musical imagery (reviewed by Halpern, Chapter 15, this volume). Those studies, including both behavioural-lesion experiments and functional imaging,[29–31] have indicated that a greater role is played by secondary auditory cortical areas on the right than on the left not only in perceptual tasks involving judgements of real melodies, but also in imaginal tasks in which subjects are invited to imagine a well-known tune in 'the mind's ear'. These studies, together with additional recent neuroimaging findings pointing in the same general direction,[32–34] provide additional evidence for a special role of several right-hemispheric cortical regions in various aspects of tonal pattern processing.

## Hemispheric differences in spectral processing

Much of the evidence presented above converges towards the conclusion that hemispheric specialization exists in the processing of pitch information, with right auditory cortical areas playing a more prominent role than left. Moving beyond the mere fact that differences do exist, it is critical to pose the question of why such lateralization might have evolved. One classic way of thinking about hemispheric differences, favoured, for example, by investigators such as Teuber or Sperry, is in terms of complementary functional specialization—that is, that hemispheric processing advantages emerge because they confer some advantage

over shared processing, since each hemispheric system can specialize in one domain, thus leading to an overall enhancement for the organism as a whole.

If we think about the data presented earlier in this chapter within this context, it seems obvious that the two domains for which the auditory cortices in each hemisphere have become specialized are, roughly, speech and tonal patterns. A bit of additional reflection suggests a unifying hypothesis to explain these complementary functional specializations: speech and tonal stimuli differ in their acoustic structure, and hence in their processing requirements. Whereas the analysis of speech requires good temporal resolution to process rapidly changing energy peaks (formants) that are characteristic of many speech consonants (see, for example, the work of Tallal *et al.*[35]), it can be argued that tonal processes instead require good frequency resolution. In a truly linear system, temporal and spectral resolution are inversely related, so that improving temporal resolution can only come at the expense of degrading spectral resolution and vice versa. This tradeoff naturally arises from a fundamental physical constraint in acoustic processing: better resolution in the frequency domain can be obtained only at the expense of sampling within a longer time window, hence degrading temporal resolution; conversely, high resolution in the temporal domain entails a degraded spectral representation. The auditory nervous system is, of course, a highly nonlinear and distributed system; yet, it may also respect this fundamental computational constraint, such that in the left auditory cortex the high temporal resolution needed to process speech imposes an upper limit on the ability to resolve spectral information, and vice versa for the right auditory cortex. To put it more simply, the hypothesis is that there may be a tradeoff in processing in temporal and spectral domains, and that auditory cortical systems in the two hemispheres have evolved a complementary specialization, with the left having better temporal resolution, and the right better spectral resolution.

According to the foregoing idea, therefore, the hemispheric differences we see in the tonal perception literature reviewed above might reflect this more fundamental level of functional organization. If this hypothesis is correct, then we ought to be able to obtain evidence for differential response within left and right auditory cortices by manipulating temporal and spectral parameters of an auditory stimulus, even if it is not particularly perceived as speech or music. A recent functional imaging study from our lab[36] set out to do just that using a parametric approach. Rather than looking for differences in cerebral blood flow between a control and an activation condition, we examined the functional changes in the brain that correlated with a given input parameter. This approach can be particularly powerful since it helps to isolate brain activity that is specifically related to the parameter of interest. We first created nonverbal stimuli that varied independently and systematically along two dimensions, one temporal, the other spectral. The stimuli were merely a series of pure tones that varied in frequency and duration; in one set of conditions the frequency change was held constant and the temporal rate became faster across scans, while in the other set of conditions the rate was held constant and the frequency differences became finer across scans. We predicted that increasing the rate of temporal change would preferentially recruit left auditory cortical areas, while increasing the number of spectral elements would engage right auditory cortical regions.

The results of greatest relevance for the present discussion were that cerebral blood flow in a region of the left auditory cortex showed a greater response to increasing temporal than spectral variation, whereas a symmetrical area on the right showed the reverse

pattern. A third area in the right superior temporal sulcus also showed a significant response to the spectral parameter but showed no change to the temporal parameter. Thus, the data supported the hypothesis that corresponding regions of the auditory cortex in the two hemispheres have different sensitivity to temporal and spectral information. Neurophysiological recordings in macaque monkeys have shown that auditory cortical neurons are highly sensitive to both spectral and temporal features of sounds simultaneously, and that they exhibit complex response functions.[37–39] It has also been shown that neurons in the belt region of the macaque auditory cortex are optimally responsive to stimuli of specific bandwidths, hence demonstrating sensitivity to a spectral parameter.[40] As well, Eggermont[41] noted that in two primary fields of the cat there was an inverse relation between bandwidth and temporal resolution, consistent with our proposal.

Additional evidence that left auditory cortical units have higher temporal resolution comes from recent depth-electrode recordings within Heschl's gyri[42] showing that the responses within the left auditory cortex encode the voice-onset time of a consonant, whereas the right auditory cortex did not show sensitivity to this temporal parameter. Even more striking consistency is offered by additional data from the same investigators (Chapter 10, this volume), who showed that intracortically recorded auditory evoked potentials were more sharply tuned to frequency in the right auditory cortex than in the left, as predicted by the hypothesis that resolution in the frequency and time domains are different in the two hemispheres.

This hypothesis has also received support from a PET study by Belin *et al.*,[43] in which the rate of formant transitions in a pseudospeech sound—not perceived as speech—was varied from 40 to 200 ms. CBF changes in the left auditory region were found for both types of stimuli, indicating a capacity for processing spectral change over a wide range of durations; by contrast, regions in the right auditory cortex responded only to the slower rates, and not to the faster rate. Thus, the right auditory cortex seems unable to respond to fast formant transitions, whereas the homologous area on the left is better able to track rapidly changing acoustic information that would be relevant for speech processing. Finally, it is also relevant to recall the findings of Robin *et al.*,[9] who studied auditory discrimination in brain-damaged patients. They too found that damage to association cortices in the right hemisphere resulted in spectral but not temporal processing deficits, while the converse was observed after left-hemispheric damage.

## Anatomical considerations

One advantage of the hypothesis presented above is that it offers a unifying framework to understand some of the functional characteristics of the auditory nervous system that allow us to be able to process speech and tonal patterns. This model raises an important question, however: How are these functional differences instantiated in the brain? That is, can we find any evidence in the structure of the auditory cortex that would not only support the existence of processing differences but would also suggest how these differences are implemented?

Some answers to these questions are provided by recent anatomical data from our laboratory[44] and from other investigators. Unlike the functional techniques described above, the aim of this research was to characterize the structural features (shape, volume, and position) of the human primary auditory cortical region *in vivo*. The approach taken was to use MRI scans from groups of normal right-handed volunteer subjects and to label the region of HG

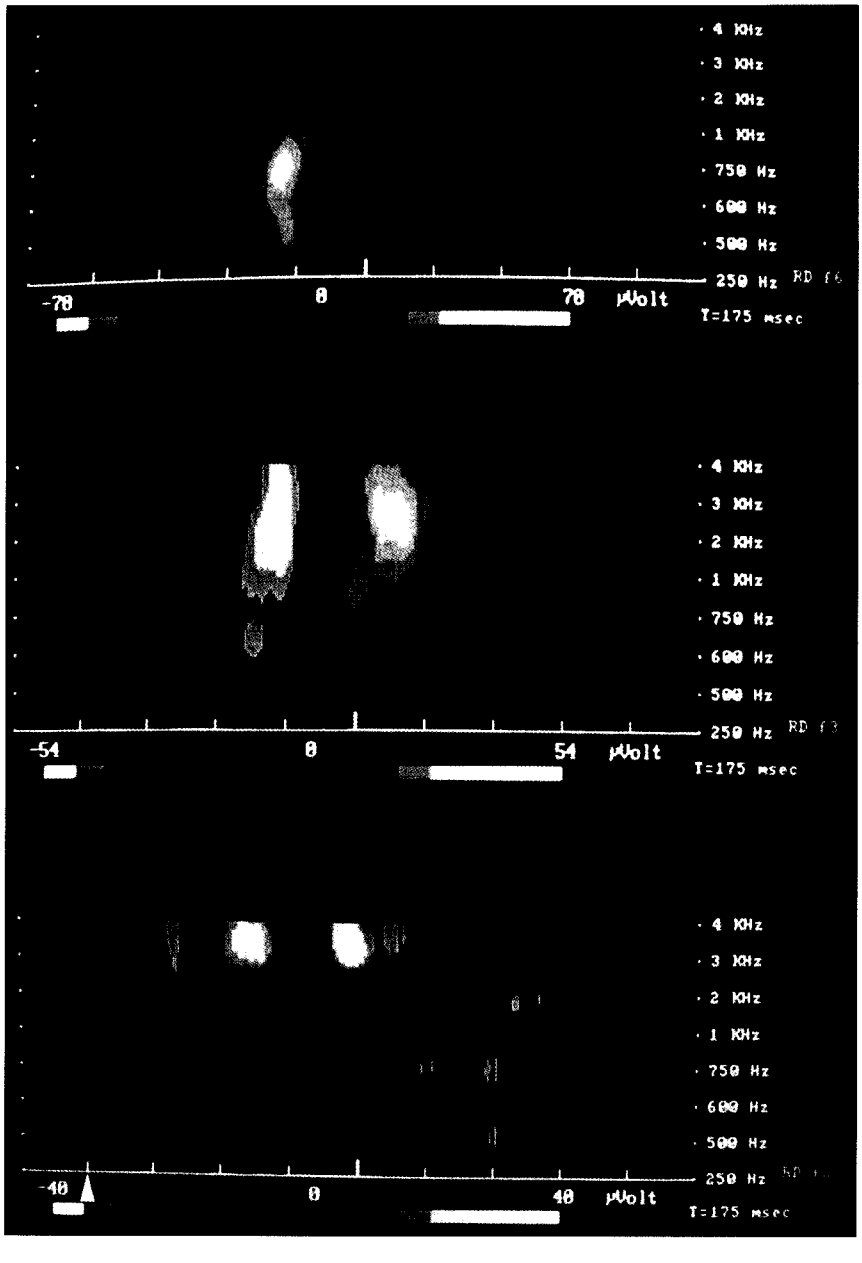



**Plate 1** Temporospectral maps of amplitude variations of potentials recorded from the right primary auditory cortex (leads 2, 3, and 6). Time, related in abscissae, is reduced to 175 ms, full scale. On the ordinate axis, each yellow dot represents the frequency. Amplitude is represented by colour. A cubic spline interpolation is used in order to obtain a continuous potential line between the different frequencies. Each spatial /reconstruction gives a continuous map both in time and in frequency. Dark colors are values close to the neutral potential, and bright colors are used for the activity—blue represents negativity and red represents positivity. The white arrow indicates the beginning of the sound.

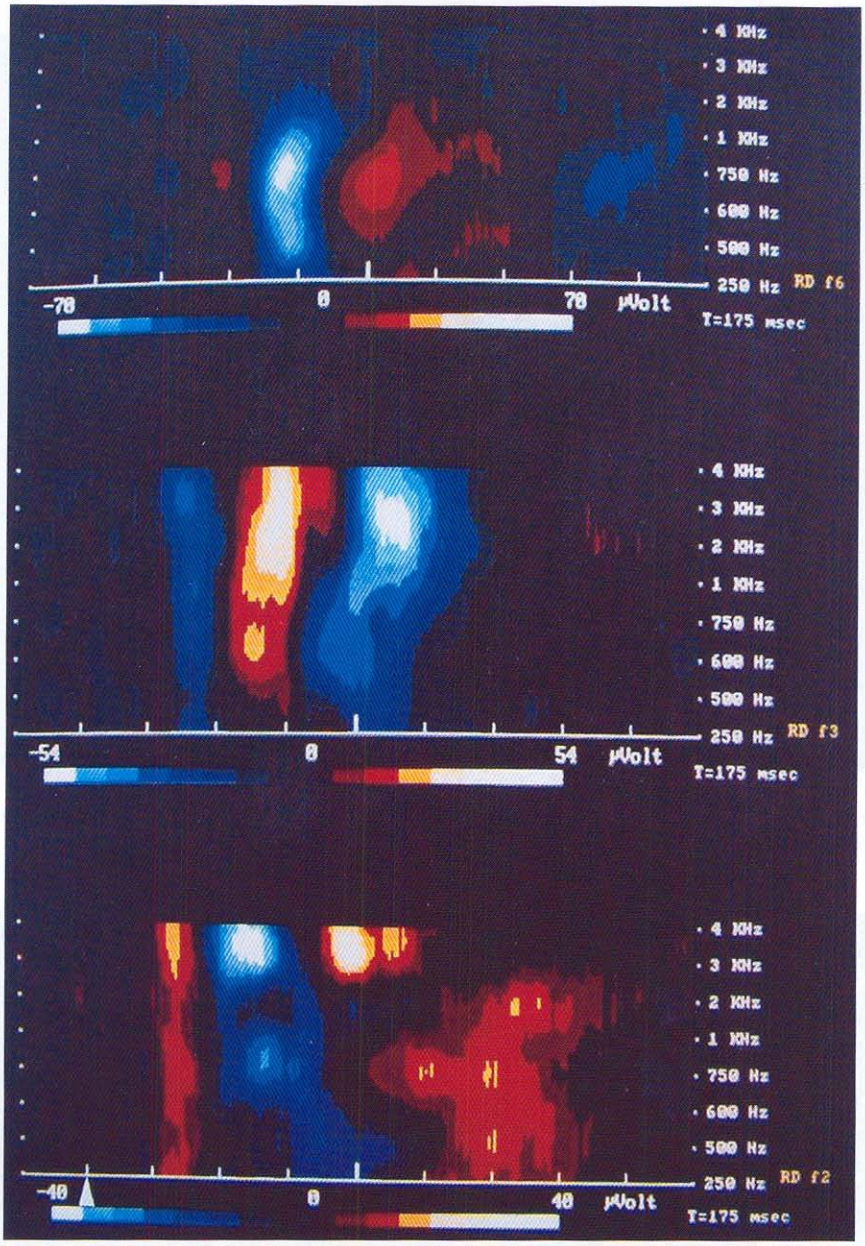



**Plate 1** Temporospectral maps of amplitude variations of potentials recorded from the right primary auditory cortex (leads 2, 3, and 6). Time, related in abscissae, is reduced to 175 ms, full scale. On the ordinate axis, each yellow dot represents the frequency. Amplitude is represented by colour. A cubic spline interpolation is used in order to obtain a continuous potential line between the different frequencies. Each spatial /reconstruction gives a continuous map both in time and in frequency. Dark colors are values close to the neutral potential, and bright colors are used for the activity—blue represents negativity and red represents positivity. The white arrow indicates the beginning of the sound.

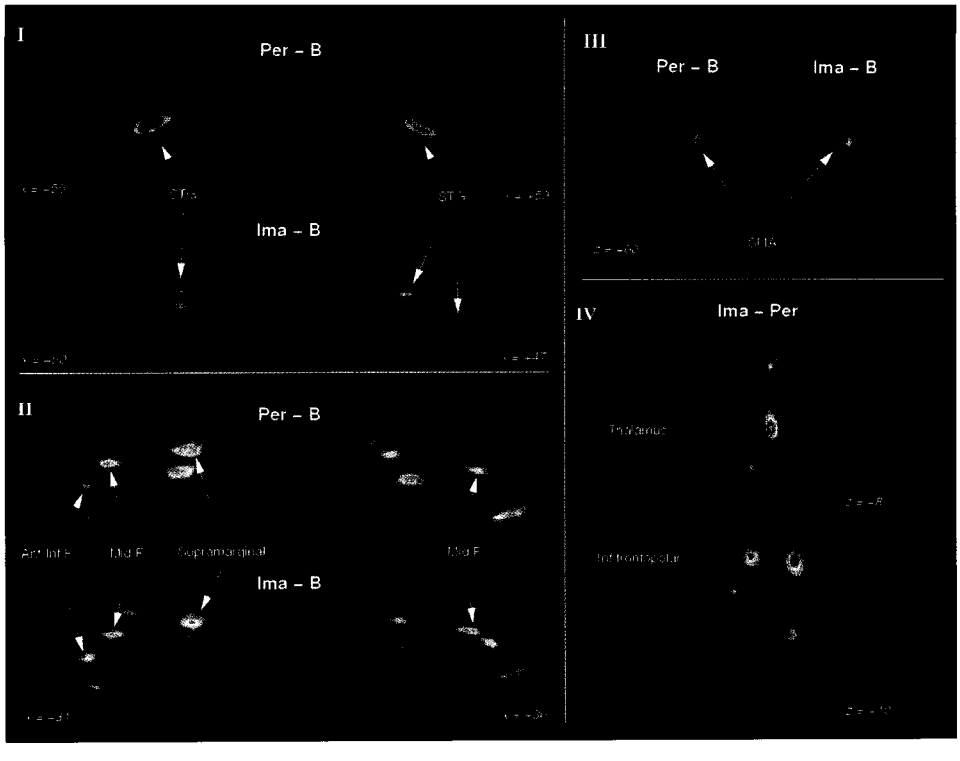



Plate 2 Brain areas showing activation in the perception minus baseline, imagery minus baseline, and imagery minus perception conditions of the verbal tune PET study.

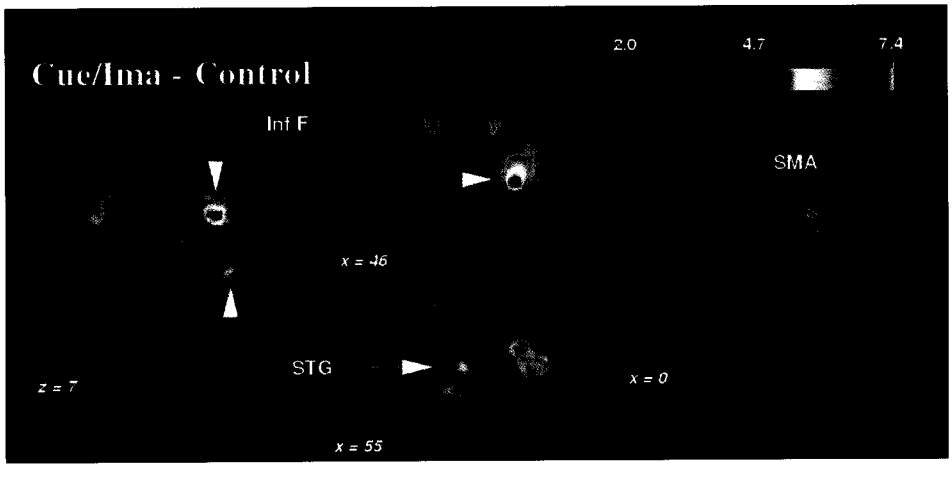



Plate 3 Brain areas showing activation in the cue/imagery minus control subtraction in the nonverbal tune PET study.

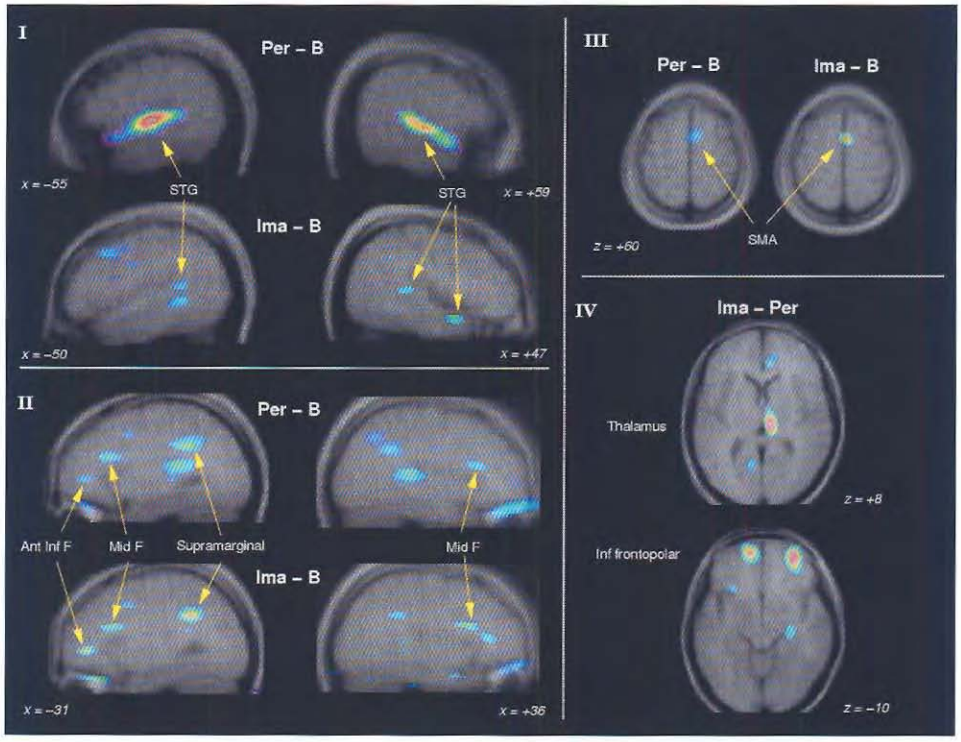



Plate 2 Brain areas showing activation in the perception minus baseline, imagery minus baseline, and imagery minus perception conditions of the verbal tune PET study.

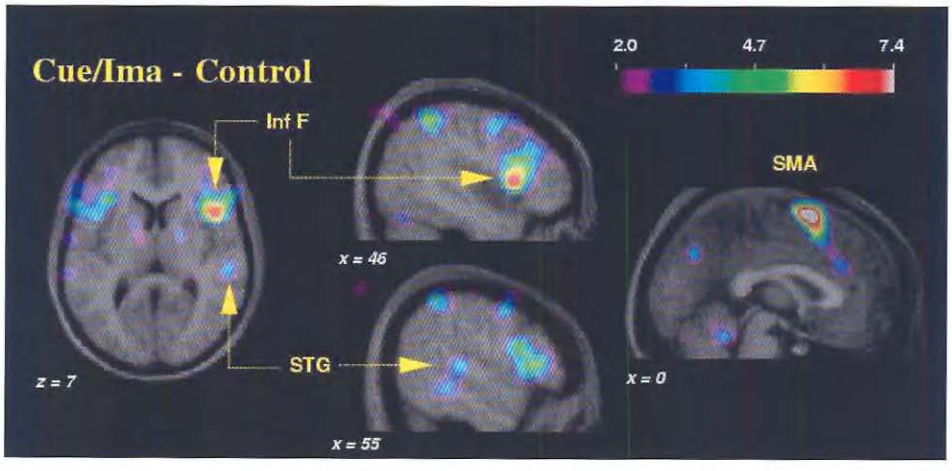



Plate 3 Brain areas showing activation in the cue/imagery minus control subtraction in the nonverbal tune PET study.

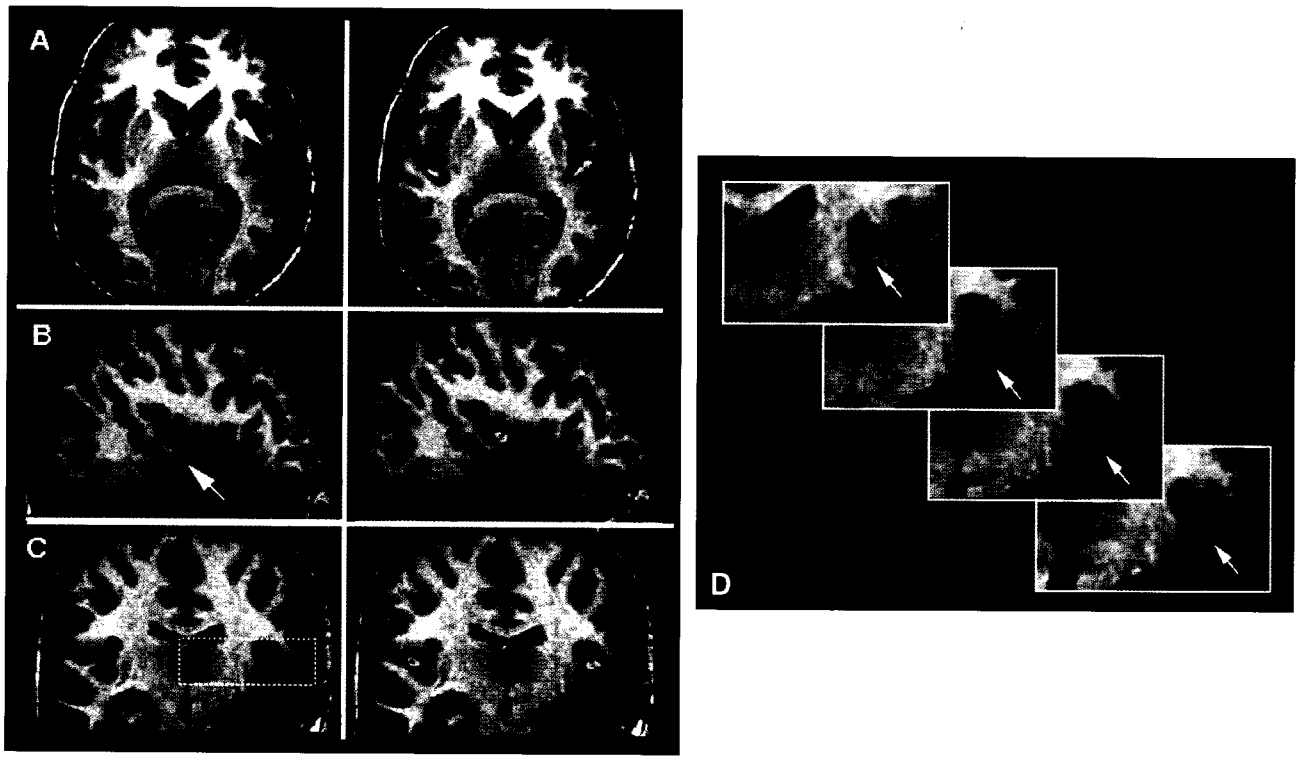



**Plate 4** MRI scan of a patient with a surgical excision in the right HG; illustrated are planes of section oriented horizontally (A), sagittally (B), and coronally (C). The left side of panels A, B, and C show the patient's scan alone, with an *arrow* indicating the region of excision/undercutting. The images on the *right side* of panels A, B, and C show the patient's scan coregistered with an anatomical probabilistic map of HG derived from normal individuals (Ref. 44). The *crosshairs* indicate the same position in the standardized space as the *arrow*. Note the correspondence between the position of HG as determined from the map and the patient's partially excised HG region. The *yellow box* in panel C indicates the region of the removal pictured in close-up in panel D, which illustrates the transition from intact, to undercut, to fully excised tissue (coronal sections taken at 3-mm intervals, posterior to anterior). *Arrows* again correspond to the *crosshairs* in the other panels and indicate the location of the HG region.

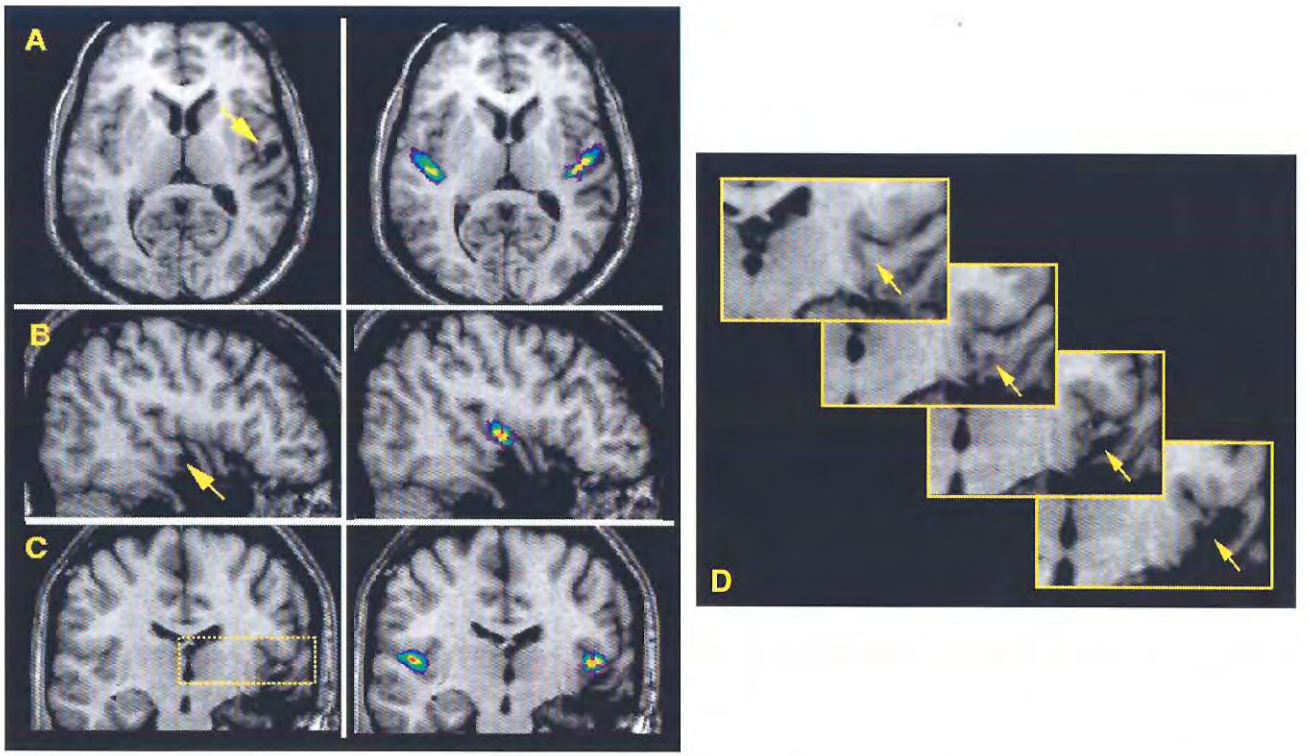



**Plate 4** MRI scan of a patient with a surgical excision in the right HG; illustrated are planes of section oriented horizontally (A), sagittally (B), and coronally (C). The left side of panels A, B, and C show the patient's scan alone, with an *arrow* indicating the region of excision/undercutting. The images on the *right side* of panels A, B, and C show the patient's scan coregistered with an anatomical probabilistic map of HG derived from normal individuals (Ref. 44). The *crosshairs* indicate the same position in the standardized space as the *arrow*. Note the correspondence between the position of HG as determined from the map and the patient's partially excised HG region. The *yellow box* in panel C indicates the region of the removal pictured in close-up in panel D, which illustrates the transition from intact, to undercut, to fully excised tissue (coronal sections taken at 3-mm intervals, posterior to anterior). *Arrows* again correspond to the *crosshairs* in the other panels and indicate the location of the HG region.

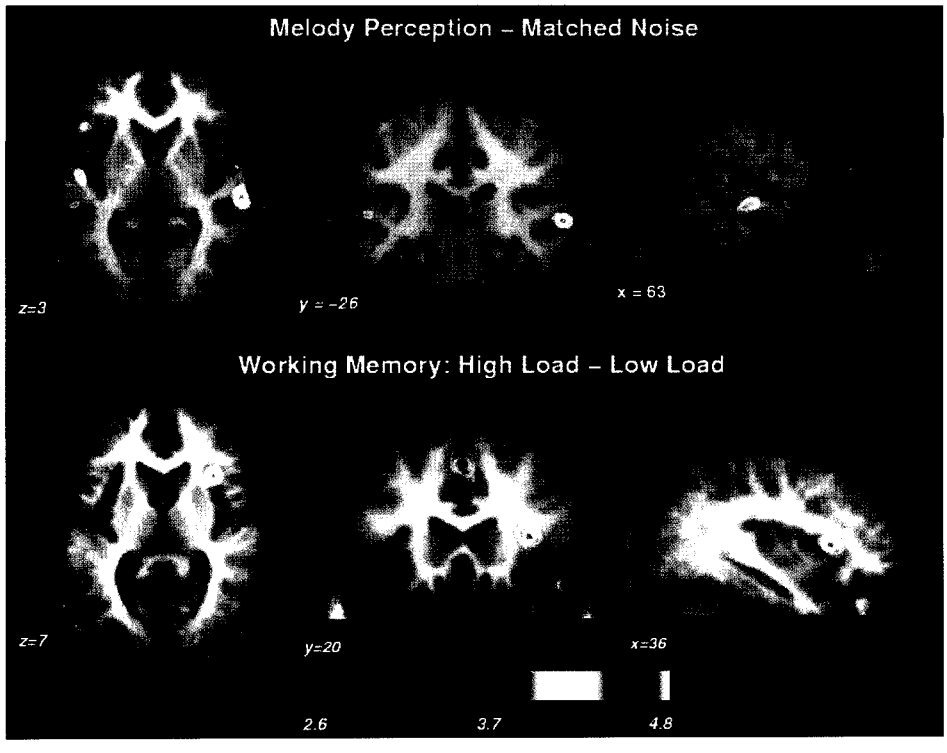



Plate 5 Selected PET CBF activation sites associated with processing of tonal melodies. The figures show averaged PET subtraction images superimposed on a corresponding section of an average anatomical MRI. (*Top*) Horizontal, coronal, and sagittal views through a focus located in the right superior temporal gyrus representing a significant blood flow increase while subjects listened to a series of short, unfamiliar tonal melodies, as compared with a baseline condition in which they listened to noise bursts that were acoustically equated to the melodies. Also visible in the horizontal section is a region of significant activity in the left temporal cortex. (*Bottom*) Horizontal, coronal, and sagittal views through a focus in the right frontal/opercular region showing activity associated with making judgements about the pitch of the first and last notes of the tonal sequences as compared to making judgements about the first two uotes. Data reanalysed from Zatorre *et al.*

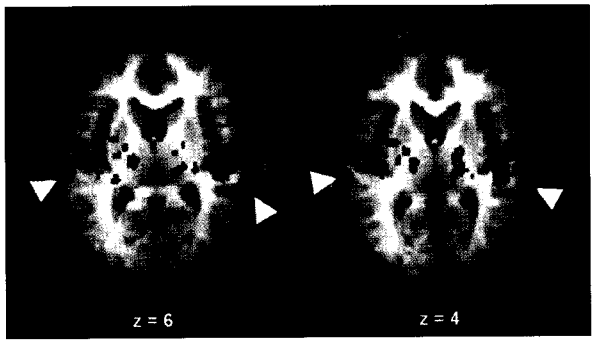



Plate 6 Significant blood flow activations in auditory (temporal) cortex specific to playing the Bach score (*red*), specific to playing scales (*blue*), and common to both performances (*cyan*). These are logical contrasts of Bach vs rest and scales vs rest shown in group-averaged PET images overlaid on anatomical magnetic resonance images (MRIs). The *z*-values indicate the axial height of the brain volume relative to the Talairach and Tournoux stereotactic atlas. Throughout, left side of brain images shows the left side of the brain, and vice versa.

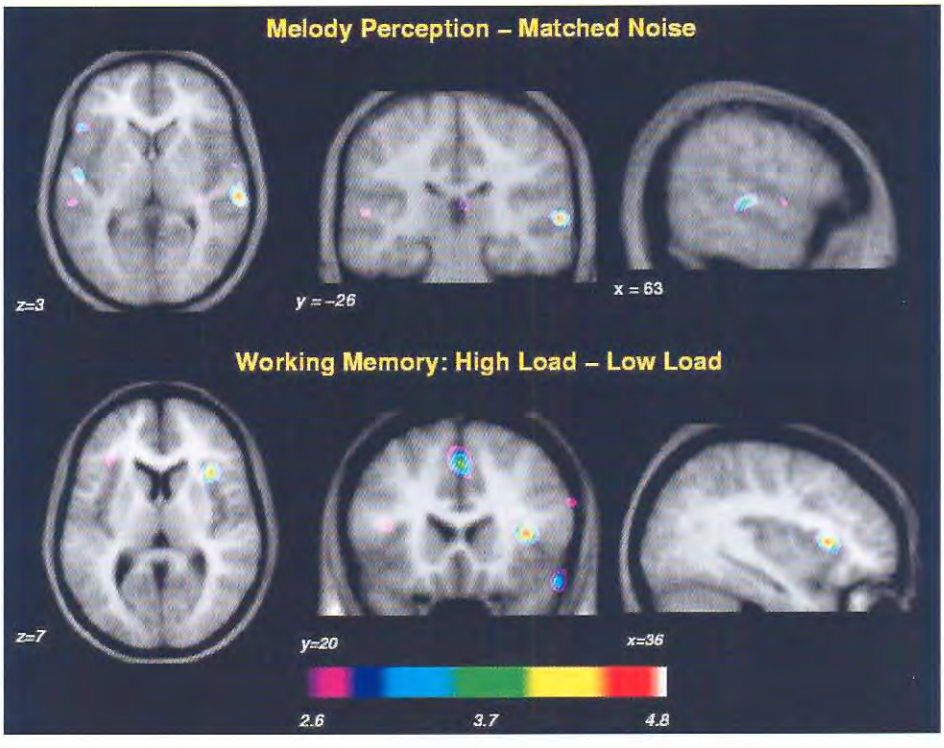



**Plate 5** Selected PET CBF activation sites associated with processing of tonal melodies. The figures show averaged PET subtraction images superimposed on a corresponding section of an average anatomical MRI. (*Top*) Horizontal, coronal, and sagittal views through a focus located in the right superior temporal gyrus representing a significant blood flow increase while subjects listened to a series of short, unfamiliar tonal melodies, as compared with a baseline condition in which they listened to noise bursts that were acoustically equated to the melodies. Also visible in the horizontal section is a region of significant activity in the left temporal cortex. (*Bottom*) Horizontal, coronal, and sagittal views through a focus in the right frontal/opercular region showing activity associated with making judgements about the pitch of the first and last notes of the tonal sequences as compared to making judgements about the first two notes. Data reanalysed from Zatorre *et al.*

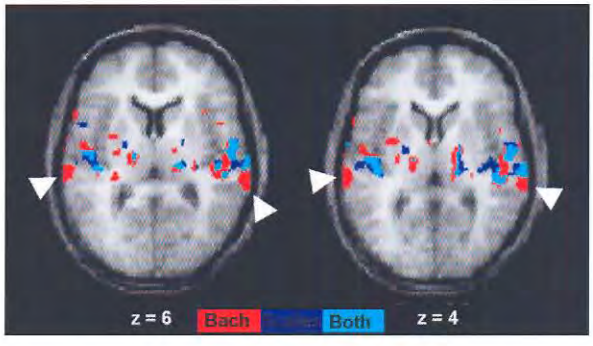



**Plate 6** Significant blood flow activations in auditory (temporal) cortex specific to playing the Bach score (*red*), specific to playing scales (*blue*), and common to both performances (*cyan*). These are logical contrasts of Bach vs rest and scales vs rest shown in group-averaged PET images overlaid on anatomical magnetic resonance images (MRIs). The *z*-values indicate the axial height of the brain volume relative to the Talairach and Tournoux stereotactic atlas. Throughout, left side of brain images shows the left side of the brain, and vice versa.

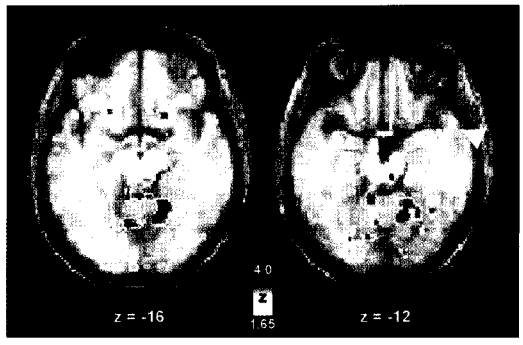



**Plate 7** Correlated blood flow activations in right anterior auditory cortex and left cerebellum (see *arrows*) as subjects played the Bach score. These are group-averaged PET images for Bach (contrasted with rest) overlaid on anatomical MRIs. PET data are *z*-scores displayed on a colour scale ranging from 1.65 (*yellow*; $p < 0.1$) to 4.0 (*red*; $p < 0.0001$).

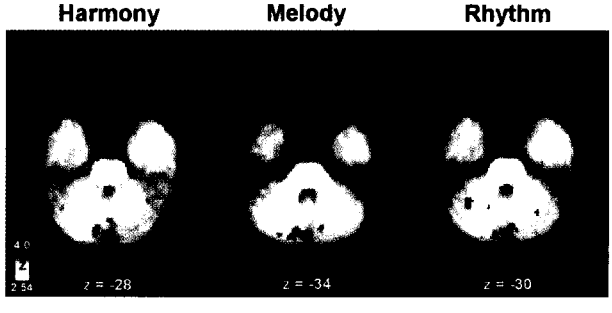



**Plate 8** Significant blood flow activations in the cerebellum as musicians sight- read a J. S. Bach chorale score and listen for performance errors in harmony, melody, or rhythm. These are group-averaged PET images for each task (contrasted with passive listening control) overlaid on anatomical MRIs. PET data are *z*-scores displayed on a colour scale ranging from 1.96 (*yellow*; $p < 0.05$) to 4.0 (*red*; $p < 0.0001$).

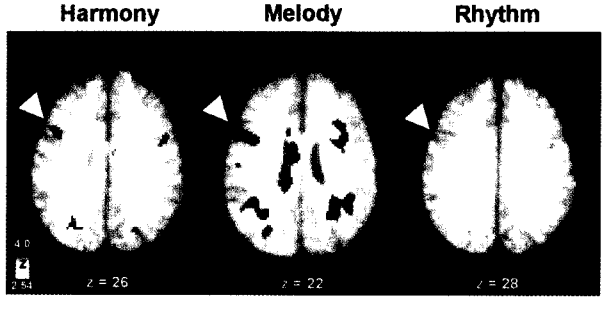



**Plate 9** Activations in different subareas of a major cytoarchitectonic area of frontal cortex (BA 44) as musicians sight-read a J. S. Bach chorale score and listen for performance errors in harmony, melody, or rhythm. These are group mean PET images for each task (contrasted with passive listening control) overlaid on MRIs. PET data are *z*-scores displayed on a colour scale ranging from 2.54 (*yellow*; $p < 0.01$) to 4.0 (*red*; $p < 0.0001$).

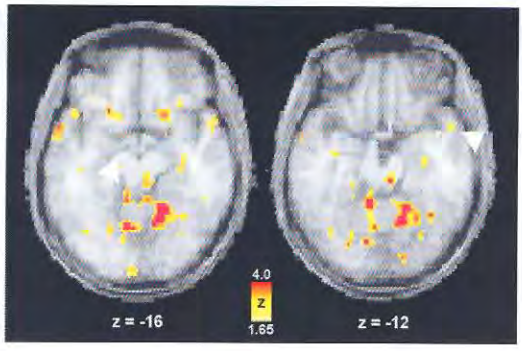



**Plate 7** Correlated blood flow activations in right anterior auditory cortex and left cerebellum (see *arrows*) as subjects played the Bach score. These are group-averaged PET images for Bach (contrasted with rest) overlaid on anatomical MRIs. PET data are *z*-scores displayed on a colour scale ranging from 1.65 (*yellow*; $p < 0.1$) to 4.0 (*red*; $p < 0.0001$).

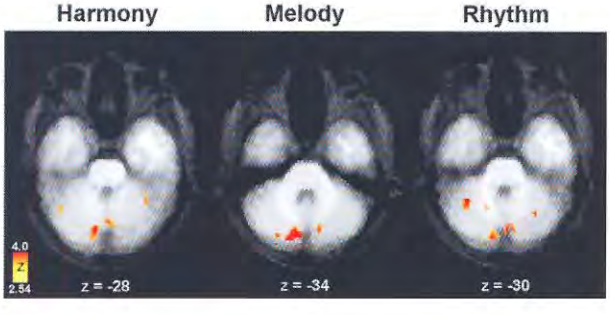



**Plate 8** Significant blood flow activations in the cerebellum as musicians sight- read a J. S. Bach chorale score and listen for performance errors in harmony, melody, or rhythm. These are group-averaged PET images for each task (contrasted with passive listening control) overlaid on anatomical MRIs. PET data are *z*-scores displayed on a colour scale ranging from 1.96 (*yellow*; $p < 0.05$) to 4.0 (*red*; $p < 0.0001$).

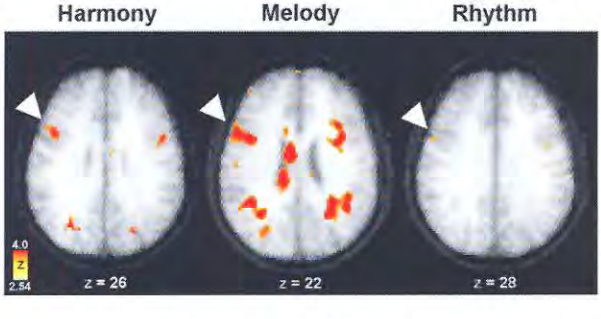



**Plate 9** Activations in different subareas of a major cytoarchitectonic area of frontal cortex (BA 44) as musicians sight-read a J. S. Bach chorale score and listen for performance errors in harmony, melody, or rhythm. These are group mean PET images for each task (contrasted with passive listening control) overlaid on MRIs. PET data are *z*-scores displayed on a colour scale ranging from 2.54 (*yellow*; $p < 0.01$) to 4.0 (*red*; $p < 0.0001$).

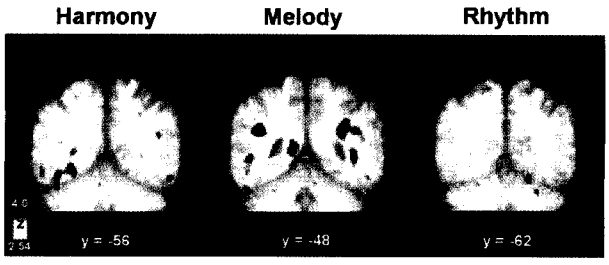



Plate 10 Activations in different subareas of auditory cortex as musicians sight-read a J. S. Bach chorale score and listen for performance errors in melody, harmony, or rhythm. These are group mean PET images for each task (contrasted with passive listening control) overlaid on MRIs. PET data are *z*-scores displayed on a colour scale ranging from 2.54 (*yellow*; $p < 0.05$) to 4.0 (*red*; $p < 0.0001$).

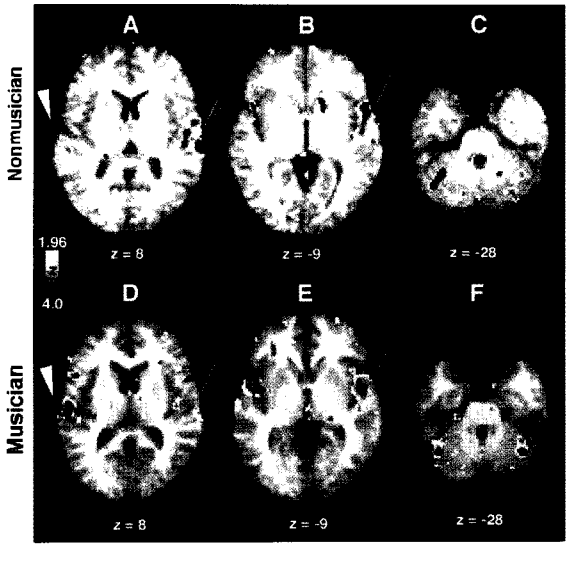



Plate 11 Activations in superior and middle temporal cortex (A, B, D, and E) and in cerebellum (C and F) as nonmusicians and musicians discriminate pitch sequences. These images show group mean PET data for each task (contrasted with rest) overlaid on MRIs. PET data are *z*-scores displayed on a colour scale ranging from 1.96 (*yellow*; $p < 0.1$) to 4.0 (*red*; $p < 0.0001$).

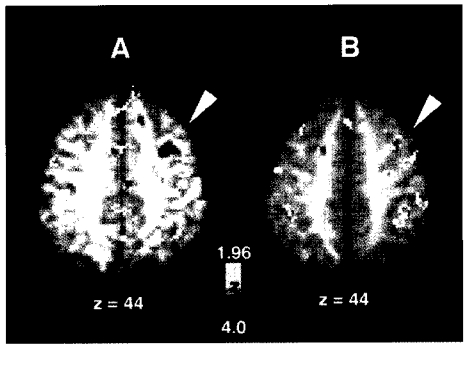



Plate 12 Activations in left medial frontal cortex (BA 9) for nonmusicians (A) as compared to musicians (B) during the discrimination of musical tempi. Shown are group mean PET images for each task (contrasted with rest) overlaid on MRIs.

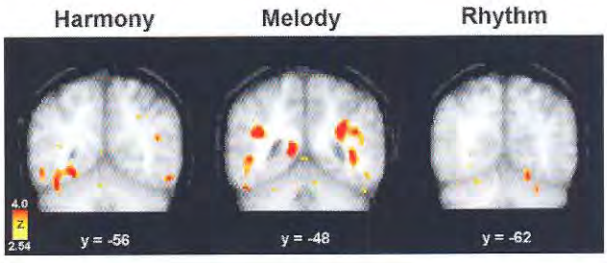



**Plate 10** Activations in different subareas of auditory cortex as musicians sight-read a J. S. Bach chorale score and listen for performance errors in melody, harmony, or rhythm. These are group mean PET images for each task (contrasted with passive listening control) overlaid on MRIs. PET data are *z*-scores displayed on a colour scale ranging from 2.54 (*yellow*; $p < 0.05$) to 4.0 (*red*; $p < 0.0001$).

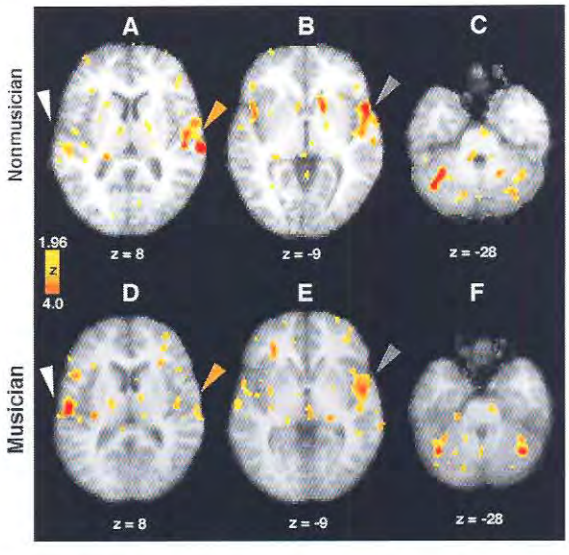



**Plate 11** Activations in superior and middle temporal cortex (A, B, D, and E) and in cerebellum (C and F) as nonmusicians and musicians discriminate pitch sequences. These images show group mean PET data for each task (contrasted with rest) overlaid on MRIs. PET data are *z*-scores displayed on a colour scale ranging from 1.96 (*yellow*; $p < 0.1$) to 4.0 (*red*; $p < 0.0001$).

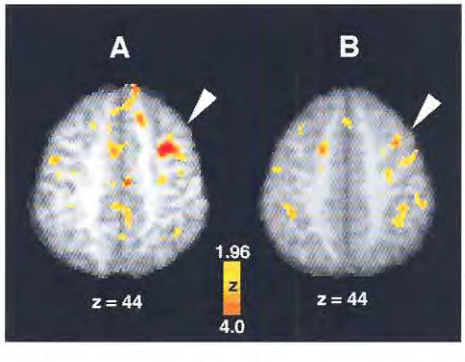



**Plate 12** Activations in left medial frontal cortex (BA 9) for nonmusicians (A) as compared to musicians (B) during the discrimination of musical tempi. Shown are group mean PET images for each task (contrasted with rest) overlaid on MRIs.

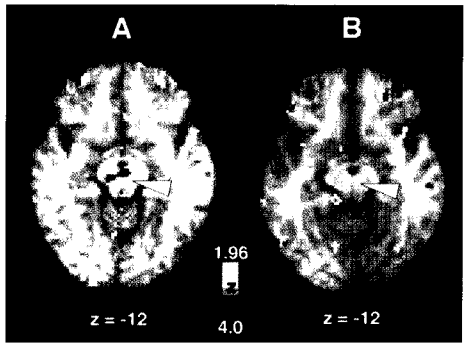



Plate 13 Activations in midbrain for nonmusicians (A) as compared to musicians (B) during the discrimination of musical pattern. These are group-averaged PET images for each task (contrasted with rest) overlaid on anatomical MRIs.

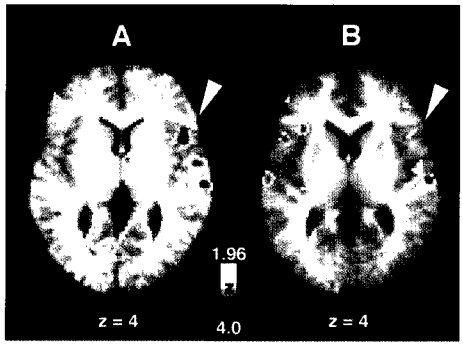



Plate 14 Activations in right inferior frontal cortex (BA 47) for nonmusicians (A) as compared to musicians (B) during discrimination of musical meter. Shown are group-averaged PET images for each task (contrasted with rest) over-laid on anatomical MRIs.

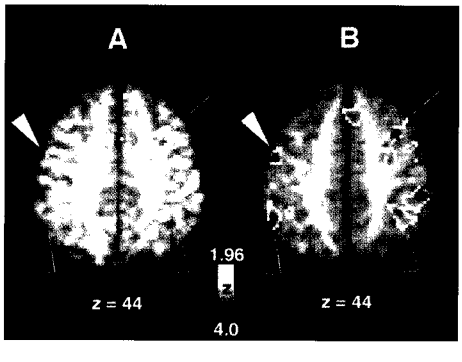



Plate 15 Activations in bilateral medial frontal cortex (BA 9) and bilateral inferior parietal cortex (BA 40) for musicians (B) as compared to nonmusicians (A) during the discrimination of sequence durations. Shown are group-averaged PET images for each task (contrasted with rest) overlaid on anatomical MRIs.

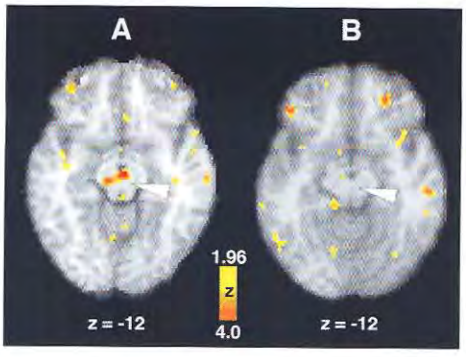



Plate 13 Activations in midbrain for nonmusicians (A) as compared to musicians (B) during the discrimination of musical pattern. These are group-averaged PET images for each task (contrasted with rest) overlaid on anatomical MRIs.

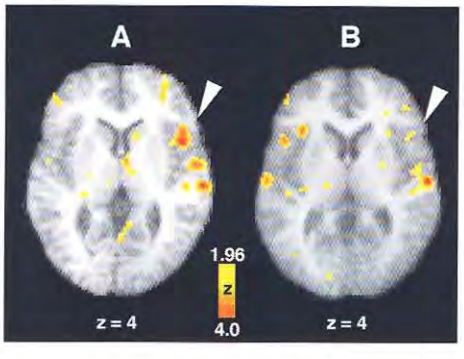



Plate 14 Activations in right inferior frontal cortex (BA 47) for nonmusicians (A) as compared to musicians (B) during discrimination of musical meter. Shown are group-averaged PET images for each task (contrasted with rest) over-laid on anatomical MRIs.

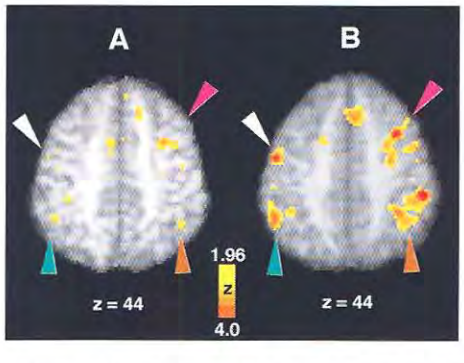



Plate 15 Activations in bilateral medial frontal cortex (BA 9) and bilateral inferior parietal cortex (BA 40) for musicians (B) as compared to nonmusicians (A) during the discrimination of sequence durations. Shown are group-averaged PET images for each task (contrasted with rest) overlaid on anatomical MRIs.

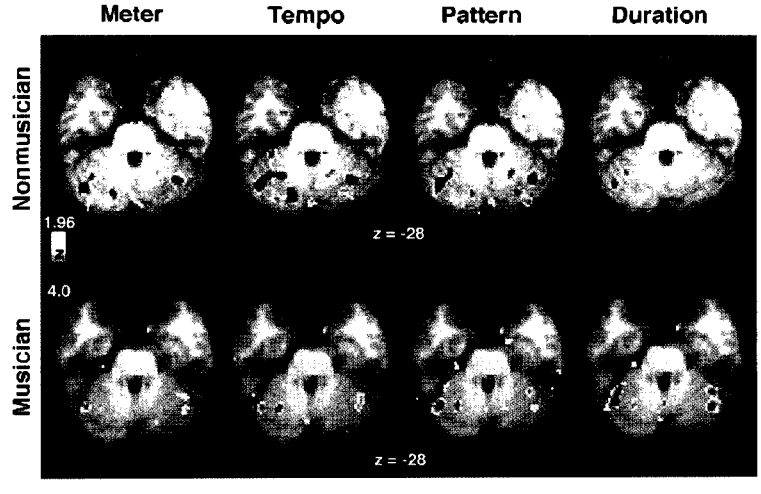



**Plate 16** Activations in posterior cerebellum as nonmusicians and musicians discriminate meter, tempo, pattern, and duration. Shown are group-mean PET images for each task (contrasted with rest) overlaid on MRIs.

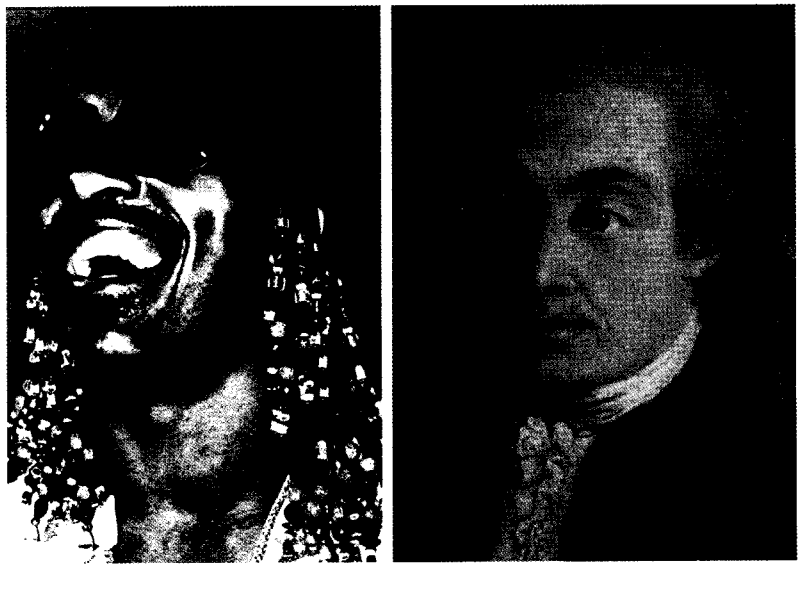


**Plate 17** Anecdotal evidence reports superior musical abilities in blind individuals. *Left*: Stevie Wonder. From: *The Rolling Stone. Illustrated History of Rock and Roll.* A. DeCurtis, J. Henke, H. George-Warren (eds) New York: Random House, p. 294. *Right*: Wolfgang Amadeus Mozart, though not blind, possibly suffered from strabismic amblyopta, reducing vision in one or both eyes by an unknown amount. Portrait attributed to Kymli of Mannheim, Germany, 1783 (held in a private collection).

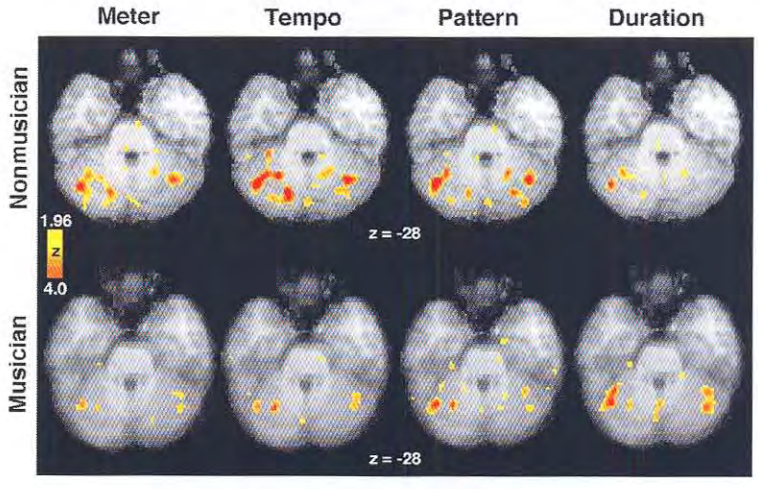



Plate 16 Activations in posterior cerebellum as nonmusicians and musicians discriminate meter, tempo, pattern, and duration. Shown are group-mean PET images for each task (contrasted with rest) overlaid on MRIs.

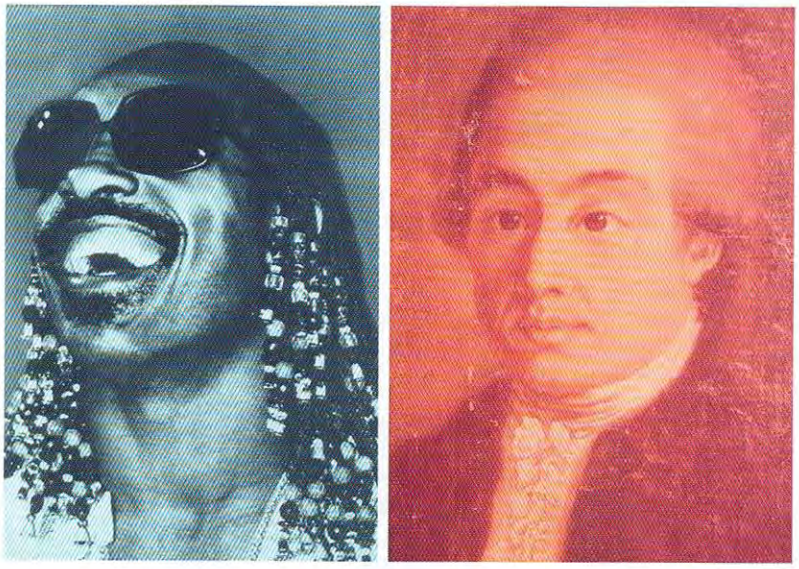


Plate 17 Anecdotal evidence reports superior musical abilities in blind individuals. *Left*: Stevie Wonder. From: *The Rolling Stone. Illustrated History of Rock and Roll.* A. DeCurtis, J. Henke, H. George-Warren (eds) New York: Random House, p. 294. *Right*: Wolfgang Amadeus Mozart, though not blind, possibly suffered from strabismic amblyopta, reducing vision in one or both eyes by an unknown amount. Portrait attributed to Kymli of Mannheim, Germany, 1783 (held in a private collection).

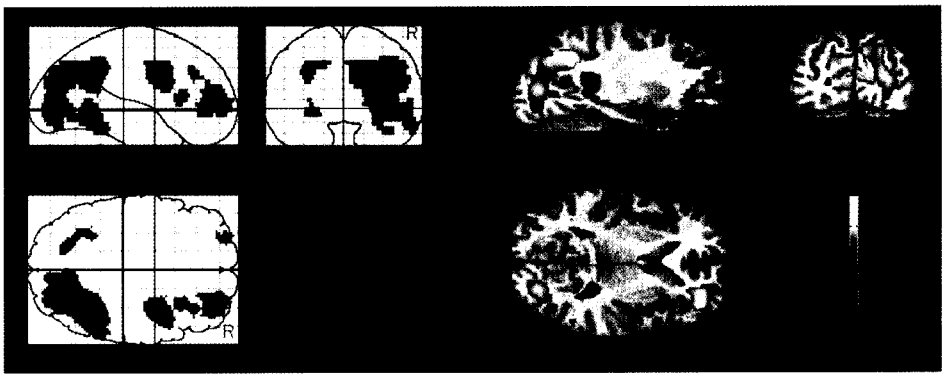


**Plate 18** Results of PET scanning in congenitally blind subjects during localization of sounds presented via headphones in virtual auditory space (data from Ref. 19). Areas in the occipital cortex that are driven by visual stimuli in sighted individuals are now activated during sound localization. Note that the right hemisphere seems to profit more from the expansion of auditory into visual territory. The 'rewiring' seems to be mediated by the parietal cortex, consistent with studies on visually deprived animals.

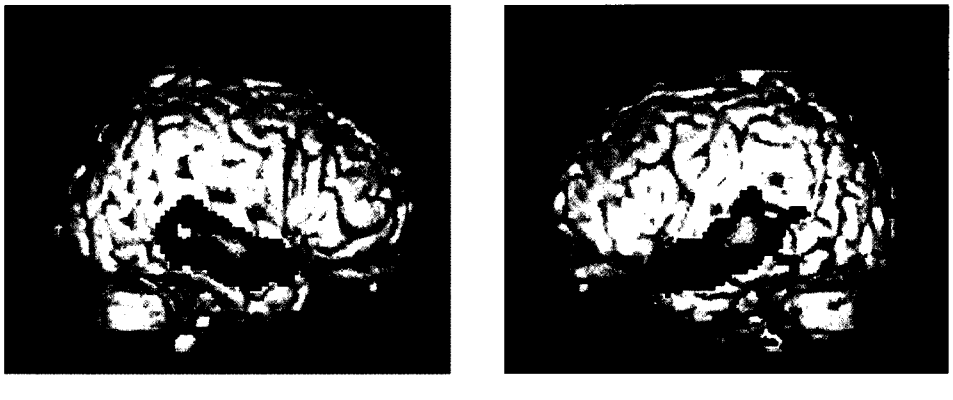


**Plate 19** Auditory cortical activation while listening to music. Averaged group results of functional magnetic resonance imaging are shown. While the processing of music by the right cerebral hemisphere has often been emphasized (as opposed to language, which is thought to be processed primarily on the left), it is now generally accepted that different aspects of music are processed by *both* hemispheres.

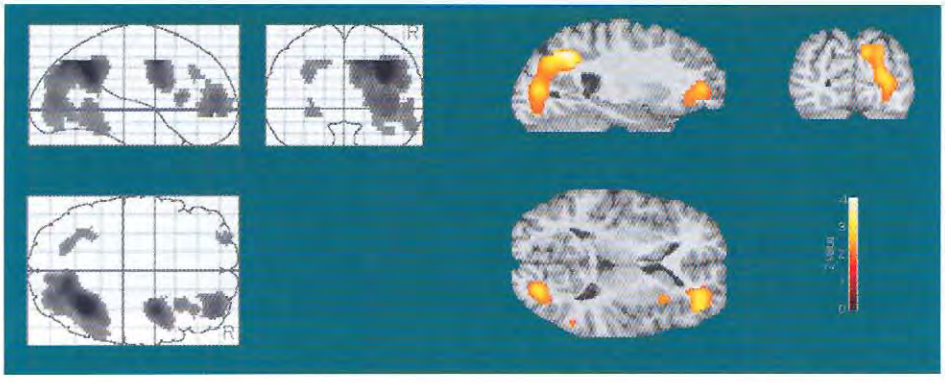


Plate 18 Results of PET scanning in congenitally blind subjects during localization of sounds presented via headphones in virtual auditory space (data from Ref. 19). Areas in the occipital cortex that are driven by visual stimuli in sighted individuals are now activated during sound localization. Note that the right hemisphere seems to profit more from the expansion of auditory into visual territory. The 'rewiring' seems to be mediated by the parietal cortex, consistent with studies on visually deprived animals.

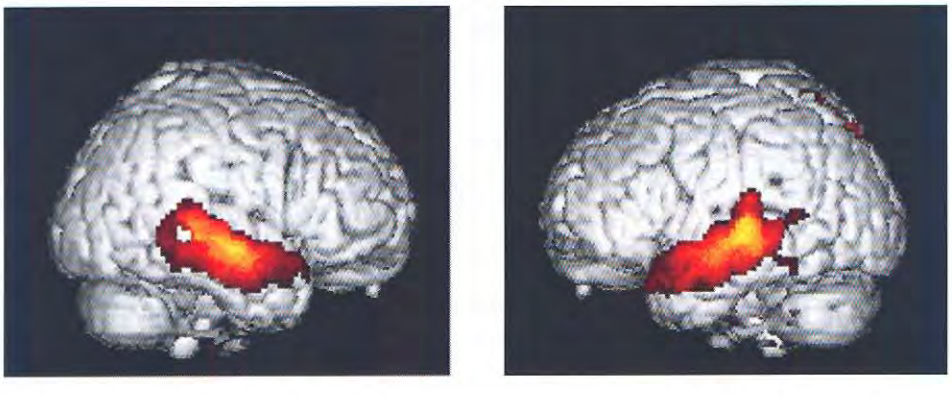


Plate 19 Auditory cortical activation while listening to music. Averaged group results of functional magnetic resonance imaging are shown. While the processing of music by the right cerebral hemisphere has often been emphasized (as opposed to language, which is thought to be processed primarily on the left), it is now generally accepted that different aspects of music are processed by *both* hemispheres.

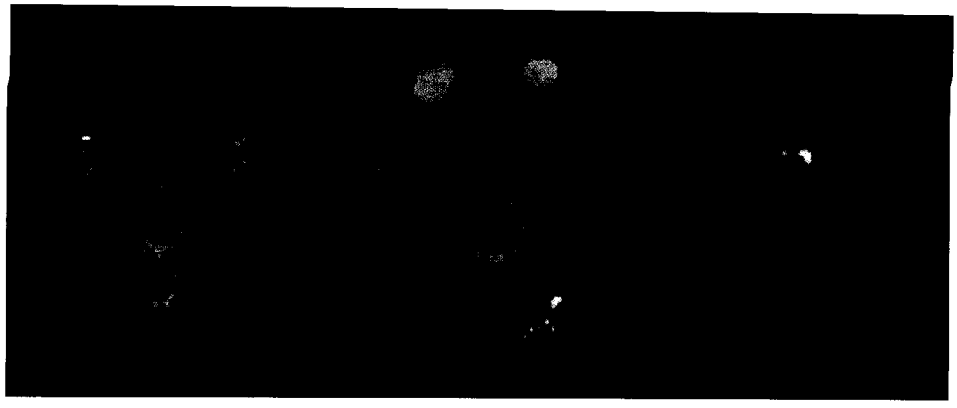


**Plate 20** Results of functional magnetic resonance imaging during anticipatory imagery of music. The axial sections from individual subjects are displayed in radiological coordinates showing the left side of the brain on the right and vice versa. Imagery leads to activation of anterior regions of auditory cortex in the right hemisphere (in conjunction with right inferior frontal regions; panel on the left of the figure). This emphasizes the role of sensory cortices in mental imagery. However, brain regions outside auditory cortex, such as the cerebellum (left cerebellar hemisphere; middle panel) and the anterior cingulate bilaterally (right panel), are also activated, demonstrating their role in cognitive processing.

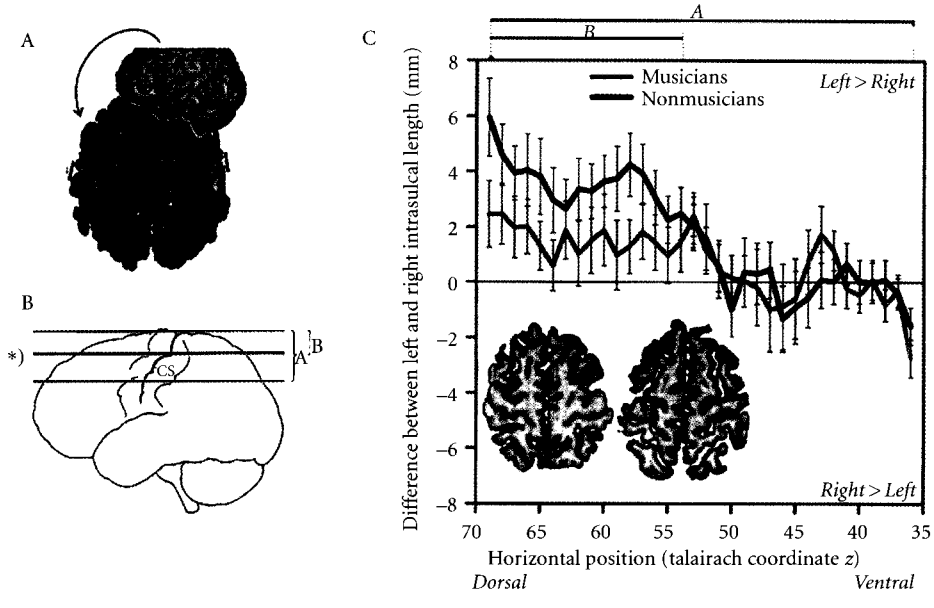



**Plate 21** (A) The intrasulcal length of the posterior bank of the precentral gyrus (ILPG) was measured in horizontal slices (parallel to the AC–PC plane) from the deepest point of the central sulcus following the contour line of the precentral gyrus to a lateral surface tangent that connected the crests of the pre- and postcentral gyrus. (B and C) We differentiated a dorsal and a ventral subregion. The dorsal subregion roughly corresponds to the location of the functional hand motor area. We found prominent differences in the intrasulcal length of the precentral gyrus in the dorsal subregion between musicians and nonmusicians. Musicians had more symmetrical lengths which was due to a disproportional increase in length of the nondominant hemisphere.

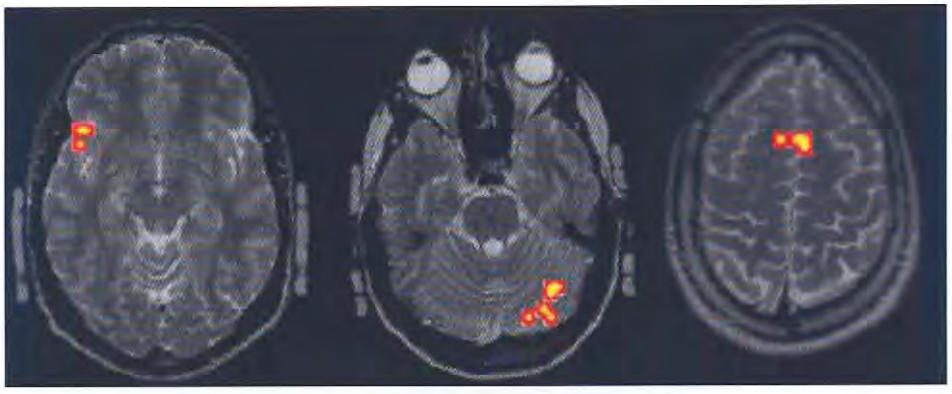


**Plate 20** Results of functional magnetic resonance imaging during anticipatory imagery of music. The axial sections from individual subjects are displayed in radiological coordinates showing the left side of the brain on the right and vice versa. Imagery leads to activation of anterior regions of auditory cortex in the right hemisphere (in conjunction with right inferior frontal regions; panel on the left of the figure). This emphasizes the role of sensory cortices in mental imagery. However, brain regions outside auditory cortex, such as the cerebellum (left cerebellar hemisphere; middle panel) and the anterior cingulate bilaterally (right panel), are also activated, demonstrating their role in cognitive processing.

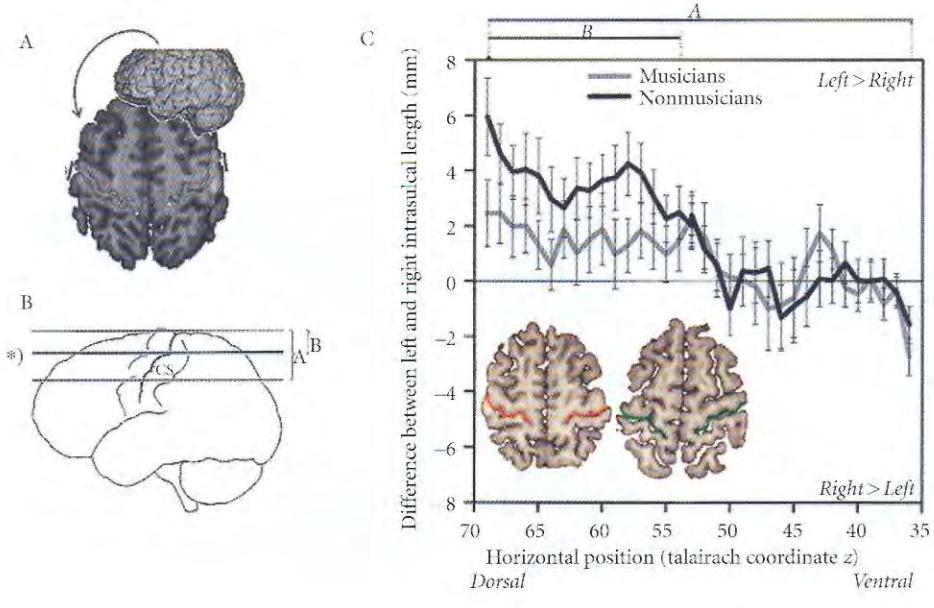



**Plate 21** (A) The intrasulcal length of the posterior bank of the precentral gyrus (ILPG) was measured in horizontal slices (parallel to the AC–PC plane) from the deepest point of the central sulcus following the contour line of the precentral gyrus to a lateral surface tangent that connected the crests of the pre- and postcentral gyrus. (B and C) We differentiated a dorsal and a ventral subregion. The dorsal subregion roughly corresponds to the location of the functional hand motor area. We found prominent differences in the intrasulcal length of the precentral gyrus in the dorsal subregion between musicians and nonmusicians. Musicians had more symmetrical lengths which was due to a disproportional increase in length of the nondominant hemisphere.

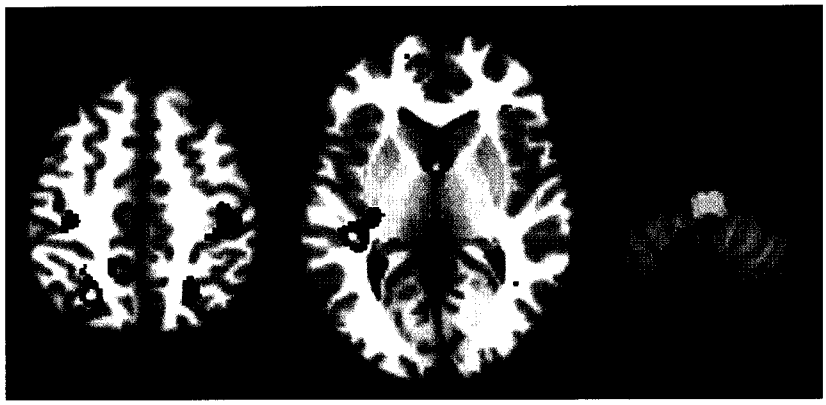


Plate 22 Preliminary findings of a voxel-by-voxel morphometric study comparing a group of 15 male musicians with 15 male nonmusicians. Three axial slices show significant differences ($p < 0.05$ corrected) between musicians and nonmusicians superimposed on corresponding T1-weighted structural images. Yellow and red voxels indicate more gray-matter concentration in the musicians, blue voxels indicate more gray matter concentration in the nonmusicians.

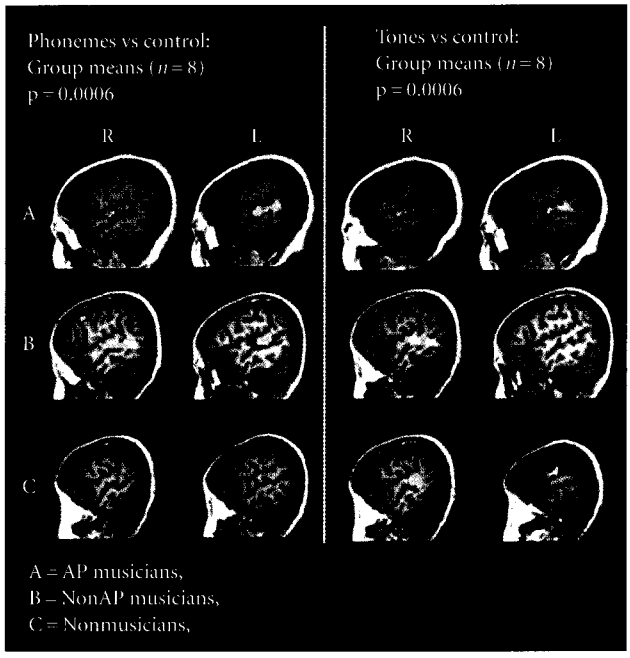



**Plate 23** Results of a fMRI experiment aimed to investigate the functional significance of the increased leftsided anatomical PT asymmetry in AP musicians. Three groups of subjects were examined: AP musicians, nonAP musicians, and nonmusicians. A phoneme memory task and a tone memory task were contrasted with a rest condition. Suprathreshold voxels showing significant task related activity changes were superimposed on a representative sagittal anatomical image of right and left hemisphere. AP musicians showed prominent left superior temporal lobe activation during the tone memory task that was not seen in the other two groups. The cerebral activation pattern of the phoneme memory task was similar between the three groups.

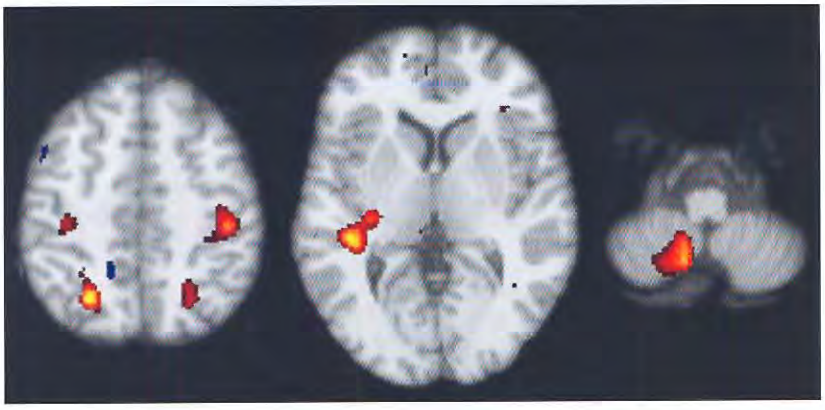


Plate 22 Preliminary findings of a voxel-by-voxel morphometric study comparing a group of 15 male musicians with 15 male nonmusicians. Three axial slices show significant differences ($p < 0.05$ corrected) between musicians and nonmusicians superimposed on corresponding T1-weighted structural images. Yellow and red voxels indicate more gray-matter concentration in the musicians, blue voxels indicate more gray matter concentration in the nonmusicians.

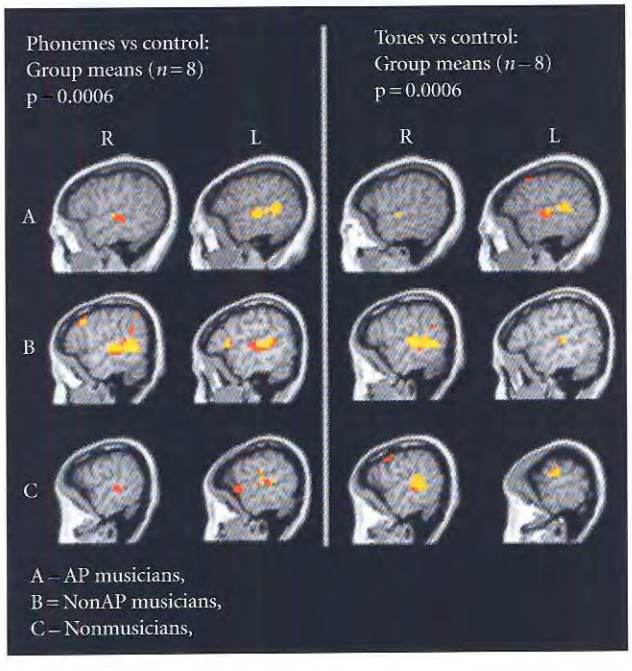



Plate 23 Results of a fMRI experiment aimed to investigate the functional significance of the increased leftsided anatomical PT asymmetry in AP musicians. Three groups of subjects were examined: AP musicians, nonAP musicians, and nonmusicians. A phoneme memory task and a tone memory task were contrasted with a rest condition. Suprathreshold voxels showing significant task related activity changes were superimposed on a representative sagittal anatomical image of right and left hemisphere. AP musicians showed prominent left superior temporal lobe activation during the tone memory task that was not seen in the other two groups. The cerebral activation pattern of the phoneme memory task was similar between the three groups.

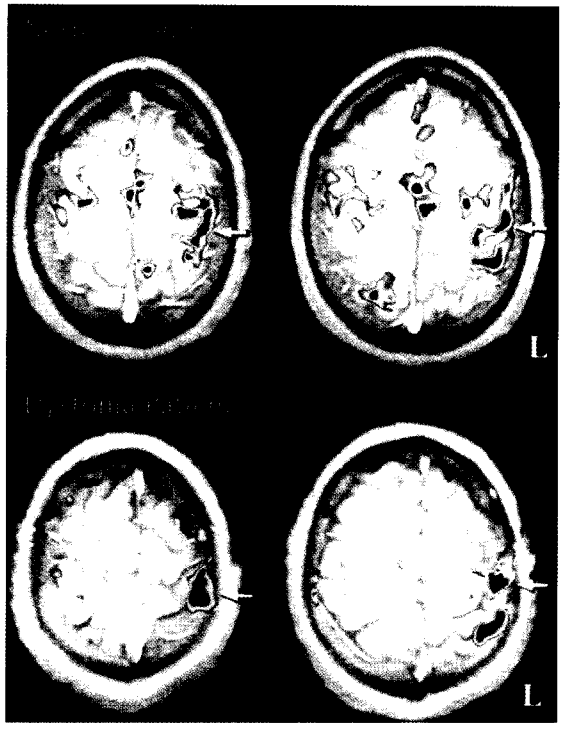



**Plate 24** The bold fMRI images of a normal and dystonic guitar player executing right hand arpeggios in the scanner are displayed. Note the greater activation of the sensorimotor cortex (*arrows*) and the lack of activation of premotor and supplementary motor cortices in the dystonic patient. Modified from Ref. 40.

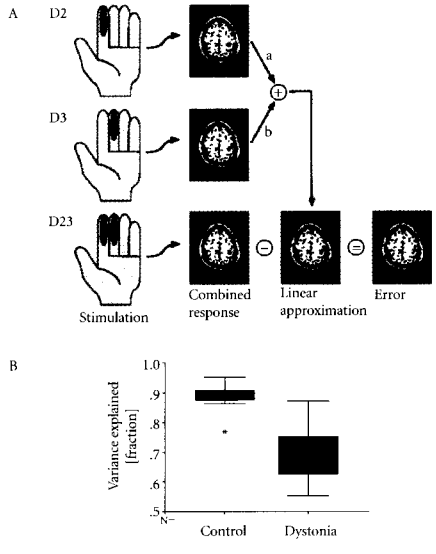



**Plate 25** (A) Schematic representation of the experimental design and representative fMRI data in one subject. Sensory stimulation was applied to the index (D2) or middle (D3) finger alone or to both fingers at the same time (D23) while the fMRI bold signal was recorded. A linear approximation of the fMRI signal associated with combined stimulation of D2 and D3 was calculated and substracted from the measured signal, giving rise to the 'error' of 'variance'. (B) Variance results in graphic form for all control subjects and patients with dystonia. Modified from Refs 46 and 47.

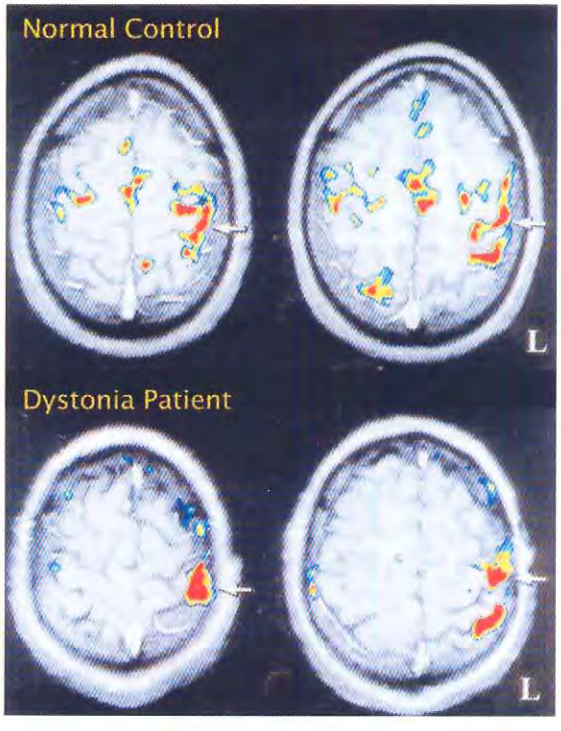



**Plate 24** The bold fMRI images of a normal and dystonic guitar player executing right hand arpeggios in the scanner are displayed. Note the greater activation of the sensorimotor cortex (*arrows*) and the lack of activation of premotor and supplementary motor cortices in the dystonic patient. Modified from Ref. 40.

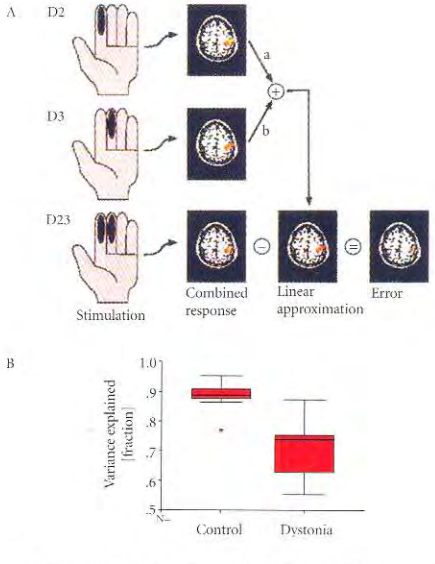



**Plate 25** (A) Schematic representation of the experimental design and representative fMRI data in one subject. Sensory stimulation was applied to the index (D2) or middle (D3) finger alone or to both fingers at the same time (D23) while the fMRI bold signal was recorded. A linear approximation of the fMRI signal associated with combined stimulation of D2 and D3 was calculated and substracted from the measured signal, giving rise to the 'error' of 'variance'. (B) Variance results in graphic form for all control subjects and patients with dystonia. Modified from Refs 46 and 47.

(Figure 16.1) using interactive pixel-marking software that permits simultaneous viewing in all three planes of the section. The result of greatest relevance pertains to the estimates of the volume of HG, which was found to be significantly greater on the left than on the right in two independent samples of 20 subjects each.[44] Most interesting of all was that the differences in volume were found to be confined to the white matter underlying HG, and not to the volume of cortical tissue (grey matter) within the structure (this effect may be seen qualitatively in the MRI shown in Figure 16.1). This finding suggests an anatomical asymmetry that arises from a difference in the volume of fibres that carry information to and from the primary auditory cortex and surrounding regions.

Although an MRI cannot reveal details of the underlying neuronal organization, the greater degree of white matter on the left could be consistent with the notion of greater speed of processing on the left, if the white-matter volume measures are related to degree of myelination. That is, a greater degree of left-sided myelination could lead to faster transmission of acoustically relevant information, because it is well established in neurophysiology that more heavily myelinated fibres transmit neuronal impulses more quickly. Confirmation that the left auditory cortical regions are indeed more heavily myelinated comes from a recent postmortem electron microscopy study showing that the myelin sheath is thicker in the left than in the right posterior temporal cortex.[45] As well, Hutsler and Gazzaniga[46] reported larger left- than right-layer IV pyramidal cells in the human auditory cortex, which would also lead to faster time constants.

In addition to the idea that left auditory cortical fibres are more heavily myelinated, there is also evidence for other hemispheric differences in the structural organization of the auditory cortex that are broadly consistent with the hypothesis presented above. For example, Seldon[47] reported that cortical columns in the left auditory cortex were more widely spaced than those on the right. Galuske *et al.*[48] recently found that clusters of neurons were spaced more widely in the left posterior temporal cortex, with more long-range intrinsic connections between these clusters. These features could be compatible with a greater degree of integration across tonotopically organized cortical areas on the left than on the right, hence leading to a relative degradation of spectral resolution on the left. Conversely, neurons in the right auditory cortex would have structural features that would enhance spectral resolution (much as shown by Liégeois-Chauvel, Chapter 10, this volume) because they would be smaller, with fewer myelinated inputs, therefore perhaps more tightly packed together; and would integrate over narrower frequency regions. These characteristics would also result in poorer temporal resolution.

## Concluding comments

We have seen how a variety of different methods and techniques point towards a general conclusion about the way in which the auditory nervous system processes information relevant for our ability to process speech and tonal sounds. These two domains of auditory processing are, of course, at the core of our uniquely human abilities to communicate. It is perhaps no coincidence that abilities to process both speech and music arise precociously in development and follow relatively fixed developmental sequences (see Chapter 1, this volume). It is perhaps also not irrelevant that both speech communication and tonal patterns appear to be ubiquitous across all human cultures (see Chapters 4 and 5, this volume).

The general approach we have defended, to view music within a biological framework, thus would seem to make sense of these disparate observations, if musical perceptual abilities—and by extension musical production abilities—are seen as emerging from functional and structural features of our nervous system. There is a final point that should perhaps be addressed, however, and this concerns the origins of music. Without entering into the controversies and speculations of this field of scholarship (see Chapter 5, this volume), perhaps it is sufficient to point out that the approach offered here would provide a relatively parsimonious account. In order to understand the emergence of music, it might be sufficient to assume only one set of evolutionary pressures: the need to communicate verbally. If the specialization of the left auditory cortex discussed above reflects the nervous system's adaptation to this selective pressure, then all we need to assume is that the complementary structural and functional changes of the two hemispheres arose as a consequence of this specialization (see also Ref. 49 for further discussion).

The ideas presented in this chapter, or indeed in this entire volume, will ultimately be shown to be correct, or at least partially so, or in error by the usual slow and meandering path of science. Regardless of the details, however, the hope is that by viewing music as a phenomenon worthy of scientific study, we will enhance both our understanding of music within our species, as well as the workings of our nervous system that make music possible.

## Acknowledgements

I wish to thank all my colleagues and students who contributed to the research described in this chapter. Funding for the work has been provided by grants from the Canadian Institutes of Health Research, the McDonnell-Pew Cognitive Neuroscience Program, and the International Foundation for Music Research.

## References

1. **Evarts, E. V.** (1952) Effect of auditory cortex ablation on frequency discrimination in monkey. *J. Neurophysiol.* 15, 443–8.
2. **Heffner, H. E.** and **B. Masterton** (1978) Contribution of auditory cortex to hearing in the monkey (*Macaca mulatta*). In D. J. Chivers and J. Herbert (eds) Recent Advances in Primatology, vol. 1, *Behaviour.* New York: Academic Press.
3. **Jerison, H. J.** and **W. D. Neff** (1953) Effect of cortical ablation in the monkey on discrimination of auditory patterns. *Fed. Proc.* 12, 237.
4. **Johnsrude, I. J., V. B. Penhune,** and **R. J. Zatorre** (2000) Functional specificity in right human auditory cortex for perceiving pitch direction. *Brain* 123, 155–63.
5. **Penhune, V., R. Zatorre,** and **W. Feindel** (1999) The role of auditory cortex in retention of rhythmic patterns in patients with temporal-lobe removals including Heschl's gyrus. *Neuropsychol.* 37, 315–31.
6. **Zatorre, R. J.** (1988) Pitch perception of complex tones and human temporal-lobe function. *J. Acoust. Soc. Am.* 84(2), 566–72.
7. **Milner, B. A.** (1962) Laterality effects in audition. In V. Mountcastle (ed.) Interhemispheric Relations and Cerebral Dominance. Baltimore, MD: Johns Hopkins Press, 177–95.

8. Divenyi, P. and A. Robinson (1989) Nonlinguistic auditory capabilities in aphasia. *Brain Lang.* 37, 290–326.
9. Robin, D. A., D. Tranel, and H. Damasio (1990) Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. *Brain Lang.* 39, 539–55.
10. Samson, S. and R. J. Zatorre (1994) Contribution of the right temporal lobe to musical timbre discrimination. *Neuropsychol.* 32, 231–40.
11. Sidtis, J. J. and B. T. Volpe (1988) Selective loss of complex-pitch or speech discrimination after unilateral lesion. *Brain Lang.* 34, 235–45.
12. Perry, D. W., R. J. Zatorre, M. Petrides, *et al.* (1999) Localization of cerebral activity during simple singing. *NeuroReport* 10, 3979–84.
13. Perry, D., R. Zatorre, and A. Evans (1996) Co-variation of CBF during singing with vocal fundamental frequency. *NeuroImage* 3(3), S315.
14. Liégeois-Chauvel, C., I. Peretz, M. Babaï, *et al.* (1998) Contribution of different cortical areas in the temporal lobes to music processing. *Brain* 121, 1853–67.
15. Samson, S. and R. J. Zatorre (1988) Melodic and harmonic discrimination following unilateral cerebral excision. *Brain Cogn.* 7, 348–60.
16. Zatorre, R. J. (1985) Discrimination and recognition of tonal melodies after unilateral cerebral excisions. *Neuropsychol.* 23, 31–41.
17. Peretz, I., R. Kolinsky, M. Tramo, *et al.* (1994) Functional dissociations following bilateral lesions of auditory cortex. *Brain* 117, 1283–301.
18. Zatorre, R. J. and S. Samson (1991) Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain* 114, 2403–17.
19. Deutsch, D. (1970) Tones and numbers: specificity of interference in short-term memory. *Science* 168, 1604–5.
20. Semal, C., L. Demany, K. Ueda, *et al.* (1996) Speech versus nonspeech in pitch memory. *J. Acoust. Soc. Am.* 100, 1132–40.
21. Zatorre, R. J., A. C. Evans, and E. Meyer (1994) Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 14(4), 1908–19.
22. Zatorre, R. and J. Binder (2000) Functional and structural imaging of the human auditory system. In A. Toga and J. Mazziota (eds) *Brain Mapping: The Systems.* Los Angeles: Academic Press, 365–402.
23. Perry, D. W., M. Petrides, B. Alivisatos, *et al.* (1993) Functional activation of human frontal cortex during tonal working memory tasks. *Soc. Neurosci. Abstr.* 19, 843.
24. Petrides, M. (1995) Functional organization of the human frontal cortex for mnemonic processing: evidence from neuroimaging studies. *Ann. N.Y. Acad. Sci.* 769, 85–96.
25. Zatorre, R. J., A. C. Evans, E. Meyer, *et al.* (1992) Lateralization of phonetic and pitch processing in speech perception. *Science* 256, 846–9.
26. Démonet, J.-F., F. Chollet, S. Ramsay, *et al.* (1992) The anatomy of phonological and semantic processing in normal subjects. *Brain* 115, 1753–68.
27. Démonet, J.-F., C. Price, R. Wise, and R. S. J. Frackowiack (1994) A PET study of cognitive strategies in normal subjects during language tasks. *Brain* 117, 671–82.
28. Binder, J., J. Frost, T. Hammeke, *et al.* (1997) Human brain language areas identified by functional magnetic resonance imaging. *J. Neurosci.* 17, 353–62.
29. Zatorre, R. J. and A. R. Halpern (1993) Effect of unilateral temporal-lobe excision on perception and imagery of songs. *Neuropsychol.* 31(3), 221–32.

30. **Zatorre, R. J., A. R. Halpern, D. W. Perry, *et al.*** (1996) Hearing in the mind's ear: a PET investigation of musical imagery and perception. *J. Cognit. Neurosci.* 8(1), 29–46.

31. **Halpern, A. R.** and **R. J. Zatorre** (1999) When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. *Cereb. Cortex* 9, 697–704.

32. **Griffiths, T. D., I. S. Johnsrude, J. L. Dean, *et al.*** (1999) A common neural substrate for the analysis of pitch and duration pattern in segmented sound? *NeuroReport* 10, 3825–30.

33. **Hugdahl, K., K. Bronnick, S. Kyllingsbaek *et al.*** (1999) Brain activation during dichotic presentations of consonant-vowel and musical instrument stimuli: a 15O-PET study. *Neuropsychol.* 37, 431–40.

34. **Tervaniemi, M., S. Medvedev, K. Alho, *et al.*** (2000) Lateralized automatic auditory processing of phonetic versus musical information: a PET study. *Hum. Brain Map* 10, 74–9.

35. **Tallal, P., S. Miller,** and **R. Fitch** (1993) Neurobiological basis of speech: a case for the preeminence of temporal processing. *Ann. N.Y. Acad. Sci.* 682, 27–47.

36. **Zatorre, R.** and **P. Belin** (2001) Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–53.

37. **de Charms, R. C., D. T. Blake,** and **M. M. Merzenich** (1998) Optimizing sound features for cortical neurons. *Science* 280, 1439–43.

38. **Phillips, D. P.** (1993) Representation of acoustic events in the primary auditory cortex. *J. Exp. Psychol.: Hum. Percept. Perform.* 19(1), 203–16.

39. **Steinschneider, M., C. Schroeder, J. Arezzo, *et al.*** (1995) Physiologic correlates of the voice onset time boundary in primary auditory cortex of the awake monkey: temporal response patterns. *Brain Lang.* 48, 326–40.

40. **Rauschecker, J. P., B. Tian,** and **M. Hauser** (1995) Processing of complex sounds in the *Macaque* nonprimary auditory cortex. *Science* 268, 111–4.

41. **Eggermont, J.** (1998) Representation of spectral and temporal sound features in three cortical fields of the cat. Similarities outweigh differences. *J. Neurophysiol.* 80, 2743–64.

42. **Liégeois-Chauvel, C., J. de Graaf, V. Laguitton, *et al.*** (1999) Specialization of left auditory cortex for speech perception in man depends on temporal coding. *Cereb. Cortex* 9, 484–96.

43. **Belin, P., M. Zilbovicius, S. Crozier *et al.*** (1998) Lateralization of speech and auditory temporal processing. *J. Cognit. Neurosci.* 10, 536–40.

44. **Penhune, V. B., R. J. Zatorre, J. D. Macdonald, *et al.*** (1996) Interhemispheric anatomical differences in human primary auditory cortex: probabilistic mapping and volume measurement from magnetic resonance scans. *Cereb. Cortex* 6, 661–72.

45. **Anderson, B., B. D. Southern,** and **R. E. Powers** (1999) Anatomic asymmetries of the posterior superior temporal lobes: a postmortem study. *Neuropsychiatry Neuropsychol. Behav. Neurol.* 12, 247–54.

46. **Hutsler, J.** and **M. Gazzaniga** (1996) Acetylcholinesterase staining in human auditory and language cortices—regional variation of structural features. *Cereb. Cortex* 6, 260–70.

47. **Seldon, H.** (1981) Structure of human auditory cortex. II: axon distributions and morphological correlates of speech perception. *Brain Res.* 229, 295–310.

48. **Galuske, R., W. Schlote, H. Bratzke, *et al.*** (2000) Interhemispheric asymmetries of the modular structure in human temporal cortex. *Science* 289, 1946–9.

49. **Justus, T.** and **J. Bharucha** (2002) Music perception and cognition. In H. Pashler and S. Yantis (eds) *Steven's Handbook of Experimental Psychology: Vol. 1. Sensation and Perception.* New York: John Wiley and Sons.

CHAPTER 17

# EXPLORING THE FUNCTIONAL NEUROANATOMY OF MUSIC PERFORMANCE, PERCEPTION, AND COMPREHENSION

LAWRENCE M. PARSONS

## Abstract

This chapter highlights findings by my colleagues and me in four neuroimaging and neurological studies of music performance, perception, and comprehension. These investigations elucidate the neural subsystems supporting musical pitch, melody, harmony, rhythm, tempo, meter, and duration. In a positron emission tomography (PET) study of pianists, a memorized performance of a musical piece was contrasted with that of scales to localize brain areas specifically supporting music. A second PET study assayed brain areas subserving selectively the comprehension of harmony, melody, and rhythm. Musicians sight-read a score while detecting specific melodic, harmonic, or rhythmic errors in its heard performance. In a third PET study, musicians and nonmusicians discriminated pairs of rhythms with respect to pattern, tempo, meter, or duration. Data in these studies implicated the cerebellum in nonmotor, nonsomatic, sensory, or cognitive processing. In a fourth study, neurological patients with degeneration of the cerebellum were found to be impaired in fine discrimination of pitch. Overall, these data suggest that the neural systems underlying music are distributed throughout the left and right cerebral and cerebellar hemispheres, with different aspects of music processed by distinct neural circuits. Also discussed are key issues for interpreting the role in music of brain areas implicated in neuroimaging studies.

**Keywords:** Cerebellum; Neuroimaging studies; Musical performance; Harmony; Melody; Rhythm

In the last decade researchers in the cognitive sciences, experimental psychology, behavioural neurology, ethology, and neuroimaging have significantly advanced the scientific understanding of music, after a long period of slower progress.[1–14] It is not possible to review these important and exciting advances here (see other chapters in this volume for a sampling of such work). This chapter complements these advances by reviewing findings

by my colleagues and me in four recent neuroimaging and neurological studies of music performance, perception, and comprehension. These investigations were attempts to elucidate the neural subsystems supporting major components of music—specifically, pitch, melody, harmony, rhythm, tempo, meter, and duration.

## Functional neuroanatomy of musical performance

Our first experiment focused on musical expression or performance. This experiment, conducted in collaboration with Peter Fox, Donald Hodges, and Justine Sergent, attempted to clarify and extend a pioneering functional neuroimaging study of right-handed piano performance of a sight-read musical score.[15] In this early study, positron emission tomography (PET) was used to measure regional cerebral blood flow, an index of neural activity, in performing pianists. Among the main findings of this study was the observation that cortical areas distinct from, but adjacent to, those underlying language operations were activated during sight-reading. The relationship between music and language has continued to be a target of fruitful investigations.[16–20]

In our study, which built upon the foregoing results, we used PET to study bimanual piano performance of memorized music.[21] By recording brain activity when both hands were equally and concurrently producing music, we examined neural systems in both cerebral hemispheres when the left and right sides were fully involved in performance. In the prior study, although the score was read with both eyes and piano output was heard with both ears, only the right hand performed the music. This feature of the study left unclear which particular right hemispheric areas may be involved in music performance.

In addition, by eliminating musical score reading from scanned task performance, we examined brain activation during a more purely musical performance. Sight-reading a score during performance adds an additional cognitive load not directly related to music *per se.* Indeed, there is a strong belief among musicians that a fully memorized piece, one performed without score reading, engenders a distinctly deeper understanding of the music and more satisfying realization of the piece in performance.

We furthermore contrasted brain activity during the piano performance of a musical piece by J. S. Bach with that during the two-handed performance of scales. Our intent was to reveal the outline of brain areas that are specifically involved in the cognitive, perceptual, and emotional representation and performance of effective, strongly engaging music *per se.* The Bach and scales performances here required movements of nearly comparable frequency and complexity from each hand. (However, the Bach required somewhat more complicated, finely controlled fingering than the scales, and we expected this to be reflected in different patterns of activity when Bach and scales brain states were directly contrasted.)

After giving informed consent, eight right-handed professional musicians participating in three PET conditions performed three times in pseudorandom order (Figure 17.1). In the Bach condition, the third movement of the Italian Concerto (BMV 971) was performed from memory. Scales were executed synchronously (from memory) with both hands at a pace approximating that of the Bach performance. Subjects performed on a full-sized Yamaha P-132 electronic piano and heard the sounds they produced. In all conditions, subjects' eyes were closed and covered.



Figure 17.1 A pianist in a positron emission tomography (PET) scanning session[21] performing the third movement of the Italian Concerto by J. S. Bach. Photograph by Dr Stephan Elleringmann (Aurora/Bilderberg).

PET images, obtained by the $^{15}0$-bolus method, were coregistered on each participant's MRI image and spatially normalized.[22–27] The grand-mean image for each task was compared to control, forming grand-mean images of the task-induced changes in brain activity ($p < 0.001$). Grand-mean images were converted to $z$-score images using the population variance.[28] Activation loci were identified by 3-D center-of-mass address, peak-voxel intensity, and $z$-value, cluster size, and statistical significance.[29] Response coordinates were interpreted[30] in accordance with the Talairach and Tournoux stereotaxic atlas[23] and confirmed with coregistered group-mean MRI.

Only a brief summary of some of the principal findings can be presented here. I will focus on the auditory cortex (the temporal lobe). This is where we detected the primary difference in neural activation for Bach and for scales, as indicated in the analysis of the direct contrast between activity in the two conditions. As illustrated in Figure 17.2, many of the activations were in the anterior secondary auditory association areas, in Brodmann's areas (BA) 22, 21, and 20. Overall, the performance of the Bach activated the temporal lobe more strongly than did the performance of scales. Scales performance activated middle temporal areas bilaterally, with more on the left than on the right (Figure 17.2). By contrast, the performance of the Bach activated superior, middle, and inferior temporal areas bilaterally, with more on the right than on the left.



**Figure 17.1** A pianist in a positron emission tomography (PET) scanning session[21] performing the third movement of the Italian Concerto by J. S. Bach. Photograph by Dr Stephan Elleringmann (Aurora/Bilderberg).

PET images, obtained by the $^{15}$0-bolus method, were coregistered on each participant's MRI image and spatially normalized.[22–27] The grand-mean image for each task was compared to control, forming grand-mean images of the task-induced changes in brain activity ($p < 0.001$). Grand-mean images were converted to *z*-score images using the population variance.[28] Activation loci were identified by 3-D center-of-mass address, peak-voxel intensity, and *z*-value, cluster size, and statistical significance.[29] Response coordinates were interpreted[30] in accordance with the Talairach and Tournoux stereotaxic atlas[23] and confirmed with coregistered group-mean MRI.

Only a brief summary of some of the principal findings can be presented here. I will focus on the auditory cortex (the temporal lobe). This is where we detected the primary difference in neural activation for Bach and for scales, as indicated in the analysis of the direct contrast between activity in the two conditions. As illustrated in Figure 17.2, many of the activations were in the anterior secondary auditory association areas, in Brodmann's areas (BA) 22, 21, and 20. Overall, the performance of the Bach activated the temporal lobe more strongly than did the performance of scales. Scales performance activated middle temporal areas bilaterally, with more on the left than on the right (Figure 17.2). By contrast, the performance of the Bach activated superior, middle, and inferior temporal areas bilaterally, with more on the right than on the left.

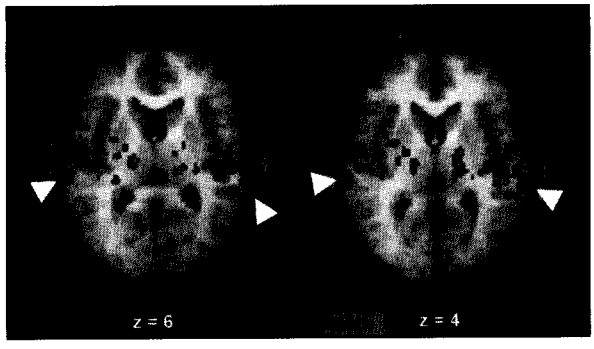



Figure 17.2 Significant blood flow activations[21] in auditory (temporal) cortex specific to playing the Bach score (*red*), specific to playing scales (*blue*), and common to both performances (*cyan*). These are logical contrasts of Bach vs rest and scales vs rest shown in group-averaged PET images overlaid on anatomical magnetic resonance images (MRIs). The *z*-values indicate the axial height of the brain volume relative to the Talairach and Tournoux stereotactic atlas.[23] Throughout, left side of brain images shows the left side of the brain, and vice versa. (See Plate 6 in colour section.)

The stronger activation in the right auditory cortex for the performance of the Bach is consistent with a variety of neuropsychological and neuroimaging results indicating that areas in right auditory association cortex are involved in the reception and expression of melody.[31–36] These findings suggest that some of the important areas specifically representing the higher-order representations of musical meaning, particular in performance, lie in these regions of temporal cortex.

Unexpectedly, there was also strongly correlated activation during the performance of the Bach (Figure 17.3) between right anterior temporal areas (BA 22, superior temporal gyrus) and left posterior lateral cerebellum ($r = +0.77$, $p < 0.001$). Because left cerebellum has its primary connectivity with right cerebral cortex, we hypothesize that during the performance of the Bach, activation in right auditory cortex is related to that in left cerebellum. There was a corresponding, but less robust, correlation between left temporal areas and right posterior cerebellum. This contralateral coactivation of the auditory temporal cortex and lateral cerebellum suggests that they form a distributed circuit of processing. This cerebellar activity appears to be effectively dissociated from motor components and thus appears to contradict the traditional assumption that the cerebellum possesses only motor functions. It is unclear from the present data what function this temporal-cerebellar circuit performs for higher-order representations of music. Interestingly, there is a rapidly growing set of neuroimaging and neurological findings indicating that the cerebellum possesses functions other than those classically attributed to it. These new data suggest that the cerebellum is involved in some way in a variety of sensory and cognitive tasks.[37–45] The present findings, and others discussed below, appear to extend this new view of cerebellar function to the perceptual representation of music.

In summary, these data assist in outlining some of the neural systems that support musical performance and understanding. These findings appear to confirm or extend other neuroimaging and neurological observations and pave the way for new, more focused, research.




Figure 17.2 Significant blood flow activations[21] in auditory (temporal) cortex specific to playing the Bach score (*red*), specific to playing scales (*blue*), and common to both performances (*cyan*). These are logical contrasts of Bach vs rest and scales vs rest shown in group-averaged PET images overlaid on anatomical magnetic resonance images (MRIs). The *z*-values indicate the axial height of the brain volume relative to the Talairach and Tournoux stereotactic atlas.[23] Throughout, left side of brain images shows the left side of the brain, and vice versa. (See Plate 6 in colour section.)

The stronger activation in the right auditory cortex for the performance of the Bach is consistent with a variety of neuropsychological and neuroimaging results indicating that areas in right auditory association cortex are involved in the reception and expression of melody.[31–36] These findings suggest that some of the important areas specifically representing the higher-order representations of musical meaning, particular in performance, lie in these regions of temporal cortex.

Unexpectedly, there was also strongly correlated activation during the performance of the Bach (Figure 17.3) between right anterior temporal areas (BA 22, superior temporal gyrus) and left posterior lateral cerebellum ($r = +0.77$, $p < 0.001$). Because left cerebellum has its primary connectivity with right cerebral cortex, we hypothesize that during the performance of the Bach, activation in right auditory cortex is related to that in left cerebellum. There was a corresponding, but less robust, correlation between left temporal areas and right posterior cerebellum. This contralateral coactivation of the auditory temporal cortex and lateral cerebellum suggests that they form a distributed circuit of processing. This cerebellar activity appears to be effectively dissociated from motor components and thus appears to contradict the traditional assumption that the cerebellum possesses only motor functions. It is unclear from the present data what function this temporal-cerebellar circuit performs for higher-order representations of music. Interestingly, there is a rapidly growing set of neuroimaging and neurological findings indicating that the cerebellum possesses functions other than those classically attributed to it. These new data suggest that the cerebellum is involved in some way in a variety of sensory and cognitive tasks.[37–45] The present findings, and others discussed below, appear to extend this new view of cerebellar function to the perceptual representation of music.

In summary, these data assist in outlining some of the neural systems that support musical performance and understanding. These findings appear to confirm or extend other neuroimaging and neurological observations and pave the way for new, more focused, research.

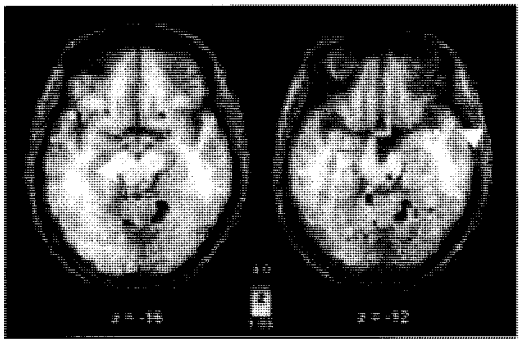


Figure 17.3 Correlated blood flow activations in right anterior auditory cortex and left cerebellum (see *arrows*) as subjects played the Bach score.[21] These are group-averaged PET images for Bach (contrasted with rest) overlaid on anatomical MRIs. PET data are *z*-scores displayed on a colour scale ranging from 1.65 (*yellow*; $p < 0.1$) to 4.0 (*red*; $p < 0.0001$). (See Plate 7 in colour section.)

## Functional neuroanatomy of the comprehension of harmony, melody, and rhythm

We pursued these results in our second functional neuroimaging study, in which we attempted to map the brain basis of the comprehension of individual principal components of music *per se*.[46] Using the same PET methodology as in our first study, my colleagues and I imaged eight right-handed, male Ph.D. faculty musicians (none of whom participated in our first study). These subjects are experts at monitoring musical scores while listening for accurate performance. Each subject was competent on several instruments, with primary specializations in either winds, strings, voice, or keyboard.

During the experiment, these subjects detected errors in a computer-generated instrumental performance of unfamiliar Bach chorales as they read the corresponding score. The subjects made no overt response during a trial; rather they acknowledged covertly the location and presence of specific errors. The errors, which we implanted on each trial, were exclusively either rhythmic, harmonic, or melodic. The errors occurred unpredictably every two to three beats. On rhythm trials, errors occurred only in the performed rhythm of a note in one of the four voices but were otherwise correct. The rhythm errors changed either the duration or onset of a written note only (leaving unchanged the other aspects of rhythm such as meter and tempo). On melody trials, errors perturbed a written note in the highest voice by raising or lowering it a chromatic half step; otherwise they were accurate. On harmony trials, the errors could occur in any one of the four voices but were otherwise correct. The harmony errors perturbed the written note by raising or lowering it a chromatic half step. Pilot studies were used to develop stimuli that equated for difficulty across the three conditions. In prescan and postscan performance, subjects were accurately detecting approximately 80 per cent of the errors in each task.

After practicing the tasks, each subject was scanned performing each of the following five trials twice in pseudorandom order. On harmony, rhythm, and melody trials, subjects

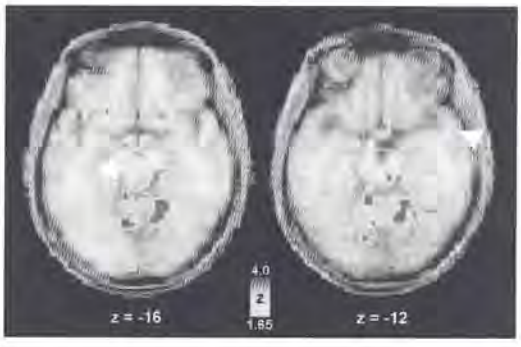



**Figure 17.3** Correlated blood flow activations in right anterior auditory cortex and left cerebellum (see *arrows*) as subjects played the Bach score.[21] These are group-averaged PET images for Bach (contrasted with rest) overlaid on anatomical MRIs. PET data are *z*-scores displayed on a colour scale ranging from 1.65 (*yellow*; $p < 0.1$) to 4.0 (*red*; $p < 0.0001$). (See Plate 7 in colour section.)

## Functional neuroanatomy of the comprehension of harmony, melody, and rhythm

We pursued these results in our second functional neuroimaging study, in which we attempted to map the brain basis of the comprehension of individual principal components of music *per se*.[46] Using the same PET methodology as in our first study, my colleagues and I imaged eight right-handed, male Ph.D. faculty musicians (none of whom participated in our first study). These subjects are experts at monitoring musical scores while listening for accurate performance. Each subject was competent on several instruments, with primary specializations in either winds, strings, voice, or keyboard.

During the experiment, these subjects detected errors in a computer-generated instrumental performance of unfamiliar Bach chorales as they read the corresponding score. The subjects made no overt response during a trial; rather they acknowledged covertly the location and presence of specific errors. The errors, which we implanted on each trial, were exclusively either rhythmic, harmonic, or melodic. The errors occurred unpredictably every two to three beats. On rhythm trials, errors occurred only in the performed rhythm of a note in one of the four voices but were otherwise correct. The rhythm errors changed either the duration or onset of a written note only (leaving unchanged the other aspects of rhythm such as meter and tempo). On melody trials, errors perturbed a written note in the highest voice by raising or lowering it a chromatic half step; otherwise they were accurate. On harmony trials, the errors could occur in any one of the four voices but were otherwise correct. The harmony errors perturbed the written note by raising or lowering it a chromatic half step. Pilot studies were used to develop stimuli that equated for difficulty across the three conditions. In prescan and postscan performance, subjects were accurately detecting approximately 80 per cent of the errors in each task.

After practicing the tasks, each subject was scanned performing each of the following five trials twice in pseudorandom order. On harmony, rhythm, and melody trials, subjects

sight-read the score while listening to its errorful performance. On passive listening trials, subjects listened to a chorale performed without errors; they were not reading a score. On each scanned trial of the harmony, rhythm, melody, and passive listening tasks, subjects heard a different Bach chorale. On rest baseline control trials, subjects fixated on a fixation point (neither reading a score nor listening to music). The image data during the primary three tasks were contrasted with those from listening to the errorless performance of a Bach chorale without reading a score, as well as to the rest baseline control. Analyses were conducted as described earlier for our preceding PET study.

To highlight some of our principal findings, the melody, harmony, and rhythm conditions each showed a distinct pattern of distributed, statistically significant brain activity compared to passive listening control ($p < 0.001$). Often, each condition activated different subareas of a particular major brain area. In broad anatomical terms, score reading and comprehension of melodic, rhythmic, and harmonic features of a piece were supported by processes in both cerebral hemispheres. More specifically, melodic comprehension activated each hemisphere equally, whereas harmonic and rhythmic comprehension activated more of the left hemisphere than the right one. In addition, the rhythm condition activated comparatively few brain areas outside the cerebellum.

All three tasks strongly activated bilateral cerebellum, primarily the lateral hemispheres, with little overlap in activated area across the three tasks (Figure 17.4). Given the lack of explicit (and apparently implicit) movement, this cerebellar activity is likely related to sensory or cognitive processing, not to motoric processing (see below). Furthermore, activation of cerebellum in the rhythm condition was twice that in harmony and melody conditions. An involvement of the cerebellum in some way in processing rhythm is consistent with recent neurological and neuroimaging data.[42,47–49] Cerebellar activations during the processing of melody and harmony are relatively novel findings but are consistent with data above from our first study and with other recent studies showing lateral cerebellar activation for processing of tonal information (see later).

One area commonly activated by the harmony, melody, and rhythm tasks was in right fusiform gyrus. Because score reading was a central process common to all three conditions,

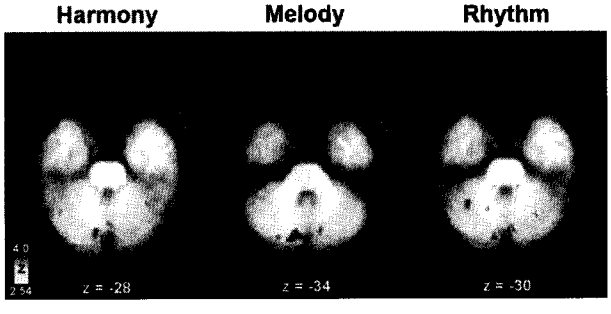



**Figure 17.4** Significant blood flow activations in the cerebellum as musicians sight-read a J. S. Bach chorale score and listen for performance errors in harmony, melody, or rhythm.[46] These are group-averaged PET images for each task (contrasted with passive listening control) overlaid on anatomical MRIs. PET data are *z*-scores displayed on a colour scale ranging from 1.96 (*yellow*; $p < 0.05$) to 4.0 (*red*; $p < 0.0001$). (See Plate 8 in colour section.)

sight-read the score while listening to its errorful performance. On passive listening trials, subjects listened to a chorale performed without errors; they were not reading a score. On each scanned trial of the harmony, rhythm, melody, and passive listening tasks, subjects heard a different Bach chorale. On rest baseline control trials, subjects fixated on a fixation point (neither reading a score nor listening to music). The image data during the primary three tasks were contrasted with those from listening to the errorless performance of a Bach chorale without reading a score, as well as to the rest baseline control. Analyses were conducted as described earlier for our preceding PET study.

To highlight some of our principal findings, the melody, harmony, and rhythm conditions each showed a distinct pattern of distributed, statistically significant brain activity compared to passive listening control ($p < 0.001$). Often, each condition activated different subareas of a particular major brain area. In broad anatomical terms, score reading and comprehension of melodic, rhythmic, and harmonic features of a piece were supported by processes in both cerebral hemispheres. More specifically, melodic comprehension activated each hemisphere equally, whereas harmonic and rhythmic comprehension activated more of the left hemisphere than the right one. In addition, the rhythm condition activated comparatively few brain areas outside the cerebellum.

All three tasks strongly activated bilateral cerebellum, primarily the lateral hemispheres, with little overlap in activated area across the three tasks (Figure 17.4). Given the lack of explicit (and apparently implicit) movement, this cerebellar activity is likely related to sensory or cognitive processing, not to motoric processing (see below). Furthermore, activation of cerebellum in the rhythm condition was twice that in harmony and melody conditions. An involvement of the cerebellum in some way in processing rhythm is consistent with recent neurological and neuroimaging data.[42,47–49] Cerebellar activations during the processing of melody and harmony are relatively novel findings but are consistent with data above from our first study and with other recent studies showing lateral cerebellar activation for processing of tonal information (see later).

One area commonly activated by the harmony, melody, and rhythm tasks was in right fusiform gyrus. Because score reading was a central process common to all three conditions,

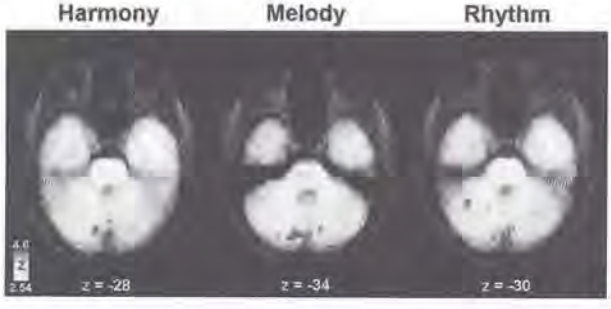



**Figure 17.4** Significant blood flow activations in the cerebellum as musicians sight-read a J. S. Bach chorale score and listen for performance errors in harmony, melody, or rhythm.[46] These are group-averaged PET images for each task (contrasted with passive listening control) overlaid on anatomical MRIs. PET data are *z*-scores displayed on a colour scale ranging from 1.96 (*yellow*; $p < 0.05$) to 4.0 (*red*; $p < 0.0001$). (See Plate 8 in colour section.)

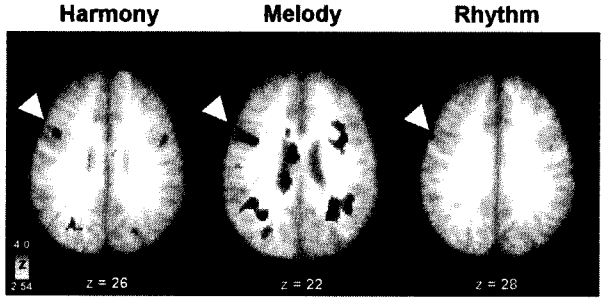



Figure 17.5 Activations in different subareas of a major cytoarchitectonic area of frontal cortex (BA 44) as musicians sight-read a J. S. Bach chorale score and listen for performance errors in harmony, melody, or rhythm.[46] These are group mean PET images for each task (contrasted with passive listening control) overlaid on MRIs. PET data are *z*-scores displayed on a colour scale ranging from 2.54 (*yellow*; $p < 0.01$) to 4.0 (*red*; $p < 0.0001$). (See Plate 9 in colour section.)

the observed right fusiform activation may implicate that area in the visual processing of musical notes. Such a function would be analogous to visual word processing by left fusiform.[50–53] These implied parallel locations for comparable reading functions in music and language are consistent with other functional and structural correspondences.[16–20]

Each of the three conditions activated different subareas of two major areas in frontal cortex, a lateral region, BA 6, and an inferior lateral region, BA 44/45 (Figure 17.5). These activations were bilateral, although stronger on the left side. In both of these major areas, the rhythm task activated superior subareas, the melody task activated inferior subareas, and the harmony task activated subareas between those two activations. These two major areas are currently associated with sequencing motor behaviour and with recognizing actions.[54–59] However, there was no activation detected in supplementary motor area (SMA) medial BA 6, as has been found in neuroimaging studies of song imagery.[7] In addition, there was no activation in primary motor areas. Therefore, the activation of these areas in the present conditions does not seem related to motor preparation as such, whether for subvocal singing or for imagined playing of an instrument. It is possible that these activations are involved in translating the notes to abstract motor-auditory codes. Such a translation process may be a natural component of expert score reading.

The melody, harmony, and rhythm tasks each also produced distinct patterns of activity in auditory cortex (Figure 17.6). During the melody condition, there were strong activations in secondary auditory association areas, specifically in bilateral superior (BA 22) and bilateral, strongly right, middle temporal (BA 21) cortex. The harmony task produced bilateral activation, more on the left side than the right, in middle (BA 21 and 20), posterior (BA 37), and superior (BA 22) temporal cortex. The rhythm condition produced only minor activation in left inferior (BA 20) and bilateral middle (BA 21) temporal cortex. Superior temporal activation for melodic and harmony tasks, but not rhythmic processing, is consistent with neuropsychological data showing that superior temporal lesions spare rhythmic processing but cause impaired melodic perception.[6,11,60] The activation of such secondary auditory regions bilaterally during these comprehension tasks is also in accord

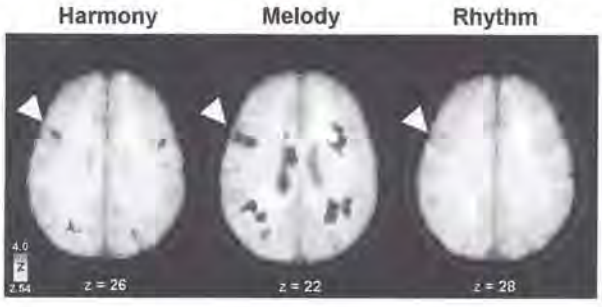



**Figure 17.5** Activations in different subareas of a major cytoarchitectonic area of frontal cortex (BA 44) as musicians sight-read a J. S. Bach chorale score and listen for performance errors in harmony, melody, or rhythm.[46] These are group mean PET images for each task (contrasted with passive listening control) overlaid on MRIs. PET data are *z*-scores displayed on a colour scale ranging from 2.54 (*yellow*; $p < 0.01$) to 4.0 (*red*; $p < 0.0001$). (See Plate 9 in colour section.)

the observed right fusiform activation may implicate that area in the visual processing of musical notes. Such a function would be analogous to visual word processing by left fusiform.[50–53] These implied parallel locations for comparable reading functions in music and language are consistent with other functional and structural correspondences.[16–20]

Each of the three conditions activated different subareas of two major areas in frontal cortex, a lateral region, BA 6, and an inferior lateral region, BA 44/45 (Figure 17.5). These activations were bilateral, although stronger on the left side. In both of these major areas, the rhythm task activated superior subareas, the melody task activated inferior subareas, and the harmony task activated subareas between those two activations. These two major areas are currently associated with sequencing motor behaviour and with recognizing actions.[54–59] However, there was no activation detected in supplementary motor area (SMA) medial BA 6, as has been found in neuroimaging studies of song imagery.[7] In addition, there was no activation in primary motor areas. Therefore, the activation of these areas in the present conditions does not seem related to motor preparation as such, whether for subvocal singing or for imagined playing of an instrument. It is possible that these activations are involved in translating the notes to abstract motor-auditory codes. Such a translation process may be a natural component of expert score reading.

The melody, harmony, and rhythm tasks each also produced distinct patterns of activity in auditory cortex (Figure 17.6). During the melody condition, there were strong activations in secondary auditory association areas, specifically in bilateral superior (BA 22) and bilateral, strongly right, middle temporal (BA 21) cortex. The harmony task produced bilateral activation, more on the left side than the right, in middle (BA 21 and 20), posterior (BA 37), and superior (BA 22) temporal cortex. The rhythm condition produced only minor activation in left inferior (BA 20) and bilateral middle (BA 21) temporal cortex. Superior temporal activation for melodic and harmony tasks, but not rhythmic processing, is consistent with neuropsychological data showing that superior temporal lesions spare rhythmic processing but cause impaired melodic perception.[6,11,60] The activation of such secondary auditory regions bilaterally during these comprehension tasks is also in accord

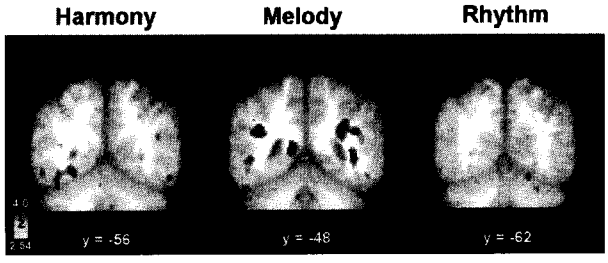



Figure 17.6 Activations in different subareas of auditory cortex as musicians sight-read a J. S. Bach chorale score and listen for performance errors in melody, harmony, or rhythm.[46] These are group mean PET images for each task (contrasted with passive listening control) overlaid on MRIs. PET data are *z*-scores displayed on a colour scale ranging from 2.54 (*yellow*; $p < 0.05$) to 4.0 (*red*; $p < 0.0001$). (See Plate 10 in colour section.)

with bilateral activation of those areas during the perception or performance of musical pieces and scales in our prior PET study.

The distributed pattern of activations observed in this investigation delineates some of the neural systems subserving musical comprehension. This study may provide an early glimpse of the functional neuroanatomy of expert musicians' comprehension of musical structure. These data confirm or extend prior understanding of the distributed system of brain areas supporting pitch and melody perception, auditory mental imagery, unimanual piano performance from a sight-read score, and bi-manual piano performance of a memorized score. These findings are also consistent with and complement those reported in a contemporaneous neuroimaging study by Platel, Price, Baron *et al.*[10] in which nonmusicians attended to either the timbre, melody, or rhythm properties of random-order, nonmusical tone sequences. Interestingly, the latter study reported that musical timbre, not studied here, was subserved by areas in right superior frontal, middle frontal, and precentral cortex.

The foregoing results suggest that music as a whole is represented in mechanisms widely distributed throughout the brain. However, different parts of the brain are activated when one pays close attention to different aspects of a piece of music. This observation is perhaps not surprising since music is a complex, multifaceted stimulus and activity. At the same time, the data illustrate the power of directed attention to change brain activity. The data also suggest that the specific subareas of major areas, such as cerebellum, and frontal and temporal cortex, perform different functions for particular subtasks of musical comprehension. More generally, the distributed neural system for musical comprehension revealed here can be compared in further studies to the better-known distributed neural systems for language tasks. Such studies can begin to identify independent and shared neural mechanisms for music and language and evaluate the scope of application for specific neural mechanisms.

## Functional neuroanatomy of musical rhythm

The next study builds upon the findings of the preceding studies and aims to provide a more complete account of the functional neuroanatomy of the mental representation of

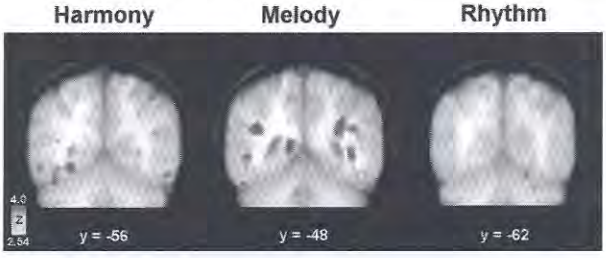



**Figure 17.6** Activations in different subareas of auditory cortex as musicians sight-read a J. S. Bach chorale score and listen for performance errors in melody, harmony, or rhythm.[46] These are group mean PET images for each task (contrasted with passive listening control) overlaid on MRIs. PET data are *z*-scores displayed on a colour scale ranging from 2.54 (*yellow*; $p < 0.05$) to 4.0 (*red*; $p < 0.0001$). (See Plate 10 in colour section.)

with bilateral activation of those areas during the perception or performance of musical pieces and scales in our prior PET study.

The distributed pattern of activations observed in this investigation delineates some of the neural systems subserving musical comprehension. This study may provide an early glimpse of the functional neuroanatomy of expert musicians' comprehension of musical structure. These data confirm or extend prior understanding of the distributed system of brain areas supporting pitch and melody perception, auditory mental imagery, unimanual piano performance from a sight-read score, and bi-manual piano performance of a memorized score. These findings are also consistent with and complement those reported in a contemporaneous neuroimaging study by Platel, Price, Baron *et al.*[10] in which nonmusicians attended to either the timbre, melody, or rhythm properties of random-order, nonmusical tone sequences. Interestingly, the latter study reported that musical timbre, not studied here, was subserved by areas in right superior frontal, middle frontal, and precentral cortex.

The foregoing results suggest that music as a whole is represented in mechanisms widely distributed throughout the brain. However, different parts of the brain are activated when one pays close attention to different aspects of a piece of music. This observation is perhaps not surprising since music is a complex, multifaceted stimulus and activity. At the same time, the data illustrate the power of directed attention to change brain activity. The data also suggest that the specific subareas of major areas, such as cerebellum, and frontal and temporal cortex, perform different functions for particular subtasks of musical comprehension. More generally, the distributed neural system for musical comprehension revealed here can be compared in further studies to the better-known distributed neural systems for language tasks. Such studies can begin to identify independent and shared neural mechanisms for music and language and evaluate the scope of application for specific neural mechanisms.

## Functional neuroanatomy of musical rhythm

The next study builds upon the findings of the preceding studies and aims to provide a more complete account of the functional neuroanatomy of the mental representation of

musical rhythm. We examined the selective perception of individual principal components of musical rhythm—that is, pattern (phrasing), tempo, meter, and duration. In addition, in order to explore the effect of musical talent and training on the neural systems supporting musical rhythm perception, expert musicians were compared to nonmusicians.

Studies of brain-damaged patients have reported that the perception of rhythm can be impaired without affecting melodic processing, suggesting that there is a distinct neural system for the representation of rhythm.[61] Some studies have been in-conclusive as to whether rhythm was a left hemispheric or bilateral neural process.[62,63] Interestingly, recent fMRI data suggest that memory for familiar, regular rhythms (with whole-number interval ratios) is supported by left frontal (BA 6) and parietal cortex and right anterior cerebellum.[48] On the other hand, these data suggested that memory for more unfamiliar rhythms (with fractionated interval ratios) was supported by right prefrontal, frontal (BA 6), and parietal areas, along with bilateral posterior cerebellum.

The involvement of the cerebellum in the processing of musical rhythm is consistent with data from our second PET study (see earlier) and those of two other studies.[10,42] This implication is congruent with neurological findings that cerebellar patients are impaired at the fine perception of temporal features of auditory stimuli.[47]

Involvement of parietal and prefrontal cortex in rhythm processing[48] is also consistent with data from our second PET study. In addition, this activity is congruent with recent neurological data implicating an inferior parietal-prefrontal circuit that supports time-dependent attention and working memory functions necessary for perception of duration of auditory stimuli.[64] Of particular note in this context is a study of unilateral temporal corticectomy patients in which evidence was found for a dissociation between the areas processing meter as opposed to rhythm pattern, with anterior superior temporal areas implicated for meter.[11]

In the present study, Michael Thaut and I explored the principal components of rhythm: duration, pattern, tempo (dynamically increasing or decreasing rate), and meter (e.g. the differing periodicities of 3/4, 4/4, 5/4, 5/8, 7/8, and 9/8). Using the same PET methods as in our preceding studies, we imaged nonmusicians and musicians, making covert discriminations of pairs of rhythmic auditory patterns. Each scanned set of trials required the subject to focus on a single one of four principal components of rhythm. We also included a pitch discrimination control task in which subjects compared pairs of auditory sequences. In the pitch discrimination task it was possible for a single note to vary in pitch in one member of the stimulus pair. The pitch discrimination stimuli possessed a metronomic rhythm (i.e. all notes had the same duration). As a control task, the pitch stimuli required comparable auditory perception, working memory, and comparison and decision processes as the rhythm tasks, without requiring the processing of rhythmic features of duration, meter, tempo, or pattern.

Participating in the study were five musicians, with (at least) an undergraduate university degree in music, and five nonmusicians, with no music training or performance experience beyond childhood. All 10 subjects were right-handed adult males. The stimuli (apart from the pitch control) were always 440-Hz computer-generated piano timbre sounds of 231- or 462-ms duration. The interval between tones in a stimulus sequence was a multiple of 231 ms. The stimuli on the duration, pattern, and pitch control trials were modeled

on the standardized Seashore Test.[65] The stimuli on the tempo and meter trials were modeled on the Gordon Musical Aptitude Profile.[66] These stimulus materials were adjusted (i.e. in tempo or accentuation) so that the stimuli on the tempo, phrasing, duration, meter, and pitch control trials were as similar as possible. Different conditions varied with respect to the exact nature of the task being performed, with stimuli being in almost all respects constant across conditions. Although tempo, pattern, and meter stimuli had more tones per sequence, between 9 and 12 events, approximately the same total number of tones per scanned interval were presented in the duration task as in the tempo, pattern, and meter tasks. A 'different' trial contained the same number of events (and accents) and was of the same total duration as the 'same' trials (by playing them at different tempi), except in the duration condition. In the pitch task, the tone alternated between 415, 440, and 466 Hz in sequences of 12 462-ms (quarter note) tones, without rests. Across all conditions, the intersequence interval was 1000 ms, and the intertrial interval was 1750 ms; the trial time was 14 s on average.

Stimuli were adjusted to produce comparable mean accuracy of musicians and of nonmusicians on each task. Each subject subsequently performed 11 PET trials: two trials each of the pattern task, duration task, meter task, tempo task, and pitch control task, as well as one rest trial. The subject's eyes were closed on all trials. During each 60 s of task there were four covert discrimination judgments. For half of the trials, the pair of stimuli was identical. After the PET session, subjects replicated the trials from the PET session, overtly indicating their responses. Their average performance across the latter trials confirmed that task difficulty was equated across conditions and across groups.

To highlight some of our principal findings, for both nonmusicians and musicians, distinct patterns of brain activity were detected for the discrimination of pitch and each musical rhythm. There were also distinct differences between nonmusicians and musicians during each discrimination.

For pitch (contrasted with rest), superior temporal areas (BA 21/22) were activated on the right in nonmusicians and on the left in musicians (Figure 17.7A and D). However, middle and inferior temporal areas (BA 21/20) were activated on the right in both groups (Figure 17.7B and E). These temporal activations confirm prior neurological and neuroimaging studies, including those for the comprehension of melody above.[6,7,11,34–36]

There were also strong cerebellar activations during pitch (Figure 17.7C and F), again confirming our earlier PET data. The cerebellar activations are not likely related to motor activity since no overt motor activity occurred and there was little or no activity detected in neocortical motor areas. Instead, the cerebellar activity seems to support perceptual or cognitive processing of pitch (or melody). Interestingly, the nonmusicians showed strong predominantly left-sided activation in posterior lateral cerebellar hemispheres (Figure 17.7C). The left cerebellum projects to the right cerebral cortex, so the activation in left cerebellum is likely supporting the right temporal activation for pitch. However, musicians showed bilateral cerebellar activation (Figure 17.7F), apparently supporting bilateral temporal activation.

The key rhythm-specific activations were as follows. First, certain activations were generally present for each rhythm task and for both nonmusicians and musicians, such as basal ganglia and cingulate cortex. The cingulate activations are likely involved in attention functions required in the task.[67–70]

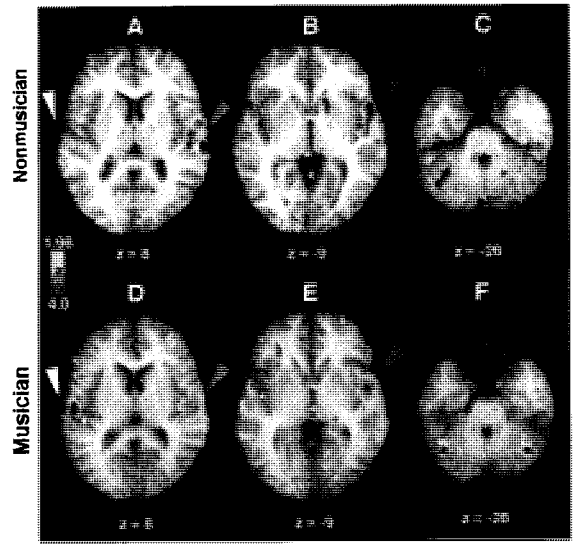



Figure 17.7 Activations in superior and middle temporal cortex (A, B, D, and E) and in cerebellum (C and F) as nonmusicians and musicians discriminate pitch sequences.[49] These images show group mean PET data for each task (contrasted with rest) overlaid on MRIs. PET data are *z*-scores displayed on a colour scale ranging from 1.96 (*yellow*; $p < 0.1$) to 4.0 (*red*; $p < 0.0001$). (See Plate 11 in colour section.)

For the tempo, pattern, and meter discriminations, there was a tendency for activation to be detected in nonmusicians that was not present for musicians. For tempo discrimination, nonmusicians showed activation in right medial frontal cortex (BA 9) and bilateral posterior lateral cerebellum, but this activation was very much weaker or absent in musicians (Figures 17.8 and 17.12). For pattern discrimination, nonmusicians showed activation in midbrain and bilateral posterior lateral cerebellum; however, this activation was absent and weaker in musicians (Figures 17.9 and 17.12). For meter discrimination, nonmusicians showed activation in right inferior frontal cortex (BA 47) and bilateral posterior lateral cerebellum (Figure 17.10 and 17.12), with little or no such activation in these areas for musicians. By contrast, during duration discrimination musicians showed bilateral medial frontal cortex (BA 9), bilateral inferior parietal cortex (BA 40), and strong activation in bilateral posterior lateral cerebellum (Figures 17.11 and 17.12). These activations were absent or very much weaker in nonmusicians.

As with pitch discrimination, the strong cerebellar activations during rhythm discriminations are not likely due to motoric processing *per se* but to a role in supporting nonmotor sensory or cognitive processing.[38–45] In addition, in the meter, tempo, and pattern discriminations, cerebellar activity is much stronger for nonmusicians than for musicians, but the reverse is true for the duration discrimination. Interestingly, musicians found discriminating the duration of rhythmic sequences more novel than discriminating meter, tempo, and pattern. By contrast, the nonmusicians found discriminating the overall duration of rhythmic sequences less novel than discriminating other rhythm features. Thus, the

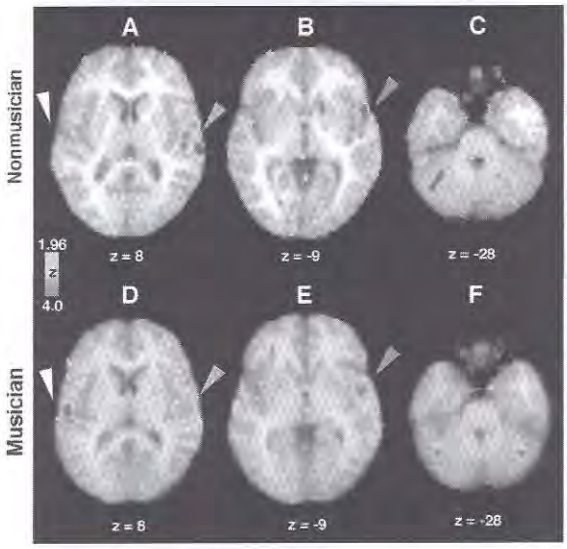



**Figure 17.7** Activations in superior and middle temporal cortex (A, B, D, and E) and in cerebellum (C and F) as nonmusicians and musicians discriminate pitch sequences.[49] These images show group mean PET data for each task (contrasted with rest) overlaid on MRIs. PET data are *z*-scores displayed on a colour scale ranging from 1.96 (*yellow*; $p < 0.1$) to 4.0 (*red*; $p < 0.0001$). (See Plate 11 in colour section.)

For the tempo, pattern, and meter discriminations, there was a tendency for activation to be detected in nonmusicians that was not present for musicians. For tempo discrimination, nonmusicians showed activation in right medial frontal cortex (BA 9) and bilateral posterior lateral cerebellum, but this activation was very much weaker or absent in musicians (Figures 17.8 and 17.12). For pattern discrimination, nonmusicians showed activation in midbrain and bilateral posterior lateral cerebellum; however, this activation was absent and weaker in musicians (Figures 17.9 and 17.12). For meter discrimination, nonmusicians showed activation in right inferior frontal cortex (BA 47) and bilateral posterior lateral cerebellum (Figure 17.10 and 17.12), with little or no such activation in these areas for musicians. By contrast, during duration discrimination musicians showed bilateral medial frontal cortex (BA 9), bilateral inferior parietal cortex (BA 40), and strong activation in bilateral posterior lateral cerebellum (Figures 17.11 and 17.12). These activations were absent or very much weaker in nonmusicians.

As with pitch discrimination, the strong cerebellar activations during rhythm discriminations are not likely due to motoric processing *per se* but to a role in supporting nonmotor sensory or cognitive processing.[38–45] In addition, in the meter, tempo, and pattern discriminations, cerebellar activity is much stronger for nonmusicians than for musicians, but the reverse is true for the duration discrimination. Interestingly, musicians found discriminating the duration of rhythmic sequences more novel than discriminating meter, tempo, and pattern. By contrast, the nonmusicians found discriminating the overall duration of rhythmic sequences less novel than discriminating other rhythm features. Thus, the

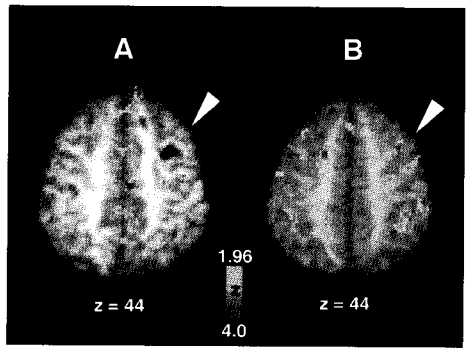



Figure 17.8 Activations in left medial frontal cortex (BA 9) for nonmusicians (A) as compared to musicians (B) during the discrimination of musical tempi.[49] Shown are group mean PET images for each task (contrasted with rest) overlaid on MRIs. (See Plate 12 in colour section.)

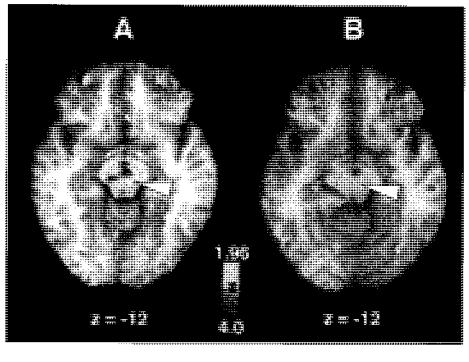



Figure 17.9 Activations in midbrain for nonmusicians (A) as compared to musicians (B) during the discrimination of musical pattern.[49] These are group-averaged PET images for each task (contrasted with rest) overlaid on anatomical MRIs. (See Plate 13 in colour section.)

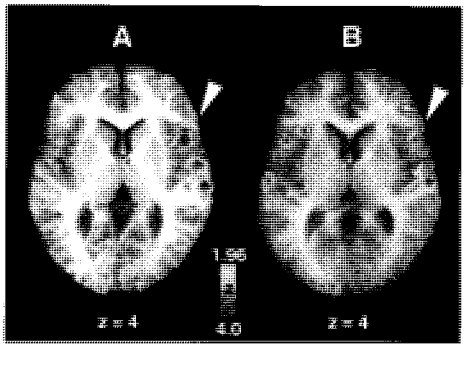



Figure 17.10 Activations in right inferior frontal cortex (BA 47) for nonmusicians (A) as compared to musicians (B) during discrimination of musical meter.[49] Shown are group-averaged PET images for each task (contrasted with rest) over-laid on anatomical MRIs. (See Plate 14 in colour section.)

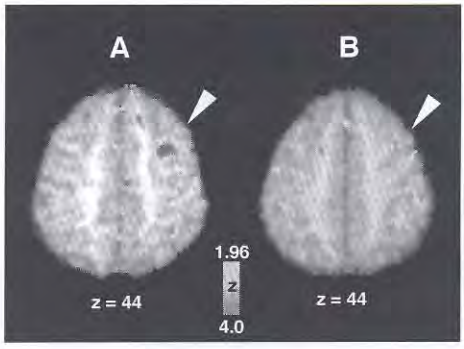



**Figure 17.8** Activations in left medial frontal cortex (BA 9) for nonmusicians (A) as compared to musicians (B) during the discrimination of musical tempi.[49] Shown are group mean PET images for each task (contrasted with rest) overlaid on MRIs. (See Plate 12 in colour section.)

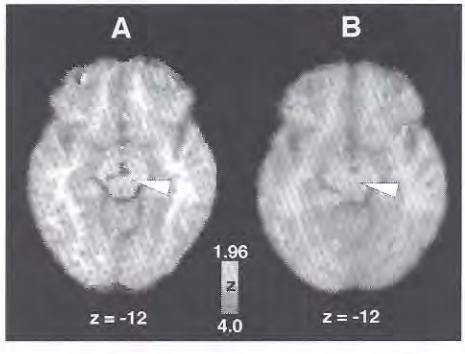



**Figure 17.9** Activations in midbrain for nonmusicians (A) as compared to musicians (B) during the discrimination of musical pattern.[49] These are group-averaged PET images for each task (contrasted with rest) overlaid on anatomical MRIs. (See Plate 13 in colour section.)

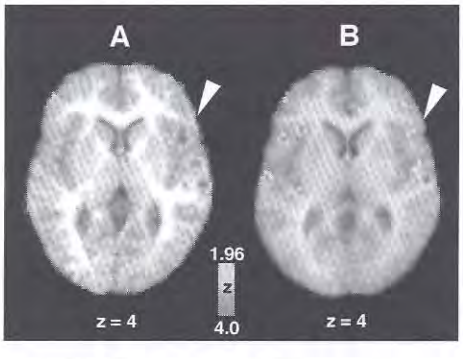



**Figure 17.10** Activations in right inferior frontal cortex (BA 47) for nonmusicians (A) as compared to musicians (B) during discrimination of musical meter.[49] Shown are group-averaged PET images for each task (contrasted with rest) over-laid on anatomical MRIs. (See Plate 14 in colour section.)

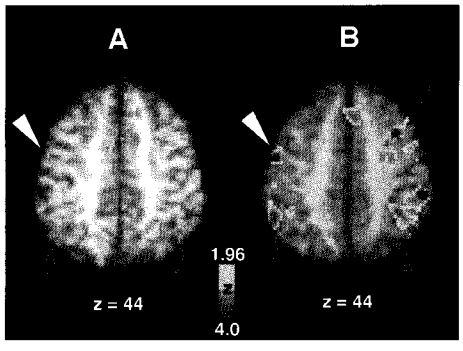



Figure 17.11 Activations in bilateral medial frontal cortex (BA 9) and bilateral inferior parietal cortex (BA 40) for musicians (B) as compared to nonmusicians (A) during the discrimination of sequence durations.[49] Shown are group-averaged PET images for each task (contrasted with rest) overlaid on anatomical MRIs. (See Plate 15 in colour section.)

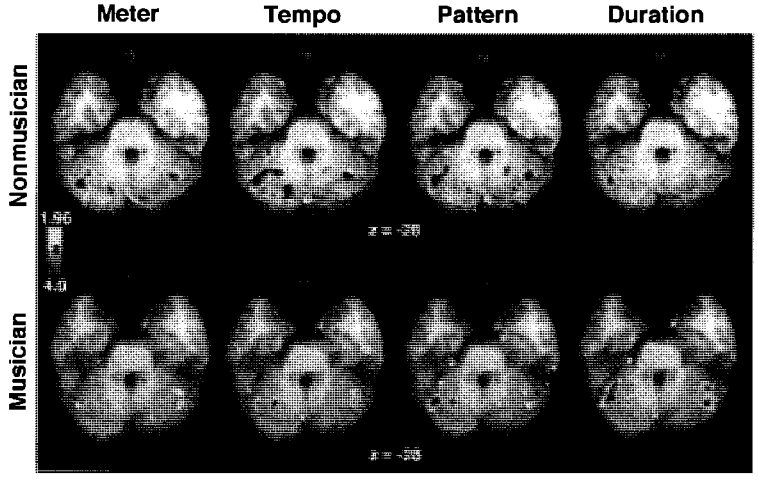



Figure 17.12 Activations in posterior cerebellum as nonmusicians and musicians discriminate meter, tempo, pattern, and duration.[49] Shown are group-mean PET images for each task (contrasted with rest) overlaid on MRIs. (See Plate 16 in colour section.)

increase in cerebellar activation appears to be associated with greater novelty of rhythm discrimination, consistent with other neuroimaging findings[71] and new hypotheses of cerebellar function.[37]

The activations in prefrontal cortical areas BAs 9, 10, and 47 are likely supporting working memory processes.[48] The activation in prefrontal and inferior parietal areas follows similar indications in the recent neurological study of tone duration discrimination.[64] We speculate that activations in temporal, frontal, and midbrain may be involved in processing the temporal grouping structure of rhythms.

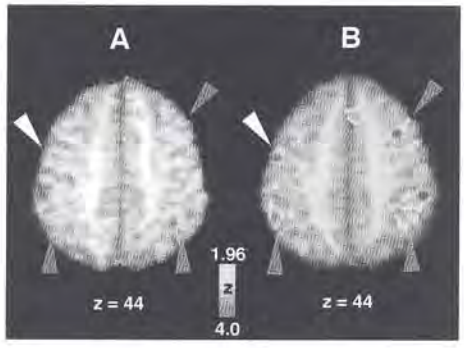



**Figure 17.11** Activations in bilateral medial frontal cortex (BA 9) and bilateral inferior parietal cortex (BA 40) for musicians (B) as compared to nonmusicians (A) during the discrimination of sequence durations.[49] Shown are group-averaged PET images for each task (contrasted with rest) overlaid on anatomical MRIs. (See Plate 15 in colour section.)

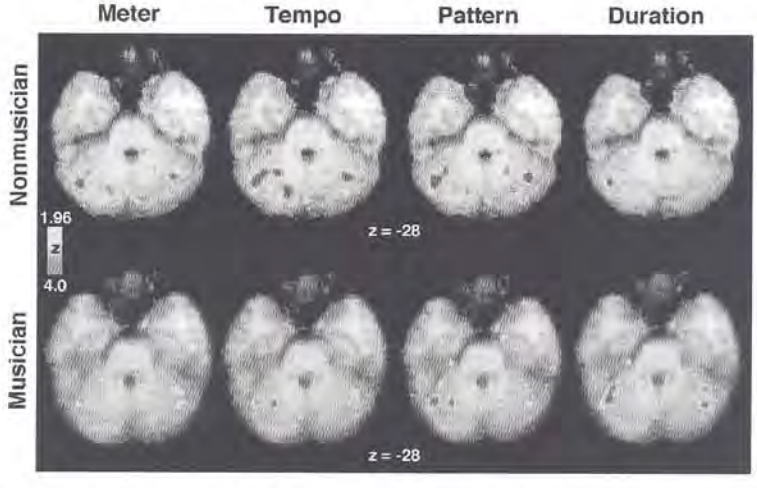



**Figure 17.12** Activations in posterior cerebellum as nonmusicians and musicians discriminate meter, tempo, pattern, and duration.[49] Shown are group-mean PET images for each task (contrasted with rest) overlaid on MRIs. (See Plate 16 in colour section.)

increase in cerebellar activation appears to be associated with greater novelty of rhythm discrimination, consistent with other neuroimaging findings[71] and new hypotheses of cerebellar function.[37]

The activations in prefrontal cortical areas BAs 9, 10, and 47 are likely supporting working memory processes.[48] The activation in prefrontal and inferior parietal areas follows similar indications in the recent neurological study of tone duration discrimination.[64] We speculate that activations in temporal, frontal, and midbrain may be involved in processing the temporal grouping structure of rhythms.

These findings suggest some key brain areas in the distributed neural circuits underlying the perceptual/cognitive representation of central aspects of musical rhythm. These data suggest that distinct brain areas subserve musical meter, tempo, pattern, and duration. In addition, the results confirm that the neural subsystems underlying pitch discrimination are distinct from those subserving the elements of rhythm discrimination. The differences in activated brain areas for nonmusicians and musicians likely reflect differences in strategy, perceptual skill, and cognitive representation of the components of musical rhythm and pitch.

## Neurology of pitch discrimination

In the foregoing review of neuroimaging and neurological studies of music, the cerebellum appears to be involved in some way in nonmotor, nonsomatic sensory or cognitive processing. This kind of conclusion is supported by many other recent neuroimaging and neurological findings.[37–45,73] Such findings challenge classical motor theories of cerebellar function and have created an environment in which researchers are increasingly questioning the extent to which the cerebellum is involved in motor control. Although in most cases, new alternative proposals are described as an extension of the role of the cerebellum in motor coordination itself, others, who have been driven primarily by new neuroimaging results, are proposing radically different functions for this structure.

James Bower and I, and our colleagues, have been using neuroimaging of tactile and cutaneous sensory and motor tasks to test the hypothesis that the cerebellum is involved in monitoring and optimizing the acquisition of information in various sensory modalities.[72–74] A direct prediction of this hypothesis is that damage to the cerebellum will impair performance on sensory tasks, especially those in which successful performance depends on sensory information processing at a very fine time scale. We evaluated this prediction for auditory perception by examining for the first time the effect of global cerebellar degeneration on auditory pitch discrimination in (nonmusician) patients.[75] This strictly auditory task does not depend on overt motor behaviour, as do other sensory tasks.

Fifteen patients and 15 healthy control subjects (matched in age and education) performed a pitch discrimination task measuring difference thresholds. The patients suffer from pancerebellar degeneration caused by hereditary, idiopathic, paraneoplastic, or postinfectious pancerebellitis. The degree of cerebellar degeneration in these patients, as confirmed by examinations of pancerebellar ataxia and anatomical imaging, ranged from minimal to severe, and included mild- and moderate-grade cases of ataxia. Patients and controls mostly had normal loudness thresholds; some had mild hearing loss in the 4000–8000 Hz range.

On each pitch discrimination trial, subjects decided whether a comparison tone was higher or lower than a preceding 500 Hz standard. All tones had suprathreshold loudness (80 dB) and 400 ms duration; the standard was followed 400 ms later by the comparison pitch. The tones were composed of the fundamental, three times the fundamental frequency at a 50 per cent amplitude, and five times the fundamental at a 25 per cent amplitude. The difference threshold estimation procedure used was PEST.[76,77] Subjects performed 60 trials, with threshold trials from above and below the standard pitch, mixed

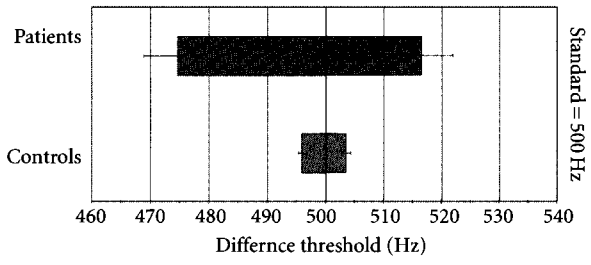



Figure 17.13 Average lower and upper difference discrimination thresholds for auditory pitch in pancerebellar degeneration patients and healthy controls.[75]

together in pseudorandom order. Each subject also performed two control tasks, the WAIS-III Digit Memory Span task, assessing auditory working memory capacity, and a simple auditory detection reaction time task, assessing simple motor reaction time for detecting audible sounds.

As shown in Figure 17.13, patients' pitch discrimination thresholds were on average 5.5 times that for controls ($p < 0.0001$). The degree of impairment was correlated ($r = +0.70$, $p < 0.003$) with the severity of patients' pancerebellar ataxia. Furthermore, the sizes of patients' pitch discrimination thresholds were uncorrelated with their loudness thresholds for the standard tone ($r = +0.02$, $p < 0.79$) and for the wider range from 125 to 8000 Hz ($r = +0.06$, $p < 0.22$). Patient and control groups did not differ on digit span memory or auditory detection reaction time.

As a group, then, patients with pancerebellar degeneration, but no other neurological conditions, show strongly impaired pitch discrimination. However, the degree of impairment for individuals is proportional to the severity of their pancerebellar ataxia. The severity of the pancerebellar ataxia (e.g. with symptoms in gait, speech, limb, and eye behaviour) is assumed to be indicative of the degree of pancerebellar atrophy. Although these data do not indicate the location of auditory processing regions in the cerebellum, it is assumed that atrophy in auditory regions of the cerebellum is proportional to the pancerebellar ataxia. Such atrophied auditory regions are assumed to be responsible for the pitch discrimination impairment. These cerebellar patients' auditory abilities are intact in a fundamental way, as indicated by reasonably normal auditory loudness thresholds. This observation is consistent with the idea that primary sensory processing is subserved in appropriate regions of cerebral cortex, which are spared in these patients. However, patients' loudness thresholds (ranging from normal hearing to mild hearing loss) are uncorrelated with the degree of their pitch discrimination impairment. Central to the idea of cerebellar involvement in sensory data acquisition is the suggestion that cerebellar effects would be most clearly evident in tasks that require highly optimized sensory data. In the current case, it seems reasonable to assume that the simple detection of the presence of a tone is unlikely to require the same quality sensory data as the fine discrimination of pitch.

Thus, these results suggest that the cerebellum has a fundamental supporting role in the information processing necessary for fine auditory discrimination. This conclusion is consistent with cerebellar activations reported in recent neuroimaging studies involving pitch

or melody processing.[46,49] The pattern of coupled pitch and ataxia measures, and decoupled pitch and loudness measures, suggests that the role of the cerebellum is to monitor and optimize the acquisition of information in sensory modalities (audition). Patients and controls did not differ in digit span memory. This indicates that the differences between patients and controls in pitch discrimination were very likely not due to any working memory deficits that may possibly be associated with cerebellar damage.

An important independent corroboration of the implications of these data is provided by a recent study examining cerebellar degeneration patients' performance on a variety of visual tasks.[43] The authors report that the patients were impaired selectively for the detection of visual motion, speed, and direction, as dissociated from eye movement control and from any other attentional, memory, or cognitive task components. It should be noted that there are also neurological studies suggesting that the cerebellum may be instrumental in other nonmotor behaviours. These behaviours include judging the timing of events, solving perceptual and spatial reasoning problems, and generating words according to a semantic rule.[47,78–83] In terms of the sensory acquisition hypothesis we are testing, however, impairments in these specific tasks may be more related to the dependence of these tasks on well-controlled information processing (e.g. sensory acquisition) than to different specific roles for the cerebellum in each of those behaviors.

## General issues in interpreting brain activations

Given the rapid emergence of promising new neuroimaging methods for the study of music performance, perception, and cognition, it may be useful in the present volume to discuss current and general methodological issues for interpreting the role of activated brain areas in the processing of music. Mapping elementary operations onto particular brain areas and mapping those operations/areas onto system-level models of specific musical and other cognitive tasks requires a convergence across paradigms (behavioural tasks) and independent experiments. No single task or paradigm used in neuroimaging or neurological investigations can isolate an elementary operation. Likewise, no single group of subjects (i.e. one study's sample) can fully capture the location and distribution of a brain area in a population. Thus, mapping localized operations entails sifting through multiple studies.

In functional neuroimaging, for example, the analysis of multiple studies, or metanalysis, can be rigorously and quantitatively performed[84] because the data are customarily reported in stereotaxic coordinates.[23,30] A growing number of rigorous metanalyses with important implications for mapping operations and systems, as well as for psychological models, are being conducted. For example, such analyses have been performed in motor processes and supplementary motor areas,[85] in oculomotor control and frontal eye fields,[86] in speech motor-sensory processes and the mouth area of primary motor-sensory cortex,[87,88] in visual information processing,[89–91] in visual attention,[91] and in memory.[92] Improved methods for such meta-analyses are currently being developed and validated.[84,87,88,93] Using such quantitative stereotactic-based metanalyses will yield important insights into the relationship between neural mechanisms supporting music and those supporting other cognitive processes.

As an aside, it should be noted that mapping neural mechanisms and systems will be advanced more efficiently if, whenever possible, the coordinate-based analyses just described replace analyses based only on lower resolution measures such as hemisphere, lobe, sulcus, gyrus, and, eventually, Brodmann area. Interestingly, close analysis suggests that much of the seemingly large variability in structure-function correlations remarked upon by some researchers using surface-based anatomical landmarks may vanish when the frame of analysis shifts to stereotactic coordinate-based analysis.[88]

Be that as it may, hypothesizing an operation (or computation) for an area requires a principled analysis of cognitive and behavioural processes and tasks. In spite of efforts in the cognitive sciences, there is no formal language in which to capture component operations that fractionate a task, including musical ones. A useful first step in that development would be for circumscribed communities of researchers studying a particular operation, area, or paradigm(s) in music processing to settle on an effective common vocabulary, with an eye towards establishing a working standard vocabulary with researchers studying other specific operations, areas, or paradigms. At the present time, this is rarely done, and apparently never in a concerted fashion. Standardized component models will be of great benefit for the process of mapping operations down to brain areas and up to system-level tasks in order to compare music and other cognitive processes.

Another significant point with regard to these goals is that whenever possible the operation hypothesized for a brain area should be consistent with its underlying microcircuitry. This is also rarely done in cognitive neuroscience. Take, for example, the case of hypothesizing operations for the cerebellum. The cerebellum has a uniform microcircuitry throughout, so until proved otherwise, we should assume that it performs a uniform operation, even if that operation may be performed on different kinds of information (as consistent with connectivity). Of course, even the most widely accepted facts about microcircuitry in an area can change from time to time. In the case of the cerebellum, David Marr's model of cerebellar function[94] was shown to be inconsistent with the underlying microcircuitry some time after its publication (although it continued to be influential).

Finally, there is a misguided tendency when interpreting neuroimaging data to assume that every activated area is directly involved in the cognitive or behavioural task eliciting it. A parallel tendency leads to the assumption that every cognitive or behavioural process affected by damage to a particular brain area necessarily implicates direct involvement of that area in those processes. As our sophistication in the computational and neural sciences grows, we will likely recognize that there are a range of indirect roles a mechanism might assume.

Consider again the cerebellum. The cerebellum is thought by various researchers to be directly involved in a variety of cognitive and behavioural processes. However, the consequences of injury to, or absence of, the cerebellum are often not fully appreciated. Although 70–80 per cent of the neurons in the brain are packed tightly in the cerebellum, the symptoms are mild and transient in cases when nearly all of it is removed. Likewise, persons born without a cerebellum show few impairments. These observations imply that the function of the cerebellum is subtly supportive and not necessary. On the hypothesis of cerebellar function that my colleagues and I are testing,[37,73] a brain with a healthy cerebellum is assumed to be optimized for operations necessary for fine sensory acquisition.

Furthermore, a brain without a cerebellum is viewed as subtly suboptimal, whereas a brain with a damaged cerebellum can range from mildly to very suboptimal, depending on degree and location of damage. In the more debilitating cases, the assumption is that a damaged cerebellum is causing other brain areas to operate on the basis of corrupted, noisy, poorly acquired and controlled sensory information. More generally, there may be a number of brain areas (such as basal ganglia or cingulate) with comparably indirect supporting roles in other cognitive or behavioural processes. Clearly, careful neurological studies (and, in some cases, transcranial magnetic stimulation studies) can assist in dissociating shades of involvement of a brain area in a particular psychological process.

## Conclusion

Musical experiences intricately and deeply engage the mind and the whole brain. Details of this process are being unveiled by new scientific advances well represented in this volume. Although many challenges lie ahead for fully understanding the biological foundations of music, the tools for meeting at least a fair share of those challenges appear to be in hand.

## Acknowledgements

I am grateful to the following collaborators for their stimulating, intelligent assistance with the studies discussed in this chapter: Peter Fox, Donald Hodges, James Bower, Jeremy Schmahmann, Stephen Grill, Michael Thaut, Bruno Repp, Michael Martinez, and John Tennison. I also thank our subjects, musicians, and patients for their gracious participation in support of basic research. The work described in this article was supported by the EJLB Foundation, the National Association for Music Merchants (now the International Foundation for Music Research), the Texaco Foundation, a State of Colorado Center of Excellence Award, and NIH (NINDS) Grant NS3710901A1.

## References

1. **Lerdahl, F.** and **R. Jackendoff** (1983) *A Generative Theory of Tonal Music.* Cambridge, MA: MIT Press.
2. **Dowling, W.** and **D. Harwood** (1986) *Music Cognition.* Orlando, FL: Academic Press.
3. **Krumhansl, C.** (1990) *Cognition Foundations of Musical Pitch.* Oxford: Oxford University Press.
4. **Repp, B. H.** (1992) Probing the cognitive representation of musical time: structural constraints on the perception of timing pertubations. *Cognition* 44, 241–81.
5. **Sergent, J.** (1993) Mapping the musician brain. *Hum. Brain Map.* 1, 20–38.
6. **Peretz, I., R. Kolinsky, M. Tramo, *et al.*** (1994) Functional dissociations following bilateral lesions of auditory cortex. *Brain* 117, 1283–301.
7. **Zatorre R. J., A. R. Halpern, D. W. Perry, *et al.*** (1996) Hearing in the mind's ear: a PET investigation of musical imagery and perception. *J. Cognit. Neurosci.* 8, 29–46.
8. **Deliege, I.** and **J. Sloboda** (1997) *Perception and Cognition of Music.* Wicyita, KS: Psychology Press.
9. **Palmer, C.** (1997) Music performance. *Annu. Rev. Psychol.* 48, 115–38.

10. **Platel, H., C. Price, J. C. Baron, *et al.*** (1997) The structural components of music perception. *Brain* 120, 299–43.
11. **Liegeois-Chauvel, C., I. Peretz, M. Babai, *et al.*** (1998) Contribution of different cortical areas in the temporal lobes to music processing. *Brain* 121, 1853–67.
12. **Pantev, C., R. Oostenveld, A. Engelien, *et al.*** (1998) Increased auditory cortical representation in musicians. *Nature* 392, 811–14.
13. **Blood, A. J., R. J. Zatorre, P. Bermudes, *et al.*** (1999) Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat. Neurosci.* 2, 382–7.
14. **Wallin, N. L., B. Merker,** and **S. Brown** (eds) (1999) *The Origins of Music.* Cambridge, MA: MIT Press.
15. **Sergent, J., E. Zuck, S. Terriah, *et al.*** (1992) Distributed neural network underlying musical sight-reading and keyboard performance. *Sci.* 257, 106–9.
16. **Besson, M.** and **F. Faita** (1995) An Event Related Potential (ERP) study of musical expectancy: comparison of musicians with non-musicians. *J. Exp. Psychol.: Human Percept. Perform.* 21, 1278–96.
17. **Patel, A. D.** and **I. Peretz** (1997) Is music autonomous from language? A neuropsychological appraisal. In I. Deliege and J. Sloboda (eds) *Perception and Cognition of Music.* Wichita, KS: Psychology Press.
18. **Besson, M., F. Faita, I. Peretz, *et al.*** (1998) Singing in the brain: independence of lyrics and tunes. *Psychol. Sci.* 9, 494–8.
19. **Patel, A. D., E. Gibson, J. Ratner, *et al.*** (1998) Processing syntactic relations in language and music: an event-related potential study. *J. Cognit. Neurosci.* 10, 717–33.
20. **Patel, A. D., I. Peretz, M. Tramo, *et al.*** (1998) Processing prosodic and musical patterns: a neuropsychological investigation. *Brain Lang.* 61, 123–44.
21. **Parsons, L. M., P. T. Fox, J. S. Sergent, *et al.*** Neural systems underlying musical piano performance from memory. Submitted.
22. **Fox, P. T., M. A. Mintun, M. E. Raichle, *et al.*** (1984) A noninvasive approach to quantitative functional brain mapping with $H_2{}^{15}O$ tissue activity. *J. Nucl. Med.* 30, 141–9.
23. **Talairach, J.** and **P. Tournoux** (1988) Co-planar Stereotaxic Atlas of the Human Brain. New York: Thieme Medical Publishers.
24. **Fox, P. T., M. A. Mintun, E. M. Reiman, *et al.*** (1988) Enhanced detection of focal brain responses using inter-subject averaging and change-distribution analysis of subtracted PET images. *J. Cereb. Blood Flow Metab.* 8, 642–53.
25. **Woods, R. P., S. R. Cherry,** and **J. R. Mazziotta** (1992) Rapid automatic algorithm for aligning and reslicing PET images. *J. Comp. Assist. Tomogr.* 17, 536–46.
26. **Woods, R. P., J. C. Mazziotta,** and **S. R. Cherry.** (1993) MRI-PET registration with automated algorithm. *J. Comput. Assisted Tomogr.* 17, 536–46.
27. **Lancaster, J. L., T. G. Glass, B. R. Lankipalli, *et al.*** (1995) A modality-independent approach to spatial normalization of tomographic images of the human brain. *Hum. Brain Map.* 3, 209–23.
28. **Fox, P. T.** and **M. A. Mintun** (1989) Noninvasive brain mapping by change-distribution analysis of averaged PET images of $H_2{}^{15}O$ tissue activity. *J. Nucl. Med.* 30, 141–9.
29. **Xiong J., J. H. Gao, J. L. Lancaster, *et al.*** (1996) Clustered pixels analysis for functional MRI activation studies of the human brain. *Hum. Brain Map.* 3, 287–301.
30. **Lancaster, J. L., M. G. Woldorff, L. M. Parsons, *et al.*** (2000) Automatic Talairach labels for functional brain mapping. *Hum. Brain Map.* 10, 120–31.

31. Zatorre, R. J. (1988) Pitch perception of complex tones and human temporal-lobe function. *J. Acoust. Soc. Am.* 84, 566–72.
32. Zatorre, R. J. and S. Samson (1991) Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain* 114, 2403–17.
33. Zatorre, R. J., A. C. Evans, E. Meyer, *et al.* (1992) Lateralization of phonetic and pitch discrimination in speech processing. *Sci.* 256, 846–9.
34. Zatorre, R. J., A. C. Evans, and E. Meyer (1994) Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 14, 1908–19.
35. Griffiths, T. D., C. Buchel, R. S. J. Frackowiak, *et al.* (1998) Analysis of temporal structure in sound by the human brain. *Nature Neurosci.* 1, 422–7.
36. Perry, D. W., R. J. Zatorre, M. Petrides, *et al.* (1999) Localization of cerebral activity during simple singing. *NeuroReport* 10, 3979–84.
37. Bower, J. M. (1997) Is the cerebellum sensory for motor's sake, or motor for sensory's sake: the view from the whiskers of a rat? In C. I. de Zeeuw, P. Strata, and J. Voogd (eds) *Progress in Brain Research.* New York: Elsevier, pp. 483–516.
38. Jueptner, M., S. Ottinger, S. J. Fellows, *et al.* (1997) The relevance of sensory input for the cerebellar control of movements. *Neuroimage* 5, 41–51.
39. Parsons, L. M. and P. T. Fox. (1997) Sensory and cognitive functions. In J. D. Schmahmann (ed.) *The Cerebellum and Cognition.* New York: Academic Press, pp. 255–72.
40. Schmahmann, J. D. (ed.) (1997) *The Cerebellum and Cognition.* New York: Academic Press.
41. Desmond, J. and J. Fiez (1998) Neuroimaging studies of the cerebellum: language, learning, and memory. *Trends Cognit. Sci.* 2, 355–62.
42. Penhune, V. B., R. J. Zatorre, and A. C. Evans (1998) Cerebellar contributions to motor timing: a PET study of auditory and visual rhythm reproduction. *J. Cognit. Neurosci.* 10, 752–65.
43. Thier, P., T. Haarmeier, S. Treue *et al.* (1999) Absence of a common functional denominator of visual disturbances in cerebellar disease. *Brain* 122, 2133–46.
44. Parsons, L. M., D. Denton, R. Shade *et al.* (2000) Neuroimaging evidence implicating cerebellum in support of sensory/cognitive processes associated with thirst. *Proc. Natl. Acad. Sci. USA* 97, 2332–6.
45. Tesche, C. D. and J. Karhu (2000) Anticipatory cerebellar responses during somatosensory omission in man. *Hum. Brain Map.* 9, 119–42.
46. Parsons, L. M., D. A. Hodges, and P. T. Fox The brain basis of the comprehension of musical harmony, rhythm, and melody. Submitted.
47. Ivry, R. and S. Keele (1989) Timing functions of the cerebellum. *J. Cognit. Neurosci.* 1, 136–52.
48. Sakai, K., O. Hikosaka, S. Miyauchi *et al.* (1999) Neural representation of a rhythm depends on its interval ratio. *J. Neurosci.* 19, 10074–81.
49. Parsons, L. M. and M. Thaut Functional neuroanatomy of musical rhythm, meter, tempo, and duration. Submitted.
50. Petersen, S. E., P. T. Fox, M. I. Posner *et al.* (1989) Positron tomographic studies of the processing of single words. *J. Cognit. Neurosci.* 1, 153–70.
51. Mazoyer, B. M., N. Tzourio, and V. Frak *et al.* (1993) The cortical representation of speech. *J. Cognit. Neurosci.* 5, 467–79.
52. Stromswold, K., D. Caplan, N. Alpert *et al.* (1996) Localization of syntactic comprehension by positron emission tomography. *Brain Lang.* 52, 452–73.

53. Price, C. (1998) The functional anatomy of word comprehension and production. *Trends Cognit. Sci.* 2, 281–8.

54. Kurata, K. (1993) Premotor cortex of monkeys—set-related and movement-related activity reflecting amplitude and direction of wrist movements. *J. Neurophysiol.* 73, 836–54.

55. Passingham, R. E. (1993) *The Frontal Lobes and Voluntary Action.* New York: Oxford University Press.

56. Kurata, K. (1994) Information processing for motor control in primate premotor cortex. *Behav. Brain Res.* 61, 135–42.

57. Jackson, S. R. and M. Husain (1996) Visuomotor functions of the lateral pre-motor cortex. *Curr. Opin. Neurobiol.* 6, 788–95.

58. Rizzolatti, G., L. Fadiga, M. Matellii, *et al.* (1996) Localization of grasp representations in human by PET: I. Observations versus execution. *Exp. Brain Res.* 111, 246–52.

59. Carey, D. P., D. I. Perrett, and M. W. Oram (1997) Recognizing, understanding, and reproducing action. In F. Boller and J. Grafman (eds.) *Handbook of Neuropsychology,* vol. 11. Amsterdam: Elsevier, pp. 111–29.

60. Thompson, W. F. (1994) Sensitivity to combinations of musical parameters: pitch with duration, and pitch pattern with durational pattern. *Percept. Psychophys.* 56, 363–74.

61. Peretz, I. and R. Kolinsky (1993) Boundaries of separability between melody and rhythm in music discrimination: a neuropsychological perspective. *Q. J. Exp. Psychol.* 46, 1283–301.

62. Peretz, I. (1990) Processing of local and global musical information by unilateral brain-damaged patients. *Brain* 113, 1185–205.

63. Polk, M. and A. Kertesz (1993) Music and language in degenerative disease of the brain. *Brain Cognit.* 22, 98–117.

64. Harrington, D. L., K. Y. Haaland, and R. T. Knight (1998) Cortical networks underlying mechanisms of time perception. *J. Neurosci.* 18, 1085–95.

65. Seashore, C. E. (1938) *The Psychology of Music.* New York: McGraw-Hill.

66. Gordon, E. E. (1965) *Musical Aptitude Profile.* Boston: Houghton-Mifflin.

67. Corbetta, M., F. M. Miezin, S. Dobmeyer, *et al.* (1991) Selective and divided attention during visual discriminations of shape, color, and speed: funtional anatomy by PET. *J. Neurosci.* 11, 2383–402.

68. Grasby, P. M., C. D. Frith, K. J. Friston, *et al.* (1993) Functional mapping of brain areas implicated in auditory/verbal memory function. *Brain* 116, 1–20.

69. Shallice, T., P. Fletcher, C. D. Frith, *et al.* (1994) Brain regions associated with acquisition and retrieval of verbal episodic memory. *Nature* 368, 633–5.

70. Doyon, J. (1997) Skill learning. In J. D. Schmahmann (ed.) *The Cerebellum and Cognition.* New York: Academic Press pp. 273–94.

71. Corbetta, M., G. L. Shulman, F. M. Miezin, *et al.* (1995) Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. *Science* 270, 802.

72. Liu, Y., Y. Pu, J. H. Gao, *et al.* (2000) The human red nucleus and lateral cerebellum in cooperative roles supporting sensory discrimination. *Hum. Brain Map.* 10, 147–59.

73. Parsons, L. M, J. M. Bower, J. H. Gao, *et al.* (1997) Lateral cerebellar hemispheres actively support sensory acquisition and discrimination rather than motor control. *Learn. Mem.* 4, 49–62.

74. Gao, J. H., L. M. Parsons, J. M. Bower, *et al.* (1996) Cerebellum implicated in sensory acquisition and discrimination rather than motor control. *Science* 272, 545–7.

75. **Parsons, L. M., J. D. Schmahmann, S. Grill, *et al.*** Neurological evidence implicating the cerebellum in fine auditory discriminations. Submitted.

76. **Taylor, M.** and **C. Creelman.** (1967) PEST: efficient estimates of probability functions. *J. Acoust. Soc. Am.* 41, 782–7.

77. **Pentland, A.** (1980) Maximum likelihood estimation: the best PEST. *Percept. Psychophys.* 28, 377–9.

78. **Bracke-Tolkmitt, R., A. Linden, G. M. Canavan, *et al.*** (1989) The cerebellum contributes to mental skills. *Behav. Neurosci.* 103, 442–6.

79. **Fiez, J. A., S. E. Petersen, M. K. Cheney, *et al.*** (1992) Impaired nonmotor learning and error-detection associated with cerebellar damage—a single case study. *Brain* 115, 155–78.

80. **Grafman, J., I. Litvan, S. Massaquoi, *et al.*** (1992) Cognitive planning deficits in patients with cerebellar atrophy. *Neurology* 42, 1493–6.

81. **Courchesne, E., J. Townsend, N. A. Akshoomoff, *et al.*** (1994) Impairment in shifting attention in autistic and cerebellar patients. *Behav. Neurosci.* 108, 848–65.

82. **Horak, F. B.** and **H. C. Diener** (1994) Cerebellar control of postural scaling and central set in stance. *J. Neurophysiol.* 72, 479–93.

83. **Daum, I., S. Graber,** and **H. Ackermann** (1998) Cognitive Neuroscience Abstracts Annual Meeting Program: Supplement, *J. Cognit. Neurosci.* 113, 15.

84. **Fox, P. T., L. M. Parsons,** and **J. L. Lancaster** (1998) Beyond the single study: function/location metanalysis in cognitive neuroimaging. *Curr. Opin. Neurobiol.* 8, 178–87.

85. **Picard, N.** and **P. L. Strick** (1996) Motor areas of the medial wall: a review of their location and functional activation. *Cereb. Cortex* 6, 342–53.

86. **Paus, T.** (1996) Location and function of the human frontal eye-field: a selective review. *Neuropsychologia* 34, 475–83.

87. **Fox, P. T., J. L. Lancaster, A. Huang, *et al.*** (1999) Functional volume modeling: scaling for group size in averaged images. *Hum. Brain Map.* 8, 143–50.

88. **Fox, P. T., A. Huang, L. M. Parsons, *et al.*** (2000) Location-probability profiles for the human primary-motor mouth representation. *Neuroimage* 13, 196–201.

89. **Shulman, G. L., M. Corbetta,** and **R. L. Buckner, *et al.*** (1997) Top-down modulation of early sensory cortex. *Cereb. Cortex* 7, 193–206.

90. **Shulman, G. L., M. Corbetta,** and **R. L. Buckner, *et al.*** (1997) Common blood flow changes across visual tasks: I. Increases in subcortical structures and cerebellum but not in nonvisual cortex. *J. Cognit. Neurosci.* 9, 623–46.

91. **Shulman, G. L., J. L. Fiez, M. Corbetta, *et al.*** (1997) Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *J. Cognit. Neurosci.* 9, 647–62.

92. **Buckner, R. L.** and **S. E. Petersen** (1996) What does neuroimaging tell us about the role of prefrontal cortex in memory retrieval? *Semin. Neurosci.* 8, 47–55.

93. **Fox, P. T., J. L. Lancaster, L. M. Parsons, *et al.*** (1997) Functional volume modeling. *Hum. Brain Map.* 5, 306–11.

94. **Marr, D.** (1969) A theory of cerebellar cortex. *J. Physiol. (Lond.)* 202, 437–70.

CHAPTER 18

# COMPARISON BETWEEN LANGUAGE AND MUSIC

MIREILLE BESSON AND DANIELE SCHÖN

## Abstract

Similarities and differences between language and music processing are examined from an evolutionary and a cognitive perspective. Language and music cannot be considered single entities; they need to be decomposed into different component operations or levels of processing. The central question concerns one of the most important claims of the generative grammar theory, that is, the specificity of language processing: do the computations performed to process language rely on specific linguistic processes or do they rely on general cognitive principles? Evidence from brain imaging results is reviewed, noting that this field is currently in need of metanalysis of the available results to precisely evaluate this claim. A series of experiments, mainly using the event-related brain potentials method, were conducted to compare different levels of processing in language and music. Overall, results favour language specificity when certain aspects of semantic processing in language are compared with certain aspects of melodic and harmonic processing in music. By contrast, results support the view that general cognitive principles are involved when aspects of syntactic processing in language are compared with aspects of harmonic processing in music. Moreover, analysis of the temporal structure led to similar effects in language and music. These tentative conclusions must be supported by other brain imaging results to shed further light on the spatiotemporal dynamics of the brain structure–function relationship.

**Keywords:** Language and music; Music and language; Event-related brain potentials; Brain imaging

## Introduction

> *Once, a long time ago [...] it so happened that people took to uncivilized ways, were ruled by lust and greed, behaved in angry and jealous ways with each other [...]. Seeing this plight, Indra and other gods approached god Brahma and requested him to give the people a toy, but one which could not only be seen, but heard, and this should turn out a diversion (so that people gave up their bad ways).*
>
> A. Rangacharya, *Introduction to Bharata's Natya-Sastra*[1]

One main function of language is communication. Linguistic communication encompasses widely diverse uses of language, from the running of everyday life and basic interactions between individuals ('*Give me the salt, please*'.) to the esthetics of words, their

combination in poetry ('*Come what come may/Hours and time run/Through the roughest day*', Shakespeare), and the telling of stories. Language is also necessary for expression of rational thought and organization of human societies. It may well have evolved from the need for social bonding between individuals belonging to the same group.[2] Language also permits projections into the past and into the future and is necessary for the transmission of knowledge.[3] While these characteristics, among others, make language specific to *Homo sapiens*, they also seem to contribute to the splendid isolation of the linguistic function among the other human cognitive activities. Largely because of the enormous impact in the cognitive sciences of the generative grammar theory, developed by Chomsky,[4] language is most often considered as relying on specific cognitive principles. Bickerton,[5] for instance, argues that the principles that govern language 'seem to be specifically adapted for language and have little in common with general principles of thought or other apparatuses that might be attributable to the human mind' (p. 158). The specificity of the computations involved in language processing is the main question that we would like to address in this chapter. To try to delineate which computations are specific to language and which rely on more general cognitive principles, if any, we chose to compare language with another well-organized cognitive function that, although very different in many respects, nevertheless presents interesting similarities with language: music.

We start by reviewing some of the evidence in favour of the similarities and differences between language and music, based, first, on an evolutionary perspective, and second, on a cognitive perspective. We then report a series of experiments directly aimed at comparing different aspects of language and music processing.

## Similarities and differences between language and music

### Evolutionary perspective

While discussions on the origin of language were banished from the *Société de Linguistique* of Paris in 1866, and although the question of the anthropological foundations of language and music certainly remains difficult, there is renewed interest in evolutionary linguistics and musicology, as can be seen, for instance, from the recent and provocative book edited by Wallin *et al.*,[6] *The Origins of Music*. The question of a common or separate origin of language and music, which was at the centre of hot debates between philosophers and scientists from the seventeenth to the nineteenth century,[7–10] is now being examined using new tools and technologies. In our opinion, one of the most promising avenues is functional brain imaging. By offering an excellent spatial and/or temporal resolution, brain imaging methods allow us to examine the brain regions that are activated by different aspects of information processing and how these processes unfold in time. Even if such methods may not help solve the problem of the common or separate origin of language and music, they provide invaluable information on the question of the similarities and differences between these two systems. Before reviewing some of the relevant brain imaging data, let us quickly consider how ideas on the origin of language and music have evolved throughout the centuries.

**'The old masters'** Rousseau[9] was a strong advocate of the view that music and language share some common ancestor and that language evolved out of music for the sake of a

rational organization of human societies. Darwin[7] also argued for a common origin, but considered that music evolved out of the primate's reproductive calls and that language was first. Interestingly, most of the authors who addressed this issue in the book edited by Wallin *et al.*[6] also seem to share the opinion of a common ancestor to language and music. The concept that the first basic function of both language and music was to express emotive meaning through variations in the intonation of the voice (intonational prosody) and rhythm also seems to be an object of consensus.[10–12] In both language and music, emotional excitement is expressed through fast, accelerating, and high-registered sound patterns. In this again they join Rousseau,[9] who considered that the first languages were sung, not spoken; they were aimed at expressing emotions, love, hate, and anger. They were passionate before being rational.

**Music and culture** In Western cultures, music has evolved to become more and more isolated from other expressive forms. Moreover, most studies in music cognition are concerned with music perception, and music performance has received much less attention (but see Ref. 13). By contrast, in other cultures, in which magical thought is still alive, the bounds among music, song, dance, poetry, and rite have not been lost.[14,15]

Furthermore, ethnomusicological studies have often emphasized music as a voluntary act: music is the acoustic result of action. Kubik,[16] for instance, pointed out that African music is not just sound; action is an intrinsic part of musical performance. Motor patterns are themselves sources of aesthetic pleasure, independent from the sound that they are associated with. This strong intertwining between music and action is even reflected in language, the same word being used in several African languages to refer to music and dance.

Blacking[14] defined music as 'sound that is organized into socially accepted patterns'. Moreover, he argues that every piece of music has its own inherent logic, as the creation of an individual reared in a particular cultural background. However, his claim that patterns of sound reflect patterns of social organization seems somewhat coarse and in need of further elaboration. Still, in much the same way that a context-sensitive grammar is a more powerful analytical tool than a context-free grammar, the cognitive systems underlying different styles of music shall be better understood if music is considered in context. Different musical styles should therefore be considered not as 'sonic objects' but as humanly organized sound whose patterns are related to the social and cognitive processes of a particular society and culture.

## Cognitive perspective

**Structural aspects** Many definitions have been (and are still to be) proposed for language and music. The amazing point, however, is that the definition given for music will often apply to language as well, and vice versa. This is striking when we consider the comparison between language and music from both a structural and a functional perspective. Arom,[17] for instance, proposed two structural criteria to define music. One is rhythm and the temporal ratios that delineate a piece of music by the formalized segmentation of time. The other is that all cultures have divided the sound continuum into discrete pitches that form the musical scales. These two criteria may apply to language as well. Language is also composed of sequential events that unfold in time with a specific rhythm and specific segmental

(phonemes) and suprasegmental (prosody) information. Moreover, the speech continuum is divided into discrete phonemes, the basic phonological unit. More generally, it is clear that both language and music are conveyed by sounds, are ubiquitous elements in all cultures, are specific to humans, and are cultural artifacts that do not correspond to natural objects.[12] They are rule-based systems composed of basic elements (phonemes, words, notes, and chords) that are combined into higher-order structures (musical phrases and sentences, themes and topics) through the rules of harmony and syntax. Therefore, there may be a musical grammar, and the experimental results to be described indeed point to the similarity of the brain's response to some specific violations of syntax in both language and music.

Functional aspects From a functional perspective, several similarities also exist between language and music. In this respect, it is interesting to examine the following citation from Pinker[18] in his book, *The Language Instinct*: 'Language is a complex, specialized skill, which develops in the child spontaneously, without effort or formal instruction, is deployed without awareness of its underlying logic, is qualitatively the same in every individual, and is distinct from more general abilities to process information or behave intelligently' (p. 18). Except for some constraints regarding music production, we could substitute 'music' for 'language' and the characteristics would apply as well. Both language and music rely on intentionality: All music implies 'an act of creation that actualizes an intention',[17] and this is true of language as well. In other words, both language and music require a theory of mind.[19] Both develop with specific learning according to more or less standardized procedures depending on the linguistic or musical culture. Even though perception precedes production in both domains, children acquire musical and linguistic rules in a similar, effortless way. Early on, children are able to create new musical and verbal sentences by applying a rule system that they have been able to abstract without conscious intentions. Both language and music involve memory; adults can recognize and reproduce learned melodies, words, poetry, and songs.

Similarities or differences? A question of grain of analysis? It should be noted, however, that the comparison between language and music might highlight either their similarities or their differences depending on the grain chosen for analysis. Thus, similarities at one level of processing may be interpreted as differences at another level. For example, while temporal structure and rhythmic organization play a fundamental role in both language and music, the metric structure is specific and consistent throughout a given musical piece, but the suprasegmental prosodic structure of language is less specific and more variable. Similarly, the segmentation of the sound continuum into discrete units (pitches or phonemes) is found in all music and languages. However, if we can eventually make an analogy between phonemes and intervals of a musical scale, we must also be aware of their differences. In fact, although the number of pitches by octave (degrees of the scale) is very similar across musical cultures (seven or fewer notes), the number of phonemes largely differs between languages (from 11 in Polynesian to 141 in the language of the Bushmen,[18] with 44 phonemes in English and 36 in French). Furthermore, some of the perceptual properties of the basic elements in music have no equivalent in language, as, for instance, the fact that octaves are perceived as equivalent in almost all cultures. This effect is linked with the finding that two notes separated by an octave are related by a simple frequency

ratio of 2 : 1. Generally, the relationships between different pitches in a musical piece are much simpler than the relationships between different phonemes in a linguistic sentence.

As a last example, all languages are organized according to a syntactic structure that may be universal.[20,21] Indeed, verbs and nouns are always present. However, the order in which those elements are presented varies among languages: Subject–Verb–Object (French, English, etc.); Subject–Object–Verb (German, Japanese, etc.); Verb–Object–Subject (Malgache, etc.).[22] Even if it is common to refer to a musical syntax or a musical grammar, the extent to which this analogy extends beyond a simple metaphor remains to be determined. Music perception shares universal laws of auditory perception. For instance, the perception of a musical phrase is automatically influenced by factors such as the grouping of discrete notes into sequences (i.e. the melodic contour) and the feeling of closure that accompanies the playing of a cadence at the end of a phrase. Some of these factors are universally shared and others, just as verbal language, are culturally shared. However, even if there is such a thing as a musical grammar, the rules seem more flexible and ambiguous than the syntactic rules used in language. Ambiguity is a key element of the grammar and aesthetics of music.[23] There are always several ways to perceive and enjoy a musical piece. Finally, musical elements are most often played simultaneously, and each element may have its own 'syntax'. This vertical dimension of musical structure, commonly referred to as harmony, is not present in language. While different words sung at the same time may melt in a sublime combination of rhythm, melody, and harmony (as in the polyphonic madrigals of Monteverdi), different words produced at the same time by different speakers will only create an unpleasant cacophony, like that in a political debate.

**Meaning and expectancy** Even if the similarities and differences between language and music depend on the level of details considered for the analysis, one fundamental difference nevertheless remains. Whereas the meaning of words is understood in relation to an extralinguistic designated space, music is considered mostly self-referential.[24–27] This does not mean that music is asymbolic. However, while the meaning of words is defined by an arbitrary convention relating sounds to meaning, notes or chords have no extramusical space in which they would acquire meaning. The internal sense of music may be conceived as something that goes beyond any objective reference structure and the possibilities of verbal language.[28] Much as Wittgenstein[29] who asked: 'Describe the coffee aroma!', music is the kingdom of the ineffable. As stated by Meyer in his wonderful book *Emotion and Meaning in Music*,[25] 'Music means itself. That is, one musical event (be it a tone, a phrase, or a whole section) has meaning because it points to and makes us expect another musical event' (p. 35). Interestingly, this statement not only highlights one of the most important differences between language and music, that is, the unsolved question of musical semantics, but also emphasizes their strongest similarity: in their own way, both systems generate strong expectancies. Just as a specific word is expected within a specific linguistic context, specific notes or chords are expected at a given moment within a musical phrase. Either these expectations are fulfilled, giving rise to resolution or satisfaction, or they are not fulfilled, giving rise to tension or surprise.

We should not believe, however, that expectation may 'bear the entire burden of deriving affect'.[30] Other factors such as tempo, volume, and nonmechanical interpretation of music certainly influence musical emotions. Still, the structure of music has intrinsic

points of instability that tend to resolve, and the tension/resolution phenomenon results in affects. Moreover, tensions are perceived at different levels depending on the analysis performed. Jackendoff,[30] in analogy to language, points out that a modular and informational encapsulated parser might be at work independently from conscious memory. This independence from memory may explain why we keep enjoying a piece on repeated hearings 'in spite of the consequences of an intuitive theory of affect based on expectation'. In fact, an autonomous parser will keep analysing and recreating whatever structure is retrieved from memory. Then 'surprise will still occur within the parser'.[30]

If rather than asking ourselves what is the meaning of music, we make a more fruitful reflection on 'what can I do with sounds'? we may discover that music is, first of all, a set of choices. The flow of these choices might possibly become visible as a musical thought. Behind all this, the image of children playing appears. When the child plays with small wood blocks, we could say that the game is a way of answering the question: what can I do with my small wood blocks? Then, from the pleasure of playing we get directly into the aesthetic pleasure.

Concluding this brief and necessarily incomplete excursus, it is important to keep in mind that musical meaning is the sum of analytic approaches (musical parser), individual and/or cultural associations to the external/internal world (during some periods in the last centuries 'music was conceived as conveying precise emotional and conceptual meanings, established by codes, or at least, *repertoires*'[31]), and aesthetic reaction. The importance of the aesthetic component of music becomes evident in considering that 'the form of a work of art gains its aesthetics validity precisely in proportion to the number of different perspectives from which it can be viewed and understood'.[32]

Levels of processing From a cognitive perspective, language and music cannot be considered as single entities. To be analysed and compared they need to be reduced to their constitutive elements. Within music, one classically differentiates the temporal (metre and rhythm), melodic (contour, pitch, and interval), and harmonic (chords) aspects. Each aspect most likely involves different types of processing, so that the processes called into play to process rhythm may differ from those involved in the processing of pitch and melodic intervals. Similarly, within language, at least four different levels of processing have been taken into consideration. The phonetic-phonological level, which comprises both segmental (phonemes) and suprasegmental (prosody) levels; the morphosyntactic level, which encompasses the combination of phonemes into morphemes and of morphemes into words; the syntactic level, which governs the relations between words; and the lexicosemantic level, with access to the meaning of words and sentences. Finally, while often ignored in psycholinguistic and neurolinguistic experiments, the pragmatic level that comprises discourse organization and contextual influences represents an essential aspect of language organization.

Insofar as we agree with the concept that language and music cannot be considered as wholes but need to be subdivided into their component operations, it becomes unrealistic, for instance, to view the linguistic function as localized in the left hemisphere and music in the right. Rather, some aspects of language processing may preferentially involve left cerebral structures, whereas others require structures on the right. The same remark applies to

music as well. With this view in mind, the task of the cognitive neuroscientist is to delineate the different computations performed within one level of processing, to understand the mechanisms that underlie these computations, and to localize where in the brain these mechanisms are implemented. This task is fraught with both philosophical and methodological problems,[19] but science is advancing rapidly, and new methods are now available to track these issues. In the second part of this presentation, we summarize the results of our research on the comparison between different levels of processing in language and music. Before going into the details of the experiments, we briefly review the position of linguistic theories on the question of the specificity of language processing.

## Specificity of the computations involved in language processing?

### The generative grammar theory

One of the most important claims of the generative grammar (GG) theory is that language is autonomous from other cognitive functions.[4,18,21,33,34] Language is considered a computational module that entails its own functional and neural architecture.[12] Moreover, the linguistic module comprises different submodules, each responsible for different aspects of language processing, phonology, morphology, syntax, semantics, and pragmatics. Each submodule is encapsulated,[35] so that the processing of information in a module is performed independently of the processing of information in another submodule. Phonological processing, for instance, is realized without being influenced by the morphological, syntactic, semantic, or pragmatic aspects of language processing. Thus, the computations required to process language are specific to language, and the computations in one module are performed independently of those in the other modules.

Another basic claim of the GG theory is that languages are defined by their deep syntactic structure: syntax plays a dominant role in the structural organization of language. Moreover, from a functional perspective, syntax is first.[36] Logico-mathematic computations are first performed on symbols that have no intrinsic meaning; they only acquire meaning in a second step. Therefore, the chain of computations necessary to process language is considered to be serially and hierarchically organized.

### Other linguistic theories

It should be noted, however, that other linguistic theories have been developed in the last 20–30 years that advocate very different views of language structural and functional organization.[22,37] Although it is beyond the scope of this presentation to go into the details of these different linguistic theories, which differ from each other in many respects (functional grammar,[38,39] cognitive grammar,[40–42] and linguistic functional typology[43–45]), the important point is that these theories call into question the two basic claims of the GG theory just summarized. First, they reject the idea that language is an autonomous function, relying on its own structural and functional architecture. By contrast, they consider that languages rely on general cognitive principles, linked with perceptual and sensorimotor motor experiences.[46,47] Second, they reject the syntacticocentrism of the GG theory and the idea of the autonomy of syntax relative to phonology, morphology, semantics, and pragmatics.[22]

Following Langacker,[41] for instance, semantics, morphology, and syntax form a continuum with specific meaning associated with lexicosemantic units and schematic meaning associated with grammatical units. Thus, in contrast to the GG view that grammatical units are semantically empty, all grammatical elements have meaning. Moreover, linguistic units are not static but constructed through a dynamic process influenced by the context of enunciation[48] and the interactions between individuals in a situation of communication.[49] Therefore, language should be studied not in isolation but in relation to other cognitive functions, specifically, attention and short-term and episodic memory.[45]

### Success of generative grammar theory in cognitive neurosciences

Several reasons may explain the success of the GG theory in both linguistic and cognitive sciences, but two are of particular interest. First, the cognitive stakes of the GG theory have been clearly explained. It has therefore been possible to make predictions and design experiments to test these predictions.[21,47] Second, the modular organization of the functional aspects of language processing is clearly neurocompatible. The concept that language is organized in submodules, each responsible for one specific processing stage, finds strong support in the localizationist views of cerebral organization. The recent development of brain imaging methods, together with older data from the neuropsychological literature, largely contributes to the idea that specific functions are implemented in specific brain structures. A brief review of the literature shows that while this concept of the mapping of basic sensory functions into the organization of primary, sensory brain areas is probably correct, the story certainly becomes more complicated when trying to localize such higher-order cognitive abilities as language or music.

### Evidence from brain imaging

Brain imaging methods are aimed at understanding the functional activity of the brain either directly through measures of the electrical activity of single neurons (intracellular recordings), or of neuronal populations (electroencephalography, EEG), or through the magnetic activity that is coupled with the electrical activity (magnetoencephalography, MEG), or indirectly through the measures of brain metabolic activity (positon emission tomography, PET, and functional magnetic resonance imaging, fMRI). Overall, direct methods have excellent temporal resolution and relatively poor spatial resolution, whereas the reverse is true for indirect methods. Elegant works have been conducted using these different methods to demonstrate, for instance, the retinotopic organization of the visual cortex using fMRI[50] and the tonotopic organization of the auditory cortex using intracellular recordings (see Chapter 9 this volume), MEG,[51] or fMRI.[52] Hence, there is strict mapping between the organization of the receptor fields at the periphery, in either the retina or the cochlea, and the functional organization of the primary visual and auditory cortex. Aside from extending to humans previous discoveries in animals, these findings validate the use of such complex methods as fMRI to study human perception and cognition.

To address the specificity of the brain structures involved in language processing, we would need metanalysis of the results obtained across the many experiments aimed at localizing the different aspects of language processing. We would then need to do the same

for music or for any other cognitive function of interest and then compare the results of these metanalyses. Although such metanalyses are being performed for some aspects of language processing, such as language production[53] or prelexical and lexical processes in language comprehension,[54] data in the neuroimaging of music are still too scarce for such an enterprise. Moreover, assuming that such metanalyses are performed for music as well, it still remains extremely difficult to compare results of experiments that were not directly designed to compare language and music processing. Indeed, leaving aside the theoretical problem of which level of processing in language is best compared with which level of processing in music, the choice of the task to be performed on the stimuli, its difficulty, as well as experimental factors, such as the mode (blocked vs mixed) and rate of stimulus presentation, stimulus repetition, and data analysis (e.g. subtraction method, correlative analyses), have been shown to exert a predominant influence on the results obtained. With these remarks in mind, it is nevertheless interesting to mention some of the results found for language and music to determine the extent to which the brain structures that are activated are similar or different.

Few experiments have been designed to directly compare language and music using brain imaging methods. Binder *et al.*[55] compared tones and word processing in an fMRI study. Results showed that several brain structures, including the left superior temporal sulcus, middle temporal gyrus, angular gyrus, and lateral frontal lobe, showed stronger activation for words than tones. However, both types of stimuli activated the Heschl gyrus and the superior temporal plane, including the planum temporale. The investigators concluded that whereas the planum temporale is similarly involved in the auditory processing of words and tones, other broadly distributed areas are specifically involved in word processing. Gandour *et al.*[56] conducted a PET study in which both Thai and English participants were required to discriminate pitch patterns and Thai lexical tones derived from accurately filtered Thai words. Results of the tone minus pitch subtraction indicated that only native Thai speakers showed activation of the left frontal operculum (BA 44/45). This finding was taken as evidence that Thai lexical tones are meaningful for native Thai speakers but not for English speakers. However, for our purposes, it is also interesting that for both Thai and English speakers, several structures, including the left anterior cingulated gyrus (BA 32), the left and right superior temporal gyrus (BA 22), and the right cerebellum, were activated in both pitch and tone tasks.

More generally, results have shown that primary auditory regions (BA 41 and BA 42) respond in similar ways to speech and music.[57] Secondary auditory regions (BA 22) are activated by hearing and understanding words[58] as well as by listening to scales,[59] auditory imagery for sounds,[60] and access to melodic representations.[61] The supramarginal gyrus (BA 40) seems involved in understanding the symbolism of language[58] and the reading of musical scores.[59] Broca's area is known to be involved in motor activity related to language and was also shown to be active when playing music[59] and when musicians were engaged in a rhythmic task.[61] The supplementary motor areas (BA 6) and the right cerebellum are also active when playing and imaging playing music.[59,62] Although this list is far from exhaustive, it nevertheless suffices to show that some of the most important language areas are clearly involved in music processing as well. Some brain structures also seem to be specifically or preferentially involved in language processing,[63] and the converse is true for

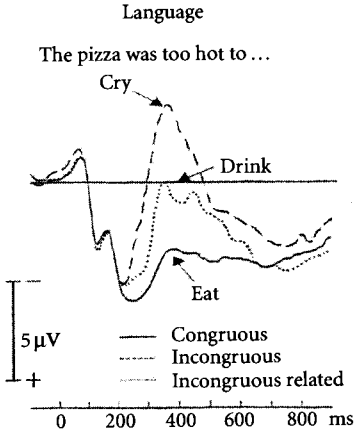



Figure 18.1 ERPs elicited by final sentence words at the central recording site (Cz) for congruous and incongruous words and for incongruous words that are semantically related to the best sentence completion. The amplitude of the negative component, peaking at 400 ms postfinal word onset (N400) is largest for incongruous words, intermediate for incongruous words related to the best sentence completion, and smallest for congruous words. In this and subsequent figures, amplitude (μV) is represented on the ordinate, with negative voltage up, and time (ms) on the abscissa. (Adapted from Ref. 66.)

music.[64] Results of metanalysis of the type just mentioned should clearly help in drawing maps of the structure–function relationships known for language and music.

## Semantics, melody, and harmony

A starting point in the study of the neurophysiological basis of language processing was the discovery of the N400 component by Kutas and Hillyard.[65] This negative component of the event-related brain potentials (ERPs), peaking around 400 ms after word onset, is elicited by words that are semantically unexpected, incongruous, and within a linguistic context (e.g. 'The pizza was too hot to cry'; Figure 18.1). Further results have shown that N400 amplitude is modulated by semantic priming, so that an unexpected word related to the best sentence completion (e.g. 'drink' when the expected word is 'eat'; Figure 18.1) elicits a smaller N400 than a completely unexpected word (e.g. 'cry'[66]). These results, together with those issued from a large number of experiments, have led to the consensus that the N400 is a good index of the integration process of a word within its linguistic context.

The first experiments that we designed were aimed at finding out whether an N400 component would also be elicited when melodically and harmonically unexpected notes were presented within a melodic context.[67–69] We presented both familiar and unfamiliar monodic musical phrases to musicians and nonmusicians. The familiar melodies were chosen from the classical repertoire of Western occidental music from the eighteenth and nineteenth centuries, and the unfamiliar musical phrases were composed by a musician following the rules of tonal harmony (Figure 18.2). These melodies were ended by the congruous or most expected note, by a note out of the tonality of the musical phrase (non diatonic incongruities perceived as wrong notes), or by a note within the tonality but not

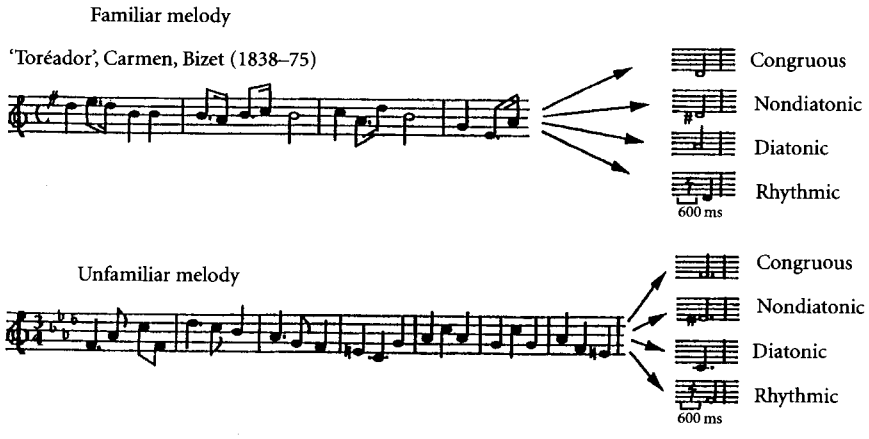



Figure 18.2 Examples of the stimuli used in the experiment. (Adapted from Ref. 69.)

the most expected ending (melodic or diatonic incongruities). Thus, we created a degree of musical incongruity from diatonic to nondiatonic.

Results clearly showed that both types of unexpected notes elicited the occurrence of late positive components, peaking around 600 ms (P600). As demonstrated for the N400 component, P600 amplitude was shown to depend on the degree of musical incongruity: it was larger for the most unexpected, nondiatonic wrong notes than for the less unexpected, diatonic incongruities. Moreover, the amplitude of the P600 was larger for familiar than for unfamiliar musical phrases and for musicians than for nonmusicians (Figure 18.3). These findings clearly demonstrate not only that specific notes are expected within a musical phrase, but also that such expectations depend on the familiarity of the musical excerpts and the expertise of the listener. Thus, an interesting similarity between language and music, just mentioned, their ability to generate strong expectancies, is supported by empirical evidence. However, our results also show that the processes that govern semantic expectancy and that are reflected by a negative component, peaking around 400 ms, the N400, are qualitatively different from those involved in musical expectancies and reflected by a positive component peaking around P600 ms, the P600. While, to our knowledge, the functional significance of positive vs negative polarities in the ERPs is not clearly established, our results, by demonstrating qualitative differences between language and music processing, nevertheless strongly argue for the specificity of the processes involved in computing the semantic aspects of language. Thus, one of the most important differences between language and music, outlined in the introduction, the fact that, in contrast to language, music has no intrinsic meaning and is a self-referential system, seems to find some support in these experimental findings.

### Semantics and harmony in opera

Opera is perhaps the most complete art form, as it calls upon music, language, drama, and choreography. It originated in Italy at the end of the sixteenth century with *Dafne*, set to

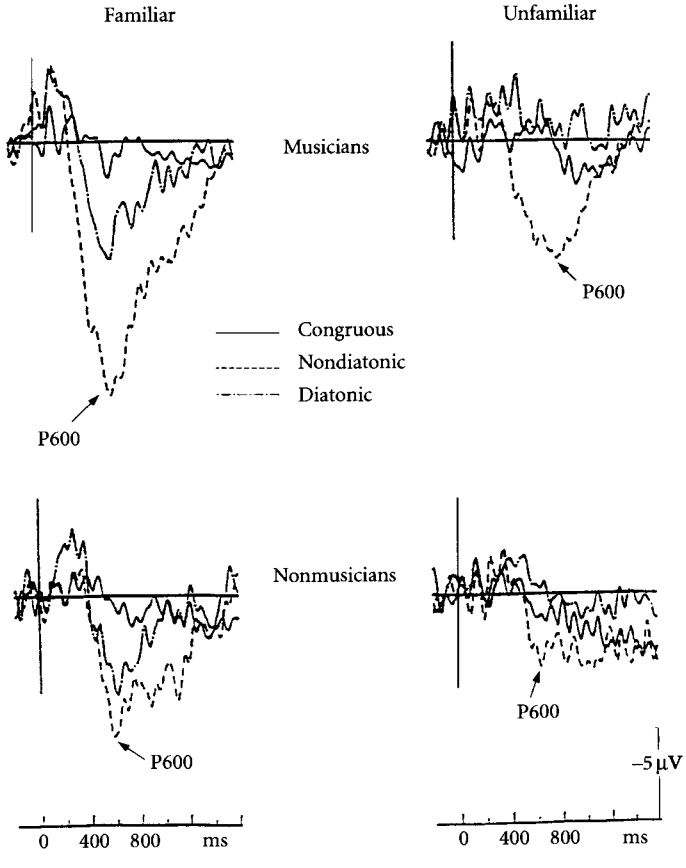



**Figure 18.3** ERP results for musicians and nonmusicians are presented separately for familiar and unfamiliar musical phrases. The vertical lines mark the onset of the final note. Results are from one typical recording site, the parietal location (Pz). The amplitude of the positive component, P600, is larger for nondiatonic than for diatonic incongruity, for musicians than for nonmusicians and for familiar than for unfamiliar musical phrases. (Adapted from Ref. 69.)

music in 1594 by the Florentine composers Corsi and Peri. The first opera to survive intact is probably *Euridice*, set to music by Peri and Caccini and performed in 1600 as a wedding gift to Maria de Medici and Henri IV. Opera, as a new art form, then spread to other Italian courts with the better-known *Orfeo* of Monteverdi in 1604. Since that time, a question that has interested both music analysts and composers has been to determine which of the words or the music plays the most important role in opera. In his *Life of Rossini*, Stendhal (1783–1842) argued that music is most important: 'its function is to animate the words'. Later, ethnomusicologists, such as Levman,[70] pointed out that the lyrics are subordinate to the music in tribal songs and rituals. By contrast, Richard Wagner (1813–83) considered that both aspects are intrinsically linked: 'Words give rise to the music and music develops and reinforces the language', an opinion shared by Pierre Boulez:[71] 'The text is the centre

and the absence of the musical piece'. Richard Strauss (1864–1949) even composed an opera, *Capriccio* (1940), to illustrate the complementarity of words and music.

To determine, based on scientific grounds, if the words or the music is most important when we listen to opera, we selected 200 excerpts from French operas of the nineteenth and twentieth centuries.[72] Each excerpt lasted between 8 and 20 s and was sung *a capella* by a female singer under each of four experimental conditions, that is, the final word of the excerpt was (1) semantically congruous and sung in tune, (2) semantically incongruous and sung in tune, (3) semantically congruous and sung out of tune, and (4) both semantically incongruous and sung out of tune (Figure 18.4).

Based on previous results,[65] it was of interest to determine whether semantically incongruous words will also elicit an N400 component when they are sung. Similarly, it was of interest to determine whether congruous words sung out of tune will also elicit a P600 component.[69] Of most interest was the double incongruity condition: will semantically incongruous words sung out of key elicit both an N400 and a P600 component? If language plays the most important role when we listen to opera, then results may show an N400 but not a P600. Conversely, if music is the cornerstone of opera, then results may show a P600 without an N400. Maybe both effects will be elicited, however; they may then be additive (i.e. equal to the sum of the effect associated with each type of incongruity alone) or interactive. To answer these questions, we recorded the ERPs associated with the final words of each excerpt, from 16 professional musicians from the opera company in Marseille.

To summarize, results demonstrated that sung incongruous words did elicit an N400 component, thus extending to songs results previously reported for written and spoken language[65,73] (Figure 18.5A). Moreover, words sung out of tune did elicit a P600 component, thus extending to songs results previously reported for out-of-tune notes[69,74–76] (Figure 18.5B). Most interesting are the results in the double incongruity condition. They show that incongruous words sung out of tune elicit both an N400 and a P600 component (Figure 18.5C). Interestingly, the N400 occurred earlier than the P600, which is taken as evidence that the words were processed faster than the music. Finally, effects in the double incongruity condition were not significantly different from the sum of the effects observed in each condition of simple incongruity (see Figure 18.6). This finding provides a strong argument in favour of the independence (i.e. the additivity) of the computations involved in processing the semantic aspects of language and the harmonic aspects of music. Therefore, when we listen to opera, we process both the lyrics and the tunes in an independent fashion, and language seems to be processed before music.

**Influence of attention** We tracked these results further by conducting another series of experiments aimed at studying the effect of attention, again testing some professional musicians from the opera company in Marseille.[76] We hypothesized that if lyrics and tunes are processed independently, listeners should be able to focus their attention only on the lyrics or only on the tunes, depending on the instructions. Without going into the details of the results, an N400 component was elicited to sung incongruous words, and a P600 was associated with congruous words sung out of tune, thus replicating our previous findings.[72] Most interestingly, the N400 to incongruous words completely vanished when participants focused their attention on the music (Figure 18.7). Thus, musicians were able not to process the meaning of words; they did not notice whether the terminal word made sense

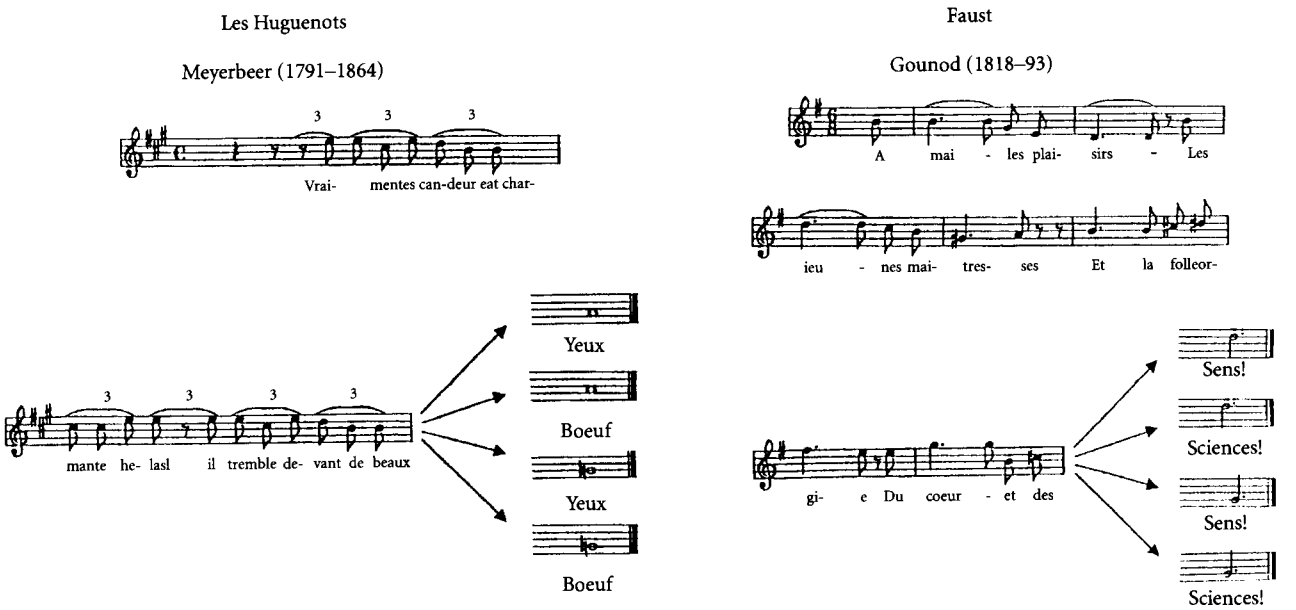



**Figure 18.4** Example of the opera's excerpts used in the experiment. Approximate translation of the excerpts, from *Les Huguenots* (Meyerbeer): 'Really, his naïvity is charming. However, he trembles in front of beautiful eyes', and from *Faust* (Gounod): 'For me, the pleasures and young mistresses, the crazy orgy of the heart and the senses'. Note that in French, the final incongruous words 'boeufs' and 'sciences' rhyme with the expected completions 'yeux' and 'sens'. The final note of the excerpt is in or out of tune. (Adapted from Ref. 72.)

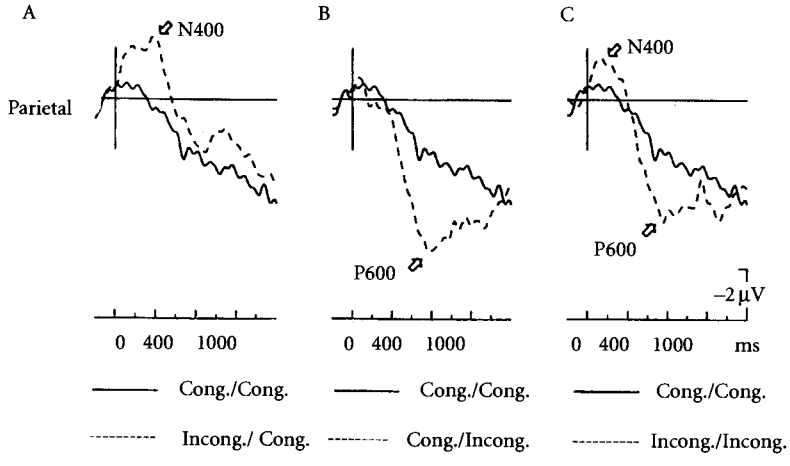



**Figure 18.5** ERPs results averaged across 16 professional musicians and recorded from the parietal electrode (Pz). Terminal congruous words sung in key (Cong./Cong.) are compared to (A) semantically incongruous words sung in tune (Incong./Cong.), (B) semantically congruous words sung out of tune (Cong./Incong.), and (C) semantically incongruous words sung out of tune (Incong./Incong.). The *vertical lines* mark the onset of the final word of the excerpts. A large N400 component develops in the 50–600 ms that follow the presentation of semantically incongruous words (A). In marked contrast, a P600 develops in the 400–1200 ms that follow the presentation of words sung out of tune (B). Most importantly, both an N400 and a P600 develop in response to the double incongruity (C). (Adapted from Ref. 72.)

within the linguistic context when they only listened to the music. Conversely, P600 amplitude was significantly reduced when musicians focused attention on language, so that they did not hear that the final word was sung out of tune. Taken together these results again provide strong arguments in favour of the independence of lyrics and tunes. There is some limit to such processing independence, however. Results in the double incongruity condition showed that the presence of one type of incongruity influenced the processing of the other type. When words were both semantically incongruous and sung out of tune, musicians could not help but hear the musical incongruity, even if they were asked to focus their attention on language.

### Syntax and harmony

The rules of harmony and counterpoint are often described as the grammar of tonal music. As syntax is used to extract the fundamental structure of an utterance by assigning different functions to different words, the rules of harmony allow us to specify the different elements, notes and chords, that fulfill a specific harmonic function. Results of experiments manipulating the harmonic function of target chords have shown that violations of harmonic expectancies are associated with P600 components.[76,77] Interestingly, research on syntax using ERPs has also shown that different types of syntactic violations, such as violations of gender, word order or noun-verb agreement, elicit a positive component, peaking around 600 ms.[78–80] Moreover, both components show a similar parietal

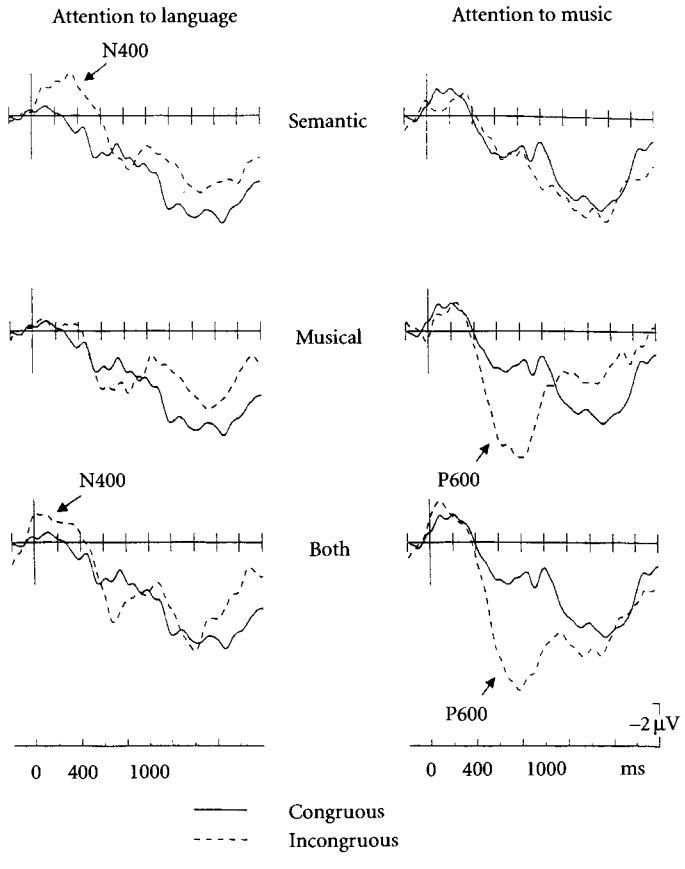



Figure 18.6 Overlapped are the ERPs to congruent and incongruent endings, recorded at the central recording site (Cz), when participants paid attention only to the language (*left column*) or only to the music (*right column*) of the opera's excerpts. A large N400 effect is generated when participants focus their attention on language. This effect completely vanishes when attention is focused on music (*top row*). Similarly, the P600 effect is much greater when participants paid attention to music than when they paid attention to language (*medium row*). Finally, when words are both semantically incongruous and sung out of tune, the N400 effect is greater when participants paid attention to the language, and the P600 effect is greater when they paid attention to the music (*bottom row*). (From Regnault and Besson, in preparation.)

distribution over the scalp, which, together with their similar polarity and latency, seems to indicate that they reflect qualitatively similar processes.

To further test this hypothesis, Patel and collaborators[81] conducted an experiment directly aimed at comparing the P600 components elicited by harmonic and syntactic violations. ERPs associated with a word within a grammatically simple, complex, or incorrect sentence were compared to those associated with the presentation of a chord that belonged to the tonality induced by the chord sequence, a nearby, or a distant tonality. Results showed that aside from early morphologic differences in the ERPs to words and chords, due to the differences in the acoustic characteristics of these two types of auditory signals, the effects associated with the

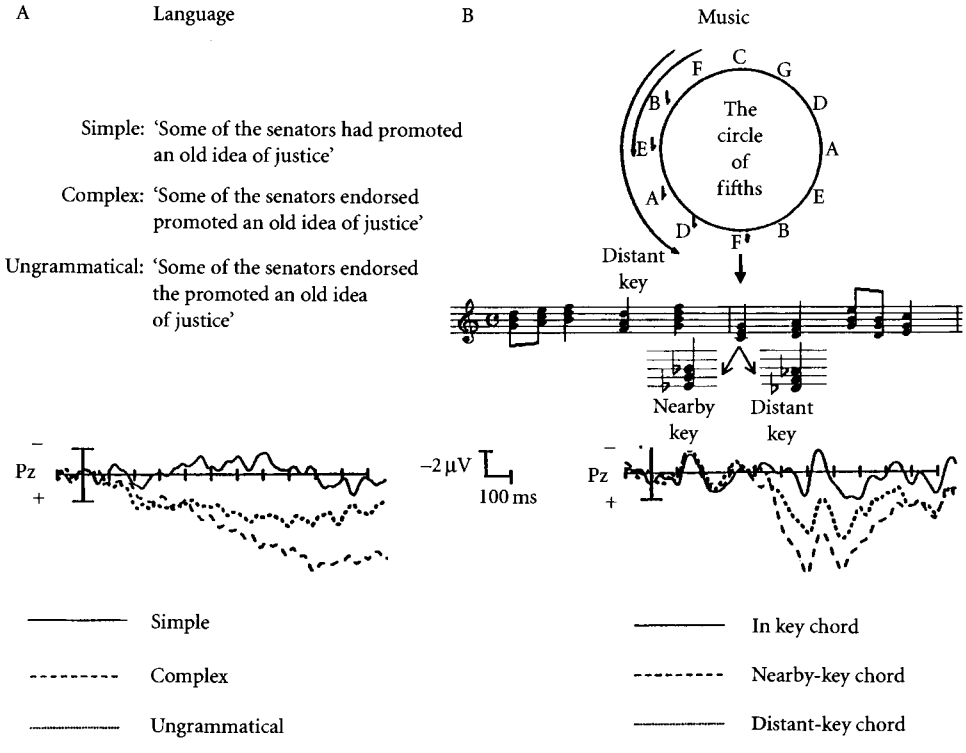



Figure 18.7 (A) Examples of the sentences presented in the auditory language experiment. Results showed increased positivity from simple to ungrammatical sentences. (B) Representation of the circle of fifths. Examples of the stimuli used in the experiment. The target chord, shown by the downward-pointing *vertical arrow*, is the congruous chord. The two *arrows* below the musical notation point to moderately incongruous (nearby key) and highly incongruous (distant key) target chords. Results also showed increased positivity from the in-key chords to the distant-key chords. (Adapted from Patel Ref. 81.)

violation of syntactic and harmonic expectancies were not significantly different (Figure 18.8). Therefore, these results raise the interesting possibility that a general cognitive process is called into play when participants are asked to process the structural aspects of an organized sequence of sounds, be it language or music. Finally, an early right anterior negativity was found at around 300–400 ms in response to a chord belonging to a distant tonality. These results paralleled those obtained in language experiments showing that an early left anterior negativity is also associated with some syntactic violations.[82] Whereas these two negative components showed a different distribution over the scalp, with a left predominance for language and a right predominance for music, they may reflect functionally similar processes.

### Temporal structure

Spoken language, as music, is composed of acoustic events that unfold in time. Because of the temporal structure inherent in both language and music, specific events are expected at specific times. The main question addressed in the next series of experiments was to

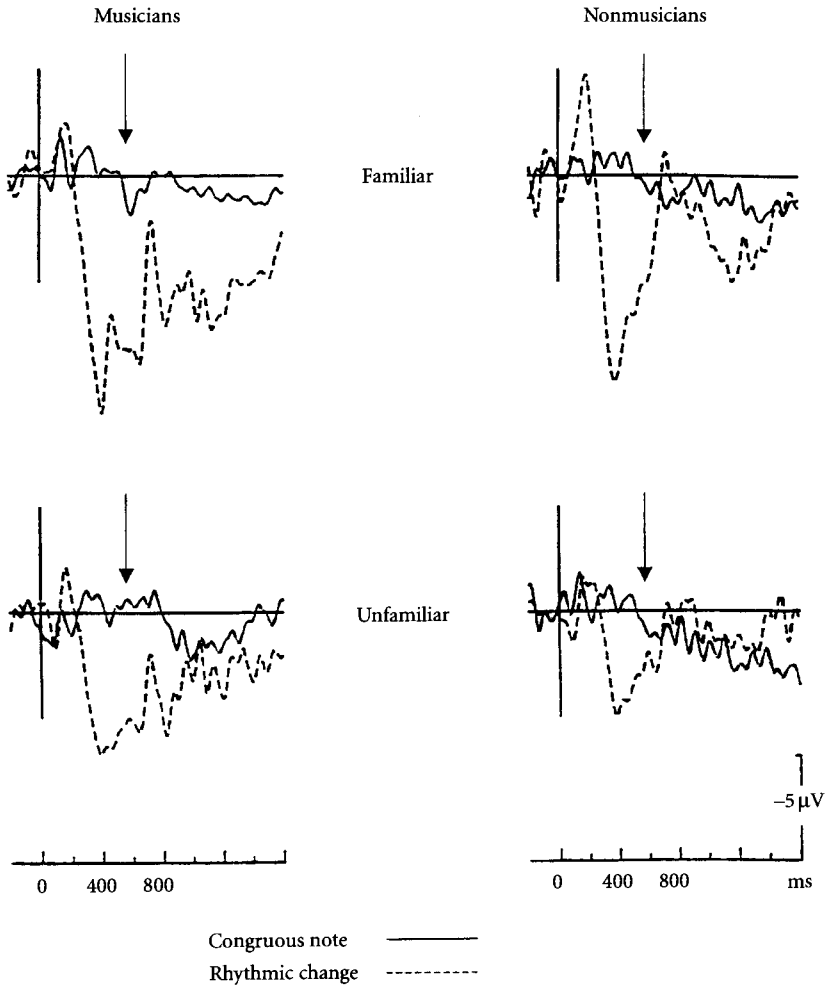



**Figure 18.8** Overlapped are the ERPs to congruous notes and to the rhythmic incongruities ending familiar and unfamiliar musical phrases for musicians and nonmusicians. Recordings are from the parietal electrode (Pz). Large emitted potentials are elicited when the final note should have been presented (*vertical bar*) but was delayed by 600 ms. The *arrow* points to the moment in time when the final note was presented. (Adapted from Ref. 69.)

determine whether the processes involved in analysing temporal structures rely on general cognitive mechanisms or differ as a function of the specific characteristics of the materials to be processed. We used both the ERP and the MEG methods to analyse the time course of the effects of temporal structure violations in language and music, and fMRI to localize the cerebral structures activated by these violations. We hypothesized that if a general mechanism is responsible for processing the temporal structures in language and music, qualitatively similar effects should be revealed in the ERP and MEG recordings, and similar brain areas should be shown to be activated by temporal violations. By contrast, if

processing temporal information in both systems relies on different mechanisms, qualitatively different effects and different brain areas should be found in language and music.

In previous experiments,[69] we introduced an unexpected silence between the penultimate note and the last note of a musical phrase (Figure 18.2). Results showed that a large biphasic, negative then positive, potential, the emitted potential,[83] was elicited when the final note should have been presented but was not, because it was delayed by 600 ms. The amplitude of this effect was similar in musicians and nonmusicians, but it was larger for familiar than unfamiliar melodies (Figure 18.9). These findings clearly indicate that both musicians and nonmusicians could anticipate the precise moment when the final note was to be presented and were surprised when it was not. Moreover, known melodies allowed participants to generate more precise expectancies than did unfamiliar melodies. Therefore, these results indicate that the occurrence of an emitted potential is a good index of temporal expectancy.

It was then of interest to determine whether similar results would be found for spoken language.[84] To this aim, we presented both familiar (e.g. proverbs) and unfamiliar auditory sentences to participants. In half of the sentences, final words occurred at their normal position, while in the other half, they were delayed by 600 ms. Results showed that an emitted potential, similar to the one described for temporal ruptures in music, developed when the final word should have been presented (Figure 18.9). Therefore, these ERP results indicate that qualitatively similar processes seem to be responsible for temporal processing in language and music.

To strengthen this interpretation, it was important to determine whether the same brain structures are activated by the processing of temporal ruptures in language and music. As already mentioned, fMRI allows localization of brain activation with excellent spatial resolution. Moreover, the MEG permits localization of the generators of the effects observed on the scalp more precisely than the ERP method, while offering an excellent temporal resolution. Therefore, in collaboration with Heinze and his research team, we conducted three experiments in which we presented both auditory sentences and musical phrases.[85] These experiments used a blocked design in which only sentences or musical phrases without temporal ruptures were presented within a block of trials, and only sentences or musical phrases with temporal ruptures at unpredictable positions were presented within another block of trials. The ERP method was used in the first experiment to replicate, within subjects, the results found previously with two different groups of subjects,[69,84] and the fMRI and the MEG methods were used, respectively, in the other two experiments, trying to localize the effects of interest.

Overall, the ERP results replicated, within subjects, those previously found in music and language separately (i.e. an emitted potential). However, comparison of the conditions with and without temporal violations revealed a different pattern of activation using the MEG and fMRI methods. Source localization based on MEG data revealed that the underlying generators of the biphasic potential recorded on the scalp were most likely located in the primary auditory cortex of both hemispheres. By contrast, fMRI results showed activation of the associative auditory cortex in both hemispheres as well as some parietal activation. Several factors may account for these differences,[85] but the main point is that similar brain areas were activated by temporal violations in both language and music. Therefore,

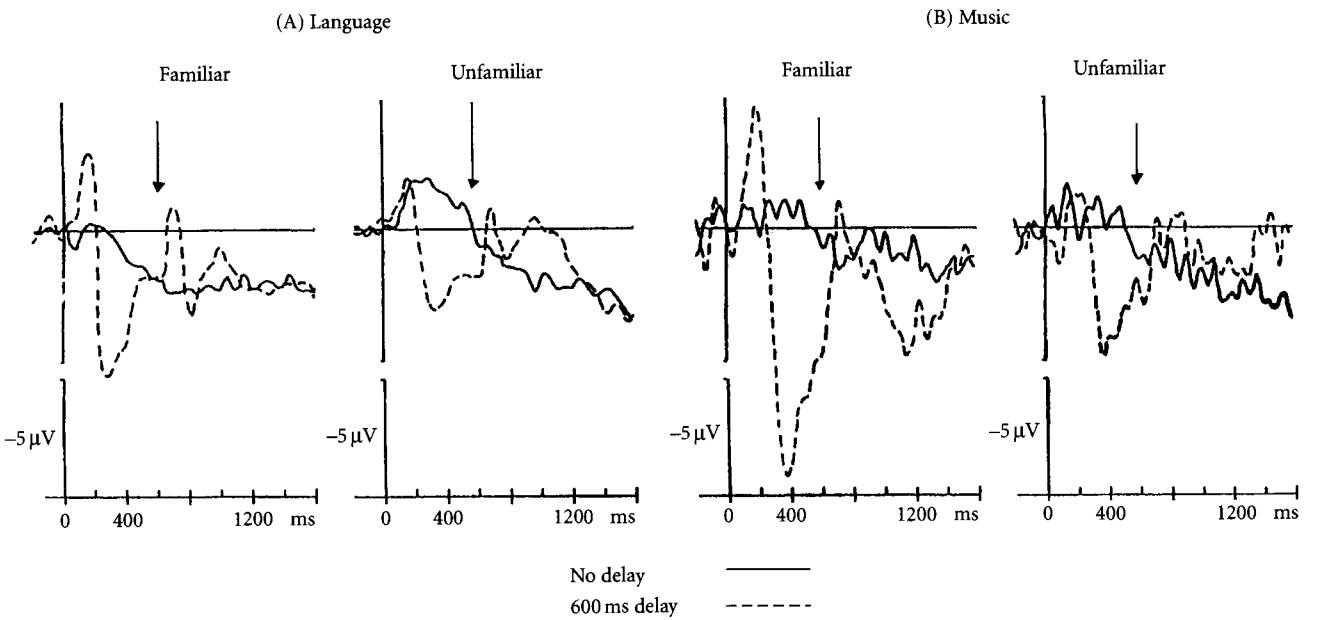



**Figure 18.9** Comparison of the effects of temporal violations in (A) language and (B) music. Recordings are from the parietal electrode (Pz). (A) Overlapped are the ERPs to congruous words and to the temporal disruptions ending familiar and unfamiliar sentences. In both language and music, large emitted potentials are elicited when the final event should have been presented (*vertical bar*) but was delayed by 600 ms. Note that the amplitude of the emitted potential is greater in music than in language, but that, in both cases, its amplitude is greater for familiar than for unfamiliar materials. (Adapted from Ref. 84 (A) and Ref. 69. (B).)

taken together our results suggest that processing temporal information in both language and music relies on general cognitive mechanisms.

## Conclusion

We have addressed one of the central questions of human cognition, the specificity of language processing. Is language an autonomous system, independent from other human cognitive abilities or does language rely on general cognitive principles? To address this question, we have conducted several experiments aimed at comparing some aspects of language processing with some aspects of music processing. We mainly used the ERPs method, which offers excellent temporal resolution and therefore permits study of the time course of information processing and determination of whether the processes involved in language and music are qualitatively similar or different.

Taken together, results have shown that the semantic computations required to access the meaning of words, and their integration within a linguistic context, seem to be specific to language. Indeed, whereas unexpected words within a sentence context are associated with the occurrence of an N400 component, unexpected notes or chords within musical phrases elicit a P600 component. By contrast, words that are unexpected on the basis of the syntactic structure of the sentence, and chords that are unexpected as a function of the harmonic structure of the musical sequence, elicit similar effects in both cases, namely, a P600 component. Early negative effects, that is, left and right anterior negativity, which developed between 200 and 300 ms, have also been reported in experiments manipulating syntax and harmony, respectively. Although their different scalp distributions seem to reflect involvement of different brain structures, more research is needed to further track their functional significance. Finally, violations of temporal structure within language and music also elicit similar effects, a biphasic negative–positive complex, the emitted potential. The occurrence of the emitted potential shows that in both language and music, words and notes or chords are expected at specific moments in time. Therefore, when we listen to language and music, not only do we expect words or chords with specific meaning and function, but we also expect them to be presented on time!

The question of the specificity of language processing has broad implications in our understanding of the human cognitive architecture and, even more generally, for the fundamental problem of the relationship between structures (different brain regions) and functions (e.g. language, music). Although the research reported here sheds some light on certain aspects of language processing and highlights some similarities and differences with music processing, more research is clearly needed in this fascinating research domain. Of utmost interest is the use of brain imaging methods that offer complementary information about the spatiotemporal dynamics of brain activity in order to pinpoint the networks of cerebral structures that are involved in two of the most human cognitive abilities: language and music.

## References


1. **Rangacharya, A.** (1966) *Introduction to Bharata's Natya-Sastra.* Bombay, India: Popular Prakashan.
2. **Nadel, S.** (1930) The Origin of Music. *Musical Q* 16, 531–46.
3. **Leroi-Gourhan, A.** (1988) *Le geste et la parole. I: La mémoire et les rythmes. Sci. d'aujourd'hui.* Paris: Albin-Michel.
4. **Chomsky, N.** (1957) *Syntactic Structures.* The Hague, the Netherlands: Mouton and Co.
5. **Bickerton, D.** (2000) Can biomusicology learn from language evolution studies? In N. L. Wallin, B. Merker and S. Brown (eds) *The Origins of Music.* Cambridge, MA: MIT Press, pp. 153–64.
6. **Wallin, N. L., B. Merker, and S. Brown** (eds) (2000) *The Origins of Music.* Cambridge, MA: MIT Press.
7. **Darwin, C.** (1871) *The Descent of Man, and Selection in Relation to Sex.* London: Murray.
8. **Descartes, R.** (1618/1987) *Abrégé de musique. Compendium musicae. Épiméthée.* Paris: P.U.F.
9. **Rousseau, J. J.** (1781/1993) *Essai sur l'origine des langues.* Paris, France: Flammarion.
10. **Spencer, H.** (1857) The origin and function of music. *Fraser's Mag.* 56, 396–408.
11. **Brown, S.** (2000) The 'musilanguage' model of music evolution. In N. L. Wallin, B. Merker, and S. Brown (eds) *The Origins of Music.* Cambridge, MA: MIT Press, pp. 271–300.
12. **Molino, J.** (2000) Toward an evolutionary theory of music and language. In N. L. Wallin, B. Merker, and S. Brown (eds) *The Origins of Music.* Cambridge, MA: MIT Press, pp. 165–76.
13. **Sloboda, J. A.** (2000) Individual differences in music performance. *Trends Cognit. Sci.* 4, 397–403.
14. **Blacking, J.** (1973) *How Musical is Man?* Seattle: University of Washington Press.
15. **Schön, A. and D. Schön** (1999) Il potere del suono e della musica. *Fuga a più voci. Psiche* 1–2, 159–65.
16. **Kubik, G.** (1969) Composition techniques in Kiganda xylophone music. *Afr. Music J.* 4, 22–72.
17. **Arom, S.** (2000) Prolegomena to a biomusicology. In N. L. Wallin, B. Merker, and S. Brown (eds) *The Origins of Music.* Cambridge, MA: MIT Press, pp. 27–30.
18. **Pinker, S.** (1994) *The Language Instinct: How the Mind Creates Language.* New York: Harper Perennial.
19. **Pacherie, E.** (1999) Philosophie et sciences cognitives. In J. F. Mattéi (ed.) *Encyclopédie Philosophique Universelle.* Paris: PUF.
20. **Chomsky, N.** (1988) *Language and Problems of Knowledge: The Managua Lectures.* Cambridge, MA: MIT Press.
21. **Chomsky, N.** (1991) Linguistics and cognitive science: problems and mysteries. In A. Kasher (ed.) *The Chomskyan Turn.* Cambridge, MA: Basil Blackwell, pp. 26–53.
22. **François, J.** (1998) Grammaire fonctionnelle et dynamique des langues—de nouveaux modèles d'inspiration cognitive et biologique. *Verbum* XX, 233–56.
23. **Aiello, R.** (1994) Music and language: parallels and contrasts. In R. Aiello and J. Sloboda (eds) *Musical Percep.* New York, NY: Oxford University Press, pp. 40–63.
24. **Jakobson, R.** (1973) Essais de linguistique générale. II. Rapports internes et externes du langage. Editions de Minuit, Arguments. Paris.
25. **Meyer, L.** (1956) *Emotion and Meaning in Music.* Chicago, IL: University of Chicago Press.
26. **Kivy, P.** (1991) *Music Alone. Philosophical Reflection on the Purely Musical Experience.* New York: Cornell Paperbacks.

27. **Boucourechliev, A.** (1993) *Le langage musical. Collections les chemins de la musique.* Paris: Fayard.
28. **Piana, G.** (1991) *Filosofia della Musica.* Milan: Guerini e associati.
29. **Wittgenstein, L.** (1953) *Philosophical Investigations.* In G. E. M. Anscombe and R. Rhees (eds). Oxford: Blackwell, pp. 209.
30. **Jackendoff, R.** (1991) Musical parsing and musical affect. *Music Percept.* 9, 199–230.
31. **Eco, U.** (1979) *Trattato di Semiotica Generale.* Milano: Bompiani.
32. **Eco, U.** (1962) *Opera Aperta.* Milan: Bompiani.
33. **Jackendoff, R.** (1997) *The Architecture of the Lang. Faculty.* Cambridge, MA: MIT Press.
34. **Pollock, J. Y.** (1997) *Langge et cognition. Introduction au programme minimaliste de la grammaire générative.* Paris: PUF.
35. **Fodor, J.** (1983) *Modularity of Mind.* Cambridge, MA: MIT Press.
36. **Frazier, L.** (1987) Sentence processing: a tutorial review. In M. Coltheart (ed.) *Attention and Performance. XII.* Hillsdale, NJ: Erlbaum, pp. 559–86.
37. **Victorri, B.** (1999) Le sens grammatical. *Languages* 136, 85–105.
38. **Dik, S.** (1997) *The Theory of Functional Grammar.* Berlin: Mouton De Gruyter.
39. **Van Valin, R. D. and R. J. LaPolla** (1997) *Syntax Structure, Meaning and Function.* Cambridge: Cambridge University Press.
40. **Lakoff, G.** (1987) *Women, Fire and Dangerous Things.* Chicago, IL: University of Chicago Press.
41. **Langacker, R. W.** (1987) *Foundations of Cognitive Grammar. Vol. I: Theoretical Prerequisites.* Stanford, CA: Stanford University Press.
42. **Talmy, L.** (1988) The relation of grammar to cognition. In B. Rudzka-Ostyn (ed.) *Topics in Cognitive Linguistics.* Amsterdam: Benjamins.
43. **Croft, W.** (1995) Autonomy and functionalist linguistics. *Lang.* 71, 490–532.
44. **Greenberg, J.** (1995) The diachronic typological approach to language. In M. Shibatani and T. Bynon (eds) *Approaches to Language Typology.* Oxford: Clarenton, pp. 145–66.
45. **Givón, T.** (1995) *Functionalism and Grammar.* Amsterdam: Benjamins.
46. **Fuchs, C.** (1997) Diversité des représentations linguistiques: Quels enjeux pour la cognition? In C. Fuchs and S. Robert (eds) *Diversité des langues et représentations cognitives.* Paris: Ophrys, pp. 5–24.
47. **Robert, S.** (1997) Variation des représentations linguistiques: des unités à l'énoncé. In C. Fuchs and S. Robert (eds) *Diversité des langues et représentations cognitives.* Paris: Ophrys, pp. 25–39.
48. **Culioli, A.** (1999) *Pour une linguistique de l'énonciation.* Paris: Ophrys.
49. **Fauconnier, G.** (1997) *Mappings in Thought and Language.* Cambridge, MA: Cambridge University press.
50. **Tootell, R. B. H., N. K. Hadjkhani, J. D. Mandola, *et al.*** (1998) From retinotopy to recognition: fMRI in human visual cortex. *Trends Cognit. Sci.* 2, 174–83.
51. **Pantev, C., M. Hoke, B. Luetkenhoener, *et al.*** (1989) Tonotopic organization of the auditory cortex: pitch versus frequency representation. *Sci.* 246, 486–8.
52. **Strainer, J. C., J. L. Ulmer, F. Z. Yetkin, *et al.*** (1997) Functional magnetic resonance imaging of the primary auditory cortex: analysis of pure tone activation and tone discrimination. *Am. J. Neuroradiol.* 18, 601–10.
53. **Indefrey, P. and W. J. M Levelt** (2000) The neural correlates of language production. In M. S. Gazzaniga (ed.) *The New Cognitive Neurosci.* Cambridge, MA: MIT Press, pp. 845–65.

54. **Norris, D.** and **R. Wise** (2000) The study of prelexical and lexical processes in comprehension: psycholinguistics and functional neuroimaging. In M. S. Gazzaniga (ed.) *The New Cognit. Neurosci.* Cambridge, MA: MIT Press, pp. 867–80.

55. **Binder, J. R, J. A. Frost, T. A. Hammeke, *et al.*** (1996) Function of the left planum temporale in auditory and linguistic processing. *Brain* 119, 1239–47.

56. **Gandour, J., D. Wong** and **G. Hutchins** (1998). Pitch processing in the human brain is influenced by language experience. *Neuroreport* 9, 2215–119.

57. **Zatorre, R. A., C. Evans, E. Meyer, *et al.*** (1992) Lateralization of phonetic pitch discrimination in speech processing. *Sci.* 256, 846–9.

58. **Falk, D.** (2000) Hominid brain evolution and the origin of music. In N. L. Wallin, B. Merker, and S. Brown (eds) *The Origins of Music.* Cambridge, MA: MIT Press, pp. 197–216.

59. **Sergent, J., E. Zuck, S. Terriah, *et al.*** (1992) Distributed neural network underlying musical sight-reading and keyboard performance. *Sci.* 257, 106–9.

60. **Zatorre, R. J., A. Halpern, D. W. Perry, *et al.*** (1996) Hearing in the mind's ear: a PET investigation of musical imagery and perception. *J. Cognit. Neurosci.* 8, 29–46.

61. **Platel, H., C. Price, J. C. Wise, *et al.*** (1997) The structural components of music perception. *Brain* 120, 229–43.

62. **Chen, W., T. Kato, X. H. Zhu, *et al.*** (1996) Functionnal mapping of human brain during music imagery processing. *NeuroImage* 3, S205.

63. **Grabowski, T. J.** and **A. R. Damasio** (2000) Investigating language with functional neuroimaging. In A. W. Toga and J. C. Mazziotta (eds) *Brain Mapping: The Systems.* London: Academic Press, pp. 425–58.

64. **Zatorre, J.** and **J. R. Binder** (2000) Functional and structural imaging of the human auditory system. In A. W. Toga and J. C. Mazziotta (eds) *Brain Mapping: The Systems.* London: Academic Press, pp. 365–402.

65. **Kutas, M.** and **S. A. Hillyard** (1980) Reading sens.ess sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–05.

66. **Kutas, M.** and **S. A. Hillyard** (1984) Event-related brain potentials (ERPs) elicited by novel stimuli during sentence processing. *Ann. N.Y. Acad. Sci.* 425, 236–41.

67. **Besson, M.** and **F. Macar** (1987) An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiol.* 24, 14–25.

68. **Besson, M., F. Faïta,** and **J. Requin** (1994) Brain waves associated with musical incongruity differ for musicians and non-musicians. *Neurosci. Lett.* 168, 101–5.

69. **Besson, M.** and **F. Faïta** (1995) An event-related potential (ERP) study of musical expectancy: comparison of musicians with non-musicians. *J. Exp. Psychol. Human Percept. Perform.* 21, 1278–96.

70. **Levman, B. G.** (1992) The genesis of music and language. *Ethnomusicology* 36, 147–70.

71. **Boulez, P.** (1966) Relevés d'apprenti. Editions du Seuil. Paris.

72. **Besson, M., F. Faita, I. Peretz, *et al.*** (1998) Singing in the brain: independence of lyrics and tunes. *Psychol. Sci.* 9, 494–8.

73. **MacCallum, W. C., S. F. Farmer,** and **P. V. Pocock** (1984) The effects of physical and semantic incongruities on auditory event-related potentials. *Electroencephal. Clin. Neurophysiol.* 59, 477–88.

74. **Paller, K. A., G. McCarthy,** and **C. C. Wood** (1992) Event-related potentials elicited by deviant endings to melodies. *Psychophysiology* 29, 202–6.

75. **Verleger, R.** (1990) P3-evoking wrong notes: unexpected, awaited or arousing? *Int. J. Neurosci.* 55, 171–9.

76. **Regnault, P., E. Bigand,** and **M. Besson** (2001) Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: evidence from auditory event related brain potentials. *J. Cognit. Neurosci.* 13, 1–17.

77. **Janata, P.** (1995) ERP measures assay the degree of expectancy violation of harmonic contexts in music. *J. Cognit. Neurosci.* 7, 153–64.

78. **Hagoort, P., C. Brown,** and **J. Groothusen** (1993) The syntactic positive shift as an ERP-measure of syntactic processing. *Lang. Cognit. Processes* 8, 439–83.

79. **Osterhout, L.** and **P. J. Holcomb** (1992) Event-related brain potentials elicited by syntactic anomaly. *J. Mem. Lang.* 31, 785–804.

80. **Friederici, A. D.** (1998) The neurobiology of language comprehension. In A. D. Friederici (ed.) *Lang. Comprehension: A Biological Approach.* New York: Springer, pp. 263–301.

81. **Patel, A. D., E. Gibson,** and **J. Ratner** (1998) Processing syntactic relations in language and music: an event-related potential study. *J. Cognit. Neurosci.* 10, 717–33.

82. **Friederici, A. D., E. Pfeifer,** and **A. Hahne** (1993) Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations: *Cognit. Brain Res.* 1, 182–92.

83. **Soutton, S., M. Braren, J. Zubin,** and **E. R. John** (1967) Evoked potential correlates of stimulus uncertainty. *Sci.* 150, 1187–8.

84. **Besson, M., F. Faïta, C. Czternasty,** and **M. Kutas** (1997) What's in a pause: event-related potential analysis of temporal disruptions in written and spoken sentences. *Biol. Psychol.* 46, 3–23.

85. **Weyert, H., M. Besson, C. Templemann,** *et al.* (2001) An analysis of temporal structure in language and music using ERPs, MEG and fMRI techniques. In preparation.

CHAPTER 19

# MUSICAL SOUND PROCESSING: EEG AND MEG EVIDENCE

MARI TERVANIEMI

## Abstract

Even during the performance of a simultaneous task unrelated to the sounds, the human auditory cortex is able to model precisely the invariances of the acoustic environment. Data acquired in mismatch negativity (MMN) paradigm have shown that temporally and spectrally complex sounds as well as their relations are automatically represented in the human auditory cortex. Furthermore, the MMN data indicate that these neural sound representations are spatially distinct between phonetic and musical sounds within and between the cerebral hemispheres. Majority of the MMN studies were conducted in pitch dimension but also temporal aspects of sound processing are under increasing experimentation. Up to some extent, also musical expertise is reflected in sound representation accuracy as indexed by the MMN.

## Introduction

Sound perception and cognition can be conceptualized as a process, in which the first stages are the least and the latest stages the most dependent on the attentional efforts of the listener.[1–3] For a long time, due to micro- and macro-level complexity of the acoustic signal, music sound encoding was assumed to require attentional effort of the listener. In the present review, some of these assumptions will be challenged.

In the majority of the studies to be reviewed below, the mismatch negativity (MMN) paradigm was utilized.[2] The MMN is evoked by an infrequently presented sound ('deviant') differing from the frequently occurring stimuli ('standard') in one or several parameters. Its presence implies that the invariant parameters of the standard sound were neurally encoded, thus producing a neural mismatch with those of the deviant sound. In other words, the MMN is an index for the neural traces of short-term auditory memory.

The MMN can be recorded even when the subject is performing a task unrelated to the stimulation under interest such as reading a book or playing a computer game. Thus it offers a direct measure for the similarity of neural sound representations, without being contaminated by differences, for instance, in attentional or motivational involvement of the subject during an experimental session or between the subject groups.

Yet, several data indicated that the MMN parameters closely correlate to the indicators of the subjects' perceptual accuracy as determined in a separate behavioural session. For instance, the MMN amplitude and latency reflect perceptual accuracy, as determined by

tests of musical aptitude (for a review, see Ref. 4) and by hit rates and reaction times.[5–8] In addition, the degree of perceptual similarity between different musical instrument timbres highly correlates with the MMN amplitude.[9] This correlation between the MMN parameters and behavioural responses imply that preattentive neural functions determine the accuracy of the subsequent attentive processes.[2,3,10]

Since the MMN is elicited without attention, it received particular interest during past years among clinicians who need to evaluate the integrity of neurocognitive functions in patients unable or unwilling to participate in a neuropsychological (behavioural) testing. For instance, it was found that the MMN elicitation predicts the recovery of consciousness in comatose patients[11,12]) (for reviews, see Refs 13, 14). In addition, the deficits in MMN elicitation might help target perceptual training or rehabilitation in milder perceptual disorders. For instance, recently it was revealed that automatic discriminative functions of dyslexic adults are worse than those of control subjects.[15–18] In addition, the MMN is also correlated with the improvement of the reading performance in dyslexic children.[19]

Since differences in subjects' attention or motivation might confound studies in music psychology especially in between-group comparisons, the MMN recordings offer a promising means in determining the accuracy of neural sound representations prior to involvement of the perceptual and cognitive functions also in music studies. In the following, an overview on studies using musical sounds and sound successions to investigate the automatic neural sound processing will be given. In addition, recent studies comparing musicians and nonmusicians will be reviewed.

## MMN generators for speech vs musical sounds

MMN is generated mainly in the primary auditory cortex or in its immediate vicinity. This has been indicated by several brain research methods such as magnetoencephalography (MEG),[20,21] intracranial electrode recordings,[22,23] positron emission tomography (PET),[24] and functional magnetic resonance imaging (fMRI).[25–27] In addition, this is reflected in electric recordings by polarity reversal above the Sylvian fissure from the fronto-central negative maximum to positivity in mastoids leads (when the nose reference is used) (for a review, see Ref. 28).

In late 1990s, a whole-head MEG experiment was conducted to determine whether the auditory cortex has spatially distinct areas for encoding musical vs phonetic sounds.[29] The subjects, while watching a silent movie with subtitles, were presented with frequent and infrequent phonemes (/e/ vs /o/) or chords (A major vs A minor). These phonetic and musical stimuli were matched in complexity as well as in the magnitude of the frequency change embedded in them. It was found that the source of the MMNm elicited by the deviant phoneme or chord was located posteriorly to the source of an earlier P1m component (see footnote a). In addition, the MMNm source for a phoneme change was located superiorly to that of the chord change. On the contrary, P1m sources did not differ between

[a] Early exogenous ERP components and their MEG equivalents are mainly determined by the physical sound characteristics, for instance, the intensity and rise time, while endogenous ERP/MEG components with longer onset latencies are modulated by cognitive and attentional factors.[2]

phonemes and chords. These data thus suggest that there are distinct cortical areas specialized in representing phonetic and musical sounds in both hemispheres. In other words, the functional specialization does not only cover stimulus complexity (as shown by Alho *et al.* comparing MMNm generator loci between single tones and chords)[30] but also the informational content (phonetic vs musical).[29] However, this specialization is not present prior to memory-related auditory processing as suggested by the dissociation between the present P1m and MMNm data.

Additionally, the data reviewed above indicated that in the right hemisphere, the MMNm was larger in amplitude for changes among chords than among phonemes. However, in the left hemisphere, no corresponding dominance for phoneme changes was found when compared with chord changes. This result was obviously discrepant with the previous results.[31,32] To solve this discrepancy, a further study was conducted with PET technique (see footnote b).

This PET study[24] used the phonetic and musical sounds developed for the MEG study described above.[29] However, the phoneme duration was prolonged from originally used 200 ms until 400 ms, since in pilot studies, presence of deviant phonemes elicited no changes in brain metabolism with phonemes of 200 ms duration. The subjects were concentrating on classifying the gender of visually presented words while they were presented with sound sequences consisting of (1) both deviant and standard sounds or (2) standard sounds only (phonemes and chords in separate sequences). The data, obtained after subtracting the activity elicited by the standard sounds only from that elicited by both deviant and standard sounds when intermixed, showed that the change from vowel /e/ to /o/ was processed in the left auditory cortex, more specifically, in the middle and supratemporal gyri. In a mirror-like manner, the change from A major to A minor chord was processed in the right auditory cortex, in the supratemporal gyrus. These data thus indicate that hemispheric specialization for phonetic vs musical processing, previously seen in dichotic-listening studies and brain-imaging studies using an active task, may be present even during the performance of a task unrelated to the sound stimulation. However, this phenomenon is very vulnerable to changes in stimulus parameters as indexed by the importance of an adequate sound duration[24,29] (see also Refs 36–38 for recent evidence about the effects of acoustic noise on lateralization of phonetic processing).

## Sound complexity and perceptual accuracy

All natural speech and instrumental sounds consist of several parallel harmonic partials, which, despite their wide frequency range, produce a percept of one single sound. The presence of harmonic partials is known to facilitate pitch naming in the possessors of absolute pitch (for a review on absolute pitch, see Ref. 39) as well as the tuning of music instruments.[40] Only recently it was systematically investigated whether harmonic partials facilitate pitch discrimination in ignore and attend conditions in nonmusicians.

[b] Localizing the active neural population with MEG (and EEG) necessarily involves inverse modelling of the active neural source. Accuracy of this operation depends from the adequacy of the head model available.[33] This source of uncertainty can be avoided by using brain-imaging techniques such as PET and fMRI. Those methods detect directly the locus of activation change in brain metabolism, presumably caused by increased neural activity.[34,35]

In the first study, the subjects were presented with pitch changes of 2.5, 5, and 10 per cent at 500 Hz sounds, first, during reading a book and, second, during a target-detection task.[41] It was found that they detected more accurately the sounds which consisted of three lowest partials than pure sinusoidal tones. Correspondingly, the frequency-MMN amplitude was larger and latency shorter in spectrally rich tones than in pure tones. This result suggests that pitch discrimination, under attentional control as well as prior to that, is facilitated by the presence of the harmonic partials.

Subsequently it was determined whether adding harmonic partials or prolonging the sound would further facilitate pitch discrimination.[42] To this aim, the frequency-MMN was recorded to a 2.5 per cent pitch change in pure tones with only one sinusoidal frequency component (500 Hz) and in spectrally rich tones with three (500... 1500 Hz) and five (500... 2500 Hz harmonic partials. Stimuli were 100 and 250 ms in duration (in separate blocks). During these recordings, subjects concentrated on watching a silent movie. The data show that the MMN amplitude was enhanced with spectrally rich sounds when compared with pure tones, with no difference between spectrally rich tones having three or five partials. Moreover, the prolonged sound duration did not significantly enhance the MMN. This suggests, first of all, that relatively few harmonic partials (at least when they

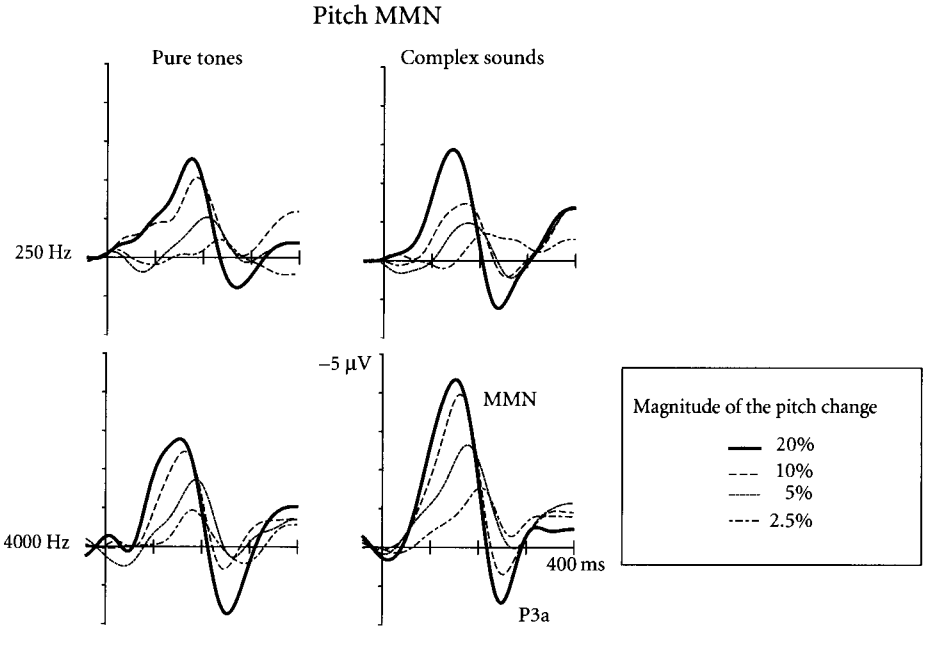



**Figure 19.1** The MMN elicited by 2.5, 5, 10, and 20 per cent pitch changes in pure sinusoidal tones (left) and in spectrally rich sounds (right) at 250 Hz (top) and 4000 Hz (bottom). The lines denote difference waves in which the event-related potential elicited by the standard tone has been subtracted from the deviant-tone ERP. The MMN was consistently larger when the pitch change was presented among spectrally rich sounds. Its amplitude also systematically reflected the magnitude of the pitch change. Adapted from Ref. 43.

fall inside the dominance region of pitch perception) are sufficient to sharpen the neural representations underlying pitch discrimination. Second, prolonging sound duration does not further facilitate pitch discrimination with spectrally rich sounds, at least when duration of 100 ms is exceeded.

Most recently, the neural and behavioural accuracy of frequency discrimination across different frequency ranges (250–4000 Hz) with pure sinusoidal vs spectrally rich sounds was compared.[43] Also magnitude of the frequency change (2.5, 5, 10, 20 per cent) was varied in both MMN and behavioural paradigms. The data showed that, in general, the spectrally rich sounds elicited larger MMN across the whole frequency range. Additionally, the changes at the middle frequencies elicited larger MMN, which, in parallel, had a shorter latency than at the lowest and highest frequencies. Replicating earlier findings, the MMN amplitude and latency reflected the magnitude of the frequency change.[5] These neural indices were mirrored in behavioural performance as well, the hit rate being the most accurate and the reaction time fastest at the middle frequency range, with the widest deviants, and with harmonically rich sounds. To summarize, the facilitation in pitch discrimination caused by the spectrally rich sound structure exceeds relatively broad frequency range.

### Sequential stimulation

Behavioural evidence shows that, in addition to spectral complexity, also temporal proximity of adjacent sounds may facilitate pitch encoding in its several forms. For instance Dewar *et al.*[44] presented subjects with standard tonal or atonal sequences followed by two comparison tones in the no-context condition or by two comparison sequences in the full-context condition. The listeners recognized the target tone more accurately under the full-context condition than under the no-context condition. In addition, they performed more accurately when the target tone was embedded in tonal rather than atonal sequences and musicians outperformed nonmusicians in all recognition tasks. This and related findings inspired us to investigate if the presence of Western, familiar sound context facilitates pitch processing even at the preattentive level and, further, whether it might be enhanced by musical expertise.[45]

To this end, 10 musicians and 10 nonmusicians were presented with a 144-Hz frequency change in three contexts. First, the frequency change was among single sounds (554 vs 698 Hz). Second, it was embedded among sound sequences in which the subsequent sounds belonged to the Western musical scale (440–493–**554**–587–659 vs 440–493–**698**–587–659 Hz). Third, the change was presented within sound sequences, which were composed of arithmetically determined intervals, compromising between Western semitones and whole-tone steps (446–467–**499**–547–622 vs 446–467–**643**–547–622 Hz). During the recordings, the subjects were (once again) reading a book of their own choice.

The MMN data indicate that, first of all, MMN amplitude was larger when the frequency change occurred among temporally complex sound sequence with the familiar scale than when it occurred among the unfamiliar scale and, further, larger in unfamiliar scale than among single tones (see Figure 19.2). This suggests that musical context facilitates pitch discrimination most effectively when it is familiar to the subjects. Second, with both familiar and unfamiliar complex sound sequences, the musicians had a generalized facilitation of the MMN in terms of a shorter MMN latency when compared to that by nonmusicians.

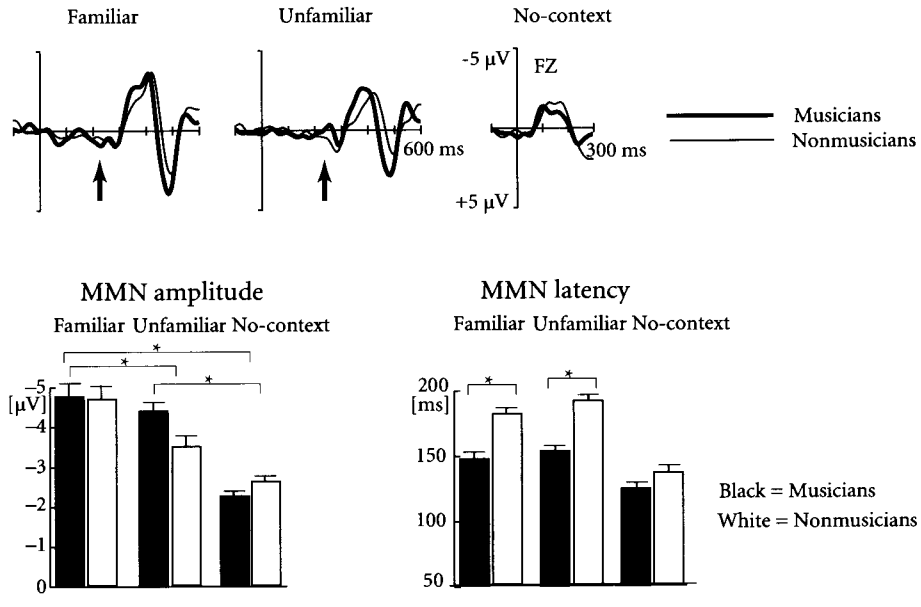



Figure 19.2 Upper panel: The MMN elicited by a pitch change within familiar (tonal) pattern (left), Unfamiliar (atonal) pattern (middle) and in pure sinusoidal tones (right) in musicians (thick line) and nonmusicians (thin line) at Fz electrode. The arrows denote the onset of the pitch change in deviant patterns. The lines denote difference waves in which the event-related potential elicited by the standard tone has been subtracted from the deviant-tone ERP. Bottom panel: the MMN mean amplitude (left) and latency (right) these conditions in musicians (black) and nonmusicians (white) (s.e.m. at the top of each bar). The values were averaged across 10 frontocentral electrodes.

In both subject groups, the MMN amplitude was larger in familiar than in unfamiliar condition and, further, smallest in no-context condition. In addition, in musicians, the MMN latency was shorter in musicians than in nonmusicians. Adapted from Ref. 45.

This suggests that the musicians can adopt the unfamiliar sound environment more readily than nonmusicians.

## Abstractions and expectations

Despite the emphasis given in above-described ERP studies on sound processing accuracy at the sensory level in audition, music perception and cognition cannot be reduced merely to a successful sensory-level processes. For instance, despite transpositions and modulations within a musical piece, the 'core' of the melody can be easily remembered and recognized, in other words, abstract sound and melody features are implicitly extrapolated from the continuously changing acoustic input. Consequently, it was questioned whether abstract sound relations and primitive concepts such as 'ascending' or 'descending' interval, pitch continuum, or chord cadences could also be neurally encoded during a performance of a parallel task.

### Abstractions

The first studies investigated the interval perception using tone pairs with a whole-tone frequency change within a tone pair.[46,47] To determine whether the direction of a sound change within a pair is encoded by the auditory cortex irrespectively of the absolute sound frequency, the pitch of the standard tone pairs was randomly varied between five different frequency levels (AB, CD, BC, ..., EF, DE, AB, ...).[47] Deviant pairs had either reversed frequency change (BA, ..., FE) or a frequency repetition (AA, BB, ..., EE), also at five different frequency levels. It was found that despite this complexity in stimulation and the parallel reading task given to the subjects, the deviant sound pairs elicited MMN, indicating that the change in the direction of the interval was discriminated in the auditory cortex. This line of research has been continued by Paavilainen and his colleagues.[48] They found that in addition to the direction of the tone interval, also the interval size is automatically encoded in the auditory cortex. This was evidenced by the MMN elicitation by smaller or larger intervals among minor seventh standard intervals (all intervals were randomly presented at 10 frequency levels). Furthermore, the cortical interval representation may also be formed on the basis of converging information from both ears as indicated by the MMN elicitation when the two successive sounds of an interval are delivered to separate ears.[49]

In sum, the present results show that the human auditory cortex can automatically encode highly complex and abstract auditory information (see Ref. 50 for a review).

### Expectations

In the two studies to be summarized below, the regularity of the sound change itself created a rule about the sound continuum, in other words, an expectation for the next sound of the sequence either in terms of pitch or sound location. An MMN elicitation by abrupt changes in such sound continuums would strengthen the above-documented conclusions about the 'intelligence' of automatically formed neural sound representations in extracting invariant rules from the continuously varying sound information.

Tervaniemi *et al.* presented reading subjects with regularly spaced sound sequences consisting of regular pitch decrease of one semitone as the standard event.[51] In two conditions, the sounds were Shepard sounds and in two conditions they were pure sinusoidal tones. Randomly, a pitch repetition or an ascending pitch interrupted this line of regularly descending pitch. The MMN was elicited by ascending sounds, be them Shepard or sinusoidal sounds, and by pitch repetition when they were Shepard sounds. These data suggested the existence of primitive expectation formation even without conscious attention being directed to the sounds.

To confirm and expand the previous data obtained by using sound sequences based on regular pitch change, the follow-up experiment was conducted in which the regularity was created by virtual sound movement.[52] To create a subjective feeling of sound movement, sounds were presented via nine loudspeakers extending in a horizontal symmetrical arc in front of the subject. Short noise bursts were presented about three times a second, continuously moving back and forth between the two extreme loudspeakers. Occasionally the regular virtual sound movement was broken by the sound skipping two loudspeakers,

which is experienced as a jump in the movement. ERPs recorded from reading subjects showed that 'jumps' elicited a large centrally negative response resembling the MMN. This suggests that expectation about the source location of the next sound was automatically formed (for related data about automatic sound source discrimination, see Refs 53–55).

### Chord cadences

How syntactic rules of major-minor tonal music are represented in the brain has been addressed by a chord-cadence paradigm.[56] Cadences of five chords (first four chords 600 ms, fifth chord 1200 ms in duration) were continuously presented at several randomly varying frequency levels. The majority of the chords were presented with synthesized piano timbre, the task of the subject being to detect the chords presented with a different timbre (e.g. celesta, marimba). The study was conducted to reveal whether chords violating musical regularities would elicit differential ERPs compared to 'music-syntactically' appropriate chords, although the subject's attention was directed away from these regularity-violating chords.

The ERPs, recorded from subjects without any formal training in music, were found to differ at two latencies between harmonically appropriate chords and Neapolitan chords violating the context established by the previous chords. The first difference was labelled as the early right anterior negativity (ERAN) according to its relatively early peak latency at around 200 ms, right preponderant distribution, and negative polarity (its linguistic counterpart being termed as the left anterior negativity, ELAN[57]). (See also Ref. 58 for previous related data in music neurocognition.) The second difference was called the N5 (negative polarity and peak latency at around 500 ms; without hemispheric dominance in distribution). The authors concluded that musically untrained subjects may have several neural mechanisms activated for extrapolating the tonal harmony and musical regularities of musical sequences, and, further, that this may occur despite the direction of their conscious attention. In fact, the second argument was confirmed in a subsequent experiment. There, the subjects were reading a book during the stimulation (Reading condition) or listening to the cadences and indicating the presence of a Neapolitan chord (Attend condition).[59] The ERAN did not differ between these two conditions. MEG evidence suggests that the ERAN is generated in the inferior fronto-lateral cortex (in the left hemisphere known as Broca's area), suggesting the involvement of the language areas in music-syntactic processing.[60]

Taken together, the auditory system appears to be able to encode automatically regularity of the sound environment. For this regularity extraction, acoustic repetitiveness of the sound material is not necessary; instead, also more abstract sound relations (direction of the sound change, interval size, chord progression) are encoded.

## Musical expertise

### Pitch encoding

Musical expertise is reflected in early auditory ERPs even 100 ms after the sound onset, at least when the attention of the subject is not directed away from the sounds.[61] Likewise, it is reflected in the cognitive late-latency ERPs such as P3 and LPC when the subjects are

required to attend to the sound stimulation.[62–64] Recently it was also investigated whether musical expertise may be reflected in the auditory ERPs during a performance of a parallel task in another modality. The first set of these studies addressed the accuracy of the violin players' and, more generally, the musicians' auditory system to process slight pitch changes. In the first study, 3-part major chords with a perfect major were used as the standard stimulus.[65] The deviant stimulus was the same chord as the standard stimulus, except that the middle tone of the chord was marginally mistuned (<1 per cent), leading the chord towards a minor chord. This stimulation was presented to subjects while they were reading a book (the first and third blocks of the experiment; reading condition) and while they were asked to detect the deviant chords (the second block; discrimination condition). In the reading condition, the deviants elicited the MMN only in musicians. In the discrimination condition, nonmusicians detected about 10 per cent and musicians about 80 per cent of the deviant chords. In this condition, the parallel ERP recordings showed a significant MMN followed by N2b and P3b deflections in musicians. This N2b–P3 complex reflects higher cognitive processes concerned with the conscious detection and evaluation of deviants. Nonmusicians had a small (but statistically significant) MMN without a subsequently elicited N2b or P3. The third block was presented to see whether the intermediate attentive task facilitated subjects' automatic pitch processing. This, however, was not the case. Musicians showed a MMN that did not differ from that elicited in the first block while nonmusicians showed no MMN. These results suggest that professional violin players automatically are able to detect tiny pitch changes in auditory information, which were undetectable for nonmusicians and, further, that these automatic functions could not be modified by attentional manipulations. Interestingly, the MMN evoked by a small (<1 per cent) or large (10 per cent) pitch change in pure sinusoidal tones did not differentiate violinists and nonmusicians. This suggests that the superior pitch processing accuracy of the violinists was activated only when the pitch change was presented among musical sounds.

In the subsequent experiment the superiority of musicians in automatic and attentive pitch discrimination, originally observed in violin players,[65] was further investigated.[66] Thirteen professional musicians (guitarists, pianists, and wind instrument players) and 13 nonmusicians were presented with harmonically rich sounds of 300 ms in duration. Standard sounds were of 528 Hz in fundamental frequency while deviant sounds were 0.7, 2, or 4 per cent higher. It was found that in the discrimination condition, the musicians detected the pitch changes faster and more accurately than the nonmusicians. This was reflected also in their P3b component, which was significantly larger in musicians than in nonmusicians. However, the event-related brain responses recorded during the reading condition showed a different data pattern—the MMN or P3a amplitude or latency did not differ between the musicians and nonmusicians.

To summarize, the data reviewed above on pitch discrimination in violin players suggest that musical expertise is reflected at preattentive cortical auditory processing especially when the sound structure is musically relevant and when the stimulus change is of musically relevant magnitude. However, in musicians with other instrumental training, despite their superiority in pitch discrimination during attentional listening, preattentive neural mechanisms were not enhanced when compared with those of nonmusicians. This implies that automated neural functions in pitch discrimination are most sensitive in musicians who most consistently need such ability. In other musicians, such a sensitivity to pitch

changes exists but is activated only during attentive listening. Whether we here evidence an effect of differential training on brain plasticity or a result of appropriately chosen musical instrument, remains unresolved on the basis of the present evidence. However, recent evidence obtained by using both structural and functional methods of cognitive neuroscience promotes the importance of training in reorganizing the brain functions.[61,67,68]

### Temporal encoding

Very recently, also temporal domain of sound processing has attracted attention in musician–nonmusician comparisons in MMN paradigm. This is valuable especially if the temporal accuracy of the event-related potentials is taken into account—it allows one to track the processing of very minute changes also in temporally complex sound sequences.

The first study investigated the accuracy of musicians vs nonmusicians to preattentively encode regularity in relatively fast sound presentation rate and to react to infrequent sound omissions in that sound stream.[69] In separate conditions, the sinusoidal sounds of 50 ms were delivered every 100, 120, 180, and 220 ms to the subjects who concentrated on a reading task. 15 musicians and 15 subjects without any formal training in music were employed. Sound omissions occurred with the probability of 3 per cent. It was found that while the sound omissions evoked the MMN in musicians in all conditions, in nonmusicians the MMN was present only with two shortest stimulation rates. This suggests that musicians have a prolonged window of sound integration.[70] Furthermore, the results of the second experiment of these authors suggest that the temporal processing of musicians within this window is also more accurate than in nonmusicians. In that study, musicians and nonmusicians were presented with tone pairs. In the standard pair, the sound onsets were separated by 150 ms whereas in the deviant pairs, they were separated by 100 or 120 ms. It was found that the MMN was elicited in musicians by both deviant pairs but in nonmusicians only by 100 ms pair.

The second set of related experiments investigated the accuracy and efficacy of musicians vs nonmusicians to preattentively encode sound sequences up to 1.5 s and to detect violations in the regularities established by the sequences.[71] In the first of these experiments, the sound sequences were either 750 ms or 1.5 s in duration (in separate conditions). Four tones, ascending in pitch, were presented in a looped manner. The deviant sequence consisted of five tones, thus being one tone longer than the standard sequence. The MMN elicitation implied that musicians were able to encode the length and frequency content of such a subsequence with both sequence durations. In contrast, in nonmusicians, no MMN was elicited. In the second experiment, the standard sound sequences were also either 750 ms or 1.5 s in duration but were made very simple in their pitch structure: they consisted of four repetitions of the same pitch. After four tones, the pitch of all of them was changed. Also here, the deviant sequences were one tone longer. It was found that the MMN was elicited by the tone sequence lengthening in both subject groups with 750 ms sequences and in musicians also with 1.5 s sequences. These data suggest that grouping of sequential sounds can take place on a preattentive processing level irrespective of musical skill with moderate sequence durations, but that professional musicians have more advanced preattentive ability to group sequential sounds than nonmusicians when sequences are lengthy. In other words, in its simple forms, auditory grouping does not differ between subjects with different levels of expertise. However, with more complicated

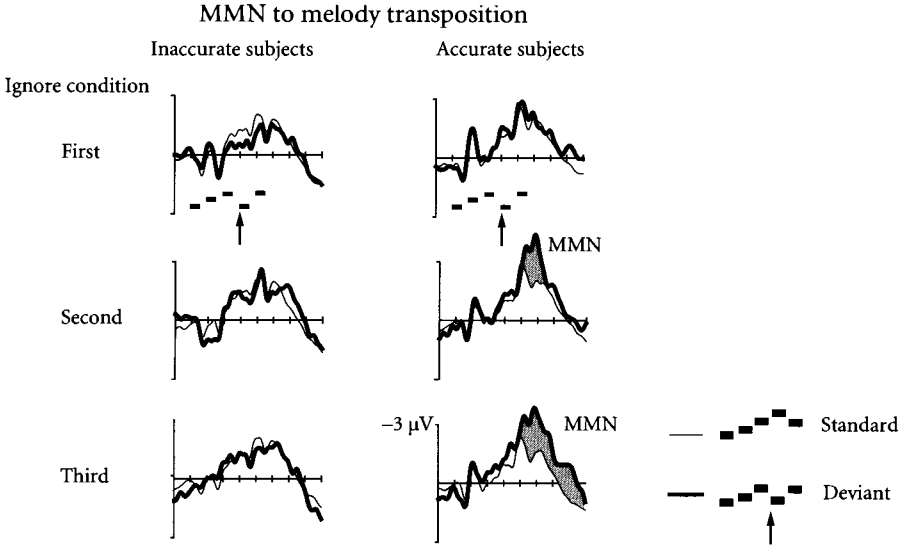



**Figure 19.3** Event-related potentials elicited by the standard (90 per cent; thin line) and the deviant (10 per cent; thick line) melodic patterns. In the accurate subjects, an MMN was elicited during the second and third Ignore conditions following the first attend condition (right column). In the inaccurate subjects, no MMN was observed at any phase of the experiment despite intermediate attend conditions (left column). Adapted from Ref. 72.

pitch information or longer sound sequences, the musicians have more advanced encoding ability even during a performance of a parallel task.

Recently it was also studied whether musicians and nonmusicians may encode melodic transpositions without attentional effort.[72] The subjects were presented with short melodic patterns of 0.5 s at 12 different frequency levels within one octave. In alternating order, the experimental session consisted of three Ignore conditions (during which the subjects watched a silent video with subtitles) and three Attentive conditions (during which they detected a target with a different melody contour). The data indicated that especially musicians who perform music mainly without a score easily learn to detect contour changes in the pitch transposed melodic patterns (see Figure 19.3). However, even those musicians had no MMN during the first Ignore condition, in other words, also they needed some conscious, explicit effort to learn to extract the melodic contour from the constantly varying sound material. After learning, their auditory cortex detected violations from the standard melodic contour even when their attention was directed away from the sounds. Interestingly, all non-musicians and the 'score-dependent musicians' were not assisted by the presence of the active task.

## Conclusions

The data reviewed above was acquired by electromagnetic and hemodynamic recordings on healthy adult subjects. In majority of those studies, an experimental paradigm was

employed in which the subject performed a parallel task outside the auditory modality during the stimulation and brain recordings. Thus, these findings offer fundamental insight to the brains's ability to encode and differentiate acoustically complex sounds despite the focus of the listener's attention.

The results indicated, first, that the cortical networks are able to automatically encode temporally and spectrally complex musical sounds as well as their abstract relations. Moreover, it was shown that pitch changes among spectrally complex sounds and temporally complex sound patterns are more readily discriminated than equivalent pitch changes among pure single sinusoidal sounds. These findings might rephrase our views on the ability of the human brain to represent the continuously changing sound environment beyond the 'listener's' attentional focus. In other words, although music and music sounds are acoustically and cognitively demanding entities, majority of relevant sound information is neurally encoded even if we are not consciously aware of it.

Second, the loci of the cortical sound representations were found to differ between musical and phonetic sounds both within and between the cerebral hemispheres. This suggests that the auditory cortex is able to represent not only the acoustical but also the informational sound content. Further, the data reviewed above imply that the tonality processing as probed by neurocognition of chord cadences involves also the brain regions primarily devoted to language, that is, the Broca's area. These two lines of research may first seem to be in conflict. However, the level of auditory material and thus the level of processing is not comparable. While isolated sounds represent one fragment of an auditory scene, a chord cadence is actually already an auditory scene of its own. In future, it will hopefully be possible to further illuminate the degree and timing of the neurocognitive modularity of music and language processing.

Third, in several experimental paradigms, musical expertise was reflected in automatically elicited brain responses to violations in pitch and temporal sound features as a response enhancement and/or latency decrease. It was also indicated that musical expertise does not guarantee facilitated neuronal processing of music sounds or sound sequences. Based on the above reviewed findings (see also Ref. 73), it might be argued that the commonly used differentiation between 'musicians' when contrasted with 'nonmusicians' does not necessarily provide a researcher with an accurate enough categorization. Preferably, we should identify and subdivide musicians more accurately on the basis of their major instruments (e.g. whether just or fixed tuning is used); the importance of pitch vs timing cues in the genre and instrument, and their relationship to musical score. Also the quantity and quality of the musical activities the 'nonmusical' subjects have, for instance, in terms of dance, sports, or music listening, should be taken into account. Hopefully, by optimizing the contrasts between the subject groups in their relationship to music in its broadest meaning, the purest forms of neurocognitive expertise in musicians could be more readily unveiled.

To summarize, during past decade, plenitude of electromagnetic evidence together with increasing number of hemodynamic data underlined the importance of the automatically activated neural networks in the auditory cortex also in musically relevant forms of cognition. In future, our task is to further illuminate the degree of modularity in audition for

music vs speech sounds and also the role of automatically activated sound representations in musical expertise.

## Acknowledgements

The author thanks the Academy of Finland, University of Helsinki, and the Marie Curie Fellowship of the European Commission. Special thanks to Elvira Brattico, M.A., for her comments on an earlier version of the manuscript.

## References

1. **Shiffrin, R. M. and W. Schneider** (1977) Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychol. Rev.* 84, 127–90.
2. **Näätänen, R.** (1992) *Attention and Brain Function.* Hillsdale: Lawrence Erlbaum Associates.
3. **Näätänen, R. and I. Winkler** (1999) The concept of auditory stimulus representation in cognitive neuroscience. *Psychol. Bull.* 125, 826–59.
4. **Tervaniemi, M.** (1999) Pre-attentive processing of musical information in the human brain. *J. New Music Res.* 28, 237–45.
5. **Tiitinen, H., P. May, K. Reinikainen, *et al.*** (1994) Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature* 372, 90–2.
6. **Kraus, N., T. McGee, T. D. Carrell, *et al.*** (1996) Auditory neurophysiologic responses and discrimination deficits in children with learning problems. *Science* 273, 971–3.
7. **Tremblay, K., N. Kraus, and T. McGee** (1998) The time course of auditory perceptual learning: neurophysiological changes during speech-sound training. *Neuroreport* 9, 3557–60
8. **Amenedo, E. and C. Escera.** (2000) The accuracy of sound duration representation in the human brain determines the accuracy of behavioural perception. *Euro. J. Neurosci.* 12, 2570–4.
9. **Toiviainen, P., M. Tervaniemi, J. Louhivuori, *et al.*** (1998) Timbre similarity: convergence of neural, behavioral, and computational approaches. *Music Percept.* 16, 223–41.
10. **Novak, G. P., W. Ritter, H.G. Vaughan Jr, *et al.*** (1990) Differentiation of negative event-related potentials in an auditory discrimination task. *Electroencephal. Clin. Neurophysiol.* 75, 255–75.
11. **Kane, N. M., S. H. Curry, C.A Rowlands, *et al.*** (1996) Event-related potentials—neurophysiological tools for predicting emergence and early outcome from traumatic coma. *Intensive Care Med.* 22, 39–46.
12. **Fischer, C., D. Morlet, P. Bouchet, *et al.*** (1999) Mismatch negativity and late auditory evoked potentials in comatose patients. *Clin. Neurophysiol.* 110, 1601–10.
13. **Gene-Cos, N., H. A. Ring, R. C. Pottinger, *et al.*** (1999) Possible roles for mismatch negativity in neuropsychiatry. *Neuropsychiatry, Neuropsychol. Behav. Neurol.* 12, 17–27.
14. **Näätänen, R. and M. Hoke** (eds) (2000) Special issue on mismatch negativity and its clinical applications. *Audiol. Neuro-Otol.* 5, 101–246.
15. **Schulte-Korne, G., W. Deimel, J. Bartling, *et al.*** (1998) Auditory processing and dyslexia: evidence for a specific speech processing deficit. *Neuroreport* 9, 337–40.
16. **Schulte-Korne G., W. Deimel, J. Bartling, *et al.*** (1999) Pre-attentive processing of auditory patterns in dyslexic human subjects. *Neurosci. Lett.* 276, 41–4.

17. Baldeweg, T., A. Richardson, S. Watkins, *et al.* (1999) Impaired auditory frequency discrimination in dyslexia detected with mismatch evoked potentials. *Ann. Neurol.* 45, 495–503.
18. Kujala, T., K. Myllyviita, M. Tervaniemi, *et al.* (2000) Basic auditory dysfunction in dyslexia as pinpointed by brain-activity measurements. *Psychophysiology* 37, 262–6.
19. Kujala, T., K. Karma, R. Ceponiene, *et al.* (2001) Plastic neural changes and reading improvement caused by audio–visual training in reading-impaired children. *PNAS* 98, 10509–14.
20. Sussman, E., I. Winkler, and M. Huotilainen (2002) Top–down effects on stimulus-driven auditory organization. *Cogn. Brain Res.* 13, 393–405.
21. Tiitinen, H., K. Alho, M. Huotilainen, *et al.* (1993) Tonotopic auditory cortex and the magnetoencephalographic (MEG) equivalent of the mismatch negativity. *Psychophysiology* 30, 537–40.
22. Kropotov, J. D., R. Näätänen, A.V. Sevostianov, *et al.* (1995). Mismatch negativity to auditory stimulus change recorded directly from the human temporal cortex. *Psychophysiology* 32, 418–22.
23. Halgren, E., K. Marinkovic, and P. Chauvel (1997) Generators of the late cognitive potentials in auditory and visual oddball tasks. *Electroencephal Clin. Neurophysiol.* 106, 156–64.
24. Tervaniemi, M., S. V. Medvedev, K. Alho, *et al.* (2000) Lateralized automatic auditory processing of phonetic versus musical information: a PET study. *Hum. Brain Map.* 10, 74–9.
25. Opitz, B., A. Mecklinger, D. Y. von Cramon, *et al.* (1999) Combining electrophysiological and hemodynamic measures of the auditory oddball. *Psychophysiology* 36, 142–7.
26. Opitz, B., A. Mecklinger, A. D. Friederici, *et al.* (1999) The functional neuroanatomy of novelty processing: integrating ERP and fMRI results. *Cerebr. Cortex* 9, 379–91.
27. Opitz, B., T. Rinne, A. Mecklinger, *et al.* (2002) Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. *Neuroimage* 15, 167–74.
28. Alho, K. (1995) Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear Hear* 16, 3–50.
29. Tervaniemi, M., A. Kujala, K. Alho, *et al.* (1999) Functional specialization of the human auditory cortex in processing phonetic and musical sounds: a magnetoencephalographic study. *Neuroimage* 9, 330–6.
30. Alho, K., M. Tervaniemi, M. Huotilainen, *et al.* (1996) Processing of complex sounds in the human auditory cortex as revealed by magnetic brain responses. *Psychophysiology* 33, 369–75.
31. Näätänen, R., A. Lehtokoski, M. Lennes, *et al.* (1997) Language-specific phoneme representations revealed by the electric and magnetic brain responses. *Nature* 385, 432–4.
32. Rinne, T., K. Alho, P. Alku *et al.* (1999) Hemispheric asymmetry of cortical activation as reflected by the mismatch negativity reveals when a sound is processed as speech. *Neuroreport* 10, 1113–17.
33. Hämäläinen, M., R. Hari, R. J. Ilmoniemi, *et al.* (1993) Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.* 65, 413–97.
34. Tervaniemi, M. and T. L. van Zuijen (1999) Methodologies of brain research in cognitive musicology. *J. New Music Res.* 28, 200–8.
35. Rugg, M. D. (1999) Functional neuroimaging in cognitive neuroscience. In C. M. Brown and P. Hagoort (eds) *The Neurocognition of Language.* Oxford: Oxford University Press, pp. 15–36.
36. Shtyrov, Y., T. Kujala, J. Ahveninen, *et al.* (1998) Noise-induced shift in hemispheric lateralization of speech processing. *Neurosci. Lett.* 251, 141–4.
37. Shtyrov, Y., T. Kujala, R. J. Ilmoniemi, *et al.* (1999) Noise affects speech-signal processing differently in the cerebral hemispheres. *Neuroreport* 10, 2189–92.

38. **Herrmann, C. S., U. Oertel, Y. Wang, *et al.*** (2000) Noise affects auditory and linguistic processing differently: an MEG study. *Neuroreport* 11, 227–9.

39. **Takeuchi, A. H. and S. H. Hulse** (1993) Absolute pitch. *Psychol. Bull.* 113, 345–61.

40. **Sundberg, J.** (1991) *The Science of Musical Sounds.* San Diego: Academic Press.

41. **Tervaniemi, M., T. Ilvonen, J. Sinkkonen, *et al.*** (2000) Harmonic partials facilitate pitch discrimination in humans: electrophysiological and behavioral evidence. *Neurosci. Lett.* 279, 29–32.

42. **Tervaniemi, M., E. Schröger, M. Saher, *et al.*** (2000) Effects of spectral complexity and sound duration in complex-sound pitch processing in humans—a mismatch negativity study. *Neurosci. Lett.* 290, 66–70.

43. **Tervaniemi, M., N. Novitski, M. Huotilainen, *et al.*** (2002) Frequency discrimination at different frequency levels as reflected by electrophysiological and behavioral indices. Submitted.

44. **Dewar, K., L. L. Cuddy,** and **D. J. K. Mewhort** (1977) Recognition memory for single tones with and without context. *J. Exp. Psychol.* 3, 60–7.

45. **Brattico, E., R. Näätänen,** and **M. Tervaniemi** (2002) Context effects on pitch perception in musicians and non-musicians: evidence from ERP recordings. *Music Percept.* 19, 199–222.

46. **Tervaniemi, M., J. Saarinen, P. Paavilainen, *et al.*** (1994) Temporal integration of auditory information in sensory memory as reflected by the mismatch negativity. *Biol. Psychol.* 38, 157–67.

47. **Saarinen, J., P. Paavilainen, E. Schröger, *et al.*** (1992) Representation of abstract stimulus attributes in human brain. *Neuroreport* 3, 1149–51.

48. **Paavilainen, P., M. Jaramillo, R. Näätänen, *et al.*** (1999) Neuronal populations in the human brain extracting invariant relationships from acoustics variance. *Neurosci. Lett.* 265, 179–82.

49. **Paavilainen, P., M. Jaramillo,** and **R. Näätänen.** (1998) Binaural information can converge in abstract memory traces. *Psychophysiology* 35, 483–7.

50. **Näätänen, R., M. Tervaniemi, E. Sussman, *et al.*** (2001) 'Primitive intelligence' in the auditory cortex. *Trends Neurosci.* 24, 283–8.

51. **Tervaniemi, M., S. Maury,** and **R. Näätänen** (1994) Neural representations of abstract features in the human brain as reflected by the mismatch negativity. *Neuroreport* 5, 844–6.

52. **Winkler, I., W. Teder, M. Tervaniemi, *et al.*** (2002) Event-related potential (ERP) correlates of unexpected jumps of a continuously moving auditory stimulus. In preparation.

53. **Paavilainen, P., M. L. Karlsson, K. Reinikainen, *et al.*** (1989) Mismatch negativity to changes in the spatial location of an auditory stimulus. *Int. J. Psychophysiol.* 7, 342–3.

54. **Schröger, E.** and **C. Wolff** (1996) Mismatch response of the human brain to changes in sound location. *Neuroreport* 7, 3005–8.

55. **Winkler, I., M. Tervaniemi, E. Schröger, *et al.*** (1998) Processing of spatial cues used for auditory lateralization as revealed by the mismatch negativity. *Neurosci. Lett.* 242, 49–52.

56. **Koelsch, S., T. C. Gunter, A. D. Friederici, *et al.*** (2000) Brain indices of music processing: 'non-musicians' are musical. *J. Cogn. Neurosci.* 12, 520–41.

57. **Hahne, A.** and **A.D. Friederici** (1999) Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. *J. Cogn. Neurosci.* 11, 194–205.

58. **Patel, A. D., E. Gibson, J. Ratner, *et al.*** (1998) Processing syntactic relations in language and music: an event-related potential study. *J. Cogn. Neurosci.* 10, 717–33.

59. **Maess, B., S. Koelsch, T. C. Gunter, *et al.*** (2001) Musical syntax is processed in Broca's area: an MEG study. *Nature Neurosci.* 4, 541–5.

60. **S. Koelsch, E. Schröger,** and **T. C. Gunter** (2002) Music matters: preattentive musicality of the human brain. *Psychophysiology* 39, 38–48.

61. **Pantev, C., R. Oostenveld, A. Engelien, *et al.*** (1998) Increased auditory cortical representation in musicians. *Nature* 392, 811–14.

62. **Verleger, R.** (1990) P3-evoking wrong notes: unexpected, awaited, or arousing? *Int. J. Neurosci.* 55, 171–9.

63. **Besson, M., F. Faïta,** and **J. Requin** (1994) Brain waves associated with musical incongruities differ for musicians and non-musicians. *Neurosci. Lett.* 168, 101–5.

64. **Crummer, G. C., J. P. Walton, J. W. Wayman, *et al.*** (1994) Neural processing of musical timbre by musicians, nonmusicians, and musicians possessing absolute pitch. *J. Acoust. Soc. Am.* 95, 2720–7.

65. **Koelsch, S., E. Schröger,** and **M. Tervaniemi** (1999) Superior attentive and pre-attentive auditory processing in musicians. *Neuroreport* 10, 1309–13.

66. **Tervaniemi, M., V. Just, S. Koelsch, *et al.*** (2002) Pitch-discrimination accuracy in musicians vs. non-musiciansan event-related potential and behavioral study. In preparation.

67. **Schmithorst, V. J.** and **M. Wilke** (2002) Differences in white matter architecture between musicians and non-musicians: a diffusion tensor imaging study. *Neurosci. Lett.* 321, 57–60.

68. **Pantev, C., L. E. Roberts, M. Schulz, *et al.*** (2001) Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport* 12, 1–6.

69. **Rüsseler, J., E. Altenmüller, W. Nager, *et al.*** (2001) Event-related brain potentials differ in musicians and non-musicians. *Neurosci. Lett.* 308, 33–6.

70. **Yabe, H., M. Tervaniemi, J. Sinkkonen, *et al.*** (1998) The temporal window of integration of auditory information in the human brain. *Psychophysiology* 35, 615–19.

71. **van Zuijen, T., I. Winkler, E. Sussman, *et al.*** (2002) Pre-attentive grouping of sequential sounds-an event-related potential study comparing musicians and non-musicians. Submitted.

72. **Tervaniemi, M., M. Rytkönen, E. Schröger, *et al.*** (2001) Superior formation of cortical memory traces for melodic patterns in musicians. *Learning Mem.* 8, 295–300.

73. **Kishon-Rabin, L., O. Amir, Y. Vexler, *et al.*** (2001) Pitch discrimination: are professional musicians better than non-musicians? *J. Basic Clin. Physiol. Pharmacol.* 12, 125–43.

CHAPTER 20

# PROCESSING EMOTIONS INDUCED BY MUSIC

L. J. TRAINOR AND L. A. SCHMIDT

## Abstract

Can music induce emotions directly and, if so, are these emotions experienced similarly to emotions arising in other contexts? We explore these questions from the perspective of neuroscience. Emotional processing has a deep evolutionary history because it is essential for self-regulation and social approach vs withdrawal. In humans, emotional processing involves autonomic, subcortical, and cortical structures. Despite the fact that music does not appear to have an obvious survival value for modern adults, research indicates that listening to music does activate autonomic, subcortical, and cortical systems in a manner similar to other emotional stimuli. We propose that music may be so intimately connected with emotional systems because caregivers use music to communicate emotionally with their infants before they are able to understand language.

## Introduction

Music is often referred to as the language of emotions, but there is considerable controversy as to whether music actually induces emotions directly in listeners and, if so, whether these musically induced emotions are the same as other emotional experiences. In this chapter, we examine evidence from neuroscience to explore these questions.

The question of meaning in music has been a subject of philosophical debate for a very long time.[1] Music is peculiar because in large part it appears to be a closed system, with musical meaning defined only in terms of music itself. Music can be referential in the sense of having an intended outside meaning (e.g. it can be about a thunderstorm, a war, or a love story), but this specific reference to events or things in the world is usually far from transparent to the listener. Music can also come to have meaning, in terms of convention (e.g. particular turns of phrase in Baroque music may be meaningful for listeners familiar with these conventions; particular *rags* may signal particular moods or moral qualities in Indian listeners educated in this tradition) or through association with events in the world (e.g. a particular song can signal bedtime for a child; lovers may associate a particular song with their courtship). However, conventional and associational referential meanings do not seem to account for the majority of meaningful responses to music. Rather, meaning appears to arise largely through the unfolding of sounds over time in relation to musical expectations.[2] In other words, to a large extent the meaning appears to be simply in the musical relations themselves.

Empirical psychomusicological research suggests that both adults and young children readily discriminate musical emotions.[3,4] The acoustic cues that differentiate emotions cover virtually all aspects of musical structure and include both structural (what is given by the composer) and performance characteristics.[5–8] For example, *sadness* is conveyed by quiet, slow, legato articulation and large deviations from metrical timing, whereas *happiness* or *joy* is conveyed by high-pitched, fast, staccato features and small variations from metrical timing.[9] However, music appears to express some emotions more precisely than others. Listeners generally agree on whether music is happy or sad, but there is less agreement when it comes to other emotions.[3,10] Acoustically, the differences between, for example, tenderness and sadness are rather subtle. According to Juslin,[9] both are quiet, have slow tempos, legato articulation, and large timing variations. Although they may differ in other subtle features, it does not appear that musical structure is set up to differentiate these emotions robustly. The notion that music may express readily only certain emotions must lead us to further question the nature of musical emotions, and how they relate to emotions in general.

It is likely surprising to psychologists to learn that some philosophers have argued that music does not express emotion, given the empirical data showing substantial agreement between listeners as to the emotion expressed in particular pieces of music. For example, a century and a half ago, Hanslick[11] proposed that music appreciation had nothing to do with emotion. Half a century ago, Langer[12] proposed that music bears some relation to emotion in that the rise and fall of tension in music, the interplay between uncertainty and resolution, mimics the time course of emotional experience. In this view, music does not express emotion, but we understand music through its similarity to emotional dynamics. On the other hand, Meyer[2] has argued that music does express emotions. According to his argument, musical emotions are induced through musical uncertainty, expectation for what will follow, and the way in which this uncertainty is resolved, similarly to how emotions are induced by other stimuli and events in the world.

Emotions can be classified in a number of ways. One approach is to extract the underlying dimensions of emotion, and to situate each specific emotion in this multi-dimensional space. Such analyses reveal two main dimensions of emotion: valence (negative to positive) and intensity (low to high).[13] Positive emotions are associated with approach behaviours and negative emotions with withdrawal behaviours. The ventromedial prefrontal cortex is thought to play a role in approach/withdrawal evaluations.[14] Although emotional processing in general tends to be more lateralized to the right than to the left,[15] there is also evidence that lateralization follows valence, with left prefrontal areas specialized for positive emotions and right prefrontal areas for negative emotions.[16–19] Despite the common view that 'music is the language of emotions', these models were developed without reference to musical stimuli. Music is an interesting stimulus in this regard, because in most cases it does not lead to overt approach or withdrawal behaviour. In fact, people often 'approach' sad music, in that they find it beautiful. Because music does not seem to be directly connected with approach/withdrawal behaviour, it is possible that music is *about* emotion, that music *communicates* emotional information, but that music does not directly *induce* emotions. One interesting question, then, is whether musical stimuli generate activity in

the prefrontal cortex, and whether differential activation can be seen for music expressing different emotions.

There is currently much debate among psychologists and neuroscientists as to the nature of emotional experience in general and its relation to cognition, behaviour, consciousness, and the sense of self. According to one view, cognitive evaluation must take place first, and the emotional response is generated subsequent to this.[20] An opposite view posits that emotions correspond to unconscious activity in the autonomic nervous system, and that our conscious feelings are interpretations of this activity.[21,22] According to Damasio,[23] emotions are induced in a relatively small number of brain sites, most of which are below the cerebral cortex. These comprise various components of the limbic system that have long been implicated in emotion,[24] including the amygdala (particularly the central nucleus), the hypothalamus, and the basal forebrain. In this view, emotion evolved through mechanisms of evolutionary adaptation, and performs an essential role in life regulation.[25] An animal needs to know when to approach a stimulus or situation and when to avoid a stimulus or situation, and it is through innate emotional biases and learned emotional associations that the animal does this. Thus, according to Damasio,[26] all experiences are emotionally tagged and influence emotional reactions in future situations, even though the person or animal may no longer have conscious access to the original experience. For example, an early unpleasant experience with a dog may induce a fear response to dogs, even though the person may not remember the origin of the association.

Although it has been argued that music may be an evolutionary adaptation serving social group cohesion (e.g. Chapter 5, this volume), it has also been argued that music does not have any survival function at all.[27,28] However, if music engages both phylogenetically old and newer emotional systems, this would suggest that music evolved alongside emotion in humans. In this chapter, we will first explore whether music engages the autonomic nervous system, subcortical emotion networks, and cortical areas involved in the emotional processing of other types of stimuli. Second, we will consider whether emotional reactions to music are simply cultural conventions by asking whether and how infants process musical emotions.

## Physiological responses to music

The subcortical emotion-processing parts of the brain affect the rest of the body through two basic mechanisms: the release of chemical molecules into the blood that act on various parts of the body; and the spread of neural activation to various brain centres and muscles. Through these mechanisms, the experience of an emotion is connected with a myriad of physiological responses, from muscle contractions, to changes in breathing and heart rate, to changes in blood flow in various parts of the body, to sweating. If music is simply about emotion, but does not induce emotion, one would not expect listening to music to activate the autonomic system, and physiological changes should not be evident. However, studies using both self-report[29] and direct measures of autonomic function[30,31] have now shown that listening to music does indeed produce autonomic changes associated with emotional processing. For example, adults report shivers down the spine, laughter, tears, and 'lump in the throat' as some physiological responses to music.[29]

Nyklicek *et al.*[31] presented listeners with musical excerpts expressing happiness, sadness, serenity, or agitation. These four emotions cover the extremes of an intensity/valence matrix: happiness and agitation are intense, whereas sadness and serenity are not; happiness and serenity are positively valenced, whereas sadness and agitation are negatively valenced. They found that changes in respiration clearly followed the intensity dimension. Respiration rate was higher, and inspiration and expiration times shorter, for the happy and agitated excerpts than for the sad and serene excerpts. This is consistent with the findings of Krumhansl,[30] who found respiration rate to be higher during fear and happy excerpts than during sad excerpts. The valence dimension, however, is less clearly seen in the autonomic measures. Using discriminant analysis, Nyklicek *et al.*[31] found that an arousal dimension accounted for 63 percent of the variance across a number of physiological variables related to respiration and cardiac function, whereas a dimension related to valence accounted for only 10 percent of the variance. Krumhansl[30] did find that respiration depth decreased more during happy excerpts than sad or fear excerpts, and that finger temperature decreased less for happy excerpts than for sad or fear excerpts. However, as a low-intensity, positive-valence emotion was not present in this study, it is difficult to determine definitively whether these measures reflect valence *per se*. It is clear, then, that music induces physiological changes consistent with the processing of emotional intensity, and perhaps also with emotional valence.

One approach to determining the relation between emotions induced by music and those induced through other means is to consider whether excerpts of music conveying different emotions produce distinctive autonomic signatures, and whether these signatures are consistent with emotions induced in other ways. However, there are inconsistencies in the autonomic signatures of specific emotions across studies using varying stimuli such as static visual stimuli, films, and directed facial action tasks. Thus, the idea that specific emotions are uniquely represented by specific autonomic patterns is still controversial.[22] Distinctive autonomic patterns for specific emotions are found when the directed facial action task is used.[32,33] However, there is little correspondence between these autonomic signatures and those found by Krumhansl[30] and Nyklicek *et al.*[31] For example, heart rate tends to be higher for sad than for happy facial actions and finger temperature does not differ, whereas heart rate tends to be slower for musically-induced sadness than for musically induced happiness, and finger temperature is higher for happy than for sad musical emotions. Facial action studies have been criticized as perhaps reflecting the difficulty of producing the facial expression rather than the emotion expressed.[34] In fact, many factors may influence the specific autonomic responses observed. Emotion is complex (e.g. there are many different kinds of happiness and fear) and differential effects may arise when overt responses are made or blocked.[35] Interestingly, musically induced autonomic responses are most consistent with those measured in studies employing manipulations that extend over time, such as watching a film or listening to a radio play.[30]

One difference between musical emotion and emotion arising in other domains is that overt action is not normally required in response to musical emotion. With development, people learn to control the overt expression of emotion to some extent. They may be able to stop themselves from uttering hurtful or angry words, but controlling their autonomic responses is more difficult. In the case of musical performers, however, feeling the music

too much, and allowing autonomic responses free rein, could lead to an inability to perform. A singer with a tight throat will not have optimal sound control, and a violin player with sweaty, shaking fingers will not be able to play well. Damasio[23] gives a wonderful example of a pianist, Maria Joao, who claimed that she could cut off the flow of emotion to her body at will. Damasio's skepticism turned out to be unfounded, as laboratory tests confirmed that she could allow or block physiological responses during music listening. This leads to the interesting question of whether musical emotions are somehow more subject to cognitive control than other emotions, or whether all emotional autonomic responses can be controlled through learning.

In summary, music does induce emotion directly. However, more research is needed before we can answer the question of whether musical emotions have the same autonomic signatures as emotions induced in other situations. Similarly, we can only speculate at present as to whether musical emotions are more under cognitive control than emotions induced through other means.

## Central nervous system responses to emotion in music

Music stimulates wide networks of bilateral activity across the brain, and specific areas within these networks appear to be specialized for the perception of various aspects of music such as melody, rhythm, and timbre (see Chapter 17, this volume). In particular, auditory cortex and frontal regions appear to be essential for musical processing.[36,37] Changes in pitch intervals and pitch contour are processed in auditory cortex even in the absence of attention,[38] and right frontal areas are activated during tasks involving pitch memory.[37] While right-hemisphere dominance is often found for pitch-based musical tasks (e.g. Chapter 16, this volume), laterality can be moved around by instruction (analytic vs holistic) and degree of musical training.[39,40] In contrast to the *perception* of music, much less is known about the networks for processing *emotional* aspects of music.

It is generally agreed that emotional processing of all types involves wide networks of central nervous system activity, including limbic and sensory areas as well as those related to cognition and consciousness.[41] However, consensus on the specific brain regions that are activated for each specific emotion is far from achieved.[22] This is likely due in part to the employment of different methodologies (e.g. PET, fMRI, EEG) across studies. However, it is also likely due to large effects on neural activity of the *contents* of emotional experience (e.g. which sensory systems are involved; the specific stimulus triggering the feelings), the salience of an emotion or extent to which the emotion is felt, different shadings of emotion within one category such as 'happiness', and, very importantly, differences between the induction of emotion, the feeling of emotion, and the conscious memory or discrimination of emotions. For example, the amygdala is strongly implicated in the induction and learning of fear responses,[41] but does not appear to be involved when fear states are recalled.[25]

Does emotion induced by music activate the same cortical structures as emotion induced by other stimuli? How is the processing of musical emotion affected by the contents of the stimulus? Given that research into the neural processing of emotion in music is just beginning, and given the lack of consensus on a model of emotional processing in

general, definitive answers to these questions cannot be given. However, a few very interesting studies are providing a starting point. We will review here three approaches that have been taken. First, lesion cases involving impairment of musical ability can give a hint as to the brain regions involved. Second, PET imaging studies in normal adults who listen to music varying in its emotional content can also suggest the brain regions involved. Third, EEG studies examining general patterns of regional activation can be compared across conditions of musical and nonmusical induction of emotion.

Peretz *et al.*[42] presented a case study of a patient, IR, who showed a dissociation between perceptual and emotional processing of music, a dissociation that appears to parallel that found between the identification of faces and the processing of emotional expression in faces.[43] IR suffered bilateral damage to her auditory cortices, such that in the left hemisphere Heschl's gyrus and the anterior portion of the planum temporale are completely destroyed, the superior temporal gyrus is infracted, and the damage extends to adjoining regions. On the right, Heschl's gyrus and the planum temporale are spared, but the anterior and superior portions of the superior temporal gyrus are damaged, and there is a large frontal lesion that includes the precentral and inferior frontal gyri and part of the orbitofrontal gyri. Despite the left hemisphere damage, IR's language is normal. However, her musical perception abilities are severely damaged, and she is unable to recognize or discriminate melodies. Despite these perceptual deficits, IR is able to tell whether musical excerpts are happy or sad.[44] Furthermore, she can do this based only on the mode and tempo of the excerpts. Thus, at least in this one case, not only are music and language dissociated, but perceptual and emotional aspects of music are as well. Further research is needed, however, to determine whether IR uses modality-general emotional processing areas for determining musical emotion.

As an initial step into examining musical emotions, Blood *et al.*[45] used PET imaging to identify regions whose activity correlated with changes in musical stimuli along the consonance/dissonance dimension. This dimension is defined such that tones sounding pleasant or smooth together are said to be consonant whereas those sounding unpleasant or rough together are said to be dissonant.[46] Dissonance is interesting as it relies on the peripheral structure of the auditory system: the critical band structure of the basilar membrane[46,47] and the firing patterns in the auditory nerve (see Chapter 9, this volume). In brief, tones with harmonics that are close enough in frequency to be difficult to resolve along the tonotopic organization of the basilar membrane are perceived to be dissonant. Likely because of its peripheral origins, infants as young as two months of age are sensitive to this dimension, preferring to listen to simultaneous tones in consonant relations over those in dissonant relations.[48,49] This dimension is also critical to musical structure, as virtually all musical systems rely on dissonance to generate tension, and consonance to resolve that tension.

Blood *et al.*[45] found that some regions involved in affective processing from previous nonmusic studies were also activated with changes in consonance and dissonance, including right parahippocampal gyrus, right precuneus, bilateral orbitofrontal, medial subcallosal singulate, and right frontal polar regions. Furthermore, the degree of activation of certain areas differed across consonance and dissonance presentation, suggesting that music conveying positive and negative valence may have distinct cortical signatures. While interesting and suggestive, a caveat is necessary before making strong conclusions about

musical emotional processing on the basis of this study. Although pleasant/unpleasant ratings correlated with activity in right parahippocampal gyrus, these dimensions did not correlate with ratings of happy/sad. Thus, while this study appears to capture something about the processing of musical valence, simply varying the amount of dissonance across a musical phrase is not the equivalent to inducing happiness or sadness. Dissonance without resolution likely corresponds more to irritation, which is a high-arousal negative emotion, rather than sadness, which is a low-arousal negative emotion. Furthermore, there is evidence for the cortical separation of consonance and dissonance from happiness and sadness. IR, described above, is unable to discriminate consonance and dissonance although she can discriminate happy from sad music.[50] In any case, this study suggests that musically induced emotions are processed in brain regions overlapping those involved in general emotional processing.

In a second study, Blood and Zatorre[51] explored intense positive emotional reactions to music by having each subject choose an except that gave them 'chills', and listen to it while PET responses were recorded. Compared to control conditions, the music induced increased cerebral blood flow in the left ventral striatum, dorsomedial midbrain areas, and paralimbic regions, areas that have been associated with euphoria, pleasant emotion, and cocaine administration in other studies. In addition, music listening was associated with decreases in blood flow to the amygdala, hippocampus, and ventral medial prefrontal cortex, which is also consistent with the experience of intense emotions in other contexts. In sum, these PET studies suggest that the experience of positive emotional responses to music activate the same brain circuits as positive emotions induced in other contexts.

There are two related lines of EEG research on emotion that may shed light on how musical emotion is processed in the brain. One line concerns the pattern of EEG activity that is observed in anterior regions during the processing of emotion. For example, a number of EEG studies have found evidence that frontal activation patterns differ across emotions.[16,17,52,53] A second line of research concerns the pattern of resting frontal EEG activity recorded during baseline or neutral states and its relation to individual differences in affective style or personality. In the personality literature, current thinking is that different personality styles emerge from how emotions are regulated.[16,17,52] For example, depression is thought to be characterized and maintained by an inability to experience positive affect; shyness and anxiety are thought to be characterized by an inability to regulate fear; sociability and extraversion are thought to be characterized and maintained by the experience of positive affect and the ability to control negative affect.

Interestingly, the pattern of resting frontal EEG activity has been shown to discriminate these various profiles, suggesting that the resting frontal EEG metric is indexing the predisposition to experience different affect states. This in turn suggests that the neural correlates of transient emotion may be similar to those shown for different personality types. One model has associated the intensity of emotion with overall amount of frontal activation.[54–56] For example, infants separated from their mothers show an increase in frontal activation.[57] A second model posits that emotions with positive valence, such as joy, interest, and happiness, show greater left than right activation whereas those with negative valence, such as fear, disgust, and sadness show greater right than left activation.[16,17,52,54,58] (It should also be noted that Blood *et al.*[45] also found that the positively valenced consonance excerpts

tended to activate left frontal areas more than the negatively valenced dissonant excerpts.) Evidence for the frontal activation/emotion valence model has been found across a number of stimulus modalities including pictures, films, and taste[16,17,52] and in infants as well as in adults.[52] Furthermore, baseline or neutral EEG asymmetries have been linked to personality types, with outgoing individuals tending to exhibit greater relative left activation and shy individuals greater relative right activation.

Because the frontal activation/emotion valence model is based on the idea of a dichotomy between approach and withdraw behaviours with a deep evolutionary history, it is not clear a priori whether it would apply to music. However, we[56] found clear evidence that it holds for music as well. We presented listeners with music expressing emotions at the extremes of the intensity/valence matrix. Specifically, we used orchestral excerpts expressing joy (high intensity, positive valence), happiness (low intensity, positive valence), fear (high intensity, negative valence), and sadness (low intensity, negative valence) that had been rated by adults along the valence and intensity dimensions. Evidence was found supporting both the intensity and the frontal activation/emotion valence models. The measure of activation was the inverse of alpha band activity (8–13 Hz), as energy in this band decreases with increased activation.[58] The amount of frontal activation was correlated with adults' ratings of the intensity of their emotional response to the excerpts (from least to most intense: sad, happy, joy, fear), supporting the intensity model (Figure 20.1). Both joy and happiness showed greater relative left frontal activation whereas both fear and sadness showed greater relative right activation (Figure 20.1), supporting the frontal activation/emotion valence model. Thus, we[56] found evidence that emotion induced by music activates frontal circuits in the brain similar to those activated by other emotional stimuli.

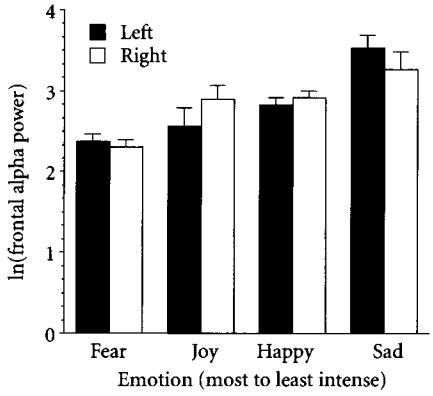



**Figure 20.1** Differences among fear, joy, happy, and sad left and right frontal EEG alpha power. Note that EEG power is thought to be inversely related to activity, thus lower power reflects more activity. Overall activation is related to intensity, and positive emotions show greater relative left activation whereas negative emotions show greater relative right activation. Error bars represent the standard error of the mean. Reprinted with permission from Schmidt and Trainor.[56]

In summary, although there are rather few studies to date addressing how musical emotion is processed in the brain, the preliminary evidence suggests that, despite the fact that music has questionable evolutionary advantages, and despite the fact that music normally does not motivate approach or withdraw behaviours, music does appear to activate the same emotional circuits as other stimuli.

## Development of emotional responses to music

The strongest argument that music is an important evolutionary adaptation arguably comes from the developmental perspective. Caregivers around the world sing to infants, and young infants are responsive to such music.[59] Singing directed at infants is rendered in a style that is distinct from other types of singing,[60,61] and infants prefer to listen to infant-directed over non-infant-directed renditions of the same song. The function of infant-directed singing remains somewhat elusive. However, one of the main theories is that caregivers use infant-directed singing to express emotional information, and to regulate their infant's state.[59–61] Young infants are, of course, not good at state regulation, and require intervention in order to calm down when upset. Mothers sing in two distinct styles, a lullaby style and a playsong style.[61–63] Adults can discriminate playsong and lullaby styles easily, and they rate playsongs as more rhythmic, brilliant, clipped, and smiling in character, and less soothing and airy than lullabies.[63] Furthermore, infants react differently when exposed to lullabies and playsongs, focussing their attention inward during the former and outward during the latter.[63] Thus, music has the power to affect an infant's state.

On the basis of the evidence that music is universally used in caretaking contexts for emotional expression and state regulation, and that infants react differently to different musical styles, it is possible that singing to infants serves an important adaptive function in development. Specifically, music may provide one route into learning about social interaction and self-regulation before infants understand any language.

Children are also able to distinguish different emotions in music. Cunningham and Sterling[64] showed that 4- to 6-year-olds could discriminate happy, sad, angry, and fearful musical excerpts. Trainor and Trehub[4] found that children as young as 4 years could reliably associate excerpts from Prokofiev's Peter and the Wolf and Saint Saens' Carnival of the Animals with pictures of the animals, giving emotion-laden justifications for their responses such as that the wolf excerpt sounded 'scary'.

The bases for early emotional reactions to music have not been investigated widely, but they likely involve interpretations based on pitch, tempo, and timbre characteristics of the music. One aspect of pitch structure has been investigated. As discussed previously, consonant intervals are associated with positive emotion and dissonant intervals with negative emotion. Using a visual looking paradigm in which infants control how long they listen to consonant vs dissonant chords by how long they fixate on a visual target, Trainor and colleagues[48,49] have demonstrated that infants as young as two months prefer to listen to consonant over dissonant musical intervals. These results complement previous findings that older infants also prefer consonance over dissonance.[65–67]

Given the behavioural evidence that young infants respond to the emotional content of music, the same question that we addressed previously with adults arises with infants: Does

musical emotion stimulate the same brain circuits as other emotional experiences in infants? A number of EEG studies of nonmusical emotional processing in infants support the presence of frontal asymmetries related to valence.[52] For example, infants tend to show greater relative right frontal EEG activation during the processing of negatively valenced stimuli (i.e. faces displaying distress or fear, tastes that are sour) and greater relative left frontal EEG activation during the processing of positively valenced stimuli (i.e. faces displaying joy, smiling, and sweet tastes). As well, infants who show greater right than left resting frontal EEG activation tend to show heightened distress to novel events and maternal separation.

In order to test whether infants' EEG responses would show the same asymmetry effects for emotion induced by music, we[68] recently recorded EEG and ECG as 3-, 6-, 9-, and 12-month old infants listened to the joy, fear, and sad orchestral excerpts used in Ref. 56. Infant alpha band activity (4–8 Hz) was recorded, and activation was taken as the inverse of this measure. There were interesting changes across age (Figure 20.2). Compared to baseline, the presence of music significantly increased brain activity at 3 months of age, had little effect at 6 and 9 months of age, and significantly attenuated brain activity at 12 months of age (Figure 20.2),

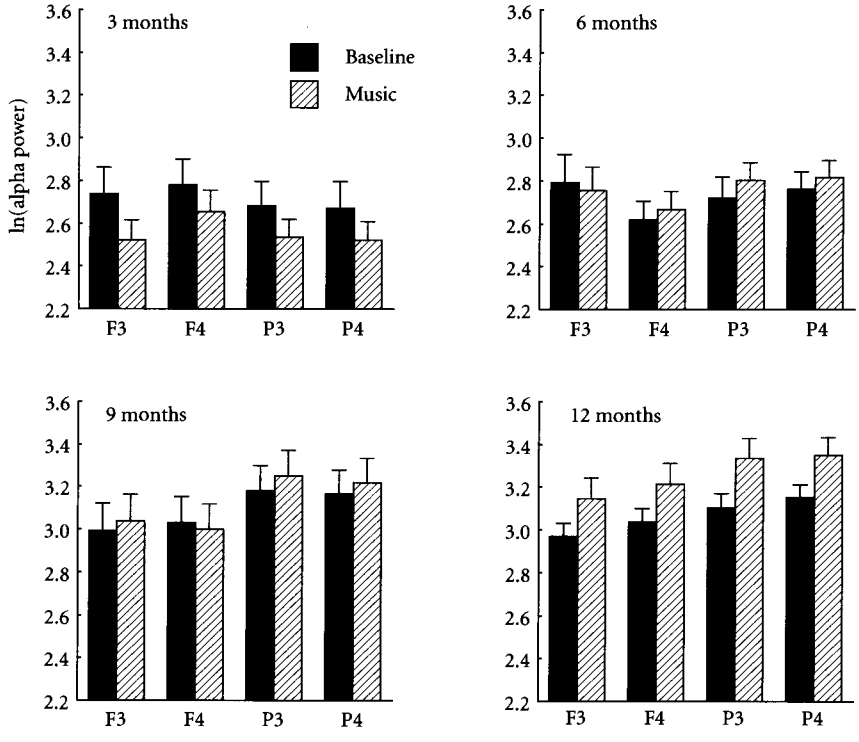



Figure 20.2 Infant alpha power for frontal (F3, F4) and parietal (P3, P4) regions and left (F3, P3) and right (F4, P4) hemispheres. Note that music has opposite effects compared to baseline at 3 and 12 months of age. Error bars represent the standard error of the mean. Reprinted with permission from Schmidt, Trainor, and Santesso.[68]

suggesting that music is having a calming effect by 12 months as infants are increasingly able to regulate their sensory input. However, no left/right asymmetries were found in any condition.

There are a number of possible reasons for the lack of asymmetric responses to musical emotion in infants. It is possible that musical emotion is processed differently than other emotions in infancy, but this would be surprising given the behavioural evidence for a central role of musical emotion in everyday infant caregiving activities. Another possibility is that the orchestral musical excerpts used were simply too complex for infants. While it is advantageous to use the same musical excerpts with infants and adults, this music is certainly not very similar to typical infant-directed singing. Still another explanation is that sufficient frontal lobe maturation for the cognitive appraisal of musical stimuli is not yet developed in the first year of life.

To address these possibilities, studies are underway to measure infant EEG with more ecologically valid stimuli. In particular, we are using vocal singing, infant-directed speech, and infant-directed singing conveying various emotions. At this point it is not possible to answer the questions of whether infants process musical emotions with similar brain circuits to those used for other emotions, or whether infants and adults use similar brain circuits to process musical emotion. However, methodologies now exist to ask these questions, and hopefully more data will be forthcoming.

## Conclusions

The popular notion that music elicits powerful emotions appears to be close to the truth. That music is not only about emotion, but that it elicits emotion directly, can be seen clearly from the physiological responses it induces. The changes in heart rate, respiration rate, blood flow, and skin conductance are clear indications that music activates the phylogenetically old parts of the nervous system, and that music elicits a cascade of subconscious activity. Music also appears to activate the cortical systems associated with emotion, including circuits in the frontal lobes. Further research is needed, however, to address the questions of how discrete emotions are processed, and how various dimensions of emotion such as valence and intensity are encoded in the brain. There are many unanswered questions with respect to music, such as why happiness and sadness are so easily expressed in musical structure whereas emotions such as anger are so much more difficult to express.

Despite the fact that, for adults in the modern world, music does not command the same approach/withdrawal reactions as other emotion-laden stimuli, much evidence suggests that music does activate the same cortical, subcortical, and autonomic circuits as other emotions. Perhaps the answer to the puzzle of why music appears to activate the essential survival circuits of the nervous system—when music does not appear to serve any obvious survival function—lies in the evolution of development and child care. Human infants are particularly helpless for an extended period of time, and are reliant on their caregivers for survival. An emotional bond and the communication of positive and prohibitionary emotional information is essential for survival. Perhaps music evolved in order to further emotional communication between infants and caregivers. And perhaps infant-directed singing and infant-directed 'musical' speech[69] are intimately connected with approach and

withdrawal behaviour on the part of the infant. Taking this one step further, as children develop other means of communication such as language, and as they learn to keep overt expressions of emotion under control, music may go 'underground'. Music still elicits emotion, but more direct and precise methods of emotional communication can occur with language. Perhaps, however, music retains a survival role in adults in that it allows the 'practice' of feeling emotions without having to risk the consequences of acting on these emotions. Thus, we are free to laugh and cry to music, to feel our heart race, to feel chills, and to hold our breath, and perhaps once again to feel like the infant who is exquisitely tuned to the emotion in its mother's voice.

## Acknowledgements

This work was supported by grants from the Canadian Institutes of Health Research and the National Science and Engineering Research Council of Canada. We thank Joanne Leuzzi for technical support, and Terri Lewis and Judy Plantinga for comments on an earlier draft.

## References

1. **Juslin, P. N.** and **J. A. Sloboda** (eds) (2001) *Music and Emotion: Theory and Research.* Oxford: Oxford University Press.
2. **Meyer, L. B.** (1956) *Emotion and Meaning in Music.* Chicago: University of Chicago Press.
3. **Terwogt, M. M.** and **F. Van Grinsven** (1991) Musical expression of moodstates. *Psychol. Music* 19, 99–109.
4. **Trainor, L. J.** and **S. E. Trehub** (1992) The development of referential meaning in music. *Music Percept.* 9, 455–70.
5. **Gabrielsson, A.** and **E. Lindstrom** (2001) The influence of musical structure on emotional expression. In P. N. Juslin and J. A. Sloboda (eds) *Music and Emotion.* Oxford: Oxford University Press.
6. **Hevner, K.** (1935) The affective character of the major and minor modes in music. *Am. J. Psychol.* 47, 103–18.
7. **Hevner, K.** (1936) Experimental studies of the elements of expression in music. *Am. J. Psychol.* 48, 246–68.
8. **Balkwill, L. L.** and **W. F. Thompson** (1999) A cross-cultural investigation of the perception of emotion in music: psychophysical and cultural cues. *Music Percept.* 17, 43–64.
9. **Juslin, P. N.** (2001) Communicating emotion in music performance: a review and a theoretical framework. In P. N. Juslin and J. A. Sloboda (eds) *Music and Emotion.* Oxford: Oxford University Press, pp. 309–37.
10. **Gabrielsson, A.** and **P. N. Juslin** (1996) Emotional expression in music performance: between the performer's intention and the listener's experience. *Psychol. Music* 24, 68–91.
11. **Hanslick, E.** (1957) (originally published in 1854). *The Beautiful in Music.* New York: Liberal Arts Press.
12. **Langer, S.** (1942) *Philosophy in a New Key: A Study in the Symbolism of Reason, Rite, and Art.* Cambridge, MA: Harvard University Press.

13. Watson, D. and A. Tellegen (1985) Toward a consensual structure of mood. *Psychonomic Bulletin* 98, 219–35.
14. Davidson, R. J., K. R. Scherer, and H. H. Goldsmith (2003) *Handbook of Affective Sciences.* New York: Oxford Unversity Press.
15. Bryden, M. P. and R. G. Ley (1983) Right-hemispheric involvement in the perception and expression of emotion in normal humans. In K. M. Heilman and P. Satz (eds) *Neuropsychology of Human Emotion.* New York: Guilford Press.
16. Davidson, R. J. (2000) Affective style, psychopathology, and resilience: brain mechanisms and plasticity. *American Psychologist* 55, 1196–214.
17. Davidson, R. J. (1993) The neuropsychology of emotion and affective style. In M. Lewis and J. M. Haviland (eds) *Handbook of Emotion.* New York: The Guilford Press, pp. 143–54.
18. Fox, N. A. (1994) Dynamic cerebral processes underlying emotion regulation. *Monographs of the Society for Research in Child Development* 59, 152–66, 250–83.
19. Silberman, E. K. and H. Weingartner (1986) Hemispheric lateralization of functions related to emotion. *Brain and Cognition* 5, 322–53.
20. Lazarus, R. S. (1991) Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist* 46, 819–34.
21. James, W. (1950) (originally published in 1890). *The Principles of Psychology* vol. 2. New York: Dover Publications.
22. Zajonc, R. B. and D. N. McIntosh (1992) Emotions research: some promising questions and some questionable promises. *Psychol. Sci.* 3, 70–4.
23. Damasio, A. (1999) *The Feeling of What Happens: Body and Emotion in the Making of Consciousness.* San Diego, CA: Harcourt Inc.
24. Papez, J. W. (1937) A proposed mechanism of emotion. *Archives of Neurology and Psychiatry* 38, 725–43.
25. Damasio, A. R., T. J. Grabowski, A. Bechara, *et al.* (2000) Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nat. Neurosci.* 3, 1049–56.
26. Damasio, A. R. (1994) *Descartes' Error: Emotion, Reason, and the Human Brain.* New York: Putnam.
27. Pinker, S. (1997) *How the Mind Works.* New York: William Morrow.
28. Pinker, S. (1994) *The Language Instinct.* New York: William Morrow.
29. Sloboda, J. A. (1991) Music structure and emotional response: some empirical findings. *Psychol. Music* 19, 110–20.
30. Krumhansl, C. L. (1997) An exploratory study of musical emotion and psychophysiology. *Canadian J. Exp. Psychol.* 51, 336–52.
31. Nyklicek, I., J. F. Thayer and L. J. P. Van Doornen. (1997) Cardiorespiratory differentiation of musically-induced emotions. *J. Psychophysiol.* 11, 304–21.
32. Levenson, R. W., P. Ekman and W. V. Friesen (1990) Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiol.* 27, 363–84.
33. Levenson, R. W. (1992) Autonomic nervous system differences among emotions. *Psychol. Sci.* 3, 23–7.
34. Boiten, F. (1996) Autonomic response patterns during voluntary facial action. *Psychophysiology* 33, 123–31.
35. Gross, J. J. and R. W. Levenson (1993) Emotional suppression: physiology, self-report, and expressive behavior. *J. Pers. Soc. Psychol.* 64, 970–86.

36. Zatorre, R. J. and S. Samson (1991) Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain* 114, 2403–17.
37. Zatorre, R. J., A. C. Evans, and E. Meyer (1994) Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 14, 1908–19.
38. Trainor, L. J., K. L. McDonald, and C. Alain (2002) Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *J. Cog. Neurosci.* 14, 430–42.
39. Peretz, I. and J. Morais (1987) Analytic processing in the classification of melodies as same or different. *Neuropsychologia* 25, 645–52.
40. Peretz, I., J. Morais, and P. Bertelson (1987) Shifting ear differences in melody recognition through strategy inducement. *Brain and Cognition* 6, 202–15.
41. LeDoux, J. E. (1996) *The Emotional Brain: The Mysterious Underpinnings of Emotional Life.* New York: Simon and Schuster.
42. Peretz, I. and L. Gagnon (1999) Dissociation between recognition and emotional judgements for melodies. *Neurocase* 5, 21–30.
43. Peretz, I. (2001) Listen to the brain: a biological perspective on musical emotions. In P. N. Juslin and J. A. Sloboda (eds) *Music and Emotion.* Oxford: Oxford University Press.
44. Peretz, I., L. Gagnon, and B. Bouchard (1998) Music and emotion, perceptual determinants, immediacy, and isolation after brain damage. *Cognition* 68, 111–41.
45. Blood, A. J., R. J. Zatorre, P. Bermudez, and A. C. Evans (1999) Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat. Neurosci.* 2, 382–7.
46. Plomp, R. and W. J. Levelt (1965) Tonal consonance and critical bandwidth. *J. Acoust. Soc. Am.* 38, 548–60.
47. Schellenberg, E. G. and L. J. Trainor (1996) Sensory consonance and the perceptual similarity of complex-tone harmonic intervals: tests of adult and infant listeners. *J. Acoust. Soc. Am.* 100, 3321–8.
48. Trainor, L. J. and C. D. Tsang (2000) Paper presented at the International Conference on Infant Studies. England: Brighton.
49. Trainor, L. J., C. D. Tsang, and V. H. W. Cheung (2002) Preference for musical consonance in 2-month-old infants. *Music Perception*, 20, 185–92.
50. Peretz, I., A. J. Blood, V. Penhune, and R. Zatorre (2001) Cortical deafness to dissonance. *Brain* 124, 928–40.
51. Blood, A. J. and R. Zatorre (2001) Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc. Natl. Acad. Sci.* 98, 11818–23.
52. Fox, N. A. (1991) If it's not left, it's right: electroencephalogram asymmetry and the development of emotion. *American Psychologist* 46, 863–72.
53. Canli, T. (1999) Hemispheric asymmetry in the experience of emotion: A perspective from functional imaging. *The Neuroscientist* 5, 201–7.
54. Dawson, G. (1994) Frontal electroencephalographic correlates of individual differences in emotional expression in infants. *Monographs of the Society for Research in Child Development* 135–51.
55. Schmidt, L. A. and N. A. Fox (1999) Conceptual, biological, and behavioural distinctions among different types of shy children. In L. A. Schmidt and J. Schulkin (eds) *Extreme Fear, Shyness and Social Phobia: Origins, Biological Mechanisms, and Clinical Outcomes.* New York: Oxford University Press, pp. 27–46.

56. **Schmidt, L. A.** and **L. J. Trainor** (2001) Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions. *Cognit. Emotion* 15, 487–500.

57. **Dawson, G. W., L. Grofer-Klinger., H. Panagiotides, D. Hill,** and **S. Spieker** (1992) Frontal lobe activity and affective behavior of infants of mothers with depressive symptoms. *Child Dev.* 63, 725–37.

58. **Davidson, R. J.** and **A. J. Tomarken** (1989) Laterality and emotion: an electrophysiological approach. In F. Boller and J. Grafman (eds) *Handbook of Neuropsychology.* Amsterdam: Elsevier Science.

59. **Trehub, S. E.** and **L. J. Trainor** (1998) Singing to infants: lullabies and play songs. In C. Rovee-Collier, L. P. Lipsitt, and H. Hayne (eds) *Advances in Infancy Research.* Stamford, Connecticut: Ablex Publishing Corporation, pp. 43–77.

60. **Trehub, S. E., A. M. Unyk,** and **L. J. Trainor** (1993) Maternal singing in cross-cultural perspective. *Infant Behav. Dev.* 16, 285–95.

61. **Trainor, L. J.** (1996) Infant preferences for infant-directed versus noninfant-directed playsongs and lullabies. *Infant Behav. Dev.* 19, 83–92.

62. **Trainor, L. J., E. D. Clark, A. Huntley,** and **B. A. Adams.** (1997) The acoustic basis of preferences for infant-directed singing. *Infant Behav. Dev.* 20, 383–96.

63. **Rock, A. M. L., L. J. Trainor,** and **T. L. Addison** (1999) Distinctive messages in infant-directed lullabies and play songs. *Developmental Psychol.* 35, 527–34.

64. **Cunningham, J. G.** and **R. S. Sterling** (1988) Developmental change in the understanding of affective meaning in music. *Motivation and Emotion* 12, 399–413.

65. **Trainor, L. J.** and **B. J. Heinmiller** (1998) The development of evaluative responses to music: infants prefer to listen to consonance over dissonance. *Infant Behav. Dev.* 21, 77–88.

66. **Zentner, M. R.** and **J. Kagan** (1998) Infants' perception of consonance and dissonance in music. *Infant Behav. Dev.* 21, 483–92.

67. **Crowder, R. G., J. S. Reznick,** and **S. L. Rosenkrantz** (1991) Perception of the major/minor distinction: V. Preferences among infants. *Bull. Psychon. Soc.* 29, 187–8.

68. **Schmidt, L. A., L. J. Trainor,** and **D. L. Santesso** Development of electroencephalogram (EEG) and autonomic (ECG) responses to affective musical stimuli during the first 12 months of postnatal life. *Brain and Cognit.* (in press).

69. **Fernald, A.** (1992) Meaningful melodies in mothers' speech to infants. In H. **Papousek** and U. Jurgens (eds) *Nonverbal Vocal Communication.* Cambridge, UK: Cambridge University Press, pp. 262–82.

CHAPTER 21

# A NEW APPROACH TO THE COGNITIVE NEUROSCIENCE OF MELODY

ANIRUDDH D. PATEL

## Introduction

What is a musical melody? At the physical level, it is simply an organized succession of individual tones. At the perceptual level, however, it is much more. One perceptually oriented definition of melody has been suggested by the ethnomusicologist Simon Shaheen: a melody is 'a group of notes that are in love with each other'.[1] A related definition, which I hope to illustrate in this chapter, is that a musical melody is a tone sequence in which the individual tones are processed in terms of multiple structured relationships.[a] This definition emphasizes the active role of the mind in melody perception. That is, a melody (like a spoken sentence) depends on a listener's perceptual system to convert a 'mere sequence of sounds' into a meaningful mental experience.

This chapter has two goals. The first is to illustrate the cognitive richness of melody perception. The second is to introduce a new method for the neural study of melody. To illustrate melody perception I have chosen one particular melody from the Western European tradition, a children's song indexed as K0016 in the Essen database of Bohemian folk melodies.[2,3,b] Figure 21.1 shows the melody in Western music notation and in 'piano-roll' notation with each tone's pitch plotted as a function of time (this melody and all sound examples in this chapter can be heard at www.nsi.edu/users/patel/sound_examples/pz_chapter). The reasons for choosing this melody are threefold. First, it is historically recent yet likely to be unfamiliar to readers of this chapter: thus it follows familiar musical conventions while being free of specific memory associations. Second, as a folksong this melody has been through a period of preservation and transmission by ear rather than by music notation, so that it has survived a certain amount of natural testing for its perceptual coherence. Third, as a children's song the melody is short and illustrates basic structural relations in relatively simple form. Beyond these constraints, there is nothing special about K0016, and any number of melodies would have served the same purpose.

[a] Throughout this chapter the term 'melody' refers to monophonic melody, that is, melody without accompaniment.

[b] The database is available from the Center for Computer Assisted Research in the Humanities (Stanford University) and at www.esac-data.org (K0016 is in the kinder0 file of this database).

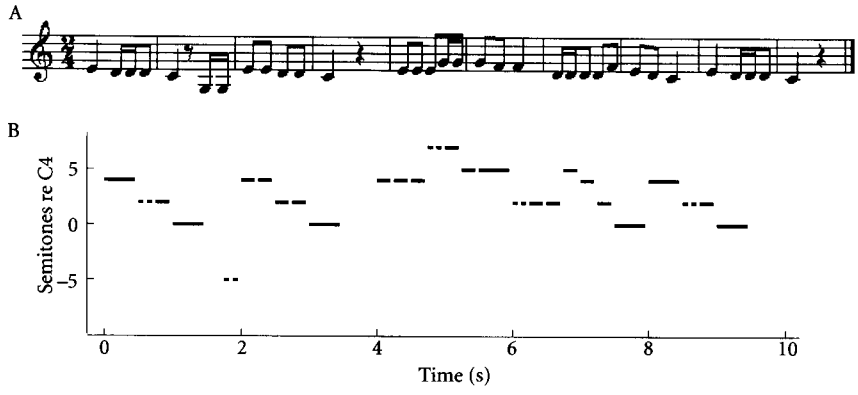



**Figure 21.1** A children's melody from the ESAC database (K0016). The melody has been transposed to the key of C. (A) Western music notation. (B) Piano roll notation, with each tone's fundamental frequency (in semitones from C4 or 261.63 Hz) plotted as a function of time. A semitone is a ratio of ~1.06 in fundamental frequency, and corresponds to the distance between a neighbouring white and black piano key. K0016 ranges over 12 semitones (one octave, a doubling in fundamental frequency), that is, −5 semitones below C4 to 7 semitones above it.

## Thirteen ways of hearing a melody

One way to appreciate the cognitive richness of melody perception is to enumerate the different things a listener hears during melodic processing. The following list is based on perception of a simple melody by a listener familiar with Western European tonal music. The list is not exhaustive, but merely touches on some of the more obvious features of melodic perception. Readers who wish to delve deeper into the structure of melody may consult Leonard Meyer's *Explaining Music*[4] and Eugene Narmour's two volumes on the analysis and cognition of melodies.[5,6]

### Instrument identity

An immediate and salient aspect of melody perception is recognition of the instrument producing the tones (e.g. piano vs violin). Instrument recognition takes place very rapidly (instruments can often be recognized from single tones), and relies heavily on timbre or 'sound quality', which is a complex feature relying on the spectral and temporal shape of sound.[7]

### Grouping

Grouping refers to the segmentation of a melody into units larger than single tones.[8] This can be illustrated with K0016, which consists of 33 separate tones. These tones define 33 minimal 'pitch-time units', each defined by a particular pitch and duration. In listening to this melody, there is a clear sense of larger pitch-time units between the levels of the individual pitches and the entire melody. For example, the melody is perceived as divided into phrases, schematically marked in Figure 21.2. The boundaries of the first two phrases are

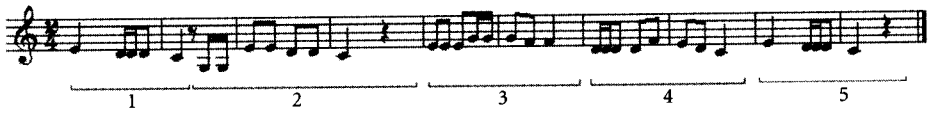



Figure 21.2 K0016 segmented into melodic phrases.

marked by silences (musical rests). Of greater interest are the boundaries at the end of the third and fourth phrases, which are not marked by any physical discontinuity in the tone sequence. Instead, these boundaries are marked by local durational lengthening and lowering of pitch, cues which also serve a boundary-marking role in speech.[9]

## Beat and metre

A beat refers to 'a psychological, more-or-less isochronous, pulse train—that provides the stimulus for synchronization'.[10] K0016 readily evokes a beat: if asked to tap the beat of this melody, listeners are likely to tap at the positions marked in Figure 21.3 by the dots at the level labelled '1x'.[c] Note that three of these taps occur during silent intervals, showing that the beat is not a simple marking of physically accented events. Indeed, in the sound examples all tones in K0016 have the same amplitude: there are no accents based on the intensity of tones. Instead, the beat is inferred from the pattern of durations and pitches.[11]

When beats are grouped into larger repeating units the result is metre, that is, periodicity at multiple time scales.[12] In K0016 the beats are organized into higher-level groups of two (as indicated by the time signature) and four (examination of Figures 21.2 and 21.3 shows that each phrase has four beats).[d] Melodic sequences frequently exhibit these sorts of multiple periodicities, which can be demonstrated by asking listeners to tap at different rates on successive listenings to a piece. Musically experienced listeners can tap at different rates related by simple integer ratios (e.g. 1/2 or 2× the basic beat), suggesting that the rate at which one spontaneously taps the beat is simply one level (the 'tactus') in a hierarchy of beats, as shown in Figure 21.3.[8,13]

## Scale structure

The numerical ratios between the fundamental frequencies of successive musical tones define the set of pitch intervals used in a melody. The intervals used in any given melody typically adhere to culture-specific schemata, for example, the scales of Western tonal music.[14,15] The intervals of K0016 are drawn from the major (Ionian) scale, characterized by a particular pattern of frequency steps between adjacent tones within an octave (in semitones, step sizes are: 2 2 1 2 2 2 1; a semitone is a ratio of approximately 1.06 in fundamental frequency). Listeners are sensitive to the structure of pitch intervals relative to the schemata of one's

[c] The tempo of K0016 has been set to 120 beats per minute (bpm) in the sound examples, that is, one beat every 500 ms. Any tempo between between ~75 and 150 bpm would likely be acceptable.

[d] This helps explain why the music notation for this melody includes a rest in the final position, rather than simply ending with the final note: this rest occurs on a beat. The psychological reality of this rest can be demonstrated by memorizing K0016 and singing it twice in immediate succession. There will be a natural tendency to pause after the final sung note before beginning again.

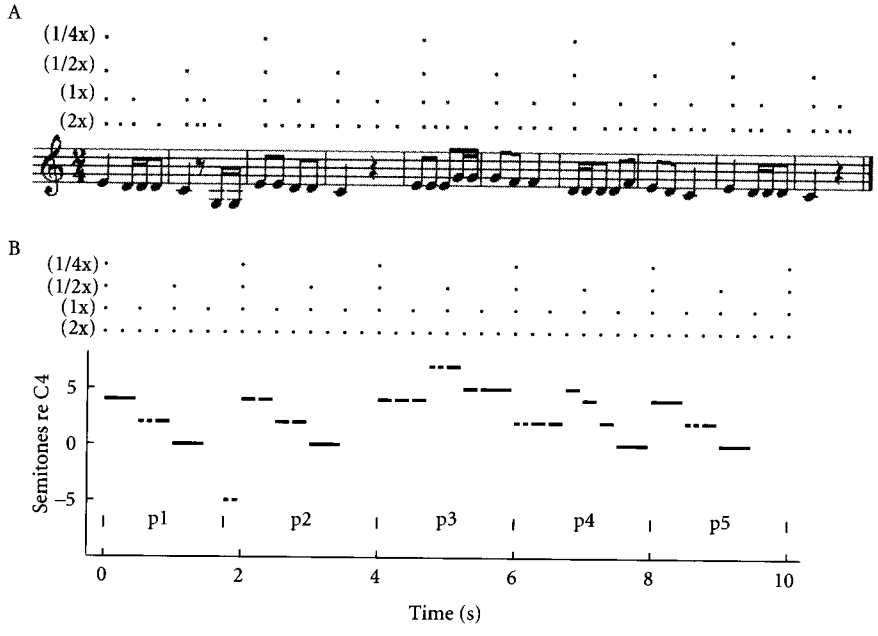



**Figure 21.3** Metrical structure of K0016, indicated as a hierarchy of beats above (A) Western music notation. (B) Piano roll notation. To help orient the reader, phrase boundaries are marked below the pitch sequence in (B), and phrases are labelled as p1, p2, etc. (cf. Figure 21.2).

culture.[14] Violations of these schemata are quite salient (sound example 2). In fact, insensitivity to such violations is diagnostic of true 'tune deafness' (congenital amusia).[16]

### Contour

The overall pattern of ups and downs created by a sequence of pitch intervals, regardless of precise interval size, defines a pitch contour. A pitch contour and its temporal pattern together define a melodic contour,[15] a perceptually salient aspect of melodies, which plays an important role in memory for unfamiliar melodies.[17] Sensitivity to melodic contour emerges early in infancy[18] and is likely to be related to the importance of intonation in speech perception.[19]

### Parallelism

Parallelism refers to motivic/thematic similarity between different parts of a melody. For example, the first and last phrase of K0016 are identical, which helps to provide this melody with a sense of closure. Parallelism need not be so literal: for example, it can be based on similar melodic contours. The study of parallelism is based on the measurement of structural and perceptual similarity. Since similarity is a matter of degree, and is influenced by many factors, the study of parallelism has lagged behind other aspects of music which have more discrete and easily measurable characteristics (e.g. grouping, metre, and interval/tonal

relations (see below)). Thus despite its fundamental role in melodic perception, parallelism and its perception is only beginning to be investigated in a quantitative framework.[20]

### Intervallic implications

Listeners are sensitive to the patterning of successive pitch intervals in a melody. One source of evidence for this is that if a melody is stopped midstream, listeners typically have well-defined expectations of how it will continue.[21] Interval-based expectancies are the focus of Narmour's 'implication-realization' (IR) model.[5,6] A flavour for the model's principles can be gleaned from Schellenberg,[22] who investigated melodic expectancy in British and Chinese melodies using American and Chinese listeners. Two principles emerged as important factors cross-culturally: 'pitch proximity' (an overall expectancy for small intervals) and 'pitch reversal' (an expectation that after a large interval, the pitch direction will change). The source of these bottom-up principles may be auditory Gestalt principles[23] or experience with the intonation contours of speech.

### Tension vs resolution

A remarkable aspect of melody perception is an unfolding sense of tension vs resolution over the course of the entire melody. That is, at each point a listener has some sense of whether the melody must continue or could naturally come to a halt. Bigand[24] has empirically measured this aspect of melody by having participants listen to increasingly long fragments of a melody and judge the stability of each fragment's end.[e] His results confirm that perceived tension varies widely during the course of a single melody. This is easily illustrated with K0016: consider two fragments which stop at the ends of phrase 3 and 4, respectively (sound examples 3 and 4). These differ dramatically in the degree of tension experienced at the final tone, even though both fragments stop at the ends of phrases.

Two factors that contribute to the degree of perceived tension are 'tonal weight' and tone duration.[24] The latter is simple (longer tone = less tension), but the former requires a brief explanation. In 'tonal' music of the Western European tradition different pitches take on distinct structural roles in the fabric of the music. Consider Figure 21.4, in which each note of K0016 has been marked with a number representing each tone's position (or 'scale degree') in the scale from which the melody is built (e.g. tone 1 is do, 2 is re, 3 is mi, etc.). Music perception research shows that once a melody is perceived as being in a particular key, certain scale degrees act as perceptual reference points,[25] especially scale degree 1 (also known as the 'tonic'). K0016, for example, is in the key of C, and the central role of the C note (scale degree 1) is marked in several ways. It is the final tone of phrases 1, 2, 4, and 5, and thus serves as a resting point within the melody and at the melody's end. Furthermore, it always occurs on a beat and is always long in duration. Thus scale degree 1 acts as a perceptual 'centre of gravity' for the entire melody, a role which can be contrasted with that of another tone in the melody (D), marked by the numeral 2 in Figure 21.4. Tone 2 is far more

[e] Bigand[24] (p. 811) defined high stability at the end of a fragment as 'the feeling that the melody could naturally stop at this point... [while] low stability [is] the feeling that there must be a continuation of the melody'. Listeners rated stability on a scale of 1–7. My use of 'tension' is equivalent to Bigand's 'low stability'. Bigand[24] reported stability profiles, which can easily be converted to tension profiles by taking their inverse.

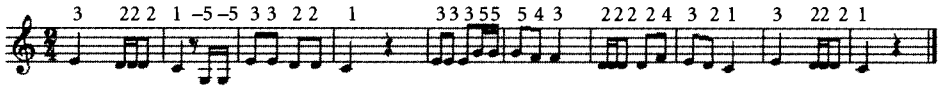



**Figure 21.4** Scale degree of each tone in K0016. 1 = do, 2 = re, 3 = mi, etc. (−5 = so in the octave below).

common than tone 1 in the melody (13 vs 4 occurrences), yet tone 2 never serves as a resting point: instead, it almost always leads back to tone 1. Furthermore, it is never long in duration, and frequently occurs off the beat. Thus scale degrees 1 and 2, though neighbours in frequency, play very different structural roles in the melody.

Bigand's 'tonal weight' is simply the degree of structural stability or centrality of each tone in a key, as quantified in perception experiments by Krumhansl.[25] Krumhansl has demonstrated an empirical hierarchy of stability for the notes in a particular key, with scale degree 1 being the most stable, followed by scale degrees 5 and 3 (G and E in the key of C) while degrees 2, 4, 6, and 7 have lower stability (D, F, A, & B in the key of C).

### Ornamentation

The concept that some pitches serve to elaborate or ornament others is central to Western European music theory (e.g. the theories of Schenker,[26] Meyer,[4] and Lerdahl & Jackendoff;[8] cf. Cook[27]), and is also found in the music theory of many non-Western cultures, including China and India. For example, in one Chinese folk tradition musicians speak of ornamenting a melodic skeleton in terms of 'adding flowers' to a melody.[28] The ability to distinguish structural from ornamental pitches is thought to play an important role in melody perception, for example, in a listener's ability to recognize one passage as an elaborated version of another.[8] An important issue for the study of melodic ornamentation is understanding the basis upon which pitches are perceived as structural vs ornamental. One relevant study in this regard is that of Bharucha,[29] who used a memory experiment to show that the salience of a tonally unstable note was influenced by its serial position relative to the following note. Specifically, Bharucha demonstrated that an unstable note which is immediately followed by a tonally stable pitch neighbour (e.g. B–C in a C major context) is less prominent/detectable than an unstable note which is not 'anchored' in this way. It is as if the stable tone subordinates the preceding tone as a local ornament and makes it less conspicuous than if the same preceding tone were inserted randomly into the sequence. This suggests that the tonal hierarchy is involved in perceived elaboration relations in music.

### Implicit harmony

While melodies present tones one at a time (i.e. 'horizontally'), chords present tones in combination (i.e. 'vertically'). The organization of chords in musical sequences is the domain of harmony. As with the tones of a melody, different chords play distinct structural roles in the fabric of the music, reflecting some of the same relationships which applied to the roles of individual tones. For example, different chords are associated with different degrees of stability, with one chord acting as a point of greatest structural stability: the triadic chord built on the first degree of the scale. This 'tonic chord', together with the chords built on the fourth and fifth scale degrees (the 'subdominant' and 'dominant' chords,

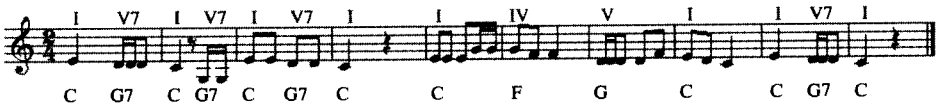



Figure 21.5 Harmonization of K0016. The letters below the staff give the chords in the key of C major, while the Roman numerals above the staff indicate chord functions (e.g. I = tonic, IV = subdominant, V = dominant).

respectively), form a structural core of chords in any key.[25] Although unaccompanied melodies do not have explicit harmony, they can have a strong implicit harmony based on the principle of 'broken chords', whereby chord tones occur successively (often with other intervening pitches) rather than simultaneously.[30] Thus K0016 has a strong underlying dominant-tonic harmonic organization (Figure 21.5), which can easily be perceived by harmonizing this melody with additional tones (sound example 5).

### Expression

'Mechanical' performances of a melody (as in sound example 1, which was produced by a computer) involve strict adherence to the relative time values of notes as indicated by the musical score, and lack any amplitude variation. Human performances are never mechanical in this sense, but feature systematic variations in timing and amplitude, which help communicate the structure of the music and the performer's artistic and emotional interpretation of it.[31–33] In addition, there are fluctuations in timing and amplitude due to the limits of sensorimotor control. Listeners are quite sensitive to expressive variations. They can easily distinguish an expressive performance from a mechanical one, and can grasp an emotional valence intended by a performer.

### Complexity

In listening to a melody a listener has intuitions concerning its structural complexity. K0016, for example, is relatively simple (informally evidenced by the fact that it can be learned by heart after one or two hearings). Recently, Eerola and North have proposed a computational model of perceived melodic complexity, which successfully predicts data from complexity-rating experiments[34] (Eerola and North, submitted). Their model suggests that perceived complexity is related to the degree to which tonally stable pitches are emphasized and to the strength of a beat, among other factors.[f]

### Meta-relations

In processing a melody a listener hears basic perceptual relations (such as grouping and metre) as well as relations between these relations, that is, meta-relations.[4,37] For example, the slight misalignment of grouping and beat at the onset of phrase 2 in K0016 adds rhythmic energy to the melody (the presence of a note or notes before the first beat of a phrase

[f] It is worth noting that perceived complexity has also been computationally modelled in sentence processing.[35,36] Sentence complexity appears to be related to the number and distance of dependent syntactic heads in a sentence. It would be instructive to contrast computational theories of complexity in the melodic and linguistic domains.

is termed anacrusis or upbeat). Another meta-relation illustrated by K0016 is the relation between contour and beat. Consider phrase 4, where the highest tone (F) does *not* occur on a beat: thus this tone is melodically prominent (as a contour peak) but rhythmically weak. This creates a syncopation (i.e. accent on an off-beat tone), which adds rhythmic interest to the melody. As suggested by these two examples, beat is an important player in forming meta-relations with other aspects of melody.

### Summary: 'relational richness' in musical melody

This brief review illustrates that musical melodies are a prime example of 'rich relations arising from modest means',[38] and as such are of considerable interest for aesthetic theory as well as for cognitive science. One way to appreciate this relational richness is to compare musical melody to speech intonation. Speech intonation is often referred to as 'speech melody' by linguists and phoneticians.[39] This is reasonable because voice pitch in speech is used in an orderly fashion to convey a variety of information, including semantic focus, pragmatic category (e.g. marking statements vs questions), emotional valence, and even lexical identity (in tone languages). Thus melody is not unique to music. Musical melody is quite different from speech melody, however, as evidenced by the fact that people rarely hum or whistle intonation contours (i.e. these contours seldom attract independent interest).[g] In other words, speech melodies rarely draw attention to themselves and away from the semantic message being conveyed. Indeed, one might think of speech melodies as pitch sequences designed to carry out certain quotidian functions without themselves being memorable auditory patterns. In contrast, musical melodies are designed to attract the attention and interest of listeners, and as a result can remain in memory for decades. This difference between musical and linguistic melodies arises from the nature of the perceptual relations between the basic elements of pitch sequences. These relations are far more intricate in musical melody. If a musical melody is 'a group of tones in love with one another', then a speech melody is a loose affiliation of tones (or pitch movements) that work together to get a job done.[h]

## The cognitive neuroscience of melody

The many-faceted nature of melody means that even simple tunes (such as K0016) engage a rich set of mental operations, which serve to structure the relations between tones. Furthermore, these operations occur quite rapidly. For example, K0016 is approximately 10 s long, yet during this time at least 13 aspects of this melody are processed in parallel by a listener. Clearly, then, one desideratum for the cognitive neuroscience of melody is the ability to 'tap into the moment-to-moment history of mental involvement with the music'.[43] In this section I briefly review different approaches to the neural study of melody

[g] The Czech composer Leos Janáček (1854–1928) is an interesting exception. He notated speech intonation contours in music notation and incorporated 'musicalized' versions of these contours into his compositions.[40]

[h] I am referring to speech melody in non-tonal languages such as English. In languages with elaborate tone systems (such as many African languages),[41,42] the organization of tones can be quite complex, though I would argue that they still do not attain the relational richness of musical melodies.

and then introduce a new technique for measuring cortical activity during the perception of melodies.[44]

### Existing approaches to the neuroscience of melody

The study of melody and the brain has relied primarily on three approaches. The first is the neuropsychological approach. This approach focuses on melodic perception deficits in individuals with localized brain damage (e.g. from stroke or surgery), in order to determine if the affected brain regions are involved in aspects of melody perception.[45,46] For example, such studies have suggested that left superior temporal cortex is involved in the processing of pitch interval information, while analogous regions in the right hemisphere have similar capacities but also play a special role in the perception of melodic contour.[19,47]

The second method for studying the neuroscience of melody is the evoked potential or ERP (event-related potential) approach. This approach is based on extracting stimulus-locked population-level neural activity from the EEG or electroencephalogram.[48–50] ERPs provide excellent time resolution—on the order of 10s to 100s of milliseconds—but due to the spatial spreading of bioelectric currents by the skull and scalp it is extremely difficult to localize the brain sources of ERPs. Thus ERPs are not used for localization, but to examine the neural response to individual tones in melodies in fine temporal detail. Initially it might seem that ERPs are ideally suited to studies of melodic processing, since they can resolve the details of neural responses to single tones. However, the low signal-to-noise ratio of evoked potentials means that they must be extracted from the EEG by averaging, that is, by repeating a similar stimulus many (e.g. 30–60) times and averaging the brain's response time-locked to stimulus onset. Thus ERP studies have not been used to study the brain's response to successive tones in a single melody, but have focused instead on responses to a particular tone in multiple repetitions of similar melodies (e.g. an out-of-key note at the end of a melody).[i] Despite this limitation, ERPs have proved a valuable tool: for example, they have demonstrated the influence of musical training on the brain's processing of tonality relations.[51]

The third approach to melody and the brain is the haemodynamic approach, based on techniques such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI). These methods provide indirect measures of neural activity via the measurement of metabolism and/or blood flow in different regions of the brain. In contrast to the ERP method, these techniques provide excellent spatial resolution, and can identify specific brain regions whose activity is modulated by particular types of perceptual processes. For example, research by Zatorre and colleagues points to regions in the right inferior frontal gyrus and the right superior temporal gyrus, which are involved in the maintenance of a tone in memory during the perception of a melodic sequence.[52] Unfortunately, both PET and fMRI have relatively poor temporal resolution since blood

[i] Theoretically, there is no reason why the responses to each tone in a repeated melody could not be examined. However, this would require a listener to hear the same melody dozens of times in an experimental session. This could lead to perceptual/neural responses which are quite different from those occurring in natural listening situations.

flow and metabolism are sluggish signals, which grow and decay over many seconds in response to the physiological demands of the underlying neural tissue. These techniques thus do not operate at the rapid time-scale of melody perception.

Each of the approaches outlined above continues to provide valuable information about melody and the brain, yet none of them fulfils the desideratum mentioned at the beginning of this section. That is, none of them follows brain activity as perception unfolds in time over the course of individual melodies. The ERP method, for example, 'zooms in' to examine the temporal details of neural responses to individual tones, while PET and fMRI 'zoom out' to examine the overall response of various brain regions to entire tone sequences. Melodic processing in the brain, however, is more than the sum of responses to single tones, and more than an average response to an entire sequence. It is a dynamic process of building mental relations between tones during the course of individual melodies. For this reason, techniques are needed that measure patterns of neural activity as perception unfolds within individual sequences. The study described below is one attempt in this direction.

### A new approach for the neural study of melody: the dynamic aSSR (auditory steady-state response) method

**Background** Patel & Balaban[44] set out to examine the temporal evolution of stimulus-related brain activity during perception of tone sequences. To accomplish this, we used a brain signal known as the aSSR.[53] The aSSR is a sinusoidal neural oscillation produced in primary auditory cortex in response to an acoustic stimulus with constant amplitude modulation (AM). aSSR frequency equals the acoustic AM rate, and it is strongest when the AM is in the 40 Hz range.[54,55] Figure 21.6A shows an acoustic signal with constant carrier frequency (400 Hz) and constant AM (40 Hz), and Figure 21.6B shows the power spectrum of a brain signal recorded with an MEG (magnetoencephalography, see later) sensor over auditory cortex while an individual listened to this tone for 1 min. A clear energy peak is visible at 40 Hz. When the same tone is presented without AM, no peak at 40 Hz is visible in the power spectrum (Figure 21.6C). Thus the aSSR is a frequency-specific, stimulus-related brain response with a high signal-to-noise ratio. Unlike an evoked potential, the aSSR represents continuous cortical activity, that is, the oscillation is present as long as the stimulus is on.

How can this seemingly esoteric brain signal be of use to the cognitive neuroscience of melody? The answer to this question has two parts. The first is an empirical observation, namely that the frequency specificity of the aSSR is preserved even when the carrier frequency of a tone changes in a step-wise fashion over the course of two octaves, as long as the AM rate stays constant.[44] Thus one can impose a constant AM on a melody and extract stimulus-related neural activity from the brain by measuring activity at the AM rate. Second, while aSSR frequency does not vary in response to changes in carrier frequency, the amplitude and phase (relative to the stimulus AM) of this oscillation *do* change dynamically over time during the perception of musical sequences.[56] It is these latter variables (the 'dynamic' aspect of the aSSR) which are of primary interest. In particular, the key question is whether aSSR amplitude and phase are sensitive to cognitively relevant processes during melody perception, such as expectancy. If aSSR dynamics are sensitive to cognitive

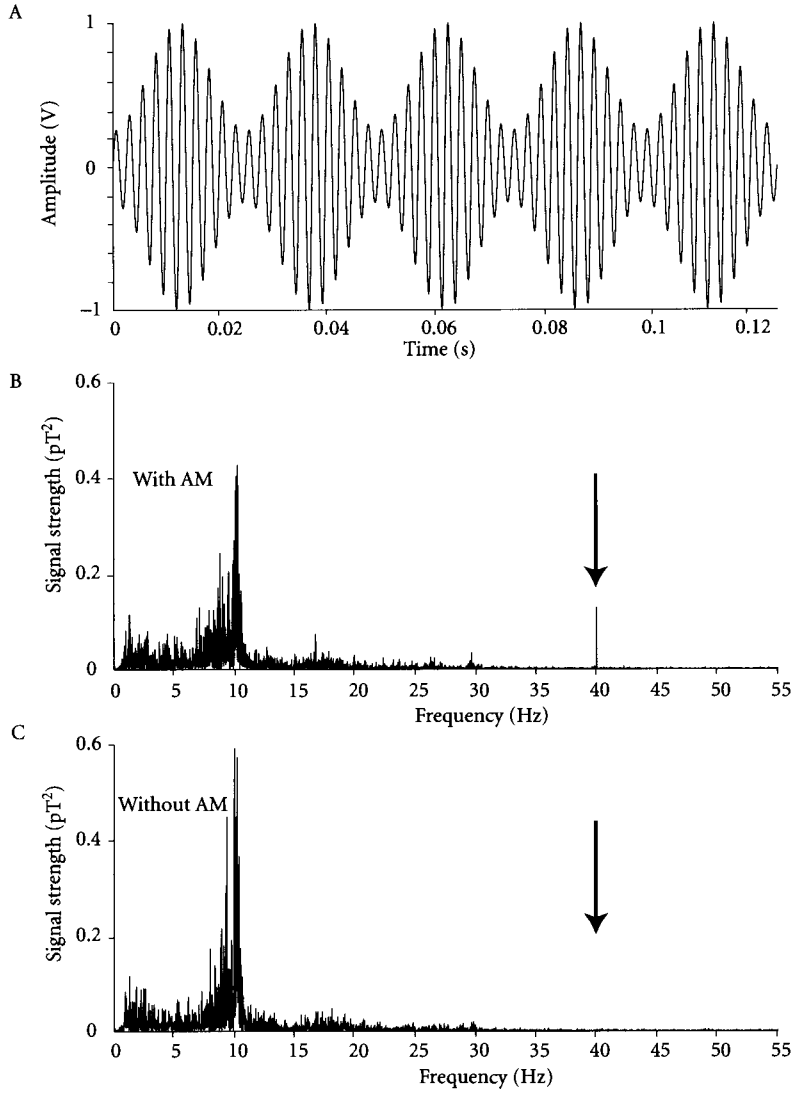



**Figure 21.6** The aSSR. Panel (A) shows a 400 Hz pure tone with 40 Hz AM. Panel (B) shows the power spectrum of the neural response to the tone, recorded from an MEG sensor located over auditory cortex in an adult human listener. Note the strong peak at 40 Hz (arrow) marking the aSSR and indicating the presence of stimulus-related cortical activity. Panel (C) shows a power spectrum of the brain signal when the listener hears a 400 Hz tone without 40 Hz AM. Stimulus-related activity can no longer be distinguished.

processes, it is presumably because cognitive processes influence the state of auditory cortex (e.g. its balance of excitatory and inhibitory activity), which in turn influences how the cortex responds to the AM oscillation. Thus the aSSR may serve as an ongoing probe of cortical function, that is, an indirect measure of cognitive processes over time.

**Details of the tone sequences employed** To explore brain dynamics during tone sequence perception, we used statistically generated tone sequences. This allowed us to generate novel stimuli which lay on a spectrum from random to deterministic in structure. We elected to use statistical tone sequences rather than precomposed melodies so that the sequences would be unfamiliar to participants, easily generated in quantity, and mathematically well characterized. The latter two points were of particular importance because we were employing a novel brain imaging technique and wanted to have good control over the stimuli.

All tone sequences were approximately 1 min long, consisting of ~150 pure tones (415 ms each) with no temporal gaps. Sequences were diatonic, and ranged between A3 (220 Hz) and A5 (880 Hz) in pitch. Four structural categories of sequences were employed: random, deterministic (musical scales), and two intermediate 'fractal' categories of constrained variation more reminiscent of musical melody. These categories were given mathematical names in accordance with the technique used to generate them:[j] $1/f$ ('one over $f$') and $1/f^2$ ('one over $f$ squared').[57]

A qualitative understanding of these categories is possible without delving into the underlying mathematics. In random sequences each successive pitch is chosen independently of the previous one, and there are no long-term pitch trends. Deterministic sequences represent the opposite case: they consist entirely of long-term pitch trends (predictable stair-like patterns) with no short-term unpredictability. The fractal sequences are intermediate. $1/f$ sequences have a hint of long-term pitch trends but still have much unpredictable variation from one pitch to the next. $1/f^2$ sequences are strong in long-term pitch trends, but retain a small amount of unpredictability in the behavior of successive tones (Figure 21.7, sound examples 6–9).

**Perceptual task and brain recordings** Participants ($n = 5$ right-handed males, 2 with musical training) were familiarized with the different stimulus categories in a training session where examples of each category were presented along with an arbitrary category label (the numbers 1–4). Participants quickly learned to identify the different categories, and during the experiment, classified novel sequences by their category with little difficulty. The experiment consisted of 28 such sequences (7 per category) presented in random order. Stimuli in each category were equally distributed among seven Western diatonic modes (ionian, dorian, phrygian, lydian, mixolydian, aeolian, and locrian). Each participant heard a unique set of stimuli for the random, $1/f$, and $1/f^2$ conditions (all heard the same set of scales).

During stimulus presentation, neural data were recorded using 148-channel whole head MEG. MEG measures magnetic fields produced by electrical activity in the brain, providing a signal with similar time resolution to EEG but with certain advantages relating to source localization and independence of signals recorded from different parts of the sensor array.[58]

**aSSR measurements** We used the aSSR to detect stimulus-related neural activity. Each sequence was given a constant rate of AM (41.5 Hz). This AM gave the tone sequences a slightly warbly quality, without disrupting their perceived pitch pattern: listeners heard them as sequences of pitches at the underlying pure tone frequencies (sound examples

[j] Inverse Fourier transforms of power spectra with different slopes. See Patel and Balaban[44] for details.

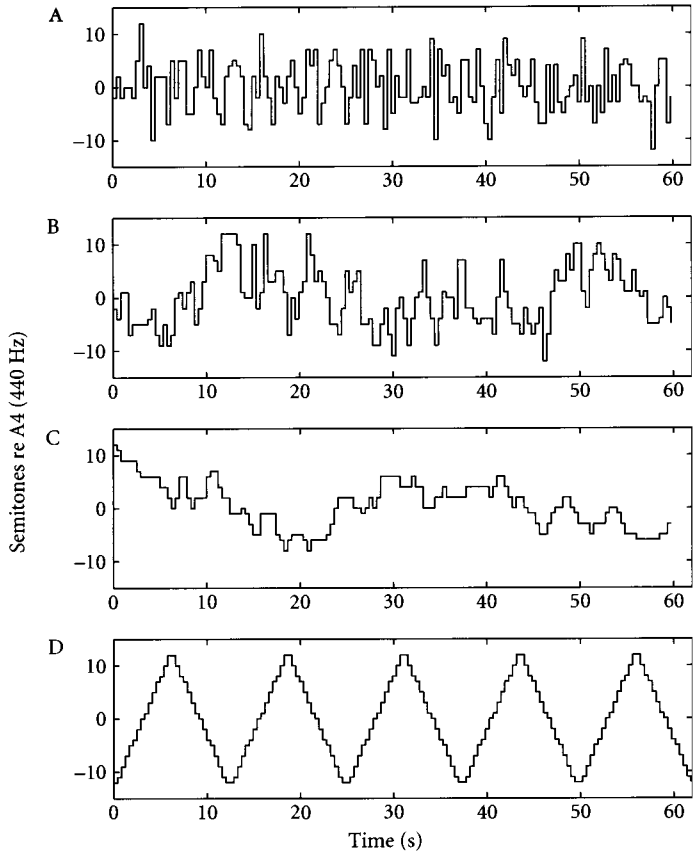



Figure 21.7 Examples of pitch contours from the different conditions used by Patel and Balaban[44] (A) Random sequence, (B) $1/f$ sequence, (C) $1/f^2$ sequence, (D) scales. All sequences had ~150 tones of 415 ms each, with no pauses between tones. The sequences in this figure can be heard in sound examples 6–9.

10–13). For each sequence heard by a subject we studied the amplitude and phase (relative to the acoustic AM) of the aSSR in contiguous 2-s epochs from each channel.[k] That is, one amplitude and phase value of the aSSR was obtained from each successive 2-s epoch of the channel's brain signal via a Fourier transform, yielding approximately 30 amplitude and phase values ×148 channels per sequence.

Since a good deal of our analysis concerns phase information, it is worth giving a brief conceptual explanation of phase. By amplitude modulating our tone sequences at 41.5 Hz, we are introducing an oscillatory signal into the brain at that same frequency. This causes an oscillatory response (the aSSR) at that frequency in certain brain regions. The relative

[k] Two-second epochs were chosen on the basis of pilot analysis suggesting that this was the shortest epoch we could study and still maintain a desired signal to noise ratio.

degree of lag between the time-referenced input signal and the oscillatory brain response is measured by the phase of the aSSR.

**Results I: phase tracking** Our first finding was that the phase of the measured brain signal varied with the pitch of the tone sequence. As pitch increased, phase advanced (corresponding to a *decreased* lag between stimulus and brain response), and vice-versa. This relationship between aSSR phase and carrier frequency was suggested by early work,[53] and has been confirmed by recent studies,[55,59] but had not previously been studied in a dynamic fashion. The underlying mechanism for this phenomenon likely involves the tonotopic layout of the basilar membrane in the cochlea of the human ear, where higher frequencies are closer to the oval window and hence are stimulated earlier than lower frequencies. However, the degree of phase change observed in the aSSR cannot be solely explained by peripheral neural mechanisms (Patel and Balaban, in preparation).

In our study, variation of phase with stimulus carrier frequency manifested itself in 'phase tracking', that is, in a correlation between the shape of the phase-time contour and the stimulus carrier frequency-time contour (Figure 21.8). As just mentioned, this tracking may be largely driven by peripheral neural mechanisms, and we make no claim that phase tracking has a causal relationship to tone sequence processing (indeed, it is only present because of the AM we have imposed on the tone sequences). Rather, the interesting question is whether phase tracking is modulated by perceptual aspects of the stimulus. We found evidence that this may be the case, in that phase tracking improved as sequences became more predictable in structure. Examples of phase-time contours (solid lines) overlaid on their corresponding pitch-time contours (dashed lines) are shown in Figure 21.8A–D, showing how tracking improves across the stimulus conditions. The best tracking occurred for musical scales, which have a completely predictable structure.

Each subject showed a number of sensor locations where this 'phase tracking' of pitch was observed. Across participants, these locations tended to be in fronto-temporal regions, with a slight right-hemisphere bias. A similar set of locations was identified when we looked for sensors where the amplitude of the aSSR was strong. However, we found no evidence that the amplitude time series of the aSSR correlated with the heard pitch contour (see Patel and Balaban[44] for details).

**Results II: phase coherence** Knowing that the phase of the brain response reflected stimulus properties (i.e. pitch contour), we then turned to looking at patterns of phase coherence between different brain regions. Phase coherence does not measure the lag between an oscillatory signal and brain response but rather the stability of the phase *difference* between oscillatory activity in different brain areas. Thus phase coherence measures the degree of temporally correlated activity in distinct brain regions. If two brain areas show greater phase coherence during certain conditions, this is suggestive of a greater degree of functional coupling between regions under those conditions.[60,1]

[1] In studying phase coherence at the aSSR frequency, we are not assuming that the aSSR is itself part of the causal chain of melodic processing, but rather that its dynamics are influenced by ongoing brain processes which *are* involved in melody perception. Thus aSSR coherence serves as a proxy for coherent brain activity, which is related to melodic processing. (It should be noted that the idea that phase coherence between brain regions has some functional significance is a topic of active research in modern neuroscience, which enjoys increasing evidence but which has yet to be conclusively proved.)

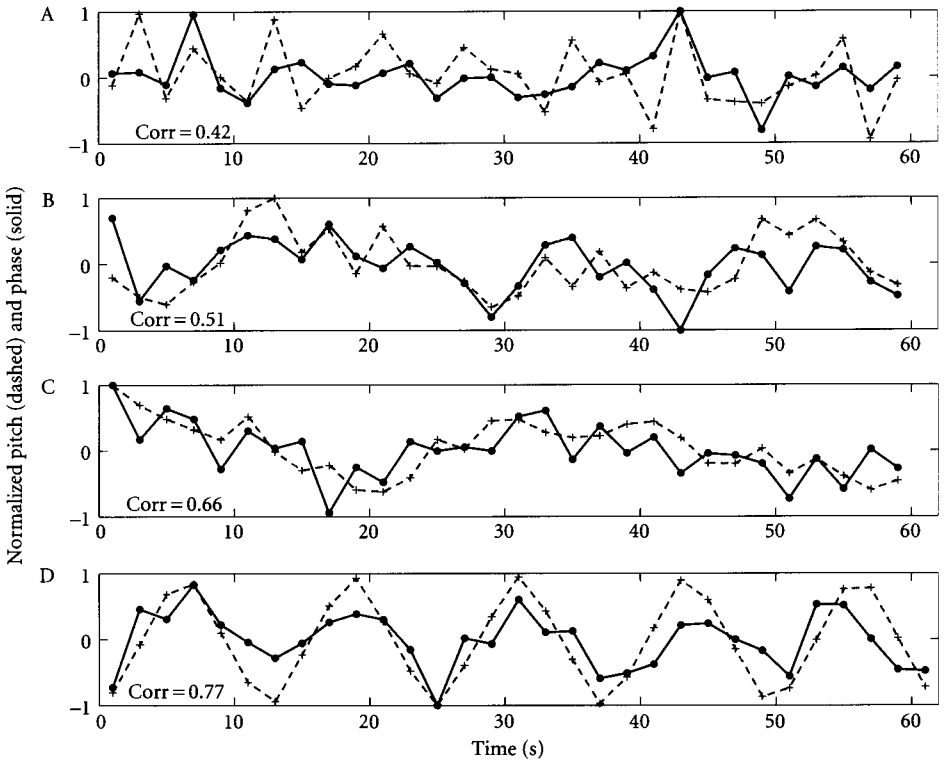



**Figure 21.8** Examples of phase-time contours (solid lines) extracted from brain signals recorded with MEG sensors over auditory cortical regions. Each phase-time contour is overlaid on its associated pitch-time contour (dashed lines: these contours have been down sampled to show the average pitch during each 2-s epoch in which aSSR phase was measured). These contours represent single channel, single-trial data recorded from one listener. Correlations between each phase-time contour and its associated down-sampled pitch-time contour are shown in the inset to each graph. Full details of correlation between phase and pitch across sensors, conditions, and subjects is given in Patel and Balaban.[44]

We found that across participants, the different conditions were characterized by differing degrees of phase coherence. Random sequences generated less phase coherence than all other categories, and among the structured categories, scales (which showed the best phase tracking) did *not* generate the highest coherence. Rather, $1/f^2$ sequences were associated with the greatest degree of phase coherence (Figure 21.9A). Interestingly, statistical research on Western music suggests that melodies have approximately $1/f^2$ statistics,[61,62] suggesting that melody-like sequences generated more brain interactions than sequences which were either overly random or overly deterministic.

To better understand the nature of these interactions, we examined topographic patterns of phase coherence, subdividing the brain into four quadrants (anterior and posterior left and right). We found that the greater phase coherence of $1/f^2$ sequences was driven by interactions between the left posterior hemisphere and the rest of the brain, including the

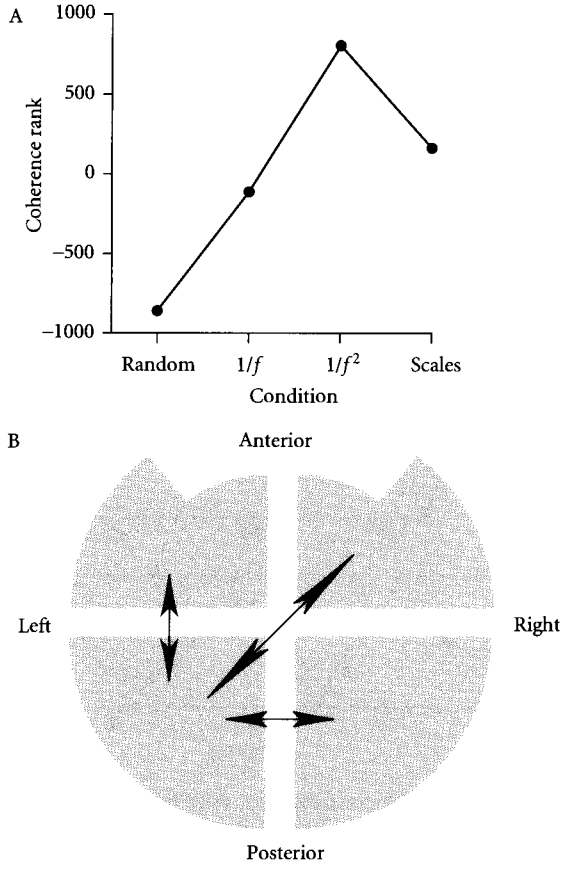



**Figure 21.9** Results of phase coherence analysis conducted in order to study the interaction of brain regions during the perception of different sequence types. Panel (A) shows that coherence is lowest for random sequences, and that among the structured sequences, is greatest in the $1/f^2$ sequences (which most closely resemble real musical melodies in their statistical structure). Panel (B) shows the regional interactions underlying this result. The relatively higher coherence in $1/f^2$ sequences results from interactions between the left posterior brain quadrant and other brain quadrants. The coupling between left posterior and right anterior and posterior brain regions may represent a dynamic neural correlate of the integration of pitch interval and pitch contour perception, that is, local and global pitch processing (see text and Patel and Balaban[44] for details).

two right hemisphere quadrants (Figure 21.9B). This is of interest because neuropsychological studies of brain-damaged patients suggests that left superior temporal regions are involved with the discrimination of precise interval sizes, while right fronto-temporal circuits have an additional role in the perception of more global contour patterns.[19,47] Thus the observed pattern of coherence may reflect the dynamic integration of local and global pitch perception, and suggests that this integration is greatest when tone sequences resemble musical melodies.

## Conclusions

The study of melody perception provides both a challenge and an opportunity for cognitive neuroscience. Melodies, like sentences, are complex structures, and unraveling the neural mechanisms by which they are understood is a project which is still in its infancy. However, melodies do afford some attractive properties to the cognitive neuroscientist. Their raw materials can be extremely simple (e.g. pure tones instead of phonemes), and their complexity can be systematically varied using quantitative principles.[34] Thus they provide a valuable tool with which to study basic issues in cognitive neuroscience, such as the interaction of bottom-up and top-down processes in sequence perception. Indeed, it is possible that neural principles learned from using melodies may prove useful in understanding brain processing of other types of complex sequences, such as sentences.

This chapter has described a new approach to the neuroscience of melody, based on a method which monitors stimulus-related cortical activity over time during the perception of individual tone sequences.[44,m] Using the aSSR, we have demonstrated that it is possible to extract a signal from the human cerebral cortex which reflects the pitch contour an individual is hearing. The accuracy with which this signal reflects the pitch contour improves as the pitch sequence becomes more predictable. Thus there may be top-down influences of musical expectancy which influence this brain signal, suggesting that future studies may be able to use aSSR dynamics to monitor how expectancy is structured in time.

The basis of aSSR phase tracking is *temporal* information in cortical activity. When the *amount* of activity was examined, no relationship with pitch contour was observed. This suggests that dynamic imaging techniques have an important role to play in the study of music perception, complementing techniques sensitive to the amount of neural activity but insensitive to the fine temporal structure of that activity (e.g. PET & fMRI).

Dynamic imaging techniques also offer the opportunity to study how brain areas interact during perception. It is clear from decades of neural research that the brain is divided into different regions, each of which has a special role to play in perception and cognition. Yet it is also clear that these brain areas must interact to form coherent and unified percepts. Complex patterns such as music and speech engage multiple brain regions, and sequences with different perceptual properties may be distinguished by the pattern of brain interactions they engender rather than by the particular brain regions which respond to them. Using phase coherence, we examined brain interactions as a function of stimulus structure and found that sequences with melody-like statistics were associated with the greatest degree of neural interactions. In particular, we found evidence for strong functional coupling between left posterior hemisphere and right hemisphere regions during the perception of melody-like sequences. This may reflect the perceptual integration of local and global pitch patterns, and suggests that one neural signature of melody is the dynamic integration of brain areas which process structure at different time scales.

Two obvious directions to pursue with the dynamic aSSR approach are the examination of tonality's influence on brain dynamics and the dynamic neural correlates of rhythm perception. The dynamic aSSR approach may also provide a way to examine brain interactions

[m] For a related approach based on the visual steady-state response, see Harris and Silberstein.[63]

as a function of stimulus structure in real melodies. This may allow an objective quantification of perceptual coherence of melodies in individuals who cannot easily give details of their perception, such as young children and infants. Without question, future studies of melody and the brain will benefit from the integration of existing approaches (e.g. neuropsychology, ERPs, fMRI, and dynamic aSSR), as well as the development of new approaches. No matter what the approach, the goal is the same: to understand how the mind converts 'mere sequences of tones' into meaningful cognitive and aesthetic experiences, which can reside in memory for a lifetime.

## Acknowledgements

This work was supported by Neurosciences Research Foundation as part of its program on music and the brain at The Neurosciences Institute, where A.D. Patel is the Esther J. Burnham fellow. I thank Joseph Daniele and John Iversen for technical assistance with figures and sound examples, and W. Jay Dowling, John Iversen, and Bruo Repp for comments which improved the manuscript.

## References

1. Hast, D. E., J. R. Cowdery, and S. Scott (eds) (1999) *Exploring the World of Music.* Dubuque: Kendall Hunt. (Quote from interview with Simon Shaheen in video program #6: Melody.)
2. Schaffrath, H. (1995) In D. Huron (ed.) *The Essen Folksong Collection in the Humdrum Kern Format.* Menlo Park, CA: Center for Computer Assisted Research in the Humanities.
3. Selfridge-Feld, E. (1995) *The Essen Musical Data Package.* CCARH (Center for Computer Assisted Research in the Humanities) Technical Report No. 1. Menlo Park, CA: CCARCH.
4. Meyer, L. B. (1973) *Explaining Music: Essays and Explorations.* Berkeley: University of California Press.
5. Narmour, E. (1990) *The Analysis and Cognition of Basic Melodic Structures: The Implication-realization Model.* Chicago: University of Chicago Press.
6. Narmour, E. (1992) *The Analysis and Cognition of Melodic Complexity: The Implication-realization Model.* Chicago: University of Chicago Press.
7. Hajda, J. M., R. A. Kendall, E. C. Carterette, and M. L. Harshberger (1997) Methodological issues in timbre research. In I. Deliège and J. Sloboda (eds) *Perception and Cognition of Music.* Hove, UK: Psychology Press, pp. 253–306.
8. Lerdahl, F. and R. Jackendoff (1983) *A Generative Theory of Tonal Music.* Cambridge, MA: MIT Press.
9. Jusczyk, R. and C. Krumhansl (1993) Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *J. Exp. Psychol.: Hum. Percept. Perform.* 19(3), 627–40.
10. Large, E. (2000) On synchronizing movements to music. *Hum. Movement Sci.* 19, 527–66.
11. Snyder, J. and C. L. Krumhansl (2001) Tapping to ragtime: cues to pulse finding. *Music Percept.* 18(4), 455–89.
12. Cooper, G. W. and L. B. Meyer (1960) *The Rhythmic Structure of Music.* Chicago: University of Chicago Press.

13. **Drake, C., M. R. Jones, and C. Baruch** (2000) The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition* 77, 251–88.
14. **Dowling, W. J.** (1978) Scale and contour: two components of a theory of memory for melodies. *Psychol. Rev.* 85(4), 341–54.
15. **Dowling, W. J.** (2001) Perception of music. In E. B. Goldstein (ed.) *Blackwell Handbook of Perception.* Malden, MA: Blackwell.
16. **Ayotte, J., I. Peretz, and K. Hyde** (2002) Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain* 125, 238–51.
17. **Dowling, W. J., S. Kwak, and M. Andrews** (1995) The time course of recognition of novel melodies. *Percept. Psychophys.* 57(2), 136–49.
18. **Trehub, S. E., D. Bull, and L. A. Thorpe** (1984) Infant's perception of melodies: the role of melodic contour. *Child Dev.* 55, 821–30.
19. **Patel, A. D., I. Peretz, M. Tramo, and R. Labrecque** (1998) Processing prosodic and musical patterns: a neuropsychological investigation. *Brain Lang.* 61, 123–44.
20. **Schmuckler, M. A.** (1999) Testing models of melodic contour similarity. *Music Percept.* 16, 295–326.
21. **Krumhansl, C. L., J. Louhivuori, P. Toiviainen, T. Järvinen, and T. Eerola** (1999) Melodic expectancy in Finnish folk hymns: convergence of behavioral, statistical, and computational approaches. *Music Percept.* 17, 151–96.
22. **Schellenberg, G.** (1996) Expectancy in melody: Tests of the Implication-realization Model. *Cognition* 58, 75–125.
23. **Bregman, A.** (1990) *Auditory Scene Analysis: The Perceptual Organization of Sound.* Cambridge, MA: MIT Press.
24. **Bigand, E.** (1997) Perceiving musical stability: The effect of tonal structure, rhythm and musical expertise. *J. Exp. Psychol.: Hum. Percept. Perform.* 23, 808–12.
25. **Krumhansl, C. L.** (1990) *Cognitive Foundations of Musical Pitch.* New York: Oxford University Press.
26. **Schenker, H.** (1969) *Five Graphic Music Analyses.* New York: Dover.
27. **Cook, N.** (1987) *A Guide to Musical Analysis.* Oxford: Oxford University Press.
28. **Jones, S.** (1995) *Folk Music of China.* NY: Oxford University Press.
29. **Bharucha, J. J.** (1984) Event hierarchies, tonal hierarchies, and assimilation: a reply to Deutsch and Dowling. *J. Exp. Psychol.: Gen.* 113(3), 421–5.
30. **Povel, D.-J. and E. Jansen** (2001) Perceptual mechanisms in music processing. *Music Percept.* 19(2), 169–97.
31. **Repp, Bruno H.** (1995) Expressive timing in Schumann's 'Traeumerei': an analysis of performances by graduate student pianists. *J. Acoust. Soc. Am.* 98(5, Pt 1), 2413–27.
32. **Palmer, C.** (1997) Music performance. *Ann. Rev. Psychol.* 48, 115–38.
33. **Juslin, Patrik N.** (2000) Cue utilization in communication of emotion in music performance: relating performance to perception. *J. Exp. Psychol.: Hum. Percept. Perform.* 26(6), 1797–813.
34. **Eerola, T. and A. North** (2000) Expectancy-based model of melodic complexity. In C. Woods, G. Luck, R. Brochard, F. Seddon, and J. A. Sloboda (eds) *Proceedings of the 6th International Conference on Music Perception and Cognition.* Keele, UK: Keele University Department of Psychology.
35. **Gibson, E.** (1998) Linguistic complexity: locality of syntactic dependencies. *Cognition* 68(1), 1–76.

36. **Gibson, E.** (2000) The dependency locality theory: a distance-based theory of linguistic complexity. In Y. Miyashita *et al.* (eds) *Image, Language, Brain.* Cambridge, MA: MIT Press, pp. 95–126.
37. **Jones, M. R.** and **P. Q. Pfordresher** (1997) Tracking musical patterns using Joint Accent Structure. *Can. J. Exp. Psychol.* 51(4), 271–90.
38. **Meyer, L. B.** (2000) *The Spheres of Music: A Gathering of Essays.* Chicago: University of Chicago Press.
39. **'t Hart, J., R. Collier,** and **A. Cohen** (1990) *A Perceptual Study of Intonation: An Experimental-Phonetic Approach to Speech Melody.* Cambridge: Cambridge University Press.
40. **Wingfield, P.** (ed) (1999) *Janáček Studies.* Cambridge: Cambridge University Press.
41. **Connell, B.** (2000) The perception of lexical tone in Mambila. *Lang. Speech* 43(2), 163–82.
42. **Hyman, L.** (2001) Tone systems. In M. Haspelmath (ed.) *Language typology and language universals: an international handbook.* Berlin: W. de Gruyter.
43. **Sloboda, J.** (1985) *The Musical Mind: The Cognitive Psychology of Music.* Oxford: Oxford University Press.
44. **Patel, A. D.** and **E. Balaban** (2000) Temporal patterns of human cortical activity reflect tone sequence structure. *Nature* 404, 80–4.
45. **Peretz, I.** (1990) Processing of local and global musical information in unilateral brain-damaged patients. *Brain* 113, 1185–205.
46. **Steinke, W. R., L. L. Cuddy,** and **L. S. Jakobson** (2001) Dissociations among functional subsystems governing melody recognition after right hemisphere damage. *Cognit. Neuropsychol.* 18, 411–37.
47. **Liégeois-Chauvel, C., I. Peretz, M. Babaï, V. Laguitton,** and **P. Chauvel** (1998) Contribution of different cortical areas in the temporal lobes to music processing. *Brain* 121, 1853–67.
48. **Besson, M.** (1997) Electrophysiological studies of music processing. In I. Deliège and J. Sloboda (eds) *Perception and Cognition of Music.* Hove, UK: Psychology Press, pp. 217–50.
49. **Tervaniemi, M.** *et al.* (2001) Superior formation of cortical memory traces for melodic patterns in musicians. *Learning Memory* 8, 295–300.
50. **Trainor, L. J., K. L. McDonald,** and **C. Alain** (2002) Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *J. Cognit. Neurosci.* 14(3), 430–42.
51. **Besson, M.** and **F. Faïta** (1995) An event-related potential (ERP) study of musical expectancy: comparison of musicians with non-musicians. *J. Exp. Psychol.: Hum. Percept. Perform.* 21, 1278–96.
52. **Zatorre, R. J., A. C. Evans,** and **E. Meyer** (1994) Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 14, 1908–19.
53. **Galambos, R., S. Makeig,** and **P. J. Talmachoff** (1981) A 40-Hz auditory potential recorded from the human scalp. *Proc. Natl. Acad. Sci. USA* 78, 2463–647.
54. **Gutschalk, A.** *et al.* (1999) Deconvolution of 40 Hz steady-state fields reveals two overlapping source activities of the human auditory cortex. *Clin. Neurophys.* 110, 856–68.
55. **Ross, B., C. Borgmann, R. Draganova, L. Roberts,** and **C. Pantev** (2000) A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. *J. Acoust. Soc. Am.* 108, 679–91.
56. **Galambos, R.** and **S. Makeig** (1988) Dynamic changes in steady-state responses. In E. Başar (ed.) *Dynamics of Sensory and Cognitive Processing by the Brain.* Berlin: Springer-Verlag, pp. 103–22.

57. **Schmuckler, M. A.** and **D. L. Gilden** (1993) Auditory perception of fractal contours. *J. Exp. Psychol. Hum. Percept. Perform.* 19, 641–60.

58. **Lewine, J. D.** and **W. W. Orrison** (1995) Magnetoencephalography and magnetic source imaging. In W. W. Orrison *et al.* (ed.) *Functional Brain Imaging.* St Louis: Mosby, pp. 369–417.

59. **John, M. S.** and **T. W. Picton** (2000) Human auditory steady-state responses to amplitude modulated tones: phase and latency measurements. *Hearing Res.* 14, 57–79.

60. **Bressler, S.** (1995) Large-scale cortical networks and cognition. *Brain Res.: Brain Res. Rev.* 20(3), 288–304.

61. **Nettheim, N.** (1992) On the spectral analysis of melody. *Interface* 21, 135–48.

62. **Boon, J. P.** and **O. Decroly** (1995) Dynamical systems theory for music dynamics. *Chaos* 5, 501–8.

63. **Harris, P.** and **R. Silberstein** (1999) Steady-state visual evoked potential (SSVEP) responses correlate with musically trained participants' encoding and retention phases of musical working memory task performance. *Aust. J. Psychol.* 51(3), 140–6.

CHAPTER 22

# HOW MANY MUSIC CENTRES ARE IN THE BRAIN?

ECKART O. ALTENMÜLLER

## Abstract

When reviewing the literature on brain substrates of music processing, a puzzling variety of findings can be stated. The traditional view of a left–right dichotomy of brain organization—assuming that in contrast to language, music is primarily processed in the right hemisphere—was challenged 20 years ago, when the influence of music education on brain lateralization was demonstrated. Modern concepts emphasize the modular organization of music cognition. According to this viewpoint, different aspects of music are processed in different, although partly overlapping neuronal networks of both hemispheres. However, even when isolating a single 'module', such as, for example, the perception of contours, the interindividual variance of brain substrates is enormous. To clarify the factors contributing to this variability, we conducted a longitudinal experiment comparing the effects of procedural vs explicit music teaching on brain networks. We demonstrated that cortical activation during music processing reflects the auditory 'learning biography', the personal experiences accumulated over time. Listening to music, learning to play an instrument, formal instruction, and professional training result in multiple, in many instances multisensory, representations of music, which seem to be partly interchangeable and rapidly adaptive. In summary, as soon as we consider 'real music' apart from laboratory experiments, we have to expect individually formed and quickly adaptive brain substrates, including widely distributed neuronal networks in both hemispheres.

**Keywords:** Music; Brain; Neuromusicology; Music centres

## Changing concepts in neuromusicology

During the past two decades, the concepts of brain substrates underlying music processing have changed. Although never unequivocally supported by classical lesion studies,[1,2] traditional theories proposed a simple right-vs-left–hemisphere dichotomy, with music being processed in the right brain, language in the left. This simple viewpoint—still represented in many textbooks—could not be held any longer, when in 1974 Bever and Chiarello[3] were able to demonstrate the influence of professional training on hemispheric lateralization during music processing, nonmusicians exhibiting right, professionals left hemispheric preponderance. In the following years, results of several brain imaging studies supported this idea. In our laboratory, for example, brain activation patterns were investigated during

demanding harmonic and melodic discrimination tasks in a large group of nonmusicians, amateurs, and professional musicians. Professional musicians processed these tasks primarily in the left frontotemporal lobes, whereas amateurs as well as nonmusicians bilaterally activated the frontal lobes and the right temporal lobe.[4] It was assumed that professional musicians had access to different cognitive strategies ascompared to amateurs and nonmusicians. The left hemisphere activation in professionals was attributed to covert inner speech, since trained musicians—as a consequence of hundreds of lessons of ear training and solfège—reported that they named the intervals and harmonies more or less automatically during processing of the task. In other words and speaking more generally, access to 'auxiliary' representations of music acquired during training was discussed as accounting in part for the variability in brain activation patterns during music listening. Consequently, brain substrates of music processing were supposed to reflect the *way of listening and processing* rather than more-or-less fixed 'music centres'.

## Listening to music: concepts of perceptive modules and hierarchies

Before considering brain substrates of music processing, we have to clarify what we term 'music' in this context. To our understanding, music is not a mere acoustic structure in time, or a stimulus created in a laboratory to fit a well-controlled experimental design, but a phenomenon of subjective human experience. Such an experience is not based on a uniform mental capacity but on a complex set of perceptive and cognitive operations represented in the central nervous system. These operations act interdependently in some parts, independently in others. They are integrated in time and linked to previous experiences with the aid of memory systems, thus enabling us to perceive, or better, to 'feel' a sort of meaning while listening. Neuromusicology has been profoundly influenced by the idea of the *modularity of musical functions*.[5,6] According to Fodor,[5] a module corresponds to a specialized computational device that is devoted to the execution of some biologically important function. Applied to music, this concept has been put forward by the groups of Isabelle Peretz and Robert Zatorre, demonstrating convincingly the neuropsychological fractionation of different musical subfunctions in patients following brain lesions.[7–12] Taking together the results of these studies, we see that a complex pattern of distinct dissociation syndromes with isolated loss of cognitive subunits of music processing following a lesion emerges. For example, there is evidence that separate modules are processing time or pitch structures of complex musical stimuli. Generally speaking, time structures seem to be processed to a greater extent in the left temporal lobe, whereas pitch structures may be processed primarily in right temporal lobe networks. According to recent results, a predominance of the posterior portions of the right supratemporal lobe for processing of pitch structures may exist.[13]

Such a modular organization concerning processing of segregated physical (temporal or pitch) properties of musical structures could account for the involvement of separated, in part, overlapping, neuronal substrates. However, the situation becomes more complex when one considers that perception of music may occur on different hierarchical levels. With respect to temporal structures, for example, two levels of organization may be distinguished: metre and rhythm. Rhythm is defined as the serial relation of durations between

different acoustical events in a train of sounds—that is, rhythm represents a serial durational pattern—whereas metre involves a temporal invariance in terms of the regular recurrence of pulses or beats marking off equal durational units, which can be organized as measures. Metre therefore represents a more complex acoustical 'gestalt', since its perception and production require information on sound intensity (accented and unaccented events) and on periodicity of rhythmical events, the latter based on integration of information over longer time periods. Again, by means of neuropsychological fractionation of perceptive subfunctions in patients with brain lesions, it was possible to isolate neuronal networks processing metre and rhythm, demonstrating spared metric judgement but disrupted rhythmic discrimination.[7] It is beyond the scope of the present contribution to discuss the complex issue of local vs global gestalt processing in the auditory domain. Whereas in the visual modality, a dissociation of neuronal substrates concerned with local (left parietal) or global (right parietal) processing can clearly be demonstrated,[14] in the auditory modality such a clear-cut neuroanatomical distinction is still under debate. This may be related to the methodological problem of fractionating time structures in appropriate perceptive units allowing for a clear and interindividually consistent discrimination between local or global processing units. In fact, most experimental designs have not been very convincing in excluding variable time fractionation based on rapidly changing individual auditory habits.

## 'Music centres' in the brain reflecting the auditory biography

Although lesion studies conducted in larger groups of patients suffering from well-defined lesions support the idea of modularity and hierarchy and yield relatively clear results, it still remains a scientific challenge to further clarify the nature and function of these hypothesized modules and hierarchies and to delineate their degree of autonomy and specificity to music. Another, in our opinion, even more urgent problem arises when taking a look at the individual data. In a recently published study,[15] where procedures were performed on patients suffering from small unilateral ischemic lesions of the temporal, parietal, or frontal lobe, respectively, a surprising heterogeneity of the patterns of impairment with respect to modular subfunctions of auditory processing emerged. Whereas none of the patients had a deficit in basic auditory perception as, for example, in pitch discrimination, in some individuals right hemispheric lesions produced deficits in more complex temporal or pitch organization tasks, including in all instances, however, a combined deficit of local as well as global processing stages. Following left hemispheric stroke, surprisingly dissociated impairments of rhythm, metre, interval, or contour processing occurred, irrespective of whether the lesion was localized anterior or posterior to the central sulcus. In summary, it emerged that individually variable varieties of brain regions were necessary to ensure complex auditory functions, including parts of the posterior parietal lobes and the frontal lobes. It therefore becomes evident that not only the temporal lobes, as suggested in the comprehensive study of Liégois-Chauvel and co-workers,[13] but widespread and individually developed neuronal networks may underlie music processing. Without going into methodological details, it should be mentioned in this context that our patients were investigated relatively early—7–14 days following the ischemic stroke—before profound plastic changes compensating

for impaired functions had reached their full extent. The discrepancies with earlier studies applying similar test batteries can in part be explained by the early timing of the investigation.

When considering individual factors as possible sources of variability in brain substrates of music processing, the next step is to further delineate the nature of these factors. In order to clarify whether the *way of learning music* plays an important role, we investigated the impact of music education on brain activation patterns in a group of 13- to 15-year-old students[16] in close cooperation with the music educator Wilfried Gruhn. The hypotheses were that (1) learning music and acquiring a new mental representation of music changes brain activation patterns while listening to music, and that (2) different ways of music learning may cause various mental representations that are reflected in different cortical activation patterns.

The task was to judge formal aspects of symmetrically structured phrases, so-called musical periods that consist of corresponding parts, 'antecedent' and 'consequent'. Students had to distinguish between correct and incorrect (balanced or unbalanced) phrases. Whereas the antecedent phrase ends in a weak cadence on the dominant, suspending the expected tonic (half cadence), the consequent phrase leads to a stable ending on the tonic (perfect cadence). This different quality of cadences and the balance of the two melodic parts can easily be recognized merely by an internal feeling of musical balance and the tension of the cadence. For training, subjects were divided into three subgroups: (1) a 'declarative' learner group that received traditional instructions about the antecedent and consequent and their tonal relationship with respect to the closing on a complete or incomplete cadence (the instructions included verbal explanations, visual aids, notations, verbal rules, and some musical examples that were played for the subjects, but never sung or performed); (2) a 'procedural' learner group that participated in musical experiences for establishing genuine musical representations by singing and playing, improvising with corresponding rhythmic and tonal elements, or performing examples from the music literature; and (3) a control group of nonlearners who did not receive any instruction about or in music. Low-frequency DC shifts of the electroencephalogram (EEG) were measured prior to learning and after a five-week training period.

In Figure 22.1, the main results of the study are summarized. After learning, in the verbally trained 'declarative' group, music processing produced an increased activation of the left frontotemporal brain regions, which probably reflects inner speech and analytical, step-by-step processing. By contrast, the musically trained 'procedural' group showed increased activation of the right frontal and of bilateral parietooccipital lobes, which may be ascribed to a more global way of processing and to visuospatial associations. These results demonstrate for the first time directly that musical expertise influences auditory brain activation patterns and that changes in these activation patterns depend on the teaching strategies applied. In other words, *brain substrates of music processing reflect the auditory learning 'biography', the personal experiences accumulated over time.* Listening to music, learning to play an instrument, formal instruction, and professional training result in multiple, in many instances multisensory, representations of music, which seem to be partly interchangeable and rapidly adaptive.

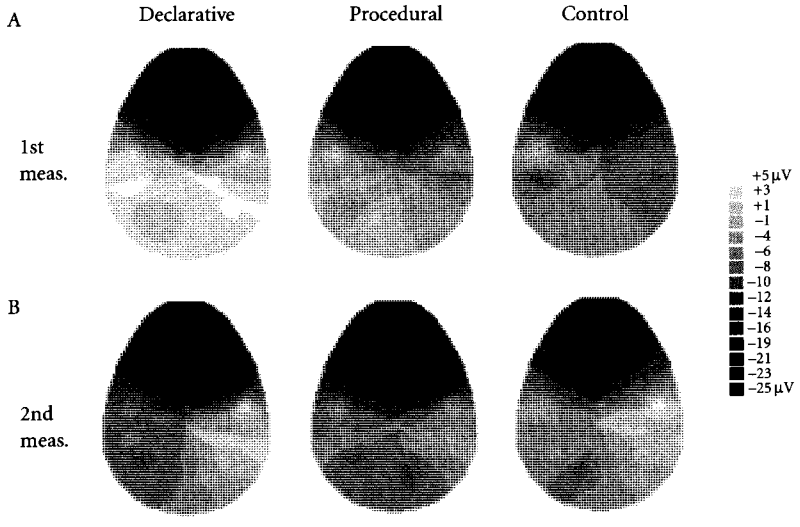



**Figure 22.1** Brain maps demonstrating cortical activation patterns (A) before and (B) after learning in the 'Declarative' learning group, in the 'Procedural' learning group, and in the 'Control' group. Group averages are displayed. Activation is *dark*, inactivation is *white* (see microvolt scale on the *right*). The brain diagrams are displayed as *top views*, frontal regions up, left hemisphere on the *left*, right hemisphere on the *right*. As can be recognized, declarative, mainly verbally mediated training leads to an increase in brain activity over the left frontal areas, whereas procedural, genuinely musical training produces an increase in activity over right frontal and bilateral parietooccipital regions. In controls, overall activity decreased slightly. (Modified from Ref. 16.)

## A tentative model on brain substrates of music processing

It is not a particularly fruitful approach to look at brain substrates of music processing from such a solipsistic viewpoint, stating that presently around six billion different auditory biographies may produce the same number of different 'music centres' in the brain. Many aspects of auditory processing in general and music processing in particular are necessarily bound to fixed neuronal substrates common to all humans and mainly located in the superior temporal gyrus. As a sort of division of labour, hemispheric specialization may take place at a very early stage of auditory processing: the left superior temporal gyrus seems to be specialized for very rapid processes requiring high temporal resolution such as the identification of different phonemes, whereas the right superior temporal lobe is specialized for spectral analyses of sound.[17] Even at such an early stage, however, the brain substrates underlying auditory processing remain adaptive and subject to plastic changes as a consequence of conditioning or training.[18,19]

In Figure 22.2, a tentative model illustrating the interrelationship between the complexity of neuronal networks involved in music processing ($y$-axis) and the complexity of auditory processing ($x$-axis) is outlined. An additional dimension can be added on the $z$-axis, accounting for effects of acculturation. The complexity of neuronal networks increases

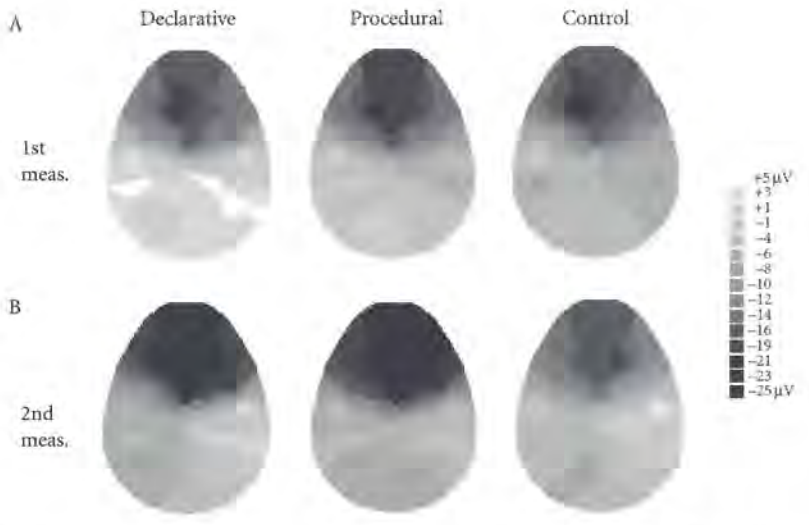



**Figure 22.1** Brain maps demonstrating cortical activation patterns (A) before and (B) after learning in the 'Declarative' learning group, in the 'Procedural' learning group, and in the 'Control' group. Group averages are displayed. Activation is *dark*, inactivation is *white* (see microvolt scale on the *right*). The brain diagrams are displayed as *top views*, frontal regions up, left hemisphere on the *left*, right hemisphere on the *right*. As can be recognized, declarative, mainly verbally mediated training leads to an increase in brain activity over the left frontal areas, whereas procedural, genuinely musical training produces an increase in activity over right frontal and bilateral parietooccipital regions. In controls, overall activity decreased slightly. (Modified from Ref. 16.)

## A tentative model on brain substrates of music processing

It is not a particularly fruitful approach to look at brain substrates of music processing from such a solipsistic viewpoint, stating that presently around six billion different auditory biographies may produce the same number of different 'music centres' in the brain. Many aspects of auditory processing in general and music processing in particular are necessarily bound to fixed neuronal substrates common to all humans and mainly located in the superior temporal gyrus. As a sort of division of labour, hemispheric specialization may take place at a very early stage of auditory processing: the left superior temporal gyrus seems to be specialized for very rapid processes requiring high temporal resolution such as the identification of different phonemes, whereas the right superior temporal lobe is specialized for spectral analyses of sound.[17] Even at such an early stage, however, the brain substrates underlying auditory processing remain adaptive and subject to plastic changes as a consequence of conditioning or training.[18,19]

In Figure 22.2, a tentative model illustrating the interrelationship between the complexity of neuronal networks involved in music processing ($y$-axis) and the complexity of auditory processing ($x$-axis) is outlined. An additional dimension can be added on the $z$-axis, accounting for effects of acculturation. The complexity of neuronal networks increases

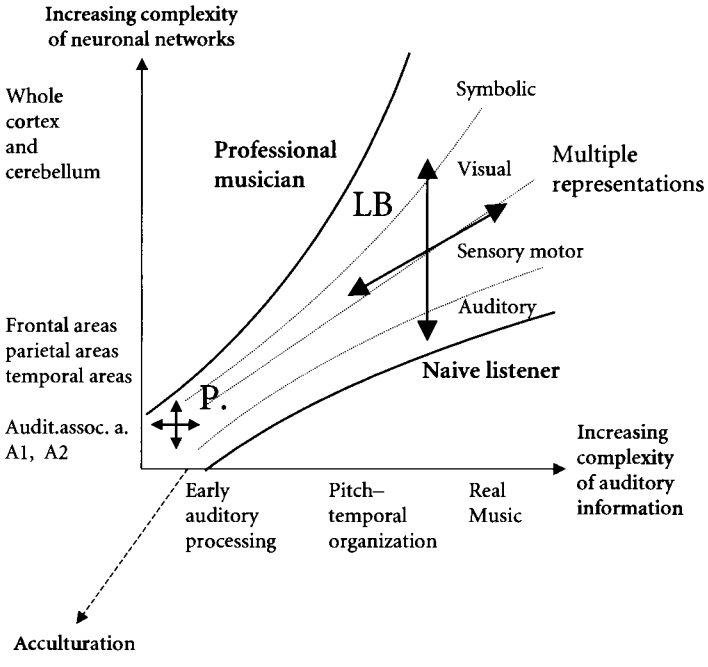



Figure 22.2 A simplified model demonstrating the interdependency between increasing complexity of auditory information processing (*x*-axis) and increasing complexity of neuronal networks (*y*-axis). *P.* = plasticity, *L.B.* = learning biography. *A1*, *A2*, and *audit. assoc. a* stand for primary auditory cortex, secondary auditory cortex, and auditory association cortices. For further explanation, see text.

with the complexity of processing demands. Training and practice add additional mental representations of music that rely on different brain substrates. Therefore, professionals presumably use larger and more complex neuronal networks during music processing than do nonprofessionals. The actual network engaged in a defined processing task is not fixed, but is subject to short- and long-term plastic changes, allowing at early processing stages for adaptations and compensations in case of damage to certain brain substrates. For example, a bilateral lesion of the primary auditory cortices with complete cortical deafness may be compensated for.[20] The vertical axis of the small cross '*P*'. (for plasticity) on the left symbolizes such a reciprocal replacement of brain functions by other brain structures. The horizontal axis relates to the fact that listeners may add or reduce complexity of auditory processing by adapting their listening strategies. When entering more advanced processing stages, including many neuropsychological laboratory experiments, the learning biography, determining the multiplicity of different auditory representations may influence the actual network used for music processing. According to actual demands or as a reaction to brain lesions, processing strategies may lead to a simplified or to a more complex way to listen to the stimuli or music, respectively. Learning biography (*LB*) and auditory strategies are symbolized by the larger cross on the right in the diagram. For the sake of clarity, other

variables influencing brain activation patterns during music processing have been omitted: guidance of attention, emotions, working memory, procedural and explicit short- and long-term memory systems, and effects of acculturation play an important role, rendering the brain substrates of music processing still more complex.

In summary, as soon as we consider 'real music' apart from laboratory experiments, we have to expect individually formed and quickly adaptive brain substrates, including widely distributed neuronal networks in both hemispheres. In our laboratories, we are just beginning to face the enormous challenges linked to the clarification of rules determining this puzzling variety of findings, determining the complexity and transitoriness of neuronal interactions during music processing. Two hundred years ago, the German poet Johann Wolfgang von Goethe recognized and expressed this problem in a masterful way in the 'scene in the laboratory':

> *Homunculus*: That is the way that things are apt to take:
> The cosmos scarce will compass nature's kind
> But man's creations need to be confined. (Johann Wolfgang von Goethe, *Faust*, Part II, Act II. Scene in the Laboratory. Translated by Philip Wayne)

## References


1. **Henschen, S. E.** (1920) Über Amusie. In *Klinische und anatomische Beiträge zur Pathologie des Gehirns*, vol. 5. Stockholm: Nordiska Bokhandeln, pp. 137–213.
2. **Uvstedt, H. J.** (1937) The method of examination in amusia. *Acta Psychiatr. Neurol.* 12, 447–55.
3. **Bever, T. G. and R. I. Chiarello** (1974) Cerebral dominance in musicians and non-musicians. *Science* 185, 537–40.
4. **Altenmüller, E.** (1989) Cortical DC-potentials as electrophysiological correlates of hemispheric dominance of higher cognitive functions. *Int. J. Neurosci.* 47, 1–14.
5. **Fodor, G. A.** (1983) *The Modularity of Mind.* Cambridge, MA: MIT Press.
6. **Gardner, H.** (1983) *Frames of Mind. The Theory of Multiple Intelligences.* New York: Basic Books.
7. **Peretz, I.** (1990) Processing of local and global musical information by unilateral brain-damaged patients. *Brain* 113, 1185–205.
8. **Peretz, I.** (1993) Auditory agnosia: a functional analysis. In S. McAdams and E. Bigand (eds) *Thinking in Sound: The Cognitive Psychology of Human Audition.* Oxford: Clarendon Press, pp. 199–230.
9. **Peretz, I., R. Kolinsky, M. Tram, *et al.*** (1994) Functional dissociations following bilateral lesions of auditory cortex. *Brain* 117, 1283–301.
10. **Peretz, I. and L. Gagnon** (1999) Dissociation between recognition and emotional judgements for melodies. *Neurocase* 5, 21–30.
11. **Zatorre, R. J. and S. Samson** (1991) Role of right temporal neocortex in retention of pitch in auditory short-term memory. *Brain* 114, 2403–17.
12. **Zatorre, R. J., A. C. Evans, and E. Meyer** (1994) Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 14, 1908–19.
13. **Liégeois-Chauvel, C., I. Peretz, M. Babaï, *et al.*** (1998) Contribution of different cortical areas in the temporal lobes to music processing. *Brain* 121, 1853–67.

14. **Heinze, H. J., G. R. Mangun, W. Burchert, *et al.*** (1994) Combined spatial and temporal imaging of brain activity during visual selective attention in humans. *Nature* 372, 543–6.
15. **Schuppert, M., T. F. Münte, B. M. Wieringa, *et al.*** (2000) Receptive amusia: evidence for cross-hemispheric neural networks underlying music processing strategies. *Brain* 123, 546–59.
16. **Altenmüller, E., W. Gruhn, D. Parlitz, *et al.*** (1997) Music learning produces changes in brain activation patterns: a longitudinal DC-EEG-study. *Int. J. Arts Med.* 5, 28–34.
17. **Zatorre, R. J.** (2001) Neural specialization for tonal processing. In R. Zatorre and I. Peretz (eds) *The biological foundations of music.* Annals of the New York Academy of Sciences, Vol. 290, pp. 193–210.
18. **Pantev, C., A. Wollbrink, L. E. Roberts, *et al.*** (1999) Short-term plasticity of the human auditory cortex. *Brain Res.* 842, 192–9.
19. **Weinberger, N. M. and J. S. Bakin** (1998) Learning-induced physiological memory in adult primary auditory cortex: receptive field plasticity, model and mechanisms. *Audiol. Neuro-Otol.* 3, 145–67.
20. **Tramo, M. J., J. J. Bharucha, and F. Musiek** (1989) Music perception and cognition following bilateral lesions of auditory cortex. *J. Cognit. Neurosci.* 2, 195–212.

PART V

# MUSICAL EXPERTISE/BRAIN PLASTICITY

CHAPTER 23

# FUNCTIONAL ORGANIZATION AND PLASTICITY OF AUDITORY CORTEX

JOSEF P. RAUSCHECKER

Remarkable as our cerebral cortex is in many respects, its ability to self-organize in response to extrinsic stimuli is perhaps the most astonishing property of all. The three chapters in this section are all impressive testimony to this ability. The effects of musical experience can be found on various levels: in the gross-anatomical differences between professional musicians with absolute pitch[1] (Chapter 24, this volume) to the subtle functional differences after musical training found with magnetoencephalography[2] (Chapter 25, this volume) or transcranial magnetic stimulation[3] (Chapter 26, this volume). Of course, even the ability to learn and memorize a simple tune is an expression of the brain's ability to change with musical experience. In this case, the changes would have to be sought in even finer modifications of synaptic strength in distributed cortical networks.

## Basic principles of cortical plasticity

The tenet that learning and memory is based on changes of synaptic efficacy was propagated, among others, by Hebb.[4] More specifically, he postulated that the conjunction of pre- and postsynaptic activity contributes to a strengthening of synaptic connections. Hebb's ideas were applied to many different areas of neuroplasticity, but perhaps most prominently so in the field of visual cortical plasticity. Developing visual cortex became a model system for the study of neural plasticity when Wiesel and Hubel had established many of the fundamental facts, such as ocular dominance plasticity after monocular lid closure.[5] The development of orientation selectivity was also proposed early on as an example of experience-dependent plasticity at the single-neuron level[6] and was later shown to follow Hebbian rules.[7,8]

## Improved auditory abilities in the blind

Deprivation in one sensory modality, such as vision, can also lead to dramatic reorganization in other modalities, such as hearing or touch.[9] The mechanisms of cross-modal plasticity of

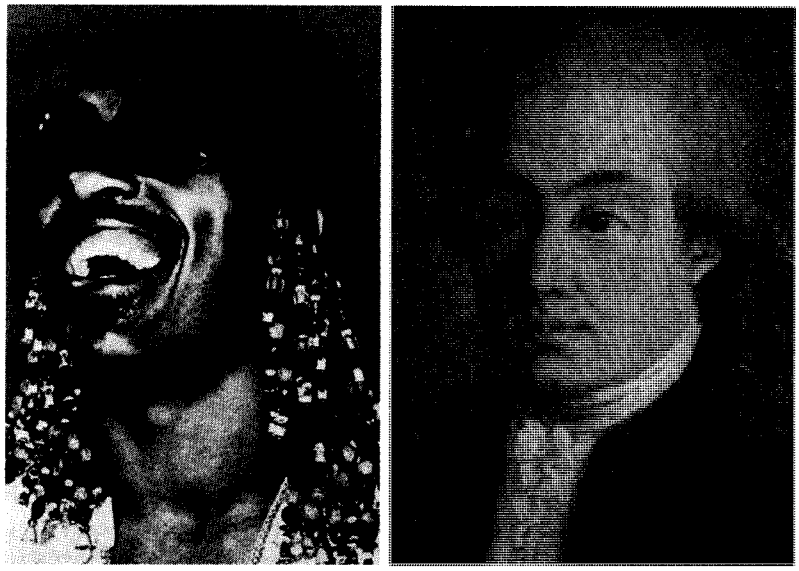


**Figure 23.1** Anecdotal evidence reports superior musical abilities in blind individuals. *Left*: Stevie Wonder. From: *The Rolling Stone. Illustrated History of Rock and Roll.* A. DeCurtis, J. Henke, H. George-Warren (eds) New York: Random House, p. 294. *Right*: Wolfgang Amadeus Mozart, though not blind, possibly suffered from strabismic amblyopta, reducing vision in one or both eyes by an unknown amount. Portrait attributed to Kymli of Mannheim, Germany, 1783 (held in a private collection). (See Plate 17 in colour section.)

the auditory cortex appear to be the same as those of visual cortical plasticity.[10] Compensatory improvement of auditory abilities in the blind had long been postulated, claiming, among other things, improved musical abilities (Figure 23.1).

While it is often difficult to quantitatively assess musical abilities, other auditory perceptual capacities can be tested more easily. An ability of blind people to localize sounds with greater precision has been found in a number of recent studies.[11,12] These studies followed on the footsteps of work in visually deprived cats[13] and ferrets[14] that had also demonstrated improved sound localization abilities in these animals. Animal studies can additionally be used to find the neural bases of these behavioural improvements. In blind cats, visually deprived by binocular lid suture from birth, an area in the parietal cortex (the anterior ectosylvian visual area [AEV]) that is normally innervated by visual fibres originating in extrastriate cortex and visual pulvinar[15] is completely taken over by auditory and somatosensory inputs.[16] At the single unit level, auditory cortical neurons show greater selectivity for the position of sounds in space.[17] Taken together, these findings suggest a greater sampling density of auditory space in the cerebral cortex as the basis for the greater precision in sound localization of blind individuals.[9] Furthermore, the expansion of auditory territory in the cerebral cortex would appear to enable other refinements of auditory perception as well.

Modern neuroimaging techniques, such as positron emission tomography (PET), can be used to look for such crossmodal map changes directly in blind humans. During localization



**Figure 23.1** Anecdotal evidence reports superior musical abilities in blind individuals. *Left*: Stevie Wonder. From: *The Rolling Stone. Illustrated History of Rock and Roll.* A. DeCurtis, J. Henke, H. George-Warren (eds) New York: Random House, p. 294. *Right*: Wolfgang Amadeus Mozart, though not blind, possibly suffered from strabismic amblyopta, reducing vision in one or both eyes by an unknown amount. Portrait attributed to Kymli of Mannheim, Germany, 1783 (held in a private collection). (See Plate 17 in colour section.)

the auditory cortex appear to be the same as those of visual cortical plasticity.[10] Compensatory improvement of auditory abilities in the blind had long been postulated, claiming, among other things, improved musical abilities (Figure 23.1).

While it is often difficult to quantitatively assess musical abilities, other auditory perceptual capacities can be tested more easily. An ability of blind people to localize sounds with greater precision has been found in a number of recent studies.[11,12] These studies followed on the footsteps of work in visually deprived cats[13] and ferrets[14] that had also demonstrated improved sound localization abilities in these animals. Animal studies can additionally be used to find the neural bases of these behavioural improvements. In blind cats, visually deprived by binocular lid suture from birth, an area in the parietal cortex (the anterior ectosylvian visual area [AEV]) that is normally innervated by visual fibres originating in extrastriate cortex and visual pulvinar[15] is completely taken over by auditory and somatosensory inputs.[16] At the single unit level, auditory cortical neurons show greater selectivity for the position of sounds in space.[17] Taken together, these findings suggest a greater sampling density of auditory space in the cerebral cortex as the basis for the greater precision in sound localization of blind individuals.[9] Furthermore, the expansion of auditory territory in the cerebral cortex would appear to enable other refinements of auditory perception as well.

Modern neuroimaging techniques, such as positron emission tomography (PET), can be used to look for such crossmodal map changes directly in blind humans. During localization

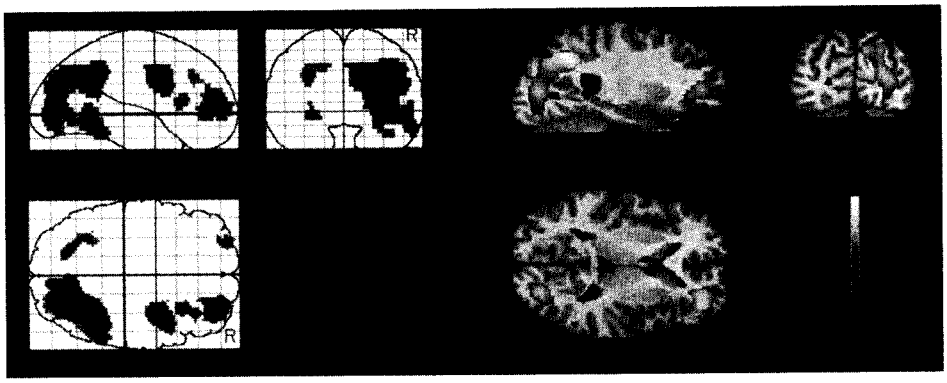


**Figure 23.2** Results of PET scanning in congenitally blind subjects during localization of sounds presented via headphones in virtual auditory space (data from Ref. 19). Areas in the occipital cortex that are driven by visual stimuli in sighted individuals are now activated during sound localization. Note that the right hemisphere seems to profit more from the expansion of auditory into visual territory. The 'rewiring' seems to be mediated by the parietal cortex, consistent with studies on visually deprived animals.[9] (See Plate 18 in colour section.)

of sounds in virtual auditory space via headphones, sighted humans normally activate parts of their posterior parietal cortex in an area separate from known visual areas involved in spatial analysis.[18] Blind humans, in addition to activating a much more extended area in the same region, also activate large parts of the occipital lobe, which is normally used for seeing[19] (Figure 23.2). The results of cross-correlating activities in different brain regions suggest that the auditory input to the occipital cortex is relayed via posterior parietal cortex, but a sustenance or strengthening of direct connections from auditory to visual cortex, as they occur especially in young animals,[20–22] cannot be excluded. These surprising results warrant further testing in animals with neurophysiological and neuroanatomical techniques.

## 'What' and 'where' in auditory cortex

Another fundamental tenet of Hebbian models of long-term memory is that information gets stored in the same places where it is processed. We are far from understanding how lengthy sequences of musical melodies or even more complex pieces of music are stored in our brains, although it is clear that music is a powerful stimulus that activates large parts of auditory cortex in both hemispheres (Figure 23.3). What we are slowly beginning to understand is that complex sounds are first broken down into their frequency components in the periphery of the auditory system and in primary auditory cortex. Then, at later stages of cortical processing this tonal information is integrated into more complex response properties,[23,24] and single units respond to more and more complex contents. The hierarchical processing of such 'what' information seems to take place in more antero-lateral portions of the superior temporal gyrus (STG),[25–28] all the way to the temporal pole. By

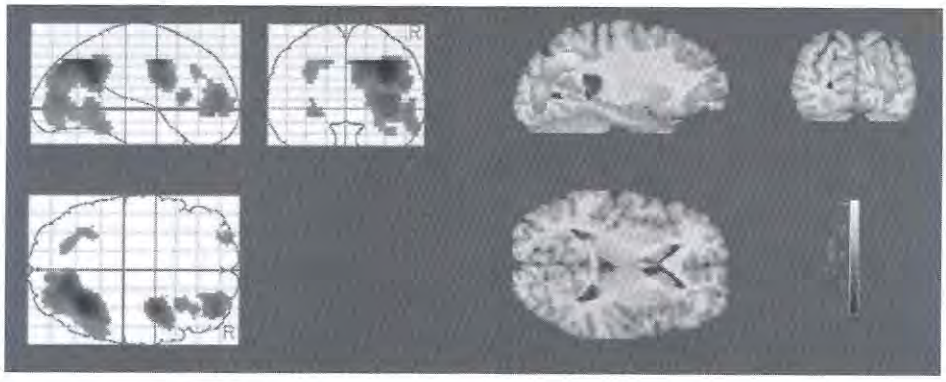


**Figure 23.2** Results of PET scanning in congenitally blind subjects during localization of sounds presented via headphones in virtual auditory space (data from Ref. 19). Areas in the occipital cortex that are driven by visual stimuli in sighted individuals are now activated during sound localization. Note that the right hemisphere seems to profit more from the expansion of auditory into visual territory. The 'rewiring' seems to be mediated by the parietal cortex, consistent with studies on visually deprived animals.[9] (See Plate 18 in colour section.)

of sounds in virtual auditory space via headphones, sighted humans normally activate parts of their posterior parietal cortex in an area separate from known visual areas involved in spatial analysis.[18] Blind humans, in addition to activating a much more extended area in the same region, also activate large parts of the occipital lobe, which is normally used for seeing[19] (Figure 23.2). The results of cross-correlating activities in different brain regions suggest that the auditory input to the occipital cortex is relayed via posterior parietal cortex, but a sustenance or strengthening of direct connections from auditory to visual cortex, as they occur especially in young animals,[20–22] cannot be excluded. These surprising results warrant further testing in animals with neurophysiological and neuroanatomical techniques.

## 'What' and 'where' in auditory cortex

Another fundamental tenet of Hebbian models of long-term memory is that information gets stored in the same places where it is processed. We are far from understanding how lengthy sequences of musical melodies or even more complex pieces of music are stored in our brains, although it is clear that music is a powerful stimulus that activates large parts of auditory cortex in both hemispheres (Figure 23.3). What we are slowly beginning to understand is that complex sounds are first broken down into their frequency components in the periphery of the auditory system and in primary auditory cortex. Then, at later stages of cortical processing this tonal information is integrated into more complex response properties,[23,24] and single units respond to more and more complex contents. The hierarchical processing of such 'what' information seems to take place in more antero-lateral portions of the superior temporal gyrus (STG),[25–28] all the way to the temporal pole. By

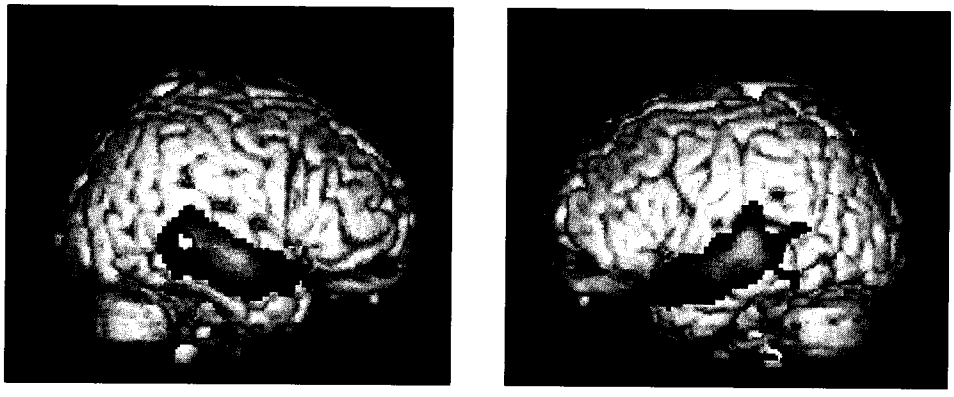


Figure 23.3 Auditory cortical activation while listening to music. Averaged group results of functional magnetic resonance imaging are shown. While the processing of music by the right cerebral hemisphere has often been emphasized[32] (as opposed to language, which is thought to be processed primarily on the left), it is now generally accepted that different aspects of music are processed by *both* hemispheres.[33] (See Plate 19 in colour section.)

contrast, location and movement of sounds are determined in more posterior aspects of the STG.[27,29,30] The latter regions project to the posterior parietal cortex,[31] also known to be involved in auditory spatial processing, as mentioned above.[18,34,35] Anterior STG projects to orbitofrontal cortex, which plays a role in working memory for objects,[36] whereas the caudal STG projects to dorsolateral prefrontal cortex, directly and via parietal cortex, again highly consistent with its role in spatial processing.[37]

## Musical learning and auditory imagery

If the anterior STG and orbito-frontal cortex are responsible for the processing of complex auditory 'images', patterns, or objects, they should also be the storage houses of musical memories. Indeed, when we try to imagine music, we activate parts of the auditory cortex that is anterior to Heschl's region (primary auditory cortex) plus areas in frontal cortex.[38] Even when we simply anticipate familiar music, inferior frontal cortex lights up.[39] This may correspond to the involvement of left frontal regions in the retrieval of complex memories, as predicted in a model put forward by Tulving and coworkers.[40] Other brain regions, such as the cerebellum and the anterior cingulate cortex, are also active during anticipatory imagery of music (Figure 23.4). This justifies the idea that mental rehearsal, by activating some of the same brain regions as during a real performance, helps to practice and memorize music, even though we do not actually play the instrument at that time. It is also well known that professional conductors are not only able to conduct hundreds of pieces from memory but also rehearse them mentally, stressing again the partial equivalence of brain activation through attention and mental exercise.

'What' and 'where' information in music come together for conductors of large orchestras. In these situations, the content of what is being played is of utmost importance, but

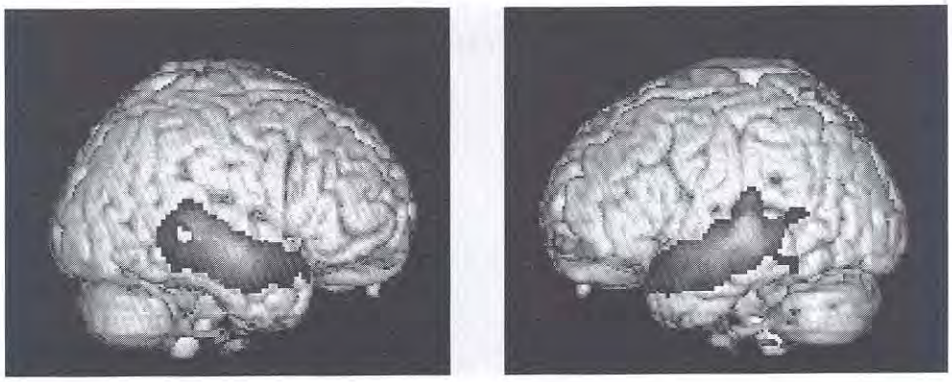


**Figure 23.3** Auditory cortical activation while listening to music. Averaged group results of functional magnetic resonance imaging are shown. While the processing of music by the right cerebral hemisphere has often been emphasized[32] (as opposed to language, which is thought to be processed primarily on the left), it is now generally accepted that different aspects of music are processed by *both* hemispheres.[33] (See Plate 19 in colour section.)

contrast, location and movement of sounds are determined in more posterior aspects of the STG.[27,29,30] The latter regions project to the posterior parietal cortex,[31] also known to be involved in auditory spatial processing, as mentioned above.[18,34,35] Anterior STG projects to orbitofrontal cortex, which plays a role in working memory for objects,[36] whereas the caudal STG projects to dorsolateral prefrontal cortex, directly and via parietal cortex, again highly consistent with its role in spatial processing.[37]

## Musical learning and auditory imagery

If the anterior STG and orbito-frontal cortex are responsible for the processing of complex auditory 'images', patterns, or objects, they should also be the storage houses of musical memories. Indeed, when we try to imagine music, we activate parts of the auditory cortex that is anterior to Heschl's region (primary auditory cortex) plus areas in frontal cortex.[38] Even when we simply anticipate familiar music, inferior frontal cortex lights up.[39] This may correspond to the involvement of left frontal regions in the retrieval of complex memories, as predicted in a model put forward by Tulving and coworkers.[40] Other brain regions, such as the cerebellum and the anterior cingulate cortex, are also active during anticipatory imagery of music (Figure 23.4). This justifies the idea that mental rehearsal, by activating some of the same brain regions as during a real performance, helps to practice and memorize music, even though we do not actually play the instrument at that time. It is also well known that professional conductors are not only able to conduct hundreds of pieces from memory but also rehearse them mentally, stressing again the partial equivalence of brain activation through attention and mental exercise.

'What' and 'where' information in music come together for conductors of large orchestras. In these situations, the content of what is being played is of utmost importance, but



Figure 23.4 Results of functional magnetic resonance imaging during anticipatory imagery of music. The axial sections from individual subjects are displayed in radiological coordinates showing the left side of the brain on the right and vice versa. Imagery leads to activation of anterior regions of auditory cortex in the right hemisphere (in conjunction with right inferior frontal regions; panel on the left of the figure). This emphasizes the role of sensory cortices in mental imagery. However, brain regions outside auditory cortex, such as the cerebellum (left cerebellar hemisphere; middle panel) and the anterior cingulate bilaterally (right panel), are also activated, demonstrating their role in cognitive processing. (See Plate 20 in colour section.)

the location of a particular instrument on the stage helps the conductor to pinpoint the source of errors. Training this ability obviously helps, as has been found in a recent study with event-related brain potentials.[41] Conductors show greater differentiation between positions in peripheral auditory space than nonmusicians or pianists. As attention has been shown to increase brain activation in nonprimary auditory cortex,[42] attentional modulation during rehearsal would be expected to sharpen conductors' skills by modulating the activity of brain regions encoding auditory space. This situation is not unlike that of blind individuals focusing completely on nonvisual events, although in this situation attention only provides the initial vehicle that leads to permanent change. At the single-neuron level, cells have been found in area CL of the monkey's nonprimary auditory cortex that are selective to both the content and the location of complex sounds.[29] Such cells could well provide the neural substrate for the above effects.

## Sensitive periods of auditory cortical plasticity

Auditory cortical plasticity can be demonstrated in everyday processes as musical learning, and it becomes even more overtly obvious in the reorganization of auditory cortex across areal borders after early blindness. Age certainly plays an important role in these plastic changes, as the age of onset of musical training as well as the age of onset of blindness have both been shown to be critical for the extent of reorganization. Learning a musical instrument



**Figure 23.4** Results of functional magnetic resonance imaging during anticipatory imagery of music. The axial sections from individual subjects are displayed in radiological coordinates showing the left side of the brain on the right and vice versa. Imagery leads to activation of anterior regions of auditory cortex in the right hemisphere (in conjunction with right inferior frontal regions; panel on the left of the figure). This emphasizes the role of sensory cortices in mental imagery. However, brain regions outside auditory cortex, such as the cerebellum (left cerebellar hemisphere; middle panel) and the anterior cingulate bilaterally (right panel), are also activated, demonstrating their role in cognitive processing. (See Plate 20 in colour section.)

the location of a particular instrument on the stage helps the conductor to pinpoint the source of errors. Training this ability obviously helps, as has been found in a recent study with event-related brain potentials.[41] Conductors show greater differentiation between positions in peripheral auditory space than nonmusicians or pianists. As attention has been shown to increase brain activation in nonprimary auditory cortex,[42] attentional modulation during rehearsal would be expected to sharpen conductors' skills by modulating the activity of brain regions encoding auditory space. This situation is not unlike that of blind individuals focusing completely on nonvisual events, although in this situation attention only provides the initial vehicle that leads to permanent change. At the single-neuron level, cells have been found in area CL of the monkey's nonprimary auditory cortex that are selective to both the content and the location of complex sounds.[29] Such cells could well provide the neural substrate for the above effects.

## Sensitive periods of auditory cortical plasticity

Auditory cortical plasticity can be demonstrated in everyday processes as musical learning, and it becomes even more overtly obvious in the reorganization of auditory cortex across areal borders after early blindness. Age certainly plays an important role in these plastic changes, as the age of onset of musical training as well as the age of onset of blindness have both been shown to be critical for the extent of reorganization. Learning a musical instrument

is easier during childhood not only because our motor abilities are more malleable during these early phases, but also because our auditory cortex has sensitive periods for self-organization. This does not mean that we cannot acquire these abilities any longer later in life, but an equivalent outcome will require greater effort. Similar sensitive periods apply to language learning: Learning a second language after the age of eight is definitely still possible, but a foreign accent will usually remain.[43] Again, this effect may have its ultimate cause in a decreasing malleability with age of recipient auditory cortical structures for auditory perception and discrimination.[44]

## Conclusions

The connection between brain activation and musical perception is becoming clearer as more and more refined techniques become available to image neural activity. Brain imaging can be performed while subjects are listening to music or imagining it. One can measure the immediate correlates of music as well as its effects on brain development and later brain organization. And, although animals do not have music (just as, in a strict sense, they do not speak), one can still get a grasp of the exact neural mechanisms of music perception by measuring the responses of single neurons in animal studies to various types of complex sounds. An analysis, for instance, of how sound sequences are processed and stored by neural networks would go a long way towards understanding the neural basis of music perception. Other aspects, such as timbre and rhythm perception, could equally be tackled by simpler paradigms in animal models.

Cortical plasticity has to be considered in at least two ways in the context of studies on music and the brain. First, early experience with the world of sound will inevitably shape our brain in an indelible fashion. Enhanced experience with sound, as it occurs with musical training and, inevitably, in blind individuals, can profoundly alter the circuitry of the part of the brain that processes complex sounds in general and music in particular. As we experience these brain changes resulting from cortical plasticity from studies in both humans and animals, we learn to appreciate how expanded areas of the brain lead to enhanced perceptual abilities in the corresponding senses.

## Acknowledgements

Supported in part by a grant from the National Institute of Deafness and Other Communication Disorders (R01-DC03489).

## References

1. **Schlaug, G., L. Jäncke, Y. Huang,** and **H. Steinmetz** (1995) In vivo evidence of structural brain asymmetry in musicians. *Science* 267, 699–701.
2. **Pantev, C., R. Oostenveld, A. Engelien, B. Ross, L. E. Roberts,** and **M. Hoke** (1998) Increased auditory cortical representation in musicians. *Nature* 392, 811–14.

3. Pascual-Leone, A., D. Bartres-Faz, and J. P. Keenan (1999) Transcranial magnetic stimulation: studying the brain-behaviour relationship by induction of 'virtual lesions'. *Philos. Trans. Roy. Soc. Lond. B. Biol. Sci.* 354, 1229–38.
4. Hebb, D. O. (1949) *The Organization of Behavior.* New York: Wiley.
5. Wiesel, T. N. and D. H. Hubel (1963) single-cell responses in striate cortex of kittens deprived of vision in one eye. *J. Neurophysiol.* 26, 1003–17.
6. Blakemore, C. and G. F. Cooper (1970) Development of the brain depends on the visual environment. *Nature* 228, 477–8.
7. Rauschecker, J. P. (1991) Mechanisms of visual plasticity: hebb synapses, NMDA receptors and beyond. *Physiol. Rev.* 71, 587–615.
8. Singer, W. (1995) Development and plasticity of cortical processing architectures. *Science* 270, 758–64.
9. Rauschecker, J. P. (1995) Compensatory plasticity and sensory substitution in the cerebral cortex. *Trends in Neurosci.* 18, 36–43.
10. Rauschecker, J. P. (1999) Auditory cortical plasticity: a comparison with other sensory systems. *Trends in Neurosci.* 22, 74–80.
11. Lessard, N., M. Pare, F. Lepore, and M. Lassonde (1998) Early-blind human subjects localize sound sources better than sighted subjects. *Nature* 395, 278–80.
12. Röder, B., W. A. S. Teder-Salejärvi, F. Rösler, S. A. Hillyard, and H. J. Neville (1999) Improved auditory spatial tuning in blind humans. *Nature* 400, 162–6.
13. Rauschecker, J. P. and U. Kniepert (1994) Auditory localization behaviour in visually deprived cats. *Eur. J. Neurosci.* 6, 149–60.
14. King, A. J. and C. Parsons (1999) Improved auditory spatial acuity in visually deprived ferrets. *Eur. J. Neurosci.* 11, 3945–56.
15. Norita, M., L. Mucke, G. Benedek, B. Albowitz, Y. Katoh, and O. D. Creutzfedlt (1986) Connections of the anterior ectosylvian visual area (AEV). *Exp. Brain Res.* 62, 225–40.
16. Rauschecker, J. P. and M. Korte (1993) Auditory compensation for early blindness in cat cerebral cortex. *J. Neurosci.* 13, 4538–48.
17. Korte, M. and J. P. Rauschecker (1993) Auditory spatial tuning of cortical neurons is sharpened in cats with early blindness. *J. Neurophysiol.* 70, 1717–21.
18. Bushara, K. O., R. A. Weeks, K. Ishii, M.-J. Catalan, B. Tian, J. P. Rauschecker, and M. Hallett (1999) Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nat. Neurosci.* 2, 759–66.
19. Weeks, R., B. Horwitz, A. Aziz-Sultan, B. Tian, C. M. Wessinger, L. Cohen, M. Hallett, and J. P. Rauschecker (2000) A positron emission tomographic study of auditory localization in the congenially blind. *J. Neurosci.* 20, 2664–72.
20. Clarke, S. and G. M. Innocenti (1990) Auditory neurons with transitory axons to visual areas form short permanent projections. *Eur. J. Neurosci.* 2, 227–42.
21. Falchier, A., L. Renaud, P. Barone, and H. Kennedy (2001) Extensive projections from the primary auditory cortex and polysensory area STP to peripheral area V1 in the macaque. *Soc. for Neurosci. Abstracts* 31, 511.21.
22. Rockland, K. S. and Ojima (2001) Calcarine area V1 as a multimodal convergence area. *Soc. for Neurosci. Abstracts* 31, 511.20.
23. Rauschecker, J. P. (1998) Cortical processing of complex sounds. *Curr. Opin. Neurobiol.* 8, 516–21.

24. **Wessinger, C. M., J. Van Meter, B. Tian, J. Van Lare, J. Pekar, and J. P. Rauschecker** (2001) Hierarchical organization of human auditory cortex revealed by functional magnetic resonance imaging. *J. Cognitive Neurosci.* 13, 1–7.

25. **Belin, P., R. J. Zatorre, P. Lafaille, P. Ahad, and B. Pike** (2000) Voice-selective areas in human auditory cortex. *Nature* 403, 309–12.

26. **Binder, J. R., J. A. Frost, T. A. Hammeke, P. S. Bellgowan, J. A. Springer, J. N. Kaufman, and E. T. Possing** (2000) Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex* 10, 512–28.

27. **Rauschecker, J. P. and B. Tian** (2000) Mechanisms and streams for processing of "what" and "where" in auditory cortex. *PNAS* 97, 11800–6.

28. **Scott, S. K., C. C. Blank, S. Rosen, and R. J. Wise** (2000) Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123 (12), 2400–6.

29. **Tian, B., D. Reser, A. Durham, A. Kustov, and J. P. Rauschecker** (2001) Functional specialization in rhesus monkey auditory cortex. *Science* 292, 290–3.

30. **Zatorre, R. J., A. C. Evans, and E. Meyer** (1994) Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 14, 1908–19.

31. **Peretz, I.** (1990) Processing of local and global musical information in unilateral brain damaged patients. *Brain* 13, 1185–205.

32. **Warren, J. D., B. A. Zielinski, G. G. Green, J. P. Rauschecker, and T. D. Griffiths** (2002) Perception of sound-source motion by the human brain. *Neuron* 34, 139–48.

33. **Lewis, J. W. and D. C. Van Essen** (2000) Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *J. Comp. Neurol.* 428, 112–37.

34. **Griffiths, T. D., G. Rees, A. Rees, G. G. Green, C. Witton, D. Rowe, C. Buchel, R. Turner, and R. S. Frackowiak** (1998) Right parietal cortex is involved in the perception of sound movement in humans. *Nat. Neurosci.* 1, 74–9.

35. **Weeks, R. A., A. Aziz-Sultan, K. O. Bushara, B. Tian, C. M. Wessinger, N. Dang, J. P. Rauschecker, and M. Hallett** (1999) A PET study of human auditory spatial processing. *Neurosci. Lett.* 262, 155–8.

36. **Goldman-Rakic, P. S.** (1996) The prefrontal landscape: implications of functional architecture for understanding human mentation and the central executive. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 351, 1445–53.

37. **Romanski, L. M., B. Tian, J. Fritz, M. Mishkin, P. S. Goldman-Rakic, and J. P. Rauschecker** (1999) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci.* 2, 1131–6.

38. **Halpern, A. R. and R. J. Zatorre** (1999) When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. *Cereb. Cortex* 9, 697–704.

39. **Van Lare, J. E., B. A. Zielinski, and J. P. Rauschecker** (1999) Anticipatory musical imagery: functional MRI studies of the human brain. In *Intl. Soc. for Sys. and Comparative Musicol.* Norway: Oslo.

40. **Tulving, E., S. Kapur, F. I. Craik, M. Moscovitch, and S. Houle** (1994) Hemispheric encoding/retrieval asymmetry in episodic memory: positron emission tomography findings. *Proc. Natl. Acad. Sci. USA* 91, 2016–20.

41. **Münte, T. F., C. Kohlmetz, W. Nager, and E. Altenmüller** (2001) Superior auditory spatial tuning in conductors. *Nature* 409, 580.

42. **Grady, C. L., J. W. VanMeter, J. M. Maisog, P. Pietrini, J. Krasuski,** and **J. P. Rauschecker** (1997) Attention-related modulation of activity in primary and secondary auditory cortex activation. *Neuroreport* 8, 2511–6.

43. **Newport, E. L.** (1990) *Cognitive Sci.* 14, 11–28.

44. **Kuhl, P. K., J. E. Andruski, I. A. Chistovich, L. A. Chistovich, E. V. Kozhevnikova, V. L. Ryskina, E. I. Stolyarova, U. Sundberg,** and **F. Lacerda** (1997) Cross-language analysis of phonetic units in language addressed to infants. *Science* 277, 684–6.

CHAPTER 24

# THE BRAIN OF MUSICIANS

GOTTFRIED SCHLAUG

## Introduction

Musicians perform complex physical and mental operations such as translating musical symbols into complex motor operations, performing independent movements of fingers and hands, remembering long musical phrases, improvising music, and identifying tones absolutely without a reference tone. The neural correlates of most of these musical operations are not fully understood yet.[1] Several studies have shown that musicians differ from nonmusicians in certain aspects of brain function and structure,[2–14] but questions remain as to whether 'musical' skills and functions are related to these functional and structural differences. The commencement of motor and auditory training for students of music starts at a time when the brain and its components are in a critical period of development. This early commencement combined with long-term training makes musicians ideal subjects in which to investigate whether the almost daily practice of auditory and motor skills over many years leads to functional and/or structural cerebral adaptations. On the one hand, the brain may be highly plastic, capable not only of changes in functional brain networks,[15,16] but also in structural components as a response to increased use.[17–19] On the other hand, the brains of musicians may be 'atypical' prior to training, both anatomically and functionally, and hence a special anatomy or brain function would be a prerequisite for musical skill acquisition rather than its consequence.

Several animal studies have shown that microstructural changes such as increases in number of synapses, glial cells number and volume, and capillary density occur after long-term motor exercises and motor learning in rats within the cerebellum and primary motor cortex.[20–23] More recently, studies have also shown evidence for neurogenesis, mainly in the hippocampus, in rats reared in a complex environment[24] or as a result of long-term motor activity (i.e. running in a running wheel) which was associated with behavioural advantages.[25] It is unknown whether the sum of these microstructural changes could amount to structural differences detectable on a macrostructural level. Only a few studies examined the results of presumed structural plasticity in the human brain.[26,27] The only study in humans besides our own studies on musicians that suggested plastic structural changes in response to environmental demands is a study by Maguire and colleagues.[27] In this study a larger posterior hippocampal volume was found in experienced taxi drivers who are under intense visual-spatial demands of driving a taxi in a large metropolitan region.

Functional reorganization of adult mammalian sensory and motor cortical representations after peripheral or central stimulation or as an adjustment after injury has been found in many different experimental animal models of brain plasticity.[28–31,19] Similar adaptive

changes in the cortical organization after skill learning and as an adjustment after injury have been found in humans using electrophysiological or neuroimaging methods.[32–40] These findings have advanced the understanding that functional properties of central nervous system neurons as well as the neural circuitry either within the same or different brain areas, are malleable and retain a significant degree of functional plasticity which could lead to microstructural changes. The term plasticity is a broad term that can mean an adjustment or adaptation of a sensory or motor system to environmental stimuli or performance requirements or a compensation of some cerebral structures for others that are impaired due to injury or deafferentiation.[15–17]

Three human studies are particularly relevant to our studies on structural adaptation which show differences in the sensorimotor representation maps as a response to skill learning and acquisition of new skills. Pascual-Leone *et al.*[35] showed that as subjects learned a five-finger exercise on the piano over the course of 5 days, the cortical representation area targeting the long finger flexor and extensor muscles enlarged. Karni *et al.*[38] showed that a few minutes of daily practice of a sequential finger opposition task induced large, incremental performance gains over a few weeks of training, which was associated with changes in cortical movement representation within the primary motor cortex. These authors argued that the changes in the primary motor cortex reflected the setting up of a task-specific motor processing routine while the subsequent changes indicated the consolidation of a motor program. Although several reports have stressed the rapid reversal of representational changes, other studies have found persistent representational changes in response to the early acquisition of fine sensorimotor skills such as having a larger sensory finger representation in the left hand of string players.[7]

## Structural brain differences between musicians and nonmusicians

In the search for a morphological substrate of musicianship, several cross-sectional studies comparing adult musicians with nonmusicians were done. A priori defined anatomical regions were selected based on the relevance for musical functions and on data derived from human developmental studies as well as animal experimental studies suggesting a high degree of plasticity.

### Subjects, neuroimaging methods, and analysis

All of our study participants were either righthanded, classically trained professional musicians or nonmusicians. A professional musician was defined as someone who was either enrolled as a student in a full time music program in music school or music conservatory or someone who was a graduate of a music program and had his or her main income derived from a professional career in music. The majority of musicians in our studies were keyboard players. We also had a subgroup of string players who were mostly keyboard players as well. For our studies, we contrasted different groups of musicians with groups of nonmusicians who were matched for age, gender, and handedness. A non-musician was defined as someone who did not have any formal training in music and never played a musical instrument for any reasonable period of time. Hand preferences was typically assessed with the 12-item Annett questionnaire.[41] Consistent right-handedness corresponded to

performance of all 12 tasks with the right hand with up to two 'either hand' preferences being acceptable.[8,42] In addition to hand preference, we also assessed distal hand motor skills using an index finger tapping test.[43] A laterality coefficient was calculated according to the following formula (R−L)/(R+L), where L (left) and R (right) were the number of finger taps with either hand within 20 s. A subgroup of our subjects also underwent a battery of behavioural and cognitive tests including subsets taken from the Wechsler Adult Intelligence Scale tests (WAIS-R, Ref. 44) as well as other test of verbal intelligence (Shipley-Hartford Scale) and visual-spatial tests such as a mental rotation task and the block design test from the WAIS-R inventory as well as other memory tests and tests of attention.

Neuroimaging studies were performed using a whole-body Siemens 1.5 T MR machine acquiring volumetric T1-weighted brain images with a typical voxel resolution of $1^3$ mm. Functional MR imaging studies were done using a T2*-weighted MR sequence that was sensitive to changes in the local concentrations of oxy- and deoxyhemoglobin. Changes in the relative concentrations of oxy-/deoxyhemoglobin are indirect markers of regional blood flow changes that again are indirect markers of changes in regional neuronal synaptic activity. Analysis of fMRI data was done using SPM99, AFNI, and other custom made software. All morphometric analyses were done using custom made software running on HP workstations and mostly implemented in the Advanced Visual System (AVS) image analysis package. Morphometric studies were typically done by two independent investigators who were blinded to the identity as well as to the hemisphere of each brain. Interobserver correlations for the anatomical measurements were usually higher than 0.9.

### Corpus callosum

There is evidence that the functional and possibly structural maturation of the corpus callosum extends into late childhood and early adolescence and coincides with the termination of its myelination cycle.[45,46] In comparing 30 professional musicians and 30 age-, sex-, and handedness-matched non-musician controls, we found that the anterior half of the corpus callosum was significantly larger in musicians (Figure 24.1; Table 24.1), particularly in musicians who started training early (<7 years old) compared to those who started late

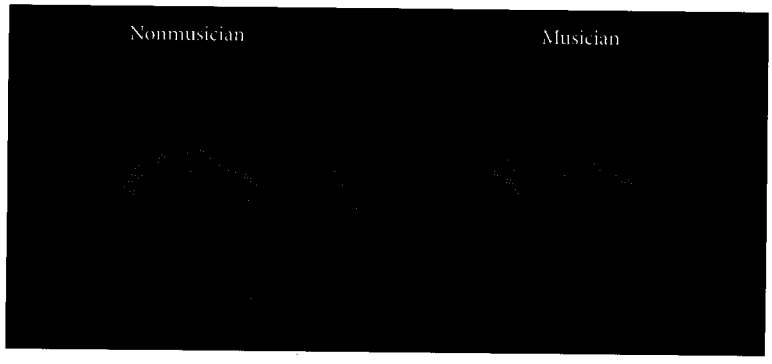



**Figure 24.1** *Two MR images showing the mid-sagittal area of the corpus callosum in a musician and a nonmusician. The anterior part of the corpus callosum was significantly larger in musicians compared to nonmusicians.

performance of all 12 tasks with the right hand with up to two 'either hand' preferences being acceptable.[8,42] In addition to hand preference, we also assessed distal hand motor skills using an index finger tapping test.[43] A laterality coefficient was calculated according to the following formula (R−L)/(R+L), where L (left) and R (right) were the number of finger taps with either hand within 20 s. A subgroup of our subjects also underwent a battery of behavioural and cognitive tests including subsets taken from the Wechsler Adult Intelligence Scale tests (WAIS-R, Ref. 44) as well as other test of verbal intelligence (Shipley-Hartford Scale) and visual-spatial tests such as a mental rotation task and the block design test from the WAIS-R inventory as well as other memory tests and tests of attention.

Neuroimaging studies were performed using a whole-body Siemens 1.5 T MR machine acquiring volumetric T1-weighted brain images with a typical voxel resolution of $1^3$ mm. Functional MR imaging studies were done using a T2*-weighted MR sequence that was sensitive to changes in the local concentrations of oxy- and deoxyhemoglobin. Changes in the relative concentrations of oxy-/deoxyhemoglobin are indirect markers of regional blood flow changes that again are indirect markers of changes in regional neuronal synaptic activity. Analysis of fMRI data was done using SPM99, AFNI, and other custom made software. All morphometric analyses were done using custom made software running on HP workstations and mostly implemented in the Advanced Visual System (AVS) image analysis package. Morphometric studies were typically done by two independent investigators who were blinded to the identity as well as to the hemisphere of each brain. Interobserver correlations for the anatomical measurements were usually higher than 0.9.

### Corpus callosum

There is evidence that the functional and possibly structural maturation of the corpus callosum extends into late childhood and early adolescence and coincides with the termination of its myelination cycle.[45,46] In comparing 30 professional musicians and 30 age-, sex-, and handedness-matched non-musician controls, we found that the anterior half of the corpus callosum was significantly larger in musicians (Figure 24.1; Table 24.1), particularly in musicians who started training early (<7 years old) compared to those who started late

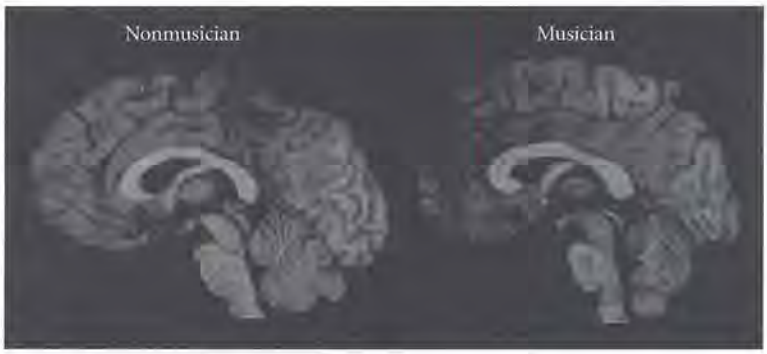



Figure 24.1 *Two MR images showing the mid-sagittal area of the corpus callosum in a musician and a nonmusician. The anterior part of the corpus callosum was significantly larger in musicians compared to nonmusicians.

Table 24.1 Midsagittal area of the corpus callosum

| Group | Size of corpus callosum ($mm^2$ (mean + SD)) | Anterior part | Posterior part |
|---|---|---|---|
| Musicians ($n = 30$) | 687 (85) | 371 (46)* | 314 (43) |
| $\leq 7$ years ($n = 21$) | 709 (81) | 384 (42)+ | 321 (44) |
| $> 7$ years ($n = 9$) | 637 (77) | 340 (43) | 297 (38) |
| Nonmusicians ($n = 30$) | 649 (88) | 344 (48)*+ | 305 (43) |

The anterior part of the corpus callosum differs insignificantly between musicians and nonmusicians (*). This difference was mainly due to musicians who started before 7 years of age with their musical training (+).

(>7 years old).[8] The differences in callosal size might be due to more fibres crossing through the corpus callosum, a larger proportion of thicker myelinated fibres with faster interhemispheric transfer, or fibres with thicker axons or more axon collaterals. The anterior part of the corpus callosum contains mainly fibres from frontal motor related regions and prefrontal regions,[47] and the anterior corpus callosum matures the latest of all callosal subregions. The anatomical difference in the mid-sagittal area of the anterior corpus callosum has to be seen in the context of a requirement for increased inter-hemispheric communication subserving complex bimanual motor sequences in musicians. This structural difference could have been triggered by performing and continuously practicing complicated and independent bimanual finger movements.

### Motor cortex

The exact histological extents of the primary motor cortex cannot be defined on current MR images. Therefore, we used a gross anatomical marker of the motor cortex, the intrasulcal length of the posterior bank of the precentral gyrus (ILPG). This marker corresponds roughly with the histological extent of the primary motor cortex.[48] All brains were spatially oriented and normalized to the coordinate system of Talairach and Tournoux. The ILPG was measured in horizontal slices (parallel to the AC–PC plane) from the deepest point of the central sulcus following the contour line of the posterior bank of the precentral gyrus to a lateral surface tangent which connected the crests of the pre- and postcentral gyrus. Results revealed a greater symmetry of the ILPG in the musician group in a dorsal subregion of the motor cortex (Talairach coordinates $z = 69$ to $z = 55$; Figure 24.2).

In addition to the greater symmetry, both left and right hemispheric measures were larger in musicians in this region. The greater symmetry was due primarily to a much larger size of the ILPG in the right hemisphere controlling the nondominant left hand. In addition, we found strong correlations (Figure 24.3) between the age of commencement of musical training and the mean right as well as left ILPG size.[10]

### Cerebellum

Several animal studies have provided compelling evidence that the cerebellum shows microstructural changes (e.g. increased number and density of synapses, glial cells, and

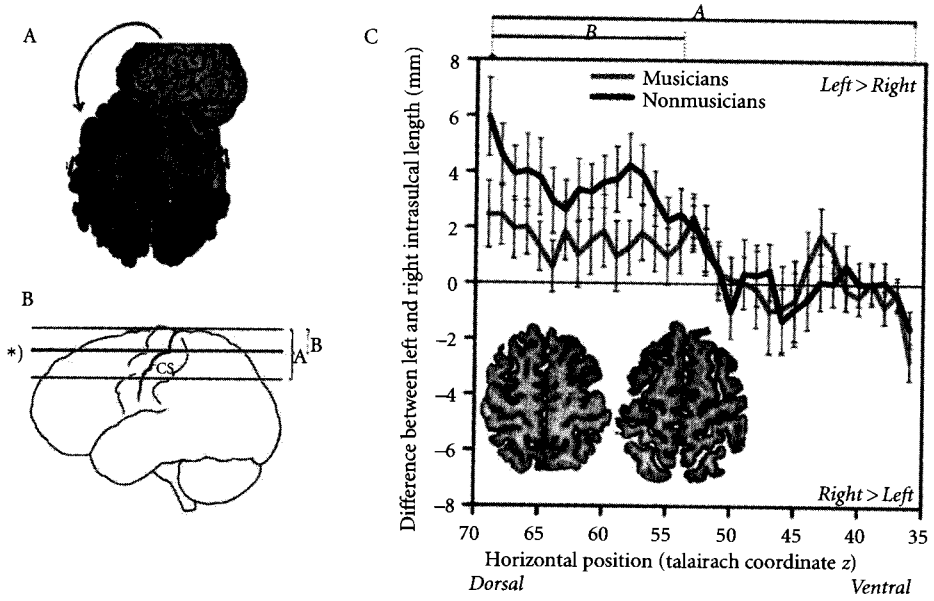



**Figure 24.2** (A) The intrasulcal length of the posterior bank of the precentral gyrus (ILPG) was measured in horizontal slices (parallel to the AC–PC plane) from the deepest point of the central sulcus following the contour line of the precentral gyrus to a lateral surface tangent that connected the crests of the pre- and postcentral gyrus. (B and C) We differentiated a dorsal and a ventral subregion. The dorsal subregion roughly corresponds to the location of the functional hand motor area. We found prominent differences in the intrasulcal length of the precentral gyrus in the dorsal subregion between musicians and nonmusicians. Musicians had more symmetrical lengths which was due to a disproportional increase in length of the nondominant hemisphere. (See Plate 21 in colour section.)

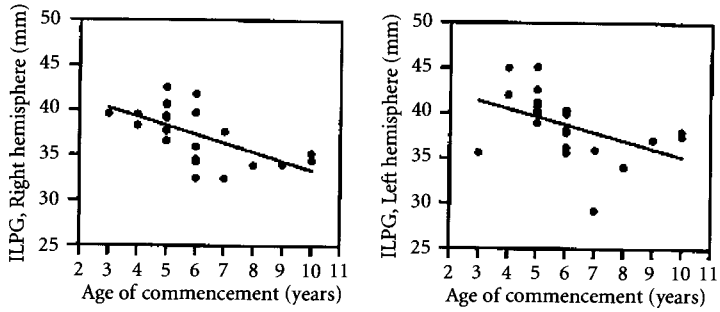



**Figure 24.3** Significant correlations between age of commencement of musical training and the intrasulcal length of the right and left precentral gyrus which was chosen as a marker of primary motor hand area.

capillaries) after long-term motor activity and motor skill acquisition. These findings combined with the role of the cerebellum in movement coordination, timing of sequential movements, and possibly other cognitive functions,[49] lead us to examine whether the cerebellum of adult musicians differs structurally from those of nonmusicians.[12] From a

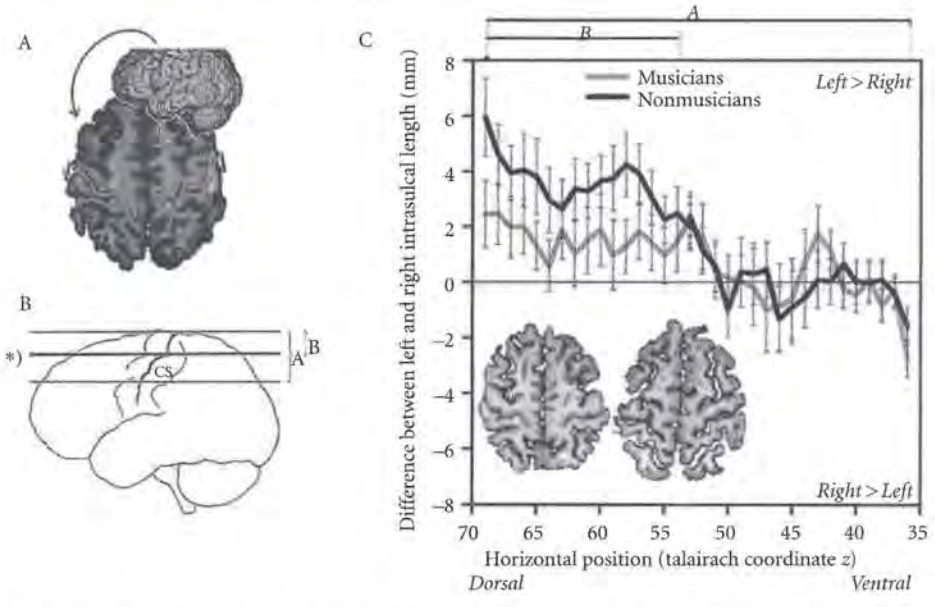



**Figure 24.2** (A) The intrasulcal length of the posterior bank of the precentral gyrus (ILPG) was measured in horizontal slices (parallel to the AC–PC plane) from the deepest point of the central sulcus following the contour line of the precentral gyrus to a lateral surface tangent that connected the crests of the pre- and postcentral gyrus. (B and C) We differentiated a dorsal and a ventral subregion. The dorsal subregion roughly corresponds to the location of the functional hand motor area. We found prominent differences in the intrasulcal length of the precentral gyrus in the dorsal subregion between musicians and nonmusicians. Musicians had more symmetrical lengths which was due to a disproportional increase in length of the nondominant hemisphere. (See Plate 21 in colour section.)

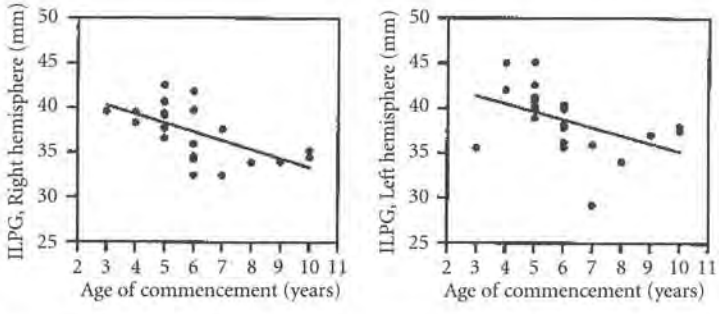



**Figure 24.3** Significant correlations between age of commencement of musical training and the intrasulcal length of the right and left precentral gyrus which was chosen as a marker of primary motor hand area.

capillaries) after long-term motor activity and motor skill acquisition. These findings combined with the role of the cerebellum in movement coordination, timing of sequential movements, and possibly other cognitive functions,[49] lead us to examine whether the cerebellum of adult musicians differs structurally from those of nonmusicians.[12] From a

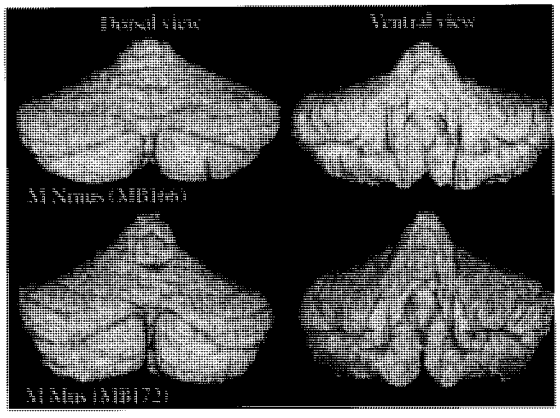


Figure 24.4 Dorsal and ventral views of the cerebellum. Cerebellar volume was normalized to whole brain volume to reduce intersubject variability. The relative cerebellar volume of male musicians was about 5 per cent larger than that of matched nonmusicians. There was a discrete positive correlation between cerebellar volume and intensity of musical training.

Table 24.2 Absolute brain volume and absolute and relative cerebellar volume in musicians and nonmusicians. There was a significant difference in relative cerebellar volume comparing male musicians to male nonmusicians ($p < 0.05$), but no difference in relative cerebellar volume comparing female musicians to female nonmusicians. Female musicians showed a trend for total brain volume differences ($p = 0.08$)

| Group | Brain volume (cc) | Rel. cerebellar volume (%) | Abs. cerebellar volume (cc) |
|---|---|---|---|
| Musicians (males; $n = 30$) | 1412.0 (84.5) | 10.4 (0.7)* | 146.2 (9.3) |
| Nonmusicians (males; $n = 30$) | 1405.2 (120.0) | 9.9 (0.7)* | 139.9 (14.4) |
| Musicians (females; $n = 26$) | 1308.4 (92.3) | 10.5 (0.6)+ | 136.7 (9.6) |
| Nonmusicians (females; $n = 26$) | 1266.6 (76.7) | 10.6 (0.7)+ | 134.8 (11.4) |

prospective databank of anatomical MR images of musicians and nonmusicians, we selected MRIs of 56 classically trained, professional musicians matched by gender, age, and handedness to a control group of nonmusicians. In order to compensate for the high intersubject variability in cerebellar volume, absolute cerebellar volume was normalized as a percentage of total brain volume. We found a significant gender effect with female subjects having a higher relative cerebellar volume although males showed a trend for a higher mean absolute cerebellar volume and total brain volume. Male musicians had a significantly higher mean relative cerebellar volume (10.4 vs 9.9 per cent; Figure 24.4; Table 24.2) compared to male nonmusicians. This was not due to a difference in total brain volume since male musicians and male nonmusicians did not show any significant difference in total brain volume. No significant difference in relative cerebellar size was found when female musicians were compared with their nonmusician counterparts. However, female musicians had a trend towards a higher absolute brain volume (Table 24.2).

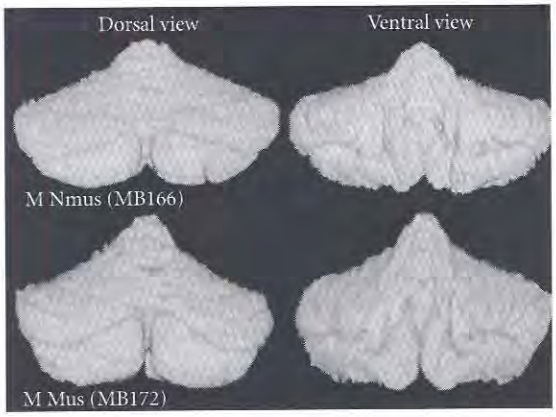



**Figure 24.4** Dorsal and ventral views of the cerebellum. Cerebellar volume was normalized to whole brain volume to reduce intersubject variability. The relative cerebellar volume of male musicians was about 5 per cent larger than that of matched nonmusicians. There was a discrete positive correlation between cerebellar volume and intensity of musical training.

**Table 24.2** Absolute brain volume and absolute and relative cerebellar volume in musicians and nonmusicians. There was a significant difference in relative cerebellar volume comparing male musicians to male nonmusicians ($p < 0.05$), but no difference in relative cerebellar volume comparing female musicians to female nonmusicians. Female musicians showed a trend for total brain volume differences ($p = 0.08$)

| Group | Brain volume (cc) | Rel. cerebellar volume (%) | Abs. cerebellar volume (cc) |
|---|---|---|---|
| Musicians (males; $n = 30$) | 1412.0 (84.5) | 10.4 (0.7)* | 146.2 (9.3) |
| Nonmusicians (males; $n = 30$) | 1405.2 (120.0) | 9.9 (0.7)* | 139.9 (14.4) |
| Musicians (females; $n = 26$) | 1308.4 (92.3) | 10.5 (0.6)+ | 136.7 (9.6) |
| Nonmusicians (females; $n = 26$) | 1266.6 (76.7) | 10.6 (0.7)+ | 134.8 (11.4) |

prospective databank of anatomical MR images of musicians and nonmusicians, we selected MRIs of 56 classically trained, professional musicians matched by gender, age, and handedness to a control group of nonmusicians. In order to compensate for the high intersubject variability in cerebellar volume, absolute cerebellar volume was normalized as a percentage of total brain volume. We found a significant gender effect with female subjects having a higher relative cerebellar volume although males showed a trend for a higher mean absolute cerebellar volume and total brain volume. Male musicians had a significantly higher mean relative cerebellar volume (10.4 vs 9.9 per cent; Figure 24.4; Table 24.2) compared to male nonmusicians. This was not due to a difference in total brain volume since male musicians and male nonmusicians did not show any significant difference in total brain volume. No significant difference in relative cerebellar size was found when female musicians were compared with their nonmusician counterparts. However, female musicians had a trend towards a higher absolute brain volume (Table 24.2).

The results in the male subgroup may suggest microstructural adaptations in the human cerebellum in response to early commencement and continual practice of complicated bimanual finger sequences, similar to findings in animal experiments.[50,20,22] The absence of an observable effect in the cerebellum in the female subgroup may be due to several reasons. First, females seem to reach adult cerebellar size much earlier than males during development.[51] Second, the relative cerebellar volume of females is larger than that of males, possibly indicating that a ceiling effect might have been reached. Third, synaptic up-and-down regulation during the female menstrual cycle has been described and could have diminished a group difference.[52] Fourth, there was a trend for a difference in total brain volume between female musicians and nonmusicians indicating that other brain regions outside the cerebellum might show structural differences. This could indicate gender related differences in presumed structural plasticity.

### Regional differences in gray matter volume

A voxel-by-voxel morphometric technique was used to detect local differences in gray matter volume and concentration between musicians and nonmusicians across the entire brain space (SPM99, Functional Imaging Laboratory, Institute of Neurology, London). This method involves spatial normalization of all images to the same anatomical space, extracting the gray matter from the normalized images, and analysing group differences in local concentration of gray matter on a voxel-by-voxel basis.[53] Professional musicians showed higher gray matter concentrations (Figure 24.5) compared to nonmusicians in the perirolandic region, the premotor region, the posterior superior parietal region, the posterior mesial perisylvian region bilaterally, and the cerebellum.[54]

This study replicates and significantly expands upon our previous studies using more traditional morphometric techniques describing size and volume differences in the motor

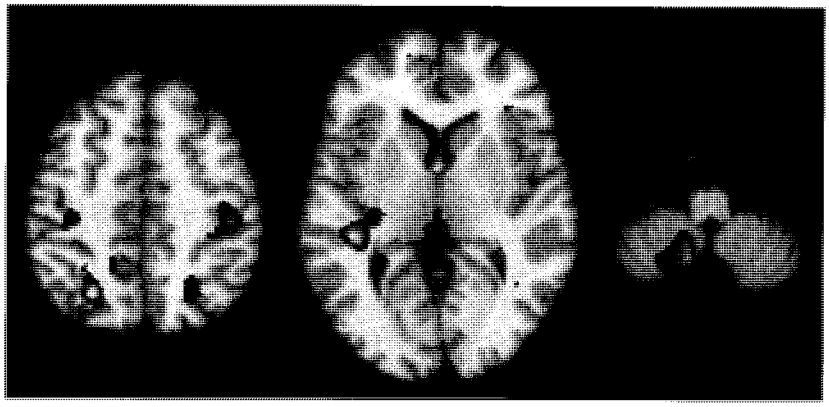


**Figure 24.5** Preliminary findings of a voxel-by-voxel morphometric study comparing a group of 15 male musicians with 15 male nonmusicians. Three axial slices show significant differences ($p < 0.05$ corrected) between musicians and nonmusicians superimposed on corresponding T1-weighted structural images. Yellow and red voxels indicate more graymatter concentration in the musicians, blue voxels indicate more gray matter concentration in the nonmusicians. (See Plate 22 in colour section.)

The results in the male subgroup may suggest microstructural adaptations in the human cerebellum in response to early commencement and continual practice of complicated bimanual finger sequences, similar to findings in animal experiments.[50,20,22] The absence of an observable effect in the cerebellum in the female subgroup may be due to several reasons. First, females seem to reach adult cerebellar size much earlier than males during development.[51] Second, the relative cerebellar volume of females is larger than that of males, possibly indicating that a ceiling effect might have been reached. Third, synaptic up-and-down regulation during the female menstrual cycle has been described and could have diminished a group difference.[52] Fourth, there was a trend for a difference in total brain volume between female musicians and nonmusicians indicating that other brain regions outside the cerebellum might show structural differences. This could indicate gender related differences in presumed structural plasticity.

### Regional differences in gray matter volume

A voxel-by-voxel morphometric technique was used to detect local differences in gray matter volume and concentration between musicians and nonmusicians across the entire brain space (SPM99, Functional Imaging Laboratory, Institute of Neurology, London). This method involves spatial normalization of all images to the same anatomical space, extracting the gray matter from the normalized images, and analysing group differences in local concentration of gray matter on a voxel-by-voxel basis.[53] Professional musicians showed higher gray matter concentrations (Figure 24.5) compared to nonmusicians in the perirolandic region, the premotor region, the posterior superior parietal region, the posterior mesial perisylvian region bilaterally, and the cerebellum.[54]

This study replicates and significantly expands upon our previous studies using more traditional morphometric techniques describing size and volume differences in the motor

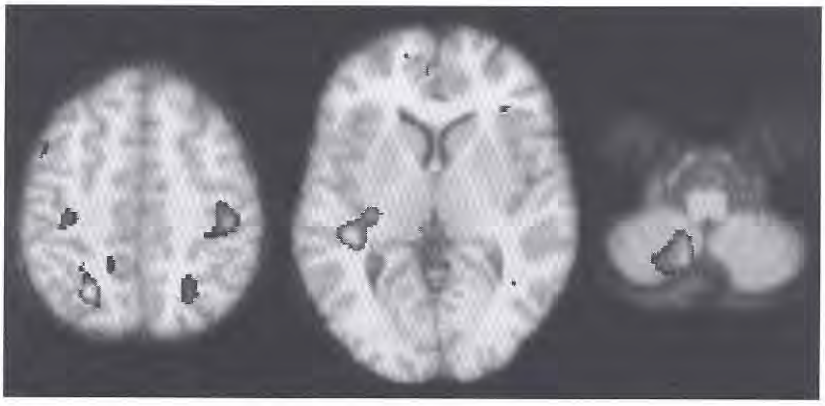


**Figure 24.5** Preliminary findings of a voxel-by-voxel morphometric study comparing a group of 15 male musicians with 15 male nonmusicians. Three axial slices show significant differences ($p < 0.05$ corrected) between musicians and nonmusicians superimposed on corresponding T1-weighted structural images. Yellow and red voxels indicate more graymatter concentration in the musicians, blue voxels indicate more gray matter concentration in the nonmusicians. (See Plate 22 in colour section.)

cortex, the cerebellum and the posterior perisylvian regions. In addition, this study showed differences in gray matter concentrations in regions not expected to show differences, such as the posterior superior parietal cortex. The superior parietal cortex does play an important role in music performance, since it may serve as an integrator of visual and auditory information with motor planning activities. Its importance in musical sight-reading has already been shown.[55]

## Functional brain differences between musicians and nonmusicians

Most pronounced functional differences between musicians and nonmusicians have been found in perisylvian brain regions with various perceptual tasks ranging from listening to musical pieces, pitch discrimination and memory, harmony, melody, and rhythm tasks.[4–6,11,56,57] It appears from these studies that musicians and nonmusicians process music in a different way, leading to more left hemispheric activation with increased musical sophistication. The reason for this is unclear and the historic explanation of a more emotional or holistic approach to music by the nonmusician vs a more analytic approach to music by the trained musician is an oversimplification of the underlying processes. However, music is certainly not just a right or left hemisphere issue. There may be aspects of music that will be processed more on the right hemisphere by both musicians and nonmusicians (e.g. melodic contour tasks) while there are others that will be more processed on the left by both groups (e.g. rhythmic tasks). Of interest are of course those aspects of music that may be processed differently. One example for this may be the processing of pitch information.

A region that seems to play an important role in the processing of musical stimuli is the posterior perisylvian region. This region is known to be very asymmetric in normal right-handers. The planum temporale has long been used as a marker of this asymmetry and in general as a marker of cerebral dominance, at least for language and handedness.[58] It has been found that the majority of right-handers have a leftward planum temporale asymmetry whereas the majority of left-handers have either a symmetric planum temporale or show a rightward asymmetry. The planum temporale is not only a structural marker of left-hemispheric dominance for language, it is also involved in auditory processing and as such is of great interest for studies investigating laterality effects of auditory processing. Recent studies have found associations between PT surface area and functional dominance when listening to stories[59] and the PT has been found to be activated in various musical tasks.[60–64,57]

In an anatomical MRI study designed to test the hypothesis that musicians might have a different degree of hemispheric dominance than nonmusicians, we found that musicians differed significantly from nonmusicians by having an increased leftsided asymmetry of the planum temporale (Figure 24.6). The surprising finding was that a subgroup of musicians, those with absolute pitch, explained all the difference between the musician and nonmusician group. The absolute pitch musician subgroup showed an increased leftsided asymmetry of the planum temporale.[9] Absolute pitch is the ability to identify a tone in the absence of a reference tone. The incidence varies between different studies, but is presumed to be between 5 and 20 per cent among musicians. In a subsequent study, Keenan *et al.*[13] compared

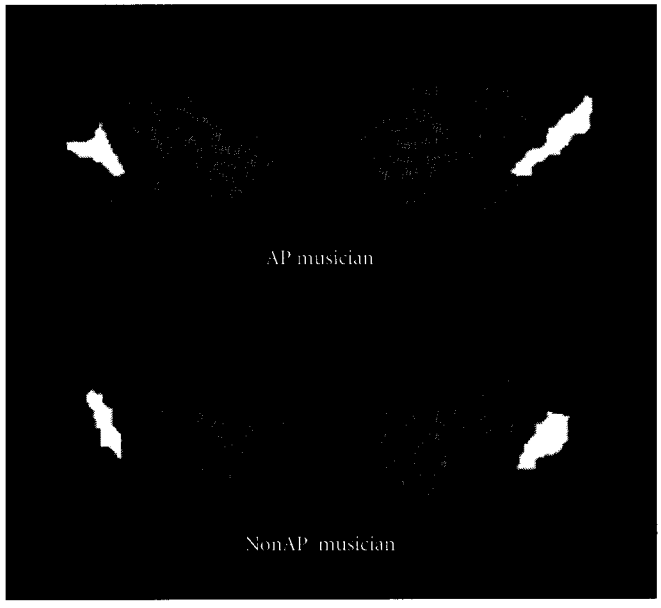



**Figure 24.6** Two representative brain renderings. The PT is highlighted in each brain. The AP musician has a pronounced leftward asymmetry that is not seen in the nonAP subject.

subgroups of AP and nonAP musicians with a similar early commencement of music training. Only the AP musicians showed extreme leftward PT asymmetry. Furthermore, Keenan *et al.*[13] found that a decrease of the right PT rather than an increase in the left PT determined the leftward asymmetry of the PT. From these and other studies[65,66] it appears that two factors are of importance in acquiring AP: age of commencement of musical training and an extreme PT asymmetry. In our sample and in other large samples, it is extremely seldom that someone developed absolute pitch if they did not start musical training or were exposed to music before the age of seven.

Although the functional significance of the increased leftsided PT asymmetry in AP musicians is not clear, it has to be examined in the context of their ability to assign any pitch to a pitch class or category in the absence of a reference tone.[67] Siegel[68] demonstrated the influence of possession of verbal labels on recognition memory for pitch. In her study, AP subjects were able to assign different verbal labels to tones that belonged to different pitch classes resulting in a better performance than nonAP subjects. Comparing different tones that belonged to the same pitch class did not result in a performance difference between AP and nonAP subjects since both groups were supposedly using a sensory coding strategy. These results could be taken as evidence for the categorical nature of absolute pitch.[68] We designed a functional MRI experiment that aimed to investigate the functional significance of the increased left-sided PT asymmetry in AP by contrasting two conditions with each other, a tone and a phoneme memory task. In the tone condition, subjects heard a string of sine wave tones (each tone was 500 ms long). A two-alternative forced choice

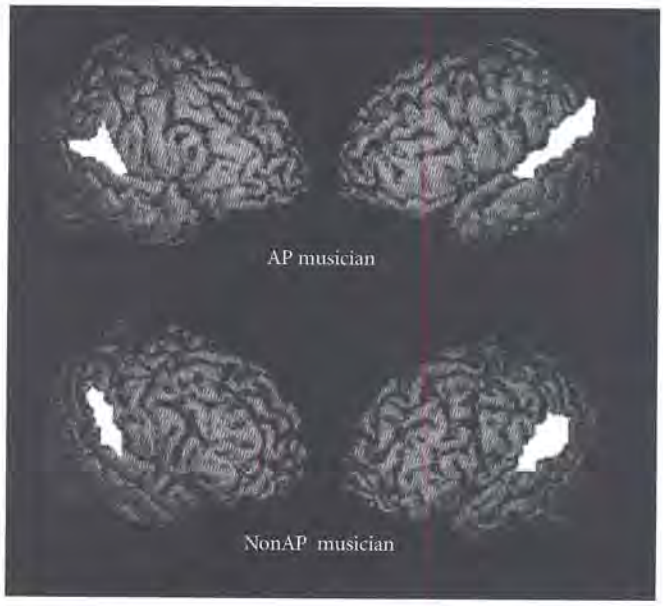



**Figure 24.6** Two representative brain renderings. The PT is highlighted in each brain. The AP musician has a pronounced leftward asymmetry that is not seen in the nonAP subject.

subgroups of AP and nonAP musicians with a similar early commencement of music training. Only the AP musicians showed extreme leftward PT asymmetry. Furthermore, Keenan *et al.*[13] found that a decrease of the right PT rather than an increase in the left PT determined the leftward asymmetry of the PT. From these and other studies[65,66] it appears that two factors are of importance in acquiring AP: age of commencement of musical training and an extreme PT asymmetry. In our sample and in other large samples, it is extremely seldom that someone developed absolute pitch if they did not start musical training or were exposed to music before the age of seven.

Although the functional significance of the increased leftsided PT asymmetry in AP musicians is not clear, it has to be examined in the context of their ability to assign any pitch to a pitch class or category in the absence of a reference tone.[67] Siegel[68] demonstrated the influence of possession of verbal labels on recognition memory for pitch. In her study, AP subjects were able to assign different verbal labels to tones that belonged to different pitch classes resulting in a better performance than nonAP subjects. Comparing different tones that belonged to the same pitch class did not result in a performance difference between AP and nonAP subjects since both groups were supposedly using a sensory coding strategy. These results could be taken as evidence for the categorical nature of absolute pitch.[68] We designed a functional MRI experiment that aimed to investigate the functional significance of the increased left-sided PT asymmetry in AP by contrasting two conditions with each other, a tone and a phoneme memory task. In the tone condition, subjects heard a string of sine wave tones (each tone was 500 ms long). A two-alternative forced choice

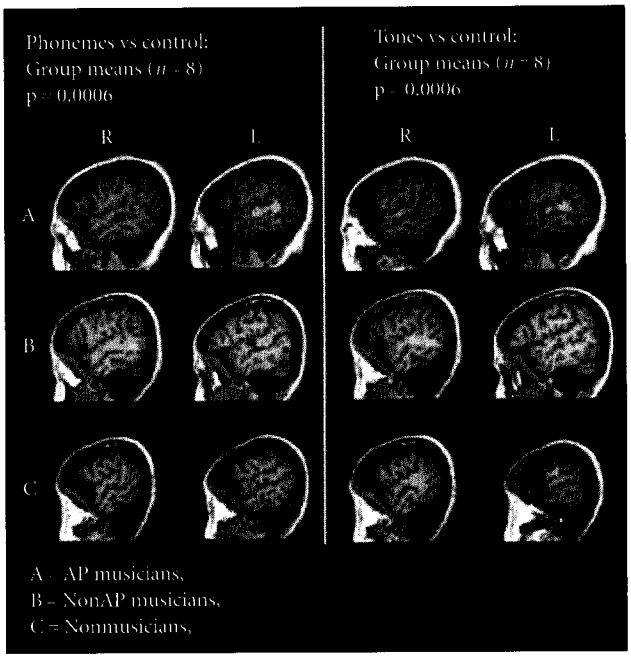



Figure 24.7 Results of a fMRI experiment aimed to investigate the functional significance of the increased leftsided anatomical PT asymmetry in AP musicians. Three groups of subjects were examined: AP musicians, nonAP musicians, and nonmusicians. A phoneme memory task and a tone memory task were contrasted with a rest condition. Suprathreshold voxels showing significant task related activity changes were superimposed on a representative sagittal anatomical image of right and left hemisphere. AP musicians showed prominent left superior temporal lobe activation during the tone memory task that was not seen in the other two groups. The cerebral activation pattern of the phoneme memory task was similar between the three groups. (See Plate 23 in colour section.)

paradigm was used requiring the subjects to indicate whether each new tone was 'same' or 'different' compared to two tones ago. In the phoneme task five different consonant-vowel syllables were presented in the same way as in the tone memory task. AP musicians, nonAP musicians and nonmusicians showed pronounced left-sided temporal lobe activation in the phonemes vs control task comparison. Similar activations were seen in the tone memory condition for the AP musicians. However, the nonAP musicians as well as the nonmusicians showed either symmetric or right-sided activation of the superior temporal lobe in the tone memory task (Figure 24.7). It appeared that in AP musicians, the perception of phonemes and tones is mediated by the same superior temporal lobe structures on the left hemisphere. Nonabsolute pitch control subjects showed a differential involvement of the superior temporal lobe depending on the stimuli. The left superior temporal lobe may be a mediator of a specialized perception that enables the categorization of auditory stimuli.[57]

Recently, Ohnishi *et al.*[64] were able to demonstrate the preferential involvement of the PT in music perception in trained musicians, with particularly strong activation of the left PT in AP musicians. Zatorre *et al.*[61] did find left hemisphere activation when contrasting

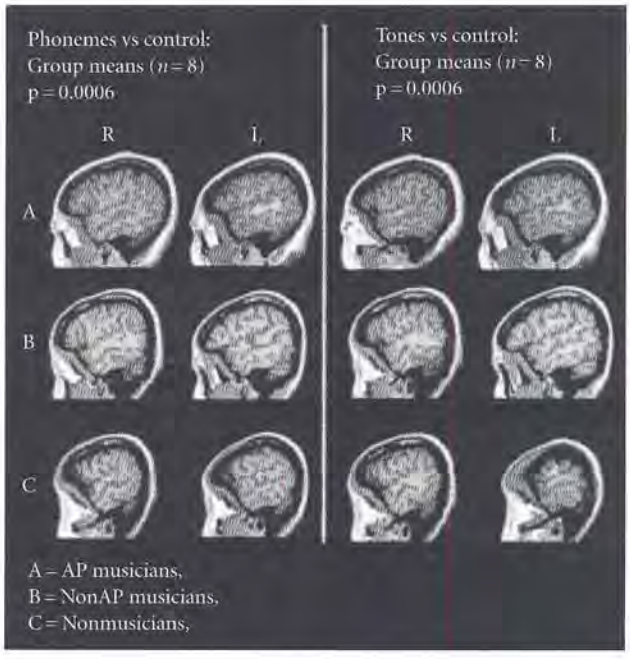



**Figure 24.7** Results of a fMRI experiment aimed to investigate the functional significance of the increased leftsided anatomical PT asymmetry in AP musicians. Three groups of subjects were examined: AP musicians, nonAP musicians, and nonmusicians. A phoneme memory task and a tone memory task were contrasted with a rest condition. Suprathreshold voxels showing significant task related activity changes were superimposed on a representative sagittal anatomical image of right and left hemisphere. AP musicians showed prominent left superior temporal lobe activation during the tone memory task that was not seen in the other two groups. The cerebral activation pattern of the phoneme memory task was similar between the three groups. (See Plate 23 in colour section.)

paradigm was used requiring the subjects to indicate whether each new tone was 'same' or 'different' compared to two tones ago. In the phoneme task five different consonant-vowel syllables were presented in the same way as in the tone memory task. AP musicians, nonAP musicians and nonmusicians showed pronounced left-sided temporal lobe activation in the phonemes vs control task comparison. Similar activations were seen in the tone memory condition for the AP musicians. However, the nonAP musicians as well as the nonmusicians showed either symmetric or right-sided activation of the superior temporal lobe in the tone memory task (Figure 24.7). It appeared that in AP musicians, the perception of phonemes and tones is mediated by the same superior temporal lobe structures on the left hemisphere. Nonabsolute pitch control subjects showed a differential involvement of the superior temporal lobe depending on the stimuli. The left superior temporal lobe may be a mediator of a specialized perception that enables the categorization of auditory stimuli.[57]

Recently, Ohnishi *et al.*[64] were able to demonstrate the preferential involvement of the PT in music perception in trained musicians, with particularly strong activation of the left PT in AP musicians. Zatorre *et al.*[61] did find left hemisphere activation when contrasting

pitch perception in musicians with and without absolute pitch, however, in the critical comparison, only a dorsal posterior frontal region became significant.

Pantev *et al.*[11] found evidence for an increased auditory cortical representation of piano tones in highly skilled musicians which was not seen for pure tones. There was a correlation between the enhanced representation of tones in the auditory cortex and age of commencement of musical training. No significant difference was seen between musicians with and without absolute pitch. This may have been due to the fact that mainly the N1 which occurs up to 100 ms after tone onset was examined. In a subsequent report, a functional difference was found between musicians with and without absolute pitch. AP musicians showed a distinct activity more posterior in the left superior temporal lobe than the nonmusicians.[69]

## Implicit musical ability of the human brain

Although there are differences in the way that musical stimuli are processed between musicians and nonmusicians, it has also become clear over the last years that nonmusicians possess some implicit ability to process musical stimuli. This then leads to the question whether the differences between musicians and nonmusicians observed in neurophysiological studies are either due to training and experience in processing complex auditory stimuli or are due to innate anatomical or functional differences.

Although enhanced cortical responses in the N1[11] as well as in the transient N1/P2 complex have been seen comparing musicians with nonmusicians, there have been other studies showing that early cortical responses to musical stimuli (such as contour violations, etc.) may be processed similarly between musicians and nonmusicians.[70] Koelsch and colleagues[63] examined early brain potentials in subjects listening to violations of musical syntax. The mapping of these changes indicated that neural activity covaried with these events in regions of the brain that were previously thought to play an important role in language processing.[71] Furthermore, these results were obtained in naïve subjects without explicit musical knowledge. On the contrary, Ruesseler *et al.*[72] showed that professional musicians exhibited a mismatch negativity (MMN), a frontal negative wave in the event-related potential (ERP), for tones anticipated by as little as 20 ms within a series of regularly spaced tones while nonmusicians did not.

Thus, there is evidence for an implicit ability of the human brain to process musical stimuli. It will need to be determined in future studies how training and experience changes the neural elements subserving the perceptual processing of musical stimuli and subserving the planning and execution of highly complex motor action requiring the integration of multimodal sensory information and precise monitoring and feedback.

## Summary

Musicians are an ideal human model to investigate possible functional and structural neural changes due to the acquisition and continuous practice of complex perceptions and actions. Experimental animal studies strongly support the existence of microstructural

plasticity. Many studies have provided evidence for functional and structural differences comparing musicians with nonmusicians. Although self-selection for musicianship by individuals with innate functional and structural brain differences cannot be completely ruled out, there is now more and more evidence to support the notion that musical training will lead to changes in brain function and structure.

## Acknowledgements

I would like to thank all the subjects, musicians and nonmusicians, who participated in our studies and allowed us to learn more about the brain that makes music. I like to acknowledge the enormous contributions that have been made during various parts of these ongoing studies by my former mentor Helmuth Steinmetz as well as by my friends, colleagues and coworkers Lutz Jaencke, Katrin Amunts, Ben Martin Bly, Andrea Halpern, Michael Charness, Ven Thangaraj, Julian Keenan, Siobhan Hutchinson, Yi Chen, and Christian Gaser. Furthermore, this work would have not been possible without the invaluable help, dedication, and efforts of my students Leslie Lee, Maria Picone, Katrin Schulze, Eileen Lüders, and Nadine Gaab. I am also very thankful to my wife, Lisa Jones, for her constant support and her many invaluable contributions to this work.

## References

1. **Tramo, M. J.** (2001) Music of the hemispheres. *Science* 291, 54–6.
2. **Bever, T. G.** and **R. J. Chiarello** (1974) Cerebral dominance in musicians and nonmusicians. *Science* 185, 537–9.
3. **Meyer, A.** (1977) The search for a morphological substrate in the brains of eminent persons including musicians: a historical review. In M. Critchley and R. A. Henson (ed.) *Music and the Brain.* London: W. Heinemann, pp. 255–97.
4. **Mazziotta, J. C., M. E. Phelps, R. E. Carson,** and **D. E. Kuhl** (1982) Tomographic mapping of human cerebral metabolism: auditory stimulation. *Neurology* 32, 921–37.
5. **Altenmueller, E.** (1986) Electrophysiological correlates of music processing in the human brain. *Euro. Archives Psychiatry Neurol. Sci.* 235, 342–54.
6. **Besson, M., F. Faita,** and **J. Requin** (1994) Brain waves associated with musical incongruities differ for musicians and non-musicians. *Neurosci. Lett.* 168, 101–5.
7. **Elbert, T., C. Pantev, C. Wienbruch, B. Rockstroh,** and **E. Taub** (1995) Increased cortical representation of the fingers of the left hand in string players. *Science* 270, 305–6.
8. **Schlaug, G., L. Jaencke, Y. Huang,** and **H. Steinmetz** (1995) In vivo evidence of structural brain asymmetry in musicians. *Science* 267, 699–701.
9. **Schlaug, G., L. Jaencke, Y. Huang,** and H. **Steinmetz** (1995) Increased corpus callosum size in musicians. *Neuropsychologia* 33, 1047–55.
10. **Amunts, K., G. Schlaug, L. Jaencke, A. Dabringhaus, H. Steinmetz, A. Schleicher,** and **K. Zilles** (1997) Motor cortex and hand motor skills: structural compliance in the human brain. *Hum. Brain Map.* 5, 206–15.

11. **Pantev, C., R. Oostenveld, A. Engelien, B. Ross, L. E. Roberts,** and **M. Hoke** (1998) Increased auditory cortical representation in musicians. *Nature* 392, 811–14.
12. **Schlaug, G.** (2001) The brain of musicians: a model for functional and structural adaptation. In R. J. Zatorre and I. Peretz (ed.) *The Biological Foundations of Music*, vol. 930, New York: Annals of the New York Academy of Sciences, pp. 281–99.
13. **Keenan, J. P., V. Thangaraj, A. Halpern,** and **G. Schlaug** (2001) Planum temporale and absolute pitch. *Neuroimage* 14, 1402–8.
14. **Hutchinson, S., L. H. L. Lee, N. Gaab,** and **G. Schlaug** (2003) Cerebellar volume in musicians. *Cerebral Cortex*, in press.
15. **Rauschecker, J. P.** (1995) Compensatory plasticity and sensory substitution in the cerebral cortex. *Trends Neurosci.* 18, 36–45.
16. **Hallett, M.** (1995) The plastic brain. *Ann. Neurol.* 38, 4–5.
17. **Zilles, K.** (1992) Neuronal plasticity as an adaptive property of the central nervous system. *Annals of Anatomy* 174, 383–91.
18. **Zheng, D.** and **D. Purves** (1995) Effects of increased neural activity on brain growth. *Proc. Natl. Acad. Sci. USA* 92, 1802–6.
19. **Nudo, R. J., G. W. Milliken, W. M. Jenkins,** and **M. M. Merzenich** (1996) Use-dependent alterations of movement representations in primary motor cortex of adult squirrel monkeys. *J. Neurosci.* 16, 785–807.
20. **Black, J. E., K. R. Isaacs, B. J. Anderson, A. A. Alcantara,** and **W. T. Greenough** (1990) Learning causes synaptogenesis whereas motor activity cause angiogenesis, in cerebellar cortex of adult rats. *Proc. Natl. Acad. of Sci. USA* 87, 5568–72.
21. **Isaacs, K. R., B. J. Anderson, A. A. Alcantara, J. E. Black,** and **W. T. Greenough** (1992) Exercise and the brain: angiogenesis in the adult rat cerebellum after vigorous physical activity and motor skill learning. *J. Cerebr. Blood Flow and Metabolism* 12, 110–19.
22. **Anderson, B. J., X. Li, A. Alcantara, K. R. Isaacs, J. E. Black,** and **W. T. Greenough** (1994) Glial hypertrophy is associated with synaptogenesis following motor-skill learning, but not with angiogenesis following exercise. *Glia* 11, 73–80.
23. **Kleim, J. A., E. Lussnig, E. R. Schwarz, T. A. Comery,** and **W. T. Greenough** (1996) Synaptogenesis and FOS expression in the motor cortex of the adult rat after motor skill learning. *J. Neurosci.* 16, 4529–35.
24. **Kempermann, G., H. G. Kuhn,** and **F. H. Gage** (1997) More hippocampal neurons in adult mice living in an enriched environment. *Nature* 386, 493–5.
25. **Praag, H., B. R. Christie, T. J. Sejnowski,** and **F. H. Gage** (1999) Running enhances neurogenesis, learning and long-term potentiation in mice. *Proc. Natl. Acad. Sci. USA* 96, 13427–31.
26. **Dettmers, C., J. Liepert, T. Adler, *et al.*** (1999) Abnormal motor cortex organization contralateral to early upper limb amputation in humans. *Neurosci. Lett.* 263, 41–4.
27. **Maguire, E. A., D. G. Gadian, I. S. Johnsrude, *et al.*** (2000) Navigation-related structural change in the hippocampi of taxi drivers. *Proc. Natl. Acad. Sci. USA* 97, 4398–403.
28. **Merzenich, M. M., G. Recanzone, W. M. Jenkins, *et al.*** (1988) Cortical representational plasticity. In P. Rakic and W. Singer (ed.) *Neurobiology of Neocortex*. New York: John Wiley Sons Limited, pp. 41–7.
29. **Recanzone, G. H., T. T. Allard, W. M. Jenkins, *et al.*** (1990) Receptive-field changes induced by peripheral nerve stimulation in S1 of adult cats. *J. Neurophysiol.* 63, 1213–25.

30. **Jacobs, K. M.** and **J. P. Donoghue** (1991) Reshaping the cortical motor map by unmasking latent intracortical connections. *Science* 251, 944–7.

31. **Wang, X., M. M. Merzenich, K. Sameshima,** ***et al.*** (1995) Remodeling of hand representation in adult cortex determined by timing of tactile stimulation. *Nature* 378, 71–5.

32. **Cohen, L. G., S. Bandinelli, T. W. Findeley,** ***et al.*** (1991) Motor reorganization after upper limb amputation in man. *Brain* 114, 615–27.

33. **Jenkins, I. H., D. J. Brooks, P. D. Nixon,** ***et al.*** (1994) Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation. *J. Neurosci.* 14, 3775–90.

34. **Pascual-Leone, A., J. Grafman,** and **M. Hallett** (1994) Modulation of cortical motor output maps during development of implicit and explicit knowledge. *Science* 263, 1287–9.

35. **Pascual-Leone, A., N. Dang, L. G. Cohen,** ***et al.*** (1995) Modulation of muscle responses evoked by transcranial magnetic stimulation during the acquisition of new fine motor skills. *J. Neurophysiol.* 74, 1037–44.

36. **Pascual-Leone, A., E. M. Wassermann, N. Sadato,** ***et al.*** (1995) The role of reading activity on the modulation of motor cortical outputs to the reading hand in Braille readers. *Ann. Neurol.* 38, 910–15.

37. **Schlaug, G., U. Knorr,** and **R. J. Seitz** (1994) Inter-subject variability of cerebral activations in acquiring a motor skill: a study with positron emission tomography. *Exp. Brain Res.* 98, 523–34.

38. **Karni, A., G. Meyer, P. Jezzard, M. M. Adams, R. Turner,** and **L. G. Ungerleider,** (1995) FMRI evidence for adult motor cortex plasticity during motor skill learning. *Nature* 377, 155–8.

39. **Hund-Georgiadis, M.** and **Y. von Cramon** (1999) Motor-learning-related changes in piano players and non-musicians revealed by functional magnetic-resonance signals. *Exp. Brain Res.* 125, 417–25.

40. **Charness, M.** and **G. Schlaug** (2000) Cortical activation during finger movements in concert pianists, dystonic pianists, and non-musicians. *Neurology* 54, A221

41. **Annett, M.** (1970) A classification of hand preference by association analysis. British *J. Psychol.* 61, 303–21.

42. **Jancke, L., G. Schlaug,** and **H. Steinmetz** (1997) Hand skill asymmetry in professional musicians. *Brain and Cognit.* 34, 424–32.

43. **Peters, M.** and **B. M. Durding** (1978) Handedness measured by finger tapping: a continuos variable. *Canadian J. Psychol.* 32, 257–61.

44. **Wechsler, D.** (1981) Wechsler Adult Intelligence Scale – revised. New York: The Psychological Corporation.

45. **Yakovlev, P. I.** and **A. R. Lecours** (1994) The myelogentic cycles of regional maturation of the brain. In A. Minkowski (ed.) *Regional Development of the Brain in Early Life.* Oxford: Blackwell, pp. 3–70.

46. **Pujol, J., P. Vendrell, C. Jungue,** ***et al.*** (1993) When does human brain development end? Evidence of corpus callosum growth up to adulthood. *Ann. Neurol.* 34, 71–5.

47. **Pandya, D. N.** and **B. Seltzer,** (1986) The topography of commissural fibers. In F. Lepore, M. Ptito, and H. H. Jasper (ed.) *Two Hemispheres – One Brain. Functions of the Corpus Callosum,* New York: Alan R Liss, pp. 47–73.

48. **Amunts, K., G. Schlaug, A. Schleicher, H. Steinmetz, A. Dabringhaus, P. E. Roland,** and **K. Zilles** (1996) Asymmetry in the human motor cortex and handedness. *Neuroimage* 4, 216–22.

49. **Schmahmann, J. D.** and **J. C. Sherman** (1998) The cerebellar cognitive affective syndrome. *Brain* 121, 561–79.

50. **Psych, J. J.** and **G. M. Weiss,** (1979) Exercise during development induces an increase in purkinje cell dendritic tree size. *Science* 206, 230–1.

51. **Caviness, V. S., D. N. Kennedy, C. Richelme, J. Rademacher,** and **P. A. Filipek** (1996) The human brain age 7–11 years: a volumetric analysis based on magnetic resonance images. *Cerebr. Cortex* 6, 726–36.

52. **Woolley, C. S.** and **B. S. McEwen** (1992) Estradiol mediates fluctuation in hippocampal synapse density during the estrous cycle in the adult rat. *J. Neurosci.* 12, 2549–54.

53. **Ashburner, J.** and **K. J. Friston** (2000) Voxel-based morphometry – the methods. *Neuroimage* 11, 805–21.

54. **Gaser, C.** and **G. Schlaug** (2001) Brain structures differ between musicians and non-musicians. *Neuroimage* 13, S1168.

55. **Sergent, J., E. Zuck, S. Terriah,** and **B. MacDonald** (1992) Distributed neural network underlying musical sight-reading and keyboard performance. *Science* 257, 106–9.

57. **Bly, B. M., C. Gaser,** and **G. Schlaug** The functional anatomy of pitch perception, submitted.

58. **Steinmetz, H., J. Volkmann, L. Jäncke,** and **H. J. Freund** (1991) Anatomical left-right asymmetry of language-related temporal cortex is different in left- and right-handers. *Ann. Neurol.* 29, 315–19.

59. **Tzourio, N., B. Nkanga-Ngila,** and **B. Mazoyer** (1998) Left planum temporale surface correlates with functional dominance during story listening. *Neuroreport* 9, 829–33.

60. **Zatorre, R. J., A. C. Evans,** and **E. Meyer** (1994) Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 14, 1908–19.

61. **Zatorre, R. J., D. W. Perry, C. A. Beckett, C. F. Westbury,** and **A. C. Evans** (1998) Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. *Proc. Natl. Acad. Sci. USA* 95, 3172–7.

62. **Rauschecker, J. P., B. Tian,** and **M. Hauser** (1995) Processing of complex sounds in the macaque non-primary auditory cortex. *Science* 268, 111–14.

63. **Koelsch, S., T. Gunter,** and **A. D. Friederici** (2000) Brain indices of music processing: 'nonmusicians' are musical. *J. Cognitive Neurosci.* 12, 520–41.

64. **Ohnishi, T., H. Matsuda, T. Asada,** ***et al.*** (2001) Functional anatomy of musical perception in musicians. *Cerebr. Cortex* 11, 754–60.

65. **Sergeant, D.** (1969) Experimental investigation of absolute pitch. *Journal of Research in Music Education* 17, 135–43.

66. **Baharloo, S., P. A. Johnston, S. K. Service, J. Gitschier,** and **N. B. Freimer** (1998) Absolute pitch: an approach for identification of genetic and nongenetic components. *Am. J. Hum. Genetics* 62, 224–31.

67. **Takeuchi, A. H.** and **S. H. Hulse** (1993) Absolute pitch. *Psychological Bulletin* 113, 345–61.

68. **Siegel, J. A.** (1974) Sensory and verbal coding strategies in subjects with absolute pitch. *J. Exp. Psychol.* 103, 37–44.

69. **Hirata, Y., S. Kuriki,** and **C. Pantev** (1999) Musicians with absolute pitch show distinct neural activities in the auditory cortex. *Neuroreport* 10, 999–1002.

70. **Trainor, L. J., R. N. Desjardins,** and **C. Rockel** (1999) A comparison of contour and interval processing in musicians and non-musicians using event-related potentials. *Australian Journal of Psychology* 51, 147–53.

71. **Maess, B., S. Koelsch, T. C. Gunter,** and **A. D. Friederici** (2001) Musical syntax is processed in Broca's area: an MEG study. *Nature Neurosci.* 4, 540–5.

72. **Ruesseler, J. E., W. Altenmuller, C. Nager, T. F. Kohlmetz, and T. F. Munte** (2001) Event-related brain potentials to sound omissions differ in musicians and non-musicians. *Neurosci. Lett.* 308, 33–6.

CHAPTER 25

# REPRESENTATIONAL CORTEX IN MUSICIANS

C. PANTEV, A. ENGELIEN, V. CANDIA, AND T. ELBERT

## Abstract

The lifelong ability to adapt to environmental needs is based on the capacity of the central nervous system for plastic alterations. In a series of neurophysiological experiments, we studied the impact of music and musical training in musicians on the specific functional organization in auditory and somatosensory representational cortex. In one such study, subjects listened to music from which one specific spectral frequency was removed. This led to rapid and reversible adaptation of neuronal responses in auditory cortex. Further experimental evidence demonstrated that long years of practice and training by professional musicians to enable them to reach their capacity is associated with enlarged cortical representations in the somatosensory and auditory domains. This tuning of neuronal representations was specifically observed for musical tones and was absent when pure sinusoidal tones were used as stimuli. In the somatosensory cortex, plastic changes proved to be specific for the fingers frequently used and stimulated. These changes were not detected in the fingers of the hand that were not involved in playing the particular instrument. Neuroplastic alterations also may be driven into a domain where they may become maladaptive. The clinical syndrome of focal hand dystonia that may occur in musicians who engage in forceful practice may be one such consequence. We will discuss the possibilities of reversing maladaptive responses leading to the successful treatment of focal hand dystonia, which relies on basic research about cortical reorganization. This example elucidates how neuroscientific progress can guide the development of practice guidelines and therapeutic measures for the benefit of professional musicians.

## Introduction

The structural and functional organization of the human brain becomes increasingly differentiated during child development. In higher mammals, including humans, neurons are formed prenatally as are some of their interconnections into neural networks. For many years the prevailing opinion suggested that network connections between neurons are built primarily during cerebral maturation processes in childhood, with the exception of only those structures that were directly involved in memory. It was thought that this network pattern, almost like a connection diagram, would not change later. However, humans respond with considerable flexibility to new challenges throughout their entire life. Since the early 1980s, increasing experimental evidence demonstrated that the connectivity of the adult brain is in fact only partially determined by genetics and early development, and may be substantially modified through sensory experiences, even during adulthood.

Practising and performing music entails intense auditory and somatosensory peripheral sensory excitations, which are transmitted via specific receptors and fibre systems to the

corresponding specialized regions of the cerebral cortex. Because the functional organization of representational cortex has been intensely studied, the examination of sensory cortical areas in the somatosensory and auditory systems provides an excellent model for studying the plastic changes that are associated with being a musician. The sensory cortex of both systems has a known topographical order of neuronal representations: a homuncular mapping of the body surface on somatosensory cortex[1–3] and a tonotopic mapping of acoustic frequency in auditory cortex.[4–7] Thus, modifications of representations can be specifically assessed. For example, different fingers of a given hand may not be equally used (and thus not equally stimulated) when playing certain musical instruments.

Based on W. James's suggestion at the end of the nineteenth century that learning may alter synaptic connectivity, the prominent Canadian neuropsychologist Donald Hebb formulated an important and innovative theory to view the brain not as a static but as a dynamic system. Hebb's rule asserts that effective connections between neurons are formed depending on synchronous activation: 'Cells that fire together, wire together'.[8] The development of new scientific methods for recording neuronal activity has made it possible to prove the hypothesis of plasticity in functional neuronal networks. The electrical activity of single neurons can only be recorded invasively in animals, but derived potentials that reflect the activity of a group of neurons can be recorded noninvasively on the scalp surface by means of electroencephalography (EEG). The magnetic counterpart of the EEG is the magnetoencephalography (MEG), which has become an established method for noninvasive study of the activity of the human cortex.[9] The main sources of cortical evoked magnetic fields are the pyramidal cells, which produce currents flowing tangentially to the surface of the head. Although MEG measurements provide only a macroscopic view of the function of the brain, the spatial resolution achieved with this technique is sufficient to give indications of functional organization and reorganizational plasticity of the human cortex by localizing the sources of evoked magnetic fields, which are elicited by defined and standardized peripheral excitation.

The first studies to clearly demonstrate reorganizational plasticity of the adult cortex were performed in the 1980s in 'deafferentation' experiments. Typically, in these studies the afferent information influx specific to certain cortical areas was eliminated.[10–15] One mechanism of change observed was that neurons that had lost their regular input were recruited by neighbouring regions. For example, neurons that were specialized for processing information from the fifth digit (little finger) may process information coming in on the second digit after the fifth digit was amputated. Similarly, neurons that were specific to a certain acoustic frequency may respond to neighbouring spared frequencies after the part of the cochlea transmitting their originally preferred input was destroyed.

Experimental studies in monkeys have demonstrated that intensive sensory stimulation may lead to an expansion of the corresponding cortical area.[16,17] Here we briefly review the experimental paradigm for one such study in the somatosensory modality concerning finger representation in adult monkeys. The animals were trained to touch a rotating disk with the tip of their index and middle fingers for 15 s. The surface of that disk was an irregular grid, so that it caused a sensation (and somatosensory excitation) specifically for those two fingertips. Anytime the monkeys touched the disk for 15 s, they received a reward (chips with banana taste). After about 600 stimulation periods, the representation of the

hand in the somatosensory cortex was measured. The representation for the two stimulated fingers had become enlarged. Recanzone *et al.* trained owl monkeys for 60–80 daily sessions to make fine-pitch discriminations in selected regions of the auditory frequency spectrum (these regions differing among animals).[17] Tonotopic mapping carried out invasively afterward showed that the cortical area tuned to the trained frequencies was enlarged by a factor of 2–3 compared to untrained monkeys or to animals that experienced the same acoustic stimuli passively while being trained on a somatosensory discrimination task. Thus, the organization of the brain seems capable of significant change to adapt to the changing demands of the organism's environment. Synchronized sensory stimulation may be of particular importance for such changes. The common observation captured in the German proverb *Übung macht den Meister* [practice makes perfect] could therefore have its neurobiological correlate in an augmented simultaneous stimulation of neurons, which entrains a reorganization of the functional neuronal network.

Musicians practise and train for many years before they achieve their professional skills. Coordination and synchronization of somatosensory and motor control on the one hand, and audition on the other hand, are crucial when playing a musical instrument. We therefore hypothesized that musical aptitude and the training to fulfill it may be associated with plastic changes of neuronal representations in the cortical organization. We will describe below a series of experiments that support this hypothesis for the somastosensory and auditory domains.[18–20]

Neuronal representations in the given cortical field of interest were determined by measuring evoked magnetic fields contralateral to the stimulated side with a 37-channel BTi Magnes system. For the somatosensory modality, a brief pneumatic stimulation was applied to the fingertips; for the auditory modality, a variety of different auditory stimuli were delivered to the right ear. The following sections describe some experimental examples in more detail.

## Short-term plasticity effect of the auditory cortex induced by notched music

Most studies of deafferentation-induced cortical reorganization have investigated cortical reorganization on a timescale of days to weeks, or longer. However, other more recent studies have documented rapid changes in cortical dynamics following deafferentation. These studies have shown that neurons broaden and shift their receptive fields to sensory surfaces near or beyond the edge of the lesioned zone within minutes of deafferentation in the somatosensory[21] and visual systems,[22] and within hours in the auditory one.[23] Rapid retuning of sensory neurons has also been observed following reversible 'functional' deafferentation in which sensory input from the environment is altered by procedures such as artificial scotomota[24] or digit ligation[25] rather than by permanent lesions of the receptor organs.

A study of 'functional deafferentation' was carried out to determine whether plastic changes of frequency representation occur on a short timescale of a few hours when the adult human auditory cortex is deprived of sensory input. Ten normal test subjects were asked to provide three favourite CDs from their CD collection. The music was manipulated

in such a way that a notch between 0.7 and 1.3 kHz, centred around 1 kHz, was produced using a band rejection filter (Bessel, 96 dB/oct) in the broad-band spectrum of the music (cf. Figure 25.1a). Although this manipulation initially produced a clearly noticeable change in perception, subjects reported that they quickly adapted to the modified sound and that their appreciation of the music during the listening time was unchanged. The subjects were asked to listen attentively to the music for a continuous period of 3 h. Due to the notch filtering during this period, any afferent input to cortical neurons tuned to frequencies around 1 kHz was abolished. In order to measure the effect of notching of the music on the neuronal representation of 1 kHz, MEG recordings to the test stimulus (band-passed noise bursts centred at 1 kHz) were compared to MEG recordings to the control stimulus (band-passed noise bursts centred at 0.5 kHz; cf. Figure 25.1a). The experiment was repeated three times in each subject on consecutive days in order to address the time course and the reversibility of cortical remodelling induced by this procedure. These repetitions served to investigate the dynamics of cortical reorganization over a period of 24 h. The auditory evoked fields (AEF, channels 12 and 35, depicting the maximum and minimum of the AEF, respectively) obtained for test and control stimuli before and after listening to notched music are shown for one representative subject in Figure 25.1b. Whereas the AEF amplitudes measured before and after listening to notched music are almost the same for the control stimulus, the AEF amplitude for the test stimulus is about 10 per cent smaller after listening to notched music than before. Figure 25.1c demonstrates the differences of the strength of the cortical sources before and after listening to the notched music for the test and for the control stimulus, averaged over all subjects and days. Whereas the strength of the cortical source decreased significantly after the listening to the notched music for the test stimulus ($p < 0.01$), this value did not change appreciably for the control stimulus. In order to provide information on the reversibility of the notching effect within the time period of 24 h, the measurements taken before listening to notched music on each day were compared. Between days, no significant differences between the strength of the cortical sources were observed for these measures, a result that proves the reversibility of the short-term plasticity effect provided by the 'functional deafferentation' with respect to listening to the notched music.

Taken together, our results suggest that reorganization of cortical representations can occur within time periods as short as a few hours following functional deafferentation of the adult human auditory cortex. The temporal properties of the notching effect are consistent with animal studies that have shown that cortical neurons deafferented by cochlear lesions display elevated response thresholds initially and then shift their tuning preferences away from the lesioned area to frequencies near the edge of the deafferented region over a period of 1–3 h or longer.[23] Several interrelated mechanisms appear to contribute to cortical remodelling induced by deafferentation, including (1) changes in the efficacy of existing excitatory synapses unmasked by lesioning,[26] (2) modification of synaptic efficacy by transcription of immediate early genes,[27] and (3) the sprouting of new connections.[28] Of these mechanisms the first and second appear to be probable candidates to account for our findings. Synaptogenesis may occur within hours,[29] but axonal sprouting and dendritic growth may require more time and hence would be less likely to be involved.

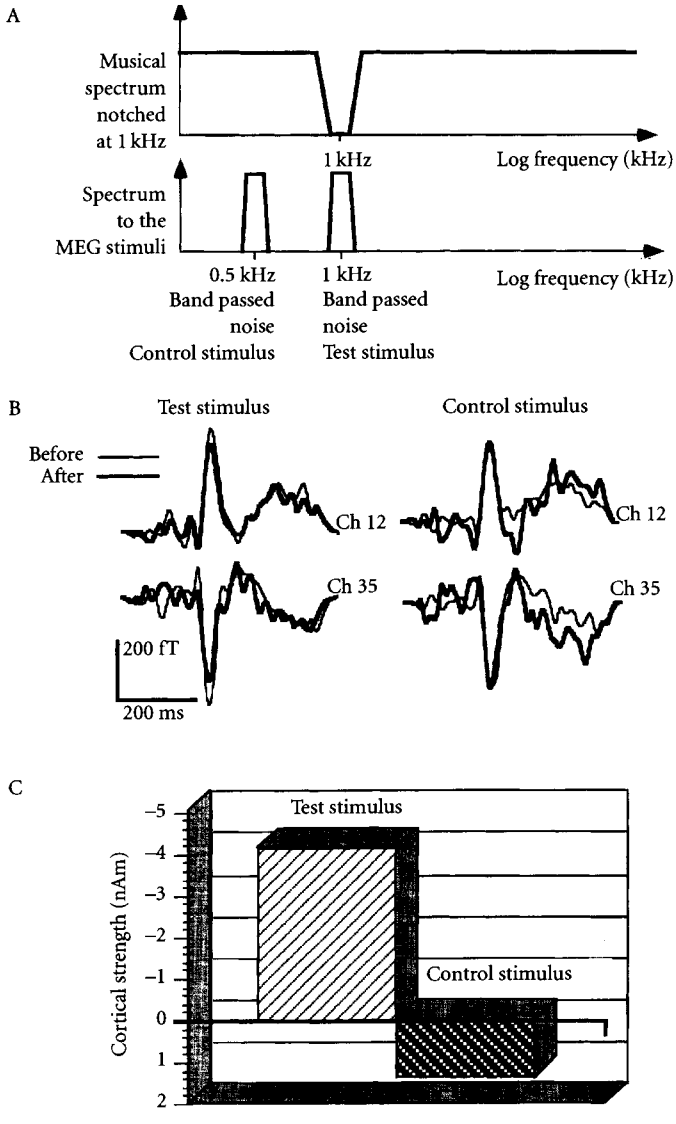



**Figure 25.1** Short-term plasticity effect of the auditory cortex induced by notched music. (A) Music spectrum notched at 1 kHz (*top*) and spectral characteristics of the test and control stimuli for MEG recording (*bottom*). (B) Auditory evoked fields (AEF) of a single subject recorded for test stimulus (1 kHz, *left panel*) and control stimulus (0.5 kHz, *right panel*) before and after listening to notched music. (C) Change in cortical representation or respectively in the strength of the cortical source produced by listening to the notched music and calculated as the moment (Q, nAm) of the equivalent current dipole fitted to the Nlm component of the AEF for test and control stimuli averaged over subjects and days. (Modified from Ref. 20.)

## Cortical plasticity and musical training

### Somatosensory representations

Two groups of subjects were examined. The first group consisted of nine musicians who played string instruments: six violinists, two violoncellists, and one guitarist. The subjects' average age was 24 ± 3 years. They had started to play their instruments 12 years earlier on average, ranging between 7 and 17 years. They practised on their instruments for an average of 9–10 h per week. The second group consisted of six control subjects who had never played a musical instrument and did not frequently carry out tasks that involved a systematic and rhythmic finger stimulation such as typing (on a typewriter or computer). All subjects were right-handed. The fingers D1 (the thumb) and D5 (the little finger) of both hands were consecutively excited with a brief, nonpainful standardized pneumatic pressure.

After excitation of the left-hand fingers D1 and D5, the strength of the cortical sources as determined with MEG were stronger in musicians than the corresponding sources in control subjects (Figure 25.2). Thus, the cerebral representation was increased representing the excitation of those fingers that are intensively used in string instrument musicians. This effect was particularly pronounced for the fifth digit (D5). The cortical representation of the left thumb was also enhanced, but not as strongly as the one for left D5. Cortical representations obtained for right-hand stimulation did not differ between control subjects and musicians.

The amount of increase in somatosensory cortical representations of left-hand fingers in musicians depended on the age at which the musicians had started to play their instrument (Figure 25.3). The cortical response for stimulation of D5 was greater in those musicians who had begun to play their instrument earlier. Among the musicians who started to play violin or violoncello later (after the age of 13 years), the cortical representations for D5

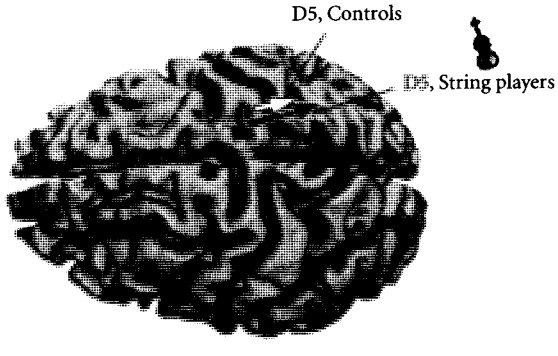



Figure 25.2 Equivalent current dipoles (ECD) elicited by stimulation of the digit 5 (D5) of the left hand of control subjects and string players, superimposed onto an MRI-reconstruction of the cerebral cortex of a control subject, selected to provide anatomical landmarks for the interpretation of the MEG-based localization. The arrows represent the location and orientation of the ECD vector for D5 averaged across musicians (*black arrow*) and control subjects (*white arrow*). The length of the arrows represents the mean magnitude of the dipole moment for D5 in each group. The dipole moment is larger for the musicians' D5 as indicated by the greater magnitude of the black arrow. (Modified from Ref. 39.)

## Cortical plasticity and musical training

### Somatosensory representations

Two groups of subjects were examined. The first group consisted of nine musicians who played string instruments: six violinists, two violoncellists, and one guitarist. The subjects' average age was 24 ± 3 years. They had started to play their instruments 12 years earlier on average, ranging between 7 and 17 years. They practised on their instruments for an average of 9–10 h per week. The second group consisted of six control subjects who had never played a musical instrument and did not frequently carry out tasks that involved a systematic and rhythmic finger stimulation such as typing (on a typewriter or computer). All subjects were right-handed. The fingers D1 (the thumb) and D5 (the little finger) of both hands were consecutively excited with a brief, nonpainful standardized pneumatic pressure.

After excitation of the left-hand fingers D1 and D5, the strength of the cortical sources as determined with MEG were stronger in musicians than the corresponding sources in control subjects (Figure 25.2). Thus, the cerebral representation was increased representing the excitation of those fingers that are intensively used in string instrument musicians. This effect was particularly pronounced for the fifth digit (D5). The cortical representation of the left thumb was also enhanced, but not as strongly as the one for left D5. Cortical representations obtained for right-hand stimulation did not differ between control subjects and musicians.

The amount of increase in somatosensory cortical representations of left-hand fingers in musicians depended on the age at which the musicians had started to play their instrument (Figure 25.3). The cortical response for stimulation of D5 was greater in those musicians who had begun to play their instrument earlier. Among the musicians who started to play violin or violoncello later (after the age of 13 years), the cortical representations for D5

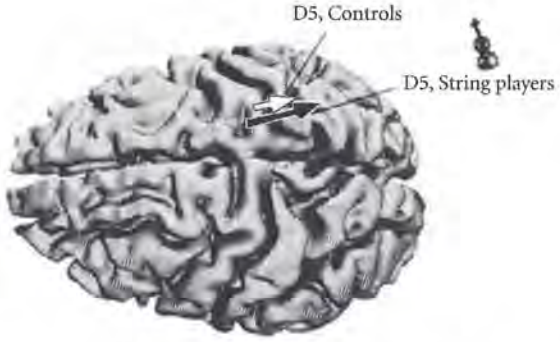



**Figure 25.2** Equivalent current dipoles (ECD) elicited by stimulation of the digit 5 (D5) of the left hand of control subjects and string players, superimposed onto an MRI-reconstruction of the cerebral cortex of a control subject, selected to provide anatomical landmarks for the interpretation of the MEG-based localization. The arrows represent the location and orientation of the ECD vector for D5 averaged across musicians (*black arrow*) and control subjects (*white arrow*). The length of the arrows represents the mean magnitude of the dipole moment for D5 in each group. The dipole moment is larger for the musicians' D5 as indicated by the greater magnitude of the black arrow. (Modified from Ref. 39.)

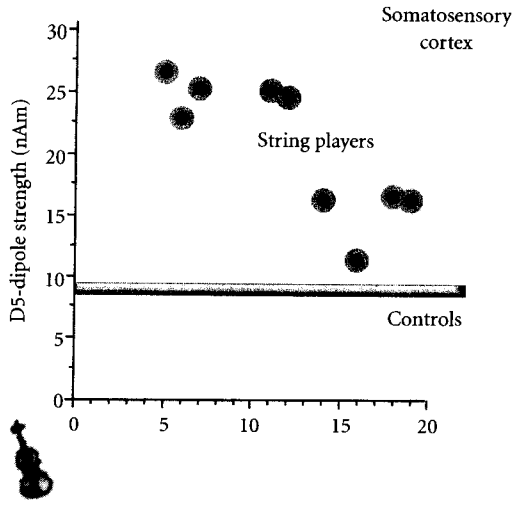



Figure 25.3 The magnitude of the D5 cortical strength as a function of the age of inception of musical practice: string players (*circles*), mean value for control subjects (*line*). Note the larger dipole moment for individuals beginning musical practice before the age of 12. (Modified from Ref. 39.)

were somewhat smaller, but still larger than those observed in the control subjects who had never played an instrument.

### Auditory representation

Three groups of subjects were examined. The first group consisted of musicians with absolute pitch ($n = 9$), the second group of musicians with relative pitch ($n = 11$), and the third group of subjects were controls who had never played a musical instrument ($n = 13$). In the first group, the mean number of years the musicians had played their instruments was $21 \pm 6$, in the second group $15 \pm 3$ years. The musicians were recruited from the conservatory in Münster. The musicians with absolute pitch reported practising their instruments for $27 \pm 14$ h per week, and the musicians with relative pitch $23 \pm 12$ h per week, during the 5 years preceding participation in the study. The musicians who claimed to have absolute pitch were subjected to a testing procedure that was developed according to criteria established in the musical psychological literature.[30–34] A random sequence of 35 piano tones from H2 to C7 was presented to the musicians. For our study, in order to be considered a musician with absolute pitch, correct recognition of at least 90 per cent of the tones was required. Musicians with relative pitch were either self-identified ($n = 9$) or were subjects who failed the test for absolute pitch ($n = 2$). The mean age in the three groups was $29 \pm 6$ years for musicians with absolute pitch, $26 \pm 5$ years for musicians with relative pitch, and $26 \pm 4$ years for control subjects. All participants were right-handed, as established by the Edinburgh handedness questionnaire[35] and did not have recent or past audiology, otology, or neurology medical history. Audiological status was normal at the time the study was performed for each subject.

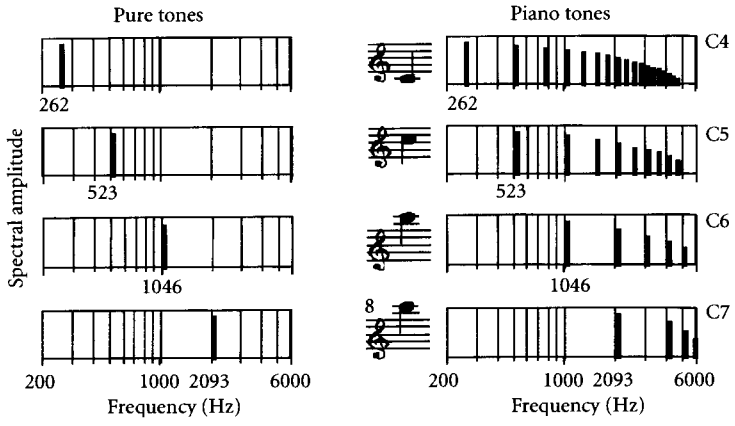



Figure 25.4 Spectra of piano tones C4–C7 (American notation, having fundamental frequency at 262, 523, 1046, and 2093 Hz, respectively) and four pure tones of the same fundamental frequencies. (Modified from Ref. 19.)

The acoustic stimulation consisted of a pseudorandom sequence of four piano tones (C4–C7; with the fundamental frequencies of 262, 523, 1046, and 2093 Hz), and four pure (sinusoidal) tones matching the fundamental frequencies (Figure 25.4). Each stimulus was presented 128 times to determine the location and strength of the electrical source in the auditory cortex.

In all musicians larger neuronal excitation was noted for piano tones as compared to pure tones. Our measure for the cortical neuronal representation, the strength of the electrical dipole moment in both musician groups, was significantly 25 per cent stronger for the piano tones as compared to the one for the pure tones ($p < 0.001$). By contrast, in the control subjects, who never played an instrument, no such difference was noted (Figure 25.5).

This finding demonstrates an increase in neuronal representation specific for the processing of the tones of the musical scale in musicians. Similar to the described enhancement of somatosensory representation, the degree of this increase was also dependent on the age at which the musicians had started to play their instruments. The earlier the initiation of musical practice, the stronger the neuronal response to the piano tones. In this study a marked difference was found between those who had begun practising music before and after 9 years of age.

The cortical representations for processing information from the fingers and the ears of musicians give an impressive demonstration that not only deafferentation, but also intensive training, can trigger a functional adaptation of the cortical organization. In general, this suggests that training can induce plastic changes of the adult human brain. In the musicians proficient in playing stringed instruments, the representation of the frequently used left-hand fingers is enlarged relative to control subjects. The more a given finger is stimulated, the larger the expansion of the cortical response. The representation for the thumb, which is not as frequently used as the other fingers, is not increased as much. However, the amount of practice is not the only factor that influences the organization of the somatosensory cortex. The plastic change of cortical representation is not directly

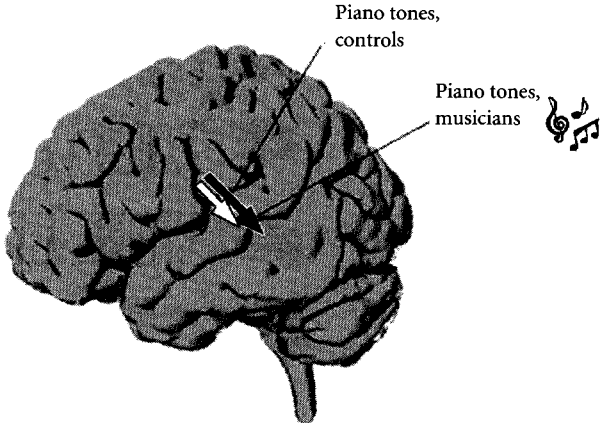



**Figure 25.5** Mean values of the strength of cortical activation during the auditory N1 evoked field for piano tones as compared to pure tones in control subjects and musicians with absolute or relative pitch. (Modified from Ref. 19.)

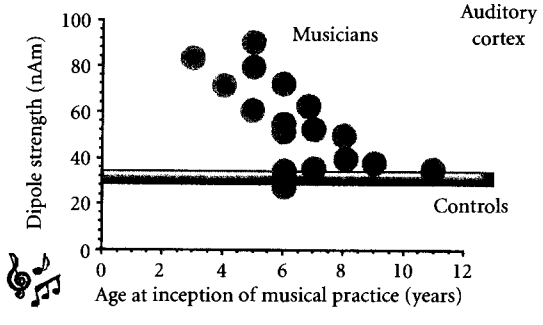



**Figure 25.6** Mean strength of cortical activation for the different piano tones as a function of the age at which musical training began in musicians with absolute or relative pitch. The line denotes mean dipole moment in control subjects who never played a musical instrument. (Modified from Ref. 19.)

correlated to the number of hours played per week, but rather to the age at which the training was started in childhood.

For the auditory system, the representation of pure tones, which are not part of our natural acoustic environment and also not encountered in musical training and practice, does not differ between musicians and control subjects. By contrast, the representation of piano tones, which in control subjects is of the same strength as the one for pure tones, is specifically enhanced in musicians. The analysis of reports of these musicians revealed that neither the passive exposure to music as a child nor the amount of time spent listening to music passively as an adult significantly influences functional cortical organization. It seems that active practice is necessary to induce plastic alterations. This reorganization of the auditory system also exhibits a trend towards more pronounced brain plasticity with a younger age of starting to play the musical instrument (Figure 25.6). The musicians who started playing before they were 9 years old demonstrated the largest cortical representations.

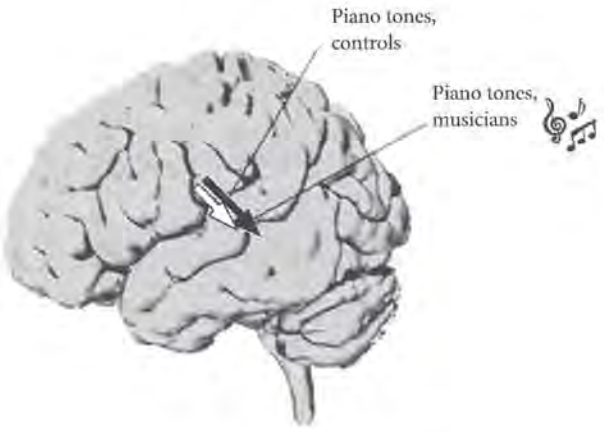



Figure 25.5 Mean values of the strength of cortical activation during the auditory N1 evoked field for piano tones as compared to pure tones in control subjects and musicians with absolute or relative pitch. (Modified from Ref. 19.)

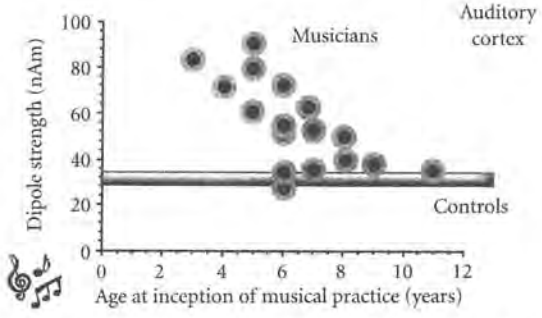



Figure 25.6 Mean strength of cortical activation for the different piano tones as a function of the age at which musical training began in musicians with absolute or relative pitch. The line denotes mean dipole moment in control subjects who never played a musical instrument. (Modified from Ref. 19.)

correlated to the number of hours played per week, but rather to the age at which the training was started in childhood.

For the auditory system, the representation of pure tones, which are not part of our natural acoustic environment and also not encountered in musical training and practice, does not differ between musicians and control subjects. By contrast, the representation of piano tones, which in control subjects is of the same strength as the one for pure tones, is specifically enhanced in musicians. The analysis of reports of these musicians revealed that neither the passive exposure to music as a child nor the amount of time spent listening to music passively as an adult significantly influences functional cortical organization. It seems that active practice is necessary to induce plastic alterations. This reorganization of the auditory system also exhibits a trend towards more pronounced brain plasticity with a younger age of starting to play the musical instrument (Figure 25.6). The musicians who started playing before they were 9 years old demonstrated the largest cortical representations.

The fact that the degree of expansion of the neuronal representations depends on the age training is started suggests that the enormous capacity of plasticity in the adult brain is, however, more limited than the plasticity of the infant brain. This corresponds to the prescientific observation that it is more difficult for adults to learn to play a musical instrument. To achieve an equivalent musical aptitude, adults can still adapt their cortical organization but they have to work harder to do so.

## The dark side of neuroplasticity in musicians

As mentioned above, neuronal networks are thought to be particularly plastic during 'sensitive periods' in the development of cortical structures. However, they maintain the ability to alter their architecture and function to afferent input throughout life. Perceptual and behavioural correlates of this reorganization indicate that neuroplasticity can be adaptive, as illustrated previously. Another type of reorganization includes the smearing of representational zones. An example has been described in blind multifinger Braille readers where overlap of digit representation may aid the ability to process different types of digital input simultaneously, thus enabling more rapid reading.[36] However, cortical reorganization can also be maladaptive, as is indicated by its association with phantom limb pain,[37] tinnitus,[38] and focal hand dystonia.[39,40] Intense training, as found in professional musicians, employs mechanisms of cortical plasticity with the inherent danger that maladaptive processes may also occur. Evidence from several laboratories suggests that one such danger is focal hand dystonia. Focal dystonia of the hand falls into a class often termed occupational hand cramp.[41,42] It involves manual discoordination occurring in individuals engaging in extensive or forceful use of the fingers. The disorder involves a disorganization of control of the digits such that the movements of some fingers on a hand become involuntarily linked to those of others, something particularly incapacitating to professional instrumental musicians.

Repetitive and behaviourally relevant stimulation to several digits performed with healthy subjects results in a fusion and in a disordered arrangement of the representation of individual digits.[36,43] At the same time subjects in these experiments consistently mislocalize light pressure stimuli applied to the fingertips. Consistent with earlier work in animals, source imaging in humans demonstrates that synchronous stimulation of the digits may create a fusion of cortical representational zones, whereas asynchronous stimulation leads to separation. Could such a process of cortical reorganization become maladaptive and thus contribute to focal dystonia? By means of magnetic and neuroelectric source imaging, an overlap or smearing of the homuncular organization of the representation of the digits in the primary somatosensory and the motor cortex was indeed observed (Figure 25.7).[39,44,45]

Work by N. Byl, M. Merzenich, and colleagues in New World monkeys has indicated that lack of digital motor coordination resulting from digital overuse is associated with an induced disorder in the representation of the digits in the somatosensory cortex.[42,46] These findings suggest that overuse-dependent central nervous system (CNS) plasticity is the basis of the focal hand dystonia. Therefore, because evidence suggests that behavioural usage and a CNS plastic response to this usage could be the cause of both the cortical

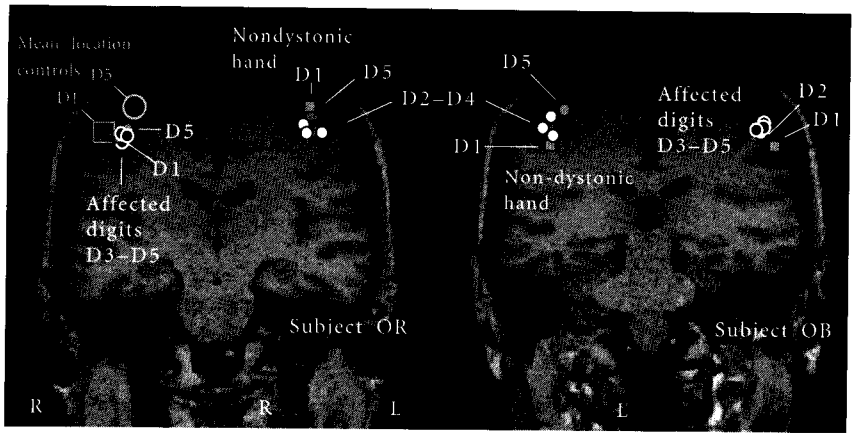



**Figure 25.7** A coronal MRI section is shown through the somatosensory cortices of an organ player (*left*) and an oboist (*right*), both suffering from focal hand dystonia. The symbols indicate cortical representations (dipole locations) of digits 1–5 (D1–D5) resulting from contralateral stimulation of the finger tips. The large *open symbols* in the right hemisphere of subject O.R. indicate the mean location of dipoles for D1 and D5 in normal control subjects. (Modified from Ref. 39.)

disorder and the involuntary lack of movement coordination, it was thought that a behavioural intervention focusing on movement could be of value in reducing or eliminating these conditions. Any such treatment approach should include practice in a behaviourally relevant setting. A number of studies have shown that the behavioural relevance of the stimulation is the critical factor for reorganization to occur, whereas simple passive stimulation seems to be ineffective in altering cortical connectivity.[17] For example, Kilgard *et al.* suggested that the cholinergic nucleus basalis might play an important role in labelling stimuli as behaviourally relevant because nucleus basalis activation facilitated cortical plasticity induced by correlated stimulation in the periphery.[47] Further evidence points to the additional involvement of dopaminergic pathways. It seems possible that blockade of dopamine, released by a potentially rewarding condition, for example, through neuroleptics, will not allow cortical reorganization to occur.

On the basis of these considerations, Candia *et al.* have developed a successful therapy for focal hand dystonia.[40] Professional musicians (pianists, guitarists, and wind players), with long-standing symptoms, who had previously received a variety of treatments, practised their instrument according to rules derived from insights into the mechanisms of cortical reorganization. The therapy involved immobilization by splints of one or more of the digits other than the focal dystonic finger. The focal dystonic finger was required to carry out repetitive exercises in coordination with one or more of the other digits for 1.5–2.5 h a day over an intended period of eight consecutive days under supervision of a therapist. The patients were instructed to continue the exercises using the splint at home for 1 h every day for 1 year posttreatment. The wind players, who, in effect, constituted placebo controls, did not improve substantially. However, each of the pianists and guitarists showed marked and significant improvement in spontaneous repertoire performance without the splint at the

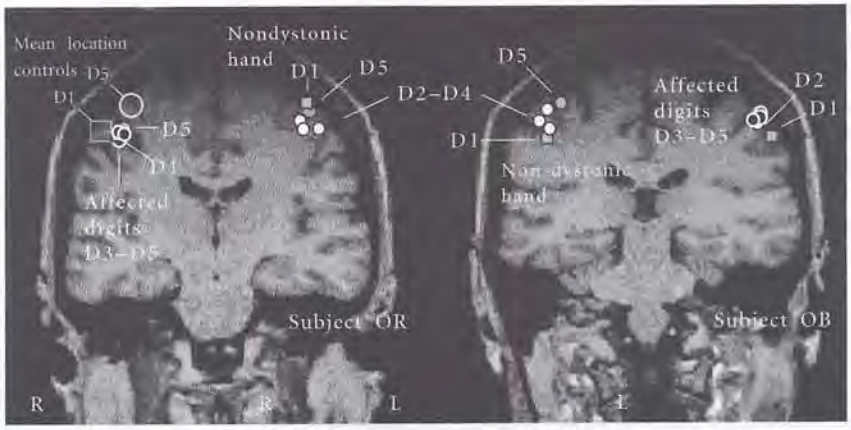



**Figure 25.7** A coronal MRI section is shown through the somatosensory cortices of an organ player (*left*) and an oboist (*right*), both suffering from focal hand dystonia. The symbols indicate cortical representations (dipole locations) of digits 1–5 (D1–D5) resulting from contralateral stimulation of the finger tips. The large *open symbols* in the right hemisphere of subject O.R. indicate the mean location of dipoles for D1 and D5 in normal control subjects. (Modified from Ref. 39.)

disorder and the involuntary lack of movement coordination, it was thought that a behavioural intervention focusing on movement could be of value in reducing or eliminating these conditions. Any such treatment approach should include practice in a behaviourally relevant setting. A number of studies have shown that the behavioural relevance of the stimulation is the critical factor for reorganization to occur, whereas simple passive stimulation seems to be ineffective in altering cortical connectivity.[17] For example, Kilgard *et al.* suggested that the cholinergic nucleus basalis might play an important role in labelling stimuli as behaviourally relevant because nucleus basalis activation facilitated cortical plasticity induced by correlated stimulation in the periphery.[47] Further evidence points to the additional involvement of dopaminergic pathways. It seems possible that blockade of dopamine, released by a potentially rewarding condition, for example, through neuroleptics, will not allow cortical reorganization to occur.

On the basis of these considerations, Candia *et al.* have developed a successful therapy for focal hand dystonia.[40] Professional musicians (pianists, guitarists, and wind players), with long-standing symptoms, who had previously received a variety of treatments, practised their instrument according to rules derived from insights into the mechanisms of cortical reorganization. The therapy involved immobilization by splints of one or more of the digits other than the focal dystonic finger. The focal dystonic finger was required to carry out repetitive exercises in coordination with one or more of the other digits for 1.5–2.5 h a day over an intended period of eight consecutive days under supervision of a therapist. The patients were instructed to continue the exercises using the splint at home for 1 h every day for 1 year posttreatment. The wind players, who, in effect, constituted placebo controls, did not improve substantially. However, each of the pianists and guitarists showed marked and significant improvement in spontaneous repertoire performance without the splint at the

end of treatment, up into the normal range. Neuroimaging results obtained in successfully treated subjects indicated normalization of the cortical representational maps, highlighting again the continued role of plasticity in the adult human brain for successful adaptation. The outcome demonstrates that learning-induced alterations in the functional architecture of the brain can be maladaptive, and that the resulting pathology can be treated using behavioural techniques that are based on learning principles that take into account recent research on neuroplasticity.

## Conclusions

By use of MEG, it was shown that it is possible to noninvasively study mechanisms of cerebral organization and reorganization in humans. Listening to music attentively for as little as 3 h can have a temporary influence on auditory cortical representation, as was shown in our study with music notched at a specific frequency. It was further shown that musical education and training is reflected in the organization of auditory and somatosensory representational cortex in musicians. From the results obtained, the following questions arise: Can we draw conclusions on the behavioural plasticity after knowing about the neuronal one? How do alterations in functional cortical organization affect musical accomplishment? Do the best virtuosos have the most atypical cortical organization?

The results of the studies presented here do not allow final conclusions. However, as shown in the case of focal dystonia, current research can lead to recommendations about musical practice and help design methods of treatment. The observed correlation between the age of initiation of musical practice and the amount of cortical change should not be overinterpreted from a music-pedagogical perspective. The complex relationships between genetic and familial, as well as educational pedagogical factors, would have to be studied by an interdisciplinary team of scientists. Such a team should comprise neuroscientists, as well as musicians and experts in music psychology, because it would be necessary to evaluate subtle differences in musical achievement to address these questions and establish causal links.

## Acknowledgements

This research was supported by the Deutsche Forschungsgemeinschaft. The work presented here was made possible through multiple collaborations with Drs Larry Roberts, Brigitte Rockstroh, Bernhard Ross, Edward Taub, and Christian Wienbruch. We wish to thank Deborah Urdang and Dr Harvey Salomon for comments on the manuscript.

## References

1. **Clark, S. A., *et al.*** (1988) Receptive fields in the body-surface map in adult cortex defined by temporally correlated inputs. *Nature* 332, 444–5.
2. **Narici, L., *et al.*** (1991) Neuromagnetic somatosensory homunculus: a non-invasive approach in humans. *Neurosci. Lett.* 121, 51–4.

3. **Elbert, T., *et al.*** (1995) The separation of overlapping neuromagnetic sources in first and second somatosensory cortices. *Brain Topogr.* 7, 275–82.
4. **Merzenich, M. M., P. L. Knight,** and **G. L. Roth** (1975) Representation of cochlea within primary auditory cortex in the cat. *J. Neurophysiol.* 38, 231–49.
5. **Pantev, C., *et al.*** (1988) Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. *EEG Clin. Neurophysiol.* 69, 160–70.
6. **Pantev, C., *et al.*** (1989) Tonotopic organization of the auditory cortex: pitch versus frequency representation. *Science* 246, 486–8.
7. **Pantev, C., *et al.*** (1995) Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *EEG Clin. Neurophysiol.* 94, 26–40.
8. **Hebb, D. O.** (1949) *The Organization of Behavior.* Chichester, UK: John Wiley and Sons.
9. **Hari, R.** (1990) Magnetic evoked fields of the human brain: basic principles and applications. *EEG Clin. Neurophysiol.* 41 (Suppl.), 3–12.
10. **Merzenich, M. M., *et al.*** (1984) Somatosensory cortical map changes following digit amputation in adult monkeys. *J. Comp. Neurol.* 224, 591–605.
11. **Merzenich, M. M., *et al.*** (1983) Topographic reorganization or somatosensory cortical areas 3b and 1 in adult monkeys following restricted deafferentation. *Neuroscience* 8, 33–55.
12. **Elbert, T., *et al.*** (1994) Extensive reorganization of the somatosensory cortex in adult humans after nervous system injury. *Neuroreport* 5, 2593–7.
13. **Elbert, T., *et al.*** (1997) Input-increase and input-decrease types of cortical reorganization after upper extremity amputation in humans. *Exp. Brain Res.* 117, 161–4.
14. **Rajan, R., *et al.*** (1993) Effect of unilateral partial cochlear lesions in adult cats on the representation of lesioned and unlesioned cochleas in primary auditory cortex. *J. Comp. Neurol.* 338, 17–49.
15. **Irvine, D. R. F.** and **R. Rajan** (1995) In G. A. Manley *et al.* (eds) *Advances in Hearing Research.* Singapore: World Scientific Publishing Co. pp. 3–23.
16. **Jenkins, W. M., *et al.*** (1990) Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation. *J. Neurophysiol.* 63, 82–104.
17. **Recanzone, G. H., C. E. Schreiner,** and **M. M. Merzenich** (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J. Neurosci.* 13, 87–103.
18. **Elbert, T., *et al.*** (1995) Increased cortical representation of the fingers of the left hand in string players. *Science* 270, 305–7.
19. **Pantev, C., *et al.*** (1998) Increased auditory cortical representation in musicians. *Nature* 392, 811–14.
20. **Pantev, C., *et al.*** (1999) Short-term plasticity of the human auditory cortex. *Brain Res.* 842, 192–9.
21. **Doetsch, G. S., *et al.*** (1996) Short-term plasticity in primary somatosensory cortex of the rat: rapid changes in magnitudes and latencies of neuronal responses following digit denervation. *Exp. Brain Res.* 112, 505–12.
22. **Gilbert, C. D.** (1997) Cortical dynamics. *Acta Paediatr. Suppl.* 422, 34–7.
23. **Robertson, D.** and **D. R. Irvine** (1989) Plasticity of frequency organization in auditory cortex of guinea pigs with partial unilateral deafness. *J. Comp. Neurol.* 282, 456–71.
24. **Gilbert, C. D., *et al.*** (1996) Spatial integration and cortical dynamics. *Proc. Natl. Acad. Sci. USA* 93, 615–22.

25. Rossini, P. M., *et al.* (1994) Short-term brain 'plasticity' in humans: transient finger representation changes in sensory cortex somatotopy following ischemic anesthesia. *Brain Res.* 642, 169–77.
26. Ziemann, U., M. Hallet, and L. G. Cohen (1998) Mechanisms of deafferentation-induced plasticity in human motor cortex. *J. Neurosci.* 18, 7000–7.
27. Kaczmarek, L. and A. Chaudhuri (1997) Sensory regulation of immediate-early gene expression in mammalian visual cortex: implications for functional mapping and neural plasticity. *Brain Res. Brain Res. Rev.* 23, 237–56.
28. Florence, S. L., H. B. Taub, and J. H. Kaas (1998) Large-scale sprouting of cortical connections after peripheral injury in adult macaque monkeys. *Science* 282, 1117–21.
29. Engert, F. and T. Bonhoeffer (1999) Dendritic spine changes associated with hippo-campal long-term synaptic plasticity. *Nature* 399, 66–70.
30. Bachem, A. (1937) Various types of absolute pitch. *J. Acoust. Soc. Am.* 9, 146–51.
31. Miyazaki, K. (1988) Musical pitch identification by absolute pitch possessors. *Percept. Psychophys.* 44, 501–12.
32. Neu, D. M. (1947) A critical review on the literature on 'absolute pitch'. *Psychol. Bull.* 44, 249–66.
33. Takeuchi, A. and S. H. Hulse (1993) Absolute pitch. *Psychol. Bull.* 113, 345–61.
34. Ward, W. D. (1998) In D. Deutsch (ed.) *The Psychology of Music.* San Diego, CA: Academic Press.
35. Oldfield, R. C. (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neurophysiologia* 9, 97–113.
36. Sterr, A., *et al.* (1998) Perceptual correlates of changes in cortical representation of fingers in blind multifinger Braille readers. *J. Neurosci.* 18, 4417–23.
37. Flor, H., *et al.* (1995) Phantom limb pain as a perceptual correlate of massive cortical reorganization in upper limp amputees. *Nature* 375, 482–4.
38. Muhlnickel, W., *et al.* (1998) Reorganization of auditory cortex in tinnitus. *Proc. Natl. Acad. Sci. USA* 95, 10340–3.
39. Elbert, T., *et al.* (1998) Alteration of digital representations in somatosensory cortex in focal hand dystonia. *Neuroreport* 9, 3571–5.
40. Candia, V., *et al.* (1999) Constraint-induced movement therapy for focal hand dystonia in musicians. *Lancet* 353, 42.
41. Lederman, R. G. (1988) Occupational cramp in instrumental musicians. *Med. Probl. Perform. Art* 3, 45–51.
42. Byl, N. N., M. M. Merzenich, and W. M. Jenkins (1996) A primate genesis model of focal dystonia and repetitive strain injury: I. Learning-induced dedifferentiation of the representation of the hand in the primary somatosensory cortex in adult monkeys. *Neurology* 47, 508–20.
43. Braun, C., *et al.* (2000) Differential activation in somatosensory cortex for different discrimination tasks. *J. Neurosci.* 20, 446–50
44. Elbert, T. and A. Keil (2000) Imaging in the fourth dimension. *Nature* 404, 29–31.
45. Bara-Jimenez, W., *et al.* (1998) Abnormal somatosensory homunculus in dystonia of the hand. *Ann. Neurol.* 44, 828–31.
46. Byl, N. N., *et al.* (1997) A primate model for studying focal dystonia and repetitive strain injury: effects on the primary somatosensory cortex. *Phys. Ther.* 77, 269–84.
47. Kilgard, M. P. and M. M. Merzenich (1998) Cortical map reorganization enabled by nucleus basalis activity. *Science* 279, 1714–18.

CHAPTER 26

# THE BRAIN THAT MAKES MUSIC AND IS CHANGED BY IT

ALVARO PASCUAL-LEONE

## Abstract

Playing a musical instrument demands extensive procedural and motor learning that results in plastic reorganization of the human brain. These plastic changes seem to include the rapid unmasking of existing connections and the establishment of new ones. Therefore, both functional and structural changes take place in the brain of instrumentalists as they learn to cope with the demands of their activity. Neuroimaging techniques allow documentation of these plastic changes in the human brain. These plastic changes are fundamental to the accomplishment of skillful playing, but they pose a risk for the development of motor control dysfunctions that may give rise to overuse syndromes and focal, task-specific dystonia.

**Keywords:** Brain plasticity; Musical training

## Introduction

> *The most intricately and perfectly coordinated of all voluntary movements in the animal kingdom are those of the human hand and fingers, and perhaps in no other human activity do memory, complex integration, and muscular coordination surpass the achievements of the skilled pianist.*
>
> Homer W. Smith, *From Fish to Philosopher*

Playing a musical instrument requires more than factual knowledge about the musical instrument and the mechanics of how it is played. For example, given complete information about hand position, finger motions, and sequence of keys to push for how long and with what force, I would still be unable to play even the simplest piano sonata. The central nervous system has to acquire and implement a 'translation mechanism' to convert knowledge into action. These translation capabilities constitute the skill that enables the pianist to act on memory systems, select the relevant facts, choose the proper response goals, activate the necessary sensorimotor structures, and execute the sonata successfully. We generally think of such a skill as being acquired with practice. The pianist confronted with a new composition, after understanding the task and its demands, develops a cognitive representation of it and initiates a first, centrally guided response that results in sensorimotor feedback and movement correction. It seems certain that both sensory and motor aspects have

to be exquisitely coordinated. At the beginning, the limbs move slowly, with fluctuating accuracy and speed, and success requires visual, proprioceptive, and auditory feedback. Eventually, each single movement is refined, the different movements chained into the proper sequence with the desired timing, a high probability of stability in the ordered sequence attained, and a fluency of all movement developed. Only then can the pianist shift his or her attentional focus away from the mechanical details of the performance towards the emotional content of the task. We can think of the acquisition of such a skill as the conversion of declarative knowledge (facts) into procedural knowledge (actions, skills).[1–3]

Learning and memory might be considered integral parts of all the operations of any neural circuit, a concept for which Fuster[4] recently made an eloquent and convincing argument. In this view, 'perception and action are phenomena of memory and, conversely, memory is an integral part of perceptual and motor processing' (Ref. 4, p. 21). The nervous system comes to be viewed as a dynamic, dialectic organization in which plasticity is an intrinsic property that relates to the acquisition of new memories and skills as an obligatory consequence of perceptions and motor actions. To play an instrument, the nervous system is modified as a consequence of practice to yield the necessary changes in ability. We refer to this experience-dependent modification in neural structure as plasticity. These changes take place both in sensory and motor systems as well as in their interface. The consequence of this notion is that these changes do not necessarily represent behavioural benefits to the subject but might in fact be misguided and functionally deleterious. The development of noninvasive imaging and neurophysiologic techniques enables us to pursue the study of such changes in humans.

## Learning to play the piano changes your brain

> *... the work of a pianist... is inaccessible for the untrained human, as the acquisition of new abilities requires many years of mental and physical practice. In order to fully understand this complicated phenomenon it is necessary to admit, in addition to the strengthening of pre-established organic pathways, the establishment of new ones, through ramification and progressive growth of dendritic arborizations and nervous terminals.... Such a development takes place in response to exercise, while it stops and may be reversed in brain spheres that are not cultivated.*
>
> Santiago Ramón y Cajal, *Textura del Sistema Nervioso del Hombre y de los Vertebrados*

Normal subjects were taught to perform with one hand a five-finger exercise on a piano keyboard connected by a MIDI interface to a computer.[5] Subjects did not play any musical instrument, did not know how to typewrite using all fingers, and held jobs not demanding skillful hand and finger activities. The exercise required pressing a piano key sequentially with thumb (C), index finger (D), middle finger (E), ring finger (F), little finger (G), ring finger (F), middle finger (E), index finger (D), thumb (C), index finger (D), and so forth. The subjects were instructed to attempt to perform the sequence of finger movements fluently, without pauses and without skipping any key, while paying particular attention to keep the interval between the individual key presses constant and the duration of each key press the same. A metronome gave a tempo of 60 beats per minute, which the subjects were asked to aim for. Subjects performed the exercise under auditory feedback. They were studied on five

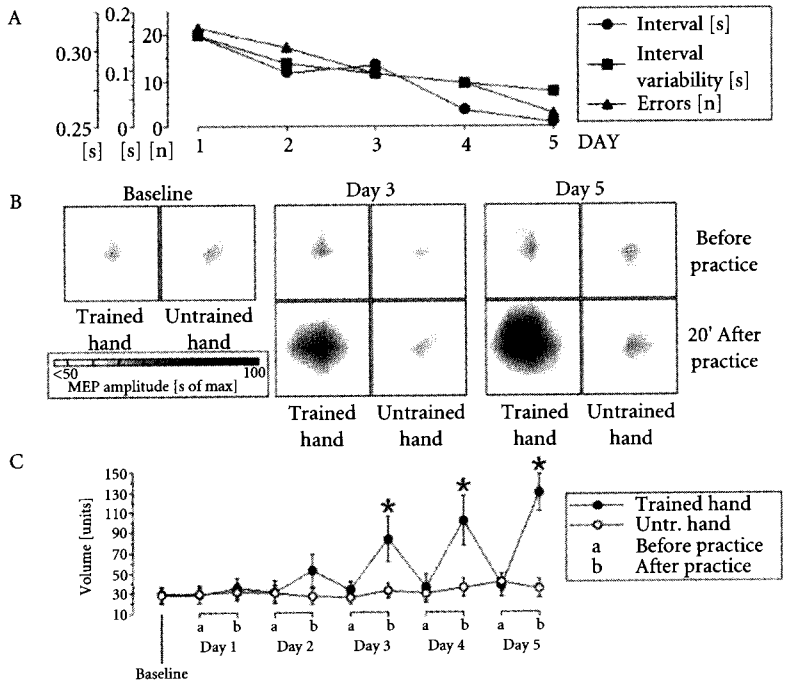



Figure 26.1 (A) Behavioural results of the week-long practice of the five-finger exercise on the piano. Interval between key presses, variability of these intervals, and number of errors in 20 repetitions of the exercise all show a highly significant decrease. This documents learning for all subjects studied. (B) Cortical output maps for the finger flexors of the trained and untrained hands of a representative subject. (See text and Ref. 5 for details on mapping method.) Note the marked changes of the output maps for the trained hand and the lack of changes for the untrained hand and the maps obtained for either hand before the daily practice sessions. (C) Graphic display of mean (± SD) volumes of the cortical output maps for all subjects studied. Note the significant (*star*) changes in cortical output maps for the trained hands after practice sessions on days 3–5.

consecutive days, and each day they had a two-hour practice session followed by a test. The test consisted of the execution of 20 repetitions of the five-finger exercise and was eventually analysed with the computer for the exact sequence of key presses, the interval between keys, and the duration and velocity of each key press. Following each test the subjects were given feedback about their performance and tips about how to improve. Figure 26.1 illustrates the great improvement in the subjects' playing skill over the course of the five study days. The number of sequence errors and the duration and variability of the intervals between key presses (as marked by the metronome beats) all decreased significantly.

The aim of the experiment was to correlate skill acquisition with changes in motor cortical output. Therefore, before the first practice session on the first day of the experiment, we used focal transcranial magnetic stimulation (TMS) to map the motor cortical areas targeting long finger flexor and extensor muscles bilaterally. Thereafter, this mapping was repeated daily. The baseline maps can be compared with those obtained during the five

days. As the subjects' performance improved, the threshold for TMS activation of the finger flexor and extensor muscles decreased steadily; even taking into account this change in threshold, the size of the cortical representation for both muscle groups increased significantly (Figure 26.1).[5] However, this increase could be demonstrated only when the cortical mapping studies were conducted following a 20- to 30-minute rest after the practice (and test) session. No such modulation in the cortical output maps was noted when maps were obtained before each daily practice session (Figure 26.1).[6]

When a near-perfect level of performance was reached at the end of a week of daily practice, subjects were randomized into two groups (Figure 26.2). Group 1 continued daily practice of the same piano exercise during the following four weeks. Group 2 stopped practising. During the four weeks of follow-up, cortical output maps for finger flexor and extensor muscles were obtained in all subjects on Mondays (before the first practice session of that week in group 1) and on Fridays (at the end of the last practice session for the week in group 1). In the group that continued practising (group 1), the cortical output maps obtained on Fridays showed an initial plateau and eventually a slow decrease in size despite continued performance improvement (Figure 26.2). On the other hand, maps obtained on Mondays before the practice session and after the weekend rest showed a small change from baseline with a tendency to increase in size over the course of the study (Figure 26.2). In group 2, who stopped practising after one week, the maps returned to baseline after the first week of follow-up and remained stable thereafter.

This experiment reveals that acquisition of the necessary motor skills to perform a five-finger movement exercise correctly is associated with reorganization in the cortical motor outputs to the muscles involved in the task. There are two main mechanisms to explain this reorganization: establishment of new connections, or sprouting, and unmasking of previously existing connections. The rapid time course in the initial modulation of the motor outputs, by which a certain region of motor cortex can reversibly increase its influence on a motoneuron pool, is most compatible with the unmasking of previously existing connections.[7–10] Supporting this notion, the initial changes are transient, demonstrable after practice, but return to baseline after a weekend of rest. We suggest that such flexible, short-term modulation represents a first and necessary step, leading to longer-term structural changes in the intracortical and subcortical networks as the skill becomes overlearned and automatic.

It is important to realize that our TMS mapping technique demonstrates a trace or memory of the activation of the motor cortical outputs that took place during the performance of the task rather than the activation during the task itself, as would be the case with neuroimaging studies. Long-term potentiation has been demonstrated in the motor cortex,[11,12] and our results might reveal a similar phenomenon. During the learning of the task, the cortical output maps obtained after task performance show a progressive increase in size, suggesting that skill acquisition is associated with a change in the pattern of activation of the executive structures. These changes are not demonstrable before the task performance, and we might hypothesize that 'learning' to activate the cortical outputs appropriately (in this case, activating a progressively larger cortical output map) constitutes the neurophysiological correlate of performance improvement.

As the task becomes overlearned over the course of five weeks, the pattern of cortical activation for optimal task performance might change as other neural structures take a

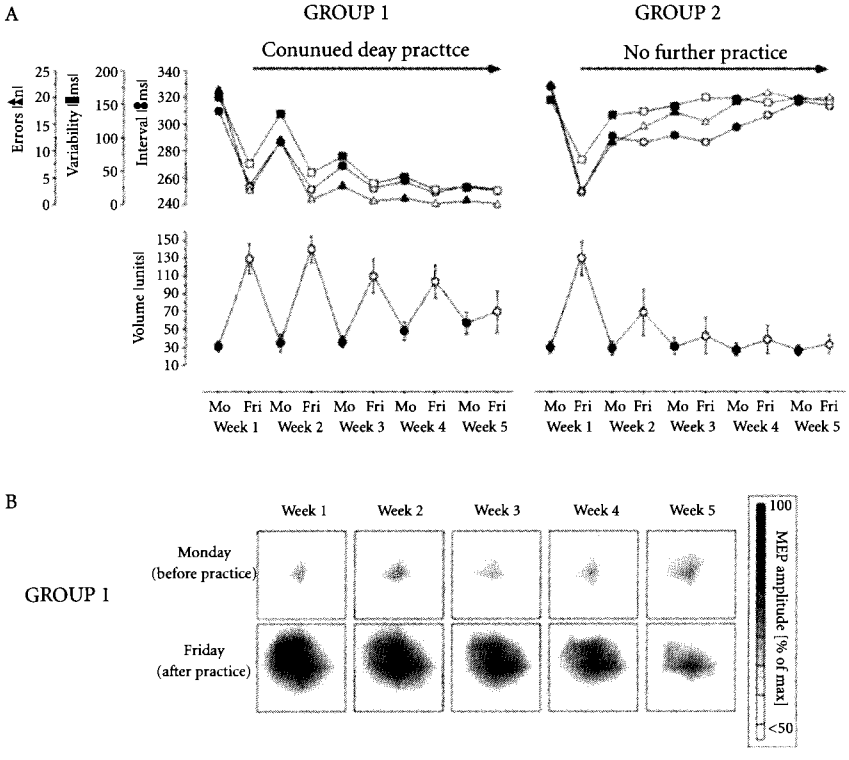



**Figure 26.2** (A) Behavioural results for subjects in groups 1 (continued practice after 1 week) and 2 (practice limited to 1 week) learning the five-finger exercise on the piano. Interval between key presses, variability of these intervals, and number of errors in 20 repetitions of the exercise are shown. Data from Monday and Friday of each week are shown. (B) Cortical output maps for the finger flexors of the trained hand of a representative subject. (See text and Ref. 5 for details on mapping method.) Note the marked changes in the output maps. There appear to be two parallel processes in action, one accounting for the rapid modulation of the maps from Monday to Friday and the other responsible for the slow and more discrete changes in Monday maps over time.[5,6]

more leading role in task performance. This might result in the decreasing size of the cortical output maps after practice and the increasing size before practice. For example, if basal ganglia play a more important role in driving task performance, the changed activity in the basal ganglia motor circuit might enhance thalamocortical connections. This may account for changes in the cortical motor representation, which would be more stable than that observed during the initial acquisition of a skill. Greenough *et al.*[13] showed that motor training is associated with changes in the dendritic branching patterns of motor and sensory cortical cells involved in the performance of a task. Sprouting may account for plastic changes in such situations, as likely occurred in monkeys deafferented for 10 years,[14] and represent the correlate of long-standing 'memories'. Neuroimaging studies suggest a similar phenomenon.[15–22] Related changes in the pattern of cortical activation and the resulting

modulation of cortical outputs might be induced by changes in the subject's strategy as explicit learning mechanisms are engaged.[23]

Our findings stress the role of the primary motor cortex (M1) in skill acquisition. It is not unreasonable to expect plastic changes in M1 during motor skill learning, because M1 is clearly involved in movement, and its cells have complex patterns of connectivity, including variable influences on multiple muscles within a body part.[24,25] Recent animal studies also illustrate the importance of the M1 in skill learning. For example, increases in excitability of primary motor cortex neurons have been found during conditioning,[26,27] and repeated activation of somatosensory inputs into the motor cortex results in long-term potentiation of motor neurons.[11,12] Even so, at the low stimulus intensities used in our mapping studies, TMS activates cortical cells largely transsynaptically.[28,29] Therefore, the demonstrated modulation in motor cortical outputs might be conditioned by changes in premotor areas projecting to M1, rather than by changes in M1 itself.

## If you cannot do it, at least think about it: mental practice

In the quote reproduced above, Cajal talks about rapid and slower plastic changes in the brain in the context of practice. Our results confirm his intuitions. Furthermore, Cajal writes about physical and *mental* practice. Might the latter result in plastic brain changes similar to those induced by the former?

Mental practice is the imagined rehearsal of a motor act with the specific intent of learning or improving it, without overt movement output. Mental practice can be viewed as a virtual simulation of behaviour by which the subject develops and 'internally' rehearses a cognitive representation of the motor act. When confronted with a new motor task, the subject must develop a cognitive representation of it and initiate a centrally guided response, which secondarily can be improved using sensorimotor feedback. Mental practice may accelerate the acquisition of a new motor skill by providing a well-suited cognitive model of the demanded motor act in advance of any physical practice.[30,31]

Mental practice has found wide acceptance in the training of athletes.[32] Musicians also have long recognized the benefit of mental rehearsal. Harold Schoenberg, in his fascinating books on virtuoso intrumentalists,[33,34] provides ample information on this topic. For example, Horowitz is supposed to have practised mentally before concerts to avoid disturbing his motor skills by the feedback of pianos other than his own Steinway. Rubinstein, eager to enjoy life to its fullest and dedicate as little time to practise as possible, found mental rehearsal the best way to minimize the number of hours spent sitting in front of the piano while maintaining his skill.

Using the same experimental design as described above for the five-finger exercise, we studied subjects who, instead of practising at the keyboard 2 h daily for five days spent time at the keyboard visualizing, rather than executing, the movements.[5] They were told to repeat the movement mentally, as if they were playing, but without moving their fingers. They could rest their fingers on the piano keyboard, but the lack of voluntary movements was monitored using electromyography and video. Such mental practice resulted in a reorganization of the motor outputs to finger flexor and extensor muscles similar to the one observed in the group of subjects who physically practised the movements and led to

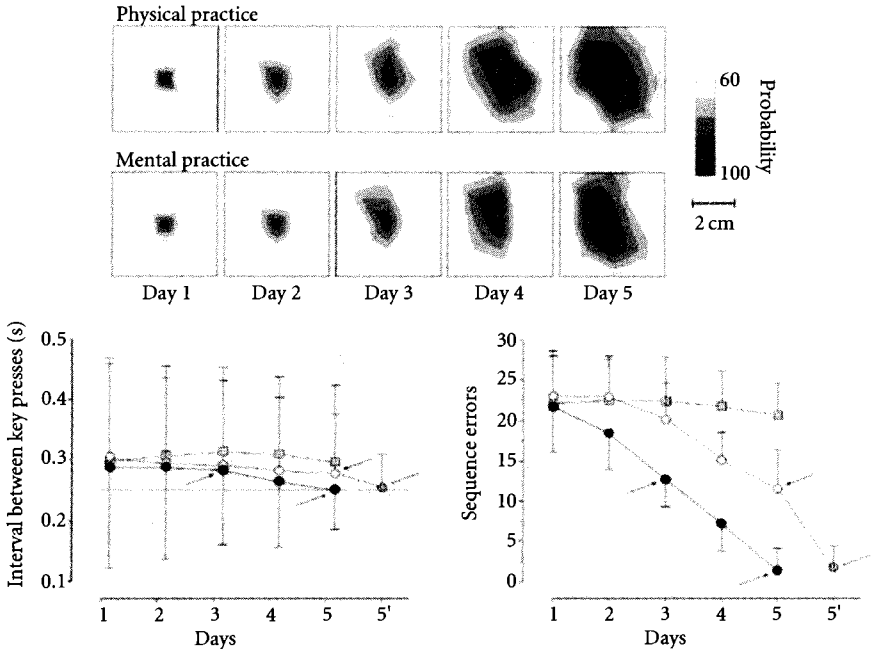



**Figure 26.3** (A) Cortical output maps for the finger flexors of the trained hand in a representative subject with physical *vs* mental practice alone. Modified from Ref. 5; see also text for details. Note the parity of changes with either form of practice. (B) Graphic representation of behavioural data in a group of control subjects (no practice, *squares*) and a group of test subjects performing mental practice (*open circles*) vs physical practice (*filled circles*). Note the behavioural advantage of the physical practice group, but the significant improvement of the mental practice group as compared with controls. The last data point (5′) represents the performance achieved by the mental practice group after 5 min of physical practice at the end of day 5. Note the apparent advantage of mental practice, because subjects go from a performance equivalent to day 3 of the physical practice group to matching their day 5 performance with only 5 min of practice. Modified from Ref. 5.

similar improvement in their ability to perform the five-finger movement exercise (Figure 26.3).

Studies of regional cerebral blood flow (rCBF) suggest that the prefrontal and supplementary motor areas, basal ganglia, and cerebellum are part of the network involved in the mental simulation of motor acts.[35–38] Therefore, mental simulation of movements activates some of the same central neural structures required for the performance of the actual movements. In so doing, mental practice alone seems to be sufficient to promote the modulation of neural circuits involved in the early stages of motor skill learning. This modulation not only results in marked improvement in performance, but also seems to place the subjects at an advantage for further skill learning with minimal physical practice.[5,6] The combination of mental and physical practice leads to greater performance improvement than does physical practice alone, a phenomenon for which our findings provide a physiological explanation.

## The risk of the change: sensorimotor mismatch and task-induced dystonia

As we have seen, skill acquisition requires plastic changes in the brain. This plastic reorganization is driven by efferent demand and afferent input. However, a system capable of such flexible reorganization harbours the risk of unwanted change. Increased demand of sensorimotor integration poses such a risk. We can postulate that faulty practice may result in unwanted cortical rearrangement and set the stage for motor control problems such as overuse syndrome and focal, task-specific dystonias. The style of piano playing—for example, the Russian vs the German school—seems to play a critical role in the risk of development of motor control problems. Forceful playing with the fingers bent and executing hammer-like movements is more frequently associated with overuse syndrome and dystonia than is softer playing with extended fingers 'caressing' the keys. This stresses the importance of proper, well-guided practice and illustrates the need for greater understanding of the neurobiology underlying music playing to define what proper practice actually is.

Focal hand dystonia in musicians is a strongly task-related movement disorder that can end an instrumentalist's career. Typically, symptoms become manifest only when players execute specific overpracticed skilled exercises on their instrument. Suddenly, a finger moves involuntarily, voluntary motor control is lost, the muscles tense up excessively, and pain develops. Playing is disturbed. For years, focal, task-specific dystonias were thought to be psychiatric in nature. It seemed too bizarre that involuntary muscle contraction might occur when playing a certain passage but not with any other activity or when playing on a certain instrument (organ) but not another (pianoforte). We now know that dystonias are neurologic involuntary movements due to disturbances in motor programs.[39] What, however, is their underlying pathophysiology?

We examined five guitarists with functional magnetic resonance imaging (fMRI) during dystonic symptom provocation by means of an adapted guitar inside the magnet.[40] As reference, we used the activation pattern obtained in the same subjects during other hand movements and in matched guitar players without dystonia during execution of the same guitar playing exercises. A 1.5-Tesla system equipped with echo-speed gradients and single-shot echoplanar imaging (EPI) software was used. Data acquisition was centred on the cortical motor system encompassed in eight contiguous slices.

Dystonic musicians compared in both control situations showed significantly greater activation of the contralateral primary sensorimotor cortex, which contrasted with conspicuous bilateral underactivation of premotor areas (Figure 26.4). Our results agree with studies of other types of dystonia in that they show abnormal recruitment of cortical areas involved in the control of voluntary movement. They do suggest, however, that rather than being hypoactive in idiopathic dystonic patients, the primary sensorimotor cortex may be overactive when tested during full expression of the task-induced movement disorder.

Although the primary manifestation of dystonia is abnormal motor function, evidence is increasing for a dysfunction of sensory processing that may be an associated or contributing factor.[39,41–43,46,47] In fact, dedifferentiation of the normally independent sensory representations of multiple digits may be a causative element in the etiology of dystonia.[44,45] For example, in musicians, extensive practice of co-ordinated hand postures in which

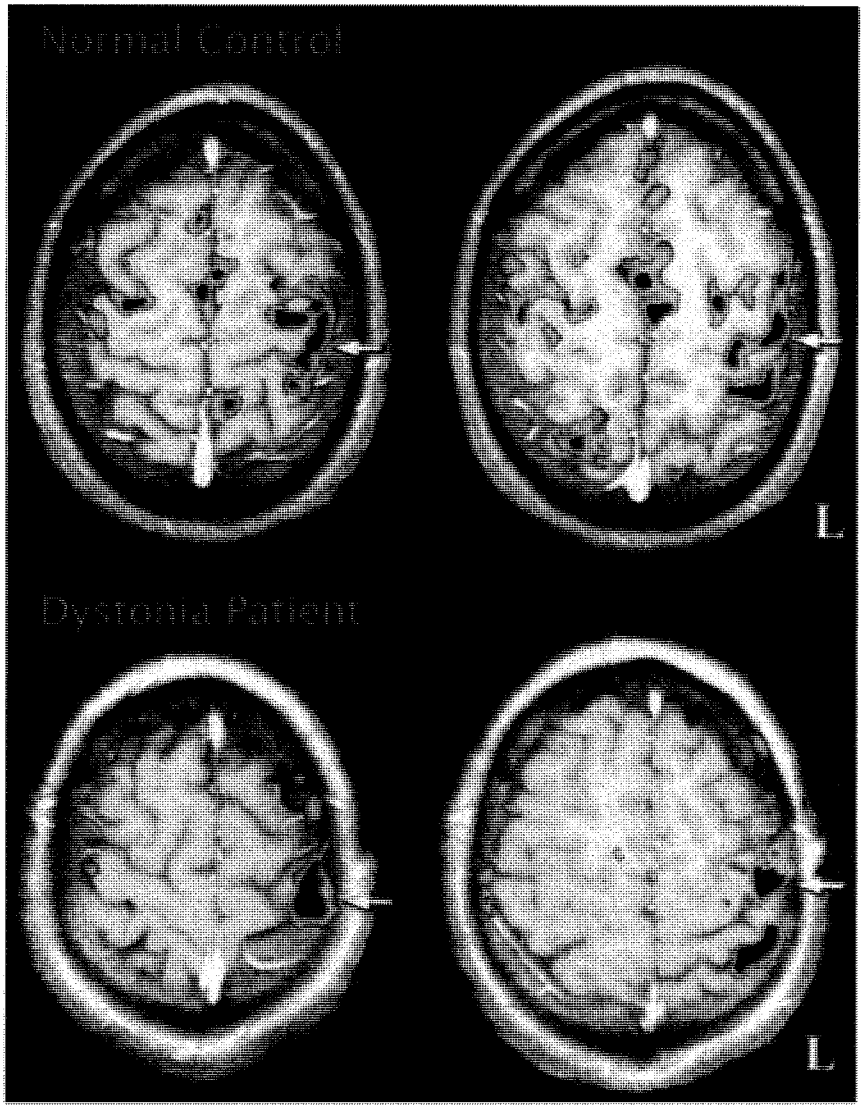



**Figure 26.4** The bold fMRI images of a normal and dystonic guitar player executing right hand arpeggios in the scanner are displayed. Note the greater activation of the sensorimotor cortex (*arrows*) and the lack of activation of premotor and supplementary motor cortices in the dystonic patient. Modified from Ref. 40. (See Plate 24 in colour section.)

various digits function as a unit, such as arpeggios, could eventually induce changes in the sensory representation of the hand with blurring of the segregation of different digits. This might be the case particularly when small repeated traumas are added, as in forceful, 'hammer-finger' piano playing. Disorganization and consequent confusion of sensory

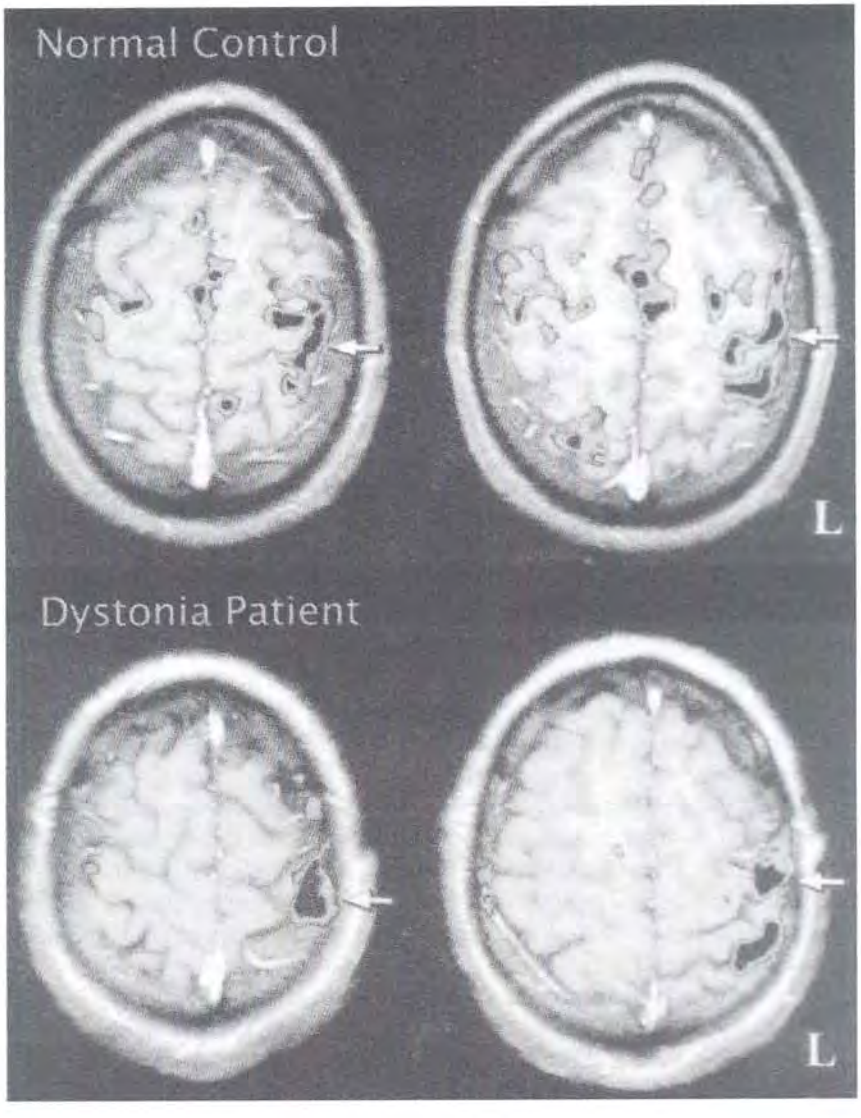



**Figure 26.4** The bold fMRI images of a normal and dystonic guitar player executing right hand arpeggios in the scanner are displayed. Note the greater activation of the sensorimotor cortex (*arrows*) and the lack of activation of premotor and supplementary motor cortices in the dystonic patient. Modified from Ref. 40. (See Plate 24 in colour section.)

various digits function as a unit, such as arpeggios, could eventually induce changes in the sensory representation of the hand with blurring of the segregation of different digits. This might be the case particularly when small repeated traumas are added, as in forceful, 'hammer-finger' piano playing. Disorganization and consequent confusion of sensory

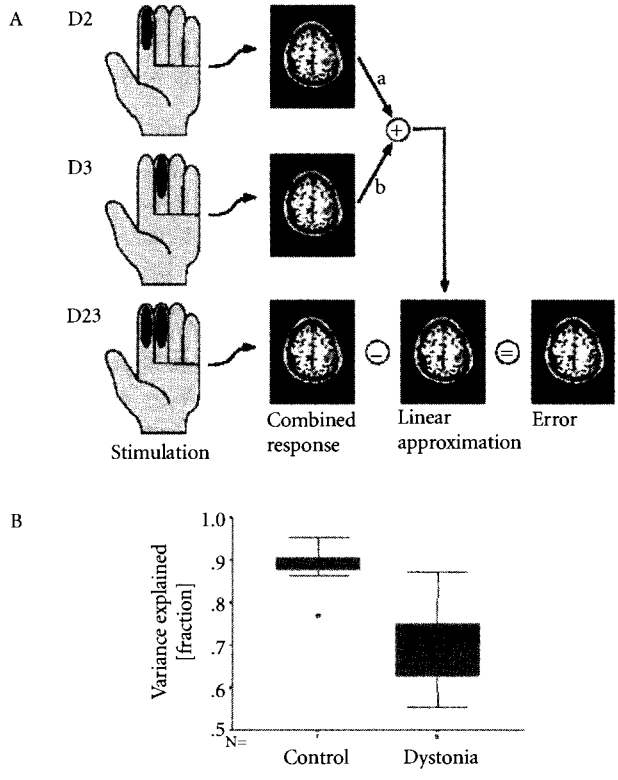



**Figure 26.5** (A) Schematic representation of the experimental design and representative fMRI data in one subject. Sensory stimulation was applied to the index (D2) or middle (D3) finger alone or to both fingers at the same time (D23) while the fMRI bold signal was recorded. A linear approximation of the fMRI signal associated with combined stimulation of D2 and D3 was calculated and substracted from the measured signal, giving rise to the 'error' of 'variance'. (B) Variance results in graphic form for all control subjects and patients with dystonia. Modified from Refs 46 and 47. (See Plate 25 in colour section.)

inputs could potentially lead to poorly differentiated control of motor representations and be the mechanisms underlying the risk of faulty motor control in some instrumentalists.

Sanger *et al.*[46,47] investigated whether evidence from fMRI supports abnormal receptive fields in primary somatosensory cortex that span more than a single finger surface. We developed a new technique for investigating overlap of sensory cortical receptive fields (Figure 26.5) and hypothesized that the combined metabolic demands of two spatially separated populations of cells, when activated simultaneously, would be approximated by the sum of the metabolic demands of each population individually. Therefore, if the populations of cells activated by the second and third digits in normal subjects are distinct, we expect that the task-related component of the blood oxygenation level–dependent (BOLD) signal obtained by simultaneous activation of the two digits will be a linear combination of the task-related BOLD signals when each finger is activated individually. Conversely, if

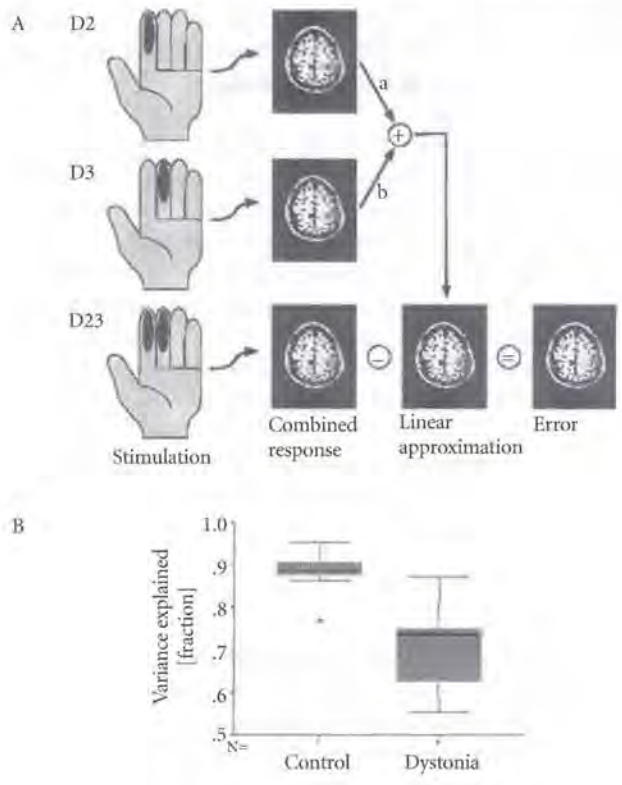



**Figure 26.5** (A) Schematic representation of the experimental design and representative fMRI data in one subject. Sensory stimulation was applied to the index (D2) or middle (D3) finger alone or to both fingers at the same time (D23) while the fMRI bold signal was recorded. A linear approximation of the fMRI signal associated with combined stimulation of D2 and D3 was calculated and substracted from the measured signal, giving rise to the 'error' of 'variance'. (B) Variance results in graphic form for all control subjects and patients with dystonia. Modified from Refs 46 and 47. (See Plate 25 in colour section.)

inputs could potentially lead to poorly differentiated control of motor representations and be the mechanisms underlying the risk of faulty motor control in some instrumentalists.

Sanger *et al.*[46,47] investigated whether evidence from fMRI supports abnormal receptive fields in primary somatosensory cortex that span more than a single finger surface. We developed a new technique for investigating overlap of sensory cortical receptive fields (Figure 26.5) and hypothesized that the combined metabolic demands of two spatially separated populations of cells, when activated simultaneously, would be approximated by the sum of the metabolic demands of each population individually. Therefore, if the populations of cells activated by the second and third digits in normal subjects are distinct, we expect that the task-related component of the blood oxygenation level–dependent (BOLD) signal obtained by simultaneous activation of the two digits will be a linear combination of the task-related BOLD signals when each finger is activated individually. Conversely, if

there is overlap of the receptive fields so that the same sensory populations respond to two different fingers, we do not necessarily expect linearity, because the metabolic demands of a single population stimulated by two different fingers may be determined by a complex function of the stimulus strength.

BOLD contrast was measured with fMRI during tactile stimulation of the index finger, the middle finger, or both simultaneously in five patients with focal dystonia and seven control subjects. In the control subjects, a linear combination of activation patterns for individual finger stimulation predicts an average of 88 per cent of the variance in the pattern of activity for combined stimulation. In patients with dystonia, the linear combination predicted only 70 per cent of the combined stimulation pattern ($p = 0.008$; Figure 26.6). Therefore, our results suggest that in patients with dystonia, the same region of sensory cortex may respond to tactile stimuli on more than one finger. Disorganization of sensory representations appears to be part of the pathophysiology of focal dystonia and may contribute to motor abnormalities. Hence, emphasis needs to be placed in the sensory as well as on the motor aspects of skill acquisition and practice in musicians. Perhaps all instrumentalists should learn Braille to enhance the functional segregation of individual fingers and minimize the risk of dystonia.

## Conclusion

Tools are now available to study the neurophysiological correlates of skill learning in humans. The motor cortex plays an important role in motor skill learning, but so does the sensory cortex. The sensorimotor cortex changes as a consequence of skill acquisition. These plastic changes, which probably include functional and structural components, place the subjects at great advantage for skillful task performance, but harbour the risk of the development of motor control disorders. The results of our studies may be useful in understanding not only the physiology of skill acquisitions, but also the pathophysiology of movement disorders in skilled performers. Further work along these lines may lead to helpful insight into the appropriate teaching/learning technique for fine motor skills.

## Acknowledgements

Part of the work summarized in this chapter was done in collaboration with Mark Hallett, Nguyet Dang, Leonardo Cohen, José María Tormos, Josep Valls-Solé, Terrance Sanger, Gottfried Schlaug, Daniel Tarsy, Jesus Pujol, and Jaume Roset-Llobet. This work was supported in part by the General Clinical Research Center at Beth Israel Deaconess Medical Center (National Center for Research Resources MO1 RR01032) and grants from the National Institute of Mental Health (MH57980 and MH60734) and the National Eye Institute (EY12091).

## References

1. **Critchley, M.** and **R. A. E. Henson** (1977) *Music and the Brain: Studies in the Neurology of Music.* Southampton: The Camelot Press.

2. Holding, D. H. E. (1989) *Human Skills*, 2nd ed. Chichester, New York: John Wiley & Sons.
3. Wilson, F. R. (1989) Acquisition and loss of skilled movement in musicians. *Sem. Neurol.* 9, 146–51.
4. Fuster, J. M. (1995) *Memory in the Cerebral Cortex. An Empirical Approach to Neural Networks in the Human and Nonhuman Primate.* Cambridge, MA: The MIT Press.
5. Pascual-Leone, A., D. Nguyet, L. G. Cohen, *et al.* (1995) Modulation of muscle responses evoked by transcranial magnetic stimulation during the acquisition of new fine motor skills. *J. Neurophysiol.* 74, 1037–45.
6. Pascual-Leone, A., F. Tarazona, and M. D. Catala (1999) Applications of transcranial magnetic stimulation in studies on motor learning. *Electroencephalogr. Clin. Neurophysiol. Suppl.* 51, 157–61.
7. Jacobs, K. M. and J. P. Donoghue (1991) Reshaping the cortical motor map by unmasking latent intracortical connections. *Science* 251, 944–7.
8. Sanes, J. N., J. Wang, and J. P. Donoghue (1992) Immediate and delayed changes of rat motor cortical output representation with new forelimb configurations. *Cerebr. Cortex* 2, 141–52.
9. Brasil-Neto, J. P., A. Pascual-Leone, J. Valls-Solé, *et al.* (1992) Focal transcranial magnetic stimulation and response bias in a forced-choice task. *J. Neurol. Neurosurg. Psychiatry* 55, 964–6.
10. Brasil-Neto, J. P., J. Valls-Solé, A. Pascual-Leone, *et al.* (1993) Rapid modulation of human cortical motor outputs following ischaemic nerve block. *Brain* 116, 511–25.
11. Iriki, A., C. Pavlides A. Keller, and H. Asanuma (1989) Long-term potentiation in the motor cortex. *Science* 245, 1385–7.
12. Iriki, A., C. Pavlides, A. Keller, and H. Asanuma (1991) Long-term potentiation of thalamic input to the motor cortex induced by coactivation of thalamocortical and corticocortical afferents. *J. Neurophysiol.* 65, 1435–1.
13. Greenough, W. T. (1984) Structural correlates of information storage in the mammalian brain: a review and hypothesis. *Trends Neurosci.* 7, 229–33.
14. Pons, T. P., P. E. Garraghty, A. K. Ommaya, and J. H. Kaas (1991) Massive cortical reorganization after sensory deafferentation in adult macaques. *Science* 252, 1857–60.
15. Seitz, R. J., E. Roland, C. Bohm, *et al.* (1990) Motor learning in man: a positron emission tomographic study. *Neuroreport* 1, 57–60.
16. Grafton, S. T., J. C. Mazziotta, S. Presty, *et al.* (1992) Functional anatomy of human procedural learning determined with regional cerebral blood flow and PET. *J. Neurosci.* 12, 2542–8.
17. Grafton, S. T., R. P. Woods, and J. M. Tyszka (1994) Functional imaging of procedural motor learning: relating cerebral blood flow with individual subject performance. *Human Brain Mapping* 1, 221–34.
18. Grafton, S. T., E. Hazeltine, and R. Ivry (1995) Functional mapping of sequence learning in normal humans. *J. Cognit. Neurosci.* 7, 497–510.
19. Jenkins, I. H., D. J. Brooks, P. D. Nixon, *et al.* (1994) Motor sequence learning: a study with positron emission tomography. *J. Neurosci.* 14, 3775–90.
20. Karni, A., G. Meyer, P. Jezzard, *et al.* (1994) Where practice makes perfect: an fMRI study of long-term motor cortex plasticity associated with the acquisition of a motor skill (abstr). *Soc. Neurosci. Abstr.* 20, 1291.
21. Karni, A., G. Meyer, P. Jezzard, *et al.* (1995) Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature* 377, 155–8.

22. Karni, A., G. Meyer, C. Rey-Hipolito, *et al.* (1998) The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. *Proc. Natl. Acad. Sci. USA* 95, 861–8.
23. Pascual-Leone, A., J. Grafman, and M. Hallett (1994) Explicit and implicit learning and maps of cortical motor output. *Science* 265, 1600–1.
24. Schieber, M. H. (1990) How might the motor cortex individuate movements. *Trends Neurosci.* 13, 440–5.
25. Schieber, M. H. and L. S. Hibbard (1993) How somatotopic is the motor cortex hand area? *Science* 261, 489–92.
26. AOu, S. A., C. D. Woody, and D. Birt (1992) Increases in excitability of neurons of the motor cortex of cats after rapid acquisition of eye blink conditioning. *J. Neurophysiol.* 12, 560–9.
27. Woody, C. D. (1986) Understanding the cellular basis for learning and memory. *Annu. Rev. Psychol.* 37, 433–93.
28. Pascual-Leone, A., J. M. Tormos, J. Keenan, *et al.* (1998) Study and modulation of human cortical excitability with transcranial magnetic stimulation. *J. Clin. Neurophysiol.* 15, 333–43.
29. Pascual-Leone, A., D. Bartres-Faz, and J. P. Keenan (1999) Transcranial magnetic stimulation: studying the brain-behaviour relationship by induction of 'virtual lesions'. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 354, 1229–38.
30. Mendoza, D. W. and H. Wichman (1978) 'Inner' darts: effects of mental practice on performance of dart throwing. *Percept. Motor Skills* 47, 1195–9.
31. McBride, E. R. and A. L. Rothstein (1979) Mental and physical practice and the learning and retention of open and closed skills. *Percept. Motor Skills* 49, 359–65.
32. Dennis, M. Visual imagery and the use of mental practice in the development of motor skills. *Can. J. Appl. Sport Sci.* 10, 4S–16S.
33. Schonberg, H. (1987) *Great Pianists.* St. Louis, MO: Fireside Books.
34. Schonberg, H. (1988) *The Virtuosi.* New York: Vintage-Random House.
35. Ingvar, D. H. and L. Philipson (1977) Distribution of the cerebral blood flow in the dominant hemisphere during motor ideation and motor performance. *Ann. Neurol.* 2, 230–7.
36. Roland, P. E., L. Ericksson, S. Stone-Elander, and L. Widen (1987) Does mental activity change the oxidative metabolism of the brain? *J. Neurosci.* 7, 2373–89.
37. Decety, J. and D. H. Ingvar (1990) Brain structures participating in mental simulation of motor behavior: a neuropsychological interpretation. *Acta Psychol. (Amst.)* 73, 13–34.
38. Decety, J., D. Perani, M. Jeannerod, *et al.* (1994) Mapping motor representations with positron emission tomography. *Nature* 371, 600–2.
39. Hallett, M. (1998) The neurophysiology of dystonia. *Arch. Neurol.* 55, 601–603.
40. Pujol, J., J. Roset-Llobet, D. Rosines-Cubells, *et al.* (2000) Brain cortical activation during guitar-induced hand dystonia studied by functional MRI. *Neuroimage* 12, 257–67.
41. Hallett, M. (1995) Is dystonia a sensory disorder? *Ann. Neurol.* 38, 139–40.
42. Bara-Jimenez, W., M. J. Catalan, M. Hallett, and C. Gerloff (1998) Abnormal somatosensory homunculus in dystonia of the hand. *Ann. Neurol.* 44, 828–31.
43. Elbert, T., V. Candia, E. Altenmuller, *et al.* (1998) Alteration of digital representations in somatosensory cortex in focal hand dystonia. *Neuroreport* 9, 3571–75.
44. Byl, N. N. and M. Merzenich (1997) The neural consequences of repetition: clinical implications of a learning hypothesis. *J. Hand Ther.* 10, 160–74.

45. **Byl, N. N., M. M. Merzenich,** and **W. M. Jenkins** (1996) A primate genesis model of focal dystonia and repetitive strain injury. I. Learning-induced dedifferentiation of the representation of the hand in the primary somatosensory cortex in adult monkeys. *Neurology* 47, 508–20.

46. **Sanger, T., A. Pascual-Leone, D. Tarsy,** and **G. Schlaug** (2001) Non-linear sensory cortex response to simultaneous tactile stimuli in writer's cramp. *Movement Dis.* In Press

47. **Sanger, T., D. Tarsy,** and **A. Pascual-Leone** (2001) Abnormalties of spatial and temporal sensory discrimination in writer's cramp. *Movement Dis.* In press.

PART VI

# RELATION OF MUSIC TO OTHER COGNITIVE DOMAINS

CHAPTER 27

# THE SOUNDS OF POETRY VIEWED AS MUSIC

FRED LERDAHL

## Abstract

An extended parallel is developed between musical and prosodic structures, using the author's cognitively oriented music theory and recent work in generative phonology. As illustration, the sounds of a short poem by Robert Frost are treated entirely in musical terms. The poem is assigned a phonological stress grid and then musical grouping and meter. These structures enable a durational realization. Phonological stress also helps assign the poem's normative melodic contour. Finally, the similarities and differences in sound repetition are given hierarchical structure by means of musical prolongational theory. These formal parallels suggest a corresponding realization in brain localization and function. Evidence from the neuropsychological literature is cited in support of this view. The picture emerges that grouping, metre, duration, contour, and timbral similarity are mind/brain systems shared by music and language, whereas linguistic syntax and semantics and musical pitch relations are systems not shared by the two domains.

**Keywords:** Poetry and music; Music and poetry; Phonology, generative; Frost, Robert

## Introduction

Comparisons between music and language have traditionally been couched in terms of syntax or rhetoric.[1] The more substantive parallels, however, are between musical structures on the one hand and phonological and prosodic structures on the other, and they derive from the fact that music and language both consist of sounds organized in time. Here I explore these parallels by developing a representational and partly computational account of the sounds of a poem treated entirely as music. This approach builds on work from previous collaborations and consultations.[2–5] As I hope to show, the analysis is suggestive for the neurobiology of music.

## A musical-poetic analysis

Consider in some detail the first two lines Robert Frost's short lyric 'Nothing Gold Can Stay', given in its entirety in Figure 27.1. Grouping structure, a fundamental component in music theory, segments a musical surface hierarchically into motives, phrases, and sections. Phonologists have developed a comparable concept, the prosodic hierarchy.[6] Hence linguistics now offers a second kind of tree structure, a prosodic as well as a syntactic one.

Nature's first green is gold,
Her hardest hue to hold.
Her early leaf's a flower;
But only so an hour.
Then leaf subsides to leaf.
So Eden sank to grief.
So dawn goes down to day.
Nothing gold can stay.

**Figure 27.1** 'Nothing Gold Can Stay', by Robert Frost.

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [C | Nature's] | [C | first] | [C | green] | [C | is gold] | [C | Her hardest] | [C | hue] | [C | to hold] |
| [P | Nature's] | [P | first | | green] | [P | is gold] | [P | Her hardest | | hue] | [P | to hold] |
| [I | Nature's | | first | | green | | is gold] | [I | Her hardest | | hue | | to hold] |
| [U | Nature's | | first | | green | | is gold | | Her hardest | | hue | | to hold] |

**Figure 27.2** Bracketed, stratified representation of the prosodic analysis of the first couplet. C = clitic group, P = phonological phrase, I = intonational phrase, and U = utterance.

Even where the two coincide in their segmentations, their meanings differ. Syntactic trees represent parts of speech and syntactic phrases; phonological trees represent groupings of spoken sound.

At the lowest levels of the prosodic hierarchy are the syllable and the word. Words categorize as either content words, which carry major semantic content, or function words, which mainly fill a syntactic role. For example, in the first line of the poem, 'Nature's first green is gold', all the words are content words except 'is'. At the level above the word is the clitic group, which coincides either with a content word or combines content and function words that, together, sound as if they were one word. In the latter case, the stressed syllable of the content word is the clitic host, the sound to which the others are drawn. In 'Nature's first green is gold', 'is gold' sounds as one word, with 'gold' as clitic host. Likewise in the second line, 'Her hardest hue to hold', the content words are 'hardest', 'hue', and 'hold'; 'Her hardest' and 'to hold' are clitic groups, with 'hard-' and 'hold' as clitic hosts. Above the clitic group is a grouping of words called the phonological phrase, for example, 'first green'. Above that is the intonational phrase, which conveys the melody of speech. The entire line, 'Nature's first green is gold', is an intonational phrase. Finally, there is the utterance, which usually corresponds to a sentence. 'Nature's first green is gold,/Her hardest hue to hold' is an utterance.

Figure 27.2 displays the grouping analysis of the first two lines of the poem. The syllable and word levels are omitted for convenience. The bracketing categories are indicated in boldface. A category can repeat from one level to the next. This representation abstractly resembles a parsing of a musical phrase into subgroups and motives. As in music, pauses are more likely to occur between, rather than within, these segments, with greater pauses between superordinate boundaries.

Now consider linguistic stress, whose musical equivalent is phenomenal accent.[2] In both, the perception, whether of a syllable or of a pitch event, is one of relative sonic prominence

within its immediate context. Linguists have proposed a so-called metrical grid to represent linguistic stress in words and phrases.[7] From a musical perspective, this designation is misleading; it acts as a stress grid, for its purpose is to represent not hierarchical periodicities but hierarchical patterns of stress. Within a polysyllabic word, such a pattern is given lexically, as in the first word of the poem, 'Náture's' (in English, one would not say 'Natúre's'). Within a phrase, the stress pattern obeys the nuclear stress rule, which, after the categorization of words into content and function words, assigns global stress to the strongest syllable of the last content word of a prosodic unit.[8] Thus, we say 'first gréen', rather than 'fírst green', because 'green' is the last stressed syllable of its phonological phrase. This principle can be overridden by the rhythm rule, which ameliorates clashes in stress between adjacent or nearby syllables. According to nuclear stress, one says 'hardest húe', but we might prefer to adjust it to 'hárdest hue', so that the nearby words 'hue' and 'hold' are not comparably stressed. The rhythm rule implicitly invokes metrical periodicity: strong beats should be more or less evenly spaced. At this level, then, the choice is between 'Her hardest húe to hóld', which obeys nuclear stress but challenges periodicity, and 'Her hárdest hue to hold', which violates nuclear stress in order to distribute the stresses. Finally, nuclear stress can be overridden by nonnormative focus on a particular word, to bring out a semantic nuance. Hence if we say 'Nature's fírst green is gold', the implication is that nature's second green is not gold. Generally, however, nuclear stress holds.

Setting the rhythm rule and focus to one side, a stress grid is constructed by cyclic generation of stresses from level to level in the prosodic hierarchy, observing nuclear stress starting at the phonological grouping level. As illustrated in Figure 27.3a, the procedure is first to assign an *x* to every syllable, second to assign an *x* in the next row to a lexically relatively stressed syllable in a polysyllabic word or to the host of a clitic group. For the third row, shown in Figure 27.3b, an *x* for the nuclear stress is added at the level of the phonological phrase. Finally, in Figure 27.3c, an *x* for the nuclear stress is assigned at the level of

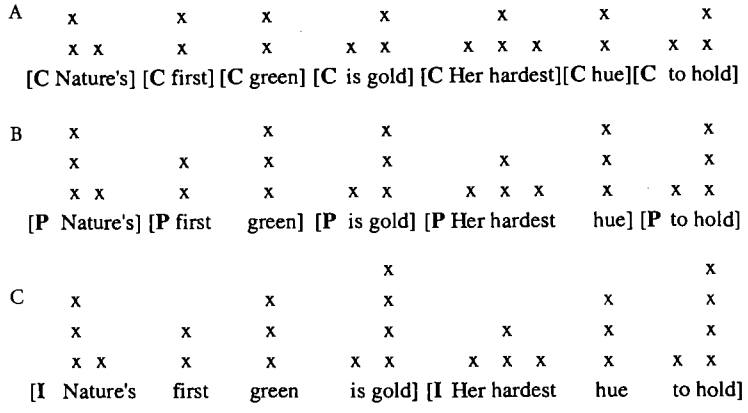



Figure 27.3 Assignment of the stress grid in conjunction with the prosodic hierarchy.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *level d:* | [Na] | [ture's] | [first] | [green] | [is] | [gold] | [Her] | [hard] | [est] | [hue] | [to] | [hold] |
| *level c:* | [Na- | ] | [first] | [green] | [ | gold] | [ | hard- | | [hue] | [ | hold] |
| *level b:* | [Na- | ] | [ | green] | [ | gold] | [ | | | hue] | [ | hold] |
| *level a:* | [ | | | | | gold] | [ | | | | | hold] |

**Figure 27.4** The stress pattern in Figure 27.3 put in reductional format.

the intonational phrase. There is little point in continuing to yet another level, for, as in music, distinctions in linguistic stress attenuate over long spans.

Figure 27.4 recasts the stress pattern in Figure 27.3c into an equivalent music-reductional format. Level *d* brackets the syllables and corresponds in content to the bottom row of the stress grid in Figure 27.3c. Level *c* reduces out the syllables and monosyllabic words that in Figure 27.3c are assigned only one *x*. Similarly, levels *b* and *a* keep only those units that carry three and four *x*s, respectively. This is the same layout as in musical time-span reduction, which selects relatively important pitch events at successively larger levels of rhythmic segmentation.[2] These relationships could alternatively be displayed by a tree representation with right and left branching, again as in the corresponding music theory.

Turn now to metre. In both music and poetry, metrical structure consists of hierarchically related periodicities inferred from the signal. It might be objected that periodicities in language, unlike those in music, do not really exist because syllabic and phrasal durations are so variable. Yet it would be misleading to say that musical durations are invariable. Expressive musical performance depends on deviations from isochrony. Like metre itself, temporal precision is a mental construct. While it is true that durations in verse are usually more variable than in music, many poetic idioms, from nursery rhymes to sophisticated traditions, demonstrate considerable regularity.[9] As in music, these verbal idioms approach periodicity as a framework against which expressive deviations take place.

Once periodicity is understood as a relative matter, poetic and musical meter may be regarded as formally and cognitively equivalent. This has, in essence, long been recognized by music theorists in their occasional borrowing of prosodic foot notation for the representation of musical metre. Now, however, it is more usual in music theory to employ a grid notation, which represents the metrical hierarchy directly and does not confuse features of metre with those of grouping.[2] A similar consensus has emerged in phonological theory.

A poetic or musical metre exists when the perceiver infers conceptually regular levels of beats from the signal, such that a beat that is strong at one level is also a beat at the next larger level. The perceptually most prominent metrical level is called the tactus; it occurs at a moderate tempo, is usually at an intermediate level within a grid, and is often more restricted in regularity than are smaller or larger levels. Beats at a given level are two or three beats apart at the next smaller level. (Even pentametre lines have intermediate strong beats, resulting in patterns of two and three.) Some poetic or musical idioms, while allowing a mixture of two or three beats apart across levels, discourage the mixture of two or three strong beats apart within a level. Thus, European classical tonal music uses compounds of two and three across levels as in Figure 27.5a and b, but it regards as atypical the grid in Figure 27.5c, which alternates two and three beats within a single level. Other musical cultures, as in the Balkan countries, would take Figure 27.5c as normative. The important point

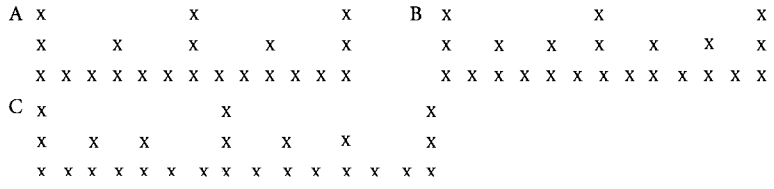



**Figure 27.5** Three metrical grids: (A) 6/8 metre; (B) 3/4 metre; (C) 7/8 metre (2 + 2 + 3).

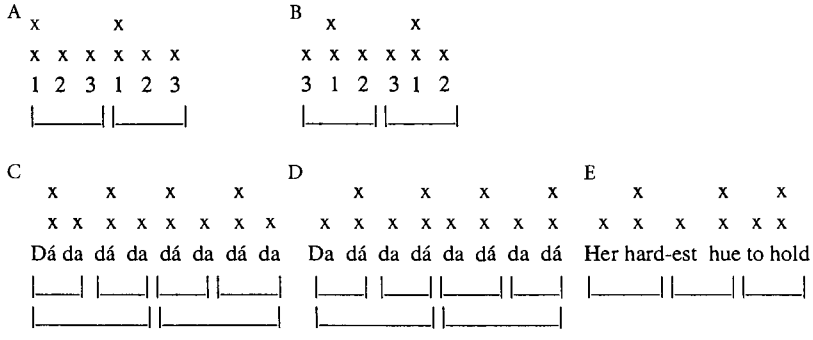



**Figure 27.6** Interaction of grouping and metre: (A) 3/4 metre with afterbeats; (B) 3/4 metre with an upbeat and an afterbeat; (C) trochaic tetrametre; (D) iambic tetrametre; (E) conflict between word and foot boundaries at 'hardest'.

is not that different cultures prefer different metrical grids, but that the possible combinations of two and three, across or within levels, are very small. Cultural variation on poetic and musical metres is intrinsically limited.

In constructing a metrical grid, the perceiver seeks an optimal fit between the stresses provided by the acoustic signal and the culturally available repertory of metrical grids. To hear a musical phrase in 3/4 time or a poetic line in pentametre is to select that particular grid because, within the repertory, that grid gives the minimal number of conflicts between stress and beat. Where the fit is imperfect, the effect is one of syncopation or, to use terms from traditional scansion, of inversion, substitution, or truncation. If the fit is seriously imperfect, the perceiver switches to another available grid or perhaps abandons metrical understanding altogether.

The interaction of grouping and metre causes intuitions of upbeat and afterbeat in relation to a downbeat. The metre is 3/4 in Figure 27.6a and b, but the two differ in that Figure 27.6a has afterbeats while Figure 27.6b has an upbeat and an afterbeat. The terminology of traditional scansion recognizes this upbeat-afterbeat distinction. In current notation, Figure 27.6c and d represents the contrasting groupings of trochaic and iambic tetrametre, respectively. Traditional foot scansion becomes troublesome, however, in cases where the presumed foot contradicts the prosodic hierarchy. Supposing our lyric to be in iambic trimetre, in the line, 'Her hardest hue to hold', the foot-grouping of '-est' is not with its own word stem, 'hard-', but counterintuitively with 'hue', as shown in Figure 27.6e. It is unclear

in such cases what function the poetic foot serves. In the present view, iambs, trochees, dactyls, and so on, arise not from a predetermined schema but directly from the prosodic groupings of beats. Moreover, the concept of the foot lacks a musical counterpart. While recognizing its historical practice, let us ignore the foot as an independent unit of rhythmic analysis.

Having set to one side the iambic dimension of our poem, I would now suggest that it is not in trimetre either. The punctuated pauses at the end of each line and, more elusively, the historical style that it evokes, suggest that it is in truncated tetrametre. Many short lyrics, such as those of Emily Dickinson, employ so-called common metre, comprising couplets in a four-plus-three beat pattern with an implied silent final beat: da-dá, da-dá, da-dá, da-dá; da-dá, da-dá, da-dá, —. Common metre is a variant of standard tetrameter, the final silent beat reinforcing the grouping structure. Our poem is in turn a variant of common metre. Frost uses it here because of the ineluctable decline that is the poem's theme. Every line falls into silence.

Given this paradigmatic metre, the first two lines of the poem receive the structural description in Figure 27.7. Above the lines appears the stress grid, taken from Figure 27.3c. Beneath, in boldface and inverted for visual convenience, is the metrical grid for the tactus and larger levels, showing four basic beats per line and their relative metrical strengths. The tactus is represented by upper-case Xs. To clarify the grid notation, the beats are numbered as two 4/4 measures, grouped into two intonational phrases. Only stresses of three or four *xs* count for the establishment of the tactus. Since the greatest stresses occur on 'gold' and 'hold', these syllables are assigned the strongest beats. The periodicity of the metrical grid requires that 'Na-' and 'hard-' receive the next strongest metrical position, with 'green', 'hue', and the silent beats getting the weaker beats. This treatment creates a slight syncopation at 'hue'. Observe that the strongest metrical beats take place at the end of each line, creating an out-of-phase relationship between the grouping boundaries and the spans between the strongest beats. This is a consequence of nuclear stress and is characteristic of poetry in English. In classical tonal music, by contrast, the strongest beat usually comes at or near the beginning of a phrase. For its ending rhythmic articulation, tonal music relies instead on the cadence, for which there is no linguistic counterpart. (In Balinese gamelan the strongest beat usually arrives at the end of a phrase.[10])

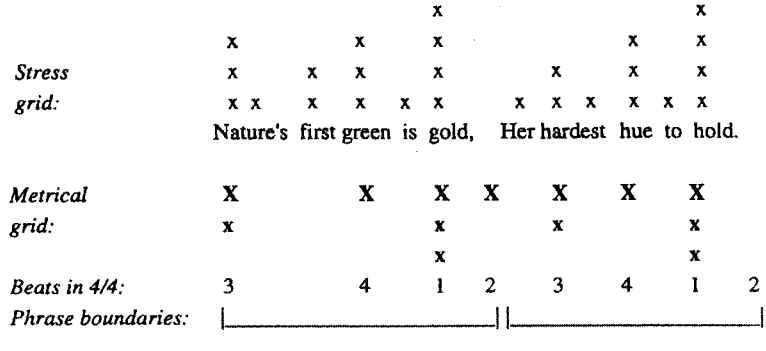



Figure 27.7 Stress and metrical analysis for the first two lines of the Frost poem.

This analysis indicates conceptually equal durations between tactus beats. To derive subtactus levels, the two requisite principles are to seek the best possible fit (1) between stress and beat, and (2) between duration and the prosodic hierarchy. Condition 1 has already been discussed; it is a matter of matching the metrical with the stress grid, starting with the tactus level. Condition 2 invokes the Gestalt principle that proximate objects tend to group together. It follows that grouping boundaries tend to exist between objects in a string that are relatively nonproximate. The prosodic hierarchy is meant to reflect perceptual groupings of sound. Therefore, contextually longer durations—or, more precisely, distances between attack points—should occur between, rather than within, the constituents of the prosodic hierarchy. Once these subtactus levels are established, the metrical positions of all the syllables of the couplet can be translated into standard musical notation.

Let us apply these two conditions to the first three words of the poem, 'Nature's first 'green', and derive their durations. Figure 27.8a fully satisfies condition 1: 'Na-' and 'green' provide the tactus frame, and in both grids '-ture's' receives a single *x* and 'First' receives two *x*'s. Translated into metrical durations, the result is a duple division with two sixteenth notes plus an eighth note. But this solution does not satisfy condition 2: the greater distance is between 'first' and 'green', weakening the perception of the phonological group 'first green'. The alternative in Figure 27.8b, an eighth plus two sixteenths, is more problematic still, for it both violates condition 1 and breaks up the word 'Nature's' with the longer duration. Figure 27.8c effects the preferable solution by smoothing out the durations in a triple-beat realization. Condition 1 is almost met, and there is no implication of an unwanted grouping boundary.

On the assumption that triple metre continues at the eighth-note level, Figure 27.9 completes the subtactus level for Figure 27.7 and realizes the rhythm for the two lines. The match between stress and beat is complete, and the durations reflect the constituents of the prosodic hierarchy: 'is gold', 'Her hardest', and 'to hold' are all clitic groups, grouped musically by temporal proximity and surrounded on either side by longer distances between attack points.

If this derivational procedure is carried out for the entire poem, the rhythmic derivation is as in Figure 27.10, cast in 12/8 metre with two-beat anacruses for each line. Minor adjustments with duple divisions appear in the second, third, and fourth lines. The first three of these have to do with syllable length. For example, in Figure 27.9 'hard-' receives an eighth note and '-est' a quarter note, even though 'hard-' is a longer syllable than '-est'. To reverse the values, however, so that 'hard-' is assigned a quarter note and '-est' an eighth note,

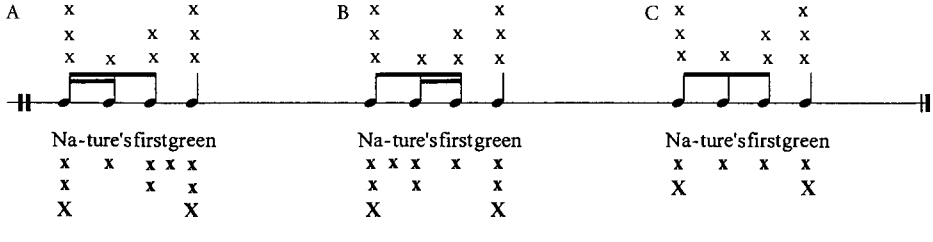



Figure 27.8 Duple and triple subtactus metrical realizations for the beginning of the first line.

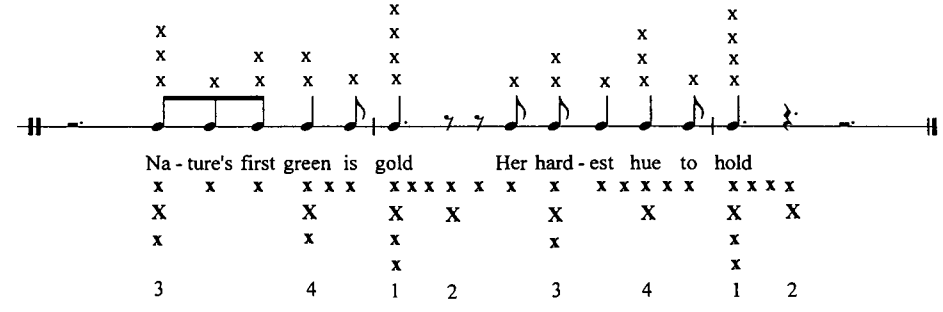



Figure 27.9 Addition of the subtactus level to Figure 27.7, with translation into musical durations.

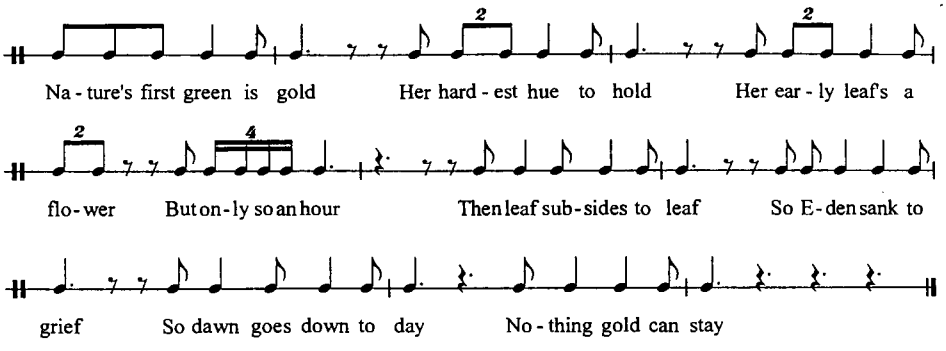



Figure 27.10 Durational reading of the poem.

would violate condition 2, weakening the word boundary. A compromise is effected in Figure 27.10 by a duple realization—two equal eighth notes—for 'hard-est'. The rhythm of 'But only so an hour' is caused by a different factor: the only content words are 'only' and 'hour', so that the monosyllabic words 'so' and 'an' are cliticized. The absence of any stress between 'on-' and 'hour' permits a quick setting. 'Hour' arrives a beat earlier than it otherwise would, and consequently there is a longer pause between the two quatrains than between the lines internal to each quatrain. These changes in lines two, three, and four induce considerable rhythmic variety in the first quatrain. This variety in turn sets up the second quatrain's rhythmic monotony, a reflection of its sense of futility.

We turn now by way of stress to contour. Three acoustic variables contribute to stress: intensity (or loudness), pitch height, and duration.[11] While duration is an independent variable, intensity and pitch height covary, for the muscles that enable greater intensity also tend to raise pitch. Even when pitch height is artifically isolated, relative height induces the perception of relative stress. Thus, pitch height is a strong projector of stress. Since pitch height is a matter of contour, stress patterns can provide a basis for the derivation of speech melody. There are many acceptable contours, however, for an intonational phrase or utterance, each having its own nuance. If our concern were grammaticality, in keeping with

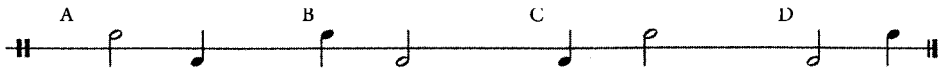



Figure 27.11 Four normative contour frameworks.

much generative linguistics, well-formedness constraints would underdetermine the options for contour realizations. A different goal, and one congenial with a musical perspective, is to ask whether a grammatical contour is normative or not. One need then derive only normative contours from stress grids, letting the other grammatical contours be seen as nuanced variants.

To pursue this question, let us initially restrict contour features to a binary classification of high and low, and apply it to the two most stressed syllables in an intonational phrase. Contour theory in linguistics has represented this classification by 'H' and 'L', with accompanying diacritics to indicate hierarchical function.[12,13] However, it is in the spirit of the approach taken here to employ musical notation, which after all is designed to show pitch height. As is common in musical reduction theory, hierarchy can be represented by note values that stand not for durations but for relative structural importance. In Figure 27.11a the first syllable is high and the second low, with the half-note first syllable acting as superordinate to the quarter-note second syllable. The high–low pattern continues in Figure 27.11b but with a superordinate low syllable. Figure 27.11c and d reverses the high–low pattern. These four cases exhaust the possibilities. The prototypical declarative contour is that in Figure 27.11a and b, a rise and fall whose musical analogue is the tensing–relaxing pattern of a standard tonal phrase. The difference between the two is whether the high or the low syllable dominates. In Figure 27.11c and d, the high syllable is to the right of the low one; this is the typical shape of an interrogative.

The claim that a higher pitch causes stress would seem to disqualify Figure 27.11b and d as normative, since in these cases the low syllable dominates. As a consequence of nuclear stress, however, a superordinate low syllable can indeed function in the position of the last clitic host of an intonational phrase. In other words, this syllable's location at the close of a major phrase boundary compensates for its lower pitch. Hence Figure 27.11b is normative. By the same token, Figure 27.11d is not normative, because in this case the dominating low syllable is not in a location to benefit from major nuclear stress. As there are no interrogatives in our poem, we can also ignore Figure 27.11c. This leaves Figure 27.11a and b as the remaining options. This very restricted repertory will provide the framework for mapping stress importance onto contour peaks and valleys.

In this presentation we take for granted but do not notate that there is sliding from one pitch level to another (in musical terms, there is a glissando from one pitch to another). We also posit that a given syllable has a single pitch level, ignoring cases in which a syllable slides between two pitch levels, especially at a phrase boundary. For example, in the utterance 'He has gone!', 'gone' normally begins high because it is stressed but ends low in order to convey its declarative status. In 'He has gone?', however, 'gone' begins low but ends high in order to convey its interrogative status. Finally, we bypass the possibility that an extended sequence of highs and lows can gradually ramp up or (more commonly) down, so that the

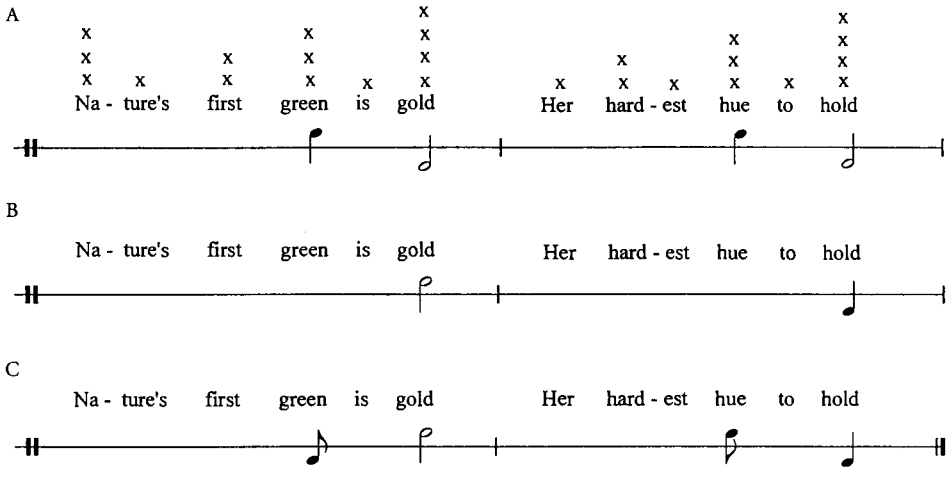



Figure 27.12 Global contour assignments for the first couplet.

pitch levels do not stay absolute but are relative. In short, our contour model is highly idealized. Even so, it generates plausible shapes derived from musical principles.

The stress analysis in Figure 27.3b and c assigns three *x*s for the most stressed syllable in each phonological phrase and four *x*s for the most stressed syllable in each intonational phrase, the latter occurring at the end of the line. Hence, as shown in Figure 27.12a, the intonational contours for the first two lines of the poem fall in the pattern of Figure 27.11b. The two lines together, however, form an utterance, leading to the assignment in Figure 27.12b of high and low to four *x*s at this level, at 'gold' and 'hold', respectively. The assignment for 'gold' is thereby reversed from low in Figure 27.12a, to high in Figure 27.12b. This step does not yet determine whether 'gold' or 'hold' dominates; for reasons beyond the present purview, 'gold' is chosen. The next step is to combine the shapes at the utterance and intonational levels. In Figure 27.12a, 'green' was realized as high, an assignment that is no longer preferred because it would conflict with the high 'gold' in Figure 27.12b. Therefore 'green' receives a low position within its intonational phrase, as shown in Figure 27.12c; the eighth-note value signifies its subordinate status. In the next line, 'hue' remains high, for it combines at the intonational level with the superordinate low 'hold', although it too receives a subordinate eighth-note value. This is the scaffolding within which more local contour shapes take place.

A two-place categorization of high and low does not provide enough distinctions as more syllables come into play. Four levels appear to be optimal: three do not afford sufficient variety, and more than four would be fussy in view of the limited ability to categorize degrees of stress. In the musical notation, there are now two lines with spaces above and below, so that the highest position is placed above the top line, the second highest on the top line, the second lowest between the lines, and the lowest on the bottom line. On the basis of the salience of registral extremes, the two most prominent syllables, again represented by half and quarter notes, are placed in the top and bottom positions.

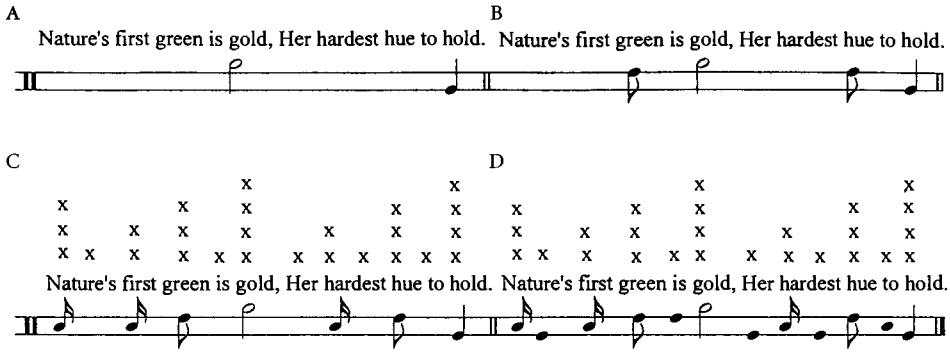



**Figure 27.13** Assignment of pitch levels to the first couplet: (A,B) from the analysis in Figure 27.12; (C) for unassigned syllables with three two *x*'s; (D) for syllables with one *x*.

Figure 27.13a performs this step for the first couplet. Figure 27.13b appends pitch heights for the syllables in the third hierarchical position, again represented by eighth notes, and places them one slot lower than their adjacent superordinate high syllable. These locations preserve the rising–falling shape of the contour schema in Figure 27.11a and b. Except where a superordinate low syllable occurs at the right edge of an intonational phrase, the contour reflects the shape of the stress grid, with its peaks and valleys of *x*s. Thus, in Figure 27.13c the syllables with three and two *x*s, notated by sixteenth notes, appear one slot below the positions of the syllables that were already given three *x*s in Figure 27.13b. Then in Figure 27.13d come the syllables with one *x*, represented by unstemmed noteheads. These syllables are cliticized and are typically spoken as schwas (the default neutral vowel for an unstressed syllable). They are placed one step below their clitic hosts, except when the latter are in the lowest slot, in which case they are located one step above the clitic host. Hence '-ure's' is one step below 'Na-' in 'Nature's', and 'is' one step below 'gold', but 'to' is one step above 'hold', since the latter is in the bottom slot.

This derivation of contour is musical not only in its attention to normative shapes but in its hierarchical treatment. Rather than generate contours from left to right, as if without projection or memory, global highs and lows are established according to a few paradigmatic shapes and the remaining syllables in the speech melody are filled in at successive levels of lesser stress. Figure 27.14 shows the result, combining the contour analysis of Figure 27.13 with the metrical/durational analysis of Figure 27.10, but now with these methods applied to the entire poem. Contained within this seemingly transparent musical notation are the structures of phonological stress, the prosodic hierarchy, the metrical grid, duration, and pitch height. The success of the derivation can be judged informally by its naturalness as a reading.

A usual prosodic analysis would have stopped before this point. From a musical perspective, there is yet another step to take, an analysis of the hierarchical patterns of recurrence of sounds. Traditional poetic analysis treats verbal recurrences such as rhyme, alliteration, and assonance as primitive sequential patterns: *aabb, abab*, and so forth. Music theory, in contrast, has a highly developed approach to recurrence in the form of prolongational structure.

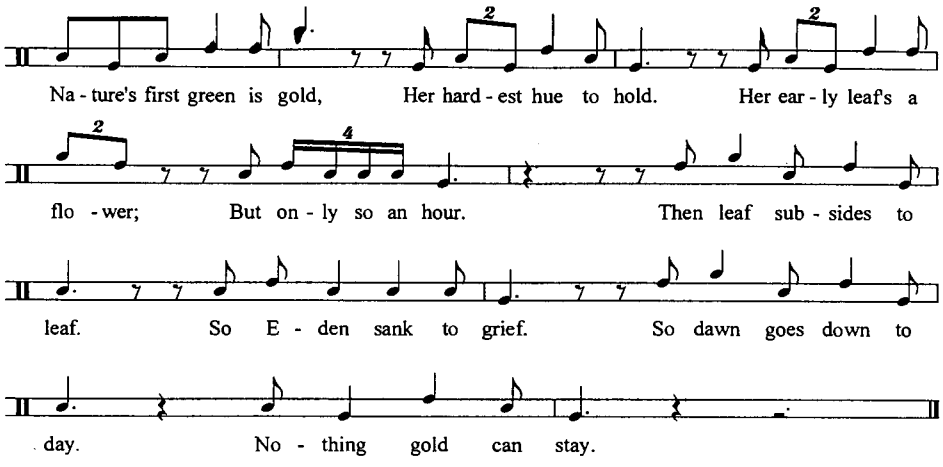



Figure 27.14 Metre and durations (from Figure 27.10) combined with a contour analysis of the entire poem.

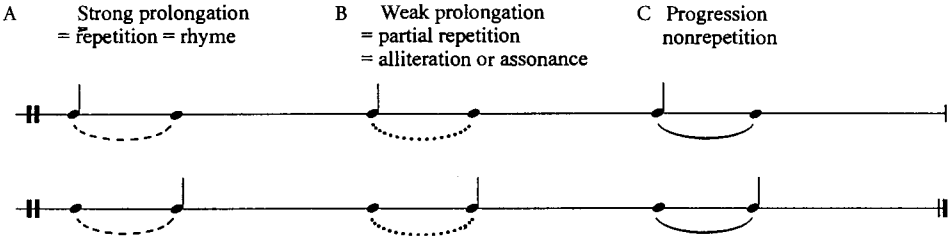



Figure 27.15 Repertory of prolongational connections.

Figure 27.15 gives three kinds of prolongational connection: (1) strong prolongation or repetition, represented by a dashed slur; (2) weak prolongation or partial repetition, represented by a dotted slur; and (3) progression or nonrepetition, represented by a solid slur. In each case the superordinate element can precede or succeed the elaborating element.

Musical prolongational connections are derived through a combination of time-span segmentation and tonal pitch stability.[2,14] Poetry incorporates the former through the prosodic hierarchy but lacks any counterpart to the tonal system. The better musical analogy, then, is to atonal music, for which event salience largely replaces event stability.[14] That is, when hearing atonal surfaces, and given the absence of acoustic or idiom-specific criteria to select stable from unstable events within hierarchical time spans, the listener tends to organize embellishing events in relation to events that have greater surface prominence within the framework of the time-span hierarchy. In phonological terms, relative surface prominence is equivalent to relative stress. Hence prosodic prolongation can be viewed as a kind of atonal prolongation, in which the elements being prolonged are not pitch structures but degrees of timbral similarity mediated by stress within the prosodic hierarchy. (This approach extends beyond my earlier treatment, in which synthesized vowels are

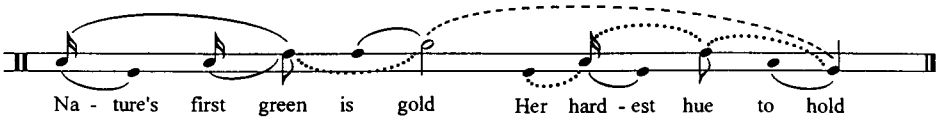



Figure 27.16 Prolongational structure for the first couplet. The contours are retained from Figure 27.14.

related prolongationally by similarity but without the incorporation of nested patterns of stress.[15]) The three prolongational categories continue intact: the repetition of rhyme equals strong prolongation; the weaker repetition of alliteration or assonance equals weak prolongation; and syllabic nonrepetition equals progression.

The syllables connected in the prolongational analysis of the first couplet follow the parsing of the prosodic hierarchy in Figure 27.2 and the hierarchical contour in Figure 27.13d. Each connection forms a progression, weak prolongation, or strong prolongation, according to the degree of repetition of sound. Figure 27.16 illustrates, retaining the structural notation of Figure 27.13d. In the first line, 'ture's' attaches as a progression from 'Na-', 'first' to 'green', and 'is' to 'gold'. At the next level, 'Na-' attaches to 'green' because both have three *x*s. Then 'green' attaches to 'gold', as a weak prolongation because of the alliteration between 'green' and 'gold'. The second line creates a similar pattern, but with more internal alliteration. Finally, at the level of the couplet, the two syllables with four *x*s, 'gold' and 'hold', form a rhyming strong prolongation. Note that it is timbral similarity rather than pitch that is connected prolongationally, for 'gold' is in the highest pitch category while 'hold' is in the lowest. The resulting graph shows not isolated instances of alliteration and rhyme, as in traditional prosodic analysis, but the richer relationship of partial repetitions nested within superordinate rhymes—'green' is to 'gold' as 'hue' is to 'hold'. Progressions between dissimilar syllables occur as relatively local connections. This approach can be extended to the poem as a whole.

This account of prosodic prolongational structure has skipped some derivational stages involving technicalities of the theory of musical reduction. The main point is that we can indeed transplant most of the mechanisms of the musical theory to the poetic realm.

## Connections

Let us step back from the derivation of the Frost analysis and assume that the music-analytic methods outlined here have validity for modelling the sonic organization of poetry in general. To what extent do these methods extend to the analysis of ordinary speech? Briefly, ordinary speech possesses phonological stress, the prosodic hierarchy, and contour, so the application of these components should continue intact. The phonological stresses of ordinary speech, however, are irregular enough to inhibit the inference of a metrical grid. Similarly, ordinary speech lacks significant patterns of syllabic repetition, so that the sense of a timbral prolongational structure is weakened. These attenuated perceptions are reflected in the model by weakened metrical and prolongational derivations.

A complementary approach to the study of the relationship between music and poetry would be to develop a rigorous theory of textsetting.[16] That listeners have strong intuitions

about whether a poetic line is well- or ill-set to music is in itself an indication of the parallels between the two media. Generally, the stronger the correlation between the common formal structures of poetry and music, the more idiomatic the textsetting. That is, a textsetting tends to match up the stresses, metrical positions, durations, phrasings, contours, and formal patterns of a poem and its musical realization. The idiom-specific traditions of *Sprechstimme* in early 20th-century German musical expressionism and of hip hop in current American popular music show how far music and language can merge along these dimensions.

The overall thrust of the Frost analysis is that the mental representation of the sounds of metrical, rhymed poetry and of music, whether texted or not, share a good deal more organization than has usually been supposed. Perhaps this commonality has not been adequately recognized because in certain basic respects the two media are unlike: poetry has words and phrases with propositional meaning; music has a hierarchical, multidimensional pitch space.[14,17] The analysis of the Frost reflects these differences by omitting references both to the poem's imagery or meaning and to musical concepts such as specific pitches and intervals, scales, harmony, key, voice leading, counterpoint, or tonal tension and relaxation.

It seems reasonable to suppose that well-founded theoretical parallels and distinctions have counterparts in brain structure and function. Appropriate theory can thus guide and interpret neurological findings. We may inquire, then, to what extent the commonalities and differences discussed here are instantiated in brain structure and function. The present analysis predicts, as one would expect, that there is not a single music area and a single language area, but that musical and linguistic functions are spread out in different structural and functional components. It further predicts that those brain modules that process rhythm, contour, and timbral relationships are the same in music and language, while those that process purely pitch structures and purely linguistic structures occupy different parts of the brain. Figure 27.17 illustrates.

Current neuropsychological evidence offers some support for this view. Studies of patients with brain lesions that affect musical and linguistic processing indicate that musical and phonological contours are processed in the same area but that musical pitch intervals are processed in different structures, with an apparent one-way channel from contour processing to intervallic pitch processing.[18–20] If so, individuals with deficits that inhibit their ability to process musical or linguistic contour are also unable to process specific pitch relations. This makes intuitive sense, for if a listener cannot tell up from down, how could he or she distinguish between a minor third and a perfect fourth? However, individuals unable to process musical pitch might still be able to process contour. People who are musically tone-deaf do not speak in a monotone. (In this connection, tone languages such as Chinese depend not on fixed pitch categories and intervals but on relative pitch height and contour.[21] Thus, a person could remain competent in speaking a tone language despite an inability to distinguish musical pitch relationships.)

Neuropsychological evidence also suggests a bifurcation in musical processing between pitch and rhythm. There is partial justification for this from music theory, in which grouping and metre are distinct components from both tonal reduction and tonal pitch space.[2,14] Support for this division comes not only from neuropsychology but from experiments in music cognition.[22] In the details, however, the story is bound to be more complex. From a

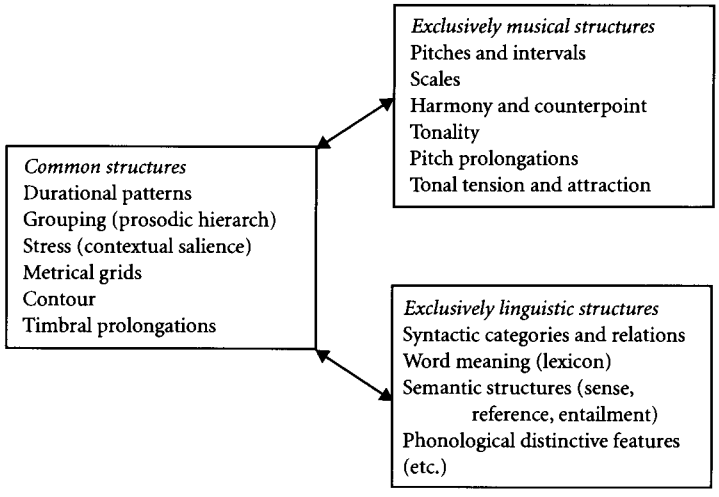



**Figure 27.17** Hypothesized brain organization of musical and linguistic structures.

music-theoretic perspective, pitch is one of the inputs to the derivation of grouping and meter. For example, a tonal cadence signals a grouping boundary, and harmonic rhythm is a strong determinant of metre. Similarly, the derivation of pitch reduction depends on the position of pitches and chords within nested time spans based on grouping and metre. Yet tonal space itself—the system (or systems) of pitches, intervals, scales, chords, tonal regions, tonal attractions, degrees of sensory and cognitive consonance, and degrees of mutual proximity and stability—remains, at a theoretical level, independent of duration, grouping, and metre. Figure 27.17 projects this overall picture: except for contour and timbral prolongations, all of the items listed under 'common structures' belong to the domain of rhythm, which music and poetry share, and most of the items listed under 'exclusively musical structures' belong to tonal space.

Note that the category 'exclusively linguistic structures' in Figure 27.17 includes not just syntactic and semantic features but some phonological aspects as well, such as distinctive features.[8] Moreover, the perceptual organization of speech may not fully submit to a general, Gestalt-based auditory account.[23] These qualifications nevertheless do not undermine the convergence between theoretical and empirical treatments for the domains listed under 'common structures.'

Presumably these exclusive and common structures are a consequence of human evolution. In this view, the roots of music and language are the same, in the form of premusical and prelinguistic communicative and expressive auditory gestures involving shapes of duration, stress, contour, timbre, and grouping.[24] We still communicate with infants and higher mammals in this manner. These elementary shapes appear to lie at the basis of expressive utterance in language and of musical expression (an idea that goes back at least to Rousseau[25]). With evolution came specialization: music and language diverged in their most characteristic features, pitch organization in music and word and sentence meaning

in language. Poetry straddles this evolutionary divergence by projecting, through the addition to ordinary speech of metrical and timbral patterning, its common heritage with music.

## References


1. **Winn, J.** (1981) *Unsuspected Eloquence: A History of the Relations between Poetry and Music.* New Haven: Yale University Press.
2. **Lerdahl, F.** and **R. Jackendoff** (1983) *A Generative Theory of Tonal Music.* Cambridge, MA: MIT Press.
3. **Lerdahl, F.** and **J. Halle** (1991) Some lines of poetry viewed as music. In J. Sundberg, L. Nord, and R. Carlson (eds) *Music Language, Speech, and Brain.* Wenner-Gren International Symposium Series. Longon: Macmillian.
4. **Lerdahl, F.** and **X. Chabot** (1991) A theory of poetry as music and its exploration through a computer aid to composition. *Proc. Intl Computer Music Conf.*
5. **Hayes, B.** Personal communication.
6. **Hayes, B.** (1989) The prosodic hierarchy in poetry. In P. Kiparsky and G. Youmans (eds) *Phonetics and Phonology: Rhythm and Meter.* New York: Academic.
7. **Liberman, M.** and **A. Prince** (1977) On stress and linguistic rhythm. *Linguistic Inquiry* 8, 249–36.
8. **Chomsky, N.** and **M. Halle** (1968) *The Sound Pattern of English,* second printing. Cambridge, MA: MIT Press.
9. **Oehrle, R. T.** (1989) Temporal structures in verse design. In P. Kiparsky and G. Youmans (eds) *Phonetics and Phonology: Rhythm and Meter.* New York: Academic.
10. **Becker, J.** and **A. Becker** (1979) A grammar of the musical genre *Srepegan. J. Music Theory* 23, 1–44.
11. **Handel, S.** (1989) *Listening.* Cambridge, MA: MIT Press.
12. **Pierrehumbert, J.** (1980) *The Phonology and Phonetics of English Intonation.* MIT Ph.D. Dissertation.
13. **Ladd, D. R.** (1996) *Intonational Phonology.* Cambridge, UK: Cambridge University Press
14. **Lerdahl, F.** (2001) *Tonal Pitch Space.* New York: Oxford University Press.
15. **Lerdahl, F.** (1987) Timbral hierarchies. *Contemporary Music Rev.* 1, 135–60.
16. **Halle, J.** and **F. Lerdahl** (1994) A generative textsetting model. *Curr. Musicol.* 55, 3–26.
17. **Krumhansl, C. L.** (1990) *Cognitive Foundations of Musical Pitch.* New York: Oxford University Press.
18. **Peretz, I.** (1993) Auditory agnosia: a functional analysis. In S. McAdams and E. Bigand (eds) *Thinking in Sound: The Cognitive Psychology of Human Audition.* Oxford: Oxford University Press.
19. **Peretz, I.** (2001) Le cerveau musical. In J. J. Nattiez (ed.) *Enciclopedia Einaudi della musica.*
20. **Patel, A. D.** and **I. Peretz** (1997) Is music autonomous from language? A neuropsychological appraisal. In I. Deliége and J. Sloboda (eds) *Perception and Cognition of Music.* Hove, UK: Psychology Press.
21. **Yung, B.** (1991) The relationship of text and tune in Chinese opera. In J. Sundberg, L. Nord, and R. Carlson (eds) *Music, Language, Speech, and Brain.* Wenner-Gren International Symposium Series. London: Macmillan.

22. **Palmer, C.** and **C. L. Krumhansl** (1987) Independent temporal and pitch structures in determination of musical phrase. *J. Exp. Psychol. Hum. Percept. Perform.* 13, 116–26.

23. **Remez, R. E., P. E. Rubin, S. M. Berns, *et al.*** (1994) On the perceptual organization of speech. *Psychol. Rev.* 101, 129–56.

24. **Brown, S.** (2001) The 'Musilanguage' model of music, In N. L. Wallin, B. Merker, and S. Brown (eds) *The Origins of Music.* Cambridge, MA: MIT Press.

25. **Rousseau, J.-J.** (1960) Essai sur l'origine des langues. In J. H. Moran and A. Gode (trans.) *On the Origin of Language.* Chicago: University of Chicago Press.

CHAPTER 28

# DOES EXPOSURE TO MUSIC HAVE BENEFICIAL SIDE EFFECTS?

E. GLENN SCHELLENBERG

## Abstract

Reports that exposure to music causes benefits in nonmusical domains have received widespread attention in the mainstream media. Such reports have also influenced public policy. The so-called 'Mozart effect' actually refers to two relatively distinct phenomena. One concerns short-term improvements in spatial abilities that are said to occur after listening to music composed by Mozart. The other refers to the possibility that formal training in music yields nonmusical benefits. A review of the relevant findings indicates that the short-term effect is small and unreliable. Moreover, when it is evident, it can be explained by between-condition differences in listeners' arousal level or mood. By contrast, the effect of music lessons on nonmusical aspects of cognitive development is still an open question. Several studies have reported positive associations between formal music lessons and abilities in nonmusical (e.g. linguistic, mathematical, and spatial) domains. Nonetheless, compelling evidence for a causal link remains elusive.

## Music and nonmusical abilities

The present report evaluates claims that exposure to music produces benefits in nonmusical domains. These claims began to influence public policy as soon as they came to public notice.[1] For example, Zell Miller, the former Governor of Georgia, budgeted for the distribution of classical music recordings to each infant born in state. Moreover, Florida mandates daily doses of classical music in state-run preschools.

Researchers (e.g. Ref. 2) and journalists (e.g. Refs 3–6) have generated confusion by failing to clarify the distinction between the short-term consequences of music listening and the long-term consequences of formal training in music. Indeed, results from both types of studies have been merged to yield the dictum, 'music makes you smarter'. But passive listening to music, a ubiquitous activity, bears little resemblance to formal training, which involves lessons and systematic practice (see also Ref. 7). Thus, separate evaluation of the short-term benefits of musical exposure and the long-term side effects of music lessons could help to clarify the issues.

Examination of the effects of previous experience on learning and behaviour has a rich tradition in the history of psychology. 'The *transfer* of training from old to new situations is part and parcel of most, if not all, learning. In this sense the study of transfer is coextensive with the investigation of learning'[8] (p. 1019, emphasis added). In addition, hundreds of

psychological examinations of *priming* have investigated how prior exposure to a stimulus affects subsequent processing of the same stimulus or a closely related stimulus.[9,10] If exposure to music causes welcome side effects, we would expect such effects to arise from transfer or priming. Moreover, some researchers who argue for the nonmusical benefits of exposure posit a specific neuropsychological basis for such benefits.[2,11–13] Presumably, this hypothesized cortical process would be compatible with other cortical processes that are demonstrably relevant to music.

## Transfer and priming

Transfer and priming occur in positive and negative forms. Positive transfer occurs when previous experience in problem solving makes it easier to solve a new problem,[8,14] typically by accelerating learning. As such, positive transfer describes successful generalization of a process or strategy. One example involves reasoning with analogies.[15] Previous exposure to analogies (e.g. *Lawyer* is to *client* as *doctor* is to ????; the correct answer is *patient*) can lead to greater success at finding the missing piece in new analogies. Similar research is available on metaphor and the transfer of skills.[14] A common theme across transfer effects is *similarity*;[14,16] positive transfer is more likely to occur when there are more similarities between the old and new problems.

Negative transfer is the opposite of its positive counterpart; previous experience interferes with solving a new problem.[8] Negative transfer, which is often called *interference*, can occur proactively or retroactively. Proactive interference is evident when previous learning makes subsequent learning relatively difficult. For example, a new problem is approached with an old mental set that is inefficient or inappropriate for the new context. By contrast, retroactive interference refers to difficulty accessing mental representations because of intervening experience between initial encoding and retrieval.

Roughly speaking, priming can be considered the 'short term' or 'low level' relative of transfer. Anderson[17] defines priming as 'an enhancement of the processing of a stimulus as a function of prior exposure' (p. 459). In a classic experiment,[18] participants were asked to identify words presented briefly in the visual modality. Performance was superior for words that were seen prior to the word-identification test. The 'low level' nature of this sort of priming is evident in greater priming effects following open-ended instructions (e.g. study the word) compared to compulsory semantic processing (e.g. generate an antonym), the latter condition involving 'deeper' levels of processing.[19] It is clear that priming does not require conscious awareness, as reflected in the priming effects observed in amnesics.[20]

Negative priming refers to situations in which the processing of a 'target' stimulus is *inhibited* by prior exposure.[21] For example, when participants are presented with two words (a 'target' and a 'distractor') and required to name only one (the target), performance on subsequent trials is relatively slow when the target word was previously a distractor. Most priming studies examine *repetition* priming, or subsequent processing of an identical stimulus.[10] Nonetheless, cross-modal and cross-language priming effects are also observable. For people who are bilingual in Spanish–English, auditory presentation of a partial sentence in Spanish (the priming stimulus) can facilitate visual recognition of a target English word, provided that the target was implied by the sentential prime.[22] There are

higher level, *associative*[17] or *semantic*[9] priming effects, in which word processing (e.g. 'butter') is facilitated by previous presentation of an associated word (e.g. 'bread').[23]

This brief review of transfer and priming provides a context for evaluating specific claims about exposure to music. These claims posit remarkable positive side effects of exposure to certain types of music, side effects that, in principle, are closely related to transfer and priming. My intention is to situate claims that *music makes you smarter* in the context of cognitive psychology, which will permit a review and evaluation of such claims with reference to well established cognitive phenomena. A secondary goal is to situate the neuronal mechanism advanced as the basis of associations between music and nonmusical abilities in the domain of cognitive neuropsychology.

## The Mozart effect

The current debate about musical exposure and its side effects was inspired, in part, by Rauscher *et al.*[24] who reported that brief exposure (10 min) to a Mozart sonata generates short term increases in spatial-reasoning abilities (the *Mozart effect*). Each participant in their study was tested in three conditions. Participants in one condition listened to a Mozart sonata before completing three tests of spatial abilities. Participants in the other two conditions listened to a relaxation tape or sat in silence before completing the tests. Performance on the first spatial test (but not the next two) was superior in the 'Mozart' condition. This finding attracted considerable attention because it appeared in a highly prestigious journal, *Nature*, and because the investigators translated their finding into an IQ-score improvement of approximately eight points (i.e. half a standard deviation). Indeed, the popular conclusion that 'music makes you smarter' followed directly from this IQ translation.

Closer examination of the method of Rauscher *et al.*[24] raises questions about the validity of their findings. The choice of comparison conditions is particularly problematic. Sitting in silence or listening to a relaxation tape for 10 min is less arousing or interesting compared to listening to Mozart. Moreover, mood-states are known to influence performance on problem-solving tasks, with superior performance associated with positive affect.[25–27] Thus, the effect could have arisen from differences in arousal or mood rather than from exposure to Mozart.

Because the Mozart effect is at odds with the literature on priming and transfer, alternative explanations of the source of the effect (i.e. the arousal/mood hypothesis) seem all the more credible. Improved spatial skills following exposure to a Mozart sonata do not represent an instance of repetition priming (i.e. the priming stimulus was not repeated). Nor are they an instance of associative priming. How is passive listening to a musical stimulus 'associated' with performance on a visually presented test of spatial skills? Evidence for associative priming typically involves pairs of words with an obvious semantic association (nurse–doctor, bread–butter). How, then, could an auditory (musical) stimulus prime performance on a task with no obvious link to music? Transfer as an explanatory framework also raises more questions than it answers. Transfer typically involves applying a learned skill or strategy to a new context. But what is learned by listening passively to a piece of music? Something about the music, no doubt, but it is difficult to rationalize how the transfer of such knowledge could yield improved performance on a spatial task.

In short, the Mozart effect is a radical claim about cognitive processes that is difficult to reconcile with known principles and findings in cognitive psychology. It comes as no surprise, then, that attempted replications have produced mixed results (for reviews see Refs 11, 28–30). Although many published studies have failed to replicate the effect, a meta-analysis that included some unpublished studies concluded that the effect was moderate but robust.[29] Because the Mozart effect studies have been reviewed elsewhere, the present report focuses on the issues raised by selected studies.

Consider the replication reported by Rauscher *et al.*,[13] who pretested their participants with a Paper-Folding-and-Cutting (PF&C) test (one of the spatial tests used in the original study). Participants were then divided into three groups of equivalent abilities. One group heard Mozart during three subsequent test sessions. A second group sat in silence during the three sessions. A third group heard a minimalist piece by Philip Glass during the first session, an audio-taped story in the second session, and a repetitive piece of dance music in the third session. After each session, the PF&C test was administered again. Although the Mozart group showed a significantly larger improvement in performance than did the other two groups after the first session, there was no difference between the Mozart and comparison groups after the next two sessions. The advantage of Mozart over silence and Glass conditions in the first test session did not extend or clarify the original finding. Participants may find repetitive, minimalist music as boring or unarousing as silence. The null findings in the second and third sessions also raise doubts about the reliability of the effect.

Rauscher[11,31] suggests that the numerous replication failures can be explained primarily by differences in the spatial tasks that have been used as outcome measures. She claims that the effect can be obtained with 'spatial-temporal' tasks (e.g. the PF&C task and other tasks involving mental transformation of visual images), but not with 'spatial-recognition' tasks. This distinction is based on the idea that perceiving and remembering music involves identifying changes and systematic transformations in musical patterns (e.g. motives) that occur over time. Thus, 'transfer' from music listening to the spatial domain should be limited to tasks involving mental manipulation of visual images, which also takes time. Indeed, the time required is linearly related to the amount of manipulation.[32] This distinction is curious in light of the original findings,[24] which indicated that the effect was identical across spatial tasks, temporal or otherwise. In a subsequent reanalysis of the original data,[11] however, the advantage of the Mozart effect proved to be significant on only one of the three spatial tests that were administered, the 'temporal' PF&C task, but not on the two nontemporal tests. Nonetheless, mean scores were highest in the Mozart condition across tests, and the design precluded tests of the two-way interaction between the listening conditions and the spatial tests. In other words, despite their conclusion and interpretation, the data did not support Rauscher's hypothesis (i.e. that the influence of Mozart's music on spatial abilities depends on the temporal nature of the tasks). Moreover, the temporal/nontemporal distinction cannot explain why several attempts to replicate the original findings failed to do so, even though the outcome measure was a task that met Rauscher's criteria for spatial-temporal status.[33–35] Finally, the distinction does not address the problem that the effect, when evident, may be a consequence of differences in arousal or mood.

In *all* cases in which the Mozart effect has been evident, comparison conditions involved repetitive music, sitting in silence, or listening to relaxation tapes. As noted, these comparison

conditions might seem boring to participants (compared to listening to music), promoting relatively low levels of cognitive arousal or negative mood states. As a first attempt to address this possibility, Nantais and Schellenberg[36] replicated and extended the original findings. In their first experiment, each participant was tested on the PF&C task twice, once after listening to 10 min of music and once after sitting in silence for 10 min. For some participants, the music was the same Mozart piece used by Rauscher and her colleagues. For others, a piece by Schubert (from the same compact disk as the Mozart piece, performed by the same pianists) was used instead. This experiment was also the first to use a computer-controlled procedure administered to participants individually. Indeed, the potential impact of group dynamics on the results of earlier studies is unknown.[13,24,33,35,37] (Imagine a classroom of undergraduates being required to sit in silence for 10 min!)

As shown in Figure 28.1 (upper left panel), performance on the PF&C test was better after listening to Mozart than after sitting in silence. In other words, the Mozart effect was

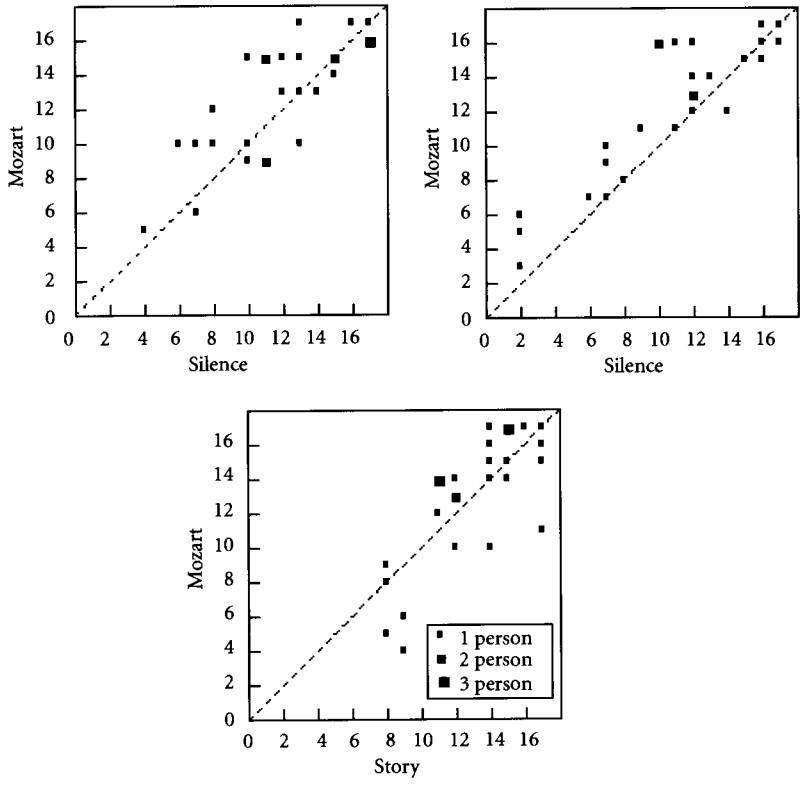



Figure 28.1 Scores on the paper-folding-and-cutting task for participants tested by Nantais and Schellenberg.[36] Each participant was tested twice. The upper left panel illustrates scores after listening to Mozart or sitting in silence. A piece by Schubert was used instead of Mozart for other participants (upper right panel). A third group was tested after listening to either Mozart or a narrated story (lower panel). The line on the diagonal represents equivalent performance across conditions.

replicated. Nonetheless, an identical effect was evident when the Mozart composition was substituted with the piece by Schubert (Figure 28.1, upper right panel). One would predict such a 'Schubert effect' if the comparison condition (silence) was depressing levels of performance. For both groups, performance also improved from the first to the second testing session, revealing a simple practice effect. In a second experiment, a Mozart condition was contrasted with a comparison condition that involved listening to a narrated short story (potentially as engaging as listening to music) instead of sitting in silence. The Mozart effect disappeared (Figure 28.1, lower panel), as one would predict if the experimental (Mozart) and comparison (story) conditions were equally engaging, and if the source of the Mozart effect stemmed from differences in arousal or mood. Perhaps even more important was the finding that performance interacted with listeners' preferences. Those who preferred Mozart over the story performed better on the PF&C test after listening to Mozart. Those who preferred the story performed better after listening to the story (Figure 28.2). These findings provide support for the suggestion that short-term effects of music on tests of spatial abilities stem from differences in arousal or mood rather than from listening to Mozart. Although the figure implies that participants who preferred Mozart performed better regardless of condition, the main effect of preference was marginal ($p = 0.09$).

Further support for the 'arousal or mood' hypothesis comes from a meta-analysis of 20 Mozart-silence comparisons.[28] Successful replications of the Mozart effect were attributed to cognitive arousal, which is predominantly a right-hemisphere function,[38–40] as are tests of complex spatial abilities.[40,41] This view helps to explain why the Mozart effect tends to be slightly larger when the control condition consists of relaxation instructions, which are designed to reduce arousal, instead of sitting in silence.[28]

Another way to interpret the Mozart effect is provided by a new theory based on a large body of findings on the association between mood and cognition.[42] The theory proposes that positive mood states increase circulating levels of the neurotransmitter dopamine.

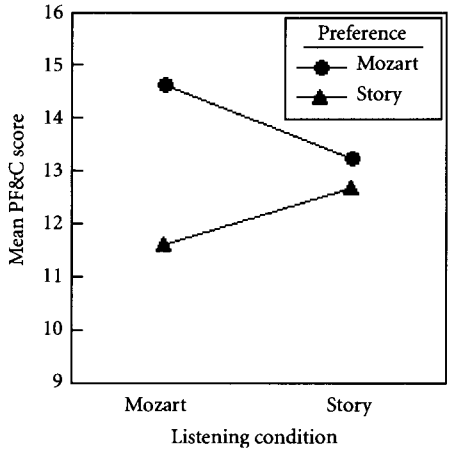



**Figure 28.2** Paper-folding-and-cutting (PF&C) scores from Nantais and Schellenberg's[36] participants as a function of testing condition (Mozart or story) and preference (Mozart or story).

During periods of positive affect, dopamine is released from the ventral tegmental area, which has projections to the prefrontal cortex. A variety of cognitive tasks that show improvement when positive affect is induced[26] may be influenced by the effects of dopamine on prefrontal function. It is possible, then, that the Mozart effect is another way in which positive affect influences performance in a problem-solving task. In short, although the seemingly mysterious Mozart effect may indeed have a neuropsychological explanation, listening to music is just one of many ways to induce arousal or positive affect.

The meta-analysis presented by Chabris[28] and the results of Nantais and Schellenberg[36] are consistent with the idea that differences in arousal or mood are the actual source of the Mozart effect, but neither report tested this hypothesis directly. Thompson, Schellenberg, and Husain[43] attempted such a test using the PF&C task as their outcome measure. Each of their participants was tested once in a music condition and once in a silence condition (as in, Ref. 36, Experiment 1). Arousal and mood were measured after listening to the music using the Profile of Mood States[44] and a subjective rating scale. Participants were also asked to rate how much they enjoyed the music. For some participants, the music condition consisted of the same Mozart piece used in the original Mozart-effect study; for others, a piece by Albinoni was used instead. Albinoni's 'Adagio' was selected because it is considered to be a stereotypical example of slow, sad-sounding music.[45] By contrast, the Mozart sonata is pleasant and happy sounding. Hence, the prediction was that increases in performance on the PF&C task would be evident for music compared to silence in the Mozart group but not in the Albinoni group.

This prediction was upheld by the data. As shown in Figure 28.3, the Mozart group showed a robust improvement in PF&C scores after listening to Mozart compared to sitting in silence. By contrast, the Albinoni group performed more-or-less identically in the music and silence conditions. More importantly, the advantage of the music over the

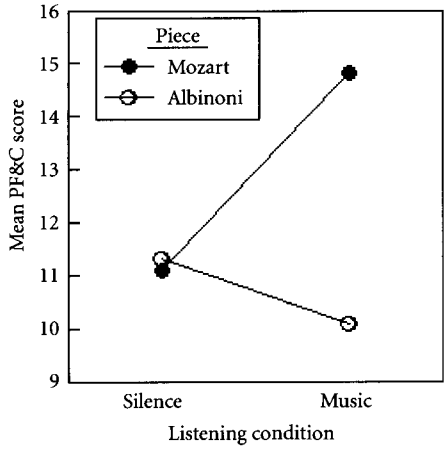



Figure 28.3 Scores on the paper-folding-and-cutting (PF&C) task for participants tested by Thompson *et al.*[43] Each participant was tested twice, once after listening to music and once after sitting in silence. For some participants, the musical piece was the Mozart sonata used by Rauscher *et al.*[24] For others, it was Albinoni's Adagio.

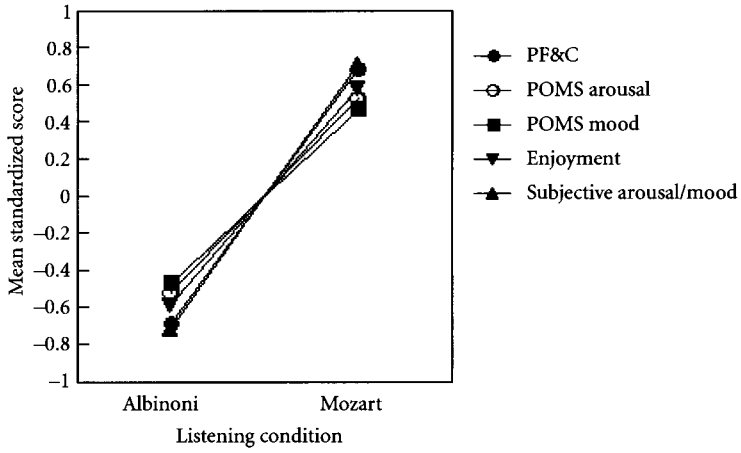



Figure 28.4 Standardized scores for Thompson *et al.*'s[43] participants after listening to a piece composed by Mozart or a piece composed by Albinoni. The variables include the paper-folding-and-cutting (PF&C) task, two subscales from the Profile of Mood States that measured arousal and mood, enjoyment ratings, and a subjective measure of arousal and mood.

silence condition in the Mozart group disappeared when the arousal, mood, or enjoyment variables were held constant. As shown in Figure 28.4, differences between the Mozart and Albinoni conditions on the post-test measures of arousal, mood, and enjoyment were virtually identical to between-condition differences in performance on the PF&C task. In short, the results were completely consistent with the notion that the Mozart effect is an epiphenomenon of arousal or mood.

The theoretical framework that Rauscher and her colleagues use to explain the Mozart effect is called the 'Trion model'.[46,47] The model states that specific cortical firing patterns are present over large areas of the cortex when one performs, composes, or listens to music. Because these patterns are considered to be spatial-temporal in nature, they are said to be highly similar to patterns evident during spatial-temporal reasoning. Both processes involve perceiving and thinking about rule-governed transformations that occur over time. The model describes more than a simple associative or connectionist network, in which one group of neurons is connected to another group. Rather, it posits actual similarities in cortical firing patterns for (1) passive listening to music and (2) actively participating in a task requiring spatial-temporal reasoning.

If we examine the neuropsychological research on music processing, however, the basic tenets of the Trion model seem implausible. The research of Peretz and her colleagues is particularly relevant. Peretz has shown that much of music perception and cognition is relatively modular, and, moreover, that individual aspects of music cognition are relatively modularized and independent of other aspects.[48] For example, melody and rhythm are processed independently and in different parts of the brain,[49–51] lyrics are processed independently of tunes[52] and perceiving musical emotion is independent of memory for music.[53,54] Most importantly, Peretz has studied brain-damaged patients with amusia and

none has exhibited accompanying deficits in spatial abilities. For example, one of her amusic patients could not discriminate tones that differed by gross differences in pitch yet she continued to drive safely around Montréal. In short, there is substantial evidence for modularity of music processing and for independence of various aspects of music. Such evidence is inconsistent with the notion that cortical activity is similar across a variety of musical activities (performing, composing, and listening), and that such patterns of activation are identical during spatial-temporal reasoning.

### Long-term side-effects of music lessons

Although the short-term Mozart effect appears to be independent of Mozart in particular and of music in general, it is still possible that positive, relatively long-term cognitive side effects result from taking music lessons. Indeed, the two issues may be orthogonal. To anticipate the conclusion, the relevant findings reviewed below are consistent with the idea of an association between musical training and nonmusical benefits (see also Ref. 7), but they fall far short of being conclusive.

**Musical abilities and nonmusical abilities** Several studies have examined whether musical *ability* (rather than musical *training*) is correlated with other kinds of abilities. Positive associations imply that improving one's musical ability through formal lessons would be accompanied by nonmusical benefits. In correlational designs, however, it is always impossible to make firm conclusions about the direction of causation when associations are discovered. It is also impossible to rule out the possibility that the association stems from a third, unidentified variable.

Gromko and Poorman[55] examined children between the ages of 4 and 13 who were enrolled in a private school. Their goal was to determine whether musical aptitude is related to children's ability to use symbols. In an initial testing session, children completed the tonal subtest of Gordon's[56,57] musical aptitude measures. During a second session, children were tested on two tasks, one that required them to match short melodies with graphic representations and another that required them to draw graphic representations of the contour of short melodies. Performance on all three measures improved with age, and each measure was significantly correlated with the other two. These findings confirm that children's musical aptitude is predictive of their ability to interpret and produce symbolic representations of music. Because each of the outcomes was associated with age, however, it is impossible to determine whether the associations would still be in evidence if differences in age were held constant (i.e. the authors did not report partial correlations).

In an examination of performance on musical and spatial tasks that required analogical reasoning, children from 6 to 12 years of age were tested on their ability to transfer a given relation between one pair of stimuli to a novel pair.[58] As the age of the children increased, performance on both tasks improved. Moreover, age-related improvements were virtually identical across tasks. As with the study by Gromko and Poorman,[55] however, the association between the music and spatial tasks could be a consequence of the fact that older children performed better on both tasks.

Lamb and Gregory[59] studied the association between reading and musical abilities in a sample of 5-year-old children. Reading abilities and phonemic awareness were positively

associated with pitch-discrimination abilities but not with the ability to discriminate timbres. These associations remained in evidence when differences in age and nonverbal intelligence were held constant. Virtually identical associations between reading abilities and musical abilities (with differences in age and IQ held constant) were reported for a sample of 9-year-old children.[60] Although these findings do not address the issue of causation, they provide evidence of an association between reading and musical abilities that is independent of age or general intelligence.

Douglas and Willatts[61] tested a sample of 8-year-olds to examine whether literacy and musical ability are associated. Pairs of tones were presented in a pitch-discrimination task that required children to identify whether the second tone was higher, lower, or the same as the first. A rhythm-discrimination test required children to respond 'same' or 'different' to pairs of sequences played on a wood block. Literacy was measured with tests of reading and spelling. All measures showed significant pairwise correlations. When differences in receptive vocabulary were held constant, however, reading and spelling measures were associated with rhythm-discrimination abilities but not with pitch-discrimination abilities. Whereas these findings suggest that rhythm-discrimination abilities are better than pitch-discrimination abilities at predicting literacy, the results of Lamb and Gregory[59] imply that pitch-discrimination abilities are a better predictor than timbre-discrimination abilities.

Finally, Lynn, Wilson, and Gault[63] examined the association between musical aptitude and general intelligence (Spearman's *g*) in groups of children 10 years of age. Children were administered rhythm- and pitch-discrimination tasks as well as tests of general intelligence. Each of the music measures was positively associated with each of the measures of intelligence. These results suggest that musical aptitude is a function of general intelligence. Alternatively, musical aptitude may be a valid estimate of *g*. Although the association between musical *aptitude* and intelligence is provocative, it remains to be seen whether music *lessons* actually promote improvements in cognitive abilities.

**Music lessons and nonmusical abilities: correlational and quasiexperimental studies** Other researchers have tested the possibility that music lessons are associated with nonmusical abilities. Again, because we can never be sure that those with and without musical training are identical on other potentially relevant dimensions (e.g. socioeconomic status and overall IQ), unequivocal determinations of causation are impossible.

A classic example of a relevant quasiexperiment is Chan, Ho, and Cheung's[62] study of female college students in Hong Kong (mean age of 20 years). The authors compared the verbal and visual memory abilities of women with no musical training to those of women who had taken 6 years of music lessons before the age of 12. Although the groups did not differ on the visual-memory task, the musically trained group outperformed the untrained group on the verbal-memory task. Unfortunately, despite the authors' claim that the groups were matched according to years of education (with alpha = 0.01), closer inspection of the findings revealed that the musically trained group had significantly more education (with alpha set to a standard 0.05 value). In other words, it is impossible to determine whether the verbal advantage stemmed from music lessons rather than from additional years of education. Indeed, we would predict better verbal skills to accompany higher levels of education.

Hassler, Birbaumer, and Feil[64] examined verbal fluency and visual-spatial abilities in children 9–14 years of age, some of whom were taking music lessons. The children were classified into one of three groups: (1) musically talented and capable of composing or improvising, (2) musically talented but not capable of composing or improvising, or (3) nonmusicians. The groups did not differ on a test of spatial relations, but significant differences were found on tests of verbal fluency and visualization abilities, with the musically talented children outperforming the nonmusicians. At a follow-up test two years later, significant differences were found for each of the three outcome variables.[65] Nonetheless, students in the composing/improvising group had more music lessons than the other musically talented group, yet no differences between these groups on the outcome measures were evident. As such, this study provides equivocal support for the idea that music lessons are accompanied by advantages in nonmusical domains.

Two studies compared the nonmusical abilities of children enrolled in a Kodály music program with those of a comparison group who were not taking music lessons.[66,67] The Kodály program is known for intensive training and placing great emphasis on singing and the development of sequential skills. The program also incorporates clapping, the use of hand signs, and simple musical notation. Hurwitz and his colleagues examined the sequencing and spatial skills of a group of 7-year-olds. Children in the Kodály group had taken music lessons for approximately 7 months, with 40-minute lessons 5 days per week. The sequencing task involved tapping mechanical keys in a regular manner, or in time with a metronome after the metronome was turned off or its rate had been changed. Children were also given tests of spatial abilities, plus a Stroop-like test of interference. The Kodály group outperformed the comparison children on the Stroop test and on some of the spatial tests. In a separate examination of children who had completed 1 year of Kodály instruction, the Kodály group performed better than a comparison group on a reading test even though the two groups had performed identically a year earlier. A subsequent study of 4- and 5-year-olds' understanding of prenumber concepts showed a benefit of Kodály training only for 5-year-old girls.[67] These results suggest that training in music may lead to nonmusical improvements, yet it is impossible to ascertain whether nonmusical aspects of Kodály training or preexisting differences between groups may have influenced the results.

Schellenberg[68] used a correlational approach to examine whether music lessons are predictive of intellectual development in a group of 147 children ranging in age from 6 to 11 years. The outcome measure was a standard IQ test (Wechsler Intelligence Scale for Children—Third Edition). His approach is noteworthy because he also measured three variables that are likely to be confounded with music lessons: socioeconomic status (measured as family income), parental education (which would be correlated with parental IQ), and time spent in nonmusical extra-curricular activities. Of the four variables he measured, music lessons had the strongest association with IQ ($r = 0.38$, $p < 0.0001$). The data are illustrated in Figure 28.5. Moreover, when the other three potentially confounding variables were held constant, the partial association between music lessons and IQ remained significant. Although these findings are consistent with the proposal that music lessons confer nonmusical benefits, it is impossible to rule out two alternative explanations: (1) children with higher IQs may be more likely to take music lessons, and (2) an as-yet-unidentified variable could be influencing IQ *and* the likelihood of taking music lessons.

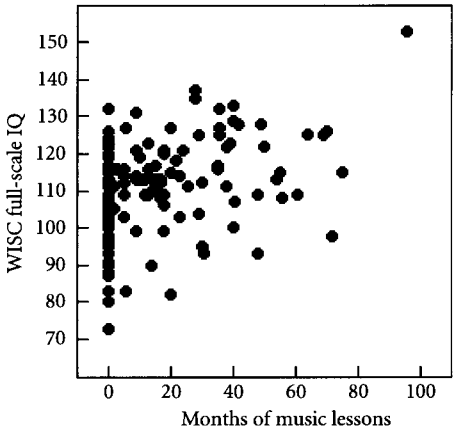



Figure 28.5 IQ scores from children tested by Schellenberg[68] as a function of months of music lessons.

**Music lessons and nonmusical abilities: experimental studies** The next group of studies had more-or-less random assignment of participants to experimental conditions. Thus, provided that comparison conditions were selected appropriately, we should be able to determine whether music lessons actually 'cause' nonmusical cognitive advantages. As with most of the short-term (Mozart effect) studies, however, none of the studies in this group used comparison conditions that preclude the possibility of alternative explanations for the findings.

For example, 6-year-old children who were taught music for 7 months by means of the Kodály method showed improvements in mathematical and reading abilities that surpassed those of children without such training.[69] The researchers' goal was to examine possible by-products of a 'test arts' (Kodály) program that was implemented in some first-grade classes but not in others. They examined two first-grade classes in each of two schools that were designated as 'test arts' classrooms, and another two from both schools that were 'standard arts' classrooms. If we assume that the classrooms were assigned to the two arts programs at random, we can consider the design to approximate a 'true' experiment. The reported advantage for the test-arts classes is remarkable when we consider that in the previous year, children in the test-arts classes were actually *behind* the standard-arts children in terms of the proportion who had reached the national average grade level. Although these results are promising, children in the standard-arts classrooms did not participate in activities focusing on 'sequenced skill development' as did children in the test-arts (Kodály) classrooms. Again, this confounding makes it impossible to attribute the remarkable recovery and achievements of the test-arts classrooms to training in music per se, rather than to other nonmusical aspects of the Kodály program.

In another study, 4-year-old children who received individual 10-min piano lessons once or twice a week for 6–8 months performed better on a test of spatial skills than children assigned to comparison conditions.[12] Nonetheless, other aspects of the design question the reliability of the effect. For example, some of the children had 33 per cent more lessons

than other children, yet this additional training in music had no effect on performance. Moreover, the primary comparison condition involved playing with commercial software programs on a computer. Although a computer instructor provided one-on-one instruction about how to use the computer and open the programs, the software (not the instructor) was designed to teach the children basic skills in reading and arithmetic. As such, superior levels of performance in the piano group could be the consequence of additional instruction from an adult.

Standley and Hughes[70] found that children in prekindergarten classes (4–5 years of age) who took 15 music lessons over a period of 2 months showed enhanced prereading and writing skills compared to other children. Children in the comparison condition were exposed to the regular prekindergarten curriculum but had no additional lessons of any kind. Again, it is impossible to determine whether the observed numerical and verbal benefits arose specifically from music instruction or from pedagogical differences that were independent of musical training. The investigators noted that 'it was also apparent from the children's reaction that the music activities provided pleasure and excitement about academic participation, possibly generating long range motivation for reading and writing' (p. 83). Nonmusical activities that generate similar levels of pleasure and excitement could generate similar increases in motivation.

Gromko and Poorman's[55,71] study of 3- and 4-year-old children enrolled in a private Montessori school is similar to Standley and Hughes'[70] study described above. Children in the music group were provided with weekly group music lessons in addition to the regular curriculum, but the comparison group received no additional lessons of any sort. As such, the modest gains in nonverbal IQ witnessed for the music group relative to the comparison group can be attributed simply to additional educational instruction from an adult.

Three recent experimental studies suffer from similar methodological problems. Each compared young children enrolled in music-education programs with children in 'control' groups who had no comparable extraschool activities.[72–74] One study provided 3 years of piano lessons free of charge to children in the fourth to sixth grade.[73] These 'piano' children performed better than children in a control group on a comprehensive test of cognitive abilities after the first and second years, but the difference disappeared after the third year. Between-group differences during the first 2 years stemmed solely from differences in spatial abilities. In another study, kindergartners were provided with group keyboard lessons for 8 months.[74] The keyboard children showed greater improvement than a control group on tests of spatial abilities, but there was no difference between groups on a test of recognition. A third study examined the influence of a 30-week structured music curriculum on cognitive development.[72] Treatment and control groups of 6-year-olds were administered six subtests from the Stanford-Binet Intelligence Test before and after the curriculum. The treatment group showed relatively larger gains on a single subtest that measured capacity of short-term memory (Bead Memory).

Another recent study examined possible side-effects of group keyboard lessons that were provided free of charge to children 6–8 years of age.[75] A control group had computer lessons with a commercial software program designed to improve English-language skills. Both groups were also given lessons intended to enhance spatial abilities by playing with a software program designed by the researchers. Unfortunately, the main outcome variable

consisted of scores on a testing version of the same spatial software, which has unknown reliability and validity. Moreover, aggregate scores on the outcome tests did not differ between groups. The investigators reported a significant advantage for the keyboard group on a subtest of mathematical fractions and proportions, and they concluded that improved musical and spatial skills lead to improved mathematical abilities. These results would be more convincing if they had been obtained with standardized tests, and if the piano group had performed better overall, or at least on subtests for which clear predictions were made a priori.

In an initial attempt to rectify some of the shortcomings of earlier experiments, Thompson, Schellenberg, and Husain[76] tested whether music lessons influence children's ability to interpret the emotions expressed by nonlinguistic cues (i.e. *prosody*) in speech. Six-year-olds were assigned randomly to a year of music lessons, a year of drama lessons, or no lessons. After the lessons were completed, the children were tested on their ability to label the emotion (happy, sad, fearful, or angry) conveyed by semantically neutral sentences spoken in English (e.g. *the chair is made of wood*) or in Tagalog, and by tone sequences that mimicked the prosody of the sentences. The children with music lessons performed better than the no-lessons children on these tasks. In fact, their performance was equivalent to children who took drama lessons, which focused specifically on using prosody to convey emotions. A separate test of adult participants suggested that such effects are long-lasting. Specifically, adults who started music lessons during childhood performed better on prosody-decoding tasks than did adults with no musical training.

The studies just reviewed provide consistent *suggestive* evidence that music lessons have positive nonmusical side effects. Nonetheless, specifics of the reported associations vary widely from study to study. If we suspend our disbelief, however, and assume that music education affects abilities in other areas, how could we account for this influence?

A number of neurological studies describe ways in which music lessons affect cortical development. Compared to nonmusicians, accomplished players of string instruments show increased representation in the cerebral cortex for the fingers of their left hand,[77] which implies that musical training can alter patterns of cortical organization. Indeed, cortical representations are especially large for those who begin music lessons at an early age when the brain is relatively plastic. Although the size of the corpus callosum is larger in musicians than in nonmusicians, this effect is particularly notable in musicians who began taking lessons before the age of seven.[78] Relatively large brain asymmetries are also evident among musicians who have absolute (perfect) pitch,[79] and this relatively rare ability to name and produce pitches in isolation is evident predominantly among musicians who begin lessons in early childhood.[80] Moreover, the representation of piano tones in the auditory cortex differs in musicians than in nonmusicians,[81] although genetic factors or simple exposure to music could also play a role.[82] Finally, specific cortical areas in the right hemisphere are activated when reading a musical score but not when reading one's primary or secondary language.[83]

Consequences of an enriched environment on other species (e.g. rats and mice) include denser patterns of dendritic branching and a greater number of hippocampal neurons.[84,85] If music education represents an enrichment of a child's environment, such enrichment could promote neurological development, which could, in turn, influence abilities in other

domains. Music, however, is simply one of many ways to enrich a child's environment. Moreover, music education is a complex process that involves many different dimensions.

We know that schooling improves a wide variety of cognitive skills and that this association is not simply a by-product of maturation.[86–89] For young children in particular, schooling is more effective in smaller classes.[90–92] Reviews of intervention programs for children who are at risk of academic failure suggest that extended one-on-one contact with a supportive adult is a common feature of successful interventions.[93,94] Thus, music lessons, which are typically taught individually or in small groups, may confer nonmusical benefits for children by providing close and extended contact with an adult other than a parent or teacher. If this is the case, then similar side effects should be evident with other types of lessons that provide similar levels of contact (e.g. chess, drawing).

Music lessons may be unique, however, because of their focus on a particular combination of factors, such as hours of individual practice, learning to read music, attention and concentration, timing, ear training, sight reading, constructive feedback from the instructor, and exposure to music.[95] Thus, positive transfer effects to nonmusical domains, such as language, mathematics, or spatial reasoning, could be similarly unique for individuals who take music lessons. On the other hand, music lessons are likely to improve many *general* skills, such as the ability to attend to rapidly changing temporal information, skills relevant to auditory stream segregation, the ability to detect temporal groups, sensitivity to signals of closure and other gestalt cues of form, emotional sensitivity and expressiveness, and fine motor skills. These general skills should be particularly likely to transfer to a variety of nonmusical domains.

As someone who took music lessons from the age of five and practised regularly for the next 11 years, I feel changed—probably for the better—in ways that seem specific to my involvement with music. It remains to be seen, however, whether this personal observation will withstand the test of rigorous experimental investigation.

## Acknowledgements

Preparation of this article was supported by a grant from the International Foundation for Music Research. Bill Thompson and Sandra Trehub provided helpful comments on an earlier draft.

## References

1. Kolata, G. (8 August 1999). Muddling fact and fiction and policy. *New York Times* WK5.
2. **Shaw, G. L.** (2000) *Keeping Mozart in Mind.* San Diego: Academic Press.
3. **Calamai, P.** (26 August 1999) Mozart theory creates discord in the scientific community. *Toronto Star* A28.
4. **Ingram, J.** (11 May 1997) More proof music lifts young IQs. *The Toronto Star* F8.
5. **Martin, S.** (October, 1994) Music lessons enhance spatial reasoning skills. *APA Monitor* 5.
6. **Motchuk, A.** (15 March 1997) Can Mozart make maths add up? *New Scientist* 17.

7. Hetland, L. (2000) Learning to make music enhances spatial reasoning. *J. Aesthetic Education* 34(3/4), 179–238.
8. Postman, L. (1971) Transfer, interference and forgetting. In J. W. Kling and L. A. Riggs (eds) *Woodworth and Schlosberg's Experimental Psychology*, 3rd edn. New York: Holt, Rinehart and Winston, pp. 1019–32.
9. Neely, J. H. (1991) Semantic priming effects in visual word recognition: a selective review of current findings and theories. In D. Besner and G. W. Humphreys (eds) *Basic Processes in Reading: Visual Word Recognition* Hillsdale, NJ: Erlbaum, pp. 264–336.
10. Tulving, E. and D. L. Schachter (1990) Priming and human memory systems. *Science* 247, 301–6.
11. Rauscher, F. H. and G. L. Shaw (1998) Key components of the Mozart effect. *Perceptual and Motor Skills* 86, 835–41.
12. Rauscher, F. H., G. L. Shaw, L. J. Levine, E. L. Wright, W. R. Dennis, and R. L. Newcomb (1997) Music training causes long-term enhancement of preschool children's spatial-temporal reasoning. *Neurol. Res.* 19, 2–8.
13. Rauscher, F. H., G. L. Shaw, and K. N. Ky (1995) Listening to Mozart enhances spatial-temporal reasoning: towards a neurophysiological basis. *Neurosci. Lett.* 185, 44–7.
14. Robins, A. (1996) Transfer in cognition. *Connection Science*, 8, 185–203.
15. Gick, M. L. and K. J. Holyoak, (1983) Schema induction and analogical transfer. *Cognitive Psychol.* 15, 1–38.
16. Gentner, D., M. J. Ratterman, and K. D. Forbus (1993) The roles of similarity in transfer: separating retrievability from inferential soundness. *Cognitive Psychol.* 25, 524–75.
17. Anderson, J. R. (1995) *Cognitive Psychology and its Implications*, 4th edn. New York: Freeman.
18. Jacoby, L. L. (1983) Perceptual enhancement: persistent effects of experience. *J. Exp. Psychol.: Learning, Memory, Cognit.* 9, 21–38.
19. Craik, F. I. M. and R. S. Lockhart (1972) Levels of processing: a framework for memory research. *J. Verbal Learning Verbal Behav.* 11, 671–84.
20. Graf, P., L. R. Squire, and G. Mandler (1984) The information that amnesic patients do not forget. *J. Exp. Psychol.: Learning, Memory, Cognit.* 10, 164–78.
21. May, C. P., M. J. Kane, and L. Hasher (1995) Determinants of negative priming. *Psychological Bulletin* 118, 35–54.
22. Hernandez, A. E., E. A. Bates, and L. X. Avila (1996) Processing across the language boundary: a cross-modal priming study of Spanish-English bilinguals. *J. Exp. Psychol.: Learning, Memory, Cognit.* 22, 846–64.
23. Meyer, D. E. and R. W. Schvaneveldt (1971) Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *J. Exp. Psychol.* 90, 227–34.
24. Rauscher, F. H., G. L. Shaw, and K. N. Ky (1993) Music and spatial task performance. *Nature* 365, 611.
25. Benedict, R. H. B., M. Dobraski, and M. Z. Goldstein (1999) A preliminary study of the association between changes in mood and cognition in a mixed geriatric psychiatry sample. *J. Gerontology* 54B, 94–9.
26. Isen, A. M. (1999) Positive affect. In T. Dalgleish and M. Power (eds) *The Handbook of Cognition and Emotion*. New York: Wiley, pp. 521–39.
27. O'Hanlon, J. F. (1981) Boredom: practical consequences and a theory. *Acta Psychologica* 49, 53–82.

28. **Chabris, C. F.** (1999) Prelude or requiem for the 'Mozart effect'? *Nature* 400, 826–7.
29. **Hetland, L.** (2000) Listening to music enhances spatial-temporal reasoning: evidence for the "Mozart effect". *J. Aesthetic Education* 34(3/4), 105–48.
30. **Jenkins, J. S.** (2001) The Mozart effect. *J. Royal Society of Medicine* 94, 170–2.
31. **Rauscher, F. H.** (1999) Prelude or requiem for the 'Mozart effect'? *Nature* 400, 827–28.
32. **Shepard, R.** and **J. Metzler** (1971) Mental rotation of three-dimensional objects. *Science* 171, 701–3.
33. **Steele, K. M., K. E. Bass,** and **M. D. Crook** (1999) The mystery of the Mozart effect: failure to replicate. *Psychol. Sci.* 10, 366–9.
34. **Steele, K. M., K. M. Brown,** and **J. A. Stoecker** (1999) Failure to confirm the Rauscher and Shaw description of recovery of the Mozart effect. *Perceptual and Motor Skills* 88, 843–8.
35. **Steele, K. M., S. Dalla Bella, I. Peretz, T. Dunlop, L. A. Dawe, G. K. Humphrey, R. A. Shannon, K. L. Kirby, Jr,** and **C. G. Olmstead** (1999) Prelude or requiem for the 'Mozart effect'? *Nature* 400, 827.
36. **Nantais, K. M.** and **E. G. Schellenberg** (1999) The Mozart effect: an artifact of preference. *Psychol. Sci.* 10, 370–3.
37. **Steele, K. M., T. N. Ball,** and **R. Runk** (1997) Listening to Mozart does not enhance backwards digit span performance. *Perceptual and Motor Skills* 84, 1179–84.
38. **Heller, W.** and **J. B. Nitschke** (1997) Regional brain activity in emotion: a framework for understanding cognition in depression. *Cognit. Emotion* 11, 637–61.
39. **Robertson, I. H., J. B. Mattingley, C. Rorden,** and **J. Driver** (1998) Phasic alerting of neglect patients overcomes their spatial deficit in visual awareness. *Nature* 395, 169–72.
40. **Tucker, D. M., A. Hartry-Speiser, L. McDougal, P. Luu** and **D. deGrandpre** (1999) Mood and spatial memory: emotion and right hemisphere contribution to spatial cognition. *Biol. Psychol.* 50, 103–25.
41. **Ditunno, P. L.** and **V. A. Mann** (1990) Right hemisphere specialization for mental rotation in normals and brain damaged subjects. *Cortex* 26, 177–88.
42. **Ashby, F. G., A. M. Isen,** and **A. U. Turken** (1999) A neuropsychological theory of positive affect and its influence on cognition. *Psychol. Rev.* 106, 529–50.
43. **Thompson, W. F., E. G. Schellenberg** and **G. Husain** (2001) Arousal, mood, and the Mozart effect. *Psychol. Sci.* 12, 248–51.
44. **McNair, D. M., M. Lorr,** and **L. F. Droppleman** (1992) *The Profile of Mood States*, 3rd ed. San Diego: Educational and Industrial Testing Service.
45. **Krumhansl, C. L.** (1997) An exploratory study of musical emotions and psychophysiology. *Canadian J. Exp. Psychol.* 51, 336–52.
46. **Leng, X.** and **G. L. Shaw** (1991) Toward a neural theory of higher brain function using music as a window. *Concepts in Neurosci.* 2, 229–58.
47. **Leng, X., G. L. Shaw,** and **E. L. Wright** (1990) Coding of musical structure and the trion model of the cortex. *Music Percept.* 8, 49–62
48. **Peretz, I.** (2001) Music perception and recognition. In B. Rapp (ed.) *The Handbook of Cognitive Neuropsycho.* Hove, UK: Psychology Press, pp. 521–40.
49. **Peretz, I.** (1996) Can we lose memories for music? A case of music agnosia in a nonmusician. *J. Cognitive Neurosci.* 8, 481–96.
50. **Peretz, I.** and **R. Kolinsky** (1993) Boundaries of separability between melody and rhythm in music discrimination: a neuropsychological perspective. *Quarterly J. Exp. Psychol.* 46A, 301–25.
51. **Peretz, I.** and **J. Morais** (1989) Music and modularity. *Contemporary Music Rev.* 4, 277–91.

52. Besson, M., F. Faïta, I. Peretz, A.-M. Bonnel, and J. Requin (1998) Singing in the brain: independence of lyrics and tunes. *Psychol. Sci.* 9, 494–8.
53. Peretz, I. and L. Gagnon (1999) Dissociation between recognition and emotional judgements for melodies. *Neurocase* 5, 21–30.
54. Peretz, I., L. Gagnon, and B. Bouchard (1998) Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognit.* 68, 111–41.
55. Gromko, J. E. and A. S. Poorman (1998) Developmental trends and relationships in children's aural perception and symbol use. *J. Res. Music Education* 46, 16–23.
56. Gordon, E. (1979) *Primary Measures of Music Audiation.* Chicago: G. I. A. Publications.
57. Gordon, E. (1982) *Intermediate Measures of Music Audiation.* Chicago: G. I. A. Publications.
58. Nelson, D. J. and A. L. Barresi (1989) Children's age-related intellectual strategies for dealing with musical and spatial analogical tasks. *J. Res. Music Education* 37, 93–103.
59. Lamb, S. J. and A. H. Gregory (1993) The relationship between music and reading in beginning readers. *Educational Psychol.* 13, 19–27.
60. Barwick, J., E. Valentine, R. West, and J. Wilding (1989) Relations between reading and musical abilities. *British J. Educational Psychol.* 59, 253–7.
61. Douglas, S. and P. Willatts (1994) The relationship between musical ability and literacy skills. *J. Res. Reading* 17, 99–107.
62. Chan, A. S., Y. C. Ho, and M. C. Cheung (1998) Music training improves verbal memory. *Nature* 396, 128.
63. Lynn, R., R. G. Wilson, and A. Gault (1989) Simple musical tests as measures of Spearman's *g*. *Personality and Individual Differences* 10, 25–8.
64. Hassler, M., N. Birbaumer, and A. Feil (1985) Musical talent and visual-spatial ability: a longitudinal study. *Psychol. Music* 13, 99–113.
65. Hassler, M., N. Birbaumer, and A. Feil (1987) Musical talent and visual-spatial ability: onset of puberty. *Psychol. Music* 15, 141–51.
66. Hurwitz, I., P. H. Wolff, B. D. Bortnick, and K. Kokas (1975) Nonmusical effects of the Kodály music curriculum in primary grade children. *J. Learning Disabilities* 8, 167–74.
67. Neufeld, K. A. (1986) Understanding of selected pre-number concepts: relationships to a formal music program. *Alberta J. Educational Res.* 32, 134–9.
68. Schellenberg, E. G. (2003) *Music lessons and IQ.* Paper presented at the biennial meeting of the Society for Research in Child Development, Tampa, FL.
69. Gardiner, M. F., A. Fox, F. Knowles, and D. Jeffrey (1996) Learning improved by arts training. *Nature* 381, 284.
70. Standley, J. M. and J. E. Hughes (1997) Evaluation of an early intervention music curriculum for enhancing prereading/writing skills. *Music Therapy Perspectives* 15, 79–85.
71. Gromko, J. E. and A. S. Poorman (1998) The effect of music training on preschoolers' spatial-temporal task performance. *J. Res. Music Education* 46, 173–81.
72. Bilhartz, T. D., R. A. Bruhn, and J. E. Olson (2000) The effect of early music training on child cognitive development. *J. Applied Dev. Psychol.* 20, 615–36.
73. Costa-Giomi, E. (1999) The effects of three years of piano instruction on children's cognitive development. *J. Res. Music Education* 47, 198–212.
74. Rauscher, F. H. and M. A. Zupan (2000) Classroom keyboard instruction improves kindergarten children's spatial-temporal performance: a field experiment. *Early Childhood Research Quarterly* 15, 215–28.

75. **Graziano, A. B., M. Peterson**, and **G. L. Shaw** (1999) Enhanced learning of proportional math through music training and spatial-temporal reasoning. *Neurol. Res.* 21, 139–52.
76. **Thompson, W. F., E. G. Schellenberg**, and **G. Husain** (2002) *Decoding speech prosody: Do music lessons help?* Manuscript under review.
77. **Elbert, T., C. Pantev, C. Wienbruch, B. Rockstroh**, and **E. Taub** (1995) Increased cortical representation of the fingers of the left hand in string players. *Science* 270, 305–7.
78. **Schlaug, G., L. Jäncke, Y. Huang, J. F. Staiger**, and **H. Steinmetz** (1995) Increased corpus callosum size in musicians. *Neuropsychologia* 33, 1047–55.
79. **Schlaug, G., L. Jäncke, Y. Huang**, and **H. Steinmetz** (1995) In vivo evidence of structural brain asymmetry in musicians. *Science*, 267, 699–700.
80. **Takeuchi, A. H.** and **S. H. Hulse** (1993) Absolute pitch. *Psychol. Bull.* 113, 345–61.
81. **Pantev, C., R. Oostenveld, A. Engelien, B. Ross, L. E. Roberts**, and **M. Hoke** (1998) Increased auditory cortical representation in musicians. *Nature* 392, 811–3.
82. **Monaghan, P., N. B. Metcalfe**, and **G. D. Ruxton** (1998) Does practice shape the brain? *Nature* 394, 434.
83. **Nakada, T., Y. Fujii, K. Suzuki**, and **I. Kwee** (1998) 'Musical brain' as revealed by highfield (3 tesla) functional MRI. *Neuroreport* 9, 3853–6.
84. **Greenough, W. T.** and **F. R. Volkmar** (1973) Pattern of dendritic branching in occipital cortex of rats reared in complex environments. *Exp. Neurol.* 40, 491–504.
85. **Kemperman, G., H. G. Kuhn**, and **F. H. Gage** (1997) More hippocampal neurons in adult mice living in an enriched environment. *Nature* 386, 493–5.
86. **Ceci, S. J.** (1991) How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Dev. Psychol.* 27, 703–22.
87. **Ceci, S. J.** and **W. M. Williams** (1997) Schooling, intelligence, and income. *American Psychologist* 52, 1051–8.
88. **Morrison, F. J., E. M. Griffith**, and **D. M. Alberts** (1997) Nature-nurture in the classroom: entrance age, school readiness, and learning in children. *Dev. Psychol.* 33, 254–62.
89. **Morrison, F. J., L. Smith**, and **M. Dow-Ehrensberger** (1995) Education and cognitive development: a natural experiment. *Dev. Psychol.* 31, 789–99.
90. **Blatchford, P.** and **P. Mortimore** (1994) The issue of class size for young children in schools: what can we learn from research? *Oxford Rev. Education* 20, 411–28.
91. **Ehrenberg, R. G., D. J. Brewer, A. Gamoran**, and **J. D. Willms** (2001) Class size and student achievement. *Psychol. Sci. Public Interest* 2, 1–30.
92. **Mosteller, F.** (1995) The Tennessee study of class size in the early school grades. *Future of Children* 5, 113–27.
93. **Barnett, W. S.** (1995) Long-term effects of early childhood programs on cognitive and school outcomes. *Future of Children* 5, 25–50.
94. **Lazar, I.** and **R. Darlington** (1982) Lasting effects of early education: a report from the consortium for longitudinal studies. *Monographs of the Society for Research in Child Development* 47(2–3), Serial No. 195.
95. **Lamont, A.** (1998) Responses to Katie Overy's paper: respondent 3. *Psychol. Music* 26, 201–4.

# INDEX


absolute (perfect) pitch 6, 34, 38, 390, 443
adaptive value of music 4, 13, 27, 52–3, 60
  nonadaptive pleasure seeking 59–60
  *see also* evolution
age differences
  duration ratios 29
  grouping 25
  regularity 26
  rhythm perception 27
  temporal processing 28
  *see also* infants; children
agraphia 184
alexia 184
amusia 171, 172, 183–5, 193, 205
  congenital 192, 195–8
  expressive instrumental 185
  motor 183
  sensory 183
anisochrony discrimination 207–10
aphasia 183–5, 193
apraxia 185
archaeological findings 62–3
artificial neural networks 109–23
ascending auditory pathway *see* auditory cortex
Asperger autism 68–9
attack 173
auditory agnosia 172
auditory cortex 152–67, 170, 359–60
  ascending pathway 170
  functional imaging 173–5
  lesions of 144–6, 170–3
  plasticity of 357–65
  recordings from 156–60
  *see also* cortical plasticity
auditory evoked potentials 152–3, 155–60
auditory hallucinations 182
auditory imagery 217, 219
  and musical learning 360–1
autism 68–9, 195
  Asperger type 68–9

Bach, Johan Sebastian
  *C Minor Prelude* 102
  *Italian Concerto* 248–51
Bartok, Béla, *Music for Strings, Percussion and Celesta* 85
beat 22, 44–5, 327
  *see also* rhythm; temporal analysis
Beethoven, Ludwig van, *Symphony No. 5* 182
beneficial effects of music 430–48
  long-term side-effects of music lessons 438–44
  Mozart effect 432–8
  on nonmusical abilities 439–44
  transfer and priming 431–2
best frequency 153
blind persons, improved auditory abilities in 357–9
Boulez, Pierre, *Structures* 87
brain lesions 170–3, 185–6, 220–1
  acquired 193–5
  auditory cortex 144–6, 170–3
  auditory evoked potentials 152–67
  congenital 195–8
  cortex 170–3
  epilepsy 181–91, 208–10
  and musical imagery 220–1
  and pitch processing 232–6
brain of musicians 366–81
  functional brain differences 373–6
  structural brain differences 367–73
    cerebellum 369–72
    corpus callosum 368–9
    gray matter volume 372–3
    motor cortex 369
brain plasticity *see* cortical plasticity
brain specialization 192–203
  hemispheric 231–46, 345–7
    in infants 9
    in musicians 373–6
  music-specific networks
    content of 201
    location of 199–200
brain stimulation 198–9
brain substrates *see* cerebral substrates
brain-damaged patients *see* brain lesions

cerebellum 247–68, 277, 369–72
cerebral blood flow *see* magnetic resonance imaging; positron emission tomography

cerebral substrates 204–16
children
  emotional responses to music 318–20
  musical play 51–2
  temporal processing 25–9
  *see also* infants
chords 106, 112, 128, 274
  chord cadences 301
chromatic scales 131
cingulate gyrus 277
complex sounds 27–8, 168–77, 296–9
congenital amusia *see* amusia
connectionist modelling 109–23
consonance 128, 130, 135–6, 140, 142, 144
consonant intervals *see* perfect fifth; perfect fourth
contour 5, 274, 328, 420–2
corpus callosum 368–9
cortical deafness 171
cortical plasticity 351, 357–65, 391–3, 396–409
  *see also* auditory cortex
counterpoint 283
critical bandwidth 138
cultural differences *see* ethnomusicology

dissonance 128, 130
dissonant intervals *see* minor second; tritone
duration 5, 28–9, 81, 424
dyslexia 196
dysphasia 196
dystonia 403–6

electroacoustic music 80, 83
electroencephalography 276, 294–309, 333–4, 383
emotional context of music 310–24
epilepsy
  auditory evoked potentials 152–67
  musicogenic seizures 181–91, 199
  and temporal processing 208–10
ethnomusicology 25, 44, 62, 64–6, 201, 271
event-related potentials *see* electroencephalography
evolution 42–56
  of language 66–7, 270–1
  of music 57–75, 270–1
    survival value 58–9, 181–2
expectations 273–4, 300–1
expertise 301–4
  *see also* cortical plasticity; musicians

Finnish spiritual folk hymns 115
form-bearing dimensions 86–7
function of music *see* evolution
functional aspects of music and language 272
functional magnetic resonance imaging *see* magnetic resonance imaging

gap detection 172
generative grammar theory 275, 276
generative phonology 413–29
Gordon Musical Aptitude Profile 256
gray matter volume 372–3
group identity 51
grouping 24–5, 326–7, 417

handedness in musicians 183, 367–8, 373
harmonic intervals 128, 129, 130, 131
harmony 116, 127–51, 272, 278–9, 283–5
Hebb's rule 357, 383
hemispheric specialization 231–46, 345–7
  in infants 9
  in musicians 373–6
  spectral processing 240–2
  *see also* brain specialization
Heschl's gyrus 173, 186, 200, 277, 360
  electrode recordings of 242
  emotional processing of music 315
  magnetic resonance imaging 233, 234
hierarchical processing 84, 175, 250–4, 257–9, 347–8, 421

imagery
  musical 217–30
  visual 218, 219
infants
  emotional response to music 318–20
  infant-caregiver interactions 47–8
  maternal singing 10–13
  musical memory 32–41
  musical predisposition 3–20, 26, 27, 29
  temporal processing 9, 25–9
  vocal capacities 47–8
instrument recognition 326
intervals 127–51, 131, 274, 329
  duration 23
  infants 5–7
  inter-onset 207–13
  perfect fifth 6, 7, 132–3, 134
  perfect fourth 6, 7, 132, 133, 134
  tritone 6, 7, 132, 133, 135, 141
intonational phrases 414
intracerebral evoked potentials 152–67
invariance in music 84–6
irregularity 26, 27–8

key distances *see* tonal cognition
key-finding models *see* tonal cognition
Kodaly music program 440–1

language
  brain imaging studies 276–8

cognitive perspective 271–5
compared to music 269–93
syntax and harmony 283–5
temporal structure 285–9
evolutionary origin 66–7
evolutionary perspective 270–1
generative grammar theory 275, 276
linguistic stress 414
similarities and differences 270–5
learning music *see* musical training
lesion studies *see* brain lesions
Ligeti, György
*Atmosphères* 80
*Lux Aeterna* 80

magnetic resonance imaging
focal hand dystonia 403, 404
Heschl's gyrus 233, 234
musical imagery 222
musical processing 187
pitch discrimination 233, 234
magnetoencephalography 173–5, 383
auditory cortex 153
language processing 276
mismatch negativity 294–309
making music *see* musical training
maternal music 9–13
melodic contour *see* contour
melodic intervals *see* intervals
melody 278–9, 325–45
memory 23–4, 32–41
mental practice 401–2
Messiaen, Olivier, *Modes de Valeurs et d'Intensités* 85, 86
metre 274, 327, 416
and grouping 417
in poetry 424
minor second 132, 133, 141
*see also* intervals
mismatch negativity 294–309
*see also* electroencephalography
modern music 65
motor cortex 369
motor skill development theory 61
Mozart effect 432–8
MRI *see* magnetic resonance imaging
music agnosia 194
music centres 348–50
music industry 57–75
music-savant syndrome 195, 198
music-specific networks 199–201
Musical Aptitude Profile 186
musical consonance *see* consonance
musical hallucinations 199
musical imagery *see* imagery
musical instruments 62–3, 326
musical intervals *see* intervals
musical performance 3, 247–68
functional neuroanatomy 248–51
comprehension 251–4
musical rhythm 254–60
musical predisposition 3–20
*see also* musical ability
musical scales 7–9, 131, 271, 327–8
*see also* intervals; tonal cognition
musical training 349, 396–409
and cortical plasticity 387–91
auditory representation 388–91
somatosensory representations 387–8
and IQ score 441
learning to play piano 397–401
and nonmusical abilities 430–44
musicians, brain of 366–81, 382–95
functional brain differences 373–6
implicit musical ability 376
structural brain differences 367–73
musicogenic seizures 181–91, 199
*see also* epilepsy

neural coding
pitch relationships 132–8
roughness 138–44
neural processing 152–67, 168–77
neuroimaging 173–5, 247–68
magnetoencephalography *see* magnetoencephalography
MRI *see* magnetic resonance imaging
PET *see* positron emission tomography

octaves 130–1

Penderecki, Krzysztof, *Threnody for the victims of Hiroshima* 80
perfect fifth 6, 7, 132–3, 134
perfect fourth 6, 7, 132, 133, 134
performance *see* musical performance
PET *see* positron emission tomography
physiological response to music 312–14
pitch 271, 274
absolute (perfect) 6, 34, 38, 390, 443
infants' perception of 5
maternal singing 10
mental presentation of 218
modularity in perceived similarity 89–91
neural coding 132–8
perception 152–67
in poetry 421–2, 423
processing 232–40

pitch (*cont.*)
relational processing 5
transposition of 82
variation in 81
Western tonal music 114
pitch discrimination 235, 297
encoding 301–3
neurology of 260–2
patterns 236–40
proximity 329
planum temporale 174, 373–6
plasticity *see* cortical plasticity
poetry and music 413–29
polyphonic music 102–3
positron emission tomography 35–9, 153, 173–5, 187, 219, 221–5, 236, 248–50, 276
prosody 272, 274, 424
psychoacoustics 130–2

Ravel, Maurice
aphasia 184–5
*Boléro* 81
reading, and musical ability 438–9
regularity 25–7, 112–13
Reynolds, Roger
*Archipelago* 80, 83
SPIRLZ algorithm 83
SPLITZ algorithm 83
*The Angel of Death* 81
*The Behavior of Mirrors* 80
*Variation* 83
rhythm 271, 274
comprehension of 251–4
functional neuroanatomy 251–60
modularity in perceived similarity 89–91
perception 26–7
and musical skill 27
reproduction 28
*see also* temporal processing
Rimsky-Korsakoff, Nicholai 183
Risset, Jean-Claude, *Sud* 83
roughness 128, 130, 138–45
*see also* consonance

scales *see* musical scales
Schoenberg, Arnold
*Fourth String Quartet* 87
*Wind Quintet* 87
Scriabin, Alexander 183
Seashore Test of Musical Abilities 185
semantics 278–83
Shebalin, Vissarion 184, 193
similarity in music 84–9
social bonding function of music 67–71
social cohesion theory 61
social context of music 46
sound complexity *see* complex sounds
specialization of brain *see* brain specialization
stress 97
metric 415, 416, 418
superior colliculus 170
superior temporal gyrus 222, 236, 277, 333
supplementary motor area 222–3, 253, 277
supramarginal gyrus 277
survival value of music 58–9, 181–2
synesthesia 183
syntax 272, 273, 283–5

tactus 44–5, 416, 419
*see also* beat
tempo 10, 21, 26, 82
temporal analysis 168–77, 303–4
cerebral substrates 204–16
hierarchical processing 175
in infants 9
intercultural aspects 21–3
predisposition towards regularity 25–6
predisposition towards simple duration ratios 28–9
segmentation and grouping 24–5
temporal microstructure of familiar tunes 210–13
temporal processing 27–8, 168–9
temporal structure 285–9
temporal lobe *see* auditory cortex
tension and resolution 329–30
timbre 5, 18, 173
timing *see* rhythm
tonal cognition 95–108
acculturation 109–23
key-finding models 95, 102, 105–6
tonal hierarchies 97–9
tonal processing 231–46
tonal stability 112
tonality 95–108
tone-deafness *see* congenital amusia
training *see* musical training
transcranial magnetic stimulation 220, 225–8
transposition 5, 86
triads 128
tritone 6, 7, 132, 133, 135, 141

Webern, Anton, *Sechs St, cke* 82–3
Wernicke's aphasia 184
Western tonal music 113–15
Williams syndrome 68–9
Wishart, Trevor, *Vox 5* 81, 83

Xenakis, Iannis, *Pithoprakta* 80

Parncutt

THE SCIENCE & PSYCHOLOGY

OXFORD

# THE SCIENCE & PSYCHOLOGY OF MUSIC PERFORMANCE

## CREATIVE STRATEGIES FOR TEACHING AND LEARNING

*Edited by*
Richard Parncutt &
Gary E. McPherson

OXFORD
UNIVERSITY PRESS
2002

OXFORD
UNIVERSITY PRESS

Oxford New York
Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Singapore Taipei Tokyo Toronto

and an associated company in Berlin

Copyright © 2002 by Oxford University Press, Inc.

Published by Oxford University Press, Inc.
198 Madison Avenue, New York, New York 10016

www.oup.com

Oxford is a registered trademark of Oxford University Press.

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
electronic, mechanical, photocopying, recording, or otherwise,
without the prior permission of Oxford University Press.

Library of Congress Cataloging-in-Publication Data
The science and psychology of music performance : creative strategies for teaching and
learning / Richard Parncutt and Gary E. McPherson, eds.
p. cm.
Includes bibliographical references and index.
ISBN 0-19-513810-4
1. Music—Performance—Psychological aspects. 2. Music—Instruction and study.
I. Parncutt, Richard, 1957– II. McPherson, Gary E.
ML3838 .S385 2002
781.4—dc21 2001036292

5 7 9 8 6 4

Printed in the United States of America
on acid-free paper

# Acknowledgments

The first draft of each chapter was independently reviewed by four people: the two editors, another author from the book, and an anonymous external reviewer. Regarding the latter, we are grateful to: Margaret Barrett, Colin Durrant, Jane Ginsborg, Thomas W. Goolsby, Donald Hall, David Hargreaves, Daniel Kohut, Gunter Kreutz, Leopold Mathelitsch, Jeff Pressing, James Renwick, Bruno H. Repp, R. Keith Sawyer, Stanley Schleuter, Emery Schubert, and William Forde Thompson for making their expertise available to us and to the authors.

We take this opportunity to thank the various representatives of Oxford University Press with whom we came in contact during the long process from submission to publication. We especially thank our acquiring editor, Maribeth Payne, and our production editor, Ellen Guerci, for their support and guidance.

Most of all we, the editors, are grateful to the authors for their patience in the face of all those guidelines, bulletins, reviews, comments, recommendations, and deadlines that we sent around by e-mail. At times they must have seemed endless! Now that the authors can see their chapters in the context of the whole book, we hope that they agree with us that it was worth the effort.

*December 2001*

R. P.
G. E. McP.

# Contents

Introduction ix

I. THE DEVELOPING MUSICIAN

1. Musical Potential 3
*Anthony E. Kemp & Janet Mills*

2. Environmental Influences 17
*Heiner Gembris & Jane W. Davidson*

3. Motivation 31
*Susan A. O'Neill & Gary E. McPherson*

4. Performance Anxiety 47
*Glenn D. Wilson & David Roland*

5. Brain Mechanisms 63
*Eckart Altenmüller & Wilfried Gruhn*

6. Music Medicine 83
*Alice G. Brandfonbrener & James M. Kjelland*

II. SUBSKILLS OF MUSIC PERFORMANCE

7. From Sound to Sign 99
*Gary E. McPherson & Alf Gabrielsson*

8. Improvisation 117
*Barry J. Kenny & Martin Gellrich*

9. Sight-Reading 135
*Andreas C. Lehmann & Victoria McArthur*

10. Practice 151
*Nancy H. Barry & Susan Hallam*

11. Memory 167
*Rita Aiello & Aaron Williamon*

12. Intonation 183
*Steven J. Morrison & Janina Fyk*

13. Structural Communication 199
*Anders Friberg & Giovanni Umberto Battel*

14. Emotional Communication 219
*Patrik N. Juslin & Roland S. Persson*

15. Body Movement 237
*Jane W. Davidson & Jorge Salgado Correia*

III. INSTRUMENTS AND ENSEMBLES

16. Solo Voice 253
*Graham F. Welch & Johan Sundberg*

17. Choir 269
*Sten Ternström & Duane Richard Karna*

18. Piano 285
*Richard Parncutt & Malcolm Troup*

19. String Instruments 303
*Knut Guettler & Susan Hallam*

20. Wind Instruments 319
*Leonardo Fuks & Heinz Fadle*

21. Rehearsing and Conducting 335
*Harry E. Price & James L. Byo*

Contributors 353

Author Index 363

Subject Index 373

# Introduction

How fluently do music psychologists, music educators, and practicing musicians communicate with one another? Circumstances do not always favor an easy interaction. Each group has its own language, and the specialist ways of communicating that exist within each group may not always work across group boundaries. Moreover, cross-disciplinary interactions are not always explicitly encouraged by the institutions within which music researchers and practitioners work. These everyday limitations help to explain why music educators and practicing musicians have not benefited as much as they could from the past few decades of music psychology research and why music psychologists often neglect to cite, and so benefit from, relevant studies published in the mainstream research journals of music education.

In this book, we attempt to bridge the interdisciplinary gaps that currently separate music psychologists, music educators, and practicing musicians by developing new approaches to teaching, learning, and making music that are informed and inspired by the results of recent research in music psychology, music education, and acoustics—whether that research is published in music psychology or in music education journals. In this way, we aim to produce something fresh—a book that is without precedent in either music psychology or music education.

To help achieve this aim, each chapter is coauthored by two internationally recognized scholars: one a scientist (psychologist, acoustician, physiologist, or physician) and the other a performer or music educator. Authors who are well versed in *both* music education and music psychology are coupled with collaborators with relevant complementary expertise in the area of the chapter. These collaborations, many of which are entirely new (some coauthors have not, at the time of writing, met face to face!), have at times been inspirational and at other times frustrating. The artistic authors wondered why the scientific authors sometimes indulged in virtuosic scientific methodology and terminology with little or no meaning for practicing musicians; referred so often to the literature as to give an impression of insecurity; made simplistic generalizations in areas that obviously involved complex cognitive, artistic, and judgmental processes; or

stated obvious musical facts as if they had just been discovered. For their part, the scientists wondered why the artists sometimes felt they could develop such complex, impressive-sounding arguments on the basis only of anecdote, intuition, and experience or why they were so ready to rely on their own judgment without taking into account the judgments of their peers. We as editors made similar errors and were not seldom corrected by the authors (for which we thank them).

The book addresses all educational levels, from elementary through to universities and conservatories—although some chapters, depending on their subject matter, are by necessity oriented toward a specific level. The book also strives to be international in its approach and scope by focusing on international commonalities in music education and research and avoiding extended treatment of materials or practices that are specific to a given country. The two authors of each chapter often live and work in different countries; in cases where they do not or where the chapter seemed biased toward a particular country, we had the chapter anonymously reviewed by an expert from another country.

Each chapter begins by surveying relevant psychological and scientific research that may be unfamiliar to many teachers and performers. After that, the chapter authors creatively explore the implications of this research for music performance and education. Of course, considerable caution, creativity, experience, and intuition are required when developing teaching and performance applications on the basis of scientific research results and theories. In some cases the transition is easy, while in others contextual complexities that were neglected by the scientific researchers need to be taken into account and their theories modified accordingly. Likewise, the processes and practices common to music teaching and learning are not always based on developed theoretical models that have been modified and refined as a result of systematic scientific enquiry. One of our jobs as editors was to try to resolve issues of these kinds as they came up in their own special way in each chapter.

The various chapters are organized into three main parts. The first covers the personal and environmental influences that shape the learning and performance of music during a musician's life span, the second surveys the essential subskills of musical performance, and the third looks at specific instruments, techniques, and ensembles. The first and second parts draw mainly on psychologically oriented literature; the third, due to its different subject matter, on a mixture of psychology and acoustics. The music education literature is referred to throughout the book.

The book is broad in its scope but is not intended to be comprehensive. When we first started discussions, we soon realized that it would be impossible to cover all important aspects of music performance in one volume. We especially became aware of deficiencies in various parts of the literature and of the need to focus our efforts on the performance of Western tonal "art music"—mainly because the bulk of research currently available focuses on this genre of performance and also because it currently dictates how music performance is currently taught in most formal institutions around the world. Even within these confines, we have neglected important issues such as gender (Green, 1997; Maidlow & Bruce,

1999; O'Neill, 1997), motor control and technique (Gellrich, 1998; Rosenbaum & Collyer, 1998; Schmidt & Lee, 1999), music technology (Berz & Bowman, 1994; Webster, 2002; Williams & Webster, 1999), and percussion instruments (Fletcher & Rossing, 1998); and within each chapter, the authors have had to select what they perceive to be the most important scientific and artistic material that corresponds to their topic. This book cannot even be regarded as a general exposition of *scientific* aspects of music performance—although of course many such aspects are covered. Instead, it is primarily about *applying* scientific findings. In this sense it is more exploratory than encyclopedic.

The book is intended for a range of different groups of readers. We are writing first for music educators with a specific interest or expertise in music psychology. The second target group comprises practicing musicians who are interested in the practical implications of scientific and psychological research for music performance. The third group consists of advanced undergraduate and postgraduate students of music education and music psychology—especially those whose instructors recognize the potential of music psychology to inform music education and vice versa. Fourth, we are writing for researchers in the area of music performance who consider it important for the results of their research to be practically useful for music educators and musicians. Almost every chapter poses questions, either directly or between the lines, that have not yet been adequately answered in the empirical literature and could be the topic of future research projects. In that sense, music psychologists may benefit at least as much as music educators from the book.

As in any interdisciplinary effort, problems of terminology arise (cf. Clarke, 1989). We have aimed for a linguistic style that can be understood across the various subdisciplines of music psychology, music education, and music performance and have tried to avoid terms that are not understood in all the main disciplines of each chapter. In some cases, we felt that one discipline could benefit from adopting the terminology of another or from adopting a more clearly defined terminology than that currently in general use. In any case, we have tried to be consistent. For example, we consistently use the word *note* to refer to dots on a page and *tone* to refer to the physical sound that is produced. Following standard practice in auditory psychophysics (psychoacoustics, music perception), we distinguish physical measurements of musical tones (frequency, SPL, spectrum) from corresponding perceptual attributes (pitch, loudness, or timbre). We have adopted the metric system and its standard abbreviations, including, for example, *ms* for milliseconds and *Hz* for hertz (cycles per second). When referring to octave register, we follow the American Standards Association (1960): $C_4$ = middle C (261 Hz), $A_4$ = 440 Hz, and so on. We follow U.S. usage for note values (quarter notes rather than crotchets, etc.). Interval sizes are expressed in either equally tempered semitones or cents (hundredths of a semitone).

### *References*

American Standards Association. (1960). *USA standard acoustical terminology.* New York: American Standards Association.

Berz, W. L., & Bowman, J. (1994). *Applications of research in music technology.* Reston, VA: Music Educators National Conference.

Clarke, E. F. (1989). Mind the gap: Formal structures and psychological processes in music. *Contemporary Music Review, 3*, 1–15.

Fletcher, N. H., & Rossing, T. D. (1998). *The physics of musical instruments* (2nd ed.) New York: Springer.

Gellrich, M. (1998). Über den Aufbau stabil-flexibler Fertigkeiten beim Instrumentalspiel. In G. Mantel (Ed.), *Ungenutzte Potentiale—Wege zu konstrukitvem Üben* (pp. 131–151). Mainz: Schott.

Green, L. (1997) *Music, gender, education.* Cambridge: Cambridge University Press.

Maidlow, S., & Bruce, R. (1999). The role of psychology research in understanding the sex/gender paradox in music–plus ça change. *Psychology of Music, 27*, 147–158.

O'Neill, S. A. (1997). Sex and gender. In D. J. Hargreaves & A. C. North (Eds.), *The social psychology of music* (pp. 46–63). Oxford: Oxford University Press.

Rosenbaum, D. A., & Collyer, C. E. (Eds.) (1998). *Timing of behavior: Neural, computational, and psychological perspectives.* Cambridge, MA: MIT Press.

Schmidt, R. A., & Lee, T. D. (1999). *Motor control and learning: A behavioral emphasis* (3rd ed.). Champaign, IL: Human Kinetics.

Webster, P. (2002). Computer-based technology and music teaching and learning. In R. Colwell and C. Richardson (Eds.), *New handbook of research on music teaching and learning* (pp. 416–439). New York: Oxford University Press.

Williams. D., & Webster, P. (1999). *Experiencing music technology* (2nd ed.). New York: Schirmer.

# PART I

## THE DEVELOPING MUSICIAN

# 1

# Musical Potential

ANTHONY E. KEMP & JANET MILLS

This chapter discusses the identification of musical potential in childhood and considers the impact of an environment that is nurturing and stimulating. It addresses the place of musical ability tests as well as personality assessment and suggests that while these are areas for the teacher's consideration, they may prove to be of minor significance in any formal selection processes. The chapter describes the manner in which parents and teachers might consider musical potential, not as something finite but as something that may gradually emerge during childhood. The question of children's potential to play various instruments is considered as well as the existence of certain stereotypes prevalent in musical circles. Suggestions are offered to parents and teachers so they may become aware of a child's potential as soon as it becomes observable and nurture it in a facilitating environment.

"This child is musical, you know." How often have we heard an adult—a proud parent, an elderly relation, a friend of the family—say something like that! What do they mean? And are they right?

The chances are that the adult is trying to say something positive about the child, something that marks him or her out as more special than other children in a manner that is wholesome or at least harmless. Being a *musical* child is, in the eyes of an interested adult, normally good. The adult thinks that something has been spotted in the child that destines him or her for greatness as some kind of musician or at least indicates that it is worth giving the child opportunities that may be denied to other children, for example, lessons on a musical instrument. In other words, the adult is suggesting that the child has what we shall describe in this book as *musical potential*: a latent, but as yet unrealized, capacity to do something musical—for example, play the flute.

But is the adult right? Is it possible to spot that one child has more potential than another to play flute or piano or trombone before either has been given a

chance to try doing any of these things? The biographies of musicians sometimes give an impression that verges on that of a baby emerging from the womb with the musical skills of a very competent adult. But clearly that cannot be the case.

## Research into the Identification of Musical Potential

Some children who seem much like other children nevertheless turn out to be gifted musicians. No child is a blank sheet with respect to music, but the experiences that children have of music vary widely with respect to quantity and quality. The signs of musical potential that adults think they spot are manifestations of *musical achievement.* These are the results of musical experience and learning, formal or informal, that some children have had, but others lack. And there has to be someone there to spot these signs.

A child may move in time to music, play large numbers of pieces of Bach from memory, sing along with a parent, or become engrossed when an older sibling plays the viola. However, a child cannot sing along with a parent if the parent does not sing. Neither can a child play Bach on the piano if there is no access to a piano or the child has never heard any Bach. Nor can a child be seen moving in time to music if no one has time to spot this happening. It is likely that virtually all children have the potential to do all manner of things musical that they never have the chance to do.

Musical potential is something that all children have, although arguably some may have more of it than others, and musical potential may come in different shapes and forms. Musical behavior such as joining in singing means that a child has responded favorably to an opportunity to learn in music but does not necessarily mean that a child has more potential for music than any other child. Neither does it necessarily mean that the child will show more *aptitude* if offered the chance to learn an instrument, that is, that he or she will realize the potential to develop the *ability* to play it effectively more speedily than any other child.

This chapter encourages the reader to take a broad view concerning the identification of musical potential. Howe (1990a) argued that high achievement in any field is fuelled by nonintellectual qualities. He cites Shuter-Dyson's (1985) account of how, in selecting pupils for places in a specialist music school, a headmaster placed principal emphasis on the pupil's insatiable appetite for music that was so strong that without it he or she would have felt deprived. Howe (1990b) went on to identify these personal qualities of the high achiever:

> What are these qualities? They seem to be ones that are only weakly governed by abilities as such. They appear to depend more on motivation and on temperament and personality. Gaining an all-important sense of direction or purpose seems to involve the combination of particular abilities and these other attributes. (p. 68)

Despite the more psychometric focus of his work, Seashore (1938) made much the same claim. He maintained that in considering the "musical mind" we should view it as "a total personality . . . functioning in a total situation" and empha-

sized the possession of essential capacities in the hearing, feeling, understanding, and expressing of music that result in a "drive or urge toward music" (pp. 1–2).

Like Howe, Cattell (1971) also maintained that the close analysis of any ability is impossible without an understanding of three "modalities": *ability*, *personality*, and *motivation*. While he claimed that these were conceptually distinct, nevertheless they were intertwined in reality. Each of these areas will be discussed later in this chapter.

### *Musical Ability Tests*

Some research into the musical abilities of very young children provides insights into the extent to which many people may ultimately underachieve in music. Bridger (1961) noticed that some babies less than five days old moved when the pitch of a note that had been playing for a long time was changed by only a major third. It is, of course, possible that they spotted changes of pitch even smaller than a major third—they just did not move. Chang and Trehub (1977b) found that babies aged five months noticed simple rhythmic changes, for example, ** **** instead of **** **. They also found that the babies appeared to have grasped the concept of melodic shape. When a six-note melody to which they had been listening repeatedly was transposed to another key, their heart rate remained constant: when the order of the six notes was changed, their heart rate decreased (Chang & Trehub, 1977a).

Over the last century there have been a number of attempts to find a scientific way of determining children's musical potential that saves the expense of giving them some lessons to see how they get on. In 1919 Carl Seashore published the first version of his Measures of Musical Talents. These were based on the psychometric techniques of the time and tried to establish the limits of a person's hearing, for example, the highest note that he or she could perceive or the smallest pitch interval that could be recognized. The theory behind these tests was that better hearing implied higher levels of musical talent.

The limitations of such tests were highlighted by Seashore (1938) himself in a case of a professional musician and his apparently nonmusical brother. It transpired that the brother performed significantly better than the musician on the battery of tests of musical capacity, causing Seashore to ponder whether the brother was musical or not. The fact that the brother had not shown any interest in pursuing music rather suggests that he possessed all the necessary auditory skills but that he lacked the motivation and temperament to mobilize these in any musical activity.

Nevertheless, subsequent researchers have continued to devise test batteries. Over the last 35 years Edwin Gordon (1965, 1971, 1982) has developed tests that are used mainly in the United States, and range from Audie, a game for children aged 3–4 years, through the Music Aptitude Profile, which is intended for children aged 11–18 years, to the Advanced Measures of Music Audiation, which are designed to be used as college entrance tests. Arnold Bentley developed his Measures of Musical Abilities (1966) that were issued on LP records and used

for a number of years in many schools in the United Kingdom. Many teachers find it very helpful to have access to test results that they can use when deciding which children might be offered special provision, for example, instrumental lessons. However, these tests still make demands that children can meet more easily if they have had prior musical experience. In addition, Bentley's tests were too demanding of the recording technology that was available at the time. For example, in two of his questions one is intended to notice that two tones are of identical pitch, rather than going *down* or *up*. However, on listening to the recording it is clear to many teachers or pupils with a good ear that the second tone is slightly lower than the first, a fact that was subsequently verified using later computer technology (Mills, 1983). Yet children who wrote *down* instead of *same* for these questions lost two marks.

Another danger with tests of this type was that teachers sometimes used them outside the context for which they were intended. For example, the manual that accompanied Mills's *Group Tests of Musical Abilities* (1988) stated clearly that their main purpose was that of providing additional positive information relating to a child about whom a teacher knew very little. They were never to be used to say no to a child who wanted to do something musical. However, almost every question that was received about them once they were published related to their being used to do precisely that.

### *Biographical Data*

Research that has pursued young musicians' biographical details encourages us to consider the importance of environmental factors as being far more significant than inborn talent. Bloom (1985), Sosniak (1990), and Sloboda and Howe (1991) suggest that the most important factor in developing a child's musical talent is providing a stimulating environment from infancy onward, encouraging his or her first musical responses if and when they occur. Coupled with this is the necessity for appropriate musical opportunities and support from parents and teachers who are caring and encouraging (see chapter 2).

Research that attempted to identify early manifestations of "natural" talent by interviewing the parents of talented young musicians failed to provide much in the way of specific indicators. Howe et al. (1995) questioned parents about musical behaviors in the early years of talented musicians in comparison to a less talented group and found that the only behavior that emerged was singing from an early age. These children may have been fortunate enough to grow up in an environment in which singing was tolerated or possibly encouraged. Other factors such as moving to music, showing a liking for music, being attentive to music, and requesting musical involvement did not occur at higher levels in the more talented.

Sloboda, Davidson, and Howe (1994) carried out research with a group of adult musicians who were asked about their early musical experiences. It emerged that many recalled "deeply felt and intensely positive early experiences to the 'internal' aspect of musical events, which seemed to lift them outside the normal state of awareness" (p. 353). We shall return to this again later in a discussion

that concerns children's musical motivation. Here we should note that such powerful experiences will tend to occur in those homes that provide an appropriate kind of stimulating environment in the form of either recorded music or music performed live by parents and older siblings.

### *Personality*

Enough has been said earlier to suggest that the identification of musical potential may involve a much broader profile of personal factors than merely auditory skills and musical precocity. This raises the question of whether certain personal attributes might somehow sustain musical development and motivation, and certainly personality studies of musicians tend to support this notion. However, it must not be overlooked that some aspects of personality may be the result of musical engagement. Research has not, as yet, disentangled the causal direction, but it is generally accepted that both effects may be operative (see Kemp, 1996).

Research into the personalities of talented musicians suggests that they have a distinct pattern all their own and that this pattern, if not fully in place in young musicians, is well on the way to being developed. Kemp (1996, 2000) has shown, for example, that young musicians as an overall group display significant levels of introversion that remain in place through childhood and into adulthood. However, it must be emphasized that the musician's form of introversion takes a slightly different nature from that normally found in general populations, simply because it does not involve the usual element of shyness. The most prevalent aspect of the young musician's introversion is a significant level of self-sufficiency coupled with an element of aloofness or detachment particularly located in the very talented.

One might ask what purpose these traits might have in the development of musical potential. An answer to this question relates to the demands that music increasingly makes upon the young child in terms of the amount of time spent practicing in seclusion. More extroverted children will find periods of separation from their friends more difficult to deal with and may well drop out of instrumental tuition during the early stages, preferring activities that are more social and group-oriented. Clearly, some instruments make more demands in terms of practice hours than others before an acceptable sound is produced. One might, therefore, expect an interrelationship to occur between instrument and level of introversion (see Kemp, 1996). We will return to this later in a discussion that concerns instrumental choice.

In passing, another point needs to be made that concerns the nature of introversion. The introverted child is one who tends to engage in considerable private and imaginative activity, developing an internal world of ideas and symbolic thought (Kemp, 1996). This was certainly Jung's (1923) view: introverts tend to direct their energies inward toward a more personal and subjective response to what goes on around them. This can take the form of painting, writing stories, making up rhymes, or improvising. Such a child may respond naturally to musical activity, engaging readily in all its emotional nuances. These tenden-

cies toward creative thinking and activity blend well with another very significant aspect of the musician's personality, which involves a special form of sensitivity. The musician's sensitivity tends to incorporate a level of feeling and intuition that the musician needs in the performance of music if it is to become personal and cherished. The point must not be overlooked, however, that the more a child engages in activities of this kind, the more likely it is that levels of introversion and sensitivity will be developed. This notion is borne out by the buildup of these traits in adult musicians (Kemp, 1996).

One additional feature of the adult musician's temperament should be mentioned at this point. Although most young musicians do not display the anxiety found in music undergraduates and professionals, it has been located in the very talented who attend special music schools (Kemp, 1996). Much has been written about the performance anxiety suffered by adult musicians, and one is led to consider that this is acquired during the maturational process. Musicians find themselves increasingly investing much of themselves in their music, very often finding it difficult to separate their personal identity from that of being a successful musician. When self-esteem is so closely intertwined with the musician's persona, any kind of setback, be it a substandard performance rating, criticism, or examination failure, is perceived as a direct attack on the person him- or herself.

## Identifying and Developing Musical Potential

The issues briefly discussed earlier raise several questions about the upbringing of young musicians that need considering by both parents and teachers. It has to be said at this point that there can sometimes be conflict between parents and teachers about what is in the very best interests of the child. This is especially true when parents and teachers invest too much of themselves in the vision that they have for the child's musical future.

### *Environment Versus Innate Talent Controversy*

The point was made earlier that certain influential researchers have stressed the environmental factors that they maintain are of paramount importance in the realization of a child's musical potential. However, the question of the existence of innate talent also needs to be addressed. In reviewing the literature, during which they adopt a pro-environment position, Howe, Davidson, and Sloboda (1998) maintained that there was little evidence for the early emergence of musical talent. What evidence that existed was in the main anecdotal and retrospective and thus, in their view, inadequate. In summing up they did not claim to have full or precise answers to the question and conceded that individual differences in some special abilities might have genetic origins.

The peer commentary concluded the paper of Howe et al. demonstrated the degree of controversy that surrounds the issue, particularly on the part of those who wish to take up polarized position. In his criticism of these researchers Csikszentmihalyi (1998) maintained that it is doubtful whether talent can be

explained exclusively by either position. Cattell (1973) would have supported this view from his perspective of personality development. Those personality characteristics that, it is claimed, in part influence musical development, are themselves inherited to some degree. Cattell (1973) maintained that, roughly speaking, personality development is fairly equally influenced by genetic and environmental factors. In other words, a child who is perceived as having inherited a parent's introversion may be more predisposed to be comfortable in a potentially solitary pursuit such as practicing a difficult instrument. However, the more extroverted child may find the more social setting that singing provides more suitable.

That point, having been made, does not in any way negate the onus placed on parents and teachers to provide the right kind of environment for a child's interest in music to develop. This, as suggested earlier, is dependent upon parents being vigilant in noticing the child's earliest musical responses. Engaging with these at a playful level will ensure that the child sees that he or she is valued.

The playful element, so often overlooked by parents and by many teachers of young children, was stressed in research carried out by Sosniak (1990) that describes in some detail the early stages of learning an instrument. Initially, lessons should be fun and engaged in with a teacher who is both warm and enthusiastic. The teacher should provide safe and contained environment in which musical play takes place. Sosniak (1990) describes the process in this way:

> For relatively little effort, the learner got more than might be expected. The effect of the early years of playful, almost romantic involvement . . . seemed . . . to get the learner involved, captivated, "hooked"—motivated to pursue the matter further. (p. 155)

The question that arises, of course, is whether there are a sufficient number of teachers who are prepared to teach beginners in this fashion and also whether those parents who are ambitious for their child's musical progress are prepared to pay the fees for such a permissive kind of teaching!

However, once "hooked" in the way described by Sosniak, the child will proceed naturally and gently to the next stage, in which the teaching becomes more focused and systematic. Although playfulness is not totally left behind, objectives become more narrowly defined and there is a gradual expectation that appropriate skills and knowledge will be acquired.

### *Developing Motivation*

The process described earlier assumes that the first manifestations of a child's musical responses are internally motivated—they are, one might suggest, manifestations of the child's first musical needs. In small infants these can develop into a repertoire of enjoyable activities in which carers can engage and of which the child will learn to request more. These behavioral indications of the child's motivation toward music belong to the child and must not be commandeered by the parent. Any effort on the part of a parent to pressurize development will tend to destroy the child's sense of motivation.

In a sense, this should remain true throughout all music learning. The child's intrinsic motivation for music can so easily be swamped by a parent's and, later, a teacher's driving efforts toward the achievement of further skills and knowledge. In other words, inappropriate forms of extrinsic motivation may have the effect of reducing the child's sense of commitment and internal drive (see chapter 3).

Let us briefly return to the research of Sloboda et al. (1994) in which they suggested that many musicians have had childhood experiences at very deep levels in which the sound of an instrument or a particular piece of music exerted a powerful and lasting response. Jacqueline du Pré experienced such an event before her fifth birthday that she later recalled with clarity:

> I remember being in the kitchen at home, looking up at the old-fashioned wireless. I climbed onto the ironing board, switched it on, and heard the introduction to the instruments of the orchestra. . . . It didn't make much of an impression on me until they got to the cello, and then . . . I fell in love with it straightaway. Something within the instrument spoke to me, and it's been my friend ever since. (In Easton, 1989, p. 26)

This, albeit anecdotal, evidence, represents a not uncommon experience reported by adult musicians. Walters and Gardner (1992) have described these as *crystallizing experiences* and attempt to explain them as overt reactions to a quality or feature of an occurrence that yields an immediate change in individuals' concept formation, their performance within it, and self-concept. Such powerfully charged experiences appear to have a lasting impact upon the individual, leaving him or her in a highly motivated state that may last a lifetime. As in the case of Jacqueline du Pré, this may take the form of an insatiable attraction to an instrument's qualities. Clearly, the environment that is rich with all manner of musical experiences is more likely to stimulate such responses than one that is bereft of cultural stimuli. Once an experience of this kind has occurred, parents and teachers need to recognize, harness, and develop this form of personal commitment and take care not to swamp it with invasively close monitoring and unnecessary forms of extrinsic motivation.

### *Choosing an Instrument*

Common sense tells us that the best way of seeing whether someone has potential to play an instrument or sing is to offer tuition and see how the young person progresses. Common sense tells us also that if he or she does not progress this may be the fault of the teacher rather than the student or, simply, the instrument is not the most suitable one. Few pupils could make a success of absolutely any instrument, and so it makes sense to give serious consideration to the question of what instrument a child might learn. Gordon (1991) claims that when his instrument Timbre Preference Test is used with children, they have a 75% better chance of continuing on the instrument of their choice. However, the basis on which children are steered toward a particular instrument is frequently far less scientific, and not informed by research.

In cases like that cited in the previous section, where the pupil arrives with a clear motivation to pursue a particular instrument or sing, little consultation may be necessary. If the drive is very strong toward a particular instrument, parents and teachers may be advised not to attempt to redirect the child toward an alternative, even in a case where the choice appears to be inappropriate. For other children certain general aspects of their personality and abilities might be kept in mind (see Kemp, 1981, 1996) but should not, in any circumstances, be allowed to dictate the choice of instrument. For example, string playing may frequently attract the quieter, more introverted and studious child, whereas brass playing may appeal to the more socially outgoing and extroverted. Singers also tend to be more extroverted, sensitive, and imaginative types. Keyboard players tend to be more extroverted but not as extroverted as the brass players and singers. Clearly, the kinds of demands that instruments make on the learner, particularly in terms of the relative ease with which reasonably pleasing sounds can be produced in the early stages, will help determine whether the beginner will persevere or, in time, discontinue instruction.

Quite apart from personality, another aspect that we might consider is a young person's physique. As long ago as 1935, Charles Lamp and Noel Keys, a teacher and a university lecturer both working in San Francisco, wrote of the selection methods of instrumental teachers as follows: "In the absence of scientifically validated conclusions, music instructors have been forced to resort to *a priori* reasoning and uncontrolled observation" (p. 587). They went on to show that some of the popular theories of the day about the physical characteristics that suited particular instruments were of very little value. These included theories that are still heard today, for example, that long, slender fingers are needed for string instruments, even teeth are required for woodwind and brass instruments, and thin lips are better suited to brass instruments with small mouthpieces and vice versa. The statistical correlations that Lamp and Keys found between the supposedly favorable characteristics and the progress that pupils made on instruments were positive but small. In other words, pupils with thin lips were slightly more likely to make a success of the French horn, but there were plenty of examples of pupils with thick lips who succeeded at the French horn and pupils with thin lips who did not. The length and slenderness of pupils' fingers predicted their success on the violin only very slightly more effectively than the evenness of their teeth!

Fifty years later, Mills (1983) obtained some very similar findings. Seventy-four instrumental teachers completed checklists of the physical characteristics of a total of 299 instrumental pupils aged 7–18, and these were compared with ratings of their instrumental performance. Again, the correlations were positive but small. There were many high-achieving pupils with supposedly undesirable physical attributes, for example, buck-teethed bassoonists, clumsy percussionists, and thin-lipped tuba players. Some of these pupils have gone on to make careers as instrumentalists.

Mills (1995) identified a physical characteristic that predicted instrumental achievement better than any other criterion. This related to the comparative length of the ring and index fingers of the left hand. She found that violinists

and viola players frequently possessed left hands where the ring finger was longer than the index finger (in common with 70% of the general population). However, other instrumentalists generally possessed the less common left hand shape.

So does this mean that pupils should be screened for a digital formula and steered toward an instrument accordingly? Not at all; a few pupils with the less desirable digital formula do succeed, and others could as well. What it means is that teachers may need to be more alert to the need to adapt their teaching of technique to suit the shape of the pupils, perhaps particularly if a pupil has a different shape from his or her teacher. Wagner (1988) has devised a "biomechanical hand profile" that can help piano teachers with this, because the hand shape of an adult student may be compared with that of a professional reference group and also that of his or her teacher.

However, it is stressed here that there are many exceptions to the generalized findings described earlier that raise serious questions about their predictive power. The music profession abounds with forms of folklore that stereotype different kinds of instrumentalists, and these have found their way into general parlance. For example, research has shown that some instruments like the violin and flute are perceived as being feminine in character, while others such as the brass are gendered as masculine (see, for example, Abeles & Porter, 1978). These perceptions concerning instruments may well result in playground ridicule, leading to high dropout rates for some instruments, particularly among boys (O'Neill, 1997). However, none of the preceding issues should be adopted by parents and teachers in selecting or rejecting children for particular instruments. The only exception to this might be where a child is not strong enough to hold an instrument or where hand span will not allow a comfortable position on an instrument's keys. Clearly, the maturational process will remedy this in time.

### *Teaching Methods and Pupils' Individual Differences*

The research carried out by Sosniak (1990) and Manturzewska (1990) clearly suggests modes of approach for young beginners, who, it has been seen, require a warm, nurturing, and playful climate in the early months. From then on, the sensitive and imaginative teacher will modify his or her teaching according to the individuality of the child. Pupils who do not receive this kind of teaching and who instead are met with formal approaches and teaching methods may not respond and may wish to discontinue lessons after a short period. These kinds of teachers may well consider the child to be unmusical and feel justified in rejecting him or her without realizing that the fault lies in the approach.

The research that suggests that young musicians tend to be introverted and sensitive should not be overlooked. "Still waters run deep" is a saying that can often be applied to the child who keeps his or her musical enthusiasms hidden. Sometimes these can be ignored by the rigid teacher as lying outside the "requirements" of tuition. For example, a child with a particular enthusiasm such as improvising, jazz playing, or scat singing may be rejected as unsuitable for tuition. In other words, the process of identifying musical talent is a much more complex task than is often realized. Talent is often left unrecog-

nized because a teacher's perception of what constitutes talent is too narrowly or poorly defined.

Unlike classroom learning, individual tuition with a sensitive and responsive teacher—a kind of mother figure—can be a haven in which the introverted child may feel listened to and valued in a way not experienced anywhere else. The converse may also be true: the extroverted child may well respond to the kind of group teaching that prevails in the early stages, particularly in brass tuition and, of course, in choirs. Such individuals may also be encouraged to undertake some of their practice together.

Research also suggests that young musicians as an overall group tend to be dependent types, depending upon their parents and teachers in the early stages. However, this research also indicates that those individuals who make most headway and who are likely to emerge as professional musicians are very much more inclined to be independent (Kemp, 1995). There is a lesson here for parents and teachers who, while exercising an important role of encouragement in the early stages of musical development, need to be able to let go of talented pupils and allow them to assume more responsibility for personal decision-making. The biographies of many young musicians report personal problems that occurred in later adult life and were caused by parents and teachers who had overprotected them, projecting too much of their own ambitions onto young, talented musicians, in some cases living through their accomplishments in a vicarious fashion. Strong-minded children who have a clear idea about how they want to learn and what pieces they wish to play, while perhaps being less comfortable for the teacher to deal with, may well be the very ones who succeed in the long term. Roe (1967) once remarked that too much loving and too little neglect does not produce a creative child. While this may overstate the case, it might be said that a child's creativity requires a certain degree of personal space if it is to thrive.

Finally, the research referred to earlier that suggests very talented young musicians as well as more mature performers display anxiety needs to be addressed in passing. Whereas certain levels of anxiety may well be of significance for all performing artists in order to facilitate performance at optimal levels, its etiology is clearly more complex than that. Many professional musicians suffer levels of acute debilitating anxiety before and during performance, and one is forced to consider whether this is a learned response (see chapter 4). One possible explanation relates to instrumental teaching that makes undue demands upon the pupil in terms of the kind of graded performance examinations offered in many parts of the world. Parents wish to see their children making swift progress, and sometimes there can be a competitive element involved as parents draw their offspring into a grade race. Such parents will sometimes place the teacher under pressure to enter pupils for examinations prematurely, proceeding to the next immediately after the completion of the last. This clearly has a narrowing effect on the amount and variety of repertoire being learned. Perhaps more important, however, it may cause pupils to be entered for examinations before pieces have been adequately prepared or before the pupils have developed an acceptable margin of safety, thus reducing the examination to an anxiety-inducing event rather than what could be a pleasurable and aesthetic experience. Such regular encounters with performance

anxiety may result in what was initially a temporary form of anxiety (*state anxiety*) taking on the more permanent manifestation of anxiety (*trait anxiety*) of the kind found in more mature musicians.

## Conclusion

Based on a broad range of literature, this chapter has argued that the identification of musical potential in young children needs to occur in the child's natural surroundings as an ongoing process. It has been suggested that identifying musical potential should not be a unique occurrence in which ready-made musical aptitude tests are applied. Such a practice will sometimes overlook other important factors. This chapter has suggested that musical potential is a much more complex phenomenon that involves a number of factors besides abilities. While aural capacities are no doubt of some significance in musicianship, they are not of themselves a reliable indicator of musical potential.

This chapter has emphasized additional factors that may assist parents and teachers in developing a more comprehensive view of what constitutes musical potential. Among these a child's natural inner motivation is clearly of paramount importance. After all, there is little musical future for a child who possesses all the necessary aural skills yet lacks the personal interest in harnessing them in musical activity. A similar point can be made concerning personality. Certain personality traits appear to support musical involvement and, it is claimed, without these musical progress, particularly for some instruments, may be more problematic, although not impossible.

Motivation and personality are phenomena that are best assessed over time, and it is therefore emphasized here that the identification of potential is an ongoing process that will occur in the home from birth onward and during the early years of any more formal tuition. Because musical potential takes many forms and will occur at different levels, each individual child requires individual consideration. For this reason, a parent or teacher may occasionally find it helpful to use a published test to identify a child's relative strengths or weaknesses in some aspect of musicianship. However, it is the quality of the nurturing environment that is critical, and the onus for this is placed firmly on parents and teachers, by not pushing this process but allowing the child to manifest the form and direction that development might take. It is these people who need to provide not only a musically stimulating environment but also one in which the child's enthusiasms are noticed, listened to, and responded to with sensitivity and imagination.

### *References*

Abeles, H. F., & Porter, S. Y. (1978). The sex-stereotyping of musical instruments. *Journal of Research in Music Education, 26*(2), 65–75.

Bentley, A. (1966). *Measures of musical abilities.* London: Harrap.

Bloom, B. S. (Ed.) (1985). *Developing talent in young people.* New York: Ballantine.

Bridger, W. H. (1961). Sensory habituation and discrimination in the human neonate. *American Journal of Psychiatry, 117*, 991–996.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.

Cattell, R. B. (1973). *Personality and mood by questionnaire: A handbook of interpretive theory, psychometrics, and practical procedures*. San Francisco, CA: Jossey-Bass.

Chang, H. W., & Trehub, S. E. (1977a). Auditory processing of relational information by young infants. *Journal of Experimental Child Psychology, 24*, 324–331.

Chang, H. W., & Trehub, S. E. (1977b). Infants' perception of temporal grouping in auditory patterns. *Child Development, 48*, 1666–1670.

Csikszentmihalyi, M (1998). Fruitless polarities. *Behavioral and Brain Sciences, 21*, 411.

Easton, C. (1989). *Jacqueline du Pré: A biography*. London: Hodder & Stoughton.

Gordon. E. (1965). *Music aptitude profile*. Boston: Houghton Mifflin.

Gordon. E. (1971). *Iowa tests of music literacy*. Iowa City: Bureau of Educational Research and Service, University of Iowa.

Gordon. E. (1982). *Intermediate measures of music audiation*. Chicago: GIA.

Gordon. E. (1991). A study of the characteristics of the Instrument Timbre Preference Test. *Bulletin of the Council for Research in Music Education, 110*, 33–51.

Howe, M. J. A. (1990a). *The origins of exceptional abilities*. Oxford: Blackwell.

Howe, M. J. A. (1990b). *Sense and nonsense about hothouse children: A practical guide for parents and teachers*. Leicester: British Psychological Society.

Howe, M. J. A., Davidson, J. W., Moore, D. M., & Sloboda, J. A. (1995). Are there early childhood signs of musical ability? *Psychology of Music, 23*(2), 162–176.

Howe, M. J. A., Davidson, J. W., & Sloboda, J. A. (1998). Innate talents: Reality or myth? *Behavioral and Brain Sciences, 21*, 399–407.

Jung, C. G. (1923). *Psychological types* (H. G. Baynes, Trans.). London: Routledge & Kegan Paul.

Kemp, A. E. (1981). Personality differences between players of string, woodwind, brass and keyboard instruments, and singers. *Council for Research in Music Education Bulletin, 66–67*, 33–38.

Kemp, A. E. (1995). Aspects of upbringing as revealed in the personalities of musicians, *Quarterly Journal of Music Teaching and Learning, 5*(4), 34–41.

Kemp, A. E. (1996). *The musical temperament: Psychology and personality of musicians*. Oxford: Oxford University Press.

Kemp, A. E. (2000). The education of the professional musician: Its psychological demands and outcomes. *Musical Performance, 2*(3), 93–110.

Lamp, C. J., & Keys. N. (1935). Can aptitude for specific instruments be predicted? *American Journal of Educational Psychology, 26*, 587–596.

Manturzewska, M. (1990). A biographical study of the life-span development of professional musicians. *Psychology of Music, 18*(2), 112–139.

Mills. J. (1983). *Identifying potential orchestral musicians*. Unpublished doctoral thesis, University of Oxford.

Mills. J. (1988). *Group tests of musical abilities*. Slough: NFER-Nelson.

Mills. J. (1995). *Music in the primary school*. Cambridge: Cambridge University Press.

O'Neill, S. A. (1997). Sex and gender. In D. J. Hargreaves and A. C. North (Eds.), *The social psychology of music* (pp. 46–63). Oxford: Oxford University Press.

Roe, A. (1967). *Parent–child relations and creativity.* Paper prepared for conference on child-rearing practices for developing creativity. Macalester College: St. Paul, MN, November 2–4.

Seashore, C. E. (1919). *Manual of instructions and interpretations of measures of musical talent.* Chicago: Stoelting.

Seashore, C. E. (1938). *The psychology of music.* New York: McGraw-Hill.

Shuter-Dyson, R. (1985). Musical giftedness. In J. Freeman (Ed.), *The psychology of gifted children* (pp. 159–83). Chichester, UK: Wiley.

Sloboda, J. A., Davidson, J. W., & Howe, M. J. A. (1994). Is everyone musical? *Psychologist, 7*(8), 349–355.

Sloboda, J. A., & Howe, M. J. A. (1991). Biographical precursors of musical excellence: An interview study. *Psychology of Music, 19*(1), 3–21.

Sosniak, L. A. (1990). The tortoise, the hare, and the development of talent. In M. J. A. Howe (Ed.), *Encouraging the development of exceptional skills and talents* (pp. 149–64). Leicester: British Psychological Society.

Wagner. C. (1988). The pianist's hand: Anthropometry and biomechanics. *Ergonomics, 31,* 97–132.

Walters, J., & Gardner, H. (1992). The crystallizing experience: Discovering an intellectual gift. In R. S. Albert (Ed.), *Genius and eminence* (2nd ed.) (pp. 135–155). Oxford: Pergamon.

# 2

# Environmental Influences

HEINER GEMBRIS & JANE W. DAVIDSON

Environmental and genetic factors affect individual development from fetus to adult, both generally and in the case of music. We consider the difference between shared and nonshared environmental influences, and different types of interaction between the individual and environment. Parents, teachers, and peers strongly influence this development. Early nonverbal interactions between child and mother or caretaker, and parental support for music activities in childhood, seem to be of particular importance. These and other influences (e.g., exposure to music through the media) occur in the more general framework of the societal, historical, and generational context. Environmental conditions for musical development may be optimized by paying more attention to shared music experiences between child and parents (e.g., parent-baby singing), and exposing the child to a wide variety of music.

The origin of human abilities has been discussed since antiquity, and the possession and development of musical skills in Western culture has intrigued educationalists and psychologists for well over a century. Theories and beliefs about the relative influence of nature and nurture have dominated the discussion. When we assess the literature on the topic, it appears that dominant cultural ideologies have had a strong influence on the development and persistence of ideas about musical ability. For instance, it appears that the notion of genius that emerged in the eighteenth and nineteenth centuries is responsible for the fact that up to the present day in public opinion musical ability is often still considered to be a special gift that is relatively independent from environmental influences and learning processes. Biological evidence indicates that genetic factors influence general development in three broad ways: maturational staged development, physical capacity, and mental capacity (Bee, 1992; Plomin & DeFries, 1999). Clear examples of each can be found in musical contexts. For example, there is a gradual development of hand and eye dexterity as a child grows, so an

eight-year-old child is better able to coordinate bow and string in violin playing than a three-year-old. People with wide hand spans have a physical advantage to develop as pianists over those with small hand spans. Generally, some people can perform mental tasks such as problem solving more readily than others. They may be able, therefore, to identify musical patterns quickly and thus carry out aural discrimination tasks more rapidly.

Thus an innate component is clearly featured in musical ability. Most researchers (e.g., Gordon, 1986) agree that the development of musical abilities is based on the interaction between innate capacities and environment. Depending on environmental conditions, the relative importance of innate differences changes. It would seem that in domains where culture exerts a relatively homogeneous influence on development innate differences would be especially visible. If someone grows up in a society that attempts successfully to develop musical abilities of all its members comprehensively and effectively with the same type of training, emerging individual differences in musical abilities should be attributed to innate factors. Conversely, in nonhomogeneous surroundings (e.g., with greatly varying musical training) the relative importance of the environment for causing individual differences in musical abilities increases, and the relative contribution of innate predispositions declines (e.g., Bouchard et al., 1990).

The current chapter will focus upon environmental influences, since these are the ones that we understand more fully at the present time. In addition, as Davies (1994) has argued, in light of the psychological evidence to date, there is an accompanying moral incentive to try to optimize the environmental circumstances that may aid maturity, development, and learning processes in as many people as possible, and this is something important for teachers to explore.

## Nature and Nurture

### *Definitions*

The broadest range of environmental factors that influence musical development are the sum of all external conditions that bear on musical experiences and activities. These include: (1) sociocultural systems such as the music culture and technological culture, (2) institutions such as the home and school, and (3) groups such as classes and peers. In a more narrow sense, and borrowing from the psychology of personality (cf. Asendorpf, 1999, p. 249), the (musically) relevant environment can be construed as all the regularly recurring situations that impact on musical experience and activities. Those regularly recurring situations, to which the individual is exposed, constitute the current environmental conditions of the individual. Important characteristics of a person's environment are the frequency and duration with which the person is exposed to a particular situation (*situational exposition*), for example, the frequency and duration with which he or she listens to music, watches video clips, or sings in a choir. The environment and its characteristics provide the framework in which musical

socialization takes place. No unified definition exists for the term *socialization* in social psychology or in sociology. However, in this context we assume that musical socialization is the learning process through which the individual grows into a musical culture, developing and adjusting his or her musical abilities, activities, ways of experiencing, and values in interrelation with the social, cultural, and material environment (cf. Gembris, 1998, p. 190).

### *Shared and Nonshared Environmental Influences*

Based on twin studies and adoption studies, population genetics has demonstrated that environmental influences contribute primarily to the *dissimilarity* of children growing up in the same family, while genetic influences contribute to the similarity (cf. Plomin et al., 1997). The influence of environmental influences consists of two different parts. On the one hand, there are such environmental influences as social strata, education, and parents' musical interests, which are shared by all siblings of the same family (shared environmental influences). On the other hand, there are environmental influences that are not the same for all children of one family, for example, the position of siblings in the family, the varying behavior of parents and teachers to the siblings, different friendship groups, and divergent musical experiences (nonshared environmental influences). So we can say that the environmental influences on siblings do vary not only between different families but also within the same family. There is clear evidence that, with the exception of the IQ in childhood, the nonshared environmental influences are more important for the development of personality characteristics (IQ, extroversion, neuroticism, etc.) than the shared environmental influences (see Reiss et al., 1994; Plomin et al., 1997).

It seems possible that these findings also apply to musical development. Casual observations show that the musical development of siblings may take distinctly different ways despite the children's being exposed seemingly to the same family situations. Take, for instance, the fact that the siblings of the famous cellist Jacqueline du Pré grew up in the same home and also received musical training, but in contrast to their renowned sister they did not attain international careers. Such differences need not necessarily be attributed to the effects of a different genome but may also be explained by the effects of the nonshared environment. The differentiation of shared and nonshared environmental influences is a new and promising way to explain differences in musical interests, motivation, and development.

### *Interaction of Environment and Individual*

The impact of the environment on the individual and his or her development is by no means unidirectional in the sense that the individual passively endures the influence of the environment. Instead, the individual interacts with his or her surroundings in three different ways (Plomin, DeFries & Loehlin, 1977). There is first a so-called passive covariation between an individual's genetic material and environment via the shared genes of parents and children, because parents

and other relatives create a certain way of life and family environment (including professional status and education, interests, musical activities), which in turn provides influential circumstances for the child (e.g., presence of musical instruments and parental music listening and performance). The interaction of this environmental configuration with the child's innate potential causes musical children to grow up in a stimulating atmosphere simply due to their parents' own musical potential. A second, so-called reactive or evocative nature–nurture covariation arises from the correspondence between the social environment and the genetically influenced personality or temperament and needs of the child. A child who is eager to learn will thus evoke more learning opportunities from his or her environment than a less eager one, and a musically interested child will provoke more musical offers than a musically less interested one. For instance, if parents realize that their child is interested in music, they may buy instruments for the child or provide opportunities to hear music. The third, so-called active nature–nurture covariation exists if the child actively selects offers from his or her environment or tries to shape the environment consistent to his or her genome (e.g., the child may ask for music lessons or choose for friends peers who are also interested in music).

The relative importance of those three genotype-environments does not remain the same throughout life but modulates. The passive nature–nurture relationship is relatively most important during childhood, and then the importance of this type of nature–nurture relationship diminishes. At the same time, the evocative and active nature–nurture relationship increases with chronological age, because children and adults evocatively and actively take charge in shaping their surroundings (cf. Plomin et al., 1997).

## Home Environment and the Family

### *Beginnings of Environmental Influences*

The environment starts to influence the individual before birth. Recent studies demonstrate that the conditions under which the fetus develops in the womb can already affect a fetal programming that may have lasting impact on the health and susceptibility to illnesses of the adult individual. A major factor in this is the stress hormones, growth hormones, and nutrients that are absorbed by the child through the placenta (Seckl, 1998). Infants are able to hear music and speech several weeks prior to their birth and can recognize these after birth (e.g., Satt, 1984; DeCaspar & Spence, 1991; Hepper, 1991; Lecanuet, 1996). Trevarthen (1999–2000) has argued that the fetal experience is musical in the sense that the fetus feels the rhythms of the mother's body and hears the inflections—the melodic contours—of external noises, including speech and music. At the present time, it is somewhat difficult to assess the importance of this kind of musical stimulation on later development. The commercial music market offers pregnant women a variety of special music programs that are supposed to have a positive effect on the state of health of the mother and

unborn child. However, there is a lack of decisive evidence about these programs' influence on the fetus.

In early infancy, parents and the family provide the most important influences for the musical development. During the infant phase, the nonverbal communication (mother–infant cooing or *motherese*) between child and mother or caretaker is crucial for communication. The expressive communicative content of the parent–child interaction (e.g., emotions, needs, mental states) is mediated primarily by musical parameters such as pitch, melody, rhythm, tempo, and dynamics (Noy, 1968; Papoušek, 1996). This "communicative musicality" (Malloch, 1999–2000, p. 31) enables the attunement and sharing of mental states, emotions, and meanings. By use of preverbal, quasi-musical interaction between parents and child in the course of child care, musical competencies are almost incidentally developed, and as a result and without being aware of it parents provide their infants with a type of elementary music education that stops with the acquisition of language (Papoušek & Papoušek, 1995). This kind of musical communication is presumably the most important environmental influence on musical development in early childhood (see Noy, 1968).

### *Family and Musical Development*

Although there is clear evidence from everyday experience and from empirical data that the family environment exerts an important influence on musical development, there is a lack of comprehensive theories that connect the results of the existing studies. Badur (1999) examined the results of studies that dealt with the relationship between the family environment on the one hand and music abilities, skills, attitudes, and preferences, as well as musical careers and excellence, on the other. The result of the analysis shows that music-related activities of the family support the musical development of children in many ways. Such activities are primarily singing and making music together. Additional activities such as attending concerts with the family, talking about music, and practicing in the presence of parents may also be useful. Other characteristics of the family such as the music abilities and attitudes of the parents or other family members or the presence of musical instruments only seem of importance if they lead to musical activities in which the child is involved.

In an extensive biographical study on the careers of outstanding Polish musicians, Manturzewska (1995) also examined the characteristics of the family backgrounds. She concluded that the family background constitutes the most influential factor for a musical career in the realm of classical music. Among the characteristics of the outstanding musician's home environment (apart from the aforementioned features) Manturzewska (1995, p. 15) mentions:

- Child-centered attitude of the parents with emphasis on the musical education of the child
- Deliberate organization and channeling of child's interests, time, and activities
- At least one person in the family believing in the potential of the future musician and encouraging the child

- Music being a genuine value in family life
- Emphasis not being placed on a musical career but on enjoying making music
- Praise and rewards even for small successes
- A positive emotional atmosphere for musical activity
- Careful selection of teachers and monitoring of musical development
- Conscious and active organization of a supportive and understanding network for the child, including personal contacts to professional musicians and music teachers
- Willingness to invest considerable time and effort in musical activities

Similar findings have emerged from the studies of other authors (e.g., Bastian, 1989; Sloane, 1985; Sosniak, 1985). Since the parental behavior serves as a model for children, parental musical involvement (instrumental or vocal) may be motivating for children.

The importance of parental support is also evident in the study by Davidson, Sloboda, and Howe (1995/1996; see also Sloboda & Davidson, 1996; Zdzinski, 1996). The authors interviewed 257 children between the ages of 8 and 18 who had instrumental lessons and performed to varying levels of achievement with regard to the role that parents and teachers played. However, the findings indicated that children in the highest achieving group were supported the most and most consistently by their parents, up to the age of 11. Thereafter the parents' support diminished while the children were increasingly driven by intrinsic motives to practice regularly by themselves. Conversely, for the children in the groups with lower levels of achievement and the dropouts, parental support was initially weak and set in forcefully during the teenage years in order to motivate the children. The authors interpreted this result as the parents' last (and often unsuccessful) attempt to keep their children playing. Note that the active parents were not necessarily professional musicians but rather interested amateurs and music buffs. The least musically interested were the parents of those children who had stopped playing and taking lessons. In sum, the results underscore the decisive importance of early and continuous parental support. According to the authors, high levels of musical achievement are most likely unattainable without such supportive parental involvement.

Although the influence of family environment may be crucial, there are examples of successful musicians, such as the jazz trumpeter Louis Armstrong, who received very little parental care and almost no formal education. However, a detailed reading of Armstrong's biography (see Collier, 1983) shows five critically important factors in his musical development:

1. Early frequent and casual exposure to musical stimuli
2. Opportunities over an extended period of time to explore the jazz musical medium
3. Early opportunities in music to experience intense positive emotional or aesthetic states
4. An opportunity to amass large numbers of hours of practice
5. A number of externally motivating factors such as role model performers whom Armstrong used to watch to gain "informal tuition"

These environmental factors have been fairly systematically studied by music psychologists and have now been identified as the key factors in musical development (see Sloboda & Davidson, 1996). This example suggests that a successful career in music without noteworthy support of the family background is possible but exceptional.

### *The Role of the Teacher*

Teachers are perhaps the most important early influence besides the parents, not only because teachers transmit musical abilities but also because they more or less influence musical tastes and values and are role models and hold a key position with regard to motivation—for good or for bad. Results of a study by Davidson et al. (1998) suggest that the first teacher is of special and critical benefit. Professional qualifications, along with personality characteristics and interpersonal skills, come into play (e.g., Davidson, Howe & Sloboda, 1997; Olsson, 1997). In the study by Davidson et al. (1995/1996), the first teacher was perceived very differently by the young learners they interviewed. The students with the highest achievements found their teachers to be entertaining, friendly, and proficient musicians, whereas the lowest-achieving students remembered their teachers as unfriendly and incompetent. With increasing age this combination of teacher characteristics did not change for the lowest group, but the higher-achieving students started to distinguish between their teachers' personality (e.g., friendly, encouraging) and the professional qualities. As the good and best students become older and more self-motivated, it is the professional quality that becomes most relevant. Children who had dropped out of their lessons did not make this distinction. When the teacher was personally pleasant, he or she was also perceived as professionally satisfactory. We can see here how crucial the personal aspects are in addition to the professional ones; especially at young ages, they may motivate students to play. This highlights the importance of the emotional climate that surrounds musical experiences. Children who develop outstanding instrumental achievements tend to have learned in a positive emotional atmosphere that was enjoyable and free of anxiety. The learning context of children who drop out tends to be negative and characterized by anxiety. Other influences of socialization in terms of personality are discussed by Kemp and Mills in chapter 1 and O'Neill and McPherson in chapter 3.

## Societal Influences

### *Historical and Generational Influences*

The general (including musical) education system of a country and its cultural institutions and traditions are essential determinants of the musical development of its people. A comprehensive overview and description of these factors in different countries is provided by Hargreaves and North (2001). The importance of the sociocultural framework as an environmental factor that influences

musical development becomes more evident if we look at these factors from a historic point of view. In a transhistorical time-series analysis that covers several centuries, Simonton (1975, 1977) found that political circumstances such as political instability affected the creative development of composers.

In a biographical study, Gembris (1997) explored the influence of time- and generation-specific factors in musical biographies. The interviewees were professional musicians, music teachers, music amateurs, and nonmusicians born between 1903 and 1975 who lived in (the former West) Germany. The analysis of biographies showed that the development of musical experiences, attitudes, preferences, careers, and so forth, was more or less influenced by time- and generation-specific factors such as World War I, National Socialism, World War II, the spreading of rock-and-roll and pop music, and the spreading of portable radios and cassettes players and videos.

For example, during the period of the Nazi government, between 1933 and 1945, in Germany, Jewish musicians were persecuted. They could not practice their trade, and their lives and careers changed dramatically. The political system controlled and restricted the musical experiences and activities of virtually all people living at that time. Similarly, in the former socialist Eastern Bloc countries the freedom to produce music and to develop musical activities was constrained and regulated by the state. People who were considered politically unreliable due to their beliefs were excluded from music academies and universities. A quite different kind of generation-specific influence in the Western nations has led to considerable differences in the musical experiences, attitudes, and preferences between those generations who grew up with rock and pop music and those who grew up in the decades before the emergence of rock and pop music in the 1950s–1960s. Today we cannot predict the effect of the spreading of music videos, portable music players, computer technology, and the constant background of music in everyday life on the musical development of young people.

### *Mass Media and Peer-Group Influences*

According to a representative study (*KidsVerbraucherAnalyse 99*, 1999), 93% of 6- to-9-year-old children hear music in their leisure time and this proportion grows to 98% among 10- to-13-year-olds. Brown, Campbell, and Fischer (1986) estimated that 12- to-14-year-old American adolescents listen to music more than seven hours a day on average. If the trend for the continuous spreading of media such as music videos, CDs, MiniDiscs, MP3-players, and so on, continues, young people will be exposed to more and more music. It is certain that young people hear much more music outside their music lessons than in them. This, together with other factors (e.g., the strongly emotional nature of music listening, influences of the peer group on music preferences), makes it probable that the development of listening behaviors and preferences will be far more influenced by such environmental factors outside the school than by music lessons. But because of methodological issues, it is very difficult to assess these environmental factors with any accuracy and to calculate the extent of their possible effects. For instance, the validity of assessment may be problematic and depends on the

precise definition of the meaning of *music listening*. Concerning music listening, Lull (1987) differentiates among *exposure* (simply being exposed to music), *consumption* (what is learned and remembered because of the contact with music), and *use* of music (conscious listening to and use of music). Most investigations do not take these differences into consideration. Furthermore, there are extreme interindividual differences as well as large variations in the amount of time children watch and listen to music (Huston & Wright, 1998). Musical development and musical biographies cannot be explained without consideration of these kinds of influences. Unfortunately, there is far too little research on these important topics.

## Educational Implications

From the literature discussed earlier, it is obvious that environmental factors take effect from very early in the developmental process. Certainly in the United States and increasingly throughout the Western world, efforts are being made to capitalize on opportunities that environmental stimulation might provide for individual development and advancement in a number of domains. There are clear advantages of such concern, for as Radford (1991) points out, children placed in highly enriching environments will progress faster than the traditionally understood stages of cognitive development like those proposed by Piaget. This suggests that children can only move to higher levels of conceptual assimilation after having worked with specific experiences allied to their maturational state (see Gardner, 1982, for a useful discussion of Piagetian approaches to understanding mental development).

However, in agreement with Radford (1991), we are cautious about recent ideas that might be interpreted as attempts to create exceptional musical achievers. It is possible to suggest ideal social and emotional environmental circumstances that have emerged from the research literature, but we are reluctant to say that these will have broad-ranging consequences. In fact, it is fairly well established (see Sloboda, Davidson & Howe, 1994, for a detailed discussion of these issues) that *hothouse environments* give young music learners no significant advantage over their peers when they reach adulthood. Indeed, Burland and Davidson (submitted) worked with the young adults who presented themselves as outstanding young musicians in the original study of young musicians in specialist musical education (see Howe et al., 1995; Davidson et al., 1996). These students were initially interviewed in 1991; now, one decade later, many of them are teachers, lawyers, and administrators, as well as outstanding musicians. So early indications of advancement and success do not necessarily mean there will be adulthood advantages.

The results of available studies suggest that the following environmental factors influence musical development.

Fetus. The research of Hepper (1991) suggests it could be worth playing music to the fetus during pregnancy. A range of CDs are now available for pregnant

women—some to assist the mother to relax and some to relax and stimulate the fetus. But any other music may be suitable if the mother likes it.

**Infant and Toddler.** Surely one of the most important things is the intensive nonverbal, quasi-musical communication between the mother or other caretakers and the child. This communication, together with holding, rocking, and singing to the child, connects the loving care of the parents to the experience and the active use of musical parameters and their meaning. Because many parents have forgotten how to sing, some music schools and music teachers offer special parent–baby singing courses in which parents can learn how to sing with their child. This is a relatively new but very important approach of music education for several reasons: First, the emotional communication between parents and child is supported. Second, skills in the use of musical elements and song singing are improved. Third, this may be an opportunity for the parents to become interested in their own musical activity and development, which in turn may produce a greater attention to the musical development of their child.

Any kind of music exposure and musical enrichment of the environment can be for the good of the child. Gordon (1990) suggests playing a great variety of music (e.g., different styles, instrumentation, etc.) to provide as rich sources of musical stimulation as possible. Early exposure to music helps the child to naturally assimilate the rules of the prevailing harmonic and tonal language.

For instrumental learning, the Suzuki approach (see Suzuki, 1969, 1981; chapter 7 in this volume) and the Pace Method (see Pace, 1999) are excellent ways to promote very early engagement and to encourage practice and creativity—but it seems important to begin before the age of 5 years. These programs provide the children with the key adult support necessary to engage and sustain their interest in music. In both cases, teachers take on roles of a parental nature, gently coaxing the child, but within a structured framework.

**Parents.** In the aforementioned programs parents are encouraged to attend lessons and learn themselves or, alternatively, to take note of what has been done and what should be done in home practice. This teacher–parent interaction was found to be of critical importance in all the studies undertaken by Davidson, Sloboda, Howe and Moore throughout the 1990s. The parent should be prepared to sit with his or her child in the early stages of practice and attend lessons to find out what needs to be done at home. Then, as the child's skills begin to develop, a more autonomous practicing strategy can be promoted, but regular encouragement of and listening to practice activities can help enormously.

**Teacher.** From what is known of teachers and their strategies, it seems that a warm and friendly first teacher is essential. Then as the child advances a more demanding taskmaster can be sought. Here it seems that parents should take note of how their children feel about their music teachers. The one-to-one traditional learning situation can be highly rewarding, but it can also be very intimidating, so the child needs to feel safe with the teacher.

**Diversity.** It seems that teachers, parents, and others engaged in promoting musical activities for young people should be prepared to provide opportunities for experimentation, a broad range of creating and participating experiences. Here the critical task seems to be to ask questions in order to stimulate the imagination and general motivation of the child as well as to provide technical expertise and ideas about expression and musicality (Hallam, 1998). Of course, children will often perform to a standard of behavior expected of them (Rosenthal & Jacobsen, 1968; Blatchford et al., 1989). It is for this reason that in very large classroom contexts children often underachieve, as the pace is set for the average child in the group, not the individual.

**Challenge.** It is important for the child to always be sufficiently challenged and stretched without being overwhelmed or threatened. Because of the individuality of every child it is not possible to give general statements that concern the optimal challenge; finding this out is one of the most subtle tasks of the teacher.

All of the preceding assumes the availability of sufficient time and money to develop these skills. Unfortunately, in Western culture music is an expensive pursuit, and we are aware that much of what we have said points toward the middle and upper classes. Consider again the case of Louis Armstrong, who developed into one of the world's leading jazz trumpeters without money or formal lessons or, for that matter, a supportive mother. But his environmental conditions did conspire to provide him with many of the necessary ingredients that the middle-class Western child who is learning classical music requires in order to develop musically.

By following the preceding points, parents and teachers can create a positive and stimulating environment that systematically nurtures musical skill acquisition. However, the research domain itself is far from complete, with many questions left unanswered. The following need more research:

- Nonshared environmental influences
- Musical development outside the arena of so-called classical music—for example, the careers of rock, pop, and folk musicians
- The real impact of leisure time factors such as youth and fan cultures and the media on musical development

Finally, methodologies could still be explored, improved, and developed. For practical reasons, we cannot trace all aspects of the musical involvement and development of every individual in an experimental sample. However, the methodological problem of measuring the influence of the environment on musical development needs to be tackled if we are to understand more fully the implications of environmental influences on how people develop musical skills.

### *References*

Asendorpf, J. B. (1999). *Psychologie der Persönlichkeit* (2nd ed.). Berlin: Springer.

Badur, I.-M. (1999). Musikalische Sozialisation in der Familie. Ein Forschungsüberblick. In C. Bullerjahn, H. J. Erwe, and R. Weber (Eds.), *Kinder-Kultur.*

*Ästhetische Erfahrungen–Ästhetische Bedürfnisse* (pp. 131–158). Opladen: Leske + Budrich.

Bastian, H. G. (1989). *Leben für Musik: Eine Biographie-Studie über musikalische (Hoch-)Begabungen.* Mainz: Schott.

Bee, H. (1992). *The developing child* (6th ed.). New York: Harper Collins.

Blatchford, P., Burke, J., Farquhar, C., Plewis, I., & Tizard, B. (1989). Teacher expectations in infant school: Associations with attainment and progress, curriculum coverage and classroom interaction. *British Journal of Educational Psychology, 59*(1), 19–30.

Bouchard, T. J., Lykken, D. T., McGue, M., Segal, N., & Tellegen, A. (1990, October). Sources of human psychological differences: The Minnesota study of twins reared apart. *Science, 250,* 223–228.

Brown, E. F., Campbell, K., & Fischer, L. (1986). American adolescents and music videos: Why do they watch? *Gazette, 37,* 19–32.

Burland, K., & Davidson, J. W. (submitted). Training the talented.

Collier, J. L. (1983). *Louis Armstrong: An American genius.* New York: Oxford University Press.

Davidson, J. W., Howe, M. J. A., Moore, D. M., & Sloboda, J. A. (1996). The role of parental influences in the development of musical ability. *British Journal of Developmental Psychology, 14*(4), 399–412.

Davidson, J. W., Howe, M. J. A., Moore, D. M., & Sloboda, J. A. (1998). The role of teachers in the development of musical ability. *Journal of Research in Music Education, 46*(1), 141–160.

Davidson, J. W., Howe, M. J. A., & Sloboda, J. A. (1997). Environmental factors in the development of musical performance skill over the life span. In D. J. Hargreaves and A. C. North (Eds.), *The social psychology of music* (pp. 188–206). Oxford: Oxford University Press.

Davidson, J. W., Sloboda, J. A., & Howe, M. J. A. (1995/1996, Winter). The role of parents and teachers in the success and failure of instrumental learners. *Bulletin of the Council for Research in Music Education, 127,* Special Issue: The 15th ISME International Research Seminar, 40–44.

Davies, J. B. (1994). Seeds of a false consciousness. *Psychologist, 7*(7), 355–357.

DeCasper, A. J., & Spence, M. J. (1991). Auditorily mediated behaviour during the perinatal period: A cognitive view. In M. J. S. Weiss and P. R. Zelazo (Eds.), *Newborn attention: Biological constraints and the influence of experience* (pp. 142–176). Norwood, NJ: Ablex.

Gardner, H. (1982), *Human development* (2nd ed.). Boston: Little, Brown.

Gembris, H. (1997). Time specific and cohort specific influences on musical development. *Polish Quarterly of Developmental Psychology, 3*(1), 77–89.

Gembris, H. (1998). *Grundlagen musikalischer Begabung und Entwicklung.* Augsburg: Wissner.

Gordon, E. E. (1986). *The nature, description, measurement, and evaluation of music aptitudes.* Chicago: GIA.

Gordon, E. E. (1990). *A music learning theory for newborn and young children.* Chicago: GIA.

Hallam, S. (1998). *Instrumental teaching.* Oxford: Heinemann.

Hargreaves, D. J., & North, A. C. (Eds.) (2001). *Musical development and learning: The international perspective.* London: Cassel.

Hepper, P. G. (1991). An examination of foetal learning before and after birth. *Irish Journal of Psychology, 12,* 95–107.

Howe, M. J. A., Davidson, J. W., Moore, D. M., & Sloboda, J. A. (1995). Are there early childhood signs of musical ability? *Psychology of Music, 23*(2), 162–176.

Huston, A. C., & Wright, J. C. (1998). Mass media and children's development. In W. Damon, I. E. Sigel, and K. A. Renninger (Eds.), *Handbook of child psychology*: Vol. 4. *Child psychology in practice* (5th ed., pp. 999–1058). New York: Wiley.

*KidsVerbraucherAnalyse 99* (1999). Bergisch-Gladbach: Bastei-Verlag, Gustav H. Lübbe GmbH, Axel Springer Verlag, & Verlagsgruppe Bauer.

Lecanuet, J.-P. (1996). Prenatal auditory experience. In I. Deliège and J. A. Sloboda (Eds.), *Musical beginnings: Origins and development of musical competence* (pp. 3–34). Oxford: Oxford University Press.

Lull, J. (1987). Listeners' communicative uses of popular music. In J. Lull (Ed.), *Popular music and communication* (pp. 140–174), Newbury Park, CA: Sage.

Malloch, S. N. (1999–2000). Mothers and infants' communicative musicality. *Musicae Scientiae*, Special Issue: Rhythm, Musical Narrative, and Origins of Human Communication, 29–57.

Manturzewska, M. (1995). Das elterliche Umfeld herausragender Musiker. In H. Gembris, R.-D. Kraemer, and G. Maas (Eds.), *Musikpädagogische Forschungsberichte 1994* (pp. 11–22). Augsburg: Wissner.

Noy, P. (1968). The development of musical ability. *Psychoanalytic Study of the Child, 23*, 332–347.

Olsson, B. (1997). The social psychology of music education. In D. J. Hargreaves and A. C. North (Eds.), *The social psychology of music* (pp. 290–305). Oxford: Oxford University Press.

Pace, R. (1999). *The essentials of keyboard pedagogy: I. Sight-reading and musical literacy*. Chatham, NY: Lee Robert's Music.

Papoušek, H., & Papoušek, M. (1995). Beginning of human musicality. In R. Steinberg (Ed.), *Music and the mind machine: The psychophysiology and psychopathology of the sense of music* (pp. 27–34). Berlin: Springer.

Papoušek, M. (1996). Intuitive parenting: A hidden source of musical stimulation in infancy. In I. Deliège and J. A. Sloboda (Eds.), *Musical beginnings: Origins and development of musical competence* (pp. 88–112). Oxford: Oxford University Press.

Plomin, R., & DeFries, J. C. (1999). The genetics of cognitive abilities and disabilities. In S. J. Ceci and W. M. Williams (Eds.), *The nature-nurture debate* (pp. 178–196). Oxford: Blackwell.

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype–environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin, 84*, 309–322.

Plomin, R., DeFries, J. C., McClearn, G. E., & Rutter, M. (1997). *Behavioral genetics* (3rd ed.). New York: Freeman.

Radford, J. (Ed.). (1991). *Talent, teaching and achievement*. London: Jessica Kingsley.

Reiss, D., Plomin, R., Hetherington, E. M., Howe, G., Rovine, M., Tyron, A., & Stanley, M. (1994). The separate worlds of teenage siblings: An introduction to the study of the nonshared environment and adolescent development. In E. M. Hetherington, D. Reiss, and R. Plomin (Eds.), *Separate worlds of siblings* (pp. 63–109). Hillsdale, NJ: Erlbaum.

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom*. New York: Holt, Rinehart & Winston.

Satt, B. J. (1984). *An investigation into the acoustical induction of intrauterine learning.* Doctoral dissertation, California School of Professional Psychology, Los Angeles.

Seckl, J. R. (1998). Physiologic programming of the fetus. *Clinical Perinatology, 25*(4), 939–962.

Simonton, K. D. (1975). Sociocultural context of individual creativity: A transhistorical time-series analysis. *Journal of Personality and Social Psychology, 32*(6), 1119–1133.

Simonton, K. D. (1977). Creative productivity, age, and stress: A biographical time-series analysis of 10 classical composers. *Journal of Personality and Social Psychology, 35*(11), 791–804.

Sloane, K. D. (1985). Home influences on talent development. In B. S. Bloom (Ed.), *Developing talent in young people* (pp. 439–476). New York: Ballantine.

Sloboda, J. A., Davidson, J. W., & Howe, M. J. A (1994). Is everyone musical? *Psychologist, 7*(8), 349–355.

Sloboda, J., & Davidson, J. (1996). The young performing musician. In I. Deliège and J. A. Sloboda (Eds.), *Musical beginnings: Origins and development of musical competence* (pp. 171–190). Oxford: Oxford University Press.

Sloboda, J. A., Davidson, J. W., Howe, M. J. A., & Moore, D. M. (1996). The role of practice in the development of expert musical performance. *British Journal of Psychology, 87*(2), 287–309.

Sosniak, L. A. (1985). Learning to be a concert pianist. In B. S. Bloom (Ed.), *Developing talent in young people* (pp. 19–67). New York: Ballantine.

Suzuki, S. (1969). *Nurtured by love.* New York: Exposition.

Suzuki, S. (1981). *Ability development from age zero.* Athens, OH: Ability Development Associates.

Trevarthen, C. (1999–2000). Musicality and the intrinsic motive pulse: Evidence from human psychobiology and infant communication. *Musicae Scientiae,* Special Issue: Rhythm, Musical Narrative, and Origins of Human Communication, 155–215.

Zdzinski, S. F. (1996). Parental involvement, selected student attributes, and learning outcomes in instrumental music. *Journal of Research in Music Education, 44*(1), 34–48.

# 3

# Motivation

SUSAN A. O'NEILL & GARY E. McPHERSON

Research on motivation in music seeks to understand how children develop the desire to pursue the study of a musical instrument, how they come to value learning to play an instrument, why they vary in the degree of persistence and the intensity they display in achieving their musical goals, and how they evaluate and attribute their success and failure in different achievement contexts. Current theories view motivation as an integral part of learning that assists students in acquiring the range of adaptive behaviors that will provide them with the best chance of achieving their own personal goals. We review the literature on these topics and provide a framework for understanding the complex range of thoughts, feelings, and actions that either sustain or hinder musicians through the many years that it takes to develop musical skills.

Why do some children seek the challenges of learning and persist in the face of difficulty, while others, with seemingly equal ability and potential, avoid challenges and withdraw when faced with obstacles or difficulties? Over the past 20 years, this fundamental question has underpinned an enormous body of educational and psychological research. The findings have contributed to our understanding of motivation by clarifying what initiates a desire to pursue certain goals, by explaining why certain goals are valued over others, by describing how students vary in the degree of persistence and intensity they display in achieving their own goals, and by specifying how children evaluate and attribute causes for their success and failure in different achievement contexts (see Pintrich & Schunk, 1996, for a comprehensive overview). An important outcome of this research is that motivation is no longer viewed as a distinct set of psychological processes but as an integral part of learning that assists students to acquire the range of behaviors that will provide them with the best chance of reaching their full potential.

With this in mind, our chapter begins with a review of key theories and research that can help musicians understand the complex nature of motivation as it applies to music performance achievement. This is followed by a discussion of teaching strategies that can be used to help music students establish and sustain positive motivation as they work toward achieving their goals.

## Motivation Theories

### *Expectancy-Value Theory: Why Do I Want to Play an Instrument?*

Expectancy-value theory, one of the most important strands of motivation research, seeks to provide a framework for understanding why individuals are interested in or care about an activity to a sufficient degree that they believe it might be important to them in the future (Pintrich & Schunk, 1996). The model describes four components relevant to students' expectations and valuing of an activity. The first, *attainment value*, refers to how important a student believes it is to do well on a task. For example, playing well at a music recital will be important for a student whose self-concept involves a strong sense of identity in terms of being a capable musician. Second, the feeling of enjoyment an instrumentalist has when performing for the sheer pleasure of making music is defined as *intrinsic motivation.* Third, students form perceptions of the *extrinsic utility value* of learning their instrument according to its usefulness to their future goals, including career choices. A student who plays an instrument exclusively for the pleasure of performing with an ensemble will value music performance differently from a student whose intention is to become a professional musician. Fourth, the perceived negative aspects of learning an instrument, such as the amount of practice needed to continue improving, are defined as the *perceived cost* of engaging in the activity. Children may decide that the cost of practicing each day is not worth the effort, because it does not leave sufficient time for other things in their life, such as sports or social activities (Eccles, Wigfield & Schiefele, 1998).

Using this framework, researchers who studied academic achievement in subjects like mathematics and reading have found that even very young children are able to distinguish between what they like or think is important for them and perceptions of their own personal competence in a particular field (Wigfield, 1994; Wigfield & Eccles, 1992) and that these beliefs predict how much effort they will exert on the task, their subsequent performance, and their feelings of self-worth, even after previous performance has been taken into account (see also Eccles & Wigfield, 1995; Eccles et al., 1998; Wigfield et al., 1997). There is also evidence that individuals' subjective task values (i.e., the degree of importance and interest a particular activity has to an individual) are important predictors of activity choices in adolescent years (Eccles et al., 1983).

Research in music is consistent with findings in academic disciplines and provides important clues to our understanding of how young people come to value musical learning and view it as important to themselves and their goals.

For example, in Britain approximately 48% of 5- and 6-year-old children express an interest in learning to play an instrument, but by age 7 this desire has halved to about 25%, where it remains constant until the age of 11, when it declines to only 4% of nonplaying 14-year-olds (Cooke & Morris, 1996). In the United States, opportunities to learn an instrument in a school instrumental program typically become available at a time when children's interest in learning music and beliefs about its usefulness and importance are rapidly declining (Eccles et al., 1983, 1993). Eccles et al. (1993) found that even children who have had very little experience with instrumental music appear to have just as reliable and differentiated self-beliefs about their ability and the importance they place on this activity as for activities (such as reading and sports) with which they have considerably more experience. Not surprisingly, children who devalue instrumental instruction and consider themselves lacking in musical ability will be more likely to engage in it for a short time and then stop (Wigfield et al., 1997; Wigfield, O'Neill & Eccles, 1999).

McPherson (2000) studied 133 young beginning instrumentalists between the ages of seven and nine and found that children bring to their music instruction expectations and values that potentially shape and influence their subsequent development. For example, interviews with the children before they began instruction show that they were able to make clear statements about their valuing of and expectations for music learning without a great deal of previous experience. They could differentiate among their interest in learning a musical instrument, the importance of being good at music, whether they believed their learning would be useful to their short- and long-term goals, and the cost of participation, in terms of the effort needed to continue improving. For many, learning an instrument was no different from participating in a team sport, taking up a hobby, or pursuing other recreational activities. Most of the children were intrinsically interested in learning an instrument but did not see it as important to their long-term future careers. Others were less intrinsically motivated but recognized the importance or utility value of learning in terms of their overall education. For the majority of children, learning an instrument was something useful to do while they were in school but of far less value in later life. Only a handful viewed their involvement as something that could possibly lead to a future career. Even before commencing, many children were also able to provide a definite view of their own potential compared to those of their peers.

One of the more interesting findings from the McPherson study was the relationship between the children's commitment to learning their instrument before they started lessons and their achievement nine months later. Students who predicted that they would play their instrument for only a few years progressed the slowest, irrespective of the amount of practice they did at home. Students who predicted that they would play their instrument until the end of their schooling achieved higher performance levels, which increased according to the amount of their practice during the period studied. The highest achieving students were those who expressed a long-term commitment to playing, coupled with high levels of practice. These students, who indicated that they planned to play their instrument for most of their lives, were typically more inclined to express in-

trinsic reasons (e.g., "I've always liked music for as long as I can remember") rather than extrinsic reasons (e.g., "I wanted to learn because all my friends were joining the band") for wanting to play.

### *Self-Efficacy: How Well Can I Perform?*

Related to children's valuing of an activity is their sense of competence or *self-efficacy*. Competence beliefs have been described as evaluations of how good one is at an activity (e.g., Eccles et al., 1983; Nicholls, 1984, 1990), expectancy beliefs about how good one's performance will be in the future (e.g., Eccles et al., 1983), or self-efficacy beliefs that one can produce the desired outcome (e.g., Bandura, 1997). In other words, self-efficacy is associated with the degree to which a musician believes in his or her own ability and capacity to achieve certain goals (Stipek, 1998). According to Pajares (1996), "Self-efficacy beliefs act as determinants of behavior by influencing the choices that individuals make, the effort they expend, the perseverance they exert in the face of difficulties, and the thought patterns and emotional reactions they experience" (p. 325). Perceptions of personal competence are so powerful that they are theorized to influence a student's motivation and future decision to continue developing his or her skill in an area (Hackett, 1995).

Research in music supports these assertions. McPherson and McCormick (1999) studied 190 pianists aged between 9 and 18 who completed a questionnaire immediately before undertaking an externally graded performance examination. Results indicated that the level of the pianists' self-efficacy before they entered their performance examination helped to predict their subsequent examination result. This finding is consistent with educational research that suggests that students who display high self-efficacy will tend to perform at a more advanced level in examinations than their peers who display the same level of skill but lower personal expectations (Pintrich & Schunk, 1996). This is because high levels of self-efficacy strengthen confidence and ensure persistence, often in spite of difficulties (Pajares, 1996).

According to Eccles, Wigfield, and Schiefele (1998), more work is needed on the links between competence beliefs and values and whether they relate differently to performance and choice. In music, for example, it is important to establish whether these beliefs are associated with the amount of time students spend practicing and the choice of easy or difficult musical tasks. Yoon (1997) studied 849 children between the ages of 8 and 11 to investigate the extent to which competence and values predicted activity choice (sports versus music) and frequency of practice. Results indicated that self-perceptions of competence (i.e., "How good are you at playing a musical instrument?") and subjective task values (i.e., "How important is learning to play an instrument?") played a key role in predicting both activity choice and amount of practice. In other words, the findings suggest that children are more likely to engage in musical activities when they feel capable in music and value it.

However, findings from a recent study by O'Neill (1999b) suggest that valuing a musical activity may be even more important in sustaining motivation for prac-

tice than believing in one's ability to succeed. O'Neill found that the extent to which 60 young musicians aged between 12 and 16 valued their practice sessions (i.e., "Was this practice important to you?"; "How important was this practice in relation to your overall goals?") was a significant predictor of the amount of time they spent practicing. Interestingly, frequency of practice was not predicted by the students' beliefs about how competent they were (i.e., "Were you succeeding at what you were doing?") and whether they were living up to their own and others' expectations. One interpretation is that the young musicians who valued their practice highly did so because it provided them with the opportunity to demonstrate aspects of their competence, which in turn reinforced the value they placed on learning to play an instrument and motivated them to achieve their goals. According to Eccles and her colleagues (e.g., Eccles, 1987; Eccles & Harold, 1992), values are linked to more stable beliefs about the characteristics of one's overall self-concept and sense of identity, whereas competence beliefs are more unstable and likely to fluctuate in different situations and with the demands of different tasks. In addition, it is possible the young musicians were not interpreting the notion of importance in the same way (e.g., they may have been more or less concerned with intrinsic or extrinsic reasons for doing the task), and this may have contributed to differences in amount of practice. It would be useful if future research examined the extent to which self-perceptions of competence and values might vary in relation to different practice activities (e.g., formal, deliberate practice versus informal, fun practice; easy tasks versus demanding tasks) and how this might relate to the amount of time spent practicing different tasks.

### *Flow Theory: Matching Challenge to Skill*

Researchers have also conceptualized motivation in terms of the changes that take place within the person when a student is actually involved in learning music. From this perspective, researchers study the types of activities that students find intrinsically motivating and compare these with the types of activities that result in less efficient learning. This is the essential ingredient of Csikszentmihalyi's (1990) *flow theory*, which suggests that optimal experience requires a balance between roughly equal levels of perceived challenge and skill in a situation that involves intense concentration. According to this explanation, activities are seen as pleasurable when the challenge is matched to the person's skill levels. If an activity is too easy and skill levels are high, boredom will develop; if an activity is too difficult and skill levels are low, anxiety will result; if both challenge and skill levels are low, students feel apathy. To remain in flow, the complexity of the activity must increase by developing new skills and taking on new challenges. Flow experience is characterized by the presence of clear goals and unambiguous feedback, focused concentration, a sense of outcomes under the person's own control, a distorted sense of time (e.g., an hour of practice seems to go by quickly), losing a sense of self-awareness, and experiencing the activity as intrinsically rewarding (see also Csikszentmihalyi, Rathunde & Whalen, 1993).

O'Neill (1999a) examined the extent to which flow experiences accounted for differences in the amount of time spent practicing among 60 young musi-

cians between ages 12 and 16 who varied in their levels of performance achievement. The moderate achievers from a specialist music school reported fewer flow experiences when practicing than their higher-achieving peers at the same school and the musically active young people attending a nonspecialist state school. Although further research is needed, the results suggest that the evaluative context of a specialist music school contributed to the reduced flow experiences of the students who were considered less musically able than their peers. During individual interviews, the moderate achievers at the music school tended to regard the high-achieving students as competitive and criticized the lack of opportunity the school provided for them to engage in nonmusical activities such as sports and other arts programs. However, both the high achievers at the specialist music school and the musically active young people at the state school were more likely to describe their peers as supportive and encouraging and that they valued the social opportunities afforded by their school to engage in music making with other like-minded peers. An implication of these findings is that music educators need to consider different ways of fostering the motivation of music students who are considered less competent in order to ensure that the quality of their musical experiences remains intrinsically rewarding.

### *Attribution Theory: Why Did I Succeed or Fail?*

Beliefs about the causes of success and failure can influence a variety of future achievement behaviors, expectancies, self-perceptions, and other emotional responses (Austin & Vispoel, 1998). Expectations are central to continuing motivation, so understanding how students attribute the causes of their success or failure, particularly in a stressful environment such as a music performance, is an important part of understanding variation in students' performance. In attribution research, theorists explore the impact of these beliefs on expectations for future success and individuals' perceptions of both their own abilities and the difficulty of the various tasks. Weiner (1986, 1992), the researcher who has made the greatest contribution to this area, believes that it is not success or failure per se but the causal attributions made for these outcomes that influence students' expectations of whether they think they will continue doing well or poorly.

Attribution researchers have focused on the perceived relationship between children's achievement and the types of reasons they give to explain their performance. According to Stipek (1998), the most common attributions are ability ("I did well because I'm a good musician") and effort ("I did well because I practiced hard"). Less common attributions include luck ("I had a lucky day"), task difficulty ("The examiner asked me the easiest scales"), and strategy ("I practiced the hard part in small sections"). Weiner (1986, 1992) identified three dimensions of causal attributions: (1) whether the cause is internal or external to the student, (2) the extent to which the cause remains the same or changes, and (3) the extent to which the individual can control the cause. For example, there is an important distinction between how children view ability and how they view effort. Ability is viewed as something internal, stable, and beyond a student's control ("I can't

do this because I'm not a good musician"), whereas effort is seen as internal, unstable, and controllable ("If I do more practice I'll be able to play this piece"). In addition, students who perceive their success as being due to internal reasons such as effort are more likely to have a higher sense of self-worth (self-esteem) than students who believe their success was due to external reasons, such as luck.

A detailed review of available evidence by Eccles et al. (1998) shows that young children view ability and effort as complementary and are not always able to distinguish between these two constructs, but that their understanding changes over time. By the age of 11 or 12, children have come to understand these processes more deeply and begin to realize that instrumentalists with less ability need to try harder and practice more in order to achieve similar results to those of students with greater ability. The realization that you need to do more practice than someone who shows greater ability means that you will come to view yourself as a less capable musician.

In music, Vispoel and Austin (1993) found that junior high school students who attributed the failure of a fictitious music student to insufficient effort or poor learning strategies were more likely to anticipate improved future performance than students who attributed failure to a lack of ability. McPherson and McCormick (2000) conducted a study of 349 instrumentalists between the ages of 9 and 18 who were completing graded performance examinations. The results demonstrated that over 50% perceived their performance examination result to be a consequence of how much effort they had given to preparing for the examination or how hard they tried during the examination. The majority of students went into their examination reporting healthy attributions. If they did well, then they could attribute their success to having prepared thoroughly. If they did poorly, then they could blame their result on not having done enough preparation or not trying hard enough during the examination.

However, the findings also revealed that 12.4% of the beginners, 9.9% of the intermediate-level players, and 19.5% of the advanced musicians attributed their result to either luck, how hard the exam might turn out to be (i.e., task difficulty), or their overall ability. Other research suggests that students who believe that their success or failure is a result of ability tend to approach a task differently from their peers who associate success and failure with effort and that low achievers often make maladaptive or unhelpful attributions in comparison with high achievers (Arnold, 1997; Pintrich & Schunk, 1996). The latter students are more likely to attribute success to luck and their inability to complete a task successfully to such factors as ability. They are also less likely to feel that an increase in effort will have any positive benefits for their development or capacity to achieve at a higher level.

### *Mastery Motivation: How Confident Do I Feel?*

Weiner's work on attribution theory helped pave the way for research focusing on specific learning styles, such as the differences between adaptive and maladaptive student motivational patterns. Research by Dweck and her colleagues

(Dweck, 1986, 2000; Dweck & Leggett, 1988; Henderson & Dweck, 1990) demonstrated that children's motivational patterns influence their behavior and performance when they encounter difficulty or failure. For example, students who display adaptive *mastery-oriented* patterns tend to remain high in their persistence following failure and appear to enjoy exerting effort in the pursuit of task mastery. In short, students of this type remain focused on trying to achieve in spite of difficulties that may interfere with their progress. Conversely, maladaptive *helpless* patterns are associated with failure to establish reasonable valued goals or to attain goals that are within one's reach. Because helpless students feel that the situation is out of their control and that nothing can be done about it, they tend to avoid further challenges, lower their expectations, experience negative emotions, give up, and perform more poorly in the future. What is especially interesting about these two motivational patterns is that helpless children are often initially equal in ability to mastery-oriented children. Indeed, some of the brightest, most skilled children exhibit helpless behavior (Dweck, 2000).

O'Neill (1997) investigated the influence of these motivational patterns on the development of children's musical performance achievement after the first year of learning to play an instrument. The findings indicated that children who displayed mastery-oriented behavior after experiencing failure on a problem-solving task prior to beginning instrumental music instruction made more progress after one year of instrumental lessons than the children who initially displayed helpless behavior. However, when the amount of time spent practicing of those children who had demonstrated mastery and helpless behavior was compared, the results indicated that helpless children were practicing roughly twice as much as mastery children to reach the same level of moderate performance achievement. This suggests that although some helpless children were spending large amounts of time practicing, they appeared to be using their time less efficiently. For example, helpless children may avoid practicing pieces that pose particular difficulties, spend more time practicing items they can already play well, and use strategies that are unlikely to improve their performance. However, mastery children tend to seek challenges and adapt their practicing strategies in order to increase competence and improve their performance.

It is possible that motivational patterns may influence children's skill development much sooner in instrumental music learning than in other subject areas. During the initial stages of learning to play a musical instrument there are many challenges to overcome (e.g., posture, position, rhythm, tone, notation). As a result, many children experience confusion, difficulty, and failure very early in their instrumental training. In addition, unlike academic school subjects where children are rarely given the choice of not pursuing an activity, learning to play a musical instrument requires a great deal of autonomy, because it is often up to the child to decide when and where to practice and also whether new or difficult material is practiced or avoided. As O'Neill's (1997) study suggests, it is in these circumstances that helpless children may differ quite markedly from mastery-oriented children in their use of practice time.

## Implications for Teaching

What can music teachers do in order to provide a teaching environment that maximizes student motivation? As we suggested in the opening of this chapter, one of the most important elements is viewing motivation as an integral part of all teaching and learning. On the one hand, students who lack motivation will not expend the necessary effort to learn. On the other hand, highly motivated students are eager to come to music lessons and learn. In a review of instructional strategies and achievement motivation, Ames (1992) examined how different approaches to teaching and evaluating performance can influence students' motivational patterns in areas such as intrinsic motivation, attributions that involve effort-based strategies, and active engagement. In the following section, we explore these ideas further. Based on the research already discussed in the chapter, our aim is to highlight what we consider to be some of the most important dimensions of music teaching that can help foster students' confidence and commitment.

### *Potential*

As explained in chapter 1 by Kemp and Mills, one of the most widely debated issues in music is the notion of talent. Folklore would have us believe that those who are talented are more likely to achieve and therefore display more adaptive motivation. However, this is not always the case. As discussed earlier, research has demonstrated that some of the most skilled individuals who display high levels of initial competence exhibit helpless behavior, which can result in failure to reach their full potential (Dweck, 2000). Beliefs about ability influence both the goals students choose to pursue and their achievement behavior. Dweck (2000) distinguishes between two different self-theory beliefs: *entity* beliefs, where ability is viewed as a fixed trait (e.g., musical talent is something you are born with), and *incremental* beliefs, where ability is viewed as something malleable that can be increased through effort. Students who hold entity beliefs are more likely to develop an overconcern with proving their competence, avoid challenges, and show an inability to cope with failure or difficulty. One can imagine how difficult it would be to teach strategies for improving a poor performance to students who believe they simply do not have the potential to do it (i.e., "It either comes naturally or it doesn't"). In contrast, students who possess incremental beliefs thrive on challenges and view performance opportunities as providing chances to learn new things rather than merely display their ability. For students who consider performance achievement to be the end point of a learning process, a poor performance can be viewed as a problem that needs to be solved. They are therefore more receptive to learning new strategies for improving their performance.

Stipek (1996) provides the following strategies to improve students' self-efficacy (i.e., the belief that they can master a situation and produce positive outcomes) by emphasizing incremental beliefs and mastery motivation:

1. Teach students specific strategies (such as dividing a difficult section or piece into smaller parts) that can improve their ability to focus on their tasks.
2. Guide students in setting their own long-term goals and the short-term goals that will help achieve them.
3. Communicate your expectations that your students have the potential to achieve and provide them with supportive statements such as "You can do this," "You will improve this with practice," and "Keep going; you will get there". However, know your students well enough to understand how much skill support they need and get them to stretch their abilities realistically. It is important for teachers to encourage students who feel less able to take personal responsibility for their behavior, including setting appropriate goals in which they are challenged and perceive themselves as having the necessary skills to cope with the demands of the task.
4. Ensure that students are not overly anxious about their performance achievement. Although researchers have found that many successful students show moderate levels of anxiety (e.g., Bandura, 1997), some students have consistently high levels of anxiety, which can undermine their ability to succeed. Some of children's high anxiety is related to unrealistic achievement expectations and pressure by parents, teachers, or peers. Many students experience higher levels of anxiety as they reach higher skill levels where they face more frequent evaluation, social comparison, and, for some, experiences of failure (Eccles, Wigfield & Schiefele, 1998).
5. Provide students with positive adult and peer models. For example, students who observe teachers and peers coping effectively and mastering challenges often adopt the role models' behaviors. Modeling is especially effective in promoting self-efficacy when students observe success by peers who are similar in ability to themselves (see Schunk & Zimmerman, 1996).

### *Success and Failure*

Findings of attribution research reinforce the need for teachers to help their students develop adaptive attributional responses so that they can motivate their students for both short-term musical development and long-term musical involvement. To do this teachers need to monitor their students' attributions by spending time talking with them about what they have achieved and where they can improve and helping them to map out strategies that ensure that their musical practice is well organized and efficient (see chapter 10). Of vital importance is the need to encourage students to attribute any failure to controllable causes such as a lack of effort. Students who report effort attributions tend to display greater persistence and stronger emotional reactions, such as a feeling of pride for a high result or shame following failure (Austin & Vispoel, 1998). Thus stressing the importance of effort, in contrast to ability, is one way of diminishing the effects of counterproductive attributions with which students often explain their own achievement.

### *Enjoyment*

Research has begun to clarify important motivational differences when children practice repertoire assigned by their teacher rather than pieces they have chosen themselves. In a study of 257 students drawn from various levels of music training, Sloboda and Davidson (1996) found that high-achieving musicians tend to do significantly greater amounts of formal practice, such as scales, pieces, and technical exercises, than their less successful peers. Interestingly, they are also likely to report more informal practice, such as playing their favorite pieces by ear, "messing about," or improvising. Sloboda and Davidson conclude that these informal ways of practicing contribute to musical success because the highest achieving students are able to find the right balance between freedom and discipline in their practice.

In a detailed case study of a young beginning clarinettist, Renwick and McPherson (2000) discovered an elevenfold increase in the time spent practicing a piece that she chose to learn herself, as compared to repertoire assigned by the teacher. In addition to this remarkable difference there were also major differences in the quality of the girl's practice. When practicing the teacher-assigned pieces, the girl almost exclusively used a play-through approach, playing her pieces from beginning to end with little attention to correcting mistakes. With the piece she wanted to learn herself, there was a marked increase in the way she monitored and controlled her performance, as evidenced in greater use of silent fingering, silent thinking, singing, and more varied strategies for correcting wrong notes. Other useful techniques that teachers can explore to encourage students to vary their practice are concentrating on different repertoire on selected days, maintaining a log book to identify aspects that need revision and refinement, and continually varying habits, all of which help alleviate the daily grind of practicing.

### *Engagement*

Intrinsic motivation is maximized when students are given repertoire that requires a reasonable amount of effort to be learned. Repertoire that is completed with little effort or that causes confusion or frustration can result in low engagement during practice. Students will be more intellectually involved when practicing repertoire that requires higher order or divergent thinking and active problem solving than when practicing mundane drill and practice exercises. Encouraging students to come to the next lesson with five different ways of playing an exercise or to devise a set of similar exercises to overcome a technical problem will be more intrinsically motivating than merely telling them to practice an exercise until it has been mastered. McPherson and McCormick (1999) showed that a distinguishing feature of musicians who do greater amounts of practice is the level of their cognitive engagement while they practice. In their study, students who reported higher levels of practice tended to be more inclined to rehearse music in their minds (mental practice) and to make critical ongoing judgments on the success or otherwise of their efforts. They were also more capable of organizing their practice in ways that provided for efficient learning, such as practicing the pieces that needed

most work and isolating difficult sections of a piece that required further refinement. Anything that a teacher can do to encourage more active cognitive engagement as a student learns independently will have positive benefits for the pace of their improvement and subsequent intrinsic desire to continue learning.

### *Goals*

According to Covington and Omelich (1979), maintaining one's sense of competence is critical for maintaining a positive sense of self-worth. However, evaluations of performance, competitions, and social comparisons can make it difficult for some children to maintain such positive competence beliefs. Stipek, a leading educational psychologist, believes that stressing evaluation "focuses attention on performance goals, engenders a feeling of being controlled, and destroys whatever intrinsic interest students might have in a task" (1998, p. 172).

In British Commonwealth countries, North America, and other regions where music competitions and/or graded performance examinations flourish, music educators disagree about the real worth of evaluations that are based on some sort of ranking or assessment. Austin (1991), for example, believes that competitions curtail achievement, especially in situations where repeated failure will tend to diminish self-efficacy and self-determination: musicians "would derive greater pleasure and benefit from performing more frequently and in settings that (a) balance emotional risk with support from teachers, family, and peers, and (b) provide for detailed instructional feedback beyond numerical indicators of *musical greatness*" (p. 156). Massie (1992), however, argues that competitions encourage skill development (see further, Asmus, 1994).

Since formal evaluation of solo and group music performance is so common in schools and community music programs, there can be no simple answer to the question of whether it hinders or curtails motivation. Obviously, the context and nature of a music evaluation will vary considerably among different learning settings. However, success in any form of music evaluation is likely to depend not only on the amount of time students devote to practicing but also on the way they feel about the examination and the degree to which they believe in themselves and feel confident enough in their own ability to give the performance evaluation their best effort. Above all, students need to feel that they are participating in the evaluation because they want to, rather than because their teacher or parents want them to (see, for example, Davidson & Scutt, 1999).

Music examinations and competitions can provide opportunities for structuring the learning environment, especially when teachers use the requirements to make their own short- and long-term goals clear to the student. But it is also important to encourage students to identify their own goals and aspirations by allowing them some choice in the works they are to learn and the pieces they are to prepare for their performance. Any evaluation of a child's progress can be valued as an independent form of assessment, which can be used to define future goals. Information obtained from music evaluations therefore needs to be seen by students as helpful and suggestive, not definitive and critical. Many students thrive in an environment that includes the completion of grade ex-

aminations or participation in competitions in their instrument or voice. However, equally valid is the claim that not all students are suited to this type of stressful environment and that a good teacher will be perceptive enough to realize when a student is not ready or emotionally capable of being formally assessed and compared with his or her peers.

## Conclusion

In this chapter we have outlined research according to five theoretical perspectives. First, expectancy-value theory informs us about ways that children can value learning an instrument, as well as beliefs that concern what they think they might gain from their learning. Second, self-efficacy constructs allow us to understand how confident learners feel about their ability to perform on an instrument, especially when faced with a stressful situation such as a public music performance. According to the third perspective, flow theory, activities are more pleasurable when the challenge of the task and the musician's skill are optimally balanced. Fourth, attribution theory addresses the different reasons a musician will give to explain a good or poor performance. Fifth, mastery motivational patterns tell us about what occurs during the event and how some learners, despite their ability, give up when faced with difficulties.

Much work has been carried out on the role of socializing agents in assisting students to maintain sufficient motivation to practice and achieve (chapters 1 and 12 in this volume; McPherson & Davidson, in press). Indeed, there is little doubt that motivation to continue instrumental training is inextricably linked to the social and cultural environment, and so it is also important to consider how motivation for playing an instrument might be influenced by external factors such as parents and teachers. Important as these factors may be, no amount of parental support is likely to make a child without some intrinsic interest engage in the long-term effort required to succeed at even modest levels of musical competence. Consequently, understanding how students think about themselves, the task, and their performance is important if teachers are to establish and sustain a stimulating and challenging learning environment. Students need to feel that their involvement in learning to play an instrument provides them with a sense of personal choice and responsibility for reaching the goals that they set themselves. With this in mind, the challenge for teachers is to be receptive to each child's perspective on his or her own learning and to develop an understanding of the complex range of thoughts, feelings, and actions that either sustain or hinder the children through the many years that it takes to develop their musical skills.

### *References*

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*, 261–271.

Arnold, J. A. (1997). A comparison of attributions for success and failure in instrumental teaching among sixth-, eighth-, and tenth-grade students. *Update: Applications of Research in Music Education, 15*(2), 19–23.

Asmus, E. (1994). Motivation in music teaching and learning. *Quarterly Journal of Music Teaching and Learning, 5*(4), 5–32.

Austin, J. (1991). Competitive and non-competitive goal structures: An analysis of motivation and achievement among elementary band students. *Psychology of Music, 19*(2), 142–158.

Austin, J. R., & Vispoel, W. P. (1998). How American adolescents interpret success and failure in classroom music: Relationships among attributional beliefs, self-concept and achievement. *Psychology of Music, 26*(1), 26–45.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Cooke, M., & Morris, R. (1996). Making music in Great Britain. *Journal of the Market Research Society, 28*(2), 123–134.

Covington, M. V., & Omelich, C. L. (1979). Effort: The double-edged sword in school achievement. *Journal of Educational Psychology, 71*(2), 169–182.

Csikszentmihalyi, M. (1990). *Flow.* New York: Harper & Row.

Csikszentmihalyi, M., Rathunde, K., & Whalen, S. (1993). *Talented teenagers: The roots of success and failure.* Cambridge: Cambridge University Press.

Davidson, J. W., & Scutt, S. (1999). Instrumental learning with exams in mind: A case study investigating teacher, student and parent interactions before, during and after a music examination. *British Journal of Music Education, 16*(1), 79–95.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist, 41*, 1040–1048.

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality and development.* Philadelphia: Psychology Press.

Dweck, C. S., & Leggett, E. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256–273.

Eccles, J. S. (1987). Gender roles and women's achievement related decisions. *Psychology of Women Quarterly, 11*, 135–172.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J., & Midgley, C. (1983). Expectances, values and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75–146). San Francisco, CA: Freeman.

Eccles, J. S., & Harold, R. D. (1992). Gender differences in educational and occupational patterns among the gifted. In N. Colangelo, S. G. Assouline, and D. L. Amronson (Eds.), *Talent development: Proceedings from the 1991 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development* (pp. 3–29). Unionville, NY: Trillium Press.

Eccles, J. S., & Wigfield, A. (1995). In the mind of the achiever: The structure of adolescents' academic achievement related beliefs and self-perceptions. *Personality and Social Psychology Bulletin, 21*, 215–225.

Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self and task perceptions during elementary school. *Child Development, 64*, 830–847.

Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In W. Damon (Series Ed.) and N. Eisenberg (Vol. Ed.), *Handbook of child psychology*: Vol. 3. *Social, emotional and personality development* (5th ed., pp. 1017–1095). New York: Wiley.

Hackett, G. (1995). Self-efficacy in career choice and development.. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 232–258). New York: Cambridge University Press.

Henderson, V. L., & Dweck, C. S. (1990). Motivation and achievement. In S. S. Feldman and G. R. Elliott (Eds.), *At the threshold: The developing adolescent* (pp. 308–329). Cambridge, MA: Harvard University Press.

Massie, D. L. (1992). Band olympics: Musical muscle. *Music Educators Journal, 79*(4), 48–49.

McPherson, G. E. (2000). Commitment and practice: Key ingredients for achievement during the early stages of learning a musical instrument. *Proceedings of the XXIV International Society for Music Education Research Commission, Salt Lake City, USA, July 10–15, 2000.* (To be published in a forthcoming issue of *Bulletin of the Council for Research in Music Education.*)

McPherson, G. E., & Davidson, J. W. (in press). Musical practice: Mother and child interactions during the first nine months of learning a musical instrument. *Music Education Research.*

McPherson, G. E., & McCormick, J. (1999). Motivational and self-regulated learning components of musical practice. *Bulletin of the Council for Research in Music Education, 141,* 98–102.

McPherson, G. E., & McCormick, J. (2000). The contribution of motivational factors to instrumental performance in a music examination. *Research Studies in Music Education, 15,* 31–39.

Nicholls, J. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review, 91,* 328–346.

Nicholls, J. (1990). What is ability and why are we mindful of it? A developmental perspective. In R. J. Sternberg and J. Kolligian (Eds.), *Competence considered* (pp. 11–40). New Haven, CT: Yale University Press.

O'Neill, S. A. (1997). The role of practice in children's early musical performance achievement. In H. Jørgensen and A. C. Lehmann (Eds.), *Does practice make perfect? Current theory and research on instrumental practice* (pp. 53–70). Oslo: Norges Musikhøgskole.

O'Neill, S. A. (1999a). Flow theory and the development of musical performance skills. *Bulletin of the Council for Research in Music Education, 141,* 129–134.

O'Neill, S. A. (1999b). The role of motivation in the practice and achievement of young musicians. In S. W. Yi (Ed.), *Music, mind and science* (pp. 420–433). Seoul: Seoul National University Press.

Pajares, F. (1996). Self-efficacy beliefs and mathematical problem-solving of gifted students. *Contemporary Educational Psychology, 21,* 325–344.

Pintrich, P. R., & Schunk, D. H. (1996). *Motivation in education: Theory, research and applications.* Englewood Cliffs, NJ: Prentice-Hall.

Renwick, J., & McPherson, G. E. (2000). "I've got to do my scale first!": A case study of a novice's clarinet practice. In C. Woods, G. B. Luck, R. Brochard, F. Seddon, and J. A. Sloboda (Eds.), *Proceedings of the Sixth International Conference on Music Perception and Cognition.* Keele, UK: Keele University, Department of Psychology. CD-ROM.

Schunk, D. H., & Zimmerman, B. J. (1996). Modeling and self-efficacy influences on children's development of self-regulation. In J. Juvonen and K. R. Wentzel (Eds.), *Social motivation* (pp. 154–180). Cambridge: Cambridge University Press.

Sloboda, J. A., & Davidson, J. W. (1996). The young performing musician. In I. Deliège and J. A. Sloboda (Eds.), *Musical beginnings: The origins and development of musical competence* (pp. 171–190). Oxford: Oxford University Press.

Stipek, D. J. (1996). Motivation and instruction. In D. C. Berliner and R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 85–113). New York: Macmillan.

Stipek, D. J. (1998). *Motivation to learn: From theory to practice* (3rd ed.). Boston: Allyn & Bacon.

Vispoel, W. P., & Austin, J. R. (1993). Constructive response to failure in music: The role of attribution feedback and classroom goal structure. *British Journal of Educational Psychology, 63*, 110–129.

Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer.

Weiner, B. (1992). *Human motivation: Metaphors, theories, and research*. Newbury Park, CA: Sage.

Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review, 6*, 49–78.

Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review, 12*, 265–310.

Wigfield, A., Eccles, J. S., Suk Yoon, K., Harold, R. D., Arbreton, A. J. A., Freedman-Doan, C., & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology, 89*, 451–469.

Wigfield, A., O'Neill, S. A., & Eccles, J. S. (1999, April). *Children's achievement values in different domains: Developmental and cultural differences*. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Albuquerque, NM.

Yoon, K. S. (1997, April). *Exploring children's motivation for instrumental music*. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Washington, DC.

# 4

# Performance Anxiety

GLENN D. WILSON & DAVID ROLAND

Performance anxiety is a common problem among both amateur and professional musicians. It afflicts individuals who are generally prone to anxiety, particularly in situations of high public exposure and competitive scrutiny, and so is best understood as a form of social phobia (a fear of humiliation). Some degree of tension adds electricity to a performance, but pessimistic self-talk and feelings of panic can seriously affect it. The most effective psychological treatments seem to be those that combine relaxation training with anxiety inoculation (developing realistic expectations of what will be felt during performance) and cognitive restructuring (modifying habitual thoughts and attitudes that are self-handicapping, regardless of their origins). Preliminary research with hypnotherapy and the Alexander Technique suggests that these might also be effective in reducing performance anxiety.

Performance anxiety, sometimes called stage fright, is an exaggerated, often incapacitating, fear of performing in public. As in any other kind of phobia, the symptoms are those produced by activation of the body's emergency system, the sympathetic branch of the autonomic nervous system, including all the well-known effects of increases of adrenaline in the bloodstream (Fredrikson & Gunnarsson, 1992). The changes observed would have an adaptive function in relation to a physical threat, preparing us for an athletic response (fighting or fleeing). Unfortunately, running from or attacking an audience is seldom appropriate and the aftereffects of the alarm system can interfere with musical performance. For example, the increased heart pumping intended to supply additional oxygen to the muscles is felt as distressing palpitations. The increased activity of the lungs and widening of airways produces a feeling of breathlessness. The sharpening of vision has an aftermath in visual disturbances such as blurring. The diversion of resources away from digestion produces "butterflies" in the stomach. The redirection of body fluids such as saliva into the bloodstream pro-

duces a dry mouth sensation, and the activation of the body's cooling system produces sweaty palms and forehead.

These alarm reactions would enhance our survival if we were confronted with a tiger in a jungle, but they are less useful when we are faced with an audience who is expecting us to entertain them. Of course, we do not really fear physical attack if our performance is substandard, but human pride is apparently such a powerful motive that the fear of humiliation or disgrace can produce a similar degree of emotional panic. This is bound to be counterproductive to finely tuned vocal or instrumental performance, which requires a clear head and steady hand and voice.

### *Incidence*

The incidence of performance anxiety is considerable. Wesner, Noyes and Davis (1990) found 21% of the students and faculty at an American University School of Music reported "marked distress" that arose out of anxiety, while another 40% experienced "moderate distress." "Marked impairment" of performance was reported by 17% and "moderate impairment" by another 30%. The most troublesome symptoms were specified as poor concentration (63% of those claiming impairment), rapid heart rate (57%), trembling (46%), dry mouth (43%), sweating (43%), and shortness of breath (40%). Other commonly reported symptoms were flushing, quavering voice, nausea, and dizziness. Nine percent said they often avoided performance opportunities because of anxiety, and 13% had interrupted an actual performance on at least one occasion.

Many other studies support such high incidences. Van Kemanade, Van Son, and Van Heesch (1995) polled all the members of professional orchestras in the Netherlands and found that 91 out of 155 respondents, 59%, had been affected professionally or personally by stage fright and 10% suffered anticipatory anxiety for weeks before significant performances. In a comparative study of different types of performers, Marchant-Haycox and Wilson (1992) found musicians to be most affected by anxiety (47%), followed by singers (38%), dancers (35%), and actors (33%). Performance anxiety is clearly a problem of some magnitude, and it is not restricted to the inexperienced amateur.

### *Associated Factors*

Given that performance anxiety represents a fear of negative evaluation by others, it is not surprising that it relates closely to other forms of social phobia (Steptoe & Fidler, 1987; Cox & Kenardy, 1993). To some extent it also relates to personality traits that create anxiety. Perfectionism relates to having an unrealistically high expectation of yourself and others. It also relates to an overconcern with small flaws and mistakes, with a tendency to focus on what's wrong and discount what's right (Bourne, 1995). Perfectionists tend to be very self-critical and as a consequence suffer low self-esteem. Another personality trait leading to anxiety is the need to have excessive personal control (Bourne, 1995). This tendency leads the person to being uncomfortable and feeling unable to succeed in unpredictable circumstances.

Mor, Day and Flett (1995) have studied the interaction of perfectionism, personal control, and performance anxiety in professional musicians, actors, and dancers. They distinguished between self-oriented perfectionism (self-imposed high standards) and socially prescribed perfectionism (high standards imposed by others). Their findings suggest that high personal and social standards together with low personal control were most strongly associated with debilitating performance anxiety. Socially prescribed perfectionism was more strongly associated with debilitating performance anxiety than self-oriented perfectionism. Those performers who exhibited higher levels of perfectionism overall and lower personal control also experienced less satisfaction with their performances. Mor et al. (1995) conclude that cognitive-behavioral interventions designed to reduce perfectionist attitudes and to improve a sense of personal control may be effective in reducing severe performance anxiety.

There is some reduction of performance anxiety with age and experience, but many professional musicians battle with it throughout their career. Evidence that concerns sex differences is equivocal; some researchers (e.g., Wesner et al., 1990) find a female prevalence in performance anxiety, consistent with higher levels of anxiety and phobias in general, while others (e.g., Van Kemanade, Van Son & Van Heesch, 1995) find an equal incidence in males and females.

In their discussion of social anxiety Beck and Emery (1985) state that the anxious person's perception of a threat to him or her is the trigger for their anxiety response. This perception of a threatening event is created by:

1. Overestimating the probability of a feared event
2. Overestimating the severity of the feared event
3. Underestimating coping resources (what you can do about it)
4. Underestimating rescue factors (what other people can do to help you)

Any situation that increases the performer's sense of threat will increase his or her level of performance anxiety (Wilson, 2002; Roland, 1994b; Hamann, 1982; Abel & Larkin, 1990; Cox & Kenardy 1993). For this reason solo performance is usually much more stressful than performing a duet, which is in turn more stressful than playing in a trio, and so on; the relationship between stage fright and number of coperformers is that of a steeply declining curve. (Beyond a certain number of coperformers, adding more makes little difference.) Public performance is more anxiety-evoking than private performance, though the degree of anxiety does not have such a simple relationship to the size of the audience as it does to the number of coperformers. While it might be supposed that performing on TV to an audience of 10 million would be far more frightening than performing in a small auditorium, this is not necessarily the case. The actual proximity of an audience (e.g., being able to see the expressions on their faces) may be more significant than their numbers, and the status relationships between performer and audience are critical.

Nearly all performers report that auditions are the most stressful performance situations because they combine scrutiny and evaluation on the one hand with a socially inferior situation on the other. The evaluation of the audition panel will have a direct effect upon career progress. Other competitive (judged) situa-

tions are also more stressful than performances given for entertainment, again depending on the relative status of performer and listener. Competitors for a top international prize are typically less susceptible than students being graded at early levels of music education. There is also an interaction between the performer's personality and the effect that an audience will have on him or her; those most susceptible to stage fright (people who are inclined to be anxious and introverted—i.e., shy and withdrawn) tend to perform worse under scrutiny, whereas the presence of an audience may actually extract the best from those who are more sociable and confident, stable extroverts (Graydon & Murphy, 1995).

## Optimal Arousal

A certain degree of arousal actually helps the quality of performance as judged by an audience, and this applies to a level that is experienced by the performer as unpleasant, stressful, and presumed to be detrimental (Gellrich, 1991; Wilson, 1994). In other words, performers themselves are not always good judges as to how anxiety is influencing their performances and awareness of this fact may itself be helpful.

It is widely accepted that the quality of performance is related to arousal with an inverted-U (rainbow-shaped) curve. That is, very low levels of arousal are insufficiently motivating and give rise to lackluster performances, while excessive arousal interferes with performance because concentration is disrupted, memory blocks occur, and there is a loss of steadiness in hands and voice. This rainbow-shaped function is called the Yerkes-Dodson Law after the researchers who first described it. This law further states that the peak of the curve will be reached earlier for difficult than for easy tasks. In other words, more complex tasks deteriorate more easily under stress than simple ones. Although this effect was first established in laboratory rats, there is ample evidence that it applies to human subjects and within the context of musical performance (Anderson, 1994; Hamann & Sobaje, 1983).

Wilson (2002) proposed a three-dimensional extension of the Yerkes-Dodson Law, grouping sources of stress into three major categories:

1. *Trait anxiety:* any personality characteristics, constitutional or learned, that mediate susceptibility to stress
2. *Situational stress:* environmental pressures such as public performance, audition, or competition
3. *Task mastery:* ranging from performances of simple, well-rehearsed works to those of complex, underprepared material

These three sources of stress vary independently; hence whether anxiety is beneficial or detrimental to performance depends upon their interplay. This model accounts for a great many research findings and has practical implications for the performer. For example, highly anxious individuals perform best when the work is well mastered and the situation relaxed, whereas low-anxiety individuals rise to a challenge and perform better with a more demanding audi-

ence. Anxiety-prone musicians should pick easy, familiar works to perform, especially for auditions or important public performances. However, hard preparation may turn a difficult work into a relatively easy one, thus reducing performance anxiety.

Hardy and Parfitt (1991) have suggested that a *catastrophe model* is more appropriate than the Yerkes-Dodson rainbow function in describing the sudden plunge in performance quality once arousal has passed a certain stress point. They note that excessive arousal can lead to a precipitous crash in performance rather than a gentle tailing off and small reductions in anxiety at this point are not capable of pulling the performer back to the optimal point on the curve. According to Hardy and Parfitt, it is necessary to distinguish cognitive (mental) anxiety from somatic (bodily) agitation. It is primarily the mental component that is likely to show the catastrophic decrement in performance.

This viewpoint is consistent with research that seeks to clarify what is going on in the heads of performers who are most susceptible to stage fright. Steptoe and Fidler (1987) identified various types of self-talk that were common among performers, and the type most associated with stage fright they called catastrophizing. This refers to an imaginary or anticipated disastrous outcome (e.g., "I think I am going to faint" or "I'm almost sure to make a dreadful mistake and that will ruin everything"). The most healthy cognitive strategy, called realistic self-appraisal, consisted of self-comments such as "I'm bound to make a few mistakes, but so does everyone" and "The audience wants me to play well and will make allowance for a few slips." The first type of statement would render a performer panic-prone, while the latter is generally optimistic and incorporates some inoculation against errors.

## Treatment Approaches

### *Drugs*

Many performers attempt self-treatment with anxiety-reducing drugs such as alcohol, Valium, and cannabis (Wills & Cooper, 1988). These may get a performer through a performance but are ultimately destructive in that they create dependence and the fine edge of the performance is lost. Two controlled studies of the effects of tranquilizers on performance have been reported: one found buspirone to be largely ineffective (Clark & Agras, 1991) and the other found benzodiazepines to be detrimental (James & Savage, 1984).

*Beta-adrenergic blockers*, which act specifically to inhibit peripheral autonomic symptoms, are more promising in that they supposedly leave the head clear while at the same time eliminating problems such as tremor and butterflies. Evidence is mixed as to whether they improve musical performance as judged by the outside observer. In any case, they seem less than ideal because of possible side effects such as loss of sexual potency, nausea, tiredness, and blunting of affect. In asthmatics the blockers are particularly dangerous and occasionally precipitate heart failure. Although many musicians take them regularly (some

estimates suggest about one-quarter), in the United States and some other countries, their use for performance anxiety is not sanctioned by medical authorities. They may be helpful in breaking strongly conditioned anxiety cycles and getting performers back onstage when they have been frightened off, but psychological procedures aimed at restoring self-control to the performer would seem preferable, because performers are then relying on their own resources.

### *Psychoanalysis*

Various psychoanalytic interpretations of performance anxiety have been offered (e.g., Nagel, 1993; Gabbard, 1979). These attribute the development of performance anxiety to early childhood experiences such as conflicts associated with genital exhibition and fear of parental punishment for masturbation. Such notions are difficult to verify, and although analysts sometimes provide descriptions of successfully treated case studies, these do not satisfy scientific criteria of evaluation. While early experiences may well contribute to performance anxiety, it is not clear that these need to be revisited as part of the treatment procedure; methods that focus on the here and now (notably the behavioral and cognitive methods to be discussed next) show more evidence of effectiveness.

### *Behavioral Therapy*

Some standard behavioral approaches for the treatment of phobias have been applied to musical performance anxiety. The best known is *systematic desensitization*, which involves training in muscular relaxation followed by having the client imagine increasingly sensitive scenes that relate to the conditions that typically evoke anxiety (progressive exposure to a *fear hierarchy*). For example, an anxious pianist may begin by imagining playing an easy piece to relatives on a friendly family occasion. Once comfortable with that image, the pianist can move on to a scenario that includes some strangers, and so on, until a major concert is imagined. The theory is that phobias are maintained by the relief one obtains by avoiding the phobic object; hence what is necessary is a method for persuading the client to encounter that object (in this case a critical audience) in gradual, stepwise stages. Some studies suggest that classic desensitization does help with performance anxiety and speech anxiety (Wardle, 1975; Appel, 1976; Allen, Hunter & Donahue, 1989). However, exposure by itself does not necessarily extinguish performance anxiety, since many musicians perform for years without ever conquering their fears spontaneously (Steptoe & Fidler, 1987; Wesner et al., 1990).

### *Cognitive-Behavioral Therapy*

Targeting a phobic person's behavior is sometimes insufficient for treating the underlying anxiety because the way a person thinks about his or her situation is critical to the onset of the anxiety. Since negative self-talk often mediates per-

formance anxiety (Lloyd-Elliott, 1991; Steptoe & Fidler, 1987), it follows that some form of cognitive restructuring, such as reorganizing the individual's habitual ways of thinking during performance, might be helpful.

An important target for most cognitive procedures is the performer's focus of attention. Kendrick et al. (1982) showed the effectiveness of a procedure called attention training with a group of 53 pianists who sought help for performance anxiety. This consisted of the identification of negative and task-irrelevant thoughts during piano playing and training in substituting for them optimistic, task-oriented self-talk. Improvement was compared against behavior rehearsal (playing repeatedly before friendly, supportive audiences) and a wait-list control. Both treatments were clearly superior to the wait-list control, but on some assessment criteria attention training was superior to behavior rehearsal. The authors admitted that their treatment included many elements, such as verbal persuasion, modeling, instruction, performance accomplishments, group influence, and homework assignments, but nevertheless felt that the key element to their treatment was the modification of maladaptive thoughts.

Subsequent research provides some backing for this conclusion. In a trial of cognitive therapy versus antianxiety drugs, Clark and Agras (1991) found that the beta-blocker buspirone did no better than placebo, but cognitive restructuring was clearly superior to both. Sweeney and Horan (1982) reported that both cognitive therapy and *cue-controlled relaxation training* were individually superior to a control condition that consisted of musical analysis training, but the combination of cognitive therapy and relaxation was best of all.

A particular form of cognitive restructuring that shows promise is that called *stress inoculation* (Meichenbaum, 1985; Salmon, 1991). The idea is that it is important to implant realistic expectations about what will be experienced during performance as well as promoting optimistic self-comments. Performers are taught to anticipate the symptoms of anxiety that are bound to arise before important public appearances and to befriend them, that is, reframe them as less threatening, even desirable, reactions. For example, the adrenaline effects (e.g., pounding heart, faster breathing) are reappraised as normal emotional responses that are not conspicuous to an audience and can provide energy, thus contributing to a more lively, exciting musical performance.

In an interview study with successful professional performers, Roland (1994a) found that they viewed anxiety or nervousness before a performance as a normal and even a beneficial part of performance preparation. They described experiencing a hyped-up, excited feeling before performing, an increase in mental focus, and sometimes inspiration. Similarly, Hanton and Jones (1999) found that elite swimmers viewed precompetition anxiety as helpful to a far greater extent than nonelite swimmers. Clearly, perceiving preperformance anxiety as a normal and helpful aspect of performing can be an important cognitive strategy.

In an attempt to improve upon the standard cognitive-behavioral treatment for musicians with performance anxiety, Roland (1994a) developed some modifications. He compared two treatment groups with a wait-list control group. The

standard cognitive-behavioral treatment incorporated relaxation techniques (breathing awareness, progressive muscle relaxation, and mental suggestions and imagery to produce a relaxed state) and cognitive techniques (normalizing the experience of anxiety and developing positive self-talk). The modified cognitive-behavioral treatment group incorporated the same relaxation techniques, normalizing the experience of anxiety but without training in positive self-talk. It also included training in task-oriented thinking, setting performance goals, mental rehearsal, and developing a preperformance routine. The training sessions were conducted over four two-hour weekly sessions. The modified treatment showed no superiority over the standard procedure (though both were effective relative to controls).

### *Hypnotherapy*

If confident self-talk is the key to alleviating performance anxiety, then verbal suggestions delivered under hypnosis might also be an effective treatment. Stanton (1993) describes a two-session hypnotherapy procedure that combines success imagery with rational-emotive therapy (overriding crippling myths such as the idea that anything short of perfection is a disaster); three cases were apparently successfully treated in this way. In a more controlled study, Stanton (1994) paired music students according to their scores on a performance anxiety questionnaire and assigned one of each pair to hypnotherapy and control groups. Hypnotherapy consisted of two 50-minute sessions, one week apart, which included relaxation suggestions, induction of slow breathing, pleasant visual imagery (clouds and a lake), and verbal suggestions that linked these images to increased mental control. The control group met twice at the same interval for similar-length discussion sessions. The hypnotherapy group (but not controls) showed a significant reduction of performance anxiety immediately after treatment and further gains six months later.

Although hypnotherapy appeared effective in Stanton's studies, there are reasons that these results should be treated as preliminary. First, his procedure included relaxation and it may be that this would be equally effective without trance induction. Second, doing better than nothing is not a great achievement. What is needed is a comparison of the cost-effectiveness of hypnotherapy and that of standard practice (currently cognitive-behavior therapy).

### *Alexander Technique*

Although not specifically developed as a treatment for performance anxiety, the Alexander Technique (AT) deserves mention because it is widely used by musicians for this purpose. For example, Watson and Valentine (1987) found that more than half of British orchestral musicians used some form of complementary medicine for anxiety reduction and of these AT was the most common (43%).

Named after the Australian actor Fred Alexander, who devised it, AT is characterized as a form of kinesthetic reeducation that uses a mixture of verbal in-

structions and hands-on demonstration to correct postural misuse. AT has become very popular among performers and others, though it is only recently that scientific evaluation has been attempted.

Valentine et al. (1995) assigned 25 musicians to either 15 sessions of AT or a no-treatment control, assessing them at audition, in class before and after treatment, and at a final recital. Measures included physiological and self-report indices of anxiety and ratings of videotaped musical performance by judges who were blind to the treatment assignment. The AT group showed more improvement in musical skill, a more positive attitude to performance, and lower levels of anxiety than the controls. However, the benefits were confined to the low-stress (classroom) situation and did not transfer to recital. Furthermore, the improvements were not related to positive changes in postural habits as rated by the AT teachers, suggesting that, such as they were, they must have been due to some other mechanism. A likely candidate would be some kind of cognitive restructuring (e.g., distraction from anxiety cues or destructive self-talk). This preliminary research suggests that AT has some beneficial effect but perhaps not for the reasons its proponents believe. Again, a comparison of its efficacy with that of cognitive therapy would be useful.

## Implications and Strategies

The research described earlier suggests that a cognitive-behavioral approach is the treatment of choice for performance anxiety. There are a number of strategies that can be employed at a cognitive and behavioral level.

### *Cognitive Strategies*

Cognitive strategies include normalizing the experience of preperformance anxiety, increasing positive self-talk, mental rehearsal of the performance, and goal setting.

**Viewing Anxiety as Positive.** The first step is for performers to view the experience of some preperformance anxiety as a normal and even helpful part of performing. However, this view is not one that performers or athletes automatically or easily arrive at. Both Roland (1994b) and Hanton and Jones (1999) found that performers and athletes usually go through some unpleasant performance and competition experiences before arriving at a more positive view of performance anxiety. It can be concluded from this that there is no easy way for the beginning performer to become acclimatized to live performing other than by actually doing it repeatedly and in many different types of performance situations.

**Positive Self-Talk.** In the interview study with successful professional performers, Roland (1994b) found that 69% used positive self-talk prior to performing. The earlier literature review referred to studies that have attempted to train performers in the use of realistic, positive self-talk. In this approach, unrealistic

negative self-statements are challenged and analyzed for their veracity. Typically overanxious performers will catastrophize about their chances of failure. The aim is to help performers realistically appraise their perception of a performance and adopt more positive and helpful self-statements. Doing this reduces the sense of threat the performers feel and increases their sense of self-control. An example of this reappraisal process would be asking the musician, "What's the worst thing that can really happen to you in this performance?" Often the process of answering this question, assessing the actual likelihood and implications of its outcome, will be sufficient to provide the performer with a more realistic appraisal of the situation. This will in turn reduce the sense of threat and fear experienced by performers (see further Bourne; 1995, Seligman, 1995; Meichenbaum, 1985).

**Mental Rehearsal and Imagery.** Mental rehearsal is another strategy for preparing for performance (Roland, 1994a; Murphy & Jowdy, 1992). Hanton and Jones (1999) and Orlick (1990) found a similar use of mental rehearsal strategies in elite athletes. Taylor and Taylor (1995) have developed imagery techniques for dancers. Mental rehearsal requires performers to imagine as vividly as they can, going through their performance in the ideal way they would like it to go. In imagining this they can draw on all senses—sound, sight, touch, taste, smell, and kinesthetic. They can include in their rehearsal the positive self-talk they wish to experience. By doing this performers can imagine themselves watching their own performance as an audience member might (external focus) or watching their performance as if they were actually performing (internal focus). The effect of mental rehearsal appears to be that it provides a form of neuromuscular programming so that the performer is more likely to automatically behave in the preferred way during the actual performance (Roland, 1997). This process is similar in nature to self-hypnosis.

**Goal Setting.** Goal setting has become a standard cognitive strategy in sports and in work environments but appears to have been little applied in musical settings (Burton, 1992; Roland, 1997). Research in these settings indicates that working out short- and long-term goals improves the quality of performance of the goal setter. Goals can be broken into process goals and outcome goals. Process goals include the aspects of the performance that the performer wishes to achieve to create a good performance; for example, these could relate to technique or the interpretation of a piece. Outcome goals relate to more observable achievements such as securing a position in an orchestra and learning a certain repertoire. In the case of children, the type of goals parents set for their children can be a major influence on the child's attitude to performing. On the one hand, parents who set learning- or process-oriented goals are more likely to encourage their children to enjoy the process of performing. On the other hand, parents who set performance outcome goals are more likely to encourage perfectionist tendencies in their children, thus encouraging performance anxiety (Ablard & Parker, 1997). This suggests that it is healthier to place an emphasis on process-oriented goals most of the time.

### *Behavioral Strategies*

Behavioral strategies include relaxation, adopting a preperformance routine, following an anxiety hierarchy, and adopting supportive lifestyle habits.

**Relaxation.** Training in relaxation skills has become a standard approach in anxiety management in general (e.g. Bourne, 1995) and in the treatment of performance anxiety in particular (Nagel, Himle, & Papsdorf, 1989). The regular practice of relaxation helps to reduce the physiological response to stress, prevent the cumulative effect of stress, improve memory and concentration, increase energy and productivity levels, and reduce muscle tension (Bourne, 1995). Roland (1994b) found that 71% of the professional performers he interviewed used breathing relaxation prior to performing and some performers used other forms of relaxation activity prior to performing. Hanton and Jones (1999) found a high use of precompetition relaxation strategies in elite swimmers.

There are two main ways in which relaxation strategies can be applied. One is on a regular basis, while the other is as part of a preperformance routine. Roland (1997) has suggested that performers need to practice relaxation training outside of performance conditions before they can successfully practice it prior to a performance. Relaxation strategies can include progressive muscle relaxation, mental suggestions and imagery to produce a relaxed state, meditation, breathing awareness, yoga, tai chi, stretching, and AT.

**Preperformance Routines.** A preperformance routine appears to be an important element in helping performers to achieve an optimal mental and physical state for their performance (Roland, 1994b). A preperformance routine might include a warm-up on the instrument, the use of positive self-talk, a focus on performance goals, the use of a relaxation strategy, controlling the type and amount of interaction with others, a nap earlier in the day, and monitoring food and fluid intake. Ultimately, all performers need to work out through experimentation a standard routine that helps them to achieve a state of readiness for performing.

**Anxiety Hierarchies.** For performers who have had a bad experience of performing that has greatly undermined their confidence or for the new performers who are just starting out, the use of an anxiety hierarchy can be very useful. The performer is helped to work out a hierarchy of performance situations that range from the situation that induces the least amount of anxiety to the situation that induces extreme anxiety. These can be rated from 0 (no anxiety) to 100 (extreme anxiety). Performers then commence carrying out the performances that have a low rating until they become comfortable. They then work progressively up the hierarchy, becoming comfortable with each stage before moving on to the next. This structured approach provides an indication of the stage of performance the performer is ready for; thus the likelihood of the performer being overwhelmed by a performance situation is greatly reduced. As performers progress through the hierarchy they practice their anxiety management skills such as relaxation and positive self-talk so that these become an integral part of the approach.

**Supportive Lifestyle.** There are standard lifestyle habits that are valuable for general stress management. These include regular physical exercise, adequate sleep, and a healthy diet. Regular exercise is likely to help prevent the buildup of stress-related chemicals in the body. Avoiding high-caffeine, high-sugar, and spicy foods prior to performing and instead eating easily digestible complex carbohydrates, fruits, and vegetables is likely to produce a sustained release of energy during the performance. This will reduce the chance of fluctuations in concentration (Stanton, 1988; Robson, Davidson, & Snell, 1995). Other factors sometimes outside of the performers' control can affect their response to the stress of performing. These can include family, financial, and health difficulties. A first step is acknowledgment by the performers of the effect these may be having on their performance.

**Flow State.** As a final consideration in examining performance anxiety, performers can aim for the experience of *flow* in performing. Flow is the state in which performers perform to their optimum, the experience of which is usually described by the performer as exhilarating (Csikszentmihalyi, 1992; see chapter 3). Csikszentmihalyi and other researchers such as Jackson (1995) have attempted to determine the conditions that are likely to precipitate the flow experience. It is these conditions that music educators and performers can be mindful of in their approach to performance. One of the critical conditions for flow is that there needs to be a skill-challenge balance such that the performers are being challenged during practice or performance but not too far beyond their skills at that time. A second condition is that performers need to have a clear goal or purpose. A third condition for the repeated occurrence of flow experiences is that the performers look for feedback before, during, and after performing. This allows performers to assess their progress and to make corrective adjustments to their future performances.

## Conclusion

This chapter has provided an overview of the incidence and nature of performance anxiety in music, an issue that is clearly of importance to every musician and music educator. In our opinion, the best approach for the treatment of debilitating performance anxiety is cognitive-behavior therapy, although drug therapy, clinical hypnosis, and the Alexander Technique also appear to be beneficial to some degree. A key implication of our review is that there are a number of positive strategies that can be employed by every performer, no matter how experienced or inexperienced he or she may be, in order to overcome the types of anxiety problems that all musicians face as they learn how to perform confidently.

### *References*

Abel, J. L., & Larkin, K. T. (1990). Anticipation of performance among musicians: Physiological arousal, confidence and state anxiety. *Psychology of Music, 18*(2), 171–182.

Ablard, K., & Parker, W. (1997). Parents' achievement goals and perfectionism in their academically talented children. *Journal of Youth and Adolescence, 26*(6), 651–667.

Allen, M., Hunter, J. E., & Donahue, W. A. (1989). Meta-analysis of self-report data on the effectiveness of public speaking anxiety treatment techniques. *Communication Education, 38*, 54–76.

Anderson, K. J. (1994). Impulsivity, caffeine and task difficulty: A within subjects test of the Yerkes-Dodson Law. *Personality and Individual Differences, 16*, 813–829.

Appel, S. S. (1976). Modifying solo performance anxiety in adult pianists. *Journal of Music Therapy, 13*, 2–16.

Beck, A., & Emery, G. (1985). *Anxiety disorders and phobias: A cognitive perspective*. New York: Basic Books.

Bourne, E. (1995). *The anxiety and phobia workbook*. Oakland, CA: New Harbinger.

Burton, D. (1992). The Jekyll/Hyde nature of goals: Reconceptualising goal setting in sport. In T. Horn (Ed.), *Advances in sport psychology* (pp. 221–250). Champaign, IL: Human Kinetic.

Clark, D. B., & Agras, W. S. (1991). The assessment and treatment of performance anxiety in musicians. *American Journal of Psychiatry, 148*, 598–605.

Cox, W. J., & Kenardy, J. (1993). Performance anxiety, social phobia and setting effects in instrumental music students. *Journal of Anxiety Disorders, 7*, 49–60.

Csikszentmihalyi, M. (1992). *Flow: The psychology of happiness*. London: Rider.

Fredrikson, M., & Gunnarsson, R. (1992). Psychobiology of stage fright. *Biological Psychology, 33*, 51–61.

Gabbard, G. O. (1979). Stage fright. *International Journal of Psychoanalysis, 60*, 383–392.

Gellrich, M. (1991). Concentration and tension. *British Journal of Music Education, 8*(2), 167–179.

Graydon, J., & Murphy, T. (1995). The effect of personality on social facilitation whilst performing a sports related task. *Personality and Individual Differences, 19*, 265–267.

Hamann, D. L. (1982). An assessment of anxiety in instrumental and vocal performers. *Journal of Research in Music Education, 30*(2), 77–90.

Hamann, D. L., & Sobaje, M. (1983). Anxiety and the college musician: A study of performance conditions and subject variables. *Psychology of Music, 11*(1), 37–50.

Hanton, S., & Jones, G. (1999). The acquisition and development of cognitive skills and strategies: 1. Making the butterflies fly in formation. *Sport Psychologist, 13*, 1–21.

Hanton, S., & Jones, G. (1999). The effects of a multimodal intervention program on performers: 11. Making the butterflies fly in formation. *Sport Psychologist, 13*, 22–41.

Hardy, L., & Parfitt, G. (1991). A catastrophe model of anxiety and performance. *British Journal of Psychology, 82*(2), 163–178.

Jackson, S. (1995). Factors influencing the occurrence of flow state in elite athletes. *Journal of Applied Sport Psychology, 7*, 138–166.

James, I., & Savage, I. (1984). Beneficial effects of nadolol on anxiety-induced disturbances of performance in musicians. *American Heart Journal, 108*(4.2), 1150–1155.

Kendrick, M. J., Craig, K. D., Lawson, D. M., & Davidson, P. O. (1982). Cognitive and behavior therapy for musical performance anxiety. *Journal of Consulting and Clinical Psychology, 50*, 353–362.

Lloyd-Elliot, M. (1991). Witches, demons and devils: The enemies of auditions and how performing artists make friends with these saboteurs. In G. D. Wilson (Ed.), *Psychology and performing arts* (pp. 211–218). Amsterdam: Swets & Zeitlinger.

Marchant-Haycox, S. E., & Wilson, G. D. (1992). Personality and stress in performing artists. *Personality and Individual Differences, 13*, 1061–1068.

Meichenbaum, D. (1985). *Stress inoculation training.* New York: Pergamon.

Mor, S., Day, H., & Flett, G. (1995). Perfectionism, control, and components of performance anxiety in professional artists. *Cognitive Therapy and Research, 19*(2), 207–225.

Murphy, S., & Jowdy, D. (1992). Imagery and mental practice. In T. Horn (Ed), *Advances in sport psychology* (pp. 221–250). Champaign, IL: Human Kinetic.

Nagel, J., Himle, D., & Papsdorf, J. (1989). Cognitive-behavioural treatment of musical performance anxiety. *Psychology of Music, 17*(1), 12–21.

Nagel, J. J. (1993). Stage fright in musicians: A psychodynamic perspective. *Bulletin of the Menninger Clinic, 57*, 492–503.

Orlick, T. (1990). *In pursuit of excellence: How to win in sport and life through mental training* (2nd ed.). Champaign, IL: Leisure.

Robson, B., Davidson, J., & Snell, E. (1995). "But I'm not ready yet." *Medical Problems of Performing Artists, 10*(1), 32–37.

Roland, D. (1994a). The development and evaluation of a modified cognitive-behavioural treatment for musical performance anxiety (Doctoral dissertation, University of Wollongong, Australia). *Dissertion Abstracts International* (University Microfilms No. 9424349), UMI, Ann Arbor, MI.

Roland, D. (1994b). How professional performers manage performance anxiety. *Research Studies in Music Education, 2*, 25–35.

Roland, D. (1997). *The confident performer.* Sydney: Currency.

Salmon, P. (1991). Stress inoculation techniques and musical performance anxiety. In G. D. Wilson (Ed.), *Psychology and performing arts* (pp. 219–229). Amsterdam: Swets & Zeitlinger.

Seligman, M. (1995). *The optimistic child.* Sydney: Random House.

Stanton, H. E. (1993). Alleviation of performance anxiety through hypnotherapy. *Psychology of Music, 21*(1), 78–82.

Stanton, H. E. (1994). Reduction of performance anxiety in music students. *Australian Psychologist, 29*, 124–127.

Stanton, R. (1988). *Eating for peak performance.* Sydney: Allen & Unwin.

Steptoe, A., & Fidler, H. (1987). Stage fright in orchestral musicians: A study of cognitive and behavioral strategies in performance anxiety. *British Journal of Psychology, 78*(2), 241–249.

Sweeney, G. A., & Horan, J. J. (1982). Separate and combined effects of cue-controlled relaxation and cognitive restructuring in the treatment of musical performance anxiety. *Journal of Counselling Psychology, 29*(5), 486–497.

Taylor, J., & Taylor, C. (1995). *Psychology of dance.* Champaign, IL: Human Kinetic.

Valentine, E. R., Fitzgerald, D. F. P., Gorton, T. L., Hudson, J. A., & Symonds, E. R. C. (1995). The effect of lessons in the Alexander Technique on music performance in high and low stress situations. *Psychology of Music, 23*(2), 129–141.

Van Kemanade, J. F., Van Son, M. J., & Van Heesch, N. C. (1995). Performance anxiety among professional musicians in symphonic orchestras: A self-report study. *Psychological Reports, 77,* 555–562.

Wardle, A. (1975). Behavior modification by reciprocal inhibition of instrumental music performance anxiety. In C. K. Madsen, R. D. Greer, and C. H. Madsen Jr. (Eds.), *Research in music behavior: Modifying basic behavior in the classroom* (pp. 191–205). New York: Teachers College.

Watson, P., & Valentine, E. (1987). The practice of complementary medicine and anxiety levels in a population of musicians. *Journal of the International Society for the Study of Tensions in Performance (ISSTIP), 4,* 26–30.

Wesner, R. B., Noyes, R., & Davis, T. L. (1990). The occurrence of performance anxiety among musicians. *Journal of Affective Disorders, 18,* 177–185.

Wills, G. D., & Cooper, C. L. (1988). *Pressure sensitive: Popular musicians under stress.* London: Sage.

Wilson, G. D. (2002). *Psychology for performing artists* (2nd ed.). London: Whurr.

# 5

# Brain Mechanisms

ECKART ALTENMÜLLER & WILFRIED GRUHN

Neurological foundations of music perception, performance, and learning rely on individually variable, widely distributed neuronal networks in both hemispheres. Music performance is a complex voluntary sensorimotor behavior that becomes automated during extensive practice with auditory feedback. It involves all motor, somatosensory, and auditory areas of the brain. Practicing a musical instrument results first in a temporary and later in a stable increase in the amount of nerve tissue devoted to the various component tasks. Overuse of movement patterns may degrade motor memory and voluntary control of movements (musician's cramp). Neuronal networks established during music learning may depend on teaching strategies. Brain regions corresponding to specific subtasks of music performance are larger in musicians with early training, which may account for their superior capacity to acquire complex musical sensory-motor and auditory skills.

Music performance at a professional level is one of the most demanding tasks for the human central nervous system. It involves the precise execution of very fast and, in many instances, extremely complex physical movements under continuous auditory feedback. A further aspect of music performance—although not specifically addressed in this chapter—is the involvement of emotional experiences.

Extensive practice is required to develop new skills and carry out these complex tasks. Motor skills, on the one hand, can only be automated by countless repetitions; aural skills, on the other hand, are developed through a broad variety of listening experiences. These skills are not represented in isolated brain areas but rather depend on the multiple connections and interactions established during training within and among the different regions of the brain. The general ability of our central nervous system to adapt to both changing environmental conditions and newly imposed tasks during its entire life span is referred to as *plasticity:* in music, learning through experience and training is accompanied

by development and changes that not only take place in the brain's neuronal networks, for example as a strengthening of neurons' connections, but also occur in its overall gross structure.

The aim of this chapter is to give a short review of current knowledge about the brain mechanisms involved in music perception, music production, and music learning. We believe that a basic understanding of the enormously complex neurobiological processes that underlie the musician's training and performance will eventually stimulate new insights into the practice and theory of music education. So far, the results of laboratory experiments have been, by necessity, restricted to very limited aspects of music making. Consequently, the brain mechanisms that underlie the rich universe of accomplished musicianship are mostly still inaccessible to brain research.

To understand neural substrates of music performance it is first necessary to understand some basic neuroanatomy and neurophysiology. In the next section we introduce essential general information for musical readers.

## Neurophysiological Background

### *Brain-Imaging Methods*

During the last decade, rapid improvements in brain-imaging methods have enabled researchers to carry out substantial new investigations into the biological foundations of music performance. The general term *functional brain imaging* covers the various methods of objectively monitoring neuronal activity during music perception, musical reasoning, and music production. These methods allow documentation of the dynamics of developing brain circuitry during the acquisition of new mental representations. They show our brain at work.

There are two principal approaches that can be used to assess brain activity. The first of these involves methods such as *electroencephalography* (*EEG*) and *magnetoencephalography* (*MEG*), which make it possible to directly measure the electrical activity of the neurons in the cerebral cortex. The second enables the assessment of brain metabolism, cerebral blood flow, and oxygen consumption of nerve cells and allows for indirect analysis of neuronal activity based on the close links between oxygen consumption and the firing activities of nerve cells. The need for oxygen is reflected in local increases of blood flow in the nervous tissue, which can, in turn, be assessed by measuring the local concentration of radioactively labeled marker substances in the blood such as oxygen or glucose. While *positron emission tomography* (*PET*) uses this kind of data collection, certain inconveniences arise with the use of this method due to the application of low dosages of radioactivity. Much safer is the method of *functional magnetic resonance imaging* (*fMRI*), which uses the magnetic properties of oxygen in blood cells in order to calculate oxygen consumption.

The main advantages of EEG and MEG are their excellent temporal resolution, which enables the monitoring of rapid processes and changes. These methods also allow communication flows among different areas of the brain to be ana-

lyzed through the calculation of so-called *coherence* between activated neuronal cell assemblies. In contrast, the main advantages of PET and fMRI are their excellent *spatial resolution*, which allows particular tasks to be related to specific brain structures. Nevertheless, the temporal resolutions of PET and fMRI are still relatively poor (ranging from six seconds to one minute), meaning that more rapid cognitive processes cannot be tracked. A further factor is cost: MEG, PET, and fMRI rely on extremely complex and expensive technologies. Only EEG is affordable enough to be used outside specialized brain-imaging centers.

### *General Structure of the Brain*

The human brain can be subdivided into three parts: the *hindbrain*, the *midbrain*, and the *forebrain* (see Figure 5.1). The hindbrain consists of the *medulla*, *pons*, and *cerebellum*. The hindbrain and the midbrain together constitute the *brain stem*, which is phylogenetically (i.e., in terms of evolution) the oldest part of the brain. The brain stem regulates all vital functions such as breathing, heartbeat, arousal, body temperature, and equilibrium. Furthermore, the brain stem controls many sensory and motor functions such as eye movements and the coordination of visual and auditory reflexes. The cerebellum lies behind the pons and mainly processes body equilibrium and the accurate timing of movements. It is involved in the learning of motor skills and is particularly relevant to the learning skills required in musical performance. The midbrain lies above the pons and contains two structures. The first of these—the *thalamus*—transmits incoming information from all sensory systems to the cerebral cortex and so acts as a gateway to the cortex. The latter—the *hypothalamus*—regulates *autonomic* and *endocrine* functions. Finally, the forebrain consists of the two outer cerebral hemispheres and three deep-lying structures: the *basal ganglia*, the *hippocampus*, and the *amygdaloid nucleus*. The last two structures lie at the inner border of the temporal lobe and are not shown in Figure 5.1. The basal ganglia participate in regulating motor performance, the hippocampus is involved with aspects of memory storage, and the amygdaloid nucleus coordinates autonomic and endocrine responses in conjunction with emotional states.

All cognitive functions are governed by the cerebral cortex—the outer part of the brain—which is the most complex organ in the human body. According to recent estimates, the cerebral cortex consists of about 100 billion neurons, which are interconnected by a dense web of nerve fibers. Each nerve cell can communicate with approximately ten thousand other cells. Small, prominent buttons (*synapses*) in the nerve fibers form these connections. The cerebral cortex is divided into two hemispheres that are interconnected by a large fiber bundle that contains about 100 million fibers. This is known as the *corpus callosum*.

Four important features characterize the organization of the cortex. First of all, each hemisphere is concerned primarily with sensory and motor processes on the opposite (*contralateral*) side of the body. Second, although appearing to be similar, these hemispheres are neither completely symmetrical in structure nor equivalent in function. Third, the cortex is hierarchically organized, with distinct primary, secondary, and tertiary (or associative) sensory or motor re-

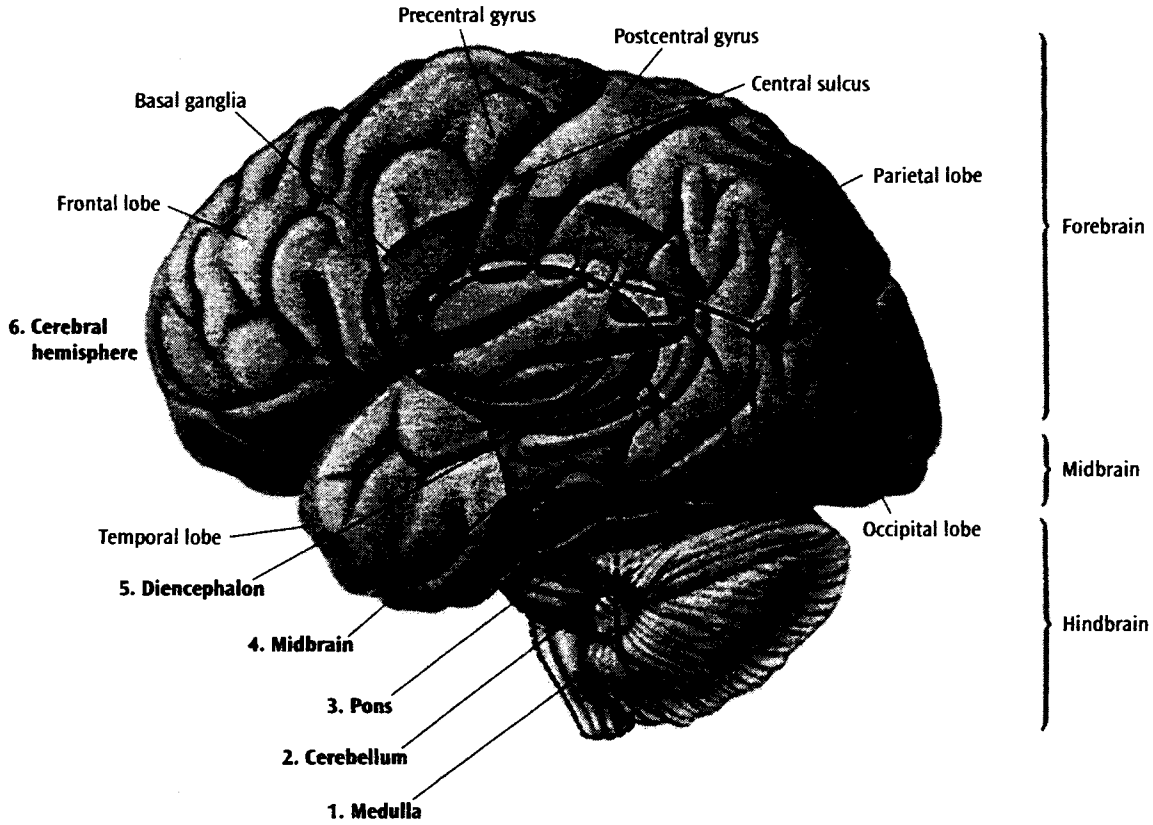



Figure 5.1. Overview of the anatomical structures of the brain. Left hemisphere is shown.

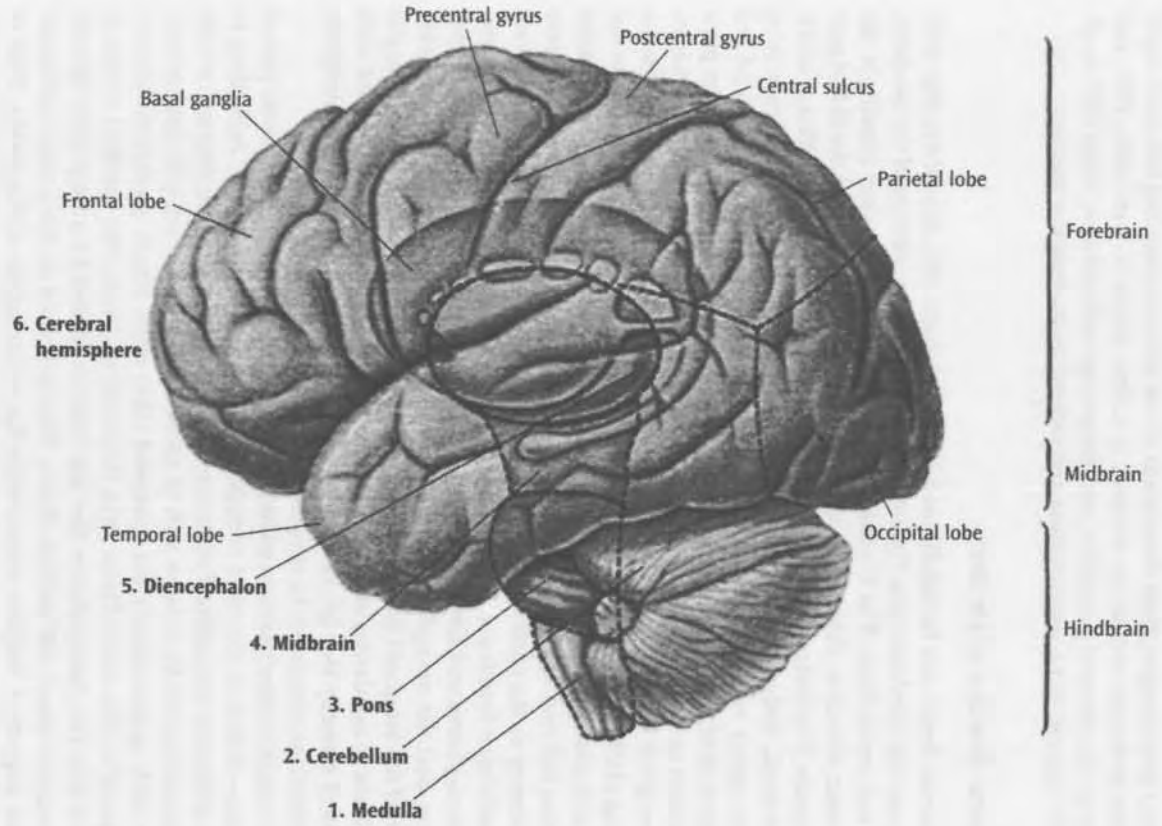



Figure 5.1. Overview of the anatomical structures of the brain. Left hemisphere is shown.

gions. *Primary sensory regions* (or *areas*) are directly linked to the sensory organs via the thalamus. *Primary motor regions* are directly linked to the spinal chord. *Secondary* and *tertiary* sensory and motor areas are adjacent to the primary areas and process more complex stimulus features. Fourth, early intense training processes, starting before the age of about ten years, may lead to enlargement of the cortical areas involved in the trained faculty.

Each hemisphere is divided into four anatomically distinct cortical lobes called the frontal (front), temporal (side), parietal (upper back), and occipital (back) lobes (Figures 5.1 and 5.2). The frontal lobes are largely concerned with the planning of future action and the control of movement. The parietal lobes, which are located behind the frontal lobe and separated by a deep fissure known as the *central sulcus*, are mainly concerned with the processing of somatic sensation and body image. The occipital lobes, which are responsible for processing vision, lie behind the parietal lobes at the back of the brain. The temporal lobes are separated from the frontal and the parietal lobes by a further deep fissure—the *lateral sulcus*. These temporal lobes deal not only with hearing but also with other aspects such as cross-modal learning, memory, and emotion.

## Functional Neuroanatomy of Music

### *Perception*

Music perception involves the *primary* and *secondary auditory areas* (A1, A2) and the *auditory association areas* (AA) in the two temporal lobes (Figure 5.2). The primary auditory area, localized in the upper portion of the temporal lobes just above the ears, receives its main input from the ears via the thalamus and can be regarded as the first station of cortical auditory processing. It is mainly involved in basic auditory processing such as pitch and loudness perception, perception of time structures, and spectral decomposition. The left side of the primary auditory cortex is specialized in the rapid analysis of time structures, such as differences in voice onset times when articulating *da* or *ta*. The right deals primarily with the spectral decomposition of sounds. The secondary auditory areas surround the primary area in a beltlike formation. Here more complex auditory patterns such as timbre are processed. Finally, in the auditory association areas, auditory *gestalt* perception takes place. Auditory gestalt can be understood, for example, as pitch-time patterns like melodies and words. In right-handers and in about 95% of all left-handers, the *Wernicke region* in the back portion of the upper part of the temporal lobe in the left hemisphere (above and behind the left ear) is specialized in language decoding.

In contrast to the *early auditory processing* of simple acoustic structures, listening to music is a far more complex task. Music is experienced not only as an acoustic structure in time but also as patterns, associations, emotions, expectations, and so on. Such experiences are not based on a uniform mental capacity but on a complex set of perceptive and cognitive operations represented in the central nervous system. In some parts, these operations act interdependently;

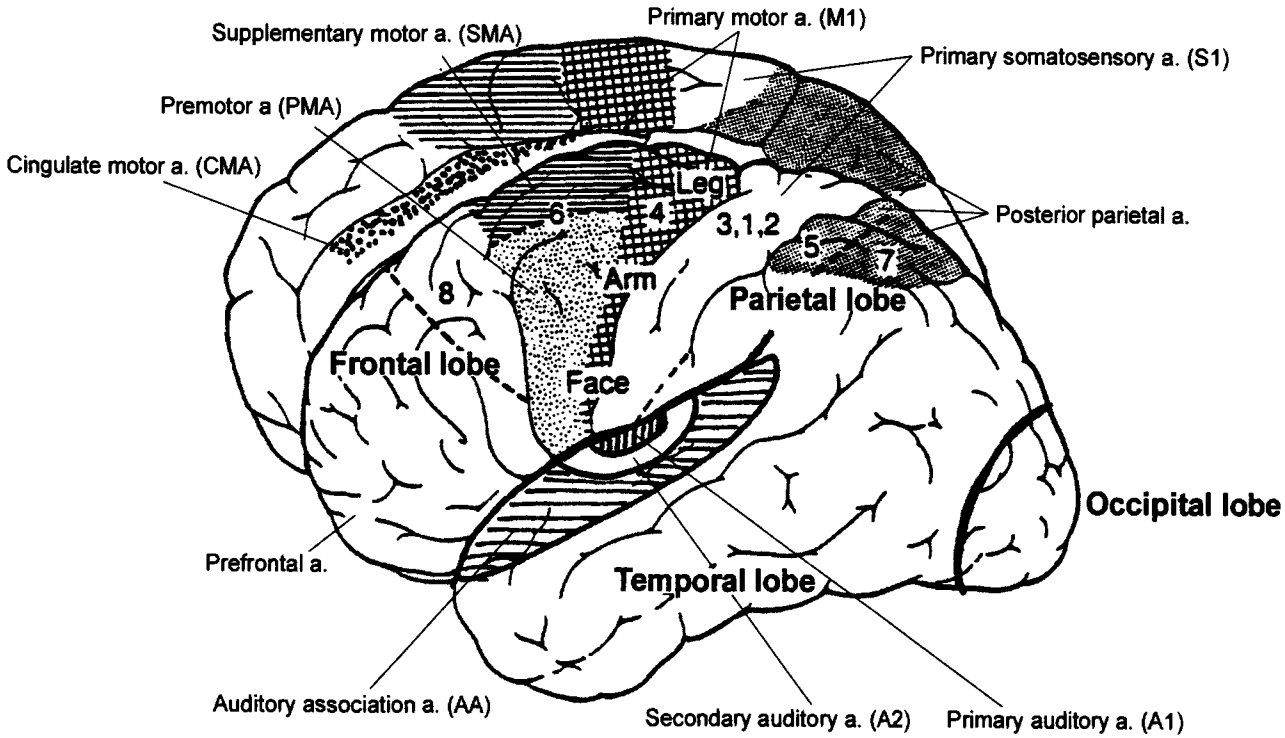



Figure 5.2. Brain regions involved in sensory and motor music processing. (The abbreviation "a." stands for "area.") Left hemisphere is shown in the foreground (lower right), right in the background (upper left).

in others, they function independently. Integrated in time and linked to previous experiences through the aid of memory systems, they enable us to experience music as meaningful. This feeling of meaning takes place on a nonverbal level and is closely linked to emotions.

The neuronal basis of the respective operations may be compared to modules working partly as isolated neuronal processing units. Time and pitch structures, for example, seem to be processed, at least in part, by separate modules, since they can be selectively impaired following brain lesions. Generally speaking, musical time structures are largely processed in the left temporal lobe, whereas pitch structures are processed primarily through networks in the right temporal lobe.

The situation becomes more complicated when considering that the perception of music may occur at different hierarchical levels. With respect to melodic structures, interval-based listening through step-by-step analysis of a melody must be distinguished from contour-based listening in which the more general features of a melody are analyzed. The first type of listening is regarded as a *local* mode of cognitive processing; the latter, as a *global* or *holistic* mode. By analogy, note-by-note rhythm processing may be considered as a local task, in contrast to meter processing, which is a more global task. From studies of brain lesions and functional imaging studies we have learned that local processing relies, to a great extent, on structures in the brain's left hemisphere. Global processing, in contrast, is reliant upon those structures in the right hemisphere (e.g., Peretz, 1990). Since music listeners may switch from one mode to another, it is evident that the neuronal networks involved in music processing are adaptive, rapidly changing, and not fixed in definite music centers. It should be emphasized that as soon as music is processed extended neuronal networks are activated that reflect the individual's way of listening and processing. These include not only the temporal but also the frontal, parietal, and occipital lobes of both hemispheres of the brain (Schuppert et al., 2000). Some factors that determine the generation of such networks will be discussed later.

### *Performance*

Music performance on a professional level requires extremely refined motor skills that are acquired over many years of extensive training and that have to be stored and maintained through further regular practice. Auditory feedback is needed to improve and perfect performance. Music making, therefore, relies primarily on a highly developed auditory-motor integration capacity, which can be compared to the oral-aural loop in speech production. In addition, somatosensory feedback constitutes another basis of high-level performance. Here the kinesthetic sense, which allows for control and feedback of muscle and tendon tension as well as joint positions and which enables continuous monitoring of finger, hand, and lip position in the frames of body and instrument coordinates (for example, the keyboard, the mouthpiece, etc.), is especially important. In a more general context, the motor system of music performance can be understood as a subspecialty of the motor systems for planned and skilled voluntary limb movements.

Voluntary skilled limb movements involve four cortical regions in both hemispheres (Figure 5.2): the *primary motor area* (M1) located in the *precentral gyrus* directly in front of the central sulcus; the *supplementary motor area* (SMA) located anterior to the M1 of the frontal lobe and the inner (medial) side of the cortex; the *cingulate motor area* (CMA) below the SMA and above the corpus callosum on the inner (medial) side of the hemisphere; and the *premotor area* (PMA), which is located adjacent to the lateral aspect of the primary motor area. SMA, CMA, and PMA can be described as *secondary motor areas*, processing movement patterns rather than simple movements. In addition to the cortical regions, the motor system includes the cerebellum and the subcortical structures of the basal ganglia. The sensory areas are necessary in order to maintain the control of movements. Their steady kinesthetic feedback information is required for any guided motor action. The sensory areas are located in the *primary somatosensory area* (S1) behind the central sulcus in the parietal lobe. The parietal lobe is involved in many aspects of movement processing. It is an area where information from multiple sensory regions converges. In the posterior parietal area, the body coordinates in space are monitored and calculated and visual information is transferred into body coordinates. As far as musicians are concerned, this area is prominently activated during sight-reading and the playing of a complex piece of music (Sergent, 1993).

The *primary motor area* (M1) represents the movements of body parts in a separate but systematic order. The representation of the leg is located on the top and the inner side of the hemisphere, the arm in the upper portion, and the hand and mouth in the lower portion of M1. This representation of distinct body parts in corresponding brain regions is called somatotopic or homuncular order. Just as the *motor homunculus* is represented upside down, so, too, is the *sensory homunculus* on the other side of the central sulcus. The proportions of both the motor and the sensory homunculus are markedly distorted since they are determined by the density of motor and sensory innervation of the respective body parts. For example, control of fine movements of the tongue requires many more nerve fibers that transmit the information to this muscle, as compared to muscles of the back. Therefore, the hand, the lips, and the tongue require almost two-thirds of the neurons in this area. However, as will be pointed out later, the representation of body parts may be modified by usage. Moreover, the primary motor area does not simply represent individual muscles: multiple muscular representations are arranged in a complex way so as to allow the execution of simple types of movements rather than the activation of a specific muscle. This is a consequence of the fact that a two-dimensional array of neurons in M1 has to code for three-dimensional movements in space (for a review see Rouiller, 1996). Put simply, our brain represents movements, not muscles.

The *supplementary motor area* (*SMA*) is mainly involved in the coordination of the two hands, in the sequencing of complex movements and in the triggering of movements based on internal cues. It is particularly engaged when the execution of a sequential movement depends on internally stored and memorized information. The SMA can be subdivided into two distinct functional areas. In the anterior SMA, it would seem that the planning of complex movement

patterns is processed. The posterior SMA seems to be predominantly engaged in two-handed movements and, in particular, in the synchronization of both hands during complex movement patterns. Electrical stimulation of this area during open-brain surgery can produce an interruption of two-handed piano playing (for a brief review see Marsden, Deecke, & Freund, 1996).

The function of the *cingulate motor area* (*CMA*) is still under debate. Electrical stimulation and brain-imaging studies demonstrate its involvement in movement selection based on reward, with reference to the close links between the cingulate gyrus and the emotion-processing limbic system. It seems clear that the CMA plays an important role in mediating cortical cognitive functions and limbic-emotional functions.

The *premotor area* (*PMA*) is primarily engaged when externally stimulated behavior is being planned and prepared. It is involved in the learning, execution, and recognition of limb movements and seems to be particularly concerned with the visual information necessary for movement planning.

The *basal ganglia*, located deep inside the cerebral hemispheres, are interconnected reciprocally via the thalamus to the motor and sensory cortices, thus constituting a loop of information flow between the cortex and the basal ganglia. They are indispensable for any kind of voluntary actions that are not highly automated. Their special role consists in the control of voluntary action by selecting appropriate motor actions and by comparing the goal and course of those actions with previous experience. In the basal ganglia, the flow of information between the cortex and the limbic emotion system, in particular the amygdala, converges. It is, therefore, assumed that the basal ganglia process and control the emotional evaluation of motor behavior in terms of expected reward or punishment.

Finally, the *cerebellum* contributes essentially to the timing and accuracy of fine-tuned movements.

### *Musicians' Brains are Different*

The organization of motor systems described earlier can be applied to skilled movements in general. However, in the following section we expand this to include literature that focuses on the unique qualities of the musicians' brain.

Comparison of the brain anatomy of skilled musicians with that of nonmusicians shows that prolonged instrumental practice leads to an enlargement of the hand area in the motor cortex (Amunts et al., 1997). This enlargement appears to be particularly prominent in all instrumentalists who have started to play prior to the age of 10 years. Furthermore, in professional musicians the normal anatomic difference between the larger, dominant (mostly right) hand area and the smaller, nondominant (mostly left) hand area is less pronounced than that in nonmusicians. These results suggest that functional adaptation of the gross structure of the brain occurs during training at an early age.

Similar effects of specialization have been found with respect to the size of the corpus callosum (Schlaug et al., 1995a). Professional pianists and violinists tend to have a larger anterior (front) portion of this structure, especially those

who have started prior to the age of 7. Since this part of the corpus callosum contains fibers from the motor and supplementary motor areas, it seems plausible to assume that the high demands on coordination between the two hands and the rapid exchange of information may either stimulate the nerve fiber growth—the *myelination* of nerve fibers that determines the velocity of nerve conduction—or prevent the physiological loss of nerve tissue during aging.

However, it is not only motor areas that are subject to anatomical adaptation. By means of MEG, the number of nerve cells involved in the processing of sensory stimulation in individual fingers can be monitored. Using this technique, professional violinists have been shown to posses enlarged sensory areas that correspond to the index through to the small (second to fifth) fingers of the left hand (Elbert et al., 1995). Their left thumb representation is not different from that of nonmusicians. Again, these effects were most pronounced in violinists who started their instrumental training prior to the age of 10.

A further example of functional specialization reflected by changes in gross cortical anatomy can be found in musicians who possess absolute pitch. In these musicians, the upper back portion of the left temporal lobe (*Wernicke region*) is larger in than that of musicians without absolute pitch (Schlaug et al., 1995b). Using MEG, the functional specialization of the auditory cortex has recently been demonstrated by Pantev et al. (1998). The number of auditory nerve cells involved in the processing of piano tones but not of pure sinusoidal tones was about 25% greater in pianists than in subjects who had never played an instrument. But it is not only cortical structures that are enlarged by early and prolonged instrumental training. Subcortical structures also seem to be highly affected. In professional musicians, the cerebellum, which contributes significantly to the precise timing and accuracy of motor commands, is also enlarged.

In summary, available evidence suggests that the central nervous system adapts to the challenging demands of professional musicianship during prolonged training. These adaptations are understood as brain plasticity. When training starts at an early age (before about 7 years), this adaptation affects brain anatomy in terms of the enlargement of certain brain structures involved in the respective skill. When training starts later, it modifies brain organization by rewiring neuronal webs and involving adjacent nerve cells to contribute to the required tasks. These changes result in enlarged cortical representations of, for example, specific fingers or sounds within existing brain structures.

## Changes in Brain Organization While Practicing Music

### *Auditory Plasticity in Musicians*

As mentioned earlier, the brain has the potential to reorganize its neural networks, in both the short and the long term, in response to experience and learning. Of the entire sensory system, the auditory system seems to be particularly adaptive. This is illustrated by the adaptation that results from routine ear training of musicians or the adaptation that follows experience with atypical acoustic stimuli.

Cerebral auditory plasticity has been reported in many research studies. The effects of plasticity are not restricted to a critical period in early life but also modulate functional auditory organization in adults. However, in general, the changes are more readily produced and are more stable in younger persons. This kind of dynamic neural plasticity can be shown in an experiment with a manipulated acoustic environment. In one study, adult subjects listened to music distorted by removal of a frequency band centered at one kilohertz for three hours. Immediately after the experiment, the number of auditory nerve cells found to respond to frequencies around one kilohertz was significantly diminished (Pantev et al., 1999). This demonstrates that even after a short exposure to unusual sounds the readiness to respond and the sensitivity of auditory nerve cells may be altered.

Just as the plastic changes mentioned earlier concern more basic auditory processes, musical learning is similarly accompanied by adaptive changes of neuronal networks. Furthermore, different teaching methods and learning attitudes during musical instruction are reflected in specific brain activation patterns. In order to clarify whether the *way music is learned* has any significant, contributory effect, we investigated the changes in brain activation patterns in a group of students who were learning to distinguish between correct and incorrect (balanced and unbalanced) phrases (i.e., the so-called musical periods that consist of corresponding parts, *antecedent* and *consequent*). Subjects were divided into three groups:

1. A *declarative (verbal) learner group* that received traditional instructions about the antecedent and consequent, as well as their tonal relation with respect to the closure of a complete or incomplete cadence, and whose instructions included verbal explanations, visual aids, notation, verbal rules, and some musical examples that were played to the subjects but never sung or performed
2. A *procedural learner group* that participated in musical experiences for establishing genuine musical representations by singing and playing, improvising with corresponding rhythmic and tonal elements, or performing examples from music literature
3. A *control group* of nonlearners who did not receive any instruction about or in music

Figure 5.3 shows the main results of the study. After learning, music processing in the verbally trained declarative group produced an increased activation of the left fronto-temporal brain regions, which may reflect inner speech and analytical, step-by-step processing. In contrast, the musically trained procedural group showed increased activation of the right frontal and bilateral parieto-occipital lobes, which may be ascribed to a more global way of processing and to visuo-spatial associations (Altenmüller & Gruhn, 1997). These results suggest that musical expertise influences auditory brain activation patterns and that changes in these activation patterns depend on the teaching strategies applied. In other words, the brain structures involved in music processing reflect the auditory learning biography—personal experiences accumulated over time. Listening to music, learning to play an instrument, formal instruction, and profes-

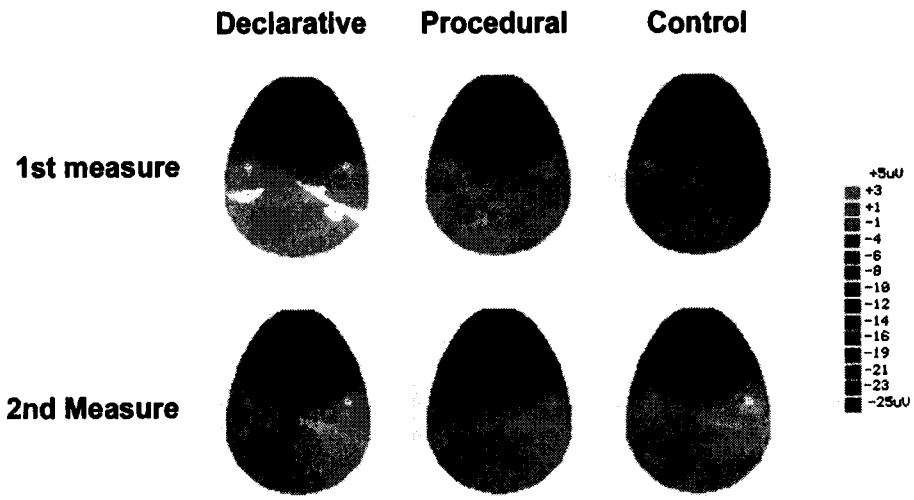



Figure 5.3. Brain maps that demonstrate cortical activation patterns before (upper row) and after (lower row) learning in the declarative learning group, the procedural learning group, and the control group. Group statistics are displayed. Activation is dark; inactivation is white (see microvolt scale on the right). The brain diagrams are displayed as top views, frontal regions up, left hemisphere on the left, right hemisphere on the right. As illustrated, *declarative* (mainly verbally mediated) training leads to an increase in brain activity over the left frontal areas, whereas the *procedural* (mainly musically mediated) training produces an increase in activity over right frontal and bilateral parieto-occipital regions. In controls, overall activity decreased slightly.

sional training result in multiple and in many instances multisensory representations of music. These seem to be partly interchangeable and rapidly adaptive.

### *Sensory-Motor Plasticity and Sensory-Motor Learning*

Our knowledge of the regions and mechanisms of the brain involved in sensory-motor learning is still incomplete. According to current concepts, all structures involved in motor control participate in the acquisition of new sensory-motor skills. Besides the motor areas in the cerebral cortex, the basal ganglia and the cerebellum also play an important role. It has been assumed since the nineteenth century that the cerebellum plays an important role in the acquisition of new motor skills. Such information was based on studies in which patients who were suffering from lesions of the cerebellum were found to be unable to increase the speed of a sequence of complex finger movements after practice. More recent evidence demonstrates that the cerebellum is involved in the selection, the sequence, and the timing of movements. Another functional system, which is equally important for the development and learning of fine finger movements, is the basal ganglia. Patients suffering from Parkinson's disease have deficits in

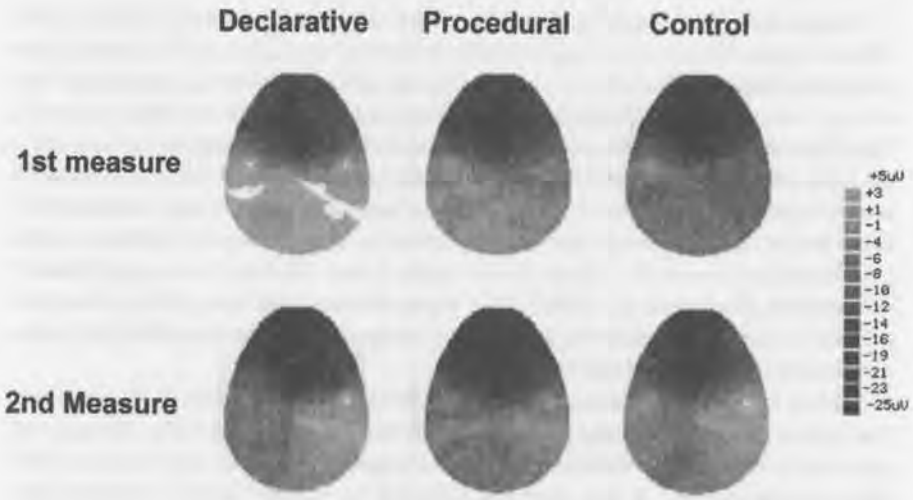



Figure 5.3. Brain maps that demonstrate cortical activation patterns before (upper row) and after (lower row) learning in the declarative learning group, the procedural learning group, and the control group. Group statistics are displayed. Activation is dark; inactivation is white (see microvolt scale on the right). The brain diagrams are displayed as top views, frontal regions up, left hemisphere on the left, right hemisphere on the right. As illustrated, *declarative* (mainly verbally mediated) training leads to an increase in brain activity over the left frontal areas, whereas the *procedural* (mainly musically mediated) training produces an increase in activity over right frontal and bilateral parieto-occipital regions. In controls, overall activity decreased slightly.

sional training result in multiple and in many instances multisensory representations of music. These seem to be partly interchangeable and rapidly adaptive.

### *Sensory-Motor Plasticity and Sensory-Motor Learning*

Our knowledge of the regions and mechanisms of the brain involved in sensory-motor learning is still incomplete. According to current concepts, all structures involved in motor control participate in the acquisition of new sensory-motor skills. Besides the motor areas in the cerebral cortex, the basal ganglia and the cerebellum also play an important role. It has been assumed since the nineteenth century that the cerebellum plays an important role in the acquisition of new motor skills. Such information was based on studies in which patients who were suffering from lesions of the cerebellum were found to be unable to increase the speed of a sequence of complex finger movements after practice. More recent evidence demonstrates that the cerebellum is involved in the selection, the sequence, and the timing of movements. Another functional system, which is equally important for the development and learning of fine finger movements, is the basal ganglia. Patients suffering from Parkinson's disease have deficits in

learning new motor tasks, and—although they can improve the speed of complex movement patterns during practice sessions—they do not learn as quickly, and do not reach the level of performance of normal controls (Salmon & Butters, 1995). Caveat: the findings mentioned earlier are based on data obtained from patients who were suffering from brain lesions and do not take into account the acquisition of the extremely complex movement patterns necessary for professional musicianship.

It has been known for a long time that, with increasing complexity of finger movement sequences, the activity in the SMA and in the premotor area are enhanced (Roland et al., 1980). Using fMRI, Karni et al. (1995) investigated the learning of complex finger sequences similar to those necessary for piano playing. After 30 minutes of practice, the representation of the fingers in the primary motor cortex was increased. However, without further training, this effect diminished after one week with the hand representation returning to its previous size. In contrast, continuous practice resulted in a stable enlargement of the hand area in M1. This effect was specific for the daily trained sequence of complex finger movements and did not occur when the subjects improvised complex finger movements that were not subsequently repeated. Parallel to the enlargement of the hand area in the primary motor cortex, the size of the cerebellar hand representation diminished, suggesting that the cerebellum plays an important role only in the initial phase of motor learning.

Recent investigations have compared brain activation in professional pianists and nonmusicians. Hundt-Georgiadis and Cramon (1999), for example, studied brain activation during 35 minutes of short-term motor learning of a complex finger-tapping task that used the dominant right hand. They found that activation of the primary motor area during motor learning only increased among the pianists. Nonmusicians developed a temporary increase in activity during the first 7 to 14 minutes of training but, subsequently, a rapid attenuation of this activation. Furthermore, compared with nonmusicians, the motor learning of these pianists was accompanied by only small contributions from the supplementary motor areas and the cerebellar cortices. This finding is compatible with the idea that the anterior SMA is essential for setting up and executing complex motor programs prior to automatic performance. The pianists were able to rapidly and thoroughly learn the complex finger sequence, reaching a high degree of automaticity during the first minutes of motor learning.

Most interestingly, the motor cortex of the *untrained* hand was, at the same time, contributing to motor learning. This resulted in improved performance of the motor task in the untrained left hand in both pianists and nonmusicians. One of the most fascinating features of the human sensory-motor system relates to this phenomenon. Despite the clear somatotopic organization of the motor cortex, a movement can be learned with one extremity and performed with the other. Rijntjes and colleagues (1999) investigated subjects who wrote their signature with their right index finger and with their right big toe. The results of the fMRI study show that the movement parameters for these highly trained movements are stored in premotor and supplementary motor cortex adjacent to the right hand area in the primary motor cortex but are also accessible for the foot area. Thus

somatotopy in secondary motor areas (SMA, PM) seems to be defined functionally—as abstract movement information independently of the executing limb (*Bewegungs-Ideen*)—and not on the basis of anatomical representations. Whereas these studies refer to explicit motor learning by trial and error and sensory feedback, implicit motor learning seems to be processed in a different way. When subjects were unaware of a motor learning task—because their attention was drawn to a different problem—activity of the basal ganglia correlated with motor learning.

None of the studies mentioned earlier take into account the special quality of musicianship and, in particular, the strong coupling of sensory-motor and auditory processing. Practicing an instrument means assembling, storing, and constantly improving complex sensory-motor programs through prolonged and repeated execution of motor patterns under control monitored of the auditory system.

Many professional pianists report that their fingers move more or less automatically when they are listening to piano music played by a colleague. In a cross-sectional experiment, we demonstrated that as a result of many years of practice, a strong linkage between auditory and sensory-motor cortical regions develops (Bangert, Parlitz, & Altenmüller, 1998). Furthermore, in a longitudinal study it was possible to follow up the formation of such neuronal multisensory connections along with piano training in beginner piano players. Nonmusicians who had never played an instrument before were trained on a computer piano twice a week over a period of 5 weeks. They listened to short piano melodies of a 3-second duration played in a five-tone range. They were then required, after a brief pause, to replay the melodies with the right hand as accurately as possible. After 10 minutes of training, listening to piano tunes produced additional activity in the central and left sensory-motor regions. In turn, playing on a keyboard produced additional activity in the auditory regions of both temporal lobes. These early signs of cortical plasticity during the first training session were not stable but stabilized within the subsequent 5 weeks of training. In the movement task, the most remarkable effect after 5 weeks was the development of additional activation of the right anterior temporal and frontal lobe. Since it has been demonstrated that this area is involved in the perception of pitch sequences, such activation might reflect the auditory imagery of sounds while moving the fingers on a soundless keyboard. In this context, it should be mentioned that the results of the experiment support the idea of the direct effectiveness of mental training on subtle sensory motor activation patterns represented in the central nervous system.

Many questions that concern the brain mechanisms of sensory-motor learning and processing during music performance remain to be clarified. It is still unclear, for example, how and where the rapid adjustment of the sensory-motor system to different spatial coordinates is processed when musicians switch between instruments of different sizes (e.g., from alto flute to piccolo). This phenomenon is referred to as response size or movement schema (Schmidt & Lee, 1999). Besides the aforementioned limb-independent storage of movement information in the secondary motor areas, the cerebellum may calculate a magnification or diminution factor. Another unsolved problem is the neuronal basis

of the transition from guided slow movements, which are performed under steady sensory control, to fast, ballistic movements, which have to be performed without on-line sensory feedback. It is assumed that different brain regions produce these two types of movements and that the transition from one type to the other may be incomplete. This might explain why practicing guided movements while slowly and systematically increasing the tempo may finally hamper the execution of this movement at a very fast tempo. Although this has been recognized as a problem by many instrumental teachers, no research data are available on this topic. From an empirical standpoint, we would recommend, even at an early stage, rehearsing small segments of the movement pattern at a fast tempo. However, at the same time, the precise automation of difficult movements has to be practiced in a precisely guided slow tempo. For examples, see Cortot (1948).

### *Loss of Motor Control in Highly Trained Musicians*

Approximately 1 in 200 professional musicians suffers from a loss of voluntary control of the extensively trained, refined, and complex sensory-motor skills—a condition generally referred to as *focal dystonia*, violinist's cramp, or pianist's cramp. In most cases, focal dystonia is so disabling that it prematurely ends the artist's professional career (Altenmüller, 1998; chapter 6 in this volume). Subtle loss of control in fast passages, finger curling (cf. Figure 5.4), lack of precision in forked fingerings in woodwind players, irregularity of trills, sticking fingers on the keys, involuntary flexion of the bowing thumb in strings, and impairment of control of the embouchure in woodwind and brass players in certain registers are the various symptoms that can mark the beginning of the disorder. At this stage, most musicians believe that the reduced precision of their movements is due to a technical problem. As a consequence, they intensify their efforts, but this often only exacerbates the problem.

Although the neurobiological origins of this disorder are not completely clarified, at present focal dystonia is understood as a cortical sensory-motor mislearning syndrome, caused by development and stabilization of disadvantageous nerve-cell connections in the sensory and motor brain regions, a condition generally referred to as *abormal plasticity* or *dysfunctional plasticity*. A study with trained monkeys (Byl, Merzenich, & Jenkins, 1996) demonstrated that chronic overuse and repetitive strain injury in highly stereotyped movements can actively degrade the cortical representation of the somatosensory information that guides the fine motor hand movements in primates. A similar degradation of sensory feedback information and concurrent fusion of the digital representations in the somatosensory cortex was confirmed in an MEG study conducted in musicians with focal dystonia, although these musicians had no history of chronic pain (Elbert et al., 1998). Therefore, additional factors such as a genetic predisposition and a certain susceptibility appear to play an important role in the development of focal dystonia

The most important and up to now unsolved question that concerns the therapy of musicians' dystonia is why this condition cannot be easily overcome by retraining and establishing new and appropriately functioning sensory-motor movement patterns.

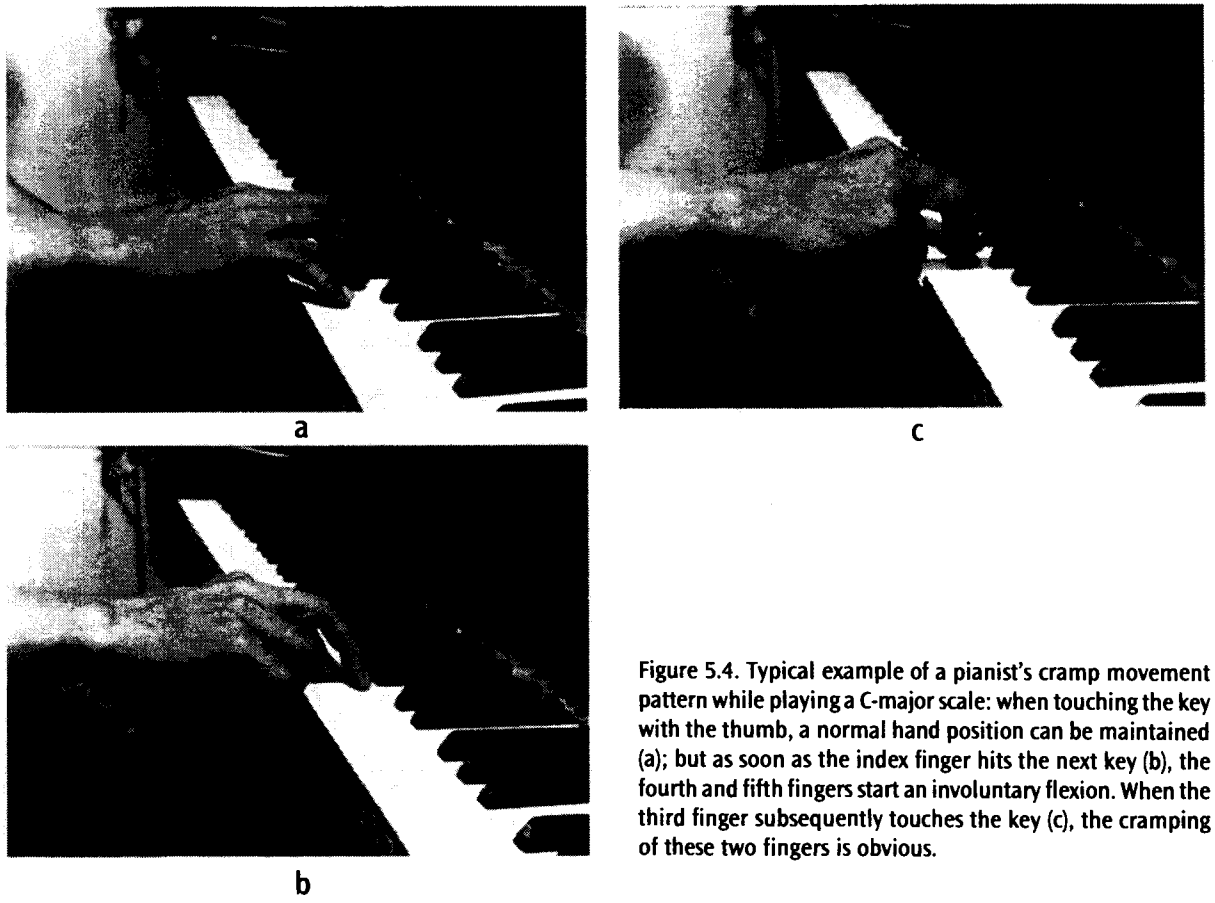



Figure 5.4. Typical example of a pianist's cramp movement pattern while playing a C-major scale: when touching the key with the thumb, a normal hand position can be maintained (a); but as soon as the index finger hits the next key (b), the fourth and fifth fingers start an involuntary flexion. When the third finger subsequently touches the key (c), the cramping of these two fingers is obvious.

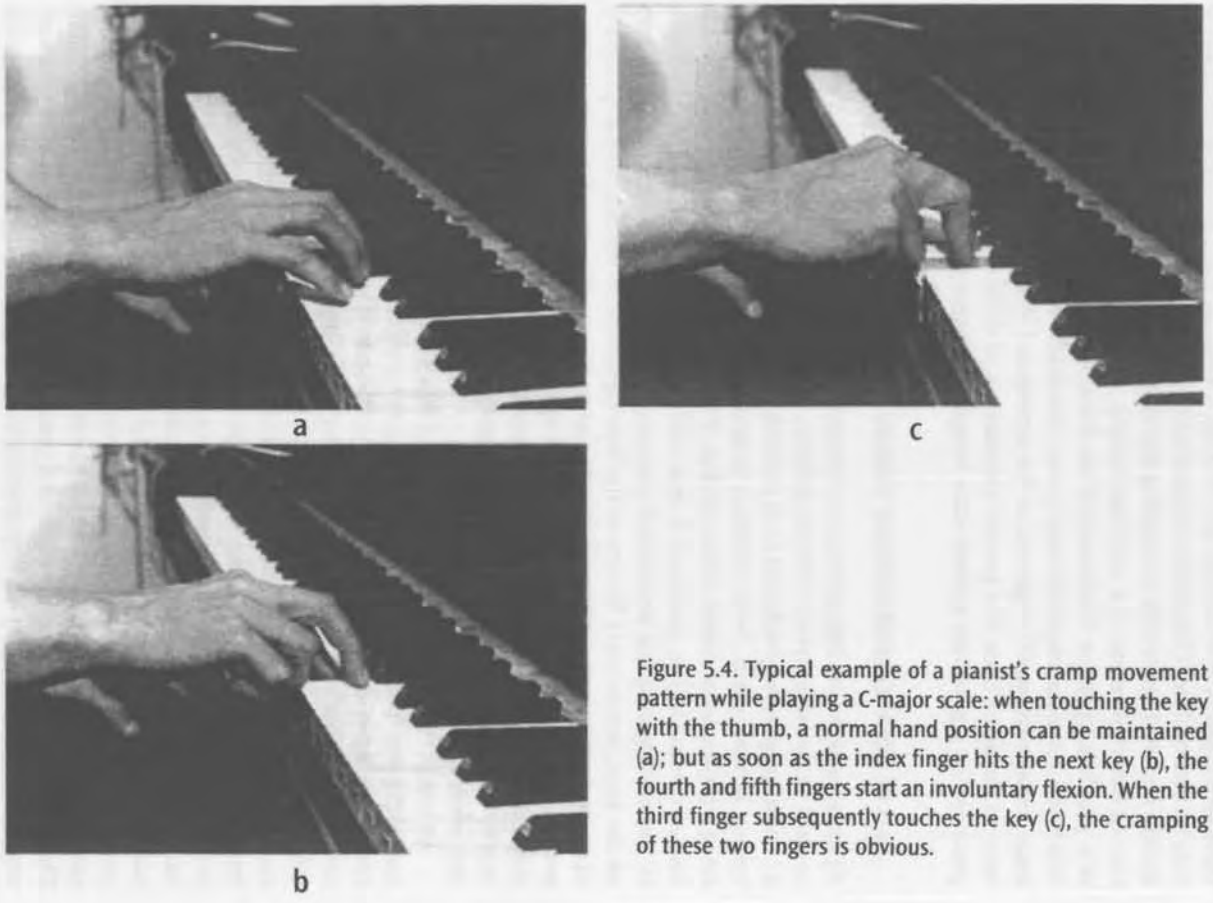



Figure 5.4. Typical example of a pianist's cramp movement pattern while playing a C-major scale: when touching the key with the thumb, a normal hand position can be maintained (a); but as soon as the index finger hits the next key (b), the fourth and fifth fingers start an involuntary flexion. When the third finger subsequently touches the key (c), the cramping of these two fingers is obvious.

## Conclusions and Implications

At a time when brain research has become a key issue in science, the public is tempted to expect clear recommendations from brain studies that can be applied to teaching and learning. It is essential to be fully aware of the distinction between descriptions of how brains work and "think" (Calvin 1996; Pinker 1997) and prescriptions and recommendations on how to teach. As Gardner (1999, p. 60) has stated: "We could know what every neuron does and we would not be one step closer to knowing how to educate our children," since decisions in education are built upon value judgments. Science, however, deals with mindful theories, explanatory models, and empirical observations with no immediate link to educational values.

Nevertheless, research findings of changes in brain organization such as those described earlier may guide teachers' awareness toward optional directions of teaching and learning. In instrumental training, the main focus is generally placed upon the development of motor skills and how to automate patterns of movement. The main point here is that the brain is most flexible (or plastic) during the early years of childhood; later it becomes increasingly difficult to compensate for an underdeveloped disposition for motor skills and fine motor reflexes. Furthermore, as research has shown, the brain learns best when it is actively involved in exploring and experiencing the physical dimension of musical materials and actions. In "learning music musically" (Gruhn, 1997; Swanwick, 1999)—that is, in establishing genuine musical representations—singing and moving should be involved to enhance the aural-oral loop. As demonstrated by Bangert, Parlitz, and Altenmüller (1998), learning results in a network of coactivation where, at the very least, visual, sensory-motor, and aural representations are linked together. Learning strategies, therefore, depend on where the focus lies. If the main goal is concerned with the automation of motor skills, countless repetitions will support that process. The development of genuine musical representations that immediately represent musical properties as musical units but not in terms of visual or verbal features calls for a musical thinking that attributes intrinsic musical meaning to musical sound. This ability is called *audiation* by Gordon (1980). Mental musical representations define the physiological correlate of and prerequisite for music audiation. Audiation is the process by which one activates already-established familiar musical patterns that are stored as mental representations. Therefore, any learning efforts should be directed to establish mental images of sound prior to the training of mere motor or reading and writing skills.

Aural imagery is closely linked with musical practice, and the ability to do this depends on the level and amount of preceding practical experience. This is often misunderstood by parents who are impatient for immediate progress in expected practical skills. This may also happen when children do not read music literally (through associating letter names or syllables with notes) but rather look at the notation and play by ear along with what they imagine from the score in front of them. What may appear to be a failure of instruction actually manifests different approaches to music learning, which is best accomplished when students relate symbols (i.e., the notation) to sounds. In this manner, the perceived sound (i.e.,

the musical ear) guides the finger movements and instructs the students about what to play. Seen from this perspective, learning how to perform music musically not only is due to the effects of conditioning the best possible execution of motor skills but also integrates different types of representation (aural, visual, sensorimotor) into a more complex neuronal network. In the future, brain activation studies may help to develop educational strategies that foster a stronger connectivity among different brain areas by exploiting cross-modal representations.

Acknowledgments. We are grateful to the coworkers of the Institute of Music Physiology and Performing Arts Medicine for stimulating ideas and many discussions on the topics of this contribution. The piano-learning experiment reported in the text has been conducted by Marc Bangert. The experiments on ear training and brain activation were carried out in collaboration with Dietrich Parlitz, to whom we owe Figure 5.3. Wendy Archer improved the English expression.

### *References*

Altenmüller, E. (1998). Causes and cures of focal limb dystonia in musicians. *Journal of the International Society for the Study of Tensions in Performance (ISSTIP), 9,* 13–17.

Altenmüller, E., & Gruhn, W. (1997). *Music, the brain, and music learning.* Chicago: GIA.

Amunts, K., Schlaug, G., Jäncke, L., Steinmetz, H., Schleicher, A., Dabringhaus, A., & Zilles, K. (1997). Motor cortex and hand motor skills: Structural compliance in the human brain. *Human Brain Mapping, 5,* 206–215.

Bangert, M., Parlitz, D., & Altenmüller, E. (1998). Audio-motor integration in beginner and expert pianists: A combined DC-EEG and behavioral study. In N. Elsner and R. Wehner (Eds.), *Göttingen Neurobiology Report* (p. 753). Stuttgart: Thieme.

Byl, N. N., Merzenich, M. M., & Jenkins, W. M. (1996). A primate genesis model of focal dystonia and repetitive strain injury. *Neurology, 47,* 508–520.

Calvin, W. H. (1996). *How brains think.* New York: Basic Books.

Cortot, A. (1948). *F. Chopin, Etudes Op. 10.* Paris: Salabert.

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science, 270,* 305–307.

Elbert, T., Candia, V., Altenmüller, E., Rau, H., Sterr, A., Rockstroh, B., Pantev, C., & Taub, E. (1998). Alterations of digital representations in somatosensory cortex in focal hand dystonia. *NeuroReport, 9,* 3571–3575.

Gardner, H. (1999). *The disciplined mind.* New York: Simon & Schuster.

Gordon, E. E. (1980). *Learning sequences in music.* Chicago: GIA.

Gruhn, W. (1997). Music learning: Neurobiological foundations and educational implication. *Research Studies in Music Education, 9,* 36–47.

Hundt-Georgiadis, M., & Cramon, D. Y. (1999). Motor-learning related changes in piano players and non-musicians revealed by functional magnetic resonance signals. *Experimental Brain Research, 125,* 417–425.

Karni, A., Meyer, G., Jezzard, P., Adams, M. M., Turner. R., & Ungerleider, L. G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature, 377,* 155–158.

Marsden, C. D., Deecke, L., & Freund, H.-J. (1996). The functions of the supplementary motor area: Summary of a workshop. In H. Lüders (Ed.), *Advances in neurology: Vol. 70. Supplementary sensorimotor area* (pp. 477–487). Philadelphia: Lippincott-Raven.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature, 392,* 811–814.

Pantev, C., Wollbrink, A., Roberts, L. E. , Engelien, A., & Lutkenhoner, B. (1999). Short-term plasticity of the human auditory cortex. *Brain Research, 842,* 192–199.

Peretz, I. (1990). Processing of local and global musical information in unilateral brain-damaged patients. *Brain, 113,* 1185–1205.

Pinker, S. (1997). *How the mind works.* New York: Norton.

Rijntjes, M., Dettmers, C., Buchel, C., Kiebel, S., Frackowiak, R. S., & Weiller, C. (1999). A blueprint for movement: Functional and anatomical representations in the human motor system *Journal of Neuroscience, 19,* 8043–8048.

Roland, P., Larsen, B., Lassen, N. A., & Skinhoi, E. (1980). Supplementary motor area and other cortical areas in organization of voluntary movements in man. *Journal of Neurophysiology, 43,* 118–136.

Rouiller, E. M. (1996). Multiple hand representations in the motor cortical areas. In A. M. Wing, P. Haggard, and J. R. Flanagan (Eds.), *Hand and brain* (pp. 99–124). San Diego, CA: Academic Press.

Salmon, D. P., & Butters, N. (1995). Neurobiology of skill and habit learning. *Current Opinion in Neurobiology, 5,* 184–190.

Schlaug, G., Jäncke, L., Huang, Y., & Steinmetz, H. (1995a). Increased corpus callosum size in musicians. *Neuropsychologia, 33,* 1047–1055.

Schlaug, G., Jäncke, L., Huang, Y., & Steinmetz, H. (1995b). In vivo evidence of structural brain asymmetry in musicians. *Science, 267,* 699–701.

Schmidt, R. A., & Lee, T. D. (1999). *Motor control and learning: A behavioral emphasis* (3rd ed.). Champaign, IL: Human Kinetics.

Schuppert, M., Münte, T. F., Wieringa, B. M., & Altenmüller, E. (2000). Receptive amusia: Evidence for cross-hemispheric neural networks underlying music processing strategies. *Brain, 123,* 546–559.

Sergent, J. (1993). Mapping the musician brain. *Human Brain Mapping, 1,* 20–38.

Swanwick, K. (1999). *Teaching music musically.* London: Routledge.

# 6

# Music Medicine

ALICE G. BRANDFONBRENER & JAMES M. KJELLAND

Most of the medical problems of musicians are the shared consequence of the specific instrument, performance technique, and repertoire interacting with the physical and psychological nature of the individual. The incidence of problems is greater for those instruments requiring more repetitious actions over a longer period of time and in all the risks that are increased by stress. Most frequent are musculoskeletal pain problems such as tendinitis, which typically resolve with simple measures like reduced activity, anti-inflammatory medication, and icing. Prevention is preferable to treatment for all these conditions but more research is needed to validate the techniques to be employed and to more precisely identify the causal factors. This requires close collaboration between medicine and music education.

Many, if not most, occupations are associated with some degree of personal medical risk, and music making is no exception. But it was not until the 1980s that attention became focused on music medicine. From the musicians' viewpoint, there was a certain degree of fear behind their reluctance to seek medical advice, fear of losing a career and fear of uninformed medical care. Most musicians know someone whose career was compromised or ended due to injury. Likewise, they may have themselves experienced or have heard of a colleague's negative medical experiences, which often resulted in a justifiable distrust of physicians. From the physicians' viewpoint, musicians with medical complaints often showed little or no observable pathology, and so the symptoms were often judged to be exaggerated and no diagnosis was made. Few physicians saw sufficient numbers of musicians to understand and properly assess their symptoms, especially in the context of their occupational demands, with the result that they frequently did not receive adequate or appropriate treatment.

In the 1980s this situation changed when musicians began to more openly discuss their medical problems. This was evident when two well-known and successful concert pianists, Leon Fleisher and Gary Graffman, disclosed to the public that they had medical problems that were severely affecting their ability to perform. This served to provide a kind of implicit permission for other musicians to acknowledge problems and to seek help without shame and with less apprehension. A case can also be made for an apparent increase in frequency of problems, but this is difficult to prove, since it is only recently that incidence data have become available. Concurrently, members of the medical community, many of whom were themselves musicians, came to recognize that the often-subtle complaints of musicians represented real pathology, albeit requiring special expertise to diagnose and treat appropriately. It was this coming together of many factors that led to the establishment and legitimizing of the new medical specialty of performing arts medicine.

Performing arts medicine has grown worldwide since the early 1980s. As of the beginning of the twenty-first century, there are clinical and research centers, national and international meetings and workshops, professional societies, and, in the United States, a peer-reviewed medical journal devoted entirely to the problems of performing artists (Harman, 1995). This is not to say that the evolution of performing arts medicine has been problem-free or that there is adequate financial support for its clinical or research endeavors. The growth of the field has been increasingly challenged by the money-saving efforts of medical insurers through limiting patient access to speciality consultations.

Because of its intrinsically multispecialty approach, there is no authorized training program or certification for becoming a performing arts medicine specialist and, therefore, there is no institutionalized quality control for its practice. Many medical students, therapists, and doctors in training spend elective time at one or another arts clinic or in a research lab to hone their skills. These experiences can obviously vary in quality and quantity. Among the goals for this new field are the development of criteria and standards for the practice and the teaching of performing arts medicine. In the meantime, those who are looking for information must keep in mind that performing arts medicine remains a discipline in the making; constructive criticism and questioning by practitioners, educators, and consumers are all necessary and appropriate.

While all musicians share certain occupational risks, there is a significant difference between the manifestations, diagnosis, and treatment of pathology in vocal versus instrumental musicians. For most of their medical problems the care of vocalists requires a different set of specialists than does that of instrumentalists. In fact, the level of science in the care of vocalists is more advanced than has been thus far achieved for other musicians. Singers have by necessity been more aware of their need for medical assistance, and problems of the vocal mechanism has long been a recognized area for subspecialization by laryngologists. There is a significant body of literature based on clinical experience and research, and a sophisticated technology has been developed for use in the diagnosis, assessment, and treatment of vocal problems and in vocal pedagogy (see chapter 16).

### *Incidence of Problems*

With heightened interest in musicians' medical issues, a number of incidence studies of problems among different musical populations have been undertaken. Zaza (1998) reviewed 24 studies, of varying sizes and among different types of musical populations, from 1980 to 1996, but concluded that only 7 were methodologically reliable. Her review reports a prevalence that ranged from 39 to 87% of injuries in adult musicians and 17% in secondary school students. Two of the best studies were completed by Fishbein, Middlestadt, Ottani, Straus, and Ellis (1988) with players in professional U.S. orchestras and by Cayea and Manchester (1998), with conservatory students. Fishbein surveyed 2,212 musicians in 48 International Conference of Symphony and Opera Musicians (ICSOM) orchestras and reported that 76% had at least one serious medical problem that affected their playing. This study also included data on the relative incidence of problems in different orchestra sections, with the most problems occurring in strings, followed by woodwind, brass, and percussion. The data also reveal a higher incidence among female than male players and evidence that as the size of the string instrument grew, so did the likelihood of problems. While this was a factor especially in women, it also affected men. Although pianists are underrepresented among orchestral musicians and there has been no comparably large survey of pianists, the incidence of their problems appear to be around the same as in string players (Cayea & Manchester, 1998).

The Cayea and Manchester study of conservatory students was based on a review of the number of visits by music students to the health service for playing-related complaints during 14 academic years (1982–1996) and showed 8.3 injuries per 100 registered music performance students in an academic year. They also reported a higher incidence among women—8.9 as opposed to 5.9 in men, both per 100 students. The increased incidence in women may have several explanations, including that women more often seek medical help, have smaller hands, have hyperextensible joints, and have more responsibilities for nonoccupational activities than do men. From these data, as well as others, it is clear that youth is no protection against most of the medical problems experienced by musicians and, correspondingly, that being a musician carries risks at all ages.

## Specific Medical Conditions and Their Treatment

Clinically, the medical problems that arise from playing instruments divide into three categories. Most common are *musculoskeletal pain syndromes* (which include *tendinitis*), followed by *nerve entrapments* and, last, *focal dystonias* (occupational cramps). The reported frequency varies with the nature of the practice. Larger clinics and those with a significant referral practice will see a proportionately greater number of patients with focal dystonias than those that see mostly students or those that have a small patient volume. The latter would see proportionately more pain syndromes, fewer nerve entrapments, and very rarely focal dystonias.

### *Muscoloskeletal Pain Syndromes*

It is generally accepted among music medicine practitioners that the most frequent symptom that causes musicians to seek medical advice is pain, both acute and chronic. Hochberg et al. (1983) reported that 21 of 49 patients they reviewed complained of pain as their primary symptom, while Lederman (1994) reported that from 45 to 69% of 672 patients with musculoskeletal pain disorders, depending on the instrument played. In order of frequency the sites involved in musicians' pain syndromes are the arm, neck, and back, although it is also common to have problems concurrently in several sites, especially when the pain is chronic. Most of these disorders occur after overplaying of some sort, and are severe enough to affect playing but often are not symptomatic when not playing. Symptoms that persist more than a few days or progress should be evaluated by a physician to see if some specific treatment is indicated.

Appropriate treatment of muscle tendon pain problems is almost always medical rather than surgical, and while specifics will depend on the factors that led to the problem, much of this requires common sense as much as medical expertise. Often the use of antiinflammatory medications, such as ibuprofen, icing (ice packs), and some degree of rest are prescribed and will suffice. The need for rest, the hallmark of treatment, may be variable. After 30 years of experience, in most instances and for several reasons, the first author of this chapter has become a proponent of continuing some degree of customary musical activity rather than prohibiting all musical activity. This obviously does not apply to nonmusical trauma, in which there may have been actual tissue disruption, such as in fractures or severe sprains. There is no clinical or experimental data that is convincing of the merits of absolute rest, nor is there any rule of thumb for dictating the length of time for rest. My experience is that for both physical and psychological reasons it is more successful overall to allow a musician to play a little than not to play at all. Physically it helps retain facility so that when the symptom has resolved there is less need for excessive practice in order to return to the previous level of playing. Also, tissues retain elasticity, flexibility, and vascularity when used somewhat rather than when totally at rest. These syndromes do not usually lead to more serious problems, and the majority resolve in a matter of days or weeks at the most.

### *Nerve Entrapments*

Nerve entrapments occur when a nerve is trapped in a particular location by some other tissue such as surrounding scar tissue or by inflamed and swollen tendons. These usually cause pain, sometimes quite severe, in the distribution of the affected nerve rather than only at the site of the entrapment (referred pain) and often are accompanied by a pins-and-needle sensation, sometimes by decreased sensation, and occasionally by weakness and atrophy (diminution in bulk) of affected muscles. However, some of these same symptoms may also occur in other conditions. Symptoms of nerve entrapments in musicians may be subtle or dramatic, especially when the problem is of long duration. Once a diagnosis is made,

treatment depends on the degree of nerve impairment present. Often nonsurgical treatment will be successful, including rest, nonsteroidal antiinflammatories, occasionally cortisone, and splinting (see "Misdiagnosis and Mistreatment") for absolute rest. Whether or not surgery seems necessary, when and of what type is a question about which different well-qualified specialists may often be contradictory. This may necessitate obtaining multiple opinions. The most frequently entrapped nerve is the *median nerve* in the *carpal tunnel*, which lies beneath the fold where the palm of the hand meets the wrist. The carpal tunnel is a rigid structure, the floor and sides of which are wrist bones and the roof of which is formed by the transcarpal ligament, an unyielding piece of soft tissue. The contents of the canal include nine flexor tendons and the median nerve. When the tendons become inflamed and swollen they put varying degrees of pressure on the median nerve, causing the symptom complex known as *carpal tunnel syndrome*. Next in frequency of nerve entrapments is *cubital tunnel syndrome*, or entrapment of the *ulnar nerve*, at or in some cases slightly above the elbow. This is a less defined structure than the carpal tunnel, but essentially the nerve may become compressed by soft tissues in an area of relatively tight passage of the ulnar nerve. While these can often be diagnosed on clinical presentation alone, most practitioners rely on electrodiagnostic tests (*electromyograms* or *EMGs* and *nerve conduction velocity* or *NCVs*) for confirmation or refutation, as well as to follow the course of a particular entrapment. Carpal and cubital tunnel syndromes are typically treated with a degree of rest, often intermittent splinting, nonsteroidal antiinflammatory agents, occasionally cortisone injections, and physical or occupational therapy. If these methods fail or there is evidence of significant progression, operative intervention with release is indicated. Here there is usually a choice of the procedure as well as of the surgeon and, since this is always an elective procedure, musicians do well to fully inform themselves and to seek several opinions.

There is yet another diagnosis that may theoretically involve nerve entrapment, *thoracic outlet syndrome* (*TOS*), but in musicians especially this is an area of often-heated medical dispute, primarily because it does not fit the usual definition of nerve entrapment. *Thoracic outlet* refers to the area inside of the upper chest through which traverse the nerves and blood vessels that ultimately supply the arms. In some individuals this space may be compromised, due to the configuration of their shoulders, due occasionally to an extra rib, and most frequently due to a constantly changing factor, which is posture. When the space reduction is due to posture, unlike other nerve entrapments, the symptoms are intermittent based on position primarily of the shoulders. In some cases the symptoms of this postural TOS may be due to transient vascular rather than nerve compression. A diagnosis of TOS in musicians is typically made on the basis of often-vague pain that may be in the shoulder, back, or arm, with or without tingling; and, as it occurs in musicians, is rarely if ever accompanied by muscle weakness. Electrodiagnostic testing in these patients does not show nerve damage. Therefore, while physicians may at times suggest surgical release for documented carpal or cubital tunnel syndromes, only an occasional surgeon does not share the view of most physicians, which is to rarely, if ever, recommend sur-

gery for the postural TOS that is typically seen in musicians. Conservative treatment consists of physical therapy with postural exercises, designed to help keep the shoulders back and down.

### *Focal Dystonias*

The most serious medical problem that occurs in musicians is called focal dystonia, and its occurrence may well end or at least significantly impair professional performance careers. It is one of a group of conditions known as occupational cramps. *Focal* means "localized," and *dystonia* refers to abnormal muscle tone. The affected musician will note an inability to control a finger or fingers, primarily in one hand, which curl into the palm, often accompanied by a reciprocal extension of one or more of the other fingers. Among pianists, dystonias tend to affect fingers of the right hand, or the hand involved with the most activity. This occurs only while playing the instrument and does not affect other unrelated hand activities. In woodwind and brass players the condition may present as loss of control of muscles of the embouchure rather than in the fingers. The problem in the embouchure is often difficult to visualize, although sometimes one sees the curling up or pulling away of one side of the upper lip. The incidence among musicians is not known, but it is less common than the previous categories of medical problems discussed. In different consultative specialty practices dystonia patients have reportedly represented from 8 to 27% of the musician patients seen (Lederman, 1991; Brandfonbrener, 1995). The frequency is significantly greater among men than women, and it tends to occur during midlife but may occur earlier, although it is relatively uncommon before the midtwenties. The focal dystonias appear to represent an abnormality of the motor cortex in the brain rather than of the peripheral nerves (chapter 5). It does not extend to other areas of the body or affect the general health of the patient. In most instances all nonmusical activity is unaffected and one cannot detect any abnormality on routine examination. The cause is unknown, although there have been many theories, which range from a history of a previous, nonspecific injury to excess stress and tension in playing. While there are anecdotal reports of partial or complete resolution of symptoms, it is generally believed to be a permanent condition. To date, there is no specific therapy, although for some the injection of botulinum toxin (Botox) into the affected muscles has provided some temporary relief, but this needs to be repeated every few months to maintain even its partial effect (Cohen, Hallett, Geller, & Hochberg, 1989). Rest does not affect the symptoms of dystonia, nor does continuing to practice. Other treatments include the use of anti-Parkinson's medications as well as many alternative, nontraditional therapies, but none has yielded a cure or typically even enough relief to allow return to the previous performance level. A new treatment, called *constraint-induced movement therapy*, is reported to show promising results, but this was in a small series (five) of patients and needs further verification (Candia et al., 1999). Depending on the severity of their particular affliction, some musicians with dystonia can continue to play using selected repertoire, alternative fingerings, or modified instruments. All instrumental

groups may be affected, but there is evidence that it occurs most, as do most other medical problems, in association with instruments that require the most repetitive muscle actions (Lederman, 1991).

## Other Considerations in Assessment and Treatment

### *Psychosomatic Aspects*

In virtually all matters of health there is a psychological component, in their cause as well as in their resolution. How we feel affects our health and our recovery for better and for worse. Depression is a major factor in injuries, both in cause and effect, as it is also an impediment to healing, especially in the context of chronic pain. Practitioners who see large numbers of musician patients note that a significant number of these patients have symptoms that are chronic and, as such, must be treated differently from symptoms that are more acute, that is, problems of more recent onset. These patients in particular need multidisciplinary physician input that includes those trained in physical medicine and rehabilitation and in psychiatry. The judicious use of tricyclic antidepressants is effective in musicians with chronic pain, as it is in chronic pain patients in general, along with exercise, psychotherapy, and often a period of musical retraining.

### *Misdiagnosis and Mistreatment*

Some of the blame for the frequency of chronic problems appears due to inappropriate treatment or mistreatment of the acute problems. Musicians are sufficiently concerned and obsessed when they face a medical problem without physicians giving them further reasons to feel impaired by lengthy prohibitions against playing. In addition, musicians are likely to be noncompliant if the medical treatment seems unduly punishing, so that in the interests of compliance, working out a reasonable plan for reduced playing is usually both viable and pragmatic. Formulas for reduced playing can be devised, but individualization is more logical, depending on all the circumstances. Sometimes it may be advisable to splint the affected part, but there are a variety of ways to do this. Resting splints while useful for providing absolute rest have the disadvantage of causing stiffness and even atrophy if left on more than absolutely needed. Intermittent, or nocturnal splinting, is often sufficient and better. There are also dynamic splints that can be designed for the individual, allowing protection while playing. Splints worn only at night or intermittently, rather than constantly during the day, are less likely to cause weakness or stiffness. Some physicians favor the use of steroid injections (cortisone-type medications) in certain cases of tendinitis and other conditions. Other physicians, including the first author of this chapter, are reluctant to do this because of the possible long-range effects on tissue integrity, especially with multiple injections. Rarely, if ever, should steroid injections be used until more conservative methods have proven inadequate.

### *Contextual Diagnosis*

Perhaps the single most important innovation utilized by practitioners of performing arts medicine is viewing medical problems in their performance context. Musicians often express their symptoms in musical terms, referring to problems playing octaves or double-stops, reduced endurance, change of range, or loss of tone quality or accuracy of pitch. Practitioners of performing arts medicine have learned that speaking and understanding musical language is often crucial. Most important, examining the patient playing the instrument is what defines a specialist in this field, in that most are expected to be familiar with the basics of positioning and technique requisite for playing each instrument. This sometimes may stretch beyond medical expertise into musical expertise, and many physicians and therapists welcome the input of music educators as participants or advisers in patient assessment. All of this takes more time than most traditional office visits, but it is this time and care upon which successful care of musicians is often predicated. Collaboration between medical and musical expertise is therefore essential to the healing process in music medicine. Due to the subtle and idiosyncratic nuances of the process of making music, the medical specialist cannot be expected to know and understand all of these aspects. This is analogous to sports medicine, where diagnosis and treatment are viewed from both medical and performance perspectives.

With musician patients it is essential to take a full medical history as well as a musical history. The former is because one cannot assume that an injury is due to a patient's occupation simply because he or she is a musician. Because musicians tend to be so aware of gradual decreases in function, nonoccupational pathology may present initially as a change in performance facility. Most problems seen among musicians do not stem from a single episode but are cumulative and caused by the interaction of several factors, such as the hours practiced, the intensity of practice, the instrument, the physical and mental properties of the musician, technique, and more. It is necessary to understand what musical maneuvers are compromised, the duration of symptoms, and about any musical changes (instrument, teacher, repertoire, schedule) that related in time to the occurrence of the symptom. Physicians are often at a loss to arrive at a precise diagnosis of musicians because of the lack of objective findings, such as swelling, loss of strength, sensation, or range of motion (ROM). Local tenderness may or may not be present. Examining the whole patient is essential, because playing musical instruments depends on the entire body; the problem in an arm may reside, for example, in weakness or tension in the more proximal musculature or joints or in the back. However, a diagnosis can usually be made on the basis of the history and physical diagnosis. Further testing is useful only to rule out other pathology. While magnetic resonance imaging (MRI) may show tissue swelling in tendinitis and other forms of musculoskeletal pain syndromes, it is rarely legitimate to resort to such extensive and expensive testing if clinical expertise alone will suffice. Likewise, electrodiagnostic tests, such as EMGs and NCVs, are useful only to rule out nerve entrapments or other pathology that can in many cases be done based on ex-

amination. Occasionally there may be a question of an underlying joint disorder, such as rheumatoid arthritis or osteoarthritis, or the effects of trauma, in which case X rays and/or blood tests may be indicated.

### *Risk Factors*

Prevention of musicians' medical injuries and other problems largely involves understanding the risk factors. The risk factors can be partially divided into those that stem from the instrument and those that stem from the musician.

**Instrument-Specific Risks.** Factors that relate to the instrument include the posture with the instrument, the effort of supporting the static load of the instrument, the number of repetitions to depress keys or strings, and the force of the required airflow through the instrument. For a pianist, sitting too close or too far may be dictated by the demands of the technique or the body's proportions, but sitting too close or too far may affect the medical biomechanics as well as the technical. Clarinetists and oboists frequently have problems with their right thumbs, especially if they have small hands or if the thumb joints are hyperlax (double-jointed). The player then uses greater muscle tension to hold the instrument, which over time can result in overuse of the thumb muscles, as well as causing increased tension in the other fingers and the entire upper extremity. This hyperlaxity of finger joints can also adversely affect all instrumentalists (Brandfonbrener, 1990, 2000). Poor posture affects wind players by increasing the work of breathing. Many wind players have jaw pain, which is diagnosed as temperomandibular joint disease, or TMJ. While the problem may ultimately reside in the joint, prior to this the symptoms can usually be attributed to excess muscle tension in the masseter (chewing) and other embouchure muscles. Sometimes this is due to how the musicians position their jaws while playing and sometimes to problems with dental alignment, or both. In any problems of wind players that involve the temperomandibular joints of the jaw or the teeth, it is important to find expert dental advice, because poor advice can result in musically, if not dentally, disastrous procedures. Once again, when one is dealing with woodwind or brass players multiple opinions are appropriate before even seemingly routine extractions of third molars (wisdom teeth), or orthodenture to adjust malocclusions (Howard, 1989).

**Risks That Stem from the Individual.** The risks that the musician brings to playing may be of genetic origin, including body proportions, joint mobility, and finger length. Musicians who practice hours per day and are sedentary have a greater risk of upper extremity and back problems because the shoulder and back muscles, crucial to maintaining posture and arm support, are inadequately developed. As muscles fatigue, they become physically shorter, less flexible, and more subject to injury. Musicians who are emotionally tense tend to have tighter muscles, even before they start to play. There is evidence that how much time is spent in practice and its intensity are equally important as risk factors. Rapid increases in playing time, as at summer music festivals, for competitions, and

for amateurs at workshops, have been documented as predisposing musicians to pain (Newmark & Lederman, 1987).

Repertoire that is technically more challenging, with long series of difficult chords and octaves for pianists, double-stops or thumb position for string players, and higher and longer sustained notes for wind players, is among risk factors that fall between those of the instrument and those intrinsic to the player. Finally, although this area is more complicated and difficult to subject to measurement of relative risk, based on clinical experience there is no question that the musician's overall ability and emotional status may be among the most critical risk factors for injuries to bear in mind.

## Preventive Strategies

To date there are few studies that document the effectiveness of preventive measures. Spaulding (1988) reported work with conservatory students in Norway on posture and education that yielded promising results, even if on a modest scale. On a larger scale, Brandfonbrener and associates undertook a year-long prevention intervention with nine ICSOM (International Conference of Symphony and Opera Musicians) orchestras in 1989 (Brandfonbrener, 1997). Although the program was initially well attended, the participants gradually dropped out over the course of the year, so that while much interesting information was gathered, there were too few subjects left at the end to make any definitive comments about the effects on prevention. The difficulty in maintaining motivation in prevention of health problems is a universal fact of life.

Logically, the best time to attend to the prevention of musicians' injuries is during the early stages of education, so that efficient performance techniques and user-friendly postural habits, a positive attitude, and a healthy lifestyle can be instilled from the earliest lessons. Practice techniques such as pacing (e.g., taking frequent breaks), avoiding excessive repetitions, paying attention to posture, remaining aware of fatigue and tension, and maintaining an exercise routine, such as swimming, are helpful.

### *Physical Tension and Its Psychological Component*

Excessive tension, muscular and emotional (often inseparable), is the most important risk factor for musicians' injuries. While both muscular and emotional tension can be prevented, they may have complicated roots that do not allow for simple solutions. On a routine basis, warming up, cooling down, and stretching help but should not be regarded as providing solutions for all playing-related problems. Yoga, the Feldenkreis method (see Nelson, 1989), and Alexander technique (see Mayers & Babits, 1987; Grey, 1991), as well as other schools of body awareness, are also helpful but frequently not enough. It is often essential, with musicians of any age, to attempt to understand the stresses that are contributing to emotional and muscular tension that have their origins in both musical and nonmusical sources. It will come as no surprise, then, that musicians themselves

freely admit to perceiving great occupational stress (Middlestadt & Fishbein, 1988). This is not a call for amateur psychiatrists but rather a call for attention on the part of all involved in the music education process—including teachers, pedagogues, parents, and health-care professionals—to the role of psychological stress as a cause of musicians' medical problems. Although individuals are motivated to learn and achieve in many different ways, students should be taught about effective, energy-preserving methods of practice, warming up and cooling down, and stretching. Furthermore, where the pressure of competition might motivate one promising student, it could easily discourage another, with serious consequences. In ensemble classes there should also be periodic breaks and avoidance of long rehearsals. In addition, music directors should be implored to be sensitive to the long-term effects on morale and health of players when subjected to recurring demeaning, negative, and personal criticisms. Teachers and parents need to be aware of all these variables and how best to monitor them in the interests of the development of physically and psychologically healthy young musicians. How an instrument is played, how the individual has been taught, and the musician's state of health are all essential elements in good music education in both preventive and remedial conditions. Accordingly, music education meets music medicine.

### *Measuring Muscle Tension*

As stated earlier, excess tension, regardless of its origins or reasons, is at the core of musicians' performance difficulties. Since muscle tension can be measured and followed through the technology of electromyography (EMG)—the measurement of electrochemical activity in muscles)—it would seem expedient to pursue research applications in this area further (Guettler, 1992; Heuser & McNitt-Gray, 1991, 1993, 1994; Kjelland, 2000; Levy, Lee, Brandfonbrener, Press, & Levy, 1992; Philipson, Sorbye, Larsson, & Kaladjev, 1990; Theim, Greene, Prassas, & Thaut, 1994). Furthermore, when use of EMGs is combined with various therapeutic interventions such as biofeedback, muscle tension can be significantly reduced (Druckman & Swets, 1988; Kjelland, 1986; Koehler, 1995; LeVine, 1988). The EMG biofeedback process involves the visual or auditory representation of muscle activity at a level that is beyond the subject's perceptibility. This representation can in turn be manipulated by the subject in a desired direction, thus altering muscle tension previously undetectable. Such technology and therapeutic intervention will always require the guidance and insight of an experienced teacher who is able to sort out fact from artifact in the process of nurturing genuine musical and technical development. Relaxation by itself without observable improvement in performance and/or physical comfort, and so on, is of no practical interest to music performance, just as it would not be to sports medicine. In fact, relaxation per se can be maladaptive if not carefully monitored in a performance context (Druckman & Swets, 1988; Kjelland, 1986). Studies such as that of Heuser and McNitt-Gray (1998) that involve, rather than exclude or bypass, the teacher in feedback process show much promise as a realistic adaptation of the traditional studio.

In this paradigm, the electromyograph contributes vitally important information for the teacher's use in diagnosing performance problems and prescribing alternative solutions.

## Applications in Teaching and Performance

The future of performing arts medicine resides in the interaction between medical science and pedagogy if the health maintenance and care of musicians are to become more consistent and effective. With a solid foundation of scientific data, injuries from misuse of the body can be consistently minimized, and injuries from well-intended but sadly misinformed teaching and outmoded schools of technique will yield to validated healthful teaching principles. (Many performance injuries unfortunately can and do happen because students follow the teacher's directions to the letter instead of doing what might work best for them individually.) Obviously there is no teacher who would seek to injure a student, but with scientific knowledge the teaching profession can be more confident of the validity of certain pedagogical directives. Without such systematic examination of methodology, the music-teaching profession is destined to perpetuate its quasi-religious debates over which approach is better, further dividing the profession into various pedagogical schools of thought that impede meaningful long-term progress.

Such scientific validation does not and will not come quickly or easily. However, much is already known through the efforts of the relatively young field of music medicine research. The information in this chapter alone furnishes much for teachers to benefit from and be conscious of on behalf of their students. Just the fact that medical problems are independent of age should alert the teaching profession to the need for a greater awareness and sense of responsibility for detecting and referring students at the inception of significant symptoms. Validation of teaching practice must start with a thorough examination of those performers who are proven to be successful models (Druckman & Swets, 1988). From this, norms can be established for comparison to assess the physiological and psychological needs of those still developing and/or relearning musical skills and habits. This in turn will lead to the development of sound physiological, biomechanical, and psychological practices in music teaching.

### *References*

Brandfonbrener, A. G. (1990). Joint laxity in instrumental musicians. *Medical Problems of Performing Artists, 5*, 117–119.

Brandfonbrener, A. G. (1995). Musicians with focal dystonia: A report of 58 cases seen during a ten-year period at a performing arts clinic. *Medical Problems of Performing Artists, 10*, 115–120.

Brandfonbrener, A. G. (1997). Orchestral injury prevention study. *Medical Problems of Performing Artists, 12*, 9–14.

Brandfonbrener, A. G. (2000). Joint laxity and arm pain in musicians. *Medical Problems of Performing Artists, 15*, 72–74.

Candia, V., Elbert, T., Altenmüller, E., Rau, H., Schafer, T., & Taub, E. (1999). Constraint-induced movement therapy for focal hand dystonia in musicians. *Lancet, 353* (9146), 42–43.

Cayea, D., & Manchester, R. (1998). Instrument-specific rates of upper extremity injuries in music students. *Medical Problems of Performing Artists, 13*, 19–25.

Cohen, L. G., Hallett, M., Geller, B. D., & Hochberg, F. (1989). Treatment of focal dystonias of the hand with botulinum toxin injections. *Journal of Neurology, Neurosurgery and Psychiatry, 52*, 355–363.

Druckman, D., & Swets, J. A. (Eds.) (1988). *Enhancing human performance: Issues, theories, and techniques.* Washington, DC: National Academy Press.

Fishbein, M., Middlestadt, S. E., Ottani, V., Straus, S., & Ellis, A. (1988). Medical problems among ICSOM musicians: Overview of a national survey. *Medical Problems of Performing Artists, 3*, 1–8.

Grey, J. (1991). *The Alexander technique.* New York: St. Martin's Press.

Guettler, K. (1992). Electromyography and muscle activities in double bass playing. *Music Perception, 9*(3), 2303–2309.

Harman, S. (1995). The evolution of performing arts medicine. In R. T. Sataloff, A. G. Brandfonbrener, and R. J. Lederman (Eds.), *Performing arts medicine* (2nd ed.) (pp. 1–18). San Diego, CA: Singular Press.

Heuser, F., & McNitt-Gray, J. L. (1991). EMG potentials prior to tone commencement in trumpet players. *Medical Problems of Performing Artists, 6*(2), 51–56.

Heuser, F., & McNitt-Gray, J. L. (1993). EMG patterns in embouchure muscles of trumpet players with asymmetrical mouthpiece placement. *Medical Problems of Performing Artists, 8*(3), 96–102.

Heuser, F., & McNitt-Gray, J. L. (1994). EMG changes in a trumpet player overcoming tone commencement difficulties: A case study. *Medical Problems of Performing Artists, 9*(1), 18–24.

Heuser, F., & McNitt-Gray, J. L. (1998). Enhancing and validating pedagogical practice: The use of electromyography during trumpet instruction. *Medical Problems of Performing Artists, 13*, 155–159.

Hochberg, F. H., Leffert, R. D., Heller, M. D., & Merriman, L. (1983). Hand difficulties among musicians. *Journal of the American Medical Association, 249*, 1869–1871.

Howard, J. A., & Lovrovich, A. T. (1989). Wind instruments: Their interplay with orofacial structures. *Medical Problems of Performing Artists, 4*, 59–72.

Kjelland, J. M. (1986). The effects of electromyographic biofeedback training on violoncello performance: Tone quality and muscle tension (Doctoral dissertation, University of Texas at Austin). *Dissertation Abstracts International*, 47A, 459A (University Microfilms No. DDJ86/09527)

Kjelland, J. M. (2000). Application of electromyography and electromyographic biofeedback in music performance research: A review of literature since 1985. *Medical Problems of Performing Artists, 15*, 115–118.

Koehler, W. K. (1995). The effect of electromyographic feedback on achievement in bowing technique (Doctoral dissertation). *Dissertation Abstracts International*, 55A, 2758A-9A (University Microfilms No. DA9502472)

Lederman, R. J. (1991). Focal dystonia in instrumentalists: Clinical features. *Medical Problems of Performing Artists, 6*, 132–136.

LeVine, W. R. (1988). Biofeedback in violin and viola pedagogy. In F. L. Roehmann and F. R. Wilson (Eds.), *The biology of music making: Proceedings of the 1984 Denver conference* (pp. 196–200). St. Louis: Magna Music Baton.

Levy, C. E., Lee, W. A., Brandfonbrener, A. G., Press, J., & Levy, A. E. (1992). Electromyographic analysis of muscular activity in the upper extremity generated by supporting a violin with and without a shoulder rest. *Medical Problems of Performing Artists, 7*(4), 103–109.

Mayers, H., & Babits, L. (1987, November). A balanced approach: The Alexander technique. *Music Educators' Journal*, 51–53.

Middlestadt, S. E., & Fishbein, M. (1988). Health and occupational correlates of perceived occupational stress in symphony orchestra musicians. *Journal of Occupational Medicine, 30*, 687–692.

Nelson, S. H. (1989). Playing with the entire self: The Feldenkreis method and musicians. *Seminars in Neurology, 9*(2), 97–104.

Newmark, J., & Lederman, R. J. (1987). Practice doesn't necessarily make perfect. *Medical Problems of Performing Artists, 2*, 142–144.

Philipson, L., Sorbye, M. D., Larsson, M. D., & Kaladjev, M. A. (1990). Muscular load levels in performing musicians as monitored by quantitative electromyography. *Medical Problems of Performing Artists, 5*(2), 79–82.

Spaulding, C. (1988). Before pathology: Prevention for performing artists. *Medical Problems of Performing Artists, 3*, 135–139.

Theim, B., Greene, D., Prassas, S., & Thaut, M. (1994). Left arm muscle activation and movement patterns in cellists employing a playing technique using rhythmic cueing. *Medical Problems of Performing Artists, 9*, 89–96.

Zaza, C. (1998). Playing-related musculo-skeletal disorders in musicians: A systematic review of incidence and prevalence. *Canadian Medical Association Journal, 158*(8), 1019–1025.

# PART II

## SUBSKILLS OF MUSIC PERFORMANCE

# 7

# From Sound to Sign

GARY E. McPHERSON & ALF GABRIELSSON

One of the most contentious issues in music pedagogy concerns when and how to introduce notation to a beginning instrumentalist. Most current teaching introduces musical notation very early in the process, perhaps because many teachers believe that beginners who are taught by ear will never reach the same level of reading proficiency as children who are introduced to notation from their earliest lessons. In contrast, proponents of the *sound before sign* approach argue that children will have difficulty learning to read notation unless their musical knowledge is sufficiently developed for them to be able to relate the sound of what they can already play with the symbols used to represent them. Our review of literature results in the identification of six principles that can be used to develop the complex range of skills needed for a child to become musically literate.

A curious contradiction in music pedagogy is that teaching practice is often in conflict with theories of instrumental teaching about how to introduce notation to a child. Whereas most children learning an instrument in Western styles of education are introduced to musical notation from their very early lessons, prominent instrumental teachers throughout history have advocated that ear playing should be emphasized before the introduction of notation.

## Historical Overview of Teaching Practice

Up until the mid-nineteenth century the teaching of instruments was regarded as a craft whereby knowledge was passed from one generation to the next by word of mouth, often through a form of musical apprenticeship. During this time, composers and teachers did not separate technical practice from more general musical skills; rather, the goal was to develop an all-round musician by inte-

grating technique with other aspects of general musicianship (Gellrich & Parncutt, 1998). Scales and arpeggios served as a means of learning the common vocabulary of musical language, and when passages were practiced in isolation from the music it was for the purpose of developing a range of musical skills, such as sight-reading, improvisation, and composition. For beginners, passages were often learned by ear rather than directly from reading musical notation. This process involved the rote learning of previously unknown pieces by imitating the teacher's model or the reconstruction of familiar music that had been internalized by repeated hearings or singing. (It should be noted that playing by ear is quite distinct from playing music from memory, which involves performing a piece that has been memorized as a result of repeated rehearsal of the notation. See further McPherson, 1995a, 1995b.) As skills developed, teachers would encourage their pupils to invent their own passages to develop the expressive and technical skills needed to master the musical language of the repertoire being learned (Gellrich & Sundin, 1993).

But changes occurred rapidly from around 1850 onward. The lithograph and high-speed printing machines were invented in 1818, and by 1830 it was possible to mass-produce relatively cheap scores in large quantities. Whereas few books of exercises had been published during the eighteenth century, from the second half of the nineteenth century they became common, and as a result the nature of learning to play changed quite dramatically. With access to new printed material the emphasis shifted from the development of skills in interpretation, improvisation, and composition to music as a reproductive art, with its resultant emphasis on technique and interpretation (Gellrich & Parncutt, 1998).

Once composers started to publish their exercises in method books, the oral tradition that had served musicians for centuries beforehand became irrevocably broken, and it was not long before pianists began to practice for long periods to develop specific technical skills that were often never applied when performing the literature. An unfortunate consequence was that exercises were often repeated endlessly during practice. Some students succeeded despite such mindless practice, while others, as we still see today, found it boring and uninspiring.

These comments highlight the changing emphasis in the use of technical exercises and how these were used by pianists to develop their craft. But this influence was soon felt on the full range of orchestral instruments. By the late 1800s, most beginning instrumental method books emphasized drill and technical material such as scales, rhythms, articulations, and finger exercises, with very little material of real melodic interest (Schleuter, 1997). One of the most consistent trends was to organize beginning method books according to the proportionality of note values whereby students were first taught a series of whole notes and whole note rests before graduating to half notes/rests, quarter notes/rests, and eventually eighth and sixteenth notes/rests (Schleuter, 1997). Modern method books place much more emphasis on learning repertoire, but the philosophy still tends to be based on the re-creation of music through a process of learning to read and develop technical skill from the very first lesson. The problem, according to Schleuter (1997), is that even today most instrumental method books associate fingerings with notation rather than fingerings with

sound, perpetuate the mathematical relationships of proportional note values, emphasize note naming and theoretical concepts more than perceptual understandings, and separate technical skill from the process of learning to play actual music. In essence, they reinforce the notion of performance as a specialist craft in which technical development and knowledge of notation are valued above all else.

### *Advocates of* Sound before Sign

Johann Heinrich Pestalozzi (1746–1827), who taught Swiss orphans after the French destruction of the canton of Unterwalden, was one of the first educators to profess that concepts should be taught through direct experience before the introduction of names or symbols (Abeles, Hoffer, & Klotman, 1994). Pestalozzi's ideas were introduced into North American schools by Joseph H. Naef, who had worked with Pestalozzi before emigrating around 1806. Naef founded an elementary school in 1809 that included music instruction as a basic subject and in 1830 outlined his "Principles of the Pestalozzian System of Music" at a meeting of the American Institute of Instruction. These included seven recommendations for teachers:

1. To teach sounds before signs and to make the child learn to sing before he learns the written notes or their names;
2. To lead him to observe by hearing and imitating sounds, their resemblances and differences, their agreeable and disagreeable effect, instead of explaining these things to him—in a word, to make active instead of passive in learning;
3. To teach but one thing at a time—rhythm, melody, and expression, which are to be taught and practiced separately, before the child is called to the difficult task of attending to all at once;
4. To make him practice each step of these divisions, until he is master of it, before passing to the next;
5. To give the principles and theory after the practice, and as induction from it;
6. To analyze and practice the elements of articulate sound in order to apply them to music; and
7. To have the names of the notes correspond to those used in instrumental music. (Abeles, Hoffer & Klotman, 1994, p. 11)

Other prominent music educators who were influenced by Pestalozzi's ideas include Lowell Mason, whose work as America's first public school music teacher helped to revolutionize music teaching. Mason was convinced that children should first experience music before learning to read notation. In England, similar calls have been a recurring theme (Pitts, 2000). One of the most articulate was Yorke Trotter (1914), who believed that a child "must first have the effect in his mind before he knows the symbols that should be used to express that effect" (p. 76).

From the second half of the twentieth century, Suzuki advocated that a child's musical education must begin as early as possible and be based on rote learning

instead of note reading during the early stages of training. Grounding his method on how children learn their *mother tongue*, Suzuki developed a highly sequenced system of teaching whereby children learn to "speak" (i.e., play) before learning to read. In Suzuki's view, notation is only appropriate when it can be beneficial, such as when the music becomes so complex that children need a symbol system to help prepare their performance (Landers, 1980).

Kohut's (1985) *natural learning process* is in many ways similar to the *mother tongue* approach of Suzuki. According to Kohut, young infants learn how to walk by watching and observing their parents and then attempting to imitate them. Through a process of trial-and-error practice children eventually learn to coordinate the muscles and cognitive processes necessary to walk with minimum conscious effort. In Kohut's view, the same natural learning process that children use to develop skills in everyday tasks can be used for learning an instrument. The first step is for the children to develop a repertoire of *mental blueprints* through listening to various models, whether they be a peer or teacher or model they hear on a recording. These mental images are then retrieved when the children attempt to imitate the image formed mentally. It is this natural, innate inclination to learn by listening and reproducing by ear, rather than from notation, that Kohut believes is the key to effective musical learning (see also Schleuter, 1997).

Other prominent instrumental teachers argue that because young children have difficulty learning more than one thing at a time, it is pointless to both expose them to the complex variety of technical skills needed to manipulate an instrument while at the same time asking them to read and comprehend notation (e.g., Schleuter, 1997). Gordon (1997) even recommends delaying the introduction of notation for as long as it takes for a student to develop an extensive aural vocabulary of tonal and rhythmic patterns according to his *music-learning theory*. He believes that poor sight-reading ability does not result from poor instrumental technique but poor aural skills. This is because notation does not always look the way fingers move when reading notation and that a premature emphasis on notation actually prevents a student from learning to hear and comprehend music internally. In Gordon's approach, playing by ear and improvising are essential ingredients for effective learning during the early prenotation stages of a child's musical development.

## Review of Psychological Literature

### *Ear Playing as Preparation for Music Literacy*

James Mainwaring, a fellow of the British Psychological Society and prominent music educator, was among the first researchers to study the acquisition of ear-playing skills. During his career, Mainwaring attempted to clarify the cognitive processes associated with musical ability, and his studies in this area led him to question teaching practice, which he believed placed far too much emphasis on technical skill and the mechanized recall of music from the printed score (1931, 1947). He was convinced that learning an instrument should "proceed from sound

to symbol, not from symbol to sound" (1951c, p. 12), based on his belief that the development of music literacy should involve processes similar to learning to speak and then read a language.

At the heart of Mainwaring's (1941, 1951a, 1951b, 1951c) concept of musicianship is the capacity of being able to "think in sound," which occurs when a musician is able to produce the mentally imagined sound, whether by playing by ear, improvising, or reading from musical notation. When reading musical notation, "thinking in sound" involves an ability to inwardly hear and comprehend notation separately from the act of performance. This concept is depicted in Figure 7.1, which highlights the important distinction between seeing notation and responding mechanically to produce the notated sound (i.e., working from symbol to action to sound) in contrast to seeing the musical notation and being able to hear the notation inwardly before reproducing it on an instrument (i.e., working from symbol to sound to action). The former method Mainwaring (1951b) believed to be typical of most instrumental teaching practice. However, it was the latter way of working that he advocated as the most efficient and effective means for developing a young player's overall musicianship .

For Mainwaring (1951a), the most musical way of teaching an instrument is to continually link sound with action. In the beginning stages of development, this means encouraging students to reproduce simple known tunes by ear, before they learn to read these songs from musical notation.

Mainwaring's ideas can be contrasted with teaching that introduces notation from the earliest lessons. Children as young as three can develop a basic understanding of the pitch elements of musical notation and relate this to the piano keyboard (Tommis & Fazey, 1999), so the issue is not whether young children can learn to read notation at an early age but whether they should learn it so early.

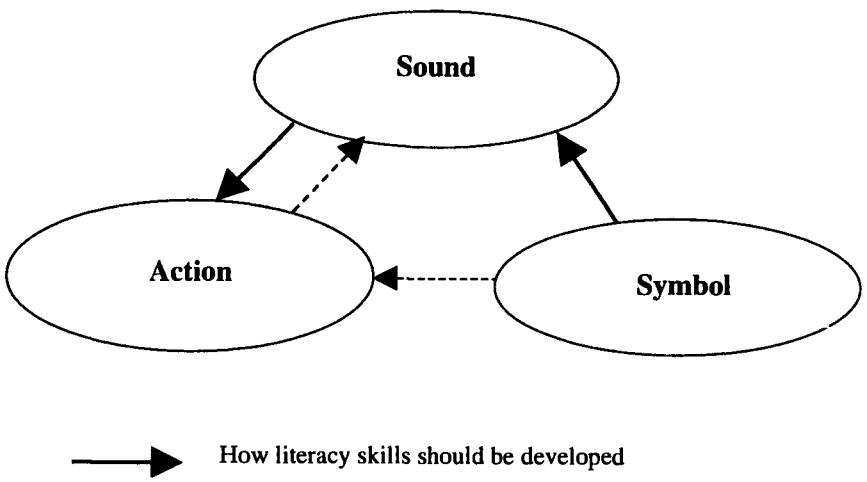



Figure 7.1. Model of literacy development (adapted from Mainwaring, 1951b, p. 20).

One way of addressing this issue is to compare the acquisition of literacy skills in language with music, given that prominent methods such as the Suzuki approach are based on the premise that learning to speak and then read a language involves similar processes to learning to become literate in music. From a psychological perspective, Sloboda (1978) provides the most definitive starting point when he maintains that

> no-one would consider teaching a normal child to read while he was at a very early stage of learning spoken language. Yet it seems the norm to start children off on reading at the very first instrumental lesson without establishing the level of musical awareness already present. (p. 15)

According to Sloboda, musical sensitivity should be developed before the introduction of notation because "without some musical knowledge a beginner has no expectancies which can be used in reading" (p. 15).

By the time children enter first grade, they typically possess a solid command of the mechanics of their own language and are able to perceive and use an extensive vocabulary in excess of 5,000 words. They are therefore capable of beginning the process of learning to associate familiar words with the symbols used to represent them (Bruning, Schraw, & Ronning, 1999). By sounding out individual letters, spelling patterns, and words, children begin to recognize the spelling, sound, and meaning of particular words and over time are able to comprehend the meaning of complex sentences (Adams, 1990). If these processes are duplicated in music, then the process that leads to the introduction of musical notation would involve teaching children to perform familiar tunes by ear before they learn to sound out tunes they already know using notation. This is why Mainwaring (1951c) and others (e.g., Gordon, 1997; Kohut, 1985; Schleuter, 1997; Suzuki, 1983) advocated teaching known tunes by ear in familiar keys before the notation of these melodies are introduced to the child.

However, studies with children who are learning to read highlight the importance of metalinguistic capabilities—"knowledge about the uses of print, how print represents sounds, how words are formed, how sentences are put together, and how sentences become stories or reports" (Bruning, Schraw, & Ronning, 1999, p. 242). This form of awareness enables children to map the visual symbols of written language onto oral language and to thereby create meaning. Learning how to comprehend musical notation involves similar processes in that it requires knowledge about the use of notation, how notes represent sounds, how phrases are formed, how melodies are put together, and how melodies become compositions. Like reading text, learning to become musically literate involves more than linking visual cues with instrumental fingerings; otherwise instrumentalists become button pushers "to whom notation only indicates what fingers to put down rather than what sounds are desired" (Schleuter, 1997, p. 48).

This line of reasoning is underpinned by psychological literature (e.g., Dowling & Harwood, 1986) that shows that listeners do not perceive music on a note-to-note basis but as patterns according to well-known gestalt principles for perceptual organization. These principles have been observed in various types of musical performance. For example, when constructing a piece by ear, musicians

work according to structural units, by rewinding their mental tape recorders to the beginning of a phrase every time they make a mistake (Bamberger, 1996, 1999; McPherson, 1993). Working in structurally meaningful units (e.g., phrases) is also how advanced instrumentalists rehearse music from notation (Gabrielsson, 1999; Gruson, 1988; chapter 10 in this volume).

Based on her extensive research with children, Bamberger (1996, 1999) concludes that young learners must first gain experience in playing musical patterns before they learn to decontexualize these patterns into individual notes, because the "units of perception" that novices intuitively attend to when processing music are "structurally meaningful entities such as motives, figures, and phrases," not individual notes (p. 42). Bamberger's advice can be compared with how beginners normally correct performance errors when reading notation. In the very early stages of learning a new piece, they will often stumble over individual notes and continue playing, sometimes so slowly or hesitantly that they no longer perceive the music they are attempting to play as a complete phrase or melody (McPherson, 1993; McPherson & Renwick, 2000). Beginning instrumentalists find it baffling and frustrating to learn music in this way, especially in situations where they are taught to "start out by listening for, looking at, and identifying the smallest, isolated objects" and to "classify and measure with no context or functional meaning" (Bamberger, 1996, p. 43). Asking beginners to focus on notation too early, according to Bamberger (1996), means asking them to "put aside their most intimate ways of knowing–figures, felt paths, context and function" (p. 44). Too early an emphasis on notation can therefore lead to a decreased aural sensitivity to the natural unified patterns that children spontaneously observe when listening to music.

Related evidence shows that notation should not be taught in isolation of perception because of the danger that children will be forced to make choices beyond their competence and will tend to do so in contradiction of their perception. This was demonstrated by Davidson, Scripp, and Welsh (1988), who asked a group of 12-, 15-, and 18-year-olds to sing and then notate songs such as "Happy Birthday." Students who were receiving private instrumental tuition plus music theory classes tended to work conceptually. They defied their perception of the song to produce a transcription, for example, that started and finished on middle C (rather than F) because they believed that if "Happy Birthday" started on C, then it must be in C major and should therefore end on C. As these students got older, they relied more on what they know theoretically about music rather than what they heard and knew perceptually, with the result that they made surprisingly inaccurate transcriptions of familiar melodies. In contrast, students exposed to an approach to instrumental learning that integrated aural development with literacy skills were more inclined to work perceptually and their notations were more accurate and generally improved across the age groups.

To extend these findings we can consider how easily beginners are able to attend to and integrate knowledge from a number of sources. Like reading text, learning to decode musical notation is a complex skill that in the beginning stages involves processes that are not instant and automatic but rather conscious and

deliberate (Samuels, 1994). However, unlike reading, where a novice can focus on decoding the visual symbols, learning to perform music involves two competing and nonautomated tasks: learning to read musical notation at the same time as learning to manipulate an instrument. In this type of situation, there are constraints on the amount of information beginners can think about at any one time, how long they will be able to hold it in their mind before it is lost, and how quickly they can process new information (Shaffer, 1999; Sweller, van Merrienboer, & Paas, 1998). To complicate this even further, vision tends to dominate and inhibit the processing of signals from other modalities (Posner, Nissen, & Klein, 1976; Smyth, 1984). Consequently, if children's attention is focused on reading notation, they may have few cognitive resources left to devote to manipulating their instrument and listening to what they are playing.

A common criticism of "sound before sign" approaches is that children will have difficulty integrating the knowledge required to read music when notation is introduced and therefore never achieve the same level of reading fluency as children exposed to notation from their first lesson. However, this folklore view may be misguided, given the literature cited earlier and also studies that show that children who learn "by ear" do not become less successful music readers (Glenn, 1999; McPherson, 1993; Sperti, 1970). For example, Glenn (1999) studied the progress of a group of sixth-grade beginning string students over a period of a full school year. The beginners were randomly assigned to one of two groups. The first group was taught without notation (i.e., using modeling and verbal explanation) for the first three months, after which notation was introduced, even though some time continued to be devoted to playing repertoire "by ear." The second group was taught with notation after a brief 2-week introduction to their instrument. For this group, the emphasis was on matching note names with instrumental fingerings and working through a method book. At the end of the study, the students were videotaped to compare their performance achievement. Results indicate that students in the "ear-playing" class performed as well as or better than students in the notation-based class on all performance evaluations, including the sight-reading measure. Importantly, however, students who were taught by ear in the initial stages were more likely to continue learning their instrument and were reported by their teacher to be more motivated and to enjoy their playing more than the other group. Based on her results, which support literature on how difficult it can be for beginning instrumentalists to learn two competing and nonautomated tasks at the same time, Glenn concludes that students should achieve a degree of automaticity in their playing before they are introduced to musical notation.

### *An Integrated Model*

To this point, our review has focused on studies that provide evidence of the importance of ear playing in the beginning stage of learning an instrument. Clearly, this literature has many gaps, but there is one project that helps clarify the beneficial effects of ear playing on subsequent development. Based on Mainwaring's (1941, 1951a, 1951c) assertion that ear playing is the most funda-

mental of all the performance skills and that playing by ear can exert a positive influence on other styles of performance, McPherson (1993, 1995a, 1995b, 1996, 1997; McPherson, Bailey, & Sinclair, 1997) undertook a 3-year longitudinal study with a sample of 101 instrumentalists to examine relationships among five different aspects of musical performance and 16 environmental variables that prior research had suggested may influence their development. At the start of his study, the students were in school grades 7 to 12 (ages 11 to 18) and had been learning their instrument for approximately 2 to 6 years.

In order to reduce the large amount of information to a manageable and interpretable form, factor analysis was used to collapse the environmental variables into four distinct categories. First, *early exposure* described variables related to the quality and quantity of the students' early exposure to music, such as when and how they started learning, how long they had been playing, and whether they had learned another instrument before commencing. Second, *enriching activities* grouped variables associated with the students' reports of how frequently they played by ear and improvised, as well as whether they were electing classroom music at school in which composing was an important component. Students who reported higher average levels of daily practice were more likely to report higher levels of these types of informal activities. Third, *length of study*, included variables associated with the period the students had been playing their instrument and taking private lessons. Fourth, *quality of study* included the subjects' report of their interest and participation in various forms of singing, the number of ensembles they were performing with, and a report of how often they mentally rehearsed music in their minds away from their instrument.

McPherson reviewed available literature in order to develop a theoretical model that he could test empirically using results from the high school students' performances and interviews (see Figure 7.2). Results from the path analysis indicate that the model fits nicely with the data he obtained. This supports evidence of a number of causal relationships among the five skills and four environmental factors. As indicated in Figure 7.2, the model implies both direct and indirect effects. Direct effects are the unmediated relation between two variables (e.g., the influence of playing by ear on improvising is shown by a direct arrow). Indirect effects are the relation between two variables that are mediated by one or more other variables (e.g., the influence of playing by ear on improvising is also mediated through its effect on sight-reading and playing from memory). (A complete explanation is available in McPherson, 1993, and McPherson et al., 1997).

*Early exposure* exerted a positive influence on the instrumentalists' ability to play by ear and no direct influence on their ability to sight-read. As expected, *enriching activities*, which included variables associated with how often the instrumentalists played by ear and improvised during their practice, had a stronger impact on the skill of playing by ear and subsequently on the musicians' ability to improvise, than it did on sight-reading. *Quality of study* had only a weak influence on the abilities of playing by ear and sight-reading, while *length of study* had a moderate influence on the instrumentalists' ability to sight-read, as well as a strong direct influence on their ability to perform rehearsed music.

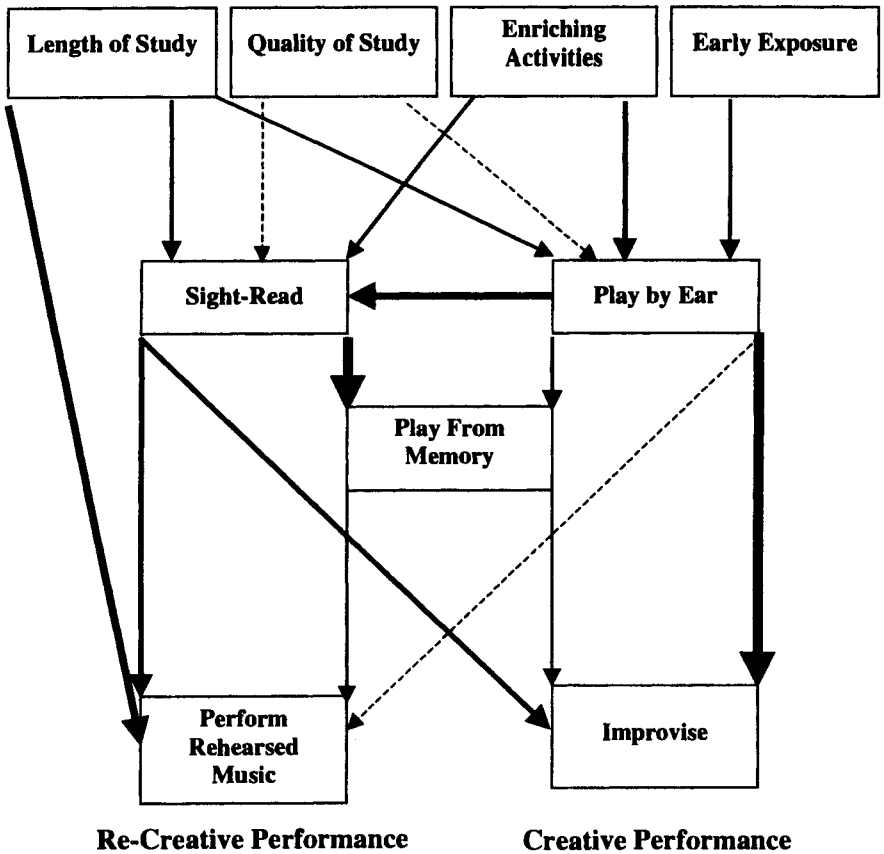



Figure 7.2. Path analysis of the McPherson theoretical model. In the simplified path diagram, path coefficients have been replaced by lines of differing thickness. Thicker lines represent stronger influences and thinner lines weaker relationships from one variable to the next as shown by the direction of the arrow.

Path analysis indicates that the skill of performing rehearsed music was most heavily influenced by the variables associated with *length of study*, plus the players' ability to sight-read. In contrast, the skill of improvising was most strongly influenced by the capacity of the musicians to play by ear (which had a strong direct path straight to improvising and a moderate indirectly effect via sight-reading and playing from memory).

Analysis of comments made by the less capable players show that they tended to use strategies that were independent of their instrument when preparing to play by ear (e.g., "While I listened to the tape I was trying to think in terms of what letter names the notes might be"). In contrast, more capable musicians were able to connect what they heard on the taped performances with the instrumen-

tal fingerings needed to execute these thoughts (e.g., "I just sing it in my head while I finger it through on my instrument"). Interestingly, McPherson's findings are in many ways similar to Davidson et al.'s (1988) explanation of how conceptual information can cloud perceptual awareness. In addition, the strong relationship McPherson reports between the students' ability to play by ear and sight-read is consistent with previous evidence (Luce, 1965; Priest, 1989) that ear playing may be more beneficial to the development of musicianship than sight-reading and also more receptive to training.

The results of McPherson's research support psychological and pedagogical assertions that the skill of playing by ear helps student musicians learn to coordinate ear, eye, and hand and to perform on their instrument what they see in notation and hear or imagine in their mind. Although continued refinement and testing of alternative models is needed to confirm the findings, this evidence suggests that teachers should recognize the importance of ear playing as an important facet of training that enhances overall musical growth and that provides for more enjoyable and meaningful learning.

## Implications for Teaching

### *Developmental Level*

There are a number of unresolved issues that arise from the discussion so far. Clearly, the dynamics of teaching play an important role in how quickly children will become literate, as do a variety of other intrapersonal and environmental catalysts that shape their development (see related chapters in this volume). However, since children start learning an instrument at widely different ages, some comments are appropriate about developmental issues. At the heart of arguments for the sound-before-sign sequence is the premise that sensory and/or motor experiences should always precede the learning and use of symbols. Prominent theories of human development (e.g., Piaget, Bruner, Gardner) are in accord with this view, as are leading classroom music education methodologies (e.g., Dalcroze, Orff, Kodály). Hargreaves (1986), summarizing philosophies in music education that go back as far as Rousseau (1762), reinforces this view when he states:

> The intuitive experience and enjoyment of music should come first, such that the latter acquisition of formal musical skills occurs *inductively*, that is, as an integral growth of the child's experience. A good deal of traditional music education has worked deductively: the formal rules have been taught in the abstract, for example, through verbal description or written notation, rather than in the practical context of making the sounds themselves. (p. 215)

Bruning, Schraw, and Ronning (1999) explain that there are obvious developmental limits to how early children can be taught to read their language, but that most children start formal reading around the age of 6, after they have gained

a great deal of prior experience in the use of their native tongue and reached the mental age necessary to maximize their success as readers. By this time their home experiences "will have prepared them with about 25 hours of storybook experience and perhaps 200 hours of general guidance about the form and nature of print" (Adams, 1994, pp. 89–90).

Obviously, a great deal of music learning takes place prior to formal lessons, and even very young children will come to their lessons having heard a vast amount of music in their everyday lives. However, the extent of a 6-year-old child's knowledge about how music can be represented in notation will typically be nowhere near the level of his or her understanding of how language can be represented in print. For this reason, in addition to the psychological literature cited earlier, there appears to be enough evidence to speculate that an emphasis on reading musical notation should not occur until a child is at least 6. Up until this age, teachers should emphasize rote teaching of pieces learned by ear in order to establish the important ear-to-hand coordination skills that provide a foundation for introducing notation later. In general, we believe that this rule should be extended to older age groups, although the length of time before introducing notation can be shortened, depending on how quickly they develop the necessary aural awareness and technical security on their instrument for them to be able to comprehend notation.

### *Prenotation Activities*

In our view, two types of activities provide the foundation from which literacy skills can then be developed. First, during the early months (or weeks for older learners) of training, familiar pieces should be taught by rote and by repeated listening (many current methods books now include recorded examples of the literature being studied). In this stage the emphasis should be on encouraging beginners to sing and then play familiar repertoire on their instrument. Singing (either mentally or out aloud) is useful because it helps to establish a correct mental model that can guide children as they translate what has been memorized into the instrumental fingerings needed to perform it on an instrument. Singing should be a common and natural part of all early lessons. However, it should be noted that students need to progress from singing to linking singing with the fingerings needed to perform the piece on their instrument. Mentally and outwardly rehearsing (singing and fingering) a piece by ear develops a child's ability to experience sounds through gestures and touch, which Galvao and Kemp (1999) believe help integrate the body with the mind. If a teacher feels competent, then solfège can also be used to reinforce this transition from memorization (i.e., singing to a neutral syllable such as *la*) to comprehension (i.e., singing using sol-fa syllables and/or instrumental fingerings).

After these initial experiences, the notation of familiar pieces that children can already perform should be introduced gradually and in ways that challenge them to consider how and why notation is used in music. Asking students to invent their own notation for a piece they can already play provides a useful window from which to view their conceptual development. Such activities also

"prompt the classification, organisation, and connections that enable the child to transform the concrete experience into one that can be represented in icons or symbols" (Gromko, 1994, p. 146). In general, a child's earliest experiences in learning notation should occur in isolation from the act of playing. Gradually, and stimulated by a variety of challenging activities, children should start to make connections between the sounds they can already play and the traditional notation used to represent them. Scanning and pointing to the melody of a piece being performed by the teacher, chanting or clapping rhythm patterns, and singing or fingering (either mentally or out aloud) phrases or melodies can each help develop the aural awareness needed for students to comprehend notation in meaningful ways.

In general, acquiring competence in reading and interpreting musical notation is best achieved via a three-way process of gaining fluency in playing music, then reading it, and then putting the two together.

### *Introducing Notation*

After these steps have been reinforced through various activities, a child can begin to integrate playing with reading. It is probably best to start by introducing the notation of pieces a child can already perform, then move to pieces that he or she has not learned to play but already knows, before moving on to new, unfamiliar repertoire that demand more-sophisticated levels of processing.

Six general principles can guide teachers as they strive to develop the wide range of skills needed for their students to become musically literate. These principles are modeled on the work of Bruning et al. (1999, pp. 258–259), whose comments on teaching children to read are reinterpreted for beginning instrumental instruction:

1. *Approach the reading of notation as a meaningful activity.* Reading musical notation is a meaning-making act. Although children need to learn to decode notation and know, for example, that a quarter note is one beat in length, their developing skills need at all times to be linked to structurally meaningful entities such as phrases and melodies rather than individual notes. Children can lose track of the purpose of learning to read notation when the focus of their attention is on decoding signs and visual representations in terms of individual note values and theoretical relationships (e.g., understanding the proportions of note values or scale-pitch relationships). They need to be continually reminded of the basic reasons for acquiring music-reading skills—learning, communicating, and enjoyment.
2. *Take a broad perspective on literacy development.* Functional literacy, in terms of knowing where to put one's fingers as a result of seeing a visual cue on the score, represents only a limited understanding of how to comprehend musical notation. The goal should be to instill competence in the use of notation for all forms of thinking and expression. Focusing the child's attention on the flow of a melody, the expressive detail in a composition, or simply how to make a passage sound different helps to foster a more sophisticated awareness of the broader pur-

pose of musical notation and how it can be used to help an instrumentalist think, reason, and communicate musically. An important aspect of good teaching, therefore, is to focus on meaningful tasks that help develop both musical and cognitive competencies.

3. *Help young readers move toward automatic decoding.* As shown in the sight-reading chapter in this volume, skilled readers are able to quickly and accurately discern key features in the visual appearance of the score and then instigate the required motor action to perform them on their instrument. When these actions are automatic, it is easier to direct attention to higher level expressive elements that enhance one's performance. Explicit instruction in how to decode notation is important. Replaying passages to improve fluency is obviously important for improving performance ability, but it does not follow that drilling activities and endless repetition will lead to automaticity. Guided meaningful reading of a range of musical pieces in which the teacher directs a student's attention to the relevant technical and expressive characteristics embedded in the work itself helps to instill confidence in performance and knowledge that can be applied when performing other literature.
4. *Draw on children's domain and general knowledge.* All children have knowledge about music before they commence instruction. Reading notation is more likely to be meaningful when it connects with a child's prior experiences of music. For example, introducing the notation for a piece a child already knows and wishes to learn can provide a powerful incentive that helps the student learn to read more quickly, as well as decipher any new elements that are being introduced by the teacher. Encouraging young beginners to invent their own notations to represent pieces they can already perform provides them with the metamusical awareness that will enhance their progress toward understanding why traditional notation looks and works the way it does. Teachers should provide multiple ways for young learners to interact with and learn about notation so that it actually means something to them.
5. *Encourage children to develop their musical knowledge.* Effective teachers draw their students' attention to the many dimensions of musical notation and encourage them to reflect on their performance and to come up with different interpretations. Learning to read musical notation is a multifaceted process that hinges on knowledge activation, memory, and attention. Aural awareness enables students to map the visual symbols of the musical notation into the required actions that produce the sounds on their instrument. Memory facilitates comprehension of the larger scale structures of what is being performed, and attention enables the musician to focus on individual signs, phrases, and expressive detail in order to perform musically.
6. *Expect children to vary widely in their progress toward fluent reading.* No two children will ever be at exactly the same level of musical development. Some will struggle with learning to read notation, while others will pick up notational skills relatively easily. Unfortunately, children who do not pick up reading skills early in their development are often not given the types of remedial attention needed to correct their deficiencies. This does not mean, however, that they cannot

develop into competent musicians. As with any sophisticated skill, learning a musical instrument involves a complex mix of abilities, and every student will have strengths and weaknesses that impact on the rate at which, and, in turn, the likelihood that, they will continue learning.

## Conclusion

The focus in this chapter has been on ear playing, which we believe teachers may be underestimating in terms of its benefits for developing literacy skills. Based on evidence cited in this chapter, we conclude that emphasizing notational skills too early can lead to a decreased sensitivity to the unified patterns that children spontaneously observe when listening to music. Notational skills should therefore never be taught in isolation from perception. In our view, stressing notation, with few opportunities to perform music by ear, or rote learning, with equally few opportunities to develop reading fluency, restricts overall musicianship and the types of skills needed for a musician to succeed long-term. Rather, we advocate a more integrated approach, where performing music by ear serves as preparation for literacy development in the beginning stages of musical involvement, and where performing with and without notation is encouraged during all subsequent levels of development.

**Acknowledgments.** We thank Daniel Kohut, Stanley Schleuter, and Jack Taylor, who each provided very helpful suggestions and insightful comments on an earlier version of this chapter.

### *References*

Abeles, H. F., Hoffer, C. F., & Klotman, R. H. (1994). *Foundations of music education* (2nd ed.). New York: Schirmer.

Adams, M. J. (1994). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Bamberger, J. (1996). Turning music theory on its ear. *International Journal of Computers for Mathematical Learning, 1*(1), 33–55.

Bamberger, J. (1999). Learning from the children we teach. *Bulletin of the Council for Research in Music Education, 142,* 48–74.

Brunning, R. H., Schraw, G. J., & Ronning, R. R. (1999). *Cognitive psychology and instruction* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.

Davidson, L., Scripp, L., & Welsh, P. (1988). "Happy Birthday": Evidence of conflicts of perceptual knowledge and conceptual understanding. *Journal of Aesthetic Education, 22*(1), 65–74.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition.* Orlando, FL: Academic Press.

Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.) (2nd ed.), *The psychology of music* (pp. 501–602). San Diego, CA: Academic Press.

Galvao, A., & Kemp, A. (1999). Kinaesthesia and instrumental music instruction: Some implications. *Psychology of Music, 27*(2), 129–137.

Gellrich, M., & Parncutt, R. (1998). Piano technique and fingering in the eighteenth and nineteenth centuries: Bringing a forgotten method back to life. *British Journal of Music Education, 15*(1), 5–24.

Gellrich, M., & Sundin, B. (1993). Instrumental practice in the 18th and 19th centuries. *Council for Research in Music Education, 119*, 137–145.

Glenn, K. A. (1999). Rote vs. note: The relationship of working memory capacity to performance and continuation in beginning string classes. (Doctoral dissertation, the University of Northern Colorado). *Dissertation Abstracts International*, 60/04, 1010. (University Microfilms No. 9927738)

Gordon, E. E. (1997). *Learning sequences in music: Skill, content and patterns.* Chicago: GIA.

Gromko, J. E. (1994). Children's invented notations as measures of musical understanding. *Psychology of Music, 22*(2), 136–147.

Gruson, L. M. (1988). Rehearsal skill and musical competence: Does practice make perfect? In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 91–112). Oxford: Clarendon Press.

Hargreaves, D. J. (1986). *The developmental psychology of music.* London: Cambridge University Press.

Kohut, D. L. (1985). *Musical performance: Learning theory and pedagogy.* Englewood Cliffs, NJ: Prentice-Hall.

Landers, R. (1980). *The talent education school of Shinichi Suzuki: An analysis.* Athens, OH: Ability Development.

Luce, J. R. (1965). Sight-reading and ear-playing abilities as related to instrumental music students. *Journal of Research in Music Education, 13*(2), 101–109.

Mainwaring, J. (1931). Experiments on the analysis of cognitive processes involved in musical ability and in music education. *British Journal of Educational Psychology, 1*(2), 180–203.

Mainwaring, J. (1941). The meaning of musicianship: A problem in the teaching of music. *British Journal of Educational Psychology, 11*(3), 205–214.

Mainwaring, J. (1947). The assessment of musical ability. *British Journal of Educational Psychology, 17*(1), 83–96.

Mainwaring, J. (1951a). Psychological factors in the teaching of music: Part 1: Conceptual musicianship. *British Journal of Educational Psychology, 21*, 105–121.

Mainwaring, J. (1951b). Psychological factors in the teaching of music: Part II: Applied musicianship. *British Journal of Educational Psychology, 21*(3), 199–213.

Mainwaring, J. (1951c). *Teaching music in schools.* London: Paxton.

McPherson, G. E. (1993). Factors and abilities influencing the development of visual, aural and creative performance skills in music and their educational implications. (Doctoral dissertation, University of Sydney, Australia). *Dissertation Abstracts International*, 54/04-A, 1277. (University Microfilms No. 9317278).

McPherson, G. E. (1995a). The assessment of musical performance: Development and validation of five new measures. *Psychology of Music, 23*(2), 142–161.

McPherson, G. E. (1995b). Redefining the teaching of musical performance. *Quarterly Journal of Music Teaching and Learning, 6*(2), 56–64.

McPherson, G. E. (1996). Five aspects of musical performance and their correlates. *Council for Research in Music Education, 127*, 115–121.

McPherson, G. E. (1997). Cognitive strategies and skills acquisition in musical performance. *Council for Research in Music Education, 133*, 64–71.

McPherson, G. E., Bailey, M., & Sinclair, K. (1997). Path analysis of a model to describe the relationship among five types of musical performance. *Journal of Research in Music Education, 45*(1), 103–129.

McPherson, G. E., & Renwick, J. (2000). Self-regulation and musical practice. In C. Woods, G. B. Luck, R. Brochard, F. Seddon, and J. A. Sloboda (Eds.), *Proceedings of the Sixth International Conference on Music Perception and Cognition*. Keele, UK: Keele University, Department of Psychology. CD-ROM.

Pitts, S. (2000). *A century of change in music education: Historical perspectives on contemporary practice in British secondary school music*. Aldershot, UK: Ashgate.

Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review, 83*(2), 157–171.

Priest, P. (1989). Playing "by ear": Its nature and application to instrumental learning. *British Journal of Music Education, 6*(2),173–191.

Rousseau, J. J. (1762). *Emile, or On education*. London: Dent.

Samuels, S. J. (1994). Toward a theory of automated information processing in reading, revisited. In R. B. Ruddell, M. R. Ruddell, and H. Singer (Eds.) *Theoretical models of processes of reading* (4th ed). (pp. 816–837). Newark, DE: International Reading Association.

Schleuter, S. (1997). *A sound approach to teaching instrumentalists* (2nd ed.). New York: Schirmer.

Shaffer, D. R. (1999). *Developmental psychology: Childhood and adolescence* (5th ed.). Pacific Grove, CA: Brooks/Cole.

Sloboda, J. A. (1978). The psychology of music reading. *Psychology of Music, 6*(2), 3–20.

Smyth, M. M. (1984). Perception and action. In M. M. Smyth and A. M. Wing (Eds.), *The psychology of human movement* (pp. 119–152). London: Academic Press.

Sperti, J. (1970). Adaptation of certain aspects of the Suzuki method to the teaching of the clarinet: An experimental investigation testing the comparative effectiveness of two different pedagogical methodologies. (Doctoral dissertation, New York University). *Dissertation Abstracts International*, 32/03, 1557. (UMI No. 7124833)

Suzuki, S. (1983). *Nurtured by love* (2nd ed.). New York: Smithtown.

Sweller, J., van Merrienboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychological Review, 10*(3), 251–296.

Tommis, Y., & Fazey, D. M. A. (1999). The acquisition of the pitch elements of music literacy skills by 3–4 year-old pre-school children: A comparison of two methods. *Psychology of Music, 27*(2), 230–244.

Yorke Trotter, T. H. (1914). *The making of musicians*. London: Herbert Jenkins.

# 8

# Improvisation

BARRY J. KENNY & MARTIN GELLRICH

Depending upon its sociocultural function, the term improvisation incorporates a multiplicity of musical meanings, behaviors, and practices. A feature common to all improvisation, however, is that the creative decisions of its performers are made within the real time restrictions of performance itself. Improvisation is therefore considered to be a performance art *par excellence*, requiring not only a lifetime of preparation across a broad range of musical and nonmusical formative experiences, but also a sophisticated and eclectic skills base. The chapter reflects on psychological models and their attempts to simulate improvising processes and constraints, the means by which improvisers acquire performance skills, improvisation as part of a larger, co-collaborative creative endeavor, recent studies highlighting the benefits of improvisation in a learning situation, and improvisation as a means of revitalizing Western education. Practical implications and an integrated model for learning to improvise are discussed in the final section.

When improvisers talk about their music, they often draw upon linguistic metaphors grounded in communication or rhetoric (Berliner, 1994; Monson, 1996). The culturally agreed upon constraints that make this spontaneous rhetoric possible distinguishes improvisation from most other forms of music making. Of these constraints, the most important is time itself, which determines that improvised creation must occur simultaneously with its performance (see Pressing, 1998). Such temporal constraints necessitate a series of efficient mechanisms designed to facilitate improvising in real time. From a psychological perspective, these constraints fall into two broad categories—*internally* (i.e., psychologically) and *externally* (i.e., socioculturally) generated.

Aside from the more obvious cognitive (i.e., memory) and physiological (i.e., motor skills) constraints that affect improvisation, the most important *internal*

constraint is the *knowledge base* (see Figure 8.1). This warehouse of previously learned material is what the performer knows and brings to the performance, such as the "musical materials and excerpts, repertoire, sub skills, perceptual strategies, problem-solving routines, hierarchical memory structures and schemas, generalized motor programmes" (Pressing, 1998, p. 53) that have been acquired and developed through conscious deliberate practice. The knowledge base used by improvising musicians typically involves the internalization of source materials that are idiomatic to individual improvising cultures. Examples of such pedagogic source material include the Persian *Radif*—a "guide to improvisatory techniques, formal patterns, and overall structure of performances" (Nettl, 1998, p. 14)—and in jazz the transcribed solos of distinguished musicians.

Referents, however, are associated with or specific to a particular performance: the external, culturally supplied forms that assist with the transmission of improvised ideas. These points of departure include a range of musical and nonmusical (i.e., graphics) stimuli that, whether sounded or not, ultimately become deeply embedded in a musician's internalized creative resources (Nettl, 1974). The musical referents of jazz, for example, are its cyclical, often 32-bar song structures (i.e., jazz standards), its chords (and rules that govern treatment of their extensions), and its characteristic rhythmic patterns (Pressing, 1998). Two of the referent's most important functions are its ability to limit improvisational choices according to appropriate guidelines and its role in building perceptual paradigms for listener appreciation (Sloboda, 1985).

The latter of these two functions is particularly important, in that most improvisations are filtered through formal structures already familiar to listeners. In contrast to knowledge bases, which performers are not typically aware of during performance (because they are internalized and automated), referents influence improvisers more directly, providing the formal and musical material unique to each improvisation. However individual one artist's interpretation of the jazz standard "Body and Soul" may be, for example, it is still likely to share many similarities with another artist's version, thereby providing a perceptual degree of commonality for listeners. The same cannot be said for each artist's knowledge base, which may be as unique as each musician's experiences and personalities.

Skilled improvisers are adept at manipulating listeners' predetermined expectations of referents for expressive purposes through, for example, ironic interpretation or adherence to or denial of expectation. The gratification or frustration of these expectations in turn generates musical emotion, which in itself may play a key role in determining musical meaning (Meyer, 1973; Narmour, 2000). The perceptual frameworks that govern listener expectation in improvised music may range anywhere between concrete notated compositional forms and the amorphous musical shapes and sounds that are deeply embedded in any culture's collective unconscious (Monson, 1996). As Westendorf (1994) observes, the perceptual frameworks of jazz improvisation, for example, include not only groups of individual notes but also generalized contour shapes, where any number of melodic fragments can potentially trigger standardized patterns of expectation (see also Dowling & Harwood, 1986).

As suggested in Figure 8.1, the two key constraints of improvisation—knowledge bases and referents—work together to generate new musical structures. This diagram also illustrates the spillage between the two. Referents, for example, are likely to become part of knowledge bases through prolonged exposure and repetition. Similarly, listeners make synchronous connections between their perception of the initial referent (through prior exposure) and its modified variant—the improvisation itself.

### *Flow States, Risk Taking, and Kinesthesia*

> Pepper forgot everything, just blew and blew, shaking all over—a state reminiscent of the horses of the gods—being possessed to the extent that they became their specific orisha, assuming all their particular dance movements and behaviours. This forgetting of oneself is a state that many improvisers strive to attain. (Floyd, 1995, p. 139)

The ecstatic state described in this quote by Floyd, where improvising artists surrender to the creative moment itself, is well documented in the psychological literature and by no means unique to musical creativity. Jazz pianist and practicing psychologist Denny Zeitlin admits this to be his state of mind when playing the piano (Csikszentmihalyi & Rich, 1997). As Csikszentmihalyi and Rich explain, these *peak experiences* or *flow states* assist improvisers, to move not only beyond the literal texts of referents but also beyond their own cognitive limits in non–flow states. Furthermore, this quasi-narcotic flow state may be one of the most important reasons that motivate improvising musicians to persevere with their craft, despite the often-adverse conditions it is produced under. Once possessed by the moment, musicians begin to forget personal problems, lose critical self-consciousness, lose track of time, and eventually feel that the activity in which they are engaged is worth doing for its own sake (Csikszentmihalyi

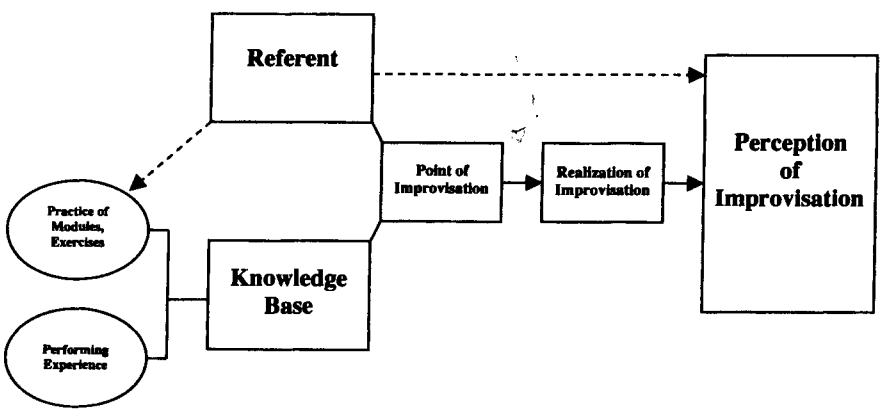



Figure 8.1. Referents, knowledge bases, and listener expectation.

& Rich, 1997). Aside from unlocking creative improvising in one of its most emancipated forms, flow states may therefore also play a key role in motivation and hence a predisposition or inclination toward further artistic development.

As a creative endeavor that occurs in real time, improvisation often involves the necessary disguising and making musical sense of mistakes. The old jazz adage that it's not a mistake if you play it twice is symptomatic of this approach. Mistakes suggest a more pervasive undercurrent that informs all improvisational creativity—risk taking. For many improvisers, risk taking provides a self-induced state of uncertainty where repetition and predictable responses become virtually impossible. In a landmark publication on this issue, Sudnow (1978) documented the frustrating process he underwent when acquiring and applying improvising skills. On his journey to become a professional musician, Sudnow reflects on the difficulty of acquiring knowledge bases from aural sources, the technical constraints of particular instruments (i.e., the black and white keys of a piano), the effect these constraints have on improvised response, and the relationship between spontaneously created material and improvised filler.

Sudnow's most important finding was that a conscious application of his internalized knowledge base resulted in what he terms "frantic" playing, where each distinct chord in the harmonic series of the song being improvised on triggered its own preconceived strategic plan, a plan inefficiently spilling over into and frustrating the plans of the next adjacent chord. Sudnow found that one way to unlearn this triggered response was to jettison his knowledge base, to let his hand go wherever it wished in a process where the intuitive shape of his hands and ears guided musical responses. Once he started taking more risks, Sudnow found more "right" notes falling under his fingers and his playing at last began to emulate the relaxed, idiomatic, and coherent sound typical of more-experienced players.

While Sudnow's introspective methodology is open to criticism, his findings nevertheless draw attention to significant psychological processes that complement and inform the automation of knowledge bases. Performance that incorporates flow states and risk taking may in fact hold the key to achieving optimal levels of musical communication, providing a clue as to why some musicians are able to access their knowledge bases more fluidly and creatively than other similarly skilled but less inspired improvisers. Berliner (1994) refers to the attainment of this heightened level of performance as being "within the groove [where] improvisers experience a great sense of relaxation, which increases their powers of expression and imagination. They handle their instruments with athletic finesse, able to respond to every impulse" (p. 389).

Surprisingly little research has been undertaken in the related area of *kinesthesia*, the sense of where parts of the body are with respect to one another. Our current understanding of complex muscular interactions and their relationship to instrumental performance is at best rudimentary; what little is known suggests a far subtler integration of mind and body than previously thought, where "what appears to be a simple act of throwing or catching an object actually involves the interaction of several feedback mechanisms" (Galvao & Kemp, 1999, p. 133). This suggests that some of the research on improvisation that makes

connections between simple motoric movement and musical structures, thereby often informing technical pedagogy, may have to be rethought in terms of a more holistic conception of technique. Only then may we more fully understand why some musicians are able to move beyond technical automation to arrive at a more direct and meaningful form of communication.

### *Theoretical and Generative Models of Improvisation*

What are improvisers thinking about at the precise moment of creation? In short, we still do not really know. As Johnson-Laird (1988) observes, in order to improvise efficiently and idiomatically, the subconscious knowledge base processes that generate improvisation need to be sufficiently automated and submerged "without any internal representation of an intermediate form" (p. 211). The fact that improvisers themselves cannot access their own subconscious processes at the moment of creation poses enormous practical problems for researchers. In addition, the creative impetus for improvisation often depends on volatile performance variables (e.g., interaction with audience, fellow musicians, acoustic considerations), all of which are extremely difficult to replicate under controlled experimental conditions or reliably account for with postevent analysis. To date, such practical considerations have precluded any serious examination of the thinking-aloud verbal protocols of improvising musicians.

In order to better understand and replicate the theoretical constructs that generate improvisation, researchers have attempted to model its salient features. Johnson-Laird's model (1991), itself based on a computer simulation of an improviser's knowledge base, sheds light on the mechanisms that assist with spontaneous creation. He argues that if the knowledge base is sufficiently internalized (in long-term memory) and automated (through practice and performance experience), the resources used to generate surface melody are ultimately freer to focus on developing coherence and structural unity. One of his model's strengths is that its ideology is grounded in cognitive load theory (i.e., the finite cognitive processing capacity of multiple tasks), which, given the multiple constraints that govern improvisation, seems to make good sense.

Significantly, Johnson-Laird's model shares a number of conceptual similarities with Chomsky's (1968) linguistic models. At the first or deepest level, improvisers commit basic structures (i.e., chord theory, prelearned formulas) to memory. At the second level, improvisers make aesthetic feedback decisions that concern the structure of the referent, such as which significant notes are to be targeted. At the third or surface level, improvised melody is generated (Johnson-Laird, 1991). In a computer program designed to test the theory, Johnson-Laird found that surface-level functions (i.e., improvised melody) required less computational processing power than the internalized functions that generated them.

Clarke's (1991) three-stage cognitive model of improvisation outlines a hierarchy of thought processes. These are employed proportionately according to the level of structure demanded by the improvising genre and/or the artistic inclinations of the improviser. These conceptual thought processes articulate various proportions of freedom and constraint and can be extended from a gen-

eralized understanding of jazz genres to include most global improvising genres. Clarke's three categories, which loosely resemble Kernfeld's (1981) much-cited theoretical genres of jazz, can be summarized as follows.

1. *Repertoire selection*: Formulaic improvising characteristic of bebop. In its crudest form, improvisers automatically associate prelearned formulas with either particular chords (i.e., C major seventh) or chord classes (i.e., major sevenths in general). When integrated with other types of improvisation, repertoire selection provides a potent means for binding disparate thematic material, a momentary resting point for improvising musicians (when inspiration momentarily fails) and a unique body of recurring motivic material that identifies particular musicians.
2. *Hierarchical*: Song form–generated improvising structures (i.e., referents) characteristic of both bebop and hard bop. These fixed forms supply not only the chord sequences that frame most improvised responses in Western improvised music but also the melodic structures associated with their original form (i.e., the original melody of "Body and Soul"). In its most emphatic form—*melodic paraphrase*—improvised responses are tantamount to variations or permutations of the original referent.
3. *Motivic*: Chain-associative improvising characteristic of modal and free jazz. In motivic improvising, motives develop linearly, with each new unit of improvisation drawing upon improvised material produced either immediately before the improvised event or within recent memory. The most overt form of motivic improvising is a series of melodic sequences in transposition.

It is interesting to read Clarke's hierarchical categories of improvisation in the light of Johnson-Laird's model. Taken together, they not only account for the basic generative mechanisms of improvisation but also further illuminate our current understanding of what constraints govern different improvising styles and forms. Figure 8.2 represents an attempt to combine these two theories together with the first author's own concerns with performance and group variables. These additions include the obvious starting point for the improvisation—the referent—in addition to other factors that affect initial input, such as performance variables (i.e., audience participation, venue, acoustics) and group creative input.

In this combined model, it is evident that *repertoire selection* processes the initial referent as a series of disjointed sections, with each section drawing a selected response from the basic structures. Both *hierarchic* and *motivic* improvising initially processes the referent (i.e., R) more holistically. This composite entity naturally requires more intermediary consideration than repertoire selection (which essentially bypasses the evaluation stage). It is also much more likely to frame improvisational responses to the degree that they will be perceived by listeners as similar to the original referent (i.e., R1). The main difference between hierarchic and motivic improvising occurs at the point of production. Once the referent has been processed hierarchically, new material (often another chorus

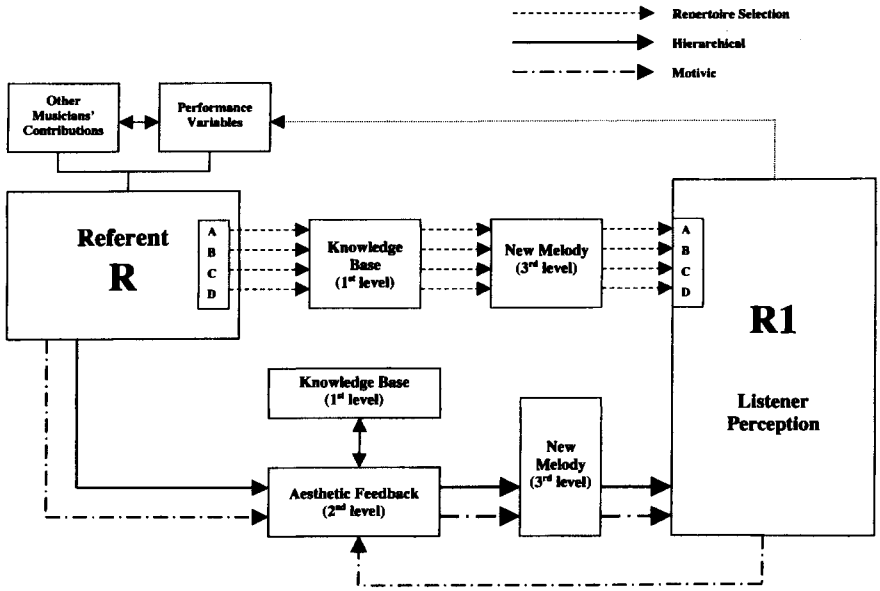



Figure 8.2. A combined generative model of improvisation.

of the same referent) is required to continue the process. In motivic improvising, any material generated as an end-point response—no matter how dissimilar to the original referent—can potentially feed back into the evaluative stage, thereby generating new improvisations. The diagram also suggests that all improvisation perceived by listeners in some way feeds back into the overall process through audience feedback, which in turn affects the future creative decisions of group participants.

### *A Model of Mental Processes During Improvisation*

As the preceding models demonstrate, improvisation operates most effectively when realized through an efficient network of constraints, all of which assist improvisers in making effective, appropriate, and meaningful choices within the restrictions of real-time improvising. But how is the potential to improvise further constrained by our own finite physical and cognitive resources? As Pressing (1998) observes, improvisation operates as "an interruptible associative process based on the ongoing evaluation of previous musical events" (p. 56). What physical and cognitive constraints, then, is this feedback mechanism subject to?

The second author of this chapter has developed a speculative model of cognitive processes during improvisation where any combination of eight potentially

different types of processes can be observed to occur. Improvisers typically shift from one process to another but cannot combine two or more simultaneously (see Heuer, 1996).

1. *Short-term anticipation*: At any point in the improvisation, musical events are anticipated within a time interval we estimate to be around 1 to 3 seconds. However, these anticipated notes cannot be sounded for a minimum of 0.3 second after the decision has been made (Gellrich, 2001b; see also psychological research on the refractory period: Welford, 1952; and from reading musical scores: Sloboda, 1985; Goolsby, 1994).
2. *Medium-term anticipation*: Musical events that occur within a 3- to 12-second time span (i.e., the next phrase or period) may be anticipated and projected into the future. (Again, these times are estimates and have not been substantiated by experimental evidence. The length of the time span will also depend on the length of the next musical phrase or period.)
3. *Long-term anticipation*: Projection of long-term plans for the remainder of the improvisation.
4. *Short-term recall*: Musical events that have occurred over the last few seconds can be recalled, in a process where concentration is focused on prior events (Gellrich, 2001a).
5. *Medium-term recall*: Musical events that have occurred within the last 4, 8, or 16 measures can be recalled so as to provide an accurate recollection of the previous musical phrase.
6. *Long-term recall*: Improvisers are able to recall the entire improvisation from its genesis up to the present moment.
7. *Flow status*: Improvisers are able to concentrate solely on what is being created at that particular moment.
8. *Feedback processes*: Musical ideas for future projected improvisation may be gathered from that which can be previously recalled. An example is an initially unintended ("wrong") note, recalled from previous performance, that the improviser decides to reiterate. Such recollections may further include a substantial amount of musical material held in medium- and long-term recall. This concept of feedback may be further extended to include the ongoing evaluation of musical events in the light of information held in medium- and long-term recall. An example of this kind of feedback occurs in jazz improvisation, where tones (of the scale) associated with the next adjacent chord in the series are sounded over the present chord. However temporarily dissonant such a practice may at first appear, it proves to be an effective means of preparing and linking adjacent chords.

In the course of performance, improvisers potentially draw on any number of these eight cognitive processes, provided that their decisions are made quickly and as a series (i.e., from one note to the next but not simultaneously). The second author's unpublished interviews with expert improvisers and personal analysis of his own improvisations suggest that the most commonly occurring activities are short- and medium-term anticipation and flow status. The other five cognitive processes generally occur when the improviser is able to master enough conscious control for their execution—for example, during

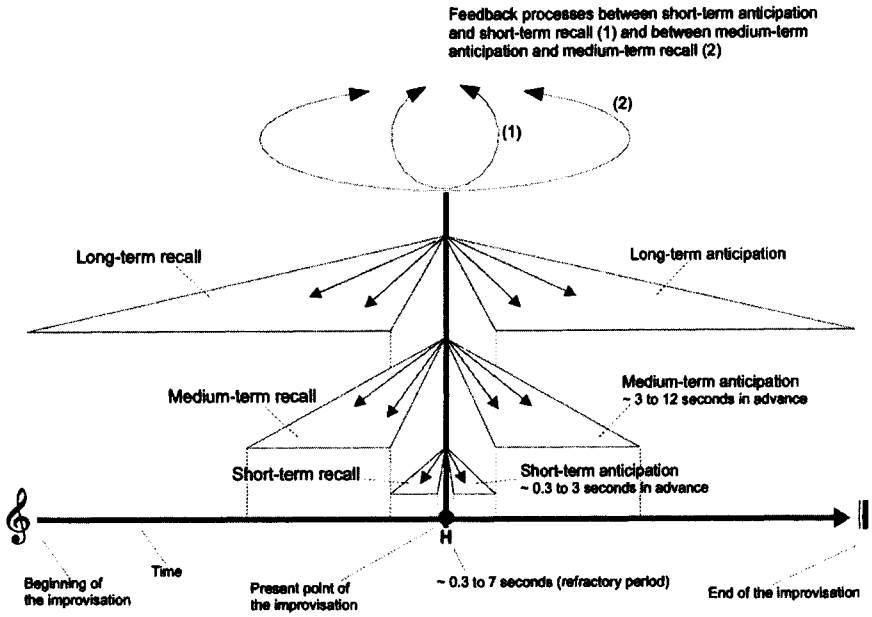



Figure 8.3. A model of mental processes during improvisation. Feedback processes between anticipation and short-term recall (1) and between medium-term musical concept and medium-term recall (2).

slower phrases or phrases with pauses and when prelearned patterns are articulated automatically.

## Practical Implications

### *The Challenges of Teaching Improvisation*

How does one teach a student to move beyond the text of knowledge to the fluid context in which it can be most fruitfully applied? Two pedagogical approaches—*deliberate practice* and *transcendence*—help answer this question. Much of the existing research on the pedagogy of improvisation so far has been concerned with deliberate practice. Transcendence can be understood as a heightened state of consciousness that moves beyond the confines of (thereby often jettisoning) the accumulated knowledge base itself. It is a state of consciousness that, like deliberate practice, can be encouraged and cultivated at the outset of an improviser's development; it need not be delayed until the final stages of an artist's development, as deliberate-practice research implies. While transcendence states, akin to the flow states discussed earlier, are more difficult to define and research, they nevertheless provide a cogent alternative to

deliberate practice and play an important role in any well-rounded practice regime (see concluding section of this chapter).

Placing deliberate practice and transcendence states in opposition is not to say, however, that one does not require or ultimately lead to the other. As Berliner (1997) observes, improvisation involves a "lifetime of preparation in the rigors of musical thinking" so that musicians are able to "respond artfully, as well as spontaneously, when improvising" (p. 37). The skill acquisition and developmental processes detailed in Berliner (1994), however, move well beyond the *individual* learning of knowledge bases to include a wider, collaborative learning environment. In this environment, social interaction produces true innovation and musical meaning. The problem for educators is how to replicate these complex sociocultural phenomena in institutionalized educational settings.

### *Deliberate Practice*

The traditional approach has been to instill as much improvising technique as possible (usually in a one-on-one setting) in the hope that it might equip *individuals* with sufficient material to cope with the unpredictable nature of *group* improvising. Much evidence supports deliberate practice as a necessary means of acquiring improvising skills, and hence expertise (Weisberg, 1999; Pressing, 1998; Lehmann & Ericsson, 1997). Through the correct levels of motivation and challenging situations, such individually tailored practice provides scaffolded targets for the accruement of improvisational skills. Ways in which a musician may practice deliberately include "working with a teacher in a directed situation, but also by aural absorption of examples of expert performance, study of theory and analysis, and interactive work in peer group ensembles during rehearsal and performance" (Pressing, 1998, p. 48).

One of the main aims of deliberate practice is to encourage improvisational expertise through the intensive development of internalized knowledge bases. The difference between expert and nonexpert improvisers is in how sophisticated, automated, and personalized these structures become (Kratus, 1989). Novice improvisers, for example, tend to access materials from the knowledge base in a diachronic and literal fashion by repeating prelearned motives parrot fashion or out of context, whereas experienced improvisers are able to make sophisticated hyperconnections between prelearned material. Borko and Livingston (1989) demonstrate this principle effectively in their observations of expert teachers of mathematics. Not only are the knowledge bases of these teachers more synthetic and dynamic than those of less experienced teachers, but they are also able to better anticipate questions and genuinely respond to the dynamics of the larger learning situation. For musicians, these hyperconnections can be effectively developed through well-established deliberate-practice routines—for example, practicing chord voicings in all inversions and spacings or motivic formulas practiced in all keys (Pressing, 1998).

Jazz is still the primary method of teaching improvisation in Western education and the chord-scale formulaic method the most widely practiced means of achieving this. It is therefore worthwhile to assess the success of this method as

a case study for teaching improvisation in Western educational settings as a whole. To put it simply, the chord-scale method attempts to constrain an improviser's choice of individual melodic notes to an array of scales or modes suggested by the predefined chord sequence of the referent. For example, minor seventh chords—such as D minor seventh (D, F, A, C)—are said to best suggest Dorian modes, such as D Dorian (D, E, F, G, A, B, C), simply because Dorian modes represent supersets of this particular chord type (i.e., minor sevenths). The methodology's basic aim, like deliberate practice itself, is to provide improvisers with as much working material as possible. Significantly, its ideological precepts support the ideology that informs generative models of jazz improvisation based upon multiple constraints, where chords are seen to elicit any number of potential internalized responses. As Birkett (1995) explains, however, "While [the chord-scale method] certainly gives students notes to play, it does not seem to offer any reasons for playing anything in particular" (p. vi). Offering a knee-jerk approach to each successive chord in a series, the methodology's randomness significantly contradicts the "tension and resolution relationships" suggested by extended harmonic passages (Birkett, 1995, p. 25) and, as Sudnow's (1978) experience shows, ultimately only produces a frantic and disconnected style of playing.

This methodology's often-uncritical acceptance points to a more serious trend in the use of improvisation in educational settings—that of theoretical modeling inhibiting improvised response (Kenny, 1999). It is evident that a great deal of improvisation that emanates from many performing institutions today appears to be more concerned with preparing improvisers for every foreseeable eventuality than with developing individual improving voices. The effects that 30 years of this type of pedagogy has had on the language of jazz, for example, is articulated by Lou Donaldson:

> All players are sounding alike today. They're all working out of Oliver Nelson's book. They play mechanical sequences of changes that will fit anything. When they get to a chord change, they skate through it. They work out clusters of notes, whole-tone patterns and things, to get through it. . . . They don't have a feeling for tonal centres in music anymore, or they just improvise on the harmony in ways that have nothing to do with the song. (Berliner, 1994, p. 280)

Donaldson's quote illustrates the downside of repetition and imitation as an appropriate means of acquiring improvising skills. As Martinez, Malbran, and Shifres (1999) assert, once a mental image has been shaped "incorrectly" for the first time, subsequent repetitions fail to clarify its structure, instead only serving to reimprint the original unidiomatic errors. In other words, initial learning experiences may play a crucial role in determining how creatively such knowledge bases are ultimately applied. The implication for improvisational pedagogy is that it may be difficult to unlearn material once it has been compounded through countless repetitions and private practice. A solution to this seeming impasse is not simply to build connections between already-internalized material but also to encourage individual improvising voices from the outset (Birkett,

1995). As Berliner's classic anthropological study of the jazz scene demonstrates, extensive aural immersion, semistructured experimentation and active participation in improvising genres almost always take place prior to the systematic acquisition of theoretical principles (Berliner, 1994).

### *Children's Play and Group Improvising*

Many of the mechanisms used by human beings in everyday problem-solving tasks are improvisatory in nature, a concept that has recently been articulated as "everyday creativity" (Sawyer, 1999). Children's play provides a fascinating window into improvised creation in one of its most unmediated forms. For example, Baker-Sennett and Matusov (1997) asked six Grade 2 and 3 girls to make up their own version of *Snow White* with minimal authoritarian intervention or guidance. Left to their own devices, these children first set the material itself aside to concentrate on the social dynamics that would ultimately facilitate the improvising process, such as the procedures they might use to resolve conflicts. Once a collaborative atmosphere had been established, the children proceeded to cooperatively improvise much of the play's structure in character. Baker-Sennett and Matusov found that a lack of authoritarian intervention produced as much cohesion and efficiency as a teacher-led control situation, if not more. Aside from the enormous educational benefits of involving children in the creation of new knowledge, these children were also aware of their privileged role as creative participants. Displaying similar artistic temperaments to adult creators, they paid special attention to the dramatic consequences of their choices and endlessly debated and workshopped the best possible solution to a problem.

A similar relaxation of authoritarian control in improvisation is discussed by Smith (1998), who investigated Miles Davis's creation of a ritualized performance space. Davis's success as a mentor and bandleader was based on similar principles to those exhibited by the children, especially his ability to exploit the semistructured possibilities of group creativity. Just as a lack of predictable control provided a point of focus for the children making up *Snow White*, musicians in Davis's groups were impelled, through Davis's refusal to provide certainty, to engage in a heightened form of group cohesion and creativity. In the absence of traditional hierarchical (top-down) leadership structures, on the one hand these musicians were freer to actively participate in creative contributions, while on the other they needed to listen and defer to one another's projections more closely than before. Not surprisingly, these interchanges gave rise to a subtle and efficient form of communication that paradoxically focused even greater attention on Davis himself than before, cementing his pivotal role as group mentor and instigator of new ideas.

Baker-Sennett and Matusov's and Smith's research draws attention to the social interactive variables of improvisation, without which the development of individual knowledge bases is relatively ineffectual. This research also demonstrates that improvisation can provide a potent means of harnessing intrinsic motivation in group (educational) settings. As Sternberg (2000) observed, educators may be selling many apparently unpromising students short by unduly

emphasizing "memory and analytical abilities" above "creative and practical abilities" (p. 255). By concentrating more on their given abilities (as opposed to capabilities), students may start to perform better across a broad range of educational and personal areas "because they can use their abilities more effectively, and because the greater interest of the material better motivates them to learn" (Sternberg, 2000).

The overtraining of predictable learned responses is perhaps the greatest shortcoming when learning to improvise. Most improvised performance, however, takes place in dynamic, group-based settings where

> each has to listen and respond to the others, resulting in a collaborative, and intersubjectively generated performance . . . [where] no one acts as the director or leader, determining where the performance will go; instead, the performance emerges out of the actions of everyone working together. (Sawyer, 1999, p. 194)

One of the suggestions that emanated from these findings is that group performance activities must complement solitary practice. While far greater amounts of solitary practice may prove beneficial in fostering the technical and theoretical principles of notated music, improvisation requires a greater emphasis on performance (and group experience) itself. After all, improvisational creativity most often ultimately takes place in a performance environment, not in the practice room, and the ability to react to and generate music from dynamic and unpredictable variables is one of the distinguishing features of improvisation.

### *An Integrated Model for Learning to Improvise*

One of the greatest challenges that face improvising musicians is their need to attend to several motoric and musical aspects simultaneously while improvising. Among other things, these aspects include harmony, patterns, melodies, form, musical expression, coordination of both hands (piano), and rhythm.

Although psychological research suggests the possible division of conscious control between two different aspects (Heuer, 1996; Pashler & Johnston, 1998), many improvisers that the second author has interviewed report that they normally can only consciously monitor one aspect at a time. This suggests that while one aspect is monopolizing conscious attention, the others must proceed unconsciously in the background. As improvisations unfold, musicians shift concentration from one aspect to another (Heuer, 1996), with conscious focus held on each aspect for only a fraction of a second. Such findings have profound consequences for the teaching and learning of improvisation. Because of these limitations on conscious control, the teaching of improvisation needs to be divided into different areas, all of which must be developed systematically and in parallel (Gellrich, 1995). Only after having mastered the ability to consciously control each aspect separately can improvisers control all aspects simultaneously and unconsciously with the added ability to switch between them.

Two basic stages in the acquisition of improvising skills can be distinguished, both of which can be understood in terms of a linguistic analogy. In the first stage

of learning a language words and grammatical rules are acquired, and in the second students explore their various possibilities of combination and application. Improvisers similarly need to first master the hardware of improvisation: patterns, parts of melodies, chord progressions, modulations, voicings, counterpoint, and the coordination between chord progressions and melodic patterns. Only then can the software of improvisation be developed—systematic rules that assist with constructing melodies, phrases, and larger musical ideas, working with motifs, and establishing relationships among different parts of the improvisation. Both the hardware and the software of improvisation, which together play a key role in the formation of the knowledge base, must be practiced systematically and separately. The more musical equivalents of words and grammatical rules an improviser is able to acquire and master, the richer the ensuing language of the improvisations will be.

Once a musician has assimilated newly acquired material such as a motive or a chord progression into the knowledge base, it is necessary to apply it as soon as possible in a practical context. Each particular musical aspect of improvisation requires different working time and attention to detail. Players of melodic instruments, for example, need to intensively practice melodic patterns in different tonalities (Gellrich, 1992; Gellrich & Parncutt, 1998). This not only builds more complex connections between preexisting materials already in the knowledge base but also is an effective means of kinesthetic reinforcement (Gellrich, 1992; Gellrich & Parncutt, 1998). These patterns ultimately lie under the fingers to the extent that musicians are able to divert attention away from technique altogether. Similar precepts can be applied to guitar, organ, and piano players and their automation of chord voicings.

For improvisation to remain vital and truly spontaneous, it is important not only that the knowledge base is constantly updated and sophisticated but also that improvisers learn to transcend it. Only then are improvisers able to unconsciously avoid predictable responses and react spontaneously to less predictable variables such as other musicians' knowledge bases and audience variables. In order to avoid predictable responses, musicians need to devote practice time to exercises and activities that encourage creativity and risk taking and, most important, replicate the improvising environment, where mistakes and disaster recovery occur on a regular basis. For example, if a pianist accidentally lands on a dissonant or nonfunctional chord, this can become the starting point for a series of exercises that revolve around resolving such chords with the optimal voice leading. In this way mistakes ultimately become the catalysts for creativity, where new accidental figures and chord progressions potentially enter the ever-increasing richness and complexity of the knowledge base. One of the best forums for risk taking and self-challenge is group improvisation, where improvisers need to constantly reassess projected responses in relation to the creative contributions of other individuals and the collective group.

A further aspect worth exploring is associative improvisation. Inspiration for improvisation can be derived from a number of artistic resources other than music, such as dance, movement, poetry, films, comics, and pictures. There

are clearly a multiplicity of factors beyond the purely musical that affect and shape the creative impulses of great improvisers. Seminal tenor saxophonist Joe Henderson illustrated this with his holistic experience of improvisation:

> I've probably been influenced by nonmusical things as much as musical things. I think I was probably influenced by writers, poets? . . . You know how to use quotation marks. You know how you quote people as a player. You use semicolons, hyphens, paragraphs, parentheses, stuff like this. I'm thinking this when I am playing. I'm having a conversation with somebody. (Floyd, 1995, p. 141)

Aside from its self-evident risk-taking benefits, such nonmusical, associative improvising may prove a useful tool in integrating improvisational hardware and software.

One of the greatest challenges that face improvisers in the current climate of global improvisation is to develop an individual improvising voice. With so many competing and culturally diverse styles to choose from, the question is where to begin. In the era before global communication, such problems of identity were never encountered by, for example, cultures where improvisational practices developed in comparative geographic isolation and across much greater historical time frames. So as to avoid the disorienting effect of competing influences, we suggest that in the initial stages students learn to improvise in one particular style, in a similar way that children naturally gravitate toward expressing themselves in certain musical forms. In this way, developing improvisers learn to initially speak in their mother tongue. Students should also study compositions in this style and learn to use elements of these as a means of increasing the range of ideas for their own improvisations. The skills honed under such a control situation can then be extended to improvising across a number of different styles. After having learned to improvise first in a particular style and then in a number of different styles, improvisers learn to develop both an individual and richly eclectic voice. Figure 8.4 demonstrates the connection between the different areas of teaching and learning improvisation discussed here.

## Conclusion

The implications that the latest research in music psychology have for the pedagogy of improvisation are wide-ranging and interdisciplinary. This research suggests an integrated learning approach, combining the best aspects of deliberate practice theory with established cultural practices such as risk taking and group creativity. As for the discipline of music psychology itself, there are promising signs of interdisciplinary integration and cooperation, manifest in two recent monographs dedicated to improvisation, both of which reflect its dynamic, multifaceted, and interdisciplinary concerns (Nettle & Russell, 1998; Sawyer, 1997). A possible area for future improvement could be a greater level of cooperation and information sharing between music theory and music psychology, both of which are essentially responsible for generating theories on improvisation, theories that

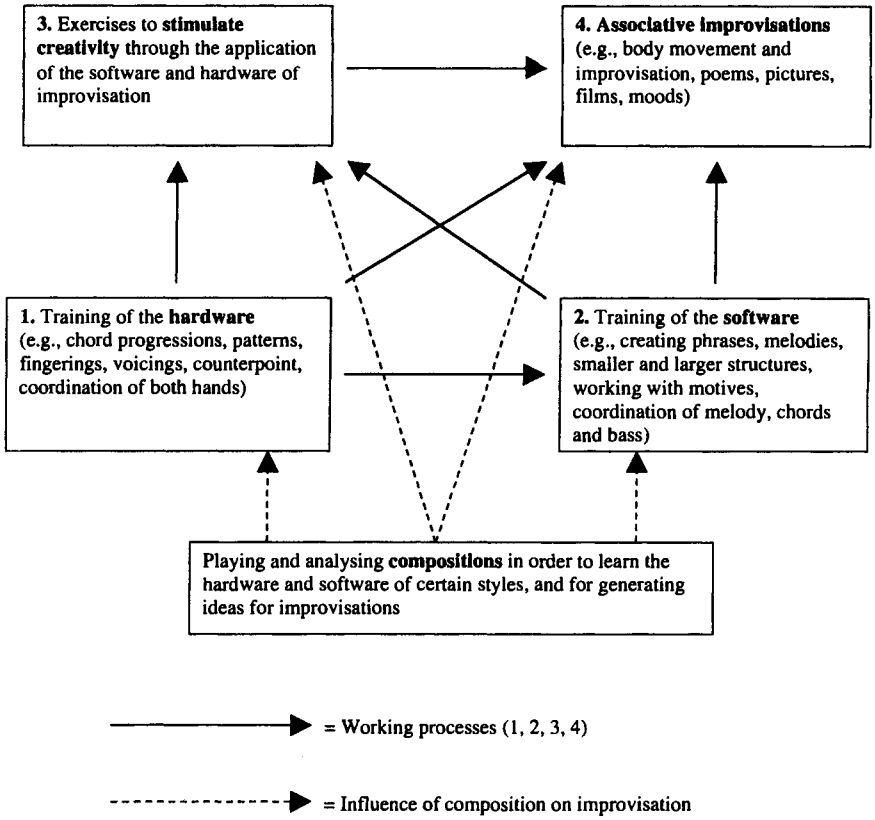



Figure 8.4. An integrative model of learning improvisation.

often translate into pedagogy. Another promising area is the psychological study of group dynamics. As improvisation is essentially about the collaborative creation of shared rhetoric, further research in this area can only improve the little we currently know about the cognitive aspects of group creativity.

### *References*

Baker-Sennett, J., & Matusov, E. (1997). School performance: Improvisational processes in development and education. In R. K. Sawyer (Ed.), *Creativity in performance* (pp. 197–212). Greenwich, CT: Ablex.

Berliner, P. (1994). *Thinking in jazz : The infinite art of improvisation.* Chicago: University of Chicago Press.

Berliner, P. (1997). Give and take: The collective conversation of jazz performance. In R. K. Sawyer (Ed.), *Creativity in performance* (pp. 9–41). Greenwich, CT: Ablex.

Birkett, J. (1995). *Gaining access to the inner mechanisms of jazz improvisation.* Doctoral dissertation, Open University, Milton Keynes.

Borko, H., & Livingston, H. (1989). Cognition and improvisation: Differences in mathematics instruction by expert and novice teachers. *American Education Research Journal, 26,* 473–498.

Chomsky, N. (1968). *Language and mind.* New York: Harcourt, Brace & World.

Clarke, E. F. (1991). Generative processes in music. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 1–26). Oxford: Clarendon Press.

Csikszentmihalyi, M., & Rich, G. (1997). Musical improvisation: A systems approach. In R. K. Sawyer (Ed.), *Creativity in performance* (pp. 43–66). Greenwich, CT: Ablex.

Dowling, W. J., & Harwood, D. W. (1986). *Music cognition.* New York: Academic Press.

Floyd, S. A., Jr. (1995). *The power of black music: Interpreting its history from Africa to the United States.* New York: Oxford University Press.

Galvao, A., & Kemp, A. (1999). Kinaesthesia and instrumental music instruction: Some implications. *Psychology of Music, 27*(2), 129–137.

Gellrich, M. (1992). *Üben mit Lis(z)t–Wiederentdeckte Geheimnisse aus der Werkstatt der Klaviervirtuosen.* Frauenfeld: Im Waldgut.

Gellrich, M. (1995). Umrisse zu einer Methode der Improvisation. *Ringgespräch über Gruppenimprovisation, 61,* 5–10.

Gellrich, M. (2001a). Psychologische Aspekte von Wahrnehmungsprozessen beim Instrumentalspiel. In M. Gellrich (Ed.), *Neue Wege in der Instrumentalpädagogik* (pp. 320–394). Regensburg: ConBrio.

Gellrich, M. (2001b). Über den Aufbau und die Koordination stabil-flexibler Spielbewegungen beim Instrumentalspiel. In M. Gellrich (Ed.), *Neue Wege in der Instrumentalpädagogik* (pp. 250–319). Regensburg: ConBrio.

Gellrich, M., & Parncutt, R. (1998). Piano technique and fingering in the eighteenth and nineteenth centuries: Bringing a forgotten method back to life. *British Journal of Music Education, 15*(1), 5–24.

Goolsby, T. (1994). Eye movement in music reading. Effects on reading ability, notational complexity, and encounters. *Music Perception, 12,* 77–96.

Heuer, H. (1996). Doppeltätigkeiten. In O. Neumann and A. F. Sanders (Eds.), *Aufmerksamkeit* (pp. 163–222). Göttingen: Hogrefe.

Johnson-Laird, P. N. (1988). Freedom and constraint in creativity. In R. J. Sternberg. (Ed.), *The nature of creativity* (pp. 202–219). Cambridge: Cambridge University Press.

Johnson-Laird, P. N. (1991). Jazz improvisation: A theory at the computational level. In P. Howell, R. West, and D. Cross (Eds.), *Representing musical structure* (pp. 291–325). New York: Academic Press.

Kenny, B. (1999). Jazz analysis as cultural imperative (and other urban myths): A critical overview of jazz analysis and its relationship to pedagogy. *Research Studies in Music Education, 13,* 56–80.

Kernfeld, B. (1981). *Adderley, Coltrane and Davis at the twilight of bebop: The search for melodic coherence, Vols. 1 and 2.* Doctoral dissertation, Cornell University.

Kratus, J. (1989). A time analysis of the compositional processes used by children aged 7–12. *Journal of Research in Music Education, 37*(1), 5–20.

Lehmann, A., & Ericsson, K. A. (1997). Research on expert performance and

deliberate practice: Implications for the education of amateur musicians and music students. *Psychomusicology, 16*(1–2), 40–58.

Martinez, I., Malbran, S., & Shifres, F. (1999). The role of repetition in aural identification of harmonic sequences. *Bulletin for the Council of Research in Music Education, 141*, 93–97.

Meyer, L. B. (1973). *Explaining music.* Berkeley: University of California Press.

Monson, I. (1996). *Saying something.* Chicago: University of Chicago Press.

Narmour, E. (2000). Music expectation by cognitive rule-mapping. *Music Perception, 17*(3), 329–398.

Nettl, B. (1974). Thoughts on improvisation: A comparative approach. *Musical Quarterly, 60*, 1–19.

Nettl, B. (1998). An art neglected in scholarship. In B. Nettl and M. Russell (Eds.), *In the course of performance: Studies in the world of musical improvisation* (pp. 1–23). Chicago: University of Chicago Press.

Nettl, B., & Russell, M. (Eds.). (1998). *In the course of performance: Studies in the world of musical improvisation.* Chicago: University of Chicago Press.

Pashler, H., & Johnston, J. C. (1998). Attentional limitations in dual-task performance. In H. Pashler (Ed.), *Attention* (pp. 155–189). East Sussex: Psychology Press.

Pressing, J. (1998). Psychological constraints on improvisational expertise and communication. In B. Nettl and M. Russell (Eds.), *In the course of performance: Studies in the world of musical improvisation* (pp. 47–67). Chicago: University of Chicago Press.

Sawyer, R. K. (Ed.). (1997). *Creativity in performance.* Greenwich, CT: Ablex.

Sawyer, R. K. (1999). Improvised conversations: Music, collaboration, and development. *Psychology of Music, 27*(2), 192–216.

Sloboda, J. A. (1985). *The musical mind: The cognitive psychology of music.* Oxford: Clarendon Press.

Smith, C. (1998). A sense of the possible: Miles Davis and the semiotics of improvised performance. In B. Nettl and M. Russell (Eds.), *In the course of performance: Studies in the world of musical improvisation* (pp. 261–289). Chicago: University of Chicago Press.

Sternberg, R. J. (2000). In search of the zipperump-a-zoo. *Psychologist, 13*(5), 250–255.

Sudnow, D. (1978). *Ways of the hand: The organisation of improvised conduct.* Cambridge, MA: Harvard University Press.

Weisberg, R. (1999). Creativity and knowledge: A challenge to theories. In R. Sternberg (Ed.), *The handbook of creativity* (pp. 226–250). Cambridge: Cambridge University Press.

Welford, A. T. (1952). The "psychological refractory period" and the timing of high speed performance: A review and a theory. *British Journal of Psychology, 43*(1), 2–19.

Westendorf, L. (1994). *Analysing free jazz.* Doctoral dissertation, University of Washington.

# 9

# Sight-Reading

ANDREAS C. LEHMANN & VICTORIA McARTHUR

From a psychological viewpoint, sight-reading involves perception (decoding note patterns), kinesthetics (executing motor programs), memory (recognizing patterns), and problem-solving skills (improvising and guessing). Sight-reading skills seem to be highly trainable and differences in sight-reading ability can be explained through differences in the amount of relevant experience and the size of the knowledge base (e.g., repertoire). The ability to perform with little or no rehearsal may be regarded as a reconstructive activity that involves higher-level mental processes. These are primarily initiated by visual input but also by conceptual knowledge and specific expectations. Common problems in sight-reading of pitches, rhythm, articulation, and expression are enumerated along with suggestions for their remediation through the use of technical equipment, practice of isolated parameters, and strategic preparations for playing.

When musicians speak of sight-reading, not all of them have the same activity in mind. Some might consider only the very first time one reads or plays through an unfamiliar piece to be true sight-reading, while others would allow the definition of sight-reading to encompass play-throughs after more extensive preparation. A conductor might even consider sight-reading to be the activity of silently reading through a score while imagining or performing appropriate conducting movements.

Rather than attempting a definitive circumscription of sight-reading, it is more helpful to picture a continuum of rehearsal, where playing a piece of music without any preparation would be located on one end of the scale while performing it after rehearsal to the point of overlearning would be located at the other end. One could limit the description of sight-reading somewhat by requiring that the music be physically played (gestured, softly voiced, or otherwise sounded) at an acceptable tempo and with appropriate expression, thereby excluding the mere deciphering of the notation, especially the tediously slow grop-

ing for notes with nothing more in mind than to internalize the piece. One could restrict sight-reading even further to those activities where the performer intends to approximate the final product as closely as possible.

The preceding description of sight-reading situates the skill of sight-reading within its naturally occurring music-making context. For example, during entrance auditions at music schools professional accompanists are confronted with scores that they may have already encountered on previous occasions. Also, professional musicians in studio orchestras, where music is being produced for soundtracks or other recordings, frequently play unrehearsed scores. At both professional and amateur levels, sight-reading commonly occurs when a choir or ensemble embarks on a new piece and the conductor wants to gain a rough impression of how it will sound before engaging in detailed rehearsal. Sometimes musicians might select new repertoire by sight-reading several works to ascertain which are the most suitable for their purposes.

While most musicians in the Western tradition sight-read to some extent, they often forget that it is the gap between each person's ordinary level of rehearsed performance and the same person's ability to perform at first sight that is the problem. The smaller this gap is, the better the sight reader.

### *History*

Sight-reading is not exclusively a phenomenon encountered in traditional Western art music but has been part of any culture that possesses music notation and literate musicians. Only exclusively oral cultures have no need for sight-reading, and some musical traditions that are otherwise improvisatory, such as those of India and Indonesia, use notation only for didactic purposes.

It is important to remember that systems of musical notation are highly culture-specific and what they capture and which instructions they contain vary considerably. Since these systems only partially indicate what needs to be played, musicians must use their extensive knowledge and skills to accurately interpret the symbols. The reasons for the emergence of notations and their historical developments are well documented (see any music dictionary or encyclopedia).

In the nineteenth century, composer-performers such a Felix Mendelssohn Bartholdy and especially Clara Schumann began the tradition of publicly performing what we now call repertoire, namely, well-rehearsed (often to the point of memory) existing pieces by other composers. Prior to this, solo and ensemble performers were accustomed to playing new scores at first sight (*prima vista*) and extensive rehearsals like those of today were uncommon. This was due to several reasons: the musical idioms were familiar, most music was not performed more than a few times, composers were afraid of plagiarism by orchestral musicians, and so on. As a result, works written by someone other than the performer were generally sight-read.

Not only was sight-reading a prerequisite skill then, but outstanding sight-reading abilities, especially among child performers, have always had the aura of unexplainable feats. Extreme proficiency was considered a characteristic of prodigies and a marker of high levels of musical aptitude. For example, Lord

Barrington described W. A. Mozart's exceptional sight-reading abilities in a scientific report. Unusually good sight-reading skills are also attributed to Felix Mendelssohn-Bartholdy, Carl Czerny, Franz Liszt, and many others. For various reasons sight-reading has lost its place in public recitals, but it is to this day an indispensable part of any serious audition and subsequent training for musicians; piano accompanists and studio orchestra musicians are proficient at it, and it is even the focus of a few lesser known competitions.

### *Research Issues*

The most interesting and simple question for the individual musician is: why are some people good sight readers and others not? Furthermore: how is sight-reading ability acquired and how can we improve sight-reading performance based on what we know about its development? Yet before we can answer these questions, we have to study the exact nature of sight-reading. Is it a simple matching of notes to fingers, or is the process more complex? Also, we need to identify what distinguishes good from less good readers.

Research on sight-reading has been a recurring topic since empirical research in music emerged, and it is also a common topic in publications aimed at music teachers. Some authors align themselves with research on reading, some with other psychological and educational aspects. As of late, the *development* of sight-reading skills has also attracted some attention (for reviews see Sloboda, 1984; Lehmann & Ericsson, 1996; Waters, Underwood, & Findlay, 1997; see also numerous music education dissertations). Our review of the literature will first cover the mental and physical task demands and the process of sight-reading and then outline the acquisition and development of sight-reading abilities.

## Basics on Looking and Perceiving

We know from vision and reading research that people do not perceive the world around them the way a camera would, namely, as a coherent sharp picture of the whole field of vision (see Smith, 1994, for a good introduction). Instead, our field of vision has a small area (*fovea*) where the objects are in focus and then a blurrier circle of peripheral vision (*parafovea*), which by no means covers the entire field of vision. Research indicates that some information in the parafoveal area can be processed in addition to that of the focal area. Depending on the viewing distance from eye to object (e.g., the printed page), the area of focus varies in size but covers less than two degrees of the field of vision. In order to view an entire image, our eye performs larger and smaller discrete movements—called *ocular saccades*—at a rate of about four to six per second. With each saccade the focus moves from one point of *fixation* to the next; picture a flashlight being pointed to different locations in a dark room. Our brain then assembles the picture and creates the perfect illusion of a coherent environment. It creates meaningful connections between isolated elements by so-called *Gestalt principles* that apply equally well to vision and to sound (Bregman, 1990; or see any introductory psychology text).

Our perceptual systems (auditory and visual) develop early in life and operate on the basis of two kinds of processes. One is a *data-driven* or *bottom-up* process, which enables us to perceive the physical properties of an object such as shapes, sizes, and pitches, and one is a *conceptually driven* or *top-down* process, which interfaces with things we have learned and stored in long-term memory. In music sight-reading, the data-driven process would automatically take care of the decoding of pitch and duration information, with unusual complex stimuli requiring (and attracting!) more attention than familiar and simple ones, while the conceptually driven processes would allow the musician to develop hypotheses about the structure of the stimulus and anticipate continuations. Thus where we look (or don't look) is determined partly by the physical properties of the stimulus itself, which is what *bottom-up* means, and partly by what we expect or need to see in order to process the score, which is what *top-down* refers to.

Given the speed demands in sight-reading, it is doubtful that these two kinds of note-reading processes can be consciously controlled. In fact, even top–down processes such as expectations are likely to become automated in skilled performance. The two processes and their interaction are likely to develop with training. As a parallel, one learns that the tonic generally follows the dominant seventh chord in common practice, and one's perceptual system learns to look for indicators of such progressions. In this way, dominant seventh chords start to attract attention to themselves and facilitate subsequent note recognition and processing.

## Subskills of Sight-Reading

Similar to other complex skills such as playing tennis, sight-reading is not a unitary skill but should rather be regarded as a collection of subskills that can be discussed and studied separately (e.g., Lehmann & Ericsson, 1993, 1996; Sergent et al., 1992; Waters, Townsend, & Underwood, 1998). Let us assume that sight-reading comprises certain perceptual, kinesthetic, memory, and problem-solving skills.

### *Perceptual Skills*

In visual tasks, eye-movement recordings have been used to uncover individual differences in attentional processes that in turn can account for differences in level of performance (eye movements are recorded by means of special infrared cameras). In general, better readers will scan (or parse) the page more efficiently than less skilled readers. Although it is difficult to compare the various studies on eye movements during reading music because they have employed markedly different experimental methods and tasks, we will attempt a cautious summary of results.

Better sight readers require shorter and fewer fixations to compare or encode material for execution because they are able to grasp more information in one

fixation (e.g., Waters et al., 1997). As a result, readers do not fixate on all notes but also fixate on blank areas between two notes. One study demonstrates that better sight readers' fixations were also directed across line and phrase boundaries, whereas less proficient readers tend to focus on individual notes (Goolsby, 1994). Also, better readers' eye movements go further ahead in the score, but also check on themselves by returning to the current point of performance. Less proficient readers search around for information and try to make sense out of what they see, while efficient readers seem to know what to look for (Goolsby, 1994).

Earlier studies demonstrated that readers of contrapuntal music tend to follow individual melodic lines more horizontally, while in homophonic music chords are scanned vertically, chord by chord (Weaver, 1943). Of course, the exact visual patterns while reading vary somewhat from one person to another. Other researchers have found that staves in piano music are read separately (Furneaux & Land, 1997). Surely the eye does not know the difference between the two types of music (contrapuntal and homophonic), but the reader has to invoke different strategies to read the music. Eye-movement data also show that, with increasing familiarity with the music fixations become shorter and eye movements longer, presumably because part of the information is already known.

These eye-movement studies can help us to better understand data from experiments about looking ahead (*perceptual span*). One of the first things investigated was how long subjects can continue playing after removal of the printed page, a measure that has been termed *eye-hand span*. This perceptual span for single-line melodies encompasses about six or seven notes for good readers but only three or four notes for less-skilled readers (Sloboda, 1984; Goolsby, 1994; for reviews), corresponding to roughly one second of music (Furneaux & Land, 1997). Once the notes are deciphered by the perceptual system, the associated storage devices in memory (*buffers*) organize their content in musically meaningful units (*chunks*). Note that those chunks are not necessarily encoded with one single eye fixation! Larger perceptual spans point to larger chunks in memory, which means that the players can devote more attention to the solution of other performance problems, for example, reassuring themselves of unclear passages or planning fingerings and jumps.

Sloboda (1984) found that the eye-hand span is not a fixed measure but grows and shrinks according to the musical structure. If a musical phrase exceeds the average span by a few notes, the span will grow to encompass the end of the phrase; if the end of a musical phrase lies closer than the average span, it will shrink. Thus the eye-hand span tends to coincide with phrase boundaries, which shows how higher level expectations ("end of phrase coming up") influence the way notes are grouped together when sight-reading ("get notes to end of phrase").

A recurring finding, not only with regard to perceptual performance in sight-reading, is that more expert subjects are more accurate and rapid than novices in many tasks related to the perception, the comparing of visual information, and so on—and that this difference is greater when the structure of the stimulus material is meaningful or familiar to the experts. "Meaningful" here means that the person can store a note sequence (e.g., C–E–G–E–C) by relating elements within this se-

quence in a sensible way (e.g., triad on C, up and down) instead of having to remember each note separately. This phenomenon is strongly related to the chunking process referred to earlier. When the information is scrambled or random and no meaning is apparent, even experts become slower and less accurate. Novices usually do not show this differentiated response to the stimulus properties, because they cannot extract the meaning. This means that prior knowledge about triads facilitates their recognition (visually, kinesthetically, and aurally).

A rarely investigated aspect of sight-reading is the influence of auditory perception (feedback) on performance. Theoretically, auditory feedback should not play a role in sight-reading, since no possibility for error correction exists. That would appear to explain why the removal of auditory feedback does not seem to cause sight-reading performance to deteriorate significantly (Banton, 1995). In an experiment, pianists sight-read music and at the same time repeated information that was presented over headphones. If sight-reading were reliant on creating an internal auditory representation of the music, this task should have caused interference—but it did not. Thus, in a self-paced performance, sight readers relied on their visual senses (i.e., the printed score) and did not need the auditory ones (i.e., the output they are producing). However, it has already been mentioned that the perceptual performance is only in part dependent on the physical properties of the printed stimulus but instead is heavily driven by expectations and prior knowledge of musical form and of the purely musical properties of the (meaningful) stimulus. The truth probably lies in the middle: playing without feedback is essentially possible when the music is not very complicated, but feedback probably serves as a cue for creating expectations about the next notes, leading to a more musical performance.

### *Kinesthetic Skills*

An often encountered belief about good sight-reading is that better sight readers have specific tactile or kinesthetic skills that allow them, at least on the piano, to orient themselves on the instrument without the help of visual monitoring. (Other instrumentalists cannot easily look at their instruments anyway, even when playing rehearsed music.)

Some authors have addressed eye and head movements made by poor sight readers, because these may interrupt visual contact with the score. What happens if subjects do not look at their hands? Researchers in typing have shown that even skilled typists' performance deteriorates when visual feedback (opportunity to monitor hand movements) is removed. It is therefore not surprising to find similar results in piano music sight-reading, namely, that the number of errors increases significantly when visual feedback is inhibited (Banton, 1995; Lehmann & Ericsson, 1996). However, pianists who are good at accurately performing large jumps on the keyboard tend to do so regardless of whether visual feedback is present or not (Lehmann & Ericsson, 1996). This suggests that kinesthetic ability is acquired in the course of piano training and is not a prerequisite for sight-reading and that avoiding unnecessary glances at the keyboard will improve performance.

### *Recall and Memory*

It is almost paradoxical to talk about memory in sight-reading, since the very nature of sight-reading would seem to contradict the importance of recall. Yet when a piece is sight-read several times, memory becomes important. Remember that sight-reading is performance with little or no rehearsal; thus the repeated playing of a piece by an accompanist would *not* generally violate this basic assumption about the nature of sight-reading. In this case, good recall would facilitate improvement on subsequent play-throughs, and even for repeating or highly similar sections within a piece.

In fact, over repeated play-throughs with expert pianists performances on the first and subsequent play-throughs indicated that better sight readers learned the piece more quickly than less-skilled readers (Lehmann & Ericsson, 1996). This faster learning can be attributed to their superior ability to grasp the structure. Also, when subjects performed an unexpected fill-in-the-blanks task that required them to remember what they had just sight-read, recall scores were consistent with scores on the sight-reading task. The performance on the fill-in-the-blanks task was also correlated with results from a subsequent memory experiment with the same subjects. These findings agree with those from other studies that showed that experts display superior recall on domain-specific material, even when they are not explicitly asked to memorize (cf. Waters et al., 1998, for another sight-reading study).

### *Problem Solving*

Musical problem solving is especially necessary in sight-reading, where not always all the notes can be clearly identfied and where external demands on the musician (for example, when playing with an accompanist or soloist) divert some attention away from the reading task. In such instances, the reader has to guess, simplify, or improvise. These processes have in common that the reader is actively reconstructing the musical material rather than simply taking it off the page and duplicating it on the instrument. Reconstruction of incomplete patterns is a fundamental perceptual-psychological, or Gestalt-psychological, process. For example, you can easily recognize an object when it is seen through a picket fence provided you are familiar with the everyday appearance of the hidden object. Your perception seems to put back the missing parts of the pattern that are obscured by the pickets.

One striking early finding in this respect is *proofreader's error*, a phenomenon we know from language reading that Sloboda (1976) investigated in music. A proofreader's error is the oversight of a mistake in a (highly) familiar word. By the way, did you notice the misspelling in the word *identified* in the previous paragraph? Sloboda gave subjects a simple tonal piano part to sight-read in which several notes had been altered by a step; if played as written it would have sounded rather awkward. Yet subjects who performed the music at sight unintentionally corrected the altered notes to match their expectations, which meant back to the original pitches. Moreover, the number of falsely corrected notes

increased with additional trials, although one would have hypothesized the opposite.

Another piece of evidence comes from an experiment where pianists were asked to sight-read a piece in which several notes had been erased (Lehmann & Ericsson, 1996). Again, similar to the fill-in-the-blanks task described earlier, the performances were scored for correctly inferred (improvised) notes. Not only did subjects who scored highly on the inference task also tend to be better sight readers, but also the number of correctly inferred notes increased statistically on a second attempt of filling in the blank. Despite their complete lack of information as to the required notes, better sight readers somehow assimilated the structure of a piece with its redundancies and were able to create appropriate passages on the fly.

Unlike computer keyboardists, whose fingers are consistently mapped to the keys, music keyboardists have to cope with the problem of the inconsistent mapping of fingers to keys. It appears that better sight readers use more appropriate fingerings at first sight and that on consecutive play-throughs the fingerings are optimized with regard to the upcoming musical context, which is (naturally) unknown at the first encounter (Lehmann & Ericsson, 1995; Sloboda et al., 1998).

## Individual Differences in Musicians' Ability to Sight-Read

From anecdotal accounts in biographies and from casual observations, we know that there have existed and still exist large individual differences in sight-reading abilities. Some authors even talk of virtuoso pianists who cannot read music fluently, while outstanding sight readers are found even among mediocre pianists (e.g., Wolf, 1976). This observation probably applies to most instruments.

### *Relation to Other Musical Skills*

Contrary to the aforementioned anecdotal evidence, the research literature suggests that better sight readers tend to be better performers (see Lehmann & Ericsson, 1993, for a discussion). In fact, moderate to high associations (correlations) are found among sight-reading performance, playing from memory, and performing rehearsed music in students of clarinet and trumpet (McPherson, 1995), sight-reading and rehearsed performance in clarinet students (Watkins, 1942), and sight-reading and amount of music committed to memory after one hour of deliberate memorization (Nuki, 1984). However, other authors failed to find such relations (Lehmann & Ericsson, 1993, 1996).

We could speculate that it is possible to find associations among specific subskills of sight-reading at one expertise level but not at another. Since the task demands of sight-reading differ from those of other musical activities, strong relationships among different activities are unlikely. However, especially among beginners there may be limiting factors in basic musical skills that constrain performance of all activities. For example, for a trumpet student who does not match notes and valve combinations fluently, even proficient reading or great

improvisatory powers will not be beneficial—he just can't get it out! Yet at higher skill levels, where basic skills are no longer an issue, divergent musical backgrounds and long-term training differences (see later) may lead to an idiosyncratic skill profile, where sight-reading may be well developed in some people but not in others. For nonkeyboard instrumentalists, sight-reading opportunities may not be as plentiful as for pianists, leading to less of a gap between people who can and people who cannot.

### *Acquisition of Sight-Reading Skills*

Educators and musicians in general are interested in whether or not sight-reading abilities can be acquired and, if so, how we can optimize this learning process. The first issue is an empirical one, and we will present some evidence that suggests that training and experience are important predictors for sight-reading achievement. The second question will be addressed in more detail in the last part of this chapter.

Banton (1995) found that frequency of self-reported sight-reading practice coincided with better sight-reading results in the expected direction, namely, that more is better. Another author (Kornicke, 1992) found that self-reported sight-reading experience was moderately associated with the results of a sight-reading test with college pianists. Unfortunately, neither of the authors took into account the level of general music experience or assessed the type of practice activity in which subjects might have engaged. Still, these findings support music teachers' advice to improve sight-reading by doing it frequently (see Lehmann & Ericsson, 1996, for references).

A more recent study was done in which college-level pianists sight-read an accompaniment to a prerecorded solo part and were interviewed about their musical backgrounds (Lehmann & Ericsson, 1996). When trying to explain individual differences in sight-reading performance, the authors found that sight-reading proficiency was primarily associated with the number of hours engaged in accompanying-related activities up to the point of the interview and with the current size of the pianists' accompanying repertoire. Conversely, age, accumulated piano practice (a proven indicator of level of performance among experts), and performance on isolated subskills (e.g., improvisation, recall, and kinesthetic ability) did not predict performance. This suggests that even performance on the named subskills can be viewed as a result of training and that sight-reading is not enabled by a fortuitous predisposition (talent) (chapter 1).

Thus experience is important, but other evidence suggests that this is not the whole story. We can observe that when someone sight-reads extensively but only at a particular, unchanging level of difficulty, then this person will not improve despite the accumulation of experience. This is the case among amateurs, be it in music or in sports. However, when this experience is combined with deliberate training efforts, be it through the choice of increasingly difficult reading material or through the strategic building up of a suitable accompanying repertoire, then the correlation between experience and level of sight-reading can be expected to increase (see Lehmann & Ericsson, 1996, for details).

## Sight-Reading as a Reconstructive Process

The various research strands enable us to develop a cognitive model of sight-reading that is both consistent with research findings and of practical interest to music practitioners. The research suggests that we should view sight-reading as a twofold process. First, the performer encodes the printed music in terms of its structural and technical properties in the course of a complex reconstructive process, and second, he or she produces appropriate physical actions. The ability to execute these two processes does not emerge suddenly but rather as the result of extended training.

For the initial stages of the reconstructive process, the performer visually scans the music, relying partly on highly automated mental processes that allow the almost effortless encoding of typical features of the music. During this process musical features are recognized as patterns and matched to pertinent information already stored in long-term memory. One example would be the left-hand patterns in classical piano music called Alberti basses: here the pianist simply has to identify the notes involved and for how long the pattern is sustained. Gestalt principles such as continuation and proximity are likely to guide and facilitate this recognition process. The motor (physical) component necessary to perform the recognized pattern is well entrenched, and the pianist can focus on expression, accuracy, and the synchronization with other players.

Further processing only becomes necessary when the musician deliberately attempts to read information that is not easily matched with preexisting patterns or when visual complexity impairs the automatic deciphering of the music. In this case, the sight reader has to actively reconstruct the music based on partially extracted information and prior knowledge. The importance of prior knowledge and expectancies can be judged from proofreader's errors and the ability to guess correct continuations.

It may be that the musicians' main problem in sight-reading is to supply enough patterns and rules from memory for the described semiautomatic deciphering and pattern-matching process, so that most of the music is executed effortlessly. The remaining attentive and cognitive resources can then be deployed in strategic ways to cope with those notated events that remained unprocessed or only partially processed. Experts are better at this because they have optimized this intricate process and they have the necessary knowledge base of rules and patterns.

Of course, some cognitive resources must also be allocated to the coordination of timing and expression in much the same way that they would be in rehearsed music performance. To coordinate their timing, musicians have to integrate movement and note processing, and to create musical expression they rely on a set of standard devices that they intuitively apply (such as slowing at the end of phrases: chapter 13). Any remaining attentive capacity can then be used to improve other aspects of the performance, such as the monitoring of internal states (anxiety, expressive ideas) or the conscious listening to oneself and other musicians (building expectations, noting mistakes for future correction).

## Pedagogical Applications

There are roughly a hundred sight-reading studies in the music education literature. Mapping these to psychological research is difficult, because educators have often implemented and validated their intuitions about teaching methods without detailing why their methods work (or fail to work). In the following suggestions, which are informed by the large amount of available literature, we draw most heavily on the idea that sight-reading has its unique skill demands and that learning to sight-read should emphasize those demands.

### *Practicing for Performance Versus Sight-reading*

In a previous paragraph we addressed the relation between sight-reading skills and other musical activities. Wolf (1976) found only a weak association between sight-reading achievement and general performance achievement. In light of the fact that general instrumental performance and sight-reading require different subskills, it may be practical to show less experienced students some important differences between their sight-reading practice and their practice for performance. Consequently, teaching sight-reading should encourage the simulation of representative situations in which sight-reading occurs. Table 9.1 summarizes some important differences that we have identified as well as others when practicing a rehearsed piece and when practicing sight-reading.

Some of the aspects mentioned in Table 9.1 pertain to the task demands of sight-reading (points 1, 2, and 4) and some to general musical considerations (points 3 and 5). A good bit of improving one's sight-reading is forming habits that are in line with the performance conditions. For example, not stopping and not looking (unless something bad happened) are mostly a matter of breaking old habits. In practicing for performance, we look at our hands even when the notes are correct. One should not do that in sight-reading because it interrupts the visual contact with the score. The big picture is not something we get the first time, because we have to rely on previous experiences, thus repetition, in order to know what the big picture could be like. Generating such an overview may be done by listening to the piece (or similar pieces) in order to help build correct expectations that in turn will guide our perceptual system. This is especially true for twentieth-century music.

Table 9.1 Comparison of strategies used by pianists when practicing for rehearsed performance and for sight-reading.

| | Practicing for Performance | Practicing Sight-reading |
|---|---|---|
| 1 | Correct your mistakes. | Maintain rhythm and meter. |
| 2 | Look at hands when playing. | Avoid looking at hands. |
| 3 | The details are important. | The big picture is important. |
| 4 | Correct fingering is crucial. | Get to notes however you can. |
| 5 | Avoid errors and omissions. | Errors and omissions are OK. |

### *Allocating Attention*

Figure 9.1 may exemplify the different levels of attention and processes during sight-reading that we have addressed earlier. We are focusing on technical difficulties rather than artistic (or aesthetic) difficulties that are more prominent for rehearsed performance. Panel (a) of the figure describes what may be a *low-demand* sight-reading task—that is, one in which the note patterns, as well as other features in the context such as the key and meter, frequently occur in many other pieces of music. It is likely that our brain will quasi-automatically decode and implement these sequences with low cognitive involvement. Panel (b) depicts a *moderate-demand* sight-reading task. In this example, the sight reader is constrained regarding his or her autonomous control of timing as well as choice of expressive features, given the additional task of playing in an ensemble. The performer may now be less able to aurally anticipate what is coming up, due to the need to consciously attend to the additional player's performance. Panel (c) represents a *high-demand* sight-reading task. The patterns as well as other features in the musical context are unusual for Western art music; thus they are unlikely to have been rehearsed to the point of automaticity by the performer (see Waters et al., 1998, for a discussion of the cognitive underpinnings of these

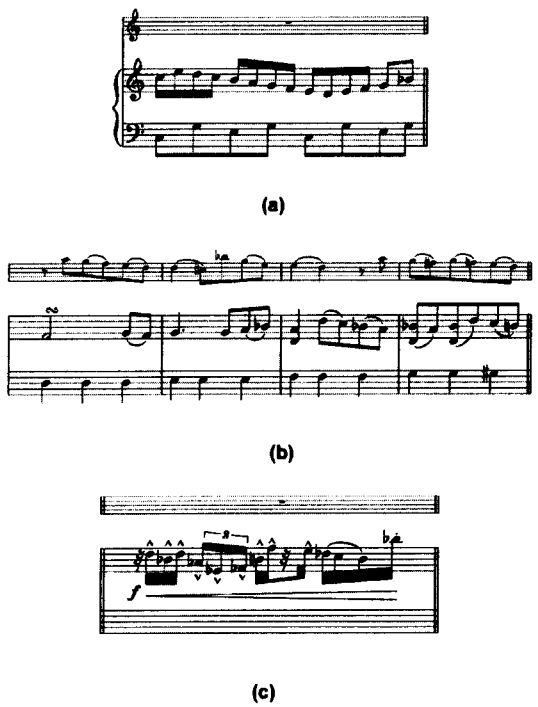



Figure 9.1. Illustrative examples of (a) low-demand, (b) moderate-demand, and (c) high-demand sight-reading for piano.

processes). It is not difficult to imagine that each task (corresponding to panels a to c) imposes different demands on perception and cognition.

Obviously, these examples are simplified and real-life music will combine aspects of all three examples. Reading music is somewhat like reading words: we have automated certain processes and actively engage in building mental models of the situation. Actively engaging in sight-reading in situations that mirror real life will ultimately create all the necessary memory traces and speed up relevant retrieval and reconstruction processes. The most important aspect, and we know this from research on expert performance, is keeping the practice tasks challenging. The following section will outline some sight-reading problems and solutions, which, applied to increasingly difficult reading material, will satisfy what we know about skill acquisition.

### *Specific Sight-Reading Problems and Solutions*

Few instructors teach sight-reading explicitly, and those who do have methods based on their intuition (and their personal reading problems). Not surprisingly, the traditional goals of most sight-reading pedagogy (Burmeister, 1991; Colwell, 1992; McArthur, 1996; Spillman, 1990) generally match the research areas described earlier. Teachers address problems that relate to visual pattern recognition (often termed music theory) and/or motor action generation (the practice of technique) with novice sight readers. In addition, teachers traditionally have urged less skilled sight readers to simply sight-read as much as possible (e.g., accompanying, participating in ensembles, etc.) in order to acquaint themselves with both the music literature as well as appropriate sight-reading strategies and techniques.

Common pattern perception problems often involve misjudging sizes of melodic and harmonic intervals. Suggested strategies for remediation involve verbalizing interval names and scales prior to playing, isolating the interval identification problem via flash-card or computer-generated drill, playing pitches only without regard to rhythm, and identifying melodic patterns prior to playing.

There appears to be a strong association between rhythmic ability and sight-reading performance (Boyle, 1970; Elliott, 1982). The inability to accurately perform rhythm (also maintain tempo and meter) may be tackled via tapping or clapping of rhythm alone (or other forms of rhythmic body movement), writing counts in the music score, drawing vertical lines that indicate note alignment, reading practice that involves the metronome or a MIDI playback device such as a sequencer, or playing in accompanying or ensemble situations with living, breathing performers.

The inability of some sight readers to perform articulation and/or dynamics can have multiple causes. Sometimes especially younger players simply choose not to attend to these seemingly minor details. When this occurs, having players actually notate in the score the expressive features exhibited in either a live or recorded performance enhances sensitivity to these features in their own performance. If the cause for unexpressive performance is attention overload, often simply slowing the tempo or only performing pitch or rhythm notation will solve the problem.

Probably the most common problem seen in sight readers is that of *stuttering*, or backing up to correct omissions or errors. The solutions offered by pedagogues for this problem all have to do with forcing the performer to continue playing: for example, performing only the notes that occur on designated beats, thus forcing the eyes to arrive on time at future beats, and reading while another person covers up the notation immediately as it is played both make it impossible for eye movement to regress. Playing with live or electronic (metronome, MIDI recordings) ensemble members is an aid here also. Another way of improving the flow of performance is to pair up with another musician and sight-read increasingly difficult material.

Actually taking one's eyes off the music momentarily in order to glance down at the keyboard is a disputable point in sight-reading lore. Some advocate training the sense of touch by consciously practicing the music with the goal of playing it perfectly without the benefit of sight; others simply advocate keeping the chin and head level when glancing down. Regardless, losing one's place through loss of visual connection with the music is the cause of many errors in sight-reading and a relatively simple problem to remedy—just don't look.

While the foregoing examples represent remedies to local problems, they will in our opinion ultimately lead the reader to develop good intuitions about the musical grammar, frequent and infrequent patterns, and how it feels to play them. This learning process will have to be supplemented with a lot of music listening that provides stylistic and idiomatic templates against which the musical score can be compared. Contrary to what teachers hope to achieve through offering look-ahead exercises to their students (e.g., Burmeister, 1991), the larger perceptual span and all other characteristics of expert sight readers are a result of repeatedly solving low-level problems with great concentration and trying to contextualize them.

## Conclusion

Our review shows that neither is the juggling of multiple tasks during sight-reading a completely unexplainable feat nor should outstanding sight-reading abilities be considered an indicator of some extraordinary musical talent. But, rather, we know that expert sight readers have had extensive experience with sight-reading tasks and have acquired a large knowledge base (repertoire of rules and patterns) to draw from. In the course of their development as sight readers, they have come to know large amounts of patterns (visual, kinesthetic, aural), have solved countless musical problems (reading, fingering, ensemble coordination), and have developed the ability to manage all situational demands of performing.

Educators should focus more attention on sight-reading pedagogy and continue to develop materials and training methods that enable students to narrow the gap between their level of rehearsed performance and ability to sight-read. Despite the fact that sight-reading is not in high demand in our public concert life and that it is only applicable to musical traditions that possess notated music,

it does offer musicians a possibility to survive in an economic situation that demands the ability to learn music quickly, if not at sight. And it also enables musicians of all levels to perform together for their own enjoyment as well as that of others.

Acknowledgment. We thank Christoph Wünsch for the musical excerpts.

### *References*

Banton, L. (1995). The role of visual and auditory feedback during the sight-reading of music. *Psychology of Music, 23*(1), 3–16.

Boyle, J. D. (1970). The effect of prescribed rhythmical movements on the ability to read music at sight. *Journal of Research in Music Education, 18*, 307–318.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Burmeister. E. (1991). *Keyboard sight reading.* Mountain View, CA: Mayfield.

Colwell, R. (Ed.). (1992). *Handbook of research on music teaching and learning.* New York: Schirmer.

Elliott, C. A. (1982). The relationship among instrumental sight reading ability and seven selected predictor variables. *Journal of Research in Music Education, 30*(1), 5–14.

Furneaux, S., & Land, M. F. (1997). The role of eye movements during music reading. In A. Gabrielsson (Ed.), *Proceedings of the 3rd Triennial ESCOM Conference* (pp. 210–214). Uppsala, Sweden: Uppsala University.

Goolsby, T. (1994). Profiles of processing: Eye movements during sightreading. *Music Perception, 12*, 97–123.

Kornicke, L. E. (1992). An exploratory study of individual difference variables in piano sight-reading achievement. (Doctoral dissertation, Indiana University). *Dissertation Abstracts International, 53*, 12A (University Microfilms No. 9301458)

Lehmann, A. C., & Ericsson, K. A. (1993). Sight-reading ability of expert pianists in the context of piano accompanying. *Psychomusicology, 12*(2), 182–195.

Lehmann, A. C., & Ericsson, K. A. (1995). Fingerings in piano performance as evidence for the mental representation of music: A preliminary study. In D. Wessel (Ed.), *Proceedings of the 1995 SMPC conference* (p. 2). Berkeley: University of California Press.

Lehmann, A. C., & Ericsson, K. A. (1996). Structure and acquisition of expert accompanying and sight-reading performance. *Psychomusicology, 15*, 1–29.

McArthur, V. H. (1996, March/April). New music for study: Piano methods and sight-reading, Part I. *Piano and Keyboard, 179*, 59–61.

McPherson, G. E. (1995). The assessment of musical performance: Development and validation of five new measures. *Psychology of Music, 23*(2), 142–161.

Nuki, M. (1984). Memorization of piano music. *Psychologia, 27*, 157–163.

Sergent, J., Zuck, E., Terriah, S., & MacDonald, B. (1992). Distributed neural network underlying musical sight-reading and keyboard performance. *Science, 257*, 106–109.

Sloboda, J. A. (1976). The effect of item position on the likelihood of identification by interference in prose reading and music reading. *Canadian Journal of Psychology, 30*, 228–236.

Sloboda, J. A. (1984). Experimental studies of music reading: A review. *Music Perception, 2*, 222–236.

Sloboda, J. A., Clarke, E. F., Parncutt, R., & Raekallio, M. (1998). Determinants of fingering choice in piano sight-reading. *Journal of Experimental Psychology: Human Perception and Performance, 24*(1), 185–203.

Smith, F. (1994). *Understanding reading: A psycholinguistic analysis of reading and learning to read* (5th ed.). Hillsdale, NJ: LEA.

Spillman, R. (1990). *Sight-reading at the keyboard.* New York: Schirmer.

Waters, A. J., Townsend, E., & Underwood, G. (1998). Expertise in musical sight-reading: A study of pianists. *British Journal of Psychology, 89*, 123–149.

Waters, A. J., Underwood, G., & Findlay, J. M. (1997). Studying expertise in music reading: Use of a pattern-matching paradigm. *Perception & Psychophysics, 59*(4), 477–488.

Watkins, J. G. (1942). *Objective measurement of instrumental performance.* New York: Teachers College Press.

Weaver, H. (1943). A study of visual processes in reading differently constructed musical selections. *Psychological Monographs, 55*, 1–30.

Wolf, T. (1976). A cognitive model of musical sight-reading. *Journal of Psycholinguistic Research, 5*, 143–171.

# 10

# Practice

NANCY H. BARRY & SUSAN HALLAM

Musicians practice to gain technical proficiency, learn new repertoire, develop musical interpretation, memorize music, and prepare for performances. Based on available empirical research, we describe appropriate practicing and learning strategies that can be incorporated into regular music teaching to encourage students to become autonomous learners. Research demonstrates that practice is more effective when musicians engage in metacognition (reflecting upon their own thought processes); employ mental practice in combination with physical practice; approach practice in an organized, goal-oriented manner; study and analyze scores; plan relatively short and regular practice sessions; are intrinsically motivated; and listen to appropriate musical examples including professional recordings and/or teacher demonstrations. Students may also benefit from understanding the relationship between time spent practicing and achievement, and the nature and the importance of motivation. The old adage *practice makes perfect* may not necessarily be true, because repetition of ineffective practice strategies can yield disappointing results.

*If I don't practice for one day, I know it; if I don't practice for two days, the critics know it; if I don't practice for three days, the audience knows it.*

Ignacy Jan Paderewski,
*An Encyclopedia of Quotations About Music*

Practice is defined as "repeated performance or systematic exercise for the purpose of learning or acquiring proficiency" (Cayne, 1990, p. 787). In many contexts, such as sports and psychology, *training* and *practice* are often used synonymously. However, since *practice* is the term traditionally used by musicians to describe systematic rehearsal, that term will be used in this chapter.

The notion that practice is required for skill development is of course very old. Scientific interest in the nature of practice and its relationship to the transition from novice to expert performance has only emerged within the past 100 years or so, with a growing body of studies across diverse disciplines (Hallam, 1997b, 1997c; Jørgensen & Lehmann, 1997). Research has shed light upon the ways that expert musicians practice, but, at present, no longitudinal data on how different styles of practice lead to better performance are available, although there is some evidence that professional musicians can attain equally high standards of performance by adopting a variety of approaches to practice (Hallam, 1995b). While practice is certainly a relevant topic for all musical genres, this chapter is limited to Western/European musical traditions, because until now most psychological research has concentrated on Western music.

## Managing Practice Time

To acquire musical skill it is essential to practice. The question that has concerned researchers is, therefore, not whether practice is necessary but to what extent it is necessary. Some researchers have argued that attainment simply increases with practice and, consequently, that accumulated practice time can directly predict achievement (Ericsson, Krampe, & Tesch-Römer, 1993; Sloboda et al., 1996).

### *Total Duration of Practice*

Achieving high levels of musical expertise requires considerable practice. Typically, 16 years of practice are required to achieve levels that will lead to international standing in playing an instrument (Sosniak, 1985). The individual usually begins to play at a very early age, with about 25 hours of practice being undertaken weekly by adolescence, increasing to as much as 50 hours. However, there is considerable individual variation (Ericsson et al., 1993; Sloboda et al., 1996). Some research (Hallam, 1998a; Williamon & Valentine, 2000) has suggested that while cumulative practice may be a good predictor of the overall level of expertise attained, it may not predict the quality of performance at any point in time. There is also evidence that the practice undertaken and required for different instruments varies (Jørgensen, 1997b).

### *Organization of Practice Time*

Psychological literature on distribution of practice time for motor skills suggests (Oxendine, 1984):

- Practice distributed over time is generally more efficient for learning and performance than massed practice (i.e., a great deal of practice in a limited time frame, such as cramming just before a lesson).
- Proficiency developed over a long period of time is retained better than proficiency developed within a short time period.

- Relatively short practice sessions are generally more effective than longer practice sessions. Of course, this varies with the age and skill level of the musician.
- Short practices are best for short, simple tasks, but longer and more complex tasks require longer practice sessions.
- A high level of motivation enables one to benefit from longer and more concentrated practice than would be possible with less motivation.
- Individuals who are more competent in a particular activity can effectively practice that activity for longer periods than individuals who are less competent. Also, older children can practice longer than younger children. However, even very advanced musicians are subject to mental and physical fatigue and are advised to distribute practice over time.
- The optimal practice duration is longer for some group activities (such as ensembles) than for individual activities because in group activities the individual may not be playing the entire time.

Some researchers (Mumford et al. 1994) have concluded that practice must be distributed across time to allow for deep, higher level thinking in which the individual actively constructs ideas about how to approach a particular task. Distributed practice may be essential when educators work with novices who lack well-developed knowledge structures and when students must begin to generate and apply new concepts. However, this study also revealed that massed practice may prove useful under two conditions: if the goal of training is a particular behavior rather than understanding and for experts who already possess the requisite knowledge structures.

## Cognitive Strategies

### *Mental Practice*

Mental practice involves cognitive rehearsal of a skill without physical activity. Oxendine (1984) described three different forms of mental practice: (1) review that immediately precedes, follows, or coincides with the performance, (2) formal or informal rehearsal between periods of physical practice, and (3) decision making relating to the strategy-making phases of the activity that occur during (or between) periods of physical practice or performance. A review of the mental practice literature (Weinberg, 1982) revealed the following principles of mental practice for motor skill acquisition. Mental practice is more effective:

- If the learner has some prior experience with the task or with a similar task
- During the early stages of learning when the student is just beginning to formulate ideas about the task and during later stages when more complex strategies have developed
- If the learner can express the task verbally
- If the learner has been taught proper techniques of mental practice such as focused concentration and visualization
- When used in combination with physical practice
- If mental practice sessions are brief

- When individuals also imagine responses in the muscles that would actually perform the movement

In music, mental practice techniques are most effective when combined with physical practice (Coffman, 1990; Ross, 1985; Rubin-Rabson, 1941c). However, more research is needed to ascertain the role of mental practice in relation to interpretation and expression.

### *Analysis*

Analytical study of the music prior to physical practice, in which aspects such as key, meter, and familiar patterns are noted, may increase performance accuracy (Barry, 1992), and reduce the number of physical trials required to achieve technical proficiency (Rubin-Rabson, 1941b). Studies of sight-reading support the notion that mental analysis and rehearsal, such as scanning the music to identify possible obstacles and making oneself consciously aware of the key and time signatures, help musicians to play more accurately (e.g., McPherson, 1994). Teachers should encourage their students to apply a variety of analysis strategies before playing, particularly in the early stages of learning a new piece.

### *Metacognition*

*Metacognition* refers to the learner's knowledge about learning itself (i.e., thinking about thinking). This is central to practice. Metacognitive skills are concerned with the planning, monitoring, and evaluation of learning, including knowledge of personal strengths and weaknesses, available strategies (task-oriented and person-oriented), and domain knowledge to assess the nature of the task and evaluate progress toward the goal (Hallam, 2001b).

The level at which a musician is able to engage in metacognition seems to be a function of musical expertise. Expert musicians have well-developed metacognitive skills that encompass technical matters, interpretation, performance, learning, concentration, planning, monitoring, and evaluation (Nielson, 1999). Novice musicians demonstrate less metacognitive awareness, with the amount and structure of their practice tending to be determined by external commitments such as examinations (Hallam, 1997a).

Novice musicians are often unaware of errors. When they can identify errors, they initially correct them by repetition of the single wrong note. Later small sections (a half-bar or a bar) are repeated. Finally, error correction changes to a focus on difficult sections that are then worked on as units. In the early stages of learning, novices have problems identifying difficult sections and tend to practice by simply playing through the music. This has been confirmed in a range of settings (Gruson, 1988; Hallam, 1997a; Renwick & McPherson, 2000).

Beginners demonstrate little or no self-regulation in their practice (McPherson & Renwick, 2000). Initially, they tend to focus on playing notes that are at the correct pitch, but as their expertise develops attention is then directed to rhythm, other technical aspects of playing, and later dynamics, interpretation, and the expressive elements of playing (Hallam, 2001a).

Students become aware of specific strategies for practice before adopting those strategies spontaneously. This is known as a *production deficiency* (Flavell, Beach, & Chinsky, 1966). Changes in strategy use seem to be more closely linked to developing expertise than age (Hallam, 1997a, 1997c), in particular the development of appropriate aural schemata. Prior to this, learning about the nature of particular strategies and how to adopt them is not productive because students have no conception of when they are making errors. Once a certain level of expertise has been attained, however, students develop approaches to practice that mirror those of professionals. There is a greater concern with interpretation, and individual differences emerge in approaches and orientations to practice (Hallam, 1997a).

### *Individual Differences*

Several researchers have studied differences in learning styles and approaches to practice. Hallam (1997a) found differences in orientation to practice, approach to detailed practice, and interpretation in professional musicians, all of which were equally effective. McLaughlin (1985) concluded that there was no single method that professional musicians adapted when transferring embouchure control, air support, tonguing, and vibrato production from one woodwind instrument to another, and Cantwell and Millard (1994) identified deep and surface approaches to learning in 14-year-old musicians. A later study (Sullivan & Cantwell, 1999) analyzed the relationships among high levels of planning, strategy use, and deep and surface approaches to learning. In two tasks, one with traditional notation, the other with a nontraditional graphic score, a deep approach was consistently linked with high-level cognitive strategies and a high level of planning. The deep approach was consistently more effective.

Different types of structure or practice organization may be best for optimal skill development in particular students (Kane, 1984). A study of practice strategies, individual differences in cognitive style, and gender (Barry, 1992) found that students who followed a detailed practice procedure (structured practice) made more improvement in the accuracy and musicality of their musical performances than did those students who practiced without benefit of a specific structure. This was particularly evident in field-dependent males (students having difficulty identifying an image embedded within a complex visual field).

## Practice Activities

### *What Is the Purpose of Practice?*

In music, practice is necessary to enable musicians to acquire, develop, and maintain aspects of technique, learn new music, memorize music for performance, develop interpretation, and prepare for performance. A key purpose of practice is to enable complex physical, cognitive, and musical skills to be performed fluently with relatively little conscious control, freeing cognitive processing capacity for higher order processing (e.g., communicating interpretation).

Three stages are generally recognized in the development of a motor skill (Fitts & Posner, 1967). In the cognitive-verbal-motor stage, learning is largely under cognitive, conscious control, requires effort, is deliberate, and may require verbal mediation. During the associate stage, the learner begins to put together a sequence of responses to produce a desired outcome. This becomes more fluent over time. In the autonomous stage, the skill becomes automated and appears to be carried out without conscious effort.

Recently the term *deliberate practice* has been used to specify the type of practice associated with the development of expert skills in a variety of areas such as computers, sports, and music (Ericsson, 1997; Ericsson et al., 1993; Pranger, 1999). According to this conceptualization, deliberate practice requires the identification of a specific goal at an appropriate difficulty level for the individual, meaningful feedback, and opportunities for repetition and correction of errors (Ericsson, 1997). Deliberate practice is highly focused and requires great effort and concentration.

### *Learning a New Piece of Music*

Another purpose of practicing for all musicians is to learn new repertoire. Several studies have addressed the issue of how musicians learn new music. These studies (Chaffin & Imreh, 1997; Hallam 1995a, 1995b, 1997a; Miklaszewski, 1989, 1995; Wicinski, 1950) taken together suggest:

- Most musicians tend to acquire an overview of the music they are to learn in the early stages of practice of a new work, in a way that depends on their ability to develop an internal aural representation of the music from examination of the score alone.
- The structure of the music determines how it is divided into sections for practice.
- The more complex the music, the smaller the sections.
- As practice progresses, the units become larger.
- A hierarchical structure appears to develop, in which the performer's notions about the ideal performance are gradually integrated into a coherent whole. This plan is guided by musical rather than technical considerations.
- There is considerable individual diversity in the ways that musicians practice.
- The way that practice progresses also depends on the nature of the task. Musicians approach the task of learning contemporary music differently from that of learning older music. They generally find it more difficult and place greater emphasis upon cognitive strategies (specific strategies about how to approach the task) (Miklaszewski, 1995; Hallam, 1995b).

### *Developing Interpretation*

Such research as there is suggests that musicians adopt two main approaches to developing interpretation: intuitive and analytic. If an intuitive approach is adopted, interpretation evolves during the course of learning to play the piece

and is based on intuition. When an analytic approach is adopted interpretation is based on extensive listening to music, comparison of alternative interpretations, and analysis of the structure of the music. Here interpretation can be developed with little actual physical practice taking place. Some musicians adopt both approaches to developing interpretation, although they tend to exhibit a preference for one (Hallam, 1995a).

### *Memorizing Music*

A further purpose of practice can be to memorize music for performance. The evidence suggests that while much memorization takes place as practice proceeds, musicians may have to adopt additional strategies toward the end of the learning process to consolidate their learning, particularly if the music is long and complex (Chaffin & Imreh, 1997; Miklaszewski, 1995).

The strategies adopted depend on the nature of the task and the context within which the music is to be performed (Hallam, 1997b). Musicians report being more likely to rely on automated aural memory without resort to cognitive analysis if the piece is short, simple, and to be performed in a relatively unthreatening environment, such as an informal concert. The evidence also suggests that changes in memorization practice strategies occur as expertise develops (chapter 11 in this volume; Hallam, 1997b; McPherson, 1996).

### *Preparing for Performance*

While most practice is in preparation for performance, some musicians adopt specific practice strategies to prepare, such as playing through as if in performance conditions, recording or videotaping themselves. We know relatively little about the effectiveness of such strategies, although there is considerable evidence that being well prepared is important (Bartel & Thompson, 1994).

## Teaching Students to Practice Effectively

### *The Teacher*

Music teachers acknowledge the importance of practice and report that they discuss specific practice techniques with their students. A survey of 94 applied music teachers (Barry & McArthur, 1994) found that most teachers encouraged students to use different approaches to practice, to begin a piece slowly and gradually increase the tempo, and to analyze a new piece before playing it. Teachers also tended to encourage students to mark their music, set specific practice goals for each practice session, have two or more short daily practice sessions instead of one longer session, and practice with the metronome. Responses that regarded other aspects of practice were varied, suggesting the lack of a shared systematic pedagogy.

A study (Barry, 2000) of the relationship between student–teacher interactions in the college music lesson and subsequent individual student practice

behaviors revealed that student practice habits were influenced to a limited degree by their teacher's advice, but the most powerful influence was the teacher's instruction style. What the teachers actually did during the lessons (e.g., the teacher demonstrating a particular technique or having the student try a particular approach) had a more profound influence upon their students' practice than what they said.

### *Providing Examples*

Hearing high-quality examples may prove beneficial in guiding practice for student musicians who have not internalized an appropriate ideal for musical performance. Music educators have stressed the importance of teacher demonstration to provide students with aural examples. This can be very useful for beginning students, but at a later stage the provision of a single example may force students into stereotypical performances and deter them from developing their own interpretations (Hallam, 1998a).

Teacher demonstration can certainly provide a useful resource for students, but many music teachers may lack the skills necessary to demonstrate on every instrument that they are called upon to teach (Kohut, 1973). Teachers may not have sufficient time to demonstrate every instrument during every class and cannot be available to demonstrate during students' home practice sessions (Linklater, 1997). There is considerable interest, therefore, in the effects of providing students with taped performance models (examples) to use in conjunction with individual practice. One experiment (Zurcher, 1975) provided beginning instrumentalists with cassette tapes that contained instructions, reminders, and model play-along performances of the music. Model-supported practice was significantly more effective than traditional practice on gross pitch discrimination, pitch matching, rhythmic discrimination, and time spent in practice. Additional studies (Dickey, 1992; Kendall, 1990) also support the use of aural models as an aid to musical development. In contrast, Anderson's (1981) study of tape-recorded aural examples for home practice failed to reveal any significant differences between sight-reading and performance skills of young clarinet students who used taped examples and those with no taped examples.

A more recent study (Linklater, 1997) investigated the effects of home practice that used three different types of examples (model videotape, model audiotape, audiotape with nonmodeling/accompaniment only) on the performance achievement of beginning clarinet students. The students in the videotape group scored significantly higher on visual and physical performance criteria immediately after practicing with examples and also scored higher on performance criteria based on tone quality and intonation in subsequent delayed longitudinal assessments than did students in the nonmodeling group.

### *Motivating Students to Practice*

There are complex relationships among motivation, achievement and practice (Hallam, 1997c; 1998b). McPherson, Bailey, and Sinclair (1997) found that stu-

dents who reported higher levels of daily practice were more likely to improvise, play music by ear, compose, and elect to take music classes at school. Likewise, O'Neill (1997) observed that higher achieving beginning instrumental music students practiced more during a 2-week diary period than lower achieving students. She also found a relationship between children's motivational processes and effective music practice. Achievement is related not only to the length of time spent practicing but also to the quality of that practice.

Given that increasing the amount of time engaged in practice can contribute to an increase in musical achievement, musicians and researchers have investigated a number of different motivational approaches. On the one hand, a study (Rubin-Rabson, 1941a) that compared the effects of verbal encouragement or cash payment with a control group (with no special incentive to practice) found no significant differences among the three conditions in the amount of students' practice time. On the other hand, Wolfe (1987) found that behavioral contracts in which students enter into a formal agreement with parents and the teacher could be an effective way to motivate students to practice.

Student motivation to practice may be increased by allowing students to make certain choices. One project (Brandstrom, 1995) allowed piano students to set their own goals regarding improvement, repertoire, and time. These university students also engaged in self-evaluation of the same three elements through written comments and reflections. Project evaluation indicated that most of the students held favorable attitudes about the project, believing that they had become better pianists and had developed greater independence and planning ability. Greco (1997) found that using student-selected repertoire arranged in appropriate keys and tessitura resulted in an increase in practice time and a decrease in attrition rate for second- and third-year instrumental music students in Grades 5 and 6 (see also chapter 3 in this volume).

### *Supervised Practice*

There is evidence to suggest that for the young or novice musician supervised practice can be more effective than practice carried out without supervision. Research indicates a strong relationship between youngsters' musical development and the direct involvement of parents and nurturing, supportive teachers (O'Neill, 1997; Sosniak, 1985). Studies (Brokaw, 1983; Davidson, Howe, Moore, & Sloboda, 1996; chapter 2 in this volume) have also revealed a significant relationship between the achievement of young instrumentalists and the amount of parental supervision of home practice. Supervision by other adults may also be helpful. A study of adolescent instrumentalists (Barry, 1992) compared the musical performance improvement (melodic accuracy, rhythmic, accuracy, and musicality) of students who followed a structured practice program under adult supervision with that of students who practiced with no supervision. The supervised group achieved greater gains for melodic accuracy, rhythmic accuracy, and musicality. While supervised practice can be quite helpful for some individuals, it can be detrimental as students become more mature and seek greater independence.

### *Structured Practice*

Practice is most effective when it is governed by an appropriate framework or structure. However, student musicians may not have sufficient knowledge to determine an optimal framework for practice. One study (Barry, 1990) compared student instrumentalists' performance improvement under three different practice conditions. Students who used a structured approach to practice (teacher-designed and self-designed) were able to correct more performance errors than students who did not employ any structured practice. Students who used the teacher-designed approach had the highest scores of all.

## Conclusion

Efficient and effective practice is central to the development of musical expertise. Musicians engage in practice for a number of reasons, such as gaining technical proficiency, learning new repertoire, developing musical interpretation, memorizing music, and preparing for performance. Fluency is achieved through effective practice, with complex tasks eventually becoming automated.

Practice is most effective when it is deliberate and mindful. Acquiring metacognitive skills is central to effective and efficient practice. The literature suggests that these skills cannot be developed independently of musical skills and knowledge and that in the early stages of learning the two are irrevocably intertwined.

Teachers have a crucial role to play not only in assisting students in acquiring musical expertise but also in developing knowledge of appropriate practicing and learning strategies and supporting students in becoming autonomous independent learners. The ways that this can be achieved will depend to some extent on the age of the student (Hallam, 1998a). The younger they are, the more directed supervision they need. As their expertise increases, it is important to encourage autonomy in learning and the acquisition of a range of strategies that will assist in developing effective practice and preparation for performance. The evidence to date, however, suggests that most musicians do not feel that their training has assisted them in developing metacognitive skills (Hallam, 1992; Jørgensen, 1997a).

The following suggestions for musicians who wish to practice more effectively (and teachers who wish to guide their students toward greater achievement) are based upon evidence from a wide range of literature sources:

- Engage in metacognition—become mindful about practicing and related physical and mental processes. Be consciously aware of your own thought processes.
- Approach practice systematically. Do not go about practice haphazardly. Practice is more effective when it is structured and goal-oriented.
- Engage in mental practice (cognitive rehearsal) in combination with physical practice.

- Invest time in score study and analysis, particularly when beginning a new piece.
- Plan regular practice sessions with several relatively short sessions distributed across time.
- Acknowledge the relationship between time spent practicing and achievement and set out to invest the time necessary.
- Be aware of the importance of motivation. When teachers and parents allow students to make some choices about goals and repertoire, student motivation is likely to increase.
- Listen to high-quality models of musical performance. This is particularly important for beginning musicians. Parents and teachers should invest in a library of fine recordings and, if capable, play and/or sing often for their charges.
- Support and nurture young musicians. Parents and teachers should demonstrate keen interest and involvement in music study and practice.

Musicians may take for granted the old adage that "practice makes perfect." However, literature in both psychology and music indicates that not only the amount of time invested in practice but also the manner in which one approaches practice will have a bearing on an individual's level of musical development. While this may be stating the obvious, the important element here is that musicians stand to improve their skills by paying careful attention to increasing their own ability to learn how to learn.

### *References*

Anderson, J. N. (1981). Effects of tape recorded aural models on sight-reading and performance skills. *Journal of Research in Music Education, 29*(1), 23–30.

Barry, N. H. (1990). The effects of different practice techniques upon technical accuracy and musicality in student instrumental music performance. *Research Perspectives in Music Education, 1*, 4–8.

Barry, N. H. (1992). The effects of practice strategies, individual differences in cognitive style, and gender upon technical accuracy and musicality of student instrumental performance. *Psychology of Music, 20*(2), 112–123.

Barry, N. H. (2000, February). *Behind closed doors: What really goes on in the practice room.* Paper presented at the meeting of the Southern Chapter of the College Music Society, Lafayette, LA.

Barry, N. H., & McArthur, V. (1994). Teaching practice strategies in the music studio: A survey of applied music teachers. *Psychology of Music, 22*(1), 44–55.

Bartel, L. R., & Thompson, E. G. (1994). Coping with performance stress: A study of professional orchestral musicians in Canada. *Quarterly Journal of Music Teaching and Learning, 5*(4), 70–78.

Brandstrom, S. (1995). Self-formulated goals and self-evaluation in music education. *Bulletin of the Council for Research in Music Education, 127*, 16–21.

Brokaw, J. P. (1983). The extent to which parental supervision and other selected factors are related to achievement of musical and technical-physical characteristics by beginning instrumental music students. (Doctoral dissertation, University of Michigan). *Dissertation Abstracts International*, 43/10-A, 3252. (University Microfilms No. 0534271)

Cantwell, R. H., & Millard, Y. (1994). The relationship between approach to learning and learning strategies in learning music. *British Journal of Educational Psychology, 64*(1), 45–63.

Cayne, B. S. (Ed.). (1990). *The new Lexicon dictionary of the English language.* New York: Lexicon.

Chaffin, R., & Imreh, G. (1997). "Pulling teeth and torture": Musical memory and problem solving. *Thinking and Reasoning, 3*(4), 315–336.

Coffman, D. D. (1990). Effects of mental practice, physical practice, and knowledge of results on piano performance. *Journal of Research in Music Education, 38*(3), 187–196.

Davidson, J. W., Howe, M. J. A., Moore, D. G., & Sloboda, J. A. (1996). The role of parental influences in the development of musical performance. *British Journal of Developmental Psychology, 14*, 399–412.

Dickey, M. R. (1992). A review of research on modeling in music teaching and learning. *Bulletin of the Council for Research in Music Education, 113*, 27–40.

Ericsson, K. A. (1997). Deliberate practice and the acquisition of expert performance: An overview. In H. Jørgensen and A. C. Lehmann (Eds.), *Does practice make perfect? Current theory and research on instrumental music practice* (pp. 9–51). Oslo: Norges Musikkhøgskole.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100(3)*, 363–406.

Fitts, P. M., & Posner, M. I. (1967). *Human performance.* Belmont, CA: Brooks/Cole.

Flavell, J. H., Beach, D. R., & Chinsky, J. M. (1966). Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development, 37*, 283–299.

Greco, V. (1997). *Investigation of the effects of student-selected repertoire on the practice habits of instrumental music students.* M. A. Action Research Project, Saint Xavier University. (ERIC Document Reproduction Service No. ED418049).

Gruson, L. M. (1988). Rehearsal skill and musical competence: Does practice make perfect? In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 91–112). Oxford: Clarendon Press.

Hallam, S. (1992). *Approaches to learning and performance of expert and novice musicians.* Unpublished doctoral dissertation, Institute of Education, University of London.

Hallam, S. (1995a). Professional musicians' approaches to the learning and interpretation of music. *Psychology of Music, 23*(2), 111–128.

Hallam, S. (1995b). Professional musicians' orientations to practice: Implications for teaching. *British Journal of Music Education, 12*(1), 3–19.

Hallam, S. (1997a). Approaches to instrumental music practice of experts and novices: Implications for education. In H. Jørgensen and A. C. Lehmann (Eds.), *Does practice make perfect? Current theory and research on instrumental music practice* (pp. 89–108). Oslo: Norges Musikkhøgskole.

Hallam, S. (1997b). The development of memorisation strategies in musicians: Implications for education. *British Journal of Music Education, 14*(1), 87–97.

Hallam, S. (1997c). What do we know about practicing? Toward a model synthesizing the research literature. In H. Jørgensen and A. C. Lehmann (Eds.),

*Does practice make perfect? Current theory and research on instrumental music practice* (pp. 179–231). Oslo: Norges Musikkhøgskole.

Hallam, S. (1998a) *Instrumental teaching: A practical guide to better teaching and learning.* Oxford: Heinemann.

Hallam, S. (1998b). The predictors of achievement and dropout in instrumental music tuition. *Psychology of Music, 26*(2), 116–132.

Hallam, S. (2001a) The development of expertise in young musicians: Strategy use, knowledge acquisition and individual diversity. *Music Education Research, 3*(1), 7–23.

Hallam, S. (2001b). The development of metacognition in musicians: Implications for education. *British Journal of Music Education, 18*(1), 27–39.

Jørgensen, H. (1997a). Teaching/learning strategies in instrumental practice: A report on research in progress. In J. A. Taylor (Ed.), *Transatlantic roads of music education: World views* (pp. 47–51). Tallahassee, FL: CMR.

Jørgensen, H. (1997b). Time for practicing? Higher level students' use of time for instrumental practicing. In H. Jørgensen and A. C. Lehmann (Eds.), *Does practice make perfect? Current theory and research on instrumental music practice* (pp. 123–140). Oslo: Norges Musikkhøgskole.

Jørgensen, H., & Lehmann, A. C. (Eds.). (1997). *Does practice make perfect? Current theory and research on instrumental music practice.* Oslo: Norges Musikkhøgskole.

Kane, M. (1984). Cognitive styles of thinking and learning: Part two. *Academic Therapy, 20*(1), 83–92.

Kendall, M. J. (1990). A review of selected research literature in elementary instrumental music education with implication for teaching. *Journal of Band Research, 25*, 64–82.

Kohut, D. L. (1973). *Instrumental music pedagogy: Teaching techniques for band and orchestra directors.* Englewood Cliffs, NJ: Prentice-Hall.

Linklater, F. (1997). Effects of audio- and videotape models on performance achievement of beginning clarinetists. *Journal of Research in Music Education, 45*(3), 402–414.

McLaughlin, D. B. (1985). An investigation of performance problems confronted by multiple woodwind specialists. (Doctoral thesis, Columbia University Teachers College, New York). *Dissertation Abstracts International*, 46/09-A, 2610. (University Microfilms No. 8525495)

McPherson, G. E. (1994). Factors and abilities influencing sightreading skill in music. *Journal of Research in Music Education, 42*(3), 217–231.

McPherson, G. E. (1996). Five aspects of musical performance and their correlates. *Bulletin of the Council for Research in Music Education, 127*, 115–121.

McPherson, G. E., Bailey, M., & Sinclair, K. E. (1997). Path analysis of a theoretical model to describe the relationship among five types of musical performance. *Journal of Research in Music Education, 45*(1), 103–129.

McPherson, G. E., & Renwick, J. M. (2000). Self-regulation and musical practice: A longitudinal study. In C. Woods, G. B. Luck, R. Brochard, F. Seddon, and J. A. Sloboda (Eds.), *Proceedings of the Sixth International Conference on Music Perception and Cognition.* Keele, UK: Keele University, Department of Psychology. CD-ROM.

Miklaszewski, K. (1989). A case study of a pianist preparing a musical performance. *Psychology of Music, 17*(2), 95–109.

Miklaszewski, K. (1995). Individual differences in preparing a musical composition for public performance. In M. Manturzewska, K. Miklaszewski, and A. Biatkowski (Eds.), *Psychology of music today* (pp. 138–147). Warsaw: Fryderyk Chopin Academy of Music.

Mumford, M. D., Costanza, D. P., Baughman, W. A., Threlfall, K. V., & Fleishman, E. A. (1994). Influence of abilities on performance during practice: Effects of massed and distributed practice. *Journal of Educational Psychology, 86*(1), 134–144.

Nielsen, S. (1999). Learning strategies in instrumental music practice. *British Journal of Music Education, 16*(3), 275–291.

O'Neill, S. A. (1997). The role of practice in children's early musical performance achievement. In H. Jørgensen and A. C. Lehmann (Eds.), *Does practice make perfect? Current theory and research on instrumental music practice* (pp. 53–70). Oslo: Norges Musikkhøgskole.

Oxendine, J. B. (1984). *Psychology of motor learning* (2nd ed.). New York: Appleton-Century-Crofts.

Pranger, H. M. (1999). How adults develop computer skills: An extension of deliberate practice theory. (Doctoral dissertation, University of Connecticut.). *Dissertation Abstracts International*, 59/09-A, 3321. (University Microfilms No. 9906558)

Renwick, J. M., & McPherson, G. E. (2000). "I've got to do my scale first!": A case study of a novice's clarinet practice. In C. Woods, G. B. Luck, R. Brochard, F. Seddon, and J. A. Sloboda (Eds.), *Proceedings of the Sixth International Conference on Music Perception and Cognition*. Keele: Keele University, Department of Psychology. CD-ROM.

Ross, S. L. (1985). The effectiveness of mental practice in improving the performance of college trombonists. *Journal of Research in Music Education, 33*(4), 221–230.

Rubin-Rabson, G. (1941a). Studies in the psychology of memorizing piano music: IV. The effect of incentive. *Journal of Educational Psychology, 32*, 45–54.

Rubin-Rabson, G. (1941b). Studies in the psychology of memorizing piano music: V. A comparison of pre-study periods of varied length. *Journal of Educational Psychology, 32*, 101–112.

Rubin-Rabson, G. (1941c). Studies in the psychology of memorizing piano music: VI. A comparison of two forms of mental rehearsal and keyboard overlearning. *Journal of Educational Psychology, 32*, 688–696.

Sloboda, J. A., Davidson, J. W., Howe, M. J. A., & Moore, D. G. (1996). The role of practice in the development of performing musicians. *British Journal of Psychology, 87*(2), 287–309.

Sosniak, L. A. (1985). Learning to be a concert pianist. In B. S. Bloom (Ed.), *Developing talent in young people* (pp. 19–67). New York: Ballantine.

Sullivan, Y., & Cantwell, R. (1999). The planning behaviours of musicians engaging traditional and non-traditional scores. *Psychology of Music, 27*(2), 245–266.

Weinberg, R. S. (1982). The relationship between mental preparation strategies and motor performance: A review and critique. *Quest, 33*(2), 195–213.

Wicinski, A. A. (1950). Psichologiceskii analiz processa raboty pianista-ispolnitiela and muzykalnym proizviedieniem (Psychological analysis of piano performer's process of work on musical composition). *Izviestia Akademii Piedagogiceskich Nauk Vyp, 25*, 171–215.

Williamon, A., & Valentine, E. (2000). Quantity and quality of musical practice as predictors of performance quality. *British Journal of Psychology, 91*(3), 353–376.

Wolfe, D. E. (1987). The use of behavioral contracts in music instruction. In C. K. Madsen and C. A. Prickett (Eds.), *Applications of research in music behavior* (pp. 43–50). Tuscaloosa: University of Alabama Press.

Zurcher, W. (1975). The effect of model-supportive practice on beginning brass instrumentalists. In C. K. Madsen, R. D. Greer, and C. H. Madsen (Eds.), *Research in music behavior: Modifying music behavior in the classroom* (pp. 131–138). New York: Teachers College Press.

# 11

# Memory

RITA AIELLO & AARON WILLIAMON

There is extensive biographical and anecdotal information on the memory of exceptional musicians, but only recently has there been systematic psychological research, and this has mostly focused on pianists. Historical reasons for performing from memory can be traced to Clara Wieck Schumann and Franz Liszt. General theories of expert memory can help us understand how expert musicians memorize music. Auditory, kinesthetic, and visual information contribute to musical memory. Recent psychological research suggests the importance of explicitly analyzing the score. Memory strategies depend on the skill of the performer and the style and difficulty of the music to be memorized. The ability to memorize seems to be enhanced by studying music theory and analysis. Learning to improvise in the style of the music could also be helpful.

During the last century, several pianists and piano pedagogues have written on how to memorize music. Matthay (1913, 1926), Hughes (1915), and Gieseking and Leimer (1932/1972) described three principal ways in which performers can learn music when preparing for a memorized performance: *aurally*, *visually*, and *kinesthetically*. Aural memory (i.e., auditory memory) enables individuals to imagine the sounds of a piece, including anticipation of upcoming events in the score and concurrent evaluations of a performance's progress. Visual memory consists of images of the written page and other aspects of the playing environment. Pianists and other keyboard players, for instance, may remember positions of the hand and fingers, the look of the chords as they are struck, and the patterns made upon the keyboard as they are played. Kinesthetic memory (i.e., finger, muscular, or tactile memory) enables performers to execute complex motor sequences automatically. For pianists, it is facilitated by extended training of the fingers, wrists, and arms and can exist in two forms: (1) position and movement from note to note and (2) sense of key resistance (Hughes, 1915). All these pianists and teachers stressed that no really intelligent memorizing is possible

without a knowledge of musical structure, including harmony, counterpoint, and form. They insisted that aural, visual, and kinesthetic memory could not function properly without this knowledge (Gieseking & Leimer, 1932/1972; Hughes,1915; Matthay, 1913, 1926).

In addition to the study of structure, Gieseking and Leimer emphasized the importance of mental rehearsal when memorizing a piece. They recommended that students learn to memorize pieces by visualizing them through silent reading and prepare for their technical execution through visualization before beginning to play them at the keyboard. The authors illustrated their method with detailed examples from the music of J. S. Bach and Beethoven.

Clearly, performances from memory involve complex interactions among the aforementioned methods of memorizing music. This chapter explores systematic and psychological research that has evaluated such combinations and approaches. Overall, the existing literature on how to memorize music for performance is scarce, and what has been written has focused mainly on pianists. Nevertheless, other instrumentalists can relate to this information in terms of their own memory performance strategies. For singers the interaction of music and text is central in memorizing their repertoire (Ginsborg, 2000). Although space limitation does not allow us to evaluate this interaction adequately, singers could find this chapter useful as a general background.

## Why Perform from Memory?

Until the nineteenth century, when musicians played in public they either improvised the music or read from the score. While it would have been preposterous during the time of Bach, Haydn, and Mozart for musicians to perform their music in public from memory, the Romantic period was marked by a trend toward the individuality of the solo performer, and it is in this era that acclaimed pianists such as Clara Wieck Schumann (1819–1896) and Franz Liszt (1811–1886) began to play in concerts without the score (Schonberg, 1963). Since the early twentieth century the majority of solo pianists and violinists have not used the score when giving performances of the standard repertoire. It is widely acceptable, however, for a contemporary work to be performed using the score.

Performing from memory can be a difficult and anxiety-provoking task. So why do performers insist on doing it? A number of performers and pedagogues have argued that there are fulfilling musical justifications for this tradition. Some of these arguments have been purely practical in their nature. For example, memorizing music dispenses with cumbersome page turns and enables musicians to monitor visually aspects of their performance such as posture and hand positions. Other arguments have focused on specific musical and communicative advantages to memorizing music, claiming for instance that memorization allows performers to develop more freely their own expressive ideas and to communicate those ideas more effectively to audiences (see Bernstein, 1981; Hallam, 1995; Hughes, 1915; Matthay, 1913, 1926).

In an investigation of some of the practical benefits of performing from memory, Williamon (1999b) asked a group of listeners to evaluate videotaped performances of a cellist playing the Preludes from Bach's *Cello Suites* nos. I, II, and III. The performances were divided into five conditions that differed in terms of memorization and visual information available to audience members. In some memorized performances, for example, an empty music stand was used, suggesting to the audience that the performer was reading from the score. The study revealed that audiences preferred memorized performances to nonmemorized performances. Moreover, audience members with musical training rated memorized performances higher than nonmusicians in terms of communicative ability, thus suggesting that listeners who were musicians may have been better equipped to pick up subtle communication cues embedded in performances from memory. In addition, the findings are consistent with the notion that enhanced visual communication (e.g., a view of the performer that is not obstructed by a music stand) augments audiences' experiences at large (Davidson, 1993, 1994).

## General Psychological Research on Memory

### *Classic Studies*

Current knowledge and theories of memory rely heavily on the findings and terminology of classic studies by Ebbinghaus and by Bartlett. With himself as the sole subject, Ebbinghaus (1885/1964) learned lists of nonsense syllables (e.g., MIB, DAK, BOK) over a period of several years. From a set of some two thousand of these stimuli, he constructed lists of a dozen or so syllables and would read them aloud in time with the beat of a metronome. After reading a given list, he would attempt to recall the items at various time intervals, measuring the number of times he had to repeat a list until recall was faultless. Not surprisingly, Ebbinghaus found that the longer the list, the more he needed to repeat the trials to reach perfect recall. Moreover, he found that the longer the interval of time between the original learning and its subsequent recall, the greater the amount of forgetting that occurred. This relationship between memory retention and time after learning is known as Ebbinghaus's forgetting curve. It shows a sharp decline in retention for up to 10 hours after the original learning and then a much more gradual decline across following weeks. The immediate practical implication of this for any student is that material should be rehearsed soon after it is newly learned to avoid such rapid forgetting.

Bartlett (1932) became interested in the kinds of information that people remember rather than the number of trials needed to memorize that information. Instead of asking individuals to learn lists of nonsense syllables, Bartlett used meaningful stimuli such as words, objects, and stories and measured the extent to which people's memories of these changed over time. His method of repeated reproduction, for instance, required participants to read an unusual story from folktale or myth and later recall the story after various periods of time. Bartlett found that recall was often characterized by omissions, simplifications, and trans-

formations, thus demonstrating that memory is often vague and incomplete. For the most part, participants' errors changed the story into more familiar and conventional forms. Bartlett was able to show, therefore, how individuals are able to fill gaps in memory by making logical inferences.

In more recent research, memory has broadly been classified into short-term memory (STM) and long-term memory (LTM). STM is, simply put, the type of memory used when retaining information for only a short period of time (e.g., for 10 seconds). The short-term store can be loosely equated with the concepts of attention and consciousness, and has a working memory component that can be used to manipulate information (see Baddeley, 1986). Working memory is essential for such cognitive tasks as speaking or reading (Baddeley, 1979; Fletcher, 1986) and, in the case of music, for sight-reading.

Across a wide variety of conditions, STM is limited to about seven plus or minus two chunks (Miller, 1956), with a chunk being a unit of information that functions as a single meaningful stimulus. For example, Simon (1974) reported that he was unable to recall the following list correctly after just one presentation because it represented nine chunks of information, an amount that typically exceeds STM capacity: Lincoln, Milky, Criminal, Differential, Address, Way, Lawyer, Calculus, Gettysburg. However, when Simon rearranged the words into four chunks of information—Lincoln's Gettysburg Address, Milky Way, Criminal Lawyer, Differential Calculus—the words were more easily recalled because the four chunks fit comfortably within STM.

LTM is a more-or-less permanent repository of information. It allows us to retain and recall information over long periods of time. Moreover, it is this store that lies at the heart of the exceptional feats of memory often displayed by skilled musicians and expert performers in other fields as discussed later.

### *Expert Memory*

Several theories have been put forth to explain how experts are able to develop and maintain their sometimes-extraordinary memory abilities. Chase and Simon's (1973) chunking theory proposes that superior memory abilities are underpinned by a vast knowledge base specific to the activity. Information in this knowledge base is continually collected and stored into chunks that often become associated with specific physical actions and commands.

The skilled memory theory of Chase and Ericsson (1982) was subsequently proposed to address certain shortcomings of chunking theory. Skilled memory theory asserts that remarkable displays of memory result from the creation and efficient use of mechanisms called *retrieval structures*. These mechanisms can only be acquired under restricted circumstances. First, individuals must be able to store information in LTM rapidly. This requires a large body of relevant knowledge and patterns for the specific type of information involved. Second, the activity must be familiar, so that individuals can anticipate future demands for the retrieval of relevant information. Third, individuals must associate the information to be recalled with appropriate retrieval cues. This association permits the activation of a particular retrieval cue at a later time, and thus partially

reinstates the conditions of learning so that the desired information can be retrieved from LTM. Only after individuals organize the set of retrieval cues can a retrieval structure be formed, thereby enabling them to "retrieve stored information efficiently without lengthy search" (Ericsson & Staszewski, 1989, p. 239).

Extensive research that supports skilled memory theory has been carried out to study the memory capacities of chess players at various levels of expertise. A close investigation of what master chess players do reveals that they abstract meaningful units from the very same material that less experienced players view as separate, single events. In fact, if chess pieces are placed randomly (i.e., in positions that could not have been reached using permissible moves in a game of chess), both experienced and inexperienced chess players score poorly in remembering the positions on the board (Chase & Simon, 1973). Let us illustrate this point within the context of reading and remembering words (Lindsay & Norman, 1977). In reading and remembering the words *Well Tempered Clavier*, inexperienced readers might attempt to learn or process 19 separate letters or three words, while musically experienced readers would immediately recognize that what they read is the title of one work by J. S. Bach. However, when reading and remembering *Llew Derepmet Reivalc* (these are the same three words as *Well Tempered Clavier* but written in reverse) both experienced and inexperienced readers would be at a similar disadvantage.

As a musical example, while inexperienced pianists might strive to remember all the individual notes that occur in the right and left hands in a certain measure, experienced pianists might easily recognize that those notes constituted a cadential formula in a particular tonality. Their knowledge of harmony, together with their advanced keyboard skills, would enable them to play these notes with ease. Basic examples of such chunks in music include scales and arpeggios. Musicians spend hours practicing these so that they can easily recognize and execute them when sight-reading or committing a piece to memory. This process of chunking allows for rapid categorization of domain-specific patterns and accounts for the speed with which experts recognize the key elements in a problem situation. Skilled memory theory has commonly been accepted as accounting for exceptional memory (Anderson, 1990; Baddeley, 1990; Carpenter & Just, 1989; Ericsson & Kintsch, 1995; Newell, 1990; Schneider & Detweiler, 1987). Although subsequent psychologists (e.g., Ericsson & Kintsch, 1995) have extended skilled memory theory to address some of its limitations, the description of the mechanisms that underlie exceptional feats of memory displayed by skilled performers has remained unaltered. For example, a musician who is performing a composition from memory will rely heavily on a hierarchically organized set of preformed retrieval cues (based perhaps on the music's formal structure) to ensure that information is retrieved reliably and efficiently. This retrieval structure may form and develop throughout the course of extensive practice of the piece.

A multitude of issues, however, must be addressed before the skilled memory theory, or any theory of general expertise for that matter, can adequately explain the cognitive processes that govern musical memory. Recent research on music performance has begun to address these issues.

## Music-Psychological Research on Memory

### *Mnemonics, Mental Representation, and Structure*

Musicians use mnemonics extensively. Glancing through any music theory book one sees how frequently musicians use mnemonics to remember the formal structure of a piece. For example, a binary form is outlined as AB, a ternary form ABA, and a theme and variations is A, A', A", A'", and so on. The formal structure of J. S. Bach's *Goldberg Variations* can be remembered according to the plan used by the composer: two variations in free style are followed by a canonic variation, and the final variation is a *quodlibet.*

Music psychologists have used the terms *internal* and *mental representation* to describe the cognitive mechanisms used by musicians. Clarke (1988) asserted that performers retrieve and execute compositions using internal representations. He proposed that to play from memory performers must understand a piece at many different levels, from a complete overall hierarchical memory of the entire piece down to the smallest detail. However, given that it would be impossible for anyone to access the memory of an entire piece at any one time during performance, it is likely that a performer will activate his or her memory of a piece one section at a time, shifting between sections as the performance progresses. As a general rule, "the depth to which the generative structure is activated is directly related to the structural significance of phrase boundaries lying close to, or at, the player's current musical location" (p. 5). In the middle of a deeply embedded musical phrase, for instance, a performer may primarily be concerned with the detailed structure of connections within the phrase itself. Therefore, only a region of low-level generative connections would be active, rather than high-level structural information. Conversely, at a phrase boundary a performer may need to know how the previous and subsequent phrases relate to one another and to the overall structure of the piece. At such a moment, "a small area of low-level structural connections may be active, sufficient to specify the immediate succession of events to be played, together with a section of the higher levels of generative structure specifying larger-scale relationships" (p. 4). Therefore when learning a piece, performers must strive to develop their memory of the overall structure of the composition, and of the way phrases and sections follow and relate to one another.

Lehmann and Ericsson (1995) have provided evidence that internal representations are not only formed and used during performance but can be manipulated. They measured general memorization ability by counting the number of trials required to perform two note-perfect renditions of excerpts from Schubert's Sonata for Violin and Piano, Op. 137. The researchers found significant correlations between memorization ability and ability with other musical tasks (i.e., playing faster and slower than the notated tempo, playing right and left hands individually, and transposing into other keys). They suggested that these abilities are mediated by an underlying mental representation that permits encoded information to be reproduced and manipulated accurately.

Chaffin and Imreh (1994, 1996a, 1996b, 1997) systematically observed the practice of Imreh, a concert pianist, to determine whether she formed an inter-

nal representation when memorizing J. S. Bach's Italian Concerto (Presto) and whether that representation was hierarchically ordered. They examined the pianist's videotaped practice on the piece and her concurrent and retrospective commentary on her practice. In doing so, they confirmed Ericsson and Kintsch's (1995) prediction that skilled performers use hierarchically ordered retrieval schemes to recall encoded information. Chaffin and Imreh found—through their direct observations of practice and the pianist's commentary—that she organized her practice and subsequent retrieval of the Presto according to its formal structure (i.e., Italian rondo form); in her practice she started and stopped more frequently at structural boundaries than in the middle of sections, and she compared the various repetitions of the A and B themes. The pianist's knowledge of the movement's formal structure allowed her to remember when repetitions occurred and the subtle changes between each successive occurrence.

Chaffin and Imreh's research is the first to demonstrate that principles of expert memory (see Chase & Ericsson, 1982; Ericsson & Oliver, 1989) apply to concert soloists. Several issues remain to be resolved with respect to how these cognitive mechanisms relate to musical skill. In particular, how does the formation and use of such structures change as musicians acquire greater levels of overall competence? How do retrieval structures change across the practice process as musicians progressively learn a given composition for performance?

A recent paper by Williamon and Valentine (in press) addressed these questions by examining the practice of 22 pianists—spread across four levels of skill—as they prepared an assigned composition for a memorized performance. In general, the results indicate that the pianists segmented their assigned composition into meaningful sections and reported using those sections in both practice and performance. Empirical examinations of the pianists' practice confirmed their reports in that they, like the concert soloist in Chaffin and Imreh's study, used the music's structure to guide their practice in preparing for the required memorized performance. Despite individual differences in the pianists' identification of structure, the findings held true for all participants. Furthermore, they were greatest for those at higher levels of skill, they increased over the practice process, and their use correlated positively with assessments of performance quality (as judged by expert evaluators). Therefore, the identification and continued use of meaningful structure in practice—regardless of what that structure may be—seems to be an ability that develops with musical competence.

In investigating how experienced pianists and intermediate piano students reported memorizing the same pieces from the piano repertoire, Aiello (2001) found that experienced pianists were more able than the piano students to describe how they memorized piano compositions in terms of the musical structure and reported more often than the piano students that they had memorized the music using some aspects of visual, auditory, or kinesthetic memory. The piano students had difficulties explaining how they had memorized the pieces at all and reported that they relied mostly on rote memory. Experienced pianists tend to conceptualize a composition into independent but linked sections that create a coherent (i.e., meaningful) musical structure, and they organize their practice to start and stop at the beginnings of sections rather than at any point

within the piece. Novice pianists, instead, are more likely to approach a composition as an amorphous whole or as a series of independent notes (Gruson, 1988).

Three interview studies have demonstrated the extent to which professional musicians employ the analysis of the score when memorizing. Hallam (1997) interviewed 22 freelance professional orchestral musicians and 55 novice string players. She found that the strategies adopted by the professional musicians to memorize a composition were related in part to the difficulty of the piece. When memorizing a short, simple piece, some of them felt confident in relying on automated processes, but to memorize longer, more complex works such as a concerto, all of them tended to adopted a more analytic approach. However, the inexperienced musicians reported memorizing by using only repetition and automated process. None of them reported analyzing the music to memorize it. Similarly, Aiello (1999) examined the methods used by professional pianists to memorize piano pieces. They all emphasized that analyzing the score carefully and having a clear idea of the musical structure were the most important and reliable aids to memorization. When concert pianists were asked to give suggestions on how to memorize Bach's Prelude in C Major from Book I of the *Well Tempered Clavier*, they all recommended analyzing the chord progression upon which this piece is based (Aiello, 2000a, 2000b).

In sum, the data of Chaffin and Imreh (1994, 1996a, 1996b, 1997), Chaffin, Imreh, and Crawford (2002), Williamon and Valentine (in press), Williamon (1999a), Hallam (1995,1997), and Aiello (1999, 2000a, 2000b), alongside the theoretical proposals of Chase and Ericsson (1982) suggest that for experienced performers studying the structure of a given composition in detail will provide the most secure foundation for memorizing a work.

Music educators and music theorists have long stressed to performers the importance of studying the structure of musical compositions. Cook (1989), for example, has argued that

> the ability to set aside details and see large-scale connections appropriate to the particular musical context, which is what analysis encourages, is an essential part of the musician's way of perceiving musical sound. For the performer, it is obvious that analysis has a role to play in the memorization of extended scores, and to some extent in the judgment of large-scale dynamic and rhythmic relationships (p. 232).

### *The Role of Musical Style*

When playing from memory, performers draw on not only their knowledge of the composition they are performing but their general knowledge of the musical style in which the piece is written. The style of a composition seems to influence how performers approach memorizing the music. When memorizing a contemporary work, some concert pianists report relying even more on their analytic memory to recall the unique patterns of the composition. Others, instead, depend more on kinesthetic memory because they find that they have to repeat the piece more to have it in their fingers. The pianist Seymour Bernstein articulated a number of difficulties that can be found when memorizing contempo-

rary and serial music and concluded that unless composers incorporate into their writing logical motivic developments, performers are forced to invent their own melodic associations or, indeed, to rely exclusively on their automatic pilot (Bernstein, 1981, p. 258).

Atonal pieces tend to be based on patterns and sonorities that are unique to the piece itself more so than tonal compositions. Not surprisingly, the research with concert pianists shows that they find atonal compositions more difficult to memorize than tonal music (Aiello, 1999; Miklaszewski, 1995). These findings are in agreement with the remarks made by the acclaimed pianists Claude Frank and Rudolf Firkusny during interviews many years ago. Claude Frank noted, "I do not memorize music easily that I do not hear thoroughly. For example, some contemporary music. I can force myself to memorize, but it's hard work, and I tend to forget easily" (Marcus, 1979, p. 59). And Rudolf Firkusny commented, "When you are memorizing complex modern works, the harmonies are more complicated and anything but what you expected. Then you need much more concentration" (Noyle, 1987, p. 84). As psychological research has shown, the more a sequence of items is presented in an unpredictable order, the more difficult it will be for a subject to remember it (Chase & Simon, 1973; Simon, 1974).

### *Different Kinds of Memory*

The information that we process when hearing music, reading a score, and playing an instrument creates memories. But are auditory memory, visual memory, and kinesthetic memory equally valuable in helping us to perform without the score? Which of these memories should we concentrate on to be sure to achieve our goal? The systematic psychological data on this question are scarce. Although most professional pianists report creating dependable memory strategies by relying mainly on the analysis of the musical structure, some also describe using aspects of auditory, visual, or kinesthetic memory to memorize. Others reveal using combinations of these types of memories (Marcus, 1979; Noyle, 1987). Research in cognitive psychology would support the notion that the more ways in which musical information is encoded, the more associations and connections will be formed to that information and, therefore, the more likely an individual is to remember it.

In their methods for memorizing piano music Matthay (1913, 1926) and Gieseking and Leimer (1932/1972) emphasized the importance of auditory and visual memory above kinesthetic memory. Gieseking and Leimer explained: "The fingers are the servitors of the brain, they perform the action the brain commands. If, therefore, by means of a well-trained ear, it is clear to the brain how to execute correctly, the fingers will do their work correctly" (p. 20). In fact, kinesthetic memory seems to be the one that is most questioned by concert pianists. André Watts noted, "I'm very mistrustful of tactile memory. I think it's the first that goes" (Noyle, 1987, p. 147). Watts's comments are in agreement with Leon Fleisher's opinion: "I think probably the least reliable, in terms of public performance, is finger memory, because it's the finger that deserts one first" (Noyle, 1987, p. 97). Kinesthetic memory may be the most helpful to enable children to

perform without the score, and is also what adult music students and inexperienced musicians report relying on to memorize (Aiello, 2000a, 2001; Hallam, 1997). Concert pianists, however, report depending on it only when playing fast virtuoso passages, because at a very fast tempo the automated response is occurring faster than the performer can think about it (Aiello, 1999).

## Implications for Performers and Teachers

By and large, detailed discussions on how to memorize music tend not to be an integral part of the music lesson. Most piano teachers do not spend much time discussing how to memorize the repertoire they assign to their students. Many piano students simply memorize by rote even after they have achieved quite a good level of technical proficiency. Often students do not question at all how they memorize. And piano teachers, seeing that the students have memorized the assigned piece, do not question how the goal was obtained (Aiello, 1999, 2000a, in preparation).

If the students' instrumental technique is more advanced than their theoretical understanding of what they are memorizing, how can they memorize by applying reflection and analysis? Rote learning can bring immediate results. For example, children trained in the Suzuki method can perform pieces from memory after just a few months of music training. And this is certainly a very rewarding and valuable musical experience for the child.

But there comes a time when the benefits of rote learning may fall short of the desired outcome because rote learning does not require a rich, profound understanding of the material (Aiello, 1999, in preparation; Hallam, 1997; chapter 7 in this volume). For memory to become reliable and long-lasting, a deeper understanding is essential. Therefore, we strongly recommend an active dialogue between the teacher and the student on the memory strategies being used by the student and how they may be improved. Students should be encouraged to question how they memorize so that they can better understand how memory works. Taking into account the students' learning styles, teachers can help them achieve this goal.

While many professional musicians use the analysis of the music structure as their primary basis for memorizing a piece, it is not unusual for some students to start memorizing a composition without having analyzed the overall plan of the piece. Students will benefit from studying a composition's structure as the basis for developing their memory strategies (Aiello, 1999, in preparation; Chaffin & Imreh, 1997; Gieseking & Leimer, 1932/1972; Hallam, 1995, 1997; Marcus, 1979; Matthay, 1913, 1926). They should come to value the importance of describing and analyzing all the theoretical aspects of the pieces to be memorized. When asked to discuss the music they practice and perform from memory, some students reveal that they have not fully integrated into their piano playing what they learned in music classes other than their piano lessons. It seems as if some students tend to compartmentalize what they learn in theory classes, analysis classes, and their piano lessons into separate domains, not seeing that there

is a common denominator to musical knowledge. Those students who do not fully integrate the various components of their musical knowledge fail to make the inherent connection between music performance and music theory, the same connection, in fact, that has been shown to be an integral part of how concert soloists memorize successfully (Aiello, 1999, 2000b). On the one hand, this is a sad commentary on education at large. On the other hand, it is encouraging to know that, as teachers, we can improve our students' abilities to memorize by helping them integrate what they have learned from various sources. The instrumental teacher can help students understand that *there really is* a relationship between performance and theory. Some students may benefit from hearing their instrumental teachers give very explicit explanations about the musical structure of the piece they play and analyzing the piece along with the teacher during the instrumental lesson.

In addition, teachers might encourage their students to develop their auditory memory, explaining that above all, music *is* sound. Gieseking and Leimer placed training the musical ear as a foundation of their method and wrote, "Listening to one self is the most important factor of the whole of music study" (p. 10). Matthay, too, emphasized the fundamental value of developing the ear when learning to memorize. He warned: "There is nothing more fatal for our musical sense, than to allow ourselves—by the hour—to *hear* musical sounds without *listening* to them" (1913, p. 5). He explained how effective listening implies *pre*-listening all the time as to what the sounds should be. Whether the music to be performed is heard inwardly from memory, or is read from the score, or is heard externally, students must be aware that all playing always involves the ear (Priest, 1989; chapter 7 in this volume).

Furthermore, learning how to improvise can be helpful for memorizing music because improvisation requires having internalized the characteristics of a particular musical style to the point of being able to create a novel piece spontaneously. While in the past keyboard performers were expected to know how to improvise well, today relatively few classically trained piano students are expected to develop advanced improvisatory skills (Gellrich & Parncutt, 1998; chapter 7 in this volume). In interviews conducted with seven concert pianists, six of them considered their improvisatory skills to be helpful in the process of memorization; interestingly, even the pianist who reported that she did not know how to improvise thought that improvisation could be helpful for others (Aiello, 1999). The relationship between improvisatory skills and memorization skills could be researched further and could be addressed more frequently within the context of the piano lesson. In sum, students may be taught that memory is based on knowledge, on information that has been internalized meaningfully, and on the connections that are made between different aspects of what one knows and senses.

But where can a piano teacher start when students are not aware of the various types of memories that they could develop and are not able to integrate effectively the various components of their musical knowledge? A way to encourage a student become more cognizant of memory strategies could be for the teacher to ask: "What specific suggestions would you give to a technically proficient fellow student to help him or her memorize this piece?" In her teaching,

Aiello has found that asking this question helps students become more aware of what strategies they use when memorizing a piece. This question can make them channel their own memorization strategies into specific instructions that must be explicit enough for a fellow piano student to understand and to follow. Based on the teaching experience of the first author of this chapter, the following suggestions have been found to be helpful in teaching students how to memorize.

### *Suggestions Based on the Analysis of the Piece*

- Describe and analyze the piece in terms of its macrostructure (the overall form) and microstructure (movements, sections, major themes, sequences, characteristic intervals, modulations, key areas, rhythmic patterns, dynamics, phrasing, etc.) (see Matthay 1913, 1926; Gieseking & Leimer, 1932/1972).
- Learn the landmarks of the piece. Describe where they occur in the piece and why. What leads to them? What occurs after them?
- Describe in detail how the various sections of each movement (or of the piece) are linked together.
- Highlight and describe the melodic and rhythmic patterns in the piece. Explain what they contribute to the music.
- Use markers of different colors to highlight the various themes or voices and their recurrences in the piece. Describe what you did and why.
- Discuss in detail the harmonic structure of the piece (modulations, key regions).
- Mark the closures and the points of tension and resolution in the piece. Describe them.
- Describe the gestures and the dynamics of the piece and how they are related to the overall structure of the composition.
- Based on the analysis and the characteristics of the piece, describe what strategies would be best for memorizing this particular piece. Explain why.
- Memorize in sections. Use the formal structure of the piece to create logical sections.

### *Suggestions Based on the Performance of the Piece*

- Practice each hand separately, and describe what each hand is supposed to play.
- Rehearse the piece mentally. Practice away from the piano visualizing the score, visualizing the keyboard, and most of all *hearing* the music in your mind (Gieseking & Leimer, 1932/1972).
- Sing the various themes or voices. Sing one voice while playing another; sing the melody while playing the accompaniment.
- Play at a slow tempo, *reflecting* upon the structure and the various patterns in the piece.
- Move to the rhythm, the tempo, and the gestures of the music.
- Learn to improvise in the style of the music to be memorized. Through the development of improvisatory skills students can acquire additional means of encoding information based on the characteristics of a particular musical style. The knowledge and the instrumental dexterity required

to improvise according to a given style can help students in developing their memory strategies and in overcoming a memory lapse during a performance. Overall, knowing how to improvise in the style of a memorized piece can contribute to a musician's knowledge and therefore can add to a performer's sense of confidence.

These suggestions may be of more use to some performers than to others. Furthermore, some teachers may realize that only some of these strategies may be suited for a particular student, given his or her learning style. Regardless, we want to stress that by learning to reflect on how they memorize and by acquiring multiple systems to remember a piece students may gain a deeper understanding of how their memories work. Teachers can be the catalysts in this important educational process.

### *References*

Aiello, R. (1999). Strategies for memorizing piano music: Pedagogical implications. Work in progress presented at the Eastern Division of the Music Educators National Conference, New York, NY.

Aiello, R. (2000a). The analysis of the score as a basis for memory. Poster presented at the Music Cognition / Music Pedagogy Group, Society for Music Theory, Toronto, Canada.

Aiello, R. (2000b). Memorizing two piano pieces: The recommendations of concert pianists. In C. Woods, G. Luck, R. Brochard, F. Seddon, and J. A. Sloboda (Eds.), *Proceedings of the Sixth International Conference on Music Perception and Cognition*. Keele, UK: Department of Psychology, Keele University.

Aiello, R. (2001). Playing the piano by heart: From behavior to cognition. In R. J. Zatorre and I. Peretz (Eds.), *The Biological Foundations of Music*. Annals of the New York Academy of Sciences, *930*, 389–393.

Aiello, R. (in preparation). Reflections on learning music.

Anderson, J. (1990). *Cognitive psychology and its implications*. San Francisco, CA: Freeman.

Baddeley, A. D. (1979). Working memory and reading. In P. A. Kolers, M. E. Wrolstad, and H. Bouma (Eds.), *Processing of visible language*, Vol. 1. New York: Plenum.

Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon Press.

Baddeley, A. D. (1990). *Human memory: Theory and practice*. Boston: Allyn & Bacon.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.

Bernstein, S. (1981). *With your own two hands: Self-discovery through music*. New York: Schirmer.

Carpenter, P. A., & Just, M. A. (1989). The role of working memory in language comprehension. In D. Klahr and K. Kotovsky (Eds.), *Complex information processing* (pp. 31–68). Hillsdale, NJ: Erlbaum.

Chaffin, R., & Imreh, G. (1994). *Memorizing for performance: A case study of expert memory*. Paper presented at the Third Practical Aspects of Memory Conference, College Park, MD.

Chaffin, R., & Imreh, G. (1996a). Effects of difficulty on practice: A case study of

a concert pianist. Poster presented at the Fourth International Conference on Music Perception and Cognition, McGill University, Montreal, Canada.

Chaffin, R., & Imreh, G. (1996b). Effects of musical complexity on expert practice: A case study of a concert pianist. Poster presented at the Thirty-seventh Meeting of the Psychonomic Society, Chicago.

Chaffin, R., & Imreh, G. (1997). "Pulling teeth and torture": Musical memory and problem solving. *Thinking and Reasoning, 3*(4), 315–336.

Chaffin, R., Imreh, G., & Crawford, M. (2002). *Practicing perfection: Memory and piano performance.* Mahwah, NJ: Erlbaum.

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. H. Bower (Ed.), *Psychology of learning and motivation*, Vol. 16. New York: Academic Press.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215–281). New York: Academic Press.

Clarke, E. F. (1988). Generative principles in music performance. In J. A. Sloboda (Ed.), *Generative processes in music* (pp. 1–26). Oxford: Clarendon Press.

Cook, N. (1987). *A guide to musical analysis.* New York: W. W. Norton.

Davidson, J. W. (1993). Visual perception and performance manner in the movements of solo musicians. *Psychology of Music, 21*, 103–113.

Davidson, J. W. (1994). Which areas of a pianist's body convey information about expressive intention to an audience? *Journal of Human Movement Studies, 26*, 279–301.

Ebbinghaus, H. (1885/1964). *Memory.* New York: Dover.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211–245.

Ericsson, K. A., & Oliver, W. (1989). A methodology for assessing the detailed structure of memory skills. In A. M. Colley and J. R. Beech (Eds.), *Acquisition and performance of cognitive skills* (pp. 193–215). Chichester, UK: Wiley.

Ericsson, K. A., & Staszewski, J. J. (1989). Skilled memory and expertise: Mechanisms of exceptional performance. In D. Klahr and K. Kotovsky (Eds.), *Complex information processing* (pp. 235–267). Hillsdale, NJ: Erlbaum.

Fletcher, C. R. (1986). Strategies for the allocation of short-term memory during comprehension. *Journal of Memory and Language, 25*, 43–58.

Gellrich, M., & Parncutt, R. (1998). Piano technique and fingering in the eighteenth and nineteenth centuries: Bringing a forgotten method back to life. *British Journal of Music Education, 15*(1), 5–23.

Gieseking, W., & Leimer, K. (1932/1972). *Piano technique.* New York: Dover.

Ginsborg, J. (2000, August). *Off by heart: Expert singers' memorization strategies and recall for the words and music of songs.* Paper presented at the Sixth International Conference on Music Perception and Cognition, Keele University, UK.

Gruson, L. M. (1988). Rehearsal skill and musical competence: Does practice make perfect? In J. A. Sloboda (Ed.), *Generative processes in music* (pp. 90–112). Oxford: Clarendon Press.

Hallam, S. (1995). Professional musicians' approaches to the learning and interpretation of music. *Psychology of Music, 23*(2), 111–128.

Hallam, S. (1997). The development of memorisation strategies in musicians: Implications for education. *British Journal of Music Education, 14*(1), 87–97.

Hughes, E. (1915). Musical memory in piano playing and piano study. *Musical Quarterly, 1*, 592–603.

Lehmann, A. C., & Ericsson, K. A. (1995). Expert pianists' mental representation of memorized music. Poster presented at the Thirty-sixth Annual Meeting of the Psychonomic Society, Los Angeles, CA.

Lindsay P., & Norman, D. (1977). *Human information processing: An introduction to psychology.* New York: Academic Press.

Marcus, A. (1979). *Great pianists speak.* Neptune, NJ: Paganiniana.

Matthay, T. (1913). *Musical interpretation: Its laws and principles, and their application in teaching and performing.* Boston, MA: Boston Music.

Matthay, T. (1926). *On memorizing and playing from memory and on the laws of practice generally.* Oxford: Oxford University Press.

Miklaszewski, K. (1995). Individual differences in preparing a musical composition for public performance. In M. Manturzewska, K. Miklaszewski, and A. Bialkowski (Eds.), *Psychology of music today* (pp. 138–147). Warsaw: Fryderyk Chopin Academy of Music.

Miller, G. A. (1956). The magic number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–93.

Newell, A. (1990). *Unified theories of cognition.* Cambridge, MA: Harvard University Press.

Noyle, L. (1987). *Pianists on playing: Interviews with twelve concert pianists.* Metuchen, NJ: Scarecrow.

Priest, P. (1989). Playing by ear: Its nature and application to instrumental learning. *British Journal of Music Education, 6*(2), 173–191.

Schneider, W., & Detweiler, M. (1987). A connectionist / control architecture for working memory. In G. H. Bower (Ed.), *Psychology of learning and motivation,* Vol. 21 (pp. 54–119). New York: Academic Press.

Schonberg, H. (1963). *Great pianists from Mozart to the present.* New York: Simon & Schuster.

Simon, H. A. (1974). How big is a chunk? *Science, 183,* 482–488.

Williamon, A. (1999a). *Preparing for performance: An examination of musical practice as a function of expertise.* Unpublished doctoral dissertation, University of London.

Williamon, A. (1999b). The value of performing from memory. *Psychology of Music, 27,* 84–95.

Williamon, A., & Valentine, E. (in press). The role of retrieval structures in memorizing music. *Cognitive Psychology.*

# 12

# Intonation

STEVEN J. MORRISON & JANINA FYK

Rather than an isolated ability, intonation is an amalgam of several sub-skills including pitch discrimination, pitch matching, and instrument tuning. Success at these skills depends on many factors including musical experience and the nature of the task presented. However, ability in any one of these areas is not clearly related to ability in the others. The skill that musicians demonstrate at isolated intonation-related tasks does not necessarily reflect their performance within a typical, complex musical setting. Within a real musical context, mature performers deviate considerably from common standards of pitch measurement, suggesting that absolute pitch accuracy should not be the ultimate goal in this area of musical development. The ability to identify and produce the most appropriate pitch within a given musical context may emerge parallel to general musical development as students' awareness of the performers around them increases and they develop their own concepts of ideal performances.

*Intonation* is an imprecise term. When teachers or performers consider intonation, they may be addressing one or more of several skills that fall under this general heading.

At the most basic aural level, they may be referring to *pitch discrimination*, which can be defined as the ability to distinguish between two successive pitches or two dissimilar examples of a single pitch. For example, students may be presented with a C and a C♯ and asked to identify the higher pitch. Or they may be given an A at 440 Hz and an A that is 15 cents flat (436.2 Hz) and asked to decide if there is a difference.

Or, combining aural and performance skills, focus may be on *pitch matching*, where one attempts to reproduce exactly a given pitch. This may involve isolating one pitch at a time or, occasionally, a sequence of individual pitches. Though the target is usually at unison, students may be requested to produce an

octave or an intervening interval such as a fourth or fifth. Tuning may be regarded as a sort of ritualized pitch matching that usually takes place near the beginning of instrumental rehearsals. Procedures for tuning differ from ensemble to ensemble. Some may be very teacher-directed, others more student-centered. Some may be assessed by machine, others by the teacher's ear, and still others by the students' ears. The target may be an A (for orchestras), a B♭ or F (for bands), or, for the tuning of string instruments, an interval series.

Finally, we arrive at *intonation*. For the purposes of this chapter, this will refer to the manipulation of pitches and intervals within a real musical context. It is often measured relative to the *just*, *Pythagorean*, or *equally tempered* systems. Pythagorean tuning is theoretically derived from combinations of perfect octaves (with frequency ratios of exactly 1:2) and perfect fifths (with frequency ratio 2:3); just intonation from combinations of octaves, fifths, and major thirds (with frequency ratio 4:5); and equal temperament by dividing the octave into 12 equal semitones; for a thorough examination of tuning systems, see Barbour (1951). Intonation is necessarily culture-bound, assessed relative to the norms and according to the criteria of a given musical tradition. Because much of the research literature has focused on Western classical music, the scope of this chapter is necessarily limited to that one tradition. This chapter will survey what research has revealed about each of the component skills identified earlier (pitch discrimination, pitch matching, and intonation), with discussions of how they are demonstrated by developing and advanced musicians, and how findings may impact music teaching.

## An Overview of Research

### *Pitch Discrimination*

Good intonation is characteristic of a musically *sensitive* performance; it supports the beauty and expressive qualities of the sound. But it is not a prerequisite for a musically *comprehensible* performance, one in which the basic musical elements are discernible to a listener. For a performance to be comprehensible to a listener, it must convey reasonably accurate and specific information that includes relationships among simultaneous pitches (as in chords) and successive pitches (as in melodies). For this information to be understood, listeners must be capable of pitch *discrimination*. (For a detailed examination of perceptual research related to intervals, scales, and tuning, see Burns, 1999).

The most minute level of discrimination ability is described through the measurement of *just noticeable differences of frequency*, or *JND*. Magnitude of the JND threshold can vary depending on a listener's musical experience and training and the testing procedures used. Given favorable experimental conditions, listeners have been found to identify differences smaller than 2 cents (0.02 semitone) (Fyk, 1982; Rakowski, 1978; Vos, 1982). Placing pitches in a musical context—a chord or a melody, for example—affects discrimination accuracy, though exact results have varied. Accuracy may be impaired by the increased

information to which the listener must attend (Watson, Kelly & Wroton, 1976). Alternatively, the presence of more musical information may allow the listener to make a more thorough evaluation (Wapnick, Bourassa, & Sampson, 1982).

In order to understand the musical information being conveyed in a performance, listeners must not only identify pitches as same or different but must also sort pitches according to melodic or harmonic function. Such sorting requires the listener to disregard strict pitch accuracy and focus instead on the intervallic relationship between pitches according to learned conventions. Within a given scalar convention several pitches quite close in frequency might be identified as the same pitch insomuch as we would call them by the same note name. Similarly, within a familiar harmonic or melodic context several intervals quite close in size might be assessed as the same interval. Researchers have labeled this phenomenon *categorical* or *zonal* perception (Fyk, 1995; Garbuzov, 1948; Siegel & Siegel, 1977).

This phenomenon has been likened to perception of vowels and consonants in speech (Burns & Ward, 1978; Rakowski, 1990). When conversing with other people, we can comprehend their meaning even if their pronunciation of *th* or *ay*, for example, differs from our own. Similarly in music, we will usually recognize the identity of an interval or tone even if it is considerably out of tune. For example, we give intervals of 280 and 320 cents the same identity (minor third). The pitches 280 and 320 cents above C are both identified as E♭/D♯. This is true even though we perceive that the two intervals or pitches are not identical. In this way we are able to recognize "Ode to Joy" whether it is performed by a beginning band or the Berlin Philharmonic.

Pitch discrimination ability can be affected by variables that include intensity (Sergeant, 1973), timbre (e.g., simple vs. complex synthesized tones in Platt & Racine, 1985), direction of error (Geringer & Witt, 1985; Kelly, 2000), tempo, and tone duration (Fyk, 1985). Tone quality in particular has been found to play a significant role in listeners' perceptions of absolute pitch and interval size. Listeners tend to misinterpret tone quality errors as pitch inaccuracy and pitch errors as poor tone quality (Geringer & Worthy, 1999; Madsen & Geringer, 1976, 1981; Wapnick & Freeman, 1980). Clearly performance skills such as embouchure, breath support, and even posture—skills not directly associated with pitch manipulation (i.e., finger placement, slide position)—contribute significantly to what listeners recognize as an in-tune performance. We may also speculate that the very term *in-tune performance* is less precise than it would seem. To many listeners, in-tuneness might be a more global assessment of musical quality than just pitch precision.

### *Pitch Matching*

Perception of small differences among pitches is a separate skill from manipulation of pitches to minimize those differences. While discrimination requires a listener to assess pitch relationships from a more or less detached vantage point, *pitch matching* requires an individual to take an active role in the shaping of sounds. The individual still acts as a listener, of course, but the focus becomes

assessment of the results of one's own actions rather than the examination of an external sonic artifact.

Skill at both pitch discrimination and pitch matching has been found to be better among musicians than in nonmusicians (Brown, 1991; Fastl & Hesse, 1984; Spiegel & Watson, 1984) and among more experienced musicians rather than those with less experience (Madsen, Edmonson, & Madsen, 1969; Platt & Racine, 1985; Tieplov, 1947). This might be expected since matching pitch using an instrument or the voice is a task performing musicians would be required to perform regularly. Nonperformers, however, may infrequently or never have to face this challenge.

Despite these findings, no significant relationship has been found between pitch discrimination skills and success at pitch matching. Preschool and fourth-grade students who were grouped according to success on a pitch discrimination pretest subsequently completed a vocal pitch-matching task (Geringer, 1983). Except for a moderate positive correlation found among the highest ability fourth-graders, no significant relationship was found between results of the two tasks. Similarly, only the most accurate of 28 10-year-old singers demonstrated a relationship between pitch-matching accuracy and understanding of pitch concepts (Fyk, 1985).

Geringer and Witt (1985) asked secondary-level, postsecondary, and professional string performers to tune to a given pitch and then provide a verbal assessment of the accuracy of the tuning stimulus compared to an equally tempered standard. Though the more experienced players demonstrated more frequent agreement between their tuning and verbal judgment, neither group showed strong agreement between the two tasks. In another study, undergraduate and graduate woodwind players performed an eight-note melodic fragment along with a stimulus tape, listened to recordings of the resulting duets, and attempted to identify out-of-tune tone pairs (Ely, 1992). Again there was no significant correlation between the accuracy of their performances and their success at identifying pitch errors.

Even relationships between seemingly similar tasks have proven elusive. In a series of studies that tested middle school, junior high, and senior high wind players, subjects attempted to match a single target pitch using both their instrument and a variable-pitch electronic keyboard (Yarbrough, Karrick, & Morrison, 1995; Yarbrough, Morrison, & Karrick, 1997). Though performance on both tasks was more accurate for each successive level of experience, no correlation was found between subjects' instrument and keyboard scores. A similar procedure was used with choir students when junior high school boys attempted to match a target pitch with both their voice a variable-pitch electronic keyboard (Demorest, 2000). Again, no significant relationship was found between accuracy of vocal responses and accuracy of keyboard responses. Much research of this type has necessarily focused on broad groups of school-aged musicians. Perhaps, as some results suggest (Fyk, 1985; Geringer, 1983), clearer relationships among the various skills related to pitch will be found by concentrating future research efforts on the most advanced performers. Though a purely speculative question, might it be that the *interactions* of skills such as pitch discrimination and pitch matching may be part of what *defines* advanced performers?

### *Adding the Musical Context*

It is unlikely that musicians at any level approach the singing of a song or the playing of a melody as an extended series of pitch-matching tasks. When a performer is faced with placing pitches within a musical setting an array of considerations is introduced that may impact pitch accuracy. (For an extended program of research in this area, see Geringer and Madsen, 1987). For less-experienced performers, the number of competing demands may be so great that attention is not available to monitor fine pitch adjustments. For more experienced performers, especially the most advanced artists, not only can they monitor fine pitch adjustments, but they can also make those adjustments to serve what they believe to be the best and most expressive musical ends (Fyk, 1995).

Experienced musicians demonstrate a tendency toward sharp performance when compared to an equally tempered standard in both solo and accompanied settings (Geringer, 1978; Geringer & Witt, 1985; Kantorski, 1986; Yarbrough & Ballard, 1990). This tendency appears to emerge among students over the first four years of performance study (Yarbrough, Karrick, & Morrison, 1995). It is most frequently observed in the performance of intervals of a third or larger, while intervals smaller than a third are often compressed (summarized in Burns, 1999). Other evidence suggests that even listeners prefer performances in which pitches are sharper and intervals are wider than an equally tempered standard (Sundberg, 1982). The octave in particular has been the focus of considerable attention. Both performers and listeners tend to favor octaves that are larger than the pure 1:2 ratio (e.g., Burns & Campbell, 1994), a phenomenon referred to as octave stretching.

It seems clear that the pitch choices advanced performers make do not conform to a specific tuning system. Nevertheless, these choices are considered acceptable and even preferable (Karrick, 1998). Most recent evidence suggests that intonation in Western music is typically closer to equal temperament (Karrick, 1998; Rakowski, 1990) though some results find the Pythagorean system to be an equally good fit (Loosen, 1993). Given these and other similar results, it is probably inappropriate to say that performers play *in* a given system. Rather, it is more accurate to identify the system closest to a given performance.

Among less-advanced performers, pitch accuracy appears to be compromised by the demands of a musical setting. Primary and secondary school band students performed an eight-measure melody in unison along with a prerecorded equally tempered model (Morrison, 2000). Four target pitches, identical in pitch class, were extracted from each response—one pitch approached from above, one from below, one repeated, and one of extended duration. There were no differences in accuracy among the four pitches. There was also a high positive correlation among the four pitches. Many of the students were also asked to tune their instrument to a single target pitch before playing the melody, similar to the tuning procedure they carry out in their daily band class. Students at all levels performed the single pitch significantly more accurately. Though the exact context of each pitch did not appear to have an impact on student performances, the overall addition of a musical context significantly affected student's accuracy.

Other findings suggest that melodic direction does have a significant effect on pitch accuracy. Advanced string instrumentalists have been observed to perform descending whole-tone tetrachords significantly sharper than ascending tetrachords when compared to an equally tempered standard (Kantorski, 1986; Sogin, 1989). Conversely, Yarbrough and Ballard (1990) found that advanced string players performed significantly sharper when ascending than descending. Overall pitch deviation was assessed by Madsen (1966), who found that primary, secondary, and postsecondary vocalists, pianists, and violinists, sang descending scales more closely to an equally tempered standard than ascending patterns. Along with melodic direction, melodic range may significantly impact pitch accuracy. Advanced performers have demonstrated greater pitch variation across registers than between each other (Fyk, 1995).

### *Improving Pitch-Matching and Intonation Skills*

Several common characteristics have been identified among performers who demonstrate good pitch accuracy. First, as discussed earlier, accuracy appears to improve with experience. Beginning instrumental students have demonstrated difficulty accurately matching a single pitch with their instrument (Yarbrough et al., 1995). As shown in Figure 12.1, from that point improvement in accuracy appears to continue for at least seven years (Yarbrough et al., 1997). It was also noted that students who took private instrumental lessons in addition to participating in an ensemble performed more accurately.

Several issues remain unclear. First, it is not known whether private instruction improves pitch skills or students who study privately spend more time performing and, therefore, develop accuracy more quickly. Alternatively, students who are more successful at pitch-related tasks may find performance rewarding enough to seek individual instruction. Second, in light of the overall trend toward improved accuracy with experience, it remains unclear whether experience itself contributes toward improved pitch skills or students who are not successful at such tasks tend to quit their school music programs earlier in their performance careers. This conclusion may be supported by studies that did not find that experience alone resulted in more accurate performance (Brittin, 1993; Duke, 1985).

The placement and characteristics of target pitches also affect accuracy. Corso (1954) observed no difference in accuracy when advanced instrumental performers attempted to tune to simultaneous or preceding target pitches. Cassidy (1989) found that students tuned simultaneous octaves at least as well as unisons. While instrumentalists appear to be little affected by the timbre of the target pitch (Ely, 1992), timbre has a greater effect on the pitch-matching accuracy of young vocalists. Primary school children have been found to be more successful performing a minor third (G–E) following a child or female model rather than a male model (Yarbrough, Green, Benson, & Bowers, 1991). However, a male falsetto model has been found to be effective (Yarbrough et al. 1995). Among young students, at least, it appears advantageous to present students with unison pitch targets.

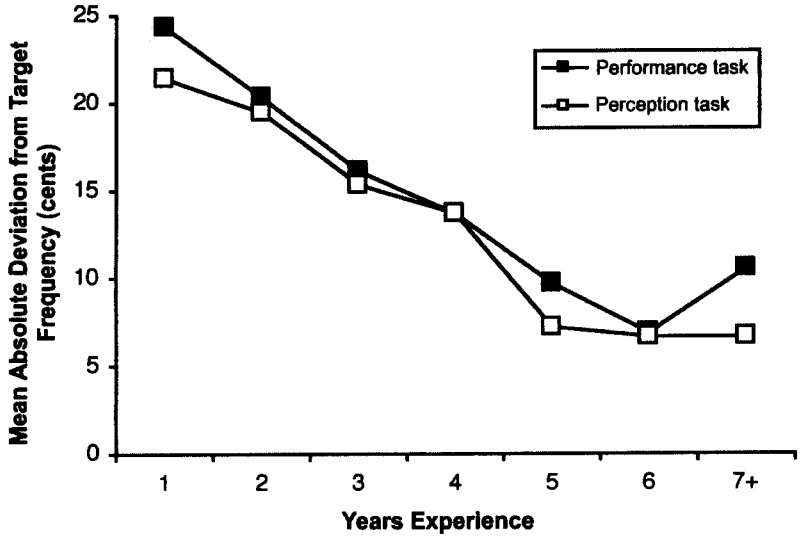



Figure 12.1. Beginning, intermediate, and advanced wind players' mean absolute cent deviation scores when matching a target pitch with their instrument (dark squares) and an electronic keyboard (white squares). Students with greater experience show more accuracy on both tasks. (Adapted from Yarbrough, Morrison, & Karrick, 1997. Copyright 1997 the University of Illinois, used by permission.)

The presence of vibrato in a target stimulus has also been found to affect the accuracy of pitch responses. Brown (1991) asked musicians and nonmusicians to match pitches presented with wide and narrow vibrato. Responses of the two groups were significantly different, with musicians' responses generally above the mean pitch and nonmusicians' responses at or below mean pitch. This contradicts other research in which the mean frequency was identified as the best match within a vibrato tone (Brown & Vaughn, 1996; Shonle & Horan, 1980). At the early stages of music instruction, the presence of vibrato may be a source of confusion. Yarbrough, Bowers, and Benson (1992) found that both certain and uncertain singers in kindergarten through Grade 3 responded more accurately to a nonvibrato adult female vocal model than to a child vocal model or an adult female model using vibrato. Uncertain singers particularly benefited from the nonvibrato model.

In an effort to improve pitch accuracy, ensembles generally tune at the outset of a rehearsal or performance. At the beginning of this chapter we identified tuning as unique from other pitch-related tasks, though its prominence within the rehearsal context suggests belief in a strong link between an ensemble's accuracy at tuning and their pitch accuracy in performance. To examine students' ability to transfer pitch information from the tuning process to a performance situation, Morrison (2000) recorded high school wind instrumentalists individually performing a familiar eight-measure melody along with a prerecorded model. Before per-

forming, the students either (1) tuned their instrument to a target pitch, (2) were verbally instructed to perform as in tune as possible, or (3) received no information or tuning opportunity. No differences were found in the accuracy of their performances. Similar results were reported by Geringer (1978), who found that cautioning undergraduate and graduate music students about tuning did not significantly improve the accuracy of their performance. Surprisingly, even rehearsal did not improve the pitch accuracy of primary, secondary school, and undergraduate students' performances of ascending and descending scales (Madsen, 1966).

Specific feedback, rather than instructions alone, has been found to improve pitch-matching accuracy. Among 32 musically experienced and inexperienced adults, visual and aural feedback given across four treatment sessions was found to improve subjects' ability to adjust an electronic tone to match a target pitch (Platt & Racine, 1985). First-year college students also demonstrated improved pitch-matching skills after taking part in aural skills coursework that emphasized instructional feedback (Fyk, 1987).

### *Development of Intonational Skill*

As stated earlier in this chapter, the musical characteristic we refer to as intonation represents pitch manipulation within a real musical context. Perhaps the improvement in this skill observed over time and experience reflects developing musicians' ability to interact within larger and larger musical environments. The following model proposes a view of how development of pitch manipulation skills parallels the growth of a young musician's musical context. Within this model the extent of young performers' musical awareness corresponds with the areas of musical skill to which they can attend. These, in turn, help the teacher to identify appropriate learning objectives.

Young musicians at the most basic level of performance operate within a very limited context, focusing solely on themselves and their actions. This is not surprising given the array of cognitive challenges they are suddenly facing. As depicted in Figure 12.2, attention is often consumed with their *internal self*—the feeling of manipulating an instrument or the process of producing a sound. These experiences are novel and, at first, feel awkward and unnatural. Members of beginning ensembles often wander off into their own worlds of tempo, melody, and rhythm, oblivious to the ensemble's disjointed results. Even the sounds that emerge from their own instruments (including the voice) are often overlooked in favor of attending to the physical demands of basic sound production.

At the next level, students become more aware of their *external self*—the different kinds of sounds they are able to produce, the results of their physical processes. They are able to select from and perform a variety of pitches that have become part of their repertoire and are able to discriminate between what is generally right and what is definitely wrong. Trumpet players can make judgments about which partial should be played; violinists can make decisions about correct finger placement. These judgments may be at a gross level—should that be an F or an F♯?—but they show a significant shift of attention toward self-generated musical information.

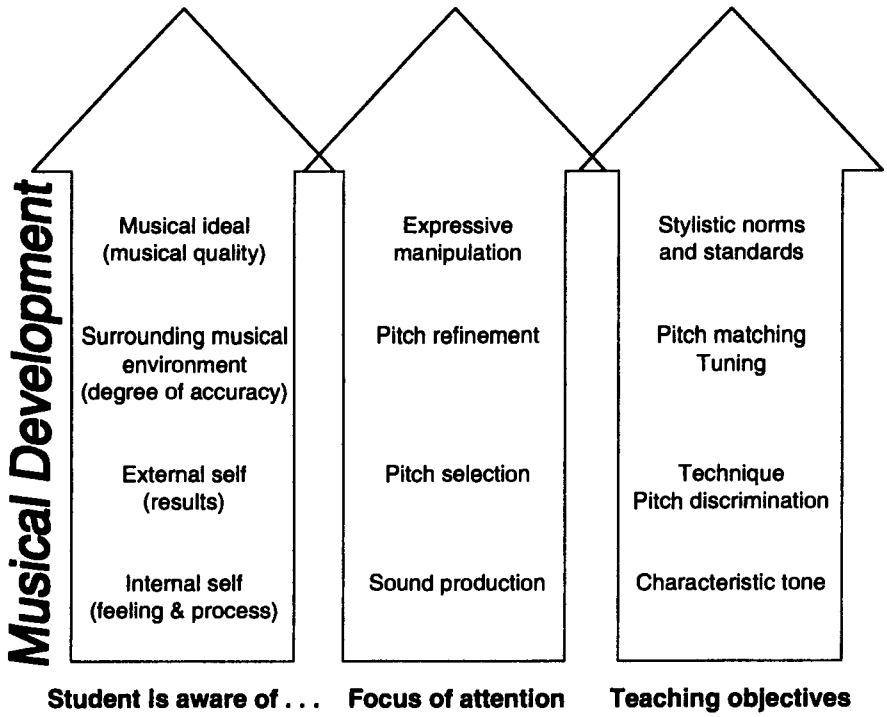



Figure 12.2. Model that describes the growth of pitch manipulation skills within the overall context of musical development.

The expansion of the musical context continues as young musicians match the sounds they are making to the sounds happening around them. They are comfortable enough with their own technique that they can attend to the *surrounding musical environment.* Accuracy is assumed as degree of accuracy becomes the focus. Not only can performers at this level select the proper pitch, but they can refine that pitch so that it matches a reference point established by other members of the group or some other external sound source (e.g., an electronic tuner).

At the most advanced stage, students can not only compare the relationship between their sounds and that of others around them, but they can also compare that comparison to an abstract standard, a *musical ideal.* They can manipulate their own actions so that the result is not only correct but also expressive. At this point, the musical context has extended beyond their immediate environment to include an awareness of historical and cultural—in short, professional—standards. Taken together with key findings that emerge from research surveyed earlier in this chapter, this model offers potential guidance for music teaching and performance.

## Implications for Teaching

### *Tone Quality Affects Perception of Pitch Accuracy*

If intonation is a process that demands at least three levels of awareness—the performers' own actions, the actions of those around them, and an abstract model of ideal performance—then it would seem well beyond the grasp of novice performers. However, that is if we limit our view of intonation to the active process of pitch manipulation. Let us instead view intonation as a result of other active processes—particularly production of a characteristic tone—that, when executed successfully, result in a more in-tune performance. Certainly intonation-as-action is one of the ultimate goals of music performance instruction, but intonation-as-result can be a realistic outcome along the way.

The development of good, characteristic tone quality is a skill within reach of even the most inexperienced beginner. As a result of the development of good tone quality, performers and performing ensembles are likely to produce performances that sound more in tune. Tone quality is a direct consequence of skills that reside at the foundation of performance instruction: correct posture, breath technique, vowel formation, embouchure, bow grip, and instrument carriage. Most important, students can successfully practice and demonstrate these skills by focusing solely on their own actions, a key characteristic of early musical development.

### *Pitch Discrimination and Manipulation Skills Are Not Clearly Related*

It seems reasonable to suggest that it is the interaction among pitch discrimination, pitch matching, and melodic and harmonic pitch placement (the task to which we have specifically applied the term *intonation*) that results in an in-tune musical performance. Indeed, the model presented earlier depicts these subskills as sequential building blocks toward performance proficiency. However, it may not be accurate to say that all students must demonstrate high achievement in each of these areas in order to produce a successful final result.

Students' skills at pitch discrimination or isolated pitch matching are often used as predictors of their contribution to a performance ensemble or future success in music. As only limited correlation has been found among these skills, it may be inappropriate to speculate on achievement in one area by assessing achievement in another. For example, the flute player who can usually match any given target pitch may struggle to play a unison melodic line with the rest of the woodwind section. The vocalist who can perform a vast repertoire of songs with impeccable intonation may be less successful at distinguishing between slightly flat or slightly sharp tones if sung by someone else. The only relationships that have emerged are among students at the highest level of musical achievement. The first-chair trumpet player in the school's top band and the soprano section leader in the school's top choir would be the most likely individuals to demonstrate strong skills in all areas. But even in these cases, results suggest that such assumptions cannot be made with complete confidence.

### *Pitch Accuracy Is Affected by the Presence of a Musical Context*

One key to effectively addressing pitch issues is to correctly match the teaching strategy with the desired outcome. Take, for example, a trombone section experiencing difficulty with a unison countermelody. One strategy might be to ask the section to tune each individual pitch in isolation, removing the variables of tempo, rhythm, dynamics, articulation, and pitch function. Another strategy might be to ask them to identify the location, source, and direction of the pitch discrepancies.

Unfortunately, these strategies may only succeed in developing the students' skills at pitch matching or discrimination. Morrison (2000), after observing that instrumentalists were more accurate performing an isolated B♭—a traditional band tuning pitch—than performing a G as part of a simple melody, suggested that students become very accurate at the specific skills that they are directed to practice. In this case, students probably had considerable experience matching isolated pitches in general and the tuning B♭ in particular. Would results have been different if students tuned to other pitches during rehearsal or if the tuning B♭ was directly transferred into a short melody?

A better strategy would be to simplify only a minimum of performance parameters, thereby retaining as much of the surrounding musical environment as possible. Slowing the tempo, eliminating dynamic contrasts, or thinning the texture by reducing the number of other parts playing might be enough simplification to allow the players to focus their attention on pitch accuracy. Identifying certain individual notes as anchors or landmarks may also serve the same function. However, consider a final, sustained unison at the conclusion of a choral work. In this case, pitch matching is exactly the skill in question. In this example, taking the pitch out of context and isolating it in time and function might, indeed, be an appropriate strategy for improving accuracy.

### *Pitch Accuracy Is Aided by Clear Models*

To successfully and accurately place a pitch within a musical context, it is first necessary to identify what that pitch should be. For the least-experienced performers the target pitch can be obscured by obstacles that, in any other situation, would be appropriate. Expecting inexperienced singers to echo back a melodic passage presented with vibrato one octave higher than they are asked to respond presents a sequence of challenges—finding a pitch within the vibrato, then displacing the pitch into their own range, then matching it—that may not easily be mastered. Students have the best chance to respond successfully if the model presented is as close as possible to the response they are expected or able to make.

### *Specific Feedback Improves Pitch Skills*

The successful implementation of any effective teaching or practice strategy necessarily includes feedback. Feedback may be in the form of a comment or

gesture from the teacher, it may be the swaying needle or blinking light of an electronic tuner, or it may be performers' own assessment of themselves. In the short term, feedback can help students make adjustments that result in accurate pitch placement. On a more permanent level, continued constructive feedback over time allows students to learn the norms and standards against which their performance is judged, what accurate tuning *is*, how good intonation *sounds*. Experience may have a profound effect on students' pitch sensitivity not only because they have spent so much time mastering their instrument but also because they have gained an understanding of the musical goals toward which they are working.

### *The Most Accurate Pitch Is Not Always the Most Preferred Pitch*

Though it would seem that pitch *accuracy* is the overall goal addressed in this chapter, it might be better to say that pitch *control* is the ultimate goal toward which musicians should strive. Performance practices of the most highly trained musicians and listening preferences of the most discriminating audiences demonstrate that pitch discrepancies are an accepted and even desirable component of the best musical performances. In short, truth is subjective when it comes to pitch.

A teacher's or performer's overreliance on an external standard such as an electronic tuner or its corresponding theoretical tuning system may limit the type of pitch flexibility characteristic of expert performances. The conductor who subjects every pitch in a melodic passage to electronic analysis may be choosing the wrong analysis tool, since the assessment of good intonation appears to be more a product of listener evaluation than mathematical precision. That is not to say that attention to pitch accuracy should be minimized. It may be more desirable that musicians have the *ability* to match a target pitch exactly (with a 0 cent deviation from a presented model) than that they *perform* it that way every time. In a real musical context, intonation appears to be more negotiation than conformity.

Intonation cannot be addressed by rehearsal strategies alone. Along with ensemble skills, performance experience helps students gain knowledge of their own instrument, be it a trombone, a viola, or a voice. Each instrument, other than those with fixed pitches, has its own tendencies and eccentricities. As students progress, success at pitch accuracy will be influenced by knowledge of such details. Nevertheless, execution of an in-tune performance—regardless of which definition of the term we are using—ultimately depends on the ability of performers to assess what is happening around them and to respond accordingly.

### *References*

Barbour, J. M. (1951). *Tuning and temperament: A historical survey.* East Lansing: Michigan State College Press.

Brittin, R. V. (1993). Effects of upper- and lower-register accompaniment on intonation. *Journal of Band Research, 29*(1), 43–50.

Brown, J. C., & Vaughn, K. V. (1996). Pitch center of stringed instrument vibrato tones. *Journal of the Acoustical Society of America, 100*(3), 1728–1735.

Brown, S. F. (1991). Determination of location of pitches within a musical vibrato. *Bulletin of the Council for Research in Music Education, 108*, 15–30.

Burns, E. M. (1999). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 215–264). San Diego, CA: Academic Press.

Burns, E. M., & Campbell, S. L. (1994). Frequency and frequency ratio resolution by possessors of relative and absolute pitch: Examples of categorical perception? *Journal of the Acoustical Society of America, 96*(5), 2704–2719.

Burns, E. M., & Ward, W. D. (1978). Categorical perception—phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America, 63*(2), 456–468.

Cassidy, J. W. (1989). The effect of instrument type, stimulus timbre, and stimulus octave placement on tuning accuracy. *Missouri Journal of Research in Music Education, 26*, 7–23.

Corso, J. F. (1954). Unison tuning of musical instruments. *Journal of the Acoustical Society of America, 26*(5), 746–750.

Demorest, S. M. (2000, March). *Pitch-matching performance of junior high boys: A comparison of perception and production.* Paper presented at the MENC National Biennial In-Service Conference, Washington, DC.

Duke, R. A. (1985). Wind instrumentalists' intonational performance of selected musical intervals. *Journal of Research in Music Education, 33*(2), 101–112.

Ely, M. C. (1992). Effects of timbre on college woodwind players' intonation performance and perception. *Journal of Research in Music Education, 40*(2), 158–167.

Fastl, H., & Hesse, A. (1984). Frequency discrimination for pure tones at short durations. *Acustica, 56*, 41–47.

Fyk, J. (1982). Perception of mistuned intervals in melodic context. *Psychology of Music*, Special Edition, 36–41.

Fyk, J. (1985). Vocal pitch-matching ability in children as a function of sound duration. *Bulletin of the Council for Research in Music Education, 85*, 76–89.

Fyk, J. (1987). Duration of tones required for satisfactory precision of pitch matching. *Bulletin of the Council for Research in Music Education, 91*, 38–44.

Fyk, J. (1995). *Melodic intonation, psychoacoustics, and the violin*. Zielona Góra, Poland: Organon.

Garbuzov, N. A. (1948). Zonal nature of the pitch perception. In *Problems of physiological acoustics*, Vol. 1 (pp. 138–152). Moscow: USSR Academy of Sciences.

Geringer, J. M. (1978). Intonational performance and perception of ascending scales. *Journal of Research in Music Education, 26*(1), 32–40.

Geringer, J. M. (1983). The relationship of pitch-matching and pitch-discrimination abilities of preschool and fourth-grade students. *Journal of Research in Music Education, 31*(2), 93–100.

Geringer, J. M., & Madsen, C. K. (1987). Programmatic research in music: Perception and performance of intonation. In C. K. Madsen and C. A. Prickett (Eds.), *Applications of research in music behavior* (pp. 244–253). Tuscaloosa: University of Alabama Press.

Geringer, J. M., & Witt, A. C. (1985). An investigation of tuning performance and perception of string instrumentalists. *Bulletin of the Council for Research in Music Education, 85*, 90–101.

Geringer, J. M., & Worthy, M. D. (1999). Effects of tone-quality changes on intonation and tone-quality ratings of high school and college instrumentalists. *Journal of Research in Music Education, 47*(2), 135–149.

Kantorski, V. J. (1986). String instrument intonation in upper and lower registers: The effects of accompaniment. *Journal of Research in Music Education, 34*(3), 200–210.

Karrick, B. (1998). An examination of the intonation tendencies of wind instrumentalists based on their performances of selected harmonic intervals. *Journal of Research in Music Education, 46*(1), 112–127.

Kelly, N. E. (2000, August). *Detecting intonation errors in familiar melodies.* Paper presented at the Sixth International Conference on Music Perception and Cognition, Keele University, UK.

Loosen, F. (1993). Intonation of solo violin performance with reference to equally tempered, Pythagorean, and just intonation. *Journal of the Acoustical Society of America, 93*(1), 525–539.

Madsen, C. K. (1966). The effect of scale directon on pitch acuity in solo vocal performance. *Journal of Research in Music Education, 14*(4), 266–275.

Madsen, C. K., Edmonson, F. A., & Madsen, C. H. (1969). Modulated frequency discrimination in relationship to age and musical training. *Journal of the Acoustical Society of America, 46* (6, pt. 2), 1468–1472.

Madsen, C. K., & Geringer, J. M. (1976). Preferences for trumpet tone quality versus intonation. *Bulletin of the Council for Research in Music Education, 46*, 13–22.

Madsen, C. K., & Geringer, J. M. (1981). Discrimination between tone quality and intonation in unaccompanied flute/oboe duets. *Journal of Research in Music Education, 29*(4), 305–313.

Morrison, S. J. (2000). Effect of melodic context, tuning behaviors, and experience on the intonation accuracy of wind players. *Journal of Research in Music Education, 48*(1), 39–51.

Platt, J. R., & Racine, R. J. (1985). Effect of frequency, timbre, experience, and feedback on musical tuning skills. *Perception and Psychophysics, 38*(6), 543–553.

Rakowski, A. (1978). *Categorical perception of the pitch in music.* Warsaw: State Higher School of Music.

Rakowski, A. (1990). Intonation variants of musical intervals in isolation and in musical contexts. *Psychology of Music, 18*(1), 60–72.

Sergeant, D. (1973). Measurement of pitch discrimination. *Journal of Research in Music Education, 21*(1), 3–19.

Shonle, J. I., & Horan, K. E. (1980). The pitch of vibrato tones. *Journal of the Acoustical Society of America, 67*(1), 246–252.

Siegel, J. A., & Siegel, W. (1977). Categorical perception of tonal intervals: Musicians can't tell sharp from flat. *Perception and Psychophysics, 21*(5), 399–407.

Sogin, D. W. (1989). An analysis of string instrumentalists' performed intonational adjustments within ascending and descending pitch set. *Journal of Research in Music Education, 37*(2), 104–111.

Spiegel, M. F., & Watson, C. S. (1984). Performance on frequency-discrimination tasks by musicians and nonmusicians. *Journal of the Acoustical Society of America, 76*, 1690–1695.

Sundberg, J. (1982). In tune or not? A study of fundamental frequency in music practice. In C. Dahlhaus and M. Krause (Eds.), *Tiefenstruktur der Musik* (pp. 69–97). Berlin: Technical University of Berlin.

Tieplov, B. M. (1947). *Psychology of musical aptitude.* Moscow: Academy of Pedagogical Sciences of the RSFSR.

Vos, J. (1982). The perception of pure and mistuned musical fifths and major thirds: Thresholds for discrimination, beats, and identification. *Perception and Psychophysics, 32*(4), 297–313.

Wapnick, J., & Freeman, P. (1980). Effects of dark-bright timbral variation on the perception of flatness and sharpness. *Journal of Research in Music Education, 28*(3), 176–184.

Wapnick, J., Bourassa, G., & Sampson, J. (1982). The perception of tonal intervals in isolation and in melodic context. *Psychomusicology, 2*(1), 21–37.

Watson, C. S., Kelly, W. J, & Wroton, M. W. (1976). Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty. *Journal of the Acoustical Society of America, 60*, 1176–1186.

Yarbrough, C., & Ballard, D. (1990). The effect of accidentals, scale degrees, direction, and performer opinions on intonation. *Update: Applications of Research in Music Education, 8(2)*, 19–22.

Yarbrough, C., Bowers, J., & Benson, W. (1992). The effect of vibrato on the pitch-matching accuracy of certain and uncertain singers. *Journal of Research in Music Education, 40*(1), 30–38.

Yarbrough, C., Karrick, B., & Morrison, S. J. (1995). Effect of knowledge of directional mistunings on the tuning accuracy of beginning and intermediate wind players. *Journal of Research in Music Education, 43*(3), 232–241.

Yarbrough, C., Morrison, S. J., & Karrick, B. (1997). The effect of experience, private instruction, and knowledge of mistunings on the tuning performance and perception of high school wind players. *Bulletin of the Council for Research in Music Education, 134*, 31–42.

Yarbrough, C., Morrison, S. J., Karrick, B., & Dunn, D. E. (1995). The effect of male falsetto on the pitch-matching accuracy of uncertain boy singers, grades K–8. *Update: Applications of Research in Music Education, 14*(1), 4–10.

Yarbrough, C. M., Green, G. A., Benson, W., & Bowers, J. (1991). Inaccurate singers: An exploratory study of variables affecting pitch-matching. *Bulletin of the Council for Research in Music Education, 107*, 23–34.

# 13

# Structural Communication

ANDERS FRIBERG & GIOVANNI UMBERTO BATTEL

The communication of structure in musical expression has been studied scientifically by analyzing variations in timing and dynamics in expert performances. The underlying principles have been extracted and models of the relationship between expression and musical structure formulated. For example, a musical phrase tends to speed up and get louder at the start and to slow down and get quieter at the end; mathematical models of these variations can enhance the quality of synthesized performances. We overview the dependence of timing and dynamics on tempo, phrasing, harmonic and melodic tension, repetitive patterns and grooves, articulation, accents, and ensemble timing. Principles of structural communication (expression) can be taught analytically by explaining the underlying principles and techniques with computer-generated demonstrations, or in traditional classroom or lesson settings by live demonstration.

Variations in timing and dynamics play an essential role in music performance. This is easily shown by having a computer perform a classical piece exactly as written in the score. The result is dull and will probably not affect us in any positive manner, although there may be plenty of potentially beautiful passages in the score. A musician can, by changing the performance of a piece, totally change its emotional character, for example, from sad to happy (chapter 14). How is this possible, and what are the basic techniques used to accomplish such a change? The key is how the musical structure is communicated. Therefore, a good understanding of structure—whether theoretic or intuitive—is a prerequisite for a convincing musical performance.

This chapter surveys the basic principles and techniques that musicians use to convey and project music structure (see also overviews in Gabrielsson, 1999; Palmer, 1997). We will only consider auditory communication and leave out

visual cues in concert performances (see chapter 15). Another issue only briefly covered is perception—the extent to which subtle variations in performance are perceived by the listener. Our focus will be on topics that have been the subject of systematic research and that can be useful for music students and teachers. Largely for practical reasons, this research has tended to focus more on traditional classical music than other styles, timing more than other parameters such as dynamics and articulation, and piano performance more than that of other instruments. Nonetheless, most of the results obtained, and hence most of the principles presented here, seem to have a universal character and may be applied to a wide variety of genres and instruments.

### *Concepts and Terms*

How can music performance be studied scientifically? We base our analysis primarily on information available in the sound alone. All cues for musical communication are contained in sound, which in turn can be described and quantified in terms of physical variables such as the duration and sound level of each tone.

*Tone duration* is the time interval between the physical start of the tone (*onset*) and the end of the same tone (*offset*), that is, the sounding duration of a tone. More important for timing in music is the *interonset interval* (*IOI*), defined as the time interval between the onset of the tone and the onset of the immediately following tone. In other words, IOI is the sum of a tone's physical duration and the pause duration between the offset of the tone and the onset of the next (see Figure 13.1). IOI is easy to measure in MIDI recordings and can also be estimated from audio recordings using computer software.

To find out if a tone is lengthened or shortened in performance, measured IOI values are compared with the note values given in the score. A *nominal*

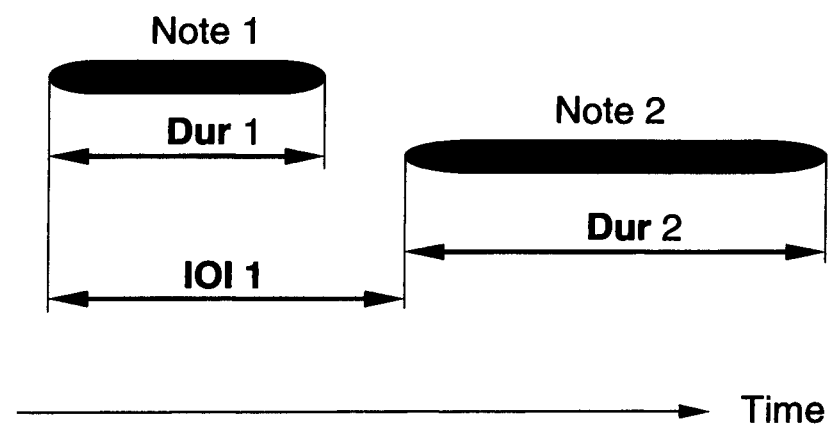



Figure 13.1. Definition of interonset interval (IOI) and duration (Dur) for two successive tones.

*performance* (or deadpan performance) is defined as a direct translation of the score into physical variables, where all notes of the same note value have the same *nominal IOI*, derived from the *global tempo* (tempo marking or mean tempo). In a nominal performance, for example, an eighth note is always exactly half as long as a quarter note. A nominal performance often serves as a reference point for research on musical timing. Timing variations relative to a nominal performance can be analyzed either tone by tone or as changes in *local tempo*, conceived as a continuously varying function of time (*tempo curve*).

Conventional Western notation developed historically within the physical and cognitive constraints of performance and sight-reading. It deliberately fails to describe music in too much detail, since that would make it too difficult to read. This means that it is not possible to define nominal values for articulation and dynamics in conventional scores. Another consequence is the development of *performance conventions*—standard (but generally style-specific) interpretations of notational symbols that are not evident in the score itself. For example, strings of eighth notes in a jazz score are typically performed unevenly in long-short patterns (*swing*).

What are the basic building blocks of musical structure? Most tonal music has a hierarchical phrase structure, sometimes simply called *grouping*. The slowest level is the entire piece, which is then divided and subdivided into sections, phrases, subphrases, and melodic groups. Superimposed upon this is usually a *metrical hierarchy*: the beat or *tactus* (corresponding to when you tap your feet) is grouped, usually in groups of two or three, into measures and groups of measures. The beat can also be divided into subbeats. Phrasing and meter are theoretically independent, although phrase and metric boundaries often coincide, reinforcing the overall perceived grouping (Lerdahl & Jackendoff, 1983) (see Figure 13.2).

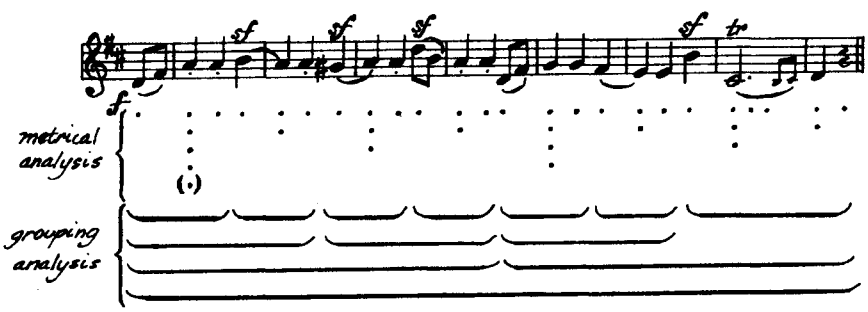



Figure 13.2. Phrase structure and metrical structure in an excerpt from Haydn's Symphony no. 104. The dots represent the hierarchical metrical structure. The top level in the figure is the beat level, the second the measure level, and the third and fourth are hypermetric levels. The hierarchical phrase structure is shown with brackets. The top level in the figure is the fastest, with only a few tones in each phrase or group. The slowest level corresponds to the whole phrase. (From Lerdahl & Jackendoff, 1983. Copyright 1983 MIT Press, used by permission.)

### *Commonalities Among Performers and Between Repetitions*

What are the similarities and differences in separate performances of the same piece? When discussing interpretation, the emphasis is often placed on performance differences. However, even musicians who are considered to interpret music quite differently may produce remarkably similar patterns of timing and dynamics (Repp, 1992).

Musical structure is reflected in physical variables in a number of ways. Figure 13.3 illustrates the variation of IOI and dynamics in two performances of the same piece. We see that both pianists express the phrase structure by lengthening and softening the tones (*ritardando* and *diminuendo*) at the end of each melodic gesture, in measures 4 and 8. The slowing and softening are more pronounced at the end of the whole phrase and are quite substantial and thus clearly perceptible. The differences in interpretation between the pianists are largely seen in variations within phrases and on a note-to-note level.

It is often argued that a repeated passage should be performed differently in both cases. This is, however, not generally confirmed in measurements. On the contrary, there are often striking similarities between the first and second presentation of a thematic group. This is also true for the repetition of a whole piece on different occasions. In Figure 13.3, the differences between the two performances are very small and, at least for timing, are below the perceptual limits in most cases.

## Basic Principles and Techniques

### *Tempo*

Global tempo can vary substantially in different renditions of the same piece by different performers. In commercial recordings of both Schumann's "Träumerei" and Chopin's Etude in E Major, Op. 10, No. 3, the fastest tempo was about twice as fast as the slowest (Repp, 1992, 1998b). Clearly tempo is influenced not only by tempo indications but also by the performer's interpretation, in particular by the intended motional and emotional character (chapter 14).

Collier and Collier (1994) investigated tempo in jazz by analyzing a large number of commercial jazz recordings. They found that *double time* (doubling of tempo, sometimes introduced into an improvisation to produce a contrasting passage or to increase the intensity) corresponded to a tempo ratio of 2.68:1—well above a mathematical doubling of tempo.

When the global tempo of a performance is changed, patterns of local timing variations may also change. For example, there may be a tendency toward more expressive timing variation (relative to tempo) at slower tempi (Repp, 1995). Also, the perceptual or motor limits of tone duration may alter the expressive pattern.

### *Phrasing*

In music from the Romantic period, large variations in local tempo are an essential part of the performance tradition. Phrases often start slow, speed up in the

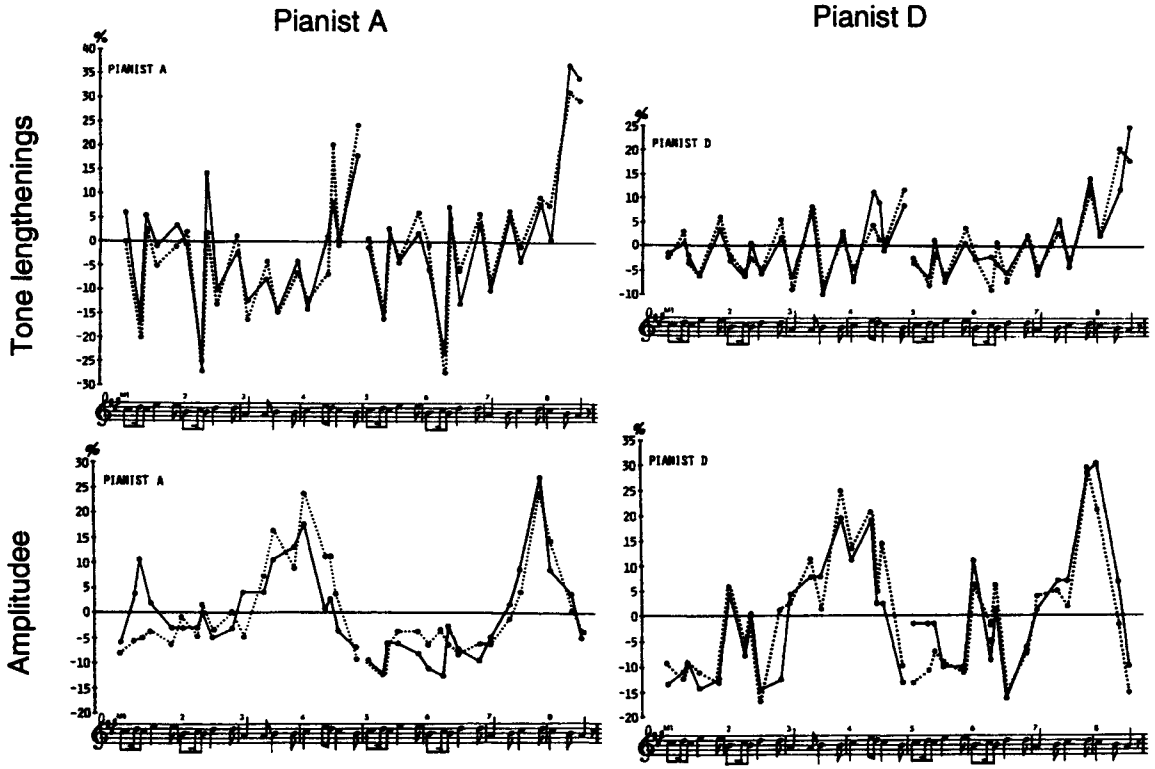



Figure 13.3. The first eight measures of Mozart's Piano Sonata K. 331, as performed by two pianists. The upper graphs illustrate the timing, with the vertical axis showing the deviation in IOI of each tone relative to nominal duration. The lower graphs illustrate the dynamic variation, with the vertical axis showing the peak amplitude of each tone relative to the mean of all tones in each performance. The first performance is represented by full lines and the repetition by dashed lines. (Adapted from Gabrielsson, 1987. Copyright 1987 Royal Swedish Academy of Music, used by permission.)

middle, and slow down again toward the last tone (e.g., Henderson, 1936; Repp, 1992). Dynamic variations tend to follow a similar pattern: soft in the beginning, loud in the middle, and softer toward the end of the phrase (see Figure 13.3) (Gabrielsson, 1987).

These typical shapes of timing and dynamics are observed in a majority of performances of Romantic music and are important for conveying the basic phrase structure to the listener. The *ritardando* at the end can communicate the phrase level, with typically a more pronounced *ritardando* at the end of a musical unit of longer duration or at a slower hierarchical level. This is clearly the case in Figure 13.3, where pianists A and D both lengthen the tones more at the end of the example than in the middle. The dynamic variation follows a similar pattern. In this way, not only the phrase boundaries but also their hierarchical level—and hence the hierarchical phrase structure of the whole piece—can be communicated, just by changing tempo and dynamics. Similar principles are found in speech, where lengthening is used to communicate phrase and sentence boundaries.

The exact amount and shape of the variation over the phrases is an important issue for the performer to decide. In Figure 13.3, both pianists slow down at the phrase boundaries but follow different tempo curves. For example, pianist A lengthens the final tones in each subphrase more than pianist D. Different shapings of local tempo or dynamics can entirely change the character of a performance, signaling different expressive intentions (Battel & Fimbianti, 1998).

Several models of these typical tempo variations have been developed. The first computational model was presented by Todd (1985, 1989; see also Windsor & Clarke, 1997; Penel & Drake, 1998) and accounted for the variation of measure duration over phrases. Todd (1992, 1995) later developed a revised model based on a different mathematical function, which he argued was more closely related to a metaphor of physical motion. Friberg (1995) modified Todd's first model so that it may be applied at the note level, introducing several extra parameters to account for individual variations. In Figure 13.4, this *phrase arch* model is fitted to three different piano performances of Schumann's "Träumerei" measured by Repp (1992). As can be seen in the figure, the model catches most of the individual variation regarding phrasing but misses local variations on the note level.

Phrasing in Baroque music typically involves smaller variations in local tempo than in Romantic music. Baroque music tends to have a more motoric, metrical character (as does most contemporary jazz and pop), suggesting the metaphor of a mass moving at a constant speed, creating a kind of musical momentum. This natural coupling of motion and music has been investigated in an intuitive way in a number of publications (overview in Shove & Repp, 1995). Looking for direct couplings between the physical world and music, Friberg and Sundberg (1999) discovered a close connection between how a runner stops and how Baroque music stops at a final *ritardando*. They found that the average velocity of the runner and the average local tempo closely followed the same curve: their model could account for the deceleration both of individual runners and of individual music performances. Moreover, participants in listening experiments

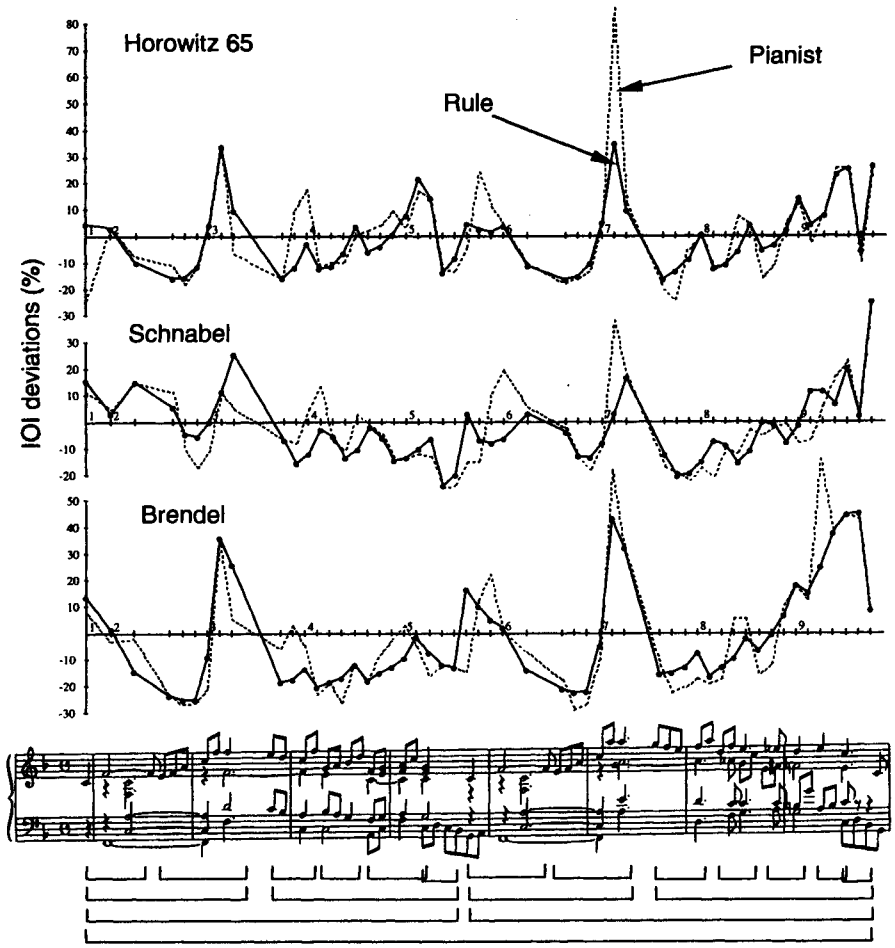



Figure 13.4. The dotted lines show the deviations in IOI from nominal values for performances of Schumann's "Träumerei" by three different pianists. The solid lines show predicted IOI deviations according to the phrase arch model. The predictions have been fitted to the three performances by adjusting the parameters of the model. The brackets below indicate the grouping analysis used in the model.

preferred final *ritardandi* that corresponded to the stopping runners, suggesting, in this case, a close coupling to physical motion.

The fastest level in the phrase hierarchy consists of small melodic units of a few notes each. Grouping (i.e., segmentation) at this level tends to be quite ambiguous, often with several possible interpretations. So communication of this structure can be subject to more individual interpretation than, say, communication of longer phrases. One example is found in Mozart's A-major sonata (Figure 13.3), where many performers chose the first five notes as a group while others

chose the first four notes, giving the second group an upbeat. This ambiguity may be due to contradictory perceptual cues from different aspects of the musical structure, such as the melodic contour or the meter, and can be resolved in performance by inserting a micropause between the last tone of one phrase and the first of the next, which both interrupts the sound and delays the onset of the following tone (Friberg, Sundberg, & Frydén, 1987; Friberg, Bresin, Frydén, & Sundberg, 1998; Clarke, 1988).

A deceptively simple performance principle is *the higher, the louder* (Sundberg, Friberg, & Frydén, 1991). The origin of this principle appears to be physical: wind instruments (including the voice) tend to produce louder tones at higher pitches, even though effort or input pressure is held constant. Often the most important tone in a phrase is also the highest in pitch. In this case, the high-loud principle produces natural-sounding phrasing (cf. Windsor & Clarke, 1997; Palmer, 1996a; Krumhansl, 1996).

### *Harmonic and Melodic Tension*

The notion of musical tension is common to both music theory and music psychology. Many authors claim that the contrast between tension and release is a major source of musical interest. Tension is coupled to expectancy; an unexpected tone or chord creates tension (Krumhansl, 1990). A number of sources that contribute to harmonic and melodic tension have been identified. A partial list is given in Table 13.1 (cf. Bigand, Parncutt, & Lerdahl, 1996).

Several models have been developed to address tension. For case 2(i), Sundberg, Friberg, and Frydén (1991) defined the *harmonic charge* of a chord as a weighted

Table 13.1. A list of tonal relationships that contribute to melodic and harmonic tension.

| Case | Relationship | Increased Tension For: |
|---|---|---|
| 1 | between keys | (i) a modulation to a key that is distant on the circle of fifths<br>(ii) a modulation to a scale with few tones in common with the original scale |
| 2 | between chords; | (i) chords more distant from the key on the circle of fifths chord relative to key (comparing roots)<br>(ii) chromatic chords, or chords that include tones foreign to the prevailing scale<br>(iii) successive chords that have few tones in common |
| 3 | tone relative to chord | (i) tones that are more distant on the circle of fifths from the root of the chord<br>(ii) tones foreign to the diatonic scale associated with the chord |
| 4 | simultaneous tones in a chord | chords that contain more dissonant intervals |
| 5 | melodic contour | unexpected melodic turns |

sum (root most important) of the distance between the tones of the chord and the prevailing tonic on the circle of fifths. They then allowed local tempo to slow down and sound level to increase in areas of high harmonic charge (Friberg, 1991), creating a kind of harmonic phrasing (see Figure 13.5).

For case 3(i), Sundberg et al. (1991) defined *melodic charge* as an increasing function of the distance on the circle of fifths. The performance model increases the sound level, IOI, and—if applicable—the extent of vibrato in proportion to the calculated melodic charge (Friberg, 1991).

A complex model that takes into account most of the aspects in Table 13.1 was formulated by Lerdahl (1996). This model explained the majority of variation in subjects' ratings of perceived tension in a Mozart sonata (Krumhansl, 1996).

The most common way to communicate tension seems to be to emphasize notes or areas of relatively high tension, as in the models of harmonic and melodic charge described earlier. However, it is difficult to trace the origins of variations of timing and dynamics measured in real performances, since the various tension concepts often are coupled to each other and to the phrasing structure. For example, chords more distant from the key are more often found in the middle of phrases, while chords close to the key are more often found in the beginning or in the end of the phrase. Also, phrasing tends to dominate performance expression, which makes it hard to isolate the more subtle details such as the expression of melodic or harmonic tension.

Analyzing a performance of a Mozart sonata, Palmer (1996a) found a weak correspondence between predictions of Lerdahl's model and observed timing (IOI): positions of high tension were emphasized by lengthening. An additional finding was that relatively dissonant chords were often performed by delaying the melody by as much as 100 ms. The purpose of this performance strategy could

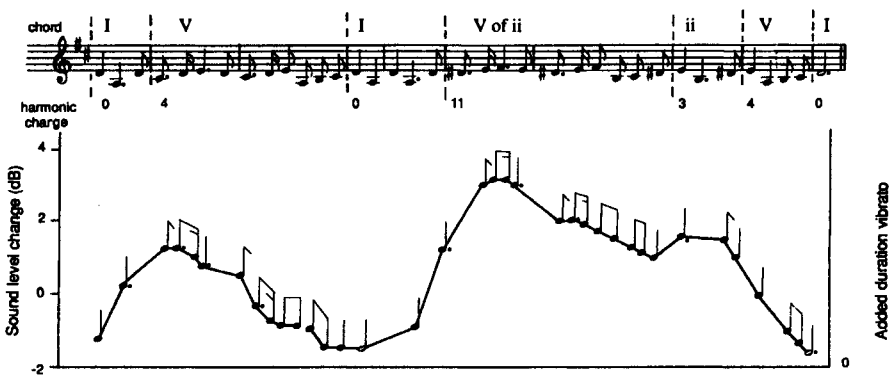



Figure 13.5. The harmonic charge model applied to a theme from Schubert's "unfinished" symphony. Sound level and vibrato rate increases and tempo decreases with increasing harmonic distance from the tonic. Note the peak in harmonic charge at the most distant chord (V of II) on the circle of fifths.

either be to reduce the perceptual dissonance of the chord (as Palmer suggested) or to emphasize the melody tone by delaying it (see "Accents" later).

### *Repetitive Metrical Patterns and Grooves*

Music that is rhythmically regular often exhibits consistent patterns of timing and dynamics within metrical units such as the measure. For example, if the first beat in each measure is accentuated, a dynamic pattern is formed that is repeated in each measure. This kind of patterning is often associated with dance, suggesting that these patterns serve to characterize the motional character of the piece.

**Patterns in Triple Meter.** In performances of both a Beethoven minuet (Repp, 1990) and a Chopin nocturne (Henderson, 1936), the pianists played the second beat late and the third beat early, forming a long–short–long pattern.

In patterns of a half note followed by a quarter note in $\frac{3}{4}$ time, or a quarter note followed by an eighth note in $\frac{6}{8}$, the ratio of the IOIs of the long and short tones is usually in the range from 1.7:1 to 1.9:1—consistently smaller than the nominal 2:1. This has been found in performances of Swedish folk music as well as Mozart and Chopin (Gabrielsson, 1987; Gabrielsson, Bengtsson, & Gabrielsson, 1983; Henderson, 1936), (see Figure 13.3; note the characteristic zigzag patterns in the timing graphs). Sundberg et al. (1991) called this effect *double duration* and incorporated it into a performance model.

A different pattern is found in Viennese waltzes: they are usually performed with an early second beat, thus forming a short–long–intermediate timing pattern for the three beats in the measure. This pattern was first systematically investigated by Bengtsson and Gabrielsson (1983), who also observed a hypermetric timing pattern spanning two measures that reflects how these waltzes are danced, as well as the underlying harmonic structure.

**Two-note Patterns.** In jazz, Baroque, and folk music, it is common to apply a long–short pattern to consecutive eighth notes (or the shortest prevalent note value). In Baroque music, this performance convention is referred to as *notes inégales* (Hefling, 1993). Drummers playing typical comping patterns in jazz perform different ratios between consecutive eighth notes depending on global tempo. At slow to medium tempi, the ratio tends to be about 3:1 (dotted eighth + sixteenth), dropping to 1:1 (even eighth notes) at fast tempi (Friberg and Sundström, in press). At the same time, the soloist uses smaller ratios than 2:1. Surprisingly, the ratio of 2:1, implied by the often-mentioned triplet feel of swing, was not observed in these experiments.

**Duration Contrast.** Contrary to the double duration principle described earlier, the contrast between long and short tones is often *increased* by a lengthening of comparatively long notes and a corresponding shortening of comparatively short notes (Taguti, Mori, & Suga, 1994). However, duration contrast applies to *all* notes, while double duration applies only to the 2:1 pattern. In the case of a re-

peated rhythmical figure, duration contrast will appear as a repetitive timing pattern. Duration contrast can also be applied backward: depending on the performer's intention, the contrast may instead be decreased (chapter 14). A model was formulated by Sundberg et al. (1991).

### *Articulation*

Articulation—at least in the sense of *staccato* versus *legato*—may be defined mathematically as the ratio of tone duration to IOI. Articulation strongly affects motional and emotional character (De Poli, Rodà, & Vidolin, 1998; Battel & Fimbianti, 1998; chapter 14 in this volume).

The duration of *staccato* tones has been found to correspond on average to about 40% of the IOI (or note value) in typical performances of a Mozart andante movement for piano (from Sonata K. 545; Bresin & Battel, 2000). When pianists were asked to play the same piece *brillante* or *leggero*, the duration decreased to 25% (*staccatissimo* range), while in *pesante* performances the duration approached 75% of IOI (*mezzo staccato* range). The typical value of 40% is in nice agreement with C. P. E. Bach's (1753/1957) observation that *staccato* notes should be played with less than 50% of their nominal duration.

In *legato* playing on the piano, successive tones often overlap—both keys are down for a short period of time. The amount of overlap (in milliseconds) increases with IOI and pitch interval size (Repp, 1997; Bresin and Battel, 2000). Bresin and Battel also found that the overlap time was used for expressive purposes: the pianists used more overlap when instructed to play *appassionato* than *piatto* (flat). Overlap is also dependent on direction of melodic motion; descending melodic patterns are usually played with more overlap than ascending patterns (Repp, 1997; Bresin, 2000).

Bresin (2000) also found an interesting link between articulation and physical locomotion. The overlap time of the feet in walking was found to vary qualitatively in the same way as in *legato* articulation. Flying time in running (the time during which neither foot is in contact with the ground) was similarly found to be related to detached time in *staccato* articulation. Based on these measurements, Bresin (2000) was able to formulate models for *staccato*, *legato*, and tone repetition.

### *Accents*

The meaning of the term *accent* varies considerably in music-theoretic literature. Two main categories can be identified: immanent accents and performed accents (cf. Parncutt, in press; Lerdahl & Jackendoff, 1983). An *immanent accent* is evident from the structure of the score itself, meaning that even if the score is performed nominally, these positions will be perceived as accented. For example, immanent accents may occur on notes in metrically strong positions, on comparatively long notes, on the second tone of an upward leap, on the top tone of a melodic turn, or at increased harmonic tension (melodic immanent accents: Thomassen, 1982; Huron & Royal, 1996).

A *performed accent* is added by the performer relative to the nominal performance. This definition is closer to the common use of the term. Performed accents seem primarily to be used to reinforce immanent accents. This was confirmed in simple melodies (Drake & Palmer, 1993). However, in more complex (real) music, only weak couplings between performed and immanent accents have been found: only when positions interacting with the grouping structure were disregarded was it possible to get statistical significance in piano performances of Schumann's "Träumerei" (Penel & Drake, 1998). This finding may indicate that the concept of melodic accent is only a cue that leads to the forming of melodic groups. In fact, cues for immanent accents and cues for making automatic melodic grouping are similar (e.g., Friberg et al., 1998; Cambouropoulos, 1998).

Perhaps the most obvious way to perform an accent is to increase loudness. In the case of instruments where the loudness envelope can be varied (such as winds, strings, and voice), there are several possibilities, including a loudness increase at the beginning of the tone and a shorter attack time. Timing can also be used in several ways to emphasize a tone: (1) by lengthening the tone; (2) by delaying the onset, that is, lengthening the preceding tone (IOI) and possibly inserting a micropause before the accented note; and (3) by playing the accented note more *legato* (Henderson, 1936; Clarke, 1988; Drake & Palmer, 1993). In addition, a change in articulation such as one *legato* note surrounded by *staccato* notes may signal an accent (cf. Lerdahl & Jackendoff, 1983).

### *Ensemble Timing*

In polyphonic classical music, it is often important to highlight the melody. An obvious way is to play it louder. However, timing is also an effective method. If two tones are presented at almost the same time, the first will be perceptually emphasized (Rasch, 1978). This technical device, commonly referred to as *melody lead*, has been observed in string and wind trios, as well as in piano performances. The melody is typically played about 20 ms ahead of the other voices, within the range of about 7 to 50 ms (Rasch, 1979; Palmer, 1996b; Vernon, 1936).

The opposite often happens in jazz: soloists deliberately play behind the beat. At each quarter-note beat, the soloist was delayed relative to the ride cymbal by up to about 100 ms at slow tempi (Ellis, 1991; Friberg & Sundström, in press). At the same time, the soloist and drummer were synchronized at the upbeats, that is, on the eighth notes between the beats. The purpose here is probably not to highlight the soloist, which in this style is clearly audible, due to either large spectral differences or the use of microphones. Rather, this timing combination creates both the impression of the laid-back soloist often strived for in jazz and at the same time an impression of good synchronization (see Figure 13.6).

## Classification and Purpose of Performance Variations

Non-notated variations in timing and dynamics (deviations from the nominal performance) can be divided into two main types (cf. Juslin, Friberg, & Bresin,

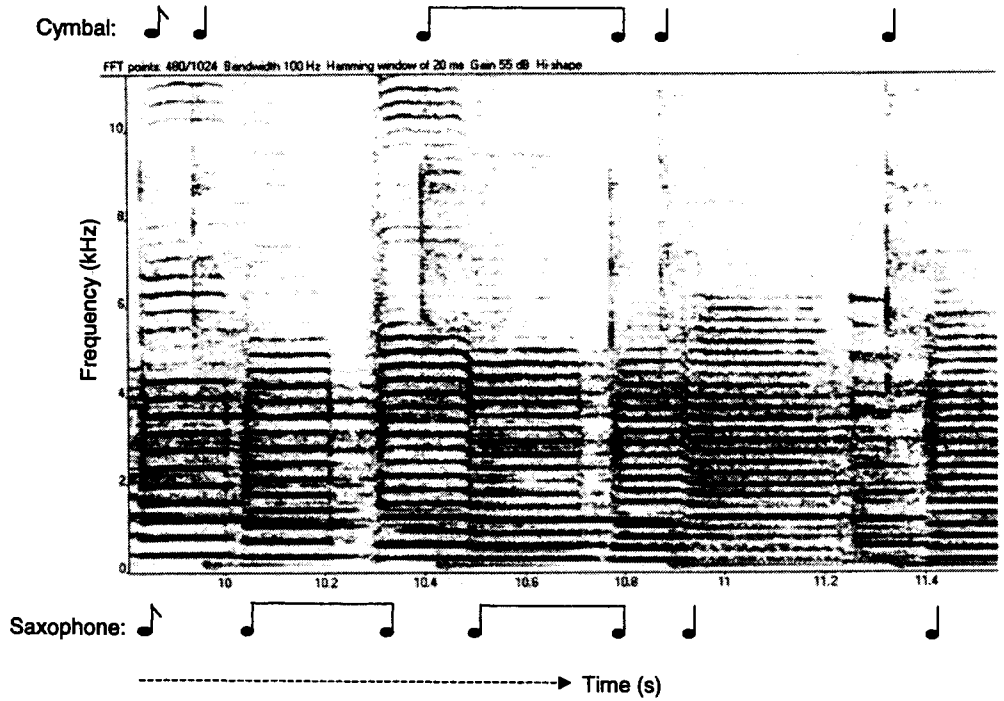



Figure 13.6. A spectrogram of a short passage of "My Funny Valentine" performed by Miles Davis Quintet, 1964, that illustrates the timing relation between the ride cymbal and the soloist in jazz. The cymbal onsets appear as vertical lines in the high-frequency part of the graph. The saxophone onsets appear as breaks, or vertical shifts, in the horizontal lines in the lower part of the graph. In this example, cymbal is being played with a swing ratio of about 4:1 and the saxophone with a swing ratio of approximately 3:2. The downbeat saxophone tones are delayed relative to the cymbal by about 100 ms, but on the upbeats the cymbal and the saxophone are synchronized.

in press). *Expressive variations* are deliberately meaningful or communicative (but not necessarily conscious). *Nonexpressive variations* are of two kinds: *variations due to technical limitations* of the instrument and the performer and *random variations* (including imperfections in the perceptual timing and motor system). Neither of these last two have a deliberate communicative function; they may tell us something about the performer's abilities, but this is of course unintentional. Nevertheless, they can be important for the naturalness of a performance.

Expressive variations can be classified according to their apparent communicative purpose. They may either communicate the *structure* of the music or express its *character*. We use the term *character* to refer to either the *emotional* (happy, sad, etc.) or *motional* (urgent, calm, swingy, etc.) implications of the music. Emotional character is dealt with in more detail in chapter 14.

Sundberg (2000) identified two main underlying principles for communicating musical structure. The first involves the *differentiation* of pitch and duration. Categorical perception is improved by increasing the difference between categories, such as stretching the frequencies of scale tones or playing short notes even shorter (e.g., duration contrast, high loud). The second principle involves the *grouping* of notes in phrases, metrical units, or harmonic areas. For example, phrases are often performed with a *diminuendo* at the end. This increases efficiency of the musical communication by introducing redundancy; the phrase boundaries are often recognized even without this cue.

Structure and character are not necessarily independent. Character can be characterized in terms of *how* the structure is communicated. In fact, most of the cues described in chapter 14 for communicating different emotions can be realized by the techniques for structural communication described in this chapter (Bresin & Friberg, 2000; Battel & Fimbianti, 1998).

## Applications in Music Pedagogy

As outlined in this chapter, music psychology research has over the past years accumulated a substantial body of knowledge about music performance that could be incorporated into music teaching. Knowledge of the basic principles of structural communication is possibly more important in earlier stages of instrumental learning; once these principles are mastered, both in theory and in performance, it may be easier to develop an individual voice and allow the more interesting development of an artist's individual personality to come to the fore (see Figure 13.7) (Andreas C. Lehmann, personal communication).

### *Teaching Theory of Structural Communication*

Each of the basic principles and techniques described earlier can be explained and discussed in terms of generality and individual differences. Consider phrasing as an example. A teacher might explain how performers introduce small *ritardandi* or *diminuendi* at the end of phrases and increase their extent at the end of sections. This archetypal phrasing is not so easy to hear from just listen-

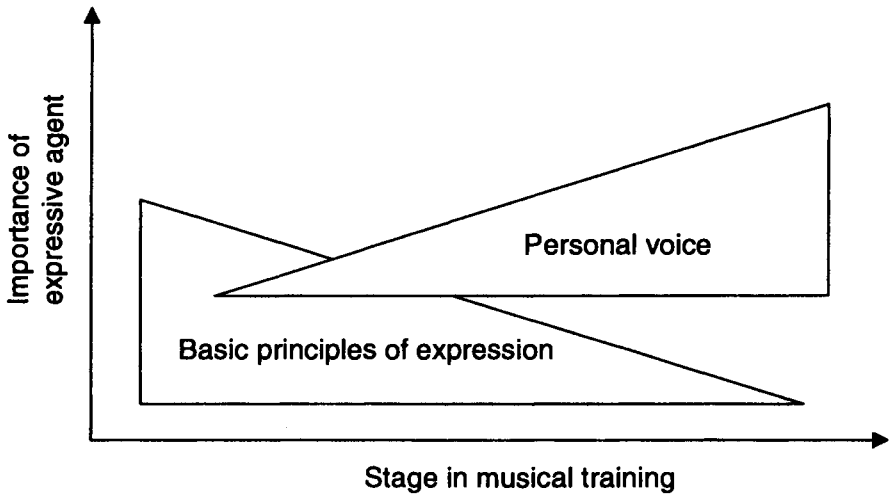



Figure 13.7. Relative importance, at different stages during performance studies, of the basic principles of communicating musical structure and character and the musician's personal voice or individual approach to musical interpretation (adapted from Andreas C. Lehmann, personal communication).

ing to recordings and can come as a surprise, even to musicians. It is also useful to show measurements of performances by well-known musicians such as in Figure 13.3 and 13.4 so that students understand that the same principles are also used in top-level performances. Other aspects of structural communication, such as articulation and timing patterns, can be taught in a similar way.

The use of sound examples generated by computer-based performance models (e.g., www.speech.kth.se/music/performance/) allows each aspect (e.g., phrasing) to be studied and listened to in isolation. Graphs that show the variation of dynamics and timing also help the student to hear and understand what is happening. A computer program such as *Director Musices* (Friberg et al., 2000) allows a teacher to directly apply phrasing rules to any music example, with the flexibility of changing the extent and parameters—all in real time and all in the classroom. Students can then judge for themselves how much variation of timing and dynamics is appropriate for the interpretation of a given phrase. Regardless of their musical training or ability to read notation, students can listen to differences between different renditions of a given passage and learn to focus their auditory attention on one aspect. Thus the suggested pedagogical approach may be useful not only for teaching performance but also for teaching relevant auditory skills.

### *Performance Studies*

Once the theoretical concepts have been explained, the various performance principles could be practiced separately. If musical tension is selected, a stu-

dent or teacher might first make an analysis of tension–relaxation patterns in a piece and think about different ways in which tension could be appropriately communicated or expressed in a particular performance on a given instrument. Then the student could practice emphasizing the tension–relaxation patterns in different ways (e.g., timing, dynamics). For example, one could, as an exercise, emphasize harmonic tension but otherwise play without expression. This can extend the student's expressive vocabulary and facilitate the creation of different emotions or patterns of implied movement (cf. chapter 14).

The importance of feedback for efficient learning is well known. Using computer analysis tools, the variations of timing and dynamics in students' performances can be measured and shown graphically. This could help students to evaluate, for example, how well their variations of dynamics and timing communicate intended patterns of tension and relaxation. Commercial recordings can also be analyzed in this way, but the methods of analysis and the required software tools depend on which instrument and structural aspect is being studied.

A majority of the principles of structural communication can be studied with just a microphone connected to a computer using commercial or free software. Ensemble timing, such as observing the amount of lead or lag for each performer, can be illustrated in a spectrogram program (e.g., Soundswell, www.hitech.se; Wavesurfer, www.speech.kth.se/wavesurfer). An example was given in Figure 13.6. Illustrations of dynamics and articulation can be obtained by computing a smoothed RMS value of the audio signal, a feature found in several audio wave editors.

Studying interonset timing and local tempo variations is currently a little bit more complicated. Here it is necessary to compute for each note the deviations of IOI relative to the nominal value. The IOIs and durations can be obtained from a MIDI-equipped instrument (e.g., a Disklavier piano from Yamaha, Zeta string instruments, an electric guitar MIDI controller, or an acoustic instrument with an audio-to-MIDI converter) connected to a computer. A program such as POCO (Desain, Honing, & Heijink, 1997; Honing, 1990; stephanus2.socsci.kun.nl/mmm/) can match the values to the score so that deviations are obtained.

Pianists may find an instrument such as the Yamaha Disklavier useful for the study of structural communication. This is essentially an acoustic piano in which the timing and dynamics of each key pressure can be registered on a computer. The computer can also control the instrument so that a recorded performance can be listened to directly on the instrument. The Disklavier has been used regularly by the second author of this chapter in teaching advanced courses in piano performance (Battel et. al., 1998). One such course consisted of two parts: (1) music performance analysis, and (2) analytical methods for performance. The participants played the first 10 bars of Mozart's Sonata K. 333 in B♭ Major on the Disklavier. Variations in dynamics and timing produced by the players were displayed on a computer screen. The data were compared to those of the second author's performance recorded earlier on a Disklavier and to those produced by a performance rule system implemented in MELODIA (Bresin, 1993). Finally, the musical meaning of performance variables in the historical period, style, and performance tradition of the composer was analyzed and discussed.

In summary, recent research on music performance has opened up a range of new possibilities for teaching expression in the sense of structural communication. The development of new software tools for this purpose would further facilitate the use of the computer as a complement in the classroom.

### *References*


Bach, C. P. E. (1753/1957), edited by L. Hoffman-Erbrecht. *Versuch über die wahre Art, das Clavier zu spielen.* Leipzig: Breitkopf & Härtel.

Battel, G. U., Canazza S., Fimbianti, R., & Rodà, A. (1998). III. Modalités de l'expérience. In Gruppo Analisi e Teoria Musicale: Quand une analyse schenkerienne est-elle utile aux interprètes? *Revue Belge de Musicologie 52,* 331–334.

Battel, G. U., & Fimbianti, R. (1998). How to communicate expressive intentions in piano performance. In A. Argentini and C. Mirolo (Eds.), *Proceedings of the XII Colloquium on Musical Informatics* (pp. 67–70). Gorizia: Associazione di Informatica Musicale Italiana.

Bengtsson, I., & Gabrielsson, A. (1983). Analysis and synthesis of musical rhythm. In J. Sundberg (Ed.), *Studies of music performance* (pp. 27–60). Stockholm: Royal Swedish Academy of Music.

Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics, 58*(1), 125–141.

Bresin, R. (1993). MELODIA: A program for performance rules testing, teaching, and piano score performance. In G. Havs and I. Pighi (Eds.), *Proceedings of the X Colloquium on Musical Informatics* (pp. 325–327). Milano: Associazione di Informatica Musicale Italiana.

Bresin, R. (2000). *Virtual virtuosity: Studies in automatic music performance.* Unpublished doctoral dissertation, Royal Institute of Technology, Stockholm.

Bresin, R, & Battel, G. U. (2000). Articulation strategies in expressive piano performance. Analysis of legato, staccato, and repeated notes in performances of the Andante movement of Mozart's Sonata in G major (K 545). *Journal of New Music Research, 29,* 211–244.

Bresin, R., & Friberg, A. (2000). Emotional coloring of computer controlled music performance. *Computer Music Journal, 24*(4), 44–62.

Cambouropoulos, E. (1998). *Towards a general computational theory of musical structure.* Doctoral Thesis, Faculty of Music and Department of Artificial Intelligence, University of Edinburgh.

Clarke, E. F. (1988). Generative principles in music performance. In J. A. Sloboda, (Ed.), *Generative processes in music* (pp. 1–26). Oxford: Clarendon Press.

Collier, G. L., & Collier, J. L. (1994). An exploration of the use of tempo in jazz. *Music Perception, 11*(3), 219–242.

De Poli, G., Rodà, A., & Vidolin, A. (1998). Note-by-note analysis of the influence of expressive intentions and musical structure in violin performance. *Journal of New Music Research, 27*(3), 293–321.

Desain, P., Honing, H., & Heijink, H. (1997). Robust score-performance matching: Taking advantage of structural information. In *Proceedings of the 1997 International Computer Music Conference* (pp. 337–340). San Francisco, CA: CMA.

Drake, C., & Palmer, C. (1993). Accent structures in music performance. *Music Perception, 10*, 343–378.

Ellis, M. C. (1991). An analysis of "swing" subdivision and asynchronization in three jazz saxophonists. *Perceptual and Motor Skills, 75*, 707–713.

Friberg, A. (1991). Generative rules for music performance: A formal description of a rule system. *Computer Music Journal, 15*(2), 56–71.

Friberg, A. (1995, May). Matching the rule parameters of Phrase arch to performances of "Träumerei": A preliminary study. In A. Friberg and J. Sundberg (Eds.), *Proceedings of the KTH Symposium on Grammars for Music Performance* (pp. 37–44). Stockholm: Royal Institute of Technology.

Friberg, A., Bresin, R., Frydén, L., & Sundberg, J. (1998). Musical punctuation on the microlevel: Automatic identification and performance of small melodic units. *Journal of New Music Research, 27*(3), 271–292.

Friberg, A., Colombo, V., Frydén, L., & Sundberg, J. (2000). Generating musical performances with Director Musices. *Computer Music Journal, 24*(3), 23–29.

Friberg, A., & Sundberg, J. (1999). Does music performance allude to locomotion? A model of final *ritardandi* derived from measurements of stopping runners. *Journal of the Acoustical Society of America, 105*(3), 1469–1484.

Friberg, A., Sundberg, J., & Frydén, L. (1987). How to terminate a phrase: An analysis-by-synthesis experiment on the perceptual aspect of music performance. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (pp. 49–55). Stockholm: Royal Swedish Academy of Music.

Friberg, A., & Sundström, A. (in press). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. *Music Perception.*

Gabrielsson, A. (1987). Once again: The theme from Mozart's piano sonata in A major (K. 331). A comparison of five performances. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (pp. 81–103). Stockholm: Royal Swedish Academy of Music.

Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.), *The psychology of music* (2nd ed.) (pp. 501–602). New York: Academic Press.

Gabrielsson, A., Bengtsson, I., & Gabrielsson, B. (1983). Performance of musical rhythm in $\frac{3}{4}$ and $\frac{6}{8}$ meter. *Scandinavian Journal of Psychology, 24*, 193–213.

Hefling, S., E. (1993). *Rhythmic alteration in seventeenth- and eighteenth-century music:* Notes inégales *and overdotting.* New York: Schirmer.

Henderson, M. T. (1936). Rhythmic organization in artistic piano performance. In C. E. Seashore (Ed.), *Objective analysis of musical performance, University of Iowa Studies in the Psychology of Music*, Vol. 4 (pp. 281–305). Iowa City: University of Iowa Press.

Honing, H. (1990). POCO: An environment for analysing, modifying, and generating expression in music. In *Proceedings of the 1990 International Computer Music Conference* (pp. 364–368). San Francisco, CA: CMA.

Huron, D., & Royal, M. (1996). What is melodic accent? Converging evidence from musical practice. *Music Perception, 13*, 489–516.

Juslin, P., Friberg, A., & Bresin, R. (in press). Toward a computational model of performance expression: The GERM model. *Musicae Scientiae.*

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch.* New York: Oxford University Press.

Krumhansl, C. L. (1996). A perceptual analysis of Mozart's piano sonata K282: Segmentation, tension, and musical ideas. *Music Perception, 13*, 401–432.

Lerdahl, F. (1996). Calculating tonal tension. *Music Perception, 13*, 319–364.
Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.
Palmer, C. (1996a). Anatomy of a performance: Sources of musical expression. *Music Perception, 13*(3), 433–453.
Palmer, C. (1996b). On the assignment of structure in music performance. *Music Perception, 14*(1), 23–56.
Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48*, 115–138.
Parncutt, R. (in press). Accents and expression in piano performance. *Systemische Musikwissenschaft. Festschrift Jobst P. Fricke zum 65. Geburtsdag.* Frankfurt: Peter Lang.
Penel, A., & Drake, C. (1998). Sources of timing variations in music performance: A psychological segmentation model. *Psychological Research, 61*, 12–32.
Rasch, R. A. (1978). The perception of simultaneous notes such as in polyphonic music. *Acustica, 40*, 21–33.
Rasch, R. A. (1979). Synchronization in performed ensemble music. *Acustica, 43*, 121–131.
Repp, B. H. (1990). Patterns of expressive timing in performances of a Beethoven minuet by nineteen famous pianists. *Journal of the Acoustical Society of America, 88*(2), 622–641.
Repp, B. H. (1992). Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's "Träumerei." *Journal of the Acoustical Society of America, 92*(5), 2546–2568.
Repp, B. H. (1995). Quantitative effects of global tempo on expressive timing in music performance: Some perceptual evidence. *Music Perception, 13*(1), 39–57.
Repp, B. H. (1997). Acoustics, perception, and production of legato articulation on a computer-controlled grand piano. *Journal of the Acoustical Society of America, 102*(3), 1878–1890.
Repp, B. H. (1998a). The detectability of local deviations from a typical expressive timing pattern. *Music Perception, 15*(3), 265–289.
Repp, B. H. (1998b). A microcosm of musical expression: I. Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America, 104*, 1085–1100.
Shove, P., & Repp, B. H. (1995). Musical motion and performance: Theoretical and empirical perspectives. In J. Rink (Ed.), *The practice of performance: Studies in musical interpretation* (pp. 55–83). Cambridge: Cambridge University Press.
Sundberg, J. (2000). Grouping and differentiation. Two main principles in the performance of music. In T. Nakada (Ed.), *Integrated human brain science: Theory, method application (music)* (pp. 299–314). Amsterdam: Elsevier.
Sundberg, J., Friberg, A., & Frydén, L. (1991). Common secrets of musicians and listeners—an analysis-by-synthesis study of musical performance. In P. Howell, R. West, and I. Cross (Eds.), *Representing musical structure* (pp. 161–197). London: Academic Press.
Taguti, T., Mori, S., & Suga, S. (1994). Stepwise change in the physical speed of music rendered in tempo. In I. Deliège (Ed.), *Proceedings of the 3rd International Conference on Music Perception and Cognition, Liège* (pp. 341–342). Liège, Belgium: ESCOM.
Thomassen, M. T. (1982). Melodic accent: Experiments and a tentative model. *Journal of the Acoustical Society of America, 71*, 1596–1605.

Todd, N. P. McA. (1985). A model of expressive timing in tonal music. *Music Perception, 3*, 33–58.

Todd, N. P. McA. (1989). A computational model of rubato. *Contemporary Music Review, 3*, 69–88.

Todd, N. P. McA. (1992). The dynamics of dynamics: A model of musical expression. *Journal of the Acoustical Society of America, 91*(6), 3540–3550.

Todd, N. P. McA. (1995). The kinematics of musical expression. *Journal of the Acoustical Society of America, 97*(3), 1940–1949.

Vernon, L. N. (1936). Synchronization of chords in artistic piano music. In C. E. Seashore (Ed.), *Objective analysis of musical performance, University of Iowa Studies in the Psychology of Music*, Vol. 4 (pp 306–345). Iowa City: University of Iowa Press.

Windsor, W. L., and Clarke, E. F. (1997). Expressive timing and dynamics in real and artificial musical performances: Using an algorithm as an analytical tool. *Music Perception, 15*(2), 127–152.

# 14

# Emotional Communication

PATRIK N. JUSLIN & ROLAND S. PERSSON

To communicate specific emotions to listeners, performers simultaneously manipulate a range of musical parameters. The research findings on emotional expression in music may be organized according to a theoretical framework that describes the communicative process in terms of E. Brunswik's (1956) *lens model*. We discuss traditional strategies for teaching expression including the use of metaphors, aural modeling, and felt emotion and conclude that these strategies rarely provide informative feedback to the performer. A new, empirically based approach to teaching expression called *cognitive feedback* is outlined and its efficacy evaluated. The goal is to provide performers with the tools they need to develop their own personal expression.

*With regard to musical performances, experience has shown that the imagination of the hearer is in general so much at the disposal of the [performer] that by help of variation, intervals, and modulation he may stamp what impression on the mind he pleases.*

Francesco Geminiani, cited in Meyer,
*Emotion and Meaning in Music*

Of all the subskills that make up music performance, the ones associated with emotional communication are often viewed as the most elusive. They go right to the core of why people engage in musical behavior, either as performers or as listeners. The performance of a piece of music is crucial in shaping its emotional expression. Thus the emotional impact of particularly expressive performers—for example, C. P. E. Bach, Niccolò Paganini, and Jimi Hendrix—has always been a source of great fascination. What is the origin of their expressiveness? How is it achieved?

It should be noted at the outset that there are different uses of the word *expression* in the literature. In studies of music performance, *expression* has been used to refer to the systematic variations in timing, dynamics, timbre, and pitch that form the *microstructure* of a performance and differentiate it from another performance of the same music (Palmer, 1997; chapter 13 in this volume). The word *expression* has also been used to refer to the *emotional* qualities of music as perceived by listeners (Davies, 1994). These two senses of the word are of course related in that performers use systematic variations in performance parameters to convey emotions to listeners.

Much has been written about expression in music—far more than we can hope to cover in this chapter. A review of the literature, from antiquity to modern times, reveals a variety of ideas about what music is able to express: emotion, beauty, motion, expressive form, energy, tension, events, religious faith, personal identity, and social conditions. In this chapter, we have chosen for pragmatic reasons to focus on expression of emotion and also limit our review to psychological research. One final limitation reflects the focus of this volume: while not wishing to deny the crucial role of the composer, we shall focus on how the *performer* contributes to emotional expression in music. (For a discussion of the composer, see Juslin & Sloboda, 2001, section 3.)

Questions about music and emotion have occupied humans ever since antiquity. The ancient Greeks argued that specific musical features are associated with specific emotions. This notion received its most precise formulation in the *doctrine of the affections* during the seventeenth and eighteenth centuries (cf. Mattheson, 1739/1954). Since then, conceptions of musical emotions have changed, throughout history, along with changing conceptions of emotions in general (e.g., eighteenth-century affect as a rationalized emotional state in contrast to nineteenth-century emotion as personal and spontaneous expression, see Cook & Dibben, 2001). Influential modern theories of music and emotion include Langer (1942), Meyer (1956), Cooke (1959), and Clynes (1977). Empirical research on emotional expression in music has been conducted for a century, including the pioneering studies by Hevner (1936). This work explored the expressive properties of various musical features, such as pitch or mode (for a review, see Gabrielsson & Juslin, in press). From this research we have learned much about how different aspects of a musical composition may influence emotional expression but less about how different aspects of a *performance* may influence expression.

Few studies have investigated how performers conceptualize and assign meaning to a piece of music while preparing it for performance. However, both empirical research (e.g., Persson, 1993; Woody, 2000) and biographical accounts (e.g., Blum 1977; Menuhin, 1996; Schumacher, 1995) confirm that performers often conceive of performance in terms of emotions and moods. Many performers consider expressivity to be one of the most important aspects of performance (Boyd & George-Warren, 1992; Persson, 1993).

Despite the strong emphasis on expressiveness among musicians, a number of studies suggest that expressive aspects of performance are neglected in music education. Specifically, teachers tend to spend much more time and effort on

technical aspects than on expressive or aesthetic aspects (Persson, 1993; Tait, 1992). As a result, students may come to focus on expressive aspects fairly late in their artistic development (Woody, 2000). Critics often complain that young musicians acquire a high technical skill without being able to induce an emotional experience in the listener (e.g., Dubal, 1985). Consequently, music educators have been encouraged to devote more attention to emotion and expression in music (Reimer, 1989). Such a change of emphasis should lead to an increased concern with the actual strategies used to teach expressive skills of music performance. Teaching strategies form the *how* of music teaching; they involve vocabulary usage, various forms of modeling, and management procedures (Tait, 1992).

The objective of this chapter is to illustrate how psychological research on emotional communication in music performance might contribute to the development of more efficient teaching strategies aimed at the expressive aspects of performance. First, we summarize relevant research. Then, we describe a framework for organizing the findings. Finally, we consider how theory and research on expression might inform day-to-day teaching practice.

## Research on Emotional Expression in Performance

Performers and listeners often discuss and compare musical performances in terms of their expressive aspects, and there are many treatises on performance practices that can be used to enhance emotional expressivity (e.g., Hudson, 1994). Therefore, it may come as a surprise that it is only recently that researchers have began to study this phenomenon. This may reflect the dominant influence of cognitive science in psychology in general and music psychology in particular, which has led to a focus on cognitive aspects of performance, such as structural representation (for a review, see Gabrielsson, 1999). However, renewed interest in emotion has led to increased interest in musical emotion also (Juslin & Sloboda, 2001), even in performance research.

Emotions are difficult to define and measure. Yet most researchers would probably agree that emotions consist of many components: cognitive appraisal, subjective feeling, physiology, expression, and action tendency (Oatley & Jenkins, 1996). For example, we may appraise an event as harmful that evokes feelings of fear and physiological reactions in our body; we may express this fear verbally or nonverbally and may also act in certain ways (e.g., running away) rather than others. There are two major approaches to conceptualizing emotions. According to the *categorical* approach, people experience emotions as categories that are distinct from one another (Ekman, 1992). The categorical approach thus focuses on the characteristics that distinguish emotions from one another. In contrast, the *dimensional* approach focuses on identifying emotions based on their placement on a small number of dimensions like *valence* (positive/negative) and *activation* (high/low) (Russell, 1980).

What is emotional communication? We reserve this term for situations where the performer intends to communicate an emotion to the listener. Accurate com-

munication takes place to the extent that the performer's expressive intention is understood by the listener. This approach implies that expression and recognition of emotions should be studied in an integrated fashion. Accordingly, most studies in the domain have used the following procedure: The performer is asked to play a piece of music to express various emotions (e.g., sadness) chosen by the investigator. The performances are first recorded and later analyzed according to their acoustic characteristics. The performances are also judged by listeners to see whether they perceive the expression in accordance with the performer's intention. Thus, the focus in this research is on *perception* (rather than induction) of emotion.

### *Accuracy*

The first question to face any researcher concerned with emotional communication in music is whether such communication is possible at all. In a pioneering study, Kotlyar and Morozov (1976) asked 10 opera singers to sing phrases from various pieces of music in such a way that they would communicate joy, anger, sorrow, and fear to listeners. Results show that the 10 musically trained listeners who judged the emotional expression of each performance were consistently successful at recognizing the intended expressions. Since then, several studies have confirmed that professional musicians are able to communicate emotions to listeners (Behrens & Green, 1993; Gabrielsson & Juslin, 1996; Juslin, 1997b, 2000; Juslin & Madison, 1999; Sundberg, Iwarsson, & Hagegård, 1995), but that there are considerable individual differences in expressive skill.

Few studies have reported the results in a manner that makes more precise estimates of communication accuracy possible. Juslin (1997c) used a forced-choice format to enable comparison with the results reported in studies of vocal expression of emotion. He found that the communication accuracy was 75% correct—about four times more accurate than would be expected by chance alone. This suggests—at least under ideal circumstances—that the accuracy of communication of emotion in music performance may approach the accuracy of facial and vocal expression (Johnstone & Scherer, 2000). Furthermore, the communicative process seems to be reliable regardless of the particular response format used to collect listeners' judgments (Juslin, 1997a).

### *Code Usage*

A number of studies have attempted to describe the *code* (i.e., the acoustic means) that performers use to communicate emotions to listeners (Gabrielsson & Juslin, 1996; Juslin, 1997b; 2000; Juslin & Madison, 1999; Rapoport, 1996). One of the main findings is that the performer's expressive intention affects almost every aspect of the performance; that is, emotional expression in performance involves a sizable array of *cues* (i.e., pieces of information) that are used by performers and listeners. A summary of the code usage established in investigations so far is shown in Table 14.1. Included are the five emotions—happiness, sadness, anger, fear, tenderness—that have been studied most extensively. These emo-

tions represent a natural point of departure because all are regarded as typical emotions by laypeople (Shields, 1984) and as *basic emotions* (i.e., innate and universal emotions) by scientists (Plutchik, 1994). The same emotions also appear occasionally as expression marks in musical scores (e.g., *festoso, dolente, furioso, timoroso, teneramente*).

The expressive cues shown in Table 14.1 include tempo, sound level, timing, intonation, articulation, timbre, vibrato, tone attacks, tone decays, and pauses. Both the mean level of a cue and its variability throughout the performance may be important for the expression. For example, sadness is associated with slow tempo, low sound level, *legato* articulation, small articulation vari-

Table 14.1. Cue utilization in communication of emotions (based on material presented in Juslin, 2001).

| Emotion | Cue Utilization | Activity, Valence |
|---|---|---|
| Happiness | fast tempo, small tempo variability, *staccato* articulation, large articulation variability, high sound level, bright timbre, fast tone attacks, small timing variations, increased durational contrasts between long and short notes, rising microintonation, small vibrato extent | high, positive |
| Sadness | very slow tempo, *legato* articulation, small articulation variability, low sound level, dull timbre, large timing variations, reduced durational contrasts between long and short notes, slow tone attacks, flat or falling microintonation, slow vibrato, final *ritardando*, phrase *decelerando* | low, negative |
| Anger | high sound level, sharp timbre, spectral noise, fast tempo, *staccato* articulation, abrupt tone attacks, increased durational contrasts between long and short notes, no *ritardando*, sudden accents, accents on tonally unstable notes, *crescendo*, phrase *accelerando*, large vibrato extent | high, negative |
| Tenderness | slow tempo, slow tone attacks, low sound level, small sound-level variability, *legato* articulation, soft timbre, moderate timing variations, intense vibrato, reduced durational contrasts between long and short notes, final *ritardando*, accents on stable notes | low, positive |
| Fear | *staccato* articulation, very low sound level, large sound-level variability, fast tempo, large tempo variability, very large timing variations, bright spectrum, fast, shallow, irregular vibrato, pauses between phrases, sudden syncopations | moderate, negative |

*Note:* This is a simplified description of the relationships established in studies of performers' and listeners' cue utilization in expression of emotion via music performance. For information about the precise definition and measurement of each cue, see Juslin (1999). The rightmost column shows each emotion's position on two emotion dimensions: *activity* (high vs. low) and *valence* (positive vs. negative) (e.g., Plutchik, 1994, chapter 3).

ability, slow tone attacks, and soft timbre (for a detailed review, see Juslin, 1999). Also, different tones in the melodic structure may be emphasized depending on which emotion the performer intends to express (Lindström, 1999). It is possible to simulate emotional expression in synthesized performances based on the findings in the table in such a way that listeners can accurately decode the intended emotion (Juslin, 1997c; Bresin & Friberg, 2000).

While Table 14.1 is limited to a few emotion categories, one can easily imagine how these categories may be combined or mixed in different ways and also how emotional expression may be altered during a performance (Juslin, 2001). Some of the cues shown in the table may be common knowledge to most performers, whereas other cues (e.g., those that have to do with the microstructure of the performance) are less obvious. The number of cues available, of course, depends on the instrument used. Furthermore, cue utilization is not completely consistent across performers, instruments, or pieces of music (Juslin, 2000).

Most studies in the field so far have focused on cues that are applied in much the same way and to much the same extent throughout a piece. However, it would seem intuitively that much of music's expressiveness lies in the way cue use changes during the course of the performance. Indeed, there seem to exist certain *expressive contours* (i.e., patterns of changes in tempo and dynamics) that are characteristic of specific emotions. Listening tests that employed systematic manipulations of synthesized (Juslin, 1997c) or real performances (Juslin & Madison, 1999) have shown that expressive contours can be used by a listener to decode emotional expression.

### *Performer Insight*

An important question is whether music performers have any insight regarding their own cue utilization. It seems generally agreed that performers are usually not aware of the details of how their musical intentions are realized in performance (Sloboda, 1996). In the case of emotional communication, performers are not normally conscious of how they use the cues listed in Table 14.1 (Juslin & Laukka, 2000).

However, there are large individual differences among performers in this regard. Some performers seem to imagine themselves being in the emotional state that their performance is intended to express and just let things happen (Juslin & Laukka, 2000; Persson, 2001). For example, the Russian violist Yuri Bashmet commented, "Identify with the emotions and the notes—fearful as they are—will look after themselves" (Seckerson, 1991, p. 26). Other performers take a more analytical approach, explicitly pondering how to vary different cues. But to the extent that cues are used *implicitly*, this presents a problem for the teaching of expression, which relies predominately on verbal instruction (Tait, 1992; Woody, 2000).

## A Theoretical Framework

One reason for the neglect of emotion in both performance research and music education may be the lack of relevant theories. In the following, we describe the

only elaborated theoretical framework aimed specifically at emotional communication via performance, namely, the *functionalist perspective* (Juslin, 1997b; 2001).

### *Origin of the Code*

Music performers are able to communicate emotions to listeners. What is it that makes this communication possible? Or, more specifically, what is the origin of the acoustical code used by performers? Arguably, the code reflects contributions from both nature and nurture.

Regarding nature, there is evidence for *innate brain programs* for vocal expression of emotion. Studies of the neurological substrates that underlie spontaneous vocalizations of emotion in monkeys and humans with brain lesions provide evidence of brain circuits that function to initiate innate affect vocalizations (Jürgens, 1992). We hypothesize that music performers communicate emotions to listeners by using the same acoustic code as is used in vocal expression. This hypothesis is supported by similarities in cue utilization between music performance and vocal expression (Juslin, 1999). For example, vocal expression of sadness involves slow speech rate, low voice intensity, and little high-frequency energy in the spectrum of the voice. Similar acoustic cues are used to express sadness in music performance (Table 14.1). Cross-cultural studies of vocal expression (Johnstone & Scherer, 2000) and music performance (Juslin, 2001) show that there are cross-cultural similarities. Thus there seems to exist an innate code for acoustical communication of emotion, which could explain why emotional expression is often regarded by music teachers as instinctive. If musicians are to apply this innate code to their performance, they need to understand the parallels between human voices and musical instruments and to learn sufficient technique to express emotions in accordance with the vocal code.

The second factor that governs emotional expression in performance is social learning or specific memories. This is a lifelong process that begins with the early interaction between mother and infant. When mothers talk to their infants, for example, if they want to calm their infant, they reduce the speed and intensity of their speech and talk with slowly falling pitch contours. If mothers want to express disapproval toward some unfavorable activity they employ brief, sharp, and *staccato*-like contours (Papoušek, 1996). Although the code used by mothers seems to be innate, the particular expressive style of the mother modulates the expressive style of the infant. This modulation of the expressive skills continues throughout life as one accumulates experience. Performers learn links between cues and extramusical aspects (e.g., motion, body language) through analogies (Persson, 1993; Sloboda, 1996; see also chapter 15 in this volume), which implies that extramusical life experiences are crucial in learning expressivity in performance (Woody, 2000).

### *Description of the Code*

One way of capturing the crucial characteristics of the communicative process is to conceptualize it in terms of a variant of Brunswik's (1956) lens model. The

lens model was originally intended as a model of visual perception that described the relationship between an organism and distal cues. However, it was later used mainly in studies of human judgment where the goal was to relate the judge's judgment strategy to a description of the judgment task (Cooksey, 1996).

The modified lens model (Figure 14.1) illustrates how performers *encode* (i.e., express) emotions by means of a set of cues (e.g., variations of tempo, sound level, and timbre) that are *probabilistic* (i.e., uncertain) and partly *redundant*. The emotions are *decoded* (i.e., recognized) by listeners who use these same cues to infer the expression. The cues are probabilistic in that they are not perfectly reliable indicators of the intended expression. Performers and listeners have to *combine* the cues for reliable communication to occur. However, this is not simply a matter of pattern matching, because the cues contribute in an additive fashion to listeners' judgments: Each cue is neither necessary nor sufficient, but the larger the number of cues used, the more reliable the communication. The redundancy among cues partly reflects how sounds are produced on instruments (e.g., a harder string attack produces a tone that is both louder and sharper in timbre).

In Figure 14.1, *accuracy* refers to the correlation between the performer's intention and the listener's judgment. *Matching* refers to the degree of similarity between the performer's and the listener's cue utilization, respectively. For

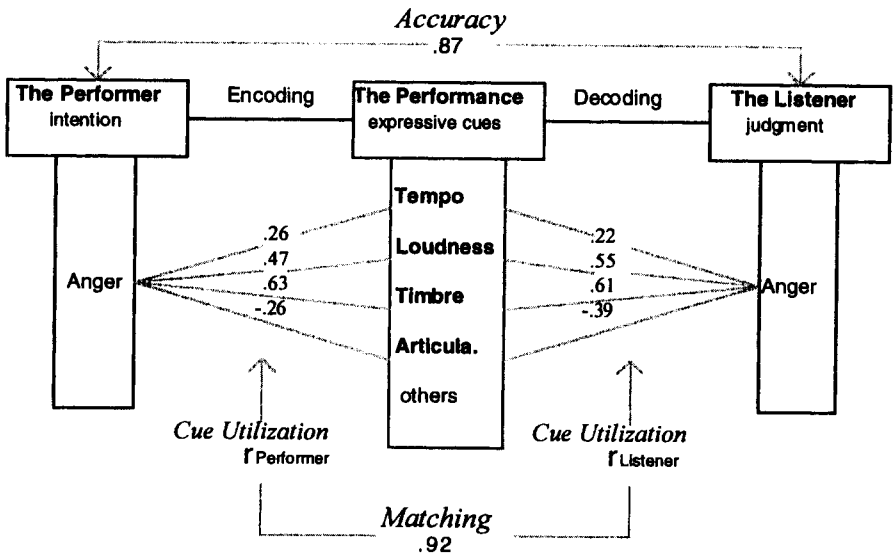



Figure 14.1. A Brunswikian lens model of emotional communication in music performance (adapted from Juslin, 1997b). The cue weights should be interpreted as follows: Positive (as opposed to negative) signs indicate, respectively for each cue, fast (vs. slow) mean tempo, high (vs. low) sound level, sharp (vs. soft) timbre, and *legato* (vs. *staccato*) articulation (for details regarding methodology, see Juslin, 2000).

successful communication to occur, the performer's cue utilization must be as similar as possible to the listener's cue utilization. For illustrative purposes, unpublished data have been inserted into Figure 14.1. As can be seen, there is a high correlation between intention and judgment, which means that the performer was successful in communicating anger to the listener. This is explained by the high level of matching between the performer's and the listener's cue utilization.

Brunswik's lens model has a number of important implications. First, we cannot expect perfect accuracy of communication. If cues are only probabilistic, that means that accuracy can only be probabilistic, too. Second, to understand why the communication is successful or not in a particular situation we have to describe both expression and recognition in terms of the same concepts. Third, because the cues are partly redundant, many different cue utilization strategies may lead to the same level of accuracy (Juslin, 2000). There is a virtue associated with this redundancy. Because there is no pressure toward uniformity in cue utilization, performers can communicate successfully with listeners without compromising their unique playing styles.

## Applications in Music Education

### *Can Emotional Expression Be Learned?*

Expertise in music performance is commonly seen as the synthesis of technical and expressive skills (chapter 7). However, technical aspects of playing are often regarded as learnable skills, whereas expressive aspects are regarded as being more instinctive. Many teachers view expression as something that cannot be taught—a view that is sometimes shared by their students: "There is no technique to perform expressively. You have to use your soul" (cited in Woody, 2000, p. 21).

This view probably stems from certain myths about artistic expression. One such myth is that expression is entirely subjective and passive in its genesis and has nothing to do with understanding (Howard, 1989). In this view, consistent with the romantic notion of art and artists as a kind of mystery, expression is seen as a hands-off affair that is best left alone. Hence, expression is wrongly believed to reflect a divine endowment—or talent—beyond learning and development (cf. Sloboda, 1996).

Although we often think of artistic expression in terms of individual talent, the criteria for what is regarded as high art ultimately reflect social construction. Nowhere is this more apparent than in an educational setting (Kingsbury, 1988). Part of the problem the music teacher faces resides in the tension between the subjective world of individual performers and the social and objective requirements of the educational setting. Hence, an important goal of any teaching strategy aimed at developing expressive skills should be to *relate* the subjective world of the performer (e.g., imagery, metaphor, emotion) to objective features of performance (e.g., articulation).

Emotional expressiveness in music is sometimes hard to describe in words, making it somewhat elusive to both research and teaching practice. Furthermore, it is probably true that expressive skills to some extent reflect the emotional sensitivity of the performer. But this does not imply that it is impossible to learn expressive skills through training. Studies that have addressed this issue have demonstrated that expressive skills *can* be improved by training (Marchand, 1975; Juslin & Laukka, 2000; Woody, 1999).

The idea that the learning of musical expression is best left untouched by conscious thought reflects a misunderstanding that pervades commonsense teaching (i.e., teaching that relies not on empirically derived models and knowledge but rather on tradition and folklore) and is typical of master performers who may never have received training on how to teach musicians (Persson, 1996). It is not realized that goal-directed strategies that initially are willfully applied normally undergo *automation* as a result of practice (Hallam, 1997). That is, although a performer may initially have to use cues in a conscious manner, soon the associations among cues and emotions become internalized by the performer and no longer require conscious control.

### *Traditional Strategies for Teaching Expression*

**Aural Modeling.** Perhaps the most frequently used tool in conveying to students how something should be played is aural modeling. The teacher's performance provides a model of what is desired from the student, and the student is usually required to learn by imitating the teacher. Such strategies are usually achieved on the instrument or using the voice (e.g., singing a certain phrase). Although modeling is useful, it has some limitations. One is that the student is required to pick up the relevant aspects of the model. It may be difficult for a student to know what to listen for and how to represent it in terms of specific skills (Lehmann, 1997). Davidson and Scripp (1992) note that masterful performances are so compact and seamless that it is difficult to observe the subskills that support a fluent production.

**Experiential Strategies.** There are, however, a number of experiential teaching strategies also, which instead aim at conveying the subjective aspects of performing to a student. One such strategy is the use of *metaphors* to focus the emotional qualities of the performance by creating an emotional state within the performer (Davidson & Scripp, 1992). A teacher may, for example, say: "Well, I don't think the lover and his lass had much fun, do you? If a young girl and her boy go into the cornfields, they have other things in mind than playing cards surely!" (comment to a student singing Finzi's "It was a Lover and His Lass," in Persson, 1993, p. 328). Or a teacher may say: "Strange . . . myths . . . ghosts . . . dragons . . . eerie . . . the unknown . . . spine-chilling . . . a drama . . . Weber was an opera composer!" (comment to a student playing an instrumental piece by C. M. Weber, in Persson, 1993, p. 328). Although metaphors can be effective (Kohut, 1985), there are some problems with them. One problem is that metaphors depend on the performer's personal experience with

words and images. Since different performers have different experiences, metaphors are often ambiguous.

Another teaching strategy endorsed by some teachers is to focus on the performer's *felt emotions*, trusting that these emotions will naturally translate into appropriate sound properties (Woody, 2000). Felt emotion, however, is no guarantee that the emotion will be successfully conveyed to listeners; nor is it necessary to feel the emotion in order to communicate successfully. As observed by Sloboda (1996), students rarely monitor the expressive outcomes of their own performances. Instead, they monitor their own intentions and "take the intention for the deed" (p. 121).

**The Problem of Feedback.** Based on an extensive review of a century of research on skill acquisition, Ericsson, Krampe, and Tesch-Römer (1993) proposed three elements that are required in a learning task for it to qualify as *deliberate practice*: (1) clear task definition (2) informative feedback, and (3) opportunities for repetition and correction of errors. Our view is that traditional strategies for teaching expression are problematic mainly in that they do not provide informative feedback to the performer.

Technically, feedback can be defined as "the process by which an environment returns to individuals a portion of the information in their response output necessary to compare their present strategy with a representation of an ideal strategy" (Balzer, Doherty, & O'Connor, 1989, p. 412). It follows from this definition that many kinds of responses that teachers give their students (e.g., "put some expression into it!") do not actually constitute feedback. Thus we agree with Tait's (1992) astute conclusion that "teaching strategies need to become more specific in terms of tasks and feedback" (p. 532). Teachers need a pedagogical model that can guide the development of expressivity and emphasize its significance for successful communication with an audience.

## *Alternative Approaches to Teaching Emotional Communication*

**Teaching Theory of Emotional Communication.** One fairly straightforward way of taking advantage of performance research would be to teach the code description shown in Table 14.1. This description could be used as a springboard for exploring different interpretations and reflecting on what makes for a stunning or lackluster performance of a specific piece of music. The value of reflection and adoption of multiple stances toward a piece of art is often emphasized. True enough, simple application of the expressive principles in Table 14.1 may not result in living expression, but these principles can serve as the core from which subtler interpretations are developed with knowledge about the particular instrument, piece, style, and composer.

We believe that the ultimate goal should be to provide performers with the tools they need to develop their own personal expression. The performer may wish the interpretation of a given piece to be clear or ambiguous, stable or variable, specific or general, depending on the composition. Knowledge about the

relationships between expressive cues and their emotional effects will help performers to reliably achieve desired listener responses. As in the visual arts, a performer may need to know the underlying principles and conventions in order to know how to vary them in an aesthetically pleasing manner.

Theory of emotional communication could usefully be taught to music students, based on the framework discussed earlier. A deeper understanding of the communicative process may be beneficial for the learning process. Rabinowitz (1988) argues that "knowledge about the how, where, and why of strategy use [is] important if students are to take control of their cognitive processing" (p. 234). Such knowledge about expressive skills makes it more likely that students retain the new knowledge and are able to apply it to new situations.

**Cognitive Feedback.** Recall that emotional communication in music performance involves a number of acoustic cues that are used by both performers and listeners. Both expression and recognition of emotions are made by *integrating* these cues. Such integration requires knowledge about the relationships among performers, cues, and listeners. The notion of *cognitive feedback* (hereafter, *CFB*) is to allow the performer to compare a model of his or her cue utilization to an *optimal* model of cue utilization. This is, of course, a fundamental feature of feedback as such. However, how this is achieved is crucial. CFB was originally developed in studies of cognitive judgment (Cooksey, 1996). To illustrate how CFB works in a musical context, we summarize a study that evaluated the efficacy of the CFB strategy (for details, see Juslin & Laukka, 2000).

In the first phase of the study, eight amateur guitarists were asked to play a piece of music in such a way that they would communicate four different emotions to listeners. Their performances were recorded and analyzed in terms of acoustic characteristics. Listeners also rated each performance on scales that corresponded to the intended emotions. Then statistical analysis was used to model the relationships between (1) performers' intentions and cues and (2) cues and listeners' judgments. Finally, the performer and listener models were quantitatively related to each other (by means of the lens model equation; Cooksey, 1996). This made it possible to directly *compare* how performers and listeners used the cues in the performances (cf. Figure 14.1).

In the second phase of the study, half of the performers returned to the laboratory to receive CFB. This consisted of a description of the performer's cue utilization, the listeners' cue utilization, and the degree of matching between performer and listeners. Instances of mismatching between performer and listeners were explained to individual performers. For example, a performer would be told that "you used *staccato* articulation in your sadness expressions, but the listeners associated *legato* articulation with sadness. You should therefore try to play your sadness expressions with more *legato* articulation." In other words, the performers were given specific directions on how to alter their cue utilization to make it more similar to the listeners' cue utilization. The other half of the performers acted as control group and did not receive any feedback.

In the final phase, all performers returned to the laboratory to repeat the first task once more. That is, they were once again asked to perform a piece of music

to communicate emotions to listeners. The goal was to see whether the performers who had received CFB had improved their communication accuracy more than the control group. Two kinds of criteria were used to evaluate the efficacy of CFB: (1) behavioral criteria (did CFB improve the performers' accuracy?) and (2) reaction criteria (how did the performers react to CFB?).

Figure 14.2 shows the results. As can be seen, *before* feedback there was no significant difference in accuracy between the groups. However, *after* feedback there was a significant difference in accuracy. CFB yielded about a 50% increase in accuracy after a single session. Furthermore, the performers reacted positively to the CFB and believed that it had really improved their skills. In contrast, the control group did *not* improve its accuracy, suggesting that simple task repetition leads to little improvement in this context. Improvement requires informative feedback.

There is often disagreement about whether teachers' instructions should address sound properties of the performance or experiential concepts. CFB resolves this problem because it *relates* the acoustical vocabulary to the experiential vocabulary. When using CFB, one is consensually establishing the extent to which a performance is successful in communicating a certain emotion, rather than relying on the judgment of a single person. The teacher's advice is provided automatically, as it were, from the statistical analysis of the communicative process. The analysis renders transparent much of the previously opaque process of communication—in particular, the combined and unique effects of the acoustic cues. Those effects are extremely difficult for a listener or performer to infer.

### *Applications of Cognitive Feedback*

Among the benefits of CFB are that (1) its efficacy has been empirically demonstrated, (2) it is not subject to the changing ideals of individual teachers, and (3) it is consistent with empirically confirmed features of deliberate practice. Among the drawbacks of CFB—in the form described—are that CFB is complicated and time-consuming to carry out and requires special equipment for acoustic analyses. With both benefits and drawbacks in mind, there are various ways, present and future, of implementing CFB in teaching practice.

**CFB as Computer-Assisted Teaching.** Applications of computer feedback for music performance have been around for some time, concerning, for instance, piano technique and intonation in singing (Hallam, 1997). However, none of the applications concern the expressive aspects of performance. Through the use of algorithms for automatic analysis of acoustic cues in performances and statistical models that simulate emotion judgments of various listener populations, CFB could be applied in teaching more swiftly. User-friendly software for this purpose is currently being developed in a research project at the Uppsala University led by the first author of this chapter. Computer-assisted teaching could serve as a complement to traditional teaching, allowing performers to experiment freely with interpretive ideas.

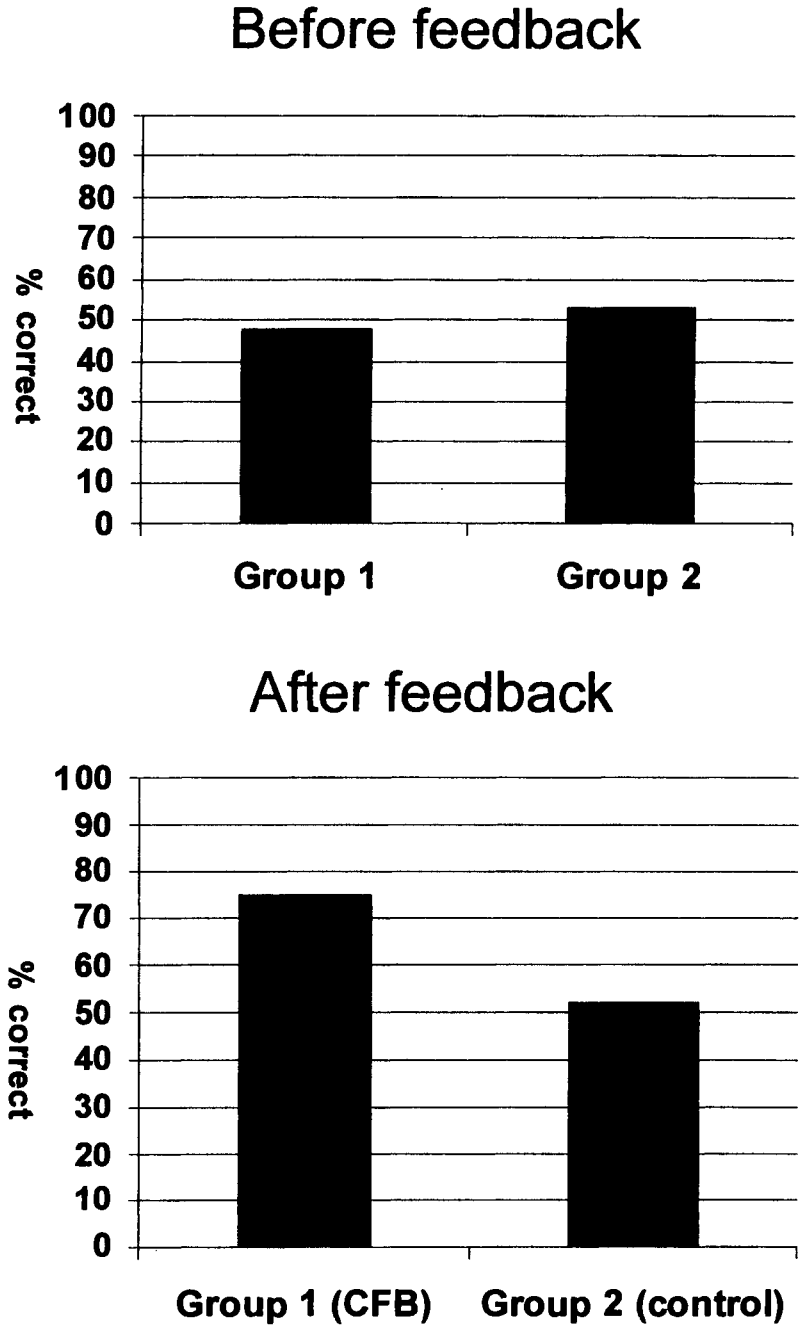



Figure 14.2. Between-subjects comparisons of communication accuracy (% correct) for Group 1 (cognitive feedback, CFB) and Group 2 (control) before and after feedback (from Juslin & Laukka, 2000).

**CFB in a Noncomputerized Setting.** Prior to the production of appropriate software, the CFB strategy can still serve as a pedagogical tool, although without the exact and objective feedback provided by computer-assisted procedures. No statistics or computer programs are required when simply asking a group of students to perform a certain piece of music with the intention of communicating a number of emotions to a group of listening students, and then asking the listeners how they perceived the different performances. A discussion could follow that highlights the possible reasons that some performers seemed more successful in conveying their intentions than others using the cue utilization list (Table 14.1). The purpose of this exercise could be to make students *aware* of the many possibilities of expressiveness.

### *Toward an Integration of Technique and Expression*

The world of Western art music, in particular, has tended to preserve what Renshaw (1986) terms a "museum culture," which has prompted many educators to focus on the technical aspects of performance. But it is likely that uncovering the nature and function of expression by research will help to restore the balance between technique and expression. However, we must not forget that expression is also a matter of learnable technique. Some of the differences among individuals in expressive ability reflect technical differences, but it is technique that allows artistic flexibility and is not likely to alienate the performer from the subjective experience of the music studied and performed.

In this chapter, we have not only presented possible teaching strategies in relation to developing expressive skills, but we have also endeavored to show that science and arts are neither opposites nor contradictory. However, in regard to finding specific applications of cognitive feedback, further research is needed. We therefore propose that qualitative studies of the subjective world of performers and quantitative studies of the communicative process should be pursued in combination. To achieve this, psychologists, educators, and musicians will need to work in close collaboration. This book may itself contribute to this very goal.

**Acknowledgment.** The writing of this chapter was supported by the Bank of Sweden Tercentenary Foundation.

### *References*

Balzer, W. K., Doherty, M. E., & O'Connor, R. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin, 106*, 410–433.

Behrens , G. A., & Green, S. B. (1993). The ability to identify emotional content of solo improvisations performed vocally and on three different instruments. *Psychology of Music, 21*, 20–33.

Blum, D. (1977). *Casals and the art of interpretation*. Berkeley: University of California Press.

Boyd, J., & George-Warren, H. (1992). *Musicians in tune: Seventy-five contemporary musicians discuss the creative process*. New York: Fireside.

Bresin, R., & Friberg, A. (2000). Emotional coloring of computer-controlled music performance. *Computer Music Journal, 24*(4), 44–62.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments.* Berkeley: University of California Press.

Clynes, M. (1977). *Sentics: The touch of emotions.* New York: Anchor/Doubleday.

Cook, N., & Dibben, N. (2001). Musicological approaches to emotion. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 45–70). New York: Oxford University Press.

Cooke, D. (1959). *The language of music.* London: Oxford University Press.

Cooksey, R. W. (1996). *Judgment analysis.* New York: Academic Press.

Davidson, L., & Scripp, L. (1992). Surveying the coordinates of cognitive skills in music. In R. Colwell (Ed.), *Handbook of research on music teaching and learning* (pp. 392–413). New York: Schirmer.

Davies, S. (1994). *Musical meaning and expression.* Ithaca: Cornell University Press.

Dubal, D. (1985). *The world of the concert pianist: Conversations with 35 internationally celebrated pianists.* London: Victor Gollancz.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6,* 169–200.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363–406.

Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.), *The psychology of music* (2nd ed.) (pp. 501–602). San Diego, CA: Academic Press.

Gabrielsson, A., & Juslin, P. N. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music, 24,* 68–91.

Gabrielsson, A., & Juslin, P. N. (in press). Emotional expression in music. In R. J. Davidson, H. H. Goldsmith, and K. R. Scherer (Eds.), *Handbook of affective sciences.* New York: Oxford University Press.

Hallam, S. (1997). What do we know about practising? Towards a model synthesising the research literature. In H. Jørgensen and A. C. Lehmann (Eds.), *Does practice make perfect? Current theory and research on instrumental practice* (pp. 179–231). Oslo: Norges Musikhøgskole.

Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology, 48,* 246–268.

Howard, V. A. (1989). Expression as hands-on construction. In H. Gardner and D. Perkins (Eds.), *Art, mind, and education* (pp. 133–141). Chicago: University of Illinois Press.

Hudson, R. (1994). *Stolen time: The history of tempo rubato.* Oxford: Clarendon Press.

Johnstone, T., & Scherer, K. R. (2000). Vocal communication of emotion. In M. Lewis and J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed.) (pp. 220–235). New York: Guilford.

Jürgens, U. (1992). On the neurobiology of vocal communication. In H. Papoušek, U. Jürgens, and M. Papoušek (Eds.), *Nonverbal vocal communication* (pp. 31–42). Cambridge: Cambridge University Press.

Juslin, P. N. (1997a). Can results from studies of perceived expression in musical performances be generalized across response formats? *Psychomusicology, 16,* 77–101.

Juslin, P. N. (1997b). Emotional communication in music performance: A functionalist perspective and some data. *Music Perception, 14*, 383–418.

Juslin, P. N. (1997c). Perceived emotional expression in synthesized performances of a short melody: Capturing the listener's judgment policy. *Musicae Scientiae, 1*, 225–256.

Juslin, P. N. (1999). *Communication of emotion in vocal expression and music performance: Different channels, same code?* Manuscript submitted for publication.

Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1797–1813.

Juslin, P. N. (2001). Communicating emotion in music performance: A review and a theoretical framework. In P. N. Juslin and J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 309–337). New York: Oxford University Press.

Juslin, P. N., & Laukka, P. (2000). Improving emotional communication in music performance through cognitive feedback. *Musicae Scientiae, 4*, 151–183.

Juslin, P. N., & Madison, G. (1999). The role of timing patterns in recognition of emotional expression from musical performance. *Music Perception, 17*, 197–221.

Juslin, P. N., & Sloboda, J. A. (Eds.). (2001). *Music and emotion: Theory and research*. New York: Oxford University Press.

Kingsbury, H. (1988). *Music, talent, and performance: A conservatory system*. Philadelphia: Temple University Press.

Kohut, D. L. (1985). *Musical performance: Learning theory and pedagogy*. Englewood Cliffs, NJ: Prentice-Hall.

Kotlyar, G. M., & Morozov, V. P. (1976). Acoustic correlates of the emotional content of vocalized speech. *Soviet Physics: Acoustics, 22*, 370–376.

Langer, S. K. (1942). *Philosophy in a new key*. Cambridge, MA: Harvard University Press.

Lehmann, A. C. (1997). Acquired mental representations in music performance: Anecdotal and preliminary empirical evidence. In H. Jørgensen and A. C. Lehmann (Eds.), *Does practice make perfect? Current theory and research on instrumental practice* (pp. 141–163). Oslo: Norges Musikhøgskole.

Lindström, E. (1999, August). *Expression in music: Interplay between performance and melodic structure*. Paper presented at the Meeting of the Society for Music Perception and Cognition, Evanston.

Marchand, D. J. (1975). A study of two approaches to developing expressive performance. *Journal of Research in Music Education, 23*, 14–22.

Mattheson, J. (1739/1954). *Der vollkommene Capellmeister*. Bärenreiter: Basel.

Menuhin, Y. (1996). *Unfinished journey*. London: Methuen.

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Oatley, K., & Jenkins, J. M. (1996). *Understanding emotions*. Oxford, UK: Blackwell.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48*, 115–138.

Papoušek, M. (1996). Intuitive parenting: A hidden source of musical stimulation in infancy. In I. Deliège and J. A. Sloboda (Eds.), *Musical beginnings: Origins and development of musical competence* (pp. 88–112). Oxford: Oxford University Press.

Persson, R. S. (1993). *The subjectivity of musical performance: A music-psychological real world enquiry into the determinants and education of musical reality.* Doctoral dissertation, Huddersfield University, UK.

Persson, R. S. (1996). Brilliant performers as teachers: A case study of commonsense teaching in a conservatoire setting. *International Journal of Music Education, 28,* 1–15.

Persson, R. S. (2001). The subjective world of the performer. In P. N. Juslin and J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 275–289). New York: Oxford University Press.

Plutchik, R. (1994). *The psychology and biology of emotion.* New York: HarperCollins.

Rabinowitz, M. (1988). On teaching cognitive strategies: The influence of accessibility of conceptual knowledge. *Contemporary Educational Psychology, 13,* 229–234.

Rapoport, E. (1996). Emotional expression code in opera and lied singing. *Journal of New Music Research, 25,* 109–149.

Reimer, B. (1989). *A philosophy of music education.* Englewood Cliffs, NJ: Prentice-Hall.

Renshaw, P. (1986). Towards the changing face of the conservatoire curriculum. *British Journal of Music Education, 3,* 79–90.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161–1178.

Schumacher, M. (1995). *Crossroads: The life and music of Eric Clapton.* New York: Hyperion.

Seckerson, E. (1991, June). Yuri Bashmet as interviewed by Seckerson. *Gramophone,* 26–27.

Shields, S. A. (1984). Distinguishing between emotion and non-emotion: Judgments about experience. *Motivation and Emotion, 8,* 355–369.

Sloboda, J. A. (1996). The acquisition of musical performance expertise: Deconstructing the "talent" account of individual differences in musical expressivity. In K. A. Ericsson (Ed.), *The road to excellence* (pp. 107–126). Mahwah, NJ: Erlbaum.

Sundberg, J., Iwarsson, J., & Hagegård, H. (1995). A singer's expression of emotions in sung performance. In O. Fujimura and M. Hirano (Eds.), *Vocal fold physiology: Voice quality control* (pp. 217–229). San Diego, CA: Singular Press.

Tait, M. (1992). Teaching strategies and styles. In R. Colwell (Ed.), *Handbook of research on music teaching and learning* (pp. 525–534). New York: Schirmer.

Woody, R. H. (1999). The relationship between explicit planning and expressive performance of dynamic variations in an aural modeling task. *Journal of Research in Music Education, 47,* 331–342.

Woody, R. H. (2000). Learning expressivity in music performance: An exploratory study. *Research Studies in Music Education, 14,* 14–23.

# 15

# Body Movement

JANE W. DAVIDSON & JORGE SALGADO CORREIA

Body movement plays a role in the construction, execution, and perception of musical performances. We explore the interface between technical matters of physical control and the expressive components of physical gestures and discuss the bodily origins of musical meaning, expressive performance, and musical skill acquisition. For example, bodily gesture and rhythm in proto-musical mother-child exchanges influence the development of thought and knowledge, and expressive slowing in music (ritardando) corresponds to the deceleration of runners coming to a halt. Specific movement gestures in music performance function as illustrative and emblematic cues and clearly indicate the focus of the performer's attention, whether on the narrative content of a song or on showing off to the audience. Thus, through body movement thoughts and concerns are communicated to the audience. Performers, educators, and students can use this knowledge to enhance their performing, teaching, and learning capacities.

All musicians use their bodies to interact with their musical instruments when performing music. The intention of this chapter is to show that the body is not only essential to the physical manipulation of the instrument for the accurate execution of music, but it is also vital in the generation of expressive ideas about the music. In addition, the body seems to be critical in the production and perception of information about the performer's concerns to coordinate with coperformers and audience and to engage in extramusical concerns on stage, for instance, preening display movements.

To begin the exploration of musical performance and the body, it is important to recall that parents rhythmically pat, bounce, and caress their babies and melodically vocalize softly to them in *motherese* (infant-directed speech). In essence, the movements and sounds are based on exchanges that utilize varying elements associated with expressive musical performance. The motherese de-

velops into play and conversational turn taking (Papoušek, 1996). There is imitation, matching, elaboration, with all kinds of emotions being played out through pitch, rhythm, and movement variation for the giver and the receiver of the information. Trevarthen (1999–2000) has argued that these *proto-musical* experiences are critical in establishing physiological and emotional attunement so that the interactive patterns become salient and, therefore, meaningful.

Given Trevarthen's ideas, it may be that proto-musical behavior is vital for the general development of meaningful thoughts and actions. It would follow that a concrete and practical way of understanding music can be explored through bodily movement. To generate music, the formulation and execution of expressive ideas and qualities in movement is certainly necessary. By implication, this would suggest that the understanding of musical performance could be clarified or enhanced when the movements that produce the musical sounds are given due attention by both the performer and the audience.

In light of the preceding, it seems no accident that physical metaphors are commonly used in reference to music: "a punchy performance," a "flowing *legato*," for example. Johnson (1999), Clarke (2001), and Damasio (1999), from the disciplines of philosophy, music psychology, and neuroscience respectively, all concur that these metaphors are helpful in assisting us to relay how we experience music. For performers, teachers, and students, it seems that the movement metaphor may be a powerful tool in the development of a musical performance. We shall return to this and other points about strategies for developing meaningful performances later in the chapter when considering how musical expression may be taught through body movement. But after a brief presentation of a justification for the central importance of the body in the production and perception of musical performances it seems critical to examine both elements in detail.

## Production of Body Movement in Performance

A motor program is conceived of as a hierarchical structure that translates information input into performed action. When we learn to play a musical instrument and then work on a specific piece of music, these programs are constructed and strengthened with rehearsal. It is now well documented that it takes in excess of 10,000 hours of technical and repertoire practice (see Ericsson, Krampe, & Tesch-Römer, 1993) to refine an instrumental technique to reach expertise (the professional standard to play in a symphony orchestra); and during this practice the development of motor programs interfaces with that of other cognitive skills such as memorization in order for the whole learning and playing process to become automated.

Examining different skilled performances made by typists and pianists, Shaffer (1982, 1984) explored how motor programs are organized. He discovered that in typing, sequences of key presses have definite timing profiles across performances of the same word sequences. He concluded that the movement control program elicits effects on the performance that are inevitable but expressive only

of the individual's production of words. In piano performance, Shaffer (1984) concluded that some of the timing deviations from the metric pulse of the musical rhythm are, similarly, the inevitable expression of features of the motor program. However, he went on to illustrate that other timing profile variations are flexible in tone production. Though these features are themselves integrated into the motor program, they are seen as timing modifications intended to enhance the musical effect. These intentional variations can be increased or diminished but never totally removed. Thus two elements of motor programming coexist in musical performance: timing profiles directly related to simple execution and timing deviations for musically enhancing effects.

In line with Shaffer's evidence, it is essential to recognize that dynamics, intonation, and timbre, as well as timing, are manipulated according to an individual's performance interpretation, and indeed, it has been demonstrated that differently intentioned performances contain different quantities of dynamic variation that are consciously used and recognized by performers as expressive devices (Nakamura, 1987). As stated earlier, it seems that the expressive effects applied to musical structures may be, in some part at least, related to bodily sources. For instance, a mathematical model produced by Todd (1985) indicated that phrase-related tempo changes—the acceleration and deceleration of tones in a musical phrase—are directly proportional to their position in the musical phrase. For example, slowing at a phrase boundary occurs in a systematic and seemingly rule-governed manner. Later Todd (1992, 1995) modified the proposal, basing the new model of deceleration toward a musical phrase end on velocity in physical motion. Recently it has been discovered that the expressive device of slowing toward a musical phrase end may be linked to the human movement properties found when adults slow down and stop after running (Friberg & Sundberg, 1999).

So, considering the body in the production of music, it is critical that motor programs are well established. However, the body is much more than a mechanistic source of input and output. There is a vast array of movements and gestures that serve many purposes. For example, one study tracked how individual points of the body move in space each 0.20 s during a piano performance, investigating the playing of the same piece of music with three different expressive intentions labeled deadpan, normal, and exaggerated (Davidson, 1994; in press). The study revealed a relationship between the movement size and the intensity of the musical sound expression—the more exaggerated the expressive intention of the music, the larger the movement. This suggested a direct correlation between the expressive movement and the production of musical intensity, since these performances were regarded as sounding as well as looking more expressive. In addition, it was revealed that the performer made a constant rocking movement when playing, though this movement was barely visible in the deadpan condition.

There are several possible accounts for the different types of movement used by pianists. First, they could be regarded as being responsive to the sounds produced—that is, reflecting the sounds. In the case of the rocking, there may even be movements analogous to those of self-stimulation observed in speech (Ekman

& Friesen, 1969)—movements used by the performer to make him- or herself feel at ease in the social context of the performance and/or enjoy the musical sounds being made. Support for this idea comes in the finding that although the rocking movement is allied to the timing and some structural features of the music (Clarke & Davidson, 1998), it seems to be not strictly bound to the musical structure, so it may be connected to the performer's individual concerns as well as a reaction to the music.

However, in line with research that suggests that the body provides a source for the generation of musical expression (see Friberg & Sundberg, 1999), it seems that the pianist's movements may be indicators of the mental and physical intentions necessary to generate the expressive performance. Although research on the origin of physical expression is scant, by drawing a parallel between the research on general locomotion (all kinds of traveling patterns from walking to running) and movement in musical performance Davidson (in press) was able to develop a theory of physical expression in playing. Although musical expression (that is, a specifically performed and intended communication) is not equivalent to locomotion, the locomotion research examined how the characteristics of gender and identity were displayed and used for communicative ends and thus expressed in the movement patterns. It is on these grounds that a parallel was made between the research findings and movements of musical expression.

Cutting, Proffitt, and Kozlowski (1978) and Kozlowski and Cutting (1977) were interested in how observers used the visual information available in the locomotion patterns. The researchers used bands of reflective tape attached to all the major body joints and filmed in highly contrasting light conditions to produce observation data that showed movement patterns only, avoiding surface information like dress, hairstyle, and so forth. From these movement displays onlookers were initially asked to identify the walker's sex. Then, in a series of studies where the onlookers knew all the walkers, they were asked to identify the walkers by name. Onlookers were very good at determining sex and naming individuals from these minimal visual displays.

The researchers showed that any part of the walking movement cycle and any body joint shown as a two-second excerpt provides similar information expressive of either gender or identity. The results were explained by demonstrating that there is a point within the walker's body referred to as the *center of moment* that acts as a reference around which movements in all parts of the body are organized, and this center is different for men and women (Cutting et al., 1978). Davidson proposed that such a principle may be in operation in musical performance, the description providing a framework from which we can begin to understand the importance of the rocking movement in the pianist's performances. The idea is that there is a physical center for the expression of musical information, the rocking being a critical part of the generation of the musical expression.

From the analysis of the pianist's movements, Davidson noted that the rocking emanated from the hip region. Given the sitting position of a pianist, the hips represent the fulcrum for the pianist's center of gravity and therefore provide the pivotal point for all upper torso movements. This center of gravity seemed

to be the central location for the generation of this overall physical expression: thus the hips may be the center of moment. Although only a single pianist was used in a detailed case study (Davidson, in press), previous studies had indicated that a range of individual performers make similar types of movements and modifications to them across different expressive interpretations, suggesting that generalizable findings can be assumed from the case-study material (Davidson, 1993, had studied the perception of violinists' movements).

Although some musicians appear to remain absolutely still when performing, data from many studies (Davidson 1991, 1993, 1994, 1997, in press) have shown that despite individual differences, when interviewed, many of the players said that the musical ideas they had were connected with the concept of a repetitive whole-body motion. It was an imagined sense of motion. Thus it could be that the movement may be internalized as well as externalized, and so it may be present to some degree. This idea links with that of Alexander Truslit, the German pedagogue:

> Musical motion is internal and encompasses the whole human being. It is not only an emotion [*Gemüts-Bewegung*] but also a true motion sensation. It must be distinguished from acoustic vibrations, from sympathetic resonance, from technical movements in playing an instrument, from the sequence of tones (which is only the outward manifestation of the inner process), and from conducting movements (though they merge partially with it). Musical motion can be likened to an invisible, imaginary dance. (Truslit, 1938, translated and presented by Repp, 1993, p. 51)

For Truslit, the movement metaphor for understanding and experiencing movement is of key importance. With the center-of-moment principle, the claim is that real or imagined motion is necessary to generate musical expression. This is a contentious idea, but nonetheless it brings together several allied propositions to produce a coherent theoretical proposal.

Related to the center-of-moment principle and in addition to the rocking motion, individually identifiable expressive locations have been discovered (see Davidson, in press). When these specific movements were closely examined in relation to the musical score, it was found, first, that they only occurred when a particular body part—the hand, for instance—was free to move, that is, when it was not tied to the technicalities of executing the music. Examples of these locations would be structural features like rests and sustained pedal notes. Second, the movements were all of a rotational nature—circling wrist movements and shoulder rotations, for instance. Third, these movements seemed to form a vocabulary of gestures in that they appeared across performances of different styles of music. They were not, however, allied in a one-to-one manner with musical features or pieces. There was a tendency for the performer to use the gestures in a variety of contexts and manners—sometimes small and slow, other times large and fast.

Further studies had refined the description of the center of moment by suggesting that different parts of the body convey similar expressive information at different hierarchical levels (Cutting & Proffitt, 1981). This finding could account

for results obtained earlier that indicated that more localized movement is of a similar expressive type to that of the overall rocking motion.

Gellrich (1991) has explored how specifically learned gestures can furnish a musical performance with an expressive intention. From his perspective, it is apparent that many of the gestures are culturally learned, as are their associated meanings. It is certain, in all forms of music, that some elements of physical gesture are learned from shared cultural codes. They may be picked up from teachers, peers, or even observation of someone on TV. Technically, the gestures may be *superfluous* to the production of the music—a surface level of movement or a kind of rhetoric that the performer adds to the performance. It could be that the deep-level source of expression is the rocking movement and that the local expressive gesture is a means of adding to that.

While Davidson's is case-study work, it does seem that the findings may be generalizable, since the growing research literature on nonverbal gesture in conversation and social interaction indicates that gestures often have particular purposes and these are commonly understood and used within a specific social and cultural context (see Ellis & Beattie, 1986; Argyle, 1988). For instance, *emblems* (such as a raising of the eyebrows or a shoulder shrug) are typically used to reinforce what is being said. The movement vocabulary of gestures used in music could have some similar functions. For instance, within the context of jazz performance saxophonists often raise the bell of the instrument and close their eyes as a symbol of intense expression, usually getting louder. Some saxophonists make the physical gesture even when they cannot actually play any louder.

So it seems that the bodily production of music involves general motion and specific localized culturally learned gestures for technical, musical, and socially communicative purposes. Beyond production, it would appear that from a perceptual perspective, body movements can assist the degree of coordination between coperformers and aid audience comprehension of the musical performance.

## Perception of Body Movement in Performance

It is now well documented that audiences can detect finely grained information about musical expression (timing, pitch, and dynamics modifications to structural features of the music) and intention (the emotional mood of the performer and the piece) from a musician's body movements when he or she is playing (Davidson, 1993, 1994, 1995). So embodied musical meaning seems to be both perceptually available and comprehensible to audiences. In fact, when audiences are not skilled as listeners they rely entirely on the visual information they receive, not being able to differentiate between musical intentions and aural information (Davidson, 1994).

In an attempt to assess whether all or parts of the movements of a musical performance are meaningful, Davidson (in press, b) explored whether the movement information of a pianist was available to observers in a continuous stream or it was limited to particular moments within a performance. It was discovered,

on the one hand, that audiences reported the repetitive body rocking as being a continuously present and key source of expressive information, but on the other hand clips of only two seconds of visual material revealed that some moments were more obvious indicators of expressive intention than others, especially those moments that contained particular expressive gestures.

So it would appear that those movements used to produce a performance are informative of the musical meaning, as well as being generative of the musical intention itself. Consider the following anecdote about Glenn Gould (cited by Delalande, 1990) that explores the link between audience perception and the performer's use of his body. Gould's career fell into two distinct phases: an early period during which he gave concert tours as well as making sound recordings and a late period when he only worked in the recording studio. Looking at video footage from both periods, Delalande discovered that in the live performances Gould's movements had a degree of unpredictability but showed great fluency. In the sound-recording sessions, the movements were far more repetitive and fixed. It is possible that the earlier movements were for direct communicative intentions: literally demonstrating ideas to the audience. In the later performances, perhaps Gould was less socially motivated and was simply focusing on the best way to achieve a sound performance. Of course, this is only one possible interpretation of Gould's movements, but it nevertheless raises issues about the role of social interaction in shaping the production of the performance. Davidson and Good (1997; in press) discovered that when playing in small ensembles performers use their movement patterns to coordinate timing, dynamics and other expressive effects with their coperformers. So perception of coperformers has an influence on shaping the movement patterns, too.

The nature of the social mediation that goes on among performer, coperformer, and audience in the construction of the performance and the critical role of the body in shaping this have received very little attention in the psychology research literature. One relevant study is on the singer Annie Lennox (Davidson, 2001). Analysis showed that in addition to the specific movements related directly to the communication of the song—such as coordination signals and expressive gestures about the narrative content of the songs—some other types of movements were used purely for audience display or showing-off purposes. These involved a deliberate attempt to involve audience participation and had nothing to do with the song's narrative. For instance, there were moments when Annie Lennox danced playfully with her coperformers and came to the front of the stage, asking the audience members near the stage to clap along in time with the singing. Frith's (1996) analysis of pop singers' behavior showed that the pop singer constantly negotiates a communication of the song's narrative, a pop-star role with associated culturally defined behaviors, and his or her own individuality in the public forum.

Annie Lennox plays the role of a star in her use of showing-off or display gestures. She is a narrator-interpreter in her use of illustrative and emblematic gestures with the coperformers and audience. She is a coworker in her use of regulatory movements to coordinate musical entrances and exits. Internal self-concern also seems to be presented in Annie's adaptive gestures such as the self-

stimulation involved in touching her face or caressing her body with her arms as she sings. The particular illustrative, emblematic, regulatory, and display gestures are specific to Annie and her personal experiences and thus also indicate critical information about an unique individual. However, the adaptive gestures seem to display inner personal states or characteristics, whereas the display, regulatory, and illustrative (emblematic) gestures are more audience (externally) oriented.

### *Summary*

Presenting the findings of the body movement work in summary, it seems that musical skills involve issues of fluency and technique, yet in the live performance critical concerns appear to be how the performer presents him- or herself as:

- A communicator interacting with coperformers to regulate the performance so that it remains unified
- An individual interpreter of the narrative or expressive/emotional elements of the work being played
- A self with individual experiences and behaviors
- A public figure with a clear aim to interact with and entertain the audience

All of these factors together highlight the tremendous range of skills involved in performance and undoubtedly provide important insights into understanding *what* information is being communicated in a performance.

The literature surveyed shows how critical gestural movements are in the production of and perception of a performance and particularly underscores the role of social interaction in the construction of a performance through the discovery that different types of gestures are used for specific coperformer and audience engagement. It could well be that the presence of others promotes the use of these communicative gestures.

Another issue to emerge from this research is the juxtaposition of movements connected with self and self-adaptation and those connected with pure social display. It seems that different aspects of self and self-projection are revealed in the movements.

Thus far, the work presented has demonstrated that musical performance movements are of the following types: purely biomechanical (they can only occur when the body is free and ready to use them), individual (each person has his or her own style), and culturally determined (some of the movements are learned through imitation of others' behaviors in specific contexts and so have common presentations within the cultural context). Many of these performance movements have clear functions and meanings: to communicate the expressive intention (for instance, a sudden surge forward to facilitate the execution of a loud musical passage or a high curving hand gesture to link sections of the music during a pause); to communicate directly with the audience or coperformers about issues of coordination or participation (for example, nodding the head to beckon the audience to join in a chorus of a song or exchanging glances for the coperformer to take over a solo in a jazz piece); to signal extramusical concerns

(for example, gesturing to the audience to remain quiet); to present information about the performer's personality, with his or her individualized characteristics providing important cues (muted contained gestures or large extravagant gestures, for example); and to show off to the audience.

In addition to the findings presented earlier, some performance movements with no specific value to the audience as interpretative cues are nevertheless the by-products of psychophysical, social, and cultural practices that surround performance—for example, standing rooted on both feet to be able to sing better. Although the audience may not read such cues in their assessment of the performance, taken together it is likely that such movements do add to the overall style and content of the performance.

A final point is that the body movements need to be presented at an optimal level. For example, it seems that too many movements might create an overly exaggerated performance and too few movements might make the performance appear stilted. Recall, for example, that although grand flourishing gestures were fashionable in nineteenth-century musical performance, Liszt was referred to by Glinka as an "exaggerator of nuance" (cited in Morgenstern, 1956, p. 129).

Bearing all of the findings discussed earlier in mind, the final section of this chapter considers their educational implications.

## Working with the Body in Music Performance Teaching

Contemporary method books and study guides include diagrams of the correct positions and postures for holding musical instruments in order to play (see, e.g., Miller, 1996), but few of these texts provide information about how to use the body effectively to achieve a musically expressive performance or a charismatic and highly personalized interaction with the audience. However, there has been a historical interest in such matters. For instance, Pierre Baillot in his *L'art du violon* of 1834 (cited in Stowell, 1985) suggests that different types of movement are necessary to produce different musical tempi. For instance, the *adagio* tempo requires "more ample movements" than the *allegro*, where the notes are "tossed off," whereas in *presto* there is "great physical abandon."

Where an interest in the use of bodily gesture for musical interpretation has been expressed, it has typically been found when the author has an interest in the domain of dance and body alignment, with the most commonly cited source of inspiration for the practical being the teachings of Emile Jacques-Dalcroze (1865–1950), who points to body movement as crucial to the process of unifying the musical elements and focusing on musical expression. The Dalcrozian approach addressed the whole issue of musical meaning and musical communication:

*What is the source of music? Where does music begin?*
Human emotions are translated into musical motion.
*Where do we sense emotions?*
In various parts of the body.

*How do we feel emotions?*
By various sensations produced by different levels of muscular contraction and relaxation.
*How does the body express these internal feelings to the external world?*
In postures, gestures, and movements of various kinds. Some of these are automatic, some are spontaneous, others are the results of thought and will.
*By what instrument does a human being translate inner emotions into music?*
By human motion. (Jacques-Dalcroze writing in 1931, cited in Choksy, Abramson, Gillespie, & Woods, 1986, p. 31).

Some individual teachers may not agree with the Dalcrozian approach to teaching and learning music, but it does centralize the role of the body, connecting sounds to particular physical gestures. Recently Pierce (1994) has made a significant impact in the United States by adapting Dalcrozian principles for work with advanced musicians. Her approach involves teaching rhythm by experiencing the beat of pulse of music through pendular, swinging movements away from the instrument. The point of an exercise like this is to embody the full motion required to produce the attack point on the beat. That is, the student can feel the approaching downbeat and the surrounding moments in the body swing. So quite an abstract musical idea can be played out through the body. In another example, in order to achieve a *legato* line and to assist in the detection of a melody line that may be submerged in a harmony passage, students are asked to trace with the hand an analog to the shape of the melody, attending to weight and flow of the line to assist or draw attention to direction of melody and the surrounding harmonies and other textures.

Although we cannot comment on the efficacy of Pierce's work or, indeed, that of Dalcroze and others, we believe that these teachers draw on principles based on the human body and human nature that can assist development of a deeper base from which music technique and expression can be explored and understood.

In line with the range of ideas presented earlier, it becomes evident that there is not just one way to explore the body in music in order to improve technique, musical expression, and the presentation of self in performance, but rooted in the research discussed earlier in the chapter we now present some of our own methods for working with students to train their bodies. More detail about these practical ideas can be found in Correia (1999a, 1999b) and Davidson, Pitts, and Correia (2000), where experimental evidence that verifies the efficacy of the methods is presented.

**Training Fluency of Movement.** Without doubt, musicians need to be able to be aware of the muscular tensions or bad postural habits that may interfere with their playing. Thus, as a general principle, attempting to find the point of best equilibrium for head, neck, spine, pelvis, and legs is recommended as a good practical grounding for students to develop self-awareness in the overall posture of their bodies when preparing to play. These principles emerge out of the teachings of Alexander (1869–1943) and his principles on body alignment—the Alexander Technique.

**Exploring Expressive Potential: The Narrative of the Musical Work.** Based on drama-style exercises (see Boal, 1992), students are asked to express different characters, affects, or states of mind suggested by the narrative of the piece they are learning. They do this first through physical gesture and then in their musical sounds.

**Working on the Intention of Each Musical Phrase.** Once engaged and acquainted with the style and the overall atmosphere of the piece, the students are then asked to find *action metaphors* for each phrase or effect, trying to build up a nonverbal narrative in which every passage of the piece becomes meaningful and thus communicable through movement or physical gesture. Here the concepts of weight and flow are adapted from the movement analyst and choreographer Laban (1960), who attracted the student's attention to all the dynamic qualities of each movement as a means of understanding how forces act upon it and to give as broad an expressive palette as possible.

**Directing an Interpretation: Conducting the Musical Expression of the Teacher.** Another means of using the body to gain insights into the musical expression is for the student to conduct the expression in the playing of the teacher, using the body as the means through which to get the teacher to make expressive sound changes to his or her playing. Trials and discussion are used to explore all the expressive nuances by gestures and/or movement.

**Finding the Inner and Outer Personality in Movement.** Stategies to develop a sense of a public versus a private communication of one's self on stage has been central to the work of the actor (see Boal, 1992). It seems likely that some of the best performers are able to connect with the communicative principles necessary to achieve a good public presentation, while other individuals need to be aided. In music performance, teachers often ask their students to make their bodily gestures larger and more obvious to the audience so that the presentational component of the performance—its publicness—is clear. Difficulties in terms of both music production and perception seem to arise when the performer's behavior is either not sufficiently public for the concert environment or, on occasion, too large-scale and overt. Here issues of historical style and audience opinion can be explored.

**Teacher's Body Movement Strategies to Develop Musical Expression in the Student.** A wide range of verbal *metaphors*—especially for situations where the expression has to do with inner motions or reactions (like contraction or tenderness) can be used by teachers to stimulate the student's imagination and bodily meaning. These metaphors usually captured physical qualities. For instance, Correia (1999b) reported one teacher as having said, "Play it as if you were throwing pebbles into a pond, and feel the ripples pulsing out of the center of your body to the bow."

## Concluding Comments

The evidence presented in this chapter has shown the significance of body movement in the development, performance, and training of musicians. Inter-

facing empirical findings, theory, and some practical applications, we hope to have inspired readers to reappraise their use of movement in the training of musicians and in their own performances. The work we have covered is in no way meant to be prescriptive, rather a starting point from which further developments can and we believe should be made.

### *References*

Argyle, M. (1988). *Bodily communication* (2nd ed.). London: Methuen.

Boal, A. (1992). *Games for actors and non-actors* (2nd ed.). London: Routledge.

Choksy, L., Abramson, R., Gillespie, A., & Woods, D. (1986). *Teaching music in the twentieth century*. Englewood Cliffs, NJ: Prentice-Hall.

Clarke, E. F. (2001). Meaning and specification of motion in music. *Musicae Scientiae, 5*(2), 213–234.

Clarke, E. F., & Davidson, J. W. (1998). The body in performance. In W. Thomas (Ed.), *Composition–performance–reception* (pp. 74–92). Aldershot: Ashgate.

Correia, J. S. (1999a). Embodied meaning: All languages are ethnic . . . *Psychology of Music, 27*, 96–101.

Correia, J. S. (1999b, September). *Making meaning: Teaching performers musical interpretation when working from a score.* Lucerne, Switzerland: European Society for the Cognitive Sciences of Music Conference on Research Relevant to Music Colleges.

Cutting, J. E., & Proffitt, D. R. (1981). Gait perception as an example of how we may perceive events. In R. D. Walker and H. L. Pick (Eds.), *Intersensory perception and sensory intelligence* (pp. 61–87). New York: Plenum.

Cutting, J. E, Proffitt, D. R., & Kozlowski, L. T. (1978). A biomechanical invariant for gait perception. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 357–372.

Damasio, A. (1999). *The feeling of what happens.* Orlando, FL: Harcourt Brace.

Davidson, J. W. (1991). The perception of expressive movement in music performance. Unpublished doctoral dissertation, City University, London.

Davidson, J. W. (1993). Visual perception and performance manner in the movements of solo musicians. *Psychology of Music, 21*, 103–113.

Davidson, J. W. (1994). What type of information is conveyed in the body movements of solo musician performers? *Journal of Human Movement Studies, 6*, 279–301.

Davidson, J. W. (1995). What does the visual information contained in music performances offer the observer? Some preliminary thoughts. In R. Steinberg (Ed.), *Music and the mind machine: Psychophysiology and psychopathology of the sense of music* (pp. 105–114). Heidelberg: Springer.

Davidson, J. W. (1997). The social in musical performance. In D. J. Hargreaves and A. C. North (Eds.), *The social psychology of music* (pp. 209–228). Oxford: Oxford University Press.

Davidson, J. W. (2001). The role of the body in the production and perception of solo vocal performance: A case study of Annie Lennox. *Musicae Scientiae, 5*(2), 235–256.

Davidson, J. W. (in press). Understanding the expressive movements of a solo pianist. *Jahrbuch Musikpsychologie.*

Davidson, J. W., & Good, J. M. M. (1997, June). Social psychology of performance.

In A. Gabrielsson (Ed.), *Proceedings of the Third Triennial ESCOM Conference* (pp. 329–332). Uppsala, Sweden: University of Uppsala.

Davidson, J. W., & Good, J. M. M. (in press). Towards social psychology of string quartets. *Psychology of Music.*

Davidson, J. W., Pitts, S. E., & Correia, J. S. (2000). Reconciling technical and expressive elements in young children's musical instrument learning. *Journal of Aesthetic Education.*

Delalande, F. (1990, June). *Human movement and the interpretation of music.* Paper presented at the Second International Colloquium on the Psychology of Music, Ravello, Italy.

Ekman, P., & Friesen, W. V. (1969). The repertory of nonverbal behaviour: Categories, origins, usage, and coding. *Semiotica, 1,* 49–98.

Ellis, A., & Beattie, G., (1986). *The psychology of language and communication.* London: Weidenfield & Nicolson.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363–406.

Friberg, A., & Sundberg, J. (1999). Does music performance allude to locomotion? A model of final *ritardandi* derived from measurements of stopping runners. *Journal of the Acoustical Society of America, 105*(3), 1469–1484.

Frith, S. (1996). *Performance rites.* Oxford: Oxford University Press.

Gellrich, M. (1991). Concentration and tension. *British Journal of Music Education, 8,* 167–179.

Johnson, M. (1999). Something in the way she moves: Musical motion and musical space. http://www.hf.uio.no/CMI-99/.

Kozlowski, L. T., & Cutting, J. E. (1977). Recognising the sex of a walker from a dynamic point-light display. *Perception and Psychophysics, 21,* 575–580.

Laban, R. (1960). *The art of movement and dance* (2nd ed.). London: Macdonald & Evans.

Miller, J. (1996). *The good brass guide.* London: Guildhall School of Music and Drama.

Morgenstern, S. (1956). *Composers on music.* London: Faber & Faber.

Nakamura, T. (1987). The communication of dynamics between musicians and listeners through musical performance. *Perception and Psychophysics, 41,* 525–533.

Papoušek, M. (1996). Intuitive parenting: A hidden source of musical stimulation in infancy. In I. Deliège and J. A. Sloboda (Eds.), *Musical beginnings: Origins and development of musical competence* (pp. 88–112). Oxford: Oxford University Press.

Pierce, A. (1994). Developing Schenkerian hearing and performing. *Integral, 8,* 51–123.

Repp, B. H. (1993). Music as motion: A synopsis of Alexander Truslit (1938) *Gestaltung und Bewegung in der Musik. Psychology of Music, 21,* 48–72.

Shaffer, L. H. (1982). Rhythm and skill. *Psychological Review, 89,* 109–122.

Shaffer, L. H. (1984). Timing in solo and duet piano performances. *Quarterly Journal of Experimental Psychology, 36,* 577–595.

Stowell, R. (1985). *Violin technique and peformance practice in late eighteenth and early nineteenth centuries.* Cambridge: Cambridge University Press.

Todd, N. P. McA. (1985). A model of expressive timing in tonal music. *Music Perception, 3,* 33–58.

Todd, N. P. McA. (1992). The dynamics of dynamics: A model of musical expression. *Journal of the Acoustical Society of America, 91*(6), 3540–3550.

Todd, N. P. McA. (1995). The kinematics of musical expression. *Journal of the Acoustical Society of America, 97*(3), 1940–1949.

Trevarthen, C. (1999–2000). Musicality and the intrinsic motive pulse: Evidence from human psychobiology and infant communication. *Musicae Scientiae*, Special Issue: Rhythm, Musical Narrative, and Origins of Human Communication, 155–215.

# PART III

## INSTRUMENTS AND ENSEMBLES

# 16

# Solo Voice

GRAHAM F. WELCH & JOHAN SUNDBERG

One of the principal challenges for the singer is to acquire an understanding of how to develop and maintain a particular set of culturally specific musical behaviors using an instrument that is not visible and in which the functional components change physically across the lifespan. The singer's instrument has three components. The respiratory system is responsible for variations in loudness; changes in the pattern and frequency of vocal fold vibration are perceived as variations in pitch and voice quality; and changes in the configuration of the vocal tract are linked to resonance and carrying power. Although often interrelated (particularly in the untrained vocalist), these three functional characteristics are susceptible through education to focused development and conscious control.

The ability to use the voice to express music through singing is a characteristic of all known musical cultures. Each culture tends to value particular vocal timbres in the performance of its art music, such as the vocal sounds commonly associated with classical opera in the West, with the throat music of Tuva, or with blues. These vocal timbres are a selection from a much wider potential variety, indicating the inherent plasticity of the human voice. So one of the central issues in any consideration of solo voice performance is the development and maintenance of a particular set of culturally specific behaviors that use an instrument that is not visible and in which the functional components change physically across the life span. This chapter will examine the basic design of the vocal instrument (its anatomy, physiology, and psychological correlates), its function in singing, and the pedagogical implications for both teachers and performers.

## Structure and Function

The vocal instrument (Figure 16.1) consists of three essential components: the *respiratory system*, the *vocal folds* and the *vocal tract* (the cavity formed by the spaces above the larynx, namely, the pharynx and the mouth, which are sometimes complemented by the nasal cavity) (cf. Sundberg, 1996). Together, their combined action determines the characteristics of perceived vocal sound in speech and singing.

The respiratory system is the energy source for human voice production. This system compresses the lungs to provide an upward-flowing air stream that sets the *vocal folds* (situated within the larynx) into vibration. The vibrating vocal folds chop the air stream into a pulsating airflow (Figure 16.1; *transglottal airflow*) called the *voice source*. This consists of a complex of simultaneously sounding pure tones, or *partials* of different frequencies. Since the relationship between the partials in voiced sounds corresponds exactly to the harmonic series, the partials are also called harmonics and the lowest is called the fundamental. These partials decrease in strength with their ascending frequencies (see glottal source spectrum in the figure). However, the voice source is filtered by the spaces of the vocal tract above the vocal folds. Depending on its configuration, the vocal tract enhances or suppresses particular partials and so forms (shapes) the output sound (see the radiated spectrum in the figure). Consequently, these vocal tract resonances are called *formants* (see figure). The properties of the voice source, plus the frequencies of the formants, enable us to perceive and label particular output sounds as vowels. Formants also characterize consonants, both voiced (e.g., *d*) and unvoiced (e.g., *t*). The combined action of the vocal folds and the vocal tract also creates the distinctive personal timbre that we perceive in a voice.

### *The Vocal Tract*

At the lower end of the vocal tract is the *larynx*. The two vocal folds are located inside the larynx and are made up of muscles shaped as folds. Their horizontal and vertical dimensions change in relation mainly to pitch. Covered by a mucous membrane, the folds are approximately 3 mm long in newborn infants and grow at an average rate of 0.4 mm per year for females and 0.7 mm per year for males. With a disproportionate growth during puberty in males, adult female vocal folds are about 9 to 13 mm, compared to 15 to 20 mm in adult males (Titze, 1994). Vocal fold length and vibrating mass are crucial to the pitch range of the voice: the longer and thicker the vocal folds, the lower the pitch range. Consequently, because their vocal folds are smaller, young children have a higher vocal pitch range than adults. Although boys tend to have slightly longer vocal folds than girls, this difference appears not to be reflected in customary pitch ranges, as these are very similar. However, once puberty begins (at around the age of eight in some girls) and the relative physical dimensions increase, voice pitch lowers and diverges between male and female voices. This results in adult male voices' being approximately one fifth to one octave lower than those for adult

females, depending on the singer's voice classification. (For example, the average spoken pitch of a typical soprano is near $B_3$ (247 Hz), a semitone below middle C or $C_4$, but the average spoken pitch of a typical tenor is about a fifth lower at $E_3$ (165 Hz), a sixth below middle C: Titze, 1994, p. 188.) Vocal fold length does not significantly depend on body height but is more related to the circumference of the neck (a function of body build). At the limits of vocal pitch range, there is an overlap between men and women in their sung vocal pitch ranges, and this is greater between altos and tenors than sopranos and basses (Titze, 1994, p. 187; see "Voice Source" later).

The vocal folds run horizontally from front to back within the *thyroid cartilage*. They are attached together at the front to the inside of the thyroid cartilage (Adam's apple) and at the back to two mobile cartilages (the *arytenoids*). When one is breathing silently, the arytenoid cartilages are moved apart (abduction) by the *opener* set of muscles (*posterior cricoarytenoid* muscles), allowing the air stream to flow freely. The gap between the two folds is known as the glottis. When one is speaking or singing, voiced sounds are produced when the *closer* muscles (the lateral *cricoarytenoid* and *interarytenoid* muscles) bring the arytenoid cartilages together and hence draw the posterior ends of vocal folds together (adduction), so that their edges vibrate in the air stream. The folds are abducted for unvoiced sounds (such as certain consonants) and adducted for voiced sounds (such as vowels). The muscles responsible for opening and closing the glottis are mainly located at the back of the vocal folds near the spine. As well as assisting in basic vocal fold vibration, this set of opener and closer muscles is also involved when changing the loudness of the voice.

The air pressure in the lungs is the main tool for varying vocal loudness: the higher the pressure, the louder the voice. The underlying mechanism is that a greater air pressure throws the vocal folds apart with a greater force, with the result that they then also close more quickly. If the closer muscles are fully activated and the opener muscles are relaxed, the vocal folds are kept tightly closed together. This requires much more effort from the air stream to push them apart. Eventually, even when the folds are firmly adducted, at high lung pressures it is usually possible to get some air between the vocal folds. The effort, however, bursts the folds apart and then snaps them back together, creating a bigger disturbance of the air, which is perceived as a louder voice and *pressed* phonation (also termed *hyperadduction* or overly tight vocal fold adduction: Harris et al., 1998).

The laryngeal *lengthener* and *shortener* muscles are essentially responsible for pitch change. The shortener muscles (the *thyroarytenoids*) are located within the vocal folds themselves. They have a horizontal orientation (front to back). When these contract, the result is a fatter, bulkier, shorter vibrating tissue that produces a lower pitch. In contrast, the lengthener muscles (the *cricothyroids*) have more of a vertical orientation and are attached near the front of the larynx, stretched between the thyroid and cricoid cartilages. The effect of the contraction of the lengthener muscles on the cartilages is to tilt and slide the thyroid cartilage forward and so stretch (lengthen) the vocal folds. In so doing, the fundamental frequency is raised.

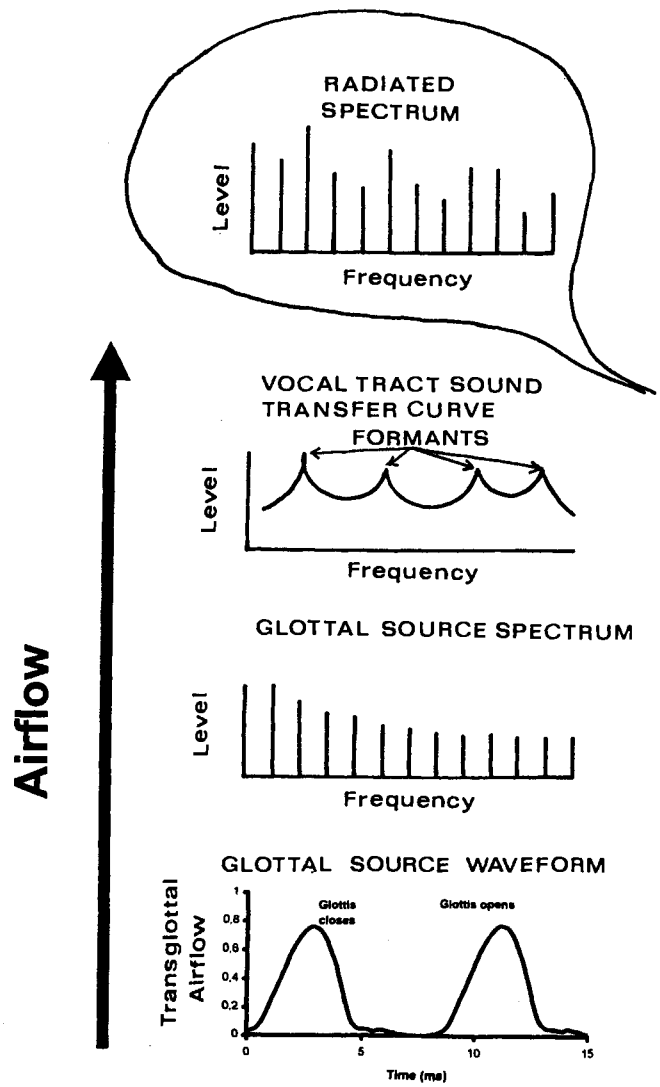
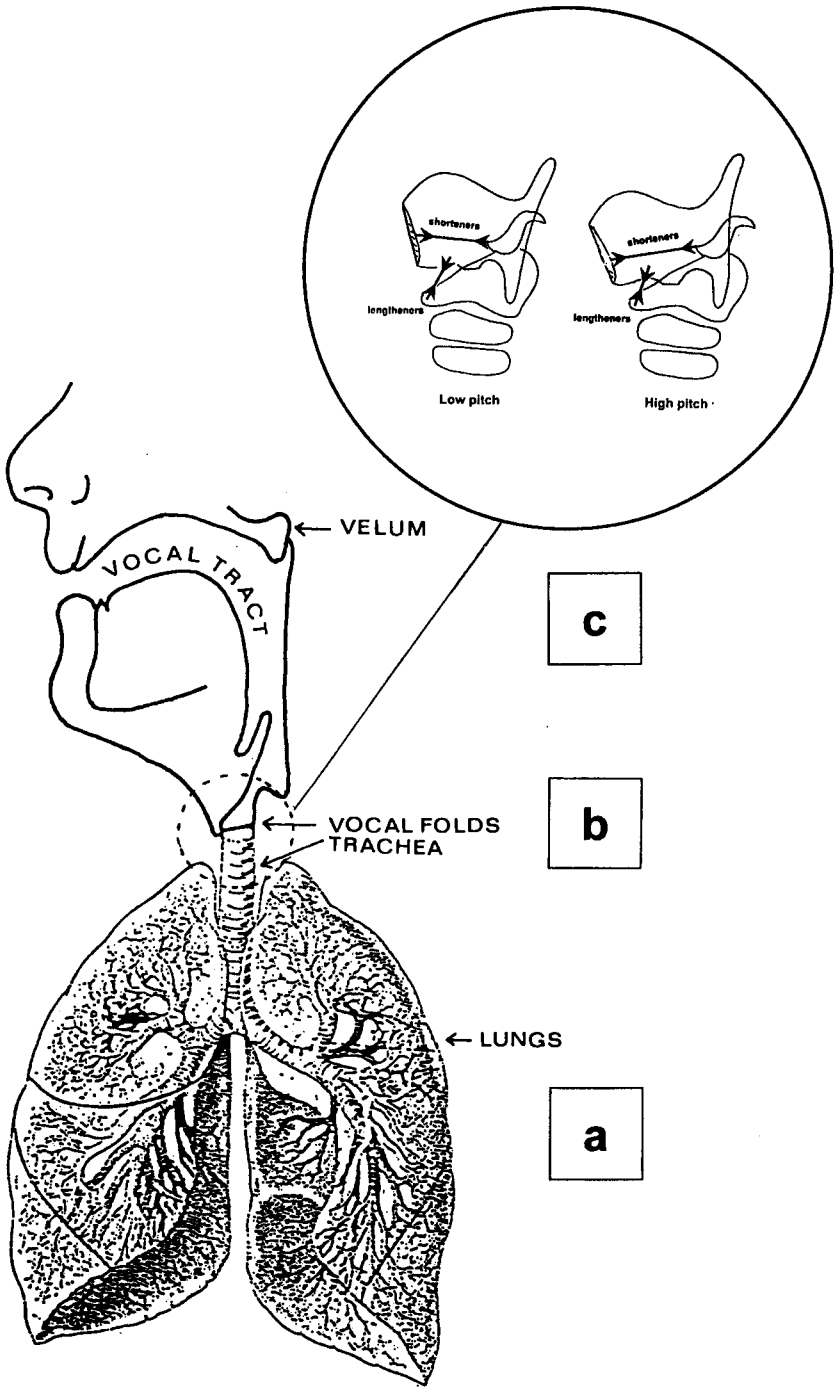



Figure 16.1. (a) The respiratory system (lungs), (b) the vocal folds, and (c) the vocal tract. (Adapted from Sundberg, 1987. Copyright 1987 University of Illinois Press, used by permission.)


shorteners
lengtheners
Low pitch
shorteners
lengtheners
High pitch
VELUM
VOCAL TRACT
c
VOCAL FOLDS
TRACHEA
b
LUNGS
a

Effective coordinated action of each of these shortener and lengthener muscles is required for skilled voicing. When a voice becomes louder, the activation of the closer muscles is increased, while the activation of the lengthener muscles is reduced. The reason is that the increased airflow creates the louder snapping together of the vibrating vocal folds and so the lengthener and shortener muscles have to work harder in combination to resist the pressure from the lungs, either to maintain or to change pitch as well as intensity (see "Voice Source" later and also Thurman & Welch, 2000, pp. 394–408).

A few millimeters above the vocal folds there is another pair of folds, also covered by mucous membrane. These are called the *false vocal folds* or the *ventricular folds*. Between the vocal folds and the ventricular folds there is a small cavity known as the *laryngeal ventricle* (sometimes termed *epilarynx*). This links into the much larger pharyngeal space above. The ceiling of the pharynx is the *soft palate* (*velum*), which also serves as the gateway to the nasal cavities. Narrow channels in the ceiling of the nasal cavity lead up to other cavities, the maxillary and frontal sinuses, which are located in the bone structure of the skull. The pharynx leads into the oral cavity. Both these cavities have their shapes altered by the movement of the tongue and the jaw. Thus the pulsating air stream flows through a tube that consists of a group of linked chambers to produce particular vocal timbres.

With regard to differences between the sexes, an adult female vocal tract is about 15% to 20% shorter than its adult male counterpart. In particular, the female laryngeal ventricle (the space between the vocal folds to the false folds) and pharynx are shorter (Story, Titze, & Hoffman, 1997) than those of the male, while the mouth length difference between the sexes is less marked (Nordström, 1977).

### *Breathing*

The lungs are elastic, spongy structures suspended in a vacuum in a sac within the rib cage. Because of the vacuum, their volume is enlarged. They contain a great number of small cavities linked by a system of air pipes, which join in the trachea, a tube that leads up to the vocal folds. Inhalation and exhalation correspond to decompression and compression of the lungs, which is performed by the respiratory system through forces generated by muscles, elasticity, and gravitation. Through muscle contraction, we can expand and contract the rib cage. As most of the lung surface is covered by the rib cage, small rib cage movements result in substantial compression or decompression. A large inspiratory muscle sheet, the diaphragm, constitutes the boundary between the rib cage and the abdomen. It originates along the lower edge of the rib cage and is vaulted into it. When contracting, it lowers the bottom of the rib cage by pushing the abdominal content downward and forward, which in turn expands the abdominal wall. Conversely, by contracting the abdominal wall muscles the abdominal content can be pushed back into the rib cage, thus compressing the lungs.

After a completely relaxed sigh, lung volume typically reaches the *functional residual capacity* (*FRC*), generally near 40% of maximum lung volume in an

upright position. In singing and speech alike, we usually phonate at lung volumes above FRC, whereas phonating below FRC can feel uncomfortable. When trained singers learn new songs, they learn how to inhale and sing in order to avoid lung volumes below FRC. Small lung volumes are generally used in speech, whereas in classical singing phrases are often initiated at very high lung volumes. Moreover, classical singers tend to repeat their breathing patterns very consistently when they sing the same phrase. They seem to use the rib cage mainly for changing lung volume, though some singers appear also to make use of the abdominal wall (Thomasson & Sundberg, 1999).

The overpressure of air in the trachea generates an air stream through the *glottis* (the space between the vocal folds). The pressure just below the glottis, the *subglottal pressure*, provides the driving force of the voice and is the main tool for controlling vocal loudness: the higher the pressure, the louder the voice (as mentioned earlier). This pressure variation in singing needs to be quite accurate, as pressure affects pitch; failure to match a target pressure may result in pitch errors (singing out of tune). A classical singing exercise that involves *staccato* performance (short tones interleaved with short silent intervals) of ascending and descending triad patterns seems particularly appropriate to train the singer's subglottal pressure control in coordination with the control of the laryngeal muscles that regulate the fundamental frequency. During the silent intervals the vocal folds part, leaving the airways open, so lung pressure must then be reduced to zero to avoid a waste of air. In addition, high tones need higher pressures than low tones, so each tone usually needs a different pressure. Singers also use subglottal pressure variations to demarcate the tones in rapid sequences, as in legato *coloratura* singing. Then each tone receives its own pressure pulse, so that one pressure pulse is produced for each tone in synchrony with the fundamental frequency pattern. In quick tempi, these pressure pulses may be as short as 150 milliseconds (ms).

### *Voice Source*

In voiced sounds, the voice source is a pulsating airflow (see waveform element in Figure 16.1). The voice source can be varied in several different ways with respect to fundamental frequency (perceived as pitch), amplitude of the spectrum partials (perceived as loudness), and dominance of the fundamental (perceived as type of voice production).

The perceived pitch corresponds to the frequency of vocal fold vibration. Thus when the pitch of $A_4$ is being sung, the vocal folds open and close the glottis 440 times per second ($A_4 = 440$ Hz). The main determinant of fundamental frequency is the length and vibrating mass of the vocal folds: the longer and thinner the folds, the higher the pitch (as mentioned earlier). These properties are controlled by the laryngeal musculature. The approximate ranges covered by the main singer classifications are as follows: bass singers, 80–330 Hz (pitch range $E_2$–$E_4$); tenors, 123–520 Hz ($C_3$–$C_5$); altos, 175–700 Hz ($F_3$–$F_5$); and sopranos, 260–1,300 Hz ($C_4$–$E_6$). Fundamental frequency is also influenced by subglottal pressure. So, to stay in tune while producing a *crescendo*, a singer needs to reduce the activation of

the laryngeal muscles that regulate fundamental frequency as the subglottal pressure increases. This may partly explain why less-skilled singers find it difficult to sing high tones softly. Their laryngeal muscles are not yet skillfully coordinated enough to ensure that the vocal fold tension matches the low subglottal pressure. Instead, they rely on increased subglottal pressure to raise pitch, resulting in loud and sometimes mistuned singing. As with all human control systems, the key to success is an awareness of the target in relation to possible means of achieving that target.

The voice source is also influenced by the degree of glottal adduction. Weak adduction results in a leaky or breathy voice, as in a voiced whisper. Overly strong adduction produces pressed phonation, or a tense or strangled voice, as when we phonate while lifting a heavy burden. In between these extremes is a range of adduction variations, with classically trained singers tending to have more consistent adduction across their pitch range. This enables such singers to produce a more even vocal sound across wide pitch ranges, whether singing softly or loudly. Adolescent voices tend to be rather breathier, reflecting this period of instability in vocal coordination during voice change. Untrained females tend to have breathier voices than untrained males across the life span (Linville, 1996). This may be because the gap between the female arytenoids cartilages is generally wider and their closure requires a larger gesture (cf. Södersten, 1994) or because of some underlying gender patterning of the vocal mechanism during childhood (Welch & Howard, in press) or both.

Another type of variation in the voice source corresponds to different patterns of vocal fold oscillation. These patterns are called *vocal registers*. There are at least three registers, generally referred to as *vocal fry*, *chest* (or modal), and *falsetto*. The modal register is sometimes further subdivided in singing into lower and higher subregisters (also termed as *chest* and *head* by singers because of the physical sensations involved) to create four registers.

As stated earlier, greater shortener muscle activity increases the mass of the vocal folds and lowers vocal pitch, while greater lengthener muscle activity reduces the mass of the vocal folds and raises vocal pitch. In vocal fry (pulse register), which frequently occurs at phrase endings in conversational speech, the vocal folds are thick and lax and vibrate asymmetrically; as a result, fundamental frequency is low, generally well below 100 Hz for male voices. By contrast, in the two components of the modal register (lower and upper) the folds are less lax and vibrate symmetrically. If there is a predominance of the shortener laryngeal muscles, there is more vertical tissue mass, creating a larger bottom-to-top contact area and more horizontal depth in the oscillating vocal fold tissues. The vocal folds tend to remain in contact with each other for longer in each cycle (Titze, 1994). The resultant tone sounds richer and full-bodied (often termed *chest voice*). However, if there is greater lengthener activity, the vocal folds are longer and thinner and vocal fold contact is confined only to their superior (upper) portion and outer layers. This reduced vertical mass, smaller bottom-to-top contact area, and less horizontal depth produces a vocal quality that generally sounds lighter and thinner (also termed *head voice*). In falsetto, the folds are thin and stretched and rarely close the glottis completely, unless the voice has been trained

in this register (such as that of the countertenor). The shortener muscles are assumed to relax completely, so that vocal fold length is determined by the action of the lengtheners. The vertical tissue mass and horizontal depth can still be varied such that, in a trained falsettist, it is still possible to produce a sung pitch range of approximately two octaves ($E_3$ to $E_5$).

Singers are trained to change registers gradually, so that no salient timbre differences occur, unless specifically chosen to add particular voice coloring in their performance. As much of singing principally involves the two elements of the modal register (lower and higher), register transitions (*passaggi*) may best be rehearsed by first practicing *descending* rather than ascending pitch patterns and utilizing quieter phonation. This reduces the potential overinvolvement of subglottal pressure that can occur when singing ascending pitch patterns. Descending pitches require a gradual increase in tension of the shorteners within the vocal folds and a concomitant relaxing of the lengtheners.

A voice source characteristic of the trained singing voice is *vibrato*. There seem to be at least two types of vibrato, which are produced and sound differently (Dejonckere, Hirano, & Sundberg, 1995; Titze, 1994, p. 290). The vibrato used in classical singing is normally produced by pulsations in the pitch-raising lengthener muscle (cricothyroid) and is a *frequency vibrato*. In various types of popular music, a different kind of vibrato is often used, which is generated by variations in subglottal pressure imposing an undulation of voice source amplitude. It corresponds to an amplitude modulation of the voice source and is really an *intensity vibrato* (Sundberg, 1987). Frequency vibrato corresponds to a slow, nearly sinusoidal fundamental frequency modulation. Typically, the rate is somewhere between 5.5 and 7 undulations per second, with the extent of the modulation varying from ±50 to as much as ±150 cents (where 100 cents = 1 semitone). Both rate and extent are important. If the rate goes below 5 undulations per second, which may occur in older or strained voices, no clear pitch can be perceived; it sounds as if the pitch is swaying around. If it is faster than 7 undulations per second, the tone sounds nervous. Also, the acceptable range of frequency modulation is limited. If the extent surpasses about ±100 cents, it tends to sound exaggerated. In choral and in pop singing, vibrato typically has a much smaller extent than in classical singing.

### *Formants*

As mentioned at the beginning of this chapter, the spaces above the vocal folds are a series of connected resonating chambers that filter the sounds that emanate from the voice source. Voiced sounds are acoustically rich, having many harmonics above the fundamental frequency. The tract amplifies certain component harmonics but dampens others. Formants are vocal tract resonances that appear at certain frequencies, the *formant frequencies*. Because partials of various frequencies are transmitted through the vocal tract simultaneously, those that coincide with the formant frequencies are radiated from the lip opening with a greater strength than others. Therefore, formants appear as peaks in the spectrum of the radiated sound (see Figure 16.1).

There are generally five formants relevant to singing, and they are crucial in our perception and discrimination of voiced sound. In speech communication our perception and labeling of certain sounds as vowels depends on the relationship between the lowest two formants. The frequency of the first formant is particularly sensitive to the size and shape of the pharynx and mouth cavity, which is largely influenced by the width of the jaw opening. Likewise, the second formant is particularly sensitive to the shape of the tongue body. If, for example, the first and second formants are located at 600 Hz (around $D\#_5$) and 1,000 Hz ($C_6$: soprano top C) respectively, the perceived vowel quality is /a:/ (as in *hard*); if they are located at 300 Hz ($D\#_4$) and 2,000 Hz ($C_7$) the perceived vowel quality is /i:/ (as in *heed*). The formant frequencies depend on vocal tract length and shape, which are controlled by the positioning of the lips, jaw, larynx, velum (soft palate), tongue, and pharyngeal side walls, that is, by articulation. In speech, the coordination of the vocal tract is not conscious but is a product of basic anatomy and physiology interacting through socialization with a particular sociocultural linguistic soundscape. Formant frequencies are decisive for vocal timbre. As well as determining perceived vowel quality, formants are important (alongside temporal patterns) for the perception of personal *voice quality*, permitting us to identify individual voices, such as those of favorite singers on a CD recording or family members on the telephone.

There is a close relationship between vocal tract shape and formant constellation. For example, a shortening of the tract, caused by retracting the corners of the mouth or by raising the larynx, increases all formant frequencies more or less. Both a narrowing of the pharynx and a widening of the mouth cavity increase the frequency of the first formant. This will make the sound brighter. Conversely, expanding the pharynx, protruding the lips, and lowering the larynx will lower the first two formants and make the sound darker in timbre. Moreover, if the pharynx is lengthened by lowering the larynx, the second formant is also lowered in vowels produced by a forward position of the tongue, such as /i:/ (as in *heed*). However, vocal tract length (and hence the formant frequencies of a given vowel) varies somewhat both among and within men, women, and children. Such differences explain some of the voice timbre variations among individuals. For example, it has been shown that for a given vowel tenors tend to sing with higher formant frequencies than basses (Sundberg, 1987, p. 110).

In trained classical singers, a characteristic feature of bass, baritone, tenor, countertenor, and alto voices is an exceptionally high energy peak, normally occurring somewhere between about 2.5 and 3 kHz (i.e., over one octave above $C_6$). This peak has been called the *singer's formant* (Sundberg, 1974, 1987). The energy peak allows the trained singer's voice to be heard above the sound of a full orchestra in a large concert hall. It appears in all voiced sounds and is normally a consequence of a clustering of the third, fourth, and fifth formant frequencies in a frequency range where the ear is particularly sensitive. This effect can be produced without excessive vocal effort, even when the orchestra is loud. This distinctive ring to the voice, and the source of its carrying power, is generated by a particular configuration of the vocal tract; a wide pharynx is often

combined with a lowered larynx (as compared to the rest position). The amplitude of the singer's formant varies among vowels and also with vocal loudness, since vocal loudness affects the slope of the voice source spectrum, as mentioned earlier.

The center frequency of the singer's formant seems significant to voice quality. In bass voices, the center frequency is near 2.4 kHz, in baritone voices near 2.6 kHz, in tenor voices near 2.8 kHz, and in alto voices (female) near 3.0 kHz. These variations can be assumed to reflect differences in vocal tract length and pharyngeal shape. In addition, the wide pharynx needed for clustering the higher formants to create the singer's formant also brings changes to the first two formants. For example, the second formant in front vowels such as /i/ is considerably lower in classical singing than in neutral speech.

Another phenomenon of formants relates to the perception of vowels in singing. The frequency of the first formant varies between about 250 and 1,000 Hz (roughly $C_4$–$C_6$) depending on the vowel. So many of the sung pitches produced by tenors, altos, and sopranos can have their fundamental frequency above the typical value of the first formant for a particular vowel. In order to remain (or attempt to remain) intelligible, singers tend to adjust the configuration of the vocal tract in order to raise the value of the first formant under these conditions. In the vowel /a/, this is achieved by a widening of the jaw opening. In other vowels, the lip opening and/or the tongue constriction of the vocal tract is first widened up to a certain point, and after this possibility has been exhausted the jaw opening is widened.

Once again, singers do not deliberately set out to tune the formants per se; rather, they learn (and self-monitor) a set of vocal tract coordinations that produces the desired perceptual effect, namely, a range of vocal pitches and vowels that sound cohesive (cf. Miller & Doing, 1998). In singing pedagogy, for example, much use has been made of the vowel /u:/ (as in *too*) in training both child and adult aspiring soloists. In normal phonation of this vowel, the third formant is low, so formant clustering does not normally occur. However, a rearticulation of the tongue shape, with an anterior placement of the tip of the tongue, filling out the cavity behind the lower incisors, will raise this formant and increase the perceptibility of a singer's formant.

At high pitches in the soprano range, all vowels share approximately the same formant frequencies. If the first formant is near that of the fundamental frequency (corresponding to the perceived pitch), it has the effect of making the fundamental much louder. For classical sopranos singing at pitches above 700 Hz ($F\#_5$, top of the staff), for example, such loudness increases come with no increase in vocal effort. However, since vowel quality is determined mainly by the two lowest formant frequencies, vowel intelligibility at high pitches becomes problematic. Faced with a choice between inaudible tones with normal vowel quality and audible tones with modified vowel quality, singers generally choose the latter. The intelligibility of very high singing relies almost exclusively on consonants, not vowels.

Thus formants are the principal resonances of the vocal tract. Other resonances occur also in the subglottal airways and the nasal sinus system. From the

audience's point of view, there appears to be little significant contribution to perceptual voice quality in classical singers from either the chest region (except at low pitches for basses) or the nasal cavities and sinuses. Other vocal genres, such as certain folk singing, may make greater use of nasal resonance. For the singers themselves, any kinesthetic sensations during singing from these bodily parts may assist in their self-monitoring of how they feel the voice is placed during performance.

*Choral singers* show less evidence of the singer's formant. This probably reflects the need to ensure that individual voices do not stand out in the creation of a choral blend (see also chapter 17). Similarly, *country* singing is more similar to normal speech than classical singing. Thus when one is singing loudly at higher pitches, there is often greater vocal fold adduction. Likewise, vocally untrained and inexperienced singers tend to use habitual speech coordinations when they sing. With increasing pitch, such singers tend to raise their larynxes and change their tone production toward pressed voice. This adds an edgy quality to the voice, corresponding to a weak fundamental and loud high partials, which is a characteristic also of some *folk music* styles.

Another popular singing style is the *belt voice* (Estill, 1988). This is a term that was used to describe singing in the American music theater in the 1940s and 1950s by Ethel Merman but is now also found outside the theater in the performance of other musical genres, such as gospel, spirituals, and jazz. Some popular singers (such as Whitney Houston, Gloria Estafan, Alanis Morissette, and Anita Baker) exploit this quality to create a particular emotional impact at certain moments in a song. The sound is loud and the vowel quality is more similar to that used in very loud speech than in operatic singing. It is produced with a narrow pharynx, a raised larynx, and high lung pressures.

One particularly striking use of formants in a musical genre is found in the throat singing (overtone singing) of Tuva (or Tyva), an autonomous republic within Russia on its border with Mongolia. This type of singing involves producing two distinct pitches simultaneously. One is low and sustained (such as in a bagpipe) and corresponds to the fundamental, while the other is much higher, with a flute- or whistle-type quality, and corresponds to one of the harmonics. There are different styles of throat singing, but each involves a particular voice source behavior (keeping the vocal folds closed for longer in each vibratory cycle) and shaping of the vocal tract (manipulating the tongue tip and root and its midpoint to cluster the second and third formants, with a protrusion of the jaw and a narrowing and rounding of the lips and, in some cases, using the false folds to vibrate an octave below at half the rate of the vocal folds). The result is a low sustained fundamental and a finely tuned formant peak. Depending on the pitch of the fundamental, up to 12 harmonics can be isolated individually. Enhanced in succession, they can create a melody (Levin & Edgerton, 1999). Throat singing is not taught formally but learned in the same way as language. It can be mastered also by non-native singers. This type of singing is also found in Tibet, Japan, and South America (Klingholz, 1993), and elements from such musics have been incorporated into modern Western music performance (Schaefer, 2000).

## Performance and Pedagogy

Voice is an essential element of self-identity. It helps to define who we are and how other people experience us. It conveys our inner feeling states, both to ourselves and to others (Thurman & Welch, 2000). Emotion and mood are central characteristics of voice production and reception. This is because of the integrated networking of the body's nervous, endocrine, and immune systems. Whether we are feeling elated, relaxed, stressed, or threatened, each inner state is likely to be reflected in voice behavior and to be communicated to others and to ourselves (Welch, 2000b). Hence, the voice is exceedingly useful as a musical instrument.

Solo singers in training often have vocal habits that are potentially challenging for the voice teacher for five main reasons:

- Much voice behavior is not conscious. It is habitual and not readily accessible to conscious processes. It is also typically influenced (as well as being a reinforcement) by how we feel at the moment of voice use (Thurman & Welch, 2000).
- We cannot physically see the sound source (vocal folds) or tongue shape; we have to rely on our proprioception (from muscle receptors) and auditory feedback to sense what our voice is doing.
- Highly skilled musical performance requires consistent musical behaviors. The moment-to-moment changes in articulation that are characteristic of normal conversation are often inappropriate and inimical to culturally derived classical music performance (Sundberg, 2000, p. 235).
- Vocal music is a highly complex sociocultural artifact. The words and music each present separate challenges to the performer, as does their combination (Welch, 2000a). The text is often a narrative that has its own internal rules, whereas the music makes physical production demands on the vocal system that might be contrary to those expected in normal speech articulation (Welch, 1985a).
- The grouping of tones in musical motifs and phrases is an important aspect of music performance (Gabrielsson, 1999). Changes in voice timbre are often used for marking boundaries in musical structures. This makes the elimination of timbral differences between different pitch regions in a voice an essential aspect of vocal pedagogy.

One of the first tasks for the singing teacher is to bring elements of voice production to conscious awareness, such as an awareness of the underlying anatomical structures and their function in skilled voice production. Major professional associations for singing teachers (such as the National Association of Teachers of Singing in the United States and the Association of Teachers of Singing in the United Kingdom) now include scientific articles in their journals (*Journal of Singing* and *Singing*) in order to assist singing teachers and performers to understand better the data from voice science.

For many centuries, one pedagogical strategy has been to utilize small musical building blocks (*solfeggi* patterns of pitches, such as scales, motifs, arpeggios). Such simple musical exercises allow the singer to focus on and memorize isolated or simple combinations of elements from the dominant musical culture. Solfeggi may

also help the singer to work toward a greater evenness of vocal timbre across the pitch range. In untrained voices, vocal timbre tends to automatically change with pitch. However, through the use of such exercises, a greater evenness of vocal timbre is likely, particularly if (1) consonants are reduced in time (although consonant duration is varied for expressive purposes in actual singing performance) and (2) vowels are modified from typical habitual speech patterns.

Although imagery plays an important traditional role in voice pedagogy, it has rarely been researched systematically. Callaghan's (1997) survey of Australian teachers of singing reported three different types of imagery evidenced in practice (visual, kinesthetic, and aural). However, such practices were based on teachers' craft knowledge rather than empirical data. The experiential evidence of the efficacy of imagery continues to need corroboration in terms of well-controlled formal experimentation.

The scientific study of vocal production in singing is a relatively modern phenomenon. The human voice's invisibility has required the application of new technologies to reveal its underlying activity. For example, MRI (Story, Titze, & Hoffman, 1996), X-ray photography, and high-speed imaging are just some of the more important techniques that have been adapted from clinical studies for the analysis of singing.

The basis for performance is a variety of practice and rehearsal strategies that allow the singer to shape the perceptual and physical systems toward the identified musical goal (cf. Welch, 1985b). The practice of constituent elements in isolation can facilitate conscious awareness and control of basic coordination, as well as permitting the conscious overlaying of musical stylistics according to the particular genre and narrative context. Furthermore, reaction time studies indicate that we need approximately 300 ms to make a conscious adjustment to our motor behavior (cf. Schmidt, 1975, p. 137; Welch, 1985a), yet this is rather slow in relation to the requirements of vocal musical performance (such as for rapid scalic passages or arpeggios).

Given the complexities of the voice production in (classical) singing, the implications of a sound knowledge of vocal structure and function for vocal pedagogy are that

- Feedback from the teacher should be related directly to shaping conscious awareness, such as in understanding the relationship between tongue shape and formant tuning.
- The more complex the vocal task, the greater the opportunity for the student to misinterpret teacher feedback (Welch, 1985a, 1985b). Simpler vocal tasks, therefore, are more accessible to focused change.
- The biological aging process changes structure and function, not withstanding the relative stability of voice production across four decades (ages 20 to 60: Titze, 1994, p. 182). For example, the laryngeal cartilages begin to ossify (become more bonelike) in the third decade of life, suggesting that younger voices are more suited to agile musical requirements, whereas older voices are more able to sustain vocally strenuous performance. Therefore, different musical demands are appropriate for different age groups.

In summary, successful solo singing across the life span is more likely if both teachers and performers have developed habitual singing behaviors that are underpinned by an understanding of their physical and psychological reality.

### *References*


Callaghan, J. (1997). *Voice science and singing teaching in Australia.* Unpublished doctoral thesis, University of Western Sydney, Australia.

Dejonckere, P. H., Hirano, M., & Sundberg, J. (Eds.) (1995). *Vibrato.* London: Singular Press.

Estill, J. (1988). Belting and classic voice quality: Some physiological differences. *Medical Problems of Performing Artists, 3,* 37–43.

Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 501–602). London: Academic Press.

Harris, T., Harris, S., Rubin, J. S., & Howard, D. M. (1998). *The voice clinic handbook.* London: Whurr.

Klingholz, F. (1993). Overtone singing: Productive mechanisms and acoustic data. *Journal of Voice, 7*(2), 118–122.

Levin, T. C., & Edgerton, M. E. (1999). The throat singers of Tuva. *Scientific American, 281*(3), 70–77.

Linville, S. E. (1996). The sound of senescence. *Journal of Voice, 10*(2), 190–200.

Miller, D., & Doing, J. (1998). Male *passaggio* and the upper extension in the light of visual feedback. *Journal of Singing, 54*(1), 3–14.

Nordström, P.-E. (1977). Female and infant vocal tracts simulated from male area functions. *Journal of Phonetics, 5,* 81–92.

Schaefer, J. (2000). "Songlines": Vocal traditions in world music. In J. Potter (Ed.), *The Cambridge companion to singing* (pp. 9–27). Cambridge: Cambridge University Press.

Schmidt, R. A. (1975). *Motor skills.* New York: Harper & Row.

Södersten, M. (1994). *Vocal fold closure during phonation—physiological, perceptual, and acoustic studies.* Doctoral dissertation, Department of Logopedics and Phoniatrics, Huddinge University Hospital, Karolinska Institutet, Stockholm.

Story, B. H., Titze, I. R., & Hoffman, E. A. (1996). Vocal tract area functions from magnetic resonance imaging. *Journal of the Acoustical Society of America, 100,* 537–554.

Story, B. H., Titze, I. R., & Hoffman, E. A. (1997). Volumetric image-based comparison of male and female vocal tract shapes. *National Center for Voice and Speech Status and Progress Report, 11,* 153–161.

Sundberg, J. (1974). Articulatory interpretation of the "singing formant." *Journal of the Acoustical Society of America, 55,* 838–844.

Sundberg, J. (1987). *The science of the singing voice.* DeKalb, IL: Northern Illinois University Press.

Sundberg, J. (1996). The human voice. In R. Greger and U. Windhorst (Eds.), *Comprehensive human physiology,* Vol. 1 (pp. 1095–1104). Berlin: Springer.

Sundberg, J. (2000). Where does the sound come from? In J. Potter (Ed.), *The Cambridge companion to singing* (pp. 231–247). Cambridge: Cambridge University Press.

Thomasson, M., & Sundberg, J. (1999). Consistency of phonatory breathing patterns in professional operatic singers. *Journal of Voice, 13*(4), 529–541.

Thurman, L., & Welch, G. F. (2000). *Bodymind and voice: Foundations of voice education* (2nd ed.). Iowa City: National Center for Voice and Speech.

Titze, I. R. (1994). *Principles of voice production*. Englewood Cliffs, NJ: Prentice-Hall.

Welch, G. F. (1985a). A schema theory of how children learn to sing in-tune. *Psychology of Music, 13*(1), 3–18.

Welch, G. F. (1985b). Variability of practice and knowledge of results as factors in learning to sing in-tune. *Bulletin of the Council for Research in Music Education, 85*, 238–247.

Welch, G. F. (2000a). Singing development in early childhood: The effects of culture and education on the realisation of potential. In P. J. White. (Ed.), *Child voice* (pp. 27–44). Stockholm: Royal Institute of Technology.

Welch, G. F. (2000b). Voice management. In A. Thody, B. Gray, D. Bowden, and G. F. Welch, *The teacher's survival guide* (pp. 45–60). London: Continuum.

Welch, G. F., & Howard, D. M. (in press). Gendered voice in the cathedral choir. *Psychology of Music*.

# 17

# Choir

STEN TERNSTRÖM & DUANE RICHARD KARNA

Choir singers and directors frequently find themselves grappling with acoustical issues that appear to affect their ability to perform well. Hearing one's own voice, for example, is crucial. It improves with increased singer spacing and depends also on the room acoustics. Singer preferences are diverse but on average one's own voice needs to be about 6 dB stronger than the rest of the choir. In most rooms this implies a fairly spread-out formation. Precise intonation may be jeopardized by pitch perception discrepancies and by mechanisms inherent to voice control but it can also be facilitated by acoustically informed measures such as articulatory enhancement of common partials. Researching such issues necessitates decomposition while remaining aware of the true complexity of the situation: that of many people singing together and hearing each other in a room.

Many people who choose to perform music as a satisfying pastime do so as choir singers. Singing in unison or in harmony with other people is a low-cost, enjoyable activity, which at the beginner's level demands little more than one's time and a certain dedication to the task.

During the past couple of decades, research into the acoustics of choir singing has uncovered numerous interesting effects that are of potential relevance to choral performance. Some concern voice production, others have to do with the acoustics of the stage and the auditorium, while still others are related to our sense of hearing. Many of them are rather subtle and, if taken in isolation, of minor importance. When the acoustical circumstances combine constructively, however, choral singing is certain to become easier; conversely, when they combine destructively, choral performance is likely to suffer. For this chapter, we have chosen to describe in detail some phenomena that we expect and hope to be particularly interesting and useful to choral directors and singers who are curious about acoustics.

### *Fundamentals of Room Acoustics*

We note first that choirs are particularly dependent on the acoustics of the room, a subject that is not covered by other chapters in this volume.

Sound waves are a form of energy. One adult singer emits about three milliwatts (mW) of acoustic power in *forte*. Most of the energy emitted by the singer will bounce around in the room as reverberation until it is absorbed (into heat) by the walls and furnishings. The greater the absorption, the quicker the sound energy disappears, and the shorter the reverberation time will be. A minuscule proportion of the radiated sound energy eventually reaches the eardrums of listeners. The sound that reaches the listener from the source has three components: the direct sound, the early reflections, and the diffuse field (or reverberation). The *direct sound* travels in a straight line from source to receiver; it arrives first and reveals the distance and the direction to the source. The nearer the source, the stronger the direct sound. *Early reflections* are sounds that arrive at the listener after only one or two bounces against nearby walls. At delays less than about 40 ms we do not perceive these as separate echoes, but they are still very important for our impression of the acoustical liveness of a room. The *diffuse field* is the sum of the thousands of all later reflections that rapidly die down as a small amount of energy is lost into the wall at each bounce. The *reverberation time* is defined as the time it takes for the diffuse field to become 60 decibels (dB) weaker once the sound source has been turned off. After that time, the reverberation is essentially inaudible.

The diffuse field is so called because the many reflections tend to merge into a practically uniform sound field with no direction of its own. For a given source, the intensity of the diffuse field is much the same throughout the room. (Analogy: when you have been swimming in a pool for a while, the slosh of waves won't tell you where in the pool you have been.) There exists a distance from the source such that the intensities of the direct sound and the diffuse field are equal; this is called the *reverberation radius* of the room. If we move out farther from the source, the diffuse field dominates and the total sound intensity does not diminish appreciably with distance. If instead we approach the source, the direct sound will dominate and become louder the closer we are to the source. Considering the situation inside the choir, in rooms typical for choir performance only the one or two nearest circles of neighbors in the choir are within one reverberation radius of each other; the rest of the choir is outside.

### *Physical Versus Perceptual Sound Properties*

The loudness, pitch, and timbre of a sound are *perceptual* entities that can be measured only by asking listeners what they hear. Fortunately, these three entities correspond closely—but not exactly—to three measurable *physical* entities of sound: pressure level (SPL), fundamental frequency ($F_0$), and distribution of energy over frequency. The physical properties are the main entry points of the

acoustician, who can measure them, study their evolution over time and space, and subject them to statistical analysis.

All the physical parameters of a tone can to some degree influence any or all perceptual parameters (Fletcher, 1934). The study of these dependencies is the basis of psychoacoustics. For example, a tone that has a strong fundamental partial but weak overtones is perceived as dull. When presented at very different loudness levels, such a tone may also appear to change somewhat in pitch, even though its fundamental frequency remains the same. This so-called pitch-amplitude effect is not generally known to musicians, but it may influence choral intonation in certain circumstances.

### *Sound Properties and Mechanisms as an Organizing Principle*

For any research project on choir acoustics it is convenient to constrain the issue to one or two of the measurable *properties*, usually adopting a perspective either of voice production, of room acoustics, or of hearing, which are all *mechanisms*. It must be stressed that such constraints are for practicality only, because the reality of choral performance is too complex to be studied all at once. In Table 17.1 we list some examples of topics related to choir acoustics, categorized along these lines.

The intricacy of possible topics means that there is no obvious, systematic path for us to follow in order to cover the subject of choir acoustics in this compact overview. Instead, we will state some acoustical problems commonly encountered by choirs and point to ways of understanding and negotiating these problems. In each case we will first list the relevant acoustical issues and provide some practical recommendations. This will be followed by an in-depth background discussion of the problem, with explanations of the recommendations and references to scientific research.

## Balance of Loudness

*Choral singers sometimes complain that they cannot hear their own voices. Why might this be, and what can be done about it?*

### *Relevant Factors*

1. The physical spacing between singers
2. The amount of reverberation in the room
3. The presence or absence of reflectors close to each singer
4. The vocal power of the complaining singers relative to that of their neighbors

### *Recommendations*

1. Increase the spacing between singers, and/or have the choir stand in fewer rows. If this is not possible, mix the sections.

Table 17.1. A tentative categorization of topics in choir acoustics. Topics discussed in this chapter are indicated in italics.

| | Loudness/ Balance | Intonation | Timbre | Multivoice Issues |
|---|---|---|---|---|
| Voice production | Diversity of singers' vocal power; directivity of the human voice | Breath support, vibrato in choirs; *articulatory perturbation of pitch* | Acoustic differences between S-A-T-B, between trained voice and untrained, singer's formant or not | *The role of flutter in ensemble sound*; placement of singers within the choir; does uniformity of vowels matter? |
| Room acoustics | Effects of choir size, formation, *singer spacing and room absorption on the self-to-other ratio (SOR)*; distances; risers, shells, and reflectors | Effect of the SOR on intonation ability; can the pitch of the reverberation seem to go flat? | Effects of room response on voice usage; effects of different types of absorbents; wall proximity for basses; where to hold the music folder | Choir placement relative to audience; trade-off between room reverberation and the number of singers; ensemble timing |
| Perception | *Reflection, occlusion and bone conduction of one's own voice; diversity of singer preferences as to the SOR* | The *pitch/ amplitude effect* as an intonation hazard; *the pitch can differ slightly in the left and right ears* | *Pitch is more salient with certain timbres; role of common partials*; why some voices fit together and not others | Blending ability of different voice types; differences between choirs and vocal groups; *combination tones and beats*; how close is unison? |
| Sound technology | Choir with instrumental combo; public-address systems; self-monitoring for the choir | *What devices are suitable for giving the starting pitch?* | Faithful reproduction or enhancement? | Placement of microphones and loudspeakers |

2. Decrease the amount of reverberation in the room by introducing absorbents, closing curtains, admitting the audience, or opening several windows.
3. Encourage the singers to experiment with the sheet music and folder as a small personal reflector. There is a particular angle that noticeably improves the perception of one's own voice. Remove any shells or other large reflecting surfaces at close range, because these tend to exacerbate the problem.
4. Place the quieter voices near the ends of the choir and the louder ones near the center.

## *Scientific Background*

**Self and Other.** Inside a choir we have a complex transmitters-receivers situation with many directional sound sources, all of whom are also listeners (Marshall & Meyer, 1985). Taking the perspective of an individual singer in the choir, we will refer to the sound of one's own voice as *Self* and to the sound of the rest of the choir as *Other*. The Self sound is dominated by *airborne* sound coming out through the mouth, but it also has a component of *bone-conducted* sound. The relative contribution of the bone-conducted sound is hard, however, to measure objectively (von Békésy, 1949). For example, it can vary dramatically with vowels: block your ears and sing to compare /u/ with /a/! An open mouth radiates more sound, leaving less energy to be coupled through the body's tissues.

The airborne part of Self is entirely dominated by the direct sound from the mouth. The room reflections of Self are typically 15 to 20 dBs weaker than the direct Self. This means that inside any choir they will be drowned out (*masked*) by Other.

**The Self-to-Other Ratio.** A choir singer hears the Self sound together with the Other sound; ideally these should both be discernible in some optimal proportion. The *Self-to-Other Ratio* (*SOR*) can be measured on location for any particular singer and can be expressed in decibels. When the ears of a singer are subjected to the sounds of Self and Other simultaneously, Self will usually be a certain number of decibels stronger, in which case the SOR is said to be positive. Clearly, the SOR is a predictor of how well a singer will hear his or her own voice. Typical SOR values in live choral performance range from +1 to +8 dB (Ternström, 1994).

The major factor to influence the SOR is the intersinger spacing (item 1 above). The farther apart the singers are standing (especially from their own section colleagues), the greater the SOR, and the easier it will be for them to hear their own voices–and, conversely, the harder it will be for them to hear the others. Daugherty (1996) found that both singers and auditors preferred the choral sound produced in spread formations over that produced in close formation.

The SOR is also inversely related to the amount of reverberation in the room: the less reverberation, the weaker the total sound will be, and so the easier it will be to make out the sound of one's own voice (B). For a given choir placement in a room with given physical properties, the SOR can be predicted with some accuracy. This is important because it gives the architect's acoustics consultant a tangible goal: to design for a SOR onstage in the vicinity of +6 dB (Ternström, 1999).

**Singer Preferences.** In a normal situation, the individual choir singer cannot do very much to change his or her hearing-of-self, other than to use the music folder as a personal reflector to increase SOR. The permissible loudness of one's own voice is constrained by the music's dynamics. Vocal artists who perform with adjustable amplification, however, can be very fussy about the exact amount of monitor feedback they require. Ternström (1999) did an experiment with 23 experienced singers to determine (1) whether choir singers are particular about

their preferred SOR in a choral context and (2) how large the preferred SOR might be. The singers were indeed quite precise about their personal preferences: they reproduced their individual preferred SOR to within ±2 dB, which is a remarkably close tolerance. The overall average preferred SOR was +6 dB. However, tastes varied wildly, from +15 dB down to 0 dB. In other words, some of these singers wanted to hear their own voice a lot, while others preferred the rest of the choir to be just as loud. This large variation may be due to different habitual placement within the choir (see later), to individual voice timbres, and even to certain hearing losses so small as to be unknown by the singer.

**Loudness Issues.** The choir director should be aware that singer placement is not just a question of how well or poorly different voices sound together but also one of how comfortable each singer is with the balance of loudness of voices in his or her vicinity.

Having the choir stand in many rows will cause large differences in SOR between the center and the periphery of the choir, since some singers will have many neighbors and others will have few. In single-row formations, the SOR varies less with position in the choir. Furthermore, the SOR will be lowest in the bass section and highest in the soprano section when all sections are singing (Ternström, 1995). This is because sopranos are loudest: the human voice radiates acoustic power more efficiently at high fundamental frequencies. In terms of SPL, the Other sound will thus be dominated by sopranos, except in very close choir formations.

The ability to monitor one's own voice is crucial for accurate singing. As an example, have some volunteer singers carry small tape recorders to record their own voices (without headphones!) at very close range in various choral performance situations. Ask them to assess their own performances from the tape a few hours later. If the tape reveals unexpected errors, chances are that they could not hear themselves well enough.

Unison singing—all voices in synchrony on the same tones—is the worst case for monitoring one's own voice, since Other masks Self most efficiently when the frequencies overlap. If the SOR is too low, the singers may be able to hear their own voices only by making mistakes, and the scatter in intonation is likely to increase. In very large choirs, therefore, sectional formation is likely to be problematic.

## Intonation

*We are working on a piece that often tends to go out of tune at a particular point. Why might this be?*

### *Relevant Factors*

1. The *intrinsic pitch* of vowels and consonants is an acoustical and physiological phenomenon known from phonetics. In speech, some sounds tend to have a slightly raised or lowered $F_0$ because of the way the voice organ must be configured to produce them.

2. The *pitch-amplitude effect* is a perceptual phenomenon that makes it possible for the perceived pitch of a constant-frequency tone to change slightly with changes in loudness. Some people are more susceptible to this effect than others. The effect is most relevant to high and loud soprano singing.
3. *Breath support* has a direct bearing on intonation, particularly on long sustained tones in the upper pitch ranges. This is not really a choral issue but one of vocal training, although the problem is very common in choral singers.
4. *Complex modulations* between tonally distant keys is a music theory issue that many others have described, although little of predictive use has been reported; see any text on scales and temperament.

### *Recommendations*

1. Analyze the text for intrinsic pitch effects (see later). Temporarily change the lyric to something completely different and see whether the intonation is affected.
2. If some sopranos tend to sing sharp on high, loud tones, chances are that their ears are fooling them. In this situation, conductors sometimes accuse the singers of a poor sense of pitch; however, most of the time they are doing their best and have no way of knowing they are sharp. Instead, try to ascertain whether the pitch-amplitude effect is in fact the culprit. An individual test for the pitch-amplitude effect is suggested later.
3. Have the singers learn vocal technique. The term *breath support* refers to a singer's proficiency in controlling the subglottal pressure. The fundamental frequency of vocal fold vibration is regulated mainly by the subglottal pressure (coarse, slow control issued by many large respiratory muscles) (Titze, 1989) and by the pull of the rather small *cricothyroid* muscle (fine, rapid control). If the subglottal pressure is too low, as with insufficient breath support, then the cricothyroid muscle must exert a much larger force to achieve a high target pitch, and so it tires quickly (chapter 16).

### *Acoustical and Psychoacoustical Background*

In speech, there is rarely a need for precise intonation of sustained tones; rather, it is *gestures* or *inflections* in pitch that carry important information. Pitch in speech is also closely related to vocal effort. Increased subglottal pressure will normally raise the fundamental frequency, or $F_0$ (e.g., Titze, 1989). Hence, we expect in speech a strong covariation of loudness and pitch: the louder, the higher. Singers, however, must learn to control loudness and $F_0$ separately. This is a task that requires considerable coordination.

From an engineer's viewpoint, the voice organ is an unruly and entangled conglomerate of cartilage, muscle, and soft tissue. For such a device to produce sound at a stable $F_0$, it must be controlled by use of a host of auditory and neuromuscular mechanisms with compensatory action, or *servo systems*, that continuously monitor the voice output and apply corrective signals until the out-

put is perceived to match the target. Voice research has shown that the servo-system paradigm is in fact a good model for $F_0$ control (Larson et al., 2000; Ternström & Friberg, 1989). Since the voice organ is driven mainly by lung pressure, is steered by minute muscular control in the larynx, and is monitored by hearing and other types of receptors in the body, there seem to be few aspects of vocal function and near-field acoustics that could *not* have some bearing on intonation.

Because the intonation precision needed for choral harmony is very high, even small disruptions to pitch can be significant. Here we will discuss two of these: *intrinsic pitch*, which is a voice production phenomenon; and the *pitch-amplitude effect* that was mentioned earlier, which is a pitch perception phenomenon.

**Articulatory Perturbation of Pitch (Intrinsic Pitch).** In a voice science context, the term *articulation* refers to the action of adjusting and positioning the jaw, tongue, and lips to produce a given sequence of sounds. To phoneticians it is well known that in running speech some vowels tend to be produced with a slightly lower or higher $F_0$ than the average for the speech as a whole (Ladd & Silverman, 1984). Front vowels such as /i:/ and /e:/ (with the tongue held forward) tend to receive a slightly higher $F_0$, and back vowels such as /a:/ and /e:/ tend to be lower, on the order of a semitone or two. Phoneticians call this *intrinsic pitch*. Despite its name (*pitch* normally refers to perception or experience), this is not a perceptual effect but rather a vowel-dependent bias on $F_0$ that is somehow exerted by the articulation. Several biomechanical and acoustic hypotheses have been advanced to explain these deviations, but the matter is not yet closed.

In addition, the articulation of consonants sometimes involves movements of the larynx or changes in the air pressure differential that drives the vibration of the vocal folds. This can cause the $F_0$ of voiced consonants to differ from that of the adjacent vowels. For example, say very slowly *ahdahdah* and notice how the pitch drops during the *d* consonants. In fact, this intrinsic drop in pitch is a strong cue for us to interpret the sound as the consonant *d*.

In singing, intrinsic pitch effects would presumably be compensated for, as the singer is striving to reach a target frequency for each note. But what happens if it is hard to hear one's own voice, as is often the case in a choir? Ternström, Sundberg, and Colldén (1988) asked choir singers to sustain tones over a change of vowel and measured the average $F_0$ before and after the vowel transition. The subjects wore closed headphones in which they alternately heard their own voice fed back and a masking noise that completely drowned out their own voice. Each vowel pair was sung in noise and was then immediately repeated without noise. Intrinsic pitch effects were found in *both* conditions, but they were larger in noise, as expected. On the average, the largest $F_0$ changes were observed between /i:/ (high) and /ɛ:/ (low). For this case the average effect was about thirty cents, larger in noise and smaller with normal feedback. Such a pitch change is certainly large enough to make a chord go out of tune. Figure 17.1 summarizes the *occurrence* of pitch changes for the different vowel pairs in this experiment. The vowel /a:/ was not included, but it is known to have an even lower intrinsic pitch than /ɛ:/.

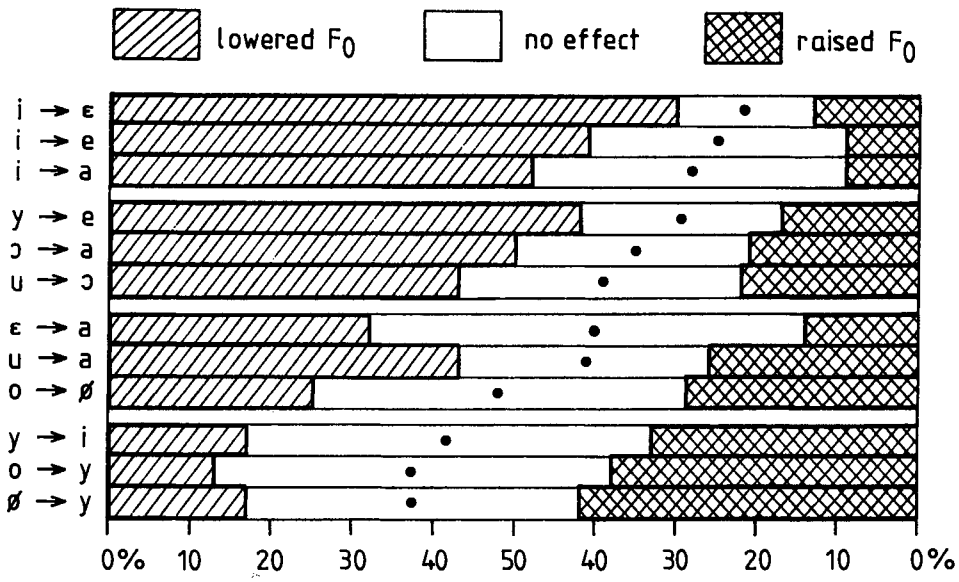



Figure 17.1. Articulatory perturbation of pitch over vowel changes. The bars show the relative occurrence of intrinsic pitch effects found when six subjects changed vowels on sustained tones (see text). The conditions with and without auditory feedback are pooled. For example, the top line is interpreted as follows: when the vowels /i:/ and /ɛ:/ were paired, /ɛ:/ was sung at a lower pitch in 70% of the cases and at a higher pitch in 13% of the cases, regardless of whether /i:/ or /ɛ:/ came first. (From Ternström, Sundberg & Colldén, 1988. Copyright American Speech-Language-Hearing Association, reprinted with permission.)

A practical consequence of this result is that intonation problems can sometimes be ascribed to a particular sequence of vowels and consonants (for example, *dies irae* or *kyrie eleison*) and that not hearing one's own voice will make things even worse.

**The Pitch-Amplitude Effect.** To test yourself for the pitch-amplitude effect, you will need a sound source that can produce a pure or nearly pure sine wave with stable loudness and frequency. Such a wave is best produced with a tone generator or a synthesizer with a Sine program or a very dull flutelike sound. Failing that, you can use a high-quality microphone and tape recorder to record somebody whistling a pure tone or have a soprano sing on a closed vowel. Try three or four tones that range from middle C (262 Hz) and two octaves up. Connect a pair of headphones to the sound source. Place the headphones on a table in front of you and adjust the volume so that you easily can tell whether the tone is playing or not, but no louder. Listen to its pitch, then put the headphones on without adjusting the volume and listen again. Many people will perceive a lower pitch when the tone thus becomes louder. To some, the drop will be large (a semitone or more); to others, it will be small or inaudible.

The pitch-amplitude effect is well known to psychoacousticians (e.g., Verschuure & van Meeteren, 1975). It is typically at its strongest in the region of 300 to 500 Hz, becomes insignificant at 1 to 2 kHz, and reverses at very high frequencies, where the pitch instead seems to increase with level. This explains why dull tones with only a few low partials are the most susceptible. The more high-frequency overtones there are in the stimulus sound, the less their combined pitch is perturbed, and the weaker the pitch-amplitude effect becomes. Ternström and Sundberg (1988) found a pronounced pitch-amplitude effect in some but not all male subjects who were attempting to sing in unison with a synthesized stimulus vowel /u:/, which is poor in high partials, but not with /a:/, which is rich in high partials. Ternström has also demonstrated the pitch-amplitude effect for soprano tones in syntheses of ensemble sounds (not published). For the effect to occur, the reference tone and the tone produced by the subject must be very different in loudness and the louder tone must have strong low partials and weak high partials. This situation is quite rare in choirs, but it does occur for sopranos who are singing high and loud. The effect will probably be more pronounced in venues with little reverberation, because then the difference in loudness will be large between one's own voice and the rest of the choir (the SOR will be high).

A soprano whose sense of hearing happens to have a steep pitch-amplitude dependency is faced with this problem: when singing high and loud she is likely to perceive her own voice as a little flat, because it is stronger than the other sounds around her and because high soprano tones are necessarily dominated by the fundamental partial (chapter 16). To compensate, she will probably produce a tone that is perceived as sharp by everyone else. Not knowing about the pitch-amplitude effect, she will feel unjustly accused of singing sharp and having a poor sense of pitch. Unfortunately, there is not much she can do to determine exactly at what pitch she should really be singing, other than gradually to learn how flat she should hear herself for the others to stop complaining. We believe choral directors should be aware that the pitch-amplitude effect exists, that it is an innate property of our auditory system, and that it has nothing to do with musical ability or talent.

A related phenomenon is the relationship between *timbre* and *pitch salience*, that is, whether the perceived pitch is clear or unclear (Terhardt, 1975). Listeners tend to agree less on the pitch of a tone if its spectrum is not balanced. For example, it is probably a bad idea to give the starting tone to the choir using a low flute stop on the organ, which has strong low-frequency partials and weak high-frequency partials, because different singers may perceive its pitch slightly differently. It is also risky to use small electronic beeper devices, which are rich in high partials but completely lacking in low partials. A voice or a piano should serve most singers well.

Singers sometimes try to hear their own voice better by blocking one ear with a fingertip. Blocking the ear does make the Self sound louder, but it also becomes muffled, while the Other sound becomes a lot weaker. This jeopardizes intonation, because such conditions increase the risk of the pitch-amplitude effect. Of course, closing one's ear would also look silly on stage. But wearing an earphone or hearing aid, for example, would have much the same effect.

Finally, we cannot even trust our two ears to perceive the same pitch. It is not uncommon to hear slightly different pitches from a tone generator—or tuning fork—that is held first to one ear and then to the other. A choir director giving pitch might need to know if one of his or her ears is slightly out of tune.

## Overtones

*In some long chords I can hear a tone that is not being sung by anyone in the choir. The tone typically sounds like a high-pitched whistle and is sometimes musically appropriate, sometimes not. What is going on?*

### *Relevant Factors*

1. One or more of the following circumstances may combine to make a particular partial tone especially loud: (a) the partial tone is a component of many of the constituent tones of the chord, (b) the formant frequency configuration of the vowel happens to enhance it, (c) many of the singers happen to agree very closely on the formation of that vowel, (d) vibrato is small or nonexistent.
2. The so-called *chorus effect* makes it easier for the ear to pick out the individual partial tones in ensemble sounds than in solo sounds.
3. In loud and/or pressed voice, in which the high-frequency partials are more intense than in soft voice, this phenomenon becomes more likely. This is in fact a strategy adopted for example by the Tuva throat singers (Levin & Edgerton, 1999).

### *Recommendations*

1. If the unsung tone is a problem, sing in a softer nuance or experiment with a different vowel to see if the problem goes away. Modify open vowels to closed or vice versa. If the unsung tone is a desired effect, modify the vowel very slowly and carefully until the effect is maximized. Decide and explain to the singers which section is to act as the intonation reference at various points in the piece.
2. If the singers are having difficulty tuning a particular harmony and if different sections have different lyrics, try making them sing the same text. Change closed vowels to open; in particular, let the basses sing on *ah* for a while. More often than not, rich overtones will facilitate accurate intonation.
3. In practice, independently audible overtones are rarely a problem but rather an interesting curiosity that often goes unnoticed by the audience, unless the composer and the performers have taken extra pains to make the effect come across. In some traditions overtones are even cultivated as an instrument in its own right. Singers often wonder about overtones and how to enhance them, and the choral director can sometimes exploit overtone enhancement in the service of intonation.

## *Background*

**Combination Tones, Beats, and Harmony.** When two stable unison tones of similar loudness are tuned slightly apart, they sound not like two tones but like one tone that *beats*. If the mistuning is gradually increased, the beating becomes too rapid to be perceived as such but rather invokes a roughness of timbre; at some point the ear will also start to resolve the two tones. Furthermore, for tones that are close in frequency to each other, distortion effects inside the ear can generate various audible *combination tones* whose frequencies are sums and differences of the frequencies of the two physical tones and their harmonics.

On instruments with stable pitch, so-called just or pure intonation sounds pure because it minimizes the beating and hides the combination tones by lining them up with the real partial tones. In choral sounds, however, things are rather different. Regular beats are never heard in a choir, for two reasons. The first and foremost is that the human voice is not stable enough in frequency for regular beats to occur. Even when singing an acceptably straight tone, most voices *flutter* up and down by 10 to 20 cents (Ternström & Friberg, 1989). This makes the beats irregular, and therefore they are not readily perceived. One interesting consequence of this is that there should be no inherent advantage of using just intonation in choir music (cf. Chapter 12). Ternström and Nordmark (1996) found that musically trained listeners who were asked to fine-tune major third intervals in synthesized sustained, nonvibrato ensemble sounds had diverse preferences but averaged between just intonation and equal temperament.

The second reason we do not hear the beats in the choir is that with more than two unison voices even the smallest $F_0$ differences between them will cause each voice to beat against each other voice, creating a great multitude of irregular beats. The beating is in fact so profuse and irregular that it sounds like something else. We might define the *chorus effect* or *ensemble effect* to be that character of sound that prevents us from hearing exactly how many voices are singing in unison. The minimum is three, which sounds distinctly different from both one and two voices (Ternström, Friberg, & Sundberg, 1988).

In vocal groups that perform close harmony with one voice to a part, the chorus effect is often intentionally avoided. Instead they strive to achieve harmonies that are so precisely tuned and so straight in pitch that the voices fuse together and we hear one instrumentlike chord rather than several part singers.

**Reinforced Partials as an Aid to Intonation.** We assume here that the reader is already somewhat familiar with the harmonic series of partial tones (fundamental plus overtones) and its relation to the musical scale.

In any sound from a single tone source, some partials will be stronger than others, but if they all vary together, the auditory brain centers will assume that they all derive from the same source and report only one percept to the listener. Therefore, we do not hear individual partial tones, unless one of them is markedly reinforced (see discussion of overtone singing in chapter 16). In ensemble sounds, however, the partial tones do not vary strictly together, as they do in

the sound of a single voice. Thus it can be easier to hear out individual partials in a sustained choral sound than in one voice.

Ternström and Sundberg (1988) did intonation experiments in which individual male singers sang fifths and major thirds together with both synthetic tones and with a prerecorded unison male choir. The experiment showed that when partial number three of the lower tone (which sounds like a fifth) was reinforced by the formant structure of the vowel, the subjects were more agreed on the intonation of fifths; similarly, when partial number five (which sounds like a major third) was strong, the subjects were more agreed on the intonation of major thirds. While the effect was not very large, it appeared even though the subjects were ignorant of the purpose of the experiment. The implication for choral directors is that accurate intonation of a particular chord can sometimes be facilitated by judiciously choosing a vowel that will emphasize (a partial tone at some octave of) the reference tone in the chord.

The formants of the vocal tract are only resonances and produce no sound of their own (chapter 16). Instead, they reinforce whatever frequencies are produced by phonation and frication. In most languages, the two lowest formants, $F_1$ and $F_2$, are sufficient to determine the identity of the vowel. The legend to Figure 17.2 discusses how this can be relevant to choral intonation. Read the words shown in Figure 17.2 in a loud whisper, without voicing, and you will probably hear the pitches of $F_1$ and/or $F_2$, albeit vaguely.

In the first example in Figure 17.2, with *who'd*, $F_2$ reinforces the tone B, which is common to the harmonic series of both the sung tones E and G; this is likely to facilitate intonation. In the *hood* example, $F_1$ emphasizes the G that is common to bass and tenor; again, this is probably good for intonation. In addition,

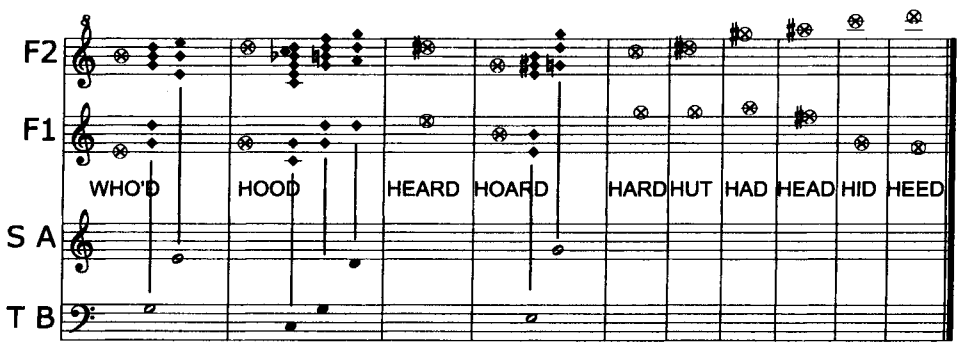



Figure 17.2. Some examples of possible interactions between formants and partials. The crossed circles indicate the approximate frequencies of the first and second formants ($F_1$ and $F_2$) for spoken English vowels as pronounced by adults (Peterson & Barney, 1952). These frequencies are vowel-specific and essentially independent of the sung pitch. Notice the *octava* on the clef for $F_2$; it needs to be notated an octave higher. The diamonds indicate the lowest few partial-tone frequencies of the corresponding sung tones on the S A and T B lines (soprano, alto, tenor, bass).

the $F_2$ of *hood* emphasizes the high D, which is common to all three voices; this is likely to reinforce the ninth of the chord. On *hoard*, however, $F_2$ might cause a problem by amplifying the clash between the $G_\sharp$ that is the fifth partial of the bass tone and the $G_\natural$ that signals the minor tonality.

A good place to learn more about intonation and overtones is at barbershop rehearsals and workshops. Because of their particular style requirements, barbershop singers usually take a keen interest in acoustical matters and often present lessons on the intricate relationship between the formant frequencies and the configuration of the common partials.

## Conclusion

With this small sampler, we hope to have stirred the reader's interest in the field of choir acoustics, which of course encompasses many more topics, such as room and stage design, choral formations, the use of choral risers and reflecting shells, the deployment of sound equipment, and so on. More references to literature by others as well as ourselves can be found at www.speech.kth.se/~sten. While many issues are already addressed by the literature, there are still many more questions that remain to be answered. In anticipation of future research (Daugherty, 2000), let us close with some intriguing questions that have *not* been investigated at the time this chapter was written:

- Does a choral piece that tends to go flat generally stay in tune if it is sung a semitone higher, and if so, why?
- How is it that one or two key singers can seem to lead a whole section, while others seem to become expert at following the leader?
- How is it that some pairs of voices sound better together than others?

**Acknowledgements.** Ternström's and Sundberg's work in choir acoustics has been funded by the Swedish Council for Research in the Humanities and Social Sciences (HSFR), the Swedish Natural Science Research Council (NFR), and the Bank of Sweden Tercentenary Foundation (RJ). Karna's work and travel was funded by the Swedish Information Service (Bicentennial Swedish-American Exchange Grant) and the Office of Sponsored Grants and Research, Salisbury State University, Salisbury, Maryland.

### *References*

Daugherty, J. F. (1996). *Spacing, formation, and choral sound: Preferences and perceptions of auditors and choristers.* Doctoral thesis, Florida State University, Tallahassee.

Daugherty, J. F. (Ed.) (2000). *International Journal of Research in Choral Singing,* on-line publication at www.choralresearch.org.

Fletcher, H. (1934). Loudness, pitch, and timbre of musical tones and their relations to the intensity, the frequency, and the overtone structure. *Journal of the Acoustical Society of America, 6,* 59–69.

Ladd, D. R., & Silverman, K. E. A. (1984). Vowel intrinsic pitch in connected speech. *Phonetica, 41*, 31–40.

Larson, C. R., Burnett, T. A., Kiran, S., & Hain, T. C. (2000). Effects of pitch-shift velocity on voice $F_0$ responses. *Journal of the Acoustical Society of America, 107*(1), 559–564.

Levin, T. C., & Edgerton, M. E. (1999). The throat singers of Tuva. *Scientific American, 281*(3), 80–87.

Marshall, A. H., & Meyer, J. (1985). The directivity and auditory impressions of singers. *Acustica, 58*, 130–140.

Peterson G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America, 24*, 175–184.

Terhardt, E. (1975). Influence of intensity on the pitch of complex tones. *Acustica, 33*, 344–348.

Ternström, S. (1994). Hearing myself with the others—sound levels in choral performance measured with separation of the own voice from the rest of the choir. *Journal of Voice, 8*(4), 293–302.

Ternström, S. (1995). Self-to-other ratios measured in choral performance. In M. J. Newman (Ed.), *Proceedings of the Fifteenth International Conference on Acoustics* (pp. 681–684). Trondheim: Acoustical Society of Norway.

Ternström, S. (1999). Preferred self-to-other ratios in choir singing. *Journal of the Acoustical Society of America, 105*(6), 3563–3574.

Ternström, S., & Friberg, A. (1989). Analysis and simulation of small variations in the fundamental frequency of sustained vowels. *Speech Transmission Laboratory Quarterly Progress and Status Report, 3*, 1–14.

Ternström, S., Friberg, A., & Sundberg, J. (1988). Synthesizing choir singing. *Journal of Voice, 1*(4), 332–335.

Ternström, S., & Nordmark, J. (1996). Intonation preferences for major thirds with non-beating ensemble sounds. Proceedings of NAM96 (Nordic Acoustical Meeting) (pp. 359–365). Helsinki: Acoustical Society of Finland.

Ternström, S., Sundberg, J., & Colldén, A. (1988). Articulatory $F_0$ perturbations and auditory feedback. *Journal of Speech and Hearing Research, 31*, 187–192.

Titze, I. R. (1989). On the relationship between subglottal pressure and fundamental frequency in phonation. *Journal of the Acoustical Society of America, 85*(2), 901–906.

Verschuure, J., & van Meeteren, A. A. (1975). The effect of intensity on pitch. *Acustica, 32*, 33–44.

von Békésy, G. (1949). The structure of the middle ear and the hearing of one's own voice by bone conduction. *Journal of the Acoustical Society of America, 21*, 217–232.

# 18

# Piano

RICHARD PARNCUTT & MALCOLM TROUP

On the basis of research on the physics and physiology of the keystroke, the acoustics and perception of piano timbre, and the psychology of piano fingering, we explain observations such as the following, and investigate their practical implications. The timbre of an isolated tone cannot be varied independently of its loudness but depends on finger-key, key-keybed, hammer-key noise, and on the use of both pedals. The timbre of a chord further depends on the balance and onset timing of its tones, whereby louder tones tend to sound earlier (melody lead, velocity artifact). Both the sustaining pedal and *una corda* can enhance *sostenuto*. Leap trajectories are curved and asymmetrical. Optimal fingering is determined by physical, anatomic, motor, and cognitive constraints interacting with interpretive considerations, and depends on expertise.

Scientific thinking, methods, and results have influenced piano performance and piano teaching for well over a century, and innumerable piano-pedagogical publications have claimed scientific validity. On the one hand, artistic writers—often great pianists and piano teachers—have tended to fashion complex pseudoscientific theories post hoc to match their beliefs, so that such theories can be controversial and unreliable. On the other hand, scientific writers tend to focus on simple hypotheses and assumptions that are easy to demonstrate and explain but are of limited interest to musicians. It is little wonder, therefore, that modern piano students are often unaware of the basic findings (Parncutt & Holming, 2000).

Ortmann (1929/1981) successfully challenged influential but unfounded assumptions on touch and relaxation. Like him, we begin with observations that are easy to demonstrate scientifically and move gradually toward more complex ideas that are more likely to be of interest to modern pianists and piano teachers. We attempt a fresh approach by combining old and new acoustical

and psychological thinking and drawing upon our own performing and teaching experience.

## Physics and Physiology of the Keystroke

### *Curved Versus Straight Fingers*

The motor independence required to perform the complex sequences of finger movements that are typical of piano performance can only be acquired over many years of concentrated practice. Pianists must control not only the order and precise timing of keystrokes but also their force and the resultant key velocity (Parlitz, Peschel, & Altenmüller, 1998).

The force required to depress a key in the middle register varies by a factor of 100: to hold down the key requires about 0.5 newton (equivalent to 50 grams), and to play *fortissimo* requires about 50 newtons (5 kilograms) (Askenfelt & Jansson, 1991). One way to accommodate this wide range is to vary the curvature of the fingers (Ortmann, 1925).

Loud, scalelike passages are often better played with curved fingers. First, this allows force to be applied vertically and so most efficiently (Ortmann, 1929/1981; Gát, 1965). Second, this reduces the distance between the fingertip and the knuckle, increasing the available force at the fingertip for a given muscle force (lever principle).

Straight fingers are often preferred in softer, slower, single-line melodies, where the fleshy pads of flatter fingers allow a bigger skin area to touch the surface of the key, appropriate for *cantabile*. Extended fingers can also move through a larger horizontal arc than curved fingers and are appropriate when big stretches are required at a low to moderate dynamic level.

### *Fingers or Arms?*

As a rule, *all* moving joints in the arms and fingers are involved to some extent in *all* piano playing; to try to prevent one joint from moving would cause unnecessary tension. Beyond that, basic physics predicts that small, lightweight limbs (e.g., the fingers) are suited for fast and/or quiet playing, while larger, heavier limbs (the upper and lower arm) are best for slow, loud playing (Ortmann, 1929/1981). According to this principle, movements of the hand (from the wrist) and forearm (from the elbow) are suitable for intermediate tempi and dynamic levels and for combinations like fast and loud (e.g., Liszt's Hungarian Rhapsody No. 6, with its right-hand presto octaves and left-hand chord leaps, all marked *sempre forte*). Similarly, scales in fast, lightweight octaves are best played from the wrist, moderately fast and loud octaves from the elbow, and loud, *bravura* octaves from the shoulder (cf. Gát, 1965), examples of all of which are also to be found in Liszt's Hungarian Rhapsody No. 6.

The underlying physical principle is Newton's second law of motion—force equals mass times acceleration (cf. Schultz, 1949)—which states that greater

masses require more force to accelerate them and exert more force when they decelerate (inertia). Thus the fingers can move back and forth more quickly, but the arms can exert more force.

This principle is mirrored in the historical evolution of keyboard technique (Gerig, 1974; Kochevitsky, 1967). In keyboard music of the eighteenth century, the fingertips did most of the work and the arms were held to the sides. In the nineteenth century, a heavier keyboard action, larger concert halls and orchestras, the new mass public, the emergence of the solo recital, and more elaborate writing for the keyboard made wrists and forearms more important (Kalkbrenner, Leschetizky). This led to Deppe's (1885) and Breithaupt's (1905) concept of arm weight and Matthay's (1903) emphasis on relaxation. Their basic idea was to relax the arm in free fall, then introduce minimal muscle tension to control the movement.

A similar pathway is still traveled by most pianists during their development from beginners to professionals. But advanced technical concepts such as arm weight can also be introduced quite early. For example, György Kurtág's *Jatekók*, a collection of children's pieces elaborated into a system of piano tuition by Kase (2001), introduces larger scale bodily movements from the start, delaying focused finger movements until later.

### *Leap Trajectories*

Performers of pieces like Brahms's *Paganini Variations* and Second Piano Concerto are constantly confronted with high-speed leaps across the keyboard. Ortmann (1929/1981) observed that the optimal trajectory for a leap is not only *curved* but also *skewed* (asymmetrical): it departs close to horizontal and lands somewhere between horizontal and vertical. This minimizes the chance of missing the target and allows the force on the target key to be applied almost vertically.

During a leap, the hand and arm must accelerate from stationary to a maximum velocity and decelerate again before reaching the target. Since acceleration is a vector (with both magnitude and direction), leap trajectories cannot include sharp corners, which would require short bursts of high acceleration (Ortmann, 1929/1981). For similar reasons, it is widely held that rounded or curved movements are preferable in all areas of piano technique: "The shortest distance between two points is not a straight line but a curve" (Neuhaus, 1973).

In general, the likelihood of missing a target increases as the trajectory becomes longer, the time available becomes shorter, and the target becomes smaller (Fitts, 1954; Huron, 2001). For the piano, the margin of error at the target is the width of the key as seen from the approaching finger—greatest if the approach is vertical (from the air) and zero if horizontal (along the keyboard). When approached from the left or bass, the widest black-key targets are the tones C♯ and F♯; from the right or treble, B♭ and E♭. The optimal trajectory toward these keys is therefore flatter, and they can be reached more accurately. According to this logic, the falling leap at the start of Beethoven's Opus 111 (E♭ down to F♯ in left-hand octaves) may be more difficult than a slightly longer leap to a bigger target (e.g., D♭ to E♭).

While performing, pianists tend to be unaware of leap trajectories. The process seems highly automatic and intuitive, with the attention focusing aurally and visually on the target, not the trajectory, and relying on a combined tactile-auditory-visual memory of the keyboard. If unnecessary tension is avoided, an appropriately curved and skewed trajectory results automatically. The kinesthetic sense of joint, tendon, and muscle movements during the trajectory can be developed either by practicing leaps out of context without depressing the target tone (just preparing it) or by practicing leaps in the dark (without visual feedback). It is difficult to practice leaps at a slow tempo, just as high jumpers cannot practice in slow motion.

### *Tone Repetitions*

High-speed repetitions are usually easier to perform by changing (alternating) fingers, because it is easier to move one finger horizontally off the key and drop another finger onto it, than to quickly move the same finger up and down. The latter requires more finger acceleration, which in turn requires more muscle force. Slow tone repetitions are often best performed *without* changing fingers, because at slower tempi the listener is more sensitive to the unevenness (variation in loudness) that can result from the differing strengths of the fingers. Moreover, at slow tempi there is less finger-key noise. Conversely, changing fingers can be useful if accentuation (i.e., *deliberate* unevenness) is required.

Fast repetitions were enabled by Sebastien Érard's *repeating action* of 1823 (Frey, 1933). The *let-off distance* (Askenfelt & Jannson, 1990) is the distance between the string and the top of the hammer after a key is slowly and silently depressed. Typically some one to three millimeters, it is also the distance traveled by the hammer in free flight after the jack is pushed away by the escapement dolly. The roughly corresponding *let-off point* of a piano key is found by slowly pressing the key until resistance is felt (as the jack escapes).

Depending on tempo and dynamics, fast repetitions may be best played *under the surface*, that is, by depressing the key to the keybed, lifting to the let-off point, and depressing again. The minimum time between repetitions is the time taken for the jack to return to its original position. A short setting of the let-off distance facilitates fast, quiet repetitions, as for example in the left-hand opening (*pp, très fondu, en trémolo*) of *Scarbo* from Ravel's *Gaspard de la nuit*. Playing from the let-off point also prevents the damper from reaching the string between keystrokes, so that a continuous sound can be obtained without pedaling.

## Acoustics and Perception of Piano Timbre

### *Relationship Between Timbre and Loudness*

We use the words *timbre* and *loudness* in the widely accepted psychoacoustical sense of how a sound is perceived or experienced by a listener (H. Fletcher, 1934). In general, the timbre of a sound–its perceived tone quality or color—depends

on at least two physical variables: its *spectral envelope* (relative amplitude of spectral partials at a given moment) and its *temporal envelope* (which rises suddenly as the hammer hits the string and then decays gradually). The critical importance of temporal envelope for piano timbre becomes clear when a piano tone is recorded and played backward. If the amplitude gradually increases instead of decreasing, with the hammer noise at the end instead of the beginning, the result sounds strangely like an organ (Houtsma, Rossing, & Wagenaars, 1987).

The experience of loudness may be distinguished from the physical measurement of *SPL*, measured in decibels, and from *dynamic level*, which is an instruction, recommendation, or suggestion from a composer to a musician. For practical purposes, loudness, SPL, and dynamic level are often the same, but they do diverge in everyday musical situations. Consider for example a thick bass chord marked *ppp*. The physics of the piano do not allow such a chord to be played with a very low SPL (as measured with an SPL meter). But a pianist playing as softly as the piano allows can manipulate other contextual variables such as timbre (pedaling, relative key velocity) and timing (tempo, delay) to enhance the impression of *ppp*.

Against the backdrop of these definitions, we may assert that the spectral and temporal envelope of an isolated piano tone cannot be changed independently of its SPL; hence, timbre cannot be changed independently of loudness (Baron & Hollo, 1935; Gát, 1965; Gerig, 1974; Hart, Fuller, & Lusby, 1934; Ortmann, 1937; Seashore, 1937). In a simple physical model of the piano action (Fletcher & Rossing, 1998, Figure 12.1), both the SPL and the spectral and temporal envelopes of a piano tone are determined uniquely by *hammer velocity*—the speed with which the hammer hits the string, which is determined in turn by key velocity (cf. Palmer and Brown, 1991). The crucial point is that the hammer hits the string in *free flight*: at (and just before) impact, there is no physical contact between the key and the hammer. Thus—apart from some interesting possibilities described later—the pianist cannot influence *how* the hammer hits the string, only the *speed* with which it does so.

Acousticians and psychologists have often wondered why, in spite of this evidence, so many pianists still believe that the timbre of a piano tone depends on *touch*—not only how fast but also *how* the key is depressed. A possible reason is that movements of a pianist's body and arms (smooth and round versus jagged and tense) seem to both performers and audiences to result in different timbres. But we cannot necessarily rely on this impression for the following reasons:

- Pianists almost never perform isolated tones without pedal, and the timbre of more complex textures clearly depends on a range of factors other than the way each individual key is depressed (see later).
- Movements that directly affect *control* may indirectly affect timbre. Consider the difference between *percussive* (or *staccato*) *touch* (when keys are hit from a distance above the surface) and *nonpercussive* or *legato touch* (when the key is depressed from the surface). Ortmann (1925) found that nonpercussive touch permits finer key control.
- Listeners can easily perceive timbral differences between piano performances but are not necessarily able to explain the origin of those differences (Heinlein, 1929b).

- Listeners cannot necessarily separate timbre from other perceptual variables. Timbre is often confused with pitch in ratings of performed intonation (chapter 12) and in other psychoacoustic experiments (Singh & Hirsh, 1992). Listeners regularly use the interdependence of timbre and loudness to deduce dynamic level in situations where sound intensity cues are lacking or unreliable, such as recordings. In most instruments (including the voice), louder tones have more jagged waveforms and hence relatively more high-frequency energy. In the piano, a sharper hammer blow produces sharper waveform peaks (Hall, 1991). The strongest perceptual cue to the original dynamic level of a piano recording is its timbre—regardless of how loudly the recording is played back.
- In general, auditory perception is influenced by visual perception (intermodal interference). Visual processing can affect the appraisal of piano performances (Behne, 1990; Shimosako & Oghushi, 1996), and an important part of the emotional and interpretive message sent from a performer to an audience is visual (chapter 15 in this volume; Rosen, 1995). For example, if an audience sees a pianist brutally hitting the keys, they may get an impression of a hard or brittle tone—beyond what the music actually sounds like.
- Pianists' perceptions of their own performances are multimodal (Galembo, Askenfelt, & Cuddy, 1998): the pianist's perception of timbre can be influenced by kinesthetic feedback from finger contact with the keys.

Askenfelt and Jansson (1991) suggested the possibility that *hammer-shank vibrations* at around 50 Hz might be greater in percussive touch and could theoretically affect piano string vibration—either by changing the angle of impact and thus shifting the striking point (only significant in the extreme treble: Galembo et al., 1998) or by allowing the hammer to rub along the string during contact. But Hart, Fuller, and Lusby (1934) and others (see Galembo et al., 1988) had already concluded that the influence of the hammer shank on the physical string motion is negligible, and when the matter was tested empirically Askenfelt and Jansson (1991) were unable to observe any rubbing motion between hammer and string. Even if effects of this kind are possible, the pianist can hardly manipulate them independently of finger-key noise (see later), so they cannot contribute *independently* to piano timbre.

Tone quality in piano performance is determined not only by the physics of individual keystrokes but also involves a complex and largely intuitive interaction among body movements, technical finesse, and musical interpretation (Kochevitsky, 1967). For example, it is possible that the exact timing of a *rubato* melodic phrase affects the global perception of timbre. The ability to produce a variety of timbres and to apply them appropriately to the interpretation of repertoire can only develop gradually over years of concentrated practice and careful listening.

### *Timbre of a Chord*

The timbre of a piano chord depends on the timing and relative loudness of the tones (Baron & Hollo, 1935), because these affect both the temporal and the spec-

tral envelope. The attack portion of the temporal envelope can be manipulated by adjusting the timing of the tone onsets. An extreme case is an arpeggiated chord, whose timbre depends on the speed and direction of the arpeggiation—an expressive strategy employed by both performers (e.g., Glenn Gould's interpretations of Bach) and composers (e.g., Boulez's Sonate No. 3, *Constellations*). The spectral envelope can be manipulated by playing some tones louder and some softer. If the louder tone takes on a singing quality, it is either because its pitch becomes more *salient* (clear, prominent, audible, able to attract attention: Terhardt, Stoll, & Seewann, 1982) and/or because the timbre of the whole sonority becomes less rough (the roughness of a beating pair of pure tones falls rapidly as the difference between their amplitudes increases; cf. Terhardt, 1974). The technique of bringing out a tone is addressed later under "Melody Lead and the Velocity Artifact."

### *Sources of Percussive Noise*

Percussive onset noise is as characteristic and integral to piano timbre as is the scraping of the bow of a stringed instrument or the breath activation of a wind instrument. This is clearly demonstrated when onset noise is electronically excluded from the tone, creating a drastic change in timbre. Normally, we hear holistically—we do not, and perhaps cannot, hear the onset noise separately. If the amount of noise increases, we hear the sound as increasingly harsh, dry, ugly, or forced and may be unaware of the physical source of the timbral change.

The three main sources of noise in piano timbre are *finger-key*, *key-keybed*, and *hammer-string*. Finger-key noise can be varied independently of string amplitude, but the biggest contributor to the noisy onset, hammer-string noise, cannot (Baron, 1958; Baron & Hollo, 1935; Hill, 1940), and, for practical purposes, neither can key-keybed noise.

**Hammer-String Noise.** The physical intensity of the hammer blow depends directly on its velocity, being most intense in *fortissimo*. For this reason, *cantabile* can sometimes be enhanced by limiting loudness. Perceptually, the situation is more complex, because the tone's partials mask the onset noise, so louder tones do not always *seem* more percussive. If higher tones seem more percussive than lower tones, it is because their noise component is more intense by comparison to the partials—not because the noise is more intense in an absolute sense.

**Finger-Key Noise.** This can contribute to the percussiveness of piano sound (appropriate, e.g., in many of Bartók's *Mikrokosmos*). It can also enhance an impression of *staccato* (Gát, 1965). Its audibility can be demonstrated by holding down a cluster of keys with one hand hitting them with the fingertips of the other hand. In practice, finger-key noise is not completely independent of hammer-string noise, because pianists tend to play louder when using percussive touch (Ortmann, 1925).

The impact between finger and key causes a sudden high acceleration and associated shock excitation (or thump) dominated by two flexing resonances of

the key at about 290 and 445 Hz (Steinway); this is transmitted through various structural parts of the piano to the soundboard (Askenfelt & Jansson, 1990, 1991). It is called *touch precursor* because it precedes the hammer blow by about twenty to forty milliseconds, depending on dynamic level (Askenfelt, 1993). Like early reflections in a concert hall (Hall, 1991; chapter 17 in this volume), the touch precursor can contribute to timbre but cannot be heard separately.

In a musical context, the tonal part of the piano sound may partially or completely mask finger-key noise, depending on dynamic level, texture, register, room acoustics, and position of the listener. Finger-key noise is more audible in higher registers, because it is less masked by the tone.

Finger-key noise may seem important to the performer but be irrelevant to the listener. Preliminary experiments suggest that listeners can discriminate percussive from nonpercussive touch, but only within some twenty centimeters of the string (Galembo et al., 1998; Koornhof & van der Walt, 1993). Pianists' perceptions of their own finger-key noise may also be influenced by the kinesthetic sensation of the finger hitting the key.

**Key-Keybed Noise.** This may be separately heard by holding the hammer against the string while pressing the key. In a musical context, key-keybed noise occurs almost simultaneously with hammer-string noise and so blends easily with it. Its low frequency gives it a deep timbre.

Theoretically, pianists can indirectly control key-keybed noise (Askenfelt & Jansson, 1991, Figures 4 and 5). In nonpercussive touch, the key velocity increases steadily during the key depression, reaching a maximum at the keybed. In percussive touch, the key velocity first increases, then decreases (as the energy is transferred to potential energy in parts of the action and kinetic energy in vibrations of the hammer shank), then increases again. So for a given dynamic level, key-keybed noise is greater in percussive than nonpercussive touch. In practice, due to limitations of sensory feedback to motor control, it is virtually impossible to vary key-keybed noise independently of the percussiveness of touch and dynamic level.

## Cantabile and the Pedals

*Cantabile* means "singable" or "songlike." The sustaining pedal enhances the singing quality of a piano melody by improving the legato and enriching the timbre. The *una corda* pedal can also enhance *cantabile*, by increasing the effective tone duration.

### *The Sustaining Pedal*

Banowetz (1985) regarded the sustaining pedal as "equivalent to the vibrato of the singer or the string player" (p. 13). Pianists do not simply depress and release the pedal but take advantage of a quasi-continuous series of intermediate positions that allow for more or fewer dampers to clear the strings in different

registers—part-pedaling (half, quarter, three-quarter, etc.), pedal squeezing, and flutter (or vibrating) pedal (K. U. Schnabel, 1950). The effectiveness of these techniques depends on a variety of physical factors (the instrument's make; wear and tear of the hammer heads; pre- and postrelease duration of tones in a given register; room acoustics), technical considerations (how many fingers are held down at a given instant), and interpretation (phrasing, tempo, dynamic and timbral shading, tonal relationships) and is controlled by highly intuitive auditory reflexes and responses (Heinlein, 1930; Repp, 1999)—which may explain the considerable differences in pedal usage observed even among expert pianists (Heinlein, 1929a).

Walter Gieseking considered that "just as one learns correct finger technique from the head and not the fingers, so one learns correct pedaling from the dictates of the ear and not the foot" (Banowetz, 1985, p. 231). The independent manipulation of pedal and fingers, as well as the ability to adjust to a variety of instruments and acoustic conditions, is one of the most difficult and important pianistic techniques to learn. This complexity makes pedaling hard to investigate scientifically.

### *Decay Time and Melody*

Is the piano a suitable instrument for the performance of melody? Two of the perceptual factors that encourage a sense of melody (across cultures) are *sustain* (slow decay) and *pitch salience* (Huron, 2001). The piano best satisfies both conditions in the middle registers: high tones decay quickly, and both very high and very low tones have low pitch salience (cf. Terhardt et al., 1982).

By convention, the decay time of a sound is the time taken for its SPL to fall by 60 dB. This corresponds roughly to its perceived or effective duration. The decay of a piano tone before the finger is lifted is called *prerelease decay*, and varies from some 15 seconds in the deep bass to 0.5 second in the highest register (Hall, 1991; Repp, 1997; Martin, 1947). *Postrelease* decay times are roughly half a second in both treble and bass (Repp 1997, Figure 4), but they vary unpredictably from tone to tone, depending on the state of the dampers and their pressure on the string. The timbral effect of damping the string is more audible in the lower registers where the strings are heavier.

A sense of melody may be created in high registers by use of the sustaining pedal and by *legatissimo*, the overlapping of successive keystrokes (Repp, 1997). Legatissimo can be acoustically effective in all but the extreme high register where there are no dampers. Finger legato is important not only where the pedal is being changed frequently but also where part- or flutter pedaling is being used, which tends to damp the upper strings more than the lower.

Another way to create an impression of melody is suggested by Bregman's (1990) theory of *auditory streaming*. Successive tones are more likely to hang together as melody if they are close in pitch and time and similar in loudness and timbre. Thus a pianist can optimize *cantabile* by holding key velocity relatively constant, compensating for the piano's acoustic discontinuity in the higher registers. If a passage calls for metrical or structural accentuation (chapter 13),

other means may be used, such as agogic accents in the melody or dynamic accents in the accompaniment.

### *The Sustaining Effect of* Una Corda

Casella (1936) considered that the *una corda* (soft) pedal "allows the executant to interpret a *cantabile* melody with greater depth of touch while never overstepping the mark" (our translation). Rubinstein (1980) remarked that he often produced his best *cantabile* by playing the melody *ff* with soft pedal.

The underlying physics was explained by Weinreich (1977; see also Hall, 1991; Fletcher & Rossing, 1998). The *una corda* pedal reduces the decay rate of the tone and so increases the effective duration. On a grand piano, the pedal shifts the mechanism sideways so that the hammer strikes only two (not one, as the Italian name suggests) of three unison strings. This reduces loudness, mellows the timbre, and modifies the hammer-string noise, because a different part of the hammer—the ridges between the grooves formed during normal playing—strikes the strings.

When the *una corda* pedal is *not* depressed, the three unison strings initially vibrate in phase, allowing them to transmit energy efficiently to the bridge and causing the sound level to fall rapidly. But the strings are never tuned to exactly the same frequency—and good tuners sometimes deliberately slightly mistune unison strings, for timbral reasons (Wead, 1921; Weinreich, 1977). So they soon get out of phase with one another, reducing the efficiency of sound transmission to the bridge. Once the phase relationship between the strings is effectively random, a second, slower decay phase—the *aftersound*—begins. With the *una corda* pedal depressed, one string begins almost at rest. Over the next few seconds, energy is transferred from the struck to the unstruck strings via the bridge. The overall decay rate is slower, because the struck and unstruck strings are out of phase from the start; the aftersound begins sooner and lasts longer (Hall, 1991, Figure 10.9).

We recommend experimenting with the sustaining effect of *una corda* in relatively quiet passages whose timbre is not adversely affected by the soft pedal. Melodic tones should be relatively long (say, a full second or more), to give the unstruck string time to start vibrating. The audibility of the effect will of course vary according to the instrument, chosen register, and room acoustics.

## Melody Lead and the Velocity Artifact

Integral to piano technique is *polyphonic touch*, or the differentiation of dynamic levels within a chord, which enables melodies to be brought out of a contrapuntal context. Pianists have been doing this since the earliest clavichords and fortepianos allowed key velocity to determine dynamic level, allowing for example the entries of Bach fugal subjects to be projected. Sophisticated exercises and studies that addressed this technique appeared in the late nineteenth century. For a later generation, Artur Schnabel suggested that students play the

cluster CDEFG repeatedly with fingering 1-2-3-4-5, each time bringing out a different tone to create a melody, for example, 5–3–3, 4–2–2, 1–2–3–4–5–5–5 ("Hänschen klein," a German nursery tune) (Wolff, 1972, p. 177). In the eighth measure of his Klavierstück Nr. 2 (1) of 1954 (London: Universal), Stockhausen notated a widely spaced five-note chord for the right hand and labeled the individual tones *ff, fff, f, mf,* and *f*—a tall order for a human performer but something that the piano is at least physically capable of.

Vladimir Horowitz recommended that "in striking a chord, in which a single note is to be accented, the effect can be produced by holding the finger which is to play the melody note a trifle lower and much firmer than the fingers which are to play the unaccented notes. The reason for holding the finger a trifle lower is only psychological in effect; in actual practice, it isn't altogether necessary" (Eisenberg, 1928). Along similar lines, Schultz (1949, p. 176) suggested that "to accent the *C* in the chord *E-G-C* . . . fingered 1-2-5, the thumb and second finger show more joint-movement than the fifth, the muscles of which contract relatively strongly. The accented tone is actually played a fraction of a second before the unaccented ones, but the difference in time is too small to be readily observable."

The asynchrony referred to by Schultz is known in modern music-psychological literature as *melody lead.* In the piano, the time taken for a single key to fall from the surface to the keybed ranges from about 25 ms in *forte* to 160 ms in *piano* (Askenfelt & Jansson, 1991). When fingers strike the key surface simultaneously but descend at different velocities, the maximum possible asynchrony is around 100 ms (remembering that, in quiet percussive touch, the hammer reaches the string up to 35 ms before the key reaches the keybed but only 2 to 5 ms in *forte*: Askenfelt & Jansson, 1990, Figure 8). In practice, melody leads are typically in the range of 20 to 40 ms, which is short enough to be inaudible (again, like early reflections in a concert hall). They are observed in performances at all levels of expertise (Palmer, 1989, 1996; Repp, 1996).

Goebl (2001) found that melody lead times could be predicted accurately by calculating key and hammer travel times on the basis of their measured velocities, confirming that in one-hand chords his pianists were striking the key surfaces almost simultaneously, regardless of emphasis. This was consistent with the pianists' apparent lack of awareness of melody lead in interviews. Melody lead may thus be regarded primarily as an *artifact* of keyboard construction.

Melody leads of similar duration also occur in ensemble performance (Rasch, 1979). This could be an artifact of a different kind: those that lead an ensemble tend to do just that and play infinitesimally earlier than the others who follow—just as less-experienced piano accompanists sometimes tend to drag. Alternatively, the effect may be a deliberate or intuitive expressive device to help bring out the melody. Pianists, too, may *deliberately* anticipate melody tones by different degrees, depending on the expressive intention (Rasch, 1978; Henderson, 1936; Palmer, 1989, 1996). But this is difficult to confirm experimentally, as one must first subtract out the effect of the velocity artifact.

Melody lead (or lag) can render a tone more salient than other tones in a chord, independent of the associated intensity differences. The anticipated tone is ini-

tially not masked by the other tones; and according to Bregman's (1990) theory and experiments, a slightly asynchronous tone is more easily heard as part of a melody. Thus, from a perceptual point of view, it may be unnecessary or even inappropriate to develop a deliberate strategy to compensate for the velocity artifact—that is, to attempt to control the relationship between key velocity and synchrony.

A pianist who wishes to do this consciously and thereby widen the range of available interpretive possibilities is faced with a technical problem. Due to intrinsic motor and physiological limitations, it is extremely difficult to directly manipulate the timing of almost-simultaneous finger movements, because the time required by sensorimotor and aural-neurophysiological feedback between the fingers and the brain would be similar to the typical duration of melody lead times (some tens of ms). The first author of this chapter gets around this by raising the finger whose onset is to be delayed and setting all fingers in motion at about the same time, applying more force and velocity to the raised finger. That finger then hits the key surface slightly later than the others—not earlier, as suggested by Horowitz, and perhaps corresponding to the "diametrically opposed" method referred to by Eisenberg (1928). By adjusting the height of the raised finger and listening carefully, the asynchrony at the keyboard can be reduced. The technical and artistic effectiveness of this technique, which follows logically from the psychological findings but seems to contradicts much pianistic practice, is yet to be systematically investigated.

We have not addressed the tendency of some pianists—especially pianists from the early days of recording—to deliberately *delay* the melody (sometimes called bass lead: Goebl, in press; Repp, 1996; Vernon, 1936). This is technically relatively easy, since the asynchrony is between rather than within hands and may be regarded as a kind of agogic accent (chapter 13).

## Psychology of Piano Fingering

> *To impose a fingering cannot logically meet the different conformations of hands . . . the absence of fingerings is an excellent exercise, suppresses the spirit of contradiction which induces us to choose to ignore the fingerings of the composer, and proves those eternal words: "One is never better served than by oneself." Let us seek our fingerings!*
>
> Debussy, *Douze Etudes*

How do pianists determine fingerings? What are the underlying criteria? How do fingerings differ among pianists? Only recently have questions such as these been regarded as amenable to psychological research methods.

In the following, we limit ourselves to psychological aspects of fingering and assume that pianists are free to choose their own. The question of whether

fingerings prescribed by composers such as Schubert, Chopin, Brahms, Liszt, Rachmaninoff, and Bartók should be followed (as, for example, Claudio Arrau has insisted) is a cultural, historical, and perhaps even ethical one and beyond our scope here.

### *Optimal Fingering*

Optimal fingering emerges from a trade-off or compromise among various physical, anatomic, motor, and cognitive constraints, in conjunction with interpretive considerations. It depends on the relative importance of these different aspects for a given pianist or musical context. The complex interaction among these various constraints and their dependence on pianist and style mean that one can rarely speak of a single best fingering for a given passage.

*Physical* constraints on fingering include the horizontal and vertical arrangement of the black and white keys. In the determination of fingerings, these interact with *anatomic* constraints and associated individual differences: hand size and shape (measured by Wagner, 1988; see photos of pianists' hands in Gát, 1965) or the maximum comfortable stretches between pairs of fingers (included in the model of Parncutt et al., 1997; see also Jacobs, 2001).

*Motor* constraints can apply either within a single hand or between hands. Within a hand, they limit finger independence. According to Ortmann (1929/1981), such motor constraints can be reduced by practice but never eliminated. Between hands, motor constraints are involved in the coordination of tremolos, two-hand trills, and shakes; the execution of seamless transitions from one hand to another within Thalberg-style thumb melodies or accompanying figures; and the sharing of technically difficult figurations between the hands.

*Cognitive* constraints determine how well we can encode and retrieve complex finger patterns and their context dependencies. Fingerings that stay the same in different keys may reduce cognitive demands but are physically and anatomically more difficult. A change of fingering at a key change may be worthwhile if the anatomic and physical advantages balance the added cognitive load. To master these difficulties, many piano pedagogues, including Liszt (and, e.g., Frey, 1933), have recommended transposing difficult passages to all keys without changing the fingering.

Interactions between the two hands are primarily limited by cognitive constraints. In sight-reading unison passages (hands an octave apart), for example, two different fingerings must be planned and executed simultaneously (Parncutt, Sloboda, & Clarke, 1999). The cognitive load is reduced in mirror-image patterns in contrary motion, where fingerings can be identical in both hands—for example, in Messiaen's *Vingt regards sur l'enfant Jésus* (Troup, 1983, 1995).

By *interpretive* considerations we mean both communication of musical *structure* (chapter 13) and *emotion* (chapter 14). Fingering choices are often determined by matching a constraint to a given interpretive intention, for example, thumb on black at the start of a measure or phrase (Clarke et al., 1997) or second (index) finger at the end of a phrase. A more general interpretive consideration

is flexibility: pianists need a fingering that allows for changes of interpretation without changes of fingering that would increase the rate of errors (Sloboda et al., 1998).

### *Procedural Versus Declarative Knowledge*

The knowledge (or memory) that underlies fingering choices may be described as episodic or semantic, or as procedural or declarative. These terms are defined in psychology (see, e.g., introductory texts) as follows. *Episodic* knowledge refers to a specific event: a pianist may finger a passage in a certain way because the passage resembles a previously encountered and fingered passage. *Semantic* knowledge is more generalized: here it includes rules and principles of fingering. Both episodic and semantic knowledge can be either declarative or procedural. *Declarative* knowledge can be verbalized (knowing *that*), but *procedural* knowledge (knowing *how*) typically refers to *automated* movements that underlie a skill, such as riding a bicycle or producing grammatically correct sentences: a five-year-old may be able to do both but is unable to describe the underlying physical or grammatical rules or principles.

All four aspects are important at all levels. Even the most experienced pianists try out fingerings at the keyboard (procedural) but can also explain their fingering choices (declarative: Clarke et al., 1997). Procedural knowledge of fingering may be acquired either by applying declarative knowledge (rules) to repertoire excerpts (episodic: see, e.g., technical methods of the 1920s and 1930s) or by improvising exercises to address specific fingering issues (semantic: e.g., Cortot, 1958; Gellrich & Parncutt, 1998).

### *Dependence on Task*

Optimal fingering depends on whether a passage of music is improvised, sight-read, played from the score after rehearsal, or performed from memory. For example, marked differences were observed between fingerings spontaneously used by pianists in the sight-reading task of Sloboda et al. (1998) and fingerings written on scores of the same music by the same participants several months previously (Parncutt et al., 1997). Standard fingerings for scales and arpeggios are of course more useful in sight-reading (where they reduce the cognitive load) than in memorized performance, where the pianist has time to automate new fingerings (Clarke et al., 1997)—although this does not necessarily apply to professional sight-readers, who have usually seen the score in advance (chapter 9).

Differences between written and sight-read (or improvised) fingerings may be explained if we assume that writing primarily taps declarative knowledge and sight-reading (or improvisation) primarily taps procedural knowledge. Only during rehearsal does a pianist have the chance to allow declarative and procedural knowledge to interact, as new fingering patterns are deliberately learned and automated. Teachers and performers may therefore develop different fingering approaches and strategies for sight-reading as apart from rehearsed per-

formance, beyond merely learning standard fingerings for scales and arpeggios (cf. Deutsch, 1959).

### *Dependence on Expertise*

Fingering depends on expertise (Clarke et al., 1997; Sloboda et al., 1998), because technique (finger independence, coordination of finger, arm, and hand movements, and so on) typically matures before interpretive ability and personal style. Beginners focus almost entirely on anatomic and physiological constraints but may not yet have the experience and knowledge necessary to choose the easiest variant. Young professional pianists (e.g., conservatory graduates) tend to use a fingering that is anatomically and physiologically optimal. Seasoned artists with consummate techniques usually focus on interpretive considerations; for example, Schnabel was famous for sacrificing digital expediency for interpretive integrity, regarding hand position ("handing") as more important than fingering (Wolff, 1972). Thus a teacher's best fingering is not necessarily best for a student, and while it is always advisable to extend the student's awareness of available fingerings, it may be unwise to expect rigorous imitation.

Acknowledgments. The first author of this chapter thanks Diana Weekes, his piano teacher from the University of Melbourne, for her copious, insightful, and challenging comments on this chapter, and another pianist and piano teacher, Patrick Holming, for ideas, discussions, and literature sources. For technical comments, we thank Anders Askenfelt, Alexander Galembo, Werner Goebl, Hans Christian Jabusch, and authors of other chapters. The second author acknowledges Alberto Guerrero, with whom he studied in Canada alongside Glenn Gould, and Walter Gieseking, with whom he studied in Germany.

### *References*

Askenfelt, A. (1993). Observations on the transient components of the piano tone. In A. Friberg, J. Iwarsson, E. Jansson, & J. Sundberg (Eds.), *Proceedings of the Stockholm Music Acoustics Conference* (pp. 297–301). Stockholm: Royal Swedish Academy of Music.

Askenfelt, A., & Jansson, E. V. (1990). From touch to string vibrations. I . Timing in the grand piano action. *Journal of the Acoustical Society of America, 88,* 52–63.

Askenfelt, A., & Jansson, E. V. (1991). From touch to string vibrations. II: The motion of the key and hammer. *Journal of the Acoustical Society of America, 90,* 2383–2393.

Banowetz, J. (Ed.) (1985). *The pianist's guide to pedaling.* Bloomington: Indiana University Press.

Baron, J. G. (1958). Physical basis of piano touch. *Journal of the Acoustical Society of America, 30,* 151–152.

Baron, J., & Hollo, J. (1935). Kann die Klangfarbe des Klaviers durch die Art des Anschlages beeinflusst werden? *Zeitschrift für Sinnesphysiologie, 66,* 23–32.

Behne, K.-E. (1990). "Blicken Sie auf die Pianisten?!": Zur bildbeeinflussten Beurteilung von Klaviermusik im Fernsehen. *Medienpsychologie, 2,* 115–131.

Bregman, A. S. (1990). *Auditory scene analysis.* Cambridge, MA: MIT.

Breithaupt, R. M. (1905). *Die natürliche Klaviertechnik.* Leipzig: Kahnt.

Casella, A. (1936). *Il pianoforte.* Rome: Ricordi.

Clarke, E. F., Parncutt, R., Raekallio, M., & Sloboda, J. A. (1997). Talking fingers: An interview study of pianists' views on fingering. *Musicae Scientiae, 1,* 87–107.

Cortot, A. (1958). *Principes rationnels de la technique pianistique.* Paris: Salabert.

Deppe, L. (1885). Armleiden des Klavierspielers. *Deutsche Musiker Zeitung,* p. 325.

Deutsch, L. (1959). *Piano sight-reading.* Chicago: Nelson-Hall.

Eisenberg, J. (1928, June). Noted Russian pianist urges students to simplify mechanical problems so that thought and energy may be directed to artistic interpretation. *Musician,* p. 11. http://users.bigpond.net.au/nettheim.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47,* 381–391.

Fletcher, H. (1934). Loudness, pitch, and timbre of musical tones and their relations to the intensity, the frequency, and the overtone structure. *Journal of the Acoustical Society of America, 6,* 59–69.

Fletcher, N. H., & Rossing, T. D. (1998). *The physics of musical instruments* (2nd ed.). New York: Springer.

Frey, E. (1933). *Bewusst gewordenes Klavierspiel und seine technischen Grundlagen.* Zurich: Hug.

Galembo, A., Askenfelt, A., & Cuddy, L. L. (1998). *On the acoustics and psychology of piano touch and tone.* Paper presented at Sixteenth International Congress on Acoustics, Seattle. Abstract: *Journal of the Acoustical Society of America, 103,* 2873.

Gát, J. (1965). *The technique of piano playing* (3rd ed.). Budapest: Corvina.

Gellrich, M., & Parncutt, R. (1998). Piano technique and fingering in the eighteenth and nineteenth centuries: Bringing a forgotten method back to life. *British Journal of Music Education, 15*(1), 5–24.

Gerig, R. R. (1974). *Famous pianists and their technique.* Washington, DC: Luce.

Goebl, W. (2001). Melody lead in piano performance: Expressive device or artifact? *Journal of the Acoustical Society of America, 110,* 563–572.

Hall, D. E. (1991). *Musical acoustics.* Pacific Grove, CA: Brooks/Cole.

Hart, H. C., Fuller, M. W., & Lusby, W. S. (1934). A precision study of piano touch and tone. *Journal of the Acoustical Society of America, 6,* 80–94.

Heinlein, C. P. (1929a). A discussion of the nature of pianoforte damper-pedalling together with an experimental study of some individual differences in pedal performance. *Journal of General Psychology, 2,* 489–508.

Heinlein, C. P. (1929b). The functional role of finger touch and damper-pedalling in the appreciation of pianoforte music. *Journal of General Psychology, 2,* 462–469.

Heinlein, C. P. (1930). Pianoforte damper-pedalling under ten different experimental conditions. *Journal of General Psychology, 3,* 511–528.

Henderson, M. T. (1936). Rhythmic organization in artistic piano performance. In C. E. Seashore (Ed.), *Objective analysis of musical performance, University of Iowa Studies in the Psychology of Music,* Vol. 4 (pp. 281–305). Iowa City: University of Iowa Press.

Hill, W. G. (1940). Noise in piano tone. *Musical Quarterly, 26*, 244–259.

Houtsma, A. J. M., Rossing, T. D., & Wagenaars, W. M. (1987). *Auditory demonstrations on compact disc.* New York: Acoustical Society of America. http://asa.aip.org/discs.html.

Huron, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception, 19*, 1–64.

Jacobs, J. P. (2001). Refinements to the ergonomic model for keyboard fingering of Parncutt, Sloboda, Clarke, Raekallio, and Desain. *Music Perception, 18*, 505–511.

Kase, S. (2001). Clusters: The essence of musicality and technique. *Piano Journal, 22*(64), 13–18.

Kochevitschy, G. (1967). *The art of piano playing: A scientific approach.* Evanston, IL: Summy-Burchard.

Koornhof G. W., & van der Walt, A. J. (1993). The influence of touch on piano sound. In A. Friberg, J. Iwarsson, E. Jansson, & J. Sundberg (Eds.), *Proceedings of the Stockholm Music Acoustics Conference* (pp. 318–324). Stockholm: Royal Swedish Academy of Music

Martin, D. W. (1947). Decay rates of piano tones. *Journal of the Acoustical Society of America, 19*, 535–541.

Matthay, T. (1903). *The act of touch.* London: Longmans.

Neuhaus, H. (1973). *The art of the piano* (K. A. Leibovitch, Trans.). London: Barrie & Jenkins.

Ortmann, O. (1925). *The physical basis of piano touch and tone.* New York: E. Dutton.

Ortmann, O. (1929/1981). *The physiological mechanics of piano technique.* New York: Da Capo.

Ortmann, O. (1937). Tone quality and the pianist's touch. *Proceedings of the Music Teachers National Association*, 127–132.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 331–346.

Palmer, C. (1996). On the assignment of structure in music performance. *Music Perception, 14*, 21–54.

Palmer, C., & Brown, J. C. (1991). Investigations in the amplitude of sounded piano tones. *Journal of the Acoustical Society of America, 90*, 60–66.

Parlitz, D., Peschel, T., & Altenmüller, E. (1998). Assessment of dynamic finger forces in pianists: Effects of training and expertise. *Journal of Biomechanics, 31*, 1063–1067.

Parncutt, R., & Holming, P. (2000, August). Is scientific research on piano performance useful for pianists? Poster presented at the Sixth International Conference on Music Perception and Cognition, Keele University, UK.

Parncutt, R., Sloboda, J. A., & Clarke, E. F. (1999). Interdependence of right and left hands in sight-read, written, and rehearsed fingerings of parallel melodic piano music. *Australian Journal of Psychology, 51*, Special Issue: Music as a Brain and Behavioural System, 204–210.

Parncutt, R., Sloboda, J. A., Clarke, E. F., Raekallio, M., & Desain, P. (1997). An ergonomic model of keyboard fingering for melodic fragments. *Music Perception, 14*, 341–382.

Rasch, R. A. (1978). The perception of simultaneous notes such as in polyphonic music. *Acustica, 40*, 21–33.

Rasch, R. A. (1979). Synchronization in performed ensemble music. *Acustica, 43*, 121–131.

Repp, B. H. (1996). Patterns of note onset asynchronies in expressive piano performance. *Journal of the Acoustical Society of America, 100*, 3917–3932.

Repp, B. H. (1997). Acoustics, perception, and production of legato articulation on a computer-controlled grand piano. *Journal of the Acoustical Society of America, 102*(3), 1878–1890.

Repp. B. H. (1999). Effects of auditory feedback deprivation on expressive piano performance. *Music Perception, 16*, 409–438.

Rosen, C. (1995). *The romantic generation*. Cambridge, MA: Harvard University Press.

Rubinstein, A. (1980). *My many years*. New York: Knopf.

Schnabel, K. U. (1950). *Modern technique of the pedal*. London: Mills Music.

Schultz, A. (1949). *The riddle of the pianist's finger*. New York: Fischer.

Seashore, C. E. (1937). Piano touch. *Scientific Monthly* (New York), *45*, 360–365.

Shimosako, H., & Oghushi K. (1996). Interaction between auditory and visual processing in impressional evaluation of a piano performance. *Journal of the Acoustical Society of America, 100*(4), pt. 2, 2779.

Singh, P. G., & Hirsh, I. J. (1992). Influence of spectral locus and $F_0$ changes on the pitch and timbre of complex tones. *Journal of the Acoustical Society of America, 92*, 2650–2661.

Sloboda, J. A., Clarke, E. F., Parncutt, R., & Raekallio, M. (1998). Determinants of fingering choice in piano sight-reading. *Journal of Experimental Psychology: Human Perception and Performance, 24*(3), 185–203.

Terhardt, E. (1974). On the perception of periodic sound fluctuations (roughness). *Acustica, 30*, 201–213.

Terhardt, E., Stoll, G., & Seewann, M. (1982). Pitch of complex tonal signals according to virtual pitch theory: Tests, examples and predictions. *Journal of the Acoustical Society of America, 71*, 671–678.

Troup, M. (1983). Regard sur Olivier Messiaen. *Piano Journal, 4*(11), 11–13.

Troup, M. (1995). Orchestral music of the 1950s and 1960s. In P. Hill (Ed.), *The Messiaen companion* (pp. 392–447). London: Faber & Faber.

Vernon, L. N. (1936). Synchronization of chords in artistic piano music. In C. E. Seashore (Ed.), *Objective analysis of musical performance, University of Iowa Studies in the Psychology of Music*, Vol. 4 (pp. 306–345). Iowa City: University of Iowa Press.

Wagner, C. (1988). The pianist's hand: Anthropometry and biomechanics. *Ergonomics, 31*, 97–132.

Wead, C. K. (1921). Acoustical notes. *Science, 54*, 467–469.

Weinreich, G. (1977). Coupled piano strings. *Journal of the Acoustical Society of America, 62*, 1474–1484.

Wolff, K. (1972). *The teaching of Artur Schnabel*. London: Faber.

# 19

# String Instruments

KNUT GUETTLER & SUSAN HALLAM

Research on the physics of bowed stringed instruments can help the string teacher to explain the underlying acoustical phenomena and to develop corresponding pedagogical strategies. The first part of this chapter surveys current and historical acoustical research, focusing on information that can be related to technique. This section discusses not only the bowed attack with specific exercises to improve performance, but also other topics such as harmonics, rosin, timbre, and aspects of room acoustics during performance. An overview of psychological research relating to the distinctive aspects of playing a bowed stringed instrument and the characteristics of string players follows. This considers the importance of well developed aural skills, practice, and conscientiousness on the part of the player to develop high levels of expertise. The need for the teacher to demonstrate and provide opportunities to develop aural schemata, and give detailed constructive feedback is also discussed.

The first serious acoustical research on the violin can be dated back to the first half of the nineteenth century, when the French physicist Felix Savart borrowed from violin maker Jean Baptiste Vuillaume free top and back plates of violins crafted by Stradivari, Guarneri, and others, to examine how they vibrated at different frequencies (Savart, 1819, 1840). Contrary to popular belief, Savart discovered that the sound post does not function primarily to transmit the vibrations from the top to the back of the violin. Rather, its main function is to impose a nearly stationary point about which the top plate can rock.

The next major step toward an understanding of the acoustics of the violin was taken by Hermann von Helmholtz (1862/1954). Through acoustical filtering he was able to break tones down into pure tone components or partials (i.e., fundamental plus overtones), whose relative amplitudes influence timbre. In this

way, he could study each partial separately. Using a microscope, vibrating with nearly the same frequency as the string, he could watch the string in slow motion and show that bowing and plucking the string produced quite different patterns of vibration. With this method, he discovered the *slip-stick* action between the bow and the string, now called *Helmholtz motion*, which underlies the sawtooth-shaped force waveform acting on the bridge.

In the first part of the twentieth century, scientists from many countries contributed to research on the violin. While most of these focused on the instrument body and its vibrations, varnish, and wood, an Indian, Chandrasekara V. Raman, was particularly interested in the parameters of bowing technique: bow speed, bowing pressure (the correct physical term being *force*), and the bow-to-bridge distance, which together control the string wave form. Raman (1920) constructed a mechanical bowing machine that enabled him to measure the minimum required bowing pressure for any given pitch on the instrument, a value he found to be inversely proportional to the square of the bow-to-bridge distance. Interestingly, Raman found that particularly resonant tones required higher bowing pressure than less resonant ones. For resonant tones, more low-frequency wave energy is drained from the string through the bridge to the instrument's body, so the string will not maintain its fundamental-frequency oscillation with the same ease as for less resonant tones. In most cases, increasing the bow pressure can compensate for this energy loss. (String players normally adjust these bowing parameters intuitively, but if the sound is unsatisfactory, intuition does not always provide guidance as to what is wrong.) When the energy loss surpasses a certain magnitude, a steady-state oscillation is impossible at any bow pressure. Such tones are commonly termed *wolf tones* due to their (sometimes) howling character.

In the early 1950s, four Americans formed a group, jokingly calling themselves the Catgut Acoustical Society. Their ambition was "to increase and diffuse the knowledge of musical acoustics and to promote construction of fine stringed instruments" (currently used as their journal's cover-page motto). The organization now has worldwide membership comprised of scientists, luthiers, and musicians, who communicate through semiannual journals. Their publications provide a wealth of information for anyone interested in learning about string-instrument acoustics.

Research in the 1970s returned to the bow-string interaction. McIntyre and Woodhouse (1979) and Schumacher (1979) made the first computer simulations of the bowed string. This led to the discovery of the important role of *string torsion* (twisting or rolling under the bow) during the *attack phase* (defined as the first part of the tone, before it has settled into a steady vibration, or the transitional state between rest and the sustained tone). Without some torsion, strings would be much harder to bow. Computer simulations also increased the demand for quantification of physical string properties (Pickering, 1985) and knowledge about different string materials' effect on tone quality (Pickering, 1990). The results contributed to the design and manufacture of modern, well-responding strings.

## Fundamentals of the Bow-String Contact

### *Helmholtz Motion*

In a typical bowed tone, the string alternatively *sticks* to and *slips* against the bow hair. There is one *sticking interval* and one *slipping interval* per period of the tone (i.e., 440 times per second on an open violin A string). After the attack phase, the string takes the shape of two almost straight lines joined in a kink or corner, which follows a lens-shaped path (see Figure 19.1). On a violin A string, the corner travels along the string at a speed of approximately 290 meters per second. As the corner passes the bow in the direction of the bridge, the bow loses its grip (stick) on the string and begins to fly back (slip) along the bow hair. During the slipping interval, the corner reaches the bridge and is reflected back in the direction of the bow. At the instant the bow is passed again, the string portion under the bow starts moving much slower relative to the hair, allowing the string to stick again and stay stuck for the remaining, longest part of the cycle period (provided no other disturbance occurs). This particular string movement is called the Helmholtz motion.

In a violin upbow, the rotation along the string (relative to the player) is counterclockwise; in a downbow, clockwise. Thus a new rotational orientation has to be developed after each bow change. Bowing technique is primarily concerned with the creation and maintenance of Helmholtz motion.

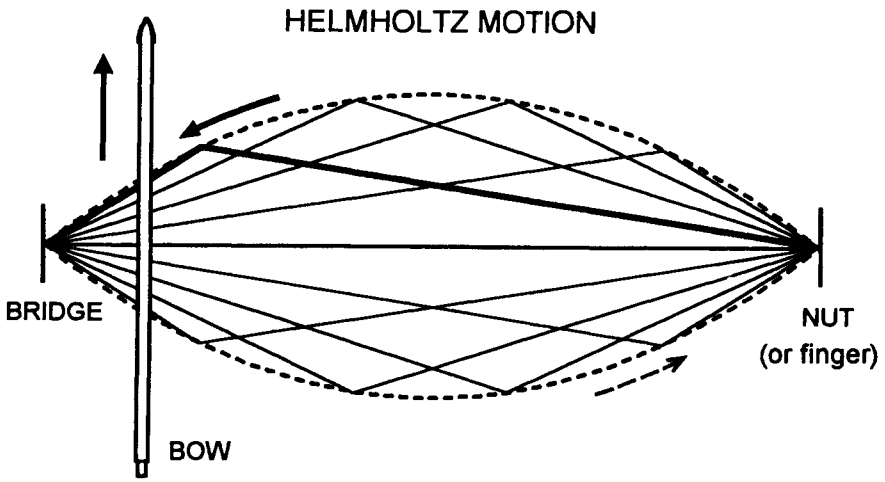



Figure 19.1. When the string is bowed in a steady state, the (idealized) string takes the form of two straight lines connected by a rotating corner. When the corner is traveling from the bow to the bridge and back, the string slips on the bow hair. In the remaining interval, the string sticks to the hair, if bowed properly.

Figure 19.2 (a) shows, in somewhat idealized form, the force exerted by the string on the bridge as a function of time. This force tries to move the bridge sideways and, as a result of the bridge's downward force shifting from leg to leg, causes the instrument body to vibrate, thereby creating sound.

### *Other Cyclic String Patterns*

The Helmholtz mode is only one of several possible modes of vibration, and the one that produces the clearest pitch and the smoothest, fullest timbre. When more than one slipping interval occurs per cycle (producing more than one corner on the string at a given moment in time), the lower partials are weakened, and the sound becomes scratchy, with increased slipping noise (see "Hiss (Rosin Noise)" later). Figure 19.2 (b) shows the force wave with one extra slip per cycle. An easy way to recognize this mode is to look at the midpoint of the string, where its amplitude appears constricted during the attack phase. In most cases, a multislip pattern soon dies out as the corners are softened by string stiffness and the following energy absorbers:

- The compliant, not-totally-reflecting bridge (transferring energy to the instrument body)
- Damping by the soft finger pad
- The internal friction of the string (e.g., friction between windings, molecular friction)
- The drag of surrounding air

The faster some high-frequency wave energy is lost in these ways, the sooner the Helmholtz motion will form, provided that the bowing speed and pressure are in the right range. The concept of energy loss can explain why open strings, strings not yet broken in (having less internal friction), and strings with great ability to sustain *pizzicato* (ditto) are more susceptible to whistling and onset noise when bowed.

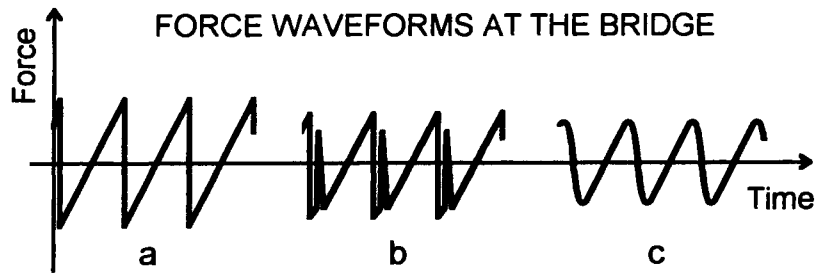



Figure 19.2. Different force waveforms acting sideways on the bridge in the plane of bowing (about three periods shown of each): (a) brilliant sound, (b) scratchy tone, (c) soft timbre that results from low bowing pressure and a rounded corner.

### *Rounding off the Helmholtz Corner*

In practice, the string can never make a perfectly sharp corner. String stiffness and the losses mentioned earlier round it off. However, the bow will, to some degree, sharpen the corner every time it passes, depending on the bowing pressure. So as long as the player does not choke the string, a higher bowing pressure means a sharper corner and thus—according to the basic mathematical theory of Fourier analysis—a higher energy content of high partials. Correspondingly, lower bowing pressure encourages rounding and a softer tone color (see Figure 19.2 c).

### *Hiss (Rosin Noise)*

During the slipping interval, an audible hiss or buzz is produced when the string slides on the rosined bow hair. Physically this consists of short bursts of noise (one per period of the fundamental frequency) and so is referred to as pulsed noise. This is an important contributor to the characteristic timbre of stringed instruments. In the scientific literature, the bow-to-bridge distance divided by the length of the vibrating string is termed $\beta$. (For example, if the bow is midway between the finger and the bridge, $\beta = 0.5$.) This ratio determines the duration of the hiss pulse relative to the cycle period. By bringing the bow nearer to the fingerboard (*sul tasto*), one increases $\beta$ and the time proportion of hiss. Very low bowing pressure (as in *flautando*) lengthens the slipping period even further, adding even more breathiness to the sound.

Noise may also be generated during the sticking interval, especially on the violin, when bowed near the bridge. This largely undesirable spiky quality is particularly audible at high bowing pressure and is caused by minor local string slips across the hair-ribbon width (McIntyre, Schumacher, & Woodhouse, 1981; Pitteroff, 1995). Tilting the bow reduces the width and minimizes the effect.

### *Getting the String into Helmholtz Motion*

So far, only the steady-state (postattack) oscillation of the string has been discussed. To get the string into Helmholtz motion quickly, a good strategy is to aim for a regular stick-slip pattern from the very onset. The violin body has a *response time* of the order of 5–15 ms, but the bowed string normally needs considerably longer to reach full amplitude (Melka, 1970). (Response time is here defined as the time required to reach 90% of full amplitude if a steady-state input signal is abruptly switched on.) The attack of bowed instruments is relatively slower than that of most wind instruments, making the string sound appear later than the wind, even when they start simultaneously. This delay is most noticeable at lower pitches (e.g., in the double basses), because transient duration varies inversely with fundamental frequency (so, for example, a tone an octave lower takes twice as long to sound fully). The string buildup time is, however, partly habitual (having to do with wrist action, etc.) and may, if desirable, be shortened through practice. We will return to this point later. First we will consider of the production of clean tone onsets.

During the attack phase, the motion of the string, for a given fundamental frequency and bow-to-bridge distance, is primarily determined by bowing *pressure* and *acceleration*. After the attack, when full amplitude has been reached, string motion is determined primarily by bowing *pressure* and *speed*. Based on computer simulations (Askenfelt & Guettler, 1998), Figure 19.3 shows schematically how the cleanness of the attack is influenced by bowing pressure and acceleration. For each relative bow-to-bridge distance ($\beta$), the noise-free region forms a (white) triangle in the figure, from which the following conclusions can be drawn:

- A certain minimum bowing pressure is required to start the proper slip-stick triggering directly (typically in the order of 0.1–0.2 Newton, comparable to 10–20 grams, for a violin).
- The required bow *acceleration* increases with the bow-to-bridge distance if the bowing pressure is kept unchanged (compare the positions of the rectangles drawn in the two panels of Figure 19.3).
- Near the minimum bowing pressure, the acceleration *tolerance* (i.e., the range of acceleration values that will produce noise-free attacks) is very small, but it increases with bow pressure, and with the bow-to-bridge distance (compare the widths of the rectangles drawn in the two panels of Figure 19.3).

In addition:

- For noise-free attacks with a given string, the required bow acceleration increases proportionally with the fundamental frequency and so does the acceleration tolerance.

The latter means that bowing can be varied more freely for high-pitched in-

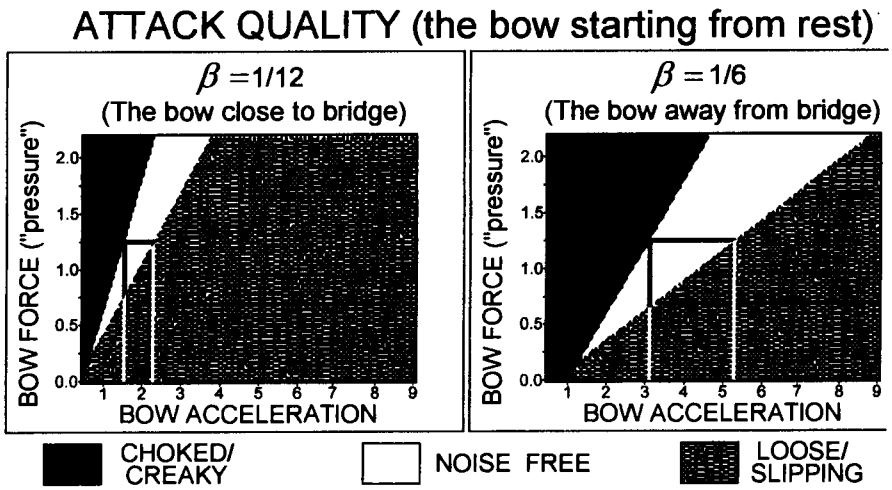



Figure 19.3. The combination of bowing pressure and acceleration (both scaled in arbitrary units) determines the quality of the attack. NOISE FREE means regular slip-stick triggering from the very onset. Within this (white) area, high acceleration gives the quickest tone buildup. Rectangle widths show the acceleration tolerances for a given bowing pressure.

struments than for their low-pitched counterparts. In order to make a quick tone buildup for a given bow pressure and position (i.e., reaching full amplitude as quickly as possible), the player must try to find the maximum (farthest to the right in the figure) acceleration within the noise-free (white) area. Only through practice and trial and error can this limit be found. The lower limit (i.e., the *lowest* noise-free acceleration) might also be practiced, as preparation for smooth bow changes.

If non-Helmholtzian waveforms are created, these will prevail for a number of *periods* before expiring, implying a longer lasting onset noise for low-pitched strings than for high-pitched ones (Guettler & Askenfelt, 1997).

### *How to Control Bow Acceleration in the Attack Phase*

*The generalizations outlined earlier need only be considered when the player is trying to account for problems in attack. For deciding what sounds best, the ear is always superior.*

Many players prefer to start the stroke from the air, deliberately producing an onset with a certain airy or slippy sound. However, for low-pitched strings such an airy introduction may be too persistent, especially in *forte*.

In practice, acceleration is not an easy concept to grasp or to teach. Newton's second law, force equals mass times acceleration, suggests the following approach: When one is playing near the bridge, an almost constant acceleration can be achieved by adding mass to the bow, so that small deviations in the frictional force will have less effect on the bowing speed. This can be achieved by temporarily reducing the suppleness of one or more joints in order to add the mass of the fingers, hand, forearm, upper arm, and torso to the mass of the bow itself. (When one is doing this, bowing pressure should not be influenced!) For low-pitched tones and/or playing close to the bridge, this technique effectively reduces the influence of a rapidly fluctuating frictional force, which could otherwise make the acceleration uneven. The technique works best on heavier strings, where the magnitude of frictional fluctuations is greatest. A posture that allows for flexibility of the upper body is essential for a consistent and well-controlled bow-arm movement on any string instrument. For a double-bass player, it is particularly important not to let the instrument's body lock the chest in a fixed position.

**Practice Suggestion:** The student should learn the arm/finger action that produces noise-free attacks when making repeated attacks close to the bridge. The following should be varied: (1) the suppleness of the arm/finger joints, (2) the dynamic level, (3) the bow's starting point (frog-middle-tip), and (4) the pitch.

### *Control of String Vibration during the Steady-State Phase*

As soon as a correct Helmholtz pattern is established, the player has a wide palette of timbres to choose from. For a given bow speed, different bow pressures may subsequently produce anything from "harmonics" and *flautando*, via "loose,"

"firm," and "pressed" timbres, to "raucous" and "creaky" sounds (Schelleng, 1970). As long as some kind of Helmholtz motion is present, the fundamental frequency will be well defined, so the pitch will be clear. When bowing pressure is far too high or the speed far too low, the sound may best be described as raucous or creaky. However, before raucousness appears, there is a point where the sound is pressed and the fundamental frequency is slightly reduced (flat), because the sticking rosin delays the string release (i.e., the transition from sticking to slipping phase) at excess bowing pressure. The effect is more pronounced for softer rosins, larger distances of the bow from the bridge, and shorter lengths of the vibrating strings, which increases the effective string stiffness (e.g., high positions on a violin G string). Moving the bow closer to the bridge hence reduces the flattening. As the bow-to-bridge distance becomes smaller, the palette of different timbres shrinks. Near the bridge, the bowing pressure must be held close to a single theoretically defined value (which depends on the chosen bow speed) if Helmholtz motion is to result. This means that when one is playing near the bridge fluctuations in bowing speed and pressure must be kept to a minimum.

### *Implications for Performance*

When one is performing a *diminuendo*, both pressure and speed of bowing must be reduced simultaneously if a consistent tone quality is to be achieved—regardless of bow-to-bridge distance. Simplified, the ideal situation occurs when bowing speed is balanced against the combination of bowing pressure, relative bow-to-bridge distance ($\beta$), and fundamental frequency. Whenever one element is changed, compensation by one or more of the remaining three is normally required.

**Example 1.** If one keeps the bow pressure and distance to the bridge fixed, playing the same pitch on alternative strings will require a higher bow speed on the lowest tuned string. This is because the relative bow-to-bridge distance is greater for the string fingered in the highest position. (In some cases, this rule of thumb will not apply to a crossing between the third and the fourth string, because the lowest string is often disproportionately heavier in order to allow for a bigger sound in the instrument's lowest register. Such strings require lower bow speed and acceleration than others.)

**Example 2.** In double-stops that span large intervals, the highest sounding string should be given a relatively lower bowing pressure to prevent it from sounding pressed or flattened. (In these cases, both $\beta$ and the fundamental frequency are higher for the upper tone, regardless of which string it is played on.)

**Practice Suggestion:** (1) Play several strokes, keeping two elements (e.g., speed and pressure) fixed while varying the third one (e.g., bow-to-bridge distance). After doing so, try to repeat the combination that gave the best sound, directly. Practice on different strings and at different pitches. (2) Play pairs of tones, two

low and two high, in intervals of an octave or larger. Notice the difference required in bow speed.

It is the bow speed that controls the string amplitude for a given bow-to-bridge distance and hence largely the physical loudness of a tone, which is mainly determined by the (loudest-sounding) lower partials. Increasing the bow pressure moderately may produce a slightly brighter sound but rarely increases the total sound level more than one or two decibels before the sound becomes pressed. However, bright sounds are usually *perceived* as loud. For example, when listening to the radio we are never in doubt of an orchestra's dynamic level, regardless of the listening volume. A change of bowing position toward the bridge increases the brilliance and may give the impression of a louder sound, even when compensated for by a lower bow speed.

### *The Smooth Bow Change*

The rotational orientation of the Helmholtz corner is determined by the stroke direction. So when one starts a new stroke in the opposite direction, the whole transient process needs to begin again. Any release of the bow pressure during this phase implies a slower buildup of the new tone, as well as reduced brilliance. This type of temporary change in timbre is less noticeable in the violin than the double bass, due to the longer transient of the latter. If a brilliant, noise-free sound is to be maintained during the entire bow change, a stable bow pressure, combined with suitable bow acceleration, must be observed.

It is important to minimize the remains of the old corner rotation before developing a reversed one to avoid slipping noise caused by interference between two opposing corners. The bow *deceleration* that precedes the change of direction can be of considerably greater magnitude than the acceleration that follows. This has been confirmed by measuring the smooth bow changes of skilled players (Williams, 1985).

Practice Suggestion: Play bow changes at different dynamic levels and different bow-to-bridge distances. Experiment to find the wrist action, or suppleness, that works best in each case.

### Spiccato *and* Ricochet

Computer simulations show that the requirements of bowing pressure and acceleration (as described in Figure 19.3) are met during clean *spiccato* attacks and that the bow quickly dampens the string prior to each new attack (Guettler & Askenfelt, 1998). If one places a small mark at the bow-stick midpoint, a horizontal figure of eight ($\infty$) should be seen when one watches the mark during rapid *spiccato*. Any other figure is likely to produce some onset noise. In *ricochet*, where several strokes are performed in the same direction, the remains of the old Helmholtz pattern can easily be picked up and refreshed without any predamping, minimizing the buildup time. The same goes for other linked strokes

like *slurred staccato* and *portato*. This is why such attacks often sound cleaner than normal *spiccato* and *détaché*.

### *"Harmonics" (Flageolet Tones)*

"Harmonics" are played by placing the finger at certain points of the string, thus filtering out the fundamental and some of its partials and allowing only some selected partials (overtones) to vibrate. (In physics, however, a *harmonic* means a frequency that is an integer multiple of the fundamental.) When harmonics are played, the string is almost motionless at certain points called nodes. Between these, the string behaves like several short strings, each of which exhibits Helmholtz motion (see Figure 19.4). The nodes are equally spaced along the string. This implies that on a string the stretch between the pressing and the touching finger in an "artificial harmonic" will remain the same whether it is performed as an octave, fifth, fourth, or major-third "harmonic" (see lower part of Figure 19.4). However, the highest position will give the shortest transient, due to earlier reflections between the bridge and the touching finger. A harmonic cannot be excited at the node and requires bowing pressure, speed, and acceleration comparable to that of a fingered tone of the same frequency. This is easily forgotten when combining harmonics with lower-pitched tones on the same string.

### *Rosins*

Rosin temporarily melts in the near vicinity of a sliding string, increasing the difference in frictional force between stick at maximum hold, and slip. It is this difference (its span often referred to as the friction delta)—not the absolute friction values—that uniquely defines the harmonic spectrum for a given bow speed and position on a mounted string. Statements such as "You should be playing with the scales," ". . . barbs," or even ". . . teeth of good-quality hair, not the rosin"

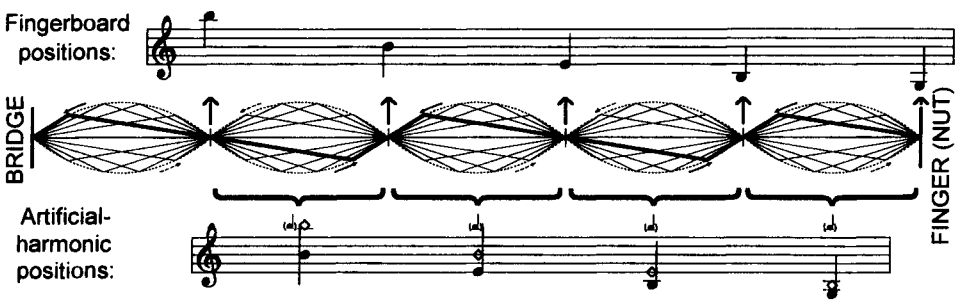



Figure 19.4. When one is playing harmonics, the string forms a series of Helmholtz figures separated by nodes. The distance between adjacent nodes equals one-half the wavelength of the harmonic frequency (here $B_5$).

are therefore scientifically unjustified. However, as advice to *limit the amount* of rosin such statements make some sense, since a thick layer of rosin tends to cause uneven friction characteristics. The hair should be judged from its ability to hold rosin evenly, rather than from its ability to cause mere friction, to which scale thickness of no more than 1/200 mm will not contribute much anyway (Rocaboy, 1990). The friction delta increases with the bowing pressure and with the softness of the rosin and decreases with increasing string weight (physical term: *wave resistance*). Because of this, soft rosins are often preferred for low-pitched instruments with heavier strings.

### *Pizzicato*

*Pizzicato* and *arco* produce completely different string movements, as *pizzicato* involves no corner rotation, only a simple side-to-side movement, symmetrical over the string's midpoint. Although different, their spectra are both functions of the relative bow- (or finger-) to-bridge distance, with a large distance giving the highest proportion of fundamental-frequency amplitude and hence the longest sound-buildup time, while a small distance gives the brightest and most direct sound. High overtones tend to fade quickly. This depends, however, on both frequency and plucking *angle*, as the mobility, thus the individual overtone absorption, of the bridge varies with both. Test it!

### *Sound Radiation*

The vibrating string does not produce any audible sound if not coupled to a body of adequate size. In fact, even the bodies of the violin family do not properly radiate the fundamental frequencies of their lowest tones. For the double bass and cello, a compliant floor sometimes comes to the rescue, as low-pitched vibrations are transmitted through their spikes much like tuning forks on a tabletop. Each string instrument has two groups of resonances: one dominated by the enclosed air and one by the wooden body parts. The main air resonance is for violin about 270 Hz ≈ $C_4$–$C\sharp_4$, viola: 240 Hz ≈ $B\flat_3$–$B_3$, cello: 115 Hz ≈ $B\flat_2$, and double bass: 65 Hz ≈ $C_2$, with their main body resonances a good fifth higher (Hutchins, 1962). (A wolf tone would most likely be situated near the latter, by which the top, and hence the bridge, moves considerably.)

The number of resonances per octave rapidly increases with frequency. In *vibrato*, many string partials temporarily coincide with body resonances and are thus well transmitted—while others fall between the resonances. This continuously varying timbre brings out a harmonic totality that is otherwise imperceptible. (One may compare it to viewing a landscape through a picket fence: unless moving, one cannot see it all.) Another interesting feature is the *directional dispersion* of high tones, by which a good instrument may produce an illusion of multiple sound sources in a reflecting room (Weinreich, 1997). In general, different-frequency bands have different radiation characteristics, which are important to bear in mind when positioning the instrument with respect to microphones or even an audience (Meyer, 1972).

When one is performing a solo in a large hall, an early first reflection increases definition and clarity. Instead of approaching the very front edge of a wide stage, stepping back might be better, because it will help to mirror the instrument for the audience in the acoustically reflective stage floor. Brilliance is crucial for perceptibility of rapid passages and should be slightly boosted, for example by bowing closer to the bridge in dark-sounding or reverberant halls. In addition, any view-obstructing music stand will also obstruct high frequencies.

## Psychology of String Instrument Performance

### *Playing a Stringed Instrument*

So far we have considered evidence that relates to how sound is produced on bowed stringed instruments and the relationship between this and acoustics. We now consider research evidence that relates to string players and what is distinctive about playing a stringed instrument. To date, research in the psychology of music has tended to neglect these issues, assuming that musicians are a relatively homogeneous group and that what applies to one musician will apply to all. Increasingly researchers are recognizing that there may be important differences between musicians, but as yet there is little research evidence.

### *What Is Distinctive about Playing a Bowed Stringed Instrument?*

As the previous section infers, the technical requirements of the string repertoire are varied and demanding, requiring the acquisition of many skills that must become automated to enable fluent professional performance. To play a bowed stringed instrument requires sensitive motor control to enable secure intonation, clear articulation of bowing, and the coordination of two very different kinds of movements being undertaken by left and right hands. In addition, well-developed aural skills are required to monitor accuracy, as any finger can be used to play almost any note. The demands of the repertoire and professional work schedules also require high-level cognitive processing skills in relation to reading music, as the demands made on string players, particularly the first violins, tend to be greater than those made on other orchestral players. So what does the research evidence tell us about string players and the way they acquire their skills?

### *The Characteristics of String Players*

Using psychometric techniques, Kemp (1995) found that orchestral string players tended to demonstrate more aloofness, conscientiousness, and considerable willpower than other instrumentalists. He suggested that these characteristics were necessary to maintain the level of practice required to master the complexity of playing a stringed instrument. In a further breakdown of the personality profiles into specific instruments, at the school and student level, Kemp found that

secondary school violinists emerged as having a tendency to be easily upset and emotional. In contrast, cellists emerged as both self-sufficient, shrewd, and aloof, demonstrating a greater social awareness not apparent in the other instrumentalists. The viola players at student level were more emotionally stable than their fellow string players, while the double-bass players presented a profile that was typical of the overall string profile. These findings must be viewed with some caution, as the number in each group was relatively small, but they are relevant to pedagogy to the extent that those who are learning to play stringed instruments may need a higher level of commitment to practice than those who are learning other orchestral instruments (see also chapter 1 in this volume).

### *Personality and Practice*

How does the research on the personality of string players relate to what we know about their practice? Although there has been little research, such as there is supports the notion that string players, especially the violinists, undertake more practice than their orchestral peers (Jørgensen, 1997). This supports the findings outlined earlier. The reasons for the violinists' undertaking more practice may relate to the larger and more demanding repertoire. We know that there is a strong, if not perfect, relationship between the amount of time spent practicing and the level of expertise attained (Ericsson, Krampe, & Tesch-Römer, 1993; Hallam, 1998; Williamon & Valentine, 2000). Similarly, differences in practice patterns across instruments seem to reflect the demands of particular instruments. As practice draws on limited physical and mental resources, the amount of practice sustainable daily is limited by the individual's ability to recover. Four hours daily seems to be the maximum sustainable over long periods without burnout occurring (Ericsson et al., 1993; Hallam, 1995; Jørgensen, 1997). To accrue the additional hours of practice to achieve professional entry to a music conservatory, string players would need to start playing at a younger age than their wind- and brass-playing colleagues and violinists at an earlier age than other string players. While there is no formal evidence to support this, pianists (who undertake the most practice) and violinists who attain status as international soloists do tend to start playing at very early ages.

### *Learning to Play a Bowed Stringed Instrument*

Given the lack of other cues to inform feedback, it is crucial for string players to develop accurate aural *schemata* (mental representations) against which to evaluate their performance. The evidence suggests that this does not occur until they have developed considerable expertise (Hallam, 2001). Teachers can provide aural representations through modeling, but they tend to spend little time in demonstration, although a large proportion of the variance in performance can be accounted for by their willingness to model aspects of their playing (Sang, 1987).

Feedback has an important role in learning to play any instrument. To be effective, it needs to be appropriate and relevant to the task. Conditions of restricted and unrestricted vision, for instance, had no differential effect on string

players' intonation, while verbal feedback produced more accurate intonation than recorded or model performance feedback in university-level string players (Salzberg, 1980). Violin posture can be improved through the teacher providing corrective feedback (Salzberg & Salzberg, 1981), and biofeedback techniques, where musicians receive through electronic monitoring direct feedback that regards muscular activity, have proved very useful in enabling string players to learn to relax specific muscles as they are playing (Irvine & LeVine, 1981; LeVine & Irvine, 1984; Morasky, Reynolds, & Clarke, 1981). These techniques can be used while students perform passages from the repertoire they are learning (Cutietta, 1986).

### *Practical Applications*

What can we learn from the research findings in this chapter that will offer specific help to string players? Playing a stringed instrument is distinctive because of its relative difficulty and its particular reliance on well-developed aural skills. String players need to start playing young, undertake a considerable amount of practice, and be given numerous opportunities for listening to music and developing aural skills. Modeling and the provision of detailed verbal feedback are very important in teaching.

### *References*

Askenfelt, A., & Guettler, K. (1998). The bouncing bow: An experimental study. *Catgut Acoustical Society Journal, 3*(6), Ser. 2, 3–8.

Cutietta, R. (1986). Biofeedback training in music: From experimental to clinical applications. *Bulletin of the Council for Research in Music Education, 87*, 35–42.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363–406.

Guettler, K., & Askenfelt, A. (1997). Acceptance limits for the duration of pre-Helmoltz transients in bowed string attacks. *Journal of the Acoustical Society of America, 101*(5), pt. 1, 2903–2913.

Guettler, K., & Askenfelt, A. (1998). On the kinematics of spiccato and ricochet bowing. *Catgut Acoustical Society Journal, 3*(6), Ser. 2, 9–15.

Hallam, S. (1995). Professional musicians' orientations to practice: Implications for teaching. *British Journal of Music Education, 12*(1), 3–19.

Hallam, S. (1998). The predictors of achievement and drop out in instrumental tuition. *Psychology of Music, 26*(2), 116–132.

Hallam, S. (2001). The development of expertise in young musicians: Strategy use, knowledge acquisition and individual diversity. *Music Education Research, 3*(1), 7–23.

Helmholtz, H. von (1862/1954). *Lehre von den Tonempfindungen* (*On the sensations of tone*). New York: Dover.

Hutchins, C. M. (1962). The physics of violins. *Scientific American, 207*, 78–93.

Irvine, J. K., & LeVine, W. R. (1981). The use of biofeedback to reduce the left hand tension for string players. *American String Teacher, 31*, 10–12.

Jørgensen, H. (1997). Time for practicing? Higher level music students' use of

time for instrumental practicing. In. H. Jørgensen and A. C. Lehmann (Eds.), *Does practice make perfect? Current theory and research on instrumental music practice* (pp. 123–140). Oslo: Norges Musikhøgskole.

Kemp, A. E. (1995). Aspects of upbringing as revealed through the personalities of musicians. *Quarterly Journal of Music Teaching and Learning, 5*(4), 34–41.

LeVine, W. R., & Irvine, J. K. (1984). In vivo EMG biofeedback in violin and viola pedagogy. *Biofeedback and Self-Regulation, 9*, 161–168.

McIntyre, M. E., Schumacher, R. T., & Woodhouse, J. (1981). Aperiodicity in bowed-string motion. *Acustica, 49*(1), 13–32.

McIntyre, M. E., & Woodhouse, J. (1979). On the fundamentals of bowed-string dynamics. *Acustica, 43*(2), 93–108.

Melka, A. (1970). Klangeinsatz bei Musikinstrumenten. *Acustica, 23*, 108–117.

Meyer, J. (1972). Directivity of the bowed string instruments and its effect on sound in concert halls. *Journal of the Acoustical Society of America, 51*, 1994–2009.

Morasky, R. L., Reynolds, C., & Clark, G. (1981). Using biofeedback to reduce left arm extensor EMG of string players during musical performance. *Biofeedback and Self-Regulation, 6*, 565–572.

Pickering, N. C. (1985). Physical properties of violin strings. *Catgut Acoustical Society Journal, 44*, 6–8.

Pickering, N. C. (1990). String tone related to core material. *Catgut Acoustical Society Journal, 1*(5), Ser. 2, 23–28.

Pitteroff, R. (1995). *Contact mechanics of the bowed string.* Unpublished doctoral dissertation, University of Cambridge.

Raman, C. V. (1920). Experiments with mechanically played violins, Vol. 6, parts 1 and 2. *Proceedings of Indian Association of Cultivation of Science*, 19–36.

Rocaboy, F. (1990). The structure of bow-hair fibres. *Catgut Acoustical Society Journal, 2d, 1*(6), 34–36.

Salzberg, R. S. (1980). The effects of visual stimulus and instruction on intonation accuracy of string instrumentalists. *Psychology of Music, 8*(2), 42–49.

Salzberg, R. S., & Salzberg, C. L. (1981). Praise and corrective feedback in the remediation of incorrect left-hand positions of elementary string players. *Journal of Research in Music Education, 29*(2), 125–133.

Sang, R. C. (1987). A study of the relationship between instrumental music teachers' modelling skills and pupil performance behaviors. *Bulletin of the Council for Research in Music Education, 91*, 155–159.

Savart, F. (1819). *Mémoire sur la construction des instruments à cordes et à archet.* Paris: Determille.

Savart, F. (1840). Cours de physique expérimentale, professé au College de France pendant l'année scolaire 1838–1839. *L'Institut.*

Schelleng, J. C. (1970). Pressure on the bowed string. *Catgut Acoustical Society Newsletter, 13*, 24–27.

Schumacher, R. T. (1979). Self-sustained oscillations of the bowed string. *Acustica, 43*(2), 109–120.

Weinreich, G. (1997). Directional tone color. *Journal of the Acoustical Society of America, 101*(4), 2338–2346.

Williamon, A., & Valentine, E. (2000). Quantity and quality of musical practice as predictors of performance quality. *British Journal of Psychology, 91*(3), 353–376.

Williams, C. E. (1985). *Violin bowing skill analysis: The mechanics and acoustics of the change in direction.* Unpublished doctoral dissertation, University of Melbourne, Australia.

# 20

# Wind Instruments

LEONARDO FUKS & HEINZ FADLE

In mouth-blown wind instruments, the energy provided by the respiratory system is converted directly into sound. In all cases a primary vibrating element, generically called *reed*, controls the airstream. The reed may be a piece of bamboo, the lips, a metallic tongue, or even the air jet (in flutes and recorders). Players control loudness, attack, intonation, and timbre by means of embouchure settings, blowing pressure, airflow, and length of the air column. The respiratory muscles perform complex and systematic movements, generating wide ranges of pressures, and coordinated oscillations that produce the vibrato effect. Intonation may be affected by the characteristics of the lung air. We address the associated sensory, physiological, and acoustical phenomena. Common controversial or misleading concepts among wind players are discussed and some simple experiments are proposed for pedagogical applications.

Mouth-blown wind instruments convert pneumatic energy, in the form of air pressure and velocity, into sound waves. They belong to the category that Sachs (1940) describes as *aerophones*. Other aerophones include the human voice, the pipe organ, and the accordion.

Wind players use their respiratory apparatus as an air compressor to supply energy to the instrument. At the mouth, air is delivered through a sophisticated interface with the instrument, which musicians refer to as their *embouchure*. Each subtle action from muscles of expiration and embouchure imprints the sound, forming part of the musical expression. The arms, hands, and fingers skillfully control a complex of tone holes, slides, rotating valves, levers, and other mechanisms. Inside the instrument, several phenomena take place, including reflection, diffraction, resonance, shock waves, damping, vortex generation, and chaotic oscillations. Body posture, respiratory techniques, and a psychological strategy to deal with stressful situations are additional key aspects of perfor-

mance. This rich environment challenges the abilities of players and educators to deal with a number of principles that are associated with performance on wind instruments. Some of the most important are described in this chapter.

### *The Reed*

Every wind instrument has some mechanism that cyclically disturbs or interrupts the air stream. This mechanism can be generically called a *reed*. Therefore, rather than a component of bamboo or other elastic material, *reed* in this context stands for a valve that controls the air passing through the instrument, as shown in Figure 20.1. This concept has been consistently adopted by several authors in music acoustics (Campbell & Greated, 1987; Sundberg, 1991; Strong & Plitnik, 1992). The tongue, the lips, and the moving structures of the respiratory system may also be used to regulate the airflow, serving as additional valves.

Winds may incorporate a *mechanical reed* (reed woodwinds), *free reed* (mouth harmonicas), *lip reed* (brass instruments), or *air reed* (flutes and recorders), where one side is in contact with the player's mouth and the other attached to the instrument. Figure 20.1 shows four simplified models of reeds, each one representing a different type of oscillation. Because instruments such as the flute involve an air reed with no solid vibrating part, they are discussed separately.

As the air stream encounters resistance in passing the reed, a difference of pressure will be created between both sides. The pressure in the upstream (before the reed) is in principle higher than that of the downstream side due to friction losses. As the reed is flexible, it will tend to bend or move due to this pressure difference. In much the same way as a spring, a reed's elasticity will allow it to return to its original position and shape. The reed also contains mass that is distributed along its body. This defines a *spring-mass* system that has its own resonance frequencies.

In a similar manner to a tuning fork, a reed can be excited to oscillate at its own frequencies. This is the case in free reeds and, exceptionally, in clarinets

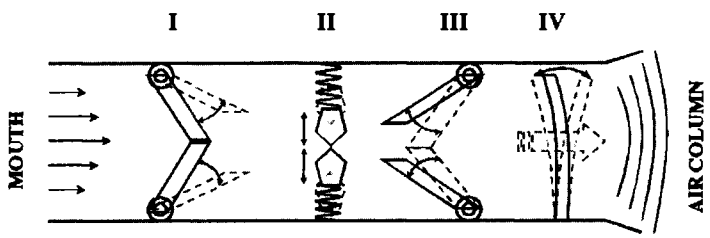



Figure 20.1. Oscillation types found in lip, mechanical, and free reeds: rest position of the reed (solid line) and extreme position during vibration (broken line). (I) outward-striking and (II) sideways- or upward-striking oscillations as found in brass instruments; (III) inward-striking oscillation, typical of reed woodwinds; and (IV) free-striking oscillations, which occur in free reeds.

that produce the so-called *squeak sounds* (Bouasse, 1929/1986). However, the tissue of the lips, forming a cushion pressed against the reeds, absorbs most of the reed energy in every cycle, with the oscillations being damped more than those of a tuning fork.

Another factor is the so-called *Bernoulli force*. When air passes through a narrow passage, the particles' velocity must increase in order to maintain a constant flow. This results in a drop in pressure past the narrow passage and a subsequent partial vacuum between air molecules. This negative pressure tends to suck the reed toward the closing position. If the closing reed blocks the air, this suction is immediately interrupted, releasing the reed to return to its original state.

In spite of being easily explained in general terms, the Bernoulli force is not important in all cases. It is certainly relevant in double reeds, especially the oboe, and for the higher range of brass instruments (Fletcher & Rossing, 1998). However, in the case of the clarinet its role is in most cases negligible, because it is not strong enough to make this instrument's reed vibrate (Worman, 1971).

### *The Air Column*

The *air column* is the length of vibrating air in the instrument during sound production. The airflow that crosses the reed creates pressure waves that travel along the air column toward its end and are then reflected back to the reed. In lip and mechanical reeds, if the reflected wave pressure is strong enough it will make the reed reverse its direction. At this point, the system is ready for a new cycle, as it establishes a chain of *self-sustained oscillations*. In an air reed, the air jet flows alternately into the column and out into the ambient, as a result of the waves set up in the air column (Benade & Gans, 1968).

The connection between the air column and the reed, technically referred to as acoustical coupling, forces the reed to vibrate at certain frequencies. When producing a steady tone, the reed vibrates with a frequency close to one particular mode of the air column, so that reflected waves arrive precisely when the reed, depending on the oscillation type, closes or opens. Benade (1990) described this situation as the reed being "enslaved" by the air column. Each configuration of tone holes, keys, pistons, valves, or slide in a wind instrument determines the length and shape of the air column.

A given air column resonates at a set of frequencies that are possible inside it. Each of these frequencies corresponds to a resonance mode. The set of tones obtained with a fixed air column should ideally coincide with those of a harmonic series. However, instruments usually deviate from the perfect geometry of a cylinder or a cone, so that the real modes also deviate from the exact harmonics. Wind players must develop the skill of matching the embouchure and blowing controls to select the desired mode of the air column, which will result in particular tones. Whenever the reed and the air column achieve a pattern of self-sustained oscillations, a *playing mode* is defined. *Overblowing* happens when the instrument jumps from a mode to a higher one.

The four models in Figure 20.1 are *pressure-controlled valves*, as pressure is the main controller of the reed movement. Only air reeds such as the flute are directly controlled by airflow (Benade & Gans, 1968).

While the instrument is being played, the sound waves generated in the reed propagate both to the tube and to the mouth cavity, toward the player's airways. However, most of this energy propagated into the mouth is lost by absorption through the multiple branches and soft walls of the respiratory apparatus. Nevertheless, the vocal tract may respond to some frequencies, as it does for the formation of the vowel sounds. In such cases, the vocal tract can be considered as an internal air column that interacts with the reed and the instrument's air column.

Some skilled players attempt to master the control of the vocal tract to improve their sound quality and stability and to help in the transition between playing modes. While some authors consider that a player's vocal tract has a negligible influence on an instrument's tone (e.g., Backus, 1985, in reed woodwinds), others (Benade, 1986; Johnston, 1987; Fletcher & Rossing, 1998) claim a considerable influence of the vocal tract in instruments such as the clarinet, saxophone, harmonica, and brass.

### *Types of Reed Oscillation*

Lip, mechanical, and free reeds perform characteristic types of oscillations according to their mechanical configuration, as shown in Figure 20.1. The airflow from the lungs crosses the reed, which is subject to forces provided by the embouchure. Helmholtz (1863/1954) referred to type I as an outward-striking and to type III as an inward-striking oscillation. Similarly, type II can be called upward- or sideways-striking and type IV a free-striking oscillation (Fletcher & Rossing, 1998).

In general, the playing frequencies in types II and III are lower than the resonant frequencies of the air column and of the reed itself, while in type I oscillations the playing frequency is higher than those of the air column and the reed alone (Fletcher & Rossing, 1998). Free reeds, corresponding to type IV, are not strongly coupled to an air column and essentially vibrate at their resonance frequencies.

### *The Embouchure*

The *embouchure* can be defined as the constellation of forces and positions in the lips, mouth region, and face that act on the wind instrument (Porter, 1967). It consists of the muscles that form and surround the mouth, the teeth, the upper jaw and mandible, and muscles responsible for mastication. The embouchure controls the reed behavior, thus affecting tone color, articulation, dynamics, and other parameters. In defining the embouchure, it is possible to include the internal mouth configuration, such as the positioning of jaw and tongue. In this sense, the mouth cavity can be regarded as a secondary air column.

In reed woodwinds, the embouchure regulates the reed parameters, such as the stiffness, the aperture for the airflow, the degree of absorption, and, to some

extent, its effective mass (Nederveen, 1969). The embouchure also provides an airtight attachment of the lips to the instrument. In brass instruments, the lips obviously serve both as a reed and as an embouchure component and include the same functions cited earlier. In the flute, the embouchure must be skillfully shaped and managed to ensure proper generation and control of the sound. In spite of the importance of the embouchure in performance and sound production, these latter aspects seem to have been unexplored and therefore deserve more research interest.

## Wind Instrument Families

### *Brass*

Stroboscopic observations of the lip vibrations in a professional trombone player indicate that the upper lip vibrates with greater amplitudes than the lower lip in all frequencies and dynamic levels (Copley & Strong, 1996). This study also showed that on average, two-thirds of the mouthpiece diameter is covered by the upper lip and one-third by the lower lip. However, considerable variations of this proportion can be observed in the general playing population, which suggests that players tend to optimize their embouchure during the process of learning.

Lip reeds correspond to types I or II in Figure 20.1. In both cases air pressure tends to open the lips, but in type II the Bernoulli forces act alone to close the lips back. Usually, the type I occurs in the lowest playing modes, with type II being more frequent in the higher modes (Yoshikawa, 1995). It is also possible that both types of motion take place in combination, as observed across the whole compass in a trombone study (Copley & Strong, 1996). Not all studies, however, found the same behavior, probably due to differences in the measurement methods that were employed and particular techniques that were adopted by individual players (Chen & Weinreich, 1996).

A common warm-up practice among brass players consists of buzzing on the mouthpiece or with lips only. This exercise is also claimed to develop flexibility and control of lip vibrations. However, according to Yoshikawa (1995), the isolated mouthpiece represents a completely different acoustical system from the full instrument bore, supporting only type I oscillations but not type II. This finding suggests that exercises of this type do not develop lip flexibility in the way often assumed by brass teachers.

The vibration frequency of the lips increases with a combination of higher muscular tension and lower mass. This *mass* refers to the section of the lip that effectively vibrates. Also, raising the pitch requires higher blowing pressure (Bouhuys, 1964). Therefore, it could be expected that the player would tend to press the lips more firmly against the mouthpiece when playing higher tones. This was corroborated by Barbenal, Davies, and Kenny (1986), who observed that the compressive forces between lips and mouthpiece increase systematically with increasing pitch in professional players. Once again, this finding contradicts some

performers' claim that it is possible to develop a rather "loose" and "fixed" embouchure, regardless of pitch and dynamic level. However, the idea of a loose embouchure may work as a mental image that helps to achieve the goal of minimizing the effort and strain of the embouchure muscles in order to preserve the necessary sensory response of the tissues.

### *Reed Woodwinds*

Mechanical reeds correspond to type III oscillations in Figure 20.1, where the airflow tends to blow them closed. In the oboe and in the bassoon, the reeds beat against each other almost all the time, completely interrupting the airflow in every oscillation. Similarly, in the saxophone, the reed tends to repeatedly beat against the mouthpiece. This occurs because the conical shape of those instruments creates comparatively high pressures where the reeds are located. In the clarinet, however, with its cylindrical shape, the reed does not necessarily beat against the mouthpiece for the relatively soft tones (Worman, 1971). This explains why softer tones of the clarinet have comparatively much stronger fundamentals than louder tones.

One major distinctive characteristic among mechanical reeds is their *hardness*, also referred to as heaviness or strength. Many players agree that harder reeds need higher blowing pressure and greater lip tension. Also, harder reeds are supposed to produce darker tones with less prominent high frequencies. In a study with oboe, bassoon, clarinet, and alto saxophone (Fuks, 1998a), two contrasting reeds were used in each, one rated as hard (heavy) and another as soft (light). In all reeds and instruments, louder tones required more airflow and higher blowing pressure. Among hard and soft reeds, the harder reeds always produced louder tones, needed more airflow, and involved higher blowing pressure.

### *Harmonicas and Similar Instruments*

Free reeds correspond to type IV oscillations in Figure 20.1. The air tends to close the free reed until it reaches the vertical line and then starts to open it. The harmonica is the most typical instrument in this category. As mentioned earlier, its usual blown frequency is about the same as if it were plucked and left to vibrate alone.

As mentioned previously, the vocal tract acts as an additional air column. The effect of the vocal tract on the sound of a diatonic harmonica can be demonstrated just by moving the tongue back and forth inside the mouth while blowing, which produces clear changes in the relative intensity of the overtones.

### *Air Reeds: Flutes and Recorders*

Air reeds contain a jet that is directed toward an air column, which generates airflow fluctuations inside the tube. In flutelike instruments, the airflow is guided through the player's lips toward a mouth hole. In recorderlike instruments, the flow is conducted through a *flue* channel that is directed toward a sharp edge called the labium.

Two different kinds of tone can be produced in such instruments: a *pipe tone* and an *edge tone*. A pipe tone results when the jet and the air column are acoustically coupled. All pipe tones coincide with the tube's modes, which roughly belong to a harmonic series and can be shifted by overblowing. However, when the jet is relatively distant from the resonant tube acoustical coupling to the pipe is poor and the jet interacts mainly with the labium. This produces edge tones with high frequencies that are proportional to the jet velocity. Therefore, they are not discrete as pipe tones but continuously varying. Edge tones may be important during the attack of a new tone. For example, if a flute player is starting a low $C_4$ (247 Hz), it takes a moment before the tone is actually initiated. During this period, the edge tones sound alone, contributing to the onset of the tone and to its sound quality. Edge tones are also generated when the pipe is damped (filled with cotton) and when the pipe is simply removed (Verge, 1995).

Air reeds are *flow-controlled* systems in which the pressure at both ends of the air column is close to the ambient pressure. For that reason, the air particles reach higher velocity close to these ends, so that they are directly affected by the airflow. Consequently, the player's vocal tract is usually weakly coupled to the flute and therefore has only a limited influence on the sound (Coltman, 1973; Fletcher & Rossing, 1998).

Boehm (1871/1964), the inventor of the modern flute, explains that the air stream strikes the sharp edge of the mouth hole and is broken (or divided) so that one part goes over or beyond the hole while the other part produces the tone. This is not a perfect description of what really happens, but it is still to be found in many textbooks. A flute can even produce a pipe tone when the edge of the mouth hole is dull. This can be demonstrated by applying a thin layer of wax or chewing gum to the edge of the mouth hole. If the wax or gum is placed in the right position, sound production might still be possible. In this case, however, the edge tones may be hampered, thus affecting the attack and overall sound quality.

## Techniques

### *Using the Respiratory Muscles*

The diaphragm is a large and thin muscle group located at the base of the lungs. It separates the thorax from the abdomen. Contrary to common belief, the diaphragm is solely an inspiratory agent that contracts when one breathes in by descending its central part, drawing the lungs downward, increasing the longitudinal dimension of the chest, and elevating the lower ribs (Roussos, 1995). These actions produce negative pressure around the lungs. During wind playing, blowing pressures range from almost 0 (e.g., *piano* low tones in the flute) to values that exceed 120 cm $H_2O$ in *fortissimo* high-pitched trumpet tones (Bouhuys, 1968; Pawłowski & Zołtowski, 1987; Fletcher & Tornapolsky, 1999). If a player completely fills the lungs and closes the airways, without any expiratory effort, a pressure between 30 and 50 cm $H_2O$ will be achieved. This is called relaxation pressure and is generated by the elasticity of the respiratory system.

Similarly, if the lungs are only partially filled—to 70% of their capacity, for example—the relaxation pressure will not exceed 20 cm $H_2O$ (Agostoni & Mead, 1964).

Whenever pressures higher then the relaxation pressure are needed, the expiratory muscles (abdominal and internal intercostal muscles) will be used. However, if pressures lower than the relaxation values are necessary, the inspiratory muscles (external intercostal muscles and the diaphragm) will be used (Bouhuys, 1968). An example of the latter is playing the recorder at high lung volume, when the relaxation pressure is relatively higher than the required blowing pressure. This may also happen at some ranges of the flute and other instruments. Diaphragm activity in skilled flutists has been assessed in at least three studies (Gärtner, 1973; Cossette, 1993; Cossette, Sliwinski, & Macklem, 2000), and all of them indicate that some players make systematic use of the diaphragm, while others do not use it most of the time. These findings suggest that excessive emphasis has been given to the diaphragm in the description of wind performance.

### *Blowing Pressures*

The pressure measured inside the player's mouth during performance is a major input parameter that has been documented in a number of studies. In reed woodwinds, blowing pressure is intimately related to airflow. Acoustical and aerodynamical considerations show that in reed woodwinds the oscillations initiate at a minimal pressure of approximately one-third of the pressure that completely closes the reed. In clarinets, for a fixed embouchure, the airflow and the sound power radiated by the instrument tend to decrease with increasing blowing pressure (Worman, 1971). If the player keeps a constant embouchure while increasing blowing pressure, the tone gets softer until eventually the reed simply closes. This apparent discrepancy explains why players must reduce the embouchure tightness when playing a *crescendo*.

Fuks and Sundberg (1999a) found that blowing pressures in the oboe, bassoon, clarinet, and alto saxophone are systematic in professional players. These authors measured the variation of blowing pressure with loudness and fundamental frequency at four dynamic levels: in isolated tones and musical arpeggios performed by two players of each instrument. Figure 20.2 displays part of the results, where each instrument presents characteristic curves, differing only slightly between players.

Vivona (1968) observed that nonprofessional trombone players use blowing pressures that increase with pitch and dynamic level. For tones $B\flat_2$, $B\flat_3$, and $B\flat_4$, the average pressures were 11, 16, and 39 cm $H_2O$ in "soft volume level" and 24, 39, and 67 cm $H_2O$ for "loud levels." This behavior is similar to reed woodwinds and should also apply to other brass instruments at different ranges.

A study of experienced trumpet players (Fletcher & Tornapolsky, 1999) showed that the minimal blowing pressure (i.e., threshold pressure) increased proportionally with the frequency of the tone played. In the first range, starting from threshold pressure, each doubling of blowing pressure caused the sound pres-

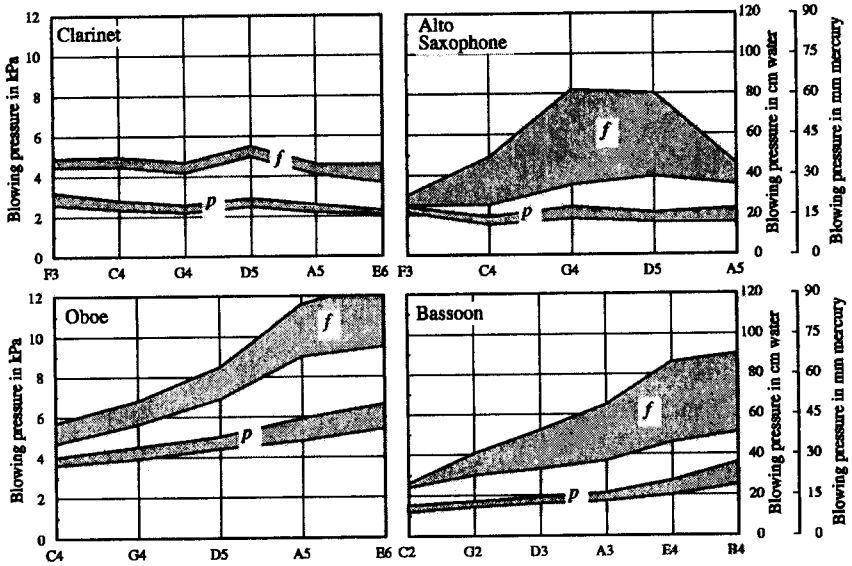



Figure 20.2. Blowing pressures in clarinet, alto saxophone, oboe, and bassoon tones, played at different dynamic levels by two professionals on each instrument. Pressure values are marked in kilopascal on the vertical axis on the left and in cm $H_2O$ and mm of mercury on the right. In all instruments, blowing pressures increased from *piano* to *forte*. In the oboe and bassoon, the pressure systematically increased with fundamental frequency. In the clarinet, there was a slight decrease of pressure with fundamental frequency. In the saxophone, the pressure tended to increase with pitch in the first octave and followed the clarinet in the second octave. (From Fletcher & Rossing, 1998, after data from Fuks & Sundberg, 1999a. Copyright Springer-Verlag 1998, used by permission.)

sure level to increase by about fifteen decibels, whereas in a second and higher range, called the saturation region, doubling of blowing pressure increased the sound level by only three decibels. In some players, extreme pressures as high as 25 kPa (250 cm $H_2O$) were measured.

In reed woodwinds, there is also an important relationship between blowing pressures and intonation. As a rule, as the blowing pressure is increased, the fundamental frequency rises. This can be explained by an increase in the reed's natural frequency. Some clarinet players believe that playing tones louder makes the pitch go flat, because a player automatically reduces the embouchure tension to increase loudness (as discussed earlier) and thus decreases the reed's natural frequency (Bak & Doemler, 1987). In other words, the effect of lip relaxation on fundamental frequency is greater than the effect of pressure.

A rather linear relationship between blowing pressures and fundamental frequency has been found for flutes and recorders (Herman, 1959; Coltman, 1968; Fletcher & Rossing, 1998). This means that an octave leap by overblowing will require approximately twice the pressure.

In flutes, blowing pressure varies between 2 cm $H_2O$ for $C_4$ and 20 cm $H_2O$ for $C_7$, with little increase between soft and loud tones (Cossette et al., 2000). A crescendo is obtained by airflow velocity increase via the reduction of the oval lip opening, in both height and width. Interestingly, professional flutists tend to use the maximal blowing pressure that a given mode can take without overblowing. They do this to ensure steadiness and quality of tone. To control the dynamic level, a player needs to change the shape and dimension of the lip opening and also move the lips forward and backward. These maneuvers affect the embouchure, the jet length, and the proportion of mouth-hole covering (Fletcher & Rossing, 1998; Cossette et al., 2000).

### *Perception of Pressure*

The results cited earlier suggest that the sensory functions in the abdominal, thoracic, and lung regions are important to effective wind instrument playing. Also, the ability of players to reach the required pressures quickly and systematically, to produce the target values with limited feedback from aural and other bodily senses, indicates that they have intuitively memorized the relevant parameters of their instrument. This assumption was confirmed by Smith et. al., (1990), who investigated the sensation of inspiratory volume and inspiratory pressure in wind instrument players. Their findings show that wind musicians have a considerably higher sensation for both inspiratory volume and pressure than nonmusical subjects. Fuks (1998b) applied a psychophysical production method to assess the sensation of blowing pressures in eight professional reed wind players. He found a proportional relationship between the measured blowing pressure and the estimated pressure level that expresses the players' perception. This approach showed that the subjects could judge quite accurately the multiplication of internal lung pressure, which seems to be a highly relevant sensory skill in wind performance. This particular ability may be trained by means of special exercises, for instance with the help of mechanical devices or a system controlled by a computer.

### *Lung Air Composition and Intonation*

The temperature inside the instrument is related to intonation, so that an increase of 10°C (18°F) will raise the pitch by about a third of a semitone, or 31 cents (a semitone contains 100 cents).

Other factors may also affect intonation. Considerable variations in the concentration of carbon dioxide ($CO_2$) and oxygen ($O_2$) can be expected during playing. Fuks (1997) measured variations in the proportion of carbon dioxide in the air exhaled by players. This typically started at 2.5% and reached 8.5% at the end of long phrases. Meanwhile, the percentage of oxygen dropped from 21% to 12%.

The speed of sound depends on the gas properties and so proportionally affects the fundamental frequency of a wind instrument (excluding the free reeds). This can be demonstrated by having a wind player take a full breath of helium before playing: the tones may be transposed by an interval that exceeds nine semitones.

Under more realistic atmospheric conditions, the effect of the changes in gas composition, together with the change in air humidity, can produce a frequency decrease of up to 15 cents (15% of a semitone) during a long phrase. Trained players tend to compensate for this tendency for the pitch to fall by adjusting their embouchure and blowing pressure during their performance.

### *Vibrato*

Vibrato consists of pulsations of pitch, usually accompanied by simultaneous pulsations of loudness and sound quality (Seashore, 1938/1967). Vibrato in winds can be produced by isolated or combined variations in blowing pressure or airflow, restrictions in the airways caused by the movement of the larynx and tongue, and movement and tension of the lips and jaw (affecting the reed vibrations). In the case of the trombone, mainly in jazz but also in some classical players, vibrato may be obtained by the variation of the instrument's length. On the trumpet, a common type of vibrato consists of a gentle shaking of the instrument, which affects the embouchure. This technique is called hand vibrato.

In general, vibrato seems to include blowing pressure oscillation, be it in flutes (Fletcher, 1975; Gärtner, 1973), reed woodwinds (Brown, 1973; Weait & Shea, 1977; Fuks, 1998a; Secan, 1997), or brass. Some studies have observed oscillations of the vocal folds during vibrato, suggesting that they play a role as a control mechanism. For intense vibrato tones produced on oboe, bassoon, and saxophone, wide pressure oscillations have been observed, on average 10 cm $H_2O$ for the oboe and bassoon and reaching values of 20 cm $H_2O$ in some cases (Fuks, 1998a). These wide undulations cannot be produced merely by the laryngeal mechanism, which otherwise would produce vocal sounds simultaneously, but require a great deal of expiratory force.

Vibrato is seldom taught, in spite of its ability to enhance the player's sonority and expressivity. Many educators consider that vibrato should be discovered by the student. Some others prescribe a series of respiratory maneuvers, most of which involve pulsation in the respiratory flow, thought to facilitate the acquisition of vibrato. A possible reason for this pedagogical attitude is that vibrato involves such a complex neuromuscular coordination that a simplified or step-by-step method for its production might produce an unsatisfactory, artificial effect.

### *Respiratory Movements*

Much attention has been given in the teaching of wind playing to respiratory techniques. Instructions on how to breathe in and out during sound production are usually given by the teacher, often while he or she is examining the student's body posture.

A few physiological methods are available to assess respiratory movement during wind performance. The technique called body plethysmography involves confining the player in a closed box, with only the head outside. Any variation in the body volume is detected and registered by the equipment. In spite of the

rather clumsy setup, this technique has been successfully applied to both singing and trombone experiments (Bouhuys, 1968). Another technique employs magnets and sensors connected to the player's thorax. These accurately record the respiratory movements. Magnetometers have been used during the playing of sustained tones in the flute (Cossette, 1993; Cossette et al., 2000). The latter study with three professional flutists found that they used similar movements when inhaling but quite different patterns of expiratory movements during expiration and that these differences did not result in perceptible differences in sound quality.

*Respiratory Inductive Plethysmography* (*RIP*) is a technique that also employs electromagnetic technology. Elastic bands that contain electrical coils are wrapped around the chest and abdomen of the player. These bands permit indirect measurement of cross-sectional areas (Strömberg, 1996). The variations of chest and abdominal areas reflect the variations in lung volume and significant changes in body position. In a study with eight professional players of the bassoon, clarinet, oboe, and saxophone, all performing different musical and respiratory tasks, the musical phrases were initiated at 55 to 87% and terminated at between 14 and 52% of the players' vital capacity, depending on instrument, piece, and phrase length (Fuks & Sundberg, 1999b). Vital capacity is the maximal air volume that a person can exhale from full lungs. The players generally showed simultaneous and in many cases equally important contributions from rib cage and abdominal wall during playing. These findings often contradicted the players' own idea of how they used their respiratory apparatus. Some of the players even claimed that their rib cage was not moving at all during inhalations and playing, while the technical equipment demonstrated quite conclusively that their rib cage was responsible for more air displacement than the abdominal wall.

The studies mentioned earlier, undertaken with professional players of the flute and reed woodwinds, show that there is no single correct way to make expiratory movements in wind playing. We nevertheless recommend a relatively constant, systematic, and automatic expiratory pattern, as this allows the player to focus on other aspects of performance, such as blowing pressure and embouchure control.

## Pedagogy and Winds

Wind playing imposes high demands in terms of learning and handling respiratory techniques, arm/hand/finger movement, tongue articulation, body posture, and embouchure control. Physiological aspects are therefore extremely important. However, it is also important for players to understand something of the acoustical behavior of the instrument they play, in terms of the function of the mechanisms and practical factors that produce the sound. In addition, psychological issues are involved, such as the aesthetic dimension of music and the emotional stress induced by intensive practice and public performance. Therefore, it is not uncommon that wind players inherit and develop an informal investigative attitude toward the physical, physiological, and psychological phenomena that they experience and try to apply these principles in teaching and

reasoning on the art. In some cases, these ideas are grounded in scientific knowledge but also include a combination of empirical conception and mental images. These images are supposed to represent an ensemble of complex tasks and sensations, which can hardly be expressed objectively (Kohut, 1985).

Physical, physiological, and psychophysical variables are continuously coordinated during performance and must be trained through systematic and repeated exercises, usually with the support of method books and under the supervision of an experienced teacher. The usual approach of most teaching methods is to combine the performance of musically significant studies with scales, arpeggios, and other drills that are used to hone the playing of complete works. Modern teaching tends to emphasize musical expression, even in studies that are relatively mechanical.

Many pedagogical methods briefly consider techniques for respiration and breathing for sound production and sometimes the underlying acoustical principles. In some cases, these methods review some basic anatomy of the respiratory apparatus, sometimes complemented by descriptions of the inspiratory and expiratory muscles, recommendations on posture during playing and on optimal inhalation strategy, and instructions on how to achieve "support" and "use the diaphragm." Some educators suggest that these terms may evoke an image of stressful playing that has led to a number of breathing and embouchure problems, even with experienced performers (Fadle & Kohut, 1999; Kohut, 1985).

Arnold Jacobs (1915–1998), a brilliant tuba player and teacher, spent his entire professional life studying respiratory and psychological aspects of winds. Jacobs was probably the most influential wind instrument teacher in the United States and other countries, particularly among brass players (Stewart, 1987; Frederiksen, 1996; Kohut, 1985). In his lessons and workshops, Jacobs used an ensemble of devices, developed for respiratory clinical purposes, to provide visual feedback to his pupils that regarded pulmonary pressure, airflow, and lung volume. Interestingly, Jacobs used a terminology that was more pedagogically oriented than based on physical facts. For instance, he emphasized *flow* as a key word for playing control, as opposed to *pressure*, which he deemed as a somewhat negative term. Because of their psychological effect, different terms tend to trigger different behaviors in players. There is another reason that Jacobs used the term *flow*. In brass instruments, a wide range of combinations of mouth pressure and lip resistance produce the same output level (Weast, 1965). If the player is able to use configurations that require lower pressures, such as lower embouchure resistance and/or stiffness, then the airflow should be maximized and the playing effort reduced (Nederveen, 1969/1998). These general principles are likely to apply also to other winds.

## Final Remarks

Knowledge acquired through scientific research cannot replace, and should not challenge, the rich imagery that inspires and guides the artist. But it may help to question, consolidate, and fill the gaps left by individual observation, thus providing an objective background for education in wind playing. Without doubt,

some research results conflict with common practice and concepts among the profession, which are often based on obsolete hypotheses that have gained credibility among musicians over many generations.

However, because measurement methods and experimental procedures are still coarser than the musical subtleties involved in performance, it is important for researchers to understand the artists' reality and to take advantage of their skills and long-term experience when designing musically relevant investigations. Collaboration among areas such as physiology, acoustics, psychology, engineering, and musicology will help to enhance knowledge of music performance in winds. As a first step, there is a need to develop some consensus concerning the meaning of technical terms, in ways that promote a better understanding and effectiveness of their application to music teaching and learning.

### *References*

Agostoni, E., & Mead, J. (1964). Statics of the respiratory system. In W. O. Fenn and H. Rahn (Eds.), *Handbook of physiology*, section 3: "Respiration," Vol. 1 (pp. 387–409). Washington, DC: American Physiological Society.

Backus, J. (1985). The effect of the player's vocal tract on woodwind instrument tone, *Journal of the Acoustical Society of America, 63*, 17–20.

Bak, N., & Doemler, P. (1987). The relation between blowing pressure and blowing frequency in clarinet playing. *Acustica, 63*, 238–241.

Barbenal, J., Davies, J. B., & Kenny, P. (1986). Science proves musical myths wrong. *New Scientist, 1502*, 29–32.

Benade, A. H. (1986). Interactions between the player's windway and the air column of a musical instrument. *Cleveland Clinic Quarterly, 53*, 27–32.

Benade, A. H. (1990). *Fundamentals of musical acoustics* (2nd ed.). New York: Dover.

Benade, A. H., & Gans, D. J. (1968). Sound production in wind instruments. *Sound Production in Man: Annals of the New York Academy of Sciences, 155*(1), 247–263.

Boehm, T. (1871/1964). *The flute and flute playing.* New York: Dover.

Bouasse, H. P. M. (1929/1986). *Instruments à vent.* Paris: Blanchard.

Bouhuys, A. (1964). Lung volumes and breathing patterns in wind instrument players. *Journal of Applied Physiology, 19*(5), 967–975.

Bouhuys, A. (1968). Pressure-flow events during wind instrument playing. *Sound Production in Man: Annals of the New York Academy of Sciences, 155*(1), 264–275.

Brown, A. F. D. (1973). *A comprehensive performance project in oboe literature with a cinefluorographic pilot study of the throat while vibrato tones are played on flute and oboe.* Unpublished doctoral dissertation, University of Iowa.

Campbell, M., & Greated, C. (1987). *The musician's guide to acoustics.* New York: Schirmer.

Chen, F. C., & Wienreich, G. (1996). Nature of the lip reed. *Journal of the Acoustical Society of America, 99*(2), 1227–1233.

Coltman, J. W. (1968). Sounding mechanism of the flute and the organ pipe. *Journal of the Acoustical Society of America, 44*, 983–992.

Coltman, J. W. (1973). Mouth resonance effects in the flute. *Journal of the Acoustical Society of America, 54*, 417–420.

Copley, D. C., & Strong, W.J. (1996). A stroboscopic study of lip vibrations in a trombone, *Journal of the Acoustical Society of America, 99*(2), 1219–1226.

Cossette, I. (1993). *Étude de la méchanique respiratoire chez les flûtistes.* Unpublished doctoral dissertation, Université de Montréal, Canada.

Cossette I., Sliwinski, P., & Macklem, P. (2000). Respiratory parameters during professional flute playing. *Respiration Physiology, 121*, 33–44.

Fadle, H., & Kohut D. (1999). *Musizieren.* Essen: Die Blaue Eule.

Fletcher, N. H. (1975). Acoustical correlates of flute performance technique, *Journal of the Acoustical Society of America, 57*, 233–237.

Fletcher, N. H., & Rossing, T. D. (1998). *The physics of musical instruments* (2nd ed.). New York: Springer.

Fletcher, N. H., & Tornapolsky, A. (1999). Blowing pressure, power, and spectrum in trumpet playing. *Journal of the Acoustical Society of America, 105*(2), 874–881.

Frederiksen, B. (1996). *Arnold Jacobs: Song and wind.* Gurnee, IL: Windsong.

Fuks, L. (1997). Prediction and measurements of exhaled air effects in the pitch of wind instruments, *Proceedings of the Institute of Acoustics, 19*(5/2), 373–378.

Fuks, L. (1998a). Aerodynamic input parameters and sounding properties in naturally blown reed woodwinds. *KTH Speech, Music, and Hearing Quarterly Progress Status Report, 4*, 1–11.

Fuks, L. (1998b). Assessment of blowing pressure perception in reed wind instrument players. *KTH Speech, Music, and Hearing Quarterly Progress Status Report, 3*, 35–48.

Fuks, L., & Sundberg, J. (1999a). Blowing pressures in bassoon, clarinet, oboe and saxophone. *Acustica / Acta Acustica, 85*(2), 267–277.

Fuks, L., & Sundberg, J. (1999b, March). Using respiratory inductive plethysmography for monitoring professional reed instrument performance. *Medical Problems of Performing Artists, 14*(1), 30–42.

Gärtner, J. (1973). *Das Vibrato unter besonderer Berücksichtigung der Verhältnisse bei Flötisten.* Regensburg: Bosse.

Helmholtz, H. von (1862/1954). *On the sensations of tone as a physiological basis for the theory of music.* New York: Dover.

Herman, R. (1959). Observations on the acoustical characteristics of the English flute. *American Journal of Physics, 27*, 22–29.

Johnston, R. B. (1987). Pitch control in harmonica playing. *Acoustics of Australia, 15*(3), 69–75.

Kohut, D. (1985). *Musical performance: Learning theory and pedagogy.* Englewood Cliffs, NJ: Prentice-Hall.

Nederveen, C. J. (1969/1998). *Acoustical aspects of woodwind instruments.* DeKalb: Northern Illinois University Press.

Pawłowski S., & Zołtowski, M. (1987). A physiological evaluation of the efficiency of playing the wind instruments—an aerodynamic study. *Archives of Acoustics of Poland, 12*(3–4), 291–299.

Porter, M. (1967). *The embouchure.* London: Boosey & Hawkes.

Roussos, C. (Ed.) (1995). *The thorax* (2nd ed.). New York: Marcel Dekker.

Sachs, C. (1940). *The history of musical instruments.* New York: Norton.

Seashore, C. E. (1938/1967). *The psychology of music.* New York: Dover.

Secan, S. (1997). *Frequency and amplitude modulation in oboe vibrato.* Master's thesis, Ohio State University.

Smith, J., Kreisman, H., Colacone, A., Fox, J., & Wolkove, N. (1990). Sensation of inspired volumes and pressures in professional wind instrument players. *Journal of Applied Physiology, 68*(6), 2380–2383.

Stewart, D. (Ed.) (1987). *Arnold Jacobs: The legacy of a master.* Northfield, IL: Instrumentalist.

Strömberg, N .O. T. (1996). *Respiratory Inductive Plethysmography (RIP): Calibration, breathing pattern analysis and external $CO_2$ dead space measurement.* Doctoral thesis, Linköping University.

Strong, W. J., & Plitnik, G. R. (1992). *Music speech audio.* Provo: Soundprint.

Sundberg, J. (1991). *The science of musical sounds,* San Diego, CA: Academic Press.

Verge, M. P. (1995). *Aeroacoustics of confined jets, with applications to the physical modelling of recorder-like instruments.* Doctoral thesis, Technical University of Eindhoven, the Netherlands.

Vivona, P. M. (1968). Mouth pressures in trombone players. *Sound Production in Man: Annals of the New York Academy of Sciences, 155,* 290–302.

Weait, C., & Shea, J. B. (1977, January/February). Vibrato: An audio-video-fluorographic investigation of a bassoonist. *Applied Radiology.* http:// www.applied radiology.com. January/February.

Weast, R. D. (1965). *Brass Performance.* New York: McGinnis & Marx Publishers.

Worman, W. (1971). *Self-sustained nonlinear oscillations of medium amplitude in clarinet-like systems.* Doctoral thesis, Department of Physics, Case Western Reserve University.

Yoshikawa, S. (1995). Acoustical behavior of brass players lips. *Journal of the Acoustical Society of America, 97*(3), 1929–1939.

# 21

# Rehearsing and Conducting

HARRY E. PRICE & JAMES L. BYO

Conducting and rehearsal behaviors play a role in establishing an appropriate and effective rehearsal atmosphere. Situations in which conductors provide predominantly positive feedback result in better attitudes, attention, and performance. Fast paced rehearsals are usually the most effective, and comprise frequent and generally brief episodes of teacher talk and ensemble performance. Enthusiastic or dynamic rehearsing features stark contrasts of behavior at optimal times—loud and soft talk, expressive and neutral conducting, group and individual eye contact. Rehearsals should be structured to include processes of diagnosis, prescription, presentation, monitoring, and feedback, with brisk paced and clear directions. Essentially, a conductor should focus on making verbalizations efficient and keeping them to a minimum, while enhancing nonverbal behaviors to include large amounts of eye contact and clear and unambiguous conducting gestures.

"From the moment a conductor steps on the podium a special world is in the process of being constructed" (Faulkner, 1973, p. 149). By definition, conductors are leaders. Leadership requires competence, credibility, and charisma, and these qualities can influence musicians' attitudes and performances (Parasuraman & Nachman, 1987).

The current role of the conductor dates back to the late eighteenth century. Not until the mid-twentieth century, though, did a traditional focus on gesture or baton technique evolve to include issues of rehearsal technique. The application of scientific inquiry to rehearsing and conducting is a recent phenomenon, and the psychology of large ensemble leadership, even in a popular sense, remains insubstantially addressed.

One might consider conducting and rehearsing as distinctly separate acts, conducting being a nonverbal physical act, and rehearsing being conductor-led

ensemble preparation (largely verbalizations) for musical performance. It is our view that conducting and rehearsing are inextricably linked. When done well, they are complementary, even indistinguishable. In the school music environment, where the focus is on the nature of the rehearsal process with the concert serving to set a context (Reimer, 2000), rehearsing and teaching are analogous. One might argue that everything involved in rehearsing and conducting can be characterized via a teaching paradigm, even in a professional ensemble environment.

## The Rehearsal

### *Rehearsal Atmosphere*

Even actions that occur before the first words are uttered or downbeat given influence ensemble members' perceptions of a conductor's abilities. A hesitant approach to the podium, fumbling of materials, low levels of eye contact, and poor posture and hand position have a negative effect on perception of conductor competence, while the converse of these behaviors has a positive effect (Fredrickson, Johnson, & Robinson, 1998).

Conductor behaviors play a role in establishing authority. The likelihood that authority figures are persuasive is in part dependent on the extent to which they are viewed as decisive. Conductors who are inspiring and persuasive can be said to be exercising a form of domination over the performers (Faulkner, 1973). This domination, however, does not preclude the necessity of collaboration between conductor and ensemble. The rehearsal atmosphere must be such that the combination of conductor persuasiveness and collaboration results in an ensemble that is responsive and receptive to the conductor's verbal and nonverbal behaviors. Indeed, performances appear to benefit when ensemble members feel a part of the learning process rather than functioning as passive recipients of information (Hamann et al., 1990).

### *Feedback*

Another aspect of conductor behavior that is relevant to rehearsal atmosphere is conductor feedback. Feedback is necessary in order to achieve the ultimate performance envisioned. Situations in which conductors provide predominantly positive feedback result in better attitudes, attention, and performance than ones in which conductors provide instruction without feedback (see Price, 1983). Indeed, positive rehearsal environments tend to result in better ensemble attentiveness and attitudes than do negative ones (e.g., Madsen & Yarbrough, 1985; Price, 1992). In contrast, there are accomplished conductors and teachers in instrumental music who use more negative than positive feedback, which results in successful performances and positive student attitudes (Cavitt, 1998; Duke & Henninger, 1998). It is important to consider that negative feedback delivered constructively is likely to function quite differently from that delivered angrily. Negative feedback is necessary and can function well if the combi-

nation of instruction and feedback, both negative *and* positive, leads to accomplishment of musical goals. If, however, the rehearsal is characterized by inordinate amounts of ensemble failures, the potential for negative feedback to be detrimental to attitude and attentiveness is increased.

### *Pacing*

In general education, fast pacing of instruction has been linked to effective teaching, though it is recognized that in seeking an optimal pace one must consider the nature and complexity of the subject matter being taught as well as the sophistication of those being taught (e.g., Brophy & Good, 1986). Slow pacing is not necessarily an indication of ineffective teaching.

There is no unequivocal definition of rehearsal pacing nor of an optimal pace of activity. However, frequency and duration of ensemble performance episodes, conductor talk, lag time between cutoff and conductor talk, and conductor rate of speech are thought to be components of pacing. Experienced observers have been asked to view videotaped rehearsal examples and rate conductor pacing on a scale from very slow to very fast, and the responses correlated fast pace to fast rate of conductor speech (Single, 1990). This association is reinforced by findings in rehearsal settings, including a professional orchestra with Bruno Walter conducting that found most teacher/conductor verbal behaviors occurred in short bursts (Yarbrough, 1988).

In an attempt to describe pacing empirically, Duke, Prickett, and Jellison (1998) selected fast- and slow-paced videotaped examples of rehearsal conducting and music teaching, based on observers' personal perceptions of pacing, not on one prescribed definition. Fast and slow paces were then described according to the frequency and duration of conductor versus ensemble activities. They found that fast paced rehearsals comprised frequent and generally brief episodes of teacher talk and student performance.

In one of the few longitudinal studies of the rehearsal process, pacing was defined as change of focus of activity between director and ensemble (Yarbrough, Dunn, & Baird, 1996). An examination of rehearsals with the lowest and highest performance ratings found that the lowest performance rated rehearsal, which was the first rehearsal of a work, had a considerably slower pace than the higher rated and later rehearsal. These findings suggest that pacing should be examined as a change in activity and a function of the relative familiarity and difficulty of the literature being rehearsed.

A conductor might endeavor to quicken the rehearsal pace by speaking more succinctly, thus reducing the duration of talk episodes. While less-accomplished conductors stop and restart ensembles frequently, as is suggested by the preceding description of fast pace, they tend to do so without providing instruction (Goolsby, 1999). In addition, expert band conductors vary the rate of alternation between conductor talk and student performance episodes according to the issue being addressed (Cavitt, 1998). Thus advising a novice conductor to quicken the pace by working faster fails to acknowledge the fact that some performance errors and rehearsal conditions simply require more time to correct than others do.

### *Performance Error Detection*

Related to conductor feedback and rehearsal pacing is the conductor's response to performance errors. How best to prepare prospective conductors to cope effectively with a "great confusion of sound" (Bruno Walter in Chesterman, 1976, p. 22) in rehearsal has been the object of study in performance error detection. There is no definitive hierarchy, by difficulty, across the different types of errors (pitch, rhythm, articulation, etc.) that may occur in rehearsals (Byo, 1997). A performance error is more or less audible depending on its type, its timbre, the texture and tempo of the music, whether the listener is conducting, and whether the listener opts for a focused or unfocused approach to the listening task (Byo & Sheldon, 2000). It may be advantageous for conductors to be aware of musicians' performance tendencies, such as rushing rather than dragging tempi and playing sharp rather than flat.

Conductors may improve their abilities to detect performance errors by practicing error detection (Byo, 1997). It is tempting to assume that tasks such as sight singing and harmonic dictation enhance skill in error detection. Research, however, shows that this may not be the case (Brand & Burnsed, 1981). It is likely that skill in error detection is most affected by practicing it directly. There seems to be pedagogical potential in structuring personal error detection practice experiences to begin with one line of music only. If one line of music is approached as the musically rich, comprehensive entity it can be (involving issues of tempo, pitch, rhythm, articulation, intonation, dynamics, and interpretation), it can present a formidable challenge to the intellect and ear of the conductor. The ability to detect errors effectively, in any context, begins with the ability to hear and evaluate one line of music completely and precisely. Further, there are indications that error detection skill increases when conductors go to rehearsal with a well-developed internal sound image of the score (Byo & Sheldon, 2000). This sound-based knowledge of the score can reduce or override aural distractions in rehearsal. Conductors should study scores in such a way that they are able to ask to what extent they hear from their ensemble what they expect to hear. Then, when the ensemble's performance fails to match their internalized images of the ideal sound, they will be able to begin the process of error detection.

### *Conductor Demeanor*

In examining the effects of conductor enthusiasm on ensemble performance, attentiveness, and attitude, one might expect an enthusiastic conductor to be more interesting, if not more effective, than a dull one. However, the musicians in Yarbrough's (1975) study, in a statistical sense, did not perform better for, or pay more attention to, the enthusiastic conductor. This result caused the investigator to look more deeply at precisely what enthusiastic conductors do. A novice conductor who is not naturally dynamic might have the same concern. Yarbrough hypothesized that the enthusiastic rehearsal conductor is one who makes high levels of eye contact with the ensemble and varies it from group to individual, frequently approaches the ensemble by moving or leaning forward, manipulates

speech across a wide range of volumes, and gestures expressively with hands/arms and face (i.e., using a variety of movements and contrasts of facial expression).

Notice a recurring theme—that is, the behaviors of an enthusiastic conductor are not static; rather, they vary. Results of a detailed analysis that used an observation form designed to isolate these behaviors made apparent what was not readily seen through more casual observation—that *variation* in conductor behavior is perhaps more salient than doing *more* of a behavior (e.g., talking louder without contrast or conducting expressively without contrast). It appears that enthusiastic conducting (dynamic conducting, if you prefer) might best be described in terms of the conductor's ability to exhibit stark contrasts of behavior at optimal times—loud and soft talk, expressive and neutral conducting, group and individual eye contact (Byo, 1990). Much as music without appropriate variety and contrast is typically uninteresting, so is a rehearsal. By analyzing one's own conducting via videotape in combination with a conductor observation form, the aspiring conductor can apply a tested research procedure in order to see more clearly the behaviors related to enthusiastic rehearsing and conducting (Madsen & Yarbrough, 1985).

Of course, effective rehearsing and conducting extend beyond the mere delivery of information to include the accuracy of information conveyed by the conductor, both verbally and nonverbally, and the quality of the interaction between conductor and ensemble (Standley & Madsen, 1987). Conductors' abilities to create and sustain an intense rehearsal atmosphere hinge on their abilities to handle the subject matters of music, rehearsal, teaching, and group psychology accurately, appropriately, and in an engaging manner. This intensity can be reduced or lost by a conductor's mistake, slow or dull delivery, or casual demeanor (Madsen, 1990). For a conductor who lacks rehearsal intensity (or whose passion is not apparent to ensemble members), observation from analysis combined with videotape review can offer specific diagnosis of the problem and suggest corrective action (e.g., Cassidy, 1990; Kaiser, 1998).

### *Rehearsal Structure*

Although originally it referred to the interaction of child with environment, Vygotsky's *Zone of Proximal Development* (ZPD) construct is pertinent to structuring rehearsals. ZPD, in the context of the rehearsal, can be thought of as the difference between what an ensemble can achieve without and with the direction of a conductor. It is the distance between what problems individuals can fix independently and the possible solutions that can be achieved in collaboration with peers or under the guidance of an authority (Vygotsky, 1978). As ensemble members become more sophisticated, they are increasingly independent of the conductor in their abilities to make appropriate decisions about the music and its performance. A group of students in a secondary school would require more basic direction about phrasing, articulation, and intonation tendencies than would a professional ensemble. It may be said that the conductor's task is to move the ZPD forward toward independence. Here the metaphor of scaffolding may be useful. It refers to establishing a situation in which musicians can achieve at

a higher level when provided external support. For example, a conductor might structure rehearsals and organize appropriate tasks so that an ensemble is better able to interpret and perform the music than if it were unaided. Studio teachers employ scaffolding strategies to help student progress across the zone by assisting students in moving from what they can accomplish dependently to reaching their potential to function independently (Kennell, 1989). With increased sophistication (independence), ensembles approach what may be comparable to self-regulation in individuals; thus less scaffolding is necessary to achieve a performance goal.

Rehearsing music is typically a process of successive approximations or small steps toward performance objectives. A rehearsal consists of any number of units or rehearsal frames in which the conductor identifies a problem in need of rehearsing, extracts it from the music, divides the problem into parts, rehearses the parts (by providing information, giving directions, asking questions, modeling, modifying a part for rehearsal purposes, providing feedback), and finishes with the problem being performed in context before the conductor moves on to other material (Duke, 1999/2000). "Accurate finger technique in the woodwind ostinato leading to letter D" might be considered a performance objective. A rehearsal frame would comprise all conductor and ensemble interactions as they relate to the development of accurate finger technique. A conductor's ability to clearly articulate and order instructional steps such that there is fairly constant improvement in the performance of this passage would likely be a major factor in an ensemble's perception of success—and perception of success by the ensemble is no trivial matter.

By viewing a videotaped rehearsal in these small units rather than as one complex whole, developing conductors are likely to see more clearly and with fewer distractions the elements involved in their decision-making processes. Knowledge of one's own sequencing of rehearsal instruction is possible through the use of formal observation tools designed to illuminate forward, failed, and repeated approximations of objectives (Duke & Madsen, 1991). Without structured videotape review of the rehearsal, it is difficult to obtain a precise evaluation of the conductor's ability to organize primarily in a forward or positive direction.

Basic to effective teaching in general education is a three-step sequence of instruction that involves (1) teacher task presentation, (2) student response, and (3) teacher feedback (Becker, Englemann, & Thomas, 1971). In music education, a parallel instructional model (conductor presentation of a task, ensemble interaction with the task, and conductor feedback) has been applied to and studied across many levels of ensemble sophistication (e.g., Arnold, 1995; Price, 1983; Tipton, 1996). Though many interaction patterns can be found in any rehearsal, the preferred complete sequential pattern is one that begins with the conductor providing a musical task for the performers, followed by the ensemble attempting to perform the task, and ending with descriptive, positive feedback from the conductor. In fact, experienced band conductors tend to use complete patterns more so than do less-experienced conductors (Goolsby, 1997). This three-step model holds promise for providing conductors with a scaffold on which to develop re-

hearsal structure. Again, videotape combined with relevant observation techniques helps one to see rehearsals in the context of these smaller units of instruction.

Rehearsing should be a process of diagnosis, prescription, presentation, monitoring, and feedback, with brisk-paced and clear directions. It can be characterized by frequent use of nonverbal communication, prioritization of rehearsal materials, task-related and contingent feedback, clear statement of rehearsal objectives, encouragement of ensemble-generated ideas, and, depending on the level of the ensemble, the teaching of musical in addition to performance skills (McCoy, 1985). Model secondary school choral conductors exhibit thorough preparation and maintain an appropriate atmosphere. Nonverbal communication tends to be positive. There is little sarcasm and a businesslike image is projected. Ensemble members talk little and disciplinary action is generally unnecessary (Fiocca, 1989).

## Verbal Communication

### *Content*

Verbal communication accounts for approximately 40 to 60% of ensemble rehearsal time (e.g., Single, 1990; Watkins, 1996; Yarbrough, 1988; Yarbrough & Price, 1989). While this is a large proportion of the rehearsal, it may be understandable, given some research evidence that indicates that ensemble members respond more accurately to verbal instruction than to conducting gestures (Skadsem, 1996, 1997).

In instrumental ensembles, conductors tend to focus on rhythm and tempo, although the sophistication of the conductor appears to have an impact on rehearsal emphasis. Experts tend to work more with overall ensemble sound, including more demonstrations, instruction on intonation, and guided listening, than do inexperienced conductors. Experienced conductors also address balance, style, and tone more than do novices (Goolsby, 1997, 1999).

Regardless of the content, less-experienced conductors tend to spend more time talking in rehearsals. They stop and restart ensembles more frequently without providing instruction and are less efficient, thus taking longer to prepare a piece for performance (Goolsby, 1999). Therefore, when conductors stop ensembles, they should make concise and substantive suggestions.

Experienced teachers not only talk less, but they provide more breaks between musical selections, spend more than half the rehearsal time performing, use more nonverbal modeling, and use less time getting started. They appear to be more proactive in that they use similar amounts of rehearsal time across pieces rehearsed, while less experienced conductors tend to spend more time on the first piece rehearsed, indicating that they are in a reactive mode. Evidently, inexperienced conductors get distracted from their rehearsal plans when responding to ensemble performance (Goolsby, 1996).

There is an evolution of focus over time as performance approaches, from fundamental issues of accuracy and precision to more general concerns of nu-

ance and interpretation. As the quality of performance increases or as a concert approaches, conductors tend to talk less and focus more on ensemble performing in rehearsals (Davis, 1998).

### *Aural Modeling*

Aural modeling alone and in combination with language is among the most frequent verbal modes in rehearsals. It generally consists of a conductor providing a demonstration—vocal, acoustical or MIDI instrument, recording—followed by ensemble kinesthetic, vocal, or instrumental response. Models can be effective in demonstrating both appropriate and inappropriate performances and can be used to minimize verbalizations. Conductors pervasively use sung or quasi-sung sounds or other means of demonstration (e.g., clapping, tapping), both imitating salient features of what was previously performed and presenting examples of what they would like to hear (Weeks, 1996).

Ensemble directors' abilities to model have a strong relationship to the quality of instrumental music student performance, and strong modelers tend to spend more rehearsal time modeling than do less skillful modelers (Sang, 1987). Modeling is critical to a rehearsal, and experienced ensemble directors use more nonverbal instruction (e.g., modeling and demonstration) and talk less than do novices (Goolsby, 1996).

### *Conceptual Teaching*

While considerable proportions of rehearsals are spent in verbal activity and modeling, principally on the part of the conductor, little of it elicits higher order or conceptual thinking on the part of the performers. General music and ensemble classes generally involve students in lower cognitive processes, emphasizing mechanics of performance almost to the exclusion of the application and accumulation of musical knowledge and the abilities to think about music (Goodlad, 1983). Conductors' efforts appear to be weighted toward providing guidance on how to make corrections or presenting exact solutions, by saying things such as "you need more air" or "the percussion need to play softer." This limits opportunities for self-correction on the part of the ensemble through slight hints or scaffolding (Weeks, 1996). In these situations, ensemble members function much like simple machinery, rendering only specific responses to specific instructions about a specific point in a specific piece of music.

Conversely, conceptual rehearsing reinforces or introduces concepts in ways that encourage the transfer of concepts from one passage to another passage or work. Statements can be as simple as "whenever you see a terraced dynamic, take care not to anticipate it with a slight crescendo or decrescendo" or as sophisticated as "this section is the recapitulation; what does that mean?" In secondary school choral settings, approximately 1% of rehearsal time is spent attempting to evoke higher order thinking, such as analysis, synthesis, or evaluation (e.g., Watkins, 1996). This same automatonlike approach to rehearsing has also been found in secondary school instrumental rehearsals, where less than 3% of

rehearsal time is spent in attempts to improve the grasp of musical concepts (Blocher, Greenwood, & Shellahamer, 1997). The literature, the bulk of which comes from the United States, is quite clear with regard to learning: unless one teaches for the transfer of ideas, there is no transfer.

Even though it appears that there is little encouragement of higher order thinking and the development of concepts in rehearsals, it would seem that planning and employing strategies to promote these throughout the rehearsal would be most effective in the long-term growth of performers and ensembles. Without these attributes, a conductor is condemned to reteach an idea every time a similar passage or concept is encountered, as opposed to musicians making connections cognitively and transferring knowledge and skills to new situations. More sophisticated performers have likely attained higher order music skills through inductive reasoning as a result of synthesis of many experiences; thus conductors of highly skilled ensembles are better able to attend to the performance nuances that help make music rapturous.

## Nonverbal Issues

### *Facial Expression*

The research in interpersonal communication indicates that the face is the primary nonverbal means of conveying six emotions—happiness, sadness, anger, fear, surprise, and disgust (Bull, 1983; Ekman & Friesen, 1975; Izard, 1997). Though it is tempting to assume that a conductor must somehow look like the music facially in order to achieve maximum effect, there is not a research base in music to support this notion. Research that isolates the conductor's face has only focused on approving, disapproving, and neutral expressions (e.g., Byo & Austin, 1994; Madsen & Yarbrough, 1985; Price & Winter, 1991). The face has been viewed as a means for conductors to respond approvingly (e.g., smile or nod) to appropriate ensemble performance or disapprovingly (e.g., furrowed brow) to inappropriate performance.

Seldom has the face been examined as a means to convey musical information (Berz, 1983; Mayne, 1993). In fact, one conductor-training technique involves mask work (covering the face) to prompt the conductor to focus on the expressive potential of the trunk and shoulders (Tait, 1985). This focus away from the face and toward the communicative power of the torso and breath in conductor training is a concept borrowed from theater and mime (Oertle, 1999). It appears that novice conductors should first develop a varied repertoire of nonverbal skills in gesture and body (see Conductor demeanor earlier). The area of greatest concern facially is eye contact.

### *Eye contact*

The communicative advantages of eye contact are well documented in several research literatures (Burgoon et al. 1984; Fredrickson, 1992). Greater conductor

eye contact with an ensemble is associated with increased attentiveness by ensemble members (Yarbrough & Price, 1981). There is evidence, however, that directors of school ensembles look at the score more than the ensemble (Fredrickson, 1992; Sherrill, 1986).

Given the potential for conductor eye contact to interact with other nonverbal behaviors to result in an intensified message (e.g., Burgoon et al., 1984; Price & Winter, 1991), it is clear that conductors should look at the ensemble more often than not. One of the variables that separates accomplished and novice conductors is the tendency of accomplished conductors to combine expressive nonverbal behaviors (eyes, arm/hands, body movement) for optimal effect (Byo & Austin, 1994). By analyzing their own conducting via videotape and in combination with research-based observation techniques, conductors can formulate a clear picture of their eye contact tendencies and, if change is desired, create a plan of action (Byo & Austin, 1994; Madsen & Yarbrough, 1985).

While there is little research that examines ensemble member eye contact with the conductor, it is likely that the extent to which ensemble members are aware of their rehearsal and performance surroundings bears heavily on whether they are in a position to receive the full impact of nonverbal messages. Byo and Lethco (2001) examined musical events and conductor-related conditions during which student musicians tend to look at the conductor. High school band musicians looked up with much greater frequency during slow music than fast music where the demands of the music superseded effects of eye contact and expressive gesture. School-based conductors might promote watching the conductor by teaching ensemble members to look up for specific reasons at predetermined places in the music, for example, at entrances, releases, rehearsal numbers or letters, changes in meter, tempo, and dynamics, and the beginning and end of the piece or movement.

Conductors should be aware of two research findings that bear on issues of eye contact by ensemble members: (1) eye contact by ensemble members increases with musical sophistication, a result that is consistent with research findings in the recognition of specific conducting gestures (Mayne, 1993; Sousa, 1988), and (2) singers performed a dynamic change most successfully following concise verbal instructions from the conductor, despite having predicted that they would respond best by watching the conductor (Skadsem, 1997).

### *Gesture*

Conductors must possess a large repertoire of nonverbal behaviors from which to choose in order to impart expressivity within the context of the structure and style of any composition. They must interpret and shape the elements that contribute to musical expression. Expressiveness/musicianship or musical effect is one of the primary factors in assessing performances (Bergee, 1995; Burnsed & King, 1987). The often-quoted statement by Seashore "beauty in music largely lies in the artistic deviation from the exact or rigid" (1938/1967, p. 249) would relate to both a conductor's interpretation and ability to demonstrate these "deviations." Otherwise, a metronome at varying tempi would serve as well as a person to lead an ensemble.

Employment of movement techniques by Rudolf von Laban, among the most influential figures in dance and movement education in the twentieth century, can result in superior performance than that achieved by typical verbal ensemble instruction (Holt, 1992). Laban characterized effort as having eight basic motions—thrusting, slashing, floating, gliding, wringing, pressing, flicking, and dabbing (Preston-Dunlop, 1980). Each of these and their permutations have analogues in conducting gestures and music performance. A conductor who has a physical command and understanding of these could draw upon them as needed; for instance, the appropriate employment of flicking, pressing, and thrusting may yield the exact representation desired of a specific *marcato* style within a passage. Given a lack of body movement coverage in conducting textbooks and traditional conducting courses, the application of principles of Laban to the pedagogy of conducting has been advocated. Laban Movement Analysis (Laban, 1975) is a framework within which conducting movement has been viewed and analyzed (Benge, 1996) and through which expressive gestures may be enhanced (Miller, 1988). A better understanding of movements and their possible interpretive meanings might well strengthen the communication between conductors and ensembles.

Whether individuals interpret other nonverbal behaviors consistently has been the focus of research on conducting emblems. Sousa (1988), adopting language from Ekman and Friesen (1969), applied the term *emblems* to conducting gestures whose meanings were interpreted reliably by musicians (e.g., changed conducting pattern size to indicate *piano*, *subito forte*, or *crescendo*; lowering of the left hand for *decrescendo*; or reboundless pattern for *staccato*). Among secondary school and college musicians, the ability to identify emblems on a paper-and-pencil test increases with years of instrumental music experience. With inexperienced musicians, even brief instruction in the recognition of emblems that reflect expressive musical characteristics such as *crescendo*, *staccato*, *legato*, and *tenuto* results in increased ability to derive meaning from gestures and perform accurately in response to them (Cofer, 1998; Kelly, 1997).

In studies that have examined the effects of conducting gestures on performance quality of high school musicians, investigators have found expressive gestures to elicit better performance quality than unexpressive ones (Grechesky, 1985; Laib, 1993; Sidoti, 1990). In research that involved young band students, however, expressive gestures did not result in better performance quality (Price & Winter, 1991). These results taken together are consistent with the positive relationship between performance experience and emblem recognition. It may be expeditious to specifically teach emblems of conducting to less-experienced musicians so they can better interpret them.

While conducting, a form of nonverbal communication, is a complex task, it can be learned outside the traditional apprenticeship and course modes. Basic conducting skills have been enhanced in research that focused on systematic self-observation of behavior—precise definitions of conducting skills, opportunities to practice, reinforcement through videotape feedback, and self-analysis (e.g., Madsen & Yarbrough, 1985; Price, 1983).

## Conclusion

Everything conductors and ensemble members do in rehearsals should be of consequence, from the moment when everyone is assembled before the first passage is rehearsed to the concert performance. Conductors therefore need to establish their authority and expertise immediately and willfully decide on the rehearsal atmosphere to be created.

A conductor must engage the ensemble in such a way that an intense rehearsal atmosphere is established and then maintained. This will not happen if conductors fail to control the nature of their interactions with ensembles, lack overt enthusiasm for the task at hand, provide inaccurate information either verbally or nonverbally, or allow the pace of the rehearsal to be slow. It is the conductor's responsibility to bridge the gap between what the ensemble can do independently and what it could do with carefully crafted rehearsals under the guidance of an expert who uses well-ordered instruction. The rehearsal must be proactive, structuring scaffolding that will allow for successive approximations in which every effort will likely result in progress.

Rehearsal models must be accurate and need to address all learning styles and modalities by providing visual, aural, and kinesthetic experiences. Providing information in a multimodal (sensory) fashion is most effective, for example, conducting (visual) while providing a singing (auditory) model or having ensemble members move (kinesthetic) while singing their parts.

The conductor must have an arsenal of verbal and nonverbal skills. Instruction is most effective when it is substantive and delivered concisely, clearly, and without sarcasm. It should help move ensemble members beyond their current knowledge and conceptions to higher order musical thinking. Conductors need to maintain and elicit eye contact and have command of a large repertoire of facial expressions, conducting emblems and movements, and other means of communicating musical expression and precision.

In short, capable conductors must be remarkably prepared and have complete knowledge of the score and how to realize it. They need to know what they should be hearing, how it differs from the current rehearsal performance, and how to get to the level needed for concert performance. All aspects of the rehearsal should be planned. Components such as teaching conceptually and giving feedback are not done in isolation; the rehearsal should be imbued with these and other activities continuously.

Not only do aspects of conducting and rehearsing develop through experience and practice, but rehearsals also need to be audio- and videotaped and reviewed by conductors. There is a consistent theme in the literature that regards the benefits of systematic self-observation and feedback in enhancing podium skills. This helps us to see ourselves objectively or as others do.

It is an extremist perspective to suggest that either conducting or verbalization is unnecessary in an ensemble setting. There are those who hold the view that if a conductor could manage to present gestures that are perfect models, lacking in any ambiguity, there would be no need for talk. Conversely, there are

those who suggest that most of the productive work in a rehearsal is done when the conductor verbally and directly tells performers what is wanted. It is true that nonverbal and verbal forms of communication must be done exceptionally well; however, both are necessary.

Visual communication (nonverbal, inaudible, symbolic, or demonstrative behavior) alone is significantly less effective in improving ensemble performance than visual communication in combination with verbalization (oral use of meaningful words to convey information) and modeling (demonstration of behaviors). The combination of verbal and modeling communication is also more effective than visual communication alone. However, the combination of verbal, modeling, and visual communication is the most effective approach (Francisco, 1994). All must be done effectively, skillfully, and efficiently to have optimal rehearsals for excellent performances.

### *References*

Arnold, J. A. (1995). Effect of competency-based methods of instruction and self-observation on ensemble director's use of sequential patterns. *Journal of Research in Music Education, 43*(2), 127–138.

Becker, W. C., Englemann, S., & Thomas, D. R. (1971). *Teaching: A course in applied psychology.* Chicago: Science Research Associates.

Benge, T. J. (1996). Movements utilized by conductors in the stimulation of expression and musicianship. (Doctoral dissertation, University of Southern California). *Dissertation Abstracts International,* 58(01), 18A (University Microfilms No. 9720183)

Bergee, M. J. (1995). Primary and higher-order factors in a scale assessing concert band performance. *Bulletin of the Council for Research in Music Education, 126,* 1–14.

Berz, W. L. (1983). The development of an observation instrument designed to classify specific nonverbal communication techniques employed by conductors of musical ensembles (Doctoral dissertation, Michigan State University). *Dissertation Abstracts International,* 44(09), 2702A (University Microfilms No. 8400532)

Blocher, L., Greenwood, R., & Shellahamer, B. (1997). Teaching behaviors of middle school and high school band directors in the rehearsal setting. *Journal of Research in Music Education, 45*(3), 457–469.

Brand, M., & Burnsed, V. (1981). Music abilities and experiences as predictors of error-detection skill. *Journal of Research in Music Education, 29*(2), 91–96.

Brophy, J. E., & Good, T. L. (1986). Teaching behavior and student achievement. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.) (pp. 328–375). New York: Macmillan.

Bull, P. (1983). *Body movement and interpersonal communication.* New York: Wiley.

Burgoon, J. K., Buller, D. B., Hale, J. L., & deTurck, M. A. (1984). Relational messages associated with nonverbal behaviors. *Human Communication Research, 10*(3), 351–378.

Burnsed, V., & King, S. (1987). How reliable is your festival rating?. *Update: Applications of Research in Music Education, 5,* 12–13.

Byo, J. L. (1990). Recognition of intensity contrasts in the gestures of beginning conductors. *Journal of Research in Music Education, 38*(3), 157–163.

Byo, J. L. (1997). The effects of texture and number of parts on the ability of music majors to detect performance errors. *Journal of Research in Music Education, 45*(1), 51–66.

Byo, J. L., & Austin, K. R. (1994). Comparison of expert and novice conductors: An approach to the analysis of nonverbal behaviors. *Journal of Band Research, 30*(1), 11–34.

Byo, J. L., & Lethco, L. (2001). Student musicians' eye contact with the conductor: An exploratory investigation. *Contributions to Music Education, 28*(2), 21–35.

Byo, J. L., & Sheldon, D. A. (2000). The effect of singing while listening on undergraduate music majors' ability to detect pitch and rhythm errors. *Journal of Band Research, 36*(1), 26–46.

Cassidy, J. W. (1990). Effect of intensity training on preservice teachers' instruction, accuracy, and delivery effectiveness. *Journal of Research in Music Education, 38*(3), 164–174.

Cavitt, M. E. (1998). A descriptive analysis of error correction in expert teachers' instrumental music rehearsals. (Doctoral dissertation, University of Texas–Austin). *Dissertation Abstracts International*, 59(06), 1958A (University Microfilms No. 9837917)

Chesterman, R. (Ed.) (1976). *Conversations with conductors.* London: Robson.

Cofer, R. S. (1998). Effects of conducting-gesture instruction on seventh-grade band students' performance response to conducting emblems. *Journal of Research in Music Education, 46*(5), 360–373.

Davis, A. P. (1998). Performance achievement and analysis of teaching during choral rehearsals. *Journal of Research in Music Education, 46*(4), 496–509.

Duke, R. A. (1999/2000). Measure of instructional effectiveness in music. *Bulletin of the Council for Research in Music Education, 143*, 1–48.

Duke, R. A., & Henninger, J. C. (1998). Effects of verbal corrections on student attitude and performance. *Journal of Research in Music Education, 46*(4), 482–495.

Duke, R. A., & Madsen, C. K. (1991). Proactive versus reactive teaching: Focusing observation on specific aspects of instruction. *Bulletin of the Council for Research in Music Education, 108*, 1–14.

Duke, R. A., Prickett, C. A., & Jellison, J. A. (1998). Empirical description of the pace of music instruction. *Journal of Research in Music Education, 46*(2), 265–280.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica, 1*(3), 49–98.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face.* Englewood Cliffs, NJ: Prentice-Hall.

Faulkner, R. A. (1973). Orchestra interactions: Some features of communication and authority in an artistic organization. *Sociological Quarterly, 14*, 147–157.

Fiocca, P. D. H. (1989). A descriptive analysis of the rehearsal behaviors of exemplary junior high and middle school choir directors. *Contributions to Music Education, 16*, 19–33.

Francisco, J. M. (1994). Conductor communication in the ensemble rehearsal: The relative effects of verbal communication, visual communication, and modeling on performance improvement of high school bands. *Dissertation*

*Abstracts International*, 56(08), 3045A (University Microfilm No. AAC 95-42638).

Fredrickson, W. E. (1992). Research on eye contact with implications for the conductor: A review of literature. *Update: Applications of Research in Music Education, 11*(1), 25–31.

Fredrickson, W. E., Johnson, C. M., & Robinson, C. R. (1998). The effect of preconducting and conducting behaviors on the evaluation of conductor competence. *Journal of Band Research, 33*(2), 1–13.

Goodlad, J. I. (1983). *A place called school.* New York: McGraw-Hill.

Goolsby, T. W. (1996). Time use in instrumental rehearsals: A comparison of experienced, novice, and student teachers. *Journal of Research in Music Education, 44*(4), 286–303.

Goolsby, T. W. (1997). Verbal instruction in instrumental rehearsals: A comparison of three career levels and preservice teachers. *Journal of Research in Music Education, 45*(1), 21–40.

Goolsby, T. W. (1999). A comparison of expert and novice music teachers' preparing identical band compositions: An operational replication. *Journal of Research in Music Education, 47*(2), 174–187.

Grechesky, R. N. (1985). An analysis of nonverbal and verbal conducting behaviors and their relationship to expressive musical performance (Doctoral dissertation, University of Wisconsin-Madison, 1985). *Dissertation Abstracts International, 46*(10), 2956A (University Microfilms No. 8513459)

Hamann, D. L., Mills, C., Bell, J., Daugherty, E, & Koozer, R. (1990). Classroom environment as related to contest ratings among high school performing ensembles. *Journal of Research in Music Education, 38*(3), 215–224.

Holt, M. E. (1992). The application of conducting and choral rehearsal pedagogy of Laban effort/shape and its comparative effect upon style in choral performance. *Dissertation Abstracts International*, 53 (02), 437A. (University Microfilm No. AAC92–14439).

Izard, C. E. (1997). Emotions and facial expressions: A perspective from differential emotions theory. In J. A. Russel and J. M. Fernandez-Dols (Eds.), *The psychology of facial expression.* (pp. 57–77). Cambridge: Cambridge University Press.

Kaiser, K. A. (1998). The effect of differentiated high- versus low-intensity teaching on band musicians' evaluation of teaching effectiveness. *Dissertation Abstracts International*, 59(03), 764A. (University Microfilms No. 9827669)

Kelly, S. N. (1997). Effects of conducting instruction on the musical performance of beginning band students. *Journal of Research in Music Education, 45*(2), 295–305.

Kennell, R. P. (1989). *Three teacher scaffolding strategies in college instrumental applied instruction.* Unpublished doctoral dissertation, University of Wisconsin.

Laban, R. (1975). *Laban's principles of dance and movement notation.* Boston: Plays.

Laib, J. R. (1993). The effect of expressive conducting on band performance (Doctoral dissertation, University of Georgia, 1993). *Dissertation Abstracts International*, 54(02), 3258A. (University Microfilms No. 9404667)

Madsen, C. K. (1990). Teacher intensity in relationship to music education. *Bulletin of the Council for Research in Music Education, 104*, 38–46.

Madsen, C. K. and Yarbrough, C. (1985). *Competency-based music education.* Raleigh, NC: Contemporary Publishing.

Mayne, R. G. (1993). An investigation of the use of facial expression in conjunction with musical conducting gestures and their interpretations by instrumental performers (Doctoral dissertation, Ohio State University, 1992). *Dissertation Abstracts International*, 53(08), 2729A (University Microfilms No. 9238229)

McCoy, C. W. (1985). The ensemble director as effective teacher. *Update: Applications of Research in Music Education*, *3*(3), 9–12.

Miller, S. (1988). The effect of Laban Movement Theory on the ability of student conductors to communicate musical interpretation through gesture (Doctoral dissertation, University of Wisconsin-Madison, 1988). *Dissertation Abstracts International*, 49(05), 1087A. (University Microfilms No. 8809040)

Oertle, E. (1999). Bud Beyer: Bodyspeak. *Journal of the World Association for Symphonic Band and Ensembles*, *6*, 19–21.

Parasuraman, S., and Nachman, S. A. (1987). Correlates of organizational and professional commitment: The case of musicians in a symphony orchestra. *Group and Organization Studies*, *12*(3), 287–303.

Preston-Dunlop, V. (1980). *A handbook for modern educational dance* (new rev. ed.). Boston: Plays.

Price, H. E. (1983). The effect of conductor academic task presentation, conductor reinforcement, and ensemble practice on performers' musical achievement, attentiveness, and attitude. *Journal of Research in Music Education*, *31*(4), 245–257.

Price, H. E. (1985). A competency-based course in basic conducting techniques: A replication. *Journal of Band Research*, *21*(1), 61–69.

Price, H. E. (1992). Sequential patterns of music instruction and learning to use them. *Journal of Research in Music Education*, *40*(1), 14–29.

Price, H. E., & Winter, S. (1991). Effect of strict and expressive conducting on performances and opinions of eighth-grade band students. *Journal of Band Research*, *27*(1), 30–43.

Reimer, B. (2000). Why do humans value music? In C. K. Madsen (Ed.), *Vision 2020: The Housewright symposium on the future of music education* (pp. 25–48). Reston, VA: MENC: The National Association for Music Education.

Sang, R. C. (1987). A study of the relationship between instrumental music teachers' modeling skills and pupil performance behaviors. *Bulletin of the Council for Research in Music Education*, *91*, 155–159.

Seashore, C. E. (1938/1967). *The psychology of music*. New York: Dover.

Sherrill, M. H. (1986). An analytical study of videotaped rehearsal and conducting techniques of selected junior high and senior high school band conductors (Doctoral dissertation, University of Rochester, Eastman School of Music, 1986). *Dissertation Abstracts International*, *47*(04), 1231A. (University Microfilms No. 8614150)

Sidoti, V. J. (1990). The effects of expressive and nonexpressive conducting on the performance accuracy of selected expressive markings by individual high school instrumentalists (Doctoral dissertation, Ohio State University). *Dissertation Abstracts International*, *51*(10), 3270A. (University Microfilms No. 9105212)

Single, N. A. (1990). An exploratory study of pacing in instrumental music rehearsals. *Contributions to Music Education*, *17*, 32–43.

Skadsem, J. A. (1996). The effect of verbal, written, gestural, and choral stimuli on singers' performance responses to dynamic changes in music. *Missouri Journal of Research in Music Education*, *33*, 28–44.

Skadsem, J. A. (1997). Effect of conductor verbalization, dynamic markings, conductor gesture, and choir dynamic level on singers' dynamic responses. *Journal of Research in Music Education, 45*(4) 509–520.

Sousa, G. (1988). Musical conducting emblems: An investigation of the use of specific conducting gestures and their interpretation by instrumental performers (Doctoral dissertation, Ohio State University, 1988). *Dissertation Abstracts International, 49*(08), 2143A. (University Microfilms No. 8820356)

Standley, J. M., & Madsen, C. K. (1987, Summer). Intensity as an attribute of effective teaching. *Quodlibet*, 15–20.

Tait, M. (1985). Striving to become a creative artist. *CBDNA Journal, 2*(1), 25–28.

Tipton, D. G. (1996). *The use of sequential patterns in select children's, youth, and adult Southern Baptist choir rehearsals.* (Doctoral dissertation, University of Alabama).

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Michael Cole, Vera John-Steiner, Sylvia Scribner, and Ellen Souberman (Eds.). Cambridge, MA: Harvard University Press.

Watkins, R. E. (1996). Nonperformance time use in high school choral rehearsals: A follow-up study. *Update: Applications of Research in Music Education, 14*(1), 4–8.

Weeks, P. (1996). A rehearsal of a Beethoven passage: An analysis of correction talk. *Research on Language and Social Interaction, 29*(3), 247–290.

Yarbrough, C. (1975). Effect of magnitude of conductor on students in selected choruses. *Journal of Research in Music Education, 23*(2), 134–146.

Yarbrough, C. (1988). Content and pacing in music teaching. *Proceedings of Symposium on Current Issues in Music Education, 13*, 9–28.

Yarbrough, C., Dunn, D. E., & Baird, S. L. (1996). A longitudinal study of teaching in a choral rehearsal. *Southeastern Journal of Music Education, 8*, 7–31.

Yarbrough, C., and Price, H. E. (1981). Prediction of performer attentiveness based upon rehearsal activity and teacher behavior. *Journal of Research in Music Education, 29*(3), 209–217.

Yarbrough, C., and Price, H. E. (1989). Sequential patterns of instruction in music. *Journal of Research in Music Education, 37*(3), 179–187.

Skadsem, J. A. (1997). Effect of conductor verbalization, dynamic markings, conductor gesture, and choir dynamic level on singers' dynamic responses. *Journal of Research in Music Education, 45*(4) 509–520.

Sousa, G. (1988). Musical conducting emblems: An investigation of the use of specific conducting gestures and their interpretation by instrumental performers (Doctoral dissertation, Ohio State University, 1988). *Dissertation Abstracts International, 49*(08), 2143A. (University Microfilms No. 8820356)

Standley, J. M., & Madsen, C. K. (1987, Summer). Intensity as an attribute of effective teaching. *Quodlibet*, 15–20.

Tait, M. (1985). Striving to become a creative artist. *CBDNA Journal, 2*(1), 25–28.

Tipton, D. G. (1996). *The use of sequential patterns in select children's, youth, and adult Southern Baptist choir rehearsals.* (Doctoral dissertation, University of Alabama).

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Michael Cole, Vera John-Steiner, Sylvia Scribner, and Ellen Souberman (Eds.). Cambridge, MA: Harvard University Press.

Watkins, R. E. (1996). Nonperformance time use in high school choral rehearsals: A follow-up study. *Update: Applications of Research in Music Education, 14*(1), 4–8.

Weeks, P. (1996). A rehearsal of a Beethoven passage: An analysis of correction talk. *Research on Language and Social Interaction, 29*(3), 247–290.

Yarbrough, C. (1975). Effect of magnitude of conductor on students in selected choruses. *Journal of Research in Music Education, 23*(2), 134–146.

Yarbrough, C. (1988). Content and pacing in music teaching. *Proceedings of Symposium on Current Issues in Music Education, 13*, 9–28.

Yarbrough, C., Dunn, D. E., & Baird, S. L. (1996). A longitudinal study of teaching in a choral rehearsal. *Southeastern Journal of Music Education, 8*, 7–31.

Yarbrough, C., and Price, H. E. (1981). Prediction of performer attentiveness based upon rehearsal activity and teacher behavior. *Journal of Research in Music Education, 29*(3), 209–217.

Yarbrough, C., and Price, H. E. (1989). Sequential patterns of instruction in music. *Journal of Research in Music Education, 37*(3), 179–187.

# Contributors

RITA AIELLO is a visiting scholar in the psychology department of New York University. She is the editor of *Musical Perceptions* (Oxford University Press, 1994). Her research addresses the mental representation of music, the effect of musical training on the perception and cognition of music, and parallels between music and language. She has taught courses on the perception of music at the Juilliard School, the Manhattan School of Music, New York University, and Queens College of the City University of New York and has been a visiting honorary adjunct professor at Teachers College, Columbia University. She was trained as a classical pianist and has extensive teaching experience.

ECKART ALTENMÜLLER is professor at the Hannover University of Music and Drama and director of its Institute of Music Physiology and Performing Arts Medicine. He graduated in medicine (1983) and music (1985) and trained as a neurologist and neurophysiologist (Habilitation 1992). Since 1985 he has been researching brain processing of music and motor learning in musicians. He has published 46 scientific articles in peer-reviewed journals and 60 book contributions. Major awards include the Young Scientists Award of the German Physiological Society (1984), the Kornmüller-Award of the German Society for Clinical Electrophysiology (1992), and the Richard Lederman Lecture, Aspen (1999).

NANCY H. BARRY holds a Ph.D. in music education and certificates in computers in music and electronic music from Florida State University. She is currently professor and coordinator of graduate music education at the School of Music of the University of Oklahoma. She serves on the editoral boards of *Update: Applications of Research in Musical Education, Professional Educator,* and *Journal of Technology in Music Learning.* She has published numerous articles in journals such as *Psychology of Music, Update, Arts and Learning,* and *Journal of Music Teacher Education.* She has given over 40 presentations at conferences and seminars in the United States, Puerto Rico, Brazil, Germany, and Australia.

She is also an experienced studio teacher and frequent adjudicator of woodwinds and piano.

GIOVANNI UMBERTO BATTEL is director of the Conservatory of Music "B. Marcello" in Venice. He holds a degree and a postgraduate diploma from the Tartini Conservatory in Trieste and the Accademia di Santa Cecilia in Rome and a degree in musicology from Bologna University. His published articles address new methodologies of performance analysis. As a pianist, he has won six Italian national competitions and prizes at seven international competitions, including first prizes in Stresa and Enna and other prizes at the Viotti and Busoni competitions. He has given concerts throughout Europe and the United States, playing with chamber and symphony orchestras conducted by De Bernant, Gavazzeni, and Steinberg. Battel has also recorded for Italian radio and television (RAI).

ALICE G. BRANDFONBRENER received her undergraduate education at Wellesley College and her doctor of medicine from Columbia University. She is assistant professor of medicine at the Northwestern University Medical School and adjunct professor of performance studies at the Northwestern University School of Music. She directs the Medical Program for Performing Artists in Chicago and is editor of the journal *Medical Problems of Performing Artists*. She founded the Aspen symposium on medical problems of musicians and dancers in 1983 and was first president of the Performing Arts Medicine Association. A frequent lecturer to medical and musical audiences, she has written many articles for music education as well as medical journals on musicians' medical problems.

JAMES L BYO is professor of music education at Louisiana State University and editor of *Update: Applications of Research in Music Education* (published by MENC: National Association for Music Education). In 1997, he completed a six-year term as chair of the International Wind Band Education Committee of the World Association for Symphonic Bands and Ensembles. In this capacity, he presented research and presided over sessions in England, Spain, Japan, and Austria. His articles on issues relevant to music teacher/conductor effectiveness appear in all major U.S. research journals in music education. He is the recipient of two distinguished professor awards at LSU.

JORGE SALGADO CORREIA is lecturer at the University of Aveiro and a professional flute soloist. He graduated in philosophy from the University of Porto and studied flute performance at Porto Conservatoire and at the University of Amsterdam. His areas of research include the investigation of teaching and learning of performance skills from historical, analytical, psychological, sociological, and philosophical perspectives and reflexivity in musical performance: defining and developing practical music making as research. His specialist performance area is contemporary music, and he has given many world premieres and has had works specially commissioned for him. His publications include a recent article in *Psychology of Music* and a number of articles written in collaboration with Jane Davidson.

JANE W. DAVIDSON studied music, dance, and education at Newcastle-upon-Tyne and London before undertaking advanced vocal studies in Canada. She completed her M.A. in performance studies at City University (London) and after a period of time teaching in schools gained her Ph.D. in the psychology of music. She worked in the psychology department at Keele University and the music department at City University before taking up her current post as senior lecturer in music at the University of Sheffield. She is a former editor of the international journal *Psychology of Music*. Besides her teaching and research work, she performs as a vocal soloist and has appeared at the Queen Elizabeth Hall, St. Paul's Cathedral, and the London International Opera Festival.

HEINZ FADLE is professor of trombone at the Hochschule für Musik, Detmold, Germany. His publications include *Looking for the Natural Way* (Piccolo-Verlag, 1996) and *Musizieren* (Die Blaue Eule, 1999). He served as president of the International Trombone Association from 1996 until 1998. Invitations to teach and lecture have led him to the United States, Russia, Kazakhstan, Britain, and Hungary. He hosted the 21st International Trombone Workshop in Detmold (1992) and also served as a jury member at international competitions. For 22 years he was first chair of the Philharmonic State Orchestra Hamburg, also performing as a member of the Bayreuth Festival Orchestra, the Berliner Philharmoniker, Radio Sinfonie Hamburg, Münchner Staatsoper, and Deutsches Sinfonieorchester Berlin.

ANDERS FRIBERG is a research assistant in music acoustics at the Royal Institute of Technology (KTH), Stockholm. He holds a master's degree in applied physics from KTH. His doctoral thesis, *A Quantitative Rule System for Musical Performance* (1995), developed the computer-based music performance simulator *Director Musices*. His published research on the technology of music performance addresses phrasing, punctuation, intonation, synchronization, swing, and links with human motion and emotion and has appeared in *Computer Music Journal*, *Journal of the Acoustical Society of America*, and *Journal of New Music Research*. He plays piano in salsa and jazz groups and holds a piano performance diploma from Berklee College of Music, Boston.

LEONARDO FUKS is associate professor of music acoustics and voice physiology at Universidade do Brasil/UFRJ in Rio de Janeiro. His musical studies were undertaken at the Villa-Lobos Music School and Uni-Rio University. He has been a professional oboist in several Brazilian orchestras and chamber music groups and holds a degree in mechanical engineering, a master of science degree in design engineering, and a doctor of philosophy in music acoustics from the Royal Institute of Technology (KTH), Stockholm. His main research topics include wind instrument performance and vocal techniques in ethnic and contemporary music. He has created a number of low-cost wind instruments for music education and is the founder and director of a bicycle orchestra, the Cyclophonica.

JANINA FYK holds a master of arts in music education from the Paderewski Academy of Music in Poznan (Poland), a doctoral degree in music theory from the

Chopin Academy of Music in Warsaw, and a habilitation in musicology from the Institute of Arts of the Polish Academy of Sciences in Warsaw. She is now professor of musical arts at the Academy of Music in Krakow. Her areas of research interests include psychoacoustics, music psychology, and music education. She is involved in research that concerns the perception of pitch, vocal, and instrumental intonation as well as absolute pitch. She is author and coauthor of numerous articles and the monographic book *Melodic Intonation, Psychoacoustics, and the Violin.*

ALF GABRIELSSON graduated from Uppsala University and the Royal University College of Music in Stockholm. His Ph.D. thesis was titled *Studies in Rhythm,* and he has subsequently published numerous research articles on music performance and music experience. He teaches general psychology and music psychology and is the head of a research group on expressive performance in music, dance, speech, and body language. He is a member of the Swedish Royal Academy of Music and the Acoustical Society of America and a vice president of the International Association for Empirical Studies of the Arts (IAEA) and was president of the European Society for the Cognitive Sciences of Music (ESCOM) from 1997 to 2000.

MARTIN GELLRICH studied piano, violoncello, secondary teaching, psychology, and musicology and holds a Ph.D. from the Hochschule der Künste und the Technische Universität Berlin in music education and psychology. His research addresses instrumental teaching, developmental psychology of music, instrumental practice, group teaching methods, relative sol-fa, improvisation, history of music education, piano technique, and musical interpretation. His publications include *Üben mit Lis(z)t* (Frauenfeld, 1992), *Impulse zu einer Veränderung der Instrumental pädagogik* (Regensburg, 2002), and articles in the *British Journal of Music Education, Üben and Musizieren,* and *Bulletin of the Council for Research in Music Education.* As a freelance piano teacher he has given over 200 presentations on instrumental teaching, particularly on improvisation.

HEINER GEMBRIS studied music education at the Music Academy of Detmold and musicology at the Technical University of Berlin. From 1991 to 2001 he was professor of musicology at the universities of Münster and Halle. Since 2001 he has been professor of music and music education at the University of Paderborn and director of the Institut für Begabungsforschung und Begabtenförderung in der Musik. He is author of *Grundlagen musikalischer Begabung und Entwicklung* (Wissner, 1998) and coeditor of the series *Musikpädagogische Forschungsberichte* (Wissner, 1992ff). He has published numerous research articles on musical development and music listening. Currently he is the president of the Deutsche Gesellschaft für Musikpsychologie.

WILFRIED GRUHN is professor and head of the music education department at the Freiburg Academy of Music. He has published several books and many articles on the history of music education, on music perception and cognition, on teach-

ing methods, and on neurobiological learning theory. He was coeditor of *Musik und Unterricht* (1980–1998) and founded the first Internet *European Music Journal* (in 1999). He has participated as clinician and presenter at various international conferences and functions as national representative of ISME in Germany. A past president of the International Research Alliance of Institutes for Music Education (RAIME), he is currently a member of the editorial board of *Research Studies in Music Education* and of *Music Education Research*.

KNUT GUETTLER is professor of double bass at the Norwegian Academy of Music and also teaches at the Royal Conservatory of the Hague, the Netherlands. His main research focus is the computer modeling of the bow–string interaction. Another area is the physical properties of bows, in collaboration with the Royal Institute of Technology, Stockholm, He has published in the *Catgut Acoustical Society Journal* (*CAS Journal*), *Journal of the Acoustical Society of America*, and *Music Perception* and is a contributing editor for the *CAS Journal*. He is a former principal player of the Oslo Philharmonic, with international experience as a lecturer on string performance. He has solo recordings on Camerata (Japan) and Musica Viva (Germany), among others.

SUSAN HALLAM trained as a violinist at the Royal Academy of Music, after which she spent 10 years as a full-time professional musician working for the BBC. After completing a B.A. in psychology externally with London University and an M.Sc. and Ph.D. in the psychology of education at the Institute of Education, London University, she took a post in the department of psychology there. After a year at Oxford Brookes University as professor of teaching and learning, she returned to London Unviersity in 2001. She is chair of the education section of the British Psychological Society, and has acted as coeditor of the *BPS Education Section Review*, and editor of *Psychology of Music*.

PATRIK N. JUSLIN holds a Ph.D. in psychology from the Department of Psychology, Uppsala University, Sweden, and currently holds a position there as researcher. He has published extensively on music and emotion, and his research is concerned with developing and testing a new theoretical framework for studies of communication of emotion via musical performance, as well as methodological developments for the field of music education. He is a member of the International Society for Research on Emotions and received ESCOM" Young Researcher Award in 1996. He has edited a book on music and emotion for Oxford University Press together with John Sloboda. In addition to Juslin's work as a researcher, he has worked professionally as a guitar player and toured internationally with blues/jazz bands.

DUANE RICHARD KARNA is director of music ministry at the First Congregational Church in Eugene, Oregon. He was director of choral activities and associate professor of music at Salisbury State University (Maryland) from 1988 to 1996 and a member of the music faculty (voice/choral) at Central Washington from 1996 to 1999. He holds a bachelor's degree in vocal performance from the Uni-

versity of Puget Sound in Tacoma, Washington; master's degrees in vocal performance and in choral conducting from Southern Methodist University in Dallas, Texas; and a doctor of musical arts in choral conducting from the University of Arizona in Tucson. He is a member of the American Choral Directors Association, the International Federation of Choral Music, and the National Association of Teachers of Singing.

ANTHONY E. KEMP is emeritus professor at the University of Reading, where he supervises research students. Originally trained as a musician and teacher, he later studied at Sussex University, gaining his D.Phil. in psychology. He is a chartered psychologist and fellow of the British Psychological Society and also senior research fellow at the University of Surrey Roehampton. He is visiting professor at the Sibelius Academy, Helsinki, where he supervises doctoral students. He was chairman of the research commission and member of the board of directors of the International Society for Music Education and has written several books and articles on music psychology, and education. His book *The Musical Temperament: The Psychology and Personality of Musicians* was published by Oxford University Press in 1996.

BARRY J. KENNY studied music at Sydney University and the University of New South Wales and holds undergraduate degrees in English literature (Sydney University) and music (University of New South Wales). He has lectured and tutored in the School of Music and Music Education, University of New South Wales, in musical analysis and jazz history and is in the process of completing his doctoral studies. Barry's published work addresses the cognitive mechanisms involved in jazz improvisation and also discuss sociocultural issues that surround the interpretation and analysis of musical forms. His work has been published in *Research Studies in Music Education* and *Annual Review of Jazz Studies.* As a professional musician, he has been actively involved as a jazz pianist, teacher, and accompanist.

JAMES M. KJELLAND holds a bachelor's degree in music education, a master of music degree from the University of Wisconsin, Madison, and a doctorate in music education from the University of Texas at Austin. He is associate professor of music education and coordinator of string pedagogy at Northwestern University. He has also presented at numerous conferences in string and orchestra pedagogy throughout the United States and is a frequent guest conductor, adjudicator, and clinician for school orchestras. His publications include his coauthored *Strictly Strings: A Comprehensive String Class Method* (Highland/Etling) and numerous articles in journals such as: *American String Teacher, Instrumentalist,* and *Bulletin of the Council for Research in Music Education.*

ANDREAS C. LEHMANN researches and teaches music psychology at the School of Music in Würzburg, Germany. He holds a degree in music education and a Ph.D. in musicology from the Hochschule für Musik und Theater, Hannover, and

worked as a researcher in the psychology department at Florida State University, Tallahassee, and in the musicology department of the University of Halle, Germany. His research addresses the structure and acquisition of high-level instrumental music performance skills (sight-reading, memory, improvisation, practice). In 1997, he received the Young Researcher Merit Award from the European Society for the Cognitive Sciences of Music (ESCOM). With Harald Jørgensen, he is coeditor of *Does Practice Make Perfect? Current Theory and Research on Instrumental Music Practice* (Norges Musikkhøgskole).

VICTORIA MCARTHUR is program director of piano pedagogy and coordinator of group piano at Florida State University. She has published in *Psychology of Music* and *Bulletin of the Council for Research in Music Education*. She is a member of the editorial board of *Piano and Keyboard*, former senior editor for FJH Music Company, Miami, and currently keyboard editor for Alfred Publishing Company, Los Angeles. She has lectured extensively and authored and edited instructional books on sight-reading. Her other research areas are practice and motor memory. Her background in piano performance, piano pedagogy, music education research methods, cognitive and motor psychology, and biomechanics has formed her ideas and approach to the teaching of piano sight-reading.

GARY E. MCPHERSON studied music education at the Sydney Conservatorium of Music before completing a master of music education at Indiana University and a doctorate of philosophy at the University of Sydney. He is an associate professor of music education at the University of New South Wales and former treasurer for the International Society for Music Education and national president of the Australian Society for Music Education. His published research addresses visual, aural, and creative aspects of musical performance in young developing musicians. McPherson is currently on the editorial boards of all flagship journals in music education, as well as being editor for *Research Studies in Music Education*. As a trumpeter, he has performed with several of Australia's leading ensembles.

JANET MILLS gained her B.A. in music and mathematics from York University and her D.Phil. at Oxford University. She began her career as a secondary music teacher and then taught in higher education before becoming an H.M. Inspector of Schools in 1990. Currently she is a research fellow at the Royal College of Music in London.

STEVEN J. MORRISON is assistant professor of music education at the University of Washington, where he teaches courses in instructional methods and research. He is an instrumental music specialist whose research includes perceptual and performance aspects of intonation and the relationship of cultural context to music preference. He is currently investigating neural aspects of cross-cultural music comprehension in a joint research project with the University of Washington medical center. His articles have appeared in the *Journal of Research in Music Education*, *Bulletin for the Council of Research in Music Education*, *Up-*

*date: Applications of Research in Music Education*, and *Music Educators Journal*. He is chair of the Perception and Cognition Special Research Interest Group of the Society for Research in Music Education.

SUSAN A. O'NEILL is senior lecturer in psychology and Associate Director of the Unit for the Study of Musical Skill and Development at Keele University. Her research interests include motivation, identity, and gender issues associated with children's and adolescents' development of musical performance skills. She has published widely in the fields of psychology and music education. As a consultant for the Associated Board of the Royal Schools of Music she has presented seminars and workshops for professional development and teacher training courses throughout the United Kingdom. She is currently project leader for the Young People and Music Participation Project, which involves a longitudinal study of young people's motivation, identity, and engagement in music. She has taught flute and has performed extensively as a soloist and chamber musician in Canada and Europe.

RICHARD PARNCUTT holds undergraduate degress in in physics and music from the University of Melbourne and a Ph.D. in physics, psychology, and music from the University of New England, Australia. He is now professor of systematic musicology at the University of Graz, Austria. He is author of *Harmony: A Psychoacoustical Approach* (Springer, 1989) and research articles on the perception of harmony, tonality, rhythm, and piano performance, including psychoacoustic and cognitive theories and models. He is a member of the editorial advisory boards of *Psychology of Music*, *Music Perception*, *Musicae Scientiae*, *Journal of New Music Research*, *Jahrbuch Musikpsychologie*, *Systematische Musikwissenschaft*, *Contemporary Music Review*, and *Research Studies in Music Education*. He is an internationally experienced pianist, piano teacher, and accompanist.

ROLAND S. PERSSON holds a master of arts in music education from Ingesund College of Music (Sweden) and a Ph.D. in psychology from Huddersfield University (England). He is currently associate professor of psychology at Jönköping University and director of its center for psychology. He is editor in chief of *High Ability Studies* (Taylor & Francis) and a former piano performer, organist, and music educator. His research interests fall within the domain of social psychology, particularly talent and giftedness, gender issues, and psychometrics, where he has written numerous scholarly articles on the training of musical performers and their psychological health and education.

HARRY E. PRICE is professor and head of music education at the University of Alabama. In addition to his current research and teaching, he was a classical and jazz performer and secondary school and university band conductor and clinician. He is the former editor of the *Journal of Research in Music Education* and is on the editorial board of *Research Studies in Music Education* and the *International Journal of Music Education*. He edited *Music Education Research: An*

*Anthology* from the *Journal of Research in Music Education* and has over 50 research articles in publications such as *Journal of Research in Music Education, Bulletin of the Council for Research in Music Education, Research Studies in Music Education,* and *Boletin de Investigación Educativo-Musica.*

DAVID ROLAND is a consultant clinical psychologist and a classical guitarist. He completed an honors degree at the University of Sydney and a Ph.D. in clinical psychology at the University of Wollongong. His area of research interest is the management of performance anxiety and the application of mental skills to performing. He is the author of *The Confident Performer* (Currency Press, 1997) and has published in *Research Studies in Music Education.* He has presented his research to the International Society of Music Education (ISME Korea, 1992). He has been on the staff of the Canberra School of Music (Australian National University); guest-lectured for many universities and conservatories, and worked with a number of performing arts groups.

JOHAN SUNDBERG is a professor of music acoustics at the Royal Institute of Technology in Stockholm, where he led its music acoustics research group from 1970 to 2001. A prolific author of peer-reviewed journal articles, his main research interests are the singing voice and music performance. As president of the Music Acoustics Committee of the Royal Swedish Academy of Music, he edited 10 volumes of proceedings of public seminars. He is a member of the Royal Swedish Academy of Music and the Swedish Acoustical Society (President 1976–1981) and a fellow of the Acoustical Society of America. He has sung in the Stockholm Bach Choir (nine years as its president) and made his public solo debut with a lieder recital on his fiftieth birthday.

STEN TERNSTRÖM holds a doctoral degree from the Royal Institute of Technology (KTH), Stockholm, and is currently senior researcher and university lecturer there. He is the author of numerous research articles on the acoustics of choir singing and other topics in music acoustics and voice, often collaborating with Johan Sundberg. Ternström is associate editor of *Logopedics Phoniatrics Vocology,* board member of the Swedish Voice Society, and member of the scientific advisory board of the Voice Foundation in Philadelphia. A choir singer and arranger, he has performed extensively with leading a cappella vocal groups in Sweden.

MALCOLM TROUP is the founding (and now emeritus) professor of the music department of City University, London. He holds a D.Phil. in music from the University of York and honorary doctorates from City University, the Memorial University of Newfoundland, and the University of Chile. He is currently the European chair of the European Piano Teachers' Association and edits its *Piano Journal.* He is also chair of the Beethoven Society of Europe, president of the annual Oxford Philomusica Piano Festival, and vice president of the World Piano Competition, London. His research addresses performance practice, history of ideas, and aesthetics. As a solo pianist he has toured five continents, received the 1998 Liszt medal, and recorded for RCA Victor and the Continuum label.

GRAHAM F. WELCH (Ph.D.) is professor of music education at the University of London Institute of Education. He is a trained singer and semiprofessional performer for over 20 years whose research and associated publications include singing development across the life span, particularly in childhood and adolescence and including the trained chorister voice. Recent publications include *Bodymind and Voice: Foundations of Voice Education*, coedited and coauthored with Leon Thurman for the National Centre for Voice and Speech, and chapters on voice management, musical development and learning in the United Kingdom, and the teaching profession. He is currently chair of the Society for Research in the Psychology of Music and Music Education and cochair of the ISME Research Commission.

AARON WILLIAMON is a research fellow in psychology of music at the Royal College of Music. He completed his Ph.D. in cognitive psychology at the University of London (Royal Holloway), where his research centered on how skilled musicians learn, rehearse, and recall information when practicing and performing. In 1998, he was awarded the Hickman Prize by the Society for Research in Psychology of Music and Music Education (SRPMME) for his work on how audiences evaluate memorized performances. In addition to his continued interest in music cognition (particularly memorization and practice), he is currently researching the nature and effects of performance anxiety and how skilled musicians work together when learning and rehearsing compositions.

GLENN D. WILSON is reader in personality at the University of London Institute of Psychiatry and adjunct professor at the University of Nevada, Reno. He is author of some 200 scientific papers and 25 books, including *Psychology for Performing Artists* (Whurr, 2002) and (with Andrew Evans) *Fame: The Psychology of Stardom* (Vision, 1999). He makes frequent radio and TV appearances and has lectured widely abroad, for example, as guest of the Italian Cultural Association and as visiting professor at several U.S. institutions, including Stanford and California State University, Los Angeles. He is also a trained opera singer with an international freelance career.

# Author Index


Abel, J. L., 49
Abeles, H. F., 12, 101
Ablard, K., 56
Abramson, R., 246
Adams, M. J., 104, 110
Agostoni, E., 326
Agras, W. S., 51, 53
Aiello, R., 173–177
Alexander, F. M., 54, 92, 246
Allen, M., 52
Altenmüller, E., 73, 76, 77, 79, 286
American Standards Association, xi
Ames, C., 39
Amunts, K., 71
Anderson, J., 171
Anderson, J. N., 158
Anderson, K. J., 50
Appel, S. S., 52
Argyle, M., 242
Arnold, J. A., 37, 340
Asendorpf, J. B., 18
Askenfelt, A., 286, 288, 290, 292, 295, 308, 309, 311
Asmus, E., 42
Austin, J., 42
Austin, J. R., 36, 37, 40, 42
Austin, K. R., 343, 344

Babits, L., 92
Bach, C. P. E., 209, 220
Backus, J., 322
Baddeley, A. D., 170, 171
Badur, I.-M., 21
Bailey, M., 107, 158
Baird, S. L., 337
Bak, N., 327
Baker-Sennett, J., 128
Ballard, D., 187, 188
Balzer, W. K., 229
Bamberger, J., 105
Bandura, A., 34, 40
Bangert, M., 76, 79
Banowetz, J., 292, 293
Banton, L., 140, 143
Barbenal J., 323
Barbour, J. M., 184
Barney, H. L., 281
Baron, J. G., 289, 290, 291
Barry, N. H., 154, 155, 157, 159, 160
Bartel, L. R., 157
Bartlett, F. C., 169
Bashmet, Y., 224
Bastian, H. G., 22
Battel G. U., 204, 209, 212, 214
Beach, D. R., 155
Beattie, G., 242
Beck, A., 49
Becker, W. C., 340
Bee, H., 17
Behne, K.-E., 290
Behrens, G. A., 222
Bell, J., 336
Benade, A. H., 321, 322
Benge, T. J., 345
Bengtsson, I., 208
Benson, W., 188
Bentley, A., 5
Bergee, M. J., 344
Berliner, P., 117, 120, 126–128
Bernouilli, D., 321–323

Bernstein, S., 168, 175
Berz, W. L., xi, 343
Bigand, E., 206
Birkett, J., 127
Blatchford, P., 27
Blocher, L., 343
Bloom, B. S., 6
Blum, D., 220
Boal, A., 247
Boehm, T., 325
Borko, H., 126
Bouasse, H. P. M., 321
Bouchard, T. J., 18
Bouhuys, A., 323, 325, 326, 330
Bourassa, G., 185
Bourne, E., 48, 56, 57
Bowers, J., 188, 189
Bowman, J., xi
Boyd, J., 220
Boyle, J. D., 147
Brand, M., 338
Brandfonbrener, A. G., 88, 91–93
Brandstrom, S., 159
Bregman, A. S., 137, 293, 296
Breithaupt, R. M., 287
Bresin R., 206, 209, 210, 212, 214, 224
Bridger, W. H., 5
Brittin, R. V., 188
Brokaw, J. P., 159
Brophy, J. E., 337
Brown, A. F. D., 329
Brown, E. F., 18
Brown, J. C., 189, 289
Brown, S. F., 186, 189
Bruce, R., x
Bruner, J., 109
Brunning, R. H., 104, 109, 111
Brunswik, E., 219, 225
Bull, P., 343
Buller, D. B., 343, 344
Burgoon, J. K., 343, 344
Burland, K., 18
Burmeister, E., 147, 148
Burns, E. M., 184, 185, 187
Burnsed, V., 338, 344
Burton, D., 56
Butters, N., 75
Byl, N. N., 77
Byo, J. L., 338, 339, 343, 344

Callaghan, J., 266
Calvin, W. H., 79
Cambouropoulos, E., 210
Campbell, K., 18
Campbell, M., 320
Campbell, S. L., 187
Candia, V., 88
Cantwell, R. H., 155
Carpenter, P. A., 171
Casella, A., 294
Cassidy, J. W. 188, 339
Cattell, R. B., 5, 9
Cavitt, M. E., 336, 337
Cayea, D., 85
Cayne, B. S., 151
Chaffin, R., 156, 157, 172, 174, 176
Chang, H. W., 5
Chase, W. G., 170, 171, 173–175
Chen, F. C., 323
Chesterman, R., 338
Chinsky, J. M., 155
Choksy, L., 246
Chomsky, N., 121
Clark, D. B., 51, 53
Clark, G., 316
Clarke, E. F., xi, 121, 172, 199, 204, 206, 210, 238, 240, 295, 297–299
Clynes, M., 220
Cofer, R. S., 345
Coffman, D. D., 154
Cohen, L. G., 88
Colacone, A., 328
Colldén, A., 276, 277
Collier, G. L., 202
Collier, J. L., 22, 202
Collyer, C. E., xi
Coltman, J. W., 325, 327
Colwell, R., 147
Cook, N., 174, 220
Cooke, D., 220
Cooke, M., 33
Cooksey, R. W., 225, 230
Cooper, C. L., 51
Copley, D. C., 323
Correia, J. S., 246, 247
Corso, J. F., 188
Cortot, A., 77, 298
Cossette, I., 326, 328, 330
Covington, M. V., 42
Cox, W. J., 48, 49
Cramon, D. Y., 75
Crawford, M., 174
Csikszentmihalyi, M., 8, 35, 58, 119
Cuddy, L. L., 290
Cutietta, R., 316
Cutting, J. E, 240, 241

Dalcroze, E. J.-, 109, 245–246
Damasio, A., 238

Daugherty, E., 336
Daugherty, J. F., 282
Davidson, J. W., 6, 8, 10, 22, 23, 25, 26, 41–43, 58, 152, 159, 169, 239, 240, 241, 242, 243, 246
Davidson, L., 105, 109, 228
Davies, J. B., 18, 323
Davies, S., 220
Davis, A. P., 342
Davis, T. L., 48
Day, H., 49
De Poli, G., 209
Debussy, C. , 296
DeCasper, A. J., 20
Deecke, L., 71
DeFries, J. C., 17, 19, 20
Dejonckere, P. H., 261
Delalande, F., 243
Demorest, S. M., 186
Deppe, L., 287
Desain, P., 214, 295
deTurck, M. A., 343, 344
Detweiler, M., 171
Deutsch, L., 299
Dibben, N., 220
Dickey, M. R., 158
Dodson, J. D., 50, 51
Doemler, P., 327
Doherty, M. E., 229
Doing, J., 263
Donahue, W. A., 52
Donaldson, L., 127
Dowling, W. J., 104, 118
Drake, C., 204, 210
Druckman, D., 93, 94
du Pré, J., 10, 19
Dubal, D., 221
Duke, R. A., 188, 336, 337, 340
Dunn, D. E., 337
Dweck, C. S., 38, 39

Easton, C., 10
Ebbinghaus, H., 169
Eccles, J. S., 32, 33, 34, 35, 37, 40
Edgerton, M. E., 264, 279
Edmonson, F. A., 186
Eisenberg, J., 295, 296
Ekman, P., 221, 239, 343, 345
Elbert, T., 72, 77
Elliott, C. A., 147
Ellis, A., 85, 242
Ellis, M. C., 210
Ely, M. C., 186, 188
Emery, G., 49
Engelien, A., 72
Englemann, S., 340
Érard, S., 288
Ericsson, K. A., 126, 137, 138, 140–143, 152, 156, 170–174, 229, 238, 315
Estill, J., 264

Fadle, H., 331
Fastl, H., 186
Faulkner, R. A., 335, 336
Fazey, D. M. A., 103
Feldenkrais, M., 92
Fidler, H., 48, 51, 52, 53
Fimbianti, R., 204, 209, 212
Findlay, J. M., 137
Fiocca, P. D. H., 341
Fischer, L., 24
Fishbein, M., 85, 93
Fitts, P. M., 156, 287
Flavell, J. H., 155
Fleischer, L., 84
Fletcher, C. R., 170
Fletcher, H., 271, 288
Fletcher, N. H., xi, 289, 294, 321, 322, 325–329
Flett, G., 49
Floyd, S. A., Jr., 119, 131
Fox, J., 328
Francisco, J. M., 347
Frederiksen, B., 331
Fredrickson, W. E., 336, 343, 344
Fredrikson, M., 47
Freeman, P., 185
Freund, H.-J., 71
Frey, E., 288, 297
Friberg, A., 204, 206–208, 210, 212, 213, 224, 239, 240, 276, 280
Friesen, W. V., 239, 343, 345
Frith, S., 243
Frydén, L., 206
Fuks, L., 324, 326–330
Fuller, M. W., 289, 290
Furneaux, S., 139
Fyk, J., 184–188, 190

Gabbard, G. O., 52
Gabrielsson, A., 105, 199, 203, 204, 208, 220–222, 265
Galembo, A., 290, 292
Galvao, A., 110, 120
Gans, D. J., 321, 322
Garbuzov, N. A., 185
Gardner, H., 10, 25, 79, 109
Gärtner, J., 326, 329
Gát, J., 286, 289, 291, 297
Geller, B. D., 88

Gellrich, M., xi, 50, 100, 124, 129, 130, 177, 242, 298
Gembris, H., 19, 24
Geminiani, F., 219
George-Warren, H., 220
Gerig, R. R., 287, 289
Geringer, J. M., 185–187, 190
Gieseking, W., 167, 168, 175, 176, 178, 293
Gillespie, A., 246
Ginsborg, J., 168
Glenn, K. A., 106
Goebl, W., 295, 296
Good, J. M. M., 243
Good, T. L., 337
Goodlad, J. I., 342
Goolsby, T. W., 124, 139, 337, 340–342
Gordon, E. E., 5, 10, 18, 26, 79, 102, 104
Gould, G., 243, 291
Graffman, G., 84
Graydon, J., 50
Greated, C., 320
Grechesky, R. N., 345
Greco, V., 159
Green, G. A., 188
Green, L., x
Green, S. B., 222
Greene, D., 93
Greenwood, R., 343
Grey, J., 92
Gromko, J. E., 111
Gruhn, W., 73, 79
Gruson, L. M., 105, 154, 174
Guettler, K., 93, 308, 309, 311
Gunnarsson, R., 47

Hackett, G., 34
Hagegård, H., 222
Hale, J. L., 343, 344
Hall, D.E., 290, 292–294
Hallam, S., 27, 152, 154–158, 160, 168, 174, 176, 228, 231, 315
Hallett, M., 88
Hamann, D. L., 49, 50, 336
Hanton, S., 53, 55, 56, 57
Hardy, L., 51
Hargreaves, D. J., 23, 109
Harman, S., 84
Harold, R. D., 35
Harris, S., 255
Harris, T., 255
Hart, H. C., 289, 290
Harwood, D. L., 104
Harwood, D. W., 118
Hefling, S. E., 208
Heijink, H., 214
Heinlein, C. P., 289, 293
Helmholtz, H. F. von, 303, 322
Henderson, A. T., 204, 208, 210
Henderson, J., 131
Henderson, M. T., 295
Henderson, V. L., 38
Hendrix, J., 219
Henninger, J. C., 336
Hepper, P. G., 20, 25
Herman, R., 327
Hesse, A., 186
Hetherington, E. M., 19
Heuer, H., 124, 129
Heuser, F., 93
Hevner, K., 220
Hill, W. G., 291
Himle, D., 57
Hirano, M., 261
Hirsh, I. J., 290
Hochberg, F. H., 86, 88
Hoffer, C. F., 101
Hoffman, E. A., 258, 266
Hoke, M., 72
Hollo, J., 289–291
Holming, P., 285
Holt, M. E., 345
Honing, H., 214
Horan, J. J., 53
Horan, K. E., 189
Horowitz, V., 295–296
Houtsma, A. J. M., 289
Howard, D. M., 255, 260
Howard, J. A., 91
Howard, V. A., 227
Howe, G., 19
Howe, M. J. A., 4–6, 8, 10, 23, 25, 152, 159
Hudson, R., 221
Hughes, E., 167, 168
Hundt-Georgiadis, M., 75
Hunter, J. E., 52
Huron, D., 209, 287, 293
Huston, A. C., 25
Hutchins, C. M., 313

Imreh, G., 156, 157, 172, 174, 176
Irvine, J. K., 316
Iwarsson, J., 222
Izard, C. E., 343

Jackendoff, R., 201, 209, 210
Jackson, S., 58
Jacobs, J. P., 297
Jacobson, L., 27

James, I., 51
Jansson, E. V., 286, 288, 290, 292, 295
Jellison, J. A., 337
Jenkins, J. M., 221
Jenkins, W. M., 77
Johnson, C. M., 336
Johnson, M., 238
Johnson-Laird, P. N., 121
Johnston, J. C., 129
Johnston, R. B., 322
Johnstone, T., 222, 225
Jones, G., 53, 55, 56, 57
Jørgensen, H., 152, 160, 315
Jowdy, D., 56
Jung, C. G., 7
Jürgens, U., 225
Juslin, P. N., 210, 220–228, 230, 232
Just, M. A., 171

Kaiser, K. A., 339
Kaladjev, M. A., 93
Kalkbrenner, F. W. M., 287
Kane, M., 155
Kantorski, V. J., 187, 188
Karni, A., 75
Karrick, B., 186, 187, 189
Kase, S., 287
Kelly, N. E., 185
Kelly, S. N., 345
Kelly, W. J., 185
Kemp, A. E., 7, 8, 11, 13, 110, 120, 314
Kenardy, J., 48, 49
Kendall, M. J., 53, 158
Kennell, R. P., 340
Kenny, B., 127
Kenny, P., 323
Kernfeld, B., 122
Keys, N., 11
King, S., 344
Kingsbury, H., 227
Kintsch, W., 171, 173
Kjelland, J. M., 93
Klein, R. M., 106
Klingholz, F., 264
Klotman, R. H., 101
Kochevitschy, G., 287, 290
Kodály, Z., 109
Koehler, W. K., 93
Kohut, D. L., 102, 104, 158, 228, 331
Koornhof, G. W., 292
Koozer, R., 336
Kornicke, L. E., 143
Kotlyar, G. M., 222
Kozlowski, L. T., 240
Krampe, R. T., 152, 229, 238, 315
Kratus, J., 126
Kreisman, H., 328
Krumhansl, C. L., 206, 207
Kurtág, G., 287

Laban, R., 247, 345
Ladd, D. R., 276
Laib, J. R., 345
Lamp, C. J., 11
Land, M. F., 139
Landers, R., 102
Langer, S. K., 220
Larkin, K. T., 49
Larson, C. R., 276
Larsson, M. D., 93
Laukka, P., 224, 228, 230, 232
Lecanuet, J.-P., 20
Lederman, R. J., 86, 88, 89, 92
Lee, T. D., xi, 76
Lee, W. A., 93
Leggett, E., 38
Lehmann, A. C., 126, 137, 138, 140, 141, 142, 143, 152, 172, 228
Leimer, K., 167, 168, 175, 176, 178
Lerdahl, F., 201, 206, 207, 209, 210
Leschetizky, T., 287
Lethco, L., 344
Levin, T. C., 264, 279
LeVine, W. R., 93, 316
Levy, A. E., 93
Levy, C. E., 93
Lindsay, P., 171
Lindström, E., 224
Linklater, F., 158
Linville, S. E., 260
Liszt, F., 286
Livingston, H., 126
Lloyd-Elliot, M., 53
Loehlin, J. C., 19
Loosen, F., 187
Luce, J. R., 109
Lull, J., 25
Lusby, W. S., 289, 290
Lykken, D. T., 18

Macklem, P., 326, 330
Madison, G., 222, 224
Madsen, C. K., 185–188, 190, 336, 339, 340, 343–345
Maidlow, S., x
Mainwaring, J., 102–104, 106
Malbran, S., 127
Malloch, S. N., 21
Manchester, R., 85
Manturzewska, M., 12, 21

Marchand, D. J., 228
Marchant-Haycox, S. E., 48
Marcus, A., 175, 176
Marsden, C. D., 71
Marshall, A. H., 273
Martin, D. W., 293
Martinez, I., 127
Massie, D. L., 42
Matthay, T., 167, 168, 175, 176, 178, 287
Mattheson, J., 220
Matusov, E., 128
Mayers, H., 92
Mayne, R. G., 343, 344
McArthur, V. H., 147, 157
McClearn, G. E., 19, 20
McCormick, J., 34, 37, 41
McCoy, C. W., 341
McGue, M., 18
McIntyre, M. E., 304, 307
McLaughlin, D. B., 155
McNitt-Gray, J. L., 93
McPherson, G. E., 33, 34, 37, 41, 43, 100, 105–107, 109, 142, 154, 157, 158
Mead, J., 326
Meichenbaum, D., 53, 56
Melka, A., 307
Menuhin, Y., 220
Merzenich, M. M., 77
Meyer, J., 273, 313
Meyer, L. B., 118, 220
Middlestadt, S. E., 85, 93
Miklaszewski, K., 156, 157, 175
Millard, Y., 155
Miller, D., 263
Miller, G. A., 170
Miller, J., 245
Miller, S., 345
Mills, C., 336
Mills, J., 6, 11
Monson, I., 117, 118
Moore, D. G., 152, 159
Mor, S., 49
Morasky, R. L., 316
Morgenstern, S., 245
Mori, S., 208
Morozov, V. P., 222
Morris, R., 33
Morrison, S. J., 186, 187, 189, 193
Mumford, M. D., 153
Murphy, S., 56
Murphy, T., 50

Nachman, S. A., 335
Nagel, J. J., 52, 57
Nakamura, T., 239
Narmour, E., 118
Nederveen, C. J., 323, 331
Nelson, S. H., 92
Nettl, B., 118, 131
Neuhaus, H., 287
Newell, A., 171
Newmark, J., 92
Newton, I., 286
Nicholls, J., 34
Nielsen, S., 154
Nissen, M. J., 106
Nordmark, J., 280
Nordström, P.-E., 258
Norman, D., 171
North, A. J., 23
Noy, P., 21
Noyes, R., 48
Noyle, L., 175
Nuki, M., 142

Oatley, K., 221
O'Connor, R., 229
Oertle, E., 343
Oghushi, K., 290
Oliver, W., 173
Olsson, B., 23
Omelich, C. L., 42
O'Neill, S. A., x, 12, 33–35, 38, 159
Oostenveld, R., 72
Orff, C., 109
Orlick, T., 56
Ortmann, O., 285–287, 289, 291, 297
Ottani, V., 85
Oxendine, J. B., 152, 153

Paas, F., 106
Pace, R., 26
Paderewski, I. J., 151
Paganini, N., 219
Pajares, F., 34
Palmer, C., 199, 206, 207, 210, 220, 295
Pantev, C., 72, 73
Papousek, H., 21
Papousek, M., 21, 225, 238
Papsdorf, J., 57
Parasuraman, S., 335
Parfitt, G., 51
Parker, W., 56
Parlitz, D., 76, 79, 286
Parncutt, R., 100, 130, 177, 206, 209, 285, 295, 297, 298, 299
Pashler, H., 129
Pawlowski S., 325
Penel, A., 204, 210
Peretz, I., 69

Persson, R. S., 220, 221, 224, 225, 228
Peschel, T., 286
Peterson, G. E., 281
Philipson, L., 93
Piaget, J., 109
Pickering, N. C., 304
Pierce, A., 246
Pinker, S., 79
Pintrich, P. R., 31, 32, 34, 37
Pitteroff, R., 307
Pitts, S. E., 101, 246
Platt, J. R., 185, 186, 190
Plitnik, G. R., 320
Plomin, R., 17, 19, 20
Plutchik, R., 223
Porter, M., 322
Porter, S. Y., 12
Posner, M. I., 106, 156
Pranger, H. M., 156
Prassas, S., 93
Press, J., 93
Pressing, J., 117, 118, 123, 126
Preston-Dunlop, V., 345
Price, H. E., 336, 340, 341, 343–345
Prickett, C. A., 337
Priest, P., 109, 177
Proffitt, D. R., 240, 241

Rabinowitz, M., 230
Racine, R. J., 185, 186, 190
Radford, J., 25
Raekallio, M., 295, 297, 298, 299
Rakowski, A., 184, 185, 187
Raman, C. V., 304
Rapoport, E., 222
Rasch, R. A., 210, 295
Reimer, B., 221, 336
Reiss, D., 19
Renshaw, P., 233
Renwick, J. M., 41, 105, 154
Repp, B. H., 202, 204, 208, 209, 241, 293, 295, 296
Reynolds, C., 316
Rich, G., 119
Rijntjes, M., 75
Roberts, L. E., 72
Robinson, C. R., 336
Robson, B., 58
Rocaboy, F., 313
Rodà, A., 209
Roe, A., 13
Roland, D., 49, 53, 55–57
Roland, P., 75
Ronning, R. R., 104, 109, 111
Rosen, C., 290
Rosenbaum, D. A., xi
Rosenthal, R., 27
Ross, B., 72
Ross, S. L., 154
Rossing, T. D., xi, 289, 294, 321, 322, 325, 327, 328
Rouiller, E. M., 70
Rousseau, J. J., 109
Roussos, C., 325
Rovine, M., 19
Royal, M., 209
Rubin, J. S., 255
Rubin-Rabson, G., 154, 159
Rubinstein, A., 294
Russell, J. A., 221
Russell, M., 131
Rutter, M., 19, 20

Sachs, C., 319
Salmon, D. P., 75
Salmon, P., 53
Salzberg, C. L., 316
Salzberg, R. S., 316
Sampson, J., 185
Samuels, S. J., 106
Sang, R. C., 315, 342
Satt, B. J., 20
Savage, I., 51
Savart, F., 303
Sawyer, R. K., 128, 129, 131
Schaefer, J., 264
Schelleng, J. C., 310
Scherer, K. R., 222, 225
Schiefele, U., 32, 34, 40
Schlaug, G., 71, 72
Schleuter, S., 100, 102, 104
Schmidt, R. A., xi, 76, 266
Schnabel, A., 295, 299
Schnabel, K. U., 293
Schneider, W., 171
Schonberg, H., 168
Schraw, G. J., 104, 109, 111
Schultz, A., 286, 295
Schumacher, M., 220
Schumacher, R. T., 304, 307
Schunk, D. H., 31, 32, 34, 37, 40
Schuppert, M., 69
Scripp, L., 105, 109, 228
Scutt, S., 42
Seashore, C. E., 4, 5, 289, 329, 344
Secan, S., 329
Seckerson, E., 224
Seckl, J. R., 20
Seewann, M., 291
Segal, N., 18

Seligman, M., 56
Sergeant, D., 185
Sergent, J., 70, 138
Shaffer, D. R., 106
Shaffer, L. H., 238, 239
Sheldon, D. A., 338
Shellahamer, B., 343
Sherrill, M. H., 344
Shields, S. A., 223
Shifres, F., 127
Shimosako, H., 290
Shonle, J. I., 189
Shove, P., 204
Shuter-Dyson, R., 4
Sidoti, V. J., 345
Siegel, J. A., 185
Siegel, W., 185
Silverman, K. E. A., 276
Simon, H. A., 170, 171, 175
Simonton, K. D., 24
Sinclair, K., 107
Sinclair, K. E., 158
Singh, P. G., 290
Single, N. A., 337, 341
Skadsem, J. A., 341, 344
Sliwinski, P., 326, 330
Sloane, K. D., 22
Sloboda, J. A., 6, 8, 10, 22, 23, 25, 41, 104, 118, 124, 137, 139, 141, 142, 152, 159, 220, 221, 225, 227, 229, 295, 297–299
Smith, C., 128
Smith, F., 137
Smith, J., 328
Smyth, M. M., 106
Snell, E., 58
Sobaje, M., 50
Södersten, M., 260
Sogin, D. W., 188
Sorbye, M. D., 93
Sosniak, L. A., 6, 9, 12, 22, 152, 159
Sousa, G., 344, 345
Spaulding, C., 92
Spence, M. J., 20
Sperti, J., 106
Spiegel, M. F., 186
Spillman, R., 147
Standley, J. M., 339
Stanley, M., 19
Stanton, H. E., 54
Stanton, R., 58
Staszewski, J. J., 171
Steptoe, A., 48, 51–53
Sternberg, R. J., 128, 129
Stewart, D., 331
Stipek, D. J., 34, 36, 39, 42
Stockhausen, K., 295
Stoll, G., 291
Story, B. H., 258, 266
Stowell, R., 245
Straus, S., 85
Strömberg, N. O. T., 330
Strong, W. J., 320, 323
Sudnow, D., 120, 127
Suga, S., 208
Sullivan, Y., 155
Sundberg, J., 187, 204, 206–209, 212, 222, 239, 240, 254, 256, 259, 261, 262, 265, 276–278, 280, 281, 320, 326, 327, 330
Sundin, B., 100
Sundström, A., 208, 210
Suzuki, S., 26, 101–104, 176
Swanwick, K., 79
Sweeney, G. A., 53
Sweller, J., 106
Swets, J. A. 93, 94

Taguti, T., 208
Tait, M., 221, 224, 229, 343
Taylor, C., 56
Taylor, J., 56
Tellegen, A., 18
Terhardt, E., 278, 291, 293
Ternström, S., 273, 274, 276–278, 280
Tesch-Römer, C., 152, 229, 238, 315
Thaut, M., 93
Theim, B., 93
Thomas, D. R., 340
Thomassen, M. T., 209
Thomasson, M., 259
Thompson, E. G., 157
Thurman, L., 258, 265
Tieplov, B. M., 186
Tipton, D. G., 340
Titze, I. R., 254, 255, 258, 260, 261, 266, 275
Todd, N. P. McA., 204, 239
Tommis, Y., 103
Tornapolsky, A., 325, 326
Townsend, E., 138
Trehub, S. E., 5
Trevarthen, C., 20, 238
Troup, M., 297
Truslit, A., 241
Tyron, A., 19

Underwood, G., 137, 138

Valentine, E. R., 54, 55, 152, 174, 315
van der Walt, A. J., 292

Van Heesch, N. C., 48, 49
Van Kemanade, J. F., 48, 49
van Meeteren, A. A., 278
van Merrienboer, J., 106
Van Son, M. J., 48, 49
Vaughn, K. V., 189
Verge, M. P., 325
Vernon, L. N., 210, 296
Verschuure, J., 278
Vidolin, A., 209
Vispoel, W. P., 36, 37, 40
Vivona, P. M., 326
von Békésy, G., 273
Vos, J., 184
Vygotsky, L. S., 339

Wagenaars, W. M., 289
Wagner, C., 12, 297
Walter, B., 338
Walters, J., 10
Wapnick, J., 185
Ward, W. D., 185
Wardle, A., 52
Waters, A. J., 137–139, 141, 146
Watkins, J. G., 142
Watkins, R. E., 341, 342
Watson, C. S., 185, 186
Watson, P., 54
Wead, C. K., 294
Weait, C., 329
Weast, R. D., 331
Weaver, H., 139
Webster, P., xi
Weeks, P., 342
Weinberg, R. S., 153
Weiner, B., 36
Weinreich, G., 294, 313
Weisberg, R., 126
Welch, G. F., 258, 260, 265, 266
Welford, A. T., 124
Welsh, P., 105, 109
Wesner, R. B., 48, 49, 52
Westendorf, L., 118
Wicinski, A. A., 156
Wienreich, G., 323
Wigfield, A., 32, 33, 34, 40
Williamon, A., 152, 169, 174, 315
Williams, C. E., 311
Williams, D., xi
Wills, G. D., 51
Wilson, G. D., 48–50
Windsor, W. L., 204, 206
Winter, S., 343–345
Witt, A. C., 185–187
Wolf, T., 142, 145
Wolfe, D. E., 159
Wolff, K., 295, 299
Wolkove, N., 328
Woodhouse, J., 304, 307
Woods, D., 246
Woody, R. H., 220, 221, 224, 225, 227, 228, 229
Worman, W., 321, 324, 326
Worthy, M. D., 185
Wright, J. C., 25
Wroton, M. W., 185

Yarbrough, C., 186–189, 336–339, 341, 343–345
Yerkes, R. M., 50, 51
Yoon, K. S., 34
Yorke Trotter, T. H., 101
Yoshikawa, S., 323

Zaza, C., 85
Zdzinski, S. F., 22
Zeitlin, D., 119
Zimmerman, B. J., 40
Zoltowski, M., 325
Zurcher, W., 15

# Subject Index


Ability
  to achieve, 34, 37
  to adjust to instrument, 293
  to produce wind tone, 328
  to recover, 315
Ability/capacity/skill
  aural/auditory, 5, 7, 14, 18, 65, 183, 190, 213, 314, 316
  communicative, 169
  conducting, 338–40, 344–5
  ear-playing, 103, 107–9
  expressive, 202, 212, 221–2, 225–30, 233, 246, 253
  imagery, 56
  improvising, 118, 128–9, 177–8
  interpretive, 299
  intonational, 189–90, 194
  kinesthetic, 110, 140
  learning, 161
  memory, 170–2
  mental/cognitive, 17–18, 67, 121, 137, 155 (*see also* intelligence quotient)
  motor, 63–5, 69, 74–6, 79–80, 88, 117, 152–3, 156
  musical, 3–5, 17–8, 31–43, 92, 102, 143, 278
  perceptual/sensory, 138, 328
  performance, 84, 112, 148
  piano, 287, 291–6, 299
  practicing, 92, 154, 157–9, 173–4
  problem solving, 41, 128, 141
  sensorimotor, 74, 77
  sight-reading, 102, 108–9, 136–7, 141–4, 147, 213
  string, 304–5, 309
  technical (technique), 56, 90–2, 94, 100–2, 121, 130, 155, 158, 176, 191, 244, 299
  tests of, 5
  vocal, 275
  *See also* audiation; mastery; pitch discrimination; task
Accent, 199, 208–10, 294–6
Achievement
  academic, 32
  and motivation, 31–42
  musical, 4, 10–1, 22–3
  and practice, 151–61
  and sight-reading, 106, 111, 120, 143–8
Acoustics, 253–332
  room, 122, 269–76, 292–4
Adam's apple, 255
Adrenaline, 47
Affect, 222, 225
Affections, doctrine of, 220
Air column, 319–25
Air pressure, 255, 276, 319, 323
  ambient, 321, 325
Alexander technique (AT), 54, 57, 92
Amplitude, 259–63, 289–91, 303–13, 323
  *See also* pitch-amplitude effect
Analysis, 154
  factor, 107
  frequency/Fourier, 319
  melodic, 69
  movement, 240, 247, 345
  music/score, 53, 126, 154–7, 167, 174–8, 214
  psycho-, 52
  statistical, 230–1, 271
  time-series, 24

Anticipation, 124, 167
Antiinflammatories, 83, 86, 87
Anxiety, 23, 35
  performance, 8, 13, 40, 47–58, 144, 168
  state versus trait, 13–14, 50
Applications. *See* strategies
Arm weight, piano, 287
Arousal, 65
  optimal, 50–1
Arpeggio, 100, 171, 265–6, 298–9, 326, 331
Articulation, 100, 135, 147, 193, 199–201, 209–14, 223, 227, 230, 262–5, 269, 276, 314, 322, 330, 338–9
Arthritis, 90
Arytenoid, 255, 260
Association
  between subskills of performance, 142, 145–7
  brain mechanisms of, 67, 73
  and emotional communication, 228
  melodic, 175
  and memory, 170, 175
Asynchrony, piano, 294–6
Attack (of a tone), 210, 223–6
  piano, 291
  violin, 303–12
  wind, 319, 325
Attainment. *See* achievement
Attainment value, 32
Attention
  and anxiety, 53
  and the body, 238, 246–7
  and the brain, 76
  and conducting, 335–8
  division of, 129
  and improvisation, 129–30
  and intonation, 187, 190, 193–4
  and medicine, 92
  and memory, 170
  and notation, 106, 111–2, 138–41, 146–7
  and pitch salience, 291
  and practice, 154
Attributes/qualities
  emotional/expressive/dynamic, 184, 220, 228, 238, 247
  personal, 4, 7, 23, 335, 343
Attribution theory, 36–40, 43
Attention, auditory
Audiation, 79, 103, 156
Audition. *See* perception, auditory
Auditions, 49–51, 55, 136–7
Awareness/consciousness
  and attention, 170
  aural, 109–12
  of body, 54, 57, 92, 246
  historical/cultural, 191
  and intonation, 192
  musical, 104, 190
  in singing, 260, 265–6
  social, 315
  state of, 125
  *See also* self
Automatic pilot, 175
Automation
  of breathing, 330
  in improvisation, 120–4, 130
  motor/kinesthetic, 75–9, 167
  in playing from memory, 175
  in practice, 228
  in sight-reading, 106, 112, 138, 144, 146
  in wind embouchure adjustment, 327
  *See also* leap, piano
Autonomy. *See* independence

Balance. *See* loudness
Bass, double, 307, 311–3
Bassoon, 11, 324–30
Bebop, 122
Beginner. *See* novice
Beliefs
  entity, 39
  expectancy, 34
  incremental, 39
  about musical ability, 17, 39
  about sight-reading, 140
  *See also* self
Beta-adrenergic blockers, 51
Biofeedback. *See* feedback
Biography, 4–6, 13, 21–5, 73, 142, 220
Biomechanics, 12, 91–4, 244, 276
Body, human
  and conducting, 343–5
  and emotion, 221, 225
  fetal experience of, 20
  and medicine, 90
  movement, 147, 237–48
  and piano, 289–90, 309, 319
  and singing, 255, 273–6
  and string playing, 309
  and wind playing, 319, 329–30
  posture/position, 69–70, 329–30
  use and abuse, 90–4
  *See also* kinesthesia
Body, of instrument, 304, 313
Botox (Botulinum toxin), 88
Brain, 63–80
  area/cortex/lobe/region, 63–77, 88
  auditory cortex, 280

cell (*see* neuron)
and emotion, 225
hind-/mid-/fore-, 65
imaging, 64
mechanisms, 63–80
organization/structure, 72–3, 79
programs, innate, 225
and sight-reading, 137, 146
stem, 65
*See also* auditory; cerebellum; frontal; motor; occipital; parietal; temporal; Wernicke
Breath support, 185, 275. *See also* pressure
Breathlessness, 47

Cantabile, piano, 286, 291–4
Capacity
lung/breathing/vital, 258, 326, 330
physical, 17
*See also* ability
Carbon dioxide, 328
Carpal tunnel, 87
Catastrophe (anxiety), 51, 56
Catgut Acoustical Society, 304
Cello, 10, 169, 313
Center of moment, 240–1
Cerebellum, 65, 70–6
Choir/choral, 13, 18, 136, 186, 192–3, 261, 264, 269–82, 341–2
Chord
choral, 279–82
intonation, 184, 188, 281–2
in jazz improvisation, 118–30
piano, 92, 286, 289–91, 294–5
progression, 130, 174
-scale method, 124–7
in sight-reading, 138–40
Chorus effect, 279–80
Chunking, 139–40, 170–1
Clarinet, 91, 142, 158, 320–30
Code
cultural, 242
emotional, 222–9
Cognition
and anxiety , 51–6
and the brain, 65–71
and conducting, 314, 343
and expression, 210, 221, 230–3
and improvisation, 117–24, 132
and intonation, 190
and memory, 170–5
meta-, 154, 160
and motivation, 41–2
and practice, 155–7
and sight-reading, 106, 112, 138, 144–7, 314
*See also* constraint; learning
Cognitive
engagement, 41–2
load, 121
process, 65, 69, 102, 121–4, 144, 155, 160, 171, 230, 314, 342
rehearsal (*see* practice, mental)
restructuring, 53–5
*See also* ability; development; feedback; skills; strategies; therapy
Coloratura, 259
Combination tones, 280
Communication, 205, 297
accuracy of, 222
emotional, 219–33
in improvisation, 117, 120–1, 128, 131
in memorized performance, 169
mother-infant (*see* motherese)
structural, 199–215
verbal/nonverbal in conducting, 341–7
*See also* expression
Competence, 4, 21–3, 32–43, 105, 110–3, 153, 173, 335–6
Composition, 100, 107
Computer
-assisted teaching, 76, 147, 215, 231–3, 328
technology, 6, 24
Computer simulation
of bowed string, 304, 308, 311
of expression, 199–200, 213–4
of knowledge base, 121
Concentration, 35, 41, 48–50, 57–8, 124, 129, 148, 153–6, 175
Condition/context/setting
acoustic, 293
educational/learning/teaching, 6, 23, 27, 31, 36, 42, 126–8, 177, 227, 233, 342
environmental, 18, 27, 63
experimental, 121, 184
medical, 77, 85–6, 88–9, 93
musical, 17, 142, 146, 174, 294, 297, 344
performance, 50, 53, 93, 119, 130, 136, 145, 157, 184–7, 193–4, 230, 274, 292
practice/rehearsal, 58, 129, 160, 189–91, 337–41, 346
sociocultural, 9, 18, 220, 240–4
Conditioning, 52, 80
Conducting, 335–47
demeanor, 338–9, 343
emblems, 345–6

Conducting (*continued*)
feedback, 336–41, 345–6
gestures/movements, 135, 241, 247, 335, 339, 341–6
modeling, 340–2, 346–7
pacing, 337–8
verbalizations, 335–47
Consciousness. *See* awareness
Consonance/dissonance, 124, 130, 207–8
*See also* roughness
Constraint
acoustical, 271
cognitive, 106, 123, 201, 297
on fingering, 297–9
on improvisation, 117–23, 127
*See also* therapy
Consumption of music, 25, 84
Contour
expressive, 224–5
melodic/pitch, 20, 69, 118, 206, 225
Convention
intonational, 185
notational, 201
performance, 201, 208, 230
Coordination
between performers (*see* ensemble)
brain mechanisms of, 65, 70–2
ear-to-hand, 109–10
expressive, 144
motor, 102
piano, 297–9
string, 18, 314
wind, 329
vocal/singing, 258–66, 275
Corpus callosum, 65, 70–1
Cortex. *See* brain
Cortisone, 87–9
Cramp, occupational/pianist's, violinist's. *See* focal dystonia
Creativity
child's, 13, 26
jazz/improvisational, 119–20, 128–32
Cricoid cartilage, 255
Cricothyroid, 261, 275. *See also* lengthener
Cubital tunnel, 87
Cues, emotional/expressive, 222, 226–7, 230

Deadpan performance, 200–1, 239
Decay, tone, 223, 289, 293–4
Decoding
emotional, 224–6
neural, 67
notational, 105–6, 111–2, 138, 146
Demonstration
conductor, 341–2, 347
teacher, 158, 315
Depression, 89
Desensitization, systematic, 52
Development
aural, 105, 155–8
artistic/musical/psychological, 7–27, 212, 221, 227–9, 238
brain, 76–9
cognitive, 25
and improvisation, 125–8, 178
and intonation, 190–2
and motivation, 37–42
musical, 7–9, 13, 18–27, 40, 102, 112, 158–61, 190–2
and practice, 155–61
and sight-reading, 100–13, 137, 148
*See also* zone of proximal development
Diagnosis, 83–90, 339–41
contextual, 90
mis-, 89
Diaphragm, 258, 325–6, 331
Discrimination, aural/auditory, 18
Dissonance. *See* consonance
Double time, 202
Drugs, 51–3
Drums. *See* percussion
Duration
contrast, 208–9, 212
double, 208
of practice/rehearsal, 152–3, 337
of tone, 138, 185–7, 200–2, 209, 212–4, 266, 292–6, 307
Dynamics
and conducting, 342–4
expressive, 99–214, 239
group, 132
and intonation, 193
and memorization, 174, 178
and piano, 286–94
and practice, 154
and sight-reading, 147
and singing, 273
and strings, 311
and winds, 323–8

Ear/aural training/development, 72, 80, 105, 177
Ear playing, 41, 79, 99–113, 120–8, 159
Echo. *See* reverberation
Edge tone, 325
Efficacy. *See* self

Effort
child's/student's, 32–43
parent's, 22
in practice, 156
in sight-reading, 143–4
in singing, 255, 262–3, 275
teacher's, 9–10
in wind playing, 325, 331
Electroencephalography (EEG), 64–5
Electromyography (EMG), 87, 90, 93
Emblems, 242–4, 345–6
Embouchure, 77, 88, 91, 155, 185, 192, 319–31
Emotion, 219–33
and articulation, 209
basic/innate/universal, 223–5
and body movement, 238, 241–6, 264–5
brain mechanisms of, 63–71
and development, 21–6
and facial expression, 343
and injury, 91–2
and motivation, 34–43
and musical structure, 212–4
and personality, 315
and voice, 264–5
Encoding
in expression, 226
in fingering, 297
in memorization, 172–5
in sight-reading, 138–9, 144
Engagement/involvement, musical, 6–7, 9, 14, 22, 26–7, 33, 39–43, 63, 146, 159–61, 244
Enjoyment, 32, 41, 109–11, 149
Ensemble
coordination, 148, 242–4
participation, 32, 94, 188
practice, 153
sight-reading, 136, 146–8
timing, 210, 214, 294
tuning, 184, 188–94, 280
*See also* choir; conducting
Ensemble effect. *See* chorus effect
Environment
and development, 6–10, 14, 17–27
human, 63, 73, 137, 339
musical performance, 36, 50, 129–30, 157, 167, 190–3, 247, 336
shared and nonshared, 19
social and cultural, 43, 191
teaching and learning, 39, 42–3, 107–9, 126, 229, 336
Equal temperament. *See* intonation
Error/mistake
and anxiety, 51
correction, 41, 105, 140, 156, 160, 337
detection, 338
in improvisation, 120, 127, 130
intonation, 185–6, 259
in practice, 48, 105, 154–5
proofreader's, 141–4
sight-reading, 140–4, 148
*See also* trial and error
Examination
medical, 88–90
performance, 8, 13, 34, 37, 42, 154
Exercises
expressive, 214, 233, 246–7
physical, 58, 88–9, 92
sight-reading, 148
technical, 41, 100, 130, 259, 265–6, 294, 298, 323, 328, 331
Expectancy/expectation
music-structural, 67, 118, 138–141, 144–5, 206
teaching/learning, 9, 32–40, 48, 53
-value theory, 32, 43
*See also* belief
Experience
extramusical, 225
musical, 4–6, 10, 18–9, 23–4, 36, 176, 184
peak, 119
proto-musical, 238
Expertise
improvisational, 124–6
medical, 84–6, 90–1
memory, 170–3
performance, 5, 73, 126, 147, 152–7, 160, 170, 194, 227, 238, 295, 299, 315, 337, 341, 346
sight-reading, 139–44, 147–8
Expression
and ability, 4
emotional, 219–33
facial, in conducting, 339, 343–6
and improvisation, 118–20, 129–31
and intonation, 184, 187, 191
and memory, 169
myths about, 227
piano, 291, 295
physical (body), 237–48
and practice, 153–4, 168
and sight-reading, 100, 111–2, 135, 144–7
structural, 199–215
and teaching, 101, 247
winds, 319, 331
vocal/singing, 225, 266, 329
*See also* emotion; motherese; verbal

Extroversion, 7–13, 19, 50
Eye
  contact, 335–9, 343–6
  coordination/dexterity, 17, 109
  movement, 17, 65, 138–40, 148
  *See also* motion; saccades; span

Failure, 8, 31, 36–42, 56
Falsetto, 188, 260
Fatigue/tiredness, 51, 91–2, 153
Feedback
  aural/auditory/sensory, 63, 69, 140, 190, 265, 273, 276, 315, 328
  bio-, 93, 316
  cognitive, 230–3
  in conducting, 335–6, 338, 345–6
  definition of, 229
  in expression, 214, 229–33
  and flow, 35, 58
  in improvisation, 120–4
  in intonation, 190–4
  kinesthetic, 69–70, 290–2, 296
  sensory, 69, 76–7, 292
  verbal/teacher/instructional, 42, 156, 190, 266, 316
  visual, 41, 140, 190, 288, 331, 336–41, 345
Fingering
  instrumental, 41, 77, 88, 100, 104–11, 120, 130, 137
  piano, 139, 142, 148, 167, 295–9
  *See also* constraint; encoding
Fingers
  brain representation of, 72, 75–7
  changing, 288
  curved versus straight, 286
  length of, 11
  tension in, 91
  versus arms, 286–7
  *See also* focal distonia
Fixation, 137–9
Flageolet (string harmonics), 309, 312
Flow theory and states, 35–6, 43, 58, 119–20, 124–5, 148
Flow-controlled instrument, 325
Flute, 3, 12, 76, 192, 264, 277–8, 320–9, 330
Focal dystonia, 77, 85, 88
Formal. *See* practice
Formant, 254, 261–6, 279–82
Fovea, 137
Frequency (acoustic)
  fundamental, 255, 259–63, 270–1, 275, 307–10, 326–9
  ratio (of interval), 184
  *See also* formant; pitch; vibrato
Frequency (incidence)
  of ensemble performance episodes, 337
  of practice, 34–5, 143
  *See also* incidence
Friction (string), 306, 312–3, 320
Frontal lobe, 67, 70, 76
Fry, vocal, 260
Functional residual capacity (FRC), 258, 259

Ganglia, basal, 65, 70–6
Generative
  mechanisms, 122, 243
  models, 121, 127
  structure, 172
Gestalt. *See* perception
Gesture
  conducting, 335, 339–46
  musical, 178, 202, 275
  physical, 110, 194, 239–47
Glottis, 255, 259–60. *See also* transglottal
Goals
  learning, 31–5, 38–43, 71, 79, 148, 153–61, 175, 194, 266, 324, 337–40
  setting, 54–8
  teaching, 99, 111, 147, 227–31
  *See also* strategy
Groove, 120, 208
Guitar, 130, 214, 230
Gyrus
  cingulate, 71
  precentral, 70

Hammer (piano), 288–95
  velocity, 289
  -shank vibrations, 290–2
  -string noise, 291–4
Hand
  profile, biomechanical, 12
  -span, 12, 18
Harmonic charge, 206–7
Harmonics. *See* flageolet; partials
Helmholtz motion. *See* string motion
Hierarchy
  cognitive, 121
  error, 338
  fear/anxiety, 52, 57
  metrical, 210
  phrase, 205
Hiss. *See* noise
Homunculus, motor/sensory, 70
Horn, French, 11

Hyperadduction. *See* pressed phonation
Hypnotherapy, hypnosis. *See* therapy

Imagery
  auditory/aural, 76, 79
  and emotion, 227
  as therapy, 54–7
  and voice teaching, 266
  and wind teaching, 331
Implications. *See* strategies
Improvisation, 117–132
  by children/students, 7, 12, 41, 73, 100–2, 107–8, 159, 177–9
  of fingering, 298
  group, 126, 128
  hierarchical, 122
  historic, 168
  referents, 118–23, 127
  structure, 118–22, 126–8
  *See also* ability; attention; automation; chords; cognition; communication; constraint; development; error; expertise; expression; feedback; transcendence
Incidence/frequency
  of medical problems, 84–9
  of performance anxiety, 48–9
Independence (autonomy)
  of ensemble, 339–40
  motor, 146, 297–9
  of student or musician, 26, 38, 151, 159–60
Individual differences, 8, 12, 18, 25, 138, 142–3, 155, 173, 222–4, 241, 297
Infant-directed speech. *See* motherese
Injury, 83–94
  chronic overuse, 77
  repetitive strain, 77
Innate. *See* potential
Instrument, choice of, 10–1
Intelligence quotient (IQ), 19
Intention, expressive/interpretive/musical, 204, 209, 222–6, 229–30, 233, 239–44, 247, 295–7
Interonset interval (IOI), 200–2, 207–10, 214
Interpretation, 56, 100, 112, 118, 154–60, 202, 205, 213, 229, 239–47, 290–3, 298, 338–44. *See also* constraint; expression
Intonation/tuning, 158, 183–94, 271, 274–82, 290, 313–6, 319–21, 327–8, 341
  equal temperament, 184, 187, 280
  just, 184, 280
  Pythagorean, 184, 187
  *See also* ensemble, formant
Introversion, 7, 9, 12, 13, 50
Intuition, 8, 105, 109, 120, 144–8, 156–7, 199, 204, 224, 288–95, 304, 328
Involvement. *See* engagement

Jazz standard, 118. *See also* bebop; chord; creativity; improvisation; repertoire
Judgment. *See* rating
Just noticeable difference (JND) of frequency, 184

Keystroke, piano, 286–93
Kinesthesia, 54–6, 69–70, 119–20, 130, 138–40, 143, 148, 167–8, 173–5, 264–6, 288–92, 342, 346
Knowledge
  base, improviser's, 118–21, 125–30, 135
  base, and memory, 170
  child's/student's, 104–6, 110–2, 153, 160, 194
  conductor's, 338, 346
  expressive, 224, 228–30
  about learning and practice (*see* metacognition)
  musical, 101, 104, 177–9
  procedural/declarative, 298–9
  sight-reading, 136, 140, 144, 148
  structural/theoretic, 168, 171, 173–4, 342–3
  teacher's, 99

Language
  acquisition, 21, 104, 130, 264
  body, 225
  brain mechanisms of, 67
  musical, 26, 90, 100, 103, 127
  written/read, 104, 109, 141
Laryngeal
  cartilage, 266
  muscle, 255, 259–60
  ventricle, 258
Larynx, 254–5, 262–4, 276, 329
  epi-, 258
Leap
  melodic, 209
  piano, 286–8
Learning
  mis-, 10, 77
  (sensori-) motor, 74–6
  natural, 102
  over-, 135

Learning (*continued*)
re-, 99
rote, 100–1, 113, 176
social, 225
style, 155, 176, 179, 356
theory, 102
*See also* ability; environment; metacognition; strategy
Legato
in expression, 209–10, 223, 230, 238, 246
piano, 289–93
vocal, 259
Lens model, 225–7, 230
Let-off distance, piano, 288
Life
-span, 63, 253, 260, 267
-style, 57–8, 92
Listening
brain mechanisms of, 67–9, 73
experiment/test, 76, 204, 224, 233
to music, 6, 24, 63, 113, 148, 157, 311
to performance model, 102, 105–6, 110, 145, 316, 341
to self, 26, 144, 177, 290, 296
Literacy, 102–5, 109–13, 135–6. *See also* sight-reading
Loudness, 270
balance of, 271, 274, 278, 285, 341
brain mechanisms of, 67
definition of, 270
envelope, 210
and expression, 210
and piano, 288–94
vocal, 255, 259, 263, 271–80
and winds, 326–9
*See also* sound pressure level

Magnetic resonance imaging (MRI), 64–5, 75, 90, 266
Magnetoencephalography (MEG), 64–5, 72, 77
Masking, psychoacoustic, 273–6, 291–2, 296
Mastery
motivation, 39, 43
orientation, 38
task, 38, 50
Maturation, 8, 12–4, 17–8, 25, 159, 299
Meaning
expressive, 212, 220, 238
gestural/motional, 242–4, 247, 345
historical, 214
linguistic, 104, 185
and memory, 169–71, 177
musical, 26, 69, 79, 105, 111–2, 118, 121–3, 126, 173, 238, 242, 245
perceptual/cognitive, 137–40
Media, mass, 24, 27
Medication, 86–9. *See also* drugs
Medicine, 83–94
Medulla, 65
Melodic
analysis, 69
charge, 207
paraphrase, 122
Melody
improvisation of, 121
lead/delay, 207–10, 294–6
performance of, 187–9, 193, 246, 292–4
teaching of, 101, 105, 111
Memory/memorization/recall, 167–79
aural/auditory, 157, 166–7, 175–7, 288
brain mechanisms of, 65–70
expert, 170–4
of fingering, 298
in improvisation, 117–8, 121–2, 129
kinesthetic, 167–8, 173–5, 178
long-/short-term, 121, 124, 138, 144, 170–1
of notated music, 4, 50, 100–2, 107, 112, 157, 238, 298
psychology of, 169–71
in sight-reading, 138–47
of sound model (rote), 110–1, 113, 176, 265
tactile, 167, 175, 288
theory of, 170–1
visual, 167–8, 173–5, 177–8, 288
*See also* chunk
Mental
practice. *See* practice
process. *See* cognitive
training. *See* practice
Metacognition, 154, 160
Metaphor, 117, 204, 227–9, 238, 241, 247, 339
Metrical hierarchy, 201
Metronome, 147–8, 157, 169, 344
Microstructure, 220, 224
Mnemonics, 172
Mode, Dorian, 127
Modeling, 40, 106, 158, 189, 219, 221, 228, 315–6, 340–2, 347
Modulation
amplitude/frequency, 261
key, 130, 178, 275
Momentum, musical, 204

Monitoring
aural, 314
by brain, 64
electronic, 316
kinesthetic, 69
in improvisation, 129
of internal state, 144
of intonation, 187
of learning, 154
of own singing, 264, 273–5
by teacher, 10, 22, 40, 93
visual, 140, 168
Mood, 220, 242, 265
Mother tongue, 102, 131
Motherese, 21, 237
Motion/movement
body, 63–71, 74–80, 121, 144, 147, 225, 237–47
conducting, 135, 339, 344–6
eye, 65, 137–40, 148
implied/abstract, 76, 204–5, 208–9, 212, 220, 239, 241
melodic, 209
musical, 237–48
piano playing, 286–90, 295–9
physical law of, 286
singing, 258, 276
string playing, 309, 314
therapy, 88–90
wind playing, 329–30
*See also* analysis; respiratory; schema; therapy
Motional character. *See* motion, implied
Motivation, 4–9, 11, 14, 19, 31–43, 126
adaptive, 37–40
and flow, 35–6, 119–20, 124–6
and health, 92
intrinsic/extrinsic, 10, 14, 32–5, 39–41, 128
and practice, 43, 153, 158–61
role of teacher, 23, 27
*See also* cognition; strategy
Motivic (improvisation), 122–3, 126
Motor
auditory-, 69
command, 72
constraint/limit, 202, 296–7
control, 74, 77, 292, 314
experience, 109
function/process/performance/action/behavior, 65, 70–1, 112, 147, 266
independence, 286
learning, 75–6, 156
movement/component/aspect, 121, 129, 144, 204
pattern/sequence, 76, 167
program/system, 69–71, 75, 118, 144, 212, 238–9
reflex, 79
region/area/cortex, 65–7, 70–7, 88
skill/task, 63–5, 69, 75, 79–80, 117, 152–6
*See also* ability; automation; brain; coordination; fingering; homunculus; independence; learning; sensory; skill
Motor area/cortex/region, 65–76, 88, 175, 212, 296
cingulate, 70–1
pre-, 70–5
primary/secondary/tertiary, 67–76
supplementary, 70–2, 75
*See also* plasticity
Mouth
cavity, 262, 322
-piece, 311, 69, 323–4
Movement. *See* motion
Muscle
atrophy/weakness, 86–7
brain representation of, 70
coordination, 102, 295
cricoarytenoid/interarytenoid, 255
lengthener/shortener, 255–61
respiratory, 319, 325–6, 331
*See also* diaphragm; embouchure; injury; laryngeal; proprioception; tendon; tension
Musicality, 27, 155, 159
communicative, 21
*See also* ability; potential
Musicianship, 14, 64, 72, 75–6, 100, 103, 109, 113, 344. *See also* ability; potential
Myelination, 72

Nasal cavity, 254, 258, 263–4
Nature. *See* potential
Nerve
auditory, 72–3
cell (*see* neuron)
entrapment, 85–90
median, 87
ulnar, 87
*See also* velocity
Nervous system, central, 47, 63, 67, 72, 76
Neuro-. *See* brain mechanisms
Neuron, 64–73, 79
Neuronal networks, 64, 69, 73, 80
Neuroticism, 19

Noise
masking, 276
in piano, 288–94
in strings, 306–11
Nominal performance, 200–1, 208–10, 214
Nose. *See* nasal
Notation. *See* sight-reading
Novice/beginner, 9–12, 37, 76, 100, 104–6, 110–2, 126, 139, 140–2, 147, 152–4, 159, 174, 192, 269, 287, 299, 337–8, 341, 342–4. *See also* expertise
Nurture. *See* environment

Oboe, 321–30
Occipital lobes, 67–9, 73
Octave
piano, 92, 286–7
tuning/stretching, 184–8
violin, 90, 311
Offset, 200
Onset, 67, 200, 206, 210, 214, 219, 291, 296, 306–11, 325. *See also* interonset interval; noise
Orchestra, 10, 48, 54–6, 85, 92, 100, 136–7, 174, 184, 238, 262, 287, 311, 314–5
Oscillation/vibration
chaotic, 319
floor, 313
hammer-shank, 290–2
lip, 323
reed, 320–1, 322–4, 329
string, 290, 304–7
violin body, 303
vocal fold, 254–5, 259–60, 275–6, 329
*See also* vibrato
Overblowing, 321, 325–8
Overtone singing. *See* throat singing
Overtones. *See* flageolet; partials

Pacing
to avoid injury, 92
conducting, 337–8
Pain
and depression, 89
disorders, 86
and focal dystonia, 77
jaw, 91
musculoskeletal, 85–6, 90
and stress, 92
and thoracic outlet syndrome, 87
Parafovea, 137
Parent, 4–14, 19–23, 26–7, 40–3, 52, 56, 79, 93, 102, 159–61, 237
Parietal lobe, 67, 70, 73
Parkinson's disease, 74, 88
Partials/harmonics/overtones, 254, 259, 261, 264, 271, 278–82, 289–91, 303–13, 324
Pathology, 83–90
Pattern recognition. *See* perception, gestalt
Pedal
sustaining, 288–9, 292–3
*una corda*, 292–4
Peers, 18–20, 24–5, 33–43, 102, 126, 242, 315, 339
Perceived cost, 32
Perception
aural/auditory, 49, 67–9, 72–9, 93, 138–41, 199, 270–1, 275–80, 290, 296, 328, 346
brain mechanisms of, 67–9, 76
of body movement, 241–3
categorical/zonal, 185, 212, 221
gestalt/pattern, 67, 104, 137, 141, 144, 147
of motion (*see* kinesthesia)
of performance, 56, 200–2, 238, 244, 314, 330
of piano, 290–6
of pressure, 328
of singing, 262–6, 272, 275–6
tactile/touch, 110, 140, 148, 167, 175, 286, 288
units of, 105
visual, 105–6, 67, 137–9, 226, 290
*See also* loudness; perception; pitch; self-; timbre
Perceptual/sensory feedback, memory, schema, span. *See* feedback; memory; schema; span
Percussion/drums, 208–10, 342, 347
Percussive
onset noise, 291
touch (piano), 289–92, 295
Perfectionism, 48–9, 56
Persistence, 31, 34, 38–40
Personality
and anxiety, 48–50
development of, 19–20
and instrument, 4, 11, 314–5
and musicality, 4–8, 14
performer's, 212, 245–7
and practice, 315
psychology of, 18
of teacher, 23
*See also* extroversion; introversion; temperament
Pharynx, 254, 258, 262–4
Phobia, 47–9, 52
Phonation, pressed, 255, 260–3, 281

Phrase
and expression, 201–7, 212–3, 222, 239, 247
hierarchy/structure, 201–5, 207
in improvisation, 124, 130
and memory, 172
para-, 122
piano, 297
in sight-reading, 104–5, 111–2, 139, 144
spoken, 260
sung, 259, 265
wind, 328–30
Phrasing, 201–7, 212–3
Physics. *See* acoustics
Physiology, 253–67, 286–8, 319–32
Piano, 3–4, 12, 53, 71–2, 75–6, 103, 119–20, 129–30, 137–44, 159, 167, 172–8, 209–10, 214, 239, 285–99
Pipe tone, 325
Pitch
absolute, 72, 185
accuracy, 90, 187–94 (*see also* intonation)
-amplitude effect, 271, 275–8
articulatory perturbation of, 276
brain mechanisms, 67–9, 72, 76
definition of, 270
discrimination, 5–6, 158, 183–6, 192
intrinsic, 274–6, 281
matching, 158, 183–93
perception, 76, 184–5, 192, 269 (*see also* frequency)
salience/clarity, 278, 291–3, 306, 310
in sight-reading, 103, 138, 141, 147, 154
of strings, 304–13
structure, 69
and timbre, 290
of voice, 254–66, 271, 275
of winds, 323–29
*See also* contour; frequency; fundamental; register; vibrato
Pizzicato/arco, 306, 313
Placement
finger/pitch (intonation), 185–94
singer/choir (acoustics), 273–4
Plasticity
neural, 63, 72–7
vocal, 253
Play, children's, 128
Plethysmography
body, 329
respiratory inductive (RIP), 330
Pons, 65
Positron emission tomography (PET), 64–5
Posture, 87, 91–2, 168, 185, 192, 245–6, 309, 316, 319, 329–31, 336
Potential/talent, genetic/innate/natural
child's, 3–14, 17–20, 31–3, 39–40, 77, 225
expressive, 247, 343
Practice (individual rehearsal, training), 151–61
and anxiety, relaxation, 21, 53–4, 57, 79
and the brain, 63–7, 72–9
of conductor, 343
definition/goal of, 151–3
deliberate/quality, 35, 41, 118, 125–7, 131, 156, 160
distributed, 152–3, 161
duration/frequency/amount of, 7, 22, 32–7, 41, 71, 86, 90–1, 100, 107, 143, 152, 229–31, 238, 286, 315–6
and ear playing, 109
early, 110, 176
and expression, 228
formal/informal, 4, 12–4, 22, 27, 35, 41, 107–10, 153
group/solitary, 9, 13, 129
long-term, 143
massed, 152–3
medical, 84
mental (cognitive rehearsal), 41, 55–6, 76, 151, 153–4, 160
and motivation, 43
and motor learning/automation, 69, 74–7, 228, 238, 297–8
motor/movement, 167, 246–7
and musicality, 18–9, 41
over-, 129
and personality, 315
piano, 140, 286–90
and pitch discrimination, 184
role of teacher/parents, 21, 26, 157, 228
and sight-reading, 102, 109, 135–8, 141–8, 154
slow, 77, 288
string, 307–11, 314–5
structure/organization/planning of, 38–41, 152–60, 168, 173
supervised, 159
technical, 99–100, 147, 261
technique/activity/task/strategy, 92–3, 101, 107, 130, 143, 147, 178, 193, 213–4, 266
wind, 323
*See also* anxiety; environment; motivation; rehearsal; relaxation

Preference
intonational, 194, 280
music, 21, 24
singer, for SOR, 273–4
timbre, 10
Pressed phonation, 255, 260
Pressure
piano key, 214
string bowing, 293, 304–13, 319
vocal subglottal/lung, 255–64, 275
wind blowing, 206, 320–32
Production deficiency, 155
Proofreader's error, 141, 144
Proprioception, 265 (*see also* kinesthesia)
Protocols, verbal, 121
Proto-musical. *See* experience
Psychoacoustics, 271, 275, 278, 288–90 (*see also* perception)
Psychoanalysis, 52
Psychometric, 4, 5, 314
Psychosomatic, 89
Puberty. *See* voice development
Punishment, 11, 71, 89. *See also* motivation
Pythagorean. *See* intonation

Qualities. *See* attributes
Quality/standard
of music medicine, 84
of performance, 8, 48–51, 56, 152, 158, 161, 173, 185, 194, 238, 342, 345
of practice, 41, 159
of study, 107
*See also* timbre

Radif, Persian, 118
Rating/judgment
of emotion, 222, 226–7, 230–1
of intonation, 190, 290
of performance, 11, 55–7,
of tempo, 337
of tension, 207
Reasoning, musical, 64
Recall. *See* memory
Recorder
flute, 320, 324–7
tape, 105, 274, 277
Reed, mechanical/free/lip/air, 319–30
Referent. *See* improvisation
Reflex, auditory, 65, 293
Register (pitch)
and intonation, 188
and piano, 286, 292–4
and strings, 310
vocal, 260–1
and winds, 77
Rehearsal (ensemble) and conducting, 335–47
atmosphere, 336
behavior, 53
duration, 93
history of, 136
and intonation, 184, 189–90, 193–4, 282
mental, 54–6, 107, 110, 55–6
plan, 341
structure, 339–41, 346
*See also* conducting; practice
Relaxation
and anxiety, 50
breathing, 57
cue-controlled, 53
and improvisation, 120, 128
lip (winds), 327
maladaptive, 93
muscular, 52–4, 57, 246
music-structural, 214
and piano, 285–7
pressure (winds), 235–6
and singing, 255, 258, 261, 265
skills/techniques, 54, 57
and strings, 316
training, 53, 57
*See also* tension
Repertoire
accompanying, 143
child's/student's, 13, 41, 56, 110–1
of conducting gestures, 343–6
and expression, 202
and injury, 88, 92
and intonation, 191–2
jazz, 118, 122
and memory, 168, 173, 176
and practice, 156, 159–61
sight-reading, 136, 143, 148
string, 314–6
Repetition
and error correction, 154–6, 229
and feedback, 231
and injury, 91–2
and improvisation, 119–20, 127
and memory, 173–4
and motor skill automation, 63, 79, 112
piano tone, 288
Respiratory
movement, 329
muscle, 275, 325
system/apparatus, 253–8, 319–25
techniques, 319

Retrieval
processes in sight-reading, 147
structures/cues/schemes in memory, 170–3
Reverberation, 270–3, 278, 347
Reward, 22, 26, 35–6, 71, 176, 188. *See also* motivation
Rhythm
brain mechanisms, 69, 73
and conducting, 341
and improvisation, 118
and movement, 246
perception by fetus/infant, 5, 20, 237–8
and sight-reading, 102, 111, 147
and timing, 208–9, 239
Risk
instrumental/medical/occupational, 83–5, 91–2
taking in improvisation, 119–20, 130–1
Rosin, 307–13. *See also* noise
Rote. *See* learning; memory
Roughness, perceptual, 280, 291
*See also* consonance

Saccades, ocular, 137
Scales
musical, 36, 100, 147, 171, 188–90, 212, 265, 275, 280, 286, 298–9, 331, 337
stretched, 206
*See also* chord; rating
Schema/representation/blueprint
aural/auditory, 140, 155–6, 315
brain, 64, 70–80, 93
internal/mental/memory, 102, 118, 121, 172–3, 229
movement, 76
musical, 79
structural, 221
visual, 111
Score study/analysis, 79, 135, 156, 161, 174, 178, 298, 338, 346. *See also* audiation
Self-
appraisal, 51
awareness, 35, 246
beliefs about competence, success/failure, 32–43
concept, 10, 32, 35
determination, 42
efficacy, 34, 39–43
esteem, 8, 37, 48
identity, 265
monitoring, 144, 263–4, 274
perceptions, 34–6
regulation, 154, 340
talk, 51–7
theory, 39
to-other-ratio (SOR), 273–4, 278
worth, 32, 37, 42
Sensitivity
aural, 105, 147, 262, 288 (*see also* discrimination)
emotional, 228
motor, 314
musical, 104, 184
neural, 73
personal, 8, 11–4,
pitch, 194
Sensorimotor
activation/processing/program/region/system, 76
feedback, 296
learning, 74–6
mislearning, 77
plasticity, 74
representation, 79–80
skill, 74, 77
Sensory. *See* perceptual
Sight-reading, 70, 79, 99–113, 135–49, 154, 158, 168–70, 175, 201, 297–8, 314. *See also* belief; brain; cognition; development; effort; encoding; ensemble; error; exercise; expertise; expression; knowledge; literacy; proofreader; retrieval; score; strategy; stuttering
Sight-singing, 338
Singing, 4–6, 9, 12, 21, 26, 41, 79, 100–1, 107, 110–1, 168, 187–9, 228, 243, 253–67, 269–82, 330, 344–346
quality (*see* cantabile)
mental, 110 (*see also* practice, mental)
Situational exposition, 18
Skill. *See* ability
Slip-stick action of bow on string, 304–12. *See also* rosin
Socialization, 19, 23, 43, 262
Soft palate. *See* velum
Solfa/solfège/solfeggi, 110, 265
Somatosensory
area, primary (S1), 70
cortex, 77
feedback, 69
information, 77

Sound
  bone-conducted, *273*
  pressure level (SPL), 270, 274, 289, 293 (*see also* loudness)
  radiation, 313
  waves, 214, 259, 270, 319, 321, 322
Span
  anticipation time-, 124
  perceptual/eye-hand, 139, 148
  *See also* hand; life
Spectral
  decomposition in brain, 67
  envelope and timbre, 289–91
Spectrum
  violin, 312
  voice, 225, 254, 259–63, 278
Splinting, 87–9
Spring-mass system, 320
Staccato, 209–10, 225, 230, 259, 289–91, 312, 345
Staccato touch. *See* percussive touch
Stage fright. *See* anxiety
Steroids, 87–9
Stick. *See* slip-stick
Strategies/applications/implications
  against anxiety, 50–7, 153
  and brain research, 79
  conducting, 335, 340, 343–5
  ear-playing, 108
  expressive, 207, 212, 221, 226–33, 238, 245–8, 291
  improvisational, 118–20, 125–7, 131
  intonational, 192
  learning/teaching, 25–7, 37, 73, 79–80, 151, 193
  memory, 167–9, 174–9
  motivational, 32, 36, 39
  practice/rehearsal, 26, 36–41, 153–7, 160, 193–4, 296
  preventive (medicine), 92–4
  sight-reading, 109, 135–9, 143–7, 298
  singing, 253, 265–6, 279–81
  string playing, 307, 316, 319, 310
  wind playing, 331–2
Streaming, auditory, 293, 296
Stress
  and Alexander technique, 21
  of auditions, examinations, competitions, 43, 49–50
  effects of, 50, 331
  hormones, 20
  inoculation, 51–3
  management, 58, 319
  occupational, 91–2
  psychological/emotional, 93, 330
  and relaxation, 57
  situational, 50
  sources of, 50–1
  of performance, 36, 43, 49, 58, 88
String
  instrument playing, 11, 18, 77, 85, 106, 174, 186–8, 303–16
  motion (Helmholtz), 290, 304–12
  piano, 288–95
  resonance, 313
  torsion, 304
  vibrato, 313
  *See also* pizzicato; rosin; slip-stick
Structure
  macro-/micro-, 178
  musical, 69, 139–42, 173, 167–8, 172–8, 199–212, 220, 224, 239–42, 265, 294, 297, 344
  pedagogical, 26, 57
  *See also* brain; cognitive; formant; generative; improvisation; metrical; phrase; pitch; practice; rehearsal; retrieval
Stuttering (sight-reading), 148
Style
  cognitive, 155
  expressive, 225
  improvising, 122, 127, 131, 177–8
  musical/historical, 26, 174, 200–1, 214, 241, 247, 297, 344
  performance/playing, 107, 227, 244–7, 299, 341
  practice, 152
  singing, 264, 282
  *See also* learning; life-; teaching
Success. *See* failure
Surgery, 71, 86–7
Syndrome. *See* carpal tunnel; cortical sensory-motor; cubital tunnel; pain (musculoskeletal); thoracic outlet

Tactile. *See* perception
Tai Chi, 57
Talent. *See* potential
Task
  dependence on, 298
  difficulty/demands, 34–7, 43, 50, 63, 101, 142, 145–6
  effort, 32
  mastery, 38, 50
  multiple, 121, 148
  orientation, 53–4, 154
  presentation, 340
  value, 32–4

Teaching/instruction
conceptual, 342
nonverbal, 342
style, 99, 158
verbal, 361
*See also* computer; environment; expression; strategy
Technique. *See* ability; technical
Temperament (personality), 5, 8, 20, 128
Temperament, equal, 184, 187–8, 275, 280
Tempo
curve, 201–4, 239
and expression/interpretation, 147, 202–4, 223–4, 289
global vs local, 201–8, 214
and implied movement, 245
jazz (*see* double time)
and kinesthetic memory, 176
and sight-reading, 135
*See also* practice, slow
Temporal
constraints in jazz, 117
envelope and timbre, 289–91, 262
lobes of brain, 65–76
resolution of PET, MRI, 64–5
Tendinitis, 85, 89–90
Tendon
flexor, 87
pain, 86
swollen, 86–7
tension, 69
Tension
lip/embouchure, 323–9
muscle/physical, 54, 57, 69, 88–93, 246, 286–8, 316, 323
musical/harmonic/melodic, 92–3, 127, 178, 187, 206–9, 213, 214, 220
vocal fold, 260–1
Theory
of music, 105, 121, 126, 131, 147, 172, 176–7, 206, 212, 275
*See* also attribution; expectancy-value; flow; learning; self
Therapy/treatment/intervention
behavioral, 52
cognitive (-behavioral), 49, 52–5, 58
constraint-induced movement, 88
hypno-, 54–6
mistreatment, 89
rational-emotive, 54
*See also* antiinflammatory; cortisone; imagery; splinting; steroids; surgery
Thoracic outlet syndrome (TOS), 87
Thorax, 325, 330
Threshold, 184, 326
Throat
music, 253
singing, 264, 279
*See also* larynx
Thyroarytenoid, 255. *See also* muscle, shortener
Thyroid cartilage, 255. *See also* cricothyroid; thyroarytenoid
Timbre (tone/sound quality)
brain mechanisms, 67
definition of, 270
expressive, 220, 223–6, 239
piano, 288–94 (*see also* cantabile)
and pitch, 158, 185, 188, 192, 278
preference test, 10
string, 90, 303–13
vocal, 253–4, 258, 260–4, 265–6, 274, 278
vowel, 262–4
wind, 322, 325, 328–30
*See also* roughness
Timing
brain mechanisms of, 65, 71–4
of chord tones, 290–1, 296
computer simulation of, 214
ensemble, 210, 214
expressive, 199–214, 220, 223, 238–43, 289
and groove/swing/waltz, 208–9
and motor control, 238–9
in piano, 286
random, 212
in sight-reading, 144–6
Tiredness. *See* fatigue
Tonality, 26, 73, 102, 127, 130, 141, 171, 175, 201, 275, 282, 293
Tone. *See* attack; combination; decay; duration; edge; offset; onset; over-; pipe; timbre
Touch
piano, 285, 289–95
polyphonic, 294–6
precursor, 292
*See also* perception; tactile
Trachea, 257–9
Transcendence, 125–6, 130
Transglottal airflow, 254
Treatment, medical, 51–7, 83–90
Trial and error, 76, 102, 309
Trombone, 3, 193–4, 323–30
Trumpet, 22, 27, 142, 190–2, 325–9
Tuba, 11, 331
Tuner, electronic, 191, 194
Tuning fork, 279, 313, 320–1

Utility value, 32–3

Valve, wind, 142, 319, 322
Velocity
  air/jet, 319, 321, 325, 328
  artifact (piano), 294–6
  of nerve conduction NCV, 72, 87, 90
  of piano hammer, 289, 291
  of piano key, 286, 289, 292–3
  and ritardando, 204, 239
Velum, 258, 262
Ventricular folds, 258
Vibration. *See* oscillation
Vibrato
  extent of, 207
  pitch of, 189, 193
  singing, 261, 264, 279–80, 293
  string, 313
  wind, 155, 329
Viola, 4, 12, 194, 313–5
Violin, 11–2, 18, 71–2, 77, 168, 172, 188–90, 241, 303–16
Vision. *See* perception; visual
Visualization, 168, 178
Vocabulary
  expressive/gestural, 214, 231, 241–2
  language, 104
  musical/aural, 100–2
  teaching, 221
Vocal folds, 254–65, 275–6, 329
  false (*see* ventrical folds)
Vocal tract, 254–64, 281, 322–5
Voice, 253–82
  belt, 264
  care of, 84
  development/puberty, 154
  individual, 127, 131, 212–3

Wernicke region, 67, 72
Wind
  instruments, 155, 189, 206, 291, 319–32
  pipe. *See* trachea

Yoga, 57, 92

Zone of proximal development (ZPD), 339

Parncutt
McPherson

THE SCIENCE & PSYCHOLOGY
OF MUSIC PERFORMANCE

OXFORD

EL ARGONAUTA
la librería de la música
0-19-513810-4
PARNCUTT, RICHARD;
THE SCIENCE & PSYCOLOGY OF
P.V.P. (euros): 47.60



ISBN 0-19-513810-4

AMS Studies in Music

# Conceptualizing Music

## Cognitive Structure, Theory, and Analysis

Lawrence M. Zbikowski

# CONCEPTUALIZING MUSIC

AMS Studies in Music

LAWRENCE F. BERNSTEIN, *General Editor*

*Editorial Board*

Joseph Auner
Scott Burnham
Richard Crawford
Walter Frisch, Chair
Sarah Fuller
Robert Judd
Janet Levy
Jessie Ann Owens
Kerala Snyder
Judith Tick
Gary Tomlinson

# CONCEPTUALIZING MUSIC

Cognitive Structure, Theory, and Analysis

*Lawrence M. Zbikowski*

OXFORD
UNIVERSITY PRESS
2002

OXFORD
UNIVERSITY PRESS

Oxford New York
Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Singapore Taipei Tokyo Toronto

Copyright © 2002 by Oxford University Press, Inc.

Published by Oxford University Press, Inc.
198 Madison Avenue, New York, New York 10016

www.oup.com

Oxford is a registered trademark of Oxford University Press

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
electronic, mechanical, photocopying, recording, or otherwise,
without the prior permission of Oxford University Press.

Library of Congress Cataloging-in-Publication Data
Zbikowski, Lawrence Michael.
Conceptualizing music : cognitive structure, theory, and analysis / Lawrence M. Zbikowski.
p. cm. — (AMS studies in music)
Includes bibliographical references (p. ) and index.
ISBN 0-19-514023-0
1. Musical perception. 2. Musical analysis. 3. Cognition. I. Title. II. Series.
ML3838 .Z25 2002
781'.11—dc21 2001058756

1 3 5 7 9 8 6 4 2

Printed in the United States of America
on acid-free paper

*dem andenken meiner Mutter:*
*Anneliese Margerite Zbikowski, 1926–1999*

# PREFACE

On picking up a book with the title *Conceptualizing Music: Cognitive Structure, Theory, and Analysis*, one might reasonably assume that it deals with music cognition and how our knowledge of that discipline can be applied to music theory and analysis. This book does not do that, or at least not in a simple way. To begin, it does not have much to say about the fairly large body of research usually placed under the rubric "music cognition." This work, having been developed out of music psychology and informed by recent research in the brain sciences and mind sciences, proceeds by carefully crafted experiments, which are subjected to closely argued statistical and logical analysis. As practiced by such eminent and able researchers as Carol Krumhansl, John Sloboda, and David Huron, the study of music cognition has told us much about how humans process sonic and musical information.

But this book proceeds in a somewhat different way. Drawing on the same body of research from the brain sciences and mind sciences that shaped studies in music cognition, it explores how basic cognitive capacities are specified for understanding music. The project takes inspiration from recent work in linguistics and rhetoric by researchers like Ronald Langacker, Gilles Fauconnier, Mark Turner, and George Lakoff, and it is based on the assumption that musical understanding relies not on specialized capacities unique to the processing of patterned sound but on the specialized use of general capacities that humans use to structure their understanding of the everyday world. The methodology, in consequence, relies not on experimental design and data analysis but on using a broad and quite extensive body of research to interpret recurrent tropes of musical understanding. These tropes involve such things as the importance to musical understanding of relatively small and compact musical phenomena like "motives," "themes," and "chords"; the use of terms grounded in nonmusical domains—terms like "space" and "depth"—to characterize musical events; and the reliance on patterns of logical inference to reason about music.

The result of this investigation is a theoretical perspective on musical organization, but one rather different from what usually counts as "music theory." To make sense of this claim requires a bit of explanation about the intellectual context of music theory, for music theory is, within the rolling seas of humanistic studies, a

rather strange fish. Put bluntly, it is clear that much of what music theory does, as a discipline, does not count as any sort of theory in modern scholarship. This is exemplified by each of the two distinct but related and intertwined strands that make up contemporary music theory. One strand is occupied with pedagogy, the other with speculative and highly systematic approaches to musical organization.

Music theory, as it is presented in the classroom, is most often engaged with a careful and often relentless explication of what, for want of a better term, we can call musical grammar. Consider the following, from Edward Aldwell and Carl Schachter's *Harmony and Voice Leading*:

> Like VII$^6$, V$\frac{4}{3}$ has $\hat{2}$ as its bass. V$\frac{4}{3}$, in fact, resembles VII$^6$ so closely that they are almost interchangeable chords. The bass of V$\frac{4}{3}$ is a more neutral tone than that of V$\frac{5}{3}$ (or, as you will see, V$\frac{2}{}$) and can move convincingly either to $\hat{1}$ or to $\hat{3}$. Consequently, V$\frac{4}{3}$, like VII$^6$, forms a natural connection between I and I$^6$ and appears very frequently as a passing chord within an extended tonic.[1]

The prose and terminology are impressively dense. But one should not be misled into thinking that the authors are concerned only with abstruse compositional techniques, for immediately after this excerpt Aldwell and Schachter refer to a passage from an impromptu by Franz Schubert that illustrates the niceties of voice leading with which they are concerned. Their assumption is that the reader is familiar with the music and counts it as typical, and it is this familiarity that provides a phenomenological anchor for what might appear to be rather thick jargon. If you know Schubert's impromptu, or (better yet) can summon it in your sonic imagination when reading the example in score, Aldwell and Schachter's point about the harmonization of the second scale-step in the bass is not just clear but even obvious.

At the heart of pedagogical music theory are familiar or typical examples of music, the mysteries of which are revealed by a music theorist (or theorists) eager to share the secrets and wonder of this music with others. As elegant and persuasive as this approach might seem, it is, within our current cultural climate, more than a little unrealistic: music by Schubert and his contemporaries is often unfamiliar to the students who read Aldwell and Schachter's text (or any of a number of similar texts) and is not typical of the music that resonates through these students' digitized and hypercommercialized environments. That this should be so is often regarded as symptomatic of an illness of the late twentieth century, an illness that leads to an insufficient engagement with the great traditions of Western culture. For Classical music (as it is so styled), the antidote is music theory. Music theory, with its careful explication of the musical grammar of Mozart and Beethoven and Schubert, thus becomes the last redoubt against the dissolution of Western culture represented by a dwindling interest in the music of eighteenth- and nineteenth-century Europe.

If, for a moment, we step back from Aldwell and Schachter's text and generalize its intent beyond the specific repertoire relative to which it is framed, we might be able to avoid this rather sanctimonious stance. We could argue that music is a highly complex and idiosyncratic mode of human communication and that having a knowl-

1. Edward Aldwell and Carl Schachter, *Harmony and Voice Leading*, 2nd ed. (San Diego: Harcourt Brace Jovanovich, 1989), 112.

edge of the grammar of this mode of communication is essential to its deeper appreciation, no matter what form music might take. The argument is a familiar one to me, not the least because I often find myself making it. And yet, something rings hollow. The grammar that music theory teaches is unavoidably tied to the repertoire to which it refers, and just how this is generalized to apply to other repertoires is not immediately apparent: I know of no theory text that explains how the grammar of Schubert's musical discourse is manifested in the music of, say, John Coltrane or Prince. Given its isolation from contemporary culture, the music theory of the classroom appears to be little more than a ghost that haunts the echoing halls of a crumbling cultural empire.

The second strand of music theory partakes of the systematic quality inherent in grammars but generalizes it away from natural language and toward a free-standing intellectual construct. As an example of this sort of theory, consider the following brief passage from David Lewin's analysis of a section of the opening of Claude Debussy's piano prelude *Reflets dans l'eau*. In this excerpt, X, Y, and Z represent specific collections of musical notes, RI refers to the compound operations retrograde and inversion, RICH is a function that effects serial transformations, and T refers to transposition:

> In measure 10 the music of measure 9 is repeated and extended. The crescendo recurs. In the melody the repetition gives rise to a rotated form of $Z_1$, marked "rot $Z_1$" on figure 10.10. Rot $Z_1$ is B♭-A♭-F-E♭; it embeds serially the original form of Y, A♭-F-E♭, and precedes this Y by its overlapping inverse-RI-chained form B♭-A♭-F. (B♭-A♭-F is $RICH^{-1}$(A♭-F-E♭).) This relationship is more or less inherent in the derivations of X, Y, T(X), their repetitions, and $Z_1$.[2]

It is, of course, something of a challenge to evaluate this passage in isolation. Not only is it just one part of a larger analysis, but also it comes late in a book occupied with various and sundry applications of formal algebra and mathematical mapping theory to music. Nonetheless, what should be clear is that more than familiarity with Debussy's prelude is required to make sense of Lewin's interpretation of the passage. The reader must also be familiar with a style of abstract thought that is bound to appear cabalistic to the uninitiated, one in which the transformation of musical entities is at least as important as the entities themselves. For some, the inaccessibility of this mode of thought is one of its charms. For others, it is proof of the hermeticism and irrelevance of music theory.

Before continuing, I should make clear that I have deep respect for the theorists whose work I have cited. I use this work in my teaching and continue to be intrigued and stimulated by it. I also want to emphasize that these excerpts by no means reflect all that there is to music theory. I take them as representative of two strands of thought that are replicated and woven together in all sorts of different ways to create the texture of contemporary theoretical practice. What is important for my purposes here is that the practice of contemporary music theory is not like that of contemporary cultural or social theory. Instead of probing the cultural or historical

2. David Lewin, *Generalized Musical Intervals and Transformations* (New Haven, Conn.: Yale University Press, 1987), 234.

context for musical utterances, or the complex networks of social interaction that give rise to musical behavior, music theory continues to focus on details of musical discourse with an obsessiveness that is both maddening and quixotic to cultural and social theorists.

Given the impressive traditions of music theory and their influence on my own thought, I cannot guarantee that what I offer here is a great improvement on this situation. One of the things I want to do in the chapters that follow, however, is to develop a somewhat different view of music theory—one that sees music theory as a response to a problem. The problem is that of musical understanding: how it is that we can make sense of sequences of nonlinguistic patterned sound, that we can do so with amazing rapidity, and that (often as not) we can return to these or similar sound sequences and find continued reward. I will argue that our understanding of music relies on a play of concepts and conceptual structures that emerges from training basic cognitive capacities on musical phenomena and that music theory and music analysis derive from this play.

This mode of inquiry is not one common in the discipline of music theory, despite its similarities to work done by Leonard B. Meyer, Eugene Narmour, and Robert Gjerdingen. It does, however, share features with the approach to music evident in Susan McClary's recent *Conventional Wisdom*. That this should be so reveals a debt on my part, for it was Professor McClary who first suggested that I read Mark Johnson's *The Body in the Mind* back in the late 1980s, and this had a profound effect on my research. In her book, McClary explores the cultural and historical forces that have shaped genres like opera and the blues and compositional practices like tonality. My focus is on a somewhat different set of forces: those that shape the way humans think. It seems inevitable that these forces are in some sort of grand, if ill-defined, dialogue with cultural and historical agencies: it is, after all, human cognitive processes and human culture and history about which we are talking. That cultural and social theory have turned a deaf ear to this dialogue is not surprising: cognitive theory has had little room for and less patience with culture, and the detail of its investigative method is no match for the epic sweep of high theoretical practice at its best. And yet it seems we must, at some point, come to terms with cognitive structure, for if we do not develop an understanding of how cognitive processes shape the basic materials of thought, we risk accepting these materials as things given by nature, just as culture and history—and music, for that matter—were once assumed to be given by nature.

A glimpse of the problem can be seen in McClary's compelling analysis of Robert Johnson's 1936 "Cross Road Blues." McClary contrasts Johnson's blues with those of Bessie Smith, noting that the influence of Johnson's music on white British blues players of the 1960s was due in part to a misconception: the British musicians believed that the idiosyncrasies of Johnson's blues style represented "authentic" blues practice. Describing the unique sound of Johnson's "Cross Road Blues," McClary writes,

> An affect of dread and entrapment pervades this tune—partly the result of his strangulated, falsetto vocals and his uncanny replication of that timbre on the guitar. Moreover, Johnson's percussive guitar pulse, which locks in at the eighth-note level, allows almost no sensual movement: even though Johnson's singing constantly strains

> against that beat, the listener's body is regulated by those short, aggressively articulated units. The guitar thus seems to represent simultaneously both oppressive outside forces and a desperate subjectivity fighting vainly for escape.[3]

I have no quarrel with McClary's analysis—indeed, in this short passage, she has captured a number of the essential features of Johnson's performance style. The difficulty comes with the ultimate justification for the affect produced by the song. Why do Johnson's vocals and guitar work yield dread and entrapment rather than joyful anticipation and a feeling of liberation? Clearly, the rhythmic framework is important, but why is it that the rigidity of Johnson's beat constrains us rather than providing a secure foundation from which we can coolly regard his plaint? These are not easy questions, all the more so because of the relatively unique character of Johnson's recording when compared with other blues recordings of the period, and because hearing "Cross Road Blues" as something other than a moving, haunting song is to misunderstand it rather thoroughly. Cultural, social, and historical context cannot, by themselves, explain the origin of our affective response to the song, for our broad agreement on the effect of Johnson's music transcends these implements of high theory (even if they have a profound influence on what we do with "Cross Road Blues" once we have heard it). I propose that exploring the way cognitive structure informs our understanding of music gives us a way to account for the source of our broad agreement on the affect that pervades Johnson's blues and can help us understand better the ways culture, society, and history reshape musical practice.

Again, the way I want to accomplish this is by reconceptualizing what it means to theorize about music. This can be done by approaching music theory from the perspective provided by recent work in cognitive science. My intellectual guideposts include not only the wealth of work done in the mind sciences and the brain sciences but also contemporary and historical ways of theorizing about music. These theories of music—now with "theory" understood in somewhat more traditional terms—capture important aspects of how it is we structure our understanding of music. At their best, they represent a technical and systematic articulation of an accord on what matters in music that is similar in kind, if not terminology, to our accord on how Robert Johnson's blues move us: we agree on how music is put together and we agree on what music means because both are structured by basic cognitive processes through which we organize our understanding of the world. And so music theory—this rather strange fish in the seas of humanistic scholarship—may yet tell us some quite interesting things about the cultural and social construction of music. Understanding the way music theory instantiates cognitive processes will also help explain its continuing value. It will offer such help whether that theory be prosaic—as within the classroom or within critical discourse (since even the most radical of cultural or sociological theorists inevitably makes recourse to basic music-theoretical constructs)—or poetic, as with the elusive and allusive constructs of abstract theory.

3. Susan McClary, *Conventional Wisdom: The Content of Musical Form* (Berkeley: University of California Press, 2000), 51.

As should be apparent from the preceding, my intent here is to address not only music theorists but also musicologists and ethnomusicologists who find the challenges of theorizing about music intriguing. What follows will also be of interest to those with either a professional or avocational interest in cognitive science, for music presents a number of interesting problems for cognitive processing, not the least of which are its embeddedness in culture and the demands it places on real-time processing. In entertaining the thought of such an audience, however, I should briefly clarify a distinction I draw between sound and music and the cognitive abilities related to each, which is based on three suppositions. First, not all sound is music. Second, an account of how humans process sound is not the same thing as how they understand music. Third, phenomena relevant to musical understanding exist at a conceptual level—that is, at a level of cognitive activity at least potentially accessible to conscious thought. I should emphasize that I regard the conceptual level as occupying only a small part of our total cognitive activity, and I am not at all opposed to efforts by music psychologists and others who try to explain how structures at the preconceptual level connect with and motivate structures on the conceptual level. For me, however, it is at the conceptual level that I find the most profoundly interesting questions, for concepts are the tools that allow us to construct the complex notions essential to musical understanding. From this perspective, conceptualizing music is fundamental to inquiries about music, whether those be from the perspective of music cognition, or ethnomusicology, or musicology, or theory.

# ACKNOWLEDGMENTS

In the course of writing this book I enjoyed the assistance of many people, all of whom improved the product immeasurably. Among the many colleagues who read drafts of material or responded to presentations I have given, are the following: Kofi Agawu, Holly Aksnes, Jeanne Bamberger, Larry Barsalou, Philip Bohlman, Candace Brower, Scott Burnham, Clif Callender, Thomas Christensen, Martin Clayton, Rick Cohn, Nick Cook, Arnie Cox, Bob Gjerdingen, Robert Hatten, Bob Holzer, Brian Hyer, Rich Janda, Carol Krumhansl, the late Jim McCawley, Marc Perlman, Anne Robertson, John Rothgeb, Janna Saslaw, Martin Stokes, and Mark Turner. Their comments, suggestions, and queries helped me strengthen and reformulate my arguments. Ben Brinner, Gilles Fauconnier, Douglas Hofstadter, Travis Jackson, Bobby Short, Sumarsam, and Susan Youens were generous in their response to questions I posed to them and provided me with information and insight I simply could not have gained elsewhere. Students in my classes and seminars at the University of Chicago listened to many of the ideas presented in this book, to which they responded with questions, challenges, and comments. Although it would be impractical to acknowledge all of them here, they contributed immensely to the shape those ideas have taken in this book.

Deborah Gillaspie of the Chicago Jazz Archive and Regenstein Library patiently endured my requests for obscure recordings and was indispensable in helping me find materials that I needed. The musical examples were expertly prepared by Jürgen Selk of Music Graphics International. Philipp Goedicke and Heinrich Jaeger helped me with the translation of difficult passages in German, and my brother Gene Zbikowski assisted in deciphering some of the more recalcitrant clauses in Rameau's and Proust's French. José António Martins and Aron Topielski served as my research assistants, ably tracking things down, organizing my sprawling files, and getting to know the photocopier, perhaps better than they wanted to. Maribeth Payne, when she served as Executive Editor for Music at Oxford University Press, listened to my ideas for the book (at least once at the risk of missing her train) and then provided valuable encouragement and assistance in the process of turning a proposal into a final manuscript.

Institutional support came through a yearlong fellowship at the Franke Institute

for the Humanities (formerly the Chicago Humanities Institute) at the University of Chicago and from the Department of Music of the University of Chicago. The latter provided not only a home for my research but financial support as well. Some of the recordings discussed in chapter 5 were made available by the Archives for Traditional Music in Bloomington, Indiana, and by the Bowling Green Sound Recordings Archives, Bowling Green, Ohio.

Finally, production of this book was facilitated by subventions from two scholarly societies. The Society for Music Theory provided a subvention to defray the costs of producing the musical examples, and the result of this assistance can be seen in the number and appearance of the examples I was able to include. The Publications Committee of the American Musicological Society was equally enthusiastic about the book and agreed to include it in its series, AMS Studies in Music. This decision not only made the book more affordable but also gave me an excellent and endlessly patient editor in Lawrence Bernstein.

My family—Vicky Long, Anna Katia, and Andrei Nicolai—had to put up with sharing me with this project, and I am afraid they got the short end of the deal more than once. Without their support, understanding, and uncritical companionship, this book would have been much the poorer.

My musical examples generally adhere to the readings of the following editions: the new *Beethoven Werke*; the Johannes Brahms *Sämtliche Werke*; the *Neue Mozart Ausgabe*; Casimiri's edition of the works of Palestrina; the *Neue Schubert Ausgabe*; Clara Schumann's edition of the *Lieder* of Robert Schumann; the new *Richard Wagner: Sämtliche Werke* (in conjunction with the C. F. Peter's score of *Tristan*); and the Collected Works of Giaches Wert for *Corpus mensurabilis musicae*. Some indications of editorial intiative in those editions—generally with respect to dynamic markings added on the basis of parallel readings—have been realized tacitly in my examples for the sake of clarity of appearance. The setting of Bernhard Klein's "Trockne Blumen" was prepared from the 1822 edition by E. H. G. Christiani in Berlin.

When dealing with historical sources written in languages other than English, I have generally followed one of two practices. Where there is no translation of the work, or where I think existing translations are not wholly satisfactory, I provide my own translation, with the original in a footnote. Where satisfactory translations exist, I use these and do not include the original in the footnote. In a few cases, I have thought readers might like to make reference to both the original and a translation and consequently have included citations for both in the footnotes.

Some of the material that follows has been previously published elsewhere and is used here by kind permission of the journals in which it originally appeared. This material includes the following: "Conceptual Models and Cross-Domain Mapping: New Perspectives on Theories of Music and Hierarchy," *Journal of Music Theory* 41/2 (© 1997, *Journal of Music Theory*); "Musical Coherence, Motive, and Categorization," *Music Perception* 17/1 (© 1999 by The Regents of the University of California); and "The Blossoms of 'Trockne Blumen': Music and Text in the Early Nineteenth Century," *Music Analysis* 18/3 (© 1999, Blackwell Publishers, Ltd.).

# CONTENTS

Introduction: Conceptualizing Music 3

PART I. ASPECTS OF COGNITIVE STRUCTURE

1. Categorization 23

2. Cross-Domain Mapping 63

3. Conceptual Models and Theories 96

PART II. ANALYSIS AND THEORY

4. Categorization, Compositional Strategy, and Musical Syntax 137

5. Cultural Knowledge and Musical Ontology 201

6. Words, Music, and Song: The Nineteenth-Century Lied 243

7. Competing Models of Music: Theories of Musical Form and Hierarchy 287

Conclusion: Cognitive Structure, Theory, and Analysis 325

Bibliography 335

Index 353

# CONCEPTUALIZING MUSIC

# CONCEPTUALIZING MUSIC

Early in the opening volume of *A la recherche du temps perdu*, Marcel Proust presents the first of a number of memorable accounts of listening to music. He describes Charles Swann's initial encounter with the andante of Vinteuil's sonata for violin and piano:

> Doubtless the notes which we hear at such moments tend, according to their pitch and volume, to spread out before our eyes over surfaces of varying dimensions, to trace arabesques, to give us the sensation of breadth or tenuity, stability or caprice. But the notes themselves have vanished before these sensations have developed sufficiently to escape submersion under those which the succeeding or even simultaneous notes have already begun to awaken in us. And this impression would continue to envelop in its liquidity, its ceaseless overlapping, the *motifs* which from time to time emerge, barely discernible, to plunge again and disappear and drown, recognized only by the particular kind of pleasure which they instill, impossible to describe, to recollect, to name, ineffable—did not our memory, like a laborer who toils at the laying down of firm foundations beneath the tumult of the waves, by fashioning for us facsimiles of those fugitive phrases, enable us to compare and to contrast them with those that follow. And so, scarcely had the exquisite sensation which Swann had experienced died away, before his memory had furnished him with an immediate transcript, sketchy, it is true, and provisional, which he had been able to glance at while the piece continued, so that, when the same impression suddenly returned, it was no longer impossible to grasp.[1]

What Proust summons in this lyrical, enchanted vignette is the awakening and initial consolidation of musical understanding. Swann's first impressions of Vinteuil's sonata are vague and unformed, his mind simultaneously struggling with and savoring the ineffability of the music. But then, with the aid of memory, patterns emerge. Although these are incomplete and subject to revision, they offer him a way to make sense of the music, even as it continues to play.

*Conceptualizing Music* provides an exploration of the process of musical under-

1. Marcel Proust, *Swann's Way* (vol. 1 of *Remembrance of Things Past*), trans. C. K. Scott Moncrieff and Terence Kilmartin (New York: Vintage Books, 1981), 227.

standing—that is, the process through which those liquid impressions spoken of by Proust are transformed into structures that make it possible to grasp music. In what follows, I argue that Swann's—or anyone's—understanding of music draws on the same cognitive processes that humans use to organize their understanding of the world as a whole. Confronted with musical sound, these processes create musical concepts, the things that enabled Swann to gain a grasp of the music. The act of conceptualizing music is the beginning of a whole chain of cognitive events that allow us to theorize about music and to analyze the things that populate our aural past, present, and future.

The notion that Swann's musings might give rise to musical concepts demands some further consideration. Concepts are often thought of as highly stable cognitive structures of considerable complexity, a view hardly commensurate with the ephemera attended to or produced by Swann. Recent work in the brain sciences and the mind sciences, however, has changed how we view concepts. There are now persuasive arguments that concepts are quite fluid, that they are not irrevocably wedded to words or to concrete representations, and that they are not even unique to our species.[2] In consequence, the provisional replicas of musical phrases that make it possible for Swann to secure a foothold in the unfamiliar terrain of Vinteuil's sonata need not automatically be excluded from the conceptual domain. In fact, they are very much like the concepts we use to structure our understanding of the everyday world.

This same work in the brain sciences and the mind sciences suggests a way to account for the apparent simplicity and immediacy of musical understanding, which seems incommensurate with the complexity of musical structure. For instance, the sonata by the fictional Vinteuil is intended to be a relatively complex contemporary work that has captured the fancy of the musical elite of the Paris salons. Nonetheless, Swann, whose connoisseurship does not extend to music, is able to gain a grasp of the work almost immediately. That he is able to do so is not simply novelistic license but is, in fact, thoroughly plausible: almost everyone has had the experience of listening to an unusual composition or exotic repertoire and being able to make *something* of it. This possibility has suggested to some a latent musicality in humans comparable to the sort of competence for language proposed by Noam Chomsky.[3] Competencies of this sort raise as many questions as they answer, however, particularly where cultural entities such as music are concerned. It seems more promising to follow the path of researchers who have rejected linguistic competence as a given

2. See, for instance, Douglas R. Hofstadter and the Fluid Analogies Research Group, *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought* (New York: Basic Books, 1995), chaps. 5–6, 8–10; Gerald M. Edelman, *The Remembered Present: A Biological Theory of Consciousness* (New York: Basic Books, 1989), chap. 8; and Donald R. Griffin, *Animal Minds* (Chicago: University of Chicago Press, 1992), chap. 6.

3. The notion of musical competence has generated a range of commentary and scholarship. See John Blacking, "Music, Culture, and Experience," in *Music, Culture, and Experience: Selected Papers of John Blacking*, ed. Reginald Byron, with a foreword by Bruno Nettl (Chicago: University of Chicago Press, 1995), 228–31, on musical competence and culture; Steven Pinker, *How the Mind Works* (New York: W. W. Norton, 1997), 528–38, on musical competence and its relation to other competencies; and Allan Keiler, "The Origins of Schenker's Thought: How Man Is Musical," *Journal of Music Theory* 33 (1989): 273–98, on a nineteenth-century conception of musical faculties akin to musical competence.

and who have set about exploring the cognitive foundations of language.[4] Their task has been to discover what processes are basic to human cognition and then to determine how they are specified for language. For my part, I would like to explore how some of these same general cognitive processes are specified for music. By this means I hope to account for the apparent ease and real rapidity with which we can conceptualize a highly complex, completely unfamiliar music on our first encounter, without having to postulate the faculty of musical competence.

At a bit more of a remove, but no less important for a complete account of cognitive processing, is the way concepts come to be organized into the more extended cognitive structures with which our thought is usually occupied. This is where theories come into play, for theories are the cognitive tools that guide the way we reason about the things we experience. At first, this might seem to be a rather specialized use of the notion of "theory," for the theory with which much current literature is occupied—within and without music scholarship—is hardly the stuff of everyday experience. Recent research has countered this view of theory by demonstrating that the elegant and abstract theories of science have much in common with the tools for reasoning used by very young children.[5] Theories are the basic means by which we make our experience coherent and guide further action. The rough-and-ready transcript that guides Swann's listening thus has something in common with the more fully articulated and systematic structures we usually associate with the idea of "music theory."

That music theory might have alliances with everyday thought processes is a provocative claim. On the one hand, music theory often manifests itself as a relentlessly practical discipline: a codification of the scales, chords, and grammatical rules proper to a highly circumscribed portion of musical discourse, assembled with the intent of rendering music comprehensible to those who would become musically literate. On the other hand, music theory can reach into the far corners of abstraction to embrace complicated mathematical concepts or the arcane symbologies of voice-leading graphs, as any reader of the *Journal of Music Theory, Music Theory Spectrum*, or *Music Analysis* will quickly discover. Nonetheless, I want to argue that music theory, in all its diverse forms, reflects the same basic processes that guide our understanding of the everyday world. Theorizing about music is an activity specialized only in its domain, not in the cognitive processes it involves.

What might these cognitive processes be, and how would they manifest themselves? To answer these questions, let us begin at a beginning, with two theories of

4. For general introductions to some of the working assumptions of cognitive linguists, see George Lakoff, "The Invariance Hypothesis: Is Abstract Reason Based on Image-Schemas?" *Cognitive Linguistics* 1 (1990): 39–74; and Michael Tomasello, "Introduction: A Cognitive-Functional Perspective on Language Structure," in *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, ed. Michael Tomasello (Mahwah, N.J.: Erlbaum Associates, 1998), vii–xxiii. In the field of cognitive linguistics, Ron Langacker's work is particularly notable for its thoroughness and its systematic approach. See Langacker, *Theoretical Prerequisites*, and *Descriptive Application*, vols. 1 and 2, respectively, of *Foundations of Cognitive Grammar* (Stanford, Calif.: Stanford University Press, 1987, 1992); idem, *Grammar and Conceptualization* (Berlin: Mouton de Gruyter, 2000).

5. See Alison Gopnik and Andrew N. Meltzoff, *Words, Thoughts, and Theories* (Cambridge, Mass.: MIT Press, 1997); and Alison Gopnik, Andrew N. Meltzoff, and Patricia K. Kuhl, *The Scientist in the Crib: Minds, Brains, and How Children Learn* (New York: William Morrow, 1999).

music from Greek antiquity. These theories, and the music to which they refer, are so unfamiliar that even many who make music theory the focus of their research have only passing knowledge of what they involve. But there is an advantage in this unfamiliarity, for the disorienting effect it can have also serves to loosen our notions about what a theory of music, or a theoretical construct, should be. Despite connections with music theory as it was practiced in Europe and (to a lesser extent) in the Arabic world, Greek music theory of antiquity is the theory of an alien society. Nonetheless, this theory was a beginning—one of the starting points for accounts of musical organization in the Western tradition. We can discern within it, therefore, the theorists of antiquity grappling with basic constructs equivalent to, if still different from, the sort of constructs taught in beginning music theory classes today. Both aspects will allow us to see the role basic cognitive structures play in our understanding of music.

## ANCIENT MUSIC THEORY AND MODERN COGNITIVE SCIENCE

Those who wrote on Greek musical practice in antiquity concerned themselves with a wide variety of topics, including the place of music in society, musical aesthetics, the construction and nature of musical instruments, and the organization of pitch materials. Specific discussions of pitch materials—the usual topic of disquisitions more directly oriented to music theory—centered around the set of pitch relationships that has come to be called the Greater Perfect System.[6] As shown in figure I.1, this consisted of a set of four tetrachords (*hypaton, meson, diezeugmenon,* and *hyperbolaion*), which, together with one additional note (called *Proslambanomenos*), provided the framework for a two-octave system of pitches basic to Greek music. The end points of the tetrachords (the notes *Hypate hypaton, Hypate meson, Mese,* and so on), together with *Proslambanomenos*, were regarded as stable, unmovable pitches. There was, however, no fixed standard for tuning—the note-names given in figure I.1 are simply for the purposes of illustration. Within the boundaries marked by each tetrachord were two other pitches, whose placement varied according to which of three different genera was understood to be in play. (In the tetrachord *hypaton*, for instance, the variable notes above *Hypate hypaton* were *Parhypate hypaton* and *Lichanos hypaton*.) The diatonic genus located the movable pitches in a manner analogous to modern diatonic scales (for instance, given the reference pitches on fig. I.1, *Parhypate hypaton* in the diatonic genus would be equivalent to F3, and *Lichanos hypaton* would be equivalent to G3), but the chromatic and enharmonic genera situated the pitches in ways that have no comfortable analogue in modern scale construction. The result was a system in which the placement of movable pitches could vary widely and in which intervals between successive pitches could be smaller than a half step and larger than a whole step.

6. For a more detailed discussion of the Greater Perfect System and its place in Greek theory of antiquity, see Andrew Barker, ed. and trans., introduction to *Harmonic and Acoustic Theory* (vol. 2 of *Greek Musical Writings*), Cambridge Readings in the Literature of Music (Cambridge: Cambridge University Press, 1989), 12–13. For a more general overview of Greek music theory and a thorough discussion of the sources, see Thomas J. Mathiesen, *Apollo's Lyre: Greek Music and Music Theory in Antiquity and the Middle Ages* (Lincoln: University of Nebraska Press, 2000), part 4.

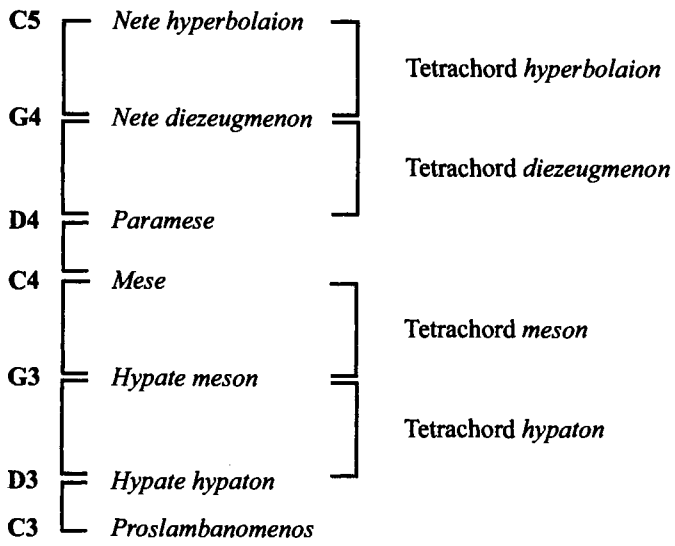



FIGURE I.1 Diagram of the Greater Perfect System of Greek music theory

There were two main theoretical approaches to presenting the elements of the Greater Perfect System. The Pythagorean approach derives from a metaphysics associated with Pythagoras of Samos, who lived during the sixth century B.C. Its hallmarks are a persistent interest in number and the deployment of numerical conceptions in cosmological contexts. In contrast to this, the Aristoxenian approach, associated with the Peripatetic school of the fourth century B.C., places little reliance on number, trusting instead in observation and reason as the means to knowledge about music.

## Pythagoras and the Blacksmiths

Nicomachus of Gerasa, a mathematician and harmonist writing around the beginning of the second century A.D., described Pythagoras's discovery of the basic principles of music theory as follows:

> He [Pythagoras] was plunged one day in thought and intense reasoning, to see if he could devise some instrumental aid for the hearing which would be consistent and not prone to error, in the way that sight is assisted by the compasses, the measuring rod and the *dioptra*, and touch by the balance and by the devising of measures; and happening by some heaven-sent chance to walk by a blacksmith's workshop, he heard the hammers beating iron on the anvil and giving out sounds fully concordant in combination with one another, with the exception of one pairing; and he recognized among them the consonance of the octave and those of the fifth and the fourth. He noticed that what lay between the fourth and the fifth was itself discordant, but was essential in filling out the greater of these intervals. Overjoyed at the way his project had come, with god's help, to fulfillment, he ran into the smithy, and through a great variety of experiments he discovered that what stood in direct relation to the difference in the

> sound was the weight of the hammers, not the force of the strikers or the shapes of the hammer-heads or the alteration of the iron which was being beaten. He weighed them accurately, and took away for his own use pieces of metal exactly equal in weight to the hammers.[7]

Nicomachus continues the story by describing how Pythagoras used the weights to conduct further experiments. After suspending the weights from identical strings, Pythagoras plucked pairs of strings and discovered the same concords as he had heard produced by the blacksmiths. He further discovered that the interval of an octave was produced by weights in a 2:1 ratio, that of the fifth by weights in a 3:2 ratio, and that of the fourth by weights in a 4:3 ratio, as shown in figure I.2a. The one discordant interval—that of a second sounded by the middle two weights—was the product of a 9:8 ratio. Additional experimentation showed that the smallest weight sounded a fourth, with the next to smallest weight (8:6 ≡ 4:3) and a fifth with the next to largest (9:6 ≡ 3:2; see fig. I.2b). The octave could thus be viewed as the product of either a fourth plus a fifth (12:9:6; fig. I.2c) or of a fifth plus a fourth (12:8:6; fig. I.2d). According to Nicomachus, Pythagoras also discovered that these ratios held constant throughout the musical domain. It made no difference whether the constituent notes of the intervals were produced through string tension, string division, beating on pots, or blowing on tubes—the relationships between these notes always reduced to the self-same ratios.[8]

The Pythagorean view of music outlined by Nicomachus assumes that music has its origins in the natural world and that the natural world has a basic (if often unseen) order that can be expressed through number. It is thus important to Nicomachus's story that the ringing of the hammers is accidental and not contrived: their harmony has everything to do with the inherent order of the world and almost nothing to do with the blacksmiths. It is also significant that the concordant intervals are immediately apparent to Pythagoras and that he can discern them even amid the discordant clang of the major second. Not only is the basis of musical order natural, but also it is manifest to all who have ears to hear. The association of these intervals with the pounding hammers provides the computational tool for

7. Nicomachus, *Enchiridion*, in Barker, *Greek Musical Writings*, 2: 256–57. The *dioptra* was a rod that was used for the indirect measurement of the height of tall objects.

8. Nicomachus, *Enchiridion*, 258. It should be noted that Nicomachus's story is a complete fiction. There is no evidence whatsoever that Pythagoras ever conducted any empirical research on the acoustic origins of harmonic relationships, with or without blacksmiths. Perhaps more important, the ratios described by Nicomachus simply do not work. To sound the intervals described in the story, the values of the weights must be squared—that is, the weights must be in the ratio 4:1 to produce the octave, 9:4 to produce the fifth, and 16:9 to produce the fourth. The ratios given by Nicomachus only work when used to segment a string into different sounding lengths. One-half the length of a string will sound an octave with the entire length of the string; two-thirds the length of the string will sound a fifth with the entire length of the string; and three-fourths the length of a string will sound a fourth with the entire length of the string. The importance of Nicomachus's story lies in its influence: in the form Boethius gave it in the sixth century, it became the standard account of basic Pythagorean principles for the Middle Ages and Renaissance; see Ancius Manlius Severinus Boethius, *Fundamentals of Music*, ed. Claude V. Palisca, trans. Calvin Bower, Music Theory Translation Series (New Haven, Conn.:Yale University Press, 1989), 17–19. A somewhat different perspective on Nichomachus's story can be found in Mathiesen, *Apollo's Lyre*, 399.

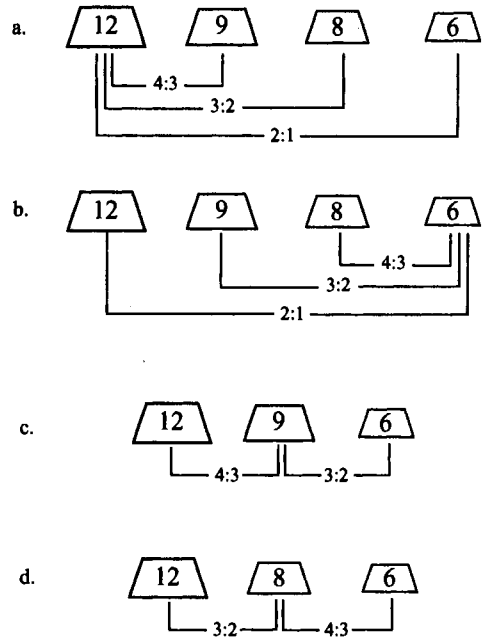



FIGURE I.2 Ratios of the octave, fifth, and fourth from Pythagorean legend

which Pythagoras had searched, for it allowed him to translate the constituent notes of the harmonic intervals into magnitudes—that is, into number.[9] These numbers then provide further proof of the order of nature, for the ratios of the concordant intervals (the octave, fifth, and fourth) are all simple (1:2, 2:3, and 3:4), while the ratio of the discordant second is relatively complex (8:9). Finally, the numbers involved in the ratios of the concords—1, 2, 3, and 4—were also those of the *tetraktys* of the decad, which Pythagoreans regarded as the "fount and root of ever-flowing nature."[10]

The account of musical organization presented by the story of Pythagoras and the blacksmiths is a model of concision. At its core are but three intervals: the octave, fourth, and fifth. The notes that make up these intervals are assigned magnitudes, the correspondence of which yields numerical ratios. These ratios come to stand for the intervals—the ratio 1:2 is the octave—and can also be used to describe relationships between them, leading to a precise account of the composi-

9. The notion of a computational tool I employ here derives from the work of Edwin Hutchins, especially that presented in the second chapter of his *Cognition in the Wild* (Cambridge, Mass.: MIT Press, 1995). Hutchins proposes that various of the navigational tools used by seafarers are in fact computational tools in that they facilitate computation by transforming analog information into digital information. This is exactly what the hammers did for Pythagoras: they transformed the analog information of sound into the digital information of hammer weights (with an intermediary stage occupied by the computational tool of a scale for measuring the weights).

10. Barker, *Greek Musical Writings*, 2: 30. A *tetraktys* is any coordinated group of four items; those of the *tetraktys* of the decad sum to 10, which is the basis of the base-10 number series used by the Greeks.

tion of intervals. The interval between the *Mese* and *Proslambanomenos* of figure 1.1 is an octave, a 1:2 ratio. This octave is made up of a fourth (from *Mese* to *Hypate meson*, a 3:4 ratio) and a fifth (from *Hypate meson* to *Proslambanomenos*, a 2:3 ratio). The fifth is in turn made up of a fourth (from *Hypate meson* to *Hypate hypaton*, another 3:4 ratio) and a tone (from *Hypate hypaton* to *Proslambanomenos*, an 8:9 ratio). With these components in place, a Pythagorean theorist could describe, through number, relationships between any of the fixed notes of the Greater Perfect System and eventually characterize the relationships that obtained among the movable notes of the various genera. The order of the cosmos, which was for Pythagoreans the order of number, thus found sounding expression in the domain of music.

## Aristoxenus and Aristotelianism

Although the Pythagorean perspective on musical order was an influential one—for instance, it informed Plato's and Aristotle's writings on music—it was not the only one available to antiquity. The alternative offered by Aristoxenus in his *Elementa harmonica* (most likely written toward the end of the fourth century B.C.) starts with a definition of the science of harmonics:

> It is to be understood as the science which deals with all melody, and inquires how the voice naturally places intervals as it is tensed and relaxed. For we assert that the voice has a natural way of moving, and does not place intervals haphazardly. We try to give these matters demonstrations which conform to the appearances, not in the manner of our predecessors, some of whom used arguments quite extraneous to the subject, dismissing perception as inaccurate and inventing theoretical explanations, and saying that it is in ratios of numbers and relative speeds that the high and the low come about. Their accounts are altogether extraneous, and totally in conflict with the appearances. Others delivered oracular utterances on individual topics, without giving explanations of demonstrations, and without even properly enumerating the perceptual data. We, on the other hand, try to adopt initial principles which are all evident to anyone experienced in music, and to demonstrate what follows from them.[11]

This account of harmonics reveals Aristoxenus to be in conflict not only with the Pythagoreans (the unnamed antagonists who dismiss perception and explain pitch relations through ratios) but also with earlier harmonic theorists whose empirical work he found deficient because they did not explain their methods of proof or properly describe their observations.

Aristoxenus's alternative was to apply Aristotle's intellectual method to music more rigorously than did Aristotle himself.[12] This entailed restricting the account of music to terms and concepts that could properly be said to belong to the domain of music. It excluded descriptions that made recourse to ratios (which are in the domain of number) or to theories about the propagation of physical sound (which

11. Aristoxenus, *Elementa harmonica*, in *Harmonic and Acoustic Theory* (vol. 2 of *Greek Musical Writings*), ed. and trans. Andrew Barker, Cambridge Readings in the Literature of Music (Cambridge: Cambridge University Press, 1989), 149–50.

12. For commentary on Aristoxenus's approach, see Barker, *Greek Musical Writings*, 2: 66–69, 119.

are in the domain of physics). Once the definition of these basic musical concepts was accomplished, an explanation of the entire domain of music could then follow.

Aristoxenus's demonstration proceeds in three steps. First, he identifies two forms of vocal motion, the continuous and the intervallic. In the continuous form, which is associated with speech, the voice appears to traverse space without stopping, until the point of silence. In the intervallic form, which is associated with singing, the voice appears to stand still at specific points, and then pass over some interval of space before coming to rest at another point.[13] The various pitches upon which the voice pauses when singing constitute musical notes; the spaces between these notes are musical intervals.

Aristoxenus's second step is to draw distinctions among the various musical intervals. The first distinction is made with regard to magnitude, which reflects the amount of space between the two notes of the interval. That such a space exists is inferred from the difference between the two pitches that adjoin the interval; the size of the space can be reckoned in terms of how many other notes could be put inside it.[14] The second distinction is made with regard to concord and discord. Aristoxenus identifies the concordant intervals as the fourth, fifth, and octave (and their octave duplications). These are the only concordant intervals he accepts as determined by the intrinsic nature of melody—all other intervals are by definition discordant.[15]

The final step toward assembling the basic definitions and principles of harmonics is the derivation and division of the tone. Aristoxenus defines the tone as the difference between the first two concords (the fourth and the fifth) and explains that it can be divided in half (yielding the semitone), in thirds (yielding the least chromatic diesis), or in fourths (yielding the least enharmonic diesis, which is the smallest interval recognized as melodic).[16] These distinctions allow him to locate various notes within the tetrachords of the different genera and thereby to specify the scalar structure of each.

With these steps, Aristoxenus lays out the basic materials for his account of musical organization. Singing involves a specific way of using the voice that creates musical notes and musical intervals. Intervals can be distinguished according to size and whether they are concordant or discordant. Concordant intervals—the fourth, fifth, and octave—are accepted as axiomatic to melody and thus representative of the basic materials from which the various intervals of the Greater Perfect System can be developed.

13. Aristoxenus, *Elementa harmonica*, 133.

14. Aristoxenus, *Elementa harmonica*, 136.

15. Aristoxenus, *Elementa harmonica*, 139. Aristoxenus further distinguishes between composite and incomposite intervals (p. 137); however, this distinction is not necessary for a basic understanding of his theory.

16. Aristoxenus, *Elementa harmonica*, 140. The different *dieses* apply to the three genera mentioned above: the half-tone is used in the diatonic genus; the third-tone is used in the chromatic genus; and the fourth-tone is used in the enharmonic genus. For further discussion of Aristoxenus's *Elementa harmonica*, see Mathiesen, *Apollo's Lyre*, 319–34.

## Greek Theory and Cognitive Structure

As mentioned, the music theories of Pythagoras and Aristoxenus belong to a world remote from our own. Not only did these theorists have to grapple with the most basic of principles, but also the music they would describe is a microtonal one that is primarily concerned with the successive notes of melody rather than the simultaneous notes of harmony. Despite this—or perhaps because of it—Pythagorean and Aristoxenian accounts of musical organization give us a glimpse into how theories are formed and, more important, the cognitive processes that are basic to these theories. In particular, three cognitive processes can be seen at work: categorization, cross-domain mapping, and the use of conceptual models.

CATEGORIZATION Our ability to categorize things is a cognitive process so basic and so pervasive that it can easily escape our notice. Were you to lift your eyes from this book and survey your surroundings, you might well see chairs, lamps, tables, and other books; were you outside, you might see trees, birds, clouds, cars, and bicycles. If you considered the other things that populate your day, you might think of friends and family members, facial expressions and gestures, actions and activities. Your recognition of these things reflects the categories through which we structure our thought: to recognize a book is to identify it as a member of the category *book*; to recognize a tree is to identify it as a member of the category *tree*. Categorization occurs in all sensory modalities and throughout the range of mental activities: we categorize smells and sounds, thoughts and emotions, skin sensations and physical movement.

Categories are not just basic to thought; they also give insight into our thought processes. At one time it was thought that categories reflected the structure of the real world, but recent research has shown that the categories humans use are shaped by their interactions with their environments. Our reasons for developing and employing a given category are part and parcel of the category itself: categories are not only *not* given by nature, but also they are subject to change and modification as our thought unfolds.

Two categories basic to Pythagorean and Aristoxenian music theory are those for consonant or dissonant intervals. Consonant intervals (such as the octave, fifth, and fourth) are fundamental to the conceptualization of Greek music: they mark the stable pitches of the Greater Perfect System and are the source of derivation for all further intervals, both consonant and dissonant. The process of categorization is also exhaustive: any interval that can be conceived belongs to one of these two categories.

This is not to say, however, that consonant and dissonant intervals are given by nature in any simple way, Pythagoras and the blacksmiths notwithstanding.[17] As one

17. I should note here that psychoacousticians distinguish between musical consonance, which is a cultural construct framed relative to a particular set of musical practices, and sensory consonance, which is a consequence of how sound waves are processed by the hearing mechanism (which involves the cochlea and the auditory cortex). Sensory consonance is thus a fairly straightforward product of nature. Although musical consonance has its basis in sensory consonance, there is some freedom in how the sensory data are interpreted.

example, consider the way Aristoxenians and Pythagoreans classified the interval of an octave plus a fourth. Aristoxenians considered the interval a consonance, since it was simply the combination of two smaller consonances. Pythagoreans, in contrast, classified intervals according to the numerical ratio formed by their constituent pitches. As explained by the anonymous (and thoroughly Pythagorean) author of the *Sectio canonis* (fourth century B.C.), consonant intervals are those whose ratios are either multiple (of the form $[mn]:n$) or epimoric (of the form $[n + 1]:n$). Dissonant intervals are those whose ratios are epimeric (of the form $[n + m]:n$, where $m$ is greater than 1 and neither equal to nor a multiple of $n$).[18] Because the octave plus a fourth had the epimeric ratio 8:3, it was regarded as a dissonance.

Another example of how categories shape our understanding of phenomena is provided by Greek theorists' treatment of thirds and sixths. Although thirds and sixths sound fairly consonant, they were nonetheless categorized as discords. Two factors bear on this classification. First, forming thirds and sixths *requires* using the movable pitches of the Greater Perfect System—at best, a third or a sixth will involve only one of the stable pitches bounding the constituent tetrachords of the system. Thirds and sixths were intervals that necessarily varied in size, and so they were placed among the dissonances. Second, in the classification of intervals Greek theory followed a tradition of dichotomous categories: there was concord, discord, and nothing else. By contrast, neither of these factors played a part in the music theory of early India. Indian music theorists were consequently free to focus on the qualitative aspect of intervals rather than on their correspondence with the fixed notes of a tuning system and to construe intervallic relationships as concordant, discordant, or neutral.[19]

These two examples show that while the categories for consonant and dissonant intervals may be basic to Pythagorean and Aristoxenian theory, just how they are defined reflects the context and goals of categorization: consonance and dissonance are not naturally occurring properties, but ways of constructing an understanding of musical organization.

Of course, there are numerous other categories important for Pythagorean and Aristoxenian music theory, including those for pitches, intervals, and numerical ratios. These categories and others are basic to the sort of systematic account of musical phenomena provided by these theories—indeed, it is simply not possible to have a theory of music, or of anything else, without first having categories.

CROSS-DOMAIN MAPPING Cross-domain mapping is a process through which we structure our understanding of one domain (which is typically unfamiliar or abstract) in terms of another (which is most often familiar and concrete). For example, one way to think about the elusive concepts of electrical conductance is in terms of a hydraulic model: flipping the light switch turns on the *juice*, and elec-

18. *Sectio canonis*, in *Harmonic and Acoustic Theory* (vol. 2 of *Greek Musical Writings*), ed. Andrew Barker, Cambridge Readings in the Literature of Music (Cambridge: Cambridge University Press, 1989), 193.

19. Lewis Rowell, *Music and Musical Thought in Early India*, Chicago Studies in Ethnomusicology (Chicago: University of Chicago Press, 1992), 157–60.

trical current *flows* to the light bulb to light the room. By this means we take what we know about a fairly concrete and familiar source domain—the flow of water and other liquids—and map it onto a rather abstract and unfamiliar target domain: that of electricity. As a wealth of research on analogy and metaphor has shown, the process of mapping structure from one domain to another is basic to human understanding.

One place cross-domain mapping is evident is in the Pythagorean and Aristoxenian construal of interval. Because musical pitches are ephemeral and virtually intangible, relationships between pitches—musical intervals—represent something of a challenge to understanding. One way to meet this challenge is to map structure from the physical world onto music, a process evident in Nicomachus's story of Pythagoras and the blacksmiths. Pythagoras hears harmonious sounds, traces their origins to the blacksmiths' hammers, and then proceeds to conduct various experiments using weights equivalent to those of the hammers. These experiments lead, among other things, to a highly pragmatic objectification of musical pitch, as pitches are translated into physical objects that can be weighed, studied, and preserved. By performing a mapping from the concrete physical domain proper to the blacksmiths' hammers onto the domain of musical sound, Nicomachus's story allows us to structure the latter domain in terms of the former. Of course, musical notes are not physical objects that can be weighed, studied, and preserved—they remain ephemeral and virtually intangible. Nonetheless, we are so accustomed to the mapping between concrete physical objects and musical sound that we sometimes have to be reminded that notes are not enduring physical objects.

Aristoxenus's construal of musical interval involves a slightly different mapping. As we have seen, according to Aristoxenus, when the voice moves intervallically, it appears to stand still at a given place (a musical pitch) and then pass over an interval of space (a musical interval) before coming to rest at another place (another musical pitch). Underlying this account is a mapping from the familiar domain of two-dimensional space onto that of music. This mapping allows us to apply the methodology of measuring space to music. The difference between two linear measures yields a third measure; similarly, the difference between the intervals of a fifth and a fourth yields the interval of a tone. Since linear measures can be easily divided into equal halves or thirds or fourths, the musical tone can be similarly divided, something impossible from the Pythagorean perspective.

On closer inspection, the Pythagorean and Aristoxenian construals of interval are indeed incommensurate. From the Pythagorean perspective, pitches are physical objects, and an interval describes the relationship between these objects. From the Aristoxenian perspective, pitches are breadthless points that simply mark out an expanse of two-dimensional space, and an interval is the expanse itself. Each mapping gives an account of interval, but each leads to a different conceptualization of musical structure. This point can be generalized for music theory as a whole: mapping structure from a nonmusical domain onto music is a way of *creating* musical structure, and different mappings will lead to different accounts of musical structure.

CONCEPTUAL MODELS Both categorization and cross-domain mapping provide the basis for fundamental ontological assertions about musical materials: this

interval *is* a consonance; the pitches of an octave *are* physical objects. They can also lead to conditional statements: *if* the interval is an octave, *then* it is a consonance; *if* a pitch is an object, *then* its properties are measurable. Propositions like this are basic to conceptual models, which act as guides to reasoning and inference. In their simplest form, conceptual models consist of concepts in specified relationships, which pertain to a specific domain of knowledge.

For an example of a conceptual model, let us return to the classification of consonant and dissonant intervals presented in the *Sectio canonis*, according to which all consonant intervals have either multiple or epimoric ratios, and all dissonant intervals have epimeric ratios. This classificatory system relies on a conceptual model that organizes concepts related to interval, concord, discord, and the three classes of ratios. The simple pattern of inference that follows from this model is that if an interval has a multiple or an epimoric ratio, it is a concord; if it has an epimeric ratio, it is a discord.

Integral to this model are the products of categorization and cross-domain mapping. Two types of categories are involved in the model: those pertaining to music (the categories of concord and discord) and those pertaining to number (the multiple, epimoric, and epimeric ratios). Cross-domain mapping correlates the two types of categories by construing musical interval as a relationship between two objects (namely, musical pitches) to which magnitudes (in the form of numbers) can be assigned. Specific classes of ratios can then be used to distinguish between the musical categories.

The robustness of this particular conceptual model is reflected in the debate over the status of the octave plus a fourth that continued into the Middle Ages. In the second century A.D., Ptolemy showed the speciousness of the correlation of concord with multiple or epimoric ratios and argued for a classification of intervals based on empirical evaluation and the postulate that a concord added to a concord produces a concord.[20] Although Ptolemy still used ratios to describe various intervals, they were no longer part of the conceptual model through which intervals were classified into concords and discords. In the sixth century, Boethius presented both the Pythagorean and Ptolemaic models but took no position on which he preferred.[21] After Boethius, when an author wished to invoke the authority of the Pythagorean approach, the Pythagorean model of intervallic classification was cited; when an author wished for a more empirically satisfying classification, the Ptolemaic model was used.[22]

Conceptual models provide the first level of organization for concepts. They are too limited and localized, however, to provide the comprehensiveness we expect from theories of music. Theories achieve this comprehensiveness by integrating

20. Ptolemy, *Harmonics*, in *Harmonic and Acoustic Theory* (vol. 2 of *Greek Musical Writings*), ed. Andrew Barker, Cambridge Readings in the Literature of Music (Cambridge: Cambridge University Press, 1989), 286–90.

21. Boethius, *Fundamentals of Music*, 81–82, 169.

22. For a discussion of these modes of reasoning, see C. André Barbera, "The Consonant Eleventh and the Expansion of the Musical Tetractys: A Study of Ancient Pythagoreanism," *Journal of Music Theory* 28 (1984): 191–223. Barbera's treatment of Aurelian's *Musica disciplina* on p. 210 is especially illuminating.

clusters of conceptual models. And as we shall see in the following chapters, conceptual models also play a role in categorization and cross-domain mapping. There, as in theories, they provide guides for reasoning and inference about specific and circumscribed domains of knowledge.

COGNITIVE PROCESSES AND MUSIC THEORY Much has been left out of this discussion of theories of music from Greek antiquity, with respect to both the theories themselves and the cognitive processes behind them. To be sure, these theories are of a different order than Swann's musings on Vinteuil's sonata. Nonetheless, the cognitive processes we have seen at work in Pythagorean and Aristoxenian theory are the same processes through which we organize our understanding of the world as a whole. Just how this occurs—how categorization, cross-domain mapping, and the use of conceptual models shape our theories of music and guide our analyses of musical works—is the subject of the remainder of this book.

## OVERVIEW

I have divided the chapters that follow into two parts. In the first, I present a detailed overview of research on the three cognitive processes highlighted in this introduction. This overview is itself framed around specific musical topics, such as motivic transformation, text painting, and the ways in which we structure our understanding of a specific musical domain. The research that has been done in cognitive science over the past three decades has been extensive and far ranging, and one of the jobs of this portion of the book is to bring this work to bear on basic issues of musical understanding. Another objective is to show in some detail how these processes relate to one another and how they form the bedrock for our thought about music. The second part of the book moves from this foundation to analytical studies of specific musical issues. These issues include relationships between categorization and musical syntax, the problem of musical ontology, text-music relations, and conceptions of musical form and musical hierarchy.

Chapter 1 begins the overview of research in cognitive science with a close look at processes of categorization. For centuries, writers in the West regarded categories as fixed and immutable, and any variation in categorization was taken as evidence of the failure of the human intellect to deal with the structure of the real world. It took the pioneering work of Eleanor Rosch and others in the 1970s to show that category structure was not as simple as first believed. In particular, some levels of categorization are preferred over others, and some members of a category are regarded as better representing the category than others (a phenomenon known as *graded membership*).

The key to how this research can be applied to music is provided by the musical motive (or, as Proust would have it, the *motif*). Motives are generally reckoned to be one of the basic building blocks of musical works, but they are also a bit slippery: the "same" motive typically assumes a number of diverse shapes over the course of a work. Thinking of a motive as a cognitive category makes it possible to account for its identity, as well as its diversity, and reveals how aspects of categorization are embodied by musical materials. These preliminary applications of categorization to

music also show ways musical materials can be organized over the course of a work and offer an explanation of how it is possible to have musical concepts that are independent of language.

If categorization can be said to be the source of musical concepts, cross-domain mapping is the means by which these concepts are placed in correlation with others. Chapter 2 examines the process of cross-domain mapping in some detail, beginning with the work of cognitive linguists who, in the 1980s, proposed that metaphor was a basic structure of understanding. This proposal gained added weight when it was shown that metaphorical projection (which is one way to accomplish cross-domain mappings) was a general process not restricted to linguistic expressions but grounded in embodied experience.

One example of cross-domain mapping that involves music in a rather immediate way is the technique of text painting, a compositional device that aims to represent in music specific images summoned by the text of a vocal work. Text painting provides a point of departure for the exploration of how cross-domain mapping is manifested in our understanding of music, as it leads to an extension of cross-domain mapping called conceptual blending. In a conceptual blend, elements from two correlated domains are projected into a third, giving rise to a rich set of possibilities for the imagination. As I show in the latter part of chapter 2, text painting can lead to such blends, as can program music.

Chapter 3, which focuses on conceptual models and theories, gets to the heart of the perspective on cognition developed in part I. My point of departure is research by Jeanne Bamberger on children's representations of musical structure. In my analysis of Bamberger's study of one specific eight-year-old boy, I show the part played by categorization and cross-domain mapping in the conceptual models used by this boy to come to terms with a musical environment. I also show how these models are combined to form a theory of music and how this theory changes in response to changes in the task at hand. This close-up glimpse of the structure and role of conceptual models and theories leads, in the middle of the chapter, to a more generalized characterization of these knowledge structures, which I connect with work on similar structures in artificial intelligence, cognitive anthropology, ethnomusicology, and developmental psychology. In the latter part of the chapter, I return to music theory and explore the role of conceptual models and theories (that is, theories framed relative to a cognitive perspective) in analyses by Jean-Philippe Rameau and Heinrich Schenker, two of the best-known music theorists of the last three hundred years.

Although the features of cognitive structure discussed in part I might seem to be relatively detailed, in truth all are associated with relatively high-level cognitive processes. My reason for focusing on this level is quite simple: it allows me to engage in issues of immediate and occasionally central importance to music scholarship and to do so in a way that connects with extensive research in cognitive psychology and cognitive linguistics. Part II explores this possibility in greater depth by considering various problems of musical understanding from the perspective on cognitive structure and music theory developed in part I.

In chapter 4, I turn to the matter of how musical materials are organized within a work—more properly, the problem of musical syntax and, by extension, musical

semiotics. Although semioticians are usually quick to grant that music has a syntax, they are more doubtful about whether its semantic level has any depth. By taking a close look at how composers make use of categories of musical events—in this case, the way Mozart and Beethoven use motives in the opening movements of three string quartets—I am able to provide insight into how musical materials are organized in the service of musical discourse, as well as how features of this organization contribute to meaning construction as a whole.

In chapter 5, I confront a somewhat larger problem—one that may seem irredeemably abstract: the problem of what counts as a work of music. I view this problem, usually called the problem of musical ontology, as one of cultural knowledge and try to show that, as opposed to being hopelessly recondite, the problem is of immediate importance for understanding music. By approaching the entire work of music as a category—a category that includes all the scores, performances, representations, and such that are said to be "of" the piece—I develop a model for the cultural knowledge upon which judgments about musical ontology are made. Determinations of what counts as an instance of a particular musical work are thus one of the ways members of a musical community construct and negotiate their identity. My examples for this chapter are two songs taken from the traditions of popular music and jazz: "I Got Rhythm" and "Bye Bye Blackbird." The latter offers an intriguing case of how the cultural knowledge relative to which determinations of musical ontology are made can become complicated when implicated in the layered discourse structures Mikhail Bakhtin called "double-voiced discourse," and which were extended to African American culture through Henry Louis Gates Jr.'s notion of "Signifyin(g)."

Chapter 6 returns to the analysis of individual musical works by pursuing one of the entailments of text painting noted in chapter 2: under certain circumstances, combinations of words and music, through the process of conceptual blending, create worlds for the imagination well beyond those that spring from words or music alone. Where only a few fairly circumscribed instances of text-music relations were considered in chapter 2, here research on conceptual blending is applied to the whole of five Lieder from the nineteenth century. These analyses offer a way to flesh out the theory of conceptual blending as it applies to music and provide a further perspective on musical syntax. In these songs, we see words and music combining to create rich domains in which the imagination can play, as well as discover how musical syntax shapes our understandings of the words themselves.

The final analytical chapter (chap. 7) turns to music theory itself, specifically to the theories of musical form and hierarchy that go back to the eighteenth and early nineteenth centuries. Accounts of the form a musical work takes, or of how its elements relate to one another, are basic to theorizing about music—indeed, we can see these emerging in the course of M. Swann's ruminations on Vinteuil's sonata—but at times it seems that theorists are talking about quite different things. For instance, there are two common ways to talk about musical form: the first approaches form as deriving from the assembly of relatively static building blocks that are combined to create the finished work; the second approaches form as an emergent property of the work, which becomes manifest only as the music unfolds over time. The first approach yields a view of musical form that is quite static, the second a view rather

more dynamic. Using the analytical framework provided by cross-domain mapping, I discuss the source of these two models for musical form, as well as two models for musical hierarchy, and explore some of the ways these models interacted over the course of the history of music theory.

In the conclusion I return to M. Swann and to his final encounter with Vinteuil's sonata after a year in which it became thoroughly intertwined with his love affair with Odette, the courtesan with whom he had become acquainted around the same time he first encountered the andante. This will provide a frame for a review of the points made in the preceding chapters and an instrumentality for drawing conclusions from the whole.

## COGNITIVE STRUCTURE, THEORY, AND ANALYSIS

A central claim of this book is that through developing an appreciation of how aspects of cognitive structure shape our understanding of music we can better appreciate the active role of theories of music in that understanding. A further claim is that our analyses of musical phenomena—from the most mundane and localized of accounts to the most abstract and comprehensive—similarly reflect cognitive structure, in that every analysis is based on some sort of theory of music. Musical analyses are in truth dialogues, and not just dialogues between the analyst and an imagined audience: musical analyses are also dialogues between the analyst and some body of theoretical knowledge. Analysis rarely, if ever, simply corroborates a theory: analysis pulls theory and pushes it, extending and changing theory just as it also extends and changes our understanding of musical phenomena.

The analyses I present throughout this book are no different, except that they engage cognitive theory as well as music theory. The intent of the analyses is to show how our understanding of particular musical phenomena can be characterized in terms of specific cognitive processes and structures and thereby connect that understanding to research in cognitive science as a whole. The analyses are not intended as definitive statements about how we can account for such understanding; they are intended to be the initiation of a dialogue with cognitive theory, a dialogue whose purpose is to expand our knowledge of both music and cognition. Analysis is thus a central concern of what follows, but, unlike cognitive structure or theory, I have not treated it as a central topic for investigation. Instead, analysis will be a fundamental tool to explore both and to provide new insight into the conceptual worlds wrought by musical sound.

PART I

# ASPECTS OF COGNITIVE STRUCTURE

CHAPTER ONE

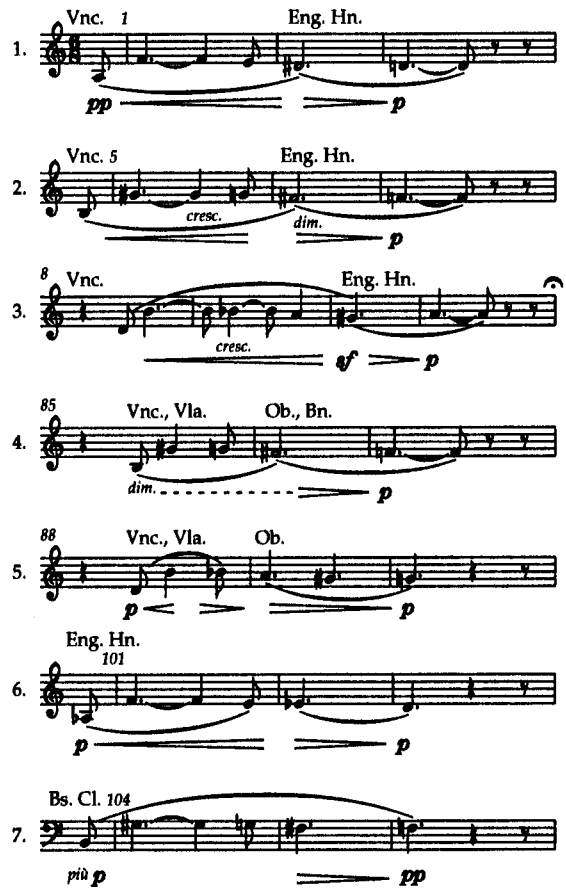
# CATEGORIZATION

In the opening eleven measures of the prelude to Richard Wagner's *Tristan und Isolde* there are three successive statements of a melodic motive that foreshadows, and then later accompanies, the tragic arc of the opera. Three times this motive emerges from silence, swells through the cellos and English horn, and then sinks back into the stillness from which it came. After the third statement, the motive disappears, then reappears a few long minutes later near the end of the prelude. Here it is stated first in the cellos and violas (and completed by the oboes and bassoons), then in the English horn, and finally in the bass clarinet. After the last of these, the prelude dissolves into the first scene of the opera. All of these statements are shown in example 1.1.

Although our first inclination is to treat these various statements as somehow the same—that is, all can be regarded as "statements of the opening motive from Wagner's *Tristan und Isolde*"—there are actually notable differences among them. There is, of course, the change in the instrumentation used for the statements at the end of the prelude, which contrasts with the distinctive timbral stamp put on the motive in the opening measures. There is a similar diversity in the pitches through which successive statements are given voice, since versions of the motive begin on either A3 (or A♭3), B3 (or B2), or D4. Although there are significant similarities in the rhythmic figuration used for the various statements, there is no pattern of duration and accentuation common to all. And there are also somewhat smaller variations: the first statement begins with a minor sixth instead of the major sixth used by the others, and the fourth is the only one shaped by both a diminuendo and a rallentando.

Despite their differences, these seven melodic fragments sound similar enough to one another to be regarded as functionally equivalent. Such equivalence was important to Wagner's compositional style, for it provided the material foundation for works uninterrupted by conventional forms. By recalling specific motives at crucial points within his musical and dramatic discourse Wagner was able to create a sense of shape and unity that seemed inevitable without being predictable. By the time of *Tristan*, however, such motives were no longer simple building blocks to be assembled and permuted but had become infinitely malleable constructs. Carl Dahlhaus, recognizing both the structural importance and mutability of these

EXAMPLE 1.1 Statements of the opening motive from Richard Wagner's *Tristan und Isolde*, Prelude




motives within the opera, likened them to woven threads that surface only to divide and then disappear from sight.[1]

To understand this music—to make sense of the sonic texture Wagner weaves—requires being able to assimilate these various musical phrases into a single cognitive construct and then recall that construct, often after an hour or more of Wagnerian effusion. Understanding Wagner—or most music, for that matter—requires being able to think in terms of categories of musical events. In this chapter, I explore what recent research into categorization can tell us about this process. First, however, I develop a slightly clearer picture of two basic aspects of understanding music, which will provide the groundwork for applying research on categorization to situations of the sort presented by the opening of Tristan. I do this by reviewing a theory of musical motive developed by one of the many students of Wagner. Not that Wag-

1. Carl Dahlhaus, *Richard Wagner's Music Dramas*, trans. Mary Whitall (Cambridge: Cambridge University Press, 1979), 63.

ner ever actually taught composition. But the influence of his music was enormous, and many lessons were taken away from his operas. Among these were lessons on compositional technique absorbed, refined, and retransmitted—in both treatise and score—by Arnold Schoenberg.

## MUSICAL COHERENCE AND MOTIVE

Schoenberg thought of himself first and foremost as a Brahmsian, but, under the influence of Alexander von Zemlinsky, he also found a passion for Wagner. Schoenberg noted that his early works showed the influence of both masters: "In my *Verklärte Nacht* the thematic construction is based on Wagnerian 'model and sequence' above a roving harmony on the one hand, and on Brahms's technique of developing variation—as I call it—on the other."[2] Wagner was also known for his harmonic innovation, although Schoenberg took pains in one of his late essays to show that Brahms's harmonies were every bit as "progressive" as Wagner's.[3] Nonetheless, it was Wagner's harmonic language that was perceived as posing the most formidable challenge to composers of the late nineteenth century, for the free treatment of dissonance integral to this language loosened prevailing conceptions of tonal organization and called into question the "laws of tonality" that had formerly held sway.

From Schoenberg's perspective, Wagner's music presented few dangers for tonality. As Schoenberg had argued in his *Harmonielehre* of 1911, tonality had its origins in acoustic phenomena.[4] But shaping music in conformance with acoustic phenomena was not all there was to musical composition: composition also required insight into how the human mind worked. In an essay that attempted to sort out some of the problems raised by the "emancipation of the dissonance" Schoenberg wrote:

> Tonality's origin is found—and rightly so—in the laws of sound. But there are other laws that music obeys, apart from these and the laws that result from the combination of time and sound: namely, those governing the working of our minds. This latter forces us to find a particular kind of layout for those elements that make for cohesion—and to make them come to the fore, often enough and with enough plasticity—so that in the small amount of time granted us by the flow of the events, we can recognize the [musical] figures, grasp the way they hang together, and comprehend their meaning.[5]

In Schoenberg's view, the laws governing the workings of our minds require the composer to write in such a way that listeners can quickly recognize musical figures and the way they cohere. The listener, upon grasping this coherence, will then be able to comprehend the work.

Key to musical coherence and comprehension were musical motives. According

2. Arnold Schoenberg, "My Evolution (1949)," in *Style and Idea: Selected Writings of Arnold Schoenberg*, ed. Leonard Stein, trans. Leo Black (New York: St. Martins Press, 1975), 80.

3. Arnold Schoenberg, "Brahms the Progressive (1947)," in *Style and Idea*, 398–441.

4. Arnold Schoenberg, *Harmonielehre* (Leipzig: Verlagseigentum der Universal-Edition, 1911); trans. as *Theory of Harmony* by Roy E. Carter (Berkeley: University of California Press, 1978), 23–25.

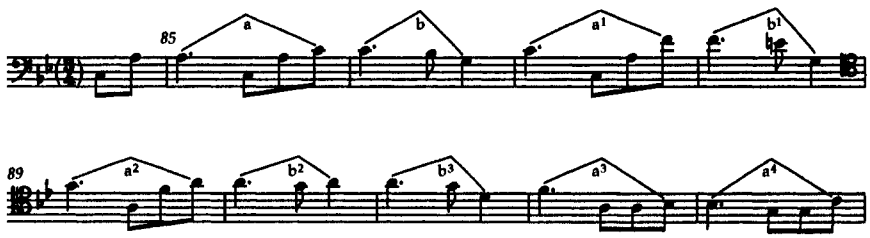
5. Arnold Schoenberg, "Opinion or Insight? (1926)," in *Style and Idea*, 259.

to Schoenberg, comprehension starts with recognition, and recognition starts with basic musical figures—that is, motives: "*Motive* is at any one time the smallest part of a piece or section of a piece that, despite change and variation, is recognizable as present throughout."[6] More specifically, motives consist of intervals and rhythmic patterns combined to produce a shape or contour that, once recognized, can be easily remembered.[7] A particularly striking illustration of Schoenberg's concept of motive is provided by his analysis of a subordinate theme from the first movement of Brahms's Sextet in B♭ major, Op. 18 (1860), shown in example 1.2.[8]

As Patricia Carpenter and Severine Neff have noted, rhythm is central to Schoenberg's concept of motive.[9] This centrality is clearly evident in his analysis of the Brahms sextet. Each family of motives is distinguished by a specific rhythmic pattern: all of the **a** motives use a ♩·♪♩♩ pattern; and all of the **b** motives use a ♩·♪♩ pattern. (Both patterns, of course, contribute to the marvelous waltz-like feel of the theme.) Contour also serves to distinguish the two families, although here some variation among the different motive forms can be noted. Motives $\mathbf{a}^1$, $\mathbf{a}^2$, and $\mathbf{a}^4$ all preserve the basic contour of motive **a**: the second note of the motive is lower than the first note, and the last note of the motive is higher than the first note. Motive $\mathbf{a}^3$ preserves only a portion of the contour of **a**: although the second note of the motive is lower than the first, so is the last note. In a similar fashion, motives $\mathbf{b}^1$ and $\mathbf{b}^3$ preserve the basic contour of motive **b** (the second note is lower than the first note, and the third note is lower than the second), while $\mathbf{b}^2$ preserves only a portion of the contour of **b** (the second note is lower than the first note, but the third note is higher than the second).

As this example shows, Schoenberg's concept of motive is broad and dynamic: a family of motives can be distinguished simply by rhythmic figure or contour; motive forms can be various and need have only some of their features in common. This conception is rather different from, for example, Heinrich Schenker's or Rudolph

EXAMPLE 1.2 Schoenberg's analysis of the Brahms Sextet, Op. 18, first movement, mm. 84–93




6. Arnold Schoenberg, *The Musical Idea and the Logic, Technique, and Art of Its Presentation*, ed. and trans. Patricia Carpenter and Severine Neff (New York: Columbia University Press, 1995), 169.

7. Arnold Schoenberg, *Fundamentals of Musical Composition*, ed. Gerald Strang and Leonard Stein (London: Faber and Faber, 1967), 8.

8. Schoenberg, "Brahms the Progressive," 417.

9. See the commentary in their edition of Schoenberg's *The Musical Idea and the Logic, Technique, and Art of Its Presentation*, 27–29.

Réti's concept of motive. Both Schenker and Réti regard specific intervallic relationships as constituent of motive; rhythm and contour are regarded as secondary aspects of motivic organization.[10] The difference between Schenker's or Réti's conception of motive and that of Schoenberg lies with Schoenberg's notion of coherence and the role it plays in making music comprehensible. *Coherence* comes about when the various parts that make up a musical entity are connected in such a way that those parts similar to other entities become prominent. The work is most *comprehensible* to the listener when the arrangement of these parts is such that their relationship to each other and to the whole is manifest.[11] Thus there is no need to restrict the distinguishing features of a family of motives to specific intervallic relationships, nor is there a need to require that all motive forms share exactly the same features. All that is required of motive is recognizability and a potential for connection to other motive forms. This potential for connection contributes to coherence, which is, in turn, the basis for comprehensibility.

According to Schoenberg, then, the process of comprehension starts with recognizable bits (motives) that are easily remembered. Motives hang together not simply because their constituent parts are connected to one another but because these connections emphasize similarities to other motives. Coherence thus reflects properties shared by collections of motives; it is not, properly speaking, a property of any one individual motive. Motive forms are *of necessity* variable, for differences between forms reveal most clearly what is typical of the collection of motives as a whole. And, although attention to coherence is important to the composer who wishes to craft a convincing work, the apprehension of coherence is essential to the listener who would make sense of that work.

Many composers, working in the long shadow cast by Ludwig van Beethoven across the nineteenth century, relied on musical motives to make their works coherent. As Edward Cone has noted, the use of thematic transformation as a unifying device became increasingly important for composers of the later nineteenth century, composers whose lengthy movements often embraced stylistic extremes of tempo, meter, texture, and mood.[12] However, the prevalence of the device also poses problems for the listener, for it is often difficult to sort out relationships among various motivic entities. Consider the relationship of derivation, in which later motive forms are derived from earlier motive forms; thus motive $\mathbf{b}^3$ in example 1.2 is said to be derived from motive **b**. As straightforward as this might appear, tracing derivational relationships is rarely a simple task: motive **b** can itself be understood to be derived from one of the motives prevalent in the preliminary subordinate theme heard starting in m. 61 of Brahms's sextet (shown in ex. 1.3), which has the same rhythmic pattern but the reverse contour. Could $\mathbf{b}^3$ then be regarded as derived directly from this earlier motive form, or is it still to be regarded as derived from **b**?

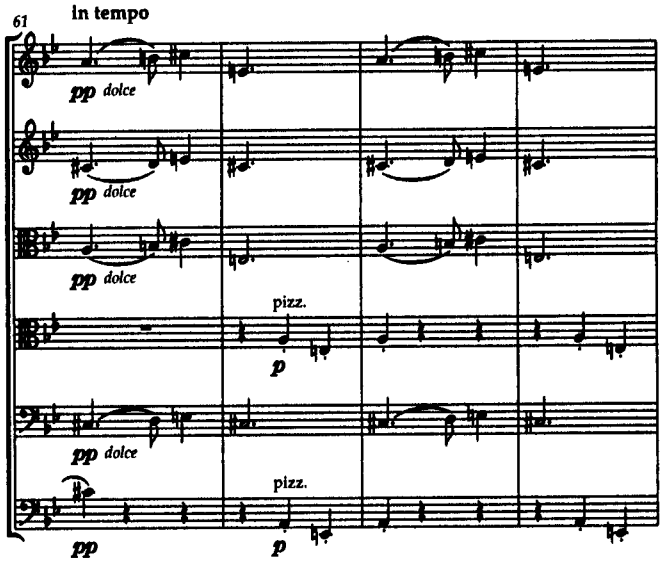
Schubert's *Wanderer Fantasie*, Op. 15 of 1822 presents an even more complicated

10. Heinrich Schenker, *Der freie Satz*, Neue musikalische Theorien und Phantasien, 3 (Vienna: Universal-Edition A.G., 1935); Rudolph Réti, *The Thematic Process in Music* (New York: Macmillan, 1951). I discuss historical construals of motive below in the first part of chap. 7.

11. Arnold Schoenberg, *Coherence, Counterpoint, Instrumentation, Instruction in Form*, ed. Severine Neff, trans. Severine Neff and Charlotte M. Cross (Lincoln: University of Nebraska Press, 1994), 20–23.

12. Edward T. Cone, "On Derivation: Syntax and Rhetoric," *Music Analysis* 6 (1987): 238.

EXAMPLE 1.3 Johannes Brahms, Sextet, Op. 18, first movement, mm. 61–64

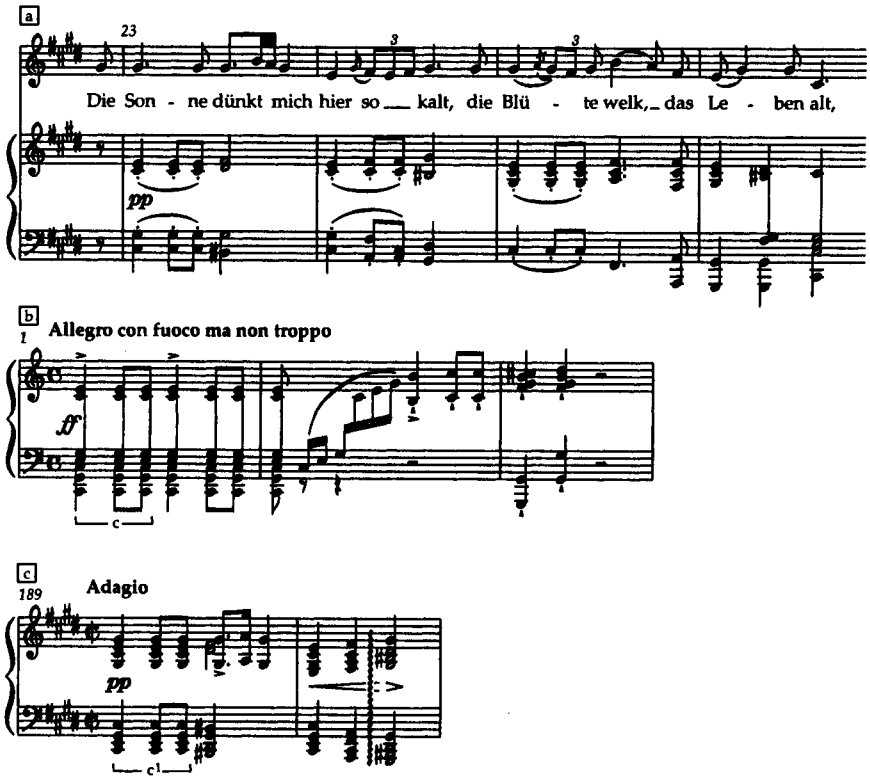



situation. The work is based on an arietta from Schubert's 1816 cantata "Der Wanderer," Op. 4/1, D. 489, which is quoted in the work's second movement Adagio; the relevant passage from the cantata appears in example 1.4a. If we have the arietta in mind, we can hear the motivic material of the first movement as derived from the arietta and then presented instrumentally in the subsequent Adagio (the motive from the first movement is labeled c in ex. 1.4b; its use in the Adagio is demonstrated in ex. 1.4c).[13] The temporal order of the process of derivation is thus reversed: c is heard as derived from $c^1$. If we are coming to the work fresh, we will most likely hear the Adagio as derived from the first movement: $c^1$ is thus derived from c. Finally, if we know the arietta but have temporarily forgotten it, its reprise in the Adagio, anticipated by the "derived" motives of the first movement, can be a moment of great rhetorical impact: the pianistic musings of the first movement have summoned forth the arietta that now stands before us.

It would seem that, as Cone argued, hearing relationships among motive forms is an ineluctably subjective affair. Following his argument still further, how we hear these relationships is important to how we respond to a work, for, as Schubert's fantasy suggests, it informs our understanding of musical rhetoric. The most dramatic case occurs in a work in which there are multiple derivations of motivic material. In the course of such a work we can lose our sense of how motive forms are connected, until the composer—often unexpectedly—reveals how the materials relate to one another by bringing them into rapprochement. The rhetorical aim of this sort of compositional strategy is what Cone calls "epiphany," and, when successful,

13. While the annotations in ex. 1.4 focus on the distinctive rhythmic figure common to each of the excerpts, it is worth noting that each passage also features a move to the dominant.

EXAMPLE 1.4 Franz Schubert, (a) "Der Wanderer," Op. 4 No. 1, D. 489, mm. 23–26; (b) *Wandererfantasie*, D. 760, mm. 1–3; (c) *Wandererfantasie*, mm. 189–90




it can compel the listener to realize a previously unsuspected—or at most unconfirmed—relationship among diverse motives.[14] Put another way, the strategy that leads to epiphany compels the listener to realize aspects of musical coherence that become evident only over the course of time.

Motives and motivic relationships are important for both the coherence and the rhetoric of music, but understanding the part they play in musical understanding poses some unique challenges. If the recognition of motives is central to the process of comprehending music, it must occur very rapidly. This rapidity has led some to argue for cognitive modules dedicated to the processing of such basic musical materials.[15] The theory of cognitive modularity, an adaptation of faculty psychology first proposed by the philosopher Jerry Fodor, bears some resemblance to the notion of competence discussed briefly in my introduction.[16] The coherence of musical

14. Cone, "On Derivation," 246.

15. See, for example, David Temperley, "Motivic Perception and Modularity," *Music Perception* 13 (1995): 141–69; and Eugene Narmour, *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model* (Chicago: University of Chicago Press, 1990).

16. Jerry A. Fodor, *The Modularity of Mind: An Essay on Faculty Psychology* (Cambridge, Mass.: MIT Press, 1983). According to the theory Fodor proposed in *The Modularity of Mind*, cognitive structure is organized into modules dedicated to specific processing tasks. There are thus modules for such things

materials, however, is an emergent property that is less well explained from a modular perspective since it involves the active comparison and evaluation of a number of motive forms.

One model that can rise to these challenges is offered by recent research into categorization, which has demonstrated that categorization is both extremely rapid and evaluative. From the perspective provided by this research, the coherence of music reflects our ability to group musical events into categories: the **a** and **b** motives of example 1.2 each constitutes a cognitive category, with five members in one category (**a**, $\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$, and $\mathbf{a}^4$) and four in the other (**b**, $\mathbf{b}^1$, $\mathbf{b}^2$, and $\mathbf{b}^3$). To describe the "laws of the mind" that Schoenberg noted—laws that force the composer to arrange musical elements so that they cohere—is to describe the process of categorization. In what follows, I shall review research on processes of categorization and show how it can be used to account for the role of motive as a starting place for higher-level cognitive processes and for relationships among diverse motive forms.

## PROCESSES OF CATEGORIZATION

For the most part, the categories through which we structure our understanding of the world are not cause for a great deal of reflection. To all appearances, these categories are simply givens. A book is a book, a tree is a tree, and any equivocal cases are just a matter of insufficient knowledge. This commonsense view of categorization was the predominant view in philosophy and psychology at least since the time of Plato: categories, because they were constitutive of knowledge, were regarded as stable and universal. It was only in the 1950s that the adequacy of this view began to be called into question, most notably by Ludwig Wittgenstein in his *Philosophical Investigations*. There, in the midst of a discussion of language games, Wittgenstein paused to consider the category of "game" itself and found that category structure was far more complicated than it first appeared.[17] Research carried out in the 1960s and 1970s showed that numerous aspects of the way humans categorize could not be reconciled with the traditional view of categorization, and by the 1980s a new view of categorization had begun to develop. The categories through which humans organize their understanding of the world came to be seen as ways of *having* knowledge, rather than as reflecting what could be known.

The research crucial to developing this new view of categorization focused on two topics. The first concerns the relationship between the process of categorization and the taxonomic structures to which it gives rise. The second concerns the

---

as vision, hearing, language processing, and, presumably, the identification of musical motives. Fodor has since made arguments against modularity; see his recent *The Mind Doesn't Work That Way* (Cambridge, Mass.: MIT Press, 2000).

17. Ludwig Wittgenstein, *Philosophical Investigations*, trans. G. E. M. Anscombe (New York: Macmillan, 1953), 31–32. For summaries of the recent history of research into categorization, see Eleanor Rosch, "Categorization," in *Encyclopedia of Human Behavior*, ed. V. S. Ramachandran (San Diego: Academic Press, 1994), 1: 513–23; and George Lakoff, *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind* (Chicago: University of Chicago Press, 1987), chap. 2.

internal structure of categories—that is, how members of a category relate to one another.

Most of the categories we use in daily life can be organized into hierarchically arranged taxonomies. The cat prowling around my feet is a *calico cat*, a *housecat*, a *feline*, a *mammal*, and a *living organism*. The levels in taxonomies of this sort are related by inclusion: *feline* includes all of the different sorts of *housecats*, a category that includes all of the different sorts of *calico cats*, a category that includes my particular cat. In categorizing things, it seems logical to work from the top of the hierarchy down, moving from the most abstract level to the most concrete. I note first that the thing before me is a living organism, and as I increase my level of discrimination determine exactly what sort of organism it is. In actually categorizing things in the world around us, however, we start neither at the most specific level ("the cat prowling around my feet") nor at the most general (*living organism*) but at a level in between; this level has come to be called the *basic level*. Research on the basic level tells us where, for practical purposes, the process of categorization begins. The first section that follows reviews research on the basic level and shows how it supports Schoenberg's contention that musical comprehension begins at the level of the motive.

It is similarly a commonplace to regard each level of a taxonomy as having clear, fixed boundaries: a creature either is or is not a *housecat*. Categories are thus unchanging anchors of understanding, their structure immutable and constant. The categories we actually *use* in daily life, however, cannot be so neatly characterized. These categories have a somewhat more complicated structure: membership is by degrees; category boundaries are relatively fluid and often hazy; and category structure can change over time. In these cases, the process of categorization is a highly dynamic one, requiring that we evaluate potential category members relative to existing category members which, together, define a standard of typicality. Categories with a variable membership are thus said to show typicality effects. The second section that follows reviews research on categories that illustrate typicality effects and shows how it can be used to model the coherence and variability of motive forms basic to Schoenberg's theory of motive.

## Basic-Level Categorization

THE BASIC LEVEL The psychologist Roger Brown was perhaps the first to note that the categories we most frequently employ are not at the lowest level of a taxonomy (and concerned with individuals), nor are they at the highest level (and concerned with broad classes); instead, they are in the middle of a taxonomy, at a level of maximum utility.[18] For instance, we typically call the thing in our purse or pocket a *dime*, even though we might also call it *money* or a *metal object* (names at the superordinate level of the taxonomy), or *my favorite 1956 dime* (which would be at the subordinate level of the taxonomy).

18. Roger Brown, "How Shall a Thing Be Called?" *Psychological Review* 65 (1958): 14–21; idem, *Social Psychology* (New York: Free Press, 1965), 319–21.

TABLE 1.1. Sample Taxonomies Showing the Basic Level

| *Superordinate* | *Basic Level* | *Subordinates* |
|---|---|---|
| Musical instrument | Guitar | Folk guitar |
| | | Classical guitar |
| | Piano | Grand piano |
| | | Upright piano |
| Fruit | Apple | Delicious apple |
| | | Mackintosh apple |
| | Peach | Freestone peach |
| | | Cling peach |
| Furniture | Table | Kitchen table |
| | | Dining room table |
| | Lamp | Floor lamp |
| | | Desk lamp |

Adapted from Eleanor Rosch et al., "Basic Objects in Natural Categories," *Cognitive Psychology* 8 (1976): 388.

Subsequent research into ethnobotanical systems of classification by Brent Berlin and his associates supported the importance of mid-level classification and gave empirical evidence that this level is psychologically basic: at this level, people name things more readily, languages have simpler names, categories have greater cultural significance, things are remembered more readily, and things are perceived holistically.[19] Research directed specifically at mid-level classification was carried out by Eleanor Rosch and her associates in the early 1970s. In their report, this level was named the *basic level*, a term which has been generally adopted in the literature on categorization.[20] Examples of sample taxonomies showing the basic level are given in Table 1.1.

Rosch and her associates suggested that two contrasting principles influence the taxonomic level at which people prefer to categorize.[21] The first is the *efficiency principle*, according to which people prefer to minimize the number of categories they

19. Brent Berlin, "Speculations on the Growth of Ethnobotanical Nomenclature," *Language in Society* 1 (1972): 51–86; idem, "Ethnobotanical Classification," in *Cognition and Categorization*, ed. Eleanor Rosch and Barbara B. Lloyd (Hillsdale, N.J.: Erlbaum Associates, 1978), 9–26; Brent Berlin, Dennis E. Breedlove, and Peter H. Raven, *Principles of Tzeltal Plant Classification: An Introduction to the Botanical Ethnography of a Mayan-Speaking People of Chiapas* (New York: Academic Press, 1974).

20. Eleanor Rosch, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson, and Penny Boyes-Braem, "Basic Objects in Natural Categories," *Cognitive Psychology* 8 (1976): 382–439.

21. This interpretation of Rosch's work is drawn from Lawrence W. Barsalou, *Cognitive Psychology: An Overview for Cognitive Scientists* (Hillsdale, N.J.: Erlbaum Associates, 1992), 181–84. See also Rosch et al., "Basic Objects in Natural Categories"; Eleanor Rosch, "Principles of Categorization," in *Cognition and Categorization*, ed. Eleanor Rosch and Barbara B. Lloyd (Hillsdale, N.J.: Erlbaum Associates, 1978), 27–48; Carolyn B. Mervis and Eleanor Rosch, "Categorization of Natural Objects," *Annual Review of Psychology* 32 (1981): 89–115; and Mary E. Lassaline, Edward J. Wisniewski, and Douglas L. Medin, "Basic Levels in Artificial and Natural Categories: Are All Basic Levels Created Equal?" in *Percepts, Concepts, and Categories: The Representation and Processing of Information*, ed. Barbara Burns, Advances in Psychology, 93 (Amsterdam: North-Holland, 1992), 327–78.

must consider in making a categorization. For instance, if you were concerned with *things in my house*, categorizing at the level of *musical instrument* would yield the most efficient category, since the number of contrasting categories at this level (such as *furniture* and *fruit*) is significantly less than at the next level (*guitar, piano, apple, peach, table, lamp*). The second principle is the *informativeness principle*, according to which people tend to maximize the informativeness of their categorizations. Since the most information about any entity is found at the most specific level of a taxonomy, you would use *grand piano* to categorize the thing sitting in the living room. Rosch and her associates proposed that the intermediate level of a taxonomy (in this case, *piano*) optimizes both efficiency and informativeness and is thus the preferred level for basic categorizations. A number of empirical operations converge at the basic level. The basic level is the highest level whose members have similar and recognizable shapes; it is also the most abstract level for which a single mental image can be formed for the category. The basic level is also the highest level at which a person uses similar motor actions for interacting with category members. The basic level is *psychologically* basic: it is the level at which subjects are fastest at identifying category members, the level with the most commonly used labels for category members, the first level named and understood by children, the first level to enter the lexicon of a language, and the level with the shortest primary lexemes.[22]

In her initial work, Rosch assumed that the basic level depended on bundles of perceived-world attributes that formed natural discontinuities and that such attributes were inherent in the real world. She later came to doubt this assumption, and subsequent research has shown that the basic level is variable. It depends on cultural context and level of expertise and on whether one is dealing with categories that occur naturally, as humans interact with their environments, or with categories created for experimental ends.[23] Other research has indicated that the basic level is strongly determined by the ways in which humans interact with their environment and that it is highly dependent on the way humans perceive the spatial configuration of physical parts.[24] In all cases, however, Roger Brown's fundamental insight has held up: the preferred categorization for most of the things we encounter occupies a maximally useful level in the middle of a taxonomy.

22. Rosch et al., "Basic Objects in Natural Categories"; Eleanor Rosch, "Human Categorization," in *Studies in Cross-Cultural Psychology*, vol. 1, ed. Neil Warren (London: Academic Press, 1977), 1–49; Barbara Tversky and Kathleen Hemenway, "Objects, Parts, and Categories," *Journal of Experimental Psychology: General* 113 (1984): 169–93.

23. On Rosch's reservations about the objectivity of the basic level, see Rosch, "Principles of Categorization"; on the relationship between expertise and the basic level, see James W. Tanaka and Marjorie Taylor, "Object Categories and Expertise: Is the Basic Level in the Eye of the Beholder?" *Cognitive Psychology* 23 (1991): 457–82; on the contrast between natural and artificial categories in this context, see Lassaline et al., "Basic Levels in Artificial and Natural Categories."

24. On relationships between the basic level and the environment, see Barbara Tversky and Kathleen Hemenway, "Categories of Environmental Scenes," *Cognitive Psychology* 15 (1983): 121–49; on the basic level and humans' interactions with physical objects, see Tversky and Hemenway, "Objects, Parts, and Categories"; Irving Biederman, "Recognition-by-Components: A Theory of Human Image Understanding," *Psychological Review* 94 (1987): 115–47; and Lawrence W. Barsalou, "Deriving Categories to Achieve Goals," *Psychology of Learning and Motivation: Advances in Research and Theory* 27 (1991): 1–64.

THE BASIC LEVEL AND MUSICAL MOTIVES In a close analogue to what are, perhaps, the most distinctive features of basic-level categories, Schoenberg's theory of musical coherence begins not with individual musical events or with four- or eight-measure phrases, but at a level somewhere in between. Although the motive is the smallest *recognizable* part of a musical work, it is in fact made up of still smaller parts—namely, its constituent pitches, intervals, and durations. The cognitive salience of the motive thus mirrors that of the basic level: in both cases, the focus is on the whole rather than on the parts. Each of Schoenberg's families of motives has a distinctive "shape" or contour that allows us to distinguish it from other motives. Similarly, the basic level is the highest taxonomic level at which category members have similarly perceived overall shapes and the highest level at which a single mental image can reflect the entire category. Finally, motive recognition is not something given directly by perception and can change with expertise: what counts as a motive, just as what counts as the basic level, is not immutable.

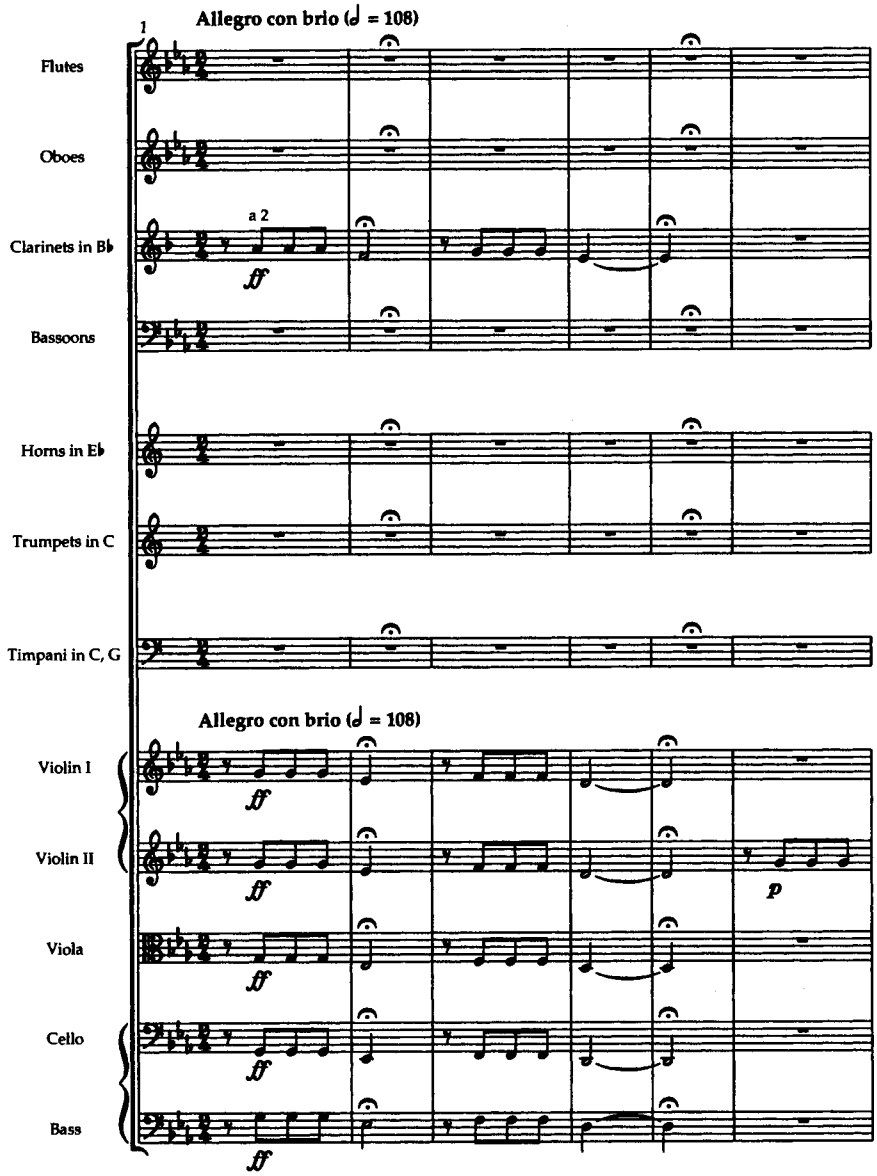
To illustrate correspondences between a musical motive and a basic-level category, let us consider the opening of Beethoven's Fifth Symphony, given in example 1.5. The passage is a familiar one, and it illustrates important matters of musical organization that we will later explore in Wagner's *Tristan*.[25] To the extent that we hear a "motive" in these opening gestures, we hear a figure that comprises four notes. The motive has a distinctive "shape" or contour that allows us to distinguish it from other motives (although such are in rather short supply in this particular example). Lastly, there is a certain minimal level of expertise involved in picking out mm. 1–2 and its ilk as "the motive." One can easily imagine an alternate hearing that would take mm. 1–2 and mm. 3–4 as the main thematic material for this section; the ensuing music would then represent a sort of fragmentation of this thematic material, rather than successive statements of various forms of the motive.[26]

SUMMARY Research on basic-level categories provides support for Schoenberg's proposal that musical comprehension begins at the motivic level, for such a level maximizes both efficiency and informativeness. In addition, there appear to be a number of correlations between musical motives and basic-level categories. Both are concerned with wholes, rather than the parts of these wholes; shape (whether as visual property or auditory analogue) is important to both; and both are subject to the influence of expertise. Finally, while entities such as "motives" may be basic to the process of comprehending music, they are actually relatively high-level cognitive constructs (as are basic-level categories). Although such constructs are fundamental to reason (and may even represent an important starting point for conscious

25. That a passage from Beethoven might be used to shed light on Wagner's compositional strategies is not accidental, for Beethoven's compositional style had a profound influence on Wagner. For discussion, see Klaus Kropfinger, *Wagner and Beethoven: Richard Wagner's Reception of Beethoven*, trans. Peter Palmer (Cambridge: Cambridge University Press, 1991); and Thomas S. Grey, *Wagner's Musical Prose: Texts and Contexts* (Cambridge: Cambridge University Press, 1995).

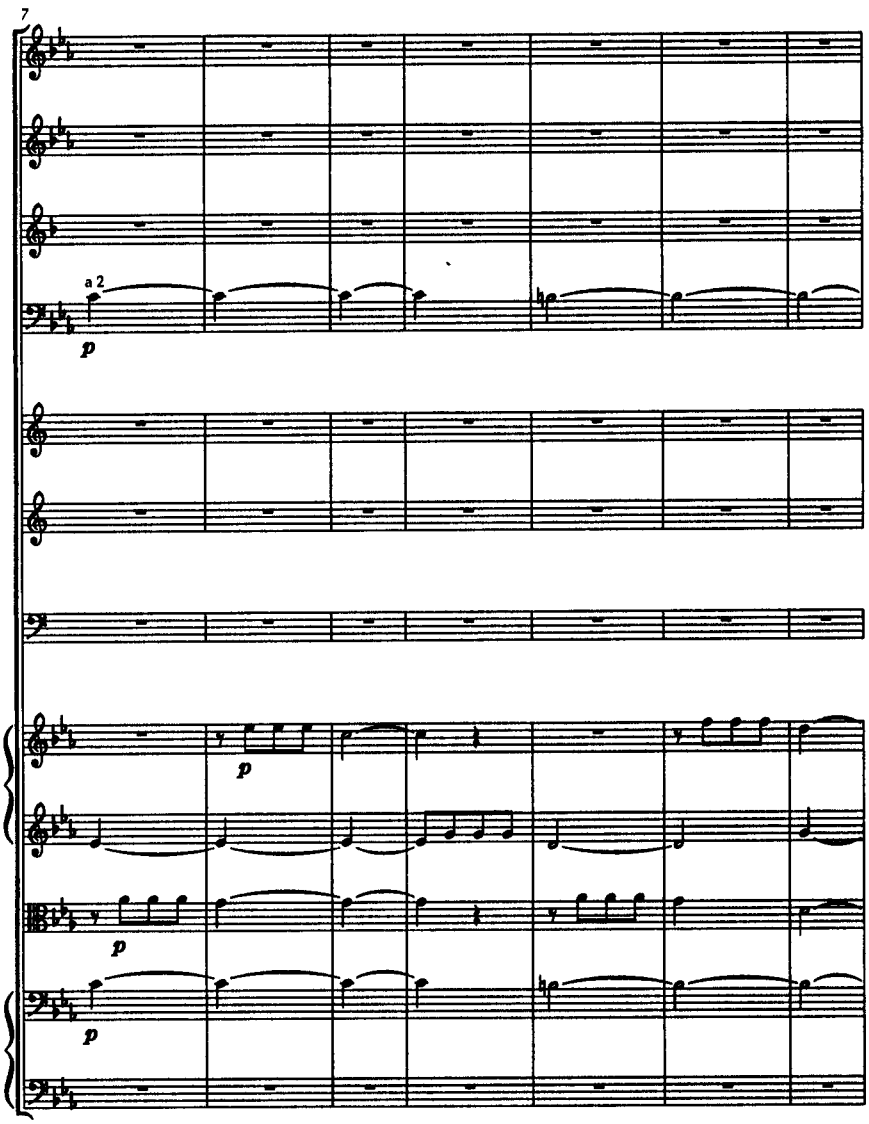
26. It is worth noting, however, the role of the fermatas in suggesting that mm. 1–2 are one entity and mm. 3–4 another, since these add time for the listener to process the musical information presented by Beethoven. This additional time may also play a role in what is considered the most "typical" form of the motive, a topic explored in greater depth later in this chapter.

EXAMPLE 1.5 Ludwig van Beethoven, Symphony No. 5 in C Minor, Op. 67, first movement, mm. 1–38




*(continued)*

EXAMPLE 1.5 *(continued)*

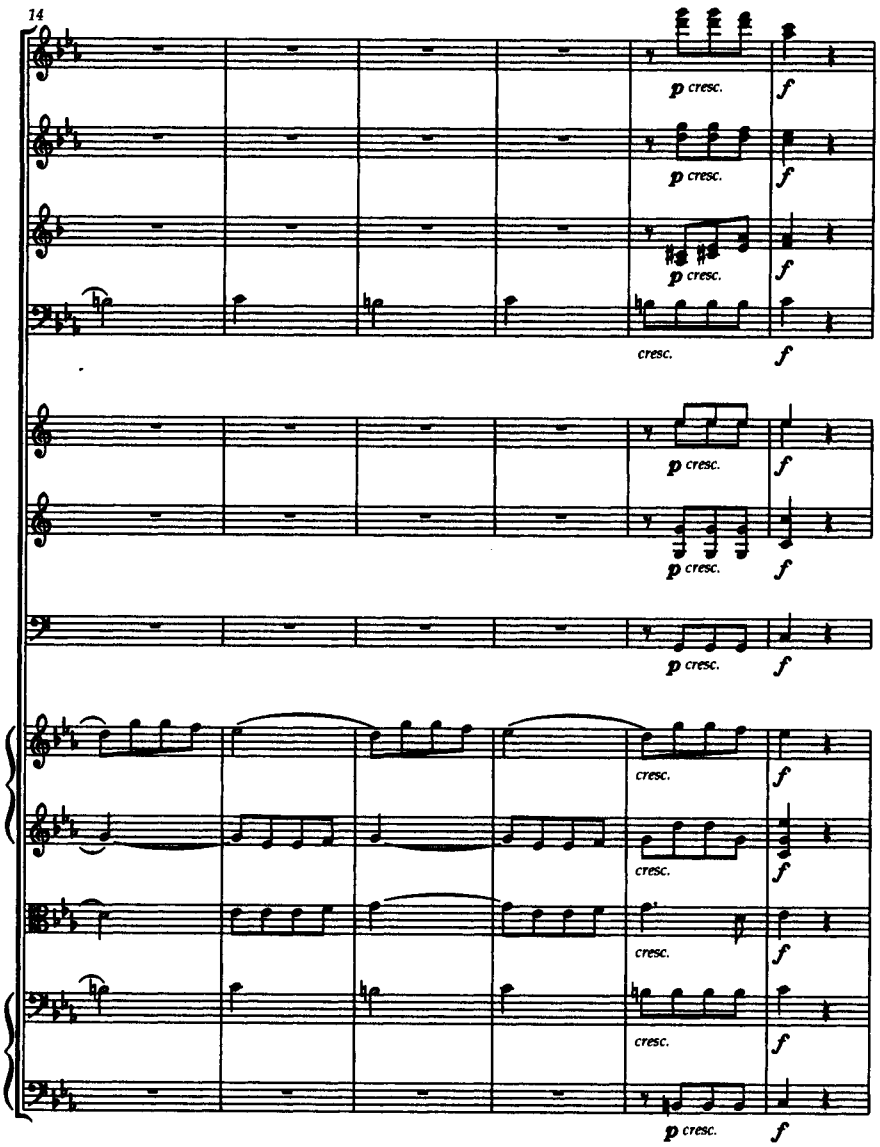



reasoning processes), they are the product of myriad other still more-basic processes that operate below the level of consciousness.

## Typicality Effects

GRADED STRUCTURE AND A TYPOLOGY OF CATEGORIES As noted, the variability of motive forms was important to Schoenberg's notion of musical coherence, for this variability tended to emphasize how motive forms were similar to

EXAMPLE 1.5 *(continued)*

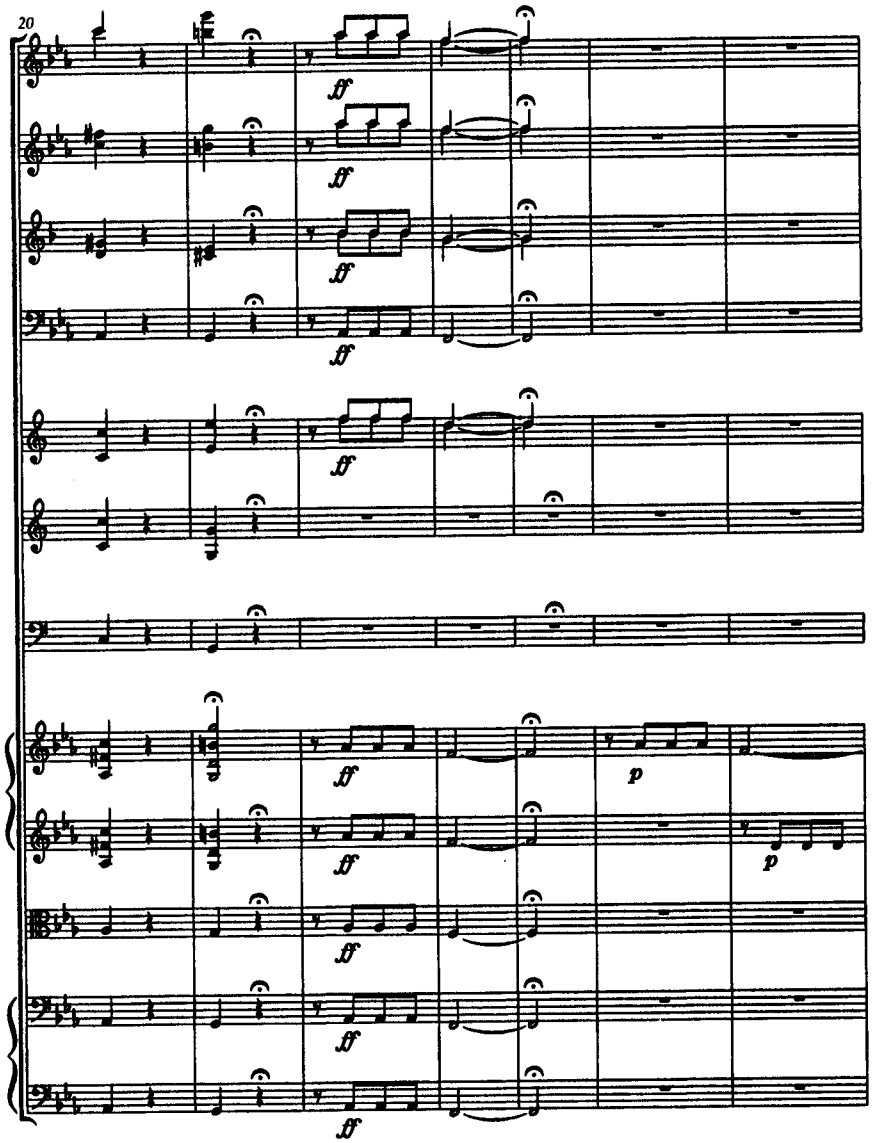



*(continued)*

one another. Research on categorization has demonstrated an analogous structure for numerous categories. Membership in these categories is not fixed but is instead graded through a dynamic process in which the attributes of potential category members are compared with the attributes most typically found within the category.

As an example of such a graded structure, consider the category *bird*. Experimental rankings show that subjects view robins and sparrows as the best examples of birds, with

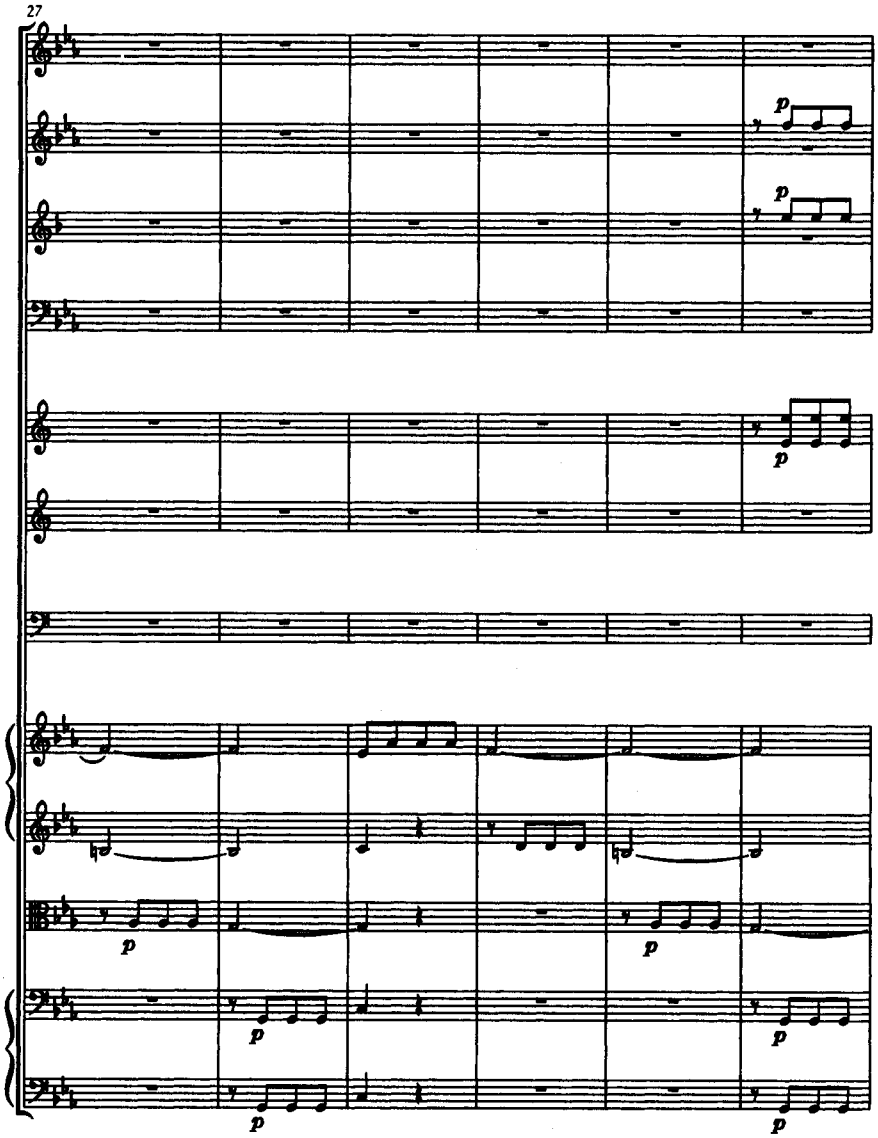
EXAMPLE 1.5 *(continued)*




owls and eagles lower down in the rankings, and ostriches, emus, and penguins among the worst examples.[27] All are considered members of the category *bird*, but some better represent the category than others. Consequently, category structure is graded according to typicality: category members range from the most typical to the least typ-

27. Eleanor Rosch, "On the Internal Structure of Perceptual and Semantic Categories," in *Cognitive Development and the Acquisition of Language*, ed. Timothy E. Moore (New York: Academic Press, 1973), 111–44; idem, "Cognitive Representations of Semantic Categories," *Journal of Experimental Psychology: General* 104 (1975): 192–233.

EXAMPLE 1.5 *(continued)*



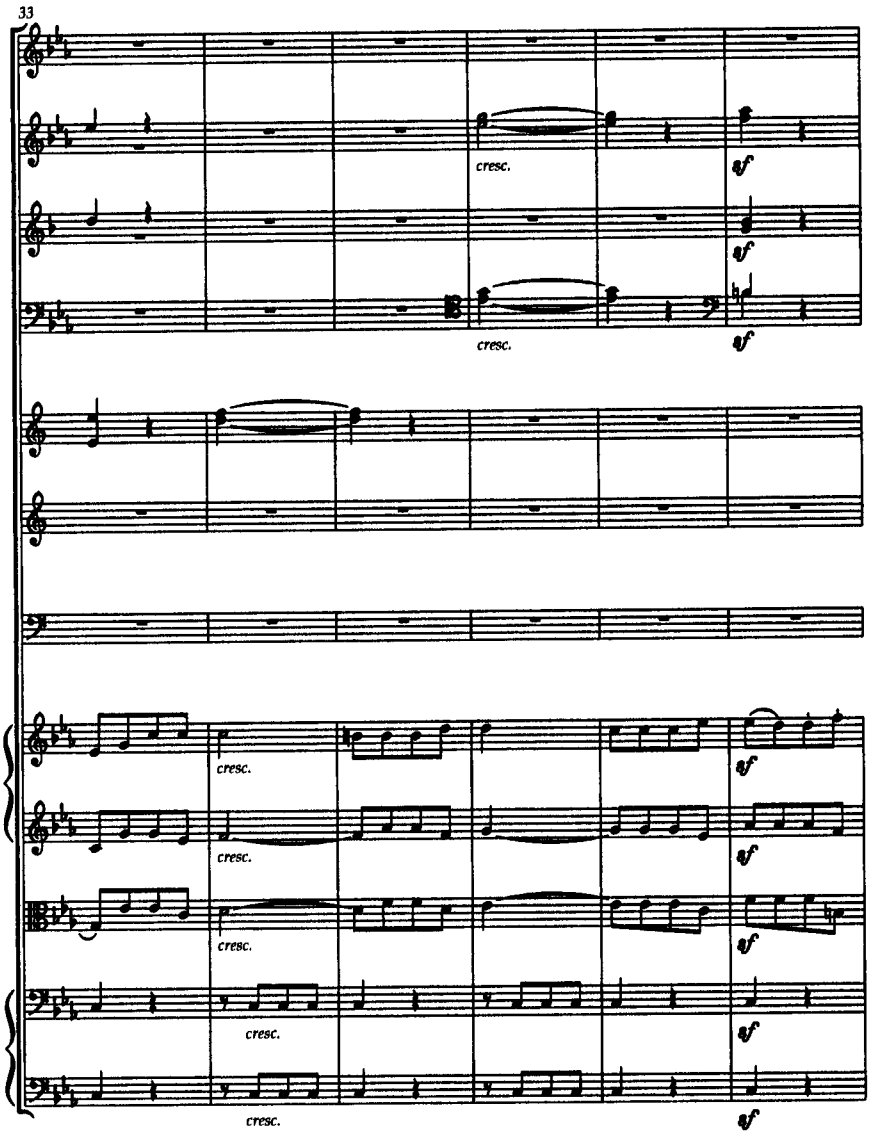

*(continued)*

ical, with the former securely inside the bounds of the category (robins and sparrows) and the latter in danger of being excluded from the category (emus and penguins).

Graded membership has been shown to be pervasive among what are conventionally called *natural* categories, so named based on their emergence from the interaction of humans with their natural environments.[28] For the sake of clarity, and

28. On the pervasiveness of natural categories, see Lawrence W. Barsalou, "The Instability of Graded Structure: Implications for the Nature of Concepts," in *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*, ed. Ulric Neisser (Cambridge: Cambridge University Press, 1987), 102. On the term "natural" applied to certain categories, see Barsalou, *Cognitive Psychology*, chap. 2.

EXAMPLE 1.5 *(continued)*




to avoid certain of the associations summoned by the term "natural," I call this sort of category a Type 1 category.

Type 1 categories stand in contrast to categories in which membership is based on individually necessary and jointly sufficient conditions, such that the members of category x must have features y and z, and if something has features y and z it is a member of category x. Because the conception of categories of this latter sort is often traced backed to the work of Aristotle, such categories have been called *clas-*

*sical* categories; because the structure of these categories is not typically an analog representation of the natural world, such categories have also been called *artificial*.[29] I call this sort of category a Type 2 category. Most evidence indicates that Type 2 categories simply represent a specialized form of Type 1 categories—take a Type 1 category, specify limits for the category through the imposition of necessary and sufficient conditions for category membership, and you create a Type 2 category.[30]

PROTOTYPE EFFECTS AND FRAME STRUCTURE Eleanor Rosch and Carolyn Mervis, in their pioneering work on Type 1 categories, suggested that such categories were organized around a stable cognitive construct called a *prototype*, which encapsulated the statistically most-prevalent features of members of the category and against which potential category members were compared.[31] Type 1 categories were thus said to exhibit "prototype effects." It is clear, however, that the prototype offers more structure than simply a record of the attributes correlated with a particular category: although *wings* and *beak* are both highly correlated with *bird*, we know that *wings* are used for locomotion and *beaks* are used for eating. Relationships between attributes, as well as the values assigned these attributes, are thus important to the function of the prototype.

As a way of capturing the relational structure inherent in most categories, Lawrence Barsalou proposed using the notion of a *frame*, as developed in artificial intelligence research, to represent category structure.[32] As Barsalou has shown, the

29. On classical or artificial categories, see Edward E. Smith and Douglas L. Medin, *Categories and Concepts* (Cambridge, Mass.: Harvard University Press, 1981), chap. 1; Sharon Lee Armstrong, Lila R. Gleitman, and Henry Gleitman, "What Some Concepts Might Not Be," *Cognition* 13 (1983): 263–308; Lila R. Gleitman, Sharon Lee Armstrong, and Henry Gleitman, "On Doubting the Concept 'Concept'," in *New Trends in Conceptual Representation: Challenges to Piaget's Theory?*, ed. Ellin Kofsky Scholnick (Hillsdale, N.J.: Erlbaum Associates, 1983), 87–110; Lakoff, *Women, Fire, and Dangerous Things*, chap. 2; Barsalou, *Cognitive Psychology*, chap. 2.

30. Although the typology of categories I introduce here places a burden on those familiar with the traditional labels, it has two advantages. First, it avoids the complications caused if "natural" is interpreted as "given by nature"; while Type 1 categories result from humans' interactions with their environment, they are not given by that environment in any simple way. Second, the typology classifies categories on the basis of their internal structure. This is especially useful when considering research on categorization, since "artificial" categories (such as *things to pack in a small suitcase*, as discussed by Barsalou in his "Deriving Categories to Achieve Goals") often have the same basic structure as do "natural" categories (such as *bird*).

31. Eleanor Rosch and Carolyn B. Mervis, "Family Resemblances," *Cognitive Psychology* 7 (1975): 573–605. See also James Hampton, "Prototype Models of Concept Representation," in *Categories and Concepts: Theoretical Views and Inductive Analysis*, ed. Iven van Mechelen, James Hampton, Ryszard S. Michalski, and Peter Theuns, Cognitive Science Series (London: Academic Press, 1993), 67–95.

32. Barsalou, *Cognitive Psychology*, chap. 7; idem, "Frames, Concepts, and Conceptual Fields," in *Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization*, ed. Adrienne Lehrer and Eva Feder Kittay (Hillsdale, N.J.: Erlbaum Associates, 1992), 21–74; idem, "Flexibility, Structure, and Linguistic Vagary in Concepts: Manifestations of a Compositional System of Perceptual Symbols," in *Theories of Memory*, ed. A. F. Collins, S. E. Gathercole, M. A. Conway, and P. E. Morris (Hillsdale, N.J.: Erlbaum Associates, 1993), 29–101.

The exploration and implementation of frame structures in work on artificial intelligence is demonstrated in Marvin Minsky, "A Framework for Representing Knowledge," in *The Psychology of Computer Vision*, ed. Patrick Henry Winston (New York: McGraw-Hill, 1975), 211–77; Marvin Minsky, *The Society of Mind* (New York: Simon and Schuster, 1985); and Douglas B. Lenat and R. V. Guha, *Building Large*

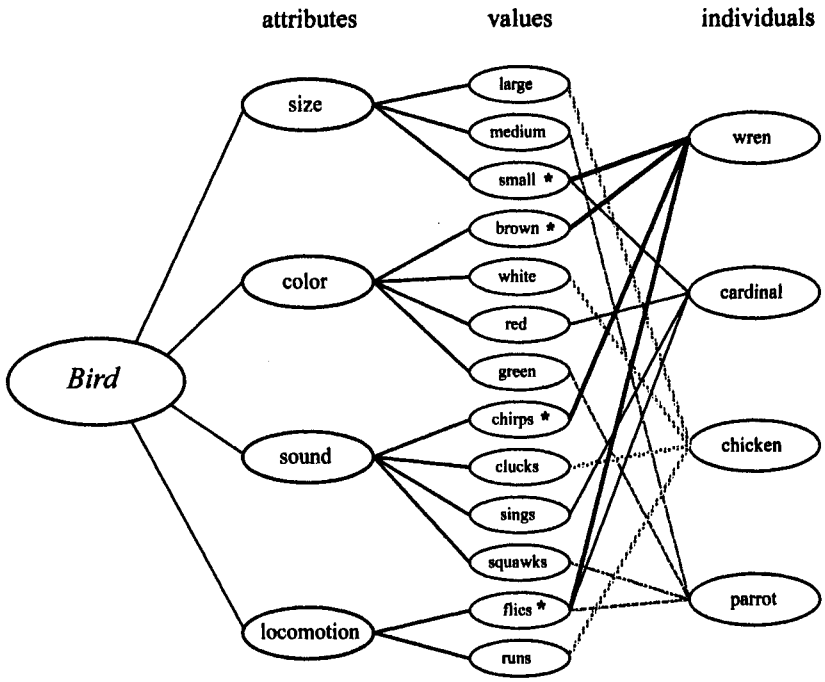



FIGURE 1.1 A partial frame for the category *bird* (asterisks indicate the values of the prototype that are stored in memory)

basic structure of a Type 1 category can be conceived of as a relational network made up of relatively abstract attributes, to which concrete values are assigned by each individual. For the category *bird*, for example, attributes might include *size*, *color*, *sound*, and *locomotion*, as shown in figure 1.1. For the attribute *size*, values might include *large*, *medium*, and *small*. Each member of the category exemplifies specific values for each of the attributes: a wren is *small*, *brown*, *chirps*, and *flies*; a chicken is *large*, *white*, *clucks*, and (for the most part) *runs*.

Of course, individuals with none of the attributes of the category would be excluded from it. Prototype effects arise from the comparison of the values represented by individuals with the values of the prototype for the category that are stored in memory (indicated with asterisks in fig. 1.1). If the values of the prototype for *bird* were *small*, *brown*, *chirps*, and *flies*, then a wren would be most typical of the category, a male cardinal (*small*, *red*, *sings*, *flies*) would be somewhat less typical, and a chicken (*large*, *white*, *clucks*, *runs*) would be least typical.

APPLICATIONS TO MUSIC Returning to the opening of the Fifth Symphony, we can see that, as suggested by Schoenberg's account of motive, there are a

---

*Knowledge-Based Systems: Representation and Inference in the Cyc Project* (Reading, Mass.: Addison Wesley, 1989).

number of different forms that Beethoven's motive takes in the first thirty-seven measures of the movement, and these vary with respect to contour, dynamics, and orchestration. Approached from the perspective of categorization, each motive form can be viewed as a member of the category *motive forms from the opening of Beethoven's Fifth Symphony*.[33] Attributes of this category might include things such as *orchestration*, *dynamic*, and *melodic profile*. In the following discussion, I show how these three attributes could play a part in determining membership in the category *motive forms from the opening of Beethoven's Fifth Symphony*; however, I should emphasize that, for present purposes, what is of concern is less the specific attributes and more the model of categorization they illustrate.

To characterize melodic profile, it is useful to account for intervallic relationships among motive forms in terms of diatonic intervals. That is, within the diatonic context of the movement, the descending third from G4 to E♭4 can be reckoned to be equivalent to the descending third from F4 to D4, even though the former is major and the latter is minor. Here I represent a descending third with **–3**, an ascending third with **+3**, and a unison with a **u**. With this in mind, the attribute-values for the first two statements of the motive would be *tutti*, *fortissimo*, and *u*, *u*, *–3*.

Although *rhythmic profile* is an equally important attribute of the category, it is not as useful for distinguishing among the various motive forms found in the opening measures shown in example 1.5. All the motive forms found there would assign the same value to this attribute: *short*, *short*, *short*, *long* (where *short* is equivalent to an eighth note and *long* varies between a quarter note and three tied half notes). In consequence, the attribute *rhythmic profile* is relatively transparent in the opening of the symphony, as is the attribute *played by instruments;* both will be omitted from the following account of category structure.

Figure 1.2 presents a frame diagram for the category *motive forms from the opening of Beethoven's Fifth Symphony;* it shows the values that thirteen instances of the motive assign to the attributes *orchestration*, *dynamic*, and *diatonic melodic profile*. The result is a rather dense diagram (owing in part to the number of individuals categorized), which gives an intriguing and rather counterintuitive picture of what is most typical of the various motive forms gathered there. Based on the evidence provided by these individuals, the most typical form of the motive would assign the values *solo*, *piano*, and *u*, *u*, *–3* to the relevant attributes. Although the last value is what we might expect, the first two are not: it is difficult to think of Beethoven's theme as anything other than *tutti* and *fortissimo*.

Were we to take into account a further fifteen instances of the motive, the profile

33. What follows is by no means the first attempt to use research on categorization to model variation among musical elements; notable are the studies of Robert L. Welker, "Abstraction of Themes from Melodic Variations," *Journal of Experimental Psychology: Human Perception and Performance* 8 (1982): 435–47; and Lucy Pollard-Gott, "Emergence of Thematic Concepts in Repeated Listening to Music," *Cognitive Psychology* 15 (1983): 66–94. The following extends this work by taking advantage of research done since the early 1980s and by endeavoring to give a more thorough account of the process of categorizing musical events. Mention should also be made of Robert O. Gjerdingen's discussion of typicality in *A Classic Turn of Phrase: Music and the Psychology of Convention*, Studies in the Criticism and Theory of Music (Philadelphia: University of Pennsylvania Press, 1988), 93–94, which anticipates many of the points made here.

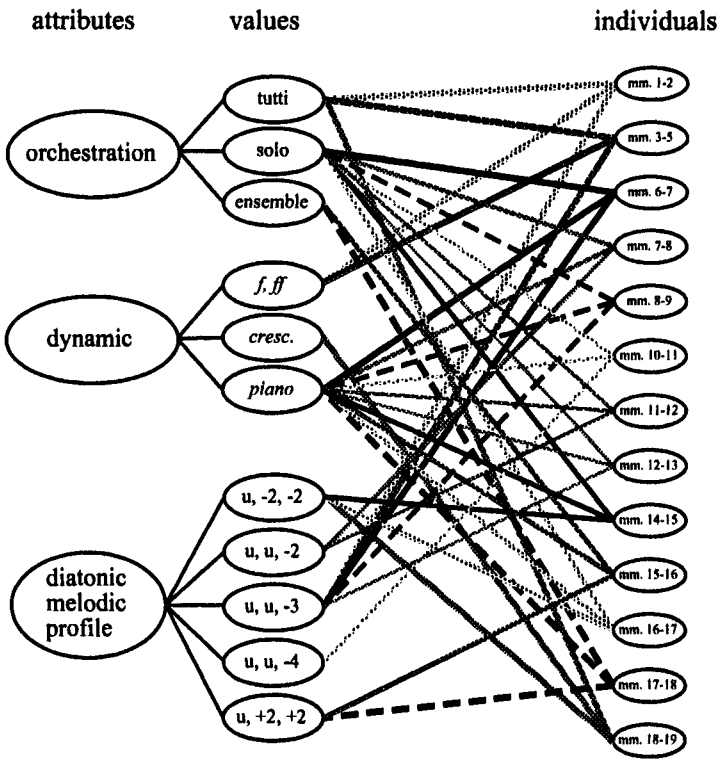



FIGURE 1.2 Diagram of the category structure for the category *motive forms from the opening of Beethoven's Fifth Symphony*, mm. 1–21

of typicality would be skewed even further away from what we might expect (this is shown in fig. 1.3, where the frame diagram has been rearranged somewhat in the interests of legibility). The difference between the evidence provided by figures 1.2 and 1.3 and our usual recollection of how Beethoven's theme sounds is due in part to a limitation of frame diagrams: they are essentially statistical representations of the most *common* values assigned attributes by individuals. They may not represent the most *important* attribute values for a given category, however.

BEYOND THE PROTOTYPE: CONCEPTUAL MODELS Recent empirical work on categorization has shown that category structure is even more variable than had been assumed by the researchers who first proposed the prototype theory of category structure.[34] In most Type 1 categories, some attributes are more important to category structure than others. As a result, some of the statistically most prevalent features of members of a category may be disregarded in formulating the

34. Armstrong et al., "What Some Concepts Might Not Be"; Gleitman et al., "On Doubting the Concept 'Concept'"; Barsalou, "The Instability of Graded Structure"; Lawrence W. Barsalou, "Intraconcept Similarity and Its Implications for Interconcept Similarity," in *Similarity and Analogical Reasoning*, ed. S. Vosniadou and A. Ortony (Cambridge: Cambridge University Press, 1989), 76–121.

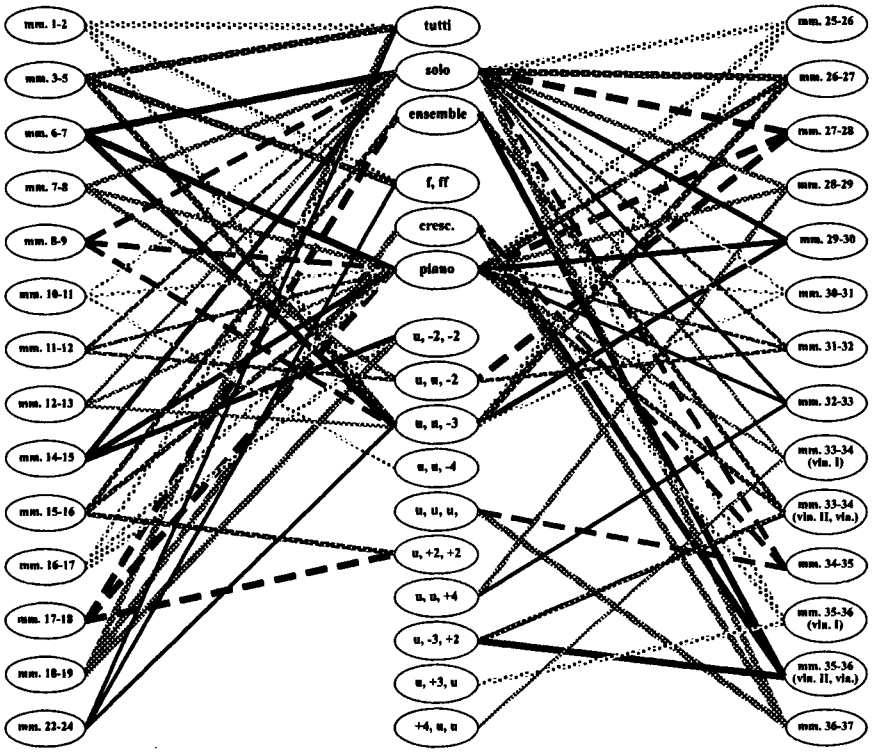



FIGURE 1.3 Diagram of the category structure for the category *motive forms from the opening of Beethoven's Fifth Symphony*, mm. 1–37

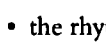
prototype. For example, attributes such as *natural environment* and *means of locomotion* have a strong influence on how we categorize whales, leading us to disregard such taxonomic attributes as *mode of respiration* and *means of reproduction*. In consequence, a whale seems closer to a prototypical fish than to a prototypical mammal, despite its proper Linnean categorization.[35] Research has also shown that judgments of typicality change over time, even when no new individuals are categorized. It appears that, rather than constituting a stable cognitive entity, the prototype for a category is actually fluid and variable and it changes with the circumstances of categorization.

An alternative view of category structure suggests that categories are organized around *conceptual models*. In brief, conceptual models consist of concepts in specified relationships; each model pertains to a specific domain of knowledge. Such models are central to reasoning and inference, and they will be discussed in greater detail in chapter 3; in what follows, the emphasis is on the role conceptual models play in categorization, rather than on the structure of such models. With regard to catego-

35. For a discussion of some of the factors involved in this particular categorization, see John Dupré, "Are Whales Fish?" in *Folkbiology*, ed. Douglas L. Medin and Scott Atran (Cambridge, Mass.: MIT Press, 1999), 461–76.

rization, the primary function of the conceptual model is to supply a guide for reasoning about accepted and potential members of the category. This is accomplished through a simplified representation of category structure that incorporates knowledge about what values are most typical for a select group of attributes for the given category. These attributes are selected according to the goals of categorization, which are themselves informed by more global conceptual models applicable to a broad range of categorization tasks.[36] The conceptual model for a given category thus reflects information of the sort summarized in frame diagrams but conditioned by knowledge about the overall goals of categorization.

As a means of developing a preliminary understanding of the concepts and relationships constituent of a conceptual model for a category, let us return once more to the opening of Beethoven's symphony and consider a conceptual model for the category *motive forms from the opening of Beethoven's Fifth Symphony*. The model, diagrammed in figure 1.4, consists of four correlated conceptual elements, characterized as the things necessary for a melodic fragment to count as an instance of the motive:

- the rhythmic pattern 
- a statement by the full orchestra
- a fortissimo dynamic
- an intervallic pattern consisting of three repeated notes followed by a descending diatonic third[37]

These elements give priority to select attributes of the motive (rhythmic pattern, orchestration, dynamics, and intervallic profile) and assign default values to these attributes—that is, they define what counts as a "typical" member of the category. This network of attributes and values then provides a guide for categorizing the various versions of the motive that occur in the opening measures. Accordingly, versions of the motive such as those of mm. 1–2 and mm. 22–24 (which instantiate all aspects of the conceptual model) would be judged most typical of the category; motive forms such as those of mm. 6–7 and mm. 29–30 (which instantiate only two aspects of the conceptual model) would be less typical; and motive forms such as those of mm. 7–8 and mm. 35–36 (which instantiate only the rhythmic aspect of the conceptual model) would be least typical of the category.

What, then, of the evidence provided by figure 1.3? That is, what can one make of the lack of correlation between the conceptual model and the individual category members, since most of these would be reckoned to be less typical members according to the conceptual model for the category? The answer to this question requires delving a bit deeper into the way conceptual models both reflect and shape our understanding of music.

36. Barsalou, "Deriving Categories to Achieve Goals"; Lawrence W. Barsalou, Wenchi Yeh, Barbara J. Luka, Karen L. Olseth, Kelly S. Mix, and Ling-Ling Wu, "Concepts and Meaning," in *Chicago Linguistics Society 29: Papers from the Parasession on the Correspondence of Conceptual, Semantic, and Grammatical Representations*, ed. Katharine Beals, Gina Cooke, David Kathman, Sotaro Kita, Karl-Erik McCullough, and David Testen (Chicago: University of Chicago, Chicago Linguistics Society, 1993), 23–61.

37. I should emphasize that the conceptual models I describe here are by no means the simplest elements of conceptual structure. Each conceptual cluster of the present model, for instance, could be represented by a still smaller conceptual model correlating the concepts that make up the cluster.

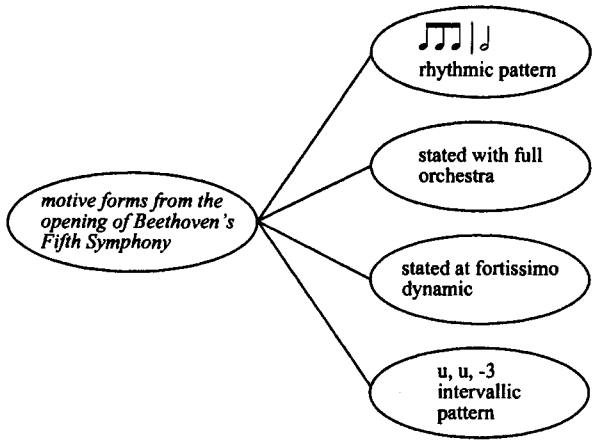



FIGURE 1.4 Conceptual model for the category *motive forms from the opening of Beethoven's Fifth Symphony*

The conceptual model shown in figure 1.4 gives undeniable priority to the statements of the motive that open the movement. The preference given the first statements of musical materials we encounter is not absolute—there are, of course, works in which the principal materials emerge only gradually—but it does reflect the fact that we must process musical materials in time and that it is often the case that significance accrues to the first events in any psychological process. Note that Beethoven departs from his opening materials only gradually; in particular, divergences from the intervallic pattern of the motive (the aspect of the typicality shown by fig. 1.3 that best matches the proposed conceptual model) occur somewhat later in the opening measures and correlate with the compositional strategies of development and transition.[38] The initial statements of the motive thus represent anchor points for the process of elaboration that Beethoven undertakes: the variations evident in figure 1.3 are variations *on* the musical materials summarized in the conceptual model of figure 1.4.

The conceptual model of figure 1.4 is also informed by our ideas about musical themes: that is, such themes should be strongly and clearly stated at privileged points within a movement (such as the beginning of the movement or after a change of key or tempo), after which they may be modified or embellished in conformance with compositional strategies. Our local model for the motive of Beethoven's symphony is thus shaped by a global model of what constitutes an appropriate musical theme. Such global models are developed through abstraction from any number of local models, and, perhaps more important, they are taken from the broader base of knowledge that constitutes culture.[39] For instance, in the case of Beethoven's symphony, the relevant global model for a musical theme reflects the influence of nineteenth-century German and Austrian musical culture and, in particular, the efforts

38. Relationships between processes of categorization and compositional strategy are developed more fully in chap. 4.

39. This perspective on culture is explored more fully in chaps. 3 and 5.

of the nineteenth-century critic and theorist A. B. Marx and others to come to terms with Beethoven's compositional style.[40]

Our global model of what counts as a theme (which applies to any number of works by any number of composers) is thus in rapport with the local models through which we organize our understanding of specific musical works. Without a global model, it can be difficult to decide what aspects of a phenomenon are relevant for categorization. Without local models, there is no way to anchor global models to specific phenomena.

It should be emphasized that the conceptual model sketched here and diagrammed in figure 1.4 represents limited features of a network of interrelated propositions that might be used to guide reasoning about how motive forms in the opening of Beethoven's symphony relate to one another. The model does not represent, in a simple way, what someone "has in mind" when confronting this music. The simplicity of the model (which reflects in part the global models relative to which it is framed) contributes to its efficiency. It is specific enough to allow for clear distinctions between the motive of the opening measures and other thematic materials in the symphony (such as the new theme that enters in m. 63). It is general enough, however, to accommodate the wide range of different forms the opening motive takes. The model also reflects our intuition that there is just one main form of the opening motive, despite evidence to the contrary. Intuitions such as this contribute to the plausibility of Type 2 categories, with their crisp boundaries and clear definitions. Features of the conceptual model for a given category will often be transformed into the necessary and sufficient conditions characteristic of Type 2 categories: take a Type 1 category and specify limits for it by regarding the constituent elements of the conceptual model relative to which it is structured as necessary and sufficient conditions for category membership, and you create a Type 2 category. In descriptions of music, this process yields "textbook" definitions of the sort that often lead to considerable debate among music theorists. Although such debates appear to be about music, they are in fact about how to define the categories through which we organize our understanding of music.

SUMMARY: TYPICALITY EFFECTS AND MUSICAL MOTIVE Categories that show typicality effects offer a way to model both the coherence and the variability of motive forms basic to Schoenberg's theory of motive. The property of coherence is analogous to the shared attributes of category members—that is, our

40. Marx's is perhaps the most important nineteenth-century voice within the tradition of analyzing the thematic aspects of Beethoven's music. Although Marx did not restrict himself to a definition of "theme" as such, preferring a broader and more organic account of musical organization, he did offer the notion of the *Satz*. Formed from motives, the Satz is a complete self-sufficient musical idea and represents the first coherent unit of musical organization. See Adolph Bernhard Marx, *Musical Form in the Age of Beethoven: Selected Writings on Theory and Method*, ed. and trans. Scott G. Burnham, Cambridge Studies in Music Theory and Analysis, 12 (Cambridge: Cambridge University Press, 1997), 68. The Satz is roughly comparable to Schoenberg's notion of a *Grundgestalt*, as discussed in Schoenberg, *The Musical Idea*, 169. It is no accident, of course, that Marx's formal theories reflect his efforts to make explicit Beethoven's compositional strategies. That is, to the extent that our notion of what constitutes a theme resonates with Marx's definition, that notion is informed by works such as the Fifth Symphony. Further discussion of Marx's notion of a motive is taken up in chap. 7.

sense that a collection of things coheres as a category reflects the attributes shared by those things. The variability Schoenberg noted is analogous to the variation among members of a Type 1 category. Just as certain motive forms are taken as the source for the derivation of further motive forms, so certain category members will be regarded as more typical of the category than others. While typicality effects reflect, to a certain extent, statistical invariances among the attribute-values demonstrated by category members, they are also a function of the goals of categorization. These goals shape the category's conceptual model, which represents what is regarded as most typical of the category and which also provides a guide for reasoning and inference about potential category members.

An important difference between collections of musical motives and cognitive categories is that the former must be processed under the temporal constraints of musical performance. Under most circumstances, these constraints will inform the conceptual model for a category of motive forms, with priority given to those forms that occur early in a work. As the example of Schubert's *Wanderer Fantasie* suggests, however, other factors may cause our concept of what constitutes the central motive of a work to change as the piece unfolds.

## CATEGORIZATION AND WAGNERIAN LEITMOTIVEN

### Wagner's Leidensmotiv

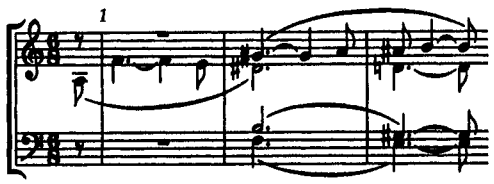
Approaching the seven statements of the opening motive from *Tristan und Isolde* given in example 1.1 as a cognitive category offers a way to explain why it is that we consider them "the same" even when they are manifestly different, and it gives some insight into their role as a point of entry into Wagner's sonic fantasy. Knowing something about the structure of cognitive categories also gives us a way to glimpse features of Wagner's compositional strategy, which both exploits and challenges our ability to categorize.

Before we explore the motive as a category, let us first situate the motive in its original context. As shown in example 1.6, the motive is accompanied by harmonies (provided by the clarinets and bassoons) and a rising chromatic line in the oboes. Hans von Wolzogen, in his classification of the motives from *Tristan und Isolde* (which he referred to as *Leitmotiven*), called the main motive the *Leidensmotiv* (the suffering motive) and the rising chromatic line the *Sehnsuchtsmotiv*, or the yearning motive.[41] The harmony on the downbeat of m. 2 is traditionally known as the "Tristan chord." As Carl Dahlhaus has noted, all of these elements are intertwined to such an extent that taking any one of them out of context distorts the whole.[42] Nonetheless, it is the

41. Hans von Wolzogen, *Thematischer Leitfaden durch die Musik zu Richard Wagner's* Tristan und Isolde*: Nebste einem Vorworte über den Sagenstoff des Wagner'schen Dramas* (Leipzig: Feodor Reinboth, 1890), 15. This work was also published in an English edition as *Guide through the Musical Motives of Richard Wagner's* Tristan and Isolde*: With a Preface on the Legend and the Poem of Wagner's Drama* (Leipzig: Feodor Reinboth, 1889).

42. Dahlhaus, *Richard Wagner's Music Dramas*, 63.

EXAMPLE 1.6 Opening motivic group from Richard Wagner's *Tristan und Isolde*



changes to the Leidensmotiv that are the most noticeable (not the least because it is the first sounding element of the ensemble), and so in the following I will focus on it, with the understanding that it is but one part—if a rather prominent one—of a larger whole.[43]

With these thoughts in mind, let me propose that the category *statements of the Leidensmotiv* is organized around a conceptual model consisting in two correlated elements, characterized as the things necessary for a melodic fragment to count as an instance of the motive:

- an ascending sixth, followed by three descending half-steps
- the rhythmic pattern 𝄋♩|♩.‿♩♩|♩.‿♩[44]

Again, this model is greatly simplified, for in addition to ignoring the materials that invariably accompany the Leidensmotiv, nothing is said of dynamics or orchestration, and the notion of a "motive" or "theme" is simply assumed (as part of the global model that frames this local model). The model is sufficient, however, to distinguish the Leidensmotiv from other motivic materials in Wagner's music drama and to evaluate the typicality of different versions that appear over the course of the work.

Turning to the versions of the motive given in example 1.1 we can note the following range of typicality: statements 1, 2, 6, and 7 are in complete conformance with the model (they are typical of the category), but statements 3, 4, and 5 conform rather less well (they are less typical of the category). Viewed from the perspective of compositional strategy, Wagner begins with two thoroughly typical members of the category and then follows them with a markedly less-typical member that adds a note and changes the rhythmic profile. After a considerable span of music, he returns to the motive, but again in a less than typical form—indeed, the two temporally compressed versions of mm. 85–90 enact the sense of reminiscence they are meant to evoke by presenting to our ears a distillation of the most salient aspects of

43. Robert P. Morgan has recently offered an analysis of the *Tristan* Prelude which focuses more on harmonic structure than on the appearance of motive forms and which offers an intriguing alternative to what I present in the following. See Morgan, "Circular Form in the *Tristan* Prelude," *Journal of the American Musicological Society* 53 (2000): 69–103. A selection of analyses of the Prelude spanning the twentieth century can be found in Robert Bailey, ed., *Prelude and Transfiguration from* Tristan and Isolde (New York: W. W. Norton, 1985).

44. The actual duration of the concluding note of the motive is variable; in all cases, however, its duration is of moderate length relative to that of the other notes that make up the motive.

the motive. Ten measures further (beginning in m. 100) there are two final appearances of the motive before the curtain goes up, each a thoroughly typical form. Wagner's compositional strategy thus falls into three stages. In the first stage (associated with statements 1 and 2), he establishes the motive. In the second stage (associated with statements 3, 4, and 5), he presents less typical forms of the motive; these are correlated with a larger set of "developmental" strategies, with which the central portion of the Prelude is concerned. Finally, in the third stage (associated with statements 6 and 7), he returns to more typical forms of the motive, preparatory to concluding the Prelude and beginning the music drama proper.[45]

Two things emerge from this initial sketch. First, in the different versions of the Leidensmotiv that appear in the Prelude, we have a category that develops over time. Because categorization is an active response to the environment, it always has a temporal dimension: a comparative categorization of birds, for instance, demands that we evaluate each in turn. With music the temporal aspect is slightly more prominent, since musical entities are thoroughly ephemeral: not only must we evaluate each musical entity in turn, but also this evaluation is practically the only evidence that the entities ever existed at all. Wagner is clearly conscious of the way the Leidensmotiv functions through time and uses it to give shape to the Prelude as a whole. Second, the view of Wagner's compositional strategy derived from comparing the versions of the motive in the Prelude, fascinating as it might be, is also misleading: the sort of circularity that it suggests is fundamentally at odds with his goal of perpetually progressive music. By looking at further instances of the Leidensmotiv over the course of *Tristan und Isolde* we can develop our understanding of both of these things and refine our interpretation of Wagner's treatment of the motive in the Prelude.

## The Leidensmotiv and Musical Syntax

The Leidensmotiv is prominent at five moments in *Tristan und Isolde*: in act 1, scene 2, mm. 26–33 (p. 36; see ex. 1.7a) and scene 5, mm. 439–41, 456–58, and 468–72 (pp. 175–78; ex. 1.7b); in act 2, scene 3, mm. 260–78 (p. 448; ex. 1.7c); and in act 3, scene 2, mm. 93–107 (pp. 594–96; ex. 1.7d) and scene 3 (near the conclusion of the opera), mm. 153–57 (p. 630; ex. 1.7e).[46] In each case, there are three entries of the motive, beginning (as did those in the opening of the Prelude) on A (or, in the case of the first entry in act 3, scene 2, A♭), B, and D.

Each group departs from the conceptual model proposed above in various ways. The versions of act 1, scene 2 have one less pitch (and, of course, one less interval) than the model and a different rhythmic profile. Those of act 1, scene 5 have a different rhythmic profile. In act 2, scene 3, the versions of the motive are quite close

45. This syntactic pattern is analogous to the Beginning–Middle–End paradigm discussed by Kofi Agawu, but here it is expressed purely in terms of categories, rather than in terms of tonal organization. See Victor Kofi Agawu, *Playing with Signs: A Semiotic Interpretation of Classical Music* (Princeton, N.J.: Princeton University Press, 1991), chap. 2.

46. Measure numbers are given by scene; page numbers refer to the Dover reprint of the C. F. Peters 1911 edition of the score: Richard Wagner, *Tristan und Isolde* (New York: Dover Publications, 1973).

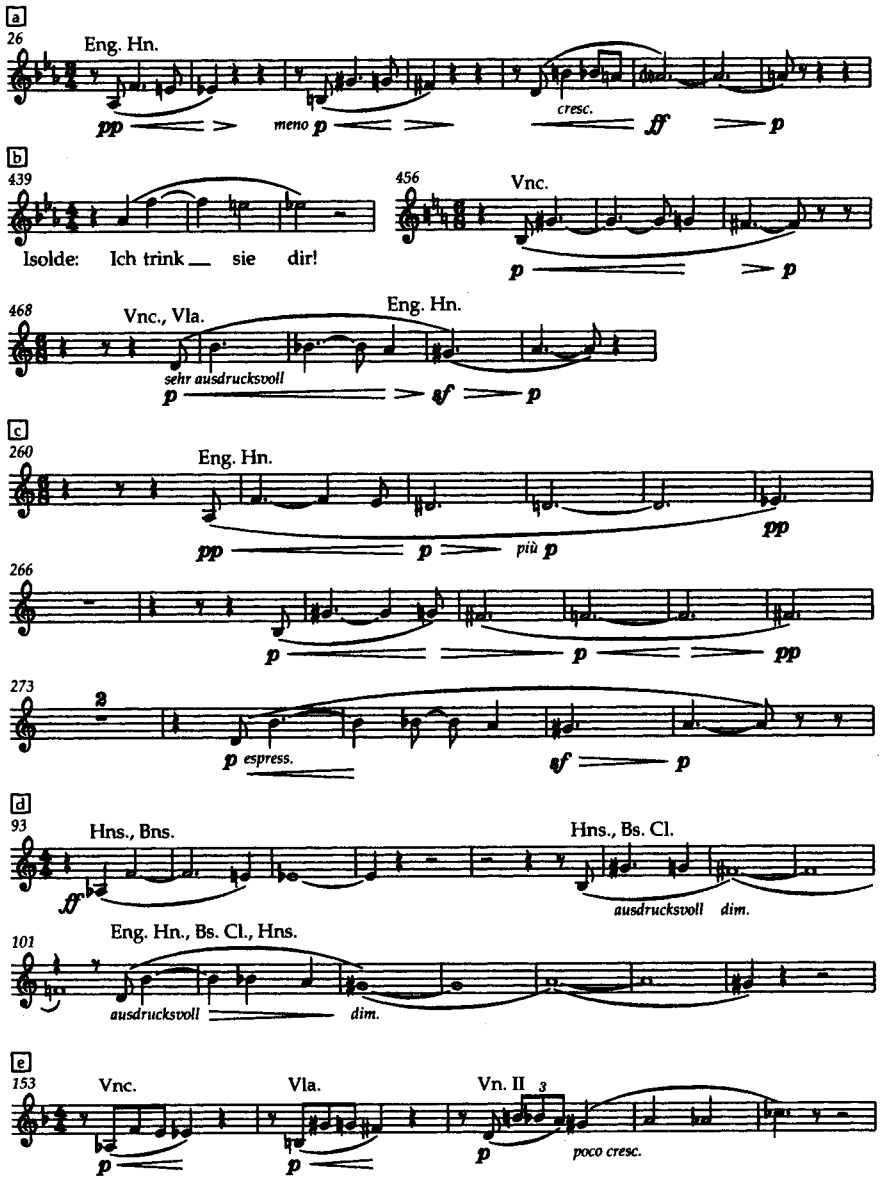
to the model, but the first two have extended endings, while those of act 3, scene 2 have an altered rhythmic profile, so that their overall duration is greatly expanded. Finally, the versions of the motive in act 3, scene 3 are greatly compressed (as are those of act 1, scene 2). Within each group, however, the versions of the motive are relatively consistent (with the exception of those of act 1, scene 5, about which more in a moment), although the last statement (beginning on D) is always a bit different from the preceding two. This consistency suggests that each group of three statements of the Leidensmotiv forms a syntactic unit, which represents its own local standard of typicality. These units are themselves members of a superordinate category (to which we could give some sort of cumbersome name like *successions of the Leidensmotiv beginning on A/A♭, B, and D*) whose members span almost the entirety of Wagner's music drama.[47]

For the most part, the statements that make up these syntactic units are conjunct, and, although they may stretch across a number of measures, the connection between the successive statements is quite obvious. The one exception comes with the three statements of act 1, scene 5, which occur at a climactic moment in the drama. Tristan has just drunk what he believes is a death potion. Isolde wrests the cup from him, sings "Verräther, ich trink' sie dir!" ("Betrayer! I drink to you!"), and then empties what remains in the cup. Unbeknownst to either, the draught they have consumed will lead not to death but will instead cause them to fall even more deeply in love. The melody for Isolde's unintentionally ironical "Ich trink' sie dir!" is the version of the Leidensmotiv that begins example 1.7b. For the fifteen measures of music that follow, the stage directions read: "She drinks, then throws the cup away. Both [Tristan and Isolde], seized with awe, in the greatest excitement but motionless, look fixedly into each other's eyes, in which the expression of defiance of death soon gives way to the glow of love." The cellos then enter with the second statement of the Leidensmotiv, beginning on B. Further stage directions follow, to accompany another ten measures of music: "Trembling seizes them: they clutch convulsively at their breasts and pass their hands over their foreheads." Unison cellos and violas then bring in the last statement of the Leidensmotiv, beginning on D. In this instance, what draws the three versions of the motive together into a single category is not so much their similarity to one another (although this is still quite audible) or their temporal proximity, but the way Wagner privileges them within his musical discourse. Not only is there nothing else to compete with them at the moment of their appearance, but also their participation in this pivotal moment suggests that they are autonomous agents within the drama itself.

We should, if but briefly, consider the notion of syntax developed in the preceding paragraphs. First, the basic idea of syntax that I employ here is what Webster's calls "a connected or orderly system." I flesh out this idea further in chapter 4; for now, it is enough to say that this appears to be how at least some music theorists

47. Robert P. Morgan has also pointed out that each successive phrase within these units takes up the same notes outlined in the top voice of the preceding one, making it seem that each phrase emerges from the one just before; see Morgan, "Circular Form in the *Tristan* Prelude," 95, 102.

EXAMPLE 1.7 Statements of the Leidensmotiv from Richard Wagner's *Tristan und Isolde*: (a) act 1, scene 2, mm. 26–33 (*Richard Wagner: Sämtliche Werke* [*WSM*], vol. 8, part 1, mm. 313–20); (b) act 1, scene 5, mm. 439–41, 456–58, 468–72 (*WSM*; vol. 8, part 1, mm. 1754–56); (c) act 2, scene 3, mm. 260–78 (*WSM*, vol. 8, part 2, mm. 1890–1908); (d) act 3, scene 2, mm. 93–107 (*WSM*, vol. 8, part 3, mm. 1301–15; (e) act 3, scene 3, mm. 153–57 (*WSM*, vol. 8, part 3, mm. 1581–85)



construe the notion.[48] Second, the structural integrity of the syntax associated with the Leidensmotiv emerges only over time. Although all of its elements are present from the opening moments of Wagner's music drama, confirmation of the syntax must wait until well into the opera. Third, this syntax is specific to *Tristan und Isolde*. Not only would we expect a different motive to be used for another work, but this motive would almost certainly be deployed in different ways. Musical syntax, at least with regard to this example, is thus created over time and is idiosyncratic to one particular composition. Finally, semantic value may become associated with syntactic structure, something evident in Wagner's manipulation of the three statements of the motive in the climactic moment that occurs in act 1, scene 5. By disrupting or modifying his musical syntax, Wagner can create a type of meaning that is thoroughly—and exclusively—musical.

## Syntactic Structure and Compositional Strategy

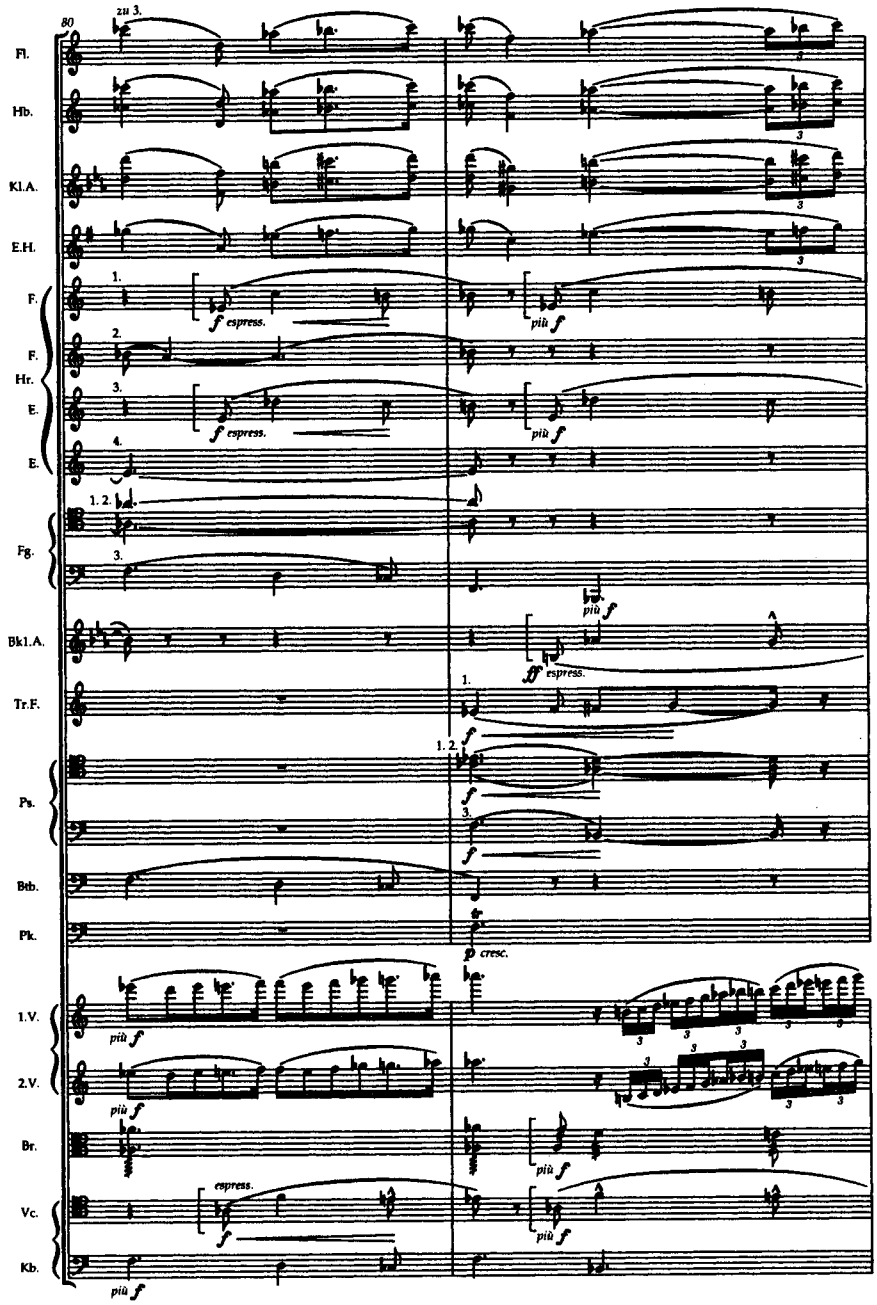
As a general type, the category *statements of the Leidensmotiv* has a temporal span coextensive with that of *Tristan und Isolde*. More specifically, Wagner carefully articulates the category over the course of the drama, with the motive appearing in syntactic units of three successive statements so that, with each entry of the motive on A (or A♭), we expect to hear two further entries, replicating the pattern set in the opening measures. Wagner relies on this expectation in act 1, scene 5 to underline both the tragedy and passion of Tristan's and Isolde's love. From this perspective, the further appearances of the Leidensmotiv in the Prelude after the initial three seem somewhat anomalous, for here they are only in pairs: statements on B and D in mm. 85–90, and on A♭ and B in mm. 100–106. But Wagner is playing a deep, even subliminal, game: buried beneath the overheated music in mm. 80–83 are three rhythmically displaced versions of the Leidensmotiv, all beginning on A♭ (as shown in ex. 1.8). Although little more than wraiths appearing in the midst of a fury of motives, these statements have their continuation—indeed, their consequence—in the entrances on B and D in the measures that follow. This hidden syntactic unit implies that the next round of statements, which start with the entrances on A♭ and B in mm. 100–106, should be followed by one additional entrance of the Leidensmotiv on D. Wagner withholds the third statement, however, and instead calls for the curtain to go up: in response to our anticipation of a third statement of the motive, Wagner gives us an entire opera.[49]

The shadowy appearances of the Leidensmotiv in mm. 80–83 of the Prelude and the somewhat less-typical versions that follow in mm. 85–90 suggest that Wag-

48. Cone, in "On Derivation," pairs the organizational system of syntax with the expressive tools of rhetoric. As another example of this sort of approach, Robert P. Morgan, in a recent essay, describes tonality as comprising "not only the abstract collections of diatonic pitches (plus their chromatic alterations), which have significant symmetrical properties of their own, but various asymmetrical *syntactical* functions, harmonic and linear, that transform these collections into a full-fledged compositional language" (Morgan, "Symmetrical Form and Common-Practice Tonality," *Music Theory Spectrum* 20 [1998]: 2, emphasis added).

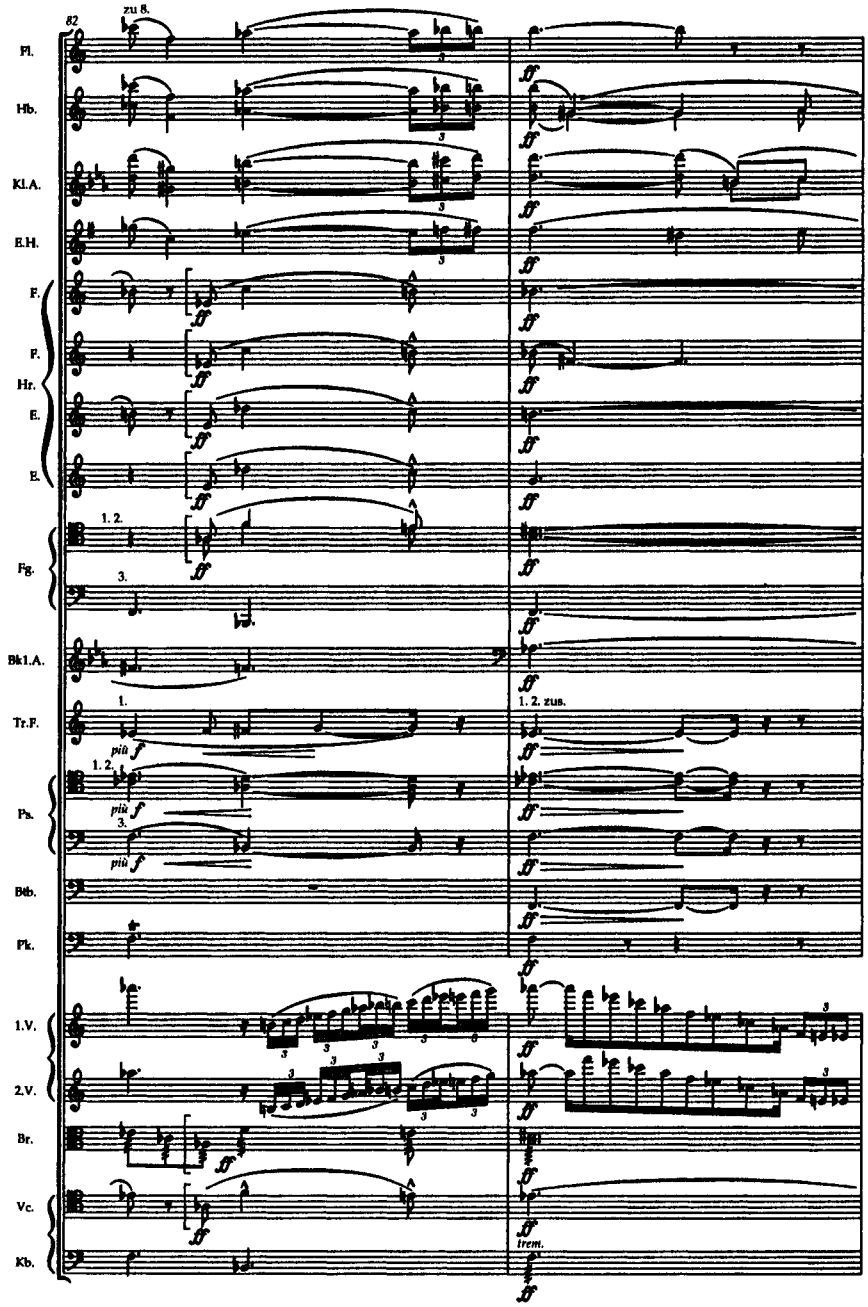
49. Morgan points out that the harmonic syntax for these final statements is also incomplete; see "Circular Form in the *Tristan* Prelude," 83.

EXAMPLE 1.8 Statements of the Leidensmotiv from Richard Wagner's *Tristan und Isolde*, Prelude, mm. 80–83




*(continued)*

EXAMPLE 1.8 *(continued)*

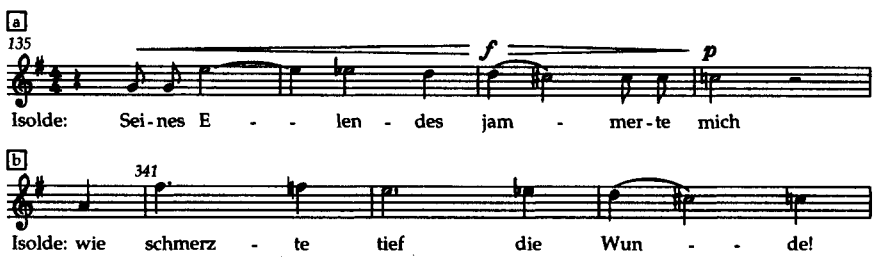


ner is both playing with and relying on our capacity to remember the motive (a capacity directly related to our ability to categorize it). But what Wagner wants is not the sort of recognition that calls attention to itself, distracting us from the flow of his drama as we become aware of its artifice, but something much more subtle. In consequence, the motive undergoes continual change, as the different versions collected in example 1.7 show, the only constant being the opening sixth and descending half steps. (Contrary to Schoenberg's concept of a motive—admittedly developed with reference to Brahms—rhythmic figuration is only weakly linked to the identity of the Leidensmotiv within the opera.) At the same time, the repetitions of the motive proper to each syntactic unit—which set up local, if ephemeral, standards for typicality—give each unit the illusion of stability and thus of properly representing the Leidensmotiv. What results is a syntax that is in constant transformation.

Further proof of Wagner's attention to musical memory can be seen in the relative atypicality of two of the rare isolated statements of the Leidensmotiv within *Tristan und Isolde*. One such statement occurs in the third scene of act 1. Isolde, recounting to Brangäne how her wrath against Tristan was turned to pity in her first encounter with him, sings "Er sah' mir in die Augen. Seines Elendes jammerte mich!—Das Schwert—ich ließ es fallen!" ("He looked into my eyes. His anguish touched my heart! The sword—I let it fall!"). As shown in example 1.9a, the middle phrase is sung to a modified version of the Leidensmotiv. This statement begins, unlike any other, on G4, and the chromatic passage that leads away from the E5 extends to four descending half-steps. And, relatively early in the long dialogue with Tristan that makes up scene 2 of act 2, Isolde sings "Wie schmerzte tief die Wunde!" ("How deeply the wound smarted"), again using a less-typical version of the Leidensmotiv (as shown in ex. 1.9b). The statement is doubled and ornamented by the strings, and its chromatic descent continues unhindered through six half steps. Both of these atypical appearances of the Leidensmotiv—atypical in their form relative to the conceptual model of the motive, and atypical in their isolation—function as remembrances of the motive rather than statements of it.

Carl Dahlhaus, in comparing the treatment of themes in the first movement of Anton Bruckner's Sixth Symphony to Brahms's method of developing variation, noted Bruckner's use of a similar strategy. When Bruckner returns to a theme, the

EXAMPLE 1.9 Isolated statements of the Leidensmotiv from Richard Wagner's *Tristan und Isolde*: (a) act 1, scene 3, mm. 135–38 (*Richard Wagner: Sämtliche Werke* [*WSM*], vol. 8, part 1, mm. 672–75); (b) act 2, scene 2, mm. 341–43 (*WSM*, vol. 8, part 2, mm. 886–88)



correspondence with the original version is often only approximate. Dahlhaus comments,

> That the one version is able to substitute for the other means, aesthetically, that instead of developing variation, where each variant represents . . . a consequence of the preceding one and a prerequisite for the next one, Bruckner makes use of an analytically elusive but clearly perceivable similarity by association, which makes the later version seem like a written-out memory of the earlier one. The logic of discourse, as conceived by Brahms, gives way to a system of approximate correspondences.[50]

The preceding analysis of transformations to the Leidensmotiv suggests that a more apt comparison would have juxtaposed Bruckner to Wagner rather than to Brahms, for it is "written-out memories" such as those shown in example 1.9 that contribute much to the power of Wagner's music dramas: we respond to such memories not with our intellect but in a way that seems far more intuitive and thus far more deeply felt.

Wagner, of course, could have had no explicit knowledge of how humans structure their understanding through categorization, since empirical evidence for this cognitive process did not begin to accumulate until a hundred years after the 1865 premiere of *Tristan und Isolde*. And yet Wagner almost assuredly had an implicit understanding of what Schoenberg called "the laws that govern the working of our minds." Wagner's music shows that he knew how to lay out his musical materials—to make them come to the fore, often enough and with enough plasticity—so that his listeners could comprehend their meaning. His music also shows that he knew how to challenge this ability, to modify the structure of categories of musical materials so that these same materials seemed new, or to have an entirely different ontological status (changing from a theme into a remembrance of a theme). An account of the way Wagner transformed the Leidensmotiv cannot exhaust his inexhaustible music, but it can give a glimpse into the processes through which we come to understand his music, as well as the way Wagner's compositional strategies both exploit and challenge that understanding.

## CATEGORIZATION AND CONCEPTUALIZING MUSIC

Schoenberg's theory of motive, shaped by his understanding of the constraints on musical composition imposed by human cognitive processing (his "laws of the mind"), offers insight into how we organize our understanding of music in terms of categories of musical events. Motives were, for Schoenberg, where musical recognition started, and the coherence of musical motives was what made musical comprehension possible. As we have seen, these two functions are replicated in cognitive categories. Our preferred level of categorization is at a level in the middle of a taxonomy—the basic level—which, like the motive, strikes a compromise between efficiency and informativeness. The categories we use most often in everyday life

50. Carl Dahlhaus, *Nineteenth-Century Music*, trans. J. Bradford Robinson (Berkeley: University of California Press, 1989), 273.

are not defined in terms of individually necessary and jointly sufficient conditions. Instead, they have variable membership: some members are more typical of the category than others, and membership in the category is graded rather than an all-or-nothing affair. Musical motives show a similar amount of variation, and it is through developing a notion of what they have in common—a conceptual model for what is typical of the category—that we begin to comprehend the musical organization of works like those that fascinated Schoenberg.

As the analysis of Wagner's use of the Leidensmotiv shows, Schoenberg's conception of a motive is not adequate for all circumstances (especially in its reliance on rhythmic figuration), nor are composers always content to establish a single model for a motive and stick with it. Indeed, they may vary it constantly (as did Wagner in the case of the Leidensmotiv), or they may combine the models for different motive forms in any number of ways (as Cone's work on the problem of motivic derivation shows). Although focusing on motives offers a place to start, it is on the more general notion of a musical category—free as it is from associations with Beethoven's compositional process, fascinating as that is—that I want to focus. While a collection of assorted motive forms is a good example of a musical category, categories can be much more various and structured around whatever set of musical relationships seems best to account for what is salient about a particular repertoire. The relevant units may be harmonic (as in a chaconne), involve a repeated bass pattern (as in a passacaglia), or combine the two (as in the linear cadential patterns that Susan McClary has shown provide a structural framework for music of the sixteenth century);[51] music of the twentieth century suggests an even greater number of possibilities. And music (construed as patterned sound) may be just one of a number of contributing factors: one could also envision musical categories informed by the text being set (of particular importance for liturgical music) or the affect that is to be summoned (a concern for some types of dramatic music). In all cases, the choice and structure of the relevant musical categories will be guided by the global conceptual models relative to which the local models for each category are framed.

My proposal, then, is that categories are where our conceptualization of music starts. If to think is to think in terms of categories, then to think of music is to think in terms of musical categories. Viewed this way, categories constitute an analogue—or perhaps replacement—for the notion of a group as it has been developed by Fred Lerdahl and Ray Jackendoff. For Lerdahl and Jackendoff, musical understanding begins with the process of segmenting a musical work into contiguous sequences of pitch-events, drum beats, or the like.[52] Each such segment constitutes a group, and groups are arranged into a strict hierarchy that ultimately yields the entire work of music.[53] Musical categories

51. Susan McClary, *Conventional Wisdom: The Content of Musical Form* (Berkeley: University of California Press, 2000), 14–16.

52. Fred Lerdahl and Ray Jackendoff, *A Generative Theory of Tonal Music* (Cambridge, Mass.: MIT Press, 1983); the process of grouping is described on p. 13; the definition of a group is given on p. 37.

53. By "strict hierarchy," I mean a hierarchy in which the boundary points at each superordinate level conform to boundary points at the level immediately subordinate. Lerdahl and Jackendoff define this sort of organization as "hierarchical"; see Lerdahl and Jackendoff, *A Generative Theory of Tonal Music*, 13, 37. Conceptions of hierarchy as they are applied to music are discussed in greater detail in the latter part of chap. 7.

can absorb many of the functions of groups—specifically, the role of the group as the first level of organization for musical phenomena—but do not require absolute contiguity or strict hierarchical organization. As a result, a theory of musical categories can easily account for the multiple strands of polyphonic music in a way that grouping theory cannot. Each of the overlapping entries of a work in canon, for instance, could be regarded as a member of a category; this sort of analysis is simply impossible with grouping theory, since overlaps between groups are not permitted.

If categories are where our conceptualization of music starts, they must be closely linked to concepts themselves. At the risk of some confusion, I would like to go further and suggest that musical categories are concepts. Now, to attempt any definition of a concept is to (at least potentially) embroil myself in a long-standing debate within philosophy and cognitive science.[54] To skirt this debate, I concentrate on a definition of "concept" that reflects recent work in the mind sciences and the brain sciences, and that can be applied to music. Although I think this definition can inform the larger debate on "concept," I shall leave the substance of the debate for another time.

The definition of "concept" I develop here is influenced by the work of Gerald Edelman, who has been interested in developing a biological approach to consciousness. Edelman's definition of a concept is developed as a description of the capacities necessary for the control of complex interactions between an organism and its environment:

> An animal capable of concepts is able to identify a particular thing or action and control its future behavior on the basis of that identification in a more or less *general* way. It must act *as if* it could make judgments based on recognition of category membership or integrate "particulars" into "universals." This recognition rests not just on perceptual categorization (although a concept may have a highly sensory content) but, to some degree, *must also be relational.* It can connect one perceptual categorization to another even in the absence of the stimuli that triggered these categorizations.[55]

Thus to have concepts involves not only the process of categorization but also recognizing relationships *between* categories.[56] What is also important is that having concepts is a capacity that is not limited to humans, a point also made by Donald Griffin.[57] Concepts are thus not necessarily tied to language.

54. For a recent and stimulating sally in this debate, see Jerry A. Fodor, *Concepts: Where Cognitive Science Went Wrong* (Oxford: Clarendon Press, 1998).

55. Gerald M. Edelman, *The Remembered Present: A Biological Theory of Consciousness* (New York: Basic Books, 1989), 141.

56. Other works that either make a strong connection between categories and concepts or assert their equivalence include James Hampton and Danièle Dubois, "Psychological Models of Concepts," in *Categories and Concepts: Theoretical Views and Inductive Analysis*, ed. Iven van Mechelen, James Hampton, Ryszard S. Michalski, and Peter Theuns, Cognitive Science Series (London: Academic Press, 1993), 11–33; Barsalou, "Flexibility, Structure, and Linguistic Vagary in Concepts"; Barsalou et al., "Concepts and Meaning"; Smith and Medin, *Categories and Concepts*; and Gregory L. Murphy and Douglas L. Medin, "The Role of Theories in Conceptual Coherence," *Psychological Review* 92 (1985): 289–316.

57. Donald R. Griffin, *Animal Minds* (Chicago: University of Chicago Press, 1992), chap. 6. A similar approach can be seen in the work of Douglas Hofstadter and his associates, which has resulted in a number of computer programs that deal with relatively compact bundles of information they call "concepts"; see Hofstadter and the Fluid Analogies Research Group, *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought* (New York: Basic Books, 1995).

With this perspective in mind, let me propose that a musical concept has three characteristics. First, it is a product of the process of categorization. A musical category, then, is quite literally where our conceptualization of music begins. Second, a musical concept is an essential part of the means through which we guide present and future actions. These actions thus constitute a sort of indirect evidence for a cognitive structure almost as ephemeral as music itself. Third, a musical concept can be related to other concepts, including concepts associated with bodily states (both physical and emotional), perceptual categories (including sound, which, after all, is not necessarily music), and linguistic constructs.

From the perspective provided by this definition, language—in the sense of natural language—is not required in order to have musical concepts. I have, of course, used language to characterize musical categories, as well as the conceptual models around which musical categories are organized. But we could imagine a listener—perhaps a particularly astute listener, but perhaps one no more skilled than Proust's M. Swann—who develops a musical category without recourse to language. While language is still nearly indispensable for communicating features of musical concepts and for developing such concepts in richer contexts, it is not required.[58] The notion of concepts independent of language is more than a little provocative, and with consequences that extend beyond music and the general debate on "concept" noted above. Where some authors have excluded music from the conceptual realm (because, by their definition, concepts require language),[59] music can join dance, the visual arts, and any of the other nonlinguistic modes of human expression as a properly conceptual activity.

This perspective on musical understanding also offers a new way to think about musical syntax. Music theory has traditionally (if not explicitly) characterized syntax in terms of relationships between categories of musical objects or events. For instance, a perfect authentic cadence, as it is usually defined, consists (in part) of a dominant chord followed by a tonic chord, with bass motion from $\hat{5}$ to $\hat{1}$, and with the soprano concluding an octave above the bass. The definition involves categories like scale step ($\hat{5}$ and $\hat{1}$), harmonic function (dominant and tonic), voice (bass and soprano), and interval (the octave). Members of these categories, when placed in specific relationships (succession and "above"), produce a component of musical syntax (the perfect authentic cadence).[60] Musical syntax can also be conceived of

58. The idea that language was not required for musical analysis was one of the ideas behind Hans Keller's wordless analyses. See Keller, "Wordless Functional Analysis no. 1: Mozart, K. 421," *The Score and I.M.A. Magazine* 22 (1958): 56–64; idem, "The Musical Analysis of Music," in *Essays on Music*, ed. Christopher Wintle (Cambridge: Cambridge University Press, 1994), 126–28; idem, "Functional Analysis no. 9a: Mozart's Piano Sonata in A Minor, K. 310," in *Essays on Music*, ed. Christopher Wintle (Cambridge: Cambridge University Press, 1994), 129–38.

59. Ray Jackendoff, *Consciousness and the Computational Mind* (Cambridge, Mass.: MIT Press, 1987); Mark DeBellis, *Music and Conceptualization* (Cambridge: Cambridge University Press, 1995).

60. It is worth noting that Type 2 categories are chiefly involved in this definition: within a given context, something either is or is not an example of $\hat{5}$; something either is or is not a dominant chord; and so on. Such categories are typical of the sorts of taxonomies established by formal music theory. However, note that at least one category (for voice) is less explicitly defined and functions more like a Type 1 category. For instance, the soprano is, by convention, the highest sounding voice within a multivoice texture, but in instrumental works it may actually sound well below the range of human sopranos (and thus be a somewhat atypical example of a soprano voice).

as an emergent property of sequences of musical events. Wagner's treatment of the Leidensmotiv offers one example of this, but the same can be said of most of the musical works that interest us: it is through the development of local levels of syntax that musical works become distinctive, and it is the multivalence of this syntax that makes repeated listening to the same work a rewarding experience. Of course, syntactic structure that is shared between musical works is also important, and in most music there is a dialogue between the two analogous to the dialogue between the local and global models relative to which musical categories are framed.[61] Elaborating this approach to musical syntax is the main goal of chapter 4. There we will see how composers rely on syntactic conventions to provide a framework for the development of more local patterns of syntax specific to individual works.

Before we can more fully investigate the role musical concepts play in understanding and the way musical syntax comes about, we need to consider two important issues. First, there is the issue of how relationships between concepts are established. This will be taken up in chapter 2, which explores how conceptual domains such as music and language are correlated through the process of cross-domain mapping. This process allows structure from one conceptual domain to be brought to bear on another and for relationships among concepts to be established. Second, there is the issue of where conceptual models come from and how they are used to guide reasoning. Chapter 3 shows how concepts and conceptual relationships are combined into conceptual models and how conceptual models are connected together to create theories of music.

61. My perspective on musical syntax owes a debt to Eugene Narmour's construal of musical style. Narmour distinguishes between intraopus style (which is specific to a given work) and extraopus style (which can be seen across a number of works). See Narmour, *The Analysis and Cognition of Basic Melodic Structures*; and idem, *The Analysis and Cognition of Melodic Complexity: The Implication-Realization Model* (Chicago: University of Chicago Press, 1992).

CHAPTER TWO

# CROSS-DOMAIN MAPPING

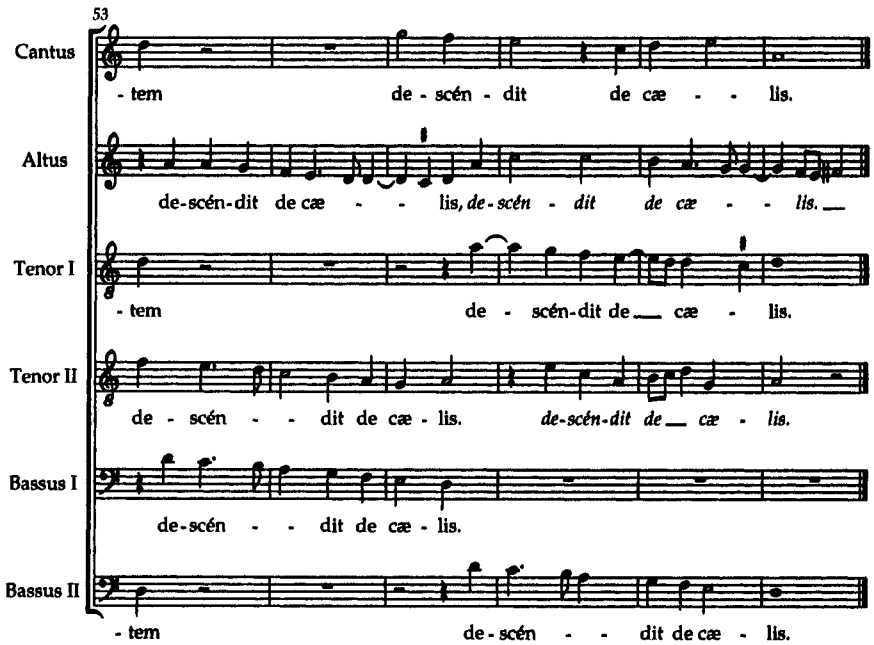
About a quarter of the way through the Credo of his *Pope Marcellus Mass* (printed 1567), Giovanni Pierluigi da Palestrina indulges in a marvelous bit of text painting with telling effect. The text Palestrina sets here is "Qui propter nos homines et propter nostram salutem descendit de cælis" ("Who for us men, and for our salvation, came down from heaven"). As shown in example 2.1, with the first statement of the word "descendit," each voice begins a scalar descent. Christ's descent from heaven is thus represented with a cascading fall through musical space, a series of overlapping movements "down" the musical scale.

Palestrina's text painting is a striking embodiment of the conventional construal of pitch as "high" and "low." This way of thinking about pitch relations has a venerable tradition. Its basic elements can be seen in Aristoxenus's description of pitch relations in terms of two-dimensional space, which was discussed in the introduction to this volume. Pitches are understood as points in space, and musical intervals are reckoned in terms of distances between these points. However, the spatial orientation crucial to Palestrina's text painting—the correlation of the "up" and "down" of physical space with specific pitch relations, such that a musical scale can "descend"—was seldom used in Greek music theory. As Andrew Barker has noted, the standard Greek for what would now be called "high-pitched" is *oxys*, which meant "sharp," "pointed," or "keen-edged"; its musical opposite was *barys*, which meant "heavy" (but not, in opposition to *oxys*, "blunt").[1] Just when "up" and "down" consistently came to be correlated with musical pitch is unclear, but the linkage was in place at least by the beginning of the tenth century, a good six hundred years before Palestrina wrote his Mass. From around the tenth century, thus, musicians in the West began writing about and depicting pitch in terms of "high" and "low," mapping structural relations from the domain of vertically oriented, two-dimensional space onto the domain of music.

Perhaps more remarkable than the long tradition of construing pitch relations in terms of "up" and "down" are the ready reminders of how arbitrary a construal it is.

1. Andrew Barker, ed., *Harmonic and Acoustic Theory* (vol. 2 of *Greek Musical Writings*), Cambridge Readings in the Literature of Music (Cambridge: Cambridge University Press, 1989), 69, n. 2.

EXAMPLE 2.1 Giovanni Pierluigi da Palestrina, Credo of the *Pope Marcellus Mass*, mm. 53–58




Consider, for instance, "up" and "down" on the piano: how can D4 be "above" C4 on the piano when they are both on the same horizontal plane? Think of playing the two notes on the cello—to play the "higher" D4, we have to move our left hand "down," so that it is closer to the ground. In fact, there are countless reminders of the artificiality of describing pitch in terms of "high" and "low." And yet Palestrina's text painting seems anything but arbitrary—there seems to be an aptness to his portrayal of the descent from heaven that goes beyond mere traditions of depiction.

In this chapter, I explore the function of cross-domain mappings of the sort that underlie this small bit of text painting in the Credo of the *Pope Marcellus Mass*. Cross-domain mapping plays two important roles in musical understanding: first, it provides a way to connect musical concepts with concepts from other domains, including those associated with language; second, it provides a way to ground our descriptions of elusive musical phenomena in concepts derived from everyday experience. Both of these contribute to the establishment of relationships between concepts, relationships that are fundamental to the prospect of theorizing about music. In the first section that follows, I provide an introduction to the theory of cross-domain mapping as it has been developed in recent work by cognitive linguists and others. This introduction provides a framework for explaining what makes mappings possible and why some mappings are more effective than others.

Cross-domain mapping also makes it possible to correlate the musical domain with others, such as the domain of physical space or of gesture. Under certain circumstances, such correlations provide the basis for rich worlds of the imagination.

Correlating musical pitches with vertically oriented, two-dimensional space, for instance, leads quite naturally to an imaginary world in which pitches become things that move through space: the successive notes of a scale gradually descend and ascend; in other passages, some notes leap, while still others fall. Within this imaginary world, each traversal of space has a specific and unmistakable sound—that is, descent sounds one way, ascent another. And this is not something limited to text painting of the sort demonstrated by Palestrina, as any number of cartoon soundtracks confirm. Never mind that actual traversals of space sound nothing like those of the hyperkinetic Roadrunner or the hapless Wile E. Coyote; if the correlation between the domains is properly established, elements from each will blend together to create novel relationships and elements. In the second section that follows, I describe the process that leads to this sort of blending and show the role conceptual blending plays in text painting and program music.

## AN INTRODUCTION TO CROSS-DOMAIN MAPPING

### Cross-Domain Mapping and Metaphor

The theory of cross-domain mapping is a product of a generalized approach to linguistic metaphor first taken by George Lakoff and Mark Johnson in 1980. Perhaps the most common conception of metaphor is of a literary device, a manifestation of the figural use of language to create colorful if imprecise images. Lakoff and Johnson accumulated a substantial body of evidence demonstrating that metaphor was not simply a manifestation of literary creativity but was, in fact, pervasive in everyday discourse.[2] As an example, consider the way *up* and *down* are used to characterize emotions, consciousness, and health:

EMOTIONS

I'm feeling *up*. My spirits *rose*. I'm feeling *down*. I *fell* into a *depression*. My spirits *sank*.

CONSCIOUSNESS

Get *up*. I'm *up* already. He *rises* early in the morning. He *fell* asleep.

HEALTH

He's at the *peak* of health. She's in *top* shape. He came *down* with the flu.

Each characterization suggests not a literal representation of the spatial domain implied by the orientation *up–down* but, instead, uses our knowledge of physical space to structure our understanding of emotions, consciousness, or health.

Based on evidence provided by a large number of similar examples of the

2. George Lakoff and Mark Johnson, *Metaphors We Live By* (Chicago: University of Chicago Press, 1980). Their work has recently been extended in Lakoff and Johnson, *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought* (New York: Basic Books, 1999).

appearance of metaphorical constructions in everyday discourse, Lakoff and Johnson proposed that metaphor was a basic structure of understanding through which we conceptualize one domain (the target domain, which is typically unfamiliar or abstract) in terms of another (the source domain, which is most often familiar and concrete). Further study has provided a wealth of empirical evidence for this proposal and contributed to the development of the field of cognitive linguistics.[3]

Fundamental to the theory of metaphor that sprang from Lakoff and Johnson's work is a distinction between conceptual metaphors and linguistic metaphors. A conceptual metaphor is a cognitive mapping between two different domains; a linguistic metaphor is an expression of such a mapping through language. For instance, the conceptual metaphor STATE OF BEING IS ORIENTATION IN VERTICAL SPACE maps relationships in physical space onto mental and physical states.[4] The cross-domain mapping wrought by this conceptual metaphor then gives rise to innumerable linguistic expressions. Some of these expressions are commonplace, such as "Maxwell seems a bit down today." Others summon a rich imagistic world, such as that of John Keats's "Ode to a Nightingale":

> My heart aches, and a drowsy numbness pains
> My sense, as though of hemlock I had drunk,
> Or emptied some dull opiate to the drains
> One minute past, and Lethe-wards had sunk.

Here the descent to the mythical river gives a physical correlate to the narcotic state of the narrator: the act of sinking is mapped onto a melancholy emotional state. Thus the same conceptual metaphor (STATE OF BEING IS ORIENTATION IN VERTICAL SPACE) is behind both linguistic metaphors, one commonplace ("Maxwell seems a bit down today"), the other poetic.

With respect to music, the "high" and "low" used to describe pitches reflect the conceptual metaphor PITCH RELATIONSHIPS ARE RELATIONSHIPS IN VERTICAL SPACE. This metaphor maps spatial orientations such as *up–down* onto the pitch continuum. The mapping yields a system of metaphors replete with possibilities for describing musical pitch. We can speak in terms of pitch contour (meaning successions of pitches, which are located at different places in pitch-space), gesture (mean-

3. For a review of the empirical evidence supporting metaphor as a basic cognitive process, see Raymond W. Gibbs, Jr., *The Poetics of Mind: Figurative Thought, Language, and Understanding* (Cambridge: Cambridge University Press, 1994). For discussion of the link between the study of metaphor as a cognitive process and the central concerns of cognitive linguistics, see George Lakoff, "The Invariance Hypothesis: Is Abstract Reason Based on Image-Schemas?" *Cognitive Linguistics* 1 (1990): 39–51.

4. By convention, conceptual metaphors are represented in capital letters. Thus LOVE IS A JOURNEY designates a conceptual metaphor of general application, and "Love is a journey" a specific linguistic expression based on that metaphor.

The conceptual metaphor STATE OF BEING IS ORIENTATION IN VERTICAL SPACE is a variant of the STATES ARE LOCATIONS conceptual metaphor discussed by George Lakoff and Mark Turner in *More Than Cool Reason: A Field Guide to Poetic Metaphor* (Chicago: University of Chicago Press, 1989). On cross-domain mapping as a general phenomenon, see Lakoff and Turner, *More Than Cool Reason*, 4; George Lakoff, "The Contemporary Theory of Metaphor," in *Metaphor and Thought*, 2nd ed., ed. Andrew Ortony (Cambridge: Cambridge University Press, 1993), 202–51; Gibbs, *The Poetics of Mind*; and Gilles Fauconnier, *Mappings in Thought and Language* (Cambridge: Cambridge University Press, 1997).

ing successions of pitches with a specific directionality and contour), and musical space (meaning a three- or four-dimensional extension of the basic two-dimensional mapping).[5] This system is given graphic representation in traditional musical notation: notes that are the result of more rapid vibrations of the sounding medium are placed higher on the page than notes that result from less rapid vibrations (with the exception of sharps and flats). The two-dimensional space of the musical page thus correlates with the spatial orientation ascribed to pitch.[6] The systematic quality that results from mapping spatial orientations onto the pitch continuum thus leads to an entire vocabulary for describing relationships among pitches that provides a rich set of possibilities for furthering our conceptualization of music.

As common as conceiving of pitches as "high" or "low" seems, not all cultures describe pitch relationships in purely spatial terms. As I noted here, Greek theorists of antiquity used *oxys* ("sharp" or "pointed") and *barys* ("heavy") to characterize pitches. And traversing historical distance is not the only way to discover alternative conceptualizations of pitch relations. Consider three examples in which it is culture, rather than time, that creates distance:

1. Steven Feld's research has shown that the Kaluli of Papua New Guinea describe melodic intervals with the same terms they use to characterize features of waterfalls. For instance, in Kaluli *sa* means "waterfall," and a *mogan* is a still or lightly swirling waterpool; *sa-mogan* is the flow of a waterfall into a level waterpool beneath it. *Sa-mogan* is also used to describe a melodic line that descends to a repeated note, the contour of which replicates that of a waterfall flowing into a pool. In contrast, there are no specific names for ascending intervals, which nonetheless do occur in Kaluli song.[7] Behind this account of musical intervals is the conceptual metaphor PITCH RELATIONSHIPS ARE WATERFALL CHARACTERISTICS, which provides the basis for a rich set of descriptive terms that capture some aspects of melody but not others.
2. In Bali and Java pitches are conceived not as "high" and "low" but as "small" and "large."[8] Here the conceptual metaphor is PITCH RELATIONSHIPS ARE RELATIONSHIPS OF PHYSICAL SIZE, a mapping that accurately

5. For further discussion of mappings between the spatial and pitch domains, see Arnie Walter Cox, "The Metaphoric Logic of Musical Motion and Space" (Ph.D. diss., University of Oregon, 1999).

6. It is worth noting that the spatial domain of the musical page is itself metaphorical, since the way the page is oriented in physical space—that is, whether it is propped on a music stand or lying flat on a table, turned right side up or upside down—does not change what we regard as the top and bottom of the page. Two factors, each independent of the actual orientation of the page in physical space, help establish the spatial domain of the page: the conventional orientation of the symbols on the page (either "rightside up" or "upside down") and the relative distance of the symbols on the page from the reader. Rightside up symbols that are farthest from the reader are at the "top" of the page; rightside up symbols that are nearest the reader are at the "bottom" of the page.

7. Steven Feld, "Flow Like a Waterfall: The Metaphors of Kaluli Musical Theory," *Yearbook for Traditional Music* 13 (1981): 30–31.

8. Personal communication from Benjamin Brinner, 8 July 1997. See also Wim van Zanten, "The Tone Material of the Kacapi in Tembang Sunda in West Java," *Ethnomusicology* 30 (1986): 85.

reflects the norms of acoustic production: small things typically vibrate more rapidly than large things. This acoustic fact is represented throughout the numerous parts of the gamelan, the collection of instruments central to the musical practice of Bali and Java.

3. The Suyá of the Amazon basin do not have an extensive vocabulary for describing pitch relationships. When they are described, however, it is in terms of age: pitches are conceived not as "high" and "low" but as "young" and "old." The conceptual metaphor that guides this mapping is PITCH RELATIONSHIPS ARE AGE RELATIONSHIPS, which accurately reflects the Suyá's beliefs that the pitch of the voice becomes deeper with age.[9]

With each of these conceptions of pitch relationships it is easy to focus on what they lack rather than what they offer. That is, it is natural (if not quite excusable) to reckon difference in somewhat chauvinistic terms: that the Kaluli do not have terms to describe ascending intervals, where we do; that the way pitch relationships are characterized in Bali and Java does not transfer into graphic representations with the same ease as do "high" and "low." It has to be remembered, however, that mapping "high" and "low" onto music has its own limitations: "high" and "low" cannot reflect the subtle play of flowing water, nor do they provide much of an explanation for how acoustic features correlate with pitch relationships. Such differences show that each mapping between domains makes some conceptualizations possible, while it disables others.

## Image Schema Theory

The variety of conceptual metaphors used to characterize pitch relations leads to the question of the ultimate grounding of the process of cross-domain mapping. Even if we grant that we understand a target domain (such as pitch relationships) in terms of a source domain (such as orientation in vertical space), how is it that we understand the source domain in the first place? Mark Johnson answered this question by proposing that meaning was grounded in repeated patterns of bodily experience.[10] These patterns give rise to what Johnson called *image schemata*, which provide the basis for the concepts and relationships essential to metaphor. An image schema is a dynamic cognitive construct that functions somewhat like the abstract structure of an image and thereby connects together a vast range of different experiences that manifest this same recurring structure.[11] Image schemata are by no means exclusively visual—the idea of an image is simply a way of capturing the organization inferred from patterns in behavior and concept formation.

As one example of an image schema, consider the VERTICALITY schema, which might be summarized by a diagram of the sort given in figure 2.1. We grasp this structure repeatedly in thousands of perceptions and activities that we experience every day. Typical of these are the experiences of perceiving a tree, our felt sense of standing upright, the activity of climbing stairs, forming a mental image of a flagpole, and

9. Anthony Seeger, *Why Suyá Sing: A Musical Anthropology of an Amazonian People* (Cambridge: Cambridge University Press, 1987), 100–02.

10. Mark Johnson, *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason* (Chicago: University of Chicago Press, 1987).

11. Johnson, *The Body in the Mind*, 2.

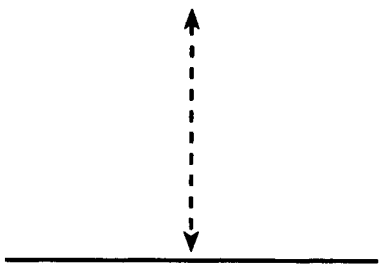


FIGURE 2.1 Diagram of the VERTICALITY schema

watching the level of water rise in the bathtub. The VERTICALITY schema is the abstract structure of the VERTICALITY experiences, images, and perceptions. Our concept of verticality is based on this schema, and this concept is in turn invoked by the various conceptual metaphors that use vertical space as a source domain through which to structure such target domains as emotions, consciousness, health, and musical pitch.

By definition, image schemata are preconceptual: they are not concepts, but they provide the fundamental structure upon which concepts are based. In consequence, it is important to emphasize that any diagram used to illustrate an image schema is intended to represent the key structural features and internal relations of the schema; it is not meant to summon a rich image or mental picture that we somehow have "in mind" and use actively to structure our thought. More directly, whatever actually occupies our thoughts is not, by definition, an image schema. We can conceive *of* image schemata, just as we can conceive of any of a number of nonconceptual or preconceptual cognitive processes. We can also note general patterns in the way concepts are structured, which can be attributed to image schemata. However, there are, by definition, no "image-schema concepts."

The relationship between the VERTICALITY schema and our characterization of musical pitch with reference to the spatial orientation *up–down* is fairly immediate: when we make *low* sounds, our chest resonates; when we make *high* sounds, our chest no longer resonates in the same way, and the source of the sound seems located nearer our head. The "up" and "down" of musical pitch thus correlate with the spatial "up" and "down"—the vertical orientation—of our bodies. The VERTICALITY schema offers a straightforward way to explain why we characterize musical pitch in terms of *high* and *low* even when the actual spatial orientation of the means through which we produce pitches either does not reinforce the characterization or runs directly counter to it.

At present, the image schema remains largely a theoretical construct. Work across a variety of fields, however, has made strong arguments for the importance of such a construct, including that by Leonard Talmy in linguistics, Gerald Edelman in neuroscience, David McNeill in psychology, and Raymond W. Gibbs Jr. and Herbert L. Colston in psycholinguistics.[12] Recently, Lawrence Barsalou and his associates have

12. Leonard Talmy, *Concept Structuring Systems* (vol. 1 of *Toward a Cognitive Semantics*) (Cambridge, Mass.: MIT Press, 2000), 409–70; Gerald M. Edelman, *The Remembered Present: A Biological Theory of Consciousness* (New York: Basic Books, 1989), chap. 8; David McNeill, *Hand and Mind: What Gestures Reveal about Thought* (Chicago: University of Chicago Press, 1992); Raymond W. Gibbs, Jr., and Herbert L. Colston, "The Cognitive Psychological Reality of Image Schemas and Their Transformations," *Cognitive Linguistics* 6 (1995): 347–78.

developed a comprehensive theory of mental representation based on perceptual symbols (which are analogical structures similar to image schemata) and have begun to provide empirical work to support the theory.[13] Antonio Damasio, working from a neurophysiological perspective, has made compelling arguments for the importance of the body to consciousness and thought.[14] Although research on image schemata and similar structures is still preliminary, it is also highly promising and offers some of the best prospects for solving some of the problems of the relationship between mind and body that have dogged cognitive research throughout this century.

## The Invariance Principle

The theory of image schemata provides a way to explain how conceptual metaphors are grounded. It does not, by itself, explain why some conceptual metaphors seem intuitively better than others. For instance, the conceptual metaphor PITCHES ARE FRUIT could provide the grounding for such expressions as "You must play the first note more like an apple, the second more like a banana." Although such mappings are possible, they are certainly not common. Pitches and fruits just do not seem to be a good match.

To account for why some metaphorical mappings are more effective than others, George Lakoff and Mark Turner proposed that such mappings are not about the *imposition* of the structure of the source domain on the target domain, but are instead about the establishment of correspondences between the two domains. These correspondences are not haphazard, but instead preserve the image-schematic structure latent in each domain. Lakoff and Turner formalized this perspective with the Invariance Principle, which Turner states as follows: "In metaphoric mapping, for those components of the source and target domains determined to be involved in the mapping, preserve the image-schematic structure of the target, and import as much image-schematic structure from the source as is consistent with that preservation."[15] According to the Invariance Principle, mapping the spatial orientation

13. Lawrence W. Barsalou, Wenchi Yeh, Barbara J. Luka, Karen L. Olseth, Kelly S. Mix, and Ling-Ling Wu, "Concepts and Meaning," in *Chicago Linguistics Society 29: Papers from the Parasession on the Correspondence of Conceptual, Semantic, and Grammatical Representations*, ed. Katharine Beals, Gina Cooke, David Kathman, Sotaro Kita, Karl-Erik McCullough, and David Testen (Chicago: University of Chicago, Chicago Linguistics Society, 1993), 23–61; Lawrence W. Barsalou, Karen Olseth Solomon, and Ling-Ling Wu, "Perceptual Simulation in Conceptual Tasks," in *Cultural, Psychological, and Typological Issues in Cognitive Linguistics: Selected Papers of the Bi-Annual ICLA Meeting in Albuquerque, July 1995*, ed. Masako K. Hiraga, Chris Sinha, and Sherman Wilcox (Amsterdam: John Benjamins, 1999), 209–28; Lawrence W. Barsalou, "Perceptual Symbol Systems," *Behavioral and Brain Sciences* 22 (1999): 577–609.

14. Antonio R. Damasio, *The Feeling of What Happens: Body and Emotion in the Making of Consciousness* (New York: Harcourt Brace, 1999), chaps. 5 and 6.

15. Mark Turner, "Aspects of the Invariance Hypothesis," *Cognitive Linguistics* 1 (1990): 254. For additional writings on the Invariance Principle (which at first was called the Invariance Hypothesis), see Lakoff, "The Invariance Hypothesis"; Mark Turner, "An Image-Schematic Constraint on Metaphor," in *Conceptualizations and Mental Processing in Language*, ed. Richard A. Geiger and Brygida Rudzka-Ostyn, Cognitive Linguistics Research, 3 (Berlin: Mouton de Gruyter, 1993), 291–306; and Mark Turner, *The Literary Mind* (New York: Oxford University Press, 1996), chap. 3.

A preliminary discussion of a similar sort of topographical invariance, with applications to music, can

*up–down* onto pitch works because of correspondences between the image-schematic structure of components of the spatial and acoustical domains. Both space and the frequency spectrum are continua that can be divided into discontinuous elements. In the spatial domain, division of the continuum results in points; in the acoustic domain, it results in pitches. Mapping *up–down* onto pitch allows us to import the concrete relationships through which we understand physical space into the domain of music and thereby provide a coherent account of relationships between musical pitches. Mapping various fruits onto musical pitches works less well because fruit does not (in any ordinary way) constitute a continuum. To employ this mapping is to highlight instead both the discontinuity among musical pitches and how they are *unlike* one another (an emphasis on difference suggested by the idiomatic phrase "like apples and oranges").

## Cross-Domain Mapping and Conceptual Models

According to current theory, cross-domain mappings are grounded in repeated patterns of embodied experience called image schemata. These schemata provide the basic structure employed in the mappings: the VERTICALITY schema is thus fundamental to our understanding of two- or three-dimensional spaces as having directionality and of musical pitch as "high" and "low." Image schemata also constrain the possibilities for mapping between two domains, a constraint reflected in the Invariance Principle. Because the VERTICALITY schema can be applied to both the spatial and the musical domains, we can use our understanding of the former to structure our understanding of the latter. It remains to be explained why one mapping would be preferred over another—why, for instance, we tend to describe pitch relations in terms of "high" and "low" rather than "small" and "large."

On the face of it, both of these mappings are equally viable. Both draw on aspects of our embodied experience: on the one hand, countless experiences with the seeming origin of our own voices; on the other hand, countless experiences with the sounds given out by physical objects in the world around us. Both mappings allow us to describe musical pitches as elements within a continuum. And each mapping can be easily understood from the perspective provided by the other. For instance, in Camille Saint-Saëns's *Le Carnaval des animaux*, the music for the elephant is played by the contrabass (*large* is *low*), that for the birds by the flute (*small* is *high*).[16] Although musicians educated in the West sense the novelty of the mapping, they can nonetheless understand it perfectly well. In a similar fashion, musicians from Bali and Java, when confronted with Western conventions for notation, have few if any problems translating "small" and "large" into "high" and "low."[17]

---

be found in Peter Gärdenfors, "Semantics, Conceptual Spaces, and the Dimensions of Music," in *Essays on the Philosophy of Music*, ed. Veikko Rantala, Lewis Rowell, and Eero Tarasti, Acta Philosophica Fennica, 43 (Helsinki: Philosophical Society of Finland, 1988), 9–27.

16. Camille Saint-Saëns, *Le Carnaval des animaux: Grande fantasie zoologique*, ed. Felix Aprahamian (Zurich: Eulenberg, 1974): "L'éléphant" (no. 5), p. 11, "Volière" (no. 10), pp. 23–27.

17. Personal communication from Sumarsam, 10 April 1998.

The reason we prefer one mapping over another has to do with the global conceptual models we absorb from culture and that supply crucial support for the preferred mapping.[18] In the West, the description of pitch relations in terms of "up" and "down" arose around the same time musicians began to develop ways of notating polyphonic compositions. These notational systems often relied, either directly or indirectly, on the physical placement of symbols on the page.[19] The attribution of "high" and "low" to musical pitches is thus correlated with a system of notation that permitted both the visualization and the preservation of musical works. In turn, this notation relied on a global model that made three important assumptions: first, pitches could be regarded as objects that were independent of the sound source that produced them; second, graphical symbols could be used to represent these pitch-objects; and third, the surface on which these symbols appeared was analogous to physical space. In Bali and Java, the performance of music was associated with the rich palette of instruments through which the music was effected. Thus the attribution of "small" and "large" to pitches correlated with characteristics of the musical instruments intrinsic to musical performance. The conception of musical pitches as physical objects relies on a global model that does not, at some fundamental level, disassociate a pitch from the sound source that produces it.[20]

As noted earlier, the characterization of pitch relations can be informed by mappings other than "high" and "low" and "small" and "large." In each case, the basic mapping relies on embodied knowledge and on the correlation of the musical domain with a more concrete domain. The specific mapping chosen within a tradition of discourse about music reflects not so much absolute musical structure as it does the broader cultural practice within which music and its understanding are embedded: mappings reflect the conceptual models that are important to culture. The cross-domain mappings employed by any theory of music are thus more than simple curiosities—they are actually key to understanding music as a rich cultural product that both constructs and is constructed by cultural experience.

18. These models and the relationship between global and local models will be discussed in greater detail in the next chapter.

19. With regard to an indirect reliance on physical placement, it was often a part of the pedagogy of *daseian* notation (developed in the ninth century, apparently with an eye toward accommodating the needs of *organum* at the fourth) to place "high" notes higher on the page, even though no such discrimination was required by the notation (since it relied strictly on a limited repertoire of letter-like symbols to indicate different pitches). See *Musica et Scolica enchiriadis una cum aliquibus tratatulis adiunctis*, ed. Hans Schmid, Veröffentlichungen der Musikhistorischen Kommission, Bayerische Akademie der Wissenschaften, 3 (Munich: Verlag der Bayerische Akademie der Wissenschaften, 1981); and the *Musica enchiriadis* and *Scolica enchiriadis*, ed. Claude V. Palisca, trans. Raymond Erickson, Music Theory Translation Series (New Haven, Conn.: Yale University Press, 1995). These treatises are currently thought to have been written no later than the tenth century and possibly as early as the middle of the ninth. For further discussion of the history of this notational convention, see Marie-Elisabeth Duchez, "La Représentation spatio-verticale du caractère musical grave-aigu et l'élaboration de la notion de hauteur de son dans la conscience musicale occidentale," *Acta musicologica* 51 (1979): 54–73.

20. It should be observed that a disassociation of pitch from sound does not obviate an image-schematic basis for understanding pitch relations. It only bears witness to a reification of pitch after its initial conceptualization.

## Cross-Domain Mapping, Music, and Music Theory

PALESTRINA'S TEXT PAINTING The theory of cross-domain mapping outlined in the preceding discussion provides a relatively simple way to account for the effectiveness of Palestrina's text painting. According to this theory, we use the VERTICALITY schema, the product of countless bodily experiences, to give physical space an "up" and a "down" analogous to the "up" and "down" of our bodies. This orientation is then mapped on to a metaphorical musical space made up of pitches and the relations (or distances) between them. Pitches that result from more rapid vibrations of the sounding medium are regarded as "higher" than pitches that result from less rapid vibrations of the sounding medium. As each of the voices in Palestrina's six-part texture takes up the word "descendit," it begins to "descend" through musical space. The notion of descent summoned by the text is thus given sounding image by a specific series of musical events.

Of course, we could also hear (through the ears of a Javanese musician) the pitches growing larger, or hear (through the ears of the Suyá) the pitches growing older, or hear (through the ears of a Greek theorist) the pitches growing heavier. Distinct from these ways of hearing, Palestrina's text painting relies on a conceptual model that characterizes pitches as objects in two-dimensional, vertically oriented space, in which "up" and "down" describe specific pitch relations. This model is strongly associated with the conventions of musical notation as developed in the West: as pitches get "lower" in sound, they are also written "lower" on the musical staff.

There is an additional reason why Palestrina's text painting is as convincing as it is, however. The image of descent created by Palestrina relies on more than just the general notion of descent, which could have been evoked in a variety of ways: by a downward leap of a single interval; by a short sequence that alternated descending thirds with ascending seconds; or through having each voice enter in turn, beginning with the highest voice and ending with the lowest. The scalar descent chosen by Palestrina, however, provides a striking analogue for the descent of our bodies through physical space (when this descent is unaided by artificial means). Such a descent involves a lessening of potential energy and a continuous action in one direction, articulated by the regular transfer of weight from one leg to another.[21]

21. Most of us find descent—especially scalar descent—well represented by the thought of walking down a staircase, but I think it could be argued that walking down a hillside works as well. The reason is that it is not so much the neat, two-dimensional image of stairs that is operative but the regular transfer of weight from one leg to another. Indirect evidence is provided by a striking anecdote related by John Hockenberry, from a time when he was a reporter for National Public Radio. In order to get to a group of Kurdish refugees on a remote edge of Iraqi Kurdistan during the aftermath of the 1991 Gulf War, Hockenberry, a paraplegic since 1976, had to abandon his wheelchair temporarily and ride on a donkey. He comments on the effect of becoming reacquainted with a non-wheeled mode of transportation through the rhythm of the donkey's gait: "It was walking, that feeling of groping and climbing and floating on stilts that I had not felt for fifteen years. I had long ago grown to love my own wheels and their special physical grace, and so this clumsy leg walk was not something I missed until the sensation came rushing back through my body from the shoulders of a donkey." Hockenberry, *Moving Violations: War Zones, Wheelchairs, and Declarations of Independence* (New York: Hyperion, 1995), 2–3.

The act of singing a descending scale correlates well with the basic structure of this event: the relaxation many singers feel as they sing a descending scale matches the lessening of potential energy; the temporary pauses on each note of the scale match the regular transfer of weight, which articulates a physical descent. The text painting of "descendit" is thus supported by our embodied knowledge of descent, as well as by the conventions of ascribing "high" and "low" to musical pitches.

In its exploitation and enrichment of a mapping between the domains of physical space and music, Palestrina's text painting also gives us a glimpse into a process of meaning construction. Conventionally, text painting is understood to operate through a crude sort of mimesis: physical descents are represented by musical descents. By contrast, I have argued that hearing a succession of musical events as a descent is an act that is thoroughly mediated.[22] What mimesis there is is highly conditioned by the choice of cross-domain mappings through which discourse about music is structured; in turn, these mappings reflect the global models of a given cultural perspective and historical moment. Palestrina's text painting is not just woven into this web of meaning construction; it also spins its own threads. Some result from specific features of physical descents summoned by the passage, some from the point within the larger musical and dramatic discourse at which this striking moment occurs, and some from the sonic attributes that get mapped back onto the notion of physical descent. The meaning constructed is not, in the final analysis, simple or direct but multivalent and contingent, and it reflects the rich set of correspondences activated by mapping between the two domains.

CROSS-DOMAIN MAPPING AND MUSIC THEORY Because it provides a way to bring an integrated system of terms and structural relations to bear on problems of musical understanding, cross-domain mapping plays an important part in theories of music. Indeed, every theory of musical organization employs cross-domain mappings of one sort or another. Often, the appeal of such mappings is strong, and the mappings seem intuitively correct (much as "high" and "low" seem intuitively correct for the characterization of pitch relations). Further investigation, however, reveals that these mappings are every bit as mediated as those that are less systematically developed.

Rudolf Louis and Ludwig Thuille's characterization of tonality in their *Harmonielehre* provides a case in point. The relevant passage, which appears at the beginning of the first chapter, runs as follows:

> The unity within the diversity of all harmony is ensured through the law of *tonality*. This asserts that any succession of harmonies can become musically understandable only when each (independently appearing) *chord* is perceived to be in a specific relationship of dependence with respect to the *principal chord* that underlies the entire harmonic context. Both melodic and harmonic succession require a *center*, a point of repose, around which everything twists and turns. For *melodic* relationships, this stationary middle point is the *tonic note*, the Tone, whose central place in the scheme of the scale is made manifest by the melodic motions that we direct away from tonic and

22. This same argument, mutatis mutandis, applies to McClary's analysis of Robert Johnson's blues mentioned above in the preface.

> then back toward it again. The same role of a fixed middle point and resting place that, within melodic unity, falls to the tonic note, is played, with respect to harmony, by the *tonic chord*: the *consonant triad* that has the tonic note as its *root*.[23]

The view set forth by this characterization is that musical organization is analogous to that of the physical world. Music consequently obeys laws that are independent of human beings—"tonality" is, as it were, a natural force.[24] Musical understanding relies on apprehending the harmonic dependencies that reflect these immutable laws. The focal point of these dependencies is the center of the system—either the tonic note or the tonic chord—which acts as both a center of gravity and an axis of symmetry.[25] To support this perspective, Louis and Thuille appeal to the empirical evidence provided by actual melodic and harmonic progressions, which they believe show that every properly constructed melody turns around a single tonic note, and that every properly constructed harmonic progression turns around a single tonic chord.

Louis and Thuille's characterization of tonality does a marvelous job of capturing the compelling coherence wrought by well-composed tonal music. Given the evidence of ethnographic and historical studies of musical cultures, however, the idea that Western European tonality is a naturally occurring force seems doubtful, as would be the notion of a science aimed at discovering the laws behind this force. Also implicit in Louis and Thuille's theory is the concept that tonality exists apart from musical syntax: successions of musical events do not give rise to the law of tonality but serve only to provide evidence of its existence. According to this perspective, successions of musical events that do not provide evidence of tonal relations cannot be understood as music. Of course, such a view places rather profound restrictions on what counts as music.

Theorizing about music requires that we bring order, even if of a tenuous sort, to an ephemeral and often intangible domain. Cross-domain mapping aids this process by bringing systems of relationships to bear on the musical domain and by

23. "Die Einheit in der Mannigfaltigkeit aller Harmonie wird gewährleistet durch das Gesetz der *Tonalität*. Dieses sagt aus, daß irgendwelche Folgen von Harmonien nur dann musikalisch verständlich wirken können, wenn jeder (selbständig auftretende) *Accord* in einem bestimmten Verhältnis der Abhängigkeit von einem dem ganzen harmonischen Zusammenhang zugrunde liegenden *Hauptaccord* aufgefaßt wird. Wie die melodische Folge, so bedarf auch die Accordfolge eines *Centrums*, eines ruhenden Pols, um den sich alles 'dreht und wendet.' Für die *melodischen* Beziehungen ist dieser feststehende Mittlepunct die *Tonica*, der Ton, dessen centrale Stellung im Schema der Tonleiter dadurch zum Ausdruck gebracht wird, daß wir deren melodische Bewegung von ihm ausgehen und zu ihm wieder zurückkehren lassen. Dieselbe Rolle eines fixen Mittle- und Ruhepuncts, die innerhalb der melodischen Einheit der Tonica zufällt, spielt in *harmonischer* Hinsicht der *Tonica-Accord*: der *consonierende Dreiklang*, der Tonica zum *Grundton* hat." Rudolf Louis and Ludwig Thuille, *Harmonielehre*, 7th ed. (Stuttgart: Carl Grüninger Nachf. Ernst Klett, 1920), 7.

24. I should point out that Louis and Thuille are not concerned with tonality as it might be broadly defined, but with tonality as it is represented in Western European music of the eighteenth and nineteenth centuries. This is made all the more clear by their characterization of tonality in harmonic terms, something that has limited applicability to a fairly wide range of music currently recognized as tonal.

25. In the pages that follow this excerpt, Louis and Thuille explore in greater detail the idea that tonic is at the center of a symmetrical scheme of pitches and harmonies. See Louis and Thuille, *Harmonielehre*, 8–10.

correlating rich, if at times complicated, images with essential musical concepts. Such mappings are a way to both structure our understanding and extend it—indeed, much work in music theory is devoted to exploring the ramifications of these mappings. Nonetheless, each mapping reflects the cultural values and imperatives relative to which it is framed, not the verities of absolute musical structure. Tonality is not, in a simple way, a centric and symmetrical system, any more than it is a reflection of logic (as Hugo Riemann would have it), psychological energetics (as Ernst Kurth would have it), or gravity (as Arnold Schoenberg would have it).[26] Instead, tonality is a way of understanding musical organization. The cross-domain mappings through which notions such as tonality are articulated are not simply essential to theories of music; in fact, they constitute what counts as music in the first place.

## Summary

Cross-domain mapping is a general cognitive process through which we structure an unfamiliar or abstract domain in terms of one more familiar or concrete. Cross-domain mapping plays two important roles in musical understanding. First, it provides a way to connect musical concepts with concepts from other domains. As we saw here, pitch relations within the domain of music have been connected with concepts associated with vertical space, waterfalls, physical size, and human aging. Each such mapping made possible systematic accounts of the ways pitches related to one another. Second, cross-domain mapping allows us to ground our descriptions of elusive musical phenomena in concepts derived from everyday experience, since the structural relations basic to cross-domain mapping have their source in repeated patterns of bodily experience—that is, in image schemata.

As we have seen, the mappings we use to structure our discourse about music are not accidental but reflect two constraints. One constraint is the Invariance Principle, which proposes that the best cross-domain mappings are those that preserve as much of the image-schematic structure of both target and source domains as possible. The other constraint is provided by the global conceptual models relative to which cross-domain mappings are framed. Taken together, these constraints suggest that cross-domain mappings not only provide a way to structure our understanding of music but also shape our ideas about what we include under the rubric "music." For an element to count as musical, it must be able to serve as a target for the cross-domain mappings that guide our discourse about music.

Because cross-domain mapping offers a way to connect what are often elusive musical concepts with concepts from more concrete domains, and because these connections give rise to integrated systems of terms and relations, cross-domain mapping is essential to our theorizing about music. Indeed, as we shall see in the

26. Hugo Riemann, *Musikalische Logik: Hauptzüge der physiologischen und psychologischen Begründung unseres Musiksystems* (Leipzig: C. F. Kahnt, 1874); Ernst Kurth, *Die Voraussetzungen der theoretischen Harmonik und der tonalen Darstellungssysteme* (Bern: Akademische Buchhandlungen von Max Dreschel, 1913); Arnold Schoenberg, *Harmonielehre* (Leipzig: Verlagseigentum der Universal-Edition, 1911); trans. as *Theory of Harmony*, by Roy E. Carter (Berkeley: University of California Press, 1978).

next chapter, the concepts derived from processes of categorization and the relations created by cross-domain mapping provide the basic materials for music theory.

## CONCEPTUAL BLENDING

Palestrina's text painting relies on our understanding of pitches as "high" and "low" and on our experiences with physical descents down staircases, slopes, and hillsides to create a vivid aural representation of the text. However, the imaginary world summoned by Palestrina also extends beyond the immediate bounds of text and music. Using the basic correlation between text and music as a point of departure, we can enter an imaginary domain in which the pitches to which "descendit" is sung become objects descending through musical space. Within this domain, every physical descent is accompanied by sounds that result from a smooth transition from very rapid to less rapid vibrations of the sounding medium. This extension of Palestrina's imaginarium cannot be predicted simply from linking the domain of text with the domain of music. It results instead from blending elements and events from these two domains to create a new one with its own structures and relations, a domain populated by such things as pitch-objects and the sound of descent.

It must be admitted that an imaginary domain populated with pitch-objects moving through musical space is a somewhat rarefied one. Even so, the process of conceptual blending through which it comes about is itself exceedingly common. For instance, concepts about humans and concepts about animals are often brought together in children's stories to produce talking animals. Such creatures are powerful devices in storytelling, for they offer narrative possibilities beyond those offered by characters with only human or animal attributes. In a like fashion, the combination of musical concepts with those from other domains creates possibilities for meaning construction that reach far beyond those of music alone. As an introduction to a methodology for exploring such conceptual blends in greater detail, let us turn to one of the talking animals from recent literature and discover what he can tell us about the process behind the imaginary world summoned by Palestrina's text painting.

### Talking Animals and Conceptual Integration Networks

> The Old Grey Donkey, Eeyore, stood by himself in a thistly corner of the forest, his front feet well apart, his head on one side, and thought about things. Sometimes he thought sadly to himself, "Why?" and sometimes he thought, "Wherefore?" and sometimes he thought, "Inasmuch as which?"—and sometimes he didn't quite know what he *was* thinking about. So when Winnie-the-Pooh came stumping along, Eeyore was very glad to be able to stop thinking for a while, in order to say "How do you do?" in a gloomy manner to him.[27]

Thus A. A. Milne introduces a new character into the stories he crafted for his son, in this case a character built around a stuffed toy donkey Christopher Robin had

27. A. A. Milne, *The World of Pooh: The Complete Winnie-the-Pooh and the House at Pooh Corner*, illustrated by E. H. Shepard (New York: E. P. Dutton, 1957), 45–46.

received as a Christmas present. Eeyore is, unquestionably, a donkey, fond of thistles and content to stay outside in all sorts of weather. But Eeyore is also endowed with human characteristics: he is able to talk, to be perversely gloomy, to work his own convoluted chains of logic, and, on occasion, to be capable of exquisite irony. In fact, Eeyore is a blend of some of the concepts associated with donkeys and with humans.

In order to study conceptual blends such as that represented by Eeyore, the rhetorician Mark Turner and the linguist Gilles Fauconnier developed the notion of *conceptual integration networks* (CINs).[28] Each CIN consists of at least four circumscribed and transitory domains called *mental spaces*. Mental spaces temporarily recruit structure from more-generic conceptual domains in response to immediate circumstances and are constantly modified as our thought unfolds.[29] For instance, Milne's sketch of Eeyore sets up two closely related mental spaces. The first is that of the Old Grey Donkey—solitary, graceless, and phlegmatic. The second is that of a somewhat morose and plodding intellect, tangled in its own thoughts and happy enough to leave them behind at the first opportunity for social interaction. Aspects of these two spaces are combined in a third space, producing the character of Eeyore. Turner and Fauconnier use CINs to formalize the relationships between the mental spaces involved in a conceptual blend, to specify what aspects of the input spaces are imported into the blend, and to describe the emergent structure that results from the process of conceptual blending.

The CIN for the conceptual blend used by Milne is diagrammed in figure 2.2. Its network involves four interconnected mental spaces, which are shown as circles. Central to the network are two correlated *input spaces*, the "donkey" space and the "human" space. The solid double-headed arrow linking these two spaces indicates

28. Research on conceptual blending and conceptual integration networks is discussed in Gilles Fauconnier and Mark Turner, "Conceptual Projection and Middle Spaces," UCSD Department of Cognitive Science Technical Report 9401 (San Diego: University of California, San Diego, 1994); idem, "Blending as a Central Process of Grammar," in *Conceptual Structure, Discourse, and Language* [based on papers presented at the First Conceptual Structure, Discourse and Language Conference, San Diego, 1995], ed. Adele E. Goldberg (Stanford, Calif.: Center for the Study of Language and Information, 1996), 113–30; idem, "Principles of Conceptual Integration," in *Discourse and Cognition: Bridging the Gap: [Second] Conceptual Structure, Discourse, and Language Conference, Buffalo, 1996*, ed. Jean-Pierre Koenig (Stanford, Calif.: Center for the Study of Language and Information, 1998), 269–83; idem, "Conceptual Integration Networks," *Cognitive Science* 22 (1998): 133–87; Mark Turner and Gilles Fauconnier, "Conceptual Integration and Formal Expression," *Metaphor and Symbolic Activity* 10 (1995): 183–204; idem, "Conceptual Integration in Counterfactuals," in *Conceptual Structure, Discourse, and Language, II*, ed. Jean-Pierre Koenig (Stanford: Center for the Study of Language and Information, 1998), 285–96; Turner, *The Literary Mind*; Fauconnier, *Mappings in Thought and Language*; Seana Coulson, *Semantic Leaps: Frame-Shifting and Conceptual Blending in Meaning Construction* (Cambridge: Cambridge University Press, 2001), 115–202. The most comprehensive study of conceptual blending as of this writing is Gilles Fauconnier and Mark Turner, *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities* (New York: Basic Books, 2002).

29. The theory of mental spaces is developed in Gilles Fauconnier, *Mental Spaces: Aspects of Meaning Construction in Natural Language*, 2nd ed., with a foreword by George Lakoff and Eve Sweetser (Cambridge: Cambridge University Press, 1994); Fauconnier, *Mappings in Thought and Language*; Gilles Fauconnier and Eve Sweetser, eds., *Spaces, Worlds, and Grammar* (Chicago: University of Chicago Press, 1996); see also Turner, *The Literary Mind*.

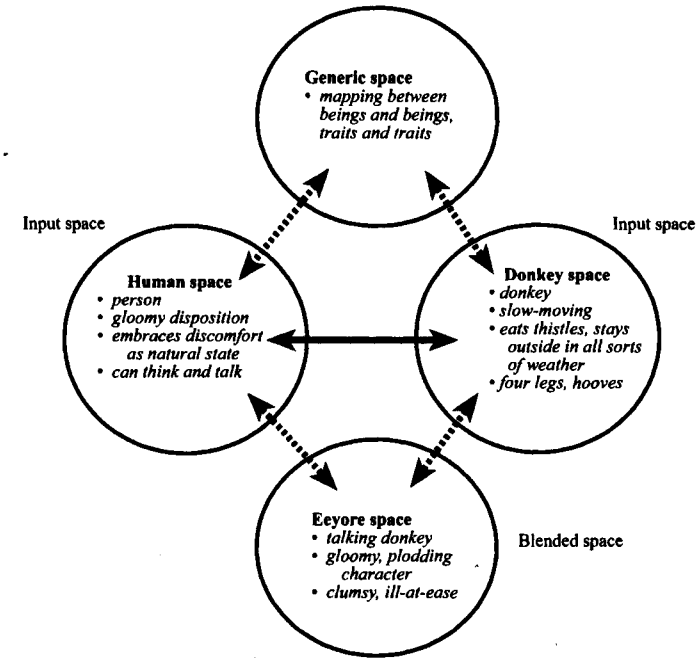



FIGURE 2.2 Conceptual integration network (CIN) for the anthropomorphic blend used for A. A. Milne's Eeyore

that elements within them serve as structural correlates: *person* is correlated with *donkey*, and *gloomy disposition* is correlated with *slow-moving*. Guiding the process of mapping between these domains is the *generic space*, which defines the core cross-space mapping and basic topography for the CIN.[30] Throughout this network, beings are mapped onto beings, and character traits onto character traits. Guided by the conceptual framework provided by the generic space, structure from each of the input spaces is projected into the fourth space, called the *blend*, which results in the anthropomorphic character of Eeyore. The mapping is only partial, however, reflecting the limitations imposed by the generic space. Since the generic space does not map between physical characteristics, we do not expect Eeyore to be of human appearance or to be human-sized.[31]

The dashed arrows linking the generic space to the input spaces, and the input spaces to the blended space, indicate the directions in which structure is projected:

30. The topography of a domain consists of the elements that populate the domain and the relations that hold between them. By definition, all of the domains in a CIN must have a uniform topography, which is represented in its abstract form by the generic space. The requirement of uniform topography within the correlational structure of a CIN can be seen as another manifestation of the Invariance Principle discussed earlier in this chapter.

31. Inasmuch as a number of the characters in Milne's stories originated as stuffed toys, they are easy to imagine as rather diminutive in size; in fact, this is how they are rendered in E. H. Shepard's illustrations. In his text, Milne is relatively silent on this matter, although various contextual clues suggest that the "animals" with which he deals are of the stuffed variety.

from the generic space to the input spaces, and from the input spaces to the blended space. The arrows are double-headed because, under certain circumstances, structure may also be projected from the blended space back into the input spaces, and from the input spaces back into the generic space. The idea of anthropomorphic animals that emerges in the blended space may thus influence the way we think about "regular" animals, leading us to talk to them and act toward them as though they had human characteristics. Similarly, our experiences with actual animals and people give life to the abstractions of the generic space. The double-headed arrows also serve as a reminder of the limitation of all of the diagrams of CINs I shall use: mental spaces are dynamic structures, as are the CINs that are built from them. Thus figure 2.2 represents a sort of analytical snapshot of this particular network, framed with the intent of capturing its essential features, but making no claim to exhausting the possibilities for description. Hints about how the CIN and its spaces may develop can be gleaned from the diagram, but a full account would require a series of such snapshots.

Although the diagram given in figure 2.2 is standard in the literature on mental spaces, it can lead to two misunderstandings. First, since the blend is at the bottom of the diagram, it gives the impression that concepts precipitate down into the blend. Second, the function of the generic space can be a bit confusing, since it does not seem to be directly involved in the blend. In the interests of clarifying these points I offer figure 2.3, which represents the essential components of a four-space CIN in a slightly different format. Here the generic space is properly represented as both the background and the foundation for the entire network. The two input spaces are concrete representations of the abstract structure represented by the generic space, and the conceptual blend is a further projection from these.

An important aspect of the topography of CINs is the basic logic established by the generic space. For the CIN of figure 2.2, the assumption is that beings visibly manifest their character traits. Old donkeys are slow-moving and balky; morose humans are given to gloomy pronouncements on the state of the world; and Eeyore talks and acts like the clumsy, gloomy character he is. The topography of the network also guides three operations—composition, completion, and elaboration—that produce emergent structure unique to the blend.

*Composition* puts together elements from the input spaces to create new entities in the blended space, yielding the character of Eeyore: although donkeys cannot actually think and talk, and humans do not have four legs or hooves, Eeyore has all these traits. *Completion* extends the image suggested by the initial mapping from the input spaces, drawing on our background knowledge of the circumstances summoned by the CIN. For instance, we know that gloomy characters expect the worst of situations. Eeyore can thus be relied on for solemn pronouncements of impending disaster on even the sunniest of days, to greet each calamity as confirmation of his estimation of the world, and to eat his thistles with little sign of relish. *Elaboration* is a more extensive operation than completion is; it develops the structure of the blended space by building on the principles and logic evinced by the blend. In effect, the input spaces decrease in importance and the focus is directed toward the rich imaginary possibilities of the blended space. At this point, we can start to write

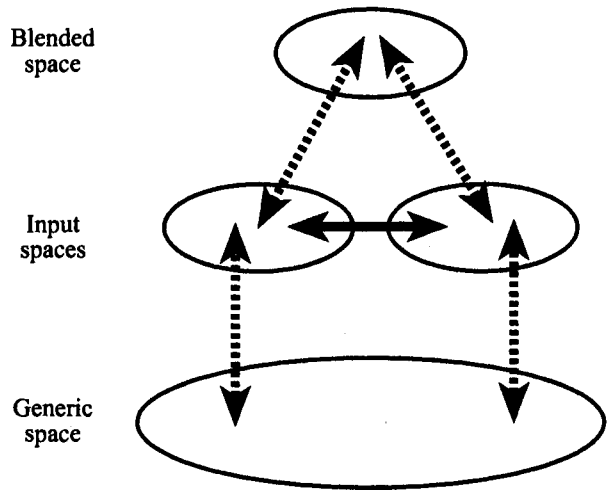



FIGURE 2.3 Alternative representation of a four-space CIN

our own stories about the Old Grey Donkey. Were we to place Eeyore aloft in an airplane, we could be sure he would profess no enjoyment of the view but would instead focus on the various aeronautical catastrophes that, from his perspective, were almost certainly imminent. Given Eeyore's clumsiness and the awkwardness of his hooves, we would be much less likely to imagine him flying the plane, leaving this to the more dexterous animals or to Christopher Robin.

Although the conceptual blend that yields anthropomorphic animals is important to Milne's story, it is not the only blend in evidence. Because Eeyore originated as a stuffed animal, he retains some of the characteristics of the species. Accordingly, in the story in which he is introduced, he loses his tail only to have it recovered by Winnie-the-Pooh and nailed back on by Christopher Robin, a sequence of events unlikely were he based solely on real donkeys. This points to yet another conceptual blend, in which one of the input spaces is structured around stuffed, rather than real, animals (and one that I shall not pursue in the present discussion).

A. A. Milne's Eeyore and most of the other characters from the Winnie-the-Pooh stories provide clear examples of conceptual blending. As indicated, however, the process of blending is not restricted to children's stories. On the one hand, conceptual blending is also common in everyday discourse. Witness statements such as "The car is being stubborn today—I just can't get her to start," which blends concepts related to the physical properties of inanimate objects and those related to the behavior of humans. On the other hand, conceptual blending also has its place in literature. Consider Proust's description of one feature of the springtime walks along the "Méséglise way" in Combray:

> We would leave town by the road which ran along the white fence of M. Swann's park. Before reaching it we would be met on our way by the scent of his lilac-trees, come out to welcome strangers. From amid the fresh little green hearts of their foliage they raised inquisitively over the fence of the park their plumes of white or mauve

> blossoms, which glowed, even in the shade, with the sunlight in which they had bathed.[32]

Here Proust blends features of the lilac trees with actions of which only animals or humans are capable, thus creating a memorable image that perfectly captures the intense sensual engagement prompted by spring.

Conceptual blending is a pervasive and often transparent cognitive process. A given situation or story may involve any number of conceptual blends, as the character of Eeyore—constructed from the attributes of both stuffed and real animals—shows. Each such blend can be described through a conceptual integration network, which permits a systematic description of specific features of the blend. Finally, the CIN associated with the character of Eeyore can be used to explain why the Old Grey Donkey has some, but not all, of the features of humans, and why we can easily imagine him taking part in some activities but not in others.

## Conceptual Blends and Text Painting

In the conceptual blends just discussed, the input spaces for each CIN were summoned by language. As Fauconnier has noted, however, mental spaces are very general and are constructed for many cognitive purposes; language is but one way to prompt the construction of a mental space.[33] Under certain circumstances, music can also prompt space construction. In the case of Palestrina's text painting, the mental space is relatively circumscribed and focuses on an orderly progression of pitches that lead resolutely toward a cadence. This space is in correspondence with that set up by the text, which focuses on Christ's movement from the heavenly to the mundane and the lessening of potential energy associated with physical descents.

Correlations between the musical and textual spaces involved in Palestrina's text painting set up the CIN shown in figure 2.4. The generic space for the CIN is structured around the notion of elements that are in directed relations conceived with respect to a teleological framework. Within the network, motions through physical and musical space are *directed* motions. As with the conceptual blend associated with the character of Eeyore, the emergent structure of the blended space can be described in terms of the operations of composition, completion, and elaboration.

Composition puts musical pitches together with the act of descent to yield pitch-objects that descend. Composition also associates descent with a specific sound, which actual descents may or may not produce. Completing this image, we might infer that the lower the pitch, the closer it is to the ground and to the mundane. We could also infer that, since descent can be given a sound, ascent can also be given a sound. Palestrina in fact confirms the latter inference in mm. 92–93 of the

32. Marcel Proust, *Swann's Way* (vol. 1 of *Remembrance of Things Past*), trans. C. K. Scott Moncrieff and Terence Kilmartin (New York: Vintage Books, 1981), 147–48. Further blends in Proust are discussed in Turner, *The Literary Mind*, 91–92, 120–23; see also my concluding chapter.

33. Personal communication from Gilles Fauconnier, 8 September 1998. This perspective is also represented, somewhat less explicitly, in Fauconnier and Turner, "Conceptual Integration Networks"; and Fauconnier, *Mappings in Thought and Language*.

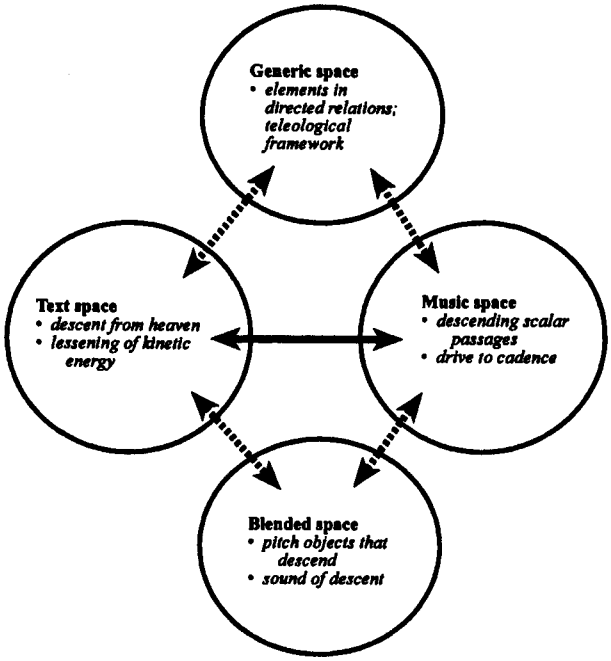
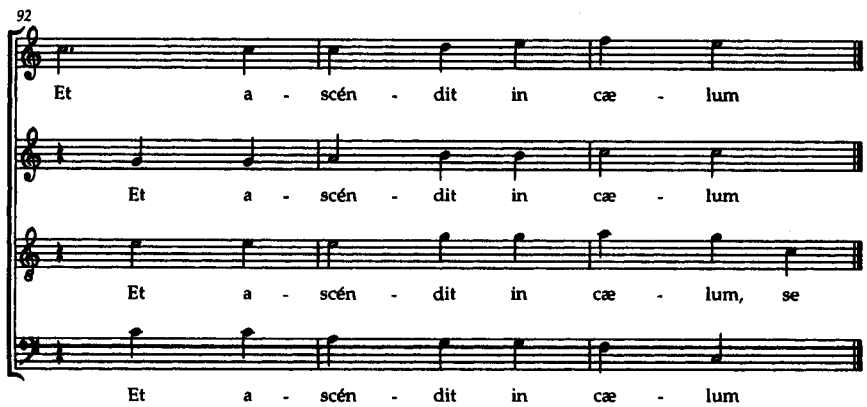



FIGURE 2.4 CIN for Palestrina's text painting

Credo, when, as shown in example 2.2, he sets "ascendit" with steadily rising pitches. Elaborating the blend, we might imagine pitch-objects doing all sorts of things, not just ascending and descending. Something like this is behind Louis and Thuille's characterization of tonality as a system of forces operating on pitches distributed throughout the field of musical space.

The extension of the conceptual blend suggested by mm. 92–93 of the Credo

EXAMPLE 2.2 Giovanni Pierluigi da Palestrina, Credo of the *Pope Marcellus Mass*, mm. 92–94



notwithstanding, Palestrina's text painting is a relatively isolated instance of meaning construction involving music and language. Indeed, text painting is usually regarded as a clearly circumscribed compositional technique, most often restricted to a single word or image. Under certain circumstances, however, music and language can combine to create rather more extended possibilities for meaning construction, as is shown by an instance of text painting in Giaches de Wert's seven-part madrigal "Tirsi morir volea," first printed in 1581.

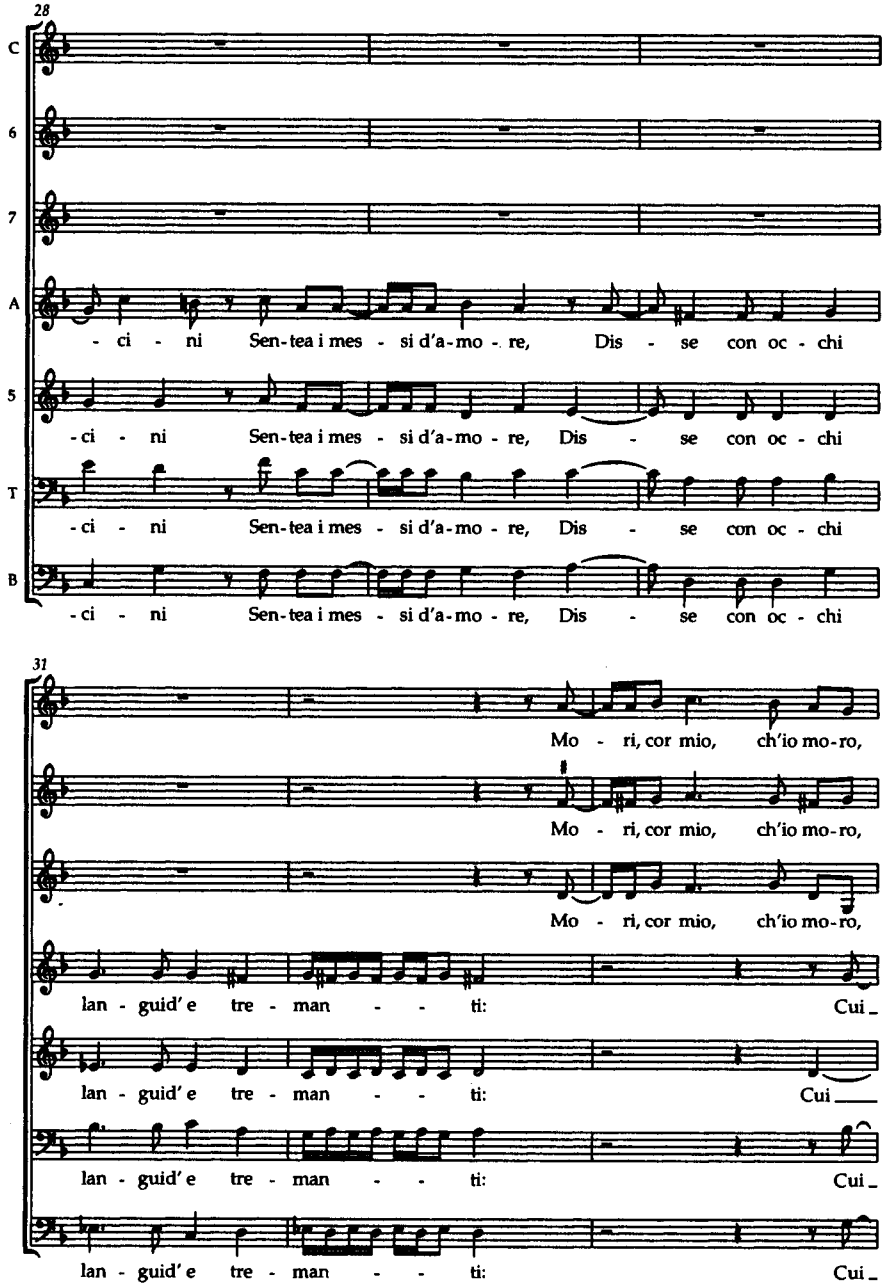
The text for Wert's madrigal is from a poem by Giovanni Battista Guarini that is probably the most popular madrigal text of the late sixteenth century. It recounts a sexual encounter between a shepherd and a nymph.[34] The passage relevant to discussion here appears in example 2.3; the text for the entire poem, with the verses that appear in example 2.3 underlined, is as follows:

| TIRSI MORIR VOLEA | THYRSIS WISHED TO DIE |
|---|---|
| Tirsi morir volea, | Thyrsis wished to die, |
| Gli occhi mirando di colei ch'adora, | gazing at the eyes of his beloved, |
| Quand'ella, che di lui non men ardea, | when she, whose ardor equaled his, |
| Li disse: "Ahimè, ben mio, | said: "Alas, my love, |
| Deh non morir ancora | do not die yet, |
| Che teco bramo di morir anch'io." | since I long to die with thee." |
| | |
| Frenò Tirsi il desio, | Thyrsis curbed the desire, |
| Ch'ebbe di pur sua vita allor finire, | which by then had almost ended his life: |
| E sentea morte, e non potea morire; | he felt death near, yet could not die; |
| E mentre il guardo suo fisso tenea | and while he kept his gaze |
| Ne' begli occhi divini, | fixed upon those eyes divine, |
| E'l nettar amoroso indi bevea, | and drank from thence the nectar of love, |
| La bella Ninfa sua, che già vicini | his pretty Nymph, who felt |
| | |
| Sentea i messi d'Amore, | Love's heralds near, |
| Disse con occhi languid' e tremanti: | said with languishing and trembling looks: |
| "Mori, cor mio, ch'io moro." | "Die my heart, for I die." |
| Cui rispose il Pastore: | At which the Shepherd replied: |
| "Ed io, mia vita, moro." | "And I, my life, die." |
| Così moriro i fortunati amanti, | Thus the happy lovers died |
| Di morte sì soave e sì gradita, | a death so sweet and pleasant, |
| Che per anco morir tornaro in vita. | that in order to die again, they returned to life. |

The text painting begins in mm. 31–32 when the word "tremanti" is set to a written-out ornament (a trill or gruppo) in all parts. "Tremanti" ("trembling") sets up a mental space focused on an intense and partially involuntary physical reaction to stress that produces repeated oscillating motions. The written-out ornament sets up a mental space focused on the rapid alternation of a "minor chord" with E♭3 in the lowest voice, and a "major chord" with F♯4 in the highest (which creates an

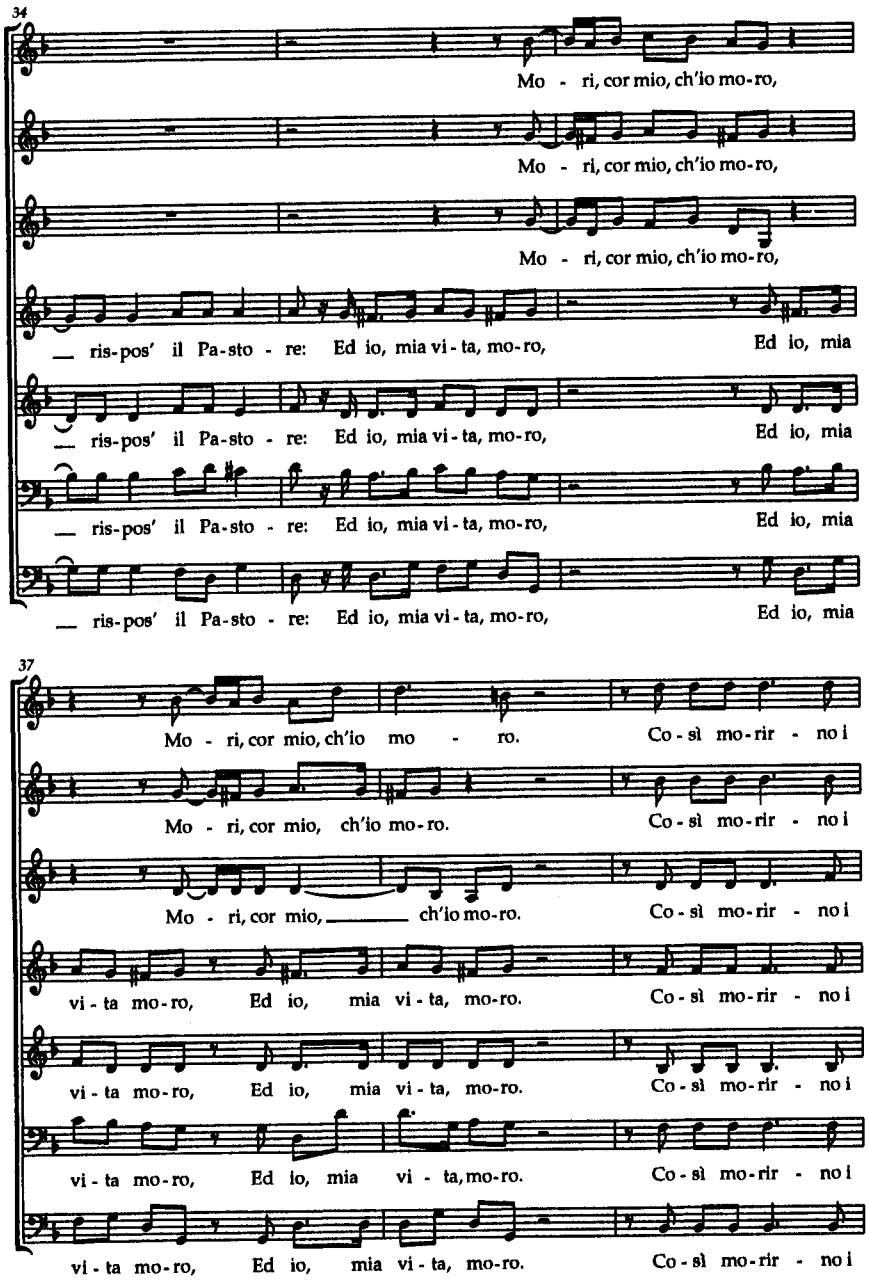
34. The context for Guarini's poem and Wert's setting of it have been discussed by Laura Macy in "Speaking of Sex: Metaphor and Performance in the Italian Madrigal," *Journal of Musicology* 14 (1996): 1–34.

EXAMPLE 2.3 Giaches de Wert, "Tirsi morir volea," mm. 28–39




*(continued)*

EXAMPLE 2.3 *(continued)*



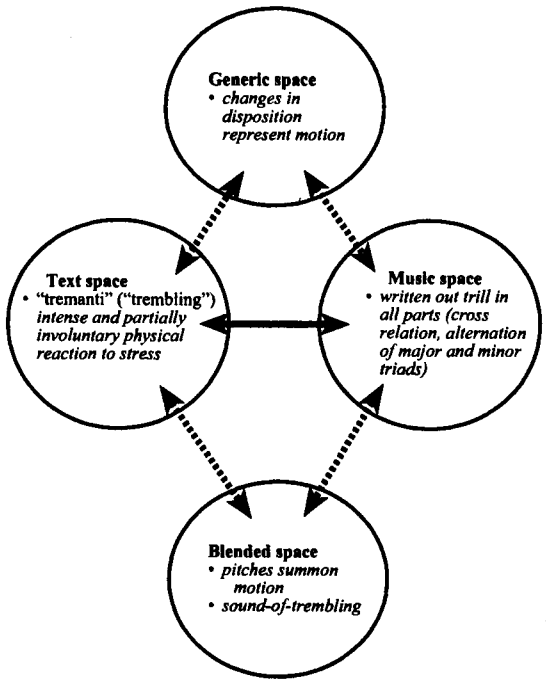



FIGURE 2.5 CIN for mm. 31–32 of Wert's "Tirsi morir volea"

augmented-second cross-relation).[35] Similarities between these spaces suggest correlations between their elements: each chord in the musical space correlates with an end-point of the physical motion summoned by the textual space, and the intensity of the physical reaction correlates with the rapid alternation of major and minor chords and the cross-relation between the outer voices.

It is the correlations between these spaces that set up the CIN shown in figure 2.5. The generic space for the CIN is structured around the idea that changes of disposition represent motion.[36] Physical trembling of the sort summoned by "tremanti" is one instantiation of this notion, and the rapid alternation of musical materials is another. Once again, the emergent structure of the blend can be described in terms of the operations of composition, completion, and elaboration. Composition combines musical pitches with specific features of trembling, so that the pitches summon a sense of physical motion even though the sound-source producing them remains relatively fixed: the singers of Wert's madrigal need not actually tremble when they sing "tremanti." Composition also suggests that trembling has a sound,

35. At the time Wert's madrigal appeared, terms such as "minor chord" and "major chord" were not conventionally used to describe simultaneous sonorities. Writers on music during the period preferred to describe sonorities in terms of their intervallic makeup. In the present context, this more accurate terminology is just a bit unwieldy, and I have consequently used somewhat more familiar and concise modern terms. I trust their anachronistic application here will not prove unduly confusing.

36. This is analogous to the phi phenomenon noted by Gestalt psychologists. See Kurt Koffka, *Principles of Gestalt Psychology* (London: Routledge and Kegan Paul, 1935), 179.

even though most actual trembling is done in relative silence. If we complete the pattern suggested by the basic mapping, we can infer that once the physical action stops the sound will stop. Finally, through elaboration we can imagine that successions of restricted groups of pitch materials could depict intense physical actions more complex than trembling, perhaps even extending to the interaction of physical bodies.

Although the elaboration of a blended space is often left to our imagination, at times it may be explicitly developed. This is indeed what Wert does in the measures immediately following the setting of "tremanti." Here are some of the salient features of the music of mm.33–38:

- Starting with the pickup to m.33, the women's and men's voices engage in a rapid alternation of entrances.
- The beginnings and endings of the entrances are elided, so that the singing is seamless.
- The spacing between entrances is gradually compressed, culminating in the joining of men's and women's voices in mm.37–38; in fact, this is the first time in the madrigal that the two vocal groups sing together.
- The setting is almost entirely homophonic within each group of voices.
- The harmonic material is highly restricted, consisting almost entirely of triadic sonorities built on G, D, and F.
- In m. 38, the music breaks off, and there is a half measure of silence.

The alternation of restricted groups of pitch material, the charged interaction of the vocal groups, the sudden breaking off of the voices in m.38, and the context provided by the blend produced by the text painting of "tremanti"—all summon an image of intense physical activity followed by a sudden suspension of that activity. It is an image that gives sounding representation to the sort of death—sexual orgasm—about which the poem obsesses. This image also offers a way to ground the double entendre exploited by Guarini, built on the paradoxical connection between life at its most intense and death. Both sexual climax and death are marked by sharp, nearly immediate contrasts: orgasm followed by quiescence; the clamor of life followed by the silence of death. Wert's music, building from the common ground shared by sexual climax and death, breathes life into a play of words that, by this point in the sixteenth century, had become a rather common commonplace.[37]

It is important to emphasize that not all combinations of text and music will yield rewarding conceptual blends. For a blend to occur, there not only needs to be some correlation between two domains, but the domains must also have a uniform topography. In some cases—certain strophic songs, for instance—the correlation between text and music is simply too general to generate a compelling blend. In

37. My interpretation of connections between sexual climax and death is somewhat different from that offered by Laura Macy. Noting that early physiological accounts saw both sex and death as an emission of spirit, Macy writes that "this link is surely at the bottom of the metaphoric use of the word death to mean sexual climax" ("Speaking of Sex," 5). While I would not want to deny the importance of this connection, my argument is that commonalities between the larger physiological processes of both orgasm and death could also serve to ground the metaphor and that Wert's music provides a musical analogue for these.

other cases—formulaic songs composed for opera or reflecting the more commercial side of popular music—the correlation between text and music may be so tenuous as to be virtually nonexistent. Also important are the conceptual models that inform the interpretation of the correlation between the input spaces for a blend. A careful listener who is nonetheless unaware of the context of Palestrina's music, for example, will almost certainly realize that something is going on in the passage cited in example 2.1 but may interpret the music as "sad" or "losing energy." If conceptual models associated with the act of descent, or with this particular portion of the Mass, are not activated, the listener may get no farther than this. Similarly, Wert's climactic moment, for a listener who does not understand the text (either for lack of knowledge of Italian or out of innocence of the double entendre), may simply be heard as "exciting." In both cases, the conceptual blends will remain latent: potential but not actualized opportunities for the construction of meaning.

As Wert's madrigal suggests, conceptual blends involving text and music may be relatively extended. Chapter 6 will consider such blends in greater depth, showing how they develop over the course of entire songs, how the same text set to new music can give rise to two very different songs, and how music alone can elaborate a blend first set up by text and music.

## Conceptual Blending and Program Music

In theory, conceptual blends may involve mental spaces associated with any domain of thought. Musical domains could thus be correlated not only with domains associated with language but also with physical gesture or color.[38] In the following, I would like to consider the conceptual blends that occur when instrumental music is associated with an extra-musical program. Although such programs are almost always invoked through language, the connection between the musical and linguistic domains is much looser, and the concepts that emerge from the blend are somewhat more variable. The generality of such blends also allows them to cover greater expanses of music—in the example with which I am concerned here, an entire symphony.

Ludwig van Beethoven's Symphony No. 6 in F major, Op. 68 ("Pastorale") is among the most well-known of instrumental works with which extra-musical programs have been associated. In the case of the Sixth Symphony, this association was aided by the descriptive titles Beethoven gave individual movements (such as

38. The correlations I refer to here are between the conceptual domains associated with music and with colors or gestures; this fits with my notion of concepts as not necessarily linguistic. However, I am not thinking of synaesthesia, where stimuli applied to one of the five senses produces responses in one or more of the other senses as well; I view this phenomenon as functionally automatic and thus not involving concepts. Synaesthesia between music and color is common enough to have warranted relatively extensive research, and no end of commentary. See, for instance, Lawrence E. Marks, *The Unity of the Senses: Interrelations among the Modalities* (New York: Academic Press, 1978); Wilson Lyle, "Colour and Music: An Introduction," *Music Review* 43 (1982): 261–64; Michael Hurte, *Musik, Bild, Bewegung: Theorie und Praxis auditiv-visueller Konvergenzen* (Bonn: Verlag für systematische Musikwissenschaft, 1982); and Jonathan W. Bernard, "Messiaen's Synaesthesia: The Correspondence between Color and Sound Structure in His Music," *Music Perception* 4 (1986): 41–68.

"Scene by the Brook"). As Richard Will has pointed out, the programmatic aspects of the Sixth Symphony created difficulties for those who sought to give an account of the work, for the simplistic representation of natural sounds seemed in conflict with the received wisdom that Beethoven's music referred to nothing outside itself.[39]

Symptomatic of this uneasiness with Beethoven's program was Donald Francis Tovey's outright rejection of any link between such natural sounds and the music of the Sixth Symphony. Tovey writes, "In the whole symphony there is not a note of which the musical value would be altered if cuckoos and nightingales, and country folk, and thunder and lightning, and the howling and whistling of the wind, were things that had never been named by any man, either in connection with music or with anything else."[40] With the correlation between the domain of nature and the domain of music sundered, Tovey paves the way to approach the work as "a perfect classical symphony" and to claim for it a place in the pantheon of Beethoven's works.

Nevertheless, Tovey's analysis of the symphony seems to speak of something other than absolute music. Describing an important moment in the slow movement, he writes,

> The deep shadow of this remote key of G flat becomes still deeper as C flat, which, changing enharmonically to B natural, swoops round to our original key B flat. At the outset of this wonderful passage the theme was that of the first subject with the murmur of the brook becoming articulately melodious in the clarinet and the bassoon. At the moment when the melody gathers itself up into a sustained phrase and makes its enharmonic modulation, there comes a phenomenon full of deep meaning. From this point nothing is left of the melody but sustained notes and bird-song trills; the whole of the rest of the return to the main key is harmonic and rhythmic. In this as everywhere else the movement remains true to type, a perfect expression of the happiness in relaxation.[41]

Here we have key areas that swoop, clarinets and bassoons that are articulate, and melodies capable of independent motion, all within a movement that expresses the happiness there is to be found in relaxation.

At first glance, Tovey's prose seems to retain the very correlation between the domains of music and nature that he had earlier rejected, and to develop a conceptual blend based on this correlation. Still present are the murmur of the brook, the bird-song trills of melody, and the general atmosphere of the pastoral, all wedded to Beethoven's music. Closer consideration, however, shows that the correlation of music and nature cannot explain the structural attributes of the imaginative domain summoned by Tovey. Although nature includes living beings, music does not; at best, music comprises sounds that occur within a temporal framework. The independent agency implied by Tovey, evident in the articulate and goal-directed entities he imagines, also seems somewhat out of step with the commonplace view of

39. Richard Will, "Time, Morality, and Humanity in Beethoven's *Pastoral* Symphony," *Journal of the American Musicological Society* 50 (1997): 271–329.

40. Donald Francis Tovey, *Symphonies* (vol. 1 of *Essays in Musical Analysis*) (London: Oxford University Press, 1935), 45.

41. Tovey, *Essays in Musical Analysis*, 1: 51. The reference is to mm. 78–85 of the slow movement.

nature as a domain in which events are guided not so much by individual volition as by larger forces working from without. In short, it is one thing to link elements of nature and musical sounds to create a domain in which the cuckoo sings with the timbre of the clarinet and the sound of a brook is summoned by undulating patterns in the strings. It is rather another to animate musical sounds and to endow them with their own volition.

As it turns out, Tovey's interpretation does not rely on a simple correlation between music and nature. Instead, Tovey activates a much larger network of mental spaces to produce his interpretation. Central to this network is one of the most common of conceptual blends, that of anthropomorphism (which played an important role in the blend that produced A. A. Milne's Eeyore). The basic correlation that underlies this blend can be seen in statements like "The car is being stubborn today —I just can't get her to start." In this mapping, the domain of inanimate objects is structured by the human domain; more specifically, *human being* is mapped onto *car*, and a volitional state proper to humans (*stubbornness*) is mapped onto the mechanical state of the car (*not-starting*).[42] A blend that exploits this mapping might then attribute additional volitional or emotive states to the vehicle ("I think the car wants to stay home," "She's mad because I haven't changed her oil"), or extend into the nonvehicular domain ("I don't think the toaster likes me, it keeps burning the bread"). In its most abstract form, this blend involves two input spaces, as shown in the CIN of figure 2.6. One recruits structure from the human domain, the other from some non-human domain. The correlation of these two spaces gives rise to the blended space of anthropomorphism, whose emergent structure reflects the counterpart correlations of the input spaces. Within this space, non-human entities become endowed with the characteristics and intellects of humans while retaining many of the features that make them distinct from humans. An anthropomorphic car may be stubborn and sullen, but it will also be able to stay outside year round (even though it may not "like" to do so).

The generic space for the CIN connects entities with entities and states with states. According to the logic proper to the network, the events that occur within each source domain have causes that follow from the rules that govern the behavior of entities that populate the domain. *Stubbornness* in the human domain reflects personal idiosyncrasies and may be abetted by external circumstances that render the current situation unsatisfactory. According to conventional wisdom, people are stubborn either because they were born that way or because they are unhappy with the way things are and will not take any action (other than being stubborn) until things change. *Not-starting* in the domain of automobile mechanics is caused by electromechanical conditions that are insufficient to initiate or produce sustained combustion. The car will not start because there is something wrong with the engine. It is important to note that, although the generic space may have a structure that is strongly image-schematic, generic spaces are often rich with detail. Generic spaces

42. Note that to ascribe a state associated with a living entity to a nonliving entity is not the same thing as to endow the latter with life. The statement "The car won't start—I think it died" does not entail a belief in animistic vehicles, but only correlates a particular state associated with living beings (*being dead*, or *being no-longer-alive*) with the mechanical state of the car (*never-to-start-again*).

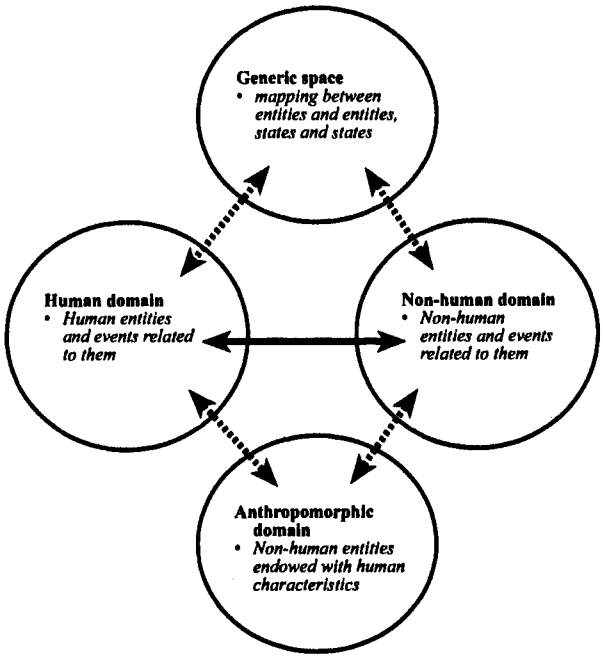



FIGURE 2.6 CIN for anthropomorphic blend

are not necessarily primitive or schematic—they only define the basic topography common to all of the spaces in the blend.

Tovey activates the mental space of anthropomorphism early in his essay when he writes, "The Pastoral Symphony has the enormous strength of some one who knows how to relax."[43] Within the blend that results from correlating the human and musical domains, musical events become actors endowed with their own characteristics and capable of their own acts of volition. Thus the transition that follows the first subject of the opening movement of the symphony "leads in three indolent strides to a second subject which slowly stretches itself out over tonic and dominant as a sort of three-part round."[44] Tovey develops an anthropomorphic perspective on nature more gradually, bringing it to fruition only in his account of the end of the slow movement. "Suddenly for a moment all is silent; we have no ears even for the untiring brook, and through the silence comes the voice of the nightingale, the quaint rhythmic pipe of the quail, and the syllabic yet impersonal signal of the cuckoo."[45] Here elements of nature (brooks and birds) become actors, each endowed with individual characteristics and capable of individual acts of volition. In both cases, the correlation of human and non-human domains gives rise to imaginative conceptual blends: symphonies that know how to relax, musical passages that stride, brooks that do not tire, and cuckoos capable of impersonal signals.

43. Tovey, *Essays in Musical Analysis*, 1: 46.
44. Tovey, *Essays in Musical Analysis*, 1: 47.
45. Tovey, *Essays in Musical Analysis*, 1: 52.

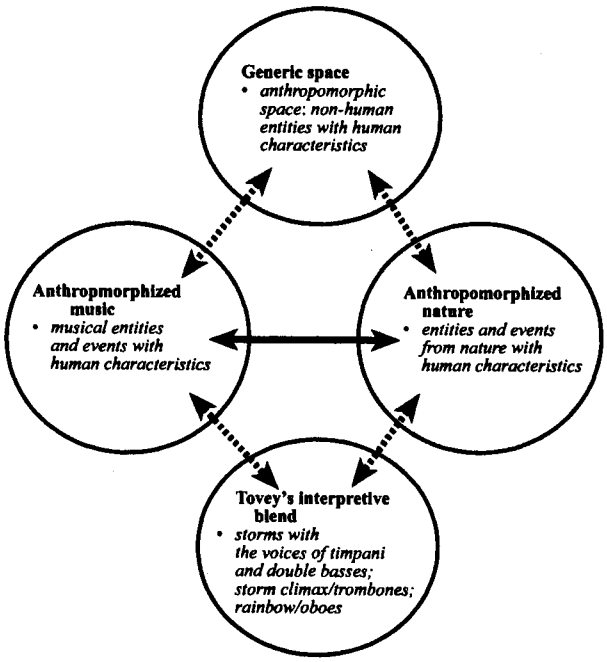



FIGURE 2.7 CIN for Tovey's analysis of the Beethoven Pastoral Symphony

Construing both music and nature in anthropomorphic terms makes it possible for Tovey to connect the two in a further blend, shown in figure 2.7. The generic space for this network is the blended space of figure 2.6. One of the input spaces is that of anthropomorphized nature, which includes natural agents who act within and respond to their native environment ("nature"). The other input space is that of anthropomorphized music, which includes musical agents who act within and respond to their native environment ("music"). In the blend, aspects of the input spaces are combined to create unique entities and relations. Here the murmur of a brook becomes an articulate melody for clarinets and bassoons; here Beethoven's music can be conceived as natural without having been reduced to nature.

The three operations that produce emergent structure in the blended space—composition, completion, and elaboration—are all evident in Tovey's account of the end of the fourth movement: "The storm moves in grand steps to its climax. This is marked by the entry of the trombones. . . . Then the storm dies away, until with the last distant mutterings of the thunder the oboes give a long slow fragment of bright sustained melody on the dominant of F. This has been aptly compared with a rainbow."[46] Composition unites the lives and interests of natural agents and musical agents. In the final twenty measures of the fourth movement, the thunder mutters with the voices of the timpani and double basses, who also prepare the tonality of the closing fifth movement precipitated by the retreat of the storm.

46. Tovey, *Essays in Musical Analysis*, 1: 55.

Completion supplies additional structure. Beethoven's terse title for the fourth movement is "Gewitter, Sturm" ("Thunderstorm, Tempest"), which says nothing of the aftermath of a storm. Nonetheless, we know that rainbows often follow storms, and so the "bright sustained melody" of the oboes can summon the image of a rainbow, a completion suggested but not necessitated by the input spaces.[47] Elaboration develops the blend by pursuing its logic. The storm, as a magnificent entity, moves with grand steps. The climax of its terror and wrath (enacted through musical agents) is marked by the entry of new musical agents (the trombones). As these passions abate, musical agents gradually disappear (a change only implicit in Tovey's prose) until only a bare few remain. Perhaps most striking about this elaboration is that a single natural agent (the storm) appears to be brought fully to life only through a multiplicity of musical agents. Such a genesis, while inexplicable in the natural domain, is fully consonant with the idiosyncrasies of the musical domain. There, individual musical entities (such as movements or entire symphonies) are often understood to emerge from the actions of the subentities (such as themes or tonal areas) that they comprise.

Tovey thus mobilizes a number of mental spaces to provide an account of Beethoven's symphony. The most active spaces are those that contribute to a blend of anthropomorphized natural and musical elements. This blend gives Tovey a way to describe music that seems naturalized but does not reduce to a simple correlation of natural and musical sounds. Less active, but still important, are spaces involved in blends fundamental to this primary network. These include spaces built up from the domain of human entities and events, along with the domains of nature and music. Tovey's description makes only sparing use of these subsidiary blends, since they cannot provide the naturalized account of music that is the goal of his essay. Far in the background is the generic space common to anthropomorphic blends, which defines the core cross-space mapping and logic that underlie the entire network.

## Summary

Conceptual blending is a dynamic process of meaning construction that involves small, interconnected conceptual packets called mental spaces, which temporarily recruit structure from conceptual domains in response to local conditions. When blending occurs, a portion of the structure from two correlated input spaces is projected into a third, blended space. As part of this process, the operations of composition, completion, and elaboration produce structure within the blend that is not found within either of the input spaces. This structure only becomes possible through the concepts and relations produced by conceptual blending.

The structure common to the mental spaces within a conceptual integration network is reflected in the generic space. This space defines the core cross-space mapping and is organized according to a basic logic that remains consistent throughout the network. Generic spaces are not necessarily primitive or image-

47. This suggestion is also followed by the animators of Walt Disney's *Fantasia*, who not only supply a rainbow to accompany this moment in Beethoven's symphony but integrate it within the multileveled, quasi-mythical story they set to the music.

schematic—they only define the basic topography common to all of the spaces in the blend.

In a basic blend, four spaces will be active (although not necessarily to the same degree): the generic space, the input spaces, and the blend. It often will be the case, however, that additional spaces will be activated, to a greater or lesser extent, as the construction of meaning proceeds. In certain passages within Tovey's account of Beethoven's Sixth Symphony, as many as six spaces may be active. In addition to the blend, these include mental spaces built up from the domains of music and nature, from the domains of anthropomorphized music and nature, and from the generic domain of anthropomorphism. This multiplicity of spaces explains the complexity of Tovey's account, as well as its interpretive richness. It also serves to explain the seeming contradiction entailed by Tovey's desire to deny the programmatic aspect of Beethoven's symphony while responding to the powerful impressions to which it gives rise.

More generally, evidence from conceptual blending lends credence to the idea that music is an independent conceptual domain. As I have shown in my analyses of text painting and program music, musical concepts combine with concepts from other domains to create blended concepts that suggest striking possibilities for the imagination. More to the point, musical syntax (from the perspective developed in chapter 1) can be seen contributing to the blended concepts associated with Wert's text painting and Beethoven's program music. In the former, Wert's deployment of musical materials sets the stage for an enactment—both by the voices of the singers and by our imaginations—of the sort of death with which Guarini's poem is concerned. As but one example from the latter, Beethoven's conclusion of the fourth movement and preparation for the fifth (the distant mutterings of the timpani notwithstanding) provide a moment of contrast and articulation that clearly supports our imagining one process ending (the storm) and another beginning (the shepherd's song of joyful thanksgiving). And musical spaces can inform our understanding of generic spaces: to hear Wert's musical climax is to get new insight into climax as a general phenomenon, be that climax sexual, cerebral, or even visual; to hear Beethoven's transition is to get a glimpse of how natural events like storms can be thought of not simply as personages but also as processes.

Each of these aspects of conceptualizing music will be taken up in more detail in chapter 6, within the context of the analysis of the nineteenth-century Lied. There it will be possible to see, to a greater extent, how processes of cross-domain mapping and conceptual blending construct and contribute to our understanding of both music and the world as a whole.

CHAPTER THREE

# CONCEPTUAL MODELS AND THEORIES

As part of her efforts to understand the everyday knowledge with which people make sense of music, the music theorist Jeanne Bamberger has often asked young children to develop ways to represent familiar melodies or simple rhythmic patterns. One way Bamberger does this is to have a child arrange a specially constructed set of bells, called Montessori bells, so that he or she can play a given tune. Bamberger then asks the child to invent instructions so someone else could play the tune on the bells as they've been arranged.

For those who have even a passing acquaintance with music, an encounter with Montessori bells can be a truly disorienting experience. Developed as part of the materials for the Montessori curriculum, the bells are designed to encourage children's exploration of sound and music. With this in mind, they are carefully constructed to be identical in appearance: the lowest-sounding bell looks just like the highest-sounding one, as do all those in between. Only by striking each bell with a small mallet can the different pitches of the various bells be discovered. Each bell is attached to a wooden stem mounted on a wooden base so that the bells are physically independent of one another. There is, in consequence, no set arrangement of the bells, just as there are no visual cues as to their pitch: how the bells might be arranged, whether in casual play or in the performance of some tune, has to be determined by each person who uses them. As a further challenge (or an additional resource), a complete set of Montessori bells contains duplicates: there are eight bells with white bases, which make up a diatonic scale from $C_4$ to $C_5$, and thirteen bells with brown bases, which make up a chromatic scale from $C_4$ to $C\sharp_5$. Thus for each white-based bell there is one brown-based bell with a matching pitch—but again, just which bells match must be discovered by the user.

In the course of her research, Bamberger discovered that children use one of two general strategies when building a tune with the bells. Children following the first strategy arrange the bells in the order in which the pitches occur in the tune. To build the first phrase of "Happy Birthday," for instance, the bells might be placed one after another to create a bell-path of the sort shown in figure 3.1, with the arrow indicating the order in which the bells are set down. (Note that two G-bells are used, one for each time the pitch G occurs anew in the tune. That is, a separate

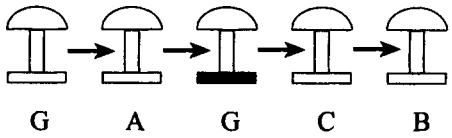



FIGURE 3.1 Bell-path no. 1 for the first phrase of "Happy Birthday" (open bases, white bells; solid bases, brown bells; arrows show direction of play)

bell is not provided to accommodate repeated notes in immediate succession. In this illustration, the second G-bell is taken from the set of brown-based bells.) Playing this much of the tune is a relatively simple matter of retracing the building process and striking each bell in turn (with the first bell getting two hits). The spatial arrangement of the bells, then, is in direct correspondence with the succession of pitches in the first phrase of the tune.

Children following the second strategy start by putting the bells in order, typically from low to high. Once this is done, they figure out the sequence in which to strike the bells in order to play the tune. Figure 3.2 illustrates the results of this strategy for the first phrase of "Happy Birthday." The bells are arranged from low to high (left to right); numbers indicate the order in which the bells should be struck, and arrows indicate the direction of motion. (Note that only one G-bell is required by this strategy; however, the path through the bells is considerably more complex.) As figure 3.2 shows, the placement of the bells no longer corresponds to the succession of pitches that make up the tune. The second strategy does not rely on a spatial analogue for the succession of pitches in the tune, as does the first. Compensation for this loss in immediacy is provided by a gain in flexibility: to play a different tune, the child needs simply to discover what new succession of hits will produce the correct results—the bells themselves do not need to be reordered.

Bamberger calls the representation produced by the first sort of strategy "figural" and that produced by the second "formal."[1] Figural representations highlight the unique "figures" (or groupings of musical materials) that make the tune what it is.

1. Jeanne Bamberger, *The Mind behind the Musical Ear: How Children Develop Musical Intelligence* (Cambridge, Mass.: Harvard University Press, 1991), 24–26; idem, "Cognitive Issues in the Development of Musically Gifted Children," in *Conceptions of Giftedness*, ed. Robert J. Sternberg and Janet E. Davidson (Cambridge: Cambridge University Press, 1986), 398–99. In personal correspondence (20 May 2000), Bamberger notes that formal strategies, as far as she has seen them among both children and adults, only occur with people who have had some training in music or who play an instrument.

Bamberger also observes that "representation" (the term I favor here) tends to imply more fully worked-out cognitive processes than a term like "description." Although I respect this distinction (especially where it applies to the activities of young children who may not really be concerned with "representing" anything), I shall retain "representation" in my discussion, but with the understanding that it applies to a quite general set of cognitive structures. Representation, defined thus, embraces image-schematic constructs, perceptual symbols, concepts, and actions—whatever is necessary to make real-world phenomena available (to re-present them) to the higher-level cognitive processes associated with theories. For further discussion of this approach to representation, see Mark Turner, "Design for a Theory of Meaning," in *The Nature and Ontogenesis of Meaning*, ed. Willis F. Overton and David S. Palermo (Hillsdale, N.J.: Erlbaum Associates, 1994), 91–108; and Antonio R. Damasio, *The Feeling of What Happens: Body and Emotion in the Making of Consciousness* (New York: Harcourt Brace, 1999), 320.

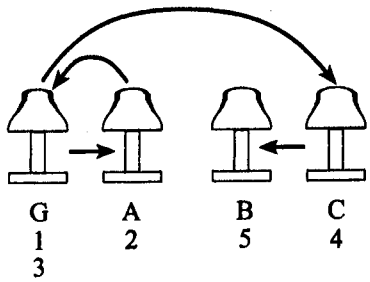



FIGURE 3.2 Bell-path no. 2 for the first phrase of "Happy Birthday"

Formal representations, in contrast, tend to obscure the unique characteristics of the tune but situate musical elements within a more general context applicable to any number of tunes. Bamberger's research suggests that figural representations are among the earliest formed by children. Formal representations develop somewhat later as the child's sense of how to organize and represent materials matures.

Bamberger's primary interest in the distinction between figural and formal representations was with the developmental changes that led to the transformation of the one into the other. Nonetheless, the contrast between these two strategies also points to different ways to construe relationships among musical concepts. That is, figural strategies offer one way to characterize these relationships, and formal strategies another. Perhaps more important, both remain options even after the passing of childhood, as is shown by the transformational networks employed by David Lewin in his analyses of late-nineteenth- and early-twentieth-century music. While Lewin's networks certainly have their abstract side—they owe much to theories of mapping that are employed in formal algebra—they also offer a way to represent unique relationships among musical events: they are figural as well as formal.[2]

These different ways to think about the essential relationships that obtain within what is putatively the same sonic domain (here, the first phrase of "Happy Birthday") demonstrate the role of two basic cognitive structures—conceptual models and theories—in our understanding of music. *Conceptual models*, which are the more basic structures, are made up of relatively limited correlations of concepts and are specific to a given domain. (We caught glimpses of such models in the discussion of categorization in chapter 1 and saw the role they played in cross-domain mapping in chapter 2.) *Theories*, in contrast, are extended structures that coordinate a number of conceptual models to provide a more comprehensive account of the world.[3] Figural

2. See David Lewin, *Generalized Musical Intervals and Transformations* (New Haven, Conn.: Yale University Press, 1987); and idem, *Musical Form and Transformation: Four Analytic Essays* (New Haven, Conn.: Yale University Press, 1993). Lewin makes explicit reference to Bamberger's work in his analysis of Karlheinz Stockhausen's *Klavierstück III*; see Lewin, *Musical Form and Transformation*, 45–47. I should note that Bamberger's work is by no means limited to early childhood development, and that she, too, takes the position that figural and formal strategies are not restricted to developmental stages in childhood. For instance, in dialogues built into *The Mind behind the Musical Ear*, she attributes figural strategies to one college-age interlocutor and formal strategies to another.

3. As is evident from this initial characterization, "theory," as I define it, embraces cognitive structures far more modest than those with which modern scientific theory is occupied. The connection between such cognitive structures and scientific theory is explored in greater detail in the next section.

strategies, on the one hand, rely on one set of conceptual models and lead to one theory of musical organization. Formal strategies, on the other hand, rely on another set of conceptual models and lead to a somewhat different theory of musical organization.

Both conceptual models and theories are central to our understanding of the world as a whole; however, both often remain implicit. It is only when they are teased out in the course of work such as Bamberger's, or as the result of deliberate reflection on our thought processes, that either conceptual models or theories come to the fore. In this chapter, I explore the structure and function of conceptual models and theories and show the part they play in both informal and formal accounts of musical organization. In the first section I examine the role conceptual models and theories play in the musical representations of one of Bamberger's subjects, an eight-year-old boy named Jeff. Jeff's encounters with music engage him in a struggle to come to terms with basic musical concepts, and they reveal music theory—that is, a theory specific to music—in the making. In the second section, I show how the cognitive structures demonstrated by Jeff's nascent theories of music connect with similar structures proposed by researchers in a wide variety of fields, from artificial intelligence and cognitive science to anthropology and ethnomusicology. This will serve to generalize the construal of conceptual models and theories beyond Jeff's specific situation and lay out some of their additional properties. In the third section, I turn to two of the more influential theories of music within the last three centuries of the Western tradition: the music theory of Jean-Philippe Rameau and that of Heinrich Schenker. My objective is to show how they employ the same basic cognitive structures that can be seen in children's figural and formal strategies for tune building.

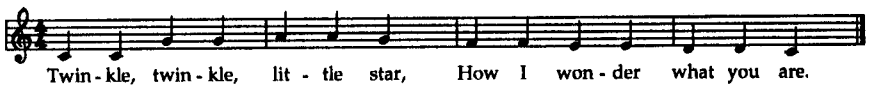
## DEVELOPING A THEORY OF MUSICAL ORGANIZATION

### Learning to Play "Twinkle" on the Bells

Jeanne Bamberger worked with Jeff for about six months in the Children's Learning Lab at MIT, starting in the fall of 1975.[4] She set him the task of building the tune "Hot Cross Buns," first with the sonic resources provided by a simple computer program and later with the Montessori bells. After Jeff achieved some facility with these tasks, Bamberger had Jeff concentrate on the tune "Twinkle, Twinkle, Little Star," using only the Montessori bells. In consequence, at the point we take up Jeff's story he has already learned some things about tunes and the properties of the bells, although "Twinkle," with its comparatively more complex structure, presents new challenges. In what follows, the focus is on the first phrase of "Twinkle" ("Twinkle, twinkle, little star, / How I wonder what you are"; the musical notation is given in ex. 3.1) and on two of the bell-paths Jeff constructed in order to play it.

4. "Jeff" is a pseudonym adopted by Bamberger to preserve the boy's anonymity. An extensive account of Bamberger's work with Jeff is given in chaps. 6–12 of *The Mind behind the Musical Ear.*

EXAMPLE 3.1 The first phrase of "Twinkle, Twinkle, Little Star"




In preparation for Jeff's work with "Twinkle," Bamberger had selected eleven bells from the complete Montessori set and put them out on a table.[5] Included were all of the white-based bells (providing a complete C major scale) and three brown-based bells: the C-, E-, and G-bells. Once Jeff had familiarized himself with the bells, Bamberger proposed that he try building "Twinkle." After beginning with the C-bell, Jeff searched among the available bells until he found one that matched the next pitch of the tune. He placed this bell next to the C-bell, and then played both to re-create the beginning of the tune (that is, the pitches that fit the words "Twinkle, twinkle"). To find the bell for "little," he once again searched among the available bells, and when he found the right one he placed it next in line. He again played what he had assembled of the tune ("Twinkle, twinkle, little") before returning to the remaining bells to find the one corresponding to the next pitch of the melody.[6]

Jeff's preference was to use the white-based bells first. He turned to the brown-based bells only when there was no white-based bell left that would provide the pitch for which he was looking; this first happens with the recurrence of the note G on the word "star." As a way of giving a visual representation of the division between the two subphrases (or "figures") of "Twinkle," Jeff left a space between the first four bells and the second four bells (marking the first four bells as one group and the second four bells as another). This produced the arrangement shown in figure 3.3. The result is a figural representation of the tune, with each successive bell corresponding to a successive pitch of "Twinkle." Performing the tune is simply a matter of beginning at one end of the bell-path and proceeding to the other (but with some bells getting two hits, and others only one).

After Jeff had completed his bell-path, Bamberger asked him to come up with instructions so someone else could play "Twinkle" on the bells as he had set them out. Jeff's initial solution was quite straightforward: he simply drew eight short vertical lines, with each line corresponding to a bell. The corresponding notation looked something like this:

| | | | | | | |

His intent was that the player should start at the beginning and play straight ahead to the end (mimicking, incidentally, his process of setting up the bell-path), some-

5. Bamberger's account of Jeff's first work with "Twinkle" can be found in *The Mind behind the Musical Ear*, 183–92.

6. The strategy of playing through as much of the tune as he had built before searching for the next pitch is one Jeff used throughout the tune-building exercises I discuss here.

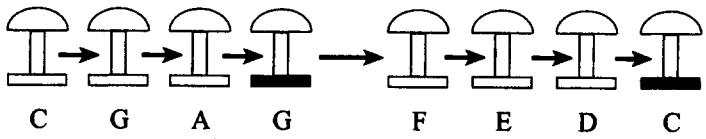



FIGURE 3.3 Jeff's initial bell-path for the first phrase of "Twinkle, Twinkle"

thing made manifest when, at Bamberger's prompting, he put in numbers to show the sequence of actions:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

With a bit of additional prompting, Jeff concluded his notation by adding another set of numbers to show how many times each bell was to be struck:

| 2 | 2 | 2 | 1 | 2 | 2 | 2 | 1 |
|---|---|---|---|---|---|---|---|
| \| | \| | \| | \| | \| | \| | \| | \| |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

With this, Jeff's first session with "Twinkle" was complete: he had constructed a bell-path with which he could play the first phrase of the tune, and he had come up with a notation that would allow another child to do the same (once the meaning of the numbers and lines had been explained).[7]

## Conceptual Models and Theories in Jeff's First Version of "Twinkle"

Jeff's actions, as he constructs a bell-path for "Twinkle" and then invents a notation for both the path and the actions to be performed on it, are based on a number of suppositions:

- The pitch-events of the tune are embodied by the bells that reproduce them.[8] In this way, each pitch-event (which is invisible and ephemeral) is given visible manifestation by a bell.
- Each of the pitch-events of the tune is unique. It follows that there will be a different bell for each pitch-event; Jeff does *not*, at this point, recognize the pitch equivalence of the two G-bells or of the two C-bells.
- The succession of temporal events that make up the tune can be represented through the spatial disposition of the bells. Since this succession is

7. It is worth noting that Jeff's notation is tied to his bell-path: were the additional space between the fourth and fifth bells eliminated, his notation would almost assuredly lead to a distortion of the temporal spacing of the pitches of "Twinkle."

8. By "pitch-event" I mean a pitched sound framed relative to some temporal context (so that it is an "event"). As will be seen, because building the tune involves arranging the bells in space, the temporal aspect of pitch-events is almost always conflated with spatial location.

conceived of as linear, the bell-path is likewise linear; similarly, the pause between the two subphrases is represented by a greater amount of space between the fourth and fifth bells.

- Playing the tune is a matter of performing actions along a path specially constructed for this tune. These actions are unidirectional—the player starts at one end of the bell-path and proceeds straight through to the other—and mimic the process of building the path. Playing the tune constitutes one large, continuous motion, articulated by the smaller motions of striking each individual bell.

These suppositions, the relationships between them, and the way they help guide Jeff's actions as he sets up his bell-path can be characterized in terms of four sorts of cognitive constructs: concepts, conceptual models, conceptual domains, and theories.

*Concepts.* Musical concepts, as noted in chapter 1, have the following attributes: they are a product of processes of categorization, they are part of the means through which we guide present and future action, and they are not necessarily linguistic. The concepts basic to Jeff's perspective include (but are not limited to) those associated with pitched sounds, with the individual Montessori bells, with the pitch-events that make up the first phrase of "Twinkle," and with the notion of performing the tune (as distinct from, say, listening to the tune).

*Conceptual models.* Conceptual models consist of concepts in specified relationships. Three conceptual models play a prominent role in Jeff's figural strategy. The first model correlates pitch-events and objects. Accordingly, the constituent pitch-events of "Twinkle" are construed as objects that endure even after the acoustic signal has faded away. This perspective is reinforced by the physical and visual evidence provided by the bells. In turn, each bell can stand for a pitch-event, even when it has not been struck.[9] The second model translates temporal relationships into spatial relationships. Events separated by equal amounts of time are interpreted as objects separated by equal amounts of space; events separated by unequal amounts of time are interpreted as objects separated by unequal amounts of space. This model informs Jeff's arrangement of the bells on the table, so that there is a greater space between the two bells that mark the end and the beginning of the two consecutive subphrases (the fourth and fifth bells) than between the bells within each subphrase. The third model links physi-

9. While I believe that the conceptual models I describe here accurately reflect the cognitive structures that guide Jeff's actions, they should be generalized only with great caution. Bamberger notes that some children believe that a bell will sound differently if it is moved to another place. Indeed, there is good evidence that Jeff is more concerned with spatial location than with pitch. For discussion of a particularly revealing moment in his development, see Bamberger, *The Mind behind the Musical Ear*, 204–09. To simplify matters I shall continue to focus on pitch-events (since these ultimately inform the "meaning" given to any particular location along the bell-path or action-path), with the understanding that the full description of an "event" requires specification of both its temporal and spatial coordinates.

cal actions with a sequence of sounds. As a result, "Twinkle" acquires a corporeal dimension to go with its acoustic dimension. Each action produces another of the pitch-events of "Twinkle," and, just as the pitches of "Twinkle" form a coherent whole, so does the sequence of physical actions.[10]

*Conceptual domains.* Each of these conceptual models sets up a conceptual domain. The first model sets up a domain of sound-objects, populated by physical embodiments of ephemeral sonic events. The second model sets up the domain of pitch-space, quite apparent in the arrangement of bells on the table. As Jeff has constructed it, pitch-space consists of temporal and structural relationships between sound-objects. The third model sets up a domain associated with a practiced sequence of actions, what Bamberger calls a "felt path."[11] It should be noted that these conceptual domains are not completely exclusive of one another: in general, conceptual domains will overlap to the extent that the conceptual models on which they are based share concepts and relationships.

*Theories.* Theories coordinate a number of conceptual models in order to guide inference and reason. In this case, the three conceptual models described above are coordinated by a theory—Jeff's theory—of "Twinkle." According to this theory, a representation of the tune can be constructed by arranging a series of unique sound-producing objects in space in order of their appearance in the tune, producing a unidirectional path that, if traversed with the appropriate actions, will yield a sounding version of the tune. Jeff's notation for "Twinkle" is a symbolic manifestation of this theory: each sound-producing object is represented by a separate line.[12] The uniqueness of these objects is indicated by the series of numbers that is placed beneath these lines; the actions that will yield the tune are indicated by the numbers that are placed above.

There will be more to say about each of these constructs, but first let us see how Jeff's theory of "Twinkle" changes as Bamberger alters one aspect of the task of constructing a representation of the tune.

## Jeff's Second Version of "Twinkle"

Jeff continued to work with "Twinkle" for a number of weeks and became quite adept at building the tune according to his first theory. One day, as an outgrowth of a pitch-matching task that she had asked Jeff to do, Bamberger changed the set of

10. In some cases, these physical actions are constitutive of the tune in a way that the arrangement of the bells is not. For instance, the bell-path gives no clues as to how many times each bell should be hit, which, of course, is an integral part of the action-path.

11. Bamberger, *The Mind behind the Musical Ear*, 213.

12. My use of "symbol" with respect to Jeff's notation is not very heavily freighted—the symbols are at best ad hoc, and under different circumstances Jeff might use quite different symbols to accomplish the same notational goals. However, I do want to say that Jeff's notation correlates with his concepts—that is, the marks he puts down stand as symbols for some of the concepts in terms of which he understands "Twinkle."

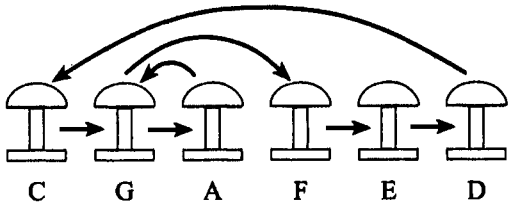



FIGURE 3.4 Jeff's modified bell-path for the first phrase of "Twinkle, Twinkle"

bells Jeff had to work with by removing the brown-based bells.[13] This left Jeff with only one bell for each pitch-type. Jeff built the tune as usual until he came to the place where, in accordance with the strategy based on his first theory, he would find and position the brown-based G-bell. When he could not find the bell, he stopped his work and said, "I need another bell."

Bamberger did not press Jeff further at this point, but a few days later she returned to the task. This time, Jeff decided to look for the needed bell among those he had already arranged in the bell-path. After first trying the C-bell, he moved to the G-bell (second in the bell-path) and discovered that it provided the pitch he was looking for. Jeff eventually was able to build up the entire tune using only the white-based bells, creating the bell-path and action-path shown in figure 3.4. While this arrangement of the bells is more parsimonious—only six bells are needed instead of eight—the action-path through the bells is more complex, requiring both switchbacks and leaps over intervening bells. As a result, the bell-path and action-path are no longer in simple correspondence with one another.

As before, Bamberger asked Jeff to come up with a notation that would enable someone else to play "Twinkle" on the bells as he had arranged them. This time, however, the complexities of the action-path required to play the tune on the new arrangement of bells posed some challenges, and it was only after a number of attempts that Jeff was able to develop the notation shown in figure 3.5. The notation shows two important innovations. First, by means of arrows, Jeff is able to show just how one negotiates the switchbacks and leaps necessary to play "Twinkle." Second, the numbers now refer to only one thing: the number of hits each bell gets. The serial notation Jeff used for his first version of "Twinkle" has been abandoned, since there is no longer a one-to-one mapping of bells to pitches.

## Conceptual Models and Theories in Jeff's Second Version of "Twinkle"

While some of the suppositions behind Jeff's construction and notation of the second version of "Twinkle" are the same as those for his first version, some are different. Unchanged are the notions that the pitch-events of the tune can be embod-

13. Bamberger did this in part to see if she could move Jeff from a figural representation of the tune toward a formal representation. For a fuller account of this aspect of Bamberger's work with Jeff, see *The Mind behind the Musical Ear*, chap. 10.

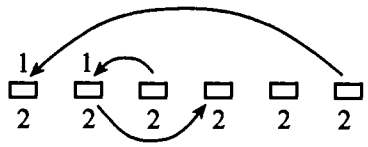



FIGURE 3.5 Jeff's notation for his modified bell-path for the first phrase of "Twinkle, Twinkle"

ied by the bells that reproduce them and that playing the tune is a matter of performing actions along a path in a specially constructed pitch-space. Changed are Jeff's ideas about pitches and about the relationship between temporal events and spatial disposition. In a conceptual shift that ultimately had profound implications for Jeff's understanding of musical organization (as Bamberger's further work with Jeff showed), he now recognizes that pitch-events are not entirely unique: some pitches will be repeated over the course of a tune. A given bell can thus serve as a sound-source for more than one pitch-event. This realization has the consequence of loosening the relationship between when pitches happen in the tune and how representatives of these pitches can be disposed in space. As a result, playing "Twinkle" is no longer a matter of following a unidirectional path through pitch-space. Even though time might still be conceived of as moving uniformly in one direction, the motions needed to play "Twinkle" are no longer unidirectional and serial: the player moves forward and backward, sometimes striking the bells serially, sometimes leaping over one or more bells to find the next pitch of the tune.

The conceptual shift associated with Jeff's changed ideas about pitches, pitch-space, and temporal order is an example of what Douglas Hofstadter calls *conceptual slippage*. As Hofstadter defines it, conceptual slippage is the context-induced dislodging of one concept by a closely related one, within the mental representation of some situation.[14] Here, the context is provided by Bamberger's removal of the brown-based bells from those Jeff had to work with in constructing "Twinkle." What has slipped is Jeff's concept of a pitch-event—specifically, how pitch-events are unique. Jeff still clearly realizes that the G that comes after the first C of "Twinkle" is different from the G that comes after the first A—after all, they occur at different points in the tune. He now knows, however, that both can be played on the same bell. This in turn means that the temporal uniqueness of pitch-events will not necessarily translate into spatial uniqueness: the two distinctly different Gs of the tune can be found in the same place on the bell-path. Jeff's concept of pitch-event as *a specific pitch in "Twinkle"* has been dislodged by a concept of pitch-event as *a general type of sonic event that is manifested at a particular point in the temporal progress of "Twinkle."*

Because all of the conceptual models through which Jeff structures his under-

14. Douglas R. Hofstadter and the Fluid Analogies Research Group, *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought* (New York: Basic Books, 1995), 198. The idea of conceptual slippage has proven quite important in Hofstadter's work in that it introduced a level of flexibility and fluidity into the computer programs his group developed, features that were absent from earlier artificial intelligence programs.

standing of "Twinkle" incorporate the concept of pitch, the slippage of this concept cascades through the system. The correlation between sounds and objects provided by the first conceptual model is now looser than before, and the domain of sound-objects it sets up is populated with generic, rather than specific, entities. In consequence, the functional distinction Jeff made between brown-based bells and white-based bells (preferring to work with the latter first) is no longer necessary—now there are only bells. The second model, which maps temporal relationships onto spatial relationships, is more profoundly affected. While temporal and spatial events are still connected (Jeff's bell-path still has a beginning and an end), the independence of temporal succession and the spatial arrangement of sound-sources is far more evident. Reflecting this, the domain of pitch-space has been reshaped: once pitches are conceived as generic entities, they start to have more to do with each other than they do with the specific sounds that make up "Twinkle." These alterations in the notion of spatial disposition then influence the third conceptual model. While the physical actions that produce "Twinkle" can still be seen as a coherent whole, it is a whole made of a larger number of disparate parts (including switchbacks and leaps). This creates a *physical* manifestation of the growing separation of the tune from the medium through which it sounds: one can now get lost along the action-path in a way that would have been unlikely with the first bell-path Jeff constructed.

As a result of these changes, Jeff's theory of "Twinkle" is changed as well. According to this modified theory, "Twinkle" can be built using various generic sound-producing objects in space in the approximate order of their occurrence in the tune. A relatively complex action-path is then imposed on this arrangement, which, if properly executed, will yield a sounding version of the tune. Again, Jeff's notation for "Twinkle" is a symbolic manifestation of this theory: there are no longer unique entities that correspond to each of the constituent pitches of the tune, and performing the tune involves a series of complex actions. In terms of the two basic strategies Bamberger observed, Jeff has begun to move from a figural representation to a formal one.

## The Function and Structure of Theories

At this point, it will be useful to pause and consider five features of theories revealed by Jeff's two approaches to constructing "Twinkle" on the bells.

First, theories serve as a guide for actions, both mental and physical. Each of Jeff's theories of "Twinkle" allows him to create an arrangement of the bells that will enable him to play the tune. Using these theories, Jeff would be able to find a way to arrange similar objects—glasses filled with varying amounts of water, marimba bars, pots and pans—so that he could play "Twinkle" on them. Should something change in one of these arrangements (for instance, if one of the bells was covertly switched for a bell with a different pitch), Jeff would most likely be able to find the problem and suggest ways of remedying it. In order to avoid confusion later, I want to broaden the characterization of this feature and say that theories serve as a guide to understanding and reasoning. Although this seems a bit grand when applied to Jeff's situation—after all, Jeff is someone who is much more predisposed to action

than to talking about action—it makes clear that the actions with which I am concerned are not of an automatic or unconscious sort but are a manifestation of either implicit or explicit patterns of reasoning.

Second, theories provide answers to conceptual puzzles. When Jeff, using his first theory of "Twinkle," could not find a brown-based bell to complete the tune, his theory told him that he could not continue building the tune, and he stopped. Of course, this is not a very satisfying answer to the puzzle, not the least because it points to the inadequacy of the theory as a guide for further actions. Jeff's second theory of "Twinkle," in which the concept of a pitch-event has slipped toward a more generic perspective, is able to provide a solution by reinterpreting the function of one of the bells already in the bell-path, thereby permitting him to continue building the tune.

Third, theories simplify reality. Jeff does not attempt to make the proportional spacing of the bells replicate the temporal intervals between the pitch-events of the tune—it is enough that there is a somewhat larger space between the bells that end the first subphrase and those that begin the second. Jeff's notation (and, presumably, his theory) also do not specify the tempo at which "Twinkle" should be played, how hard the bells should be struck, whether the player should be seated or standing, and so on.

Fourth, theories are relatively extended cognitive structures. As analyzed here, both of Jeff's theories involve at least three conceptual models, which describe relationships among a number of concepts and which set up conceptual domains. Each conceptual model can operate independently of the others: construing pitches as objects does not necessarily require arranging the objects in any particular way, nor does it necessarily involve physical actions performed on these objects (to which mental rehearsal by musicians bears testament); temporal successions (particularly those that do not involve music) can be represented spatially without involving pitch-objects, and so on. Theories, however, coordinate such models into systems for inference. The systematic quality of theories (in the sense that they involve coordinated—and thus interrelated—conceptual models) is evident in the way the slippage of Jeff's concept of pitch-event cascades through his theory of "Twinkle": it affects each of the conceptual models involved.

Fifth, theories are dynamic. When Jeff's concept of pitch-event changes, his theory of "Twinkle" changes as well. Theories thus respond to changes in circumstances and (as Bamberger's further work with Jeff shows quite clearly) can grow even more extensive as they are developed to deal with a broader range of circumstances.

Although Jeff's efforts to construct and notate bell-paths for playing "Twinkle" give a glimpse into how we structure our understanding of music, the view is limited by the specificity of his situation, which is that of a young boy struggling with basic musical concepts. In order to know what it means to have a theory, and to connect Jeff's functional and implicit theories with the abstract and explicit theories of writers like Rameau and Schenker, we need to consider in greater depth the nature and function of the cognitive structures—conceptual models and theories—basic to our understanding of the world as a whole.

## CONCEPTUAL MODELS AND THEORIES

As chapter 1 demonstrated, the last quarter of the twentieth century bore witness to a great deal of research on categorization—research that fundamentally changed our understanding of the process of categorization and its role in human cognition. And as shown by chapter 2, at about the same time, work in metaphor theory and cross-domain mapping blossomed, resulting in similar changes to our understanding of the role played by mapping structure from one domain onto another. It does not appear, however, that we will come to regard this time as one in which work on conceptual models and theories came to fruition. While much interesting work has been done—work to which I shall refer in the following—it has not converged on broad agreement as to what counts as a conceptual model (or even what such structures should be called), nor has it produced any unified view of what constitutes a theory.

Two factors have contributed to this situation. First, the functions we call "the mind" and the organ we call "the brain" are both staggeringly complex—indeed, Gerald Edelman has called the brain the most complicated material object in the known universe.[15] Further, mind and brain are not coextensive in any simple way: mind functions extend at the least to the central nervous system, and much of what the brain does is not directly involved with what we conventionally accept as mind functions.[16] Mind and brain thus admit of any number of structural descriptions, and in consequence there is great diversity in how writers have characterized constructs like conceptual models and theories. The second factor, which is related to the first, is that researchers have tended to focus either on conceptual-model-like structures or on theory-like structures, but rarely on both. The advantage of this strategy is that it limits the amount of cognitive structure that has to be explained. The disadvantage is that it makes it difficult to explain how relatively low-level structures like conceptual models coordinate with relatively high-level structures like theories.

In what follows, I give an account of cognitive structure that begins with concepts and extends through theories. In line with my analysis of the cognitive structure behind Jeff's strategies for building "Twinkle," I distinguish among concepts, conceptual models, conceptual domains, and theories. I should make clear that I assume humans rarely, if ever, make explicit reference to structures of this sort in the process of reasoning about a given situation. Instead, as is evident from Jeff's responses to the challenges presented by building tunes with the Montessori bells, we most typically are completely unaware of things like concepts and conceptual models, no matter how evident they are in our patterns of reasoning. Concepts, conceptual models, conceptual domains, and theories are, by and large, implicit cognitive structures: they can be teased out through experiment or analysis, but they need

15. Gerald M. Edelman, *Bright Air, Brilliant Fire: On the Matter of Mind* (New York: Basic Books, 1992), 17.

16. Things get even more complicated when emotions and feelings are factored into our account of cognition, for there is strong evidence that these are both "of the body" and "of the mind." Antonio Damasio has offered compelling arguments on this point; see in particular Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain* (New York: Avon, 1994); and idem, *The Feeling of What Happens.*

not be explicit to be useful to us in reasoning about the world. I also do not want to suggest that these constructs have unitary representations, either cognitively or physiologically. Indeed, recent work indicates that the lines between entities such as concepts, models, domains, and theories are more often blurry than clear and that there are no simple neurobiological explanations for any of them. My distinctions are thus pragmatic ones, intended to reflect the different sorts of cognitive work done at different levels of structural complexity and the compass of the conceptual structures that result.

In the first subsection that follows, my focus is on conceptual models and their relationships to concepts and conceptual domains. In the second subsection, I turn to theories and explore what recent research in anthropology and developmental psychology can tell us about what it means to have a theory.

## Conceptual Models

BASIC FEATURES OF CONCEPTUAL MODELS As I define it here, a conceptual model consists of concepts in specified relationships. To get a better idea what this means, let us return to one of the models used by Jeff in his initial attempts to construct a bell-path for "Twinkle," the model that translates temporal events into spatial events. The concepts required for the model are event and object. Both of these are very broad Type 1 categories: the typical event is something that happens in time; the typical object is something that persists over time. And, as always with this sort of category, there are boundary cases that challenge what is typical: a rainbow is in some respects an event (since it happens only as a result of certain atmospheric conditions), but, in other respects, an object (since it persists—if briefly—over time). In the model, events and objects are related to one another by a cross-domain mapping between pragmatic construals of time and space, which interprets durations (that is, the amount of time between events) as spatial dispositions (that is, the amount of distance between objects).[17] Large units of time are represented by distantly spaced points, small units of time are represented by closely spaced points, and equivalent temporal units get equivalent spacing. This way of interpreting time can be seen in the distribution of marks for minutes on the face of an analog clock: marks for five-minute intervals are relatively far apart, marks for one-minute intervals are relatively close together, and equivalent units get equivalent spacing.[18]

Conceptual models are drawn from two sources. The first, and perhaps most important, is culture, especially as it is represented by the people around us. We learn conceptual models by observing and imitating the actions of others and by abstracting models from the lessons we are taught. Conceptual models are thus part of the shared body of knowledge that constitutes culture. As Naomi Quinn and Dorothy

17. Modern physics, of course, does not retain the clear separation between the categories time and space that is necessary for this mapping.

18. The spatial representation of temporal units on the face of an analog clock reflects in part the mechanism that moves the hands at a steady rate across the clock face. Compare this to the face of a sundial, on which the marks for equal units of time are not spaced equally.

Holland put it, culture is "not a people's customs and artifacts and oral traditions, but what they must know in order to act as they do, make the things they make, and interpret their experience in the distinctive way they do."[19] Conceptual models may also be created through cross-domain mapping. The propositions and structure of a model framed relative to one domain are applied to another domain; if the process is successful (if it allows us to structure effectively our knowledge of some aspect of the target domain), a new conceptual model is born.

Models, once learned or created, are stored in memory and called up to organize our understanding of a given set of circumstances. Each of the models Jeff used in his theory of "Twinkle" was learned or developed before he attempted to build the tune (even if, as in the case of the model that correlates pitch-events with objects, it was learned only when he worked with "Hot Cross Buns") and was recalled from memory when needed for the new task.

Although there has been wide interest in constructs like conceptual models, there has been an almost equally wide variety of construals of mental models of this sort.[20] There is general agreement on two basic features, however. First, conceptual models are of relatively restricted extent—that is, they structure a quite limited portion of our knowledge of the world. Second, conceptual models are abstractions from experience—they do not replicate the outside world in all its detail but model

19. Naomi Quinn and Dorothy Holland, "Culture and Cognition," in *Cultural Models in Language and Thought*, ed. Dorothy Holland and Naomi Quinn, Cambridge Readings in the Literature of Music (Cambridge: Cambridge University Press, 1987), 4.

20. Analogues for conceptual models, and the relevant literature, include the following.

MENTAL MODELS: Philip Nicholas Johnson-Laird, *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness* (Cambridge, Mass.: Harvard University Press, 1983); Lawrence W. Barsalou, *Cognitive Psychology: An Overview for Cognitive Scientists* (Hillsdale, N.J.: Erlbaum Associates, 1992); Dedre Gentner and Albert L. Stevens, eds., *Mental Models* (Hillsdale, N.J.: Erlbaum Associates, 1983); Jane Oakhill and Alan Garnham, eds., *Mental Models in Cognitive Science: Essays in Honour of Phil Johnson-Laird* (East Sussex: Psychology Press, 1996).

IDEALIZED COGNITIVE MODELS: George Lakoff, *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind* (Chicago: University of Chicago Press, 1987); Robert N. McCauley, "The Role of Theories in a Theory of Concepts," in *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*, ed. Ulric Neisser (Cambridge: Cambridge University Press, 1987), 288–309.

CULTURAL MODELS: Quinn and Holland, "Culture and Cognition"; Roy G. D'Andrade and Claudia Strauss, eds., *Human Motives and Cultural Models*, Publications of the Society for Psychological Anthropology (Cambridge: Cambridge University Press, 1992); Roy G. D'Andrade, *The Development of Cognitive Anthropology* (Cambridge: Cambridge University Press, 1995); Bradd Shore, *Culture in Mind: Cognition, Culture, and the Problem of Meaning* (New York: Oxford University Press, 1996).

COGNITIVE DOMAINS: Ronald W. Langacker, *Theoretical Prerequisites* and *Descriptive Application*, vols. 1 and 2 of *Foundations of Cognitive Grammar* (Stanford, Calif.: Stanford University Press, 1987, 1992).

FRAMES: Marvin Minsky, "A Framework for Representing Knowledge," in *The Psychology of Computer Vision*, ed. Patrick Henry Winston (New York: McGraw-Hill, 1975), 211–77; idem, *The Society of Mind* (New York: Simon and Schuster, 1985); Douglas B. Lenat and R.V. Guha, *Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project* (Reading, Mass.: Addison Wesley, 1989);

KNOWLEDGE STRUCTURES: Robert P. Abelson and John B. Black, "Introduction," in *Knowledge Structures*, ed. James A. Galambos, Robert P. Abelson, and John B. Black (Hillsdale, N.J.: Erlbaum Associates, 1986).

Needless to say, these constructs are by no means equivalent to each other. However, all involve the basic features I outline for conceptual models.

aspects of it as a guide to understanding. Modifying a distinction drawn by Clifford Geertz, these are models *for* understanding a particular set of circumstances rather than models *of* the circumstances themselves.[21]

Conceptual models are thus a very basic and yet abstract construal of how the world is organized—they are examples of how concepts first become useful through being related to other concepts.[22] Because we cannot think without thinking in terms of relationships of concepts, conceptual models are part of the warp and woof of our knowledge of the world: they represent the way the world *is*. However, this is not to say that conceptual models reflect, in a simple way, an objective and unchanging reality, even if they are crucial to the way we construct *our* reality.

In earlier work, I proposed, following George Lakoff, that models of the sort I am concerned with here were idealized.[23] Idealization is a way to account for instances when there is a poor fit between a conceptual model and actual circumstances. For example, the conceptual model that sets up the notion of a sound-object runs into something of an obstacle when it comes to the human voice, since it is difficult to find the physical object that correlates with vocal sound. Most of the time this difficulty is ignored: confronted with the sounds of the human voice, we continue to think in terms of sound-objects. A conceptual model concerned with sound-objects could thus be regarded as an ideal to which actual circumstances may or may not conform.

I now think it better to regard idealization as belonging to a later stage in conceptualization. As they typically function, conceptual models are immediate and transparent: they represent our working notion of reality. Accounting for why a conceptual model does not fit a given situation belongs to a less immediate level of cognitive structure occupied with reasoning and causal explanation. This is the level of the theory. One job of theories, then, is to explain discrepancies between circumstances and conceptual models, even if this explanation is little more than the recognition that the model works in most, but not all, cases. The distinction reflected in this perspective is one I shall maintain throughout the following: conceptual models are immediate and not consciously evaluative; theories are always evaluative—they provide an interpretation of circumstances and possible courses of action based on this interpretation.

Conceptual models are a response to, not a simple reflection of, the outside world. They deal with only a small portion of that world and represent its structure

21. Geertz draws his distinction between models without reference to cognitive processing, focusing only (but quite importantly) on how models are used. A "model for" is used to guide physical relationships; a "model of" is a symbolic representation of a particular situation. See Clifford Geertz, "Religion as a Cultural System," in *Anthropological Approaches to the Study of Religion*, ed. Michael Banton (New York: Frederick A. Praeger, 1966), 7–8. While I assume that conceptual models have a symbolic representation, I do not assume these symbols are necessarily available for conscious manipulation (that is, they may be what Lawrence Barsalou calls "perceptual symbols," which were mentioned above in chap. 2).

22. In function and structure, conceptual models bear a resemblance to the very simple, image-schematic stories described by Mark Turner in *The Literary Mind* (New York: Oxford University Press, 1996), chaps. 2 and 3. Such stories may involve little more than a record of the action of throwing a ball but nonetheless are essential to our understanding of the world.

23. Lawrence Zbikowski, "Large-Scale Rhythm and Systems of Grouping" (Ph.D. diss., Yale University, 1991), chap. 4.

in ways that have more to do with cognitive efficiency than with accuracy. Nonetheless, changes in the structure of the outside world will at times require changes to a conceptual model. As we saw in the preceding discussion, the absence of the brown-based bells caused Jeff's concept of a pitch-event to slip, resulting in changes to the conceptual models basic to his theory of "Twinkle." On a more local level, biological processes will have a similar, if more incremental, effect on the structure of our conceptual models. For instance, every time we remember something, we strengthen some synaptic connections and weaken others, "re-wiring" our brain on a micro-level and subtly transforming the memory we would seek to recall. Such changes will have an influence, slight but cumulative, on our conceptual models. Conceptual models are thus dynamic at both the global and the local level, although, because of their immediacy, this fluidity is rarely evident.

REPRESENTING CONCEPTUAL MODELS My thinking about conceptual models has been shaped in important ways by research in artificial intelligence from the 1970s and 1980s—in particular, the work of Marvin Minsky and of Roger Schank and Robert Abelson[24]—and I find it useful to characterize conceptual models in terms of simple propositions linked together to create a very basic system of inference. I thus represented the conceptual models in chapter 1 as a collection of interlinked nodes and shall use similar representations in chapters 4 and 5. I should note, however, that I propose these models, and this way of representing them, much in the spirit of that early work in artificial intelligence: as a place to start a description of the processes basic to reason and inference. The models are generalized, and they are presented in a way that connects with a wide body of research. But they are not meant to represent a true, or even empirically verified, picture of what humans have in their heads when they reason about various situations. This is not to say that there is no empirical evidence for conceptual models, only that my purpose in this book is to show how conceptual models are used rather than to make an argument for exactly what they comprise or how they are structured. As I discuss in the concluding chapter, I believe much work has to be done before a detailed picture of the structure and function of conceptual models can be developed.

CONCEPTUAL MODELS AND CONCEPTUAL DOMAINS In the conceptual model for sound-objects, the concept of a sound is linked with the concept of an object through a process of cross-domain mapping. Elements from the invisible and ephemeral domain of sound are correlated with those from the visible and enduring domain of objects, and structure is mapped between the domains, giving rise to sound-objects. The domains involved in this mapping—the domain of sound and the domain of objects—are broad and relatively generic. In comparison, the domain of sound-objects is rather more specific. I shall call the more specific domains set up by conceptual models *conceptual domains*.

The specificity of conceptual domains is manifested in two ways. First, because

24. Minsky, *The Society of Mind*; Roger C. Schank and Robert P. Abelson, *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures* (Hillsdale, N.J.: Erlbaum Associates, 1977).

some of the relationships proper to a conceptual domain may not obtain in the more generic domain (or domains) from which it descends, the scope of the conceptual domain is more limited: for example, not everything that can be said of sound-objects can be said either of sounds or of objects. Second, conceptual domains are typically associated with specific phenomena or experiences. The conceptual domain of sound-objects set up by Jeff's model pertains to his experience with the Montessori bells and with the task of building a bell-path for "Twinkle." This particular domain is instantiated by the various bells used to construct the bell-path for "Twinkle," but it can be expanded to include things whose relationship to sounds or objects is more abstract, such as marks on a page or physical places on a musical instrument.

I regard conceptual domains as distinct from mental spaces; again, this is a departure from my earlier work. As described in chapter 2, mental spaces are complex and transitory sites for meaning construction and for reasoning. Furnishing a mental space requires drawing on a number of coordinated conceptual domains. One way to achieve this coordination is through the recursive application of conceptual models. Through recursion, the structuring principles that give rise to a conceptual model are subsequently applied to the *elements* of a conceptual domain originally created by the same, or another, conceptual model. For example, consider applying a conceptual model for anthropomorphism to the conceptual model for sound-objects. (Anthropomorphism, it will be recalled from the discussion of anthropomorphic blends in the second part of chapter 2, maps the properties of living beings onto objects.) This application maps the properties of living beings onto sound-objects, yielding a mental space populated with animistic pitches. Admittedly, such a space is fairly fantastic—even those with a steadfast belief in the viability of an organicist interpretation of pitch structure usually rely on conceptual blends rather than on the rather curious idea of animistic pitches—but not entirely dissimilar from mental spaces populated by pitches that obey the laws of classical physics.[25]

Another way to achieve the coordination of conceptual domains is by linking conceptual models into a system for reasoning about the world—that is, by creating a theory.

## Theories

When I outlined the structure and function of theories in the preceding discussion, I did so with reference to an eight-year-old boy's attempts to arrange some rather unusual bells so as to play the first phrase—and only the first phrase—of "Twinkle, Twinkle, Little Star." But is this really a theory in the way Darwin's theory of natural selection is a theory, or in the way Einstein's special theory of relativity is a theory?

Scientific theories, viewed from a broad perspective, have the features of theories

25. Organicist construals of pitch relationships are discussed in chap. 7. Early in his career, Jean-Philippe Rameau construed pitch relations in terms of Newtonian physics; for a discussion, see Thomas Christensen, *Rameau and Musical Thought in the Enlightenment*, Cambridge Studies in Music Theory and Analysis, 4 (Cambridge: Cambridge University Press, 1993), 7–11.

I outlined above. They serve as a guide for understanding and reasoning, and they provide answers to conceptual puzzles. Thus Darwin's theory sought to explain (among other things) why finches on various islands of the Galápagos group had different-shaped beaks and why similar variations could be observed in other species.[26] Scientific theories also simplify reality. For instance, Darwin's theory did not take into account the effect of inheritance on variation: if natural selection were the only process affecting variation, variation should become less with each successive generation until it finally disappears altogether. However, because there is a recombination of parental traits in offspring, each successive generation has new variations. Finally, scientific theories are extended cognitive structures that are subject to change: Darwinian evolutionary theory was eventually refined to take into account the effect of inheritance, as well as natural selection, on the evolution of species.

Despite these similarities, there are nonetheless significant differences between scientific theories and theories of the sort that Jeff used in his encounters with "Twinkle." Jeff's theories are of quite limited extent, pertaining only to the way bells or similar objects might be arranged to play a particular song, and they have the limited function of guiding his actions within this specific context. Scientific theories, by contrast, attempt to account for as broad a range of phenomena as possible, and they offer predictions that can be proved true or false by any of a number of researchers working independently. These features have contributed to the success of scientific theories as a type, a success demonstrated by the incredible explosion of knowledge in the hard sciences over the past two centuries. It is no wonder that explorations of the structure and function of scientific theories generated an extensive literature in the twentieth century and that scientific theories have become a standard to which all other theories are compared.

The distance, then, between Jeff's theories and scientific theories is considerable, enough so that using "theory" for the former might seem ill advised. Nonetheless, work by anthropologists and others over the past thirty years has provided evidence that scientific theories are but one sort of general theory and that the process of theorizing is both various and omnipresent. More recently, research by developmental psychologists has shown that even very young children make use of theory-like structures to guide their actions. These two lines of research, which I review briefly in the following paragraphs, do not diminish the distance between Jeff's theories and those of modern science, but they do suggest that these two types of theory represent points along a continuum rather than discrete entities.

CULTURAL THEORIES The notion of a cultural theory (as distinct from a scientific theory) came to prominence in the 1960s and reflected a growing aware-

26. One can see the puzzle that challenged Darwin emerging in his earliest writings. After noting, in *The Voyage of the Beagle* (first published in 1839, three years after the *Beagle* returned home), the gradation in the sizes of the finches's beaks, Darwin makes this prescient comment: "Seeing this gradation and diversity of structure in one small, intimately related group of birds, one might really fancy that from an original paucity of birds in this archipelago, one species had been taken and modified for different ends." Charles Darwin, *The Voyage of the Beagle*, with an introduction by Leonard Engel (Garden City, N.Y: Doubleday, 1962), 381 (chap. 17).

ness among anthropologists that Western models for representing knowledge, including those of anthropology, were not always adequate or appropriate for non-Western cultures. Here are three examples of the kind of knowledge that caused anthropologists and those working in related fields to rethink what was meant by a theory.

1. Research by Brent Berlin and his associates on ethnobotanical classification (mentioned in chapter 1 in connection with the way processes of categorization were reconceptualized during the 1970s) demonstrated that the taxonomic systems of indigenous cultures could be every bit as complex as those of Western biologists. Berlin's further work has shown that ethnobotanical classification reflects, in great part, the order that is inherent in nature and that it also achieves a measure of comprehensiveness by classifying unknown species through making reference to known species.[27]
2. For more than a thousand years, long-distance noninstrumental navigation has been practiced in Polynesia and Micronesia. Experienced navigators routinely sail their outrigger canoes up to one hundred fifty miles between islands in an area where less than 0.2 percent of the surface is land. Edwin Hutchins—perhaps uniquely qualified as someone with expertise in navigation, ethnography, and artificial intelligence—has shown that this method of navigation is based on a computational system distinct from (yet comparable to) that of Western navigation.[28]
3. The cure for snakebite among the Rama Indians of Nicaragua is a complicated affair involving ritual, diet, beliefs in the supernatural, and herbal medicines especially prepared by snakebite doctors. For instance, after someone has been bitten, the victim must make sure the snake is killed; if the snake is not killed, the Rama believe the person cannot be cured. While under cure, the patient is placed on a special diet, with foods prepared in a special way; the Rama believe that this is necessary to avoid endangering the lives of other members of his or her household. Although it is difficult to sort out just which aspects of the snakebite cure have medical efficacy and which do not, the cure appears to be reasonably effective, saving the patient over three-fourths of the time.[29]

One way to account for the stability and efficacy of the knowledge represented by each of these sets of cultural practices is to construe it as organized by a theory (a theory of botanical classification, or navigation, or snakebite cure). In that this knowledge is specific to a particular culture or ethnographic context, anthropolo-

27. Brent Berlin, *Ethnobiological Classification: Principles of Categorization of Plants and Animals in Traditional Societies* (Princeton, N.J.: Princeton University Press, 1992); idem, "How a Folkbotanical System Can Be Both Natural and Comprehensive: One Maya Indian's View of the Plant World," in *Folkbiology*, ed. Douglas L. Medin and Scott Atran (Cambridge, Mass.: MIT Press, 1999), 71–89.

28. Edwin Hutchins, *Cognition in the Wild* (Cambridge, Mass.: MIT Press, 1995), 65–93.

29. Franklin O. Loveland, "Snakebite Cure among the Rama Indians of Nicaragua," in *Medical Anthropology*, ed. Francis X. Grollig and Harold B. Haley (The Hague: Mouton, 1976), 97.

gists have designated the corresponding theories cultural theories, folk theories, or ethnotheories.[30]

One thing that distinguishes cultural theories from scientific theories is that scientific theories have an explicit tradition of verification: it has to be possible to replicate experimental results that support the theory, or the theory is cast into question. While cultural theories clearly have efficacy—the classification of a wealth of plant life, navigation between distant islands, and cure from snakebite all actually occur—they are not subjected to the open and constant scrutiny that marks the modern scientific method. Cultural theories, in contrast to scientific theories, organize a relatively stable body of knowledge that changes only gradually—and most often imperceptibly.[31] Cultural theories, as a further manifestation of cultural knowledge, thus both reflect and constitute culture.

As I construe them, cultural theories need not be explicit; indeed, theories are often manifested as patterns of behavior rather than as explicit statements describing how a specific situation should be understood or what actions are appropriate for a given set of circumstances.[32] This is particularly important for music, for theories of music may be both implicit and nonverbal.[33] A glimpse of an implicit and at least partially nonverbal theory of music was provided by Steven Feld's research on the Kaluli of Papua New Guinea, which was touched on in chapter 2. Feld showed that Kaluli terminology centers around waterfall imagery, which yields a number of highly distinctive ways to describe the contour of musical descents. By contrast, there is only a single term in Kaluli to describe musical ascents.[34] This does not mean that the Kaluli conception of musical ascent is undifferentiated—were an

30. Of course, "ethnotheory," as a sort of Third World shadow of "theory," is not a very comfortable concept: "theory" always remains the privileged term. For discussion of the related topic of the relationship between ethnopsychology and psychology, see Catherine Lutz, "Ethnopsychology Compared to What? Explaining Behavior and Consciousness among the Ifaluk," in *Person, Self, and Experience: Exploring Pacific Ethnopsychologies*, ed. Geoffrey M. White and John Kirkpatrick (Berkeley: University of California Press, 1985), 35–79.

For an illuminating discussion that also bears on this matter, framed with respect to folk models (which have the function I here ascribe to theories), see Ladislav Holy and Milan Stuchlik, "The Structure of Folk Models," in *The Structure of Folk Models*, ed. Ladislav Holy and Milan Stuchlik (London: Academic Press, 1981), 1–34.

31. Although historical evidence suggests that changes in the theories (and the corresponding conceptual models) that are constitutive of culture have often been so gradual as to escape notice, the incursion of Western commercial and technological culture into traditional societies during the past two decades has accelerated the rate of change markedly—all of the researchers I cite note the effect of this incursion into the cultural practices they study. Further discussion of this point, framed relative to musical practice, is taken up in chap. 5.

32. My construal of cultural theories as at least potentially implicit differs from that of Roy D'Andrade, who has proposed that cultural theories are always explicit. See D'Andrade, *The Development of Cognitive Anthropology*, 172–73.

33. My thinking here follows that of John Blacking, who proposed that implicit theory embraces both verbal and non-verbal discourse about music. See Blacking, "Music, Culture, and Experience," in *Music, Culture, and Experience: Selected Papers of John Blacking*, ed. Reginald Byron, with a foreword by Bruno Nettl (Chicago: University of Chicago Press, 1995), 228–31.

34. Steven Feld, "Flow Like a Waterfall: The Metaphors of Kaluli Musical Theory," *Yearbook for Traditional Music* 13 (1981): 31–32.

ascending melody played improperly, they would be quick to call attention to the mistake—but that, within the broader cultural context relative to which musical concepts are framed, ascents are not privileged and thus require only the barest of descriptive terms.[35]

In sum, cultural theories have the same general features that Jeff's theories and scientific theories have. They serve as a guide to reasoning and inference, provide answers to conceptual puzzles, and simplify reality. They are extended cognitive structures, and they are subject to change. But cultural theories suggest two extensions to the conception of theory that I have outlined thus far. First, theories should be seen as manifestations of cultural knowledge. Jeff's theories must thus be understood as developed within the context of a larger set of cultural practices, and so should those of modern science. Second, theories can be implicit and to some extent nonverbal. Without a doubt, explicit theories are immeasurably important to the advancement of knowledge, for their precepts are available for public debate, discussion, and modification. Even so, implicit theories make up a great part of our cultural and private knowledge, knowledge that becomes visible only when these theories are subjected to challenge or analysis.

THEORY IN VERY YOUNG CHILDREN The fact that theories are everywhere—in science, in culture, and in our personal construals of the world—suggests that they are basic to human understanding. This perspective has received support from recent research by Alison Gopnik and Andrew Meltzoff, who have shown that children use patterns of reasoning quite similar to those of scientists to organize their understanding of the world. Focusing on experimental evidence gathered from studies with nine- and eighteen-month-old children, Gopnik and Meltzoff have demonstrated that children's understanding of the world replicates the structure and function of scientific theories.[36] For instance, six-month-old infants do not appear to realize that spatial contact is required to move an object: they assume that if the object moved shortly after they performed some action (like kicking their legs), it was this action that caused the object to move. At around eight to ten months, however, infants begin to develop a theory of action on objects—that is, they appreciate that their actions can affect objects only if they are spatially in contact with those objects. As they grow older, children come to realize that this theory does not work in all circumstances: if an object is out of reach and she is given a toy rake, the one-year-old will use the rake to make contact with the object

35. For other work on implicit theories of music, see Simha Arom, *African Polyphony and Polyrhythm: Musical Structure and Methodology*, trans. Martin Thom, Barbara Tuckett, and Raymond Bond (Cambridge: Cambridge University Press, 1991), 139–40; idem, "New Perspectives for the Description of Orally Transmitted Music," *World of Music* 23 (1981): 43; James R. Cowdery, *The Melodic Tradition of Ireland* (Kent, Oh.: Kent State University Press, 1990); and Ruth M. Stone, *Let the Inside Be Sweet: The Interpretation of Music Event among the Kpelle of Liberia* (Bloomington: Indiana University Press, 1982).

36. Alison Gopnik and Andrew N. Meltzoff, *Words, Thoughts, and Theories* (Cambridge, Mass.: MIT Press, 1997). Those who have worked with infants and young children will appreciate the cleverness with which the experiments which produced this data were designed and the rigor with which they were carried out. For a report intended for a more general audience, see Alison Gopnik, Andrew N. Meltzoff, and Patricia K. Kuhl, *The Scientist in the Crib: Minds, Brains, and How Children Learn* (New York: William Morrow, 1999).

(where the six-month-old might simply wave the rake at random), although this contact alone will not make the object come any closer. By eighteen months of age, children refine the theory of action on objects further still, and, now realizing that just touching the object is not enough, an eighteen-month-old will adjust the rake so that it captures the object correctly.[37]

According to Gopnik and Meltzoff's view, infants and young children develop their knowledge much as scientists do. They start with an initial theory for a particular set of circumstances (the relationship between actions and objects). They conduct a variety of experiments (reaching out to grab objects, observing how objects move as a result of their actions). Then they refine and adapt the theory on the basis of the experimental results (when touching the object alone does not produce the desired results, for example, infants and young children will modify the theory to acknowledge that spatial contact must be of a specific sort). Indeed, Gopnik and Meltzoff go even further to argue that it is not so much that there is a scientist in the child as it is that there is a child in the scientist: the cognitive role of theories in earliest childhood is the same as the cognitive role of theories at maturity.

The strong claim that Gopnik and Meltzoff make in *Words, Thoughts, and Theories*—that theory (scientific or developmental) is a means of finding truth and that theorizing is an evolutionary adaptation conducive to the survival of the species—gives occasion for pause, if only because each theory frames "truth" a bit differently.[38] Nonetheless, the weaker version of the claim—that the theories of children resemble those of scientists—is somewhat more promising. Thomas Kuhn, as part of the central conceptual frame for his essay on thought experiments, noted that both children and scientists sometimes hold paradoxical views of phenomena and that both learn when these paradoxes are made evident.[39] According to Kuhn's argument, thought experiments—which introduce no new evidence and would thus seem to offer no opportunity for learning—are one way of making paradoxical views of phenomena evident and thus advancing scientific knowledge. Although only implicit in his discussion, Kuhn appears to consider the thought processes of children to be analogous to those of scientists inasmuch as both experience the same sort of "nonempirical" learning.

While we have much to learn about the cognitive structures that guide the behavior of very young children—it is hardly a pun to say that developmental psychology is still in its infancy—evidence has begun to accumulate in support of the

37. Gopnik and Meltzoff, *Words, Thoughts, and Theories*, chap. 5.

38. Paul L. Harris offers a somewhat different argument against what Gopnik and Meltzoff call the "theory theory" (that is, the theory that children have theories of the world), based on two points: first, that the theories of children and lay adults do not count as true theories; second, that the pursuit of truth is not what guides the scientist's day-to-day activities. As is evident from the preceding, I differ from Harris on the first point; however I am in broad agreement with his second point. See Harris, "Thinking by Children and Scientists: False Analogies and Neglected Similarities," in *Mapping the Mind: Domain Specificity in Cognition and Culture*, ed. Lawrence A. Hirschfield and Susan A. Gelman (Cambridge: Cambridge University Press, 1994), 294–315.

39. Thomas S. Kuhn, "A Function for Thought Experiments," in *The Essential Tension: Selected Studies in Scientific Tradition and Change* (Chicago: University of Chicago Press, 1977), 242–46.

notion that theories are a basic structure of human understanding.[40] Whether the theories of children are of exactly the same sort as those of modern science remains to be seen; however, it does seem that these two sorts of theories have a similar structure and that they are responses to the same impulse to understand the world.

COGNITIVE THEORIES AND MUSIC THEORIES Based on his work building "Twinkle" with the Montessori bells, I do not think that there is a strong case to be made for characterizing Jeff as a scientist. Jeff is not really involved in a wide-ranging search for knowledge, aided by a protocol of experiment and analysis. Jeff is instead doing his best to meet a fairly specific and, for him, relatively challenging task. He is similar to a scientist, however, in that theory is an important cognitive tool, serving to guide his actions and make the elusive world of sound a bit more comprehensible. On the basis of evidence from anthropology and developmental psychology, Jeff is not alone in using theory in this way, for theories are part of the very fabric of human thought. Given this context, the fact that Jeff should have a theory of music, primitive as it may be, is not surprising. Indeed, we should only be surprised if we could find no evidence for the sort of cognitive theories I have described. Even so, if the distance between Jeff's theories and those of modern science is considerable, the same can be said of the distance between his theories and those of music theory as a discipline. In the next section, I shrink this distance just a bit by showing how the same cognitive structure evident in Jeff's theories can be seen in the mature work of two authors who made music theory their life's work: Jean-Philippe Rameau and Heinrich Schenker.

## CONCEPTUAL MODELS, THEORIES, AND MUSIC THEORY

During the eighteenth century, Jean-Philippe Rameau (1683–1764) was widely recognized as the foremost opera composer in France. Rameau had intellectual aspirations, as well as musical ones, however, and these he first brought to fruition in his *Traité d'harmonie* of 1722. The *Traité* offered two ideas that changed thinking about musical organization throughout Europe. First, it suggested that the rich but complex lexicon of chordal sonorities used in early-eighteenth-century music could be simplified and rationalized by granting certain arrangements of notes priority (the theory of chord inversion). Second, it proposed that musical organization was governed by privileged relationships between specific harmonic entities (the theory of harmonic progression). In part because of the novelty of these ideas, in part because of Rameau's famously treacherous prose, these theories were subject to a variety of reinterpretations and misunderstandings by those who adopted them, and Rameau himself continued to revise and refine his theories through no fewer

40. For further work on the role of theories in reasoning by children, see Frank C. Keil, *Concepts, Kinds, and Cognitive Development* (Cambridge, Mass.: MIT Press, 1989); and Frank C. Keil, Daniel T. Levin, Bethany A. Richman, and Grant Gutheil, "Mechanism and Explanation in the Development of Biological Thought: The Case of Disease," in *Folkbiology*, ed. Douglas L. Medin and Scott Atran (Cambridge, Mass.: MIT Press, 1999), 285–319.

than five additional treatises and a host of other writings until almost the last years of his life.[41]

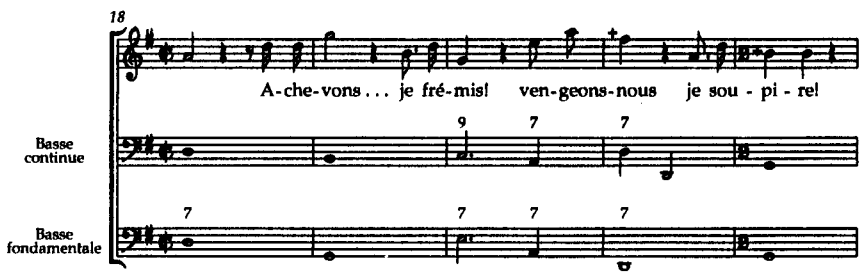
Heinrich Schenker (1867–1935) was an Austrian pianist, critic, and music theorist who supported himself primarily through giving private lessons. Schenker's writings include the journal-like series *Tonwille* and *Meisterwerk*, analyses of Beethoven's piano sonatas and Ninth Symphony, and three main treatises: *Harmonielehre* (1906), *Kontrapunkt* (1910–1922), and *Der freie Satz* (1935). Through these publications, Schenker developed two ideas that had a profound influence on Anglo-American music theory during the latter half of the twentieth century. First, he maintained that highly complex musical compositions were the result of elaborations of more basic and abstract structures. Second (and related), he held that musical works consisted of a number of hierarchical levels, proceeding from the most basic and abstract to the most complex and concrete, and that each successive level involved a further elaboration of a preceding level.[42]

The theories of music developed by Rameau and Schenker are complex, highly formalized systems for describing and analyzing music. In what follows, I consider each from the perspective on theories developed in this chapter, using short musical analyses by Rameau and Schenker to provide glimpses into their respective theories. This will allow me to show some of the cognitive substructure of their theories, and it will also reveal some important features of extended theories of music.

## Rameau on a Recitative by Lully

As do most of the music treatises of the early eighteenth century, Rameau's contain few if any analyses of actual works of music. An exception is found in his *Nouveau système* of 1726, where he provides an analysis of an entire recitative from Jean-Baptiste Lully's opera *Armide* of 1686, written on a libretto by Philippe Quinault.[43] An excerpt from the analysis is given in example 3.2. The top two staves are as they would have appeared in Lully's score: a vocal line with text, below which is the *basse*

EXAMPLE 3.2 Rameau's analysis (1726) of mm. 18–22 of Armide's recitative from act 2, scene 5 of Lully's *Armide*




41. For a more complete and more elegant account of Rameau's music theory, see Christensen, *Rameau and Musical Thought in the Enlightenment*.

42. The notion of hierarchy employed by Schenker is somewhat complex and is discussed further in chap. 7.

43. Jean-Philippe Rameau, *Nouveau système de musique théorique et pratique* (Paris: Ballard, 1726), 80–90.

*continue* part, with Arabic-numeral figures to indicate how the harmonization should be completed. Added beneath these is a staff for Rameau's *basse fondamentale*, which encapsulated his analysis of Lully's music.

Rameau's basse fondamentale had its origins in the precept that each chord had one fundamental sound (in modern terms, its root), from which the remaining notes were derived. In his analyses and descriptions of various successions of chords, Rameau would notate each of these fundamental notes on an additional staff, thus giving an account of what he considered to be the basis of any given harmony. The basse fondamentale was more than just a record of chord roots, however: it also allowed Rameau to give an account of musical syntax. The way Rameau accomplished this is more than a little complicated, but the essential points are as follows.[44] To explain why chord Y followed chord X rather than chord Z, Rameau exploited the notion of cadence prevalent in eighteenth-century music theory: chord Y followed chord X because the two were part of a normative cadence formula. As part of this formula, chord X would invariably be a major chord to which was added a minor seventh (in modern terminology, a dominant-seventh chord). Rameau came to regard the dissonance introduced by this seventh as the mechanism that led to cadence: the dissonant notes strike together, and the collision leads to the appearance of the chord that completes the cadence—what Rameau called the *tonique*. The chord that preceded the tonique, with its fundamental bass a fifth above that of the tonique, was called the *dominante-tonique*.[45] Of course, music of even the simplest sort involves more than tonic chords preceded by their dominants, but the basic pattern gave Rameau a place to begin his explanation of musical syntax. This he did by using the basse fondamentale to show the cadential relationships he took to be the motive force in music, even if this meant that the note indicated in the basse fondamentale could be related through only the most circuitous means to the chord that actually appeared in the music.[46]

Returning to the passage from Lully's recitative, there are two places where the note of the basse fondamentale is different from that of the basse continue. In m. 19, Rameau indicates that the fundamental of the chord B-D-G is G, an interpretation wholly in keeping with modern thought about chordal inversion. In m. 20, however, Rameau takes E as the fundamental of the C-E-G chord and implies that a D should be added to the chord; thus the 7 above the E of the basse fondamentale. This interpretation reflects his belief that the chord with A as its fundamental should be preceded by its dominante-tonique, which has E as its fundamental and which should be accompanied by the dissonant seventh; thus the E and D of m. 20. As an additional complication, Rameau added the figure 9 to the basse continue part of the version of Lully's score that appeared in the *Nouveau système*. In this way he was able to accept Lully's bass-note C and accommodate his own added D by

44. My discussion draws on that of Christensen in *Rameau and Musical Thought in the Enlightenment*, 106–23.

45. As mentioned in n. 25, Rameau employs a cross-domain mapping between the physical and musical domains to characterize musical progression. Consequent to this mapping, individual pitches are construed as physical objects that obey the laws of Newtonian physics.

46. For more on the function of the basse fondamentale as an extra-musical commentary on music, see Allan Keiler, "Music as Metalanguage: Rameau's Fundamental Bass," in *Music Theory: Special Topics*, ed. Richmond Browne (New York: Academic Press, 1981), 83–100.

suggesting that the latter should appear as a dissonant ninth above the bass. The fact that the E of the basse fondamentale appears *above* the C of the basse continue is not a problem for Rameau, since the C is explained by his notion of *supposition*—it functions as an ornamental note below the true bass note E.[47] From the perspective of modern music theory, Rameau's interpretation of m. 20 is viewed as completely wrongheaded, whereas his interpretation of m. 19 is viewed as essentially right. It is important to note, however, that *both* interpretations follow from Rameau's theory of musical organization and its account of musical syntax.

Rameau's treatment of Lully's recitative in the *Nouveau système* was not to be his last. In 1753, Jean-Jacques Rousseau offered a competing analysis of Lully's recitative, taking Rameau's discussion as a point of departure.[48] Rousseau's intent was to show that French musicians—even the most celebrated—had no idea of how to set text (in contrast to Italian musicians who, in Rousseau's view, did). Singling out the passage cited in example 3.2, Rousseau writes

> Certainly here is the most violent moment in the whole scene. It is here that the greatest struggle is taking place in Armide's heart. Who would believe that the Musician has left all this agitation in the same key, without the slightest intellectual transition, without the slightest harmonic distinction, in a manner so insipid, with a melody so little distinguished and so inconceivably clumsy, that instead of the last verse spoken by the Poet,
>
> *Achevons; je frémis. Vengeons-nous; je soupire*
> [End it; I tremble. Avenge myself; I sigh]
>
> the Musician says precisely this:
>
> *Achevons, achevons. Vengeons-nous, vengeons-nous.*
> [End it; end it. Avenge myself; avenge myself.][49]

Rameau responded to Rousseau's polemic with the 1754 *Observations sur notre instinct pour la musique, et sur son principe*, which set out to prove two points: that harmony is the most important aspect of music (prior to both the melody or the text that is sung) and that French music excels in the way it uses harmony.

The idea that harmony was the most important aspect of music was not new for Rameau, for he had argued the point in his 1722 *Traité*. In the *Observations*, however, the argument takes a slightly different form, based on the premise that our understanding of music begins at an instinctual level that reflects a natural principle. This principle exists, Rameau argues, in the harmony that results from the reso-

47. For further discussion, see Christensen, *Rameau and Musical Thought in the Enlightenment*, 120–29; and E. Cynthia Verba, "The Development of Rameau's Thoughts on Modulation and Chromatics," *Journal of the American Musicological Society* 26 (1973): 69–91.

48. Jean-Jacques Rousseau, "Letter on French Music," in *Essay on the Origin of Languages and Writings Related to Music*, trans. and ed. John T. Scott, Collected Writings of Rousseau, 7 (Hanover, N.H.: University Press of New England, 1998), 168–73.

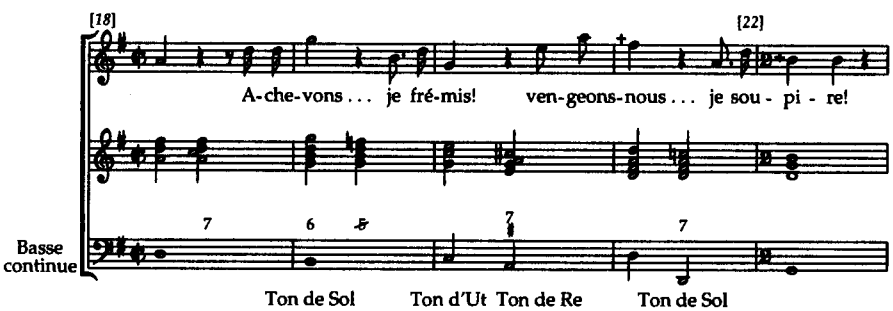
49. Rousseau, "Letter on French Music," 171–72. At this "most violent moment," the princess and sorceress, Armide, is about to take her revenge on the hero Renaud, who, enchanted, sleeps before her. However, confusion freezes Armide's vengeful hand as love casts a spell more powerful than that which grips the sleeping hero. At the point singled out by Rousseau, Armide swings back and forth between a steely determination to kill her foe and a desire to melt into his arms.

nance of every vibrating system, or *corps sonore*, including musical instruments and the human voice.[50] The essential elements of this harmony—the principle it reveals—are the intervals of the octave, twelfth, double octave, and seventeenth, or, in a more compact representation, root, third, and fifth.

According to Rameau, these intervals are generated not only above the root (as overtones) but below the root as well.[51] These "undertones" gave Rameau a way to explain the origins of the minor mode. In the *Observations*, he also used the intervals below the root to explain the affectual qualities of the subdominant in contrast with those of the dominant. The dominant is a product of the overtones; in consequence, it arouses feelings of vigor and joy. The subdominant—literally, the "underdominant"—is a product of the undertones; thus, it gives rise to feelings of weakness, gentleness, tenderness, and sadness.[52]

Having established this perspective on harmony and harmonic function, Rameau argued that Quinault's text was served perfectly by Lully's music, most particularly so in the case of the passage quoted in example 3.2. Rameau's argument, illustrated by the example given in example 3.3, runs as follows: shortly before this point, the music has been in E minor. With the D dominante-tonique of m. 18 (which sets the angry, ascending "Achevons"), the music establishes G major (Ton de Sol in ex. 3.3). At "je frémis," this G is converted into the dominante-tonique of C. This represents a move to the subdominant of G (Ton d'Ut), which is accompanied by a melodic descent. With "vengeons-nous," the dominante-tonique of D is introduced. This represents a move to the dominant of G (Ton de Re), which is accompanied by a melodic ascent. Finally, with "je soupire," D is converted into the dominante-tonique of G. Although this signals a return to G (Ton de Sol), it is also a move to the subdominant of the immediately preceding key of D, and it is accompanied by a final melodic descent. According to Rameau, Lully relies on our

EXAMPLE 3.3 Rameau's analysis (1754) of mm. 18–22 of Armide's recitative from act 2, scene 5 of Lully's *Armide*




50. Jean-Philippe Rameau, *Observations sur notre instinct pour la musique, et sur son principe* (Paris: Prault fils, 1754; facs. ed., Monuments of Music and Music Literature in Facsimile. Second Series, Music Literature, 54 [New York: Broude Brothers, 1967]), 2.

51. The claim is first made in Rameau's *Génération harmonique ou traité de musique théorique et pratique* (Paris: Prault fils, 1737; facs. ed. Monuments of Music and Music Literature in Facsimile. Second Series, Music Literature, 6 [New York: Broude Brothers, 1966]), 4–5.

52. Rameau, *Observations*, xii–xiii.

instinctual responses to these oscillations between the dominant and the subdominant to summon Armide's swings between revenge and love: each move toward the dominant arouses in us feelings appropriate to her revenge; each move toward the subdominant arouses feelings appropriate to the dissolution of her resolve. Rameau's analysis of this passage serves as the capstone to his argument in the *Observations*: the excellence of French music, the appreciation of which had eluded Rousseau, is embodied in the way it exploits our instinct for music in the service of musical expression.

Rameau's revised analysis necessitated some important changes in his basse fondamentale for this passage. In the *Observations*, he reads the second half of each of mm. 18–21 as including a dissonant seventh (which is in inversion in m. 19), reflecting the dominante-tonique chords that establish the keys of G, C, D, and finally G. Perhaps more important, the chord in the first half of m. 20 no longer serves as the dominante-tonique of the chord in the second half of the measure, for this would undercut Rameau's assertion that C major is in effect at the beginning of the measure: the note in the basse fondamentale is now C instead of E, and there is no 7 above it. As Rameau's theory changed, so did his analysis.

Rameau's intent with this analysis, as with the theory of music on which it is based, is to explicate the principles of musical organization and musical syntax. He does this by focusing on harmonies and harmonic relationships, reading out the larger implications of the basse continue—that is, not only the notes that will complete the musical texture but also the musical sense made by the bass line itself—and making them explicit through the basse fondamentale. The project is a large and complex one, and Rameau's analysis of Lully's recitative lets us glimpse only a portion of all that he would try to say about musical organization. Nonetheless, this vignette is enough to show how Rameau's music theory is a theory in the more general sense developed in this chapter, for all of the five features of theories are clearly in evidence:

THEORIES GUIDE UNDERSTANDING AND REASONING Rameau's theory gives an account of harmonic progression—that is, it provides an explanation of why one chord follows another. Given a bit of context connected with a particular chord, we could make a reasoned judgment about what the following chord would most likely be. Through the notion of chordal inversion, his theory also explains relationships between chords that comprise the same pitch classes. In the analysis shown in example 3.3, the harmony in the first half of m. 19 is "the same" as that in m. 22, even though their bass notes are different. Finally, by locating the origin of music in the naturally occurring corps sonore, Rameau's theory offers a way to account for the emotions music evokes in humans. One set of emotions (broadly, the happy ones) is associated with notes generated above the corps sonore; another set (the sad ones) is associated with notes generated below the corps sonore.

THEORIES PROVIDE ANSWERS TO CONCEPTUAL PUZZLES One of the profound puzzles that Rameau sought to solve in the *Observations* is why music moves us. This he did by proposing that the corps sonore was not simply an acoustic phenomenon but a natural one, too: we are moved by music in conformance with

natural laws. A more parochial puzzle concerns successions of bass notes that do not conform to his notions of harmonic progression (such as the C3 to A3 in m. 20 of ex. 3.2, from Rameau's first analysis in the *Nouveau système*). Rameau suggests that the given bass note has, essentially, an ornamental function and the real bass note (indicated in the basse fondamentale) is the one that conforms to the proper progression of harmonies. Of course, the "puzzle" here could be seen as simply a manifestation of the limitations of Rameau's theory—it is one that disappears when the theory changes, as it has by the time of the *Observations*. Nonetheless, the issue raised by this puzzle/non-puzzle points up something important about theories and conceptual puzzles: theories, through the way they describe situations, tend to create puzzles, for which the theory then offers solutions. When the theory changes, these puzzles often disappear, to be replaced by new puzzles specific to the new theory.[53]

THEORIES SIMPLIFY REALITY As a composer of many successful operas, Rameau knew well that music consisted of more than just successions of dominant and tonic chords. Focusing on this succession to the exclusion of others, however, allowed Rameau to give an explanation of why one harmony followed another, and thus to give an explanation of harmonic progression. And Rameau's account of the way humans respond to music greatly simplifies the emotional response to music, reducing it to an essential binarism at a variance with the range of emotion evoked by effective musical works.

THEORIES INVOLVE A NUMBER OF CONCEPTUAL MODELS Rameau's complete theory coordinates a large number of conceptual models, ranging from models for pitch-events and intervallic relationships to models for ornament and expressivity. Three models stand out in both of his analyses of Lully's recitative: the conceptual model associated with the corps sonore, which uses basic acoustic properties to explain relationships among musical events; the conceptual model for musical mechanics, which maps Newtonian physics onto musical events to explain musical syntax; and the conceptual model for chordal inversion, which describes relationships between the notes that make up a harmony in terms of a set of dependencies created by chord generation. Although the notion of the corps sonore was ultimately absorbed into general thinking about the acoustic origins of music, the latter two models had a profound effect on the music theory of succeeding generations: the success of Rameau's theory was such that later theorists either incorporated these models into their own theories or found substitutes for them in alternative conceptions of chordal progression and inversion.

THEORIES ARE DYNAMIC As is evident from the two analyses of Lully's recitative, Rameau's theory of music changes over time: in the 1726 version, the

53. This situation is somewhat more common in the humanistic disciplines than in the hard sciences. In the latter, the puzzles are typically independent of humans—they obtain whether humans are observing the phenomena or not—whereas in the former the puzzles (and the disciplines, for that matter) would not exist without humans.

chord at the beginning of m. 20 acts as a dominante-tonique for the chord that follows; in the 1754 version (expanded to account for the drama evoked by the passage), the same chord represents a temporary, but affectually important, harmonic goal. It should be said that when Rameau changed his theories, it was not always for the better: even though these changes allowed him to account for a wider range of phenomena, they often introduced complications that made his theory more unwieldy.

Although Rameau's is a complex and highly elaborate theory, embedded in history and pertaining to a relatively limited musical repertoire, it nonetheless shares significant features with the cognitive theories that are the basic tools of human understanding. To fully comprehend Rameau's theory requires our reclaiming its historical context and bringing our own conceptual models into correspondence with those on which his theory is based. Certain models, such as that for chordal inversion, pose few problems. Others, such as the conceptual model for supposition, have little in common with our own models and require much greater effort to understand. Nonetheless, the distance between his way of understanding and ours, or between either of these and Jeff's, is only a matter of the concepts and conceptual models employed. The basic cognitive structures that make this understanding possible are the same.

## Schenker on a Waltz by Schubert

Although Heinrich Schenker adopted some of Rameau's theoretical precepts, his attitude toward them changed over the course of his writings.[54] In his *Harmonielehre* of 1906, Schenker speaks approvingly of Rameau's work, but after World War I, Schenker became highly critical of all things French, Rameau included. His late essay "Rameau or Beethoven?" is a sustained diatribe in which he portrays Rameau as at best ignorant, at worst a malevolent force who contributed to the downfall of music.[55] Rameau was not the only theorist Schenker criticized, however. Even before the war, Schenker had become convinced that German music was being pulled into mediocrity and that music theorists had aided and abetted the process.[56] A central and explicit goal of Schenker's writings on music was to retard or arrest this process by showing the basis for the superiority of a select repertoire of music associated with the heritage of German-speaking countries. Schenker believed that only when the work of composers from Johann Sebastian Bach to Johannes Brahms was once again fully appreciated could humanity be saved from its own excesses.

54. For a fuller account of how Schenker's attitude toward Rameau changed over the course of his writings, see Harald Krebs, "Schenker's Changing View of Rameau: A Comparison of Remarks in *Harmony*, *Counterpoint*, and 'Rameau or Beethoven?,'" *Theoria* 3 (1989): 59–72.

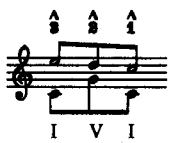
55. Heinrich Schenker, "Rameau or Beethoven? Creeping Paralysis or Spiritual Potency in Music?," in *The Masterwork in Music: A Yearbook*, trans. Ian Bent, ed. William Drabkin, Cambridge Studies in Music Theory and Analysis 5 (Cambridge: Cambridge University Press, 1994–97), 3: 1–9.

56. Heinrich Schenker, *Counterpoint: Book I: Cantus Firmus and Two-Voice Counterpoint*, ed. John Rothgeb, trans. John Rothgeb and Jürgen Thym, New Musical Theories and Fantasies, 2 (New York: Schirmer Books, 1987), xxvii–xxx.

Basic to Schenker's music theory were two ideas: first, that the forces that shape music are independent of humans; second, that these forces determine fundamental aspects of musical syntax.[57] These forces are made manifest in what Schenker called the *Naturklang*, or chord of nature, consisting of the overtones generated from a single fundamental, with priority granted to those closest to the fundamental.[58] Couching his argument in strongly organicist terms, Schenker proposed that these overtones were first witness to the desire of the tone to propagate itself.[59] This desire also gave rise to various *Ursätze*, or fundamental structures, each of which began with either the third, fifth, or octave of the fundamental. As shown in example 3.4, an Ursatz consists of a simple two-part counterpoint. The notes with upward stems make up the *Urlinie* (or fundamental line), which descends stepwise from the *Kopfton* (or head note—here, the third of the fundamental).[60] Below this, in one-to-one counterpoint, is the *Brechung* (or arpeggiation); in example 3.4, this consists of an alternation between the fundamental and fifth of the Naturklang. The Ursatz represents an idealized succession of musical materials, all of which are in clearly specified relationships with each other. It thus provides an account—abstract, minimal, but nonetheless complete—of musical syntax: that is, how musical materials are organized to produce the elements of musical discourse.

As Schenker would have it, the skilled composer, in intuitive response to the powerful urges of the Naturklang, elaborates an Ursatz further and further until, finally, a complete work is born. The possible elaborations of Ursätze are manifold

EXAMPLE 3.4 A Schenkerian Ursatz beginning on the third of the fundamental




57. These two aspects of Schenker's music theory are discussed in a bit more depth in chap. 7.

58. Although there are obvious affinities between Rameau's corps sonore and Schenker's Naturklang, the latter is more complex in both the overtones proper to the system and their application to musical structure. In general, the idea that there was an acoustical basis to musical organization was a commonplace of nineteenth-century musical thought, and there is little evidence that Schenker's thinking on this matter was directly linked to that of Rameau.

59. "It is self-evident that the urge to produce unending generations of overtones belongs to every tone in equal measure. One might also compare this urge to that of animals, for it appears in fact to be in no way inferior to the procreative urge of a living being." ["Es ist selbstverständlich, daß den Trieb, Generationen von Obertönen ins Unendliche zu zeugen, jeder Ton in gleichem Maße besitzt. Man darf, wenn man will, auch diesen Trieb einem animalischen vergleichen, denn er scheint in der Tat dem Fortpflanzungstrieb eines Lebewesens durchaus nicht nachzustehen."] Heinrich Schenker, *Harmonielehre*, Neue musikalische Theorien und Phantasien, 1 (Stuttgart: J. G. Cotta'sche, 1906), §13, p. 42.

60. Translating Schenker's terminology is a notoriously fraught issue. For a discussion, see Robert Snarrenberg, "Competing Myths: The American Abandonment of Schenker's Organicism," in *Theory, Analysis and Meaning in Music*, ed. Anthony Pople (Cambridge: Cambridge University Press, 1994), 29–56. In the following, I shall rely on Schenker's original terms (followed, at their introduction, by common translations) as presented in Schenker, *Der freie Satz*, Neue musikalische Theorien und Phantasien, 3 (Vienna: Universal-Edition, 1935).

and, for all intents and purposes, are coextensive with the repertoire Schenker sought to explain. Within these possibilities, two main types of elaboration can be distinguished. The first is a manifestation of the desire of individual components of the Ursatz to assert control over longer spans of music. This occurs through the process of *Verwandlung* (transformation or prolongation), realized through a variety of compositional stratagems.[61] The second type of elaboration replicates whole sections of the Ursatz on subsidiary levels of musical structure, thus presenting the Ursatz in a sort of structural diminution. In both cases, the Ursatz is a first level of musical syntax, whose elements are subsequently transformed or reiterated to create more extensive musical utterances. This notion of syntax is different from Rameau's in two important respects. First, rather than describing local, event-to-event relationships, Schenker's theory attempts to account for relationships over the course of an entire work. Second, where harmony was the starting point for Rameau, for Schenker, musical syntax starts with the contrapuntal voice-leading of the Ursatz.

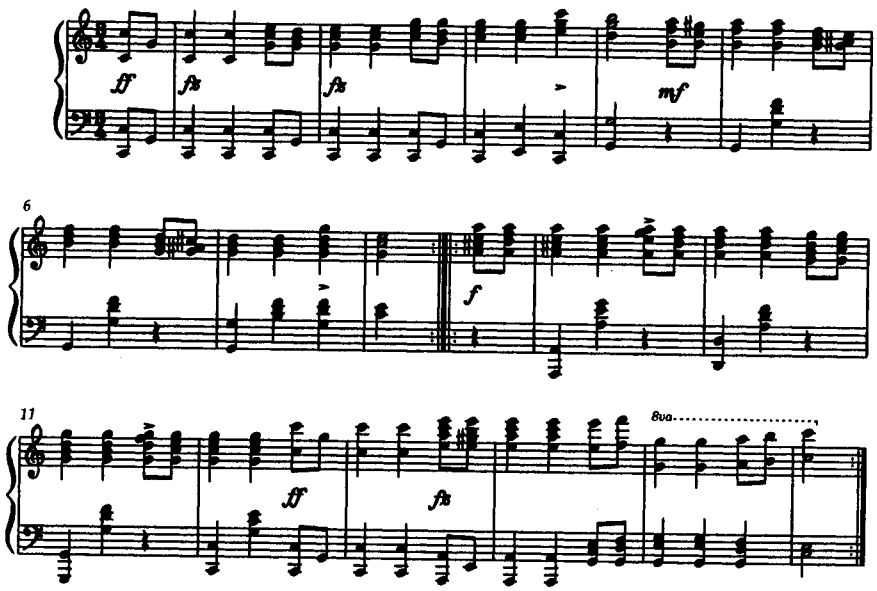
Of course, composers rarely if ever committed the successive elaborations of the Ursatz to paper. According to Schenker, creating a musical composition that both conformed to and properly expanded on an Ursatz was the work of genius. The best a nongenius could do was to come to an appreciation of the finished work by retracing the successive elaborations of the Ursatz through a series of hierarchical levels, which proceed from the background (the most minimal elaboration of the Ursatz) to the foreground (which often resembles a simplified version of the finished work).[62] The result is a striking image of musical structure that shows musical syntax operating on a number of levels and within a number of different time frames within each composition.

To represent the elaborations of an Ursatz and the hierarchical levels that resulted, Schenker developed a unique set of graphical symbols that he used in his analyses. These allowed him to show relationships among structural elements and also made manifest the replication of elements across the various levels of an elaborative hierarchy. As an example, let us turn to Schenker's analysis of the first of Franz Schubert's *Valses nobles*, Op. 77. The score for the work is given in example 3.5; Schenker's graphic analysis appears in example 3.6 (the Ursatz for the analysis is the same as that given in ex. 3.4). Although not all of the features of Schenker's graphic notation need concern us here, a few will be useful for orienting ourselves and for understanding his analysis. Open note heads in the graph (with the exception of the F5 neighbor note after the double bars) represent components of the Ursatz; generally speaking, slurs and beams mark off larger syntactic units. The graph omits the rhythmic figuration, filled-out chords, and much of the registral detail of Schubert's waltz, as well as various subsidiary passing and neighbor notes

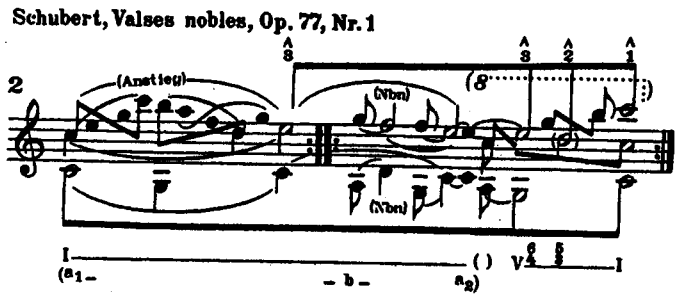
61. The transformation (or prolongation) of elements of the Ursatz is more than a little complicated, and Schenker used a number of terms in order to capture the variety of ways it is accomplished. See Schenker, *Der freie Satz*, §45, p. 49.

62. Actual musical analysis most often proceeds in the reverse direction, going from completed musical work, to foreground, to background.

EXAMPLE 3.5 Franz Schubert, *Valse noble*, Op. 77 No. 1, D. 969




EXAMPLE 3.6 Schenker's analysis of Schubert's Op. 77 No. 1




and brief arpeggiations. Nonetheless, it does not take too much imagination to see (and hear) the outlines of the opening of the waltz in the first portion of Schenker's graph.

After the double bars, the connection between Schubert's music and Schenker's graph is somewhat less immediate, and the graph becomes somewhat more complicated. Schenker's basic argument, however, is clear: mm. 9–11 serve as contrasting material that prolongs the Kopfton, and after m. 12, the music returns to a modified version of the opening material that culminates in the completion of the Ursatz. The completion of the Ursatz is accomplished in spite of two complications. First, scale degree 2 is not actually present in the music—Schenker shows this by placing the $\hat{2}$ of the Ursatz in parentheses. Second, the Ursatz is brought to com-

pletion in a higher octave than that established by the Kopfton in m. 8, which runs counter to Schenker's principle of *obligate Lage* (obligatory register).[63]

Schenker's theory, then, offers a view of Schubert's waltz as a sort of abstract discourse structure. The first eight measures (the $a_1$ section) serve as introduction, providing an *Anstieg* (initial ascent) from C5 through D5 to the Kopfton E5. The next four measures (the b section) offer a contrast, prolonging the Kopfton through a neighbor-note figure (indicated with "Nbn" in ex. 3.6). The final four measures (the $a_2$ section) return to the opening material, modified in order to complete the Urlinie and bring about the conclusion of the piece. According to Schenker, to the extent that we understand this discourse structure, we understand how the Ursatz is elaborated to create this unique composition. The force of this understanding is such that we "supply" the D6 of the Ursatz at the conclusion of the waltz, even though it is not literally present in the music.

In the hundreds of analyses that accompanied the exposition of his theory, Schenker showed evidence of similar discourse structures in the works of all of the principal composers from Germany and Austria during the eighteenth and nineteenth centuries, and a few others besides. The theory of music that emerges, not unlike the works whose structure it seeks to explicate, is fascinating in its extent and complexity. Nonetheless, it is a theory with all of the features of a cognitive theory outlined in this chapter.

THEORIES GUIDE UNDERSTANDING AND REASONING Schenker's theory provides a broad account of musical syntax by explaining how musical materials relate to one another in conformance with the rules of counterpoint and harmonic progression. His intent is to describe the entirety of musical works. Given this objective, and the organicist imperative relative to which the theory is framed, Schenkerian analysis provides an explanation for the function and meaning of even the smallest part of a musical composition: one can proceed, in an orderly fashion, from the graph of example 3.6 back to the music of example 3.5. From a larger perspective, Schenker's theory gives an account of form, understood as the various roles that different sections of music play within an overall discourse structure. The indication of a return to the opening material of Schubert's waltz in m. 12 (indicated by Schenker's $a_2$ below the staff) marks not only a reprise of the opening figure but also a return to the syntactic process initiated by the opening of the waltz. Finally, Schenker also offers an aesthetic theory: he believed that the works most worthy of his theoretical accounts were also the most valuable aesthetically.

THEORIES PROVIDE ANSWERS TO CONCEPTUAL PUZZLES One of the modest puzzles presented by Schubert's waltz concerns the attention we should pay to it. It is one of dozens if not hundreds of waltzes Schubert tossed off to gen-

63. The principle of obligate Lage is a reflection of the urge of the notes of the Ursatz to return to their source, which in all cases should involve a descent from the *Kopfton* to the octave above the fundamental (as in ex. 3.4). However, Schenker himself did not seem to regard obligate Lage as obligatory in all cases—he notes the Schubert waltz discussed here as an instance in which the principle is violated. See Schenker, *Der freie Satz*, §268, p. 175.

erate income. Does it bear the mark of his compositional genius, or is it simply another bit of doggerel churned out to fit the needs of the moment? Schenker's analysis decidedly affirms the former: whether it was written in a few minutes or over the course of several days, the waltz conforms to and properly expands on one of the Ursätze basic to music. As with Rameau's account of Lully's recitative, we also see the theory providing answers to puzzles specific to the theory. In the analysis of the Schubert waltz, these are the complications presented by the apparent departures from the Urlinie in the concluding measures.

THEORIES SIMPLIFY REALITY One of the superficial criticisms made of Schenker's theory is that his graphic analyses leave out all of the surface detail that makes a musical work distinctive and rewarding, and that register, dynamics, and rhythmic figuration are often obscured or ignored. Indeed, the graph given in example 3.6 bears this out. This is not a deficiency, however, but simply something that marks Schenker's theory as a theory: one that simplifies reality as do all theories.[64] The process of theorizing, then, is one of selecting data from a given situation and using them to build a more comprehensive understanding of that situation.

THEORIES INVOLVE A NUMBER OF CONCEPTUAL MODELS Schenker's propensity for neologisms, his attempts to vivify the realm of tones through comparisons to human and animal domains, and his adaptations of the theoretical work of earlier theorists give rise to an astounding number of conceptual models. Two in particular give his analysis of the Schubert waltz a distinctive flavor: the conceptual model associated with the Naturklang, which supplies the organicist basis for his account of musical syntax; and a conceptual model of hierarchy (also organicist, and discussed in further detail in chap. 7) that correlates numerous layers of musical syntax as a way of explaining how basic syntactic structures account for the syntax of an entire work.

THEORIES ARE DYNAMIC In as brief a summary as this, it is not possible to show in any compelling way the dynamic aspect of Schenker's theory over time. Indeed, about the best that can be done is to note Schenker's flexibility with regard to obligate Lage, which suggests that his approach to musical structure is not as rigid as he would sometimes have us believe. Nonetheless, there is a wealth of research that shows that Schenker's theory was a dynamic one, changing in response to both the repertoire and problems he engaged.

Many claims have been made for Schenker's theory of music, a theory that, in its ambition and extent, has few competitors in the Western tradition. No one to my knowledge, however, has claimed that it is cut from the same cloth as the music theory of Rameau or that of an eight-year-old boy from the Boston area. Indeed, I would not want to make such a claim, or at least not without considerable qualifi-

64. Ronald Langacker makes a similar point about the analyses produced by theories of language, noting that "there is certainly some truth in the view that analysis and description inevitably distort subject matter since they cannot be their subject matter." Langacker, *Foundations of Cognitive Grammar*, 1: 14.

cation. In the things that matter most to music theorists, Schenker's theory is simply of another order. Nonetheless, in its general features—in the cognitive work that it does, and in the way that it does it—it shares much not only with Rameau's theory of music and with Jeff's theories of "Twinkle" but also with cognitive theories as a whole.

## Features of Extended Theories

The theories of Rameau and Schenker, unlike those of Jeff, are sophisticated intellectual constructs that can account for a wide range of complex musical phenomena. Such theories, which I shall call *extended theories* (in order to differentiate them from more ad hoc theories like Jeff's), have three prominent features. First, extended theories usually engage a problem or problems of long standing, like those of musical expression or musical coherence. Ancillary to this, extended theories often reveal —or create—problems through the constructs they apply to music.

Second, extended theories, as the product of traditions of discourse that often stretch across generations, are utterly embedded in culture. This was quite apparent in the nationalism that informed—indeed, fueled—both Rameau's and Schenker's work. Although it might be easy to dismiss such cultural trappings as extraneous to the business of music theory, judging from the writings of Rameau and Schenker, it is hardly so: they see music as thoroughly implicated in the business of history and nation.

Third, theories like those of Rameau and Schenker are invested in developing systematic accounts of musical organization. This systematic quality can take the form of developing an exhaustive account of musical structure. In some measure, this is what Schenker aspired to, as did Allen Forte in his efforts to provide a theory of pitch organization for atonal music.[65] The systematic aspect of a theory, however, may be somewhat more circumscribed, especially when its focus is musical syntax. Here the goal is not so much to explain everything about musical organization but, rather, what is *s* in musical organization. This is closer to the goals of Rameau's theory, and it tempered both the repertoire Schenker chose to deal with and the problems he engaged.

The topic of musical syntax, especially as a manifestation of the systematic quality of extended theories, brings us back to the paradox that theories simultaneously treat and create problems. The very notion of musical syntax—the idea that musical discourse is in some way organized—is in part a creation of theories that seek to explain musical organization. Depending on the models music theory brings to bear on the problem, musical syntax can be seen to be quite different things: a matter of how one harmony leads to another; an account of how basic contrapuntal and harmonic frameworks are elaborated into full-fledged musical works; or a different explanation altogether. As shall be seen in the next chapter, my own approach to musical syntax is different yet, involved as it is with the way musical materials are arranged to take advantage of and to exploit our ability to categorize musical events.

65. Allen Forte, *The Structure of Atonal Music* (New Haven, Conn.: Yale University Press, 1973).

## CONCLUSIONS

Theories, as general cognitive constructs manifested in a wide range of human activity, accomplish their work by correlating a number of conceptual models, each of which sets up a conceptual domain. In some cases, these conceptual domains will be quite similar to one another, but in others, the domains may have very little in common. This means that the topography of any theory, reckoned in terms of the conceptual domains it comprises, will not necessarily be uniform. A theory may thus appear to make contradictory claims when in fact it has simply shifted focus from one conceptual domain to a second, and incommensurate, conceptual domain. Thinking about a theory in terms of its constituent conceptual models can thus be a tool for coming to understand it better, especially where these models are novel or lead the theory in surprising directions.

Theories are also part of culture. The culture specific to music theory, it must be admitted, is a quite rarefied one. Nonetheless, the broader influence of culture can be seen here, too, as was demonstrated by the nationalistic claims made by both Rameau and Schenker. These claims are not extraneous to Rameau's and Schenker's theories but are in fact integral to them—manifestations of the broader culture of which these theories are a part and relative to which they are framed. This is not to deny the uneasiness we may feel when reading the strident claims of the music theory of another era but only to say that they are not so very different from the notions about pitch relationships and chordal inversion common within the same cultural milieu.

Finally, theories of music can tell us something about cognitive structure as a whole. Because the focal domain for music theory is a nonlinguistic one, theories of music demonstrate how we have and correlate nonlinguistic concepts. The abstraction of Rameau's and Schenker's theories notwithstanding, theories of music are often action oriented. These actions can be of a quite visible kind, as demonstrated by Jeff's attempts to set up different bell-paths for "Twinkle," but they can also involve actions less apparent to the eye: theories often suggest ways to attend to or respond to music. Given their involvement with nonlinguistic entities and their concern with action, theories of music can be seen to be deeply involved with embodied understanding in ways that few other highly developed cognitive theories are.

While the preceding chapters have sketched how work in cognitive science can be applied to problems of musical understanding—or, put another way, how our accounts of musical understanding can be brought into line with what is currently known about the mind and brain—much work remains to be done to fill out this picture. The chapters that follow make a start by showing how the framework developed thus far can be used to give accounts of musical syntax, cultural knowledge, musical ontology, text-music relations, and the structure of theories. As I noted in the introduction to this volume, I regard every analysis as based on some theory or collection of theories. The theories I highlight in part II, however, are theories about how basic cognitive processes are manifested in and exploited by musical understanding. In consequence, the analyses that follow are of a somewhat different character than "musical analysis" as it is usually practiced, for my intent is

not only to shed light on specific issues of concern to music scholars but also to refine a theory of music cognition. The examples around which I build my argument are not simply case studies, but are also tools that will help us probe further into the issues discussed thus far.[66]

As should be clear, the view of cognitive structure, theory, and analysis that I have presented thus far is of necessity preliminary, and it deals with only a small part of the potential issues raised even when the domain of inquiry is restricted to the conceptual level of cognition. Nonetheless, as the following chapters will make clear, even this partial view offers striking new perspectives on musical understanding and on what it means to theorize about music.

66. This approach to the role examples play in explorations of the structure of thought is similar to that noted by Mark Turner in his discussion of cognition and rhetoric. See in particular Turner, "Figure," in *Figurative Language and Thought*, ed. Albert N. Katz, Cristina Cacciari, Raymond W. Gibbs, Jr., and Mark Turner (New York: Oxford University Press, 1998), 80.

PART II

# ANALYSIS AND THEORY

CHAPTER FOUR

# CATEGORIZATION, COMPOSITIONAL STRATEGY, AND MUSICAL SYNTAX

Early in his *Theory of Semiotics*, Umberto Eco surveyed the communicative processes and cultural systems that might be included in the field of semiotics, moving through a list worthy of Jorge Luis Borges: zoosemiotics, olfactory signs, tactile communication, codes of taste, paralinguistics, medical semiotics, kinesics and proxemics, formalized languages, written languages, unknown alphabets, secret codes, natural languages, visual communication, systems of objects, plot structure, text theory, cultural codes, aesthetic texts, mass communication, and rhetoric. Midway through this list he considered what he called "musical codes." He began by proposing that the whole of musical science since the Pythagoreans has been an effort to describe musical communication as a rigorously structured system.[1] However, Eco also observed that, in spite of this tradition of structuralism from which musical semiotics can draw, music poses serious difficulties for semiotic theory: "Music presents, on the one hand, the problem of a semiotic system without a semantic level (or a content plane): on the other hand, however, there are musical 'signs' (or syntagms) with an explicit denotative value (trumpet signals in the army) and there are syntagms or entire 'texts' possessing pre-culturalized connotative value ('pastoral' or 'thrilling' music, etc.)."[2]

Eco developed a somewhat fuller account of what he meant when he characterized music as a semiotic system without a semantic level a bit further on, just prior to setting out the combinatorial rules for semiotic codes. The problem was not that musical syntax cannot be interpreted—indeed, Eco argued that any syntactic system can be interpreted on some level—but that the interpretation of music is resolutely indeterminate. There is thus no depth to the semantic levels produced by musical syntax.[3]

1. Umberto Eco, *A Theory of Semiotics*, Advances in Semiotics (Bloomington: Indiana University Press, 1976), 10. The idea that musical communication is a rigorously structured system is, of course, another way to theorize about music. Interest in such an approach has waxed and waned along with musicians' interest in numerology, mathematics, and science as possible models for musical organization.

2. Eco, *A Theory of Semiotics*, 11.

3. Eco, *A Theory of Semiotics*, 88–90. In advancing this point, Eco engages in a dialogue with Louis Hjelmslev's distinction between semiotic and non-semiotic systems; see Louis Hjelmslev, *Prolegomena to*

I find Eco's comments interesting, not for the view of musical semiotics they provide (since the passages to which I have referred are pretty much all he has to say about music in *A Theory of Semiotics*), but for their clear focus on the importance of musical syntax for musical understanding. Musical syntax is where any semiotics of music must begin. In what follows, I build on this idea, using the perspective on musical syntax introduced in chapter 1. There, I proposed thinking of syntax as a connected or orderly system. In the case of music, organized sequences of musical events properly recognized as syntactic are typically created through a set of musical practices shared among a number of musicians. Within most traditions of music making, however, there is also some latitude in how musical materials can be arranged, and the syntax specific to a particular musical work may emerge only over time. This was the case with the Leidensmotiv from Wagner's *Tristan und Isolde*, for the importance of three successive statements of the motive was something that became evident only once we were well into the opera. In situations such as this, certain aspects of the systematic quality associated with musical syntax are determined by the expressive goals of the individual ultimately responsible for arranging the musical materials. My term for the way such expressive goals are realized is "compositional strategy."[4]

Compositional strategy, conceived in these terms, assumes an alliance between syntax and processes of meaning construction: composers arrange musical materials with relatively specific expressive or communicative goals in mind. This approach takes inspiration from recent work in cognitive linguistics, which does away with a hard-and-fast distinction between syntax and semantics and instead construes language as made up of symbolic structures that consist of form-meaning pairs.[5] A similar perspective can be seen in Mark Turner's work on the iconicity of rhetorical figures. Turner argues that, in some cases, the form of a rhetorical figure is matched to the meaning the speaker wishes to convey, connecting the image-schematic structure of the form with the image-schematic structure of the meaning. Thus a rhetorical figure based on repetition, such as anaphora (which involves the repetition of the same word or group of words at the beginning of successive clauses, sentences, or lines), can be used to summon the image of repeated blows, as in an example attributed to Longinus: "By his manner, his looks, his voice, when

---

*a Theory of Language*, rev. ed., trans. Francis J. Whitfield (Madison: University of Wisconsin Press, 1963), 111–13. Hjelmslev, for his part, leaves open the question of whether music is a semiotic system in the strict sense he develops.

4. It should be clear from the general way I have formulated the notion of "compositional strategy" that these strategies are not the exclusive property of those traditionally recognized as "composers." In certain situations, performers and even listeners may use such strategies to organize musical materials. Throughout this chapter, I shall simplify things by assuming that composers are the architects of compositional strategies, with the understanding that (mutatis mutandis) these strategies can also be implemented by those not conventionally recognized as "composers."

5. Ronald W. Langacker, "Conceptualization, Symbolization, and Grammar," in *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, ed. Michael Tomasello (Mahwah, N.J.: Erlbaum Associates, 1998), 2. Langacker's notion of a symbol is spelled out in greater detail in his *Theoretical Prerequisites* (vol. 1 of *Foundations of Cognitive Grammar*) (Stanford, Calif.: Stanford University Press, 1987), 11–12, 56–58.

he strikes you with insult, when he strikes you like an enemy, when he strikes you with his knuckles, when he strikes you like a slave." The efficacy of such a connection is straightforward enough—as Turner notes, "Involving members of the audience in the image schema of the iconic form automatically involves them in the basic structure of the meaning, thus moving them part way toward accepting the whole."[6] In what follows, I show how compositional strategies in music make use of similar connections between expressive goals and the forms expressions take. Music being music, the meaning that is conveyed is not as direct as that which concerned Longinus, nor as quotidian as that which concerns cognitive linguists. Nonetheless, I argue for a direct connection between compositional strategy and musical syntax: the organization of musical elements into systematic structures is part and parcel of meaning construction in music.[7]

The particular focus of what follows is on the part processes of categorization play in compositional strategies. As discussed in chapter 1, categories of musical events have immediate salience for listeners and are thus important guides to any understanding of how musical materials are organized. Building on the perspective developed in chapter 3, we can say that they are basic to any theory of music. Inasmuch as composers try to shape their music to conform to the capacities of listeners (trying, in Schoenberg's terms, to create musical coherence so that the work can be comprehensible), processes of categorization will also play an important part in compositional strategies. Musical categories are thus a place where the concerns of listeners and the concerns of composers meet.

In the first section of this chapter, I explore the role categories of musical events play in the compositional strategies basic to a musical game attributed to Wolfgang Amadeus Mozart. The game in question is one of a genre, which seems to represent the very antithesis of compositional strategy, since such games produce finished—or nearly finished—musical works simply through tosses of the dice. Embedded within their structure, however, is a set of rather traditional strategies for writing two-part

6. Mark Turner, "Figure," in *Figurative Language and Thought*, ed. Albert N. Katz, Cristina Cacciari, Raymond W. Gibbs, Jr., and Mark Turner (New York: Oxford University Press, 1998), 50–51. Turner's discussion of the iconicity of rhetorical form appears on pp. 49–51. The quotation attributed to Longinus is from Aristotle, *On the Sublime*, trans. W. Hamilton Fyfe, Loeb Classical Library, 199 (Cambridge, Mass.: Harvard University Press, 1995), §20, p. 190. Both a poetic and musical use of anaphora can be found in Bob Dylan's "Knockin' on Heaven's Door," which first appeared in his 1973 album *Pat Garrett and Billy the Kid*. The chorus of the song—consisting of reiterations of the line "Knock, knock, knockin' on heaven's door"—makes use of a repeated note in the melody, as well as a repeated word. Both repetitions correlate with the meaning of the text. Note as well that the meaning is at least potentially figurative, rather than literal, since nowhere in the song is it assumed that heaven actually has a physical door on which one can knock.

7. As I have characterized it here, compositional strategy bears a strong resemblance to the tactical logistics of rhetoric, and it might thus be seen as another instance of the use of ideas drawn from rhetoric to conceptualize musical design (something discussed in greater detail in the first part of chap. 7). As will become apparent, I construe the idea of compositional strategy as rather more general than its rhetorical counterpart, and more involved with issues of immediate importance to musical expression (such as how the temporal succession of musical materials should be arranged). This construal of compositional strategy seems largely consonant with the approach Edward Cone took in his article "On Derivation: Syntax and Rhetoric," *Music Analysis* 6 (1987): 237–55, which was discussed in chap. 1.

musical forms, realized through the manipulation of categories of musical events. During the eighteenth century, these same strategies were applied, on a much larger scale, to sonata form, guiding the way harmonic and thematic materials were laid out. They could also be applied to melodic material alone, however. In the second section of this chapter, I show how this was accomplished through an analysis of the syntax of the principal motive from the first movement of Mozart's "Dissonance" Quartet (K. 465). Taking the different versions of the motive as members of a Type 1 category, the pattern of typicality described by these members as they appear over the course of the movement shows strong similarities to the basic strategies for sonata form. The syntax for these motive forms nonetheless occupies a different layer than the syntax that guides the formal organization of the movement as a whole.

As with the grouping of statements of Wagner's Leidensmotiv, the syntax of the principal motive from the first movement of the "Dissonance" Quartet emerges over the course of the movement. Both the emergence of musical syntax and the independence of syntactic levels is further demonstrated by the first movement of Ludwig van Beethoven's String Quartet, Op. 18 No. 6, which is taken up in the third section of the chapter. Although there is little variation of the principal motivic material for the movement, Beethoven is nonetheless able to imbue two versions of the motive with distinctly different syntactic functions. The categories of motive forms that comprise these versions then become elements at a higher level of syntax. The nesting of syntactic functions can also be seen, in a slightly more complicated way, in the first movement of Beethoven's String Quartet, Op. 18 No. 1, which is discussed in the fourth section of the chapter. There, different versions of the main motive are distinguished by melodic structure and by compositional strategy, and the categories for each become syntactic elements within the movement as a whole. In a concluding section I review the perspective on musical syntax developed over the course of the chapter and give a bit further consideration to how syntax contributes to the construction of meaning in music.

## MUSICAL GAMES AND COMPOSITIONAL STRATEGY

### A Musical Game Attributed to Mozart

Europe in the latter half of the eighteenth century saw the blossoming of a number of musical games that allowed any person, "without the least knowledge of Musick," to compose minuets, marches, waltzes, and contredanses simply by tossing dice or spinning a top.[8] As if by magic, a series of throws or spins produced a workable piece of music that was different every time the game was played.

The game described here—the *Musikalisches Würfelspiel*—allowed its players to

8. One example is Pierre Hoegi's *A Tabular System Whereby the Art of Composing Minuets Is Made So Easy That Any Person, without the Least Knowledge of Musick, May Compose Ten Thousand, All Different, and in the Most Pleasing and Correct Manner* (London: Welcker, 1770). For further information about such games and comprehensive lists of them, see Leonard G. Ratner, "*Ars combinatoria*: Chance and Choice in Eighteenth-Century Music," in *Studies in Eighteenth-Century Music: A Tribute to Karl Geringer on His Seventieth Birthday*, ed. H. C. Robbins Landon (New York: Da Capo Press, 1979), 343–63; Stephen A. Hedges, "Dice Music in the Eighteenth Century," *Music and Letters* 59 (1978): 180–87; and Sebastian Klotz,

| | I | II | III | IV | V | VI | VII | VIII |
|---|---|---|---|---|---|---|---|---|
| 2 | 96 | 22 | 141 | 41 | 105 | 122 | 11 | 30 |
| 3 | 32 | 6 | 128 | 63 | 146 | 46 | 134 | 81 |
| 4 | 69 | 95 | 158 | 13 | 153 | 55 | 110 | 24 |
| 5 | 40 | 17 | 113 | 85 | 161 | 2 | 159 | 100 |
| 6 | 148 | 74 | 163 | 45 | 80 | 97 | 36 | 107 |
| 7 | 104 | 157 | 27 | 167 | 154 | 68 | 118 | 91 |
| 8 | 152 | 60 | 171 | 53 | 99 | 133 | 21 | 127 |
| 9 | 119 | 84 | 114 | 50 | 140 | 86 | 169 | 94 |
| 10 | 98 | 142 | 42 | 156 | 75 | 129 | 62 | 123 |
| 11 | 3 | 87 | 165 | 61 | 135 | 47 | 147 | 33 |
| 12 | 54 | 130 | 10 | 103 | 28 | 37 | 106 | 5 |

FIGURE 4.1 Number table for mm. 1–8 of the *Musikalisches Würfelspiel*

write a sixteen-measure waltz. It is one of several such games attributed to Mozart and is catalogued as K. Anh. C 30.01, but the ascription is hardly secure.[9] Mozart's authorship, in any event, is not my central concern. Instead, I want to focus on the compositional strategies that can be teased out of this particular game.

The game requires the use of a set of dice, and of two tables—a number table (the one used for the first eight-measure period of the waltz is shown in fig. 4.1)

---

*"Ars combinatoria* oder 'Musik ohne Kopfzerbrechen': Kalküle des Musikalischen von Kircher bis Kirnberger," *Musiktheorie* 14 (1999): 231–45.

9. See Wolfgang Amadeus Mozart, *Musikalisches Würfelspiel: Eine Anleitung "Walzer oder Schleifer mit zwei Würfeln zu componieren ohne Musikalisch zu seyn, noch von der Composition etwas zu verstehen,"* ed. Karl Heinz Taubert (Mainz: B. Schott's Söhne, 1956). On the matter of the authenticity of the attribution of this particular game to Mozart, see Paul Löwenstein, "Mozart-Kuriosa," *Zeitschrift für Musikwissenschaft* 12 (1930): 342–46; Otto Erich Deutsch, "Mit Würfeln komponieren," *Zeitschrift für Musikwissenschaft* 12 (1930): 595; Herbert Gerigk, "Würfelmusik," *Zeitschrift für Musikwissenschaft* 16 (1934): 359–63; Ludwig Ritter von Köchel, *Chronologisch-thematisches Verzeichnis sämtlicher Tonwerke Wolfgang Amadé Mozarts*, 8th ed., ed. Franz Giegling, Alexander Weinmann, and Gerd Sievers (Wiesbaden: Breitkopf and Härtel, 1983), 581, 910; and Hedges, "Dice Music in the Eighteenth Century," 183.

It is known from sketch studies (unrelated to the particular game under consideration) that Mozart had an interest in musical diversions of this sort; see Köchel, *Chronologisch-Thematisches Verzeichnis*, 581. And a musical dice game survives that is perhaps more reliably attributed to Mozart; it appears, in his hand, on the verso of a manuscript that contains a fragment—also in Mozart's hand—of the G minor String Quintet K. 516. See Hideo Noguchi, "Mozart—Musical Game in C Major, K. 516f," *Mitteilungen der Internationalen Stiftung Mozarteum* 38 (1990): 89–101. K. 516f is significantly different from the game to be discussed below, however, and the disparities that distinguish the two could be seen as arguing against the authenticity of the latter one. The game from K. Anh C 30.01 that produces the waltz, and that I will consider, was published under Mozart's name soon after the composer's death by Johann Julius Hummel (Amsterdam and Berlin, 1793). This is hardly a guarantee of Mozart's authorship, since his name could readily have been used simply to boost the sales of a work by another (and unknown) author.

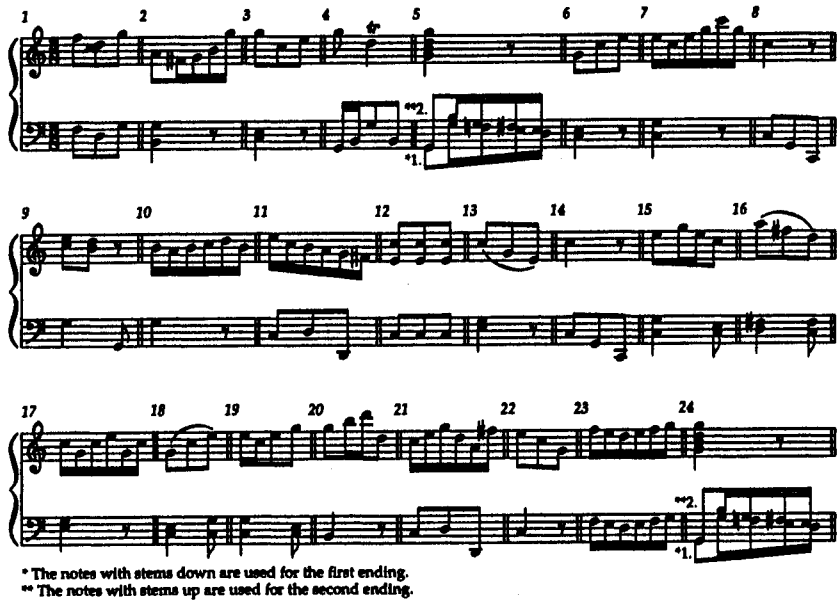


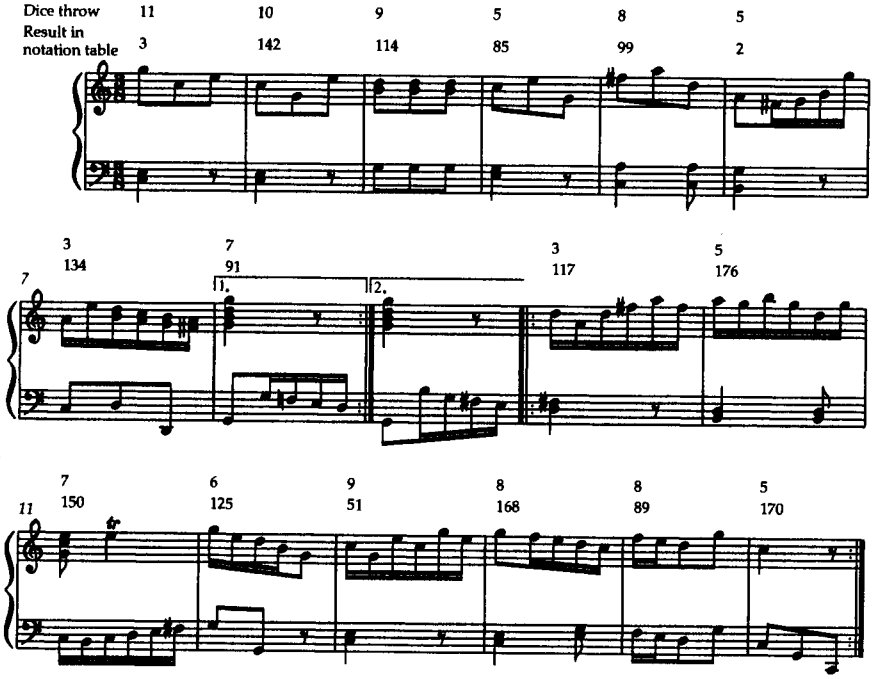

FIGURE 4.2 Portion of the notation table for the *Musikalisches Würfelspiel*

and a notation table (a portion of which is shown in fig. 4.2). For the first measure of music, the player throws the dice and uses the result to find a number in the first column of the table. Thus, the throw 11 gives the result 3 in the first column. The player then turns to the notation table, finds the music designated with a 3 there, and uses it for the first measure of the waltz. For the second measure of music, the player throws the dice again and uses the result to find a number in the second column of the number table. The throw 10 gives the result 142 in the second column; the music that corresponds to 142 on the notation table is then used for the second measure of the waltz. A complete waltz produced by the game is shown in example 4.1. Although the dance is hardly distinctive, it is nonetheless completely serviceable, equal to if not exceeding the quality of the *Tafelmusik* produced by knowledgeable amateurs of the period.

Of course, the point of such games was not the production of works of music, but diversion. There were a number of ways the game could be played: a group of players might sit in a circle and take turns throwing the dice, so that each successive measure was determined by a different player; players might each create their own waltzes, which could then be compared with one another; the entire piece might be written down and then performed; or players might take turns performing the piece as it progressed. In any event, the game offered an amusing and pleasant way to spend a portion of an evening, with the added benefit that music produced by what appeared to be total chance was orderly and coherent, no matter what the skill of the player.

In truth, chance played little part in the success of the music produced by such games. Instead, what was required of the compilers of games such as this one were

EXAMPLE 4.1 Waltz produced by the *Musikalisches Würfelspiel*

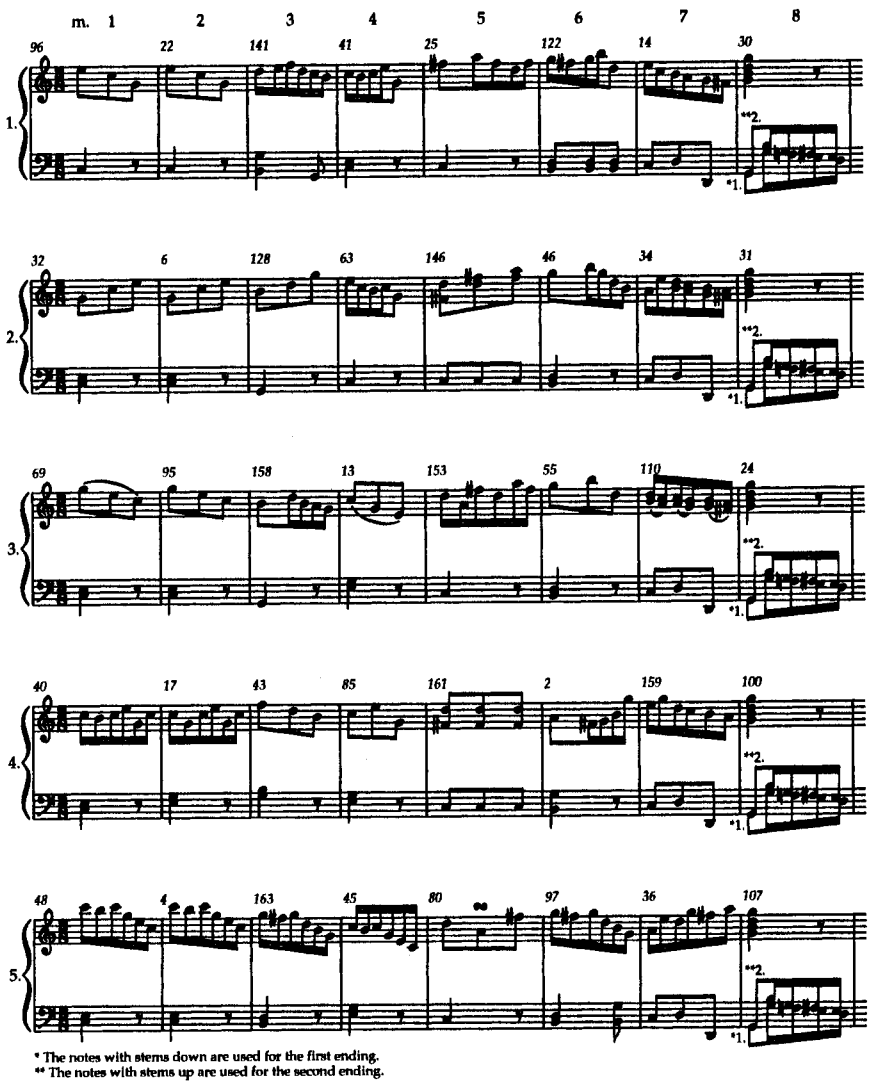



a little knowledge about how to put the game together and an understanding of the formal design of waltzes, contredanses, and the like. One first needs to construct a number table.[10] A table for a sixteen-measure waltz using two dice, for instance, starts with an 11 by 16 grid (for the eleven possible throws of the dice and sixteen measures of the piece; the table shown in fig. 4.1 provides only one-half of this grid). The numbers 1 through 176 are then randomly distributed on the 176 squares of the grid to complete the table. To construct a notation table, it is necessary to write eleven variations on a sixteen-measure waltz, preferably arranging these on a single sheet with the first variation at the top and the eleventh at the bottom. The various numbers of the number table are then mapped onto the individual measures of the eleven variations. For instance, the number table used in our *Musikalisches Würfelspiel* would map the number 96 (found at the intersection of the first row and first column) onto the first measure of the first variation, the number 22 onto the second measure of the first variation, and so on for all 176 measures of the eleven variations. An example of how this might look, derived by working backward from the number and notation tables, can be seen in example 4.2.[11] Once each measure

10. Another approach would be for a number table to be adopted from a previous publication. In fact, the number table for our waltz game is exactly the same as that for the the one published in the Weimar *Journal des Luxus und der Moden* of February 1787. See Karl Heinz Taubert (ed.) *Das Menuett: Musikalisches Wurfelspiel, Paris 1786* (Zurich: Musikhaus Pan, 1988).

11. There is no way to know whether the eleven variations shown in ex. 4.2 are actually the source variations for the game attributed to Mozart: the number found in the first square of the number table

EXAMPLE 4.2 Eleven variations with numbers derived from the number table for the *Musikalisches Würfelspiel*



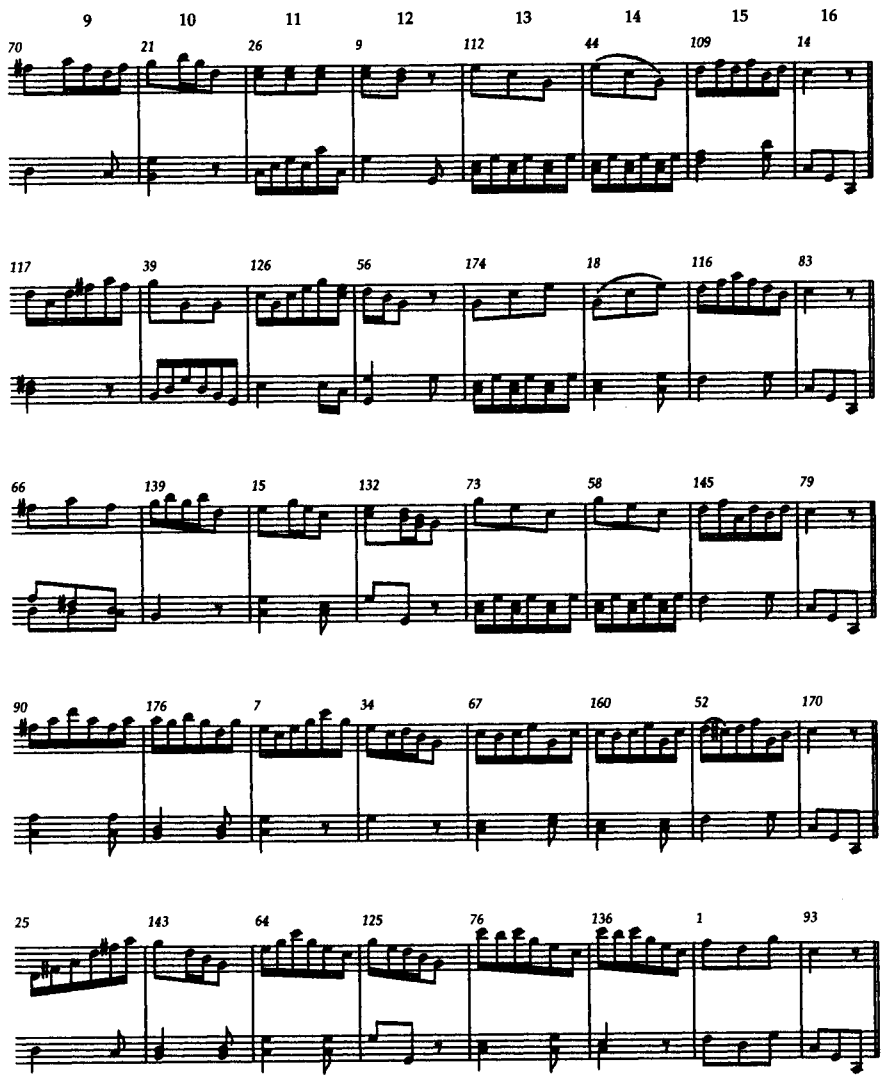

* The notes with stems down are used for the first ending.
** The notes with stems up are used for the second ending.

---

could, in theory, be mapped onto any of the variants of m. 1, and the game would still function properly. However, my focus in the following is not on establishing the original form of each of the eleven variations but on the way the material of each measure is varied. The principal function of ex. 4.2, which puts each measure into its proper column, is to aid discussion of this aspect of the musical game.

Readers who would like to try their luck at composing a piece with the *Musikalisches Würfelspiel* can use ex. 4.2 (supplying their own dice). The number corresponding to each variation can be read as the result of the throw of the dice; the columns denote the measures of the dance, as in the original number table. Of course, playing the game with the number table and the notation table collapsed together (as in ex. 4.2) takes most of the mystery and pretty much all of the fun out of the game. What remains is the somewhat thin pleasure of finding out what combinations of measures chance will dictate.

EXAMPLE 4.2 *(continued)*

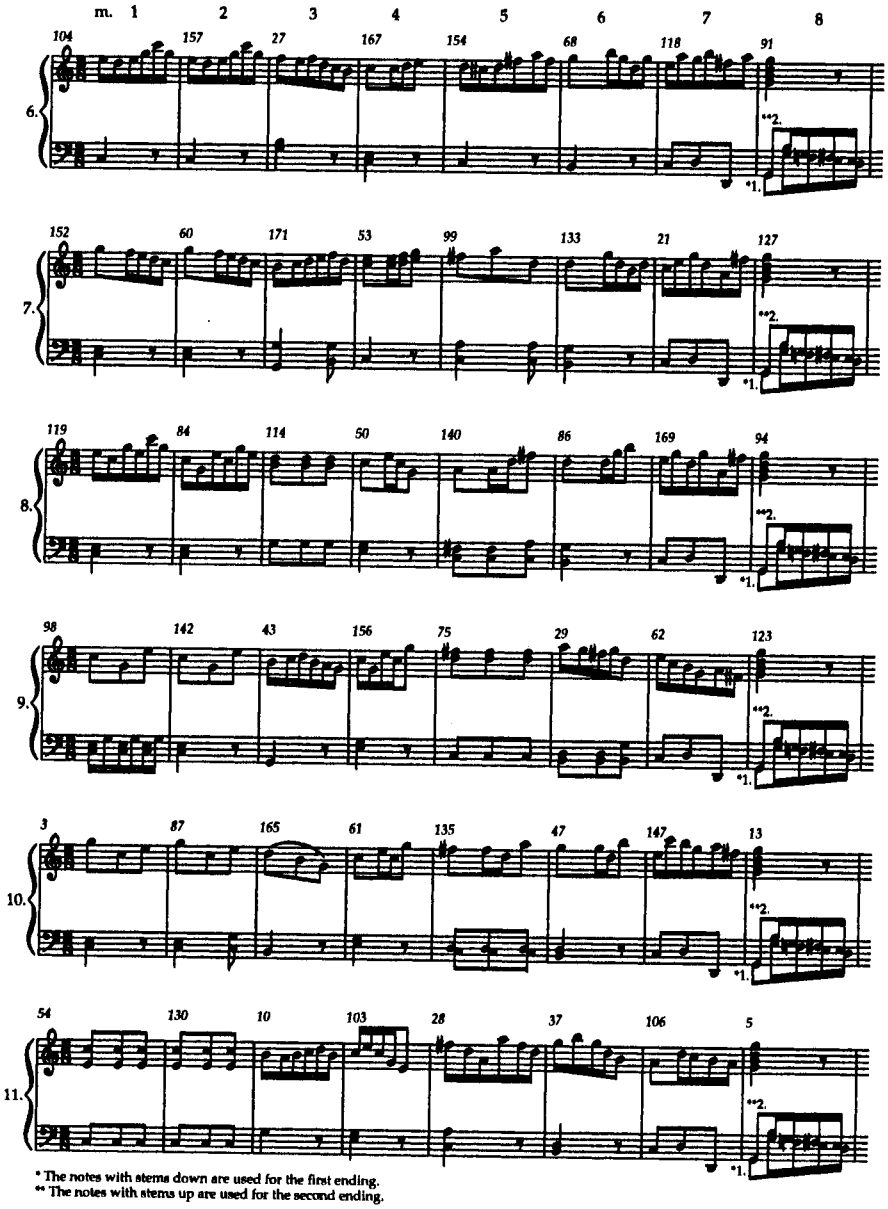



*(continued)*

has been assigned a number, all of the measures are rewritten, in ascending numerical order, to create the notation table.

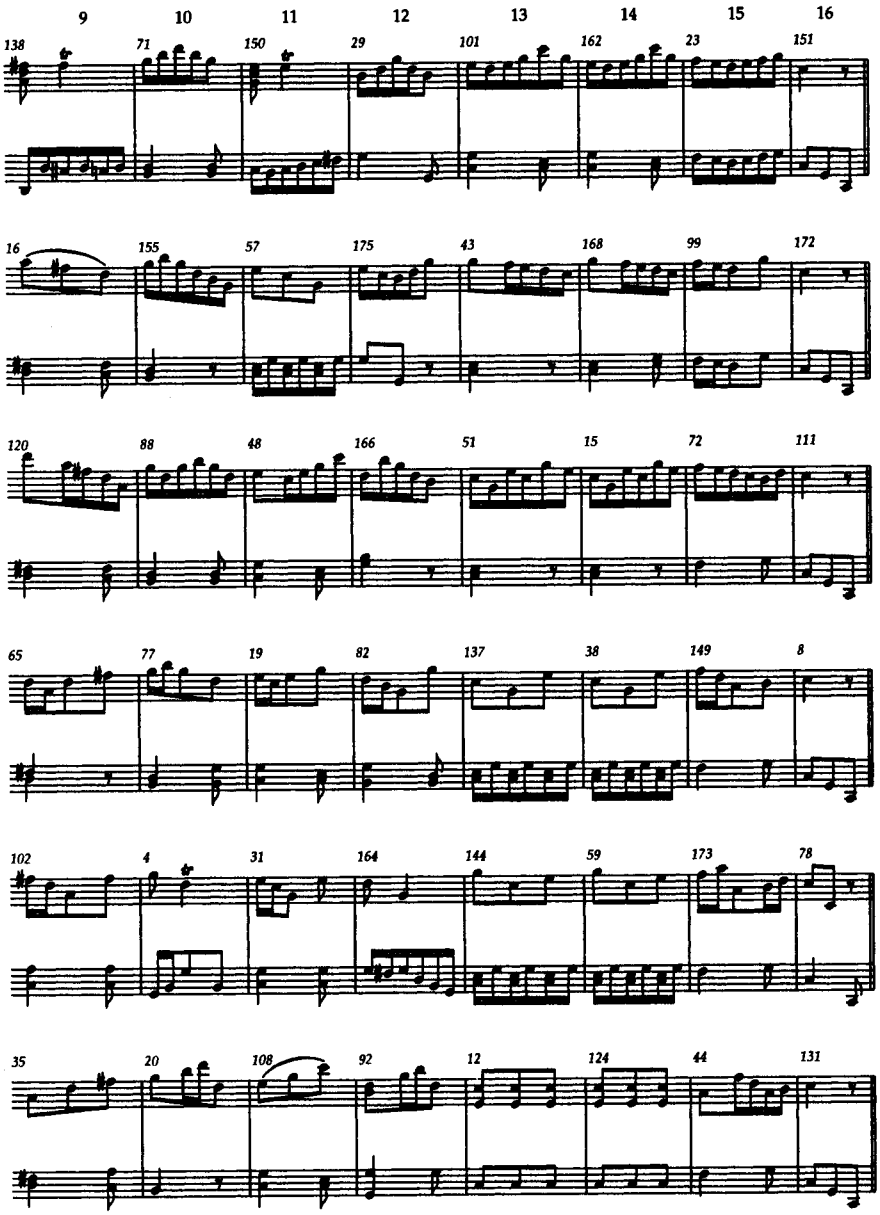
The permutational combination required by the game—each of the variants of m. 1 had to make sense with each of the variants of m. 2; each of the variants of m. 2 had to make sense with each of the variants of m. 3; and so on—placed significant constraints on what could be varied and what sort of music was produced by the game. In general, all of the variants of a given measure had to have the same har-

EXAMPLE 4.2 *(continued)*




* The notes with stems down are used for the first ending.
** The notes with stems up are used for the second ending.

EXAMPLE 4.2 *(continued)*



mony, and they typically had the same bass note. The melody for each measure was restricted to generic figures—arpeggios, an occasional scalar run, or sometimes a single note—that could easily be concatenated with other such figures. In truth, the compiler of the game wrote not so much variations on a waltz as variations on the harmonic plan for a waltz, using melodic fragments that served the minimal contrapuntal plan of the whole.[12] Rhythmic figuration offered the chief means of deriving the variants of a given measure, although changes to the beginning melodic note, overall melodic contour, or number of nonharmonic tones were also used to render each variant unique. The result of a game constructed according to this plan was a set of variations that was, for all practical purposes, inexhaustible: there are theoretically $11^{16}$—that is, 45,949,729,863,572,161—different waltzes that could be written from the materials provided in example 4.2. The waltzes would not be great music, but their formulaic melodies and predictable harmonic plan were perfectly in keeping with dance forms of the late eighteenth century.

The success of the compositions produced by such games is often attributed to the rationalization of musical materials effected during the eighteenth century, which allowed specific compositional functions to be attributed to musical materials.[13] Although such rationalization is quite evident in the work of music theorists like Joseph Riepel and Heinrich Christoph Koch,[14] and while it is supported by the interchangeability of the components of example 4.2, it is not the only thing that makes the waltz of example 4.1 believable as music. What is also at work is a compositional strategy that shapes the whole of the waltz. The strategy is evident in the overall harmonic plan for the dance, an outline of which is given in figure 4.3. The waltz is divided into four-measure sections; the first section (mm. 1–4) establishes the tonic, the second (mm. 5–8) modulates to the dominant, and then the two sections repeat. The third section (mm. 9–12) remains in the dominant, the fourth section (mm. 13–16) returns to the tonic, and these, too, are repeated. Every one of the waltzes generated by the game will have this basic harmonic syntax. However, a succession of harmonies alone does not create convincing music—Rameau's harmonic theory notwithstanding, every succession of dominant to tonic does not make for a cadence. What is crucially important for the compositional strategy embodied by the *Musikalisches Würfelspiel* is the way the material for each measure is varied. For instance, note that only one version of m. 8 appears in example 4.2

12. Mostly, the counterpoint above the bass is relatively free (within the strictures of style). In a few cases, however—mm. 9–10, for example, where the dominant in $\frac{4}{2}$ position leads to a tonic $\frac{5}{3}$, and at the cadences that conclude each period—the melodic possibilities are constrained by the overall counterpoint in the outer voices.

13. Ratner, "*Ars combinatoria*," 344–45.

14. See Joseph Riepel's *Anfangsgründe zur musicalischen Setzkunst* (1752–68), included in his *Sämtliche Schriften zur Musiktheorie*, Wiener musikwissenschaftliche Beiträge (Vienna: Böhlau, 1996); and Heinrich Christoph Koch, *Versuch einer Anleitung zur Composition* (Leipzig: Adam Friedrich Böhme; Rudolstadt: gedruckt mit Schriften der Löwischen Erben, und Schirach, 1782 [vol. 1]; Leipzig: Adam Friedrich Böhme, 1787–93 [vols. 2–3]; repr. Hildesheim: Georg Olms, 1969), esp. vol. 2. For a translation of portions of the latter, see Heinrich Christoph Koch, *Introductory Essay on Composition: The Mechanical Rules of Melody*, Sections 3 and 4, trans. Nancy Kovaleff Baker, Music Theory Translation Series (New Haven, Conn.: Yale University Press, 1983). The work of both Riepel and Koch is discussed more extensively in the first half of chap. 7 in this volume.

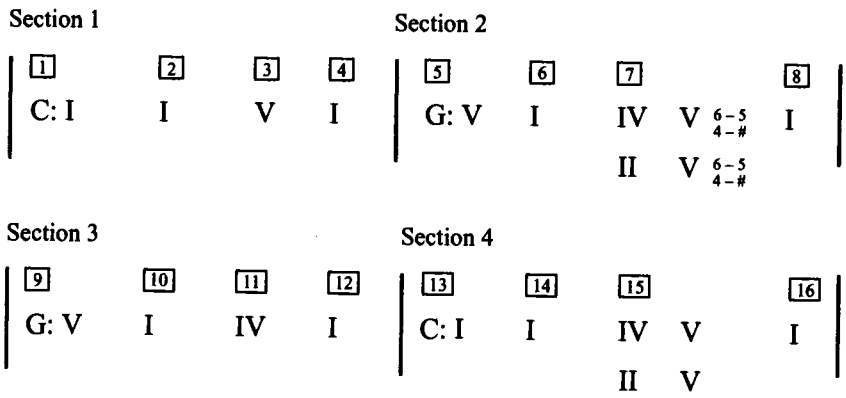



FIGURE 4.3 Basic harmonic plan for the *Musikalisches Würfelspiel*

and, save for the ending of the tenth variation, only one version of m. 16.[15] By contrast, none of the versions of m. 1 are identical. Shaping each of the trillions of waltzes that might be generated by the game is a template that regulates the degree and kind of variation that can occur within the sixteen-measure framework of the dance.

## A Template for Compositional Strategy

We can discover the plan of the template if we approach each column of measures in example 4.2 as a Type 1 category of the sort discussed in chapter 1. Each measure-category can be defined in terms of attributes (such as harmony, rhythmic figuration, beginning melodic note, and so on), to which each of the members of the category (that is, the different variants included in the category) supplies values. Each measure-category also has a distinctive profile of typicality, determined by the range of values its members assign to the various attributes and the weight given any particular set of values. For example, the members of measure-category 16 assign two different values for the attribute *rhythmic figure*, but assign one of these ten out of eleven times (as shown in fig. 4.4). With respect to the attribute *rhythmic figure*, we can say that the profile of typicality for measure-category 16 is clearly defined. By comparison, the members of measure-category 15 assign seven different values for the attribute *rhythmic figure*; one value is assigned five times, the other six are assigned once each. The profile of typicality for rhythmic figuration is thus somewhat less clearly defined for measure-category 15.

Two patterns can be glimpsed from figure 4.4. The first concerns which attribute-values change from measure-category to measure-category and which stay the same—for instance, there is no rhythmic figure in common between measure-categories 15 and 16. The second pattern concerns changes in the profile of typi-

15. The lack of variation in mm. 8 and 16 means that only $(11^{14}) \times 2$—merely 759,499,667,166,482—different waltzes can be written using this particular game.

| measure-category 15 | | measure-category 16 | |
|---|---|---|---|
| rhythmic figure | number of times used | rhythmic figure | number of times used |
|  | 5 |  | 10 |
|  | 1 |  | 1 |
|  | 1 | | |
|  | 1 | | |
|  | 1 | | |
|  | 1 | | |
|  | 1 | | |

FIGURE 4.4 Values assigned the attribute *rhythmic figure* in measure-categories 15 and 16

cality between measure-categories—as can be seen, there is a relatively sharp contrast between the profiles of measure-categories 15 and 16. These patterns, taken together, create the template that shapes each of the waltzes generated by the game. A somewhat more complete perspective is provided by looking at the values assigned to just two attributes—*harmony* and *rhythmic figure*—in the first section of the waltz, as illustrated by figure 4.5. As can be seen, there is little change in the values that members of measure-category 1 and measure-category 2 assign to these attributes: one new rhythmic figure is introduced and one is dropped, but otherwise the categories are (at least with respect to these two attributes) identical. While measure-category 3 retains most of the rhythmic figures of measure-category 2, the values assigned the harmony are completely different and markedly more various. Finally, measure-category 4 has significant changes in the values assigned both attributes: two new rhythmic figures come to prominence, and the attribute-values for harmony change once more.

| | I | I⁶ | V | V⁶ | V⁷ | V⁶₅ |  |  |  |  |  |  |  |  |  |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| m. 1 | 11 | | | | | | 4 | 4 | 1 | 1 | 1 | | | | |
| m. 2 | 11 | | | | | | 4 | 4 | 1 | | 1 | 1 | | | |
| m. 3 | | | 5 | 2 | 3 | 1 | 3 | 4 | 1 | | 1 | | 2 | | |
| m. 4 | 4 | 7 | | | | | 2 | 1 | | | | | | 4 | 4 |

FIGURE 4.5 Attribute-values for *harmony* and *rhythmic* figure in measure-categories 1 through 4 (*left side*: the values assigned to the attribute *harmony* by the members of measure-categories 1 through 4 and the number of members within each measure-category that assign that value; *right side*: the values assigned to the attribute *rhythmic figure* and the number of members within each measure-category that assign that value)

There are also changes to the profile of typicality for the measure-categories in this section. While the profile for measure-categories 1 and 2 is identical—a typical variant will have a root-position tonic harmony and use either the first or second rhythmic figure—it is less clearly defined for measure-category 3. Although the number and range of rhythmic figures do not change dramatically, there are now four possible values for the attribute harmony, and only a bit of additional weight is put on one of these attribute-values (the root-position dominant is assigned by five of the eleven members of the category). Things come back into focus with measure-category 4—the most typical variant is a first-inversion harmony using either of the last two rhythmic figures—but the profile of typicality is still not quite as clearly defined as it was in the first two measure-categories.

The syntax of this section (reckoned in terms of harmony and rhythmic figure) is one in which new materials are gradually introduced and displace previous material. At the same time, measure-categories with a fairly clear profile of typicality give way to ones that have a less clearly defined profile (a trend reversed somewhat by measure-category 4). Similar syntactic patterns structure the remaining three sections of the waltz. The most important features of the syntactic template for the waltzes produced by the game, taken as a whole, are as follows:[16]

> New attribute-values—especially for rhythmic figure—are introduced throughout the first three sections. At the beginning of section 4, this process is temporarily halted and, coincident with the return C major in m. 13, rhythmic figures are for the first time brought back.
>
> Cadences (such as those that occur in measure-categories 7–8 and 15–16) are marked by a distinctive treatment of attribute-values: values that occur in

16. In developing this characterization of the changes in attribute-values and typicality that can be found in the measure-categories collected in ex. 4.2, I considered the attributes of contour, number of non-chord tones, bass motion, and starting pitch of the melody, as well as harmony and rhythmic figure. I found, however, that focusing on the latter two attributes provides the clearest picture of typicality within—and across—the measure-categories. In consequence, my account of the syntactic template tends to emphasize these attributes.

no other measure-category occur here, and the number of different attribute-values within each measure-category is sharply attenuated. In particular, the measure-categories that correspond with harmonic arrivals (that is, 8 and 16) have the most clearly defined profile of typicality of any of the sixteen.

The syntax of sections 1 and 2 involves complementary processes: section 1 begins with measure-categories that have clearly defined profiles of typicality; the profiles of the measure-categories at the end of the section are less clearly defined. Section 2 begins with measure-categories that have less clearly defined profiles of typicality; at the end of this section, the profiles of the measure-categories are clearly defined.

The syntax of sections 3 and 4 involves contrasting processes: section 3 has the broadest range of attribute-values of any of the sections (there are, for instance, over two times as many different rhythmic figures in section 3 as there are in section 1) and the most poorly defined profiles of typicality. By comparison, the range of attribute-values is much more limited in section 4, and the clarity of typicality is perhaps second only to section 1.

The compositional strategy embodied in this syntactic template is one that would be familiar to aspiring composers of the eighteenth century, for it is in complete accord with the compositional strategies for sixteen-measure dance forms proposed in treatises of the period. Theorists like Riepel and Koch taught that the tonic should be plainly stated in the first section (the measure-categories with tonic harmony in this section have the most clearly defined profile of typicality); there should be clear cadences at the end of the second and fourth sections (in the musical game, this is accomplished by the distinctive treatment accorded the end of sections 2 and 4); and the third section should be planned so that it does not become independent from the whole—in Koch's words, "it must be formed so that it loses its completeness and is bound more closely with the following fourth melodic section"[17]—something achieved in the template by the large amount of variation among the members of measure-categories 9 through 12.

At a more abstract level, the compositional strategy incorporated in the template is, of course, similar to late-eighteenth-century conceptions of what is now conventionally called sonata form.[18] These conceptions divided sonata form into two main

17. Koch, *Introductory Essay on Composition*, 100. Koch's discussion of sixteen-measure dance forms is on pp. 78–117.

18. The classic study of eighteenth-century conceptions of sonata form is Leonard Ratner's "Harmonic Aspects of Classic Form," *Journal of the American Musicological Society* 2 (1949): 159–68. The viability of eighteenth-century conceptions of sonata form, and their relevance for recent literature on the topic of form, have been given a thorough discussion recently in Mark Evan Bonds, *Wordless Rhetoric: Musical Form and the Metaphor of Oration*, Studies in the History of Music, 4 (Cambridge, Mass.: Harvard University Press, 1991). A thoughtful critique of Bonds's work can be found in Peter Hoyt, "Review of Mark Evan Bonds, *Wordless Rhetoric*," *Journal of Music Theory* 38 (1994): 123–43.

It should be noted that recent scholarship has argued for a substantially more complicated view of sonata form than that presented by either Ratner or Bonds. For an insightful discussion, see Scott Burnham, "The Second Nature of Sonata Form," in *Music Theory and Natural Order from the Renaissance to the Early Twentieth Century*, ed. Suzannah Clark and Alexander Rehding (Cambridge: Cambridge University Press, 2001), 111–41.

parts, whose functions were determined by harmonic syntax: the *Exposition* (which set out the tonic and, in major keys, closed on the dominant); and the *Development* and *Recapitulation* (which closed on the tonic).[19] The Exposition presented the principal materials of the movement, much as sections 1 and 2 of the musical game present tonic and dominant; the Development departed from these materials, as does section 3 of the game; and the Recapitulation returned to the material of the Exposition, just as there is a reprise of previously used material in section 4 of the game.

Given these similarities, we can see that the conception of compositional strategy behind the template provides a link between modest sixteen-measure dance forms and sonata form: both rely on the same syntax for organizing musical materials.[20] This points out something rather important about the teaching of composition in the eighteenth century: the highly detailed accounts of dance forms by Koch and others did indeed set out the means by which modest utilitarian works could be generated, but they also schooled students in compositional strategies that could be applied to larger works. These larger works, then, were not simply generated by repeated applications of the compositional techniques described for minuets and the like, but by the reinterpretation of these techniques as strategies for composition that could be applied to categories of musical events. When these categories were realized over longer spans of time, extended works (such as those by Mozart and Beethoven discussed in the following sections) were born.

## Typicality Effects and Musical Syntax

What emerges from this analysis of the *Musikalisches Würfelspiel* is the significance of typicality effects for compositional strategy. Important structural points in the syntactic template—the establishment of tonic in the opening measures, the cadence on the dominant at the end of the second section, and the cadence on the tonic at the end of the fourth section—were all marked by measure-categories with clearly defined profiles of typicality. Indeed, a functional distinction between tonic and dominant can be made simply on the basis of the amount of variation present among the members of the measure-categories associated with each. Taking the template as a whole, there is less variation in (and thus a clearer profile of typicality for) measure-categories associated with the tonic (that is, measure-categories 1–2, 4, 13–14, and 16); there is comparatively more variation in (and consequently a less clear profile of typicality for) measure-categories associated with the dominant (that is, measure categories 3, 6, 8, 10, 12, and 15).[21]

19. In the following, I capitalize "Exposition," "Development," and "Recapitulation" when they refer to components of sonata form, reserving the uncapitalized versions of the terms for the compositional techniques through which musical materials are laid out, developed, and reprised. There is, of course, a certain measure of anachronism in applying these terms to eighteenth-century works, especially when sonata form is conceived as a two-part structure; nonetheless, the terms do single out compositional strategies recognized, if not so named, by eighteenth-century theorists.

20. The absence of an explicit link between the discussion of the appropriate compositional strategy for dance forms and that for sonata form has been noted by Bonds; see *Wordless Rhetoric*, 26–27.

21. I interpret the C major harmony of measure-category 11 as the subdominant of G major; note the F♮3 in the bass of the sixth variation.

What also emerged was the possibility of describing the syntactic structure for a two-part form purely in terms of a general set of compositional strategies realized through the manipulation of categories of musical events. There is no doubt that, in the case of the waltz created by the game, harmonic syntax is of central importance; nor is there any doubt that those who have described form with an emphasis on harmonic syntax do it with the assumption that this syntax also embodies compositional strategies applied to other aspects of music. Nonetheless, the account I have provided shows how certain compositional strategies—viewed as a generalized set of processes on musical materials—operate independently of the strategies associated with harmonic syntax.

With these thoughts in mind, let us turn to some actual musical compositions—as opposed to recreational games—that make use of some of the compositional strategies seen in the *Musikalisches Würfelspiel*, as well as more novel strategies that are appropriate for extended and adventuresome instrumental works.

## MOZART, STRING QUARTET IN C MAJOR (K. 465), FIRST MOVEMENT. ALLEGRO

The Allegro of Mozart's String Quartet in C major (K. 465) was one of Arnold Schoenberg's favorites for illustrating motivic development of the sort discussed in chapter 1 of this volume. He found the work appropriate for these purposes in part because of the prominence of the motive among the various musical materials and in part because it is easy to find traces of the motive even after it has left the scene (thus supporting the organic conception of musical composition of which Schoenberg was so fond).[22] Because my concern here is not with developing an account of the way a musical whole is derived from a single theme, but with the motive as an element in an unfolding musical discourse, my account approaches the syntax of the principal melodic motive as independent from syntactic levels occupied with harmony and stylistic formulae, and independent from musical signs or syntagms of the sort noted by Eco. I shall, however, consider how these are correlated with the syntax of the principal motive toward the end of my analysis.

Mozart completed the quartet in C major during January of 1785, and in September of that year it was published by Artaria in Vienna as one of a set of six quartets dedicated to Joseph Haydn.[23] The first movement of K. 465 begins with an Adagio (mm. 1–22) remarkable for its harmonic peregrinations and rather odd pitch juxtapositions, which together earned the work its nickname as the "Dissonance" Quartet. The ensuing Allegro, by contrast, has hardly anything dissonant

22. For a discussion of Schoenberg's analyses of this movement, see the Commentary by Patricia Carpenter and Severine Neff in their translation of Schoenberg's *The Musical Idea and the Logic, Technique, and Art of Its Presentation* (New York: Columbia University Press, 1995), 53. Schoenberg first developed his idea of "developing variation"—which he later, and famously, applied to the music of Brahms—in an analysis of this movement made in 1917; see Arnold Schoenberg, *Coherence, Counterpoint, Instrumentation, Instruction in Form*, ed. Severine Neff, trans. Severine Neff and Charlotte M. Cross (Lincoln: University of Nebraska Press, 1994), 41.

23. For a fuller account of the quartet and its companions, see John Irving, *Mozart: The "Haydn" Quartets*, Cambridge Music Handbooks (Cambridge: Cambridge University Press, 1998).

about it. The Allegro is cast in sonata form, and the principal melodic motive appears prominently in five places: the opening and end of the Exposition (mm. 23–51 and 91–106, respectively); the Development section (mm. 107–54); and the opening and end of the Recapitulation, extending into the Coda (mm. 155–70 and 211–46). My discussion concentrates on the first three of these, with briefer comments on the beginning of the Recapitulation and the Coda.

## Measures 23–51 of the Allegro

Example 4.3 gives the first twenty-nine measures of the Allegro. The principal motive of the movement is presented in the first violin in mm. 23–24, and then again in mm. 25–26 (symbolized with α in ex. 4.3). Using an analytical perspective that will inform each of the remaining analyses, I shall approach the various forms of this motive as a cognitive category, structured around a conceptual model. I shall leave open the question of the status of these conceptual models—that is, whether they are empirically verifiable constructs or simply convenient analytical tools—and proceed on the general understanding developed in chapter 1 and briefly reviewed at the beginning of this chapter: categories are basic structures in terms of which we understand the world, they are organized around conceptual models, and they represent a common ground where the concerns of listeners and composers meet.

The model for the category consists of three correlated conceptual elements, characterized as the things necessary for a melodic fragment to count as an instance of the motive:

- a rhythmic pattern that spans two measures of common time: 
- a rather specific diatonic contour, consisting of four elements (each is shown in ex. 4.4); in order of succession, they are

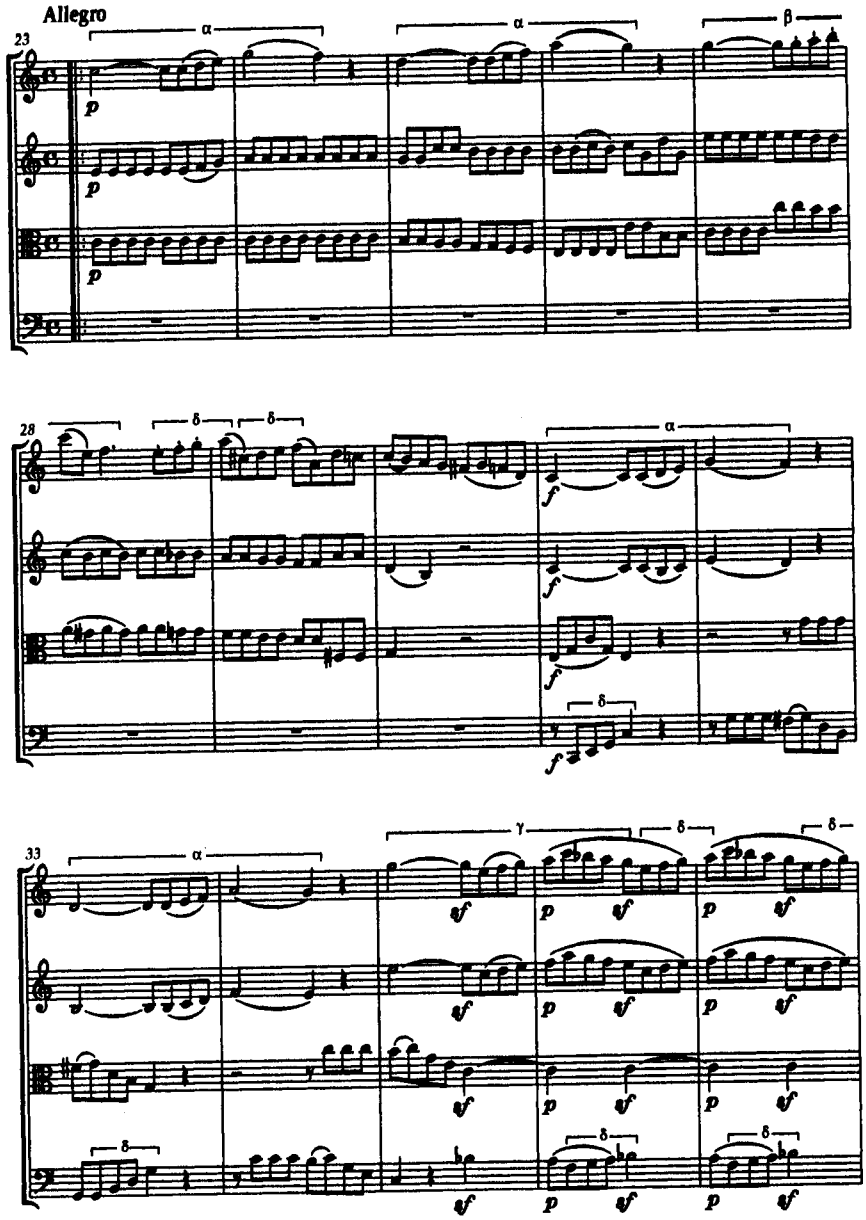
  (i) a sustained first note
  (ii) an ascending stepwise motion through a third
  (iii) an upward skip of a third
  (iv) a downward step

- an implied harmonic change from the first to the second measure of the motive

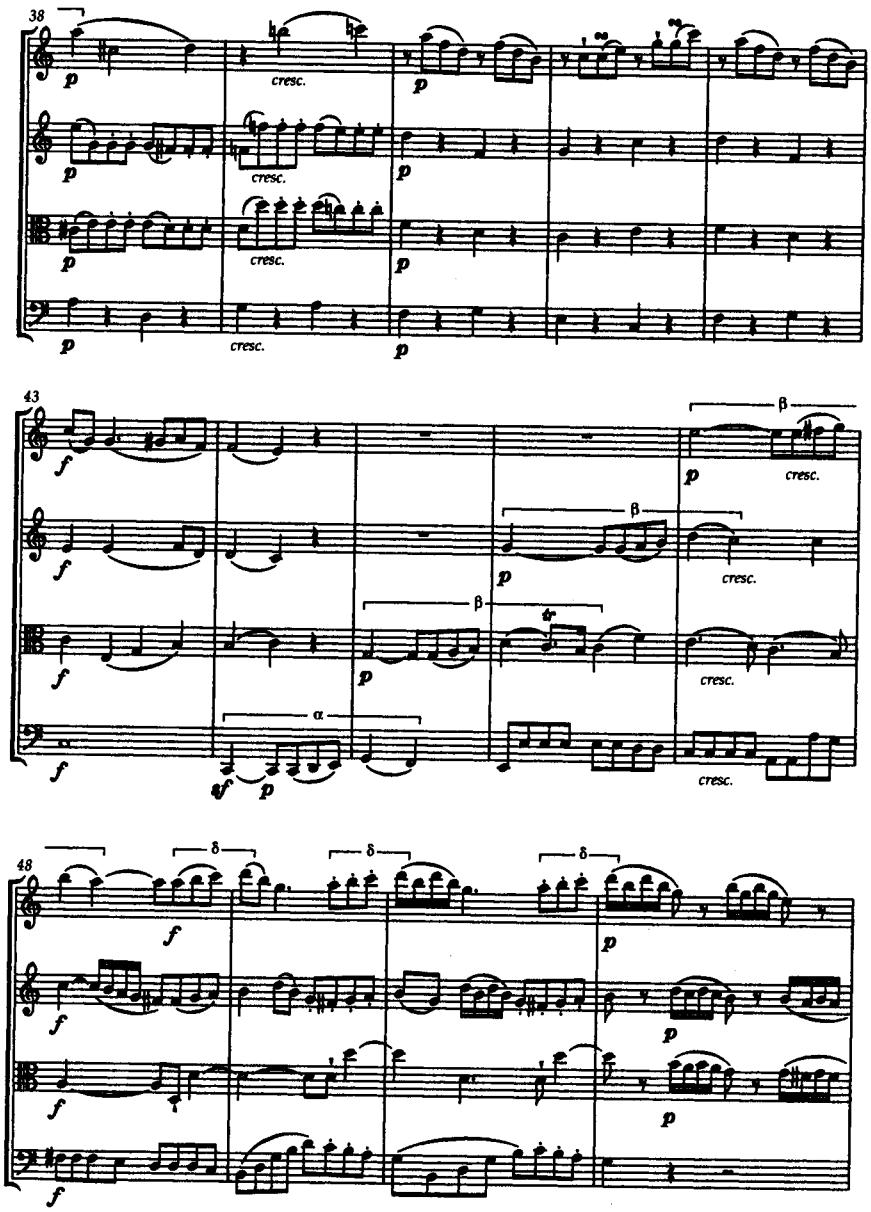
The contour of the motive gives rise to two distinctive intervals: the fifth spanned by ii and iii; and the fourth between i and the last note of iv. Although the implied change in harmony is masked in mm. 23–24 (owing to the viola's pedal point), in most other cases it is quite apparent.

After mm. 23–26, there are three further instances of the motive that correspond to this model: two in mm. 31–34 (with the motive accompanied by the second violin), and one in the cello in mm. 44–45 (with the cello's concluding half note taking the place of the quarter-note and quarter-rest of the original). Somewhat less-typical forms appear in the first violin in mm. 27–28 and in the ascending cascade of mm. 45–48 (β in ex. 4.3). In each case, the very end of the motive has been

EXAMPLE 4.3 W. A. Mozart, String Quartet K. 465, first movement, mm. 23–51



EXAMPLE 4.3 *(continued)*



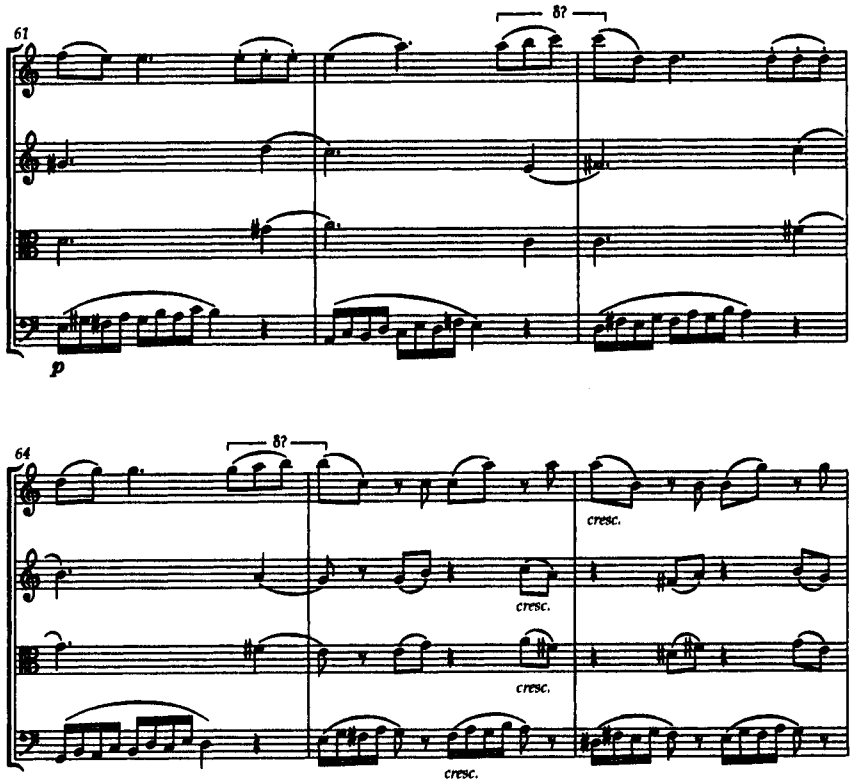
EXAMPLE 4.4 Schoenberg's analysis of the main motive from W. A. Mozart, String Quartet K. 465, first movement




modified: in mm. 45–48, the most significant change is to the end of the rhythmic figure; in mm. 27–28, both the end of the rhythmic figure and contour pattern have been changed. One additional version of the motive, heard in the first violin in mm. 35–36 (γ in ex. 4.3), is still less typical, retaining most of the rhythmic pattern of the original, but only element ii of the contour.

One other bit of musical material (shown with δ in ex. 4.3) seems to demand affiliation with the motivic category. As but a subcomponent of the motive—the ascending gesture that ends on a strong beat—it is hardly typical of the category. Indeed, hearing the fragment as a member of the category requires that it appear in close proximity to more typical members of the category. When a similar gesture occurs in mm. 62–63 and 64–65 in the midst of the first subsidiary theme (as shown in ex. 4.5), it is decidedly subservient to the alternation of ascending and

EXAMPLE 4.5 W. A. Mozart, String Quartet K. 465, first movement, mm. 61–66



descending skips that dominates the musical scene, having lost its connection to the principal motive. The change in the way we hear this fragment—in mm. 48–51, as belonging to the category that includes the principal motive; in mm. 62–65, as a rather generic bit of filler—is an example of the phenomenon of conceptual slippage noted in chapter 3, but one applicable to purely musical concepts. The context that induces this slippage is the absence of the principal motive and the presence of new material associated with the first subsidiary theme.

The melodic syntax for these opening measures, then, involves setting out the most typical members of the category first—indeed, these provide the basis for the conceptual model for the category.[24] Less typical members of the category follow (in mm. 27–28 and 35–36), and, in mm. 37–43, the motive temporarily disappears altogether (save for the δ fragments). It then reappears in its most typical form in m. 44, immediately followed by three slightly less typical versions. After this, there are further δ fragments that lead to the generic cadential material that prepares for the entrance of the first subsidiary theme in m. 56.

Taken as a whole, the motive forms presented in mm. 23–51 are a relatively normative Type 1 category. What is striking, however, is the dynamic shape of the category as it is presented over the course of time. Although typical members dominate the musical texture in the first half of the section, they momentarily disappear from the texture in the second half, to be brought back at an important moment preceding the transition to the first subsidiary theme. The importance of the principal motive for the melodic syntax of mm. 23–51, then, is less a function of its constant presence and more a function of its memorability and cognitive utility. These make possible a musical syntax of both exposition and development: exposition as the motive is first established in its most typical form; and development as less-typical variants are then introduced.

## Measures 90–106 of the Allegro

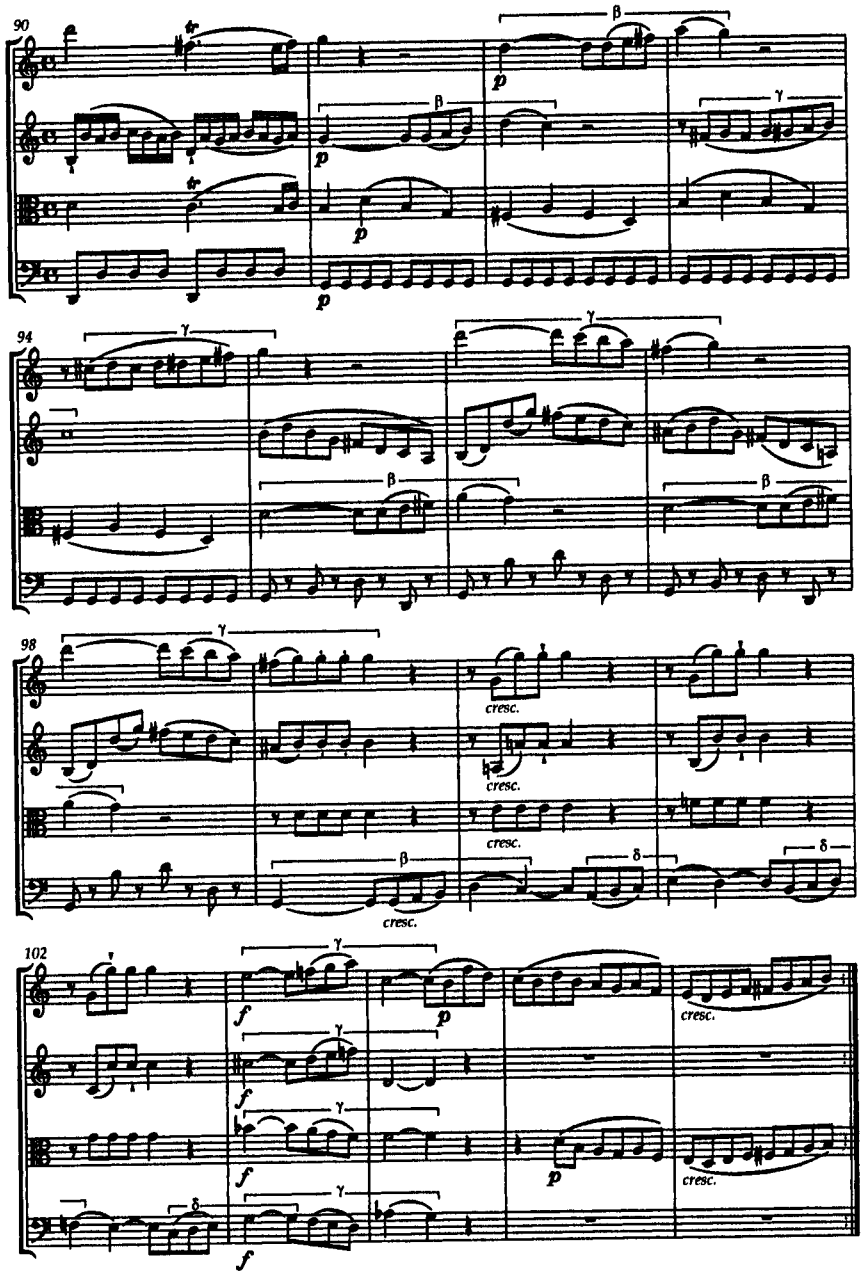
After its gradual departure in mm. 44–51, the motive disappears completely for forty measures, during which time two subsidiary themes are presented in the dominant. When the motive does return, it is at a moment of opportunity, following a final emphatic cadence of the sort that usually signals the conclusion of a section or that prepares the way for additional thematic material. By bringing back the main motive at this point, Mozart ties together the beginning and end of the Exposition. The bonds are relatively loose ones, however, for it is a variant of the main motive rather than the original that is introduced, and the compositional strategy applied to the statements that follow is closer to development than it is to exposition.

As shown in example 4.6, the motive enters in the second violin in m. 91, just as the first violin achieves its cadential G5.[25] The first violin responds by taking up

24. As noted in chap. 1, it is also possible for the conceptual model for a category of motive forms to emerge over time, although the longer the process is delayed, the more the conceptual model comes into conflict with the global model for a musical theme.

25. The sustained note at the beginning of the motive allows Mozart to introduce it under the cover of cadence, something he exploits at various developmental moments throughout the movement.

EXAMPLE 4.6 W. A. Mozart, String Quartet K. 465, first movement, mm. 90–106



the motive as well, beginning a short game of motivic tag in which some form of the motive enters in each of the succeeding measures. Here, the initial motive forms are of the β variety, highly similar to the original form but with a slightly modified ending. Although harmonic change is still evident, the constant overlap of motives limits the harmonic possibilities to the tonic and dominant of G major. The sense of harmonic progress that accompanied the original statement of the motive is markedly attenuated: contributing factors are the constant alternation between tonic and dominant, the cello's pedal point in mm. 91–95, and the halving of harmonic rhythm that follows in m. 96.

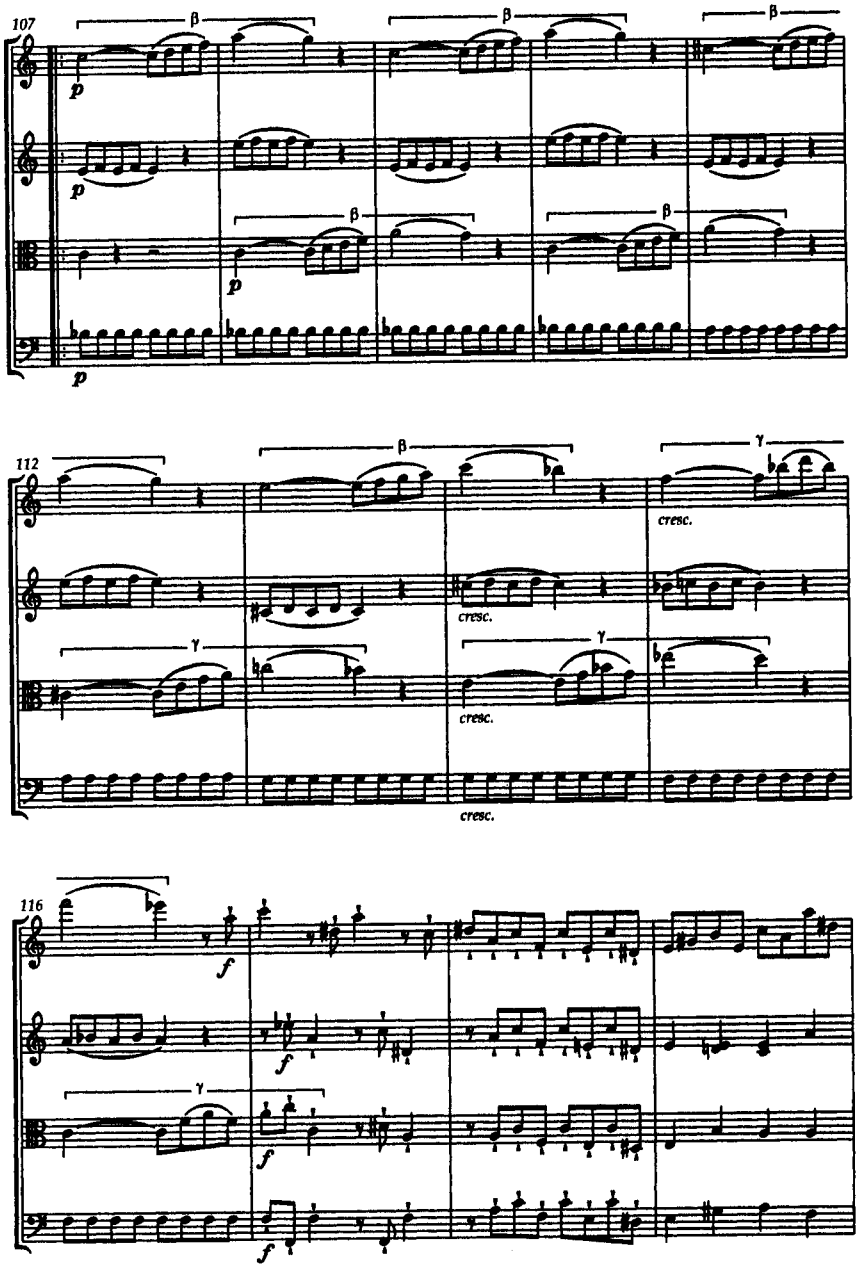
This reprise of the motive bears witness to new, and somewhat more extensively varied, members of the motivic category. In mm. 93–94 and 94–95, the second and first violins present embellished versions of the motive forms they have just stated (indicated with a γ in ex. 4.6); were it not for the temporal proximity of these to the preceding motive forms, it might be difficult to make a case for their derivation from the original. More obvious is the inverted form of the motive that appears in the first violin in mm. 96–99, and which is then used in the simultaneous presentation of motive forms in all four instruments in mm. 103–104. Finally, the δ fragment that provided a measure of unity at the opening of the exposition reappears in mm. 100–103 to serve as a link between the motive forms that precede and follow it.

Given the absence of the original form of the motive in these measures, a case could be made for modifying the conceptual model for the category so that variants of the sort shown with a β in example 4.6 would be regarded as typical (indeed, such a move would be the first step along a slippery slope of motivic development of the sort that Wagner leads us down with his transformations of the Leidensmotiv). Nonetheless, retaining the sense that we have moved away from what is typical of the category is important for what happens next. If the repeat sign of m. 106 is observed and the Exposition reprised, the typicality of the original version of the motive will be reaffirmed: we will have moved from the edges of the category (as represented by mm. 91–106) back to the center. If we instead proceed from m. 106 into the Development, we will move yet farther toward the fringes of the category.

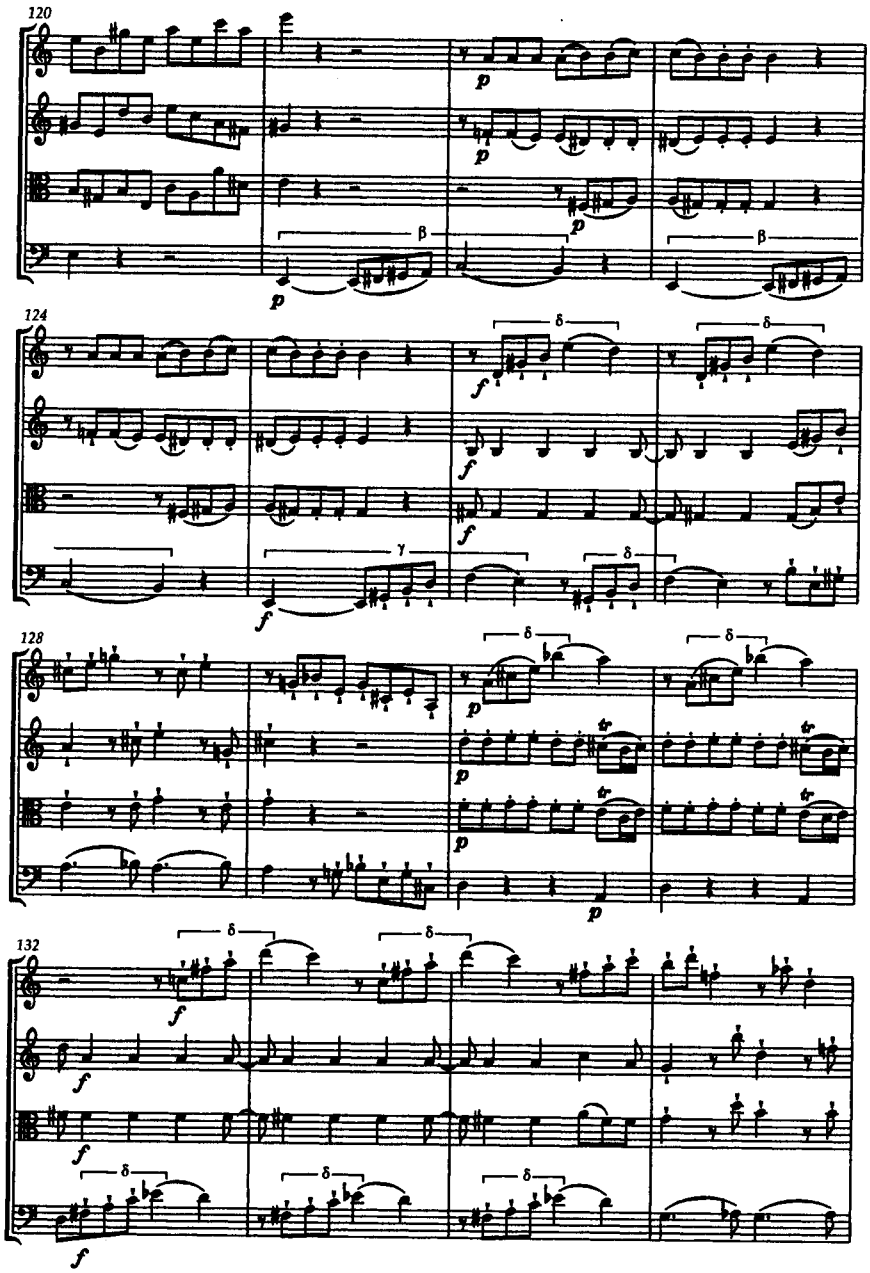
## Measures 107–35 of the Allegro

As shown in example 4.7, Mozart begins the Development with a slightly modified version of the original motive, but with changes now made to the middle of the motive rather than to the end: he starts the ascending-third figure a step higher than the sustained note, instead of repeating the sustained note at the beginning of the figure. As a result, the motive, as a whole, now spans a fifth and implies no change of harmony. The chromatic alteration in the first violin at m. 111 moves the motive a bit further away from its typical form, but it is the version in the viola at mm. 112–13 that accelerates the process of transformation. In this version (indicated with a γ in ex. 4.7), the ascending-third figure is replaced with an arpeggio, which becomes a feature of the following motive forms. At m. 117, the arpeggio figure takes over, and the subsequent dissolution of the motive leads to a momentary cadence on the dominant of A minor in m. 121.

EXAMPLE 4.7 W. A. Mozart, String Quartet K. 465, first movement, mm. 107–35



EXAMPLE 4.7 *(continued)*



At m. 121, the cello returns to the form of the motive used at the beginning of the Development (but now within the context of A minor). After two statements, the rising arpeggio figure once again replaces the ascending third within the motive, after which point it displaces the main motive or its variants altogether (as indicated by the preponderance of δ symbols after m. 125 in ex. 4.7). In m. 147 (which is not given in ex. 4.7), the form of the motive used at mm. 107 and 121 makes a final appearance (beginning with the cello's G2, but pointing toward C minor). Given the events that followed the previous two appearances of this form of the motive, this third appearance seems to suggest yet another episode of development. Nonetheless, in m. 151, the motive yields to an ascending cascade of arpeggio-figures (all outlining a G dominant-seventh), which brings the Development to a close.

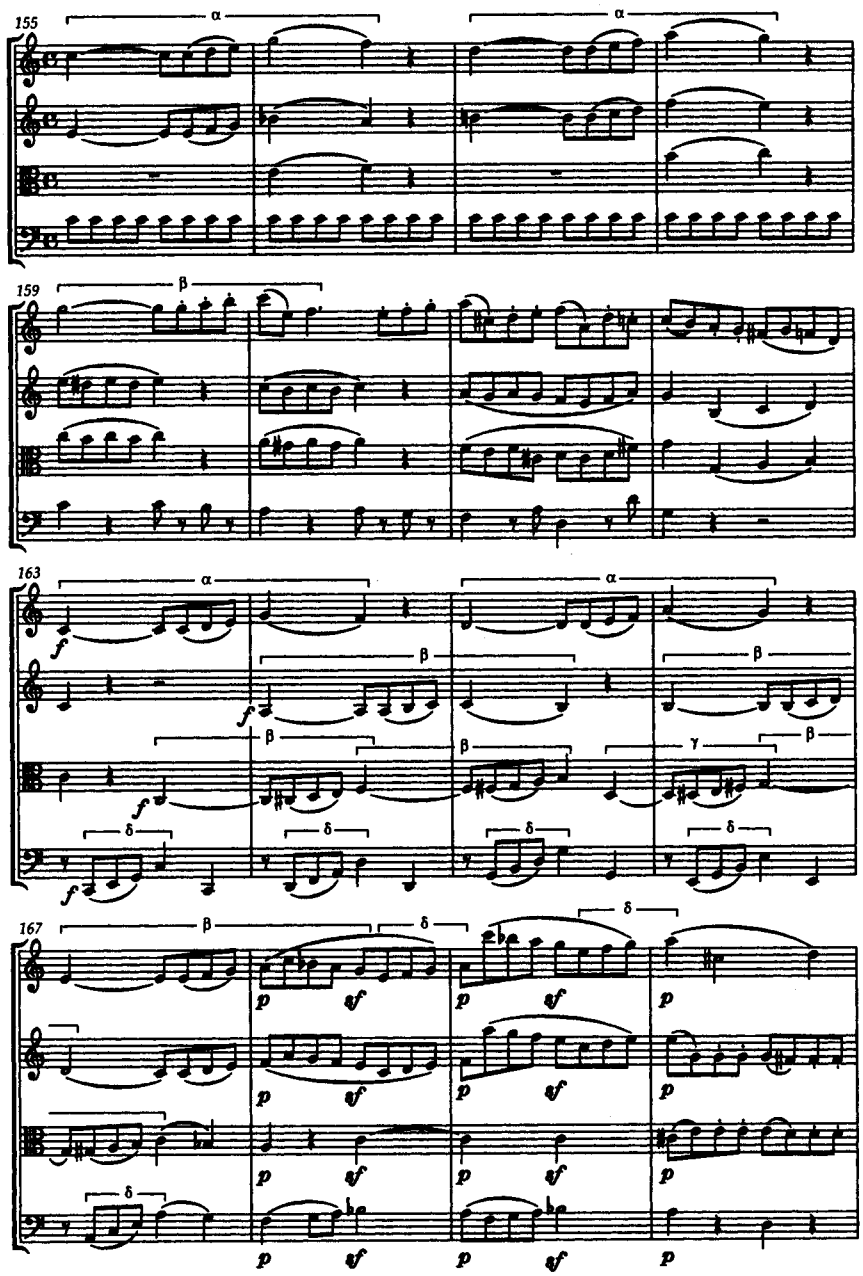
As was the case in the Exposition, in the Development more typical forms of the motive are followed by less typical forms. What is distinctive is the repetition of this pattern in close order (which occurs three times within the forty-odd measures of the Development), and the emergence of the ascending δ figure as the means for Mozart to advance his musical argument. While the slightly varied form of the motive that marks the beginning of each pattern has prominence as a point of reference for this musical argument (and thus represents what is "typical" of thematic material), it lacks the sense of balanced harmonic progression summoned by the original form of the motive. Given the variation of the motive that takes place over the course of mm. 107–35, however, just what *is* typical of the original motive is something of an open question by the time we arrive at the dominant-seventh of mm. 151–55.

### Measures 155–60 and 211–46 of the Allegro

The question of motivic identity posed by the Development is answered by the understated but nonetheless confident return of the original motive in mm. 155–58, shown in example 4.8. Although the accompaniment is somewhat different than that used in mm. 23–26, the motive itself is unchanged from its original form: this is the motive at its most typical. Mozart is not content to leave it at that, however. The accompaniment for the second statement of the paired motives, in mm. 163–66, is not only fuller, but it brings back the overlapped motive forms used at the end of the Exposition and at the beginning of the Development. These slightly varied forms of the original (indicated with β in ex. 4.8) recall the developmental cast of the earlier passages, as do the ascending δ figures in the cello. Mozart nonetheless quickly turns away from any further dissolution of the main motive and instead moves with dispatch toward the first subsidiary theme, which commences at m. 176.

The end of the Recapitulation restates the end of the Exposition almost exactly: mm. 91–106 are transposed into C major for mm. 211–26. From this point, the music proceeds either back to the beginning of the Development (if the repeat is taken) or forward to the Coda (which occupies mm. 227–46). As shown in example 4.9a, the latter grows out of the opening of the Development but uses yet another modification of the original motive. This version owes much to that used in mm. 211–13 (the transposed quotation of mm. 91–93), but applies a chromatic

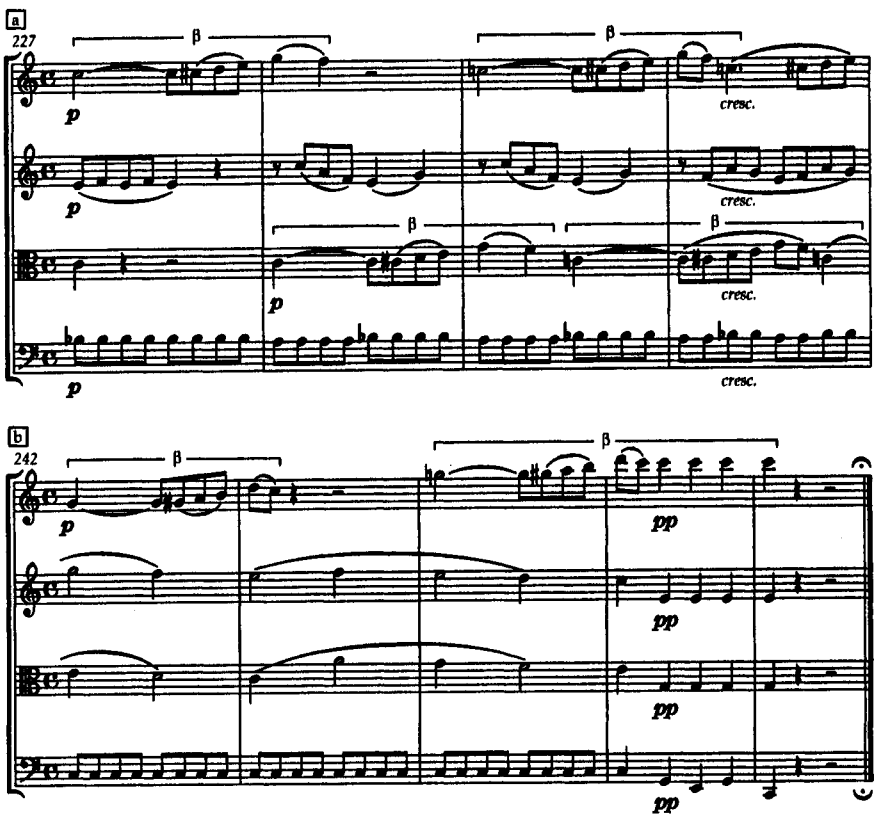
EXAMPLE 4.8 W. A. Mozart, String Quartet K. 465, first movement, mm. 155–70



inflection to the note at the beginning of the ascending-third figure. A further variant then concludes the movement, which is shown in example 4.9b: the end of the motive, shortened from the original half-note–quarter-note figure to the successive quarters of mm. 227–29, is shortened yet again to two successive eighth-notes. Although a similar figure occurs in mm. 229–30, *that* variant refers back to mm. 27–28, for the successive eighths lead immediately to additional material (in the case of mm. 229–30, a repetition of the ascending-third figure). In mm. 242–45, the eighths simply represent a compression of the original motive, and it is with this abbreviated form of the motive that the movement ends.

Although the concluding section recalls the opening of the Development, Mozart allows the prospect of further divergence from the main motive to linger for only a moment before banishing it with a vigorous and triumphant cadence in C major in mm. 234–35. Nonetheless, it cannot be said that the movement ends with an unequivocally typical form of the motive: if the original form of the

EXAMPLE 4.9 W. A. Mozart, String Quartet K. 465, first movement: (a) mm. 227–30; (b) mm. 242–46



motive represents the stable center of the cognitive category of motive forms, the conclusion does not return us to that center. Two factors inform the ending; both are related to global models, which provide a context for the conceptual model that defines what is typical of this motivic category. First, the ending is harmonically quite stable: the final version of the motive introduced in mm. 242–45 reinforces this stability by outlining scale degrees $\hat{5}$ and $\hat{1}$ and by placing the arrival on $\hat{1}$ as close as practicable to the first beat of the metric cycle while retaining the appogiatura figure characteristic of the motive. Second, the most typical version of the motive, in accordance with the global model for music themes relative to which it is framed, represents musical material to be acted upon—it is a theme in the sense of a topic for discourse. At this point in the movement, however, discourse must be concluded: what is required is not a highly typical version of the main motive but a reminiscence of the motive suited to the purposes of ending.

## Compositional Strategy, Categorization, and the Allegro

Although I chose not to emphasize it in the preceding discussion, the overall compositional strategy that shapes the Allegro (as opposed to its principal motive alone) has, not surprisingly, a high degree of conformance both with models for sonata form proposed by music theorists during the late eighteenth century and with the syntactic template for the musical game discussed above. The first main section of the work (mm. 23–106, which I have called the Exposition) moves from tonic to dominant, with themes or thematic groups for each key area. The second main section (mm. 107–204) begins by moving through a number of different key areas (the Development) and then turns to music that emphatically confirms the tonic (the Recapitulation). The syntax of this form is articulated by important cadential arrivals (such as the arrival on the tonic in m. 44 or the arrival on the dominant in m. 91), and although melodic themes are important for projecting a key and rendering a particular key area distinctive, they are not the driving force behind this portion of the musical discourse.

Running parallel to this formal syntax, but independent from it, are the compositional strategies that shape the category comprising the various versions of the principal melodic motive of the Allegro. While the most typical members of this category are associated with tonic (and, in their typicality, serve to emphasize the importance of tonic within the formal syntax), members with a high degree of conformance to the conceptual model for the category are also used to initiate presentations of the motive at points not associated with the tonic (such as the end of the Exposition or beginning of the Development). Each of these presentations of the motive (whether on or off the tonic) leads to an increasing number of less typical members of the category, a process that ends either with the introduction of a more typical form (as in m. 44 or m. 155), or with a move toward new thematic material (as happens following m. 49). Through the accumulating evidence of these processes we come to understand the motive as a basis for Mozart's musical discourse—

something to be introduced, varied, and ultimately reprised. A similar understanding does not develop in connection with the subsidiary themes first introduced in the dominant during the Exposition; instead, these are melodic materials with the highly localized syntactic function typically attributed to such themes by eighteenth-century theorists.

Understanding the principal motive as the basis for Mozart's musical discourse, however, is not the same thing as regarding it as the generator of discourse throughout the movement. As I pointed out in connection with example 4.5, snippets of material reminiscent of some other aspect of the motive can be found elsewhere in the movement.[26] Nonetheless, without the context provided by more typical members of the motivic category, we tend not to hear such fragments as related to the principal motive and thus do not hear them contributing to the discourse of which the motive is a part.

It is important to emphasize that the syntax associated with the principal motive of the Allegro is not entirely unique to this movement. As Robert Gjerdingen has shown, composers of the late eighteenth century relied on a variety of stylistic formulae for both the substance and shape of their melodic and harmonic material, and these formulae inform the compositional strategies applied to the motive.[27] However, the way this particular motivic category unfolds—how many typical members of the category are presented before variation ensues, the relative typicality of subsequent members, and the amount of time before more typical members return—is something specific to the Allegro.

The principal motive of the Allegro should also be understood as an instance of the sort of musical topics to which eighteenth-century composers regularly made recourse (and to which Eco referred in his acknowledgement of musical "signs"). Mozart's motive is a good example of the "singing style" as characterized by Leonard Ratner, and the lyrical presuppositions of this topic do much to shape just what components of the motive are called on for the purposes of variation.[28] Nonetheless, I think Mozart's motive is also more than simply a musical topic, for in its status as a remarkable musical category—remarkable in the variety of members the category comprises, and remarkable in the way the manipulation of these members contributes to Mozart's overall musical argument—it transcends topicality to become an essential element within the musical discourse peculiar to the Allegro of the C major quartet.

26. A thorough analysis that makes an argument for the omnipresence of the principal motive can be found in Bonds, *Wordless Rhetoric*, 102–10.

27. See in particular Robert O. Gjerdingen, *A Classic Turn of Phrase: Music and the Psychology of Convention*, Studies in the Criticism and Theory of Music (Philadelphia: University of Pennsylvania Press, 1988); idem, "Revisiting Meyer's 'Grammatical Simplicity and Relational Richness,'" in *Cognitive Bases of Musical Communication*, ed. Mari Riess Jones and Susan Holleran (Washington, D.C.: American Psychological Association, 1992), 225–43; and idem, "Courtly Behaviors," *Music Perception* 13 (1996): 365–82.

28. For examples and discussion of musical topics, see Leonard G. Ratner, *Classic Music: Expression, Form, and Style* (New York: Schirmer Books, 1980), 9–29. For a discussion of musical topics of this sort within the specific context of musical semiotics, see Victor Kofi Agawu, *Playing with Signs: A Semiotic Interpretation of Classical Music* (Princeton, N.J.: Princeton University Press, 1991), 17–34.

## BEETHOVEN, STRING QUARTET IN B♭ MAJOR, OP. 18 NO. 6. FIRST MOVEMENT

Joseph Kerman devoted the first quarter of his classic study of the Beethoven string quartets to the Op. 18 set (which date from 1798 to 1800), rightly believing that these early works held important keys to understanding how Beethoven transformed what had been a relatively modest genre into the ne plus ultra of chamber music.[29] But Kerman had little patience with the first movement of the sixth quartet, for he found scant evidence of the sort of motivic development that would become the hallmark of Beethoven's mature style.[30] In truth, the opening motive of the work undergoes hardly any transformation over the course of the movement: its syntax is not easily explained in terms of relative degrees of typicality among category members. What is involved instead is a category that includes two distinctly different forms of the motive, forms distinguished not only by the materials from which they are fashioned but also by the way they are used. The different uses of these two forms suggests a higher level of syntax that includes both, and this in turn gives a glimpse into the process by means of which musical topics, of the sort discussed by Ratner and alluded to by Eco, are formed.

Ratner, for his part, characterized this movement as "pure buffa, a dancing cartoon."[31] As can be seen in example 4.10, the two forms of the principal motive contribute much to the effect, providing a musical Tweedledum and Tweedledee to the listener's Alice.[32] The first of these (α in ex. 4.10) sketches a rapid ascent through a broken arpeggio on the tonic and, aided in the first case by a slightly overemphatic *fortepiano*, sets out the first beats of mm. 3 and 5 as points of arrival. The second form of the motive (β in ex. 4.10), suave and somewhat understated, introduces the dominant and is metrically more open. In mm. 7 and 9, it ends on the second main beat of the measure; in its third appearance, it ends on the first beat of the measure, but only after its ending is extended by mm. 12–13. The overall harmonic contour of the β form is also relatively flat: in mm. 5–7, for instance, it begins and ends on B♭2, while m. 6 is preoccupied with a simple alternation of the fifth and root of the dominant chord.

29. With this characterization I mean to take nothing away from the accomplishments in this genre by Haydn or Mozart, but only to suggest that it is with Beethoven that the quartet eventually becomes a site for the active contestation of musical values and that this represents a new possibility for chamber music. See, for instance, Daniel K. L. Chua, *The "Galitzin" Quartets of Beethoven: Opp. 127, 132, 130* (Princeton, N.J.: Princeton University Press, 1995). The issue for musicologists, then, is the extent to which the early quartets prepare the way for this transformation of the genre.

30. Joseph Kerman, *The Beethoven Quartets* (New York: W. W. Norton, 1966), 71–74.

31. Leonard G. Ratner, *The Beethoven String Quartets: Compositional Strategies and Rhetoric* (Stanford, Calif.: Stanford Bookstore, 1995), 89.

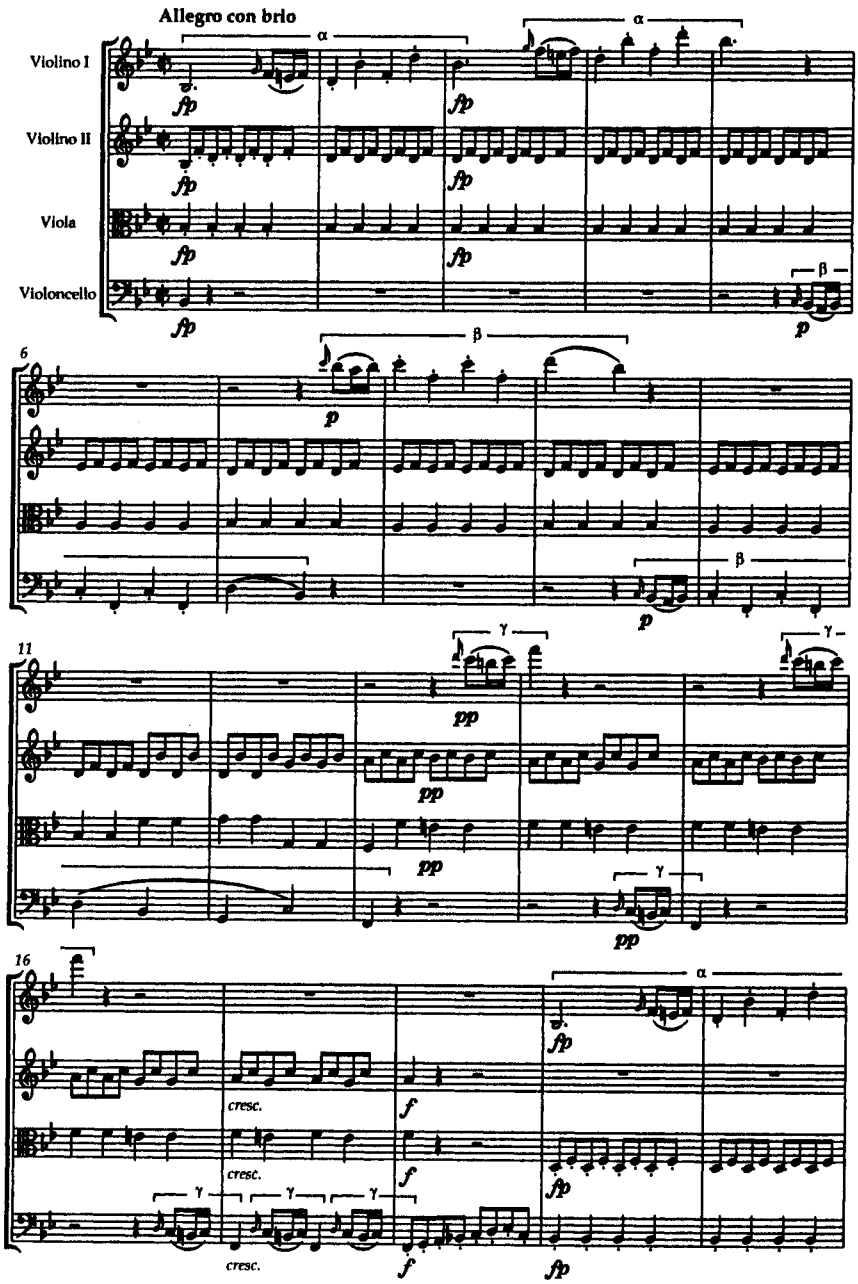
32. In *Through the Looking Glass*, Alice quotes a nursery rhyme about Tweedledum and Tweedledee, and, after hearing this, the two brothers reply:

"I know what you're thinking about," said Tweedledum; "but it isn't so, nohow."

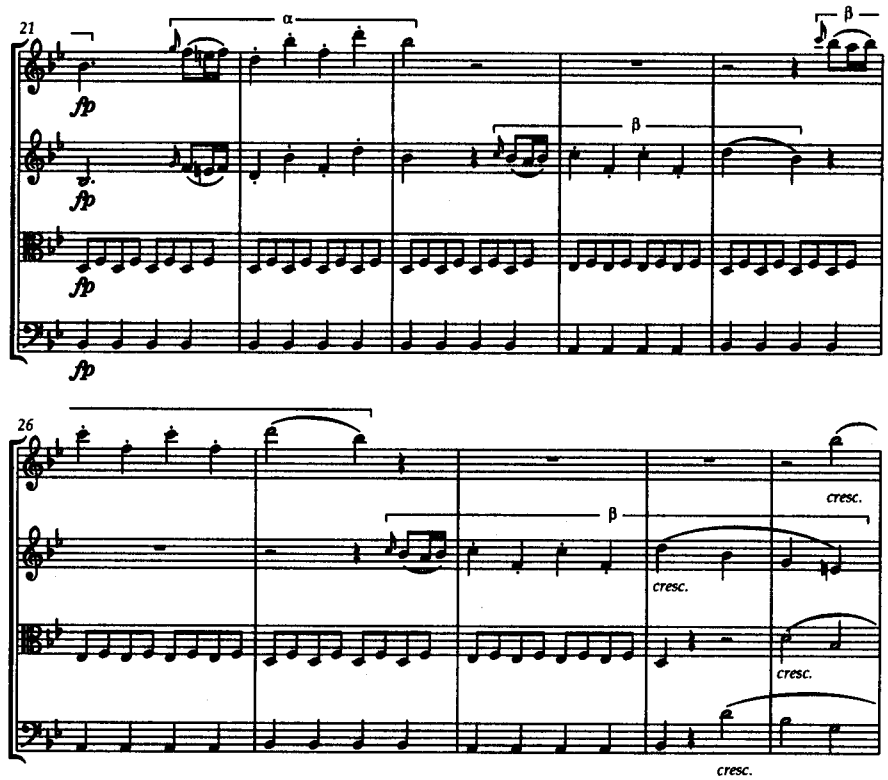
"Contrariwise," continued Tweedledee, "if it was so, it might be; and if it were so, it would be; but as it isn't, it ain't. That's logic."

Lewis Carroll, *The Annotated Alice: Alice's Adventures in Wonderland and Through the Looking Glass*, introduction and notes by Martin Gardner, illustrated by John Tenniel (New York: W. W. Norton, 2000), 181. In a note to this passage, Gardner observes that the nursery rhyme may descend from a bit of doggerel by John Byrom (an eighteenth-century hymn writer and teacher of shorthand), which recounted the rivalry between George Frideric Handel and Giovanni Battista Bononcini.

EXAMPLE 4.10 Ludwig van Beethoven, String Quartet, Op. 18 No. 6, first movement, mm. 1–30



EXAMPLE 4.10 *(continued)*




Commenting on the deliberate and somewhat stagy entrances of α and β is a fragmentary form of the motive (γ in ex. 4.10). This consists of only a turn and a leap (first ascending, then descending); its entrances reverse the expansion of the distance between strong metric accents occasioned by the extension of mm. 12–13, shrinking it first to one-measure and then (in mm. 17–18) half-measure intervals. This process of acceleration culminates in an arrival on the dominant in m. 18, embellished with a run in the cello that returns the music to the tonic and to the α form of the motive.

To better understand the distinctive features of and relationships between these three forms of the motive, as well as their progress through the movement, let us consider conceptual models for each; diagrams for the models are given in figure 4.6. The conceptual model for the category of α forms consists of five correlated conceptual elements, characterized as the things necessary for a melodic fragment to count as an instance of this version of the motive:

- the rhythmic pattern ¢ 𝅗𝅥 𝅗𝅥 𝅗𝅥 𝅗𝅥 | 𝅗𝅥 𝅗𝅥 𝅗𝅥 𝅗𝅥 | 𝅗𝅥.
- the motive begins with a turn figure (here, around $\hat{5}$)
- the continuation of the motive involves a broken arpeggio with ascending contour (here beginning on $\hat{3}$ and ascending through a sixth to $\hat{1}$)

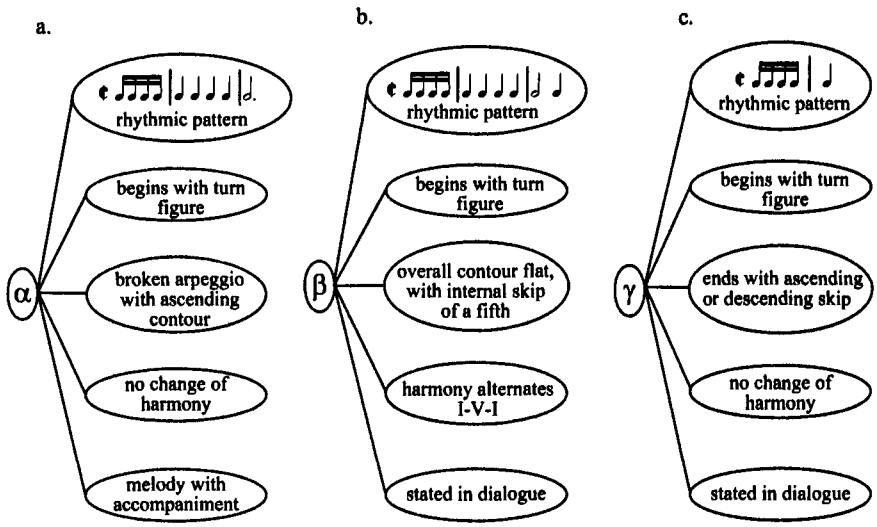



FIGURE 4.6 Conceptual models for the three forms of the principal motive from Beethoven, String Quartet, Op. 18 No. 6, first movement, mm. 1–30: (a) the α form; (b) the β form; (c) the γ form

- there is no change of harmony implied by the motive
- the motive is presented as a melody supported by accompaniment in the remaining strings[33]

Each of these elements could, of course, be broken down into smaller parts. This is particularly true of the rhythmic pattern, which is extended by the β form of the motive and truncated by the γ form. As indicated by figure 4.6b, the conceptual model for the β form retains most of the first two components but alters the last three. The γ form continues this process, retaining only the turn figure that is the motive's most characteristic feature.

Slightly varied restatements of the α and β forms of the motive in mm. 19–29—the texture thickened somewhat, and with the two violins taking part in the dialogue proper to the β form of the motive—lead to a long pedal on the dominant of F major preparatory to the statement of the second theme, which commences in m. 45. What is remarkable, from the perspective on musical syntax I have developed thus far, is the nearly complete lack of typicality effects among these materials. While it would be quite possible to include them all in a single category for the motive, framed relative to a conceptual model that combines the essential elements of those included in figure 4.6, there does not seem to be a compelling argument for regarding either the α and β form as more typical of the category. A better case could be made for conceiving of the γ form as a less typical member of the category (if only on account of its fragmentary nature), but it is a bit hard to say just

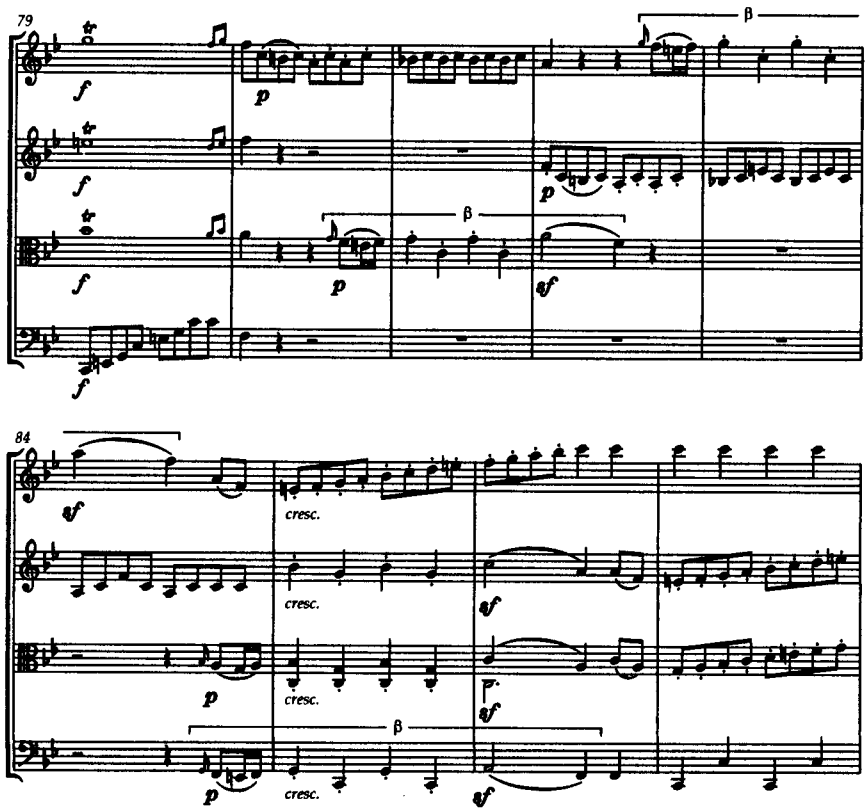
33. There are, of course, other ways the motive could be characterized. For instance, the rhythmic figure I identify does not include the very first note of the piece. My chief reason for seizing upon this figure is for comparison with other forms of the motive. Should one wish to include the first note, the remaining forms of the motive would represent even more radical departures from the standard established by the α form.

what it is less typical of. It could be derived from either the α or the β form, and without further variants it remains somewhat difficult to settle the question.

Subsequent to a cadence on F major in m. 80 that wraps up the exposition of the second theme, the β form of the motive reappears, as shown in example 4.11. Although it has been transposed to F, and its third statement in the cello is accompanied by parallel tenths in the viola and by a reprise of the passagework from the bridge in the first violin, it is otherwise exactly the same as when it appeared in the context of the first theme group. This restatement has two consequences. First, it confirms the character of the β form of the motive that was established in the opening measures—that is, it reinforces the conceptual model outlined in figure 4.6b. Second, it draws out a function of β implicit in the initial exposition of motive forms, which was to close off musical discourse. Where the α form of the motive sets out the conditions for musical discourse—establishing a two-measure hypermeter and opening the melodic range to cover more than two octaves—the β motive, with its pedestrian alternation of tonic-dominant-tonic and restricted range, provided closure. Following m. 80, the motive has a similar function, but now it serves to close off the discourse of the entire Exposition.

The syntactic functions of the α and β forms of the motive are due in part to

EXAMPLE 4.11 Ludwig van Beethoven, String Quartet, Op. 18 No. 6, first movement, mm. 79–87



their resemblance to similar figures in other compositions: α has all the features of a prototypical opening gesture within eighteenth-century style, and β has much in common with prototypical closing gestures. It should be emphasized, however, that how such motives are used is also important. As Janet Levy has pointed out, Joseph Haydn, at least on occasion, used closing gestures as his opening material.[34] The use of a closing gesture in this way is in part comic—it creates a piece that seems to end before it has begun—but it also makes possible some marvelously ambiguous seams: one is not sure whether a particular passage is a beginning or an ending. The possibility of this ambiguity testifies to the importance of the way musical materials are used within a particular composition, as distinct from how similar materials are used in other compositions. For a gesture to signify "opening," it has to be *used* in openings; for a gesture to signify "closing," it has to be *used* for closings.

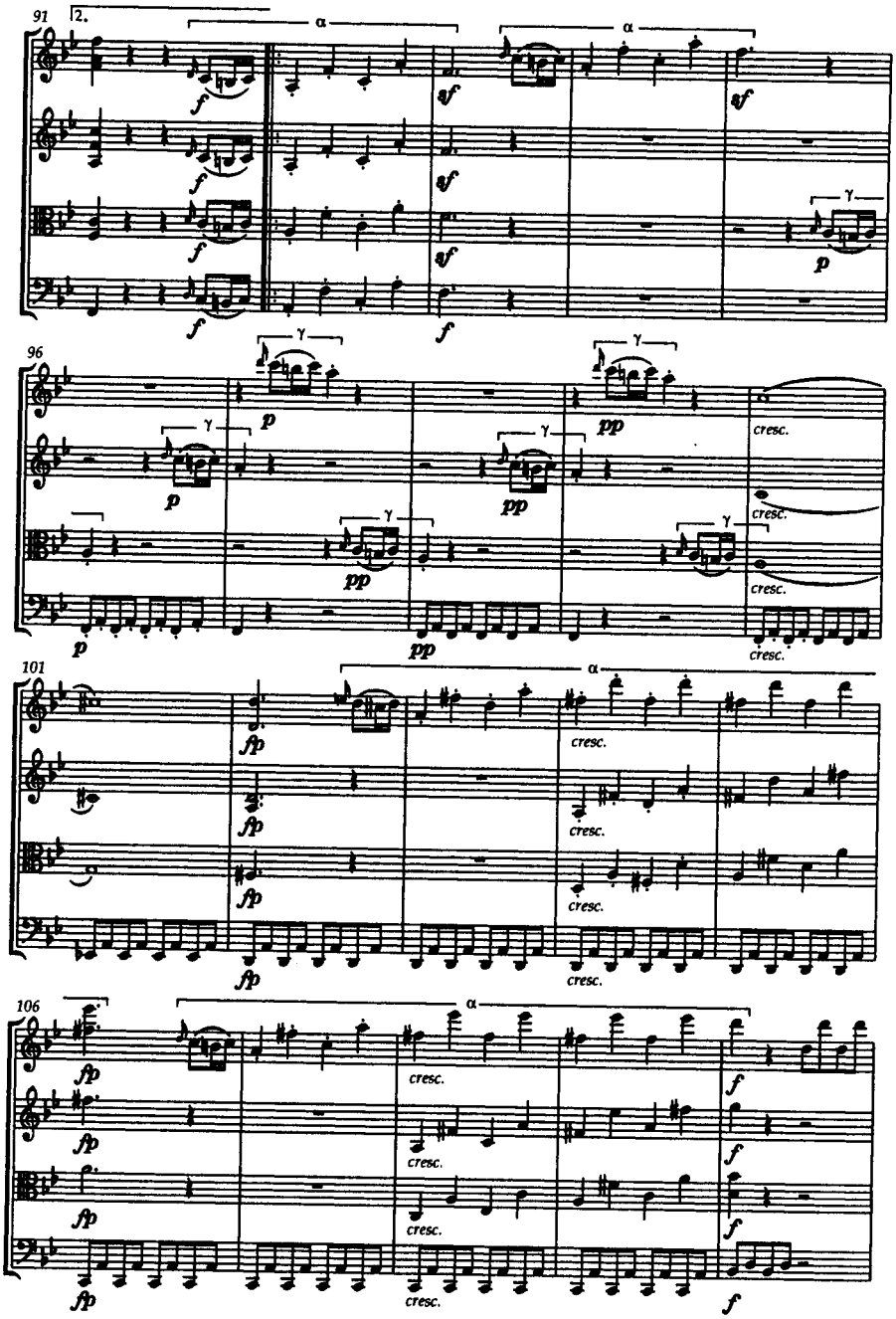
The α form of the motive's function as an opening gesture is confirmed when it initiates the Development, which begins with a tutti statement of the motive, immediately echoed by a solo statement in the first violin (as shown in ex. 4.12). These slightly less than typical versions of the motive (which has heretofore always appeared as melody with accompaniment) are followed immediately by the γ form, which tosses a repetition of the head of the α form just heard back and forth between the upper strings. In mm. 103–110, two further variants of the α form appear, which extend the motive both metrically (to four measures) and intervallically (to a compass of an eleventh, and then a twelfth). These expansions lead to G minor, which initiates a rapid cycle of keys explored through the medium of the thematic material that first served as a bridge between B♭ and F back in mm. 33–43.

In m. 139, there is a weak cadence on F, followed immediately by a reappearance of the β form of the motive (shown in ex. 4.13). Stated, as it is, in F major, it recalls the end of the Exposition; however, this particular version does not conform completely to the conceptual model diagrammed in figure 4.6b. The rhythmic profile has been changed and is now the same as that of the α form; the dialogue presentation typical of β forms is absent, replaced here by one unaccompanied and one accompanied statement. Nonetheless, although close to thirty measures of the Development remain, this entry of the β form of the motive signals the end of the harmonic exploration initiated by the variant α forms in mm. 103–10: F major having been regained, the remainder of the Development simply serves to postpone the inevitable return to the opening material. Just as it has each time before, the β form of the motive confirms a closing of harmonic structure.

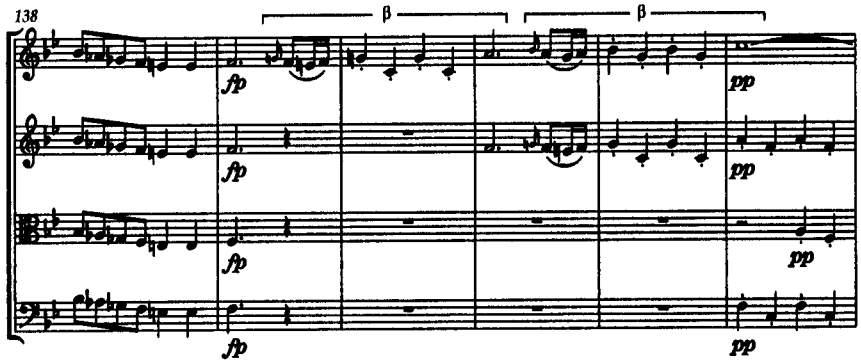
With the return to B♭ in m. 175, there is an exact reprise of the first eighteen measures of the movement, after which the γ form of the motive is presented in a series of close-order repetitions that initiate the bridge to the second theme. At the conclusion of the movement, the β form returns once more, as it did in mm. 80–86, only this time confirming a final close on B♭.

34. Janet Levy, "Gesture, Form, and Syntax in Haydn's Music," in *Haydn Studies: Proceedings of the International Haydn Conference, Washington, D.C., 1975*, ed. Jens Peter Larsen, Howard Serwer, and James Webster (New York: W. W. Norton, 1981), 355–62.

EXAMPLE 4.12 Ludwig van Beethoven, String Quartet, Op. 18 No. 6, first movement, mm. 91–110



EXAMPLE 4.13 Ludwig van Beethoven, String Quartet, Op. 18 No. 6, first movement, mm. 138–43




As Kerman noted, there is little development of the principal motive in the first movement of Beethoven's Quartet, Op. 18 No. 6. This is not to say that the motive is uninvolved in the syntax of the movement, however, only that the musical discourse is not, in any significant way, about this motive. Thus the α form of the motive signals an opening for musical discourse, but it is not discourse about α. Even when the α form makes an attempt to grab center stage at the beginning of the Development, the effort ends up being little more than comedic bluster: knowing only how to begin, it cannot quite figure out how to continue, and after two attempts (in mm. 102–10), it gives up. Similarly, the β form of the motive closes off discourse, but it is not always a discourse in which β has been involved (as in its appearances at the end of the Exposition and Recapitulation, and in the Development). Thus the two motive forms, by virtue of the way they are used, contribute to a larger syntax concerned not with them but with the progress of the movement as a whole.

Despite their similarities, I believe these two motive forms are distinct enough in their features to be regarded as separate, though related, categories (something illustrated by the conceptual models of fig. 4.6). The syntactic function that accrues to each over the course of the movement—with the α form used only for openings, and the β form used only for closings—gives further support to this distinction. Nonetheless, I regard distinguishing the motive forms on the basis of syntactic function to belong to a later and more theoretical stage of understanding. Accordingly, categorizing the motives is relatively immediate, while understanding the way they function within the work emerges only over time.

As I have noted, these two forms of the motive bear resemblance to stylized musical gestures that signal openings and closings. However, I would argue that, at least on the abstract level concerning processes that pertain to musical categories, their function within the movement does not derive from these gestures. To be blunt, I do not believe knowledge of such stylized figures is necessary to understand Beethoven's musical argument. That it would enrich this understanding I have no doubt, but this is not the same as asserting such knowledge as a precondition for making sense of the discourse of the movement. Instead, the function of the α and

β forms of the motive is an emergent property consequent to the way they are constituted (as contrasting yet similar motivic structures) and the way they are used within this movement.

The notion that Beethoven's use of these motives bestows particular functions on them further suggests how musical topics come to signify the things they do. Granted, the musical material for any topic needs to meet certain prerequisites of harmony and counterpoint and conform to general expectations about affect. But beyond this, one must acknowledge the significance of the way composers have used particular musical gestures—that is, their prevalence within a certain historical and cultural milieu—in establishing the things these gestures come to signify. Put another way, signification is not an inherent property of musical materials but derives from how they are deployed in particular works of music—that is, from the syntactic function with which they are invested.

## BEETHOVEN, STRING QUARTET IN F MAJOR, OP. 18 NO. 1, FIRST MOVEMENT

Where the discourse of the first movement of Beethoven's B♭ major quartet tends not to be about its opening motive, the discourse of the first movement of the F major quartet has difficulty being about anything else: as A. B. Marx famously observed, the opening motive occurs no fewer than 131 times within the 427 measures of the movement.[35] Within this multiplicity, two categories of motive forms can be distinguished, differentiated both by their constituent features and by the compositional strategy that is applied to them. The distinction between these two categories becomes especially plain in the Development, when a new and highly dynamic variant of the motive appears that combines features of both. The importance of this variant for developing Beethoven's musical argument is affirmed in the Coda, which revisits all three forms of the motive.

The compositional strategy associated with these various motive forms is relatively complicated, and it was not completely successful in Beethoven's first attempt at the movement, a version described in great detail in a thorough study by Janet Levy.[36] In what follows, I discuss how this strategy is manifested in the Exposition, Development, and Coda of the second (and final) version of the F major quartet, and then turn to Beethoven's first version to explore how it was brought to its final form.

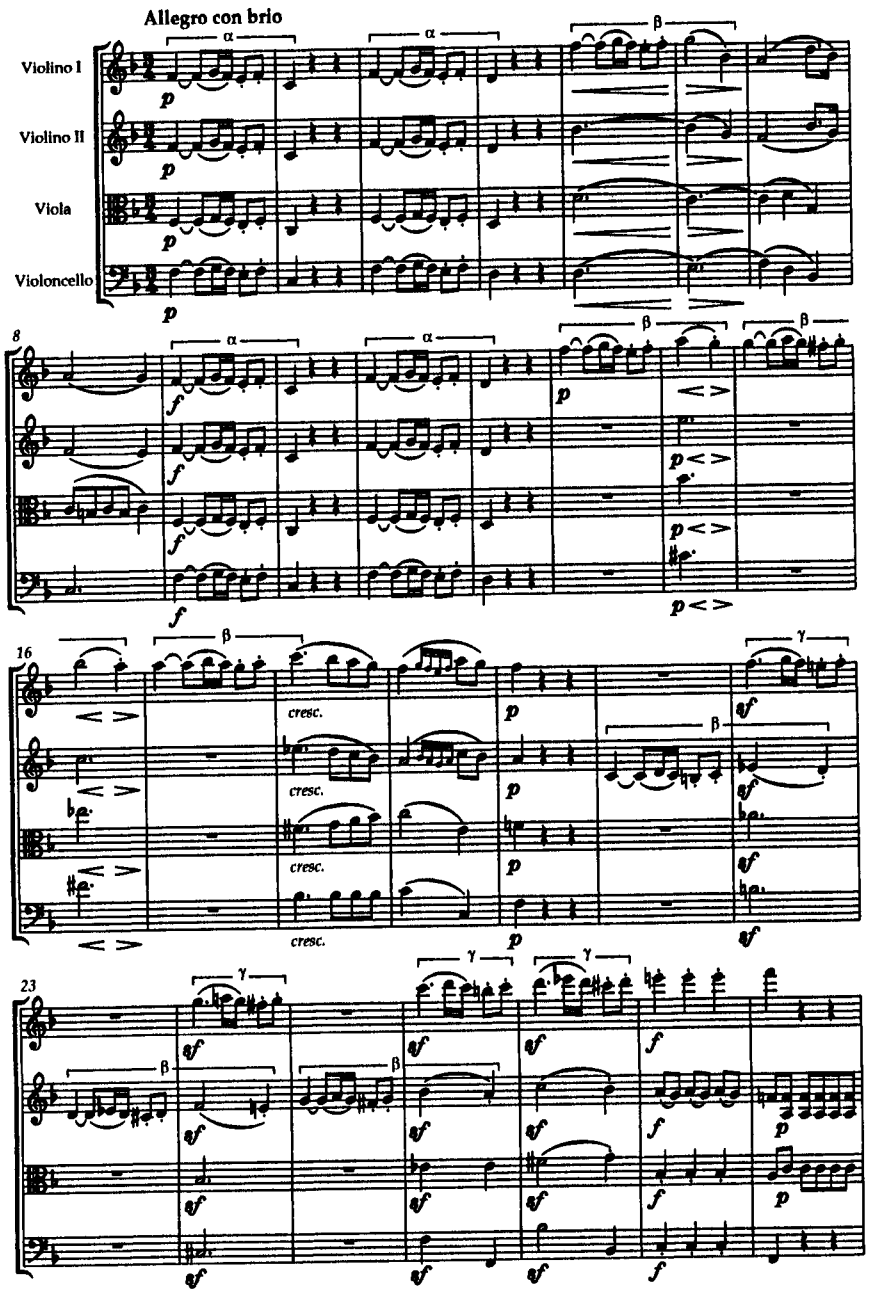
### The Exposition

In a manner similar to the first movement of the B♭ major quartet, the F major quartet opens with a principal motive that takes two distinctly different forms. The first (α in ex. 4.14) is assertive, direct, and ends on the first beat of the measure. The second (β in ex. 4.14) is voiced in a lighter manner, is harmonically fluid, and ends

35. Adolph Bernhard Marx, *Ludwig van Beethoven: Leben und Schaffen*, 6th ed., ed. Gustav Behncke (Berlin: Otto Janke, 1908), 203.

36. Janet M. Levy, *Beethoven's Compositional Choices: The Two Versions of Opus 18, no. 1, First Movement*, Studies in the Criticism and Theory of Music (Philadelphia: University of Pennsylvania Press, 1982).

EXAMPLE 4.14 Ludwig van Beethoven, String Quartet, Op. 18 No. 1, first movement, mm. 1–29



on the third beat of the measure. Joining these is a fragmentary form (γ in ex. 4.14) whose character is less distinctive, and which plays a much more subsidiary role in the Exposition. This fragment consists of little more than the stylized turn figure common to all forms of the motive (and which was almost certainly the basis for Marx's impressive tally). It is this feature of commonality that, as Ratner has noted, allows the motive to fit seamlessly into a variety of different contexts.[37]

As with the preceding analyses, we can characterize the distinctive features of and relationships between various motive forms by considering conceptual models for each; diagrams for the models are given in figure 4.7. The conceptual model for the category of α forms consists of five correlated conceptual elements, characterized as the things necessary for a melodic fragment to count as an instance of this version of the motive:

- the rhythmic pattern
- the first part of the motive outlines a stylized turn figure
- the motive ends with a small downward skip
- the motive is stated tutti
- the motive has a relatively static harmonic context

Save for the change in dynamics from the first to the second pair of statements—*piano* replaced by *forte*—all four of the statements of this form of the motive in the opening measures conform to this model. Put another way, the members of the category presented in mm. 1–4 and 9–12 assign a relatively restricted range of values to the attributes picked out by the conceptual model.

Where the members of the α motivic category are all quite similar to one another, the members of the β category are rather more diverse. In mm. 13–18, the

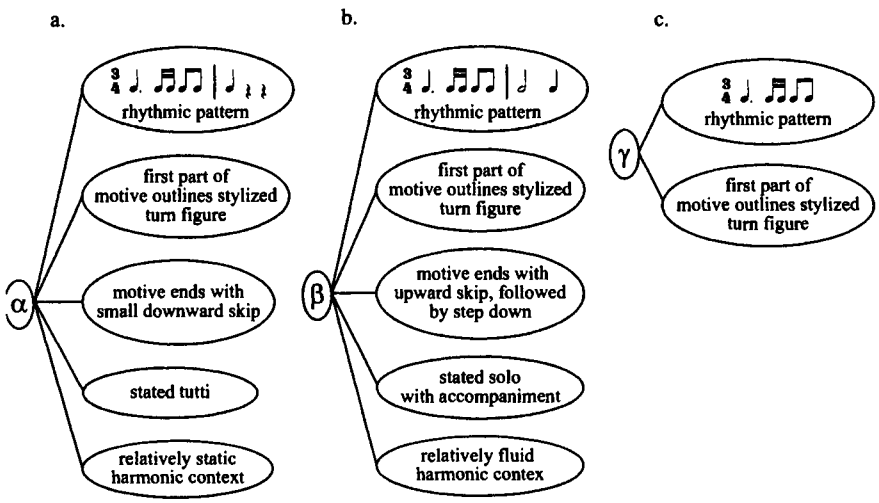



FIGURE 4.7 Conceptual models for the three forms of the principal motive from Beethoven, String Quartet, Op. 18 No. 1, first movement, mm. 1–29: (a) the α form; (b) the β form; (c) the γ form

37. Ratner, *The Beethoven String Quartets*, 12–13.

first measure of each statement of the motive is unaccompanied, an appogiatura over a fully-diminished seventh chord is introduced (in mm. 14 and 16), and the end of the last statement (in mm. 17–18) has a slightly modified rhythmic profile.[38] Versions of the motive that appear in mm. 21–26 are similar to those of mm. 13–18, but include echoes provided by the γ fragment as part of their accompaniment. The category of β motives is, as a consequence, much more diverse: category members represent a relatively wide range of attribute-values.

In the remainder of the Exposition, not only are these two categories of motives maintained but also Beethoven continues to treat them as he did in laying out the first key area. The α motives (for the most part) continue to stay close to the conceptual model around which their motivic category is structured, and the β motives (for the most part) continue to diverge in various ways from the conceptual model around which their motivic category is structured. The Exposition thus reinforces both the structure of the two categories and the compositional strategies through which they were first presented.

After the opening thematic group, the α form of the motive is prominent in four places, putting in appearances twice in the transition to C major: once in the middle of the second thematic group, and preparatory to the cadences that conclude the exposition. In each case, one of two relatively minor variants is used.

The first variant, which initiates the transition that begins in m. 30 (as shown in ex. 4.15a), differs from the model in that it is stated solo against accompaniment and in alternation with a competing figure in the first violin and then concludes with a skip down an octave. From a harmonic perspective, this variant is more static than the original α form in that it prolongs but a single pitch-class (here, C) through the combination of the turn figure and an octave skip; Ratner has noted that, in this form, the motive evokes the sound of a musette.[39]

The second variant of the α form of the motive, introduced by the viola in mm. 41–42, is similar to the first, but it concludes with a repeated note instead of an octave skip (as shown in ex. 4.15b): rather than a single pitch-class, now only a single pitch (E♭3) is prolonged. Appearing against the α form is a β form (in the first violin), whose features are discussed below.

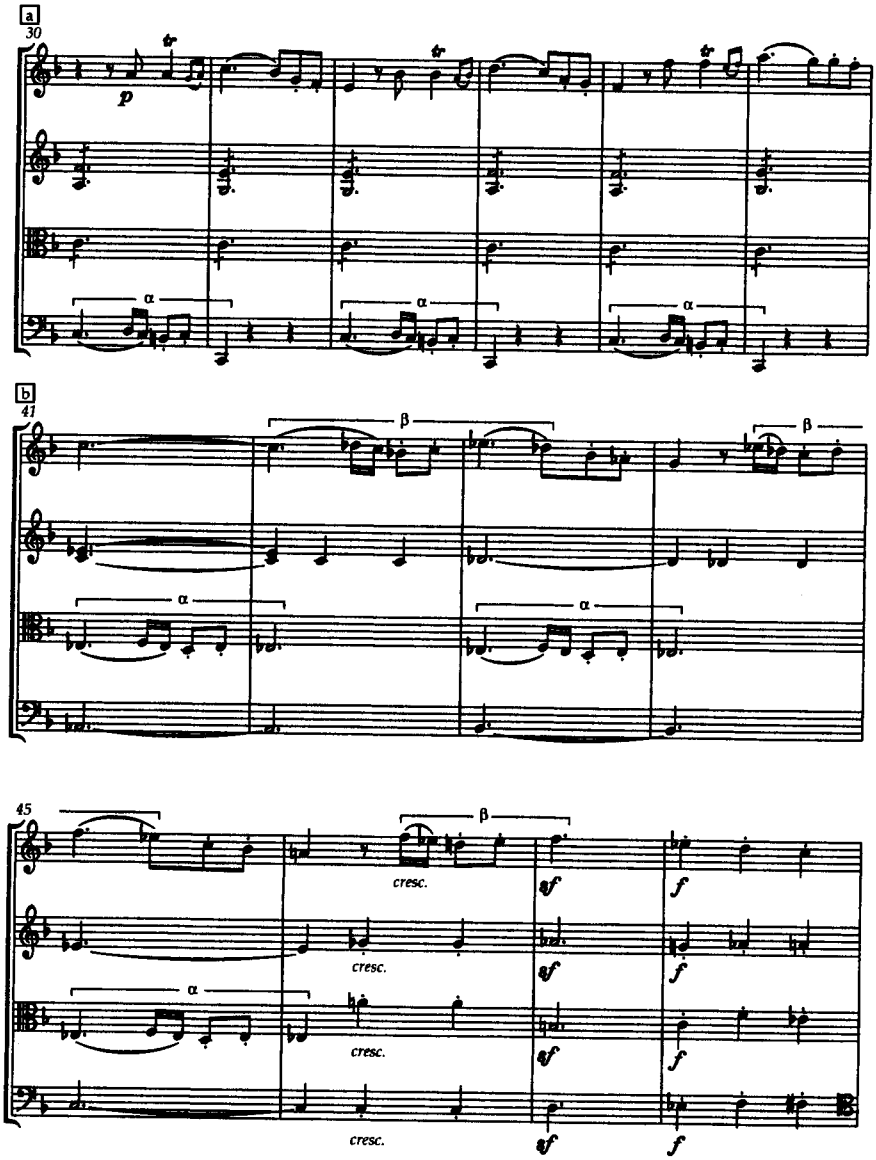
This same variant of α is used by the cello in mm. 72–77 in the midst of the second theme group (see ex. 4.15c).[40] An interesting complication here is that the figure with which this appearance of the variant competes (stated in the first violin) can easily be heard as a further variant of the α form of the motive: admittedly, only the rhythmic profile and the beginning of the turn figure are retained (and, unlike other α variants, this one is sequenced upward), but placed against the second α variant the similarities are striking.

38. In the first version of the quartet, the initial statement of the β form of the motive (in mm. 5–6) was also unaccompanied. The difference in texture, however, is less striking than the clarification of harmonic context provided by the accompaniment Beethoven added, which helped emphasize the contrast between the α and β forms.

39. Ratner, *The Beethoven String Quartets*, 14.

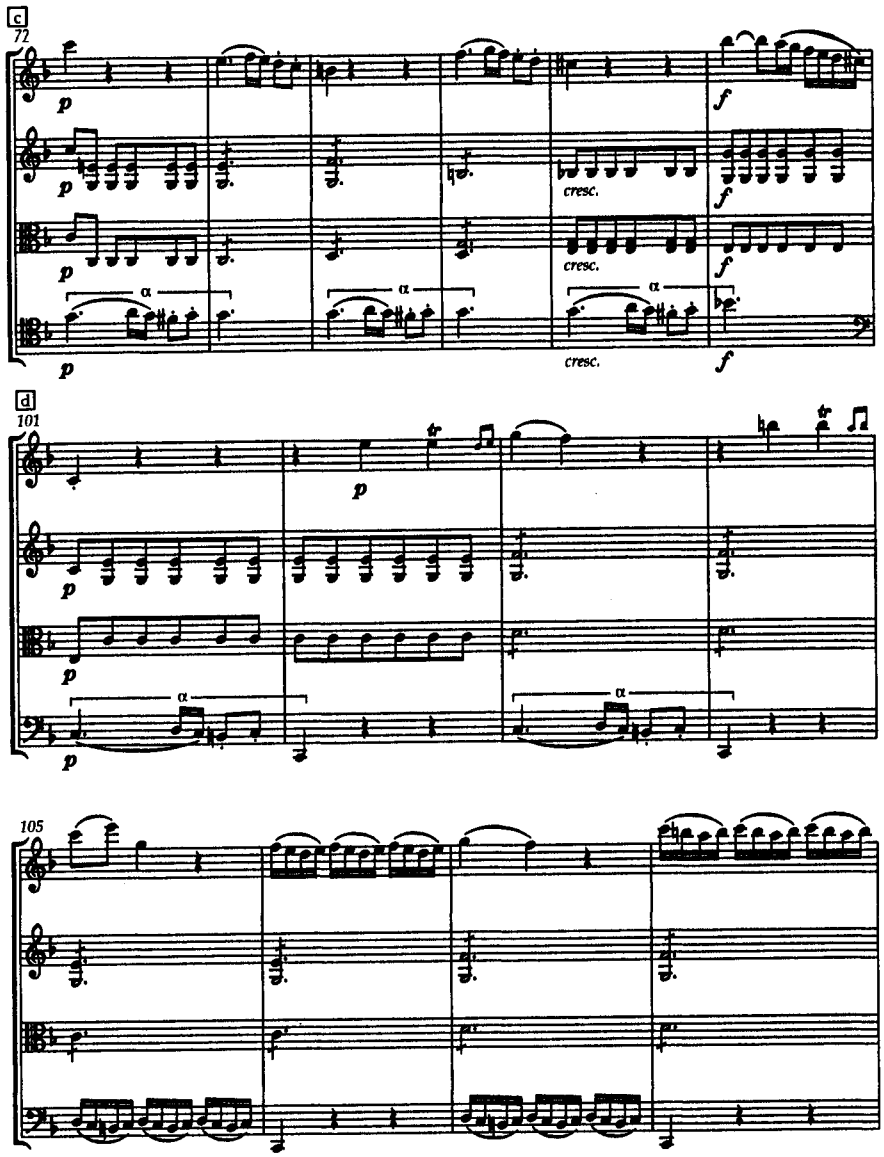
40. Note that the cello appears here an octave lower than in most published scores. This reading of the passage reflects that of the new *Beethoven Werke*.

EXAMPLE 4.15 Ludwig van Beethoven, String Quartet, Op. 18 No. 1, first movement, variants of the principal motive: (a) mm. 30–35; (b) mm. 41–48; (c) mm. 72–77; (d) mm. 101–8




*(continued)*

EXAMPLE 4.15 *(continued)*




Heard as a variant of the α form of the motive, it offers the first real challenge to the standard of typicality for the motivic category.

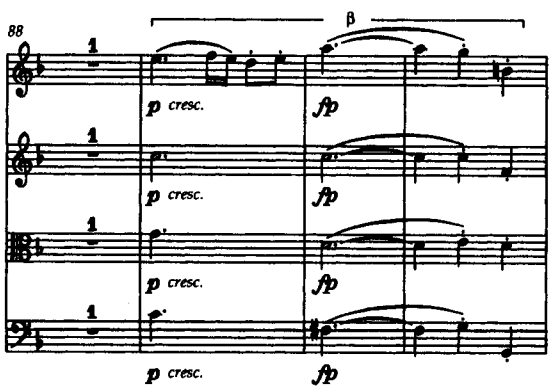
The first variant of α returns in the cello in mm. 101–104 (shown in ex. 4.15d) against a figure in the first violin that can be related to the β form of the motive; this passage is discussed in a bit more detail below.

In contrast to the ubiquity of the α form of the motive, the β form makes only two appearances after m. 30. The first, already mentioned above, is in m. 42 (see ex. 4.15b). Here the motive appears with an altered rhythmic profile (lacking its first note and using the variant ending first introduced in m. 18), and, in its role as counterstatement to the modified α form in the viola, with less independence than before. The third entrance of the variant (in m. 46) departs further still, with its ending a durational augmentation of that used in the preceding measures. A similar augmentation of the tail of the β form is heard in mm. 89–91 (as shown in ex. 4.16) after a general pause that interrupts the material from the second thematic group. Save for this augmentation, this variant of the β form of the motive is quite close to the very first version of β, heard in mm. 5–6.

The variant heard in mm. 89–91 is the last explicit reference to the β form of the motive before the conclusion of the Exposition. Nonetheless, the appogiatura figure that came to be a feature of β beginning in m. 14 figures quite prominently in the counterstatement to the final appearance of the α form in mm. 101–104 (shown in ex. 4.15d). When this counterstatement is embellished with a repeated turn figure in m. 106 (answering the embellishment of the α variant in m. 105), the connection becomes somewhat more plausible: the counterstatement could be heard as a further variant of the β form of the motive. This interpretation is supported by two things: first, there is the close relationship between the α and β forms of the motive throughout the Exposition (exemplified by their juxtaposition in mm. 41–48). Second, the turn figure typical of all forms of the motive is replicated, in compressed form, by the combination of the trill and termination on the third beat of m. 102 and the appogiatura of m. 103 (and further varied in mm. 104–105).

The possibility of extending the category of β motives to include the counterstatement of mm. 102–105, together with the connection between the category of α motives and various counterstatements (in mm. 41–48, 72–77, and 101–108), suggests a way our concept of the β form of the motive might slip further still. If we return to the transition that begins in m. 30 (ex. 4.15a), we note a similarity between

EXAMPLE 4.16 Ludwig van Beethoven, String Quartet, Op. 18 No. 1, first movement, mm. 88–91



the accompanimental figure in the first violin and the β variant from mm. 17–18: the rhythmic pattern and contour of mm. 31, 33, and 35 is the same as that of m. 18. Our return to these measures becomes literal rather than imaginary if the Exposition is repeated: hearing mm. 30–35 after the several pairings of the α form of the motive with counterstatements and the extensions of the β form that occur during the first pass through the Exposition could provide enough pressure to cause significant slippage within our concept of the β form. This, in turn, might lead us to embrace the accompanimental figure of mm. 30–35 within that form when we encounter it for the second time.

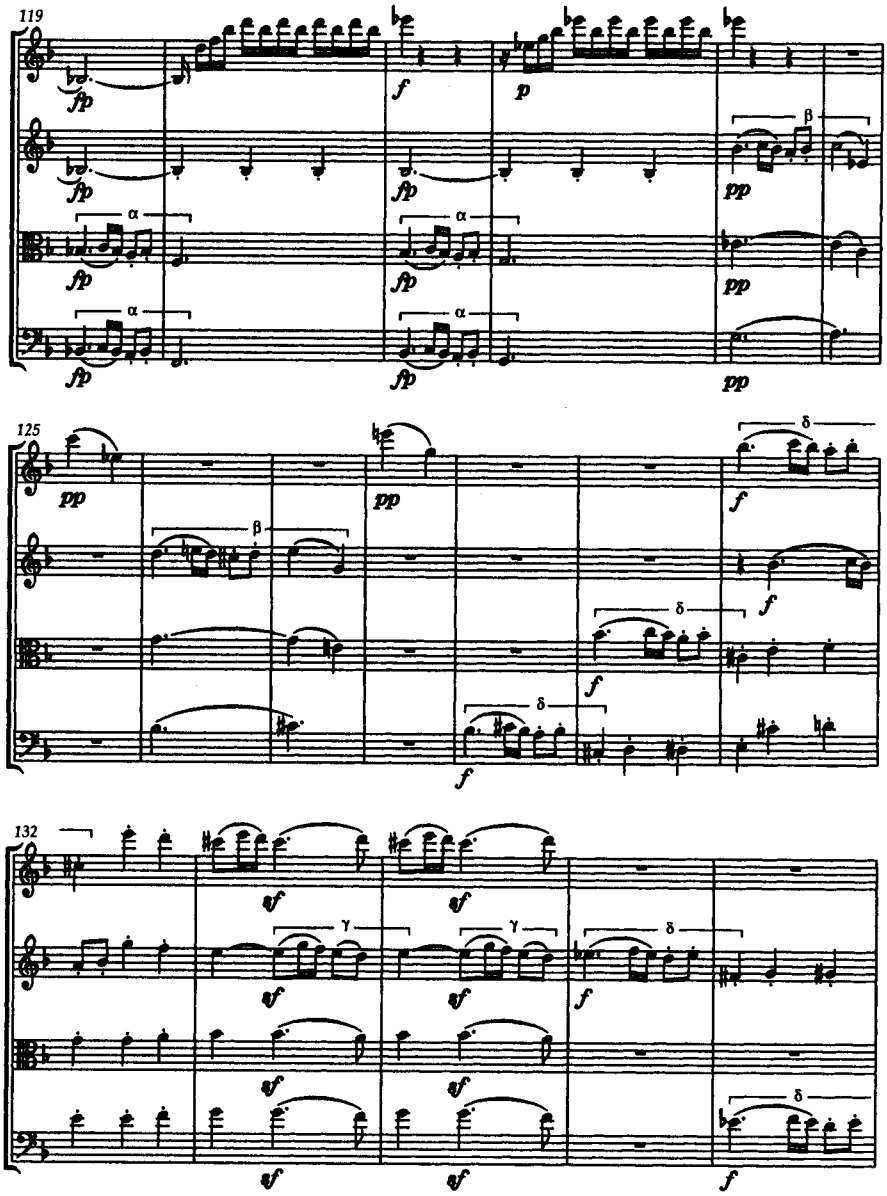
Perhaps more remarkable than the conceptual fluidity of the β form of the motive is the resistance of the α motive to change: despite numerous juxtapositions with rival materials, it remains constant. Indeed, these two categories of motive forms can be further distinguished by the compositional strategies associated with each. Over the course of the Exposition, α forms are subjected to relatively little variation, and most members of the category are close to the conceptual model outlined in figure 4.7a. In contrast, β forms are varied almost from the outset. While the conceptual model for the category of β forms still makes sense—it captures the salient features of members of the category as presented in the opening measures and serves to distinguish them from members of the category of α motives—the range of values that members assign to the attributes picked out by the conceptual model is much broader, and a given variant is rarely repeated. The difference between the compositional strategies applied to each category of motive forms lends a distinctive character to them: α as stable in the face of various challenges; β as an agent of change and transition. Although qualities such as these are part of the elemental materials of sonata form, as they are manifested in the first movement of the F major quartet they also pose a kind of problem. The sort of fluidity with which the β form of the motive is associated operates principally in contrast to the stability of the α form. By itself, it is not quite sufficient to create the dynamic energy—the sparks of real transformation—that the Development section will require. For this sort of drama, a new form of the motive is required, but one that draws on the musical materials and syntax developed over the course of the Exposition.

## The Development

After a brief A major flourish, the Development revisits α and β, recalling the opening of the Exposition (as shown in ex. 4.17; the example does not include the A major flourish). While each motive departs slightly from its respective conceptual model—the α motive includes light accompaniment in the upper strings, and the β motive acquires an echo—for the most part the conceptual model for each category is reaffirmed. Beethoven then introduces a new form of the motive that reconfigures the properties of α and β and transforms the possibilities for musical discourse.

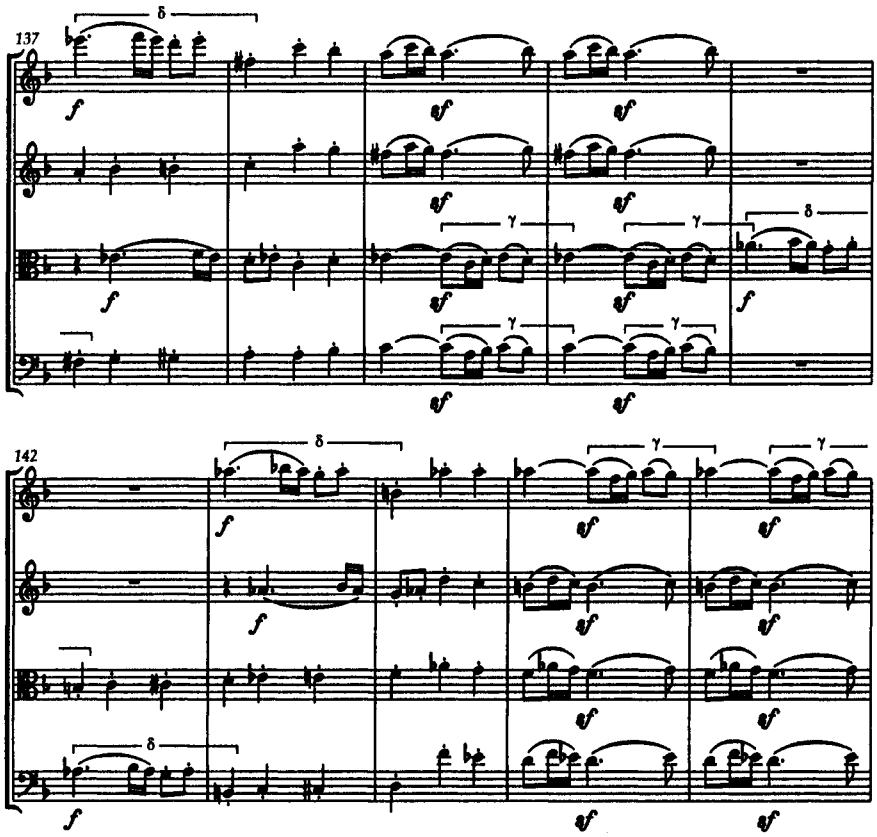
The new motive form (shown with a δ in ex. 4.17) is launched in m. 129 through a series of imitative entrances. Although motive forms have appeared at one-measure intervals before, here they differ only in the octave in which they are stated. This manner of presentation makes reference to a topic to which Beethoven often made recourse in the Development sections of his instrumental works: the

EXAMPLE 4.17 Ludwig van Beethoven, String Quartet, Op. 18 No. 1, first movement, mm. 119–46




*(continued)*

EXAMPLE 4.17 *(continued)*




learned style that includes fugue and canon.[41] This new motive form resembles a member of the category of α motives. It is new in its isolation, harmonic stasis, and clear assertion of the downbeat of each measure as a point of initiation (rather than a point of articulation within a larger gesture, as it is in the second measure of a typical member of the β category of motives). However, the dissonant interval formed by the downward leap consequent to the turn figure (a diminished seventh—B♭3 to C♯3 in mm. 129–30) demands continuation beyond the downbeat, a fluidity reminiscent of the β form of the motive.[42] This amalgam of features and the way they dominate the measures that follow suggest that δ constitutes a new category of

41. For a discussion of the learned style, see Ratner, *Classic Music*, 23–24.

42. Given both the emphasis of the large downward leap and the fact that the leap is dissonant, the δ form of the motive is more ambiguous at the hypermetric level than either the α or β forms: where the previous forms of the motive can be heard in terms of pairs of strong and weak measures, such a grouping in the case of the δ motive is much more arbitrary. This is not to say that determining which measure of the α and β forms of the motive is strong or weak at any given point is a simple matter, however. Janet Levy and Roger Sessions, for instance, give interpretations of the accentuation of mm. 1–2 that are the exact opposite of one another. See Levy, *Beethoven's Compositional Choices*, p. 10, n. 28; and Roger Sessions, *The Musical Experience of Composer, Performer, Listener* (Princeton, N.J.: Princeton University Press, 1950), 13–14.

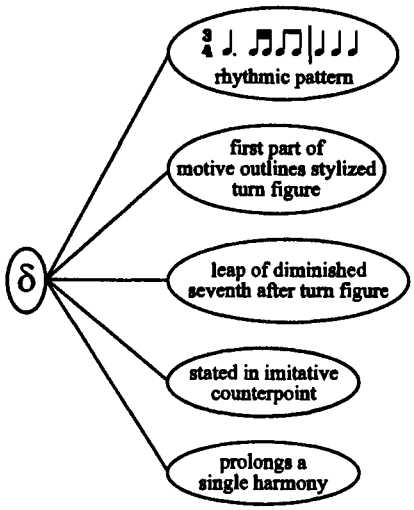



FIGURE 4.8 Conceptual model for the δ form of the principal motive from Beethoven, String Quartet, Op. 18 No. 1, first movement, mm. 129–44

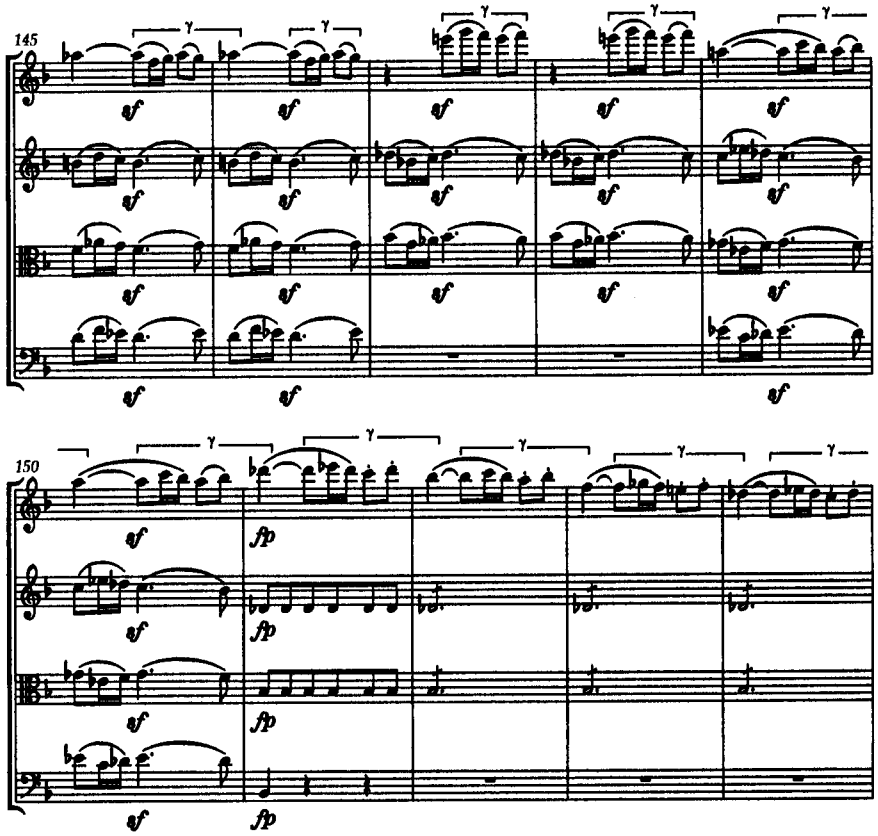
motive forms. The conceptual model for this category (diagrammed in fig. 4.8) consists of five correlated conceptual elements, characterized as the things necessary for a melodic fragment to count as an instance of this version of the motive. While this model has much in common with those diagrammed in figure 4.7, it also demonstrates enough differences to warrant its being used to distinguish the δ form from any of the other forms of the motive important in the movement.

Beethoven follows mm. 129–34 with two similar sections, the first beginning in m. 135, the second in m. 141. Each section begins with a slightly different pattern of entrances among the instruments, and each section prolongs a single fully diminished seventh chord: C♯$^{o7}$ (mm. 129–34), F♯$^{o7}$ (mm. 135–40), and B$^{o7}$ (mm. 141–46).[43] Each section also includes a variant of the γ fragment, based on the turn common to all forms of the motive but with a *sforzando* on the second beat: it first appears in the second violin in mm. 133–34, and subsequently in the viola and cello (mm. 139–40) and first violin (mm. 145–46). In truth, the variant of γ is difficult to hear, for within these imitative sections it must compete with the ♩♬♩· figure that comes to dominate the first half of each measure in which the fragment appears. Nonetheless, as can be seen in example 4.18, the fragment is gradually transformed,

43. This is in fact the same sequence of fully diminished seventh chords initiated in mm. 14–16; however, that sequence led not to a B$^{o7}$ chord but to a B♭ pedal preparatory to the first full cadence on F in mm. 19–20.

In his analysis of mm. 129ff., Ratner interprets each fully diminished seventh chord as a leading-tone harmony with a normal resolution; see Ratner, *The Beethoven String Quartets*, 15. Indeed, on the very last eighth-note of mm. 128 and 129 there is a D minor chord, as there is a G minor chord in mm. 139 and 140, a C minor chord in mm. 145 and 146, and F and B♭ minor chords in mm. 147–50. However, I hear these harmonies as simply embellishing the fully diminished seventh chords that dominate these measures. This hearing is not without its theoretical problems—it suggests the kind of dissonant prolongation that has vexed a number of music theorists over the past two decades—but it has the advantage of capturing the intensity and uncompromising nature of this passage.

EXAMPLE 4.18 Ludwig van Beethoven, String Quartet, Op. 18 No. 1, first movement, mm. 145–54




over the course of mm. 147–51, into the γ fragment as it originally appeared. So transformed, it leads in a cascading sequence to the arrival on the dominant of F major in m. 167. In one sense, however, this is not the γ fragment as it was first heard in mm. 22–27, for that fragment was one-half of a two-measure motive. By contrast, the fragment of m. 151 and following is heard as the extension, in one voice, of the imitative entrances initiated in m. 129.

With its abrupt introduction of the learned style, combination of features from both the α and β forms of the motive, and prolongation of dissonant harmonies, the δ form of the principal motive provides elements of contrast and continuity with the preceding music. And compared with the various changes visited upon α and β in the course of the Exposition, the repetitions of mm. 129–46—both on the level of the individual motives that make up the imitative passages and on the level of the separate sections that prolong three different fully diminished seventh chords—make for an island of stability within the changing fortunes of the principal motive. Nonetheless, this moment of stability provides its own disruptions, and not only because it is based around a sequence of fully diminished seventh chords. The balance that had existed between the α and β forms of the motive—one associated

with a syntax of stability, the other with a syntax of fluidity, and each necessary for the other—has now disappeared: what is left after δ has done its work is only the chattering of the γ form of the motive. This, and passagework scales over a dominant pedal, is all that remains until the α form of the motive reasserts itself at the beginning of the Recapitulation.

## Recapitulation and Coda

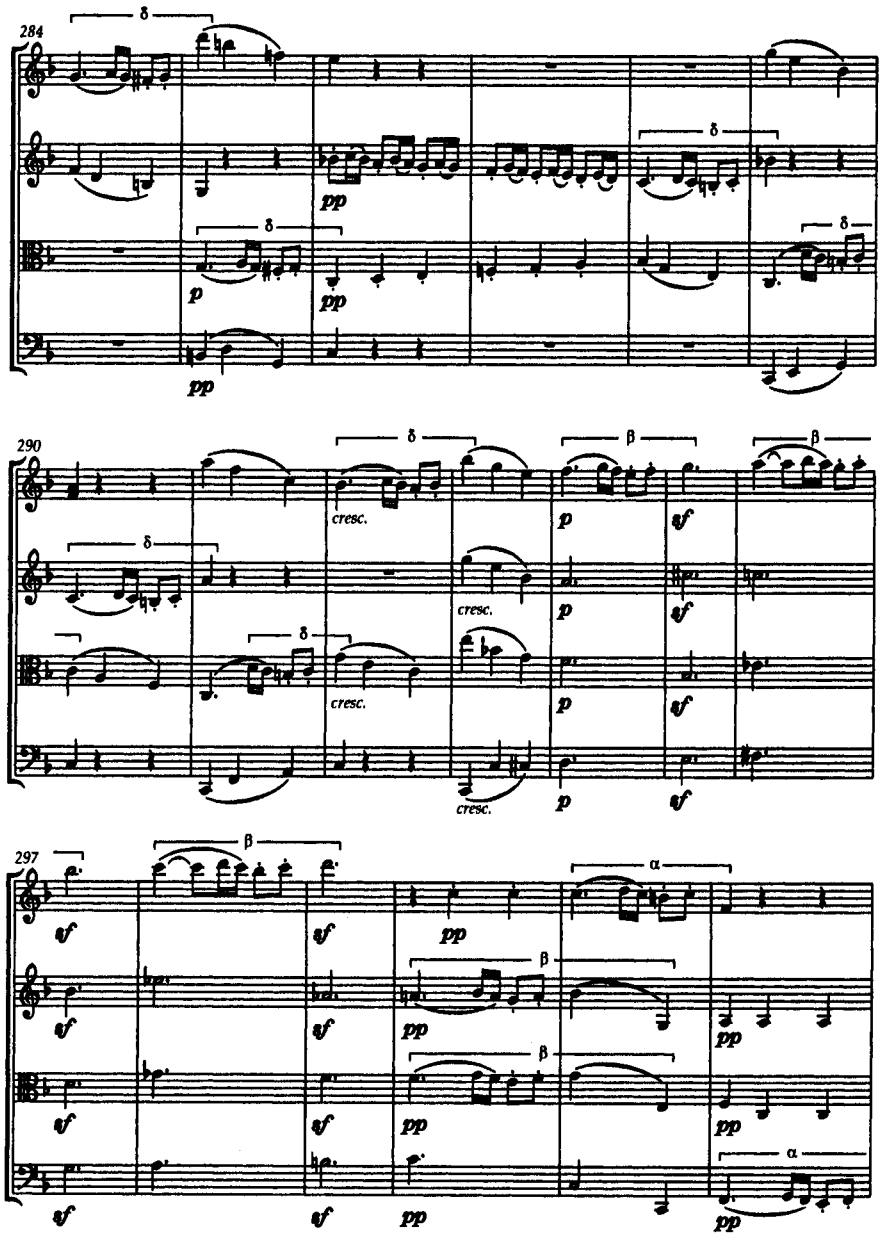
The appearance of motive forms in the Recapitulation, which begins with a literal return to the opening of the movement in m. 179, follows the same general outline as that of the Exposition. Missing, however, are almost all of the members of the β category of motives. After the initial response to the α motives (mm. 183–86, an exact repetition of mm. 5–8), β disappears, along with other vestiges of the opening twenty-odd measures, including the sequence of fully diminished seventh chords from mm. 14–16 that had their extension in mm. 129–46. Beethoven replaces mm. 9–29 with a brief episode that veers off toward G♭ major, which in its latter half makes reference to the transition that began in m. 30, and which also absorbs the function of the statement and counterstatement of motive forms in mm. 41–48.[44] The consequence of these changes is that the α form of the motive goes virtually unchallenged in the Recapitulation, for both the β form and the syntax of fluidity with which it was associated are absent. What we are left with is a clear impression of the stability—even obstinateness—of the α form of the principal motive.[45]

Beethoven does offer a final challenge to this stability in the Coda, which begins in m. 274. After a pair of short ascending passages in staccato quarter notes (mm. 274–81) that recall the quarter notes that followed δ in the Development and a brief descending figure built from the paired sixteenth-note portion of the stylized turn figure (mm. 282–83), the main motive reenters in the first violin in m. 284 (see ex. 4.19). With the viola's entrance at the octave below in m. 285, and continuation with the same sort of staccato quarter-notes that began the Coda, the reference is clear: Beethoven means to summon the δ form of the motive, although without its *fortissimo* entrances or dissonant harmonies. On the one hand, the subsequent entries (beginning in m. 288) hint at the possibility of further development as the motive is tossed back and forth between different voices; on the other hand, they all serve to prolong the dominant. The possibility of development becomes a bit more real with the move to the submediant, in m. 294, and the reprise of a variant of the β form of the motive. While the rhythmic figure is different from a typical member of the β category, the syntactic references are clear: the motive is sequenced upward (as it was in mm. 13–18, 21–26, and 123–27), and its harmonic

44. Levy notes that these changes tend to draw together various features of the original transition; see Levy, *Beethoven's Compositional Choices*, 79–83.

45. The impression of the solidity of the α form of the motive is given further support by Beethoven's decision to eliminate the repeat of the second section of the movement (that is, the Development and Recapitulation), one of the significant changes he made between the first and second versions of the quartet.

EXAMPLE 4.19 Ludwig van Beethoven, String Quartet, Op. 18 No. 1, first movement, mm. 284–302



context is thoroughly fluid, including four fully diminished seventh chords (in mm. 295, 296, 298, and 299) and a cadential dominant (in mm. 300–301). Against the resolution of the $\frac{6}{4}$ over this dominant, the α form of the motive finally enters (in the first violin in m. 301), and for the first and only time in the movement, its concluding downward leap outlines a falling $\hat{5}$–$\hat{1}$, signaling an end to the peregrinations suggested in the preceding measures.

After this closing gesture, there is one more bit of dialogue between α and β (not shown in ex. 4.19). The cello, beginning and ending on F2, repeats the variant of α that ends with a repeated note (in mm. 302–303 and 304–305, a variant first heard in mm. 41–42). It is answered by a variant of β reminiscent of that just heard in mm. 294–299 (played in thirds in the first and second violins in mm. 306–307 and 308–309). True to form, these β variants pull away from F major by introducing its $\flat\hat{6}$ (D♭), but they are answered by a cadential gesture that returns the movement, for once and for all, to F major.

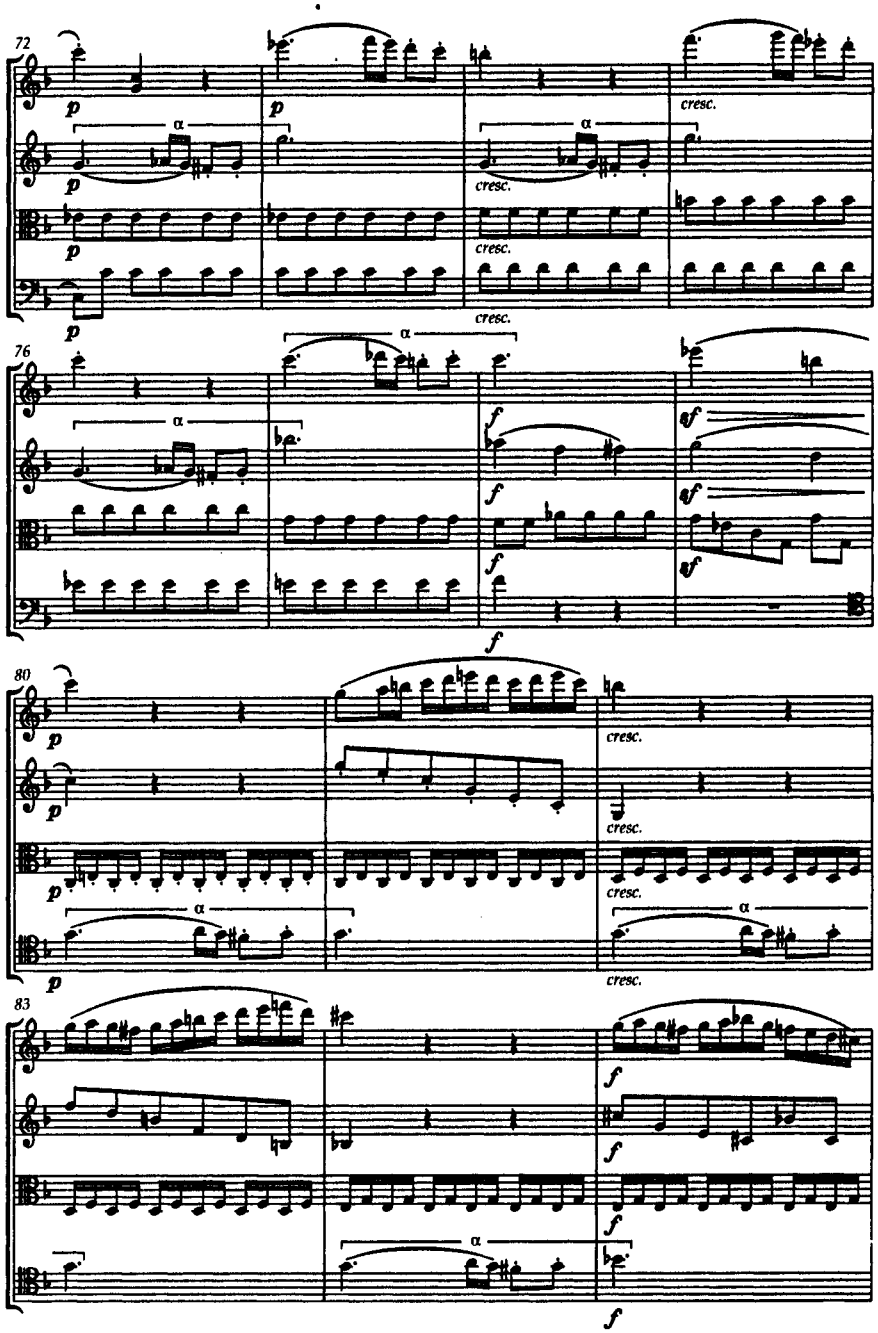
With regard to the principal motive of Beethoven's F major quartet, the Coda has a summational function that is twofold. First, the Coda revisits each of the three categories of motive forms that appeared over the course of the movement, doing so in reverse order of their initial appearance. Second, the Coda revisits the syntax with which each of these categories was associated. The Coda begins (much as did that of the first movement of Mozart's "Dissonance" Quartet) with a reference to the Development; the subsequent allusions to δ suggest a challenge to the integrity of the principal motive (through fragmentation) but also take place within a context of overall stability (with the bulk of the entrances linked to the dominant of F). The β motives that follow these δ forms exemplify the syntax with which β is most often associated: although the broad range of features exhibited by members of the category of β motives is not in evidence, the harmonic fluidity typical of the category is well represented by the fully diminished seventh chords prominent in the passage. And the final α motive, with its summoning of cadence, is the very embodiment of general, processive stability, yielded by the compositional strategies most often used for the category of which it is a member.

## Changes between the First and Second Versions of the F Major Quartet

Janet Levy observed that one of the principal differences between the first and second versions of the F major quartet is the greater concision of the second version.[46] Beethoven accomplished this in a number of ways, which together resulted in the elimination of a significant number of statements of the α form of the principal motive. Basic elements of the plan Beethoven followed in his revisions are evident in the changes made to the episode involving the α motive initiated in the midst of the second key area. The passage as originally written is given in example 4.20. In the fourteen measures spanned by the episode, there are seven clear statements of the α motive. In the second version of the movement, this episode was cut back to

46. See Levy, *Beethoven's Compositional Choices*, 2–4.

EXAMPLE 4.20 Ludwig van Beethoven, String Quartet, Op. 18 No. 1 (first version), first movement, mm. 72–85



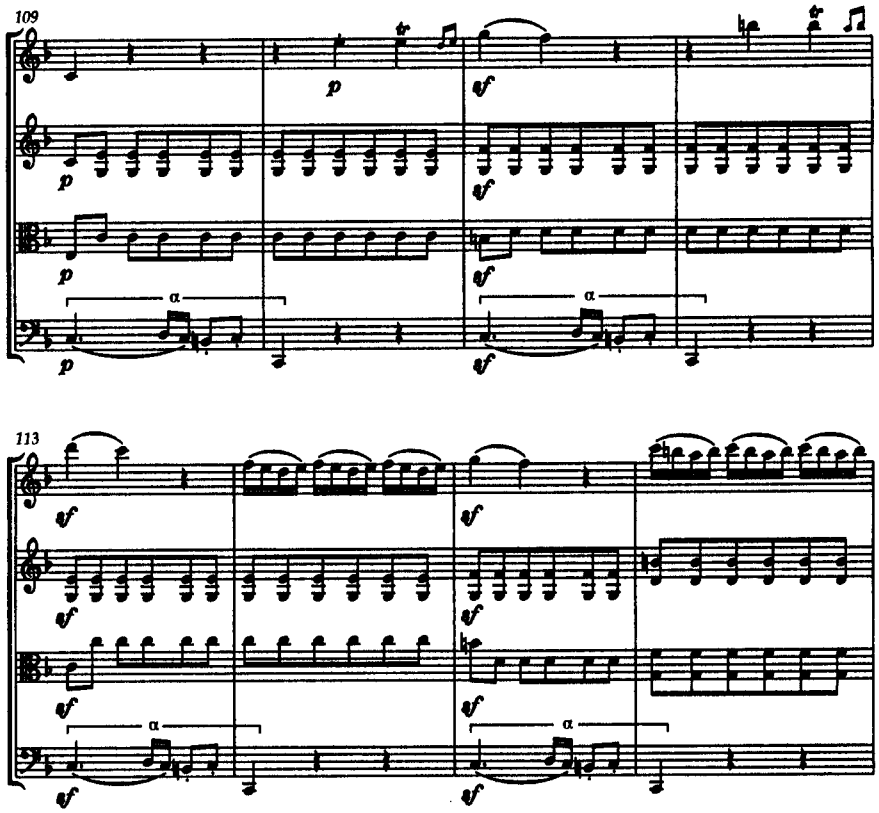
six measures (which are given in ex. 4.15c), with only three statements of the α form of the motive. But perhaps more important than simply reducing the appearances of the α motive, Beethoven appears to have rethought how it fits into the overall musical fabric, as a closer inspection of his first attempt at this episode reveals.

The original version of the passage divides into two parts, the first in C minor (mm. 72–79), the second in C major (mm. 80–86). In both parts, there are three main statements of the α motive: in the C minor section these are in the second violin, and each ends with a large ascending leap; in the C major section the statements are given in the cello and end with a repeated note. (The latter form of the α motive is heard in a prior episode, given in ex. 4.15b, which is substantially the same in both versions of the movement.) Each of the second violin's statements is answered by a counterstatement in the first violin. The counterstatements in mm. 73 and 75 begin with a turn figure and share their rhythmic profile with the α form of the motive. On the other hand, the counterstatement of m. 77 answers with the same variant of the α motive that will be heard momentarily in the cello. The cello's three statements of this α variant are not given similar counterstatements but are answered instead by passagework in the first violin, which ultimately leads out of the episode.

When revising this passage, Beethoven cut the C minor section, salvaging the first violin's counterstatements of mm. 73 and 75. These he placed an octave lower, but still in the first violin (see ex. 4.15c). He omitted the counterstatement of m. 77, changing the first violin's music so that it starts with what appears to be another counterstatement, but which then dissolves into passagework that leads out of the episode. The effect of these alterations is not just concision but a view of the α motive that is changed in three ways. First, by dropping the three statements of the motive made by the second violin in the C minor section, Beethoven eliminated a further variant—namely, one that ends with a large leap—from the category of α forms. Second, the counterstatements made by the first violin, now closer in register to the cello's statements of the motive, make a direct if limited challenge to the structure of the α motivic category. Third, changing the last counterstatement so that it leads out of the episode (rather than echoing the cello's third statement of the α motive, as the original version of m. 77 would) lets the challenge stand, albeit only for a moment or two. Thus, in the revised version of the passage Beethoven has not only achieved great concision, but he has also accomplished two strategic goals: he has tightened the boundaries of the α motivic category, and he has intimated how these boundaries might be challenged.

A similar elimination of outright statements of the α motive and modification of the structure of motivic categories occurs with the final statement of the motive at the end of the Exposition. In the first version of the passage, shown in example 4.21, the "musette" variant of the α motive (in mm. 109 and 111) is answered by a counterstatement with a trill and appogiatura (in mm. 110 and 112). When the phrase is repeated (in mm. 113–16), the counterstatement is embellished by repeated turn figures, but the α motive is not. In Beethoven's revisions, both the counterstatement and the "musette" variant are embellished when the phrase is repeated (mm. 105–108 of ex. 4.15d). Although this removes two explicit state-

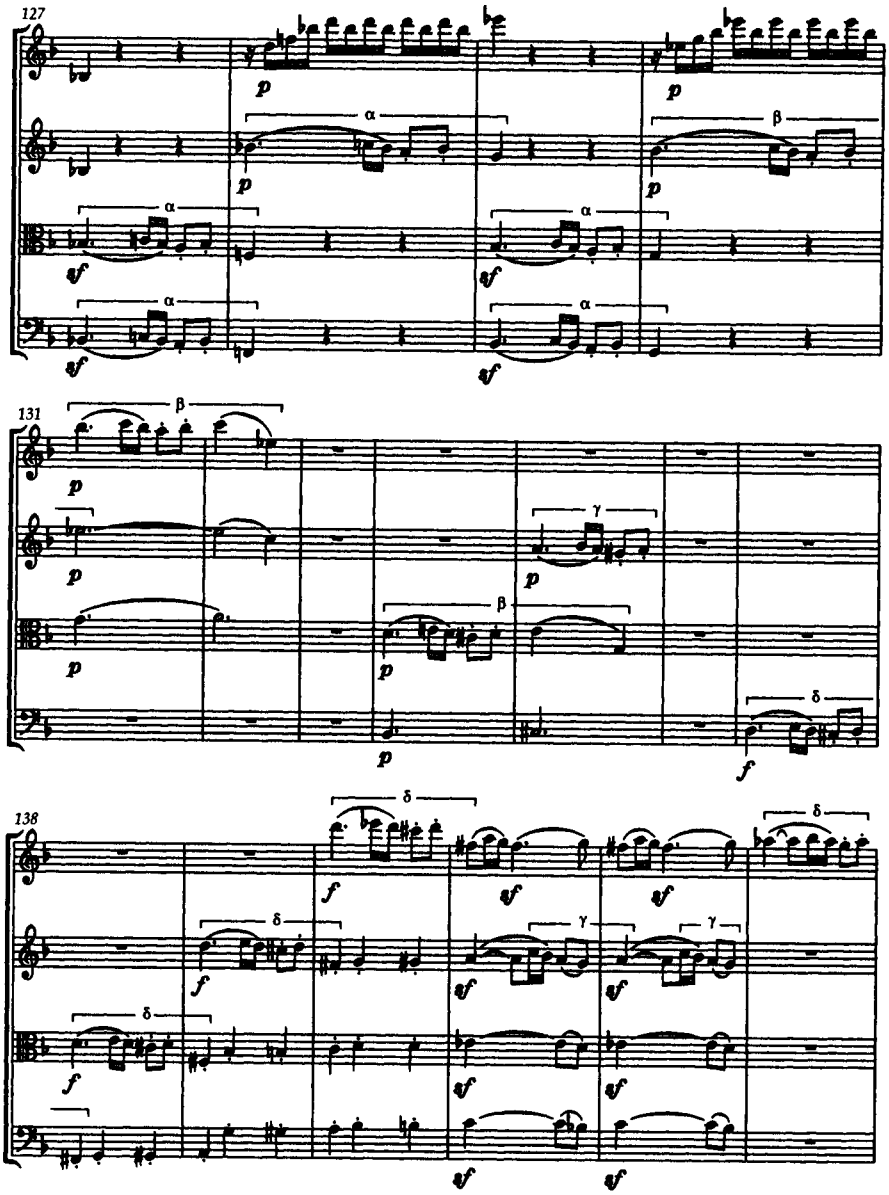
EXAMPLE 4.21 Ludwig van Beethoven, String Quartet, Op. 18 No. 1 (first version), first movement, mm. 109–16




ments of the α motive (and, consequently, could be seen to weaken the structure of that category), the change makes it possible to hear the first violin's counter-statements as further variants of the β motivic category (a possibility raised in the earlier discussion of ex. 4.15d).

A further result of Beethoven's decision to reduce the number of appearances of α in the revised version of the movement was a clearer distinction and more even balance between the α and β forms of the motive. In the Development of the first version of the movement, for instance, the two forms of the principal motive blend into one another. As shown in example 4.22, the entrances of the α form subsequent to the A major flourish that opens the Development are first answered by an α form of the motive and then by a variant of β (mm. 127–32). Appearing above the extended tail of this answering β is another β (a quite typical member of the category), which leads, after a pause, to yet one more entrance of β (itself answered by a γ fragment, as it was in m. 21 and following). Thus an episode beginning with the α form of the motive ends with the β form of the motive, blurring the distinction between the two. In Beethoven's revised version (ex. 4.17),

EXAMPLE 4.22 Ludwig van Beethoven, String Quartet, Op. 18 No. 1 (first version), first movement, mm. 127–43



the two forms are successive but do not overlap: α and β stand out more clearly as separate entities.

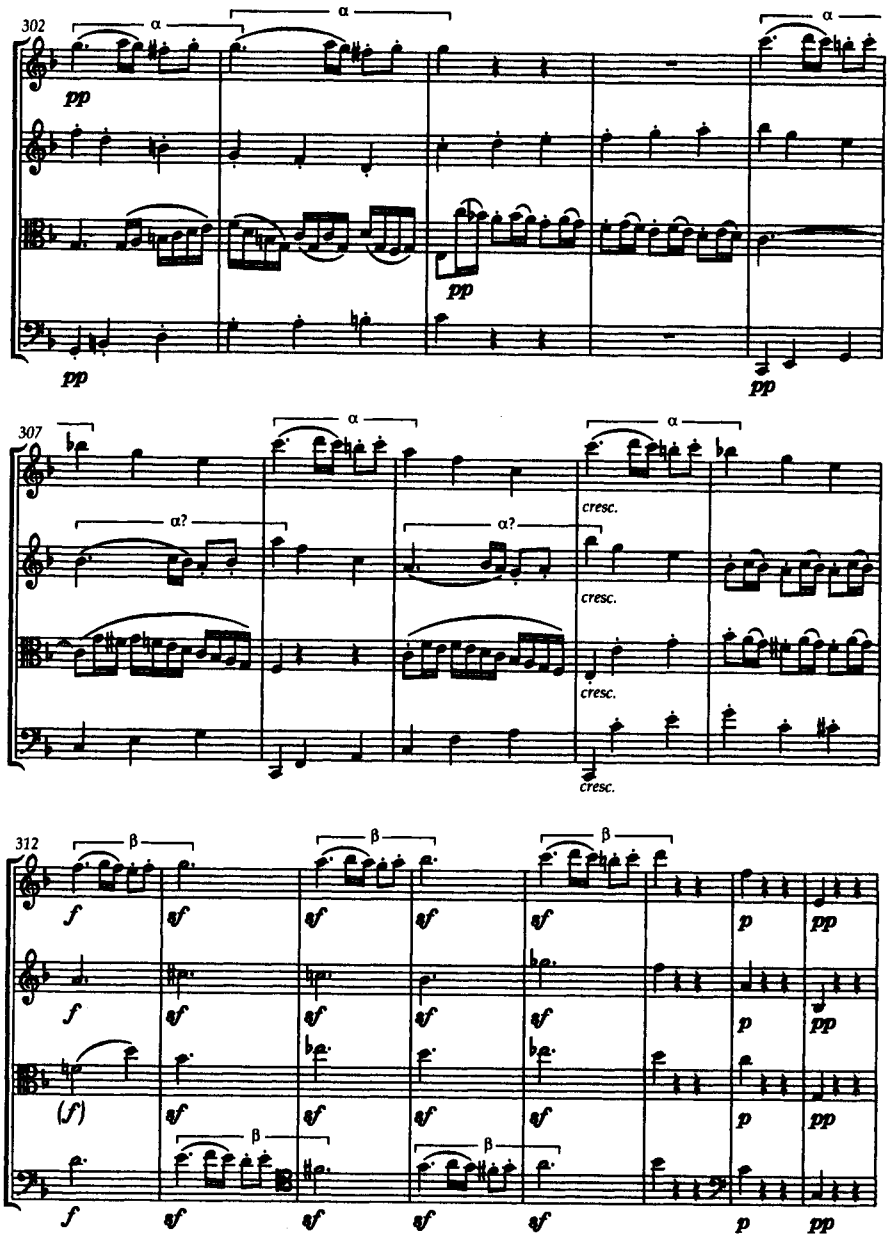
Beethoven's revisions to the Development also resulted in a sharper distinction between the δ form of the motive and the α and β forms, especially where the syntax of the motivic categories is concerned. In the original version, the answering statements by the second violin (first α in m. 128 of ex. 4.22, then β in m. 130) both enter at the octave above the viola. They thus anticipate the imitative entrances of δ in mm. 137 and following. When these answers are removed in the second version (cf. ex. 4.17, m. 119 and following), the entrances of δ seem fresher—the introduction of the learned style seems more of an innovation. The entry of δ in the first version of the movement, moreover, leads to a consonant harmony (the dominant of G minor), and the downward leap at this point is a minor sixth (ex. 4.22, mm. 137–38). In the revised version, δ leads to a dissonant harmony (ex. 4.17, mm. 129–30), and the downward leap is a diminished seventh. In sum, in the original version of the Development δ seems to grow out of a generalized idea of the principal motive (α and β blurring together); after Beethoven's revisions, the introduction of δ is more striking, and it contrasts sharply with the statements of α and β that precede it.

Revisions to the Coda offer a final bit of evidence that the changes Beethoven made were with the intent of more clearly differentiating between different versions of his principal motive. The original version of the Coda, starting at the point that references to the principal motive become explicit, is given in example 4.23. In the first half of this excerpt, Beethoven relies heavily on the α form of the motive, providing statements of it that begin variously on G and C in the first violin, and B♭ and A in the second violin. Because the latter have an ascending skip after the turn figure, they could also be heard as a variant of the β form of the motive, but the quality of the α form of the motive still pervades. In m. 312, with the shift away from the dominant to the submediant, the β form of the motive enters in earnest, with a clear reference to the successive entries between β and γ over the fully diminished seventh chords found in the Exposition. This series of β motives leads ultimately to a conventional, if deliberately underplayed, cadence.

In the second version of the Coda (ex. 4.19), the α motive is withheld, and the first references are to the δ form. As noted in the discussion of example 4.19, the syntax of the passage owes much to the fluidity associated with the process of developing musical materials and extends to the β motive that enters in m. 294. When the α form of the motive finally returns in m. 301, it is coincident with a perfect authentic cadence and signals a definitive end to the possibility of further development raised by the entrances of the δ motive.

The changes Beethoven made to the first movement of the F major quartet, then, had two important results. First, they created a more balanced relationship between the α and β forms of the motive in the Exposition. Second, the distinctions between the three main categories of motive forms are more sharply drawn. These changes make the syntax associated with each category somewhat more apparent—that is, we can more easily hear the role of each motive form in the discourse of the movement as a whole.

EXAMPLE 4.23 Ludwig van Beethoven, String Quartet, Op. 18 No. 1 (first version), first movement, mm. 302–19



## Compositional Strategy and Multiple Levels of Musical Syntax

As A. B. Marx noted, the principal motive of the first movement of Beethoven's F major string quartet is a nearly constant presence, never far from the listener's ear. The function of this dense network of motives, however, is not so much to lend the movement coherence—after all, Beethoven actually reduced the number of times the α form of the motive appeared when he revised the movement—but to set up a musical discourse separate from but coordinated with the harmonic and thematic syntax of sonata form. This discourse involves not simply a syntax specific to the principal form of the motive (as in the case of the first movement of the "Dissonance" Quartet), but higher levels of syntax of the sort seen in the sixth quartet from the Op. 18 set.

Beethoven's compositional strategy depends on clear distinctions between different forms of the motive. These distinctions are made not only on the basis of audible features of the motive (of the sort captured in the conceptual models of fig. 4.7) but also on the compositional strategies that lend a distinctive character to each category of motive forms: α as stable in the face of various challenges and β as an agent of change and transition. By the end of the Exposition, a delicate balance has been established between the two, a balance not in evidence in the first version of the quartet, where the category of α motive-forms dominated.

This balance is maintained through the first part of the Development, but it is disrupted by the introduction of the δ form of the motive and the topic of the learned style (ex. 4.17, m. 129). It is only with the Recapitulation that the focus comes back to the α form of the motive, where it stays (with little or no competition from β) until the Coda. The disruption created by δ and the recuperation facilitated by α suggest a function for the topic of the learned style within Beethoven's compositional practice: the fugato on the δ form of the motive creates the impression of "development" by temporarily suspending the normative syntax for the projection of tonal areas associated with α and β, a syntax restored only with the reprise of α at the beginning of the Recapitulation. This function of α is in evidence again at the end of the movement when, after hints at further development and harmonic fluidity in the Coda, α returns as the melodic voice of the final cadence in F major.

The distinctions among these three forms of the principal motive—distinctions that make possible a musical discourse that is focused on the changing fortunes of the motive over the course of the movement, and that were made more apparent in Beethoven's revisions to the movement—suggest that motivic analysis could be refined by attending to how motives are used, rather than to how successfully they can be traced back to some Ur-type. Such a refinement makes possible an appreciation of multiple levels of syntax—levels that interact with the overall harmonic and thematic syntax of the musical work, and which collectively engage more and less immediate levels of cognitive processing. Ultimately, an appreciation of this refinement leads to a better understanding of how compositional strategies shape our conceptions of musical organization.

## CONCLUSIONS

For Umberto Eco, one of the mysteries of semiotics was that music could have a fairly well developed syntax and yet no apparent semantic depth. Music meant something, but outside of conventionalized signs like those associated with military or pastoral music, it was quite difficult to say what this was. In this chapter, I have not tried to address this problem head on; indeed, describing the origins and scope of musical meaning is an endeavor far more ambitious than what I undertake here. I have, however, tried to argue that musical syntax and semantics are better viewed as points along a continuum than as discrete domains. More directly, composers (as well as others who bring music to life) make use of strategies that disrupt and redefine musical syntax as part of a process of meaning construction.

My path into these compositional strategies has been through categorization, on the assumption that categories of musical events represent a common ground where the concerns of composers and listeners meet. The viability of this assumption was demonstrated by the structure of the waltzes produced by the musical dice game. Where formal structure needed to be clearly defined (for instance, at the end of each half of the waltz), the range of category members within each measure-category was quite narrow, and what was typical of the relevant measure-category was quite apparent. Where formal structure could be more open (for example, at the beginning of the second half of the waltz), the range of category members was quite wide and typicality harder to determine. Typicality could thus be used to define formal syntax, a syntax that not incidentally coordinated with—and thus confirmed—harmonic syntax.

A somewhat different example of a compositional strategy organized around typicality was seen in the category for the principal motive of the first movement of Mozart's "Dissonance" Quartet. At important moments in the movement—at the beginning and close of the Exposition and Recapitulation, and in the Coda—either thoroughly or highly typical forms of the motive made their appearance. These were followed by less typical members of the motivic category, leading either to a reprise of a typical form or to new musical material. This basic process—a syntax of typicality yielding to atypicality—was repeated three times in the course of the Development, ending only with the appearance of a thoroughly typical form that signaled the start of the Recapitulation. While the compositional strategies that shaped the different versions of the principal motive ran parallel to the formal syntax of the movement, they were nonetheless independent. The motive thus emerged as a topic for musical discourse, whose progress could be followed along with the other strands of musical discourse comprised by the movement.

By contrast, the principal motive of the first movement of the B♭ major quartet from Beethoven's Op. 18 never functioned as a topic for musical discourse. Although the motive forms fell into two distinct categories, individual members never strayed far from their respective conceptual models. In consequence, the immediate level of syntax for the categories of motive forms was not particularly interesting, since members were almost always quite typical of their category. However, over the course of the Exposition each category acquired distinct functions: one ($\alpha$)

signaled the opening of musical discourse; the other (β) signaled the close of discourse. This emergent syntax was then exploited in the Development and more generally served to show how musical topics of the sort described by Ratner could come into being.

It was in Beethoven's F major quartet from Op. 18 that we saw both of these levels of syntax come together. As with the B♭ major quartet, Beethoven introduced two categories of motive forms (α and β) in the Exposition, with a third (γ) playing a decidedly subsidiary role. A compositional strategy focused on typicality was applied to both categories, but where the range of variation for α was relatively narrow, for β it was quite wide. Contextual pressures on β even suggested that the concept of what counted as a member of the category could slip to include a number of rather remote variants, a kind of slippage that never seemed to threaten α. By the end of the Exposition, and even into the beginning of the Development, α and β appeared to maintain a delicate balance, the former stable in the face of various challenges, the latter an agent of change and transition. Beethoven then introduced a new form of the motive (δ) to disrupt this balance, creating a species of development of high drama: α and β were swept from the stage, vanquished by a volatile combination of properties drawn from each. Balance within the movement as a whole (but not between α and β) was restored with the return of α at the start of the Recapitulation. Even the developmental urges of the Coda could not threaten this equilibrium. The compositional strategies applied to the principal motive, then, operate on both a local level (within a given category) and global level (among groups of categories). Again, these syntactic levels are independent of but coordinated with those for the overall harmonic and thematic structure of the movement—the syntactic strands associated with the principal motive are multiple, but they do not exhaust the syntax of the whole.

In sum, these analyses suggest that the construction of meaning in music can be achieved through the way composers choose to deploy the elements of musical syntax. Of course, compositional strategy is not the only source of meaning construction in music—I have also pointed out the contributions of musical topics and cross-domain mapping to musical semantics—but it is one to which our knowledge of categorization can be profitably applied. As the meeting place for the concerns of composers and listeners, categories of musical events are important to both compositional strategy and musical syntax, for they represent a means through which uniquely musical meaning can be created.

CHAPTER FIVE

# CULTURAL KNOWLEDGE AND MUSICAL ONTOLOGY

As part of a search for an interpretive theory of culture and consequent to a critique of early work in cognitive anthropology, Clifford Geertz momentarily and somewhat famously turned his thoughts to music as an example of culture. Taking a Beethoven string quartet as his case in point, he suggested that no one would identify the quartet with its score, with the skills and knowledge needed to play it, with the understanding of it possessed by its performers or audience, with a particular performance of it, or with some mysterious entity transcending material existence. There was one perspective, however, that Geertz thought would meet with general agreement: "That a Beethoven quartet is a temporally developed tonal structure, a coherent sequence of modeled sound—in a word, music—and not anybody's knowledge of or belief about anything, including how to play it, is a proposition to which most people are, upon reflection, likely to assent."[1]

By this example, Geertz intended to put to rest, for once and for all, the notion that culture exists as some sort of program inside people's heads. Nonetheless, if music is not knowledge or belief, then what is it? Geertz's definition of music as "a coherent sequence of modeled sound" is simply too broad, for it would have to include strings of Morse code, speech in an unknown tongue, and the cries of various animals. Indeed, from the perspective of this definition, there is, in principle, nothing to distinguish music from noise or noise from music.[2] "Music" is a category constructed by humans, not a substance or set of relations floating free in the world.

1. Clifford Geertz, "Thick Description: Toward an Interpretive Theory of Culture," in *The Interpretation of Cultures: Selected Essays* (New York: Basic Books, 1973), 11–12.

2. Much can and has been said regarding the relationship between music and noise. Three sources from the last ninety years represent the range of issues engaged by the topic: Luigi Russolo, *The Art of Noises* (1916), trans. Barclay Brown (New York: Pendragon Press, 1986); Jacques Attali, *Noise: The Political Economy of Music*, trans. Brian Massumi, with a foreword by Fredric Jameson and an afterword by Susan McClary, Theory and History of Literature, 16 (Minneapolis: University of Minnesota Press, 1985); and the improvised performances by Derek Bailey, Pat Metheny, Gregg Bendian, and Paul Wertico on *The Sign of 4* (Knitting Factory Works, [CD] KFW 197, 1997).

I indicate the media of the recordings I cite with the following abbreviations: CD = compact disk; LP = 33 RPM record; 78 = 78 RPM record.

As such, the notion of music is dependent upon and constrained by human knowledge, which presumably resides with individual humans. Having said this, it must also be said that the continued existence of a Beethoven string quartet does not depend on the personal knowledge of any one individual. Music is indeed a cultural fact, and culture is largely public and interpersonal.

Cultural knowledge, as represented by Geertz's example, is thus a bit of a paradox. On the one hand, it is not something that is determined by any one individual: culture is, as Geertz maintained, public. On the other hand, individuals need such knowledge if they are going to participate in culture, and it is through the participation of individual humans that culture—and cultural knowledge—is maintained.

The problem, as framed by Geertz's example of a Beethoven quartet, is more specifically one of musical ontology—that is, the ontological status of a work of music. In the West, it is common to regard pieces of music as having an objective status, speaking of a "work of music" as one might of a painting or sculpture. This habit of thought persists in spite of the transience of musical phenomena: unlike a painting or sculpture, a "musical work" has no enduring material existence. The enigma of "works" that are of profound cultural importance and yet are totally ephemeral has led to a number of attempts by philosophers and others to account for the objective status of the work of music and, by extension, music itself; thus the problem of musical ontology.[3] A contrasting perspective has developed among ethnomusicologists who, working chiefly with non-Western music, have noted the fluidity of musical practice and a concomitant absence of any ascription of permanence to the products of musical activity. These observations render doubtful any universal solution to the problem of the ontology of the musical work, or even "music," since many cultures do not show linguistic or practical evidence of equivalent concepts.[4] And yet, musical practice within these cultures, while fluid, is certainly not

3. Thoughtful discussions of the problem of musical ontology can be found in Lydia Goehr, *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music* (Oxford: Clarendon Press, 1992); Philip V. Bohlman, "Ontologies of Music," in *Rethinking Music*, ed. Nicholas Cook and Mark Everist (Oxford: Oxford University Press, 1999), 17–34; and Nicholas Cook, "At the Borders of Musical Identity: Schenker, Corelli and the Graces," *Music Analysis* 18 (1999): 179–233. See also the essays collected in Michael Talbot, ed., *The Musical Work: Reality or Invention?* (Liverpool: University of Liverpool, 2000).

4. Examples of cultures the languages of which lack a word equivalent to "music" or concepts equivalent to that of a "musical work" include the Hausa of Nigeria, the Macuma Shuar of Ecuador, and the Mapuche of Argentina; Charles Keil reports such a lack in over twenty languages of the African continent and gives detailed discussion of the methodological problems involved in studying such cultures. See David W. Ames and Anthony V. King, *Glossary of Hausa Music and Its Social Contexts* (Evanston, Ill.: Northwestern University Press, 1971), ix; William Belzner, "Music, Modernization, and Westernization among the Macuma Shuar," in *Cultural Transformations and Ethnicity in Modern Ecuador*, ed. Norman E. Whitten, Jr. (Urbana: University of Illinois Press, 1981), 735–36; Carol E. Robertson-DeCarbo, "*Tayil* as Category and Communication among the Argentine Mapuche: A Methodological Suggestion," *Yearbook of the International Folk Music Council* 8 (1976): 39; and Charles Keil, *Tiv Song* (Chicago: University of Chicago Press, 1979), 27–52.

I should make clear that what I find dubious is a universal solution to the problem of the ontology of a musical work. I do not think it is impossible, however, to speak about "music" in the case of a culture that does not show linguistic or practical evidence of such a concept, as long as one does not attempt to colonize that culture by insisting that its members, when they try to make sense of their own experiences and cultural practices, make use of the term.

unregulated. In general, the habit of objectifying aspects of musical practice is omnipresent; in consequence, the problem of musical ontology, while most properly one that arises only among those who believe in an objective status for the work of music, also has applications to situations where no such object is posited.

In this chapter, I propose that solutions to the problems of cultural knowledge in general, and of musical ontology in particular, can be found in what we now know about processes of categorization. In brief, knowing that Beethoven's quartet is music—or that anything else is music, for that matter—means knowing how to categorize sequences of sound-events in accordance with conceptual models shared with other members of a culture. More specifically, each "work of music"—or whatever other unit of cultural currency we choose to focus on—constitutes a Type 1 category of the sort discussed in chapter 2.[5] On the one hand, the transparency and immediacy of the conceptual models that guide the process of categorization explain why we describe works of music in the same language we use for objects: we speak of "the Beethoven quartet" just as we speak of "the red cup," even though the former (as sounding music) is a phenomenon completely unlike the latter.[6] On the other hand, because membership in Type 1 categories is by degree, a certain amount of fluidity is possible, varying in degree with the cultural context relative to which the conceptual model for the category is framed. We might include a wide variety of performances in the category for the Beethoven quartet—some more, some less competent, even some in an arrangement for four-hand piano—but we could well draw the line at a version for pitched percussion and automobile horns.

In what follows, I explore the idea that musical works can be thought of as Type 1 categories and that such categories are a manifestation of cultural knowledge. I approach this issue by taking a close look at two popular songs from the early twentieth century: "I Got Rhythm" and "Bye Bye Blackbird." Songs like this challenge our ideas about what constitutes a "work of music": while they are almost always associated with a score (as is a Beethoven quartet), they are also part of a performance tradition that involves some level of improvisation (introducing the fluidity typical of many non-Western performance traditions).

"I Got Rhythm" provides a good example of this situation, for there were multiple competing versions of the song almost from the moment it appeared on Broadway late in 1930. In the first section below, I consider a number of recordings

5. Of course, a work of music could be regarded as a Type 2 category were necessary and sufficient conditions for category membership specified. This is how I would interpret Nelson Goodman's proposal that the score is the absolute determinant of musical identity. Goodman writes, "If we allow the least deviation [from the score], all assurance of work-preservation and score-preservation is lost; for by a series of one-note errors of omission, addition, and modification, we can go all the way from Beethoven's Fifth Symphony to *Three Blind Mice*. Thus while a score may leave unspecified many features of a performance, and allow for considerable variation in others within prescribed limits, full compliance with the specifications given is categorically required." Nelson Goodman, *Languages of Art: An Approach to a Theory of Symbols*, 2nd ed. (Indianapolis: Hackett Publishing, 1976), 186–87.

6. There are, of course, aesthetic perspectives that support this habit (such as that associated with the "work-concept," as described by Lydia Goehr in *The Imaginary Museum of Musical Works*, 89–119) and technological innovations that give it credence (like music notation and sound recording), but these are not preconditions for the notion of a musical object.

of the song made during the 1930s and 1940s, explore how the model for the song changed and how it stayed the same, and show that both the fluidity and the stability of this model are straightforward consequences of our use of categorization to organize our understanding of music.

"Bye Bye Blackbird" presents a somewhat different problem, which takes us deeper into the relationship between categorization and cultural knowledge. Written as a popular song during the 1920s and quite successful when it first appeared, it never really entered the standard repertoire until a recording by Miles Davis in 1956 transformed the song into a jazz tune.[7] Because there was no tradition of gradual change, it was possible for radically different models for the song to coexist and come into conflict. In the second section below, I consider this sort of conflict in light of what Mikhail Bakhtin called double-voiced discourse and what Henry Louis Gates Jr. calls Signifyin(g). I attempt to show that a fuller account of how we categorize musical works requires that we go outside of the relatively parochial concerns of musical performance and consider the context in which musical performance takes place.

In the third section, I offer some concluding thoughts on cultural knowledge, musical ontology, and the cognitive processes through which we structure our understanding of the world. I end the chapter with some reflections about how the perspective on musical ontology I offer here can change the way we view music theory.

## CATEGORIZATION, MUSICAL ONTOLOGY, AND CULTURAL KNOWLEDGE

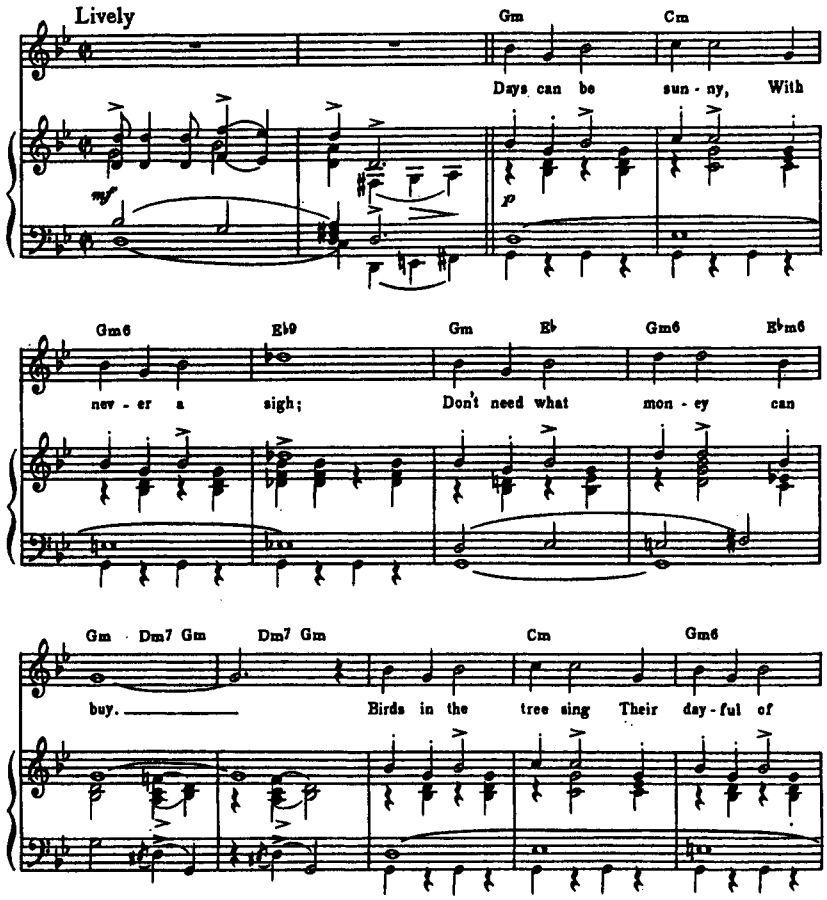
### A Multiplicity of "Rhythm"s, 1930–1946

"I Got Rhythm" was written by George and Ira Gershwin for their musical *Girl Crazy*, which opened on Broadway on 14 October 1930. The song, intended to be a show stopper, appeared near the end of the first act, and its refrain was repeated a number of times: to accompany a tap dance chorus after the final vocal refrain; as an encore for Ethel Merman, whose success with the song became legendary; as part of the mostly instrumental entr'acte before the second act; and, sung by the whole company, as the show's concluding number.

The song itself, the score for which is given in example 5.1, comprises an introduction, verse, and refrain. The melody of the verse begins with an eight-measure phrase in G minor (featuring a prominent $\flat\hat{5}$ at its midpoint), which is given a var-

7. I borrow the distinction between a song and a tune from Richard Crawford, who writes, "in the jazz tradition, we usually speak of tunes, not songs. A jazz tune is defined first and foremost by its structure: by the pattern of repetition and contrast in its melodic phrases and the harmonic framework underlying them. Second, it is defined by its ethos: by the mood it projects and the tempo at which it is played. Only third does its melody come into play, for in the jazz tradition the melody is often little more than an entrée into the performance; after being heard, it is usually discarded for free melodic invention by the performers." Richard Crawford, "George Gershwin's 'I Got Rhythm' (1930)," in *The American Musical Landscape* (Berkeley: University of California Press, 1993), 221, 225.

EXAMPLE 5.1 Version of George and Ira Gershwin's "I Got Rhythm" used in act 1 of *Girl Crazy*. "I Got Rhythm," words and music by George Gershwin and Ira Gershwin, © 1930 WB Music Corp. (Renewed). All rights reserved. Used by permission. Warner Bros. Publications U.S. Inc., Miami, FL 33014.

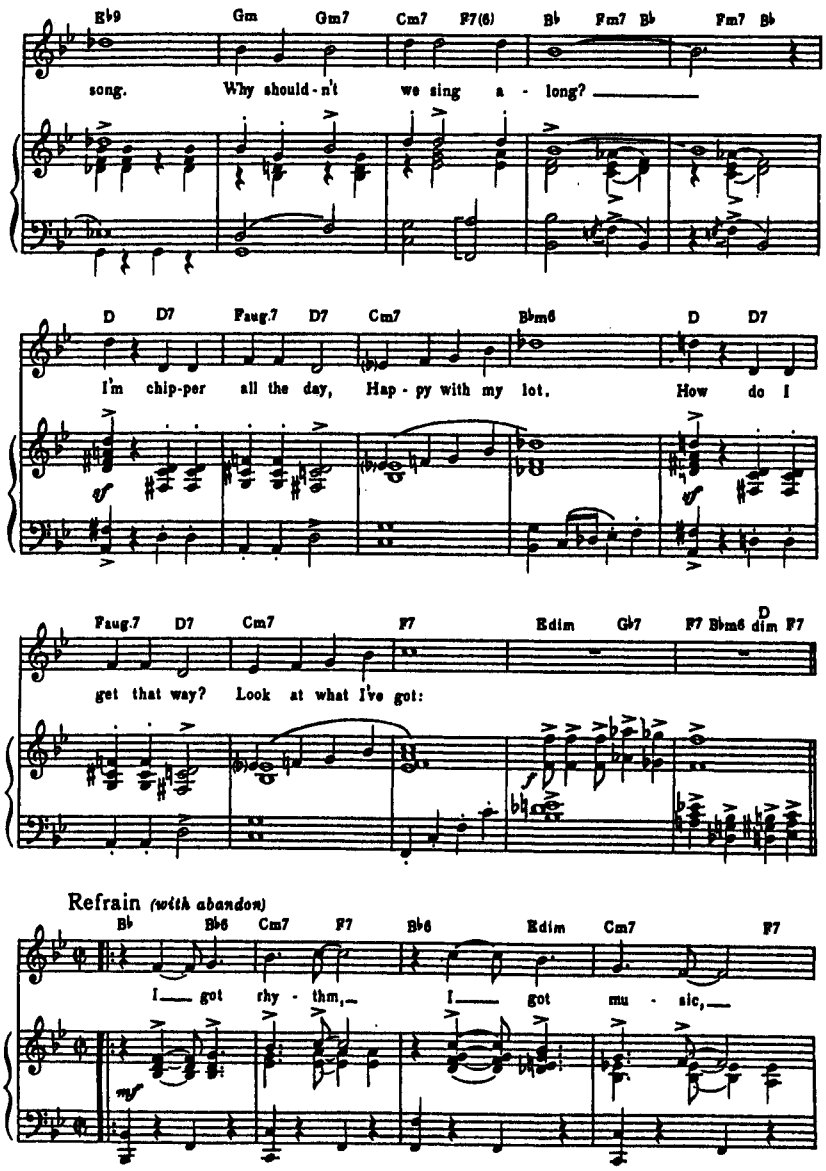



*(continued)*

ied repetition. This is followed by a four-measure phrase, which is also repeated (again, with variations), after which a two-measure transition leads to the refrain. The refrain is in the AABA form typical of popular songs of this period, but it includes an ascent to F5 near the end of the last A section. This climactic rise requires a two-measure tag to bring the melody back to B♭4 and yields a 34-measure AABA' form.

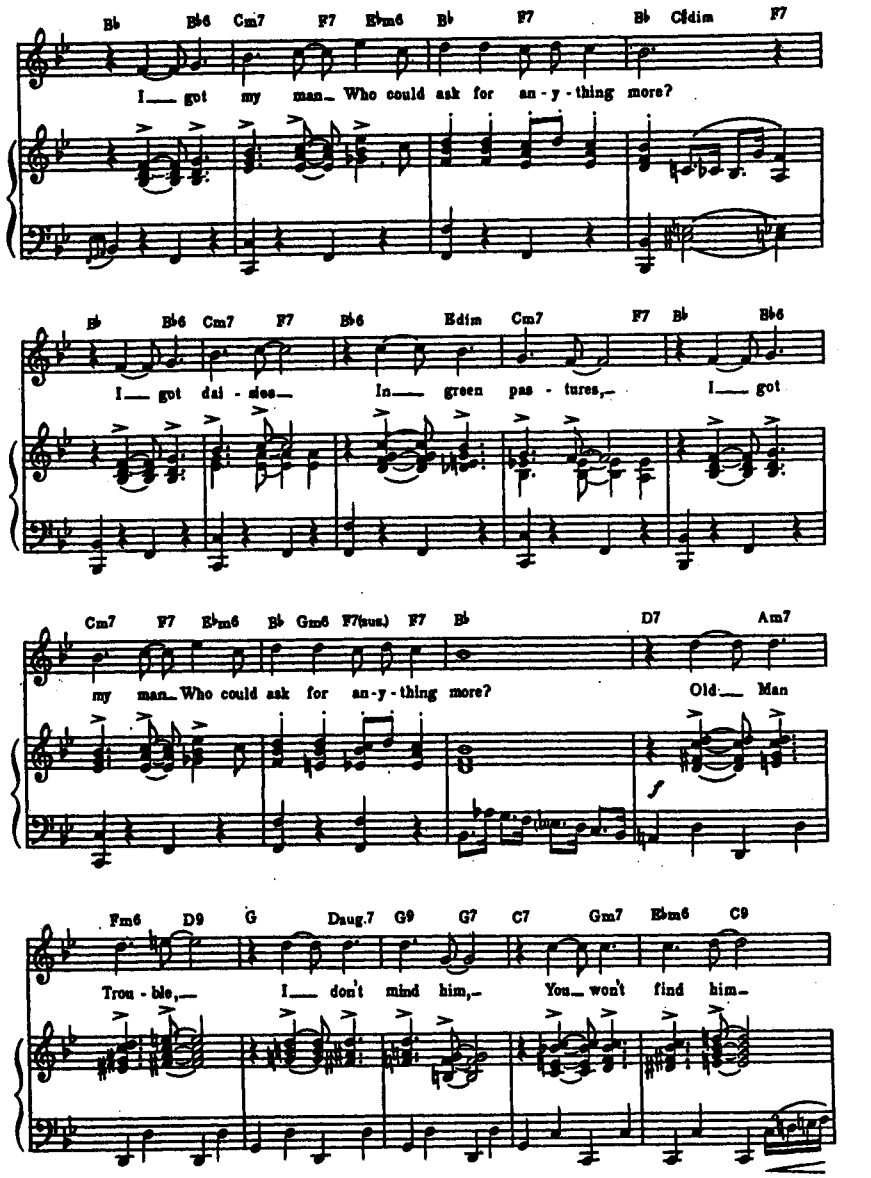
"I Got Rhythm" was only rarely recorded as it appeared in the original show. Although it contributed significantly to her professional career, Ethel Merman did not record it until 1947. Many amateurs and semi-professional musicians must have performed the song from the sheet music, which did reproduce the music as it

EXAMPLE 5.1 *(continued)*



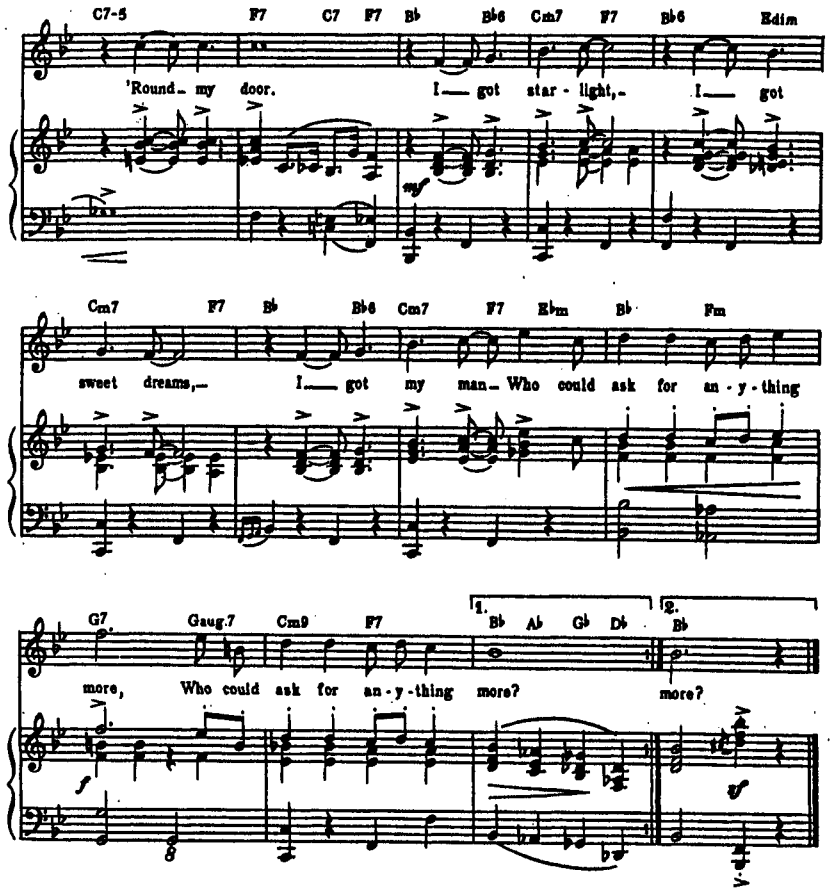

appeared in *Girl Crazy*. This was not the case, however, with professional musicians. Even in the earliest recordings (made within a few weeks of the show's opening), the verse loses its introductory (and narrative) function and is used instead as an interlude between repetitions of the refrain. For example, the recording made on 23 October 1930 by Red Nichols and His Five Pennies (a group drawn from the pit orchestra for *Girl Crazy*) begins with an instrumental version of the original refrain,

EXAMPLE 5.1 *(continued)*




*(continued)*

EXAMPLE 5.1 *(continued)*




followed by the refrain with vocal.[8] Only after this does the music for the verse enter, but without vocals. The recording concludes with two more statements of the refrain, also without vocals. A very similar arrangement was used by Fred Rich and a studio band in their recording of 29 October 1930. The overall form of their rendition consists of three refrains (the second with vocal), followed by an instrumental verse, followed by one complete and one partial statement of the refrain (the latter using only the B and A' sections to conclude the performance).[9]

In the two years following the opening of the show, almost all of the recordings that appeared made use of vocals, although none of which I am aware used the words for the verse.[10] With the instrumental version recorded by Don Redman and

8. Red Nichols and His Five Pennies, "I Got Rhythm," in *Rarest Brunswick Masters, 1926–1931: Red Nichols and His Five Pennies* (MCA Records, [LP] MCA 1518, 1982).

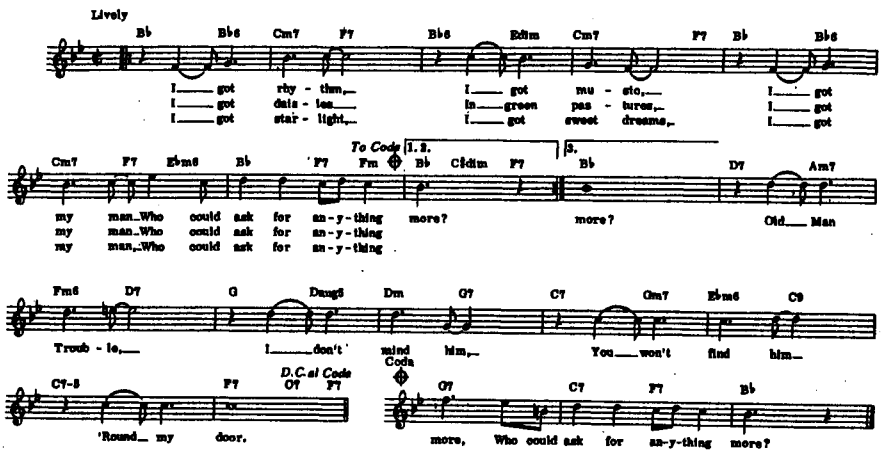
9. Fred Rich and His Orchestra, "I Got Rhythm," in *1930—You're Driving Me Crazy: Portrait of a Year in Music* (Phontastic, [CD] PHONT CD 7618, 1992).

10. A fairly comprehensive list of recordings of "I Got Rhythm" from 1930 to 1942, based on Brian Rust's *The American Dance Band Discography 1917–1942* (New Rochelle, N.Y.: Arlington House, 1975), is included in Crawford's "George Gershwin's 'I Got Rhythm,'" 222–25.

His Orchestra on 30 June 1932, however, we encounter a new approach to the song.[11] In this arrangement, the verse was eliminated completely, as was the role of the vocalist. The result is something very much like the version of "I Got Rhythm" that can be found in jazz fake books, such as shown in example 5.2.[12] After 1932, the number of recordings that include vocals drops off precipitously, and "I Got Rhythm" became a tune used by dance bands for instrumental improvisation—a jazz "standard"—as may be observed in recordings by Red Norvo and His Swing Sextette in 1936 or Chick Webb and His Little Chicks in 1937.[13] Seen from the perspective of these recordings, George Gershwin's own performance of "I Got Rhythm" as a set of variations for piano and orchestra is somewhat anomalous, for there is little if any sense of improvisation, nor is this music intended for dancing.[14]

From 1932 to 1942, recorded performances are remarkably consistent in their approach to "I Got Rhythm": all use the same thirty-four-measure form, all make

EXAMPLE 5.2 Jazz fake book version of George and Ira Gershwin's "I Got Rhythm." "I Got Rhythm," words and music by George Gershwin and Ira Gershwin, © 1930 WB Music Corp. (Renewed). All rights reserved. Used by permission. Warner Bros. Publications U.S. Inc., Miami, FL 33014.




11. Don Redman and His Orchestra, "I Got Rhythm," in *Don Redman and His Orchestra, 1931–1933* (Classics, [CD] 543, 1990).

12. Jazz "fake books" typically contain hundreds of songs. Generally, as is true of the piece in ex. 5.2, each is given as a melody, along with chord symbols that outline a basic harmonization and some minimal instructions for performance; the words for the songs may or may not be included. These materials allow players familiar with jazz performance practice to give an acceptable rendering of a song even if they have never played it before—that is, they can "fake" their way through a performance.

13. Red Norvo and His Swing Sextette, "I Got Rhythm," in *Red Norvo* (Time-Life Records, [LP] STL-J14, 1980), originally recorded 16 March 1936; Chick Webb and His Little Chicks, "I Got Rhythm," in *Spinnin' the Webb: Chick Webb and His Orchestra* (GRP Records, [CD] 513678L, 1994), originally recorded 21 September 1937.

14. George Gershwin, "I Got Rhythm," in *Gershwin Performs Gershwin: Rare Recordings 1931–1935* (Musical Heritage Society, [CD] 512923A, 1991), originally recorded 19 February 1934.

some reference to the original melody, and all preserve the affect and tempo of the original. However, an arrangement recorded by the Jimmy Dorsey Orchestra on 28 November 1942 signaled yet another approach to "I Got Rhythm," for after employing the customary thirty-four-measure format for the head, the band switched to a thirty-two-measure format for the solos.[15] In truth, this was not the first time the musical materials of "I Got Rhythm" had been associated with a thirty-two-measure form. Early on, the song had been used as the basis for a number of other compositions.[16] These derived tunes typically retain the harmonic framework, overall form, and affect of the original song but substitute a new melody and give the whole a new title. In the case of "I Got Rhythm," derived tunes allowed jazz players to keep what was useful from the song—its fast tempo, lively affect, and straightforward harmonic structure—yet escape the limitations of its simple, predominantly pentatonic melody. The earliest recorded tune derived from "I Got Rhythm," Sidney Bechet and the New Orleans Feetwarmers' "Shag" (recorded 15 September 1932), did retain the thirty-four-measure form of "I Got Rhythm," with its two-measure tag at the end of the final A section.[17] However, Fletcher Henderson's "Yeah Man," which appeared the following year, did away with the two-measure tag, yielding a thirty-two-measure form in which all of the A sections were of equal length.[18] There were two advantages to this change. First, the excision of the two-measure tag made the connection between "Yeah Man" and "I Got Rhythm" more obscure: "Yeah Man" could sound a bit more like a brand new tune. Second, jazz practice of this period was beginning to embrace longer solos. The 34-measure form of "I Got Rhythm," it will be recalled, ends with a two-measure tag added to its final A section. A musician who wanted to take two "choruses" in succession would consequently find the regular flow of eight-measure phrases in the background harmonic structure interrupted in midstream with the appearance of the ten-measure A' section. When a musician took two "choruses" on the 32-measure form, however, the succession of eight-measure phrases was constant: the improvisation could stretch seamlessly across the two refrains, allowing the player to build longer melodic lines and to shape the dynamic structure of the solo over larger spans of music. Although after "Yeah Man" the tunes derived from "I Got Rhythm" were invariably based on the thirty-two-

15. Jimmy Dorsey and His Orchestra, "I Got Rhythm," in *Jimmy Dorsey and His Orchestra*, Vol. 2: *1942–44* (Hindsight Records, [LP] HSR-153, 1980). For performances with an emphasis on improvisation that are guided only by a relatively informal arrangement, the practice among jazz performers is to state the tune at the beginning and end of the performance. In these cases, the tune functions as the "head." The head is typically followed by solos over the harmonic framework of the tune, which are, in turn, followed by a concluding statement of the tune.

16. Richard Crawford calls such derived tunes "contrafacts," borrowing the term from one of several medieval practices to which it is linked. The specific practice invoked is one in which new texts are grafted on to preexistent music, which was itself freely adapted. See Crawford, "George Gershwin's 'I Got Rhythm,'" 225, 228–29.

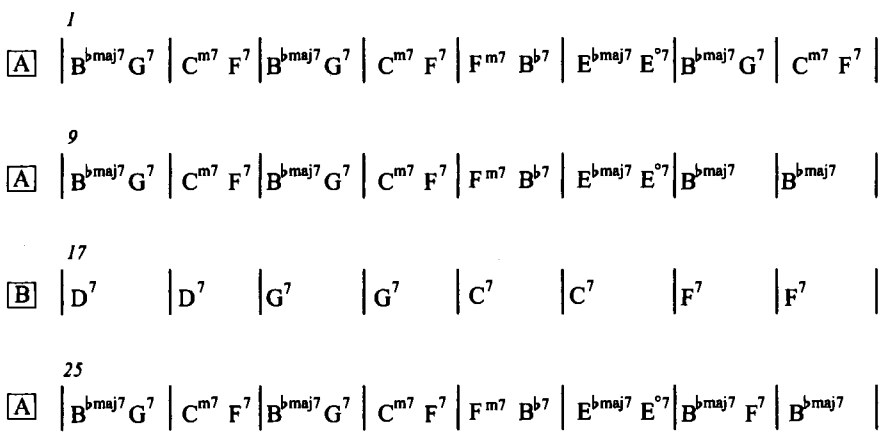
17. Sidney Bechet and the New Orleans Feetwarmers, "Shag," in *Sidney Bechet* (Time-Life Records, [LP] STL-J09, 1980).

18. Fletcher Henderson and His Orchestra, "Yeah Man," in *Fletcher Henderson, Developing an American Orchestra, 1923–1937* (Smithsonian Collection, [LP] R 006, 1977), originally recorded 18 August 1933.

measure form, it was only around 1942 that this form was brought to "I Got Rhythm" itself.

At first, the thirty-two-measure form was used only for solos—the head retained its thirty-four-measure form, complete with the two-measure tag (as in the Jimmy Dorsey arrangement). By the mid-1940s, however, even the tag for the head had fallen away. When Charlie Parker recorded "I Got Rhythm" with a group of all-star performers on 22 April 1946, the thirty-two-measure form was used for the head, as well as for the twenty-three choruses that followed.[19] The transformation of "I Got Rhythm" into a platform for improvisation is even more evident in a recording made two years earlier by Lester Young and the Kansas City Six. In this performance, not only is the thirty-two-measure harmonic pattern used throughout, but also the melody of the song is never stated.[20] In this form—that is, as a thirty-two-measure, AABA harmonic structure—"I Got Rhythm" came to be one of the basic elements of the jazz player's musical vocabulary: the "Rhythm changes" shown in example 5.3.[21]

EXAMPLE 5.3 Scheme for "Rhythm changes" (based on the harmonic structure of the thirty-two-measure form of "I Got Rhythm"). "I Got Rhythm," words and music by George Gershwin and Ira Gershwin, © 1930 WB Music Corp. (Renewed). All rights reserved. Used by permission. Warner Bros. Publications U.S. Inc., Miami, FL 33014.




19. Charlie Parker and others, "I Got Rhythm," in *Bird: The Complete Charlie Parker on Verve* (Verve/Polygram Classics, [CD] 837 142-2, 1988).

20. Lester Young with the Kansas City Six, "I Got Rhythm," in *The Tenor Sax: Lester Young, Chu Berry and Ben Webster: The Commodore Years* (Atlantic, [LP] SD 2-307, 1973), originally recorded 28 March 1944. That the performance by Young's group resulted from an arrangement employed for successive performances rather than happenstance is suggested by an alternate take of the same tune recorded the same day (included on this album), which uses precisely the same arrangement.

21. The term "changes" refers to a conceptual structure that consists of a succession of harmonies correlated with a metric and hypermetric framework. Knowing this structure—knowing where and when the harmonies change—is essential for creating convincing improvisations.

On the importance of "Rhythm changes" as a basic harmonic prototype, see Paul F. Berliner, *Thinking in Jazz: The Infinite Art of Improvisation*, Chicago Studies in Ethnomusicology (Chicago: University of Chicago Press, 1994), 76-78.

From the evidence provided by recordings, then, what counted as "I Got Rhythm" changed from 1930 to 1946 as the song was transformed from a Broadway or popular song, to a jazz standard, to a platform for improvisation. This précis is not quite accurate, however, for even as new conceptions of the song emerged, older ones remained in circulation. Benny Goodman, for instance, continued to use the thirty-four-measure form of the refrain even after others had adopted the thirty-two-measure form;[22] productions of the musical (including the 1943 movie *Girl Crazy*, starring Mickey Rooney and Judy Garland) continued to use the entire song after most other performances had omitted the verse.[23] From almost the moment of its introduction, there were multiple "Rhythm"s in circulation, and what belonged under the rubric "I Got Rhythm" depended on where you stood in the stream of American popular culture.

## Categorizing "I Got Rhythm"

Collected in table 5.1 are all of the versions of "I Got Rhythm" discussed thus far. How might someone go about categorizing these various versions were they encountered in the course of daily life rather than as a pre-selected list? That is, assuming one were not so fortunate as to have each performance identified as "a performance [or recording, or rendering] of 'I Got Rhythm,'" what would be the basis for including any or all of these in the category *performances of "I Got Rhythm"*? As shown in chapter 1, categories like this are organized around a conceptual model that sets up an idealized representation of what counts as a typical member of the category.[24]

As a place to begin, let us consider a conceptual model focused on the attributes of the song as it first appeared in *Girl Crazy* (leaving aside the variants created as the song was reprised over the course of the musical). The model (which I shall call the "pop-music" model, for reasons that will become clear below) is diagrammed in figure 5.1 and consists of five correlated conceptual elements, characterized as the things necessary for an adequate performance of the song. Two aspects of this model should be clarified. First, although the verses of popular songs were often omitted in performance, I take the fact that composers still bothered to write verses and that the verse was included, in some form or another, in early recordings as evidence of its importance to the identity of the song. Second, my distinction between variation and improvisation is based on the extent rather than the presence of musical materials not specified by the score: if these materials are relatively limited, I would tend to regard them as a product of variation; if they are extensive, I would tend to regard them as a product of improvisation. Where to draw the line involves another set of conceptual models for performance practice; nonetheless, the distinction is relatively clear in early recordings.

22. Benny Goodman Sextet, "I Got Rhythm," in *Benny Goodman* (CBS Records, [LP] P5 15536 CP, 1981), originally recorded 18 September 1945.

23. Although the plot line for the movie is completely different from that of the musical, the musical numbers are the same.

24. By "categories like this" I mean not categories involving musical works specifically, but the categories we use in daily life generally.

TABLE 5.1. Select Performances of "I Got Rhythm"

| *Performers* | *Date of Recording* |
|---|---|
| Red Nichols and His Five Pennies | 23 October 1930 |
| Fred Rich and His Orchestra | 29 October 1930 |
| Don Redman and His Orchestra | 30 June 1932 |
| George Gershwin | 19 February 1934 |
| Red Norvo and His Swing Sextette | 16 March 1936 |
| Chick Webb and His Little Chicks | 21 September 1937 |
| Jimmy Dorsey and His Orchestra | 28 November 1942 |
| Lester Young with the Kansas City Six | 28 March 1944 |
| The Benny Goodman Sextet | 8 September 1945 |
| Charlie Parker and others | 22 April 1946 |

Applying this model to the performances listed in table 5.1 yields the typicality effects characteristic of Type 1 categories. Most typical would be the performances by Red Nichols's and Fred Rich's groups; performances from the original production of the musical (which are not included in the list of recordings) would of course also be reckoned as highly typical.[25] By contrast, the performance led by Gershwin, which drops the verse and lyrics, would not conform as well with the model. Even less typical would be the performances led by Redman, Norvo, Webb, and Goodman, for all these add a substantial emphasis on improvisation. Finally, on the fringes of the category would be the performances led by Dorsey, Young, and Parker, for these not only omit the verse and lyrics but also use a thirty-two-measure form of the refrain as a platform for extended improvisations.

As we see in this example, conceiving of various renditions of a song as a Type 1 category organized around a conceptual model offers a way to explain the peculiar ontological status of musical works. Our sense that there is only one thing that should be called "I Got Rhythm" is supported by the idealization basic to the model and to the model's role as the conceptual anchor for the category. At the same time, our ability to accept as members of the category things that do not fully correspond to this model shows how the model functions as a guide to understanding the structure of the category as a whole. Although Charlie Parker's "I Got Rhythm" is not like Ethel Merman's "I Got Rhythm," it shares enough attributes with the latter that we will not want to exclude it from the category outright. Further, the model can provide a basis for describing why we would grant Parker's performance a liminal rather than a central status within the category—we might say something along the lines of "He uses too much of that crazy improvisation, so he really isn't playing 'I Got Rhythm.'"

25. There was no recording of the original production of *Girl Crazy*. However, there was a painstaking recreation of the 1930 production undertaken by Elektra Nonesuch in 1990, and Lorna Luft's performance there (in the part of Kate Fothergill originated by Merman) agrees in most respects with the version of "I Got Rhythm" recorded by Merman in 1947. See George Gershwin and Ira Gershwin, *Girl Crazy* (Elektra Nonesuch, [CD] 79250-2, 1990), recorded 26–28 February 1990.

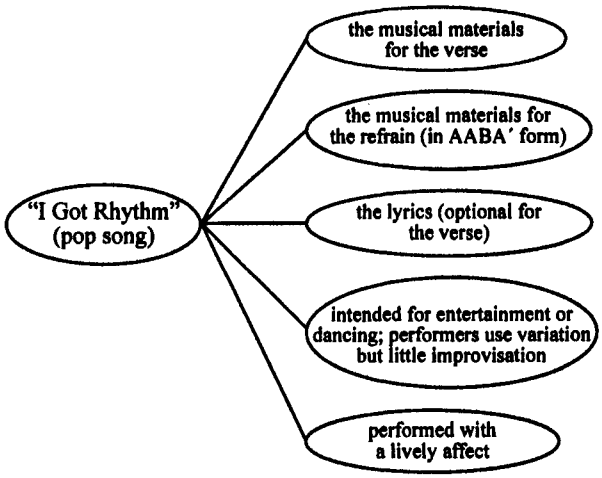



FIGURE 5.1 Pop-music model for the category *performances of "I Got Rhythm"*

Useful as this approach to musical ontology is, it demands at least a bit of further elaboration: we need to find a way to describe how the conceptual model for a musical work changes over time. After all, while Parker's 1946 performance might have been regarded as unusual by some, it is clear from the musical practice of his contemporaries that it was far from exceptional. Here Richard Crawford's account of three different approaches taken to "I Got Rhythm" during the years 1930–1942 provides a helpful guide.[26] The first approach, in which "I Got Rhythm" was treated as a popular song, is similar to that assumed by the pop-music model diagrammed in figure 5.1. The second approach described by Crawford is one in which "I Got Rhythm" is treated as a jazz standard. A conceptual model that fits this approach is diagrammed in figure 5.2 and consists of five things necessary for an adequate performance.[27] Adopting this model would, of course, change what counted as a typical version of the tune. Now the performances led by Redman, Norvo, Webb, and Goodman would be central to the category, and the performances led by Dorsey, Young, and Parker only a bit less typical. Far more atypical would be the performance led by Gershwin, with its relative lack of emphasis on improvisation, and the performances by Red Nichols's and Fred Rich's groups, with their inclusion of the verse.

The third approach to "I Got Rhythm" noted by Crawford was as a musical structure that served as the basis for various derived tunes. I would like to generalize this somewhat to accommodate the slightly expanded time frame represented by the recordings collected in table 5.1. After 1942, "I Got Rhythm" served not only as the basis for derived tunes but also as a framework for improvised performance—

26. Crawford, "George Gershwin's 'I Got Rhythm,'" 218.

27. Explicit mention of the melody in this conceptual model is intended as a hedge against confusing "I Got Rhythm" with any of the tunes derived from it.

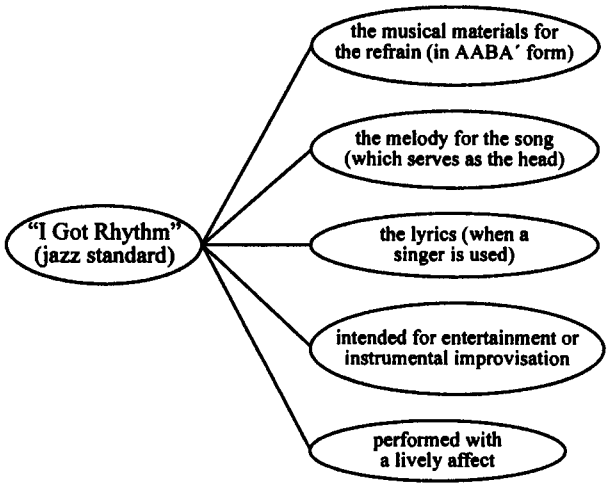



FIGURE 5.2 Jazz-standard model for the category *performances of "I Got Rhythm"*

a framework embodied in the "Rhythm changes" of example 5.3. This leads us to a third model for "I Got Rhythm," which is diagrammed in figure 5.3. This "improv" model, conceived as a skeletal vehicle for improvisation and composition, consists of four things necessary for an adequate performance. Once again, what would count as a typical version of the tune changes. From the perspective provided by the improv model, the performances led by Young and Parker would be most typical, with that led by Dorsey following close behind. The performances led by Redman, Norvo, Webb, and Goodman, with their allegiance to the original form and melody of Gershwin's song, would now be regarded as less typical, and the performances led by Gershwin, Nichols, and Rich would have to be considered as thoroughly atypical.

Let us now return to the matter of how the conceptual model for a musical work changes over time. If we take these three models as an historical sequence—the pop-music model followed by the jazz-standard model, which is followed by the improv model—we can see that each successive conceptual structure retains some of the elements of the previous structure and either modifies or drops others.

- The jazz-standard model retains from the pop-music model the elements concerning affect and the materials of the refrain. It changes those relating to function and lyrics, and it adds an element specifying the presence of the melody. (The latter is implicit in the "materials of the refrain" inherited from the pop-music model, but it is made explicit here).
- The improv model retains from its jazz-standard counterpart the element concerning affect. It changes the components of the model associated with the refrain, melody, and function, and it drops any requirement for the lyrics.

As is evident from table 5.1, these are not the only models possible, nor should my descriptions of the model be taken as comprehensive or final. One could easily

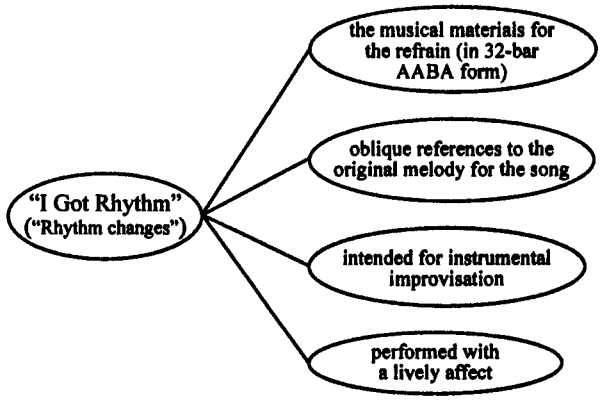



FIGURE 5.3 Improv model for the category *performances of "I Got Rhythm"*

imagine idiosyncratic or transitional models that blend features of those outlined here. Thus a model based on the performance led by George Gershwin, for instance, would borrow some features from the pop-music model and some from the jazz-standard model; a model based on the arrangement played by the Jimmy Dorsey band would incorporate some aspects from the jazz-standard model and some from the improv model. And were such models taken as the basis for our determinations as to what belonged in the category *performances of "I Got Rhythm,"* there would of course be changes in what counted as a typical or atypical member of the category.

When people share the conceptual model for a song (realizing that all such sharing is approximate), they will tend to make similar judgments about what counts as a typical or an atypical rendering of the song (or whether a succession of sounds should even be counted as an instance of the song). As I proposed in chapter 3, sharing knowledge in this way is one of the bases of culture. This becomes especially clear when we consider other roles played by the model for a musical work: as a set of cognitive resources for performance, as a basis for negotiations on how musical practice should proceed, and as a focus for how musical practice can be regulated.

## Conceptual Models and Performance

Up to this point, we have considered music that is part of a tradition of performance at least partially regulated by the score. But what of performance traditions in which scores play no part—that is, music that is part of an aural tradition? Where scores (or similar artifacts) are absent, performers must rely on a store of knowledge about music and music making, knowledge often gathered through a long and painstaking process that starts in earliest childhood and continues for many years.

Constantin Brailoiu was perhaps the first to suggest that musical performance within aural traditions relies on stored cognitive constructs. As a way of explaining the source of the divergent performances folk musicians gave of the "same" song, Brailoiu proposed that such performances were ephemeral incarnations of an ideal

archetype stored in the mind of the musician.[28] In his work on African music, Simha Arom adopted Brailoiu's idea, but he called the archetype a "model."[29] James Cowdery, whose research has focused on Irish music, proposed a similar cognitive basis for performance practice, but one derived from the tradition of inquiry centered on the tune family.[30] Cowdery called the construct a "tune model" and conceived of it as a generating pattern in the mind of the performer or musical community.[31] Finally, James Porter has developed detailed descriptions of a succession of conceptual performance models for the changing versions of the song "My Son David" sung by the Scottish singer Jeannie Robertson between 1953 and 1960.[32] For Porter, the conceptual performance model represented a holistic, self-contained concept of a particular song in the mind of the singer.[33] This model enabled the singer to give successive but nonidentical performances of the "same" song on a number of occasions, but the model was nonetheless subject to change over time—that is, the "same" song gradually became a different song as the conceptual performance model changed.

The idea behind all of these approaches is that performers who work within aural traditions base their performance of a given tune on a cognitive construct that is stored in memory and that represents essential features of that tune. In their basic features, such constructs are essentially the same as the conceptual models I described in chapter 3. I have used these same models in this chapter to explain how individuals determine whether a particular performance is of a specific tune and how typical that performance is. The difference that distinguishes my models from those of Brailoiu, Arom, Cowdery, and Porter is that the sort of constructs they pro-

28. The relevant texts in translation and the original are as follows: "Lacking an unchallengeable [musical] text, we must admit that we never collect more than variants and that latent in the singers' minds lives an ideal archetype of which they offer us ephemeral incarnations." ["A défaut de tout texte irrécusable, force nous est d'admettre que nous ne recueillons jamais que des variantes et que, dans l'esprit des chanteurs, vit, d'une vie latente, un archétype idéal, dont ils nous offrent des incarnations éphémères."] Constantin Brailoiu, "Le Folklore musicale," in *Musica aeterna: La Vie et la production musicales de tous les temps et de tous les peuples, en tenant compte particulièrement de la Suisse, de la Belgique, de la France et de la musique de nos jours*, ed. Gottfried Schmid (Zurich: M. S. Metz, 1949), 2: 319.

29. Simha Arom, *African Polyphony and Polyrhythm: Musical Structure and Methodology*, trans. Martin Thom, Barbara Tuckett, and Raymond Bond (Cambridge: Cambridge University Press, 1991), 148.

30. The idea of a tune family was first proposed by Samuel P. Bayard in his classic essay "Prolegomena to a Study of the Principal Melodic Families of British-American Folk Song," *Journal of American Folklore* 63 (1950): 1–44.

31. James R. Cowdery, *The Melodic Tradition of Ireland* (Kent, Oh.: Kent State University Press, 1990), 33.

32. James Porter, "Jeannie Robertson's *My Son David*: A Conceptual Performance Model," *Journal of American Folklore* 89 (1976): 7–26; and idem, "Context, Epistemics, and Value: A Conceptual Performance Model Reconsidered," in *Selected Reports in Ethnomusicology: Issues in the Conceptualization of Music*, 7, ed. James Porter and Ali Jihad Racy (Los Angeles: University of California, Department of Ethnomusicology, 1988), 69–97. Porter's thoughtful appraisal of what might be involved in a conceptual performance model goes well beyond the primitive models I sketch here and should be taken into account in any further application of theories of categorization to ethnographic research.

It should be noted that while Arom, Cowdery, and Porter all use the term "model" for their hypothetical cognitive constructs, they do so without reference to any of the cognitive science research on models or related constructs.

33. Porter, "Context, Epistemics, and Value," 71.

pose are used not for categorizing musical performances, as I attempt to do, but for producing them. In truth, we have already seen conceptual models used in this way, for Jeff's conceptual models for "Twinkle" not only helped him organize the Montessori bells in particular patterns but also guided his actions when he performed the song. In the present context, however, the network of knowledge with which we are concerned is far more extensive than that with which Jeff was engaged. The skills required for an adequate performance of a tune like "I Got Rhythm" (relative to the performance tradition described above) are extensive: a trumpet player, for instance, must know where to find the notes for the tune on her instrument, how to produce these notes in the right order, how to shape this production to create a convincing version of the tune, how to coordinate her playing with that of other members of an ensemble, and so on.[34] For performers, then, there will be a number of conceptual models related to the various challenges of creating musical sound, as well as the conceptual model for a particular piece.

The complexity of this network of knowledge emphasizes once again the importance of the global models relative to which local models are framed. With regard to the conceptual model for a musical work, the global model speeds the process of learning the conceptual model for a new tune and also provides a framework for discourse among musicians. For instance, among performers the jazz-standard model for "I Got Rhythm" (which was diagrammed in fig. 5.2) is framed relative to a global model for jazz practice, which includes the five elements diagrammed in figure 5.4. A comparison of figure 5.4 with figure 5.2 shows that each generic element in the former has its more specific correspondent in the latter. Confronted for the first time with a tune like "I Got Rhythm," a musician who has internalized the global model will already have a start on assembling a local conceptual model to use as a guide in performing the tune. The global model also serves as a frame for discourse (both verbal and musical) about local models. That is, approaching "I Got Rhythm" as a song that belongs to the performance practice of jazz means that features such as its harmonic structure, melody, and character are what are relevant, rather than whether there is a definitive score on which to base performance.

Conceptual models, then, are used not only to structure our understanding of

34. Of necessity, this network of knowledge will include embodied knowledge: what it feels like to make musical sounds and to coordinate this activity with other physical activities. Richard Powers touches on the importance of such knowledge in his novel *Galatea 2.2*, which centers on a hypothetical experiment to replicate aspects of human intelligence on a vast and spreading neural network, implemented via parallel processing on a massive supercomputer. At one point, the implementation, named Helen, asks "How do you sing?" The narrator responds by singing a fragment of a song he learned as a child: "Bounce me high, bounce me low, bounce me up to Jericho." To his amazement, the implementation not only absorbs this song but can later be heard singing it, over and over, like a child trying out her voice. Yet "sing" is not quite the right word, for Powers's narrator realizes that there is something missing: "Helen did not sing the way real little girls sang. Technically, she almost passed. Her synthesized voice skittered off speech's earth into tentative, tonal Kitty Hawk. Her tune sounded remarkably limber, given the scope of that mechanical tour de force. But she did not sing for the right reasons. Little girls sang to keep time for kickball or jump ropes. . . . Helen didn't have a clue what keeping time meant, never having twirled a jump rope, let alone seen one." Richard Powers, *Galatea 2.2* (New York: Farrar Straus Giroux, 1995), 205.

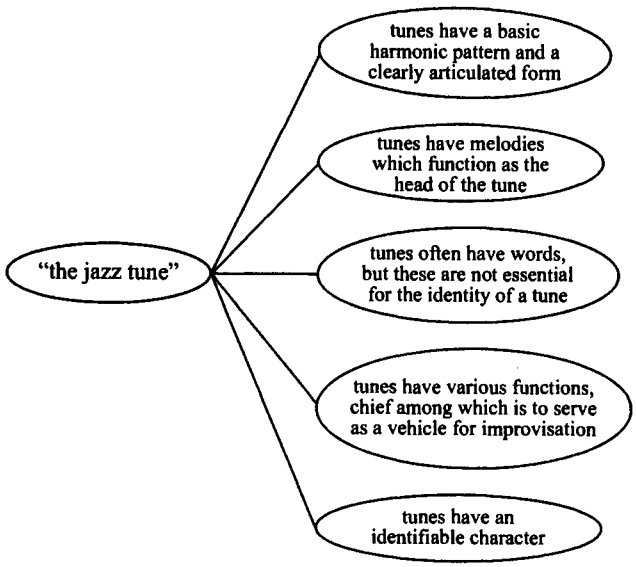



FIGURE 5.4 Partial representation of the global model of jazz practice

what we hear but also to create the things we hear. In both cases, the model is part of a network of knowledge that, through being shared by the members of a musical community and realized in sound, constitutes musical culture.

## Negotiating Conceptual Models

Musical practice is not a steady state, as may be seen in the different versions of "I Got Rhythm" collected in table 5.1. Over time, conceptual models will be modified in response to changes in cultural priorities. Rarely are such modifications the result of creative fiat by one individual (although one such case is discussed below, in connection with "Bye Bye Blackbird"). More commonly, they are the product of a gradual series of negotiations within a musical community.

The basic outlines of the process of negotiating the structure of conceptual models can be illustrated by considering two highly simplified situations. Each involves a musician who initiates the process of negotiation by suggesting changes to the constituent elements of a given conceptual model. In the first case, the negotiations occur within a group of musicians occupied with performing the same work; in the second case, the negotiations occur between the musician and an audience of non-performers.

The first situation can be conceived of as taking place in a rehearsal or a performance of the version of "I Got Rhythm" associated with the jazz-standard model. Instead of playing the usual harmonic progression for the refrain, the pianist, without comment, introduces a completely new series of harmonies for the A section and truncates the overall form to thirty-two measures. The other musicians may accept this change (by adjusting their performance to the new harmonies), reject it

(by resolutely hewing to the structure of the original model), or suggest further modifications (by offering their own interpretation of the harmonic progression). This process could also play out with the aid of language (the changed harmonic progression could be announced in advance) or symbols (a chord chart for the changed refrain could be supplied to the other musicians), or through various combinations of musical performance, language, and symbols by any or all participants.

The second situation can be conceived of as taking place during a performance of "I Got Rhythm" by a solo pianist, playing for an audience of nonperformers. Again, instead of the usual harmonic progression for the refrain, the pianist introduces a new series of harmonies and truncates the overall form to thirty-two measures. The audience may respond by accepting the change (through applause or a similar activity), rejecting it (by shouting catcalls or throwing bottles), or suggesting further modifications (by shouting "More!" or "Too much!"). (Of course, it is quite probable that an audience would not even notice changes of this sort, but for the sake of the argument, I shall assume that it did.)

In each of these imaginary situations, some elements of the conceptual model for the tune are changed and some are retained. The conceptual models thus form a framework for negotiation—that is, negotiation is conducted in terms of the conceptual model for the tune. Series of such negotiations then lead to the more substantive changes in conceptual models of the sort suggested by my account of how such models might be linked in a historical succession. Of course, the imaginary situations I have described are highly simplified: as the works of Ruth Finnegan, Paul Berliner, and Ingrid Monson have demonstrated, actual negotiations among musicians are much more complex than those outlined here.[35] Musical discourse is rich and creative, and it includes comedy and irony, as well as meaningful exchanges about the business at hand. Nonetheless, what makes this communication possible is the knowledge shared among musicians, knowledge that is organized by conceptual models.

## Limits on Negotiation

The mutability of conceptual models under the social and cultural pressures of negotiation has the potential to destabilize cultural knowledge by transforming it until it becomes alien and unfamiliar. Under certain circumstances, the interconnectedness and dispersion characteristic of cultural knowledge can counteract destabilization, a process noted by Mantle Hood.[36] Under other circumstances, however, the challenge to stability may be significant. In the case of musical practice, such challenges might come from a number of sources: from the innovations of influen-

35. Ruth Finnegan, "The Relation between Composition and Performance: Three Alternative Modes," in *The Oral and the Literate in Music*, ed. Yoshihiko Tokumaru and Osamu Yamaguti (Tokyo: Academia Music, 1986), 73–87; Berliner, *Thinking in Jazz*; Ingrid Monson, *Saying Something: Jazz Improvisation and Interaction*, Chicago Studies in Ethnomusicology (Chicago: University of Chicago Press, 1996).

36. Mantle Hood, "The Reliability of Oral Tradition," *Journal of the American Musicological Society* 12 (1959): 201–09.

tial individuals, from contact with other cultures, from a sudden increase in the complexity of musical practice, or from changes in technology. Faced with this sort of challenge, a musical community may develop strategies to eliminate or control the instability of the conceptual models upon which the practice of that community is based.

One common strategy is to represent key structural elements of a local conceptual model through artifacts such as scores or recordings. Although the score illustrations used for "I Got Rhythm" carry the aura of authority, they actually follow from the conceptual models for the song: a score is an artifactual manifestation of the elements of the conceptual model deemed most relevant to the musical practice of which the model is a part, created as a means of stabilizing the model. The score for the original version of "I Got Rhythm" (ex. 5.1) can thus be seen as a reflection of the conceptual model for pop songs (fig. 5.1). Such songs invariably had a verse, even if it was only rarely performed. Since the songs were often performed by musicians with only limited score-reading abilities, chord symbols (which often do not match the chords indicated by the music notation) are almost always included, and they serve to indicate the minimal materials necessary to accompany the voice. The jazz fake book rendering of "I Got Rhythm" (ex. 5.2) does away with the verse and any specifics about how the accompaniment is to be realized; it reflects the essential features of the jazz-standard model (fig. 5.2). And the "Rhythm changes" of example 5.3 preserve only the bare essentials, cryptic to anyone not schooled in jazz but more than enough to serve musicians as a starting point for improvisation, provided they are familiar with the concepts and relations embodied in the improv model (fig. 5.3). Thus, both in its production and interpretation, musical notation reflects (rather than generates) the local and global conceptual models that constitute musical practice.[37]

The idea that musical notation could serve as a way to control (rather than simply enable) musical practice is not a new one, having been proposed by Charles Seeger over fifty years ago.[38] With respect to Western practice, the limits on negotiation that follow from this reliance on musical notation operate on two levels. First, any given instance of musical notation, as an embodiment of "the work of music," limits negotiation to those aspects of music deemed "nonessential" to the work (such as the specifics of dynamics, timbre, and tempo): these are the things left relatively ambiguous by the score.[39] Second, the *terms* for negotiation become artifactual: any aspect of musical practice that would vie for determining the identity of a musical work must be such that it can be expressed artifactually. Arguments over

37. Leo Treitler has made similar points about the role of notation in medieval music. See Treitler, "Transmission and the Study of Music History," in *Report of the Twelfth Congress, Berkeley 1977*, ed. Daniel Heartz and Bonnie Wade (Kassel: Bärenreiter-Verlag; Philadelphia: The American Musicological Society, 1981), 202–11; and idem, "History and the Ontology of the Musical Work," *Journal of Aesthetics and Art Criticism* 51 (1993): 483–97.

38. Charles Seeger, "Oral Tradition in Music," in *Funk and Wagnalls Standard Dictionary of Folklore, Mythology, and Legend*, ed. Maria Leach and Jerome Fried (New York: Funk and Wagnalls, 1950), 2: 825–29.

39. Nelson Goodman observes that where the score attempts to capture what I have called the "nonessential" aspects of music it ceases to be, in his terms, notational. See Goodman, *Languages of Art*, 183–86.

musical identity (such as those that play out in discussions about the *Urtext* for a score) are arguments about the status and interpretation of artifacts, and any competing account of musical identity must be expressed in artifactual terms.

It bears mention that the authority of notation is not monolithic: most musicians accept notation as something to be interpreted and consider the way it is to be used to create a musical performance a matter for negotiation. And even where notation has been introduced with the explicit intent of eliminating variation within musical practice, it may be circumvented. As Ruth Davis has shown, the modified Western notation applied to Tunisian art music beginning in the 1930s for the purpose of standardizing its melodic tradition has been turned into a creative tool for reinterpreting and redefining that tradition.[40] Notation has thus slowed, but not eliminated, the fluidity of Tunisian musical practice.

## Categorization, Musical Ontology, and Cultural Knowledge

By the late 1940s, as we have seen, there were multiple versions of George and Ira Gershwin's "I Got Rhythm" in circulation. There was the show stopper and popular tune from the Broadway musical, the "hot" jazz tune played by any number of dance bands, and the relatively abstract platform for improvisation used as the basis for any number of derived tunes by jazz composers and for extended improvisations by smaller ensembles. Confronted with this multiplicity of "I Got Rhythm"s, how might a listener discriminate between them? Based on what we now know about processes of categorization, we can say that our listener would almost certainly recognize that, while all these versions belonged in the category *performances of "I Got Rhythm,"* not all would represent the category equally well. Depending on the conceptual model around which the category was organized, some versions would be regarded as quite typical of the category, others as less typical, and still others in danger of being excluded from the category altogether. Following Richard Crawford's description of three different approaches to "I Got Rhythm," I proposed three models—the pop-music model, the jazz-standard model, and the improv model—which would yield different interpretations of typicality when applied to a range of performances of the tune. The knowledge represented by such models, when combined with knowledge about musical technique and framed by global conceptual models applicable to an entire repertoire, can also serve as the basis for musical performance within aural traditions. Because they are retained by individual members of a musical community, who may wittingly or unwittingly change the model as they use it, conceptual models are subject to change over time. When such changes are brought to the attention of other members of the musical community, a series of negotiations about what constitutes the model may ensue. At times, changes to conceptual models may be perceived as a threat to the integrity of a musical community, and the community may then seek either to halt or to retard change by placing limits on what can be negotiated.

40. Ruth Davis, "The Effects of Notation on Performance Practice in Tunisian Art Music," *The World of Music* 34 (1992): 111.

Given the various conceptual models we have considered and the ways these models are interpreted and correlated with other types of knowledge, it is clear that the network of knowledge within which any musical work (or similar construct) is embedded is a complex one. And yet we have thus far restricted ourselves to an almost purely musical world. Matters only stand to become more complicated when we contemplate how this world might connect with the cultural context within which musical practice is itself embedded. In the next section, we encounter some of these complications and gain a sense of the richer understanding of musical practice to which they lead, as we explore some of the different ways "Bye Bye Blackbird" was conceived during the middle of the twentieth century.

## "BYE BYE BLACKBIRD," SIGNIFYIN(G), AND CULTURAL KNOWLEDGE

On 5 June 1956, Miles Davis, John Coltrane, Red Garland, Paul Chambers, and Philly Joe Jones went into Columbia's 30th Street Studio in New York and recorded three tunes: "Tadd's Delight," "Dear Old Stockholm," and "Bye Bye Blackbird."[41] Although the recording session took place at a time when Davis was actively shaping the sound and repertoire of his quintet, the choice of "Bye Bye Blackbird" was nonetheless somewhat odd. Unlike the other two tunes, "Bye Bye Blackbird" had never been a part of the jazz repertoire, but was simply an upbeat popular song usually played as a foxtrot. Recorded by Davis as a ballad, it acquired a melancholy and an irresistible groove that were thoroughly exceptional.

The choice was made odder still by the racial overtones associated with the tune. First published in 1926, "Bye Bye Blackbird" played a role in the racially charged election campaign for the mayoralty of Chicago in 1927. Big Bill Thompson, a Republican who had successfully cultivated the African American vote to become mayor of Chicago in 1915, returned in 1927 to challenge Democratic mayor William E. Dever. As part of their strategy to fight off this challenge, the Democrats appealed to racial prejudice, voiced specifically through "Bye Bye Blackbird." The ploy ultimately backfired, for it galvanized African American support for Thompson:

> When some of the Dever supporters sent calliopes through the streets piping the strains of "Bye Bye Blackbird," and spreading a circular which displayed a trainload of Negroes headed from Georgia with Thompson as pilot of the train, and the caption: "This train will start for Chicago, April 6, if Thompson is elected," the obvious answer of the colored leaders was: "Elect Big Bill or it's going to be bye-bye blackbirds in Chicago."[42]

41. These three tunes, together with "'Round Midnight," "Ah-leu-cha," and "All of You," were released on Miles Davis, *'Round about Midnight* (Columbia, [CD] CK 40610, 1987), originally recorded 27 October 1955 and 5 June and 10 October 1956. There are extensive transcriptions and a discussion of this recording of "Bye Bye Blackbird" in Berliner, *Thinking in Jazz*; see esp. pp. 678–88.

42. Harold F. Gosnell, *Negro Politicians: The Rise of Negro Politics in Chicago*, with an introduction by James Q. Wilson (Chicago: University of Chicago Press, 1966; originally publ. in 1935, with an introduction by Robert Park, Social Science Studies, 35), 54–55. There is a discussion of a similar cartoon from around the same time that invokes the phrase "Bye Bye Blackbird" in Lloyd Wendt and Herman Kogan, *Big Bill of Chicago* (Indianapolis: Bobbs-Merrill, 1953), 256.

In a similar vein, Jack Chambers reports that during the 1930s, the few African Americans who participated in activities dominated by whites, like the semi-professional National Football League, had the song sung to them by the all-white spectators.[43]

What seems to be reflected in Davis's adoption of the tune is the sort of double-voiced discourse first described by Mikhail Bakhtin: on one level of discourse, Davis's "Bye Bye Blackbird" is an aging pop song, with racial overtones that turn against its performers; on another level, it is a deeply felt rewriting—"felt" both in terms of melancholy and groove—that makes a mockery of the simplistic messages of the song, not least by rising above them. This process of repetition (in the sense of returning to an earlier "text") and revision is associated with a kind of intertextuality Henry Louis Gates Jr. called "Signifyin(g)," which he argued is fundamental and specific to African American rhetorical practice.[44] While Gates's primary focus was on literature, he extended his argument to other domains, music among them, and his arguments have proved compelling to musicologists. Adopting one of Gates's examples, Ingrid Monson was able to show in some detail how John Coltrane's 1960 recording of Richard Rodgers and Oscar Hammerstein's "My Favorite Things" (written in 1959) accomplishes this sort of Signifyin(g). Coltrane's interpretation completely revises the song: the focus is changed from the verse to the interludes, Rodgers and Hammerstein's waltz is transformed into a polyrhythmic extravaganza, and what was an upbeat tune becomes a brooding one.[45]

There is, however, a problem with claiming, in any simple way, that Davis is Signifyin(g) on "Bye Bye Blackbird," and that problem is profiled vividly by a recording of the tune in 1955 by Peggy Lee. In her performance, Lee adopts an exceptionally slow tempo—slower than Davis's—and completely vanquishes the upbeat sentiments of the original song.[46] Lee, too, seems to have engaged in double-voiced discourse, but would we want to say that she is Signifyin(g)? Indeed, this sort of performance practice is so common in jazz—within and without the African American community—that Signifyin(g), present everywhere, would seem to signify nothing.[47]

Two sets of issues, then, are raised by Davis's recording of "Bye Bye Blackbird." First, what is the nature of the double-voiced discourse of Signifyin(g)? Is it in any way distinct from double-voiced discourse as a whole? Second, what part does the reinterpretation of musical works—or, from the perspective developed in the pre-

43. Jack Chambers, *Milestones I: The Music and Times of Miles Davis to 1960* (Toronto: University of Toronto Press, 1983), 237.

44. Henry Louis Gates, Jr., *The Signifying Monkey: A Theory of African-American Literary Criticism* (New York: Oxford University Press, 1988), chap. 2. The parenthetical *g* in "Signifyin(g)" is meant to signal African American pronunciation, as well as to distinguish it from the more general process of "signification"; for discussion, see *The Signifying Monkey*, 46.

45. Monson, *Saying Something*, 106–21.

46. Peggy Lee, "Bye Bye Blackbird," in *Songs from* Pete Kelly's Blues (Decca, [LP] DL 8166, 1955), recorded 6 May 1955.

47. The practice of revisiting the work of other composers and musicians was also common in the music of the previous century: we thus have Beethoven Signifyin(g) on Diabelli in his Op. 120 variations and Brahms Signifyin(g) on Paganini in his Op. 35 variations.

ceding section, revisions to the conceptual model for a musical work—play in Signifyin(g)? In what follows, I begin by developing an approach to double-voiced discourse in terms of a cultural theory of information exchange proposed by the linguist Eve Sweetser. I proceed then to explore some of the conceptual models for "Bye Bye Blackbird" in circulation during the period 1926–1955, and, finally, I return to Davis's interpretation of the tune.

## Double-Voiced Discourse and a Cultural Theory of Information Exchange

Mikhail Bakhtin developed his notion of double-voiced discourse with reference to the novels of Dostoevsky, going so far as to claim that this rhetorical mode was essential to Dostoevsky's style.[48] Double-voiced discourse is actually the third of three types of discourse distinguished by Bakhtin. The first type is "direct discourse," which is essentially the unmediated disquisition of the author. The second type Bakhtin called "objectified" or "represented discourse"; its most common form is the direct speech of characters. In "double-voiced discourse" (the third type), the author takes the direct discourse of someone else and infuses it with the author's own intentions and consciousness while still retaining the original speaker's intentions. Thus two discourses and two consciousnesses are present at the same moment.[49]

With this taxonomy in place, Bakhtin further distinguished three types of double-voiced discourse. In the first, the discourse of the author and that of the other person essentially cooperate—the author inflects, but does not substantially change, the discourse of the other person. The second type of double-voiced discourse is associated with parody and irony (this is essentially the sort of discourse Monson deals with in her analysis of Coltrane's "My Favorite Things"). Here the discourse of the author is opposed to that of the other person; as Bakhtin puts it, "the second [that is, authorial] voice, once having made its home in the other's discourse, clashes hostilely with its primordial host and forces him to serve directly opposing aims. Discourse becomes an arena of battle between the two voices."[50] In the third type of double-voiced discourse, the other person's discourse remains outside the purview of the author's, which nonetheless becomes influenced by that of the other person. The result is often a hidden polemic: the author's discourse represents a reaction to the discourse of the other person, but the latter can be heard only by implication—it is never directly represented.

Each of these three types of discourse—and, most important, the three types of double-voiced discourse—can be analyzed with reference to a cultural theory of

48. Mikhail Mikhailovich Bakhtin, *Problems of Dostoevsky's Poetics*, ed. and trans. Caryl Emerson, with an introduction by Wayne C. Booth, Theory and History of Literature, 8 (Minneapolis: University of Minnesota Press, 1984), chap. 5.

49. The premise of double-voiced discourse is that the two domains of discourse remain separate and identifiable; it is thus different from the process of conceptual blending discussed previously in chap. 2 (in that blending is concerned with how elements from two distinct domains come together in a third domain), but it may be seen as a related phenomenon.

50. Bakhtin, *Problems of Dostoevsky's Poetics*, 193. I shall follow Bakhtin's practice and speak of the first voice in double-voiced discourse as that of the "other" and the second voice as that of the author.

information exchange developed by Eve Sweetser. Sweetser used the theory to explain the results of an experiment by Linda Coleman and Paul Kay that was designed to show that the English word "lie" (in the sense of "to prevaricate") functioned like a Type 1 category. That is, rather than following a clear-cut definition, some statements were regarded as typical lies, others as atypical. Coleman and Kay hypothesized that a statement is judged to be a more-or-less good example of a lie on the basis of three conditions. First, the actual statement must be false (a condition that is commonly taken as necessary and sufficient when "lie" is construed as a Type 2 category). Second, the speaker must believe the statement to be false. Third, in uttering the statement, the speaker must intend to deceive the listener.[51] To test this hypothesis, Coleman and Kay constructed eight stories that satisfied all, only some, or none of these conditions. They then asked sixty-seven subjects to determine whether, in the case of each story, a lie had been told. In support of their hypothesis, the results indicated that "lie" functioned as a Type 1 category: judgments of what constitutes a "lie" were not made comprehensively, but, rather, could be made by degree.

Although they were able to prove their hypothesis, Coleman and Kay were not able to explain satisfactorily what it was about the word "lie," or about the conditions surrounding its proper use, that made it function as a Type 1, rather than as a Type 2, category. Sweetser, for her part, proposed that judgments about lies were made relative to a cultural theory of information exchange.[52] The theory, as formulated by Sweetser, consists of three relatively straightforward propositions, which can be stated as rules:

1. *Try to help (do not harm).* This rule is combined with the belief that knowledge is helpful and that misinformation is harmful (which leads to the next rule).
2. *Give knowledge (do not misinform).* Sweetser argues that what counts as knowledge, in everyday terms, is what we believe. All things being equal, we assume that what we believe matches some state of affairs or possible state of affairs: we take what we believe to be true, to count as knowledge. Using this understanding of knowledge together with (1) and (2), we arrive at the final rule.
3. *Say what you believe (do not say what you do not believe).* If what we believe counts as knowledge, and if we wish to be helpful, we must give knowledge.

This chain of entailments represents a cultural theory of information exchange that allows us to reach the somewhat surprising conclusion that if a speaker (S) utters a

51. Linda Coleman and Paul Kay, "Prototype Semantics: The English Word *Lie*," *Language* 57 (1981): 26–27.

52. Eve E. Sweetser, "The Definition of *Lie*: An Examination of the Folk Models Underlying a Semantic Prototype," in *Cultural Models in Language and Thought*, ed. Dorothy Holland and Naomi Quinn (Cambridge: Cambridge University Press, 1987), 43–66. Sweetser's "models" have the same function as the theories described in chap. 3; I have simplified some aspects of her account in the interests of concision.

proposition (P), then that proposition is true. The causal chain can be represented schematically:

a. *S said P.* [as part of an information exchange; from rules (1) and (2)]
b. *S believes P.* [(a) combined with rule (3)]
c. *Therefore, P is true.* [(b) combined with the understanding that we believe what is true and do not believe what is false]

This is, without a doubt, a greatly simplified perspective. What is important is that the model of information exchange set out by the theory works quite a bit of the time, and when it does not work, the theory provides a way to interpret where things went wrong.

To illustrate, let us consider two situations. In the first, I claim that I saw a friend of mine walking down the street. On closer inspection, it turns out to be someone who just looked like my friend. According to the simplest definition, my claim is a lie—a statement that is not true. However, since my behavior as a whole fits with the cultural theory of information exchange (my intent was to provide helpful information), my claim will most likely not be judged a lie. In the second situation, I see someone walking down the street whom I mistake for my friend. Remembering that I owe my friend money and being a cheap sort, I claim that the person I see is not my friend. According to the simplest definition, this is not a lie, since my claim is true: the person walking down the street is not, in fact, my friend. Nonetheless, since I deliberately said what I thought to be misinformation, I have violated the theory of information exchange and my claim might well be judged a lie.

Of course, this straightforward theory of information exchange does not operate in isolation but is interpreted through other cultural theories to explain types of discourse that seem to contravene the theory. Jokes, tall tales, and fiction all may have limited truth-value, but their "lies" are permitted as part of a cultural theory that interprets certain types of exceptional behavior as entertainment. Social lies are based on the recognition that not all forms of knowledge are helpful: rarely when someone remarks "How are you?" do they really want to know how we are. The cultural theory of information exchange is thus filtered through a theory that says maintaining social order through polite behavior is more important than saying what you believe to be true.[53]

Applying Sweetser's example of a cultural theory of information exchange to Bakhtin's first two types of discourse is relatively straightforward. All things being equal, we assume that direct discourse—the voice of the author—involves the truth; we similarly give the benefit of the doubt to the objectified discourse of the author's characters.[54] With double-voiced discourse, however, the application of the

53. For a related view of how information exchanges contribute to and constitute social order, see Michael B. Bakan's account of how various Balinese musicians reported the results of a gamelan competition in *Music of Death and New Creation: Experiences in the World of Balinese Gamelan Beleganjur*, Chicago Studies in Ethnomusicology (Chicago: University of Chicago Press, 1999), chap. 5.

54. The genre of fiction, which is the focus of Bakhtin's analysis, requires an intriguing nesting of theories about information exchange. On the level of everyday life, fiction is generally recognized as nonfactual (except by the delusional). Within a given work of fiction, however, we usually apply the theory of information exchange outlined by Sweetser: we assume that characters say what is true in the con-

theory is a bit more complicated. With the first type of double-voiced discourse—where the discourse of the author and the other person essentially cooperate—the idea that both voices speak the truth is easily accommodated. What results is a sort of heterophony of discourse: two voices saying the same thing in different ways. With the second and third types—parody and hidden polemic—the contradictions and tensions between the two voices suggest that one of the voices must be violating the principles of the cultural theory of information exchange. The traditional resolution of the problem, implicit in Bakhtin's analysis of discourse as a whole, is that in double-voiced discourse, the voice of the author is to be privileged: it is the voice of authority.[55] This resolution requires modification of Sweetser's cultural theory of information exchange by the addition of one further proposition:

4. *The voice of authority is always helpful.*

On the one hand, it is difficult to ascribe this interpretation to Bakhtin without a crushing sense of irony, given his personal history. But on the other hand, it is certainly the case that, in his analyses of double-voiced discourse, the "first" (nonauthorial) voice is the one he most often distrusts.

Given this analysis, equating Gates's notion of Signifyin(g) with Bakhtin's double-voiced discourse brings us to two conclusions. First, the voices of black rhetoric become voices of authority: they are the helpful voices that speak the truth. Second, Signifyin(g) is by no means unique to the African American community—indeed, Dostoevsky and countless other "authors" are also Signifyin(g). But these conclusions are hardly satisfying, for they trivialize the points Gates wishes to make about African American literature and culture and miss the ways language can be used against oppressors.

The solution to these problems lies in two important adjustments to the cultural model of information exchange adumbrated here. First, a distinction must be made between "author" and "authority." For the African American community, the first voice in double-voiced discourse was almost always the voice of "authority." Since experience had shown that the voice of authority was almost always unhelpful, it was a voice to be distrusted. Rather than trusting one of the voices of a double-voiced discourse, it was important actively to distrust one of those voices; by default, the nonauthoritative voice of the author was to be regarded as the helpful one. Second, it did not particularly matter if the statements of the author were true or not. Indeed, as Gates points out, one of the common African American terms for Signifyin(g) is "lies."[56] What mattered more was whether the author, in some way, adhered to the first principle of information exchange: to be helpful. And it is clear that, to an oppressed people, speech that co-opts, mocks, or indirectly challenges the voice of authority, whether the statements made in the course of that speech are

---

text of the story. When a character says something that is not true within the context of the story (even though it may be true in the everyday world), we then have to fall back on this theory or its elaborations to make sense of the utterance.

55. The pun is not surprising, since "authority" and "author" share the same Latin root (*auctor*).

56. Gates, *The Signifying Monkey*, 56, 71.

true or not, is singularly helpful. With these adjustments, distinctions between double-voiced speech and Signifyin(g) become clear: Signifyin(g) relies on a modified theory of information exchange that assumes that the first voice in double-voiced discourse is the voice of an extracultural authority and that this authority is not to be believed. Under these circumstances, double-voiced discourse is, as Bakhtin characterized it, "an arena of battle between the two voices." However, it is an arena in which the author is at a distinct disadvantage and must rely on whatever combination of truths, half-truths, and lies is necessary to win the day.

My intent with this analysis is to show how Bakhtin's notion of double-voiced discourse functions relative to a highly simplified, yet quite serviceable, rendering of a cultural theory of information exchange and how the double-voiced discourse of Signifyin(g) relies on a modified version of this theory. In applying this perspective on discourse to music, I restrict my discussion to the level of the conceptual model for a work. Musical discourse, understood in this way, is thus concerned with the realization and projection of the conceptual model for a work. In terms of the cultural theory of information exchange (extended to music), our assumption is that the performer intends to project this model and that it is valid for the piece. My decision to impose this restriction is based on practical reasons. It will keep discussion focused on the main topics of this chapter, it will greatly simplify the analyses, and it requires no strong analogy between linguistic and musical discourse. There is, however, no reason that, mutatis mutandis, this treatment could not be extended to more finely grained levels of musical discourse.

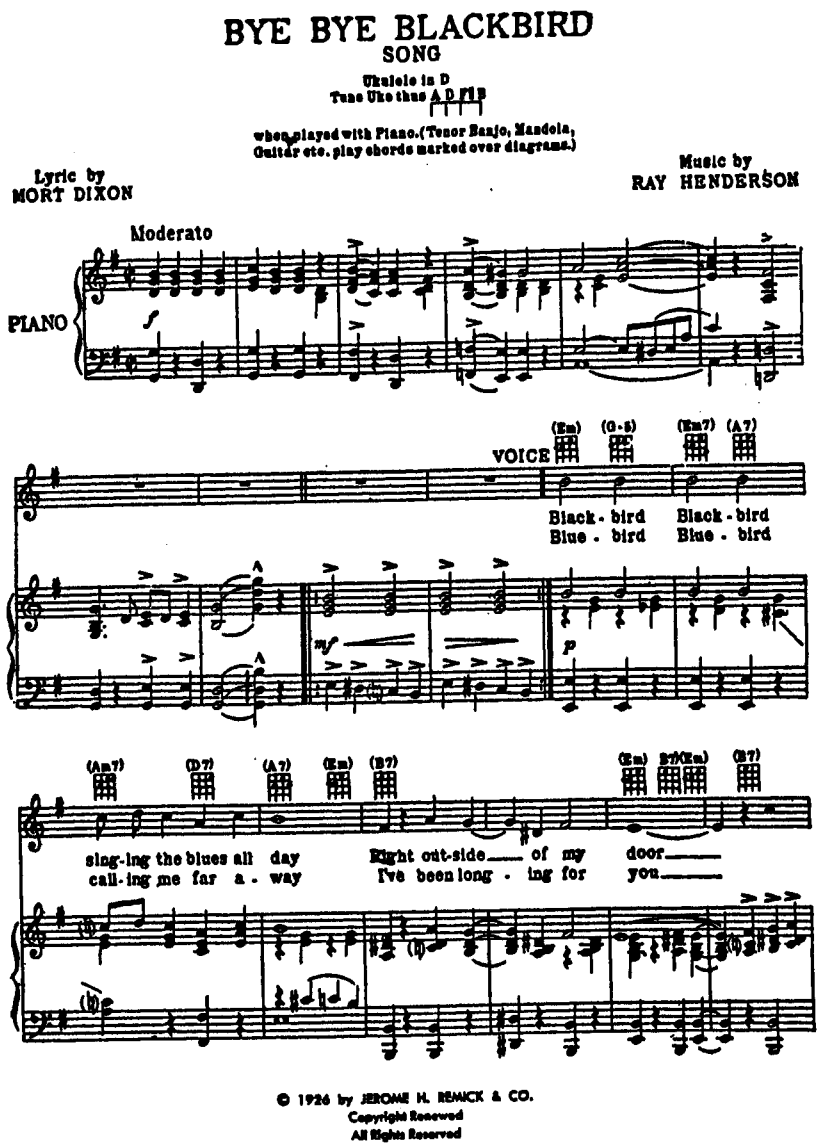
## Conceptualizing "Bye Bye Blackbird," 1926–1955

"Bye Bye Blackbird" was written in 1926 by the song-writing team of Mort Dixon and Ray Henderson. The score, given in example 5.4, reveals a song typical of the thousands written during the days of Tin Pan Alley. It consists of an eight-measure introduction (taken from the concluding phrase of the refrain); a two-measure vamp to allow the vocalist to get ready; a thirty-two-measure verse in AABA' form; and a thirty-two-measure refrain in AA'BA" form (where A' presents the material of the A section a diatonic step higher, and A" changes the original A section at midphrase so that it ends on the tonic).[57]

The general topic of the song centers around a narrator in alien or unhappy surroundings, symbolized in the first verse by the trope of winter, the blackbird's doleful pronouncements, and the use of minor mode. The narrator rejects these conditions and looks to where "there's sunshine galore," a promise held out by the refrain: a waiting loved one, a familiar bed, and a light burning in the window, all intro-

57. At least two other scores for "Bye Bye Blackbird" have been in circulation. Both of these are in F major, and both shorten the verse to a sixteen-measure form by omitting the internal AB sections. One score, with the copyright renewed and assigned to the Remick Music Corporation, has a short, four-measure introduction; the sixteen-measure verse; and the refrain. The other, copyrighted in 1948 by the Remick Music Corporation, was printed by Francis, Day and Hunter, London. It gives the refrain first, followed by the introduction and verse. I have found no recordings of performances based on either of these scores.

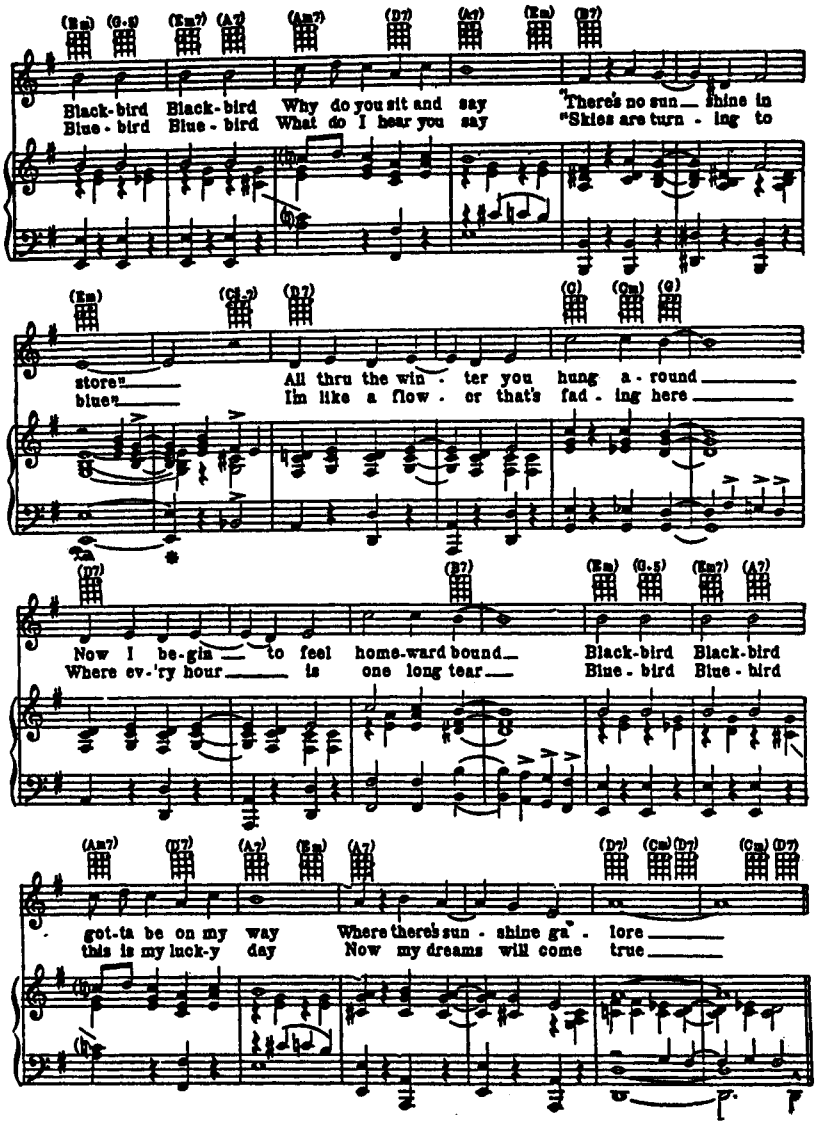
EXAMPLE 5.4 Ray Henderson and Mort Dixon's "Bye Bye Blackbird," © Fred Ahlert Music Group




duced against the backdrop of a sprightly G major. The second verse (which replaces the blackbird with a bluebird) reflects the hope offered by the refrain, conditioned nonetheless by the return to minor mode and references to longing, fading like a flower, and hours that are like "one long tear."

Recordings of the period indicate typical treatments of the song. In their version of 22 April 1926, for example, Bennie Krueger's Orchestra begins with a short introduction (which bears no resemblance to the published one) and proceeds to an

EXAMPLE 5.4 *(continued)*



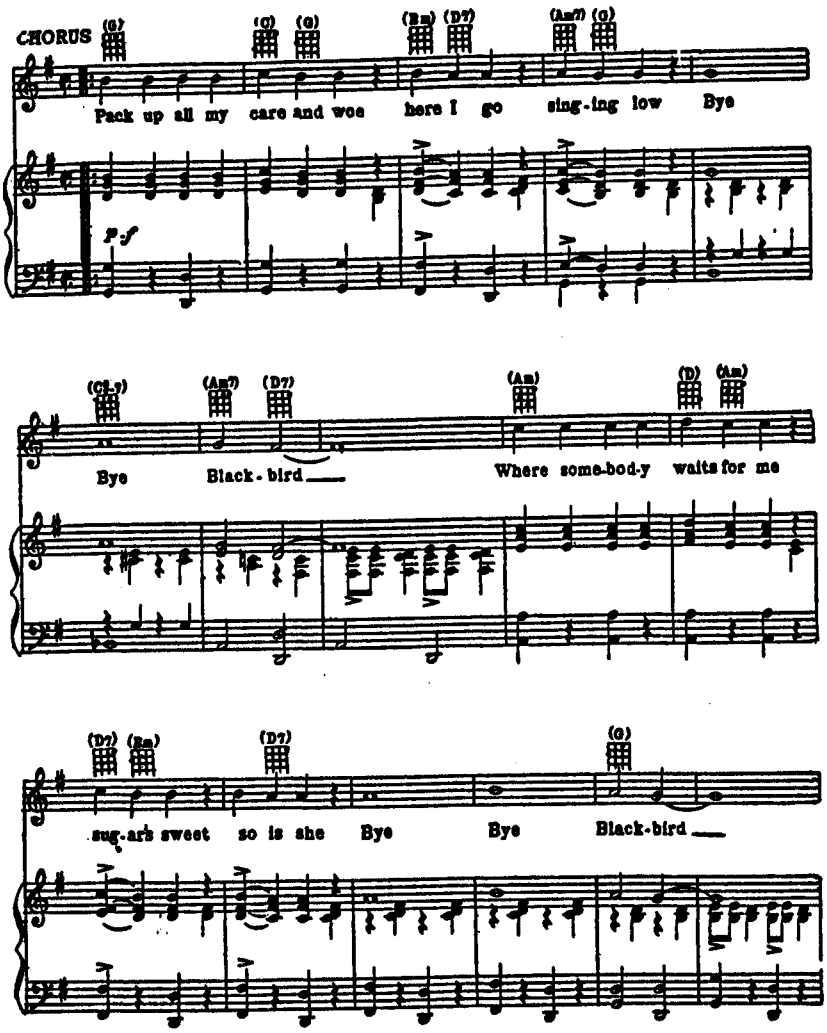

*(continued)*

instrumental version of the refrain, verse, and refrain.[58] (As with the early recordings of "I Got Rhythm," the Krueger arrangement uses the verse as an interlude between the refrains.) This is followed by a sung refrain and one final instrumental refrain. The tempo of the recording (♩ = 116–120) is a bit faster than the "Moderato" indicated by the score, but it is perfectly suitable for a foxtrot.[59] Gene Austin's

58. Bennie Krueger's Orchestra, "Bye Bye Blackbird" (Brunswick, [78] no. 3186, 1926).

59. The song is in fact identified as a foxtrot on the label of the record, which also includes the rubric "for dancing."

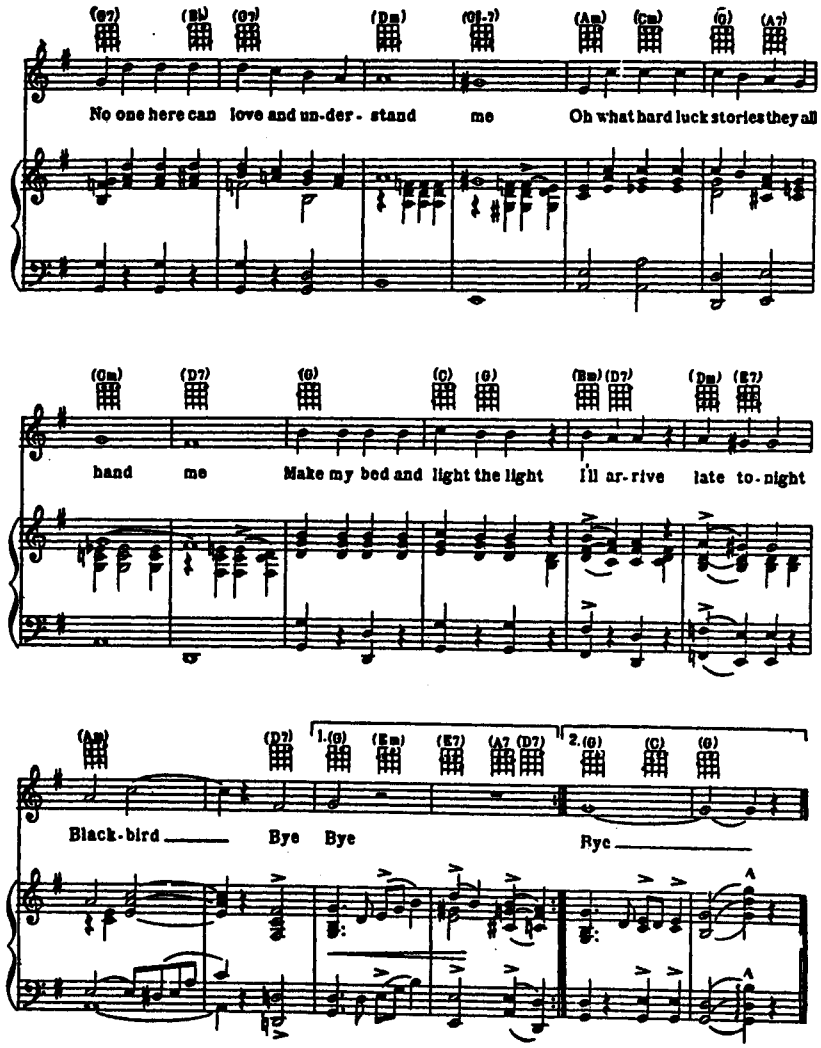
EXAMPLE 5.4 *(continued)*




recording, made a week later, follows the score rather more closely.[60] Although he adds his own introduction and piano accompaniment and varies the melody in the third phrase of the verse, his performance otherwise conforms to the score given in example 5.4, even adhering to a somewhat slower tempo than did Krueger's group (♩ = 96–100). By most standards, "Bye Bye Blackbird" was a hit: it was recorded at least fourteen times between March and June of 1926, and Austin's recording was no. 1 on the *Billboard* charts of 4 September of that year.

60. Gene Austin, "Bye Bye Blackbird" (Victor, [78] no. 20044-B, 1926), originally recorded 29 April.

EXAMPLE 5.4 *(continued)*




The various performances of "Bye Bye Blackbird" recorded in 1926 could be collected into a category organized around a conceptual model quite similar to that used for early versions of "I Got Rhythm." The model, which is diagrammed in figure 5.5, consists of five correlated conceptual elements, characterized as the things necessary for an adequate performance of the song. However, where the pop-music model for "I Got Rhythm" continued to have some currency in the years just after the first appearance of the song (in part because of its association with the musical), the model for "Bye Bye Blackbird" appears to have gone dormant for a

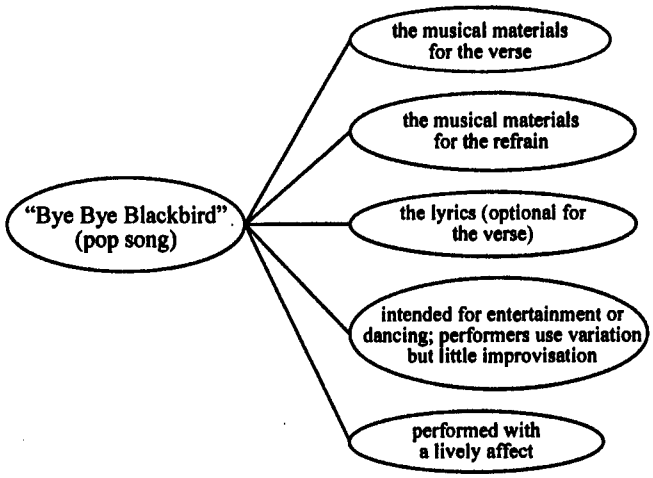



FIGURE 5.5 Model for the category *performances of "Bye Bye Blackbird"*

time: there are few signs of the song having been recorded after 1926 and throughout the 1930s.[61]

New recordings of "Bye Bye Blackbird" did begin to appear in the 1940s. There are a number of indications, however, that by then the conceptual model for the song had changed. The "novelty" recording by Freddie "Schnickelfritz" Fisher and His Orchestra (made on 22 February 1941), with its rapid succession of unlikely instruments (including solo whistling and baritone saxophone) and its "shave-and-a-haircut" ending, is just short of an all-out joke.[62] In a 1941 German propaganda recording by Charlie and His Orchestra, the words of the song are altered to celebrate early Axis victories in the Second World War and to taunt the British ("Bye Bye Blackbird" becomes "Bye Bye Empire").[63] And in a tribute to the songs of 1926 recorded on 27 May 1942 by the piano duo of Marlene Fingerle and Arthur Schutt, "Bye Bye Blackbird" is the final song in a medley that includes "I Know That You Know" and "Baby Face."[64] In each of these recordings, only the refrain of the song is used, and the tempo varies from the moderate swing of Charlie and His Orchestra (♩ = 84) to the Fisher Orchestra's foxtrot (♩ = 120–126). While there are some signs the song has entered the repertoire of "standards"—its verse has

61. Recording, of course, is not the only measure of popular interest in a song: sales of records, piano rolls, and sheet music; programming of the song in live performances; and the playing of recordings on radio broadcasts and in juke boxes would also provide an indication of continued interest in the song. Unfortunately, no information on these indicators is presently available for "Bye Bye Blackbird."

62. Freddie "Schnickelfritz" Fisher and His Orchestra, "Bye Bye Blackbird" (Decca, [78] no. 25357/3788, 1941).

63. Charlie and His Orchestra, "Bye Bye Blackbird," in *Charlie & His Orchestra: German Propaganda Swing 1941–1942* (Harlequin, [CD] HQ CD 03, 1990).

64. Marlene Fingerle and Arthur Schutt, "Medley: 'I Know That You Know,' 'Baby Face,' 'Bye Bye Blackbird,'" in *Song Hits of 1926*, Songs of Our Times (Decca, [LP] DL 5170, 1950).

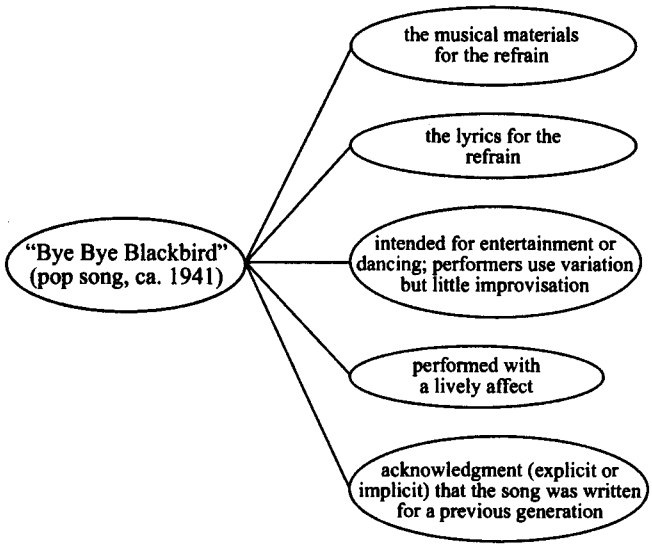



FIGURE 5.6 Model for the category *performances of "Bye Bye Blackbird," ca. 1941*

fallen away, and it is readily available for parody—it is also clearly marked as an "old" song. The conceptual model for the category *performances of "Bye Bye Blackbird," ca. 1941*, then, would consist of five correlated conceptual elements (which are diagrammed in fig. 5.6).[65]

As a consequence of a growing interest in the 1920s as an era, this model was, if anything, reinforced during the post–World War II period. Following the phenomenal success of *The Jolson Story* in the fall of 1946, Warner Brothers films contracted in 1947 for the filming of a similar story with (and about) Eddie Cantor. Although Warner Brothers ultimately dropped the project, Cantor and Sidney Skolsky (producer of *The Jolson Story*) eventually filmed *The Eddie Cantor Story* independently, and it was released in 1953.[66] As part of the soundtrack for the movie, Cantor recorded "Bye Bye Blackbird" in something approximating the style of the 1920s, although using only the refrain of the song.[67] A similar approach can be heard in

65. I do not mean to suggest, with the idea that "Bye Bye Blackbird" was "written for a previous generation," that the fifteen years between 1926 and 1941 actually fill the span of a human generation. However, I do think this period is long enough to mark a significant change in popular tastes, which is what this portion of the conceptual model aims to capture.

66. Herbert G. Goldman, *Banjo Eyes: Eddie Cantor and the Birth of Modern Stardom* (New York: Oxford University Press, 1997), 283–84.

67. Eddie Cantor, "Bye Bye Blackbird," in *Songs by Eddie Cantor from* The Eddie Cantor Story (Capitol, [LP] L-46, 1954). There is no evidence that Cantor recorded "Bye Bye Blackbird" during the 1920s or 1930s, although he might well have performed it in live shows; in fact, Goldman characterizes it as one of the standards of Cantor's repertory from the 1920s. By most accounts, *The Eddie Cantor Story* was a flop; its failure is generally attributed to the inadequacies of Keefe Brasselle (who played the role of Cantor) and to a script that was little more than a remake of *The Jolson Story*. Of the soundtrack, however, Goldman writes: "The musical arrangements were first rate, and the vocals better than might be

Johnny Maddox's recording of the song on piano (accompanied by a rhythm section), released in 1955, which interprets it in a quasi-ragtime style as part of a collection of 1920s songs intended to accompany tap dancing.[68]

During this same period the general fascination with the character and music of the 1920s was given a specific focus when there was a resurgence of interest in early jazz, variously called "New Orleans" or "Dixieland" jazz.[69] Partly, this came about because of a strike by musicians against the recording industry from 1942 to 1944, which led record companies to reissue older recordings. Around the same time, there emerged an audience who felt the swing era had passed but who were not attracted to be-bop, the new jazz style that emerged during the Second World War. In 1950, the actor Jack Webb (now chiefly remembered for his role as Sergeant Friday on the *Dragnet* radio and television series) developed a radio series around a jazz musician in 1920s Kansas City, reflecting contemporary interest in the "roaring 20s" and Webb's own enthusiasm for early jazz.[70] The radio series was eventually developed into the movie *Pete Kelly's Blues*, which Webb directed and in which he also starred. In the movie, Pete Kelly is a bandleader threatened by a gangster who is attempting to extract protection money from musicians. Kelly initially tries to resist but eventually capitulates. As part of his agreement with the gangster, Kelly's band has to take on the gangster's girlfriend, Rose Hopkins (played by Peggy Lee), as a singer. Rose is an alcoholic but can still render a moving ballad. A tragic character, Rose is eventually reduced to the mentality of a five-year-old by alcoholism and by a beating from the gangster.

As part of the promotion of *Pete Kelly's Blues* as a "jazz" movie, Lee, together with Ella Fitzgerald (who played a blues singer working in a roadhouse), recorded an album of songs associated with the movie on 10 May 1955. Although "Bye Bye Blackbird" appears in the movie, neither Lee nor Fitzgerald sings it. Instead, it is rendered by a drunken chorus of band members returning from a late-night party. On the album, Lee retains Rose's persona from the movie and sings the song (consisting only of the refrain) as a slow ballad (♩ = 72–76), with slight allusions to Billie Holiday's distinctive style of inflection.[71] The arrangement makes use of complicated harmonies typical of the treatment of ballads in 1940s and 1950s jazz and

---

expected, considering that Cantor had suffered a major heart attack less than three months earlier and that his voice had lost its former brilliance after World War II." Goldman, *Banjo Eyes*, 284.

68. Johnny Maddox, "Bye Bye Blackbird," in *Tap Dance Rhythms* (Dot, [LP] DLP 3008, 1955).

69. Chip Deffaa, *Traditionalists and Revivalists in Jazz*, Studies in Jazz, 16 (Metuchen: Scarecrow Press; Institute of Jazz Studies, Rutgers, The State University of New Jersey, 1993), 132. In the following I refer to this style of jazz as "traditional" jazz, retaining the quotation marks to highlight the function of the term as a rubric rather than a judgment about the place of this music within any historical account of jazz styles. See also Bernard Gendron, "'Moldy Figs' and Modernists: Jazz at War (1942–1946)," in *Jazz among the Discourses*, ed. Krin Gabbard (Durham, N.C.: Duke University Press, 1995), 31–56.

70. Charles Emge, "Meet 'Jazzman' Jack Webb: Swing Era Helped Make Him Music-Conscious," *Downbeat* 22 (10 August 1955): 8.

71. The character of Rose seems to have had a profound personal effect on Lee. In her autobiography, she speaks of Rose in the third person, speculates on whether a visit from James Dean was intended for her or for Rose, and writes of "moving out of Rose's trailer [on the set]" once the film was completed. See Peggy Lee, *Miss Peggy Lee: An Autobiography* (New York: Donald I. Fine, 1989), 165–67.

adds words for the second pass through the refrain that signal a new perspective on the song:

A while ago when I was young, I heard a song and heard it sung
[male chorus sings as distant echo] Bye bye blackbird
I don't know why it makes me sad, a happy song should make me glad
[male chorus] Bye bye blackbird

Singing in her persona as a dissolute, tragic chanteuse, Lee evokes a melancholy that, up to this point, had not been associated with "Bye Bye Blackbird."

*Pete Kelly's Blues* was, at best, a mixed success critically, and Lee's treatment of "Bye Bye Blackbird" might well have been forgotten (especially since she did not sing the song in the movie). Fate, however, intervened: the singer and pianist Bobby Short heard and liked Lee's recording and decided to approach "Bye Bye Blackbird" in a similar fashion, finding reflective sadness in the refrain rather than joyful anticipation.[72] In his recording of 15 September 1955, Short also adopts a slow ballad tempo (♩ = 84) and fashions his accompaniment out of rich, progressive harmonies. To provide an introduction, he starts with the B section of the refrain, the one section that retains the sense of alienation generated by the original verse: "No one here can love and understand me . . ." The effect is to cast a shadow over the remaining lyrics of the refrain, emphasizing what is negative in them—the narrator's care and woe, distant loved one, and need to travel late into the night—rather than the pleasant things that wait at home. Short also adds his own lyrics for the second pass through the refrain (with words adapted from the song "I'm a Little Blackbird Looking for a Bluebird"):[73]

I've been all over the east and the west, in search of clover to feather my nest,
Bye bye blackbird
I have a daydream the same as you do, why would a daydream have to be hoo-doo?
Bye bye blackbird
No one here can love and understand me,
Oh, what hard luck stories they all hand me. (etc.)[74]

With these two recordings by Lee and Short, "Bye Bye Blackbird" was transformed from an upbeat popular song in a fast tempo, designed for dancing—a "hop" tune, as Short has called it[75]—into a melancholy jazz ballad in a slow tempo, designed for listening. The motivation for this change was not, strictly speaking, musical, for most of the music in *Pete Kelly's Blues* kept to a "traditional" or Dixieland style.[76] Instead, the change was initially motivated (that is, in Peggy Lee's

72. Personal communication from Bobby Short, 8 March 2000.

73. "I'm a Little Blackbird Looking for a Bluebird" was written by Clarke Grant, Arthur Johnston, George W. Meyer, and Roy Turk. It was recorded by both Louis Armstrong and Sydney Bechet in the mid-1920s.

74. Bobby Short, "Bye Bye Blackbird," in *Bobby Short* (Atlantic, [LP] 1230, 1956), recorded 15 September 1955. As will be noted, Short's revised lyrics return to Dixon's original refrain at midpoint (with, "No one here can love and understand me").

75. Personal communication from Bobby Short, 8 March 2000.

76. The soundtrack for the movie (with priceless narrations by Jack Webb) consisted exclusively of

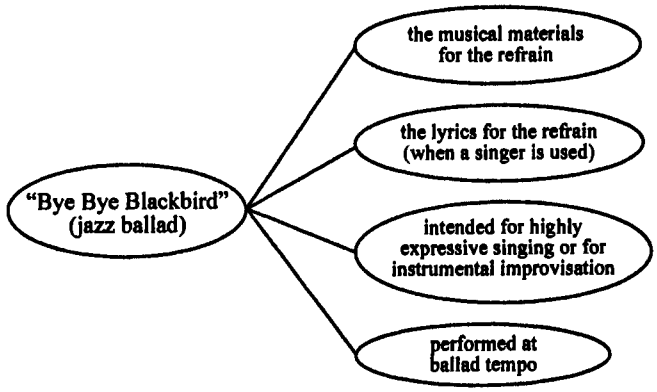



FIGURE 5.7 Jazz ballad model for the category *performances of "Bye Bye Blackbird"*

recording) by a romanticized notion of the female jazz singer, itself derived from any number of nineteenth-century fantasies about performers.[77]

The new conceptual model for "Bye Bye Blackbird" suggested by Lee's and Short's recordings, and diagrammed in figure 5.7, consists of four elements necessary for an adequate performance. Gone is the sense of history that had freighted approaches to the song since the early 1940s; it is replaced by a sense of immediacy generated by the necessarily reflective nature of the performances.

It may be useful at this point to consider differences in the way conceptions of "I Got Rhythm" and "Bye Bye Blackbird" changed over roughly the same period in time. Because "I Got Rhythm" was an enduring part of the popular music and jazz repertoire, its conceptual model changed only gradually, through a process of slow but relatively constant negotiation. In contrast, no such process is in evidence for "Bye Bye Blackbird." After its first blossoming, the tune, and its conceptual model, faded from popular consciousness. When it was recalled during the early 1940s, it was as part of a larger process of revisiting, through music, an earlier era: "Bye Bye Blackbird" was simply emblematic of "a popular tune from the 1920s." When Peggy Lee introduced her version, the change to the conceptual model was sudden and dramatic, by fiat of her character in the movie.

---

"traditional" arrangements. See "Pete Kelly and His Big Seven" (Matty Matlock and others), *Pete Kelly's Blues* (BMG Special Products, [CD] DRC12081, 1998), originally recorded 7, 9, 11 June 1955.

77. For a recent, and stimulating, discussion of some of these fantasies, see Susan Bernstein, *Virtuosity of the Nineteenth Century: Performing Music and Language in Heine, Liszt, and Baudelaire* (Stanford, Calif.: Stanford University Press, 1998), chap. 3.

My characterization of the persona developed by Lee in connection with *Pete Kelly's Blues* should not obscure the powerful presence of African American women who sang the blues and jazz, or the very real trials to which these women were subjected. (For a discussion of the role of these women in shaping the blues in particular, see Susan McClary, *Conventional Wisdom: The Content of Musical Form* [Berkeley: University of California Press, 2000], 42–49.) I do believe, however, that Lee, as a white entertainer in the United States of the 1950s, was able to choose this persona (much as she chose to emulate Billie Holiday's voice in the course of her performance of "Bye Bye Blackbird") and that part of the framework for this choice was a romanticized notion of the suffering performer.

Thus by the end of 1955 there were two conceptual models for "Bye Bye Blackbird" in circulation. The first conceptualized the tune as a somewhat dated but nonetheless upbeat pop song; the second conceptualized the tune as a slow, melancholy ballad. There may have also been still circulating a third, and more sinister model, derived from the first, that interpreted the song in racial terms—a signal for African Americans to be on guard, that they were not wanted here. It was into the context provided by these models that Miles Davis introduced his version of "Bye Bye Blackbird."

## Davis's "Bye Bye Blackbird," Double-Voiced Discourse in Music, and Signifyin(g)

The recording that Davis made with Garland, Chambers, Jones, and Coltrane was strictly instrumental, at a tempo slower than the "hop" version of the tune (♩ = 120), but notably faster than the tempos adopted by Lee and Short. After a brief introduction by Garland, Chambers, and Jones that establishes a loose, almost loping, groove, Davis enters playing the refrain. From his first notes, Davis announces the song as a jazz ballad, using a languid approach to the melody and muted trumpet to create an understated, slightly distant reflection on the song. After the head, Davis takes two choruses, followed by two choruses from Coltrane that occasionally hint at a double-time feel. Garland takes a chorus and a half, and, at the bridge of Garland's second chorus, Davis reenters with the melody and guides the performance to a conclusion.

Davis's principal contributions are to infuse "Bye Bye Blackbird" with an irresistible groove and to establish it as a venue for instrumental commentary and expansion. Although these contributions are by no means insignificant—Davis's recording virtually established "Bye Bye Blackbird" as a jazz standard, something it had never been before—one can also hear the influence of Lee's and Short's recordings in his version. There is, obviously, the slower tempo and hint of melancholy adopted from the jazz-ballad model (although the groove Davis sets up tends to move the tune back toward the pop-music model). There is also, however, the same sense of immediacy that marked Lee's and Short's performances: "Bye Bye Blackbird" is not, in Davis's hands, about nostalgia.[78] Davis has modified the conceptual model for the tune used by Lee and Short, but he has also retained enough of it that we can still recognize its main elements.

The immediate relationship between the jazz-ballad model for "Bye Bye Blackbird" and the model Davis set out in his recording suggests that double-voiced discourse can be accomplished in music when one conceptual model for a musical work is understood to make reference to another conceptual model for the same

78. The element of nostalgia has continued to cling to "Bye Bye Blackbird" as a popular tune (as distinct from a jazz standard) and has been exploited in a number of films. In Jonathan Demme's *Melvin and Howard* (1980), "Bye Bye Blackbird" is the one song Howard Hughes is able to recall when Melvin Dummar presses him—on pain of being thrown out of Dummar's truck—to sing something, and it becomes a deeply felt recollection of Hughes's childhood. In Nora Ephron's *Sleepless in Seattle* (1993), "Bye Bye Blackbird" is cast as the song the boy's deceased mother used to sing to him when he couldn't sleep.

work. If we know Peggy Lee's recording of "Bye Bye Blackbird," hear Miles Davis's, and understand the latter as making reference to the former, musical double-discourse takes place.[79] Musical double-discourse thus requires, minimally, a listener who knows at least two conceptual models for a given work. Connections between the models may be drawn out by the performer, but they may be more oblique and constructed mostly by the listener. Each of the three types of double-voiced discourse described by Bakhtin can be created: an inflection of the first conceptual model by the second (which is what Davis does); a parody of the first conceptual model by the second (which is what Monson describes Coltrane as doing); or tacit reference to the first model by the second (which is one way to conceive how derived tunes operate).

But what of Signifyin(g)? Signifyin(g) would require that the first conceptual model be recognized as the "authoritative" one and that its authority is not to be believed: that is, it is not a valid model for the work. Validity is then granted by default to the conceptual model used by the performer. A consequence is that the validity of this model depends on the rejection of the model associated with authority; it is not granted independently. Under these circumstances, a case can be made that, in his recording of "My Favorite Things," Coltrane is Signifyin(g) on Rodgers and Hammerstein's song: the validity of the Broadway-show-tune model for this song has been rejected, and it is Coltrane's model of the tune that is accorded validity. It is much less clear, however, whether Davis is Signifyin(g) on "Bye Bye Blackbird." It is difficult to ascribe authority of a sort that one should be skeptical of to either Lee's or Short's recordings, and it is unclear whether Davis was aware of, or cared about, the pop-music model for the tune (with or without its overlay of racism). This does not mean that a listener might not hear the recording by Davis as Signifyin(g) on "Bye Bye Blackbird," only that the relationship of Davis's rendition of the song to the concept of Signifyin(g) is much more equivocal.

Under what circumstances might we hear Davis as Signifyin(g) on "Bye Bye Blackbird"? First, Davis could, in the course of a performance (recorded or otherwise), make an obvious reference to one of the earlier models for the tune, perhaps by playing a fragment of the verse or by evoking another nostalgic tune from the 1920s. The other model for "Bye Bye Blackbird" would thus be brought to our attention (even if only as something anomalous in Davis's performance) to be then regarded as a voice of authority under challenge. Second, we could bring to Davis's performance a knowledge or awareness of one of the earlier models, hear that he reverses certain of the key elements of this model, and interpret this as subverting authority. Third, Davis could, through either word or deed, situate himself as in dialogue with what he regarded as an authoritative version of the tune.

With this analysis of Davis's recording of "Bye Bye Blackbird" I do not mean to suggest that Signifyin(g) is a rare occurrence in music, but only that it involves a complex matrix of cultural knowledge and that the listener must understand the process of Signifyin(g) as taking place. When this matrix of knowledge is not in place, what may be left to hear is only double-voiced discourse, or unmediated

79. Although my focus here is on recordings, the same situation would obtain were our experience with live, as opposed to recorded, performances.

authorial discourse. (Indeed, all music could be heard as involved in double-voiced discourse of a quite general kind, in that the conceptual model for each work is in dialogue with the conceptual models for other works.) Double-voiced discourse, as well as Signifyin(g), becomes interesting when tensions between the models become obvious—that is, when the tensions become as important as the models themselves.

### Signifyin(g) and Cultural Knowledge

Given the cognitive complexity of the process of Signifyin(g), it is well worth asking why the notion has become a trope among those who work on jazz and popular music. Three reasons come to mind. First, Signifyin(g) offers a clear alternative to the Saussurean model of meaning construction: in Signifyin(g), meaning is something that is not fixed but is created by the listener.[80] Second, Signifyin(g) provides a way of interpreting the cultural practice of jazz and popular music in terms developed to account for the unique profile of one of its most important influences, African American culture. Third, applying Signifyin(g) to music allows music critics to avail themselves of the power and prestige of literary criticism.

Of these three, it is perhaps only the second that has much relevance to music, as opposed to the practice of writing about music. Musical Signifyin(g) is a complex process, distinct from linguistic Signifyin(g) and thoroughly implicated in the dense matrix of cultural knowledge. To explain how it works—when it works—requires not only a theory of Signifyin(g) but also a way to account for musical ontology, an explanation of the framework for discourse provided by cultural theories of information exchange, and a theory of double-voiced discourse. With these in place we can start to see—and hear—how Signifyin(g) shapes musical discourse.

## CULTURAL KNOWLEDGE AND MUSICAL ONTOLOGY

Clifford Geertz was skeptical about using descriptions of individual cognitive processes to build an account of culture, and rightly so. Cognitive structure is incredibly complex, and the relationship between individual knowledge and cultural knowledge is not a simple one. Nonetheless, one can discern patterns of consistency in the features selected for defining and evaluating cultural products, patterns that are indicative of cultural knowledge. The approach to cultural knowledge I have developed in this chapter is meant to capture these consistencies and to facilitate a discussion of how cultural knowledge is organized, sustained, and changed. It is not meant as a literal rendering of what someone has in mind in evaluating particular cultural products. The approach is thus an interpretation of culture, but of a

80. This is essentially the approach taken by Robert Walser in "'Out of Notes': Signification, Interpretation, and the Problem of Miles Davis," in *Jazz among the Discourses*, ed. Krin Gabbard (Durham, N.C.: Duke University Press, 1995), 165–88. While my approach to Signifyin(g) is broadly consonant with this treatment of meaning construction, my argument is made from the perspective of research in cognitive science rather than structuralism.

somewhat different sort than those called for by Geertz, for it assumes that culture is what we know, rather than the specific things we do, make, or say.

As members of a culture, one of the things we know is how to categorize sequences of sound events as "music" (or whatever other term is operational in our culture). Knowing that a Beethoven string quartet is music (to return to Geertz's example) means that it conforms to our conceptual model for music, no more and no less. Similarly, John Cage's *4'33"* is music if it conforms to our conceptual model for music; if it does not, we may simply regard it as a rather strange ritual in which a group of people sit in silence for a specified period of time.[81] The conceptual models around which we organize our categories of music are also the means by which we negotiate or resist cultural change, guide performance, and engage in complex patterns of musical discourse. They are, quite simply, the stuff of musical ontology.

Quite obviously, what I offer here is only a beginning. In particular, the account of the conceptual models appropriate to cultural knowledge needs to be refined through further study. Coleman and Kay's study of the meaning of "lie," together with Sweetser's interpretation of that study, offers an interesting example of how relevant empirical research might be conducted. If conceptual models truly have value as a tool for describing cultural knowledge, then it should be possible to develop experimental protocols that test whether and how members of a culture use such models. What will result is not a simple picture of what someone has "in mind" when reasoning about a given situation (which is at best a dubious goal) but a better understanding of the elements of knowledge that make it possible to function as a member of a culture.

Another obvious extension of the work presented in this chapter is to other repertoires, both contemporary and historical. In some cases (for instance, when dealing with a remote historical period), the application may be so speculative as to be useful chiefly in providing a fresh perspective, but in others it will yield surprising results. There are processes every bit as complex and challenging as Signifyin(g) that operate in other repertories—for instance, the intricate patterns of influence and innovation that abound in the music of Gustav Mahler and Arnold Schoenberg—patterns that await only our interest and efforts to discover them.

Perhaps the most intriguing consequence of the argument put forth in this chapter is that, if musical works are cognitive categories, and if music is a cognitive category, then music theory is about the study of categories or—more typically—the conceptual models around which musical categories are organized. Understood this way, music theory immediately escapes the gravitational force field of "the text" that has at times kept the literary criticism of the previous generation earth-bound (even while remaining relentlessly recondite). Thus liberated, music theory's own traditions of high abstraction and unabashed pragmatism may yet find a place in contemporary intellectual discourse.

81. Although I have not tried (or wanted) to develop an overarching conceptual model for music here, it is worth noting in this connection that Cage's four-and-a-half minutes of silence has a score and has been presented in concert numerous times. If these things are important elements of someone's conceptual model of music, they would most likely regard Cage's piece as music.

CHAPTER SIX

# WORDS, MUSIC, AND SONG: THE NINETEENTH-CENTURY LIED

On 15 December 1822, the poet Wilhelm Müller noted the publication of a group of songs that took his poems as their texts. In a letter to the composer of the songs, Bernhard Klein, Müller observed, "indeed, my songs lead . . . only *half* a life, a paper-life, black upon white . . . until music breathes the breath of life into them, or at least, when it slumbers within, calls it out and wakens it."[1]

That music could breathe life into poetry is a sentiment also expressed in a eulogy for a composer Müller chose not to address in the foregoing letter, but whose settings have kept Müller's poems from disappearing into the obscure recesses of nineteenth-century German poetry. Speaking of Franz Schubert's Lieder, Josef von Spaun wrote:

> That which moved the poet's breast Schubert rendered true and transfigured in tone in each of his songs as none had done before him. Every one of his songs is actually a poem on the poem that he set to music—and not only for sentiment, which is doubtless at home in songs, but also for all the magic of fantasy, for its supernatural charm and its nocturnal terrors, could he find the perfect moment—yes, even for the lofty seriousness of reflection he had a language.[2]

These two remarkably similar views of the relationship between poetry and music are, without a doubt, as much a testament to the intellectual and artistic life

1. Quoted in Carl Koch, *Bernhard Klein (1783–1832): Sein Leben und seine Werke* (Leipzig: Oscar Brandstetter, 1902), 34, n. 8. The complete text runs as follows: "In der That führen meine Lieder, die zu einem declamatorischen Vortrage, wenige ausgenommen, durchaus nicht geeignet sind, nur ein *halbes* Leben, ein Papierleben, schwarz und weiß—(Ach, wie traurig sieht in Lettern Schwarz auf Weiß das Lied mich an!) bis die Musik ihnen den Lebensodem einhaucht, oder ihn doch, wenn er darin schlummert, herausruft und weckt."

2. "Was des Dichters Brust bewegte, hat Schubert in jedem seiner Lieder treu und verklärt in Tönen wiedergegeben, wie keiner vor ihm. Jede seiner Lieder-Compositionen ist eigentlich ein Gedicht über das Gedicht, das er in Musik setzte; und nicht für Empfindung nur, die im Liede wohl daheim ist, sondern auch für jeden Zauber der Fantasie, ihren überirdischen Reitz und ihre nächtlichsten Schrecken, wußte er den eigensten Augenblick zu finden, ja selbst für den hohen Ernst des Gedankens hatte er eine Sprache." Till Gerrit Waidelich, ed., *Franz Schubert: Dokumente, 1817–1830*, Veröffentlichungen des Internationalen Franz Schubert Instituts, 10 (Tutzing, Germany: Hans Schneider, 1993), 516–17.

and ideals of early nineteenth-century Germany and Austria as they are clear-eyed proposals for a regard of song.[3] And yet they capture something important about the possibilities that arise when words and music come together: under certain circumstances the result is far more than simply the sum of its parts.

The topic of song has fascinated writers on music at least for the last two centuries, and a study of the study of song could be its own book. In a 1992 essay, Kofi Agawu offered a survey of some of this work (as well as his own contribution to it) and identified four basic models for the analysis of song, of which three were fairly well entrenched in the literature on song.[4] The first model derives from the work of Suzanne Langer and operates on the premise that when words and music come together in song, the music wholly absorbs the words. As Langer put it, "Song is not a compromise between poetry and music, though the text taken by itself be a great poem; song is music."[5] The second model proposes that there is an irreducible relationship between words and music; as Lawrence Kramer has put it, "A poem is never really assimilated into a composition; it is *incorporated*, and it retains its own life, its own body, within the body of music."[6] The third model interprets song as a compound structure in which the words carry the primary semantic content and the music colors or enhances what the words signify. Agawu observed that this model is limited because it provides no independent account of the meaningfulness of music.

In reviewing these three models, Agawu was rather too polite, resisting the temptation to note that, within the tradition of song analysis, there have been relatively few attempts to come to terms with how words and music combine to create the phenomena called "song." There are thorough and painstaking analyses of song texts, with cursory mention of the music that sets the text; there are carefully crafted analyses of musical structure, supported at turns by isolated instances of how the words buttress or anticipate this structure; but there are few comprehensive analyses that try to capture how the meaning constructed by the words and the meaning constructed by the music come together to create the meaning associated with a particular song.[7] As an improvement, Agawu proposed a fourth model in which song is construed as a confluence of three independent but overlapping systems: *music* and *words* and *song*. However, he almost immediately withheld full endorsement of this model, noting that it provided no account of a concrete iden-

3. Both must also be situated within the more personal agendas of their authors. Müller's letter is part of an attempt to get the composer to set more of his poems; Spaun's eulogy is part of an attempt to preserve what he viewed as the true legacy of his departed friend.

4. Victor Kofi Agawu, "Theory and Practice in the Analysis of the Nineteenth-Century *Lied*," *Music Analysis* 11 (1992): 3–36.

5. Susanne K. Langer, *Feeling and Form: A Theory of Art Developed from Philosophy in a New Key* (New York: Charles Scribner's Sons, 1953), 152.

6. Lawrence Kramer, *Music and Poetry: The Nineteenth Century and After* (Berkeley: University of California Press, 1984), 127.

7. One analysis that develops a comprehensive account of a song is David Lewin's "*Auf dem Flusse*: Image and Background in a Schubert Song," in *Schubert: Critical and Analytical Studies*, ed. Walter Frisch (Lincoln: University of Nebraska Press, 1986), 126–52. However, Lewin's main interest in this essay is in developing an account of certain remarkable features of this particular song; his study, thus, did not set out to delineate a broad methodological approach to the analysis of song.

tity for song but, instead, allowed song an identity only by default. This led him to muse, "Perhaps, then, what the model points to is song as process, not product. What is interesting, in other words, is not what song *is*, but what it *becomes* in its perpetual striving for a concrete mode of existence."[8]

Nicholas Cook was able to steer a course between the Scylla of emphasizing the text at the expense of the music and the Charybdis of emphasizing the music at the expense of the text, as well as to give an account of a concrete identity for song, by approaching song as an instance of multimedia: words were one medium, music another.[9] However, because Cook defined multimedia as the *perceived* interaction of different media, he tended to regard songs as successful to the extent that the syntax of their words and music remained separate—indeed, in a state of contest. The approach is effective for many songs, and the breadth offered by integrating song analysis into the analysis of multimedia as a whole makes for a very comprehensive perspective. Nonetheless, what is missing is a sense of how, despite their differences, words and music work *together* in song—how music can breathe the breath of life into poetry or create a poem on a poem.

In this chapter, I explore further the nature of song and also the process of conceptual blending. As noted in chapter 2, conceptual blending is a pervasive and often transparent process in which elements from two correlated mental spaces combine in a third.[10] And, as suggested by the examples used in chapter 2, the mental spaces basic to conceptual blends can be set up by music as well as by language. In what follows, I expand on this idea, working on the assumption that music can set up complete discourse structures, which can be correlated with the discourse structure of a text to give rise to a unique domain for the imagination. What I am concerned with, then, are not isolated or very general correlations between music and text but a conceptual blend produced by an entire song.

In the first section that follows, I analyze two settings of Wilhelm Müller's poem "Trockne Blumen" (Withered Flowers), taken from the cycle of poems he called *Die schöne Müllerin* (The Beautiful Miller Maid). The first is by the Berlin composer Bernhard Klein, whom Müller addressed in the letter quoted above; the second is by Franz Schubert. These analyses will allow me to show how different musical environments interact with the same text to produce two quite different conceptual blends, and thus two different songs. In the second section, I turn to Robert Schumann's "Im Rhein" (In the Rhein), from his song cycle *Dichterliebe* (Poet's Love). In "Im Rhein," the music makes possible a conceptual blend whose emergent structure includes a narrative considerably more extensive than that of the text alone. In the third section, I consider two settings of a poem by Joseph von Eichendorff, "In der Fremde" (In Foreign Lands), the first by Schumann, the second by Johannes

8. Agawu, "Theory and Practice," 7–8.

9. Nicholas Cook, *Analysing Musical Multimedia* (Oxford: Clarendon Press, 1998); see esp. pp. 98–106 and Cook's analysis of Robert Schumann's "Wenn ich in deine Augen seh'" from *Dichterliebe* on pp. 136–43.

10. A focus on the possibilities for meaning construction that arise in this blended space is one of the fundamental differences between the idea of conceptual blending and both Cook's conception of multimedia and Bakhtin's notion of double-voiced discourse.

Brahms. Again, two different songs result, but, in this case, the reasons lie with the generic spaces that structure each conceptual blend.

## "TROCKNE BLUMEN"

### Wilhelm Müller's "Trockne Blumen"

Müller's *Die schöne Müllerin* had its genesis in the diversions of a group of young artists and intellectuals who gathered at the house of the privy councilor Friedrich August von Stägemann during the autumn and winter of 1816–17.[11] There Müller (1794–1827)—together with the Stägemanns' daughter Hedwig (1799–1891); the artist Wilhelm Hensel (1794–1861); his sister, the poetess Luise Hensel (1798–1876); the publisher Friedrich Förster (1791–1868); and others—put together an informal drama structured around the story of a pretty miller's maid and her suitors. Each member of the group took a part, for which they wrote a verse script. In the enactment of the resulting *Liederspiel*, the participants would recite their verses, using whatever gestures or changes in vocal inflection the situation required in order to tell their part of the story.

The evidence of the poems and anecdotes that survive from the Stägemann circle suggests that its Liederspiel had a thoroughly tragic ending. In what can be reconstructed of the story they developed, the miller's daughter, named Rose (played by Hedwig von Stägemann), was courted not only by the young miller (played by Müller) but also by a country squire (played by Förster), a hunter (played by Wilhelm Hensel), and a gardener (played by Luise Hensel). The miller, after prevailing briefly over the other suitors, falls into despair when the hunter wins the miller maid's affections, and commits suicide. The maid, filled with remorse, then joins him in death.

During the three years after the meetings of the Stägemann circle, Müller, apparently intrigued by the possibilities of the story of the miller maid and her suitors, went on to create his own version of the story, consisting of twenty-five poems under the title *Die schöne Müllerin (Im Winter zum lesen)*, which was first published in 1821.[12] The cycle consists of a verse prologue (in the voice of the poet), twenty-

11. My account of *Die schöne Müllerin* follows Susan Youens's excellent work on the origins and historical context of the cycle, which is reported in three of her publications: Youens, "Behind the Scenes: *Die schöne Müllerin* before Schubert," *19th Century Music* 15 (1991): 3–22; idem, *Schubert: Die schöne Müllerin*, Cambridge Music Handbooks (Cambridge: Cambridge University Press, 1992); and idem, *Schubert, Müller, and Die schöne Müllerin* (Cambridge: Cambridge University Press, 1997). Another important source is Graham Johnson's copious notes for his recording of the cycle. See Johnson, "*Die schöne Müllerin*, The Poems and Their Background," in *Die schöne Müllerin, D795* (Hyperion, [CD] CDJ 33025, 1996).

Although the works I discuss here, both literary and musical, use "trockne" in their titles, the proper word is "trockene," which may be contracted to "trock'ne." For the sake of consistency, I shall use "trockne" throughout.

12. The cycle initially appeared as part of the fancifully titled *Siebenundsiebzig Gedichte aus den hinterlassen Papieren eines reisenden Waldhornisten*. The edition of *Die schöne Müllerin* upon which I have relied and that appears below is from Wilhelm Müller, *Vermischte Schriften*, ed. Gustav Schwab (Leipzig: F. A. Brockhaus, 1830), 1: 1–56.

three poems proper to the drama, and a verse epilogue (again in the voice of the poet), together intended to take about an hour to perform.[13] In poems 1–5, we are introduced to the miller, the millstream, the miller maid, her father, and the group of young workers in his charge. In poems 6–11, we become aware of the miller's growing infatuation with the maiden and the depths of his unrequited love. In poems 12–14, the miller becomes convinced that the miller maid is, as he says in poem 12, "Mine!" Despite this belief, there is absent from the poetic narrative any explanation for what changed his perspective from one of dreamy mooning to a self-assured sense of certain love; the poems provide no evidence that the miller maid has actually returned his love. In poems 15–17, the hunter suddenly enters the scene and, in the confused understanding of the miller, steals the miller maid away. Just as suddenly, the bliss of the preceding poems is replaced by anger, jealousy, and vindictiveness. In poems 18–21, the miller, torturing himself with visions of the hunter and miller maid in fond embrace, gradually begins to dwell more and more on death. It is in poem 21, "Trockne Blumen," that the miller decides to commit suicide, bringing to culmination the thoughts of death and the grave that came to dominate the preceding poems. Poem 22 is a sad dialogue between the miller and the stream, and poem 23 is given to the stream alone, as it gently and eerily caresses the drowned miller. The cycle concludes with an epilogue, which is curiously detached from the passion of the preceding poems, and which attempts to build a bridge from Müller's intense world of fantasy back to the world of the everyday.

The basic story told in "Trockne Blumen," the text and a translation of which are given here, can be seen to proceed in three stages.[14] In the first three stanzas, the miller addresses a collection of withered, moribund flowers that are unaccountably wet. In the second three stanzas, he contemplates the ineffectuality of his tears (the tacit cause of the wetness of the flowers) and the insignificance of his existence in the face of love that has died. This group ends with a stanza that echoes the first stanza of the poem, both in its return to the image of the faded flowers and in the recurrence, in reverse order, of the rhyming pair "gab" and "Grab." Together, these suggest a circular return that closes off the first six stanzas. The final stage of the story commences with the last two stanzas of the poem. These speak of a future in which the miller lies dead in his grave, in which the miller maid recognizes the truth of his love, and in which flowers bloom once more. The poem thus involves sharp contrasts: dying flowers set against blossoming flowers, winter set against spring, and dead love set against true love.

| TROCKNE BLUMEN | WITHERED FLOWERS |
|---|---|
| Ihr Blümlein alle, | All you flowers |
| Die sie mir gab, | That she gave to me, |

13. Müller writes in the prologue, "As it's wintertime I expect you won't regret a brief hour here in the countryside" (as it is depicted in the cycle of poems). ["Erhoffe, weil es grad' ist Winterzeit, / Thut euch ein Stündlein hier im Grün nicht Leid."] Müller, *Vermischte Schriften*, 1: 1.

14. There are other ways to parse the poem, as we shall see later in this discussion. The analysis I offer here is simply for the purposes of drawing out some of the aspects of the poem important for a number of its musical settings.

| | |
|---|---|
| Euch soll man legen | You shall be laid |
| Mit mir ins Grab. | With me in the grave. |
| | |
| Wie seht ihr alle | Why do you all |
| Mich an so weh, | Look at me so sadly, |
| Als ob ihr wüßtet, | As if you knew |
| Wie mir gescheh'? | What happened to me? |
| | |
| Ihr Blümlein alle, | All of you flowers, |
| Wie welk, wie blaß? | Why so faded, so pale? |
| Ihr Blümlein alle, | All of you flowers, |
| Wovon so naß? | Why are you so moist? |
| | |
| Ach, Tränen machen | Ah, tears do not make |
| Nicht maiengrün, | The greenness of May, |
| Machen tote Liebe | They do not make dead love |
| Nicht wieder blühn. | Bloom anew. |
| | |
| Und Lenz wird kommen | And spring will come |
| Und Winter wird gehn, | And winter will go, |
| Und Blümlein werden | And flowers will |
| Im Grase stehn. | Grow in the grass. |
| | |
| Und Blümlein liegen | And flowers lie |
| In meinem Grab, | In my grave, |
| Die Blümlein alle, | All the flowers |
| Die sie mir gab. | That she gave to me. |
| | |
| Und wenn sie wandelt | And when she walks |
| Am Hügel vorbei, | Past the mound, |
| Und denkt im Herzen: | And ponders in her heart: |
| »Der meint' es treu!« | "His love was true!" |
| | |
| Dann Blümlein alle, | Then all you flowers, |
| Heraus, heraus! | Come out, come out! |
| Der Mai ist kommen, | May is here, |
| Der Winter ist aus. | Winter is over. |

One thing immediately evident is that this poem relies, as do most, on conceptual blends. When the miller addresses the flowers, for instance, it is not as plants but as anthropomorphized beings, creatures that can not only look upon the miller but do so with pathos: "Wie seht ihr alle / Mich an so weh, / Als ob ihr wüßtet, / Wie mir gescheh'?" This, of course, is another example of the anthropomorphic blends discussed in chapter 2. Another blend is set up in the fourth stanza, where Müller combines concepts related to the seasons and to romantic love. Thus "maiengrün" becomes both a harbinger of spring and a sign of love's first bloom. Yet a third blend crops up in the seventh and eighth stanzas, when concepts related to the seasons combine with those of life and death: the consummatory "Der Mai ist kommen / Der Winter ist aus" represents both the future turn of seasons the miller envisions

and his rebirth as the eternal lover of the miller maid. The poem is thus a rich discourse structure in its own right, one with intriguing possibilities for pairing with the discourse structures of music.

## Bernhard Klein's "Trockne Blumen"

The Berlin composer Bernhard Klein (1793–1832) apparently discovered Müller's poetic cycle shortly after it was published and set "Trockne Blumen" as the last of his five *Lieder und Gesänge mit Begleitung des Piano Forte*. This collection was brought out by E. H. G. Christiani in 1822,[15] and the inclusion of "Trockne Blumen" is all the more remarkable in that the texts for the first four songs involve rather conventional stories of wanderlust and relatively innocent romantic encounters. "Trockne Blumen," as Klein clearly recognized, is a text of quite another order.

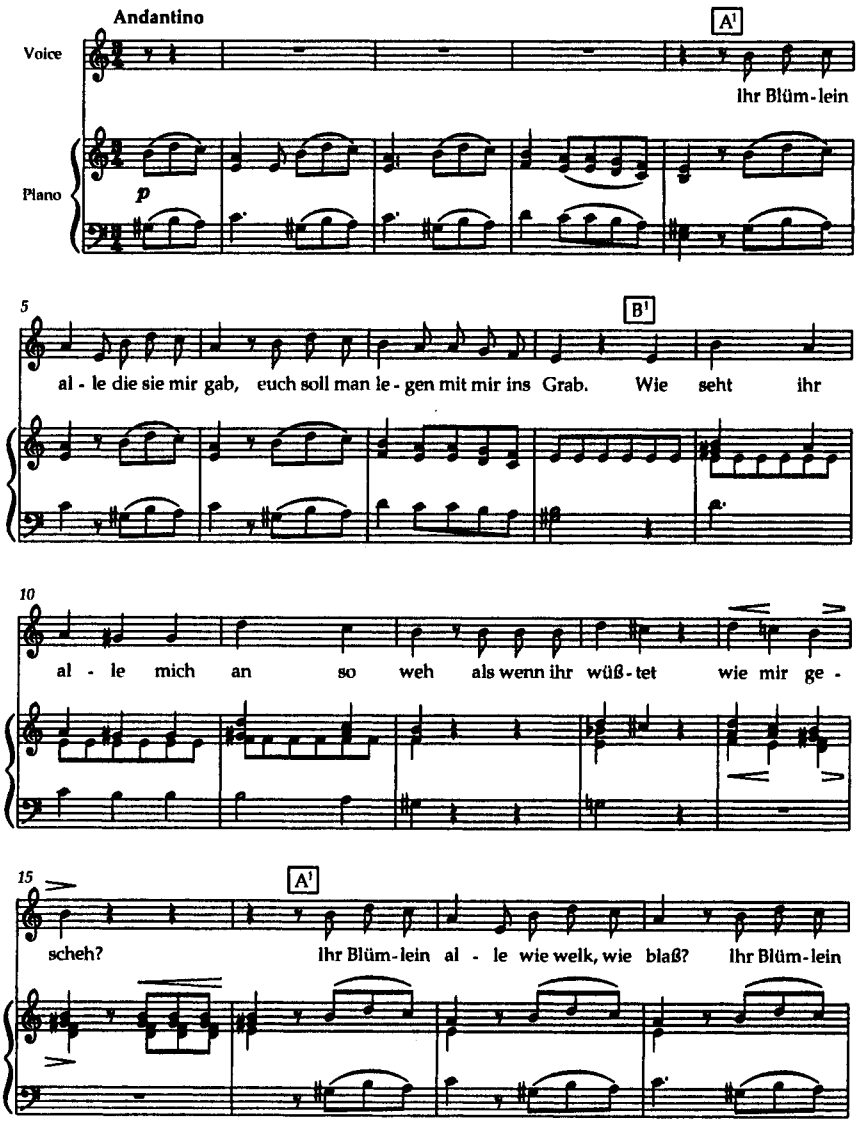
Klein's setting, the score for which is given in example 6.1, is guided by his division of the text into four strophes, each comprising two stanzas of the poem. Although the music for each strophe is different, so much material is reused that the song seems more strophic than through-composed, as the formal outline given in table 6.1 indicates.

The song begins with a brief piano introduction that sets out the material used to begin each of the first three strophes. This material, labeled $A^1$ in table 6.1, consists of three iterations of a motion from dominant to tonic, culminating in a passage in parallel sixths that returns to the dominant. The whole is couched in a spare two- and three-voice texture, with a repetitive melody built up from a compact succession of ascending and descending gestures that circulate around C5 and A4. Although A as tonic is scarcely in doubt, it falls short of attaining a full measure of stability, given the absence of root-position tonic triads in strong rhythmic positions and the emphasis on the dominant harmony at both the beginning and end of each $A^1$ section. Sections $B^1$ and $B^2$ consist entirely in prolongations of the dominant, either by temporarily transforming it into a leading-tone harmony (as in mm. 11–12 and mm. 20–24) or through a pathos-drenched passage that introduces an unprepared $VII^{\frac{4}{3}}$ of IV that leads through $IV^6$ back to VII, and then to V (as in mm. 13–16 and mm. 25–28). In each of the B sections, the melody struggles upward to D5 three times, only to fall back either to B4 or C♯5. Of these two sections, $B^2$ is slightly more intense. In it, a repeated F♮4 replaces the repeated E4 that signaled the opening of the $B^1$ section, giving rise to a $VII^{o7}$ that lasts from m. 20 to m. 24.

In the third strophe, Klein lightens the mood of the music by shifting to the par-

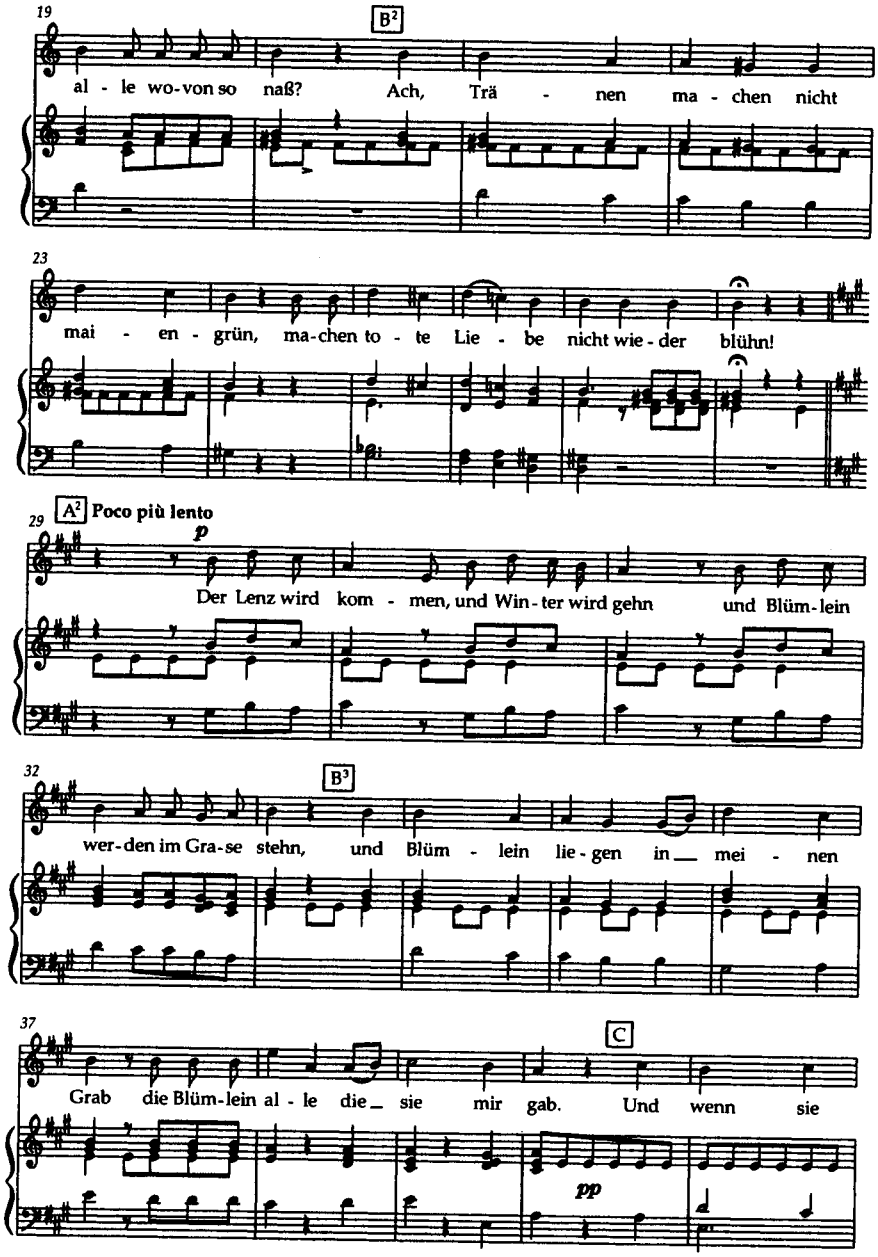
15. Müller is the author of all of the poems of Klein's *Lieder und Gesänge* of 1822; "Trockne Blumen" is the only poem Klein selected from *Die schöne Müllerin*. It is unclear what motivated Klein's choice of poet. In his letter of 15 December quoted above, Müller praised Klein's settings and asked if he would be interested in setting more of his poetry. Given the tone of the letter, it would seem unlikely that Müller was an active participant in the selection of texts for the 1822 collection of songs. I should note, as well, that another Berlin composer, Ludwig Berger, published a setting of "Trockne Blumen" in 1818; analysis of this work, as well as the context for its publication, can by found in Lawrence Zbikowski, "The Blossoms of 'Trockne Blumen': Music and Text in the Early Nineteenth Century," *Music Analysis* 18 (1999): 307–45.

EXAMPLE 6.1 Bernhard Klein, "Trockne Blumen"




allel major. The repeated notes that lent an air of intensity or obsession to the dominant prolongations of the preceding B sections now reappear to keep the music from stalling on the first and second beats of each bar of section $A^2$. Section $B^3$ begins with another prolongation of the dominant, but instead of being transformed into a leading-tone harmony in m. 36, as it had at the analogous juncture in sections $B^1$ and $B^2$, it now remains a dominant, leading initially to a $I^6$ in m. 38, and then to the first cadence on A thus far. With the arrival on the $I^6$ of m. 38, the melody finally reaches E5, preparatory to its arrival on A4 in m. 40.

EXAMPLE 6.1 *(continued)*




*(continued)*

EXAMPLE 6.1 *(continued)*

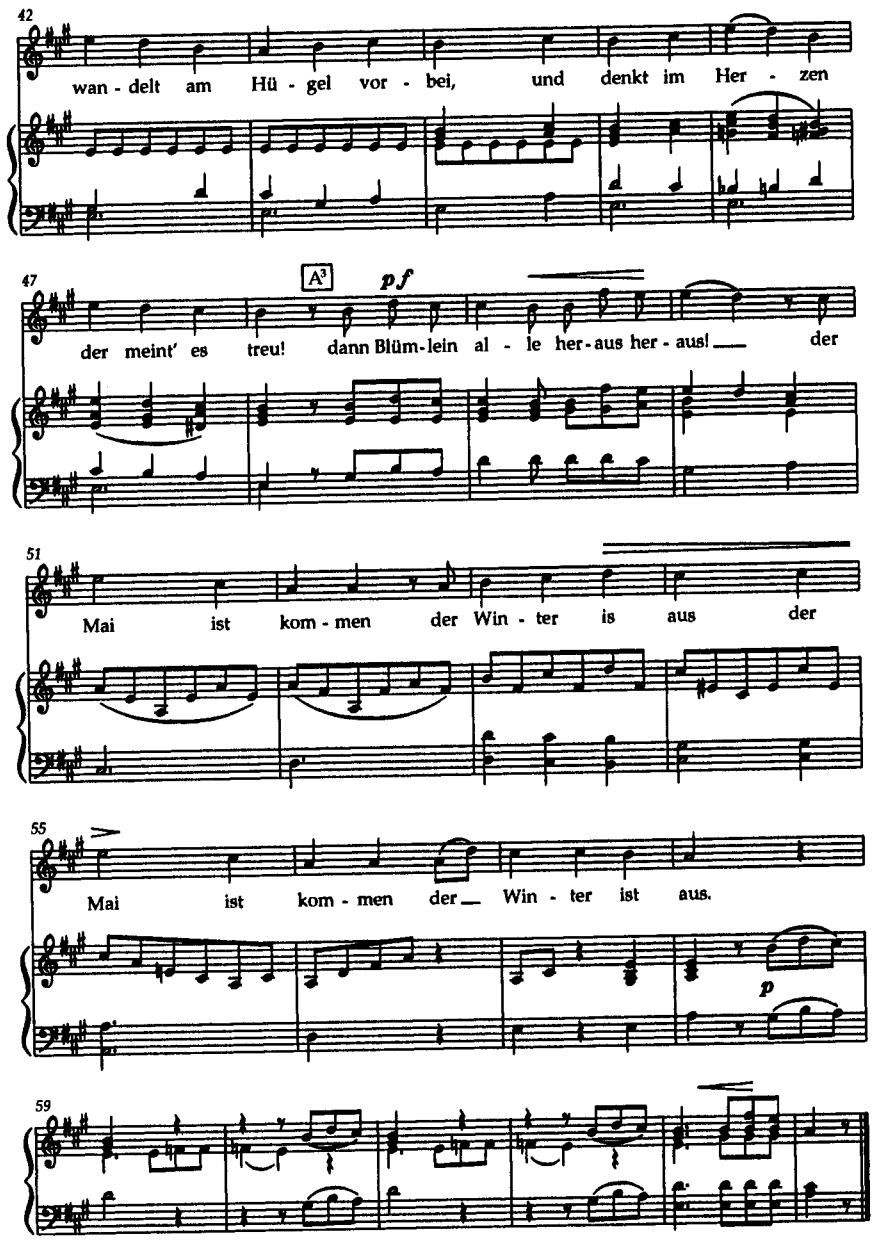


Table 6.1. Formal Outline of Bernard Klein's "Trockne Blumen"

| *Section and Key* | *Text Stanza* | *Musical Material and Melody* | *Measures* |
|---|---|---|---|
| Introduction A minor | | $A^1$, a | 0–4 |
| Strophe 1 | Stanza 1 | $A^1$, a | 4–8 |
| | Stanza 2 | $B^1$, b | 8–16 |
| Strophe 2 | Stanza 3 | $A^1$, a | 16–20 |
| | Stanza 4 | $B^2$, b | 20–28 |
| Strophe 3 A major | Stanza 5 | $A^1$, a | 29–33 |
| | Stanza 6 | $B^3$, b' | 33–40 |
| Strophe 4 | Stanza 7 | C, c | 40–48 |
| | Stanza 8 | $A^2$, a' | 48–58 |
| Postlude | | $A^2$, a' (fragment) | 58–64 |

The C section begins another prolongation of the dominant, one whose repeated notes are strongly reminiscent of sections $B^1$ and $B^2$. Absent, however, is the relatively active bass line that opened each of the latter sections; it is replaced here by a pedal on $E_3$. The melody, too, has changed, now peacefully circulating around $B_4$, with unimpeded ascents to $E_5$. After m. 45, the intensity of the section increases with the cessation of the repeated notes and the introduction of the chromatic second alto line of mm. 46–47. In m. 48, however, the melodic line that descends from $E_5$ once again comes to rest on a clear dominant harmony, which leads immediately to the reprise of the A material.

Although section $A^3$ opens with the same figure in parallel tenths that began the song (adjusted to fit the major mode), the figure now leads toward a dominant-seventh in m. 50 and then to a relatively extensive prolongation of the tonic in mm. 51–55. This prolongation prepares the way for the cadence of mm. 56–58 with which the vocal part concludes. A brief piano postlude follows; it is based on the opening of the $A^3$ section, but it includes a motion from F♮ to E that recalls the minor mode of the first half of the song.

Thus, Klein's music falls into two distinct parts. The first part, through m. 28, sets out an intense and unstable A minor. Relatively lengthy and obvious prolongations of large-scale harmonic dissonances (such as those associated with the dominant and leading-tone) and sudden interruptions mark the texture, as do persistent repeated notes. The second part, from m. 29 onward, disperses the intensity of A minor and replaces it with a goal-oriented and (for the most part) orderly A major. Here there are also prolongations of the dominant, but of a sort that seem to affirm rather than challenge the viability of tonic. What had been difficult in the first part becomes easy in the second; what had been troubled becomes calm.

This music, together with Müller's text, sets up the conceptual integration network (CIN) diagrammed in figure 6.1. Two observations should be made about the network represented by this diagram, observations that extend to all of the CINs discussed in this chapter. First, as was the case in chapter 2, the diagram represents a kind of analytical snapshot of the elements and relations that make up the network, and that contribute to the emergent structure of the conceptual blend. A

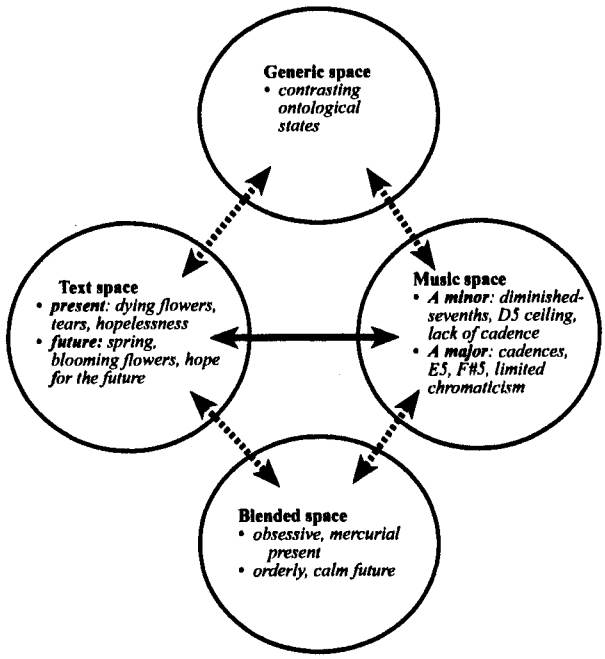



FIGURE 6.1 CIN for Klein's "Trockne Blumen"

fuller account of the way concepts combine within this network would require a series of such snapshots. Second, what I am concerned with here is the *overall* discourse set up by the text and the *overall* discourse set up by the music. Although there are interesting details at more local levels, I am most intrigued by what conceptual blending can tell us about song and by what song can tell us about conceptual blending. Let me expand on this for just a moment before returning to Klein's "Trockne Blumen."

The theory of conceptual blending assumes that there are structural invariances between the input spaces of a blend: these invariances, encapsulated in the elements and relations of the generic space, are what make conceptual blends possible. In the case of songs, the invariances are between the mental space set up by the text and the mental space set up by the music. Put another way, the fact that combinations of text and music can give rise to conceptual blends suggests that there are syntactic correspondences between linguistic and musical discourse. Although the syntax of most nineteenth-century German poetry is usually clear at or below the level of the stanza, the syntax of songs from the same milieu needs to be conceived in terms of the song as a whole, which was construed as a single, if often extended, musical utterance. Because of this, I find it most useful to focus on conceptual blends that encompass the entirety of a given song—that is, blends that arise from the overall discourse set up by the text and the overall discourse set up by the music.[16]

16. This perspective also leads me to regard the poem set by the composer as a text, rather than a poem. Susan Youens, for different reasons, advocates a similar approach; see Youens, *Retracing a Winter's*

What nineteenth-century songs can tell us about conceptual blending stems from the characteristic of musical syntax I have just described. The blend set up by a song (assuming that there are structural correlates between the mental spaces set up by the text and the music) spans at least two or three minutes, involves two different media (with two different, if related, kinds of syntax), and engages us on both emotional and intellectual levels. All these features have the potential to add immeasurably to our understanding of conceptual blending as a general cognitive process.[17]

With these thoughts in mind, let us return to the CIN set up by Klein's song. The generic space for the CIN is structured around the notion of two linked but contrasting ontological states. Such states are often temporally indexed (as past and present or as present and future), but they can be more abstractly represented as theme and variation, or prototype and copy. The delineation of two ontological states often provides the basis for a more extensive interpretive framework. This framework then furnishes an account of the cause for the change of state or an explanation for why one state is to be preferred over another.

Contrast between the two ontological states is supported in Müller's poem by an emphasis on the present in the first four stanzas and an emphasis on the future in the second four stanzas. There is also far less activity described in the first four stanzas than in the second four. Of course, there are exceptions to both of these generalizations. In the first half of the poem, the future is summoned at the end of the first stanza and the past is alluded to in the fourth; the gaze of the flowers and the tears of the miller suggest that the scene is not wholly devoid of activity. In the second half of the poem, the past is evoked briefly in the last two lines of the sixth stanza, the first two lines of which call up the embodiment of lifelessness.

In the mental space of the music, the linked but contrasting ontological states of the generic space are set up by the A minor and A major music of the two halves of the song. These states are linked less by temporal index than by conceptual frame: the A minor music functions as the theme, as it were, and the A major music as the variation. Nonetheless, the A major music *rewrites* the character and syntax of the A minor music. It exchanges the third and sixth scale steps of minor for those of major, provides the first cadences on tonic, avoids the fully diminished seventh chords omnipresent in the first half of the song, and allows the melody to ascend to E5 and F♯5. As a result, it enacts a specific shift of time frame by placing the A minor music firmly in the past and announcing the A major music as the new present. This shift is underlined by the mirrored processes that occupy each half of the song: over the course of the first half of the song, the music becomes more intense

---

*Journey: Schubert's Winterreise* (Ithaca, N.Y.: Cornell University Press, 1991), 307–12. As will also become clear, I am typically concerned with the story told by the text as a whole, a parabolic level discussed by Mark Turner in *The Literary Mind* (New York: Oxford University Press, 1996), esp. chap. 1.

17. Although it should be clear from the foregoing, let me emphasize that not all songs will set up conceptual blends. Among those songs that do set up blends, some blends will be more successful (that is, lead to more extended elaborations) than others. One example of the latter situation is provided by the third phrase of Brahms's "In der Fremde"; see the discussion of the song later in this chapter.

and darker; over the course of the second half, the music becomes less intense (with the exception of the poignant recollection of mm. 46–47) and brighter.

Correlations between these idealized representations of contrasting ontological states provide an opportunity for the combination of concepts in the blended space, which draws upon elements of the input spaces to create a portrayal of the miller's transformation from a tortured soul into a hopeful and ecstatic one. As was the case in chapter 2, the emergent structure of conceptual blends (including the present one) can be described in terms of the processes of composition, completion, and elaboration.

COMPOSITION The combining of elements from the text-space and the music-space is quite straightforward. The enervated present of the first four stanzas and the A minor music combine to create a representation of one ontological state in the blended space: a stark image of the present as dominated by obsessive thoughts, scant opportunity for rest, and a mercurial play of emotions. The active future of the second four stanzas and the A major music combine to create a representation of the second ontological state: what was shadowy has become clear; there is an orderly progress toward goals; and the surroundings are, almost without exception, calm and soothing. The transformation of one state into the other occasions a strange conceptual inversion in this part of the song: the flowers that blossom after the change to A major in m. 29 seem to be the same ones that the miller predicts will lie in his grave (cf. mm. 31–33 and mm. 34–40), even though the latter event, according to the logic of the poem, must *precede* the former.

COMPLETION Building on the context provided by this blend of concepts, we can imagine outward signs of the miller's torment in the first half of the song—a sluggish, depressed body language, accompanied by sudden expressions of anguish as his agony increases—and how, in the second half of the song, these give way to gestures and behavior associated with peacefulness and happiness as the miller contemplates his future. We can also imagine that, overall, the miller grows happier and more at peace with himself as he becomes more distant from his earlier torment. This inference is supported in part by the music of the concluding $A^3$ section, where the melody breaks through to the high F♯5 of m. 49, a moment followed by the long prolongation of tonic in mm. 51–55 that prepares the singer's final cadence.

ELABORATION Elaboration of the blend suggests that the miller will act in order to make the future he imagines a reality. His "Heraus, heraus!" thus calls forth not only the future's flowers but also the future itself. Elaboration may also provide an explanation for how the state of affairs portrayed in the first half of the song is transformed into the state of affairs portrayed in the second half. Here the music provides a useful guide, for it suggests that the present can be transformed into the future by reinterpreting it, just as the musical materials from A minor were reinterpreted in A major. Thus, the dead love of the present will become the eternal love of the future; the dying blossoms will become living ones. The strange celebration that is enacted in mm. 34–40, as the miller contemplates the flowers that will be laid

in his grave, then becomes more comprehensible, for in the second half of the song the withered flowers that he holds are transformed into symbols of his eternal love.

## Franz Schubert's "Trockne Blumen"

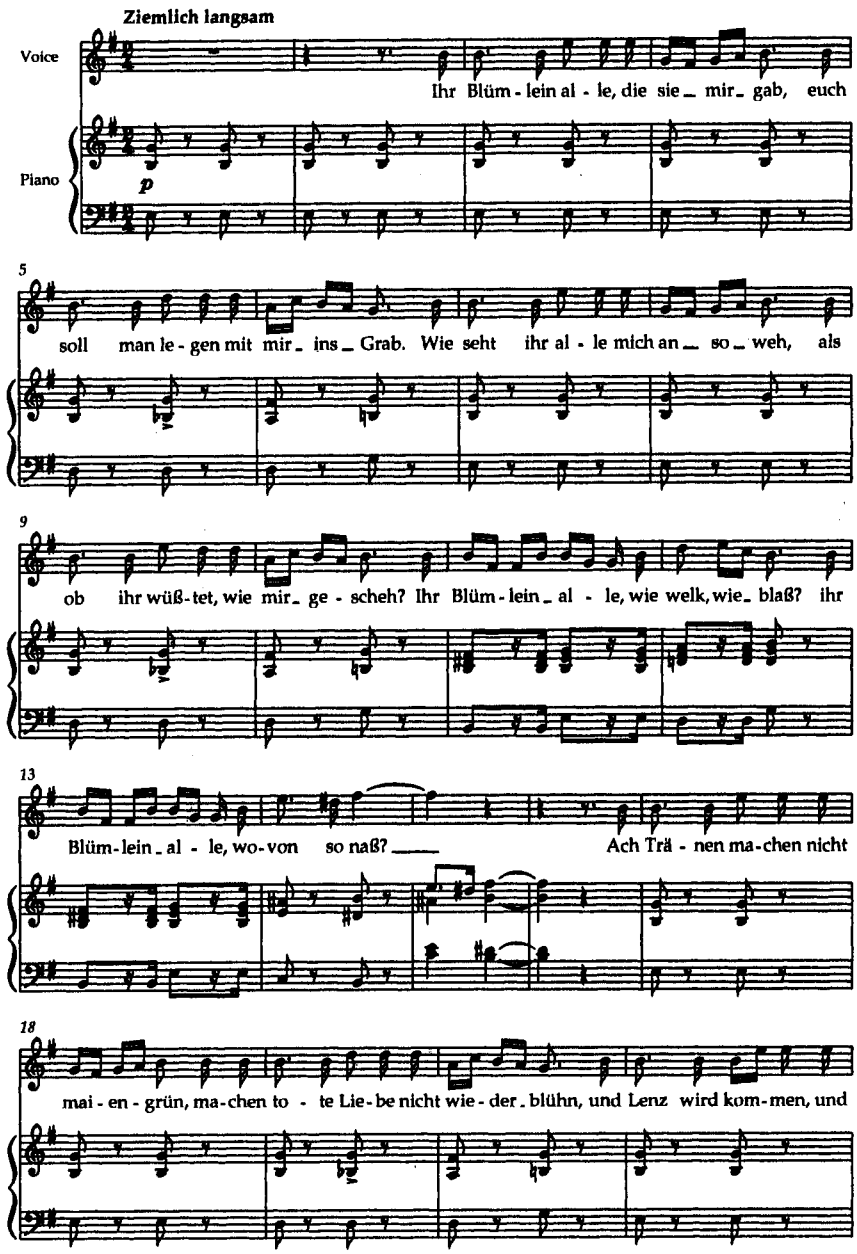
Although the exact circumstances of Franz Schubert's encounter with Müller's *Die schöne Müllerin* are not known, it appears that Schubert came upon the poetic cycle sometime in the first part of 1823 and set twenty of the twenty-five poems during the second half of the same year. The completed cycle was published the following year by the firm of Sauer and Leidesdorf. "Trockne Blumen," the score for which is given in example 6.2, is the eighteenth song in the cycle. Schubert's setting divides the eight stanzas of the poem into a 3 + 3 + 2 grouping. The music in E minor for the first three stanzas is repeated for the second three, after which there is a shift to the parallel major for the final two stanzas. Schubert does not repeat any text in his setting of stanzas 1–6, but he does repeat the last two stanzas in the pattern 7–8, 7–8, 8.

The song begins with the most minimal of piano introductions: two bars of a repeated E minor chord, played *piano*, each eighth-note chord followed by an eighth rest. This pattern continues as the voice enters with a bare melody outlining E minor. With the shift of the bass from E3 to D3 in m. 5, the E minor chord is transformed into a G chord in $\frac{6}{4}$ position. The latter leads to a cadence on G on the second beat of m. 6, coincident with the conclusion of the first stanza of text. The music of mm. 3–6 is then repeated for the second stanza in mm. 7–10, sustaining the spare effects created by the minimal melodic and harmonic materials, but adding the barest hint of forward motion when the voice closes on B4 instead of on G4, as it had in m. 6.

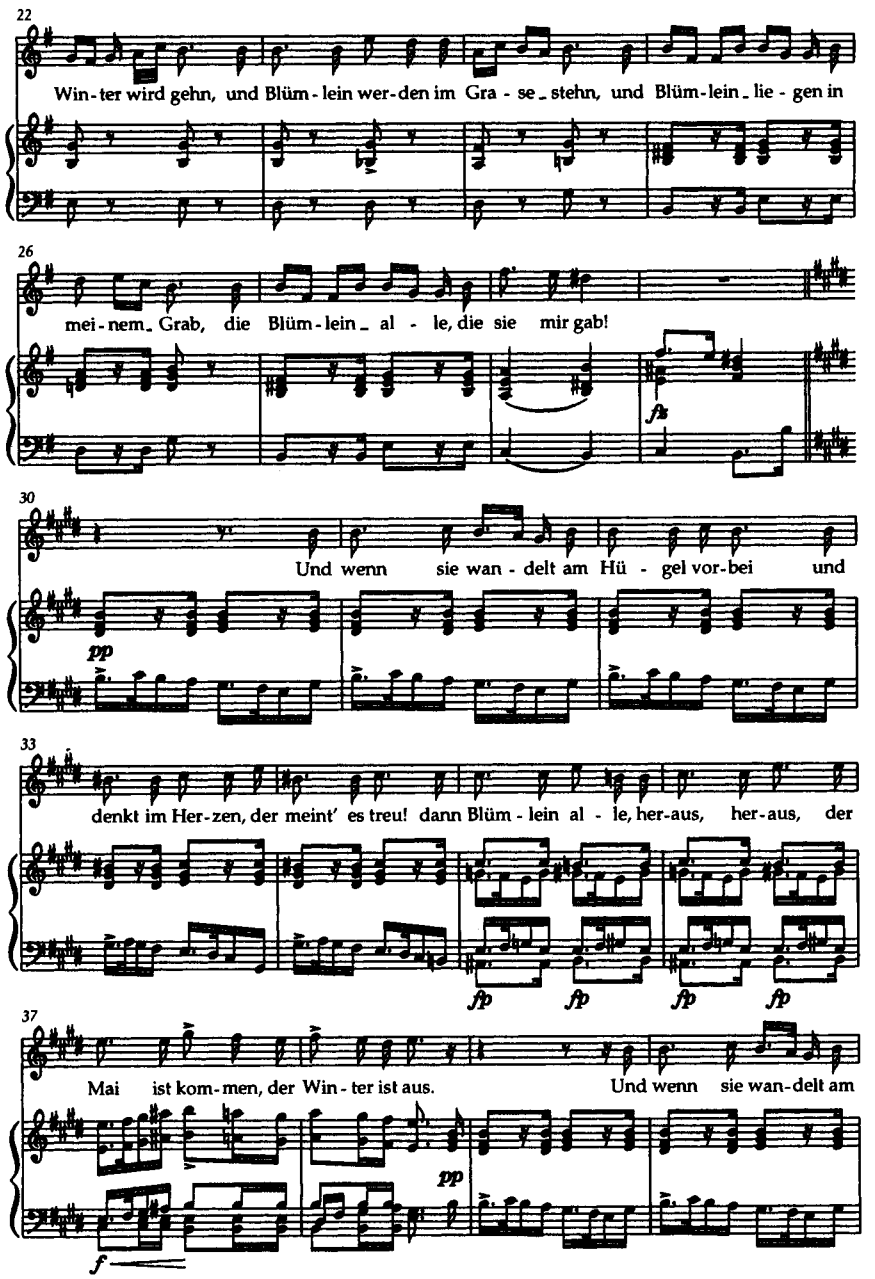
In m. 11, the music becomes more active and its harmonic material somewhat more variegated: the rhythmic pattern of the accompaniment is livelier; the harmonies change every beat; and the melody, as it follows these harmonic changes, becomes more animated. The initial harmonic strategy for the setting of the third stanza is, in compressed form, similar to that used for the first two stanzas: E minor, now preceded by its dominant in m. 11, leads again to G in m. 12. Measure 11 is repeated in m. 13, but this time it is followed by an Italian augmented-sixth chord that leads to the dominant of E minor, while the singer reaches past E5 for the first time to grasp a high F♯5. The piano then echoes this gesture, sustaining the dominant chord for a half cadence on B that lasts through m. 16. Schubert's setting of stanzas 4–6 in mm. 17–29 is virtually identical to the music he provided for the first three stanzas; the only significant change occurs in mm. 28–29, where the E5-D♯5-F♯5 of the melody is restated as F♯5-E5-D♯5.

With the shift to the parallel major in m. 30, the accompaniment becomes yet more active: Schubert adds a lightly tripping bass melody to the animated rhythmic pattern of the right hand. In m. 31, the singer takes up the bass melody, in modified form. With the momentary shift toward C♯ minor in mm. 33–34, however, the melody becomes streamlined once again. The C♯ minor chord in m. 34 yields in the following measure to another augmented-sixth chord. This leads (in m. 37) to a cadential $\frac{6}{4}$, which prepares the first cadence on E in the song. Schubert

EXAMPLE 6.2 Franz Schubert, "Trockne Blumen," from *Die schöne Müllerin*, D. 795/18

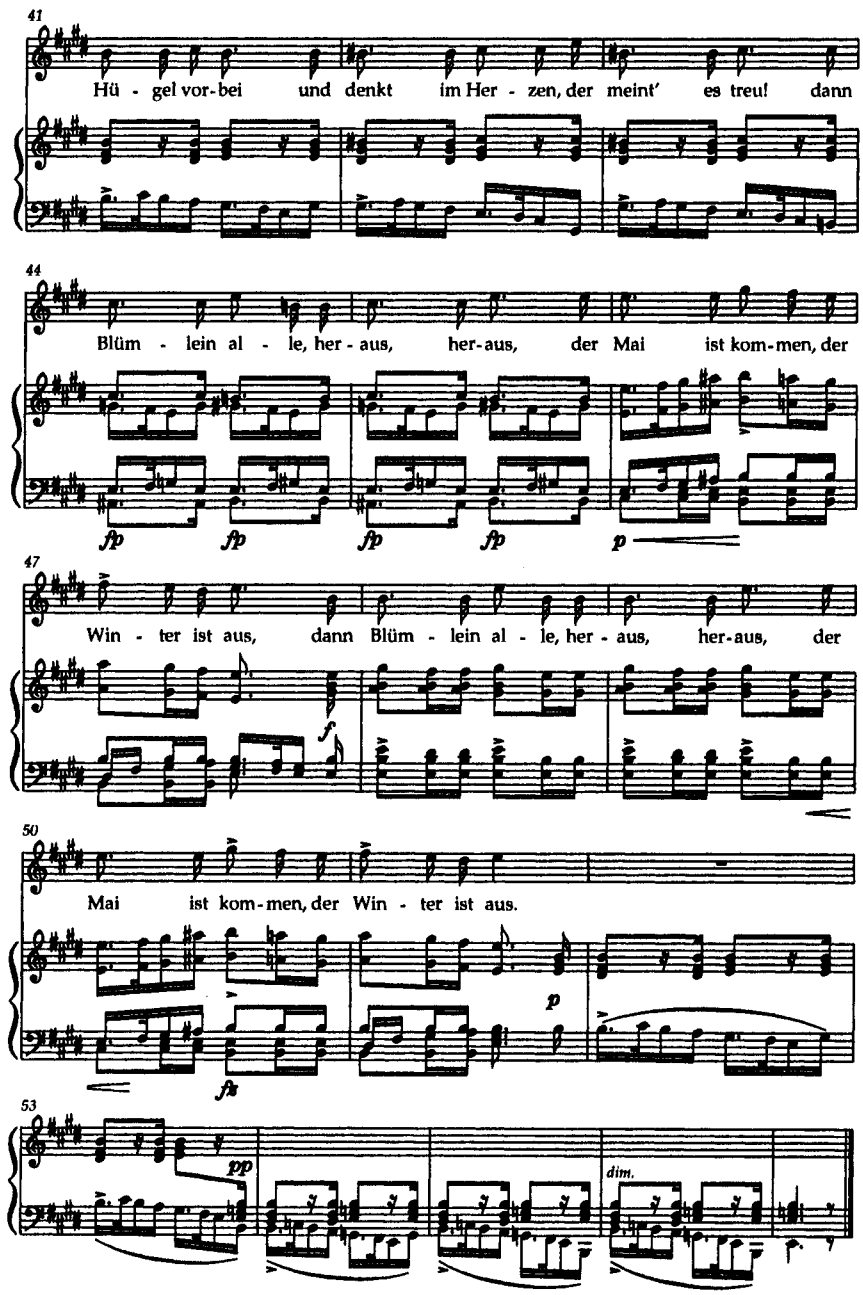


EXAMPLE 6.2 *(continued)*




*(continued)*

EXAMPLE 6.2 *(continued)*

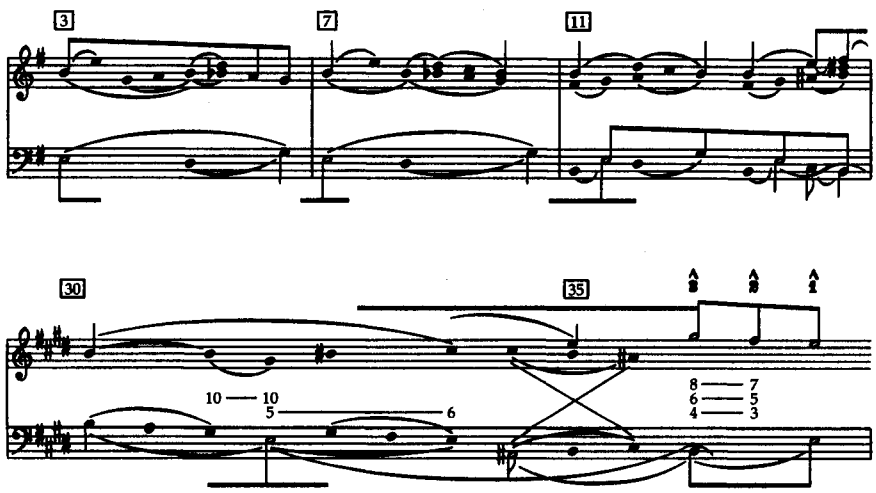


immediately repeats the whole of mm. 30–38 in mm. 39–47 and then, without pause, repeats the last four bars, with a change of bass and a slight thickening of texture. A six-bar piano postlude, reprising the accompanimental figure of m. 30, concludes the song. In the course of the postlude, the register drops an octave, and the mode shifts to minor for the last four bars.

Although this description cannot do it justice, the effect of the relatively spare musical materials that Schubert employs in this setting is profound. A somewhat more nuanced view of the overall musical discourse can be gained through a middleground sketch of the voice-leading of mm. 1–16 and 30–38, given in example 6.3. As the sketch shows, in the first portion of the song the repeated motions toward G major simply prolong E minor and the initial B4 of the Anstieg. It is only in m. 14 that the music finally begins to move, wrenched from its enervated reiteration of tonic (or the tonic-substitute mediant) by the augmented-sixth chord that leads to the dominant. The augmented-sixth chord also provides the impetus for the continuation of the Anstieg, which is temporarily left hanging on the F♯5 over the dominant. Given the long delay of the continuation of the Anstieg, almost the entirety of mm. 1–29 remains stalled on a prolongation of tonic and B4. Only at the half cadences of mm. 15–16 and mm. 28–29 does the music move away from B4 and momentarily struggle upward toward the Kopfton. With the shift to the parallel major in m. 30 the oppressive harmonic stasis of the first half of the song lifts, and the music begins to move forward, pulled along by the extended dominant preparation of mm. 33–37. Finally, in m. 37, the voice achieves the Kopfton, but as G♯5 rather than G♮5. This accomplishment is followed immediately by the descent of the fundamental line, supported by the first cadence on tonic of the entire song.

Schubert's music falls into two distinct parts in a way that parallels Klein's setting of Müller's poem. The first part, up to m. 29, presents a minimal, washed-out

EXAMPLE 6.3 Middleground sketch of mm. 1–16 and 30–38 of Schubert's "Trockne Blumen"



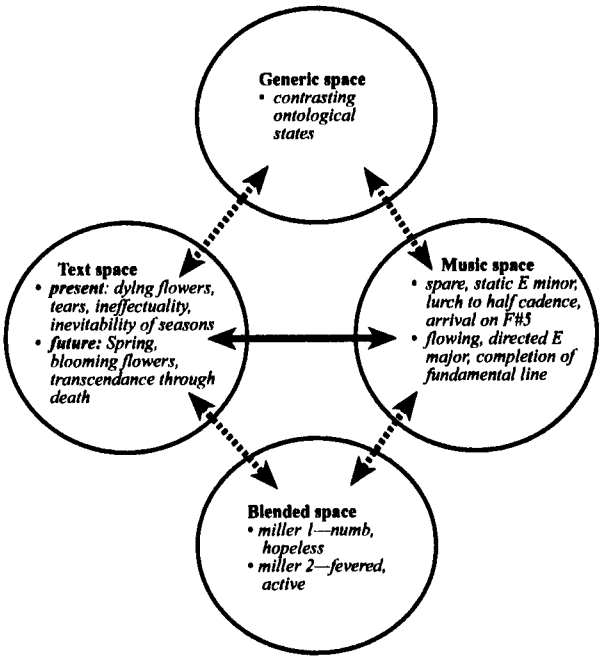



FIGURE 6.2 CIN for Schubert's "Trockne Blumen"

E minor interrupted twice by rather overwrought cadences on the dominant. The second part, from m. 30 on, moves forward, and without so much as a backward look culminates in authoritative and emphatic cadences on E major.

This music, together with Müller's text, sets up the CIN diagrammed in figure 6.2. The generic space, like that for Klein's "Trockne Blumen," is structured around the notion of two linked but contrasting ontological states. However, where in Klein's setting the emphasis is on connections between the ontological states, in Schubert's setting the emphasis is on their incommensurability. Klein's approach takes the point of view of an observer: the undesirable situation represented by the first ontological state can be improved by reconstructing it in terms of the second. Schubert's approach takes the point of view of a participant: recognizing that the first ontological state offers no hope, it is simply abandoned in favor of the second.

The change of emphasis in the generic space reflects a reinterpretation of the elements of the textual space, prompted by Schubert's division of the text into a group of six stanzas followed by a group of two stanzas (which are subsequently repeated). While the ontological states can still be characterized as present and future, the index is now emotional rather than temporal. The present is still occupied with dying flowers, thoughts of the grave, tears, and dead love, but it is also suffused with a sort of numbness, born of the miller's inability to bring about any change in his circumstances. He barely realizes that the tears that moisten the dying flowers are his own; he is powerless to stop them, just as he is powerless to stop the

coming of the seasons, or of his own death. Thus the future of "Und Lenz wird kommen / Und Winter wird gehn" is contained within the present, a leaden emotional domain circumscribed by the echoing rhymes of the first and sixth stanzas. By contrast, the future state called forth by the last two stanzas of the poem is devoid of such dark thoughts: here the miller maid wanders past the resting place of the miller; here spring has truly arrived, and the winter of despair has been banished.

Contrast between ontological states is also apparent in the musical space. Schubert's E minor music is, for the most part, spare and static. After m. 11, it gradually pulls itself into an agitated state, lurches into a half cadence, and then returns to its barren beginning. His E major music, in contrast, exploits and adds to the quickening of pulse that precedes the half cadences of the E minor music, and it drives toward a consummatory cadence that effaces the equivocation and irresolve of the first half of the song. This is music of possibility and completion, of *action*.

Correlations between these idealized representations of contrasting ontological states provide an opportunity for the combination of concepts in the blended space, which draws on elements of the input spaces to create a portrayal of two very different personae for the miller.

COMPOSITION Composition provides the concepts basic to this portrayal. The E minor music and the first six stanzas of text combine to summon a miller who sits in numb contemplation of what seems to be a hopeless predicament. When the gravity of his situation seizes him, he stirs for a moment (responding to the augmented-sixth chords), but even the intensity of torment is denied him and he sinks back into his torpor. The E major music and the last two stanzas portray a quite different miller who, with growing confidence, reaches out toward goals—most particularly, achieving immortality through true love—beyond the grasp of the miller's first persona.

COMPLETION Through completion we can build on what we know about human behavior in such extreme states: we can imagine that in his first persona the miller is virtually inert until, at the moment of his anguish, he half rises, a look of panic on his face; similarly, we can imagine the expansive gestures of the miller in his second persona and perhaps also see the somewhat fevered look in his eyes as he proclaims "Der Mai ist kommen / Der Winter ist aus."

ELABORATION While the emergent structure produced by composition and completion offers a wealth of material to the imagination, it is elaboration that offers the most striking interpretive possibilities. Elaboration of the blend suggests that the transition from the first persona to the second marks a decisive turn in the miller's story. With this transition, the miller has moved beyond merely *entertaining* the thought of death and has now *accepted* the fact of it. The miller will overcome his ineffectuality, he will transform himself, and he will do so through suicide. Supporting this interpretation are the almost frenetic repetitions of the second half of the song (as though the miller still needs to convince himself of the necessity of his

course of action) and the sad turning toward minor of the postlude as the enormity of his decision begins to sink in.[18]

## Two Settings, Two Millers

These two settings of Müller's "Trockne Blumen" give us two rather different millers. We first come upon Klein's miller turning over the shards of his shattered love affair and becoming more distraught as he realizes the hopelessness of his predicament. Then there is a sudden shift, and what was dark and foreboding becomes light and promising. The miller's burden has been lifted, for he has found a way to breathe life into dead flowers and dead love: he will welcome death rather than fear it. In contrast, Schubert's miller has two strikingly different personae. The first is emptied out, nearly paralyzed by despair, numbly aware of his misery and his inability to escape it. The second is effusive, expansive in welcoming a future in which the miller is no more. Schubert's miller is, in short, psychotic: at first immobile in the face of death and then feverishly and joyously embracing it.

Are both millers present in Müller's poem? Yes and no. Yes, in that his poem can support these different interpretations, as the CINs above have shown. And yes, in that, from the perspective of the poetic cycle as a whole, the miller is a truly complex character, a quiet psychopath destined for destruction. But no, in that it requires the agency of the music to bring forth these *specific* millers.

The differences between these two millers suggests that a review of the different compositional strategies that brought them into being is warranted so that we may better understand the role of musical syntax in breathing them to life. The similarities between these two settings of Müller's poem are striking: both employ a shift from minor to major mode; both avoid any definitive cadence on the tonic in the first portion of the song; and both make use of relatively delicate textures, especially in the first half of the song.[19] More distinctive are Klein's and Schubert's respective treatments of dominant and tonic. Klein, through extended prolongations, employs the dominant as a tension-generating device, and in consequence the dominant receives more emphasis than the tonic. By contrast, Schubert is sparing with the dominant of his home key, reserving it as a device to pull the music temporarily out of its stasis (in the first half of the song) or using it as part of a normative cadence (in the second half of the song). Tonic, by comparison, is given quite a bit of emphasis, so much so that the first portion of the song becomes threatened with immobility.

The conceptual blends created by Klein's and Schubert's songs, then, are not sim-

18. Further evidence for the moment of this decision is provided by the poetic cycle as a whole. In all of the poems up to "Trockne Blumen," the miller's presence is marked by the use of the first person. After the shift to the third person in "Trockne Blumen," the miller never appears in the first person again.

19. In addition to the similarities between Klein's and Schubert's settings of Müller's poem, both also share features with Ludwig Berger's 1818 setting of the poem, which was mentioned in n. 15; the common features are discussed in Zbikowski, "The Blossoms of 'Trockne Blumen,'" 335. Although such similarities are provocative, no research of which I am aware traces specific patterns of influence between Berger, Klein, and Schubert.

ply a matter of words and music happening at the same time (and thus being "combined" by dint of sharing the same temporal frame). For blending to take place, the mental spaces set up by the text and the music must have a shared topography, reflected in part in commonalities between the syntactic structures of the two media. This shared topography does not exhaust the resources of the input spaces. Indeed, the difference in syntax important to Nicholas Cook's conception of multimedia (discussed at the beginning of this chapter) is no less important in conceptual blending, for it ensures that the stories told by both music and text are compelling in themselves. Nonetheless, common structural features (which may obtain at a different level of structure than those that lead to contest between media) are necessary for conceptual blending to take place.

In the settings of "Trockne Blumen," words and music work together to tell stories that enhance and expand on structural features inherent in the source spaces for each conceptual blend. I now turn to a song by Robert Schumann in which the mental space set up by the music makes possible a conceptual blend whose emergent structure includes a narrative considerably more extensive than that suggested by the text alone.

## "IM RHEIN"

In 1840, having established himself as a composer of works for the piano, Robert Schumann turned his energies to song. And formidable energies they were, for by the time the year had ended, he had written over 140 songs. Schumann quite obviously worked at speed: the twenty songs in the original version of his *Dichterliebe* cycle (four were cut before its publication) were written from 24 May to 1 June. Nonetheless, among the songs written that year are some of the most profound and haunting of the nineteenth century, in part because Schumann found ways to combine words and music to create fantastic worlds that extend far beyond those that can be summoned by either words or music alone.

### Heine's "Im Rhein"

"Im Rhein," the text and translation of which are given here, is the sixth of sixteen songs that make up Schumann's *Dichterliebe*, first published by C. F. Peters in 1844.[20] The text for all of the songs was taken from Heinrich Heine's *Lyriches Intermezzo*, a collection of sixty-five poems published in 1823. Although usually termed a song cycle, *Dichterliebe* is held together not so much by common narrative as by theme: all of the poems deal with romantic love, which is transformed, splintered, or distorted by reminiscence or overheated emotions.

| IM RHEIN | IN THE RHINE |
|---|---|
| Im Rhein, im heiligen Strome, | In the Rhine, in the holy river, |
| Da spiegelt sich in den Well'n, | There, reflected in its waves, |

20. The text for "Im Rhein" given here is from the song as it appears in Clara Schumann's edition of the Lieder of Robert Schumann, but with text repetitions omitted.

| | |
|---|---|
| Mit seinem großen Dome | With its great cathedral, |
| Das große, heilige Köln. | Is the great holy city, Cologne. |
| | |
| Im Dom, da steht ein Bildnis, | In the cathedral there hangs a portrait |
| Auf goldenem Leder gemalt; | Painted on gilded leather; |
| In meines Lebens Wildnis | Into my life's wilderness |
| Hat's freundlich hineingestrahlt. | It has shone down kindly. |
| | |
| Es schweben Blumen und Englein | Flowers and cherubs float |
| Um Unsere Liebe Frau, | Around Our Dear Lady. |
| Die Augen, die Lippen, die Wänglein, | Her eyes, her lips, her cheeks |
| Die gleichen der Liebsten genau. | Are exactly like those of my love. |

The three stanzas of "Im Rhein" can be divided into two parts. The first part, which comprises the first stanza, summons a broad view of Cologne centered on its cathedral and reflected in the waves of the Rhein. There is a sense of vastness to the perspective (suggested by the repetitions of *groß*), and of the holy (provided by the cathedral and the description of Cologne as the "holy" city).[21] The second part of the poem, which comprises the second and third stanzas, moves inside the cathedral to focus on a portrait within. It is at this point that the presence of a narrator is made explicit and the shifting perspective of the poem acquires an anchor and a rationale. We are here in this city, inside this cathedral, to view a painting of singular importance to the narrator, for its radiance has proved a balm to the troubles of his soul. In the third stanza, the poetic lines become longer, and the narrator takes us inside the painting itself. He begins by describing a heavenly vision of the Madonna, but as he takes us deeper into the painting, the image of the Madonna becomes ever more worldly, until it becomes apparent that the narrator savors this image because it is the exact likeness of his beloved. The poem ends with this sudden shift from the spiritual to the carnal, and we hang in mid-air, suspended both by and within the narrator's vision.

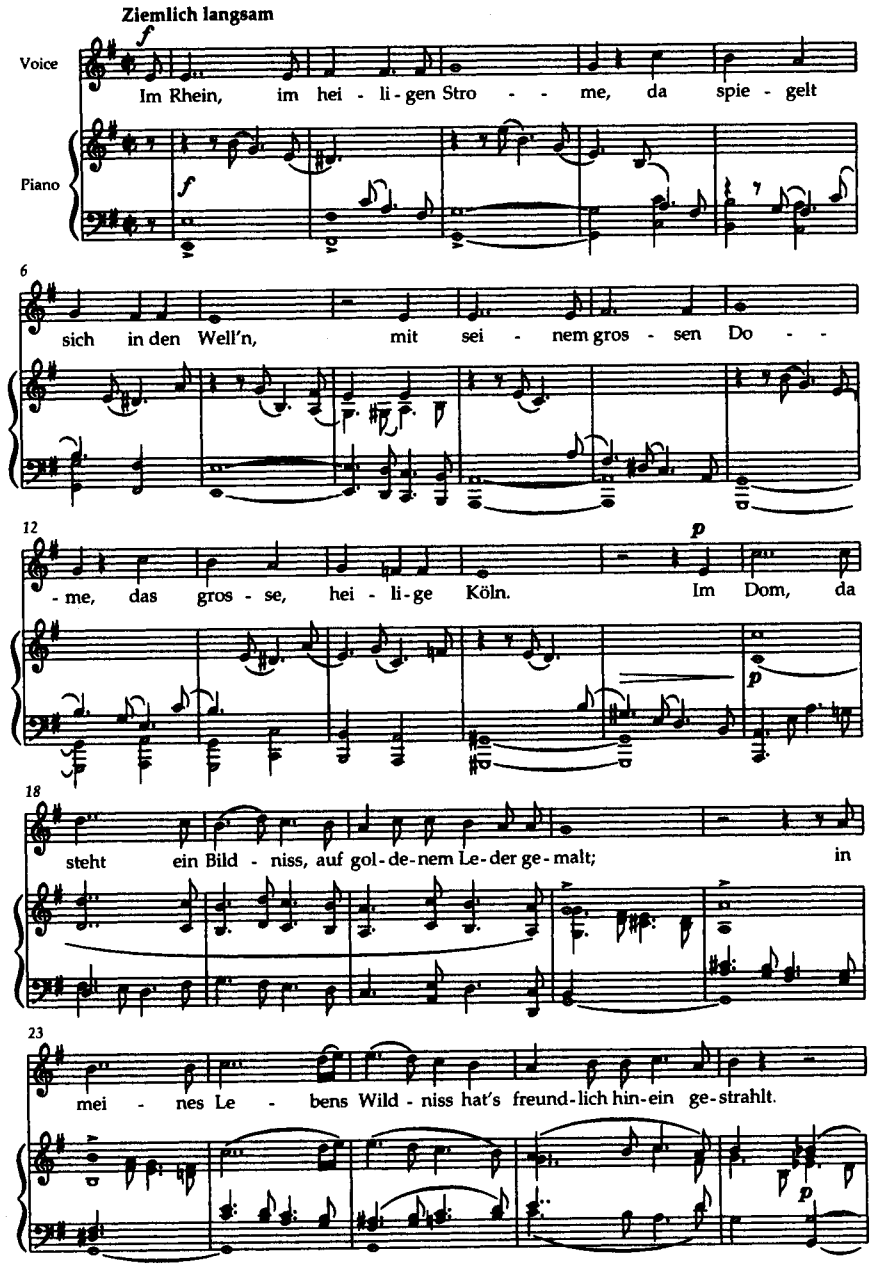
The elegance and power of Heine's unsettling poem stem from two parallel processes. The first is perspectival: we start from the largest scale—the city of Cologne and its cathedral—and then narrow the scope, first to the interior of the cathedral, then to a painting within, and then to specific features of the face portrayed in the painting. The second process is conceptual and affectual: we begin with ideas about vastness and holiness; proceed to the active solace provided by the contemplation of a religious work of art; and then are suddenly confronted with an intense, carnal passion for a lover.

## Schumann's "Im Rhein"

Schumann's setting of the poem "Im Rhein," the score for which is given in example 6.4, is in ABA' form. The first sixteen measures, which make up the A section, present a forceful, leaden E minor. Beginning *forte*, the voice (accompanied by the

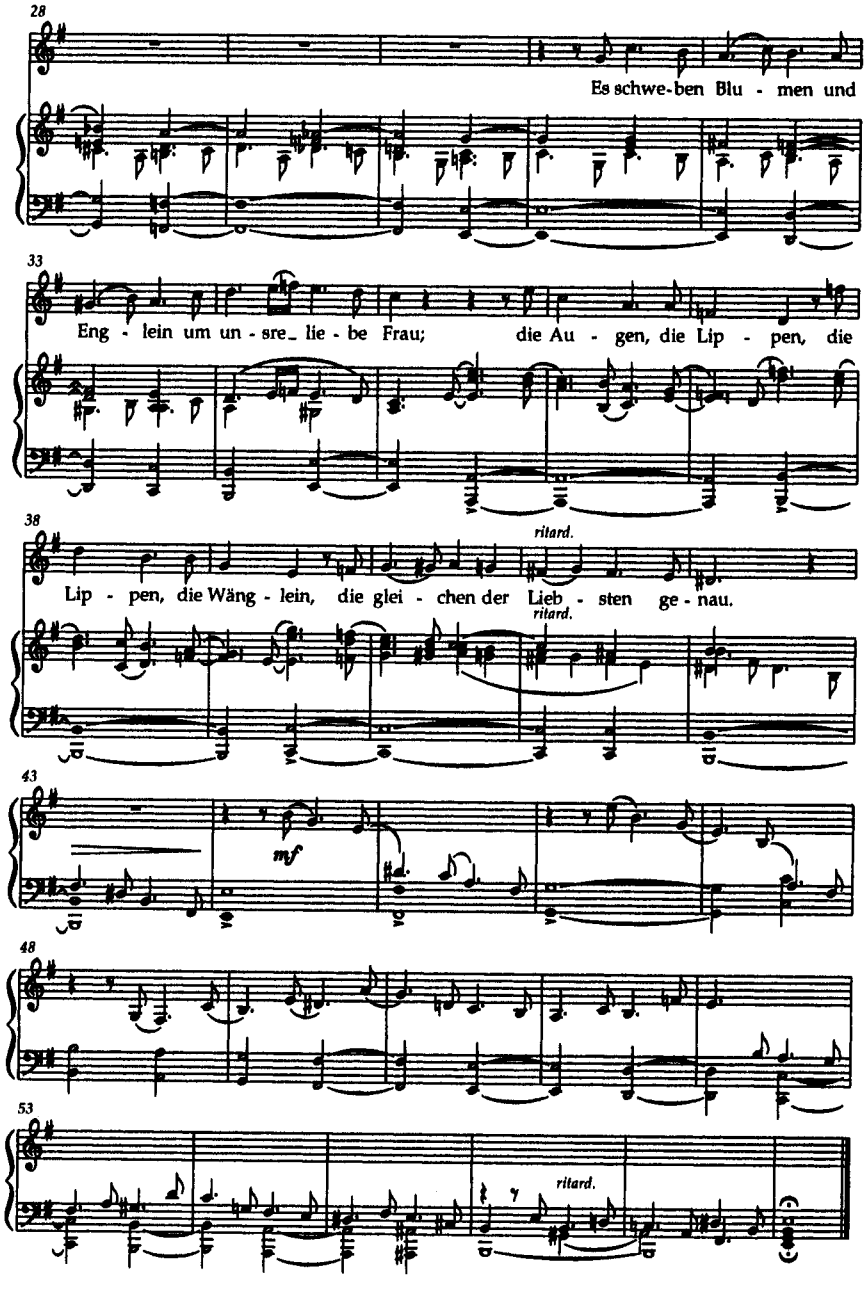
21. Schumann provided additional emphasis on the holy by changing the end of Heine's first line from "im schönen Strome" to "im heiligen Strome."

EXAMPLE 6.4 Robert Schumann, "Im Rhein," from *Dichterliebe*, Op. 48




*(continued)*

EXAMPLE 6.4 *(continued)*

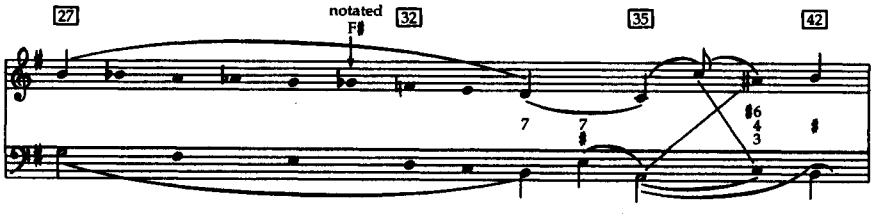


left hand of the piano in octaves) makes a stepwise ascent from E4 to G4 in m. 4, answered by a stepwise descent from C5 back to E4. At this point, voice and piano part ways. While the voice repeats its melodic material almost exactly (save for the crucial change of the penultimate F♯4 for F♮4), the piano continues the descent initiated in m. 4, sinking to the G1-G2 octave of mm. 11–13. In the second half of m. 13, the bass echoes the voice's descent from C to G of mm. 12–13 but, prompted by the important F♮4 in the voice, it arrives on a G1♯-G2♯ octave instead. This arrival coincides with the final E4 in the voice part and with the completion of the first stanza of the poem. The right hand of the piano fills out, if minimally, the bare texture created by the voice and the left hand; it does so with an accompaniment made up of cascading gestures in a halting pattern of alternating eighth- and dotted-quarter-notes.

The dominant $\frac{6}{5}$ of mm. 15–16, with which the A section concludes, leads to the A minor chord and *piano* of m. 17 that starts the B section. Here, the striding bass of the left hand disappears, replaced by a more fluid and gravitationally liberated adaptation of the eighth/dotted-quarter pattern of the right-hand accompaniment. In mm. 17–21, the voice also abandons its relentless descents and floats gently between D5 and G4; both voice and piano momentarily cadence on G major in m. 21. Although there is a brief struggle in mm. 22–25 as the voice counters the piano's descending figures with an ascent from A4 to E5, once the E5 is achieved the air clears, and the equanimity of the preceding phrase is restored, an effect confirmed by another cadence on G major in m. 27. There follows a brief piano interlude that begins a chromatically harmonized and rhythmically displaced descent through the sixth G2-B1 (the voice-leading of which is shown in ex. 6.5). The voice joins the piano in m. 31, and the completion of the sixth-descent in m. 34 is followed by a cadence on A minor in the following measure. An A minor chord, prolonged through a chromatically inflected voice-exchange (as shown in ex. 6.5), prepares the way for the return to E minor effected by the arrival on the B chord of mm. 42–43.

The return to E minor in m. 44 also marks a return to the material of the A section, which (other than the *mezzo forte* dynamic) is exactly the same as in its first appearance until the left hand of the piano reaches the F♯2-F♯3 octave in m. 49 (as it did in m. 6 of the A section). At this point, the rhythmic displacements of mm. 27–41 reappear, and the bass continues, now in augmentation, to the A1-A2 octave of m. 54. The A1-A2 octave then leads, through A♯, to the dominant of mm. 56–57 and the final cadence.

EXAMPLE 6.5 Linear analysis of mm. 27–42 of Schumann's "Im Rhein"



Schumann's music thus enacts a threefold process. The first part of this process (in the A section of mm. 1–16) establishes an austere, weighty E minor. It is well to note that the tonality is established with minimal reference to the sort of harmonic progressions and disjunct bass motions that are typical of tonal music of the period: there is no cadence on E minor in the entire section, and, moreover, the dominant of E minor is completely absent, its place taken by the leading-tone seventh-chord.[22] The second part of this process (in the B section of mm. 17–43) moves completely, if only temporarily, away from E minor. Here there are definitive cadences, but on G major and A minor. The large-scale instability of this section (a product of its avoidance of the tonal center so unequivocally established in the A section) becomes apparent only gradually. Ultimately, E minor cannot be avoided forever, and the section concludes (in mm. 42–43) with the very first appearance of the dominant of E minor in the entire song. The third part of the process (in the A' section of mm. 44–58), initiated by the definitive return to E minor, reprises the opening music. However, the music is now colored by rhythmic displacements that recall the latter half of the B section and that drive inexorably forward to a final cadence in the lowest practicable register.

This music, together with Heine's text, sets up the CIN diagrammed in figure 6.3. The generic space for the CIN is structured around the notion of exterior and interior domains. In the textual space, the exterior domain is that summoned in the first stanza of the poem: the city of Cologne, with its river and cathedral. The interior domain is that summoned in the second two stanzas: the cathedral and the painting with which the narrator is obsessed. In the musical space, the exterior domain is that of E minor, with its stark octaves and striding phrase rhythm. The interior domain is that of the contrasting (or "not-E-minor") music of the B section, with its more flowing accompanimental pattern, softer dynamic, rhythmic displacements, and large-scale instability.

Correlations between these idealized representations of exterior and interior spaces provide an opportunity for the combination of concepts in the blended space, which draws on elements of the input spaces to create a rich portrayal of the actions of the narrator and the environment within which they take place.

COMPOSITION Composition combines the exterior view of the river, city, and cathedral with the austere E minor music of the A section and suggests a sweeping and somewhat foreboding perspective, surveyed with a deliberative and relentless gaze. In turn, the interior view of the cathedral and the painting within combines with the soothing music of the B section to create a sense of refuge, a surcease from the harsh realities without.

22. It is interesting to speculate on whether Schumann himself would have differentiated between the dominant and the leading-tone chord. The distinction is of more concern to theorists of the period than to composers. Nonetheless, Schumann would most certainly recognize the sharper dissonance created by the leading-tone chord (with or without an added seventh), and it is worth noting that in the E minor portions of "Im Rhein" he favored these over the relative consonance of the dominant chord until m. 42.

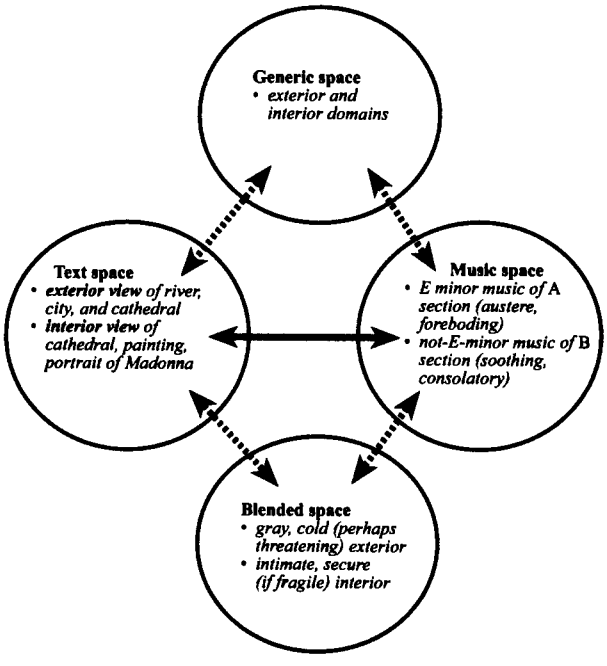



FIGURE 6.3 CIN for Schumann's "Im Rhein"

COMPLETION Completing the image, we can imagine details of each domain: an exterior world gray and cold and in some measure threatening; an interior world intimate and secure, if somewhat fragile. We can also imagine something of the physical and emotional life of the narrator. Outside the cathedral, he is almost certainly self-absorbed, a disheveled figure hurrying to escape the demons that torment him. Inside, he stands rapt, even as the tumult in his mind gradually increases to the point of becoming unbearable.

ELABORATION Our elaboration of the conceptual blend produced by Schumann's setting of Heine's poem might well seek to develop a narrative that could explain the forces that drive the narrator into the cathedral: a personality or perspective that makes life in the outside world intolerable, or a deep obsession with the beloved that renders her both an object of worship and something to be owned and controlled. However, we could also follow the lead of Schumann's music, which, in response to the denouement of the poem, returns to the E minor music of the opening. We are thus returned to the world outside the cathedral and there can see the narrator, now more disturbed than ever, staggering under the weight of his obsession with his lover, just as the music staggers with the rhythmic displacements of mm. 49–55. This furtherance of the narrative by the music suggests that it is the psychic torment of the narrator that is behind the sense of foreboding summoned by the exterior domain, as well as for the irreality of the interior domain. It is an oppressive torment that absorbs all, and among whose fires the narrator

wanders, knowing no rest. The elaboration of the blend, then, suggests a profoundly troubled mind to whom the Rhein represents more than a mirror for a great and holy place: it also offers the means to extinguish the fires that consume him.

### Summary

The dark twists and turns of love explored in Heine's *Lyriches Intermezzo* are, in their way, not so very far removed from those Müller captured in his *Die schöne Müllerin*. However, where Müller develops the topic and its ramifications over the course of twenty-three poems, it is Heine's gift to be able to accomplish this in the space of just one. This suggests a mastery of psychological detail—both in understanding his subjects and in what can be asked of his reader—that is altogether impressive. Equally impressive is Schumann's ability to exploit this level of detail by combining Heine's words with music that supports, expands on, and even offers a completion for the story told by the poem. It might be argued that Schumann's setting strips some of the ambiguities from Heine's poem, ambiguities that offer even more possibilities for interpretation than those provided by the song. The point is well taken, for the conceptual blend set up by the song emphasizes certain aspects of the topography of the text space at the expense of others. And yet the gains are considerable, for the conceptual world that emerges from syntactic invariances between Heine's text and Schumann's music is one rich in detail and striking in its immediacy. The challenges and rewards of creating such worlds no doubt fueled the burst of creativity that led to Schumann's amazing output of songs in 1840 and that produced a transformation in the way composers thought about how to organize their musical materials.

## "IN DER FREMDE"

The similarities between Klein's and Schubert's settings of "Trockne Blumen" were striking in that Schubert apparently did not know Klein's setting, and Schubert was occupied with developing a song cycle as opposed to creating a single Lied: not only are the musical details of each setting similar, but the generic spaces behind the conceptual blends that each song sets up are almost exactly the same. In Schumann's and Brahms's settings of Joseph von Eichendorff's "In der Fremde," nearly the exact opposite set of conditions holds. Brahms knew and respected Schumann's setting, but he seems to have wanted to show his friend and mentor what he could do with the same material. The result is two songs every bit as different as those of Klein and Schubert, but whose divergences stem from the generic spaces that structure the blends produced by each song.

### Eichendorff's "In der Fremde"

"In der Fremde," the text and translation of which are given here, first appeared as the text of an untitled song introduced in the course of Eichendorff's 1833 novella *Viel Lärmen um nichts*.[23] The scene in the novella is a huge and gloomy hall within

23. Joseph Freiherr von Eichendorff, *Viel Lärmen um nichts*, ed. Gerhart Baumann and Siegfried Grosse, Neue Gesamtausgabe der Werke und Schriften (Stuttgart: J. G. Cotta'sche Buchhandlung Nachf.,

a castle; outside, a storm rages, and, tucked away in a window niche, a young woman sings to her own accompaniment on the guitar. Her song is a deeply sad one, concerned with loneliness and thoughts of a solitary death far from home. The song over, however, she springs up, full of life, to meet her beloved and his friends, who have chanced upon her musings. The scene within which this poem first occurs is thus one of sharp contrasts—between the cold and forbidding interior of the castle, which makes for scant refuge from the storm outside, and the blushing and lively young woman who rushes to embrace her friends.

| IN DER FREMDE | IN FOREIGN LANDS |
|---|---|
| Aus der Heimat hinter den Blitzen rot | From home, behind the red lightning |
| Da kommen die Wolken her, | There come clouds, |
| Aber Vater und Mutter sind lange tot, | But Father and Mother are long dead; |
| Es kennt mich dort keiner mehr. | No one there knows me anymore. |
| Wie bald, wie bald kommt die stille Zeit, | How soon, how soon will that quiet time come, |
| Da ruhe ich auch, und über mir | When I too shall rest, and over me |
| Rauscht die schöne Waldeinsamkeit, | Shall rustle the forest's beautiful solitude, |
| Und keiner mehr kennt mich auch hier. | And no one here shall know me anymore. |

The poem itself offers images and impressions rather than any specific narrative structure. Dominant are the images of nature (in the form of the red lightning, approaching clouds, and promise of the forest's loneliness). Of equal importance is the narrator's profound sense of abandonment (her father and mother are long dead, and she is unknown both in her homeland and in her final place of rest). However, there are also within the poem sharp contrasts similar to those that accompanied its introduction in the novella: the comforting images of the homeland, parents, and sylvan setting are juxtaposed with the ambiguous "here" of the poem, solitude of the forest, and oppressive anonymity of the present surroundings. Nonetheless, connections between the stories told by the poem and the novella were obscured, and some of the ambiguities of the poem were muted, when it was placed by Eichendorff's friend Adolf Schöll in the "Totenopfer" section of an 1837 volume of Eichendorff's collected poetry, at which time it acquired its present title.[24]

## Schumann's "In der Fremde"

It was in Schöll's 1837 edition of Eichendorff's poetry that Schumann (or his fiancée, Clara Wieck) came upon the poem, which he began to set to music on 4 May 1840; the score is given in example 6.6.[25] Although there are indications Schu-

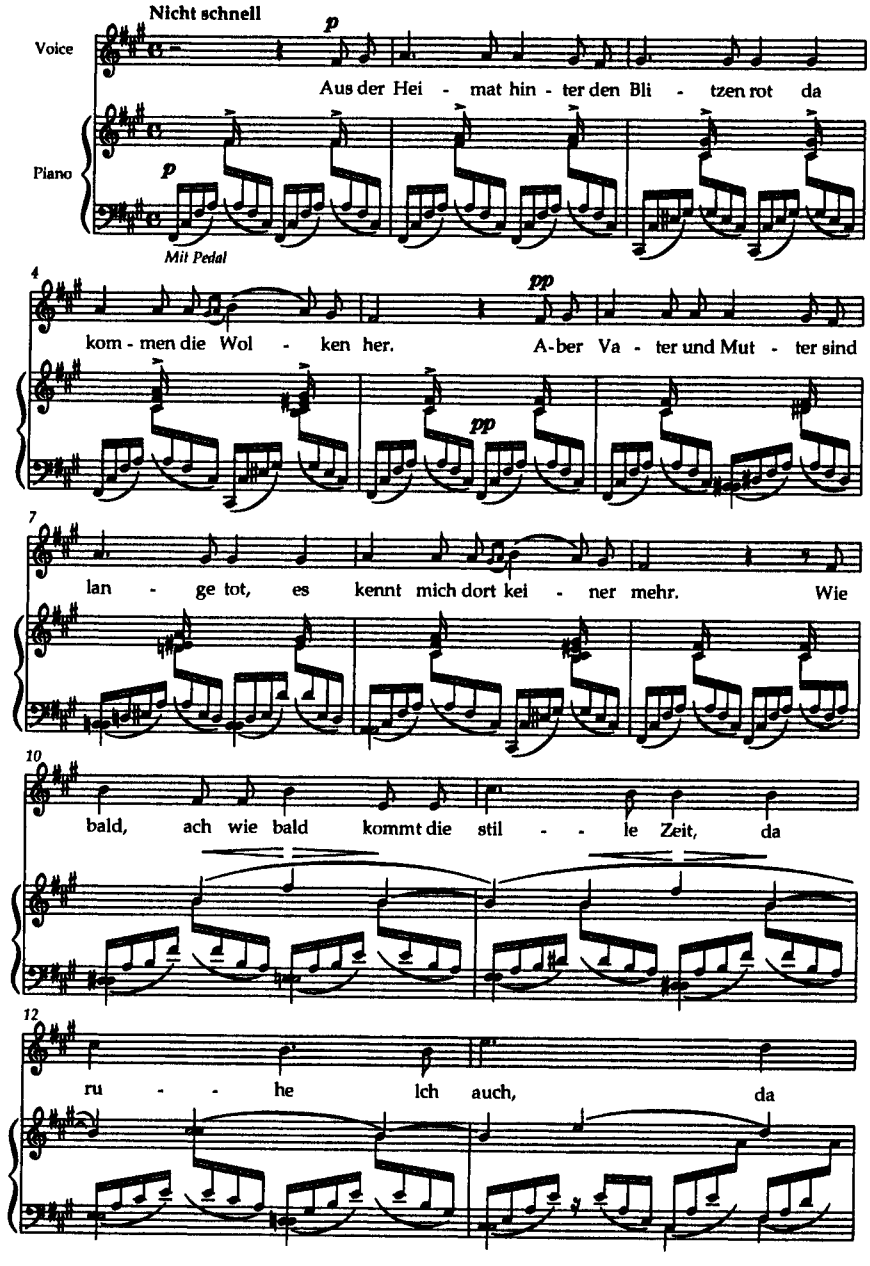
---

1957), 478. The text provided here is the same as that in the novella; some of Schumann's departures from Eichendorff's poem are discussed in the next section.

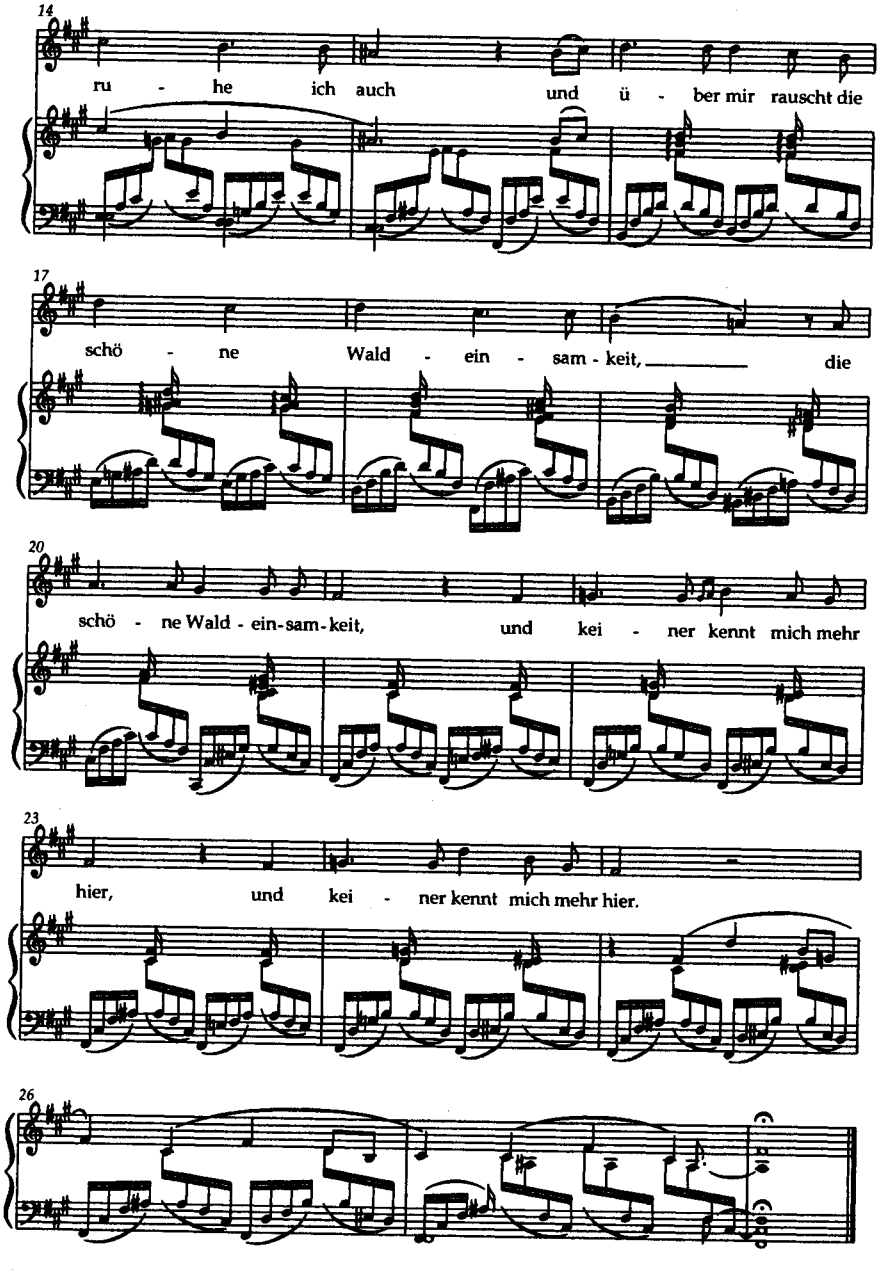
24. For a discussion of Schöll's role as editor of Eichendorff's work, see Ansgar Hillach and Klaus-Dieter Krabiel, *Eichendorff-Kommentar* (Munich: Winkler, 1971), 1: 47–48.

25. For a stimulating discussion of the songs from Schumann's Op. 39, their ordering in the cycle, and their relationship to the 1837 edition of Eichendorff's poetry, see Patrick McCreless, "Song Order in the Song Cycle: Schumann's *Liederkreis*, Op. 39," *Music Analysis* 5 (1986): 5–40. Dates of the composition of the songs are given in Robert Schumann's *Liederbücher*, which are reproduced in Viktor Ernst Wolff, *Lieder Robert Schumanns in ersten und späteren Fassungen* (Berlin: H. S. Hermann, 1913), 12–19. The same

EXAMPLE 6.6 Robert Schumann, "In der Fremde," from *Liederkreis*, Op. 39 No. 1



EXAMPLE 6.6 *(continued)*



mann intended "In der Fremde" to be the opening song of his Op. 39 *Liederkreis*, when the cycle was first published in 1842 by Tobias Haslinger, "Der frohe Wandersmann" (later published as Schumann's Op. 77 No. 1) appeared in its stead.[26] It was only with the second edition of the Eichendorff cycle, published in 1849 by Friedrich Whistling, that "In der Fremde" made its appearance as the opening song. In revisions to the song, Schumann changed Eichendorff's poem slightly to improve his melody, adding the "ach" in m. 10 and changing the last line of the poem from "Und keiner mehr kennt mich auch hier" to "Und keiner kennt mich mehr hier."[27]

Schumann's setting begins with a flowing accompaniment that continues throughout the song. Although there is no evidence that Schumann referred to Eichendorff's novella in the course of setting this poem, some characteristics of the music seem to fit the image portrayed there of a woman singing to her own accompaniment on the guitar. The running arpeggios in a relatively low register are typical of guitar music, and the melody is folk-like in its simplicity. The first five measures present a thoroughly straightforward F♯ minor, using only tonic and dominant. When the melody is repeated in mm. 6–9, the harmonization becomes a bit richer with the fully diminished seventh chords of mm. 6 and 7, but the basic harmonic framework established in the opening measures remains intact. Beginning in m. 10, however, there is a shift toward A major, and the sound brightens considerably as the higher register of the piano is called into service for the ringing B4-F♯4 fifth of mm. 10–11. A proper arrival on the $I^6$ of A in m. 13 is thwarted by the melody's temporary escape to E5, and when the outer-voice sixths are retraced in mm. 14–15 (as shown in ex. 6.7), the harmony at the point of arrival is changed to the dominant of B minor (that is, F♯ major) rather than the tonic of A.[28] This not only

---

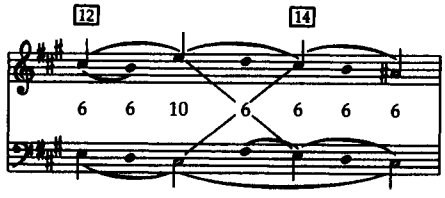
sources are correlated with the manuscripts at the Deutsche Staatsbibliothek Berlin in Reinhold Brinkmann, *Schumann und Eichendorff: Studien zum Liederkreis Opus 39*, Musik-Konzepte, 95 (Munich: Edition Text + Kritik, 1997), 80–86. See also Herwig Knaus, *Musiksprache und Werkstruktur in Robert Schumanns "Liederkreis": Mit dem Faksimile des Autographs*, Schriften zur Musik, 27 (Munich: Emil Katzbichler, 1974).

26. "Der frohe Wandersmann" and why Schumann may have included it in the 1842 edition of the *Liederkreis* are discussed in Jon W. Finson, "The Intentional Tourist: Romantic Irony in the Eichendorff *Liederkreis* of Robert Schumann," in *Schumann and His World*, ed. R. Larry Todd (Princeton, N.J.: Princeton University Press, 1994), 156–70.

27. A comparison of the two versions is given in Wolff, *Lieder Robert Schumanns*, 56–57; the revisions are not given in the facsimile provided in Knaus, *Musiksprache und Werkstruktur in Robert Schumanns "Liederkreis."*

28. In my analysis, I have emphasized the sense of departure evoked by mm. 10–14—that is, the firm motion away from minor and the implied cadence on A major. In a compelling analysis of "In der Fremde," David Ferris discusses Rufus Hallmark's reconstruction of one of Schumann's earlier versions of the song; in this version Schumann does indeed cadence on A major. See Ferris, *Schumann's Eichendorff* Liederkreis *and the Genre of the Romantic Cycle* (New York: Oxford University Press, 2000), 111–13. Working from a strongly Schenkerian perspective, Ferris also makes a convincing argument that, in terms of large-scale tonal structure, the move to A major in the final version of the song is, in fact, illusory and is subordinate to an expansion of the subdominant of F♯ minor (pp. 98–106). Although Ferris's reading makes sense of Schumann's overall harmonic and contrapuntal plan, it does not preserve the sense that with the arrival on F♯ major the promised cadence on A major has been subverted. It is this sense of subversion that my analysis, focused as it is on a more surface level of harmonic and contrapuntal activity, attempts to capture.

EXAMPLE 6.7 Linear analysis of mm. 12–15 of Schumann's "In der Fremde"




returns the music to minor but also starts the return to F♯, accomplished in the latter half of m. 19 when a B minor chord is transformed into the leading-tone chord of C♯. The cadence on F♯ minor in m. 21 is followed by eight measures with an F♯ pedal in the bass, over which there are double chromatic neighbors to F♯ (G♮ and E♯), as well as a major third (A♯).

This music, together with Eichendorff's text, sets up the CIN diagrammed in figure 6.4. The generic space, like that for Klein's and Schubert's settings of "Trockne Blumen," is structured around the notion of two contrasting ontological states. Lacking, however, is the notion that there is a smooth or ready connection between the two states: although there is a sense that they are linked in some larger scheme, they are portrayed as basically incommensurate.

Contrast between ontological states in Eichendorff's poem is supported by a marvelously subtle shift of conceptual metaphor. When reference is made to the past or the present, a spatial mapping is used: the past is "there" (the "dort" of the fourth line), the present is "here" (the "hier" of the eighth line). When reference is made to the future, an affective mapping is used: the future is "quiet" (the "stille" of the fifth line) and partakes of "forest-solitude" (the "Waldeinsamkeit" of the seventh line). Each of these mappings is further associated with somewhat more diffuse emotions. The spatial mappings of the past and present are connected with the poignant memory of departed parents and the sense of abandonment that comes with disappearing from memory, with the confused or uncomfortable feelings summoned by the red lightning and approaching clouds, and with the loneliness of an anonymous existence. The affective mappings of the future are connected with a sense of rest and of being in the presence of beauty.

Contrast between ontological states can also be seen in Schumann's music. The F♯ minor music is, for the most part, resolutely secure, yet it is, nonetheless, tinged with chromaticism. Thus, in mm. 1–9, regular cadences appear, but also the fully diminished seventh chords of mm. 6 and 7. Similarly, mm. 20–28 offer the security of the pedal point, as well as the double chromatic neighbors to F♯, which render ambiguous the assertion of major in the final measures. The music of mm. 10–19 tends to be brighter (note the higher register in the piano accompaniment that is especially audible in the right hand of mm. 10–15) and less involved with chromaticism (with the important exceptions of the melodic deflection to A♯ in m. 15 and the fully diminished seventh chord of m. 17). However, in this same passage, convincing cadences are conspicuous only for their absence.

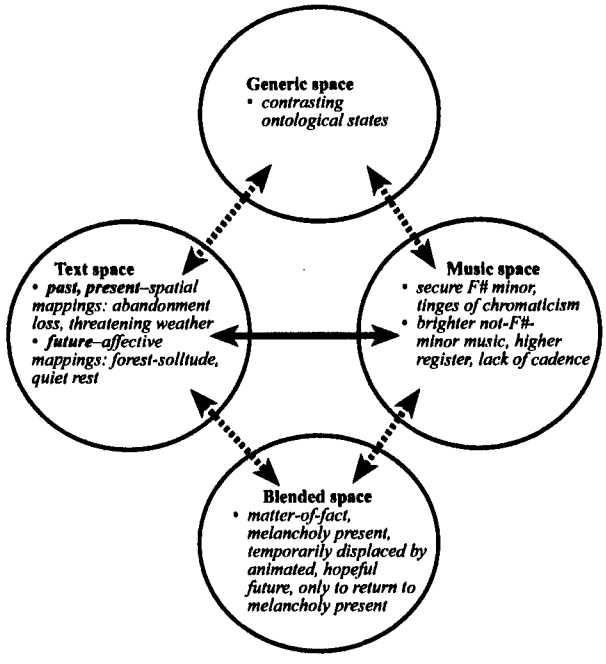



FIGURE 6.4 CIN for Schumann's "In der Fremde"

Correlations between these idealized representations of contrasting ontological states provide an opportunity for the combination of concepts in the blended space. The combination draws on elements of the input spaces to create a portrayal of two worlds, not totally unlike those of "Im Rhein," but framed in emotional rather than architectural terms.

COMPOSITION Composition combines the F♯ minor/major music with the spatially mapped past and present. Evidence that Schumann was aware of the importance of maintaining a consistent musical environment for the first four lines of the poem is provided by his manuscript, for after writing a more varied accompaniment for mm. 6–10 he scratched it out and continued the pattern established in mm. 1–5.[29] The affectively mapped future is first combined with A major and then, with the voice still near the top of its range for the song, B minor. With the repetition of "die schöne Waldeinsamkeit," however, the melody drops, and we return to F♯ and the spatially mapped present.

COMPLETION Following the structure of the blend and our knowledge that a single first-person voice is being projected, we could imagine a singer who begins with almost matter-of-fact observations but who gradually becomes more engaged

29. See the unpaginated "Faksimile des Autographs" appended to Knaus, *Musiksprache und Werkstruktur in Robert Schumanns "Liederkreis."*

and more melancholy. This melancholy is then rather abruptly displaced (in m. 10) by more animated and hopeful emotions, leading the singer to a higher register and to sing with a fuller voice (in keeping with the profusion of longer note values in the melody). The hopeful attitude cannot be sustained, however, and solitude once again becomes loneliness, the present that is reasserted in m. 21 effectively vanquishing the bright promise of the future.

ELABORATION The imaginary world of Schumann's song is one quite different from that of Eichendorff's original poem. The song sung by the young woman in the window niche is almost folk-like, its regular meter and stylized emotions disturbed only once, by the metrical awkwardness of the sixth line. But this momentary shadow of emotional depth passes, and it is possible to imagine the singer jumping up, the guitar and song flung aside, to greet her friends. Schumann's singer will not be so easily released. The almost formulaic melancholy of the song's opening reveals itself in the end to be inexorable, pulling the energy and hope from the singer and leaving her hollow and forlorn. Schumann's song leaves us with a singer for whom eternity now stretches out as flat and colorless as her present, her spirit struggling as the walls of anonymity close around it.

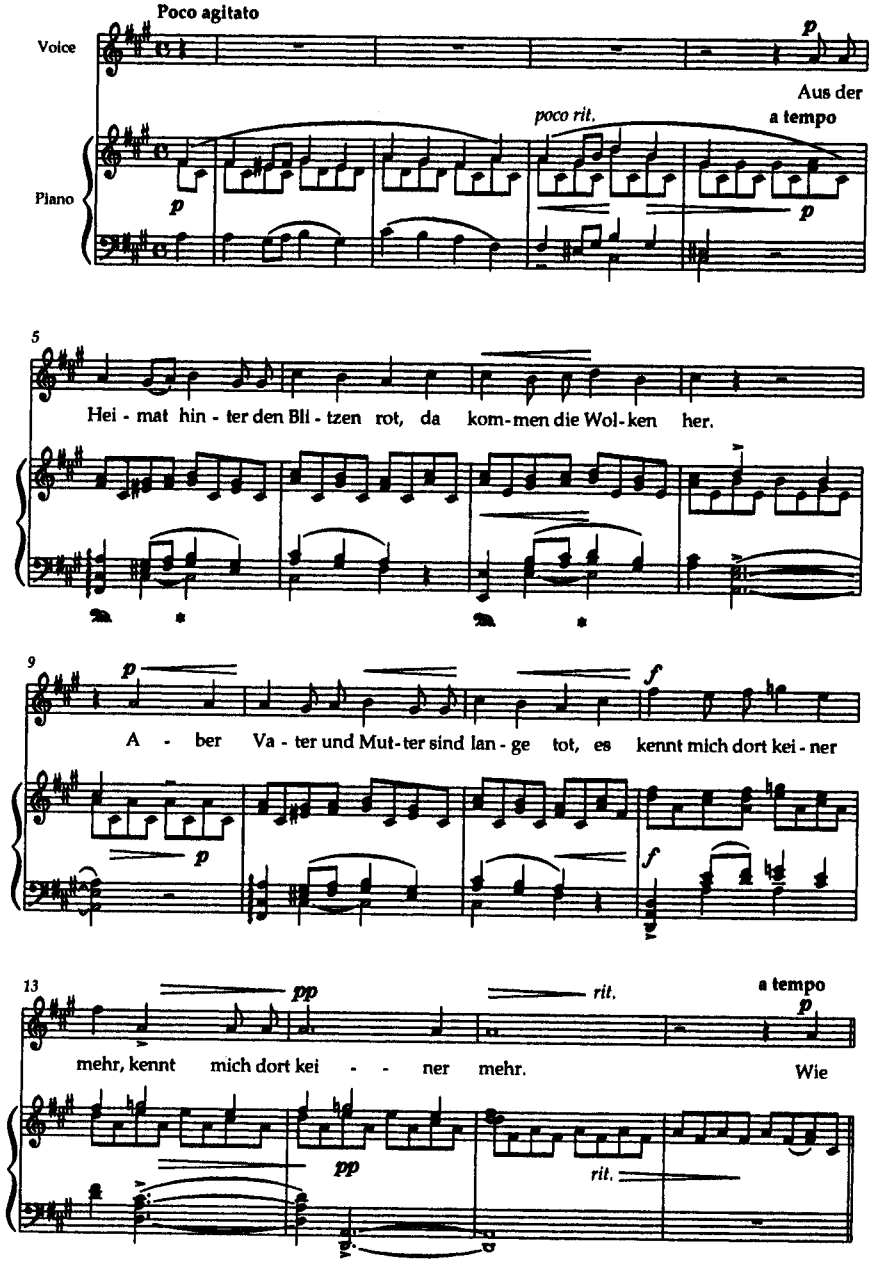
## Brahms's "In der Fremde"

Johannes Brahms was but nineteen when, in November of 1852, he made his setting of "In der Fremde," published the next year by Breitkopf and Härtel as the fifth of his *Sechs Gesänge*, Op. 3; the score is given in example 6.8.[30] "In der Fremde" was the first of the Op. 3 songs that he composed, and it shows clear debts to Schumann: Brahms adopted Schumann's version of the text, the key of F♯ minor, and some important motivic details. Nonetheless, the song is clearly Brahms's own, and nowhere is this more apparent than in the conceptual blend it sets up.

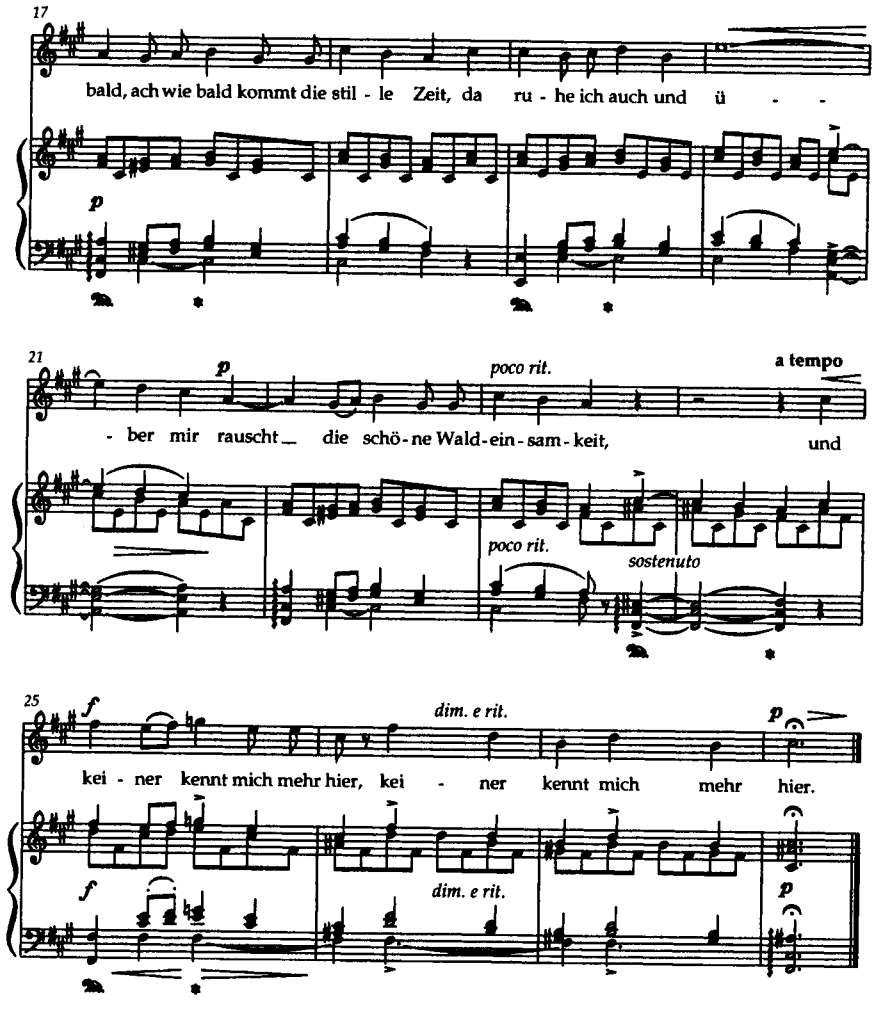
After a short introduction that anticipates the vocal melody without directly stating it (by presenting only elements of the piano accompaniment for the melody), the song proper commences, laid out in four phrases. These phrases have almost identical beginnings, both in harmony and melody, although each ends rather differently. The first touches back upon the tonic at its midpoint (m. 6) and then deflects away to A major. Arrival on A (m. 8) is unsettled by the entrance of the dominant on the second beat (over an A pedal), which supports an echo of the last three notes of the voice part. The second phrase moves at its midpoint to D major (m. 12); its cadence (on the downbeat of m. 13) is as unsettled rhythmically as that of mm. 7–8, but it gains a bit of stability when the fourth line of the poem is repeated (mm. 13–15). Deflection toward A major occurs again in the third phrase (m. 19), but the completion of the cadence is delayed by two beats and the vocal line continues, ultimately eliding with the beginning of the fourth phrase in m. 22. The fourth phrase breaks off momentarily at its midpoint (which, as in all the

30. Brief discussions of the song can be found in Lucien Stark, *A Guide to the Solo Songs of Johannes Brahms* (Bloomington: Indiana University Press, 1995), 15–16; and Eric Sams, *The Songs of Johannes Brahms* (New Haven, Conn.: Yale University Press, 2000), 36–37.

EXAMPLE 6.8 Johannes Brahms, "In der Fremde," from *Sechs Gesänge*, Op. 3 No. 5



EXAMPLE 6.8 *(continued)*




preceding phrases, is marked by an arrival on the tonic), and the mode suddenly changes to major on the last beat of m. 23. Brahms then completes the phrase by recalling the melody of m. 12, but this time harmonized in F♯ rather than D. Within this context, the G♮ in the melody at m. 25 recalls the end of Schumann's setting, an association made stronger when the repetition of the last line of the poem in (mm. 26–28) is set to a melody similar to that which appeared in m. 26 of Schumann's piano accompaniment.

Besides the obvious melodic repetitions engendered by this overall compositional scheme, two additional features bear note. First, the compass of the melody is relatively restricted (from G♯4 to G5, the smallest range of any of the Op. 3 songs). Second, the melody is constantly brought back to A4: each phrase begins on A4, it

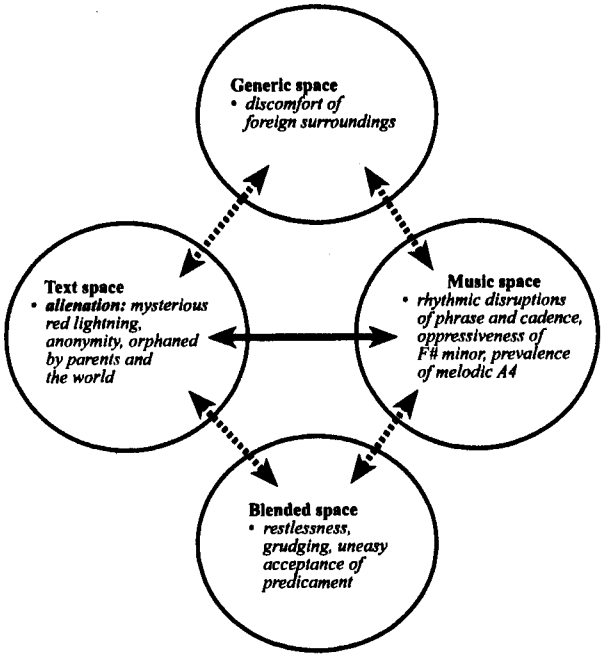



FIGURE 6.5 CIN for Brahms's "In der Fremde"

marks the arrival on the midpoint of each phrase, and the second and third phrases end on it (in the latter case, this brings about the elision of the end of the third phrase with the beginning of the fourth).

This music, together with Eichendorff's text, sets up the CIN diagrammed in figure 6.5. The generic space is structured around the feeling of discomfort attendant on being in alien surroundings. This space is rather unlike those for the other CINs discussed in this chapter in that it does not lend itself as readily to a narrative framework. However, discomfort, by its nature, suggests a dynamic situation (to the extent that we want to return to a state of comfort). This may not quite constitute a narrative in its own right, but it creates an opportunity at least for a minimal story.

Eichendorff's poem supports this sense of discomfort with the ominous portents of the clouds and red lightning, as well as with the isolation and anonymity noted in the fourth and eighth lines. Discomfort is also summoned, if negatively, through the thought that rest is to be found elsewhere.

The discomfort of alien surroundings is suggested musically by the repeated attempts in mm. 4–21 to escape the pitch A4 and the key of F♯ minor. (The accompaniment pattern for the F♯ minor sections also makes use of an insistent C♯4 that starts as a simple bit of rhythmic activation in the introduction, but soon begins to cloy.) Each attempt fails, of course, as it is undercut by insecure cadences and the relentless pull of A4. When a sort of escape is fashioned, at the song's conclusion, one senses that the forces of oppression have been held in abeyance more than defeated. Although the mode changes to major, the G♮5 summons the ♭2̂ that is

more typical of minor; although the melody ends away from A4, it has not accomplished a structural close on F♯4.[31]

Correlations between these idealized representations of discomfort at alien surroundings provide an opportunity for the combination of concepts in the blended space that draws on the elements of the input spaces to create a portrayal of life "in der Fremde."

COMPOSITION Composition combines the images of isolation and the alien in the first four lines of the poem with two successive musical phrases that begin in the same place but move away to end at a distance from the tonic and with the downbeat destabilized. The combination of the last line of the poem and the concluding music is slightly more tenuous: the return of the voice to the high end of its range and the recall of the end of phrase 2 reinforce the portrayal of discomfort, but the sonorous F♯ major chords and pedal point also suggest a sort of reconciliation outside of the basic mapping that structures the CIN.[32]

Although I have not emphasized the failure of elements from the text-space and music-space to combine in my discussions of the process of blending, it is particularly noticeable in the case of lines 5–7 of Eichendorff's poem and mm. 17–23 of Brahms's setting. At this point, the topography of the two spaces is not uniform—more specifically, the text at this point (concerned as it is with contemplation of peaceful rest in the embrace of the forest) does not conform to the structural premises of the generic space. Although words and music certainly co-occur here, they do not combine from the perspective of the CIN diagrammed in figure 6.6. This lack of correspondence may be one reason Brahms's song seems weak here: his text-painting of "über" (as a long, relatively high note that is manifestly "over" the accompaniment) offers rather insufficient compensation for the failure of words and music to work together on a larger scale.

COMPLETION It is through completion of the conceptual blend that Brahms's song takes on a dynamism that extends beyond the rhythmic disruptions and key-area juxtapositions of the music. As noted, we know that states of discomfort cry out for resolution, whether that is provided by shifting position in our chair or moving away from someone who makes us uneasy. When the music of the first half of each phrase is followed by the contrasting music in the second half, it is easy to hear this as a moving away, a physical adjustment to unpleasant circumstances. The sudden shift to major and return of the higher vocal tessitura at the end of the song suggest not so much a final resolution as an uneasy truce: the struggle of the first twenty-odd measures has yielded results, but they can hardly be called definitive.

31. From a Schenkerian perspective, the lack of a clear descent to $\hat{1}$ (through the melodic progression A–G♯–F♯) poses interesting, although by no means irresolvable, problems.

32. It is worth noting that the two lines of the poem most clearly associated with the discomfort of abandonment and alienation—lines 4 and 8—are also the only two that Brahms chose to repeat, thus providing an opportunity to emphasize the composition of elements from the text-space and music-space.

ELABORATION Entering further into the world of Brahms's song, we can see that Brahms's singer is thoroughly within the foreign lands as she sings, and, although her deflections toward major keys suggest an attempt to make herself comfortable in these alien surroundings, she remains an anonymous stranger—she cannot escape the iron grip of F♯ minor, just as she cannot escape her isolation. Nonetheless, Brahms's singer does not despair—at least not yet. Instead, she struggles against her fate, crying out against her anonymity in her highest register, attempting to use her predicament as a lever to gain recognition and a life in this foreign land. Given this text and this music, we can imagine her shaking her fist at fate; if her eyes fill with tears, they are tears of a struggle against destiny.

### Summary

With his setting of "In der Fremde" Brahms clearly wanted to demonstrate his independence as a composer, as well as to pay homage to an important friend and mentor. Yet, two other factors are important. The first is raw youth: Brahms at nineteen was no stranger to feelings of alienation and restless energy, and it is hardly surprising that he drew these from Eichendorff's poem. The second are the compositional advances made by Schumann in the *Liederjahr* of 1840: the year bore witness not only to an incredible outpouring of song but also to a new understanding of how musical syntax could be shaped to respond to the possibilities of poetry. Brahms learned the lessons of Schumann's Liederjahr, as well as the many other lessons taught by the music of his century, and he built on them. Where cadence was still a primary articulator of the musical argument for the Schumann of 1840, it became something infinitely more malleable for the Brahms of 1852. Indeed, over the course of Brahms's "In der Fremde," there is not one clear cadence; instead, there are cadences disrupted, so understated as to be almost invisible (as the cadence of m. 23), or assembled out of the bits and pieces of traditional cadential formulae (which is how Brahms brings the song to a conclusion). Brahms, of course, knew when and how to exploit traditional cadential formulae, and he did so with great success in the years that followed, but his setting of "In der Fremde" shows that, when his understanding of a text demanded it, he could create a new set of syntactic possibilities for musical discourse.

## CONCEPTUAL BLENDING AND SONG

As part of his anthropological fantasy on the origins of language, Jean-Jacques Rousseau imagined the primordial world that gave birth to the languages of the temperate south. Around the water wells necessary for life in these arid lands people gathered, and young men met young women for the first time.

> Beneath aged oaks, conquerors of years, an ardent youth gradually forgot its ferocity, gradually they tamed one another; through endeavoring to make themselves understood, they learned to explain themselves. There the first festivals took place, feet leaped with joy, eager gesture no longer sufficed, the voice accompanied it with passionate accents; mingled together, pleasure and desire made themselves felt at the same

> time. There, finally, was the true cradle of peoples, and from the pure crystal of the fountains came the first fires of love.[33]

As Rousseau would have it, the language that developed around these aqueous gathering places was not like modern language, but a fusion of words and music:

> Around the fountains of which I have spoken, the first discourses were the first songs; the periodic and measured recurrences of rhythm, the melodious inflections of accents caused poetry and music to be born along with language; or rather, all this was nothing but language itself in those happy climates and those happy times when the only pressing needs that required another's help were those to which the heart gave rise.[34]

In the myth Rousseau offers, music and language don't simply cohabitate; they commingle their syntaxes: the periodic and measured recurrences of rhythm and the inflections of accent are combined in an Ur-language, which, only later, after the Fall or its equivalent, separated into music and language.

In the end, Rousseau's fantasy has more to tell us about the role of language and music in the anthropological speculations of the eighteenth century than it does about the possibility that the two media were once one.[35] But the conviction that music represents a communicative medium to equal language—a conviction Rousseau shared with his old antagonist Rameau[36]—and the idea that the sum of music and language might be greater than its parts are concepts writers like Müller and von Spaun believed were demonstrated, in fairly specific ways, by the nineteenth-century Lied. While it is possible to dismiss such beliefs as products of overheated rhetoric, the fascination composers, audiences, and critics have had with song cannot be ignored. Conceptual blending offers a way to explain why this fascination has been so enduring: under certain circumstances, elements and relations from the mental space set up by the music of a song blend with elements and relations from the mental spaces set up by the text to create a new world for the imagination—a world of supernatural charm, nocturnal terror, and much, much more.

For conceptual blending to occur, the mental spaces that serve as inputs to the blend must have a shared topography. For the songs analyzed in this chapter I have located this shared topography in commonalities between the complete discourse

33. Jean-Jacques Rousseau, "Essay on the Origin of Languages (c. 1763)," in *Essay on the Origin of Languages and Writings Related to Music*, trans. and ed. John T. Scott, Collected Writings of Rousseau, 7 (Hanover, N.H.: University Press of New England, 1998), 314. In my comments on Rousseau's essay, I am more concerned with the idyllic state he imagines than with the overall context relative to which the image is framed. For a probing discussion of this context, see Lydia Goehr, *The Quest for Voice: Music, Politics, and the Limits of Philosophy* (Berkeley: University of California Press, 1998), 98–106.

34. Rousseau, "Essay on the Origin of Languages," 318.

35. For more on the role of language and music in speculations about human origins by French writers of the eighteenth century, see Downing A. Thomas, *Music and the Origins of Language: Theories from the French Enlightenment*, New Perspectives in Music History and Criticism (Cambridge: Cambridge University Press, 1995), chap. 2.

36. Although Rousseau and Rameau both argued that music was the equal to language, for Rousseau this meant melody and for Rameau it meant harmony. This difference of opinion (among other things) was the basis for their debate about musical expression, discussed at the end of chap. 3 in this volume.

structure set up by the music and the complete discourse structure set up by the text. That such commonalities might exist suggests deep correspondences between musical and linguistic syntax. To be sure, these correspondences are, in most instances, of a general sort, outlining as they do constructs like contrasting ontological states (as in the settings of "Trockne Blumen" and in Schumann's "In der Fremde"), exterior and interior domains (as in "Im Rhein"), or discomfort with foreign surroundings (as in Brahms's "In der Fremde"). And yet, such highly abstract constructs still have their power, for they can find realization in manifold concepts associated with both the emotions and the intellect.

Agawu ended his provocative 1992 essay with these words: "We await the development of a syntax of song."[37] The preceding analyses suggest that this syntax is to be found not "in" the song proper, but in the way the constituent media of song—words and music—relate to one another. The syntax of song, then, is a syntax of conceptual blending, through which music breathes life into poetry and poetry breathes life into music, and together they proceed as song.

37. Agawu, "Theory and Practice," 30.

CHAPTER SEVEN

# COMPETING MODELS OF MUSIC: THEORIES OF MUSICAL FORM AND HIERARCHY

Midway in her journey through Wonderland, Lewis Carroll's Alice takes leave of the Cheshire Cat and makes her way toward the house of the March Hare. There she finds a tea party under way—a mad, perpetual tea party, attended by the March Hare, the Mad Hatter, and the somnolent Dormouse. As the party proceeds, Alice's patience begins to wear thin, tried by turnabout word play and abrogated manners, and is stretched to the limit when she finds out that the Hatter's riddle—"Why is a raven like a writing desk?"—is without solution.

> Alice sighed wearily. "I think you might do something better with the time," she said, "than wasting it in asking riddles that have no answers."
>
> "If you knew Time as well as I do," answered the Hatter, "you wouldn't talk about wasting *it*. It's *him*."
>
> "I don't know what you mean," said Alice.
>
> "Of course you don't!" the Hatter said, tossing his head contemptuously. "I dare say you never even spoke to Time!"
>
> "Perhaps not," Alice cautiously replied; "but I know I have to beat time when I learn music."
>
> "Ah! That accounts for it," said the Hatter. "He wo'n't stand beating. Now, if you only kept on good terms with him, he'd do almost anything you liked with the clock."[1]

Exchanges like this are common enough in Wonderland, a place inhabited by creatures whose logic is often as fantastic as their appearance. Nonetheless, the game Carroll is playing at here is less a matter of labyrinthine logic and more a matter of different ways to structure what is putatively the same domain. Alice uses two different cross-domain mappings to reckon with time: time as something to be saved or wasted (mapping from the domain of commodities or substances onto the domain of time); and time as a set of regular motions (mapping from the domain of physiologically defined space onto the domain of time). For the Hatter's cockeyed contribution, Carroll exploits a traditional characterization of time as a human

1. Lewis Carroll, *The Annotated Alice: Alice's Adventures in Wonderland and Through the Looking Glass*, intro. and with notes by Martin Gardner, illustrated by John Tenniel (New York: W.W. Norton, 2000), 72. The tea party is the topic of chap. 7 of *Alice's Adventures in Wonderland*.

agent (as in Father Time, or the New Year as a diapered baby), an anthropomorphic mapping that exploits conceptual blends like those discussed in the latter portion of chapter 2. The merry confusion of mappings that results is part and parcel of the imaginary world Carroll constructed—indeed, Wonderland is a maze of interfiliated conceptual blends. However, the scene at the tea party also throws into relief a fundamental conflict between Alice's world and that of Wonderland, which comes to a head in the final chapter. There, Alice responds to the Queen of Hearts's peremptory "Off with her head!" ("Sentence first—verdict afterwards") by dismissing the entire court and courtroom with "Who cares for you? You're nothing but a pack of cards!" At this, Wonderland disappears, and Alice wakes to find herself lying on the bank of the river, with her head in the lap of her sister.[2]

A confusion of mappings similar to the one apparent in Alice's dialogue with the Hatter often arises in music theory. The result, while not quite as entertaining as a turn in Wonderland, can have some of the same flavor. Theorists appear to be talking about the same thing, but they structure their understanding of it in very different ways. Two examples are taken up in this chapter. The first involves musical form, the second musical hierarchy.

During the eighteenth century, the study of musical form emerged as an important basis for discourse about music. Form was regarded as dealing with the very essence of musical works: how musical materials were arranged and correlated over the course of a piece of music. As the study of form developed and was continued through the nineteenth century, theorists worked with two basic—and seemingly opposed—models of musical form, one static, the other dynamic. Musical form, viewed from a static perspective, is reminiscent of architecture (a parallel all the more ironic, given Friedrich von Schelling's characterization of architecture as "frozen music")[3] and typically consists of either a framing structure clad with musical material or relatively abstract containers filled with musical events. Musical form from a dynamic viewpoint is processive and a bit unpredictable: the musical work emerges over the course of time, and musical materials are both the substance of and raison d'être for this emergence. As a further complication, the cross-domain mappings that activated these two models changed during the nineteenth century. Consequently, there are two levels of mapping involved, both of which can breed confusion about just what is meant by musical form: a generic level associated with the two basic models (one interpreting form as static, the other as dynamic); and a specific level associated with the cross-domain mappings that activate the generic models. In the first section that follows, I explore static and dynamic models of musical form that emerged in the eighteenth century and the way these models were transformed by changes in ideology and music education that took place beginning late in that century.

2. Carroll, *The Annotated Alice*, 124.

3. "Architecture, as the music of the plastic arts, thus necessarily follows arithmetical relationships. Since it is music in space, however, in a sense, solidified music, these relationships are simultaneously geometric relationships." Friedrich Wilhelm Joseph von Schelling, *Philosophy of Art*, ed., trans., and intro. Douglas W. Stott, with a foreword by David Simpson, Theory and History of Literature, 58 (Minneapolis: University of Minnesota Press, 1989), 165. See also, on p. 177, the passage that begins "If architecture is indeed solidified music . . ."

The two models of musical hierarchy I deal with in the second of the two sections that follow are a somewhat different matter, for these models were first developed to apply to domains other than music. It was only in the early nineteenth century that they were applied to music, and then to two rather different aspects of musical organization. The first model views hierarchy as a matter of control: each level in a hierarchy (with the exception of those at the extremes) controls the next-lower level and is itself controlled by the next-higher level. This model stretches back at least to the Middle Ages and is most often used to account for tonal organization. The second model of hierarchy relies on a more componential approach: the elements of level A of the hierarchy combine to make up individual elements at the next-higher level (level B of the hierarchy); the elements of level B then combine to make up individual elements at the next higher level (level C of the hierarchy), and so on, until the account of structure is exhausted. This model emerged during the seventeenth century and is most often used to explain music's metrical structure.

When confronted by a maze of mapping like those I have just outlined, mappings that yield contradictory accounts of quite basic musical materials, our inclination may be to throw up our hands and shout the equivalent of "Who cares for you? You're nothing but a pack of cards!" But problems of musical understanding are rarely so easily dismissed. Instead, we must enter into the mad tea party that is music theory, where music theorists from different nations and different historical periods gather around the table to discuss and dispute musical concepts. If we are to make sense of the conversation there, we need to understand how cross-domain mapping shapes this discourse—indeed, as I suggested in chapter 2, how it makes music theory possible.

## TWO MODELS OF MUSICAL FORM

There was a certain enviable simplicity to the study of music in the fin de siècle Anglo-American musical world, due in no small part to its near-complete subjugation to the German musical world.[4] Music teachers in Anglo-American circles were either German or trained in Germany, the repertoire studied was almost entirely German, and the theoretical models—developed exclusively with respect to German musical practice—had been formulated by German music theorists of the nineteenth century. Thus the beginning Anglo-American music student of the late nineteenth and early twentieth century learned a tradition of musical study that was almost entirely German. This is not to say, however, that the tradition they assimilated was completely unified. Consider, for instance, the account of musical motives offered by Ebenezer Prout, a prominent music educator in England during the late nineteenth century. Having distinguished between phrases (which end with a cadence) and sections (subphrase units that do not necessarily end with a cadence), Prout makes one further subdivision:

> Every section contains at least two [metric] accents, and can be separated into smaller parts, containing only one accent each. Such parts are called MOTIVES, and correspond

4. By "German musical world," I mean the German-speaking countries of Europe.

> exactly to the *feet* in poetry. Here we have what (to borrow a scientific term) we may call the "protoplasm"—the germ out of which all music springs. A thorough knowledge of the nature of the motive is therefore indispensable to anyone who would understand the fundamental principles of musical form.[5]

Important for Prout's characterization of a motive are the dynamic context provided by metrical accentuation—he will later define motive as composed of a strongly accented note preceded by at least one unaccented note[6]—and the somewhat more abstract notion of a motive as the seed of a musical organism.

Rather different is the approach to motive taken some ten years later by Percy Goetschius, born in New Jersey but educated in Stuttgart, where he first taught. Goetschius situates the motive in an explicitly grammatical context:

> The smallest unit in musical composition is the single tone. The smallest cluster of successive tones (from two to four or five in number) that will convey a definite musical impression, as miniature musical idea, is called a Figure. Assuming the single tone to represent the same unit of expression as a letter of the alphabet, the melodic figure would be defined as the equivalent of a complete (small) word;—pursuing the comparison further, a series of figures constitutes the melodic motive, equivalent to the smallest group of words (a subject with its article and adjective, for example); and two or three motives make a Phrase, equivalent to the complete, though comparatively brief, sentence (subject, predicate, and object).[7]

In contrast with Prout's dynamic account of musical form, Goetschius relies on a highly schematized grammar of musical elements. So formalized is this grammar that it ultimately reduces to a matter of counting measures: "Melodic motives differ in length from one to four measures; by far the most common extent, however, is two measures, and the student will do wisely to accept this dimension and analyze accordingly, unless there is unmistakable evidence to the contrary."[8]

What we have in these excerpts, then, are two markedly different approaches to musical form, voiced through characterizations of motive.[9] Prout's is the more dynamic: music is likened to some sort of organism, and the motive is the seed from which it springs. Goetschius's approach, by contrast, is rather more static: music is but prose in tones, and the motive is simply the smallest intelligible syntactic unit.

5. Ebenezer Prout, *Musical Form* (London: Augener, 1893), 26.

6. Prout, *Musical Form*, 31.

7. Percy Goetschius, *Lessons in Music Form: A Manual of Analysis of All the Structural Factors and Designs Employed in Musical Composition* (Bryn Mawr, Penn.: Oliver Ditson, 1904), 19.

8. Goetschius, *Lessons in Music Form*, 23.

9. Although I have chosen to highlight the differences between Prout's and Goetschius's approaches to musical form, Robert Gjerdingen has reminded me that the two theorists knew each other's work and exchanged dedications of their texts. Moreover, not all theorists within the Anglo-American tradition agreed on the importance of motive. For instance, Stewart Macpherson, whose book on form supplanted that of Prout and who was mightily suspicious of overly thorough analyses, wrote, "As an example of a too mathematical conception of analysis may be instanced the stress laid by some theorists upon an arbitrary dissection of the music into certain metrical fragments called motives, a dissection of which usually conveys little or nothing to which the musical sense can respond." Macpherson, *Form in Music: With Special Reference to the Designs of Instrumental Music* (London: Joseph Williams, 1908), 39. A footnote makes clear that in referring to "some other theorists," Macpherson actually meant Prout.

However, two related factors complicate these seemingly neat constructions. First, neither Prout nor Goetschius eschews entirely the approach of the other. Prout is perfectly capable of rendering a crushingly static account of musical form, and Goetschius knows how to enliven his account to make music seem more like an organism than a dry recitation. Second, both of these models of form originate in mappings between language and music (a mapping more obvious in the excerpt from Goetschius, but evident as well in Prout's reference to poetic feet). To untangle matters, we need to begin with theories of musical form developed during the Enlightenment. From there we must proceed to explore how these theories were subsequently reformulated in the first half of the nineteenth century under the influence of political and ideological pressures that transformed the way music was taught.

## Grammar, Rhetoric, and Musical Form in the Eighteenth Century

Mappings between music and language go back at least to the Middle Ages. For instance, the ninth-century *Musica enchiriadis* opens with an analogy that anticipates Goetschius by over a thousand years: "Just as the elementary and indivisible constituents of speech (*vox articulata*) are letters, from which syllables are put together, and these in turn make up verbs and nouns, and from them is composed the fabric of a complete discourse, so the roots of song (*vox canora*) are *phthongi* [musical notes], which are called *soni* in Latin. The content of all music is ultimately reducible to them."[10] In the eighteenth century, mappings between language and music flourished and—as an almost inevitable consequence—became more manifold and more complicated. Enlightenment preoccupations with the origins of language and the place of music within systems of human communication led to a number of proposals for how language and music did, or did not, relate to one another.[11] The continued growth of print culture made it increasingly common for music and language to intersect—as score and text—on the printed page, a phenomenon exploited by author-printers such as Samuel Richardson.[12] Finally, a growing middle class fueled a demand for music instruction books, and, within these books, mappings between language and music provided a ready frame of reference for the explanation of fundamental musical concepts. During this period, two mappings from the domain of language onto the domain of music became particularly important for accounts of musical form: the first mapped from grammar onto music, the second mapped from rhetoric onto music.

10. *Musica enchiriadis and Scolica enchiriadis*, ed. Claude V. Palisca, trans. Raymond Erickson, Music Theory Translation Series (New Haven, Conn.:Yale University Press, 1995), 1.

11. For a discussion centered on eighteenth-century France, see Downing A. Thomas, *Music and the Origins of Language: Theories from the French Enlightenment*, New Perspectives in Muisic History and Criticism (Cambridge: Cambridge University Press, 1995).

12. Richardson's insertion of a musical score into the text of a novel is discussed by Janine Barchas in "The Engraved Score in *Clarissa*: An Intersection of Music, Narrative, and Graphic Design," *Eighteenth-Century Life* 20 (1996): 1–20.

GRAMMAR AND MUSIC Mappings between grammar and music were based on the assumption that music was language-like in its organization—that is, it had structural features analogous to those of language and a normative syntax. The mapping also relied on a familiarity with and respect for grammar acquired through the thorough grounding in the subjects of the Classical trivium—rhetoric, grammar, and logic—that was a common feature of early education in the eighteenth century. Just as every verbal construct, whether spoken or written, can be divided up into clauses, sentences, and paragraphs, the thinking went, so can musical works. This was the basis of Johann Mattheson's account of musical grammar, first proposed in his *Kern melodischer Wissenschaft* of 1737 and subsequently expanded in his *Der vollkommene Capellmeister* of 1739.[13] Mattheson began his discussion with vocal music, showing how the grammar of music had to be in rapport with that of the text: each of the significant divisions in the text had to be supported by analogous divisions in the music. This led Mattheson to a discussion of what things made for effective musical divisions (typically, cadence was the main articulative event) and how such divisions were graded from the barely noticeable to the emphatic.[14]

In a subsequent chapter of *Der vollkommene Capellmeister*, Mattheson argued that knowledge of this sort was even more important for knowing how to write instrumental music. He illustrated his discussion with the minuet shown in example 7.1.[15] Mattheson analyzes the whole as a paragraph, made up of two periods (or *Sätzen*; mm. 1–8 and 9–16), which are divided into a clause (marked by the colon below m. 4), a subclause (marked by the semicolon below m. 12), and phrases (marked by the commas below mm. 2, 6, and 10). Lacking, however, is any explanation for how these analytical decisions were made—why, for instance, mm. 1–4 constitute a clause, while mm. 9–12 constitute a subclause. The reader is left with a clear sense of the importance of grammatical structure to musical works—that grammar is, in fact, a self-evident property of musical works—but with only a few hints as to how the grammatical divisions in music are determined.

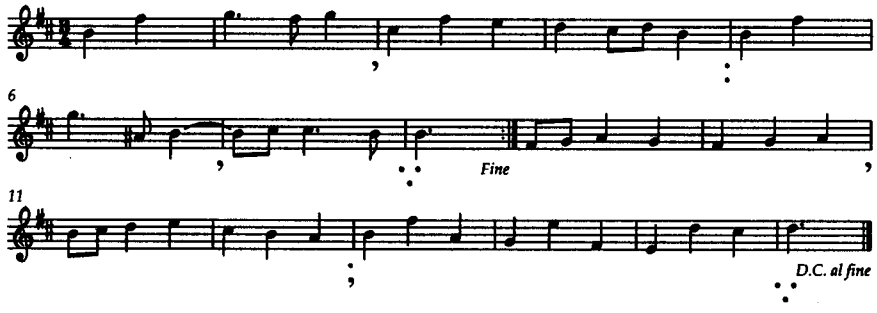
A more complete methodology for analyzing musical grammar was provided some fifteen years later by Joseph Riepel in his *Grundregln zur Tonordnung insgemein* of 1755 (the second volume of his *Anfangsgründe zur musicalischen Setzkunst*). Using a thoroughly schematic and comparatively systematic approach to musical organization (written in the dialogue style common in pedagogical treatises of the day), Riepel made explicit connections between grammatical components and musical passages, illustrated through hundreds of examples drawn from or applicable to

13. Johann Mattheson, *Kern melodischer Wissenchaft* (Hamburg: C. Herold, 1737; repr. Hildesheim: Georg Olms, 1976); idem, *Der vollkommene Capellmeister* (Hamburg: C. Herold, 1739; repr. Dokumenta musicologica, 1. Reihe, Druckschriften-Faksimiles, 5, ed. Margarete Reimann [Kassel: Bärenreiter, 1954]).

14. Mattheson, *Der vollkommene Capellmeister*, 180–95; Ernest C Harriss, *Johann Mattheson's* Der vollkommene Capellmeister: *A Revised Translation with Critical Commentary*, Studies in Musicology, 21 (Ann Arbor, Mich.: UMI Research Press, 1981), 380–404.

15. Mattheson, *Der vollkommene Capellmeister*, 223–25; Harriss, *Mattheson's* Der vollkommene Capellmeister, 451–53. For a discussion and analysis of this passage, in a slightly different context, see Ian Bent, ed., *Fugue, Form and Style* (vol. 1 of *Music Analysis in the Nineteenth Century*), Cambridge Readings in the Literature of Music (Cambridge: Cambridge University Press, 1994), 5–6.

EXAMPLE 7.1 Minuet analyzed by Johann Mattheson in *Der vollkommene Capellmeister*, part 2, chap. 13




instrumental music.[16] Riepel based his analysis of the grammatical function of a musical passage on harmonic structure and on the assumption that subdivision of the phrase into even-numbered units was the norm in instrumental music. Although his references to the grammatical structure of speech are explicit, they are also relatively few in number. Subphrase units (which Riepel called *Einschnitte*, and which were usually less than four measures in length) were marked by incomplete harmonic progressions—that is, they ended inconclusively. In contrast, the phrase (the *Satz* or, more frequently, the *Absatz*) was a four-bar unit that provided a complete harmonic progression: it had a conclusive ending. Each Absatz was further characterized on the basis of whether it ended on tonic, on the dominant, or in some other key.[17] By this means, Riepel was able to give an exhaustive (and at times exhausting) account of the different sorts of grammatical structures that might be evinced by music, based on the ways various components could be fit together and modified.

Riepel's approach to musical grammar was adopted by Heinrich Christoph Koch in his *Versuch einer Anleitung zur Composition* (1782–1793), but Koch began his account by returning to the model of speech. After noting the importance of "certain more or less noticeable resting points of the spirit" to those fine arts that attain their goal through speech—namely, poetry and rhetoric—Koch then draws the connection to music.

> Speech . . . breaks down into various sentences [*Perioden*] through the most noticeable of these resting points of the spirit; through the less noticeable the sentence, in turn,

16. Joseph Riepel, *Grundregln zur Tonordnung insgemein* (vol. 2 of *Anfangsgründe zur musicalischen Setzkunst*) (Regensburg: Johann Leopold Montag, 1755), 32–56; idem, *Sämtliche Schriften zur Musiktheorie*, ed. Thomas Emmerig, Wiener musikwissenschaftliche Beiträge, 20 (Vienna: Böhlau, 1996), 138–62.

17. For further discussion of Riepel's approach, see Justin London, "Riepel and *Absatz*: Poetic and Prosaic Aspects of Phrase Structure in 18th-Century Theory," *Journal of Musicology* 8 (1990): 505–19; Nola Reed Knouse, "Joseph Riepel and the Emerging Theory of Form in the Eighteenth Century," *Current Musicology* 4 (1986): 46–62; and John Walter Hill, "The Logic of Phrase Structure in Joseph Riepel's *Anfangsgründe zur musikalischen Setzkunst*, Part 2 (1755)," in *Festa musicologica: Essays in Honor of George J. Buelow*, ed. Thomas J. Mathiesen and Benito V. Rivera, Festschrift Series, 14 (Stuyvesant, N.Y.: Pendragon Press, 1995), 467–87.

> breaks down into separate clauses [*Sätze*] and parts of speech [*Redetheile*]. Just as in speech, the melody of a composition can be broken up into periods by means of analogous resting points, and these, again, into single phrases [*Sätze*] and melodic segments [*Theile*].[18]

However, Koch also saw limits to the applicability of mapping from the grammar of speech onto the grammar of music. After a long and thoughtful passage in which he pursues the notion that musical phrases are organized around subjects and predicates as are those of speech, Koch turns away from the model. He writes, "I am abandoning this comparison because, as already mentioned, subject and predicate cannot be differentiated enough in melodic sections. We must content ourselves with learning how to determine, through feeling, the presence or absence of resting points of the spirit and the nature of sections with regard to their completeness."[19] As did Riepel, Koch assumes subdivision of the phrase into even-numbered units to be the norm. He uses harmonic structure—specifically, the harmony with which each phrase ends[20]—to determine the divisions of musical grammar. And he builds larger musical structures by modifying and concatenating musical phrases.

Mappings between grammar and music of the sort evident in the work of Mattheson, Riepel, and Koch allowed writers in the eighteenth century to give detailed accounts of musical organization and to indicate how composers manipulated musical materials to create unique musical statements that nonetheless conformed to certain basic syntactic principals.[21] The mapping yields a characterization that emphasizes the static rather than the dynamic—what Koch came to call the "mechanical" aspect of music in the introduction to the second volume of his *Versuch*.[22] Each unit of musical grammar can be contemplated and discussed separately,

18. Heinrich Christoph Koch, *Introductory Essay on Composition: The Mechanical Rules of Melody, Sections 3 and 4*, trans. Nancy Kovaleff Baker, Music Theory Translation Series (New Haven, Conn.: Yale University Press, 1983), 1. The source from which Baker's translation is drawn is the second volume of Koch's *Versuch einer Anleitung zur Composition* (Leipzig: Adam Friedrich Böhme, 1787; repr. Hildesheim: Georg Olms, 1969). I have adapted this and the following quotation to restore Koch's somewhat unwieldy locution "resting points of the spirit" ("Ruhepuncte des Geistes"), which preserves the subjective aspect of determining music-grammatical divisions.

19. Koch, *Introductory Essay on Composition*, 6, translation adapted. In a note to this passage, Koch explains that he abandoned strict comparisons between the phrases of speech and melody because his treatise is intended for beginning musicians who wish to learn composition, and "these seldom have either grammatical knowledge of speech or familiarity with that part of logic which explains the different types of phrases and their closures."

20. Koch writes, "The essential difference between an internal phrase and a closing phrase depends on nothing more than the essential difference of ending formulas." Koch, *Introductory Essay on Composition*, 7.

21. As noted in chap. 2, it is through mappings such as these that theorists bring an integrated system of terms and structural relations to bear on particular problems of musical understanding. In chap. 2, the problems discussed had to do with ways to characterize tonal organization. Here the problems relate to how musical materials should be organized to create meaningful utterances.

22. Nancy Kovaleff Baker and Thomas Christensen, eds. and trans., *Aesthetics and the Art of Musical Composition in the German Enlightenment: Selected Writings of Johann Georg Sulzer and Heinrich Christoph Koch*, Cambridge Studies in Music Theory and Analysis, 7 (Cambridge: Cambridge University Press, 1995), 140.

and each can be fit together with other such units to create a complete statement at the level of a paragraph or higher. The perspective, then, is inherently constructivist: mapping grammar onto music isolates the building blocks of music and describes how they fit together to create coherent structures.

The account of musical form produced by this mapping is one that tends to emphasize complete, free-standing structures, as well as a synoptic view of such structures.[23] What is incomplete is ungrammatical; what is ungrammatical is incomplete. This is a view of basic materials well suited to pedagogy—mappings from grammar to music are invariably connected with lessons in composition—but one that has little to say about the process of constructing grammatical statements that are both complete and compelling.

RHETORIC AND MUSIC One of the principal uses of mappings from rhetoric to music during the eighteenth century was to describe how to go about composing music and the laying-out of musical materials in particular. Mattheson, in his *Das neu-eröffnete Orchestre* of 1713, offers the following analysis: "A [musical] composition comprises three things: Inventio (*Erfindung*), Elaboratio (*Ausarbeitung*), [and] Executio (*Ausführung* or *Aufführung*), which seem to display a close relationship to oratory or rhetoric. The last two things can be learned; the first has no competent master."[24] Of course, Mattheson is being somewhat disingenuous here, since the three techniques he mentions—and, for that matter, the very term "composition"—are all derived from the teaching of Classical rhetoric and were only secondarily applied to music. Twenty-six years later, in *Der vollkommene Capellmeister*, Mattheson expanded this plan to include a total of five stages: Inventio (*Erfindung*), Dispositio (*Einrichtung*), Elaboratio (*Ausarbeitung*), Decoratio (*Schmückung* or *Zierde*), and Executio (*Ausführung* or *Aufführung*).[25] In this more extensive application of rhetoric to music, Mattheson also drew on the structural categories of rhetorical theory to characterize the parts of a musical work (as distinct from the process of creating the work). These he identified as the Exordium (*Eingang*), Narratio (*Bericht*), Propositio (*Antrag*), Confirmatio (*Bekräfftigung*), Confutatio (*Wiederlegung*), and Peroratio (*Schluß*). Mattheson then used these categories to analyze the music of a *da capo* aria by Benedetto Marcello, but the result was at best a mixed success:

23. As I understand it, this synoptic view of the musical work does not recover what Michel Foucault calls the "simultaneity" of thought: musical works, whether "paragraphs" or "chapters," must still unfold in the temporal series specified by their grammar. The synoptic function instead reflects the taxonomy of forms basic to grammatical systems. There is an immediacy to such taxonomic units—be they phrases or paragraphs, books or symphonies—that is independent of the time it takes to traverse the individual syntactic components they comprise. On the simultaneity of thought, see Foucault, *The Order of Things: An Archaeology of the Human Sciences* (New York: Vintage Books, 1973), 83.

24. "Es gehören sonst zu einen Composition dreyerley: Inventio, (Die Erfindung) Elaboratio, (Die Ausarbeitung) Executio, (Die Ausführung oder Aufführung) welches eine ziemliche nahe Verwandschafft mit der Oratorie oder Rhetorique (Rede-Kunst) an den Tag legt; Die beyden lezten Stücke können erlernet werden; zum ersten hat sich noch kein tüchtiger Maitre." Johann Mattheson, *Das neu-eröffnete Orchestre* (Hamburg: Printed for the author, 1713; repr. Hildesheim: Georg Olms, 1993), 104.

25. Mattheson, *Der vollkommene Capellmeister*, 235; Harriss, *Mattheson's* Der vollkommene Capellmeister, 470.

the aria simply did not provide an adequate target for mappings from the traditional structure of oratory.[26]

Where Mattheson's mappings between *grammar* and music pointed to a fruitful model for characterizing musical organization, his proposed mappings between *rhetoric* and music never flourished in the same way. One problem was that the structural model of an oration used in Classical rhetoric was simply too rigid for music—indeed, there was every sign that it was coming to be regarded as too rigid for speech as well.[27] Another problem was that the flexibility of musical semantics made it difficult to develop hard and fast rules for what counted as a Propositio, a Confirmatio, or any of the other elements of an oration: it was clear to Mattheson and other theorists that music had meaning, but just how this meaning corresponded with the figures and processes of Classical rhetoric proved difficult to specify.[28]

The account of the *process* of composition afforded by rhetoric was given new promise when it was incorporated into the theory of aesthetics promulgated by Johann Georg Sulzer in his *Allgemeine Theorie der schönen Künste*, first published in two large volumes in 1771 and 1774. Sulzer drew on Classical rhetoric to develop a general characterization of the creative process, reflecting his belief that art shared with rhetoric the goal of moral improvement. According to Sulzer, the artist first discovers the ideas necessary to a work of art through invention. After the ideas have been discovered, the artist then realizes them through a three-stage process: *Anlage* (layout), *Ausführung* (realization), and *Ausarbeitung* (elaboration).[29] Of these, the latter two stages are identical with those identified by Mattheson in *Das neu-eröffnete Orchestre*. Sulzer's first stage corresponds with the second stage—Dispositio (*Einrichtung*)—of the expanded plan presented in *Der vollkommene Capellmeister*. (Of Sulzer's treatment of Mattheson's first stage—Inventio—more presently.)

The approach to composition outlined by Sulzer, while promising, was also very

26. Mattheson's analysis is in *Der vollkommene Capellmeister*, 237–39 (pp. 472–76 of the Harriss translation). For discussion, see Mark Evan Bonds, *Wordless Rhetoric: Musical Form and the Metaphor of Oration*, Studies in the History of Music, 4 (Cambridge, Mass.: Harvard University Press, 1991), 82–90; and Peter Hoyt, "Review of Mark Evan Bonds, *Wordless Rhetoric*," *Journal of Music Theory* 38 (1994): 127–29.

27. For further discussion of the limitations of rhetorical models, see Hoyt, "Review of *Wordless Rhetoric*," 130–33.

28. The problem is, if anything, more intractable at the level of individual rhetorical figures, as is evident in the complicated and occasionally redundant system set up by Joachim Burmeister in the early seventeenth century; see Burmeister, *Musical Poetics [1606]*, trans. Benito Rivera, Music Theory Translation Series (New Haven, Conn.: Yale University Press, 1993). For his part, Mattheson attempts no formalization of rhetorical figures, preferring to describe a few of them, indicate the applicability of many more, and then move on; see *Der vollkommene Capellmeister*, 243–44 (pp. 482–84 of the Harriss translation).

Although locating the specifics of rhetorical structure in musical discourse is challenging, the importance of rhetoric for all manner of thought and expression—even beyond the eighteenth century—should not be underestimated. For a discussion of the importance of rhetoric in scientific discourse, for instance, see Jeanne Fahnestock's recent *Rhetorical Figures in Science* (New York: Oxford University Press, 1999).

29. My summary of Sulzer's aesthetic theory and translation of the German terms for the stages in the compositional process are taken from Thomas Christensen's introduction to Sulzer's writings in Baker and Christensen, *Aesthetics and the Art of Musical Composition*, 3–24.

abstract, and it lacked specifics where music was concerned. These were supplied by Koch in the second and third volumes of his *Versuch*, in which he endeavored to show how Sulzer's three-stage process could be implemented with musical materials.[30] Where Sulzer was vague, having attempted to give an account of composition adequate for all the fine arts, Koch was specific, providing a detailed account of just what was involved at each stage of the compositional process, which he illustrated with an analysis of an aria from Carl Heinrich Graun's highly popular cantata *Der Tod Jesu* of 1755.[31] Making Sulzer's framework for the compositional process specific for music also provided a further justification for the technical language of musical grammar developed by Koch, for Koch believed it was by the conscious manipulation of this grammar that the beginning composer could create an aesthetically satisfying final work.

Mappings between rhetoric and music allowed music theorists like Mattheson and Koch to describe the compositional process in terms of distinct stages, which made it possible to give specific advice about how the novice composer should proceed along the way. Inasmuch as this process was construed as analogous to that undertaken in other art forms, music was placed on a parallel with sister arts to which it had formerly been but a poor cousin. Mapping from rhetoric to music also allowed music—more specifically, untexted music—to be regarded as a medium with persuasive powers similar to that of oratory. Although the trope of music as oratory was common at least since the Renaissance, preoccupations with the relationship between language and music during the Enlightenment revivified the analogy and put into relief the dynamic aspect of music: here was a type of oratory that not only unfolded over time but also was conducted in the pure language of tones—what Mattheson called *Klang-Rede* (sound-speech).

The view of the compositional process and its product afforded by mapping from rhetoric to music also highlighted the essential mystery of musical creation. This is plainest in the quotation from Mattheson's *Das neu-eröffnete Orchestre*—"invention is the process that knows no competent master"—but it is also evident in Sulzer's discussion of invention in the *Allgemeine Theorie* and in Koch's adaptation of that discussion in his *Versuch*. At first Sulzer maintains that invention is a capacity that can be improved through diligent and careful practice: "The power of invention, like the power of judgment, is a natural and inborn faculty that all men possess in proportion to their own genius. And just as one may turn to reason to aid one's power of judgment, so can one's talent for invention be augmented, in as much as it might be treated, like logic, as a part of philosophy."[32] Sulzer's initial suggestion for accomplishing this augmentation is to turn to rhetoric, urging the artist to exploit the rules and methodology that rhetoricians had developed for the express

30. Koch's realization of Sulzer's plan is discussed in Ian Bent, "The 'Compositional Process' in Music Theory 1713–1850," *Music Analysis* 3 (1984): 29–55; and by Nancy Kovaleff Baker in her introduction to Koch's writings in Baker and Christensen, *Aesthetics and the Art of Musical Composition*, 111–43. Koch was apparently unaware of Sulzer's work when he prepared his first volume, published in 1782.

31. Koch's analysis of Graun's aria can be found in translation in Baker and Christensen, *Aesthetics and the Art of Musical Composition*, 163–76.

32. Baker and Christensen, *Aesthetics and the Art of Musical Composition*, 56.

purpose of discovering useful ideas. But as attractive as such an orderly process is for the ever-moralizing Sulzer, he ultimately realizes that it is not the most productive where invention is concerned. Better to be constantly occupied with one's art, for one will surely stumble upon good ideas this way. As Sulzer ruefully admits,

> The most important inventions probably do not arise through the first deductive manner described above, but rather by the second way: the main subject appears only dimly at first to the artist; he recognizes its importance and takes time to think about its contents so it can be set in its proper light. This is how a famous composer told me he worked. He had more than once saved material that he heard by chance to use as a theme or subject for a composition. He never could have invented anything as good had he decided ahead of time to look for something having the identical character of expression.[33]

In this Koch agrees, writing, "Theory will never be able to invent a truly effective means to indicate just how the beginning composer should contrive that a beautiful melody arises in his soul."[34] However, once inspiration strikes and invention has commenced, the composer must work with haste to capture the whole in a plan of the sort Koch developed for Graun's aria. The plan in place, the composer may then proceed to the more mechanical side of musical composition.

The mystery surrounding invention, together with the focus on purely musical meaning afforded by mapping the structures of rhetoric onto music, suggests an independence for music adumbrated by Koch's reflections on the value of theory. Music, now more than simply materials to be manipulated, is something that arises in the soul of the composer. Although the full flowering of the idea that music was an autonomous medium would have to wait until the turn into the nineteenth century, we can see its seeds in eighteenth-century mappings from rhetoric to music. These provided a view of music as a powerful and highly dynamic medium with the potential to shape the thought of those who listened to it.[35]

GRAMMAR, RHETORIC, AND MUSICAL FORM It has been traditional and convenient to think of grammar and rhetoric as separate subdomains of language. Of the two, grammar is the simpler and more basic, necessary for proper communication but not the stuff to turn heads. Grammar, reckoned this way, is thus pure syntax, a careful if static ordering of linguistic units. By comparison, rhetoric is more complex and subtle, persuasive because it has true semantic force. As the means by which the passions can be moved, it is truly dynamic.[36] Nonetheless, the

33. Baker and Christensen, *Aesthetics and the Art of Musical Composition*, 61.

34. Baker and Christensen, *Aesthetics and the Art of Musical Composition*, 186.

35. For a discussion of the emergence of the idea that music was an autonomous art form, see Lydia Goehr, *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music* (Oxford: Clarendon Press, 1992), chap. 6; and Bellamy Hosler, *Changing Aesthetic Views of Instrumental Music in 18th-Century Germany*, Studies in Musicology, 42 (Ann Arbor, Mich.: UMI Research Press, 1981).

36. When Michel Foucault sought to characterize the distinct functions of rhetoric and grammar, he seemed rather deliberately to invert this traditional perspective, arguing that rhetoric is spatial and that grammar distributes that spatiality in time—that is, rhetoric is static and grammar is dynamic (Foucault, *The Order of Things*, 84). However, "spatial" here refers to the simultaneity of thought—to its essential and ongoing polyphony. Grammar distributes this spatiality in time only by freezing fragments of it so

distinctions between the two are hardly as clear-cut as this traditional perspective would have it, and nowhere is this more evident than when they are mapped onto music.

Returning briefly to Riepel, his classification of cadence types provides a case in point. In distinguishing between an Absatz that ends with a "garbled cadence" (*verstümmelter Schluß*) and one that ends with a "complete cadence" (*vollständiger Schluß*), Riepel uses the rhetorical term "enthymeme" for the first, and "syllogismus" for the second.[37] Descending momentarily into the labyrinth of rhetorical terms, we can discover that enthymeme refers to an incomplete syllogism, one in which either the major or minor premise of the syllogism is left implied. A syllogismus, by contrast, calls on the audience to draw an obvious conclusion, the major and minor premise having been stated.[38] Although the mapping between these two terms and the musical phenomena to which Riepel would have them apply is less than perfect, what is important is the implication that each of the grammatical components is imbued with semantic value—syntax has its semantic function. Indeed, as I argued in chapter 4, the manipulation of musical syntax is one of the principal means by which musical semantics is created.

Careful attention to the grain and texture of discourse about music is thus necessary to determine which of these metaphorical mappings is used to structure a given theory of music. Generally speaking, eighteenth-century writers used grammar to emphasize things that do not change in music—suitably defined, an Einschnitt is always an Einschnitt, an Absatz always an Absatz. The same writers used rhetoric to emphasize how music functions in time, either in the challenges it presents a composer attempting to deploy musical materials or in the reaction of a listener persuaded by a musical argument. Grammar gave a view of musical form that was static, and rhetoric gave a view of musical form that was dynamic. Significant, however, is the context relative to which these characterizations were made. As the dynamic view offered by rhetoric began to be more important for those who thought about music, and as the cultural context for thinking about music changed, mappings from language began to give way to other possibilities for conceptualizing music.

## Formal Theory in the Nineteenth Century

Over the course of the nineteenth century, the two basic models of musical form that emerged in the eighteenth century remained in place. Musical works could be regarded as manufactured things that comprised discrete—and inherently static—components, or as instances of quasi-abstract communication that unfolded dynamically in time. However, explicit mappings from language to music faded as they

---

that they can be ordered, much as musical materials, isolated from their becoming, are transformed into basic building blocks for the beginning composer.

37. Riepel, *Grundregln zur Tonordnung insgemein*, 59; idem, *Sämtliche Schriften zur Musiktheorie*, 153.

38. These definitions are drawn from Gideon Burton's online *Silva Rhetoricæ*, at http:/humanities.byu.edu/rhetoric/silva.htm.

were absorbed into musical nomenclature or replaced by other mappings that served a similar purpose. Instrumental in this process were the context within which music was taught and the intellectual framework relative to which music was conceptualized. Beginning in the final years of the eighteenth century, music instruction was taken over by institutions concerned as much with the advancement of the state as with education. In consequence, the move toward a complete and comprehensive curriculum that arose in furtherance of the former objective brought with it the development of integrated courses of study that employed both static and dynamic accounts of music. At around the same time, music came to be thought of in organicist terms, and, as a result, even the most basic and isolated of musical materials were imbued with dynamic characteristics. Hence, the conceptual metaphors for musical form established in the eighteenth century—form as static, or form as dynamic—remained in place, but they were activated by a number of different linguistic metaphors.

THE PARIS CONSERVATOIRE AND THE BERLIN UNIVERSITY During the eighteenth century, music education in the French- and German-speaking countries was carried out in a diffuse manner. Students typically studied with a local music master or at a school associated with a church, and there were few if any efforts at centralized music instruction. As a result of various political upheavals, this situation began to change in the late eighteenth century, and music education, together with instruction in other topics, began to be more closely controlled by the state.

The circumstances of music education changed most rapidly in France. After the 1789 revolution, numerous educational reforms were proposed, focused largely on efforts to create a utilitarian educational system and to promote deductive reasoning for the arts and sciences; some of these were at least partially enacted.[39] As a consequence of the diminishing role of the church after the revolution, the system of *maîtrises*—the music schools associated with the major cathedrals—collapsed, leaving two principal options for music instruction: the School of the National Guard (a product of the revolution) and the Old Regime's School of Singing (founded in 1784 with the intent of providing a reliable supply of singers for the Académie Royale de Musique, known familiarly as the Opéra). To realize the goal of conserving a French national music that was free from the influence of Italy and Germany, the Republic in 1795 consolidated the faculty and power of these two schools and formed the Conservatoire National de Musique et de Déclamation. The institution, as originally envisioned, was to provide free training for musicians (especially those destined for military service or the Opéra). It was also to serve as official arbiter of music used for state functions or music that was to be sanctioned by the state. The Conservatoire also became the source of the first national music theory, for its original charter included the provision that the texts used therein would be written by its faculty. Because of the political power that eventually

39. A thorough account of changes in music education after the French revolution is given in Cynthia M. Geselle, "The Institutionalization of Music Theory in France: 1764–1802" (Ph.D. diss., Princeton University, 1989), 235–306.

accrued to the Conservatoire, the texts that it adopted became the core of a national music curriculum.

The first texts adopted by the Conservatoire were models of practicality, simplifying those notions that were deemed unnecessarily complex and striving toward uniformity of method. As time went on, however, the faculty went beyond the highly utilitarian education envisioned by the original reformers and developed a sophisticated pedagogy that embraced virtually all aspects of musical thought. As an example, the pedagogy created by Antoine Reicha, who was appointed professor of counterpoint at the Conservatoire in 1817, extended from his *Traité de mélodie* of 1814 through the *Cours de composition musicale* (?1816–18) to the *Traité de haute composition musicale* of 1824–26.[40] In these texts, Reicha, with the zeal of a geometer and the passion of an artist, strove to provide a systematic approach that extended from the most basic aspects of musical structure—expressed through a grammar of music analogous to that of Koch—to a full account of the largest forms. Thus was institutionalized an approach to music that began with isolated components and culminated in large forms conceived as dramas enacted in sound.[41] Within this course of study, musical form was presented as both static and dynamic. In the early stages of the pedagogical plan, form was static, consisting of various components that could be assembled to create musical compositions. In the later stages, it became thoroughly dynamic, as this perspective yielded to one that accommodated, even if it did not completely sanction, the artistic license necessary to create works such as Mozart's Overture for the *Marriage of Figaro*.[42]

Changes to the German universities were not as sudden as those that led to the formation of the Conservatoire, but they were motivated by pragmatic and political concerns similar to those that shaped France's new educational system. There had developed in Germany over the course of the eighteenth century the perception that a university education served no useful purpose. This notion, together with efficiencies imposed by French rule and the growing financial burden presented by universities whose traditional sources of financial support had begun to erode, contributed to the disappearance of over half the German universities during the Napoleonic era.[43] As a way of countering this general disintegration, a delegation of professors from the University of Halle, which had been suspended by Napoleon in 1806, approached King Friedrich Wilhelm III in the late summer of 1807 and asked

40. Reicha's *Cours de composition musicale* ultimately replaced Charles-Simon Catel's *Traité d'harmonie*, which had been approved by the General Assembly of the Conservatoire in 1801 as the Conservatoire's main harmony text.

41. For a discussion of affinities between Reicha's account of the *grande coupe binaire* (now conventionally called sonata form) and drama, see Peter Hoyt, "The Concept of *Développement* in the Early Nineteenth Century," in *Music Theory in the Age of Romanticism*, ed. Ian Bent (Cambridge: Cambridge University Press, 1996), 141–62.

42. For a discussion of the range of Reicha's analytical perspective, which included the Overture from the *Marriage of Figaro*, see Ian Bent's commentary on one of the analyses from Reicha's *Traité de haute composition musicale*, in *Music Analysis in the Nineteenth Century*, 1: 152–53.

43. Theodore Ziolkowski, *German Romanticism and Its Institutions* (Princeton, N.J.: Princeton University Press, 1990), 228. My account of the history of German universities owes much to chap. 5 of this book, which is given over to a consideration of the institution of the university during the time of German Romanticism.

that their university be reestablished in Berlin. The king was apparently receptive to the idea and is reported to have responded that "the state must replace through intellectual powers what it has lost in the way of physical powers."[44] Loosely modeled on the ideals associated with the University of Jena when Friedrich Schiller, Johann Gottlieb Fichte, and Friedrich von Schelling lectured there, the University of Berlin was conceived as an institution combining faculties in the sciences, medicine, and the humanities. It was to be unified by a central faculty of philosophy professing an "encyclopedic" approach to knowledge. The university formally opened in October 1810 with a faculty that included Fichte and Friedrich Schleiermacher.

By the time the composer and music theorist Adolph Bernhard Marx was offered a chair at Berlin in 1830, much of the spirit of the Romantic university had faded to a memory.[45] What remained were a faculty that included Hegel in its ranks (who began lecturing at Berlin in 1818) and the notion that the university was an institution by, of, and for the state. Marx's encyclopedic approach to music theory resembled Reicha's sweeping attempt to provide a complete account of musical knowledge in his three treatises. In 1837, Marx published his four-volume *Die Lehre von der musikalischen Komposition, praktisch-theoretisch*. This comprehensive tract, which went through six full or partial editions during Marx's lifetime, presented a graduated course of study that began with the simplest of musical materials and forms and proceeded to the most complex.[46] However, where the pedagogy of Reicha's treatises was conceptualized around the models of language and drama, Marx adopted a different model, one closely associated with the same Romantic idealism that had shaped the structure of the University of Berlin. This was the model of organicism.

ORGANICISM As has often been noted, organicist models, which map features from principally vegetative domains onto inorganic or artificial domains, flourished among early Romantic philosophers and writers.[47] The model was, however, also in use during the late eighteenth century. Our friend Sulzer, as part of his more practical remarks on invention, advises the artist faced with an intractable situation to put the work to one side for a short while. Gradually, solutions will present themselves, until a way is found where none was thought to exist.

44. Ziolkowski, *German Romanticism*, 287.

45. An account of Marx's place at the University of Berlin can be had in Scott Burnham's introduction to Adolph Bernhard Marx, *Musical Form in the Age of Beethoven: Selected Writings on Theory and Method*, ed. and trans. Scott G. Burnham, Cambridge Studies in Music Theory and Analysis, 12 (Cambridge: Cambridge University Press, 1997), 5–6.

46. For overviews of the content of Marx's *Die Lehre* and discussions of the educational context within which it was embedded, see Burnham's introduction to *Musical Form in the Age of Beethoven*, 7–11; and Mark Spitzer, "Marx's 'Lehre' and the Science of Education: Towards the Recuperation of Music Pedagogy," *Music and Letters* 79 (1998): 489–526.

47. A thorough account of the role of organicist models in theories of invention during the eighteenth and nineteenth centuries (framed in terms of the memorable phrase "vegetable genius") can be found in Meyer Howard Abrams, *The Mirror and the Lamp: Romantic Theory and the Critical Tradition* (Oxford: Oxford University Press, 1953), 198–225.

> It is one of many remarkable secrets of psychology that apparently clear thoughts can, when one turns to them for deeper contemplation, refuse to be developed or comprehended in a clear way. But when they are left alone they will by themselves grow in greater clarity, much as that period in which plants germinate unnoticed and all at once burst into full bloom. Some concepts will gestate little by little in our mind, so to speak, and extract themselves from the mass of obscure ideas into the clear light. Every artist must rely upon such fortunate moments of genius, and if he cannot always find what he diligently seeks, he must await with patience that moment when his thoughts ripen.[48]

Here the mystery of invention, which Sulzer also connected with rhetoric, is explained in terms of an organicist model: ideas are seeds, and, given time and a little care, they will put forth complete and fully articulated intellectual constructs—the flowers of invention—that could never be produced by more mechanical means.

The perspective evident in Sulzer's account of unconscious psychology had a profound effect among German thinkers of the late eighteenth and early nineteenth centuries. These proponents of "nature philosophy," including Fichte, Schelling, Jean Paul Richter, and August Wilhelm Schlegel, embraced the newly emerging biological sciences, finding in them a source of concepts that could replace those derived from Cartesian or Newtonian mechanics. It was this model that Marx absorbed and adapted for the characterization of musical form.

In the first volume of *Die Lehre*, after a few preliminaries, Marx begins the journey that would culminate in a consideration of the most exalted compositional forms of his day with a brief contemplation of the relationship between a major scale and its tonic. He observes that no note within the scale provides the sense of repose that the tonic does: pausing on any other note is unsatisfying and demands further motion. This observation leads Marx to a principle basic to all music, the main terms of which are placed in typographical relief on the page:

> In this we have found the opposition that thoroughly permeates all of music:
> Rest and Motion
> Tonic and Scale.[49]

Marx extended this opposition to apply to all aspects of musical form, formulated as the dynamic impulse *rest—motion—rest*. For Marx, this basic pattern is truly comprehensive: it governs the modest four-measure phrase, as well as the entire expanse of sonata form.[50]

48. Baker and Christensen, *Aesthetics and the Art of Musical Composition*, 63.

49. "Hierin haben wir einen durch die ganze Tonkunst wirksam hindurchgehende Gegensatz,
Ruhe und Bewegung
Tonika und Tonleiter
gefunden." Adolph Bernhard Marx, *Die Lehre von der musikalischen Komposition, praktisch-theoretisch*, 4th ed. (Leipzig: Breitkopf and Härtel, 1841–47), 1: 27.

50. This characterization is from Burnham's introduction to Marx, *Musical Form in the Age of Beethoven*, 9.

It was in locating a source for this impulse that Marx turned to an organicist perspective that resonated not only with his era but also with the characterization of invention that had been offered by Sulzer some sixty-five years earlier. However, where for Sulzer the idea is the seed, for Marx it is the motive: "The motive is the primal configuration [*Urgestalt*] of everything musical, just as the germinal vesicle, that membranous sac filled with some fluid element (or perhaps with solid bodies), is the primal configuration of everything organic—the true primal plant or primal animal. The motive, this conjunction of two tones or some other unities, simply *is*."[51] Marx's mapping from the organic domain to the musical domain makes possible an account of form that extends from little more than a pair of notes to entire symphonies. Evident throughout is an approach to musical form that is not only dynamic (through the progression *rest—motion—rest*) but also motivated by the forces latent within the most basic of musical materials. According to Marx, the motive, as the essential organic entity, offers something that can connect with the composer's spirit. Should the spirit adopt the motive, the composer's intellect will then come into play, shaping the motive to create works of art.[52] The organicist perspective that Marx employs thus serves as a way to characterize the dynamic aspect of musical form and as a launching point for the pedagogy of composition. It also fulfills the promise of eighteenth-century theorists who first suggested that music was an independent domain of meaning and expression: in the organicist perspective of the early nineteenth century, the fully autonomous work of music is born.[53]

As heady as was the intellectual environment associated with the organicist perspective basic to his conception of form, Marx's *Die Lehre* was, as its full title announced, both theoretical and practical. In his discussion of the basic elements of musical form, Marx notes a certain confusion regarding the notion of a period, which is formed from two fundamental elements (the *Gang* and the *Satz*) derived from the motive: "If, in my compositional method, I associated the period with the fundamental forms, it was only for reasons of method, in order to keep this important and ubiquitously hard-working form constantly before the student's eyes from the very beginning."[54] Marx's pedagogy, shaped by the practical exigencies of a system of education intended to work the will of the state, had its static building blocks, as well as its dynamic idealism. Although Marx could always trace these building blocks back to their organic origins, they nonetheless represented basic components of a student's education and, at least for the beginner, of the musical work.

51. Marx, *Musical Form in the Age of Beethoven*, 66. This passage is from Marx's 1856 essay "Form in Music."

52. Marx, *Musical Form in the Age of Beethoven*, 66–67.

53. Although my interpretation of the origin of the fully autonomous work of music is somewhat different from that of Lydia Goehr, this is due mostly to the somewhat different focus of my approach. I am concerned here not simply with the idea of a fully autonomous work but with the complex of ideas represented by organicism that allowed music theorists to articulate the notion of an autonomous work on both a practical and a theoretical level. Goehr's interpretation can be found in *The Imaginary Museum of Musical Works*, 159–75.

54. Marx, *Musical Form in the Age of Beethoven*, 68.

That the high abstraction of Marx's theory caused problems, at least for some theorists, is suggested by the approach to motive taken by Hugo Riemann. Riemann, as editor of the ninth edition of Marx's *Die Lehre* (which appeared in 1887), was well acquainted with Marx's conception of a motive. Nonetheless, in his *Musik-Lexikon* of 1882, Riemann began his definition of a motive by mapping from yet another domain: "Motive means in music, as it does in architecture, the most basic characteristic parts of an artistic structure."[55] Riemann goes on to distinguish between rhythmic motives, melodic motives, and harmonic motives, all conceived in architectural terms. Although the dynamic impulse basic to Marx's conception of motive would appear to have vanished, in other of his works it is evident that Riemann has simply transferred it to metric structure, which he construed as having, as its essence, the impulse *upbeat-downbeat*.[56] This construal gave Riemann a way to represent the musical dynamism that Marx had developed through an organicist metaphysics in more concrete terms. To the extent that the student or reader understands the basic experience of musical rhythm, so will he or she understand the dynamic impulse basic to music.

MODELS OF MUSICAL FORM To all appearances, Ebenezer Prout and Percy Goetschius are a bit like Alice and the Mad Hatter (although I shall not venture to say which theorist takes which role), seeming to talk about the same thing but doing so in very different terms. But where Alice and the Hatter are engaged with the elusive topic of Time, Prout and Goetschius are dealing with the rather more humble matter of Motive. And where the Hatter is as surely a denizen of Wonderland as Alice is not, Prout and Goetschius draw their ideas about motive and form from the same pool of German music scholarship.

As we can now see, Prout's approach is informed by the work of Marx and Riemann: he conceives of motive in terms of both metrical accentuation and organicism, although in the end it is the former that is foregrounded in his theory while the latter remains safely in the background. Goetschius, in contrast, conceives of motive in almost exclusively grammatical terms. His approach resembles, first and foremost, that of Reicha (whose three treatises were translated into German by Carl Czerny and published in a four-volume French-German bilingual edition in 1832). To a lesser extent, it resembles that of Koch (whose work was not as widely known in the nineteenth century). Thus, Goetschius's method is shaped not only by a grammatical approach to musical form but also by a pedagogy that was thoroughly pragmatic. Just as Reicha's theory of form was part of a larger program of complete instruction in music for a nation intent on building its musical identity, so Goetschius's manual was part of a series of texts intended to fulfill the same function in the New World.

Beyond being able to sort out the lineages of Prout's and Goetschius's ideas

55. "*Motive* nennt man in der Musik wie in der Architektur die letzten charakteristischen Glieder eines Kunstgebildes." Hugo Riemann, *Musik-Lexikon* (Leipzig: Verlag des Bibliographischen Instituts, 1882), 605.

56. Hugo Riemann, *System der musikalischen Rhythmik und Metrik* (Leipzig: Breitkopf and Härtel, 1903), 13–18.

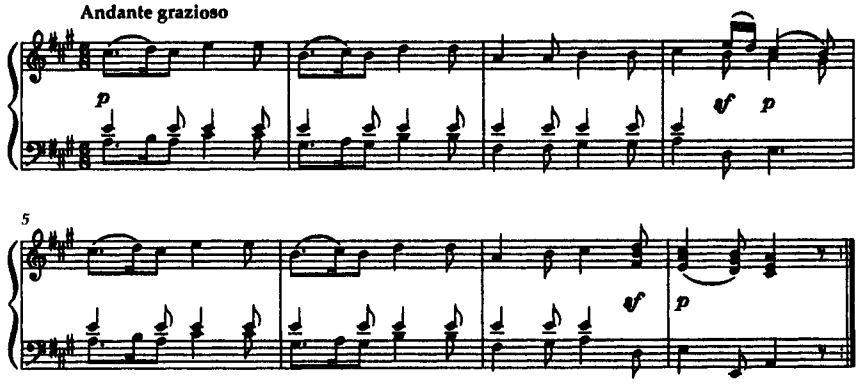
about motive and form, however, we now also have a better grasp on the different sorts of conceptual work done by the mappings they employ. Mappings that yield a static account of musical form put into relief the explicitly conceptual side of music: they rely on our ability to isolate musical events or series of musical events and put them into categories. Members of the categories can then be compared with one another, as can the categories themselves. The view that results is of music as material that can be shaped and manipulated, whose meaning is neatly contained in discrete packages. Mappings that yield a dynamic account of musical form put into relief the experiential side of music: music as something to be lived through, which may move us or not, but which in all cases resists packaging into neat conceptual containers. Here meaning is at best only imperfectly contained and is in constant danger of overflowing the crude boundaries of ontological necessity.

As the examples drawn from the ways musicians characterized form in the eighteenth and nineteenth centuries show, the cross-domain mappings we use in our theories of music are more than simple curiosities of language, for these mappings have everything to do with our understanding of how music is organized, along with what counts as music in the first place.

## TWO MODELS OF MUSICAL HIERARCHY

As I attempted to show in the preceding pages, descriptions of musical form have occasioned a number of cross-domain mappings, operating on both the specific and generic level. The situation with musical hierarchy is somewhat different, in part because hierarchy was not deemed an attribute of musical works as such but was instead used to characterize the organization of musical materials. Consequently, explicit discussions of hierarchy rarely appear in accounts of works of music but instead form the context for those accounts. As an example, let us consider three contemporary analyses of the same passage: the first eight measures of the theme from the opening movement of Mozart's Piano Sonata in A major K. 331. All three analyses deal with the rhythmic structure of the passage (the score appears in ex. 7.2).

EXAMPLE 7.2 W. A. Mozart, Piano Sonata in A major K. 331, first movement, theme, mm. 1–8



Edward Cone's analysis of the theme appeared in his *Musical Form and Musical Performance* of 1968. The analysis comes just after the introduction of a striking metaphor for the melodic and harmonic shape of a musical phrase: Cone likens the shape of a phrase to a game of catch:

> If I throw a ball and you catch it, the completed action must consist of three parts: the throw, the transit, and the catch. There are, so to speak, two fixed points: the initiation of the energy and the goal toward which it is directed; the time and distance between them are spanned by the moving ball. In the same way, the typical musical phrase consists of an initial downbeat (/), a period of motion (⌣), and a point of arrival marked by a cadential downbeat (\).[57]

This metaphor then provides the framework for his analysis of the rhythm of Mozart's theme, a portion of which is given in example 7.3. Cone rejects a simple alternation of strong and weak measures of the sort posited by Riemann on the basis of harmony. Cone asserts, rather, that the first and fifth measures of the period should be considered strong because of the firm statement of tonic. The fourth and eighth measures should also be deemed strong because of the emphasis provided by the cadences. A consideration of the motivic structure of the period enables Cone to refine this analysis. The first phrase consists of two individual sequential measures followed by a two-measure unit. The shape of the two-measure unit duplicates the shape of the phrase in miniature: it consists of two half-measure units followed by a full measure. As we see in example 7.3, the internal dynamic of the two-measure unit is the same as that of the complete phrase. The second phrase begins with the same sequential measures as the first, but in m. 7 the compression that prepares the closing cadence brings the rise from A4 to C♯5 into prominence. Cone's hearing of this rise as a third member of the sequence prompts him to assign it the same rhythmic symbol as the second member of the sequence, heard in m. 6. Voice-leading concerns and the forward energy of the *sforzando* cause Cone to group the final eighth note of m. 7 with the cadential material of m. 8; this constitutes the closing downbeat.

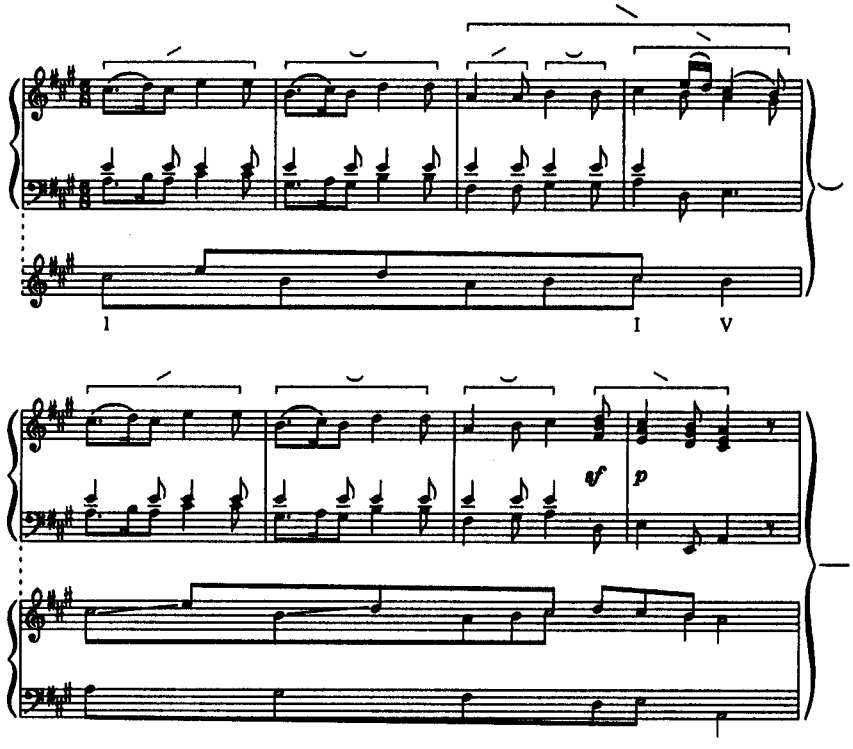
Although Leonard Meyer's analysis of this theme appeared after Cone's (he presented it first in one of the Ernest Bloch lectures he gave at the University of California at Berkeley in 1971), his methodology is essentially the same as that which he had developed with Grosvenor Cooper over a decade earlier. Cooper and Meyer understood rhythmic structure to be perceived as "an organic process in which smaller rhythmic motives, while possessing a shape and structure of their own, also function as integral parts of a larger rhythmic organization."[58] Rhythmic relationships are analyzed as patterns of beats, in which a stable accent and one or more weak beats are grouped together in different ways; these low-level, foreground patterns combine with one another in various ways to form more extended rhythmic groupings. In this way the musical surface gives rise to a hierarchy of rhythmic groupings.[59]

57. Edward T. Cone, *Musical Form and Musical Performance* (New York: W. W. Norton, 1968), 26–27.

58. Grosvenor W. Cooper and Leonard B. Meyer, *The Rhythmic Structure of Music* (Chicago: University of Chicago Press, 1960), 2.

59. Leonard B. Meyer, *Explaining Music: Essays and Explorations* (Berkeley: University of California Press, 1973; repr. Chicago: University of Chicago Press, 1978), 27–28. Meyer's Bloch Lectures of 1971 form the basis for *Explaining Music.*

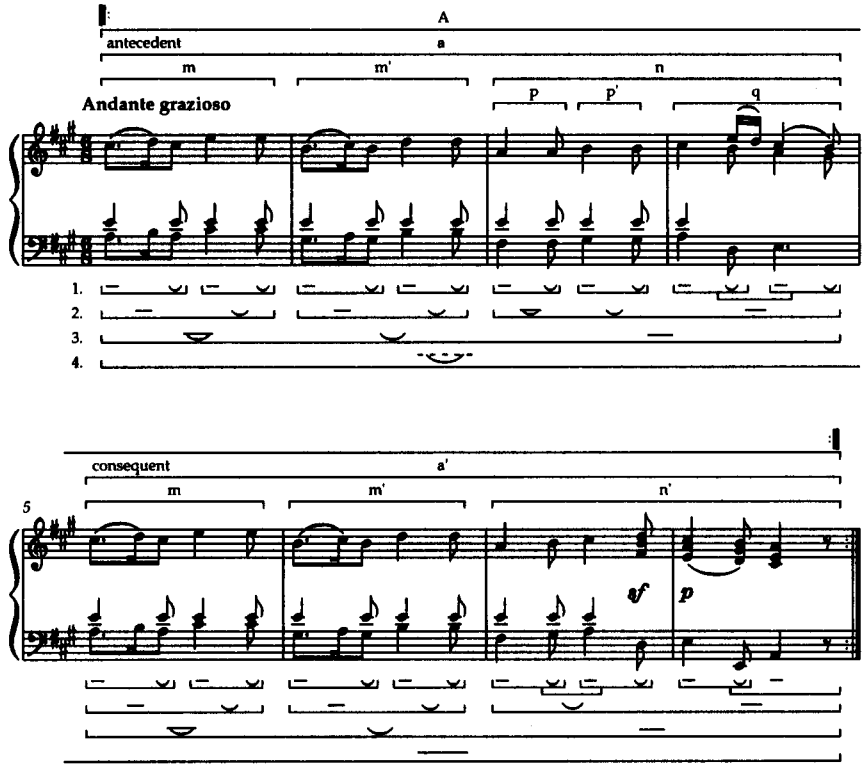
EXAMPLE 7.3 Edward Cone's analysis of the opening theme of Mozart's Piano Sonata K. 331, mm. 1–8




Meyer's analysis of the opening period of Mozart's theme is given in example 7.4. The first level of the analysis starts at the half-measure level of the music, important for Meyer because the ♩♫♫ patterns of the first two measures of each phrase do much to determine the grouping and accentual pattern of the complete bar. In mm. 3 and 4, Meyer's analysis on level 2 corresponds closely with Cone's reading: although the symbols are slightly different (Meyer's indication for a retrospectively weak accent replacing Cone's symbol for an initial downbeat), on this level as well as the third and fourth, the analyses for the most part agree. The main difference between the two is that Cone *starts* his analysis at what for Meyer is the third level of a rhythmic hierarchy.

In his article "The Theory and Analysis of Tonal Rhythm" (1978), Robert Morgan noted that, although both Cone's and Meyer's work reflected the influence of Schenker's theory of tonal music, their analyses of Mozart's theme were still beholden to certain aspects of metrical dogma. For his part, Morgan proposed considering the theme from a more consistently Schenkerian viewpoint. Morgan noted that Schenker's theory supplied a method for locating points at which structural motions originate and terminate; his idea was to correlate these points of structural origin and termination with the points of principal accent through which the larger rhythm of a passage is regulated. In Morgan's analysis, the first phrase of the theme

EXAMPLE 7.4 Leonard B. Meyer's analysis of the opening theme of Mozart's Piano Sonata K. 331, mm. 1–8



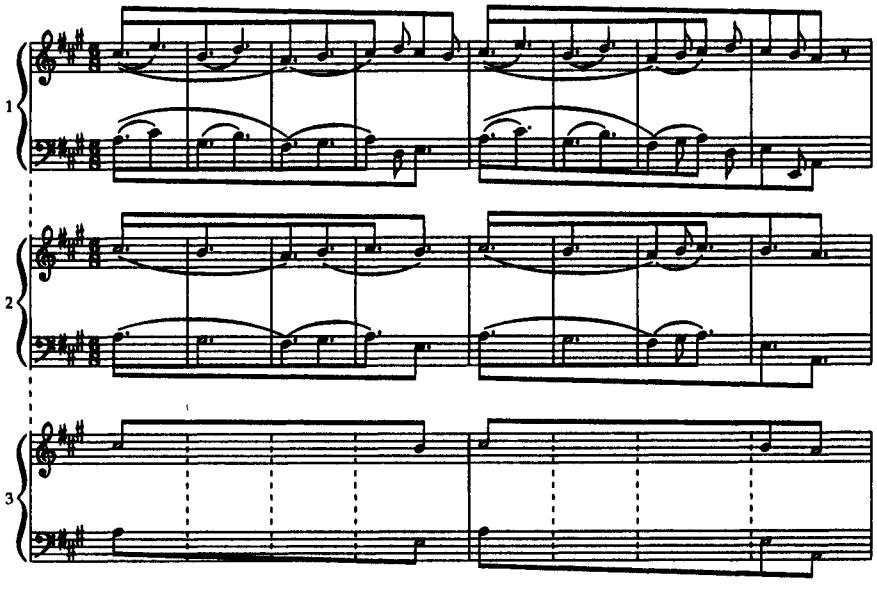

is governed by the overall motion from C♯5 and A3, in m. 1, to B4 and E3 in m. 4 (see ex. 7.5, level 3). This overall motion is articulated on the next level (level 2 of ex. 7.5) by a motion down to A4 and F♯3 in m. 3. Although the A could be seen as the completion of a third span from C♯5, the bass F♯3 and the acceleration away from A4 to B4 on the second beat of m. 3 prevent closure.

Up to this point, Morgan's analysis agrees, in most respects, with the analyses by Cone and Meyer. Morgan does not, however, believe the first beat of m. 4 should be heard as accented relative to the second beat. The acceleration of m. 3 continues into m. 4, in his view, carrying the music forward to the arrival of the bass on the dominant on the second beat. This arrival, according to Morgan's analysis, serves as the rhythmic goal and closing accent of the phrase. From Morgan's perspective, then, the rhythm of this phrase is not a matter of strong and weak beats or strong and weak measures; it is created rather by an articulated tonal process that moves between the two points defined by the outer-voice frame.[60]

The acceleration in the first phrase is answered by an even greater acceleration

60. Robert P. Morgan, "The Theory and Analysis of Tonal Rhythm," *Musical Quarterly* 64 (1978): 446.

EXAMPLE 7.5 Robert Morgan's analysis of the opening theme of Mozart's Piano Sonata K. 331, mm. 1–8




in the second, for the consequent phrase must go through the dominant to arrive on the tonic. That arrival gives a different character to the closing accent of the second phrase; for Morgan, it is the difference in character between the two closing accents, and not some analogy to two large beats, that is responsible for the upbeat-downbeat effect created by the pairing of antecedent and consequent. Implicit in this analysis is the notion that accent refers to a *point* of emphasis, which also leads Morgan to reject Cone's and Meyer's reading of the rhythm of mm. 3 and 4 as a diminution of the overall rhythm of mm. 1 through 4. If mm. 3 and 4 supply the downbeat for mm. 1 and 2, they must represent an arrival, but in fact the point of arrival occurs only at the end of the unit. Morgan does not believe m. 3 can simultaneously be both an arrival and a departure.

Although these three analyses converge on a number of points, their divergences reflect two different models of musical hierarchy. The first model, which gives rise to what I call *chain-of-being hierarchies*, does a good job of capturing the harmonic and scale-degree dependencies typical of the harmonic organization of instrumental music of the late eighteenth century. The second model, which gives rise to what I call *atomistic hierarchies*, does a good job of capturing the regularity and nesting of accentual patterns typical of musical meter during the same period. Because both models are associated with what appears to be a common set of structural relations—called "hierarchy"—we might be misled into thinking that they are simply variations of each other. Such is not the case, however, for these models developed at different times in history, and with reference to different domains. These differences are reflected in the structure of the models and also in the ways they have been applied to music.

## Historical Models of Hierarchical Organization

The principal mark of chain-of-being hierarchies is a conception of the universe as a precisely arranged system of interdependent levels or degrees of existence extending from a Supreme Being down to the lowliest organism. Although this concept is part of an incredibly rich collection of ideas—one writer has suggested that the history of chain-of-being hierarchies is, in reality, a history of Occidental thought[61]—only a few central aspects need concern us here.

The ingredients for this complex of ideas came from Plato and Aristotle; Aristotle's connection of "powers of the soul" with levels of being in *De anima* proved particularly influential. It was in Neoplatonism, however, that these ideas first appeared fully organized into a coherent general scheme.[62] Crucial to this organization was Plotinus's idea that the perfection of the One included a superabundance which, overflowing, was the source of the Many. According to Plotinus, this generation of the Many from the One could not come to an end as long as any possible variety of being in the descending series was left unrealized.

One of the principal ways the Neoplatonic cosmology was transmitted to the Middle Ages was through Macrobius's *Commentary on the Dream of Scipio* (which dates from the late fourth to early fifth centuries A.D.). In a passage expanding on the revelation presented to Scipio that minds are given to man out of the eternal and divine fire of the stars, Macrobius writes:

> Accordingly, since Mind emanates from the Supreme God and Soul from Mind, and Mind, indeed, forms and suffuses all below with life, and since this is the one splendor lighting up everything and visible in all, like a countenance reflected in many mirrors arranged in a row, and since all follow on in continuous succession, degenerating step by step in their downward course, the close observer will find that from the Supreme God even to the bottommost dregs of the universe there is one tie, binding at every link and never broken. This is the golden chain of Homer which, he tells us, God ordered to hang down from the sky to the earth.[63]

61. C. A. Patrides, "Hierarchy and Order," in *Dictionary of the History of Ideas: Studies of Selected Pivotal Ideas*, ed. Philip P. Wiener (New York: Charles Scribner's Sons, 1973), 2: 434.

62. Concerning the complex of ideas associated with the great chain of being and Neoplatonism, see Arthur O. Lovejoy, *The Great Chain of Being: A Study of the History of an Idea* (Cambridge, Mass.: Harvard University Press, 1936), 61; concerning *De anima*, see p. 58.

63. Ambrosius Aurelius Theodosius Macrobius, *Commentary on the Dream of Scipio*, ed. and trans. William Harris Stahl (New York: Columbia University Press, 1990), 145. It is also in Macrobius (pp. 185–200 in the Stahl translation) that we find a detailed discussion of the harmony of the spheres, as well as the legend of Pythagoras and the blacksmith's shop. The connection between harmony and order found in Macrobius was an important one, and it may have contributed to the application of a metaphorical chain-of-being hierarchy to music. However, it is also striking that the image of a ladder is used in both domains. In the Middle Ages, the hierarchical organization summoned by the Great Chain of Being was commonly represented by Jacob's Ladder; see Paul G. Kuntz, "A Formal Preface and an Informal Conclusion to *The Great Chain of Being*: The Universality and Necessity of Hierarchical Thought," in *Jacob's Ladder and the Tree of Life: Concepts of Hierarchy and the Great Chain of Being*, ed. Marion Leathers Kuntz and Paul Grimley Kuntz, American University Studies, Ser. V, Philosophy, 14 (New York: Peter Lang, 1987), 6. Also, some early mnemonic diagrams show the gamut as notes on a ladder; see Karol Berger, *Musica Ficta: Theories of Accidental Inflection in Vocal Polyphony from Marchetto da Padova to Gioseffo Zarlino* (Cambridge: Cambridge University Press, 1987), 9–10.

The notion of a Great Chain of Being proved a concise and powerful model through which humans could give order to their universe. It also provided a model through which the distribution of power could be accomplished and justified. The model is quite evident in ecclesiastical hierarchies: power flows from the top down, in the same way that being and soul flow from the Godhead down through the various levels of being, from highest to lowest.

Inherent in the Neoplatonic scheme is a tension between the self-sufficiency of the One and the multitudinous abundance of creation it gives rise to: it is a paradox that the One needs nothing else and yet produces everything else. Giordano Bruno, writing in the sixteenth century, reveled in this paradox.[64] Yet it was also Bruno who, by combining the concept of atomism with the comprehensive organization provided by the Neoplatonic scheme, inspired a way of thinking about hierarchical organization that would eventually compete with the Great Chain of Being.

For Bruno, atomism was a metaphysical principle that provided the basis for a demonstration of the underlying unity of all nature.[65] This view was especially influential in England, where Bruno lived from 1583 to 1585 and where two factors combined to transform his vague, metaphysical scheme into a compelling account of order. The first factor came in the form of a challenge to the model for the distribution of power provided by the application of a chain-of-being hierarchy to the social sphere. In a social chain-of-being hierarchy, the power of any individual is linked, as an essential property, to the level occupied by that individual within the system: the higher the level, the greater the power over others. A refutation of this model can be seen in Thomas Hobbes's *Leviathan* of 1651. Hobbes argued that all humans were basically equal in the faculties of body and mind: all had approximately equal power over others. Because of this equality, the natural state of individual humans was one of competition. Only by imposing a social order on humans from without, in the form of a common power "to keep them all in awe," was it possible to overcome the natural state of virtual or actual war.[66] Thus the properties ascribed to each member of a society (which result in equal measures of power for all members) are different from those according to which societal order should be established (which result in unequal measures of power): the properties of the individual are independent of the properties of the system.

The second factor important in the development of an alternative view of hierarchical organization was the increased presence of mechanical devices in people's lives.[67] These machines provided thinkers of the time with practical models of the organization of nature. Among the most important of these was the mechanical clock. As Gideon Freudenthal has observed: "In the construction of the mechanical clock it was possible to produce complicated movements by an appropriate disposition of gears and a driving force. The task of science [in the seventeenth and eighteenth centuries] can be interpreted as the attempt to discover in a limited area

64. Lovejoy, *The Great Chain of Being*, 120.

65. Robert H. Kargon, "Atomism in the Seventeenth Century," in *Dictionary of the History of Ideas: Studies of Selected Pivotal Ideas*, ed. Philip P. Wiener (New York: Charles Scribner's Sons, 1973), 1: 133.

66. Thomas Hobbes, *Leviathan, or the Matter, Forme, and Power of a Common-Wealth Ecclesiasticall and Civill* (London: Andrew Crooke, 1651), chap. 13.

67. Kargon, "Atomism in the Seventeenth Century," 1: 132.

and to a limited extent the principles of construction of the divine clock."[68] The model of a clock is one that appears consistently in the dispute between Isaac Newton and Gottfried Wilhelm Leibniz on the ultimate nature of matter. Of particular interest are the different methodological approaches used by Newton and Leibniz. Leibniz held to what had become, by the early eighteenth century, a traditional Baconian methodology: scientific inquiry should seek to disassemble the divine clock, piece by piece, never speculating beyond what could be observed. Although Newton was in sympathy with this methodology, it was inadequate for a purely mechanical account of physics, since the ultimate workings of the clock—its atomic structure—could not be observed. Newton's solution, given in the final Query appended to the fourth edition of his *Opticks*, was a practical one that divided scientific theorizing into two halves.[69] The first half is analysis: "This Analysis consists in making Experiments and Observations, and in drawing general Conclusions from them by Induction, and admitting of no Objections against Conclusions, but such as are taken from Experiments, or other certain Truths."[70] The second half is synthesis or composition, in which the various inductive generalizations produced by analysis are organized into a logical system based on a limited number of essential principles. In the case of the divine clock, synthesis of necessity played a larger role than analysis (which led Leibniz to accuse Newton of arguing from hypothesis rather than from observation). However, Newton believed this exception to his overall methodology was justified by the potential results:

> To derive two or three general Principles of Motion from Phænomena, and afterwards to tell us how the Properties and Actions of all corporeal Things follow from those manifest Principles, would be a very great step in Philosophy, though the Causes of those Principles were not yet discover'd: And therefore I scruple not to propose the Principles of Motion above-mention'd, they being of very general Extent, and leave their Causes to be found out.[71]

In other words, the scientist simply had to assume that the general principles produced by synthesis held, even on levels that could not be directly observed: the mechanical workings of the divine clock were assumed to be consistent throughout.

When combined with the independence of elemental and systemic properties proposed in the social sphere, this approach developed into a new conception of hierarchical order, which I call an atomistic hierarchy. In contrast to a chain-of-being hierarchy, an atomistic hierarchy can be built from the bottom up, its organization governed by a set of general principles that obtain no matter what the

68. Gideon Freudenthal, *Atom and Individual in the Age of Newton: On the Genesis of the Mechanistic World View*, trans. Peter McLaughlin, Boston Studies in the Philosophy of Science, 88 (Dordrecht: D. Reidel, 1986), 63. Sulzer also used the analogy of a clock, but to characterize the unity of a work of art; see Baker and Christensen, *Aesthetics and the Art of Musical Composition*, 43–46.

69. R. I. G. Hughes, "Reason and Experiment in Newton's *Opticks*: Comments on Peter Achinstein," in *Philosophical Perspectives on Newtonian Science*, ed. Phillip Bricker and R. I. G. Hughes (Cambridge, Mass.: MIT Press, 1990), 177.

70. Isaac Newton, *Opticks, or a Treatise of the Reflections, Refractions, Inflections and Colours of Light*, 4th ed. (London: W. Innys, 1730; repr. New York: Dover, 1952), 404.

71. Newton, *Opticks*, 401–02.

specific attributes of a given level. It is a concise and concrete account of organization, and one that, through applications in physics, chemistry, and biology, has had an enormous influence on how humans view their world.

## Structural Features of Chain-of-Being and Atomistic Hierarchies

Each of these models of hierarchy offers a different view of how the world is organized. The fundamental conceit of chain-of-being hierarchies is to regard a domain as a huge organism pervaded by a force, the origins of which are mysterious.[72] "Being" and "political power" are both examples of such forces. The mysterious force behind any chain-of-being hierarchy manifests itself as properties instantiated in varying degrees by the elements of the domain. In a traditional Great Chain of Being, for instance, a rock has the property of *substance*, a tree has the properties of *substance* and *life*, and an insect has the properties of *substance*, *life*, and *mobility*. The domain can be organized into a hierarchy in which any given level is distinguished by a specific set of properties that embody aspects of being. These properties include all of the properties that distinguished lower levels, plus properties unique to the given level: trees include *substance* and add *life*; insects include *substance* and *life* and add *mobility*. Inclusion within the hierarchy is thus of properties, but not elements.

The fundamental conceit of atomistic hierarchies is to regard a domain as an extended mechanism operating according to a limited set of general principles. The laws of classical physics are an example of one such set of principles. All of the actions and properties of elements within the domain, from the smallest to the largest, and from the most simple to the most complex, follow these general principles. Thus subatomic particles and planets both conform to the same principles. Classes of elements may be distinguished according to common conformance with the general principles of the system. The actions and properties of electrons, protons, and neutrons on the atomic level are by no means identical, but they can be distinguished from the actions and properties of atoms on the molecular level. In addition, all of the elements within a given class have the shared property of combining into units: electrons, protons, and neutrons combine to form atoms, and atoms combine to form molecules. The domain can be organized into a hierarchy in which each level is a conformance class whose elements combine into units that constitute the elements of the next higher level in the hierarchy. This process continues recursively until the limits of the system are reached. Inclusion within the hierarchy is thus of elements, but not properties.

The unique features of each of these models of hierarchy can be characterized in terms of relationships that are established between the general, abstract property of "energy" and the CONTAINER schema. "Energy" takes form as either the mysterious

72. While the parallels with the ideology of organicism that blossomed in the late eighteenth century are intriguing, it should be remembered that the organicism associated with chain-of-being hierarchies dates back at least to the Middle Ages. Accordingly, the mysterious origins of the power behind chain-of-being hierarchies is somewhat different from the mystery associated with Romantic organicism since the former is an essential property of the model where the latter would be better described as accidental.

force of a chain-of-being hierarchy or the power that drives the mechanism of an atomistic hierarchy. For both models, the CONTAINER schema provides the basic internal structure of the levels that make up the hierarchical system.

In a chain-of-being hierarchy, the energy proper to the system (that is, "being" or "political power") is distributed unequally throughout the system. Systemic organization is consequently based on the amount of energy (interpreted as "properties of being" or "extent of political power") each element of the system has. The containers (or hierarchical levels) that articulate this organization group elements according to their amount of energy (or systemic properties). Although chain-of-being hierarchies often offer vivid and dynamic accounts of organization, hierarchical structure (based on the contents of the containers that make up the hierarchy) is, in fact, relatively abstract.

In an atomistic hierarchy, energy is distributed throughout the system according to a fixed set of general principles. Systemic organization is consequently based on the way different elements of the system manifest these principles. The containers (or hierarchical levels) that articulate this organization group elements according to common conformance with the general principles. The generality of the organizing principles of atomistic hierarchies means that the principles can seem somewhat removed from actual experience, although hierarchical structure (based on the contents of the containers that make up the hierarchy) is, in fact, relatively concrete.

Given these structural attributes, each model has a slightly different potential as a source domain. A chain-of-being hierarchy will map these predicates onto a target domain:

- The target domain is pervaded by a mysterious force.
- This force manifests itself as properties instantiated to varying degrees by the elements of the domain.
- The domain can be organized into a hierarchy, in which each level is distinguished by a unique set of properties.
- Each level of the hierarchy includes all of the properties of the next lower level and adds new properties unique to that level.
- The higher levels of the hierarchy more completely manifest the mysterious force (and are thus more perfect) than the lower levels of the hierarchy.

Because chain-of-being hierarchies have often been used to structure natural and human domains, applying the model to other domains may suggest that elements of that domain have natural or anthropomorphic aspects, even though these are not necessary entailments of the mapping.[73]

An atomistic hierarchy will map these predicates onto a target domain:

- The target domain is analogous to an extended mechanism.
- This mechanism operates according to a limited set of general principles.
- All of the actions and properties of elements within the domain follow these general principles.

73. As the analyses at the end of chap. 2 suggest, factoring anthropomorphism into the application of chain-of-being hierarchies to different domains would quickly lead to a dense and highly interconnected network of correlated domains.

- Classes of elements may be distinguished according to common conformance with the general principles and the shared property of combining into units.
- The domain can be organized into a hierarchy, in which each level represents a conformance class.
- The elements on each level of the hierarchy combine to form units; these units constitute the elements of the next higher level in the hierarchy.
- The recursive combination of elements into units continues until the limits of the system are reached.

Because atomistic hierarchies have often been used in scientific explanations, applying the model to other domains may suggest that a "scientific" approach is being used, even though this is not a necessary entailment of the mapping.

Given the different structural attributes each model of hierarchy will map onto a target domain, it follows that each provides a different model for reasoning about the target domain. Nowhere is this clearer than in the first thoroughgoing applications of these models to the domain of music during the early nineteenth century.

## Mapping Models of Hierarchy onto Music

CHAIN-OF-BEING HIERARCHIES AND TONAL THEORY Early in his *Cours complet d'harmonie et de composition* of 1806, Jérôme-Joseph de Momigny presents a C major scale arranged as two symmetrical tetrachords:

G a b C
C d e F

Momigny reminds his readers that the three main notes of the scale are G, C, and F. He goes on to stress the centrality of tonic, providing a telling image in the process: "The tonic triumphs over all the other notes. It is [tonic] who plays the premier role in *le Ton*. It is the center of gravity, the purpose of all purposes, the end of all ends; in a word, it is to [tonic] that the scepter of the musical empire is entrusted."[74] At the beginning of the chapter in which this passage occurs, Momigny linked his concept of *Ton* to that of hierarchy;[75] with his image of tonic holding the scepter, he leaves no doubt as to what sort of hierarchy he has in mind.

74. The complete passage is as follows: "J'ai déjà dit que ce sont là les trois Notes principales du Ton. La TONIQUE l'emporte sur toutes les Notes. C'est elle qui joue le premier rôle dans le Ton. Elle est le centre de gravité, le but de tous le buts, la fin de toutes les fins; en un mot, c'est à elle que le sceptre de l'empire musical est confié." Jérôme-Joseph de Momigny, *Cours complet d'harmonie et de composition, d'après une théorie neuve et générale de la musique* (Paris: Chez l'auteur, 1806), 1: 47. The image is used again on p. 81 of the same volume.

Although the early nineteenth-century notion of *Ton* can be considered to be approximately equivalent to the modern notion of "key," there are also differences between the two concepts; specifically, *Ton* has a dynamic quality lacking in "key." For this reason, I shall not use the latter as a translation of the former.

75. "*Ton* is the hierarchy, the order established between the notes of a genus and a mode." ["Le Ton est la hiérarchie, l'ordre établi entre les Notes d'un genre et d'un mode."] Momigny, *Cours complet d'harmonie et de composition*, 1: 47.

A similar, if less colorful, image may be found in the fuller explanation of order in music that he offered twelve years later in the second music volume of the *Encyclopédie méthodique*. This order, which Momigny claims is purely metaphysical, consists in a natural hierarchy of seven notes that are arranged according to the authority of the one called tonic.[76] As Renate Groth has noted, this is essentially the same definition of order given over two decades later by François-Joseph Fétis; the most important difference is that Fétis assigned it the name *tonalité*.[77]

By mapping a chain-of-being hierarchy onto the domain of music, Momigny creates a very specific notion of *Ton*. The origin of *Ton* is mysterious—Momigny never gets more specific than claiming that it is "metaphysical."[78] The notes organized by *Ton* occupy different levels, according to their directive force; Momigny establishes these levels through various musical propositions. Perhaps most important, notes are regarded as agents with powers proportional to their status within the hierarchy: the properties characteristic of *Ton* increase as one ascends toward tonic. Thus, even though Momigny identifies dominant as the generator of the scale (based on the presence of a minor seventh in the overtone series), he argues that it must be regarded as subservient to tonic, for the authority of tonic cannot be challenged.[79]

The conception of tonal organization produced by mapping a chain-of-being hierarchy onto music was a powerful one. We see it not only in Momigny's early formulation of the concept of *Ton* but also in the persuasive account of tonal organization presented in the first and third volumes of Heinrich Schenker's *Neue musikalische Theorien und Phantasien*. As presented in Schenker's *Harmonielehre* of 1906, the mysterious force that pervades the domain of music is that of nature, manifested through the overtone series. Earlier, in chapter 3, in a discussion of conceptual models as the basis for theories, I quoted a central passage from the *Harmonielehre* on Schenker's organicist approach to the overtone series and pursued its application to his approach to musical analysis. The same passage and related discussion is directly relevant to our present concerns, and we turn to them again here. In Schenker's conception, every tone within the musical domain is possessed of the urge to produce endless overtones, much as living beings have procreative urges.[80] Beginning with a single tone, Schenker creates a community of tones arranged in

76. Jérôme-Joseph de Momigny, "Musique," in *Musique*, Series 36 of *Encyclopédie méthodique, ou par ordre de matières*, ed. Nicolas Étienne Framéry, Pierre Louis Ginguené, and Jérôme-Joseph de Momigny (Paris: Chez Madame veuve Agasse, 1818), 2: 179.

77. Renate Groth, *Die französische Kompositionlehre des 19. Jahrhunderts* (Wiesbaden: Franz Steiner, 1983), 58.

78. On this use of "metaphysical" (which can also be seen in the work of Fétis), see Thomas Christensen, "Fétis and Emerging Tonal Consciousness," in *Music Theory in the Age of Romanticism*, ed. Ian Bent (Cambridge: Cambridge Univesity Press, 1996), 37–56.

79. Momigny, *Cours complet d'harmonie et de composition*, 1: 49–51, 47–48.

80. "It is self-evident that the urge to produce unending generations of overtones belongs to every tone in equal measure. One might also compare this urge to that of animals, for it appears in fact to be in no way inferior to the procreative urge of a living being." ["Es ist selbstverständlich, daß den Trieb, Generationen von Obertönen ins Unendliche zu zeugen, jeder Ton in gleichem Maße besitzt. Man darf, wenn man will, auch diesen Trieb einem Animalischen vergleichen, denn er scheint in der Tat dem Fortpflanzungstrieb eines Lebewesens durchaus nicht nachzustehen."] Heinrich Schenker, *Harmonielehre*, Neue musikalische Theorien und Phantasien, 1 (Stuttgart: J. G. Cotta'sche, 1906), §14, p. 42.

hierarchical order, with the notes closest to the source-pitch having priority over those more remote.[81] The tonal system that results is a manifestation of the desire of the source tone to rule over others and to extend this rule as far as possible.[82]

In *Der freie Satz*, from 1935, Schenker retained this conception of tonal organization and, by emphasizing the organicism latent in this conception, built from it an account of the coherence of entire works of music. He did this by proposing that the transformation of the musical source (now the Naturklang, or chord of nature) proceeded through discrete levels from background to foreground. The background represented an abstract, initial working-out of the tendencies of the chord of nature through the Ursatz, or fundamental structure; the foreground represented the musical work as we conventionally think of it.[83] Schenker makes clear that this foreground emanates from and is controlled by the background.[84] Thus the hierarchical organization of music is not restricted to the ordering of single tones into a scalar system, but manifests itself as a series of transformations that steadily enrich musical content. In Schenker's vision of the relationship between structural levels, a version of the Neoplatonic cosmogony central to the original Great Chain of Being continues to echo: "Between fundamental structure and foreground there is manifested a rapport much like that ever-present, interactional rapport which connects God to creation and creation to God. Fundamental structure and foreground represent, in terms of this rapport, the celestial and the terrestrial in music."[85]

81. Schenker, *Harmonielehre*, §§17–18, pp. 54–55.

82. "But now, what does 'relationship' mean in the life of the tone and what in this context would be implied by the intensity of living life to the fullest? The relationships of a tone are its system. Once the egoism of the tone manifests itself in such a way that the tone (rather like a human) might rule over its fellow-tones rather than be ruled by them, then these very systems provide the means for satisfaction of this egoistic love of power. A tone rules over others when it subjugates them to the relationships outlined in the systems above." ["Nun aber, was soll denn Beziehung heißen im Leben des Tons und was soll hier wohl Intensität des Sichauslebens bedeuten? Beziehungen des Tones sind seine Systeme. Äußert sich der Egoismus des Tones darin, daß er, hierin einem Menschen ähnlich, lieber über seine Mittöne herrscht, als daß er von ihnen beherrscht wird, so sind ihm zur Befriedigung dieser egoistischen Herrschsucht eben in den Systemen die Mittel zur Herrschaft geboten. Ein Ton herrscht über die andern, wenn er sich dieselben nach den in den Systemen geschilderten Verhältnisse unterwirft."] Schenker, *Harmonielehre*, §38, pp. 106–07.

83. Robert Snarrenberg has recently pointed out what I would characterize as competing conceptual models in the usual translations of Schenker's terms (e.g., "fundamental structure" for "Ursatz"); see Snarrenberg, "Competing Myths: The American Abandonment of Schenker's Organicism," in *Theory, Analysis and Meaning in Music*, ed. Anthony Pople (Cambridge: Cambridge University Press, 1994), 45–52.

84. "All the foreground diminutions, including the apparent 'keys' arising out of the voice-leading transformations, ultimately emanate from the diatony in the background." Heinrich Schenker, *Free Composition (Der Freie Satz)*, ed. and trans. Ernst Oster, New Musical Theories and Fantasies, 3 (New York: Schirmer Books, 1979), §4, p. 11. ["Ist doch alle Vordergrund-Diminution, einschließlich der scheinbaren Tonarten aus den Stimmführungsvervandlungen, zuletzt eben aus der Diatonie im Hintergrund erflossen."] Heinrich Schenker, *Der freie Satz*, Neue musikalische Theorien und Phantasien, 3 (Vienna: Universal-Edition A. G., 1935), §4, p. 32.

85. Schenker, *Free Composition*, 160 (Passage E). "Aehnlich wie von Gott zum Geschöpf zu Gott eine Fühlungnahme waltet, stets ineinanderlaufend, stets gegenwärtig, wirkt sich eine Fühlungnahme auch zwischen Ursatz und Vordergrund aus als gleichsam einem Jenseits und Diesseits in der Musik." Schenker, *Der freie Satz*, 18.

ATOMISTIC HIERARCHIES AND METRICAL THEORY A rather different view of musical order is presented by Gottfried Weber in his *Versuch einer geordneten Theorie der Tonkunst*, which appeared in a number of editions in the early nineteenth century. In the course of introducing this particular view, Weber focused not on the organization of pitches but on the organization of musical rhythm. For Weber, rhythm depends on regularity, and the mechanical nature of this regularity is revealed in his breathless definition:

> If . . . a symmetrical measured division of the times is found in a piece of music, *i.e.* the time is distributed into exactly equal general divisions, and these are again divided into equal parts, and the latter are farther separated into equal smaller quotas or proportional parts, &c. and the duration of the tones in relation to one another is exactly measured according to such divisions of time, so that a general division always appears as a symmetrically arranged group of several smaller portions of time, and these taken together as a smaller subordinate group of yet smaller parts, and the accent is also symmetrically apportioned amongst all these divisions of time—then the music is measured and rhythmical.[86]

This mechanical notion of rhythm is accompanied by a mechanical notion of time-keeping: Weber emphasizes the importance of inventions like the metronome and devotes a lengthy footnote to instructions for constructing a pendulum that will accomplish the same thing.[87]

Weber follows his digression into the mechanics of time-keeping with a detailed account of the rules by which musical measures are constructed, first dividing the measure into its constituent parts and then building measures according to symmetrical patterns of accent. He then proceeds to further levels of rhythmic organization:

> But there is still a *higher* symmetry than this. That is to say, as parts of times taken together form small groups, so also can several groups taken together be presented as parts of a larger group, of *a greater* or *a higher rhythm*, of *a rhythm of a higher order.*
>
> We may go still farther, and to such a greater rhythm we may annex moreover a second and a third, so that these two or three together constitute again a still higher rhythm. Thus, *e.g.* in the following passage, two measures taken together constitute a small rhythm, two of these, taken together, constitute again a rhythm of a higher species, and again two of the latter taken together, constitute a capital or principal rhythm.[88]



> The construction of the members of the larger rhythms is a symmetry proceeding more by the great; it is perfectly similar to that involved in the structure of measures, except that it is simply on a larger scale. As a measure consists of two or three parts,

86. Gottfried Weber, *Theory of Musical Composition, Treated with a View to a Naturally Consecutive Arrangement of Topics*, ed. John Bishop, trans. James F. Warner (London: Robert Cocks, 1851), 1: 62 [§47].

87. Weber, *Theory of Musical Composition*, 1: 68–71 [§51].

88. Weber, *Theory of Musical Composition*, 1: 85 [§68].

> so two or three measures form the parts of a greater rhythm, and several such rhythms are again parts of a still higher group.
>
> Hence the measures are distinguished from one another in such higher rhythms, in respect to their greater or less internal weight or accentuation, in the same way the parts of the measure are distinguished among themselves; *i.e.* the heavy or accented *measures* assume a prominence above the lighter, as do the heavier *parts* of the measure above the lighter.[89]

What Weber has described, of course, is an atomistic hierarchy of metrical groups: beats combine into measures, measures combine into groups of measures, and so on up through a series of hierarchical levels.

Mapping an atomistic hierarchy onto the domain of musical rhythm emphasizes the regularity of accentual pattern often associated with the concept of meter. By turning rhythm into a machine, Weber can describe its organization with three basic principles: (1) all units of rhythm are either accented or unaccented; (2) these units are grouped into cyclic patterns of accent that form symmetrical groups; (3) the groups of one level constitute the units of the next higher level. All of the objects within this domain must conform to these principles, but not necessarily in exactly the same way. Although accent or nonaccent applies to both beats and groups of measures, with beats, the property lasts for but an instant, whereas with groups of measures, the property lasts for the duration of the group. No limit is specified for the hierarchy that results, although Weber does not develop levels beyond those shown in his example.

The approach to musical rhythm proposed by Weber became standard during the nineteenth century, although subsequent theorists often added interesting complications. Moritz Hauptmann, for instance, imposed a Hegelian interpretation on metrical groupings, which permitted him to frame the basics of meter and harmony in the same terms.[90] By the time of Riemann's *System der musikalischen Rhythmik und Metrik*, the hierarchical structure represented by Weber's "capital or principal rhythm" was presented as the normative framework (*normatives Grundschema*) of rhythmic organization: beats formed measures, which in turn formed duple groups, which combined into two half-phrases of four measures each, which joined to create the eight-measure phrase.[91]

## Reasoning about Music through Chain-of-Being and Atomistic Hierarchies

Chain-of-being and atomistic hierarchies developed at different points in the history of Western thought to explain two different domains of human experience. These differences are reflected in the basic structure of each type of hierarchy, and they result in two distinct models for reasoning about the world. When these mod-

89. Weber, *Theory of Musical Composition*, 1: 87 [§69].

90. Moritz Hauptmann, *The Nature of Harmony and Metre*, trans. W. E. Heathcote, 2nd ed. (London: Sonnenschein, 1893; repr. with a new forward by Sigmund Levarie, New York: Da Capo Press, 1991), 189–91.

91. Riemann, *System der musikalischen Rhythmik und Metrik*, 196–98.

els are mapped onto music they yield different inferences about how musical materials are organized, which lead to different ways to analyze music. If our reasoning is guided by a chain-of-being hierarchy, we infer that each musical work constitutes a domain pervaded by a mysterious force. Musical analysis consists in explicating the way different musical elements embody this mysterious force. By this means, the proper place of these elements in the musical hierarchy can be determined and an account of musical structure can be produced. If our reasoning is guided by an atomistic hierarchy, we infer that each musical work operates like an extended mechanism. Musical analysis consists in identifying the basic components of this mechanism and explicating the principles according to which it works. The view of musical structure that results is of a hierarchy of distinct yet interrelated substructures that, combined, constitute the musical work.

With these two models for reasoning in mind, let us now return to the three analyses of the rhythmic organization of Mozart's little theme discussed at the opening of this section and consider in more detail how these two models of hierarchy condition and constrain musical analysis.

## Conceptual Models and Music Theory

According to the Schenkerian paradigm Robert Morgan adopts, tonality is characterized by the property of control: within a tonal composition, every pitch controls or is controlled by other pitches. Pitches located at the foreground have only the property of being controlled. Pitches located at middleground levels control pitches at lower levels and are themselves controlled by pitches at higher levels. Pitches located at the background level are governed only by restrictions placed on the Ursatz by the chord of nature.

Morgan's correlation of rhythmic organization with this account of tonal organization has two important entailments for the conception of musical rhythm. First, rhythm becomes a matter of motion between structurally significant events. The initiation and completion of each rhythmic motion is marked by the appearance of the pitches that control a given span; the character of the rhythmic motion will reflect relationships between these controlling pitches. Thus the first rhythmic motion of Mozart's theme begins with the Kopfton in m. 1 and ends with the appearance of $\hat{2}$ in m. 4; the motion has the character of an antecedent because a complete motion from $\hat{3}$ to $\hat{1}$ has not yet been achieved (see ex. 7.5). Second, the type of accent assigned a given musical moment will tend to be unequivocal. Emphasis, as an articulation of rhythmic motion, follows from hierarchical structure: emphasized pitches will be those located at higher levels of the tonal hierarchy. Because the location of a given pitch within the hierarchy is not equivocal, the emphasis it receives will not be equivocal. Thus m. 3 cannot have both an upbeat accent and a downbeat accent at the same time but must be given a rhythmic reading that matches the overall tonal structure.

Although Leonard Meyer's treatment of Mozart's theme certainly takes account of tonal structure, his rhythmic analysis begins, as noted, at the musical surface. The accentual groups of the first (surface) level of the hierarchy become the elements of the second level of the hierarchy, the accentual groups of the second level

become the elements of the third level, and so on up through four levels of hierarchical organization (see ex. 7.4).[92] From Meyer's perspective, rhythmic organization, insofar as it is metrical, is independent of tonal organization, since the elemental patterns of metrical accent can be specified in the absence of pitch. The overall rhythmic character of a passage results from an interaction between tonal organization and patterns of metrical accent occurring on a number of hierarchical levels. The emphasis accorded a given musical event can only be reckoned in context. On level 2 of Meyer's hierarchy, m. 3 is (as a whole) accentually weak; on level 3, it is one component of a strong accent; and on level 4, it is a subcomponent of a weak accent. The property of accent is consequently different on each hierarchical level, since the accent proper to each level pertains to different spans of musical events.

The differences between these two accounts of the rhythmic structure of Mozart's theme follow in part from the way each model of hierarchy constrains reasoning about music. From the perspective of a chain-of-being hierarchy, the musical surface is the least important of all levels. Given the way musical events are assigned to hierarchical levels, any polysemous reading of such events is little more than analytical dithering. From the perspective of an atomistic hierarchy, there can be no grand plan of musical organization that does not originate with the simplest and most readily apprehended of musical materials. Nonetheless, there is a bit more to the story. The constraint these models impose on our reasoning goes beyond the inferential possibilities provided by their relational structure and penetrates to the very core of what counts as relevant phenomena for theorizing and analysis. Indirect evidence for this constraint is supplied by the way each model of hierarchy is mapped onto music. Chain-of-being hierarchies have been mapped almost exclusively onto the domain of musical pitch—they have generally not been mapped onto the domain of musical meter. Atomistic hierarchies, on the other hand, have been most often mapped onto the domain of musical meter—mapping them onto the domain of pitch has not met with wide acceptance.[93]

Explaining how this constraint works requires that we briefly revisit the procedures of cross-domain mapping discussed in chapter 2, with particular attention to the Invariance Principle. According to this principle, for those portions of the source and target domains involved in a cross-domain mapping, the mapping preserves the image-schematic structure of the target domain, and it imports as much of the image-schematic structure from the source domain as is consistent with this preservation. Mapping a model of hierarchy onto music will preserve aspects of our musical experience and structure these by importing the relationships between

92. In the complete analysis Meyer actually adds two additional levels of hierarchy, although the sixth and final level is somewhat speculative; see Meyer, *Explaining Music*, 39.

93. Eugene Narmour's recent attempts to map an atomistic hierarchy onto pitch have met with only limited success, and the application has not been widely adopted. See the in-depth treatment of the issue in Narmour, "Some Major Theoretical Problems Concerning the Concept of Hierarchy in the Analysis of Tonal Music," *Music Perception* 1/2 (1983–84): 129–99; Narmour discusses mm. 1–8 of the first movement of Mozart's Piano Sonata in A, K. 331, on pp. 160–98. In some recent analyses, Narmour reads musical structure as only quasi-hierarchical and thus better dealt with in terms of networks. See Narmour, *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model* (Chicago: University of Chicago Press, 1990).

"energy" and the CONTAINER schema embodied by this model. When we take a chain-of-being hierarchy as the source domain, we preserve our sense that some aspects of music are more important than others, and we import a model of systemic organization in which energy is distributed unequally and hierarchical levels group elements according to their amount of energy. This mapping works well for the asymmetrical pitch relations typical of tonal music but less well for the measured domain of metrical rhythm. When we take an atomistic hierarchy as the source domain, we preserve our sense that some aspects of music are regular and recurring, and we import a model of systemic organization in which energy is distributed equally and hierarchical levels group elements according to common conformance with general principles. This mapping works well for the measured domain of metrical rhythm but less well for the asymmetrical pitch relations typical of tonal music. Thus the image schemata that ground our conceptual models also facilitate mapping these models onto novel domains, subject to the limitations of the image-schematic structure of both domains.

Additional evidence for the part played by image schemata in conceptual models and cross-domain mapping is suggested by Edward Cone's analysis of Mozart's theme, which combines aspects of the approaches demonstrated by Morgan's and Meyer's analyses. For Cone, as for Morgan, musical rhythm involves motion between points of tonal stability. The first and fifth measures of the theme should be strong because of the firm statement of tonic; the fourth and eighth measures should also be strong because of the emphasis provided by arrivals on important harmonies. However, Cone, like Meyer, is also interested in the contribution of the musical surface to our understanding of musical rhythm as a whole. Cone's overall interpretation of the rhythmic structure of the opening eight measures is thus shaped by meter, motivic structure, and surface accents. What guides Cone's analysis is his concept of musical energy—it was this that prompted his comparison of a musical phrase to a game of catch. The path of energy that marks the transit of the ball in the game of catch is the analogue for the path of musical energy constituted by a given passage in a work of music. Starting with the concept of energy rather than with a model of hierarchy allows Cone to tap into both models of hierarchy and draw from each what he needs. By this means, he can forge a single account of musical rhythm that recognizes both hierarchic paradigms but which is restricted to neither. Thus, Cone's analysis demonstrates a conceptual blend, the generic space of which focuses on the notion of energy common to both models of hierarchy.

According to the approach I have developed here, the discrepancies between these three analyses of Mozart's theme reflect the role of conceptual models in theories of music. At work are two different models of hierarchical organization. One maps a chain-of-being hierarchy onto music and highlights one set of aspects of the musical domain. The other maps an atomistic hierarchy onto music and highlights a different set of musical aspects. Two of the analyses—those by Robert Morgan and Leonard Meyer—rely on relatively straightforward mappings, one from each model of hierarchy. Where the structure of the models of hierarchy diverge, there the musical analyses based on these models will also diverge. The third analysis, by Edward Cone, blends the two mappings by focusing on commonalities between their image-schematic structure. By focusing on these commonalities, Cone puts

some distance between his interpretation and notions about hierarchical organization and produces yet a third analysis of Mozart's theme. The discrepancies between the three analyses are thus not the result of different opinions about what constitutes the theme of the first movement of Mozart's sonata, nor are they a consequence of analytical inconsistencies. Rather, they follow from the structure of these two conceptual models and the way they are employed in accounts of music.

## COMPETING MODELS OF MUSICAL FORM AND HIERARCHY

The foregoing represents an attempt to sort out why musicians—reasonable, highly experienced musicians—disagree about such things as musical form and musical hierarchy. As I have endeavored to show, such disagreements can be traced to competing models of musical structure. Much as Alice and the Hatter have different ideas about what time is, so music theorists have different ideas about the properties of musical form and the structure of hierarchical relationships. The analyses I offer should help explain why this is so. If one is thinking about musical form in terms of grammar or building blocks, it may not be immediately evident how music achieves its rhetorical effects or seems to grow from within. If one is thinking about musical relationships in terms of dominance and control, musical meter may appear to be something of a cipher, for meter provides only weak demonstrations of these properties.

More interesting, however, is the way these models reflect back on the theorists who employ them, much as Alice's models of time reflect the world of a seven-year-old girl in Victorian England, while the Hatter's model of time reflects the world of Wonderland. Theorists who focus on a static model of musical form are almost invariably occupied with a very basic level of music pedagogy. Dynamic models only start to make real sense once a level of artistry has been achieved—that is, once the student can speak and understand the language of music with some fluency. Not surprisingly, then, the two models begin to overlap and intertwine in confusing and interesting ways in the early nineteenth century, for the institutionalization of music education required a bit of both perspectives. Nation building required the mechanisms of rationalized instruction and of artistic exegesis, the first to guarantee a future, the second to celebrate a past. Similarly, theorists who focus on atomistic hierarchies are convinced that we enter music from its surface and only later can penetrate to its depths by understanding the general laws that connect them both. Theorists who make recourse to chain-of-being hierarchies, on the other hand, are convinced that surfaces are illusions—the real truth of music lies hidden in mysterious powers that guide compositional creation like an unseen hand. These summaries are too reductive, of course, but they make in brief the point toward which this chapter has argued at length: the models we choose for our characterizations of music tell us at least as much about ourselves as they do about music.

# COGNITIVE STRUCTURE, THEORY, AND ANALYSIS

Let us return to Proust's *A la recherche du temps perdu* and to the encounters between listener and music described within. The introduction to this volume was framed around the moment when Charles Swann first hears the Andante of a sonata for violin and piano by Vinteuil at the bourgeois salon of the Verduins. As that moment unfurled, Swann moved from vague and unformed impressions to active understanding: he began to conceptualize music. Now, a year later, Swann again encounters the same fictitious composition, this time at a soirée at the Marquise de Saint-Euverte's. Over the course of this year, Swann has become obsessed both with the Andante—especially as represented through the fragment he calls "the little phrase"—and with Odette, the courtesan whom he had also first come to know at the Verduins. At the moment we find him, Odette has left him, and he finds himself surrounded by people who, if they knew of his love for her, would only trivialize it. Then, unexpectedly, the musicians commence the sonata, and Swann finds himself enraptured once more. Where in his first encounter the sonata had amounted to not much more than pleasant sensations, it is now a full-fledged discourse, speaking in a beautiful and paradoxical language: "Ever since, more than a year before . . . the love of music had, for a time at least, been born in him, Swann had regarded musical *motifs* as actual ideas, of another world, of another order, ideas veiled in shadow, unknown, impenetrable to the human mind, but none the less perfectly distinct from one another, unequal among themselves in value and significance."[1] The ideas that enthrall Swann—distinct yet shadowy, perfectly clear yet impenetrable—are musical concepts; the discourse to which he attends is musical discourse. What began as a sketchy, provisional transcript has blossomed into a theory of musical organization that is robust enough to reach out to the transcendental.

The process that led to this transformation (as described in chap. 1) begins with categories of musical events, of which the *motifs* (or motives) upon which Swann focuses are a good example. Musical motives show a number of similarities to basic-

1. Marcel Proust, *Swann's Way* (vol. 1 of *Remembrance of Things Past*), trans. C. K. Scott Moncrieff and Terence Kilmartin (New York: Vintage Books, 1981), 379–80.

level categories, offering an effective compromise between efficiency and informativeness. The variability of motives, as they are shaped to the exigencies of musical context and compositional process, matches perfectly the variability of Type 1 categories; our impression that there is a motive reflects the conceptual models around which such categories are organized. Of course, motives of the sort typical in late nineteenth-century music, of which the little phrase is intended to be a paragon, represent only one possibility for musical categories. Categories might also be organized around a harmonic structure or a ground bass, around a two-voice contrapuntal pattern or a persistent rhythmic figure. As the analysis of atonal music has suggested, musical categories might even involve relatively abstract relationships, which are obscure at first but which emerge with repeated listenings.[2]

As I proposed in chapter 1, we can construe musical concepts as equivalent to musical categories. Although such concepts may be affiliated with language, language is not necessary for their creation or maintenance. The paradox presented by Proust—that musical ideas are "ideas veiled in shadow, impenetrable to the human mind, but none the less perfectly distinct from one another"—is simply a reflection of the limits of language. Musical concepts are of another world, another order, because they extend into a domain that is beyond words.

One of the principal ways musical concepts are related to other concepts—musical and otherwise—is through the process of cross-domain mapping. A bit before the passage quoted above, the little phrase is first a kind of superhuman agent transformed into sound (whose presence is itself likened to the perfume of a flower), then the words this spirit speaks, then an almost corporeal lover, and finally once again a comforting spirit:

> Swann felt its presence like that of a protective goddess, a confidante of his love, who, in order to be able to come to him through the crowd and to draw him aside to speak to him, had disguised herself in this sweeping cloak of sound. And as she passed, light, soothing, murmurous as the perfume of a flower, telling him what she had to say, every word of which he closely scanned, regretful to see them fly away so fast, he made involuntarily with his lips the motion of kissing, as it went by him, the harmonious, fleeting form. He felt that he was no longer in exile and alone since she, who addressed herself to him, was whispering to him of Odette.[3]

This dense field of overlapped mappings cannot, of course, quite summon the acoustic reality of the little phrase (which exists, after all, only in Proust's imagination), but it does select certain aspects of the musical domain as targets. The little phrase is distinct enough from the surrounding materials that it can be likened to a person emerging from a crowd; it has a measure of delicacy about it, both in its

2. I am thinking especially here of the work of David Lewin, who has a knack for discovering intriguing musical relationships, characterizing them through impressively abstract formulations, and managing to create analyses of compelling immediacy. On the possibility of hearing these formulations, Lewin comments, "One should not ask of a theory, that every formally true statement it can make about musical events be a perception-statement. One can only demand that a preponderance of its true statements be potentially meaningful in sufficiently developed and extended perceptual contexts." Lewin, *Generalized Musical Intervals and Transformations* (New Haven, Conn.:Yale University Press, 1987), 87.

3. Proust, *Swann's Way*, 378.

overall aesthetic shape and in its lack of assertiveness; despite this delicacy, the substance of its communicative force is such that it evokes response not simply in words but in deeply intimate gestures. If, after this description, the little phrase still remains elusive, we can nonetheless generate a fairly clear sonic picture of what it involves. Proust's mappings have thus constructed an image of this imaginary music, selecting musical concepts—vague and ill formed, to be sure, but convincing nonetheless—from the general field of possible musical constructs.

It is through similar, if more prosaic, mappings that we structure our understanding of music, building up the integrated systems of terms and relations through which we describe our experience with and conceptions of music. Such cross-domain mappings have their ultimate grounding in embodied experience (as we observed in chap. 2). In turn, our experiences with music allow us, on occasion, to map from the musical domain onto other domains. For instance, in the passage just cited, music (as represented by Vinteuil's sonata) provides a model according to which a succession of fleeting impressions—the movement of a person's body, the scent of a flower, the brush of a kiss, a scattering of forms—can be construed as constituting an indivisible whole. Just as the sonic materials proper to a musical work appear and disappear and yet always remain part of "the work," so the fleeting impressions noted by Swann are, in truth, but some of the disparate threads out of which his love for Odette is woven.[4]

The possibility of mapping structure both to and from the musical domain contributes to the process of conceptual blending, in which elements from two correlated mental spaces are blended together in a third mental space. This offers a further interpretation of Swann's encounter with the little phrase, for what appears before him is a blend of features drawn from the human and sonic domains—a protective goddess who has disguised herself in a sweeping cloak of sound. Just such a blend seems to have been intended by Proust, who goes on to write that Swann "had no longer, as of old, the impression that Odette and he were unknown to the little phrase. Had it not often been the witness of their joys? True that, as often, it had warned him of their frailty."[5] As the story continues, the little phrase becomes fully personified, engaging in dialogue with Swann, offering sage advice, and leading him to transports of the spirit. What Proust is relying on, of course, is the sort of anthropomorphic blend he also employed in his description of Swann's lilac trees, and which Tovey employed so strategically in his analysis of Beethoven's Sixth Symphony (both discussed in this volume in the latter portion of chap. 2).

The musical concepts produced by categorization and the structural relationships produced by cross-domain mapping come together in conceptual models, which provide the basis for reasoning about music. Swann relies on two models. Of these, the first yields a view of music as a type of discourse structure: musical motives are regarded as ideas that are shaped into a communicative medium by the way they are organized in time and with respect to other musical ideas. The sec-

4. This is particularly apparent because Proust's allusion to "the perfume of a flower" is anticipated a few pages earlier by Swann's detailed recollection of the chrysanthemum Odette gave him early in their romance.

5. Proust, *Swann's Way*, 378.

ond model yields a view of the nonlinguistic as supernatural. Words, as proper to human discourse, are necessarily mortal; those things that escape words are consequently immortal; things that are immortal exist only outside the natural world. These models are then combined into a theory of music appropriate to Swann's moment of clairvoyance, following the perspective developed in chapter 3. According to this theory, musical discourse is a suprahuman mode of communication, something evident as the two musicians conclude the performance that had summoned the little phrase and Swann's rapture.

> Swann dared not move, and would have liked to compel all the other people in the room to remain still also, as if the slightest movement might imperil the magic presence, supernatural, delicious, frail, that was so soon to vanish. But no one, as it happened, dreamed of speaking. The ineffable utterance of one solitary man . . . breathed out above the rites of those two hierophants, sufficed to arrest the attention of three hundred minds, and made of that platform on which a soul was thus called into being one of the noblest altars on which a supernatural ceremony could be performed.[6]

This is a theory of music rather different from the others I have discussed, not the least because where specifics are wanted we find only sweeping generalities. Nonetheless, the theory of music that Swann has developed in his efforts to understand the effect of Vinteuil's sonata is in truth but a manifestation of a portion of the broader theory of aesthetics Proust sought to develop in *A la recherche du temps perdu.* Within this theory, music was to occupy the same place accorded it by Schopenhauer: as a pure expression of the will.[7] Necessary to this perspective is the conceit that absolute music somehow exists outside the domain of human cognition.

The perspective relative to which I have framed this book, of course, is that music, "absolute" or otherwise, is wholly within the domain of human cognition. In making this argument, I have focused not on the most basic aspects of musical cognition—the level where raw data is taken in by the auditory processing system and transformed into electro-chemical information within the brain—but on a relatively high level of cognition accessible to our conscious thoughts. My emphasis has been on how general cognitive capacities (like those for categorization, cross-domain mapping, and the construction and maintenance of conceptual models and theories) are specified for music. Indeed, understanding something as music—as opposed to the hammering of a woodpecker, or an agreeable if not particularly interesting succession of sounds, or just noise—is intimately bound up with directing these general cognitive capacities to the specific task of making sense of patterned sound.

This approach to music cognition makes it possible to give a fairly detailed account of how M. Swann (and M. Proust, for that matter) conceptualize music. It also provides a perspective on broader problems of musical understanding, four of which I treated at length in the second part of this book.

6. Proust, *Swann's Way*, 383.

7. On this point, see Jean-Jacques Nattiez, *Proust as Musician*, trans. Derrick Puffett (Cambridge: Cambridge University Press, 1989), chap. 4.

MUSICAL SYNTAX The idea that musical works could be woven from relatively limited bits of musical materials—motives—was one that gained currency around the time that composers became increasingly interested in the resources of purely instrumental music—that is, around the middle of the eighteenth century. As such, the idea that motives are important for musical discourse is hardly a new one, even if it did not assume prominence in music theory treatises until the mid-nineteenth century. Viewing a motive as a cognitive category has the practical benefit of explaining why we construe diverse motive forms as "the same," despite their manifold differences. The perspective becomes really interesting, however, when we start to consider relationships among the members of a category of motives, for these offer a way to explain the way the musical syntax specific to a given work emerges over time. As the composer deploys now more, now less typical members of the category, we not only develop a more complete sense of the structure of the category but also come to see how typicality correlates with overall compositional design. This syntactic strand could be woven together with those spun from harmonic and formal syntax to create a rich, multileveled discourse (as we saw in the analyses of works from the eighteenth century presented in chap. 4).

The same approach also provides a way to explain how the musical topics so important in eighteenth-century instrumental music emerged. As we saw in the first movements of two Beethoven string quartets, categories of motive forms were distinguished by how they were used in the course of a movement. In the case of a motivic category used only for beginnings, for instance, this property then becomes part of the conceptual model for the category of motive forms: in addition to its other attributes, the category now acquired the attribute "used only for beginnings." This change has the consequence of giving the category a semantic value: within the context of this movement, the motive means "this is a beginning." As a result, a musical topic, even if of a very local and restricted sort, has been born. For eighteenth-century music, this process is of interest chiefly for answering the riddle of how musical topics were created in the first place and for understanding the interplay between locally generated signs and established topics. For music with a less-clearly defined tradition, such as that of the twentieth century, the process is of rather more interest, for it offers a way to explain how each individual piece generates a semiotic system. Indeed, it is in part through establishing and manipulating categories of musical events that composers of the twentieth century create musical meaning. We can also see the part this process plays in helping us understand an unfamiliar repertoire: by attending to the specifics of musical syntax, we can begin to understand the process of meaning construction of which this syntax is a part. In a manner of speaking, music teaches us how we should listen to it.

The dynamic approach to musical syntax made possible by considering relationships among the members of musical categories (and among categories themselves) is one that is well suited to developing an account of large-scale rhythm—that is, rhythm at the phrase level or above. The challenge of theorizing large-scale rhythm has been to develop a model that is more than simply an adaptation of metric theory or a reinterpretation of large-scale harmonic theory. One way to develop such a model is to construe relationships among category members in dynamic terms. For instance, if we return to the compositional strategies discussed in chapter

4, we could use cross-domain mapping to correlate the typicality of category members with stability, and atypicality with instability. Applied to the opening measures of the Allegro from Mozart's "Dissonance" Quartet, what emerges is a large-scale pattern that begins with stability (in mm. 23–26), moves to instability (beginning in m. 27), returns to stability (in mm. 31–34), moves somewhat more quickly to instability (beginning in m. 35), returns (after a brief hiatus) to stability (in m. 44), and then finally dissolves into instability (after m. 48).[8] The pattern could be further applied to each subsequent episode involving the motive, producing a dynamic account of the movement as a whole.

MUSICAL ONTOLOGY That an exploration of musical ontology might yield significant insights into the structure and character of cultural knowledge might at first seem surprising, for as common as music is, it is still not the central business of human culture. However, because music is resolutely ephemeral, knowledge of it is, in most cases, preserved only in human cognitive structure. We can know what Greek pottery of the fifth century B.C. was like, because some of it has been preserved into the present era. The same cannot be said of Greek music. Theoretical and other writings give us hints about the character of Greek music, but direct experience of it is impossible. Music, where it is not preserved through some artifactual means, exists solely in the conceptual models and theories of the members of a musical community. In consequence, giving an account of the basis for determinations of musical identity—what it means for something to be considered an instance of music—means giving an account of the cultural knowledge upon which such determinations are based.

Although it is easy to hold up the Western preoccupation with "the score" as the acme of musical objectification, the notion of a musical object is not a consequence of artifacts like scores but is instead an offshoot of the conceptual models we use to structure our understanding of what we hear. Because these models are, for the most part, completely transparent, we take them for the way things are: conceptual models are part and parcel of what counts as reality (as noted in chap. 3). Our impression that "music" is an objective category, and that there are musical objects, is simply a consequence of the cognitive efficiency conceptual models represent.

In that the conceptual model relative to which we make determinations of musical ontology captures the things our culture deems most important about patterned sound (or the place patterned sound occupies in our culture), our theories of music are consequent to these models: what counts as "music" determines what counts as "music theory." One prominent example of this was provided by the Greek music theories of antiquity discussed in the introduction. However, more subtle examples are available as well: for Rameau, "music" meant French music, and it was relative to that national tradition of musical expression that his theory was framed. Similarly,

8. The changing pattern of stability and instability I have outlined very much resembles the pattern that occurs at the level of the beat, specifically in the context of what Harald Krebs calls metric dissonance. The principal difference is that the pattern as I have described it is expanded to encompass entire musical phrases. This theory of metric dissonance is most fully worked out in Krebs, *Fantasy Pieces: Metrical Dissonance in the Music of Robert Schumann* (New York: Oxford University Press, 1999).

for Schenker "music" meant music from the German-speaking world, and it was to this repertoire that his theory was directed. This conception of musical ontology can be extended to other repertoires, like jazz (as noted in chap. 5). There we observed that certain jazz musicians of the 1950s and 1960s regarded the harmonic possibilities of a tune (that is, the potential for the harmonic template of the tune to be reconstituted through the modification or substitution of harmonies) and its overall affect as more important than the bare pitch structure of its melody or the basic syntax established by its harmonies. In consequence, the theory of music that issued from this practice—a theory woven throughout the interviews and quotations that form the fabric of Paul Berliner's *Thinking in Jazz*—was centrally concerned with ways to create and layer the multiple musical discourses made possible by a particular harmonic pattern.[9] For this music, it was far less important to develop an explanation of the imperatives of melodic or harmonic syntax, since these were assumed to be almost totally dependent on local context.

MUSIC AND OTHER DOMAINS According to at least some definitions, language and music are quite separate domains. The distinction is eminently sensible: when we want to know if supper is ready, we expect a response in words, not a musical phrase ending with a deceptive cadence. Nonetheless, inasmuch as both language and music are products of human cognitive processes, it also seems sensible that they will rely on some of the same cognitive capacities. For example, evidence from conceptual blending suggests that they both employ similar discourse strategies. Admittedly, these strategies, as represented by the generic space of a conceptual integration network, are rather abstract, so elemental as to be almost simplistic. And yet there is strong evidence that such simple stories are basic to human cognition and communication—indeed, as we observed in chapter 2, Antonio Damasio has argued that they are involved with consciousness itself.[10] In situations where linguistic discourse and musical discourse are simultaneous and structured by the same basic strategy, concepts from the mental spaces set up by each can combine in a third mental space: a conceptual blend. Situations of this type were demonstrated by the analyses of songs in chapter 6. And, as we saw with Brahms's "In der Fremde," if a uniform topography between the constituent spaces in a conceptual integration network is lacking, conceptual blending will not occur. The combination of words and music in such cases may be pleasant enough, but it will not realize the full potential of these conjoined media.

A basic conceptual blend involves four coordinated mental spaces (as noted in chaps. 2 and 6). However, as Tovey's analysis of Beethoven's Sixth Symphony shows, conceptual blending may involve any number of mental spaces. It is fairly easy to imagine songs that might involve multiple spaces—for instance, if the song makes

9. Paul F. Berliner, *Thinking in Jazz: The Infinite Art of Improvisation*, Chicago Studies in Ethnomusicology (Chicago: University of Chicago Press, 1994).

10. For an overall account of the importance of abstract discourse structures to human cognition, see Mark Turner, *The Literary Mind* (New York: Oxford University Press, 1996). The approach Antonio Damasio takes in his *The Feeling of What Happens: Body and Emotion in the Making of Consciousness* (New York: Harcourt Brace, 1999) is notable for the empirical evidence and level of neurophysiological detail that he brings to bear.

reference to a well-known myth, or if, as in the case of Schubert's "Trockne Blumen," there are strong parallels between the situation described in the words and details of the composer's life.[11] Even more intriguing are the possibilities offered by opera, which can easily engage—and often revels in—simultaneous musical, linguistic, theatrical, gestural, and scenic discourse. Various kinds of musical multimedia—music and film, music and dance, music and visual images—offer further possibilities for exploring conceptual blends involving multiple spaces. In each case, where a uniform topography between the spaces set up by these media obtain, conceptual blending can occur. What is potentially intriguing about the blends that can be set up by the coordination of these diverse media are the views of discourse structure that will result, so that the syntax of music will shed light on the syntax of other media, and they on music. While differences between the media are essential to the experience of multimedia, as Nicholas Cook has suggested, the deep commonalities between the discourse structure of various media offer new ways to think about musical structure.

COMPETING MODELS OF MUSICAL STRUCTURE The theory of music developed by Charles Swann, while useful for making a reconciliation with what he imagined to be the end of his love affair, offers hardly any explanation for why the little phrase works the magic that it does or how Vinteuil's sonata can hold three hundred Parisian socialites enthralled. Where we long to discover more about its syntax, or its place within a tradition of similar works, or the means by which it sprang forth from the mind of its creator, Swann's theory gives us only obscurantism and effusion. That this should be so is a direct consequence of the conceptual models his theory employs. Given the assumption that music is a type of discourse structure (rather than, say, some sort of organism or mechanism), Swann's theory places an emphasis on the message rather than the medium and on temporal constraint as an essential element of musical experience. Music has things to tell us, but the disappearance of these messages is practically coincident with their emergence. Given the assumption that nonlinguistic discourse is supernatural (rather than something humans make use of regularly), we cannot expect to make much sense of the things music tells us; we can only be grateful that we have been able to hear them at all. Of course, Swann's—and Proust's—theory of music is a product of a particular cultural and historical moment, one that sought to explain how humans could participate in the divine even while they were forever shut off from it. Other cultural and historical moments—including our own—will seek other explanations, employing the categories, cross-domain mappings, conceptual models, and theories that resonate with them most deeply and constructing a vision of music that best fits with their vision of the world as a whole.

Theories of music, then, are not so much about truth as about a particular perspective on what aspects of patterned sound are rewarding or important to attend to; they are equally concerned with the integration of this perspective into a larger

11. Although this biographical correlation does not figure in my consideration in chap. 6 of Schubert's song, I touch on these parallels in Lawrence Zbikowski, "The Blossoms of 'Trockne Blumen': Music and Text in the Early Nineteenth Century," *Music Analysis* 18 (1999): 340.

view of human society and culture. Similarly, music does not offer us an unmediated vision of the truth—despite what M. Swann believes—but a truth assembled from the possibilities of patterned sound and human cognitive capacities.

Most of my discussion has focused on cognitive structure and theory, and I have had relatively less to say about analysis—especially analysis as the object of a separate domain of inquiry. The reason for this apparent imbalance, as mentioned in the introduction to this volume, is that while analysis is essential for making claims about musical structure explicit, it is not independent of theory: every analysis relies on some theory of music.[12] For that matter, I am somewhat doubtful about making a meaningful distinction between description (as a more or less passive reporting of the relevant features of a work) and analysis (as a thorough contextualization of aspects of musical structure), since decisions about what counts as "relevant features" and "musical structure" rely equally on some theoretical understanding about music. It seems more promising to distinguish between ad hoc theories and extended theories, as I did at the end of chapter 3. This distinction is based on the kind of problems a theory engages, its place within a tradition of inquiry, and the degree to which it is rendered systematic. Generally speaking, descriptions or analyses based on extended theories are more interesting than those based on ad hoc theories, for extended theories draw much more deeply on our cognitive capacities and cultural knowledge.

Although analysis has been a bit of a dark horse, it has hardly been out of the running. Indeed, it has been present throughout the course traced by the preceding pages: in the analyses by Schoenberg, Tovey, Rameau, Schenker, Mattheson, Meyer, Cone, and Morgan I have discussed, and in the analyses I myself have presented. In the case of earlier writers, I have been most interested in what their analyses have to say about their theories of music. My own analyses, of course, are based on the theory that our understanding of music is shaped by our cognitive capacities and that these are visible at every turn in our encounters with music. Through our theories and analyses, we thus shed light not just on music but on the fascinating mysteries of human thought.

The abiding mystery I have sought to explain is the one embodied in M. Swann's first encounter with Vinteuil's sonata: how it is that we can make some sort of sense of an unfamiliar music almost immediately and yet have this same work provide continued reward in our further experiences with it. Music, as revealed by Proust's vignette, is thus simultaneously immediate and complex, transparent and opaque. The solution I have offered is that our understanding of music relies on cognitive capacities that we use throughout our daily lives. The cognitive processes behind categorization, cross-domain mapping, conceptual models, and theories would all be operational had we never heard a note of music. Doubtless, our lives would be immeasurably impoverished had this unlikely turn of events transpired,

12. Among the first to draw this distinction was Edward T. Cone in "Analysis Today," *Musical Quarterly* 46 (1960): 172–88. This essay can also be found in Paul Henry Lang (ed.), *Problems of Music Analysis: The Princeton Seminar in Advanced Musical Studies* (New York: W. W. Norton, 1960), 34–50; and Cone, *Music, a View from Delft: Selected Essays* (Chicago: University of Chicago Press, 1989), 39–54.

but not only because we would have never been exposed to what Proust characterized as "another world, another order, ideas veiled in shadow, impenetrable to the human mind, but none the less perfectly distinct from one another." We would also have never had these cognitive capacities stretched and transformed by the challenge of listening to the play of patterned sound. Nor would we have taken what we learned in our encounters with that domain back to our understanding of the world as a whole. And we would never have had the opportunity to extend our conceptualization of music into theories of music—theories that, in the webs of relations they build both within and without music, offer their own topic for intellectual investigation. Conceptualizing music, seen in this way, is thus both a celebration and an exploration of the means by which our minds go out to meet the world.

# BIBLIOGRAPHY

*Books and Articles*


Abelson, Robert P., and John B. Black. "Introduction." In *Knowledge Structures*, edited by James A. Galambos, Robert P. Abelson, and John B. Black. Hillsdale, N.J.: Erlbaum Associates, 1986.

Abrams, Meyer Howard. *The Mirror and the Lamp: Romantic Theory and the Critical Tradition*. Oxford: Oxford University Press, 1953.

Agawu, Victor Kofi. *Playing with Signs: A Semiotic Interpretation of Classical Music*. Princeton, N.J.: Princeton University Press, 1991.

———. "Theory and Practice in the Analysis of the Nineteenth-Century *Lied*." *Music Analysis* 11 (1992): 3–36.

Aldwell, Edward, and Carl Schachter. *Harmony and Voice Leading*. 2nd ed. San Diego: Harcourt Brace Jovanovich, 1989.

Ames, David W., and Anthony V. King. *Glossary of Hausa Music and Its Social Contexts*. Evanston, Ill.: Northwestern University Press, 1971.

Aristotle. *On the Sublime*. Transated by W. Hamilton Fyfe. Loeb Classical Library, 199. Cambridge, Mass.: Harvard University Press, 1995.

Aristoxenus. *Elementa harmonica*. In *Harmonic and Acoustic Theory*. Vol. 2 of *Greek Musical Writings*. Edited and translated by Andrew Barker, 126–84. Cambridge Readings in the Literature of Music. Cambridge: Cambridge University Press, 1989.

Armstrong, Sharon Lee, Lila R. Gleitman, and Henry Gleitman. "What Some Concepts Might Not Be." *Cognition* 13 (1983): 263–308.

Arom, Simha. *African Polyphony and Polyrhythm: Musical Structure and Methodology*. Translated by Martin Thom, Barbara Tuckett, and Raymond Bond. Cambridge: Cambridge University Press, 1991.

———. "New Perspectives for the Description of Orally Transmitted Music." *World of Music* 23 (1981): 40–58.

Attali, Jacques. *Noise: The Political Economy of Music*. Translated by Brian Massumi, with a foreword by Fredric Jameson, with an afterword by Susan McClary. Theory and History of Literature, 16. Minneapolis: University of Minnesota Press, 1985.

Bailey, Robert, ed. *Prelude and Transfiguration from Tristan and Isolde*. New York: W. W. Norton, 1985.

Bakan, Michael B. *Music of Death and New Creation: Experiences in the World of Balinese Gamelan Beleganjur*. Chicago Studies in Ethnomusicology. Chicago: University of Chicago Press, 1999.

Baker, Nancy Kovaleff, and Thomas Christensen, eds. and trans. *Aesthetics and the Art of Musical Composition in the German Enlightenment: Selected Writings of Johann Georg Sulzer and Heinrich Christoph Koch*. Cambridge Studies in Music Theory and Analysis, 7. Cambridge: Cambridge University Press, 1995.

Bakhtin, Mikhail Mikhailovich. *Problems of Dostoevsky's Poetics*. Edited and translated by Caryl Emerson, with an introduction by Wayne C. Booth. Theory and History of Literature, 8. Minneapolis: University of Minnesota Press, 1984.

Bamberger, Jeanne. "Cognitive Issues in the Development of Musically Gifted Children." In *Conceptions of Giftedness*, edited by Robert J. Sternberg and Janet E. Davidson, 388–413. Cambridge: Cambridge University Press, 1986.

———. *The Mind behind the Musical Ear: How Children Develop Musical Intelligence*. Cambridge, Mass.: Harvard University Press, 1991.

Barbera, C. André. "The Consonant Eleventh and the Expansion of the Musical Tetractys: A Study of Ancient Pythagoreanism." *Journal of Music Theory* 28 (1984): 191–223.

Barchas, Janine. "The Engraved Score in *Clarissa*: An Intersection of Music, Narrative, and Graphic Design." *Eighteenth-Century Life* 20 (1996): 1–20.

Barker, Andrew, ed. *Harmonic and Acoustic Theory*. Vol. 2 of *Greek Musical Writings*. Cambridge Readings in the Literature of Music. Cambridge: Cambridge University Press, 1989.

Barsalou, Lawrence W. *Cognitive Psychology: An Overview for Cognitive Scientists*. Hillsdale, N.J.: Erlbaum Associates, 1992.

———. "Deriving Categories to Achieve Goals." *The Psychology of Learning and Motivation, Advances in Research and Theory* 27 (1991): 1–64.

———. "Flexibility, Structure, and Linguistic Vagary in Concepts: Manifestations of a Compositional System of Perceptual Symbols." In *Theories of Memory*, edited by A. F. Collins, S. E. Gathercole, M. A. Conway, and P. E. Morris, 29–101. Hillsdale, N.J.: Erlbaum Associates, 1993.

———. "Frames, Concepts, and Conceptual Fields." In *Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization*, edited by Adrienne Lehrer and Eva Feder Kittay, 21–74. Hillsdale, N.J.: Erlbaum Associates, 1992.

———. "The Instability of Graded Structure: Implications for the Nature of Concepts." In *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*, edited by Ulric Neisser, 101–40. Cambridge: Cambridge University Press, 1987.

———. "Intraconcept Similarity and Its Implications for Interconcept Similarity." In *Similarity and Analogical Reasoning*, edited by S. Vosniadou and A. Ortony, 76–121. Cambridge: Cambridge University Press, 1989.

———. "Perceptual Symbol Systems." *Behavioral and Brain Sciences* 22 (1999): 577–609.

Barsalou, Lawrence W., Karen Olseth Solomon, and Ling-Ling Wu. "Perceptual Simulation in Conceptual Tasks." In *Cultural, Psychological, and Typological Issues in Cognitive Linguistics: Selected Papers of the Bi-Annual ICLA Meeting in Albuquerque, July 1995*, edited by Masako K. Hiraga, Chris Sinha, and Sherman Wilcox, 209–28. Amsterdam: John Benjamins, 1999.

Barsalou, Lawrence W., Wenchi Yeh, Barbara J. Luka, Karen L. Olseth, Kelly S. Mix, and Ling-Ling Wu. "Concepts and Meaning." In *Chicago Linguistics Society 29: Papers from the Parasession on the Correspondence of Conceptual, Semantic, and Grammatical Representations*, edited by Katharine Beals, Gina Cooke, David Kathman, Sotaro Kita, Karl-Erik McCullough, and David Testen. Chicago: University of Chicago, Chicago Linguistics Society, 1993: 23–61.

Bayard, Samuel P. "Prolegomena to a Study of the Principal Melodic Families of British-American Folk Song." *Journal of American Folklore* 63 (1950): 1–44.

Belzner, William. "Music, Modernization, and Westernization among the Macuma Shuar." In

*Cultural Transformations and Ethnicity in Modern Ecuador*, edited by Norman E. Whitten, Jr., 731–48. Urbana: University of Illinois Press, 1981.

Bent, Ian. "The 'Compositional Process' in Music Theory, 1713–1850." *Music Analysis* 3 (1984): 29–55.

———, ed. *Fugue, Form and Style*. Vol. 1 of *Music Analysis in the Nineteenth Century*. Cambridge Readings in the Literature of Music. Cambridge: Cambridge University Press, 1994.

Berger, Karol. *Musica Ficta: Theories of Accidental Inflection in Vocal Polyphony from Marchetto da Padova to Gioseffo Zarlino*. Cambridge: Cambridge University Press, 1987.

Berlin, Brent. "Ethnobotanical Classification." In *Cognition and Categorization*, edited by Eleanor Rosch and Barbara B. Lloyd, 9–26. Hillsdale, N.J.: Erlbaum Associates, 1978.

———. *Ethnobiological Classification: Principles of Categorization of Plants and Animals in Traditional Societies*. Princeton, N.J.: Princeton University Press, 1992.

———. "How a Folkbotanical System Can Be Both Natural and Comprehensive: One Maya Indian's View of the Plant World." In *Folkbiology*, edited by Douglas L. Medin and Scott Atran, 71–89. Cambridge, Mass.: MIT Press, 1999.

———. "Speculations on the Growth of Ethnobotanical Nomenclature." *Language in Society* 1 (1972): 51–86.

Berlin, Brent, Dennis E. Breedlove, and Peter H. Raven. *Principles of Tzeltal Plant Classification: An Introduction to the Botanical Ethnography of a Mayan-Speaking People of Chiapas*. New York: Academic Press, 1974.

Berliner, Paul F. *Thinking in Jazz: The Infinite Art of Improvisation*. Chicago Studies in Ethnomusicology. Chicago: University of Chicago Press, 1994.

Bernard, Jonathan W. "Messiaen's Synaesthesia: The Correspondence between Color and Sound Structure in His Music." *Music Perception* 4 (1986): 41–68.

Bernstein, Susan. *Virtuosity of the Nineteenth Century: Performing Music and Language in Heine, Liszt, and Baudelaire*. Stanford, Calif.: Stanford University Press, 1998.

Biederman, Irving. "Recognition-by-Components: A Theory of Human Image Understanding." *Psychological Review* 94 (1987): 115–47.

Blacking, John. "Music, Culture, and Experience." In *Music, Culture, and Experience: Selected Papers of John Blacking*, edited by Reginald Byron, with a foreword by Bruno Nettl, 223–42. Chicago Studies in Ethnomusicology. Chicago: University of Chicago Press, 1995.

Boethius, Ancius Manlius Severinus. *Fundamentals of Music*. Edited by Claude V. Palisca. Translated by Calvin Bower. Music Theory Translation Series. New Haven, Conn.: Yale University Press, 1989.

Bohlman, Philip V. "Ontologies of Music." In *Rethinking Music*, edited by Nicholas Cook and Mark Everist, 17–34. Oxford: Oxford University Press, 1999.

Bonds, Mark Evan. *Wordless Rhetoric: Musical Form and the Metaphor of Oration*. Studies in the History of Music, 4. Cambridge, Mass.: Harvard University Press, 1991.

Brailoiu, Constantin. "Le Folklore musicale." In *Musica aeterna: La Vie et la production musicales de tous les temps et de tous les peuples, en tenant compte particulièrement de la Suisse, de la Belgique, de la France et de la musique de nos jours*, edited by Gottfried Schmid, vol. 2, 277–332. Zurich: M. S. Metz, 1949.

Brinkmann, Reinhold. *Schumann und Eichendorff: Studien zum Liederkreis Opus 39*. Musik-Konzepte, 95. Munich: Edition Text + Kritik, 1997.

Brown, Roger. "How Shall a Thing Be Called?" *Psychological Review* 65 (1958): 14–21.

———. *Social Psychology*. New York: Free Press, 1965.

Burmeister, Joachim. *Musical Poetics [1606]*. Translated by Benito Rivera. Music Theory Translation Series. New Haven, Conn.: Yale University Press, 1993.

Burnham, Scott. "The Second Nature of Sonata Form." In *Music Theory and Natural Order*

*from the Renaissance to the Early Twentieth Century*, edited by Suzannah Clark and Alexander Rehding, 111–41. Cambridge: Cambridge University Press, 2001.

Carroll, Lewis. *The Annotated Alice: Alice's Adventures in Wonderland and Through the Looking Glass*. Introduction and notes by Martin Gardner, illustrated by John Tenniel. New York: W. W. Norton, 2000.

Chambers, Jack. *Milestones I: The Music and Times of Miles Davis to 1960*. Toronto: University of Toronto Press, 1983.

Christensen, Thomas. "Fétis and Emerging Tonal Consciousness." In *Music Theory in the Age of Romanticism*, edited by Ian Bent, 37–56. Cambridge: Cambridge University Press, 1996.

———. *Rameau and Musical Thought in the Enlightenment*. Cambridge Studies in Music Theory and Analysis, 4. Cambridge: Cambridge University Press, 1993.

Chua, Daniel K. L. *The "Galitzin" Quartets of Beethoven: Opp. 127, 132, 130*. Princeton, N.J.: Princeton University Press, 1995.

Coleman, Linda, and Paul Kay. "Prototype Semantics: The English Word *Lie*." *Language* 57 (1981): 26–44.

Cone, Edward T. "Analysis Today." *The Musical Quarterly* 46 (1960): 172–88.

———. *Music, a View from Delft: Selected Essays*. Edited by Robert P. Morgan. Chicago: University of Chicago Press, 1989.

———. *Musical Form and Musical Performance*. New York: W. W. Norton, 1968.

———. "On Derivation: Syntax and Rhetoric." *Music Analysis* 6 (1987): 237–55.

Cook, Nicholas. *Analysing Musical Multimedia*. Oxford: Clarendon Press, 1998.

———. "At the Borders of Musical Identity: Schenker, Corelli and the Graces." *Music Analysis* 18 (1999): 179–233.

Cooper, Grosvenor W., and Leonard B. Meyer. *The Rhythmic Structure of Music*. Chicago: University of Chicago Press, 1960.

Coulson, Seana. *Semantic Leaps: Frame-Shifting and Conceptual Blending in Meaning Construction*. Cambridge: Cambridge University Press, 2001.

Cowdery, James R. *The Melodic Tradition of Ireland*. Kent, Oh.: Kent State University Press, 1990.

Cox, Arnie Walter. "The Metaphoric Logic of Musical Motion and Space." Ph.D. diss., University of Oregon, 1999.

Crawford, Richard. "George Gershwin's 'I Got Rhythm' (1930)." In *The American Musical Landscape*, 213–36. Berkeley: University of California Press, 1993.

Dahlhaus, Carl. *Nineteenth-Century Music*. Translated by J. Bradford Robinson. Berkeley: University of California Press, 1989.

———. *Richard Wagner's Music Dramas*. Translated by Mary Whitall. Cambridge: Cambridge University Press, 1979.

Damasio, Antonio R. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Avon, 1994.

———. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace, 1999.

D'Andrade, Roy G. *The Development of Cognitive Anthropology*. Cambridge: Cambridge University Press, 1995.

D'Andrade, Roy G., and Claudia Strauss, eds. *Human Motives and Cultural Models*. Publications of the Society for Psychological Anthropology. Cambridge: Cambridge University Press, 1992.

Darwin, Charles. *The Voyage of the Beagle*. With an introduction by Leonard Engel. Garden City, N.Y.: Doubleday, 1962.

Davis, Ruth. "The Effects of Notation on Performance Practice in Tunisian Art Music." *The World of Music* 34 (1992): 85–114.

DeBellis, Mark. *Music and Conceptualization*. Cambridge: Cambridge University Press, 1995.

Deffaa, Chip. *Traditionalists and Revivalists in Jazz*. Studies in Jazz, 16. Metuchen: Scarecrow Press; Institute of Jazz Studies, Rutgers, The State University of New Jersey, 1993.

Deutsch, Otto Erich. "Mit Würfeln komponieren." *Zeitschrift für Musikwissenschaft* 12 (1930): 595.

Duchez, Marie-Elisabeth. "La Représentation spatio-verticale du caractère musical grave-aigu et l'élaboration de la notion de hauteur de son dans la conscience musicale occidentale." *Acta musicologica* 51 (1979): 54–73.

Dupré, John. "Are Whales Fish?" In *Folkbiology*, edited by Douglas L. Medin and Scott Atran, 461–76. Cambridge, Mass.: MIT Press, 1999.

Eco, Umberto. *A Theory of Semiotics*. Advances in Semiotics. Bloomington: Indiana University Press, 1976.

Edelman, Gerald M. *Bright Air, Brilliant Fire: On the Matter of Mind*. New York: Basic Books, 1992.

———. *The Remembered Present: A Biological Theory of Consciousness*. New York: Basic Books, 1989.

Eichendorff, Joseph Freiherr von. *Viel Lärmen um nichts*. Neue Gesamtausgabe der Werke und Schriften, 2. Stuttgart: J. G. Cotta'sche Buchhandlung Nachf., 1957.

Emge, Charles. "Meet 'Jazzman' Jack Webb: Swing Era Helped Make Him Music-Conscious." *Downbeat* 22 (10 August 1955): 8.

Fahnestock, Jeanne. *Rhetorical Figures in Science*. New York: Oxford University Press, 1999.

Fauconnier, Gilles. *Mappings in Thought and Language*. Cambridge: Cambridge University Press, 1997.

———. *Mental Spaces: Aspects of Meaning Construction in Natural Language*. 2nd ed. With a foreword by George Lakoff and Eve Sweetser. Cambridge: Cambridge University Press, 1994.

Fauconnier, Gilles, and Eve Sweetser, eds. *Spaces, Worlds, and Grammar*. Chicago: University of Chicago Press, 1996.

Fauconnier, Gilles, and Mark Turner. "Blending as a Central Process of Grammar." In *Conceptual Structure, Discourse and Language*, edited by Adele E. Goldberg, 113–30. Stanford, Calif.: Center for the Study of Language and Information, 1996. [Based on papers presented at the First Conceptual Structure, Discourse and Language Conference, San Diego, 1995.]

———. "Conceptual Integration Networks." *Cognitive Science* 22 (1998): 133–87.

———. "Conceptual Projection and Middle Spaces." UCSD Department of Cognitive Science Technical Report 9401. San Diego: University of California, San Diego, 1994.

———. "Principles of Conceptual Integration." In *Discourse and Cognition: Bridging the Gap: [Second] Conceptual Structure, Discourse, and Language Conference, Buffalo, 1996*, edited by Jean-Pierre Koenig, 269–83. Stanford, Calif.: Center for the Study of Language and Information, 1998.

———. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books, 2002.

Feld, Steven. "Flow Like a Waterfall: The Metaphors of Kaluli Musical Theory." *Yearbook for Traditional Music* 13 (1981): 22–47.

Ferris, David. *Schumann's Eichendorff* Liederkreis *and the Genre of the Romantic Cycle*. New York: Oxford University Press, 2000.

Finnegan, Ruth. "The Relation between Composition and Performance: Three Alternative Modes." In *The Oral and the Literate in Music*, edited by Yoshihiko Tokumaru and Osamu Yamaguti, 73–87. Tokyo: Academia Music, 1986.

Finson, Jon W. "The Intentional Tourist: Romantic Irony in the Eichendorff *Liederkreis* of

Robert Schumann." In *Schumann and His World*, edited by R. Larry Todd, 156–70. Princeton, N.J.: Princeton University Press, 1994.

Fodor, Jerry A. *Concepts: Where Cognitive Science Went Wrong*. Oxford: Clarendon Press, 1998.

———. *The Mind Doesn't Work That Way*. Cambridge, Mass.: MIT Press, 2000.

———. *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, Mass.: MIT Press, 1983.

Forte, Allen. *The Structure of Atonal Music*. New Haven, Conn.: Yale University Press, 1973.

Foucault, Michel. *The Order of Things: An Archaeology of the Human Sciences*. New York: Vintage Books, 1973.

Freudenthal, Gideon. *Atom and Individual in the Age of Newton: On the Genesis of the Mechanistic World View*. Translated by Peter McLaughlin. Boston Studies in the Philosophy of Science, 88. Dordrecht: D. Reidel, 1986.

Gärdenfors, Peter. "Semantics, Conceptual Spaces and the Dimensions of Music." In *Essays on the Philosophy of Music*, edited by Veikko Rantala, Lewis Rowell, and Eero Tarasti, 9–27. Acta Philosophica Fennica, 43. Helsinki: Philosophical Society of Finland, 1988.

Gates, Henry Louis, Jr. *The Signifying Monkey: A Theory of African-American Literary Criticism*. New York: Oxford University Press, 1988.

Geertz, Clifford. "Religion as a Cultural System." In *Anthropological Approaches to the Study of Religion*, edited by Michael Banton, 1–46. New York: Frederick A. Praeger, 1966.

———. "Thick Description: Toward an Interpretive Theory of Culture." In *The Interpretation of Cultures: Selected Essays*, 3–30. New York: Basic Books, 1973.

Gendron, Bernard. "'Moldy Figs' and Modernists: Jazz at War (1942–1946)." In *Jazz among the Discourses*, edited by Krin Gabbard, 31–56. Durham, N.C.: Duke University Press, 1995.

Gentner, Dedre, and Albert L. Stevens, eds. *Mental Models*. Hillsdale, N.J.: Erlbaum Associates, 1983.

Gerigk, Herbert. "Würfelmusik." *Zeitschrift für Musikwissenschaft* 16 (1934): 359–63.

Geselle, Cynthia M. "The Institutionalization of Music Theory in France: 1764–1802." Ph.D. diss., Princeton University, 1989.

Gibbs, Raymond W., Jr. *The Poetics of Mind: Figurative Thought, Language, and Understanding*. Cambridge: Cambridge University Press, 1994.

Gibbs, Raymond W., Jr., and Herbert L. Colston. "The Cognitive Psychological Reality of Image Schemas and Their Transformations." *Cognitive Linguistics* 6 (1995): 347–78.

Gjerdingen, Robert O. *A Classic Turn of Phrase: Music and the Psychology of Convention*. Studies in the Criticism and Theory of Music. Philadelphia: University of Pennsylvania Press, 1988.

———. "Courtly Behaviors." *Music Perception* 13 (1996): 365–82.

———. "Revisiting Meyer's 'Grammatical Simplicity and Relational Richness'." In *Cognitive Bases of Musical Communication*, edited by Mari Riess Jones and Susan Holleran, 225–43. Washington, D.C.: American Psychological Association, 1992.

Gleitman, Lila R., Sharon Lee Armstrong, and Henry Gleitman. "On Doubting the Concept 'Concept.'" In *New Trends in Conceptual Representation: Challenges to Piaget's Theory?* edited by Ellin Kofsky Scholnick, 87–110. Hillsdale, N.J.: Erlbaum Associates, 1983.

Goehr, Lydia. *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music*. Oxford: Clarendon Press, 1992.

———. *The Quest for Voice: Music, Politics, and the Limits of Philosophy*. Berkeley: University of California Press, 1998.

Goetschius, Percy. *Lessons in Music Form: A Manual of Analysis of All the Structural Factors and Designs Employed in Musical Composition*. Bryn Mawr, Penn.: Oliver Ditson, 1904.

Goldman, Herbert G. *Banjo Eyes: Eddie Cantor and the Birth of Modern Stardom*. New York: Oxford University Press, 1997.

Goodman, Nelson. *Languages of Art: An Approach to a Theory of Symbols*. 2nd ed. Indianapolis: Hackett Publishing, 1976.

Gopnik, Alison, and Andrew N. Meltzoff. *Words, Thoughts, and Theories*. Cambridge, Mass.: MIT Press, 1997.

Gopnik, Alison, Andrew N. Meltzoff, and Patricia K. Kuhl. *The Scientist in the Crib: Minds, Brains, and How Children Learn*. New York: William Morrow, 1999.

Gosnell, Harold F. *Negro Politicians: The Rise of Negro Politics in Chicago*. With an introduction by James Q. Wilson. Chicago: University of Chicago Press, 1966; originally published with an introduction by Robert Park, Social Science Studies, 35. Chicago, 1935.

Grey, Thomas S. *Wagner's Musical Prose: Texts and Contexts*. Cambridge: Cambridge University Press, 1995.

Griffin, Donald R. *Animal Minds*. Chicago: University of Chicago Press, 1992.

Groth, Renate. *Die französische Kompositionlehre des 19. Jahrhunderts*. Wiesbaden: Franz Steiner, 1983.

Hampton, James. "Prototype Models of Concept Representation." In *Categories and Concepts: Theoretical Views and Inductive Analysis*, edited by Iven van Mechelen, James Hampton, Ryszard S. Michalski, and Peter Theuns, 67–95. Cognitive Science Series. London: Academic Press, 1993.

Hampton, James, and Danièle Dubois. "Psychological Models of Concepts." In *Categories and Concepts: Theoretical Views and Inductive Analysis*, edited by Iven van Mechelen, James Hampton, Ryszard S. Michalski, and Peter Theuns, 11–33. Cognitive Science Series. London: Academic Press, 1993.

Harris, Paul L. "Thinking by Children and Scientists: False Analogies and Neglected Similarities." In *Mapping the Mind: Domain Specificity in Cognition and Culture*, edited by Lawrence A. Hirschfield and Susan A. Gelman, 294–315. Cambridge: Cambridge University Press, 1994.

Harriss, Ernest C. *Johann Mattheson's* Der vollkommene Capellmeister: *A Revised Translation with Critical Commentary*. Studies in Musicology, 21. Ann Arbor, Mich.: UMI Research Press, 1981.

Hauptmann, Moritz. *The Nature of Harmony and Metre*. 2nd ed. Translated by W. E. Heathcote. London: Sonnenschein, 1893; reprinted with a new foreword by Sigmund Levarie, New York: Da Capo Press, 1991.

Hedges, Stephen A. "Dice Music in the Eighteenth Century." *Music and Letters* 59 (1978): 180–87.

Hill, John Walter. "The Logic of Phrase Structure in Joseph Riepel's *Anfangsgründe zur musikalischen Setzkunst*, Part 2 (1755)." In *Festa musicologica: Essays in Honor of George J. Buelow*, edited by Thomas J. Mathiesen and Benito V. Rivera, 467–87. Festschrift Series, 14. Stuyvesant, N.Y.: Pendragon Press, 1995.

Hillach, Ansgar, and Klaus-Dieter Krabiel. *Eichendorff-Kommentar*. 2 vols. Munich: Winkler, 1971.

Hjelmslev, Louis. *Prolegomena to a Theory of Language*. Rev. ed. Translated by Francis J. Whitfield. Madison: University of Wisconsin Press, 1963.

Hobbes, Thomas. *Leviathan, or the Matter, Forme, and Power of a Common-Wealth Ecclesiasticall and Civill*. London: Andrew Crooke, 1651.

Hockenberry, John. *Moving Violations: War Zones, Wheelchairs, and Declarations of Independence*. New York: Hyperion, 1995.

Hoegi, Pierre. *A Tabular System Whereby the Art of Composing Minuets Is Made So Easy That Any Person, without the Least Knowledge of Musick, May Compose Ten Thousand, All Different, and in the Most Pleasing and Correct Manner*. London: Welcker, 1770.

Hofstadter, Douglas R., and the Fluid Analogies Research Group. *Fluid Concepts and Creative*

*Analogies: Computer Models of the Fundamental Mechanisms of Thought*. New York: Basic Books, 1995.

Holy, Ladislav, and Milan Stuchlik. "The Structure of Folk Models." In *The Structure of Folk Models*, edited by Ladislav Holy and Milan Stuchlik, 1–34. London: Academic Press, 1981.

Hood, Mantle. "The Reliability of Oral Tradition." *Journal of the American Musicological Society* 12 (1959): 201–09.

Hosler, Bellamy. *Changing Aesthetic Views of Instrumental Music in 18th-Century Germany*. Studies in Musicology, 42. Ann Arbor, Mich.: UMI Research Press, 1981.

Hoyt, Peter. "The Concept of *Développement* in the Early Nineteenth Century." In *Music Theory in the Age of Romanticism*, edited by Ian Bent, 141–62. Cambridge: Cambridge University Press, 1996.

———. "Review of Mark Evan Bonds, *Wordless Rhetoric*." *Journal of Music Theory* 38 (1994): 123–43.

Hughes, R. I. G. "Reason and Experiment in Newton's *Opticks*: Comments on Peter Achinstein." In *Philosophical Perspectives on Newtonian Science*, edited by Phillip Bricker and R. I. G. Hughes, 175–84. Cambridge, Mass.: MIT Press, 1990.

Hurte, Michael. *Musik, Bild, Bewegung: Theorie und Praxis auditiv-visueller Konvergenzen*. Bonn: Verlag für systematische Musikwissenschaft, 1982.

Hutchins, Edwin. *Cognition in the Wild*. Cambridge, Mass.: MIT Press, 1995.

Irving, John. *Mozart: The "Haydn" Quartets*. Cambridge Music Handbooks. Cambridge: Cambridge University Press, 1998.

Jackendoff, Ray. *Consciousness and the Computational Mind*. Cambridge, Mass.: MIT Press, 1987.

Johnson, Graham. "*Die schöne Müllerin*: The Poems and Their Background." In *Die schöne Müllerin, D795*. Compact disk. Hyperion, [CD] CDJ 33025, 1996.

Johnson, Mark. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press, 1987.

Johnson-Laird, Philip Nicholas. *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge, Mass.: Harvard University Press, 1983.

Kargon, Robert H. "Atomism in the Seventeenth Century." In *Dictionary of the History of Ideas: Studies of Selected Pivotal Ideas*, vol. 1, edited by Philip P. Wiener, 132–41. New York: Charles Scribner's Sons, 1973.

Keil, Charles. *Tiv Song*. Chicago: University of Chicago Press, 1979.

Keil, Frank C. *Concepts, Kinds, and Cognitive Development*. Cambridge, Mass.: MIT Press, 1989.

Keil, Frank C., Daniel T. Levin, Bethany A. Richman, and Grant Gutheil. "Mechanism and Explanation in the Development of Biological Thought: The Case of Disease." In *Folkbiology*, edited by Douglas L. Medin and Scott Atran, 285–319. Cambridge, Mass.: MIT Press, 1999.

Keiler, Allan. "Music as Metalanguage: Rameau's Fundamental Bass." In *Music Theory: Special Topics*, edited by Richmond Browne, 83–100. New York: Academic Press, 1981.

———. "The Origins of Schenker's Thought: How Man Is Musical." *Journal of Music Theory* 33 (1989): 273–98.

Keller, Hans. "Functional Analysis no. 9a: Mozart's Piano Sonata in A Minor, K. 310." In *Essays on Music*, edited by Christopher Wintle, 129–38. Cambridge: Cambridge University Press, 1994.

———. "The Musical Analysis of Music." In *Essays on Music*, edited by Christopher Wintle, 126–28. Cambridge: Cambridge University Press, 1994.

———. "Wordless Functional Analysis no. 1: Mozart, K. 421." *The Score and I.M.A. Magazine* 22 (1958): 56–64.

Kerman, Joseph. *The Beethoven Quartets*. New York: W. W. Norton, 1966.

Klotz, Sebastian. "*Ars combinatoria* oder 'Musik ohne Kopfzerbrechen': Kalküle des Musikalischen von Kircher bis Kirnberger." *Musiktheorie* 14 (1999): 231–45.

Knaus, Herwig. *Musiksprache und Werkstruktur in Robert Schumanns "Liederkreis": Mit dem Faksimile des Autographs*. Schriften zur Musik, 27. Munich: Emil Katzbichler, 1974.

Knouse, Nola Reed. "Joseph Riepel and the Emerging Theory of Form in the Eighteenth Century." *Current Musicology* 4 (1986): 46–62.

Koch, Carl. *Bernhard Klein (1783–1832): Sein Leben und seine Werke*. Leipzig: Oscar Brandstetter, 1902.

Koch, Heinrich Christoph. *Introductory Essay on Composition: The Mechanical Rules of Melody, Sections 3 and 4*. Translated by Nancy Kovaleff Baker. Music Theory Translation Series. New Haven, Conn.: Yale University Press, 1983.

———. *Versuch einer Anleitung zur Composition*. Leipzig: Adam Friedrich Böhme; Rudolstadt: gedruckt mit Schriften der Löwischen Erben, und Schirach, 1782 (vol. 1); Leipzig: Adam Friedrich Böhme, 1787–93 (vols. 2–3); reprinted, Hildesheim: Georg Olms, 1969.

Köchel, Ludwig Ritter von. *Chronologisch-thematisches Verzeichnis sämtlicher Tonwerke Wolfgang Amadé Mozarts*. 8th ed. Edited by Franz Giegling, Alexander Weinmann, and Gerd Sievers. Wiesbaden: Breitkopf and Härtel, 1983.

Koffka, Kurt. *Principles of Gestalt Psychology*. London: Routledge and Kegan Paul, 1935.

Kramer, Lawrence. *Music and Poetry: The Nineteenth Century and After*. Berkeley: University of California Press, 1984.

Krebs, Harald. *Fantasy Pieces: Metrical Dissonance in the Music of Robert Schumann*. New York: Oxford University Press, 1999.

———. "Schenker's Changing View of Rameau: A Comparison of Remarks in *Harmony*, *Counterpoint*, and 'Rameau or Beethoven?'" *Theoria* 3 (1989): 59–72.

Kropfinger, Klaus. *Wagner and Beethoven: Richard Wagner's Reception of Beethoven*. Translated by Peter Palmer. Cambridge: Cambridge University Press, 1991.

Kuhn, Thomas S. "A Function for Thought Experiments." In *The Essential Tension: Selected Studies in Scientific Tradition and Change*, 240–65. Chicago: University of Chicago Press, 1977.

Kuntz, Paul G. "A Formal Preface and an Informal Conclusion to *The Great Chain of Being*: The Universality and Necessity of Hierarchical Thought." In *Jacob's Ladder and the Tree of Life: Concepts of Hierarchy and the Great Chain of Being*, edited by Marion Leathers Kuntz and Paul Grimley Kuntz, 3–14. American University Studies, Series V, Philosophy, 14. New York: Peter Lang, 1987.

Kurth, Ernst. *Die Voraussetzungen der theoretischen Harmonik und der tonalen Darstellungssysteme*. Berne: Akademische Buchhandlungen von Max Dreschel, 1913.

Lakoff, George. "The Contemporary Theory of Metaphor." In *Metaphor and Thought*, 2nd ed., edited by Andrew Ortony, 202–51. Cambridge: Cambridge University Press, 1993.

———. "The Invariance Hypothesis: Is Abstract Reason Based on Image-Schemas?" *Cognitive Linguistics* 1 (1990): 39–74.

———. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press, 1987.

Lakoff, George, and Mark Johnson. *Metaphors We Live By*. Chicago: University of Chicago Press, 1980.

———. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books, 1999.

Lakoff, George, and Mark Turner. *More Than Cool Reason: A Field Guide to Poetic Metaphor*. Chicago: University of Chicago Press, 1989.

Lang, Paul Henry, ed. *Problems of Modern Music: The Princeton Seminar in Advanced Musical Studies*. New York: W. W. Norton, 1960.

Langacker, Ronald W. "Conceptualization, Symbolization, and Grammar." In *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, edited by Michael Tomasello, 1–39. Mahwah, N.J.: Erlbaum Associates, 1998.

———. *Descriptive Application*. Vol. 2 of *Foundations of Cognitive Grammar*. Stanford, Calif.: Stanford University Press, 1992.

———. *Grammar and Conceptualization*. Berlin: Mouton de Gruyter, 2000.

———. *Theoretical Prerequisites*. Vol. 1 of *Foundations of Cognitive Grammar*. Stanford, Calif., Stanford University Press, 1987.

Langer, Susanne K. *Feeling and Form: A Theory of Art Developed from Philosophy in a New Key*. New York: Charles Scribner's Sons, 1953.

Lassaline, Mary E., Edward J. Wisniewski, and Douglas L. Medin. "Basic Levels in Artificial and Natural Categories: Are All Basic Levels Created Equal?" In *Percepts, Concepts and Categories: The Representation and Processing of Information*, edited by Barbara Burns, 327–78. Advances in Psychology, 93. Amsterdam: North-Holland, 1992.

Lee, Peggy. *Miss Peggy Lee: An Autobiography*. New York: Donald I. Fine, 1989.

Lenat, Douglas B., and R. V. Guha. *Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project*. Reading, Mass.: Addison Wesley, 1989.

Lerdahl, Fred, and Ray Jackendoff. *A Generative Theory of Tonal Music*. Cambridge, Mass.: MIT Press, 1983.

Levy, Janet M. *Beethoven's Compositional Choices: The Two Versions of Opus 18, No. 1, First Movement*. Studies in the Criticism and Theory of Music. Philadelphia: University of Pennsylvania Press, 1982.

———. "Gesture, Form, and Syntax in Haydn's Music." In *Haydn Studies: Proceedings of the International Haydn Conference, Washington, D.C., 1975*, edited by Jens Peter Larsen, Howard Serwer, and James Webster, 355–62. New York: W. W. Norton, 1981.

Lewin, David. "*Auf dem Flusse*: Image and Background in a Schubert Song." In *Schubert: Critical and Analytical Studies*, edited by Walter Frisch, 126–52. Lincoln: University of Nebraska Press, 1986.

———. *Generalized Musical Intervals and Transformations*. New Haven, Conn.: Yale University Press, 1987.

———. *Musical Form and Transformation: Four Analytic Essays*. New Haven, Conn.: Yale University Press, 1993.

London, Justin. "Riepel and *Absatz*: Poetic and Prosaic Aspects of Phrase Structure in 18th-Century Theory." *Journal of Musicology* 8 (1990): 505–19.

Louis, Rudolf, and Ludwig Thuille. *Harmonielehre*. 7th ed. Stuttgart: Carl Grüninger Nachf. Ernst Klett, 1920.

Lovejoy, Arthur O. *The Great Chain of Being: A Study of the History of an Idea*. Cambridge, Mass.: Harvard University Press, 1936.

Loveland, Franklin O. "Snakebite Cure among the Rama Indians of Nicaragua." In *Medical Anthropology*, edited by Francis X. Grollig and Harold B. Haley, 81–102. The Hague: Mouton, 1976.

Löwenstein, Paul. "Mozart-Kuriosa." *Zeitschrift für Musikwissenschaft* 12 (1930): 342–46.

Lutz, Catherine. "Ethnopsychology Compared to What? Explaining Behavior and Consciousness among the Ifaluk." In *Person, Self, and Experience: Exploring Pacific Ethnopsychologies*, edited by Geoffrey M. White and John Kirkpatrick, 35–79. Berkeley: University of California Press, 1985.

Lyle, Wilson. "Colour and Music: An Introduction." *Music Review* 43 (1982): 261–64.

McCauley, Robert N. "The Role of Theories in a Theory of Concepts." In *Concepts and*

*Conceptual Development: Ecological and Intellectual Factors in Categorization*, edited by Ulric Neisser, 288-309. Cambridge: Cambridge University Press, 1987.

McClary, Susan. *Conventional Wisdom: The Content of Musical Form*. Berkeley: University of California Press, 2000.

McCreless, Patrick. "Song Order in the Song Cycle: Schumann's *Liederkreis*, Op. 39." *Music Analysis* 5 (1986): 5-40.

McNeill, David. *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press, 1992.

Macpherson, Stewart. *Form in Music: With Special Reference to the Designs of Instrumental Music*. London: Joseph Williams, 1908.

Macrobius, Ambrosius Aurelius Theodosius. *Commentary on the Dream of Scipio*. Edited and translated by William Harris Stahl. New York: Columbia University Press, 1990.

Macy, Laura. "Speaking of Sex: Metaphor and Performance in the Italian Madrigal." *Journal of Musicology* 14 (1996): 1-34.

Marks, Lawrence E. *The Unity of the Senses: Interrelations among the Modalities*. New York: Academic Press, 1978.

Marx, Adolph Bernhard. *Die Lehre von der musikalischen Komposition, praktisch-theoretisch*. 4 vols. 4th ed. Leipzig: Breitkopf and Härtel, 1841-47.

———. *Ludwig van Beethoven: Leben und Schaffen*. 6th ed. Edited by Gustav Behncke. Berlin: Otto Janke, 1908.

———. *Musical Form in the Age of Beethoven: Selected Writings on Theory and Method*. Edited and translated by Scott G. Burnham. Cambridge Studies in Music Theory and Analysis, 12. Cambridge: Cambridge University Press, 1997.

Mathiesen, Thomas J. *Apollo's Lyre: Greek Music and Music Theory in Antiquity and the Middle Ages*. Lincoln: University of Nebraska Press, 2000.

Mattheson, Johann. *Kern melodischer Wissenschaft*. Hamburg: C. Herold, 1737; reprinted, Hildesheim: Georg Olms, 1976.

———. *Das neu-eröffnete Orchestre*. Hamburg: Printed for the author, 1713; reprinted, Hildesheim: Georg Olms, 1993.

———. *Der vollkommene Capellmeister*. Hamburg: C. Herold, 1739; reprinted, Dokumenta musicologica, 1. Reihe, Druckschriften-Faksimiles, 5, edited by Margarete Reimann, Kassel: Bärenreiter-Verlag, 1954.

Mervis, Carolyn B., and Eleanor Rosch. "Categorization of Natural Objects." *Annual Review of Psychology* 32 (1981): 89-115.

Meyer, Leonard B. *Explaining Music: Essays and Explorations*. Berkeley: University of California Press, 1973; reprinted, Chicago: University of Chicago Press, 1978.

Milne, A. A. *The World of Pooh: The Complete Winnie-the-Pooh and the House at Pooh Corner*. Illustrated by E. H. Shepard. New York: E. P. Dutton, 1957.

Minsky, Marvin. "A Framework for Representing Knowledge." In *The Psychology of Computer Vision*, edited by Patrick Henry Winston, 211-77. New York: McGraw-Hill, 1975.

———. *The Society of Mind*. New York: Simon and Schuster, 1985.

Momigny, Jérôme-Joseph de. *Cours complet d'harmonie et de composition, d'après une théorie neuve et générale de la musique*. 3 vols. Paris: Chez l'auteur, 1806.

———. "Musique," in *Musique*, Series 36 of *Encyclopédie méthodique, ou par ordre de matières*, edited by Nicolas Étienne Framéry, Pierre Louis Ginguené, and Jérôme-Joseph de Momigny, vol. 2, 175-94. Paris: Chez Madame veuve Agasse, 1818.

Monson, Ingrid. *Saying Something: Jazz Improvisation and Interaction*. Chicago Studies in Ethnomusicology. Chicago: University of Chicago Press, 1996.

Morgan, Robert P. "Circular Form in the *Tristan* Prelude." *Journal of the American Musicological Society* 53 (2000): 69-103.

———. "Symmetrical Form and Common-Practice Tonality." *Music Theory Spectrum* 20 (1998): 1–47.

———. "The Theory and Analysis of Tonal Rhythm." *The Musical Quarterly* 64 (1978): 435–73.

Mozart, Wolfgang Amadeus. *Musikalisches Würfelspiel: Eine Anleitung "Walzer oder Schleifer mit zwei Würfeln zu componieren ohne Musikalisch zu seyn, noch von der Composition etwas zu verstehen."* Edited by Karl Heinz Taubert. Mainz: B. Schott's Söhne, 1956.

Müller, Wilhelm. *Vermischte Schriften.* 5 vols. Edited by Gustav Schwab. Leipzig: F. A. Brockhaus, 1830.

Murphy, Gregory L., and Douglas L. Medin. "The Role of Theories in Conceptual Coherence." *Psychological Review* 92 (1985): 289–316.

*Musica enchiriadis and Scolica enchiriadis.* Edited by Claude V. Palisca. Translated by Raymond Erickson. Music Theory Translation Series. New Haven, Conn.: Yale University Press, 1995.

Narmour, Eugene. *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model.* Chicago: University of Chicago Press, 1990.

———. *The Analysis and Cognition of Melodic Complexity: The Implication-Realization Model.* Chicago: University of Chicago Press, 1992.

———. "Some Major Theoretical Problems Concerning the Concept of Hierarchy in the Analysis of Tonal Music." *Music Perception* 1/2 (1983–84): 129–99.

Nattiez, Jean-Jacques. *Proust as Musician.* Translated by Derrick Puffett. Cambridge: Cambridge University Press, 1989.

Newton, Isaac. *Opticks, or a Treatise of the Reflections, Refractions, Inflections and Colours of Light.* 4th ed. London: W. Innys, 1730; reprinted, New York: Dover, 1952.

Nicomachus. *Enchiridion.* In *Harmonic and Acoustic Theory.* Vol. 2 of *Greek Musical Writings.* Edited and translated by Andrew Barker, 247–69. Cambridge Readings in the Literature of Music. Cambridge: Cambridge University Press, 1989.

Noguchi, Hideo. "Mozart— Musical Game in C Major, K.516f." *Mitteilungen der Internationalen Stiftung Mozarteum* 38 (1990): 89–101.

Oakhill, Jane, and Alan Garnham, eds. *Mental Models in Cognitive Science: Essays in Honour of Phil Johnson-Laird.* East Sussex, U.K.: Psychology Press, 1996.

Patrides, C. A. "Hierarchy and Order." In *Dictionary of the History of Ideas: Studies of Selected Pivotal Ideas,* edited by Philip P. Wiener, vol. 2, 434–49. New York: Charles Scribner's Sons, 1973.

Pinker, Steven. *How the Mind Works.* New York: W. W. Norton, 1997.

Pollard-Gott, Lucy. "Emergence of Thematic Concepts in Repeated Listening to Music." *Cognitive Psychology* 15 (1983): 66–94.

Porter, James. "Context, Epistemics, and Value: A Conceptual Performance Model Reconsidered." In *Selected Reports in Ethnomusicology: Issues in the Conceptualization of Music,* 7, edited by James Porter and Ali Jihad Racy, 69–97. Los Angeles: University of California, Department of Ethnomusicology, 1988.

———. "Jeannie Robertson's *My Son David*: A Conceptual Performance Model." *Journal of American Folklore* 89 (1976): 7–26.

Powers, Richard. *Galatea 2.2.* New York: Farrar, Straus Giroux, 1995.

Proust, Marcel. *Swann's Way.* Vol. 1 of *Remembrance of Things Past.* Translated by C. K. Scott Moncrieff and Terence Kilmartin. New York: Vintage Books, 1981.

Prout, Ebenezer. *Musical Form.* London: Augener, 1893.

Ptolemy. *Harmonics.* In *Harmonic and Acoustic Theory.* Vol. 2 of *Greek Musical Writings.* Edited and translated by Andrew Barker, 275–391. Cambridge Readings in the Literature of Music. Cambridge: Cambridge University Press, 1989.

Quinn, Naomi, and Dorothy Holland. "Culture and Cognition." In *Cultural Models in Language and Thought,* edited by Dorothy Holland and Naomi Quinn, 3–40. Cambridge: Cambridge University Press, 1987.

Rameau, Jean-Philippe. *Génération harmonique ou traité de musique théorique et pratique.* Paris: Prault fils, 1737. Facsimile edition; Monuments of Music and Music Literature in Facsimile. Second Series, Music Literature, 6. New York: Broude Brothers, 1966.

———. *Nouveau système de musique théorique et pratique.* Paris: Ballard, 1726.

———. *Observations sur notre instinct pour la musique, et sur son principe.* Paris: Prault fils, 1754. Facsimile edition; Monuments of Music and Music Literature in Facsimile. Second Series, Music Literature, 54. New York: Broude Brothers, 1967.

Ratner, Leonard G. "*Ars combinatoria*: Chance and Choice in Eighteenth-Century Music." In *Studies in Eighteenth-Century Music: A Tribute to Karl Geringer on His Seventieth Birthday*, edited by H. C. Robbins Landon, 343–63. New York: Da Capo Press, 1979.

———. *The Beethoven String Quartets: Compositional Strategies and Rhetoric.* Stanford, Calif.: Stanford Bookstore, 1995.

———. *Classic Music: Expression, Form, and Style.* New York: Schirmer Books, 1980.

———. "Harmonic Aspects of Classic Form." *Journal of the American Musicological Society* 2 (1949): 159–68.

Réti, Rudolph. *The Thematic Process in Music.* New York: Macmillan, 1951.

Riemann, Hugo. *Musikalische Logik: Hauptzüge der physiologischen und psychologischen Begründung unseres Musiksystems.* Leipzig: C. F. Kahnt, 1874.

———. *Musik-Lexikon.* Leipzig: Verlag des Bibliographischen Instituts, 1882.

———. *System der musikalischen Rhythmik und Metrik.* Leipzig: Breitkopf and Härtel, 1903.

Riepel, Joseph. *Grundregln zur Tonordnung insgemein.* Vol. 2 of *Anfangsgründe zur musicalischen Setzkunst.* Regensburg: Johann Leopold Montag, 1755.

———. *Sämtliche Schriften zur Musiktheorie.* Edited by Thomas Emmerig. Wiener musikwissenschaftliche Beiträge, 20. Vienna: Böhlau, 1996.

Robertson-DeCarbo, Carol E. "*Tayil* as Category and Communication among the Argentine Mapuche: A Methodological Suggestion." *Yearbook of the International Folk Music Council* 8 (1976): 35–52.

Rosch, Eleanor. "Categorization." In *Encyclopedia of Human Behavior*, vol. 1, V. S. Ramachandran, gen. ed., 513–23. San Diego: Academic Press, 1994.

———. "Cognitive Representations of Semantic Categories." *Journal of Experimental Psychology: General* 104 (1975): 192–233.

———. "Human Categorization." In *Studies in Cross-Cultural Psychology*, 1, edited by Neil Warren, 1–49. London: Academic Press, 1977.

———. "On the Internal Structure of Perceptual and Semantic Categories." In *Cognitive Development and the Acquisition of Language*, edited by Timothy E. Moore, 111–44. New York: Academic Press, 1973.

———. "Principles of Categorization." In *Cognition and Categorization*, edited by Eleanor Rosch and Barbara B. Lloyd, 27–48. Hillsdale, N.J.: Erlbaum Associates, 1978.

Rosch, Eleanor, and Carolyn B. Mervis. "Family Resemblances." *Cognitive Psychology* 7 (1975): 573–605.

Rosch, Eleanor, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson, and Penny Boyes-Braem. "Basic Objects in Natural Categories." *Cognitive Psychology* 8 (1976): 382–439.

Rousseau, Jean-Jacques. "Essay on the Origin of Languages (c. 1763)." In *Essay on the Origin of Languages and Writings Related to Music*, translated and edited by John T. Scott, 289–332. Collected Writings of Rousseau, 7. Hanover, N.H.: University Press of New England, 1998.

———. "Letter on French Music." In *Essay on the Origin of Languages and Writings Related to Music*, translated and edited by John T. Scott, 141–74. Collected Writings of Rousseau, 7. Hanover, N.H.: University Press of New England, 1998.

Rowell, Lewis. *Music and Musical Thought in Early India.* Chicago Studies in Ethnomusicology. Chicago: University of Chicago Press, 1992.

Russolo, Luigi. *The Art of Noises (1916).* Translated by Barclay Brown. New York: Pendragon Press, 1986.

Rust, Brian. *The American Dance Band Discography 1917–1942.* New Rochelle, N.Y.: Arlington House, 1975.

Saint-Saëns, Camille. *Le Carnaval des animaux: Grande fantasie zoologique.* Edited by Felix Aprahamian. Zurich: Eulenberg, 1974.

Sams, Eric. *The Songs of Johannes Brahms.* New Haven, Conn.: Yale University Press, 2000.

Schank, Roger C., and Robert P. Abelson. *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures.* Hillsdale, N.J.: Erlbaum Associates, 1977.

Schelling, Friedrich Wilhelm Joseph von. *Philosophy of Art.* Edited, translated, and introduced by Douglas W. Stott, with a foreword by David Simpson. Theory and History of Literature, 58. Minneapolis: University of Minnesota Press, 1989.

Schenker, Heinrich. *Counterpoint: Book I: Cantus Firmus and Two-Voice Counterpoint.* Edited by John Rothgeb. Translated by John Rothgeb and Jürgen Thym. New Musical Theories and Fantasies, 2. New York: Schirmer Books, 1987.

———. *Free Composition (Der freie Satz).* Edited and translated by Ernst Oster. New Musical Theories and Fantasies, 3. New York: Schirmer Books, 1979.

———. *Der freie Satz.* Neue musikalische Theorien und Phantasien, 3. Vienna: Universal-Edition A. G., 1935.

———. *Harmonielehre.* Neue musikalische Theorien und Phantasien, 1. Stuttgart: J. G. Cotta'sche, 1906.

———. "Rameau or Beethoven? Creeping Paralysis or Spiritual Potency in Music?" In *The Masterwork in Music: A Yearbook,* 1–9. Translated by Ian Bent, edited by William Drabkin. Cambridge Studies in Music Theory and Analysis, 5. Cambridge: Cambridge University Press, 1994.

Schmid, Hans, ed. *Musica et Scolica enchiriadis una cum aliquibus tratatulis adiunctis.* Veröffentlichungen der Musikhistorischen Kommission, Bayerische Akademie der Wissenschaften, 3. Munich: Verlag der Bayerische Akademie der Wissenschaften, 1981.

Schoenberg, Arnold. "Brahms the Progressive (1947)." In *Style and Idea: Selected Writings of Arnold Schoenberg,* edited by Leonard Stein, translated by Leo Black, 398–441. New York: St. Martin's Press, 1975.

———. *Coherence, Counterpoint, Instrumentation, Instruction in Form.* Edited by Severine Neff. Translated by Severine Neff and Charlotte M. Cross. Lincoln: University of Nebraska Press, 1994.

———. *Fundamentals of Musical Composition.* Edited by Gerald Strang and Leonard Stein. London: Faber and Faber, 1967.

———. *Harmonielehre.* Leipzig: Verlagseigentum der Universal-Edition, 1911.

———. *The Musical Idea and the Logic, Technique, and Art of Its Presentation.* Edited and translated by Patricia Carpenter and Severine Neff. New York: Columbia University Press, 1995.

———. "My Evolution (1949)." In *Style and Idea: Selected Writings of Arnold Schoenberg,* edited by Leonard Stein, translated by Leo Black, 79–92. New York: St. Martin's Press, 1975.

———. "Opinion or Insight? (1926)." In *Style and Idea: Selected Writings of Arnold Schoenberg,* edited by Leonard Stein, translated by Leo Black, 258–64. New York: St. Martin's Press, 1975.

———. *Theory of Harmony.* Translated by Roy E. Carter. Berkeley: University of California Press, 1978.

*Sectio canonis.* In *Harmonic and Acoustic Theory.* Vol. 2 of *Greek Musical Writings.* Edited by Andrew Barker, 191–208. Cambridge Readings in the Literature of Music. Cambridge: Cambridge University Press, 1989.

Seeger, Anthony. *Why Suyá Sing: A Musical Anthropology of an Amazonian People.* Cambridge: Cambridge University Press, 1987.

Seeger, Charles. "Oral Tradition in Music." In *Funk and Wagnalls Standard Dictionary of Folklore, Mythology, and Legend*, edited by Maria Leach and Jerome Fried, vol. 2, 825–29. New York: Funk and Wagnalls, 1950.

Sessions, Roger. *The Musical Experience of Composer, Performer, Listener.* Princeton, N.J.: Princeton University Press, 1950.

Shore, Bradd. *Culture in Mind: Cognition, Culture, and the Problem of Meaning.* New York: Oxford University Press, 1996.

Smith, Edward E., and Douglas L. Medin. *Categories and Concepts.* Cambridge, Mass.: Harvard University Press, 1981.

Snarrenberg, Robert. "Competing Myths: The American Abandonment of Schenker's Organicism." In *Theory, Analysis and Meaning in Music*, edited by Anthony Pople, 29–56. Cambridge: Cambridge University Press, 1994.

Spitzer, Mark. "Marx's 'Lehre' and the Science of Education: Towards the Recuperation of Music Pedagogy." *Music and Letters* 79 (1998): 489–526.

Stark, Lucien. *A Guide to the Solo Songs of Johannes Brahms.* Bloomington: Indiana University Press, 1995.

Stone, Ruth M. *Let the Inside Be Sweet: The Interpretation of Music Event among the Kpelle of Liberia.* Bloomington: Indiana University Press, 1982.

Sweetser, Eve E. "The Definition of *Lie*: An Examination of the Folk Models Underlying a Semantic Prototype." In *Cultural Models in Language and Thought*, edited by Dorothy Holland and Naomi Quinn, 43–66. Cambridge: Cambridge University Press, 1987.

Talbot, Michael, ed. *The Musical Work: Reality or Invention?* Liverpool: University of Liverpool, 2000.

Talmy, Leonard. *Concept Structuring Systems.* Vol. 1 of *Toward a Cognitive Semantics.* Cambridge, Mass.: MIT Press, 2000.

Tanaka, James W., and Marjorie Taylor. "Object Categories and Expertise: Is the Basic Level in the Eye of the Beholder?" *Cognitive Psychology* 23 (1991): 457–82.

Taubert, Karl Heinz, ed. *Das Menuett: Musikalisches Würfelspiel, Paris 1786.* Zurich: Musikhaus Pan, 1988.

Temperley, David. "Motivic Perception and Modularity." *Music Perception* 13 (1995): 141–69.

Thomas, Downing A. *Music and the Origins of Language: Theories from the French Enlightenment.* New Perspectives in Music History and Criticism. Cambridge: Cambridge University Press, 1995.

Tomasello, Michael. "Introduction: A Cognitive-Functional Perspective on Language Structure." In *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, edited by Michael Tomasello, vii–xxiii. Mahwah, N.J.: Erlbaum Associates, 1998.

Tovey, Donald Francis. *Symphonies.* Vol. 1 of *Essays in Musical Analysis.* Oxford: Oxford University Press, 1935.

Treitler, Leo. "History and the Ontology of the Musical Work." *Journal of Aesthetics and Art Criticism* 51 (1993): 483–97.

———. "Transmission and the Study of Music History." In *Report of the Twelfth Congress, Berkeley 1977*, edited by Daniel Heartz and Bonnie Wade, 202–11. Kassel: Bärenreiter-Verlag; Philadelphia: The American Musicological Society, 1981.

Turner, Mark. "Aspects of the Invariance Hypothesis." *Cognitive Linguistics* 1 (1990): 247–55.

———. "Design for a Theory of Meaning." In *The Nature and Ontogenesis of Meaning*, edited by Willis F. Overton and David S. Palermo, 91–108. Hillsdale, N.J.: Erlbaum Associates, 1994.

———. "Figure." In *Figurative Language and Thought*, edited by Albert N. Katz, Cristina Cacciari, Raymond W. Gibbs, Jr., and Mark Turner, 44–87. New York: Oxford University Press, 1998.

———. "An Image-Schematic Constraint on Metaphor." In *Conceptualizations and Mental Processing in Language*, edited by Richard A. Geiger and Brygida Rudzka-Ostyn, 291–306. Cognitive Linguistics Research, 3. Berlin: Mouton de Gruyter, 1993.

———. *The Literary Mind*. New York: Oxford University Press, 1996.

Turner, Mark, and Gilles Fauconnier. "Conceptual Integration and Formal Expression." *Metaphor and Symbolic Activity* 10 (1995): 183–204.

———. "Conceptual Integration in Counterfactuals." In *Discourse and Cognition: Bridging the Gap: [Second] Conceptual Structure, Discourse, and Language Conference, Buffalo, 1996*, edited by Jean-Pierre Koenig, 285–96. Stanford, Calif.: Center for the Study of Language and Information, 1998.

Tversky, Barbara, and Kathleen Hemenway. "Categories of Environmental Scenes." *Cognitive Psychology* 15 (1983): 121–49.

———. "Objects, Parts, and Categories." *Journal of Experimental Psychology: General* 113 (1984): 169–93.

Verba, E. Cynthia. "The Development of Rameau's Thoughts on Modulation and Chromatics." *Journal of the American Musicological Society* 26 (1973): 69–91.

Waidelich, Till Gerrit, ed. *Franz Schubert: Dokumente, 1817–1830*. Veröffentlichungen des Internationalen Franz Schubert Instituts, 10. Tutzing, Germany: Hans Schneider, 1993.

Walser, Robert. "'Out of Notes': Signification, Interpretation, and the Problem of Miles Davis." In *Jazz among the Discourses*, edited by Krin Gabbard, 165–88. Durham, N.C.: Duke University Press, 1995.

Weber, Gottfried. *Theory of Musical Composition, Treated with a View to a Naturally Consecutive Arrangement of Topics*. Edited by John Bishop. Translated by James F. Warner. London: Robert Cocks, 1851.

Welker, Robert L. "Abstraction of Themes from Melodic Variations." *Journal of Experimental Psychology: Human Perception and Performance* 8 (1982): 435–47.

Wendt, Lloyd, and Herman Kogan. *Big Bill of Chicago*. Indianapolis: Bobbs-Merrill, 1953.

Will, Richard. "Time, Morality, and Humanity in Beethoven's *Pastoral* Symphony." *Journal of the American Musicological Society* 50 (1997): 271–329.

Wittgenstein, Ludwig. *Philosophical Investigations*. Translated by G. E. M. Anscombe. New York: Macmillan, 1953.

Wolff, Viktor Ernst. *Lieder Robert Schumanns in ersten und späteren Fassungen*. Berlin: H. S. Hermann, 1913.

Wolzogen, Hans von. *Guide through the Musical Motives of Richard Wagner's Tristan and Isolde: With a Preface on the Legend and the Poem of Wagner's Drama*. Leipzig: Feodor Reinboth, 1889.

———. *Thematischer Leitfaden durch die Musik zu Richard Wagner's Tristan und Isolde: Nebste einem Vorworte über den Sagenstoff des Wagner'schen Dramas*. Leipzig: Feodor Reinboth, 1890.

Youens, Susan. "Behind the Scenes: *Die schöne Müllerin* before Schubert." *19th Century Music* 15 (1991): 3–22.

———. *Retracing a Winter's Journey: Schubert's Winterreise*. Ithaca, N.Y.: Cornell University Press, 1991.

———. *Schubert: Die schöne Müllerin*. Cambridge Music Handbooks. Cambridge: Cambridge University Press, 1992.

———. *Schubert, Müller, and* Die schöne Müllerin. Cambridge: Cambridge University Press, 1997.

Zanten, Wim van. "The Tone Material of the Kacapi in Tembang Sunda in West Java." *Ethnomusicology* 30 (1986): 84–112.

Zbikowski, Lawrence. "The Blossoms of 'Trockne Blumen': Music and Text in the Early Nineteenth Century." *Music Analysis* 18 (1999): 307–45.

———. "Large-Scale Rhythm and Systems of Grouping." Ph.D. diss., Yale University, 1991.

Ziolkowski, Theodore. *German Romanticism and Its Institutions*. Princeton, N.J.: Princeton University Press, 1990.

*Recordings*

This reference list makes use of the following abbreviations: CD = compact disc; LP = 33 rpm record; 78 = 78 rpm record.

Bennie Krueger's Orchestra. "Bye Bye Blackbird." Brunswick, [78] no. 3186, 1926.

Benny Goodman Sextet. "I Got Rhythm." In *Benny Goodman*. CBS Records, [LP] P5 15536 CP, 1981.

Bobby Short. "Bye Bye Blackbird." In *Bobby Short*. Atlantic, [LP] 1230, 1956.

Charlie and His Orchestra. "Bye Bye Blackbird." In *Charlie and His Orchestra: German Propaganda Swing 1941–1942*. Harlequin, [CD] HQ CD 03, 1990.

Charlie Parker and others. "I Got Rhythm." In *Bird: The Complete Charlie Parker on Verve*. Verve/Polygram Classics, [CD] 837 142-2, 1988.

Chick Webb and His Little Chicks. "I Got Rhythm." In *Spinnin' the Webb: Chick Webb and His Orchestra*. GRP Records, [CD] 513678L, 1994.

Derek Bailey, Pat Metheny, Gregg Bendian, and Paul Wertico. *The Sign of 4*. Knitting Factory Works, [CD] KFW 197, 1997.

Don Redman and His Orchestra. "I Got Rhythm." In *Don Redman and His Orchestra, 1931–1933*. Classics, [CD] 543 Classics, 1990.

Eddie Cantor. "Bye Bye Blackbird." In *Songs by Eddie Cantor from* The Eddie Cantor Story. Capitol, [LP] L-46, 1954.

Fletcher Henderson and His Orchestra. "Yeah Man." In *Fletcher Henderson, Developing an American Orchestra, 1923–1937*. Smithsonian Collection, [LP] R 006, 1977.

Fred Rich and His Orchestra. "I Got Rhythm." In *1930—You're Driving Me Crazy: Portrait of a Year in Music*. Phontastic, [CD] PHONT CD 7618, 1992.

Freddie "Schnickelfritz" Fisher and His Orchestra. "Bye Bye Blackbird." Decca, [78] no. 25357/3788, 1941.

Gene Austin. "Bye Bye Blackbird." Victor, [78] no. 20044-B, 1926.

George Gershwin. "I Got Rhythm." In *Gershwin Performs Gershwin: Rare Recordings 1931–1935*. Musical Heritage Society, [CD] 512923A, 1991.

George Gershwin and Ira Gershwin. *Girl Crazy*. Elektra Nonesuch, [CD] 79250-2, 1990.

Jimmy Dorsey and His Orchestra. "I Got Rhythm." In *Jimmy Dorsey and His Orchestra*, Vol. 2: *1942–44*. Hindsight Records, [LP] HSR-153, 1980.

Johnny Maddox. "Bye Bye Blackbird." In *Tap Dance Rhythms*. Dot, [LP] DLP 3008, 1955.

Lester Young with the Kansas City Six. "I Got Rhythm." In *The Tenor Sax: Lester Young, Chu Berry and Ben Webster: The Commodore Years*. Atlantic, [LP] SD 2-307, 1973.

Marlene Fingerle and Arthur Schutt. "Medley: 'I Know That You Know,' 'Baby Face,' 'Bye Bye Blackbird.'" In *Song Hits of 1926*. Songs of Our Times. Decca, [LP] DL 5170, 1950.

Miles Davis. *'Round about Midnight*. Columbia, [CD] CK 40610, 1987.

Peggy Lee. "Bye Bye Blackbird." In *Songs from* Pete Kelly's Blues. Decca, [LP] DL 8166, 1955.

"Pete Kelly and His Big Seven" (Matty Matlock and others). *Pete Kelly's Blues*. BMG Special Products; [CD] DRC12081, 1998.

Red Nichols and His Five Pennies. "I Got Rhythm." In *Rarest Brunswick Masters, 1926–1931: Red Nichols and His Five Pennies.* MCA Records, [LP] MCA 1518, 1982.

Red Norvo and His Swing Sextette. "I Got Rhythm." In *Red Norvo.* Time-Life Records, [LP] STL-J14, 1980.

Sidney Bechet and the New Orleans Feetwarmers. "Shag." In *Sidney Bechet.* Time-Life Records, [LP] STL-J09, 1980.

# INDEX

Page numbers in italics indicate that a musical example appears on the given page. Page numbers in boldface indicate substantial analytical discussions with numerous musical examples and figures.


Abelson, Robert P., 112
*Absatz*, 293, 299
Académie Royale de Musique. *See* Opéra
Agawu, Victor Kofi, 51n45, 244–45, 286
Aldwell, Edward, viii
analysis, musical
  and music theory, 19, 134–35, 333
analysis of song, 18, 244–45
  and conceptual blending, 254–55, 264–65, 285–86, 332
*Anlage*, 296
*Anstieg*, 130, 261
anthropomorphism, 79–81, 91–95, 113, 127n59, 248, 315, 327
*Antrag* (Propositio), 295
Aristotle, 10, 311
Aristoxenian music theory, 7, 12–14, 16
Aristoxenus, 10–12, 14, 63
Arom, Simha, 217
*Aufführung* (Executio), 295
*Ausarbeitung* (Elaboratio), 295–96
*Ausführung* (Executio), 295–96
Austin, Gene, 231–32

"Baby Face" (by Harry Akst and Benny Davis), 234
Bach, Johann Sebastian, 126
Bakhtin, Mikhail Mikhailovich, 18, 204, 224–25, 227–29
Bali. *See* pitch relations, descriptions of
Bamberger, Jeanne
  research on musical understanding, 96–99
  work with Jeff, 17, 99–101, 103–7
Barker, Andrew, 63
Barsalou, Lawrence W., 41, 69
Bechet, Sidney, 210
Beethoven, Ludwig van, 18, 27, 48, 59, 169, 201–3, 242
  Fifth Symphony (Symphony No. 5 in C minor, Op. 67), first mvmt., 34, *35–40*, 42–44, 46–48
  Ninth Symphony (Symphony No. 9 in D minor, Op. 125), 120
  piano sonatas, 120
  Sixth Symphony (Symphony No. 6 in F major, Op. 68; "Pastorale"), 89–94, 327, 331
  String Quartet, Op. 18 No. 1, first mvmt., **177–91**, 191–94, 196, 198, 200, 329
  String Quartet, Op. 18 No. 1 [first version], first mvmt., 177, **191–97**
  String Quartet, Op. 18 No. 6, first mvmt., **169–77**, 199, 329
*Bekräfftigung* (Confirmatio), 295
*Bericht* (Narratio), 295
Berlin, Brent, 32, 115
Berliner, Paul F., 220, 331
Boethius, Ancius Manlius Severinus, 15
Brahms, Johannes, 126, 279
  compositional technique, 25, 57–58, 284
  "In der Fremde" (Op. 3 No. 5), **279–84**, 286, 331
  Sextet in B♭ major, first mvmt., Op. 18, *26*, 26–27, *28*
Brailoiu, Constantin, 216–17
*Brechung*, 127
Brown, Roger, 31, 33
Bruckner, Anton
  Sixth Symphony, 57–58

Bruno, Giordano, 312
"Bye Bye Blackbird" (by Mort Dixon and Ray Henderson), 18, 203–4, 219, 223–25, **229–39**, 240
  racial associations, 223–24, 239–40
  recording history, 230–39

cadence, 121, 151–52, 284
  complete (*vollständiger Schluß*), 299
  garbled (*verstümmelter Schluß*), 299
  perfect authentic, 61
Cage, John
  *4'33"*, 242
Cantor, Eddie (Edward Israel Iskowitz), 235
Carroll, Lewis (Charles Dodgson)
  *Alice's Adventures in Wonderland*, 287–88, 305, 324
  *Through the Looking Glass*, 169
categories, 12, 16, 326, 332–33
  artificial (*see* categories: Type 2)
  basic-level, 31–33
  basic-level, and musical motives, 34, 36
  classical (*see* categories: Type 2)
  natural (*see* categories: Type 1)
  Type 1, 39–42, 44, 48–49, 61n59, 149, 326
  Type 2, 40–41, 48, 61n59
  typicality effects in, 36–40
  typicality effects in, and musical motive, 43–44, 46–49
categories, musical, 59–62, 168, 176, 242, 326
  and compositional strategy, 139–40, 150–52, 159, 167, 191, 199–200
  and musical ontology, 203, 213, 222–23
  typicality effects in, and compositional strategy, 47, 50–51, 149, 153, 200
categories, musical motives as, 34, 43–44, 46–52, 58–59
  in Beethoven's String Quartet Op. 18 No. 1, 177, 179–80, 182–84, 186, 188–89, 191, 193–94, 196, 198
  in Beethoven's String Quartet Op. 18 No. 6, 169, 171–74, 176
  in Mozart's String Quartet K. 465, 155, 158–60, 164, 166–67
categorization, 12–17, 24, 30–33, 37, 43–46, 48–49, 51, 58, 60–61, 77, 102, 108, 203–4, 222, 327–28, 333
  and compositional strategy, 139, 199–200
  research on, 30–31, 108
Catel, Charles Simon, 301n39
Chambers, Jack, 224
Chambers, Paul, 223, 239
changes (i.e., harmonic progressions; in jazz performance), 211n21
Charlie and His Orchestra, 234
Chomsky, Noam, 4
chorus (in jazz performance), 210
CINs. *See* conceptual integration networks
cognitive capacities, 328, 331, 334
cognitive domain, 110n20
cognitive linguistics, 5, 17, 64, 66, 138–39
cognitive structure, 4–6, 17, 19, 108, 111, 119, 133–34, 330, 333
cognitive theory, 19
Coleman, Linda, 226, 242
Colston, Herbert L., 69
Coltrane, John, 223–25, 239–40
competence, linguistic, 4
competence, musical, 4
compositional strategy, 47, 54, 138–41, 148–49, 152–54, 159, 167–68, 177, 180, 184, 191, 198–200, 264, 329
  and musical syntax, 138–40, 148, 153–54, 167, 198–200, 264
concepts, 108–9
  categories and, 60–61
  musical, 17, 64, 102, 326–27 (*see also* categories, musical)
conceptual blending, 17–18, 288, 323, 331
  in Brahms's "In der Fremde," 284
  in Klein's "Trockne Blumen," 254, 256, 264–65
  in Müller's "Trockne Blumen," 248–49
  and program music, 89–94
  role in text painting, 65, 82–84, 87–89
  in Schubert's "Trockne Blumen," 263–65, 286
  in Schumann's "Im Rhein," 270–72
  in Schumann's "In der Fremde," 278–79
  theory of, 77–82
  *See also* analysis of song
conceptual domains, 103, 108–9
  basic features of, 112–13
  and mental spaces, 113
  and theories, 113
conceptual integration networks (CINs), 78–80, 82, 331
  for anthropomorphism, 91–92
  for Brahms's "In der Fremde," 282
  for Klein's "Trockne Blumen," 253–56
  for Palestrina's text painting, 82–83
  for Schubert's "Trockne Blumen," 262–63
  for Schumann's "Im Rhein," 270
  for Schumann's "In der Fremde," 277
  topography of, 79–80, 254, 265, 272, 283, 285–86
  for Tovey's analysis of Beethoven's Sixth Symphony, 92–93
  for Wert's "Tirsi morir volea," 87–88
conceptual models, 14–17, 102, 108–9, 324, 327, 332–33

basic features of, 109–12
cross-domain mapping and, 71–72
global, 46–48, 50, 59, 62, 72, 74, 76, 112, 167, 218, 221–22
idealization of, 111
Jeff's development and use of, 101–2, 104–7, 109, 218
local, 48, 50, 59, 62, 112, 218, 221–22
and music theory, 332
musical motives and, 46–48, 50–52, 57–58
and musical notation, 221–22
and musical ontology, 203, 213–14, 242, 330
and musical performance, 216–23
negotiation of, 219–22
representing, 112
role in categorization, 44–48
and theories, 15–17, 62, 98–99, 102–3, 107–9, 111, 125, 131, 133, 317, 323–24
conceptual models for musical categories
for Beethoven's Fifth Symphony, 46–48
for Beethoven's String Quartet Op. 18 No. 1, 179–80, 184, 187, 198
for Beethoven's String Quartet Op. 18 No. 6, 171–72
for "Bye Bye Blackbird," 233, 235, 238–39
for "I Got Rhythm," 212–16
for Mozart's String Quartet K. 465, 155, 160, 167
conceptual slippage, 105–6, 112, 159, 183–84
conceptualization of music, 12, 14, 59–61, 67, 334
concords, 8–9, 11, 13, 15
Cone, Edward T., 333
analysis of Mozart Piano Sonata K. 331, 307–10, 323–24
on motivic derivation, 27–28, 54n48, 59
Confirmatio (*Bekräfftigung*), 295
Confutatio (*Wiederlegung*), 295
Conservatoire (Conservatoire National de Musique et de Déclamation), 300–301
consonant intervals. *See* concords
Cook, Nicholas, 245, 265, 332
Cooper, Grosvenor W., 307
*corps sonore*, 123, 127n59
Cowdery, James R., 217
Crawford, Richard, 204n7, 210, 214, 222
cross-domain mapping, 13–17, 19, 71, 73, 287–89, 291, 295–97, 326–27, 332–33
accounts of music and, 64, 76
between grammar and music, 291–92, 294–96, 299, 301
between language and music, 291–92, 299
between rhetoric and music, 295, 297–99
image-schema theory and, 68–69
metaphor theory and, 65–66
music theory and, 72–76
and musical categories, 330
role in conceptual models, 109–10
*See also* hierarchy; metaphor: contemporary theory of
cultural knowledge, 109–10, 115–16, 202–4, 219–20, 222–23, 240–42
cultural model, 110n20
cultural theory of information exchange, 225–29
Czerny, Carl, 305

Dahlhaus, Carl, 23–24, 49, 57–58
Damasio, Antonio R., 70, 108n16, 331
Davis, Miles, 223–25, 239–40
Davis, Ruth, 222
Debussy, Claude
*Reflets dans l'eau*, ix
Decoratio (*Schmückung* or *Zierde*), 295
Dever, William E., 223
*diezeugmenon*, 6
discords, 8–9, 11, 13, 15
Dispositio (*Einrichtung*), 295–96
dissonance, 25, 121
dissonant intervals. *See* discords
Dixon, Mort, 229
Dorsey, Jimmy (James), 210, 213–16
double-voiced discourse, 204, 224–25, 227–29, 239–41
Dylan, Bob (Robert Allen Zimmerman), 139n73

Eco, Umberto, 137–38, 154, 168–69, 199
Edelman, Gerald M., 60, 69, 108
Eichendorff, Joseph Freiherr von
"In der Fremde," 272–73, 276–77, 279, 282–84
*Viel Lärmen um nichts*, 272–73, 276
*Eingang* (Exordium), 295
*Einrichtung* (Dispositio), 295–96
*Einschnitt*, 293, 299
Elaboratio (*Ausarbeitung*), 295–96
embodied experience. *See* image schemata
enthymeme. *See* cadence: garbled
*Erfindung* (Inventio), 295
ethnobotanical classification, 32, 115–16
Executio (*Aufführung* or *Ausführung*), 295–96
Exordium (*Eingang*), 295

fake books, jazz, 209
*Fantasia* (Walt Disney movie), 94n47
Fauconnier, Gilles, 78, 82
Feld, Steven, 67, 116
Ferris, David, 276n28
Fétis, François-Joseph, 317
Fichte, Johann Gottlieb, 302–3
Fingerle, Marlene, 234

Finnegan, Ruth, 220
Fisher, Freddie "Schnickelfritz," 234
Fitzgerald, Ella, 236
Fodor, Jerry A., 29, 60n53
form, musical, 18, 288. *See also* grammar, and musical form; sonata form; theories of musical form
Förster, Friedrich, 246
Forte, Allen, 132
Foucault, Michel, 295n23, 298n36
frames, 41–44, 110n20
Freudenthal, Gideon, 312–13
Friedrich Wilhelm III (king of Prussia), 301

games, musical, 139–40
Garland, Judy (Frances Gumm), 212
Garland, Red (William M.), 223, 239
Gates, Henry Louis, Jr., 18, 204, 224, 228
Geertz, Clifford, 111, 201–2, 241–42
German universities, 301–2. *See also* music education: German
Gershwin, George, 204, 209, 213–16. *See also* "I Got Rhythm"
Gershwin, Ira, 204
Gibbs, Raymond W., Jr., 69
*Girl Crazy* (George and Ira Gershwin musical), 204–6, 212, 213n25
Gjerdingen, Robert O., 43n33, 168
Goetschius, Percy, 290–91, 305
Goodman, Benny (Benjamin), 212–15
Goodman, Nelson, 203n5
Gopnik, Alison, 117–18
grammar and music, 290, 292–95, 297–99. *See also* cross-domain mapping: between grammar and music
grammar, and musical form, 298–99
Graun, Carl Heinrich
  *Der Tod Jesu*, 297
Greater Perfect System, 6–7, 10–13
Greek
  music, 6, 12, 330
  theories of music, 6–16, 330
Griffin, Donald R., 60
Groth, Renate, 317
Guarini, Giovanni Battista, 84

Hammerstein, Oscar, 224
"Happy Birthday" ("Good Morning to All," by Mildred J. and Patty Smith Hall), 97–98
Harris, Paul L., 118n39
Hauptmann, Moritz, 320
Haydn, Joseph, 154, 174
head (in jazz performance), 210n15
Hegel, Georg Wilhelm Friedrich, 302
Heine, Heinrich
  "Im Rhein," 265–66, 272
  *Lyriches Intermezzo* (1823), 265, 272
Henderson, Fletcher, 210
Henderson, Ray, 229
Hensel, Luise, 246
Hensel, Wilhelm, 246
hierarchy
  atomistic, 310, 313, 320
    mapping on to music, 320–23
    structure of, 314–16
  chain-of-being, 310–13, 320
    mapping on to music, 316–18, 321–23
    structure of, 314–15
  models of, 289, 310, 314, 316, 324
  musical, 18–19, 59–60, 120, 128, 131, 306–8, 323–24
Hobbes, Thomas, 312
Hockenberry, John, 73n21
Hofstadter, Douglas R., 105
Holland, Dorothy, 109–10
Hood, Mantle, 220
"Hot Cross Buns" (folk song), 99
Hutchins, Edwin, 9n9, 115
*Hypate hypaton*, 6, 10
*Hypate meson*, 6, 10
*hypaton*, 6
*hyperbolaion*, 6

"I Got Rhythm" (by George and Ira Gershwin) 18, 203, **204–12**, 213–16, 218–22, 231, 233, 238
  recording history, 205–6, 208–12
  "Rhythm changes," *211*, 331
  tunes derived from, 210
"I Know That You Know" (by Vincent Youmans), 234
"I'm a Little Blackbird Looking for a Bluebird" (by Grant Clarke, Arthur Johnston, George W. Meyer, and Roy Turk), 237
idealized cognitive model, 110n20
"Im Rhein." *See* Heine, Heinrich; Schumann, Robert
image schemata, 68–71, 314–15, 322–23, 327
  and rhetorical figures, 138–39
  theory of, 68–70
"In der Fremde." *See* Brahms, Johannes; Eichendorff, Joseph Freiherr von; Schumann, Robert
India, music theory of, 13
intervals. *See* pitch relations, descriptions of
Invariance Principle, 70–71, 322–23
Inventio (*Erfindung*), 295
invention (process of), 295–97, 302–3

Jackendoff, Ray, 59
Java. *See* pitch relations, descriptions of
Jeff (experimental subject). *See* Bamberger, Jeanne: work with Jeff
Johnson, Mark, 65–66, 68
Johnson, Robert
  "Cross Road Blues," x–xi, 74n22
Jones, Philly Joe (Joseph Rudolph Jones), 223, 239

Kaluli (Papua New Guinea). *See* pitch relations, descriptions of
Kay, Paul, 226, 242
Keats, John
  "Ode to a Nightingale," 66
Keller, Hans, 61n57
Kerman, Joseph, 169, 176
Klein, Bernhard, 243
  "Trockne Blumen," **249–57**, 264–65, 286
knowledge structures, 110n20
Koch, Heinrich Christoph, 148, 152, 293–94, 297–98, 301
*Kopfton*, 127, 129–30n64, 261, 321
Kramer, Lawrence, 244
Krueger, Bennie, 230–32
Kuhn, Thomas S., 118
Kurth, Ernst
  characterization of tonality, 76

Lakoff, George, 65–66, 70, 111
Langer, Suzanne, 244
language and music, 5, 17, 61, 64, 84, 89, 220, 245, 285, 297, 326, 331. *See also* cross-domain mapping: between language and music
language, origins of, 284–85, 291
Lee, Peggy (Norma Deloris Egstrom), 224, 236–40
Leibniz, Gottfried Wilhelm, 313
*Leidensmotiv*. *See* Wagner, Richard: *Leidensmotiv* from *Tristan und Isolde*
Lerdahl, Fred, 59
Levy, Janet, 174, 177, 191
Lewin, David, ix, 98, 326n2
*Lichanos hypaton*, 6
Longinus, 138–39
Louis, Rudolf, and Ludwig Thuille
  characterization of tonality, 74–76, 83
Lully, Jean-Baptiste, 120
  recitative from *Armide*, *120*, *123*

McClary, Susan, x–xi, 59, 74n22
McNeill, David, 69
Macpherson, Stewart, 290n9
Macrobius, Ambrosius Aurelius Theodosius, 311
Maddox, Johnny, 236
Mahler, Gustav, 242
Marcello, Benedetto, 295
Marx, Adolph Bernhard (A. B.), 48, 177, 302–5
  definition of musical motive, 304
  *Die Lehre von der musikalischen Komposition*, 302–5
Mattheson, Johann, 292, 294–97, 333
  analysis of anonymous minuet, 292, *293*
Meltzoff, Andrew, 117–18
*Melvin and Howard* (Jonathan Demme movie), 239n78
mental models, 110n20
mental spaces, 78, 80, 82, 89, 331. *See also* conceptual domains: and mental spaces
Merman, Ethel (Ethel Agnes Zimmerman), 204–5, 213
*Mese*, 6
*meson*, 6
metaphor
  conceptual, 66–70, 300
  contemporary theory of, 14, 17, 65–70, 108
  linguistic, 66, 300
Meyer, Leonard B., 307–10, 321–23, 333
Milne, A. A. (Alan Alexander), 77
Minsky, Marvin, 112
modularity, theory of, 29–30
Momigny, Jérôme-Joseph de
  theory of tonality, 316–17
Monson, Ingrid, 220, 224–25, 240
Montessori bells, 96
Morgan, Robert P., 50n43, 52n47, 54nn48–49, 333
  analysis of Mozart K. 331 theme, 308–10, 321, 323
*motif*. *See* motive, musical
motive, musical, 3, 16, 18, 325
  as category, 16
  as dynamic construct, 289–90, 305
  Goetschius's definition of, 290
  Marx's definition of, 304
  and musical form, 289–91
  and process of derivation, 27–29
  Prout's definition of, 289–90
  Riemann's definition of, 305
  as static building block, 290
  *See also* theories of musical motive
Mozart, Wolfgang Amadeus, 18
  *Marriage of Figaro*, Overture, 301
  musical game (*see Musikalisches Würfelspiel*)
  Piano Sonata in A major, K. 331, first mvmt., **306–10**, 321–24
  String Quartet K. 465 ("Dissonance"), first mvmt., **154–68**, 199, 330
Müller, Wilhelm, 243, 285
  *Die schöne Müllerin*, 246–47, 272
  "Trockne Blumen," 247–49

music education
demand for instruction books, 291
French, 300–301, 305
German, 289, 302, 304–5
institutionalization of, 324
*Musica enchiriadis* (anonymous treatise), 291
musical notation, 67, 72, 330
Jeff's development of, 100–101, 103–4, 106–7
and the negotiation of conceptual models, 221–22
musical ontology, 18, 202–4, 214, 222, 241–42, 330
*Musikalisches Würfelspiel* (attributed to W. A. Mozart), **140–54**
organization of game, 141–43, 145, 148
template for compositional strategy, 149–54
"My Favorite Things" (by Oscar Hammerstein and Richard Rodgers), 224–25, 240

Napoleon I (Napoleon Bonaparte), 301
Narmour, Eugene, 62n60, 322n93
Narratio (*Bericht*), 295
*Naturklang*, 127, 131, 318
navigation in Polynesia and Micronesia, 115–16
Neoplatonic cosmology, 311–12, 318
Neoplatonism, 311
Newton, Isaac, 313
Nichols, Red (Ernest Loring Nichols), 206, 213–15
Nicomachus of Gerasa, 8–9
Norvo, Red (Kenneth Norville), 213–15
notation, musical. *See* musical notation

*obligate Lage*, 130–31
Odette (from Proust's *A la recherche*), 19, 325–27
Opéra (Académie Royale de Musique), 300
organicism, 113, 127, 130–31, 300, 302–5, 314n71, 317–18

Palestrina, Giovanni Perluigi da
*Pope Marcellus Mass*, Credo, 63–*64*, 73, 82–*83*
use of text painting, 63–64, 73–74, 77, 82–84
*Parhypate hypaton*, 6
Parker, Charlie (Charles), 211, 213–15
Pastoral Symphony. *See* Beethoven, Ludwig van: Sixth Symphony
perceptual symbols, 70, 111n21
Peroratio (*Schluss*), 295
*Pete Kelly's Blues* (Jack Webb movie), 236
pitch relations, descriptions of
Aristoxenus, 11, 14, 63
Bali, 67–68, 71–72
basis in image schemata, 69
Greek, 14, 63, 67
Java, 67–68, 71–72
Kaluli (Papua New Guinea), 67–68, 116–17
Suyá (of the Amazon basin), 68
pitch relations, musical notation and, 67, 72
pitch-event, 101–5, 107, 110, 112
Plato, 10, 311
Plotinus, 311
Pollard-Gott, Lucy, 43n33
Porter, James, 217
Powers, Richard, 218n34
Propositio (*Antrag*), 295
*Proslambanomenos*, 6, 10
Proust, Marcel, 16
*A la recherche du temps perdu*, 3–4, 81–82, 325–27, 334
aesthetic theory, 328
Prout, Ebenezer, 289–91, 305
Ptolemy, 15
Pythagoras of Samos, 7–9, 12, 311n62
Pythagorean music theory, 7–8, 10, 12–14, 16

Quinault, Philippe, 120
Quinn, Naomi, 109–10

Rama Indians (Nicaragua)
snakebite cure, 115–16
Rameau, Jean-Philippe, 17, 119, 285, 330, 333
analysis of recitative by Lully, 120–24
theory of music, 121–24, 132, 148
theory of music as a cognitive theory, 124–26
ratio (applied to music), 8–10, 13, 15
epimeric, 13, 15
epimoric, 13, 15
multiple, 13, 15
Ratner, Leonard G., 168–69, 179–80, 200
Redman, Don, 208, 213–15
Reicha, Antoine, 301–2, 305
Réti, Rudolph
theory of motive, 26–27
rhetoric, 139n74, 225, 228, 292–93, 303
and music, 296–98
musical, 288–89, 324
and musical form, 298–99
*See also* cross-domain mapping: between rhetoric and music
rhetorical figures, 138–39, 295–96, 299
Rich, Fred, 208, 213–15
Richardson, Samuel, 291
Richter, Jean Paul, 303
Riemann, Hugo, 305
characterization of tonality, 76
definition of musical motive, 305
theory of meter and rhythm, 320
Riepel, Joseph, 148, 152, 292–94, 299
Robertson, Jeanne, 217
Rodgers, Richard, 224

Rooney, Mickey (Joe Yule, Jr.), 212
Rosch, Eleanor, 16, 32–33, 41
Rousseau, Jean-Jacques
  analysis of recitative by Lully, 122
  on the origins of language, 284–85

Saint-Saëns, Camille
  *Le Carnaval des animaux*, 71
*Satz*, 48n40, 293
Schachter, Carl, viii
Schank, Roger S., 112
Schelling, Friedrich von, 288, 302–3
Schenker, Heinrich, 17, 119–20, 331, 333
  analysis of Schubert's waltz Op. 77 No. 1, 128–30
  theory of hierarchy, 317–18
  theory of music, 127–30
  theory of music as a cognitive theory, 130–32
  theory of musical motive, 26–27
Schiller, Friedrich, 302
Schlegel, August Wilhelm, 303
Schleiermacher, Friedrich, 302
*Schluss* (Peroratio), 295
*Schmückung* (Decoratio), 295
Schoenberg, Arnold, 242, 333
  characterization of tonality, 76
  theory of musical motive, 25–27, 30, 34, 42, 57–59, 139, 154
Schöll, Adolf, 273
Schopenhauer, Arthur, 328
Schubert, Franz, 243, 332
  "Der Wanderer" (Op. 4/1, D. 489), 28, *29*
  "Trockne Blumen," from *Die schöne Müllerin* (D. 795), **257–64**, 265, 332
  *Valses nobles*, Op. 77 No. 1, 128–*29*, 130
  *Wanderer Fantasie* (Op. 15), 27–28, *29*
Schumann, Robert, 265, 272–73, 279
  "Der frohe Wandersmann" (Op. 77 No. 1), 276
  *Dichterliebe* (Op. 48), 265
  "Im Rhein" from *Dichterliebe*, **266–72**, 278, 286
  "In der Fremde" (Op. 39 No. 1), **273–79**, 281, 286
  *Liederjahr* of 1840, 265, 284
Schutt, Arthur, 234
*Sectio canonis* (anonymous treatise), 13, 15
Seeger, Charles, 221
semantics, musical, 17–18, 138, 199–200, 329
  musical syntax and, 54, 199–200, 299, 329
semiotics, musical. *See* semantics, musical
Short, Bobby, 237–39
Signifyin(g), 18, 224, 229, 240–41
Skolsky, Sidney, 235
*Sleepless in Seattle* (Nora Ephron movie), 239n78
sonata form, 152–53
Spaun, Joseph von, 243, 285
Stägemann, Friedrich August von, 246
Stägemann, Hedwig von, 246
Sulzer, Johann Georg, 296–98, 302–4, 313n67
Suyá (of the Amazon basin). *See* pitch relations, descriptions of
Swann, Charles (from Proust's *A la recherche*), 3–5, 16, 18–19, 61, 81, 325–28, 333
  theory of music, 332
Sweetser, Eve E., 225–27, 242
syllogismus. *See* cadence: complete
symbols, 72, 103, 106. *See also* perceptual symbols
synaesthesia, 89n38
syntax, musical, 17–18, 52, 54, 75, 95, 132–33, 137–38, 140, 199–200
  in Beethoven's String Quartet Op. 18 No. 1, 189, 191, 196, 198
  in Beethoven's String Quartet Op. 18 No. 6, 172–73, 176
  dynamic approach to, 329
  formal, 143, 167, 199, 329
  harmonic, 124–25, 130, 146, 153–54, 198–200, 329, 331
  melodic, 140, 154, 159, 167–68, 331
  in Mozart's String Quartet K. 465, 154, 159, 167–68
  multiple levels of, 198
  musical categories and, 51–52, 54, 61–62, 329
  and the *Musikalisches Würfelspiel*, 151–54
  Rameau's theory of music and, 121–22, 124–25
  and rhetoric, 54n48
  Schenker's theory of music and, 127–28, 130–31
  and song texts, 254–55, 284–86
  thematic, 198, 200

Talmy, Leonard, 69
tetrachords, 6
*tetraktys* of the decad, 9
text painting, 17, 63–65, 73–74, 77, 82–84, 87–88, 95
theme, musical
  conceptual models and, 47–48
theories, 5, 15–17, 62, 98–99, 102–3, 108–9, 119, 125, 131, 133, 317, 332–33
  ad hoc, 132
  cultural, 114–17
  extended, 132
  function and structure, 106–7, 111, 113–14
  implicit, 116–17
  scientific, 113–14, 117
  of very young children, 117–19
  *See also* conceptual models: and theories
theories of meter and rhythm, 307–10, 319–23

theories of music, 5, 15, 121–24, 127–30, 132, 148, 323–24, 332, 334
as cognitive theory, 119, 124–26, 130–34
of early India, 13
Greek, 6–16
Jeff's development of, 101–2, 104–7
and syntax, 121–22, 124–25, 127–28, 130–31
*See also* Aristoxenian music theory; Pythagorean music theory
theories of musical form, 300–306
theories of musical motive, 25–30, 34, 42, 57–59, 139, 154, 289–90, 304–5
theories of tonality, 74–76, 83, 316–18
Thompson, William (Big Bill), 223
Thuille, Ludwig. *See* Louis, Rudolf, and Ludwig Thuille
*Ton*, 316–17. *See also* Momigny, Jérôme-Joseph de: theory of tonality
tonality
characterizations of, 76
as metaphysical, 317
*See also* theories of tonality
topics, musical, 168, 174, 176–77, 184, 198, 200, 329
topography. *See* conceptual integration networks (CINs)
Tovey, Donald Francis, 90–94, 327, 331, 333
"Trockne Blumen." *See* Klein, Bernhard; Müller, Wilhelm; Schubert, Franz
Tunisia, music of, 222
Turner, Mark, 70, 78
on rhetorical figures, 138–39
"Twinkle, Twinkle Little Star" (folk song), *100*
Jeff's work with, 99–108

University of Berlin, 302
University of Halle, 301
University of Jena, 302
*Urlinie*, 127, 130–31
*Ursatz*, 127–31, 318, 321

*Verwandlung*, 128
Vinteuil's sonata (from Proust's *A la recherche*), 3–4, 16, 18–19, 325, 327–28, 332–33

Wagner, Richard
compositional strategy, 49–51, 58
*Leidensmotiv* from *Tristan und Isolde*, 49–52, 54, 57–59, 62, 138
*Sehnsuchtsmotiv* from *Tristan und Isolde*, 49
*Tristan und Isolde*, 23–24, *24*, **49–58**
Webb, Chick (William Henry Webb), 209, 213–15
Webb, Jack, 236
Weber, Gottfried, 319–20
Welker, Robert L., 43n33
Wert, Giaches de
"Tirsi morir volea," 84, *85–86*, 87–89
Wieck, Clara, 273
*Wiederlegung* (Confutatio), 295
Will, Richard, 90
Wittgenstein, Ludwig, 30
Wolzogen, Hans von, 49
work of music. *See* musical ontology

Young, Lester, 211, 213–15

Zemlinsky, Alexander von, 25
*Zierde* (Decoratio), 295

Printed in the United Kingdom
by Lightning Source UK Ltd.
103748UKS00001B/154



"Zbikowski has written a book that bears comparison to some of the great works of psychologically oriented, humanistic music theory, including Meyer's *Emotion and Meaning in Music* and Kurth's *Musikpsychologie*."
—Robert Gjerdingen, Professor of Music, Northwestern University

ALSO OF INTEREST

*Tonal Pitch Space*
Fred Lerdahl

*Musical Perceptions*
Edited by Rita Aiello, with John A. Sloboda

*Analysing Musical Multimedia*
Nicholas Cook

*Aural Skills Acquisition*
Gary S. Karpinski

ISBN 0-19-514023-0
90000
9 780195 140231

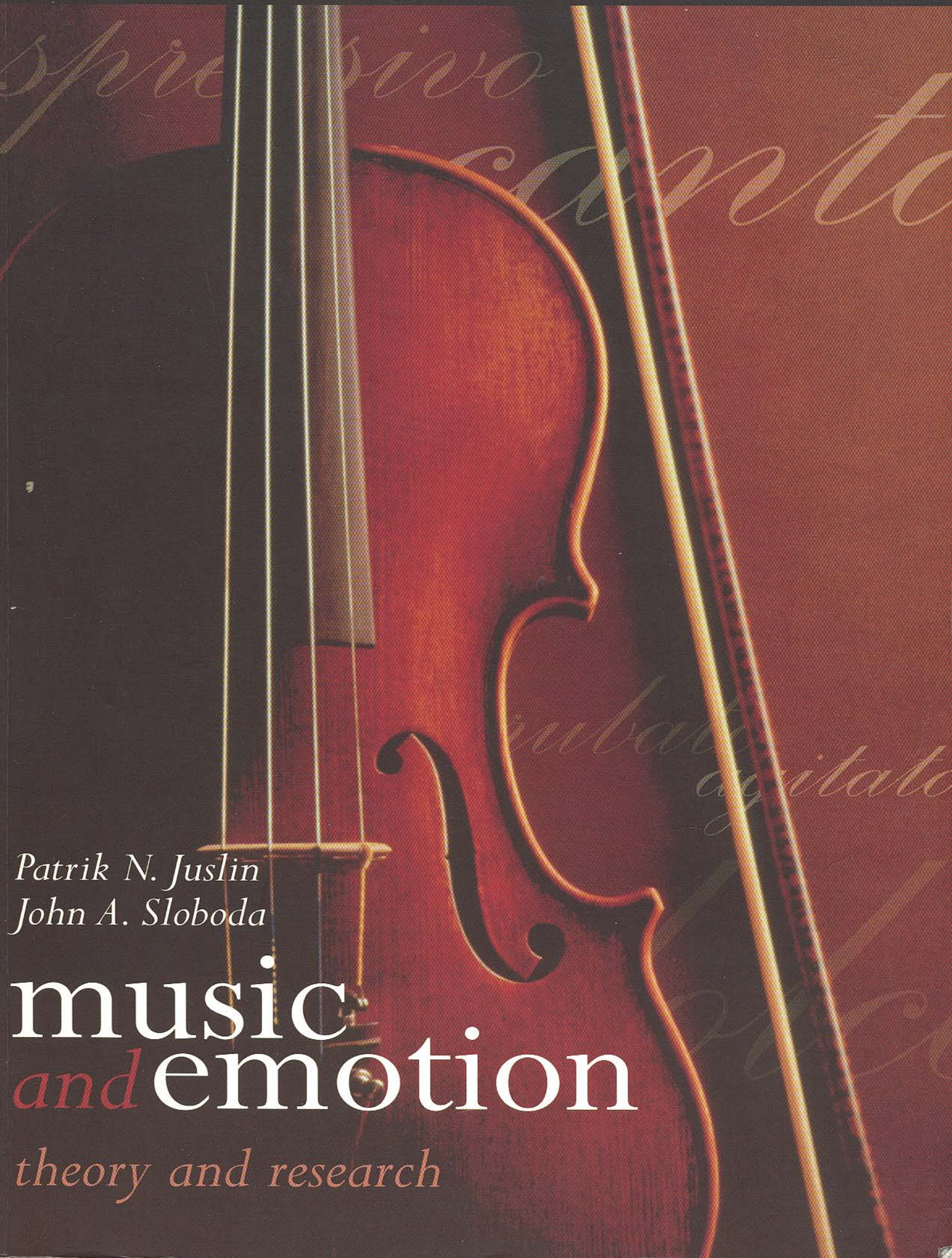
EL ARGONAUTA
la librería de la música
0-19-514023-0
ZBIKOWSKI, LAWRENCE M.
CONCEPTUALIZING MUSIC.
P.V.P. (euros): 51.80

Zbikowski

# Conceptualizing Music

OXFORD

SERIES IN AFFECTIVE SCIENCE



*Patrik N. Juslin*
*John A. Sloboda*

# music and emotion

## *theory and research*

# MUSIC AND EMOTION

## Series in Affective Science

*Series Editors*
Richard J. Davidson
Paul Ekman
Klaus R. Scherer

*The Nature of Emotion: Fundamental Questions*
edited by Paul Ekman and Richard J. Davidson

*Boo!: Culture, Experience, and the Startle Reflex*
by Ronald Simons

*Emotions in Psychopathology:*
*Theory and Research*
edited by William F. Flack Jr. and James D. Laird

*What the Face Reveals:*
*Basic and Applied Studies of*
*Spontaneous Expression Using the*
*Facial Action Coding System (FACS)*
edited by Paul Ekman and Erika Rosenberg

*Shame: Interpersonal Behavior,*
*Psychopathology, and Culture*
edited by Paul Gilbert and Bernice Andrews

*Affective Neuroscience: The Foundations of*
*Human and Animal Emotions*
by Jaak Panksepp

*Extreme Fear, Shyness, and Social Phobia:*
*Origins, Biological Mechanisms, and*
*Clinical Outcomes*
edited by Louis A. Schmidt and Jay Schulkin

*Cognitive Neuroscience of Emotion*
edited by Richard D. Lane and Lynn Nadel

*The Neuropsychology of Emotion*
edited by Joan C. Borod

*Anxiety, Depression, and Emotion*
edited by Richard J. Davidson

*Persons, Situations, and Emotions: An Ecological Approach*
edited by Hermann Brandstätter and Andrzej Eliasz

*Emotion, Social Relationships, and Health*
edited by Carol D. Ryff and Burton H. Singer

*Music and Emotion: Theory and Research*
edited by Patrik Juslin and John Sloboda

# MUSIC AND EMOTION

*Theory and research*

Edited by

PATRIK N. JUSLIN,

*Uppsala University, Sweden*

and

JOHN A. SLOBODA

*University of Keele, UK*

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Athens Auckland Bangkok Bogotá Buenos Aires Cape Town
Chennai Dar es Salaam Delhi Florence Hong Kong Istanbul Karachi
Kolkata Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
Paris São Paulo Shanghai Singapore Taipei Tokyo Toronto Warsaw

and associated companies in Berlin Ibadan

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press, Inc., New York

© Oxford University Press, 2001

The moral rights of the author have been asserted

Database right Oxford University Press (maker)

First published 2001

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above.

You must not circulate this book in any other binding or cover
and you must impose this same condition on any acquirer

A catalogue record for this title is available from the British Library

Library of Congress Cataloging in Publication Data

Music and Emotion : theory and research / edited by Patrik N.Juslin and John A. Slodoba
(series in affective science)
Includes bibliographical references and index.
1. Music–Psychological aspects. 2. Emotions. I. Juslin, Patrik N II. Sloboda, John A. III. Series.
ML3830 .M965 2001 781.11'038—dc21 2001036068

ISBN 0–19–263188–8 (Pbk)
ISBN 0–19–263189–6 (Hbk)

10 9 8 7 6 5 4 3 2 1

Typeset by J&L Composition Ltd, Filey, North Yorkshire
Printed in Great Britain
on acid free paper by
Biddles Ltd., Guildford & King's Lynn

# CONTENTS

List of contributors *page vii*
Acknowledgments *page ix*
Foreword *page x*

## Introduction

1 Music and emotion: introduction *page 3*
*Patrik N. Juslin and John A. Sloboda*

## Multidisciplinary Perspectives

2 Philosophical perspectives on music's expressiveness *page 23*
*Stephen Davies*

3 Musicological approaches to emotion *page 45*
*Nicholas Cook and Nicola Dibben*

4 Psychological perspectives on music and emotion *page 71*
*John A. Sloboda and Patrik N. Juslin*

5 Listen to the brain: a biological perspective on musical emotions *page 105*
*Isabelle Peretz*

6 Anthropological perspectives on music and emotion *page 135*
*Judith Becker*

7 Aesthetic agency and musical practice: new directions in the sociology of music and emotion *page 161*
*Tia DeNora*

8 Music and emotion: perspectives from music therapy *page 181*
*Leslie Bunt and Mercédès Pavlicevic*

## The Composer

9 Emotion and composition in classical music: historiometric perspectives *page 205*
*Dean Keith Simonton*

10 The influence of musical structure on emotional expression *page 223*
*Alf Gabrielsson and Erik Lindström*

11 Music as a source of emotion in film *page 249*
*Annabel J. Cohen*

## The Performer

12 The subjective world of the performer *page 275*
*Roland S. Persson*

13 Negative emotions in music making: the problem of performance anxiety *page 291*
*Andrew Steptoe*

14 Communicating emotion in music performance: a review and a theoretical framework *page 309*
*Patrik N. Juslin*

## The Listener

15 Music and emotion: distinctions and uncertainties *page 341*
*Leonard B. Meyer*

16 Emotional effects of music: production rules *page 361*
*Klaus R. Scherer and Marcel R. Zentner*

17 Continuous measurement of self-report emotional response to music *page 393*
*Emery Schubert*

18 Emotions in everyday listening to music *page 415*
*John A. Sloboda and Susan A. O'Neill*

19 Emotions in strong experiences with music *page 431*
*Alf Gabrielsson*

## Postlude

20 Music and emotion: commentary *page 453*
*John A. Sloboda and Patrik N. Juslin*

Author index *page 463*

Subject index *page 475*

# LIST OF CONTRIBUTORS

**Judith Becker** School of Music, University of Michigan, 1100 Baits Drive, Ann Arbor, Michigan 48109 – 2085, USA

**Leslie Bunt** Faculty of Health and Social Care, St. Matthias Campus, University of the West of England, Bristol BS16 2UP, UK

**Annabel J. Cohen** Department of Psychology, University of Prince Edward Island, Charlottetown PE C1A 4P3, CANADA

**Nicholas Cook** Department of Music, University of Southampton, Highfield, Southampton SO17 1BJ, UK

**Stephen Davies** Department of Philosophy, University of Auckland, Private Bag 92019, Auckland, NEW ZEALAND

**Tia DeNora** Department of Sociology, University of Exeter, Amory Building, Rennes Drive, Exeter EX4 4RJ, UK

**Nicola Dibben** Department of Music, University of Sheffield, Sheffield S10 2TN, UK

**Alf Gabrielsson** Department of Psychology, Uppsala University, Box 1225, SE – 751 42 Uppsala, SWEDEN

**Patrik N. Juslin** Department of Psychology, Uppsala University, Box 1225, SE – 751 42 Uppsala, SWEDEN

**Erik Lindström** Department of Psychology, Uppsala University, Box 1225, SE – 751 42 Uppsala, SWEDEN

**Leonard B. Meyer** Department of Music, University of Pennsylvania, 201 SO. 34th Street, Philadelphia, PA 19104, USA

**Susan A. O'Neill** Department of Psychology, University of Keele, Newcastle, Staffordshire ST5 5BG, UK

**Mercédès Pavlicevic** Music Therapy Programme, Department of Music, University of Pretoria, Pretoria 0002, SOUTH AFRICA

**Isabelle Peretz** Department of Psychology, University of Montreal, CP 6128, Succ. Centre-ville, Montreal, Québec H3C 317, CANADA

**Roland S. Persson** School of Education and Communication, Jönköping University, Box 1026, SE – 55 111 Jönköping, SWEDEN

**Klaus R. Scherer** Department of Psychology, University of Geneva, 40 Boulevard du Pont d'Arve, CH – 1205 Geneva, SWITZERLAND

**Emery Schubert** School of Music and Music Education, University of New South Wales, Sydney, NSW 2052, AUSTRALIA

**Dean Keith Simonton** Department of Psychology, University of California, Davis, CA 95616 – 8686, USA

**John A. Sloboda** Department of Psychology, University of Keele, Newcastle, Staffordshire ST5 5BG, UK

**Andrew Steptoe** Department of Epidemiology and Public Health, University College London, London WC1E 6BT, UK

**Marcel R. Zentner** Department of Psychology, University of Geneva, 40 Boulevard du Pont d'Arve, CH – 1205 Geneva, SWITZERLAND

# ACKNOWLEDGMENTS

The editors would like to thank the authors, all of whom entered into this project with great enthusiasm, professionalism, and efficiency. They did more than write their individual contributions. They willingly engaged with other authors and ourselves to develop and refine a coherent and cross-referring set of papers which, as a result, is considerable greater than the sum of the parts. Crucial to this process was the contribution of a number of referees. Each chapter was reviewed by a least two experts in the field, and their perceptive comments helped authors to refine and update their contributions. Among those who participated in the review process were Emmanuel Bigand, Anne Blood, Warren Brodsky, Claudia Bullerjahn, Alfred Cramer, Ian Cross, Jane Davidson, Walter Dowling, Andrew Gregory, David Hargreaves, Antoine Hennion, Anthony Kemp, Carol Krumhansl, Andreas Lehmann, Jerrold Levinson, Justin London, Clifford Madsen, Peter Martin, Eugene Narmour, Bruno Repp, John Rink, Martin Stokes, William Forde Thompson, Magnus Widesheim, Suzanna Widmer, and Dolf Zillman. In addition, several of the authors provided comments on parts of the book.

We are particularly grateful for the help and support of colleagues at Oxford University Press, for receiving the initial proposal warmly, and encouraging the project at all stages. Martin Baum, our Editor, has taken a close personal interest in all aspects of the project, and his constant attentiveness and readiness to act on behalf of the project has meant a great deal to us.

We acknowledge the support of our respective institutions. Patrik Juslin's work was supported by the Bank of Sweden Tercentenary Foundation. John Sloboda was enabled to devote time and resources to the project by the award of a semester of sabbatical leave under the Keele University Research Investment Scheme.

Finally, we wish to acknowledge two individuals, now retired from full-time academic positions, without whom this project would have been inconceivable. Leonard B. Meyer's brilliant work in the 1950s is the foundation on which almost everything else in this book rests, directly or indirectly. We are delighted that he has supported this project with a contribution which is as fresh and innovative as anything that he has written. Alf Gabrielsson is, in many ways, the father of modern psychological investigations into musical emotions. He was publishing pioneering empirical work relating to emotion and musical expression in the early 1970s, long before any other contributor to this book. He has been profoundly influential for many researchers, not least the editors of the current volume—Juslin through many years of direct collaboration at the Uppsala University, and Sloboda through two decades of collegial support in activities such as the organization of scientific meetings, the nurturing of scholarly scientific bodies, and tireless contributions to the editorial work of books and journals. We salute these two giants, and hope that this book will add to the recognition of their important contributions to the furthering of knowledge.

# FOREWORD

It has often been claimed that music is the language of the emotions. One can interpret this claim in either of two senses. On the one hand, it could be taken to mean that music provides an iconic representation of the affective 'movements of the soul' which is more appropriate to the dynamic flow of emotional experience than symbolic description through verbal labels (which imply steady states). On the other hand, the claim that music is the language of the emotions can also be understood as commentary on the nature of music, specifically the suggestion that music cannot be composed, performed, or listened to without affective involvement. In consequence, the privileged relationship between emotion and music has been the object of much scholarly discussion as well as empirical research. Not surprisingly, then, the topic has been addressed by many different disciplines - philosophy, musicology, literature, psychology, acoustics, anthropology, and neurosciences - as well as by practising musicians, composers, and performers alike.

Given its multifaceted nature, the relationship between music and emotion is a topic of central concern for the emerging domain of affective science. The current volume, edited by Patrik Juslin and John Sloboda, will have a major role in establishing the topic as a core research area in this multidisciplinary domain of scientific inquiry. It attempts no lesser feat than bringing together the various concerns with the affective nature of music that have so far been spread over a wide variety of journals and book publications in different disciplines. For all practical purposes, this material has been inaccessible to the non-specialist. In addition, the volume provides a state-of-the-art overview of current thinking and a synthesis of the major research results to date. The editors succeed admirably on both counts.

Following an overview of the various perspectives that have been adopted by the disciplines concerned with the emotional expressiveness and the affective impact of music, the contributions are grouped with respect to their focus on the composer, the performer, and the listener. In an extensive introduction, that helps to structure the issues and define fundamental concepts, as well as in a final commentary, Juslin and Sloboda pull together the various strands and work toward an integration of the different themes. Among the contributors, all established experts in their respective areas, are two of the eminent pioneers in the study of music-emotion relationships - Alf Gabrielsson and Leonard Meyer. This volume is a fitting tribute to their vision.

Given the multitude of perspectives provided in this book and the ensuing enlargement of the horizon for scholars and researchers in the highly specialized disciplines concerned with this topic, one can expect a quantum jump in the sophistication of affective science research on music and emotion. It is equally to be hoped that the fascinating questions and hypotheses detailed in the pages of this volume will recruit a new generation of multidisciplinary-minded researchers to the study of a phenomenon that has intrigued mankind since the dawn of time.

Klaus R. Scherer

PART I

# INTRODUCTION

CHAPTER 1

# MUSIC AND EMOTION: INTRODUCTION

PATRIK N. JUSLIN AND JOHN A. SLOBODA

Questions about music and emotion have occupied human beings ever since antiquity. It is a topic of considerable interest to laymen and experts alike. Some sort of emotional experience is probably the main reason behind most people's engagement with music. Emotional aspects of music should thus be at the very heart of musical science. It is all the more remarkable then that this topic has been seriously neglected during the last decades. None of the recent books on music psychology, or emotion psychology, have treated emotional aspects of music other than sparingly, despite the fact that some emotion theorists have argued that studies of music might be helpful also in understanding emotions more generally (Gaver & Mandler 1987). The most authoritative handbook of music psychology to appear thus far, *The psychology of music* (Deutsch 1999), does not include a chapter on emotion. Similarly, the most extensive handbook on emotion yet to appear, *Handbook of emotions* (Lewis & Haviland-Jones 2000), also does not contain a chapter on music.

Still, it is probably true that most people experience music—somehow, somewhere—every day of their lives, often with an accompanying affective response of some sort (e.g. nostalgic recognition of a favourite song on the radio while driving a car, frustration directed at the music at the shops, joy when listening to an excellent performance at an evening concert, a sad mood created by the soundtrack of a late night movie). As noted by Dowling and Harwood (1986), 'music arouses strong emotions in people, and they want to know why' (p. 202). However, despite the ubiquity of emotional responses to music, it seems that, for a long time, such reactions have defied psychological explanation.

Without wishing to forestall the different conceptualizations of emotion in this book, it may be useful to provide a preliminary description of how affect is differentiated from other psychological phenomena that occur in relation to music. When an individual engages with music, either as performer, listener, or composer, a very broad range of mental processes and contents may be engaged. The present book focuses on those musical processes and mental contents that have an *evaluative* component. We contrast such phenomena to those processes and contents that are primarily descriptive or *representational.*

Examples of representational processes would include the determination, or awareness, of such properties of a piece of music as metre, rhythm, tonality, harmony, melody, form, and style (whether accomplished explicitly or implicitly). Such representational processes are central to the recognition, identification, and performance of

music. They are the focus of the study of music perception and cognition, and, as such, have dominated the scientific literature on musical activity for the last century or more (Deutsch 1999; Sloboda 1985).

Examples of evaluative processes would include the determination, or awareness, of music as eliciting liking or disliking; preference; emotion and mood; and aesthetic, transcendent, and spiritual experiences. We use the term evaluative because it is typical of such processes that they are both *valenced* and *subjective*. Valence refers to the tendency to attach differing degrees of positivity or negativity to whatever is being experienced or processed. Subjectivity refers to the fact that such experiences are affected by labile personal factors such as attitudes, associations, and goals. There is, of course, no absolute dichotomy between representational and evaluative processes. Rather they represent extremes of a continuum. Representation can incorporate evaluative components (particularly in situations of ambiguity, with which music abounds). For instance, hearing the lower rather than the upper line of a two-part invention as 'foreground' may involve preference grounded in subjectivity (this particular listener perhaps prefers focusing on lower parts because he or she is a bass player). Affect can also incorporate representational aspects, especially where some aspect of what is being experienced is deeply embedded culturally or biologically. For example, listeners within a particular culture may have no option but to hear organ music as 'churchy'—with all the emotional connotations that brings—thus appearing to be an 'objective' characteristic of the music. One of the key issues for musical science is to better understand the relationship between the representational and affective processes. Do affective responses always take representations as their starting point, and if so, what other factors moderate between representation and outcome? Unfortunately, musical affect has been the poor relative of contemporary musical science (Sloboda 1995).

## 1.1 The slow progress in the study of music and emotion

Why has the field of music and emotion been so slow to catch up with the other domains of musical science, such as music cognition or musical development? We believe that there are several reasons for this neglect—some that the domain of music and emotion has in common with the field of emotion as a whole, and some that are particular to the music domain. One problem, which has impeded progress in understanding emotions in general, is the difficulty involved in studying emotions in the laboratory (Plutchik 1994, Chapter 1). This is also true of emotional reactions to music, which are difficult to observe under laboratory conditions.

Another reason for the slow progress in the field of music and emotion is the dominant influence of so-called *cognitive science* (Gardner 1985) on psychology in general and music psychology in particular, which has led to a strong emphasis on cognitive aspects of musical behaviour—both in perception (Deutsch 1999; Dowling & Harwood 1986) and performance (Gabrielsson 1999; Palmer 1997). This is illustrated through the titles of some of the main scholarly organizations for music that exist around the world. Most of them contain the word 'cognition'. None contain the word 'emotion' (e.g. Society for Perception and Cognition of Music, European Society for the Cognitive Sciences of Music, International Conference for the Perception and Cognition of

Music). And whereas studies of music cognition have been guided by a common 'paradigm'—the information-processing approach—the study of music and emotion has not yet benefited from any unifying paradigm. Indeed, a number of authors have expressed doubts that current theories of emotion are adequate for dealing with music. Thus it seems to us that the study of musical emotion is pretty much in its 'pre-paradigmatic' phase, to borrow from Kuhn's (1962) terminology. There is still plenty of disagreement in the field about many issues, and at present, it seems that the field is best served by a multiplicity of approaches. A lack of a unifying theoretical framework may explain why no psychologist has attempted to write a book on the topic: it has been too difficult to present a coherent story based on the available literature. However, one important goal for future research could be to try to reach some degree of consensus concerning a common set of concepts, definitions, and methodologies for the study of music and emotion. Indeed, we hope that this very book may contribute to that goal by summarizing a variety of theoretical perspectives and methods in a systematic fashion, highlighting areas of agreement whenever possible.

One reason why many researchers have been reluctant to study musical emotions is that it is a very difficult topic. Most of the early studies in the field were concerned with matching emotion labels to excerpts of music, and for about 50 years, it appeared as if the field would not move beyond such studies. There was no sense of direction in this early work, despite the pioneering studies by Kate Hevner (1935) and others. Perhaps frustrated by a lack of theories and by some difficulties inherent in describing responses to music in words, many researchers simply abandoned the field. Part of the reason for this was also the behaviourist era of the 1940s and 1950s, during which the study of emotions and other 'inner' mental processes were seen as less than scientific. Since then, we believe, there has been a tendency for many researchers to think that the problem of music and emotion is so complex that they would prefer to avoid it altogether—rather than dealing with the problem inadequately, it is better not dealing with it at all. This view has, of course, not improved the state of affairs. As a result, there has been limited progress in the field.

A final reason for the neglect of emotion comes from the fact that much academic study of music takes as paradigmatic a very particular way of listening to music and talking about it that is enshrined in the narrow 'classical concert culture' where, as described by Frith (1996) and others (Cook 1998; Small 1999), audiences are taught to listen 'silently and respectfully' with minimum bodily movement or emotional expression (until the end). 'Appreciation' of music is often taken to mean having an intellectual understanding of the history and form of the musical composition, rather than an articulated emotional response. Even where emotions are valued, they tend to be those rarefied (transcendent or spiritual) forms that are related to 'higher' abstract and aesthetic properties of works, rather than the everyday or full-blooded emotions. As a consequence, there have been few academic discourses available in which to frame and articulate a relevant understanding of the emotions. Such discourses have mainly been relegated to 'the discourse of the hallway' (e.g. Frith 1996), those informal 'off-duty' moments when academics allow themselves to say what they feel about the music they study. It is in the context of increasingly viewing the whole music of the world's cultures as worthy of study (including taking account of the discourses and practices surrounding music making in these diverse

cultures) that the possibility of making emotion more central to disciplined talk about music may be increasingly enabled.

## 1.2 Purpose and organization of this book

The present book is, to the best of our knowledge, the first scientific anthology ever devoted specifically to musical emotions. Why should it appear now? Along with a renewal of interest in emotional phenomena in general (e.g. Ekman & Davidson 1994; Lewis & Haviland-Jones 2000) there has been a new surge of interest in music and emotion. Music and emotion is once again at the forefront of music psychology, as reflected in innovative and exciting research all over the world. After a long time of regarding music and emotion as an 'off-limits' topic, scientists are returning to the field to address what may be the most exciting aspect of musical behaviour: its affective consequences. Today, it is almost becoming fashionable again to study emotion. For instance, in August 2000, 94 out of 310 papers (30 per cent) presented at the Sixth International Conference of Music Perception and Cognition (Woods *et al.* 2000) contained the word 'emotion' in the text (as verified by applying a search engine to the CD rom version of the proceedings). By contrast, visual inspection of the proceedings of the First International Conference in the same series (Japanese Society of Music Perception and Cognition 1989) yields only 11 out of 94 papers (12 per cent). The peak of this renaissance is yet to come, we believe.

This book can be characterized by two words. The first word is 'change'. The concept of emotion is intimately related to changes—in particular to significant changes in the social environment (cf. Oatley 1992). Music, too, is about changes. Music takes place in time, in a constant flux. Emotional responses to music are often induced by particular kinds of changes in the music. The field of music and emotion itself has gone through some dramatic changes over the last decade. One of the most significant changes is the increasing diversity of topics, approaches, and methods involved. Another significant change is the increasing concern with theories of music and emotion. Both of these trends are evident in the present volume.

The second word that can be used to characterize this book is 'interface'. Emotions have been conceptualized as the interface between the inner world of goals and plans, and the outer world of objects and events (cf. Scherer 1985). Emotion has also been described as the point where nature meets nurture (Oatley 1992, p. 14)—this is evident in many of the contributions to this volume. Moreover, much of what is interesting about music and emotion takes place at the interface between performance and perception (Juslin 2000). Last but not least, this book represents an interface between researchers from different disciplines. Recent years have seen some important developments in diverse approaches to music and emotion.

Unfortunately, current work is largely scattered across different publications. It is thus difficult to get a sense of the progress in the field, or to grasp what questions researchers are currently debating. The purpose of the present book is to bring together some of the leading researchers of the field for an up-to-date look at the relationship between music and emotion. Fundamental questions addressed by this book include: Why does music induce emotions in its listeners? Are the emotions we experience in

relation to music different from the emotions we experience in everyday life? Why are different pieces of music associated with different emotions? Are performers able to communicate specific emotions to listeners? Do emotional responses to music vary as a function of the cultural context? How do emotional responses to music affect the brain and body of listeners?

The purpose of the book is both summative and formative. We aim to summarize what is currently known about music and emotion, but we also aim to stimulate further research in promising directions that have been little researched. The focus is on theory and research on music and emotion rather than practical applications. We believe that, with a few exceptions in the context of music therapy, performance anxiety, and film music, the field of music and emotion has not yet reached a stage where it can provide a consistent body of knowledge for application to a wide range of practical problems (this is incidentally another reason why we prefer to think of the field as being in a 'pre-paradigmatic' phase). However, we believe that the present book covers every major approach to the study of music and emotion.

This volume is organized into four sections following this introduction. The first section provides multidisciplinary perspectives on music and emotion from philosophy, musicology, psychology, biology, anthropology, sociology, and therapy. The next three sections of the book provide research reviews that focus on the composer, the performer, and the listener, respectively. Although, in practice, most musical situations involve some combination of composition, performance, and listening, research has largely tended to focus on one or two of these aspects. (Most research has focused on the listener, and the least on the composer.) Therefore, we have assigned the remaining chapters into sections according to whether their primary focus lies on the composer, the performer, or the listener.

### 1.2.1 Multidisciplinary perspectives

Arguably, emotion is first and foremost a psychological concept (e.g. given the 'mentalistic' nature of the construct and its relation to behaviour). It is not surprising then that the majority of studies on music and emotion to date are psychological. However, the study of music and emotion often transcends the traditional lines of enquiry, and we believe that much can be learned from other perspectives as well. The chapters in Part II of the book are intended to stimulate further research and collaboration in areas that—in some cases at least—have been little studied previously. The chapters are ordered to reflect both historical primacy (philosophy and musicology are arguably the oldest disciplines to consider emotion and music) and also a progression from the more individualist perspectives (as epitomized by psychology and biology) to social and cultural perspectives (as reflected in anthropology and sociology).

One important problem that has plagued theory and research on music and emotion is conceptual confusion. Therefore, an important contribution to the study of music and emotion from philosophy could lie in its ability to resolve conceptual problems. In Chapter 2, Davies provides a discussion of philosophical perspectives on music and emotion. After outlining the nature of academic philosophy and discussing philosophical perspectives on emotions, Davies considers three philosophical problems with respect to music and emotion: The first problem concerns the fact that music is often perceived as

expressive of emotion—despite the fact that music is not a sentient being. The second problem concerns the listener's emotional response to music in such instances where the listener's response mirrors the music's, despite a lack of beliefs that usually underlie such a response. The third problem concerns why listeners enjoy and revisit pieces of music that, on their own account, incline them to feel sad. In addressing these problems, Davies scrutinizes the major philosophical theories of music's expressiveness—music as symbol, expression theory, arousal theory, music as virtual persona, contour theory—involving the works of Langer, Kivy, Levinson, and Davies himself. Davies further explores the idea that music is a universal language of emotions, and mentions some critical issues that need to be considered in regard to this notion. (This touches upon one of the recurring themes of the book, i.e. the relationship between biological universals and cultural specifics.) Davies' philosophical discussion is framed by a set of constraints that an acceptable theory of music's expressiveness should satisfy, according to Davies. One of these constraints is that the theory must explain how—despite their manifest differences—our non-primary use of emotion words to describe responses to music relates to these words' normal application to the emotions of sentient creatures.

Philosophical concerns about emotion in music have been expressed for almost as long as people have thought about music. Musicological thinking, surveyed by **Cook and Dibben in Chapter 3**, also has historical roots stretching back for many centuries (indeed, it could be argued that musicological and philosophical approaches have only been separable in the last three centuries or so). Cook and Dibben's treatment of musicological approaches to music and emotion in Western art music and popular music focuses on theoretical frameworks aimed at understanding how music can embody, express, arouse, or otherwise signify emotional states. Cook and Dibben begin with a historical sketch, tracing some of the developments in thinking about music and emotion that still exert an influence on modern thinking, for instance the ancient notion that music can both represent and induce emotions. An important observation is that our views on musical emotions, throughout history, have been much shaped by contemporary conceptions of emotion in general (e.g. the eighteenth-century affect as a rationalized emotional state, in contrast to the nineteenth-century emotion as personal and spontaneous expression). Cook and Dibben then review a selection of classic music-theoretical approaches to musical emotion (many of which were influenced by linguistic models), including the 'doctrine of affections' and the theories by Cooke, Tagg, Meyer, and others. This leads them to propose that the fundamental problem in musicological approaches to music and emotion is 'to talk about both the emotion and music in a sustained manner'. Cook and Dibben suggest that Hatten's neo-semiotic approach is the most comprehensive attempt to integrate expressive and structural analysis within a coherent critical practice. Cook and Dibben conclude by discussing a 'performative' approach in which music-analytical practices are understood less in terms of how the music *is* and more in terms of how it *might* be heard; the implication is that musicological knowledge actually may shape our perceptions of music. Thus, as Cook and Dibben note, the very act of saying that a certain piece of music is expressive in a certain way well lead us to hear it in that way.

Contemporary psychological approaches to music and emotion have been very much influenced by philosophical and musicological debates. In **Chapter 4, Sloboda and**

Juslin describe the essentials of a psychological approach to the study of music and emotion. They suggest that a central aim of psychology is to understand the mechanisms that intervene between music reaching a person's ears and an emotion being perceived, or experienced, by that person as a result of hearing that music. Other important aims involve understanding the roles of emotion in composing and performing music. First, Sloboda and Juslin review the major approaches to conceptualizing emotion in psychology: categorical, dimensional, and prototype approaches. They also show that there are certain underlying and objectively observable characteristics that are present in the set of phenomena we normally describe as 'emotions', and that these same characteristics seem to apply also in the case of music. Then, they consider two primary sources of emotion in music, namely *intrinsic* and *extrinsic* sources. Intrinsic sources are those that are non-arbitrarily embedded in structural characteristics of the music and that contribute to the creation, maintenance, confirmation, or disruption of schematic expectations. Extrinsic sources are of two kinds: (i) iconic sources which come about through some formal resemblance between a musical structure and some event or 'agent' carrying emotional tone; and (ii) associative sources which are premised on arbitrary and contingent relationships between the music being experienced and a range of non-musical factors, which also carry emotional messages of their own. Sloboda and Juslin suggest that failure to see the many ways that music may represent and induce emotions has contributed to the controversies that have characterized the field of music and emotion. No single theory is likely to be able to account for all emotional responses to music. Sloboda and Juslin conclude by suggesting some directions for future research, noting in particular that the psychological approach has tended to ignore the social context of musical behaviour.

Chapter 5 by **Peretz** outlines a biological approach to music and emotion. This approach is very much rooted in psychology, and uses data from the brain to shed light on the cognitive organization of musical experience. It should be noted that a biological approach to music and emotion may entail evolutionary perspectives, physiological perspectives, or neuropsychological perspectives. It is the last of these that is the focus of Peretz' chapter (an evolutionary perspective is provided by Juslin in Chapter 14, and physiological evidence is reviewed by Scherer and Zentner in Chapter 16). Neuropsychological studies are burgeoning at the moment, with a variety of brain-imaging techniques that can be used to investigate the effects of music on the brain, such as positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and event-related potentials (ERPs). Peretz reviews studies that may shed light on the brain substrates underlying emotional responses to music. Peretz also suggests promising future directions for research, showing what questions, in particular, such studies of music and emotion are most likely to answer. Peretz concludes that emotional appreciation of music is best conceived as the product of a specialized cortical arrangement that is shared by most members of a given musical culture.

Psychology and neuropsychology focus on emotion as experienced and perceived by the individual. The remaining chapters of this section consider musical emotion from a broader social perspective, beginning in Chapter 6 where **Becker** presents perspectives on music and emotion from music anthropology: the study of modes of music making as a means of cultural expression. Becker challenges some of the unexamined

'ideologies' and presuppositions that have dominated Western theorizing about music and emotion and that—as she demonstrates—do not apply for most of the world. This includes the common view of the listener as inwardly focused and isolated. Becker introduces the concept of 'habitus' of listening to describe the listening habits and expectations that listeners within a particular culture 'accumulate largely unawares'. These habits reveal implicit assumptions in the listening context about what music 'means', what it is for, how it is to be perceived, and what might be 'appropriate' responses. Becker goes on to consider anthropological views on emotion, and reviews evidence of cultural differences in the expression of, motivation for, and interpretation of emotion in relation to musical events. However, she also notes that, through experience and empathy, we can come to understand differing emotional reactions to different kinds of music. Becker then presents examples of different kinds of habitus of listening from her fieldwork and that of others. Becker suggests that autonomic arousal is a human universal related to music making. This leads her to the main theme of her chapter: a plea for a move to transcend the dichotomy between 'scientific universalism' and 'humanistic particularity'. In particular, Becker suggests that biological and cultural approaches to music and emotion may be brought together by understanding music listening as a process in which the relationship between music and emotion extends beyond the minds and bodies of individuals. She argues that this process can be viewed as 'biological', since both performers and listeners undergo a learning process in which they imitate physical and mental gestures that transform both their inner structures and their relationships to everything beyond the boundaries of the body.

The social context of musical behaviour is also central to **Chapter 7** in which DeNora presents a sociological perspective on music and emotion. Sociology is concerned with investigating forms of social organization and the relations of these forms to social action, whether 'macro' (large-scale institutions) or 'micro' (interactional processes). DeNora first considers the recent rediscovery of emotion in sociology, which has been linked to sociology's traditional concern with social contexts. (A concern with the social context of music making is another recurring theme in the present book.) Then she describes how the focus of sociological studies of music has changed over the last decade, from preoccupations with music's production—music as an object shaped by social relations—to concerns with how it is consumed and what it 'does' in social life. Within this 'new' sociology of music, she observes, music is conceptualized as a device for the constitution of emotive action in and across a range of social settings. (DeNora traces the origins of this development to the writings of Theodor Adorno and ethnographic studies in the 1970s.) One of the most promising trends in the 'new' sociology, according to DeNora, is the discovery of how music can be used by, and on behalf of, actors as a resource to construct self-identity and to create and maintain a variety of emotions. DeNora illustrates this point by reviewing some recent work by Hennion and herself concerning the uses of music in everyday life. Based on in-depth interviews, these studies reveal that listeners are not merely passively affected by music, but rather are actively constructing their own ability to be moved. Moreover, the studies reveal how people use music to regulate, enhance, and change qualities and levels of emotion. They show considerable awareness about the music they 'need' to hear in different

situations and at different times to influence aspects of themselves and their self-concepts. Such uses of music are based on the listeners' perceptions of what they believe the music to 'afford', for instance associations that listeners have made between certain musical materials and extramusical features. DeNora suggests that there may be a useful bridge here between sociology's focus on associations that listeners forge between music and emotions, and musicology's traditional focus on ways in which musical structures lend themselves to particular interpretations.

Ending this section is a contribution which shows how music embedded in a context of social interaction can become an agent of therapy or healing. **Chapter 8** by **Bunt and Pavlicevic** considers the role of music in music therapy, with a particular focus on improvisational music therapy. First, they give a general background to music therapy, discussing different forms of music therapy and showing how the different sources of emotion described by Sloboda and Juslin in Chapter 4 may be observed in clinical practice. A critical aspect to the generating and receiving of emotion in music therapy is the direct experiencing of emotion within the listener or player, as well as the recognition of emotion within the music itself. Therapists' abilities to make emotion judgments from musical improvisations were explored in a study summarized by Bunt and Pavlicevic. The therapists listed a large number of musical parameters that aided the process of judging the emotions (Bunt and Pavlicevic note that these are common features of many musical improvisations made by patients of all ages). Moreover, Bunt and Pavlicevic show how recent work in developmental child psychology by Trevarthen, Trehub, Papoušek, and others, and in particular Stern's so-called 'vitality affects', can provide reference points in understanding the subtle interplay between the personal and the musical in improvisational music therapy. They suggest that the primary locus of any therapeutic change lies in the interpersonal musical relationship built up between therapist and patient. Examples from their own practice show how Pavlicevic's concept of dynamic form can be used to understand how the relationship between the therapist and the patient is defined in real-time improvisations. This chapter complements the scientific perspectives offered by the previous chapters by providing unique insights from clinical practice.

We believe that this multidisciplinary section makes at least two things clear. First, it suggests that multidisciplinary approaches are called for—perhaps even necessary—to make real progress in understanding the relationship between music and emotion in all its richness and complexity. Second, it calls for some humility. Given the broad range of domains that intersect with the study of music and emotion—from brain imaging to sociological studies to analyses of musical structure—we cannot expect any one individual to be able to address the topic from all angles simultaneously. The key word is 'collaboration'. The first section of the book presents a challenge to the reader, with its wide range of disciplines, theories, concepts, and methods. Many of the traditional views on music and emotion are certainly challenged. However, we believe that the reader who is willing to accept this challenge will be rewarded with a fresh look on the topic, and that he or she may be able to look at the other contributions of the book in the light of these multidisciplinary statements.

### 1.2.2 The composer

In Part III, the focus is on the composer and—in particular—the materials or 'building blocks' that the composer can work with. In subsequent sections, the contributions examine emotion as communicated and experienced by the performer, and finally by the end-recipient, the listener.

One of the fundamental emotion dimensions which studies of music listening have identified is *arousal*. The notion of arousal is crucial to **Simonton's Chapter 9**, which links this concept to the act of music composition. In other chapters of the book, the authors have reviewed evidence based on such methods as philosophical enquiry, experiment, observation, brain imaging, interview, and structural analysis. In this chapter, Simonton adopts yet another approach called 'historiometrics', which involves the adaptation of psychometric methods to analyse data about historic individuals. In the case of classical composers, Simonton includes both biographical information about their lives and content-analytical information about their works. Simonton has been using this particular technique to study creativity and aesthetics in classical music. Although this research programme was not directed specifically at investigating musical emotion, some of the empirical results clearly have a bearing on this topic. Simonton begins by defining his concept and measure of 'melodic originality' in musical compositions, and goes on to show that this simple measure—which incidentally seems to encompass more compositional aspects than just melody alone—may be related to a piece's 'arousal potential' as perceived by listeners. Pieces of music scoring high on 'melodic originality' should sound more unpredictable, interesting, or complex (and thus yield more uncertainty, to elaborate on Meyer's views in Chapter 15) than melodies scoring lower. Simonton refers to experimental work that supports this hypothesis, and there is a clear link to Meyer's theory that emotional arousal reflects the violation of musical expectations. Simonton goes on to demonstrate that the relation between arousal potential and aesthetic impact is curvilinear, in accordance with the 'Wundt curve', and that the most popular classical works in the real world are those that evoke an optimum arousal level. However, due to 'habituation', the response to an arousing stimulus will often diminish. Therefore, as Simonton illustrates, composers have a tendency to increase their use of melodic originality as their careers progress. Moreover, Simonton argues (boldly) that the degree of melodic originality also reflects the emotional life of the composer at the time of a work's composition, such that the melodic originality (and hence the 'arousal potential') is larger during periods of life stress. Indeed, Simonton finds a positive relationship between measures of melodic originality and an index of life stress for ten eminent composers of classical repertoire (after controlling for confounding factors). Simonton also discusses the 'Swan song' phenomenon (i.e. the idea that composers exhibit a 'late life' compositional style). Simonton concludes that 'melodic originality', at least to some extent, forms part of the composer's tools for expressing emotions in music. However, he also observes that composers have a great deal more in their 'expressive tool kit' than just the choice of pitches.

Additional aspects of the expressive tool kit that composers can use to express various emotions are dealt with in **Chapter 10**, in which **Gabrielsson and Lindström** present a review of studies—from the end of the nineteenth century to the present day—that have attempted to specify the expressive properties of various elements of

musical structure usually indicated in the musical notation of a piece of music (performance aspects are reviewed in Chapter 14 by Juslin). This is, in fact, the domain within which most of the traditional studies of music and emotion tend to fall. Gabrielsson and Lindström first describe different methodological approaches used in studies of expression in music since the late nineteenth century, including different response formats: free description, adjective checklists, adjective ratings, and continuous response (reviewed by Schubert in Chapter 17). Studies also differ with regard to musical stimuli, ranging from brief and simple sound sequences to recordings of entire musical compositions. Then, Gabrielsson and Lindström review findings about the expressive properties of different musical elements, such as interval, mode, rhythm, tempo, melodic direction, harmony, timbre, and form. This review uncovers a number of relationships between musical elements and particular emotions, but also indicates several gaps and ambiguities in the results. In particular, few investigations have considered potential interactions among different elements. Gabrielsson and Lindström conclude by suggesting directions for future research, emphasizing the need to explore more sophisticated manipulations of melody, rhythm, timbre, and harmonic progression. Even so, the findings reviewed by Gabrielsson and Lindström provide a rich source of music-emotion relationships that have been used, or could be used, in a variety of practical contexts such as music therapy or film music.

In Chapter 11, Cohen considers an applied compositional domain that has a substantial history: the use of music in film. This complements Chapters 8 and 13 which explore applications of knowledge about emotion in the context of music therapy and performance anxiety. The purpose of Cohen's chapter is to address the question of what contribution music makes to the emotional impact of film. Cohen argues that music is one of the strongest sources of emotion in film and intends to open doors to further empirical work that explains why this is so. Cohen begins by establishing a context for discussing emotion in music and film. Then, she presents a historical perspective on film music, from the use of music in silent film to the advent of the sound film in the late 1920s. Cohen also reviews empirical research that has investigated the effects and functions of music in film. These functions of film music include masking extraneous noises, providing continuity between shots, directing attention to important features, inducing mood, communicating meaning and furthering the narrative, enabling the symbolization of past and future events through the technique of leitmotiv, heightening the sense of reality, and adding, as an art form, to the aesthetic effect of the film. Cohen also presents a cognitive framework for understanding musical phenomena of relevance to soundtracks. Finally, she considers the role of the composer as the ultimate source of musical emotion in film, discussing conditions and requirements associated with composing music for film. A hypothesis of Cohen's chapter is that the capacity of music to accomplish its emotional task may be based on the ability of music to simultaneously carry several kinds of emotional information in its harmony, rhythm, melody, timbre, and tonality. Real life often entails multiple emotions, simultaneously and in succession. Miraculously, yet systematically, Cohen argues, this 'emotional polyphony' can be represented by the musical medium.

### 1.2.3 The performer

In Part IV of the book, the focus is on the role of emotion in the performance of music. In **Chapter 12, Persson** makes an excursion into the subjective world of the performer, outlining exploratory research into the 'phenomenology' of emotion in musical performance. This includes such issues as performance motivation and interpretation, and the learning and conceptualization process of constructing musical meaning. Persson acknowledges that there is little previous research in this area, but notes that the present chapter might serve as a basis for future research and provide a preliminary understanding of how performers construe the subjective aspects of musical performance. Persson's overview is based on qualitative and quantitative research by himself and others, as well as on published interviews with famous performers. Persson considers the motivational basis for becoming a performer, asking what made certain individuals choose a musical career. Although a variety of reasons are suggested, a significant finding from Persson's research is that many performers are motivated primarily by a 'hedonic' motive—treasuring music as a means to control and induce positive emotional experiences by means of playing (this theme is also evident in Chapter 6 by Becker). Persson then considers the roles of emotion in interpretation, discussing imagery, conceptualization of musical meaning, and so-called self-induction to altered states of awareness. He concludes by suggesting directions for (much needed) further research concerning the subjective world of performers.

As noted by Persson in Chapter 12, one important reason for performing music is that it induces positive emotions in the performer. In **Chapter 13, Steptoe** considers the flip side of the coin: While performance of music often generates strong positive emotions, it may also induce *negative* emotions in the performer, such as distress and anxiety. Such emotions may impair the quality of performance, and could also have debilitating effects on the performer's career. Steptoe reviews the psychological and medical literature on performance anxiety and also considers other negative emotions related to music performance as a profession, such as social tension, worries about job security, boredom, feeling separated from family, and so forth. Furthermore, Steptoe argues that research on performance anxiety would benefit from regarding performance anxiety as a form of situational stress response, which is similar in certain respects to problems such as social phobia or academic examination stress. By analogy with other work on stress, Steptoe suggests, it may be useful to separate distinct components of performance anxiety. One such component is affect, which forms the central experience of performance anxiety for many musicians. However, there are other important aspects as well, including cognitive responses (e.g. disturbances in information processing), changes in behaviour, and physiological responses. Steptoe reviews some promising treatment methods that have been devised for stress management, and argues that the incorporation of these principles into the training of musicians could be very beneficial.

Numerous studies have investigated emotional expression in pieces of music (these studies are reviewed by Gabrielsson and Lindström in Chapter 10). However, only recently have researchers begun to study how different aspects of the *performance* affect the emotional expression of a piece of music. This topic is reviewed by **Juslin** in **Chapter**

14. He provides a review of empirical studies of communication of emotion in music performance, and outlines a theoretical framework suitable for organizing the findings. The review of studies focuses on two aspects of the communicative process, namely how well performers succeed in conveying a particular emotional interpretation to a listener, and what acoustic means they use to do this. The relationships between variables in the performance and emotion categories (e.g. sadness) and dimensions (e.g. arousal, activity) are described, with emphasis on 'expressive contours' and so-called 'mixed' emotions. Then, Juslin outlines a theory to account for the origin of the 'code' used by performers. Following the biological approach to music and emotion discussed by Peretz in Chapter 5, Juslin suggests an evolutionary perspective on performers' expressive skills. He argues that the acoustic code of performers reflects both innate 'brain programmes' for vocal expression of emotions, and social learning (for example, learning of links between cues and extramusical aspects through analogies). He further argues that the code used by performers today may be linked—in functional ways—to important cultural activities of the past. Juslin also describes a Brunswikian 'lens model' than can serve to illustrate how music performers communicate emotions to listeners by combining a number of uncertain but partly redundant cues in the performance. The lens model makes it possible to quantitively relate performance to perception in order to better understand the functional relationships between expression and recognition of emotion. Finally, possible mechanisms for the induction of emotion by expressive performance are discussed by Juslin.

### 1.2.4 The listener

In the final part of the book, the focus is on the listener. The contributions are ordered to reflect the progression of Part II (multidisciplinary perspectives). This progression begins with a musicological (and philosophical) contribution, moving to two chapters concerned with the explanation and measurement of individual response, and ends with two contributions which emphasize the effect of social context on listener response.

One of the most influential analyses of musical emotion was given by Meyer (1956) in his book, *Emotion and meaning in music.* In Chapter 15, Meyer revisits his theory to provide an updated version of his view on music and emotion. The focus of Meyer's essay is on the role of uncertainty in listeners' emotional responses to music. First, Meyer provides a number of distinctions for the field of music and emotion, observing that the field has lacked conceptual specificity and precision (another theme recurring throughout this book, see, for example, Chapters 4 and 16). Then, he points out that researchers have tended to focus on emotions as states rather than processes. Consequently, Meyer notes, researchers have underestimated the importance of succession in explaining emotional responses to music. In fact, the heart of the emotional experiences of music lies, he argues, in processes that can be understood in terms of creation and minimization of uncertainty (the role of uncertainty is also discussed by Juslin in Chapter 14, but whereas Juslin focuses on the 'textural' resolution of uncertainty via parallel 'cues' to emotion, Meyer focuses on the 'successive' resolution of uncertainty in musical 'episodes'). Meyer distinguishes between 'native' processes (which

reflect constraints that are inherent in the human mind and body, e.g. the gestalt laws) from 'syntactic' processes (which depend on learning, having internalized the constraints of a particular style). Both 'native' and 'syntactic' processes generate uncertainty (about what will happen, when it will happen, and how it will happen). Uncertainty, in turn, generates arousal and schematic implications that are disrupted, delayed, or confirmed. Ultimately, Meyer observes, this process reflects the human need to envisage in order to choose. But whereas we try hard to avoid uncertainty in everyday life, in the arts—including music—we frequently cultivate and relish a considerable amount of uncertainty. But uncertainty itself is not an emotion, according to Meyer. Uncertainty is rather one of the sustaining conditions for emotions. The uncertainty is a function of the amount of information the mind is required to process. Therefore, as the number and strength of implied musical alternatives increase, so does the uncertainty.

Scherer and Zentner in **Chapter 16** present a theoretical framework for studying emotional effects of music in which they include very stringent criteria for distinguishing different kinds of affective response: mood, emotion, preference, attitude, personal trait, and interpersonal stance. Scherer and Zentner propose that we should theoretically define the effects of music on the listener in the form of 'rules', which specify the relative contribution of different features (such as musical structure, performance, listener personality and mood, and context). The bulk of their chapter deals with production rules, outlining those mechanisms that can be theoretically expected to produce, or induce, changed emotional states in the listener (e.g. appraisal, memory, empathy, or proprioceptive feedback). Scherer and Zentner suggest that it is important to clarify exactly which type of state can be produced by music and what criteria should be used to determine the presence of the respective states empirically. Then, they review the literature pertinent to this issue, and examine the effects of music on cognition, physiological arousal, motor expressive behaviour, action tendencies, and subjective experience. Scherer and Zentner argue that, ideally, the claim that music can induce emotions should be instantiated by studies showing that listening to a piece of music reliably leads to a *synchronization* of cognitive, physiological, expressive, motivational, and experiential processes. Few studies of this kind are currently available, although Scherer and Zentner review a large number of studies of separate components, which provide preliminary evidence of emotional responses. One theme in the chapter is that researchers should be open to new ways of conceptualizing affective responses to music. Scherer and Zentner also suggest that an important goal for future research might be to develop a vocabulary of emotion specifically for the domain of music. They conclude by making suggestions for theoretically based and methodologically sound future research on emotion production.

One of the themes of Meyer (Chapter 15) is that researchers have tended to underestimate the importance of succession in explaining emotional responses to music. Following this lead from Meyer, **Schubert** (in **Chapter 17**) presents an overview of recent attempts to address this problem by means of 'continuous response methodology'. The comprehensive research agenda proposed by Scherer and Zentner (Chapter 16) also requires the development of such methodologies. First, Schubert considers different response formats used in continuous response research—checklists, open-ended formats, and rating scales—noting some of the problems associated with particular

formats. He argues that rating scales based on emotion dimensions like *valence, activity,* and *arousal* are most useful in continuous measurement of emotional response. Schubert goes on to discuss the different uses of continuous response methods in the literature: validation of various response formats, comparative investigations (e.g. comparing the responses of musically trained and untrained listeners), stimulus-response investigations (i.e. exploring which musical parameters predict changes in response), and system-dynamics investigations (i.e. studying the dynamic nature of emotional responses over time). Schubert also provides a critical discussion of problems concerning the design, validity, and reliability of continuous response instruments (e.g. the coding of musical features, response coding, synchronization of music and response, sampling rate, lag structure, and statistical analysis). Schubert argues that researchers using continuous response methods must become more aware of the analytical problems created by continuous response data, and suggests some solutions for dealing with these problems. He concludes that the main advantages of continuous response methodology—the abilities to study moment-to-moment changes in the relationship between musical structures and emotional responses and the 'lag' structure of listeners' response to music—are yet to be fully exploited by researchers.

Scherer and Zentner argue in Chapter 16 that future work on the emotional effects of music must pay greater attention to motivational constructs that are specific to music (e.g. 'why do people listen to music?'), because these are the motivational urges that need to be taken into account when we attempt to assess goal conduciveness or obstruction in cognitive appraisals of musically induced emotions. A fresh perspective on this motivational terrain, the uses and functions of music in everyday life, is provided in **Chapter 18 by Sloboda and O'Neill.** They build on the foundation laid out by DeNora in Chapter 7, pointing out that music is a cultural resource in the social construction of emotions. Music, they observe, is heard in a particular place and time, with or without other persons being present, and with other activities taking place at the same time. Thus, an account of emotional responses to music must recognize the interaction among these different factors. The purpose of Sloboda and O'Neill's chapter is to discuss research on habitual and routine modes of engagement with music, including as much of the social context as possible. By definition, everyday uses of music are difficult to study in the laboratory. Further, they tend to be unmemorable and, hence, unsuitable for retrospective report of the kind used by Gabrielsson (in Chapter 19). The solution adopted by Sloboda and O'Neill is called the 'experience sampling method'. This involves the use of electronic pagers that subjects carry with them during the waking hours, along with a booklet that they fill in on each paging. A study by Sloboda and coworkers using this method indicated that there was a 44 per cent likelihood of music being experienced in any 2-hour period—confirming the common claim that music pervades everyday life. More importantly, the study showed how music was used in a wide range of everyday activities. One of the working assumptions made by Sloboda and O'Neill is that music can have different 'emotional functions' in different situations, and they present evidence that listeners use music to manage their moods in a number of different ways. Indeed, this deliberate use of music to achieve specific psychological outcomes can be described in terms of 'self-therapy'—a practice partly corresponding to the professional uses of music to that same end in music therapy (see Bunt and

Pavlicevic, Chapter 8). One notable finding from Sloboda and O'Neill's pager study is that the majority of episodes that involved music resulted in the listeners feeling more positive and alert. That is, music made them feel better. However, Sloboda and O'Neill also review evidence of negative emotions in memories of childhood experiences, often related to an evaluative context and hence reminiscent of the negative emotions experienced by some adult performers (Chapter 13). Sloboda and O'Neill conclude that musical activities are enmeshed in a social and cultural world which involves a number of 'hidden' practices and assumptions about musical activity. Only by uncovering the 'hidden' meanings of these practices, Sloboda and O'Neill argue, can we hope to gain a better understanding of emotional responses to music in everyday life.

In **Chapter 19, Gabrielsson** reviews findings about emotional reactions to music based on free phenomenological description of strong experiences of music. Gabrielsson begins by discussing the notion of *peak experience*, made famous by the humanistic psychologist Abraham Maslow. Gabrielssson then describes his research project on strong experiences of music (SEM), the purpose of which was to get some sense of what phenomena are contained in such experiences, and to study how SEM are related to the properties of the music, the individual, and the situation. Gabrielsson is also concerned with the long-term consequences of such experiences. To address these questions, Gabrielsson adopts an approach involving written self-reports by listeners of their strongest experiences of music. Content analysis of several self-reports yielded an overall classification of SEM into physical and behavioural aspects, perceptual aspects, cognitive aspects, emotional aspects, motivational aspects, existential/transcendental aspects, and developmental aspects. In the present chapter, Gabrielsson explores in depth the emotional aspects, providing a large number of examples of emotional reactions to music from the listeners' reports. The examples include so-called 'basic' emotions as well as 'aesthetic' and 'transcendental' emotions. These descriptions provide fascinating glimpses into the peak experiences that many people treasure. Gabrielsson notes that positive emotions dominate in the reports, often in the form of euphoria that borders on ecstasy, but that there are also examples of strong negative emotions. Finally, Gabrielsson considers some implications of these findings for theories of emotional responses to music, and also mentions certain problems associated with retrospective, verbal description of music experience, such as memory distortions and individual differences in verbal fluency.

## 1.3 Concluding remarks

The peer-reviewed contributions of the present book will reveal how much progress there has been in this field since Meyer's (1956) classic book, *Emotion and meaning in music*. Besides providing research overviews, many of the chapters also set forth significant paths for future theory and research. *Music and emotion* may thus serve many functions. First, it may provide a research source for what is currently known about music and emotion. Second, it may serve as a starting point for investigating the field. Third, and as already noted, it can be a source of interdisciplinary dialogue. The chapters in this volume represent the history of the study of music and emotion, its present state, and its future. This scope is reflected in the age range of the contributors. At the

time of writing, the youngest of the contributors is 30 years old and the oldest is 82 years old. Throughout the work on this book we have often been struck by the enthusiasm with which the authors have shared their ideas about music and emotion; we hope that this enthusiasm will come through in their contributions.

Last but not least, we would like to describe one of our personal reasons for bringing together this anthology. We are dedicating this book to Alf Gabrielsson to coincide with his retirement in June 2001. Alf has made an enormous contribution to the field of music research—both as a researcher and as a teacher—as many of his colleagues would be willing to attest. Of more direct relevance for the present volume, Alf was one of the (relatively few) people who repeatedly stressed the importance of studying emotional aspects of music, long before it became fashionable to do so. We sense an increase of interest in affective processes in music at the moment and hope that the twenty-first century will be the time when the domain of music and emotion finally catches up with the other domains of musical science. This would surely help us to gain a better understanding of the manifold ways that music enriches all of our lives.

## References


Cook, N. (1998). *Music: A very short introduction.* Oxford, UK: Oxford University Press.

Deutsch, D. (ed.) (1999). *The psychology of music,* (2nd edn). San Diego, CA: Academic Press.

Dowling, W. J. & Harwood, D. L. (1986). *Music cognition.* New York: Academic Press.

Ekman, P. & Davidson, R. J. (ed.) (1994). *The nature of emotion: Fundamental questions.* New York: Oxford University Press.

Frith, S. (1996). *Performing rites: On the value of popular music.* Oxford, UK: Oxford University Press.

Gabrielsson, A. (1999). The performance of music. In *The psychology of music,* (2nd edn) (ed. D. Deutsch), pp. 501–602. San Diego, CA: Academic Press.

Gardner, H. (1985). *The mind's new science. A history of the cognitive revolution.* New York: Basic Books.

Gaver, W. W. & Mandler, G. (1987). Play it again, Sam: On liking music. *Cognition and Emotion,* 1, 259–82.

Hevner, K. (1935). Expression in music: A discussion of experimental studies and theories. *Psychological Review,* 42, 186–204.

Japanese Society of Music Perception and Cognition (1989). *Proceedings of the First International Conference on Music Perception and Cognition.* Kyoto City, Japan: University of Arts, Department of Music.

Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance,* 26, 1797–813.

Kuhn, T. S. (1962). *The structure of scientific revolutions.* Chicago: University of Chicago Press.

Lewis, M. & Haviland-Jones, J. M. (ed.) (2000). *Handbook of emotions* (2nd edn). New York: The Guildford Press.

Meyer, L. B. (1956). *Emotion and meaning in music.* Chicago: University of Chicago Press.

**Oatley, K.** (1992). *Best laid schemes. The psychology of emotions.* Cambridge, MA: Harvard University Press.

**Palmer, C.** (1997). Music performance. *Annual Review of Psychology*, **48**, 115–38.

**Plutchik, R.** (1994). *The psychology and biology of emotion.* New York: Harper-Collins.

**Scherer, K. R.** (1985). Vocal affect signaling: A comparative approach. In *Advances in the study of behavior*, (Vol 15) (ed.J. Rosenblatt, C. Beer , M-C. Busnel, & P. J. B. Slater), pp. 189–244. NewYork: Academic Press.

**Sloboda, J. A.** (1985). *The musical mind. The cognitive psychology of music.* Oxford, UK: Oxford University Press.

**Sloboda, J. A.** (1995). Psychology of music today: The need for applicable psychology. In *Psychology of music today*, (ed. M. Manturzewska, K. Miklaszzewski, & A. Biatkowski), pp. 19–24. Warsaw, Poland: Frederyk Chopin Academy of Music.

**Small, C.** (1999). Musicking: the meanings of performance and listening. A lecture. *Music Education Research*, 1, 9–21.

**Woods, C., Luck, G., Brochard, R., Seddon, F., & Sloboda, J. A.** (ed.) (2000). *Proceedings of the Sixth International Conference on Music Perception and Cognition, August 2000* (CD rom). Keele, UK: University of Keele, Department of Psychology.

PART II

# MULTIDISCIPLINARY PERSPECTIVES

CHAPTER 2

# PHILOSOPHICAL PERSPECTIVES ON MUSIC'S EXPRESSIVENESS

STEPHEN DAVIES

## 2.1 Philosophy and its method

Sometimes philosophy seems to psychologists to be psychologizing in a fashion that is uninformed and unrestrained by empirical data. (That's alright, sometimes psychology looks to philosophers like unskilled philosophizing!) As this is the only chapter by a philosopher, I begin with an introduction outlining the nature of academic philosophy.

For most questions, one (or more) of the following strategies supplies the answer: fact finding, scientific theorizing, calculating, voting, and legislating. The natural and social sciences rely on the first three of these methods. Some questions are not satisfactorily resolved by their use, however. The issues these questions present are neither empirical nor merely matters of opinion—they are 'philosophical'. 'Philosophy' comes to connote the recalcitrant residue of questions which remain after all the (other) branches of knowledge and taste have taken the questions they are equipped to answer. The fact that philosophical questions are immune from empirical data suggests, perhaps, that they arise from deep conceptual confusions or subtleties, and that it is the purpose of philosophy to untangle these knotted skeins.

Some of the questions of philosophy are distinctive to it, such as many of those addressed by metaphysics, ethics, and logic. But, also, philosophical queries are generated by consideration of foundational issues in all areas and disciplines, and so there is philosophy of art, of science, of mathematics, of feminism, of medicine, of mind, and so on.

Some classic examples of philosophical questions are these: What is truth? Is causality a natural relation or, instead, merely how constant conjunction sometimes is interpreted by human observers? What is identity? Is more than one consideration (such as merit and need) relevant to justice and, if so, are these different aspects mutually commensurable? How is the good to be characterized? What is time? A few moment's reflection should make clear that the answers cannot be settled by an investigation looking at the sorts of facts tested by science. The usual scientific methods of enquiry typically presuppose certain answers to such questions, and the use of such approaches thereby prejudges the outcome of the study. To take another case, how could science analyse the nature of scientific facts without begging the outcome? And another: if empirical facts

underdetermine the explanatory theories we apply to them, we cannot choose between theories, as we will want to do in distinguishing pseudo-science and psycho-babble from legitimate disciplines, solely by citing the facts, because what counts as a fact and how it is significant is disputed between the contested theories. Also, if we are indeed enmeshed in some subtle confusion about a topic, we may be confused in how to describe and respond to (further) empirical evidence.

To say that philosophical method is non-empirical is not to imply that philosophy is indifferent to the facts of science. Philosophical analyses must be consistent with the facts, or with interpretations of what these are. But philosophical analyses must go beyond the facts in resolving the problems, paradoxes, and inconsistencies they can seem to generate. What is needed often is not more facts, but a clarification of the issues raised by those that are available. The most familiar notions can produce conceptual puzzles, and then it is not more facts but a deepening of our current understanding, or new interpretations of the resident data, that is needed. Sometimes philosophical investigations are suggestive of new empirical questions, and here science takes over again. At other times, a grasp of unexpected conceptual connections or distinctions enables us to overcome the mental cramps that formerly afflicted us.

To bring out their non-empirical nature, philosophical puzzles often describe two things that are alike in their appearance or empirical features but which strike us as conceptually very different, and then asks what makes that difference. For instance, Descartes wondered how he knew he was not presently dreaming, given that he could dream that he was dreaming. Others have explored how we are to describe the distinction between one's arm's being moved and the action of one's moving one's arm, or how Marcel Duchamp's ready-mades could be artworks where their lookalike, unappropriated counterparts were not.

It could be that nothing hangs on how we answer the philosophical questions before us. In that case, we will come to see that it is for us to decide which way to go, since there is no truth of the matter to be discovered. And it can be that, on examination, philosophical questions turn out to be nonsensical. But it cannot be assumed from the beginning, however, that all philosophical questions are empty, trivial, or silly. In many cases, how we propose to resolve a philosophical question can have far-reaching implications for other cases, for overall consistency within our folk practices, for the power and fecundity of our explanatory models, and for coherence and consistency in our theories.

Not all philosophical questions are weighty. 'Is the karaoke singer a co-performer with the people on the accompanying DVD?' is a philosophical question that provokes us to consider the individuating conditions of performances and recordings, but it is not in the same league as 'what is the meaning of life?' And not all weighty questions are as exciting as this last. Also, the significance of the topic under discussion is often not apparent in the examples in terms of which it is discussed. These frequently are mundane or commonplace, either because it is appropriate to begin with uncontroversial paradigms or to avoid illustrations that attract attention to themselves, rather than to what is at issue. Philosophers who discuss the expression of emotion in music regularly offer the slow movement of Beethoven's *Eroica* as illustrating musical sadness. This does not mean they deny that other pieces are expressive, or that other emotions can be

expressed in music. Instead, it indicates that even such a humdrum case already presents in a graphic form the puzzle that troubles the philosopher.

Sometimes, Anglo-American analytical philosophy, which is the approach of this chapter, is contrasted with continental philosophy. The differences are largely matters of degree or style. In terms of the philosophy of music, Continental approaches sometimes describe an entire metaphysical system and consider music in order to locate it within that system; or they view music in terms of larger sociopolitical or psychological models of human behaviour, such as Marxist or psychoanalytic ones; or, again, they approach it from the perspective of general theories that are semiotic, structuralist, or deconstructionist. And when they do focus on music as such, they are likely to emphasize the subjective response and the phenomenology of musical experience. By contrast, analytic philosophy is less theory-driven and politically motivated. It is centred on perceived 'problems' or 'questions', rather than on a canon of works by great philosophers. The approach is (often) piecemeal, tackling issues one by one as their relevance emerges. The focus falls on interpersonal judgments and public criteria, rather than the idiosyncratic or personal.

In what follows, I concentrate on philosophical discussions of music's expressive character. An obvious question asks how music could be expressive of emotion, which is how we seem to experience many pieces, when it is non-sentient. Depending on how we answer this first query, we are led to others. For instance, if we argue that expressive predicates apply only metaphorically to music, we might then be puzzled at the strength of the response music is capable of eliciting from the listener.

The topic on which I focus, that of music's expressiveness, has been a dominant one in music aesthetics, but many subjects not covered here have also attracted detailed examination by philosophers. These include the ontology of musical works (Goehr 1992; Gracyk 1996; Levinson 1990), performances and their relation to works (Godlovitch 1998; Thom 1993), the notion of 'authentic' performance (Davies 1987; Kivy 1995; Levinson 1990), musical representation (Davies 1994; Kivy 1988), ways in which music can be said to be meaningful and the kind of understanding it invites (Davies 1994; Kivy 1990; Levinson 1990, 1997*a*), music's value (Budd 1995; Davies 1994; Goldman 1995*a*; Higgins 1991; Kivy 1990, 1997), and music's social significance and ethical dimension (Goehr 1998; Higgins 1991). For recent collections covering many of these topics, see Alperson (1994, 1998), Robinson (1997), and Hjort and Laver (1997).

## 2.2 Ways in which music's expression of emotions is philosophically problematic

In this chapter I consider three puzzles, the first of which has dominated philosophical discussions about music and emotion over the past two decades. It observes that purely instrumental music is not the kind of thing that can express emotions.[1] Music is not

[1] The focus falls on instrumental music simply because the problem of music's expressive powers is at its most acute where music is divorced from words, narrative, and drama.

sentient and neither is its relation to occurrent emotions such that it could express them. The second problem concerns the listener's response, where this mirrors the music's expressive character. When listeners are saddened by the music's sadness, apparently they lack the beliefs that normally underpin such a reaction; for example, they do not think the music suffers, or that the music's expressiveness is unfortunate and regrettable. The third perplexity concerns negative responses elicited by music, such as the sad one just mentioned. Why do listeners enjoy and revisit works that, on their own account, incline them to feel sad? (For a summary of philosophical approaches to these three topics and more besides, see Levinson 1997*b*, 1998.) Before addressing these puzzles, it is useful to examine philosophical theories regarding the nature of the emotions.

## 2.3 Theories of the emotions

It was once thought, by Descartes for instance, that emotions involve the subject's awareness of the perturbations of his or her animal spirits. It was the dynamic structure of this inner motion, along with the feeling of pleasure or displeasure with which it was apprehended, that distinguished the various emotions. Call this the 'hydraulic theory' of the emotions. In this view, emotions are experiences passively undergone by the subject; they are only contingently connected to their causes and to their behavioural manifestations; they are essentially non-cognitive.

In the latter half of the twentieth century, an alternative account, usually called the 'cognitive theory', has been developed (see Lyons 1980; Gordon 1987; Solomon 1976).[2] This allows that emotions possess a phenomenological profile, but regards this as only one element among several, all of which are necessary and none of which is sufficient alone for an emotion's occurrence. Emotions may be characterized by physiological changes, but, more importantly, they are object-focused. Emotions are directed toward their objects. This means they are usually outward facing, as when I fear the lion that is before me, though the emotion's object also may be one's own sensations or emotions, as when I am alarmed by how tense I feel or where I am ashamed that I am angry. Moreover, emotions involve the categorization of their objects; for instance, if the emotion is one of fear its object must be viewed as harmful, and if the emotion is one of envy its object must be viewed as something both desirable and not already controlled or possessed. In addition, they include attitudes toward their objects; for example, though I judge you to be injured, my emotional response will depend on whether this is a source of concern, satisfaction, or indifference to me. Also, particular emotions find expression in typical behaviours; if I pity you I will try to comfort you and to change your situation for the better, and if I fear you I will fight, flee, or seek protection.

Philosophers disagree in the versions of the cognitive theory they espouse. Some insist that emotions require a belief in their objects' existence, while others think that cognitive

[2] The term 'cognitive' has a somewhat different meaning in philosophical theories of emotion than it has in psychological theories. In the latter, the term implies a focus on underlying information-processing mechanisms, whereas in philosophical theories, it refers to beliefs, imaginings, thoughts, intentions, desires, and like states of consciousness.

attitudes, such as make belief, also can play the role of securing the emotion's object. Some hold that emotions can be individuated in terms of their sensational patterns, without reference to the emotion's propositional component, whereas others regard its cognitive ingredient as crucial to its identification. (For a useful summary of the literature, see Deigh 1994.)

Many of these disputes can be resolved by acknowledging that emotions do not constitute a homogenous class. For instance, some, such as disgust, might be primitive, automatic responses that are not susceptible in their operation to changes in the subject's cognitive state, whereas others, such as patriotism, are marked more by their self-conscious, intellectual content than their sensational character. Even if there is a continuum of cases between these extremes, it is useful to distinguish between emotions (such as jealousy, hope, and remorse) in which the cognitive elements are prominent, malleable, and sophisticated from those (such as lust, fear, and disgust) in which the cognitive elements may not be present to awareness and the reaction is inclined to be automatic and inflexible.[3]

Some further distinctions that may be useful are those between emotions and moods, and between emotions and mere sensations (or mere feelings). These distinctions are drawn roughly within folk psychological discourse along the following lines. Moods are not object-directed and involve rather general feelings. There can be moods of dread, depression, and happiness, but not of embarrassment or remorse, because these latter lack a distinctive experiential character and are distinguished more in terms of what is cognized about their objects. Meanwhile, emotions may involve bodily sensations but are not reducible to them. A person who sits too close to the fire might experience exactly the same sensations as another who is acutely embarrassed, but it is only the latter who feels an emotion.

## 2.4 Constraints on the theory

Before outlining philosophers' theories of musical expressiveness, it is helpful to consider the desiderata that an acceptable theory must satisfy.

We could not account for the interest and value of expression in music, or for the emotional responses music calls from the listener, unless terms like 'sad' and 'happy' retain their usual meanings in connection with music's expressiveness. So a principal

[3] Griffiths (1997) argues that the emotions do not form a 'natural kind' and that the cognitive theory applies to only some. He is impressed by experimental data supposedly showing that the neural structures dealing with the emotions are modular and cognitively opaque, thereby generating evolutionarily adaptive quick, dirty, and conservative reactions to affective stimuli. (For work on this wavelength, see LeDoux 1998.) But it seems to me that the cognitive theory is not so easily dismissed. Propositional attitudes—beliefs, desires, intentions—can be in play without being held before the mind. Yesterday I believed that the earth is round, though I did not bring the relevant thought to consciousness at any time. Nevertheless, though I suspect the experimental data referred to by Griffiths might be interpreted in ways consistent with cognitivism, as I have just indicated, later I argue that some of the most common emotional responses to music's expressiveness do not fit the cognitive model.

task will be to indicate how, despite their manifest differences, music's expressing an emotion parallels the default case in which a person expresses an emotion they feel. In other words, an account explaining and justifying our attribution to music of predicates such as 'sad' and 'happy' must make clear how this non-primary use relates to these words' normal application to the occurrent emotions of sentient creatures.

As I see it, this constraint quickly rules out three approaches to the topic. It will not do to attempt to reduce music's expressiveness to a catalogue of technicalities and compositional devices. Even if it is true that all, and only, music in minor keys sounds sad, it cannot be that 'sounds sad' *means* 'is in a minor key'. Even if one can make sad music by composing it in the minor key, there must be more to the analysis of music's expressiveness than acknowledgment of this, for it is by no means clear how the music's modality relates to the very different kinds of things that make it true in the standard case that a person is expressing sadness. Musical features ground music's expressiveness, and it is interesting to discover what features those are, but identifying them is, at best, only an initial step toward an informative theory of musical expressiveness.

Another of the disallowed strategies claims that music's expressiveness is metaphorical and declines to unpack the metaphor.[4] The claim here is not merely that, as a figure of speech, music can be described metaphorically. It is, rather, that the music itself is metaphorically expressive. While this last assertion obviously locates expressiveness squarely with the music, its meaning is quite mysterious. The idea that musical expression is metaphorical must itself be a metaphor, since metaphor primarily is a linguistic device depending on semantic relations for which there is no musical equivalent. This approach indicates what is puzzling about music's expression of emotions—that it is hard to see how emotion terms could retain their literal sense when predicated of music, though clearly their application to music trades somehow on their literal meaning—but it offers no solution to that puzzle.

Also unacceptable is the theory insisting that music's expressiveness is *sui generis*; that is, of its own kind and not relevantly comparable to the default case in which occurrent emotions are expressed. That approach is not offering a theory, but, rather, is rejecting the philosophical enterprise that seeks one. I do not deny that, when it comes to expressiveness, music does its own thing. This is only to be expected: its medium is that of organized sound, not that of a biological organism evolved and educated to engage emotionally with its environment. What I repudiate is the suggestion that an analysis of music's expressiveness can avoid addressing if and how the musical medium realizes a kind of expressiveness that is equivalent to the biological one.

## 2.5 The qualified listener

One assumption common to the theories discussed below should be made explicit: listeners must be suitably qualified if they are to be capable of detecting and appreciating music's expressiveness. (Unprepared listeners may miss, or misidentify, the music's

[4] In its strongest version, the theory denies the possibility of analysing the crucial metaphor. For theories to which this claim is central, see Goodman (1968) and Scruton (1997).

expressive character.) Qualified listeners are at home with the type of music in question, with its genre, style, and idiom. They know when the melody begins and ends, and when the piece is over. They can recognize mistakes and can distinguish predictable from unusual continuations. They may not be able to articulate this knowledge, most of which is acquired through unreflective exposure to pieces of the relevant kind. Indeed, the majority of qualified listeners have no formal music education and are not familiar with the musicologist's technical vocabulary (Davies 1994; Kivy 1990; Levinson 1996).

## 2.6 First problem: the expression of emotion in music

When we say that something expresses an emotion, usually we mean that it publicly betrays or indicates a state that it feels. People's tears express their sadness only if they are experiencing sadness. Therefore, only sentient creatures can express emotions. Musical works are not sentient, so emotions cannot be expressed in them. Yet many of them do express emotions such as sadness and happiness. How could that be?

### 2.6.1 Music as a symbol

A first theory suggests that music operates as a symbol or sign, the import of which is purely associative and conventional. Though it bears no natural relation to an emotion, it comes to denote or refer to an emotion, and then to characterize it, by virtue of its place within a system. In this view, music picks out and conveys something about emotions after the manner of linguistic utterances; that is, through combining elements according to rules with the function of generating and communicating a semantic or propositional content (Coker 1972). Musical signs, like linguistic ones, are both unlike and opaque to their referents.

Quasi-vocabularies sometimes have been described for music in its relation to the emotions, and music is highly organized according to quasi-syntactic rules governing the well-formedness of musical strings (Lerdahl & Jackendoff 1983; Meyer 1956), yet there is not a semantics in music. Without that, the parallel with linguistic and other symbol systems collapses. It is not the case that music points or refers to emotions which it then goes on to describe. There are no plausible equivalents in music to predication, to propositional closure, or to any of the other functions and operators that are essential to the meaningful use of linguistic and other truth-functional systems.

An alternative theory would have it that music refers to the emotions not within the framework of a symbol system, but as a result of *ad hoc*, arbitrary designations and associations. For instance, certain musical gestures or phrases happen to be linked saliently with texts expressive of a given emotion and retain that connection over many years, so that purely instrumental music comes to be heard as expressive when it includes the relevant gesture or phrase (Cooke 1959). Or music of certain kinds is linked with rites or events that otherwise are emotionally charged, and these ties persist, becoming commonplaces of musical expressiveness. In this theory, expressiveness involves techniques like those followed by Wagner in his use of leitmotiv, except that the relevant conventions are available to many composers and occur in many works, so widespread and entrenched are the associations that underpin them.

There is no denying that some aspects of music's expressiveness—for instance, the links between instruments and moods, as between the oboe and bucolic frames of mind, the organ and religiosity, or the trumpet and regality or bellicosity—seem to be arbitrary and conventional in ways that may depend on historical associations. Such cases notwithstanding, this last account is no more plausible or attractive than the first. It reduces music's expressiveness to something like brute naming; it indicates how music might refer to an emotion but not how it could characterize it.

These theories regard expressive music as referring beyond itself. As with language or signs relying on arbitrary associations, features intrinsic to the music are of interest only in so far as they happen to be relevant to its role as a symbolic vehicle. Though the music mediates contact with the emotion that is symbolized, listeners should not be distracted by its intrinsic qualities from pursuing its referential target. Because it is radically different from the emotions it symbolizes, it is opaque with respect to them, yet the music is of interest only in its symbolic import.

This account is seriously at odds with the phenomenology of listeners' experiences of music's expressiveness. Registering music's expressiveness is more like encountering a person who feels the emotion and shows it than like reading a description of the emotion or than like examining the word 'sad'. While the dinner bell might, through association, lead us to salivate, we do not think of it as tasty. By contrast, we experience the sadness of music as present within it. Emotion is transparently immediate in our experience of music and our awareness of its expressiveness is not separable from, or independent of, our following the music's unfolding in all its detail. Moreover, the listener's connection is not with some general, abstract conception of the emotions, but with a specific and concrete presentation.

Any theory of musical expressiveness must acknowledge and respect the phenomenological vivacity and particularity with which music presents its expressive aspect. Here, then, is a further constraint on acceptable theories of music's expressiveness, and it is one that is failed by the theories discussed so far. Music is not merely a vehicle for referring beyond itself in a fashion that largely ignores the intrinsic and unique character of its individual works.

The semiotic theory can respond to this objection if the link between music and emotion is transparent because it is natural, not arbitrarily conventional. Here is a first suggestion: there is a synaesthetic quality to certain timbres. The trumpet's upper notes are bright and the clarinet's low register is dark; the tone of the celesta is ethereal, while high string harmonics are brittle. Even if these connections are widely made, however, they lack the temporally extended complexity that could account for music's expressiveness. They might contribute to the work's emotional ambience, but they could not generate it.

A stronger form of natural connection is that of similarity, and this is emphasized in theories regarding music as an iconic or exemplificatory symbol. If music vividly resembles the emotions it expresses—indeed, if it depicts them in virtue of these resemblances—then it would be natural to respond to the symbol much as we respond to that for which it stands. Iconic symbols (such as representational paintings) are more transparent to their referents than are signs that rely on arbitrary associations or symbol systems (such as that of a natural language) to establish the connection. We regularly talk of pictures as if we are in the presence of what they depict, though this is not to say

we are deceived by them. We do not react to linguistic descriptions in the same way. Both Langer (1942) and Goodman (1968) have suggested that music is symbolic precisely because it is experienced as resembling or exemplifying what it denotes.

What is it about the emotions that music resembles? Not their thought components if, as was just argued, purely instrumental music is not equipped to convey the contents of proposition. It has been suggested that expressive instrumental music recalls the tones and intonations with which emotions are given vocal expression (Kivy 1989), but this also is dubious. It is true that blues guitar and jazz saxophone sometimes imitate singing styles, and that singing styles sometimes recall the sobs, wails, whoops, and yells that go with ordinary occasions of expressiveness. For the general run of cases, though, music does not sound very like the noises made by people gripped by emotion. A more plausible source of resemblance lies in the music's dynamic structure than in its sound as such. We experience movement and pattern in music; we hear in music a terrain shaped by ongoing interactions between its parts, which vary in their highness, complexity, teleological impetus, energy, texture, inertia, tension, and so on. If music resembles an emotion, it does so by sharing the dynamic character displayed either in the emotion's phenomenological profile, as Addis (1999) maintains, or in the public behaviours through which the emotion is standardly exhibited.

The first of these suggestions assumes that the phenomenological profile of some emotions is distinctive enough to provide for their individuation. I am doubtful both that cognitively rich emotions, like hope or jealousy, survive being divorced from their cognitive elements and that there is anything to distinguish the internal dynamics of bursting with joy from blowing one's top. Moreover, to suggest that music symbolizes the 'general form of emotions' (Langer 1942), not particular kinds or their instances, enfeebles the account. To those who have abandoned the hydraulic in favour of the cognitive theory of the emotions, it is more promising to compare music with the outward expressions of emotions than with their experiential shape. A number of emotions have standard behavioural expressions that are partly constitutive of their nature, rather than dispensable concomitants, and these have distinctive dynamic physiognomies. A downcast bearing and slow movements go with sadness, whereas joy is upbeat and lively. Sometimes we can tell what a person is feeling from the carriage of their body, without knowing the cause of their feeling, their cognitive state, or the object of their emotion.

A fatal problem remains in explaining music's expressiveness in terms of this or any other resemblance between music's features and properties displayed by emotions: in the normal case, the pertinent behaviours are expressive only if they stand in the relevant relation to an instance of the appropriate emotion. Someone might always display the behaviour without feeling the way their behaviour leads us to suppose. In that case, no occurrent emotion is expressed. And if a given physiological state is not accompanied by relevant thoughts, attitudes, desires, or behavioural dispositions, the experience of that state would not normally be regarded as an emotion. No matter how powerful the resemblance, the analogy fails to go through, since it cannot be supposed that music experiences or undergoes the emotions expressed in it.

Theories regarding music as a sign or symbol referring to the emotions accept the conclusion of the argument with which I commenced: occurrent emotions cannot be expressed in musical works. They look for some other, more abstract, way music can

connect with the affective life. But semiotic theories inevitably leave a gap between music and emotion. In consequence, they do not do justice to the direct and unmediated fashion in which emotional expression imposes itself on our experience of the music.

### 2.6.2 Experiencing subjects: composers, listeners, and imagined personas

Most theorists accept that only sentient creatures can express occurrent emotions, but deny that this counts against music's expressiveness. They hold that, when emotion is expressed in a piece of music, that piece stands to a sentient being's occurrent emotion as expressing it. Accordingly, they seek a sentient being whose emotion is given expression by the music. The prime candidates are the composer (or performer) or a persona represented in the music. Alternatively, they maintain it is the occurrent feelings of the listener, ones caused by his or her attention to the music, that license the judgment that the music is expressive.

The *expression theory* analyses the music's expressiveness as depending on the composer expressing his or her occurrent emotion through the act of composition. The chief difficulty for this theory is conveyed by O. K. Bouwsma's aphorism: 'The sadness is to the music rather like the redness to the apple, than it is like the burp to the cider' (1950, p. 94). In other words, we experience music's expressiveness not as a residue of feelings discharged in the compositional process but as resident in its nature.

The expression theory seems to be empirically false as not all expressive music is written by composers who feel emotions and try to express them. A more philosophical point is this: in the default case, sadness is expressed by weeping and the like, not by musical composition. The connection between the composer's emotions and the work he or she writes is by no means as natural or transparent as that between his or her emotions and the behaviours, like weeping or whooping, that vent them. So, even if composers sometimes express their emotions in the works they write, this fact, rather than accounting for the music's expressiveness, needs to be explained. Indeed, in the most plausible account, the composer appropriates the music's expressiveness in order to make the connection with his or her own emotions. In other words, the composer is like the person who expresses his or her feelings, not by showing them directly, but by making a mask that wears an appropriate expression. Just as the mask is expressive whether or not it is used in this sophisticated act of self-expression, so too is the music. If composers occasionally match the expressiveness of the music to their own feelings, that is possible only because the music can present expressive aspects apart from its being appropriated in this fashion.[5] (For further criticism of the expression theory, see Davies 1987, 1994; Goldman 1995*b*; Kivy 1989; Tormey 1971.)

The *arousal theory* explains the music's expressiveness as its propensity to evoke the corresponding emotion in the listener. What makes it true that grass is green is that it

[5] Similar arguments can be ranged against the version of the expression theory that identifies the performers as the ones who express their emotions through their rendering of the music. Performers need not feel the emotions they present, and when they do, there is matching rather than direct expression.

arouses certain experiences in (human) observers under standard conditions; grass's greenness is its causal power to bring about appropriate experiences. Similarly, what makes it true that music is sad or happy is its causal power to bring about these or related responses in the listener (Matravers 1998).

I doubt that the correspondence between listeners attributing sadness to music and their experiencing feelings or emotions of sadness in response to it is sufficient to make the arousal theory plausible. In the case of colour, the experience inevitably goes with the judgment and the two are pulled apart only when the observer or the conditions of observation are abnormal. The 'standard conditions' for music to produce its effects are those in which a qualified listener pays attention to the music. Those conditions are often satisfied. When they are, the arousal of a response in listeners who correctly judge the music to be expressive is not nearly as regular as the arousal theory requires. And it is unconvincing to claim that the relevant feelings, or dispositions to them, can be so weak as to escape the listener's notice. In fact, we have a clear sense of the music's expressive character as quite distinct from our (very variable) responses to it. This is not to deny that the music sometimes can cause an emotional reaction. What is denied is that this reaction is what makes it true that the music is expressive. Normally, we regard the connection as reversed: it is because the expressiveness is apparent in the work that we are moved by the music.

Many theorists (but cf. Beever 1998) would subscribe to the following proposition: if we were never moved by music, we would not find it expressive. This involves no commitment to the arousalist's programme for analysing music's expressiveness, though. Usually the conditional is regarded as reversible: if we never found music expressive, we would not be moved by it. In other words, it identifies the close and mutual dependence of our experience of music and the judgments we make concerning its features; it does not imply that one takes explanatory precedence over the other.

Expression and arousal theories go hunting for an experiencing subject to whom the music might stand, either as the expression of his or her (the composer's) occurrent emotion or as the cause of his or her (the listener's) emotion-like response. Instead of actual persons and emotions, perhaps we should consider imagined ones. In the case of works generating fictional worlds, such as novels and films, we engage imaginatively with characters inhabiting that world. Maybe music's expressiveness connects to fictional or make-believe experiences of emotion. There are two possibilities. In the first, listeners imaginatively ascribe emotions to themselves on the basis of their make-believe engagement with the world of the work. In the second, listeners make believe that the work generates a fictional world to which they are external observers; they imagine of the music that it presents a narrative concerning the emotional life of a persona.

Both views are presented by Walton (1988), but it is his version of the first that I will consider. He suggests that a passage is expressive of sadness if the listeners imagine of their hearing of it that it is a cognizance of their own feeling of sadness. Listeners take their awareness of their auditory sensations to be an awareness of their own feelings, and it is these feelings that the music can be said to express.

Even if one charitably allows that awareness of music's expressiveness could be as self-centred and introspective as this, the theory remains implausible. Reflecting on one's auditory sensations is not plainly similar to experiencing emotions, so it is difficult

to see how what one imagines can be connected back to and controlled by the music, so that, ultimately, it is the music's expressiveness that is revealed.

The thesis that, in hearing expressiveness in music, we sometimes imagine a persona who is subject to a narrative that unfolds in the music, is widely supported (Budd 1985; Karl & Robinson 1995; Ridley 1995; Vermazen 1986; Walton 1988, 1990). The idea could be offered as a heuristic—as a way of helping people recognize the music's expressiveness—or as a claim applying only to particular works. The strongest position insists that this manner of hearing is always required for appreciation of the music's expressiveness. Levinson (1996) comes nearest to the strong position by defining musical expressiveness such that a passage is expressive of an emotion if, and only if, it is heard (by appropriately experienced listeners) as the expression of that emotion by an imagined human subject, the music's persona.

A first objection denies that all qualified listeners imagine a persona as a condition of their awareness of the music's expressiveness. They might be able to say what it would be suitable to imagine, even if they do not imagine it themselves, but they do this in terms of an awareness of the music's expressive character that is not mediated by the imagination. Besides, I contend that what the listener imagines is too little constrained by the course and detail of the music to provide a theory regarding music's expressiveness as an objective property, which is what Levinson intends. In the case of novels and films, a great deal of information about the fictional world is conveyed to the audience, even if its members must entertain the reality of this world. Those data control what is to be imagined, and why and how, in following the story. Because it does not convey a definite propositional or depicted content, and hints at such things (if at all) only in the vaguest and most general fashion, purely instrumental music cannot direct and channel the content of the listener's imagining (Davies 1997). For instance, what is to determine how many personas he or she should make believe, or the background of relations that might hold between different personas? Inevitably, what is imagined reveals more about the listener than about the music's expressiveness.

### 2.6.3 The contour theory

A final view, the *contour theory*,[6] abandons the attempt to analyse music's expressiveness as depending on its connection to occurrent emotions. It observes that certain behaviours, comportments, and physiognomies are experienced as expressive without giving expression to, or being caused by, occurrent emotions. Some faces, gaits, or movements are happy-looking. They present an emotion characteristic in their appearance. St Bernards are sad-looking dogs, but this is to say nothing about how they feel. The use of emotion terms to name the expressive characteristics of appearances is secondary, but it bears an obvious connection to those terms' primary use: the behaviours that display an emotion characteristic unconnected with an occurrent emotion are the same (or very similar) to the

[6] Kivy's version of the theory (1989) often is called 'cognitivism' as a way of acknowledging his commitment to a cognitive theory of the emotions, especially as these concern the listener's response. His theory of music's expressiveness does not invoke the cognitive theory of the emotions, however, because it denies that the music expresses occurrent emotions.

ones that, where the emotion is occurrent, give direct and distinctive expression to it. Only those emotions that can be recognized solely on the basis of the outward expressions that betray them have corresponding emotion characteristics in appearance.

Turning now to music, the contour theory proposes that pieces present emotion characteristics, rather than giving expression to occurrent emotions, and they do so by virtue of resemblances between their own dynamic structures and behaviours or movements that, in humans, present emotion characteristics. The claim is not that music somehow refers beyond itself to occurrent emotions; music is not an iconic symbol of emotions as a result of resembling their outward manifestations. Rather, the claim is that the expressiveness is a property of the music itself. This property resides in the way the music sounds to the attuned listener, just as happy-lookingness can be a property displayed in a creature's face or movements. Because music is a temporal art, its expressive character is revealed only gradually, and can be heard only through sustained attention to its unfolding. It takes as long to hear the music's expressive properties as it takes to hear the passages in which those properties are articulated.[7]

Consider Fig. 2.1. The car and the puppet are happy looking, and the dog and the weeping willow are sad looking. These attributions apply to the appearances the depicted items present, not to occurrent emotions. Only the dog is sentient, and there is no reason to think it feels as it looks. (Besides, dogs do not display feelings of sadness, when they have them, in their *faces*.) These looks present emotion characteristics because they resemble bearings or expressions which, were they shown by people under appropriate circumstances, would express those people's occurrent emotions. I maintain that, when we attribute emotions to music, we are describing the emotional character it presents, just as we do when we call the willow sad or the car happy. In the case of music, this 'appearance' depends on its dynamic topography, as this unfolds through time. In general, music resembles gaits, carriages, or comportments that are typically expressive of human occurrent emotions, rather than facial expressions.

In discussing the theory that regards music as an iconic symbol or depiction of emotions, I have already considered objections to the view that music resembles expressions of the emotions. I concluded that the resemblance claim is at its most plausible when it compares music's dynamic pattern to that apparent in non-verbal, behavioural expressions of emotion. Yet even if this is accepted, a further objection notes that resemblance alone could not ground music's expressiveness. Resemblances, which are symmetrical and, anyway, can be found between music and many things besides expressive appearances, are insufficient to explain why we experience music as powerfully expressive of emotion.

One might reply, as Kivy does, that we are evolutionarily programmed to 'animate' what we perceive. Or one might simply say 'yet this is how we hear it', without committing oneself to an account of the mechanisms and triggers that underlie the response. Not just music, but many things, are experienced as redolent of emotions, despite lacking

[7] In outline, this is the theory presented in Kivy (1989), as well as in Davies (1994), though the terms in which it is formulated are closer to my account than Kivy's. As observed earlier, Kivy is more inclined than I to find an expressive resemblance between music and the human voice (see also Juslin, this volume). For further discussion of our differences, see Davies (1994) and Kivy (1999) and for critical commentaries, see Goldman (1995*b*), Levinson (1996, 1997*b*), and Madell (1996).

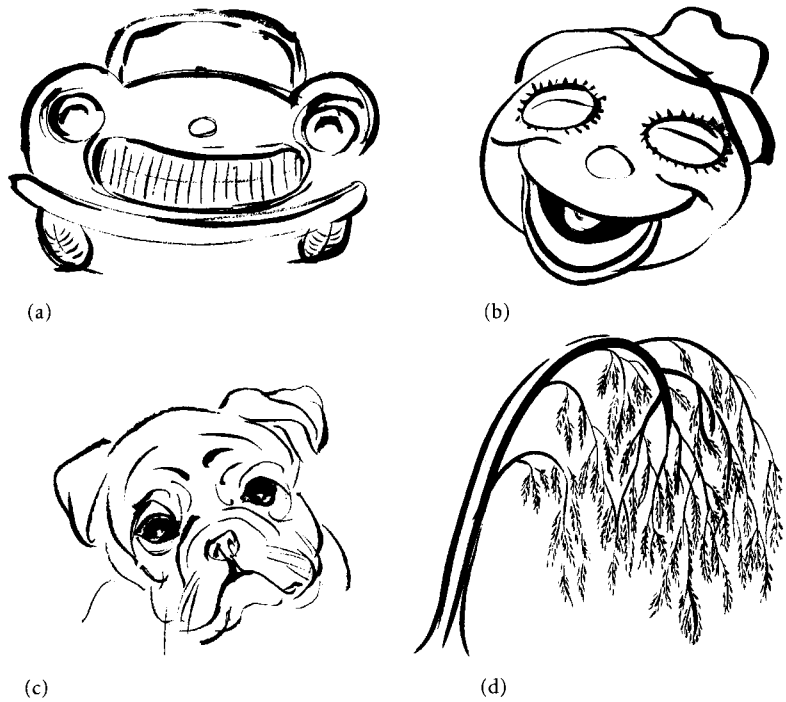



**Figure 2.1** Appearances with various emotion characteristics: (a) car, (b) puppet, (c) dog, and (d) weeping willow.

the feature one would assume to be crucial; namely, sentience. There can be no denying that crude representations of the human face can be emotionally compelling in their expressive power, though such responses are not strictly entailed by the resemblances that can be found. Consider the masks of comedy and tragedy, or a simple drawing such as Edvard Munch's 'scream' face.

If these last observations are not fully satisfying, that does not reflect more badly on the contour theory than on other analyses. For instance, the arousalist is reduced ultimately to saying 'simply, this is how music affects us'; and philosophers who regard music as an iconic symbol, or as calling on us to make believe a narrative about a persona, are no better equipped than the contour theorist to go beyond the perceived resemblances that are central to their accounts of music's expressiveness.

A different line of objection doubts that the contour theory can explain the significance we attach to expressiveness in music, or the energy with which music engages our emotions. What can we learn from, and why should we be moved by, mere appearances of emotion that are not expressions of occurrent, deeply felt emotions? One answer draws attention to the fact that music is intentionally and ingeniously designed to be as it is. Though expressiveness is a property of the piece's sounds, we encounter it not as an accident of nature but as deliberately created and used, which adds considerably to its potential importance. Another response could question if it is true that music is valued

as a source of knowledge about the emotions, rather than for the experience it provides, where this experience takes in much more than only its expressiveness.

### 2.6.4 Universalism

The contour theory, more than any other, lends itself to the idea that music is a universal 'language of the emotions'; that is, to the suggestion that expressiveness can be recognized cross-culturally. If, as some psychologists have claimed (Ekman 1980), certain emotions have characteristic appearances that are universally understood, and if music is experienced as expressive as a result of its recalling these same appearances in its dynamic character, then cross-cultural appreciation of music's expressiveness should be possible. And perhaps it is sometimes. When the musical systems of different cultures are parallel (for instance, in their principles of scalar organization and modalities), there may be sufficient transparency to allow members of one culture to correctly recognize expressiveness in the music of the other culture. Many Westerners can access sub-Saharan African music, and not only because it provided the seeds from which a number of popular Western musical types emerged.

Often, though, the music of one culture is expressively opaque to outsiders. There are several reasons why this can be so. The emotions appropriate to given circumstances can differ, so that one group sees death as an occasion for sadness where another views it as a cause for joyous celebrations. Until one appreciates the belief systems that determine the significance of the social settings in which emotions are situated, and then recognizes the connection of music with all this, it will not be a simple matter to read off expressiveness from foreign music. Even if music's expressiveness implicates 'natural' resemblances to behaviours that are transcultural in their import, these then are structured according to historically malleable musical conventions of genre and style, so that they are no longer apparent to those who lack familiarity with the culture's music. To take a crude example, whether a given pitch is 'high', 'middle', or 'low' depends on the range that is deemed available for use, and that can vary arbitrarily from musical type to type. The contour theory, no less than other analyses, supposes that qualified listeners can become such only by immersing themselves in the kinds of music that are their focus, and that listeners have no guaranteed access to the properties of foreign music, including its expressive ones, until they become appropriately experienced.

## 2.7 Second problem: mirroring responses to music's expressiveness

People often respond emotionally to musical works. While there is nothing odd about a listener being moved by the work's beauty, it is strange that he or she should respond with sadness to the sadness it expresses. The listener's sad response appears to lack beliefs of the kind that typically go with sadness. When I am sad because the dog has died, or because it is raining on your parade, or because you are depressed, I believe the death of the dog, or the rain, or your depression, are unfortunate occurrences, but when the sadness of the music makes me feel sad I do not believe there is anything unfortunate about the music. Moreover, the response to another's emotion often does not mirror it. Another's anger is as likely to produce in me fear, or disappointment, or irritation, as

it is likely to precipitate my anger. Yet the listener is not as liable to feel pity, or compassion, or evil delight at the music's sadness as he or she is liable to feel sadness. How is the listener's response appropriate to the music?

The problem is not a general one. Many of our emotional reactions to music conform neatly to the cognitivist model. We can marvel at the music's complexity, and be shocked by its discordant novelty. These responses, in taking the music as their object, involve beliefs or thoughts of the kinds that normally accompany marvelling and shock. The problem case is the one in which listeners mirror in their reactions what the music expresses; where they are saddened by sad music, or cheered by happy music.

Kivy (1989) denies the problem's existence: people are mistaken when they claim to be saddened by sad music.[8] They are moved by the music, certainly, but not to sadness. This explains why concert audiences neither display nor act as if they are sad about the music; simply, that is not how they feel. People are not often wrong about the identity of their emotions (cf. Griffiths 1997), however, and Kivy's position will fail so long as some people sometimes react to the music's expressiveness by mirroring it in their own feelings. For these reasons, and by appeal to their own experience, most philosophers reject Kivy's stance (for discussion, see Davies 1994; Goldman 1995*b*). Any alternative theory that can deal with the problem without denying the phenomenon will be preferable.

If the listener believed music expresses a sadness felt by its composer (or performer), there would be no special puzzle about his or her reaction, for such beliefs are appropriate to a sad response. In this case, however, the object of the response would not be the music, but the composer or performer. When we react to a person's emotional state, our response is directed to them, not to their expressive behaviour as such, even if it was this behaviour that alerted us to their condition. This account does not after all address the problem case, that in which the listeners' responses are solely to the music's expressive character. To have this reaction, they need not believe that the music expresses emotions experienced by its composer; it can be sufficient that they acknowledge the music's expressive appearance, without supposing this to be connected to anyone's occurrent emotions.

According to the theory in which a persona is the human subject of the imaginary act of expression we hear as going on in the music, the problem response can be approached as follows: if that response is directed to the persona, then it will be targeted at the music, for it is in the world of the music that the persona is imagined to exist. And if we hear the persona as undergoing the emotional vicissitudes outlined in the music, then we entertain thoughts about the situation of the persona that are appropriate for mirroring reactions. Admittedly, these thoughts are make believed, not believed, but if this presents no special difficulty in accounting for our reactions to fictional characters (as argued in Carroll 1998), then the response also is unproblematic in the musical case. So long as the cognitive theory of the emotions allows that the cognitive connection between the emotion and its object can be secured by the imagination in some cases, as

[8] Addis (1999) agrees that listeners are not aroused to sadness by sad music. He differs from Kivy, though, in holding both that they are aroused to a uniquely musical experience—a kind of stirring—and that this response does not take the music as its intentional object, though it is caused by the music.

unfortunate and the rest. Nevertheless, if I am roused to an emotion under those circumstances, it will be a mirroring one, because, in the absence of relevant cognitions, it is only through a kind of contagion or osmosis that my feelings are engaged.[9]

## 2.8 Third problem: negative responses

Yet if we accept that music expressive of negative emotions sometimes produces an echo in feelings experienced by the listener, another problem emerges. People avoid sad experiences where they can, because these are unpleasant. Those who are under no duty to listen to sad music often choose to do so. They report that such music gives rise to a negative emotional response, yet they offer this in praise of the music. Rather than fleeing, they are attracted to the music, and they willingly return to it, despite predicting that it will again make them feel sad. Given that music lovers are not masochists, how is this to be explained?

For Kivy (1990) there is again no problem. Listeners to sad music do not experience negative feelings, or if they do, these are of the ordinary kind—as when one is disappointed in the poverty of the work's ideas, or by its execrable execution—and provide reason for avoiding the emotion's object, the work, or the performers, in the future. Those who think that music can lead the listener to a negative, mirroring response cannot avoid the issue, though. Three argumentative strategies are available.

The first notes that there can be much to enjoy about musical works that arouse negative emotions; for instance, the work's beauty, the composer's treatment of the medium, and so forth. In addition, because it lacks 'life implications', one can savour and examine one's response, thereby coming to understand the emotion better while being reassured of one's own sensitivity (Levinson 1982). In this view, the negative elements are outweighed by positive ones. We listen to music that arouses negative emotions because it also does much more, and the overall balance is on the credit side.

The position is not entirely convincing in its present formulation. If we can get the same or similar benefits from works that do not make us feel unhappy, we should prefer them. We should shun skilful, interesting works that make us feel sad in favour of equally skilful and interesting works that make us feel happy. To reply to this objection, the original view can be developed (as in Levinson 1982) by arguing that at least some of the benefits cannot be obtained from works other than those that are liable to induce negative feelings. The Aristotelian position, according to which we are better off for purging negative feelings in the context of art, pursues this line, as does the theory that our experiences of artworks educate us about the emotions in a setting that insulates us from the practical demands and dangers of

[9] The view presented here requires rejection or revision of the cognitive theory of the emotions sketched earlier, since it countenances emotions that lack the appropriate beliefs or make beliefs, the desires, and behavioural dispositions that would follow from these, and the relevant emotional object. As indicated, Matravers (1998) and Addis (1999) also argue that the mirroring response takes the music as its emotional object. Of course, this is not to deny that the music is the focus of attention and perceptual object of the response, which is a point apparently missed by Madell (1996).

the real world. In this connection, it is also often held that the feelings experienced in regard to artworks are muted and undemanding compared to equivalents provoked by real-world situations.

These ways of addressing the objection are more convincing in the discussion of our reactions to narrative and representational artworks, not instrumental music, I find. If the response to music lacks the cognitive content of emotions, it is difficult to see how it could be a source of education or insight, or how it is easier to tolerate than similarly unpleasant feelings caused by real-world phenomena. If music does not generate a contentful fictional world, the reaction to its expressive properties is not less a response to 'real-world' features than is, say, that in which an especially vivid shade of lime green induces sensations of dyspepsia in its observer.

The second approach to the issue derives from Hume (1912), who argued of the experience of tragedy that its negative aspects are transformed to positive ones through the delight taken in the narrative's construction, the natural attractiveness of representation, and so on. It is far from clear, though, what is the character of this conversion, or how feelings such as sad ones could remain sad while becoming intrinsically pleasant. Perhaps what Hume was driving at is better articulated by the third strategy, which offers the strongest possibility for justifying the interest of someone whose sensitivities incline him or her to negative feelings on hearing music in which negative emotions are presented.

Even if we accept that the negative aspects of experience are unpleasant, and that this gives a reason for avoiding them, it is plain that, for many, this reason is not always overriding. For music, that which is negative often is integral to the whole. Provided our desire to understand and appreciate the work is strong enough, we may be prepared to face those negative elements. The experience that results is not just good on balance; it is not as if the work would be better if we could ignore its negative aspects, for then we would not be engaging with it as such. In other words, experience of, and reaction to, music's negative expressiveness, where that expressiveness is important to the work, is something to be accepted if our goal is to understand and, through understanding, to appreciate the music. There is nothing irrational in pursuing that goal, though the experience to which it gives rise can be unpleasant in parts.

There is a different way of getting at the same point. It simply is not true that people always duck the avoidable negative aspects of life. These are recognized as essential components of many things we like and value. They come along with the territory, not solely as something to be endured but also as contributing to its being the territory it is. This is true of the most important components of our lives: intimate personal relationships, child rearing, self-realization, career. To achieve a fulfilling life, the individual must honestly and seriously face these in all their dimensions, both positive and negative. Yet it also is true of the way we live generally, even apart from the big issues of survival and flourishing. Thousands of amateurs train for endurance races, such as marathons and triathlons. Other hobbies and activities, in which the challenge of the negative is no less central, are pursued with the same passionate commitment by other people, though they are under no compulsion to do so. Against this background, surely it is safe to deny there is a special problem about the fact that people willingly engage with something so rewarding as music, though they know that doing so will expose

them to expressions of negative emotions, which are liable to cause feelings that are unpleasant to experience.[10]

## References


**Addis, L.** (1999). *Of mind and music.* Ithaca, NY: Cornell University Press.

**Alperson, P.** (ed.) (1994). *What is music? An introduction to the philosophy of music.* University Park, PA: Pennsylvania State University Press.

**Alperson, P.** (ed.) (1998). *Musical worlds: New directions in the philosophy of music.* University Park, PA: Pennsylvania State University Press.

**Beever, A.** (1998). The arousal theory again? *British Journal of Aesthetics*, **38**, 82–90.

**Bouwsma, O. K.** (1950). The expression theory of art. In *Philosophical analysis*, (ed. M. Black), pp. 71–96. Englewood Cliffs, NJ: Prentice-Hall.

**Budd, M.** (1985). *Music and the emotions: The philosophical theories.* London: Routledge.

**Budd, M.** (1995). *The values of art: Pictures, poetry, and music.* London: Penguin Press.

**Carroll, N.** (1998). *A philosophy of mass art.* New York: Oxford University Press.

**Coker, W.** (1972). *Music and meaning: A theoretical introduction to musical aesthetics.* New York: Free Press.

**Cooke, D.** (1959). *The language of music.* London: Oxford University Press.

**Davies, S.** (1987). Authenticity in musical performance. *British Journal of Aesthetics*, 27, 39–50.

**Davies, S.** (1994). *Musical meaning and expression.* Ithaca, NY: Cornell University Press.

**Davies, S.** (1997). Contra the hypothetical persona in music. In *Emotion and the arts*, (ed. M. Hjort & S. Laver), pp. 95–109. Oxford, UK: Oxford University Press.

**Deigh, J.** (1994). Cognitivism in the theory of the emotions. *Ethics*, **104**, 824–54.

**Ekman, P.** (1980). Biological and cultural contributions to body and facial movements in the expression of the emotions. In *Explaining emotions*, (ed. A. O. Rorty), pp. 73–101. Los Angeles: University of California Press.

**Godlovitch, S.** (1998). *Musical performance: A philosophical study.* London: Routledge.

**Goehr, L.** (1992). *The imaginary museum of musical works: An essay in the philosophy of music.* Oxford, UK: Clarendon Press.

**Goehr, L.** (1998). *The quest for voice: On music, politics, and the limits of philosophy.* Oxford, UK: Clarendon Press.

**Goldman, A. H.** (1995*a*). *Aesthetic value.* Boulder, CO: Westview Press.

**Goldman, A. H.** (1995*b*). Emotion in music (a postscript). *Journal of Aesthetics and Art Criticism*, **53**, 59–69.

**Goodman, N.** (1968). *Languages of art.* Indianapolis, IN: The Bobbs-Merrill Company.

**Gordon, R.** (1987). *The structure of the emotions.* Cambridge, UK: Cambridge University Press.

**Gracyk, T. A.** (1996). *Rhythm and noise: An aesthetics of rock music.* Durham, NC: Duke University Press.

**Griffiths, P.** (1997). *What emotions really are.* Chicago: University of Chicago Press.


[10] I thank the editors, Jerrold Levinson, and an anonymous referee for comments on this chapter, and Vivian Ward for help with the artwork.

**Hatfield, E., Cacioppo, J. T., & Rapson, R. L.** (1994). *Emotional contagion.* New York: Cambridge University Press.

**Higgins, K.** (1991). *The music of our lives.* Philadelphia, PA: Temple University Press.

**Hjort, M. & Laver, S.** (ed.) (1997). *Emotion and the arts.* Oxford, UK: Oxford University Press.

**Hume, D.** (1912). Of tragedy. In *Essays moral, political and literary,* (Vol. 1) (ed. T. H. Green & T. H. Grose), pp. 258–65. London: Longmans, Green, & Co (originally published 1777).

**Karl, G. & Robinson, J.** (1995) Shostakovitch's tenth symphony and the musical expression of cognitively complex emotions. *Journal of Aesthetic and Art Criticism,* 53, 401–15.

**Kivy, P.** (1988). *Osmin's rage: Philosophical reflections on opera, drama and text.* Princeton, NJ: Princeton University Press.

**Kivy, P.** (1989). *Sound sentiment.* Philadelphia, PA: Temple University Press.

**Kivy, P.** (1990). *Music alone: Philosophical reflection on the purely musical experience.* Ithaca, NY: Cornell University Press.

**Kivy, P.** (1995). *Authenticities: Philosophical reflections on musical performance.* Ithaca, NY: Cornell University Press.

**Kivy, P.** (1997). *Philosophies of art: An essay in differences.* New York: Cambridge University Press.

**Kivy, P.** (1999). Feeling the musical emotions. *British Journal of Aesthetics,* 39, 1–13.

**Langer, S. K.** (1942). *Philosophy in a new key.* Cambridge, MA: Harvard University Press.

**LeDoux, J. E.** (1998). *The emotional brain.* London: Weidenfeld & Nicolson.

**Lerdahl, F. & Jackendoff, R.** (1983). *A generative theory of tonal grammar.* Cambridge, MA: MIT Press.

**Levinson, J.** (1982). Music and negative emotion. *Pacific Philosophical Quarterly,* 63, 327–46.

**Levinson, J.** (1990). *Music, art, and metaphysics.* Ithaca, NY: Cornell University Press.

**Levinson, J.** (1996). *The pleasures of aesthetics.* Ithaca, NY: Cornell University Press.

**Levinson, J.** (1997*a*). *Music in the moment.* Ithaca, NY: Cornell University Press.

**Levinson, J.** (1997*b*). Emotion in response to art: A survey of the terrain. In *Emotion and the arts,* (ed. M. Hjort & S. Laver), pp. 20–34. Oxford, UK: Oxford University Press.

**Levinson, J.** (1998). Emotion in response to art. In *Routledge Encyclopedia of Philosophy,* (Vol. 3) (ed. E. Craig (chief ed.) & M. Budd (subject ed.)), pp. 273–81. London: Routledge.

**Lyons, W.** (1980). *Emotion.* New York: Cambridge University Press.

**Madell, G.** (1996). What music teaches about emotion. *Philosophy,* 71, 63–82.

**Matravers, D.** (1998). *Art and emotion.* Oxford, UK: Clarendon Press.

**Meyer, L. B.** (1956). *Emotion and meaning in music.* Chicago: University of Chicago Press.

**Ridley, A.** (1995). *Music, value and the passions.* Ithaca, NY: Cornell University Press.

**Robinson, J.** (ed.) (1997). *Music and meaning.* Ithaca, NY: Cornell University Press.

**Scruton, R.** (1997). *The aesthetics of music.* Oxford, UK: Clarendon Press.

**Solomon, R. C.** (1976). *The passions.* Garden City, NY: Anchor Press/Doubleday.

**Thom, P.** (1993). *For an audience: A philosophy of the performing arts.* Philadelphia, PA: Temple University Press.

**Tormey, A.** (1971). *The concept of expression.* Princeton, NJ: Princeton University Press.

**Vermazen, B.** (1986). Expression as expression. *Pacific Philosophical Quarterly,* 67, 196–224.

**Walton, K. L.** (1988). What is abstract about the art of music? *Journal of Aesthetics and Art Criticism*, **46**, 351–64.

**Walton, K. L.** (1990). *Mimesis as make-believe: On the foundations of the representational arts.* Cambridge, MA: Harvard University Press.

CHAPTER 3

# MUSICOLOGICAL APPROACHES TO EMOTION

NICHOLAS COOK AND NICOLA DIBBEN

## 3.1 Defining musicology

Writing about music may well go back as far as writing itself, but to speak of 'musicology' before about 1800 is to court anachronism. In the centuries before then, there were in the West two broadly distinct traditions of writing about music: on the one hand more or less practical manuals addressed in general to (and generally by) composers and performers, and on the other hand more or less speculative writings that, to modern eyes, can at times resemble cosmology with a musical vocabulary. The emergence of 'musicology' (a term not used in English until the twentieth century, but borrowed from the earlier French term *musicologie*) can, then, be traced to two circumstances: first, the new relationship between compositional theory and practice that resulted from a general growth in historical awareness and the establishment of a canon of 'masterworks'; and second, the increasing systematization of what in the course of the nineteenth century became known (especially in Germany) as the human sciences.

In its original form, and to some extent still in common British parlance, 'musicology' means simply the study of and knowledge about all aspects of music, taking in, for example, the systematic approaches to musical organization collectively known as 'music theory'; American usage of the term is distinctly narrower: in Joseph Kerman's words, 'the study of the history of Western music in the high-art tradition' (Kerman 1985, p. 11). In the 1990s, however, the self-imposed limitations of this version of musicology were seriously challenged by the so-called 'new' musicology (a development foreshadowed by Kerman's critique of the traditional discipline's formalist and positivist tendencies). There is, then, no single, universally accepted definition of the discipline's scope.

This poses problems in determining the scope of this chapter. A genuinely inclusive treatment of musicological approaches to emotion would not restrict itself to the work of 'musicologists' in any narrow sense, but might encompass the views of performers, composers, aestheticians, and others who wrote on music before 1800; similarly, it might take in not only present-day musicological writing but also the ethnomusicological, philosophical, psychological, or sociological perspectives that inform it. (It might also recognize the existence of other traditions of systematic thinking about music, other musicologies, outside the Western orbit.) But that would take a book and not a chapter! More modestly, then, we discuss here a selection of musicological or

music-theoretical approaches to the expression of emotion in Western 'art' and popular music, most of them written since 1950, emphasizing in particular those that set out general frameworks within which to understand how specific pieces of music embody, express, arouse, or otherwise signify emotional states. (We recognize, and readers should be aware, that other authors might well have made a different selection.) Bearing in mind the likelihood that many of our readers will be unfamiliar with the discipline, however, we precede this with a whistlestop tour of thinking about music and emotion since around 1600: this will serve as much to contextualize musicology as to identify issues and points of reference that will be drawn on later in the chapter.

## 3.2 A historical sketch

While comparatively little is known of the nature and practice of ancient Greek music, classical writings about music have formed a constant point of reference for subsequent theorists and have accordingly exerted a lasting influence. One could almost frame a history of thinking about music and emotion in terms of the two principal functions which the Greeks ascribed to music: on the one hand *mimesis* (the imitation or transformation of an external reality) and, on the other, *catharsis* (the purification of the soul through emotional experience). The first term values music for its representational function, in this sense embracing it within the theory of knowledge, whereas the second locates music's value in the effect it makes upon the experiencing subject. At issue are two fundamentally different conceptions of what sort of thing music is. As might be expected, however, such philosophical concerns did not greatly impinge upon the practice of working musicians, who were less likely to discuss issues of music's emotional powers than simply to take them for granted; a comprehensive history of pre-musicological approaches to emotion would have to be extracted from a multitude of oblique references hidden in the nooks and crannies of a wide range of texts.

The rise of opera around 1600, with which we begin this overview, was associated with a long period of ascendancy in the idea of music as *mimesis*. (It is perhaps not irrelevant that the Florentine *camerata* responsible for this development saw themselves, however misguidedly, as reconstructing the practices of ancient Greek drama.) The function of music in baroque opera is to reflect or heighten the expression of the emotions signified by the words and presented by means of staged action; we can talk of the music 'expressing' emotion just to the extent that the emotion is located not within the music as such, but within an external reality (whether actual or imagined) which the music references. Music, in short, represents reality, just as language does. It is not surprising, then, that baroque musicians borrowed wholesale from contemporary thinking about language, seeking to understand even instrumental music as a form of discourse and, in particular, applying to it the principles of classical rhetoric: an eighteenth-century example is Mattheson (1981), who traced the operation in music of such rhetorical devices as repetition, variation, or counterstatement, seeing them as serving to project or heighten emotional content in the same way that, in language, the devices of rhetoric colour or inflect propositional content. And in each case the aim was persuasion: just as legal or political orators would seek to persuade their listeners of the case being presented, so musicians

sought to convince audiences of their emotional veracity. Like actors, musical performers had themselves to feel the emotions they wanted to convey to their listeners: as C. P. E. Bach put it in his *Essay on the true art of playing keyboard instruments*, 'A musician cannot move others unless he too is moved' (Bach 1974, p. 152).

During the seventeenth and eighteenth centuries, vocal music—that is, music allied to words—was understood as the paradigm case for music in general. It was through heightening verbal signification that music itself acquired meaning; purely instrumental music by contrast had no meaning, or at best signified at second hand, functioning as a more or less pale reflection of texted music. Unsurprisingly, then, music fared badly in the formulation of the eighteenth-century discipline of aesthetics; music can arouse emotions, Kant admitted, but it amounts to no more than 'a play with aesthetic ideas . . . by which in the end nothing is thought'. (He added that the pleasure it creates 'is merely bodily, even though it is aroused by ideas of the mind, and . . . consists merely in the feeling of health that is produced by an intestinal agitation corresponding to such play'.[1]) It is in this light that the much vaunted rise around 1800 of autonomous instrumental music has to be understood: the point is not that people did not play purely instrumental music up to that time, but that it was not the paradigm case for music that it became by the middle of the nineteenth century. Lydia Goehr has characterized this conceptual transformation as a conjunction of a '*formalist* move which brought meaning from the music's outside to its inside' and a '*transcendent* move from the worldly and the particular to the spiritual and the universal' (Goehr 1992, p. 153).

The formalist move to which Goehr refers is most famously represented in the writings of the Viennese critic Eduard Hanslick, the first edition of whose short book *On the musically beautiful* appeared in 1854. In describing the 'content' of music as 'tonally moving forms' (Hanslick 1986, p. 29), Hanslick was claiming that questions about musical meaning, which had long been formulated in terms of representation, should be reformulated as questions about the intrinsic properties of the music itself. The clearest indication of what is at stake here is to be found in those later writers, such as the turn-of-the-century Viennese theorist Heinrich Schenker, who sought to build systems of musical analysis on Hanslickian foundations: commenting on C. P. E. Bach's prescriptions of how the improviser is to excite or soothe the passions, Schenker claims that 'One must not seek in Bach's word 'passions' [*Leidenschaften*] what certain aestheticians of the doctrine of affections bring to it . . . [Bach] means by it simply the consequences of a change in diminution: pure musical effects which have nothing in common with the amateurishly misunderstood and so grossly exaggerated ideas of the aestheticians' (Schenker 1994, p. 5).[2] (This development of the scope of musical analysis, in effect elaborating on the structural potential of 'tonally moving forms', might be seen as a definitive rebuttal of Kant's dismissive reduction of music to the level of intestinal

[1] This passage (from Kant's *Critique of judgement*) is quoted and discussed by Kivy, who scornfully encapsulates Kant's image of music as 'the sonic counterpart of Tums for the tummy' (Kivy 1993, pp. 258–9).

[2] The passage in Bach's *Essay on the true art of playing keyboard instruments* to which Schenker refers is on pp. 438–9 of Mitchell's translation (Bach 1974).

agitation.) It is in this sense that, as Goehr puts it, meaning was brought from the music's outside to its inside.

Released from its representational function, music is no longer seen as describing specific emotions within specific situations (the Count's mounting jealousy in Act 1 of Mozart's *Marriage of Figaro*, say), but rather as abstracting the essence of phenomena; this is the move from the particular to the transcendental of which Goehr speaks. Just as in the work of early romantic artists like Caspar David Friedrich, personal emotions are subsumed within a generalized aesthetic yearning for the universal and the spiritual; for the influential writer E. T. A. Hoffmann, writing in the early years of the nineteenth century, music's power lay precisely in its ability to transcend the conditions of ordinary existence, a view echoed in Hanslick's statement that the composer creates 'something which has no counterpart in nature and hence none in the other arts, indeed none in this world' (Hanslick 1986, p. 74). Seen in this light, music, once the poor relation of the visual and literary arts, could be considered uniquely privileged within the aesthetic hierarchy: the very lack of specificity that formerly consigned it to a subordinate role was now construed as an infinite suggestiveness. It is precisely because of its lack of direct reference to the specifics of everyday existence that the art historian Walter Pater (1910, p. 135) wrote that 'All art constantly aspires to the condition of music'.

But it is at this point that the story takes a series of unexpected turns, which between them can be seen as accounting for the fragmented and in many ways contradictory nature of present-day thinking about music, and particularly about musical meaning. For no sooner had music made itself independent of the word than the word began to re-establish itself, only this time outside rather than inside the musical text. It was during the nineteenth century that the consumption of music began to be routinely surrounded by a variety of informative and interpretative texts: programme notes, music appreciation texts, and, today, CD liners, broadcasts, and magazines. This development was largely associated with the reception of Beethoven's instrumental music, and in particular the increasingly idiosyncratic, not to say erratic, works of his late period. Some contemporary listeners put their hermetic style down to the composer's now profound deafness. Others felt that, behind the music's apparent contradictions (such as the consistent mixing of genres in the Ninth Symphony) there must lie some intention, some deeper message. And so there emerged a plethora of commentaries on the music, each attempting to work from the musical text to the meaning assumed to lie behind it. The most influential of these commentaries understood the music in autobiographical terms. In 1828, for example, Franz Joseph Fröhlich published a review of the Ninth Symphony, interpreting it as the expression of Beethoven's struggle with deafness and his ultimate transmutation of suffering into joy (Wallace 1986). Fröhlich even brought this interpretation to bear upon the bar-by-bar details of the music: bars 1–30 of the first movement, he explained, represent in succession tender longing, heroic strength, pathos, and a vision of joy—all of them complementary aspects of Beethoven's own, complex personality.

Commentaries of this kind, which maintained their popularity among the concert-going public during the remainder of the nineteenth century, set up a number of striking paradoxes. The first we have already referred to by implication: in Scott Burnham's words, 'music no longer in need of words now seems more than ever in need of words' (Burnham 1999, p. 194). Secondly, and related to this, music now vindicated as the most

abstract of the arts is explained in the most concrete terms; it is worth observing how the music programme tradition which runs from Berlioz and Liszt to Richard Strauss (and from there, arguably, into film music) in effect brought such explanations back within the music itself, so setting itself up in opposition to the prevailing aesthetic of autonomous instrumental music. Thirdly, meaning now understood as located within the musical text (the very term applied to such commentaries, 'hermeneutics', is borrowed from traditions of biblical exegesis) is explained through the construction of an authorial persona whose experiences the music represents. Most telling of all, however, is the manner in which music is now construed as an expression of what might be termed bourgeois subjectivity.

As presented in baroque and classical opera, emotion belongs to the public sphere; to put it another way, it is conceived dramatically. The paradigm of nineteenth-century literature, by comparison, is the novel, where emotion subsists in private, subjective experience, to which the various characteristic devices of the novel give the reader access. In the same way, the emotion which nineteenth-century commentators read into Beethoven's music belongs to the private sphere, unmistakably so when Wagner describes the Ninth Symphony as Beethoven's attempt to reach out from the solitude of his deafness: 'When you meet the poor man, who cries to you so longingly, will you pass him on the other side if you find you do not understand his speech at once?' (Wagner 1899, p. 203). Music is still understood as expression, just as under the principle of *mimesis*, but this now takes the form of self-expression, the attempt of the artist to make contact with others and, so to speak, with himself or—more rarely—herself. However dated its origins, this is a concept that has never really lost currency in the years since the 1820s, though it is nowadays perhaps most evident in popular music: music videos, for instance, promise the fan a means of intimate access to the star, while the values of authentic self-expression are central to the carefully constructed persona of an artist like Bruce Springsteen (Cavicchi 1999).[3] It is in the contrast which Adorno once drew between the emotional qualities of Schoenberg's music and the earlier model of dramatic representation, however, that the idea of music as self-expression reaches its apogee: 'Passions are no longer simulated', he wrote, instead, 'genuine emotions of the unconscious—of shock, of trauma—are registered without disguise through the medium of music' (Adorno 1973, pp. 38–9). For the composer, then, and perhaps for the listener too, music provides something akin to the purification of the soul through emotional experience that the Greeks associated with it, and so assumes the function of *catharsis*.

Nowadays, formalism and hermeneutics are seen as opposite ends of the critical spectrum. But as we have explained, it was not always so: formalism, the idea that music is autonomous and should be understood in its own terms, was originally associated with the sense of music's other-worldliness and consequent spiritual value. Nor did Hanslick seek to deny music's emotional powers, although later formalists often read him this way. (Hence his statement in the foreword to the eighth edition of *On the musically*

[3] For a study of the bourgeois subject that takes in both romantic *Lieder* and rock, see Bloomfield (1993); for an example of popular music which problematizes this expressive norm, see Dibben (2001).

*beautiful*—published in 1891, 37 years after the book's first appearance—that 'I share completely the view that the ultimate worth of the beautiful is always based on the immediate manifestness of feeling' (Hanslick 1986, p. xxii).) He did, however, seek to qualify its ability to represent specific emotions, putting forward an early version of what is nowadays called the cognitive theory of the emotions (Kivy 1993, pp. 270, 284; see also Davies, this volume). You cannot feel love, Hanslick claims, without thinking of the loved one; the emotion is inextricably tied up with its object. But depending on the circumstances 'its dynamic can appear as readily gentle as stormy, as readily joyful as sorrowful, and yet still be love'. His point is that 'music can only express the various accompanying adjectives and never the substantive, e.g. love itself' (Hanslick 1986, p. 9). But how does it do this? In order to answer this question, Hanslick asks another (p. 11):

> What, then, from the feelings, can music present, if not their content? Only that same dynamic . . . It can reproduce the motion of a physical process according to the prevailing momentum: fast, slow, strong, weak, rising, falling. Motion is just one attribute, however, one moment of feeling, not the feeling itself.

His conclusion, then, is that it is these purely musical motions, and not the feelings with which they may be associated, that must form the focus of any meaningful theory or criticism of music.

It would be possible to argue that what Hanslick means is that only when you can fully explain music in *motional* terms is there any point in trying to do so in *emotional* ones. Because you can never fully explain anything in the arts, however, that is a recipe for infinite deferral. At all events, and whatever Hanslick himself may have intended, he laid the foundations for a formalism that grew in strength during the later part of the nineteenth century and survived through the greater part of the twentieth, according to which issues of musical meaning and expression were regarded as off limits for purposes of serious academic discussion. This change in the nature of formalism coincided with the development of increasingly abstract and powerful methods for musical analysis (through the work of not only Schenker but also Schoenberg and Hindemith in the first half of the twentieth century, for example, and Babbitt, Forte, and Lerdahl in the second half), the aim of which was to provide complete and self-sufficient explanations of music in exclusively structural terms. That did not necessarily entail a denial that music might have emotional powers. Lerdahl and Jackendoff (1983, p. 8) specifically state that

> To approach any of the subtleties of musical affect, we assume, requires a better understanding of musical structure. In restricting ourselves to structural considerations, we do not mean to deny the importance of affect in one's experience of music. Rather we hope to provide a stepping stone toward a more interesting account of affect than can at first be envisioned.[4]

But the practical effect is, of course, the indefinite postponement of consideration of music's expressive qualities to which we have already referred. In what follows, we shall consider the viability of the separation that this implies between structural analysis and expressive interpretation. For now it is worth just noting its apparent similarity to the traditional piano teacher's advice formulated by Ralph Kirkpatrick as 'Learn the notes

[4] Cf. Edward T. Cone's claim that 'If verbalization of true content—the specific expression uniquely embodied in a work—is possible at all, it must depend on close structural analysis' (Cone 1982, p. 233).

and then put in the expression'; tellingly, Kirkpatrick continues, 'My admonition is to learn the notes and understand their relationships, and then to draw the expression out' (Kirkpatrick 1984, p. 128).

By the 1980s there was a growing view that this narrowness of analytical purview, coupled to a parallel reluctance to enter upon aesthetic interpretation on the part of music historians, had resulted in an apparently unbridgeable schism between the concerns of professional musicologists on the one hand, and those of practically everyone else who had an interest in music on the other: as Kerman put it, 'Along with the preoccupation with structure goes the neglect of other vital matters—not only the whole historical complex . . . but also everything else that makes music affective, emotional, expressive' (Kerman 1985, p. 73). In this context one can see the 'new' musicology that followed as less new than revisionary, aiming to reinstate something of the more generous purview of musical interpretation as practised in the middle of the nineteenth century. Indeed one of the vehicles of the 'new' musicology was the revival of a self-styled hermeneutic criticism tacitly modelled on the nineteenth-century version.[5] Leo Treitler has made the acid observation that such criticism is 'not different in form or verisimilitude from the sort of nineteenth-century hermeneutic that interpreted Beethoven's Ninth Symphony in images drawn from Goethe's *Faust*' (Treitler 1999, p. 370).[6]

There is a crucial difference, however, for the 'new' musicological revival of hermeneutics was supplemented by a strain of social critique deriving primarily from the work of Adorno. It was most conspicuous in the work of Susan McClary, who gained a great deal of notoriety through her critical interpretations of the war-horses of the symphonic repertory: most famously, she likened parts of the Ninth Symphony to a rapist's murderous fantasy (McClary 1991, pp. 128–9), contrasting what she saw as the violent and misogynous emotions expressed by Beethoven's music with the altogether more accommodating and socialized expression found in the music of the contemporary but (arguably) homosexual Schubert. One might see this as a frivolous or mischievous application of traditional hermeneutical procedures, involving the construction of a clearly unhistorical compositional persona (and McClary has been criticized on precisely these grounds). But that is to miss her substantial point, which is that music serves to express not only personal emotions but also social ideologies, in this case a characteristic attitude to women or, more generally, to sociocultural differences. While many aspects of McClary's work have been subject to critique, the broadened musicological purview represented by her work has been widely accepted. And as a result, the long marginalized topic of musical meaning has become central to the discipline.

## 3.3 Some classic approaches

As we have seen, one strand which has dominated thinking about musical expression is a reliance on linguistic models. The notion of a shared musical 'language' is particularly

[5] See, for instance, Kramer (1990), along with McClary (1991), one of the key texts of the 'new' musicology; for a brief overview, focusing on issues of gender representation, see Cook (1998a, Chapter 7).

[6] The reference is to the programme note that Wagner drew up for his 1846 performance of the symphony in Dresden.

prominent in baroque music treatises which, influenced by newly rediscovered Greek and Roman doctrines of oratory and rhetoric, provided instruction on how the composer was to 'move' the 'affections' or emotions of the listener.

A primary aesthetic goal in baroque music was to achieve stylistic unity based on the representation of a single emotion—a phenomenon exemplified by the da capo aria, so called because its final section repeated the first one da capo, 'from the top'. Affections were conceived as typified and static attitudes of mind, expressed in music through specific figures and compositional techniques identified and discussed in contemporary theoretical sources (one of the earliest was Burmeister's *Musica Autoschediastike* of 1601). The approach derived from the rhetorical concept of *decoratio*, that is, rules and techniques for embellishing the ideas of an oration and infusing it with passion. Applied to music, devices such as melodic repetition, fugal imitation, and the use of dissonance were invoked in order to explain and justify irregular contrapuntal writing, in much the same way that nineteenth-century commentators proposed hermeneutical explanations for the idiosyncracies of Beethoven's style. Gesturing was also part of rhetorical delivery (itself an object of study for politicians, ministers, and actors), and was theorized in the form of treatises on gestures in everyday life, and on rhetorical delivery and opera acting (Fig. 3.1) (Solomon 1989). The

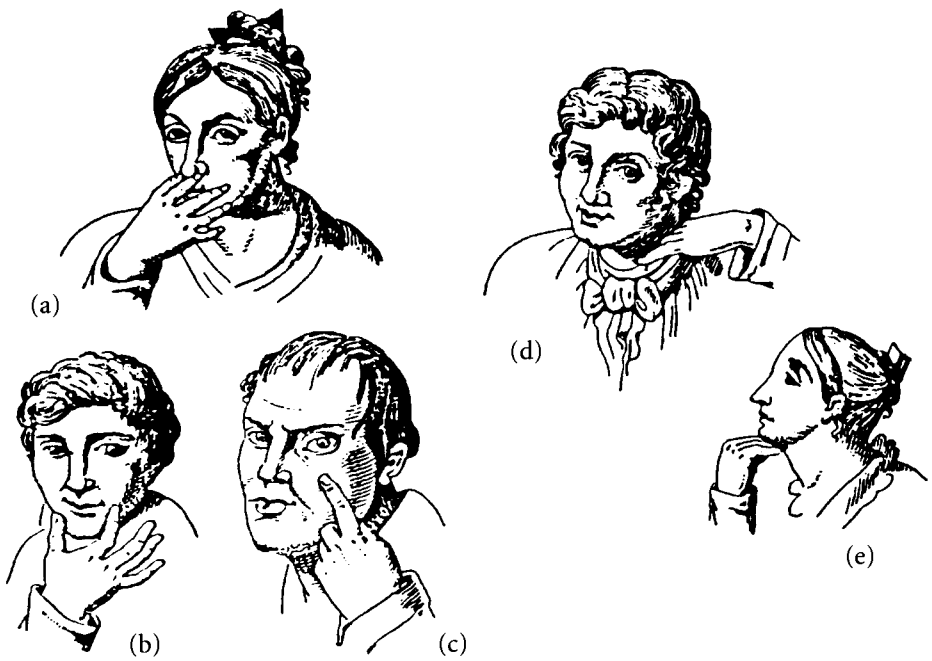



Figure 3.1 Gestures with the hand and face: (a) astuteness, (b) beauty, (c) carefulness, (d) deception, and (e) the negative (from Andrea di Joria, *La mimica degli antichi investigata nel gestire napoletano*, Napoli, 1832, reproduced in Solomon 1989).

precise bodily gestures and movements associated with the representation of specific attitudes of mind outlined in these texts seem perversely static and unreal from the standpoint of nineteenth- and twentieth-century constructions of subjectivity; in the same way, the da capo aria has long been criticized for the artificial nature of its return to the opening, as if emotions could be repeated to order. All this reflects a fundamental shift of conception: whereas the eighteenth-century affection is a rationalized attitude of mind, the nineteenth-century understood emotion as a personal and spontaneous expression or experience. (The listeners of Fig. 3.2, eyes turned away from the performers or head buried in hands, embody the migration of emotion from the outer to the inner world.)

In the later part of the eighteenth century, there was a shift from the earlier emphasis on *decoratio* to a lexicon of affective types or 'topics', as they have been termed by Ratner (1980); this was linked to an increasingly dynamic view of musical expression, whereby a single piece might encompass different emotions, or the transition between one emotion and another. (For contemporary listeners this was one of the most striking and innovatory features of Mozart's operatic arias and ensembles.) Basing his work on contemporary treatises as well as the study of musical texts, Ratner showed how classical compositions evoked, often in kaleidoscopic succession, a more or less fixed repertory of melodic, rhythmic, textural, and other types, each of which had its own specific connotation. Such types might refer, for example, to particular dances (minuet, gigue),



Figure 3.2 Beethoven playing for his friends (lithograph by Albert Graefle, c. 1877).

styles (the 'brilliant' style, the 'Turkish' style), compositional procedures (the 'learned' style), or extramusical associations (the hunt).

This notion of an affective lexicon has been further developed by means of semiotic theory, in relation to both 'art' music (Agawu 1991; Allanbrook 1983) and popular music (Tagg 1982); it has also been interpreted in terms of cognitive theory (Gjerdingen 1988). All such approaches understand expressive meaning as the product of historical usage: music signifies through reference to established syntactical types (the V–I progression and motivic liquidation that encode closure in a Mozart string quartet, say), to other music and its circumstances of performance (the parade ground, for instance, in the case of an orchestral march), or to some kind of extramusical reality. It is crucial to such approaches that these meanings were deliberately employed by composers and understood as such by contemporary audiences. For example, Agawu (1991, p. 31) cites a letter from Mozart to his father in which he discusses his compositional strategy in the opera *Die Entführing aus dem Serail*, and the likely impact of particular features on the audience. Mozart refers specifically to his use of the 'Turkish' style to create a comic effect during the scene in which Osmin expresses his rage, as well as to the use of syntactical features such as a change of key and metre at the end of a particular aria in order to surprise his audience.[7]

Because of the assumption that musical meaning results from historical sedimentation, the application of any given lexical approach of this kind is necessarily restricted to a specific repertory. By contrast, the aim of Deryck Cooke's *The language of music*, first published in 1959, is to provide not only a lexicon of but also an explanatory basis for the emotional meanings of Western 'art' music in general (and perhaps even, as we shall see, for music as a universal phenomenon). In Cooke's words, his book 'attempts to show that the conception of music as a language capable of expressing certain very definite things is not a romantic aberration, but has been the common unconscious assumption of composers for the past five-and-a-half centuries at least' (Cooke 1959, p. xi). He locates his basic expressive lexicon in the tonal and intervallic 'tensions' embodied in specific scale steps and the patterns of motion between them, seeing other musical dimensions such as timbre and texture as 'characterising agents', that is, as merely modifying the tensions established through pitch, time, and dynamics (Cooke 1959, pp. 37–8). He substantiates his theory by citing large numbers of extracts taken from vocal works, the texts of which embody similar expressive meanings. The assumption is that the emotional associations of particular musical formations apply equally when these formations occur in autonomous instrumental music. After going through this process with each of the different scale steps, he draws up a summary of what he terms 'the basic

[7] Such lexicons had a kind of after-life in early cinema history, in the form of the cue sheets, manuals, and handbooks that provided guidance to the performers, usually pianists, who accompanied silent films. These volumes contained both transcriptions from the classical repertoire and originally composed music, categorized according to the narrative situations and emotions engendered by it (see also Cohen, this volume). One of the earliest anthologies (*Motion picture piano music: Descriptive titles to fit the action, character or scene of moving pictures*, cited in Marks 1992, p. 68) contains 51 short pieces with titles such as 'Aged Colored Man, Aged Persons, Ancient Dance, Andante, Antique Dance, Apparitions', listed alphabetically.

expressive functions of all twelve notes of our scale', beginning as follows (Cooke 1959, pp. 89–90):

> *Tonic:* Emotionally neutral; context of finality. *Minor Second:* Semitonal tension down to the tonic, in a minor context: spiritless anguish, context of finality. *Major Second:* As a passing note, emotionally neutral. As a whole-tone tension down to the tonic, in a major context, pleasurable longing, context of finality. *Minor Third:* Concord, but a 'depression' of natural third: stoic acceptance, tragedy. *Major Third:* Concord, natural third: joy'

Philip Tagg's (1982) application of semiotic theory to popular music is reminiscent of Cooke's approach in both concept and method. He identifies a range of signs in popular music, equating particular emotions, moods, and meanings with particular harmonies, melodies, timbres, and so forth, and referring to them as 'musemes' or minimal units of musical meaning; the term is based on 'phoneme', and is thus a further illustration of the indebtedness of theories of musical emotion to linguistic models.[8] Tagg verifies his interpretations in two ways. The first of these is what he calls 'inter-subjective comparison', by which he simply means establishing the consistency of attributions of meaning to the same music played to different respondents. The second and better known is what he terms 'inter-objective comparison' (IOC): this means the substitution for a given museme of others drawn from comparable expressive contexts, especially ones involving words, in order to confirm their semantic equivalence. (Again the reference to texted music provides a link with Cooke's method.) Tagg demonstrates his theory through the detailed analysis of specific pieces of music, the most extended examples being his analyses of the *Kojak* TV theme tune and the Abba hit *Fernando.* The readings that result are more complex than Cooke's in that they do not simply bring to light emotional contents but attempt to define an attitude towards the ostensible meaning of the lyrics: in the case of *Fernando,* for instance, Tagg concludes that the music gives the lie to the revolutionary sympathies expressed by the words.

There are a number of criticisms which can be levelled equally at Cooke and Tagg.[9] One is that both their methods are heavily dependent upon extramusical aspects, particularly accompanying lyrics or other programmatic elements. Although such elements are obviously important where present, how do we know what happens in music from which they are absent and the connotations of which are therefore less clear (Middleton 1990, p. 234)? How confident can we be that what applies in the one case applies in the other? (Does music derive its emotional meanings from purely historical associations with certain texts, or were those texts associated with the music because of some

[8] Tagg's concept of sign, however, is broader than Cooke's: as well as 'style indicators', which refer to compositional 'norms', and 'genre synecdoche', referring to a 'foreign' style and hence to the genre and culture to which it belongs, Tagg identifies 'anaphones', which share a structural homology with the sonic, kinetic, or tactile events that they signify, and 'episodic markers', which designate structural functions (Tagg 1992).

[9] See, for example, Middleton (1990), Davies (1994, with many references), and Shepherd and Wicke (1997). Scruton (1997, pp. 203–8) offers a trenchant critique of the parallel between music and language implicit in Cooke's approach.

emotional quality inherent in the music?) Related to this is an objection which applies to all lexical approaches to musical meaning: the emotional or other signification of musemes may vary according to the musical contexts within which they are articulated (in other words, it may not be possible to draw so confident a distinction between the 'semantic' and the 'syntactical' aspects of music).[10] Lexical approaches, in short, may endow the music with a false semblance of semantic fixity.

Other criticisms apply to one method but not the other. For example, one problem with Tagg's IOC technique is that it is so laborious, if it is to be done properly, as to verge on the impractical. As Allan Moore (1993, p. 158) puts it, 'it is so time-consuming that I have never seen it properly undertaken in equivalent studies' (other than by Tagg, that is to say). Again, Tagg has been criticized for drawing comparisons between contemporary popular styles and the European 'art' repertory, so implicitly assuming that they form part of a common cultural tradition (Middleton 1990, p. 234). This may be a reasonable assumption in the case of some film and television music: several commentators have traced the appropriation of semantic elements from late nineteenth- and early twentieth-century 'art' music by the Hollywood film industry (Eisler [and Adorno] 1947; Flinn 1992). Analyses of popular music also sometimes reveal the influence of 'art' traditions. An example is Walser's linkage of heavy metal music with the ideologies and even some of the performance practices of nineteenth-century Romanticism (Walser 1993). However, it would be clearly wrong to claim that traditions such as blues, rock, rap, or dance music derive primarily from 'art' music. So it becomes necessary to define just what the relevant frame of reference is for any particular repertory or piece, and that of course raises questions of who is doing the defining and to what ends. Nor is that the only problem. If you set out to identify the different stylistic traditions within which different musemes operate, how are you going to decide where to stop? (Are there specific musemes applicable to popular music? Or to pop but not rock? Or just to dance music? house? garage? jungle? Bristol jungle?)

Such problems may not be insoluble in principle, but they are certainly hard to solve in practice. In such a context, there is something very tempting about Cooke's attempt to derive the emotional expression of music directly from natural principles. As we have seen, he explains expression in terms of intervallic tensions, and he explains intervallic tensions in terms of basic acoustic principles, specifically the overtone series. 'That the major third should be found to express pleasure should surprise no one', he writes, since it appears 'early on in the harmonic series: it is nature's own basic harmony, and by using it we feel ourselves to be at one with nature' (Cooke 1959, p. 51). He admits that tradition builds on tradition, that is, that there is a historical element in musical meaning, but he insists that traditions arise in the first place because of the natural principles built into the music (Cooke 1959, pp. 40–1); for him, then, there was emotional mean-

[10] Cf. Jenefer Robinson's criticism (also voiced, for example, by Anthony Newcomb) that 'most philosophical theorists of musical expression have either ignored or underemphasized . . . the fact that the musical expression of complex emotions is not a function of a few isolated measures here and there, as in Kivy's examples in *The corded shell*; rather it is very often a function of the large-scale formal structures of the piece as a whole. We cannot understand the expression of complex emotions in music apart from the continuous development of the music itself' (Robinson 1998, p. 19).

ing in music before texts were associated with it. And, of course, principles like the overtone series do not vary from one culture to another. It is in this sense that, although limiting his examples to Western 'art' traditions, Cooke implicitly sets out a universal model of emotion in music: the laws of nature apply in China and in Africa as they do in Europe and America, and therefore the same principles should govern musical expression. Such thinking has become deeply unfashionable, because of its potentially essentializing and ethnocentric nature, and few present-day scholars would probably be prepared to sign up to this aspect of Cooke's theory. (Approaches to emotion and meaning in the ethnomusicological literature, while falling outside the scope of this chapter, are almost invariably rooted in the 'thick' description of specific cultures and contexts.[11]) Perhaps the most telling symptom of Cooke's ethnocentricity, however, lies in the way in which he unquestioningly identifies meaning with the expression of personal emotion—in which he understands it, in short, in terms of bourgeois subjectivity. A strong argument could be made that Cooke's approach, valid as it may be for Schubert or Schumann, is misleading in relation not just to Chinese or African music, but to Monteverdi's and Handel's too.

Despite the difficulties to which we have referred, Cooke's work was important in placing emotion firmly within the musicological agenda, and in attempting to do so on a principled basis. Much the same might be said of Leonard B. Meyer (1956), whose first book, *Emotion and meaning in music*, appeared 3 years before Cooke's (indeed, Kivy has described it, perhaps rather sweepingly, as 'the book that taught many of us for the first time that you can talk about music without talking nonsense' (Kivy 1987, p. 153)). Whereas Cooke attempted to ground the emotional expression of music in natural-scientific principles, thereby as it were short-circuiting cultural mediation, Meyer drew upon a range of sources in psychology. His overall interpretation of musical structure, which he has continued to develop in the ensuing 40 years (cf. Meyer, this volume), is based on gestalt principles: listeners are drawn to perceive musical patterns as wholes, as tending towards closure, and compositional techniques can be understood as ways of both stimulating and challenging such perceptions. In essence, Meyer sees music as setting up expectations in the listener (or, to put it another way, implications in the music) which are in general fulfilled or realized, but often only after postponement or apparent diversion. It is here that Meyer's approach to emotion comes in. The basic principle, which Meyer also drew from contemporary psychology, is that 'Emotion or affect is aroused when a tendency to respond is arrested or inhibited' (Meyer 1956, p. 14). To take a very simple example, a dominant seventh chord implies resolution; the longer resolution is postponed the more affect will be created, and if when it comes the resolution is to *vi* (an interrupted cadence), further postponing the return to *I*, there will be a further heightening of affect. Because the implications created by any given piece of music, and the manner in which they are or are not realized, fall clearly within the domain of structural analysis, the result is that emotion is incorporated within the latter. In effect, the analysis of musical emotion plugs into that of musical structure.

[11] For representative examples—one chapter-length, one book-length—see Tolbert (1994) and Feld (1982). See also Becker (this volume).

For all its neatness, however, this approach creates a number of difficulties. One concerns the relationship between emotions as normally understood—hope, joy, grief, and so on—and the kind of undifferentiated affect or feeling tone (in effect a unidimensional variable) that Meyer's theory predicates. Some commentators have doubted whether Meyer's theory can properly be considered as pertaining to the emotions at all (e.g. Budd 1985). Davies (1994, p. 288), by contrast, observes that 'for Meyer the possibility of the listener's differentiating the feeling tone is a real one', and he elaborates and justifies Meyer's position by arguing that the feeling tone can itself take on a kind of mimetic function:

> There are, within any culture, groups of emotions having a common pattern of behavioral expression. For instance, the dynamics of the behavioral expressions of sadness, grief, disappointment, and regret are similar and can be distinguished from the behavioral expressions of joy, happiness, and enthusiasm. Music can be heard as imitating the dynamics of behavior. As listeners we hear these dynamics and imagine that the music is through them expressive of, say, sadness. In so doing we (imaginatively) make the music the emotional object in terms of which our feeling tone can be differentiated.

In saying this, Davies is invoking the Hanslickian idea that feelings are characterized by specific dynamic qualities. As we saw above, Hanslick argues that music cannot specify love, but through its motional qualities it can specify such adjectival properties as the joyful or sorrowful quality of love. In effect, Davies is suggesting that Meyer's undifferentiated affect conveys such adjectival properties, which we are led to attribute to the music itself (that is why it is common, if arguably misleading, to talk of the music being happy or sad, as if it were a sentient being). Since Hanslick's day the idea that music expresses emotion through somehow mirroring the dynamics of our inner life has become a commonplace in writing on the topic; the classic formulation is Langer's claim that 'there are certain aspects of the so-called "inner life"—physical or mental—which have formal properties similar to those of music—patterns of motion and rest, of tension and release, of agreement and disagreement, preparation, fulfilment, excitation, sudden change, etc.' (Langer 1942, p. 228).[12] When we said that music 'somehow' mirrors the dynamic of inner life, however, we meant to signal a crucial vagueness in such thinking: without entering into details, the problem lies in specifying what sort of entity the 'dynamics of our inner life' might actually be, or how it might be measured. In the absence of adequate answers to these questions, the appeal to such impalpable dynamics has little explanatory force.

A second line of criticism is that Meyer's approach is too narrow and exclusive. Davies, for instance, complains that it 'overemphasizes the status of the unpredictable at the expense of the structural significance of similarity' (Davies 1994, p. 289), while Keil (1994) similarly complains that it ignores the significance of repetition and performance skills. Underlying these complaints is a suspicion that, through incorporating them within a psychologically based theory, Meyer is lending a spurious

[12] Davies (1994, pp. 123–4) offers a critique, along with an extensive though by no means complete listing of similar claims by other writers (p. 230, n. 33). The basic approach, of course, links with earlier ideas of music 'moving' listeners' affections.

generality to principles grounded in the modernist aesthetic, with its valorization of innovation and the unique. This charge would be not altogether fair, because the link he makes between psychological principles and musical meaning is not a simple or direct one. For one thing, Meyer recognizes the extent to which listeners' expectations reflect not only the properties of the musical stimulus but also the culturally specific stylistic codes that are acquired through enculturation. For another, he took pains to explain on the very first page of *Emotion and meaning in music* (1956) that the kind of musical meaning we have so far discussed is not the only one: he distinguishes 'embodied' meaning, which is the result of the realization or frustration of expectation, from 'designative' meaning, which involves reference to objects, concepts, or events outside music (a distinction corresponding to Jakobson's 'introversive' and 'extroversive semiosis'; see Jakobson 1971, pp. 704–5), emphasizing that the latter reflects not just culture-specific associations but the listener's personal experiences and disposition, too. Nevertheless, embodied meanings play a far larger role in Meyer's work than designative ones, and it is not hard to see why, for his theory *explains* musical meaning (rather than simply acknowledging its existence) only to the extent that it can be reformulated in terms of relationships between musical elements. Or to put it another way, explaining meaning means translating it into music-structural terms; it is a one-way process. There is something symbolic in the fact that Meyer's first book is the only one he wrote specifically about musical emotion and meaning, the others all focusing on issues of structure and style.

## 3.4 Structural and expressive vocabularies

The basic problem we have diagnosed in Meyer is that of talking about both the emotion and the music in a sustained manner: you may start off talking about the emotion, but you seem to end up talking about the music in more or less the usual way. Of the admittedly limited amount of musicological or music-theoretical work in this area since 1956, much is open to the same criticism. So it is of interest to consider a more recent critique of such work by a professional philosopher (Kivy 1993). Kivy takes as his starting point a critical reading of Schumann's Second Symphony by a writer loosely associated with the early stages of the 'new' musicology (Newcomb 1984). Newcomb's essay is a good example of the revival of a hermeneutical approach to which we previously referred, and indeed sets out to reconstruct a nineteenth-century perception of Schumann's symphony (which was more highly rated in its own time than is generally the case today). Newcomb attempts to achieve this through consideration of the expressive trajectory of the piece, which he characterizes as 'suffering leading to healing or redemption', and relates these emotional characteristics to the difficult circumstances of Schumann's life as he sketched the work. He also points out that many other compositions from around the same period exhibit a similar trajectory (an example to which we have already referred being the 'transmutation of suffering into joy' in Beethoven's Ninth Symphony), and suggests that we might think of this as a 'plot archetype', in other words a generic category in terms of which individual instances may be understood.

Kivy (1993) applauds Newcomb's aims (indeed he sees in them the seeds of a 'new music criticism'), but identifies what he sees as serious methodological failings. In

particular, Kivy objects that Newcomb is trying to add an additional, interpretative layer to structural analysis, whereas the whole project of interpreting autonomous instrumental music is built on a fallacy. 'Pure' music cannot represent anything, Kivy claims, and therefore the symphony cannot contain the meaning Newcomb ascribes to it; in other words, as soon as Newcomb started talking about Schumann's biography, he stopped talking about the music. (As a musicologist rather than a philosopher, Newcomb might very well reply that nineteenth-century listeners clearly *did* interpret the music in just this way, and that the question of whether they were philosophically justified in doing so is irrelevant to an exercise in historical reconstruction.) But in that case, you might ask, how *are* we to build emotional expression into music criticism? Kivy (1993, p. 316) replies:

> The new way is not to amplify criticism by adding interpretation to analysis but, rather, by amplifying analysis itself. For once one ceases to see expressive properties of music as semantic or representational properties, it becomes clear that they are simply *musical* properties: they are phenomenological properties of music, and as such a proper subject of musical analysis.

In other words, he is saying, when we analyse music we should analyse its expressive properties alongside its structural ones. And he offers some examples of what he has in mind. In Mozart's 'Dissonance' Quartet, he says, the relationship between the dark, anguished character of the introduction (which gives the quartet its soubriquet) and the light, sunny character of the allegro that follows creates an effect of strong, expressive resolution; take away the expression and you weaken the resolution. Or again, there is an expressive difference between a Haydn symphony movement that goes from tonic minor to relative major to tonic minor (that is, from dark to light to dark), and one that ends in the tonic major (so that its overall trajectory is from dark to light). Because these are directly experienced aspects of the music, Kivy concludes, 'an analysis . . . that leaves out the expressive contrast, is incomplete in a musically non-trivial way' (1993, p. 319). But does this solve the problem of how to talk about both the music and the emotion in a sustained manner? The answer turns on whether Kivy has identified expressive properties as distinct from structural ones, or whether he is simply talking about structural properties but giving them expressive names. In the Mozart case, is the effect of resolution created by the passage from dark to light actually different from that created by the passage from the dissonance of the opening to the diatonicism of the main movement? Or are these just different ways of talking about the same thing? (If so, then in what sense are you leaving out the expression in speaking only of dissonance and diatonicism?) Again, in the Haydn case, just what does it mean to speak of the passage from dark to light over and above what it means to speak of the passage from minor to major? If the conclusion is that the one terminology can be simply substituted for the other, then there must be a suspicion at least that Kivy's 'new music criticism' is just the old music criticism with an updated (some music theorists might say outdated) vocabulary.

Of course, it might just be that Kivy applies his expressive terminology in too literal a manner. That, at any rate, is suggested by the comparison with another study, this time a collaboration between a philosopher and musicologist, which interprets Shostakovitch's Tenth Symphony as 'a progression from dark to light or struggle to victory' (Karl &

Robinson 1997, p. 166). In using these terms, the authors are not referring to anything as simple as a major or minor mode, or dissonance versus consonance. They are referring to the composite properties of a complex musical structure, and, indeed, their main purpose is to show how the large-scale musical unfolding of the third movement is central to the construction of its expressive message, which they see as one of hope (and more than that, false hope). While their argument begins with structural analysis and ends with expressive analysis—the same sequential approach that Kivy criticized in Newcomb—it would not be fair to describe Karl and Robinson as simply building expression onto structure, like icing on a cake. The very characterization of complex, large-scale passages as 'light' or 'dark' stimulates the attempt to explain these expressive qualities in terms of the music's structural properties; again, the authors set out what they call the 'cognitive content' of hope (which involves looking forward to, and striving towards, a happier but uncertain future) and then match the structural properties of the music against these criteria, in effect using them as an analytical framework. In this way their interpretative strategy involves what they describe as 'a complex interplay among a variety of different sorts of observations' (1997, p. 170), both formal and expressive. There is, however, a certain ambiguity at this point. On the one hand, the authors write that 'we consider this dichotomy [between the formal and the expressive] ill-conceived, because often the formal and expressive threads of a work's structure are so finely interwoven as to be inextricable' (1997, p. 176). On the other, they claim that 'the formal and expressive elements of musical structure are so thoroughly interdependent that the formal function of particular passages can often be accurately described only in *expressive* terms' (1997, p. 177). The second statement seems to start by saying the same as the first but to end up saying the opposite: if you can describe a particular passage in expressive but not structural terms, or for that matter vice versa, then this seems to reinstate the dichotomy between them (an implication which becomes still clearer when the authors continue 'there is no 'strictly formal' or purely musical explanation for why our focal passage unfolds as it does in the central section of the third movement').[13] The relationship between structure and expression remains unclear.

This issue has been explicitly addressed by another musicologist, Fred Maus. In an article entitled 'Music as drama' (Maus 1988), he launches a systematic attack on the sequential, structure-to-expression approach, and then proceeds to offer an analysis of the opening of Beethoven's String Quartet, Op. 95, combining analytical and expressive

[13] Karl and Robinson make a related point in a footnote on p. 173, where they explain that 'a listener schooled in the epic symphonic tradition might expect that the tension of the third movement should be swept away at the beginning of the finale ... Upon finding this expectation frustrated, he or she may then feel tense, impatient, or bewildered ... The listener's feelings of impatience or frustration, while poorly mirroring the expressive structure of the work, may nonetheless provide the initial clues that something is amiss ... providing an impetus to interpretation'. In this way the frustration of expectation which Meyer saw as the source of affect functions as an interpretative cue: implicit in this is a distinction between the listener's emotional response and the expressive properties of the music *per se*. (This distinction bulks large in philosophical writing on the subject, particularly Kivy's, but falls outside the scope of this chapter.)

vocabulary and commenting self-consciously on his analytical strategies as he does so.[14] The conclusion that Maus draws from this analysis—as much an analysis of analysis as of Beethoven—is that music has neither structural nor emotional content. Rather it consists of a series of *events*, which we make sense of by regarding them as the *actions* of imaginary agents, so attributing intentions or motivations to them. His Beethoven analysis, Maus claims, '*explains* events by regarding them as *actions* and suggesting *motivation*, *reasons* why those actions are performed, and the reasons consist of combinations of psychological states' (1988, p. 67). It is this imagined dramatic content that both structural and expressive vocabularies seek to represent, each in their own way: as Maus (1988, p. 69) explains,

> the technical language and the dramatic language offer descriptions of *the same events*. That event at the opening of the quartet, according to the analysis, is an outburst and it is also a unison passage with an obscure relation between metrical hierarchy and pitch hierarchy. The technical vocabulary of the analysis describes the actions that make up the piece.

While referring to the same musical content, then, the two vocabularies construct it in different ways, filter it differently, and the analyst can exploit the asymmetries between them in order to gain purchase on the music. It is for this reason, presumably, that Maus observes that 'Both sorts of description—"technical" or music-theoretical, "dramatic" or anthropomorphically evocative—belong, interacting, to the analysis' (1988, p. 63), and Robert Hatten uses the same word when he writes of 'the interaction of expressively significant . . . thematic, motivic, or topical events with structurally significant voice-leading events' (Hatten 1994, p. 320, n. 8). Hatten's book is, to date, the most comprehensive attempt to incorporate expressive and structural analysis within a coherent critical practice, and it brings together many of the approaches we have described so far. As his reference to topics suggests, Hatten builds on the documentary evidence that historical listeners (his book focuses on Beethoven) not only identified topical references in the music they heard but also saw them as one of the main ways in which it conveyed emotional or other meaning; to this extent Hatten's work, like Agawu's, links with the tradition of understanding emotion as a public, rationalized state, which as we have seen goes back to the baroque affections. Unlike Agawu's work, however, Hatten's also embraces nineteenth-century conceptions of emotion as a private, subjective experience, and indeed the book is easily read as a contribution to the hermeneutic tradition (like Newcomb, Hatten claims to be reconstructing historical practices of listening, and his concept of 'expressive genre' corresponds closely to Newcomb's 'plot archetype'). More centrally, it brings together extroversive and introversive approaches within a unified theoretical framework drawn from present-day semiotic theory. The question on which we wish to focus, however, is the extent to which Hatten succeeds in theorizing the interaction between structure and expression to which he refers.

The immediate impression of Hatten's working procedure is that, once again, he starts with structural analysis, and then adds in the interpretation. The first ten pages of his analysis of the first movement of the Quartet, Op. 130, for instance, consist of structural

[14] For detailed discussion of the relationship between structural and expressive vocabularies in critical writings which draw analogies between music and narrative, see Maus (1997).

analysis, while the final page offers an expressive interpretation in terms of a quasi-Hegelian synthesis of opposed emotions (Hatten 1994, pp. 134–45). And while he does sometimes begin with a provisional and intuitive identification of the music's expressive content, following this with structural analysis and then a refinement of the expressive interpretation based on that analysis, it is significant that Hatten never goes on to the next stage (that is, returns to and refines the structural analysis on the basis of the expressive one). To this extent the relationship between structure and expression is still an unequal one: as usual in musical analysis, the cards are stacked in favour of structure. In other contexts, however, such as his account of the *Cavatina* from the same quartet, Hatten's treatment is more even-handed. He makes use of vocabulary that applies equally well to structure or expression (undercutting, irruption, or reversal, for example). He mingles structural and expressive vocabulary (1994, p. 213):

> A strong wedge motion in m. 5 expands registrally to the apex on the downbeat of m. 6, which is negated as a climax by the more intensely expressive crux created by the unexpected reversal (registral collapse) on the second beat. The 'willed' (basically stepwise) ascent takes on a hopeful character supported by the stepwise bass'.

At other times, he offers parallel structural and expressive interpretations; on p. 214 he writes of bar 12 that 'As the vi region of E♭ major, C minor would appear to be an appropriate choice in light of the harmonic emphasis it has already received in m. 6. The region could also be motivated by the expressive value of the relative minor as an emotionally troubled opposition to the serenity of the major'.

But most revealing is the way in which he stresses what, in 'purely musical' terms, might be seen as points of abnormality or incoherence, and siezes upon them as interpretative opportunities. For instance, he says of the repeated $iv^7$ chords on the second beats of bars 15 and 16, following anomalously on cadential six-four chords, that they support (1994, p. 215)

> a construal as 'parenthetical' or outside the normal course of time, thus akin to 'insight' in that it comes without logical necessity or sequential train of thought . . . Interactively interpreting these inflections of meaning, one might understand the resulting trope as 'tragically weighted insight or reflection, expanding an instant that seems frozen within the flow of time'.

Expanding the same principle to the scale of the movement's formal unfolding, he provides a figure showing 'unconvincing formal analyses' (1994, p. 209, Fig. 8.1): the very incoherence of the movement when understood in purely formal terms becomes a source of its overall expressive meaning, which Hatten sees as a kind of metaphorical blend of tragic insight and faith, each illuminating the other. In this way, expression is not seen as simply a gloss on structure; it works in counterpoint with it.

Writers such as Karl and Robinson, Maus, and Hatten demonstrate that it is possible to use structural and expressive vocabularies together in a way that does not simply map the one onto the other, in the manner of Kivy's 'new music criticism'. It might be going too far to say that any of them succeeds in talking in a sustained manner about both the music and the emotion (even Hatten sometimes gives the impression of lurching from one to the other). But they manage to do so in flashes, and the result is illuminating.

## 3.5 Conclusion: discourse and emotion in music

In the work of some writers, the distance between technical and 'anthropomorphically evocative' interpretation, to borrow Maus's phrase, becomes much greater. A good illustration is provided by a pair of curiously similar essays: by Charles Fisk on the first movement of Schubert's Sonata in B♭ major, D. 960 (Fisk 1997), and by Marion Guck on—among other things—the second movement of Mozart's G minor Symphony, K. 550 (Guck 1994). Each begins with a single, portentous moment in the piece:[15] the left hand trill in the seventh bar of Schubert's sonata, and the C♭ in bar 53 of the symphony movement. And each constructs out of this moment a similar story: that of an alien element, an 'outsider' who after a succession of vicissitudes becomes accepted within the larger context represented by the piece—is accepted, in other words, as an 'insider'. The link that each author makes between music on the one hand, and on the other the narrative of rejection and acceptance, alterity and assimilation, builds on the expressive properties of the music and at the same time brings new expressive properties, new connotations of positive or negative emotion, to bear upon it. Neither author, of course, is claiming that the original music was in any sense 'about' immigration. Like Newcomb in the case of Schumann's Second Symphony, and Karl and Robinson in the case of Shostakovitch's Tenth, Fisk invokes the circumstances of Schubert's life to add credibility to, or at least to deepen, his interpretation: maybe, he says, the (again arguably) homosexual Schubert felt himself to be an outsider in his own society and so 'took solace, at least unconsciously' from his own parable of integration (Fisk 1997, p. 200). Whereas Karl and Robinson give the impression of speaking from a position of interpretative authority, however ('Shostakovich, ever mistrustful of happy endings, undercuts its optimistic qualities' (Karl & Robinson 1997, p. 178)), Fisk claims to be doing no more than telling a story, offering 'a naively poetic description of what happens in the music' (Fisk 1997, p. 195), and the personal and provisional nature of his interpretation is highlighted in the title: 'What Schubert's last sonata might hold' (further glossed, on the final page of the article, as 'what Schubert's last sonata holds for me'). Guck similarly introduces her account as 'a story about C♭' (Guck 1994, p. 67).

Then what, you might ask, is the point of all this? Fisk tells his story in order to substantiate an argument that every listener projects his or her own emotional experiences into the music, which in turn moulds those experiences so that they become in some sense purely musical emotions (Fisk 1997, p. 182). The limitation of his approach, however, is that he begins with technical analysis, goes on to unfold his narrative (showing how it can be fitted to the framework of the analysis)—and then stops. Guck, by contrast, structures her article through a process of oscillation between technical and expressive interpretation. She begins with her perception of the C♭'s portentousness, an observation which (as she puts it) 'incorporates a conceptual structure that is covert' (Guck 1994, p. 71). The purpose of constructing the parallel narrative, she continues, is that 'the conceptual structure of the immigrant's tale is revealed in the process of its telling'; in other words, the story renders the structure explicit, and so initiates the

[15] Both authors use this word 'portent'/'portentous': Fisk (1997, p. 184) and Guck (1994, p. 63).

process of answering the analytical question '*why* is the C♭ portentous?'. But to complete the answer it is necessary to return to the text, to the plane of technical analysis, understanding and experiencing the music in light of the narrative: as Guck puts it, her story 'suggests a strategy for hearing not only the highlighted events but also the lines of development in which they participate, which is to say that it provides a means of codifying and enriching the hearing of the whole piece' (1994, p. 71). The point of the analysis, in short, is not just to describe how you experience the music, but to *change* how you experience it, both in technical terms and emotionally.

It is not only Fisk's and Guck's analyses which do this. We mentioned above that Hatten presents his work as, in essence, an exercise in the reconstruction of historical listening practices. However, this is not the only nor necessarily the most productive way of reading it, and indeed Hatten does not embark upon the systematic kind of historical reconstruction of period perception, based on contemporary documentation, that has been attempted for instance by Johnson (1995). The alternative is to read Hatten's work as a sustained argument about how we might most fruitfully hear Beethoven's music today. Understood this way, Hatten's book becomes an attack on the etiolated listening practices brought about by what he refers to as the 'errors of formalism' (Hatten 1994, p. 228); by contrast, he is offering a richer, more complex, more human way to hear the music, and one which (as a bonus) perhaps recreates something of the manner in which Beethoven's first listeners heard it. Another way to express this is that it is not simply a matter of hearing emotions out of the music—that is, hearing the meanings that were always there within it—but, so to speak, of hearing them *into* it. Understood this way, analytical and critical practices are less concerned with how the music *is* than how it might be heard to be; they aim less at proof than at persuasion (as Guck puts it, 'Truth is replaced by the plausibility of the narrative'; 1994, p. 72). This is the basis of what we term a 'performative' approach to music and expressive meaning.[16]

There is a sense in which writing about music has always been performative, and one of the most common ways of misinterpreting the documents of music history is by reading them as descriptions of actual practice rather than interventions in it. This is obvious in the case of the treatises of the baroque and classical periods, which prescribed how their readers were to compose, perform, or understand music (and in some cases the music they were to compose, perform, or understand, too).[17] Equally, musicological knowledge can shape perceptions: Spitzer (1987) has shown how divergent accounts by different critics of the same piece of music (*Sinfonia Concertante* in E♭,

[16] The concept of the performative derives from speech act theory (Austin 1962) and is particularly influential in performance and gender studies (see, respectively, Schechner 1988 and Butler 1990). Central to it is the idea that meaning is constructed through performative acts: whereas the meaning of 'the sky is blue' is referential (it refers to an external reality), when you say 'I promise' you are not referring to something but actually *doing* it by virtue of what you say. The term therefore has no direct link with performance in the musical sense (which is not to say that performative approaches do not have much to offer the study of musical performance).

[17] For an example of the role of performance pedagogy in the construction of musical experience, see Blasius (1996). Blasius argues that, along with other discursive domains, early nineteenth-century piano method is a realization of the eighteenth-century project to systematize affect and rhetoric.

attributed to Mozart) are not simply due to differences in taste, but reflect debate over the work's authenticity. (Not only do critics tend to write more positively when they think the music is by Mozart, Spitzer found, but they tend to focus on the same passages and use the same metaphors in describing them.) And one might equally claim that, to misquote Judith Butler, all the approaches to musical emotion outlined in the previous sections have helped to bring about the very responses and practices which they purport to describe.[18] Or to use the language of social constructionism, discourse about music (by which we mean principally, but certainly not exclusively, verbal discourse) actively constructs rather than simply reconstructs its expressivity. The very act of saying that a certain piece of music is expressive in a certain sort of way leads you to hear it that way, and so one can think of dominant interpretations of particular compositions (Tovey's interpretation of the first movement recapitulation of the Ninth Symphony as the embodiment of cosmic catastrophe, for instance, or for that matter McClary's interpretation of the same passage in terms of a rapist's fantasy) as exercises in the creation and not merely the reporting of expressive meaning.

There was a strong awareness of the performative value of criticism among certain writers associated with the 'new' musicology (in particular, Kramer and Abbate[19]). More generally, and reflecting parallel trends in cultural studies, the ethnomusicologist Philip Bohlman (1993) argued in his significantly named article 'Musicology as a political act' that the discipline as a whole is as much involved in the generation as the analysis of the values and meanings that are ascribed to music. Attempts to develop explicit theoretical models for the attribution of meaning to music, however, are a more recent development and arise out of a promising intersection of music theory, psychology, and cognitive science. Zbikowski (1999), for instance, has drawn on conceptual blending theory (itself representing the intersection of linguistics and literary theory) in order to analyse three separate settings of the same poem, Müller's *Trockne Blumen*. Without entering into details, the basic idea is that there are aspects of shared structure (what Zbikowski calls a 'common topography') between the words and the music, which enables us to map the one onto the other as we listen to the song.[20] The result is a blending of the attributes of each, which gives rise to emergent meaning—meaning, that is, which is not present in either words or music separately but arises out of their interac-

[18] '[Gender] identity is performatively constituted by the very 'expressions' that are said to be its results' (Butler 1990, p. 25).

[19] See, in particular, Kramer (1990) and Abbate (1991). More recently, Monson (1996) has applied the same principle to the analysis of jazz performance, arguing that 'emotion is constituted or constructed through social practices . . .. Music, it seems to me, is a particularly powerful constructor of emotion . . . a powerful activity that can produce a 'community of sentiment' binding performers and audiences into something larger than the individual' (p. 178).

[20] The analysis involves the creation of a graphic diagram or 'conceptual integration framework' (CIN) which shows four interrelated 'spaces'. These are the 'text space' and 'music space', which show the corresponding features on which the mapping is based; the 'generic space', which indicates the dimension within which it takes place; and the 'blended space', where attributes of the text and music spaces are combined to create new meaning. For further explanation of CINs, which derive from the work of Mark Turner and Gilles Fauconnier, see Zbikowski (1999, pp. 310–14).

tion. In this way the emotional associations of the words are transferred to the music: as Zbikowski says, 'We can easily imagine outward physical expressions of the miller's torment—facial contortions, hand wringing, pacing—mirroring his internal conflict' (1999, p. 309). The point is that these become as much properties of the music as of the words,[21] so that once again there is a sense in which we are talking about emotional content even as we analyse musical structure. Up to now such approaches have been principally applied to the analysis of multimedia texts (Cook 1998b builds a general theory of multimedia on a similar basis), but there is no reason why this should be the case: you can equally well use them to compare Tovey's and McClary's interpretations of the first movement recapitulation of the Ninth Symphony, demonstrating the manner in which each interpretation creates new meaning (Cook, forthcoming). Used this way, the conceptual blending approach reveals how meaning is constructed through the interaction between musical texts and the invisible web of discourse, ranging from books and CD liners to television and concert-interval chatter, that surrounds them.

Coming from a different direction but converging with the conceptual blending approach is another, drawn this time from psychology—specifically, from James Gibson's (1979) 'ecological' approach to perception (which musicologists and music theorists are only now beginning to explore). Traditional approaches to music perception see it as built up in succession from basic elements such as tones, scale patterns, chords, and so on, with expressive or social meaning representing, to repeat our earlier phrase, the icing on the cake. (This corresponds to the sequential approach we have criticized: first analysing the structure, and then interpreting the meaning.) By contrast, the ecological approach seeks to understand perception in terms of its function, that is to say the means by which any organism grasps and interacts with its environment, and from this perspective the semantic properties of any stimulus, including its emotional properties, are just as basic—just as much properties of the music—as any acoustic ones. Emotional properties, in short, are not deduced through interpretation of the music, but directly specified by it. But this does not represent a reversion to earlier ideas that meaning is 'locked into the sounds', to borrow Moore's phrase (1993, p. 157), just waiting to be unlocked by listeners, because meaning emerges from a mutual relationship between perceiver and perceived in which any number of personal, historical, or critical influences come into play. It is this mutual relationship that becomes the object of analysis. As most conspicuously represented by Clarke's (1999) study of subject position in music by Frank Zappa and P. J. Harvey, the aim is very much the same as we have described in the work of Zbikowski: to find a way of talking about music which articulates the emotional properties that keep us listening to it, while at the same time conforming to the requirements for intersubjective intelligibility in the absence of which writing about music becomes a critical free-for-all where anything goes.

New as they may be, at least to musicology, such approaches address an old problem: how to speak about music and emotional meaning at the same time, without changing the subject. In addition, they emphasize the intrinsic reflexivity of critical activity—

[21] This links with Fisk's claim, mentioned in the previous section, that listeners' emotions are so moulded by music as to become purely musical emotions.

what we say about music's meaning contributes to bringing that meaning into being. If the further development of musicological approaches to emotion depends on continued cross-fertilization with other disciplines, then (to misquote Butler again) this book may itself contribute to the very thinking we have attempted to describe.

## References

**Abbate, C.** (1991). *Unsung voices: Opera and musical narrative in the nineteenth century.* Princeton, NJ: Princeton University Press.

**Adorno, T. W.** (1973). *Philosophy of new music.* London: Sheed & Ward.

**Agawu, V. K.** (1991). *Playing with signs.* Princeton, NJ: Princeton University Press.

**Allanbrook, W. J.** (1983). *Rhythmic gesture in Mozart.* Chicago: University of Chicago Press.

**Austin, J. L.** (1962). *How to do things with words.* Cambridge, MA: Harvard University Press.

**Bach, C. P. E.** (1974). *Essay on the true art of playing keyboard instruments* (trans. W. J. Mitchell). London: Eulenberg Books (originally published 1959–1962).

**Blasius, L. D.** (1996). The mechanics of sensation and the construction of the Romantic musical experience. In *Music theory in the age of Romanticism*, (ed. I. Bent), pp. 3–24. Cambridge, UK: Cambridge University Press.

**Bloomfield, T.** (1993). Resisting songs: Negative dialectics in pop. *Popular Music*, 12, 13–31.

**Bolhman, P.** (1993). Musicology as a political act. *Journal of Musicology*, 11, 411–36.

**Budd, M.** (1985). *Music and the emotions: The philosophical theories.* London: Routledge.

**Burnham, S.** (1999). How music matters: Poetic content revisited. In *Rethinking music* (ed. N. Cook & M. Everist), pp. 193–216. Oxford, UK: Oxford University Press.

**Butler, J. P.** (1990). *Gender trouble: Feminism and the subversion of identity.* New York: Routledge.

**Cavicchi, D.** (1999). *Tramps like us: Music and meaning among Springsteen fans.* New York: Oxford University Press.

**Clarke, E. F.** (1999). Subject-position and the specification of invariants in music by Frank Zappa and P. J. Harvey. *Music Analysis*, 18, 347–74.

**Cone, E. T.** (1982). Schubert's promissary note. *19th–Century Music*, 5, 233–41.

**Cook, N.** (1998a). *Music: A very short introduction.* Oxford, UK: Oxford University Press.

**Cook, N.** (1998b). *Analysing musical multimedia.* Oxford, UK: Clarendon Press.

**Cook, N.** (forthcoming). Theorizing musical meaning. *Music Theory Spectrum*, 23.

**Cooke, D.** (1959). *The language of music.* London: Oxford University Press.

**Davies, S.** (1994). *Musical meaning and expression.* Ithaca, NY: Cornell University Press.

**Dibben, N.** (2001). Pulp, pornography and voyeurism. *Journal of the Royal Musical Association*, 126, 83–106.

**Eisler, H. [and Adorno, T. W.]** (1947). *Composing for the films.* London: Dennis Dobson.

**Feld, S.** (1982). *Sound and sentiment: Birds, weeping, poetics, and song in Kaluli expression.* Philadelphia, PA: University of Pennsylvania Press.

**Fisk, C.** (1997). What Schubert's last sonata might hold. In *Music and meaning*, (ed. J. Robinson), pp. 179–200. Ithaca, NY: Cornell University Press.

Flinn, C. (1992). *Strains of utopia: Gender, nostalgia, and Hollywood film music.* Princeton, NJ: Princeton University Press.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton-Mifflin.

Gjerdingen, R. O. (1988). *A classic turn of phrase.* Philadelphia, PA: University of Pennsylvania Press.

Goehr, L. (1992). *The imaginary museum of musical works: An essay in the philosophy of music.* Oxford, UK: Clarendon Press.

Guck, M. A. (1994). Rehabilitating the incorrigible. In *Theory, analysis and meaning in music,* (ed. A. Pople), pp. 57–73. Cambridge, UK: Cambridge University Press.

Hanslick, E. (1986). *On the musically beautiful: A contribution towards the revision of the aesthetics of music* (trans. G. Payzant). Indianapolis, IN: Hackett (originally published 1854).

Hatten, R. (1994). *Musical meaning in Beethoven: Markedness, correlation, and interpretation.* Bloomington, IN: Indiana University Press.

Jakobson, R. (1971). *Language in relation to other communication systems. Selected writings,* (Vol. 2). The Hague: Mouton.

Johnson, J. H. (1995). *Listening in Paris: A cultural history.* Berkeley, CA: University of California Press.

Karl, G. & Robinson, J. (1997). Shostakovich's Tenth Symphony and the musical expression of cognitively complex emotions. In *Music and meaning* (ed. J. Robinson), pp. 154–78. Ithaca, NY: Cornell University Press.

Keil, C. (1994). Motion and feeling through music. In *Music grooves* (ed. C. Keil & S. Feld), pp. 53–76. Chicago: University of Chicago Press.

Kerman, J. (1985). *Musicology.* London: Fontana Press. (= *Contemplating music: Challenges to musicology.* Cambridge, MA: Harvard University Press).

Kirkpatrick, R. (1984). *Interpreting Bach's 'Well-Tempered Clavier': A performer's discourse of method.* New Haven, CT: Yale University Press.

Kivy, P. (1987). How music moves. In *What is music? An introduction to the philosophy of music* (ed. P. Alperson), pp. 149–62. University Park, PA: Pennsylvania State University Press.

Kivy, P. (1993). *The fine art of repetition: Essays in the philosophy of music.* Cambridge, UK: Cambridge University Press.

Kramer, L. (1990). *Music as cultural practice, 1800–1900.* Berkeley, CA: University of California Press.

Langer, S. K. (1942). *Philosophy in a new key.* Cambridge, MA: Harvard University Press.

Lerdahl, F. & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.

McClary, S. (1991). *Feminine endings: Music, gender and sexuality.* Minnesota, MN: University of Minnesota Press.

Marks, M. M. (1992). *Music for the silent film: Contexts and case studies 1895–1924.* Oxford, UK: Oxford University Press.

Mattheson, J. (1981). *Der vollkommene Capellmeister* (trans. E. C. Harris). Ann Arbor, MI: UMI Research Press (original work published 1739).

Maus, F. E. (1988). Music as drama. *Music Theory Spectrum,* **10**, 56–73.

Maus, F. E. (1997). Narrative, drama and emotion in instrumental music. *Journal of Aesthetics and Art Criticism,* 55, 293–302.

Meyer, L. B. (1956). *Emotion and meaning in music.* Chicago: University of Chicago Press.

**Middleton, R.** (1990). *Studying popular music.* Milton Keynes, UK: Open University Press.

**Monson, I.** (1996). *Saying something: Jazz improvisation and interaction.* Chicago: University of Chicago Press.

**Moore, A. F.** (1993). *Rock: The primary text.* Buckingham, UK: Open University Press.

**Newcomb, A.** (1984). Once more 'between absolute and program music': Schumann's Second Symphony. *19th-Century Music*, 7, 233–50.

**Pater, W.** (1910). *The Renaissance: Studies in art and poetry.* London: Macmillan (Original work published in 1873).

**Ratner, L. G.** (1980). *Classic music: Expression, form and style.* New York: Schirmer.

**Robinson, J.** (1998). The expression and arousal of emotion in music. In *Musical worlds: New directions in the philosophy of music* (ed. P. Alperson), pp. 13–22. University Park, PA: Pennsylvania State University Press.

**Schechner, R.** (1988). *Performance theory* (rev. edn). New York: Routledge.

**Schenker, H.** (1994). The art of improvisation (trans. R. Kramer). In *The masterwork in music: A yearbook. Volume 1 (1925)*, (ed. W. Drabkin), pp. 2–19. Cambridge, UK: Cambridge University Press (originally published 1925).

**Scruton, R.** (1997). *The aesthetics of music.* Oxford, UK: Clarendon Press.

**Shepherd, J. & Wicke, P.** (1997). *Music and cultural theory.* Cambridge, UK: Polity Press.

**Solomon, N.** (1989). Signs of the times: A look at late 18th-century gesturing. *Early Music*, 17, 551–61.

**Spitzer, J.** (1987). Musical attribution and critical judgement: The rise and fall of the Sinfonia Concertante for Winds, K. 279b. *Journal of Musicology*, 5, 319–56.

**Tagg, P.** (1982). Analysing popular music: Theory, method and practice. *Popular Music*, 2, 37–67.

**Tagg, P.** (1992). Towards a sign typology of music. In *Secondo Convegno Europeo di Analisi Musicale*, (ed. R. Dalmonte & M. Baroni), pp. 369–78. Trento, Italy: Università degli studi di Trento.

**Tolbert, E.** (1994). The voice of lament: Female vocality and performative efficacy in the Finnish-Karelian *itkuvirsi*. In *Embodied voices: Representing female vocality in Western culture*, (ed. L. C. Dunn & N. A. Jones), pp. 179–94. Cambridge, UK: Cambridge University Press.

**Treitler, L.** (1999). The historiography of music: Issues of past and present. In *Rethinking music*, (ed. N. Cook & M. Everist), pp. 356–77. Oxford, UK: Oxford University Press.

**Wagner, R.** (1899) *Prose works*, (Vol. 8) (trans. W. A. Ellis). London: Reeves.

**Wallace, R.** (1986). *Beethoven's critics: Aesthetic dilemmas and resolutions during the composer's lifetime.* Cambridge, UK: Cambridge University Press.

**Walser, R.** (1993). *Running with the Devil: Power, gender and madness in heavy metal music.* Hanover, NH: Weslyan University Press.

**Zbikowski, L. M.** (1999). The blossoms of 'Trockne Blumen': Music and text in the early nineteenth century. *Music Analysis*, 18, 307–45.

CHAPTER 4

# PSYCHOLOGICAL PERSPECTIVES ON MUSIC AND EMOTION

JOHN A. SLOBODA AND PATRIK N. JUSLIN

## 4.1 The psychological approach

In the context of a multidisciplinary work, it may be appropriate for us to briefly set out what we consider to be the essentials of a psychological approach, so as to show its distinct orientation, alongside the other disciplines represented here. Yet, we are also mindful of the apocryphal tale of a young colleague of William James who was asked by the master to take an introductory course in psychology while James was away, and thought he would begin by defining psychology, and was still struggling to do that when James returned 12 weeks later!

So as not to delay our treatment of emotion for too long, let alone its intersection with music, we shall confine our remarks on psychology to a set of rather dogmatic statements. We believe that in broad outline they describe what most psychologists do, although there is much dispute over the details, with differing schools and periods placing differing emphasis on the various subcomponents (Leahey 1987). More importantly for the present chapter, perhaps, these statements reflect our own perceptions of what it is to do psychology, of both music and emotion, and form some sort of underlying framework for the material we present here.

1. Psychology is concerned with the explanation of human behaviour. Behaviour includes overt action as well as 'inner' behaviour, such as thought, emotion, and other reportable mental states. It can include behaviour of which the agent is not fully or even partly aware, such as the dilation of the pupils of the eye.

A psychological approach to music and emotion therefore seeks an explanation for how and why we experience emotional reactions to music, and how and why we experience music as expressive of emotion.

2. The explanations sought by psychologists are essentially causal and organismic. They are causal in the sense that they seek both to discover antecedents to the behaviours requiring explanation and uncover or postulate mechanisms whereby the antecedents interact to bring about the observed behaviours. They are organismic in that the primary interest of psychology is in mechanisms which are internal to the organism. External causes, such as historical or social factors, are only relevant to a

psychological framework where, even if only in sketched in principle, there is a route by which these may impact on the internal mechanisms.

In the study of music and emotion, a central aim of psychology is to understand the mechanisms that intervene between music reaching a person's ears and an emotion being experienced or detected by that person as a result of hearing that music. Other important aims involve understanding the roles of emotion in composing and performing music.

3. There are different levels at which organismic explanations can operate. These include the physiological and the phenomenological. A physiological explanation attempts to describe the causal pathway in terms of mechanisms within the nervous system, the endocrine system, and other biological constellations. A phenomenological explanation, on the other hand, attempts to describe the causal pathway in terms of the intentions, meanings, and interpretations that the individual concerned places on his or her own behaviours. There is, however, an explanatory level intermediate between these extremes, which is the core concern of psychology. A term commonly used for this level of explanation is 'cognitive'. Another term is 'computational'. The concern of a cognitive explanation is more in the functional logic of the mechanism (what it does—the 'software') than the precise neural location and instantiation (how it does it—the 'hardware'). However, this is not simply a redescription of the phenomenological level, because some of the most significant advances made by the psychological approach relate to aspects of behaviour which are pre-conscious and therefore could not—even in principle—be incorporated at the phenomenological level.

We shall try to be explicit about the level of explanation that is being used at various points in the psychological arguments that we describe and critique.

4. Psychology is a materialist science. Psychology is materialist because it holds some version of reductionism (or mind-brain identity). Every event describable at a phenomenological level (e.g. an intention) is, at the same time, assumed to be describable as a set of physical events (e.g. a particular constellation of biological events). Psychology is a science because it only allows explanations to survive in the long term when they have withstood repeated, critical interactions with observed data. The organization and interpretation of such interactions are collectively known as 'the scientific method' (Neale & Liebert 1986).

Emotional responses to music are material phenomena whose components are correlated with observable phenomena in law-governed ways. We aim in this chapter to demonstrate the relationships between scientific observations and the development of theory.

5. Whilst psychology operates within a materialist framework, it is—at least tactically—not reductionist in its programme. Many psychologists believe that seeking links between biological events and behavioural or phenomenological events does not always lead to scientific progress. For instance, showing a lawful correspondence between a set of behavioural and biological events (which is rarely easy to do) does not in itself yield an explanation of how the biology and the behaviour

are causally linked. Psychological explanations have sufficient explanatory and predictive power that few psychologists currently operate on the ambition to render psychological explanations obsolete in favour of biological ones. Indeed, psychological explanations allow certain kinds of understanding that biological explanations cannot easily support (see next point).

In this chapter, we shall centre our account at the psychological level, leaving detailed treatment of other levels to other authors.

6. Psychological explanations are particularly successful at incorporating teleological or functional perspectives. Understanding a behaviour requires more than knowing how it was caused. It also requires knowing something about the ends that the behaviour serves. What, in other words, does the behaviour and the psychological mechanisms that support it allow the organism to achieve? Functional explanations may also be offered at different levels. At the most distal, functions can be considered in terms of the survival of the individual, or even of the species. However, more proximally, functions may be considered in terms of individual goals and plans, which may be very short term and malleable.

An adequate psychological theory of emotional responses to music should, therefore, explain how such emotions reflect the ways that emotions serve to fulfil psychological goals and outcomes. It is not enough to simply demonstrate law-like relationships between musical antecedents and emotional consequents.

Before considering the enduring issues of music and emotion from a psychological perspective, we briefly review emotion concepts and major approaches to emotion. This serves to anchor our subsequent discussion of music and emotion.

## 4.2 What is an emotion?

Emotion is one of the most pervasive aspects of human existence, related to practically every aspect of human behaviour—action, perception, memory, learning, and decision making. It is thus all the more remarkable that the study of emotion has been neglected throughout much of psychology's brief history. Although the scientific study of emotions dates back to the nineteenth century (Darwin 1872; James 1884; Wundt 1897), studies of emotion have often been eclipsed by studies concerned with the 'higher' forms of mental processes, such as reasoning, problem solving, and decision making. As noted by Lazarus (1991, p. 4), 'Failure to give emotion a central role puts theoretical and research psychology out of step with human preoccupations since the beginning of recorded time.' Almost every great philosopher has been concerned with the nature of emotions. Writers, artists, and musicians have always attempted to appeal to the emotions, to influence and move the audience through emotional communication.

What is an emotion? This question, which incidentally was the title of one of William James's (1884) most famous papers, has not yet received a definitive answer. There are many reasons for this state of affairs. One reason is that emotions are difficult to define and measure. As pointed out by Fehr and Russell (1984, p. 464), 'everyone knows what an emotion is, until asked to give a definition'. Another reason may be the 'disruptive' role ascribed to emotions as motivators of human behaviour. Although our

view of the emotions has changed considerably throughout history, one of the prevailing views has been that rationality can be hijacked by the pirates of emotion (Cacioppo & Gardner 1999). Only recently have researchers seriously reconsidered the role played by emotions in human rationality (Damasio 1994; De Souza 1987; Johnson-Laird & Oatley 1992; Simon 1967). This has reached the point where scientists today readily speak of 'emotional intelligence' (e.g. Mayer & Salovey 1993), a notion that would have been inconceivable during the heyday of behaviourism!

Emotion is both an everyday concept and a scientific construct. As such it involves both an implicit and an explicit body of knowledge (e.g. Plutchik 1980). The implicit knowledge is embodied in the so-called 'folk theories' of emotion. We all know something about emotions. We think that they are powerful forces that affect our behaviour and thoughts in powerful ways, that some emotions feel good and some bad, and that some people are more 'emotional' than others. The explicit knowledge is based on studies of problems such as how emotions develop, whether emotions are associated with physiological changes, and how well we judge emotions from facial or vocal expressions. Emotion as a scientific construct is inferred from three kinds of evidence: (a) self-reports; (b) expressive behaviour; and (c) physiological measurement.

The most common—and deceptively simple—way to measure emotional responses in human adults is by means of *self-report*: adjective checklists, rating scales, questionnaires, or free descriptions. This approach is associated with numerous problems, such as the imperfect relationship between emotions and words that denote emotions, and the problem of choosing which words to include in checklists or scales. At the same time, self-reports are probably the most direct form of evidence, and it is conceivable that certain aspects of emotional life (e.g. 'peak' experiences, see Gabrielsson, this volume) cannot be reached in any other way.

Because the use of self-reports is not always possible—or reliable, some would say—another approach has been to measure different forms of *expressive behaviour*, or products of expressive behaviour (such as drawings; see, for instance, Smith & Williams 1999). This may include facial expressions, vocalizations, and body language (e.g. gestures). The most famous example is perhaps Ekman's (1973) cross-cultural studies of facial expression of emotion. One problem with this form of evidence is that emotions are not always accompanied by expressive behaviours. Furthermore, expressive behaviours often occur in the absence of emotions, because people use such behaviours intentionally to communicate information to other individuals.

The third kind of evidence that is used to infer emotions involves various *physiological measures* of emotion. This kind partly reflects the early work of William James, who claimed that emotion was basically a perception of internal bodily changes. Since then, researchers of emotion have used a number of physiological indices to measure emotion, including heart rate, respiration, skin conductance, muscle tension, electrocadiagram (EKG), blood pressure, and electroencephalograph (EEG) (for a review, see Cacioppo *et al.* 1993). There are many problems with physiological measures, the most serious of which is probably that autonomic changes often occur even in the absence of emotions (primarily because each physiological system has many different functions within the body). Therefore, it is difficult to establish clear-cut relationships between emotional states and physiological responses.

Based on these three kinds of evidence, theorists have proposed a number of different definitions of emotion during the past 100 years. For example, Kleinginna and Kleinginna (1981) identified 92 definitions found in textbooks, articles, dictionaries, and other sources. Based upon a review of these definitions they proposed the following consensual definition (Kleinginna & Kleinginna 1981, p. 355):

> Emotion is a complex set of interactions among subjective and objective factors, mediated by neural/hormonal systems, which can (a) give rise to affective experiences such as feelings of arousal, pleasure/displeasure; (b) generate cognitive processes such as perceptually relevant effects, appraisals, labeling processes; (c) activate widespread physiological adjustments to the arousing conditions; and (d) lead to behavior that is often, but not always, expressive, goal-directed, and adaptive.

Many researchers like to think of emotion as a sequence of events: emotions are triggered by our cognitive appraisals of significant events; these appraisals evoke strong reactions of most of our bodily systems; the bodily reactions generate subjective experiences of feeling; and the subjective experiences yield action tendencies and expressive behaviour. However, researchers often disagree about the precise sequence of events, and about where such an episode begins or ends (Scherer 1993).

As even a superficial reading of the literature on emotion would show, scientists have used a number of different terms to denote the phenomena of interest in this chapter—affect, emotion, mood, feeling, arousal, etc. We think that much of the confusion and controversy in the field of music and emotion has been caused by the lack of clear definitions of these terms. Whereas most researchers of music appear to have used the terms affect, emotion, and mood interchangeably—lay people certainly do that—there is a growing consensus among emotion researchers that the three terms should be differentiated from one another. Affect is seen as a more general term than emotion or mood (Oatley & Jenkins 1996, p. 124). Affect refers to the positive or negative valence of the emotional experience. Many researchers consider affective valence as the most basic feature of emotional life, and believe that affect is phylogenetically and ontogenetically more primitive than emotion or mood (cf. Batson *et al.* 1992). Emotions are often said to differ from moods in three ways. First, emotions are said to be brief, whereas moods last much longer. Second, many researchers think that emotions usually have an identifiable stimulus event, whereas moods do not. Third, some researchers have proposed that emotions are accompanied by distinct facial expressions, whereas moods are not (Ekman & Davidson 1994, pp. 51–96).

Davidson (1994), however, suggests that it is more useful to differentiate emotion and mood in terms of their functions. He proposes that 'emotions bias action, whereas moods bias cognition' (p. 54). Emotions arise most often when adaptive action is needed. Moods, on the other hand, tend to shift the 'mode' of information processing in ways that influence memory, decision making, and evaluative judgments (see Bless & Forgas 2000). In this way moods can alter the probability that particular emotions will be triggered. Thus, emotions can be viewed as *phasic* perturbations that are superimposed on the *tonic*, affective background provided by the mood (Davidson 1994).

*Arousal* is often considered to be an important aspect of emotional reactions, and this term has figured prominently in many theories of emotion (e.g. Mandler 1984), and

also in many accounts of music and emotion (Berlyne 1971; Gaver & Mandler 1987; Meyer 1956; see also Simonton, this volume). Sometimes, the term is used more generally to refer to the level of emotional intensity of the response, other times it is used to refer specifically to the sympathic activation of the nervous system. The relationship between the two meanings of the term is not clear. For example, there is evidence that people can feel intense emotion, without showing autonomic arousal (Harrer & Harrer 1977).

A major element in both folk theories and scientific theories of emotion is that emotions are subjective feelings that people experience—the word 'emotion' comes from a Latin word which means 'move' or 'to stir up' (Plutchik 1994). But how many emotions are there, and how do different emotions relate to each other?

## 4.3 Major approaches to emotion

The ways that emotions are experienced and expressed help both scientists and lay people to conceptualize emotions and differentiate among them. In this section, we describe the most prominent approaches to conceptualizing emotion: the categorical approach, the dimensional approach, and the prototype approach. We also consider an additional class of affects—the vitality affects—which we believe could be useful in explaining listeners' responses to music.

### 4.3.1 The categorical approach

According the categorical approach, people experience emotions as categories that are distinct from each other. Essential to this approach is the concept of *basic emotions*; that is, the idea that there is a limited number of innate and universal emotion categories from which all other emotional states can be derived (see Ekman 1992; Izard 1977; Oatley 1992; Plutchik 1994; Power & Dalgleish 1997; Tomkins 1962). Each 'basic' emotion can be defined functionally in terms of a key appraisal of goal-relevant events that have occurred frequently during evolution (Table 4.1). Thus, basic emotions are regarded as adaptive in dealing with life emergencies without resorting to time-consuming processing: 'When no fully rational solution is available for a problem of action, a basic

Table 4.1 Key appraisals for basic emotions adapted from Oatley[1] (1992) and Lazarus[2] (1991)

| Emotion | Juncture of plan[1] | Core relational theme[2] |
|---|---|---|
| Happiness | Subgoals being achieved | Making reasonable progress towards a goal |
| Anger | Active plan frustrated | A demeaning offense against me and mine |
| Sadness | Failure of major plan or loss of active goal | Having experienced an irrevocable loss |
| Fear | Self preservation goal threatened or goal conflict | Facing an immediate, concrete, or overwhelming physical danger |
| Disgust | Gustatory goal violated | Taking in or being close to an indigestible object or idea (metaphorically speaking) |

emotion functions to prompt us in a direction that is better than a random choice' (Johnson-Laird & Oatley 1992, p. 201). Basic emotions solve problems with speed rather than precision. They may be viewed as 'fast and frugal algorithms' that cope with conditions of limited time, knowledge, or cognitive capacities (Gigerenzer & Goldstein 1996).

Some general criteria for distinguishing basic emotions from 'secondary' or 'complex' emotions have been suggested. The most important of these are that basic emotions: (a) have distinct functions that contribute to individual survival; (b) are found in all cultures; (c) are experienced as unique feeling states; (d) appear early in development; (e) are associated with distinct patterns of physiological changes; (f) can be inferred in other primates; and (g) have distinct emotional expressions. A considerable body of evidence supports the case for a concept of basic emotions as defined by these criteria (see Ekman 1992; Izard 1977; Oatley 1992; Panksepp 1992; Power & Dalgleish 1997). 'Secondary' emotions are usually believed to be founded on basic emotions, but involve 'blends' of basic emotions (Plutchik 1994), or particular cognitive evaluations that occur together with a basic emotion (Oatley 1992).

The notion of basic emotions has been criticized on a number of grounds, most notably because different researchers have come up with different sets of basic emotions (Ortony & Turner 1990). However, as noted by Oatley (1992) 'this issue is not about voting, it is about theory' (p. 104). It all depends on how we *define* emotions. Therefore, if one takes a closer look at the researchers who consider emotions in terms of their functional significance, one finds that there is reasonable agreement among researchers with regard to at least five basic emotions: happiness, sadness, anger, fear, and disgust (Kemper 1987; Plutchik 1994; Power & Dalgleish 1997). For an example of how the basic emotions approach has been applied in the case of music, see Juslin (this volume).

### 4.3.2 The dimensional approach

The categorical approach focuses mainly on the characteristics that distinguish emotions from one another. In contrast, the dimensional approach focuses on identifying emotions based on their placement on a small number of dimensions, such as *valence*, *activity*, and *potency*. The search for the dimensional structure of emotions dates back to Spencer (1890; cited in Izard 1977), but more influential formulations of this same notion were provided by Wundt (1897), Woodworth (1938), and Schlosberg (1941). A dimensional structure can be derived from any type of response data, although the most common sources are similarity judgments of emotion words or facial expressions that are analysed using factor analysis or multidimensional scaling.

Russell's (1980) *circumplex model* has generated the most research (Fig. 4.1). This model consists of a two-dimensional, circular structure involving the dimensions of *activation* and *valence*. Within this structure, emotions that are across the circle from one another, such as sadness and happiness, correlate inversely. The circumplex model captures two important aspects of emotions: that they vary in their degree of similarity and that certain emotions (e.g. happy, sad) are often thought of as bipolar. About the same circular structure has been found in a large number of different domains—including music (Thayer 1986)—suggesting that the circumplex model really captures something fundamental about emotional responses.

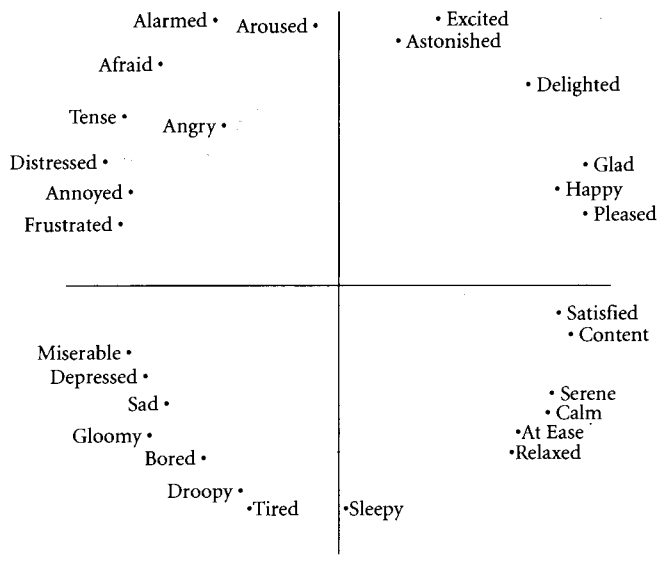



**Figure 4.1** A circumplex model of emotion (from Russell 1980). (Copyright ©1980 by the American Psychological Association. Reprinted with permission.)

Among the strengths of the circumplex model is that it suggests a simple yet powerful way of organizing different emotions in terms of their affect appraisals (pleasant or unpleasant) and physiological reactions (high or low arousal). On the other hand, dimensional approaches have been criticized on the grounds that they blur important psychological distinctions and therefore obscure important aspects of the emotion process (cf. Lazarus 1991). In Russell's circumplex model, for example, emotions that are placed in the same position in the circular matrix may, in fact, be very different. Thus, anger and fear are two emotions that are highly correlated within the model, because they are both high in activity and unpleasantness. Yet, they are very different in terms of their implications for the organism (see Table 4.1).

Notably, many researchers regard categorical and dimensional approaches to emotion as complementary to each other (Nyklicek *et al.* 1997) and both receive some support from neurophysiological data (Damasio 1994). Dimensional models may be especially useful for capturing the continuous changes in emotional expression that occur during a piece of music. For examples of response methods based on the dimensional approach to emotion, see Schubert (this volume).

### 4.3.3 The prototype approach

The prototype approach to emotions is based on the work by Rosch (1978). Essential to this approach is the idea that language, and knowledge structures associated with language, shape how people conceptualize and categorize information. A basic principle is that membership in a particular category is determined by resemblance to prototypical exemplars. The prototype itself is an abstract image that consists of a set of weighted features that represent the exemplar of a family of instances. Thus, for example, robins

are presumably more prototypical of birds than ostriches. Similarly, some emotions (e.g. elation) are probably better exemplars of joy than others (e.g. relief), depending on related features such as elicitors, cognitive appraisals, and physiological reactions.

The prototype approach provides an interesting compromise between categorical and dimensional approaches to emotion—it addresses both the contents of individual categories and the hierarchical relationships among categories. An example of a prototype structure is shown in Fig. 4.2. The vertical dimension of the structure shows the hierarchical relationships among categories. The most general level is the *superordinate* level, which is defined by the positive or negative valence of the emotions within a particular category. The middle level represents the basic-level categories, or *prototypes*, of emotions, which anchor our mental representations of all emotions within a given category. The *subordinate* level consists of all other emotions related to a particular prototype. Finally, the horizontal dimension shows the relations among members of the same prototype category (Shaver *et al.* 1987).

Critics of the prototype approach have argued that people's accounts of emotions are insufficient for capturing the underlying structure of emotions. For instance, the prototype structure presented in Fig. 4.2 includes surprise among the prototypes, despite the fact that many researchers would not consider surprise an emotion (cf. Oatley 1992; Plutchik 1994). Critics also complain that there is disagreement about which emotions qualify as basic-level prototypes, and that emotions may cross boundaries in the prototype structures. However, supporters of the prototype approach accept that boundaries between emotion categories are sometimes 'fuzzy', and argue that emotions cannot be defined in terms of a set of necessary and sufficient conditions.

### 4.3.4 The vitality affects

Apart from the more traditional approaches to emotion, there is another class of qualities that, although their status as emotions may be called into question, seems to be particularly relevant in the case of music. Stern (1985) introduced the concept of *vitality affects* to describe a set of elusive qualities related to intensity, shape, contour, and movement. These characteristics are best described in dynamic terms such as crescendo, fleeting, explosive, diminuendo, etc. These qualities are not emotions, but rather abstract 'forms' of feeling that occur both together with, and in the absence of, proper emotions. The vitality affects are 'amodal' in the sense that they are common to all modes of expression. Stern (1985) suggests that the vitality affects are of a particular importance in the early communicative acts of mother and infant. Mother and infant respond to one another by constantly adapting and adjusting the intensity, timing, and contour of their expressive acts. This process of constant matching of gestural events is referred to as *attunement.* Lack of capacity to enter into such interpersonal synchrony is often regarded as symptomatic of mental illness (Bunt & Pavlicevic, this volume).

Although the notion of vitality affects is admittedly vague, it seems to capture something important about music's expressiveness. Often when researchers speak about affect in music, they seem to be thinking about something akin to the vitality affects rather than true emotions. Vitality affects, with their dynamic changes and patterns of events, are somewhat reminiscent of philosopher Susanne Langer's (1951)

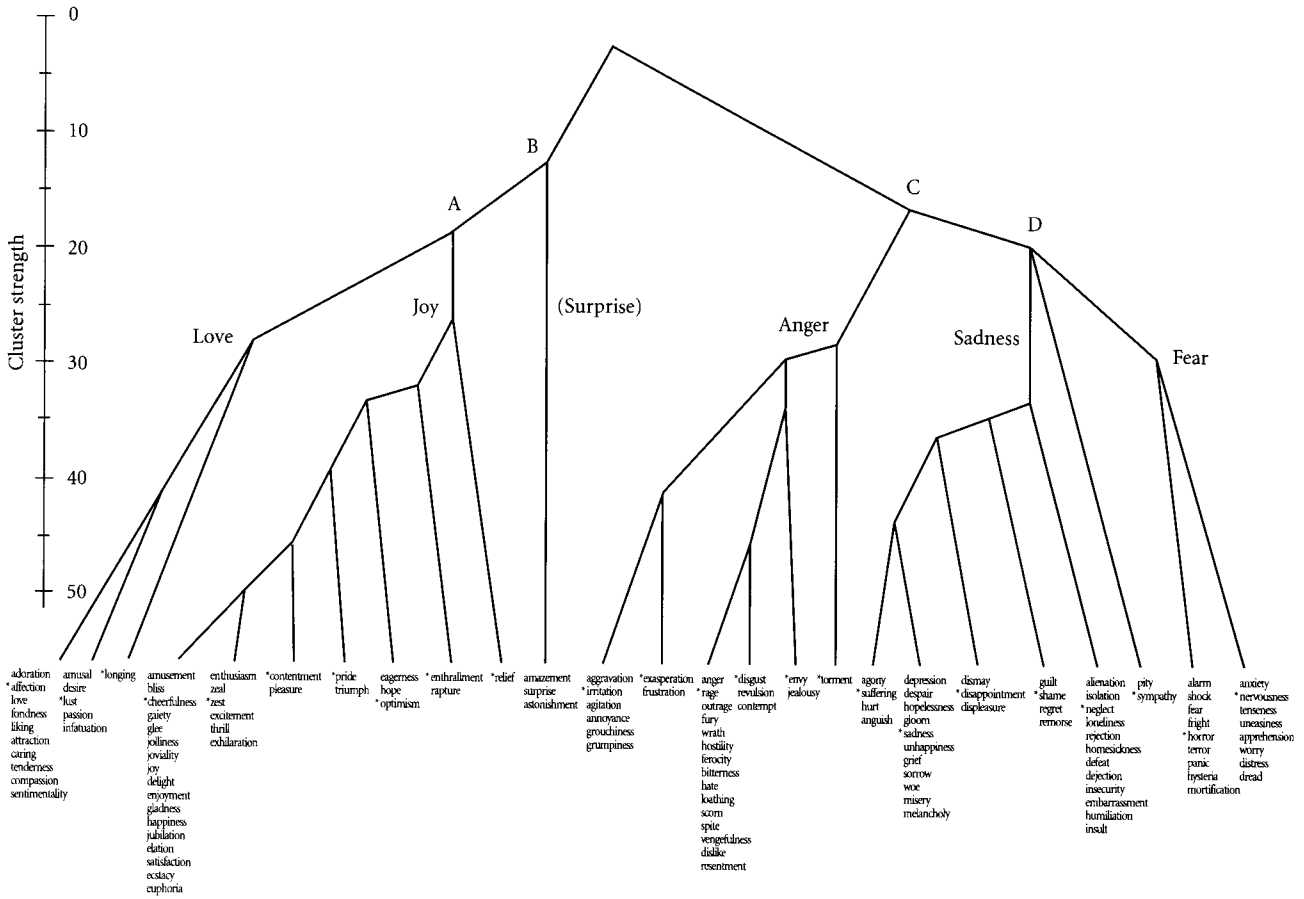



Figure 4.2 A prototype analysis of emotion words (from Shaver et al. 1987). (Copyright ©1987 by the American Psychological Association. Reprinted with permission.)

notion of music as expressive of the forms of feeling. Music, suggests Langer, represents the *dynamic form* of emotional life, not specific emotions. When Langer observes that 'some musical forms seem to bear a sad and happy interpretation equally well' (Langer 1951, p. 202), this might be because vitality affects are not emotions themselves, and because some vitality affects are common to many different emotions.

Of all art forms, no one may be better equipped to capture vitality affects than music, in particular classical music. For an example of how the notion of vitality affects can be applied to a phenomenological description of music experience, see Imberty (1997). Imberty argues that musical style is ultimately 'an architecture of vitality affects', and goes on to discuss the role of vitality affects in the protonarrative character of musical time. The notion of vitality affects is also important in accounts of emotional processes in music therapy, in which dynamic form in improvisation is used to communicate a statement of the self (Bunt & Pavlicevic, this volume).

Having briefly reviewed the major approaches to emotion, we are ready to look more closely at the relationship between music and emotion.

## 4.4 The study of emotion in relation to music

Although emotions have been studied scientifically for a century, treatises on emotion rarely consider emotional reactions to music (for a few notable exceptions, see Fridja 1989; Lazarus 1991; Oatley 1992; Zajonc 1994). This may reflect the assumption that emotions experienced while listening to music are in some sense different from other emotions. To be sure, there are important differences between musical emotions and other emotions, both in antecedents and consequences, although this in itself does not imply that the emotions themselves are different. Another reason for the neglect of music in the emotion literature could be that such emotions are somehow considered less important—perhaps because scientists have failed to realize that music fulfils vital functions in many people's daily lives (Sloboda & O'Neill, this volume).

Aesthetic emotions—emotional reactions to films, drama, music, or paintings—have never been in the 'mainstream' of psychology (Lazarus 1991). Although psychologists have acknowledged that art may evoke strong emotional reactions despite its illusory nature, few scientists seem to have considered this problem worthy of study. Even in contemporary music psychology, the study of emotion has occupied a peripheral existence (cf. Gabrielsson 1993). However, we sense an increased interest in music and emotion at the moment, and today it is almost fashionable among music scientists to move into the emotion field. We believe that this interest may reflect an increasing awareness that music psychology has made limited progress regarding the problem that should be closest to its heart—how people experience music.

Musical emotions are, at a first glance, of two kinds: first, there are those emotions that concern the aesthetical value of music. Psychologists have, on the whole, embarrassingly little to say about the aesthetics of music, whereas philosophers have been discussing this topic for centuries (Budd 1985). It has been suggested that an aesthetic response is an intense personal experience that provides insight into the nature of life, and that involves emotional, cognitive, and social components (Konečni 1979). In the field of 'new' empirical aesthetics, however, researchers have focused on more mundane

relations between liking, or preference, and such informational properties of the stimulus as its complexity, expectedness, or familiarity (e.g. Berlyne 1971; North & Hargreaves 1997; for a review of musical preferences, see Russell 1997).

Second, there are those emotions that are induced or expressed by the music, more or less apart from the aesthetical value of the music. These are the emotions with which we will mainly concern ourselves in this chapter. However, the two classes of emotion are not wholly independent of each other, and we believe that an understanding of psychological aspects of music and emotion requires that researchers make the connection between the two explicit.

The study of emotion in relation to music is problematic in many ways. First, emotional reactions are commonly understood in terms of their adaptive functions related to biological survival. Since emotional reactions to music are not as directly related to biological survival, they seem to be difficult to explain within existing theories of emotion (Dowling & Harwood 1986). Second, scientists who wish to study emotional reactions to music are faced with the problem that there is great variability between individuals, and across time within individuals (Sloboda 1996). Third, experiments that attempt to measure listeners' affective responses to music may impact so much on the listening process that the task destroys the very thing it is supposed to measure—the problem of *reactivity* (Neale & Liebert 1986).

On the other hand, some authors have suggested that music is an excellent medium for studying emotion. For example, Gaver and Mandler (1987) have argued that using musical stimuli to study emotion is ecologically valid, because people are used to making judgments about music and their affective responses to music. They also note that music has a structure that is rich, and yet fairly well known, and that understanding the structure of an event may help in understanding an emotional reaction to it. Yet another advantage of using music as a stimulus is that it avoids many of the ethical concerns associated with other kinds of stimuli.

A review of the literature on music and emotion reveals the atheoretical nature of much of this work. Music psychologists have been strikingly reluctant to turn to emotion psychology for theoretical guidance (Juslin 1997*a*). One reason for this could be that the field of emotion has presented such a confusing picture that many researchers have found it difficult to distil something useful from this research. Another reason may be that researchers have considered current theories of emotion inadequate for explaining emotional responses to music—a view also evident in the work of some philosophers (e.g. Kivy 1990). In our view, however, the problem does not lie in current theories of emotion, but in the fact that most researchers have not made serious attempts to apply the theories to their problems. It is significant that one of the most influential analyses of music and emotion to date, namely that of Meyer (1956), was partly based on a theory of emotion (Dewey 1894).

One of the primary questions in the field of emotion is whether there are some underlying and objectively observable characteristics that are invariably present in the class of phenomena we normally describe as emotions. In the following section, we describe typical characteristics of emotion and show how these apply in the case of music.

## 4.5 Typical characteristics of emotions

### 4.5.1 Emotions are functional despite their apparent non-instrumentality

Fridja (1986) has provided one of the most influential psychological analyses of the behavioural characteristics that lead one to impute emotion to a person. His analysis is particularly valuable because it is not circular—it does not smuggle in the very concepts it is supposed to explain. He says the following (Fridja 1986, p. 2):

> At some moments when observing behaviour, that behaviour seems to come to a stop. Effective interaction with the environment halts and is replaced by a behaviour that is centered, as it were, around the person himself, as in a fit of weeping or laughter, anger or fear. Or interaction with the environment may go on but seems peculiarly ineffective. When someone smashes the dinner plates, the broken plates would hardly seem to be the end result the person had in mind. Other behaviour that invites emotion words seems to contain a surplus that is not needed for the end result: superfluous emphasis in speed and scope of movement, or hesitation and undue toning down, or a smile that, in someone who is stroking a child, does not add to the tenderness of the touch.

To capture this aspect of emotion Frijda uses the notion of 'non-instrumentality', in that the emotional phenomena do not seem to be linked in any direct way towards the achievement of immediate goals and plans. A major purpose of emotion science is therefore to explain what functionality emotions do have. Are they useful to us, and if so how? Most researchers seem to believe they are (Keltner & Gross 1999). There is some consensus that the primary function of emotions is to guide behaviour; emotions evolved because they enabled successful interaction with the environment. But how does the idea of functionality apply in the case of music?

First, the view of emotions as functional adaptations has important implications for how we should understand and explain emotional responses to music (Juslin 1997*a*). Specifically, the functional architecture of emotions should *constrain* our responses to music, such that our responses to music can be explained by the ways that emotions have been shaped by biological evolution. Thus, for example, some of our emotional responses to music might be understood in terms of selective pressures favouring our ability to employ acoustic cues in our environment—particularly the social environment—to make useful inferences about the probable behaviours of other individuals (e.g. Juslin, 2001). Such mechanisms have presumably been favoured by biological evolution because they contribute to individual survival.

A second, and by no means less important, sense in which the functionality of emotions should be considered is the functions that music serve in everyday life, an aspect which is still obscure and poorly understood (Sloboda & O'Neill, this volume). Until we have a clear notion of what a person's involvement in music is for, then we are not going to know whether or not the emotions it evokes are or are not instrumental or functional, and in what way. It has been suggested that music may serve a mood-optimizing function in people's lives (Zillman 1988). However, the mere fact that emotional responses to music are so puzzling—even to the very people who often experience and value such emotions—suggests that musical emotions share in the apparent non-instrumentality of other emotions.

### 4.5.2 Emotions have behavioural, physiological, and experiential components

Although it may seem self-evident that we experience emotions to music, it is really an open question whether the conditions for ascribing emotion in music are the same as in other areas of emotional life. We noted in Section 4.2 that emotion is usually inferred from three kinds of evidence—self-reports, expressive behaviour, and physiological behaviour. It seems pertinent to ask, then, to what extent these kinds of evidence can be obtained also in connection to music.

The most basic observations arguably come from studies that use self-reports. These studies have yielded evidence that people experience emotions while listening to music, and that they perceive music as expressive of emotions. An interesting example is Pike's (1972) phenomenological analysis of music experience. Pike asked a number of musically untrained participants to listen to different pieces of music. After hearing each piece, which lasted no more than 5 minutes, the participants were asked to write down their responses. Through content analysis, the listeners' responses were reduced to a small set of common experiential factors (listed in the order of their frequency of occurrence): (a) feeling of pleasure (96 per cent); (b) perception of stable moods (86 per cent); (c) feeling of oneness with the music (83 per cent); (d) perception of spontaneous and transient emotional states (72 per cent); and (e) feeling of movement (65 per cent). This finding suggests that emotional responses to music—like emotional responses to other stimuli—include stable moods, transient emotions, and feelings of pleasure. Many other studies provide evidence from self-reports that listeners experience emotions in relation to music (e.g. Behne 1997; Gabrielsson 1991; Krumhansl 1997; Sloboda 1991, 1992; Waterman 1996).

We also have some evidence of expressive behaviour in responses to music. For example, Gabrielsson (1991), Waterman (1996), and Sloboda (1991) have shown that people really do cry when listening to some music. Furthermore, studies using facial electromyography (EMG) have suggested that people react with subliminal facial expressions towards expressive music (Witvliet & Vrana 1996). Kenealy (1988) found that a music mood-induction procedure yielded significant effects on behavioural measures, such as decision time, distance approximation, and writing speed (behaviours believed to be affected by moods). Moreover, there are many social contexts in which music provides a rich stimulus for expressive behaviours (e.g. rock concerts).

There is also evidence of physiological reactions to music (Bartlett 1996). Krumhansl (1997) has reported that the constellation of physiological responses occurring while people listen to music generally judged to be sad, fearful, or happy, under laboratory conditions (i.e. still and silent) are similar, although not always identical, to those constellations identified by other researchers as present when people experience these emotions to non-musical events—particularly in situations involving behavioural suppression (cf. Cacioppo *et al.* 1993; Gross & Levinson 1993). Nyklicek *et al.* (1997) have reported evidence of cardiorespiratory differentiation of music-induced emotions (happy, serene, sad, agitated). Finally, a few studies have explored a special kind of intense emotional responses to music called 'thrills' or 'chills', indicating that such responses to music are quite common, and are more prevalent for sad than for happy pieces of music (see Goldstein 1980; Panksepp 1995). For a review of physiological studies, see Scherer and Zentner (this volume).

Regardless of the particular method used, all these studies have provided evidence that emotional responses to music involve what Scherer (1993, p. 3) has termed the classic 'reaction triad' of subjective feeling, expressive behaviour, and physiological reaction.

### 4.5.3 Emotions have proximal elicitors

Frijda (1986, p. 4) again gives an influential expression to this feature: 'Emotions are elicited. The eliciting events appear to fulfill a special role: they are not just stimuli. They appear to act through their significance, their meaning, their rewarding or aversive nature.'

This characteristic of emotions seems to be necessary in order to distinguish them from other classes of affective phenomena. For instance, sadness which is pervasive and prolonged tends to be seen as a mood or a disposition rather than an emotion, according to its duration and the relative difficulty of isolating a specific proximal stimulus. However, the distinctions are necessarily imprecise. Clearly, weeping on first news of the death of a loved one is part of an emotional reaction to that event. A year later, a predominant and pervasive sense of sadness would not easily be characterized as an emotional reaction to the death, but rather a mood (although even at that distance a specific instance of weeping might well be viewed as an emotion if a proximal reminder of the loved one could be identified as a proximal elicitor).

A particularly interesting set of issues arise with respect to the assertion that eliciting stimuli should have significance or meaning to the person experiencing the emotion. Is it necessary that this person should be able to know, consciously, what meanings mediate the emotion? Is it, in other words, plausible to attribute emotion in the case that a stimulus occurs, the experience of sadness follows, the person believes that the stimulus has caused or triggered his or her response, and yet the person experiencing this response has no knowledge of why this stimulus should cause the reaction it does? This is an important issue for the application of emotion to the musical domain, because a great deal of what looks like emotional response to music is not accompanied by any strong sense of where the significance lies. A person can be reduced to tears by a particular passage of music, yet be completely unable to specify, even in outline, any objective feature of the music which would account for its grief-inducing qualities.

Fortunately, music is not unique in this respect. Research has shown that people often perceive and process emotional information in a pre-conscious or automatic manner. LeDoux (1996), for example, has recently described a neurophysiological system consisting of parallel-processing perceptual mechanisms, or modules, that react to simple stimulus features. LeDoux has found a direct neural link from auditory nuclei in the thalamus to the 'fear effector system' in the amygdala, which provides immediate information about simple features of emotionally relevant auditory stimuli. This 'quick and dirty' route of transmission bypasses the thalamocortical pathway, which gives full meaning to the stimuli. The perceptual system outlined by LeDoux (1996) provides some support for Zajonc's (1980) controversial slogan that 'affect precedes inference', and shows how emotions can be elicited outside conscious awareness.

Of particular interest is the fact that people implicitly perceive emotional expressions, which subsequently influence their behaviour (Niedenthal & Showers 1991). This means that implicit perception of the emotional expression of a piece of music may

affect us emotionally although we are not consciously aware of this process. Much has been made by some music psychologists of the resemblance of certain musical sounds to signals that human infants may be biologically 'pre-programmed' to respond emotionally to (see Juslin, this volume).

Another class of implicit stimuli are those which obtain their emotional charge through resemblance to characteristics of events which the person can no longer remember (whether through Freudian repression, or normal processes of forgetting). One of the present authors has an overwhelming fear of being submerged in water, but no memory for any event which could account, even non-rationally, for that fear. The fear survives successful contemporary non-damaging experiences of submersion. In a similar fashion, musical stimuli might arouse strong emotions through conditioned responses, the origin of which we cannot remember.

### 4.5.4 Emotions are intrinsically social

Although emotions can occur when a person is alone, their full manifestation very often seems to require other people. For instance, it is well known that young children (e.g. pre-verbal) will look to a nearby adult for a cue concerning how to respond to some events, such as a fall or graze. If the adult looks upset and concerned, the infant will cry. If the adult appears calm and confident, the infant is less likely to cry. This process is referred to as *social referencing* (Sorce *et al.* 1985), and illustrates that others' emotional expressions are important sources of information which may influence our behaviour.

In another well-known series of studies, Schacter and Singer (1962) placed participants whose arousal levels had been artificially elevated into different social situations, where actor confederates of the experimenter appeared to express particular strong emotions. In general, the participants experienced the emotion displayed by the confederate. All these studies indicate that we are extremely ready to 'catch' ourselves the emotions we observe in others. Emotions are 'contagious' (Hatfield *et al.* 1994).

This tendency may be internalized in the self-regulation of emotions. For instance, a person may ask him or herself 'what would X feel in this situation' and use the stored answer to such a question to influence how he or she responds. In other situations, we identify another person's emotions not so much to align our emotional state with them, but to determine how we should respond to them. If I detect another person as uncontrollably angry, I may well become afraid of what that person might do to me, and withdraw. I might become afraid of what that person might do to someone else, become apprehensive for that person's welfare, and so attempt to help her avoid hurt.

For such reasons, identification of the emotion present in other people is a key priority for us in most situations, and therefore we have very well-developed mechanisms for scanning the environment for any cues that would allow us to both attribute emotion to others and determine whether we should feel that same emotion ourselves. These cues do not necessarily come from direct observation of another person's behaviour. We may deduce something about a person's emotional state from the way that they have affected the environment, for instance, arranged the contents of a room or written a note. People in some cultures have read emotions of unseen beings (such as gods) into such natural phenomena as storms. An angry human may throw small objects around, and so when big objects (such as trees) get thrown around, it is natural to ascribe anger to a

superhuman being. It is not only facial expression, gesture, and vocalizations that suggest emotions to us, but material phenomena which may be the outcome of, or in some sense resemble, those expressions, gestures, and vocalizations.

Music, it could be argued, has many of the requisite qualities that incline us to ascribe emotions to it. There is a great deal of evidence on this point. In general, the experimental technique has been to present a participant with a musical extract and then ask him or her to select, from a smaller or larger list of emotion-related words, those words which most closely match the extract (e.g. Gabrielsson & Juslin, in press). All of these studies confirm that there are some extract-emotion combinations which generate high agreement among listeners, even when they may not have engaged explicitly in such tasks previously. Levinson (1996) has given particularly strong form to this argument through his claim that pieces of music are often construed by listeners as a 'virtual persona'.

Sloboda (2000, p. 227) has suggested that music probably creates an environment in which the distinction between attributing detected emotion to oneself as opposed to an outside source or agent is particularly fluid: 'If I detect some piece of music as containing elements appropriate to anger, then there is a sense in which I am pretty free to ascribe that anger where I will. I may ascribe it to the composer, or to the performer, to myself, to God. None of these ascriptions need to bring me up against any of the immediate constraints or contradictions provided by a real-life event in which my own interests and relationships figure.'

### 4.5.5 Emotions invoke action tendencies

Emotions change the probabilities associated with subsequent behaviours (Fridja *et al.* 1989). For instance, fear renders it more probable that one will avoid (or seek to avoid) the eliciting stimulus. Frijda (1986) uses the term 'action tendency' rather than 'intention' or 'plan' because the emotion does not contain within it a specific outcome, but rather directs the person experiencing it towards one category of behaviours rather than another.

So, for instance, a fearful person may simply close her eyes, so as not to see the fear-inducing stimulus, feeling that it cannot be avoided; she may run away from it (but with no specific end-location in mind); she may move towards a known place of safety; or she may develop plans to destroy or render harmless the fearful stimulus. These actions are distinct from the set of actions which might be primed by an emotion of tenderness towards the same eliciting stimulus (e.g. a mouse). In both cases, however, the precise action that occurs is not specified from within the emotion itself. This will depend on subsequent appraisals, and the intentions, skills, and knowledge of the subject.

From this analysis comes the most generally accepted theory of what emotions are for, that is why we have developed them. Frijda's (1986, pp. 77–8) formulation is, once again, seminal:

> Action tendencies ... originate in discrepancies: in mismatch or in perception of potential match. Action tendencies terminate in unresolvable mismatch or in match. Mismatch, potential match, and match must be assumed to give rise to signals indicating that action is called for, is likely to succeed, or can terminate. We may identify these signals with feelings of pain, pleasure, and desire. These are the signals that initiate, maintain, and terminate action tendency. They constitute the direct source of 'being set' for change or continuation of the given situation. At the same

time, they confer an 'aim' upon action tendency: they provide such tendency with an end point, the functional equivalent of a goal. Action tendency continues until the mismatch signals that initiate and maintain it die down; action tendency remains aligned when potential match signals resound, and terminates where match signals indicate success.

In effect, emotions are biologically embedded mechanisms that ensure that, in general, our psychological energies are directed to the meeting of primary needs, both physical and psychological. They are 'non-negotiable' and force our attention onto the eliciting stimulus until the action tendency is realized. But emotions by themselves do not guarantee effective solutions to life's challenges, particularly in the complex and rapidly changing material and social environments that human beings have constructed for themselves. Such solutions are developed by learning and the establishment of a repertoire of problem-solving strategies.

Understanding the links between one's emotions and effective behaviour has recently been described as an important aspect of 'emotional intelligence' (Mayer & Salovey 1993). The development of effective responses to emotionally challenging situations can be helped through various kinds of 'off-line' activities, such as role playing (observed as spontaneous 'pretend play' in children), and story telling. Stories present the observer with a series of situations which elicit emotions in the protagonist, and by empathy or identification, in the reader or listener. By exploring the interactions between situations, emotions, and actions in a simulated environment, we can develop and test different strategies that might be deployed in real situations, but without the potentially harmful consequences of 'getting it wrong'.

It is in the context of such considerations that emotional responses to many art forms are rendered partly explicable. Novels, plays, operas, and films can provide opportunities to feel emotions with and for the protagonists, and explore consequences of different ways of acting on these same emotions. This is true even in the Western art culture, where the passive and immobile 'respectful silence' of the audience has such paradigmatic status. There is a strange paradox in the typical public performance, where much emotion is experienced (as evidenced by self-report of audience members) but almost none is displayed.

On the other hand, many other social contexts in which music and drama has occurred through the ages has provided opportunities for, and even encouraged, action. Pop concerts, for instance, provide a typical context and stimulus for exuberant and joyful bodily and vocal expressions among audience members. Indeed, the urge to participate seems to be a natural and early occurring phenomenon. Young children seem to spontaneously break into motion when they hear certain types of music (Moog 1976). Even when suppressing an overt response to music, listeners may still be engaging in subliminal physical action (e.g. Fraisse *et al.* 1953; Harrer & Harrer 1977). There is also evidence that soothing music yields more helpful behaviour in subjects than music that is aversive in nature (Fried & Berkowitz 1979). So action tendencies are by no means absent in emotional responses to music.

Perhaps more puzzling is the question of the underlying motivation for the action tendencies. If Frijda's analysis is correct, then music-engendered emotions, just like any other emotions, should involve the awareness of match, potential match, or

mismatch between the current situation and some envisaged situation. Emotion should be initiated, maintained, or terminated by key events or actions which bring about transitions from match to mismatch or vice versa. A central problem for music psychology is to characterize and explain the various ways in which musical events can be interpreted by the human cognitive system as marking these transitions, while not directly 'telling a story'. Not only does music achieve this, but it often achieves it spectacularly well. The emotions felt by many people to music are equal in intensity to the emotions they experience to life events. And in some circumstances, people have reported that these emotions have propelled them towards significant personal change (Bunt & Pavlicevic, this volume; Gabrielsson & Lindström 1993; Sloboda 1992).

### 4.5.6 Emotions change during the course of human development

All researchers who study emotion seem to agree that emotions do change with development (e.g. Ekman & Davidson 1994). However, researchers differ among themselves with regard to their views on what changes occur, and why these changes occur. Although the number of proposed changes in emotions as a function of development is too large to permit a thorough overview, they include maturational changes of the nervous system, changes in the stimulus conditions eliciting emotions, changes in emotion regulation and coping skills, changes in the relationship between cognition and emotion, changes in expressive behaviour as a consequence of mothers' communication styles, and changes associated with cultural influences, including language. Theories of emotional development typically describe how emotions develop from pleasure and displeasure to successively more differentiated emotional states (Harris 1989).

To what extent does emotional development apply to the case of music? Unfortunately, individual differences in emotional responses to music within age cohorts have by and large not been systematically examined. However, there is a stream of developmental research that has concerned the development of perception of emotional expression in music (e.g. Cunningham & Sterling 1988; Dolgin & Adelson 1990; Kastner & Crowder 1990; Kratus 1993; Terwogt & van Grinsven 1988). These studies used pictures of facial expressions and asked children to select the face that matches the music's emotional tone most closely. The results indicate that even children 3 or 4 years of age can recognize certain emotions—joy, sadness, anger, and fear—in music with an accuracy better than chance. This ability seems to increase with age, however, particularly up to the age of six.

Developmental studies show that increasingly sophisticated music-emotion associations are acquired through childhood, and suggest that recognition of emotional expression in music depends on the general development of emotions (Gembris 1995). Very little is known about the development beyond childhood. However, the simple fact that listeners' preferences for music change over the life span (Russell 1997) could partly reflect the fact that certain types of music resonate better with certain phases of life in terms of the associated emotions. Rock music, for instance, with its focus on emotions like sexuality, anger, and rebellion may have a special appeal to adolescents, for which these same emotions are especially salient.

### 4.5.7 Emotions involve cognitive appraisals of organism-environment relationships

For an organism to be able to seek pleasure and avoid displeasure, it has to be able to evaluate its immediate environment. Thus, most theorists believe that emotions involve some kind of cognitive appraisal. However, theorists differ in their views on how much cognitive processing is minimally required for an emotion to occur (see, for example, Lazarus 1982; Zajonc 1980). Early theorists of emotion paid limited attention to the precise analyses of the elicitors of emotion, and the 'cognitions' associated with these elicitors. However, a number of recent 'appraisal theories' (Scherer 1999) hold that the precise nature of the emotional reaction is a result of appraisal processes undertaken on the elicitor of the emotion, and the way that elicitor is construed.

Ortony *et al.* (1988) have suggested a tripartite division of elicitors into events, agents, and objects. Developing one of their examples, one could consider a person's emotional reactions to a recently acquired new car which has broken down. As object, the car may elicit valenced reactions as a result of its aesthetic qualities (shape, colour, design). These would elicit liking or disliking. As event, the breakdown would elicit pleasure or displeasure as a result of its perceived consequences for self or others. Thus, for instance, I might be pleased with the breakdown if it prevented me from making an unwelcome journey. As agent, the car would be attributed with personality or intention, and so I might feel angry with the car for not working (irrational but common!).

On such a view, the basic emotional categories would be those given in Table 4.2. Note that the emotions contained there are relatively non-specific. Ortony *et al.* place other emotions in subcategories of the basic categories in respect of further more specific qualifying conditions of the appraisal. Thus, for instance, anger turns out to be one of the more complex emotions, being, at the very least, 'disapproval of the action of an agent in causing an event which has negative consequences for oneself'.

**Table 4.2** Basic emotional appraisal categories (adapted from Ortony *et al.* 1988)

| | Proximal elicitor construed as | | |
|---|---|---|---|
| Valence | Event | Agent | Object |
| Positive | Pleasure | Approval | Liking |
| Negative | Displeasure | Disapproval | Disliking |

Waterman (unpublished, reported in Sloboda 1991) derived a list of emotion words designed to exhaustively sample the emotion space proposed by Ortony *et al.*, and asked his participants to identify which of them they had ever felt to music, and if they had not felt it, how plausible it would be for a piece of music potentially to elicit it. These data suggested that music is capable of eliciting both positive and negative emotions relevant to all three types of construals—music can be appraised as event, action of agent, and object.

However, the more complex and specific the qualifying conditions, the less likely the emotion was to be experienced (or judged as possible for experiencing) to music. So, for instance, music elicits happiness quite easily, but almost never elicits gloating, although both emotions involve positively valenced reactions to something appraised as 'event'. Gloating denotes pleasure in the very specific context of the belief that someone else has

well as by belief in others, the listener's response can be seen to be consistent with the cognitive theory of the emotions.

The claim that attitudes other than belief can play the cognitively central role in emotions is not accepted by all who support the cognitive theory of the emotions. And it might be thought that it is one thing to imagine of the emotion's object that it has emotion-pertinent features that one does not believe it to have, yet quite another to make believe that the emotion's object exists when one does not believe it to do so. In addition, there is the concern mentioned earlier: that it is not clear that what is entertained is sufficiently controlled by what happens in the music to count as belonging to the world of the work.

The arousalist maintains that what makes it true that the music is sad is that it arouses sadness in the listener; the listener's response is not to some expressive property possessed independently by the music. While the arousalist might deny that the listener's responses *mirror* an expressiveness that is independent of their reaction, still he or she must hold that the response correlates with the music's expressiveness by licensing the judgment that the music is expressive of what is felt by the listener. Given this, and also the fact that the causal relation between the music and the listener's response need not be informed by cognitions beyond those involved in tracking the unfolding of the music, a problem remains for the arousalist in characterizing the listener's reaction as emotional.

In a recent defence of arousalism, Matravers (1998) acknowledges that the crucial response is a feeling, not an emotion as such, because it lacks the cognitive contents that characterize the emotions. For instance, the response feels like sadness or pity, and this makes it true that the music expresses sadness, but the response is not an object-directed, cognitively founded emotion. This explains why listeners are not strongly inclined to act on their feelings; the prime motivators for action are beliefs and desires directed to an emotional object, but these are absent in the musical case. Because only a few feelings have distinctive phenomenologies, music can arouse only rather general feelings, and thereby is capable of expressing only a limited range of emotions.

I endorse this approach, which can be disassociated from arousalism: if the listener does mirror the music's expressiveness, that response is caused by and tracks the music, but does not take the music, or any other thing, as its emotional object. This is not to agree, however, with the arousalist's claim that it is the listener's reaction that licenses the judgment that the music is sad.

My account can appeal to one resource that is not available to the arousalist. Earlier I suggested that inanimate appearances often strike us as expressive. To this it can be added that sometimes we find expressive characteristics in appearances highly evocative of responses of the mirroring kind (Davies 1994), not only in the musical case but in others (see Hatfield *et al.* 1994). Whether through empathy or sociality, we often 'catch' the mood prevailing around us. Both high spiritedness and despondency can be 'contagious'. The same applies sometimes, I claim, when we are confronted with powerfully expressive appearances that are not connected to occurrent emotions. There is no reason why appearances of sadness should make me feel gloomy if I do not think they show how anyone feels (and often they do not do so); which is to say, mere appearances of sadness are not a suitable object for sadness, since they are not thought to be

been less fortunate than oneself. It is this precise cognitive context that defines the emotion. Music seems ill-suited for carrying such specific messages.

Rather clear examples of the three major categories of construal are provided by the study of Waterman (1996). He asked participants to listen to prerecorded pieces of music in a comfortable and quiet laboratory situation. They were asked to press a button each time they 'felt something'. After the music ended, the recording was played back, and stopped at each point where they had pressed, with a request to explain what had been going through their minds when they pressed. Event-related appraisals often concerned structural awareness (e.g. 'I knew the end was coming'). Action-of-agent appraisals often related to the performer or composer (e.g. 'her technique fascinated me'). Object-related appraisals included evaluations of the piece as a whole (e.g. 'I liked it') or of specific aspects of it (e.g. 'I didn't like the singer's voice'). What is particularly interesting about Waterman's study is that, although different listeners tended to press the 'felt something' button at the same point in the music, their stated reasons for pressing were very different from one another. This suggests that in music there may be a partial decoupling between the mechanisms that determine intensity of affect and those that determine emotional content, the former being predominantly determined by structural characteristics of the music (*intrinsic emotion*), the latter being determined more strongly by contextual factors, including the memories, associations, and priorities of the person hearing the music (*extrinsic emotion*).

Having established that responses to music involves true instances of emotion, we move on to consider the sources of emotion in more detail.

## 4.6 Sources of emotion in music

### 4.6.1 Intrinsic emotion

There is now an accumulating body of knowledge that shows that there is a lawful relation between the intensity of emotional qualities experienced in music and the specific structural characteristics of the music at a particular point in time. The intensity of emotional response to a piece of music often rises and falls as the music unfolds. Musical discourse, both formal and informal, talks of climaxes and points of repose, tension, and relaxation. In other words, there are peaks, where intense emotions (or other affective sensations) are prone to be experienced, and troughs, where the intensity is weak.

Sloboda (1991, 1992) has identified a set of structural characteristics that are associated with the elicitation of bodily and behavioural manifestations of emotion, such as weeping or piloerection ('thrills' or 'shivers') (Goldstein 1980; Panksepp 1995). These characteristics include syncopations, enharmonic changes, melodic appoggiaturas, and other music-theoretical constructs, which have in common their intimate relationship to the creation, maintenance, confirmation, or disruption of musical expectations. The positing of a central role for such constructs in generating emotional responses to music has characterized the accounts of a number of influential musicological theorists (e.g. Meyer 1956; Narmour 1991; see also Meyer, this volume). It is only in the last decade, however, that their intuitions have been substantiated by scientific data.

The term 'intrinsic' seems appropriate because it is only with reference to other musical events and structures that particular events become more or less expected. These expectations may reflect learning, or may operate at the level of primitive perceptual processes, such as the 'gestalt laws' of perception, where, for instance, movement in a particular direction creates expectation for further movement in that direction (law of good continuation) (Meyer 1956). Narmour (1990) provides a particularly detailed theoretical working out of these ideas in the context of melodic perception and expectancy. Other expectations may arise as a function of exposure to a large body of music sharing a set of structural regularities. Some expectations relate to the entire body of, for instance, tonal music; others may relate to specific styles or genres (e.g. AABA song form, sonata form).

Sloboda (2000) has discussed some of the issues that need to be resolved before strong predictions can be made about the intensity of emotional responses on the basis of structural analysis. Thus, for example, almost every note in a piece of music confirms or violates *some* expectancy. Therefore, in order to predict actual emotional responses, further parameters must be added to the basic model. These may include density of eliciting events, positioning within the compositional architecture, and what Sloboda has called 'asynchrony of levels'—an event which simultaneously confirms an expectation on one level while violating it on another. None of these factors has yet been the subject of scientific investigation. However, some preliminary models of musical expectancy have recently been tested (Eerola & North 2000).

If musical expectation is really the key to emotional intensity, how is it that we can feel emotions to music we are highly familiar with? First, many of the violations of expectations may occur on a subconscious level, involving modules for music processing that are denied conscious access (cf. Fodor 1983). Second, even when the musical 'narrative' is familiar to us, we may still be able to enjoy it. Much like when re-watching a great movie, we can appreciate the twists and turns of a good story in music, even though we already know how the story is going to end. Third, iconic and associative sources of emotion, such as emotional contagion and memories, may remain much the same throughout repeated listening to the same piece of music (many of our emotional responses to music may simply reflect the fact that we find the music beautiful, and this beauty may be perceived each time we listen to the piece). Fourth, a basic principle says that familiarity with an object itself mght increase our liking of that object up to a certain point (Gaver & Mandler 1987). Finally, it is possible that some effects of music processing are executed by a processor, whose responses are 'hard-wired' in regard to certain perceptual primitives. If this processor is 'sealed' from other more labile parts of the system, then it may be possible that however well one knows the piece, expectation, suspense, and surprise can still occur within the processor, because the processor is always hearing the piece 'for the first time' (Jackendoff 1992).

A theory of emotional responses to music in terms of 'musical expectations' has clear explanatory value in relation to Frijda's (1986) notion of emotions as a function of monitoring match and mismatch. Most compositional systems, such as the tonal system, provide a set of dimensions that establish psychological distance from a 'home' or 'stability point'. Proximity or approach to this resting point involves reduction of tension; distance or departure involves increase of tension. Distance can be measured on a number of dimensions, including rhythm and metre (strong beats are stable, weak beats

and syncopations are unstable), and tonality (the tonic is stable, non-diatonic notes are unstable). There is now much experimental evidence that the human listener is sensitive to such features, and represents music in relation to them (e.g. Bharucha 1994; Krumhansl 1990). These features provide reference points against which the emotional system can plausibly compute match or mismatch in terms of envisaged end-points.

One problem with this approach is that it cannot without some difficulty explain why people experience different emotions in relation to pieces of music. The interplay of tension, release, surprise, and confirmation does not amount yet to a full-blown emotion. It is better characterized as 'proto-emotion', because it has a strong tendency to grow into emotions through the addition of further mental content (e.g. appraisal, valence). For instance, if the surprise is appraised as pleasurable, then an emotion of joy or elation may emerge (a similar point has been made by Ortony *et al.* 1988). What is needed to turn the structure-induced proto-emotions into full-blown emotions is semantic content. They need to be about something. Theory and evidence suggest two sources of content—iconic and associational. Because they both refer outside the music, we call these sources extrinsic.

### 4.6.2 Extrinsic emotion

*Iconic sources of emotion*

Iconic relationships come about through some formal resemblance between a musical structure and some event or agent carrying emotional 'tone' (e.g. Dowling & Harwood 1986). So, for instance, loud and fast music shares features with events of high energy and so suggests a high energy emotion such as excitement.

Iconic relationships are closely associated with the abundant literature on expression in music. Throughout history, there has been a number of different views on what music is able—or unable—to express. Music has, for example, been regarded as expressive of: (a) emotion; (b) motion; (c) beauty; (d) Christian faith; (e) tension and release; (f) things and events; (g) human character; and (h) political and social conditions. Empirical research has touched on only a few of these aspects. At some level, the common denominator for all these aspects may be human feeling, of one kind or another. Thus, it is perhaps not so surprising that psychological research has focused mainly on iconic relationships between musical features and particular emotions (for an extensive review, see Gabrielsson & Juslin, in press).

Empirical studies have shown that iconic relationships may specify particular emotions, and that they thus supply emotional content to the non-specific sensations of surprise, tension, and arousal engendered by the listener's engagement at the purely structural level of the music. Examples of studies of such iconic correspondences between musical variables and emotions are seen in Bruner (1990), Hevner (1935), Juslin (1997*b*), Rigg (1964), Scherer and Oshinsky (1977), and Wedin (1972). One recent strand in this work has been the suggestion that certain musical devices directly suggest gestural and other expressions of emotions by the human body (e.g. Clynes 1977; Scherer 1995). Another development has been studies of how performers communicate emotions to listeners through their performance of a piece of music (Gabrielsson & Juslin 1996; Juslin 1997*a*, 2000).

Unfortunately, most of these studies have measured only a few musical parameters and their main effects, whereas intercorrelations and possible interaction effects have been largely unexplored. Generally, the amount of study carried out on a particular musical parameter is inversely related to the difficulty of quantifying it. Therefore, most studies have not taken into consideration the dynamic changes in emotional expression that take place during a piece of music. Only recently have music researchers devised new techniques to capture the moment-to-moment changes in emotional expression (for a review, see Schubert, this volume). For a review of associations between composed structure and emotional expression, see Gabrielsson and Lindström (this volume). For a review of associations between performance features and emotional expression, see Juslin (this volume).

Study of emotional expression in music is probably the most common of all approaches to music and emotion in psychology—perhaps because of the relative ease with which such studies are carried out. A number of generalizations have emerged from this research. First, listeners seem to find it natural to attach emotion labels to pieces of music. Second, listeners are often consistent in their judgments and agree about the emotional expression of the music. Third, the veridicality of the judgments is rarely studied due to a lack of sufficient criteria of composers' intentions. Fourth, iconic representation of emotions in music seems to operate on a broad level of emotion categories, perhaps corresponding to the 'basic' emotion categories (cf. Dowling & Harwood 1986; Juslin 1997*c*; Krumhansl 1997). Fifth, listeners' judgments of emotion are influenced by such musical parameters as tempo, dynamics, rhythm, timbre, articulation, pitch, mode, tone attacks, and harmony.

The term 'iconic' seems particularly appropriate for this set of relationships, because the resemblances between a musical event and its non-musical referent is, in some sense, obvious to anyone who is familiar with the non-musical referent. Therefore, as a number of studies have shown (e.g. Juslin 1997*a*), recognition of iconic emotional meaning does not seem to require specific musical training. Non-musicians access iconic meanings just as easily as do musicians.

A second feature of iconic relationships is that, like the intrinsic relationships described in the previous section, they are non-arbitrarily embedded in specific structural characteristics of the music. If you change the structures then, perforce, the iconic meanings change. This means that iconic sources mediate moment-to-moment changes in emotional experience as a piece of music unfolds. Iconic relationships contribute to the ebb and flow of music-entrained emotion.

*Associative sources of emotion*

Associative sources of emotion are those that are premised on arbitrary and contingent relationships between the music being experienced and a range of non-musical factors which also carry emotional messages of their own. A memorable formulation of a key exemplar of associative sources is that of Davies (1978) who identifies associative theories as 'Darling, they're playing our tune' theories.

There are certain types of stimuli (e.g. music, smells, and tastes) which seem to become associated in human memory with particular contexts or events in earlier life, and provide a trigger to the recall of these events. This seems particularly so when the

earlier events were, in themselves, occasions of strong emotion (Dutta & Kanungo 1975). Many investigators (e.g. Gabrielsson 1991; Sloboda 1991) have found examples of specific pieces of music that trigger strong emotion in this way. Such emotions tend to lead attention away from the present music onto the remembered past event.

Waterman (1996) has demonstrated that even when music does not directly trigger past experiences, many of the emotional processes are self-referring in some way (e.g. 'I should have recognized that', 'this is not my type of music'). Because these feelings are linked to the life histories of individuals they are often completely idiosyncratic. However, common cultural experiences can sometimes lead to shared emotions which still are fundamentally extrinsic, for example, the extreme negative emotions felt by many Jews after the Second World War on hearing the music of Richard Wagner; the strong emotional identification of generational cohorts with the popular music prevalent in their teenage years (e.g. Holbrook & Schindler 1989); and the cultural associations formed by film-music pairings, such as Johann Strauss' 'Blue Danube' waltz, with the spaceship docking sequence in Stanley Kubrick's film *2001: A Space Oddesey.*

Unlike the intrinsic and iconic sources of emotion, associative sources are often bound non-specifically to entire pieces or passages, and do not require any particular feature of the music for their binding (i.e. other than some means of identifying the music in question as the relevant music rather than a different piece or genre). Although associative sources of emotion are obvious and pervasive, there is little scientific research into them. This may be because the intrinsic idiosyncrasy of these associations renders control and generalizability problematic.

### 4.6.3 Interactions between different sources of emotion

In the foregoing, we noted that emotional responses to music reflect both intrinsic emotion that is locally focused, and extrinsic emotion that is more globally context-dependent. Furthermore, music both induces emotions in listeners and is perceived by listeners as expressive of emotion. This means that emotional responses to music may come about in a large number of different ways, and failure to recognize this fact may account for some of the past disagreements in the field. In each listening situation, (a) all, (b) only some, or (c) none of the different sources of emotion may be present. Much of the richness of emotional responses to music comes from this multitude of sources of emotion. Moreover, different sources of emotion may interact in interesting ways. For example, the mood of a particular piece of music (iconic signification) may evoke mood-congruent memories (associative signification) through so-called 'spread of activation' in neural networks associated with emotional reactions (Bower 1981). Similarly, the manner in which a particular section of a piece of music builds up tension through musical expectations (intrinsic affect) may remind us of the patterns of tension and release in a certain extramusical context (extrinsic affect). Furthermore, different sources of emotion may support or contradict one another, thus creating either clarity or ambiguity in the emotional impact of the music. The emotional ambiguity of certain pieces of music is, of course, reminiscent of the mixed emotions we often experience in many other domains of life (Oatley & Duncan 1994).

Music (instrumental music, at least) often suggest emotions in a vague fashion. Sloboda (2000, p. 226) notes that music:

can suggest or resemble certain types of human gestures and actions. Given shared experience and understandings, it is easy to see how a group of people might be able to extrapolate specific emotions from such cues. It is equally easy to see how there is room for ambiguity and imprecision. Very often we feel that there is an emotion present, we know it is of one general type rather than that of another, but we cannot quite tie it down. In such a state of ambiguity and cue-impoverishment we may well expect the profound and semi-mystical experiences that music seems to engender. Our own subconscious desires, memories, and preoccupations rise to the flesh of the emotional contours that the music suggests. The so-called 'power' of music may very well be in its emotional cue-impoverishment. It is a kind of emotional Rorschach blot.

## 4.7 Implications and future directions

This chapter has suggested a multitude of ways in which emotional processes may occur in relation to music. Yet, this chapter has provided only a small glimpse of all the work that has been carried out within the psychological approach to music and emotion. One implication of the foregoing overview is that emotional responses to music have very much in common with emotional responses to other stimuli. However, in the case of a non-musical stimulus, emotion is usually generated via appraisal of the capacity of the stimulus to affect the interests, values, and goals of the perceiver. Because music has no direct capacity to further or block such goals, a challenge for theorists of music and emotion has been to provide an alternative but plausible account of the casual path between musical features and experienced emotion, for instance in terms of musical expectations or various extramusical associations.

Given the emphasis on emotional responses to music in this chapter, we should perhaps note that music does not always arouse emotional responses in listeners. Often, we may listen to a piece of music without feeling anything in particular. In fact, some people claim that they rarely, if ever, experience emotional reactions to music. Unfortunately, there is little research so far on the *epidemiological* aspects of musical emotions; in other words, how often people experience emotions to music under various circumstances. This represents a crucial area for future research. We should attempt to delineate the conditions (including the music, person, and situation) under which a listener is most likely to react emotionally to music. This might involve the kind of diary approach that have been used in earlier research on epidemiological aspects of emotions (Oatley & Duncan 1994).

Psychological studies of music and emotion have largely ignored the interaction among three factors: the listener, the music, and the context (cf. Gabrielsson, this volume). Previous research has focused mainly on factors in the music. We know quite a lot about the ways that factors in the music can induce and represent emotions. However, the fact that different styles of music may yield very different emotional responses in listeners has received little attention. Therefore, it should be acknowledged that theoretical explanations of emotional responses relevant to one particular style of music (e.g. classical music) may not apply equally well to a different style (e.g. blues).

A fair amount of research has focused on factors in the individual that might influence emotional responses to music. Listeners' affective responses to music are likely to depend to some extent on their age, gender, personality, training, current mood, and so

forth (Abeles & Chung 1996). Interactions between music and listener may involve the listener's familiarity with the music (North & Hargreaves 1997) and his or her musical preferences (Holbrook & Schindler 1989; Russell 1997).

The factor that has been mostly ignored in previous research is the context—including everything from the concrete situation in which the musical activity takes place, to the wider social and cultural context. Konečni (1982) has strongly criticized previous empirical studies for having treated aesthetical responses as if they occur in a social, emotional, and cognitive 'vacuum'. Emotional responses to music are likely to depend on what the music is used for. To understand emotional responses to music, we should consider the social functions of the music in its particular context. Hargreaves and North (1999) have suggested that the social functions of music manifest themselves in three main ways: (a) the formulation and expression of self-identity (our responses to some musical styles may partly reflect what they stand for in the larger social context); (b) the establishment and maintenance of personal relationships (e.g. prestige effects and conformity); and (c) mood management (music tastes reflect situationally determined goals of 'mood optimization'). Studies of music and emotion have recently begun to address these issues (e.g. DeNora 2000, this volume; North & Hargreaves 1997; Sloboda & O'Neill, this volume).

One example of the strong social influence is the kind of emotion that arises from what one has been told, and therefore may come to believe, about the music in question. A seminal study by Weick *et al.* (1973) manipulated the reputation of the composer of a jazz orchestra piece by telling some players it was by an established jazz composer and others that it was by a dance band arranger (not exactly a high prestige genre for jazz players). This manipulation clearly affected emotional and motivational involvement with the piece. Players who thought the piece was by an established composer both liked the piece more and devoted more effort to learning it. As suggested by this study, almost every piece of music is given a cultural value by members of the community within which it is experienced, produced, and discussed. Therefore, there is little music that we come to 'blind'. Our hearing of a Beethoven Symphony cannot now be dissociated from the ubiquitous discourse (available on programme notes, CD cases, radio stations, etc.) that assigns him a pre-eminent position of 'greatness' in the canon of masterworks. Because greatness has been ascribed to Beethoven, our emotional response to hearing a famous Beethoven work is conditioned by cultural attitudes appropriate to 'greatness' (e.g. reverence). We may rebel against these tendencies, but we cannot ignore them.

Such considerations lead us into the heart of contemporary debates in musicology (see Cook 1998). Traditional musicology has become increasingly under fire for being unreflexive about the cultural and social sources of its own attitude toward the music it was studying, an attitude which could lead to the assumption that Beethoven's greatness was 'for all times and all places' and had nothing to do with the particularities of nineteeth- and twentieth-century European societies into which his music has been received and championed. The increasing self-critical reflexivity of the musicological community augurs well for developments in psychology. Many psychologists have been 'victims' of the same cultural unawareness that dominated mid twentieth-century musicology, and this may be one reason why this area has not

been well researched. Psychologists have not really grasped the nature of the problem. To put the problem is its most acute form, one of the most prevalent cultural discourses about music concerns its emotional power. This discourse surrounds all of us from our earliest years. We, therefore, approach any new music already expecting (even willing) it to have an emotional content. One answer to the question 'why is music so emotionally powerful' is that we have decided to construe it thus. At one very important level of analysis, music is not inherently emotional. It is the way we hear it that makes it so (see Becker, this volume).

Because the concept of emotion itself relies on research from such diverse domains as philosophy, psychology, sociology, biology, anthropology, and physiology, we believe that interdisciplinary collaborations provide the royal road to progress in this field. Juslin (1997*c*, pp. 15–16) notes that music researchers have approached questions of music and emotion with empirical rigor but that:

> this has unfortunately not been matched by an equal sophistication concerning conceptual problems. There are often vague references to the 'musical emotions', without further clarification of the meaning of this concept. Philosophers may, of course, contribute to these conceptual problems through thoughtful analyses, but too often they are not familiar with the latest developments in emotion psychology, which means that their ground premises may be wrong. In the end, it may turn out that we can make progress in this area only through the combined efforts of different disciplines.

We believe that psychologists must learn to draw from the knowledge currently available in related disciplines—philosophy, sociology, anthropology, biology, and musicology.

Given the problems of the study of music and emotion, we think it is also desirable for music researchers to follow closely the developments in the emotion domain more generally. Recent debates in emotion psychology have involved: (a) the role and limits of self-reports in the study of emotion; (b) cross-cultural agreement in emotion judgments; (c) the nature and existence of basic emotions; (d) individual differences in emotion disposition; (e) the tracking of phasic aspects of emotion; (f) linguistic analyses of emotion; and (g) relationships between emotion and cognition (Cacioppo & Gardner 1999), all of which have a direct bearing on the study of emotion in music. The implication is that music researchers cannot escape from the theoretical and methodological problems confronted by other researchers who study emotion. Studies of music and emotion that measure emotions in multiple channels simultaneously (self-report, physiology, expressive behaviour), and over time, is clearly an important goal for future research. Such an endeavour may have to involve a switch from the traditional agrarian-based inferential statistics to non-linear, dynamic models that are better suited to the time-dependent nature of emotional responses to music.[1]

[1] We thank Carol Krumhansl and Klaus Scherer for useful comments on a preliminary version of this chapter.

## References


Abeles, H. F. & Chung, J. W. (1996). Responses to music. In *Handbook of music psychology* (2nd edn) (ed. D. A. Hodges), pp. 285–342. San Antonio, TX: IMR Press.

Bartlett, D. L. (1996). Physiological reactions to music and acoustic stimuli. In *Handbook of music psychology* (2nd edn) (ed. D. A. Hodges), pp. 343–85. San Antonio, TX: IMR Press.

Batson, C. D., Shaw, L. L., & Oleson, K. C. (1992). Differentiating affect, mood, and emotion: Toward functionally based conceptual distinctions. In *Review of personality and social psychology: Emotion*, (ed. M. S. Clark), pp. 294–326. Newbury Park, CA: Sage.

Behne, K. E. (1997). The development of 'musikerleben' in adolescence: How and why young people listen to music. In *Perception and cognition of music*, (ed. I. Deliége & J. A. Sloboda), pp. 143–59. Hove, UK: Psychology Press.

Berlyne, D. E. (1971). *Aesthetics and psychobiology.* New York: Appleton Century Crofts.

Bharucha, J. J. (1994). Tonality and expectation. In *Musical perceptions*, (ed. R. Aiello, & J. A. Sloboda), pp. 213–39. New York: Oxford University Press.

Bless, H. & Forgas, J. P. (ed.) (2000). *The message within. The role of subjective experience in social cognition and behavior.* Hove, UK: Psychology Press.

Bower, G. H. (1981). Mood and memory. *American Psychologist*, 36, 129–48.

Bruner, G. C. (1990). Music, mood, and marketing. *Journal of Marketing*, 54, 94–104.

Budd, M. (1985). *Music and the emotions. The philosophical theories.* London: Routledge.

Cacioppo, J. T. & Gardner, W. L. (1999). Emotion. *Annual Review of Psychology*, 50, 191–214.

Cacioppo, J. T., Klein, D., Berntson, G. G., & Hatfield, E. (1993). The psychophysiology of emotion. In *Handbook of emotions*, (ed. M. Lewis & J. M. Haviland), pp. 119–42. New York: Guildford Press.

Clynes, M. (1977). *Sentics: The touch of emotions.* New York: Doubleday.

Cook, N. (1998). *A very short introduction to music.* Oxford, UK: Oxford University Press.

Cunningham, J. G. & Sterling, R. S. (1988). Developmental changes in the understanding of affective meaning in music. *Motivation and Emotion*, 12, 399–413.

Damasio, A. (1994). *Descartes' error. Emotion, reason, and the human brain.* New York: Grosset/Putnam.

Darwin, C. (1872). *The expression of the emotions in man and animals.* London: John Murray.

Davidson, R. J. (1994). On emotion, mood, and related affective constructs. In *The nature of emotion: Fundamental questions*, (ed. P. Ekman & R. J. Davidson), pp. 51–5. New York: Oxford University Press.

Davies, J. B. (1978). *The psychology of music.* London: Hutchinson.

DeNora, T. (2000). *Music in everyday life.* Cambridge, UK: Cambridge University Press.

De Sousa, R. (1987). *The rationality of emotions.* Cambridge, MA: MIT Press.

Dewey, J. (1894). The theory of emotions. I. Emotional attitutes. *Psychological Review*, 1, 553–69.

Dolgin, K. & Adelson, E. (1990). Age changes in the ability to interpret affect in sung and instrumentally-presented melodies. *Psychology of Music*, 18, 87–98.

Dowling, W. J. & Harwood, D. L. (1986). *Music cognition.* New York: Academic Press.

Dutta, S. & Kanungo, R. N. (1975). *Affect and memory: A reformulation.* New York: Pergamon Press.

Eerola, T. & North, A. C. (2000). Expectancy-based models of melodic complexity. In *Proceedings of the Sixth International Conference on Music Perception and Cognition, August 2000*, (CD rom) (ed. C. Woods, G. Luck, R. Brochard, F. Seddon, & J. A. Sloboda). Keele, UK: Keele University.

Ekman, P. (ed.) (1973). *Darwin and facial expression.* New York: Academic Press.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6, 169–200.

Ekman, P. & Davidson, R. J. (ed.). (1994). *The nature of emotion: Fundamental questions.* New York: Oxford University Press.

Fehr, B. & Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General*, 113, 464–86.

Fodor, J. A. (1983). *The modularity of the mind.* Cambridge, MA: MIT Press.

Fraisse, P., Oleron, G., & Paillard, J. (1953). Les effets dynamogeniques de la musique. *L'Anée Psychologique*, 53, 1–34.

Fridja, N. H. (1986). *The emotions.* New York: Cambridge University Press.

Fridja, N. H. (1989). Aesthetic emotions and reality. *American Psychologist*, 44, 1546–7.

Fridja, N. H., Kuipers, P., & ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, 57, 212–28.

Fried, R. & Berkowitz, L. (1979). Music that charms ... and can influence helpfulness. *Journal of Applied Social Psychology*, 9, 199–208.

Gabrielsson, A. (1991). Experiencing music. *Canadian Journal of Research in Music Education*, 33, 21–6.

Gabrielsson, A. (1993). Music and emotion. *ESCOM Newsletter*, 4, 4–9.

Gabrielsson, A. & Juslin, P. N. (1996) Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music*, 24, 68–91.

Gabrielsson, A. & Juslin, P. N. (in press). Emotional expression in music. In *Handbook of affective sciences*, (ed. R. J. Davidson, H. H. Goldsmith, & K. R. Scherer). New York: Oxford University Press.

Gabrielsson, A. & Lindström, S. (1993). On strong experiences of music. *Jarbuch der Deutschen Gesellschaft für Musikpsychologie*, 10, 114–25.

Gaver, W. W. & Mandler, G. (1987). Play it again, Sam: On liking music. *Cognition and Emotion*, 1, 259–82.

Gembris, H. (1995). The development of perception of emotional expression in music: A review. Paper presented at the ESCOM Conference, Jahrestagung Deutsche Gesellschaft für Musikpsychologie, Bremen, 15–17 September 1995.

Gigerenzer, G. & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650–69.

Goldstein, A. (1980). Thrills in response to music and other stimuli. *Physiological Psychology*, 8, 126–9.

Gross, J. J. & Levinson, R. W. (1993). Emotion suppression: physiological, self-report, and expressive behaviour. *Journal of Personality and Social Psychology*, 64, 970–86.

Hargreaves, D. J. & North, A. (1999). The functions of music in everyday life: Redefining the social in music psychology. *Psychology of Music*, 27, 71–83.

Harrer, G. & Harrer, H. (1977). Music, emotion, and autonomic function. In *Music and the brain. Studies in the neurology of music*, (ed. M. Critchley & R. A. Henson), pp. 202–16. London: William Heinemann Medical Books.

Harris, P. L. (1989). *Children and emotion.* Oxford, UK: Basil Blackwell.

**Hatfield, E., Cacioppo, J. T., & Rapson, R. L.** (1994). *Emotional contagion.* New York: Cambridge University Press.

**Hevner, K.** (1935). Expression in music: A discussion of experimental studies and theories. *Psychological Review*, 42, 186–204.

**Holbrook, M. B. & Schindler, R. M.** (1989). Some exploratory findings on the development of musical tastes. *Journal of Consumer Research*, 16, 119–24.

**Imberty, M.** (1997). Can one seriously speak of narrativity in music? In *Proceedings of the Third Triennial ESCOM Conference, 1997*, (ed. A. Gabrielsson), pp. 13–22. Uppsala, Sweden: Department of Psychology, Uppsala University.

**Izard, C. E.** (1977). *The emotions.* New York: Plenum Press.

**Jackendoff, R.** (1992). Musical processing and musical affect. In *Cognitive bases of musical communication*, (ed. M. R. Jones & S. Holleran), pp. 51–68. Washington, DC: American Psychological Association.

**James, W.** (1884). What is an emotion? *Mind*, 9, 188–205.

**Johnson-Laird, P. N. & Oatley, K.** (1992). Basic emotions, rationality, and folk theory. *Cognition and Emotion*, 6, 201–23.

**Juslin, P. N.** (1997*a*). Emotional communication in music performance: A functionalist perspective and some data. *Music Perception*, 14, 383–418.

**Juslin, P. N.** (1997*b*). Perceived emotional expression in synthesized performances of a short melody: Capturing the listener's judgment policy. *Musicae Scientiae*, 1, 225–56.

**Juslin, P. N.** (1997*c*). Can results from studies of perceived expression in musical performances be generalized across response formats? *Psychomusicology*, 16, 77–101.

**Juslin, P. N.** (2000). Communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 1797–1813.

**Juslin, P. N.** (2001). A Brunswikian approach to emotional communication in music performance. In *The essential Brunswik: Beginnings, explications, applications*, (ed. K. R. Hammond & T. R. Stewart), pp. 426–430. New York: Oxford University Press.

**Kastner, M. P. & Crowder, R. G.** (1990). Perception of the major/minor distinction: IV. Emotional connotations in young children. *Music Perception*, 8, 189–202.

**Keltner, D. & Gross, J. J.** (1999). Functional accounts of emotions. *Cognition and Emotion*, 13, 465–6.

**Kemper, T. D.** (1987). How many emotions are there? Wedding the social and the autonomic components. *American Journal of Sociology*, 93, 263–89.

**Kenealy, P.** (1988). Validation of a music mood induction procedure: Some preliminary findings. *Cognition and Emotion*, 2, 41–8.

**Kivy, P.** (1990). *Music alone: Philosophical reflections on the purely musical experience.* Ithaca, NY: Cornell University Press.

**Kleinginna, P. R. & Kleinginna, A. M.** (1981). A categorized list of emotion definitions, with a suggestion for a consensual definition. *Motivation and Emotion*, 5, 345–71.

**Konečni, V. J.** (1979). Determinants of aesthetic preference and effects of exposure to aesthetic stimuli: Social, emotional, and cognitive factors. In *Progress in experimental personality research*, (Vol. 9) (ed. B. A. Maher), pp. 149–97. Orlando, FL: Academic Press.

**Konečni, V. J.** (1982). Social interaction and musical preference. In *The psychology of music*, (ed. D. Deutsch), pp. 497–516. San Diego, CA: Academic Press.

**Kratus, J.** (1993). A developmental study of children's interpretation of emotion in music. *Psychology of Music*, **21**, 3–19.

**Krumhansl, C. L.** (1990). *Cognitive foundations of musical pitch.* Oxford, UK: Oxford University Press.

**Krumhansl, C. L.** (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology*, **51**, 336–52.

**Langer, S.** (1951). *Philosophy in a new key.* (2nd edn). New York: New American Library (original work published 1942).

**Lazarus, R. S.** (1982). Thoughts on the relations between emotion and cognition. *American Psychologist*, **37**, 1019–24.

**Lazarus, R. S.** (1991). *Emotion and adaptation.* New York: Oxford University Press.

**Leahey, T. H.** (1987). *A history of psychology. Main currents in psychological thought*, (2nd edn). Englewood Cliffs, NJ: Prentice Hall.

**LeDoux, J.** (1996). *The emotional brain.* New York: Simon & Schuster.

**Levinson, J.** (1996). *The pleasures of aesthetics.* Ithaca, NY: Cornell University Press.

**Mandler, G.** (1984). *Mind and body. Psychology of emotions and stress.* New York: Norton.

**Mayer, J. D. & Salovey, P.** (1993). The intelligence of emotional intelligence. *Intelligence*, 17, 433–42.

**Meyer, L. B.** (1956). *Emotion and meaning in music.* Chicago: Chicago University Press.

**Moog, H.** (1976). *The musical experience of the pre-school child.* London: Schott.

**Narmour, E.** (1990). *The analysis and cognition of basic melodic structures.* Chicago: University of Chicago Press.

**Narmour, E.** (1991). The top-down and bottom-up systems of musical implication: Building on Meyer's theory of emotional syntax. *Music Perception*, 9, 1–26.

**Neale, J. M. & Liebert, R. M.** (1986). *Science and behavior.* Englewood Cliffs, NJ: Prentice Hall.

**Niedenthal, P. M. & Showers, C.** (1991). The perception and processing of affective information and its influences on social judgment. In *Emotion and social judgments*, (ed. J. P. Forgas), pp. 125–43. Oxford, UK: Pergamon Press.

**North, A. C. & Hargreaves, D. J.** (1997). Experimental aesthetics and everyday music listening. In *The social psychology of music*, (ed. D. J. Hargreaves & A. C. North), pp. 84–103. Oxford, UK: Oxford University Press.

**Nyklicek, I., Thayer, J. F., & van Doornen, L. J. P.** (1997). Cardiorespiratory differentiation of musically-induced emotions. *Journal of Psychophysiology*, **11**, 304–21.

**Oatley, K.** (1992). *Best laid schemes. The psychology of emotions.* Cambridge, MA: Harvard University Press.

**Oatley, K. & Duncan, E.** (1994). The experience of emotions in everyday life. *Cognition and Emotion*, **8**, 369–81.

**Oatley, K. & Jenkins, J. M.** (1996). *Understanding emotions.* Oxford, UK: Blackwell.

**Ortony, A, Clore, G. L., & Collins, A.** (1988). *The cognitive structure of the emotions.* New York: Cambridge University Press.

**Ortony, A. & Turner, T. J.** (1990). What's basic about basic emotions? *Psychological Review*, 97, 315–31.

**Panksepp, J.** (1992). A critical role for affective neuroscience in resolving what is basic about basic emotions. *Psychological Review*, **99**, 554–60.

Panksepp, J. (1995). The emotional sources of 'chills' induced by music. *Music Perception*, **13**, 171–208.

Pike, A. (1972). A phenomenological analysis of emotional experience in music. *Journal of Research in Music Education*, **20**, 262–7.

Plutchik, R. (1980). A general psychoevolutionary theory of emotions. In *Emotion: Theory, research, and experience. Vol 1. Theories of emotion*, (ed. R. Plutchik & H. Kellerman), pp. 3–33. New York: Academic Press.

Plutchik, R. (1994). *The psychology and biology of emotion*. New York: Harper Collins.

Power, M. & Dalgleish, T. (1997). *Cognition and emotion. From order to disorder*. Hove, UK: Psychology Press.

Rigg, M. G. (1964). The mood effects of music: A comparison of data from earlier investigations. *Journal of Psychology*, **58**, 427–38.

Rosch, R. (1978). Principles of categorization. In *Cognition and categorization*, (ed. E. Rosch & B. B. Loyd), pp. 27–48. Hillsdale, NJ: Erlbaum.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, **39**, 1161–78.

Russell, P. A. (1997). Musical tastes and society. In *The social psychology of music*, (ed. D. J. Hargreaves & A. C. North), pp. 141–58. Oxford, UK: Oxford University Press.

Schacter, S. & Singer, J. E. (1962). Cognitive, social, and physiological determinants of emotional states. *Psychological Review*, **69**, 379–99.

Scherer, K. R. (1993). Neuroscience projections to current debates in emotion psychology. *Cognition and Emotion*, **7**, 1–41.

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice*, **9**, 235–48.

Scherer, K. R. (1999). Appraisal theories. In *Handbook of cognition and emotion*, (ed. T. Dalgleish & M. Power), pp. 637–63. Chichester, UK: Wiley.

Scherer, K. R. & Oshinsky, J. S. (1977). Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion*, **1**, 331–46.

Schlosberg, H. (1941). A scale for the judgment of facial expressions. *Journal of Experimental Psychology*, **29**, 497–510.

Shaver, P. R., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology*, **52**, 1061–86.

Simon, H. (1967). Motivational and emotional controls of cognition. *Psychological Review*, **74**, 29–39.

Sloboda, J. A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music*, **19**, 110–20.

Sloboda, J. A. (1992). Empirical studies of emotional response to music. In *Cognitive bases of musical communication*, (ed. M. Riess-Jones & S. Holleran), pp. 33–46. Washington, DC: American Psychological Association.

Sloboda, J. A. (1996). Emotional responses to music: A review. In *Proceedings of the Nordic Acoustical Meeting*, (ed. K. Riederer & T. Lahti), pp. 385–92. Helsinki: The Acoustical Society of Finland.

Sloboda, J. A. (2000). Musical performance and emotion: Issues and developments. In *Music, Mind, and Science*, (ed. S. W. Yi), pp. 220–38. Seoul, Korea: Western Music Research Institute.

Smith, L. D. & Williams, R. N. (1999). Children's artistic responses to musical intervals. *American Journal of Psychology*, **112**, 383–410.

**Sorce, J., Emde, R., Campos, J, & Klinnert, M.** (1985). Maternal emotional signalling: Its effect on the visual cliff behavior of one-year olds. *Developmental Psychology*, **21**, 195–200.

**Stern, D.** (1985). *The interpersonal world of the infant. A view from psychoanalysis and developmental psychology.* New York: Basic Books.

**Terwogt, M. M. & van Grinsven, F.** (1988). Musical expression of mood states. *Psychology of Music*, **19**, 99–109.

**Thayer, J. F.** (1986). Multiple indicators of affective responses to music. *Dissertation Abstracts International*, **47**, 12.

**Tomkins, S.** (1962). *Affect, imagery, and consciousness: The positive affects.* New York: Springer.

**Waterman, M.** (1996). Emotional responses to music: Implicit and explicit effects in listeners and perfomers. *Psychology of Music*, **24**, 53–67.

**Wedin, L.** (1972). A multi-dimensional study of perceptual-emotional qualities in music. *Scandinavian Journal of Psychology*, **13**, 1–17.

**Weick, K. E., Gilfillian, D. P., & Keith, T. A.** (1973). The effect of composer credibility on orchestra performance. *Sociometry*, **36**, 435–62.

**Witvliet, C. V. & Vrana, S. R.** (1996). The emotional impact of instrumental music on affect ratings, facial EMG, autonomic response, and the startle reflex: Effects of valence and arousal. *Psychophysiology Supplement*, **91**.

**Woodworth, R. S.** (1938). *Experimental psychology.* New York: Holt.

**Wundt, W.** (1897). *Outlines of psychology* (trans. C. H. Judd). Lepzig: Englemann.

**Zajonc, R. B.** (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, **35**, 151–75.

**Zajonc, R. B.** (1994). Emotional expression and temperature modulation. In *Emotions: Essays on emotion theory*, (ed. S. H. M. Van Goozen, N. E. Van de Poll, & J. A. Sergeant), pp. 3–27. Hillsdale, NJ: Erlbaum.

**Zillman, D.** (1988). Mood management: using entertainment to full advantage. In *Communication, social cognition, and affect*, (ed. L. Donohew, H. E. Sypher, & E. T. Higgens), pp. 147–71. Hillsdale, NJ: Earlbaum.

CHAPTER 5

# LISTEN TO THE BRAIN: A BIOLOGICAL PERSPECTIVE ON MUSICAL EMOTIONS

ISABELLE PERETZ

The functioning of the brain is fascinating. It fascinates for the obvious reason that the brain is the commander of all our actions, thoughts, and motivations. By studying its functioning, the hope is to obtain crucial information about the biological determinants of human cognition and emotion. Neuropsychology is the discipline concerned with these questions. As its name indicates, neuropsychology aims to relate neural mechanisms to mental functions. It is an old discipline, dating back to the discovery that speech was related to the functioning of a small region of the left brain by Broca in 1861. Following this discovery, the neural correlates of musical abilities were similarly scrutinized (for a recent review, see Marin & Perry 1999). Although the neuropsychological approach to music is a century old, progress has been slow. However, the recent advances made in brain-imaging techniques, as well as the current trend of viewing most human activities from a biological perspective, has intensified research activities in the field.

As a result of this biological trend, neuropsychology has been recently renamed 'cognitive neuroscience'. This change in terminology reflects the intention to include neuropsychology within the vast domain of neuroscience. Since neuroscience covers a large spectrum of disciplines, going from the physiological study of single neurons in the turtle retina (Sernagor & Grzywacz 1996) to the brain organization subserving the sense of humour in humans (Shammi & Stuss 1999), 'cognitive' has qualified the 'neuroscience' term. The qualifier is unfortunate because 'cognitive' usually excludes emotion. Cognition is often seen as antagonistic to emotion. For example, Kivy (1990) in his influential essay on the meaning of music coined the term 'cognitivist' to refer to the position that holds that music simply expresses emotions without inducing them. The opposite position, called 'emotivist', holds that music elicits emotional responses in listeners.

The antagonism between 'cognitivists' and 'emotivists' is not limited to music psychology. It has a long history starting with Descartes' early separation of emotion and reason. Still, today, the majority of experimental psychologists and neuropsychologists are 'cognitivists' by default. They have simply ignored emotions. This neglect partly reflects the information-processing approach that started in the early 1960s (e.g. Neisser 1967) and that used the computer as a metaphor for mental functions. Following this approach, the brain is a machine, devoid of emotions.

Recently, neuropsychologists have become more concerned with emotions. Although neuropsychologists do distinguish between emotional and cognitive processes, they no longer reject the emotional part as being too obscure or subjective to be scientifically studied. On the contrary, emotions are now studied for their own sake. In their two popular books published in 1994 and 1996, respectively, Damasio and LeDoux have greatly contributed to the *rapprochement* between cognition and emotion in neuropsychology. Reason is no longer seen as the human-specific activity that controls emotional irrationality. As Damasio (1994) has shown, emotional processes are an integral part of decision making and are no longer confined to subcortical brain structures that are shared with other animals. Emotions recruit portions of the frontal lobes, which are the largest and latest brain structures to develop in the human brain. Thus, neuropsychologists are currently studying emotions in the same way as any other mental function worthy of enquiry.

The objective of this chapter is to present current knowledge about musical emotions from a biological perspective. I will introduce this topic by first covering neuropsychological data accumulated on facial expressions of emotions. This area of investigation has made remarkable progress recently. An outline of the major findings will allow me to illustrate a number of principles that characterize the field of neuropsychology in general. After this excursion into the facial processing domain, I will cover what is currently known about the neural correlates of musical emotions. Then, I will address the possibility that musical emotions follow similar paths to vocal emotions (the emotional expressions produced by non-verbal inflections of the human voice) by reviewing the relevant literature. Although the neuropsychological study of musical emotions lags behind, drawing parallels across domains should benefit the musical endeavour and, hopefully, promote cross-fertilization.

## 5.1 An illustration of the rise of affective neuroscience: facial expressions

Emotions have generally been considered too personal, elusive, and variable to be studied scientifically. *A fortiori*, emotions could hardly be conceived as subserved by a neuroanatomical arrangement that can be shared by all members of the same species. The work of Paul Ekman on human facial expressions has contributed to convince the scientific community that the above view is inadequate. In what follows, I will summarize his work and the subsequent research done on the neural substrates of human facial expressions, because it has spurred my own interest in studying the neural basis of musical emotions. Above all, it will allow me to introduce basic concepts about the neuroanatomical substrates of emotions and the major technologies by way of an example rather than as a disembodied general model associated with a series of techniques without objects.

### 5.1.1 Facial expressions: universality and innateness

Like music, facial expressions are non-verbal communicative displays, which can be very subtle but none the less quickly and easily appreciated by humans. Infants show a precocious ability to respond to different facial expressions of emotions (Field *et al.*

1982). Darwin (1872) was the first to hypothesize the innateness and universality of facial expressions of emotions, but his proposal received little attention. It was only in the late 1960s, with the work of Ekman (Ekman 1994; Ekman *et al.* 1969) and Izard (1971), that Darwin's position regained credibility. At the time Ekman and Izard began their studies, it was thought that facial expressions of emotions are socially learned and therefore variable from culture to culture. Ekman and his collaborators dismissed the idea that facial expressions are pure social constructs, in showing that a variety of emotions are expressed cross-culturally by the same facial movements (even though there might be culture-specific variations as captured by Ekman's concept of 'display rules').

For example, Ekman and collaborators (1987) reported a study in which facial expressions were examined in Japanese and Americans while they watched stress-inducing films (e.g. body mutilations) and neutral films (nature scenes). When the subjects in each culture watched the films alone, unaware of a hidden camera, the same negative facial responses were emitted regardless of culture. However, when a scientist was present while they watched the films, the Japanese more than the Americans masked their expressions with smiles. This study illustrates how universal expressions of disgust on faces can manifest themselves, beyond cultural differences, in their display. In other studies, the same group of researchers has convincingly demonstrated that facial expressions of 'basic' emotions (for a further explanation of this concept, see Sloboda & Juslin, this volume), such as happiness, sadness, anger, and fear, are easy to recognize by members of widely different and isolated cultures. The general rules and processes subserving facial expressions appear universal.

The universality of emotion expressions is necessary but not sufficient for considering them to be innate or biologically determined, as initially posited by Darwin (1872). Universality can arise as common learning experiences because all infants are exposed to the same situations, such as pain, being left alone, or reassured, across cultures. However, as summarized by Oster and Ekman (1978), even premature newborns are capable of producing all the discrete facial muscle movements involved in the expressions of basic emotions such as disgust, endogenous smiling, and sadness. Children born deaf and blind show similar facial expressions to those of hearing and seeing children (e.g. Eibl-Eibesfeldt 1973). Before 1 year of age, infants demonstrate meaningful relations between events and expressions of joy, sadness, fear, disgust, and anger. It is unlikely that young infants could have learned all these complex muscular patterns of facial expressions so quickly. The notion that facial expressions are learned lacks plausibility. Rather, facial expressions appear constrained by innate mechanisms. Thus, the innateness and universality of facial expressions appears sufficiently robust to consider Darwin's hypothesis established.

### 5.1.2 Facial expressions: brain specialization

The fact that facial expressions of emotion exhibit universality and innateness is important from a neuropsychological perspective. More specifically, brain organization is expected to differ for functions that respond to cultural pressure and for functions that meet biological requirements. Functions resulting from social conventions might be implemented in neural networks that vary across individuals. Depending on the quality and quantity of familiarization with a skill, various brain regions might be mobilized to

different degrees in individuals. Driving a car, an obvious cultural and recent acquisition in the human behavioural repertoire, would be expected to be subserved by variable, not specialized, neural networks. In order to drive a car, a person would be expected to borrow functional principles and neural resources from other skills, such as navigating, and to adjust them to their processing needs. In contrast, biologically important functions are likely to be subserved by a fixed and specialized neural architecture. This is probably the case for both expression and recognition of emotions in faces, since they exhibit universality and precociousness. Hence, the neural basis of facial expressions is expected to exhibit the same functional and neuroanatomical specialization across human brains. Accordingly, specification of the functional neuroanatomy underlying facial expressions should be relatively easy.

The first neural property that facial expressions should possess is brain specialization. That is, recognition of emotional expressions should be isolable in the brain. This property is indeed verified. Selective losses in the ability to recognize emotions from faces have been reported after focal brain damage (e.g. Calder *et al.* 1996; Etcoff 1984; Parry *et al.* 1991; Sprengelmeyer *et al.* 1996; Young *et al.* 1993). Such patients can recognize all known faces and infer the age of unknown faces, but are unable to see variations in emotional expressions. Conversely, other patients remain capable of discriminating and identifying various facial expressions, while being entirely unable to recover the identities of familiar faces of relatives and even of themselves (e.g. Tranel *et al.* 1988).

The study of the effects of brain lesions on a particular mental function is the oldest method used in neuropsychology. The observed disorders reveal both how the function under study is organized functionally and what part of the brain is *necessarily* involved in its processing. The idea is that the disorder is not arbitrary, but rather reveals the organization of the brain prior to the accident. Assumptions and methods have become extremely sophisticated over the past 20 years, so the patient-based approach remains one of the major sources of information in cognitive neuroscience (see Rapp 2001, for a new textbook devoted to just this: what deficits reveal about the human mind). One of the landmarks of this approach is its power to reveal functional and neural specificity, often referred to as 'modularity'. It is this phenomenon of functional specificity that is illustrated here in the recognition of facial expressions of emotion relative to the recognition of the identity of the face.

The neural segregation of facial emotion and facial identity that emerges from lesion studies has been confirmed by functional brain imaging. For example, Sergent and collaborators (1994) have shown that different brain regions are involved in the analysis of identity and expression by measuring variations of regional cerebral blood flow in normal brains with the positron emission tomography (PET) technique.

The PET technique and the functional magnetic resonance imaging (fMRI) technique are two recent tools that are most often used today for studying the functioning of *normal* brains. PET is more invasive than fMRI because it involves the injection of radioactive substance. However, PET scanning is less noisy, and hence more appropriate for studying music perception. Fortunately, new means have been developed to facilitate auditory testing in the MRI machine. One such means consists in clustering volume acquisition (and hence noise) by acquiring all images of the brain in rapid succession for about 1 second, every 10 seconds, and so providing a period of quiet during which the stimuli may be presented uninterrupted and uncontaminated

by cortical responses to the noise (e.g. Edmister *et al.* 1999). Both PET and fMRI are valuable neuropsychological tools because they allow the measurement of variations of cerebral blood flow in three dimensions in the brain of a participant who is performing a task. Local variations in blood flow, or oxygen consumption, respond to different needs at any given time during the performance. However, measurements of cerebral blood flow changes are slow relative to the speed of information processing—requiring 2–3 minutes or 4–5 seconds with PET and fMRI scanning methods, respectively. All measures are taken after transformation of the magnetic resonance images of each individual into a standardized stereotaxic space. This manipulation eliminates any differences attributable to global differences in overall brain size or shape.

Using the PET technique with visual presentations of faces, Sergent and her co-workers were able to show that the processing of facial identity is associated with the ventromesial region of the right hemisphere (i.e. the fusiform gyrus and the lingual gyrus) and the parahippocampal gyrus, whereas facial emotion (covering happiness, sadness, fear, disgust) mostly engages the lateral part of the right occipital cortex and the cingulate gyrus (located in the limbic system). The regions are indicated in Fig. 5.1. Similarly, in the monkey, cells selective for facial identity and expression are found in different brain areas; cells which respond to face identity tend to be located in the inferior temporal gyrus, whereas cells whose responses are modulated preferentially by expression are located within the monkey's superior temporal sulcus (e.g. Hasselmo *et al.* 1989).

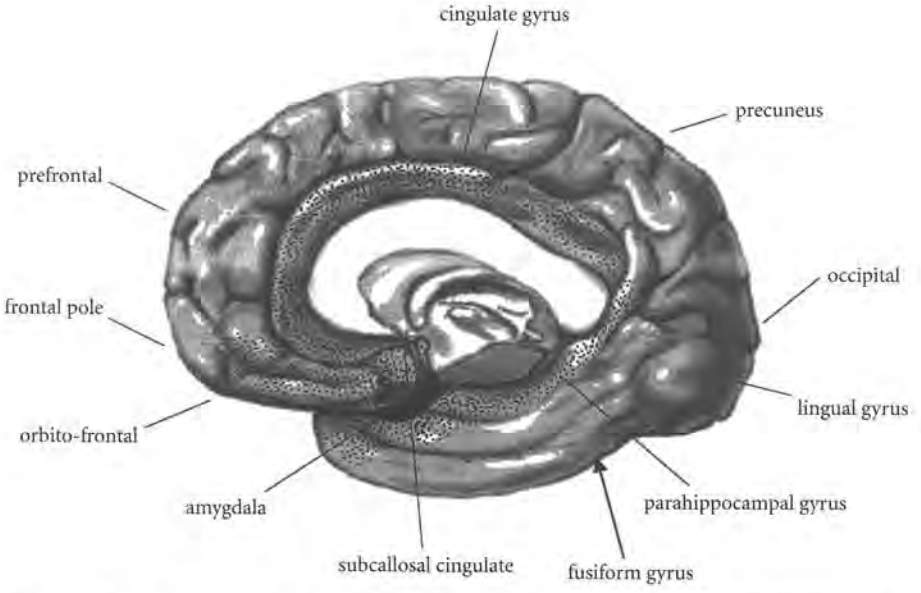



Figure 5.1 A medial inside view of the brain. The stippled area is the classic limbic system. The fusiform gyrus is hidden inside the temporal lobe. The amygdala, represented in black, is found deep inside the medial part of the temporal lobe.

### 5.1.3 Facial emotions: the role of the amygdala

We have seen that the processing of facial emotions is isolable in the brain. This does not mean facial emotions are subserved by a unique neural pathway. Rather, neural systems appear specialized for the recognition of specific emotions. The clearest evidence comes from studies of patients with damage to the amygdala, a brain structure long thought to play an important role in emotion. The amygdala is located in the limbic system (see Fig. 5.1). This system, named 'le grand lobe limbique' by Broca (1878), corresponds to subcortical structures that appeared early in evolution and remained relatively stable across species. The limbic system was designated as the substrate of emotions by Broca over a century ago, and later also by Papez (1937). Since then, the concept has gradually developed to include many more regions, both cortical and subcortical (Damasio 1994; LeDoux 1996). Emotions are no longer confined to the functioning of the limbic system, although the latter retains its fundamental role in emotion.

In the limbic system, the amygdala has been identified as the brain structure most implicated in fear. Bilateral amygdala damage impairs the recognition of unpleasant emotions, especially fear, in facial expressions (Adolphs *et al.* 1994, 1995; Broks *et al.* 1998; Calder *et al.*, 1996; Young *et al.* 1995); but not always (see Hamann *et al.* 1996). Such patients can see happiness and sadness in faces, but have severe problems with fearful expressions.

Additional support for the selective involvement of the amygdala in perceiving fear in facial expressions has been gathered in recent functional imaging studies of normal brains. Increased neural activity has been observed within the amygdala when subjects are viewing facial expressions of fear as opposed to happiness (Breiter *et al.* 1996; Morris *et al.* 1996; Whalen *et al.* 1998), disgust (Phillips *et al.* 1997), and anger (Blair *et al.* 1999).

The involvement of the amygdala is not limited to facial expressions of fear. The amygdala also plays a key role in conditioned fear in both humans and rats (Bechara *et al.* 1995; LeDoux 1996), as well as in fear-related behaviours such as uncontrollable rage reactions (Davis 1992). Moreover, direct stimulation of the human amygdala with in-depth electrodes evokes similar reactions (Halgren *et al.* 1978), further suggesting that the amygdala is critical to the processing of a class of emotions, namely those related to threat and danger.

The amygdala is strongly connected to the autonomic nervous system, which induces changes in physiological measures such as heart rate, blood flow, and respiration. This connectivity, coupled with the selectivity, of the emotional response, suggests that the amygdala is appropriately located in the brain to enable reactions to threat and danger, without the necessity of higher-level processing. The idea is that an expression of fear on the face of a conspecific may trigger an automatic and quick response to potential danger via the amygdala.

The role of the amygdala is highly selective. It does not participate in the evaluation of all negative emotions expressed by faces. For instance, the amygdala is not as actively engaged in perceiving disgust. Disgust rather involves the basal ganglia and the anterior insula. Increased activation of the insula has been observed in a functional brain-imagery study (Phillips *et al.* 1997), while impairments in disgust have been reported in patients with Huntington's disease, which is a neurological disorder that initially affects the basal ganglia (Gray *et al.* 1997; Sprengelmeyer *et al.* 1996). Disgust is another distinct

facial expression that has clear survival value, as does fear. Recognition of disgust allows avoiding contamination by offensive stimuli, and, in particular, the avoidance of ingestion of potentially harmful food. The neural response to facial expressions of disgust appears closely related to another, but related, sensory organ, namely taste. In humans, the anterior insula is particularly activated by aversive tasting (e.g. Zald *et al.* 1998). Altogether, the localization of the neural structures involved in disgust confirms common sense—that disgust is closely related to the perception of unpleasant tastes and smells.

The major implication of the relative engagement of the amygdala in perceiving fear relative to disgust in facial expressions is that it raises the possibility that different basic emotions have different underlying neural substrates. Basic emotions serve important and distinct functions, and the brain appears to be organized so as to respond to these distinct functions with distinct neural systems.

### 5.1.4 Facial expressions: the contribution of the cortical systems

The cortical systems are the neural structures that are evolutionarily more recent, and are particularly developed in the human brain. These cortical systems involve the occipital, temporal, parietal, and frontal cortex (Fig. 5.2). Another characteristic of these cortical structures is that they exhibit functional specialization within and across the two cerebral hemispheres. For example, it is well established that regions of the left hemisphere of the brain are much more essential for speech comprehension and expression than the right side of the brain.

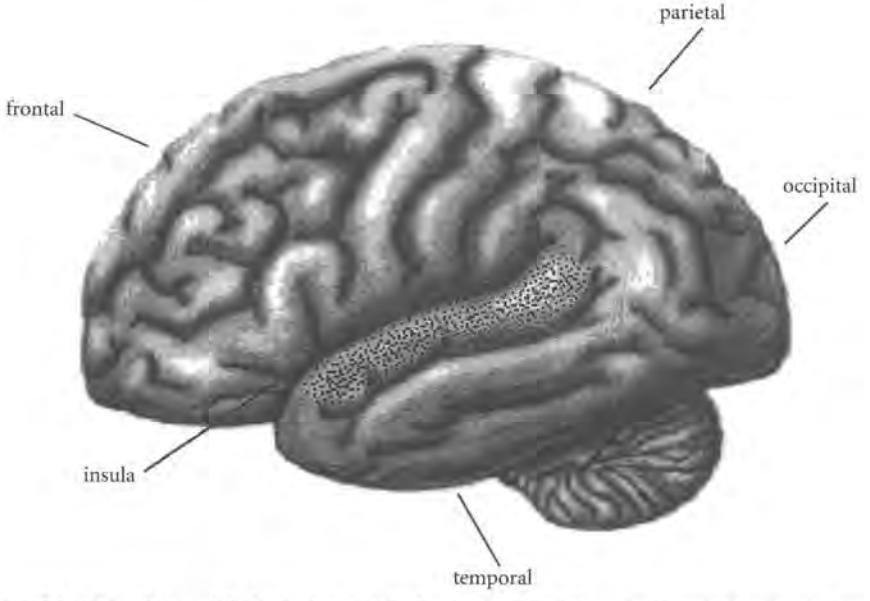



**Figure 5.2** A lateral view of the brain showing the four lobes of the left cerebral hemisphere in which several structures of interest are indicated. The stippled area corresponds to the superior temporal gyrus (see text for explanation).

With respect to emotions, there is currently a debate involving two alternative views. One view is that the right hemisphere is specialized for processing *all* emotions, by containing all 'modules' for non-verbal affect computation (e.g. Bowers *et al.* 1993). This position is known as the 'right hemisphere hypothesis'. The opposing view is the 'valence hypothesis' according to which the right hemisphere is more involved in negative emotions and the left hemisphere more engaged in positive emotions (e.g. Davidson 1992). Both hypotheses have been discussed with regard to the expression and recognition of emotion, as well as the experience of emotion.

The first observations related to the emotional competence of each cerebral hemisphere go back to Jackson (1878), who noticed that emotional language is often preserved in aphasic patients. He attributed the source of the preserved affective utterances to the contribution of the right hemisphere. This dissociation between propositional and affective language has often been reported since. However, it is difficult to reconcile with another classic pattern in clinical neurology that identifies each hemisphere with a distinct affective style. A lesion in the right hemisphere often produces indifference (anosognosia) and a tendency to joke. In contrast, an injury in the left hemisphere often leads to depressive–catastrophic states. The former is obviously less comprehensible than the latter as a reaction to the brain accident. The two emotional modes of responding have been associated with differences in 'affective styles' of the cerebral hemispheres. The left and right frontal cortex would be mediating approach and avoidance, respectively (for a recent review, see Davidson & Irwin 1999).

The valence hypothesis accounts more convincingly for expression and experience of emotions than for perception of emotions. For the latter, the bulk of the evidence seems to favour the right hemisphere hypothesis. In the study of normal adults' perceptual asymmetries, a left visual field (i.e. right hemisphere) superiority is typically found for both positive and negative expressions (see Best *et al.* 1994, for a recent review and additional data on adults' perception of infants' emotional expressions).

Perceptual asymmetries can be easily measured in normal right-handed subjects, in all modalities. In the visual modality, which is used for face presentation, facial expressions are typically and quickly presented to the left or to the right of a central fixation point. A left visual field advantage is inferred if the subjects process more accurately or more quickly the facial expressions presented on the left side of the screen than the expressions presented on the right side. Since the left visual field is directly projected to the right cerebral hemisphere, the advantage is taken to reflect a right hemispheric superiority for the task under study. The same logic applies to the auditory modality with which we will be concerned in the next section. This technique, often referred to as *dichotic listening*, consists of presenting different auditory stimuli in each ear, simultaneously. A right ear advantage is interpreted as a left hemispheric superiority, because it is well established that the information that enters in one ear reaches mostly the primary auditory cortex of the opposite hemisphere. It should be noted that although laterality measures are convenient and non-invasive tools for testing hemispheric specialization, they are less used nowadays because they have low sensitivity and low reliability, particularly in the auditory modality.

Lesion studies are compatible with both the right hemisphere hypothesis and the valence hypothesis. In a recent study involving a large sample of brain-damaged patients, Adolphs and his collaborators (1996) found that lesions of the right hemisphere led to impairments in recognizing negative emotions, such as fear and sadness, but spared

recognition of happy expressions. However, there are fewer positive than negative emotions regarded as basic emotions. Positive emotions, such as happiness, may simply be easier to distinguish from negative expressions than are negative emotions from each other. For example, a single prototypical feature (smile) distinguishes happiness from all other emotions.

Although the nature of each hemispheric contribution to the perception of facial expression remains to be determined, the presence of hemispheric advantages highlights the contribution of non-limbic cortical structures. It is plausible that the cortical contribution is related to both the demands of the perceptual analysis of the input (e.g. detailed facial features) and the contribution of memory, such as factual knowledge about emotions.

### 5.1.5 Facial expressions: a lesson and challenge

We have seen that recognition of facial expressions is the result of an adapted behavioural system that is implemented in a distributed neural network involving the limbic system, the amygdala in particular, but that it is not limited to this system. Recent research suggests additional and differential implication of the left and right prefrontal cortex, the anterior cingulate (see Fig. 5.1), and the insular cortex (see Fig. 5.2). Each of these regions appears to play a distinct function in emotional appraisal, with different distributed networks underlying different emotions such as fear and disgust.

Just as cognitive neuroscience has demonstrated the necessity to fractionate global functions, such as face processing or music recognition (Rapp 2001), into more elementary constituents whose neural substrates can be identified, so too modern research in the neuroscience of emotions suggests fractionation of affective processes. This conclusion highlights what is often not obvious to psychologists. It shows why the study of neural correlates in general, and brain localization in particular, is of importance. Examination of brain correlates can shed light on more general assumptions, by requiring that complex brain functions be decomposed into simpler processes so that simple components can be localized anatomically and studied in relative isolation. This fractionation into elementary localized mechanisms can then serve to test current models of emotional functioning or contribute to the building of new models.

## 5.2 Neural correlates of musical emotions

We have seen how facial emotions have become a tractable problem in the neurosciences. The central question is now to examine whether the neuropsychological study of musical emotions will hold similar promises.

### 5.2.1 Which musical emotions and what for?

Like facial expressions, certain musical emotions can be categorized as happiness, sadness, anger, and fear. The happy and sad emotional tones tend to be among the easiest ones to communicate in music (e.g. Gabrielsson & Juslin 1996; Krumhansl 1997). Like facial expressions, these emotions might be expressed by similar structural

features across musical styles and cultures. But the claim cannot be made forcefully, since the search of universals in the expression of musical emotions has just begun (Balkwill & Thompson 1999; but see also Juslin, this volume). In contrast, there is increasing evidence that musical emotions are quickly and easily perceived by members of the same culture. For example, ordinary adult listeners need less than a quarter of a second of music (e.g. one chord or a few notes) to reliably distinguish the tone of the whole musical excerpt as happy or sad (Peretz *et al.* 1998a). Moreover, and more generally, we were struck by the systematicity and reliability of the judgment data collected with ordinary listeners. Emotional judgments exhibit a high degree of consistency, suggesting that perception of emotions in music is natural and effortless for the large majority of listeners.

*Developmental perspectives*

Perception of emotion in music is also a skill that appears early in life. Recent empirical research suggests that music plays an important role in emotion regulation and emotional communication between care-givers and infants (e.g. Trehub & Trainor 1999). From birth, care-givers around the world sing to their infants, with the intuition (or instinct?) that music has the power to regulate the infant's state (e.g. comforting) or the quality of interaction (e.g. attention getting). The care-givers nicely mirror infants' perceptual abilities by singing more slowly, at higher pitch, with exaggerated rhythm, and in a more loving or emotionally engaging manner than when singing alone (e.g. Trainor *et al.* 1997). Responsiveness to such infant-directed singing appears inborn. Two-day-old hearing infants, born from congenitally deaf parents (who sign and do not sing or speak), prefer infant-directed singing to adult-directed singing (Masataka 1999).

Care-givers also speak to infants in a special, singing manner, so-called 'baby talk' or 'motherese' (see Bunt & Pavlicevic, this volume). Nevertheless, infants seem to discriminate between infant-directed speech and infant-directed music. Trehub (2001) reports a study in which 6-month-old infants viewed videotaped performances of their own mothers (recorded while singing or speaking to their infants). The infants showed more sustained attention to their mothers' singing episodes than to their speaking episodes. Infants were 'hypnotized' by these sung performances, remaining 'glued' to the monitor for extended periods. Mothers' speaking was not as engaging as their singing. The fact that infants are particularly responsive to the emotional messages of music directed toward them suggests an adaptive value of maternal singing. Mother–infant emotional communication is crucial to survival.

It is also remarkable how skilled the young children are in emotional perception. From the age of 3 years, they show the ability to recognize happiness in elaborate, art music of their culture, and by the age of 6 years they show adult-like abilities to identify sadness, fear, and anger in music (Cunningham & Sterling 1988; Terwogt & van Grinsven 1988, 1991; but for later emergence, see Dolgin & Adelson 1990). Furthermore, the child competence can be related to abstraction of specific musical features. At 5 years of age, children are able to discriminate between happy and sad excerpts by relying on tempo differences (fast vs. slow). At 6, children show evidence of using both tempo and mode (major vs. minor) like adults do. Although the results suggest that sensitivity

to tempo precedes sensitivity to mode, it is remarkable that by the age of 6 years, children show full knowledge of the rules that govern the happy–sad character of surrounding music. This ability seems to remain generally unchanged over a lifetime (Dalla Bella *et al.* 2001; Gerardi & Gerken 1995; Gregory *et al.* 1996; Kastner & Crowder 1990; Kratus 1993). Therefore, with consistency and precociousness, musical emotions resemble facial emotions.

*Evolutionary perspectives*

Musical emotions also seem to respond to important needs of the organism, even though considerable controversy surrounds the questions of the origin, biological significance, and function of music (e.g. Pinker 1997). In other words, music may have adaptive significance. Two main evolutionary explanations have been offered. The initial account was provided by Darwin himself (1871), who proposed that music serves to attract sexual partners. This view has been recently revived by Miller (2000), who reminds us that music making is still a young male trait. However, the dominant view of the adaptive value of music lies at the group level rather than at the individual level, with music helping to promote group cohesion. The initial step for this bonding effect of music could be the mother–infant interactive pattern created through maternal singing. The individual- and group-level roles attributed to music do not need to be mutually exclusive. As pointed out by Kogan (1994), individuals taking the lead in gatherings by virtue of their musical and dance prowess can achieve leadership status in the group, a factor that contributes to reproductive success (for other formulations of these evolutionary issues, see Wallin *et al.* 2000).

In support of the contention that music has an adaptive value, particularly for the group, is the fact that music possesses two design features that reflect an intrinsic role in communion (as opposed to communication, which is the key function of speech). Pitch intervals allow harmonious voice blending when sounding together, and temporal regularity facilitates motor synchronicity. These two musical features are highly effective in promoting simultaneous singing and dancing (Brown 2000). The design is specific to music; it is certainly not shared with speech, which requires individuality for its intelligibility. These special features fit with the important criterion, discussed by Buss and collaborators (1998), that for a system to qualify as 'adaptive' it must offer effective solutions to a problem. The system must have a 'special design'. The bonding problem in the case of music is to override 'selfish genes' for the benefit of the group.

For neuropsychologists, this evolutionary perspective on music is essential. As mentioned earlier, the fact that musical emotions are viewed as adaptive responses that can be aroused similarly in every human being is best conceived as the product of neural structures that are specialized for their computation. Furthermore, if biologically important, these emotions can even be seen as *reflexes* in their operation. That is, musical emotions would occur with rapid onset, through automatic appraisal, and with involuntary changes in physiological and behavioural responses. This conception would correspond to the fact that 'we often experience emotions as happening to us, not as chosen by us. We do not simply decide when to have or not have a particular emotion' (Ekman 1994, p. 17).

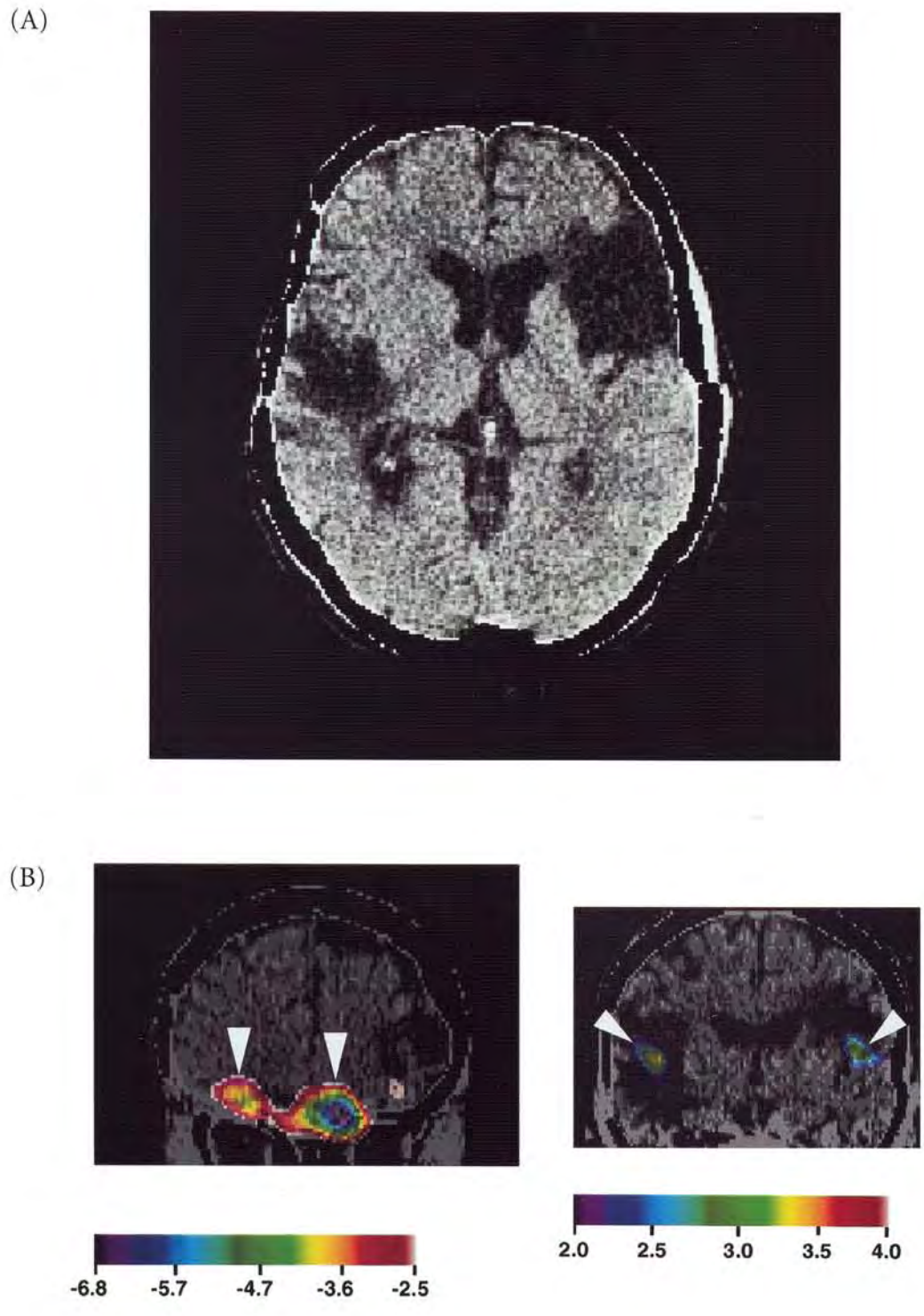
### 5.2.2 Musical emotions: brain specialization

Support for the existence of specialized neural networks for music processing is presently compelling (for reviews, see Peretz 2001*a*, 2001*b*). However, the evidence relies on the detailed study of music perception and memory, that is on the processing of music as a special auditory structure, not as an emotional medium. Emotional appreciation of music is a new research avenue in neuropsychology. Hence, the data are scarce, albeit consistent and promising.

First of all, musical emotions appear isolable in the human brain. By isolable, I mean that musical emotions can be selectively either lost or spared, usually as a consequence of a brain accident occurring at an adult age. There are a number of case reports that describe musicians who retained their musical skills, but who complained that they lost interest in music because it sounded 'flat' or without emotion (Mazzoni *et al.* 1993; Mazzuchi *et al.* 1982). Unfortunately, the reports are anecdotal. No attention was paid to the emotional complaints of the two patients because of 'the highly subjective nature of the symptom concerned' (Mazzuchi *et al.* 1982, p. 646) or because 'disturbances of this type are, unfortunately, difficult to view objectively because of their highly subjective nature' (Mazzoni *et al.* 1993, p. 322).

However, recent progress made in affective neuroscience in general, and in the study of facial expressions in particular, has encouraged us to go beyond anecdotal reports and study emotional responses in the laboratory. In doing so, we have been able to discover mirror cases, that is, patients who manifest selective sparing of musical emotions. The case of one patient, IR, who suffers from longstanding bilateral brain damage to the auditory cortex (see panel A in Fig. 5.3) is given here. The case is remarkable in that 15 years after the brain damage, IR still experiences severe difficulties with music while her language abilities and her general intellectual and memory abilities are normal (Peretz *et al.* 1997; Peretz & Gagnon 1999). Despite her severe musical deficits, she reports that she still enjoys music, listening regularly to prerecorded music in the car. This possible preservation of emotions for music appeared to us as paradoxical, given the severity of IR's perceptual disorders with music. Therefore, we decided to study her emotional judgments in detail. Three independent studies have been completed with her that will be summarized here.

The goal of the first study (Peretz & Gagnon 1999) was to verify experimentally IR's insights. To do so, IR was presented with popular melodies, half of which are familiar in Québec where the study was conducted. IR performed two judgments: in the emotional judgments, she classified the melodies as 'happy' or 'sad'. In the non-emotional 'control' task, she performed a 'familiar–unfamiliar' classification task. The results confirm her report. She is able to classify melodies as 'happy' and 'sad' as normal controls do, while she fails to reach normal performance in classifying these same melodies for familiarity. For example, when presented with the melody of 'Happy Birthday' without its lyrics, IR would say 'I don't know that tune but it sounds happy'. Similarly, CN, another patient with severe recognition problems for melodies that were once highly familiar to her (Peretz 1996), had an interesting reaction. When listening to the famous adagio of Albinoni taken from her own record collection, CN first said that she had never heard that piece before. Suddenly, she said: 'it makes me feel sad . . . the feeling makes me think of Albinoni's adagio' (R. Kolinsky, personal communication). CN, who was totally




CT scan of I.R.'s brain, transformed into the standardized stereotaxic space. For a more detailed caption see Fig. 5.3 on p.117.

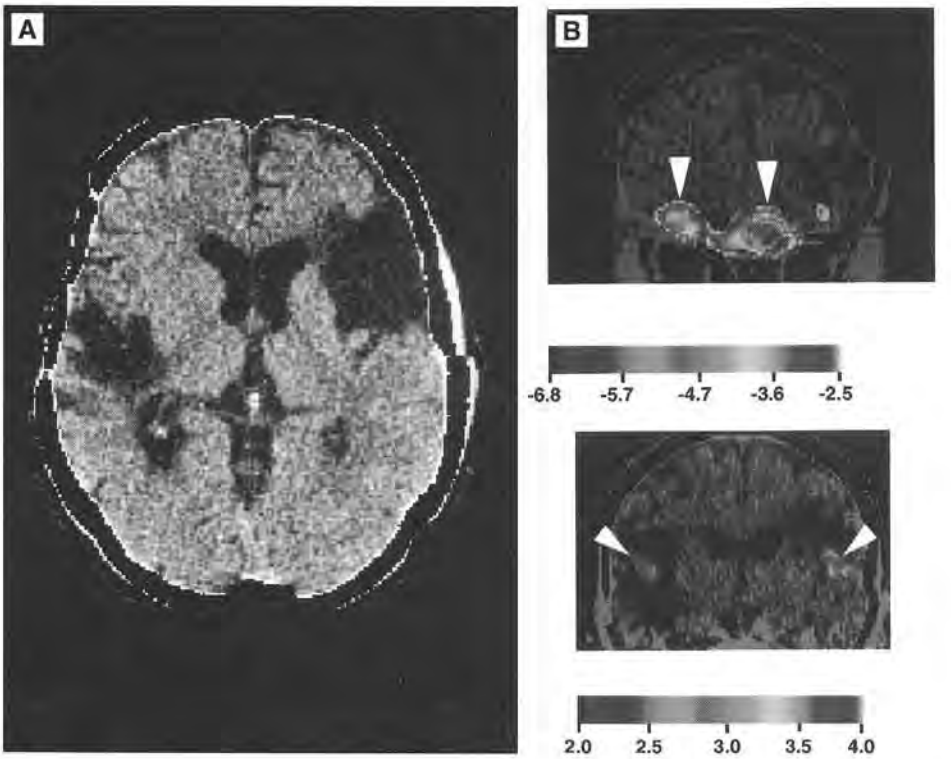



**Figure 5.3** CT scan of IR's brain, transformed into standardized stereotaxic space. The right side of the scan corresponds to the right side of the brain. (A) Horizontal slice showing the differing locations of temporal and frontal lobe damage (i.e. darker areas correspond to damaged tissue) in the left and right hemispheres, and (B) coronal slices showing PET data from a group of normal non-musicians (Blood *et al.* 1999*a*) superimposed on the CT scan of IR's brain. The top panel shows areas whose cerebral blood flow (CBF) changes are related to pleasantness judgments; the arrowheads show the subcallosal area and the left orbitofrontal areas. None of the paralimbic or neocortical regions recruited by the dissonance manipulation in normals overlapped with the lesions in IR's brain. The panel below illustrates areas of significant CBF increase in the subtraction of a highly consonant stimulus condition with a matched noise baseline condition. This illustrates the areas of activity within the superior temporal gyri bilaterally (indicated by arrowheads). Note that these regions overlapped significantly with the lesions in IR's brain. The *t*-statistic ranges are coded by the colour scale below the corresponding panel. (Adapted from Peretz *et al.* 2001.) (See also colour section).

unable to recognize melodies that were highly familiar to her before the brain accident, was able to do so via her emotional responses. This is probably not an isolated finding. However, we did not pursue our work in that direction. Rather, we wished to establish more firmly the dissociation observed in IR between emotion and perception.

In a follow-up study of IR (Peretz *et al.* 1998a), we were interested in explaining how she could derive the emotional tone from melodies, while her ability to recognize and discriminate melodies is so limited. One intriguing possibility, derived from the animal

work of LeDoux (1996) and from the theoretical position of Zajonc (1984), is that emotions need no cortical mediation. That is, basic emotional responses might function like subcortical reflexes. These emotional responses would be immediate but superficial. They would not require the additional time taken by elaborate processing of the signal in cortical structures. Since IR's lesions are vast, but limited to cortical structures, perhaps her intact limbic system is responsible for her spared ability to distinguish happy from sad music.

To assess the nature of this spared emotional pathway in IR, and to further assess its independence from the rest of her processing system for music, we tested her in a series of experiments employing the same set of excerpts taken from the classical repertoire (e.g. Albinoni's adagio). These were selected to unambiguously convey a 'happy' or 'sad' tone, and were presented under various transformations and with different task demands. At first sight, IR's results seemed to conform to the separate pathways view. As with normals, IR was able to use both the mode and the tempo characteristics to derive the 'happy' or 'sad' tone of the music. Moreover, as with normals, IR's judgments were immediate. In contrast with the relatively sophisticated emotional processing of the musical excerpts, IR showed impaired performance in her non-emotional assessment. Specifically, she performed well below normal controls in the discrimination of these musical excerpts in a 'same–different' classification task that was very easy for controls. Likewise, she failed to detect most errors purposely inserted on either the pitch or time dimension of the musical excerpts, although these mistakes were obvious to control subjects. Thus, IR's pattern of results suggests that perceptual analysis of the musical input is intact for emotional purposes and defective for non-emotional ones.

The relative independence of the emotional pathway from non-emotional processing of music is consistent with previous data of ours that were obtained with neurologically intact students. We showed in a series of experiments that exposure effects differed for recognition and liking judgments (Peretz *et al.* 1998b). Similarly, we showed that IR can process melodies for emotion recognition, but not for identity recognition (Peretz & Gagnon 1999) or discrimination (Peretz *et al.* 1998a). However, IR's spared emotional processing cannot be easily accommodated with the isolated functioning of a subcortical pathway. IR is able to use a change of mode, from major to minor and vice versa, in elaborate musical structures. This ability can hardly be conceived as the result of the operation of a primitive subcortical system. Furthermore, as we will see later, there is presently little support for the involvement of a subcortical 'short cut' in emotional processing of music. Rather, musical emotions seem to require cortical mediation. Therefore, all we can conclude from these studies with IR and CN is that severe deficits in perceptual processing and memorization of music can leave emotional judgments of music unimpaired. Such a spectacular isolation of emotional judgments of music suggests the presence of a specialized cortical arrangement for it.

### 5.2.3 Musical emotions: hemispheric specialization

As mentioned above, there are indications that emotional judgments require cortical mediation. The participation of cortical structures seems to depend on the cerebral hemisphere concerned, as observed for facial expressions. Likewise, the available data are divided between the two current hypotheses regarding the nature of this specialization.

Support for the valence hypothesis has been obtained in two recent studies measuring brain electrical activity—the electroencephalogram (EEG)—from scalp electrodes in normal listeners (Schmidt & Trainor 2001; Tsang *et al.* 2001).

EEG methods are non-invasive measurements of the ongoing electrical activity that can be captured by electrodes located on the scalp of the listener. Various measurements are made possible with the recent increases in computer power. The EEG method used in the studies by Schmidt and Tsang that are cited here consists of recording the EEG variations over extended periods of time, in the order of 60 seconds, while the subject is listening to a piece of music. The EEG signal is then decomposed into particular frequency bands, such as the alpha bands taken to reflect variations in activation. Another EEG method that is increasingly exploited in neuropsychology is the evoked response potential (ERP) (for a review of its use with regard to music, see Besson 1997). This method consists of averaging EEG signals time-locked to a particular category of events. The advantage of the ERP is that it is an on-line measure of processing that has a high temporal resolution window, of the order of a few milliseconds. By increasing the number of electrodes over the scalp, and by exploiting sophisticated algorithms to locate the neural source responsible for the observed response, it is now possible to obtain information about brain localization.

Asymmetrical frontal EEG activity has been found to distinguish the valence of musical excerpts taken from the classical repertoire. Subjects exhibit greater relative left frontal activity to music expressing joy and happiness and greater relative right frontal EEG activity to music expressing fear and sadness (Schmidt & Trainor 2001). Subsequently, the same group of researchers (Tsang *et al.* 2001), using the same EEG technique, were able to replicate and relate the EEG findings to musical structure, using our set of musical excerpts used with IR (Peretz *et al.* 1998a). They observed that changes in both tempo and mode in the happier direction resulted in greater relative left frontal activation, whereas changes in both tempo and mode in the sadder direction resulted in greater relative right frontal activation.

In a similar vein, Gagnon and Peretz (2000) have observed ear asymmetries in normal listeners that are compatible with the notion that the left hemisphere contributes more to the perception of positive emotions and the right hemisphere more to negative emotions in listeners. Non-musicians were required to classify tonal and atonal melodies as pleasant and unpleasant in one condition. In a non-affective 'control' condition, the same subjects were required to judge if the melodies sounded correct or not. Listeners exhibited a left-ear superiority effect, taken to reflect the predominance of the right hemisphere, when judging atonal melodies as unpleasant while they displayed a tendency towards the right ear when judging tonal melodies as pleasant. This pattern of ear asymmetries was specific to emotional judgments since a different pattern was obtained when the same melodies had to be classified for 'correctness'.

However, the valence account of cerebral asymmetries is not systematically observed. Two studies are more supportive of the right hemisphere hypothesis. The first one measured ear asymmetries in normal subjects who were judging minor and major melodies as eliciting positive or negative emotions (Bryden *et al.* 1982). In that study, an overall left ear (i.e. right hemisphere) advantage was observed across positive and negative emotional judgments. The second study measured changes in regional cerebral blood flow

with the PET technique as a function of pleasantness judgments for chord sequences varying in degree of dissonance (Blood *et al.* 1999*a*). In that study, the activated neural structures by pleasantness judgments were found primarily in the right hemisphere. Moreover, reciprocal activation was found between particular neural structures and the valence of the musical stimulus. However, these reciprocal activations did not occur between hemispheres but rather between the parahippocampal and frontal regions.

### 5.2.4 Musical emotions: contribution of distributed neural structures

Up to now, we have reviewed data suggesting that emotional appreciation of music is subserved by a distinct neural pathway that probably requires cortical mediation. Cortical mediation appears to involve primarily the right hemispheric structures, with a possible contribution of the left frontal regions, depending on the valence of the perceived emotion. More specific localization suggestions have been made by Blood and collaborators (1999*a*), using dissonance as the emotion-eliciting stimulus and PET scanning as the brain-recording measure.

In that PET study, regional cerebral blood flow changes were obtained in several distinct paralimbic and neocortical regions of the brain. Most activated areas depend on the valence of emotion. Activity in the right parahippocampal gyrus and precuneus regions is associated with increasing dissonance, whereas activity in the bilateral orbitofrontal, medial subcallosal cingulate, and right frontal polar cortex correlates with decreasing dissonance (or increasing consonance). In contrast, activity in the superior temporal cortices is observed bilaterally, independently of dissonance level. All implicated regions are indicated in Fig. 5.1.

One powerful means to determine which neural region is essential in a given function is to compare brain-imaging data with brain-lesion data. To this aim, we tested IR, our case with bilateral lesions to the auditory temporal cortex, with consonant and dissonant music (Peretz *et al.* 2001). In fact, IR is unable to distinguish consonant from dissonant versions of the same musical excerpts in pleasantness judgments, although she is able to judge the 'happy–sad' character of the music. This is a striking disorder. Preference of consonant over dissonant music is widely shared by ordinary listeners, beginning with 4-month-old infants (see, for example, Trainor & Heinmiller 1998; Zentner & Kagan 1996). IR has lost this ability. Localization of the responsible damage cannot be easily determined in her brain. As can be seen in Fig. 5.3, IR's lesions are extensive and asymmetrically localized in her cerebral hemispheres. Therefore, it is not clear which of these damaged areas can be taken as responsible for the observed deficit in responding to dissonance. By combining information about IR's lesion localization with normal cerebral regions that are implicated in judging dissonance, we might be able to isolate the critical neural regions.

Therefore, we co-registered images of IR's brain with the activated brain regions found in normal brains. We found that the complex network involving paralimbic and neocortical regions associated with the affective evaluation of dissonance did not overlap with IR's lesions (see top panel B in Fig. 5.3). In contrast, significant overlap was found between IR's lesions and the pattern of activity elicited by all musical stimuli, irrespective of dissonance level (see bottom panel B in Fig. 5.3). The damaged areas in IR's brain coincided with the regions located in the superior temporal gyri that were

involved in the perceptual analysis, including those related to dissonance, of the chord sequences in normal brains. Therefore, we can infer that the critical networks for dissonance appraisal are located bilaterally in the superior temporal gyri. Furthermore, we can ascribe IR's indifference to dissonance to a perceptual defect.

That IR's problem with dissonance lies at the perceptual level is consistent with a number of independent observations. First, in general, IR's emotional processes appear fairly normal. Second, the neural network specifically associated with emotional evaluation of dissonance, which involves the subcortical–frontal areas, seems largely spared in IR's brain. Third, IR was unable to discriminate the different versions of the same musical sequence that only differed in degrees of dissonance. Therefore, a plausible account of IR's lack of responsiveness to dissonance is simply that she does not perceive it. Had she perceived it correctly, she would have judged its emotional valence accordingly. If this interpretation is correct, affective responses to dissonance are mediated via an obligatory cortical perceptual relay.

In summary, the present findings suggest a particular functional architecture underlying emotional interpretation of music. Taken together, the results suggest that the musical input first reaches the superior temporal gyri, where perceptual organization is taking place. The perceptual output is then relayed to emotional systems in the paralimbic structures or more frontal areas, depending on its valence. This two-stage model suggests that emotion and perception are not taking place along two parallel and independent pathways as some models (LeDoux 2000; Zajonc 1984) posit. However, caution is required because this model only derives from the study of a single emotional dimension, relating dissonance to unpleasant judgments.

## 5.3 Fractionation of musical emotions

As seen with facial expressions, neuropsychological methods are most productive for dissecting emotions and their neural substrates into elementary components. Unfortunately, research has just begun with musical emotions. Hence, the power of the methods cannot be fully appreciated in the musical domain. In what follows, I will raise a few questions that, I believe, are important to address using neuropsychological methods. Since these questions do not have an answer yet, their outlining should be regarded as guidelines for future research.

### 5.3.1 Single or multiple emotional systems for music?

In theory, lesion studies and functional brain imagery should provide evidence regarding the possibility that there is not a single but multiple emotional systems that can be aroused by music, depending on its content. Although parsimonious, the notion of a single emotional system is difficult to hold given the diversity of emotions and their differential functionality for the organism. This was illustrated earlier with the neuropsychological separation of fear from disgust with regard to facial expressions. The problem with musical emotions is that their functionality is not settled, and *a fortiori* the principles along which musical emotions may fractionate remain open. But neuropsychology may, in principle, provide relevant cues.

For example, there are indications that musical emotions can be studied in relative isolation from each other. The 'happy–sad' character of music can be studied independently from a valence ('pleasant–unpleasant') dimension. As shown in the study of IR, discrimination of happiness and sadness in music can be preserved in the presence of a lost sensitivity to dissonance. Although this dissociation has a perceptual rather than an emotional origin, it highlights one clear advantage of studying musical emotions over other domains, such as facial and vocal expressions. In music, it is easier to manipulate structural characteristics systematically so that the same music can express different emotions gradually. For example, tempo increase and decrease are easy to apply without altering other aspects of the musical structure (e.g. Gagnon & Peretz 2001, within certain limits). We can also manipulate two emotional dimensions orthogonally, as previously done with the same set of stimuli, by modifying mode (from major to minor and vice versa) and dissonance, independently (Peretz *et al.* 2001). There is such a close relationship between structural characteristics and certain emotional expressions in music that the domain lends itself to experimentation. This is hardly a new finding (see Gabrielsson & Lindström, this volume; Juslin, this volume). Yet, it is often overlooked in the literature.

Finally, musical emotions present interesting peculiarities. Sadness in music is a good example. It is pleasurable in music while perceived as unpleasant in most other channels. By studying the neural correlates of certain musical emotions, like sadness, the generalization and neuropsychological validity of various taxonomies of emotions can be advantageously tested. It may turn out that some musical emotions are unlike any other type of emotion. If that is the case, the evidence will in all likelihood come from neuropsychology, since the discipline is perfectly suited to reveal functional specificity.

### 5.3.2 Are emotional systems domain-specific?

As indirectly suggested, musical emotions might differ from other emotions in numerous aspects. However, there is as yet no theoretical or empirical reason for assuming such specificity. Inversely, musical emotions are frequently compared to vocal expressions (Juslin, this volume), with the idea that the two expressive channels may derive ontogenetically and phylogenetically from a unique and common mode of communication. In this context, it may seem curious that up to now I have only considered the advances made in the domain of facial expressions. The reason is simply that progress has been much slower in the domain of vocal expressions of emotions.

Perhaps vocal expressions of emotions have not yet received sufficient theoretical attention (but see Johnstone & Scherer 2000, for a recent summary of promising findings). Despite the obvious biological significance of vocal emotions, there is no consensus over an evolutionary-based theory of it. This lack of theoretical development, compared to, say, the domain of facial expressions, might in turn explain why the topic of vocal expressions has had a limited impact in the recent shaping of the neuropsychology of emotions. This is certainly not due to lack of interest. The literature is abundant, as summarized below.

#### *Vocal emotions: neural correlates*

Just as musical expressions of emotions can be conveyed by variations in pitch and articulation, emotions can be expressed by modulations of the tone of the voice.

Emotion expressions through vocal sounds can take two different forms in humans. It can be conveyed by the tone of voice in speaking or by non-verbal expressions such as in laughing, crying, and screaming. There has been a greater interest in the first category of vocal expressions, and it is usually referred to as *affective prosody*. However, there is as yet no empirical or theoretical basis for distinguishing these two classes of vocal expressions of emotions, and thus they will not be treated separately in what follows.

The most remarkable neuropsychological property of the system underlying both kinds of vocal expressions of emotions is its specificity. The neural system appears highly specialized, both functionally and neuroanatomically. Firstly, despite the obvious link between facial and vocal expression of emotions, emotional evaluation of the tone of voices is dissociable from the emotional evaluation of faces. Some patients can infer the emotional tone of vocal expressions but not of facial expressions, whereas others show the opposite pattern (e.g. Hornak *et al.* 1996). This dissociation is not due to differences related to the modality of input, because the patient's ability to process other aspects of faces and voices is spared. The separability of vocal from facial emotions is impressive, given that they typically signal emotions in a consistent manner at any given time. Yet, the two modes of expressions seem to be captured by independent neural systems.[1]

Secondly, within the speech signal, the recognition of vocal cues that are emotionally meaningful is clearly dissociable from the recognition of vocal cues that are semantically informative. As mentioned previously, Jackson (1878) was the first to notice the separability of affective prosody from propositional speech. The distinction has been made several times in perception as well. For example, in dichotic listening studies with normal adults, Ley and Bryden (1982) have found a left ear (right hemisphere) advantage for the perception of the emotional tone of voice, and a right ear (left hemisphere) superiority for the recognition of the actual content of the words from the same set of sentences. Similarly, patients who can no longer understand speech, but who are still able to infer the emotion conveyed by vocal cues, have been reported (e.g. Barrett *et al.* 1999). These findings show that emotional processing of speech signals can take different routes, one determined by semantic content and one determined by vocal form.

Finally, the tone of the voice is not analysed just for affective purposes. Prosodic cues can provide important person-specific information (e.g. identity, age, and sex) and linguistic, non-semantic information (i.e. differentiating a question from a statement). Once again, the use of these non-affective cues can dissociate from their emotional use. Brain lesions can interfere with the identification of vocal emotions of happiness, sadness, and fear and spare the identification of prosodic differences marking questions, exclamations, and assertions in spoken sentences (e.g. Heilman *et al.* 1984). Similarly, functional imaging of neurologically intact brains points to the involvement of distinct neural regions in the recognition of emotions and of the speaker's identity in spoken sentences (Imaizumi *et al.* 1997). Together, these findings highlight the fine tuning and specialization of the neural system underlying the emotional interpretation of vocal sounds.

[1] This conclusion requires more systematic work, realized in more comparable testing situations. Computer facilities allow the testing of participants with dynamic events where facial expressions and vocal emissions can be studied jointly. To my knowledge, such testing situations have not yet been exploited with brain-damaged patients or with functional brain imaging.

This highly specific system mostly recruits cortical structures of the right hemisphere. Right hemispheric lateralization has been frequently reported after the seminal work of Heilman and collaborators (1975) in brain-damaged patients (for a review, see Etcoff 1989). It has been observed in normal listeners tested either with laterality paradigms (e.g. Ley & Bryden 1982), as described above, or with functional brain-imaging techniques (e.g. George *et al.* 1996; but see Morris *et al.* 1999, for negative findings relative to hemispheric asymmetry). Within the right hemisphere, and at times in the left hemisphere as well, various cortical and subcortical structures have been identified as contributing to the emotional evaluation of vocal sounds. These structures are the temporal cortex, the insula, the basal ganglia, the caudate nucleus, and the ventral prefrontal cortex (for example, for recent brain-imaging data, see Morris *et al.* 1999, and for recent lesion data, see Hornak *et al.* 1996). Thus, the neural basis of vocal emotions appears rather distributed and complex. Specification of the role of each contributing structure remains a task for the future.

Nevertheless, it is worth noting that attempts to find evidence for the involvement of the amygdala have been negative (Adolphs & Tranel 1999; Anderson & Phelps 1998; but see also Scott *et al.* 1997). Patients with selective bilateral damage to the amygdala, who show evidence of impaired recognition of fear in faces, show preserved recognition of fearful voices. This lack of convergence in recruiting the amygdala across non-verbal communication channels in response to danger is consistent with the general conclusion that the vocal emotional system is subserved by distinct neural structures.

*Are vocal emotions separable from musical emotions?*

Because the neural organization underlying the recognition of vocal expression of emotions appears highly specialized, I find it difficult to conceive how it might also serve musical emotions. However, this is a highly speculative and controversial issue. Juslin (this volume) argues that if the neural circuitry used in the recognition of emotion in vocal expression is used also in music, it can be only for those aspects of music's expressiveness that are common to speech and musical performance, that is, cues like speed, intensity, and timbre. However, music's expressiveness does not derive solely from such cues, but also from intrinsic sources of emotion (cf. Sloboda & Juslin, this volume) having to do with the internal structure (e.g. harmonic progression), and so it would not be surprising to find that perception of musical emotions involves neural substrates over and above those involved in the perception of vocal expressions. Panksepp (1995) sees similarities in the emotional 'chills' evoked by sad music and those engendered by 'distress calls' of young children and animals. Yet, Blood and collaborators (1999*b*) suggested the implication of the septum, a region typically associated with pleasure, in emotional 'chills' for self-selected music. In any case, neuropsychological methods are ideally suited for comparing vocal and musical emotions. As a result of advances in neuroscientific procedures and improved digital-processing techniques, findings obtained in the coming years should prove especially informative for our understanding of emotional processing through the vocal and musical channel.

### 5.3.4 Is emotion recognition dissociable from emotion experience?

A crucial question that I have left for the end is to what extent all the evidence reviewed in prior sections is related to emotional experience. What was said above mainly derived from tasks requiring emotion recognition, rather than emotion experience. Thus, results may just reflect emotional knowledge, not felt emotions. In theory, all emotions, including fear, may be perceived, recognized, known, described, and mimicked, without being felt. I can see fear in the face displayed on the screen, I can recognize that the Albinoni's adagio expresses sadness in an exquisite manner. Yet, it may leave me cold, particularly in impoverished experimental settings that are of typical use in current research. In this perspective, the study of emotions remains 'cognitivist', contributing little to the building of affective neuroscience as a new discipline.

The debate around the separability of emotion recognition from emotion experience has a long history in the musical domain (cf. Cook & Dibben, this volume). Although the problem seems easy to conceptualize, it is difficult to solve empirically. One solution is to exploit more 'objective' or 'direct' measures of emotions. By objective measures, it is generally meant those measures that cannot be under the voluntary control of the subject. The measures typically pertain to the domain of psychophysiology. In fact, it is well established that emotions are accompanied by physiological changes that occur automatically without voluntary control. It takes talent or intensive training to control physiological expressions of emotions. Musical emotion is no exception. Music may be particularly powerful in eliciting such changes, particularly 'chills' or 'shivers' (e.g. Goldstein 1980). More commonly, music can elicit changes in heart beat, electrodermal responses, respiration, skin temperature, etc. (e.g. Krumhansl 1997). A review of physiological changes is provided by Scherer and Zentner (this volume; see also Bartlett 1999).

Physiological measures do not solve all problems, however. They have their own limitations. First, their 'objectivity' should not be equated with 'specificity'. Just because physiological changes occur without voluntary control this does not mean that they reveal emotions unequivocally. Physiological measures are notorious for their lack of selectivity in addition to their limited sensitivity (e.g. it is common to reject half the subjects because they are 'non-responders'). Physiological changes do not easily discriminate among emotions. For instance, different types of music (happy, sad, fearful) may all produce physiological changes, but the changes may be in the same direction compared with baseline (Krumhansl 1997; but see Nyklicek *et al.* 1997, for a recent example where physiological measures did discriminate among emotions). Nevertheless, physiological changes represent an interesting complement to other brain-related measurements of emotion. Convergence of measurements is highly desirable, probably more in the emotional domain than in any other field.

Does this mean that if we combine behavioural judgments with physiological measurements and both converge, we can safely conclude that the emotion under study was not only recognized but also experienced? Probably. However, the setting is costly in terms of sample size and of technological sophistication. Fortunately, there are other solutions worthy of investigation. One promising alternative is to study emotions *indirectly* or implicitly.

Musical emotions have been typically studied in a direct and explicit way, by asking listeners to judge, label, or categorize music in emotional terms, or by inviting performers to interpret a musical piece with a particular emotion. The impact of musical emotions is rarely assessed indirectly by the way it affects other tasks or behaviours, without awareness. There are different means to probe emotional judgments indirectly. One procedure that has been successfully used with facial expressions was to require a gender decision (e.g. 'Is this a face of a woman or of a man?', which can be difficult to answer when hair and accessories are removed from the picture). The key variable was that some of the faces displayed an emotion, such as fear or disgust, with various degrees of intensity. The indirect influence of the facial expression can then be measured in a graded fashion on both the timing of the gender decision and the resulting changes in cerebral blood flow as measured by functional brain imagery (for a review of their work done with this procedure, see Dolan & Morris 2000).

The advantage of indirect measurement is that emotional responding is studied as a spontaneous, automatic response of the organism. This is what emotions are: spontaneous responses that are difficult to disguise and that may occur without awareness. However, the question is to what extent these spontaneous responses arise from the operation of the same mechanisms as those implicated in emotional judgments or rather reflect the functioning of different systems. Curiously, this question of consciousness, which has been addressed in various domains such as memory and face recognition, has as yet attracted little attention in affective neuroscience.

My intuition, and experience as illustrated below, is that musical emotions typically occur without consciousness or willingness. For example, a few years ago, I felt obliged to accompany a friend to a Broadway show, featuring 'Mrs Butterfly'. I do not like those shows; I find them boring and caricatural. Yet, at the end of the show, I found myself in tears. The show was not of any better quality than any other one, I found. However, my disliking was probably a cultural response, whereas my emotional reaction was a biological reflex.

To summarize, it is possible, at least in theory, that emotions can be recognized without being experienced. To bypass this problem, one can take different routes. These include brain-related and physiological measurements but also consideration of *indirect* measures. There is an obvious need to go beyond conventional means of studying emotions.

## 5.4 Other issues related to musical emotions

There are a number of issues related to the neural correlates of musical emotions that have not been addressed in this chapter because of insufficient relevant data. One of these issues is the possible exploitation of *neurochemical correlates* of musical emotions. Neurochemicals are neurotransmitters and hormones that alter the response properties of sets of neurons. Music is apparently effective in eliciting such responses, as suggested by the action of the antagonists of endorphins (Goldstein 1980) and cortisol measures (Trehub 2001). The study of these neurochemicals would provide yet another neuropsychological avenue to better understanding of the nature and brain organization of musical emotions (for reviews, see Buck 1999; Panksepp 1998).

A completely different neurological condition that can help in understanding the nature of musical emotions is the study of individuals suffering from *autism* and *Williams' syndrome*. Both disorders are usually associated with intellectual deficiency and a relatively high level of musical proficiency (Don *et al.* 1999; Levitin & Bellugi 1998; Miller 1989). However, the two disorders diverge in several ways. The most apparent distinctive feature concerns sociability. Autistic individuals are socially handicapped, avoiding contacts and interactions with others. In contrast, individuals with Williams' syndrome are very friendly and avid of social exchanges. Since both forms of neurological disorders spare musical skills to a large extent, it would be interesting to compare the consequences of the two disorders for the ability to recognize and express musical emotions.

My prediction is that only individuals with Williams' syndrome, not autistic individuals who are matched for cognitive and musical skills, will be able to interpret music emotionally. Specifically, music is above all seen as a function serving social ends. Because autistic individuals are handicapped at this level,[2] it seems reasonable to expect them to have failed to develop an emotionally responsive system for music. Music would have a different meaning for autistic individuals. Although music performance by 'musical savants' has been qualified at times as 'mechanical' (Mottron *et al.* 1999; Sloboda *et al.* 1985), it has never been experimentally assessed. The only empirical study that has been undertaken in relation to musical emotions in autism has reported negative results (Heaton *et al.* 1999). That is, autistic children were found to be able to distinguish musical excerpts as 'happy' or 'sad'. However, performance was generally poor in that study and measured in a single test. Further investigation is worth undertaking, preferably by coupling different types of measurements, as advocated earlier.

## 5.5 Conclusion

The neuropsychology of musical emotions is in its infancy. One of the most vexing issues in this respect is the lack of consensus of what music is for (see, for example, Sloboda & O'Neill, this volume). Knowing better the functionality of music would certainly help neuroscientists to identify the musical emotions that are most probably hard-wired in the brain. Nevertheless, this lack of knowledge has not prevented researchers from making progress in understanding what are the neural bases underlying both our intuitions of what emotions are and what music appreciation might be.

The current evidence is pointing to the existence of a specific neural arrangement for certain musical emotions, such as happiness and sadness. It is also apparent that there is not a single, unitary emotional system underlying all emotional responses to music. For instance, most of the neural pathway underlying pleasantness judgments as related to consonance has been delineated, and involves a complex and distributed system in the brain. Yet, this pleasant-consonant pathway is distinguishable from the happy–sad

[2] One collorary assumption is that music deprivation should lead to social deficiencies. This prediction might be difficult to test since music deprivation almost never occurs. Even deaf individuals manage to dance based on what they can perceive as musical.

neural system. Finally, both forms of emotional judgments require cortical mediation and confirm recent conceptualization of emotions as being subserved by evolved brain structures.

Important advances are to be expected in the near future. This progress will be mostly due to the flourishing of new brain-imaging techniques, but will also reflect the new interest in the domain of emotions in neuroscience. The motivation for understanding the biological foundations of music is also currently high in the scientific community (see, for instance, the two target articles bearing on the biology of music in *Science*, 5 January 2001). There is increasing awareness of the social value of music, particularly with respect to its effectiveness in communicating emotions. Nowadays, music plays an even more powerful and unique role in human life than ever, with wide-ranging effects on many aspects of functioning besides its obvious social function. Music has become such a key element in the human behavioural repertoire that it might be considered as a defining human attribute. In fact, music is so highly valued that very few people are willing to acknowledge a lack of emotional responsiveness to music.[3]

## References

**Adolphs, R., Damasio, H., Tranel, D., & Damasio, A.** (1996). Cortical systems for the recognition of emotion in facial expressions. *Journal of Neuroscience*, **16**, 7678–87.

**Adolphs, R. & Tranel, D.** (1999). Intact recognition of emotional prosody following amygdala damage. *Neuropsychologia*, **37**, 1285–92.

**Adolphs, R., Tranel, D., Damasio, H., & Damasio, A.** (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, **372**, 669–72.

**Adolphs, R., Tranel, D., Damasio, H., & Damasio, A.** (1995). Fear and the human amygdala. *Journal of Neuroscience*, **15**, 5879–92.

**Anderson, A. & Phelps, E.** (1998). Intact recognition of vocal expressions of fear following bilateral lesions of the human amygdala. *Neuroreport*, **9**, 3607–13.

**Balkwill, L.-L. & Thompson, W. F.** (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception*, **17**, 43–64.

**Barrett, A., Crucian, G., Raymer, A., & Heilman, K.** (1999). Spared comprehension of emotional prosody in a patient with global aphasia. *Neuropsychiatry, Neuropsychology and Behavioral Neurology*, **12**, 117–20.

**Bartlett, D.** (1999). Physiological responses to music and sound stimuli. In *Handbook of music psychology*, (ed. D. Hodges), pp. 343–85. San Antonio, TX: IMR Press.

**Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., & Damasio, A.** (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science*, **269**, 1115–18.

**Besson, M.** (1997). Electrophysiological studies of music processing. In *Perception and cognition of music*, (ed. I. Deliège & J. A. Sloboda), pp. 217–50. Hove, UK: Psychology Press.

[3] I wish to acknowledge the assistance of Nathalie Gosselin and Daniel Saumier in creating the figures. My research summarized in this chapter has been supported by research grants from the Natural Sciences and Engineering Research Council of Canada and from the Canadian Institutes of Health Research.

Best, C., Womer, J., & Queen, H. (1994). Hempispheric asymmetries in adults' perception of infant emotional expressions. *Journal of Experimental Psychology: Human Perception and Performance,* **20**, 751–65.

Blair, R., Morris, J., Frith, C., Perrett, D., & Dolan, R. (1999). Dissociable neural responses to facial expressions of sadness and anger. *Brain,* 122, 883–93.

Blood, A., Zatorre, R., Bermudez, P., & Evans, A. (1999*a*). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience,* **2**, 382–7.

Blood, A., Zatorre, R., & Evans, A. (1999*b*). Intensely pleasant emotional responses to music correlate with CBF modulation in paralimbic and other subcortical brain regions. *Society of Neuroscience Abstracts,* 25, 2146.

Bowers, D., Bauer, R., & Heilman, K. (1993). The nonverbal affect lexicon: Theoretical perspectives from neuropsychological studies of affect perception. *Neuropsychology,* 7, 433–44.

Breiter, H., Etcoff, N., Whalen, P., Kennedy, W., Rauch, S., Buckner, R. *ET AL.* (1996). Response and habituation of the human amygdala during visual processing of facial expression. *Neuron,* 17, 875–87.

Broca, P. (1861). Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie (Perte de la parole) [Remarks on the seat of the faculty of articulate language, followed by an observation of aphemia]. *Bulletin de la Société Anatomique,* **6**, 330–57.

Broca, P. (1878) Anatomie comparée des circonvolutions cérébrales. *Revue d'Anthopologie,* **1**, 385–498.

Broks, P., Young, A., Maratos, E., Coffey, P., Calder, A., Isaac, C. *ET AL.* (1998). Face processing impairments after encephalitis: Amygdala damage and recognition of fear. *Neuropsychologia,* 36, 59–70.

Brown, S. (2000). The 'musilanguage' model of music evolution. In *The origins of music,* (ed. N. Wallin, B. Merker, & S. Brown), pp. 271–300. Cambridge, MA: MIT Press.

Bryden, P., Ley, R., & Sugerman, J. (1982). A left-ear advantage for identifying the emotional quality of tonal sequences. *Neuropsychologia,* **20**, 83–7.

Buck, R. (1999). The biological affects: A typology. *Psychological Review,* **106**, 301–36.

Buss, D., Haselton, M., Shackelford, T., Bleske, A., & Wakelfield, J. (1998). Adaptations, exaptations, and sprandels. *American Psychologist,* 53, 533–48.

Calder, A., Young, A., Rowland, D., Perrett, D., Hodges, J., & Etcoff, N. (1996). Facial emotion recognition after bilateral amygdala damage: Differentially severe impairment of fear. *Cognitive Neuropsychology,* 13, 699–745.

Cunningham, J. G. & Sterling, R. S. (1988). Developmental change in the understanding of affective meaning in music. *Motivation and Emotion,* **12**, 399–413.

Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition,* **80**, B1–B10.

Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain.* New York: Avon Books.

Darwin, C. (1871). *The descent of man, and selection in relation to sex.* London: John Murray.

Darwin, C. (1872). *The expression of the emotions in man and animals.* London: John Murray (reprinted Chicago: University of Chicago Press, 1965).

Davidson, R. J. (1992). Emotion and affective style: Hemispheric substrates. *Psychological Science,* 3, 39–43.

**Davidson, R. J. & Irwin, W.** (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Sciences*, 3, 11–21.

**Davis, M.** (1992). The role of the amygdala in fear and anxiety. *Annual Review of Neuroscience*, 15, 353–75.

**Dolan, R. & Morris, J.** (2000) The functional anatomy of innate and acquired fear: Perspectives from neuroimaging. In *Cognitive neuroscience of emotion*, (ed. R. Lane & L. Nadel), pp. 225–41. New York: Oxford University Press.

**Dolgin, K. G. & Adelson, E. H.** (1990). Age changes in the ability to interpret affect in sung and instrumentally-presented melodies. *Psychology of Music*, 18, 87–98.

**Don, A., Schellenberg, E.G., & Rourke, B.** (1999). Music and language skills of children with Williams syndrome. *Child Neuropsychology*, 5, 154–70.

**Edmister. W., Talavage, T., Ledden, P., & Wesikoff, R.** (1999). Improved auditory cortex imaging using clustered volume acquisitions. *Human Brain Mapping*, 7, 89–97.

**Eibl-Eibesfeldt, I.** (1973). The expressive behaviors of the deaf-and-blind-born. In *Social communication and movement*, (ed. M. von Cranach & I. Vine), pp. 163–94. New York: Academic Press.

**Ekman, P.** (1994) All emotions are basic. In *The nature of emotions. Fundamental questions*, (ed. P. Ekman & R. J. Davidson), pp. 15–19. New York: Oxford University Press.

**Ekman, P., Friesen, W., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K. *ET AL.*** (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, 53, 712–17.

**Ekman, P., Sorenson, E., & Friesen, W.** (1969). Pan-cultural elements in facial displays of emotions. *Science*, 164, 86–8.

**Etcoff, N.** (1984). Selective attention to facial identity and facial emotion. *Neuropsychologia*, 22, 281–95.

**Etcoff, N.** (1989). Asymmetries in recognition of emotion. In *Handbook of neuropsychology*, (Vol. 3) (ed. F. Boller & J. Grafman), pp. 363–82. New York: Elsevier.

**Field, T., Woodson, R., Greenberg, R., & Cohen, D.** (1982). Discrimination and imitation of facial expressions by neonates. *Science*, 218, 179–81.

**Gabrielsson, A. & Juslin, P. N.** (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music*, 24, 68–91.

**Gagnon, L. & Peretz, I.** (2000). Laterality effects in processing tonal and atonal melodies with affective and non-affective task instructions. *Brain and Cognition*, 43, 206–10.

**Gagnon, L. & Peretz, I.** (2001). Mode and tempo relative contributions to 'happy–sad' judgments in equitone melodies. *Cognition and Emotion*, (forthcoming).

**George, M., Parekh, P., Rosinsky, N., Ketter, T., Kimbrell, T., Heilman, K. *ET AL.*** (1996). Understanding emotional prosody activates right hemisphere regions. *Archives of Neurology*, 53, 665–70.

**Gerardi, G. M. & Gerken, L.** (1995). The development of affective response to modality and melodic contour. *Music Perception*, 12, 279–90.

**Goldstein, A.** (1980). Thrills in response to music and other stimuli. *Physiological Psychology*, 8, 126–9.

**Gray, J., Young, A., Barker, W., Curtis, A., & Gibson, D.** (1997). Impaired recognition of disgust in Huntington's disease gene carriers. *Brain*, 120, 2029–38.

**Gregory, A., Worral, L., & Sarge, A.** (1996). The development of emotional responses to music in young children. *Motivation and Emotion*, 20, 341–9.

**Halgren, E., Walter, R., Cherlow, D., & Crandall, P.** (1978). Mental phenomena evoked by electrical stimulation of the human hippocampal formation and amygdala. *Brain,* **101**, 83–117.

**Hamann, S., Stefanicci, L., Squire, L., Adolphs, R., Tranel, D., Damasio, H. *ET AL.*** (1996). Recognizing facial emotion. *Nature,* 379, 497.

**Hasselmo, M., Rolls, E., & Baylis, G.** (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioural Brain Research,* 32, 203–18.

**Heaton, P., Hermelin, B., & Pring, L.** (1999). Can children with autistic spectrum disorders perceive affect in music? An experimental investigation. *Psychological Medicine,* 29, 1405–10.

**Heilman, K., Bowers, D., Speedie, L., & Coslett, H.** (1984). Comprehension of affective and non-affective prosody. *Neurology,* 34, 917–21.

**Heilman, K., Scholes, R., & Watson, R.** (1975). Auditory affective agnosia. *Journal of Neurology, Neurosurger and Psychiatry,* 38, 69–72.

**Hornak, J., Rolls, E., & Wade, D.** (1996). Face and voice expression identification in patients with emotional and behavioral changes following ventral frontal lobe damage. *Neuropsychologia,* **34**, 247–61.

**Imaizumi, S., Mori, K., Kiritani, S., Kawashima, R., Sugiura, M., Fukuda, H. *ET AL.*** (1997). Vocal identification of speaker and emotion activates different brain regions. *Neuroreport,* **8**, 2809–12.

**Izard, C.** (1971). *The face of emotions.* New York: Appleton-Century-Crofts.

**Jackson, J.** (1878). On the affections of speech from disease of the brain. *Brain,* **1**, 304–30.

**Johnstone, T. & Scherer, K. R.** (2000). Vocal communication of emotion. In *Handbook of emotions,* (2nd edn) (ed. M. Lewis & J. M. Haviland-Jones), pp. 220–35. New York: Guilford Press.

**Kastner, M. P. & Crowder, R. G.** (1990) Perception of the major/minor distinction: IV. Emotional connotations in young children. *Music Perception,* **8**, 189–202.

**Kivy, P.** (1990). *Music alone. Philosophical reflections on the purely musical experience.* Ithaca, NY: Cornell University Press.

**Kogan, N.** (1994). On aesthetics and its origins: Some psychobiological and evolutionary considerations. *Social Research,* **61**, 139–65.

**Kratus, J.** (1993). A developmental study of children's interpretation of emotion in music. *Psychology of Music,* **21**, 3–19.

**Krumhansl, C. L.** (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology,* 51, 336–53.

**LeDoux, J.** (1996). *The emotional brain.* New York: Simon & Schuster.

**LeDoux, J.** (2000). Cognitive–emotional interactions: Listen to the brain. In *Cognitive neuroscience of emotion,* (ed. R. Lane & L. Nadel), pp. 129–55. New York: Oxford University Press.

**Levitin, D. & Bellugi, U.** (1998). Musical abilities in individuals with Williams' syndrome. *Music Perception,* **15**, 357–90.

**Ley, R. & Bryden, P.** (1982). A dissociation of right and left hemispheric effects for recognizing emotional tone and verbal content. *Brain and Cognition,* **1**, 3–9.

**Marin, O. & Perry, D.** (1999). Neurological aspects of music perception and performance. In *The psychology of music,* (2nd edn) (ed. D. Deutsch), pp. 653–724. San Diego, CA: Academic Press.

**Masataka, N.** (1999). Preference for infant-directed singing in 2-day-old hearing infants of deaf parents. *Developmental Psychology,* **35**, 1001–5.

**Mazzoni, M., Moretti, P., Pardossi, L., Vista, M., Muratorio, A., & Pugliolo, M.** (1993). A case of music imperception. *Journal of Neurology, Neurosurgery and Psychiatry,* **56**, 322–4.

**Mazzuchi, A., Marchini, C., Budai, R., & Parma, M.** (1982). A case of receptive amusia with prominent timbre perception defect. *Journal of Neurology, Neurosurgery and Psychiatry,* **45**, 644–7.

**Miller, G.** (2000). Evolution of human music through sexual selection. In *The origins of music,* (ed. N. Wallin, B. Merker, & S. Brown), pp. 329–60. Cambridge, MA: MIT Press.

**Miller, L.** (1989). *Musical savants. Exceptional skill in the mentally retarded.* Hillsdale, NJ: Erlbaum.

**Morris, J., Frith, C., Perrett, D., Rowland, D., Young,** A., Calder, A., & Dolan, R. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature,* 383, 812–15.

**Morris, J., Scott, S., & Dolan, R.** (1999). Saying with feeling: Neural responses to emotional vocalizations. *Neuropsychologia,* 37, 1155–63.

**Mottron, L., Peretz, I., Belleville, S., & Rouleau, N.** (1999). Absolute pitch in autism: A case-study. *Neurocase,* 5, 485–501.

**Neisser, U.** (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

**Nyklicek, I., Thayer, J. F., & van Doornen, L. J. P.** (1997). Cardiorespiratory differentiation of musically-induced emotions. *Journal of Psychophysiology,* 11, 304–21.

**Oster, H. & Ekman, P.** (1978). Facial behavior in child development. In *Minnesota Symposia on Child Psychology,* (Vol. 11) (ed. W. Collins), pp. 231–76. Hilldsdale, NJ: Erlbaum.

**Panksepp, J.** (1995). The emotional sources of 'chills' induced by music. *Music Perception,* 13, 171–208.

**Panksepp, J.** (1998). *Affective neuroscience.* New York: Oxford University Press.

**Papez, J.** (1937). A proposed mechanism for emotion. *Archives of Neurology and Psychiatry,* 38, 725–43.

**Parry, F., Young, A., Saul, J. S., & Moss, A.** (1991). Dissociable face processing impairments after brain injury. *Journal of Clinical and Experimental Neuropsychology,* 13, 545–58.

**Peretz, I.** (1996). Can we lose memories for music? The case of music agnosia in a nonmusician. *Journal of Cognitive Neurosciences,* 8, 481–96.

**Peretz, I.** (2001*a*). Music perception and recognition. In *The handbook of cognitive neuropsychology,* (ed. B. Rapp), pp. 519–40. Hove, UK: Psychology Press.

**Peretz, I.** (2001*b*). Brain specialization for music: New evidence from congenital amusia. In *The biological foundations of music,* (ed. R. Zatorre & I. Peretz), forthcoming. New York: Annals of the New York Academy of Sciences.

**Peretz, I., Belleville, S., & Fontaine, F. S.** (1997). Dissociations entre musique et langage après atteinte cérébrale: Un nouveau cas d'amusie sans aphasie. *Revue Canadienne de Psychologie Expérimentale,* 51, 354–67. [Dissociation between music and language after brain damage: a new case of the amusia without aphasia. *Canadian Journal of Experimental Psychology.*]

**Peretz, I, Blood, A., Penhune, V., & Zatorre, R.** (2001). Cortical deafness to dissonance. *Brain,* 124, 928–940.

**Peretz, I. & Gagnon, L.** (1999). Dissociation between recognition and emotional judgment for melodies. *Neurocase,* 5, 21–30.

**Peretz, I., Gagnon, L., & Bouchard, B.** (1998a). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition,* 68, 111–41.

Peretz, I., Gaudreau, D., & Bonnel, A.-M. (1998b). Exposure effects on music preference and recognition. *Memory and Cognition,* **26**, 884–902.

Phillips, M., Young, A., Senior, C., Brammer, M., Andrew, C., Calder, A. *ET AL.* (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature,* **389**, 495–8.

Pinker, S. (1997). *How the mind works.* New York: Norton.

Rapp, B. (ed.) (2001). *The handbook of cognitive neuropsychology. What deficits reveal about the human mind.* Philadelphia, PA: Psychology Press.

Schmidt, L. & Trainor, L. (2001). *Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions.* Manuscript submitted for publication.

Scott, S., Young, A., Calder, A., Hellawell, D., Aggleton, J., & Johnson, M. (1997). Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature,* **385**, 254–7.

Sernagor, E. & Grzywacz, M. (1996). Influence of spontaneous activity and visual experience on developing retinal receptive fields. *Current Biology,* 6, 1503–8.

Sergent, J., Ohta, S., MacDonald, B., & Zuck, E. (1994). Segregated processing of facial identity and emotion in the human brain: A PET study. *Visual Cognition,* 1, 349–69.

Shammi, P. & Stuss, D. (1999). Humour appreciation: A role of the right frontal lobe. *Brain,* 122, 657–66.

Sloboda, J. A., Hermelin, B., & O'Connor, N. (1985). An exceptionnal musical memory. *Music Perception,* 3, 155–70.

Sprengelmeyer, R., Young, A., Calder, A., Karnat, A., Lange, H., Hömberg, V. *ET AL.* (1996). Loss of disgust: Perception of faces and emotion in Huntington's disease. *Brain,* 119, 1647–65.

Terwogt, M. M. & van Grinsven, F. (1988). Recognition of emotions in music by children and adults. *Perceptual and Motor Skills,* **67**, 697–8.

Terwogt, M. M. & van Grinsven, F. (1991) Musical expression of moodstates. *Psychology of Music,* 19, 99–109.

Trainor, L., Clark, E., Huntley, A., & Adams, B. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior and Development,* **20**, 383–96.

Trainor L. & Heinmiller, B. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behavior and Development,* 21, 77–88.

Tranel, D., Damasio, A., & Damasio, H. (1988). Intact recognition of facial expression, gender and age in patients with impaired recognition of face identity. *Neurology,* **38**, 690–6.

Trehub, S. (2001). Musical predispositions in infancy. In *The biological foundations of music,* (ed. R. Zatorre & I. Peretz), forthcoming. New York: Annals of the New York Academy of Sciences.

Trehub, S. & Trainor, L. (1998). Singing to infants: Lullabies and play songs. In *Advances in infancy research,* (ed. C. Rovee-Collier), pp. 43–77. Greenwich, CT: Ablex.

Tsang, C., Trainor, L., Santesso, D., Tasker, S., & Schmidt, L. (2001). Frontal EEG responses as a function of affective musical features. In *The biological foundations of music,* (ed. R. Zatorre & I. Peretz), pp. 00–00. New YorK: Annals of the New York Academy of Science.

Wallin, N., Merker, B., & Brown, S. (ed.) (2000). *The origins of music.* Cambridge, MA: MIT Press.

Whalen, P., Rauch, S., Etcoff, N., McInerney, S., Lee, M., & Jenike, M. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience,* **18**, 411–18.

Young, A., Aggleton, J., Hellawell, D., Johnson, M., Broks, P., & Hanley, J. (1995). Face processing impairments after amygdalotomy. *Brain,* **118**, 15–24.

**Young, A., Newcombe, F., de Haan, E., Small, M., & Hay, D.** (1993). Face perception after brain injury. Selective impairments affecting identity and expression. *Brain*, **116**, 941–59.

**Zald, D., Lee, J., Fluegel, K., & Pardo, J.** (1998). Aversive gustator stimulation activates limbic circuits in humans. *Brain*, **121**, 1143–54.

**Zajonc, R.** (1984) On the primacy of affect. In *Approaches to emotion*, (ed. K. R. Scherer & P. Ekman), pp. 259–70. Hillsdale, NJ: Erlbaum.

**Zentner, M. R. & Kagan, J.** (1996). Perception of music by infants. *Nature*, **383**, 29.

CHAPTER 6

# ANTHROPOLOGICAL PERSPECTIVES ON MUSIC AND EMOTION

JUDITH BECKER

The study of anthropology is the study of the modes of being human, while the study of ethnomusicology is the study of the modes of music making as a means of cultural expression and as sources of meaning. Both are concerned with human behaviour and the cultural ideologies that inform behaviour; both tend to rely upon ethnographic research and techniques of participant observation, with the results later written and published as a monograph. Within each discipline there are 'universalists' who tend to look for commonalities of behaviour and beliefs, and 'particularists' who are more interested in how we differ one from another, one group from another group.

Whatever the difference of perspective, nearly all anthropologists and ethnomusicologists subscribe to the dictum that one must attempt to understand behaviours and belief systems from within, that is, the necessity for making the effort to understand other musics and other styles of emotional expression from the perspective of their owners. This is related to, but not the same thing as, what is called 'cultural relativism,' which is taken to mean a refusal to make value judgments about beliefs and practices of peoples outside one's own group. A commitment to trying to comprehend the motivations and underlying paradigms of another's actions certainly delays or impedes a 'rush to judgment'. In the case of music and emotion, the ethnographic approach involves both looking closely at the cultural construction of emotion and at the cross-cultural differences in the uses, purposes, and meanings of musical expression.

In this chapter, I will present a theoretical frame for the cross-cultural exploration of music and emotion, provide specific examples, and conclude with a section on the possibility of conjoining humanistic and scientific approaches.

## 6.1 Culturally inflected listening

Studies on music and emotion conducted by Western scholars and scientists nearly always presume a particular image of musical listeners: silent, still listeners, paying close attention to a piece of music about which they communicate the type of emotion evoked by the piece to an attendant researcher. Communication of the emotion may be immediate in a laboratory setting, or retrospective, recalling the emotion after the musical event.

What is wrong with this image?

Nothing, if the intent is to describe the emotional responses to decontextualized performances by middle-class American or European listeners to music while seated quietly, in the second half of the twentieth century or the early twenty-first century. The laboratory situation may reflect the habit of many current Western listeners, possibly even more so now than in the past as so much musical listening takes place attached to headphones, rather than at live performances. Silent, still, focused listening is also the habit in some other musical traditions, notably the north Indian Hindustani 'classical' tradition, where one sits quietly, introspectively listening to the gradually developing filigree of the musical structure of a *raga*, played, perhaps, on a sitar. Thoughts and feelings are turned inward. The setting is intimate, conducive to introspection, and a distancing from one's fellow listeners.

But if the intent is to delineate something more general about the relationships between musical event and musical affect, the image of an inwardly focused, isolated listener is inadequate. This portrayal of listener and listening presents a set of unexamined ideologies and presuppositions that would not apply for most of the world. The unasked questions include: What constitutes 'listening' to music? What are the appropriate kinds of emotions to feel? What kind of subjectivity is assumed? Who is it that is 'having' the emotions? How is the event framed?

Anthropology and ethnomusicology have contributed to the study of music and emotion an expanded notion of the possibilities that constitute the relationship between the musical event and the listener, and the degree to which emotional response to musical events is culturally inflected. Musical performances and listeners find themselves in a relationship in which they define each other through continuous, interactive, ever-evolving musical structures and listener responses. Meaning resides in the mutual relationship established at any given moment in time between particular listeners and musical events. A group of listeners develops a 'community of interpretation' (Fish 1980), not necessarily uniform, but overlapping in some salient features. This community of listening means that when presented with a musical event, this community will approach the music with a pre-given set of expectations, a 'forestructure of understandings' (Gergen 1991, p. 104). Every hearer occupies a position in a cultural field not of his or her own making: every hearing is situated.

We accumulate our listening habits and expectations largely unawares; only when confronted with an alternate kind of listening are we likely to reflect upon our own conventionalized mode. Listeners can shift modes in different contexts, such as the ways in which one listens to music at a chamber music recital, or at a rock concert, or a jazz club, or a cinema, or at a salsa club (Becker 1983). Cross-culturally, modes of listening may add features not shared by us, or may not involve features that we take for granted. What is appropriate to say about musical affect, what one does *not* say, what one feels, and what one does *not* feel may reveal underlying assumptions surrounding musical listening. What is *not* assumed in one mode (such as bodily movement in Western classical listening) may become central in another mode (such as dancing while listening to a salsa band). To sit quietly focused on musical structure at a salsa concert is as inappropriate as swing dancing to a Schubert quintet.

We need a term to express the temporal and spacial situatedness of the hearer which is the aural equivalent of the visual term for modes of seeing; that is, *the gaze*. Frequently a feminist challenge to the dominance of *the male gaze* in literature (Kern 1996), film (Zizek 1991, p. 88), painting (Hebdige 1995), photography (Slater 1995), television (Morley 1995), psychoanalysis (Zizek 1996, p. 90), medicine (O'Neill 1995), and advertising (Barnard 1995), the term *gaze* is now used in a wide variety of contexts to exemplify the situatedness of looking, the historical and psychological specificity of any one visual approach, and the complex imbrication of modes of seeing with rhetorical and institutional structures and beliefs (Brennan & Jay 1996; Gamman & Marshment 1988; Jenks 1995).

Modes of *looking* imply habits of seeing that change not only across space (Mitchell 1986), but also at different historical periods within a single culture (Baxandall 1974; Goldhill 1996). Similarly, modes of *listening* vary according to the kind of music being played, the expectations of the musical situation, and the kind of subjectivity that a particular culture has fostered in relation to musical events (Johnson 1995). Even more than modes of looking, modes of listening implicate not only structures of knowledge and beliefs, but intimate notions of personhood and identity. Listening addresses interiors; listening provides access to what is hidden from sight.

Styles of audition, of aural perception, of aural awareness, of *listening* in response to musical events directly impinge upon studies of music and emotion. A given community will foster a particular comportment to listening; a comportment not only of attitude, affect, and expectation, but also bodily gesture. Emotional responses to music do not occur spontaneously, nor 'naturally', but rather, take place within complex systems of thought and behaviour concerning what music means, what it is for, how it is to be perceived, and what might be appropriate kinds of expressive responses. We need a word like Bourdieu's (1977) *habitus*, coined as an alternative to terms such as 'culture' which seemed too static, and sometimes seemed to imply a rigidity, an all-inclusiveness that obscured individual, idiosyncratic, or innovative modes of thought and behaviour. Still left with the need to refer to the ways in which beliefs and behaviours seem relatively stereotypical within a given society, Bourdieu (1977, p. 72) proposed the term habitus to do the theoretical work formerly carried by the word 'culture':

> The structures constitutive of a particular type of environment (e.g. the material conditions of existence characteristic of a class condition) produce *habitus*, systems of durable, transposable *dispositions* . . . that is, as principles of the generation and structuring of practices and representations which can be objectively 'regulated' and 'regular' without in any way being the product of obedience to rules, objectively adapted to their goals without presupposing a conscious aiming at ends or an express mastery of the operations necessary to attain them and, being all this, collectively orchestrated without being the product of the orchestrating action of a conductor.

Bourdieu (1977, p. 214, f. 1) goes on to define what he means by *dispositions*:

> The word *disposition* seems particularly suited to express what is covered by the concept of habitus defined as a system of dispositions. It expresses first the *result of an organizing action*, with a meaning close to that of words such as structure; it also designates a *way of being, a habitual state* (especially of the body and, in particular, a *predisposition, tendency, propensity*, or *inclination*.

Habitus is an embodied pattern of action and reaction, in which we are not fully conscious of why we do what we do; not totally determined, but a *tendency* to behave in a certain way. Our habitus of listening is tacit, unexamined, seemingly completely 'natural'. We listen in a *particular* way without thinking about it, and without realizing that it even is a particular way of listening. Most of our styles of listening have been learned through unconscious imitation of those who surround us and with whom we continually interact. A habitus of listening suggests, not a necessity or a rule, but an inclination, a disposition to listen with a particular kind of focus, to expect to experience particular kinds of emotion, to move with certain stylized gestures, and to interpret the meaning of the sounds and one's emotional responses to the musical event in somewhat (never totally) predictable ways. Scholars working within the disciplines of anthropology or ethnomusicology typically assume that the stance of the listener is not a given, not *natural*, but necessarily influenced by place, time, the shared context of culture, and the intricate and unreproduceable details of one's personal biography.

The term I have adapted from Bourdieu, habitus of listening, underlines the interrelatedness of the perception of musical emotion and learned interactions with our surroundings. Our perceptions operate within a set of habits gradually established throughout our lives and developed through our continual interaction with the world beyond our bodies, the evolving situation of *being in the world.*

## 6.2 Emotion as a cultural construct

> But recognition of the fact that thought is always culturally patterned and infused with feelings, which themselves reflect a culturally ordered past, suggests that just as thought does not exist in isolation from affective life, so affect is culturally ordered and does not exist apart from thought. (Rosaldo 1984, p. 137)

This view, while in the ascendancy among cultural anthropologists and ethnomusicologists, has not gone unchallenged within the discipline of anthropology:

> I can make no sense of a line of thought which claims that 'passions' are culturally defined. From my prejudiced position as a social anthropologist this passage reveals with startling clarity the ultimately radical weakness of the basic assumption of cultural anthropology, namely, that not only are cultural systems infinitely variable, but that human individuals are products of their culture rather than of their genetic predisposition. (Leach 1981, p. 16)

These two quotations, starkly put and differentiated by nationality, gender, and perspective—the American, female, cultural anthropologist and the British, male, social anthropologist—state baldly the issue at hand: the cultural relativism of emotion and thought, or the universality of emotion and thought. (Leach was not a racist, but rather a believer in the commonality of mankind. See also Brown 1991; Goodenough 1970, p. 122; Levi-Strauss 1962, p. 161; Spiro 1984; Turner 1983). These contesting views have elicited penetrating dialogues on both sides of the divide and continue to evoke considerable emotion in their defenders. There are good reasons for these passions: much is at stake.

Informing our beliefs about the universality or, conversely, the culturally conditioned aspects of music and emotion, the path of Western intellectual history leads in both

directions; one direction taken by the sciences that stresses general laws and instances, the other taken by the humanities and cultural anthropology that stresses cases and interpretations (Geertz 1983*a*).

One way to understand the divide between the scientific and the cultural approach is to look at its development from the eighteenth-century Enlightenment onwards. One of the great contributions of the Enlightenment was the propagation of the idea (already taught by religion, but little observed by society) that humankind shared a basic natural state, independent of geography, chronology, or personality ('We hold these truths to be self evident: that all men are created equal'; Jefferson 1994). In its day, this doctrine was dazzlingly liberal and liberating. It led, in spite of all the detours to the contrary, to the end (almost) of institutionalized slavery in the Western world. A sentence from a history book of the eighteenth-century vividly presents a theatrical metaphor for the belief in the commonality of human life and human nature.

> The stage-setting [in different periods of history] is, indeed, altered, the actors change their garb and their appearance; but their inward motions arise from the same desires and passions of men, and produce their effects in the vicissitudes of kingdoms and peoples. (J. J. Mascou, *Geschichte der Teutschen*, quoted in Lovejoy 1948, p. 173)

While elegantly simple, and a vast improvement over earlier views concerning 'The Great Chain of Being' (Lovejoy 1964), this view, in practice, led to the assumption that all peoples everywhere thought and felt like educated, male, Europeans and Americans. Post-colonial studies (Appadurai 1996; Spivak 1988) and gender studies (McClary 1991; Solie 1993) have brought home the bias in such views. We have come to appreciate the nuanced differences of emotion with different stage settings, garbs, and appearances. The 'desires and passions of men' have come to be seen as *not* producing identical 'inward motions'. Partly through the cumulative effects of works by cultural anthropologists writing in the 1970s (Geertz 1973, p. 36; Myers 1979), the 1980s (Levy 1984; Lutz 1986, 1988; Rosaldo 1984; Shweder & Bourne 1984), and the 1990s (Irvine 1990; Lutz & Abu-Lughod 1990), scholars who have championed the concept of the cultural construction of emotion, social scientists and psychologists are increasingly sensitive to the cultural component in the categorizing of, the interpretation of, and the expression of emotion (Davidson 1992; Ekman 1980, p. 90; Ortony *et al.* 1988, p. 26; Russell 1991*a*, 1991*b*).

If we accept Leach's version of the uniformity of human passions, we condone the silences imposed upon subalterns of all times and places whose feelings were assumed to be isomorphic with those of the persons who controlled the writing of history, and we ignore the developing body of data supporting the cultural inflection of the emotions. If we accept the idea of the social construction of knowledge, of morality, and emotion, we seem to be abandoning the idea of a *human* nature, a bond of mind, emotion, and meaning that enfolds us all, and binds us to one another. We may also be in danger of losing sight of the individual as he or she slips into the constructed conventionality of cultural appearance, behaviour, beliefs, and desires and disappears altogether. Persons may become exemplars, instances of this or that cultural model.

I would like to propose that both approaches have incontrovertible empirical support, and that rather than choose sides, we need to accept the paradox that, in fact, we

cannot do without either perspective (Needham 1981; Nettl 1983, p. 36; Shweder 1985; Solomon 1984). Cultural difference in the expression of, the motivation for, and the interpretation of emotion in relation to musical events has been persuasively demonstrated over and over again; for example in music from South Africa (Blacking 1973, p. 68), Liberia (Stone 1982, p. 79), Brazil (Seeger 1987, p. 129), New Guinea (Feld 1982, p. 32), Peru (Turino 1993, p. 82), South India (Viswanathan & Cormack 1998, p. 225), Java (Benamou 1998), and Arabic music (Racy 1998, p. 99). Likewise, the fact that most of us can, with experience and empathy, come to understand differing expressive reactions to different kinds of music as reasonable and coherent demonstrates some level of commonality and universality in relation to both music and emotion. It may be that we come into the world with the full range of human emotional expression available to us (Geertz 1974, p. 249). Through continual patterns of interaction with (primarily) close family members in the early years, we develop particular patterns of emotional feelings and expressions in relation to the events of our lives. For the most part, habituated responses and actions delimit the range and type of any one person's emotional responses. Yet, it would appear that we can imaginatively enter into a much wider palette of human emotional possibilities. We need to make a Hegelian move and transcend the dichotomy between scientific universalism and humanistic particularity and embrace both as necessary to the study of music and emotion.

## 6.3 Person as a cultural construct

> The 'subjectivity' we are discussing here is the capacity of the speaker to posit himself as 'subject'. It is defined not by the feeling which everyone experiences of being himself (this feeling, to the degree that it can be taken note of, is only a reflection) but as the psychic unity that transcends the totality of the actual experiences it assembles and that makes the permanence of the consciousness. (Benveniste 1971, p. 224)

Benveniste has argued that the sense of person is a product of 'languaging', as have many others, most notably, Buddhist philosophers. (The gerund 'languaging' is a form favoured by scholars who wish to differentiate language activity from the structure of language, see Becker 1995, p. 9; Maturana & Varela 1987, p. 234; Smith & Ferstman 1996, p. 52). While the sense of personhood seems inextricably tied to the development of language, one's sense of bodily boundaries—our sense of what (somatically) belongs to us and where our bodies end—seems to escape the linguistic formation of person and takes shape in areas of the brain not directly involved with languaging (Damasio 1999, p. 108; Melzack 1992). The literature on subjectivity is vast and cannot be dealt with here. None the less, an important aspect of any particular habitus of listening is the way in which the listener thinks of him or herself as a person, how he or she establishes identity in relation to other persons and the phenomenal world.

A classic definition of Western identity is that formulated by Clifford Geertz (1983*b*, p. 59):

> The Western conception of the person as a bounded, unique, more or less integrated motivational cognitive universe, a dynamic center of awareness, emotion, judgment, and action organized into a distinctive whole and set contrastively both against other such wholes and against its social

and natural background, is, however incorrigible it may seem to us, a rather peculiar idea within the context of the world's cultures.

Among others, the philosopher Charles Taylor (1989) has written about the development of this style of personhood in the West, from Plato to the present day (see also Gergen 1991, p. 18):

> Our modern notion of the self is related to, one might say constituted by, a certain sense (or perhaps a family of senses) of inwardness . . . The unconscious is for us within, and we think of the depths of the unsaid, the unsayable, the powerful inchoate feelings and affinities and fears which dispute with us the control of our lives, as inner . . . But as strong as this partitioning of the world appears to us, as solid as this localization may seem, and anchored in the very nature of the human agent, it is in large part a feature of our world, the world of modern, Western people. The localization is not a universal one, which human beings recognize as a matter of course, as they do for instance that their heads are above their torsos. Rather it is a function of a historically limited mode of self-interpretation, one which has become dominant in the modern West and which may indeed spread thence to other parts of the globe, but which had a beginning in time and space and may have an end. (Taylor 1989, p. 111)

It is almost a truism of contemporary cultural anthropology that the nature of subjectivity, of the sense of self, varies cross-culturally. This is a profoundly anti-intuitive notion, and one that has strong objectors in philosophy from St Augustine (von Campenhausen 1964, pp. 226–7) to Bertrand Russell (Seckel 1986, p. 96). As with the issue of 'emotion', to say that subjectivity is culturally constructed seems to deny the basic humanity of mankind. Yet, differences seem to remain in how persons think of themselves in relation to other persons, and these differences are often markedly cultural (Shweder & Bourne 1984, p. 191).

## 6.4 Person and emotion in the habitus of listening

The particular subjectivity of the listeners described in psychological studies, the prototypical Western, middle-class listener to music is likely to be some variant of the following: an individual with a strong sense of separateness, of uniqueness from all other persons, an individual whose emotions and feelings are felt to be known in their entirety and complexity only to him or herself, whose physical and psychic privacy is treasured, and whose emotional responses to a given piece of music are not felt to be in relation to anything outside of his or her own particular self-history and personality: The emotion, for us, belongs to the individual, not to the situation or to relationships. Emotion is the authentic expression of one's being, and is, in some sense, natural and spontaneous. The emotion is interior, may or may not be shared with anyone else, and may be a guide to one's inner essence. Leo Treitler (1993, p. 48) has written about one way in which this style of subjectivity can relate to listening to music:

> It is that interaction of the selves of the listener with those in the music—no, better put: the awareness of the self (selves) in the music through its (their) interaction with the listener's self—that interests me here; musical communication as a function of the interaction of identities.

The differing identities of the listening subject, and those projected by the music become the focus of interest for Treitler. One common variant of this Western kind of subjectivity while listening to music is to identify with the different identity projected by the music. Listening to music offers the opportunity to temporarily be another kind of person than one's ordinary, everyday self. This interpretation segues into theories concerning one's fantasy life and seems to be a fruitful approach to the kind of emotion associated with music that one finds, for example, among adolescents (Frith 1987, p. 143; Shepherd & Giles-David 1991), and may help to explain their profound identification with the popular music of their times. Musical listening may offer the opportunity of experiencing relief from one's own presentation of everyday self (Goffman 1959) by trying on another self-presentation (see also DeNora, this volume).

### 6.4.1 Being a griot

How different is the subjectivity, the habitus of listening of the Western listeners to the Wolof griots of Senegal as described by Judith Irvine (1990). Among the Wolof, the musical expression of emotion is dialogical and situational, not personal and interior. Griots, low-caste individuals, are believed to be highly expressive, highly excitable, 'volatile and theatrical'. The nobles are believed to be the opposite: composed, 'cool', detached, and somewhat bland in affect. It is the duty of the griots to stimulate the nobles to action, nobles who might otherwise be given over to lethargy. The highly expressive, emotional performances of the griots are intended, in part, to provide the energy to the nobles that they might carry out their governing duties. Irvine relates that a frequent image in Wolof oral poetry and epic narratives involves the playing of an ensemble of drums and iron clappers to awaken the king 'lest his royal duties go unfulfilled' (Irvine 1990, p. 134). The emotion of the musicians and dancers is contrasted to the lack of emotion of their primary audience, the nobles. Both the emotion of the griots and its absence in the nobles are public, dialogical, and situational; not private and hidden. One of the primary manifestations of subjectivity among Europeans and Americans, emotions are not personal attributes for the Wolof griots and nobles in the same way. For us, emotional responses to music are not considered to be assigned by virtue of one's class and profession, but to be an inalienable characteristic of a bounded, inviolate individual. This is not at all to say that Wolof nobles do not feel strong personal emotions, or that a griot is necessarily always highly extroverted and volatile. It is only to say that these are the *dispositions*, the *inclinations* that are likely to be fulfilled more often than not. The key participants in these musical events exhibit a habitus of listening and a type of subjectivity that largely conforms to Wolof cultural expectations.

### 6.4.2 The habitus of listening to the sitar, tambura, and tabla

Earlier, mention was made of the similarity of the habitus of listening of the listener of Hindustani 'classical' music and the Western middle-class listener in terms of physical stillness, focused attention, and inner withdrawal. Furthermore, both traditions would claim that music can *represent* emotion. In his autobiography, Ravi Shankar (1968, pp. 23, 27) describes his own subjectivity as a musician, a style of personhood that seems concordant with what I or most readers of this essay might feel as performing musicians:

A *raga* is an aesthetic projection of the artist's inner spirit; it is a representation of his most profound sentiments and sensibilities, set forth through tones and melodies . . . I may play *Raga Malkauns*, whose principal mood [*rasa*] is *veera* [heroic], but I could begin by expressing *shanta* [serenity] and *karuna* [compassion] in the *alap* and develop into *veera* and *adbhuta* [astonishment] or even *raudra* [anger] in playing the *jor* or *jhala*.

The inner spirit of the artist, functioning within culturally constructed categories of emotion, is made manifest in the outward expressions of his musical presentation. Bringing yet another dimension of similarity to the two listening situations is the fact that emotion (*rasa*) experienced in listening to Hindustani classical music is distanced and impersonal. One can feel the emotion without the troublesome immediacy and consequences of an emotion that compels action (Abhinavagupta in Gnoli 1968, pp. 82–5; Masson & Patwardhan 1969, p. 46). June McDaniel (1995, p. 48), writing of *rasa* in the Indian province of Bengal, uses the metaphor of the glass window separating the experiencer from the emotion while still allowing a clear view:

*Bhava* is a personal emotion; *rasa* is an impersonal or depersonalized emotion, in which the participant is distanced as an observer. Why is a depersonalized emotion considered superior to a personal one? Because the aesthete can experience a wide range of emotions yet be protected from their painful aspects. Emotion is appreciated through a glass window, which keeps out unpleasantness. Though the glass is clear, thus allowing a union of sorts with the observed object, the window is always present, thus maintaining the dualism.

The following description is a Western mirror-image of the way to experience *rasa*:

When people listen, say, to 'Questi i campi di Tracia' from Monteverdi's *Orfeo* . . . they do not directly perceive the anguish and guilt of the twice-widowed singer. Listeners only hear a *representation* [italics mine] of the way his voice moves under the influence of his emotions. Nevertheless, this can give listeners important insights into a type of emotional response. Monteverdi skillfully displays the contortions through which Orpheus's voice goes. When listeners know how a voice moves under the influence of an affect, they are given (if they are familiar with the conventions of the music, and otherwise qualified) an immediate demonstration of something about the affect. A good performance of this aria immediately demonstrates to a sensitive audience something about what it is like to feel guilt, remorse, and despair. (Young 1999, p. 48)

In both the Hindustani and the Western classical listener, an emotion evoked by listening to music can be contemplated with a certain deliberation and calmness. But at some point the congruencies between the habitus of listening of each breaks down. The Western observer may well, as Treitler suggests, be involved with comparing identities, or with constructing a more glamorous self in relation to the music heard, or with contemplating 'what it is like to feel guilt, remorse, and despair'. The Indian, however, may be performing a very different act, a somewhat strenuous religious exercise, a kind of refining of emotional essence, a distillation of his or her emotion that will lead to a transformation of consciousness to a higher level of spirituality. Listening to music for the Hindustani classical music devotee should not be, according to canon, a passive act, but requires the active will and mind of the listener to carry consciousness to a higher plane, closer to the divine (Coomaraswamy 1957, p. 39). Western references to *inspiration* or *genius*, which at one time perhaps indicated a holy possession, have become

largely metaphoric rather than literal expressions of sacred connections between musical performance and musical listening. In Indian classical traditions, the pursuit of emotion, of *rasa*, in relation to listening to music, may be a path to greater awareness, leading one to cosmic insight. Dance can share as well in this configuration of emotion as a stepping stone to a higher gnosis. One of India's most revered Bharata Natyam dancers has written the following:

> It is here that Bharata Natyam, the ancient and holy art of Indian dance, cuts deeply into the conscious and subconscious levels and revealingly brings to the forefront the fact that it is ultimately and intimately oriented to the nucleus, *atman* [the universal self]. It is a revelation not only to the performing artist, but in an equal measure to the audience as well . . . By the inexplicable power born of the union of melody, lyric, rhythm and gesture, the emotions are released from their limited secular locus and are expanded to universal proportions where there is only pure spirit with nothing of the sensual. (Balasaraswati 1985, pp. 2–3)

According to the Bengali version of the theory of *rasa*, intense emotions, not milder 'aesthetic emotions', are the appropriate vehicle to lead one to mystical knowledge of life's meaning and purpose. Emotions are compared to water that can best be understood by immersion in an ocean rather than by the delicate feel of a raindrop (McDaniel 1995, p. 51). What we might call emotional excesses become pregnant possibilities for greater spiritual attainment for an Indian music listener. While the quiet stance and introverted demeanor of the listener in the prototypical Western case and the Hindustani classical listener is similar, the understanding and interpretation of what is supposed to happen in each case differs. In one case, the listener may be exploring the emotional nuances of his or her inner self or identifying with the emotional interiors presented by the music: In the other, the listener is trying to bring about a kind of 'sea' change, a different self altogether, one that comes closer to divinity.

## 6.5 Arousal: a human universal

Particularities relating to the cultural construction of personhood and of emotion may obscure certain physiological constants that are the correlates of listening to music. Arousal, defined in its most narrow sense as stimulation of the autonomic nervous system (ANS), is one of the most important aspects of musical performance and plays into nearly all studies of music and emotion (see Meyer, this volume; Scherer & Zentner, this volume; Sloboda & Juslin, this volume). The heart beats faster, the pulse rises, breathing becomes shallower, the skin temperature rises, and the pattern of brain waves becomes less regular. All these changes have been observed without any necessary reference to the affective, interpretive component of arousal. They may occur in relation to sexual activity, to exercise, or as a result of drugs or alcohol, as well as to musical listening.

In his famous and controversial theory of the emotions, William James claimed that the physiological component of arousal is primary and precedes the interpretation of the subsequent emotion (James 1950, pp. 442–85). The 'feeling' of anger is the feeling that results from an angry facial expression, bodily stance, shallow breathing, etc. Anger, the emotion, *is*, according to James, what one *feels* when enacting this display. Following

James, but with much more knowledge of the neurophysiology of emotion, Damasio (1999, p. 79) claims that the term 'emotion' should be applied to ANS arousal, and that 'feeling' which follows 'emotion' should be applied to the complex cognitive, culturally inflected interpretation of 'emotion'. Both the Polish Grotowski school of acting (Grotowski 1968) and the Indian Kathakali theatre tradition (Schechner 1988, p. 270) seem to support the theories of James and Damasio. Training a Kathakali actor, or a Western actor of the Grotowski tradition, begins with *mimesis*, recreating the bodily gestures of an emotion, not with delving into one's memories to recreate a 'feeling', as in Stanislavsky's school of 'method' acting (Stanislavsky 1958, p. 1977).

In any case, it is useful analytically to separate physiological arousal ('emotion' for James and Damasio) from the more cognitive concept of 'feeling'. Unlike feelings that directly relate to judgments and beliefs learned within a cultural context and rely upon linguistic categories which are often incommensurable across languages, arousal is more clearly a universal response to musical listening. Combined with a concentrated religious focus, musical arousal can contribute to extreme states of emotion.

Happiness is the emotion most frequently associated with musical listening and may constitute one of the 'universals' of cross-cultural studies of music and emotion. From the 'polka happiness' of the Polish-American parties of Chicago (Keil 1987, p. 276), to the !Kung of the Kalahari desert: 'Being at a dance makes our hearts happy' (Katz 1982, p. 348), to the Basongye of the Congo who 'make music in order to be happy' (Merriam 1964, p. 82), to the extroverted joy of a Pentecostal musical service, music has the ability to make people feel good (cf. Gabrielsson, this volume; Sloboda & O'Neill, this volume). The happiness of listening to music, however one construes 'happiness', is in part the simple result of musical arousal. We tend to feel better when we are musically aroused and excited. The emotion may be attributed to some other aspect of the event, such as the text or our own dancing; none the less, the musical stimulation should not be minimized. Music can be a catalyst for a changing state of consciousness (cf. Persson, this volume).

The following section includes three examples of extreme emotion contextually situated within a religious ceremony in which music is an essential element; all three are sites of the author's fieldwork.

### 6.5.1 Music and ecstasy: the Sufis

The strongest version of happiness in relation to musical listening and an example of extreme arousal is ecstasy. (See Rouget 1985 for descriptions of music and trancing in many time periods and geographical locations.) Usually associated with religious rituals, ecstasy, as extreme joy, almost by definition involves a sense of the sacred (although musical ecstasy can justly be claimed by some attendees at secular musical events such as rock concerts). The degree to which Muslim Sufi orders have formalized and institutionalized musical ecstasy has seldom been exceeded. The works of the eleventh-century Persian Sufi mystic al Ghazzali about music and ecstasy are still basic pedagogical texts for contemporary Sufis in Iran, Afghanistan, Pakistan, and north India.

> The heart of man has been so constituted by the Almighty that, like a flint, it contains a hidden fire which is evoked by music and harmony, and renders man beside himself with ecstasy. These

harmonies are echoes of that higher world of beauty which we call the world of spirits, they remind man of his relationship to that world, and produce in him an emotion so deep and strange that he himself is powerless to explain it. (Ghazzali 1991, p. 57)

Sufi musicians, called *qawwal* in Pakistan and north India, play for religious ceremonies in which the devotees may reach toward a pinnacle of ecstasy that will bring them into close communion with Allah. The *qawwals* move from song to song, eliciting reactions from the devotees such as moaning, sighing, swaying, even rising up and turning in place as a certain text line speaks directly to the spiritual condition of a particular devotee (Qureshi 1986, p. 119). For a devout Sufi, his ecstasy is a preview, a foretasting of his ultimate union: and though he may weep, he weeps from excess of joy.

Sufi doctrine interprets the music as supportive, as secondary to the all important text, the religious poetry (Qureshi 1986, p. 83). Yet the question remains: to what degree is the arousal stimulated by the sensual overload of intensifying rhythms and soaring phrases sung over and over again at the top of the *qawwal's* range? (For examples, listen to any of the many CDs of Nusrat Fateh Ali Khan.)

A Sufi habitus of listening at a musical religious ceremony involves a sequence of feeling and action that could be called a 'script' (Russell 1991*a*, 1991*b*). Initially, while sitting quietly and reverently, the listener may hope that he or she will be touched by a particular line of text that seems directly applicable to his or her personal situation. If the lead *qawwal* catches the subtle indications of arousal in a listener, he will begin repeating the verse over and over again. The Sufi script then calls for swaying, weeping, rising up, moving to the centre of the room, and slowly turning in place (Fig. 6.1). The emotion of the script, when fully acted out, is the ultimate joy of a direct and personal knowledge of Allah. Musical emotion in the listener changes both its form and its intensity as the script progresses.

### 6.5.2 Music and rage: Balinese bebuten trancing

In Bali, an exorcist ceremony is performed when misfortune befalls a village in which the negative powers of the divine witch Rangda must be neutralized for the well-being of the community. The elaborate mask of the witch Rangda will be brought from its storage place in the temple, blessed and infused with spirit in a potent ceremony conducted by a priest. An exceptionally tall and spiritually strong man will be chosen to undergo the trance of *becoming* the witch. A number of men will volunteer to undergo another kind of trance called *bebuten*, from the root *buta*, which means a creature of base instincts, low on the ladder of sentient beings, often translated as *demon*. These men will confront the witch and ultimately neutralize her power with the help of a magical beast called Barong. *Bebuten* trance is not ecstasy. It is a feeling of rage directed toward the witch Rangda, and may leave the trancer feeling embarrassed later by his behaviour during the trance and with an exhaustion that may last for several days (Eiseman 1989, p. 153). Yet, the trancing is a social obligation which he fulfills voluntarily, surrendering his own comfort for the betterment of his community. He experiences a kind of homicidal rage (or, rather, 'theo-cidal' rage). One experienced trancer (I Wayan Dibia) passionately described his feelings while listening to the music of the gamelan (Fig. 6.2) upon encountering the witch:

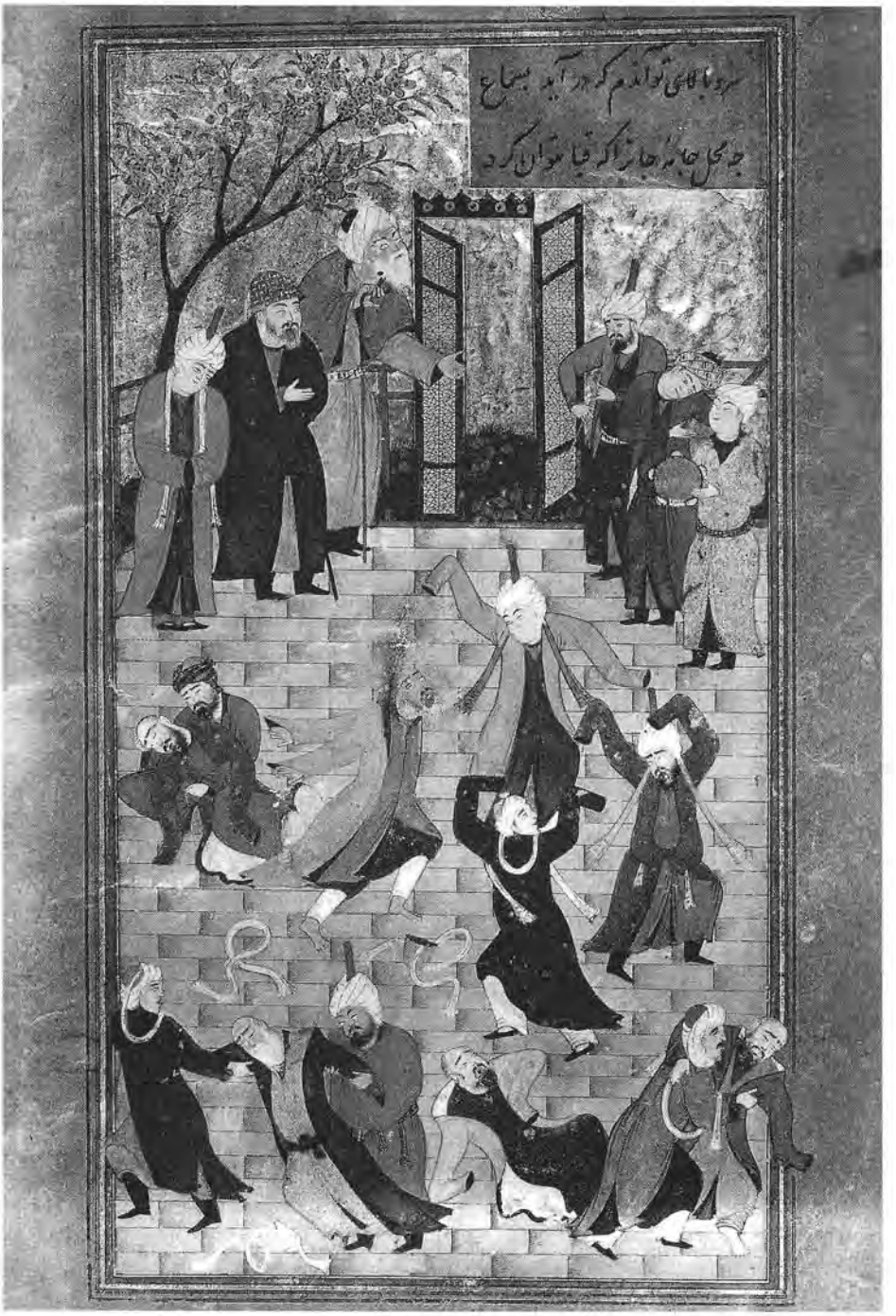


Figure 6.1 Dancing Sufis From the Divan (Book of poems) of Hafiz (Walters Art Gallery, Baltimore).

When I come up to that tower, when the curtain opens like that, as soon as I step up to approach Rangda, I see a strong fire coming from her eyes! I just—Oh! [slaps his hands together]—I do this! I just jump! I feel myself floating—because of the excitement! . . . Whenever they pick up the music [sings 'jangga jangga jangga'], whenever you sing that song [sings a bit of the gamelan piece that accompanies the trance] people just go crazy! I want to attack Rangda!!!' (I Wayan Dibia, interview, 1996)



**Figure 6.2** Gamelan musicians accompanying a Rangda/Barong ritual (photo by Judith Becker).

Neither the trance, nor the ceremony, nor the pacification of the witch can happen without the gamelan music (Fig 6.3). The habitus of listening of I Wayan Dibia upon encountering Rangda would have to include a description of not only the gamelan music and the presence of the witch, but a complex of beliefs about the negative forces of the universe, their effects upon human communities, the embodiment of these forces in the divine witch Rangda, and the methods by which she may be contained and controlled. His own emotion, culturally constituted but felt interiorly, is a necessary component of the maintenance of community well-being. His murderous passion has little to do with an interiorized self, with the identity of I Wayan Dibia. His rage, in part musically induced, is in the service of his community. Like the Wolof griots, musical emotion for the trancer in a Balinese Rangda exorcism is public, situational, predictable, and culturally sanctioned.

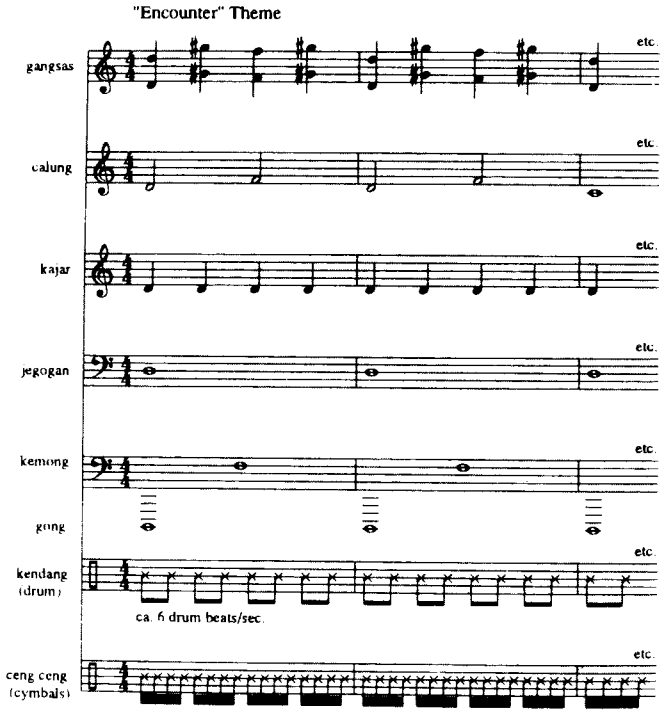



Figure 6.3 Musical transcription of the 'encounter' theme from a Rangda/Barong ritual.

### 6.5.3 Pentecostal arousal: 'music brought me to Jesus' (Cox 1995, p. 139)

Pentecostal religions were formally constituted at the beginning of the twentieth century, and are now rivalled only by Islam in their spread to all parts of the globe (Cox 1995, p. 15). Pentecostalism constitutes a faith that is dependent upon music to structure its religious services and to validate its system of beliefs by provoking intense emotional reactions within its most devote practitioners, leading them to 'testify'. To dance in The Spirit, to be possessed by the Holy Ghost, is demonstration that one is accepted into the congregation of those who will experience the final act of history, the reappearance of Jesus Christ and the establishment of the Kingdom of God on earth (Abell 1982). Music is the driving force for this emotional apotheosis. From softly played passages underlining a sermon or a prayer, to swinging, driving choruses sustaining a wave of religious emotionalism, music is rarely absent from the hours-long services. Pentecostal musical offerings shape a musical/emotional/religious arc that carries the congregation along with it. A service will begin with quiet, slow, soothing music: 'Music gets people in the attitude of worship. It helps them to forget outside influences and to focus on the Lord' (Interview with Jerry Trent, Church of God, Willow Run, Michigan, 1996).

Pentecostal churches may use a wide variety of musical instruments, often including piano, electric organ, synthesizer, guitars, and drumset to back up their driving, repetitive gospel hymns. As the music becomes louder, more rhythmic, and more repetitive, its driving quality supports, propels, and sustains the hand-waving, hand-clapping, foot-stomping choruses of 'Amen!' as the emotional temperature of the congregation gradually rises. High on the trajectory of the musical arc, worshippers may come forward to the altar to pray, and some may dance or trance. As in the Sufi services, the musical support will continue at a high intensity until all worshippers have worked through their transport and regained their normal composure. Religious ecstasy is always accompanied by great joy as it is a confirmation of the salvation of the worshipper. He or she has become a part of the historical narrative of millenarial Christianity and will join fellow believers and Jesus at the last day. The music never flags as some members are moved to tears, to dance, to quiver and jerk in the uncoordinated gestures of some religious trances. As religious passions subside, so does the music, until every last ecstatic has become quiet—exhausted and joyful. 'And it was terrific', exclaimed one worshipper at a service marked by intense, sustained, high-energy music, dancing, and trancing, 'and we really *got down* here. I mean we really *had church*' (Cox 1995, p. 268).

The musical listener at a Pentecostal service is expectant, alert, and waiting for the song leader, the musicians, the singing congregation, and the pastor to transport him or her to the point of experiencing the Holy Spirit in his or her own body. The attitude is one of surrender and openness. No critical irony, no mental busyness, or aesthetic evaluation can be allowed to interfere with a process of holy possession called 'Slain in the Spirit'.

Similar to the Sufi listener and the Balinese *bebuten* trancer, for a Pentecostal listener, the habitus of listening involves a scripted sequence of actions, emotions, and interpretations. Within each of these three scripts, musical, behavioural, and emotional events will occur within a certain predictable frame. Simultaneously, each individual event will be unique and non-repeatable. The listeners to a Wolof griot, an Indian sitar player, a group of *qawwali* singers, a Balinese *bebuten* trancer, or a Pentecostal worshipper have all developed habits of mind and body in response to specific musical events. These habits are acquired throughout their life experiences of interaction with others in similar situations. The emotions are private *and* public, interior *and* exterior, individual *and* communal.

## 6.6 Music, emotion, and 'being': conjoining anthropology, ethnomusicology, psychology, and neurophysiology

It seems clear to this author that a common ground needs to be explored between the more humanistic, cultural anthropological approach and the more scientific, cognitive psychological approach. I see the bringing together of the scientific and cultural approaches to the study of music and emotion as one of the great challenges of our fields. While the styles of argument and the criteria for evidence may remain distinct, the conclusions need to be comparable and not incommensurable. Both disciplinary areas may have to give up some of their established scholarly practices.

### 6.6.1 Single brain/body approaches

For those approaching the study of the emotional effect of musical listening on the brain, the unit of analysis is usually the brain or body of a single listener. From the early studies of music and the brain, such as Neher's early work in the 1960s (Neher 1961, 1962) analysing the brain rhythms of subjects listening to drum beats in a laboratory, to the more recent works of Clynes (1986), Wallin (1991), Peretz (this volume), and many others, musical scholars have adopted the scientific model and studied single brains or bodies in isolation. Studies in music cognition have largely followed the empirical models of cognitive studies in general in their efforts to model the algorithms of single brains processing music (Lerdahl & Jackendoff 1983; Raffman 1993). Neurophysiologists and psychologists of consciousness likewise follow the dictates of their scientific training, isolating and narrowing the problem as much as possible in order to have some control over the variables. In fact, some of the most exciting advances in the understanding of the mind are coming from scientists and philosophers who have focused on single minds/brains (Churchland 1986; Dennett 1991; Edelman 1992). The development of technologies such as magnetic resonance imaging (MRI) and positron emission tomography (PET) scanning have reinforced the practice of the study of single minds in the attempt to unravel the mysteries of conscious experience (see Peretz, this volume). The choice of a single human brain as the unit of analysis has larger implications as well. The emphasis upon the primacy of the self-contained individual has been a treasured credo of Western intellectual pursuits in all fields for many centuries. Contemporary examples include the cognitive sciences, Piaget's genetic epistemology, and Freudian psychoanalysis.

Without intending to denigrate any of these fields, and while fully acknowledging their enormous contributions, I am suggesting that the scripts of music and emotion, the habitus of listening, can be helpfully understood as a process which is supra-individual (Nunez 1997), in which the relationship between music and emotion needs to be understood as extending beyond the minds and bodies of single musicians and listeners, that is as a contextually situated social practice. Emotions relating to music are culturally embedded and socially constructed and can usefully be viewed as being about an individual within a community, rather than being exclusively about internal states. First-person descriptions of music and emotion are rife with tropes of interiority, yet the understanding of how music affects interiors takes place within consensual, shared views of what makes up 'reality'. Musical events set up an aural domain of coordination that envelops all those present.

### 6.6.2 Rhythmic entrainment

A familiar example of a changed interior, personal consciousness in a musical domain of coordination is the phenomenon of *rhythmic entrainment* (Chapple 1970, p. 38; Condon 1986; Hall 1977, p. 61; McNeill 1995, p. 6; Neher 1961, 1962; Rider & Eagle 1986; Vaughn 1990; Walter & Walter 1949). Bodies and brains synchronize gestures, muscle actions, breathing, and brain waves while enveloped in musicking. (The gerund 'musicking' is preferred for the same reasons as 'languaging'. The activity, the

process, the experiencing of music is foregrounded; see Small 1998.) Many persons, bound together by common aims, may experience revitalization and general good feeling. The situation is communal and individual, the music descends upon all alike, while each person's joy is his or her own.

> . . . it is no wonder that our words fumble when seeking to describe what happens within us when we dance or march. The initial seat of excitement is far removed from our verbal capabilities. It centers instead in those parts of the nervous system that function subconsciously, maintaining rhythmic heartbeat, digestive peristalsis, and breathing, as well as all the other chemical and physiological balances required for the maintenance of ordinary bodily functions. (McNeill 1995, p. 6)

### 6.6.3 Habitus of listening: culture and biology

An enactive approach to meaning and cognition based in biology has recently been proposed by Maturana and Varela (1987), Varela *et al.* (1991), Nunez (1997), and Foley (1997)—a group including biologists, linguists, and psychologists. While Bourdieu (and others) developed the analytical tool of habitus from an anthropological and sociological perspective, these scholars are reaching toward phenomenology to develop a comparable biological model which transcends the division between the cultural and the biological; between phenomenology and scientific materialism. The two approaches (Bourdieu and this group) are strikingly similar. While wishing to examine the relationships between music and emotion as supra-individual, one may still view the scripts associated with the habitus of listening as a biological process.

A biological process is a process that has an autopoietic organization (i.e. it is self-producing), a process that has boundaries between itself and the enveloping environment, the components of which (people) are dynamically related in a network of ongoing interactions that change the structure of the individuals as well as the structure of their interactions. Further criteria for biological processes are that they are autonomous, that is, that they can specify their own laws, and that they only allow changes to occur dependent upon maintaining the viability of their own structures (Maturana & Varela 1987, pp. 47–52). Groups of people who are focused on a common event and who share a common history of that event act, react, and to some extent think in concert, without sacrificing their bounded personal identities.

Viewing music and emotions as scripts that involve supra-individual biological processes, one can describe the changes that occur in these time sequences as having evolved over time through a historical process called 'structural coupling' (Maturana & Varela 1987, pp. 75–80). 'Structural coupling' describes the internal, structural changes that occur within an organism as a result of interaction with other organisms and with a world; changes of one's being or one's ontology. These changes become new domains of knowledge, knowledge gained through interactive behaviours, through doing. Music listeners as well as musicians undergo a learning process in which they imitate physical and mental gestures that ultimately transform both their inner structures as well as their relations to everything beyond the boundaries of their skins. Music and emotion are part of a larger processual event that includes many other people doing many other things while the whole event unfolds as a unity that

has been organized and reorganized over time by small structural changes within the participants.

We are accustomed to thinking of biological processes as extending from a single cell, to multicellular organisms such as a tree, to animals, and to a single human being. When groups of people are acting in some kind of accord, we tend to say that the phenomenon is social, or psychological, or political, but not biological. If we think of music making and music listening within a process that may also include quiet introspection or, conversely, include preaching, glossolalia, and dancing, or the murmuring of mantras, as all part of a biological process that has had a long history of structural coupling, or continual self-recreation, we can begin to think of the music, the emotion, and every other aspect of the process as contributing to bringing forth the activities of each other, as bringing forth a world or a reality in which certain emotions and actions are expected and appropriate, and in which the reality brought forth by all is enacted by all (Dreyfus 1991). A musical event is not just in the minds of the participants, it is in their bodies; like a vocal accent in speaking, emotion in relation to musical listening is personally manifested, but exists supra-individually. Each person, both musicians and listeners, seem to be acting as self-contained, bounded individuals, and indeed they experience whatever they experience as deeply personal and emotional, but the event as a whole plays itself out in a supra-individual domain.

> This embodied understanding, including dispositions in the habitus, is tacit knowledge and hence one can only be subsidiarily aware of it; this accounts for why so much of the habitus is preconscious and unable to be reflected on or modified. This conservatism leads to the practices generated by the dispositions of the habitus being transmitted from generation to generation, in other words they are potentially *cultural practices*. Culture in this view is that transgenerational domain of practices through which human organisms in a social system communicate with each other. These practices may be verbal or non-verbal, but they must be communicative in the sense that they occur as part of ongoing histories of social structural coupling and contribute to the viability of continued coupling . . . Culture, then, consists of the things people do to communicate in ongoing transgenerational histories of social interaction. (Foley 1997, p. 14)

Emotion is an enactment, not a representation in the mind. It is a way of *being-in-the-world*, not a way of thinking about the world.

> The experience of anything out there is validated in a special way by the human structure, which makes possible 'the thing' that arises in the description. This circularity, this connection between action and experience, this inseparability between a particular way of being and how the world appears to us, tells us that *every act of knowing brings forth a world* . . . All this can be summed up in the aphorism *All doing is knowing, and all knowing is doing.* (Maturana & Varela 1987, p. 26)

Language, music, and dance become one system of ontogenic coordination of actions. Together, they bring about changes in being and changes in the music event involved. While it is the individual who experiences the emotion, it is the group and its domain of coordinations that triggers the emotion. The changes in the neurophysiology of the listener are not attributable simply to the brain/body of a self-contained

individual. They occur through the group processes of recurrent interactions between co-defined individuals in a rhythmic domain of music.

Thinking of the relationship of music and emotion as a biological process with a co-defined, historically enacted ontology, as a group creation in which self-contained individuals have undergone structural changes through their interaction with other self-contained individuals helps to provide an embodied analysis of the relationship of music and emotion and the mysteries of musical affect. The domain of musicking and emotion is intrinsically social, visibly embodied, profoundly cognitive, and biologically consistent with the created domains of many other kinds of living creatures.

This type of analysis is aided by technologies such as film and video. One needs to be able to see everything at once, and to have the opportunity to replay the event. Qureshi's skilful score-like transcriptions of the multiple events at a Sufi religious ceremony demonstrate one method of researching music and emotion that does not focus on individual emotions, but on individual emotions within the context of many other individual emotions (Qureshi 1986, pp. 143–16). One can study her scores and literally *see* how the musicians interact with the congregants, and one can imagine how the emotions of one listener may have stimulated the emotions of another, leading to the ultimate reaction, an enactment of trancing (Fig 6.4).

### 6.6.4 What is gained by claiming that these processes are biological?

We may claim that special modes of conscious experience such as emotional responses to music listening, a habitus of listening, emerges as a phenotypic feature of humankind, as a biological development in the evolution of the species, emerging from both genetic and environmental influences. Music and emotion, viewed as evolving together in the interaction of each individual with performances throughout his or her life, dissolves intractable dichotomies concerning nature versus culture, and scientific universalism versus cultural particularism.

Furthermore, for those of us who want to use the word 'culture' but are inhibited by all the recent attacks upon the term relating to essentialism, historical stasis, and reified, abstracted psychology, culture can be restored as a more precise, more useful term. Culture (redefined) can be understood as a supra-individual biological phenomenon—a transgenerational history of social interactions that become embodied in the individual and transmitted through future actions.

No habitus of listening is entirely stable nor entirely fluid. As interactions change, so do the interiors of those who are interacting. We may have developed a habitus of listening that inclines us to introspectively contemplate the colourful fluidity of changing emotional states in response to Bach's *Musical Offering*, but if it would not disturb the viability of our own structures, we could learn to dance ecstatically in The Spirit with the Pentecostals.

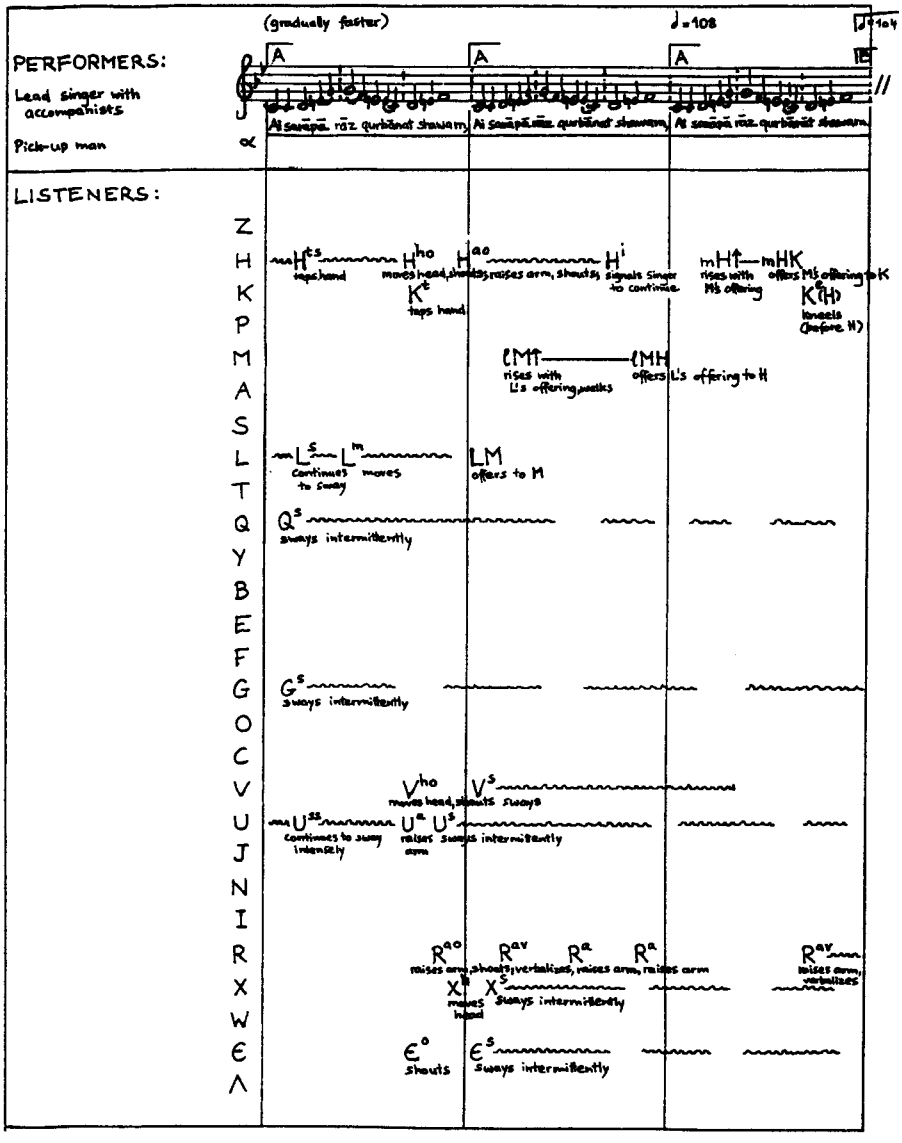



**Figure 6.4** Transcription (excerpt) of music/participant interactions in a north Indian Sufi ritual (from Qureshi 1986, p. 171).

## References


Abell, T. D. (1982). *Better felt than said: The Holiness-Pentecostal experience in southern Appalachia.* Waco, TX: Markham Press.

Appadurai, A. (1996). *Modernity at large: Cultural dimensions of globalization.* Minneapolis, MN: University of Minnesota Press.

Balasaraswati, T. (1985). The art of bharata natyam: a personal statement. In *Dance as cultural heritage, 2, Dance Research Annual XV,* (ed. B. T. Jones), pp. 1–7. New York: Congress on Research in Dance.

Barnard, M. (1995). Advertising: The rhetorical imperative. In *Visual culture,* (ed. C. Jenks), pp. 26–41. London: Routledge.

Baxandall, M. (1974). *Painting and experience in fifteenth century Italy: A primer in the social history of pictorial style.* New York: Oxford University Press.

Becker, A. L. (1995). *Beyond translation: Essays toward a modern philology.* Ann Arbor, MI: University of Michigan Press.

Becker, J. (1983). 'Aesthetics' in late 20th century scholarship. *World of Music,* 25, 65–80.

Benamou, M. (1998). *Rasa in Javanese musical aesthetics* (PhD dissertation: Musicology). Ann Arbor, MI: UMI Company.

Benveniste, E. (1971). Subjectivity in language. In *Problems in general linguistics* (trans. M. E. Meek), pp. 223–30. Coral Gables, FL: University of Miami Press.

Blacking, J. (1973). *How musical is man?* Seattle, WA: University of Washington Press.

Bourdieu, P. (1977). *Outline of a theory of practice* (trans. R. Nice). Cambridge, UK: Cambridge University Press.

Brennan, T. & Jay, M. (ed.) (1996). *Vision in context: Historical and contemporary perspectives on sight.* London: Routledge.

Brown, D. E. (1991). *Human universals.* Philadelphia, PA: Temple University Press.

Chapple, E. D. (1970). *Culture and biological man: Explorations in behavioral anthropology.* New York: Rinehart & Winston.

Churchland, P. S. (1986). *Neurophilosophy: Toward a unified science of the mind-brain.* Cambridge, MA: MIT Press.

Clynes, M. (1986). When time is music. In *Rhythm in psychological, linguistic and musical processes,* (ed. J. Evans & M. Clynes), pp. 169–224. Springfield, IL: C. C. Thomas.

Condon, W. S. (1986). Communication: Rhythm and structure. In *Rhythm in psychological, linguistic and musical processes,* (ed. J. Evans & M. Clynes), pp. 55–77. Springfield, IL: C. C. Thomas.

Coomaraswamy, A. K. (1957). *The dance of Shiva.* New York: Noonday Press.

Cox, H. (1995). *Fire from heaven: The rise of Pentecostal spirituality and the reshaping of religion in the twenty-first century.* Reading, MA: Addison Wesley.

Damasio, A. R. (1999). *The feeling of what happens: Body and emotion in the making of consciousness.* New York: Harcourt Brace & Co.

Davidson, R. J. (1992). Prolegomenon to the structure of emotion: Gleanings from neuropsychology. *Cognition and Emotion,* 6, 245–68.

Dennett, D. (1991). *Consciousness explained.* Boston, MA: Little, Brown & Co.

Dreyfus, H. L. (1991). *Being-in-the-world: A commentary on Heidegger's being and time, Division I.* Cambridge, MA: MIT Press.

Edelman, G. (1992). *Bright air, brilliant fire: On the matter of the mind.* New York: Basic Books.

Eiseman, F. B. (1989). *Bali: Sekala and niskala. Vol. 1: Essays on religion, ritual, and art.* Singapore: Periplus Editions.

Ekman, P. (1980). Biological and cultural contributions to body and facial movement in the expression of emotions. In *Explaining emotions*, (ed. A. Rorty), pp. 73–101. Berkeley, CA: University of California Press.

Feld, S. (1982). *Sound and sentiment: Birds, weeping, poetics, and song in Kaluli expression.* Philadelphia, PA: University of Pennsylvania Press.

Fish, S. (1980). *Is there a text in this class? The authority of interpreting communities.* Cambridge, MA: Harvard University Press.

Foley, W. A. (1997). *Anthropological linguistics: An introduction.* Oxford, UK: Blackwell.

Frith, S. (1987). Towards an aesthetic of popular music. In *Music and Society: The politics of composition, performance and reception*, (ed. R. Leppert & S. McClary), pp. 133–49. Cambridge, UK: Cambridge University Press.

Gamman, L. & Marshment, M. (ed.) (1988). *The female gaze: Women as viewers of popular culture.* London: Women's Press.

Geertz, C. (1973). The impact of the concept of culture on the concept of man. In *The interpretation of cultures*, pp. 33–54. New York: Basic Books.

Geertz, C. (1983*a*). Blurred genres: The refiguration of social thought. In *Local knowledge: Further essays in interpretive anthropology*, pp. 19–35. New York: Basic Books.

Geertz, C. (1983*b*). 'Native's point of view': Anthropological understanding. In *Local knowledge: Further essays in interpretive anthropology*, pp. 55–70. New York: Basic Books.

Geertz, H. (1974). The vocabulary of emotion: A study of Javanese socialization processes. In *Culture and personality: Contemporary readings*, (ed. R. LeVine), pp. 249–64. Chicago: Aldine.

Gergen, K. J. (1991). *The saturated self: Dilemmas of identity in contemporary life.* New York: Basic Books.

Ghazzali, A. H. M. (1991). *The alchemy of happiness* (trans. C. Field & Rev. E. Daniel). London: M. E. Sharpe.

Gnoli, R. (1968). *The aesthetic experience according to Abhinava Gupta.* Varanasi: Chowkhamba Publications.

Goffman, E. (1959). *The presentation of self in everyday life.* Garden City, NY: Doubleday/Anchor Books.

Goldhill, S. (1996). Refracting classical vision: Changing cultures of viewing. In *Vision in context: Historical and contemporary perspectives on sight*, (ed. T. Brennan & M. Jay), pp. 17–28. London: Routledge.

Goodenough, W. H. (1970). *Description and comparison in cultural anthropology.* Chicago: Aldine.

Grotowski, J. (1968). *Towards a poor theatre.* New York: Simon & Schuster.

Hall, E. (1977). *Beyond culture.* Garden City, NY: Anchor Books.

Hebdige, D. (1995). Fabulous confusion! Pop before pop? In *Visual culture*, (ed. C. Jenks), pp. 96–122. London: Routledge.

Irvine, J. T. (1990). Registering affect: heteroglossia in the linguistic expression of emotion. In *Language and the politics of emotion* (ed. C. Lutz & L. Abu-Lughod), pp. 126–61. Cambridge, UK: Cambridge University Press.

James, W. (1950). *The principles of psychology*, (Vol. 2). New York: Dover Publications (Originally published in 1890).

**Jefferson, T.** (1994). *Declaration of independence and the constitution of the United States of America: the texts.* Washington, DC: National Defense Press (Originally published in 1776).

**Jenks, C.** (ed.) (1995). *Visual culture.* London: Routledge.

**Johnson, J. H.** (1995). *Listening in Paris.* Berkeley, CA: University of California Press.

**Katz, R.** (1982). Accepting 'boiling energy': The experience of !Kia healing among the !Kung. *Ethos: Journal of the Society for Psychological Anthropology*, **19**, 348.

**Keil, C.** (1987). Participatory discrepancies and the power of music. *Cultural Anthropology*, **2**, 275–83.

**Kern, S.** (1996). *Eyes of love: The gaze in English and French painting and novels: 1840–1900.* London: Reaktion Books.

**Leach, E.** (1981). A poetics of power. *New Republic*, **184**, 14.

**Lerdahl, F. & Jackendoff, R.** (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press

**Levi-Strauss, C.** (1962). *The savage mind.* London: Wiedenfeld & Nicholson.

**Levy, R. I.** (1984). Emotion, knowing, and culture. In *Culture theory: Essays on mind, self, and emotion*, (ed. R. Shweder & R. LeVine), pp. 214–37. Cambridge, UK: Cambridge University Press.

**Lovejoy, A. O.** (1948). *Essays in the history of ideas.* Baltimore, MD: John Hopkins Press.

**Lovejoy, A. O.** (1964). *The great chain of being: a study of the history of an idea.* Cambridge, MA: Harvard University Press (Originally published in 1936).

**Lutz, C. A.** (1986). Emotion, thought, and estrangement: Emotion as a cultural category. *Cultural Anthropology*, **1**, 287–309.

**Lutz, C. A.** (1988) *Unnatural emotions: Everyday sentiments on a Micronesian atoll and their challenge to Western theory.* Chicago: University of Chicago Press.

**Lutz, C. A. & Abu-Lughod, L.** (1990). Introduction: emotion, discourse, and the politics of everyday life. In *Language and the politics of emotion*, (ed. C. A. Lutz & L. Abu-Lughod), pp. 1–23. Cambridge, UK: Cambridge University Press.

**McClary, S.** (1991). *Feminine endings: Music, gender, and sexuality.* Minneapolis, MN: University of Minnesota Press.

**McDaniel, J.** (1995). Emotion in Bengali religious thought: Substance and metaphor. In *Emotions in Asian thought: A dialogue in comparative philosophy* (ed. J. Marks & R. T. Ames), pp. 39–63. Albany, NY: State University of New York Press.

**McNeill, W. H.** (1995). *Keeping together in time: Dance and drill in human history.* Cambridge, MA: Harvard University Press.

**Masson, J. L. & Patwardhan, M. V.** (1969). *Santarasa.* Poona: Bhandarkar Oriental Research Institute.

**Maturana, H. & Varela, F. J.** (1987). *The tree of knowledge: The biological roots of human understanding.* Boston: New Science Library.

**Melzack, R.** (1992). Phantom limbs. *Scientific American*, **266**, 120–6.

**Merriam, A. P.** (1964). *The anthropology of music.* Evanston, IL: Northwestern University Press.

**Mitchell, W. J. T.** (1986). *Iconology: Image, text, ideology.* Chicago: University of Chicago Press.

**Morley, D.** (1995). Television: Not so much a visual medium, more a visible object. In *Visual culture*, (ed. C. Jenks), pp. 170–89. London: Routledge.

**Myers, F. R.** (1979). Emotions and the self: A theory of personhood and political order among Pintupi aborigines. *Ethos: Journal of the Society for Psychological Anthropology*, 7, 343–70.

**Needham, R.** (1981). Inner states as universals. In *Circumstantial deliveries*, pp. 53–71. Berkeley, CA: University of California Press.

Neher, A. (1961). Auditory driving observed with scalp electrodes in normal subjects. *Electroencephalography and Clinical Neurophysiology*, 13, 449–51.

Neher, A. (1962). A physiological explanation of unusual behavior in ceremonies involving drums. *Human Biology*, 34, 151–60.

Nettl, B. (1983). *The study of ethnomusicology: Twenty-nine issues and concepts.* Urbana, IL: University of Illinois Press.

Nunez, R. E. (1997). Eating soup with chopsticks: Dogmas, difficulties and alternatives in the study of conscious experience. *Journal of Consciousness Studies*, 4, 143–66.

O'Neill, J. (1995). Foucault's optics: The (in)vision of mortality and modernity. In *Visual culture*, (ed. C. Jenks), pp. 190–201. London: Routledge.

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions.* Cambridge, UK: Cambridge University Press.

Qureshi, R. B. (1986). *Sufi music of India and Pakistan: Sound, context and meaning in qawwali.* Cambridge, UK: Cambridge University Press.

Racy, A. J. (1998). Improvisation, ecstasy, and performance dynamics in Arabic music. In *In the course of performance: Studies in the world of musical improvisation*, (ed. B. Nettl), pp. 95–112. Chicago: University of Chicago Press.

Raffman, D. (1993). *Language, music, and mind.* Cambridge, MA: MIT Press.

Rider, M. S. & Eagle, C. T. (1986). Rhythmic entrainment as a mechanism for learning in music therapy. In *Rhythm in psychological, linguistic and musical processes*, (ed. J. Evans & M. Clynes), pp. 225–48. Springfield, IL: C. C. Thomas.

Rosaldo, M. Z. (1984). Toward an anthropology of self and feeling. In *Culture theory: Essays on mind, self, and emotion*, (ed. R. Shweder & R. LeVine), pp. 137–57. Cambridge, UK: Cambridge University Press.

Rouget, G. (1985). *Music and trance: A theory of the relations between music and possession* (trans. B. Beibuyck). Chicago: University of Chicago Press.

Russell, J. A. (1991*a*). Culture and the categorization of emotions. *Psychological Bulletin*, 110, 326–450.

Russell, J. A. (1991*b*). In defense of a prototype approach to emotion concepts. *Journal of Personality and Social Psychology*, 60, 37–47.

Schechner, R. (1988). *Performance theory* (rev. edn). London: Routledge.

Seckel, A. (ed.) (1986). *Bertrand Russell on God and religion.* Buffalo, NY: Prometheus Books.

Seeger, A. (1987). *Why Suya sing: A musical anthropology of an Amazonian people.* Cambridge, UK: Cambridge University Press.

Shankar, R. (1968). *My music, my life.* New Delhi: Vikas Publications.

Shepherd, J. & Giles-David, J. (1991). Music, text and subjectivity. In *Music as social text*, pp. 174–85. Cambridge, MA: Polity Press.

Shweder, R. A. (1985). Menstrual pollution, soul loss, and the comparative study of emotions. In *Culture and depression: studies in the anthropology and cross-cultural psychiatry of affect and disorder*, (ed. A. Kleinman & B.Good), pp. 182–215. Berkeley, CA: University of California Press.

Shweder, R. A. & Bourne, E. J. (1984). Does the concept of the person vary cross-culturally? In *Culture theory: Essays on mind, self, and emotion* (ed. R. Shweder & R. LeVine), pp. 158–99. Cambridge, UK: Cambridge University Press.

Slater, D. (1995). Photography and modern vision: The spectacle of 'natural magic'. In *Visual culture*, (ed. C. Jenks), pp. 218–37. London: Routledge.

**Small, C.** (1998). *Musicking: The meanings of performing and listening.* Hanover: Wesleyan University Press.

**Smith, J. C. & Ferstman, C.** (1996). Knowledge and the languaging body. In *The castration of Oedipus: Feminism, psychoanalysis, and the will to power,* (ed. J. Smith), pp. 52–79. New York: New York University Press.

**Solie, R. A.** (1993). *Musicology and difference: Gender and sexuality in music scholarship.* Berkeley, CA: University of California Press.

**Solomon, R. C.** (1984). The Jamesian theory of emotion in anthropology. In *Culture theory: Essays on mind, self, and emotion,* (ed. R. Shweder & R. LeVine), pp. 238–54. Cambridge, UK: Cambridge University Press.

**Spiro, M. E.** (1984). Some reflections on cultural determinism and relativism with special reference to emotion and reason. In *Culture theory: Essays on mind, self, and emotion,* (ed. R. Shweder & R. LeVine), pp. 323–46. Cambridge, UK: Cambridge University Press.

**Spivak, G. C.** (1988). *In other worlds: Essays in cultural politics.* New York: Routledge.

**Stanislavsky, K.** (1958). *Stanslavski's legacy: A collection of comments on a variety of aspects of an actor's art and life,* (ed. and trans. E. R. Hapgood). New York: Theatre Arts Books.

**Stone, R.** (1982). *Let the inside be sweet: The interpretation of music events among the Kpelle of Liberia.* Bloomington, IN: Indiana University Press.

**Taylor, C.** (1989). *Sources of the self: The making of the modern identity.* Cambridge, MA: Harvard University Press.

**Treitler, L.** (1993). Reflections on the communication of affect and idea through music. In *Psychoanalytic explorations in music, Second Series,* (ed. S. Feder, R. Karmel, & G. Pollock), pp. 43–62. Madison, CT: International Universities Press.

**Turino, T.** (1993). *Moving away from silence: Music of the Peruvian altiplano and the experience of urban migration.* Chicago: University of Chicago Press.

**Turner, V.** (1983). Body, brain, and culture. *Zygon,* **18**, 221–45.

**Varela, F. J., Thompson, E., & Rosch, E.** (1991). *The embodied mind: Cognitive science and human experience.* Cambridge, MA: MIT Press.

**Vaughn, K.** (1990). Exploring emotion in sub-structural aspects of Karelian lament: Application of time series analysis to digitalized melody. *Yearbook for Traditional Music,* 22, 106–22.

**Viswanathan, T. & Cormack, J.** (1998). Melodic improvisation in Karnatak music: The manifestation of raga. In *In the course of performance: Studies in the world of musical improvisation,* (ed. B. Nettl & M. Russell), pp. 219–33. Chicago: University of Chicago Press.

**Von Campenhausen, H.** (1964). *The fathers of the Latin church* (trans. M. Hoffman). London: Adam & Charles Black.

**Wallin, N. L.** (1991). *Biomusicology: Neurophysiological, neuropsychological, and evolutionary perspectives on the origins and purposes of music.* Stuyvesent, NY: Pendragon Press.

**Walter, V. J. & Walter, W. G.** (1949). The central effects of rhythmic sensory stimulation. *Electroencephalography and Clinical Neurophysiology,* 1, 57–86.

**Young, J. O.** (1999). The cognitive value of music. *Journal of Aesthetics and Art Criticism,* 57, 1–54.

**Zizek, S.** (1991). *Looking awry: An introduction to Jacques Lacan through popular culture.* Cambridge, MA: MIT Press.

**Zizek, S.** (1996). 'I hear you with my eyes'; or, the invisible master. In *Gaze and voice as love objects,* (ed. R. Saleci & S. Zizek), pp. 90–126. Durham, NC: Duke University Press.

CHAPTER 7

# AESTHETIC AGENCY AND MUSICAL PRACTICE: NEW DIRECTIONS IN THE SOCIOLOGY OF MUSIC AND EMOTION

TIA DENORA

## 7.1 The sociological approach

If psychology is concerned with human behaviour (Sloboda & Juslin, this volume), sociology is concerned with forms of social organization and the relation of these forms to social *action*, by which is usually meant, to paraphrase Max Weber (1978, Vol. 1, p. 4), conduct (real or hypothetical) that is meaningfully orientated in its course and outcome. The scope of material that falls within sociology's compass is wide, and perhaps it is because of its substantive breadth that sociology is a diverse—some would argue, fragmented—field.

Sociologists have traditionally employed a range of empirical strategies and theoretical perspectives, and have tended to engage in a high degree of intradisciplinary opposition such that is probably impossible to speak of 'a' sociological paradigm. For example, some sociologists seek to identify and quantify causes of collective and/or individual behaviour and to predict behavioural trends, while others seeks to understand meanings and meaning systems using ethnographic and qualitative methods of observation. The fissures within sociology are legion—*macro* (concern with large-scale institutions and trends) versus *micro* (focus on interactional process), *quantitative* versus *qualitative*, *structuralist* (focus on social forces and trends) versus *interpretive* (focus on meaning-making activity), to name a few. With that caveat, it is none the less possible to caricature the field by saying that sociology takes as its substantive focus a concern with action, intersubjectivity, the production of culture and knowledge, institutions, organizations and their conventions, and the implications of all these things for the social lives and life chances of individuals.

Whereas some psychologists might suggest that these topics are 'external' to the human organism, some sociologists, particularly those concerned with the sociology of the body, would counter that the nature/culture divide is ripe for dissolution (Birke 1992; Featherstone *et al.* 1991; Haraway 1985), to be replaced with a focus on how organismic features, cultural and material-cultural systems are reflexively linked.

Similarly, Foucauldian-inspired sociologists who have taken inspiration from post-structuralism and post-modern theory often suggest that our images of psychological 'interiority' owe much to post-Enlightenment discourses and discourse practices through which psychological subjects are constituted (Giddens 1991; Law 1994; Smith 1992). Put plainly, interior experience (here understood as the perception and self-perception of that experience) is shaped with reference to conventional, culturally, and historically specific categories that help to articulate and contain that experience. To be sure, these intra- and interdisciplinary conundrums can be explored most interestingly through the topic of emotion; adding music to the equation is even more useful because it allows for core sociological issues that centre around *agency* (here defined as a capacity for, and ability to, formulate action and experience) to be explored and advanced in a focused manner, in particular issues that concern agency's aesthetic and emotional dimension.

To that end, this chapter begins by considering the status of the emotions within sociology's paradigms. It moves from there to a discussion of music's position within sociology and focuses in particular on the changing conception of music's link to social life that has emerged within sociology of music over the past decade. From there, I survey classic works in the sociology of music as these illuminate music as a 'device' for the constitution of subjectivity and emotion in a range of settings. Throughout these discussions, my aim is to demonstrate how new approaches within the sociology of music help to highlight a congenital feature of sociology writ large, namely, the concern with social action, its sources, and its structures.

## 7.2 The rediscovery of emotion within sociology

There has been a longstanding interest within sociological theory in the emotions. One strand has developed out of Tönnies' (1957) concepts of *Gemeinschaft* and *Gessellschaft*. Tönnies used these terms to distinguish between social bonds built upon communal, emotional, traditional, and personal grounds versus those produced via rational and administrative procedures. The distinction, and with it a concern with the 'feeling' bases of social organization, has persisted and has established itself across a range of sociological perspectives. Max Weber's concern with the affective action (Weber 1978, Vol 1, pp. 24–8), Max Scheler's (1992) concern with the emotional flashes of insight that lead to knowledge (Remmling 1967, p. 33), Vilfredo Pareto's discussion of sentiments and non-logical aciton (Pareto 1963, p. 161), the 'human relations management' of Elton Mayo (Mayo 1933), Charles Horton Cooley's 'looking glass self' and its focus on emotion as the outcome of imaginative co-operation with external images (Cooley 1983), and Randall Collins' treatment of emotion in relation to rational choice theory (Collins 1993) have all highlighted emotion's role in relation to action and social structure.

And yet, despite this interest (and despite the formation of a thriving section on sociology of the emotions within the American Sociological Association), the affective dimension of human social being was, throughout the 1980s, mostly relegated to sociology's periphery. Particularly in the United States, where the discipline was dominated by rational actor models, and an attendant preoccupation with rule following, choice, and free will (Barnes 2000), the sociology of emotions was typically viewed as an off-shoot of 'micro' sociology and so disconnected from structuralist and 'macro' concerns. Within

sociology's dominant disciplinary frameworks, phenomena such as revolution and war, diplomacy, occupations, social and class rivalry, political and economic activity, organizational and institutional behaviour, the rise and fall of social movements, and the exercise of social control were all portrayed as if they took place in passionless corridors, executed by agents who possessed reason but who did not *feel*.

Over the past 5 years or so, more sensitive sociological portraits of action have begun to emerge. With them, the status of the emotions as a topic has been elevated. From a number of subdisciplinary directions has come a new concern with the 'feeling' component of social action. This emphasis can be seen in current approaches to the sociological study of social and political movements and the emerging focus on the affective character of identification with a movement (Melucci 1996*a*, 1996*b*) and 'structures of feeling' as these are entered into, adopted, and adapted in the course of identity politics and movement activity (Hetherington 1998). This emotion renaissance is also evident in new work on political affiliation, for example on the processes by which citizens transfer feelings of 'belonging' from nations to global entities (Berezin 1999). It has been further fuelled by a rapidly growing sociology of the body (Featherstone *et al.* 1991; Turner 1984) and embodied experience (Bourdieu 1998; Williams 1996, 1998).

Thus, sociology is now much more closely affiliated with concerns and topics traditionally lodged within the purview of social psychology. But this trend toward motivation and emotion has simultaneously been linked to sociology's abiding concern with social contexts—situations, occasions, institutions, and organizations—and with the hitherto 'macro' sociological terrain of collective and institutional action. The result is that sociology has begun to eradicate a bifurcation that has been both its hallmark and impediment for more than a century, namely, the distinction between 'micro' and 'macro' levels of analysis. In place of this division, prominent figures within sociology have moved toward a position that brings the feeling component of action—and hence some of the non-rational bases of social order—to the fore. As Alberto Melucci (1996*b*, p. 1) has expressed it:

> Each and every day we make ritual gestures, we move to the rhythm of external and personal cadences, we cultivate our memories, we plan for the future. And everyone else does likewise. Daily experiences are only fragments in the life of an individual, far removed from the collective events more visible to us, and distant from the great changes sweeping through our culture. Yet almost everything that is important for social life unfolds within this minute web of times, spaces, gestures, and relations. It is through this web that our sense of what we are doing is created, and in it lie dormant those energies that unleash sensational events.

The playing out of social change, politics, social movements, and relations of production are experienced and renewed from within this 'web', as Melucci calls it; it is from within the matrix of 'times, spaces, gestures and relations' that these 'larger' things are realized. Put differently, it is on the platform of the mundane and the sensual that the specificity of social experience is rendered and patterns are renewed.

A common thread running throughout nearly all of the recent focus on emotions within sociology has consisted of a focus on the interplay between emotions in lived experience and cultural forms. Inspired by the emphasis within post-structuralist and post-modernist theory on discourse and by the concurrent emphasis on texts and

artefacts as they imply readers, users, and 'subject positions' (particular and conventionally understood emotional stances, styles, or niches), sociologists across a wide range of specialist areas have devoted themselves to the question of how material-cultural and aesthetic media may be understood to provide models or candidate structures for the production and achievement of emotion and feeling within specific social settings. This focus has been clear within areas such as gender and sexuality and its concern with the cultural construction of desire and sexual practice (Jackson 1996, 1999), the study of spoken discourse in real time encounters (Frazer & Cameron 1989), organizational behaviour (Witkin 1995), and the sociology of material culture (Akrich 1991; Latour 1991; Law 1994; Moore 1997; Woolgar 1997).

It is from within this framework that the sociology of music can be seen to interact with the broader sociological focus on emotion. Having undergone significant transformation in the past decade, socio-music studies have moved from preoccupations with music's production to concerns with how it is consumed and what it 'does' in social life. In this regard, developments within the sociology of music merge with developments in the sociology of the arts more widely (Bowler 1994). Recent developments within the field have examined connections between musical consumption and musical experience as a means of producing and sustaining social ordering in real time, over the life course, and with reference to organizational and collective spheres of action. These developments have helped to return the sociology of music to the systematic concerns laid out early on in the area by figures such as Max Weber (1958), T. W. Adorno (1967, 1973, 1976, 1999), and Alfred Schutz (1964)—who is of particular interest because of his focus on music's relation to, and ability to, configure 'inner time' (pp. 172, 174). They have helped to renew the sociology of music's emphasis on fundamental questions concerning music's role as an active ingredient of social formation, its traditional and common sense role as the medium *par excellence* of emotion construction. In short, there is now a place at the top of sociology's table for the study of music.

Such a project involves considerable sweep: it encompasses matters that are usually desegregated—a focus on systems of music production and distribution, on organizational ecology and management, and on subjective experience as it is configured in relation to musical media.

## 7.3 Music and social structures: the movement toward conceptual symmetry

From the 1970s to fairly recently, the sociology of music and the sociology of culture more broadly have been concerned with how music is *produced*, with its occupational–politics–distribution systems, and with semiotic readings of musical forms. The emphasis was upon music as an object shaped by social relations. What remained tacit was any symmetrical attempt to conceive of music as a potentially dynamic medium, or to consider what music may 'do' in, to, and for the social relations in which it is embedded. Thus, only one half of the equation was present, albeit a crucial half, focusing on how musical works take shape in ways that are enabled and constrained by their cultures and worlds of production. As a result, apart from some notable exceptions discussed below, the matter of musical experience, and music's links to emotion, passion, and energy,

and to the phenomenological and existential features of social life, were all uncharted waters within sociology's paradigm. Even when sociologists considered music consumption, the focus was directed less to the matter of musical experience than to the ways in which tastes, musical values, and listening practices served as symbolic boundaries for status groups and status differences (Bourdieu 1984; Bryson 1997; DeNora 1995; DiMaggio 1982; Lamont 1992; Peterson & Simkus 1992).

As Peter Martin has observed (1995, p. 1), music is ubiquitous in modern societies, yet the significance of music's ubiquity often goes unnoticed within social sciences. In recent years, music's role as an active ingredient of social formation and subjectivity has been restored. Within this 'new' music sociology of emotions, it is possible to conceptualize music as a device for the constitution of emotive action in and across a range of social settings. This concern with *emotion construction* (through musical practice) offers an avenue into sociology's new concerns with the non-cognitive dimensions of agency (e.g. its emotional dimension) and also illuminates the relationship between culture and agency. In the remainder of this chapter, I seek to develop this perspective. I begin, in the next section, with reference to two, originally distinct, strands within the discipline that can be understood to provide the bedrock for more recent work: Adorno's theory of music as a medium of social organization and the more overtly empirical and ethnographic studies of music and lifestyle practice as propounded by Paul Willis in the 1970s.

## 7.4 Music, emotions, and subjectivity: sociology's classic models

For Theodor Adorno, music was inextricably connected to habits of mind, to social structural organization, and to modes of subjectivity. The music of Arnold Schoenberg, for example, he viewed as able to foster critical consciousness because its material organization went against the grain of musical convention and cliché. Conversely, in the age of 'total administration', music was also a medium that 'trains the unconscious for conditioned reflexes' (Adorno 1976, p. 53). Jazz, Tin Pan Alley, and nearly every other popular music genre was linked, in Adorno's view, to forms of regression and infantile dependency (Adorno 1990).

Adorno's work explores the idea that music interiorizes and is able to instigate forms of social organization in and through the ways that it works on and configures its subject-recipient. Put bluntly, Adorno conceived of music as active, indeed formative, in relation to consciousness. In this regard, his work makes some of the strongest claims on behalf of music's power in any discipline. But because Adorno provided no machinery for conceptualizing music's power empirically, his work is also frustrating; it never operationalizes the claim that music 'trains' the unconscious, nor points the way toward how these matters might be observed at the grounded level of musical practice. Adorno's work remained in the safe haven of theory and semiotic 'readings' of (primarily art music) works; it ventured but little into the realm of empirical analysis and the sphere of music consumption (indeed, Adorno's comments on music reception are mainly devoted to delineating a hierarchy of listener types). In its refusal to go out and look at the world, Adorno's work suffers at times from parochialism (e.g. his views on jazz, his unilateral condemnation of popular culture, and his tendency to prescribe musical taste).

During the 1970s, more empirically orientated work on 'music and society' provided an alternative to Adorno's hypothetical approach. The originally British tradition of cultural studies, ethnographically conceived, was directed at exploring connections between music and lived experience—identity creation, feeling, and the social constitution of embodied action and its parameters. Moreover, these explicitly ethnographic works may be read, with hindsight, to have presaged current developments.

Within the classic studies of young people and their intimate involvement with music, in books such as Paul Willis' *Profane culture* (1978), and Simon Frith's early monographs, *Sound effects* (1981) and *The sociology of rock* (1978), music's social presence was illuminated. Rereading these works with their characteristic focus on the experience of music, it is possible to see music providing a resource in and through which agency, identity, and peer culture are produced. Indeed, these studies can be seen to be compatible with Adorno's concern with music's link, and ability to instigate, social being. But this time, the music–social structure nexus was specified in a manner amenable to observation and in terms of how music was actually consumed and used. Music's structuring properties were understood as actualized in and through the practices of musical use, through the ways music was used and referred to by actors in their ongoing attempts to produce their social situations and themselves.

For example, in his report on the culture of the 'bikeboys', Willis noted that the boys' preferred songs were fast-paced and characterized by a strong beat, a pulsating rhythm. Willis showed how the boys established connections between music and social life. Structural similarities—homologies—between music and social behaviour were forged through the bikeboys' cultural practices and classifications.

Willis's focus was directed at the question of how particular actors make connections or, as Stuart Hall (1980, 1986) later put it, 'articulations' between music and social phenomena and how this process of linking music with forms of social life is part of actors' ongoing constitution of their life worlds and themselves. Although he never described his project as such, Willis's work on the bikeboys and on music's role as part of 'profane culture' more widely served to establish no less than an interactionist version of Adorno's original vision, that is an approach to the question of music's link to subjectivity, to cognitive style, and to actors' ontological orientations. Key, however, was that Willis's approach captured music's social impact as it was achieved by the ways in which actors themselves made 'articulations' or links between forms of music and forms of social life. It drew upon the basic interactionist tenet within sociology that agents attach connotations to things and orientate to social circumstances on the basis of perceived meanings. And it highlighted, through ethnographic description, the way in which social forms of organization are made through reference to aesthetic materials. This focus on the social process of world making (and the role of cultural media in this process) signalled an important shift in focus from aesthetic objects and their 'content' (static) to the cultural practices in and through which aesthetic materials were appropriated and used (dynamic) to produce social life. It simultaneously served to justify the need for 'symmetrical' conceptualizations of the relationship between music and social structure.

The bikeboys' music, as Willis noted, did not leave its recipients 'just sit[ting] there moping all night' (1978, p. 69). On the contrary, it invited, perhaps incited, movement.

As one of the boys put it, 'if you hear a fast record you've got to get up and do something, I think. If you can't dance any more, or if the dance is over, you've just got to go for a burn-up [motorcycle ride]' (1978, p. 73). Willis's work was thus pioneering in its demonstration of how music does much more than 'depict' or embody values. It portrayed music as active and dynamic, as constitutive not merely of 'values' but constitutive of trajectories and styles of conduct in real time. As one of the boys observed, 'you can hear the beat in your head, don't you . . . you go with the beat, don't you?' (1978, p. 72). This insight was germane to more recent developments in the field, for it points to the ways in which music, as it is heard or imagined, may serve to organize or 'configure' its users. Willis showed, for example, how music was implicated in the temporal structure of the bikeboys' evening, how it took them from—or provided the fulcrum for—one state (sitting around) to another (dancing as the music plays) to another (riding as the music plays in memory). Viewed in this way, music can be conceived of as a kind of aesthetic technology, or an instrument of social ordering.

It is precisely this interactionist conception that has been rediscovered by younger sociologists and sociologically minded scholars of music and subjectivity (it also lives on in the recent work of its pioneers; see Frith 1996; Willis 2000). For example, Sarah Cohen was one of the first to call for a renewal of an ethnography of musical practice in her insistence that 'focus upon people and their musical practices and processes rather than upon structures, texts or products illuminates the ways in which music is used and the important role that it plays in everyday life and in society generally' (Cohen 1993, p. 127). Similarly, Georgina Born argued, in her ethnography of IRCAM, that it is necessary to focus on 'the actual uses of technologies [she could just as well have said 'musics'], which are often depicted in idealised, unproblematic, and normative ways' (Born 1995, p. 15). In common with all instruments and technological devices, then, music needs to be understood in terms of its (non-verbal) capacities for enabling and constraining its user(s). Within current sociology, this focus has been directed at how individuals come to appropriate musical forms, and how musical forms feature within social settings as organizing materials of action. How, then, can this idea be developed and how can music's structuring powers be illuminated at the level of social experience? In the remaining sections of this chapter, I present two more examples of recent work concerned with music's role as a resource for the constitution of emotional experience. This discussion should be read in conjunction with the chapter by Sloboda and O'Neill (this volume) on emotions in everyday listening to music.

## 7.5 Music and emotion work: producing feeling in daily life

One of the most promising trends in current music sociology has been the focus on how music may be used—with varying degrees of conscious awareness by and on behalf of actors as a resource—to construct self-identity and to create and maintain a variety of feeling states. This work clearly connects with pioneering efforts within social psychology (De Las Heras 1997; Neilly 1995; Sloboda 1992, 1999) and ethnomusicology (Crafts *et al.* 1993; see also Becker, this volume). But it also speaks to an explicitly sociological concern with organizational aesthetics and with the social distribution of emotion and conventional notions concerning feeling structures within particular social

settings (for example, matters such as who is entitled to feel what type of emotion, where, when, and how).

The distinctively sociological quality of this concern is its focus on how individuals are produced or produce themselves as social agents with attendant styles of feeling and emotional stature. Thus, from within sociology, emotions are regarded as 'socially constituted' so that the study of music reception and consumption connects with recent discussions within sociology of 'aesthetic reflexivity' (Lash & Urry 1994, Chapter 4) and 'aesthetic agency' (DeNora 1997, 2000; Witkin & DeNora 1997). These terms highlight the consumption of aesthetic media as a means for self-interpretation and self-constitution, for the doing, being, and feeling that is the matrix of social experience. They simultaneously call for an overtly ethnographic turn in the study of aesthetic reception and thus interact with developments in the sociology of media and the plastic arts (Press 1994; Radway 1988; Tota 1997, 1999).

### 7.5.1 Music as a technology of emotion construction

Two recent studies illustrate this point. The first focuses upon music consumption in France (Gomart & Hennion 1999; Hennion & Maisonneuve 1998), the second on England and the USA (DeNora 1999, 2000). Both draw on roughly 50 in-depth interviews and both have concentrated on music's role in relation to the achieved character of feeling, to the ways in which actors' adoption of emotional stances, subject positions, and states are simultaneously the objects of actors' cultural practices in specific settings. In this regard, these works have begun to pave the way for a new sociology of music and emotions, one that conjoins an ethnomethodological focus (on the procedures and practical activities in and through which social orderings are accomplished in real time and in particular social contexts; Garfinkel 1967; Heritage 1984; Law 1994) with a focus on emotions in terms of how they are crafted and experienced within social situations.

In a piece that compares the love of music with the love of taking drugs, and which draws upon in-depth interviews with music lovers and with drug addicts, Gomart and Hennion (1999) examine actors as they engage in 'techniques of preparation' that produce forms of attachment so as to clarify the mechanisms that produce 'disposition'. In so doing, they shift the focus of their particular theoretical persuasion, 'actor network theory', from action to the study of interaction between people and things, and with this shift they make a concomitant move away from the overly general category of 'action' to the more specific concept of 'event'. They highlight 'what occurs' when, as part of the musical—or narcotic—experience, the self is 'abandoned' or given over to sensation and/or emotion. They are interested in no less than an ethnographic sociology of the production of passion.

With regard to the music lovers, they delineate the various practices involved in preparing for aesthetic experience. For example, Gomart and Hennion describe how their interviewees describe becoming 'ready in one's head' for the ear to hear and the body to respond (1999, p. 232), or how they employ particular listening strategies and rhythms so as to be ready to respond in preferred or expected ways. This process is akin to 'tuning in' or attempting to produce, through fine tuning practices, a signal's power and clarity. Listeners, they suggest, like drug users, 'meticulously *establish conditions* [emphasis in original]: active work must be done in order to be moved' (1999, p. 227).

Listeners are by no means simply 'affected' by music but are, rather, active in constructing their 'passivity' to music—their ability to be 'moved'. The music 'user' is thus deeply implicated as a producer of his or her own emotional response—that is, one who:

> strives tentatively to fulfil those conditions which will let him be seized and taken over by a potentially exogenous force. 'Passivity' then is not a moment of inaction—not a lack of will of the user who suddenly fails to be a full subject. Rather passivity adds to action, potentializes action. (Gomart & Hennion 1999, p. 243)

Gomart and Hennion were attuned to the question of how 'events' of musical passion and emotional response—the being 'taken over' by music—is reflexively accomplished by music lovers. Similarly, DeNora's study deals with music's role in the day-to-day lives of American and British women as they used music to regulate, enhance, and change qualities and levels of emotion. Nearly all of these women were explicit about music's role as an ordering device at the 'personal' level, as a means for creating, enhancing, sustaining, and changing subjective, cognitive, bodily, and self-conceptual states. Levels of musical training notwithstanding, the respondents exhibited considerable awareness about the music they 'needed' to hear in different situations and at different times, drawing upon elaborate repertoires of musical programming practice, and were sharply aware of how to mobilize music to arrive at, enhance, and alter aspects of themselves and their self-concepts. Part of their criteria for the 'right' music was how well it 'fitted' or was suitable for the purpose or situation they wished to achieve, or for achieving a particular emotional state (see the discussion of 'fit' in Sloboda and O'Neill, this volume). For example, as one respondent, Lucy, put it describing how she came to take time out to listen to some of the Schubert *Impromptus* (quoted in DeNora 2000, p. 16).

> I was feeling very 'stressed' this morning 'cause we're in the throes of moving house and its, you know, we're not, we haven't sold our house yet, and it's moving, you know, and, um, so I actively decided to put on Schubert's *Impromptus* because they were my father's favourite—you might want to come along to that again, because Schubert's *Impromptus* have a long history with my life—and I thought, my husband had just gone off to work and I thought well, about half an hour before I come up here [to her place of paid work], I'll just listen to them. So, the speakers are [she gestures] there and there on either side of what used to be the fireplace and I sit in a rocking chair facing them, so I get the sound in between the speakers, and I just sat there and listened [sighs, gentle laughter].
>
> Q. [gentle laughter]
>
> A. But I needed it. It was only ten minutes or so, you know, I didn't listen to them all. I just listened to the bits I wanted to listen to.'

Here, the music's powers to instigate emotion are constituted by Lucy herself; the music's power derives from the ways she interacts with it (Fig 7.1). The *Impromptus*, for Lucy, bear meaning on many levels. Their character as physical sound structures and their relation to a body of musical-stylistic convention is interpolated with, for Lucy, equally important biographical connotations and with a history of use such that the pieces calm her not only because they embody musical calm, but because they restore to Lucy a sense of her own identity. First, for Lucy, the works are associated with comfort; they are bound up with a complex of childhood memories and associations. Her late father, to whom she was close, used to play the piano after dinner and these works,

wafting up the stairs, were ones Lucy used to hear as she was falling asleep. Second, the material culture of listening is also an accomplice, in this example, of music's power to shift Lucy's mood on the morning she describes. Lucy's listening is conducted in a quiet room. She sits in a rocking chair placed between the speakers and so is almost nestled in the, as she perceives it, calm and nurturing music. (The vocabulary of nurturing is Lucy's. As she puts it, music 'soothes me', 'I retreat into music when I can't bear the rest of the world, you know', 'you can go into [music] and have it around you or be in it', quoted in DeNora 2000, p. 42.)

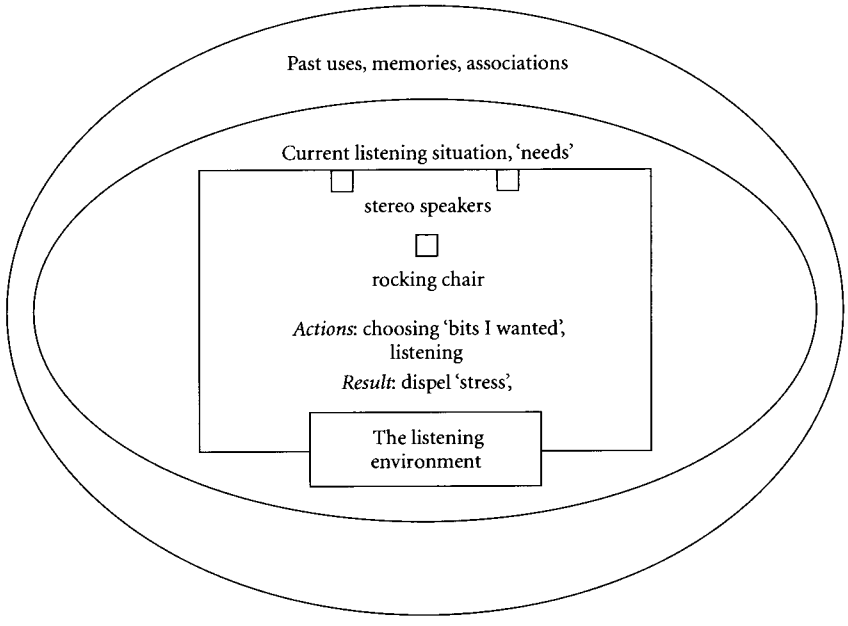



Figure 7.1 Ten to 15 minutes spent with music: 'Lucy' and Schubert's *Impromptus*.

This example typifies the kind of practical knowledge respondents exhibited repeatedly in their accounts of (often tacit) music practices. Through these practices, respondents produced themselves as coherent social and socially disciplined beings. As in the French data reported by Gomart and Hennion (1999), the British and American respondents routinely engaged in various practices of 'tuning in', of producing a musical event that would be capable of 'moving' them, and this production of passivity in the face of music, and its subsequent emotional 'effect' was achieved through an assemblage of musical practices. These included the choice of specific recordings, volume levels, material-cultural and temporal environments of listening (for example choosing to listen in bed, in a rocking chair, in the bath in the evening, in the morning, or while preparing to go out), and the pairing and compiling of musical works, memories, and previous and current contexts of hearing, such that the respondents could often be conceived as—and spoke of themselves as—disk jockeys to themselves (Fig 7.2).

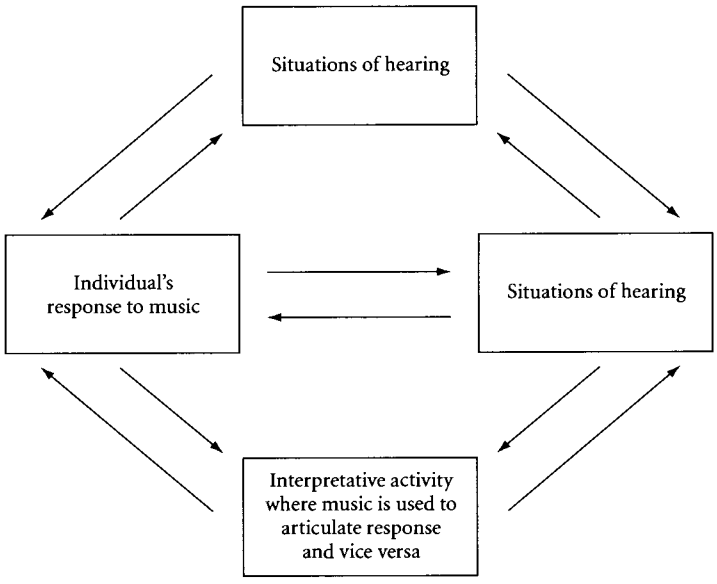



Figure 7.2 The reflexive constitution of emotional states or events through reference to music in social situations.

Thus, in both the French and the British/American studies, music can be seen as a medium or device for achieving 'emotional work'. First coined by Arlie Hochschild (1983) to describe practices within the workplace where workers were required to produce not only material goods or services, but also to produce themselves as types of emotional agents acting under the organizational cultural auspices (and bearing a strong resemblance to Cooley's conception of emotion constitution in his discussion of 'looking glass self', see above), the concept of emotional work has since been adopted more widely within sociology, where emotion is conceived as a 'bodily co-operation with an image, a thought, a memory—a co-operation of which the individual is aware' (Hochschild 1979, p. 551, quoted in Williams 1996, p. 129). Respondents described how they used music both routinely and in exceptional circumstances to regulate moods and energy levels, to enhance and maintain desired states of feeling and bodily energy (e.g. relaxation, excitement), and to diminish or modify undesirable emotional states (e.g. stress, fatigue).

For example, in their own words, the British and American women interviewed by DeNora described how music's specific properties such as rhythms, gestures, harmonies, and styles provided them with referents or representations of 'where' they wish to 'be' or 'go', emotionally, physically, and so on:

Like with my R&B, um, most of the time I listen to it when I'm, you know, trying to relax. I'm gonna sleep, sometimes I'll throw on a few tracks to wake me up, nice 'n slow and then I'll throw on, something else. And then, sometimes, you know, if I'm not really not in that relaxed mood,

I'm like, you know, 'I don't wanna listen to that' and I'll throw something fast on, or, something *fast is playing and I'm like 'that's too chaotic for me right now, I have to put something slow on'.* (Latoya, New York City, quoted in DeNora 2000, p. 49)

Respondents made, in other words, articulations between musical works, styles, and materials on the one hand, and desired modes of agency on the other, and then used music to presage, inspire, elaborate, or remind them of those modes of agency and their associated emotional forms. When respondents chose music as part of this care of self, they often engaged in self-conscious articulation work, thinking ahead about the music that might 'work' for them. And their articulations were made on the basis of what they perceived the music to afford. This perception is in turn shaped by a range of matters. Among these are previous associations respondents have made between particular musical materials and other things (biographical, situational), their understandings of the emotional implications of conventional musical devices, genres, and styles, and their perceived parallels (articulations/homologies) between musical materials/processes and social or physical materials/processes.

With regard to this last issue, the popular music scholar Richard Middleton (1990) has provided a useful bridge between sociology's focus on the articulations actors forge between music and feeling and musicology's focus on the ways in which temporal, rhythmic, and harmonic structures may lend themselves to particular interpretations, and how they may afford or instigate particular responses (see also Cook & Dibben, this volume). Middleton discusses what he refers to as 'exosemantic correspondences' between music and spatial or emotional structures/forms (e.g. slow and quiet: 'relaxed'; minor mode: 'depressed'), and develops semiotic concepts of 'primary' and 'secondary' signification (1990, pp. 220–32). The former refers to music's internal references and relationships, its arrangement of structural elements, and to the ways in which its relationships become terms in which to map, frame, and configure (in this case) emotions: '. . .as this note is to that note, as tonic is to dominant, as ascent is to descent, as accent is to weak beat (and so on), so X is to Y' (1990, p. 223). The latter refers to music's connotative level, its ability to invoke or symbolize images, emotions, and ideas.

While, as Middleton (1990) rightly emphasizes, music's material specificity comes to be implicated in emotion construction, music's capacity to engender emotion is by no means preordained. A characteristic of the perspectives on music and emotions discussed so far is their circumvention of, on the one hand, essentialist conceptions of subjectivity (as if feeling arises from internal processes) and, on the other, the notion that music is a 'cause' of emotional states. By contrast, these 'subject-object' dichotomies are bypassed by considering music as part of—a resource for—the reflexive constitution or working up of emotional states or events (see Fig 7.2). Here, music is not conceived as a 'stimulus' nor does it, *pace* Langer (1942), merely 'resonate' with some pre-existing subjective or mental state to which it exists in parallel. Indeed, this view, of which Langer is an exponent, has been subject to criticism from within sociology by Shepherd and Wicke (1997, pp. 12–14) who, though they endorse the idea that homologies exist or are forged between emotions and musical structures, suggest that the insights from musicology on the ways in which music's structures symbolize emotional structures do not 'stretch to an explanation of the mechanisms through which musical processes and processes of subjectivity are involved with one another' (1997, p. 13). In contrast with

musicological conceptions that confine analytical focus to, for the most part, the musical 'object', most of the sociologists discussed so far tend to conceive of music as providing a *resource* for emotional states and their achievement. Neither a cause nor a parallel of emotional states, music is conceived within sociology as providing a candidate simulacrum, or contrast structure, against which feeling may be formulated in real time. To speak of music in this way is to speak of it as a material against which 'how one feels' may be identified, elaborated, and made into an object of knowledge. One may, for example, say in reference to a musical example, '*this* is how I feel', and thereby grow tense and relax as the music does, when the music does, or in ways that are modelled, often without conscious awareness, on the music. It is here, then, that music's vital, and indeed, moral dimension—as declaimed by classical philosophers such as Plato ('in music that the guardians will build their guardhouse'; Plato 1966, p. 72) and Aristotle ('what we have said makes it clear that music possesses the power of producing an effect on the character of the soul'; Aristotle 1980, p. 24) can be seen: musical structures and representations are drawn, by actors themselves, into the vortex of feeling production; they literally 'get into' action and subjectivity.

In short, recent studies of music in relation to the achievement of emotional states and events points to music's use in real social settings as a device that actors employ for entraining and structuring feeling trajectories. Music is a resource to which agents turn so as to regulate themselves as aesthetic agents, as feeling, thinking, and acting beings in their day-to-day lives. Achieving this regulation requires a high degree of practical reflexivity. The respondents in both the French and the English/American case studies show how actors often perceive their 'need' for this regulation and know the techniques of auto-emotion work. These techniques may be found inadvertently (e.g. something is tried once and 'works' and so is later repeated and thus becomes part of an individual repertoire), or they may be suggested through culture and the media (and so be, at least initially, imitative), or they may be handed down by associates (and so exist as part of group or family culture), or initially encountered in a social setting—all of these technique locations were mentioned by respondents.

The natural history of the processes through which feeling states are identified and 'expressed' (i.e. enacted to self or other over time) is, moreover, a topic to be developed as it concerns the question of how aesthetic agency is configured in real time, as passion is choreographed and entrained. Holding on to this focus, but widening it from the individual experience of culture (and the social regulation of emotion in and through reference to cultural materials) to music's role in relation to the organization of collective action and its emotional component is but the next step within this programme. In the remainder of this chapter, I survey classic and recent work devoted to music's regulatory role in relation to emotional conduct between two actors engaged in intimate conduct, and between more than two actors, using the case of social movement activity.

## 7.6 Music and emotions in close interaction

Studies of gender and musical taste have long directed their focus to music's connections to love and to courtship. Shepherd and Giles-Davis (1991), for example, have reported on the consumption and use of music by a small number of middle-class

English-speaking teenage girls in Montreal, showing how music provides a kind of template for the formation of sexuality and emotional state. More recently, focus has shifted to the question of how music is actually used during intimate occasions and how it may draw bodies into temporal trajectories (rhythms, pulses, corporeal grammars, movement styles, and feeling grammars). Here, music is viewed as a non-verbal accomplice for particular intimate action forms (DeNora 1997), such that 'who puts what on the record player' as a prelude or backdrop for intimacy is overtly conceived as a question of love's real-time construction and its politics.

Many of the interviewees in DeNora's study (in particular the two younger age groups) described, through a rich series of examples from their daily lives, how they sought to structure both the grammar and style of intimate interaction through musical means and also to reproduce feeling styles and forms by using music that had served as the soundtrack on prior intimate occasions. The typical category for this activity was the concept of 'background' or 'mood' music and the interviewees were unanimous on one point: the music should not be fast paced, loud, or jolting. On the contrary, respondents spoke of how they valued the temporal structures and embodied practices commensurate with a slow pace, with leisured intimate activity. Their responses overlapped with what has already been observed in relation to gender and musical taste (Bryson 1997; Frith & McRobbie 1990; Peterson & Simkus 1992). 'Setting is very important' as one respondent put it, because getting the music 'right' is simultaneously a way of trying to make the action 'right', not merely in the embodied and technical sense, but as a way of prospectively calling out forms of emotional and embodied agency that are comfortable and preferable, that 'feels' right in emotional and embodied terms. (Such agency, of course, includes specific material practices, e.g. a 'light' touch or an 'unhurried' approach, and the music can be perceived as providing 'primary significations' of these things and hence employed as a model or blueprint for action):

> I think, last night, it was really funny, it was like 'mood setting' in a way. Cause, like, he had Enya on, and like, as people call that, 'chick music' [i.e. 'women's' music that young men may choose to play when entertaining women because they think it is what women prefer] . . . and he was trying to produce a relaxed atmosphere and um, I think in a way it does promote physical, or just intimacy in general because it's just like, certain music's more calming and, I remember like I think Stigma or Hyper came on and we were like, 'no no no, we don't want that!' and we like tried to get this piece, like I had him play the *First Night* sound track which I love, and there's like, a love song I, there, that's so beautiful, but everything else is like, 'bu bu de bah' [she sings here a triplet followed by a whole note the interval of a fifth higher than the triplet figure] and I'm like, 'no, no, this is not good' but I do, I think it was just very, it's very calming, very intimacy . . .
>
> Q. So it's part of what creates an intimate atmosphere?
>
> Melinda: Yeah, definitely. I think it's, setting is very important, and like music is a very big part of that. (quoted in DeNora 2000, p. 112)

## 7.7 Music and collective action: the case of social movements

Outside the sphere of intimacy, music can also be seen to serve as a model for emotional agency. One of the key areas where this perspective has been developed is

within the area of social movements, where new work has sought to draw together the sociology of emotions and the sociology of collective action. Taking inspiration from anthropologist Victor Turner's (1981) concern with 'action paradigms' (cultural materials such as narratives and imagery that may be used in ways that 'emplot' courses of action), Ron Eyerman and Andrew Jamieson (1998, p. 23) have attempted to specify how music may be understood to play an organizing role in the structure of collective action:

> To the categories of action discussed by sociologists we wish to add the concept of exemplary action. As represented or articulated in the cognitive praxis of social movements, exemplary action can be thought of as a specification of the symbolic action discussed by Melucci and others. The exemplary action of cognitive praxis is symbolic in several senses; but it is also 'more' than merely symbolic. As real cultural representations—art, literature, songs—it is artefactual and material, as well. What we are attempting to capture with the term is the exemplary use of music and art in social movements, the various ways in which songs and singers can serve a function akin to the exemplary works that Thomas Kuhn characterised as being central to scientific revolutions: the paradigm-constituting entities that serve to realign scientific thinking and that represent ideal examples of fundamentally innovative scientific work (Kuhn, 1970). The difference between culture and science, however, is that the exemplary action of music and art is lived as well as thought: it is cognitive, but it also draws on more emotive aspects of human consciousness.

As with the examples of individual feeling work discussed by DeNora (1999, 2000) and Hennion (Gomart & Hennion 1999; Hennion & Masisonneuve 1998), Eyerman and Jamieson (1998) describe how music may serve, within the context of social movement activity, as a prescriptive device of agency. Within music's structures, its perceived connotations, and its sensual parameters (dynamics, sound envelopes, harmonies, textures, colours, etc.), actors may 'find' or compose themselves as agents with particular capacities for social action.

For example, Eyerman and Jamieson (1998) describe how Todd Gitlin, a president of the Students for a Democratic Society (SDS) in the 1960s, described the SDS's identification with the music of Bob Dylan ('we followed his career as if he were singing our song; we got in the habit of asking where he was taking us next'; p. 116). In this sense, music may be conceptualized as akin to what Robert Witkin (1974) calls a 'holding form'—a set of motifs that proceed, and serve as a reference point for, lines of feeling and lines of conduct over time. Viewed as such, music plays a role that is similar to that of memory artefacts, as discussed by Urry (1996) and Radley (1990): musical motifs are orientated to for the ways they encapsulate and provide a container for what might otherwise pass as a momentary impulse to act, or a momentary identification of some kind. Holding forms thus provide a touchstone to which actors may return as they engage in collective expressive activity. They are the templates within which agency takes shape and to which actors may refer to renew themselves as types of agents and emotional agents.

Seen in this light, the study of how music is appropriated for emotion work illuminates a fundamental topic within sociology: it provides a way of exploring how social reality, with its forms and relations of feeling, are produced in real time and within specific social milieu. Critical then is the issue that held Adorno and which concerns

music's role in modern societies, the matter of how the aesthetic environments that come to afford agency's production are themselves produced. And so the study of consumption returns full circle to the study of production and dissemination as complementary enterprises. Especially with regard to the public spaces where agency is produced, music's role here has grown massively over the past two decades. If music is a device of social ordering, if—in and through its manner of appropriation—it is a resource against which holding forms, templates, and parameters of action and experience are forged, if it can be seen to have 'effects' upon bodies, hearts, and minds, *then* the matter of music in the social space is, as I have discussed above, an aesthetic-political matter. This is why matters of musical taste have, throughout history and across culture, often been contentious.

## 7.8 Conclusion

It seems fair to characterize current sociology of music as involved in two projects with regard to the emotions. The first is to produce portraits of music 'in action', that is, ethnographic documentaries of actors' musical practices and of the role these practices play in the production of emotion in real time and specific social settings. The second is to use these ethnographic reports to generate grounded hypotheses of how music functions in social life, what it does, and how it is reflexively implicated in the creation and recreation of social realms, events, and states.

With regard to this second aim, current work on music as a dynamic medium in social life has highlighted music's role as providing a structure or container for feeling, one whose specific properties contribute to the shape and quality of feeling as it is articulated and sustained in real time within and between individuals as part of the fabrication of ongoing social existence. In short, music is a resource for the practical constitution of entities we know as 'selves' and also for emotions and states that we refer to as 'intersubjective'. It is also a resource for the prospective structuring of action and a material that actors may employ as a referent or template as they elaborate and fill in, to themselves and to others, the modes of agency and subjective stances and identities that are structure-in-action. New work in the sociology of music and the emotions has been devoted to revealing music as it 'gets into' and is incorporated into action and to illuminating this process as it is achieved in and through actors' musical practices. This sociology of the 'musical composition' of subjectivity should be distinguished from more general post-structuralist or post-modernist theories of subjectivity and its 'cultural construction' that deal with emotions and subjectivity in monolithic ways and often leave little space for actors' creative and adaptive practices of emotion work. By contrast, the new sociology of music is devoted to a task that far exceeds assertions about music and the subject positions it may place on offer. It is concerned with specifying these matters at the level of action and action's 'heart' where actors are involved in the production of their social 'experience' and also in the mobilization of cultural and aesthetic materials for being, doing, and feeling. Anything less than this cannot address or begin to account for the mechanisms through which cultural materials 'get into' social psychological life and provide building materials of action.

## References


Adorno, T. W. (1967). *Prisms* (trans. S. Weber & S. Weber). London: Neville Spearman.

Adorno, T. W. (1973). *Philosophy of modern music* (trans. W. Blomster). New York: Seabury.

Adorno, T. W. (1976). *Introduction to the sociology of music* (trans. E. B. Ashby). New York: Seabury.

Adorno T. W. (1990). On popular music. In *On record: Rock, pop and the written word,* (ed. S. Frith & A. Goodwin), pp. 301–14). London: Routledge (originally published 1949).

Adorno, T. W. (1999). *Sound figures.* Stanford, CA: Stanford University Press.

Akrich, M. & Latour, B. (1991). A summary of a convenient vocabulary for the semiotics of human and nonhuman assemblies. In *Shaping technology/building society: Studies in sociotechnical change,* (ed. W. E. Bijker & J. Law), pp. 259–64. Cambridge, MA: MIT Press.

Aristotle (1980). *The politics.* Oxford, UK: Oxford University Press.

Barnes, B. (2000). *Understanding agency: the institution of responsible action.* London: Sage Publications.

Berezin, M. (1999). Emotions unbound: Feeling political incorporation in the new Europe. Paper presented at special session on Emotions and Macrosociology, American Sociological Association, August 1999.

Birke, L. (1992). In pursuit of difference: Scientific studies of men and women. In *Inventing women: Science, technology and gender,* (ed. G. Kirkup & L. Smith Keller), pp. 81–102. Cambridge, UK: Polity Press.

Born, G. (1995). *Rationalizing culture: IRCAM, Boulez, and the institutionalization of the musical avant-garde.* Berkeley, CA: University of California Press.

Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste* (trans. R. Nice). Cambridge, UK: Polity Press.

Bourdieu, P. (1998). *La domination masculine.* Paris: Seuil.

Bowler, A. (1994). Methodological dilemmas in the sociology of art. In *The sociology of culture,* (ed. D. Crane), pp. 247–66. Oxford, UK: Blackwell.

Bryson, B. (1997). 'Anything but heavy metal': Symbolic exclusion and musical dislikes. *American Sociological Review,* 61, 884–99.

Cohen, S. (1993). Ethnography and popular music studies. *Popular Music,* 12, 123–38.

Collins, R. (1993). Emotional energy as the common denominator of rational choice. *Rationality and Society,* 5, 203–30.

Cooley, C. H. (1983). *Human nature and the social order.* New Brunswick, NJ: Transaction (originally published 1902).

Crafts, S., Cavicchi, D., & Keil, C. (1993). *My music.* Hanover, NH: Wesleyan University Press.

De Las Hera, V. (1997). *What does music collecting add to our knowledge of the functions and uses of music?* Unpublished MSc Dissertation, Department of Psychology, Keele University.

DeNora, T. (1986). How is extra-musical meaning possible? Music as a place and space for 'work'. *Sociological Theory,* 4, 84–94.

DeNora, T. (1995). *Beethoven and the construction of genius: musical politics in Vienna 1792–1803.* Berkeley, CA: University of California Press.

DeNora, T. (1997). Music and erotic agency—sonic resources and social-sexual action. *Body and Society,* 3, 43–65.

DeNora, T. (1999). Music as a technology of the self. *Poetics: Journal of Empirical Research on Literature, the Media and the Arts*, **26**, 1–26.

DeNora, T. (2000). *Music in everyday life.* Cambridge, UK: Cambridge University Press.

DiMaggio, P. (1982). Cultural entrepreneurship in nineteenth-century Boston: The creation of an organizational base for high culture in America. Parts 1 and 2. *Media, Culture and Society*, **4**, 35–50, 303–22.

Eyerman, R. & Jamieson, A. (1998). *Music and social movements: Mobilizing tradition in the 20th century.* Cambridge, UK: Cambridge University Press.

Featherstone, M., Hepworth, M., & Turner, B. S. (ed.) (1991). *The body: Social process and cultural theory.* London: Sage Publications.

Frazer, E. & Cameron, D. (1989). On knowing what to say. In *Social anthropology and the politics of language*, (ed. R. Grillo), pp. 25–40. London: Routledge.

Frith, S. (1978). *The sociology of rock.* London: Constable.

Frith, S. (1981). *Sound effects: Youth, leisure, and the politics of rock 'n' roll.* New York: Pantheon.

Frith, S. (1996). *Performing rites: Evaluating popular music.* Oxford, UK: Oxford University Press.

Frith, S. & McRobbie, A. (1990). Rock and sexuality. In *On record: Pop, rock and the written word*, (ed. S. Frith & A. Goodwin), pp. 371–89. London: Routledge.

Garfinkel, H. (1967). *Studies in ethnomethodology.* Cambridge, UK: Polity Press.

Giddens, A. (1991). *Modernity and self identity.* Cambridge, UK: Polity Press.

Gomart, E. & Hennion, A. (1999). A sociology of attachment: music amateurs, drug users. In *Actor network theory and after*, (ed. J. Law & J. Hazzart), pp. 220–47. Oxford, UK: Blackwell.

Hall, S. (1980). Recent developments in theories of language and ideology: A critical note. In *Culture, media, language: Working papers in cultural studies 1972–79*, (ed. S. Hall, D. Hobson, A. Lowe, & P. Willis), pp. 157–62. London: Hutchinson.

Hall, S. (1986). On postmodernism and articulation: An interview with Stuart Hall. *Journal of Communication Inquiry*, **10**, 45–60.

Haraway, D. (1985). A manifesto for cyborgs: Science, technology and socialist feminism in the 1980s. *Socialist Review*, **80**, 5–107.

Hennion, A. & Maisonneuve, S. with Gomart, E. (1998). *Figures de l'amateur: Formes, objects et pratiques de l'amour de la musique aujourd'hui.* Paris: Ministere de la culture.

Hennion, A. & Meadel, C. (1989). Artisans of desire. *Sociological Theory*, 7, 191–209.

Heritage, J. (1984). *Garfinkel and ethnomethodology.* Cambridge, UK: Polity Press.

Hetherington, K. (1998). *Expressions of identity: Space, performance, politics.* London: Sage Publications.

Hochschild, A. (1979). Emotion work, feeling rules and social structure. *American Journal of Sociology*, **85**, 551–75.

Hochschild, A. (1983). *The managed heart.* Berkeley, CA: University of California Press.

Jackson, S. (1996). Heterosexuality as a problem for feminist theory. In *Theorizing heterosexuality*, (ed. D. Richardson), pp. 20–38. Milton Keynes, UK: Open University Press.

Jackson, S. (1999). *Questioning heterosexuality.* London: Sage Publications.

Lamont, M. (1992). *Money, morals, and manners.* Chicago: University of Chicago Press.

Langer, S. (1942). *Philosophy in a new key.* Cambridge, MA: Harvard University Press.

Lanza, J. (1995). *Elevator music: A surreal history of muzak, easy-listening and other moodsong.* London: Quartet Books.

Lash, S. & Urry, J. (1994). *Economies of signs and space.* London: Sage Publications.

Latour, B. (1991). Where are the missing masses? A sociology of a few mundane artefacts. In *Shaping technology/building society: Studies in sociotechnical change,* (ed. W. E. Bijker & J. Law), pp. 225–58. Cambridge, MA: MIT Press.

Law, J. (1994). *Organizing modernity.* Cambridge, UK: Polity Press.

McElrea, H. & Standing, L. (1992). Fast music causes fast drinking. *Perceptual and Motor Skills,* 75, 362.

Martin, P. (1995). *Sounds and society: Themes in the sociology of music.* Manchester, UK: Manchester University Press.

Mayo, E. (1933). *The human problems of an industrialized civilization.* Cambridge, MA: Harvard University Graduate School of Business.

Melucci, A. (1996*a*). *Challenging codes: Collective action in the information age.* Cambridge, UK: Cambridge University Press.

Melucci, A. (1996*b*). *The playing self: Person and meaning in the planetary society.* Cambridge, UK: Cambridge University Press.

Middleton, R. (1990). *Studying popular music.* Milton Keynes, UK: Open University Press.

Milliman, R. E. (1986). The influence of background music on the behavior of restaurant patrons. *Journal of Consumer Research,* 13, 286–9.

Moore, L. J. (1997). It's like you use pots and pans to cook with. *Science, Knowledge and Human Values,* 22, 434–71.

Moores, S. (1990). *Interpreting audiences.* London: Sage Publications.

Neilly, L. (1995). *The uses of music in people's everyday lives.* Unpublished undergraduate dissertation, Department of Psychology, Keele University, UK.

Pareto, W. (1963). *Treatise on general sociology.* New York: Dover.

Peterson, R. & Simkus, A. (1992). How musical tastes mark occupational status groups. In *Cultivating differences: Symbolic boundaries and the making of inequality,* (ed. M. Lamont & M. Fournier), pp. 152–86. Chicago: University of Chicago Press.

Plato (1966). *The Republic,* (ed. and trans. I. A. Richards). Cambridge, UK: Cambridge University Press.

Press, A. (1994). The sociology of cultural reception: notes toward an emerging paradigm. In *The sociology of culture,* (ed. D. Crane), pp. 221–46. Oxford, UK: Blackwell.

Radley, A. (1990). Artefacts, memory and a sense of the past. In *Collective remembering,* (ed. D. Middleton & D. Edwards), pp. 00–00. London: Sage Publications.

Radway, J. (1988). Reception study: ethnography and the problems of dispersed audiences and nomadic subjects. *Cultural Studies,* 2, 359–76.

Remmling, G. (1967). *The road to suspicion.* New York: Appleton-Century-Crofts.

Roballey, T. C., McGreevy, C., Rongo, R. R., Schwantes, M. L., Steger, P. J., Winninger, M. A., & Gardner, E. B. (1985). The effect of music on eating behavior. *Bulletin of the Psychonomic Society, 23,* 221–222.

Scheler, M. (1992). *On feeling, knowing, and valuing: selected writings,* (ed. and introduction H. J. Bershady). Chicago: University of Chicago Press.

Schutz, A. (1964). Making music together. *Collected papers. Vol 2.* The Hague: Martinus Nijhoff.

Shepherd, J. & Giles-Davies, J. (1991). Music, text and subjectivity. In *Music as Social Text,* pp. 174–87. Cambridge, UK: Polity Press.

Shepherd, J. & Wicke, P. (1997). *Music and cultural theory.* Cambridge, UK: Polity Press.

**Sloboda, J. A.** (1992). Empirical studies of emotional response to music. In *Cognitive bases of musical communication*, (ed. M. Riess-Jones & S. Holleran), pp. 33–46. Washington, DC: American Psychological Association.

**Sloboda, J. A.** (1999). Everyday uses of music listening: a preliminary study. In *Music, mind and science*, (ed. S. W. Yi), pp. 354–69. Seoul, Korea: Western Music Research Institute.

**Smith, D.** (1992). *Texts, facts and femininity: Exploring the relations of ruling.* London: Routledge.

**Tönnies, F.** (1957). *Community and society.* New York: Harper (originally published 1887).

**Tota, A. L.** (1997). *Etnografia dell'arte: Per una sociologia dei contesti artistici.* Rome: Logica University Press.

**Tota, A. L.** (1999). *Sociologie dell'arte: Dal museo tradizionale all'arte multimediale.* Rome: Carocci.

**Turner, B. S.** (1984). *The body and society.* Oxford, UK: Blackwell.

**Turner, V.** (1981). Social dramas and stories about them. In *On narrative*, (ed. W. J. T. Mitchell), pp. 137–64. Chicago: University of Chicago Press.

**Urry, J.** (1996). How societies remember the past. In *Theorizing museums: Representing identity and diversity in a changing world*, (ed. S. Macdonald & G. Fyfe), pp. 45–68. Sociological Review Monograph. Oxford, UK: Blackwell.

**Weber, M.** (1958). *The rational and social foundations of music.* Carbondale, IL: Southern Illinois University Press (originally published 1921).

**Weber, M.** (1965). *The sociology of religion* (trans. E. Fischoff). London: Methuen.

**Weber, M.** (1978). *Economy and society*, (ed. G. Roth & C. Wittich and trans. E. Fischoff *et al.*). Berkeley, CA: University of California Press.

**Williams, S. J.** (1996). The 'emotional' body. *Body and Society*, 2, 125–39.

**Williams, S. J.** (1998). Modernity and the emotions: Reflections on the (ir)rational. *Sociology*, 32, 747–69.

**Willis, P.** (1978). *Profane culture.* London: Routledge.

**Willis, P.** (2000). *The Ethnographic Imagination.* Cambridge, UK: Polity Press.

**Witkin, R. W.** (1974). *The intelligence of feeling.* London: Heineman.

**Witkin, R. W.** (1995). *Art and social structure.* Cambridge, UK: Polity Press.

**Witkin, R. W.** (1998). *Adorno on music.* London: Routledge.

**Witkin, R. W. & DeNora, T.** (1997). Aesthetic materials and aesthetic action. *Culture: The Newsletter of the American Sociological Association*, October, 1–5.

**Woolgar, S.** (1997). Configuring the user: Inventing new technologies. In *The machine at work*, (ed. K. Grint & S. Woolgar), pp. 65–94. Cambridge, UK: Polity Press.

**Yelanjian, M.** (1991). Rhythms of consumption. *Cultural Studies*, January, 91–7.

CHAPTER 8

# MUSIC AND EMOTION: PERSPECTIVES FROM MUSIC THERAPY

LESLIE BUNT AND MERCÉDÈS PAVLICEVIC

This chapter focuses on the relationship between music and emotion from within what is generally called 'improvisational music therapy' and aims to explore how a music therapist begins to generate and receive patients' emotional responses to the music in music therapy. We do not attempt to duplicate existing literature by reviewing the clinical and research literature of music therapy, nor do we provide an overview of the profession and the various applications of music as therapy. Instead, the chapter is divided into three main sections. We begin with some general background to music therapy and how connections with emotions are observed in clinical practice. Then follows a summary of a study that investigated how music therapists are able to judge the intended presentation of a series of basic emotions by listening to short improvisations played live by colleagues. This hopefully sheds light on the complex range of emotions within the more natural interplay of the patient's and therapist's improvised music. In the final section, we explore this world of co-created music with patients in music therapy by examining recent work on early infant development and elaborating on the connections between the musical and personal in improvisational music therapy using the theoretical concept of dynamic form.

## 8.1 Some background to music therapy

Music as a therapy provides patients of all ages with an effective means of exploring and communicating a wide range of emotions. As a contemporary discipline, music therapy has developed consistently over the last 50 years, with rapid strides over the last couple of decades in Europe, North and South America, and more recently in the Far East, Australia, New Zealand, and South Africa. Historically, the music therapy profession emerged from work with children and adults where difficulties in communication resulted from profound disabilities or mental health problems (e.g. Alvin 1975; Nordoff & Robbins 1971, 1977; Priestley 1975). A steady profusion of texts during the 1990s indicates how the range of work has developed in various settings: preschool centres and nurseries for children; special schools and units for children with varied learning difficulties; hospitals and special units for adults with learning difficulties, physical disabilities, mental health and neurological problems; hospital units and centres for older

people and for people with visual or hearing impairments; hospices and specialist centres for people living with terminal illness; and the prison and probation service (Aldridge 1996, 1999, 2000; Ansdell 1995; Bruscia 1991; Bunt 1994; Heal & Wigram 1993; Lee, 1995, 1996; Pavlicevic 1999*a*; Wigram & de Backer 1999*a*, 1999*b*; Wigram *et al.* 1995).

Examples of research into the clinical efficacy of music therapy with different populations are reported in these texts, as well as, for example, in the *British Journal of Music Therapy* and the US-based *Journal of Music Therapy*. These journals regularly present the results of both outcome and process-based studies, including, for example, studies on music therapy with preadolescents (Montello 1998), autistic children (Edgerton 1994), adult psychiatric patients (Heaney 1992; Pavlicevic & Trevarthen 1989, 1994), as well as the use of music therapy in paediatric oncology settings (Standley & Hanser 1995). Current research uses a proliferation of research techniques drawn from both quantitative and qualitative perspectives (for a review, see Wheeler 1995). Some researchers consider a qualitative approach as being more in keeping with the nature of music therapy practice (for a review, see Langenberg *et al.* 1996). Others propose that music therapy researchers need to continue using both, with different contexts and questions necessitating different perspectives (e.g. Smeijsters 1997). It is interesting that this debate has evolved out of a context of a more established clinical base, with practice preceding research.

### 8.1.1 Active and receptive approaches

Psychological responses to music are the bedrock of much that takes place in music therapy. People of all ages with wide-ranging health care needs come to music therapy and make fundamental connections between their emotions and the music that they experience either as active participants or as listeners. In music therapy, we can observe how music can be used to both express and arouse a wide gambit of emotions.

In an essentially active approach, patients are encouraged to articulate their emotions externally by forming musical gestures and structures. This process can be observed most clearly in improvised music making where the roles of composer and performer merge. Here there are also very close connections in the chain of communication between the originator of the music and the recipient. In a receptive approach, various emotions can be aroused while listening to pre-composed music played live by the therapist or on a recording—and this could include recording of previous improvisations.

Further levels of complexity are inherent in music therapy because emotions emerge alongside the facilitating presence of a music therapist who engages with the patient(s) in active musical involvement, by playing music to the patient(s), with the patient(s), or by listening to music together. Any of these scenarios finds a context within individual or group work. The overlapping links between music and emotions, therapeutic relationship, and the various needs of patients are featured in many definitions of music therapy, for example: 'Music therapy is the use of sounds and music within an evolving relationship between child or adult and therapist to support and encourage physical, mental, social and emotional well-being' (Bunt 1994, p. 8).

The complex interactions between the various musical and personal dynamics within music therapy can be illustrated with a quotation from Francis, a musician and patient

who lived with AIDS. His voice is a crucial part of Lee's (1996) extensive case study. Discussing the importance of his music therapy Francis said:

> I feel this is almost like a testament. It is the only expression I have of a spiritual journey. The only time that I feel that I am living and communicating. The only time that I feel I'm living the time I have left is when I am improvising. The rest of the time I'm wondering what the hell I'm doing. Time is a limited factor. When I'm doing these sessions with you I am actually living a moment; I'm actually living with somebody and producing something and revealing myself. When I'm improvising, when we are improvising together, I feel that I am saying something and living—and it's terribly important to me. (Lee 1996, pp. 77–8)

Francis talks here of the importance of living in the present moment of making music (cf. Sloboda and O'Neill, this volume). The music therapy sessions are a means for Francis to gain more understanding about his emotions, and to contribute towards a greater sense of self. It is important that his music is heard not in isolation but in the context of sessions that involve a relationship with the therapist. The quote from Francis relates in several ways to Christopher Small's use of the term 'musicking' (1999, p. 13):

> The act of musicking brings into existence among those present a set of relationships, and it is in those relationships that the meaning of the act of musicking lies. It lies not only in the relationships between the humanly organized sounds that are conventionally thought of as being the stuff of music, but also in the relationships that are established for the duration of the performance between the participants within the performance space.

Small observes how these relationships can model relationships in the world outside the performance space: 'relationships between person and person, between individual and society, between humanity and the natural world and even perhaps the supernatural world' (Small 1993, p. 13). Parallels can be drawn here with therapist and patients in improvisational music therapy: the musical relationship that develops between the players is personally intimate and relationally significant—for both therapist and client. Moreover, it has direct links with patients' lives outside the therapeutic space, offering patients opportunities to explore, extend, and gain meaning into their emotional lives, both within and outside the therapeutic relationship, and to make connections between their individual and social lives.

### 8.1.2 Musical material in therapy

The range of musical material used to generate a therapeutic relationship is, naturally, context-dependent. One day a music therapist may be facilitating a group improvisation that evolves naturally from a song of emotional significance sung by an elderly patient. The next day may see the same music therapist with a profoundly disabled child working with very elemental sounds or short phrases of music. Here, in addition to the limited musical gestures, any responses and communication of emotions may be via eye blinks, changes in breathing, or subtle shifts in body posture. The therapist's musical and personal response to, and matching of, a patient's musical and personal presentation is in keeping with a well-established principle used in music therapy, namely the 'iso-principle' (Altshuler 1954).

In improvisational music therapy, the improvised music emerging from a session with a group of young children with language problems and limited attention will be

qualitatively different from a session with a group of adults with mental health problems (and will also have different emotional significance). In terms of explaining the efficacy of music therapy, the group of language-delayed children may be so motivated to play the instruments and vocalize that their shifts in levels of excitement, quality of relationship, and emotional experiences and expressions may generalize outside of the sessions into further use of vocal sounds, gestures, and a corresponding rise in self-esteem. Likewise, the group of adults with mental health problems may be able to explore the shifting patterns of relating to one another and to the therapist in music and, with the therapist's help, draw analogies to more personal systems of interaction—both within and outside the music therapy setting.

Each of these scenarios presents a complex and multilayered generating, expressing, and conveying of emotional feelings and is dependent upon the music therapist's capacity to generate, judge, and read emotional responses to the music by patients within therapy. Understandably, the development of attentive listening, awareness, and empathy are some of the core qualities needed to work as a music therapist. But before exploring in more detail the various aspects of emotion in music therapy, we first make a brief detour to explore some of the recent thinking about music and emotion highlighted in this volume. Such thinking is potentially enriching to music therapy thinking and practice.

## 8.2 Sources of emotion in music therapy

Sloboda and Juslin (this volume) review different sources of emotion in music. There are examples of these meeting points between music and emotion in music therapy.

### 8.2.1 Associative connections

We are familiar with this connection, the 'Darling, they're playing our tune' notion (Davies 1978), in passive or receptive music therapy. Here, the music is a very powerful trigger that can set off a whole range of associations with specific events and places, memories of certain people, particularly if the experiences or people have been significant in a patient's life. The associations are generally highly individual, making it difficult to predict how different people will associate with different musical repertoires (Pavlicevic 1997). This necessitates much work on the music therapist's part to discover the musical preferences and history of the individual patient. This kind of emotional association is often the focus of work with elderly patients, when music can be used as a kind of reminiscence therapy (Odell-Miller 1995). It is also apparent in hospices and palliative care work where sensitive discovery of a patient's personal musical history is often a starting point of the treatment programme (Aldridge 1999; Lee 1995, 1996).

### 8.2.2 Iconic connections

Listening to a particular sound or series of sounds, whether improvised or pre-composed, can link musical characteristics to some external musical event or human feeling. For example, a group of children with learning problems can relate a general build-up of loud and fast sounds with a growing storm at sea or some other such high

energy event. Improvising to an external theme provides members of a group in a cancer unit, for example, with opportunities to channel and shift levels of energy, to begin to release deep-set levels of tension and to work together towards a common aim (Bunt & Marston-Wyld 1995; Burns *et al.*, 2001). Such a group may wish to improvise to a given theme or pair of emotions, for example to represent shifts from darkness to light, sorrow to joy, etc. Patients may decide to give a specific feeling-based label to an improvisation after hearing a recording of the piece. Not all patients are comfortable making such connections. Some patients are drawn to the very ambiguous nature of music, preferring to perceive the musical experience as an event in itself. They may acknowledge the emotional power of a musical improvisation or listening experience, but wish to contain the experience within the music and not wish to explore any external associations.

### 8.2.3 Intrinsic connections

There are many examples in music therapy practice—in both active and receptive approaches—when connections can be observed between the emotional experiences of the patients and both surface and deep structural aspects of the music. For example, we can observe how intrinsic features of the musical structures evoke different emotions in the specific listening approach, guided imagery in music (GIM):

> This does not happen in any simplistic one to one symbolic or causal correspondence between one musical gesture and the creation of an image but in an on-going organic and interactive way. In any series of GIM sessions we can observe similarities between different musical structures and processes and certain structures and processes in the imagery. At the same time of hearing the music the client is also creating an instantaneous link between image and music. It is as if the client is experiencing the music at different levels from the surface to the deep structural, reaching into the music to find the level that matches and resonates with the image. (Bunt 2000, p. 46)

GIM, created by Helen Bonny, involves listening to specifically programmed sequences of recorded music in a deeply relaxed state. The 'traveller' is invited to share the experience and any images or feelings with the therapist/guide (e.g. Bonny & Savary 1973; Bruscia 1995; Clark 1999; Erdonmez Grocke 1999). The images evoked by the music may be visual in nature but also body-based; they may open up a particular pattern of thought or various levels of emotional response. In many ways GIM demonstrates the interaction between the three kinds of connections discussed in this section. A particular melodic detail or surface musical detail may trigger a number of personal associations, memories, and images. Iconic associations may be made between aspects of the music and non-musical events. GIM therapists also witness patients making intrinsic connections between layers of musical expectations and deep structures and the forms and patterns of their internal emotions. The selection of the music used in GIM is indicative of the attention paid to the potential for these kinds of interactive connections.

One of the authors recalls a powerful emotional experience working through Bonny's programme entitled 'peak experience.' The sequence of music is: the 'Adagio' from Beethoven's Fifth Piano Concerto; the 'Et in Terra Pax' from Vivaldi's *Gloria*; the 'Adagio' from Bach's Toccata, Adagio and Fugue in C; the 'In Paradisum' from Fauré's *Requiem*, and the Prelude to Act I of Wagner's *Lohengrin*. A review of the written

transcript of the session indicates how certain moments in the music triggered personal associations, memories, and evoked external images. In addition to these links there were many intrinsic connections. The overall order of the pieces created a mood of expectancy and deepening feelings of awe and wonder. Body-based reactions—changes in breathing, shudders, tension in the stomach, or feelings of lightness, etc.—were closely entrained with the formal aspects of the music—phrasing, changes in pitch, sudden harmonic shifts, shifts in loudness levels, density of the textures, etc. The climactic moments of the Prelude to *Lohengrin* corresponded with the most illuminating and profoundly moving moments of the session, the powerful memory of which has proved long lasting and in many ways transformative.

As a link to the next two sections of the chapter it is worth recalling Bruscia's (1989) useful distinction between 'active' music therapy, generally understood as based on the playing of music—whether pre-composed or improvised—and 'receptive' music therapy, generally understood as listening to music. We have already discussed some receptive approaches and shall now comment further on the use of active music therapy.

## 8.3 Active music therapy: the example of improvisational music therapy

Improvisational music therapy is a generic term for those approaches in which the act of improvisation is understood as the locus of the therapeutic encounter between therapist and patient, in either individual or group work. Patients in music therapy improvise; the music therapist improvises. The participants in any improvisation make personal variations within the musical forms, the process of deciding what to play and how to shape musical events changing moment by moment. The music therapist values and attends to all the musical communications made by the individual or group member—there is no 'right' or 'wrong' way of playing. In the psychologically and physically safe context of a therapeutic setting, the different forms of improvisation can also help an individual to try out different aspects of relating both within the self (intramusical and intrapersonal) and between people (intermusical and interpersonal) (e.g. Bruscia 1987, 1989). We can view improvisation as a means of negotiating both old and new meanings.

For example, in a group therapy setting with adult patients, instruments such as drums, cymbals, marimbas, xylophones, guitars, and bongos focus the initial musical curiosity of a group of chronically ill patients referred to a music therapy group by the hospital's psychiatrist. The adults begin to explore the instruments and try out the different sounds. The therapist creates an initial sense of security by suggesting an opening improvisation based on a repeating steady pulse. Gradually, members of the group improvise patterns around the pulse. The group members grow in confidence as they explore freer kinds of improvisation. They pass musical messages to one another and play in small groups. The improvisations release a wide range of feelings and verbal exchange takes place in between the improvisations. After a lengthy discussion, the group decides to improvise on a theme, one that moves from despair to hope. Further discussion ensues exploring the emotions brought to the surface during the long group improvisation. A final improvisation adds a sense of closure for the group.

Music therapists respond instantaneously to the sounds that emerge in improvisations. How will the therapist be able to understand the sound gestures? How will the child or adult be able to use the instruments as a direct and authentic expression of different emotions? How can any communications be read, interpreted, and met or engaged in the therapist's authentic music? Can there be any sense of shared meaning? These are some of the challenges that daily confront a practising music therapist. The therapist will also bring to the music therapy room a range of previous experiences, both personal and musical, that will inform how such musical gestures are read. In addition, these gestures will most likely be framed within a particular philosophical, psychological, or therapeutic perspective which could include anthropological, behavioural, cognitive, developmental, humanistic, medical, musicological, psychodynamic, or transpersonal approaches (see overviews by Aigen 1998; Aldridge 1996; Ansdell 1995; Bruscia 1987; Bunt 1994, 1997; Pavlicevic 1997; Ruud 1980). A therapist working within a psychodynamic framework, for example, will be making use of the dynamics of transference and countertransference in understanding the various relationships between the patient and the music, the patient and the therapist, and, in group work, the patient and other patients (see Bruscia 1998; Priestley 1994).

We shall now focus in more detail on aspects of the generating and receiving of emotion through music in the realms of clinical improvisation and in Pavlicevic's development of her theory of dynamic form in music therapy.

## 8.4 Generating and judging emotion in short improvisations: a small study

A critical aspect to the generating and receiving of emotion in music therapy is the direct experiencing of emotion within the listener/player, as well as the recognition of emotion within the music itself.

### 8.4.1 Background to the study: improvisation and emotions

Research by Gabrielsson and Juslin (1996) which investigated the connections between the performer's expressive intention and the listener's experience was the inspiration behind a simple classroom exercise to explore the kind of instantaneous observations music therapists make to judge the intended emotional content of short improvisations. We explored how a limited number of emotions—happiness, sadness, tenderness, anger, and fear—were judged. These emotions have been regarded as 'basic' by many researchers (see Sloboda and Juslin, this volume).

Deciding to use short examples of improvised music played live by individual music therapists to the listening groups was an essential feature of this study. Would the music therapists playing the short improvisations be able to communicate the intended emotion? Would the therapists listening to the improvisations be able to judge the correct emotional intention? What musical parameters would the listeners use to decode the emotional expression of each improvisation? Behrens and Green (1993) have pointed out some advantages of using improvisations as compared to pre-composed music in judging the affective content of musical examples. There is no question of familiarity;

the emotional intention of each composer/improviser is known in advance; the length can be controlled, and each presentation can represent a complete musical idea. The use of live, as opposed to prerecorded, examples in our study added additional variables, namely the influence of visual and other non-verbal cues (for example, facial expressions, body movements, posture) in communication of the intended emotion.

### 8.4.2 Procedure

Two groups of music therapists took part in this study. The first was a group of thirteen music therapy students in their final stages of training, the second a group of eleven music therapy researchers and teachers. The student group was divided into two groups of six and seven and the researchers worked together as one group. The same procedure was used for both the student and researcher groups (referred to as S and R, respectively, in Table 8.1). Descriptions of the five emotions were written on cards and placed face down on a table. Individual therapists from both groups were invited to make a random selection of a card and to improvise freely within the prevailing emotion indicated for a period up to 2 minutes. The improviser could choose from within a wide range of tuned and untuned percussion instruments (e.g. xylophone, drum, cymbal, bells) as well as the therapist's own instrument (e.g. flute, piano). The therapist was asked to spend a few moments trying to access the particular emotion before selecting an appropriate instrument or group of instruments. Some therapists chose also to sing. The improvising therapist then played or sang to the listening group with the intention of expressing the emotion from the card, the title only being shown to the study's convenor.

After each improvisation, the listening music therapists were asked to rate each improvisation for the expression of each emotional state using the scale: 0 = improvisation does not express the state to 4 = improvisation does express the state (the same range as used in the Behrens and Green study). A score was required for each of the five emotions for each musical improvisation since overlaps were expected in decoding certain emotions. For later descriptive analysis, the listeners were also asked to list both musical and other reasons for their judgment of the main emotional state being communicated. There was a total of 26 improvisations (Table 8.1).

### 8.4.3 Results

The columns in Table 8.1 are used to designate the intended emotional expression of each improvisation and the rows the corresponding adjectives used by the listeners to rate each improvisation. The mean ratings of the respondents are used for each improvisation and total means are tabulated. *F*-tests were used to test for intended expression. The results were highly significant and provide strong evidence that the responses vary between descriptions.

Looking at the mean responses, clearly happy and angry descriptions correspond to happy and angry intended expressions. We can note that there were some overlaps in the ratings of certain emotions, for example sadness/tenderness and angry/fearful. The music therapists also listed the musical parameters and other observations that aided the process of decoding the intended emotions (Table 8.2).

**Table 8.1** Mean ratings for each improvisation and group totals

| Description | Intended expression | | | | |
|---|---|---|---|---|---|
| | **Happy** | **Sad** | **Tender** | **Angry** | **Fearful** |
| | S(a) **3.6** | S(g) 0 | S(j) 0 | S(p) 0.7 | S(r) 1.6 |
| | S(b) **4.0** | S(h) 0 | S(k) 0 | S(q) 0.5 | S(s) 0 |
| | S(c) **3.8** | S(i) 0 | S (l) 1.6 | R(e) 0 | S(t) 0.3 |
| Happy | S(d) **3.6** | R(b) 0 | S(m)**1.0** | | R(f) 0 |
| | S(e) **4.0** | R(c) 0.1 | S(n) 0.2 | | |
| | S(f) **4.0** | R(d) 0 | S(o) 0.2 | | |
| | R(a) **3.3** | | | | |
| | $t = 3.76$ | $t = 0.02$ | $t = 0.5$ | $t = 0.4$ | $t = 0.47$ |
| | 0 | **3.0** | 1.6 | 0.2 | 0 |
| | 0 | 1.0 | 2.2 | 0 | 0.2 |
| | 0 | **3.3** | **0.8** | 0.2 | 0.5 |
| Sad | 0 | **2.0** | 1.7 | | 1.0 |
| | 0 | **3.7** | 2.7 | | |
| | 0 | **3.1** | 2.2 | | |
| | 0 | | | | |
| | $t = 0$ | $t = 2.68$ | $t = 1.87$ | $t = 0.13$ | $t = 0.42$ |
| | 0 | 2.8 | **2.8** | 0 | 0 |
| | 0.2 | 3.2 | **4.0** | 0 | 0.6 |
| | 0.6 | 2.0 | **3.6** | 0 | 0 |
| Tender | 0 | 1.1 | **3.7** | | 0.5 |
| | 0.2 | **1.5** | **3.0** | | |
| | 0.2 | 1.8 | 3.6 | | |
| | 0 | | | | |
| | $t = 0.17$ | $t = 2.07$ | **t = 3.45** | $t = 0$ | $t = 0.27$ |
| | 0.2 | 0.2 | 0.2 | **3.5** | 3.0 |
| | 0.2 | 0 | 0 | **4.0** | 0 |
| | 0 | 0 | 0 | **3.8** | 3.0 |
| Angry | 1.0 | 0.1 | 0 | | 1.0 |
| | 0 | 0 | 0 | | |
| | 0 | 0.6 | 0 | | |
| | 0.4 | | | | |
| | $t = 0.26$ | $t = 0.15$ | $t = 0.03$ | $t = 3.77$ | $t = 1.75$ |
| | 0.2 | 1.8 | 1.0 | 1.2 | 1.0 |
| | 0 | 1.2 | 0.6 | 0.8 | **4.0** |
| | 0 | 2.0 | 0 | 1.0 | 3.7 |
| Fearful | 0 | 2.1 | 0 | | 3.5 |
| | 0 | 0.4 | 2.2 | | |
| | 0 | 1.8 | 0.4 | | |
| | 0 | | | | |
| | $t = 0.03$ | $t = 1.55$ | $t = 0.7$ | $t = 1.0$ | $t = 3.05$ |

R, research improviser; S, student improviser; *t*, total mean for each series of improvisations. The bold type indicates the correct judgment of intended emotion.

**Table 8.2** Various parameters used by music therapists in decoding basic emotions

*Happy*
Tempo and rhythm: fast, flowing tempi; lively, skipping, dotted rhythms; firm pulse
Pitch and melody: high pitches, rising melodies
Tonality and harmony: major tonality, clear structured harmony
Timbre, texture and style: bright textures and timbre, much staccato
Loudness level: no major shifts, mainly middle ground
Phrasing and structure: clear, predictable, many short 'flourishes'
Non-verbal communication: smiling, free body movements, foot tapping, bouncing, clapping, upright posture with no apparent tension, a 'twinkle in the eye'
Other: feeling of abandonment, much communication of high levels of energy, meeting cultural expectation of feeling happy

*Sad*
Tempo and rhythm: slow pulse (absence at times), gentle rhythms
Pitch and melody: wandering, hesitant and unfocused melodies
Tonality and harmony: minor tonality, small intervals, some dissonances, no sustained tension
Timbre, texture and style: sustained, legato, use of single sounds, flat timbre with blurring of sounds, slow vibrato
Loudness level: predominantly soft, small range from *pp* to *mf*
Phrasing and structure: falling phrases
Non-verbal communication: little movement, flat expression, slow breathing, concentration in the eyes, head hanging low
Other: feeling of being unresolved, general look of sadness and introspection

*Tender*
Tempo and rhythm: gentle, slow pulse, lilting rhythms
Pitch and melody: gradual rise and fall to the melody, frequent glissandi, lyrical
Tonality and harmony: major and minor alternation; use of modes, supportive accompaniment with broken chords and gentle arpeggios
Timbre, texture and style: a light touch and texture, sustained
Loudness level: soft
Phrasing and structure: legato, use of pauses
Non-verbal communication: stillness, enclosed posture, head hanging low
Other: communication of relaxed and reflective state

*Angry*
Tempo and rhythm: quick tempi, jerky rhythms, sudden changes in rhythmic patterns with accents
Pitch and melody: low sounds, unpredictable or absence of melody, clusters
Tonality and harmony: minor tonality, atonality, feeling incomplete
Timbre, texture and style: fragmented sounds; strident, harsh and percussive timbres
Loudness level: very loud, distorted sounds, strong and forceful
Phrasing and structure: some repetitions, absence of order, shape and form
Non-verbal communication: aggressive body movements
Other: much construed from strong level of energy and shape of opening musical gesture

*Fearful*
Tempo and rhythm: rapid and jerky tempi and rhythms
Pitch and melody: low sounds, minor tonality or lacking a sense of tonality
Tonality and harmony: random and tense dissonant harmonies, strong chords and clusters
Timbre, texture and style: fragmented, heavy textures
Loudness level: sudden bursts of loud playing building to climaxes
Phrasing and structure: pauses creating feeling of unrest, general lack of organization, no clear pattern or shape
Non-verbal communication: jumpy movements, weighty feeling to the body
Other: cultural stereotypes, e.g. tremolandi; sense of confusion, chaos, hesitation and disconnection; fear read in players' faces

### 8.4.4 Discussion

Clearly, the music therapists who played or sang these short improvisations were able to communicate the intended emotions to the listeners. The listening music therapists were able for the most part to judge accurately the intended emotional expression. The detailed musical parameters were used almost instantaneously as evidence before selection of the main intended emotion was made: evidence preceding the inference. The improvisers were able to communicate some of the very basic emotions: happiness, sadness, tenderness, anger, and fearfulness. These are very common features of many musical presentations made by patients of all ages. An interaction was observed between tender and sad, alerting music therapists to the need to be particularly cautious in jumping to quick judgments with these and other more than basic emotions. We may describe an improvisation as tender and can pick up the general feeling, but it is hard to be more specific. The students found close links between both sad and tender and between angry and sad. These mixed responses are a feature of the interaction effects of even the most basic of emotions (for further elaboration, see Juslin, this volume). There were also discrepancies picked up when a player appeared to find it difficult to articulate a particular emotion. To some extent, this study was also a judgment on how well each improviser was able to communicate the intended expression.

The musical parameters and other behaviours used to describe how each intended emotion was judged indicate the subtle interaction between musical and other cues. Clearly, visual cues are highly relevant in adding meaning to an emotional gesture. It would be interesting to repeat the exercise with the listeners closing their eyes. Music therapists also often enter into the musical world of their patients by playing music with them. This provides a far closer and resonant means of relating directly with the expression of a specific emotion. The emotion may then be embodied in the joint action of interactive music making. The therapeutic situation also differs from the classroom setting of this study in the fact that a more spontaneous and natural expression of emotions can be observed.

In the next section of this chapter, the concept of dynamic form in music therapy is presented, exploring the personal/emotional and musical/emotional interface in clinical improvisation.

## 8.5 Dynamic form

### 8.5.1 Understanding improvisation in music therapy

Explaining the emotional nature of improvisation in improvisation music therapy models (Bruscia 1989) presents complex issues for music therapists and music psychologists because of the musical, psychological, and relational dimensions of the improvisatory act. Music therapists generally understand the spontaneous act of 'improvisation' as being 'more than'—or 'other than'—music improvisation as an 'artistic' event; and as having personal and interpersonal significance. Such improvisation is labelled 'clinical' and is defined by the Association of Professional Music Therapists (UK) as: 'Musical improvisation with a specific therapeutic meaning and purpose in an environment facilitating response and interaction' (quoted in Odell-Miller 1988, p. 54).

At the same time, however, the act of improvisation in music therapy may well draw stylistically from the 'musical' repertoire (for example, the therapist, being a skilled musician, may well use idioms from jazz, blues, modes, boogie-woogie, romantic or baroque music, or free improvisation). This duality of the 'personal' and 'musical' in clinical improvisation has given rise to much debate between therapists, especially with regard to whether or not verbal interpretations of psychological material are necessary, and how words can explain the musical act itself (e.g. Aigen 1999; Ansdell 1995; Pavlicevic 1997; Streeter 1999).

As well as drawing directly from allied theoretical frameworks, as stated earlier, music therapists also describe and explain clinical improvisation using a variety of discourses, including those from behavioural theory (Madsen *et al.* 1968); musicological theory (Ansdell 1997; Lee 1989, 1990); psychoanalytic theory (Brown 1999; Priestley 1994; Robarts 1996; Streeter 1999; Tyler 1998); the medical literature (Aldridge 1996); psychology (Aigen 1999; Bunt 1994; Nordoff & Robbins 1971; Pavlicevic 1997); anthropology (Ruud 1998*a*, 1998*b*); and semiotics (Styge 1999).

### 8.5.2 A brief scenario

Andrew, a teacher suffering from depression as a result of a family bereavement, arrives for a first therapy session and is presented with a room full of various percussion instruments. He has no formal musical training, and last played music at school. He has been referred to music therapy by his general practitioner, to help alleviate Andrew's depression. At the therapist's invitation, Andrew chooses to play an African Conga drum, and begins to beat it somewhat tentatively, erratically, using both hands in a rather clumsy manner. The music therapist listens closely to what Andrew is playing, and hears the tentativeness in his rhythmic patterns, the dull timbre, the slow and erratic pulse, the uneven dynamic level: somewhere between mezzo-piano and mezzo-forte, with the occasional sforzando thud.

Soon she begins to play with him, on the vibraphone, meeting his playing on the Conga. Even though Andrew's beating is erratic, she can detect some fundamental pulse, which she reflects in her own spontaneous playing. She also matches his dynamic level and reflects aspects of his rhythmic patterns, attempting to establish a therapeutic relationship through musical sounds. Soon after beginning to play with Andrew, the dorian mode comes into the therapist's mind, influenced, perhaps, by an aspect of their joint rhythmic forms which remind her of a folk tune. She begins to improvise in this style, while at the same time continuing to meet Andrew's tempo, rhythm, dynamic level, and so on.

To an uninformed listener, their joint playing might well sound like a 'musical event'. Music therapists, however, are able to distinguish between music therapy improvisation (in which the emphasis is on the interpersonal, emotional communication between the players) and 'non-clinical' music improvisation (in which the emphasis is on the musical content and processes). Indeed, an exploratory study reports that 'blind' music therapist listeners asked to identify music therapy and non-music therapy improvisation are able to identify each of these correctly. Moreover, they report being able to 'hear therapeutic thinking' when listening to music therapy improvisations. In contrast, they describe non-clinical improvisation as being musically driven (Brown & Pavlicevic 1997). But how is

this distinction to be understood? What is its theoretical foundation? The next section offers a brief explanation for understanding clinical improvisation as a personal–relational act between therapist and patient: one that elicits and generates emotion.

### 8.5.3 The musical as personal

During Andrew's opening moments on the Conga drum, the therapist listens to two things—as one. As a musician, she groups the musical elements of rhythm, pulse, dynamic level, pitch, and timbre, creating a musical event in her listening mind. Based on this, a melody enters her mind in the *dorian* mode, which she begins to play on the vibraphone: this is the 'musical' aspect of the therapeutic event. The other listening is a 'personal' one—and, more specifically, an emotional listening to Andrew. Here, it is the qualities of Andrew's spontaneous playing—that is, the *erratic* pulse, the *tentative* rhythmic patterns, the *dull* thudding of the drum, the *unsustained* dynamic—that reveal Andrew to her as a person (rather than as a musician). When the therapist joins in with his playing, it is with the intention of meeting him directly, as a person (rather than meeting only his music)—through musical sounds. She meets his erratic, tentative, unsustained playing—without, initially, attempting to 'rectify' it in the musical-aesthetic sense—immediately establishing emotional rapport with him. This is the 'interpersonal' and 'musical' duality of clinical improvisation: an event that instantly interfaces musical form and basic expressive qualities. This duality has gained rich understanding by drawing from the psychological literature that has explored the innate musicality of infant communicative behaviour—explored briefly below.

### 8.5.4 Infant musicality and mother–infant communication: a blueprint for understanding emotion in music therapy improvisation?

Psychological literature on early infant development has shown that newborn infants are neurologically predisposed to identifying, and responding to, contours and rhythms of movements, gestures, and vocalizations in their mothers' gestures, vocal sounds, and facial expression (Malloch 1999; Robb 1999). In addition, they 'tune in' to subtle shifts in vocal timbre, tempo, and volume variations, and with their mothers negotiate and share a flexible musical pulse between them, constantly adapting their tempo, intensity, motion, shape, and contour of their sounds, movements, and gestures in order to 'fit' and to communicate with one another (Beebe *et al.* 1985; Papoušek 1989; Papoušek & Papoušek 1991; Stern 1985; Trevarthen 1993, 1999; Trevarthen & Aitken 1994). Mother and infant develop and share a rich musical 'code' that has interactive significance (Papoušek 1996).

These sensitivities can be understood as an innate susceptibility to 'musical' patterning, which is critical for the infant's biological survival in terms of communicating and expressing fundamental needs such as hunger, distress, discomfort, frustration, and contentment to its mother (or primary caretaker). Thus, infants receive, elicit, and respond to their mothers' movements, gestures, and acts, not as musical or temporal events or 'behaviours' in themselves, but rather, as being *personally* expressive and communicative; that is, as the basis for forming human relationships (Papoušek 1996). Adults too seem to intuitively fine-tune to infants with nuances of rhythm, tempo, and

intonation in 'infant-directed' speech, which contrasts with the 'musical' nuances of 'adult-directed' speech (Trehub & Trainor 1993; Unyk *et al.* 1992). Adults seem to show an innate capacity for enabling and inviting the infant to enter into interactional synchrony with them: that is, as adults we do not need to be 'trained' or to 'learn' how to communicate with babies. The concept of 'interactional synchrony' (Bernieri & Rosenthal 1991; Brown & Avstreih 1989; Condon & Ogston 1966) explains what can be seen as a dance between persons, in which each adapts and shifts eye gaze, head movements, body movements, and vocal sounds in order to personally 'fit' and 'exchange' with the communicating partner (Feldstein & Welkowitz 1978).

Stern's concept of 'vitality affects' (Stern 1985), as well as other mechanisms of non-verbal communication are offered in support of this understanding. Stern explains 'vitality affects' as the qualities of shape or contour, intensity, motion, and rhythm—'amodal' properties that exist in our minds as dynamic and abstract, not bound to any particular feeling or event—enabling us to 'make sense' of the world. Thus, a 'bursting' into tears (or laughter), a bursting watermelon, a musical sforzando, or an athlete's final 'burst' of speed may well share—in our minds—the same 'vitality affects' in terms of intensity, motion, contour, and rhythm. Qualities of, for example, surging, fading away, fleeting, or being drawn out, also underpin the way that we move, talk, gesture, and express feelings, and their presence in our minds enables us to receive the acts and gestures of another as a unified whole—and as carrying emotional meaning.

The concept of 'vitality affects' also explains our capacity to 'recognize', in pre-composed music, what Langer (1942, 1953) calls the 'morphology' of feelings: we hear the swooping, brittleness, tentativeness, and increases in intensity because these very qualities exist in our minds as 'amodal' forms, common both to music and to feelings. From a music therapy perspective, Stern's 'vitality affects' can be understood as being essentially 'musical' in character, realized in clinical improvisation in music therapy (Pavlicevic 1990, 1995, 1997). Because 'vitality affects' are abstract forms that can be expressed through any of our acts (e.g. we may talk, walk, play music, and show anger or joy in a quickly shifting, 'loud', and 'fast' manner), and because 'vitality affects' can be understood as essentially 'musical' in character, then we can consider spontaneous musical acts in music therapy to be personal—as well as musical (for further discussion of 'vitality affects', see Sloboda & Juslin, this volume). But how does improvisation in music therapy 'work' as an interpersonal, emotional event?

### 8.5.5 Improvisation as an interpersonal event: dynamic form in clinical improvisation

The concept of dynamic form represents this duality of the musical and the personal in music therapy improvisation. Whilst not denying or diminishing the *musical* basis of the therapeutic event in music therapy (after all, clinical improvisation frequently draws from the *musical* stylistic repertoire), dynamic form is grounded in Stern's 'vitality affects'. However, it is more than the direct realization of 'vitality affects' through sound. Rather, dynamic form is at the same time musical *and* relational (or interpersonal) in character. This distinguishes it from Langer's (1942, 1953) morphology of feelings in music, and from the deeply intrapersonal experience of listening to, or performing pre-composed music. Its communicative and relational agility also distinguishes it from free improvisation—which resonates primarily with stylistic and grammatical rules of

music. Dynamic form is elicited within, and as part of, the clinical–musical relationship, using the communicative and expressive mechanisms of basic emotions, that is, those of intensity, contour, tempo, rhythm, timbre, and dynamics. Dynamic form is generated by both therapist and patient, who relate intimately and directly with one another through jointly created sounds (Pavlicevic 1990, 1997, in press). These musical sounds correspond to the mechanisms of 'non-verbal communication', enabling the therapist to 'receive' and directly (and viscerally) experience patients' musical utterances as a presentation of themselves—and also as a presentation of their clinical pathology (Pavlicevic & Trevarthen 1989; Pavlicevic *et al.* 1995). Adult psychiatric patients suffering from conditions that interfere with their capacity to express and communicate emotion—and form relationships with others—portray this collapsed relational capacity in clinical improvisation. Both self and interactional synchrony can be regarded as measures of health with illness and disease occurring when a person becomes out of synchrony—both physically and mentally—with the self and the surrounding environment (Capra 1983; Pavlicevic 1991).

Although clinical improvisation may be heard as musical, the therapist's skill is not solely a 'musical' one, as has been assumed (Schögler 1998). Rather, music therapists' skills lie in their capacity to interface the personal and music in music therapy improvisation; to 'read' music therapy improvisation as an interpersonal event (in the way that mothers and babies read one another's acts not as musical or temporal, but as emotionally expressive and communicative); and to support, develop, and extend the jointly created improvisation according to personal and therapeutic, rather than musical–aesthetic, needs and dictates. It is this direct knowing by one person of the other that enables the therapist to gain a direct experience of the patient, uncluttered by words—and replicating, to some extent, the psychotherapeutic dyad, in which the relationship between therapist and patient is itself the fertile ground for the patient's self-growth and knowing.

If we now return to the opening moments of Andrew's clinical improvisation with the therapist, we can think of each of them as presenting themselves through spontaneously created sounds, using the very ingredients of non-verbal communication—which are essentially musical in character. In the first moments of playing, Andrew portrays aspects of his 'vitality affects' through the musical relationship, using the very modality that directly captures, and portrays to another person, the essence of who he is. However, the jointly created clinical improvisation (once the therapist joins in with him on the vibraphone) elicits much more than this. By joining in with him, the therapist provides the basis for an interpersonal therapeutic relationship, enabling Andrew to express himself, to release deep-seated feelings to which he may, as yet, have no awareness, and to communicate through a powerful, non-verbal, medium: that of music. This musical experience bypasses (and complements) his 'talking' about his feelings—and enables him to gain new insights about himself—which may, or may not, need to be interpreted through words. This direct therapeutic musical event will hopefully enable him to resolve, and be released from, his feelings of loss, and to help him shift from his depressed state (Champion 1992; Pavlicevic 1999*b*). Although, like most music therapy patients, Andrew has no idea about how clinical improvisation 'works', his innate, biological, and neurological preverbal musicality—present from birth—is engaged by the

therapist's clinical musicianship: enabling and inviting his sounds to become personal and interpersonal.

Critical, here, is that not just any musical improvisation will do. The therapist needs to elicit dynamic form between herself and Andrew. Were they simply to play music together, however spontaneously, their focus would be on the sounds, tunes, and rhythms of their music, rather than on playing themselves. Through his depression, Andrew's 'vitality affects' have become narrowed and somewhat stuck: this is revealed in his playing which is erratic, tentative, and somewhat clumsy. His depression means that his experience of the world is slow, cloying, dark, and unshifting. These qualities are experienced directly by the therapist when she plays with him, and she then works with the dynamic form, using clinical techniques to stabilize his beating, extend his tempo, introduce rubatos, accelerandos, diminuendos and crescendos, and so on. This is not in order to make him into a better musician, but rather, to grow his expressive, emotional vocabulary, to 'loosen him up' emotionally, and to help him to access the feelings underlying his depression. At this point, music therapy may (or may not) move into verbal interpretation, with the therapist helping Andrew towards gaining insight into his life, depression, and bereavement. However, many music therapy clients are unable to speak because of their conditions, and even with verbal clients like Andrew, their work might be essentially in the musical event—the gaps between meaning and emotion in clinical improvisation and in words are far from being convincingly explicated in the music therapy literature (Ansdell 1996).

However, a case for the emotional richness of clinical improvisation is made in the final part of this chapter: a transcribed musical excerpt of dynamic form shows, through a brief microanalysis, the interactive mechanisms of communication and relationship between the therapist and client. Neither of their playing can be taken in isolation from the other's—hence the insistence of dynamic form being a feature of their joint playing, rather than of simply the patient's in clinical improvisation.

### 8.5.6 A musical example

Figure 8.1 shows 19 s of an improvisation by a music therapist (T) with a depressed patient (P) who is an adult amateur musician. The therapist and patient are seated opposite one another with a marimba between them, taking turns to play—hence the one melodic line. The instrument faces the patient, so that the therapist is playing 'upside down'. Although the excerpt shows the two players reflecting aspects of one another's rhythmic motifs (the excerpt has a compound rhythm), melodic line, dynamic levels, and timbre, the duration of the therapist's eleventh statement (T11) is 300 cs, while the patient's eleventh statement of 602 cs is discrepant in duration. This discrepancy exists despite both T11 and P11 comprising ten notes each, and despite the patient's other musical features (i.e. melodic contour, dynamic level, timbre), which clearly show an awareness by the patient of the therapist's preceding statement. The discrepancy is caused by the patient delaying the fourth note of her statement (i) as well as the conclusion of her statement at (ii). The delays within P11 'unbalances' the improvisation, resulting in an asynchronous interaction—this is somewhat perplexing if we consider that the patient is an amateur musician, but not unusual for someone suffering from depression.

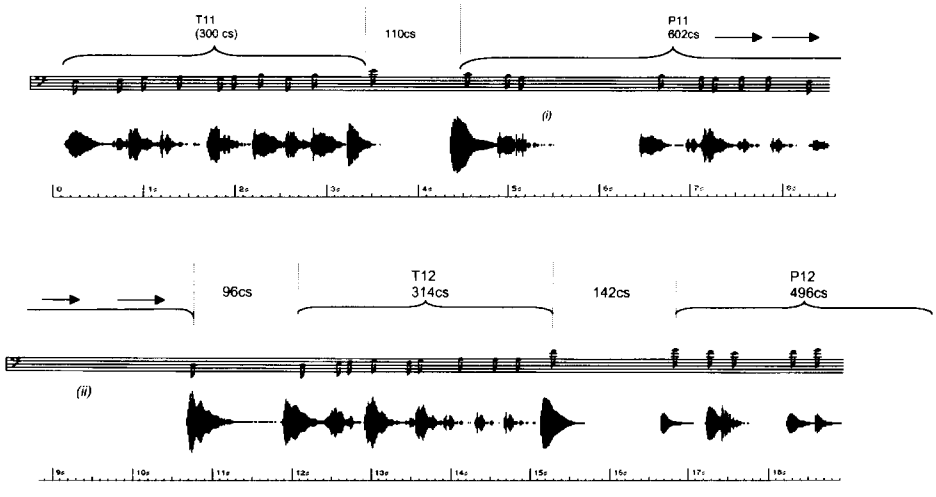



Figure 8.1 Turn-taking on the marimba by a therapist (T) and depressed patient (P).

Other interesting aspects in this brief excerpt are that when the therapist plays again at T12, her statement acknowledges that the patient has reflected some of the material in T11, which encourages the therapist to offer a similar statement to T11 in her twelfth statement (T12). Also of interest is that the patient is contracting her phrase to 496 cs in P12, suggesting some interpersonal responsiveness and flexibility—despite her depression. The therapist's statement at T12 is almost a repeat of her eleventh statement—and this can be understood as her offering the patient a stable and unambiguous musical structure, with predictable rhythm and pulse: a temporal beacon for the patient. Were the therapist to simply reflect the patient's uneven playing, and offer some rhythmic variation in T12, this might have the effect of destabilizing the patient. Instead, the therapist offers the patient a potential relational and communicative environment to expand into, leading, possibly, to a more fluid interpersonal (and personal) expressiveness. Here is dynamic form as an observable, explicit phenomenon with all the mechanisms of interpersonal timing, contour, motion, shape, and phrasing. It is the 'music between' (Ansdell 1995) the therapist and patient that generates and reveals emotion in music therapy—through the therapist's skill at eliciting, and working with, clinical improvisation.

## 8.6 Conclusion

An emotional response to music is a fundamental aspect of music therapy. It appears as a central link for people of all ages with wide-ranging clinical needs. Music therapists demonstrate the capacity to judge effectively the intended basic emotions of short

improvisations played live. The musical parameters used in their judgments are similar to those used when listening to recorded pre-composed music. Recent work in developmental child psychology, in particular Stern's (1985) 'vitality affects', provides reference points in understanding the subtle interplay between the personal and the musical in improvisational music therapy. Here is perhaps where we have the primary locus of any therapeutic change, in the interpersonal musical relationship built up between therapist and patient. Tuning in to our patients with a focused attention and sense of listening to every emotional nuance being expressed are prerequisites of the work of a vigilant music therapist.[1]

## References


**Aigen, K.** (1998). *Paths of development in Nordoff–Robbins music therapy.* Gilsum, NH: Barcelona.

**Aigen, K.** (1999). The true nature of music-centered music therapy theory. *British Journal of Music Therapy*, 13, 77–82.

**Aldridge, D.** (1996). *Music therapy research and practice in medicine.* London: Jessica Kingsley.

**Aldridge, D.** (1999). *Music therapy in palliative care.* London: Jessica Kingsley.

**Aldridge, D.** (2000). *Music therapy in dementia care.* London: Jessica Kingsley.

**Altshuler, I. M.** (1954). The past, present and future of music therapy. In *Music therapy*, (ed. E. Podolsky), pp. 24–35. New York: Philosophical Library.

**Alvin, J.** (1975). *Music therapy.* London: John Clare.

**Ansdell, G.** (1995). *Music for life: Aspects of creative music therapy with adult clients.* London: Jessica Kingsley.

**Ansdell, G.** (1996). Talking about music therapy: a dilemma and a qualitative experiment. *British Journal of Music Therapy*, **10**, 4–16.

**Ansdell, G.** (1997). Musical elaborations: what has the new musicology to say to music therapy? *British Journal of Music Therapy*, 11, 36–44.

**Beebe, B., Jaffe, J., Feldstein, S., Mays, K., & Alson, D.** (1985). Interpersonal timing: the application of an adult dialogue model to mother–infant vocal and kinesic interactions. In *Social perception in infants*, (ed. T. M. Field & N. A. Fox), pp. 217–47. Norwood, NJ: Ablex.

**Behrens, G. A. & Green, S. B.** (1993). The ability to identify emotional content of solo improvisations performed vocally and on three different instruments. *Psychology of Music*, **21**, 20–33.

**Bernieri, F. J. & Rosenthal, R.** (1991). Interpersonal coordination: behavior matching and interactional synchrony. In *Fundamentals of nonverbal behavior*, (ed. R. S. Feldman & B. Rim), pp. 401–32. Cambridge, UK: Cambridge University Press.

**Bonny, H. L. & Savary, L. M.** (1973). *Music and your mind.* New York: Station Hill.

**Brown, J. J. & Avstreih, Z. A. K.** (1989). On synchrony. *Arts in Psychotherapy*, **16**, 157–62.

**Brown, S.** (1999). Some thoughts on music, therapy, and music therapy. *British Journal of Music Therapy*, 13, 63–71.

**Brown, S. & Pavlicevic, M.** (1997). Clinical improvisation in creative music therapy: musical aesthetic and the interpersonal dimension. *Arts in Psychotherapy*, 23, 397–405.


[1] We are grateful to Mollie Gilchrist, Faculty of Health and Social Care, University of the West of England, for help with the statistical analysis.

**Bruscia, K. E.** (1987). *Improvisational models of music therapy.* Springfield, IL: Charles C. Thomas.

**Bruscia, K. E.** (1989). *Defining music therapy.* Spring City, PA: Barcelona.

**Bruscia, K. E.** (1991). *Case studies in music therapy.* Phoenixville, PA: Barcelona.

**Bruscia, K. E.** (1995). Modes of consciousness in guided imagery and music (GIM): A therapist's experience. In *Listening, playing, creating: essays on the power of sound,* (ed. C. B. Kenny), pp. 165–99. Albany, NY: State University of New York.

**Bruscia, K. E.** (1998). *Dynamics of music psychotherapy.* Gilsum, NH: Barcelona.

**Bunt, L.** (1994). *Music therapy: An art beyond words.* London: Routledge.

**Bunt, L.** (1997). Clinical and therapeutic uses of music. In *The social psychology of music,* (ed. D. J. Hargreaves & A. C. North), pp. 249–67. Oxford, UK: Oxford University Press.

**Bunt, L.** (2000). Transformational processes in guided imagery and music. *Journal of the Association for Music and Imagery,* 7, 44–69.

**Bunt, L. & Marston-Wyld, J.** (1995). Where words fail music takes over: A collaborative study by a music therapist and a counsellor in the context of cancer care. *Music Therapy Perspectives,* **13**, 46–50.

**Burns, S. J., Harbuz, M. S., Huckelbridge, F., & Bunt, L.** (2001). A pilot study into the therapeutic effects of music therapy at a cancer help centre. *Alternative Therapies in Health and Medicine,* 7, 48–56.

**Capra, F.** (1983). *The turning point.* London: Fontana.

**Champion, L.** (1992). Depression. In *Adult psychological problems,* (ed. L. A. Champion & M. J. Power), pp. 26–49. London: Falmer Press.

**Clark, M.** (1999). The Bonny method of guided imagery and music and spiritual development. *Journal of the Association for Music and Imagery,* **6**, 55–62.

**Condon, W. S. & Ogston, W. D.** (1966). Sound film analysis of normal and pathological behavior patterns. *Journal of Nervous and Mental Diseases,* **143**, 338–47.

**Davies, J. B.** (1978). *The psychology of music.* London: Hutchinson.

**Edgerton, C. L.** (1994) The effect of improvisational music therapy on the communicative behaviors of autistic children. *Journal of Music Therapy,* **31**, 311–62.

**Erdonmez Grocke, D.** (1999). The music which underpins pivotal moments in guided imagery and music. In *Clinical applications of music therapy in psychiatry,* (ed. T. Wigram & J. De Backer, J.), pp. 197–210. London: Jessica Kingsley.

**Feldstein, S. & Welkowitz, J.** (1979) A chronography of conversation: In defense of an objective approach. In *Nonverbal behavior and communication,* (ed. A. W. Siegman & S. Feldstein), pp. 329–278). Hillsdale, NJ: Erlbaum.

**Gabrielsson, A. & Juslin, P. N.** (1996). Emotional expression in music performance. Between the performer's intention and the listener's experience. *Psychology of Music,* **24**, 68–91.

**Heal, M. & Wigram, T.** (ed.) (1993). *Music therapy in health and education.* London: Jessica Kingsley.

**Heaney, C. J.** (1992). Evaluation of music therapy and other treatment modalities by adult psychiatric inpatients. *Journal of Music Therapy,* **29**, 70–86.

**Langenberg, M., Aigen, K., & Frommer, J.** (ed.) (1996). *Qualitative music therapy research: Beginning dialogues.* Gilsum, NH: Barcelona.

**Langer, S.** (1942). *Philosophy in a new key.* Cambridge, MA: Harvard University Press.

**Langer, S.** (1953). *Feeling and form.* London: Routledge.

Lee, C. A. (1989). Structural analysis of therapeutic improvisatory music. *Journal of British Music Therapy*, 3, 11–19.

Lee, C. A. (1990). Structural analysis of post-tonal therapeutic improvisatory music. *Journal of British Music Therapy*, 4, 6–20.

Lee, C. A. (ed.) (1995). *Lonely waters.* Oxford, UK: Sobell.

Lee, C. A. (1996). *Music at the edge: The music therapy experiences of a musician with AIDS.* London: Routledge.

Madsen, C. K., Cotter, V. M., & Madsen, C. H. (1968). A behavioural approach to music therapy. *Journal of Music Therapy*, 5, 69–71.

Malloch, S. N. (1999). Mothers and infants and communicative musicality. *Musicae Scientiae*, Special Issue 1999–2000, 29–54.

Montello, L. (1998). Effects of active versus passive group music therapy on preadolescents with emotional, learning, and behavioral disorders. *Journal of Music Therapy*, 35, 49–67.

Nordoff, P. & Robbins, C. (1971). *Therapy in music for handicapped children.* London: Gollancz.

Nordoff, P. & Robbins, C. (1977). *Creative music therapy.* New York: John Day.

Odell-Miller, H. (1988). A music therapy approach in mental health. *Psychology of Music*, 16, 52–61.

Odell-Miller, H. (1995). Approaches to music therapy in psychiatry with specific emphasis upon a research project with the elderly mentally ill. In *The art and science of music therapy: A handbook*, (ed. T. Wigram, B. Saperston, & R. West), pp. 83–111. Langhorne: Harwood Academic Press.

Papoušek, M. (1989). Determinants of responsiveness to infant vocal expression of emotional state. *Infant Behavior and Development*, 12, 505–22.

Papoušek, M. (1996). Intuitive parenting: A hidden source of musical stimulation in infancy. In *Musical beginnings*, (ed. I. Deliège & J. A. Sloboda), pp. 88–112. Oxford, UK: Oxford University Press.

Papoušek, H. & Papoušek, M. (1991). Innate and cultural guidance of infants' integrative competencies: China, the United States and Germany. In *Cultural approaches to parenting*, (ed. M. H. Bornstein), pp. 23–44. Hillsdale, NJ: Erlbaum.

Pavlicevic, M. (1990). Dynamic interplay in clinical improvisation. *British Journal of Music Therapy*, 4, 5–9.

Pavlicevic, M. (1991). *Music in communication: Improvisation in music therapy.* Unpublished doctoral dissertation, University of Edinburgh, Scotland.

Pavlicevic, M. (1995). Music and emotion: aspects of music therapy research. In *Art and music: Therapy and research*, (ed. A. Gilroy & C. Lee), pp. 51–65. London: Routledge.

Pavlicevic, M. (1997). *Music therapy in context: Music, meaning, and relationship.* London: Jessica Kingsley.

Pavlicevic, M. (1999*a*). *Music therapy—intimate notes.* London: Jessica Kingsley.

Pavlicevic, M. (1999*b*). With listeners in mind: creating meaning in music therapy dialogues. *Arts in Psychotherapy*, 26, 85–94.

Pavlicevic, M. (2000). Improvisation in music therapy I: Human communication in sound. *Journal of Music Therapy*, xxxvii, 270–85.

Pavlicevic, M. & Trevarthen, C. (1989). A musical assessment of psychiatric states in adults. *Psychopathology*, 22, 325–34.

Pavlicevic, M., Trevarthen, C., & Duncan, J. (1995). Music therapy in the rehabilitation of persons suffering from chronic schizophrenia. *Journal of Music Therapy*, 31, 86–104.

Priestley, M. (1975). *Music therapy in action.* London: Constable.

**Priestley, M.** (1994). *Essays on analytical music therapy.* Phoenixville, PA: Barcelona.

**Robarts, J.** (1996). Music therapy for children with autism. In *Children with autism,* (ed. C. Trevarthen, K Aitken, D. Papoudi, & J. Robarts), pp. 134–60. London: Jessica Kingsley.

**Robb, L.** (1999). Emotional musicality in mother–infant vocal affect, and an acoustic study of postnatal depression. *Musicae Scientiae,* Special Issue 1999–2000, 123–51.

**Ruud, E.** (1980). *Music therapy and its relationship to current treatment theories.* St Louis: Magnamusic-Baton.

**Ruud, E.** (1998*a*). *Music therapy: Improvisation, communication, and culture.* Gilsum, NH: Barcelona.

**Ruud, E.** (1998*b*). Science as metacritique. *Journal of Music Therapy,* 35, 218–24.

**Schögler, B.** (1998). Music as a tool in communications research. *Nordic Journal of Music Therapy,* 7, 40–9.

**Small, C.** (1999). Musicking: the meanings of performance and listening. A lecture. *Music Education Research,* 1, 9–21.

**Smeijsters, H.** (1997). *Multiple perspectives: A guide to qualitative research in music therapy.* Phoenixville, PA: Barcelona.

**Standley, J. M. & Hanser, S. B.** (1995). Music therapy research and applications in pediatric oncology treatment. *Journal of Pediatric Oncology Nursing,* 12, 3–8.

**Stern, D.** (1985). *The interpersonal world of the infant.* London: Academic Press.

**Streeter, E.** (1999). Finding a balance between psychological thinking and musical awareness in music therapy theory—a psychoanalytic perspective. *British Journal of Music Therapy,* 13, 5–20.

**Styge, B.** (1999). Perspectives on meaning in music therapy. *British Journal of Music Therapy,* 12, 20–8.

**Trehub, S. E. & Trainor, L. J.** (1993). Listening strategies in infancy: The roots of music and language development. In *Thinking in sound: The cognitive psychology of human audition,* (ed. S. McAdams & E. Bigand), pp. 278–327. Oxford, UK: Oxford University Press.

**Trevarthen, C.** (1993). The self born in intersubjectivity: An infant communicating. In *Ecological and interpersonal knowledge of the self,* (ed. U. Neisser), pp. 121–73. New York: Cambridge University Press.

**Trevarthen, C.** (1999). Musicality and the intrinsic motive pulse: evidence from human psychobiology and infant communication. *Musicae Scientiae,* Special Issue 1999–2000, 155–211.

**Trevarthen, C. & Aitken, K. J.** (1994). Brain development, infant communication, and empathy disorders: Intrinsic factors in child mental health. *Development and Psychopathology,* 6, 599–635.

**Tyler, H.** (1998). Behind the mask: An exploration of the true and false self as revealed in music therapy. *British Journal of Music Therapy,* 12, 20–8.

**Unyk, A. M., Trehub, S. E., Trainor, L. J., & Schellenberg, E. G.** (1992). Lullabies and simplicity: A cross-cultural perspective. *Psychology of Music,* 20, 15–28.

**Wheeler, B.** (ed.). (1995). *Music therapy research: Quantitative and qualitative perspectives.* Phoenixville, PA: Barcelona.

**Wigram, T. & De Backer, J.** (ed.) (1999*a*). *Clinical applications of music therapy in develpmental disability, paediatrics and neurology.* London: Jessica Kingsley.

**Wigram, T. & De Backer, J.** (ed.) (1999*b*). *Clinical applications of music therapy in psychiatry.* London: Jessica Kingsley.

**Wigram, T., Saperston, B., & West, R.** (ed.) (1995). *The art and science of music therapy: A handbook.* Langhorne: Harwood Academic Press.

PART III

# THE COMPOSER

CHAPTER 9

# EMOTION AND COMPOSITION IN CLASSICAL MUSIC: HISTORIOMETRIC PERSPECTIVES

DEAN KEITH SIMONTON

The compositions that constitute the classical music tradition of the Western world probably have many different functions. Yet among those functions is probably some kind of emotional expression or communication. That is, when aficionados of classical music attend a concert or turn on their stereo, they usually expect to be 'moved'. Furthermore, listeners often believe that these emotional reactions are what the composer intended—that the work is a vehicle for emotional communication. As Beethoven once put it, 'Coming from the heart, may it go to the heart' (quoted in Scherman & Biancolli 1972, p. 951).

What makes these expectations somewhat surprising is the fact that classical music constitutes a fairly abstract form of aesthetic expression. Indeed, the concert halls are often dominated by purely instrumental forms—such as the symphony, concerto, quartet, and sonata—in which not a single word or programme gives any clue what the piece is actually about. Moreover, looking at a music score does not help render a classical composition any less abstract. On the contrary, musical notation is far more precise, rarified, and refined than can be found in any other form of artistic expression. Superficially at least, a score has more in common with a mathematical proof than with a painting, poem, or sculpture. This mathematical abstraction even holds for vocal forms, such as opera, oratorio, cantata, and song, where the words to be sung occupy a relatively small space on the page relative to all those meticulously placed symbols indicating pitch, duration, rhythm, dynamics, and other essential features of a composition. No wonder Claude Debussy could claim that 'music is the arithmetic of sounds as optics is the geometry of light' (quoted in *Who said what when*, Anon 1991, p. 252).

These observations thus lead to the following question: how do the abstractions of classical music evoke emotional reactions? There are many useful routes to addressing this issue. One option is to engage in a detailed musicological analysis, such as Leonard Meyer (1956) did in his important book on *Emotion and meaning in music*. Alternatively, one can conduct a laboratory experiment in which the emotional responses of listeners are assessed directly. An excellent example is Krumhansl's (1997) psychophysiological

study of the reactions of college students to musical excerpts from the classical repertoire. These musicological and experimental enquires have contributed a great deal to our understanding of the emotional impact of classical music. Yet they by no means exhaust the available methodological approaches.

Here, in fact, I adopt a distinctive analytical strategy that is far less common than either musicological analysis or laboratory experiments. This approach entails the historiometric study of the composers and compositions that define the classical repertoire. By historiometrics, I mean the adaptation of psychometric methods to examine data about historical individuals (Simonton 1990, 1999*b*; cf. Woods 1909, 1911). In the case of classical composers, these data can include both: (a) biographical information about their lives; and (b) content-analytical information about their works. In a series of articles published since 1977, I have been using this specific technique to study creativity and aesthetics in classical music (e.g. Simonton 1977*b*, 1991, 1995). This research was part of a broader series of studies devoted to understanding the psychology of creative genius. Although this research programme was not directed specifically at investigating emotional expression in music, some of the empirical results might possibly shed a little light on how this phenomenon operates, at least within the classical music tradition.

I begin by describing an objective and quantitative measure of a composition's melodic originality. I then show that scores on this computerized measure have certain aesthetic consequences; consequences apparently consistent with the assumption that the measure gauges something of a composition's 'arousal potential'. These content-analytical scores are next shown to relate to the composer's biography in a manner seemingly compatible with the same assumption. I then conclude with a general evaluation about what this research programme may possibly suggest about music and emotion.

## 9.1 Melodic originality

The original impetus for this research came from an unexpected direction: the issue of musical style. Those with lots of performing or listening experience with classical music eventually learn to identify the distinctive styles of particular composers. Not only does Beethoven's music sound different in comparison to that of Monteverdi, Handel, Tchaikovsky, or Bartók, but it is even recognizably different from fellow composers of the classical period, such as Haydn and Mozart. But what makes one composer noticeably different from another, even when they both are composing in the same period style?

To answer this question, Paisley (1964) developed an idea taken from those art critics who try to identify the creators of certain unsigned Renaissance paintings. What the critics learn to look for are certain 'minor encoding habits', such as the way a particular painter depicts ears, hands, or stones. Because these objects are rather secondary in importance relative to the composition as a whole, they will often be executed in a mechanical, almost mindless manner. Minor encoding habits have also proven useful in identifying the unknown authors of various literary texts (Holsti 1969). These habits include favourite function words, such as preferred prepositions and conjunctions.

Paisley wished to determine whether classical composers betrayed their identity through analogous encoding habits.

Paisley's (1964) specific solution was to perform a content analysis of the melodies or themes that make up almost any musical composition. Taking the first four notes of each theme, he then determined which combinations of notes were the most common for particular composers. Paisley found that different composers had distinctive ways of constructing their thematic material, and that these encoding habits could be used to predict the identity of the composer of anonymous test samples. Using this technique, Beethoven's work could be easily distinguished from that of other classical composers, including the works of Haydn and Mozart.

If composers can reveal their identity from just the first four notes of their melodies, what else can be learned from the content analysis of their melodic ideas? I decided to develop Paisley's approach to an objective and quantified assessment of an important property of music, namely its melodic originality.

### 9.1.1 Computerized measurement: two-note transition probabilities

The first task was to obtain a large and representative sample of classical composers. The sampling began with all those whose music could be found in a two-volume dictionary of musical themes (Barlow & Morgenstern 1948, 1950). This dictionary included all the classical music that was recorded up to the middle of the twentieth century. After deleting anonymous compositions and other problematic works, the sample consisted of 15 618 themes by 479 composers (Simonton 1980*b*). This could hardly be called a 'sample' at all, given its size. In fact, according to one study based on performance data, only 100 composers account for 94 per cent of the music that has any place whatsoever in the standard repertoire of classical music (Moles 1968; see also Simonton 1991). So the sample was exhaustive when it came to major composers. In addition, the sample was fairly exhaustive when it came to the principal compositions by the tradition's major figures. For instance, among the 15 618 themes were those coming from *all* of Beethoven's symphonies, concertos, string quartets, and piano sonatas (for a complete listing, see Barlow & Morgenstern 1948, 1950).

Following Paisley (1964), the thematic dictionaries that generated the sample also provided the foundation for the content analysis. For each theme, the dictionaries give the beginning notes, all transposed to a C tonic. In other words, all melodies in the major mode had been transposed into C major, and all those in the minor mode had been transposed into C minor. Although Paisley accomplished a lot with only the first four notes, I decided to use the first *six* notes of each of the 15 618 melodies. These data were all entered into a huge machine-readable data file. The next step was to have a computer calculate the two-note transition probabilities. That is, the first six notes contain five two-note transitions (first to second, second to third, etc.). The partial results of this computerized content analysis are shown in Table 9.1 (for a more complete table, see Simonton 1984, Table 1).

Clearly, some two-note transitions are far more common than others. On the one hand, the dominant–dominant (GG), tonic–tonic (CC), and dominant–tonic (GC) together account for about 17 per cent of all two-note combinations. On the other hand, combinations involving sharps and flats are much more rare—all occurring less

**Table 9.1** Two-note transition probabilities from the first six notes of 15 618 classical themes

| |
|---|
| $P \geq 0.06$: GG (0.067) |
| $P \geq 0.05$: CC (0.053) |
| $P \geq 0.04$: GC (0.049) and CD (0.044) |
| $P \geq 0.03$: CB (0.032), CG (0.032), GF (0.031), EE (0.030), ED (0.030), and DC (0.030) |
| $P \geq 0.02$: GE (0.029), GA (0.029), EF (0.028), EG (0.026), DE (0.024), BC (0.023), CE (0.022), FG (0.021), FE (0.021), GA (0.021), and AG (0.020) |
| $P \geq 0.01$: E♭D (0.018), EC (0.016), DE♭ (0.014), AB (0.012), BA (0.011), A♭G (0.011), and DD (0.011) |

Probabilities given in parentheses are averaged across the first five two-note transitions (extracted from the more complete listing in Simonton 1984).

than 2 per cent of the time. Moreover, those that do appear, namely E♭ and A♭, are not chromatic, but rather are representative of the minor mode. Truly chromatic notes (those that depart from either the major or the minor mode) are even more improbable, only F# occurring with any frequency whatsoever.

It should be pointed out that the probabilities presented in Table 9.1 were averaged across all five two-note transitions. Slightly different results obtain when the two-note probabilities are calculated separately for each of the five transitions. For instance, two-note combinations involving the tonic and dominant decline in frequency from the first to the sixth note. Thus, over 11 per cent of the time the first transition contains the dominant–dominant (GG) pairing, but this figure decreases to less than 5 per cent by the fifth transition. In contrast, those transitions that involve other notes of the scale increase in frequency. Apparently, when composers first begin constructing a melody, they usually start by defining the theme's key with the most definitive notes of the scale before incorporating the other notes of the diatonic scale, such as the mediant, subdominant, and leading tone.

These latter, transition-specific probabilities were then used to construct a quantitative measure of melodic originality (Simonton 1980*b*). For each of the 15 618 themes, the probability of each of its five transitions was calculated, and then an average taken. The result is a number that represents the mean percentage of times a particular theme's five two-note transitions appear in the entire repertoire. To illustrate, the second movement of Haydn's Symphony No. 94 (the 'Surprise') opens with an extremely simple theme that begins CCEEGG. All of the component two-note transitions are extremely commonplace: the average transition probability is 0.040. A similar degree of predictability is found in the main theme of the concluding movement of Beethoven's 'Waldstein' Sonata, which begins GGEDGC, yielding a mean transition probability of 0.041. By comparison, the opening theme of the introduction to Mozart's 'Dissonant' Quartet commences with AGF#GAB♭, with a mean probability of 0.005. The theme that initiates Liszt's Faust Symphony is even more unlikely: A♭GBE♭F#B♭, yielding an average two-note probability of less than 0.001 (for more examples, see Simonton 1984, Table 3). In any case, given these numbers, a measure of each theme's melodic originality can be defined as one minus the mean probability (i.e. $1-\sum P_i/5$). The higher this number, the higher is the assigned score for melodic originality.

This computerized measure may seem very crude. To begin with, the measure concentrates solely on the melody, thereby completely ignoring instrumentation,

counterpoint, harmony, formal structure, and text (if any). The implicit assumption being made is that 'melody is the main thing', to quote Haydn (Landowska 1964, p. 336). But even with respect to melody alone, the computer-generated score ignores the theme's key, its rhythmic structure, and many other significant features. Instead, the measure concentrates on the relationships between consecutive pitches. Nevertheless, it turns out that this objective and quantitative score is not only perfectly reliable—the reliability coefficient is necessarily unity—but in addition it enjoys enough validity for research purposes.

### 9.1.2 Measure validation

The melodic originality measure was validated several ways. One very direct validation involved the direct comparison of the measure based on two-note transition probabilities with an alternative measure based on three-note transition probabilities (see Table 2 in Simonton 1984). These two measures correlate so highly that they yield practically the same results (Simonton 1980*a*). Perhaps this should not be too surprising, given that the most likely three-note transitions tend to be made up of the most common two-note transitions (e.g. GGG and CCC). In addition, originality scores based on the probabilities specific to each consecutive two-note transition produce almost the same outcomes as those based upon probabilities that are averaged across all transitions for all 15 618 themes (Simonton 1984). Hence, this gauge of melodic originality is not contingent on the particular operational definition adopted.

Another route to validation was to see if melodic originality correlates with other characteristics of the composition in a manner consistent with expectation. These correlations can be grouped into two categories: general correlates with other characteristics of the composition and those correlates that may more specifically link the measure with emotional expressiveness (Simonton 1994).

*General correlations with compositional characteristics*

The melodic originality of the themes in a composition is associated with other aspects of that composition. These linkages suggest that the computer content analysis has captured something that has musical significance. In particular, consider the following four empirical findings:

1. Originality is higher for those themes that are also more metrically complex (Simonton 1987). By metric complexity I mean melodies in less common metres, such as 3/2 or 9/8 rather than 2/4 or 4/4 time. Thus, as the two-note transitions become more unpredictable, the beats that contain the corresponding pitches become less simple. This hints that the originality measure might indirectly encompass more compositional effects than just melody alone.
2. Originality is higher for themes from instrumental works than those from vocal works (Simonton 1980*b*). Part of this relation may reflect the superior flexibility of most musical instrumental instruments relative to the human voice (e.g. the virtuoso violin versus a coloratura soprano). But another part may be ascribed to the fact that in vocal compositions, the text has part of the responsibility for communicating content (including any emotional expression).

3. Originality is higher for themes from chamber compositions than those from orchestral compositions, the latter in turn having more original themes than those from theatrical works, such as ballet (Simonton 1980*b*). In chamber compositions, such as sonatas and quartets, largely the thematic material carries the burden of maintaining listener interest. As the instrumental resources increase, as in symphonies, tone poems, and overtures, the opportunities rendered by orchestration can assume more importance in maintaining such appeal and attention. In theatrical compositions, extramusical stimulation becomes prominent as well, such as costumes, sets, and choreography. This renders melodic originality even less necessary.
4. Originality is higher in the outer movements of multimovement compositions than in the inner movements (Simonton 1987). For instance, symphonies and quartets will most often contain four movements: the first usually in sonata–allegro form, the second a slow movement in song or theme-variation form, a third a minuet (or scherzo) and trio, and a fourth that is in sonata–allegro, rondo, or variation form. The first movement contains the themes with the highest melodic originality, followed by the last movement. The middle movement, or movements, features the themes with the lowest melodic originality. Evidently, once the composer has got the listener's attention, originality can relax momentarily, until the intensity picks up again in the work's conclusion.

*Specific correlations with emotional expression and impact*

Although the melodic originality measure was not designed to tap the emotional content of a composition, two empirical findings suggest that some connection might exist:

1. Originality is higher for themes in minor keys than for those in major keys (Simonton 1987). As is apparent from Table 9.1, transitions containing the notes that define the minor mode have lower probabilities than those from the major mode, which gives minor-key themes their higher originality scores. This fits the subjective experiences of many music listeners, who tend to find themes in the minor mode more exotic, strange, or unusual in some way (Meyer 1956). That impression reflects the greater unpredictability of melodies that necessarily flow by less predictable transition probabilities. Perhaps as a consequence, composers frequently choose the minor mode for emotionally expressive compositions. Representative examples include Chopin's F Minor Fantasia, Mozart's G Minor String Quintet, Rachmaninoff's C Minor Piano Concerto No. 2, or Tchaikovsky's B Minor *Pathétique* Symphony. 'The minor mode is not only associated with intense feeling in general but with the delineation of sadness, suffering, and anguish in particular', observed Meyer (1956, p. 297). Later in this chapter, Meyer's assertion will receive some historiometric support.
2. Scores on the melodic originality measure can be directly compared with the subjective experiences of human listeners. Melodies that score higher in melodic originality should sound more unpredictable, interesting, or complex than melodies that score much lower. Anyone familiar with the themes mentioned ear-

lier should agree that this is the case. The Haydn theme sounds very much like a simple nursery tune (*Twinkle, Twinkle, Little Star*, in its English version), whereas the Mozart theme was so incomprehensible to Mozart's contemporaries (including Haydn) that the quartet which it opens with has been known as the 'dissonant' ever since. More important, one experimental investigation found that the originality scores correspond with subjective assessments of a theme's 'arousal potential' (Martindale & Uemura 1983). The originality scores were calculated using the two-note transition probabilities for themes by 252 composers, with one theme per composer. Naïve listeners then made an independent assessment of the arousal potential of each theme after listening to recordings made by a professional violinist. Those themes that scored higher on the objective measure of melodic originality tended to be those that also scored higher on the subjective measure of arousal potential. In other words, melodic originality predicts whether a theme will be perceived as exciting, stimulating, or arousing. Admittedly, the correlation was only 0.21 ($P < 0.001$), but this coefficient was probably attenuated by two factors. The first factor is that the transition probabilities were based on only 252 themes rather than 15 618. The second attenuating factor is that the variance in melodic originality would probably be truncated, given that only 252 themes by 252 composers were examined instead of 15 618 themes by 479 composers. If this correlation is then taken as the lower bound for the true correlation, it seems reasonable to conclude that, within the entire classical repertoire, those themes that contain more original two-note transition probabilities will be more likely to evoke emotional reactions in listeners.[1]

## 9.2 Aesthetic consequences

If the hypothesized linkage between melodic originality and arousal potential has some justification, then the content-analytical measure should bear a theoretically meaningful association with: (a) a composition's aesthetic impact on listeners; and (b) changes in aesthetic style within a composer's career and within the history of classical music. The reasons for these expectations will become clear below.

### 9.2.1 Compositional impact

Research in experimental aesthetics has suggested that the aesthetic success of any work of art is in part a function of its capacity to evoke emotional arousal (e.g. Berlyne 1971, 1974). The arousal potential of a given work is associated with such things as the complexity, novelty, surprisingness, and ambiguity of the aesthetic stimulus. The more complex, novel, surprising, or ambiguous the stimulus, the higher the magnitude of arousal elicited. However, the functional relation between arousal potential and aesthetic

[1] Tuomas Eerola (personal communication, 26 March 1999) informed me that he had found a correlation of 0.93 between the melodic originality measure and an alternative measure of melodic complexity derived from research on music perception. This constitutes another independent validation of the computer-generated scores.

impact is not linear, but rather curvilinear. That is, the relation is described by what has been called the 'Wundt curve' (Berlyne 1971). The most successful works are those that evoke an optimal level of arousal. In contrast, compositions with low arousal potential will prove to be predictable, commonplace, or obvious—and hence boring. And compositions that provoke excessive arousal will induce a state of stress, anxiety, or even fear (e.g. the notorious debut of Stravinsky's *Rite of Spring*). A number of investigations in experimental aesthetics have in fact found curvilinear inverted-U relations between measures of arousal potential and aesthetic preferences (e.g. Kammann 1966; Steck & Machotka 1975; Vitz 1964).

However, most of these experimental studies have used artificial 'art-like' stimuli, which were then rated by college students in somewhat unnatural laboratory conditions (Berlyne 1974). This may help explain why the results are not always consistent with theoretical expectation (e.g. Martindale & Moore 1989). Therefore, it behoves us to ask whether the same Wundt curve can be found under more natural circumstances. In particular, is the aesthetic success of real compositions in the classical repertoire partially determined by the melodic originality of the themes those compositions contain? If this content-analytical measure can really be taken as an approximate gauge of a theme's arousal potential, then scores should predict a composition's aesthetic impact. This is in fact the case for both objective and subjective indicators:

*Objective ratings* were based on the actual success of the composition in the classical repertoire (Simonton 1980*a*, 1980*b*, 1983). That is, these ratings gauged the frequency that a musical product is likely to be heard in the concert hall, opera house, and recording studio (for a psychometric assessment of such indicators, see Simonton 1998*a*). The relationship between this objective measure and melodic originality is neither linear positive nor linear negative but rather curvilinear. In particular, the relationship is best described as an inverted backward-J curve, as shown in Fig. 9.1 (Simonton 1980*b*). The most popular works are those that have moderate levels of melodic originality, whereas the most unpopular works are those with the highest levels of melodic originality—the pieces with low melodic originality having more middling popularity. This curvilinear function demonstrates that the Wundt curve provides an adequate description of what happens in the real world of music listening. At least, this function holds for the 15 618 themes that define the classical repertoire. The same function has also been shown to hold for the 593 themes that Beethoven contributed to the classical repertoire (Simonton 1987).[2]

[2] Here, as in the rest of this chapter, I am more precisely speaking of *repertoire* melodic originality, that is, the unpredictability of the sequence of pitches when compared with the entire standard repertoire of classical music. Another kind of originality has been defined, namely *Zeitgeist* melodic originality (Simonton 1980*b*). In this case, originality is gauged according to whether a particular theme's construction fits the stylistic norms at the time the work was composed. This distinction is crucial, because the two kinds of melodic originality do not correlate the same way with other variables. For example, the popularity of a composition is a U-shaped function of its *Zeitgeist* melodic originality (Simonton 1980*b*). The most successful works are those that depart (in either direction) from the level of repertoire melodic originality most typical of the day.

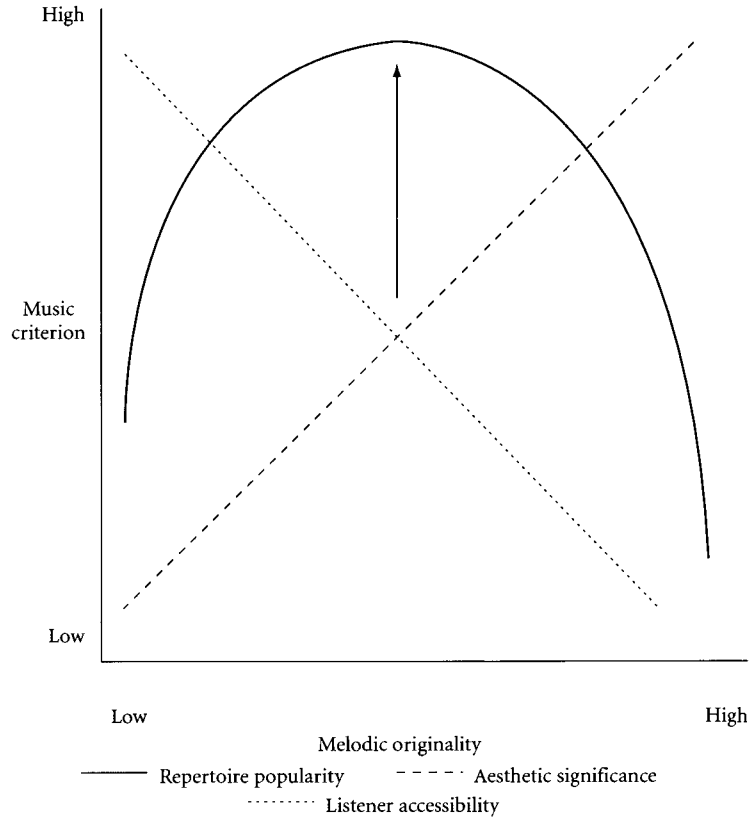



**Figure 9.1** The graph summarizes the empirical results regarding how a theme's melodic originality (as computed from two-note transition probabilities) is associated with three musical criteria: objective repertoire popularity (a composite measure of the performance and recording frequency of the composition which contains the theme) and subjective assessments of its aesthetic significance and listener accessibility (according to Halsey 1976). The arrow indicates the explanatory hypothesis that the peak of the curvilinear function might reside where the product of aesthetic significance and listener accessibility may be maximized.

*Subjective ratings* were taken from Halsey (1976), who rated thousands of classical compositions on two dimensions, namely aesthetic significance and listener accessibility. Melodic originality was correlated with both of these measures, albeit in opposite directions, as also indicated in Fig. 9.1 (Simonton 1986). Aesthetically significant compositions are those that were deemed sufficiently profound that they could withstand repeated listening. They cannot be fully appreciated in a single, superficial hearing. Such pieces score high on the melodic originality of their thematic material. Accessible compositions, in contrast, can be understood at once, and thereby lend themselves to music appreciation courses taught at schools and colleges. Not surprisingly, works that score high on listener

accessibility tend to score low on melodic originality. Highly unpredictable themes would only interfere with making an immediately favourable impression.

Note that the functions for the objective ratings were curvilinear, whereas those for the subjective ratings were linear and in opposed directions. Do these results contradict each other? I do not think so. Although aesthetic significance and listener accessibility correlate differently with melodic originality, they both correlate in the same positive direction with repertoire popularity. Accordingly, one could argue that the peak of the Wundt curve represents the point that maximizes the joint contribution of aesthetic significance and listener accessibility to repertoire popularity. At low levels of melodic originality, the music is very predictable and thus accessible, but the music is also considered of trivial aesthetic importance. At high levels of melodic originality, the aesthetic significance is not so much in doubt as the ability of listeners to appreciate the composition. At middle levels of originality, the works are still reasonably accessible while at the same time they are above the norm in significance. The vertical arrow in Fig. 9.1 represents the hypothesis that the highest popularity is to be found where the product of aesthetic significance and listener accessibility is maximal.

Interestingly, because low melodic originality is preferred over high melodic originality in terms of repertoire popularity, aesthetic significance appears to have less weight than listener accessibility. On the whole, at least with respect to emotional arousal, listeners would rather be quietly entertained than anxiously provoked.

### 9.2.2 Stylistic transformations

If the relation between aesthetic impact and melodic originality is mediated by the emotional arousal that highly original compositions evoke, then this relation could be moderated by another basic psychological process: *habituation*. With repetition, the response to an arousing stimulus will often diminish (Berlyne 1971). As a consequence, over time, increasingly higher levels of melodic originality may be required in order to elicit the same optimal level of arousal in music listeners (Martindale 1990; for literary illustration, see Kamman 1966). According to trend analyses of all 15 618 themes by 479 composers, this upward movement manifests itself in two ways (Simonton 1997).

First, composers tend to steadily increase their use of melodic originality as their careers progress (Simonton 1980*b*). Originality is lowest at the beginning of the career, but at once begins to increase. The only exception to this upward growth occurs at around the 56th year of life, where originality tends to reach a peak. After this a modest decline sets in. Beethoven enthusiasts may find the peak at 56 the most intriguing. Beethoven died at that age, which means that he should have been publishing his most original work then, if his career followed the same pattern as 478 other classical composers. A hearing of his Late Quartets suggests that this was indeed the case.

Second, there exists an overall tendency for melodic originality to increase from the Renaissance to the twentieth century (Simonton 1980*b*). This upward trend suggests that each generation of composers feels compelled to produce work that overcomes habituation to a given level of originality, just as predicted by Martindale's (1990) theory of stylistic change. However, just as an exception to the general positive trend was found for composer's careers, so are curious departures found across the history of the classical music tradition. Specifically, superimposed over the positive linear trend are

cyclical fluctuations. Melodies in the Renaissance were highly predictable, but by the time of Monteverdi and Gesualdo, originality had reached a maximum point, with a degree of chromaticism that was to be unmatched for nearly three centuries. After that, themes became more predictable again, albeit not as predictable as during the Renaissance. The trough occurred during the time of the 'classical period'—Haydn, Mozart, and Beethoven—when thematic material would be constructed from basic chords (e.g. the opening of Beethoven's *Eroica* Symphony). With the advent of the 'romantic period', melodic originality increased, most notably with the chromaticism of Chopin, Brahms, and Wagner (e.g. the famous prelude to *Tristan und Isolde*). This trend reached a new height about the time of the First World War. It was this peak that witnessed the emergence of the atonal and serial music of Arnold Schoenberg, such as his *Pierrot Lunaire*. Then melodic originality saw another decline, although again not going as low as the two previous lows. The music of Aaron Copland and Sergei Prokofiev is indicative of this retrogression.

The superimposed oscillation would seem to contradict the basic proposition that arousal potential has tended to increase over time. Yet, if this cyclical pattern is examined more closely, it becomes evident that the momentary declines in melodic originality tend to come during those periods in which composers acquire new ways to increase the arousal potential of their works besides increasing the novelty, complexity, surprisingness, and ambiguity of the thematic material. Rather than augment the magnitude of chromaticism, for example, the composer may have access to novel forms, rhythms, and orchestration. In partial support for this conjecture, not only is melodic originality lower in those forms giving the composer greater resources (as noted earlier), but also the relation between melodic originality and repertoire popularity is moderated by the size of the work (Simonton 1980*b*). The larger the form in which the composer writes, the weaker is the relation between originality and the composition's impact.

In any event, once alternative techniques for evoking emotional reactions become fully incorporated in the repertoire, the composers must again resort to melodic originality as a primary means to overcome the pressures produced by jaded tastes. Furthermore, it is essential to recognize that as the history of classical music unfolds, the troughs in originality get less deep and the crests become more elevated. Hence, the musical history of melodic originality never really repeats itself.

## 9.3 Biographical antecedents

For a chapter that purports to deal with music and emotion, the results so far may seem a bit cold. Granted, the objective and quantified indicator of melodic originality may indeed correspond with a theme's capacity to stimulate arousal—and that stimulation may shape a composition's aesthetic impact. But emotion in music is often seen as entailing much more than the simple activation of the human sympathetic nervous system. Emotion is also seen as a form of expression, as a means of communicating emotions between a creative artist and an audience. Yet so far we have treated melodic originality as some compositional quality that a composer must manipulate in order to keep listeners from falling asleep—or from running out of the concert hall in disgust or terror.

Fortunately, research using computerized measures has provided ample reason for believing that melodic originality may reflect the emotional state of the composer at the time of a work's composition. For example, the composer's use of melodic originality is noticeably altered when a composition is created during traumatic political events, such as international war (Simonton 1986, 1987; see also Brook 1969; Cerulo 1984). In particular, compositions penned during times of warfare exhibit much greater variation in melodic originality, extremely unpredictable and extremely predictable themes being found in the same work. It is almost as if such highly stressful events polarize emotional reactions, inspiring vast mood swings between resignation and despair, hope and fear. Shostakovitch's notable (and notorious) *Leningrad Symphony* provides a graphic auditory illustration of the magnitude of this polarization of trite and profound.

More significant, however, is the impact of highly personal events in the life of the composer. Below I treat two potential instances of this personal connection: biographical stress and late-life effects.

### 9.3.1 Biographical stress

Of special interest to any emotional expression hypothesis is the relation between the amount of stress occurring in a composer's personal life and the amount of melodic originality appearing in work created during the same period. If 'music sounds the way emotions feel' (Pratt 1954, p. 296, italics removed), and if stressful events provoke strong emotions, some correspondence between the two should be found. Moreover, the measure of melodic originality would seem to capture some of the attributes of music that contribute to its emotional expressiveness. An excellent example is chromaticism, or the use of notes beside those that define the theme's key. Chromatic notes are often used to express emotions, the unpredictability of the melodic line reflecting the turmoil of our internal emotional state (Meyer 1956). Indeed, this usage is not even confined to the classical music tradition, as the expressive use of chromatic 'blue notes' in blues and jazz illustrates (La Rue 1970).

To test the hypothesized linkage between life stress and melodic originality, I took advantage of an earlier data set that had been compiled for a different purpose (Simonton 1977*a*). The careers of classical composers were scrutinized to determine the factors that affect a composer's output across the lifespan. The particular composers were the ten most eminent according to a large survey of musicologists (Farnsworth 1969), namely, Bach, Handel, Haydn, Mozart, Beethoven, Schubert, Chopin, Wagner, Brahms, and Debussy. Among the potential predictors of compositional productivity was the occurrence of stressful or traumatic events in the composer's life. To obtain a measure of these events, a chronology of such events was first compiled from the extensive biographies available for these illustrious contributors to the classical repertoire. The next step was to convert these raw facts into a quantitative measure of the degree of stress experiences by each composer during consecutive periods of his life. This conversion was achieved by adapting the social readjustment rating scale, which has proven its utility in research on psychosomatic illnesses (Holmes & Holmes 1970; Holmes & Rahe 1967; see also Simonton 1998*b*). Table 9.2 shows some of the biographical events and the weights they were each assigned in the resulting tabulations.

Table 9.2 Biographical stress coding scheme designed for ten classical composers

*Legal difficulties*: Litigation or lawsuit (30); Detention in jail or exile to avoid arrest (63)

*Economic problems*: Major loan (20); Trouble with creditors (30); Aversive change in financial state or business readjustment (38)

*Educational changes*: Change in schools (20); Begin or terminate formal schooling (26)

*Vocational changes or problems*: Job change (20); Trouble with employer or superior (23); Change in responsibilities at work (29); Begin or end work but not fired or retired (36); Retirement (45); Fired from work (47)

*Mobility*: City or town of permanent residence changed (30 per move); Nation of permanent residence changed (40 per move)

*Interpersonal problems*: Duels, fights, and other physical confrontations (10); Argument with friend (10); Disappointed or unrequited love (15); Onset or termination of a reciprocated love affair (30); Death of a close friend (37)

*Family problems*: Gain of a new family member, including adoption (39); Change in health or behaviour of family member (44); Death of close family member, except children under 3 (63)

*Marital difficulties*: Marital reconciliation (45); Marriage (50); Marital separation (65); Divorce (73); Death of spouse, unless separated (100)

The points assigned to each life-change stress are given in parentheses.

Given these previously calculated assessments, it was a relatively simple matter to estimate the relation between a theme's melodic originality and the amount of stress the composer had experienced during the period of composition (tabulated in cumulative 5-year periods). The association was found to be positive, even after introducing controls for other confounding factors, such as the composer's age at the time a work was conceived (Simonton 1980*a*). This finding was replicated and extended in a second study that concentrated just on the career of Beethoven (Simonton 1987). Besides replicating the positive association between stress and originality, this investigation tabulated a measure of the lifespan fluctuations in Beethoven's physical health (cf. Porter & Suedfeld 1981; Simonton 1998*b*). This latter measure was also found to correspond to the melodic originality found in Beethoven's compositions. When the composer was experiencing robust health, originality tended to be relatively low, whereas when the composer's health failed, originality was likely to be high. Given that physical illness can be considered a major source of life stress, this finding reinforces what was found for biographical stress. The thematic material found in a composer's works may indeed communicate to the listener the composer's underlying emotional life at the time of their composition.

### 9.3.2 Late-life effects

Biographical stress and physical illness serve to heighten levels of melodic originality. Yet sometimes certain life events can operate in a contrary manner. This contrast is found in a study conducted of what was called the 'swan-song phenomenon' (Simonton 1989). There has been much speculation about whether artistic creators exhibit a 'late-life' or 'old-age' style (e.g. Arnheim 1986; Münsterberg 1983), and even some empirical evidence suggests that these speculations have captured a grain of truth (Lindauer

1993). Analogous speculations have been advanced regarding supposed last-work effects in classical music (Einstein 1956). So, do late-life compositions exhibit any consistent stylistic shifts? This question was addressed by examining melodic originality in 1919 works by 172 classical composers, and a striking developmental pattern was identified (Simonton 1989). In the years immediately preceding death, a composer's compositions exhibit rapidly declining levels of melodic originality. In other words, the thematic material becomes much more predictable. Significantly, this effect survived statistical control for the composer's age. Hence, the decline is a consequence of the proximity of death rather than old age. Composers who died young, such as Mozart and Schubert, could thus display the same swan-song effects as those who died at relatively advanced ages, like Vaughan Williams and Stravinsky.

But what is the psychological basis for this swan-song phenomenon? One possibility is to dismiss this effect as an example of the mental deterioration that too often occurs towards the end of any person's life (e.g. Suedfeld & Piedrahita 1984). Consistent with this interpretation, not only do the melodies tend to be less original in last works, but in addition the works themselves are of shorter duration on average (Simonton 1989). Even so, this negative interpretation does not seem very compatible with other empirical results: the closer to the composer's death a work appears, the more frequently it is performed in the classical repertoire and the higher the rating it receives in aesthetic significance. Hence, swan songs may be shorter and have more simple melodies, but they are also more successful in terms of popular and critical appraisal!

Given the sum total of these findings, it seems more likely that the lower melodic originality in late works is deliberate rather than negligent. If high originality is associated with highly emotional states, then low originality may reveal low levels of emotionality. For a composer who may feel that the end is quickly and inevitably approaching—as perhaps indicated by terminal illnesses or the frailties of old age—that low arousal state may hint that he has resigned himself to his fate. That resignation, that acceptance then takes the form of powerful and expressive swan songs that communicate the most with the least. This communication of internal peace can itself be considered a form of emotional expression. To get a direct sense of that feeling, we need only listen to the final chorale of Bach's *The Art of the Fugue*, Mozart's *Ave Verum Corpus*, Brahms' Four Serious Songs, or Richard Strauss' Four Last Songs.

## 9.4 Conclusion

The research just reviewed has many limitations, two of which are especially critical here. First, none of these studies were specifically dedicated to scrutinizing the relation between music and emotion. Instead, their primary purpose was to comprehend more generally the creativity of composers and the aesthetic impact of their compositions. In only one investigation was emotional impact directly assessed, and then only in the rather stripped-down form of 'arousal potential' (Martindale & Uemura 1983). The differentiated emotional states seen in other investigations, such as happiness and sadness, have no direct representation in this historiometric literature (cf. Krumhansl 1997). Furthermore, emotional reactions were mostly introduced as intervening variables to predict or explain the empirical relations between observed variables. Thus, the positive association

between biographical stress and melodic originality was predicted under the assumptions that: (a) stress produces emotional states; and (b) such emotional states influence the composer's melodic creativity (e.g. the use of chromaticism). Likewise, the expected curvilinear relation between melodic originality and repertoire popularity was directly based on Berlyne's (1971) optimal-arousal model of the relation between collative variation and aesthetic preference—and the Wundt curve that it theoretically predicts.

Fortunately, none of the empirical findings contradicted the two-fold supposition that melodic originality both influenced listeners' emotional response and reflected the composer's emotional states. In addition, a strong agreement exists between many of these findings and what Meyer (1956) argued on the basis of a detailed, sophisticated analysis of a select number of representative scores (e.g. the relation between emotion and the violation of melodic expectations). None the less, to make a stronger case for this supposition, these historiometric results should be tied more closely with findings from laboratory experiments. For instance, I have previously suggested using the two-note transition probabilities to construct artificial musical stimuli (Simonton 1984). Experiments can then gauge whether these stimuli influence both physiological responses and subjective appraisals of participants (e.g. as in Krumhansl 1997; see also Scherer & Zentner, this volume). This extension would introduce the experimental controls so obviously lacking in historiometric enquiries.

The second limitation concerns the fact that the correlations between melodic originality and other significant factors are far from perfect. At no time can we say that even a quarter of the variance is accounted for by any one relationship, and more often the effects hover around one or two per cent. Yet we should not expect it to be otherwise. Musical creativity, like other forms of creative behaviour, is a complex, multidetermined phenomenon (Simonton 1999*a*). It probably would take hundreds of variables to fully describe what happens when a composer puts notes on a piece of paper—or when a listener emotionally responds to their performance. In addition, it must be reiterated that the computerized assessment of melodic originality remains only a crude approximation. Composers have a great deal more in their expressive toolkit than the choice of pitches. It certainly makes a difference in emotional impact whether the tempo is adagio or presto, the rhythm regular or syncopated, the dynamics pianissimo or fortissimo, the melody line played on a violin or tuba (cf. Gabrielsson & Lindström, this volume). Indeed, when the current scheme is compared with various alternative content-analytical measures, its simplicity becomes quite obvious (see, for example, Cerulo 1988, 1989). A certain price has been paid to quantify efficiently and objectively so many themes and their corresponding compositions.

All that admitted, the fact remains that the discovered relationships, whatever their limitations, hold for all of the works that dominate the classical repertoire. Certainly, 15 618 themes by 479 composers constitute a far more representative sample than seen in any other investigation, whether musicological or experimental. Therefore, it may be worthwhile to lend serious consideration to the findings that connect the computer-generated melodic originality with a provocative array of variables. Scores on this content-analytical attribute have been empirically linked with various characteristics of classical works, including their key, rhythm, medium, form, and compositional structure. Even more pertinent here, melodic originality seems to bear some connection with a

theme's emotional expressiveness. This latter consequence, however modest, has implications for understanding: (a) how compositions vary in their repertoire popularity, listener accessibility, and aesthetic significance; and (b) how melodic originality changes during the development of a composer's career and during the evolution of musical history. At the same time, the melodic originality of particular themes appears to have some foundation in the life experiences prevailing at the time the works were created by the composer. Stressful and even traumatic events tend to intensify the level of melodic originality in concomitant compositions.

All of these empirical relationships imply that melodic originality, at least to some degree, forms part of the composer's tools for expressing emotion. It seems most likely that more sophisticated methods of assessing the expressive qualities of music will reinforce rather than contradict this conclusion.

## References

**Anon** (1991). *Who said what when: A chronological dictionary of quotations.* New York: Hippocrene Books.

**Arnheim, R.** (1986). *New essays on the psychology of art.* Berkeley, CA: University of California Press.

**Barlow, H. & Morgenstern, S.** (1948). *A dictionary of musical themes.* New York: Crown.

**Barlow, H. & Morgenstern, S.** (1950). *A dictionary of vocal themes.* New York: Crown.

**Berlyne, D. E.** (1971). *Aesthetics and psychobiology.* New York: Appleton-Century-Crofts.

**Berlyne, D. E.** (ed.) (1974). *Studies in the new experimental aesthetics.* Washington, DC: Hemisphere.

**Brook, B. S.** (1969). Style and content analysis in music: The simplified 'Plaine and Easie Code'. In *The analysis of communication content,* (ed. G. Gerbner, O. R. Holsti, K. Krippendorff, W. J. Paisley, & P. J. Stone), pp. 287–96. New York: Wiley.

**Cerulo, K. A.** (1984). Social disruption and its effects on music: An empirical analysis. *Social Forces,* 62, 885–904.

**Cerulo, K. A.** (1988). Analyzing cultural products: A new method of measurement. *Social Science Research,* 17, 317–52.

**Cerulo, K. A.** (1989). Variations in musical syntax: Patterns of measurement. *Communication Research,* 16, 204–35.

**Einstein, A.** (1956). *Essays on music.* New York: Norton.

**Farnsworth, P. R.** (1969). *The social psychology of music,* (2nd edn). Ames, IW: Iowa State University Press.

**Halsey, R. S.** (1976). *Classical music recordings for home and library.* Chicago: American Library Association.

**Holmes, T. S. & Holmes, T. H.** (1970). Short-term intrusions into the life style routine. *Journal of Psychosomatic Research,* **14**, 121–32.

**Holmes, T. S. & Rahe, R. H.** (1967). The social readjustment rating scale. *Journal of Psychosomatic Research,* 11, 213–18.

**Holsti, O. R.** (1969). *Content analysis for the social sciences and humanities.* Reading, MA: Addison-Wesley.

**Kammann, R.** (1966). Verbal complexity and preferences in poetry. *Journal of Verbal Learning and Verbal Behavior,* 5, 536–40.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology,* 51, 336–52.

La Rue, J. (1970). *Guidelines for style analysis.* New York: Norton.

Landowska, W. (1964). *Landowska on music,* (ed. and trans. D. Restout). New York: Stein & Day.

Lindauer, M. S. (1993). The old-age style and its artists. *Empirical Studies and the Arts,* 11, 135–46.

Martindale, C. (1990). *The clockwork muse: The predictability of artistic styles.* New York: Basic Books.

Martindale, C. & Moore, K. (1989). Relationship of musical preference to collative, ecological, and psychophysical variables. *Music Perception,* 6, 431–46.

Martindale, C. & Uemura, A. (1983). Stylistic evolution in European music. *Leonardo,* 16, 225–8.

Meyer, L. B. (1956). *Emotion and meaning in music.* Chicago: University of Chicago Press.

Moles, A. (1968). *Information theory and esthetic perception* (trans. J. E. Cohen). Urbana, IL: University of Illinois Press (originally published 1958).

Münsterberg, H. (1983). *The crown of life: Artistic creativity in old age.* San Diego, CA: Harcourt-Brace-Jovanovich.

Paisley, W. J. (1964). Identifying the unknown communicator in painting, literature and music: The significance of minor encoding habits. *Journal of Communication,* 14, 219–37.

Porter, C. A. & Suedfeld, P. (1981). Integrative complexity in the correspondence of literary figures: Effects of personal and societal stress. *Journal of Personality and Social Psychology,* 40, 321–30.

Pratt, C. C. (1954). The design of music. *Journal of Aesthetics and Art Criticism,* 12, 289–300.

Schermann, T. K. & Biancolli, L. (ed.) (1972). *The Beethoven companion.* Garden City, NY: Doubleday.

Simonton, D. K. (1977*a*). Creative productivity, age, and stress: A biographical time-series analysis of 10 classical composers. *Journal of Personality and Social Psychology,* 35, 791–804.

Simonton, D. K. (1977*b*). Eminence, creativity, and geographic marginality: A recursive structural equation model. *Journal of Personality and Social Psychology,* 35, 805–16.

Simonton, D. K. (1980*a*). Thematic fame and melodic originality in classical music: A multivariate computer-content analysis. *Journal of Personality,* 48, 206–19.

Simonton, D. K. (1980*b*). Thematic fame, melodic originality, and musical zeitgeist: A biographical and transhistorical content analysis. *Journal of Personality and Social Psychology,* 38, 972–83.

Simonton, D. K. (1983). Esthetics, biography, and history in musical creativity. In *Documentary report of the Ann Arbor Symposium,* Session 3, pp. 41–8. Reston, VA: Music Educators National Conference.

Simonton, D. K. (1984). Melodic structure and note transition probabilities: A content analysis of 15,618 classical themes. *Psychology of Music,* 12, 3–16.

Simonton, D. K. (1986). Aesthetic success in classical music: A computer analysis of 1935 compositions. *Empirical Studies of the Arts,* 4, 1–17.

Simonton, D. K. (1987). Musical aesthetics and creativity in Beethoven: A computer analysis of 105 compositions. *Empirical Studies of the Arts,* 5, 87–104.

Simonton, D. K. (1989). The swan-song phenomenon: Last-works effects for 172 classical composers. *Psychology and Aging,* 4, 42–7.

Simonton, D. K. (1990). *Psychology, science, and history: An introduction to historiometry.* New Haven, CT: Yale University Press.

**Simonton, D. K.** (1991). Emergence and realization of genius: The lives and works of 120 classical composers. *Journal of Personality and Social Psychology,* **61**, 829–40.

**Simonton, D. K.** (1994). Computer content analysis of melodic structure: Classical composers and their compositions. *Psychology of Music,* **22**, 31–43.

**Simonton, D. K.** (1995). Drawing inferences from symphonic programs: Musical attributes versus listener attributions. *Music Perception,* **12**, 307–22.

**Simonton, D. K.** (1997). Products, persons, and periods: Historiometric analyses of compositional creativity. In *The social psychology of music,* (ed. D. J. Hargreaves & A. C. North), pp. 107–22. Oxford, UK: Oxford University Press.

**Simonton, D. K.** (1998*a*). Fickle fashion versus immortal fame: Transhistorical assessments of creative products in the opera house. *Journal of Personality and Social Psychology,* **75**, 198–210.

**Simonton, D. K.** (1998*b*). Mad King George: The impact of personal and political stress on mental and physical health. *Journal of Personality,* **66**, 443–66.

**Simonton, D. K.** (1999*a*). *Origins of genius: Darwinian perspectives on creativity.* New York: Oxford University Press.

**Simonton, D. K.** (1999*b*). Significant samples: The psychological study of eminent individuals. *Psychological Methods,* **4**, 425–51.

**Steck, L. & Machotka, P.** (1975). Preference for musical complexity: Effects of context. *Journal of Experimental Psychology: Human Perception and Performance,* **104**, 170–4.

**Suedfeld, P. & Piedrahita, L. E.** (1984). Intimations of mortality: Integrative simplification as a predictor of death. *Journal of Personality and Social Psychology,* **47**, 848–52.

**Vitz, P. C.** (1964). Preferences for rates of information presented by sequences of tones. *Journal of Experimental Psychology,* **68**, 176–83.

**Woods, F. A.** (1909). A new name for a new science. *Science,* **30**, 703–4.

**Woods, F. A.** (1911). Historiometry as an exact science. *Science,* **33**, 568–74.

CHAPTER 10

# THE INFLUENCE OF MUSICAL STRUCTURE ON EMOTIONAL EXPRESSION

ALF GABRIELSSON AND ERIK LINDSTRÖM

The expressive qualities of music have been a matter of discussion by philosophers and music theorists ever since antiquity and, more recently, by music psychologists. Many kinds of expression have been proposed, but expression of emotion is undoubtedly the most frequent. Likewise, most empirical research has focused on emotional expression in attempts to find out, on the one hand, which emotions can be reliably expressed in music (reviewed in Gabrielsson & Juslin, in press) and, on the other hand, which factors in music contribute to the perceived emotional expression. The latter refer to factors in the composed musical structure represented in the musical notation, such as tempo, loudness, pitch, mode, melody, rhythm, harmony, and various formal properties.

While it is a popular conception, and sometimes perhaps true, that composers express their present feelings in their compositions (cf. Cook & Dibben, this volume; Simonton, this volume), a more plausible view is that composers try to use various structural factors in order to achieve certain intended expressions, different in different works, with little or no direct connection to their present feelings. 'A composer . . . knows the forms of emotions and can handle them, "compose" them' (Langer 1957, p. 222). Tchaikovsky, usually considered a very 'emotional' composer, wrote that 'Those who imagine that a creative artist can—through the medium of his art—express his feelings at the moment when he is *moved*, make the greatest mistake. Emotions—sad or joyful—can only be expressed *retrospectively* . . . a work composed in the happiest surroundings may be touched with dark and gloomy colors' (cited in Fisk 1997, p. 157). For further discussion of these questions, see, for instance, Davies (1994, Chapters 4–6; see also Davies, this volume).

The purpose of this chapter is to review empirical research concerning the influence of different factors in musical structure on perceived emotional expression. However, because listeners usually judge perceived expression of composed music as realized in performance, there is a confounding of the properties of the composed structure and properties of the actual performance. As a rule, performance involves various modifications of the notated structure, for instance, variations of tempo, loudness, articulation, intonation, deviations in note timing, etc. (for a review, see Gabrielsson 1999). Perceived expression is thus dependent both on factors in the composed structure and factors in the performance. Only recently have the influence of the latter been subject to systematic

investigations (see Juslin, this volume). Listeners' perception of emotional expression—for instance, to perceive a piece of music as 'happy'—should be distinguished from listeners' own emotional reactions—for instance, to feel happy. Listeners' emotional reactions are treated elsewhere in this volume (Gabrielsson; Scherer & Zentner; Sloboda & O'Neill).

In Section 10.1, we review different methodological approaches to study the influence of musical structure on perceived emotional expression. The results are summarized in Section 10.2, and implications for future research are discussed in the final section.

## 10.1 Different methodological approaches

Empirical studies of musical expression had a cautious start in late nineteenth century, and it was not until the 1930s that they became more common. Subjects listened to pieces of music or other tonal stimuli and reported the perceived expression by either: (a) free phenomenological descriptions; (b) choice among descriptive terms, adjectives, or nouns, provided by the investigator; or (c) ratings of how well such descriptive terms applied to the music in question. Free descriptions were subjected to content analysis. Listeners' choice among descriptive terms was analysed regarding the frequency with which each term was chosen and regarding intersubject agreement. Ratings were usually analysed by multivariate techniques, such as factor analysis, cluster analysis, and multidimensional scaling in order to find a limited number of fundamental descriptive dimensions. Recently, various techniques for continuous and/or non-verbal recording of perceived expression have been developed.

Stimuli were usually recordings of selected pieces of music. However, before common use of modern sound recording the pieces were performed live by one or more musicians (e.g. Downey 1897; Gilman 1891; Hevner 1935*a*; Rigg 1937). The relationship between the composed structure and perceived expression was studied *post hoc* by analysing the musical scores (e.g. Gundlach 1935; Nielsen 1983) or having musical experts judge the selected pieces with regard to structural properties (e.g. Kleinen 1968; Watson 1942; Wedin 1972*c*).

Studying emotional expression using real music (see Section 10.1.1) means good ecological validity; on the other hand, conclusions regarding the effects of separate structural factors can only be tentative since they are confounded. In contrast to this, there are studies in which the researcher systematically varies one or more structural factors (e.g. loudness, pitch level) in short sound sequences without musical context (see Section 10.1.2). Ecological validity is then limited, but the systematic variation allows more definite conclusions regarding the effects of separate factors. A compromise between these two approaches is to use systematic manipulation of various factors within a musical context (see Section 10.1.3), for instance, to systematically vary tempo, mode, or melodic direction in real pieces of music and thus try to combine the advantages of the before-mentioned strategies. However, some types of manipulation may result in musically unnatural stimuli, thus jeopardizing ecological validity.

### 10.1.1 Studying emotional expression using real music

*Free descriptions and choice among descriptive terms*

The earliest reported empirical investigations on musical expression may be those by Gilman (1891, 1892) and Downey (1897). Their subjects listened to live performances on the piano, or piano and violin, of classical music and gave free phenomenological reports on perceived expression. A variety of emotions was reported and some tentative relationships were noted, for instance, descending thirds in an aria by Handel to indicate sadness.

Gundlach (1935) presented 40 musical phrases from different classical compositions to listeners, who were instructed to determine 'what mood or attitude the composer has succeeded in expressing' (p. 628) by choosing one or more out of 17 descriptive terms. From his analysis Gundlach concluded that speed was by far the most important factor for perceived expression, followed by rhythm, interval distribution, orchestral range, loudness, mean pitch, and melodic range.

Rigg (1937) used 20 short phrases played on the piano that were supposed to express either joy, lamentation, longing, or love. Listeners answered the question 'What emotion is suggested to you by each passage?' either by giving free descriptions or by choosing among the above emotion terms plus anger, fear, and disgust. Rigg discussed the advantages and disadvantages of these procedures. With free descriptions, responses may be scattered, but coherence in listeners' responses would be of more significance than if they merely checked terms in a list. With choice among descriptive terms, responses may be nicely bunched and easier to treat but will partly be a product of the given terms. For instance, love was not often mentioned in free descriptions but more frequently chosen when given as an alternative among the emotion terms. Expression of joy was convincingly indicated by the listeners, and some musical characteristics of the joy excerpts were rapid tempo, major mode, simple harmony, staccato notes, and forte dynamics.

Watson (1942) had 20 musical experts mark which out of 15 selected adjectives were appropriate for each of 30 musical examples. They also judged the examples, using a five-step scale, regarding pitch (low–high), loudness (soft–loud), tempo (slow–fast), sound (pretty–ugly), dynamics (no quick changes in loudness–very many changes), and rhythm (regular–irregular), thus enabling a study of the relationships between these factors and perceived expression. For instance, high pitch and fast tempo tended to express happiness and excitement; low pitch and slow tempo, sadness; high loudness, excitement; and small dynamic range, expression of dignity, sadness, and peacefulness.

*Ratings and multivariate analysis techniques*

In studies by Kleinen (1968), Wedin (1969, 1972*a*), Gabrielsson (1973, Experiments 4–5), and Nielzén and Cesarec (1981, 1982) subjects judged selected pieces of music using a large number of rating scales. The correlations between the scales were subjected to factor analysis in order to obtain a limited number of fundamental factors. These were given similar interpretations in all studies: tension/energy, gaiety–gloom, and solemnity–triviality (Wedin 1969, 1972*a*); intensity–softness, pleasantness–unpleasantness, and solemnity–triviality (Wedin 1972*b*, 1972*c*; in these studies using

multidimensional scaling techniques); tension–relaxation, gaiety–gloom, and attraction–repulsion (Nielzén & Cesarec 1981); gay/vital–dull and excited–calm (Gabrielsson 1973); '*Heiterkeit-Ernst*' (cheerful–serious) and '*Robustheit-Zartheit*' (strong/powerful-soft/tender) (Kleinen 1968).

Musical experts rated the respective pieces of music with regard to structural properties. For example, Wedin (1972*c*) used five-point scales for rating intensity (*pp–ff*), pitch (bass–treble), rhythm (outstanding–vague), tempo (fast–slow), rhythmic articulation (firm–fluent, staccato–legato), harmony (dissonant/complex–consonant/simple), tonality (atonal–tonal), modality (major–minor), melody (melodious–unmelodious), type of music (serious–popular), and style (date of composition).

To avoid listeners' dependence on descriptive terms chosen by the investigator, Imberty (1979) had 80 non-musical subjects freely choose any adjectives which came to their mind while listening to excerpts from Debussy's *Préludes* for the piano. Out of a total of 1063 adjectives, 172 were retained for analysis by correspondence analysis, resulting in a joint representation of the musical excerpts and the adjectives on a number of dimensions. The first two dimensions were *Les schèmes de tension et de détente* (tension vs. relaxation) and *Les schèmes de résonances émotionelles* (positive vs. negative emotions). Imberty went on to construct indices for formal complexity and *dynamisme général* in the used musical stimuli by combining, in various ways, variables as note duration, *intervalle métrique* (duration between two accented notes), density of notes per time unit, loudness, accents, syncopations, and certain characteristics of melodic, harmonic, and rhythmic motives. Using these indices, he claimed that: (a) low complexity combined with average dynamism means formal integration and expression of positive emotions; (b) high formal complexity combined with low dynamism means formal disintegration and expression of melancholy and depression; and (c) high formal complexity combined with high dynamism means formal disintegration and expression of anxiety and aggressiveness. This is one of the few attempts to investigate how emotional expression is influenced by properties of musical form. However, the indices are complex, and require much analytical work and further validation.

With regard to the dimensional structure of emotional expression, the above studies show similar results converging on two main emotion dimensions: valence (pleasantness–unpleasantness, gaiety–gloom) and activity/arousal (tension–relaxation, excited–calm).

### *Continuous recording of emotional expression*

As type and intensity of emotional expression usually vary during the course of a piece of music, researchers may want to use a technique for continuous recording of perceived expression. This may also allow a more detailed study of how factors in the musical structure affect perceived expression. Nielsen (1983, 1987) pioneered the development of such a technique for continuous recording of perceived tension in music. While listening to the music, his subjects continuously pressed a pair of tongs in proportion to the perceived tension in the music—the more tension, the harder the press, and vice versa. The pressures were registered on a polygraph showing 'pressure curves' with waves of heightened tension alternating with periods of relaxation. Tension peaks were mainly conditioned by high intensity (fortissimo), but increased tension could also be

related to ascending melody, increased note density, dissonance, harmonic complexity, rhythmic complexity, and formal properties such as repetition of various units, condensation of musical material, sequential development, and pauses.

Later studies of perceived tension using other response devices include Madsen and Fredrickson (1993), Fredrickson (1997), and Krumhansl (1996). Krumhansl found that tension peaks followed by a rapid decrease occurred at the end of large-scale sections and that judgments of new musical ideas co-occurred with low tension levels. Continuous recording of perceived emotional expression was also studied by Namba *et al.* (1991), Krumhansl (1997), Madsen (1997), and Schubert (1999) using various computerized techniques (for a more extensive review, see Schubert, this volume).

*Non-verbal responses*

Some methods for studying tension have been mainly non-verbal (e.g. Nielsen 1983), but use of non-verbal responses to indicate emotional expression is rare. Clynes (1977) introduced finger pressure on a so-called *sentograph*—a device for recording the pressure exerted by a finger upon a small disc—to study how subjects expressed different emotions. He found different pressure patterns (sentograms) with regard to duration and dynamics for different emotions (joy, grief, anger, hatred, reverence, love, sex; depicted in Clynes 1977, p. 29 onwards) and then tried to trace corresponding patterns in music (Clynes & Walker 1982). De Vries (1991) had 30 subjects use a sentograph while listening to ten pieces of music and found that sentograms were different for different pieces, and that there were similarities between the sentograms for pieces with a certain emotional expression (anger, grief, joy, love) and Clynes's sentograms for the corresponding emotions. Some evidence in the same direction was found in Gabrielsson (1995) and Gabrielsson and Lindström (1995). (Another example of non-verbal responses—drawings to describe intervals—appears in 10.1.2.)

*Specially composed music*

Thompson and Robitaille (1992) asked composers to compose short monophonic melodies to express joy, sorrow, excitement, dullness, anger, and peace. Their scores were transformed to sounding music using computer software and sampled grand piano sound. Listeners, moderately trained in music, rated the pieces on corresponding emotion scales. In general they perceived the intended expressions. Analysis of the scores showed that joyful melodies were strongly tonal and rhythmically varied. Sad melodies were slow with implied minor or chromatic harmony. Melodies for excitement were fast, and contained intervallic leaps and high pitches. Dull melodies were tonal in a stepwise motion. Angry melodies were rhythmically complex and with implied chromatic harmony or atonality. Melodies for peacefulness were tonal, slow, and often involved stepwise motion leading to melodic leaps.

### 10.1.2 Manipulation of structural factors without musical context

*Intervals*

A distinction must be made between harmonic (simultaneous) intervals (two-tone chords, bichords) and melodic (successive) intervals. Costa *et al.* (2000) used all 12 equal-tempered harmonic intervals (organ sound) contained in an octave, one set in a

low register (geometric mean 185 Hz) and another in a high register (geometric mean 1510 Hz) and had listeners rate them on 30 bipolar scales, reflecting 'emotional evaluation' (e.g. happy–sad, pleasant–painful, friendly–hostile), 'activity' (e.g. tense–relaxed, dissonant–consonant, worried–calm), and 'potency' (e.g. strong–weak, vigorous–languid) in accordance with the three dominant dimensions of meaning according to Osgood *et al.* (1957). High register bichords were rated higher in 'activity' (e.g. more unstable, restless, tense) and in 'potency' (stronger) than low register bichords, and low register bichords were evaluated more negatively than high register bichords. Dissonant bichords were judged as more negative, more active, and stronger than consonant bichords; minor bichords as more dull and weaker than major bichords. Females showed greater polarization of scores and judged bichords as more active and tense than males. Difference in musical experience had no effect. Generally, intervals differed more in emotional evaluation and activity than in potency. Earlier, Maher (1980), using 14 different harmonic intervals in just intonation with triangular waveforms in two sets, one high pitched (geometric mean frequency 500 Hz) and one low pitched (250 Hz), obtained some similar results. High-pitched intervals were rated more happy and powerful than low-pitched intervals. Minor and major second, minor and major seventh, and minor ninth were the most displeasing.

Maher and Berlyne (1982) investigated the perception of 12 melodic intervals, sine-wave intervals in tempered intonation (minor second to octave, rising and falling, with the common 500 Hz geometric mean frequency). Larger intervals were judged as more 'powerful' than smaller ones; minor second was considered the most 'melancholic' interval; whereas octave, fourth, fifth, major sixth, and minor seventh were the most 'carefree'.

Smith and Adams (1999) had South African children listen to six ascending intervals performed on a flute and to represent them in drawings. The thematic content of the drawings was assessed by independent raters and also rated by university students using adjective scales. The primary distinction in the findings was between consonance and dissonance, the latter being associated with unpleasant events, the former with home, family, and stability. Major seventh and augmented fourth tended to be associated with danger and violence, perfect fifth with activity, and perfect fourth with being desolate and old. The octave was perceived as positive and strong, and major third as neutral with no particular meaning. The children's responses were influenced by their level of schooling, race, and living environment.

Earlier, Huber (1923) presented short pitch patterns to musical listeners who were asked to freely describe perceived expression. The reports were classified into: (a) mood impressions; (b) impressions of human character, for instance an elderly man; (c) emotionally coloured announcements, such as a call, question, or request; (d) impressions of movement; and (e) various inner images. Referring to Huber, Langer (1957, p. 231) remarked that 'The entire study shows effectively how many factors of possible expressive virtue are involved even in the simplest musical structure'.

### *Mode*

Heinlein (1928) studied listeners' perception of major and minor chords in all keys, at different intensities and different pitch levels. Musically trained and untrained subjects described the perceived expression, choosing among adjectives such as bright, happy,

joyful (supposed to reflect expression of major mode), melancholy, sad, and soothing (minor mode). None of the 30 subjects got all responses 'correct': there were many sad-type responses to major chords and happy-type responses to minor chords. Loud chords and chords at a high pitch level evoked more happy-type responses than soft chords and chords at a low pitch level, irrespective of mode. Crowder (1984) showed in a reanalysis of Heinlein's data that all but one of the 30 listeners in fact gave more happy-type responses to major chords than to minor chords, that is, in accordance with the conventional association. Crowder (1985) found the conventional major–happy and minor–sad associations in young adults who listened to sine-wave triads in three tonalities and two inversions of each chord. The major–happy and minor–sad associations were also found in many investigations using real music (Hevner 1935*a*; Kleinen 1968; Krumhansl 1997; Nielzén & Cesarec 1982; Peretz *et al.* 1998; Wedin 1972*c*).

*Rhythm and tempo*

Motte-Haber (1968) had ten rhythm patterns, differing in metre, sound event density, and homogeneity (e.g. uniform or dotted pattern), presented at three different metronomic tempos (ratios 1 : 2 : 4) to listeners, who rated them on a large number of bipolar scales (*Polaritätsprofile*). Metronomically rapid rhythms were rated happier (*fröhlicher*) than slow rhythms. Happiness ratings were strongly correlated with ratings of *subjective* tempo, which was related not only to metronomic tempo but also to sound event density and other rhythm characteristics in interaction with metronomic tempo. Similar relations appeared in Gabrielsson (1973, Experiments 1–3) using other rhythm patterns.

*Melodic properties*

Gabriel (1978) investigated the perceived expression of 16 short tone sequences, 'basic terms', which music theorist Deryck Cooke (1959) claimed recur with the same expression in Western music from the late Middle Ages up to the twentieth century. For instance, an ascending major triad (1–3–5) with possible insertions of the intervening notes (2 or 4), was said to express 'an outgoing, active, assertion of joy' (Cooke 1959, p. 115). Its counterpart in minor would be 'expressive of an outgoing feeling of pain—an assertion of sorrow, a complaint, a protest against misfortune' (Cooke 1959, p. 122). Gabriel generated the 16 basic terms using sinusoidals, constant tempo, and uniform rhythm. Students listened to these 'reduced' basic terms and had to judge whether Cooke's characteristics were adequate. The result was mainly negative. However, this experiment has in its turn been criticized regarding the choice of stimuli, lack of context, and use of musically untrained listeners (Cazden 1979; Gabriel 1979; Nettheim 1979; Sloboda 1985, pp. 60–4; for a discussion of Cooke's theory, see Cook & Dibben, this volume).

*Synthesized tone sequences*

Scherer and Oshinsky (1977) investigated the relative importance of several factors by systematic manipulations of synthesized eight-tone sequences (no musical notation was provided, which makes it difficult to judge their musical qualities). The manipulated variables were amplitude variation (small–large), pitch level (high–low), pitch contour

(up–down), pitch variation (small–large), rhythm (even–uneven), tonality (major, minor, atonal), tempo (slow–fast), envelope (round, sharp), and filtration cut-off level (intermediate–high). Psychology students rated the stimuli on three bipolar scales: pleasantness–unpleasantness, activity–passivity, and potency–weakness, and also indicated whether each sequence expressed any or more of the following emotions: happiness, sadness, anger, fear, boredom, surprise, and disgust—the last five rarely or never included in earlier studies of emotional expression in music. Multiple regression analysis was used to estimate the predictive strength of each acoustic parameter on each emotion dimension. Most of the results can be found in Table 10.2 (see below). As seen there, any acoustical variable was associated with many emotions, listed in decreasing order of associative strength, in some cases even with seemingly 'opposite' emotions (e.g. low pitch associated with boredom, pleasantness, and sadness).

### 10.1.3 Manipulation of structural factors in musical context

*Manipulation of several factors*

The earliest and still best-known investigations with systematic manipulations of various factors in real music are those by Hevner (1935*a*, 1935*b*, 1936, 1937). She arranged a large number of emotion terms in eight clusters in a circular configuration, an 'adjective circle' (Fig. 10.1). The terms within each cluster were supposed to be close in meaning, and adjacent clusters should deviate slightly by cumulative steps until reaching a contrast in the opposite position. Inspection of Fig. 10.1 suggests an implicit dimensionality similar to that described in Section 10.1.1; in other words, valence (sadness–happiness, cluster II–cluster VI) and activity/arousal (exciting/vigorous–serene/dreamy, clusters VII/VIII–clusters IV/III). The circular configuration is reminiscent of the circumplex model by Russell (1980; see also Sloboda & Juslin, this volume) and may be one of the earliest indications of this conception, earlier than that proposed by Schlosberg in 1941 (cf. Remington *et al.* 2000). Listeners were instructed to mark as many of the terms as they found appropriate for each piece. Hevner selected short pieces of tonal music, but besides the original version she also constructed a variant that differed from the original in: (a) mode, a piece in major mode was also played in minor mode; notated example in Hevner (1935*a*); (b) melodic direction (ascending vs. descending), harmony (simple consonant harmonies vs. complex dissonant harmonies), and rhythm (firm vs. flowing, that is, a firm beat with a chord on every beat vs. a flowing motion in which the chords were broken up); notated examples in Hevner (1936); and (c) tempo (fast vs. slow) and pitch level (one octave or more apart); see Hevner (1937).

Any difference in listeners' choice of descriptive terms between the respective two versions of a piece could then be ascribed to the difference in the manipulated variable. All versions were performed by an experienced pianist, and listeners were hundreds of students, most of them without special musical training. The joint conclusions of several experiments were that variables with the largest effects on listeners' judgments were tempo and mode, followed by pitch level, harmony, and rhythm, whereas melodic direction had little if any effect (Hevner 1937). This order can be seen in Table 10.1, which was Hevner's summary of results from six experiments. It shows the relative

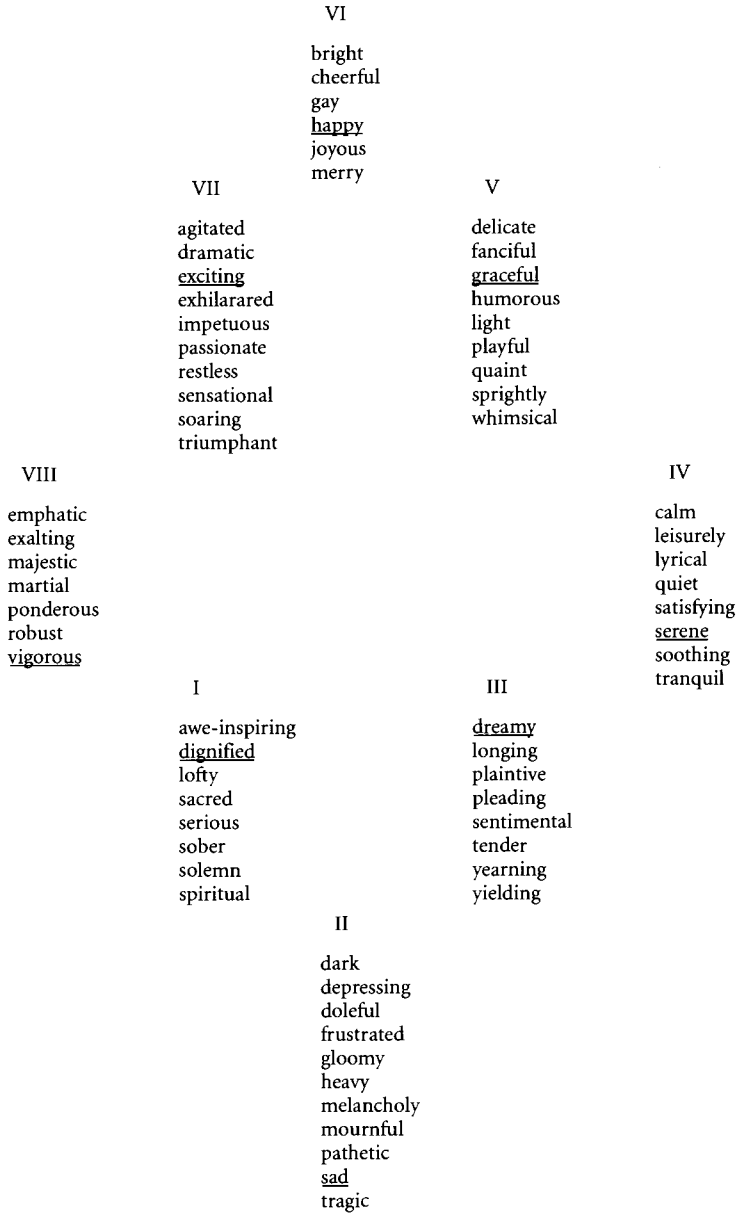



**Figure 10.1** Hevner's adjective circle. Adjectives within each cluster appear in alphabetical order. Adjectives used by Hevner (1936, 1937) to represent the respective clusters are underlined (adapted from Farnsworth 1954, p. 98).

weights of each musical factor for each emotion cluster; for instance, minor mode, low pitch, and slow tempo were most important, in that order, for the sad–heavy cluster,

whereas major mode, fast tempo, and simple harmony were most important for the happy–bright cluster.

Hevner (1935*b*, 1937) noted the difficulties of manipulating real pieces of music and still having the manipulated version sound musically acceptable (cf. Behne 1972; Peretz *et al.* 1998, p. 118, footnote). She emphasized the necessity of using pieces with similar emotional expression throughout the whole piece and remarked that the results should only be interpreted in a relative and contextual sense. For instance, although happiness was usually associated with major mode, a piece in minor mode may sound happy due to other factors, such as tempo and rhythm. She warned against drawing too far-going conclusions since results may be dependent on the selected pieces and, above all, since emotional expression usually results from several musical factors in complex interplay. These arguments should be carefully considered by any researcher on emotional expression in music.

Rigg (1939) composed five four-bar phrases supposed to represent two emotion categories: pleasant/happy or sad/serious. They were systematically manipulated regarding tempo, mode, articulation (legato, staccato), pitch level, loudness, rhythm (iambic, trochaic), and certain intervals in an impressively large number of variations. Listeners described perceived expression by choosing first among the above-mentioned alternatives and then among subcategories within each of them (e.g. hopeful longing, sorrowful longing). In a subsequent experiment (Rigg 1940*a*) the same phrases were transposed up or down an octave, to the dominant (a fifth upwards or a fourth downwards), a second down, and a minor second up. Shifts an octave upward made the phrase happier or less sorrowful, and vice versa. There were less effects for the dominant transposition and practically no effect of the smallest transpositions. Finally, Rigg (1940*b*) had the same five phrases played by a pianist at six different tempos, from 60 to 160 bpm (beat = quarter note), and found that the higher the tempo, the more pleasant/happy and the less serious/sad judgments were given. However, one of the phrases, supposed to express lamentation, still sounded sad even at the highest tempo, and for another one, supposed to express sorrowful longing, there was about the same number of happy and sad responses at the highest tempo. Both these phrases were in minor mode and had descending minor or major seconds in the upper voice. Rigg (1964) reviewed the results obtained by Gundlach (1935), Hevner (1935*a*, 1936, 1937), Rigg (1937, 1939, 1940*a*, 1940*b*), and Watson (1942).

Juslin (1997, Experiment 2) used synthesized versions of 'Nobody knows the trouble I've seen' manipulating tempo (slow, medium, fast), sound level (low, medium, high),

Table 10.1 Hevner's summary of results from six experiments (adapted from Hevner 1937, p. 626)

| Musical factor | Dignified/ solemn | Sad/ heavy | Dreamy/ sentimental | Serene/ gentle | Graceful/ sparkling | Happy/ bright | Exciting/ elated | Vigorous/ majestic |
|---|---|---|---|---|---|---|---|---|
| Mode | Minor 4 | Minor 20 | Minor 12 | Major 3 | Major 21 | Major 24 | – | – |
| Tempo | Slow 14 | Slow 12 | Slow 16 | Slow 20 | Fast 6 | Fast 20 | Fast 21 | Fast 6 |
| Pitch | Low 10 | Low 19 | High 6 | High 8 | High 16 | High 6 | Low 9 | Low 13 |
| Rhythm | Firm 18 | Firm 3 | Flowing 9 | Flowing 2 | Flowing 8 | Flowing 10 | Firm 2 | Firm 10 |
| Harmony | Simple 3 | Complex 7 | Simple 4 | Simple 10 | Simple 12 | Simple 16 | Complex 14 | Complex 8 |
| Melody | Ascend 4 | – | – | Ascend 3 | Descend 3 | – | Descend 7 | Descend 8 |

The numbers indicate the relative weight of each musical factor (left column) for each emotion cluster.

frequency spectrum (soft, bright, sharp), articulation (legato, staccato), and tone attack (slow, fast). Students rated these versions on six adjectve scales: happy, sad, angry, fearful, tender, and expressive. Multiple regression analyses showed that the variance accounted for ($R^2$) by the manipulated factors was typically 0.78–0.88, except for fearful (0.55) and expressive (0.46). Sharp spectrum (higher partials amplified) and high sound level were the most important predictors for angry; fast tempo for happy; slow tempo for sad; slow tempo, soft spectrum (higher partials attenuated), and legato articulation for tender; low sound level and staccato articulation for fearful; and legato articulation for expressive.

Lindström (1997) constructed 72 versions of 'Frère Jacques' by systematic manipulation of tonal progression (choice of notes affecting latent harmony), rhythm (even notes, 1 : 1, uneven 3 : 1 and 1 : 3 ratios, respectively), melodic contour (original and complex), and melodic direction (original and reversed). Listeners' judgments of emotional expression were most affected by changes of harmonic functions, followed by changes of rhythm and melodic contour. Inspection of interactions and most extreme ratings showed intricate interplay between several structural factors (Lindström 2000).

Peretz *et al.* (1998, Experiment 2) manipulated tempo and mode in synthesized versions of 32 excerpts of classical music and found that these factors affected both normal listeners' and a brain-injured listener's perception of happiness and sadness in similar ways. Moreover, exposition shorter than a second was enough for reliable recognition of these expressions. Kamenetsky *et al.* (1997) compared listeners' ratings of emotional expression in synthesized versions of two baroque and two romantic keyboard pieces. The 'original' version with variations in both tempo and dynamics was rated the most expressive. Removing tempo variations but keeping variations in dynamics reduced expressiveness only marginally, whereas removing variation in dynamics but keeping tempo variations significantly reduced perceived expressiveness.

Balkwill and Thompson (1999) asked two professional Hindustani musicians to perform a short *alap* (improvised opening section) from any raga they wanted to choose to evoke each of four emotions: joy, sadness, anger, and peace. Listening to these performances, Canadian students indicated which of the four emotions was most dominant in each performance and rated the degree to which the respective emotion was conveyed. They also rated the performances regarding tempo, melodic and rhythmic complexity, and pitch range. The recognition of intended emotions was successful except for peace, and data show some confusion between joy and anger and between sadness and peace. Judgments by four experts on Hindustani music showed a similar pattern. On the whole, then, Western listeners were sensitive to intended emotions in Hindustani music. Joy was associated with faster tempo and less melodic and rhythmic complexity; sadness with slower tempo and melodic and rhythmic complexity; anger with timbre of stringed instruments; and peace with slow tempo, flute timbre, and less melodic and rhythmic complexity.

### *Tempo*

Behne (1972) presented ten pieces of recorded music at three different tempos by means of a so-called Springer machine to music students who rated them using bipolar scales. With decreased tempo of the pieces, listeners rated them as more serious (*ernst*)

and complaining (*klagend*). In another experiment, six eight-bar versions of a theme were composed to represent three different basic tempos (slow, midway, rapid) and two levels of harmonic complexity (simple, complex). Each version was then performed at five different tempos by a woodwind trio and rated by 30 subjects. The (complex) results were discussed in relation to conceived tempo, performed tempo, and to melodic and harmonic information density. Hevner's observation that it is difficult to manipulate different factors independently of others in real music was confirmed. The two levels of harmonic complexity were not equivalent across the different tempos: versions with slow original tempo were in fact more harmonically complex than the versions with rapid original tempo at the (assumed) same complexity level.

*Mode*

Kastner and Crowder (1990) presented 12 short melodies, each of them in major or minor mode, to 38 children from preschool and elementary school. For each presentation they were asked to choose among four pictures of faces: happy, contented, sad, and angry. All children, even the youngest (3 years), showed a reliable positive-major and negative-minor association, not perfect but significantly different from what would happen on chance basis. In similar studies by Gerardi and Gerken (1995) and Gregory *et al.* (1996), the happy–major and sad–minor association was found in children aged 7–8 years and in young adults, but not in children aged 3–5 years. Thus, in all three studies the conventional happy–major and sad–minor association was present in older children, whereas the results for younger children are inconclusive.

Hill *et al.* (1996) studied which of two alternatives—salvation or condemnation—adults associated with a setting either in Ionian mode or in Phrygian mode of a well-known melody from early seventeenth century, variously known as *Herzlich tut mich verlangen* (promise of salvation, reward) or *Ach Herr, mich armen Sünder* (fear of condemnation, punishment), both in settings by J. S. Bach. Adults as well as children judged the Ionian mode more suitable for the 'salvation/reward' alternative and the Phrygian mode more suitable for the 'condemnation/punishment' alternative. The fact that Ionian mode is identical to major mode may have contributed to its association with the 'salvation/reward' alternative.

*Musical form*

Konečni and his co-workers (Gotlieb & Konečni 1985; Karno & Konečni 1992; Konečni 1984; Konečni & Karno 1994) demonstrated that changing the order of movements in Beethoven sonatas and string quartets, randomizing the order of variations in Bach's *Goldberg Variations*, or rearranging the order of different parts in sonata form as in the first movement of Mozart's Sympony in G Minor K. 550 had little or no effect on university students' ratings on various hedonic (e.g. beautiful, pleasing) or emotion-related (e.g. exciting, emotional) scales.

Tillman and Bigand (1996) chunked three pieces by Bach, Mozart, and Schoenberg into (musically adequate) segments of about 6 seconds and then played these either in original or in backward order to university students, who rated them on 27 (mostly) emotion-related scales. There were significant differences among the *pieces* in all scales, but only two significant differences between the two *versions* (original, backward) of

each piece. They concluded that for these subjects musical expressiveness was mainly affected by local structures within the chunks, not by global musical structure. This is in good agreement with the main theme in *Music in the moment* by Levinson (1997).

## 10.2 Effects of separate musical factors

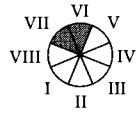
Most of the results from the studies reviewed above are summarized in Table 10.2. The first column contains different structural factors in alphabetical order, the second different levels within each factor, and the third column the corresponding emotional expression according to different studies. Authors' names are abbreviated to the initial two letters of the (first) author and publication year to the two last digits (e.g. He36 = Hevner 1936). Studies are grouped into: (A) early studies using choice among descriptive terms; (B) studies based on multivariate analyses; and (C) later experimental studies; thus offering alternative bases for interpretation of the results. As a rule, the emotions are listed in decreasing order of association with the corresponding factor. Following Rigg (1964), results of the early studies (A) are described in terms of Hevner's clusters designated by Roman numbers and the adjective that Hevner used as a label for each cluster (Hevner 1936, p. 265), with the use of shaded areas in the adjective circle.

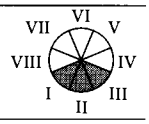
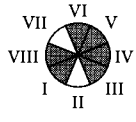
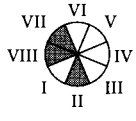
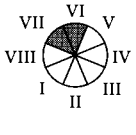
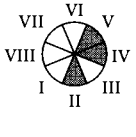
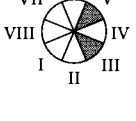
### 10.2.1. Tempo

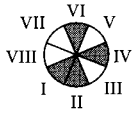
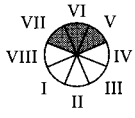
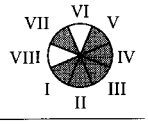
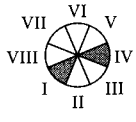
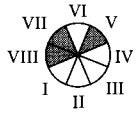
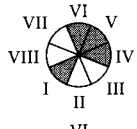
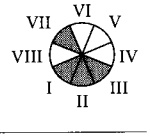
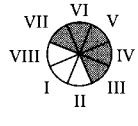
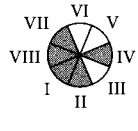
Among factors affecting emotional expression in music, tempo is usually considered the most important (cf. Gundlach 1935; Hevner 1937; Juslin 1997; Rigg 1964; Scherer & Oshinsky 1977). Composers as a rule indicate intended tempo, either exactly—for instance, MM (Mälzel metronome) or bpm (beats per minute) = 100—or with designations such as presto, allegro, moderato, andante, and adagio.

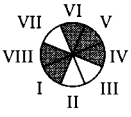
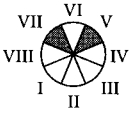
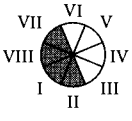
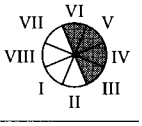
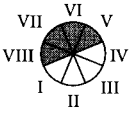
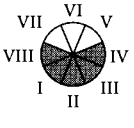
The studies under tempo in Table 10.2 indicate that fast tempo may be associated with various expressions of activity/excitement, happiness/joy/pleasantness, potency, surprise, anger, and fear. Slow tempo may be associated with various expressions of

Table 10.2 Summary of results from reviewed studies. Left two columns indicate structural factor and levels within each factor, middle column the associated emotional expression, right column the corresponding clusters in Hevner's adjective circle. Within each factor studies are grouped into (A) early studies using choice among descriptive terms, (B) studies based on multivariate analyses, and (C) later experimental studies. Authors' names are abbreviated to the initial two letters of the (first) author, publication year to the two last digits (e.g., He36 = Hevner, 1936). See text for further explanation.

| Factor | Levels | Emotional expression | Adjective circle (Hevner) |
|---|---|---|---|
| Amplitude envelope | Round | (C) Disgust, sadness, fear, boredom, potency (Sc77), tenderness, fear, sadness (Ju97) | |
| | Sharp | (C) Pleasantness, happiness, surprise, activity (Sc77), anger (Ju97) | |
| Articulation | Staccato | (A) VI: gaiety; VII: agitation (Ri39)<br>(B) Intensity/energy/activity (We72c), gaiety (Ni82)<br>(C) Fear, anger (Ju97) |  |

| Factor | Levels | Emotional expression | Adjective circle (Hevner) |
|---|---|---|---|
| | Legato | (A) I: solemn; II: melancholy, lamentation; III: longing (Ri39)<br>(B) Softness (We72c)<br>(C) Tenderness, sadness (Ju97) |  |
| Harmony | Simple/ consonant | (A) VI: happy (He36, Wa42), joy (Ri39; V: graceful; IV: serene; III: dreamy (He36); I: dignified (He36, Wa42), serious (Wa42), solemn (Ri39); VIII: majestic (Wa42)<br>(B) Gaiety, pleasantness (We72c), attraction (Ni82)<br>(C) Relaxation, tenderness (Li97) |  |
| | Complex/ dissonant | (A) VII: exciting (He36, Wa42), agitation (Ri39); VIII: vigorous (He36); II: sad (He36, Wa42)<br>(B) Gloom, unpleasantness (We72c), tension (Ni82, Kr96)<br>(C) Tension (Ni83, Kr96, Li97), fear (Kr97), anger (Li97) |  |
| Intervals | Harmonic: | | |
| | Consonant | (C) Pleasant, 'non-active' (Co00) | |
| | Dissonant | (C) Displeasing (Ma80), unpleasant, 'active', strong (Co00) | |
| | High-pitched | (C) Happy, powerful (Ma80), 'activity', potency (Co00) | |
| | Low-pitched | (C) Sad, less powerful (Ma80, Co00) | |
| | Melodic: | | |
| | Large | (C) Powerful (Ma82) | |
| | Minor 2nd | (C) Melancholy (Ma82) | |
| | Perfect 4th, perfect 5th, major 6th, minor 7th, octave | (C) Carefree (Ma82) | |
| | Perfect 5th | (C) Activity (Sm99) | |
| | Octave | (C) Positive/strong (Sm99) | |
| Loudness | Loud | (A) VII: excitement (Wa42), triumphant (Gu35); VI: joy (Ri39)<br>(B) Gaiety (Ni82), intensity (We72c), strength/power (Kl68), solemnity (We72c), tension (Ni83, Kr96)<br>(C) Anger (Ju97) |  |
| | Soft | (A) II: melancholy (Gu35); V: delicate (Gu35); IV: peaceful (Wa42)<br>(B) Softness (Kl68, We72c), tenderness (Kl68)<br>(C) Fear, tenderness, sadness (Ju97) |  |
| Loudness variation | Large | (C) Fear (Sc77) | |
| | Small | (C) Happiness, pleasantness, activity (Sc77) | |
| | Rapid changes | (A) V: playful, amusing; III: pleading (Wa42)<br>(C) Fear (Kr97) |  |

| Factor | Levels | Emotional expression | Adjective circle (Hevner) |
|---|---|---|---|
| | Few/no changes | (A) II: sad; IV: peaceful; I: dignified, serious; VI: happy (Wa42) |  |
| Melodic (pitch) range | Wide | (A) V: whimsical; VI: glad; VII: uneasy (Gu35)<br>(C) Fear (Kr97), joy (Ba99) |  |
| | Narrow | (A) I: dignified; II: melancholy; III: sentimental; IV: tranquil; V: delicate; VII: trimphant (Gu35)<br>(C) Sadness (Ba99) |  |
| Melodic direction | Ascending | (A) I: dignified; IV: serene (He36)<br>(C) Tension (Ni83, Kr96), happiness (Ge95) |  |
| | Descending | (A) VII: exciting; V: graceful; VIII: vigorous (He36)<br>(C) Sadness (Ge95) |  |
| Pitch contour | Up | (C) Fear, surprise, anger, potency (Sc77) | |
| | Down | (C) Sadness, boredom, pleasantness (Sc77) | |
| Melodic motion | Stepwise motion | (C) Dull melodies (Th92) | |
| | Intervallic leaps | (C) Excitement (Th92) | |
| | Stepwise + leaps | (C) Peacefulness (Th92) | |
| Mode | Major | (A) VI: happy (He36), joy (Ri39); V: graceful (He36); IV: serene (He36); I: solemn (Ri39)<br>(B) Happiness (Kl68, We72c), attraction (Ni82)<br>(C) Happiness (Sc77, Cr85, Kr97, Pe98) |  |
| | Minor | (A) II: sad (He36); lamentation (Ri39); III: dreamy; I: dignified (He36); VII: agitation (Ri39)<br>(B) Sadness (Kl68, We72c), tension (Ni82)<br>(C) Sadness (Cr85, Kr97, Pe98), disgust, anger (Sc77) |  |
| Pitch level | High | (A) V: graceful; IV: serene; VI: happy (He37), joy (Ri40a); III: dreamy (He37), sentimental (Gu35), pleading (Wa42); VII: triumph (Ri39), exciting (Wa42)<br>(B) Gaiety (Kl68, We72c)<br>(C) Surprise, potency, anger, fear, activity (Sc77) |  |
| | Low | (A) II: sad (He37, Wa42), melancholy (Gu35), lamentation (Ri40a); VIII: vigorous (He37); I: dignified (He37), serious (Wa42), solemn (Ri40a); VII: exciting (He37), agitation (Ri40a); IV: tranquil (Gu35)<br>(B) Serious (Kl68), sadness, solemnity (We72c)<br>(C) Boredom, pleasantness, sadness (Sc77) |  |

| Factor | Levels | Emotional expression | Adjective circle (Hevner) |
|---|---|---|---|
| Pitch variation | Large | (C) Happiness, pleasantness, activity, surprise (Sc77) | |
| | Small | (C) Disgust, anger, fear, boredom (Sc77) | |
| Rhythm | Regular/smooth | (A) VI: happiness (Wa42), glad (Gu35); I: serious, dignified, IV: peaceful (Wa42); VIII: majestic (Wa42); V: flippant (Gu35) |   |
| | Irregular/rough | (A) V: amusing (Wa42); VII: uneasy (Gu35) |   |
| | Complex | (C) Angry melodies (Th92) |   |
| | Varied | (C) Joyful melodies (Th92) | |
| | Firm | (A) I: dignified; VII: vigorous; II: sad; VII: exciting (He36) | |
| | | (C) Sad (We72c) | |
| | Flowing/fluent | (A) VI: happy; III: dreamy; V: graceful; IV: serene (He36) |   |
| | | (C) Gaiety (We72c) | |
| Tempo | Fast | (A) VII: exciting (He37, Wa42), uneasy (Gu35) agitation, triumph (Ri40b); VI: happy (He37, Wa42), glad (Gu35), gaiety, joy (Ri40b); V: graceful (He37), mischievous (Wa42), whimsical, flippant (Gu35); VIII: vigorous (He37) |   |
| | | (B) Happiness, pleasantness (Kl68, Mo68, We72c, Ga73, Ni82) | |
| | | (C) Activity, surprise, happiness, pleasantness, potency, fear, anger (Sc77), happiness, anger (Ju97), happiness (Th92, Kr97, Pe98), joy (Ba99), excitement (Th92) | |
| | Slow | (A) IV: serene (He37), tranquil (Gu35); III: dreamy (He37), longing (Ri40b), sentimental (Gu35); I: dignified (He37), serious (Wa42), dignified (Gu35), solemn (Ri40b); II: sad (He37, Wa42), lamentation (Ri40b); VIII: excited (Gu35) |   |
| | | (B) Sadness (Kl68, Be72, We72c, Ni82), solemnity (We72c) | |
| | | (C) Sadness, boredom, disgust (Sc77), sadness, tenderness (Ju97), sadness (Th92, Kr97, Pe98, Ba99), peace (Ba99) | |
| Timbre | Few harmonics | (C) Pleasantness, boredom, happiness, sadness (Sc77) | |
| | Many harmonics | (C) Potency, anger, disgust, fear, activity, surprise (Sc77) | |
| | Soft | (C) Tenderness, sadness (Ju97) | |
| | Sharp | (C) Anger (Ju97) | |
| Tonality | Tonal | (C) In joyful, dull, peaceful melodies (Th92) | |
| | Atonal | (C) In angry melodies (Th92) | |
| | Chromatic | (C) In sad and angry melodies (Th92) | |

| Factor | Levels | Emotional expression | Adjective circle (Hevner) |
|---|---|---|---|
| Musical form | Low complexity + average dynamism | (C) Relaxation/less tension (Ni83), joy, peace (Ba99)<br>(B) Positive emotions (Im79) | |
| | High complexity (melodic/harmonic/rhythmic) | (C) Tension (Ni83, Kr96), sadness (Ba99) | |
| | + low dynamism | (B) Melancholy and depression (Im79), | |
| | + high dynamism | (B) Anxiety and aggressiveness (Im79) | |
| | Repetition, condensation, sequential development, pauses etc. | (C) Increased tension (Ni83) | |
| | End of large-scale sections | (C) Tension peaks followed by rapid decrease (Kr96) | |
| | New musical ideas | (C) Low tension (Kr96) | |
| | Disruption of global form | (C) Little effect (Ko84, Go85, Ka92, Ko94, Ti96) | |

calmness/serenity, dignity/solemnity, sadness, tenderness, boredom, and disgust. Which of the different expressions is perceived in each case is highly dependent on the actual context, that is, presence and level of other factors. For instance, although faster tempo tends to increase perceived happiness, other factors may overrule this tendency, such as minor mode and descending seconds (e.g. Rigg 1940*b*; see also Section 10.1.3).

The term tempo may not always have the same meaning, as discusssed in Behne (1972), Gabrielsson (1986, 1988, 1999), and Motte-Haber (1968). Usually, tempo refers to perceived pulse rate. However, sometimes the pulse may be felt at, say, half or double pace. Perceived speed may also be affected by note density (e.g. the number of notes per second) as well as density of melodic or harmonic changes. Interplay between tempo (pulse rate) and note density may occur in an infinite number of combinations, for instance, rapid tone sequences (or high note density) but slow pulse (for examples, see Gabrielsson 1986).

### 10.2.2 Mode

Major mode may be associated with happiness/joy and minor mode with sadness, at least from 7–8 years of age. Major mode may also be associated with expressions such as graceful, serene, and solemn, and minor mode with expressions such as dreamy, dignified, tension, disgust, and anger.

What expression is perceived depends, again, on the context. For instance, loud chords and high-pitched chords may suggest more happiness than soft chords and low-pitched chords, irrespective of mode. Major mode is not a necessary condition for expression of happiness. An example may be the beginning of the last movement,

'Badinerie', in J. S. Bach's *Second Suite for Orchestra* (Fig 10.2). The theme in B minor with its forward-driving rhythmic pattern is usually played in rapid tempo and can hardly be associated with sadness—the title itself suggests happiness (French: *badinerie*, joke). However, if the same theme were transferred to B major, it may sound even happier (cf. Hevner 1935*b*, p. 202).

### 10.2.3 Loudness

Loud music may be associated with various expressions of intensity/power, tension, anger, and joy, and soft music with softness, tenderness, sadness, solemnity, and fear. Large variations of loudness (amplitude) may suggest fear, small variations happiness or activity. Rapid changes in loudness may be associated with expressions such as playful or pleading, and few or no changes with sadness, peace, and dignity (cf. Juslin, this volume).

### 10.2.4 Pitch

High pitch may be associated with expressions such as happy, graceful, serene, dreamy, and exciting, and, further, with surprise, potency, anger, fear, and activity. Low pitch may suggest sadness, dignity/solemnity, vigour, and excitement, as well as boredom and pleasantness (such a 'contradiction' may depend on the context of other factors). Large pitch variation may be associated with happiness, pleasantness, activity, or surprise; small pitch variation with disgust, anger, fear, or boredom.

### 10.2.5 Intervals

For harmonic intervals, results concerning consonance and dissonance are similar to the corresponding results under harmony (see Section 10.2.7) and concerning high and low pitch level to the corresponding results under pitch (see Section 10.2.4). For melodic intervals, results so far seem rather tentative, for instance, large intervals may be perceived as more powerful than small ones, the minor second as the most sad interval, and the octave as positive and strong.

### 10.2.6 Melody

*Melodic range*

Wide melodic range may be associated with joy, whimsicality, and uneasiness, and a narrow range with expressions such as sad, dignified, sentimental, tranquil, delicate, and triumphant.

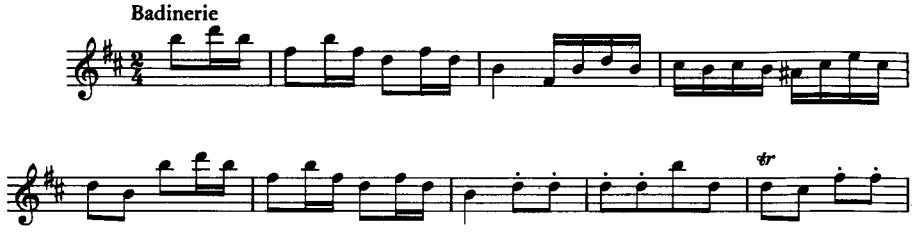



**Figure 10.2** Beginning of 'Badinerie', in J. S. Bach's *Second Suite for Orchestra.*

*Melodic direction (pitch contour)*

Ascending melody may be associated with dignity, serenity, tension, and happiness (but not for children, see Gerardi & Gerken 1995), as well as with fear, surprise, anger, and potency. Descending melody may be associated with expressions such as exciting, graceful, vigorous, and sadness (especially when combined with the minor mode; Gerardi & Gerken 1995) and, further, with boredom and pleasantness. According to Gundlach (1935), Hevner (1937), and Rigg (1939, 1964) melodic direction had no or little importance for emotional expression.

*Melodic motion*

There is little research on melodic motion except melodic direction. Stepwise motion may suggest dullness; intervallic leaps, excitement; and stepwise motion leading to melodic leaps, peacefulness. Lindström (1997) found that an interaction between melodic contour and rhythm affect judgments of happiness and sadness.

### 10.2.7 Harmony

Simple, consonant harmony may be associated with expressions such as happy/gay, relaxed, graceful, serene, dreamy, dignified, and majestic; complex and dissonant harmony with excitement, tension, vigour, anger, sadness, and unpleasantness.

### 10.2.8 Tonality

Composers used chromatic harmony in sad and angry melodies. Joyful, dull, and peaceful melodies were tonal, and angry melodies could be atonal (Thompson & Robitaille 1992).

### 10.2.9 Rhythm

Regular/smooth rhythm may be perceived as expressing happiness, dignity, majesty, and peace; irregular/rough rhythm, amusement, uneasiness, and anger; and varied rhythm, joy. Firm rhythm may be associated with expressions of sadness, dignity, and vigour, and flowing/fluent rhythm with expressions such as happy/gay, graceful, dreamy, and serene. Terminology varies among authors, which makes comparisons difficult.

### 10.2.10 Timbre

Tones with many harmonics may suggest potency, anger, disgust, fear, activity, or surprise, and tones with amplified higher harmonics may suggest anger. Tones with few, low harmonics may be associated with pleasantness, boredom, happiness, or sadness; tones with suppressed higher harmonics, tenderness and sadness. Behrens and Green (1993) found sadness best expressed by the singing voice or violin, anger by the timpani, and fear by the violin. With Hindustani music judged by Western listeners, Balkwill and Thompson (1999) found stringed instruments associated with the expression of anger, and the flute with the expression of peace.

### 10.2.11 Articulation

Staccato may be associated with gaiety, energy, activity, fear, and anger, and legato with sadness, tenderness, solemnity, and softness. Articulation is usually more discussed in connection with performance (Juslin, this volume).

### 10.2.12 Amplitude envelope

This refers to the type of attack and decay of tones. Sharp envelope (rapid attack and decay) may be associated with anger, happiness, surprise, and activity; round envelope with tenderness, sadness, fear, disgust, boredom, and potency. The shaping of amplitude envelope is usually more discussed in connection with performance (Juslin, this volume).

### 10.2.13 Musical form

The influence of various aspects of musical form has been little studied. The examples in Table 10.2 concerning the effects of high and low complexity, repetition, condensation of material, and so forth were discussed in Sections 10.1.1 (Imberty 1979; Krumhansl 1996; Nielsen 1983) and 10.1.3 (Balkwill & Thompson 1999). On the other hand, some studies (Konečni 1984; Tillman & Bigand 1996) indicate that various manipulations of musical form may not matter very much (see Section 10.1.3).

### 10.2.14 Interaction between factors

Perceived emotional expression in music is rarely or never exclusively determined by a single factor but is always a function of many factors (see, for example, Hevner 1935*b*, 1937; Rigg 1964). The influence of a certain factor may depend on how it is combined with other factors, that is, on the interaction between factors. A few examples mentioned earlier were mode × loudness and mode × pitch interactions (Section 10.1.2), and tempo × mode interactions (Sections 10.2.1 and 10.2.2). Possibilities for different types of interactions seem almost endless. Apparent contradictions in results, such that two 'opposite' emotions may be associated with the same structural factor may be resolved by examining interactions with other factors. However, interactions are primarily discussed *post hoc*, and there is practically no planned research regarding interaction between factors (Lindström 2000).

## 10.3 Some implications for future research

Despite the considerable amount of results summarized above there are still many gaps, uncertainties, and ambiguities regarding the influence of various structural factors on emotional expression. There is no systematic research on how the timbre of different musical instruments affects emotional expression—certainly something that composers pay much attention to. Research on harmony is very simplistic, usually focusing on consonance/dissonance, while there is practically nothing on chords, harmonic progressions, or implied harmony. There is surprisingly little and fairly tentative research on intervals. The study of melody is limited to the effects of range and direction, while melodic motion, that is, the specific sequences of pitches/intervals during the course of

a melody, is almost totally neglected, as is its rhythmic structure. Likewise, the influence of specific rhythm patterns—beyond their degree of simplicity or complexity—is unresearched. Questions concerning how musical form influences emotional expression are touched upon only in a few studies.

However, the most obvious gap concerns potential interactions between different factors. It has been emphasized, again and again (e.g. Hevner 1935*b*, 1937; Rigg 1964), that no factor works in isolation, its effects are dependent on what other factors and levels are present. Music abounds with interactions. To take but one example not mentioned earlier: the expressive qualities of different intervals can hardly be ascertained in a general manner. How a melodic interval is perceived is obviously dependent on in what tempo, in what direction (up, down), in what rhythm pattern, and at what loudness and pitch level the interval appears, not to mention its position within the musical context in question. Analogous examples may be easily found for other factors as well.

Moreover, most factors are studied only regarding two 'extreme' levels—fast or slow tempo, high or low pitch, loud or soft sound, ascending or descending melody, legato or staccato articulation. Intermediate levels are usually neglected, probably on the implicit assumption that their effects may be inferred from results at the extreme levels. Nor are the effects of transitions from one level to another investigated, such as crescendo or diminuendo, accelerando or ritardando, glissando, transition from minor to major mode, from diatonicism to chromaticism, or from thin to thick texture—not to mention combinations of them, for instance, imagine a glissando towards high pitch combined with diminuendo and ritardando.

The last-mentioned examples hint at another neglected aspect: how expressive qualities vary and change in dynamic ways. Most research has focused on one expression at a time: happiness, sadness, anger, etc., and it is recommended that the investigator choose musical excerpts which are homogeneous with regard to emotional expression (e.g. Hevner 1935*b*). However, many of music's most expressive qualities relate to structural *changes* across time which make possible expression of, say, shifting emotions, blending emotions, competing or conflicting emotions.

Going beyond emotions, Langer suggested that, because there are certain formal properties that are similar in music and in human feeling, we may perceive expressions such as 'growth and attenuation, flowing and stowing, conflict and resolution, speed, arrest, terrific excitement, calm, or subtle activation and dreamy lapses' (Langer 1953, p. 27), or 'patterns of motion and rest, of tension and release, of agreement and disagreement, preparation, fulfilment, excitation, sudden change, etc.' (Langer 1957, p. 228). These examples represent complex combinations of perceptual, cognitive, and emotional aspects that elude conventional psychological terminology. They remind us of what Stern (1985) called 'vitality affects', which seem intuitively appealing to the study of musical expression (see also Sloboda & Juslin, this volume). Phenomena like these are little investigated in research; the closest examples are studies on tension and release by means of continuous recording (see 10.1.1). Continuous recording of listeners' judgments is a necessary prerequisite to study these phenomena (for a review, see Schubert, this volume).

Research is also hampered by a lack of appropriate definitions and terminology, both regarding structural factors (e.g. on tempo and rhythm; see 10.2.1 and 10.2.9), and regarding categories or dimensions of emotions (see Sloboda & Juslin, this volume).

Should we try to subsume the hundreds of descriptive words for emotional expression under a few dimensions, such as evaluation, activity, and potency (Osgood *et al.* 1957), or valence and arousal in Russell's (1980) circumplex model? (See Remington *et al.* 2000, for a recent review of the circumplex model.) It seems obvious that this reduction goes too far and cannot capture all possible emotional nuances that we may perceive in music. On the other hand, we may have penetrating phenomenological descriptions abounding with suggestive descriptive terms to the limit of overwhelming the reader. In between, there may be something like Hevner's adjective clusters, although criticized and revised (e.g. Farnsworth 1954), which are still useful and referred to. However, emotions such as anger, fear, surprise, boredom, and disgust were missing and were included only in much later studies (e.g. Scherer & Oshinsky 1977).

In order to ensure both adequate control and ecological validity, future research should aim for an efficient interplay between studies using real music and studies using systematic manipulation of structural factors in as music-like contexts as possible. Many gaps, confusions, and contradictions in the present results may be resolved, slowly but safely, by careful research on the interactions between different factors. Techniques for continuous recording, maybe combined with non-verbal responses, should be used in order to better study perceived expression as a dynamic process. Choice of musical examples should be considerably broadened to overcome the present limitation to Western art (and some popular) music. Connections should be sought to studies of emotional expression in other areas, say, speech and body language, as well as to emotion psychology in general, and, furthermore, to music theory, music aesthetics, and music anthropology.

The real experts are, of course, the successful composers, past and present. They know, somehow, how to manipulate different variables in musical structure to attain an intended expression. This knowledge may be in part explicit, in part implicit, in combinations that are different for different composers. Although we may learn some grains of wisdom from the writings of certain composers—for instance, in the volume edited by Fisk (1997)—on the whole, their knowledge is rarely made available to others except for their students and other close persons. Thus, the mystery of composing music may remain—to the satisfaction of those who want to have some mystery in life left untouched.[1]

## References

**Balkwill, L. L. & Thompson, W. F.** (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception,* 17, 43–64.

**Behne, K. E.** (1972). *Der Einfluss des Tempos auf die Beurteilung von Musik.* Köln, Germany: Arno Volk Verlag.

**Behrens, G. A. & Green, S. B.** (1993). The ability to identify emotional content of solo improvisations performed vocally and on three different instruments. *Psychology of Music,* 21, 20–33.

[1] The writing of this chapter was supported by a grant from The Bank of Sweden Tercentenary Foundation.

**Cazden, N.** (1979). Can verbal meanings inhere in fragments of melody? *Psychology of Music,* 7, 34–8.

**Clynes, M.** (1977). *Sentics: The touch of emotions.* New York: Anchor Press/Doubleday.

**Clynes, M. & Walker, J.** (1982). Neurobiologic functions of rhythm, time, and pulse in music. In *Music, mind, and brain. The neuropsychology of music,* (ed. M. Clynes), pp 171–216. New York: Plenum Press.

**Cooke, D.** (1959). *The language of music.* London: Oxford University Press.

**Costa, M., Bitti, P. E. R., & Bonfiglioli, L.** (2000). Psychological connotations of harmonic musical intervals. *Psychology of Music,* **28,** 4–22.

**Crowder, R. G.** (1984). Perception of the major/minor distinction: I. Historical and theoretical foundations. *Psychomusicology,* **4,** 3–12.

**Crowder, R. G.** (1985). Perception of the major/minor distinction: III. Hedonic, musical, and affective discriminations. *Bulletin of the Psychonomic Society,* 23, 314–16.

**Davies, S.** (1994). *Musical meaning and expression.* Ithaca, NY: Cornell University Press.

**Downey, J. E.** (1897). A musical experiment. *American Journal of Psychology,* **9,** 63–9.

**Farnsworth, P. R.** (1954). A study of the Hevner adjective list. *Journal of Aesthetics and Art Criticism,* 13, 97–103.

**Fisk, J.** (ed.) (1997). *Composers on music.* Boston: Northeastern University Press.

**Fredrickson, W. E.** (1997). Elementary, middle, and high school student perceptions of tension in music. *Journal of Research in Music Education,* **45,** 626–35.

**Gabriel, C.** (1978). An experimental study of Deryck Cooke's theory of music and meaning. *Psychology of Music,* 6, 13–20.

**Gabriel, C.** (1979). A note on comments by Nettheim and Cazden. *Psychology of Music,* 7, 39–40.

**Gabrielsson, A.** (1973). Adjective ratings and dimension analysis of auditory rhythm patterns. *Scandinavian Journal of Psychology,* 14, 244–60.

**Gabrielsson, A.** (1986). Rhythm in music. In *Rhythm in psychological, linguistic and musical processes,* (ed. J. R. Evans & M. Clynes), pp. 131–67. Springfield, IL: Charles C. Thomas.

**Gabrielsson, A.** (1988). Timing in music performance and its relations to music experience. In*Generative processes in music: The psychology of performance, improvisation, and composition,* (ed. J. A. Sloboda), pp. 27–51. Oxford, UK: Clarendon Press.

**Gabrielsson, A.** (1995). Expressive intention and performance. In *Music and the mind machine,* (ed. R. Steinberg), pp. 35–47. Heidelberg: Springer.

**Gabrielsson, A.** (1999). The performance of music. In *The psychology of music,* (2nd edn) (ed. D. Deutsch), pp. 501–602. San Diego, CA: Academic Press.

**Gabrielsson, A. & Juslin, P. N.** (in press). Emotional expression in music. In *Handbook of affective sciences,* (ed. R. J. Davidson, H. H. Goldsmith, & K. R. Scherer). New York: Oxford University Press.

**Gabrielsson, A. & Lindström, E.** (1995). Emotional expression in synthesizer and sentograph performance. *Psychomusicology,* 14, 94–116.

**Gerardi, G. M. & Gerken, L.** (1995). The development of affective responses to modality and melodic contour. *Music Perception,* 12, 279–90.

**Gilman, B. I.** (1891). Report on an experimental test of musical expressiveness. *American Journal of Psychology,* 4, 558–76.

**Gilman, B. I.** (1892). Report of an experimental test of musical expressiveness (continued). *American Journal of Psychology,* 5, 42–73.

**Gotlieb, H. & Konečni, V. J.** (1985). The effects of instrumentation, playing style, and structure in the Goldberg Variations by Johann Sebastian Bach. *Music Perception*, **3**, 87–102.

**Gregory, A. H., Worrall, L., & Sarge, A.** (1996). The development of emotional responses to music in young children. *Motivation and Emotion*, **20**, 341–8.

**Gundlach, R. H.** (1935). Factors determining the characterization of musical phrases. *American Journal of Psychology*, 47, 624–44.

**Heinlein, C. P.** (1928). The affective character of the major and minor modes in music. *Journal of Comparative Psychology*, **8**, 101–42.

**Hevner, K.** (1935*a*). The affective character of the major and minor modes in music. *American Journal of Psychology*, 47, 103–18.

**Hevner, K.** (1935*b*). Expression in music: A discussion of experimental studies and theories. *Psychological Review*, 47, 186–204.

**Hevner, K.** (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology*, **48**, 246–68.

**Hevner, K.** (1937). The affective value of pitch and tempo in music. *American Journal of Psychology*, **49**, 621–30.

**Hill, D. S., Kamenetsky, S. B., & Trehub, S. E.** (1996). Relations among text, mode, and medium: Historical and empirical perspectives. *Music Perception*, **14**, 3–21.

**Huber, K.** (1923). *Der Ausdruck musikalischer Elementarmotive.* Leipzig, Germany: Johann Ambrosius Barth.

**Imberty, M.** (1979). *Entendre la musique.* Paris: Dunod.

**Juslin, P. N.** (1997). Perceived emotional expression in synthesized performances of a short melody: Capturing the listener's judgement policy. *Musicae Scientiae*, 1, 225–56.

**Kamenetsky, S. B., Hill, D. S., & Trehub, S. E.** (1997). Effect of tempo and dynamics on the perception of emotion in music. *Psychology of Music*, 25, 149–60.

**Karno, M. & Konečni, V. J.** (1992). The effects of structural interventions in the first movement of Mozart's Symphony in G Minor K. 550 on aesthetic preference. *Music Perception*, **10**, 63–72.

**Kastner, M. P. & Crowder, R. G.** (1990). Perception of the major/minor distinction: IV. Emotional connotations in young children. *Music Perception*, **8**, 189–202.

**Kleinen, G.** (1968). *Experimentelle Studien zum musikalischen Ausdruck.* Hamburg, Germany: Universität Hamburg.

**Konečni, V. J.** (1984). Elusive effects of artists' 'messages'. In *Cognitive processes in the perception of art*, (ed. W. R. Crozier & A. J. Chapman), pp. 71–93. Amsterdam: North-Holland.

**Konečni, V. J. & Karno, M.** (1994). Empirical investigations of the hedonic and emotional effects of musical structure. *Musikpsychologie. Jahrbuch der Deutschen Gesellschaft für Musikpsychologie*, 11, 119–37.

**Krumhansl, C. L.** (1996). A perceptual analysis of Mozart's Piano Sonata K. 282: Segmentation, tension, and musical ideas. *Music Perception*, **13**, 401–32.

**Krumhansl, C. L.** (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology*, 51, 336–52.

**Langer, S. K.** (1953). *Feeling and form.* London: Routledge.

**Langer, S. K.** (1957). *Philosophy in a new key*, (3rd edn). Cambridge, MA: Harvard University Press.

**Levinson, J.** (1997). *Music in the moment.* Ithaca, NY: Cornell University Press.

Lindström, E. (1997). Impact of melodic structure on emotional expression. In *Proceedings of the Third Triennial ESCOM Conference, Uppsala, June 1997*, (ed. A.Gabrielsson), pp. 292–7. Uppsala, Sweden: Uppsala University.

Lindström, E. (2000). *Impact of melodic structure on emotional expression in music* Manuscript submitted for publication.

Madsen, C. K. (1997). Emotional responses to music. *Psychomusicology*, **16**, 59–67.

Madsen, C. K. & Fredrickson, W. E. (1993). The experience of musical tension: A replication of Nielsen's research using the continuous response digital interface. *Journal of Music Therapy*, **30**, 46–63.

Maher, T. F. (1980). A rigorous test of the proposition that musical intervals have different psychological effects. *American Journal of Psychology*, 93, 309–27.

Maher, T. F. & Berlyne, D. E. (1982). Verbal and exploratory responses to melodic musical intervals. *Psychology of Music*, **10**, 11–27.

Motte-Haber, H. de la (1968). *Ein Beitrag zur Klassifikation musikalischer Rhythmen.* Köln, Germany: Arno Volk Verlag.

Namba, S., Kuwano, S., Hatoh, T., & Kato, M. (1991). Assessment of musical performance by using the method of continuous judgment by selected description. *Music Perception*, **8**, 251–76.

Nettheim, N. (1979). Comment on a paper by Gabriel on Cooke's theory. *Psychology of Music*, 7, 32–3.

Nielsen, F. V. (1983). *Oplevelse av musikalsk spœnding [Experience of musical tension].* Copenhagen: Akademisk Forlag (includes summary in English).

Nielsen, F. V. (1987). Musical 'tension' and related concepts. In *The semiotic web '86. An international yearbook*, (ed. T. A. Sebeok & J. Umiker-Sebeok), pp. 491–513. Berlin: Mouton de Gruyter.

Nielzén, S. & Cesarec, Z. (1981). On the perception of emotional meaning in music. *Psychology of Music*, **9**, 17–31.

Nielzén, S. & Cesarec, Z. (1982). Emotional experience of music as a function of musical structure. *Psychology of Music*, **10**, 7–17.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning.* Urbana, IL: University of Illinois Press.

Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, **68**, 111–41.

Remington, N. A., Fabrigar, L. R., & Visser, P. S. (2000). Reexamining the circumplex model of affect. *Journal of Personality and Social Psychology*, **79**, 286–300.

Rigg, M. G. (1937). Musical expression: An investigation of the theories of Erich Sorantin. *Journal of Experimental Psychology*, 21, 442–55.

Rigg, M. G. (1939). *What features of a musical phrase have emotional suggestiveness?* Publications of the Social Science Research Council of the Oklahoma Agricultural and Mechanical College, No. 1.

Rigg, M. G. (1940*a*). The effect of register and tonality upon musical mood. *Journal of Musicology*, 2, 49–61.

Rigg, M. G. (1940*b*). Speed as a determiner of musical mood. *Journal of Experimental Psychology*, 27, 566–71.

Rigg, M. G. (1964). The mood effects of music: A comparison of data from four investigators. *Journal of Psychology*, **58**, 427–38.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161–78.

**Scherer, K. R. & Oshinsky, J. S.** (1977). Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion*, **1**, 331–46.

**Schlosberg, H.** (1941). A scale for the judgement of facial expressions. *Journal of Experimental Psychology*, **29**, 497–510.

**Schubert, E.** (1999). *Measurement and time series analysis of emotion in music.* Unpublished doctoral dissertation, University of South Wales, Sydney, Australia.

**Sloboda, J. A.** (1985). *The musical mind. The cognitive psychology of music.* Oxford, UK: Clarendon Press.

**Smith, L. D. & Williams, R. N.** (1999). Children's artistic responses to musical intervals. *American Journal of Psychology*, **112**, 383–410.

**Stern, D. N.** (1985). *The interpersonal world of the infant. A view from psychoanalysis and developmental psychology.* New York: Basic Books.

**Thompson, W. F. & Robitaille, B.** (1992). Can composers express emotions through music? *Empirical Studies of the Arts*, **10**, 79–89.

**Tillman, B. & Bigand, E.** (1996). Does formal musical structure affect perception of musical expressiveness? *Psychology of Music*, **24**, 1–17.

**de Vries, B.** (1991). Assessment of the affective response to music with Clynes's sentograph. *Psychology of Music*, 19, 46–64.

**Watson, K. B.** (1942). The nature and measurement of musical meanings. *Psychological Monographs*, **54**, 1–43.

**Wedin, L.** (1969). Dimension analysis of emotional expression in music. *Swedish Journal of Musicology*, **51**, 119–40.

**Wedin, L.** (1972*a*). *Evaluation of a three-dimensional model of emotional expression in music.* Reports from the Psychological Laboratories, University of Stockholm, No. 349.

**Wedin, L.** (1972*b*). Multidimensional scaling of emotional expression in music. *Swedish Journal of Musicology*, **54**, 1–17.

**Wedin, L.** (1972*c*). Multidimensional study of perceptual-emotional qualities in music. *Scandinavian Journal of Psychology*, **13**, 241–57.

CHAPTER 11

# MUSIC AS A SOURCE OF EMOTION IN FILM

ANNABEL J. COHEN

Emotion characterizes the experience of film, as it does the experience of music. Because music almost always accompanies film, we may well ask what contribution music makes to the emotional aspects of film. The present chapter addresses this question.

It should be said at the outset that in spite of the integral role of music for film, film music has been largely neglected by the disciplines of both musicology and music psychology until the last decade (e.g. Cohen 1994; Marks 1998; Prendergast 1991). The reasons for the neglect are complex, arising from social, technological, economic, historical, and cultural factors. Some of these factors also account for a parallel neglect by psychology of the study of film perception (Hochberg & Brooks 1996*a*, 1996*b*). Moreover, unlike other types of popular or art music, much music for film has been composed with the understanding that it will not be consciously attended to. Countering this neglect, the present chapter takes a psychological perspective on the sublime and remarkable emotional phenomena produced by music in the context of film. This chapter has the joint intent of supporting the argument that music is one of the strongest sources of emotion in film and of opening doors to further empirical work that explains why this is so.

The chapter is divided into five sections. The first section briefly establishes a context for discussing emotion in music and film. Section 11.2 focuses on music in film, first establishing a historical perspective and then examining the role of music at the interface of the fictional and non-fictional elements of film. It continues with empirical studies of music as a source of inference and then summarizes the functions that music serves for film. Section 11.3 presents a cognitive framework for understanding musical soundtrack phenomena previously described. Section 11.4 considers the role of the composer as the origin of the source of musical emotion for film, whilst conclusions are drawn in the final section.

## 11.1 Emotion: definitions in music and film contexts

In the present chapter, the term film refers to the narrative dramas characteristic of movie theatres, television, and video with which most people are familiar as a source of entertainment. Music typically accompanies a considerable proportion of the duration of such films. Because of the relative novelty of the empirical study of film music in general, let alone the study of the emotional contribution of music in film, it would be

premature to advocate a particular way of considering emotion in the present chapter. What is more important is to show how various 'music-alone' perspectives on emotion translate in the film context. These perspectives include the contribution of music to emotional meaning (Juslin 1997; Levi 1982), the establishment of general mood (Pignatiello *et al.* 1986), and the experience of genuine, deep emotions (Gabrielsson 1998; Gabrielsson & Lindström 1993; Sloboda 1985, 1992; see also Gabrielsson, this volume).

The film context sometimes permits greater terminological clarity than the music-alone situation. For example, consider the terms 'mood' and 'emotion' which are often differentiated with respect to the presence of an object (e.g. Barrett & Russell 1999; Tan 1996). Whereas both moods and emotions may be regarded as dispositions toward appraising emotional meaning structures and a readiness to respond in a certain manner, moods do not have objects; emotions do. For example, experiencing the emotion of relief requires an object of that emotion, such as a safe arrival after a treacherous journey. Objects are not as evident in music-alone contexts, but, as argued by musicologist Nicholas Cook (1998), in a multimedia context music readily finds an object. The emotional associations generated by music attach themselves automatically to the visual focus of attention or the implied topic of the narrative. Because film content provides the object of emotion generated by music, the film helps to control the definition of the object of the emotion experienced during the presence of music.

Considering music and emotion within the context of film also has the advantage of bringing knowledge from psychological studies of film to bear on questions regarding music and emotion. Based on the emotion theorist Frijda (1986), Tan (1996), for example, addresses the question of the genuine nature of emotions in film, a topic that will be addressed later in this chapter. Thus, research on music and emotion in the film context may benefit from research insights derived from studies of film and emotion. Conversely, our understanding of emotion associated with autonomous music may shed light on emotional processes that occur in the film context. All of this information may contribute to the understanding of both the unique accomplishment of composing music for film and the extent to which music provides an important source of emotion in film.

## 11.2 Music and cinema

### 11.2.1 Historical background

*Beginnings*

From the earliest days of film, music played a part. When silent film was first introduced at the turn of the century, the film projector was anything but silent. Music was therefore enlisted to mask the extraneous noise. While serving the masking function, music also was exploited to illustrate and explain the action (Palmer 1980, p. 549). Kracauer (1960, p. 133) emphasizes that the noise problem of the film projection was relatively shortlived, and yet the importance of music remained. An entire music-for-the-silent-film industry developed to support this function of music (Limbacher 1974; Thomas

1997, pp. 37–40). It included the publication of anthologies of music to represent various emotional settings, an increased demand for pianos in the thousands by small movie theatres that sprang up, and architectural plans for movie theatres that included places for pianists and sometimes other musicians.

*Hugo Münsterberg*

The first psychologist to direct attention to the new phenomena of film was Hugo Münsterberg at Harvard University (Fig. 11.1). Between 1899 and 1916, he wrote 24 books, one of the last of which was *The photoplay: A psychological study.* In what is regarded as the first book on film theory (Anderson 1997), Münsterberg's views are enlightening. His experience of film was as fresh as a child's although acquired as a highly intelligent adult. His understanding of both introspection and scientific method encourage our confidence in his record of, and insight into, film at that time.

> Yes, it is a new art—and this is why it has such fascination for the psychologist who in a world of ready-made arts, each with a history of many centuries, suddenly finds a new form still undeveloped and hardly understood. For the first time the psychologist can observe the starting of an entirely new esthetic development, a new form of true beauty in the turmoil of a technical age, created by its very technique and yet more than any other art destined to overcome outer nature by the free and joyful play of the mind (Münsterberg 1970, pp. 232–3).



**Figure 11.1** Hugo Münsterberg.

He did not live to experience the talking film, but his film experience was not lacking in sound. There were sound effects—he describes a machine, the allefex, 'which can produce over fifty distinctive noises, fit for any photoplay emergency' (Münsterberg 1970, p. 205)—and there was music. In his view, music relieved tension, maintained interest ('keeps the attention awake'), provided comfort, reinforced emotion, and contributed to the aesthetic experience (pp. 204–5).

Münsterberg also used musical metaphor in describing the film experience (e.g. pp. 120 and 128–9). For example, he recounts a narrative cliché of the period, a rapid alternation between three scenes: a jovial boss and his secretary enjoying a private after-hours party in the office, the dismal parents of the secretary awaiting their daughter's return, and the lonely wife awaiting her husband's attention. 'It is as if we saw one through another, as if three tones blended into one chord . . . The photoplay alone gives us our chance for such omnipresence' (p. 105). Yet, to extend his metaphor of the musical chord, it is also music that can represent in rapid succession and perhaps simultaneously the 'emotional polyphony' of these multiple messages.

Possibly because he never experienced the sound film ('talkies'), Münsterberg directed attention to the importance of music within the film and to music as a means of understanding the psychological processes underlying film. He suggested that cinema is more similar to music than to photography and drama, which on the surface are arts that bear a more striking resemblance:

> . . . we come nearer to the understanding of its [film's] true position in the esthetic world, if we think at the same time of . . . the art of the musical tones. They have overcome the outer world and social world entirely, they unfold our inner life, our mental play, with its feelings and emotions, its memories and fancies, in a material which seems exempt from the laws of the world of substance and material, tones which are fluttering and fleeting like our own mental states. (Münsterberg 1970, p. 168–9)

Münsterberg's untimely death (the year of his publication of *The photoplay*) and the coincident advent of behaviourism, focusing as it did on only objectively observable behaviour, may account for the failure of psychological research in film and music to run in parallel with the technological developments associated with these media. Instead, technology developed and its psychological study lagged far behind in spite of having got off to a good start.

### *The sound film*

In 1927, approximately 10 years after the death of Münsterberg, *The Jazz Singer* signaled the advent of the 'talkies' and the demise of the film-music industry. With real voices and sound effects, music would no longer be needed to establish mood and emotional context . . . or would it? To the surprise of many, something was missing without music (Kracauer 1960, p. 138). The screen had lost part of its vitality. As Kalinak (1992, p. 45) says 'when the possibility of synchronized speech and sound effects released sound film from its reliance upon continuous musical accompaniment, it initially rejected music entirely. But the life span of the all-talking picture was brief, the need that music filled quickly reasserting itself'.

Several theorists have commented that music adds a third dimension to the two-dimensional film screen (Palmer 1990; Rosar 1994). Composers also shared this view.

Aaron Copland (1941) stated 'the screen is a pretty cold proposition'. Film composer David Raksin (in Brown 1994, p. 282) referred to Nietzsche's idea 'without music, life would not be worth living'. His statement is extreme (deaf people live worthwhile lives) but paraphrasing the maxim, few hearing people would deny that music contributes to their experience of film.

Since the early days of film, directors have recognized the contribution that film-editing made to the viewer, often referred to as montage. Viewers are typically unaware of the rapid changes in camera angle, the move from close-up to long-shot or from one part of the scene to another and back again. None the less, viewers make sense of the world depicted by these juxtaposed shots. Theories of montage concern the audience's synthesis of juxtaposed information in the film. With the advent of the sound film, Russian director Sergej Eisenstein was among the first to extend the notion of visual montage to sound, and suggested that the listener incorporates the same synthetic process in making sense of the entire audiovisual cinematic presentation.

### 11.2.2 Music and the diegesis and the non-diegesis

Film theory commonly refers to the fictional, imagined, narrative world of the film as the *diegesis*. In contrast, the *non-diegesis* refers to the objective world of the audience, the world of artefact, of film screens, projectors, proficiency of actors, and technical aspects of the film. In terms of physical reality, music as acoustic vibrations belongs to the non-diegesis. Logically—unless such sound were part of a scene portrayed in a film, as in a film about a musical instrument or the life of a great composer—these sounds of music should *detract from* rather than *add to* the sense of reality of the film. This point was well made in Mel Brooks' comedy *Blazing Saddles* (1974). A sheriff rides out on the desert—with seemingly appropriate music in the background—and meets face to face with the Count Basie Band performing the now inappropriate music *Paris in the Spring*. The fictional (diegetic) and the non-fictional (non-diegetic) realities collide and add to the humour of the scene. It is probably not coincidental that Brooks, the director and screen writer of the film, is also a composer of music, including some film scores, and so he would have been particularly sensitive to this film-score convention.

Thus, the audience selectively attends to only the part of the music that makes sense with the narrative. Selective attention is a common perceptual-cognitive operation. The recently discovered phenomena of 'inattentional blindness' is another example of it in the visual domain of film. Here it has been shown that people rarely noticed or were distracted by impossible visual aspects represented in either a film or in their real-world experience. For example, viewers did not notice that a woman in a film clip began the short scene with a scarf and ended the scene without it (Levin & Simons, 2000). In another study (Levin & Simon, 1998), an experimenter positioned on a college campus solicited directions from unsuspecting subjects. Their conversation was interrupted by two confederates carrying a door. One of the confederates changed position with the experimenter who had initially asked for directions. The conversation about directions then continued. The subject rarely realizes that there were two different people to whom he or she has been conversing. Two facts are important here. First, the visual system is blind to much available information, and this inattentional blindness (cf. Mack & Rock 1998) is equally characteristic of vision in the real and in the film world. Thus, the fact

that audiences extract the emotional information in music and fail to attend to the acoustical aspects might be described as a case of inattentional deafness, a byproduct of the fact that awareness depends on attention, and attentional capacity is limited. A better parallel can be drawn to the role of prosody in speech perception, which shares with music a greater syntactic similarity than the visual examples given. Here patterns of intensity and frequency from intonation systematically provide emotional meaning to a listener, yet the listener focuses on the meaning and is unconscious of this source of information (e.g. Banse & Scherer 1996; Murray & Arnott 1993).

A good example of the role of music as a source of emotion in the narrative is illustrated in the film *Witness*. Here, a young Amish boy is the sole observer of a violent murder in a train station. He is directed to a noisy police station in order to search through a book of photographs of suspects for a match of his memory to the actual perpetrator of the crime. While left unattended momentarily, he wanders toward a display cabinet that holds a photograph of an honoured senior officer in the force. As he views the photograph, Maurice Jarre's music replaces the background sound effects of the police station. The audience realizes that the boy becomes awestruck with the sudden recognition that the photograph of the police officer depicts the person who he saw commit the crime. The audience takes interest in this scene and is concerned now about the implications for the safety of the boy. This is one of the most critical points in the film; without it there would be no further plot. The criminal within the police force would remain unknown and the boy could continue on his trip without a care. Yet it is this crucial moment, a moment that must be comprehended by the audience, to which the unrealistic music is not merely added but is added at the expense of the more realistic, diegetic sound effects. But the audience, caught up in the drama, is unlikely to have noted this departure from reality.

Film-music scholar Claudia Gorbman (1987) has addressed the unconscious perception of music in her book *Unheard melodies: Narrative film music*. Gorbman's perspective is well captured by Jeff Smith (1996, p. 234): 'By veiling the lacks and deficiencies of other discursive structures, film music, according to Gorbman, lubricates the various cogs and pistons of the cinematic pleasure machine'. If anything, departures from reality via music makes an episode 'more real', more vivid, more emotionally relevant. To date, experimental research has not focused on the subtle uses of music in film (such as its replacement of realistic sound as in the *Witness* example). A number of studies, however, have concerned the role of music in generating inferences, and often those inferences are associated with emotional meaning.

### 11.2.3 Music and inference: empirical studies

Music presumably adds to the diegetic realism while providing non-diegetic, acoustical, information that is completely incompatible with that realism. At the interface of the diegetic and non-diegetic worlds, the use of music in this and in most other film situations is paradoxical (Cohen 1990). To escape the paradox, the analysis of the acoustical information must be regarded as a preattentive step that leads the listener to inferences consistent within the diegetic world of the film. From moment to moment, the audience member extracts information from non-diegetic sources to generate the emotional information he or she needs to make a coherent story in the diegesis. The successful

director and film-score composer provide just the right cues to guide the attentional and inferences processes.

One attempt at understanding the phenemenon of diegetic inference comes from the context of psychological situation models described as 'vicarious experiences in narrative comprehension' (Zwaan 1999, p. 15). Zwaan, for example, has focused on how literature enables readers to 'mentally leap into imagined worlds'. The information provided by the text is sufficient to enable a reader to place himself or herself at a spatial, temporal, and psychological vantage point from which events are vicariously experienced. The perspective is termed a deictic centre.

Magliano *et al.* (1996) extended the approach to the study of film. In their study, music was indirectly examined as one of six operationally defined film factors (such as montage) that might contribute to the psychological definition of the situation. In two experiments, they investigated visual, auditory, and discourse conditions that enable viewers to predict future events while viewing a James Bond movie, *Moonraker*. In experiment 1, participants were instructed to generate predictions while watching the movie. In experiment 2, participants provided think-aloud protocols at different locations in the film. In both experiments, the presence of supporting visual and discourse information led to systematic predictions by the participant. Music significantly cooccurred with other cinematic sources of support such as montage and *mise en scène*, which were found to influence inference processes.[1]

The study by Magliono *et al.* (1996) was not specifically designed to show the influence of music in the film context. Studies that have been so designed have been successful in showing this influence. Such studies typically require more conditions than do comparable studies in music-alone situations, because it is necessary to determine the effects of music alone, film alone, music judged in the context of film, and film judged in the context of music. The studies to be reported have often involved several, if not all, of these different comparison conditions.

In a study by Bullerjahn and Güldenring (1994), professional composers of film music (including Peer Raben, composer for all the Fassbinder films) created a total of five different backgrounds (e.g. crime, melodrama) for the same 10-minute film segment. Both quantitative and qualitative analysis showed that the different soundtracks led to different judgments of the appropriateness of emotional categories (e.g. sad, thrilling, sentimental, vivid), choice of genre (horror, comedy, thriller, crime, etc.), reasons for the actions of the protagonist, and expectations about the completion of the film. In some cases, these judgments and inferences could be attributed to specific aspects of the film. For example, the authors suggested that the final closure of the major chord of one of the melodramatic soundtracks accounted for the presumed reconciliation between the film characters though it is preceded by an argument.

In a series of experiments, Thompson *et al.* (1994) specifically examined the effects of musical closure on perceived closure of a film. In their first experiment, a closed soundtrack ended with a traditional 'perfect' cadence (dominant chord to tonic chord ending).

[1] When music was entered first into the regression equation, it too was found to be a significant predictor of inference, though accounting for only 6 per cent of the variance (J. P. Magliano, personal communication).

The unclosed soundtrack differed only with respect to the final bar, such that the ending was not on the tonic chord. Subjects viewed a short animation accompanied by one of the soundtracks and were asked to rate the degree of closure represented by the clip. Judgments of closure were significantly higher for the condition in which the closed ending was presented. In a second experiment, a professional composer produced closed and unclosed soundtracks to accompany a short film clip produced by one of the experimenters. Closed soundtracks ended on the tonic chord, in contrast to the unclosed soundtracks. The effect of soundtrack closure was strong for only one of the film clips, suggesting that visual factors may take precedence over musical structure in some cases. In a final study, 12 separate soundtracks were composed for 12 clips from a commercial film, initially chosen for their assumed range in degree of closure. The soundtracks were also composed to represent a range of closure, and the degree of closure need not match that of the clip. Participants in the experiment judged the degree of closure of the soundtracks, the clips alone, and the soundtrack and clips together. The influence of independently judged visual and musical closure on judged closure of the film was shown through regression analysis, with a slightly greater contribution arising from visual than musical information. In addition to demonstrating a robust direct effect of musical structure on the feeling of closure of a film narrative, the authors also reported that the role of music was almost completely implicit. When participants were asked for the basis of their judgments, they almost always attributed their judgments to the visual information.

Boltz *et al.* (1991) examined the role of music on inferences in a study that compared music that foreshadowed an outcome versus music that accompanied an outcome. Participants viewed 20 different 3–4-minute clips from feature films and television dramas. Excerpts were selected that ultimately resolved in a happy (positive) or sad (e.g. tragic) way. Music emotionally consistent or inconsistent with these endings (as determined by music-alone ratings) either preceded (foreshadowed) or accompanied the video excerpt. Thus, in some cases the foreshadowing music correctly predicted the mood of the following video event and the accompanying music was congruent with the mood of the video event, and in other cases the foreshadowing music incorrectly predicted the subsequent event and the accompanying music was incongruent with the event. Music accompanying an episode's outcome led to higher recall when the mood of the music and scene were congruent with each other. Conversely, mood-incongruent relations significantly lowered performance to a level comparable to that of the control condition in which no music had occurred. Foreshadowing, however, revealed the opposite pattern. Here, expectancy violations arising from mood-incongruent relations were significantly more memorable than were mood-congruent episodes in which viewer's expectancies were confirmed (Boltz *et al.* pp. 597–8). Boltz *et al.* (1991, p. 602) concluded that their results supported the notion that viewers rely on the emotional expression of music to either generate expectancies about future scenarios or to direct attending toward corresponding aspects of visual activities.

Marshall and Cohen (1988) also observed the ability of music to alter the interpretation of a simple visual presentation. They studied the effects of two different soundtracks on impressions about three geometric forms, a large and small triangle and a circle, in a short animation developed by Heider and Simmel (1944). In their experiment, subjects

viewed the 2-minute animation with one of two soundtracks or with no soundtrack (control condition). They then provided 12 ratings for the film overall and for the three figures. Other groups of subjects rated the music on these same scales. Each of the 12 scales represented a bipolar adjective pair (e.g. fast–slow, nice–awful) and specifically tapped one of the three dimensions of emotional meaning—activity, evaluation, and potency—comprising the semantic differential (Osgood *et al.* 1957). The activity and evaluation dimensions are understood to represent the motivation (arousal) and appraisal (valence) dimensions associated with two-dimensional theories of emotion on which many emotion theorists (Barrett & Russell 1999; Lang 1995; Storm & Storm 1987; see also Sloboda & Juslin, this volume) and music psychologists (Gregory 1998; Madsen 1997; Schubert 1998; see also Schubert, this volume) converge.

In Marshall and Cohen's (1988) study, two musical soundtracks were judged to have approximately the same activity level (measured by averaging responses on scales of fast–slow, active–passive, agitated–calm, restless–quiet). The relative activity levels of the three 'characters' in the film, however, differed for the two different musical backgrounds. For example, the large triangle was judged as the most active under one soundtrack while the small triangle was judged as the most active in the other. Marshall and Cohen (1988) argued *post hoc* that shared accent patterns in the music and in the motion of the figures operated to focus attention on the temporally congruent part of the visual scene, and subsequently associations of the music were ascribed to this focus (Fig. 11.2). A similar accent-pattern/association breakdown in the processing of film and music was proposed by Lipscomb and Kendall (1994, p. 91).

Cook (1998), in his book *Analyzing musical multimedia,* has suggested the generality of Marshall and Cohen's (1988) theory to other multimedia examples in which musical meaning alters the interpretation of events that are at the focus of visual attention. In advertisements for cars, for example, the car takes on both the vitality and the high cultural associations of the classical music in the background. Music does more than echo or provide a counterpoint to a concept already present in the film. Music can also direct attention to an object on the screen and establish emotionally laden inferences about that object.

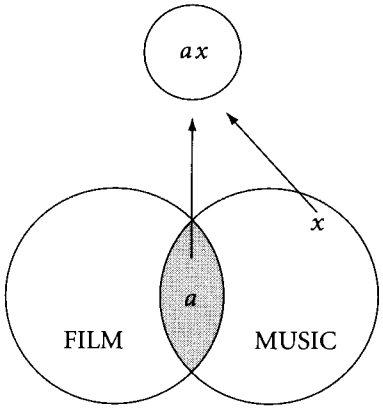



Figure 11.2 Congruence-associationist model based on that depicted by Marshall and Cohen (1988). The total meaning and structure of the music and film are presented by their respectives circles. The overlap in music and film is depicted by the intersection of the circles (*a*). Attention is directed to this area of visual overlap in the film. Other associations of the music (*x*) are ascribed to this focus of attention (*ax*). Thus, music alters the meaning of a particular aspect of the film.

Regarding the ability of music to focus attention, Bolivar *et al.* (1994), following Boltz *et al.* (1991), noted that attention to a visual object might arise not only from structural congruencies but also from semantic congruencies. Hypothetically, for example, a soundtrack featuring a lullaby might direct attention to a cradle rather than to a fish-bowl when both objects are simultaneously depicted in a scene. Subsequently, additional associations from the lullaby would be ascribed to the cradle (and conversely for Schubert's *Trout Quintet*, for those familiar with it). Empirical evidence for the ability of music to focus attention is scant and tends to appear with very simple geometric figures, but has not yet been demonstrated with more complex displays (cf. Lipscomb 1999). Part of the difficulty has been in creation of the materials to study the phenomena; however, rapid advances in the affordability of non-linear editing equipment may make such experiments more practical.

### 11.2.4 Functions of film music

Cohen (1999*a*) described eight functions of music in a film or multimedia context. First, music masks extraneous noises. Second, it provides continuity between shots, for example, when the camera alternates between close-ups of two people who are presumably looking at each other (cf. Magliano *et al.* 1996, p. 205). Third, as Marshall and Cohen (1988) and Bolivar *et al.* (1994) had argued, and as noted by Münsterberg (1970), it directs attention to important features of the screen through structural or associationist congruence. Fourth, when unassociated with a particular focus it induces mood, as often occurs during the opening credits of a film. The ability of music to induce mood has been supported in several experiments (Pignatiello *et al.* 1986) and is used in music therapy (Albersnagel 1988; Bunt & Pavlicevic, this volume). Fifth, it communicates meaning and furthers the narrative, especially in ambiguous situations (Bullerjahn & Güldenring 1994; Cohen 1993; Kalinak 1992; Levinson 1996). Sixth, through association in memory, music becomes integrated with the film (Boltz *et al.* 1991) and enables the symbolization of past and future events through the technique of *leitmotiv*. In leitmotiv, a particular musical theme is continuously paired with a character or event so that eventually the theme conjures up the concept of the character or event in its absence (Palmer 1980, p. 550). The composer Richard Wagner is typically regarded as the first to exploit this principle in opera. In an insightful article by the composer Saint-Saens (1903, p. 259), entitled 'The composer as psychologist', the author remarks that psychological principles must be responsible for the effectiveness of leitmotiv. Mood-dependent memories can also be cued with the emotions established by music (Eich 1995). Seventh, music heightens the sense of reality of or absorption in film, perhaps by augmenting arousal, and increasing attention to the entire film context and inattention to everything else (cf. discussions of reality status by Preston 1999; Qian *et al.* 1999). Finally, music as an art form adds to the aesthetic effect of the film.

## 11.3 A cognitive framework for understanding musical soundtracks

Many of the functions of film music can be explained via notions of congruence or association, because these represent two primary ways in which the brain operates:

through innate grouping principles (Bregman 1990) and by learned connections (Cohen 1993), respectively. Cohen (1999*a*, 1999*b*, 2000) presented a capacity-limited information-processing framework that represented the congruence-associationist concepts in a broad cognitive context (Fig. 11.3).

The framework consists of three parallel channels along the vertical axis, each devoted to one of the significant domains of film: speech, visual information, and music. Each channel is hierarchically organized into four processing levels, with bottom-up levels (A and B) meeting the top-down level (D) at level C, the level of conscious attention and

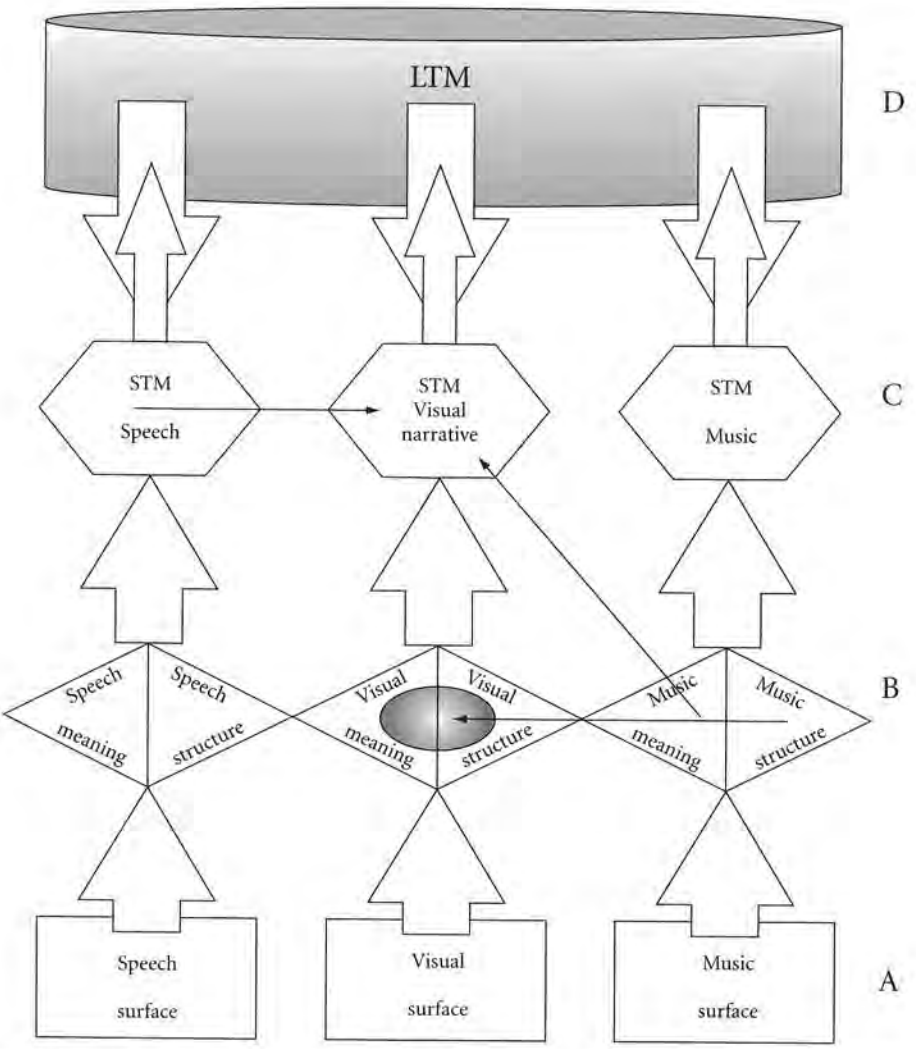



**Figure 11.3** Congruence-associationist framework for understanding film-music communication (see text for explanation). LTM, long-term memory; STM, short-term memory.

short-term memory (STM). Environmental sounds (sound effects) have not been included for the sake of simplicity and because of lack of research to date, but their own grammar and complexity calls ultimately for incorporation as a separate channel. Likewise, the visual channel might admit to further breakdowns, for example, animate/inanimate or background/foreground, but this is secondary to the main concepts presented below.

### 11.3.1 Bottom-up processes: levels A and B

Processing begins at level A with the analysis of physical features of the speech, visual, and musical surfaces into components such as phonemes, lines, and frequencies, respectively. Within each of these domains, at the next level (B), groups of features are subsequently analysed into structural (gestalt-type) and semantic (associationist) information. For music, this means assembly into temporal structures and categorization of cues (e.g. pitch height, tempo, direction) leading to emotional meaning. The outcome of analysis at level B affords the possibility for emergence of cross-modal congruencies, for example, shared accent patterns in audio and visual modalities, to lead preattention to only a portion of the visual information, shown here as the material within the oval in the visual channel.

#### *Cross-modal structural congruence*

To further explain the concept of cross-modal congruence, cognitive psychologists have typically applied gestalt principles to visual pattern (e.g. Wertheimer 1939) and later to auditory information (e.g. Bregman 1990; Meyer 1956; Narmour 1991). Rarely are the principles applied to the visual and auditory domains at once. But film music provides the necessity for such application. The simultaneous presentation of music and film automatically elicits bottom-up principles that entail perceptual grouping in both auditory and visual domains. When the auditory information and visual information are structurally congruent (e.g. share temporal accent patterns), the visually congruent information becomes the figure, the visual focus of attention (as originally argued by Marshall & Cohen 1988).

As part of a larger study, Iwamiya (1994, pp. 134–5) showed that judgments of degree of audiovisual matching of four short film clips was lower when the original video and music components were desynchronized by 500 ms (the only delay examined). Thus, gestalt theoretical ideas that are typically applied to visual or auditory domains independently can be applied to conjoint visual and auditory dynamic information. It follows that through innate gestalt grouping processes, music can define the visual figure against the audiovisual background; music can sometimes determine what is the visual focus of attention.

More recently, the gestalt account of the phenomenal sense of 'belonging together' has been seen as a forerunner of the current solution to a general problem in cognition, that of consciousness. Lately, cognitive scientists have focused on the concept of 'binding' to explain how 'the unity of conscious perception is brought about by the distributed activities of the central nervous system' (Revonsuo & Newman 1999, p. 123). It is interesting to note the significance cognitive scientists attach to neural synchronization associated with shared temporal patterning across neural ensembles. Film music that

shares patterns with visual information on the screen, may therefore contribute to attention and consciousness. Indeed, oscillations (cf. accent patterns) below 10 Hz (the typical limit for the film-score composer's click-track) have been proposed as necessary for information integration and the generation of the stream of consciousness characterizing working memory (cf. Revonsuo & Newman 1999, p. 126).

These notions of congruent patterns in perceptual and cognitive psychology relate to those in the film-music literature on sensitivity to, and effectiveness of, synchronized musical and film structures. Kalinak (1992) suggests that it is this synchronization that contributes to the inaudibility of the music. Synchronization masks the real source of the sound (like ventriloquism):

> The vocal track in classical cinema anchors diegetic sound to the image by synchronizing it, masking the actual source of sonic production, the speakers, and fostering the illusion that the diegesis itself produces the sounds. Mickey Mousing [synchronized music and visual accent] duplicates these conventions in terms of nondiegetic sound. Precisely synchronizing diegetic action to its musical accompaniment masks the actual source of sonic production, the offscreen orchestra, and renders the emanation of music natural and consequently inaudible. Musical accompaniment was thus positioned to effect perception, especially the semiconscious, without disrupting narrative credibility. (Kalinak 1992, p. 86)

### 11.3.2 Short-term memeory visual narrative: level C

Returning to the remainder of the framework of Fig. 11.3, information is transferred to STM at level C. Priority of transfer is given to visual information in the gray oval at level B, which has been selected by cross-modal audiovisual congruence. Not all of the visual sensory information that is potentially available reaches the STM. Also, information of musical meaning from B is transferred not only to a music STM but also to the visual STM. Note that visual STM is referred to as the STM visual narrative. This is a unique aspect of the framework: visual STM is construed as the location of consciousness of the film representation. This decision is supported by the evidence for visual primacy and the typical subservience of audio to vision (e.g. ventriloquism, cf. Driver 1997; see also Bolivar *et al.* 1994; Thompson *et al.* 1994). The concept is that of constructing a narrative from the information gleaned simultaneously from visual, music, and speech sources. The present depiction illustrates how music transports various packages of information, be they structural or semantic, meeting the (diegetic and non-diegetic) goals of a film director and film-music composer for the minds of the audience. The term narrative is in recognition of the aim of the visual STM process to make sense of the visual information using whatever information is at hand. Thus, the emotional meaning of the music is directed here because it is useful in determining the meaning of the visual scene. However, consciousness of this meaning arises only through correspondence with information from top-down processes based in long-term memory (LTM).

### 11.3.3 Top-down processes: level D

During bottom-up processing, some preattended information from all channels proceeds to the LTM at D and begins a top-down inference process with the goal of constructing the narrative of the film. Thus, both bottom-up and top-down processes simultaneously generate information that meets at level C, the STM. In order to achieve

consciousness in the STM of material driven by bottom-up processes, matching of this information by inferences generated from the LTM is necessary. The notion of such a matching process is found in theories of conscious attention (Grossberg's 1995 adaptive resonance theory, ART), comprehension (Kinstch 1998, construction-integration), and consciousness (Baars 1997). These theories assume that preattentive processes are sufficient to initiate the inference processes from LTM.[2]

Because everyday emotional experience associated with events is stored in the LTM, inferences based on past experience would include emotion (e.g. how a protagonist would feel in a certain situation). These inferences generated by the LTM matching process accommodate the visual and emotional information from a film, but not the acoustical properties of the musical accompaniment that are the source of the emotional information. This explains why the acoustical aspects of the music are not generally attended to: in the context of the rest of the narrative, the acoustical aspects of the music do not make sense to LTM (where is that background music coming from?) and no hypotheses would be generated easily to include it (unless, of course, music were part of the diegesis, e.g. the portrayal in the film of attending a concert or taking a music lesson).

An example of this attentional and inference process is provided by consideration of a portion of the film *Apollo 13*, the drama based on the dangerous technical difficulties within the spacecraft, Apollo 13, that prevented the planned landing on the moon and threatened a safe return to earth. Toward the end of the film, Apollo 13 hovers over the moon—so near and yet so far. The film depicts the fantasy of one of the astronauts, Jim Lovell (played by Tom Hanks) imagining his dream of having landed, taking several weightless steps, and slowly brushing his gloved fingers across the moon's surface. The audience has no trouble in inferring the anguish, awe, and exhilaration that he would have felt. Thanks to the composer, James Horner, the musical basis for such emotional information is carried by the musical soundtrack. On the other hand, the story gives no reason to predict that a full symphony orchestra is performing outside the spacecraft; hence, the acoustical aspect of the music is not transported to the STM. It is encoded by sensory memory but it is not predicted by inferences derived from the LTM and hence it is unattended to.

Thus, the main phenomenal experience at the STM visual narrative is one of a narrative with visual, verbal, and emotional components (but not music *qua* sound). Once attended in STM, the information about the narrative can itself be stored in LTM and form the basis of new inferences. In parallel, acoustical aspects of the music can be processed at a conscious level (see levels C and D in the music column, Fig. 11.3), as it is known that simultaneous tasks can co-occur (Neisser & Becklen 1975), and there is evidence that background music is remembered (Smith 1996). A similar process is envisioned for speech as well, but this is not the focus of the present chapter.

[2] In Grossberg's (1996) theory, 'when both bottom and top-down signals are simultaneously active, only the bottom-up signals that receive top-down support can remain active . . . Top-down matching hereby generates a focus of attention that can resonate across processing levels, including those that generate the top-down signals. Such a resonance acts as a trigger that activates learning processes within the system. Resonance is also proposed to be a necessary condition for conscious attention'.

### 11.3.4 Emotion

Emotion cuts through six of the eight functions of film music identified by Cohen (1999*a*): contributing to the narrative's continuity; emotional meaning of events; induction of mood; creation and activation of memory (state dependence, heightening attention to particular events, providing cues in leitmotiv); maintenance of arousal, global attention, and associated sense of reality; and, finally, aesthetic experience. Hence emotion enters at every level of the proposed framework: analysis of structure and meaning, directing emotional meaning to STM, cuing the inference process of LTM, and the matching of LTM and STM representations.

In regard to the nature of experienced emotions generated by music in the context of film, it is useful to consider six criteria of a genuine emotion that Tan (1996) outlined, based for the most part on Frijda's (1986) laws of emotion.[3]

1. *Control precedence.* Music controls emotion response (Thayer & Faith 2000; Thayer & Levenson 1983), hence, like other genuine emotions, emotion created by background music exerts control over the audience member. These effects can result simply from bottom-up analysis of the stimuli, although higher-order learned associations may also play a role.
2. *Law of concern: emotion entails identifiable concern.* When music is combined with other media, the music readily finds an object. Cook (1998) provides clear examples of this with respect to advertising. Marshall and Cohen (1988) explained that music directed attention to an object and ascribed its meaning to that object. Attention is required for concern. Music directs such attention (see horizontal arrow from music to vision at level B in Fig. 11.2).
3. *Law of situational meaning (or stimulus specificity).* Each emotion has a particular 'situational meaning structure', a set of critical characteristics of the stimulus. Characteristics of musical stimuli giving rise to particular emotions have been identified by various researchers (e.g. Juslin 1997; Juslin & Madison 1999; Krumhansl 1997; Rigg 1964), and similarities between these characteristics and visually depicted emotions through gait, posture, and speech intonation have also been noted (for a summary, see Boltz *et al.* 1991; see also Juslin, this volume). Some aspects of the emotional meaning of music transfer directly to film (Iwamiya 1994; Sirius & Clarke 1994; Smith 1999).
4. *Law of apparent reality: the stimulus must represent some reality or other* (see also Tan 1996, p. 67). Music contributes to the sense of reality of the narrative (first demonstrated in the 'talkies' in Steiner's score for *King Kong*; cf. Palmer 1990, p. 28). It accentuates important events. The contribution of only the emotional components of the music to the diegesis has been described in Section 11.2 and explained via Fig. 11.3 (see, in particular, the diagonal arrow between levels B and C).
5. *Law of change: emotion responds to changes in the situation* (see also Tan 1996, p. 56). Music creates an everchanging auditory environment that establishes expectations

[3] But see Russell and Barrett's (1999) practical guide to assessment of emotion in which they claim that films do not induce true emotion, what they refer to as emotional episodes.

and implications, some of which are realized and some of which are violated. As such it is a fertile source of emotion (see also Meyer 1956).

6. *Law of closure: an emotion tends toward complete realization of its appraisal and action tendency, and is relatively immune to outside influences such as conscious control.* Music commands interest, especially in a darkened film theatre, as described by Münsterberg (1970). The emotion generated by music is governed by the tension and resolution established by the music of which the audience is unaware (cf. Thompson *et al.* 1994) and over which one seems to have little control, although this is a matter for further empirical work. Rehearing music reproduces emotional responses regardless of prior expectations (e.g. Jourdain 1997).

Thus, having satisfied the six constraints described by Tan (1996), it can be concluded that music contributes genuine emotional experience in a film. The congruence-associationist framework has provided a perspective for understanding how different aspects of music contribute to this outcome, although much still remains to be explained.

## 11.4 Emotion and the film-score composer

It is well to say that music is a source of emotion in film, but the ultimate source is the composer. The average theatre-goer appreciates the emotion established by film music but would be hard pressed to compose this music. Whereas many classical composers have created film scores (e.g. Saint-Saens, Satie, Britten, Honegger, Milhaud, Prokofiev, Shostakovitch, Vaughan Williams, Bernstein, Copland, Schuller, and Corigliano), such composition is often regarded as a special talent and preoccupation, exemplified by George Steiner, Miklós Rósza, Erich Korngold, Bernard Hermann, Dimitri Tiomkin, John Williams, Rachel Portman, and Ennio Morricone among others. Composers known primarily for their film music, have also been recognized for classical music composition, for example, Rósza and Korngold.

According to film-score composer Victor Young, the film-score composer is characterized by exceptional exactitude, diplomacy, and patience, in addition to musical training (cited in Karlin 1994, p. 310). Music composition for film differs from music composition for its own aesthetic sake. Typically film music is music produced for the sake of the story. It is constrained by the intent of the director, narrative, time, and budget. Working within these constraints, the composer may be regarded as exploiting his or her metacognition of the operations described in the framework of Fig. 11.3. The composer must know how shared audiovisual accent patterns can focus visual attention, how musical information avoids conscious attention, how mood is established, how musical associations provoke inferences through reinforcement or counterpoint, and how inferences are cued and generated via LTM to further the diegesis.

The composer is usually called upon at the end of the film production (Palmer 1980; Rózsa 1982, p. 191; some exceptions being Eisenstein and Hitchcock classics) and may be shown the film for the first time with recorded music already in place, known as temp tracks. The temp tracks indicate the director's wishes for type and placement of music, and therefore can restrict the composer's latitude considerably. In Henry

Mancini's opinion, familiarity with the temp tracks may bias the director against new insights offered by the composer (Brown 1994, p. 301). The composer's job is to replace the temp tracks with new material that must meet some or all of a number of constraints: to time the music cue to a fraction of a second to coincide with the rhythm of the action of a particular frame of the film, to match or create the mood or spirit of the film content, to use affordable orchestration and rehearsal time, to be unheard (unless the music is part of the diegesis) but be memorable, and to never drown out the dialogue (cf. Burt 1994, Chapter 6; Rózsa 1982, pp. 69, 108, and 110). In spite of these constraints, some composers, such as classically trained John Barry, claim that composing for film can be the ultimate freedom. Within these constraints, the composer can do whatever he or she wants and is assured of exposure.

Composing for film is one way of transmitting musical culture (e.g. Rózsa 1982, p. 205), because, as shown in Fig. 11.3, although the film music serves narrative function, it is also encoded in an information-processing channel devoted solely to music. Exposure to new compositional styles can be an added aspect of the film experience. For a recent example, *The Red Violin* may provide one's first exposure to the work of the contemporary composer John Corigliano. Films provide a major source for transmission of a culture's musical conventions. Thus, composing for film is a two-way street: the composer learns to code music to match the visual and emotional information of a narrative; at the same time, the film provides the composer with an opportunity to represent this emotional information in musically novel and creative ways, often to a large audience.

For a feature film, the composer may be given only a month of intensive work to score an hour of music. This pace is faster than that of a composer of 'music alone' but the genre is often, though not necessarily, redundant and characterized by cliché. The music does not have to stand independently, yet the possibility of (recent expectation for) a lucrative soundtrack album may create a challenge to compose music that lives on its own yet hardly reaches consciousness during the film.

Knowledge of the techniques and technology of film scoring can be acquired formally through courses, books, apprenticeships, or trial and error. The art of film music, however, perhaps more than other forms of music, requires 'taking the attitude of the other' (Meyer 1956, Chapter 1). Specific messages must be communicated in an aesthetic package, but the aesthetic goals may be secondary, unlike composing music alone. Like other skills, such as chess, bridge, music performance, or knowledge of a discipline, expertise in film-music composition may follow the 10-year rule of concentrated practice (Ericsson 1996). A young McGill University student composer, Aaron Hanson, in scoring an 8-minute film for a friend, spoke of the many hours of experimentation that were entailed until he achieved the effects he wanted. Presumably, like learning any language late in life, extensive effort is required to master the syntax and vocabulary. But the film-score language differs from languages learned from scratch in that the grammar of the film-music language is already known implicitly from exposure to music and film-music conventions. The film-score composer must turn that implicit knowledge into explicit knowledge and, in other words, must become an expert of the rules.

Research by Lipscomb and Kendall (1994) corroborates the notion that the professional film-score composer has the knowledge to create a score that uniquely matches a portion of the film. Moreover, the explicit knowledge of the composer is implicitly

shared by the audience. Lipscomb and Kendall (1994) asked participants to select the best fitting of five film scores by Leonard Rosenman for a feature film, *Star Trek IV: The Voyage Home*. One of the scores had been originally composed for the excerpt and the remaining selections were by the same composer but drawn from other excerpts in the film. Confirming the effectiveness of Rosenman's music, the most frequent choice of the subjects was the actual score he had composed for the segment, although not every subject made this choice. Similarly, we recall from research previously reviewed by Bullerjahn and Güldenring (1994), that the professional film-score composer can systematically manipulate the inferences generated by the viewer/listener. Likewise, in the study by Thompson *et al.* (1994), the use of musical closure by the professional composer altered the judged closure in the film.

Some composers may be more suited to film-score composition than others in terms of both personality and motivation (which may play more of a role than talent, as is sometimes the case in regular musical achievement; see Sloboda 1996). The film composer must have a dramatic sense (Rózsa, cited in Brown 1994, p. 278), an appreciation of the visual world of film, and a sensitivity to speech nuances. Unlike many other types of composition, the creation of a film score is a collaborative process. Generally, interpersonal intelligence (Gardner 1993) would be necessary on two fronts: appreciation of the demands of socially shared cognition and the accurate assessment of common ground (Krauss & Fussell 1991), and the willingness to cooperate with the film production team (although the film composer Bernard Hermann was known to be irascible, according to Karlin 1994, p. 270). Korngold describes his positive relations with executive producers and others responsible for the film (cited in Carroll 1997, pp. 298–9) and claims that his artistry was not compromised in film composition. Similarly, Franz Waxman felt 'there was always room for fresh musical ideas in writing for the screen' (in Karlin 1994, p. 307).

## 11.5 Conclusion

Emotion characterizes the primary experience of both music (Sloboda 1985) and film (Münsterberg 1970; Tan 1996). Music typically plays an integral part of film. Kalinak's (1992, p. 87) argument for the importance of music to the emotional experience in classical narrative film finds support in much of the information presented in this chapter:

> Scenes that most typically elicited the accompaniment of music were those that contained emotion. The classical narrative model developed certain conventions to assist expressive acting in portraying the presence of emotion . . . close-up, diffuse lighting and focus, symmetrical mise-en-scene, and heightened vocal intonation. The focal point of this process became the music which externalized these codes through the collective resonance of musical associations. *Music is, arguably, the most efficient of these codes*, [italics added] providing an audible definition of the emotion which the visual apparatus offers . . . Music's dual function of both articulator of screen expression and initiator of spectator response binds the spectator to the screen by resonating affect between them.

Kalinak's statement regarding the role of music as a source of emotion is a strong one. She has claimed that music is 'the most efficient code' for emotional expression in

film. According to Kalinak (1992, p. 87): 'The lush, stringed passages accompanying a love scene are representations not only of the emotions of the diegetic characters but also of the spectator's own response which music prompts and reflects'. She is arguing that the simultaneity of both the representation and the elicitation of feeling is key. Though her analysis seems correct, more empirical research would be welcomed that compared the relative abilities of music and film to represent and elicit emotion.

That music contributes to the emotional expression and experience of film seems logical, yet surprisingly discussions of emotion in film often ignore music (e.g. Tan 1996). However, Münsterberg (1970) suggested that the psychological processes underlying film were more similar to those of music than to visual art or drama which on the surface might seem more similar. Experimental evidence since then has shown that music influences the interpretation of film narrative and that the music becomes integrated in the memory with the visual information. Music accomplishes other attentional functions through gestalt structural and associationist principles. In addition, the fact that music requires cognitive resources probably plays a role in determining absorption, arousal, and general attention. Music also contributes to the aesthetic experience of the film. More importantly, for present purposes, music in conjunction with film can satisfy Frijda's requirements, as identified by Tan (1996), of a stimulus that can support genuine emotion.

Film-music composition can be regarded as a type of problem solving that exploits knowledge of the musical rules that express and create emotion through specific musical relations. There are many goals that must be satisfied by the film-score composer: providing continuation, directing attention, inducing mood, communicating meaning, cuing memory, creating a sense of reality, and contributing to the aesthetic experience. The ultimate compositional goal is to produce sound patterns that express the emotion consistent with the narrative, the emotion that is jointly recognized and experienced by the audience, binding the spectator to the screen (Kalinak 1992, p. 87). The capacity of music to accomplish the emotional task, arguably far better than the screen itself as Kalinak has suggested, may be based on the ability of music to simultaneously carry many kinds of emotional information in its harmony, rhythm, melody, timbre, and tonality. Real life entails multiple emotions, simultaneously and in succession. Miraculously, yet systematically, these complex relations—this 'emotional polyphony'—can be represented by the musical medium. An example is Korngold's music from the classic film *Sea Hawk* that links romantic love and the spirit of childhood adventure: 'The music for the love scenes still makes an indelible impression with its sweeping heroic lyricism, characterized by arching, repeated rising sevenths that dovetail perfectly with a hypnotic and unforgettable horn call that is redolent of every schoolboy's dream of pirate adventure' (Carroll 1997, p. 254; other examples are provided by Steiner's ability to 'crystallize the essence of a film in a single theme', see Palmer 1990, pp. 29 and 48)

As depicted by the congruence-associationist cognitive framework presented here, and as argued by Cook (1998), music is strong in the representation of emotion in the abstract, and the screen is strong in representing the object to which the emotion is directed. While more research is warranted to further examine the simultaneous contribution of music to emotional meaning, mood, feeling, and absorption, there are sufficient data available now to conclude that music, owing in large part to the explicit

knowledge and skills of the composer, provides one of the strongest sources of emotion in film.[4]

## References


**Albersnagel, F.** (1988). Velten and musical mood induction procedures: A comparison with accessibility of thought associations. *Behavior Research Theory,* 26, 79–96.

**Anderson, J. D.** (1997). Introduction to the Symposium on Cognitive Science and the Future of Film Studies. In *Proceedings of Symposium Cognitive Science and the Future of Film Studies,* pp. 2–6. Lawrence, KS: University of Kansas. Retrieved 17 June 2000 from the WWW: http://www.gsu.edu./~wwwcom/ccsmi/vsnipi.htm.

**Baars, B. J.** (1997). *In the theater of consciousness: The workspace of the mind.* New York: Oxford University Press.

**Banse, R. & Scherer, K. R.** (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology,* 70, 614–36.

**Barrett, L. F. & Russell, J. A.** (1999). The structure of current affect: Controversies and emerging consensus. *Current Directions in Psychological Science,* 8, 10–14.

**Bolivar, V. J., Cohen, A. J., & Fentress, J. C.** (1994). Semantic and formal congruency in music and motion pictures: Effects on the interpretation of visual action. *Psychomusicology,* 13, 28–59.

**Boltz, M., Schulkind, M., & Kantra, S.** (1991). Effects of background music on remembering of filmed events. *Memory and Cognition,* 19, 595–606.

**Bregman, A.** (1990). *Auditory scene analysis.* Cambridge, MA: MIT Press.

**Brown, R.** (1994). *Overtones and undertones: Reading film music.* Berkeley, CA: University of California.

**Bullerjahn, C. & Güldenring, M.** (1994). An empirical investigation of effects of film music using qualitative content analysis. *Psychomusicology,* 13, 99–118.

**Burt, G.** (1994). *The art of film music.* Boston: Northeastern University Press.

**Carroll, B. G.** (1997). *The last prodigy: A biography of Erich Wolfgang Korngold.* Oregon: Amadeus.

**Cohen, A. J.** (1990). Understanding musical soundtracks. *Empirical Studies of the Arts,* 8, 111–24.

**Cohen, A. J.** (1993). Associationism and musical soundtrack phenomena. *Contemporary Music Review,* 9, 163–78.

**Cohen, A. J.** (1994). Introduction to the special volume on the psychology of film music. *Psychomusicology,* 13, 2–8.

**Cohen, A. J.** (1999*a*). The functions of music in multimedia: A cognitive approach. In *Music, mind, and science,* (ed. S. W. Yi), pp. 53–69. Seoul, Korea: Seoul National University Press.

**Cohen, A. J.** (1999*b*). Music serves as a vehicle in multimedia contexts. [CD-ROM] *Collected papers of the 137th Meeting of the Acoustical Society of America and the 2nd Convention of the European Acoustics Association* (CD rom, 4 pp). New York: Acoustical Society of America.


[4] The Social Sciences and Humanities Research Council is acknowledged for its support of the author's programme of research in the psychology of film music. Appreciation is expressed to four anonymous reviewers and colleague Thomy Nilsson for their comments on an earlier version of this manuscript. The advice of editors Patrik Juslin and John Sloboda is also gratefully acknowledged.

Cohen, A. J. (2000). Film music: Perspectives from cognitive psychology. In *Music and cinema,* (ed. J. Buhler, C. Flinn, & D. Neumeyer). Middlebury, VT: Wesleyan University Press.

Copland, A. (1941). *Our new music.* New York: McGraw-Hill.

Cook, N. (1998). *Analysing musical multimedia.* Oxford, UK: Clarendon Press.

Driver, J. (1997). Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. *Nature,* 381, 66–8.

Eich, E. (1995). Searching for mood dependent memory. *Psychological Science,* 6, 67–75.

Ericsson, K. A. (1996). The acquisition of expert performance: An introduction to some of the issues. In *The road to excellence: The acquisition of expert performance in the arts and sciences, sports and games,* (ed. K. A. Ericsson), pp. 1–50. Mahwah, NJ: Erlbaum.

Frijda, N. H. (1986). *The emotions.* Cambridge, UK: Cambridge University Press.

Gabrielsson, A. (1998). Verbal description of music experience. In*Proceedings of the Fifth International Conference on Music Perception and Cognition,* (ed. S. W. Yi), pp. 271–6. Seoul: Seoul National University.

Gabrielsson, A. & Lindström, S. (1993). On strong experiences of music. *Jarbuch der Deutschen Gesellschaft für Musikpsychologie,* 10, 114–25.

Gardner, H. (1993). *Frames of mind: The theory of multiple intelligences.* New York: Basic Books (originally publisher 1983).

Gorbman, C. (1987). *Unheard melodies: Narrative film music.* Bloomington, IN: Indiana University Press.

Gregory, A. H. (1998). Tracking the emotional response to operatic arias. In *Proceedings of the Fifth International Conference on Music Perception and Cognition,* (ed. S. W. Yi), pp. 265–70. Seoul: Seoul National University.

Grossberg, S. (1995). The attentive brain. *American Scientist,* 83, 438–49.

Grossberg, S. (1996). *The attentive brain: Perception, learning, and consciousness* (abstract). Retrieved 6 June 2000from the WWW: http://www-math.mit.edu/amc/fall96/grossberg.html.

Heider, F. & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology,* 57, 243–59.

Hochberg, J. & Brooks, V. (1996*a*). Movies in the mind's eye. In *Post-theory: Reconstructing film studies,* (ed. D. Bordwell & N. Carroll), pp. 368–87. Madison, WI: University of Wisconsin Press.

Hochberg, J. & Brooks, V. (1996*b*). The perception of motion pictures. In *Cognitive ecology,* (ed. M. P. Friedman & E. C. Carterette), pp. 205–92. New York: Academic Press.

Iwamiya, S. (1994). Interaction between auditory and visual processing when listening to music in an audio visual context. *Psychomusicology,* 13, 133–53.

Jourdain, R. (1997). *Music, the brain and ecstacy: How music captures our imagination.* New York: Morrow.

Juslin, P. N. (1997). Emotional communication in music performance: A functionalist perspective and some data. *Music Perception,* 14, 383–418.

Juslin, P. N. & Madison, G. (1999). The role of timing patterns in recognition of emotional expression from musical performance. *Music Perception,* 17, 197–221.

Kalinak, K. (1992). *Settling the score.* Madison, WI: University of Wisconsin Press.

Karlin, F. (1994). *Listening to the movies.* New York: Schirmer.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* Cambridge, UK: Cambridge University Press.

Kracauer, S. (1960). *Theory of film: The redemption of physical reality.* Oxford, UK: Oxford University Press.

Krauss, R. M. & Fussell, S. R. (1991). Constructing shared communicative environments. In *Socially shared cognition.*, (ed. L. Resnick, J. M. Levine, & S. D. Teasley), pp. 172–200. Washington, DC: American Psychological Association.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Psychology,* 51, 336–52.

Lang, P. J. (1995). The emotion probe. *American Psychologist,* 50, 372–85.

Levi, D.S. (1982). The structural determinants of melodic expressive properties. *Journal of Phenomenological Psychology,* 13, 19–44.

Levin, D. T. & Simons, D. J. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin and Review,* 5, 644–9.

Levin, D. T. & Simons, D. J. (2000). Perceiving stability in a changing world: Combining shots and integrating views in motion pictures and the real world. *Media Psychology,* 2, 357–80.

Levinson, J. (1996). Film music and narrative agency. In *Post-theory: Reconstructing film studies,* (ed. D. Bordwell & N. Carroll), pp. 248–82. Madison, WI: University of Wisconsin Press.

Limbacher, J. L. (1974). *Film music.* Metuchen, NJ: Scarecrow.

Lipscomb, S. (1999). Cross-modal integration: Synchronization of auditory and visual components in simple and complex media. *Collected papers of the 137th Meeting of the Acoustical Society of America and the 2nd Convention of the European Acoustics Association* (CD rom, 4 pp). New York: Acoustical Society of America.

Lipscomb, S. & Kendall, R. (1994). Perceptual judgment of the relationship between musical and visual components in film. *Psychomusicology,*13, 60–98.

Mack, A. & Rock, I. (1998). *Inattentional blindness.* Cambridge, MA: MIT Press

Madsen, C. K. (1997). Emotional response to music. *Psychomusicology,* 16, 59–67

Magliano, J. P., Dijkstra, K., & Zwaan, R. A. (1996). Generating predictive inferences while viewing a movie. *Discourse processes,* 22, 199–224.

Marks, M. (1997). *Music and the silent film.* New York: Oxford University Press.

Marshall, S. & Cohen, A. J. (1988). Effects of musical soundtracks on attitudes to geometric figures. *Music Perception,* 6, 95–112.

Meyer, L. B. (1956). *Emotion and meaning in music.* Chicago: University of Chicago Press.

Münsterberg, H. (1970). *The photoplay: A psychological study.* New York: Arno (originally published 1916).

Murray, I. R. & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America,* 93, 1097–108.

Narmour, E. (1991). The top-down and bottom-up systems of musical implication: Building on Meyer's theory of emotional syntax. *Music Perception,* 9, 1–26.

Neisser, U. & Becklen, R. (1975). Selective looking: Attending to visually significant events. *Cognitive Psychology,* 7, 480–94.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning.* Urbana, IL: University of Illinois Press.

Palmer, C. (1980). Film music. In *New Grove dictionary of music and musicians,* (Vol 6) (ed. S. Sadie), pp. 549–56. Washington, DC: Macmillan.

Palmer, C. (1990). *The composer in Hollywood.* New York: Marion Boyars.

**Pignatiello, M. F., Camp, C. J., & Rasar, L.** (1986). Musical mood induction: An alternative to the Velten technique. *Journal of Abnormal Psychology,* **95**, 295–7.

**Prendergast, R. M.** (1991). *Film music: A neglected art,* (2nd edn). New York: Norton.

**Preston, J. M.** (1999). From mediated environments to the development of consciousness. In *Psychology and the internet: Intrapersonal, interpersonal, and transpersonal implications,* (ed. J. Gackenbach), pp. 255–91. San Diego, CA: Academic Press.

**Qian, J., Preston, J., & House, M.** (1999). Personality trait absorption and reality status evaluations of narrative mediated messages. Paper presented at the Annual Meeting of the American Psychological Association, Boston, MA, August 1999.

**Revonsuo, A. & Newman, J.** (1999). Binding and consciousness. *Consciousness and Cognition,* **8**, 123–7.

**Rigg, M. G.** (1964). The mood effects of music: A comparison of data from four investigations. *Journal of Psychology,* **58**, 427–38.

**Rosar, W.** (1994). Film music and Heinz Werner's theory of physionomic perception. *Psychomusicology,* **13**, 154–65.

**Rózsa, M.** (1982). *Double life.* New York: Hippocrene Books.

**Russell, J. A. & Barrett, L. F.** (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology,* 76, 805–19.

**Saint-Saens, C.** (1903/1951). The composer as psychologist. In *Pleasures of music,* (ed. J. Barzun), pp. 258–64. New York: Viking Press.

**Schubert, E.** (1998). Time series analysis of emotion in music. In *Proceedings of the Fifth International Conference on Music Perception and Cognition,* (ed. S. W. Yi). pp. 257–63. Seoul, Korea: Seoul National University.

**Sirius, G. & Clarke, E. F.** (1994). The perception of audiovisual relationships: A preliminary study. *Psychomusicology,* **13**, 119–32.

**Sloboda, J. A.** (1985). *The musical mind: The cognitive psychology of music.* New York: Oxford University Press.

**Sloboda, J. A.** (1992). Empirical studies of emotional response to music. In *Cognitive bases of musical communication,* (ed. M. R. Jones & S. Holleran), pp. 33–46. Washington, DC: American Psychological Association.

**Sloboda, J. A.** (1996). The acquisition of musical performance expertise: Deconstructing the 'talent' account of individual differences in musical expressivity. In *The road to excellence: The acquisition of expert performance in the arts and sciences, sports and games,* (ed. K. A. Ericsson), pp. 107–26. Mahwah, NJ: Erlbaum.

**Smith, J.** (1996). Unheard melodies? A critique of psychoanalytic theories of film music. In *Post-theory: Reconstructing film studies,* (ed. D. Bordwell & N. Carroll), pp. 230–48. Madison, WI: University of Wisconsin Press.

**Smith, J.** (1999). Movie music as moving music: Emotion, cognition, and the film score. In *Passionate views,* (ed. C. Plantinga & G. M. Smith), pp 146–67. Baltimore, MD: Johns Hopkins University Press.

**Storm, C. & Storm, T.** (1987). A taxonomic study of the vocabulary of emotions. *Journal of Personality and Social Psychology,* 53, 805–16.

**Tan, W.** (ed.) (1996). *Emotion and the structure of narrative film: Film as an emotion machine.* Mahwah, NJ: Erlbaum.

**Thayer, J. F. & Faith, M.** (2000). A dynamical systems model of musically induced emotions: Physiological and self-report evidence. Poster presented at the New York Academy of Sciences Conference on the Biological Foundations of Music, New York, June 2000.

**Thayer, J. F. & Levenson, R.** (1983). Effects of music on psychophysiological responses to a stressful film. *Psychomusicology*, **3**, 44–54.

**Thomas, T.** (1997). *Music for the movies*, (2nd edn). Los Angeles: Silman-James.

**Thompson, W. F., Russo, F. A., & Sinclair, D.** (1994). Effects of underscoring on the perception of closure in filmed events. *Psychomusicology*, **13**, 9–27.

**Wertheimer, M.** (1939). Laws of organization in perceptual forms. In *A source book of gestalt psychology*, (ed. W. D. Ellis), pp. 71–88. New York: Harcourt Brace.

**Zwaan, R. A.** (1999). Situation models: The mental leap into imagined worlds. *Current Directions in Psychological Science*, **8**, 15–18.

PART IV

# THE PERFORMER

CHAPTER 12

# THE SUBJECTIVE WORLD OF THE PERFORMER

ROLAND S. PERSSON

'A musician cannot move others unless he too is moved', C. P. E. Bach argued in his treatise on the true art of playing keyboard instruments, 'He must of necessity feel all of the affects that he hopes to arouse in his audience' (1985, p. 152). Famed singer Janet Baker, in an interview published in *The Observer*, stated of musicians in general that '. . . [their] business is emotion and sensitivity—to be the sensors of the human race' (quoted in Crofton & Fraser 1985, p. 112).

Clearly emotions, feelings, and affects are paramount issues to musicians in communicating and understanding music. In fact, the musician's role is so intrinsically tied to the subjective world of emotion that attempts to regard a piece of music, or a performance of it, as in any way detached from a positive emotional experience of some kind, is often frowned upon by musicians. In composer Fredrick Delius' words: 'How can music ever be a mere intellectual speculation or a series of curious combinations of sound that can be classified like the articles in a grocer's shop? Music is an outburst of the soul' (quoted in Crofton & Fraser 1985, p. 49).

The reasons that music has emotional significance, however, are complex and often paradoxical. This is particularly true in regarding the interplay between the dynamics of socially determined criteria for what is considered a piece of music, or a performance, of 'good quality' and the different aspects of music cognition (Persson, 2000). Without the potential to evoke some type of emotional response, however, music as a cultural phenomenon would most probably lose its basically universal appeal (Gregory & Varney 1996).

This chapter outlines exploratory research into the *phenomenology* of emotion in musical performance, which includes performance motivation and generation, as well as the learning and conceptualization process of constructing 'musical meaning'. Most of the material providing the framework for this chapter is based on exploratory qualitative and quantitative research in combination, pursued in a real-world research setting; that is, in a tertiary British music institution (Persson 1993, 1995, 1996, 2000; Persson *et al.* 1996; Persson & Robson 1995; Robson 1993). Data were generated interactively between researchers and musicians as co-researchers, to accommodate the potential conflict between representatives of traditional science and art as reported by, for example, Hurford (1988), Kingsbury (1988), and Persson (1993).

The purpose of the presented research was to seek hitherto untried ways in music-behavioural research to explore the understanding of a limited number of expert

performers' subjectivity. Hence, the results of the research presented here should not be understood as an effort to generalize findings. They are rather the basis for further research and provide an *initial* understanding of how musicians may construe the subjective aspects of musical behaviour.

## 12.1 Why become a performer?

An obvious first question in seeking to understand the subjective world of the musical performer is to ask what made certain individuals choose a musical career. Nagel (1987), from a largely psychodynamic perspective, has suggested that musicians make such a choice on the basis of gratification of needs, which entails: (a) certain emotion-evoking qualities in the music itself; (b) the potentially positive social feedback of the musical setting: (c) a means to explore devouring and aggressive impulses by the specific motor skills involved in playing an instrument; and (d) a degree of exhibitionism and voyeuristic impulses. However, early relationships, such as the support and encouragement of parents and significant others, also play an important role. One example of this is renowned violinist Ernst Wallfisch (quoted in McDonald 1970, pp. 514–15):

> Music played such a big part in my life from the beginning . . . and I recall much . . . my father was a businessman and an amateur violinist . . . music was his big hobby . . . his circle of friends included many like him and chamber music seemed part of living . . . I can remember as a very small boy all of these people in our home talking, laughing, smoking cigars, and making music.

However, relationships in this context may also be understood as finding certain role models, or 'heroes', who possess certain attractive attributes or qualities, which the developing musician wishes to master or in some way emulate. Drummer Steve Gadd, for example, made several famous jazz musicians his role models (quoted in Boyd & George-Warren 1992, p. 125):

> When I was a young teenager . . . my father used to take me to clubs to hear a lot of organ groups, where the groove got so intense . . . The music was feeling so good you just couldn't sit still. That feeling made you want to be a part of what was creating it . . . I heard a lot of jazz groups like Dizzy [Gillespie], and Oscar Peterson, Max Roach, Gene Krupa, and I loved them all. Can remember listening to Tony Williams and Elvin Jones, listening to things over and over again . . . They touched me so much. You get something from everyone you hear; you try to take it and use it and pass it on.

Motte-Haber (1984), taking a broader theoretical view of what motivates musicians, proposed that achievement motivation, curiosity, self-actualization, the conquering of failure anxiety, as well as the fear of success should all be considered in understanding the basis of motivation specific to learning and mastering a musical instrument. Hallam (1998) puts forward self-determination as an additional variable.

Persson (1993) and Persson *et al.* (1996), in interviewing piano performers, and combining qualitative analyses of open-ended responses with factor-analytical procedures on given variables, relating to different aspects of motivation as found in the literature, established a tentative typology of motivation specific to performance motivation (Table 12.1). Although a small sample of participants ($n = 15$; of whom ten were male and five were female), the musicians were very eloquent in discussing their interest in music and their choice of a musical career.

Table 12.1 Components of musical performance motivation as elicited amongst pianists ($n = 15$). The main motive categories are listed in order of significance (from Persson 1993)

| | Type of motivation | Description (and function) |
|---|---|---|
| 1 | Hedonic motive | The search for positive emotional experience |
| 2 | Social motive | The significance of group identity and belonging |
| 3 | Achievement motives: | |
| | 3a Exhibitionism | A desire to show results of an effort |
| | 3b Independence | The means to achievement are secondary to success itself |
| | 3c Dependence | The means to achievement are important, but do not constitute the ultimate target |
| | 3d Aesthetics | The means also provide the target |
| | 3e Support | Motives are extrinsically provided, mainly by teachers and parents |

While motives differed considerably amongst the musicians, they all had in common that their musical pursuit offered, for different reasons, satisfaction. Most participants were therefore motivated primarily by a *hedonic* motive; that is, treasuring music as a means to generate positive emotional experiences mostly for one's own satisfaction. A typical response given by the participants was, for example: 'Mood is a big one for me. You have to be in the right mood, because sometimes you can be sort of in a state where you can't exactly get yourself into how the music should be played. Sometimes you sit down and play as if . . . Ahhhh! [The musician sighs profoundly] You can really *feel* everything in the music' (quoted in Persson 1993, p. 150).

Social motives were also significant in choosing a musical career. It is important to belong to a certain group that has certain desirable attributes with which one identifies. While such identification is occasionally prevented by the so-called 'maestro' function (Persson 1996, 2000), the weight of social motive amongst the musicians interviewed may also suggest that performers are not necessarily the self-sufficient 'lone rangers' that the historical and mythical tradition of the musical maestro often would tempt us to believe (cf. Braudy 1997; Osborne 1990). On the contrary, social identification with a certain musical setting, group, or mentor, as some research has shown, is very important, even a key factor in optimizing musical and artistic development (Manturzewska 1990; Ruud 1997).

Achievement motives, interestingly, tended not to be a priority in the studied group. That is not to say, of course, they do not play a significant role in any career or goal setting. Six types of achievement motivation were identified. Some were exhibitionistic and revealed a desire to show off the results of an effort (cf. Nagel 1987). One informant typically exclaimed, when asked whether he liked being alone on a stage to perform: 'Oh lovely! I enjoy that!' Others were categorized as being independently motivated (where the means to achievement are secondary to success itself):

> . . . basically, it is the attention to yourself, wanting to achieve something. Music is something I enjoy. I have always had this drive in me to do well in whatever I do, so it is not necessarily just music but any other thing . . . I enjoy cooking . . . and I know I do it well because I enjoy it. (Quoted in Persson 1993, pp. 154–5)

However, some participants were categorized as dependently motivated (where the means to achievement are important, but do not constitute the ultimate target of the

musical pursuit): 'I've got this degree. That is motivating for me but I don't wish to be famous by the end of it. I want to do something administrative, which is not so much performance-related really' (quoted in Persson 1993, p. 155).

A few were driven mainly by an aesthetic motive in which the means also provide the target itself. One example of this would be one of the participating performance teachers, who ardently argued that 'Music making is for its own sake, isn't it? You want to feel that you have measured up to what the music demands' (quoted in Persson 1993, p. 266). Well-known violinist Isaac Stern (1979) similarly remarks, in an interview published in *The New York Times Magazine*, that 'somewhere along the line the child must become possessed by music, by the sudden desire to play'.

Finally, some participants felt that the reason they chose a musical career was mainly due to extrinsic motivation. These participants were categorized as displaying a supported achievement motive (where motives are mainly extrinsically supplied by teachers and/or parents). A typical response was the following: 'When you are a child, it is difficult to separate out those [motivational aspects]. I was kind of forced into doing it then, but about the age of twelve I *wanted* to do it' (quoted in Persson 1993, p. 157).

## 12.2 Determinants of performance generation

Although there are many reasons for embracing a musical career, the importance of being able to control and induce pleasant emotional experience by means of playing appears to be the most fundamental consideration amongst the musicians participating in Persson *et al.*'s (1996) study. If emotions and the phenomenology of experiencing music are paramount to many, if not all, musicians, then, which are more specifically the potential determinants of generating and conceptualizing any particular performance?

The question of what makes a musical interpretation 'artistic' has fascinated the cognitive science of music for some time, and the pursuit of understanding just how performers manipulate musical parameters (e.g. tempo, timbre, dynamics, articulation) to obtain an optimal performance has led to the study of algorithms for computer-generated performances (e.g. Friberg 1991), generative principles as based on motor skills and the nature of their programming (Clarke 1988; Parncutt *et al.* 1997), as well as the nature of specific mental representations (Lehmann 1997). As important as these research efforts are, however, the notions of 'feeling' (the term often preferred by musicians) and 'emotional experience' (the term usually preferred by behavioural science), so important to the musicians themselves, are conspicuously absent from this literature.

While questions have been asked how music, and musicians, communicate emotions to listeners and how music-interpretational intentions are perceived by them (e.g. Clynes 1989; Gabrielsson & Juslin 1996; Juslin 1997, 2000; Ohgushi & Senju 1987; Rapoport 1996; Waterman 1996; see also Juslin, this volume), the question of individual conceptualization and its emotional significance is elusive to a traditional psychological research paradigm, and has therefore rarely been asked. Classical in its field, of course, is Meyer's (1956) treatise *Emotion and meaning in music*, but only recently have attempts been made to address the issue of musical meaning by empirical research (e.g. Sloboda 1998; Watt & Ash 1998).

To understand the nature of motor programming, mental representations, and musical meaning more fully, music psychological researchers also need to ask *why* a performer plays in a certain way but not in another, and by what means—other than motor programming strategies and melodic and harmonic memorization procedures—musicians learn a piece of music as well as recreate it in a performance situation. Although there is some controversy about whether or not emotion and cognition should be regarded as separate systems, it is nevertheless well established that cognition and emotion affect each other in different ways in different circumstances (cf. Lazarus *et al.* 1980; Zajonc 1980). Hence, emotions inevitably play a significant role in learning, conceptualizing, and performing music also (cf. Bower 1992). There is therefore a need for music psychology—and perhaps even more so for music education—to begin to understand 'the emotional frame of mind' that generates recall, certain motor patterns, and communicative potential.

### 12.2.1 Performers and imagery

In a series of studies, Persson (1993) endeavoured to explore and chart the emotional aspects of conceptualizing music for a performance. A piece of music (Russian composer Reinhold Glière's *Prelude*, Opus 31, No. 1), unknown to all of the 15 participating pianists from a tertiary British music institution, was chosen because of its manifold interpretational ambiguity, and used as a stimulus in the research design. The musicians were all asked to learn and study the piece for 2 weeks, after which it would be performed, recorded, and discussed in an in-depth interview with the researcher. All interpretational clues provided in the score by the composer, including the title of the piece, were erased. Participants were specifically asked to deal with the music in a way that felt appropriate to them, rather than to consider traditional conventions for any given musical style. In other words, musicians were encouraged to break with perceived conventions should they think it necessary to optimize the conceptualization process.

During the period of study, participants were also asked to supply a descriptive title, which they thought suited the piece and their understanding of it. The title then became the point of departure for the interview on how the performers conceptualized the music at hand. Interview transcripts were verified by the participants themselves and were subsequently subject to analysis (for an exhaustive outline of research strategy and analytical models, see Persson 1993). To prompt musicians to elicit a title was believed to reveal something about the process by which they arrived at an individual understanding of the studied music. Seashore (1938), for example, as possibly the only music psychologist who has addressed the issue, argued in his time that musical imagery is *necessary* in all forms of musical memory:

> In vivid musical memory we relive the music. The person who does not have the capacity to do so may recall in abstract terms; such as the musical notation or even the most refined logical concepts of elements in musical performance and musical criticism. But these are only the cold facts. He does not relive the music . . . The non-emotional person can recall the cold facts, but these facts are not the essence of the music, the welling up of musical emotion . . . I have developed a classification of types of musical imagination in terms of which we can readily classify musicians with whom we are acquainted. Basic types in such classification are the sensuous, the intellectual, the sentimental, the impulsive, and the motor. Any given individual may be dominantly one of

these types but ordinarily the personality represents integration of two or more. (Seashore 1938, pp. 168–72)

Not all participants in the study succeeded in finding a descriptive title, and not everyone, as Seashore (1938) would have predicted, was keen on labelling a piece of music in such a way. Following instructions, however, all supplied the piano piece with a title (Table 12.2). Even the reluctant musicians' choices were important to the research, since in spite of their reluctance, they nevertheless had to draw from the musical structure to find something appropriately descriptive. Musical structure is related to emotional response (Sloboda 1991; see also Gabrielsson & Lindström, this volume).

**Table 12.2** Participating musicians' descriptive titles, types of titles, and duration of performance ($n = 15$). A denotes student performers and B their teachers (from Persson 1993)

| Performer | Chosen title | Type of choice | Duration |
|---|---|---|---|
| A1 | By the water's edge | Mood/semblance | 2′ 14″ |
| A2 | Lazy Sunday by the river | Semblance | 1′ 48″ |
| A3 | Rain | Semblance | 1′ 47″ |
| A4 | Soaring | Mood/semblance | 1′ 53″ |
| A5 | Thoughts | Semblance | 1′ 39″ |
| A6 | Falling feelings | Mood/semblance | 2′ 13″ |
| A7 | Portrait of a lost love | Mood | 2′ 40″ |
| A8 | The typical romantic | Idiom/mood | 1′ 39″ |
| A9 | Saturday sprinting | Semblance | 1′ 06″ |
| A10 | Autumn | Mood/semblance | 3′ 20″ |
| A11 | Coming home | Mood/semblance | 1′ 48″ |
| B1 | Souvenir de Chaminade | Idiom | 1′ 58″ |
| B2 | Song without words | Idiom | 2′ 22″ |
| B3 | Melancholy (tristesse) | Mood | 2′ 00″ |
| B4 | I wonder why? Warum? Warum? Pourqoui? Pourqoui? | Structure | 2′ 13″ |

Mean performance duration was 2′ 00″.

Titles appear to have been chosen according to mainly four criteria: *semblance* prompted musicians to provide a title in relation to some extramusical association, whereas *mood* made the participants rather focus on their affective response to the music. A musician who appeared to follow a strategy of semblance when learning the music at hand would typically explain his or her choice of title in the following way (quoted in Persson 1993, p. 189):

Sitting by the river, listening to the quiet lap of water. Obviously a warm day! Perhaps going there every day; leisure on the banks, passing dykes, the occasional wildlife here and there . . . well, my descriptive title is rubbish. It is *Lazy Sunday by the river*. The title is rather corny, but—to me—that is what it conjures up . . . It wasn't really telling a story. It was more sort of a scene 'à la Debussy'.

An individual who was categorized as relying mainly on mood would rather respond, for example, in the following representative manner (quoted in Persson 1993, p. 189):

When I read the instruction on the sheet, I thought 'Oh no!', because I remember . . . I could never give pieces and compositions a title. But when I actually started to play [the provided piece], no hassle at all. I was really surprised at the title. I couldn't decide between *Falling leaves* or *Falling feelings*. I wanted 'falling' because I got the impression that the whole . . . The theme is always falling . . . I said falling because I think it is quite a sad piece—melancholy.

Some performers, however, focused on *idiom* when selecting a title as they tried to identify the music stylistically. These participants would argue that the music described nothing but itself. One performer insisted that *structure* was the most significant criterion according to which a title should be chosen. This particular performer argued that there were similarities between melodic phrases and language inflection in several languages, hence her chosen title 'I wonder why? Warum? Warum? Pourquoi? Pourquoi?'.

In further discussions with the participating musicians, the significance of imagery in learning and conceptualizing music was probed. The discussion with each individual was, of course, linked to their choice of title. It would seem that all participants made use of imagery in one way or another in order to construe understanding and meaning, although some would argue that they did not. The analysis of the interviews, however, suggests that learning a piece of music, at least for the occasion staged as part of the research project, most certainly entails a framework of memories that is more extensive than merely relating it to pitch, dynamics, timbre, and the social conventions of interpretation. It also suggests that imagery is used *differently* by musicians.

Participants clearly differentiated between *emotional memory* (associating to anything that might trigger an emotional response, that is also compatible with the qualities inherent in the musical structure) and *memory of an emotion* (recalling a certain emotion, or pattern of emotions, in order to evoke a certain state of mind, which is perceived as appropriate for the piece of music to be studied or performed, and that may, or may not, entail visual imagery; cf. Stanislavsky 1988). The musicians also prioritized one or the other: five participants regarded visual imagery (in terms of emotional memory) as their primary concern in conceptualizing or performing music and memory of emotion was secondary. Eight musicians felt that memory of emotion was primary and visual memory only secondary in importance. Four participants avoided the issue altogether and appeared to feel uncomfortable with discussing music in these terms. However, such a reaction is not to say that these participants did not conceptualize music in a like manner. Their reluctance to discuss and reflect on music in this way is likely, at some level, to be socially determined and to some extent also be influenced by type of personality, as evidenced in participants' sometimes contradictory statements on what they *wished* to do with a particular interpretation and what they felt they *had* to do.

### 12.2.2 Emotion and self-induction to altered states of awareness

One particularly important finding in the exploratory research on the generation and conceptualization of musical performance was how musicians—some more than others—very consciously manipulated recall of certain memories (emotional memories and memories of emotions) in order to 'get into the mood'. Out of 15 participating pianists, eight acknowledged some sort of self-discovered technique to arrive at a

particular emotional frame of mind for the sake of learning and/or performing. Three musicians were identified as systematically employing what could appropriately be labelled *mood induction* (that is, to remember the sensation of a certain specific emotion without necessarily conjuring up specific imagery). Five musicians, however, used induction by means of *visualization* in order to evoke a particular emotion. The aim of both strategies was to form the basis of their particular understanding of the music at hand. A musician applying mood induction would typically say:

> I find often that to get me into the mood of a piece, say the 'Pathétique Sonata' by Beethoven, and the opening of that, you think something sad. You think *sadly*, not necessarily something that has happened to you, but you think of *the experience of sadness* before you play that chord . . . [it] is a subconscious thing and you are just trying to bring it out' (quoted in Persson 1993, p. 197).

Other performers seemed to combine both mood induction and visualization, depending on circumstance, to attain a certain frame of mind for the performance:

> I very rarely resort to just [the technical]. I don't like that! To me, I'd rather have it musical. [That is], you can see a sense in things; see pictures and images, and feel things. If you're not in the mood for that [music, which you have at hand] you are obviously not going to do justice to it. I suppose being a musician you have to get into the mood of it. (Quoted in Persson 1993, p. 198)

Both approaches are reminiscent of what the musicians of the baroque and classical eras, following the influential 'doctrine of affections' (Bukofzer 1947; see also Cook & Dibben, this volume), then prescribed as a necessity for 'true' performing musicians: 'A musician cannot move others unless he too is moved' (Bach 1985). Thus, performers appear to develop strategies by which they succeed to induce themselves into an 'altered state of mind'. Such a state, in this context, and with the different theoretical schools of thought regarding altered states of consciousness in mind, is best defined broadly, such as done by Ludwig (1990, pp. 18–19), as:

> . . . any mental state(s), induced by various physiological, psychological, or pharmacological maneuvers or agents, which can be recognized subjectively by the individual himself (or by an objective observer of the individual), as representing a sufficient deviation in subjective experience or psychological functioning from certain norms for that individual during alert, waking consciousness. This sufficient deviation may be represented by a greater preoccupation than usual with internal sensations or mental processes, changes in formal characteristics of thought, and impairment of reality testing to various degrees.

These induction strategies are presumably developed intuitively by musicians, as a result of potentially emotion-evoking structures inherent in the music in combination with striving towards hedonic satisfaction (see also Zuckerman 1994). It is known, for example, that music—at least under certain circumstances—is a particularly efficient medium for inducing an altered state of consciousness (Kenealy 1988; see also Becker, this volume). A number of experiments have shown that induction to an altered state is also subject to a learning process (Kamiya 1990; Shore 1990).

Note that performance teachers never discussed such induction procedures during lessons (Persson 1993), which makes it feasible to propose that mastering such induction procedures is mainly a personal and intuitive matter; it develops over time and—because it is natural and so intrinsically tied to the way we perceive and construe music—is probably often taken for granted in teaching performers.

The pivotal significance of the altered state of mind for a creative effort, however, is fairly well documented in the literature on creativity (cf. Krippner 1999), where phenomena such as inspiration (Kneller 1965), imagination (Stein 1974), peak experiences and being cognition (Maslow 1968), flow (Csikszentmihalyi 1988), and play and fantasy (Russ 1993) are in focus. An extraordinary account of this in the realm of musical performance has been provided by Swedish cellist Frans Helmerson (1990, pp. 250–1), when asked to account for how creativity relates to his artistry (translated from Swedish by the author):

> My first point of departure is to evoke the whole [of the piece]. The understanding of the whole is a prerequisite in order to provide the music with a content. It is like assembling the parts in this entire [conceptualization] . . . In other words, it takes a creative process to fit the minute parts of this composition together . . . later in concert I want to—to put it rather solemnly—transcend into a different dimension; one which makes me one with the music. It is a dimension where boundaries between consciousness and intuition are vague and evanescent.

While this 'inspiration-seeking behaviour' certainly could be seen as a social artefact of primarily the romantic era (cf. Persson 2000), the same need for, and use of, the altered state of mind is also found amongst contemporary musicians. Boyd and George-Warren (1992, p. 156), in interviewing 75 popular musicians (such as Eric Clapton, Peter Gabriel, Keith Richards, Ravi Shankar, etc.) regarding their creative process also note, that:

> . . . by completely concentrating on the music they're playing or writing, musicians are able to open themselves up to a peak experience. It is as if an intense concentration can push the conscious mind away from 'self-consciousness' and the unconscious is allowed to filter through. The result can be songs that seem to come from nowhere, the ability to suddenly play a riff that had been too difficult before, or an on-stage 'merging' with band mates.'

### 12.2.3 Conceptual flexibility

If a performer conceptualizes a piece of music by constructing a type of affective pattern, related to the musical structure, then how stabile is this conceptual structure? Can a musician generate more than one understanding of a certain piece? These questions were also addressed in Persson's (1993) study. Participants were asked to produce two *different* presentations of the given piece and then discuss both. All participants could produce different versions and understandings of the music at hand, but all but one did not. Most argued that they found one conceptualization which they preferred, and that they did not wish to deviate from.

Above all, the content of their conceptual structure was determined by the chosen tempo (see Table 12.2), which has been shown to be the most potent structural factor in music to determine the type of resulting emotional response (e.g. Holbrook & Anand 1990; Rigg 1940; Wedin 1972; see also Gabrielsson & Lindström, this volume; Juslin, this volume). The participants also argued, however, that they could well imagine another conceptualization of the piece featuring different dynamics, tempi, and articulation, but *only* if it was played by someone else!

This response is likely to indicate that a conceptualization is often inflexible and exists within the cognitive confines of what could be functionally termed a *degree of*

*tolerance.* That is, a musician will often only accept a way of playing a certain piece of music that does not alter the musical parameters to such an extent that they trigger an emotional response different to that which constitutes the way in which they originally construed meaning and understanding in relation to the musical structure. However, there is *some* flexibility in this schema of cognitive-emotional memory structures, the extent of which is likely to be related to personality—and perhaps tolerance for ambiguity in particular (Rogers 1967).

It is possible also that the original conceptualization may change altogether over time in some cases due to accumulated experience and new discoveries that will prompt a musician to re-evaluate 'meaning', thus potentially creating an entirely new schema with a new set of parameters. Pianist Alfred Brendel, for example, argues that:

> . . . as I mold my interpretation and conception I play as instinctively as possible; only later do I attempt to understand what I am doing, why I am doing it. Then I start correcting myself whenever necessary, and from that moment on, I am reassessing my findings as often as I can (quoted in Dubal 1985, p. 91).

In the present research sample, there were apparently limits to a conceptualization, which, when transgressed, prompted the musician to either reject a certain performance or understanding—since it is no longer compatible with the construed 'meaning' of the piece at hand—or construe an entirely new understanding of the music, which takes motivation, effort, and time. However, social context is also likely to play a significant role in determining the degree of tolerance.

It is possible that circumstances and performance tradition *prevented* the participating musicians in deviating from a perceived norm. Although participants knew no details of the piece of music at hand, they still tended to project styles and conventions, typical of the romantic era, onto it—in spite of being encouraged to approach the piece more freely. For example, one performer argued the following in trying to conceptualize the given piece of music in more ways than only one: 'When I think of how my highly respected former piano teacher would have played this, I would know exactly what he would do. I would highly admire that and it would [also] govern what I would do' (quoted in Persson 1993, p. 209). A conservatoire culture is known to often not encourage creative initiatives amongst performers-to-be, but rather insists on conformity and adherence to very elusive rules of tradition and 'authenticity' (Kennedy 1991; Persson 2000; Renshaw 1986).

## 12.3 Musical reality and its further research

Musicians from all eras, and no doubt in all genres of music, are indeed 'sensors of emotion and sensitivity', and subjectivity is, in a variety of ways, most probably the basis of their musical pursuit. It is thus helpful, both for future research and in understanding the phenomenology of musical endeavours, to term the dynamic nature of this emotional basis from which musicians appear to draw motivation, construe artistic understanding, and generate performances, a *musical reality* (Persson 1995), akin to the Kellyan notion of personal constructs (Kelly 1963).

Musicians, at least as evidenced in the research surveyed in this chapter, do indeed construe elaborate meaning based on intrinsic and mainly hedonic criteria, on extrinsic and contextual demands, and on the triggers of emotional responses inherent in musical structures. Musicians further seem to uphold musical reality, pertaining to a particular repertoire, by means of a monitoring function, presumably optimizing different means of achieving a self-serving satisfaction. Haroutounian (1995) has very appropriately labelled such a function *meta perception*: the perceptual and cognitive process of internally manipulating perceptions with expression.

Needless to say, the subjective conceptualization of music needs further study. The results obtained in the discussed research project, as intriguing and suggestive as they are, are still tentative. Above all, the phenomenological excursion in to the subjectivity of musical performers has generated a number of hypotheses that need testing and further exploration in order to arrive at verifiable cognitive models, and no less important, to facilitate application in an educational setting (Juslin & Persson, in press).

While the existence of a subjective musical reality is certain, its nature and function, also in different contexts, need specific research and a theoretical framework. For example, how does musical reality more specifically relate to choice of repertoire, to determining an interpretation, to how a piece of music is rehearsed, and to the actual public performance of it? There are almost certainly differences, as well as common denominators, regarding how emotions (and by implication also the construct of musical reality) will function in these different situations. It is also quite likely, however, that there are *several* strategies available to musicians that govern musical behaviour in these different settings. This was suggested by one participant, who commented: 'Despite the fact when I'm playing to an audience, I play mainly . . . for recognition . . . [I also play at other times] for a pure love of playing' (quoted in Persson 1993, p. 175). Hence, it is one thing to play to an audience (being under public scrutiny and being assessed as artist) and occasionally another to play for enjoyment (with no sense of external scrutiny and evaluation).

Thus, there seems occasionally to exist a conflict between the internal hedonic motive (the largely self-serving positive emotional experience) and the external demand (i.e. socially imposed traditions and expectations). In some circumstances, however, the different strategies employed in different contexts should perhaps simply be understood, in a Kellyan sense, as *several* coexisting musical realities for the same repertoire, but constructed differently for a variety of performance contexts. If so, there would not necessarily exist a conflict, but the existence of several such coexisting realities would perchance be indicative of adaptive 'survival behaviour' rather than artistic persuasion.

In researching how emotions interact with decisions, choices, rehearsal behaviour, and public performance—and, of course, performance under the pressure of a perceived assessment; real or imagined—therefore, it is probably necessary to consider both possibilities: the potential conflict between self-satisfying hedonism and the imposition of context expectations and assessment, as well as the possibility of several concurrent constructed realities. Their common denominator, however, would be a framework of emotional patterns, motives, imageries, memories, and associations.

A pertinent question is how research into the subjective world of the performer might proceed. I envisage three main thrusts of necessary research efforts.

1. Cognitive functions and neurological correlates need to more solidly, and theoretically, outline the nature of the connection between emotion and musical learning.
2. There are developmental aspects to the notion of musical reality, and research questions need to include possible gender differences. How does musical reality develop over time, in what way, and is there possibly a difference between male and female performers in the way they develop and establish preferences for how to learn? Note that, for example, research on music and cognitive styles has so far not demonstrated that such individual differences have much bearing on understanding musical behaviour (cf. Barry 1992; Heitland 1982). But perhaps an extended research effort into musical performance and personality (regardless of theoretical preference) would (cf. Kemp 1996).
3. Also, given that the notion of musical reality is an appropriate heuristic for future research, one particularly interesting question inevitably emerges: how does musical reality—or rather the capacity to generate effectively an emotionally based conceptualization—relate to musical ability and talent? The understanding of musical talent—or giftedness—has not changed much during the hundred years or so that music has been studied as behaviour. The subjective features, which are presumably part of talent and giftedness, have thus far not been studied at all. Hence, to tie musical reality to the notions of expertise and talent would also seem an important future research effort.

## References

**Bach, C. P. E.** (1985). *Essay on the true art of playing keyboard instruments*, (ed. and trans. W. J. Mitchell). London: Eulenburg Books. (Originally published in 1778.)

**Barry, N.** (1992). The effects of practice strategies, individual differences in cognitive style, and gender upon technical accuracy and musicality of student instrumental performance. *Psychology of Music*, **20**, 112–23.

**Bower, G. H.** (1992). How might emotions affect learning? In *The handbook of emotion and memory: research and theory*, (ed. S-Å. Christianson), pp. 3–31. Hillsdale, NJ: Erlbaum.

**Boyd, J. & George-Warren, H.** (1992). *Musicians in tune. Seventy-five contemporary musicians discuss the creative process.* New York: Fireside/Simon & Schuster.

**Braudy, L.** (1997). *The frenzy of renown. Fame and its history.* New York: Vintage Books/Random House.

**Bukofzer, M. F.** (1947). *Music in the Baroque Era: music from Monteverdi to Bach.* New York: W. W. Norton & Co.

**Clarke, E. F.** (1988). Generative principles in music performance. In *Generative principles in music. The psychology of performance, improvisation, and composition*, (ed. J. A. Sloboda), pp. 1–26. Oxford, UK: Clarendon Press.

**Clynes, M.** (1989). *Sentics: The touch of emotions.* Bridport, UK: Prism Press.

**Crofton, I. & Fraser, D.** (1985). *A dictionary of musical quotations.* London: Routledge.

**Csikszentmihalyi, M.** (1988). The flow experience and its significance for human psychology. In *Optimal experience: Psychological studies of flow in consciousness*, (ed. M. Csikszentmihalyi & S. Selega-Csikszentmihalyi), pp. 15–35. Cambridge, UK: Cambridge University Press.

**Dubal, D.** (1985). *The world of the concert pianist. Conversations with 35 internationally celebrated pianists.* London: Victor Gollancz.

**Friberg, A.** (1991). Generative rules for music performance: a formal description of a rule system. *Computer Music Journal,* 15, 56–71.

**Gabrielsson, A. & Juslin, P. N.** (1996). Emotional expression in music performance: between the performer's intention and the listener's experience. *Psychology of Music,* 24, 68–91.

**Gregory, A. H. & Varney, N.** (1996). Cross-cultural comparisons in the affective response to music. *Psychology of Music,* 24, 47–52.

**Hallam, S.** (1998). The predictors of achievement and dropout in instrumental tuition. *Psychology of Music,* 26, 116–32.

**Haroutounian, J.** (1995). Talent identification and development in the arts: An artistic/educational dialogue. *Roeper Review,* 18, 112–17.

**Heitland, K. W.** (1982). *Cognitive styles and musical aptitudes: An exploratory study.* Unpublished doctoral dissertation, School of Education, Indiana University.

**Helmerson, F.** (1990). Att tolka musik [How to interpret music]. In *Om kreativitet och flow,* (ed. G. Klein), pp. 250–2. Stockholm: Brombergs.

**Holbrook, M. & Anand, P.** (1990). Effects of tempo and situational arousal on the listener's perceptual and affective responses to music. *Psychology of Music,* 18, 150–62.

**Hurford, P.** (1988). *Making music on the organ.* Oxford, UK: Oxford University Press.

**Juslin, P. N.** (1997). Emotional communication in music performance: A functionalist perspective and some data. *Music Perception,* 14, 383–418.

**Juslin, P. N.** (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance,* 26, 1797–813.

**Juslin, P. N. & Persson, R. S.** (in press). Emotional communication. In *The science and psychology of music performance: Creative strategies for teaching and learning,* (ed. R. Parncutt & G. E. MacPherson). New York: Oxford University Press.

**Kamiya, J.** (1990). Operant control of the EEG alpha rhythm and some of its reported effects on consciousness. In *Altered states of consciousness,* (3rd edn) (ed. C. T. Tart), pp. 600–11. San Fransisco, CA: Harper.

**Kelly, G. A.** (1963). *A theory of personality: The psychology of personal constructs.* New York: W. W. Norton & Co.

**Kemp, A. E.** (1996). *The musical temperament. Psychology and personality of musicians.* Oxford, UK: Oxford University Press.

**Kenealy, P.** (1988). Validation of a music mood induction procedure: Some preliminary findings. *Cognition and Emotion,* 2, 41–8.

**Kennedy, N.** (1991). *Always playing.* London: Weidenfeld & Nicolson.

**Kingsbury, H.** (1988). *Music, talent, and performance. A conservatory system.* Philadelphia, PA: Temple University Press.

**Kneller, G. F.** (1965). *The art and science of creativity.* New York: Holt, Rinehart & Winston.

**Krippner, S.** (1999). Altered and transitional states. In *Encyclopedia of creativity,* (Vol. 1) (ed. M. A. Runco & S. R. Pritzker), pp. 59–70. San Diego, CA: Academic Press.

**Lazarus, R. S., Kanner, A. D., & Folkman, S.** (1980). Emotions: A cognitive-phenomenological analysis. In *Emotion: theory, research, and experience. Volume 1. Theories of emotion,* (ed. R. Plutchik & H. Kellerman), pp. 189–217. London: Academic Press.

Lehmann, A. C. (1997). Acquired mental representations in music performance: Anecdotal and preliminary empirical evidence. In *Does practice make perfect? Current theory and research on instrumental music practice*, (ed. H. Jørgenssen & A. C. Lehmann), pp. 141–63. Oslo, Norway: Norwegian State Academy of Music.

Ludwig, A. M. (1990). Altered states of consciousness. In *Altered states of consciousness*, (3rd edn) (ed. C. T. Tart), pp. 18–33. San Fransisco, CA: Harper.

McDonald, M. (1970). Transitional tunes and musical development. *Psychoanalytic Study of the Child*, 25, 503–20.

Manturzewska, M. (1990). A biographical study on the life span development of professional musicians. *Psychology of Music*, **18**, 112–39.

Maslow, A. (1968). *Toward a psychology of being*, (2nd edn). New York: Van Nostrand.

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Motte-Haber, H. de la (1984). Die Bedeutung der Motivation für den Instrumentalbericht [The significance of motivation in instrumental reports]. *Zeitschrift für Musikpädagogik*, 51, 51–4.

Nagel, J. J. (1987). *An examination of commitment to careers in music: Implications for alienation from vocational choice*. Unpublished doctoral dissertation, University of Michigan, Ann Arbor, MI.

Ohgushi, K. & Senju, M. (1987). How are the player's ideas conveyed to the audience? *Music Perception*, 4, 311–24.

Osborne, R. (1990). *Conversations with Karajan*. Oxford, UK: Oxford University Press.

Parncutt, R., Sloboda, J. A., & Clarke, E. F. (1997). Interdependence of right and left hands in sight-read, written, and rehearsed fingerings of piano music in parallel octaves. In *Proceedings of the Third Triennial ESCOM Conference*, (ed. A. Gabrielsson), pp. 702–5. Uppsala, Sweden: Department of Psychology, Uppsala University.

Persson, R. S. (1993). *The subjectivity of musical performance: An exploratory music-psychological real world enquiry into the determinants and education of musical reality*. Unpublished doctoral dissertation, School of Human and Health Sciences, Huddersfield University, Huddersfield, UK.

Persson, R. S. (1995). Musical reality: Exploring the subjective world of performers. In *Song and signification. Studies in music semiotics*, (ed. R. Monelle & C. T. Gray), pp. 58–63. Edinburgh, UK: Faculty of Music, University of Edinburgh.

Persson, R. S. (1996). Studying with a musical maestro: A case study of commonsense teaching in artistic training. *Creativity Research Journal*, 9, 33–46.

Persson, R. S. (2000). Survival of the fittest or the most talented? Deconstructing the myth of the musical maestro. *Journal of Secondary Gifted Education*, **xii**, 25–38.

Persson, R. S., Pratt. G., & Robson, C. (1996). Motivational and influential components of musical performance: A qualitative analysis. In *Fostering the growth of high ability: European perspectives*, (ed. A. J. Cropley & D. Dehn), pp. 287–301. Norwood, NJ: Ablex.

Persson, R. S. & Robson, C. (1995). The limits of experimentation: On researching music and musical settings. *Psychology of Music*, 23, 39–47.

Rapoport, E. (1996). Emotional expression code in opera and lied singing. *Journal of New Music Research*, 25, 109–49.

Renshaw, P. (1986). Towards the changing face of the conservatoire curriculum. *British Journal of Music Education*, 3, 79–90.

Rigg, M. G. (1940). Speed as a determination of musical mood. *Journal of Experimental Psychology*, 27, 566–71.

Robson, C. (1993). *Real world research. A resource for social scientists and practitioner-researchers.* Oxford, UK: Blackwell.

Rogers, C. R. (1967). *On becoming a person. A therapist's view of psychotherapy.* London: Constable.

Russ, S. W. (1993). *Affect and creativity. The role of affect and play in the creative process.* Hillsdale, NJ: Lawrence Erlbaum.

Ruud, E. (1997). *Musikk og identitet [Music and identity].* Oslo, Norway: Universitetsforlaget.

Seashore, C. E. (1938). *Psychology of music.* New York: Dover Publications.

Shore, R. E. (1990). Hypnosis and the concept of the general reality orientation. In *Altered states of consciousness,* (3rd edn) (ed. C. T. Tart), pp. 281–301. San Fransisco, CA: Harper.

Sloboda, J. A. (1991). Music structure and emotional response. *Psychology of Music,* **19**, 110–20.

Sloboda, J. A. (1998). Does music mean anything? *Musicae Scientiae,* 2, 21–32.

Stanislavski, C. (1988). *An actor prepares.* London: Methuen Drama. (Originally published in 1937.)

Stein, M. I. (1974). *Stimulating creativity: Individual procedures,* (Vol. 1). New York: Academic Press.

Stern, I. (1979). An interview with Isaac Stern. *The New York Times Magazine,* 23 December.

Waterman, M. (1996). Emotional responses to music. Implicit and explicit effects in listeners and performers. *Psychology of Music,* **24**, 53–67.

Watt, R. J. & Ash, R. L. (1998). A psychological investigation of meaning in music. *Musicae Scientiae,* 2, 33–54.

Wedin, L. (1972). A multidimensional study of the perceptual-emotional qualities in music. *Scandinavian Journal of Psychology,* **13**, 241–57.

Zajonc, R. B. (1980). Feeling and thinking: preferences need no inferences. *American Psychologist,* 35, 151–75.

Zuckerman, M. (1994). *Behavioral expressions and biosocial bases of sensation seeking.* Cambridge, UK: Cambridge University.

CHAPTER 13

# NEGATIVE EMOTIONS IN MUSIC MAKING: THE PROBLEM OF PERFORMANCE ANXIETY

ANDREW STEPTOE

The performance of music induces extraordinarily positive feelings in many players and singers. Participation in a successful and exciting performance often leads to feelings of euphoria, whether it be in a school concert, a piece sung by a village choir, or a public show in a prestigious concert hall. Here is the baritone Thomas Allen (interviewed in the *Independent on Sunday* in 1990) describing his feelings at the end of a performance of *Don Giovanni* in Japan: 'I could have walked over the house tops I was so high'. When the producer came backstage to congratulate him, Allen picked him up and carried him across the stage, 'Oh, I could have run up Mount Fuji!'

The same feelings of joy of creative music making are vividly expressed in a letter written by the 21-year-old Mozart to his father in November 1777. He was playing the organ during a Sunday service in the Court chapel in Mannheim in the presence of the ruler of the state:

> I entered during the Kyrie and played the end of it; and after the priest had intoned the Gloria, I played a cadenza. Because this was all so different from what they are accustomed to here, they all turned round . . . The people had enough to amuse them. Now and then was a pizzicato, and each time I just brushed the keys. I was in the best of spirits. Instead of the Benedictus, the organist has to play throughout. I took the subject to the Sanctus and treated it as fugue. Everybody stood there and stared. The pedal is different from ours; that put me off at first, but I soon understood how it worked. (Quoted in Anderson 1988, p. 370)

Unfortunately, musical performance may also induce negative emotions, including distress and anxiety, which in some individuals can approach extreme levels of terror, impair the quality of the performance, and have debilitating effects on professional and amateur careers.

This chapter discusses musical performance anxiety, outlining the psychological and medical literature on this topic, and touches on issues such as the definition and prevalence of musical performance anxiety, its causes and consequences, and methods of alleviating severe problems. Musical performance anxiety is a complex phenomenon. It is no accident that this chapter began with descriptions of positive feelings, because there are close links between negative and positive feelings in the musical performer.

Indeed, sensations of arousal experienced by most individuals prior to a performance may be disturbing to some, but facilitate a good performance in others. As will become apparent later in the chapter, it has proved difficult to demonstrate that physiological arousal is positively correlated with performance anxiety and that musicians with problems in this area are characterized by excessive arousal. Rather it is the interplay between cognitive processes and physiological activation that is central to musical performance anxiety.

Musical performance anxiety is often regarded as synonymous with stage fright, although stage fright is a phenomenon that occurs in non-musical performers such as actors and public speakers as well. I prefer the term musical performance anxiety for several reasons. Firstly, it refers specifically to the feelings experienced by musicians. Secondly, musical performance anxiety occurs in many settings, and not just on the stage. Stage fright has connotations of distress in front of large audiences, but musical performance anxiety may be elicited in quite intimate surroundings, such as a lesson or an audition. It depends on the evaluative nature of the situation, and not on the presence of an audience. Thirdly, the term 'fright' implies a sudden fear or alarm, while musical performance anxiety may be quite predictable and develop gradually over days prior to an important occasion. Finally, the term has implications for the way in which the music is played, and not just the fear that the performer experiences. Stage fright would be bad enough if it were only a feeling of intense fear. But a major concern for many musicians is that it will have an impact on the quality of the performance, and on their ability to carry out the task of playing or singing accurately and effectively. This feature marks musical performance anxiety out from many other types of situational fear.

The empirical literature on musical performance anxiety is sparse, and relatively few investigators have gone beyond describing the phenomenon into disentangling its elements so as to understand how it develops and is maintained (Brodsky 1996). This chapter will therefore draw on research into related phenomena where appropriate, including stage fright in actors and studies of test anxiety. Other potentially negative aspects of the performer's life will also be briefly discussed, since work stress among musicians appears to be related to performance anxiety.

## 13.1 Occurrence and experience of musical performance anxiety

Musical performance anxiety is not an all or nothing experience, but a graduated phenomenon. Different musicians encounter it to a different extent, and the same individual may experience varying degrees of performance anxiety on different occasions. Surveys of the occurrence of musical performance anxiety must therefore be treated cautiously, since they typically rely on somewhat arbitrary classifications. The frequency of performance anxiety is also important, since many more musicians will have suffered occasional bouts of severe anxiety than experience them on a regular basis. The best evidence for the occurrence of musical performance anxiety comes from surveys of musicians who are not specially selected on the basis of problems in this area. Obviously, studies of musicians seeking help for difficulties associated with performance anxiety will be biased, and this may lead therapists to overestimate the magnitude of the problem.

Varying estimates of the occurrence of musical performance anxiety have emerged from different surveys, and it is difficult to know whether these are due to genuine variations or differences in questioning. In probably the largest survey published to date, Fishbein and Middlestadt (1988) studied more than 2200 musicians from 48 orchestras in the USA. Nineteen per cent of women and 14 per cent of men indicated that stage fright was a severe problem. Another large survey involving 56 orchestras throughout the world was carried out on behalf of the Fédération Internationale des Musiciens (FIM) in 1997.[1] Seventy per cent of the 1639 respondents indicated that they sometimes experienced such intense anxiety before a performance that it impaired their playing. For 16 per cent, this happened more than once a week. Fity-nine per cent of symphony orchestra players in a Dutch survey claimed to be acquainted with performance anxiety from their own experience, though only 21 per cent described it as intense or very intense (Van Kemenade *et al.* 1995). A survey of more than 300 students in a faculty from the University of Iowa School of Medicine concluded that 21 per cent experienced marked performance anxiety (Wesner *et al.* 1990). Fifty-eight per cent of the members of the Vienna Symphony Orchestra studied by Schulz (1981) indicated that their 'nervous stress' during concerts was high or very high, while 24 per cent also complained of high levels of tension before performances. These surveys indicate that musical performance anxiety is a serious problem for 15–25 per cent of musicians. Some studies have reported higher levels, probably related to differences in exactly how musical performance anxiety was defined, rather than genuine differences in prevalence (Bartel & Thompson 1995; Marchant-Haycox & Wilson 1992). A pattern of slightly higher levels among women than men is seen in several studies, as it is for other forms of self-reported distress (Mirowsky & Ross 1995).

There have been few systematic comparisons of different types of musician within the same study. However, the author and a colleague (Steptoe & Fidler 1987) compared responses from members of the Royal Philharmonic Orchestra, the London Philharmonic Orchestra, and an amateur orchestra, and orchestral instrumental students at the College of Music, using an adaptation of the state scale from the state-trait anxiety inventory (Spielberger *et al.* 1970). The mean scores were significantly higher among students, and were lowest in the professional musicians in the world class orchestras. Does this mean that performance anxiety declines with experience? Not necessarily, because it is possible that more anxious individuals hoping to become orchestral musicians drop out of the profession at an early stage. Alternatively, it is possible that more anxious musicians fail to be employed by prestigious orchestras, leaving a selected sample of less anxious individuals. Unfortunately, no longitudinal studies have yet been conducted that allow the different explanations of cross-sectional results to be teased out.

There are similar difficulties in understanding associations between musical performance anxiety and age. One might expect performance anxiety to decline with age as musicians become more experienced and familiar with different audiences and challenges. On the other hand, success comes with responsibility, and the acclaimed

[1] The FIM survey has not been published in detail. I am grateful to the British Association for Performing Arts Medicine for giving me access to the data.

mature performer may perceive additional burdens in the high public expectations on them. The violinist Nathan Milstein (quoted in the *Guardian* in 1993) recalled how carefree he was when performing in his youth, playing cards or sightseeing during the day before an evening recital. Later in life he was unable to eat lunch and remained preoccupied for several hours before a performance. Highly successful popular musicians and classical musicians, such as Arthur Rubinstein and Luciano Pavarotti, have admitted to extreme nervousness and distress prior to concerts. Perhaps the most celebrated case is that of the great actor Laurence Olivier, who suddenly succumbed to severe stage fright for several years at a late phase of his career (Olivier 1982).

### 13.1.1 The phenomenon of performance anxiety

Musicians who suffer from performance anxiety describe a wide range of experiences, including fear and terror, trembling, difficulty concentrating, disturbances of breathing, and dry mouth. There is a strong belief among musicians that performance anxiety impairs their playing or singing, and this is corroborated in studies involving objective evaluations. For example, Craske and Craig (1984) had independent judges evaluate the performance quality of piano students divided into anxious and non-anxious groups. Judgments were made from audiotapes recorded when the student played alone, or in the presence of an evaluative audience. Subjective ratings of distress were also collected. Some results are illustrated in Fig. 13.1. It can be seen that the judges' ratings of performance quality were significantly impaired in the audience condition in the anxious group, but improved in the non-anxious group. Subjective distress increased in both groups when an audience was present, but the increase was much larger in the performance anxiety group.

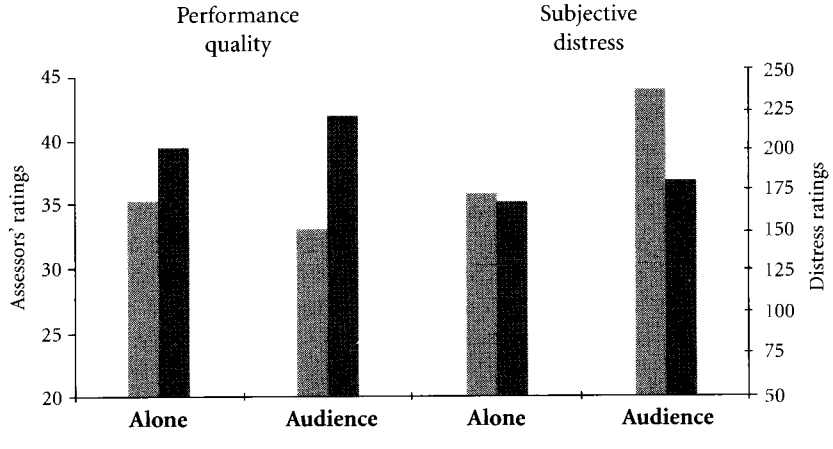



**Figure 13.1** Performance quality and subjective distress in anxious (light bars) and non-anxious (dark bars) pianists playing alone or in front of an evaluative audience. Performance quality (left-hand axis) was judged by assessors blind to the pianists' group, and higher scores indicate superior quality. Distress (right-hand axis) was assessed by combining ratings obtained before and immediately after playing, with higher scores indicating greater distress. (From Craske & Craig 1984.)

Musical performance anxiety is a complex phenomenon. By analogy with other work on stress and emotion (Steptoe 1998), it may be useful to separate four distinct elements. These four components are listed in Table 13.1, along with various examples of the ways in which they are manifested. The primary component is affect or feeling, which forms the central experience of performance anxiety for many musicians. It should also be recognized in this context that the moods and emotions evoked by music are not confined to the audience, but may also affect the performer. However, there are other important aspects as well, including cognitive responses or disturbances in information processing, changes in behaviour, and physiological responses.

**Table 13.1** The four components of musical performance anxiety

| | |
|---|---|
| *Affect* | Feelings of anxiety, tension, apprehension, dread or panic |
| *Cognition* | Loss of concentration, heightened distractibility, memory failure, maladaptive cognitions, misreading of the musical score, etc. |
| *Behaviour* | Tremor, trembling, difficulty in maintaining posture and moving naturally, failures of technique |
| *Physiology* | Disturbances in breathing pattern, perspiration, inhibition of salivation (dry mouth), high heart rate, the release of hormones such as adrenaline (epinephrine) and cortisol, gastrointestinal disturbances |

The 1997 FIM survey illustrates the importance of these different elements. Physiological symptoms were frequent before and during performances, with rapid heart beat (67 per cent), sweating hands (56 per cent), and muscle tension (56 per cent) being the most prominent. At the same time, 49 per cent complained that they often lost concentration, and 46 per cent that they experienced trembling and shaking prior to playing. These same four elements are common to many modern conceptualizations of the stress response, which are essentially offshoots of the three system model of fear formulated by Lang and others (Hugdahl 1981).

The various manifestations of musical performance anxiety do not necessarily occur together, and correlations between responses in the different domains are variable (Craske & Craig 1984; Fredrikson & Gunnarsson 1992). One person may show impairment of movement or posture without feeling very upset, while another may experience their principal difficulties in cognitive disturbance. Craske and Craig (1984) showed that concordance between components was greater in more anxious performers, but the correlations between physiological and subjective measures still did not exceed 0.38, suggesting that marked discrepancies remain. Understanding the interrelationships between these domains helps to delineate the core components of performance anxiety.

The framework outlined here and elaborated in the following sections is based on the view that musical performance anxiety is a form of situational stress response. That is, it is similar in some respects to problems such as social phobia or academic examination stress. It should be recognized that this model differs from some approaches that conceptualize performance anxiety as a unique problem related to psychological issues stimulated by performing in front of others. This latter perspective has derived primarily from the psychodynamic tradition. For example, Gabbard (1997) has argued that stage fright can emerge from shame related to genital inadequacy, or impulses that produce feelings of guilt in the unconscious in triumphing over Oedipal and sibling rivals. I

have little sympathy with such views, which lack objective verification, and would argue that bringing musical performance anxiety into the mainstream of stress research has major advantages for understanding issues such as the role of cognitive processes or the relations between physiological arousal and performance quality, while highlighting ways in which performance anxiety can be effectively managed.

## 13.2 Physiological arousal during musical performance

Perhaps the simplest model of musical performance anxiety is that it is a consequence of heightened physiological arousal. Certainly, the perception of physiological arousal figures prominently in the experience of many anxious musicians, with symptoms such as palpitations, breathlessness, and perspiration that are indicative of sympathetic nervous system activation. Subjectively, many anxious musicians report symptoms of *hyperventilation*—the pattern of dizziness, tremor, and high heart rate due to overbreathing (and subsequent reductions in arterial carbon dioxide levels) that is characteristic of patients with panic attacks (Widmer *et al.* 1997).

The physiological arousal induced by performing in public has also been established objectively. Two types of study have been carried out: (a) those in which physiological monitoring takes place under conditions specially devised and controlled by the investigator; and (b) naturalistic studies in which musicians are monitored during live performances. Thus, in the study by Craske and Craig (1984) described earlier, heart rate was higher during performances in front of an audience compared with playing alone. Fredrikson and Gunnarsson (1992) studied 19 music students, and reported that adrenaline (epinephrine) increased by an average of 139 per cent and cortisol by 66 per cent with performance in front of an audience compared with playing alone.

The classic investigation of physiological activation during live performance is the work by Haider and Groll-Knapp with the Vienna Symphony Orchestra, first published in German in 1971 and in English in 1981. The heart rate and electroencephalogram (EEG) were monitored during rehearsals and live concerts from 24 musicians. String and woodwind players showed similar responses, with a mean heart rate of 93.9 bpm during concerts, reaching an average maximum of 113 bpm. Since the resting heart rate of these musicians was 66.6 bpm, this means that heart rate increases averaged 27.3 bpm. Interestingly, the heart rate difference between rehearsals and concerts was more than 8 bpm, even though the same pieces were being performed. These substantial increases in heart rate confirmed an earlier study of musicians in concert (Schmale & Schmidtke 1965) and the more recent investigation of members of the BBC Symphony Orchestra (Mulcahy *et al.* 1990). Haider and Groll-Knapp (1981) also used EEG recordings to show high levels of cortical activation during concerts.

In the light of these findings, one possibility is that musical performance anxiety is a consequence of excessive physiological activation. The 'inverted-U' association between tension and performance is a well-established phenomenon in the experimental laboratory, and suggests that information processing and behavioural performance improves with arousal to an intermediate stage, but then deteriorates as physiological arousal rises beyond the optimum (Duffy 1962; Näätänen 1973). However, few studies have documented this pattern in performing musicians. Some evidence for such an association

emerged from an interview study I carried out with young professional opera singers (Steptoe 1983). Participants were asked to rank the quality of their performance in five settings: lesson, private practice, audition, dress rehearsal, and public performance. Then, at a different stage of the interview, ratings of emotional tension in these situations were obtained on a nine-point scale ranging from 1 = *extreme tension* to 9 = *extreme relaxation.* I then averaged the emotional tension ratings for each situation, and plotted them against the estimates of performance quality. Figure 13.2 depicts the interrelationship between these measures. It can be seen that the best singing was produced at an intermediate level of tension (in public performances), while the situations associated with greater tension (dress rehearsal and audition) were those in which performance was impaired. The results of this study are consistent with the inverted-U pattern, though they are limited by the use of self-reported measures.

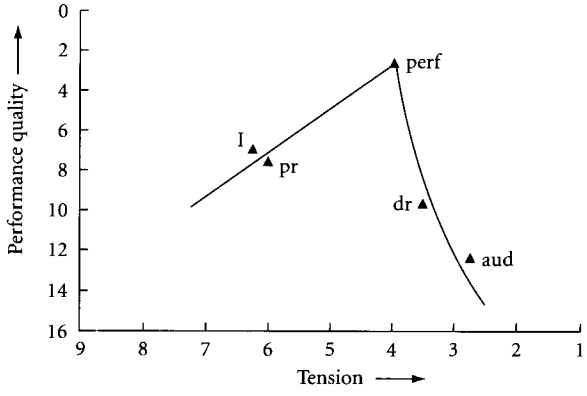



**Figure 13.2** The association between self-rated tension and quality of performance in a sample of opera singers, assessed in five situations: private practice (pr), a lesson (l), public performance (perf), dress rehearsal (dr), and in audition (aud). (For details, see Steptoe 1983.)

When it comes to objective physiological measures, there is much less evidence that musical performance anxiety is associated with excessive arousal. In fact, some studies have shown similar levels of physiological activation irrespective of musical performance anxiety. Thus, Abel and Larkin (1990) show that blood pressure and heart rate increased prior to playing in public, but that responses were not significantly correlated with performance anxiety. Similar levels of heart rate and sweat gland activity response in individuals with high and low levels of anxiety were recorded by Craske and Craig (1984), and adrenaline and cortisol response by Fredrikson and Gunnarsson (1992). The only evidence for excessive arousal to emerge thus far is the observation by Fredrikson and Gunnarsson that anxious musicians showed a mean 39.2 bpm increase in heart rate, compared with 22.6 bpm in the less anxious. It is also possible that musical performance anxiety is manifest physiologically, not in general measures of arousal such as heart rate, but in more specific correlates of breathing disturbance associated with hyperventilation (Widmer *et al.* 1997). An interesting result came from the heart-rate-monitoring studies of the BBC Symphony Orchestra (Mulcahy *et al.* 1990). It was found that the increase in heart rate during concerts was paralleled in measures

obtained from managerial and technical staff for the orchestra; even though these individuals were not actually playing music, they responded just as much at a physiological level.

These findings suggest that while physiological arousal may be a necessary condition for musical performance anxiety, it is not a sufficient explanation. Other factors are involved that relate to the cognitive responses to performing, and to perceptions of physiological arousal. However, the results of these objective studies of physiology are also relevant to the issue of the apparently beneficial or facilitatory effects of performance anxiety.

One of the statements that musicians often make about playing in public is that they 'need' stage fright, and that it helps them to perform at the optimum level. This facilitating effect is important, since it may lead musicians not to seek ways of alleviating performance anxiety, because they fear that the quality of their playing or singing may suffer. Some years ago, I interviewed one young professional opera singer whose concern was that she did not get excited enough when performing. She was worried that her lack of 'stress' may have had an adverse effect, giving the impression of little involvement in the music. Her subsequent extremely successful career has given the lie to this fear.

The notion that arousal is facilitative can be understood within a curvilinear model of performance (see Fig. 13.2). However, I would argue that this is not because performance anxiety is beneficial, but because physiological arousal is necessary for a good performance. When musicians say that performance anxiety, or stage fright, is helpful, they may be mislabelling physiological arousal as fear, rather than construing it as an independent entity. Physiological activation may be required for a good performance, but not the negative feelings associated with playing or singing in public. Such empirical work as has been done suggests that facilitating and debilitating aspects of performance anxiety are negatively correlated (Mor *et al.* 1995). The most parsimonious explanation is that the same feelings of physiological arousal are experienced, but while one performer finds these energizing, another perceives them as signs of impending disaster.

## 13.3 Cognitive factors in musical performance anxiety

The musician playing in public tends to be occupied with task-orientated thoughts. That is, the cognitive apparatus is taken up with thoughts related to effective performance. For the orchestral musician, this will include being sensitive to tempo and dynamic directions from the conductor, the volume of sound in relation to the rest of the orchestra, intonation, the particular challenges of the passage, and so on. A member of a rock band will be alert to the way in which the music is being played by colleagues, to improvisations by others, and to maintaining a regular beat. The singer or instrumental soloist often has added cognitive demands related to remembering the music, the actions on the stage, and adjusting performance to the acoustics of the venue.

A characteristic feature of musical performance anxiety is disruption of task-orientated cognitions. Several types of disruption have been identified. One common problem is *catastrophizing*, or exaggerating in imagination the likelihood and effects of negative events during the performance. For example, singers may find that their breath does not hold out in the way they hoped, or the instrumentalist makes a small error of

intonation. Musicians with musical performance anxiety tend to believe such occurrences will ruin the entire performance, and will lead to loss of control over the situation. Wolfe (1989) observed that anxious musicians were convinced that a small mistake at the beginning of a performance upset them so much that they tended to make even more simple mistakes later on. Catastrophizing also takes the form of fear of collapse, and the concern that the performer will be sick or faint on stage. We have found in both musicians and student actors that there are strong positive associations between catastrophizing and performance anxiety (Steptoe & Fidler 1987; Steptoe *et al.* 1995).

A second form of distraction comes from preoccupation with evaluation by others. Lehrer and co-workers (1990) carried out a factor analysis of thoughts and worries related to musical performance. Various factors emerged, including worry about anxiety and fear of distraction. In addition, concern about the reactions of other people (ranging from close family to critics) was a prominent feature, and was positively correlated with anxiety. By contrast, we have found that more realistic attitudes, such as the belief that the audience wants the performer to play well and will make allowance for slips, are associated with adaptive levels of moderate arousal, and not with severe performance anxiety (Steptoe & Fidler 1987). Concern about the reactions of others may be coupled with very high internal standards and a perfectionist attitude. An analysis of perfectionism in professional performers was carried out by Mor and colleagues (1995). The perfectionism scale included items such as 'one of my goals is to be perfect in everything I do', and 'the people around me expect me to succeed in everything I do'. It was found that perfectionism scores were positively correlated with debilitating performance anxiety, and that more anxious performers also experienced a lower sense of personal control in their lives.

Cognitive responses of this type tend to divert the musician from concentrating on the task itself, thus increasing the risk of errors in performance. Heightened perception of physical signs also contributes; anxious performers may interpret muscle tension or a rapidly beating heart not as a sign of excitement and involvement, but as signalling loss of control and impending collapse. They may therefore become further distracted by trying to keep themselves physically calm. Thoughts about failure during performance have been shown to predict playing below the standards of which the individual is capable of in other situations (Kivimäki 1995).

## 13.4 Other stresses in the musician's life

Negative emotional experiences among musicians are not restricted to performance anxiety. The musician's life, both during training and their subsequent professional career, is subject to other sources of stress, just as is the case with other types of job. Much of the work on these problems has been carried out in the context of occupational health research.

Parasuraman and Purohit (2000) built on the work of Schulz (1981) to identify five major sources of stress among orchestral musicians. These are issues related to the work environment (air quality, humidity, seating comfort, readability of the musical score, etc.), social tensions (such as conflict within the orchestra, feelings of being undervalued), musical performance anxiety, problems related to artistic integrity (such as

unhappiness about how the music is played, being subordinate to the will of a conductor), and concerns about the technical difficulty of the music. They found that social tension, problems of artistic integrity, and poor working conditions all contributed to job dissatisfaction among professional orchestral musicians. To these might be added worries about job security and the financial viability of music making, and boredom arising from playing the same pieces over and over again. In our own work, we found that the principle sources of stress for orchestral musicians were separation from family, irregular working hours, the monotony of rehearsals, travelling, and uncertainty about regular employment (Steptoe 1989).

The 1997 FIM survey asked orchestral musicians to rate more than 40 potential sources of stress. The ten items that were endorsed as being moderately or severely stressful by the largest proportion of players are listed in Table 13.2. It can be seen that the most common sources of stress for professional orchestral musicians were problems with conductors, conflict with fellow musicians, work environment problems, unsatisfactory rehearsal arrangements, and aspects of playing. However, set against these difficulties are the positive feelings of interest and fulfilment deriving from the sheer pleasure of playing in an orchestra, the variety in the job, and the excitement of performing to audiences.

Table 13.2 Stressful aspects of the musician's life.

| | Proportion of respondents who gave a rating of moderate or severe stress |
|---|---|
| Working with a conductor who saps your confidence | 73% |
| Playing an orchestral solo | 73% |
| Illegible music | 65% |
| Disorganized rehearsal time | 65% |
| Problems with your instrument | 63% |
| Making a mistake when performing | 63% |
| Incompetent conductor | 61% |
| Incompatible desk partner | 61% |
| Having medical problems that affect your work | 57% |
| Playing in a cold hall | 55% |

Somewhat similar problems are reported by popular musicians playing jazz and rock (Cooper & Wills 1989). The major concerns reported by a sample of popular musicians were the uncertainty of employment, periods of work overload interspersed with underload coming from being obliged to play routine music in unexciting venues, conflicts with fellow musicians, and the lack of clear structures of career development. The concerns of music students run along the same lines (Dews & Williams 1989).

There are, of course, stressful elements in any work environment, and it is not clear whether the problems of musicians are worse than those experienced by other people. Kivimäki and Jokinen (1994) compared orchestral musicians with industrial and clerical workers, and with people in the health care professions. Average levels of job satisfaction were higher among orchestral musicians than all the other groups, primarily because they enjoyed the variety of the work and the skills they could use in their jobs. However, their sense of autonomy or personal control was relatively low. This same

pattern emerged from a comparison of six occupations from a biomedical perspective carried out by Theorell and co-workers (1990). Orchestral musicians felt that they had less autonomy over their work than did all the other occupational groups, including aircraft mechanics, waiters in restaurants, and physicians. Low control at work has been highlighted over recent years as a major source of stress and as a contributor to health risk (Steptoe & Appels 1989).

What is also clear is that these other stresses in the musician's life are not entirely separate from musical performance anxiety. In our own work, we found that musicians with high performance anxiety were more likely to be concerned about other negative aspects of their careers, such as poor financial rewards, social conflict with other professionals, the competitive nature of the job, and uncertainty about employment (Steptoe & Fidler 1987). It is possible that performance anxiety serves to accentuate the individual's perception of other stresses in music making, while these same difficulties may be ignored by musicians who are not troubled with nerves and are able to enjoy their playing.

## 13.5 Coping with musical performance anxiety

Musicians engage in a variety of informal coping efforts in their attempt to manage performance anxiety. Common methods of coping immediately prior to a performance include deep breathing and muscle relaxation, meditation, immersion in the music, reading, and distraction (Steptoe 1989; Wolfe 1990). Relatively small but none the less worrying numbers report using sedatives and alcohol, while a significant proportion of professional musicians use β-blockers to help them cope with the physical manifestations of performance anxiety. Longer-term methods of coping include thorough preparation and practice, and checking on all aspects of the performance venue (lighting, access, acoustics, etc.).

An apparently paradoxical finding in this literature is that coping has been positively correlated with musical performance anxiety. Lehrer *et al.* (1990) showed that more anxious musicians engaged in more coping efforts, while Steptoe and Fidler (1987) found that anxious musicians were more likely to meditate. The use of distraction techniques has been positively correlated with performance anxiety in student actors as well (Steptoe *et al.* 1995), although Wolfe (1990) observed that musicians who use emotion-focused coping techniques felt greater confidence and competence. The explanation for these associations is not clear. One possibility is that anxious musicians are more likely to engage in deliberate efforts at coping than their less anxious colleagues. The methods of coping may be useful, even though anxiety levels still remain somewhat elevated. A second possibility is that some types of coping are effective, while others are maladaptive. That is, musicians may engage in activities that they think are helpful which do not actually alleviate performance anxiety. Such a pattern is difficult to identify, because methods of coping may be helpful on some occasions but not others. When discussing these issues with musicians, it is striking that efforts at coping may become rather rigid.

A good example is the ritual that some instrumentalists and singers carry out on the day of an important performance. They may, for example, remain silent for several hours, feel they have to eat particular foods or wear certain colours, or insist on arriving at the venue at a specific time. These ritual patterns may develop over months or years

as the musician comes to believe that he or she performs much better when these routines are fulfilled. The problem arises when these coping routines are disrupted. Perhaps an extra rehearsal is unexpectedly called, or traffic problems lead to a change in travel plans. Under these circumstances, the musician who has learnt to rely on an inflexible coping routine may be more disturbed by these occurrences than other performers, leaving them worse off than before.

## 13.6 Methods of treatment and support

Performance anxiety is not the only serious problem related to health and well-being experienced by musicians. High numbers report pain in their shoulders, backs, arms, and fingers, due partly to overuse and partly to excessive tension. Many wind players experience difficulties in lip control after many years, related to tenderness and inflammation. None the less, performance anxiety is so distressing and intense for some musicians that it seriously impairs their lives and careers. There are a range of popular and self-help guides on managing stage fright, some of which are intended for all types of public presentation (e.g. Ashley 1996; Desberg 1996). The most useful book designed for musicians that I have read is Salmon and Meyer (1998).

Professional support for musical performance anxiety provides two basic approaches: (a) pharmacotherapy; and (b) psychotherapies based on relaxation or cognitive behavioural stress management. Anxiolytic drugs, such as benzodiazepines, have been used to some extent to control performance anxiety, despite side effects and the risk of drug dependence (Lederman 1999). The non-benzodiazepine buspirone has been found to have no beneficial effect on performance anxiety in a small controlled trial by Clark and Agras (1991). Much more popular have been β-adrenergic blockers, originally developed for the treatment of cardiovascular disease. These drugs inhibit the peripheral physiological manifestations of sympathetic nervous system activation, such as rapid heart rate and excessive perspiration. They have been welcomed as peripherally acting drugs that modify the somatic elements of performance anxiety without dulling mental processes.

It is clear that many professional musicians use β-blockers, often without medical supervision. Fishbein and Middlestadt (1988) reported that 27 per cent of musicians in their survey had used β-blockers to help with stage fright, and 70 per cent of these obtained the drugs from informal sources. In the 1997 FIM survey, 20 per cent of respondents had used these drugs over the past year. Musicians who do use β-blockers certainly experience favourable effects, with a reduction in tremor and a greater sense of calm prior to and during performance (Fishbein & Middlestadt 1988).

The scientific evidence for benefits is more equivocal. Liden and Gottfries (1974) compared β-blockers and placebo in a double-blind trial. Musicians' ratings of their own symptoms decreased under both conditions, with no significant differences between them. James and Savage (1984) found no difference in mood ratings prior to performance in β-blockade, benzodiazepine, and placebo groups. However, β-blockers did have superior effects on self-ratings of bow shake and dynamics in string players, and on one of eight ratings of performance quality by professional observers. In another placebo-controlled trial in Switzerland, string players played both with and without an

audience (Neftel *et al.* 1982). Tape recordings of their playing were analysed objectively using a sonograph, which allowed subtle aspects of performance—such as the regularity of trills and vibrato of movements—to be assessed. The heart rate increase when playing in front of an audience was effectively blocked in the active drug condition, and the increase in output of adrenaline was also attenuated by β-blockade. There was evidence from the sonographic analysis that fine motor skills were impaired during performance in front of an audience in the placebo condition, but that these aspects were improved by β-blockade. At the same time, there were no significant differences in self-reported anxiety.

The impression from these and other studies is that β-blockers do indeed reduce physiological activation. Aspects of performance that are critically dependent on fine motor control, such as bow movements and vibrato, may therefore benefit. The impact on anxiety, concentration, and other aspects is less clear. The unsupervised use of such potent medications cannot be recommended.

Psychotherapeutic approaches to musical performance anxiety cover the whole range of talk therapies, only a few of which have been evaluated systematically. Psychodynamic therapies such as psychoanalysis have been recommended (Nagel 1993), but there is no evidence involving controlled trials or standardized measurement concerning their efficacy. The best evidence is for cognitive-behavioural programmes based on the model of stress outlined in this chapter.

There are four core elements of interventions that tend to be effective, either alone or in combination. The first is for the musician to acquire self-assessment skills. It is essential that performers themselves become aware of the sensations, thoughts, and reactions to audiences that they experience when they are playing in public. Without this self-awareness, progress in developing coping skills is likely to be limited. Secondly, most successful programmes include some method of helping musicians reduce or control physiological arousal. This might involve relaxation training, although *biofeedback* (providing participants with information about their physiological function) has been used as well (Nagel *et al.* 1989; Niemann *et al.* 1993; Sweeney & Horan 1982). Music-based relaxation has also been shown to be helpful in this context (Brodsky & Sloboda 1997). Thirdly, cognitive methods are used to challenge the negative thoughts that spring into the anxious musician's mind. For example, Kendrick and co-workers (1982) first elicited negative or distracting thoughts from participants, then taught them to substitute task-orientated and positive thoughts through role play, imagined performance sequences, and verbalizing thoughts out loud. The fourth ingredient is practice at using these skills in the performance setting. Controlled exposure to threatening stimuli is one of the key elements of anxiety management (Noyes & Hoehn-Saric 1998), and the same is true for musical performance. Coping skills cannot be imposed prescriptively since the musician needs to experiment with different approaches before deciding what is most useful.

In a number of randomized trials, methods that involve these elements have been shown to be superior to comparison conditions in terms of reducing anxiety and increasing self-reported confidence, and there is some evidence for favourable effects on objective assessments of comportment and playing (Brodsky & Sloboda 1997; Kendrick *et al.* 1982; Nagel *et al.* 1989; Valentine *et al.* 1995). An additional important factor is the

application of systematic learning and practice methods, such as those described by Reubart (1985) and Salmon and Meyer (1998).

Despite these benefits, it is clear that the average musician does not seek professional help for performance anxiety. In the 1997 FIM survey, for example, only 11 per cent had consulted a psychologist, psychiatrist, or counsellor over the past year, and many of these consultations were probably not for performance anxiety. This represents only a fraction of musicians who complained of performance anxiety. Fishbein and Middlestadt (1988) found that only 25 per cent of musicians in their large survey who experienced severe stage fright had sought psychological or psychiatric counselling. Reluctance to seek help may stem from three factors. The first is that consulting a psychologist or psychiatrist has connotations of mental illness, and musicians quite reasonably do not see their problems in this way. Secondly, there is ambivalence about arousal stemming from the belief that tension is beneficial for a good performance. This may lead to fears that treatment would have an adverse effect on playing. Thirdly, musicians may not be aware of the availability of effective interventions. More positive attitudes may evolve only gradually. In the meantime, advice and strategies for coping with performance anxiety could usefully be incorporated into the training of musicians in conservatories and colleges of music. Self-help materials based on proven techniques also provide musicians with alternatives to formal support (Salmon & Mayer 1998).

## 13.7 Conclusion

The negative emotions elicited in musicians during performance can be highly aversive, and may lead to the curtailment of promising careers in professionals, and the spoiling of an otherwise enjoyable activity for amateurs. Performance anxiety can be placed within the broader context of stress-related problems, and is not a unique phenomenon limited to actors and musicians. Misconceptions exist about the facilitating role of arousal and the adverse effect of performance anxiety. Promising treatment methods have been devised that are based on cognitive-behavioural stress management. The incorporation of the principles underlying these methods into the training of musicians would almost certainly be very beneficial.

## References

**Abel, J. L. & Larkin, K. T.** (1990). Anticipation of performance among musicians: Physiological arousal, confidence, and state-anxiety. *Psychology of Music*, **18**, 171–82.

**Anderson, E.** (ed.) (1988). *The letters of Mozart and his family*, (3rd edn). London: MacMillan.

**Ashley, J.** (1996). *Overcoming stage-fright in everyday life*. New York: Clarkson Potter.

**Bartel, L. R. & Thompson, E. G.** (1995). Coping with performance stress: A study of professional orchestra musicians in Canada. *Quarterly Journal of Music Teaching and Learning*, 5, 70–8.

**Brodsky, W.** (1996). Music performance anxiety reconceptualized. *Medical Problems of Performing Artists*, 11, 88–98.

**Brodsky, W. & Sloboda, J. A.** (1997). Clinical trial of a music generated vibrotactile therapeutic environment for musicians. *Journal of Music Therapy*, **34**, 2–32.

Clark, D. B. & Agras, W. S. (1991). The assessment and treatment of performance anxiety in musicians. *American Journal of Psychiatry,* **148**, 598–605.

Cooper, C. L. & Wills, G. I. D, (1989). Popular musicians under pressure. *Psychology of Music,* 17, 22–36.

Craske, M. G. & Craig, K. D. (1984). Musical performance anxiety: The three-systems model and self-efficacy theory. *Behavior Research and Therapy,* 22, 267–80.

Desberg, P. (1996). *No more butterflies.* Oakland, CA: New Harbinger.

Dews, C. L. B. & Williams, M. S. (1989). Student musicians personality styles, stresses, and coping patterns. *Psychology of Music,* 17, 37–47.

Duffy, E. (1962). *Activation and behavior.* New York: Wiley.

Fishbein, M. & Middlestadt, S. (1988). Medical problems among ICSOM musicians: Overview of a national survey. *Medical Problems of Performing Artists,* 3, 1–8.

Fredrikson, M. & Gunnarsson, R. (1992). Psychobiology of stage fright: The effect of public performance on neuroendocrine, cardiovascular and subjective ratings. *Biological Psychology,* 33, 51–62.

Gabbard, J. O. (1997). The vicissitudes of shame in stage fright. In *Work and its inhibitions: Psychoanalytic essays,* (ed. C. W. Socarides & S. Kramer), pp.209–20. Madison, CT: International Universities Press.

Haider, M. & Groll-Knapp, E. (1981). Psychophysiological investigation into the stress experienced by musicians in a symphony orchestra. In *Stress and music: Medical, psychological, sociological, and legal strain factors in a symphony orchestra musician's profession,* (ed. M. Piperek), pp. 15–34. Vienna: Wilhelm Braumüller (originally published 1971).

Hugdahl, K. (1981). The three-systems-model of fear and emotion: A critical examination. *Behaviour Research and Therapy,* **19**, 75–85.

James, I. M. & Savage, I. (1984). Beneficial effects of nadolol on anxiety induced disturbance of performance in musicians. A comparison with diazepam and placebo. *American Heart Journal,* **108**, 1150–5.

Kendrick, M. J., Craig, K. D., Lawson, D. W., & Davidson, P. O. (1982). Cognitive and behavioral therapy for musical-performance anxiety. *Journal of Consulting and Clinical Psychology,* **50**, 353–62.

Kivimäki, M. (1995). Test anxiety, below-capacity performance, and poor test performance: Intrasubject approach with violin students. *Personality and Individual Differences,* **18**, 47–55.

Kivimäki, M. & Jokinen, M. (1994). Job perceptions and well-being among symphony orchestra musicians: A comparison with other occupational groups. *Medical Problems of Performing Artists,* 9, 73–6.

Lederman, R. J. (1999). Medical treatment of performance anxiety: A statement in favour. *Medical Problems of Performing Artists,* **14**, 117–121.

Lehrer, P. M., Goldman, N. S., & Strommen, E. F. (1990). A principal components assessment of performance anxiety among musicians. *Medical Problems of Performing Artists,* 5, 12–18.

Liden, S. & Gottfries, C-G. (1974). β-blocking agents in the treatment of catecholamine-induced symptoms in musicians. *Lancet,* 2, 529.

Marchant-Haycox, S. E. & Wilson, G. D. (1992). Personality and stress in performing artists. *Personality and Individual Differences,* **13**, 1061–8.

Mirowksy, J. & Ross, C. E. (1995). Sex differences in distress: Real or artifact? *American Sociological Review,* **60**, 449–68.

Mor, S., Day, H. I., Flett, G. L., & Hewitt, P. L. (1995). Perfectionism, control and components of performance anxiety in professional artists. *Cognitive Therapy and Research,* 19, 207–25.

Mulcahy, D., Keegan, J., Fingret, A., Wright, C., Park, A., Sparrow, J. *et al.* (1990). Circadian variation of heart rate is affected by environment: A study of continuous electrocardiographic monitoring in members of a symphony orchestra. *British Heart Journal,* 64, 388–92.

Näätänen, R. (1973). The inverted-U relationship between activation and performance: A critical review. In *Attention and performance IV,* (ed. S. Kornblum), pp. 155–174. London: Academic Press.

Nagel, J. J. (1993). Stage-fright in musicians: A psychodynamic perspective. *Bulletin of the Menninger Clinic,* 57, 492–503.

Nagel, J. J., Himle, D. P., & Papsdorf, J. D. (1989). Cognitive-behavioural treatment of musical performance anxiety. *Psychology of Music,* 17, 12–21.

Neftel, K. A., Adler, R. H., Kappell, L., Rossi, M., Dolder, M., Kaser, H. E. *et al.* (1982). Stage-fright in musicians: A model illustrating the effect of beta blockers. *Psychosomatic Medicine,* 44, 461–9.

Niemann, B. K., Pratt R. R., & Maughan, M. L. (1993). Biofeedback training, selected coping strategies, and music relaxation interventions to reduce debilitating musical performance anxiety. *International Journal of Arts Medicine,* 2, 7–15.

Noyes, R. & Hoehn-Saric, R. (1998). *The anxiety disorders.* New York: Cambridge University Press.

Olivier, L. (1982). *Confessions of an actor.* London: Routledge.

Parasuraman, S. & Purohit, Y. S. (2000). Distress and boredom among orchestra musicians: The two faces of stress. *Journal of Occupational Health Psychology,* 5, 74–83.

Reubart, D. (1985). *Anxiety and musical performance.* New York: Da Capo Press.

Salmon, P. G. & Meyer, R. G. (1998). *Notes from the green room: Coping with stress and anxiety in musical performance.* San Francisco: Jossey Bass.

Schmale, H. & Schmidtke, H. (1965). *Untersuchungen uber die Psychophysiologische Belastung von Musikern in ein Kulturorchestern.* Mainz: Schott.

Schulz, W. (1981). Analysis of a symphony orchestra. In *Stress and music: Medical, psychological, sociological, and legal strain factors in a symphony orchestra musician's profession,* (ed. M. Piperek), pp. 35–56. Vienna: Wilhelm Braumüller (originally published 1971).

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the state-trait anxiety inventory.* Palo Alto, CA: Consulting Psychologists Press.

Steptoe, A. (1983). The relationship between tension and the quality of musical performance. *Journal of the International Society for the Study of Tension in Performance,* 1, 12–22.

Steptoe, A. (1989). Stress, coping and stage fright in professional musicians. *Psychology of Music,* 17, 3–11.

Steptoe, A. (1998). Psychophysiological bases of disease. In *Comprehensive clinical psychology. Volume 8: Health psychology,* (ed. D. W. Johnston & M. Johnston), pp. 39–78. New York: Elsevier Science.

Steptoe, A. & Appels, A. (ed.) (1989). *Stress, personal control and health.* Chichester, UK: Wiley.

Steptoe, A. & Fidler, H. (1987). Stage fright in orchestral musicians: A study of cognitive and behavioural strategies in performance anxiety. *British Journal of Psychology,* 78, 241–9.

Steptoe, A., Malik, F., Pay, C., Pearson, P., Price, C., & Win, Z. (1995). The impact of stage fright on student actors. *British Journal of Psychology,* 86, 27–39.

**Sweeney, G. A. & Horan, J. J.** (1982). Separate and combined effects of cue-controlled relaxation and cognitive restructuring in the treatment of musical performance anxiety. *Journal of Counseling Psychology,* **29**, 486–97.

**Theorell, T., Alberg-Hulten, G., Sigala, F., Perski, A., Söderholm, M., Callner, A., & Eneroth, P.** (1990). A psychosocial and biomedical comparison between men in six contrasting service occupations. *Work and Stress,* **4**, 51–63.

**Valentine, E.R., Fitzgerald, D. F. P., Gorton, T. L., Hudson, J. A., & Symonds, E. R. C.** (1995). The effect of lessons in the Alexander technique on music performance in high and low stress situations. *Psychology of Music,* 23, 129–41.

**Van Kemenade, J. F. L. M., Van Son, M. J. M., & Van Heesch, N. C. A.** (1995). Performance anxiety among professional musicians in symphonic orchestras: A self-report study. *Psychological Reports,* 77, 555–62.

**Wesner, R., Noyes, R., & Davis, T.** (1990). The occurrence of performance anxiety among musicians. *Journal of Affective Disorders,* **18**, 177–85.

**Widmer, S., Conway, A., Cohen, S., & Davies, P.** (1997). Hyperventilation: A correlate and predictor of debilitating performance anxiety in musicians. *Medical Problems of Performing Artists,* **12**, 97–106.

**Wolfe, M. L.** (1989). Correlates of adaptive and maladaptive musical performance anxiety. *Medical Problems of Performing Artists,* **4**, 49–56.

CHAPTER 14

# COMMUNICATING EMOTION IN MUSIC PERFORMANCE: A REVIEW AND THEORETICAL FRAMEWORK

PATRIK N. JUSLIN

Music is perhaps the most widely practised and appreciated of all art forms. Part of the appeal of music may lie in its ability to induce and represent emotions (cf. Budd 1985; Davies 1994; Meyer 1956; see also Sloboda & Juslin, this volume). Music is often regarded as an effective means of communicating emotions, and this notion pertains not only to the notated or implied structure of the music, but also to the way it is *performed.* Both empirical research (Persson 1995; Persson *et al.* 1996; Woody 2000) and biographical accounts (Blum 1977; Kennedy 1990; Schumacher 1995; Seckerson 1991) have suggested that many performers conceive of performance in terms of emotions, moods, or scenarios. A quote from one of the leading guitarists in the world, Eric Clapton, may serve to illustrate this point: 'Music . . . it's a form of communication and reassurance of feelings . . . what I get out of music is a feeling that I'm not alone' (Clapton 1998, p. 20).

The last decade has seen a steady increase of studies of how performers communicate emotions to listeners. The goal of this chapter is to: (a) review research on communication of emotions in music performance; (b) outline a theoretical framework suitable for organizing the findings; and (c) consider some implications for future research on music performance.

## 14.1 The role of the performer

Music might be viewed as part of a communication system in which composers code musical ideas in notation, performers recode from the notation to musical signal, and listeners recode from the acoustic signal to ideas (Kendall & Carterette 1990). The communicative content in music performance includes the performer's *interpretation* of the musical composition, that is, the performer's individualistic shaping of the piece according to his or her intentions (Palmer 1997). Most music performances involve some intention from the performer's side regarding what the music should 'express' to the listeners. Consequently, interpretation involves assigning some kind of meaning to the music (Persson 1995; see also Persson, this volume).

Although there has been much debate on what constitutes the 'meaning' of music, most accounts seem to involve a reference to the emotional impact of music (Budd 1985). It is not surprising, then, that a great deal of research has explored the expressive properties of various musical compositions (Gabrielsson & Lindström, this volume). However, the *performance* is also important in determining the emotional expression of a piece of music. The same notated structure can be performed in a number of different ways, and the precise way it is performed may affect the listener's impression of the music in profound ways. Both anecdotal evidence and empirical research have suggested that performers attempt to convey particular moods to listeners (cf. Persson 1995; Shaffer 1992; Woody 2000). Such expressive intentions may be more common in certain musical styles (e.g. opera, blues) than in others (e.g. serial music). Similarly, they may be more common among some performers (Fig. 14.1) than among others. Still, they occur frequently enough to be of legitimate interest to artists and scientists alike.

The process of interpretation itself is still largely shrouded in mystery. It is often argued that expertise in music performance involves the synthesis of technical and expressive aspects (Gabrielsson 1999; Sloboda 1996), but music teachers tend to spend more time and effort on the former than the latter (Tait 1992). Many music teachers still regard interpretation as 'the thing that cannot be taught'. Others make efforts to



Figure 14.1 Expressive performance: blues guitarist B. B. King in concert. King (1996) explained his relationship to the guitar: 'If I was feeling lonely, I'd pick up the guitar; feel like talking, pick up the guitar; if something's bugging me, just grab the guitar and play out the anger; happy, horny, mad or sad, the guitar was right there. It was incredible luxury to have this instrument to stroke whenever the passion overcame me' (p. 41). 'I started experimenting with sounds that expressed my emotions, whether happy or sad, bouncy or bluesy. I was looking for ways to let my guitar sing' (p. 123). (Photographs by Steve Berman.)

enhance performers' expressive skills by using tools such as metaphors, modelling, or mood induction (see Juslin & Persson, in press).

There are various traditions in musical interpretation, and emotions may enter into the process of developing an interpretation in different ways depending on the tradition. In one tradition, the performer's task is mainly to be 'a faithful carrier of the composer's message' (Sloboda 1998, p. 22). In this tradition, interpretation involves a careful study of the score, perhaps accompanied by consultation with the composer or of historical commentaries. The main factors determining the interpretation are performance convention and inferences about the expressive intentions of the composer.

In other traditions of interpretation, the performer has much more freedom. (This is, of course, particularly true in musical improvisation where the performer is also the composer.) In these traditions, the interpretation may be guided by the performer's own emotions while playing the piece (see Woody 2000). Alternatively, the performer may attempt (explicitly or implicitly) to render the performance with an emotional expression that seems suitable for the piece in question, perhaps enhancing a quality of emotion that is already latent in the piece of music. The performer's judgment of what emotion is suitable for the piece may be guided by his or her own listening to the music or by simple intuition.

In sum, the amount of freedom allowed to the performer in his or her interpretation of a piece varies greatly depending on the musical style and culture. This implies that the relative importance of the performance in determining the emotional expression of a piece of music is likely to vary considerably across different contexts. Although the emotional expression of a piece depends *both* on its composed structure and on its realization in performance, one may, in principle at least, separate the two aspects for research purposes. In fact, such a distinction has formed the basis of a series of recent studies of how performers communicate emotions to listeners. In the following, I summarize the main findings from this research and delineate a theoretical framework suitable for organizing the findings.

## 14.2 A brief history of the research

Throughout history, there has been plenty of anecdotal evidence concerning the extraordinary expressive skills of particular performers. Performers, listeners, and critics have all testified to the effectiveness of performance in communicating specific emotions to listeners. This can be illustrated by the following quote from the eighteenth-century violinist Francesco Geminiani:

> . . . with regard to musical performances, experience has shown that the imagination of the hearer is in general so much at the disposal of the [performer] that by help of variation, intervals, and modulation he may stamp what impression on the mind he pleases (quoted in Meyer 1956, p. 201).

Claims of this kind are common in the literature. However, anecdotal evidence can be unreliable. Further, in addition to those authors who have made claims about the expressive powers of music, there are others who have argued that music is unable to express anything whatsoever, like Stravinsky (Fisk 1997, pp. 280–1). The extent to which performers can communicate specific emotions to listeners has thus remained an open

question. Given that both performers and listeners regard expressive aspects of performance as important, it may be surprising to learn that research on emotional expression in performance is a fairly recent phenomenon, especially in comparison to research on emotional expression in compositions (Gabrielsson & Juslin, in press).

That this problem did not receive attention earlier is strange. One of the general findings from the early studies of performance was that the performance of a piece of music practically never corresponds to the nominal note values of the notation. In every performance, there are certain 'expressive deviations', or 'systematic variations', from what seems prescribed by the notation. These variations are not random, because performers are able to replicate them with great precision (Shaffer & Todd 1987). This gives rise to a question: what is the *function* of these variations?

One function of expressive variations may be to clarify the structure of the piece to the listener (Clarke 1988; Palmer 1989, 1997; Sundberg 1988). By clarifying the structure, the performer may try to enhance the 'inherent' emotional impact of the musical structure. In this view, then, expressive variations primarily reflect the performer's conceptual interpretation of the structure. However, this explanation has failed to account for all of the variance in musical performances. Repp (1998) analysed 115 performances of a Chopin piece in order to study the individual differences among performers. Based on the results, Repp convincingly argued that individual differences at the phrase level do not arise (mainly) from different interpretations of the musical structure but from different ways of giving 'expressive shape' to one and the same structure. Indeed, a number of researchers have suggested that expressive variations may serve to convey moods to listeners (Gabrielsson 1988; Juslin 1997*a*; Shaffer 1992).

One of the pioneers in music psychology, Carl E. Seashore, argued that 'deviation from the exact . . . is the medium for the creation of the beautiful—for the conveying of emotion' (quoted in Seashore 1937, p. 155). Seashore did not develop any theory to explain how such systematic deviations convey emotions to listeners. However, he did propose a suitable paradigm for exploring this problem. He was interested in the question of how singers express emotions and asked whether there is 'any possibility of a scientific approach' (Seashore 1947, p. 171). He gave the answer himself by describing a study of vocal expression by Fairbanks (1940), one of his colleagues at the University of Iowa. This study required actors to express different emotions in speech, which was later analysed acoustically. Seashore concluded that 'we can exactly parallel this experiment in the field of music' (Seashore 1947, p. 176).

Strangely, Seashore's plea went unheard, and he did not publish any study of that sort himself. Before the mid-1970s, there was not a single study of emotional expression in music performance. By the late 1980s, only two studies on this topic had been published (Kotlyar & Morozov 1976; Senju & Ohgushi 1987). After this modest beginning, however, there was a virtual explosion of studies in the early 1990s (for a review, see Juslin 1999). Most of these studies used the same paradigm that Seashore suggested 40 years earlier.

## 14.3 The standard paradigm

The paradigm used in most studies of emotional expression in music performance is simple: performers are asked to play a number of brief melodies to express different

emotions chosen by the researcher. The resulting performances are first recorded and then evaluated in listening experiments to see whether listeners can recognize the intended expression. Each performance is further analysed to study what acoustic means the performer used to achieve each emotional expression. (The focus in these experiments has primarily been on *represented* emotion rather than *induced* emotion.) The basic assumption is that because the melodies remain the same in different expressions, whatever effects appear in listeners' judgments or in acoustic measures should mainly be the result of the performer's expressive intention. To have performers play the same piece of music with different kinds of emotional expression might seem 'unnatural' from a musical point of view. However, this design is necessary to secure the internal validity of the experiments: If different emotions are expressed by different melodies, it is impossible to know whether the obtained effects on listener judgments or performance measures are due to the melody, the performance, or some complex interaction between the two. This paradigm is not without its problems, but I will temporarily postpone my critique until later sections (for alternatives approaches to studying emotional expression in performance, see Section 14.7.1).

There are about 30 studies of emotional expression in performance. These studies cover a broad range of musical styles, such as opera, classical music, folk music, jazz, and pop/rock. The emotions investigated have usually included happiness, sadness, anger, fear, and tenderness. A variety of musical instruments has also been included, such as the violin, flute, clarinet, electric guitar, piano, trumpet, drums, synthesizer, and singing voice. Studies have mainly focused on two aspects of the communicative process: (a) the accuracy of the communication; and (b) the code used by performers and listeners.

## 14.4 Communication accuracy

The first question of interest to any researcher concerned with emotional communication in music performance is whether such communication is at all possible. It can seem odd to talk about 'accuracy' in the context of music, but most performers are probably concerned about whether their interpretation of the music is perceived in the intended way by listeners. Thus, the extent to which a performer and a listener agree on the expression of the performance can be seen as a measure of the 'accuracy' of the communication. In a pioneering study, Kotlyar and Morozov (1976) asked ten opera singers to sing phrases from various pieces of music in such a way that they would communicate joy, anger, sorrow, and fear to listeners. Musically trained listeners were asked to judge the expression of each performance. The results showed that the listeners were highly successful at recognizing the intended expression.

Since then, many studies have confirmed that professional music performers are able to communicate particular emotions to listeners (Behrens & Green 1993; Gabrielsson & Juslin 1996; Gabrielsson & Lindström 1995; Juslin 1997*a*, 2000; Juslin & Madison 1999; Ohgushi & Hattori 1996*a*; Sundberg *et al.* 1995). However, these studies have also revealed marked individual differences in *encoding* (expression) and *decoding* (recognition) accuracy, consistent with the results from studies of other non-verbal communication channels (Buck 1984; Wallbott & Scherer 1986). Studies of amateurs have

indicated that they have a lower accuracy than experts, but that their accuracy can be improved by appropriate feedback (Juslin & Laukka 2000).

Few studies have reported the decoding accuracy in a manner that makes more precise estimates possible. However, Juslin (1997*b*) used a forced-choice format in order to make the results comparable with the results from studies of vocal expression. Juslin found that decoding accuracy was 75 per cent correct, or about four times higher than what would be expected by chance alone. This suggests that, at least under ideal circumstances, the accuracy of communication of emotion via musical performance may approach the accuracy of facial or vocal expression (cf. Ekman 1973; Scherer 1986). The examination of error distributions shows that certain pairs of emotional expressions are confused more frequently than others. For instance, sadness and tenderness expressions are typically confused with each other, whereas sadness and happiness expressions are rarely confused. On the whole, though, the communicative process is robust.

One objection to this finding could be that the accuracy of the communicative process is boosted because of the response formats used. In most studies, listeners made their judgments by means of forced-choice or adjective ratings, methods which offer only a limited number of response options. To address this problem, Juslin (1997*c*) investigated the generalizability of decoding accuracy across response formats using a *parallel enrichment procedure* (Rosenthal 1982). That is, quantitative data from forced-choice judgments of musical performances were augmented by qualitative data from free labelling of the same performances. The two methods converged on the conclusions that: (a) communication was reliable regardless of the response format; and (b) what could be communicated reliably was the basic-level emotion categories, but not specific nuances within these categories.

## 14.5 Code usage

The early studies were primarily concerned with investigating whether performers are able to communicate emotions to their listeners at all. However, if we want to study communication as a process, we must also consider its mechanisms, particularly the code used by performers and listeners. Accordingly, researchers have tried to describe the means by which performers express specific emotions (Baroni & Finarelli 1994; Gabrielsson & Juslin 1996; Gabrielsson & Lindström 1995; Juslin 1997*a*, 1997*b*, 2000; Juslin & Madison 1999; Rapoport 1996).

### 14.5.1 Expressive cues, emotion categories, and emotion dimensions

One main finding from this line of research is that the performer's expressive intention affects almost every aspect of the performance; that is, emotional expression in performance seems to involve a whole set of *cues*—or bits of information—that are used by performers and listeners. A summary of the code usage established in studies thus far is presented in Fig. 14.2. Included are the five emotions (happiness, sadness, anger, fear, love/tenderness) that have been studied most extensively. These emotions represent a natural point of departure since all are regarded as typical emotions by lay people (see Shaver *et al.* 1987; Shields 1984) and have been postulated as so-called 'basic emotions'

by scientists (Plutchik 1994, p. 58). The same emotions have also occurred (some more than others) in the expression marks of musical scores (e.g. *festoso, dolente, furioso, timorosi, teneramente*).

Figure 14.2 combines *categorical* and *dimensional* approaches to emotional expression to illustrate that the code allows performers to communicate both graded signals (e.g. level of activity) and categorical signals (e.g. happiness). Each emotional expression is placed at an approximate point in a two-dimensional emotion space constituted by *valence* and *activity level* (cf. Russell 1980). The placement of each emotional expression in the emotion space is partly based on the results from a study in which participants rated the valence and activity of 400 emotion terms (Whissell 1989). These positions of the emotional expressions are further supported by a cluster analysis of listeners' ratings of various emotional expressions reported in Juslin (1997*b*). Figure 14.2

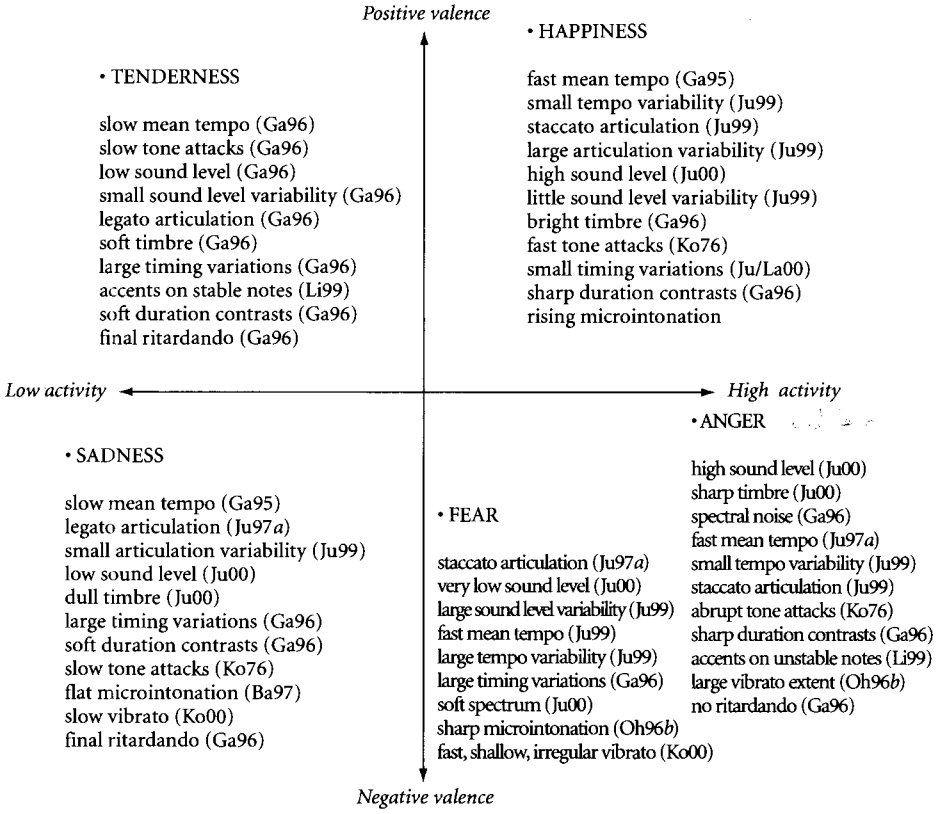



**Figure 14.2** Summary of cue utilization in performers' communication of emotion in music. (One representative study is cited for each cue. Authors' names are abbreviated to the initial two letters of the first author (and the second author if needed for clarity), and the publication year to the two last digits, e.g. Ko76 = Kotlyar & Morozov 1976.)

shows that happiness and tenderness expressions are both positive, whereas only happiness is high in activity. Similarly, both anger and sadness expressions are negative, whereas only anger is high in activity.

As seen in Fig. 14.2, the expressive cues include tempo, sound level, timing, intonation, articulation, timbre, vibrato, tone attacks, tone decays, and pauses. Both the mean level of a cue and its variability throughout the performance may be important for the communicative process. For example, sadness expressions are associated with slow tempo, low sound level, legato articulation, small articulation variability, slow tone attacks, and dull timbre, whereas happiness expressions are associated with fast tempo, high sound level, staccato articulation, large articulation variability, fast tone attacks, and bright timbre (for a more detailed review, see Juslin 1999). Although Fig. 14.2 is limited to a few seemingly 'static' emotion categories, it is easy to imagine how the emotional expression can be gradually altered during the course of a performance, along the two emotions dimensions.

Some of the cues described in Fig. 14.2 may be common knowledge to most performers, whereas other aspects, such as those having to do with the *microstructure* (e.g. timing) of the performance, are less obvious. The number of cues available depends on the instrument used. In addition, the cue utilization is not completely consistent across performers, instruments, or pieces of music (Juslin 2000). Notably, even children (4–12 years old) seem to be able to use some of these cues to express emotions in song. For instance, they use a fast tempo and a high sound level in happy expressions, whereas they use a slow tempo and a low sound level in sad expressions (Adachi & Trehub 1998).

The positions of the emotional expressions in the emotion space in Fig. 14.2 suggest that certain cues may have a larger influence on certain emotion dimensions than others. Thus, for instance, tempo, sound level, and articulation seem to be especially important for the activity level, whereas timbre (e.g. the amount of high-frequency energy in the spectrum), perhaps in combination with sound level, seems to be especially important for emotional valence. In the latter case, intermediate levels of sound level and high-frequency energy are associated with positive emotions, whereas more extreme levels (e.g. very high or very low sound level) are associated with negative emotions. This hypothesized two-factor structure can, of course, be tested in future research, using listening tests with synthesized performances.

As suggested by Fig. 14.2, research to date has primarily focused on the expressions of a set of 'basic emotions' (Sloboda & Juslin, this volume). However, it is interesting to speculate whether 'secondary' or 'complex' emotions may be developed from these expressions. Many emotion researchers believe that such emotions constitute 'mixtures' of more basic emotions. For instance, it has been proposed that pride involves the combination of happiness and anger (Plutchik 1994, p. 61). Judging from Fig. 14.2, the expression of pride would then presumably involve some kind of compromise among the cues used to communicate happiness and anger, although perhaps with a slightly lower activity level (cf. Whissell 1989).

Performance analyses have shown that performers may use a number of cues to express specific emotions. However, these analyses do not show that listeners actually *use* the same cues in their judgments. To test the validity of hypotheses derived from performance studies, it is therefore necessary to conduct listening experiments with

synthesized and systematically varied performances. Such experiments have mostly confirmed that listeners make the same associations between cues and emotion categories as performers do (Juslin 1997*b*; Scherer & Oshinsky 1977; see also Bresin & Friberg 2000; Juslin *et al.*, in press).

In addition to the results with regard to relationships between acoustic cues and emotion categories, Juslin (1997*b*) reported findings on how acoustic cues contribute to the perceived 'expressiveness' of the performance. Juslin manipulated the tempo, sound level, articulation, timbre, and tone attacks of synthesized performances in a factorial design, and asked listeners to rate the expressiveness of the resulting 108 cue combinations. The single cue combination that was most expressive according to listeners had the following characteristics (with cues in order of predictive strength): legato articulation, soft spectrum, slow tempo, high sound level, and slow tone attacks. This raised an interesting question: Is this how performers should play in order to be judged as expressive by listeners? It may be noted that the most expressive cue combination was highly similar to the cue combination that expressed sadness and tenderness best according to the listeners. This seems to suggest that there is a close connection between sadness/tenderness and expressiveness. This result was confirmed by a cluster analysis of the listeners' judgments, which showed that the ratings of sadness and tenderness correlated more strongly with the ratings of expressiveness than did the ratings of happiness, fear, or anger. It is worth noting in this context that so-called 'separation calls', associated with social loss, are believed to elicit strong sympathetic responses in listeners (Plutchik 1994, p. 331). This could perhaps explain why listeners find performances with a sad expression particularly expressive.

### 14.5.2 The importance of musical time: expressive contours

Most studies thus far have focused on cues that remain 'fixed' during the performance or on 'summary measures' (e.g. mean, standard deviation) of cues that do vary. However, from an intuitive point of view, it would seem that much of music's expressiveness lies in patterns of *changes* in cues during the course of a performance. Frick (1985) used the concept 'prosodic contours' to refer to dynamic patterns of voice cues over time found in vocal expressions. In music, this corresponds to continuously changing patterns of tempo, articulation, and sound level throughout the performance (examples of patterns of articulation for different emotions are presented in Fig. 14.3). In this chapter, I will call these patterns *expressive contours*. Are such expressive contours involved in the communicative process?

Preliminary evidence that expressive contours are really involved in the communicative process was obtained in a study by Juslin (1997*b*). This study included two kinds of stimulus. The first consisted of a set of musical performances of 'Nobody Knows' on the electric guitar, made with different expressive intentions (e.g. sadness, happiness, tenderness). The second kind of stimuli consisted of synthesized performances of the same melody, in which the same emotional expressions were simulated on the basis of hypotheses from previous performance analyses. The 'human' performances included expressive contours, whereas the synthesized performances included only cues that were 'fixed' throughout the whole performance. Both kinds of performance communicated the intended emotions with high reliability.

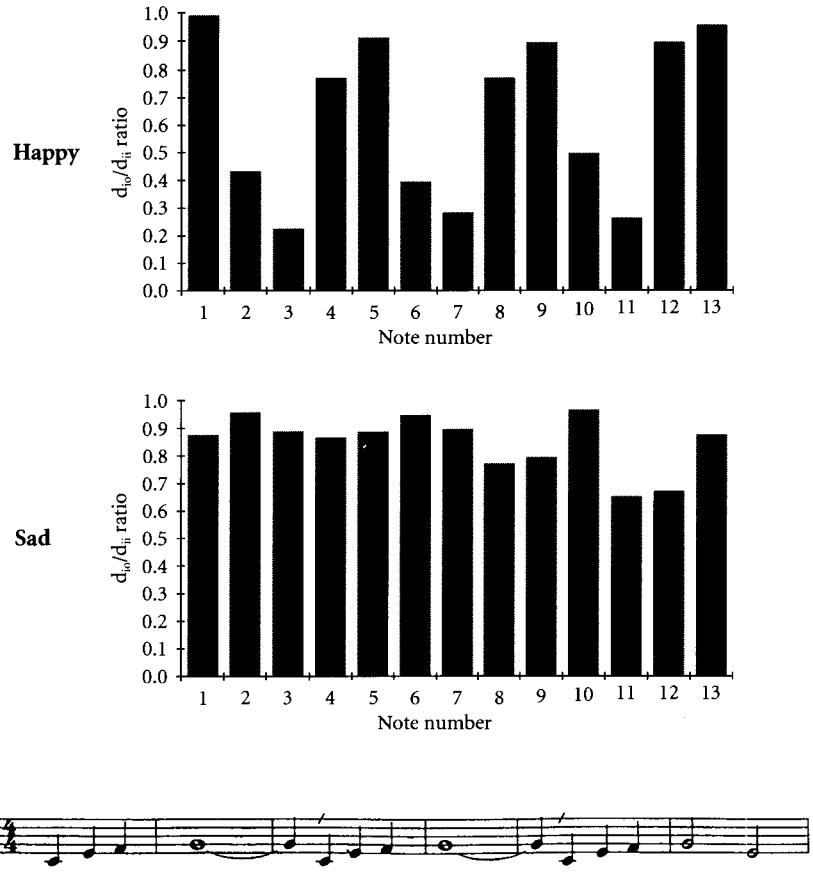



Figure 14.3 Articulation patterns for happy and sad performances of 'When the Saints' on the electric guitar (unpublished data from Juslin 1997*a*). The duration from the onset of a tone until the onset of the next tone ($d_{ii}$) and the duration from the onset of a tone until its offset ($d_{io}$) were used to calculate the $d_{io}/d_{ii}$ ratio, i.e. the articulation. Articulation close to 1.0 may be referred to as legato, whereas articulation below 0.5 may be referred to as staccato.

To study the relative importance of expressive contours, Juslin played the performances backwards and then asked listeners to decode the intended emotional expression. Playing the sequences backwards changed the expressive contours (i.e. patterns over time) but preserved the summary measures (e.g. tempo) of the performances. The results showed that the human performances suffered from this manipulation (i.e. decoding accuracy was reduced), whereas synthesized performances did not. This suggests that human performances depend, at least to some extent, on the use of expressive contours. However, this study did not provide any clues about what expressive contours were used by listeners.

Further evidence that expressive contours may be of importance comes from a study by Juslin and Madison (1999). They used a combination of analysis and synthesis

afforded by the possibility to 'resynthesize' human performances by means of computer technology. This method was used to investigate the relative importance of timing patterns in communication of emotions through musical performance. Juslin and Madison gradually removed different expressive cues (tempo, dynamics, timing, articulation) from piano performances played with various intended expressions (anger, sadness, happiness, fear) to see how such manipulations would affect a listener's ability to decode the emotional expression. The results indicated that: (a) removing the timing pattern yielded a decrease in listeners' decoding accuracy; (b) timing patterns alone were capable of communicating at least some emotions with better than chance accuracy; and (c) timing patterns were less effective in communicating emotions than were tempo and dynamics. The results further indicated that timing patterns (defined as sequences of inter-onset intervals) were particularly important for fear expressions, whereas articulation patterns were particularly important for happiness expressions. This study demonstrated that listeners indeed can use timing patterns to decode the emotional expression of a performance. However, the results did not reveal what it is in the timing patterns that provides the listeners with emotional information.

One way of exploring what it is in the expressive contours that provides the emotional information is to rely on listeners' continuous ratings of expression in performance. In recent years, researchers have increasingly been developing methods for describing the moment-to-moment changes in listeners' emotional responses to music (for a review, see Schubert, this volume). Most of the attempts have involved the use of different pieces of music. However, Sloboda and Lehmann (in press) have applied continuous response methodology specifically to performance. They focused on how music performers use various strategies to enhance the emotional intensity of a performance. Ten expert pianists were required to perform a Chopin prelude as if in a concert. The resulting performances were recorded and the performers were interviewed concerning their interpretations of the piece. Finally, musically trained listeners were asked to rate the concurrent level of perceived 'emotionality' in each interpretation by moving a mouse pointer on a continuous response computer interface.

The results indicated that the different interpretations of the piece gave rise to different profiles of perceived emotional intensity, and that at least some of the differences among the interpretations could be explained in terms of the expressive devices used by the performers. The performers used devices related to timing, dynamics, and asynchronies between melody notes and bass notes (e.g. delaying a note, lengthening a note, changing the dynamic range). Most of the expressive devices occurred at the beginnings of musical phrases. But can such devices be used by listeners to decode different emotions?

One possible answer is provided by research on how notated structure and performed structure *interact* in producing an emotional expression. In a recent study, Lindström (1999) manipulated a brief melody ('Frère Jacques') concerning tonal progression, rhythm, melodic contour, and melodic direction. The resulting versions of the melody were rated by listeners on emotion scales. On the basis of these ratings, Lindström chose 16 versions of the melody that could be regarded as 'good' or 'less good' exemplars of specific emotions (e.g. anger). Then, he asked professional pianists to play the different versions of the melody to express four different emotions (happiness,

sadness, anger, tenderness) to listeners. The performers were asked to 'think aloud' throughout the preparation of the performance.

Two findings were of particular interest. First, there was a 'compensation tendency' in the performances, such that the performers used more 'extreme' variations in the performance to achieve the intended expression (e.g. happiness) in those versions of the melody that were 'less good' exemplars of the intended emotion. Second, the results showed that the performers accentuated different features of the melodic structure depending on the intended expression. Thus, for example, if a certain note in the melody structure was seen as particularly 'happy', then the performers emphasized this note (e.g. via accents in timing, dynamics, articulation) in happiness expressions, but de-emphasized it in sadness expressions. However, the general pattern of cues was the same as in previous studies, which suggests that it is mainly the local aspects of the performance that are affected by particular melodies, whereas overall cues (e.g. mean tempo, mean articulation, mean sound level) are used pretty much the same regardless of the particular melody (see Juslin 2000). In sum, it seems that, depending on the intended expression, different parts of the melodic structure are emphasized. This might explain how listeners can arrive at judgments of the expression on the basis of expressive contours.

This section has shown that there are a number of acoustic cues that performers can use to communicate emotions to listeners. But many studies have shown that cue utilization is inconsistent both across and within performers. Similarly, listening experiments have shown that cue utilization is inconsistent both across and within listeners (cf. Juslin 1998). This presents something of a puzzle: how can the communicative process be so reliable if the cue utilization is inconsistent? To explain the characteristics of the communicative process, we need a theoretical framework.

## 14.6 A theoretical framework: the functionalist perspective

One explanation of the scarcity of studies of emotional expression in music performance prior to 1990 might be a lack of relevant theories. This has led to a neglect of expressive aspects of performance in both research and education. One candidate for a useful theory would seem to be Clynes' (1973, 1977) 'sentic' theory, in which he argued for the existence of biologically pre-programmed spatiotemporal forms, or *essentic forms*, for the communication of specific emotions. Unfortunately, independent attempts to replicate these spatiotemporal forms have yielded mixed results (e.g. Gorman & Crain 1974; Nettelbeck *et al.* 1989; Trussoni *et al.* 1988).

The only elaborated theoretical framework aimed specifically at performance aspects has been presented by Juslin (1995, 1997*a*). Juslin suggested that researchers should adopt a *functionalist perspective* on emotional communication in music performance. This involves the integration of ideas from research on emotion and non-verbal communication with Egon Brunswik's (1956) *probabilistic functionalism*. This framework provides useful questions, hypotheses, and ways of interpreting results. The functionalist perspective is concerned with three aspects of the expressive code: (a) its origin; (b) its description; and (c) its improvement. The following sections focus on the first two of these aspects. For a discussion of applications for music education, see Juslin and Laukka (2000) and Juslin and Persson (in press).

### 14.6.1 Origin of the code

The fact that performers are able to communicate emotions to listeners is something that requires an explanation: What is it that makes this possible? What is the origin of the code used by performers and listeners? According to the functionalist perspective, the key to this problem lies in considering the *functions* that communication of emotions has served—and continues to serve—in social interaction (e.g. Darwin 1872). Although music performance may be a recent addition to the human behavioural repertoire, it must still be constrained by psychological mechanisms that have been shaped by evolution (Juslin 1998). Evolutionary considerations may be particularly relevant for emotional communication, because animals also communicate emotion, for instance, via acoustic signals (e.g. Plutchik 1994, Chapter 10). Accordingly, Juslin (1998) has argued that communication of emotion in music performance reflects two factors: (a) brain programmes; and (b) social learning. Each of these factors will be discussed below.

*Brain programmes*

The first factor hypothesized to govern emotional expression in performance is innate 'brain programmes' for vocal expression of emotion. Studies of both monkeys and humans with brain lesions have investigated the neurological 'substrates' underlying spontaneous vocalizations of emotion (e.g. Jürgens & von Cramon 1982; MacLean 1978; Panksepp 1982; Penfield & Roberts 1959; Ploog 1981, 1986). These studies have yielded evidence of 'brain programmes' that function to initiate and organize 'pre-wired' vocal expressions. Electrical stimulation of particular brain structures in the amygdala and the hypothalamus produces vocalizations that observers cannot distinguish from natural expressions. In humans, at least, these 'pre-wired' vocal expressions can also be executed at will for communicative purposes, thanks to brain structures in the anterior cingulate cortex. Such 'brain programmes' might explain the reported cross-cultural agreement in encoding and decoding of vocal expressions (Frick 1985). There is further evidence that both encoding (Shapiro & Danly 1985) and decoding (Borod *et al.* 1998) of emotion through voice prosody is handled by the right hemisphere of the brain.

What, then, does voice prosody has do to with performance of music? According to the functionalist perspective, music performers are able to communicate emotions to listeners by using the same acoustic code as is used in vocal expression of emotion. The notion that there is a close relationship between music and the human voice has a long history (Helmholtz 1954; Kivy 1989; Rousseau 1986; Scherer 1995; Sundberg 1982). However, it seems unlikely that this link can explain all of music's expressiveness. I have thus suggested that the hypothesis that there is an iconic similarity between vocal expression of emotion and musical expression of emotion applies mainly to those aspects of the music that the performer can control during his or her performance, such as tempo, loudness, or timbre (Juslin, in press). Neuropsychological evidence suggests that certain aspects of music, such as timbre, share the same neural resources as speech, whereas others (e.g. tonality) draw on neural resources that are unique to music (Handel 1991; Patel & Peretz 1997). Further, the same brain hemisphere (the right) seems to be involved in perception of emotion in music and speech (Borod *et al.* 1998; Bryden *et al.* 1982).

In addition, whereas previous formulations of the link have been vague and speculative, the functionalist perspective specifies the actual neurological substrates responsible for these similarities, and leads to research that provides empirical evidence in favour of this hypothesis. Juslin (1999) carried out a systematic review of the code usage in vocal expression and music performance. The results revealed a number of similarities in code usage. For example, vocal expression of sadness is associated with slow speech rate, low voice intensity, low intonation, and little high-frequency energy in the spectrum of the voice (Juslin & Laukka, in press; Scherer 1986). The same acoustic cues are used to communicate sadness in musical performances (cf. Fig. 14.2).

Juslin also argues that encoding and decoding of emotions proceed in terms of a limited number of 'basic' emotion categories (cf. Ekman 1992; Izard 1977; Johnson-Laird & Oatley 1992; Oatley 1992), which provide decoders with the maximum information and discriminability (Ross & Spalding 1994). To be useful as guides to action, expressions are quickly decoded in terms of just a few categories of emotion related to such fundamental life problems as danger (fear), competition (anger), loss (sadness), and cooperation (happiness). It may be noted that even in the first months of life, infants are able to differentiate prototypical vocal expressions in infant-directed speech, and to respond adequately to their categorical messages (Papoušek *et al.* 1990). Indeed, infants may be able to decode vocal expression of emotion even earlier than they are able to decode facial expression (see Oatley & Jenkins 1996, Chapter 6). There is also evidence of categorical perception of emotion in vocal expression in adults (e.g. de Gelder & Vroomen 1996). This supports a conceptualization of the perception of emotion from acoustic stimuli in terms of basic emotion categories.

A related finding is that neurophysiologist and ethologist Ploog (1981, 1992) and his co-workers have managed to identify a limited number of vocalization categories in squirrel monkeys. (Note that the neurological substrates involved in vocalizations are rather similar in monkeys and humans.) The vocalization categories are associated with important events in the monkeys' lives, and include warning signals ('alarm peeps'), threat signals ('groaning'), desire for social contact signals ('isolation peeps'), and companionship signals ('cackling'). It seems plausible that vocal expressions deriving from 'brain programmes' discovered in humans have served—and still continue to serve—similar communicative functions (Panksepp 1982).

It is possible to speculate that vocal expressions associated with the 'basic' emotion categories gradually became meshed with music that accompanied related cultural activities. There is evidence that people can accurately categorize songs of different 'emotional' types (e.g. mourning, love, war, lullaby) that come from diverse cultures (e.g. Eggebrecht 1983). Moreover, there are similarities in the acoustic code used in such songs in different cultures; for example, mourning songs usually have a slow tempo, low sound level, and soft timbre, whereas festive songs have a fast tempo, high sound level, and bright timbre (Eibl-Eibesfeldt 1989). Thus, the acoustic code used by performers today may be linked—in functional ways—to important cultural activities of the past (Gregory 1997; see also Becker, this volume).

*Social learning*

The second factor governing emotional expression in performance is social learning or specific memories. This is a life-long process that begins with the socialization of

emotions. Certain aspects of music performance (e.g. those not shared with vocal expression) may be completely governed by cultural influences. Others—although based on innate mechanisms—are 'modulated' by experience. This modulation of the code begins with the early interaction between mother and infant (or even earlier, through 'biochemical imprinting' during the prenatal stage). When mothers talk to their infants, for example, if they want to calm their infant, they reduce the speed and intensity of their speech and talk with slowly falling pitch contours. If, on the other hand, mothers want to express disapproval towards some unfavourable activity they use brief, sharp, and staccato-like contours (Papoušek 1996). Thus, as Papoušek argues, the origins of expressive skills in music cannot be easily separated from the earliest stages of emotional communication.

The shaping of the expressive code continues throughout life, along with accumulated experience. Performers learn links between expressive cues and extramusical aspects (e.g. body language, biological motion) through analogies (cf. Sloboda 1996). In the professional career, the code is modulated by advice from music teachers, performance conventions, and the use of metaphors. This second factor of social learning, reflecting cultural influences, could explain: (a) the individual differences among performers; (b) why the cue utilization is inconsistent; and (c) why there are cultural differences in cue utilization. However, Juslin (1999) assumes that the 'core' of the expressive code is provided by the first factor—innate brain programmes for vocal expression. This could explain why expressive skills are often regarded as 'instinctive'.

Predictions based on the functionalist perspective that have received empirical support include: (a) that there are similarities in cue utilization between vocal expression and musical performance (Juslin 1999); (b) that basic emotions are easier to communicate than other emotions (cf. Gabrielsson & Juslin 1996; Gabrielsson & Lindström 1995; Senju & Ohgushi 1987); (c) decoding of emotion from music is quick (Peretz *et al.* 1998); (d) decoding of emotion from music performance is largely independent of musical training (Juslin 1997*a*); (e) that even children, 3–4 years old, are able to decode such basic emotions as happiness, sadness, anger, and fear in music with better than chance accuracy (see Cunningham & Sterling 1988; Dolgin & Adelson 1990; Kratus 1993; Terwogt & van Grinsven 1988); and (f) that listeners are able to decode basic emotions with accuracy better than chance even from single voice tones (Konishi *et al.* 2000).

One more prediction is that there should be cross-cultural agreement regarding some aspects, at least, of music's expressiveness. A common hypothesis is that there is universal recognition of emotions from vocal expression. For instance, Frick (1985) reviewed a large number of cross-cultural studies of vocal expression of emotion and concluded that 'there is little evidence for . . . cultural differences in the prosodic communication of emotion' (p. 414). If, as I have claimed, there are many similarities in cue utilization between vocal expression and music performance, we would expect to find some degree of cross-cultural agreement in listeners' judgments of emotional expression in musical performances. Thus, Juslin (1997*b*, p. 248) proposed that 'some aspects of music (e.g. tonality, melody, and harmony) are relatively more culture-specific, whereas other aspects . . . are more culture-independent (because they are based on nonverbal communication of emotion)'. This notion was subsequently turned into a formal model by

Balkwill and Thompson (1999). While few studies have explicitly aimed to address this issue, there is evidence of cross-cultural similarities in performers' utilization of cues from studies in India (Balkwill & Thompson 1999), Russia (Kotlyar & Morozov 1976), Italy (Baroni & Finarelli 1994), Japan (Ohgushi & Hattori 1996*b*), Sweden (Juslin 1997*a*), and Germany (Langeheinecke *et al.* 1999). These studies suggest that cues like tempo, sound level, and timbre are used in a similar manner in different cultures. When it comes to cross-cultural studies that involve other cues (e.g. harmony), there are larger cross-cultural differences (Gabrielsson & Juslin, in press).

The discussion so far has mainly been concerned with the origin of the acoustic code used in the communication of emotion in music performances. However, to truly understand the virtues and limitations of the communicative process, it is also necessary to describe the nature of the code in more detail. The following section describes a conceptual tool that may be useful in describing the essential characteristics of the communicative process.

### 14.6.2 Description of the code: the lens model

Juslin (1995, 1998, 2000) suggested that the communicative process be described in terms of a modified version of Brunswik's (1956) well-known 'lens model' (Fig. 14.4). This model is meant to illustrate how music performers *encode* specific emotions by means of a number of *probabilistic* (i.e. uncertain) albeit partly redundant *cues* in the performance (e.g. tempo, sound level, timbre, articulation). The emotions are *decoded* by listeners who use the same cues to judge the emotional expression. The cues are probabilistic in the sense that they are not wholly reliable indicators of the intended expression. Listeners thus have to *combine* the cues in order to arrive at reliable judgments of the expression. Brunswik's (1956) notion of *vicarious functioning* may be used

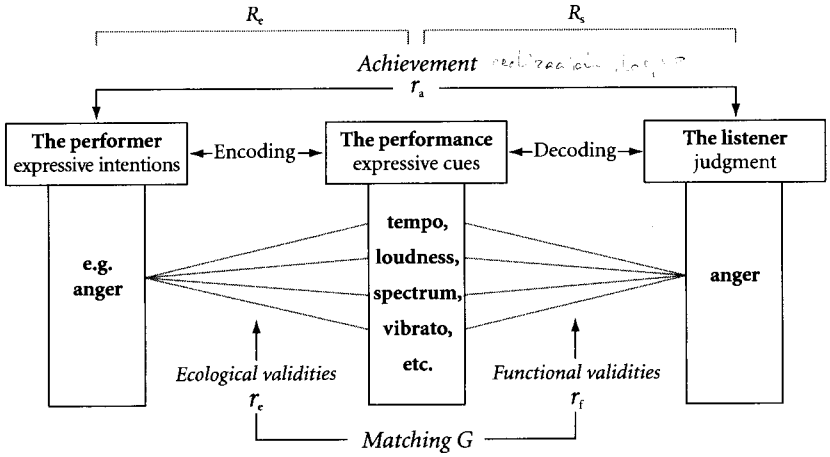



**Figure 14.4** A Brunswikian lens model for communication of emotion in music performance (adapted from Juslin 1995).

to describe how listeners use the partly interchangeable cues in flexible ways, sometimes shifting from one cue that is unavailable to another that is available.

Intercorrelations between cues partly reflect how sounds are produced on instruments. For example, hitting a string harder on the electric guitar generates a tone with increased sound level, but it also generates more high-frequency energy in the spectrum of the tone. However, intercorrelations between cues may also reflect how performers use the cues to accomplish different expressions (e.g. the use of high sound level and fast tempo in both anger and happiness). Empirical studies have confirmed that the cues in performances are probabilistically related to both performers' intentions and listeners' judgments (see Juslin 1997*a*, 1997*b*, 2000), and that there are cue intercorrelations (Juslin 1997*a*; Kratus 1993). Studies using synthesized music performances that are varied in a systematic manner have suggested that the performance cues contribute independently—in an additive fashion—to listeners' emotion judgments, as evidenced by a lack of interactions among the cues (e.g. Juslin 1997*b*). Notably, this is consistent with findings from studies of the vocal expression of emotion, which involve many of the same cues (Ladd *et al.* 1985). However, it stands in contrast to the findings from studies that manipulate aspects of the melodic structure (see Gabrielsson & Lindström, this volume).

The redundancy of the cues reduces the uncertainty of the communicative process. On the other hand, it also limits the information capacity of the communicative channel, because the same information is conveyed by many cues. In Brunswik's own words, the system aims at 'smallness of error at the expense of the highest frequency of precision' (1956, p. 146); such a system leads to 'compromise and falling short of precision, but also the relative infrequency of drastic error' (1956, p. 145). It is easy to see how a system for communication of emotions with this characteristic could have survival value. It is ultimately more important to avoid making serious mistakes (e.g. mistaking anger for joy), than to have the ability to make more subtle discriminations between emotions (e.g. detecting different kinds of joy). The characteristics of the present communicative system can thus be partly explained by evolutionary pressures of the past (Juslin, 2001).

A 'Brunswikian' conceptualization of the communicative process in terms of separate cues that are 'integrated'—as opposed to a 'Gibsonian' conceptualization in terms of more holistic 'higher-order variables' (Gibson 1979)—seems to be supported by studies of the physiology of listening. Handel (1991) points out that speech and music use similar perceptual mechanisms, and that the auditory pathways involve different neural representations for different aspects of the acoustic signal (e.g. timing, frequency) that are kept separate until later stages of analysis. Perception of both speech and music requires the integration of these different representations. (Of course, the question of exactly which acoustic cues are relevant and how they should be defined is an empirical matter.)

In the lens model, the relationship between the performer's expressive intention and a cue in the performance (e.g. articulation) describes the *ecological validity* ($r_e$) of that cue (Fig. 14.4). Ecological validity is a measure of the validity of that cue in predicting the performer's intention. The *functional validity* ($r_f$) of the same cue is indexed by the relationship between the cue and the listener's judgment. Functional validity is a measure

of the validity of the cue in predicting the listener's emotion judgment. *Achievement* ($r_a$) refers to the accuracy of the communication, as measured by the relationship between the performer's intention and the listener's emotion judgment. Finally, *matching* ($G$) refers to the extent to which ecological and functional validities are 'matched' to each other—that is, whether the performer and listener share a common code (the measurement of each of these quantities is described below).

The Brunswikian lens model has a number of important implications. First, we cannot expect perfect communication accuracy. If cues are only probabilistic, it means that accuracy can only be probabilistic. Second, to understand why the communication is successful or not in a particular situation, researchers need to describe both encoding and decoding in terms of the same concepts. Third, because the cues are intercorrelated to some extent, many different cue utilization strategies can lead to a similar level of accuracy (cf. Dawes & Corrigan 1974). There is a virtue associated with the redundancy of the communicative process: because there is no pressure toward uniformity in cue utilization, performers are able to communicate basic emotions successfully without compromising their unique styles of playing.

One problem in analysing cue utilization is that the strategies used by performers and listeners are largely *implicit* in nature (Juslin & Laukka 2000; Sloboda 1996). Therefore, we need to somehow externalize their expressive strategies. One useful approach to this problem may be to use multiple regression analysis (MRA) to describe the complex relationships between: (a) the performer's expressive intention and the cues (ecological validities); and (b) the listener's judgment and the cues (functional validities). MRA is a data-analytic method flexible enough to handle the complexity of the communicative process (e.g. multiple cues, uncertainty, intercorrelations among cues). MRA also permits us to describe performers and listeners in terms of similar concepts, so that their *cue weights* (i.e. the extent to which they rely on particular cues) can be compared (for details on how this is done, see Juslin 2000).

As demonstrated by Juslin (2000), MRA models of performers and listeners can be mathematically related to each other by means of the *lens model equation* (LME). The LME was originally presented in an influential article by Hursch *et al.* (1964) in the context of studies of human judgment. In such studies, the goal was to relate the judge's cognitive system to a statistical description of the judgment task (Cooksey 1996). However, the LME can also be used to describe communication of emotion in music performance.

The LME (eqn 14.1) embodies the fact that communication accuracy, or achievement ($r_a$), is a function of two additive components. The first component is usually called the *linear* component because it represents that component of the achievement which can be attributed to the linear regression models of the performer and the listener. The linear component shows that achievement is a function of *performer consistency* ($R_e$), *listener consistency* ($R_s$), and *matching* ($G$). Performer consistency refers to the multiple correlation of the performer model (i.e. performer's intention and cues), whereas listener consistency refers to the multiple correlation of the listener model (i.e. listener's judgment and cues). Both indices reflect the extent to which the regression models fit the cue utilization and are usually interpreted as measures of consistency, or

predictability, of the cue utilization. If $R = 1.0$, then the cue utilization is perfectly consistent. Matching ($G$) is a measure of the extent to which the beta weights of performers and listeners are 'matched' to each other; that is, whether they use the same code. This index is obtained by correlating the predicted values of the performer's regression model with the predicted values of the listener's regression model. The resulting correlation is interpreted as the extent to which the performer's beta weights and the listener's beta weights would agree if both regression models were perfect (i.e. $R_e = R_s = 1.0$).

$$r_a = G R_e R_s + C\sqrt{(1-R^2_e)}\sqrt{(1-R^2_s)} \quad (14.1)$$

If the communication under study is unsuccessful, we may ask whether this is because: (a) the performers use a different code than the listeners (indicated by a low $G$ value); (b) the performers apply their code inconsistently (indicated by a low $R_e$ value); or (c) the listeners apply their code inconsistently (indicated by a low $R_s$ value). These three factors set the upper limit of achievement (Hursch *et al.* 1964). By analysing each of them separately, it becomes possible to see how the communicative process could be improved (Juslin & Laukka 2000).

The second component of the LME is usually called the *unmodelled* component of the communicative process. It includes both unsystematic and systematic variance not accounted for by the linear component. This includes effects of inconsistent cue utilization, order effects, distractions, memory intrusions, omission of relevant cues, or *configural cue utilization* (i.e. the use of particular patterns of cue values). $(1-R^2_e)$ and $(1-R^2_s)$ refer to the residual variance of the regression models of performers and listeners, respectively. $C$—or *unmodelled matching*—represents the correlation between the residuals of the performer's model and the residuals of the listener's model. If $C$ is high it indicates: (a) a common reliance on cues in the performance not included in the regression models; (b) chance agreement between random model errors; (c) cue interactions common to both models; or (d) non-linear cue function forms common to both models (Cooksey 1996). However, research indicates that the unmodelled matching of the cue utilization is relatively small in music performance (Juslin & Madison 1999).

Analyses based on the lens model paradigm have generated a number of findings. First, linear regression models provide a good fit to the cue utilization, explaining approximately 70–80 per cent of the variance in expression and judgment (see Juslin 1997*b*, 2000). Second, achievement seems to depend mainly on the extent to which the cue utilization of the performer is 'matched' to the cue utilization of the listener (at least as regards professional performers; amateur performers may also differ with regard to the consistency of their cue utilization) (Juslin & Laukka 2000). Third, cue utilization is more consistent across different pieces of music than across different performers. That is, performers use pretty much the same code regardless of the melody, but there are considerable individual differences between performers. Fourth, two performers can communicate equally well despite slight differences in their cue utilization (see Juslin 2000). Finally, it has been shown that the lens model paradigm can be used to improve performers' expressive skills through so-called 'cognitive feedback' (Juslin & Laukka 2000).

## 14.7 Implications and directions for future research

Viewing emotional communication in music through a Brunswikian 'lens' has a number of important implications. First, previous studies have tended to focus on either performance or perception. The Brunswikian lens model provides the tool for an ultimately more interesting enterprise: how to *relate* performance to perception. Each of the two aspects can only be fully understood in relation to the other. Second, future studies should consider a whole set of cues to deal with the complete code used and thereby permit both performers and listeners to show their capacity for vicarious functioning. Third, future studies should analyse sufficiently large samples of performers, listeners, and pieces of music to allow for a greater generality and the use of multivariate methods, such as MRA. This recommendation parallels Brunswik's own insistence that we should deal with the phenomenon under study 'on an adequate level of complexity' (1952, p. 1). Lens modelling is an attempt to capture the relationships among *subjective* intentions or judgments and *physical* characteristics of music performances, and can therefore be conceived of as 'emotional psychophysics'.

### 14.7.1 Limitations of previous research

There are a number of limitations of previous research, which provide incentives for further research. First, studies of emotional communication in music performance have mainly been conducted in the laboratory. For obvious reasons, scientists have striven to obtain internally valid results using experimental methods. However, it is now high time to bring the research from the laboratory into the 'field'. We still have limited knowledge about how performers communicate emotions to listeners in real life. How reliable is a performer's communication under realistic conditions? Ways of getting closer to 'ecologically valid' settings might be to analyse existing recordings of performances (Rapoport 1996; Siegwart & Scherer 1995) or to observe performers, from the preparation of a piece to the actual performance in a concert.

Second, the research summarized in this chapter has largely rested on the assumption that the encoding process proceeds more or less independently of the musical structure, but this assumption may be questioned. Performers usually modulate their cue utilization to suit specific compositions, hence there is some inconsistency in cue utilization across different pieces of music (Juslin 2000). Future research should focus on the *interaction* between the composed structure and the performance in determining the expression of a piece of music. The studies of emotional expression in performance in the 1990s measured performance effects separately from structural effects. Now, after 10 years of research, we may be ready to start considering composition and performance together; to study the complete process. One way of doing this could be to expand the lens model further to also include cues associated with the structure of particular pieces, such as rhythm, mode, melodic contour, and tonal progression.

A number of researchers have raised the question whether performances recorded under laboratory conditions produce very different results than performances recorded under natural conditions, such as a concert (Rapoport 1996). Because we lack data on this issue, the jury is still out. However, it seems plausible that similar patterns of effects will be found in realistic settings, but that the effects will be smaller. One reason for this

may be that performers often have to coordinate their own playing with that of other performers. This places constraints on their freedom of expression. In addition, the social dynamics within orchestras or bands may influence the expression of the music in ways that merit further research (Davidson 1997).

Only a few studies have explored emotional expression in natural music performances. Siegwart and Scherer (1995) investigated emotional expression in opera singing by analysing existing recordings by famous artists. They reported that: (a) different interpretations elicited significantly different listener ratings of emotional expressiveness; (b) voice samples differed substantially with respect to acoustic cues; and (c) listeners' judgments could be successfully predicted on the basis of the acoustic cues. These are important results, because they suggest that the effects reviewed in this chapter are not confined to the laboratory.

### 14.7.2 Directions for future research: the case of emotion induction

Thus far in the chapter, I have primarily been concerned with how emotion is *represented* in music performance. This simply reflects the fact that most studies to date have investigated representation of emotion—perhaps because *induction* of emotion is more difficult to study. In everyday life, however, the question of why particular performances of music manage to induce strong emotions in listeners is just as important. In this section, I suggest some ideas with respect to how expressive performances might induce emotional reactions in listeners. These ideas are admittedly speculative, but may serve as directions for future research.

First of all, it is possible that the expression of a performance may arouse emotions in listeners through a process of so-called *emotional contagion*, which has been defined as 'the tendency to automatically mimic and synchronize facial expressions, vocalizations, postures, and movements with those of another person and, consequently, to converge emotionally . . .' (Hatfield *et al.* 1994, p. 5). There is a large amount of evidence that people easily 'catch' the emotions of others when watching their facial expressions or hearing their vocal expressions, presumably through some kind of 'motor mimicry'. The fact that certain aspects of emotional expression in music are highly similar to vocal expression of emotion suggests that we get emotionally aroused by the voice-like aspects of musical performances (Juslin, in press).

One objection to this theory could be that most performances do not sound very much like vocal expressions, at least superficially (Budd 1985, Chapter 7; see also Davies, this volume). Why, then, should we react to performances as if they were vocal expressions? One explanation may be that emotional expression is processed by simple, automatic, and independent 'brain modules', which react to certain features in the stimulus (Fodor 1983). Such modules do not 'know' the difference between vocal expression of emotion and other acoustic expression but will react in the same way as long as certain cues (e.g. high speed, loud dynamics, fast attack, many overtones) are present in the stimulus (Juslin, in press). Evidence that we indeed react emotionally to performances of music as we do to vocal expressions comes from studies that use facial electromyography and self-reports to measure emotions (see Hietanen *et al.* 1998; Lundqvist *et al.* 2000; Neumann & Strack 2000; Witvliet & Vrana 1996; Witvliet *et al.* 1998). This modular view of perception of emotion in music is supported by evidence that processing of information about

emotion is largely implicit (Niedenthal & Showers 1991), that judgments of emotion in music are quick, and that researchers have obtained brain dissociations between judgments of musical emotion and of musical structure (Peretz *et al.* 1998).

A second way in which a particular performance might arouse emotions in its listeners involves the fact that emotional reactions often occur when musical expectations are violated in some way (Meyer, this volume). It is easy to imagine this happening with regard to music performances. Meyer (1956, p. 206) suggested that expressive variations in a musical performance may serve an aesthetic function by 'delaying an expected resolution', or otherwise 'creating psychological tension'. It should be observed that such effects may be achieved in different ways. For instance, a performer may enhance listeners' emotional responses to the music by emphasizing notes that are of particular 'significance' in the composition, thereby enhancing violations of musical expectations that are already latent in the structure. Support for this idea has been obtained in recent studies (e.g. Sloboda & Lehmann, in press).

However, a performer can also arouse emotions in listeners by performing in a manner that deviates from stylistic expectations with regard to the performance of a certain structure. In this case, the deviation does not serve to enhance significant aspects of the structure, but rather to create psychological tension through deviations from performance conventions. It should be noted that the most 'extreme' performances of classical music are often made by the most well-known performers (cf. Sloboda 1998). In some cases, expressive devices of a performance may have a direct impact on the physiological processes of the listener. Harrer and Harrer (1977), for example, have claimed that it is possible to 'drive' the pulse rate of a listener by dynamic changes in volume (e.g. the crescendo or decrescendo of a rolling drum beat). Moreover, sudden syncopations in a musical performance may be capable of producing premature heart beats. The use of such effects by music performers has been described in the vast literature on rubato. For instance, the famous virtuoso violinist Niccolò Paganini (1782–1840) is known to have used tempo rubato to great effect (Hudson, 1994).

## 14.8 Conclusion

In this chapter, we have seen that performers can communicate specific emotions to listeners by using a number of probabilistic but partly redundant cues in the performance. The relative importance of the performance for the expression of the music is likely to differ depending on the performer, genre, and piece of music—ranging all the way from the faithful rendition of a piece of avant-garde music to the heartbreaking, improvised solo of a blues number.

The vicarious functioning of cues allows performers to simultaneously communicate emotions to listeners in a universally accessible manner and develop a personal expression. This communicative system therefore constitutes the happy marriage of the general and the individual, of the biological and the cultural. The communicative process is fairly robust but has a limited information capacity, ultimately leaving it to the listeners to specify the precise 'meaning' of the music. The phylogenetic roots of the communicative system can explain its powerful impact, and also why many performers regard emotions as 'the essence of musical communication' (Persson 1995, p. 60).

The subtle emotional nuances of music may seem a distant cry from the evolutionary framework of 'basic emotions' that I have described in this chapter. But I would argue that, ultimately, the real force of music's expressiveness—if not its uttermost subtlety—resides in such roots. In writing about the origins of music, Rousseau (1986, p. 65) said:

> . . . how far from understanding the power of their art are those many musicians who think of the potency of sounds only in terms of their air pressure and string vibrations. The more they assimilate it to purely physical impressions, the farther they get from its source and the more they deprive it of its primitive energy. In dropping the oral tone and sticking exclusively to the . . . harmonics, music becomes nosier to the ear and less pleasing to the heart. As soon as it stops singing, it stops speaking. And then, with all its accord and all its harmony it will have no more effect upon us.

Thus, the music that touches us the most does not necessarily come from someone who 'plays and sings with the grace of a God', but from someone 'bold enough to act as a messenger of the heart' (Schumacher 1995, p. 314).[1]

## References


Adachi, M. & Trehub, S. E. (1998). Children's expression of emotion in song. *Psychology of Music*, 26, 133–53.

Balkwill, L-L. & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception*, 17, 43–64.

Baroni, M. & Finarelli, L. (1994). Emotions in spoken language and vocal music. In *Proceedings of the Third International Conference for Music Perception and Cognition*, (ed. I. Deliège), pp. 343–5. Liége, Belgium: University of Liége.

Baroni, M., Caterina, R., Regazzi, F., & Zanarini, G. (1997). Emotional aspects of singing voice. In *Proceedings of the Third Triennial ESCOM Conference*, (ed. A. Gabrielsson), pp. 484–9. Uppsala, Sweden: Department of Psychology, Uppsala University.

Behrens, G. A. & Green, S. B. (1993). The ability to identify emotional content of solo improvisations performed vocally and on three different instruments. *Psychology of Music*, 21, 20–33.

Blum, D. (1977). *Casals and the art of interpretation*. Berkeley, CA: University of California Press.

Borod, J. C., Obler, L. K., Erhan, H. M., Grunwald, I. S., Cicero, B. A., Welkowitz, J. *et al.* (1998). Right hemisphere emotional perception: Evidence across multiple channels. *Neuropsychology*, 12, 446–58.

Bresin, R. & Friberg, A. (2000). Emotional coloring of computer-controlled music performance. *Computer Music Journal*, 24, 44–62.

Brunswik, E. (1952). *The conceptual framework of psychology*. Chicago: Chicago University Press.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley, CA: University of California Press.

Bryden, M. P., Ley, R., & Sugerman, J. (1982). A left-ear advantage for identifying the emotional quality of tonal sequences. *Neuropsychologia*, 20, 83–7.


[1] I am grateful to Bruno Repp, John Rink, and John Sloboda for helpful comments on a preliminary version of this chapter. The writing of this chapter was supported by the Bank of Sweden Tercentenary Foundation.

**Buck, R.** (1984). *The communication of emotion.* New York: Guilford Press.

**Budd, M.** (1985). *Music and the emotions: The philosophical theories.* London: Routledge.

**Clapton, E. P.** (1998). *Official tour program 1998.* New York: Warner Brothers.

**Clarke, E. F.** (1988). Generative principles in music performance. In *Generative processes in music. The psychology of performance, improvisation, and composition,* (ed. J. A. Sloboda), pp. 1–26. Oxford, UK: Clarendon Press.

**Clynes, M.** (1973). Sentics: Biocybernetics of emotion communication. *Annals of the New York Academy of Sciences,* 220, 55–131.

**Clynes, M.** (1977). *Sentics: The touch of emotions.* New York: Doubleday.

**Cooksey, R. W.** (1996). *Judgment analysis.* New York: Academic Press.

**Cunningham, J. G. & Sterling, R. S.** (1988). Developmental changes in the understanding of affective meaning in music. *Motivation and Emotion,* 12, 399–413.

**Darwin, C.** (1872). *The expression of the emotions in man and animals.* London: John Murray.

**Davidson, J. W.** (1997). The social in musical performance. In *The social psychology of music,* (ed. D. J. Hargreaves & A. C. North), pp. 209–28. Oxford, UK: Oxford University Press.

**Davies, S.** (1994). *Musical meaning and expression.* Ithaca, NY: Cornell University Press.

**Dawes, R. M. & Corrigan, B.** (1974). Linear models in decision making. *Psychological Bulletin,* 81, 95–106.

**de Gelder, B. & Vroomen, J.** (1996). Categorical perception of emotional speech. *Journal of the Acoustical Society of America,* 100, 2818.

**Dolgin, K. & Adelson, E.** (1990). Age changes in the ability to interpret affect in sung and instrumentally-presented melodies. *Psychology of Music,* 18, 87–98.

**Eggebrecht, R.** (1983). *Sprachmelodie und musikalische Forschungen im Kulturvergleich.* Doctoral dissertation, University of Munich, Germany.

**Eibl-Eibesfeldt, I.** (1989). *Human ethology.* New York: Aldine.

**Ekman, P.** (ed.) (1973). *Darwin and facial expression.* New York: Academic Press.

**Ekman, P.** (1992). An argument for basic emotions. *Cognition and Emotion,* 6, 169–200.

**Fairbanks, G.** (1940). Recent experimental investigation of vocal pitch in voice. *Journal of the Acoustical Society of America,* 11, 457–66.

**Fisk, J.** (ed.) (1997). *Composers on music.* Boston, MA: Northeastern University Press.

**Fodor, J. A.** (1983). *The modularity of the mind.* Cambridge, MA: MIT Press.

**Frick, R. W.** (1985). Communicating emotion: The role of prosodic features. *Psychological Bulletin,* 97, 412–29.

**Gabrielsson, A.** (1988). Timing in music performance and its relations to music experience. In *Generative processes in music: The psychology of performance, improvisation, and composition,* (ed. J. A. Sloboda), pp. 27–51. Oxford, UK: Clarendon Press.

**Gabrielsson, A.** (1999). The performance of music. In *The psychology of music,* (2nd edn) (ed. D. Deutsch). pp. 501–602. San Diego, CA: Academic Press.

**Gabrielsson, A. & Juslin, P. N.** (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music,* 24, 68–91.

**Gabrielsson, A. & Juslin, P. N.** (in press). Emotional expression in music. In *Handbook of affective sciences,* (ed. R. J. Davidson, H. H. Goldsmith, & K. R. Scherer). New York: Oxford University Press.

**Gabrielsson, A. & Lindström, E.** (1995). Emotional expression in synthesizer and sentograph performance. *Psychomusicology,* 14, 94–116.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton-Mifflin.

Gorman, B. S. & Crain, W. C. (1974). Decoding of 'sentograms'. *Perceptual and Motor Skills,* **39,** 784–6.

Gregory, A. H. (1997). The roles of music in society: The ethnomusicological perspective. In *The social psychology of music,* (ed. D. J. Hargreaves & A. C. North), pp. 123–40. Oxford, UK: Oxford University Press.

Handel, S. (1991). *Listening. An introduction to the perception of auditory events.* Cambridge, MA: MIT Press.

Harrer, G. & Harrer, H. (1977). Music, emotion, and autonomic function. In *Music and the brain. Studies in the neurology of music,* (ed. M. Critchley & R. A. Henson), pp. 202–16. London: William Heinemann Medical Books.

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion.* New York: Cambridge University Press.

Helmholtz, H. L. F. von (1954). *On the sensations of tone as a psychological basis for the theory of music.* New York: Dover (originally published 1863).

Hietanen, J. K., Surakka, V., & Linnankoski, I. (1998). Facial electromyographic responses to vocal affect expressions. *Psychophysiology,* 35, 530–6.

Hudson, R. (1994). *Stolen time. The history of tempo rubato.* Oxford, UK: Clarendon Press.

Hursch, C. J., Hammond, K. R., & Hursch, J. L. (1964). Some methodological considerations in multiple-cue probability studies. *Psychological Review,* **71,** 42–60.

Izard, C. E. (1977). *The emotions.* New York: Plenum Press.

Johnson-Laird, P. N. & Oatley, K. (1992). Basic emotions, rationality, and folk theory. *Cognition and Emotion,* 6, 201–23.

Jürgens, U. & von Cramon, D. (1982). On the role of the anterior cingulate cortex in phonation: A case report. *Brain and Language,* **15,** 234–48.

Juslin, P. N. (1995). Emotional communication in music viewed through a Brunswikian lens. In *Musical expression. Proceedings of the Conference of ESCOM and DGM 1995,* (ed. G. Kleinen), pp. 21–5. Bremen, Germany: University of Bremen.

Juslin, P. N. (1997*a*). Emotional communication in music performance: A functionalist perspective and some data. *Music Perception,* **14,** 383–418.

Juslin, P. N. (1997*b*). Perceived emotional expression in synthesized performances of a short melody: Capturing the listener's judgment policy. *Musicae Scientiae,* **1,** 225–56.

Juslin, P. N. (1997*c*). Can results from studies of perceived expression in musical performances be generalized across response formats? *Psychomusicology,* **16,** 77–101.

Juslin, P. N. (1998). A functionalist perspective on emotional communication in music performance. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 78.* Uppsala, Sweden: Uppsala University Library.

Juslin, P. N. (1999). *Communication of emotion in vocal expression and music performance: Different channels, same code?* Manuscript submitted for publication.

Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance,* 26, 1797–813.

Juslin, P. N. (2001). A Brunswikian approach to emotional communication in music performance. In *The essential Brunswik: Beginnings, explications, applications,* (ed. K. R. Hammond & T. R. Stewart) pp. 426–30. New York: Oxford University Press.

Juslin, P. N. (in press). Vocal expression and musical expression: Parallels and contrasts. In *Proceedings of the Eleventh Meeting of the International Society for Research on Emotions.*, (ed. A. Kappas). Quebec, Canada: ISRE Publications.

Juslin, P. N., Friberg, A., & Bresin, R. (in press). Toward a computational model of expression in music performance: The GERM model. *Musicae Scientiae.*

Juslin, P. N. & Laukka, P. (2000). Improving emotional communication in music performance through cognitive feedback. *Musicae Scientiae,* 4, 151–83.

Juslin, P. N. & Laukka, P. (in press). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. In *Proceedings of the 11th Meeting of the International Society for Research on Emotions,* (ed. A. Kappas). Quebec, Canada: ISRE Publications.

Juslin, P. N. & Madison, G. (1999). The role of timing patterns in recognition of emotional expression from musical performance. *Music Perception,* 17, 197–221.

Juslin, P. N. & Persson, R. S. (in press). Emotional communication. In *The science and psychology of music performance. Creative strategies for teaching and learning,* (ed. R. Parncutt & G. E. McPherson). New York: Oxford University Press.

Kendall, R. A. & Carterette, E. C. (1990). The communication of musical expression. *Music Perception,* 8, 129–64.

Kennedy, N. (1990). *Always playing.* London: Weidenfield & Nicolson.

King, B. B. (1996). *Blues all around me.* London: Hodder & Stoughton.

Kivy, P. (1989). *Sound sentiment.* Philadelphia, PA: Temple University Press.

Konishi, T., Imaizumi, S., & Niimi, S. (2000). Vibrato and emotion in singing voice (abstract). In *Proceedings of the Sixth International Conference on Music Perception and Cognition, August 2000,* (ed. C. Woods, G. Luck, R. Brochard, F. Seddon, & J. A. Sloboda), (CD rom). Keele, UK: Keele University.

Kotlyar, G. M. & Morozov, V. P. (1976). Acoustic correlates of the emotional content of vocalized speech. *Soviet Physics. Acoustics,* 22, 370–6.

Kratus, J. (1993). A developmental study of children's interpretation of emotion in music. *Psychology of Music,* 21, 3–19.

Ladd, D. R., Silverman, K. E. A., Tolkmitt, F., Bergmann, G., & Scherer, K. R. (1985). Evidence of independent function of intonation contour type, voice quality, and $F_0$ range in signaling speaker affect. *Journal of the Acoustical Society of America,* 78, 435–44.

Langeheinecke, E. J., Schnitzler, H.-U., Hischer-Buhrmeister, M., & Behne, K.-E. (1999). Emotions in singing voice: Acoustic cues for joy, fear, anger, and sadness. Poster presented at the Joint Meeting of the Acoustical Society of America and the Acoustical Society of Germany, Berlin, March 1999.

Lindström, E. (1999). Expression in music: Interaction between performance and melodic structure. Paper presented at the Meeting of the Society for Music Perception and Cognition, Evanston, USA, 14–17 August 1999.

Lundqvist, L. G., Carlsson, F., & Hilmersson, P. (2000). *Facial electromyography, autonomic activity, and emotional experience to happy and sad music.* Paper presented at the 27th International Congress of Psychology, Stockholm, Sweden, 23–28 July, 2000.

MacLean, P. D. (1978). Effects of lesions of globus pallidus on species-typical display behavior of squirrel monkeys. *Brain Research,* 149, 175–96.

Meyer, L. B. (1956). *Emotion and meaning in music.* Chicago: Chicago University Press.

Nettelbeck, T., Henderson, C., & Willson, R. (1989). Communicating emotion through sound: An evaluation of Clynes' theory of sentics. *Australian Journal of Psychology*, **41**, 17–24.

Neumann, R. & Strack, F. (2000). Mood contagion: The automatic transfer of mood between persons. *Journal of Personality and Social Psychology*, **79**, 211–23.

Niedenthal, P. M. & Showers, C. (1991). The perception and processing of affective information and its influences on social judgment. In *Emotion and social judgments*, (ed. J. P. Forgas), pp. 125–43. Oxford, UK: Pergamon Press.

Oatley, K. (1992). *Best laid schemes*. Cambridge, MA: Harvard University Press.

Oatley, K. & Jenkins, J. M. (1996). *Understanding emotions*. Oxford, UK: Blackwell.

Ohgushi, K. & Hattori, M. (1996*a*). Emotional communication in performance of vocal music. In *Proceedings of the Fourth International Conference on Music Perception and Cognition*, (ed. B. Pennycook & E. Costa-Giomi), pp. 269–74. Montreal, Canada: McGill University.

Ohgushi, K. & Hattori, M. (1996*b*). Acoustic correlates of the emotional expression in vocal performance. Paper presented at the Third Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, Honolulu, Hawaii, 2–6 December 1996.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 331–46.

Palmer, C. (1997). Music performance. *Annual Review of Psychology*, **48**, 115–38.

Panksepp, J. (1982). Towards a general psychobiological theory of emotions. *Behavioral and Brain Sciences*, 5, 407–67.

Papoušek, M. (1996). Intuitive parenting: A hidden source of musical stimulation in infancy. In *Musical beginnings. Origins and development of musical competence*, (ed. I. Deliége & J. A. Sloboda), pp. 89–112. Oxford, UK: Oxford University Press.

Papoušek, M., Bornstein, M. H., Nuzzo, C., Papoušek, H., & Symmes, D. (1990). Infant responses to prototypical melodic contours in parental speech. *Infant Behavior and Development*, **13**, 539–45.

Patel, A. D. & Peretz, I. (1997). Is music autonomous from language? A neuropsychological appraisal. In *Perception and cognition of music*, (ed. I. Deliége & J. A. Sloboda), pp. 191–215. Hove, UK: Psychology Press.

Penfield, W. & Roberts, L. (1959). *Speech and brain mechanisms*. Princeton, NJ: Princeton University Press.

Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, 68, 111–41.

Persson, R. S. (1995). Musical reality: Exploring the subjective world of performers. In *Song and signification. Studies in music semiotics*, (ed. R. Monelle & C. T. Gray), pp. 58–63. Edinburgh, UK: Faculty of Music, University of Edinburgh.

Persson, R. S., Pratt. G., & Robson, C. (1996). Motivational and influential components of musical performance: A qualitative analysis. In *Fostering the growth of high ability: European perspectives*, (ed. A. J. Cropley & D. Dehn), pp. 287–301. Norwood, NJ: Ablex.

Ploog, D. (1981). Neurobiology of primate audio-vocal behavior. *Brain Research Reviews*, 3, 35–61.

Ploog, D. (1986). Biological foundations of the vocal expressions of emotions. In *Emotion. Theory, research, and experience. Volume 3: Biological foundations of emotion*, (ed. R. Plutchik & H. Kellerman), pp. 173–97. New York: Academic Press.

Ploog, D. (1992). The evolution of vocal communication. In *Nonverbal communication*, (ed. H. Papoušek, U. Jürgens, & M. Papoušek), pp. 6–30. Cambridge, UK: Cambridge University Press.

**Plutchik, R.** (1994). *The psychology and biology of emotion.* New York: Harper-Collins.

**Rapoport, E.** (1996). Emotional expression code in opera and lied singing. *Journal of New Music Research,* 25, 109–49.

**Repp, B. H.** (1998). A microcosm of musical expression. I: Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America,* 102, 1085–99.

**Rosenthal, R.** (1982). Judgment studies. In *Handbook of methods in nonverbal behavior research,* (ed. K. R. Scherer & P. Ekman), pp. 287–361. Cambridge, UK: Cambridge University Press.

**Ross, B. H. & Spalding, T. L.** (1994). Concepts and categories. In *Thinking and problem solving,* (2nd edn) (ed. R. J. Sternberg), pp. 119–50. New York: Academic Press.

**Rousseau, J. J.** (1986). Essay on the origin of languages. In *On the origin of language: Two essays,* (ed. J. H. Moran & A. Gode), pp. 5–74. Chicago: University of Chicago Press (originally published 1761).

**Russell, J. A.** (1980). A circumplex model of affect. *Journal of Personality and Social Psychology,* 39, 1161–78.

**Scherer, K. R.** (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin,* **99**, 143–65.

**Scherer, K. R.** (1995). Expression of emotion in voice and music. *Journal of Voice,* 9, 235–48.

**Scherer, K. R. & Oshinsky, J. S.** (1977). Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion,* **1**, 331–46.

**Schumacher, M.** (1995). *Crossroads. The life and music of Eric Clapton.* New York: Hyperion.

**Seashore, C. E.** (1947). *In search of beauty in music. A scientific approach to musical aesthetics.* Westport, CT: Grenwood Press.

**Seashore, H. G.** (1937). An objective analysis of artistic singing. In *Objective analysis of musical performance: University of Iowa studies in the psychology of music,* (Vol. 4) (ed. C. E. Seashore), pp. 12–157. Iowa City, IA: University of Iowa.

**Seckerson, E.** (1991). Yuri Bashmet as interviewed by Seckerson. *Gramophone,* June, 26–7.

**Senju, M. & Ohgushi, K.** (1987). How are the player's ideas conveyed to the audience? *Music Perception,* 4, 311–24.

**Shaffer, L. H.** (1992). How to interpret music. In *Cognitive bases of musical communication,* (ed. M. R. Jones & S. Holleran), pp. 263–78. Washington, DC: American Psychological Association.

**Shaffer, L. H. & Todd, N.** (1987). The interpretative component in musical performance. In *Action and perception in rhythm and music,* (ed. A. Gabrielsson), pp. 139–52. Royal Swedish Academy of Music Publication No. 55. Stockholm: Royal Swedish Academy of Music.

**Shapiro, B. E. & Danly, M.** (1985). The role of the right hemisphere in the control of speech prosody in propositional and affective contexts. *Brain and Language,* 25, 19–36.

**Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C.** (1987). Emotion knowledge: Further explorations of a prototype approach. *Journal of Personality and Social Psychology,* 52, 1061–86.

**Shields, S. A.** (1984). Distinguishing between emotion and non-emotion: Judgments about experience. *Motivation and Emotion,* 8, 355–69.

**Siegwart, H. & Scherer, K. R.** (1995). Acoustic concomitants of emotional expression in operatic singing: The case of Lucia in Ardi gli incensi. *Journal of Voice,* 9, 249–60.

**Sloboda, J. A.** (1996). The acquisition of musical performance expertise: Deconstructing the 'talent' account of individual differences in musical expressivity. In *The road to excellence,* (ed. K. A. Ericsson ), pp. 107–26. Mahwah, NJ: Erlbaum.

**Sloboda, J. A.** (1998). Musical performance and emotion: Issues and developments. In *Proceedings of the Fifth International Conference on Music Perception and Cognition*, (ed. S. W. Yi), pp. 21–5. Seoul, Korea: Seoul National University.

**Sloboda, J. A. & Lehmann, A. C.** (in press). Tracking performance correlates of changes of perceived intensity of emotion during different interpretations of a Chopin piano prelude. *Music Perception.*

**Sundberg, J.** (1982). Speech, song, and emotions. In *Music, mind, and brain. The neuropsychology of music*, (ed. M. Clynes), pp. 137–49. New York: Plenum Press.

**Sundberg, J.** (1988). Computer synthesis of music performance. In *Generative processes in music. The psychology of performance, improvisation, and composition*, (ed. J. A. Sloboda), pp. 52–69. Oxford, UK: Clarendon Press.

**Sundberg, J., Iwarsson, J., & Hagegård, H.** (1995). A singer's expression of emotions in sung performance. In *Vocal fold physiology: Voice quality control*, (ed. O. Fujimura & M. Hirano), pp. 217–29. San Diego, CA: Singular Press.

**Tait, M.** (1992). Teaching strategies and styles. In *Handbook of research on music teaching and learning*, (ed. R. Colwell), pp. 525–34. New York: Schirmer Books.

**Terwogt, M. M. & van Grinsven, F.** (1988). Musical expression of mood states. *Psychology of Music*, **19**, 99–109.

**Trussoni, S. J., O'Malley, A., & Barton, A.** (1988). Human emotion communication by touch: A modified replication of an experiment by Manfred Clynes. *Perceptual and Motor Skills*, **66**, 419–24.

**Wallbott, H. G. & Scherer, K. R.** (1986). Cues and channels in emotion recognition. *Journal of Personality and Social Psychology*, **51**, 690–9.

**Whissell, C. M.** (1989). The dictionary of affect in language. In *Emotion: Theory, research, and experience: Vol. 4. The measurement of emotions*, (ed. R. Plutchik & H. Kellerman), pp. 113–31. New York: Academic Press.

**Witvliet, C. V. & Vrana, S. R.** (1996). The emotional impact of instrumental music on affect ratings, facial EMG, autonomic responses, and the startle reflex: Effects of valence and arousal. *Psychophysiology Supplement*, **91**.

**Witvliet, C. V., Vrana, S. R., & Webb-Talmadge, N.** (1998). In the mood: Emotion and facial expressions during and after instrumental music, and during an emotional inhibition task. *Psychophysiology Supplement*, **88**.

**Woody, R.** (2000). Learning expressivity in music performance: An exploratory study. *Research Studies in Music Education*, **14**, 14–23.

PART V

# THE LISTENER

CHAPTER 15

# MUSIC AND EMOTION: DISTINCTIONS AND UNCERTAINTIES

LEONARD B. MEYER

I have the gall to divide this chapter into only *two* parts:[1] 'Distinctions' and 'Uncertainty'. The first section considers conceptual/terminological problems having to do with music and emotion, and provides the background and terminology for the matters considered in the second part. The distinctions made in Section 15.1 do not pretend to constitute a coherent argument but are connected by the succession of topics discussed. Section 15.2, the core of the chapter, is concerned with the role of uncertainty in human experience generally, and in musical/aesthetic experience in particular.

## 15.1 Distinctions

I begin by proposing a number of distinctions because, despite its empirical outlook and methodology, the psychological analysis of human emotional experiences—causes and responses—as well as humanistic theory have lacked conceptual specificity and precision. This is evident in the range of behavioural states considered 'emotional'. To take a representative example, Klaus Scherer (1994, p. 31) writes that some terms for emotions 'highlight the physiological, as in the case of "aroused" or "tired." Some are quite cognitive, like "bewildered" or "curious".'

It seems curious, however, to consider 'curious' or 'tired' as being emotions. Such 'states of being'—of awareness (e.g. 'that's an interesting idea') or of feeling (e.g. 'I'm hungry')—are not generally classed as emotions.[2] But are not feelings of sadness, joy, anger, etc., which *are* considered emotions, similarly states of being? So the question is: what are the necessary and sufficient psychological and/or physiological conditions for considering a state of being—a feeling—an emotion?[3]

[1] Since my role is supposed to be that of the 'resident humanist', I have taken the liberty of playfully paraphrasing the first sentence of Caesar's *Gallic wars*: 'All Gaul is divided into three parts'.

[2] If tiredness is disturbed or curiosity thwarted, emotion may be evoked; but the 'evoking agent' is the disturbance, the thwarting, and the uncertainty this produces, not the tiredness or the curiosity.

[3] In the course of thinking about these matters, a possible distinction occurred to me: namely, that states of being such as curiosity, hunger, fatigue, cold, etc. depend upon conditions *internal* to an individual, while love and hate, fear and anger, etc. also include conditions *external* to an individual.

Though, to the best of my knowledge, there has been no satisfactory answer to that question, we go on using the concept of emotion and distinguishing different kinds: love and hate, fear and courage, etc. We conceptualize and classify our experiences of the inner as well as the outer world because envisaging and choosing, which lie at the heart of human behaviour, occur primarily in terms of such classes rather than individual instances. So this part of the chapter attempts to distinguish some of the kinds of relationships involved in the interaction between music and emotion.

### 15.1.1 Emotional states versus emotional processes

*States*

*Delineation and association: statistical parameters.* As a rule, the terms used to characterize emotions denote what I will call emotional 'states'. When listeners or critics describe music as sad, happy, angry, elated, and so on, they are referring to such states. In music, such states are essentially delineated by the action of what I have called the 'statistical' parameters. These aspects of sound vary in *amount* or *degree*—for example, register (lower–higher), dynamic level (louder–softer), speed (faster–slower), continuity (gradual–abrupt), and so forth. Although the configuration of parameters delineating an emotional state fluctuates, state-defining parameters such as register, speed, and dynamic level remain relatively constant over moderate-sized segments of music. And it is this constancy that makes *naming* possible. Indeed, when subjects in psychological experiments identify emotions their answers are, for the most part, based on the disposition of statistical parameters. Put negatively, one cannot imagine sadness being portrayed by a fast forte tune played in a high register, or a playful child being depicted by a solemnity of trombones.

The identification and naming of states is a result of the association of different dispositions of statistical parameters with feelings characteristically evoked in natural or cultural *circumstances*. That is, we deem a passage to portray 'love' or 'anger' or 'jealousy' not because of knowledge of physiological responses or psychological constraints, but because of awareness of, and often empathy with, the feelings and behaviours that occur in particular sociocultural contexts.[4] Thus we recognize the sadness expressed in Macbeth's 'Tomorrow, and tomorrow, and tomorrow' speech because we know the circumstance in which it is delivered—that is, Macbeth's learning of his wife's death. In addition, of course, his feelings are communicated by the repetition of the word 'tomorrow', which in this context signifies sadness and depression, by his performance (slow, soft tone of voice, etc.), and his physical behaviour and facial expression.[5] If this view

[4] This calls attention to a methodological question: namely, when listeners record their responses to music by, say, pressing buttons or naming emotions, or when psychologists measure the physiological changes (in blood pressure, pulse rate, etc.) of listeners, what do the 'results' represent? Often it seems unclear whether the answers are a result of the subjects' *recognition* of the 'state' delineated by the music or of their own affective *response* to the musical patterning? In this connection, see Sloboda (1992).

[5] Though overt behaviour may be the result of an innate response, behaviour is usually influenced by cultural convention. This is the case because behaviour almost always acts as a sign that is implicative

has merit, then observed behavioural circumstances play an important role in the identification and classification of emotional states.

*Classification.* It is sometimes supposed that our understanding of the world begins with the peculiarities of the individual—person, work of art, landscape, and so on. But, though experiences of phenomena may be peculiar, the peculiar is almost always qualified by, and comprehended in terms of, a conceptual class. This is one of the reasons why understanding and responding to different styles of music depends on learning. For it is largely learning that enables composers, performers, and listeners to internalize the norms governing musical relationships and, consequently, to comprehend the implications of patternings, and experience deviations from stylistic norms.

Just as we classify the phenomenal world into physical and biological realms, styles of music into classic and romantic, folk and jazz, and phrase relationships into antecedent-consequent, so we classify emotions. In most cases, the classifications are fuzzy. They are so because the classes are abstractions from individual, though similar, relationships and experiences. And one of the problems with abstractions is the tendency to treat them as though they possessed a set of objective properties independent of context and culture. But unlike tuberculosis, the biological nature of what we conceptualize as emotional experience has not been ascertained.

Nevertheless we classify emotions. We do so for the same reasons that we classify and categorize the world in general: because to choose effectively, it is necessary to imagine the probability of coming events (whether the behaviour of people or of musical patterns), and such imagining invariably entails classification. One manifestation of the need to classify is our proclivity to give 'to airy nothing/A local habitation and a *name*' (Shakespeare, *A midsummer night's dream*, V. i, emphasis added).[6] Put prosaically—because they threaten our sense of predictive power, nameless phenomena are highly disturbing.

Awareness of, and responding to, the class of some phenomenon need *not*, however, involve conscious conceptualization or naming but rather recognizing and experiencing the relationships presented—that is, knowing how the kind of patterning *works*. The point is important because, as mentioned above, conscious classification and explicit naming tend to undermine uncertainty and weaken emotional response. Something akin to 'the suspension of disbelief' (discussed below) seems at work here. That is, because enjoying music is a matter of experiencing the tension of syntactic and native processes as well as the fluctuations of statistical parameters, listeners generally avoid

(both to the individual and to others) of probable subsequent behaviour. Put in evolutionary terms, emotional behaviour was evidently preadapted for communication. There is a need for careful cross-cultural studies of emotional representation. Are the emotional states classified and represented in Western culture differentiated and recognized in non-Western cultures? If so, are they similarly represented?

[6] The fundamental importance of categorization (classification) was the subject of studies that appeared in *Science* (Vol. 291, 12 January 2001) after this essay was written. The result relevant for this essay is summarized in the section, 'This week in science': 'Categories are the fundamental building blocks of nearly all higher level cognitive functions' (p. 207).

consciously classifying and naming kinds of relationships, techniques, or strategies. Just as a reader enjoys a poem not by 'watching the metaphors go by', but by responding to the relationships implicit in the comparisons, so experienced listeners respond to the implicative tension of a gap-fill melody or an antecedent-consequent phrase without conceptualizing its relational class.

*Mapping versus relating.* The conceptualization of feelings as phenomena independent of experiential context has led to a search for the neurological basis of emotion. However, although the areas of the brain that respond to some order of stimuli have been located and mapped, neurologists have yet to explain the nature of *relationships* within or between such areas.

The situation is not unlike that of the astronomers of ancient Greece who were able to map the heavens with considerable precision, but could not *explain* the relationships between the stars, planets, and the like. Neurological studies seem in a comparable position with respect to the mind. The brain has been mapped, and the areas that respond to different kinds of stimuli have been located; and, as with the Greeks, various areas have been given appropriate names. But while the parts of the brain that respond to specific musical stimuli have been identified, the neurological nature of the relationships between such stimuli remains a mystery. For example, the places in the brain that respond to the pitch A (440) and the C# above it, and those stimulated by the pitch E (660) and the G# above it, can be identified; but why both intervals are perceived as belonging to the same relational class (i.e. major third) remains to be explained.

*Processes: native and syntactic*

As observed above, emotional states fluctuate in intensity; and human empathy with the action of the music begets analogous feelings in listeners.[7] But like other fluctuations (e.g. in temperature, energy, or the stock market) changes in intensity or amount do not give rise to processes—that is, to changes involving *functionally differentiated relationships.* And, as argued in Section 15.2 below, crucial facets of the emotional experience of music lie in the realm of processes rather than that of states.

Two kinds of musical processes, 'native' ones and 'syntactic' ones, need to be distinguished.[8] Both are subject to the cognitive constraints of the mind and the physical ones of the body. Syntactic processes, however, also involve learning the norms of a particular

[7] Some emotional states require more contextual specification than others. For instance, musical representations of grief and joy, anxiety and calm are readily recognized by competent listeners (see, for example, Gabrielsson & Lindström, this volume; Juslin, this volume); but, because their recognition is more context-dependent, emotions such as jealousy and disgust are more difficult to represent in music alone. Consequently, their representation is, I suspect, relatively uncommon and when it occurs it tends to be accompanied by a text—a vocal part or a programme.

[8] In *Style and music* (Meyer 1989), I emphasized the different roles of syntactic and statistical parameters in creating musical relationships. As will be clear from what follows, though I am keeping the term 'statistical parameters', I now want to use the concept of *process* in connection with syntactic and native relationships. Although implicit in my earlier work, the implication-realization relationship was first made explicit in *Explaining music* (Meyer 1973), Chapters VI–VIII. The relationship has since been thoughtfully analysed, extended, and modelled by Eugene Narmour (1990).

musical style. As a result, native and syntactic processes almost always complement one another in the shaping of musical experience. Those based on the syntax of a style tend to be quite specific, but are historically and culturally circumscribed. Native processes, on the other hand, are less specific. But, as limitations common to humankind, they are essentially universal.[9]

The very existence of sound entails the use of the statistical parameters. Consequently, native and syntactic processes are invariably coordinate with, and qualified by, the action of statistical parameters. Responses to these parameters are also native, rather than learned. But because they are essentially matters of amount, statistical means do not give rise to functional differentiation. Rather what they involve is the *continuation* of prevalent action.

There is, however, an important exception: when amounts—whether of information or redundancy, of speed or intensity of stimulation—become excessive in relation to human cognitive capabilities, the native implication is change rather than continuation. The question is: to what? Because relationships of amount are without specific implications, the probability of the 'what' is uncertain. And when uncertainty is strong, almost any kind of change is welcome. A prolonged crescendo may *end* in silence or a jazzy fox trot, in a soft tremolo or a trumpet fanfare. Whatever the case, the proper term for such termination is indeed 'end', *not* close.

Processes, on the other hand, are functionally differentiated. For instance, the parts of the first phrase of Harold Arlen's 'Somewhere Over the Rainbow' perform different functions. The melody begins with an upward skip of an octave, a 'gap' (native process), whose physical effort implies the relaxation of the following conjunct descent, the 'fill', whose specific pitches are governed by tonal syntax. The fill also involves gradual reduction in physical tension through diminished motor emphasis, lower volume, increased rhythmic regularity, and simplification of texture. This 'abatement' implies ending. And in a tonal piece such as Arlen's, syntax transforms an end into a close. In addition, to 'make assurance double sure'(Shakespeare, *Macbeth*, IV. i), the final pitch of the fill constitutes a return to the first structural tone of the melody.

Like the gap–fill pattern just discussed, many native processes are manifestations of the gestalt principles of pattern perception: for example, once begun processes tend to continue to a point of stability or clear articulation; the mind tends to improve shapes, making them simpler—more regular and complete. Cognitive constraints also result in other native processes: for example, rapid note-succession is associated with high registers and small size; long, slow note-succession is associated with low registers and larger sizes;[10] as mentioned earlier, very high levels either of redundancy or of information imply change; and, as discussed below, forceful motor emphasis tends to weaken the effect of syntactic process.

[9] Preoccupied with the evident individuality of different cultures, ethnomusicologists have, it seems to me, neglected the influence of shared, 'native' constraints on the development and organization of the musics of different cultures.

[10] The interrelation among size, speed, and place seems a universal in all realms of being: the physical, biological, social, and perceptual. In this connection, see Schmid *et al.* (2000).

Although syntactic and native processes often complement one another in shaping musical experience, they are essentially independent variables. Thus, while an abatement process generally implies closure, closure could easily be thwarted by syntax—for instance, by an open-ended rhythmic pattern or a non-closural harmony. Or to take a converse case, potential syntactic closure can be undermined by a process whose regularity natively implies continuation. The need to weaken the momentum of a marked continuity process explains why, when a perfect cadence is preceded by regular motion by fourths and fifths, the chord on the second step of the scale is usually inverted (vi–ii6–V–I). Strong syntactic closure is possible because the 'native' continuity process is broken by the first inversion rather than a root-position ii chord.[11]

In other cases, though not congruent, processes may work together to shape experience. Such complementary interaction is evident in the duet of Aida and Radames, 'O terra, addio' at the end of Verdi's *Aida*. Because the initial gap (native) of a major seventh (G♭ to F) strongly implies a fill, the following rise to G♭ functions not only as a poignant, yet delightful, resolution to the tonic (syntactic), but also as an escape tone that descends to D♭, which begins the fill (native).

While syntactic and native processes may complement one another, the relationship of process to amounts is not complementary. Though the statistical aspect of music is a necessary condition for both kinds of processes, its action is inversely related to theirs. That is, the more compelling the action of statistical parameters, the less effective the processes which, as we shall see, shape emotional experience. When, however, syntactic and native processes are primary, statistical parameters tend to be subservient.

### 15.1.2 Implication

The distinctions between emotional states and emotional processes suggest that different kinds of relationships evoke and shape different facets of musical experience. The statistical parameters beget the physical–somatic conditions which, through empathy, characterize emotional states in the primal present. Syntactic and native processes not only beget physical empathy with music, but generate mental arousal that, directing attention to what is still to come, gives rise to implication.

#### *Implication versus causation*

Both syntactic and native processes generate implications—guesses (feelings) about how present patterns will be continued and perhaps reach closure. But because 'What's to come is still unsure' (Shakespeare, *Twelfth night*, II. iii),[12] implications are never more

[11] The desire to enhance closure also accounts for why, instead of continuing directly to the tonic, a descending melodic line (e.g. from the fifth of the scale) often moves from the second to the third step of the scale before closing on the tonic.

[12] I have used quotations from, and references to, Shakespeare's writings throughout this essay not only because I happen to know his works quite well, but because their variety and pertinence so often illustrate (with little qualification) points I want to make.

than *probable.* Consequently, it is important to distinguish implication from causation. As I put the point in *Style and music* (Meyer 1989, p. 96):

> The possibility of alternative consequences . . . may help to account for the difference between implication and causation. For though both involve 'if–then' relationships, causation is a special case in which the antecedent (cause) is a sufficient condition for the consequent (effect). Consequently, alternatives are not possible. The difference in our understanding of the two concepts is reflected in ordinary language. Though it seems entirely proper to assert that 'X *implied* Y, but Y did not occur,' it would seem illogical to assert that 'X *caused* Y, but Y did not occur.'

*Implication: native and syntactic processes*

Native as well as syntactic processes can give rise to implications. As observed earlier, the implication created by a gap such as the one that begins 'Somewhere Over the Rainbow' is the result of a native process. At the beginning of Mozart's Piano Sonata in A Major (K. 331), however, a *learned* syntactic process generates the implications that link the antecedent to the consequent phrase. Similarly, an antecedent 'when clause', such as begins Shakespeare's Sonnet 'When to the sessions of sweet silent thought', implies a consequent closure:

> But if the while I think on thee dear friend
> All losses are restored and sorrows end.

Notice that in both music and language, the 'suspense' created by syntactic or native processes is short range, as compared with 'referential' suspense—for example, about finding the culprit in, say, a detective story. Thus, while the implications of the deceptive cadence at the beginning of Beethoven's Piano Sonata in E Flat, Opus 81a ('Les Adieux') are reactivated during the movement, suspense is experienced locally. And this is also true of language. One cannot, for instance, imagine the implicative tension (suspense) of a 'when' clause at the beginning of a chapter lasting until its close.

Not all implications arise from the presence of orderly processes. As discussed in Section 15.2, some result from the human antipathy to disorder. For instance, whether resulting from unpredictable change or incomprehensible redundancy, uncertainty leads listeners to anticipate the advent of stability and order. Thus insistent repetition (coupled with harmonic instability and a gradual crescendo), such as occurs early in the development section of the first movement of Beethoven's Sixth Symphony (mm. 151ff.), leads listeners to anticipate change. The question is to what extent the expectation of change is a result of culturally acquired goal-directed proclivities; after all, marked beat redundancy does not seem to imply change in contemporary popular music.

### 15.1.3 Emotional expression and emotional experience

The nature of emotional expression versus emotional experiences (and related matters) has been widely discussed in the psychological literature. Accordingly, I have chosen: (a) to offer an early instance (*c.* 1604) of the distinction between expression and experience; and (b) to call attention to a kind of emotional response seldom considered in the psychological literature.

*Expression versus experience*

Seems, madam! Nay, it is; I know not 'seems.'
'Tis not alone my inky cloak, good mother,
Nor the customary suits of solemn black,
Nor the windy suspiration of forc'd breath,
No, nor the fruitful river in the eye,
Nor the dejected havior of the visage,
Together with all forms, moods, shows of grief,
That can denote me truly. These indeed seem,
For they are actions that a man might play;
But I have that within which passeth show,
These but the trappings and the suits of woe.
(Shakespeare, *Hamlet*, I. ii)

*Connotations and recollections*

What is usually meant by an emotional response to music is a response made by acculturated listeners to the relationships that shape auditory experience. But, as is well known, an emotional response may be evoked not primarily by the sound patterns presented, but by allusions that those patterns are understood to make to other compositions (e.g. a Bach chorale), to music associated with a cultural activity (e.g. a waltz or a funeral march), or to sounds in nature (e.g. of the wind or of an animal). In these cases, it is not the patterning of the music *per se* that evokes emotion, but the *connotations* that it has for different listeners.

A related kind of response, which I call 'nostalgia' emotion, is not the result of a culturally shared reference. Rather, depending on the experience and susceptibility of a particular listener (coupled with the specific circumstance), emotion may be evoked by any sound whatsoever—from the soft sound of waves on the shore, to the singing of birds in the trees, to the rather crude music of 'The Old Grey Mare, she ain't what she used to be many long years ago', to which I marched during basic training 'many long years ago' (in 1943).[13]

## 15.2 Uncertainty: 'puzzles the will'

### 15.2.1 Evolution, choice, and culture

*Evolution and choice*[14]

Although evolution unquestionably influenced human behaviour through the direct action of biological constraints, what most significantly shaped human behaviour and

[13] It is important to recognize that emotional responses to the 'remembrance of things past' is dependent on the context of the act of remembering. Music that moves us in the concert hall or at a memorial service will probably be without affect in an elevator or a supermarket.

[14] I begin the first part of Section 15.2 with a paraphrase of the end of the last chapter of *The spheres of music* (Meyer 2000). I have included it not only because it constitutes the context for my concern with uncertainty, but because I believe it offers an account of how evolution affected human cultures that is significantly different from those proposed by neuropsychologists and sociobiologists.

gave rise to human cultures was not the presence but the absence of adequate innate constraints.[15] It is because evolution resulted in such an animal that human cultures became indispensable. This aspect of evolution is summarized by biologist François Jacob (1982, p. 61):

> In lower organisms behavior is strictly determined by the genetic program. In complex metazoa the genetic program becomes less constraining, more 'open' . . . This openness of the genetic program increases with evolution and culminates in mankind.

It seems reasonable to suggest that the growth of cognitive capacity was coordinate with the development of consciousness. And consciousness is aroused when choice is required. Indeed, it seems possible that the growth of consciousness was partly a result of the increased importance of choice. That is, as animal intelligence increased so did the apprehension of differences—of kinds and relationships—in the phenomenal world. This change may well have been accompanied by a diminution of the role of instinct in animal behaviour.[16]

Increased apprehension of kinds and relationships led to the invention and production of usables (from cradles to coffins), to a concomitant increase in the number of conceptual categories, and to a burgeoning of behavioural options. The consequences of these developments were momentous because the proliferation of conceptual categories and behavioural possibilities makes choosing at once inescapable and burdensome. In philosopher Peter Singer's words: 'Our ability to be a participant in a decision-making process, to reflect and to choose, is as much a fact about human nature as the effect of the limbic system on our emotions' (1981, p. 42).

The necessity of choice is a universal that lies at the very heart of the human condition. It constitutes a basis for both ethical and aesthetic judgments. In ethics, as in the law, individuals are culpable only if the behaviour in question can be shown to be a result of choice. In the arts, forgeries are banished to the basement because the relationships reproduced were not created—not *chosen*—by the forger.[17]

### *Choice*

A number of interrelated conditions combine to facilitate effective, propitious choosing: the presence of constraints limiting options; the ability to envisage the consequences of alternative choices; concepts of hierarchical structuring; and the power to realize one's choices.[18]

*Constraints.* Without controls limiting the number and ordering the priority of alternative behavioural possibilities, human beings would be caught in distressing indecision—

---

[15] Since writing this passage, I discovered that Clifford Geertz (1973, pp. 45–6) made much the same point in *The interpretation of cultures.*

[16] Put the other way around: as the power of instinct waned, choice became the main basis for human behaviour, and with the necessity of choice came the burden of uncertainty.

[17] See 'Forgery and the anthropology of art' in *Music, the arts, and ideas* (Meyer 1994), Chapter 4.

[18] Because it is not involved in the process of choosing, but in the realization of choices, 'power' is not allotted a separate heading under 'Choice', but is considered below—especially in Section 15.2.2.

an uncertainty that is not only time-consuming, but a source of unwelcome psychological tension. Growing out of and complementing innate constraints, cultures provide the learned, behavioural controls without which human beings could not choose, function, or, indeed, survive.

A single example from the history of music serves to illustrate this point. The abandonment of common-practice tonality by some early twentieth-century composers spawned so many compositional options that limitations were necessary to make choosing feasible. From this point of view, the invention of the twelve-tone method was a response to the need for constraints limiting options. This was the necessary condition for the advent of serialism; the sufficient condition for the constraints actually devised was largely a result of nurture—specifically, of the cultural belief that similarity begets unity.[19] Put succinctly, if too simply: the native need for compositional constraints created the 'problem'; cultural concepts provided a 'solution'.

*Envisaging, culture, and stability.* In order to choose intelligently, the various options available in a particular natural or cultural context must be comprehended and the probable consequences of each envisaged.[20] The need to envisage in order to choose is evident everywhere in human culture—from soothsayers to scientists, from astrological imaginings to the interpretation of dreams, from poll-taking to fortune-telling, from economic forecasting to the postulation of predictable historical processes. Similarly, faith in the existence of an established, unchanging order or, alternatively, in change that occurs in an orderly way is appealing because it seems to assure the possibility of successful envisaging.[21]

Indeed, all our institutions, laws and customs, as well as styles of art, serve to stabilize the conceptual-behavioural environment for the sake of effective envisaging and successful choosing. And the same is true not only of theories (including even those about chaos), but of styles of behaviour, whether in the arts or in culture generally. Conversely, the insane are institutionalized because we are uncertain about how to act in the face of 'irrational'—that is, frighteningly *un*predictable—behaviour.

*Hierarchies: functional versus continuous.* One mode of conceptualization that serves to stabilize our understanding of relationships in the realms of both nurture and nature is that of hierarchic structuring. Because the term 'hierarchy' has been used to refer to

[19] In this connection, see the Webern and Krenek quotations in *Music, the arts, and ideas* (Meyer 1994, p. 263). The nature of aesthetic unity has been an enduring concern of artists and aestheticians since Greek times. But the intensity of that concern and the nature of its conceptualization are culture-dependent. Thus, the belief that similarity produces unity appealed to romantic composers, because their repudiation of conventional constraints led to a significant increase in the number of available choices, and the variation and transformation of an existing pattern (motive, harmonic progression, etc.) facilitated compositional choice by limiting options. The role of similarity in creating unity is considered below, in note 36.

[20] Envisaging entails making inferences (based on both native and learned constraints) about the implications of a particular patterning or about the probable course of events in a specific set of circumstances.

[21] Thus, despite manifest differences in belief and method, science and religion have a common goal: namely, to stabilize the perceptual/cognitive world for the sake of explaining, envisaging, and choosing. In other words, both realms are ultimately ways of minimizing unwanted uncertainty.

different kinds of relationships, it seems important to distinguish between two basic types: (a) functionally differentiated hierarchies; and (b) continuous (functionally undifferentiated) ones.[22]

Because their *functional differentiation* specifies relational processes, biological organisms (plants as well as animals), cultural institutions (governmental, educational, industrial), most temporal art works (novels and dramas, sonata forms and rondos), and man-made machines (submarines to space ships) tend to enhance the stability of the behavioural environment. Wealth, power, and celebrity; geological, biological, and technological change; and the fluctuations of climate and populations, all form continuous hierarchies whose elements are related to one another not in terms of function, but of *amount*—for example, of wealth or power; strength, speed, or intensity.[23] In the absence of functional constraints, it seems doubtful whether statistical hierarchies foster stability and enhance the ability to envisage and choose successfully.[24]

### 15.2.2 Uncertainty

A critical corollary of the imperative of choice is the human abhorrence of uncertainty—for instance, the uncertainty surrounding death, which is undoubtedly a crucial reason for the prevalence of religion and for the correlative belief in an after-life. These are the focus of lines from Hamlet's famous 'To be, or not to be' soliloquy:

. . .the dread of something after death,
The undiscover'd country from whose bourn
No traveler returns, puzzles the will
And makes us rather bear those ills we have
Than fly to others that we know not of?
(Shakespeare, *Hamlet*, III.i)

[22] The difference between functionally differentiated and continuous hierarchies is analogous to that between the functioning of syntactic/native processes and the action of statistical parameters. This difference is exemplified by that between the closure created by a functionally differentiated syntactic progression (e.g. from subdominant to dominant to tonic) and that created by 'tonal centring' through the repetition of a pitch, etc. (as, for instance, in the first movement of Béla Bartók's Piano Sonata). After this section was written, one of the readers called my attention to Lawrence Zbikowski's (1997) essay on 'Conceptual models and cross-domain mapping'. That essay also distinguished between two kinds of hierarchy. The first, 'chains of being', is comparable to my 'continuous hierarchies'; but his second type does not seem to consider the role of functional differentiation.

[23] Whether a hierarchy is functional or continuous does not depend on the *kind* of phenomenon, but on the level being scrutinized. For instance, although social history constitutes a continuous hierarchy on the highest level of change, change on lower levels (e.g. of a particular political event) may be governed by functional constraints, and the same can be said of, say, the history of sonata form versus the form of a particular sonata-form movement.

[24] This does not gainsay the fact that statistical change may seem implicative: for instance, the gradual intensification of dynamics, increase in frequency, and shortening of duration may be understood to imply arrival at some kind of goal. But whatever 'arrival' (articulation) occurs necessarily involves either significant repetition of the tonicized pitch (see note 22) or a significant change in other parameters—for instance, a lengthening of the final sound, a change in orchestration, dynamic level, or register.

Here an apparent paradox arises. On the one hand, uncertainty is anathema to humankind.[25] For this reason, as noted above, we devise ways of reducing uncertainty both in the 'out-there' world and in our personal lives. On the other hand, in the arts and in other 'playful' activities such as sports, games, and gambling, we actually relish and cultivate a considerable amount of uncertainty.

*Belief, empathy, and aesthetic experience*

The paradox is possible not only because of what Coleridge called 'the willing suspension of disbelief',[26] but more importantly because of a *positive* belief in the competence, integrity, and creativity of the artist and, above all, in the significance of works of art. Such positive belief is indispensable for man-made aesthetic experience (as distinguished from the experience of beauty in the natural world), because it is a necessary condition for the empathy on which perceptual engagement and affective response depend.[27]

The importance of such belief is evident in the ability of criticism to affect the responses of listeners. For instance, the negative reception of avant-garde music was, at least in part, a result of critiques that undermined audience empathy by suggesting that the choices of composers were the result of system-based calculation or of random chance, rather than of 'inspiration' and 'inner feeling'. On the other side, consider this stunning, if preposterous, instance of the power of positive belief to produce a favourable response in an audience. In a book review, 'Fascinating rhythm', David Hajdu writes: 'In 1971, Ravi Shankar, the Indian virtuoso, performed at New York's Madison Square Garden. After hearing a few minutes of Shankar's ensemble, the audience of some 20 000 roared in approval. "Thank you," Shankar replied. "If you appreciate the tuning so much, I hope you will enjoy the playing more."'[28]

The empathy that belief begets is not solely cognitive. Almost always it involves physical behaviour—an inner performative empathy, a kind of imitative identification, with the qualities and patterns of the music.[29] Lack of motor empathy, for example, seems to

[25] Uncertainty is evidently also anathema to other animals. Elizabeth Pennisi (2000, p. 577), for example, writes that: 'Although baboons are considered quite aggressive primates by nature, violence tends to be much more prevalent when the troop's structure is unstable . . . if the troop members don't know one another, or if the hierarchy is disturbed, say, by the loss of the top male, the rush to establish rank results in scuffles that subside when everyone knows his or her place.' In short, instability, lack of familiarity, disturbance of social structuring—all create uncertainty which, in turn, begets violence.

[26] *Biographia literaria*, Chapter 14. The 'willing suspension of disbelief' is an important factor in the ability of audiences to re-experience implicative relationships—to 'make believe' that they are experiencing anew.

[27] In addition to being a source of aesthetic enjoyment, music has, since its beginnings, served to foster social cohesion—through the evocation of empathic sharing in religious and secular ceremonies, attending to or participating in theatrical performances, and marching or dancing together.

[28] *New York Review of Books*, Vol. 47, No. 12 (20 July 2000), p. 41.

[29] Many years ago (Meyer 1956*b*), I argued that belief was essential for the success of music therapy. More recently, studies described in the *Harvard Men's Health Watch* (September 1997, p. 6) have demonstrated the importance of belief for longevity and more general well-being. In addition, it seems likely that belief plays a significant role in the placebo effect (see Enserink 1999).

have played a significant role in the negative aesthetic and hence emotional response to avant-garde music, while motorically accessible musics (from Bach to rock) have found ardent audiences. Indeed, I suspect that often contemporary music has failed to find favour with concert audiences, not primarily because it tends to be dissonant, but because it is difficult to empathize with and experience motorically—not just the lack of rhythmic regularity, but such motor-related problems as the prevalence of large descending skips and intervals such as augmented fourths.

The roles of belief and empathy (physical as well as cognitive) help to explain why the uncertainty experienced in relation to works of art, games, and sports differs significantly from that evoked by 'real-life' circumstances.[30] Indeed, in the temporal arts, the arousal and resolution of uncertainty is an essential basis for aesthetic-emotional experience. Because patterns imply alternative possible continuations and realizations, implication invariably involves some degree of uncertainty—uncertainty about *what* will happen, *when* it will happen, and *how* it will happen.[31] Consequently, to paraphrase *Hamlet* again: like sorrows, uncertainties 'come not single spies/But in battalions' (Shakespeare, *Hamlet*, IV. v).

It is important to recognize that uncertainty entails at least an inkling of patterning. Because either too little or too much information precludes the perception of even a modicum of order, neither is compatible with uncertainty. That is, unless there is some degree of order, whether presumed or inferred, there is nothing to be uncertain *about*. And because even a bit of order tends to be implicative, uncertainty is so as well.

### *The succession of states*

The role of uncertainty has unfortunately been neglected in the study of the emotional response to music. Neglect has probably occurred because emotional 'states' such as love, anger, and hate are, as mentioned earlier, experienced, classified, and named in terms of the particular real-life circumstances with which they become associated.[32] But uncertainty itself is not 'circumstantial'. And, because it can be coupled with and qualify many states of being or feeling, uncertainty is not thought of as an 'emotion'. However, although not a nameable, circumstantial state, uncertainty is a necessary, though perhaps not sufficient, condition for the emotional experiences and behaviours of human beings.[33]

What I am suggesting, then, is that states of feeling (e.g. love, contentment, hunger, fatigue) become emotional experiences when they are at least tinged by uncertainty. As

[30] Another reason for this difference has to do with our relationship to 'roads-not-taken'. In the arts and other pleasure activities, 'roads-not-taken' constitute a source of experiential enrichment. In real life, on the other hand, an abundance of alternatives tends to be disturbing and perplexing.

[31] At times, implications are realized remotely and may, as a result, be recognized largely in retrospect. That is, we did not experience uncertainty when the pattern was presented earlier, but realize subsequently that we *might* have done so.

[32] I suspect that the neglect of uncertainty was, ironically, a result of the psychological and cultural craving for the control that *naming* (separate emotions) provides.

[33] This account represents a modification of that given in Meyer (1956*a*, p. 16 ff.).

observed earlier, although being 'curious' is not in itself an emotional experience, it may become so if the behavioural inclinations (the needs or goals) associated with it are thwarted. And this appears to be the case with other goal-directed needs—from the need for food to the need to choose. *Qua* feelings, they are states of being. They become emotional experiences or responses when thwarting creates uncertainty about appropriate courses of action.

Seen thus, uncertainty is the result of ignorance—ignorance about how past and present patterns will be continued, eventually reaching stability and closure. In other words, uncertainty is always more or less goal-directed, implying (perhaps unconsciously) resolution to the security of knowing.

Since uncertainty is a function of the amount of information the mind is required to process, if there is so little information that patterning is impossible, the result is a sense of powerlessness and hope-based uncertainty is abandoned. Perhaps this is why during the Vietnam War it was found that, deprived of any stimulation whatsoever, prisoners became mentally deranged. Stimulus overload (too much information), which may be a matter either of amount, or speed, or both, is almost as disturbing as stimulus deprivation. Hence audiences tend to be alienated, when (what for them is) an excessive amount of information in a piece of contemporary music precludes patterning and, consequently, a sense of control.[34]

At the end of *Emotion and meaning in music* (Meyer 1956*a*) and subsequently in *Explaining music* (Meyer 1973), I grappled with the problem of accounting for the succession of emotional states presented in a piece of music. Much later, in 'A universe of universals' (Meyer 2000), I argued that no necessary, inherent order governs the succession of emotional states or referential subjects presented in a piece of music.[35] We accept such successions because the uncertainty and lack of cognitive control created by intervening instability make the return to mentally manageable patterns satisfying. In short, it is the uncertainty–resolution (or, correlatively, the implication–realization) process, not similarity of melodic, rhythmic, or tonal pattern, that unifies the succession of emotional states presented in a piece of music.[36]

As emphasized in Section 15.1, we conceptualize and understand phenomena not in terms of individual instances, but in terms of *classes*. And classification plays a critical role in our sense of cognitive control and power.

[34] If the amount of information is inordinately increased, the result is a kind of cognitive white noise.

[35] A particular succession may, of course, be the result of a programme conceived by the composer, or devised by critics or listeners to account for a succession not linked together compositionally by the tension of syntactic uncertainty. This, together with the diminished sophistication of the audience, may help to account for the increased use of programmes during the nineteenth century.

Also, in order to enhance expressivity, there tended to be an increase in the instability and tension of putatively stable passages. As a result, the difference in degree of uncertainty between presumably stable and unstable passages diminished. When this occurred, the relatively weak level of uncertainty could not unambiguously link successive emotional states.

[36] As observed in note 19, the belief that similarity begets unity became aesthetic doctrine during the nineteenth and early twentieth centuries. But a moment's thought makes it evident that the more

*Power*

When it involves loss of control, uncertainty may also become associated with feelings of angst and terror. This perhaps partly explains why rapid tremolos, which cannot readily be processed by either mind or body, are associated with angst and terror. As in the startle reaction, shockingly unexpected stimuli engender anxiety and fear of the unpredictable and the inexplicable. Writing about the situation in Kosovo, Mark Danner observed that random, indiscriminate violence is more effective than orderly extermination: 'Imposition of terror, *the more "indiscriminate", the better,* breeds fear . . .'[37] Briefly: total uncertainty is unbearably painful.

So too is the frustration created by the impossibility of action. This may help to account for the weeping which seems to come from strong feeling, often evoked by emotion-laden associations and memories, that can find no expression in action. Feelings of hopelessness—an inferno resulting from the inability to choose and hence to act—give rise to strong emotion. For instance, when we recognize that an important goal cannot be realized—that hope must be abandoned—feelings often result in weeping.[38] The action that might-have-been is, as it were, transformed into tears.[39]

This is the source of what I have called the emotion of 'nostalgia'. A personal experience exemplifies this kind of response. At a performance of Verdi's *La Traviata*, when Germont began to sing 'Di Provenza', I began to weep, not only because my father

*alike* two patterns are, the more they seem separate—the case, for instance, with the motivic repetitions in the Scherzo of Beethoven's Symphony No. 7 (e.g. mm. 37–40 and 49–52). Confusion occurs because unity is equated with *uniformity*.

Unity, however, involves the functional connection of phenomena that differ in some significant respect; uniformity involves connection through similarity of substance and/or pattern. Thus, although the grains of sand on a beach are essentially alike, it would surely seem strange to say that the beach was unified. Similarly, while it is proper to characterize a musical ostinato or a word repetition (such as Macbeth's 'Tomorrow, and tomorrow, and tomorrow') as uniform, it would be wrong to say that they created unity. Whether in music, or in literature, or in biological organisms, unity results not from similarity, but from functional differentiation.

[37] 'Endgame in Kosovo', *New York Review of Books*, Vol. 46, No. 8 (6 May 1999), p. 8.

[38] As Sloboda (1992, p. 37) observed, weeping would seem to be a unambiguous manifestation of emotion. Yet it has been strangely neglected in the psychological literature about music. I can think of two (not mutually exclusive) reasons: (a) because, like uncertainty, weeping is circumstantially variable; and (b) because it is difficult to know whether the tears are a result of acculturation or personal experience.

[39] If weeping comes from within, laughter comes from outside us, from the incongruities of the outside world (in this connection, see note 3). Even when we laugh at ourselves, it is the out-there, 'other' self that we laugh at. Laughter (e.g. at jokes) exemplifies a related facet of uncertainty. Initially, as with any process, we are uncertain as to how stability and closure will be reached. To heighten the tension of uncertainty, the first part of the process is usually elaborated. The tension of patent, though usually playful, delay is resolved by a 'punch-line', and we laugh: first, because of the relief that comes with *knowing* and second because, though apt, the outcome is surprising and amusingly incongruous.

especially loved that aria, but because evoking memories of 'days beyond recall', and beyond empathic reciprocity, made my present powerlessness apparent.[40]

*Risk-taking and culture*

Uncertainty can also be associated with what appears to be the opposite of frustration: namely, the thrill and exhilaration of risk-taking. For what sky-diving and the virtuoso performance of a challenging cadenza, or adventurous mountain-climbing and daring modulations to a distant tonality, have in common is the thrilling uncertainty of risk-taking and the illumination of self-discovery.[41] The pleasure of, and empathy with, risk-taking is, however, dependent on the belief of all concerned (composer, performer, and audience) in the existence of expert control.

Risk-taking is in turn related to the heroic. To appreciate, and empathize with, the uncertainty of risk is to move toward and experience something of the heroic. And though it seems *passé* (and a bit theatrical), Samuel Johnson touched on an important facet of human behaviour when he observed that:[42]

> Were Socrates and Charles the Twelfth of Sweden both present in any company and Socrates to say 'Follow me and hear a lecture on philosophy'; and Charles, laying his hand on his sword, to say 'Follow me and dethrone the Czar'; a man would be ashamed to follow Socrates . . . The profession of soldiers and sailors has the dignity of danger [of risk-taking]. Mankind reverence those who have got over fear, which is so general a weakness.

But in contemporary culture there has been a significant decline in the valuing of personal honour, the heroic, and the risk-taking associated with them. Hayden White (1973, p. 232) relates the ironic mode of much twentieth-century thought to 'the passage of the age of heros and of the capacity to believe in heroism'.[43] And, according to Admiral Arthur, commander of the American naval forces during the Gulf War, even in the military 'We now have people being taught how to be cautious, not how to be bold' (Schmitt 1995). Construed in terms of the subject of this chapter, tolerance for the tensions of uncertainty has been considerably reduced. One reason for this reduction is

[40] Needless to say, not all crying is 'nostalgia-generated'. Infants cry because of unrealized needs—for food, affection, nappy (diaper)-changing, etc. But certainty can, without forcing, also be considered an unrealized need—the need for power, for confidence in our ability to act.

[41] Thus, Garry Wills observed that Robert Kennedy 'courted danger, climbing mountains, shooting rapids. What he was looking for was his true self, so long repressed, so totally held in the service of others' ('Waiting for Bobby,' a review of Ronald Steel's *In love with night: The American romance with Robert Kennedy* published in the *New York Review of Books*, Vol. 47, No. 2, 10 February 2000, p. 20).

[42] Said by Samuel Johnson in 1778 and quoted in Harry G. Summers Jr., 'The men of Company E' (a review of *Band of brothers* by Stephen E. Ambrose published in the Book Review, *New York Times*, 6 December 1992, p. 11). It could be argued that fear is basically a kind of uncertainty.

[43] Also see J. B. Freeman's 'Dueling as politics: Reinterpreting the Burr-Hamilton duel' (*William and Mary Quarterly*, 1996, pp. 289–318), discussed in Gordon S. Wood, 'An affair of honor' published in the *New York Review of Books*, Vol. 47, No. 6, 13 April 2000, pp. 67–8.

that risk-taking, honour, and the heroic are *individual* endeavours that are incompatible with egalitarian, mass culture.

But perhaps the most important reason for the changes taking place in our culture is the altered attitude toward history itself.[44] As the future-oriented perspective prevalent since the Renaissance has weakened, there has been a trend away from risky goal-directed action, in both life and art. This seems evident in the complacent *present*ational character of much minimalist music. Paradoxically, this is also the case with 'bing-bang' popular music whose aggressive force implies no future goal, but rather calls for compliance NOW.[45]

*Further thoughts about uncertainty*

*Narratives and plots.* Uncertainty is of vital importance in the aesthetic process of moving from the insecurity of ignorance to the security of knowing. This process, which Aristotle referred to as the 'reversal' of the action in a drama, is at the core of all plots. And this calls attention to the difference between a narrative and a plot.

Most histories (evolution, political, economic, and intellectual), as well as many literary and musical genres (novels and dramas; sonata forms and song cycles), are narratives.[46] But not all narratives are plots. A narrative succession is transformed into a plot when uncertainty is experienced in terms of an imagined goal. One of the problems with many narrative interpretations of music is that they are based on passages that present stable emotional states and referential associations. The result is a succession of largely separate vignettes rather than a plot in which uncertainty reaches resolution in stability.

Here another paradox arises. The human need to envisage leads us to conceptualize the world as being more or less predictable. But continual succession without functionally ordered structure (as in cosmology, evolution, and human history) creates the uncomfortable uncertainty of *un*ordered endlessness. As a result, we tend to transform what are essentially fluctuating continuities (narratives without goals) into plots. To put the paradox succinctly: the uncertainty created by pattern instability begets functional differentiation, which reduces the uncertainty produced by an unarticulated succession of emotional states.

Finally, I suspect that the psychological desire to transform narratives into plots is one reason why we tend to anthropomorphize phenomena. That is, anthropomorphism not only provides conceptual means for envisaging and choosing, but provides goals that change potentially endless successions into implication-realization patterns—that is, into plots with potential for closure.

[44] The end of historic optimism is discussed in the 'Postlude' to the second printing of *Music, the arts, and ideas* (Meyer 1994).

[45] Although my aesthetic predilections are obvious, my concern is not with making value judgments, but with trying to understand and explain some of the changes taking place in contemporary culture.

[46] Though implicative, sonnets (as illustrated above) are not narrative—nor are limericks. Both are like propositional 'arguments'. In music, the closest form to 'pure' narrative is the theme and variations which, like high-level histories, are potentially endless successions.

*Motor action, deviation, and implication.* Because it implies virtually automatic continuation, marked motor regularity and strong emotional process tend to be inversely related. And because rapid tempos are usually coordinate with forceful motor action, speed and emotion also tend to be inversely related. For instance, in the opening section of Bach's Brandenburg Concerto No. 2, rhythm, tempo, and (implicit) dynamics evoke strong motor empathy, but only minimal emotional tension.[47]

This suggests why forceful motor action is incompatible with richly implicative processes: namely, because emphatic motor action tends to minimize the imagining of alternative implications—that is, other continuations.[48] For the same reason, however, forceful motor action can be important in the generation of surprise: that is, we are most easily and forcefully surprised when we are *un*aware of the possibility of alternatives.

Mention of surprise reminds us that implications and uncertainties can be recognized, though not experienced, in retrospect. This occurs when the regularity of a pattern and its context make continuation seem highly probable. Then if an abrupt change occurs—as, for instance, in the slow movement of Haydn's 'Surprise' Symphony (No. 94)—we realize that we *should have been* less complacent about the implications of the earlier patterning. Conversely, we may in retrospect recognize that a present pattern is, in fact, the realization of an implication not noticed in a prior process.

In general, not just forceful motor action, but any pattern or behaviour that seems predictable—is stylistically normal and cognitively coherent—reduces uncertainty. By disturbing such predictability, deviation (surprise, delay, etc.) may evoke uncertainty and emotional experience.[49] In short, some degree of regularity (stability) is a necessary condition for an emotional response to music; the sufficient condition is, as I suggested 'many long years ago', the uncertainty created by some degree of deviation.[50]

*Intensity, control, and emotional experience.* As observed earlier, the intensity of uncertainty is variable, depending both on the patterning of events and what is believed to be the importance of their goal(s). In general, as the number and strength of implied alternatives increases so does uncertainty;[51] and the same holds for the richness of their

[47] Again, the distinction between an emotional state and an emotional process is critical. Although Bach's opening creates feelings of forceful, confident power, it does not, as I see it, give rise to marked emotional process.

[48] Because they tend to preclude awareness of possible alternative continuations, overly rapid tempi also obscure the relational richness potential in performance.

[49] Briefly, in 'life' we try to conceptualize the world so that it seems regular and predictable. In the arts, on the other hand, we devise regularities as a base for the deviations that, by creating uncertainty, beget emotion and shape aesthetic experience.

[50] In *Emotion and meaning in music* (Meyer 1956*a*), I argued that deviation, whether one present in the composer's score or one created by the performer, tends to evoke emotion. But I failed to see that deviation does so because it is a source of uncertainty.

[51] Janet M. Levy (1995) distinguishes the multiple meanings (implications) that beget uncertainty from ambiguity which 'gives rise to two or more specifiable meanings' (p. 152).

imagined consequences, both proximate and remote.[52] Thus, when it is said of a novel that 'the plot thickens', what is probably meant is that uncertainty has increased. In music, the degree of uncertainty that is pleasant and enjoyable depends not only on musical relationships, but, as suggested earlier, on the beliefs of the audience about the historical importance and aesthetic value of the music being played. Beliefs are a prerequisite for empathy, and empathy for attentive perception and sensitive response.

The degree of uncertainty appears to be the basis for the distinction made by psychologists between 'pleasant' and 'unpleasant' emotions.[53] This is, I suspect, because intensity influences the sense of control experienced by the listener. Put differently, emotions are felt to be pleasant, even exhilarating, when associated with confidence in the power to envisage and choose; they are experienced as unpleasant when intense uncertainty precludes the sense of control that comes from the ability to envisage with confidence.

Finally, the aesthetic importance of uncertainty lies not only in the shaping of emotional *processes*, as distinguished from the evocation of emotional *states*, but in the articulation of musical structure. When the tensions of instability are resolved to the cognitive security of stable patterning, functional relationships have at once articulated and unified musical structure.

I am, appropriately, uncertain about how to end this essay—and that makes me somewhat sad. Yet, at the same time and for the same reason, I feel glad—glad because connecting puzzled uncertainty with emotion *seems* to confirm my basic argument.[54]

## References

Enserink, M. (1999). Can the placebo be the cure? *Science*, **264**, 238–40.

Geertz, C. (1973). *The interpretation of cultures*. New York: Basic Books.

Jacob, F. (1982). *The possible and the actual*. New York: Pantheon Books.

Levy, J. M. (1995). Beginning-ending ambiguity: Consequences of performance choices. In *The practice of performance*, (ed. J. Rink), pp. 150–69. Cambridge, UK: Cambridge University Press.

Meyer, L. B. (1956*a*). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Meyer, L. B. (1956*b*). Learning, belief, and music therapy. *Music Therapy*, **5**, 27–35.

[52] There are significant differences between complexity and relational richness. In this connection, see *The spheres of music* (Meyer 2000), Essay 2.

[53] Anxiety needs to be distinguished from uncertainty. Unlike uncertainty, anxiety occurs when governing constraints are not known and cannot be inferred from the perception of overly complex stimuli. Consequently, anxiety is not implicative but endures, fluctuating in a persistent present. However, because of our beliefs about its special nature, aesthetic experience seldom gives rise to anxiety; it does so only when all sense of order and hence the possibility of cognitive control is lost.

[54] My thanks to the reviewers of this essay—especially to Nicholas Cook—for their helpful comments, criticisms, and suggestions. I am deeply indebted to my wife, Janet Levy, for her leading questions, astute observations, and cogent critiques of the various versions of this study.

**Meyer, L. B.** (1973). *Explaining music. Essays and explorations.* Berkeley, CA: University of California Press.

**Meyer, L. B.** (1989). *Style and music.* Philadelphia, PA: University of Pennsylvania Press.

**Meyer, L. B.** (1994). *Music, the arts, and ideas.* Chicago: University of Chigago Press (originally published 1967).

**Meyer, L. B.** (2000). *The spheres of music.* Chicago: University of Chigago Press.

**Narmour, E.** (1990). *The analysis and cognition of basic melodic structures: The implication-realization model.* Chicago: University of Chicago Press.

**Pennisi, E.** (2000). Personality and rank. *Science,* **289**, 577.

**Scherer, K. R.** (1994). Toward a concept of modal emotions. In *The nature of emotion,* (ed. P. Ekman & R. J. Davidson), pp. 25–31. New York: Oxford University Press.

**Schmid, Tokeshu, & Araya** (2000). Relation between population density and body size in stream communities. *Science,* **289**, 1557–9.

**Schmitt, E.** (1995) The military's getting queasier about death. *New York Times,* The Week in Review, 6 May, p. 5.

**Singer, P.** (1981). *The expanding circle: Ethics and sociobiology.* New York: Farrar, Straus & Girony.

**Sloboda, J. A.** (1992). Empirical studies of emotional response to music. In *Cognitive bases of musical communication,* (ed. M. Riess-Jones & S. Holleran), pp. 33–46. Washington, DC: American Psychological Association.

**White, H.** (1973). *Metahistory: The historical imagination in nineteenth-century Europe.* Baltimore: Johns Hopkins University Press.

**Zbikowski, L. M.** (1997). Conceptual models and cross-domain mapping: New perspectives on theories of music on hierarchy. *Journal of Music Theory,* **41**, 193–225.

CHAPTER 16

# EMOTIONAL EFFECTS OF MUSIC: PRODUCTION RULES

KLAUS R. SCHERER AND MARCEL R. ZENTNER

It is an ancient, and very pervasive, idea that music *expresses* emotion. Apart from the copious literature to this effect contributed by composers, musicologists, and philosophers, there is also solid empirical evidence from psychological research, reviewed in chapters of this book (e.g. Gabrielsson & Lindström, this volume; Juslin, this volume), that listeners often agree rather strongly about what type of emotion is expressed in a particular piece. It is also a pervasive belief that music can, at times, actually *produce* emotion in listeners. The distinction between perception and production is related to the distinction between cognitivism and emotivism proposed by philosophers in their analysis of emotion in music (e.g. Kivy 1989). Whereas 'emotivists' hold that music elicits real emotional responses in listeners, 'cognitivists' argue that music simply expresses or represents emotions. Our view is that it would be premature to prejudge the issue and that both positions may be perfectly appropriate depending on a number of factors outlined below. Our purpose in this chapter is to provide a formalization of the processes whereby music produces emotional effects in the listener that go beyond the cognitive inference of what the music can be said to express. In addition, we review the pertinent evidence to date and suggest ways in which future research might investigate these processes in a systematic fashion. We state at the outset that our discussion and the review of the available evidence are largely based on Western classical music, thereby restricting the generalizability of our claims to other kinds of music and other cultures.

Our attempt at a formalization of emotional effects of music will consist in defining the affective changes that music is supposed to produce in the listener and to identify the determinants of the listening situation (features such as the musical structure of the piece listened to, the interpretation by the performer, relevant state and trait characteristics of the listener, and the respective context). An important issue for discussion will be the relative weighting of the different determinants and the type of their interaction in producing the affective outcome. We follow the senior author's attempt to define the emotional meaning of music in analogy to Bühler's Organon model of language, postulating composition, expression, perception, and production *rules* to model the different facets of emotional meaning (Scherer, 2000*d*). In this context, the term *rule* is used to denote a certain regularity or lawfulness of the effects discussed that can be expressed in concrete predictions or hypotheses. It is not excluded that these rules can be integrated into attempts at computational modelling of the underlying mechanisms (in fact, we consider this a highly desirable option), but we do not feel that the current state of the

art provides the necessary precision in specifying the model parameters and the input and output variables that would be required to pursue this aim. Thus, the term *rules* and the pseudo-mathematical formulas to formalize them suggested below are used in a purely descriptive sense in this chapter.

A formalization of the process by which music generates emotion requires, first of all, a definition of what exactly is to serve as the *output variable*, that is, the type of emotional or affective state that is supposed to be produced by music. This is necessary since much of the confusion in the literature is due to a lack of conceptual clarity concerning the hypothetical constructs involved. Scherer (2000*a*) has suggested a design feature delimitation of various affective phenomena, reproduced in an adapted form in Table 16.1. In view of the lack of agreement in the field, we have not attempted to identify the characteristics of *aesthetic emotions*, particularly relevant to music, either as a separate category or as a subcategory of the types proposed. This is a task that remains to be accomplished in future work. We will consistently apply the distinctions suggested in Table 16.1 in this chapter, both in classifying the literature to be reviewed (which at times requires one to speculate what an author meant or what a study measured) and in formalizing mechanisms and discovery procedures. In particular, we will try to distinguish affective preferences, moods, and emotion episodes, all of which might be produced by music. Clearly, any attempt to demonstrate that these different types of states are induced by music requires that the defining characteristics of the respective state are operationalized and measured.

The next step is to operationalize the *input variables*; that is, to specify in greater detail which aspects of listening to a piece of music are involved in the inference and/or induction of emotion (for a more detailed example of such an analytical approach, see Dowling & Harwood 1986). We will distinguish structural, performance, listener, and context features, presuming that these features can, individually or in combination, produce the different affective states described above.

*Structural features*

We subdivide structural features of music—that is, all those qualities of a composer's score that a good performer needs to respect—into segmental and suprasegmental types. *Segmental features* consist of the acoustic characteristics of the building blocks of musical structure: individual sounds or tones as produced by the singing voice or specific musical instruments. In some ways one may presume that such individual sounds correspond to the isolated affect vocalizations that may lie at the source of speech and music. The use of 'Ah' and 'ahimé' in baroque operas illustrates what has been called the domestication of 'brute' affect vocalizations (Kainz 1962; Wundt 1900; for a review of this literature, see Scherer 1994). The closest equivalent in instrumental music is the individual tone, interval, or chord. The acoustic structure of such individual tone segments, corresponding to phones in speech, is described by duration, energy (amplitude), pitch (fundamental frequency), and timbre or harmonic structure (spectral composition) of the complex wave, as well as the energy and pitch envelopes and microchanges in timbre over the duration of the sound. In sound or tone sequences, this acoustic information can be aggregated, yielding measures of central tendencies (mean, median) that, in relatively homogeneous sequences, can be described as the

Table 16.1 Design feature delimitation of different affective states (adapted from Scherer (2000*a*))

| Design feature / Type of affective state: brief definition (*examples*) | Intensity | Duration | Synchronization | Event focus | Appraisal elicitation | Rapidity of change | Behavioral impact |
|---|---|---|---|---|---|---|---|
| Preferences: evaluative judgments of stimuli in the sense of liking or disliking, or preferring or not over another stimulus (*like, dislike, positive, negative*) | L | M | VL | VH | H | VL | M |
| Emotions: relatively brief episodes of synchronized response of all or most organismic subsystems in response to the evaluation of an external or internal event as being of major significance (*angry, sad, joyful, fearful, ashamed, proud, elated, desperate*) | H | L | VH | VH | VH | VH | VH |
| Mood: diffuse affect state, most pronounced as change in subjective feeling, of low intensity but relatively long duration, often without apparent cause (*cheerful, gloomy, irritable, listless, depressed, buoyant*) | M | H | L | L | L | H | H |
| Interpersonal stances: affective stance taken toward another person in a specific interaction, colouring the interpersonal exchange in that situation (*distant, cold, warm, supportive, contemptuous*) | M | M | L | H | L | VH | H |
| Attitudes: relatively enduring, affectively coloured beliefs and predispositions towards objects or persons (*liking, loving, hating, valuing, desiring*) | M | H | VL | VL | L | L | L |
| Personality traits: emotionally laden, stable personality dispositions and behavior tendencies, typical for a person (*nervous, anxious, reckless, morose, hostile, envious, jealous*) | L | VH | VL | VL | VL | VL | L |

VL = very low, L = low, M = medium, H = high, VH = very high.

dominant acoustic structure of the sound sequence. This process of aggregation is straightforward for some acoustic variables (e.g. pitch) and rather difficult for others (e.g. timbre). However, even in the more difficult cases approximations can be found (e.g. a rough aggregate measure of dominant timbre might be the mean proportion of the energy above 1 kHz).

Segmental effects on emotion inference or induction are expected to be relatively stable and universal, with the exception of random error, over all types of listeners and

performance conditions. Generally, such effects are mediated by evolutionarily evolved iconic signalling characteristics, based on physiological changes in affect vocalization and relatively independent of individual or cultural differences (cf. also Rosar 1994).

*Suprasegmental features* consist of systematic configurational changes in sound sequences over time, such as intonation and amplitude contours in speech. In music, comparable features are melody, tempo, rhythm, harmony, and other aspects of musical structure and form. While iconic coding also plays an important role (e.g. with respect to tempo, rising/falling contours), suprasegmental features seem to carry emotional information primarily through *symbolic coding*, as based on a process of historically evolved, sociocultural conventionalization (see Kappas *et al.* 1991; Scherer 1988; Sloboda 1992).

*Performance features*

These refer to the way in which a piece of music is executed by performer(s). Both the stable *identity* (physical appearance, expression, reputation) and *ability* (technical and interpretative skills) of the performer, as well as transient performance-related variables referred to here as *performance state* (interpretation, concentration, motivation, mood, stage presence, audience contact, etc.), may have a major impact on the perception and induction of emotion. The effects of performance features can be based on both iconic and symbolic coding.

*Listener features*

Listener features are based on the individual and sociocultural identity of the listener and on the symbolic coding convention prevalent in a particular culture or subculture. They can consist of interpretation rules (e.g. musical systems) that are shared in a group or culture, or of inference dispositions based on personality, prior experiences, and musical talent. These factors can be summarized into *musical expertise*, including cultural expectations about musical meaning, and *stable dispositions* unrelated to music, such as personality or perceptual habits. In addition, transient listener states such as motivational state, concentration, or mood may also affect emotional inference from music (cf. Cantor & Zillmann 1973). In the case of listener features, we suggest a third type of meaning-conferring coding—*associative coding*—in addition to the established semiotic distinction between iconic and symbolic coding. By this we mean the fact that segmental, suprasegmental, and performance cues can be associated with emotional content in an individual's memory due to learned associations and conditioning.

*Contextual features*

These refer to certain aspects of the performance and/or listening situation. Thus, the *location* of a performance and/or a listening situation may be a concert hall, church, open-air site, car, or recording studio. The dominant material surrounding the listener/performer may be wood, glass, stone, metal, or cement. In addition to the location, the particular *event* may be a wedding, a funeral, a ball, or the celebration of an outstanding achievement. The music may be transmitted through loud-speakers, headphones, or

without any technical support. The music may be heard without interruption or it may be disturbed by the sirenes of an ambulance or the coughing of a concert visitor. We submit that all these features can have an influence on the acoustics, the ambiance of the location, or the behaviour of the audience, which in turn may lead to different emotional effects due to objective features of the situation or subjective perceptions of the listeners.

## 16.1 Production rules

### 16.1.1 Predictions

We suggest that an emotion that is actually experienced by a listener while listening to music is determined by a multiplicative function consisting of several factors:

Experienced emotion = Structural features × Performance features × Listener features × Contextual features,
where
Structural features = $W_1$ (Segmental features) × $W_2$ (Suprasegmental features),
Performance features = $W_3$ (Performer skills) × $W_4$ (Performer state),
Listener features = $W_5$ (Musical expertise) × $W_6$ (Stable dispositions) × $W_7$ (Current motivational / mood state),
and
Contextual features = $W_7$ (Location) × $W_8$ (Event).

We postulate multiplicative rather than additive functions because it is unlikely that any of the constituent factors in and of themselves, in the absence of other factors, can lead to marked, reliable emotional effects. Furthermore, it seems intuitively reasonable to assume that some listener features (e.g. musical expertise) will strongly interact (in the statistical sense of conveying special strength or significance) with specific structural or performance features. Finally, we submit that the different factors carry different weight. In fact, the order in which the variables are listed above follows the results of a study in which experienced listeners, many of them experts, were asked to rate the extent to which each of the major factors listed above was perceived to influence musical emotion induction (Scherer *et al.*, in press). Clearly, more exploratory studies are needed before more subtle predictions can be made and tested.

After this initial formalization of the output and the input variables involved in emotional responses to music, and the formulation of certain predictions, we now turn to what we call *routes*, that is, particular mechanisms whereby emotion may be generated by music. We distingish: (a) *central routes* (i.e. implicating the central nervous system (CNS) in emotion generation); and (b) *peripheral routes* (based on direct effects on the peripheral, i.e. the somatic and autonomic, nervous systems (SNS and ANS) with ensuing proprioceptive feedback to central areas). In Sections 16.1.2 and 16.1.3 below, we will briefly describe these different routes and some of the underlying mechanisms. Section 16.1.4 describes the evidence to date.

### 16.1.2 Central route production

*Appraisal*

A parsimonious premise is that musical stimuli provoke emotions in a similar fashion to any other emotion-eliciting event. Thus, the mechanisms described by emotion psychologists may be also applicable to the study of emotion induction via music. There is an emerging consensus that emotion elicitation and differentiation is best understood by assuming a process of event evaluation, or appraisal, that models the way in which an individual assesses the personal significance of an event for its well-being on a number of critieria and dimensions. This process of evaluation is demonstrated in the upper portion of Fig. 16.1: an object or event is evaluated by a specific person, with respect to a number of criteria or dimensions concerning the implications of the event for needs, goals, or values of the individual and his or her ability to cope with the consequences of the event. The result of this appraisal process is an emotion, which is then expressed or externalized in physiological symptoms and, particularly, in motor expressive movements in the face, body, and voice. In this theoretical tradition, the type of ensuing emotion and the patterning of the physiological and expressive responses are seen as dependent on the specific profile of appraisal results on the pertinent criteria (see Scherer 1999; Scherer *et al.*, 2001; for alternative models of emotion differentiation, see Scherer 2000*a*). This appraisal process may occur in a rudimentary, automatic fashion at lower levels of the CNS (mostly the limbic system), especially for evolutionarily 'prepared' stimuli, or in a more elaborated and more effortful process involving the cortical association regions of the CNS as well as the lower centres (see Leventhal & Scherer 1987; Teasdale 1999; van Reekum & Scherer 1998).

The automatic, lower-level appraisal route to emotion elicitation has been well described by Öhman and his colleagues (Öhman 1988) who have suggested that a number of stimuli (e.g. snakes and spiders) can be processed by the lower centres of the brain in an extremely short period of time, outside of awareness, without any apparent

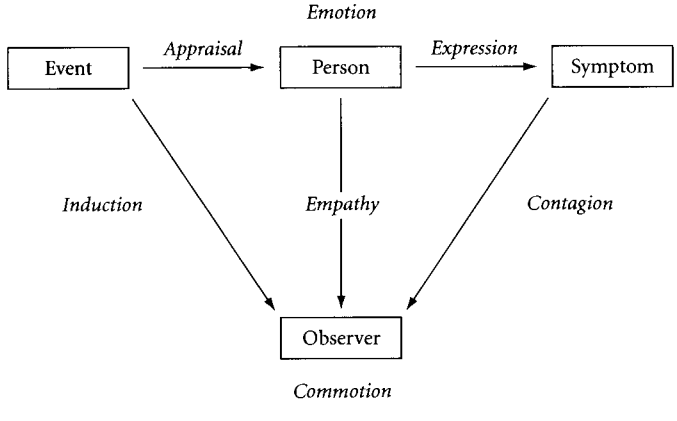



**Figure 16.1** Model of 'normal' emotion induction via appraisal (upper part) and mediated 'commotion' due to empathy or other mechanisms of emotional communication (lower part) (based on Scherer, 1998).

implication of cortical association areas. This research group has demonstrated that facial expressions of fear may have the same effect (Dimberg 1988). While empirical evidence for humans is lacking, one can assume that vocal expressions of emotions comparable to alarm calls in primates (see Jürgens 1988) may be similarly 'prepared' stimuli with respect to immediate, low-level meaning assignment. In consequence, musical stimuli sharing the acoustic characteristics of such fear vocalizations (sudden onset, high pitch, wide range, strong energy in the high frequency range) may be appraised by evolutionarily primitive but extremely powerful detection systems and may provoke, like pictures of spiders or facial expressions of fear, physiological defence responses (see Öhman 1988).

Similar low-level detection mechanisms can be demonstrated for the appraisal criteria of suddenness/novelty and intrinsic pleasantness (e.g. Scherer 1984, 1988). While the results of such rudimentary evaluation processes may not be emotions in the fully-fledged sense, as defined in Table 16.1, they may nevertheless produce positive valence reactions with respect to preferences or other types of affective phenomena short of fully-fledged emotions. There is some evidence that musical sounds can generate strong suddenness/novelty and, in particular, intrinsic pleasantness evaluations.

For example, Zentner and Kagan (1996, 1998) examined the hypothesis of an innate preferential bias favouring consonance over dissonance by exposing 4-month-old infants to consonant and dissonant versions of two melodies. Consonance and dissonance were created by composing the dissonant version in parallel minor seconds, the consonant version in parallel thirds using only two (synthetic) voices (for details regarding the stimuli, see Zentner & Kagan 1998). It was found that infants looked significantly longer at the source of sound and were less motorically active to consonant compared with dissonant versions of each melody. Further, fretting and turning away from the music source occurred more frequently during the dissonant than the consonant versions. Because no relation was found between infants' previous exposure to music (as assessed by questionnaire) and their behaviour in the experiment, the authors suggested that the human infant might possess a biological preparedness that makes consonance more attractive than dissonance. This tentative conclusion is not only consistent with perceptual work on infants' responses to consonant and dissonant intervals (e.g. Schellenberg & Trehub 1996), it is also supported by animal work which demonstrated a preference of consonant over dissonant chords in albino rats (Borchgrevink 1975).

It should be noted that similar automatic evaluation processes can occur for auditory stimuli that are not themselves evolutionarily prepared but that have been conditioned to such stimuli. Thus, LeDoux (1996) has convincingly shown that sounds conditioned to electrical shocks are processed by direct projections from the thalamus to the amygdala in rats (eliciting preparatory neuroendocrine and peripheral responses), before more extensive processing via the auditory cortex and the association areas occur. Similarly, pieces of music that have been consistently associated with evolutionarily prepared (or other powerful, unconditioned stimuli) can be evaluated at very low levels of the brain and evoke emotional responding.

One interesting issue that, so far, has not been much explored in the psychology of emotion, is that, either as a part of intrinsic valence detection or as a separate

mechanism, there may be automatic evaluation of aesthetic qualities. Thus, it cannot be excluded that there are some universal criteria of beauty that are evaluated automatically on the basis of visual and auditory stimulation and give rise to an affective response (see Etcoff, 1999). This is certainly the case for the schematic level of appraisal where learned aesthetic preference may play a powerful role.

The capacity of musical stimuli to elicit emotion via processes of higher-level appraisal are virtually unlimited. The appraisal criteria or dimensions suggested by appraisal theorists in particular concern the goal conduciveness/obstructiveness of an event, the individual's potential to cope with the consequences, as well as socionormative and self-ideal norms or standards, (for a summary, see Scherer 1999). Very simply put, appraisal theory predicts that if I have to listen to music that I tend to abhor and that in addition disturbs my concentration on an important task at hand (obstructiveness with respect to desires and goals), I will experience anger or aggravation if I believe it is possible to get the music to stop eventually, and desperation or resignation if I feel powerless.

It can reasonably be argued that music plays only a secondary role in this example, easily exchangeable with any other disturbing stimulus. Future work in this area will need to pay much greater attention to motivational constructs that are specific to music. Why do people listen to music? Why do they go to concerts? Why are some pieces chosen over others? These are the kinds of motivational urges that need to be taken into account when assessing goal conduciveness or obstruction in appraisal processes antecedent to emotions produced by music. More detailed discussions of this point can be found in the chapters by Sloboda and O'Neill, DeNora, and Gabrielsson (this volume).

The evaluation of the compatibility of a stimulus event with external standards (norms, cultural values) and internal standards (personal values) as part of emotion antecedent appraisal is highly relevant for emotion elicitation via music. There seem to be prescriptions specific to culture and/or historical periods as to what is aesthetically pleasing or beautiful and what is to be rejected as a violation of 'good taste' (e.g. Farnsworth 1969; Kenyon 1991; see also Lynxwiler & Gay 2000, for an interesting recent example). Throughout musical history, the social norm or standards criterion has been involved in powerful emotional reactions towards 'modern' music, which was seen as violating established standards of morality and decency. The well-known scandal provoked by the première of Igor Strawinsky's dissonant and polyrhythmic *Sacre du Printemps* or by the première of Edgar Varèse's surreal *Deserts*, both in Paris, are just two particularly drastic examples of strong emotional reactions to the perceived disregard of established standards.

Appraisal processes, then, are one possible route by which emotion may be generated by music. Before turning to other potential mechanisms of emotion production, let us mention certain limitations with this explanatory approach. First of all, the appraisal dimensions pertinent to an evaluation of music may only be in part the same as the dimensions implicated in the generation of everyday emotion. In fact, it has often been claimed by theorists that music does not seem to present stimulus antecedents that are usually implicated in emotion appraisal (e.g. Budd 1985; see also Davies, this volume; Sloboda & Juslin, this volume). Methodologically, examination of this route faces particular challenges, one of which is the microanalysis of specific appraisal processes that

continously precede emotional changes in *real time.* Appropriate methods for such sophisticated analyses do not exist and have yet to be invented.

*Memory*

Another central route that is well established as a mechanism of emotion induction is imagination or recall from memory. In these cases, a strong emotional reaction that an individual has experienced in the past resurges in memory, spontaneously or triggered by a specific cue, or is evoked due to an experimental instruction to vividly imagine the event. It has been suggested that expressive and physiological reaction patterns to emotion-inducing events are stored in memory together with the experiential content (Lang 1979; Lang *et al.* 1980). In consequence, it is often claimed that recall of past emotional experiences from memory and imagination can evoke similar emotional reactions as in the original experience. The empirical evidence for such memory-induced resurgence of expressive and physiological reactions is still scarce (but see Contrada *et al.* 1991; Dalton 1998; Tarrant *et al.* 1994) except for clinical evidence from the area of post-traumatic stress disorder (Pitman *et al.* 1999; van der Kolk 1997).

Music, like odours, seems to be a very powerful cue in bringing emotional experiences from memory back into awareness. This is not surprising, for two reasons: first, music is quite a pervasive element of social life and accompanies many highly significant events in an individual's life—religious ceremonies, marriage, burial rites, dancing, and other festivities, etc. Thus, there are many associations between musical elements and emotionally charged memories. Second, music, like odours, may be treated at lower levels of the brain that are particularly resistant to modifications by later input, contrary to cortically based episodic memory (e.g. LeDoux 1992).

*Empathy*

In the two routes described above, emotion elicitation was based either on the occurrence or the remembering of an event of major significance to the individual. However, emotions are sometimes elicited by just *observing* another person being affected by an event that is very important to him or her but not necessarily to us. Examples are pity with someone afflicted by malfortune or shared joy with someone who has succeeded beyond expectation. The lower part of Fig. 16.1 illustrates three possible origins of such 'commotion' (for further detail, see Scherer 1998). The arrow to the left marks a 'normal' case of emotion *induction* via appraisal: we evaluate the event that has produced the emotion in the person we observe and evaluate it in a similar fashion, with similar results. Even if we are not directly affected by the consequences of the event (e.g. an unreasonable action by a third person), we may evaluate the injustice or unreasonableness in exactly the same fashion as the person that is directly concerned and react equally with anger or irritation (possibly without knowing the reaction of the person concerned). The middle arrow illustrates the mechanism of *empathy.* Here the assumption is that we directly identify with the person concerned and feel 'with' her or him. The process of empathy requires sympathy—were we to dislike the person in question, we might actually feel the opposite emotion (e.g. joy over our enemy's anger). Finally, the right arrow illustrates a mechanism often called *emotional contagion.* Here, the essential feature is the observation of the motor expressions of the person concerned,

possibly without any knowledge of the event that caused the reaction. The assumption is that the sheer observation of strong motor expressions can produce similar muscular innervations in ourselves. This process, generally called 'motor mimicry' (Eisenberg & Strayer 1987), has been at the centre of a theory of emotional inference suggested by the German philosopher and psychologist, Lipps (1909). He argued that understanding the emotions of others occurs through *Einfühlung* (empathy), which is based on our mimicking, at least in a rudimentary fashion, the expressive patterns we see in the other (in other words, a contagion, at least of the expressive movement).

Why are these mechanisms of interest to the issues at hand? The link becomes clear when we try to extend the model in Fig. 16.1 to the situation of mass media communication such as film or TV. Assume that you are watching a character in a movie who is experiencing a very sad event, like the loss of a lover. In this case, there is no objective event, only the idea of an event that is the imaginary cause of the actor's acted feelings and the consequent expression manifestations. Assuming that the motor mimicry effect underlying contagion does not require an authentic experience, we presume that the viewer may still reproduce some of the features of the observed expression, albeit in a very weak and rudimentary fashion. In addition, it is not impossible that one will feel empathy, even though we know the emotion of the actor to be prescribed by the script. How else can we explain the tears shed in movie theatres? And, finally, we may even evaluate the imaginary event and be affected by this virtual appraisal. The adaptation of the emotion model to theatre or mass media-generated emotions is shown in Fig. 16.2 (for further detail, see Scherer 1998).

One can argue that listening to expressive music is very close to this example. Rather than an actor playing out a script, we have a musician performing a musical score. We may expect that the expressive movements in the music will lead to some kind of contagion, for example, a certain rhythm producing synchronized movements of the body. This mechanism will be described in greater detail below, because

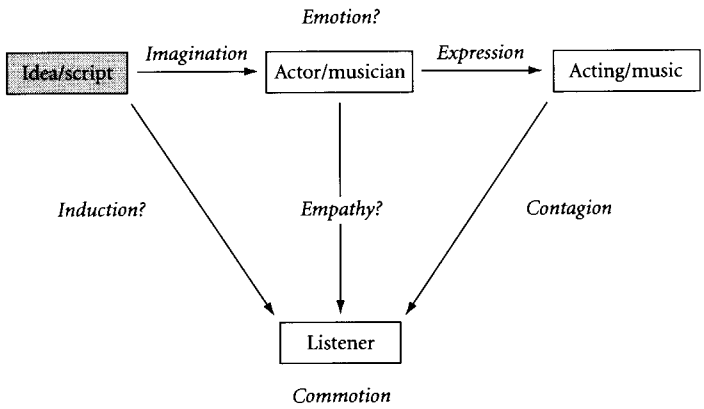



**Figure 16.2** Adaptation of the commotion model in Fig. 16.1 to the case of listening to emotionally expressive musical performances, based on the suggestions concerning emotion communication in the mass media in Scherer (1998). Grey shading indicates the potemtially virtual nature of ideas/scripts.

it may occur on a peripheral rather than a central level. However, there may also be a kind of empathy with the emotion presumed to be felt by the performer that may be construed in our imagination through an underlying 'idea' that is seen as responsible for the emotional state that is expressed (for example, the longing of the composer for his homeland, as in Dvoràk's 'New World Symphony'). Thus, emotion might be induced by the identification with a performer seen as living through an emotional experience produced by an underlying idea. We expect that such a process of musical empathy will be more likely in the case of listening to an admired performer (a listener feature, for example, an attitude) who is performing in a highly emotional manner (performance features such as skill, interpretation, and affective state). Finally, one might wonder whether, even in the absence of such performance cues, the 'underlying idea' might induce an emotional state on the basis of the evaluation of the musical structure of the piece alone. While these hypotheses are highly speculative, they do have the merit of bringing the investigation of the emotional effects of music theoretically closer to other kinds of live or mediated performances (theatre, movies, TV) that are generally considered to have the potential for emotion induction in the viewer, thereby providing opportunities for theoretical and empirical cross-fertilization.

### 16.1.3 Peripheral route production

*Proprioceptive feedback*

This route, while much less established than those described above, provides an interesting alternative, especially for some types of music. It is based on the idea that the emotion system consists of integrated components and that the system as a whole can be activated by manipulating the patterning of one of its components. For example, in a series of studies, Ekman and Levenson have provided some evidence that it may be possible to induce physiological and experiential emotion reactions by asking an individual to produce the facial patterning that is characteristic of certain emotions (without the person being aware of this purpose; see Ekman *et al.* 1983). While these findings are often contested, explaining the effects by a variety of experimental artefacts (e.g. Boiten *et al.* 1994), Ekman, Levenson, and their collaborators have systematically replied to these criticisms by adding further controls in replication experiments (P. Ekman, personal communication). The general idea is also consistent with proprioceptive feedback theories (e.g. facial feedback; see McIntosh 1996) that claim, in their strong form, that subjective feeling can be produced, or in their weak form, enhanced or intensified, by increased or uninhibited motor expression.

Thus, if music could be assumed to systematically influence one of the emotion components, peripheral mechanisms might be invoked to explain a spread to other emotion components, thus in fact producing emotion states that did not exist before. One potential candidate for such influence is rhythm. We all know the contagious effect strong musical rhythm has, at least on susceptible individuals who find it difficult not to move their heads or their legs in unison with the rhythm (e.g. in the case of dance rhythms, marches, or techno beat). Recent evidence suggests that such coupling of internal rhythms to external drivers, as originally described by Byers (1976), might be present at

a very early age (Rochat & Striano 1999). If there is, indeed, a fundamental tendency for synchronization of internal biophysiological oscillators to external auditory rhythms, such coupling may provide a promising explanatory venue for the emotion-inducing effects of music. For example, given the close relationship of respiration and cardiovascular function, a change of respiration through musical rhythm may have an impact on a variety of neurophysiological systems (Boiten *et al.* 1994), in many ways similar to emotion-induced physiological changes.

*Facilitating the expression of pre-existing emotions*

So far, we have discussed the production of emotions where there were none before. Another potential effect of listening to emotionally arousing music can be the weakening or the elimination of control or regulation efforts. Due to sociocultural display and feeling rules, or because of strategic considerations, emotional reactions are often highly controlled or regulated, both with respect to motor expression and subjective feeling (Ekman 1984; Goffman 1959, 1971; Hochschild 1983; Scherer 2000*b*). One interesting effect of music, so far reported more often in an anecdotal fashion than empirically established, seems to be a loosening of such control or regulation effects. Movie-goers often report that, while watching a moving film, when the music started they could no longer hold back their tears. This effect may be partly due to similar contagion effects as described in the section above: when a pre-existing tendency toward specific motor expressions and physiological reactions is reinforced by external stimulation, it may be much harder to keep affect under control.

This remark on facilitating mechanisms concludes our overview of some of the mechanisms that may underlie the production of emotion through music. It should be noted, however, that we do not suggest that these mechanisms function in an exclusive manner. Rather, we assume that, generally, all of these mechanisms are operative jointly but having, depending on the circumstances, more or less of an impact on the emotional state produced. Thus, the peripheral route via rhythm entertainment may be particularly powerful for certain types of musical structures (lullabies, waltzes, or marches) and for particular occasions (a ballroom setting or a carnival march). In contrast, the central route, both via appraisal, associative memory, and empathy effects may be dominant in cases where one listens to well-known music in a reflexive mood, as in a concert setting.

### 16.1.4 Evidence

While there is massive evidence that listeners consistently *attribute* specific emotions to certain types of music, it is much more difficult to find convincing evidence that music can indeed *produce* emotion through one of the mechanisms described above. This task is rendered all the more difficult by the problem of deciding on a valid criterion for the *presence* of a certain emotion. As shown below, most studies have used a verbal report criterion, that is, listeners reporting that they are actually experiencing a certain emotion or affective feeling state as a consequence of listening to a particular piece of music. However, the danger of experimental artefacts, demand characteristics, social desirability, and many other problems besetting verbal report make one wonder whether verbal report can really be considered the gold standard as a criterion for genuine emotional experience (see below).

Ideally, the criterion to be used should be linked to the psychological definition of emotion as a hypothetical construct, that is, the operationalization, accessible to empirical measurement, as to the necessary and sufficient conditions to diagnose the presence, including beginning and end, of an emotional episode. So far, there is no consensus on such an operational definition in the literature. However, there is increasing convergence on the notion that emotions need to be seen as multicomponential phenomena (Scherer 2000*a*), involving at least the components of physiological arousal, motor expressive behaviour, subjective feeling, and possibly also a motivational component in the form of action tendencies and a cognitive component consisting of emotion-constituent appraisal and reappraisal (of external and internal events, including the emotional reaction itself). In addition, there is widespread agreement on the need to conceive of emotions as constantly changing processes, or episodes, rather than steady states. Thus, Scherer (1984, 1993, 2000*c*, 2001), proposing a *component–process model of emotion*, has suggested that researchers should use the criterion of coupling or synchronization of the different emotion components as evidence that an individual *has* a certain emotion at a particular point in time (see Table 16.1). Using this strict criterion, the claim that music induces emotion would have to be instantiated by studies showing that listening to a piece of music reliably leads (at least for a clearly specified group of listeners and under specified context conditions) to a synchronization of cognitive, physiological, motor-expressive, motivational (in the sense of evoked action tendencies), and experiential (subjective feeling) processes.

Although a few musical mood induction studies have incorporated multiple measures of affect (e.g. Albersnagel 1988; Kenealy 1988; Pignatiello *et al.* 1989; see also below), it may take a while until a study that measures at least three of the different components and determines their synchronization, is designed, excecuted, and published. In consequence, we will review the available evidence on the basis of individual components, providing an inventory of empirical findings separately for each of the components.

*Cognitive changes*

The philosophical literature across the ages has credited music with multiple effects on the mind, especially the power to provoke specific images or ideas that were, in turn, assumed to produce some of the emotional effects (for recent overviews of the philosophical literature, see Budd 1985; Davies 1994). Similarly, music is often supposed to affect judgment and evaluation. Unfortunately, few of these presumed effects have been empirically validated, demonstrating significant, experimentally induced effects for appropriately operationalized variables. The most frequently examined effects of music on cognitive performance have focused on memory.

For example, Martin and Metha (1997) examined the effects of musical mood induction on recall of childhood memories and found that music had an impact on the total number of memories recalled in the happy condition, but that there was no significant effect of sad music on eliciting sad memory recall. Balch *et al.* (1999) demonstrated mood-dependent memory effects following mood induction through music (i.e. the facilitation of remembering events one has memorized under the effect of a certain emotion when a similar emotion is elicited) (Blaney 1986; Forgas 1991). Their results suggest that it seems to be mostly the valence dimension of the emotional state that triggers the effect.

In one of the few experimental studies on the effect of music on judgmental processes, Bouhuys *et al.* (1995) showed that participants who felt depressed after listening to sad music perceived more rejection/sadness in faces with ambiguous emotional expressions and less invitation/happiness in faces with clear, unambiguous expressions. Some pertinent work may also be found in the studies on the effect of music on consumer attitudes and behaviour (Hargreaves & North 1997; see also Brown & Mankowski 1993, which is discussed below under 'subjective experience').

*Physiological arousal*

The idea that music can influence the ANS (also called the peripheral or vegetative nervous system) both in an arousing and a calming fashion is a very old one, going back at least as far as Greek philosophy (Bartlett 1996). Similarly, in many contemporary, popular treatments of this issue, the power of music to affect our autonomic nervous system and even the immune functions of our body are evoked. Thus, Ornstein and Sobel (1989) suggest that music evokes a calmness that may decrease blood pressure and reduce stress, and that it may help control hormones that suppress the immune system.

As to the empirical study of this idea, the French composer Grétry (1741–1813) is credited with the first reported attempt to measure cardiovascular effects of singing by counting the pulse rate by placing a finger on an artery (Diserens 1923, cited in Bartlett 1996). Serious study of the notion that music affects the ANS started as soon as the first instruments for physiological measurement became available and has continued unabatedly for the last 120 years, the period reviewed by Bartlett (1996). In general, researchers were concerned with the arousing or stimulating versus calming or sedating effects of music. Bartlett concludes that in the majority of studies (estimated at 61 per cent), an effect of music on different ANS parameters, corresponding to the prior hypothesis (e.g. increased heart rate and muscle tension for arousing music vs. decreased heart rate, decreased muscle tension, increased skin temperature, increased skin conductance for calming music) was found. There were, however, important differences between physiological variables: while the cardiovascular predictions were confirmed in only 39 per cent of the studies, skin temperature predictions were confirmed in almost all studies.

The large majority of studies reviewed by Bartlett (1996) focused on the issue of arousal versus relaxation as emotional effects of music. Given the period of time in which most of these studies were conducted, this is congruent with the dominant view of emotion during the heydays of behaviourism in psychology. This view defined the emotions as states of the organism characterized by different degrees of psychophysiological arousal or activation (Duffy 1941).

More recently, following the work of Tomkins (1962, 1963, 1984), and particularly his followers Ekman (1984, 1992) and Izard (1971, 1994), the Darwinian notion of a limited number of 'basic' or 'fundamental' emotions serving adaptational purposes (which has been postulated by philosophers since the dawn of time) reasserted itself (see Sloboda & Juslin, this volume, for further discussion). If the induction of emotion via music is seen from this theoretical angle, the issue becomes to demonstrate which of the basic emotions are elicited by a specific piece of music, as judged from the physiological

response pattern observed. This question is consistent with the idea prevalent in musical circles ever since the baroque *Affectenlehre* (Mattheson 1739), claiming that music can express specific emotions, a position violently attacked by Hanslick (1854). One could presume that if music can express discrete emotions, it might also be able to generate or induce these same emotions (see 'production rules' above). The evidence required to prove such induction is the production, via music, of the specific patterns of physiological reactions that are postulated for the basic emotions by theorists in that tradition (Levenson 1992) and that have been demonstrated, more or less convincingly, in studies using film stimuli, imagination, or other induction techniques (Ekman *et al.* 1983; but see also Cacioppo *et al.* 1993; Stemmler 1996).

In an exploratory study, Krumhansl (1997) attempted to examine this prediction. She presented 38 Berkeley students with six excerpts from pieces of classical music that are generally expected to induce specific emotional states (and which were consensually rated, by another group of judges, to express the emotions of sadness, fear, happiness, or tension). Eleven physiological parameters, the large majority tapping respiration and cardiovascular functioning, were measured while subjects were listening to the music excerpts (average duration 3 minutes). The results supported the general prediction that music does indeed affect the ANS: during the listening periods, many of the physiological parameters had values significantly different from those during the pretrial (resting) period. Furthermore, Krumhansl was able to show statistically significant differences between the music samples (chosen to express the three emotions, sadness, fear, and happiness) in the extent to which they seemed to influence specific parameters (e.g. sadness excerpts producing large changes in heart rate, blood pressure, skin conductance, and temperature; fear excerpts producing large changes in pulse transit time and amplitude; and happy excerpts producing large changes in respiratory patterns). The author found that these changes were rather well correlated with subjectively reported experiences when analysed on the level of dynamic change over time, but not for the averaged physiological parameters and an overall judgment of emotion for a piece.

However, with respect to the central piece of evidence, the matching of the emotion-specific physiological response patterns, the results were far less encouraging. While, in some cases, the changes for a particular parameter went in a similar direction as the theoretically postulated ones, it would be difficult to argue that complete physiological response patterns, prototypical for 'discrete' or 'basic' emotions, were found. In any case, this would be a tall order, given that even psychologically highly sophisticated induction techniques, designed to produce specific emotions, so far have not yielded the kind of specificity in physiological response patterning that is expected by discrete emotion theorists (see Cacioppo *et al.* 1993; Stemmler 1996).

In a well-designed and higly controlled study using very naturalistic inductions of anger (harassment by experimenter) and fear (dangerous equipment failure), Stemmler and his collaborators (2001) were able to find systematic differences between anger and fear that can be directly linked to the differential action tendencies (i.e. fight and flight; see Frijda 1986; Scherer 1984) for these two emotions. The authors conclude that significant differences can only be expected on the basis of functional demands on behaviour and express some doubt as to the existence of highly specific patterns for many different emotions.

In another music induction study, Nyklicek *et al.* (1997) presented 26 undergraduates with 12 classical music excerpts (of 120–220 seconds' duration), with three excerpts each having been found to reliably express emotional states described as happy, serene, sad, or agitated. Physiological measurement again focused on cardiorespiratory variables. Rather than try to match theoretically defined patterns for discrete emotions, the authors employed multidimensional discriminant analysis to examine whether the reactions to the excerpts with different affective tones could be reliably distinguished on the basis of the physiological patterns. The discriminant analysis classified the reactions to the four types of excerpts with an accuracy of 46.5 per cent, the first discrimination function being largely determined by respiration parameters (reflecting an arousal or activation dimension), and the second dimension being determined by dynamic aspects of the cardiovascular processes (interbeat interval and left ventricular ejection time). While the first dimension was directly linked to subjectively perceived arousal, the second is more difficult to describe with respect to subjective meaning of the emotion space (it did not directly represent the valence dimension of subjective experience). The authors interpret their findings as evidence for both a discrete emotion and a dimensional model, arguing that the lower-order discrete emotions may be discriminated in a space formed by higher-order dimensions. This is reminiscent of the claims made by all dimensional models, starting with Wundt (1974). However, it should be noted that emotion theorists proposing dimensional models generally focus on the subjective feeling rather than the physiological or expressive components of emotion (see Scherer 2000*a*).

In a recent series of studies, Witvliet and Vrana (1996; Witvliet *et al.* 1998) looked at the effect of music on autonomic measures of skin conductance and heart rate, as well as on the somatic parameters of electromyographic change scores for the zygomatic, corrugator, and orbicularis oris muscles in the face. The music excerpts had been preselected with respect to their scores on the valence and arousal dimensions. In line with the earlier literature reviewed by Bartlett (1996), their results showed clear arousal effects on the autonomic measures with increased levels of skin conductance and heart rate, as well as increased tension of the orbicularis oculi muscle, during listening to excerpts judged as high on the activity dimension. In addition, they were among the first to show *valence* effects for music, particularly for the somatic measures. As predicted by discrete emotion and appraisal theorists (Ekman 1992; Scherer 1986, 1992), there was increased corrugator activity (used in frowning) while listening to music with negative valence and increased zygomatic activity (used in smiling) for music with positive valence (especially when combined with high arousal). This is consistent with the results of a more recent study (Lundqvist *et al.* 2000), in which facial electromyography, autonomic activity, and emotional experience were measured while subjects listened to happy and sad music. The happy music elicited larger zygomaticus activity, larger skin conductance, lower finger temperature, more (self-reported) happiness and less sadness compared with the sad music.

Limitations of the studies reviewed above include the lack of sensitivity to the dynamic unfolding of the effects of music over time (often with very rapid changes) and the lack of attention to the determining role of musical structure. Indeed, much of this research only examined physiological reactions integrated over a relatively long time interval. Furthermore, often no expert analyses of the structural features (musical

score) were done to examine, or at least derive hypotheses about, what exactly in the music might have been responsible for the reported physiological phenomena.

A pioneering study by Vaitl *et al.* (1993) took both of these aspects into account. In a field study carried out in the festival theatre of Bayreuth in the summers of 1987 and 1988, 27 listeners were psychophysiologically examined while listening to a number of Wagner operas, from *Tannhäuser* to *Parsifal.* However, only data produced by three subjects in response to the leitmotivs during *The Mastersingers of Nuremberg* were analysed in depth. Electrodermal responses, respiratory activity, and ratings of emotional arousal were continuously recorded and analysed during the leitmotivs. It was shown that physiological responses differed markedly with respect to the leitmotivs and their musical features (e.g. melody, rhythm, continuation). An important additional finding was the interindividually variable, but overall weak, correspondence between emotional arousal ratings and physiological responses. Although this case study has obvious limitations, it is a good model to follow for future studies, combining dynamic measurement, on line, of physiological changes and feelings of emotional arousal within a live performance and linking the observed values to particular features of musical structure.

*Motor expressive behaviour and action tendencies*

From ancient times, there has been an assumption that movement is a key component, if not the underlying mechanism, of emotional reactions to music. This position is, for example, articulated by Aristotle as follows

> Why is sound the only sensation that excites the feelings? Even melody without words has feeling. But this is not the case for color or smell or taste . . . But we feel the motion which follows sound . . . These motions stimulate action, and this action is a sign of feeling (cited in Helmholtz 1954, p. 251).

Among modern philosophers, it is, in particular, Edmund Gurney and Carroll Pratt who gave a central place to movement in their theories of emotional effects of music (Gurney 1880, Pratt 1931).

The fact that much of modern-day music listening occurs under ritualized conditions in the concert hall, under more relaxed conditions in one's living room or car, or via earphones while pursuing all kinds of activities, masks the fact that throughout history much music was composed with very specific action tendencies in mind—getting people to sing and dance, march and fight, or work and play together. Composers of dance music and marches attempt to entice listeners to produce the appropriate rhythmic movements, and many people actually report that they cannot sit still and have trouble suppressing rhythmic body sway and other movements when hearing some bars of a Strauss waltz. On the other hand, lullabies evolved because of their presumed ability to put children to sleep, thus leading to the opposite effect of inhibition of motor arousal (Trehub & Schellenberg 1995).

One would expect, then, to find mechanisms that can produce motor expression and action tendencies via the peripheral production route described above. In one of the few attempts at theoretical foundation, Todd (1992) suggests that expressive sounds may induce a percept of self-motion in the listener via the central vestibular system. The

argument is that expression based on elementary mechanics may sound natural because the vestibular system evolved to deal with similar kinds of motions (i.e. gravity and linear and rotational acceleration). As mentioned above, Byers (1976) reviewed potential mechanisms for rhythm induction. Clynes (1977) also argued for pervasive emotion-specific dynamic motor patterns.

Unfortunately, hard empirical, especially experimental, evidence for the production of motor patterns and action tendencies by music is rare (but for some attempts, see Fraisse *et al.* 1953; Harrer & Harrer 1977). Kneutgen (1964) discussed the biological significance of rhythm synchronization in different species and emphasized the role of music in driving biological rhythms. In an interesting pilot study (Kneutgen 1970), an Argentinian lullaby was played to four listeners in sessions lasting from 30 minutes to 2 hours (almost every day for a period of 3 months). It was found that the lullaby had the effect of markedly decreasing heart rate. In addition, the breathing rhythm became synchronized with the rhythm of the music. Interestingly, no such effects were found in a control group exposed to jazz music.

DeVries (1991) used Clynes' (1977) *sentograph* to measure affective motor responses to music and reported that the pressure patterns shown for 11 brief musical excerpts were similar across subjects and seem correlated with the expressive content of the music. However, so far convergent evidence from methodologically sophisticated studies on the validity of the sentograph are missing, and also the theoretical assumptions underlying its use seem rather elusive (but for studies in this area, see Carmean 1980; Gorman & Crain 1974; Nettlebeck *et al.* 1989; Trussoni *et al.* 1988).

Generally, there have been very few controlled studies using measures other than verbal report, on the induction of dynamic motor behaviour or specific action tendencies by music. The same is true for motor expression in the voice and face. In one of the very few studies looking at the facilitation of vocalization by music, Pujol (1994) varied the instrumentation (flute vs. bells) in two types of melodies (pentatonic vs. major) and studied the effects on physiological (respiration and pulse rate) and behavioural responses (eye movement, facial expression, vocalizations, and motor movement) of 15 profoundly handicapped persons. The music (particularly the major flute melody) stimulated the production of vocalizations, an effect which decreased over time.

Given the assumption that the face is 'the primary theatre' of the emotions (Ekman 1992; Tomkins 1984), one would expect to find much evidence for the facilitation of facial expression via music. Again, appropriate studies are rare. However, as reviewed above, Witvliet and Vrana (1996) reported evidence that different kinds of music differentially stimulate the innervation of the zygomatic muscle or corrugator muscles (see also Lundquist *et al.* 2000).

Studies on the effects of music on the tendency to perform certain actions (e.g. prosocial or aggressive acts) tend to be rare. In an experiment by Fried and Berkowitz (1979), different groups heard 7-minute-long musical selections that were either soothing, stimulating, or aversive in nature, while a fourth group was not exposed to any music. Those who heard the soothing music were most apt to show altruistic behaviour immediately afterwards (volunteering for another study). Mood ratings indicated that the soothing and stimulating music created somewhat different positive moods, while the aversive music tended to arouse negative feelings.

*Subjective experience (self-report of feeling)*

Among the methods used to study emotion induction through music, the self-report of feeling has been the most widely used approach. However, most studies having examined emotion *production* via self-report were not carried out within the area of music psychology, where the main emphasis has been on the perception of emotion in music. Assessments of emotion production via self-report is typically found in *mood induction* studies. The aim of such studies is often to examine the effects of different affect states, moods, and emotions on cognitive and evaluative processes, memory, behaviour, and physiology. Such 'instrumental' use of music as a means of emotion or mood manipulation has rapidly increased in recent years, placing music among the most widely used techniques for mood induction in psychology (for a review of mood induction studies, see Westermann *et al.* 1996).

There are two ways to examine the success of a given mood induction. One is indirect, relying on significant effects of the mood conditions on the dependent variables of interest. However, in addition, a self-report of mood or emotion is often used as a so-called manipulation check for the success of the mood induction. Typically, after the mood manipulation, the subject is given a self-report measure of mood. The results are then compared across mood conditions and significant differences among the conditions are interpreted as evidence for the success of the mood induction. Some of the most frequently used self-report instruments to assess emotion or mood production via music include: the multiple affect adjective check list (MAACL(-R-)) (Zuckerman & Lybin 1985); the Nowlis mood adjective checklist (Maslach 1979); the visual analogue scales (VAS) (Bond & Lader 1974); and the brief mood introspection scale (BMIS) (Mayer & Gaschke 1988).

The use of self-report measures as evidence for emotion production has both advantages and disadvantages. It remains the only method that allows access to the subjective emotional experience. Therefore, it should not be replaced by any other measure. However, the data collected in this way are sensitive to the following potential biases and artefacts that need to be carefully considered before conclusions are drawn. Firstly, demand characteristics can be responsible for the mood effects found. Secondly, listeners may confuse the emotions expressed in the music with what they actually feel unless they are specifically requested to distinguish between the two emotion modalities, as has been the case in our recent studies (Zentner *et al.* 2000).

To be taken as indicators of mood effects of music, then, self-report measures have to be validated against other criteria, or they have to be experimentally controlled. Following the latter avenue, Kenealy (1988) has compared self-reported mood effects in two conditions; one in which demand characteristics were intentionally created and one in which the effect of demand characteristics was minimized. She found no effects of demand characteristic on self-reported mood effects, concluding that the 'music procedure is relatively free from demand characteristics' (Kenealy 1988, p. 46). However, one could argue that the so-called 'no demand' condition was in reality also affected by demand characteristics. Subjects were left alone in a room by the experimenter who left on the pretext of having something urgent to attend to. Then, happy or sad music (depending on the condition) was played for 8 minutes. The subject, knowing they were in a psychological experiment and suddenly hearing music, might well have guessed the

purpose of the experiment even without explicit demands and even without admitting it after the experiment. Such difficulties in the interpretation of the results notwithstanding, Kenealy's research is important in suggesting possible strategies for controlling demand effects.

The second criticism, confusion of perception and induction, is particularly relevant to studies in which subjects are explicitly instructed to report what they feel in response to the music. Kivy (1989) has argued that listeners make a fundamental attribution error in that they *habitually* take the expressive properties of the music for what they feel. This position is not new. In fact, it has already been summarized by Meyer (1956, p. 8) as follows: 'it may well be that when a listener reports that he felt this or that emotion, he is describing the emotion which he believes the passage is supposed to indicate, not anything which he himself has experienced'.

To our knowledge there is little, if any, empirical evidence to back up this claim. For this reason, we recently conducted a study in which participants were instructed to describe both what they *felt* and what they *perceived* in response to different genres of music. We found that the type of instruction (felt vs. perceived) changed the results quite dramatically (Zentner *et al.* 2000). Overall, the majority of emotions were reported much less frequently when the participants were instructed to focus on the *production* rather than the *perception* of emotion. This does, of course, not suggest that particular emotions are never felt in response to music, as Kivy would have it. (As a matter of fact, for particular emotions the reverse was true—they appeared to be more readily felt than perceived.) However, it indicates that the distinction between perception and production of emotion is empirically valid and that it is essential to take this distinction into account when instructing people to report their emotional responses to music. The studies by Kenealy (1988) and preliminary evidence from our group show that, although verbal reports of moods and emotions have their limitations, there still is much to be done to improve their quality.

In addition to asking listeners directly about their emotional reactions, there is also the possibility of assessing the feeling state without asking subjects directly about their emotional experience. Little use has been made of this possibility in the present area. In a recent review of mood effects on judgments and social cognitions, Forgas (in press) discusses several paradigms that could also be easily applied to the study of music and emotion. In some of these paradigms, subjects are asked to rate something other than affect or mood, for example certain individual dispositions related to attitudes, personality, or self-esteem. In a typical study (Brown & Mankowski 1993), a musical mood induction was used to induce positive, negative, or neutral affect. Afterwards, listeners evaluated *their* specific qualities and characteristics (How smart are you? How kind are you?). It was found that subjects rated themselves more favourably after listening to happy compared with sad music (see also Bouhuys *et al.* (1995), which is discussed above under 'cognitive changes'). Interestingly, this effect was stronger for subjects with low self-esteem, suggesting greater sensitivity to mood manipulation for this particular group.

In sum, because access to the subjective emotional experience of a listener is logically impossible without verbal report, optimization and extension of these methods is indicated. Perhaps, in the future, various brain-imaging techniques can be used as additional

indicators of subjective feeling state (cf. Peretz, this volume). At present, verbal report remains the only means to assess this important component of emotional experience. Therefore, it is surprising that there has not been more effort made to systematically develop a vocabulary suited to describe what listeners feel in response to music. This is even more striking if we consider the doubts, expressed by several theorists, that emotional responses to music may not resemble everyday life or basic emotions, such as anger, fear, joy, sadness, shame, and disgust. Thus, Lippmann (1953, p. 569) suggested: 'Musical feelings have their own character: they are not the feelings we know and roughly name in our experience outside of music . . . Thus music may be an emotional experience, and still not represent emotional contexts belonging to other areas of life, for the emotions it formulates are not identical with those accompanying extra-musical experience, nor does the one kind necessarily remind us of the other'.

As is often the case in this domain, there is little empirical evidence to confirm or invalidate this conjecture. Few efforts have so far been made to examine whether the basic emotion categories derived from emotion psychology are theoretically sensible, empirically valid, or musically plausible for describing the emotional experience induced by music. The few systematic attempts at creating a taxonomy of 'musical emotions' (e.g. Hevner 1936; Rigg 1964; Wedin 1972; see also Gabrielsson & Lindström, this volume) were entirely focused on perceived, not aroused, emotions. As has been pointed out by Sloboda (1992, p. 36), 'The relevance of these studies to emotional experience, however, is not proven. It is possible to make character judgements on the basis of conventional characteristics without experiencing any emotion whatsoever'. Moreover, these attempts lacked rigour, as has been pointed out recently by Gabrielsson (1998, p. 2): 'Procedures followed for this purpose are not always explicitly described and are rarely, if ever, based on some theory, rather on the researcher's own experiences and ideas and material used by earlier researchers'. Clearly, then, adopting a systematic approach for deriving a taxonomy of musical emotions while focusing on the aroused, not the perceived, emotions, is an important task for future research in this area.

Such work requires a theoretical background that does not prejudge the issue, as is the case with emotion theories that either, like discrete emotion theories, start from the assumption of a limited number of basic or fundamental emotions or that, like dimensional theories, focus exclusively on the valence and arousal dimensions of emotional feeling (see Scherer 2000*a*). Unfortunately, neither of these make any provision for emotion processes that may be specific to music. Scherer (1984) has suggested that there may be as many different emotions as there are differentiated outcomes of appraisal processes, rejecting the very limited taxonomy that is imposed by the semantic structure of the 'basic emotion' vocabulary in a specific language. Thus, componential process approaches (Scherer 2001) are perhaps better suited to provide the theoretical underpinnings for research on the specific emotions produced by different types of music. It should be recalled that this does not mean that conscious appraisal implying all criteria is necessarily involved in emotion elicitation via music (see above). It may be necessary to discard the standard, 'fundamental' emotion terms altogether and to focus more strongly on terms for more subtle, music-specific emotions (such as longing, tenderness, awe, activation, solemnity) that seem to describe the effects of music more readily than anger, sadness, or fear (Zentner *et al.* 2000).

## 16.2 Perspectives for future research

The review of the research evidence presented above has suggested that we are far from being able to provide a clear answer to the question of how music can actually *produce* emotional states. Apart from a general lack of sufficient research evidence, particularly with respect to non-verbal measures of affect, it is difficult to avoid the impression that this research domain suffers from lack of theoretical rigour. This is a bad omen for future research, since it is to be feared that additional, isolated research efforts with little or no theoretical underpinnings are more likely to add to the current confusion than to the insight to which the researchers aspire. We believe that much of the problem is due to the lack of differentiation between the different types of affective processes as described in Table 16.1. Clearly, confusion is hard to avoid when researchers apply different criteria for the evidence they require to infer the presence of the successful induction of an affective process via music. Therefore, it would be advisable for researchers to specify more carefully what kind of affective process they attempt to induce through the presentation of musical stimuli.

We suggest the adoption of the distinctions proposed in Table 16.1 for further work in this area. In what follows, we will attempt to specify in greater detail what criteria, both verbal and non-verbal, might be chosen to infer the presence of the respective process. Since it seems unrealistic to assume that music can produce any of the long-term affect dispositions in rows 4–6 of Table 16.1, only the short-term processes in rows 1–3 are discussed. Because there is no established consensus for either the classification proposed or the criteria suggested, the recommendations made below will have to be judged on the basis of their plausibility.

*Preferences*, defined as evaluative judgments of stimuli in the sense of liking or disliking, or preferring or not over another stimulus, are the simplest form of affect manifestations and some theoreticians may deny that they have an affective as well as a cognitive basis. However, because they are treated as emotional phenomena by some theorists (Zajonc 1980; Zajonc & Markus 1984) and because they are directly linked to valence which is seen by many theorists (especially the dimensional theorists) as an essential component of emotion (see Scherer 2000*a*), it seems mandatory to include them here. The verbal criterion for the existence and strength of induced preferences is the verbal claim by a person to like or dislike a certain piece of music with a certain intensity or certainty, or, in a comparative context, to like a piece of music more than (or prefer it over) another piece. One can also ask listeners to sort pieces of music in ascending fashion for greater liking or arrange them on a bipolar scale from extreme disliking to extreme liking.

On the non-verbal level, preferences can be assessed rather precisely by observing choice behaviour, that is, the choice of listening to one type of music rather than another, the length of listening behaviour, and other behavioural measures. Another non-verbal measure of liking is facial expression, measured either through observer coding of visible behaviour (Ekman 1992) or facial electromyography measurement (Cacioppo *et al.* 1993). The presence and strength of musculus zygomaticus innervation (smiling) can be taken as an indicator of degree of liking, and of musculus corrugator innervation (frowning) as a sign of degree of dislike. On the level of physiological measurement,

Lang and his collaborators have suggested that the startle probe (e.g. Lang *et al.* 1990) can be used as a measure of valence (positive or negative feeling state). It should be noted that these behavioural changes are expected to be stimulus-driven, that is, they should only be present in the anticipation, presence, and possibly immediate aftermath of the stimulus, in this case, the music presentation. In contrast, the cognitive judgment, and verbal expression, of liking should be permanently accessible once a stimulus is identified and remembered.

Music seems to elicit preferences almost automatically, as the work on low-level automatic appraisal, reviewed in Section 16.1.2, suggests. Even if there is no comprehensive empirical evidence at present, there is little music that people find completely neutral and for which they are impartial as to whether to listen or not, if given a choice. Judging from informal evidence, most individuals seem to have little trouble in rating the degree of preference or liking (or disliking) on standard scales. However, it seems that there are powerful differences with respect to liking and disliking (e.g. Behne 1997). This seems quite natural since music that nobody likes is unlikely to survive (although detractors of modern music often claim that this is listened to for reasons other than liking). Thus, research on the induction of preference or liking may only be of interest in a differential context.

Contrary to preferences, *moods* are generally not elicited by concrete stimulus events at a particular point in time. Rather, they are defined as diffuse affect states, most pronounced as a change in subjective feeling, of low intensity but relatively long duration, often without apparent cause and thus persisting across encounters with multiple stimulus events. Because the labels for moods and emotions are often similar, care has to be taken to clearly distinguish the two in the instructions by specifying the long-term, non-event-related character of mood states (e.g. scales of good or bad mood, touchy vs. serene mood, etc.). Similarly, behavioural indices have to remain present over a lengthier period of time and not be related to an immediately preceding stimulus. This obviously rules out behavioural choice measures because these are automatically stimulus-driven. However, enduring evidence for morose versus serene facial expression, or prosocial versus aggressive acts, might serve as behavioural indicators.

Music may well be able to produce mood effects. At least this is what the music industry suggests, in particular for positive moods, by selling selections of music destined to produce relaxation or to generally improve one's mood (Bruner 1990; for concrete marketed examples see the mood-inducing compilations by Conifer Classics 1997 and Sony 1995). Similarly, claims for the effectiveness of music therapy are based on the assumption that the effect outlasts the presentation of the music (cf. Bunt & Pavlicevic, this volume). In contrast, music designed to incite customers to buy something or to calm aggression at football games, may rely on co-presence for its effect. There are only a few studies that have addressed the issue by carefully selecting valid *mood* criteria as defined above. Future research in this area, using appropriate measures, seems of prime importance, given the important implications for applied settings.

Can music induce *emotions* in the strict sense of Table 16.1, that is, defined as relatively brief episodes of synchronized responses of all or more organismic subsystems in response to the evaluation of an external or internal event as being of major significance? According to the senior author's theory, this implies that the emotion

process is directly produced by an appraisal of the musical event and that the response pattern generated is specific to the event. While it may last beyond the presence of the event, in the case of certain emotions even for relatively long periods (e.g. Frijda *et al.* 1991; Scherer & Wallbott 1994), it is generally expected to be of shorter duration and of higher intensity than moods. Thus, contrary to preferences, valence cues should be accompanied by the presence of arousal indicators in the ANS. While these might also be present for mood, they should be more intense and less stable than in the case of mood. Furthermore, emotion processes should entail motivational changes and preparation of adaptive action in the form of modes of action readiness or action tendencies (Frijda 1986; Scherer 1984).

As mentioned above, Scherer (1984, 1993) suggested that the most salient criterion of the presence of an *emotion* process (in the strict sense defined in Table 16.1) is a high degree of coupling or synchronization of all organismic subsystems, including cognition, during the emotion episode. In order to answer the question of whether music induces emotion in this strict sense, all pertinent indicators in the respective organismic subsystems need to be measured and the degree of their synchronization assessed. So far, none of the studies in this area has attempted to do this, and it is not to be denied that it is a formidable undertaking. Yet, if one agrees to define emotions in this narrow or strict sense, it will be necessary to engage in comprehensive verbal and non-verbal measurement and powerful mathematical modelling of coupling or synchronization.

It is rather probable, of course, that one will not be able to find as intense and highly synchronized response patterns as found in the case of violent rage leading to fighting, for example. Ellsworth (1994) has pointed out that music is quite a special emotion-inducing stimulus in that it is difficult to specify the underlying concerns or goals that tend to be the motor of emotion (cf. Sloboda & Juslin, this volume). One could add, that it is also difficult to see which modes of action readiness or action tendencies that, according to Frijda (1986), define the quality of emotions, can be evoked by music. The exceptions are action tendencies like singing or dancing that are quite regularly provoked by certain types of music. However, it is likely that these action tendencies are provoked by peripheral entrainment, as described above, rather than by centrally controlled motivational changes as is generally the case in the provocation of adaptive action tendencies in 'real' emotions.

We believe that progress in the area will be difficult as long as researchers remain committed to the assumption that real intense emotions must be traditional basic emotions, such as fear or anger, for which one can identify relatively straightforward action tendencies such as fight or flight. Progress is more likely to occur if we are prepared to identify emotion episodes where all of the components shown in Table 16.1 are in fact synchronized *without* there being a concrete action tendency or a traditional, readily accessible verbal label. In order to study these phenomena, we need to free ourselves from the tendency of wanting to assign traditional categorial labels to emotion processes. Take, as an interesting example, the feelings that are rather vaguely described as 'being touched' or 'being moved' for which there is not even a substantive in some languages such as English and French (but there is *Rührung* in German). Examples are the tears shed during sentimental movies in the cinema or the flash of warmth experienced when hearing about a good deed. It would seem that such experiences are what music often produces. One is 'moved' by music (accompanied by symptoms such as moist eyes, chills, thrills, or goose-

flesh) without there being a readily accessible label to describe what one is feeling or why. Of course, this does not mean that the lack of suitable verbal labels (or the inability of normal people to discriminate different states) absolves scientists from making the necessary differentiation, falling back to simple arousal models. On the contrary, the task is to develop both suitable non-verbal and verbal indicators that allow one to identify music-induced states that may not correspond to the traditional categories described in emotion textbooks. Finally, and taking this reflection a step further, we have also to be open to the possibility that, in certain instances, responses to music may not fit the states presented in Table 16.1 and may be of a thoroughly non-affective nature.

## 16.3 Conclusion

The suggestions for more principled research efforts in this area, studying several components of emotion, require substantial research investments and competencies in several areas of measurement. It may be unrealistic to expect that such research efforts will emerge in the near future. Yet, even if one is unable to launch such a sophisticated research programme, one still needs to define the theoretical suppositions in greater detail if the work in this area is to achieve a status that permits systematic comparison, replication, and eventually accumulation of findings. Following the specification of the theoretical model to be tested, the appropriate measures would need to be obtained, particularly in the non-verbal domain. This chapter provides an entire spectrum of possibilities for multimodal measurement approaches, encouraging research that does not only rely on standard self-report inventories with their well-known limitations (e.g. Nisbett & Wilson 1977; Scherer & Ceschi 2000). Thus, in future research, the criterion for successful induction should be multimodal, for example verbal report complemented by non-verbal measures such as physiological responses and expressive behaviour. Some of these measures are particularly well suited for this purpose since they provide a continuous, variable signal and are in large parts resistant to conscious regulation.

What are the requirements for measurement for proponents of different theoretical approaches, as outlined above? Researchers adopting dimensional theories would need to specify the expected values for valence and activity for the emotions to be studied and include physiological and expressive indicators in addition to verbal report measurement. Proponents of discrete emotion theories would need to clearly specify the expected emotion-specific response patterns for the emotions studied and include all pertinent variables in their measurement operations. In addition, with respect to the statistical analysis it would seem necessary to use pattern-matching procedures (e.g. profile analyses) rather than simple comparisons for differences (e.g. univariate *F*-tests) between various emotions. Finally, appraisal theorists, in addition to sharpening the predictions for the outcomes to be expected for different appraisal results and the measurement of the respective variables (Johnstone *et al.*, 2001; Pecchinenda 2001), need to indicate how appraisals can be determined in the case of music. Obviously, this requires a much greater research effort than obtaining verbal reports, and one might claim that the suggestions made in this chapter are unrealistic. Yet, similar approaches can already be found in several research domains in the affective sciences and we believe that this tendency will continue.

In brief, we have suggested that by linking research on the emotional expressiveness of music more directly to the progress in the affective sciences generally (see Davidson *et al.*, in press), it may be possible to design empirical studies that are likely to yield results that can be interpreted against a background of theoretical predictions as compared with multiplying studies that have been designed on an *ad hoc* basis and that are difficult to integrate into a coherent body of literature.[1]

## References


Albersnagel, F. A. (1988). Velten and musical mood induction procedures: A comparison with accessibility of thought association. *Behaviour Research and Theory*, 26, 79–96.

Balch, W., Myers, D. M., & Papotto, C. (1999). Dimensions of mood in mood-dependent memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 70–83.

Bartlett, D. L. (1996). Physiological responses to music and sound stimuli. In *Handbook of music psychology*, (2nd edn) (ed. D. A. Hodges), pp. 343–85. San Antonio, TX: IMR.

Behne, K-E. (1997) Musikpräferenzen und Musikgeschmack [Music preferences and musical taste]. In *Musikpsychologie: Ein Handbuch*, (ed. H. Bruhn, R. Oerter, & H. Rösing), pp. 339–53. Reinbek bei Hamburg: Rowohlt.

Blaney, P. H. (1986). Affect and memory: A review. *Psychological Bulletin*, **99**, 229–46.

Boiten, F. A., Frijda, N. H., & Wientjes, C. J. E. (1994). Emotions and respiratory patterns: Review and critical analysis. *International Journal of Psychophysiology*, 17, 103–28.

Bond, A. & Lader, M. (1974). The use of analogue scales in rating subjective feeling. *British Journal of Medical Psychology*, 47, 211–18.

Borchgrevink, H. M. (1975). Musikalske akkod-prefereanser hos mennesket belyst ved dyreforsok [Musical chord preferences in humans as demonstrated through animal experiments]. *Tidskrift for den Norske Laegefoerning*, **95**, 356–8.

Bouhuys, A. L., Bloem, G. M., & Groothuis, T. G. (1995). Induction of depressed and elated mood by music influences the perception of facial emotional expressions in healthy subjects. *Journal of Affective Disorders*, **33**, 215–26.

Brown, J. D. & Mankowski, T. (1993). Self-esteem, mood, and self-evaluation: Changes in mood and the way you see you. *Journal of Personality and Social Psychology*, **64**, 421–30.

Bruner, G. C. (1990). Music, mood, and marketing. *Journal of Marketing*, **54**, 94–104.

Budd, M. (1985). *Music and the emotions. The philosophical theories.* London: Routledge.

Byers, P. (1976). Biological rhythms as information channels in interpersonal communication behavior. In *Perspectives in ethology*, (Vol. 2) (ed. P. P. G. Bateson & P. H. Klopfer) pp. 135–64. New York: Plenum Press.


[1] The preparation of this chapter was facilitated by Grant No. 1114-049680.96 entitled 'Introduction of emotional states through music' by the Swiss National Research fund to Marcel Zentner and Klaus Scherer. The theoretical suggestions linked to perception and production rules were first presented by Klaus Scherer in the context of a keynote speech 'Music and emotional meaning: Perception and production rules' at the Sixth Conference of the International Society for Music Perception and Cognition, 5–10 August 2000, at Keele University, UK. The authors acknowledge important suggestions by a group of anonymous reviewers.

Cacioppo, J. T., Klein, D. J., Berntson, G. C., & Hatfield, E. (1993). The psychophysiology of emotion. In *Handbook of emotions*, (ed. M. Lewis & J. M. Haviland), pp. 119–42. New York: Guilford Press.

Cantor, J. R. & Zillmann, D. (1973). The effect of affective state and emotional arousal on music appreciation. *Journal of General Psychology*, 89, 97–108.

Clynes, M. (1977). *Sentics: The touch of emotions.* New York: Doubleday.

Conifer Classics (1997). *Complete Serenity.* CD No. 51309.

Contrada, R. J., Hilton, W. F., & Glass, D. C. (1991). Effects of emotional imagery on physiological and facial responses in Type A and Type B individuals. *Journal of Psychosomatic Research*, 35, 391–7.

Dalton, K. M. (1998). Relationships between anterior cerebral asymmetry, cardiovascular reactivity, and anger—expression style during re-lived emotion and coping tasks. *Dissertation Abstracts International: Section B: the Sciences and Engineering*, 59, 1406.

Davidson, R., Goldsmith, H., & Scherer, K. R. (ed.) (in press). *Handbook of affective sciences.* New York: Oxford University Press.

Davies, S. (1994). *Musical meaning and expression.* Ithaca, NY: Cornell University Press.

DeVries, B. (1991). Assessment of the affective response to music with Clynes' sentograph. *Psychology of Music*, 19, 46–64.

Dimberg, U. (1988). Facial expressions and emotional reactions: A psychobiological analysis of human social behaviour. In *Social psychophysiology and emotion: Theory and clinical applications*, (ed. L. W. Hugh), pp. 131–50. Chichester, UK: Wiley.

Dowling, W. J. & Harwood, D. L. (1986). *Music cognition.* San Diego: Academic Press.

Duffy, E. (1941). An explanation of 'emotional' phenomena without the use of the concept 'emotion'. *Journal of General Psychology*, 25, 283–93.

Eisenberg, N. & Strayer, J. (ed.) (1987). *Empathy and its development.* Cambridge, UK: Cambridge University Press.

Ekman, P. (1984). Expression and the nature of emotion. In *Approaches to emotion* (ed. K. R. Scherer & P. Ekman), pp. 319–44. Hillsdale, NJ: Erlbaum.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6, 169–200.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, 221, 1208–10.

Ellsworth, P. C. (1994). Levels of thought and levels of emotion. In *The nature of emotion: Fundamental questions*, (ed. P. Ekman & R. J. Davidson), pp. 192–6. New York: Oxford University Press.

Etcoff, N. (1999). *Survival of the prettiest. The science of beauty.* New York: Anchor Books.

Farnsworth, P. R. (1969). *The social psychology of music*, (2nd edn). Ames, IA: Iowa State University Press.

Forgas, J. P. (ed.) (1991). *Emotion and social judgments.* Oxford, UK: Pergamon Press.

Forgas, J. P. (in press). Affective infuences on attitudes and judgments. In *Handbook of affective sciences*, (ed. R. J. Davidson, H. Goldsmith, & K. R. Scherer). New York: Oxford University Press.

Fraisse, P., Oleron, G., & Paillard, J. (1953). Les effets dynamogéniques de la musique [Activation effects in music]. *L'Année Psychologique*, 53, 1–34.

Fried, R. & Berkowitz, L. (1979). Music that charms . . . and can influence helpfulness. *Journal of Applied Social Psychology*, 9, 199–208.

Frijda, N. H. (1986). *The emotions.* Cambridge, UK: Cambridge University Press.

**Frijda, N. H., Mesquita, B.,** Sonnemans, J., & Van Goozen, S. (1991). The duration of affective phenomena or: Emotions, sentiments and passions. In *International review of studies on emotion*, (Vol. 1) (ed. K. T. Strongman), pp. 187–225. Chichester, UK: Wiley.

**Gabrielsson, A.** (1998). The study of emotion in music psychology. Paper presented at the Geneva Emotion Week, May 1998, Geneva, Switzerland.

**Goffman, E.** (1959). *The presentation of self in everyday life.* Garden City, NY: Doubleday Anchor.

**Goffman, E.** (1971). *Relations in public: Microstudies in the public order.* New York: Basic Books.

**Gorman, B. S. & Crain, W. C.** (1974). Decoding of 'sentograms'. *Perceptual and Motor Skills*, 39, 784–6.

**Gurney, E.** (1880). *The power of sound.* London: Smith.

**Hanslick, E.** (1854). *Vom musikalisch Schönen.* [Of beauty in music.] Leipzig: Weigel.

**Hargreaves, D. & North, A.** (1997). *The social psychology of music.* Oxford: Oxford University Press.

**Harrer, G. & Harrer, H.** (1977). Music, emotion, and autonomic function. In *Music and the brain. Studies in the neurology of music*, (ed. M. Critchley & R. A. Henson) pp. 202–16. London: William Heinemann Medical Books.

**Helmholtz, H. L. F.** (1954). *On the sensations of tone as a physiological basis for the theory of music*, (trans. A. Ellis) (2nd English edn). New York: Dover (originally published 1863).

**Hevner, K.** (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology*, **48**, 248–68.

**Hochschild, A. R.** (1983). *The managed heart: The commercialization of human feeling.* Berkeley, CA: University of California Press.

**Izard, C. E.** (1971). *The face of emotion.* New York: Appleton-Century-Crofts.

**Izard, C. E.** (1990). Facial expression and the regulation of emotions. *Journal of Personality and Social Psychology*, **58**, 487–98.

**Izard, C. E.**(1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin*, 115, 288–99.

**Johnstone, T., van Reekum, C. M., & Scherer, K. R.** (2001). Vocal correlates of appraisal processes. In *Appraisal processes in emotion: Theory, methods, research*, (ed. K. R. Scherer, A. Schorr, & T. Johnstone) pp. 271–84). New York: Oxford University Press.

**Jürgens, U.** (1988). Central control of monkey calls. In *Primate vocal communication*, (ed. D. Todt, P. Goedeking, & D. Symmes), pp. 162–70. Berlin, Germany: Springer.

**Kainz, F.** (1962). *Psychologie der Sprache. I. Band. Grundlagen der allgemeinen Sprachpsychologie.* [Psychology of language. Vol. I. The bases of the general psychology of language.](3rd edn). Stuttgart: Enke.

**Kappas, A., Hess, U., & Scherer, K. R.** (1991). Voice and emotion. In *Fundamentals of nonverbal behavior*, (ed. R. S. Feldman & B. Rimé), pp. 200–38. Cambridge, UK: Cambridge University Press.

**Kenealy, P.** (1988). Validation of a music induction procedure: Some preliminary findings. *Cognition and Emotion*, 2, 41–8.

**Kenyon, N.** (ed.) (1991). *Authenticity and early music.* Oxford, UK: Oxford University Press.

**Kivy, P.** (1989). *Sound sentiment: An essay on the musical emotions.* Philadelphia, PA: Temple University Press.

**Kneutgen, J.** (1964). Beobachtungen über die Anpassung von Verhaltensweisen an gleichförmige akustische Reize [Observations concerning the adaptation of behavior patterns to repetitive acoustical stimuli]. *Zeitschrift für Tierpsychologie*, 21, 764–79.

**Kneutgen, J.** (1970). Eine Musikform und ihre biologische Funktion. Ueber die Wirkungsweise der Wiegenlieder [The biological function of a category of music: On the effect of lullabies]. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 17, 245–65.

**Krumhansl, C. L.** (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology*, 51, 336–52.

**Lang, P. J.** (1979). A bio-informational theory of emotional imagery. *Psychophysiology*, 16, 495–512.

**Lang, P., Bradley, M. M., & Cuthbert, B. N.** (1990). Emotion, attention, and the startle reflex. *Psychological Review*, 97, 377–95.

**Lang, P. J., Kozal, M. J., Miller, G. A., Levin, D. N., & MacLean, A.** (1980). Emotional imagery: Conceptual structure and pattern of somato-visceral response. *Psychophysiology*, 17, 179–92.

**LeDoux, J. E.** (1992). Emotion as memory: Anatomical systems underlying indelible neural traces. In *Handbook of emotion and memory: Theory and research*, (ed. S. Christianson), pp. 269–88. Hillsdale, NJ: Erlbaum.

**LeDoux, J. E.** (1996). *The emotional brain.* New York: Simon & Schuster.

**Levenson, R. W.** (1992). Autonomic nervous system differences among emotions. *Psychological Science*, 3, 23–7.

**Leventhal, H. & Scherer, K. R.** (1987). The relationship of emotion to cognition: A functional approach to a semantic controversy. *Cognition and Emotion*, 1, 3–28.

**Lippman, E. A.** (1953). Symbolism in music. *Musical Quarterly*, 39, 554–75.

**Lipps, Th.** (1909). *Leitfaden der Psychologie* [Primer of psychology]. Leipzig, Germany: Engelmann.

**Lundqvist, L. G., Carlsson, F., & Hilmersson, P.** (2000). Facial electromyography, autonomic activity, and emotional experience to happy and sad music. Paper presented at the 27th International Congress of Psychology, Stockholm, Sweden, 23–28 July 2000.

**Lynxwiler, J. & Gay, D.** (2000). Moral boundaries and deviant music: Public attitudes toward heavy metal and rap. *Deviant Behavior*, 21, 63–85.

**McIntosh, D. N.** (1996). Facial feedback hypotheses: Evidence, implications, and directions. *Motivation and Emotion*, 20, 121–47.

**Martin, M. A. & Metha, A.** (1997). Recall of early childhood memories through musical mood induction. *Arts in Psychotherapy*, 25, 447–54.

**Maslach, C.** (1979). Negative emotional biasing of unexplained arousal. *Journal of Personality and Social Psychology*, 37, 933–69.

**Mattheson, J.** (1739). *Der volkommene Capellmeister* [The perfect musical director]. Hamburg: Herold.

**Mayer, J. D. & Gaschke, Y. N.** (1988). The experience and meta-experience of mood. *Journal of Personality and Social Psychology*, 55, 102–11.

**Meyer, L. B.** (1956). *Emotion and meaning in music.* Chicago: University of Chicago Press.

**Nettlebeck, T., Henderson, C., & Willson, R.** (1989). Communicating emotion through sound: An evaluation of Clynes' theory of sentics. *Australian Journal of Psychology*, 41, 17–24.

**Nisbett, R. E. & Wilson, T. D.** (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231–59.

**Nyklicek, I., Thayer, J. F., & van Doornen, L. J.** (1997). Cardiorespiratory differentiation of musically-induced emotions. *Journal of Psychophysiology*, 11, 304–21.

**Öhman, A.** (1988). Preattentive processes in the generation of emotions. In *Cognitive perspectives on emotion and motivation*, (ed. V. Hamilton, G. H. Bower, & N. H. Frijda), pp. 127–44. Dordrecht, the Netherlands: Kluwer.

Ornstein, R. & Sobel, D. (1989). Coming to our senses. *Advances*, 6, 49–56.

Pecchinenda, A. (2001). The psychophysiology of appraisals. In *Appraisal processes in emotion: Theory, methods, research*, (ed. K. R. Scherer, A. Schorr, & T. Johnstone) pp. 301–15. New York: Oxford University Press.

Pignatiello, M., Camp, C. J., Elder, S. T., & Rasar, L. A. (1989). A psychophysiological comparison of the Velten and musical mood induction techniques. *Journal of Music Therapy*, 26, 140–54.

Pitman, R. K., Orr, S. P., Shalev, A. Y., Metzger, L. J., & Mellman, T. A. (1999). Psychophysiological alterations in post-traumatic stress disorder. *Seminars in Clinical Neuropsychiatry*, 4, 234–41.

Pratt, C. (1931). *The meaning of music.* New York: McGraw-Hill.

Pujol, K. K. (1994). The effect of vibrotactile stimulation, instrumentation, and precomposed melodies on physiological and behavioral responses of profoundly retarded children and adults. *Journal of Music Therapy*, 31, 186–205.

Rigg, M. G. (1964). The mood effects of music: A comparison of data from earlier investigations. *Journal of Psychology*, 58, 427–38.

Rochat, P. & Striano, T. (1999). Emerging self-exploration by 2-month-old infants. *Developmental Science*, 2, 206–18.

Rosar, W. H. (1994). Film music and Heinz Werner's theory of physiognomic perception. *Psychomusicology*, 13, 154–65.

Schellenberg, E. G. & Trehub, S. E. (1996). Natural musical intervals. Evidence from infant listeners. *Psychological Science*, 7, 272–7.

Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In *Approaches to emotion*, (ed. K. R. Scherer & P. Ekman), pp. 293–317. Hillsdale, NJ: Erlbaum.

Scherer, K. R. (1985). Vocal affect signalling: A comparative approach. In *Advances in the study of behavior*, (Vol. 15) (ed. J. Rosenblatt, C. Beer, M.-C. Busnel, & P. J. B. Slater), pp. 189–244. New York: Academic Press.

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, 99, 143–65.

Scherer, K. R. (1988). On the symbolic functions of vocal affect expression. *Journal of Language and Social Psychology*, 7, 79–100.

Scherer, K. R. (1991). Emotion expression in speech and music. In *Music, language, speech, and brain*, (ed. J. Sundberg, L. Nord, & R. Carlson), pp. 146–56. London: Macmillan.

Scherer, K. R. (1992). What does facial expression express? In *International review of studies on emotion*, (Vol. 2) (ed. K. Strongman), pp. 139–65. Chichester, UK: Wiley.

Scherer, K. R. (1993). Neuroscience projections to current debates in emotion psychology. *Cognition and Emotion*, 7, 1–41.

Scherer, K. R. (1994). Affect bursts. In *Emotions: Essays on emotion theory*, (ed. S. van Goozen, N. E. van de Poll, & J. A. Sergeant), pp. 161–96. Hillsdale, NJ: Erlbaum.

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice*, 9, 235–48.

Scherer, K. R. (1998). Emotionsprozesse im Medienkontext: Forschungsillustrationen und Zukunftsperspektiven [Emotion processes in the context of the media: Illustrative research and perspectives for the future]. *Medienpsychologie*, 10, 276–93.

Scherer, K. R. (1999). Appraisal theories. In *Handbook of cognition and emotion*, (ed. T. Dalgleish & M. Power), pp. 637–63. Chichester, UK: Wiley.

Scherer, K. R. (2000*a*). Psychological models of emotion. In *The neuropsychology of emotion*, (ed. J. Borod), pp. 137–62. New York: Oxford University Press.

Scherer, K. R. (2000*b*). Emotional expression: A royal road for the study of behavior control. In *Control of human behavior, mental processes, and awareness*, (ed. A. Grob & W. Perrig), pp. 227–44. Hillsdale, NJ: Erlbaum.

Scherer, K. R. (2000*c*). Emotions as episodes of subsystem synchronization driven by nonlinear appraisal processes. In *Emotion, development, and self-organization: Dynamic systems approaches to emotional development*, (ed. M. Lewis & I. Granic), pp. 70–99. Cambridge, UK: Cambridge University Press.

Scherer, K. R. (2000*d*). Music and emotional meaning: Perception and production rules. Paper presented at the Sixth Conference of the International Society for Music Perception and Cognition, 5–10 August 2000, Keele University, UK.

Scherer, K. R. (2001). Appraisal considered as a process of multi-level sequential checking. In *Appraisal processes in emotion: Theory, methods, research*, (ed. K. R. Scherer, A. Schorr, & T. Johnstone) pp. 92–120. New York: Oxford University Press.

Scherer, K. R. & Ceschi, G. (2000). Criteria for emotion recognition from verbal and nonverbal expression: Studying baggage loss in the airport. *Personality and Social Psychology Bulletin*, 26, 327–39.

Scherer, K. R. & Oshinsky, J. (1977). Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion*, 1, 331–46.

Scherer, K. R., Schorr, A., & Johnstone, T. (ed.) (2001). *Appraisal processes in emotion: Theory, methods, research*. New York: Oxford University Press.

Scherer, K. R. & Wallbott, H. G. (1994). Evidence for universality and cultural variation of differential emotion response patterning. *Journal of Personality and Social Psychology*, 66, 310–28.

Scherer, K. R., Zentner, M. R., & Schacht, A. K. (in press). Emotional states generated by music: A conceptual and empirical analysis. *Musicae Scientiae*.

Sloboda, J. A. (1992). Empirical studies of emotional response to music. In *Cognitive bases of musical communication*, (ed. M. R. Jones & S. Holleran), pp. 33–46. Washington, DC: American Psychological Association.

Sony (1995). *Rendezvous der Sinne. Die sinnlichste Classic-Collection der Musikgeschichte.* [Rendezvous of the senses. The most sensual CD collection in musical history.] CD No. 60464.

Stemmler, G. (1996). Psychophysiologie der Emotionen [Psychophysiology of emotions]. *Zeitschrift für Psychosomatische Medizin und Psychoanalyse*, 42, 235–60.

Stemmler, G., Heldmann, M., Pauls, C., & Scherer, T. (2001). Constraints for emotion specificity in fear and anger: The context counts. *Psychophysiology*, 38, 275–91.

Tarrant, M. A., Manfredo, M. J., & Driver, B. L. (1994). Recollections of outdoor recreation experiences: A psychophysiological perspective. *Journal of Leisure Research*, 26, 357–71.

Teasdale, J. (1999). Multi-level theories of cognition-emotion relations. In *Handbook of cognition and emotion*, (ed. T. Dalgleish & M. Power), pp. 665–81. Chichester, UK: Wiley.

Todd, N. P. (1992). The dynamics of dynamics: A model of musical expression. *Journal of Acoustical Society of America*, 91, 3540–50.

Tomkins, S. S. (1962). *Affect, imagery, consciousness. Vol. 1: The positive affects.* New York: Springer.

Tomkins, S. S. (1963). *Affect, imagery, consciousness. Vol. 2: The negative affects.* New York: Springer.

Tomkins, S. S. (1984). Affect theory. In *Approaches to emotion*, (ed. K. R. Scherer & P. Ekman), pp. 163–96. Hillsdale, NJ: Erlbaum.

**Trehub, S. E. & Schellenberg, G. E.** (1995). Music: Its relevance to infants. In *Annals of child development*, (Vol. 11) (ed. R. Vasta), pp. 1–24. London: Jessica Kingsley.

**Trussoni, S. J., O'Malley, A., & Barton, A.** (1988). Human emotion communication by touch: A modified replication of an experiment by Manfred Clynes. *Perceptual and Motor Skills*, 66, 419–24.

**Vaitl, D., Vehrs, W., & Sternagel, S.** (1993). Prompts—leitmotif—emotion: Play it again, Richard Wagner. In *The structure of emotion: Psychophysiological, cognitive, and clinical aspects*, (ed. N. Birbaumer & A. Öhman), pp. 169–89. Seattle: Hogrefe & Huber.

**Van der Kolk, B. A.** (1997). The psychobiology of posttraumatic stress disorder. *Journal of Clinical Psychiatry*, **58** (Suppl. 9), 16–24.

**Van Reekum, C. M. & Scherer, K. R.** (1998). Levels of processing for emotion-antecedent appraisal. In *Cognitive science perspectives on personality and emotion*, (ed. G. Matthews), pp. 259–300. Amsterdam: Elsevier.

**Wallin, N. L., Brown, S., & Merker, B.** (ed.) (2000). *The origins of music.* Cambridge, MA: MIT Press.

**Wedin, L.** (1972). A multidimensional study of perceptual–emotional qualities in music. *Scandinavian Journal of Pschology*, 13, 1–17.

**Westermann, R., Spies, K., Stahl, G., & Hesse, F. W.** (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology*, 26, 557–80.

**Witvliet, C. V. & Vrana, S. R.** (1996). The emotional impact of instrumental music on affect ratings, facial EMG, autonomic measures, and the startle reflex: Effects of valence and arousal. *Psychophysiology Supplement*, **91**.

**Witvliet, C. V., Vrana, S. R., & Webb-Talmadge, N.** (1998). In the mood: Emotion and facial expressions during and after instrumental music, and during an emotional inhibition task. *Psychophysiology Supplement*, **88**.

**Wundt, W.** (1900). *Völkerpsychologie. Eine Untersuchung der Entwicklungsgesetze von Sprache, Mythos und Sitte* (Band I. Die Sprache). [Psychology of cultures: A study of the developmental laws of language, myth, and customs. Vol. I. Language.] Leipzig: Kröner.

**Wundt, W.** (1974). *Grundzüge der physiologischen Psychologie* [Fundamentals of physiological psychology], (5th edn). Leipzig: Engelmann (originally published 1874).

**Zajonc, R. B.** (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 2, 151–76.

**Zajonc, R. B. & Markus, H.** (1984). Affect and cognition: The hard interface. In *Emotions, cognition, and behavior*, (ed. C. E. Izard, J. Kagan, & R. B. Zajonc), pp. 73–102. Cambridge, UK: Cambridge University Press.

**Zentner, M. R. & Kagan, J.** (1996). Perception of music by infants. *Nature*, 383, 29.

**Zentner, M. R. & Kagan, J.** (1998). Infants' perception of consonance and dissonance in music. *Infant Behavior and Development*, 21, 483–92.

**Zentner, M. R., Meylan, S., & Scherer, K. R.** (2000). Exploring 'musical emotions' across five genres of music. Paper presented at the Sixth International Conference of the Society for Music Perception and Cognition (ICMPC), 5–10 August 2000, Keele, UK.

**Zuckerman, M. & Lubin, B.** (1985). *Manual for the Multiple Affect Adjective Check List*, (2nd edn). San Diego, CA: Educational and Industrial Testing Service.

CHAPTER 17

# CONTINUOUS MEASUREMENT OF SELF-REPORT EMOTIONAL RESPONSE TO MUSIC

EMERY SCHUBERT

By definition music requires time to happen. Many styles of music also express or evoke different emotions as time unfolds. Logically, it follows that to understand the relationship between music and emotion, psychologists need to investigate their time-varying relationships. It is therefore surprising to find that the study of *continuous* self-reported emotion in music is less than 20 years old. Continuous response tools measure self-reported emotional responses *during* the listening process and help researchers to better understand the moment-to-moment fluctuations in responses that traditional 'asynchronous' approaches cannot. The approach means that the music does not need to be broken into small portions to enable lucid response, nor does the listener have to compress their response by giving an overall impression at the end of the excerpt. A more realistic listening experience is possible. This 'realism' (Hargreaves 1986, p. 105) contributes to the ecological validity of experimental design at the expense of experimental control.

Given that the methodology is relatively new, there is a need to examine the literature and to provide a framework for future directions. Gregory (1995) and Schmidt (1996) review one kind of continuous response device, and the reviews encompass emotional response as one aspect of a broader scope. The present chapter examines a wider range of continuous measure devices, but restricts the discussion to those studies that investigate self-report emotional response. In this chapter, I will define 'continuous', describe the response formats used for collecting continuous responses, discuss the ways in which researchers have applied continuous response methodology, and then provide an evaluation of the major issues.

## 17.1 Definition of continuous response

The term 'continuous' is arguably used incorrectly in the literature. Continuous means without break, inferring that responses are measured without interruption. Modern techniques of measuring response require some form of sampling because they are associated with digital computers. At best, the frequency of sampling may be high, but is

still interrupted. Even a rate of tens of thousands of samples per second is still not, strictly speaking, a continuous measure. Instead, the more correct term is *continual*, or perhaps *multiple equispaced*. It is apparent that the 'erroneous' terminology stems from the kymographic devices used in physiological response studies since the late nineteenth century (Dogiel 1880 cited in Radocy & Boyle 1988, p. 205; Weld 1912), which drew line plots of physiological activities as a function of time. These analogue devices were indeed continuous. And, to be sure, high sampling rates associated with modern digital equipment can give the impression of a continuous measure.

However, the patterns and latency of emotional response makes the need for strict continuity or even high sampling rates of measurement superfluous. For example, a lexical response task alone will take in the order of 1 second (Greene & Royer 1994), and the response latency in the tracking of a musical feature, such as loudness, is 1–1.5 seconds (Geringer 1995). By the same token, the term *continuous* has been used consistently in the literature, and consequently its use will be retained to mean continuous *or* continual.

## 17.2 Response formats

### 17.2.1 Open-ended format

Three of the traditional self-report formats found in the asynchronous emotion–music literature are also found in continuous response research: open-ended (or free-verbal) format, checklist, and rating scale (Gabrielsson & Juslin, in press; Sloboda 1992; see also Gabrielsson & Lindström, this volume). Open-ended approaches (e.g. one part of the study by Watson 1942) are rarely found in the literature because of the obvious problem of the response task *during* listening interfering with the listening process.

### 17.2.2 Checklist

Checklist formats have also been used infrequently (Goins 1998; Hair 1995; Hevner 1936; Mull 1949; Namba *et al.* 1991). An innovative and pioneering example is Hevner's (1936) study, which was the first to use what could be called a self-report continuous response approach. Hevner addressed the problem of collecting an overall response for a piece (that may have actually expressed several different emotions) by asking the listener to make judgments at the end of each section of music as the music unfolded. For example, if a piece of music had several sections, the participant would place the number '1' next to the selected term or terms which best described the emotion in the first section of music, and when the next section of the music began (signalled by the experimenter), a '2' would be placed next to the term characterizing that section. This process continued until all sections were heard. The pieces Hevner reported were divided into three sections and, together, these sections usually made up an entire movement.

In one experiment, Hevner collected responses to *Reflections on the Water* by Debussy. She discovered that responses to the piece were widely distributed across the eight clusters (Fig. 17.1a). The experiment was conducted again using the continuous procedure. In the first section of the piece, the responses were made predominantly in clusters 5 and 7 ('humorous' and 'exhilarated'), whereas in the second section, responses shifted

to cluster 4 ('lyrical'). In the final section, clusters 4 and 5 ('lyrical' and 'humorous') were chosen most frequently (Fig. 17.1b). These results demonstrated that different sections of music evoked different, though internally consistent responses.

Despite this and later studies, the categorical approach of the checklist does not easily lend itself to continuous methodology. Categorical responses require the participant to select one or several answers from a range of answers. Unless lengthy training time is administered (e.g. 30 minutes per participant in the Namba *et al.* (1991) study), a large range of choices will require the participant to constantly shift attention between the various response options, in addition to concentrating on the music. In a continuous response task, this may either interfere with the listening experience, or lead to the omission or guessing of some responses so as to minimize interference. This might explain why there are relatively few checklist studies that employ the continuous response approach at high sampling rates. Hevner's approach of gathering responses

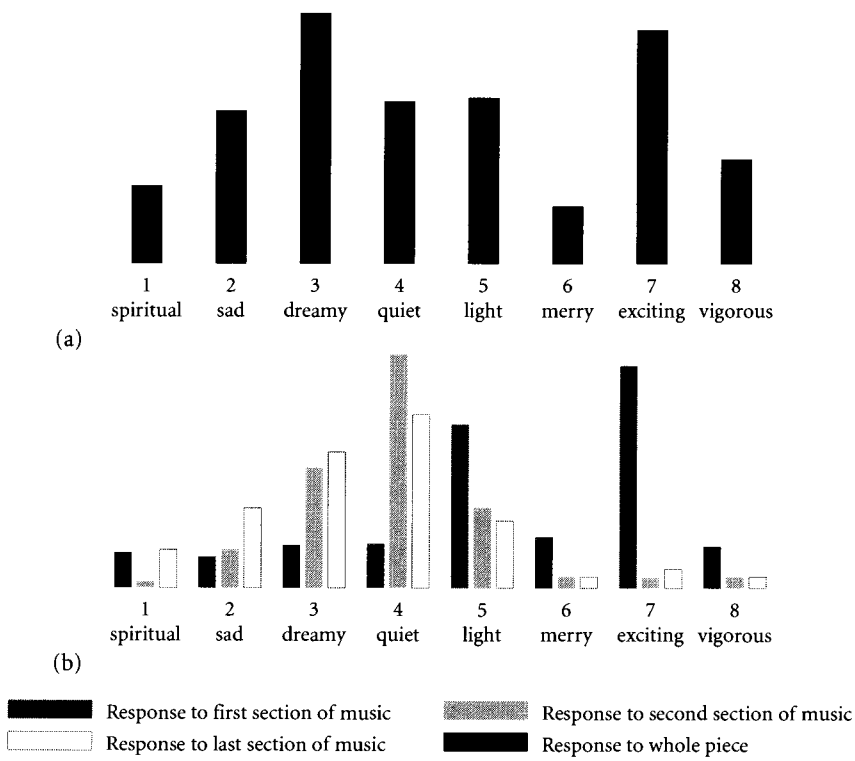



Fig 17.1 Hevner's adjective checklist responses to Debussy's *Reflections on the Water* for the whole piece and in sections (adapted from Hevner 1936, pp. 251 and 254). The numbers correspond to the cluster from the Hevner adjective circle with a representitive word from each cluster number appearing under each number (the complete adjective circle is presented in Gabrielsson & Lindström, this volume). (a) The relative frequencies of each adjective circle cluster, and (b) the responses to the same piece in a different experiment with separate responses recorded for the three sections of the music.

once only per section of music mitigates the problem, but it does not allow for the investigation of changing shades of meaning within a section of the music.

### 17.2.3 Rating scale

A useful subdivision of continuous response literature employing rating scale measures is according to the number of simultaneous dimensions measured. Most studies use one dimension, and more recently two scales have also been used simultaneously. One-dimensional studies have used a variety of labels representing either one specific aspect or dimension of emotion (such as happiness or tension) to more general labels (such as aesthetic experience or 'when the music causes something to happen to you'). A selection of studies using one-dimensional scales is shown in Table 17.1.

Table 17.1 Labels used in one-dimensional measures of continuous self-report emotional response, with scale labels and scale poles (or task) indicated

| Self-report label | Poles/task | Examples |
|---|---|---|
| Aesthetic experience | Negative-positive | Capperella-Sheldon (1992) |
| | Negative-positive | Madsen *et al.* (1993*a*) |
| Chills | Raise hand if present | Panksepp (1995) |
| Emotion | Minimum-maximum | Krumhansl (1998) |
| Emotional reaction | Positive-negative | D. Gregory (1998) |
| Emotional response | No emotional response-emotional response | Adams (1994) |
| | Negative-positive | Lychner (1998) |
| Emotionality ('the capacity of the performance at that moment to suggest, communicate, or evoke musically relevant emotion', p. 631) | High-low | Sloboda *et al.* (1997) |
| Fear | Amount of fear | Krumhansl (1997)) |
| Felt emotion | Negative-positive | Lychner (1998) |
| Happiness | Amount of happiness | Krumhansl (1997) |
| Humour | Raise hand if present | Mull (1949) |
| Intensity of emotions | Low-high | A. H. Gregory (1998) |
| Perceived artistic tension | Decrease-increase | Frego (1999) |
| Sadness | Amount of sadness | Krumhansl (1997) |
| Tension | Less-more | Byrnes (1996) |
| | Less-more | Frederickson (1995, 1997, 1999) |
| | Amount of tension | Krumhansl (1997) |
| | Less-more | Madsen & Fredrickson (1993) |
| | Amount of physical pressure on a pair of tension tongs | Nielsen (1983) |
| | Minimum tension-maximum tension | Krumhansl (1996) |
| Thrill | Raise fingers for strength and location of thrill | Goldstein (1980) |
| When the music causes something to happen to you | Frequency of responses | Waterman (1996) |

An important criticism cast against one-dimensional rating scales is that they cannot capture the complexities of emotion satisfactorily (Schmidt 1996). One solution to this problem is to track more than one dimension simultaneously. Since 1996 there has been a trend toward the measurement of two dimensions simultaneously. An early example of what could be referred to as a two-dimensional, continuous response instrument, namely the *sentograph*, was developed by Clynes (Clynes 1977, 1980; Clynes & Nettheim 1982; see also de Vries 1991; Gabrielsson & Lindström 1995). However, because it was used to measure 'inner' emotional experience and 'essentic forms' it is, strictly speaking, not a self-report instrument. Three self-report continuous, two-dimensional 'emotion spaces' (2DES) were developed independently and almost simultaneously by Schubert (1996), Tyler (1996), and Madsen (1997). For each of these instruments, the listener makes responses by moving about a two-dimensional space (computer interface), with each axis (or scale) of the space representing a separate dimension of emotion.

There are various problems associated with rating scale measures. The following sections discuss the kinds of questions asked by continuous response researchers, followed by some of the issues and concerns relevant to continuous response methodology.

## 17.3 When are continuous measures used?

Continuous measures of self-report emotion in music have been used for four broad purposes: validation, comparative investigations, stimulus–response investigations, and system–dynamics investigations.

### 17.3.1 Validation

Continuous response measures have been used to validate traditional asynchronous response formats and vice versa. Goins (1998) compared grand mean categorical continuous response digital interface (CRDI—a dial whose manipulations are recorded on a computer in real time) responses with conventional asynchronous responses using a modified version of the Hevner adjective checklist. He noted a strong similarity between the two methods. This supports the external validity of the approach. However, with just eight categories identified (with asynchronous $n = 30$, and continuous $n = 40$), it becomes difficult to identify subtle differences between the measures and to provide a quantitative index. In a study of intensity in music, Brittin and Duke (1997) found that overall, asynchronous responses tended to be higher in magnitude than the averaged continuous response. Schubert (1999*b*) used an adjective checklist *and* an asynchronous 2DES to investigate the validity of the continuous 2DES rating scales. Participants were required to respond to a piece of music in three ways: continuously, then with an overall judgment, and finally by selecting a word from a checklist. While the checklist responses provided good agreement with averaged responses on the continuous 2DES, the overall 2DES gave a result consistent with Brittin and Duke's findings. Specifically, overall responses were more exaggerated than the average of continuous responses. Continuous responses may gravitate to the centre of the scales being used, or the relationship between averaged and overall responses are in some way scaled.

### 17.3.2 Comparative investigations

An example of the use of the continuous response measure for comparative studies can be found in the Madsen *et al.* (1993*b*) study which used the CRDI to compare non-musicians' and musicians' perceived aesthetic experiences. They concluded that aesthetic experiences do not differ between musicians and non-musicians. Similarly, Lychner (1998) found no significant differences between musicians and non-musicians in their continuous ratings of three classical music extracts and a Sousa march according to felt emotion, aesthetic response, and free response. Another example is the study by Sloboda *et al.* (1997) where 'emotionality' in performances of a Chopin prelude perceived by non-pianists, pianists (a subset of the pianists who had provided the performances), and music jurors were investigated. These researchers found that the pianists responded with significantly lower emotionality scores in the first segment of the piece, and significantly higher scores in the last segment of the piece, compared with the other two groups.

### 17.3.3 Stimulus–response investigations

The use of the continuous response device to investigate the music–emotion system (M→E) is hinted at in several studies. Madsen (1996, p. 106) compared data in studies conducted by himself and John Geringer, producing correlations between musical features (melody, dynamics, timbre, rhythm, and 'everything') that were the focus of listeners' attention with aesthetic response. The highest correlation reported was with melody ($r = 0.67$), and the next highest was dynamics ($r = 0.59$), when listening to an excerpt from Puccini's *La Boheme*, and rhythm ($r = 0.69$) when listening to the first movement of Haydn's Symphony No. 104. A. H. Gregory (1998) observed that higher intensity emotions appeared to occur during vocal passages and periods of higher loudness. In a comparative study (degrees of familiarity with the test piece), Fredrickson (1999, p. 50) described a region of judged high musical tension over bars 32–45 of 'Chaconne' from Holst's *First Suite in E♭*:

> There is a *crescendo* to a dynamic peak at measure 41, where a series of running sixteenths begin in the upper woodwinds. The musicians who are unfamiliar with the music (singers and high school pre-test) indicate a higher tension level here that continues through another statement of the melodic material in the trumpets at measure 49, where the melody moves above a heavy (*pesante*) eighth-note section in the low brass. Graphs for musicians who are familiar with the music (wind ensemble and high school post-test) drop off sharply after the *crescendo* and rise only slightly before falling off along with the other groups at the *diminuendo*, which ends the section.

Since Hevner's (1936) early exploration, a small number of studies have used continuous response methodology to focus *specifically* on the issue of the relationship between musical features and emotional response. Nielsen (1983, 1987) and Krumhansl (1996) asserted that there was a relationship between structural features (based on the theories of Lerdahl & Jackendoff (1983) and Narmour (1990)) and tension response. However, graphical examination of the Nielsen (1987) study, and Krumhansl's own observation (1996; p. 415) suggests that loudness matches reasonably well with tension responses; in the case of Nielsen perhaps better than the higher-level structural analysis. Both studies

also identified specific locations where the melodic contour was related to the tension response: for Nielsen, in the development section of the first movement of Haydn's Symphony No. 104, and for Krumhansl, in the first movement Mozart's Piano Sonata in E♭ Major, K. 282, Krumhansl found that Lerdahl's pitch-space model (which provides a quantitative coding of pitch and harmonic dissonances) accounted for a significant proportion of the tension response. In another study, Krumhansl (1998) correlated high-level musical features, referred to as 'topics' (Ratner 1980; see also Cook & Dibben, this volume), with judged emotionality. While correlations were not very high, the 'gavotte' and 'alla zoppa' musical topics produced relatively strong emotional responses for the allegro from Mozart's String Quintet No. 3 in C Major K. 515 ($r < 0.2$ and $r < |-0.3|$, respectively), and the 'brilliant style' topic for Beethoven's String Quartet No. 15 in A Minor, Op. 132 ($r < 0.3$).

Sloboda (1991) constructed a list of musical structural features that produced particular physical emotional responses. For example, a harmonic descending cycle of fifths to tonic, a melodic appoggiaturas, and a melodic or harmonic sequence tended to produce a 'tears' response, and a new or unprepared harmony tended to produce a 'shivers' response. Schubert (1996, 1999*b*) found that the best predictor of arousal was loudness. He also proposed that perfect cadence harmonic progression increased valence response (Schubert 1999*a*).

Continuous response methodologies have provided an enriching alternative to the study of music as a causal agent of emotional response. While the methodology is inefficient compared with conventional, asynchronous, single-sample-per-stimulus approaches, it has facilitated an understanding of the processes of the M→E system that are difficult to isolate and may require musical context to have an effect. For example, it would be hard to determine empirically the emotional effect of the appoggiatura or the perfect cadence within an ecologically valid framework (musical context) without continuous response methodology.

### 17.3.4 System–dynamics investigations

The single most important purpose of continuous measures is in better understanding the time-dependent, dynamic nature of the M→E system. The analytical method proposed by Schubert and Dunsmuir (1999) provides an example of how lag structure can be better understood. Before continuous response methodologies were accessible, experimental music psychologists could do little more than speculate about the delay in emotional response after a 'causal' musical event. The findings and issues concerned with the dynamic nature of emotional response to music using continuous response methodology are discussed below under Sections 17.4.5 and 17.4.6.

## 17.4 Evaluation of continuous measure devices

The growing interest in continuous measure approaches to emotion in music necessitates the consideration of several issues pertaining to the design, reliability, and validity of continuous measure instruments. Important issues to emerge are the defining and coding of the stimulus and of the emotional response, synchronizing response and

stimulus, and determining the optimum and adequate response sampling rates. Pertinent issues in examining the relationship between musical features and emotional response are the response latency (or 'lag structure') and the method used to analyse data. These issues will be discussed with reference to the literature described above and with respect to possible future directions.

### 17.4.1 Stimulus coding

To determine how emotion varies in response to music, the musical stimulus must be analysed and coded. One approach is a *post hoc* analysis where once a significant or important emotional response is identified, the musical features or structures thought to be responsible are then ascertained. Sloboda (1991) and Krumhansl (1996) operated in this fashion. This approach requires a high degree of musicological skill on the part of the researcher. A more objective approach is to code musical feature variables before collecting emotional responses, and to use a statistical method to determine which combinations of musical features can explain the temporally nearby changes in emotion. This technique treats the musical stimulus as a multidimensional set of interacting musical and psychoacoustic features or structures. These features can be coded at a high level, for example in terms of their aesthetic association or 'topics' (used by Krumhansl 1998, based on Ratner 1980), or according to some coding strategy deduced from a theory of harmony or pitch dissonance (Krumhansl 1996, based on Lerdahl's theory of tonal pitch space), or grouping structure (Nielsen 1983, after Lerdahl and Jackendoff's theory of generative grammar in music). Alternatively, low-level musical features could be coded (Schubert 1999*a*), such as pitch, loudness, duration, timbre, and harmony.

The kinds and levels of variables coded will be a function of the requirements of the researcher, and the feasibility of the coding scheme. Some musical features, such as pitch and loudness, can be coded as perceptually relevant, objective multilevel variables (e.g. Cabrera 1997, 1998; Hall 1987; Moore *et al.* 1997; Rossing 1990; Schubert 1999a). Other musical features, such as harmony and rhythm, require some form of polychotimization to achieve multilevel coding. For example, Schubert (1999a) coded cadential progressions as a continuous, dichotomous variable having the value 'perfect cadence' or 'no perfect cadence'.

### 17.4.2 Response coding

A theoretical consideration that can help to determine response format and coding strategy is the paradigm of emotion. There are two broad systems of classifying emotions: categories and dimensions. Categorical classification of emotion assumes that emotions carrying different meanings, such as happy and sad, are distinct and independent entities. While Goins (1998), Hevner (1936) and Namba and associates (1991) have used the checklist format to obtain data that are essentially continuous, the categorical approach does not lend itself easily to continuous methodology. Categorical approaches require the participant to select one or several answers from a range of answers. A large range of choices will require the participant to constantly shift attention between the various response options, in addition to the experimental task. In a continuous response task, this may either interfere with the listening experience, or lead

to omission or guessing of some responses so as to minimize interference. Hevner's approach of gathering responses once only per section of music is an improvement to continual checklist evaluation, but it does not allow for the investigation of changing shades of meaning.

The dimensional classification of emotion holds that all emotions are in some way related within an *n*-dimensional *semantic space* (or, more correctly, *emotion space).* For example, the dimensional structure suggests that happy and sad are more or less opposite emotions along the *valence* dimension of emotion (Russell 1989). To distinguish distinct emotions having similar valence, such as sad and angry, a second dimension, *arousal,* may be added: Sad expresses low arousal and angry expresses high arousal (see also Sloboda & Juslin, this volume).

Many researchers have assumed that participants can manage concentrating on one dimension only while listening to music (e.g. Madsen & Fredrickson 1993; Nielsen 1983; Sloboda *et al.* 1997). Because emotion is multidimensional, the dimension to choose has been problematic. My review of the literature supports Schmidt's (1996) criticism that most unidimensional continuous response formats refer to a general, abstract construct such as 'aesthetic experience' or 'emotionality', or, even more vague, 'when the music causes something to happen to you' (see Table 17.1). A reasonably specific term was 'musical tension' (Fredrickson 1999; Krumhansl 1996; Nielson 1983), and even this is problematic—what, exactly, is musical tension? Schmidt is critical of vague instructions and non-specific criteria. However, the complexity of the theoretical concepts in some ways demands that the participant uses some implicit and intuitive judgment criterion.

The validity of choosing a non-specific dimension as a means of tapping into emotion is a major issue. For example, when preparing this chapter I compared the response of the first 37 bars of Rodrigo's *Concierto de Aranjuez* obtained from the Schubert (1999a, p. 510) study with that obtained by Waterman (1996, p. 60). The peak responses in the Waterman study ('when . . . something . . . happen[s] to you') occurred at the same locations as the peaks in arousal response gradient in the Schubert study, namely at bars 23, 31, and 33, but no such similarity was found between the valence response gradient and the Waterman data. This prompts me to propose that the more general instruction may actually be tapping into the arousal dimension as a default response, rather than some general, idealized emotional/aesthetic response. I tentatively extend this proposal to include 'tension' because of the large amount of literature which supports a correlation between tension and arousal (e.g. Berlyne 1973; Krumhansl & Schenck 1997).

Another solution attempted by researchers to the problem of measuring emotion on a rating scale is to add more response dimensions. Here, the problem of cognitive load and attentional resources must be considered carefully (Damos 1992; M. Zentner, personal communication). If enough dimensions are responded to simultaneously, researchers will obtain a more complete picture of emotion expressed but, on the other hand, the participant's capacity to execute the more complex task will lead to poorer quality of data. The general consensus in the research community appears to be that two dimensions provide an optimal balance between a parsimonious definition of emotion and limiting the complexity of the task. While there is a lack of agreement on

the exact nature of these dimensions, they agree in principle with two of the three dimensions of semantic space proposed by Osgood *et al.* (1957). Through their work on semantic differentials, they asserted that meaning space could be largely described in terms of the dimensions *evaluation*, *activity*, and *potency*. Labels implying the first two dimensions (evaluation and activity) are commonly found in the music–emotion literature (e.g. Asmus 1985; Collier 1996; Collins 1989; Wedin 1972). The labels used by Madsen, Mulder, Schubert, and Tyler are listed in Table 17.2, grouped so as to reflect the matching dimensional constructs.

Table 17.2 Two-dimensional response labels

| Evaluation | Activity | Examples |
|---|---|---|
| Ugly-beautiful | Relaxing-exciting | Madsen (1997, 1998) |
| Valence | Arousal | Schubert (1996) |
| Unhappy-happy | Not excited-excited | Mulder (1997); Tyler (1996) |

It remains to be seen whether a third dimension will be added by future researchers, what the dimension will be, and how the response format will be interfaced. Based on the literature (Russell 1980; Whissell *et al.* 1986), a possible third dimension is likely to include *potency*, with polar labels of *dominant–submissive*. The participant of two-dimensional continuous studies usually moves the mouse around a two-dimensional surface. Either a three-dimensional mouse could be employed (M. Zentner, personal communication), or the participant's unused hand could be used to provide additional information. A second pass in listening is a problematic solution, because it makes the assumption that the listener is responding on the second pass as he or she would if asked to perform the third dimension task on the first pass. However the 'second-pass' approach (with counterbalancing across listeners) seems to be the most feasible alternative.

The rating scale format is particularly suited to a computer interface. In a study on the reliability and validity of the 2DES, Schubert (1999*b*) found that participants required a resolution of approximately 10 per cent per dimension to distinguish words and pictures of faces expressing different emotions. On a 201-point scale, there was general agreement as to the emotion expressed by a word with a boundary of 20 × 20 points. Therefore, a coding that provides a higher resolution than one unit on a ten-unit scale should adequately capture the resolution of human semantic and emotional response. Given that computer displays are restricted by the number of pixels (in the order of hundreds), the issue of resolution and coding of rating scale formats using computer interfacing poses no problem to continuous response research.

### 17.4.3 Synchronization

Early studies had considerable problems with synchronization. Mull's (1949) use of a stop watch as a means of synchronizing response and stimulus was apparently unreliable and the dependent data were discarded. Hevner's (1936) ingenious solution of using numbers for different sections of music was adequate provided that the piece had clearly distinguishable moods or sections. However, it mainly served to distinguish gross characteristics of the M→E system, rather than more subtle ones.

While computer technology has mitigated problems of synchronizing the musical signal with the emotional response, there are other techniques that do not depend on computer technology. For example, Sloboda (1991) used the musical score as a response form. This method ensured precise matching of response and stimulus because the participant was responsible for locating the emotion-evoking episode on the musical map of the stimulus (a musical score). A problem with using the musical score as a response sheet is that it restricts respondents to those who are able to read the particular score in question, although in the Sloboda study respondents unable to read music could refer to recordings. Furthermore, data collection so removed from conventional pen and paper formats could be a precarious task, as the coder searches for some mark, perhaps clear, perhaps esoteric, that signifies an emotional, or some other, event. Sloboda's solution was to have the participant make a mark on the score and then provide a written explanation (J. A. Sloboda, personal communication). The techniques employed by Flowers (1983) and Waterman (1996) also used a physical link between the response and the stimulus, the former by recording the stimulus and response onto an audio tape, the latter by using one of the spare audio tracks, parallel to the stimulus storing tracks, for recording responses.

By far the most common method of synchronization availed to the contemporary researcher is the use of a computer-generated clock which the software uses to track the location of the music and is used as a tag for the participant's response. In the case of Schubert (1996, 1999*b*), the musical stimuli were taken from an audio CD that was controlled by the software. One routine in the software (Sudderth 1995) was able to read the elapsed time (in seconds) directly from the CD. This value was then tagged with the response in real time. After data collection, the experimenter converted the time into bar (measure) locations so that the concomitant musical structure could be easily identified. This conversion to bars made the analysis independent of the CD recording of the piece under investigation. Another technique of making the music independent of the performance was to employ the method of Sloboda *et al.* (1997), who recorded response after every sixth note. MIDI (Musical Instrument Digital Interface) coding of the performance lent itself to this kind of synchronization strategy.

### 17.4.4 Sampling rate

While a variety of sampling rates have been found in the self-report literature, varying from five emotional response samples per second (Madsen *et al.* 1993*a*) to one per second (e.g Brittin 1991), the majority of studies have used two samples per second. The choice of two samples per second appears to be based on intuition rather than a firm theoretical foundation. However, the lag structure of the M→E system provides a clue of the optimum sampling rate—optimum meaning the balance between minimum data and maximum information. By applying Nyquist's sampling theorem (e.g. Pohlmann 1995), that the optimum sampling rate is twice the highest frequency which is likely to be encountered, it is possible to posit a suitable sampling rate for measuring continuous emotional response to music. Although no literature was cited which indicated the fastest *emotional* response to a musical feature, the fastest response to a musical feature alone necessarily sets the upper limit (assuming that the listener is not trying to predict an emotional response). Geringer (1995) reported that the response latency

in detecting loudness information was in the order of 1.5 seconds (see also Robinson 1988, p. 64). On the basis of this information, responses should be sampled at half this interval—once every 0.75 seconds. Therefore, the popular, slightly higher sampling rate of two per second is a good choice (Gregory 1992).

### 17.4.5 Lag structure

The main benefit to be sought from continuous measures of emotion in music is to determine why responses change as a function of the stimulus. Surprisingly, very few studies have attempted to do this. Unfortunately, the kinds of issues investigated by continuous response studies verge on trivial. Nearly all the studies investigated *describe* the movement of the response time-series, but only a handful have made serious attempts to *explain* these responses in terms of the changing stimulus (e.g. Capperella-Sheldon 1992; Krumhansl 1996, 1998; Nielsen 1983; Schubert 1998, 1999*a*; Sloboda 1991).

Given the primitive state of stimulus-response (M→E) analysis, it comes as no surprise that our understanding of lag structure in emotion in music is minimal. Lag structure refers to the delay in response after an emotion-producing stimulus event. When any stimulus is presented, the response is not instantaneous. Schubert (1998, 1999*a*) proposed that responses are causally related to the stimulus. Consequently, he analysed response according to specific musical features. Using regression analysis (discussed below), he found that a dynamic relationship between emotional arousal and loudness was present. In general, positive changes in stimulus loudness led to positive changes in the arousal dimension of emotion, and these changes occur 2–3 seconds after the change in the stimulus. This is in reasonable agreement with Krumhansl's (1996) identification of a lag of up to 3.25 seconds from the time an event occurs to the time a tension response is made.

Schubert and Dunsmuir (1999) further proposed that the M→E system is non-linear. For example, sudden changes in loudness produce a faster change in arousal response (1–2 seconds instead of 2–3 seconds). In other words, lag structure is dynamic. Indeed, understanding lag structure is perhaps the single most important application of continuous response methodology. However, the work of Schubert and Dunsmuir is introductory and largely exploratory.

Continuous response methodology provides a fruitful foundation for the investigation of the M→E system dynamics. A hint of the value of this methodology has been provided in Schubert and Dunsmuir's (1999) categories of outliers responsible for altering the hypothetically linear lag structure of the M→E system (or more specifically in the case cited, the 'loudness→arousal system'). They proposed that a faster than typical linear response lag is a *startle* response, while a slower than typical linear lag is an *inertia* response. Three kinds of inertia responses were identified:

1. *Orientation*, where the participant needs time to adjust to a new piece or section of music.
2. *Overshoot recovery*, where the participant has overreacted to a change and requires extra time to recover to the linear model-predicted lag.
3. *After-glow*, where the participant still notices an emotion in the music, even though the music has finished.

An example of an overshoot recovery inertial response is shown in Fig. 17.2 (taken from Schubert 1999*a*). The system-predicted arousal response (based on regression analysis) is shown for seconds 137 to 146 (bars 166–180) of Dvorak's Slavonic Dance Op 42 No. 1. When 142 seconds of the piece had elapsed, the actual change in arousal dropped suddenly at the same time as the change in loudness and spectral centroid (which is related to perceived timbral sharpness) increased. The rate at which the model-predicted arousal gradient changed (Δ) was much slower, and the model was not able to predict the actual response satisfactorily.

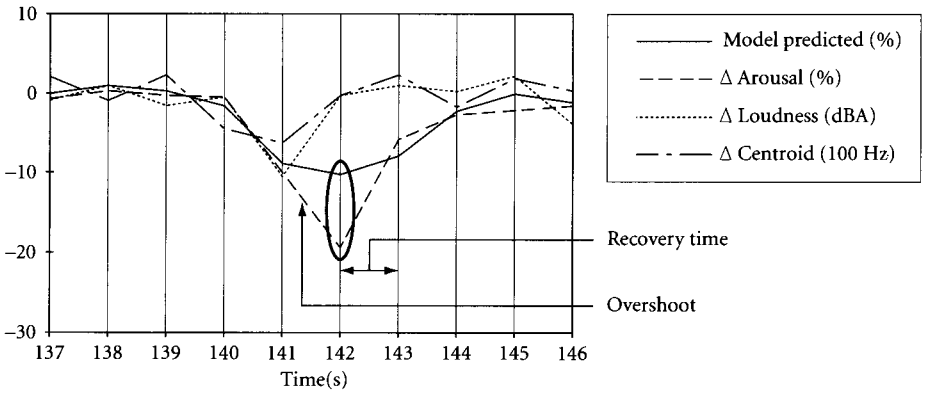



**Fig 17.2** Example of an overshoot recovery inertia response.

Another prudent area for investigating lag structure is in how individual differences affect response time. A common example is the comparison of 'musician and non-musician' responses to music. Capperella-Sheldon (1992) defined such groups and examined where aesthetic peaks and drops in responses occurred while listening to the first movement of Holst's First Suite in E♭ for Military Band. She found five peaks and drops for each group, but could not identify any consistency in response lag between the musicians and non-musicians.

### 17.4.6 Analysis

Perhaps the most grave concern faced by continuous response researchers in music-emotion is the analysis of data. While econometricians and engineers have developed traditions of sophisticated time-series analytical techniques (Box & Jenkins 1976; Hamilton 1994), in the field of self-report emotion research, music psychologists have failed to do much more than provide descriptive, interoccular tests of their continuous response findings. This is somewhat odd because music and emotion are both time series, requiring special analytical techniques. Consequently, an important advantage of continuous response measurement has been neglected, if not resisted. While general empirical psychology publications are subjected to fairly intense statistical rigour and scrutiny, the continuous emotional response to music literature has produced a stream of studies that report simple, weakly substantiated, descriptive claims. This

means that the potential benefits of continuous measure methodology are, to some extent, wasted.

Specifically, continuous response enables measurement in response to an entire piece or movement of music. Instead of listening to a few bars of a piece, the music context can be laid out for the listener, and a response can be made in a more realistic listening situation. Measurement of this kind of approach provides the advantage of ecological validity (Lipscomb 1996). In the statistical world, this realistic approach may be best described as a process involving *serial correlation*—that a response or experience at a given point in time is in some way related to something that happened at another point (or points) in time. It is this process of serial correlation that must be addressed to make continuous measures productive tools for research. The analytical approaches used by researchers can be divided into four broad classes that are, in order of (more or less) increasing validity and sophistication: interoccular analysis, correlation of means, analysis of variance, and adjustment for serial correlation.

*Interoccular analysis*

The dimensional rating-scale studies provide potential for a closer analysis of the time-variant relationship between emotional response and musical features (Schmidt 1996). However, there are numerous instances in the literature where a complete analysis will consist of general comments about the appearance of a time-series chart of mean responses. For example, Gregory (1995, pp. 206–7) concludes her review of the CRDI by stating that 'visual comparisons of graphic contours provide immediate discrimination of response similarities across total time spans and at selected samples within time spans. This may be the CRDI's singular contribution to music research instrumentation.'

While visual inspection is an important approach for commencing analysis, it would be unfortunate if this remains the 'singular contribution' of the CRDI or any other continuous response measure. Gottman (1981) points out the foibles of the interoccular test (or 'graphic analysis') that is implied by Gregory's comment. It lacks statistical rigour and is prone to biased interpretation. Further, visual analysis of a chart can be misleading, particularly when deviation scores are omitted—a regular feature of continuous response data reports in the literature under investigation.

*Correlation of means*

A common kind of analysis reported in continuous measure studies is the Pearson product-moment correlation analysis of time-series data averaged across participants. Correlation is used to describe the similarity of two time series, and is the most common statistical approach used for comparative studies (such as comparing non-musicians with musicians). With comparison of categorical data, agreement quotients have been used, where the numerical index is obtained by dividing the number of categories in which there are concomitant agreements across test–retest listenings by the total number of categorical responses (Gregory 1995). None of the studies cited have attempted to provide correlation analyses with serial correlations removed, and only a handful have considered a cross-correlation analysis (Schubert 1999*a*), which provides information about correlations at non-instantaneous lags (i.e. correlations between two time series separated by one sample or more).

Schubert (2000) demonstrated that correlation coefficients were inflated in comparison to correlation analyses made with repeated measures of independent data sets. Indeed, the literature does have a tendency to report boastfully though misleadingly high correlation coefficients, often well above 0.8. For example, Johnson (1996) compared the experimentally obtained correlation coefficient of serially correlated data with the range recommended by Anastasi (1987, p. 106, cited in Johnson 1996) for a reliable measure, without considering this problem. Schmidt (1996) has suggested that, in some cases, a Spearman rank-correlation analysis might be more appropriate because it is more conservative. There are other non-parametric approaches to correlation analyses, and many of the techniques discussed below are applicable (e.g. see Howell 1997). Schubert (2000) proposed that differencing each time series (so as to produce *changes* in response) prior to conventional Pearson product-moment correlation might produce a more valid correlation analysis.

*Accounting for deviation*

A more sophisticated approach than interoccular tests and correlation analyses are techniques that examine variations in responses. The first study of continuous response to emotion in music cited which used mathematical techniques to analyse time-series data was the de Vries (1991) study. While it used the *sentograph* (see Section 17.2.3), and so strictly speaking is not a self-report instrument, it is worth examining the techniques de Vries used. His approach was based on the method of analysis of electrocardiograms developed by Van Bemmel (1982, cited in de Vries 1991). De Vries produced five parameters from a two-dimensional response time series (force and angle of finger pressure). The parameters were: (1) mean angle of pressure direction; (2) roughness (or fluctuation) of pressure angle; (3) roughness of force; (4) pointedness of force; and (5) skewness of force. This produced five time series. Discriminant analysis resulted in three weighting functions that were produced so as to maximize the distinction in responses across musical selections. In other words, de Vries could map musical selections onto a three-dimensional music space using the transformed response parameter weightings. Together, these three functions accounted for 97 per cent of the variance. However, de Vries did not disclose if and how he accounted for serial correlation. Uncontrolled serial correlation may have inflated the explained variance.

The first published *self-report* study among those reported to use some form of inferential statistical analysis (aside from instantaneous correlations) was by Byrnes (1996), who used paired sample *t*-tests to determine the effect of the music as a 'treatment' on the first and last responses. While this is a statistically valid procedure, it seems difficult to justify the collection of so much data between the first and last samples when only the two end-points are used for the analysis.

Another 'early' attempt at using inferential statistical analysis of time-series data was presented by Sloboda *et al.* (1997), who conducted analysis of variance on the grouped responses to performances of Chopin's Prelude No. 4 in E Minor. After eye-balling the time-series plots, the piece was divided into four sections that also coincided roughly with structurally salient points in the music. The repeated measures ANOVA produced significant results, and it was found that adjacent sections were significantly different. Sloboda and associates then used this information to claim that they had captured

different levels of 'emotionality'. While it is commendable that inferential statistics were applied, the question of why the analysis was performed remains. It is not clear why it was important to have a piece with varied emotional expression within it (as distinct from across different versions of the same piece). More relevant to the research problem of the study would have been 'what produced these changes?' The notion of musical features as the cause of the response was implied: 'The present study aims to discover how score-based structural indicators, observed performance parameters, and listener characteristics influence the emotional response to music' (Sloboda *et al.* 1997, p. 630). To this end, it would have been of greater relevance if the relationship between response variations and musical structure could be identified using the data obtained.

*Accounting for serial correlation*

Krumhansl (1996) used multiple regression to model tension response as a function of musical feature variables based on an analysis by Lerdahl. While no coefficents were reported, the multiple correlation coefficent was estimated as $R_{7\,177} = 0.79$ ($P < 0.0001$). Krumhansl supported Lerdahl's predictions by leading (or 'shifting backward in time') responses by 0 to 3.25 seconds, and found that the predictions were not as smooth as the actual mean responses. Therefore, Krumhansl smoothed and lagged the predicted response, thus introducing serial correlation and lagging into the model, and hence making it more realistic. A second regression analysis produced an improved fit ($R_{14\,170} = 0.91$, $P < 0.0001$). However, it is not clear that Krumhansl overtly accounted for the serial correlation in the model: No comment was made about the serial correlation in the residual series (the time series which could *not* be explained by the regression model). This is important, because the presence of serial correlation in the residual violates several assumptions of the linear regression model (Ostrom 1990). Krumhansl's study is ground-breaking because it attempts to explain lag structure. In a later study (Krumhansl 1998), she applied a similar approach by modelling one-dimensional emotional response in terms of dichotomized musical topics. The model fit was $R_{14\,767} = 0.52$, but again no consideration of modelling serial correlation was mentioned.

Schubert (1998, 1999*a*; Schubert & Dunsmuir 1999) examined stimulus time-series features including tempo, melodic pitch, and loudness, *and* response time-series dimensions of valence and arousal. His approach was to produce a univariate, linear regression analysis, where each emotional dimension was described in terms of a linear combination of the musical features. The traditional least squares approach was not satisfactory because the assumption of independence was violated. Instead, Schubert used a three-step procedure: (1) differencing all variables; (2) lagging the musical feature variables; and (3) making an autoregressive adjustment (Box & Jenkins 1976; Gottman 1981; Ostrom 1990) to model the remaining serial correlation.

Differencing variables provided a meaningful transformation (it meant 'change in response', denoted by the Greek capital delta symbol, 'Δ') and at the same time removed a large amount of serial correlation (see Gottman 1981). In other words, relative values of variables rather than absolute values were modelled. Lagging the variables (by 1, 2, 3, and 4 seconds) enabled the examination of the dynamic lag structure in the M→E system and satisfied the assertion that emotional response is causally affected by

musical features. The remaining serial correlation component was modelled using the technique of autoregressive adjustment (Ostrom 1990). Autoregressive adjustment could be described as the addition to the mathematical model that acted as a kind of 'memory', adding a proportion of the linear-system-produced error of the previous point in time. The difference-lag-autoregressive adjusted regression model produced a series of satisfactory models, explaining up to 73 per cent of the variation in response for arousal, and up to 62 per cent of the variation in response for valence in terms of a selection of musical features alone.

The approach taken by Schubert provides researchers with a relatively simple and valid analytical technique for modelling time-series data. However, this approach will not work in all cases. Schubert assumed that the time series investigated were only first-order autoregressive (Gottman 1981; Ostrom 1990). More research is required to determine whether time-series processes in the M→E system may contain higher-order autoregressive processes, or indeed other processes altogether, heading, for example, toward the more sophisticated ARIMA (Autoregressive Integrated Moving Average) models proposed by Box and Jenkins (1976). The possibility of analysing the data in the frequency domain (Gottman 1981) or using other techniques such as neural network models (Todd & Loy 1991) and dynamic system models (Abraham *et al.* 1990) also await investigation in modelling self-report continuous emotional response data (see also Beran & Mazzola 1999).

## 17.5 Concluding remarks and future directions

The literature on self-report continuous response methodology to understand emotion in music has been growing considerably since the pioneering days of Hevner (1936) and, particularly, since Nielsen (1983). The methodology enables researchers to better understand the moment-to-moment fluctuations in emotional response caused by the musical stimulus within an ecologically valid framework (because the listener is usually listening to an uncontrolled, real piece of music rather than an excerpt or contrived auditory stimulus). This methodology has been used to find convergent evidence with traditional asynchronous approaches regarding, for example, the influence of musical features upon emotional response. However, its main advantages are yet to be fully exploited. In particular, continuous response methodology allows researchers to examine the relationship between moment-to-moment fluctuations in music and response as well as the lag structure of response—that is, how quickly people respond after a change in musical structure, or how quickly they respond in comparison to other people listening to the same music.

Important, major issues which remain unresolved are which dimensions of emotional response should be recorded, the dynamic nature of the M→E system, and which analytical techniques to use. Using two dimensions appears to be a reasonable compromise between a meaningful emotional response and cognitive overload. One such combination of emotion dimensions is *valence* (happy–sad) and *arousal* (aroused–sleepy), although the optimal solution is yet to be determined.

A relatively less important issue in continuous response methodology is the genre of musical stimulus to use. Selections have largely been based on the music of romantic

Western art music idioms (I include Holst's band music in these idioms). While it would be interesting to examine different styles, such as contemporary, pre nineteenth century, and popular styles, it is also sensible for the time-being to restrict the number of styles investigated, given that continuous response methodologies are still being established and are already associated with a potentially large number of complex, interacting variables. Restricting stimuli to the romantic music idioms is logical because these are idioms that are defined by the importance of expressing emotion (e.g. Gorbman 1987).

Continuous self-report measures have arisen from a need to provide greater ecological validity in emotion-in-music investigations and through technological developments that have made such an approach to research feasible. However, researchers using this methodology need to remain cognisant as to why they are using the technology. It seems that experimental music psychologists have required a period of exploration and discovery to more fully realize the potential of continuous response measures.

In conclusion, continuous response researchers need to become more aware of the analytical problems created by serial correlation, an inherent part of much time-series data. There are simple, effective solutions available for dealing with the problem, such as differencing, autoregressive adjustment, and other techniques associated with regression-type analysis. It is vital that researchers are willing to apply these methods and other non-parametric techniques to maximize the benefits of continuous research methodology, and in particular to exploit its capacity to provide a basis for understanding the still elusive dynamic nature of emotion in music.

## References

**Abraham, F. D., Abraham, R. H., Shaw, C. D., & Garpike, A.** (1990). *A visual introduction to dynamical systems theory for psychology.* Santa Cruz, CA: Aerial.

**Adams, B. L.** (1994). *The effect of visual/aural conditions on the emotional response to music.* Unpublished doctoral dissertation, Florida State University, Tallahassee.

**Anastasi, A.** (1987). *Psychological testing,* (4th edn). New York: Macmillan.

**Asmus, E. P.** (1985). The development of a multidimensional instrument for the measurement of affective responses to music. *Psychology of Music,* **13**, 19–30.

**Beran, J. & Mazzola, G.** (1999). Analysing musical structure and performance: A statistical approach. *Statistical Science,* **14**, 47–79.

**Berlyne, D. E.** (1973). Interrelations of verbal and nonverbal measures used in experimental aesthetics. *Scandinavian Journal of Psychology,* **14**, 177–84.

**Box, G. E. P. & Jenkins, G. M.** (1976). *Time series analysis: Forecasting and control,* (rev. edn). San Francisco: Holden-Day.

**Brittin, R. V.** (1991). The effect of overtly categorizing music on preference for popular music styles. *Journal of Research in Music Education,* **39**, 143–51.

**Brittin, R. V. & Duke, R. A.** (1997). Continuous versus summative evaluations of musical intensity: A comparison of two methods for measuring overall effect. *Journal of Research in Music Education,* **45**, 245–58.

**Byrnes, S. R.** (1996). The effect of audio, video, and paired audio-video stimuli on the experience of stress. *Journal of Music Therapy,* **33**, 248–60.

Cabrera, D. (1997). *Db&dba&centr.orc* (Computer software). Sydney, Australia.

Cabrera, D. (1998). *PsySound* (Computer software). Sydney, Australia.

Capparella-Sheldon, D. A. (1992). Self-perception of aesthetic experience among musicians and non-musicians in response to wind band music. *Journal of Band Research,* 28, 57–71.

Clynes, M. (1977). *Sentics: The touch of emotions.* New York: Anchor Press/Doubleday.

Clynes, M. (1980). The communication of emotion: Theory of sentics. In *Emotion: Theory research and experience,* (Vol. 1) (ed. R. Plutchik & H. Kellerman), pp. 271–300. New York: Academic Press.

Clynes, M. & Nettheim, N. (1982). The living quality of music: Neurobiologic basis of communicating feeling. In *Music, mind, and brain: The neuropsychology of music,* (ed. M. Clynes), pp. 47–82. New York: Plenum Press.

Collier, G. L. (1996). Affective synesthesia: Extracting emotion space from simple perceptual stimuli. *Motivation and Emotion,* 20, 1–32.

Collins, S. C. (1989). *Subjective and autonomic responses to western classical music.* Unpublished doctoral dissertation, University of Manchester, Manchester, UK.

Damos, D. (1992). *Multiple task performance.* London: Taylor Francis.

De Vries, B. (1991). Assessment of the affective response to music with Clynes's sentograph. *Psychology of Music,* 19, 46–64.

Flowers, P. J. (1983). The effect of instruction in vocabulary and listening on nonmusicians' descriptions of changes in music. *Journal of Research in Music Education,* 31, 179–89.

Fredrickson, W. E. (1995). A comparison of perceived musical tension and aesthetic response. *Psychology of Music,* 23, 81–7.

Fredrickson, W. E. (1997). Elementary, middle, and high school perceptions of tension in music. *Journal of Research in Music Education,* 45, 626–35.

Fredrickson, W. E. (1999). Effect of musical performance on perception of tension in Gustav Holt's First Suite in E-flat. *Journal of Research in Music Education,* 47, 44–52.

Frego, R. J. D. (1999). Effects of aural and visual conditions on response to perceived artisitic tension in music and dance. *Journal of Research in Music Education,* 47, 31–43.

Gabrielsson, A. & Juslin, P. N. (in press). Emotional expression in music. In *Handbook of affective sciences,* (ed. R. J. Davidson, H. H. Goldsmith, & K. R. Scherer). New York: Oxford University Press.

Gabrielsson, A. & Lindström, E. (1995). Emotional expression in synthesizer and and sentograph performance. *Psychomusicology,* 14, 94–116.

Geringer, J. M. (1995). Continuous loudness judgments of dynamics in recorded music excerpts. *Journal of Research in Music Education,* 43, 22–35.

Goins, W. E. (1998). The effect of moodstates: Continuous versus summative responses. *Journal of Music Therapy,* 35, 242–58.

Goldstein, A. (1980). Thrills in response to music and other stimuli. *Physiological Psychology,* 8, 126–9.

Gorbman, C. (1987). *Unheard melodies: Narrative film music.* Bloomington, IN: Indiana University Press.

Gottman, J. M. (1981). *Time-series analysis: A comprehensive introduction for social scientists.* Cambridge, UK: Cambridge University Press.

Greene, B. A. & Royer, J. M. (1994). A developmental review of response time data that support a cognitive components model of reading. *Educational Psychology Review,* 6, 141–72.

Gregory, A. H. (1998). Tracking the emotional response to operatic arias. In *Proceedings of the Fifth International Conference of Music Perception and Cognition*, (ed. S. W. Yi), pp. 265–70. Seoul, Korea: Western Music Research Institute, College of Music, Seoul National University.

Gregory, D. (1992). Comparisons of zone delineations, sampling rates, and graphic options of the continuous response digital interface. Paper presented at the Music Educators National Conference, New Orleans, LA.

Gregory, D. (1995). The continuous response digital interface: An analysis of reliability measure. *Psychomusicology*, 14, 197–208.

Gregory, D. (1998). Reactions to ballet with wheelchairs: Reflections of attitudes toward people with disabilities. *Journal of Music Therapy*, 35, 274–83.

Hair, H. I. (1995). Mood categories of lines, colors, words, and music. *Bulletin of the Council for Research In Music Education*, 127, 99–105.

Hall, D. E. (1987). *Basic acoustics*. New York: Wiley.

Hamilton, J. D. (1994). *Time series analysis*. Princeton, NJ: Princeton University Press.

Hargreaves, D. J. (1986). *The developmental psychology of music*. Cambridge, UK: Cambridge University Press.

Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology*, 48, 246–68.

Howell, D. C. (1997). *Statistical methods for psychology*, (4th edn). Belmont, CA: Duxbury.

Johnson, C. M. (1996). Musicians' and nonmusicians' assessment of perceived rubato in musical performance. *Journal of Research in Music Education*, 44, 84–96.

Krumhansl, C. L. (1996). A perceptual analysis of Mozart's Piano Sonata K.282—Segmentation, tension, and musical ideas. *Music Perception*, 13, 401–32.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and physiology. *Canadian Journal of Experimental Psychology*, 51, 336–52.

Krumhansl, C. L. (1998). Topic in music: An empirical study of memorability, openness, and emotion in Mozart's String Quintet in C Major and Beethoven's String Quartet in A Minor. *Music Perception*, 16, 119–34.

Krumhansl, C. L. & Schenck, D. L. (1997). Can dance reflect the structural and expressive qualities of music? A perceptual experiment on Balanchines's choreography of Mozart's divertiment No. 15. *Musicae Scientiae*, 1, 63–85.

Lerdahl, F. & Jackendoff, R. (1983). *A generative theory of tonal music*. Boston: MIT Press.

Lipscomb, S. D. (1996). The cognitive organization of musical sound. In *Handbook of music psychology*, (ed. D. A. Hodges), pp. 133–75. San Antonio, TX: IMR Press.

Lychner, J. (1998). An empirical study concerning terminology relating to aesthetic response to music. *Journal of Research in Music Education*, 46, 303–20.

Madsen, C. K. (1996). Empirical investigation of the 'aesthetic response' to music: Musicians and non-musicians. In *Proceedings of the Fourth International Conference of Music Perception and Cognition*, (ed. B. Pennycook & E. Costa-Giomi), pp. 103–10. Montreal, Canada: McGill University.

Madsen, C. K. (1997). Emotional response to music as measured by the two-dimensional CRDI. *Journal of Music Therapy*, 34, 187–99.

Madsen, C. K. (1998). Emotion versus tension in Haydn's symphony No. 104 as measured by the two-dimensional continuous response digital interface. *Journal of Research in Music Education*, 46, 546–54.

**Madsen, C. K., Brittin, R. V., & Capperella-Sheldon, D. A.** (1993). An empirical investigation of the aesthetic response to music. *Journal of Research in Music Education,* **41**, 57–69.

**Madsen, C. K., Byrnes, S. R., Capperella-Sheldon, D. A., & Brittin, R. V.** (1993). Aesthetic responses to music: Musicians vs. nonmusicians. *Journal of Music Therapy,* **30**, 174–91.

**Madsen, C. K. & Fredrickson, W. E.** (1993). The experience of musical tension: A replication of Nielsen's research using the continuous response digital interface. *Journal of Music Therapy,* **30**, 46–63.

**Moore, B. C. J., Glasberg, B. P., & Baer, T.** (1997). A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society,* **45**, 224–40.

**Mulder, S. J.** (1997). *The identification of the emotional content in choral music: A descriptive study using the two-dimensional continuous response digital interface (2D CRDI).* Unpublished doctoral dissertation, Florida State University, Tallahassee.

**Mull, H.** (1949). A study of humor in music. *American Journal of Psychology,* **62**, 560–6.

**Namba, S., Kuwano, S., Hatoh, T., & Kato, M.** (1991). Assessment of musical performance by using the method of continuous judgement by selected description. *Music Perception,* **8**, 251–76.

**Narmour, E.** (1990). *The analysis and cognition of basic melodic structures.* Chicago: University of Chicago Press.

**Nielsen, F. V.** (1983). *Oplevelse af musikalsk spænding [The experience of musical tension].* Copenhagen: Akademisk Forlag.

**Nielsen, F. V.** (1987). Musical tension and related concepts. In *The semiotic web '86. An international year-book,* (ed. T. A. Sebeok & J. Umiker-Seboek), pp. 491–513. Berlin: Mouton de Gruyter.

**Osgood, C. E., Suci, G.J., & Tannenbaum, P. H.** (1957). *The measurement of meaning.* Urbana, IL: University of Illinois Press.

**Ostrom, C. W.** (1990). *Time series analysis regression techniques.* Newbury Park, CA: Sage Publications.

**Panksepp, J.** (1995). The emotional sources of 'chills' induced by music. *Music Perception,* **13**, 171–208.

**Pohlmann, K. C.** (1995). *Principles of digital audio,* (3rd edn). New York: McGraw Hill.

**Radocy, R. E. & Boyle, J. D.** (1988). *Psychological foundations of musical behavior. Psychological foundations of musical behaviour,* (2nd edn). Springfield, IL: Charles C. Thomas.

**Ratner, L. G.** (1980). *Classic music: Expression, form and style.* New York: Schirmer.

**Robinson, C. R.** (1988). *Differentiated modes of choral performance evaluation using traditional procedures and a continuous response digital interface device.* Unpublished doctoral disseration, Florida State University, Tallahassee.

**Rossing, T. D.** (1990). *The science of music,* (2nd edn). Reading, MA: Addison-Wesley.

**Russell, J. A.** (1980). A circumplex model of affect. *Journal of Social Psychology,* **39**, 1161–78.

**Russell, J. A.** (1989). Measures of emotion. In *Emotion: Theory, research, and experience. Vol. 4. The measurement of emotion,* (ed. R. Plutchik & H. Kellerman), pp. 81–111. New York: Academic Press.

**Schmidt, C. P.** (1996). Research with the continuous response digital interface: A review with implications for future research. *Philosophy of Music Education Review,* **4**, 20–32.

**Schubert, E.** (1996). Continuous response to music using a two dimensional emotion space. In *Proceedings of the Fourth International Conference of Music Perception and Cognition,* (ed. B. Pennycook & E. Costa-Giomi), pp. 263–8. Montreal, Canada: McGill University.

**Schubert, E.** (1998). Time series analysis of emotion in music. In *Proceedings of the Fifth International Conference of Music Perception and Cognition,* (ed. S. W. Yi), pp. 257–63. Seoul, Korea: Western Music Research Institute, College of Music, Seoul National University.

**Schubert, E.** (1999*a*). *Measurement and time series analysis of emotion in music.* Unpublished doctoral dissertation, University of New South Wales, Sydney.

**Schubert, E.** (1999*b*). Measuring emotion continuously: Validity and reliability of the two dimensional emotion space. *Australian Journal of Psychology,* 51, 154–65.

**Schubert, E.** (2000). Unresolved issues in continuous response methodology: The case of time series correlations. In *Proceedings of the Sixth International Conference of Music Perception and Cognition,* (CD rom) (ed. C. Woods, G. Luck, R. Brochard, F. Seddon, & J. A. Sloboda). Keele, UK: Keele University.

**Schubert, E. & Dunsmuir, W.** (1999). Regression modelling continuous data in music psychology. In *Music, mind, and science,* (ed. S. W. Yi), pp. 298–352. Seoul, Korea: Western Music Research Institute, College of Music, Seoul National University.

**Sloboda, J. A.** (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music,* 19, 110–20.

**Sloboda, J. A.** (1992). Empirical studies of emotional response to music. In *Cognitive bases of musical communication,* (ed. M. R. Jones & S. Holleran), pp. 33–46. Washington, DC: American Psychological Association.

**Sloboda, J. A., Lehmann, A. C., & Parncutt, R.** (1997). Perceiving intended emotion in concert-standard performances of Chopin's Prelude No. 4 in E-minor. In *Proceedings of the Third Triennial ESCOM Conference,* (ed. A. Gabrielsson), pp. 629–34. Uppsala, Sweden: Department of Psychology, Uppsala University.

**Sudderth, J.** (1995). *CoreCD (version 1.4)* (Computer software). Core Development Group, Inc.

**Todd, P. M. & Loy, G. D.** (ed.) (1991). *Music and connectionsim.* Cambridge, MA: MIT Press.

**Tyler, P.** (1996). *Developing a two-dimensional continuous response space for emotions perceived in music.* Unpublished doctoral dissertation, Florida State University, Tallahassee.

**Waterman, M.** (1996). Emotional responses to music: Implicit and explicit effects in listeners and performers. *Psychology of Music,* 24, 53–67.

**Watson, K. B.** (1942). The nature and measurement of musical meanings. *Psychological Monographs,* 54, 1–43.

**Wedin, L.** (1972). Evaluation of a three-dimensional model of emotional expression in music. *Reports from the Psychological Laboratories, University of Stockholm,* 349, 20.

**Weld, H. P.** (1912). An experimental study of musical enjoyment. *American Journal of Psychology,* 23, 245–308.

**Whissell, C. M., Fournier, M., Pelland, R., Weir, D., & Makarec, K.** (1986). A dictionary of affect in language: IV. Reliability, validity, and applications. *Perceptual and Motor Skills,* 62, 875–88.

CHAPTER 18

# EMOTIONS IN EVERYDAY LISTENING TO MUSIC

JOHN A. SLOBODA AND SUSAN A. O'NEILL

There are several key assumptions underpinning our focus on emotion and 'everyday' music listening. Firstly, music is always heard in a social context, in a particular place and time, with or without other individuals being present, and with other activities taking place which have their own complex sources of meaning and emotion. The emotional response to the music is coloured, and possibly sometimes completely determined, by these contextual factors. Emotions should not be thought of as abstract entities such as 'anger' or 'elation', but rather as actual moments of emotional feelings and displays in particular situations within a particular culture. Emotional responses to music are linked to a sequence of events based on conventions and rules that depend not only on shared understanding and representations, but also a common background of knowledge and beliefs. Music is a cultural material (as is language) that provides a kind of semiotic and affective 'power' which individuals use in the social construction of emotional feeling and displays. As such, the impact of music on emotion is not direct but interdependent on the situations in which it is heard. Any meaningful account of music's role in the emotional response of individuals must involve the recognition of these complex, interdependent social factors. We therefore focus in this chapter on several investigations which attempt to preserve, or take account of, as much of the social context of music listening as possible.

Secondly, music is ubiquitous in contemporary life. The prevailing contexts in which we encounter music are, by definition, mundane. They include those contexts in which the most routine activities of life take place: waking up, washing and dressing, eating, cleaning, shopping, travelling. There are, of course, special and 'out of the ordinary' events which music can form a part of, even a crucial and defining part, but we would argue that in order to understand fully these special events and their emotional significance, we must consider them in relation to everyday experience. It is the everyday and normal which frames and helps define the special. This chapter focuses, therefore, on habitual and routine modes of engagement with music.

Thirdly, despite the growing recognition that our experience of emotion is inextricably linked to the social world and the linguistic practices used to make sense of that world, we tend to think of our emotions as personal, 'private' experiences, especially if they do not involve 'public' emotional displays. For example, when we feel angry, we might experience our anger as a private emotion which may or may not be 'acted out' in the things we do or say. However, this is not the case in all cultures. Lutz (1982) describes

the way in which the Ifaluk (Samoan and Pintupi Aborigines) use emotion words to describe their relationship to events and other people, and not as expressions of private, internal states. The closest translation of anger in the Ifaluk language is the word 'song' which means 'justifiable anger'. It is not considered a privately owned feeling but rather a moral and public account of some transgression of accepted social practices and values (cf. Burr 1995). As such, emotional feelings and displays are not considered to represent internal states of an individual (whether innate or acquired through learning and experience) that result in physiological reactions to environmental stimuli. Rather, they are meaningful displays that are taken as emotional when they are embodied expressions of judgments, and in many cases, ways of accomplishing certain social acts (Harré & Gillett 1994). Our focus in this chapter is on research which attempts to understand, at least in part, the inextricable connection between individual 'personal' emotional experience and the 'social' forms of meaning associated with music listening.

Our perspective makes it necessary to exclude any significant treatment of three types of phenomenon. The first exclusion relates to the extremely intense, sometimes transcendent, experiences to music which people characterize as defining or life-changing (e.g. Gabrielsson & Lindström 1995; see also Gabrielsson, this volume). These events rarely occur in the lives of individuals, and so are difficult to study empirically. One retrospective study by Sloboda (1989*a*), which we will discuss in more detail in a later section, provides suggestive evidence that these 'peak' events can occur in childhood, and, when they do, can have a profound influence on a person's attitudes towards, and commitment to, music. Indeed, people who devote their lives to music often include such experiences in the stories of their developing commitment (e.g. Brodsky 1995). This, then, brings us to the second exclusion of our chapter. We shall not be examining the everyday emotional world of the professional musician or the intending professional musician (i.e. the music student). The emotional experiences of these groups are covered in other chapters in this book (Persson, this volume), and we would wish to claim that the everyday contexts in which musicians experience music are so different from those of non-musicians (who form the vast bulk of the population of industrialized nations) that they require separate treatment. Finally, we do not intend to provide a comprehensive review of the literature associated with emotional response to everyday music listening. Much of this literature is covered elsewhere in this book (DeNora, this volume; Becker, this volume; Gabrielsson, this volume). In many cases, previous research in this area include laboratory studies or semiquasi experimental research, which have quite specific social features that make it unlike many other music-listening situations. Rather, our aim is to highlight a number of key features of research in this area which are based on the assumptions outlined above, using examples from several studies, many of which we have been involved in ourselves. In this way, we hope to address important questions that are left unanswered by laboratory research, stimulate further discussion and debate on the topic, as well as indicate some directions for future research.

## 18.1 Capturing everyday experience

One of the problems with studying everyday experience is that it happens outside the laboratory in a wide variety of settings. Although an investigator may make observations

(or even attempt to experimentally manipulate) musical engagement in a single setting, it is difficult to observe directly the entire range of settings in which music might be experienced, even in the course of a single day. A second problem is that the everyday is, by definition, unmemorable, and so retrospective studies (such as interviews) may not capture the richness and diversity of musical experience. The more mundane occurrences are simply forgotten or filtered out. Understanding how music interacts with everyday contexts requires a method that not only examines the phenomenon in 'real-world' situations or events, but that does so as the events are unfolding and are experienced by individuals in the course of their daily lives. For these reasons, we (Sloboda *et al.*, 2001) recently conducted a pilot study with eight individuals where we adapted a method developed by Csikszentmihalyi and Lefevre (1989) which enables everyday experiences to be studied in some detail. Our aim in presenting this research is not to make generalizable claims about the frequency of musical occurrences in everyday life. Rather, our aim is to illustrate, through this exploratory research, a method in which we might begin to capture individual subjective experience of emotion concurrent with everyday activities where music is heard, and highlight the heuristic features of the findings in order to inform future theoretical development and research.

The method we employed in our study is known as the experience sampling method (ESM). It involves participants carrying electronic pagers with them at all times during their waking hours. The pagers are connected to a computer that is programmed to call participants at preprogrammed intervals. In our study, participants were paged once in each 2–hour period between 0800 and 2200. The precise timing of a call within each 2–hour block was determined randomly. On each paging, participants were asked to stop what they were doing as soon as practicable and complete one page of a response booklet, which they were also asked to carry with them at all times. The booklet asked questions regarding the most recent experience of music listening since the previous paging. If there had been no music experienced in that period, participants completed the booklet with respect to the activity taking place when the pager sounded.

The eight participants in our study were all adult non-musicians between the ages of 18 and 40 years who were either studying or working at our institution (half of the participants were not involved in psychology). They each carried a pager with them for a 1–week period. Analyses of the data collected provided some quite revealing information about their everyday music use. First, 91 per cent of pagings resulted in a page of the booklet being filled in (we call each event described in the response booklet an 'episode'). This suggested a high level of compliance and involvement in this task over an extended period of time. It also meant that the method provided an almost complete picture of the musical and non-musical experiences of each participant over a normal week (none of the participants reported the week in question being unusual or atypical). The 9 per cent of pagings which did not result in a report sheet being recorded were primarily due to technical problems, participants not having the pager with them at the time (e.g. during swimming or showering), or participants not hearing the pager's signal (e.g. due to travelling or being in a very noisy location).

A second major finding was that 44 per cent of all episodes involved music. In other words, there was a 44 per cent likelihood of music being experienced in any 2–hour

period. It has repeatedly been claimed that music pervades everyday life. Our data confirm this and provide a specific estimate of its frequency.

Participants were asked to respond in their own words to the question: 'what was the *main* thing you were doing?' Reported activities were coded *post hoc* according to three main categories: personal, leisure, and work. Personal activities cover those everyday activities that are a necessary consequence of living, and were further divided into states of being (e.g. sleeping, waking up, being ill, suffering from a hangover), maintenance activities (e.g. washing, getting dressed, cooking, eating at home, housework, shopping), and travel (e.g. leaving home, driving, walking, going home). Leisure activities were divided into three subcategories: music listening (there were no examples of performing music), leisure-passive (e.g. watching TV/film, putting on the radio, relaxing, reading for pleasure), and leisure-active (e.g. games, sports, socializing, eating out, chatting with friends). Work activities were categorized according to whether they were primarily solitary activities (e.g. writing, computing, marking/assessing, reading for study) or primarily group-based activities (e.g. planning for a meeting, in a lecture/seminar, making appointments, in a meeting).

It is particularly significant that listening to music as a main activity accounted for only a small percentage of all episodes (2 per cent). This suggests that the concentrated, attentive, focusing on music that is paradigmatic of the classical concert or the laboratory experiment is a rather untypical activity for most listeners. Instead participants were distributing their attention across a complex situation of which music is only a part. In fact, there were three categories of activity which seemed particularly likely to involve music. These were the personal maintenance, personal travel, and active leisure categories.

Emotional reactions to the music were measured by asking participants to rate their state on eleven bipolar scales (e.g. happy–sad, irritable–generous, bored–interested) with respect to how they felt before the music started and how they felt after it had ended. We found that these scales grouped under three main factors. The first factor related to the degree of positivity or negativity, the second factor to the degree of arousal or alertness, and the third factor to the extent to which participants' attention was focused on the present situation or elsewhere (e.g. reminiscing, daydreaming, nostalgic). These factors have characteristics in common with previous research done on emotional response to music (see Sloboda & Juslin, this volume).

We found that, on average, the experience of music resulted in participants becoming more positive, more alert, and more focused in the present. Insofar as such emotional transitions are desired and beneficial, it could be concluded that in general music made these participants 'feel better'. An examination of the number of episodes resulting in change for each of the factors indicated that where there was change, arousal rarely moved in the direction of lower arousal (7 per cent of changes), and little more so in the direction of less positivity (13 per cent). However, present-mindedness showed 35 per cent of change episodes involving moves away from the 'positive pole'. It seems, in other words, as though emotional states arising from a focus on 'things and people not present', such as nostalgia, are a particularly important subcategory within everyday responses to music.

## 18.2 Functional niches within daily life

Our working assumption has been that music has different emotional functions in different contexts (cf. Merriam 1964). In our ESM study we were, therefore, interested to see whether there were any cases in which mood change factors were dissociated from one another (i.e. cases where one mood increased simultaneously with another mood decreasing). Inspection of cases showed that 16 per cent of music episodes were in this category. These cases are particularly important in beginning to identify distinct functional niches for music engagement. The largest group of such episodes ($n = 10$) involved increases in positivity along with decreases in present-mindedness.

One example of this comes from a male participant who reported being at home relaxing with a group of friends and acquaintances. The activity was being done out of choice. There was ambient music playing on a CD, although the participant had not chosen it. The participant commented that 'the music was very tranquil and relaxing', that others present were 'discussing work boringly', and that he was 'very, very tired'. This episode was also associated with a decrease in arousal during the music. It would be reasonable to assume that the participant was using the music as a means of relaxing and disengaging from the surrounding conversation.

A second example from the same category is provided by a female participant who reported being at home, tidying a bedroom as part of the normal basic routine. The participant had chosen to listen to a piece of pop/chart music on a tape. The participant commented that the music was chosen to 'enhance the wonderful experience of cleaning' and was 'very lively'. This episode was associated with an increase in arousal during the music. It seems as if the purpose of this music was to allow the participant to focus attention on the music, and away from the uninteresting domestic chore, and this focused attention was used to increase energy levels.

It was less easy to find examples of episodes where positivity decreased, but one clear episode involved a female participant at home, alone, doing the washing up as part of a basic routine. She had chosen to listen to rock music on the radio and commented that the track was 'a favourite song I had not heard for some time . . . It brought back certain memories'. The music increased this participant's nostalgia, sadness, and loneliness, at the same time as making her more alert. It is clear that this episode reminded the participant of a significant past event which brought on nostalgia. At the same time, it appears that she had chosen the music to engage and arouse during an uninteresting routine task.

In a separate study, using a stratified sample of 76 panel members from the Sussex Mass Observation Archive (Sloboda 1999; see also Sheridan, 2000), respondents were asked to write, in open-ended fashion, about the uses they made of music in their everyday lives. Table 18.1 shows the activities and functions spontaneously mentioned by members of the sample. The most frequently mentioned activities were housework and travel. This mirrored quite closely the distribution found in the ESM study. Functions mentioned were varied but had a predominantly affective character, with many participants (particularly women) explicitly mentioning music as a mood changer or enhancer. The most frequently mentioned function was essentially nostalgic. Participants found it natural to link functions to activities, often mentioning both in the same sentence (e.g. on arrival home from work 'music lifts the stress of work: it has an immediate healing effect').

Table 18.1 Percentage of mass observation respondents reporting various functions and activities chosen for music

| Activities | |
|---|---|
| To wake up to | 6 |
| While having a bath | 4 |
| Whilst exercising | 4 |
| To sing along to | 6 |
| To work to (desk work) | 14 |
| To work to (housework) | 22 |
| On arrival home from work | 2 |
| Whilst having a meal | 12 |
| Background while socializing | 4 |
| To accompany sexual/romantic events | 4 |
| Whilst reading | 6 |
| In bed/to get to sleep | 14 |
| While driving/running/cycling | 22 |
| While on public transport (Walkman) | 4 |
| *Functions* | |
| Reminder of valued past event | 50 |
| Spiritual experience | 6 |
| Evokes visual images | 2 |
| Tingles/goose pimples/shivers | 10 |
| Source of pleasure/enjoyment | 6 |
| To put in a good mood | 16 |
| Moves to tears/catharsis/release | 14 |
| Excites | 2 |
| Motivates | 2 |
| Source of comfort/healing | 4 |
| Calms/soothes/relaxes/relieves stress | 8 |
| Mood enhancement | 8 |
| To match current mood | 6 |

In an interview study involving 52 women ranging in age from 18 to 78 years, most of whom were not accomplished musicians, DeNora (1999) also found examples of emotional self-regulation involving a number of musical strategies described by participants as 'revving up' or 'calming down', 'getting in the mood' (e.g. for a particular social event), 'getting out of a mood' (e.g. to improve a 'bad' mood or to 'destress'), or 'venting' strong emotions. For the most part, as in the two previously mentioned studies, these were predominantly described at the 'personal' or intrapersonal level as a means of creating, enhancing, sustaining, and changing subjective, cognitive, bodily, and self-conceptual states. According to DeNora, the women exhibited considerable awareness of the music they 'needed' to hear in different situations and at different times, often working as 'disc jockeys' to themselves:

> They drew upon elaborate repertoires of musical programming practice, and a sharp awareness of how to mobilize music to arrive at, enhance, and alter aspects of themselves and their self-concepts. This practical knowledge should be viewed as part of the (often tacit) practices in and through which respondents produced themselves as coherent social and socially disciplined beings (DeNora 1999, p. 35; see also DeNora, this volume).

In addition to the emotional functions at the intrapersonal level, DeNora found examples in which music played a social function in communicating emotions to others. For example, a university student described repeatedly playing at full volume a song from Radiohead entitled 'We Hope You Choke!' not only as way of diffusing her anger against her boyfriend's parents when she lived with them over the summer holidays, but also as a way of communicating her anger to them. In other words, the display of anger or irritation described by the respondent expressed a judgment of the moral quality of some other person's actions. Such a display was also an act of protest directed towards the boyfriend's parents. Music provided one means by which this display was 'acted out' at an interpersonal level. However, as DeNora points out, music is not simply used to express some internal, private feeling or state, nor does it simply 'act upon' individuals, like a stimulus. 'It is a resource for the identification work of 'knowing how one feels"—a building material of 'subjectivity"' (DeNora 1999, p. 41). In this way, music becomes part of the construction of the emotion itself through the way in which individuals orientate to it, interpret it, and use it to elaborate, 'fill in', or 'fill out', to themselves and others, an emotional feeling or display.

## 18.3 Autonomy and individuality

One of the questions in Sloboda, O'Neill, and Ivaldi's ESM study asked participants to rate the music in each episode on an 11-point scale according to the degree of personal choice exercised in hearing the music (from 0 = *none at all*, to 10 = *completely own choice*). There was a significant effect of degree of choice on the degree of emotional change experienced while listening to the music. For each emotion factor this showed a similar effect—the greater the choice the greater the change. We found that high choice situations were most likely to occur when the person was alone, travelling, or working, at home or in a vehicle, or undertaking activities for duty. Low choice situations occurred more often when with others, during active leisure or personal maintenance activities, in shops, gyms, and entertainment venues, and when doing activities because one wants to. Most of these findings are not surprising, although the link between choice and activities undertaken for duty is not intuitively obvious. It may be that choosing music to accompany duties is a way of bringing some autonomy and personalization back to them. DeNora (1999) suggests that the music associated with duties is used as a catalyst to shift individuals out of their reluctance to adopt what they perceive as 'necessary' modes of agency, and into modes of agency 'demanded' by particular circumstances.

The issue of choice and individuality also permeated the responses of the mass observation panel (Sloboda 1999). Many of the musical situations described were solitary, and some participants graphically characterized a difference between the private arena, where emotional work of one sort or another could be accomplished with the help of music, and the public arena, where self-presentation or the conflicting demands of others precluded this kind of activity. One respondent wrote: 'When I'm down I listen to this [a specific track] and go down as far as I can, then I cry, I cry deep from inside. I wallow in self-pity and purge all the gloom from my body. Then I dry my eyes, and wash my face, do my hair, put on fresh makeup, and rejoin the world'. Another

reported 'the car is the only place where I can listen to it loud enough without annoying other people'.

Many participants displayed negative, or at best, ambivalent reactions to the music that they experienced in public places, such as shops, restaurants, and bars. These attitudes were sometimes associated with reports of dramatic behavioural consequences with high emotional charge (e.g. abruptly leaving a shop with disliked music, arguing with waiting staff in a restaurant about getting the music turned off). In several responses, the appropriateness of the music was a major theme. Judgments were made both about the situation itself, and about the fit between the music and the person's own identity and preferences. For example, restaurant music could be acceptable so long as it matched the ambience or mood of the venue (e.g. oriental music in an oriental restaurant, 'mellow' music during a romantic encounter), or if its general acoustic characteristics matched the listeners' needs (e.g. not so loud that conversation was difficult, but loud enough to cover potentially embarrassing silences and to prevent people at nearby tables overhearing one's conversation). In other cases, although music in public places might be generally disliked, exceptions were made where the music itself had value for the participant (e.g. 'I have discovered Waterstones Bookshop in Newcastle plays good music: the last time I was there Beethoven's 'Egmont Overture' was playing, and the trouble is that I pay more attention to the music than to finding the book I want'). In this example, because the music was judged to be 'good', the participant was prepared to let his appreciation of the music partially override his main intention of buying a book. The issue of 'fit' between music and its context has been investigated experimentally. For instance, North and Hargreaves (2000) showed that people make consistent discriminatory judgments about the music that is suitable for such activities as aerobic exercise or yoga, and that specific characteristics of the music (e.g. tempo) are implicated in these judgments.

There is evidence that resistance to music in public places increases with age. DeNora (1999) found that most of her respondents over 70 years (and, interestingly, those who were trained musicians) found it antithetical to conceive of music as 'background' to anything. Data from the Sussex Mass Observation Archive also showed that males in the 40–60-year age group reported more negative emotional reactions to music in public places than any other group. It may be hypothesized that the higher average status a group has, the less tolerant it will be of removal of autonomy. However, there are also often more complex subcurrents to do with intergenerational and intergroup stereotypes. The following example reveals much about the attitudes of one middle-aged man to the public self-presentation of some younger men: 'I also dislike the din that some cars make when they pass by, infesting the streets with their thumping noises from within. How they can drive properly with such a din in their cars, God knows. It concerns me too that they wouldn't be able to hear the sound of an ambulance or police car with such a noise going on inside their cars. It seems to me, rather like the fastest drivers, it is usually young men between the age of 17 and 25 who are the main culprits. They usually like to have their driving-seat window open, elbow leaning out of the car and looking macho.' It appears as though young people are more tolerant of, and positive towards, music in public places. Further investigation is needed of these apparent age-related differences.

One phenomenon which deserves greater study and theorizing is the unique position of the busker (street musician) in the affections of even the most hardened opponents of music in public places. Many mass observation respondents spontaneously singled out buskers as evading the opprobrium of other forms of public music. It often was reported as relevant that this music is live rather than prerecorded. Two quotes are provided as representative: 'Both my husband and I have performed on the streets. My husband is still a regular busker. This to me is musical entertainment at its purest form. The joy of busking is its spontaneity. Your audience is free to come and go as it wishes, to pay or not pay, to listen or not listen. There is a beautiful freedom about busking that I love, and I hope we never lose street entertainment.' 'The music in the streets is acceptable. Quite often the musicians are quite good and anyway the noise is dispersed.' Arguably, street musicians do not undermine the sense of agency and autonomy that people like to experience. The musician is potentially amenable to interaction (you can request your favourite song, and audience reactions can be assumed to matter to the busker). There is also a sense of groundedness in a busker. The piped music in shops is produced by unknown people in unknown places, and mediated through hidden production mechanisms (e.g. an under-the-counter sound system), for hidden purposes (e.g. suspected manipulation of buying behaviour). The busker, by contrast, is earning an 'honest' living by aiming to please and entertain through the exercise of visible craft that is the result of personal effort and investment. It may also be significant that modern busking is probably the phenomenon which comes closest to meeting the conditions under which most people in most cultures through history have interacted with music. It is live, public, improvisational, spontaneous, participatory, and social. It creates a small arena of the communal in a pervading culture of individualism and isolation. It may meet an emotional need that is quite fundamental, but which there are few opportunities to indulge in.

## 18.4 Personal and social identity

In recent years, there has been a growing questioning of traditional approaches to the study of identity, which fail to take into account the multiple dimensions of identity as a continual process of negotiation and change. Discursive psychology, which views identity as constructed out of the discourses that are culturally available to us and that we use when communicating with others, has radically altered the way social psychologists conceptualize and study identity. Music provides numerous ways in which musical materials and practices can be used as a means for self-interpretation, self-presentation, and for the expression of emotional states associated with the self. According to DeNora (1999, p. 50), a sense of self is locatable in music, in that 'musical materials provide terms and templates for elaborating self-identity'. For example, one of the respondents in her study described a preferred type of musical material ('juicy chords') as like 'me in life', associating certain musical structures with her sense of self. DeNora (1999, p. 51) suggests that 'music is a "mirror" that allows one to "see one's self". It is, also, however, a "magic mirror" insofar as its specific material properties also come to configure (e.g. "transfigure", "disfigure", etc.) the image reflected in and through its (perceived) structures.'

Nowhere is this more apparent than during adolescence, where identities are being forged, experimented with, and explored. According to Green (1999), music can offer a powerful cultural symbol, which aids in adolescents' construction and presentation of self. Several studies suggest that musical tastes are predictive of a wide range of non-musical activities and attitudes, such as clothing, media preferences, drug use, and degree of sexual activity (Hanaken & Wells 1990; Lewis 1995).

Insofar as the processes involved with the formation and maintenance of adolescent identity are key to an individual's self-image and well-being (e.g. Shotter & Gergen 1989), then we would expect emotions to be deeply implicated in this process. Because most adolescents do not have jobs or family responsibilities they have more 'discretionary' time than many adults do (estimated as up to 50 per cent of waking hours by Csikszentmihalyi and Larson 1984). Much of this time is taken up with media use, mainly TV and music. Eighty-one per cent of young people say that music is an important part of their life, and has influenced how they think about important issues (Leming 1987). This is in direct contrast to TV watching, which most adolescents do not believe has had any major influence on their lives (McCormack 1984). In an experience-sampling study of adolescent daily lives, Larson *et al.* (1989) found that music listening was associated with greater personal involvement than TV watching (which was associated with feeling less happy, less alert, more passive, and more bored than at other times). Larson and Kleiber (1993, p. 130) summarize their findings as suggesting that:

> . . . the moving lyrics of ballads and the hard-driving beat of rock appear to stimulate a level of personal involvement that is lacking in TV watching. A teenager may be lying face-down on her bed, but her mind is alive and active, thinking about friends, school, or the future'.

In a recent study by North *et al.* (2000), marked gender differences were found in adolescents' reasons for listening to music. Girls were more likely to report that music could be used as a means of mood regulation, whereas boys reported that music could be a means of creating an external impression with others. It is perhaps not surprising that more girls than boys referred to emotionality in relation to music listening, given that 'masculine' identity tends to be associated with an ability to keep one's emotions out of one's reasoning. However, the boys' reason for listening to music was particularly interesting in that the majority of respondents reported that they usually listened to music on their own. Thus, the external impression the boys sought to accomplish did not necessarily involve direct social contact with others at the time they were actually listening to music. Rather, it appears as though adolescent boys, in particular, were actively involved in the construction of their identity through the use of stereotypes and gendered role models associated with the music they listened to (e.g. the 'macho' and 'sexy' image of male pop/rock musicians playing mainly guitars and drums). In other words, the cultural stereotypes, attitudes, and opinions associated with particular styles of music appears to provide adolescent males with a vision of what it means to be 'trendy' or 'cool', which is 'acted out' or reinforced during their music listening experiences. O'Neill (1997) points out that adolescents' musical values are influenced by culturally defined stereotypes that, once learned, are extremely resistant to change and disconfirmation. In this way, musical associations or 'prevailing discourses' in a culture are both established and maintained through their use as a 'thread' out of

which we might fashion or construct aspects of the self and the expression of emotional states associated with the self.

## 18.5 The balance between context and content

Emotional responses to music are a complex outcome of the contribution of a person's reaction to the content (i.e. the musical materials themselves and their associations) and their reactions to the social context in which the music is embedded. For instance, the context may determine the amount of attention available to place on the musical content (e.g. a parent out shopping with an irritable and loudly complaining small child and being embarrassed by the hostile looks of other shoppers may barely notice the 'soothing' Mozart playing quietly in the background). At the same time, the content may be subtly changing the way the person construes the context (e.g. the Mozart excerpt suggests the subliminal message that this is a 'high-class' store for serious, 'well-behaved' adults, and small noisy children have no place here—thus intensifying the feelings in the parent of embarrassment and shame). These effects have been studied experimentally by, for instance, manipulating music–film pairings or music–picture pairings and observing how construed emotional meaning is affected (Cohen, this volume).

However, everyday life does not provide opportunities for controlled manipulation of variables, and so it is difficult to disentangle the unique contribution of content and context. The work that comes the closest to providing an insight into these issues is a study by Sloboda (1989*a*) in which 70 individuals were asked to recall any incidents from the first 10 years of life that were in any way connected with music. This period was chosen because a major aim of the study was to find connections between early music experience and later attitudes to music. For each incident recalled, participants were asked to say as much as they could about the context (who they were with, what event the music formed part of), and what meaning or significance the event had for them. This allowed each incident to be assigned a value on each of two dimensions—an internal dimension, concerned with the musical content, and an external dimension concerned with the context. On each dimension the significance could be positive, neutral, or negative.

When the 113 incidents elicited were cross-tabulated on these two dimensions, some interesting features emerged. First, there were very few cases where the musical content had a negative significance. One person recalled disliking the sound of a particular set of pipes on a church organ, but there were few other clear cases. This was not, however, the result of a rosy view of childhood where all negative memories had been erased, because there were many cases where the musical context was negative. Many people remembered situations of anxiety, pain, and humiliation, mainly connected with negative appraisal of their musicality by adult authority figures (particularly teachers). The most extreme incident was where a teacher physically beat a child for a performance error. There were, however, positive contexts too. For instance, one respondent recalled rehearsing carols for a school carol concert. What gave the incident its positive significance was the sense of enjoyment among the group, undertaking this festive preparation in place of normal academic classes.

The second important feature was the almost complete absence of incidents where the musical content was appraised positively and the context was appraised negatively.

One might interpret this finding by supposing that where the immediate context is a source of threat or emotional challenge, there is little chance of musical content capturing the emotional system. This study suggests that positive emotions derived from engagement with musical materials is only possible when the context is appraised as, at worst, emotionally neutral. The specific contexts where strong content-related emotions were felt included home, church, and the concert hall, alone or with friends and family. They tended not to include lessons at school, or situations in the direct presence of a teacher. Although hopefully not inevitably so, it seems that formal instructional settings have a tendency to be inimical to emotional engagement with music. This may be because of the emphasis within such settings on achievement, success, and failure, with the concomitant threats to self-esteem and self-worth. It has the paradoxical consequence that people may be driven to express their deepest and most personal relationships with music in private, or in supportive peer reference groups outside, and hidden from, the formal educational process.

## 18.6 Everyday music use as self-therapy

In the preceding paragraphs, we have seen much evidence of music being used deliberately and consciously to achieve psychological outcomes which are reflected in emotional change. Music clearly is used by people to make them feel better or different, or help them accomplish or attune themselves to some concurrent or anticipated activity. These kinds of activity have sometimes been, perhaps rather glibly, cited as examples of 'self-therapy' (Sloboda 1989*b*, 1992). However, therapy, as understood by therapists, is not simply about manipulation of emotions, it is about helping the individual in therapy to develop more appropriate and functional responses to the problems of living (cf. Bayne & Nicolson 1993; Bunt & Pavlicevic, this volume). There is almost nothing in the literature on everyday uses of music that would count as strong evidence that particular self-chosen music improves problem solving or decision making, as compared to no music, or different music. It would be an interesting study to ask people to deny themselves any self-chosen music for a period of time, and compare their adjustment on a number of psychological measures to that found during normal music use. ESM would seem to be particularly well suited to the measurement of the effects of such manipulations.

However, the strong claims made by users of music in a wide range of studies, supported by subjective and anecdotal evidence, make it likely that such effects do exist, although the mechanisms by which they are mediated are poorly theorized. There is, however, one very plausible class of mechanisms by which music could have a therapeutic effect. Much is made in the literature about 'everyday' psychological disorders, such as depression or the effect of cognitive set or narrowing. A depressed person is often locked in a cycle of negative and self-defeating cognitions, unable to call to mind plausible alternatives to the narrow circle of linked aversive scenarios (Blaney 1986; Pollock & Williams 1998). These cognitions are also often accompanied by anomalous states of arousal, such as high anxiety, insomnia, or lethargy. Well-chosen pieces of music may be able to help individuals break out of such cycles by the specific combination of intrinsic and extrinsic cues that they provide (Sloboda & Juslin, this volume).

Extrinsic cues may remind the person of situations, scenarios, people, and emotions that lie outside the closed loop of the pathological state. Intrinsic cues (the ebb and flow of tension, resolution, expectancy, etc.) may provide means for altering arousal states in positive directions. In addition, the unique capacity of music to engender emotional release (as in crying; see Sloboda & Juslin, this volume) may be in itself therapeutic. For reasons that are not well understood, emotional release appears to assist psychological adjustment (through a process sometimes called *catharsis*; e.g. Davis 1988). A fuller working out of some of these ideas in the context of music's role in spirituality is provided by Sloboda (2000).

## 18.7 Conclusion

It is a significant feature of many of the emotional feelings and displays that individuals experience in relation to the everyday musical scenarios that we have outlined that although they may occur in solitude, their point of reference is the relationship between the music user and others. Although viewed as essentially 'private' experiences, involving a great deal of autonomy or agency, emotional feelings and displays are deeply embedded in a social context, which exerts a powerful influence (albeit often implicitly) on our music listening. Reliving past relationships, constructing identity, using music to 'siphon off' emotions that are not for public presentation: all of these depend on, and are used to negotiate and develop, the complex web of cognitions and behaviours that constitute social life. As such, music becomes part of the construction of emotional feelings and displays that are both reflective and communicative 'embodied' judgments used to accomplish particular social acts. In other words, musical emotions are a form of social representation, which is negotiated as an interaction between cultural/ideological values of a society, the values and beliefs operating in a social grouping or subculture in that society, and the individual's own social and personal experience.

Our research highlights the importance of studying emotional feelings and displays as part of a sequence of everyday music listening based on conventions and rules of a particular social context and the unfolding episodes in which they occur. Our recent study (Sloboda *et al.*, 2001) involving electronic pagers and self-reports (experience sampling method) provides one approach which may be used to capture individuals' subjective experiences of emotions during 'real', evolving musical episodes. Another possibility is to approach the analysis of episodes involving music and emotion in a similar way to how we might analyse conversations in discursive psychology. This is not to say that we should analyse the actual vocabulary used by individuals to define emotion words in relation to music. Rather, a 'discourse' in this sense can be thought of as a kind of frame of reference, a conceptual framework in which our emotional feelings and displays can be interpreted. One method of discursive analysis which may prove particularly useful in the study of music and emotions is what has been referred to as the 'positioning triad' (cf. Harré & van Langenhove 1999). It consists of three elements: (a) the story (or content) of an evolving episode; (b) the relative positions of individuals in terms of the cultural/social conventions involved in speaking (and listening) during or about the episode, displaying judgments, and/or expressing acts; and (c) the social act the episode performs. The 'positioning triad' provides a framework for investigating the dynamics

of real-life, evolving situations that could be used by researchers seeking to identify and further our understanding of the role of music and the people engaging in emotional displays.

However, a word of caution may be necessary. Most of us are not aware of the fact that our musical activities are completely enmeshed in a social and cultural world. Our engagement with music leads us to 'forget' or become unaware of the grounds on which our feelings and behaviours are based. This 'forgetting' is the product of years of training, socialization, and the institutionalization of music. Not only have our musical practices become routine and invisible, but as musicians and psychologists we are limited in our ability to describe musical materials in a way that is free of the assumptions and biases associated with our own experiences and training. For example, there are many examples of research where specific discourses involving musicological characterizations of emotions (e.g. reference to specific musical works or to structural and symbolic features of these works) are introduced before an attempt is made to characterize the user's emotional responses in relation to them. In doing so, we continue to perpetuate the historical projections and social preoccupations of the various professional elites that have dominated the work of defining what music is and what it is for, not only for themselves but for the wider culture. By turning our attention to the role of emotional 'work' in relation to music, which takes into account individual subjective experience of *both* emotional feelings and displays as they occur in everyday evolving situations, it might be possible to retrieve some of these 'forgotten' or 'hidden' practices. Only by revealing the meanings associated with our evaluative judgments embedded within a particular context or general conceptual framework, can we hope to gain a better understanding of the complexity of emotions and music in everyday life.

## References

**Bayne, R. & Nicolson, P.** (1993). *Counselling and psychology for health professionals.* London: Chapman & Hall.

**Blaney, P. H.** (1986). Affect and memory: a review. *Psychological Bulletin,* 99, 229–46.

**Brodsky, W.** (1995). *Career stress and performance anxiety in professional orchestra musicians: A study of individual differences and their impact on therapeutic outcomes.* Unpublished doctoral dissertation, University of Keele, England.

**Burr, V.** (1995). *An introduction to social constructionism.* London: Routledge.

**Csikszentmihalyi, M. & Larson, R.** (1984). *Being adolescent.* New York: Basic Books.

**Csikszentmihalyi, M. & Lefevre, J.** (1989). Optimal experience in work and leisure. *Journal of Personality and Social Psychology,* 56, 815–22.

**Davis, P. J.** (1988). Physiological and subjective effects of catharsis: a case report. *Cognition and Emotion,* 2, 19–28.

**DeNora, T.** (1999). Music as a technology of the self. *Poetics,* 27, 31–56.

**Gabrielsson, A. & Lindström, S.** (1995). Can strong experiences of music have therapeutic implications? In *Music and the mind machine: The psychophysiology and psychopathology of the sense of music,* (ed. R. Steinberg), pp. 195–202. Berlin: Springer.

Green, L. (1999). Research in the sociology of music education: some introductory concepts. *Music Education Research,* 1, 159–69.

Hanaken, E. A. & Wells, A. (1990). Adolescent music marginals: Who likes metal, jazz, country, and classical? *Popular Music and Society,* 14, 57–66.

Harré, R. & Gillett, G. (1994). *The discursive mind.* London: Sage Publications.

Harré, R. & van Langenhove, L. (ed.) (1999). *Positioning theory.* Oxford, UK: Blackwell.

Larsen, R. & Kleiber, D. (1993). Daily experience of adolescents. In *Handbook of clinical research and practice with adolescents,* (ed. P. Tolan & B. Cohler), pp.125–45. New York: Wiley.

Larsen, R., Kubey, R., & Colletti, J. (1989). Changing channels: Early adolescent media choices and shifting investments in family and friends. *Journal of Youth and Adolescence,* 18, 583–600.

Leming, J. (1987). Rock music and the socialisation of moral values in early adolescence. *Youth and Society,* 18, 363–83.

Lewis, G. H. (1995) Taste cultures and musical stereotypes: Mirrors of identity? *Popular Music and Society,* 19, 37–72.

Lutz, C. (1982). The domain of emotion words on Ifaluk. *American Ethnologist,* 9, 113–28.

McCormack, J. (1984). *Formative life experiences and the channelling of adolescent goals.* Unpublished doctoral dissertation, University of Chicago, Chicago.

Merriam, A. (1964). *The anthropology of music.* Chicago: Northwestern University Press.

North, A. C. & Hargreaves, D. J. (2000). Musical preferences during and after relaxation and exercise. *American Journal of Psychology,* 113, 43–67.

North, A. C., Hargreaves, D. J., & O'Neill, S. A. (2000). The importance of music to adolescents. *British Journal of Educational Psychology,* 70, 255–72.

O'Neill, S. A. (1997). Gender and music. In *The social psychology of music,* (ed. D. J. Hargreaves & A. C. North), pp. 46–63. Oxford, UK: Oxford University Press.

Pollock, L. R. & Williams, J. M. G. (1998). Problem-solving and suicidal behaviour. *Suicide and Life-threatening Behaviour,* 28, 375–87.

Sheridan, D. (2000). Mass-observation revived: The Thatcher years and after. In *Writing ourselves: Mass-observation and literacy practices,* (ed. D. Sheridan, B. Street, & D. Bloome), pp. 43–78. Cresskill, NJ: Hampton Press.

Shotter, J. & Gergen, K. J. (1989). *Texts of identity.* London: Sage Publications.

Sloboda, J. A. (1989*a*). Music as a language. In *Music and child development* (ed. F. Wilson & F. Roehmann), pp. 28–43. St. Louis, Missouri: MMB Music.

Sloboda, J. A. (1989*b*). Music psychology and the composer. In *Structure and perception of electroacoustic sound and music,* (ed. S. Nielzen & O. Olsson), pp. 3–12. Amsterdam: Elsevier.

Sloboda, J. A. (1992). Empirical studies of emotional response to music. In *Cognitive bases of musical communication,* (ed. M. Riess Jones & S. Holleran), pp. 33–50. Washington, DC: American Psychological Association.

Sloboda, J. A. (1999). Everyday uses of music listening: A preliminary study. In *Music, mind and science* (ed. S. W. Yi), pp. 354–69. Seoul, Korea: Western Music Research Institute.

Sloboda, J. A. (2000). Music and worship: A psychologist's perspective. In *Creative chords: Studies in music, theology, and Christian learning,* (ed. J. Astley, T. Hone, & M. Savage), pp. 110–25. Leominster: Gracewing.

Sloboda, J. A., O'Neill, S. A., & Ivaldi, A. (2001). *Functions of music in everyday life:* An exploratory study using the experience sampling method. *Musicae Scientiqe,* 5, 9–32.

CHAPTER 19

# EMOTIONS IN STRONG EXPERIENCES WITH MUSIC

ALF GABRIELSSON

In most people's minds there is a strong association between music and emotions. Music can be felt to express emotions as well as to arouse emotions. While musical expression has been given considerable attention in empirical research (cf. Gabrielsson & Lindström, this volume; Juslin, this volume), emotional reactions to music have been less studied. The purpose of this chapter is to describe emotional reactions that occur in particularly strong experiences of music, and further to explore which factors can elicit such reactions.

Strong experiences of music may have many properties in common with other types of strong or exceptional experience, for instance, mystical experience which is described and discussed in many religious contexts. In the following, I first briefly review some earlier research on strong experiences, and then describe a research project on strong experiences of music (Section 19.3) and the associated analysis of emotional reactions (Section 19.4), followed by a discussion (Section 19.5).

## 19.1 Earlier research on strong experiences

In empirical psychology, the best known analysis of strong experiences was furnished by Abraham Maslow, one of the founders of the so-called humanistic psychology. Maslow coined the term *peak experience,* which he originally associated with self-actualization, the highest level in his well-known hierarchy of needs (Maslow 1954, Chapter 5). To explore such experiences, Maslow asked people to describe 'the most wonderful experience of your life; happiest moments, ecstatic moments, moments of rapture, perhaps from being in love, or from listening to music, or suddenly 'being hit' by a book or a painting, or from some great creative moment' (Maslow 1968, p. 71). Surveying the contents of these descriptions, he found several characteristics of generalized peak experience, for example, total attention on the object in question, complete absorption, disorientation in time and space, transcendence of ego, and identification or even fusion of the perceiver and the perceived. Peak experience is good and desirable, there is a complete loss of fear, anxiety, inhibition, defence, and control. Moreover, 'the emotional reaction in the peak experience has a special flavor of wonder, of awe, of reverence, of humility and surrender before the experience as before something great' (Maslow 1968, pp. 87–8), even a fear of being overwhelmed with more than one can bear. The experience may occasionally be described as sacred. Maslow found music,

especially the great classics, to be one of the easiest ways of getting peak experience (Maslow 1976, pp. 169–70).

As an extension of Maslow's investigations, Panzarella (1980, p. 71) gathered reports on 'intense joyous experience of listening to music or looking at visual art' from a sample of 103 persons, 51 describing musical experiences and 52 visual art experiences. His analysis (content analysis followed by factor analysis) revealed four major factors of the experience: (a) *renewal ecstasy*, an altered perception of the world ('The world is better, more beautiful than had been thought before'; Panzarella 1980, p. 73); (b) *motor-sensory ecstasy*, physical responses (changes in heart rate, breathing, posture or locomotion, presence of shivers, chills, tinglings, etc.) and quasi-physical responses (e.g. feeling 'high', 'floating'); (c) *withdrawal ecstasy*, loss of contact with both the physical and social environment; and (d) *fusion-emotional ecstasy*, merging with the aesthetic object. Motor-sensory ecstasy and fusion-emotional ecstasy were more pronounced in music reports than in visual art reports.

Earlier, Laski (1961) asked 63 individuals whether they had known an experience of 'transcendent ecstasy' and further analysed texts from selected literary sources and from books on religious experiences. Triggers of such experiences included nature, sexual love, childbirth, religion, art, science, creative work, 'beauty', and others. The contents of the descriptions were classified into four broad categories: (a) feelings of loss of something (e.g. loss of time, place, sense, self); (b) feelings of gain of something (e.g. gain of timelessness, release, satisfaction, joy, salvation, perfection, new knowledge); (c) feelings of ineffability (e.g. the experience is indescribable, eludes verbal communication); and (d) quasi-physical feelings (e.g. light and/or heat words, improvement words, pain words, calm and peace words). Classical music was the most frequently mentioned trigger among the arts.

The concept of *flow* developed by Csikszentmihalyi (e.g. Csikszentmihalyi 1990) has several aspects similar to peak experience and self-actualization. It refers to a state of intense yet effortless involvement in an activity, the experience of which is 'so enjoyable that people will do it . . . for the sheer sake of doing it' (Csikszentmihalyi 1990, p. 4). Some major characteristics of the phenomenon are full concentration on the relevant stimuli, total absorption in the activity; altered perception of time, and loss of self-consciousness. Flow may appear in connection with a manifold of activities, for example, rock climbing, sailing, chess, games, dancing, and making music. With regard to music, the flow concept seems most applicable to performance (see Persson, this volume), for instance, in mastering a difficult piece in a fully concentrated but seemingly effortless involvement, a phenomenon often reported by musicians as the most happy moments in their activities (I agree). Flow experience in listening to music is briefly mentioned as well (Csikszentmihalyi 1990, pp. 68 and 108).

Evidence from the above-mentioned studies indicate many different physical, quasi-physical, perceptual, cognitive, emotional, and transcendental phenomena in strong experiences, whether they are called peak experiences, mystic experiences, ecstasy, flow, or something else. Emotions are important ingredients and may manifest themselves, not only as experience, but also in physical responses (e.g. Plutchik 1994, p. 136).

## 19.2 Physical responses

Physical responses, such as thrills, shivers, and changes in heart rate are often mentioned in descriptions of strong experiences and are commonly associated with strong emotional arousal. From answers to a questionnaire, Goldstein (1980) concluded that thrills—a chill, shudder, tingling, or tickling—are a common response to music, primarily to musical passages with special emotional meaning for a person.

Sloboda (1991) obtained reports from 83 respondents on various physical responses—shivers, tears, lump in the throat, goose pimples, racing heart, and others—that they often experienced to music. In about a third of the cited cases, the response was located to a specific theme, phrase, motif, bar, chord, or moment, all of them belonging to pieces of classical music. Tears or lump in the throat were provoked by melodic appogiaturas, shivers or goose pimples by sudden changes in harmony, and heart race occurred in connection with acceleration and syncopation (see also comments in Sloboda 1999).

Panksepp (1995) investigated chills induced by music. In general, chills were reported more for pieces brought by the participating students themselves than for other pieces, which suggested effects of conditioning. Most examples belonged to popular music. When subjects listened to unfamiliar music, sad pieces produced significantly more chills than happy pieces for females but not for males. Chills tended to peak at the most intense and dramatic crescendos.

## 19.3 The SEM project

Evidence from the studies reviewed above suggests that music may be a common trigger of extraordinary experiences. In the following, I briefly describe a research project aiming to gather further, more detailed knowledge of this topic.

The strong experiences of music (SEM) project was initiated independently of the works described above. The main impulse was my disappointment that music psychology rarely, if ever, comes to grips with the questions that most people consider the most relevant regarding music—simply, how are we affected by music? I was also approached, typically in private, by people who told me about exceptional experiences of music, affecting them deeply in ways that they could not understand.

I started this project at the very end of the 1980s together with psychologist Siv Lindström Wik. We studied Maslow's writings on peak experiences, but found no concrete examples of what people told about their peak experiences, rather 'an impressionistic, ideal, 'composite photograph' or organization of personal interviews . . . and of written reports' (Maslow 1968, p. 71). Panzarella's (1980) pioneering study was very useful as it provided examples of subjects' reports, and procedures for obtaining and analysing them. However, his musical sample was limited to some 50 persons, and in both Maslow's and Panzarella's studies classical music seemed to be the (almost) sole elicitor of peak experiences. This was also stated by Laski (1961). We were also concerned about definitions such as 'intense joyous experience' (Panzarella 1980, p. 71). All this seemed to imply some limitations in relation to the purpose of the SEM project, that is, to obtain descriptions of any type of strong experiences (including negative

ones) in connection with any kind of music. In order not to prejudice our potential subjects, we avoided any examples or suggestions of what would define a strong experience of music or of what kind of music could be involved.

The purpose of the project was first of all descriptive, to analyse which phenomena are contained in strong experiences of music—in other words, to find out how people are affected. We also wanted to explore how SEM is related to properties of the actual music as well as to characteristics of the respondent and of the situation in which the experience occurred. Moreover, we were interested in possible consequences of SEM for the individual's future.

The very beginning of the project was briefly reported in Gabrielsson (1989). A classification of SEM phenomena was sketched in Gabrielsson and Lindström (1993), and further elaborated in Gabrielsson and Lindström Wik (2000). Examples of the potential therapeutic effects of SEM were given in Gabrielsson and Lindström (1995); and SEM in relation to musicality is discussed in Gabrielsson (in press). Given the present volume's focus on emotion, the following chapter will be centred on the description of emotional aspects of SEM. Some connections to other aspects and to underlying factors will be briefly discussed within the available space.

### 19.3.1 Data collection

Briefly, subjects were asked to describe 'the strongest, most intense experience of music that you have ever had. Please, describe your experience and reactions in as much detail as you can.' Supplementary questions concerned, for example, if this was the first time the subject had listened to the music in question and if the same strong experience had recurred at later listenings, if any, to the same music; how the respondent felt before and after the experience and what the experience had meant in a longer perspective; how often such strong experiences occurred and if similar strong experiences occurred in other situations than with music; and if the respondent had any idea about what caused the strong experience. The reports were obtained by means of interviews and, mostly, as written reports.

### 19.3.2 Subjects

All participation was on a voluntary basis, and there was no attempt to achieve strict representative sampling of some defined subject population. However, attempts were made to ensure variation with regard to gender, age, occupation, and musical preferences. The present account is based on about 400 reports obtained from some 300 persons (many subjects provided two or more reports). Gender distribution is about equal. Almost 60 per cent of participants are between 20 and 40 years old, about 30 per cent between 40 and 60 years old. About 50 per cent are amateur musicians, about one-third professional musicians, and about one-sixth non-musicians. Although musicians, professional and amateurs, are thus in the great majority, about 80 per cent of the reports refer to experiences as listeners. Musical preferences are spread over many genres.

### 19.3.3 Analysis

Subjects' SEM accounts are subjected to content analysis (see, for example, Patton 1990, Chapter 8), gathering all relevant phenomena mentioned and successively exploring

different ways of sorting them into a limited number of fundamental categories. This work was/is conducted independently by the author and his co-worker, also at times by students in connection with writing theses on SEM. After the necessarily tentative analysis of the earliest reports, the agreement between the analysts has usually been very good. However, as human experiences and reactions are hard to fit into truly exhaustive and mutually exclusive categories, there are still some problems regarding proper classification. Some double classifications seem unavoidable, and the classification process may never come to a definite end. A basic problem is, of course, the large interindividual variation in participants' vocabulary and linguistic competence. We stay as close as possible to the descriptive terms used by the subjects themselves, but a certain amount of interpretation from our side is unavoidable. For more complete accounts of methods, see Gabrielsson and Lindström (1993) and Gabrielsson and Lindström Wik (2000).

### 19.3.4 Some general results

The length and detail of the reports vary widely, from only a few lines up to eight pages. A few respondents even wrote poetry to describe their experience. Each report represents a unique case with regard to contents of the experience and influencing factors.

Briefly, the present classification of SEM aspects comprises the following basic categories: general characteristics; physical and behavioural aspects; perceptual aspects; cognitive aspects; emotional aspects; existential and transcendental aspects; and developmental aspects. Each of them contains a number of subcategories as seen in Table 19.1. For discussion and examples of these subcategories, see Gabrielsson and Lindström (1993) and Gabrielsson and Lindström Wik (2000). As noted earlier, we do not pretend that these categories are completely exhaustive, nor quite mutually exclusive. However, repeated application of this classification to lots of SEM reports shows good results. The number of doubtful cases is relatively small and decreases as the classification is successively refined. The classification has also proved useful in pilot studies on strong experiences of dance and of visual art. The present account will be limited to emotional aspects of SEM and related matters.

## 19.4 Emotional aspects of SEM

The emotional aspects of SEM have been classified into four subcategories: (a) intense emotions; (b) positive emotions; (c) negative emotions; and (d) mixtures of emotions/conflicting emotions.

### 19.4.1 Intense emotions

As could be expected, emotional reactions in SEM are typically of unusual intensity, regardless of which specific emotions are experienced. Some illustrative quotations are given here:

I had to close my eyes, it was such an enormous emotional experience.

**Table 19.1** Classification of SEM phenomena

| |
|---|
| *General characteristics of SEM* |
| Unique, fantastic, incredible |
| Hard-to-describe experience, words insufficient |
| *Physical reactions, behaviours* |
| Physiological responses |
| Behaviours |
| Quasi-physical responses |
| *Perception* |
| Auditory |
| Visual |
| Tactile |
| Kinaesthetic |
| Synaesthetic |
| Intensified perception |
| Musical perception-cognition |
| *Cognition* |
| Changed attitude |
| Changed experience of body, time, and space |
| Changed attitude to music |
| Thoughts, memories |
| Imagery |
| Cognition-emotion |
| Musical cognition-emotion |
| *Emotion* |
| Intense emotions |
| Positive emotions |
| Negative emotions |
| Mixtures of emotions/conflicting emotions |
| *Existential and transcendental aspects* |
| Meaning of existence, being |
| Transcendental states |
| Religious experiences |
| *Personal development* |
| New insights, new possibilities |
| Confirmation of identity |
| Feeling of community |
| Music: new interests, attitudes, motives |

. . . overwhelming waves of feelings that were thrown forth as reaction to the music . . . I could absolutely not control the emotional outburst . . . so ABSOLUTELY IMPOSSIBLE to change it.

This song arouses so strong feelings in me that I cannot listen to it more than once a year.

. . . and last an internal peace that words cannot describe.

Although intensity of emotions may seem mainly associated with high arousal, as in the second quotation above, it may as well refer to states of low arousal such as peacefulness (last quotation).

### 19.4.2 Positive emotions

Positive emotions dominate in SEM reports. The by far most frequently mentioned feelings, singly or in various combinations, are happiness, joy, elation, and bliss. In many cases, the feelings are described in even stronger terms, such as a kind of intoxication, rapture, or euphoria boarding on ecstasy. Below are some illustrative excerpts from SEM reports. They are designated E1, E2, etc. to facilitate finding them again when referred to later in text.

A young lady attended an outdoor concert with the artist known as Prince:

[E1] The music began before the curtain rose, and you just stood there as semi-paralysed and screaming . . . Everybody in the audience is exciting each other to a stage next to a climax, and when the artist at last comes on stage he does not have to say more than 'hi' to trigger that climax. It is very much the atmosphere in the audience that gives this concert feeling. If I was standing there all by myself looking at Prince, it would not be the same thing at all . . . One feels so free somehow. At concerts one can dance, jump, scream and sing as much as one wants. You are like a part of it all, not just a spectator. Throughout the whole concert the audience was in total ecstasy. It was only one thing that mattered: the music! You are filled with the wonderful feeling that everything is all right, that you are living only for what is happening 'right now'. You are in a way 'intoxicated with joy' . . . You don't think about what you are doing. You do what you feel like, without even thinking about it.

A young performer was rehearsing a piece by Bach and then:

[E2] . . . suddenly experienced a tremendously strong feeling that was felt both in my body and in my head. It was as if I was charged with some kind of high tension, like a strong intoxication. It made me ecstatic, inconceivably exhilarated, everything concentrated to a single now. The music flowed as by itself. I felt like I was penetrated by Bach's spirit: the music was suddenly so self-evident. There was no doubt any longer how it should be played, as if I had come to the deepest insight and found the genuine, true and correct expression. The intoxication lasted the whole piece, and I staggered out afterwards . . . What a fantastic happy experience!

A woman attended a performance by a band playing Finnish tango in a pub:

[E3] I was filled by a feeling that the music started to take command of my body. I was charged in some way . . . I was filled by an enormous warmth and heat. I swallowed all tones . . . the music became so distinct. I was captured by each of the instruments and what they had to give me . . . Nothing else existed. I was dancing, whirling, giving myself up to the music and the rhythms, overjoyed, laughing. Tears came into my eyes—however strange it may seem—and it was as a kind of liberation . . . Afterwards I was standing there intoxicated with joy. That was a real kick for me.

There are lots of reports of this kind in which happiness and joy abound. In other cases a 'happy end' is attained although actually nothing indicated such a possibility—as in this report by a depressed mother, alone at home with her little child. She felt:

[E4] . . . so tired, so alone . . . grey sky, cold . . . nobody would know if I was to die . . . For some reason, I put on Mozart's *Requiem* (I am not aware of logically thinking in any way about this act) . . . I read the sleeve notes and sink, sink, sink deeper—no time—only the sounds enveloping me in a world aside of this—sinking deeper beyond words, beyond thoughts, only the waterfall of sound cascading around and through me . . . My head is gone, body too, and yet I am here. The grey 4 square table leg is exquisite in its perfection—beyond meaning in the intensity of its

greyness, displays a perfection that defies description . . . the whole room, the whole universe is present here in perfect harmony. All colors are vivid beyond description . . .I seem to be experiencing directly and have no words to express this state of bliss, harmony, meaningless beauty. I am totally a part of it all . . . The whole world is dancing before my eyes—I am calm, tears run down my cheeks. I have no idea why . . . I am not happy, I am not sad, am not anything in particular, yet feel the potential for anything to pass through me . . . secure and warm and full of love. I want to kiss the postman, ring the whole world and tell it, it's okay, it's all okay . . . the total qualitative shift in my relating to the world, universe, whatever you can now call it, remains. That pure existence and nothing else could be so beautiful! Satisfaction. Ecstasy.

This is a good example of what Panzarella (1980) named 'renewal ecstasy'. Other commonly mentioned feelings are enjoyment, delight, loveliness, and beauty; for instance, 'I felt an enormous delight' or '. . . it was lovely. It was God and heaven and all angels'. Closely related to this type of responses are also feelings of perfection: 'It can't be better than this!'; 'I remember the feeling of "this is not true, I am dreaming, it is not possible to do like this"'; 'When I came out after the concert, I told my company that I could die happy now'. Performers may feel that they surpass their usual achievements, and manage things that they usually do not master: 'This feeling when everything fits together is wonderful'; 'The music one created was far beyond one's usual capacity'.

While the emotional reactions described so far are associated with high or rather high levels of arousal, another common class of responses reflects feelings at rather low arousal, such as feelings of calm, peace, repose, safety, warmth, and love: 'I experienced a stillness that was not like anything else . . . everything else felt irrelevant . . . the important thing was to be able to exist in the present and the stillness that felt as eternal'; 'I listened to the tape 2–3 times a day . . . it became a safety among all chaos and misunderstandings'; 'I was in a state of warmth and kindness'; 'I felt in love with this music'.

Many feelings seem to represent a mixture of cognitive and emotional aspects. They are contained within the cognition-emotion subcategory in Table 19.1. Feelings of expectancy and tension may occur both before and during the music: 'It is always thrilling to sit down and know that the orchestra is going to perform'; '. . . building up an almost insufferable tension concerning what this would lead up to'. Music may make the listener surprised, amazed, shaken, even shocked (examples of the latter appear under negative emotions): 'I was brought away by the music, amazed and surprised by its beauty'. A common phrase, almost a cliché, is to 'be hit' by the music, feel that it goes straight into oneself, bypassing or crossing all cognitive barriers: 'A pianist played Beethoven's *Appassionata*. It hit the poor, musical child who had neither radio or gramophone at home'; 'The monumental work . . . hit me like a storm'; '. . . the piece goes so directly into me. Quite mercilessly it cuts through all barriers straight into the emotions'.

Still other examples within this subcategory may be feelings of longing ('The music brought me away in dreams, longing'), wonder ('The feeling when you experience this totality is just wonder'), smallness and humility ('It was an enormous experience, I felt small and humble'), gratitude ('I felt so infinitely grateful'), honour ('It was . . . almost an honour having had the possibility to experience this'), solemnity ('I remember the solemn feeling of being present at such an evening'), and sacredness ('I felt myself standing in the midst of a sacred moment in life').

### 19.4.3 Negative emotions

A few extremely negative SEMs were reported. A young man attended an outdoor rock concert:

[E5] After a while, they announced a female's band . . . Malaria was their symptomatic name. They played a hard, heavy, fateful synthetic rock at high volume. We had been sitting quite close but it soon became unbearable. The base was roaring a heavy, deafening monotone sound, that penetrated one's body. The drummer was also beating heavily and tired . . . a saxophone screamed in anguish, completely free from melody . . . accompanied by a singer who completely freely sang the most monotonous and howling song I had ever heard. Their make-up was black and pale, they wore black clothes . . . This naked and gruesome death music that was flowing from the loudspeakers and drowning us made me feel physically ill. I just wanted to get away from the music, and got on my feet to walk away. At the same time whole groups of black-dressed guys and girls came flowing toward the stage and loudspeakers. Everybody looked so pale and listless without life. I tried to remain but felt that it only provoked a destructive feeling of sickness . . . I thought that the Malaria band and the audience, in open anxiety before life and future, allied themselves, horribly and totally, with death and destruction . . . It was a deeply negative experience, strong and gruesome.

One report refers to a special chord in Mahler's Tenth Symphony. A man sat listening to this music shortly after a beloved relative's death:

[E6] But there it was. A chord so heart-rending and ghost-ridden that I had never experienced before. A single tone . . . is added with an endless number of instruments from the orchestra. Not unlike a huge organ where you pull out every organ stop at random. A dissonance that pierced my very marrow. My brother and I reacted the same: we were both filled with such a primitive horror, almost prehistorical, that none of us could utter a single word. We both looked at the big black window and both of us seemed to see the face of Death staring at us from outside.

While the negative emotions in these reports seem to be mainly caused by the music, in other cases the negative experience is associated with various personal and/or situational factors. A female alone at home in a large house a dark evening told:

[E7] I was sitting reading with the radio on. Suddenly Stravinsky's *Rite of Spring* breaks loose without notice. I almost panicked and felt an intense discomfort and had to switch off the radio.

Another negative experience was related to a frightful experience during childhood:

[E8] The Stockholm Philharmonic Orchestra was playing *Finlandia* by Jean Sibelius. The then aged composer was sitting in an armchair a little way from the orchestra. I felt immensely split—not to say that I was in a kind of chaos that would not pass off. The reason was that the work was said to describe the varying nature of Finland, barren landscape, wind and storm—and nothing else. The first time I listened to *Finlandia* was on radio in August 1940 . . . while the war was raging. I was a child living with my grandmother, separated from my parents and did not know when I would meet them again. I was very scared by the piece because I immediately interpreted the trumpet blows as machine guns, bomb explosions and the like. I felt anguish.

Some persons reported strong feelings of sadness associated with a piece of music. Usually the music had originally occurred in connection with negative circumstances, such as unhappy love, illness, attempted suicide, death, or memories of war. Strong frustration was felt by a female in the midst of an SEM when 'the grand mother steps in and

turns off the radio telling 'You need not wear out the batteries for such noise'. It felt like she tried to stifle me, I got quite frustrated.' One subject felt anger at a funeral, because quite inappropriate music was played. There was a single report of shame; a child was so affected by the music that 'I started crying by emotion. But I felt ashamed because of this . . . and I did not want to show that I was crying.'

### 19.4.4 Mixtures of emotions and conflicting emotions

While the preceding categories have provided examples dominated by positive or negative feelings, there are many reports with mixed appearances of feelings, feelings that change during the course of listening, feelings that are opposites to one another or in conflict with each other. Musicians' nervousness before a public performance may turn into feelings of joy and inspiration:

[E9] The whole day before I was very restless and tense . . . When I sat down to begin playing I felt, beside tension, increasing feelings of joy. The whole first piece became like a joy-intoxication, particularly as the audience looked very happy and helped to beat time through the whole piece . . . Afterwards I felt both elated and glad and totally relaxed.

Listeners may feel nervous concerning musicians' possibly unsatisfactory capacity, a nervousness that may soon disappear:

[E10] During concerts it often happens that I begin to worry about how it is going to turn out . . . Only a few minutes listening was enough to make me feel quite safe, lean back and let the music speak.

Listeners may be shaken, even shocked by the music, but after a while these feelings may turn into positive ones. A mainly perceptually conditioned shock is the overwhelming loudness and volume perceived when a musician is sitting inside an orchestra for the first time:

[E11] Then came the shock. The sound shock. The sounds that filled all of me, pressed me together into a small louse. Where would I go? I must get away . . . But soon this feeling changed into a courage to grow into the sounds, become a part of the music.

Anxiety or depression may be relieved by music:

[E12] I was in a deep depression, everything was black . . . This day began as all the others. I had great difficulties to bring myself to leave the bed's hiding place and get out into the threats that the new day meant . . . I went to a church where XX as usual was going to play the organ at noon. . . Somewhere in the piece by YY the fantastic happened, my mood was totally changed. Darkness became light and my uneasy mind calmed down . . . I left the church full of appetite on life.

Mixtures of positive and negative feelings may cause confusion:

[E13] One moment I felt like an abandoned child, lost and lonely, the next moment I was like a child in its most carefree moment . . . Suddenly a terrible feeling of confusion came upon me. In what I believe to be but a few seconds, a whole series of conflicting emotions: fright, sorrow, hate, love, anger, and last an internal peace that words cannot describe.

A report that provided penetrating introspective evidence of conflicting emotions was given by a young lady, fascinated by Gustav Mahler's music, when she attended a performance of the reconstruction of Mahler's unfinished Tenth Symphony:

[E14] I was waiting in an atmosphere of terror, waiting for what was to come after the intermission . . . I was scared. In fact already long before, at home and in the subway, I was scared. Afraid of being disappointed, afraid that this work of music which I had been wondering so much about would not be that precipice that I had imagined, would not be that step toward the new music that I thought it would be . . . The intermission had come to its end . . . I had never been so full of expectancy before. The conductor arrives. Silence. He raises his baton and starts the lonely violin part. After a few bars with this monotonous, curiously ambiguous song, like a doubt, the sun arrives. The orchestra is breaking out in a warmth that is fascinatingly painful. I remember tears filling my eyes. I felt as if I understood a message, from one time to another, from one human to another. I sat as if turned to stone, with my fingers gripping the armrest. After a while I became quite dizzy. Then I realized that I had forgotten to breathe . . . At the end of the first movement the whole orchestra is joining in a terrible great organ-like sound—a cry of pain and terror. It came as a shock after quite an enduring pianissimo. I was totally overwhelmed by the violence of this dread for death. I perspired and was hardly able to breathe. For once in my life, I had been totally knocked to the ground by something so great and strong that it wiped out all those thoughts that can slip by when one listens without fully concentrating . . . The last movement opens with eleven beats of an enormous wooden club against the largest bassdrum. The effect is horrible. A flute, followed by the strings, starts hesitatingly to develop a melody, but is suddenly crushed by a beat of the wooden club. I was totally taken by surprise. Each time the club beat, I was seized by terror. Each time an instrument began to sound, doubtfully, I anticipated the hit of the wooden club, I mourned. Here, too, I felt like I was a receiver of a message . . . I felt like I had not missed a single note . . . the feeling of being spoken to, strongly and directly, lived inside me. Maybe it was because I knew the music was not composed to completion, had not been ornamented, that it was so naked—an uncensored document about a condition oscillating between the deepest fear of death and an attraction towards darkness. After the orchestra silenced, the audience sat quiet for a long while . . . I think there were many like me who felt that we had been given the opportunity to look into a soul and therefore felt joy. Then the applause broke loose. And it would not stop. The way through the symphony had been hard and frightening, even shocking, but it was a great happiness—almost an honour—having had the possibility to experience this. I was flying home, weightless. I was glad that I attended the concert alone because now all words would have seen superfluous, even my own. My thoughts had nothing to do with words.

### 19.4.5 Relations to other SEM aspects

*Physical responses*

Physiological and behavioural responses were often reported (see E14 for some examples). By far the most common response was tears, meaning anything from getting moist eyes to uninhibited crying. In most cases tears were associated with positive feelings, but there were also examples of tears in connection with sadness and grief. Next in frequency came shivers/chills, thereafter gooseflesh, to feel warm/perspire or feel cold, muscular relaxation, changed breathing, heart race, pressure over the chest, hairs standing on end, various stomach responses, muscular tension, trembles/quivers, lump in the throat, dizziness, pain, and single examples of clouded eyes and dry mouth. Among overt behaviours, the dominating response was moving or dancing to the music (e.g. E1), followed by singing, shouting, smiling/laughing and—in a certain contrast to the preceding ones—closing one's eyes, sitting absolutely still hardly breathing (e.g. E14), and being unable to sing, play, or talk because of overwhelming feelings.

*Note on other SEM aspects*

In the excerpts we note the occurrence of many other phenomena, most of them also mentioned by previous authors (see Section 19.1): *quasi-physical feelings* (E3: 'the music started to take command of my body. I was charged in some way . . . I was filled by an enormous warmth and heat', E14: 'I was flying home, weightless'); *intensified perception* (E3: 'the music became so distinct', E4: 'All colors are vivid beyond description'); *total concentration on the music* (E1: 'It was only one thing that mattered: the music!', E3: 'Nothing else existed'); *disorientation of time* (E4: 'no time'); *loss of self-consciousness* (E4: 'My head is gone, body too'); *no thoughts* (E1: 'You don't think about what you are doing'); *intense feeling of living* (E1: 'you are living only for what is happening 'right now"'); *direct experience* (E4: 'I seem to be experiencing directly. . .feel the potential for anything to pass through me', E14: 'I felt like I was a receiver of a message . . . the feeling of being spoken to, strongly and directly, lived inside me'); *merging with the music* (E2: 'I felt like I was penetrated by Bach's spirit', E3: 'giving myself up to the music and the rhythms', E4: 'only the waterfall of sound cascading around and through me'); *feeling of freedom* (E1: 'One feels so free somehow', E3: 'it was as a kind of liberation'); *gaining new insight* (E2: 'I had come to the deepest insight and found the genuine, true and correct expression'); *altered perception of the world* (E4: 'the whole room, the whole universe is present here in perfect harmony . . . That pure existence and nothing else could be so beautiful!'); *community* (E1: 'You are like a part of it all', E4: 'I am totally a part of it all'); *transcendence* (E4: 'the sounds enveloping me in a world aside of this—sinking deeper beyond words, beyond thoughts'); *experience eludes verbal communication* (E4: '[I] have no words to express this state of bliss, harmony, meaningless beauty', E14: 'My thoughts had nothing to do with words').

SEM includes many other aspects as well (cf. Table 19.1). Further analysis of these and their relationships to each other is outside the scope of this chapter.

### 19.4.6 Influencing factors

Factors influencing SEM may be classified into musical, personal, and situational factors, briefly examplified in the following.

*Musical factors*

The music featuring in these 400 reports came from many different genres. Only a crude classification is made here. Art music (anything from old church music up to electronic music) appeared in barely half of the reports. The remaining half included pop/rock, jazz, folk music, and various kinds of popular music. Thus, SEM may occur in connection with 'any' music, and one can hardly find any distinct differences in emotional reactions to music of different genres.

In the excerpts, there are some examples of specific musical factors which caused or contributed to the strong emotional reaction. The negative SEM in E5 was attributed to the high volume, heavy drumming, screaming saxophone, and monotonous and howling song. In E6, a 'heart-rending and ghost-ridden' chord in Mahler's Tenth Symphony led to a horror reaction. It is probably the same chord that also was mentioned in E14: 'the whole orchestra is joining in a terrible great organ-like sound—a cry of pain and

terror' causing a shock in the listener. The notation of this chord (Fig 19.1) reveals a fortissimo chord extending over the whole frequency range of the orchestra with intervals of (mostly) minor thirds between its component tones contributing to very pronounced dissonance (cf. dissonance in relation to critical bandwidth, Dowling & Harwood 1986, pp. 82–5). There is also hardly anything in the immediately preceding music that would make the listener expect such an outburst. The chord thus works as an icon of strong negative emotion and also violates expectations, two ways of representing or arousing emotions (Dowling & Harwood 1986, pp. 205–24). The same chord appears again, further elaborated, in the final movement of the symphony (bars 275–283). To my knowledge, such a complex chord has no earlier equivalent, and it appeared at a time around 1910 when atonal music made its appearance, also in Vienna, with Schoenberg, Webern, and Berg. The listener in E14 also referred to the beginning of the last movement in this symphony, where the beats of 'an enormous wooden club against the largest bassdrum' crushed any attempt by other instruments and caused horror.

Rather than indicating specific moments in the music, many respondents more generally referred to musical factors such as timbre (e.g. an unusual instrument or unusual combination of instruments), loudness/dynamics (e.g. crescendo, diminuendo), tempo (e.g. accelerando), mode (e.g. transition minor to major), rhythm (often in connection with dancing, as in E3), beautiful melodies and harmonies, thick texture, increasing tension followed by relaxation, and still others. It is striking, however, that although many respondents were active musicians, in their reports they did not often use conventional musical terminology but rather terms that can be used by 'anyone': 'brutal chords', 'heavenly beautiful tones', 'majestic sound', 'pure music', 'enormous power and energy', and the like. Metaphors were used now and then: 'every chord goes like waves through yourself', 'in the last bars it is like a light over my closed eyes which more and more goes out in the following diminuendo'. Respondents also often attributed their reaction to the qualities of the performance—the skill, concentration, and involvement displayed by the musician(s), including conductors as well. This was very much due to visual impressions of the musicians, the instruments, and the physical and social environment.

*Personal factors*

Besides common demographic variables such as gender, age, educational level, and profession, our classification of person-related factors influencing SEM includes a variety of factors, for example:

1. *Physical state*, for instance, feeling well, rested, tired, ill.
2. *Cognitive factors*: expectations (positive, negative, not specific, see, for example, E14); feel a need for listening to or performing music; be attentive, receptive, sensitive, open-minded, or not; having heard or performed this music earlier, or not; be familiar or not with musical style, composer, performers, etc.; use or abandon analytical attitude; have pleasant or unpleasant memories of the music (e.g. in E8 trumpet blows were associated with war).
3. *Emotional state*: be in high or low spirits, calm, relaxed, nervous (e.g. E9, E10), stressed, depressed (e.g. E12), in crisis, in love, unhappy love, etc.

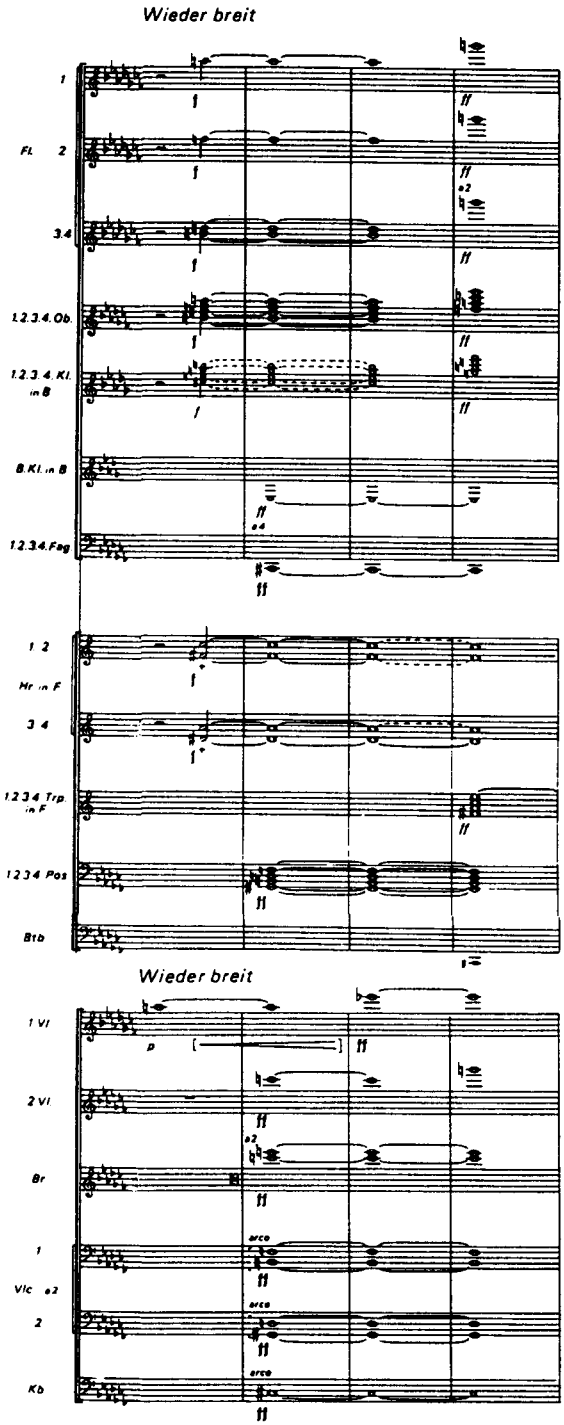



Figure 19.1 The chord in bars 203–206 of the first movement of Gustav Mahler's Tenth Symphony arranged by Deryck Cooke (1976). The chord is repeated in bar 208. Copyright ©1966 (Renewed) by Associated Music Publishers, Inc. (BMI). International Copyright Secured. All Rights Reserved. Reprinted by permission.

4. *Personality-related variables* (as suggested by respondents themselves) such as temperament, maturity, and dispositions.

*Situational factors*

Factors related to the listening or performing situation included, among others, the following:

1. *Physical factors*: acoustical conditions; listening to live or recorded music; listening or performing at home, in a concert hall, church, restaurant, outdoors, or on a Walkman, etc.; position, close or far away, in relation to performers or among co-performers (e.g. E9); visual conditions (e.g. good sight) and impressions, time of day, season, weather, etc.
2. *Social factors*: listening or performing alone or together with others; size, appearance, and behaviour of audience (see, for example, E1 and E9), of performers, or co-performers (see, for example, E5).
3. *Special occasions/circumstances*: experience during vacations, in another country, in unusual environment, etc.
4. *Performance conditions*: for instance, music well rehearsed or unrehearsed.

*Interactions among factors*

A reasonably full treatment of these factors and their effects on SEM goes far beyond the scope of this chapter. It is obvious, however, that emotional reactions in SEM, as in any music situation, depend on interactions between musical, personal, and situational factors (the latter two are sometimes indistinguishable). The relative importance of each factor varies, but none of them can ever be excluded. Even in reports where musical factors appear to be very strong determinants, there are influences of personal and/or situational factors, for instance, in E6 (grief after relative died), E7 (alone at home in a large house on a dark evening), E8 (memories of war during childhood), and E14 (expectancy and fear of disappointment before listening to Mahler symphony).

In many cases, it is obvious that personal and/or situational factors are in fact more important than the music in question, as when, for instance, a piece of music is associated with war (e.g. E8), happy or unhappy love, an exotic environment, beautiful nature, pleasant holidays, certain ceremonies, etc. (for examples, see Gabrielsson & Lindström 1993, 1995). Such cases illustrate what Dowling and Harwood (1986, Chapter 8) call indexical representation, that is, a piece of music is associated with some extramusical object or situation.

In consideration of the ever present interactions between musical, personal, and situational factors it may be appropriate to speak of strong experiences *of and with* music (Gabrielsson & Lindström Wik, 2000) or strong experiences *in connection with* music.

## 19.5 Discussion

In accordance with the title of this chapter and this volume, the focus here has been on emotional aspects of SEM, supplemented by some brief outlooks to other aspects and influencing factors. A look at Table 19.1 shows that this represents but a limited part of the

available material on SEM. Cutting out a part of a whole inevitably leads to gaps and imperfections, and the loss of relevant connections and of broader context. For further content and analysis, I refer to our earlier reports as well as to reports in press and in preparation.

### 19.5.1 Comments on emotion

A general difficulty with the concept of emotion is how to make a reasonable demarcation of emotions or emotional feelings versus other feelings. Both psychologists and people in general use terms such as feelings, emotions, and mood in very variable, sometimes idiosyncratic, ways (e.g. Plutchik 1994, Chapter 3; see also Sloboda & Juslin, this volume). This is certainly also true for the corresponding terms in Swedish (for instance, the Swedish word 'känsla' is used in a variety of contexts and meanings). We have therefore simply stayed as close as possible to the terms used by the respondents themselves, provided that they have a reasonable emotional meaning, or at least an emotional connotation. Most of these terms appear in alphabetical order in Table 19.2, divided into positive and negative emotions. No further division is made presently (however, a distinction between different levels of arousal was hinted at in Section 19.4.2). Of course, some reservations must be made regarding the proper translation of Swedish terms into corresponding English ones.

Table 19.2 Emotions mentioned in SEM reports

| |
|---|
| *Positive emotions* |
| Amazement, beauty, bliss, calm, contentment, delight, ecstasy, elation, enjoyment, euphoria, excitement, exhilaration, expectancy, fascination, freedom, gladness, gratitude, happiness, harmony, honour, hope, humility, intoxication, joy, longing, love, loveliness, overjoy, overwhelmed, peace, perfection, pleasure, rapture, relaxation, repose, sacredness, safety, satisfaction, smallness, solemnity, surprise (positive), tension (positive), thrill, warmth, wonder |
| *Negative emotions* |
| Afraid, anger, anguish, anxiety, chaos, conflict, confusion, depression, disappointment, discomfort, fear, fright, frustration, gloom, gruesome, hate, horror, mourning, nervousness, pain, panic, sadness, scare, shaken, shame, shock, sorrow, split, surprise (negative), tension (negative), terror, uneasiness, worry |

In the reports, we find numerous examples of so-called 'basic' emotions. However, opinions differ concerning which and how many basic emotions there are. The survey by Plutchik (1994, p. 58) indicates the best agreement regarding emotions like happiness, sadness, anger, and fear, but proposals by various authors also include (in alphabetical order) amazement, anxiety, elation, enjoyment, expectancy, grief, interest, love, pain, panic, pleasure, satisfaction, shame, shock, surprise, and tension, all of which are represented, to varying extents, in our material. Also proposed, but not represented in the present material, are boredom, contempt, disgust, and guilt—negative emotions which seem less plausible in connection with SEM. However, other negative emotions such as panic, shame, and shock may seem equally implausible, yet they have appeared; so too hasty exclusions should be avoided.

In previous studies on peak experience and flow (Csikszentmihalyi 1990; Laski 1961; Panzarella 1980), rather little was written about emotions. However, Maslow (1968, pp. 87–8) in his survey of 'generalized peak experience' mentioned happiness, ecstasy,

rapture, love, wonder, awe, reverence, humility, surrender before the experience as before something great, fear of being overwhelmed, and occasionally a description of the experience as sacred. These feelings are also represented in the present reports (although not all of them in the selected excerpts). None of them, except happiness, love, and fear, appears in lists of basic emotions but rather represent complex mixtures with considerable ingredients of cognitive character. Generally, the borderline between cognition and emotion is blurred, and from the standpoint of investigating strong experiences with music, this distinction may not be very important.

Whether some form of cognitive appraisal precedes emotion or not can be discussed at considerable length. In several SEM reports, respondents described their reaction as a kind of 'direct experience', something that bypasses cognitive barriers and goes 'straight into the emotions'. This suggests that there was no cognitive appraisal, at least not at a conscious level. However, there are also many reports in which the respondent first told about, say, the beautiful timbre of a singing voice, fantastic acoustics, irresistible rhythm, etc., which may be interpreted as rather deliberate cognitive appraisal mediating the emotional reaction. Of special interest are wordings referring to the beauty of the music (which we subsume under enjoyment). In some cases, there is obvious cognitive appraisal behind the beauty response—an insight into, and admiration of, say, the composer's ingenious construction of the music, its formal properties, delicate instrumentation, combination of text and music, etc.—that is, a kind of 'intellectual' enjoyment. In other cases, the response comes as a direct and spontaneous exclamation—'Oh, my God, how beautiful!'—devoid of any analysis or reflection.

The distinction between emotions *expressed* in music and emotions *aroused* by music is not always easy to make in the reports. For instance, 'the music became more and more gruesome' can be interpreted as the respondent's 'neutral' description of perceived expression, or as indicating that the respondent felt anxious, or as a combination of both possibilities.

### 19.5.2 Comments on methodological issues

Verbal reports are associated with several well-known problems (Ericsson & Simon 1993), not least with regard to retrospective reports as here. Many reports refer to experiences which happened many decades ago and are obviously sensitive to memory distortions. Furthermore, subjects differ much in linguistic competence, vocabulary, and willingness to use language to describe their experiences. Many respondents frankly stated that their SEM was impossible to describe by words. Others wrote long and penetrating accounts but still complained about the validity of words to describe what they actually felt—the experience seemed ineffable. However, presently we have no better tools available than language in attempts to approach phenomena like SEM, and I am convinced that language will remain the main instrument for a considerable time still. For some further comments, see Gabrielsson (1989, 1991) and Gabrielsson and Lindström (1993).

### 19.5.3 Continued research

Recently, Lowis (1998) tried to investigate peak experiences under experimental conditions. Subjects listened to prerecorded classical music: five pieces of 'gentle' music (e.g.

'Nimrod' from *Enigma Variations* by Elgar) and nine pieces of 'up-beat' music (e.g. the ending of *The Firebird* by Stravinsky), and used a button device to indicate 'a moment of particularly deep and profound pleasure or joy . . . the sort that produces a tingle in the spine' (Lowis 1998, p. 212). Further definitions of peak experience alluded to phenomena described by Maslow (see Section 19.1). Roughly two-thirds of the subjects pressed the button at least once. The up-beat music generated significantly more responses than the gentle music. Significant correlations were found between frequency of button pressings and ratings for the following feelings evoked by the music: joy, love/tenderness, longing, sadness, reverence/spirituality, action, and memory/thoughtfulness. Waterman (1996) and Sloboda (1999), although not directly referring to peak experiences, used other, partly similar methods for obtaining listeners' continuous and immediate indication of emotionally loaded moments in the music, thereby also increasing the possibilities of relating them to properties of the music in question.

The advantages of experimental procedures are the control of musical stimuli and the immediate, continuous response by the subjects. On the other hand, the situation is by necessity more or less artificial, and a 'button response' does not not tell much about the contents of the experience unless it is supplemented by other information, as is somewhat examplified in these reports. For continued research of strong experiences of music, whatever they are called, a flexible interplay between naturalistic and experimental approaches seems appropriate in order to combine their respective advantages and compensate for their respective drawbacks.

That music is one of the most effective triggers of strong emotional experiences is amply confirmed by several studies referred to in this chapter. 'The power of music' is a cliché that has considerable validity. It remains a fascinating challenge, although frustrating at times, to investigate how, why, and in what contexts we can be so strongly affected by music.[1]

## References

**Cooke, D.** (1976). *A performing version of Gustav Mahler's draft for the Tenth Symphony.* New York: Associated Music Publishers, Inc.

**Csikszentmihalyi, M.** (1990). *Flow. The psychology of optimal experience.* New York: Harper & Row.

**Dowling, W. J. & Harwood, D. L.** (1986). *Music cognition.* New York: Academic Press.

**Ericsson, K. A. & Simon, H. A.** (1993). *Protocol analysis. Verbal reports as data,* (2nd edn). Cambridge, MA: MIT Press.

**Gabrielsson, A.** (1989). Intense emotional experiences of music. In *Proceedings of the First International Conference on Music Perception and Cognition,* pp. 371–6. Kyoto, Japan: Japanese Society of Music Perception and Cognition.

**Gabrielsson, A.** (1991). Experiencing music. *Canadian Journal of Research in Music Education,* 33, 21–6.

[1] My sincere thanks to Siv Lindström Wik for many years of continuous and repeated analysis and discussion of our SEM material. This research is supported by The Bank of Sweden Tercentenary Foundation.

Gabrielsson, A. (in press). Musicality, music performance, and music experience. In *Multidisciplinary perspectives on musicality,* (ed. D. Coffman, K. Gfeller, D. J. Nelson, & C. Rodriguez). Iowa City, IA: University of Iowa Press.

Gabrielsson, A. & Lindström, S. (1993). On strong experiences of music. *Musikpsychologie. Jahrbuch der Deutschen Gesellschaft für Musikpsychologie,* **10** , 118–39.

Gabrielsson, A. & Lindström, S. (1995). Can strong experiences of music have therapeutic implications? In *Music and the mind machine. The psychophysiology and psychopathology of the sense of music,* (ed. R. Steinberg), pp. 195–202. New York: Springer.

Gabrielsson A. & Lindström Wik, S. (2000). Strong experiences of and with music. In *Musicology and sister disciplines: Past, present and future,* (ed. D. Greer) pp. 100–8. Oxford, UK: Oxford University Press.

Goldstein, A. (1980). Thrills in response to music and other stimuli. *Physiological Psychology,* **8**, 126–9.

Laski, M. (1961). *Ecstasy. A study of some secular and religious experiences.* London: Cresset Press.

Lowis, M. J. (1998). Music and peak experiences: An empirical study. *Mankind Quarterly,* 39, 203–24.

Maslow, A. H. (1954). *Motivation and personality.* New York: Harper.

Maslow, A. H. (1968). *Toward a psychology of being,* (2nd edn). New York: Van Nostrand Reinhold.

Maslow, A. H. (1976). *The farther reaches of human nature.* New York: Penguin Books.

Panksepp, J. (1995). The emotional sources of 'chills' induced by music. *Music Perception,* **13**, 171–207.

Panzarella, R. (1980). The phenomenology of aesthetic peak experiences. *Journal of Humanistic Psychology,* **20**, 69–85.

Patton, M. Q. (1990). *Qualitative evaluation and research metods,* (2nd edn). London. Sage Publications.

Plutchik, R. (1994). *The psychology and biology of emotion.* New York: Harper-Collins.

Sloboda, J. A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music,* 19, 110–20.

Sloboda, J. A. (1999). Musical performance and emotion: Issues and developments. In *Music, mind, and science,* (ed. S. W. Yi), pp. 220–38. Seoul: Seoul National University Press.

Waterman, M. (1996). Emotional responses to music: Implicit and explicit effects in listeners and performers. *Psychology of Music,* **24**, 53–67.

# POSTLUDE

CHAPTER 20

# MUSIC AND EMOTION: COMMENTARY

JOHN A. SLOBODA AND PATRIK N. JUSLIN

The chapters collected here have highlighted a range of dichotomies (or bi-polar dimensions) which 'open windows' to a wider view than that with which most of us will have started our consideration of emotion in music. For many of us who would be naturally drawn to this book, and the enterprise it represents, the assumptions we make about emotion are probably a result of our by now effortless (and hence often unreflexive) accommodations to the specific musical culture which we inhabit. These assumptions often embody one pole of a particular dichotomy. Progress in understanding musical emotions may require the distancing from our own perspective that an articulation of such dichotomies can provide. This postlude highlights five core 'dichotomies' (from a considerably larger potential set), calls for attempts to reach beyond these dichotomies, and concludes with some general evaluative comments about the field as a whole, based on the evidence provided in this book.

## 20.1 Emotion as reception versus emotion as construction

A theme running through many of the chapters concerns the role of the person experiencing the emotion in the causal process. At one extreme is the case where the emotion is 'induced' in someone by the 'force' of the music, in much the same way as the appearance of a snake might induce fear in someone, pretty much independently of whether or not the person concerned had it in mind to feel fear. This case motivates the concern of many contributors to this volume with fundamental emotional processes (see, especially, the chapters by Peretz, Juslin, and Scherer & Zentner). At the other extreme is the case where the person appropriates the music in order to generate an emotional experience, using the music as a resource in a more active process of emotional construction. In such a case, there would be no emotion had the person not decided that there should be. Such a case is driving the arguments of contributors such as DeNora, Cook and Dibben, and Sloboda and O'Neill.

Some of the historical confusions and perplexities engendered by the phenomenon of music's emotionality may have been contributed to by an overemphasis on the 'receptive' rather than 'constructive' pole of the dichotomy within explanatory discourses. The receptive approach has sat comfortably with the apparent role of the listener in both the concert hall and the psychological laboratory. In these settings, the superficial appearance is one of passivity. We sit still and silent, and music is 'presented'

to us by powerful others whose purpose is to influence our thoughts and feelings (and, in the case of the laboratory, record them).

Several chapters in this book suggest that the 'active' view of emotional engagement is revolutionizing many of the intellectual approaches to the music. From the 'new musicology' through 'social constructionism' in psychology to the 'interpretive' approach in sociology, a new appreciation is being elaborated. This places the person experiencing music as the crucial agent at the junction of a network of forces and influences—an agent who *makes* emotion happen by his or her 'musicking', to use Small's (1998) deeply appropriate neologism. In one sense, the still silent member of a classical audience is no less active than the performer on the stage. It is simply that the form, vectors, and boundaries of that activity are different.

## 20.2 Emotion as biology versus emotion as culture

A second dichotomy, which is strongly related to the first (in practice if not by necessity), is that between a biological and a cultural approach to musical emotion. On the one side of this dimension, we find the biological perspective on musical emotion, evident in the chapters by Peretz and, to a lesser extent, Juslin, and Scherer and Zentner. This approach recognizes that emotions have a biological basis and reflect the workings of neurological and neurochemical mechanisms developed by natural selection. Emotions have a long evolutionary history, and their neural substrates show 'homologies' with those of non-human primates. The biological perspective takes advantage of the fact that natural selection is the only force in nature that can build functional organization into organisms. Indeed, natural selection is a casual process in which a structure spreads precisely because of its functional consequences, that is, function determines structure (Cosmides & Tooby 2000). The flagship of this perspective on emotion is the 'affect program' approach represented by Paul Ekman, among others. According to this view, each emotion should be defined functionally in terms of how it is adaptive in managing a particular problem scenario that has occurred frequently during evolution.

A biological perspective on music and emotion recognizes that neuropsychological and evolutionary considerations are crucial also for the case of music. Expression, perception, and induction of emotion via acoustic signals did not begin when human beings invented the first musical instrument. Humans have always communicated emotions via the voice, and many animals communicate emotions via acoustic signals as well. Thus, as Juslin notes, it seems plausible that some aspects, at least, of how music intersects with emotion are constrained by psychological mechanisms for acoustic communication that have been 'shaped' by evolution. A biological approach assumes that music recruits essentially the same emotional circuits as other emotion-inducing events, and that we, therefore, can learn a lot about musical emotions by understanding the functional design of emotions in general. Moreover, as Peretz observes, neuropsychological studies may help to resolve some controversies and conceptual problems with regard to the nature of musical emotions.

On the other side of the biology–culture dichotomy, we find a sociocultural approach to musical emotions, represented in this book by the chapters by Becker and DeNora, and also to some degree by Sloboda and O'Neill. Although a greater understanding of

emotions in wider non-musical settings is clearly important to the progress of the musical enterprise, we have to remember that music is first and foremost a cultural artefact, and so unlike many of the other things and events to which we may experience emotions. Fearful responses to evolutionarily 'time-honoured' threatening situations (such as snakes) are, in some sense, non-optional. Our genes have designed it so that we have 'quick and dirty' responses to certain situations that impinge directly on our survival (LeDoux 1996). The behaviour of other human beings, and the things they say within practical discourse, can often point directly to these 'basic' event types, and so take on their power and immediacy.

However, cultural artefacts often require a great deal more 'mental scaffolding' to hold them in place. Although this scaffolding quickly becomes second nature within a particular 'habitus' (in the sense outlined by Becker in this book), it can take on different forms, even in superficially similar situations. If even one unnoticed element in the scaffolding is absent in a particular musical event, then the emotion achieved can be quite different, for reasons which may be obscure to the participants. Thus, for instance, the well-known phenomenon among classical performers of the unpredictability of the performance that 'really comes alive', or the perplexity of the adult listener returning to a favourite piece of music and finding that it does not exert the usual emotional power. In both of these cases, it may be a mistake to look for the difference inside the auditory musical object itself, but rather in the way that the participant is construing or appropriating the musical event.

It is important to note that mainstream psychological research on musical emotion has occupied a middle position in regard to this biology–culture dimension. In other words, most psychological studies have taken neither the biological foundation of emotions, nor the wider social context of emotions, into consideration. Thus it seems important that psychologists are willing to broaden their perspectives on musical emotion, keeping themselves informed about the latest developments in neighbouring fields like neuropsychology and sociology. This very book may represent one small step in that direction.

## 20.3 Emotion perception versus emotion induction

One further dichotomy involves the notion (known from ancient Greece) that music may both 'represent' emotions (that are perceived by the listener) and 'induce' emotions (that are felt by the listener). Such a distinction between emotion perception and emotion induction is made in many, if not all, of the chapters in this book. Some (for example, those by Gabrielsson & Lindström and Juslin) focus mainly on perception of emotion. Others (for example, those by Simonton, Meyer, Scherer & Zentner, Sloboda & O'Neill, and Gabrielsson) focus mainly on emotion induction.

It is pointed out by some authors that it is difficult sometimes to separate these different processes in empirical studies. Moreover, the interaction between perception and induction is still little investigated. Research summarized by Scherer and Zentner, however, indicates that it is important to take the distinction into consideration when designing research (this has not always been done in earlier research). Becker reviews some evidence which suggests that the perception–induction distinction may not be

unique to the Western world, but also appears in other cultures. There is also some evidence that the distinction has neurophysiological reality, because both expression and perception of emotion seem to be associated with processing in the right hemisphere of the brain, whereas induced emotion seems to be lateralized according to emotional valence—positive emotions processed in the left hemisphere, and negative emotions in the right hemisphere (Blonder 1999). It is possible that brain-imaging techniques, such as PET, fMRI, and ERPs (reviewed by Peretz, this volume), in combination with other measures, might help us to empirically separate felt emotions from perceived emotions in future research. It seems likely that different theories may be required to explain these different processes.

## 20.4 Emotion as private experience versus emotion as public expression

At the heart of some of our dilemmas in understanding how emotion in music 'works' is the common Western view that emotion is primarily a private experience. This view has made it deeply problematical for researchers to confidently identify when an emotion is taking place (cf. previous section). We seem to be required to 'infer' emotion from a variety of unreliable and incomplete cues, such as verbal and physiological data. We are not even sure to what the supposed emotion is a reaction. Within a complex and rapidly changing musical structure, our emotional response may be slightly delayed from the crucial precipitating event, or may be an integrative response to a cumulative series of transitions across time (as discussed in different contexts by Meyer, Scherer & Zentner, and Schubert, among others).

Such problems are unravelled, at least in part, when one considers traditions in which the emotional outcome can be rather more directly read from the emotional behaviour of the participants, or from agreed conventional signals of such outcomes within the musical act. Becker, perhaps, provides the most radical challenge to our common conceptions of emotion by showing how these link to conceptions of the self, which are by no means universal. In some cultures, emotional displays associated with music are the means by which social relations are codified, expressed, and given 'power'. They express an accepted social reality, and the issue of whether a particular individual really feels the emotion expressed becomes secondary. Cook and Dibben remind us that the 'individualistic' approach to emotions in music has had a fairly short history, even within the Western tonal tradition.

Having articulated an important dimension, much work is now needed to locate various types of musical 'habitus' along it. Becker's example of the Pentecostal service is particularly interesting, since the surface discourse suggests a private perspective (the goal of achieving a state of individual ecstasy), yet there would be no 'point' to the occasion if individuals who achieved this state did not display their state through relatively ritualized behaviours which tell other participants 'I've made it'. And, of course, the willingness to indulge in such affect displays creates a kind of self-fulfilling prophecy, through the mechanism that Scherer and Zentner describe as 'peripheral route production'. The mere actions involved in building up to and displaying ecstasy will, along with the 'emotional contagion' of other people providing the same affect display, in themselves help to bring about the desired inner state. Becker also emphasizes the impor-

tance of being in a properly receptive frame of mind. Critical responses will 'block' the emotional path to ecstasy. Thus the music will not 'work' for someone unless they have already decided to allow it. Those of us who find it goes against the grain to enter into such activities are only too painfully aware of the complex web of socially engendered beliefs, interpersonal obligations, and dispositions that underpin the Pentecostal 'habitus' (because they are so alien to us). Similar belief systems are, however, there in every 'habitus'. The Pentecostal example highlights the often intricate relationship between private experience and public expression of emotion in relation to musical events.

## 20.5 Emotion as separate from the musical experience versus emotion as constitutive of the musical experience

Many of the chapters in this book express a general perplexity that it has taken so long for emotion to gain even a modest foothold in the arena of academic discourse about music (from whatever disciplinary perspective). Although methodological difficulties have slowed the progress in some disciplines, notably psychology, this cannot be the whole story. The psychology of emotion in general is far more advanced than the psychology of emotion as applied to music. It is more plausible to attribute the delay in progress to two factors.

First, there is the problem of tying down musical emotion to some 'formal object'. The formal object of an emotion, according to Solomon (2000, p. 11), can be described as 'the minimum essential set of 'beliefs' defining an emotion and an emotional experience'. Emotions have intentionality: they are always 'about' something (e.g. an event, person, or state of affairs). This notion is problematic in the case of music. It is difficult to say what musical emotions are 'about'. (Besides the obvious: that they are, in some way, about the music.) As recognized by philosophers, an emotion can also be about some non-existing, merely imagined object (e.g. a 'virtual persona' in the music). An important task for the study of musical emotion (as noted by Davies) is thus to understand how music can come to express and induce emotions despite an apparent lack of 'beliefs' that usually underlie such a response. Many chapters in this book propose possible theoretical solutions to this crucial problem, for instance in terms of 'musical expectancy' (Meyer), 'emotional contagion via independent brain modules' (Juslin), 'melodic originality' (Simonton), 'emotional memories', 'proprioceptive feedback', and 'conditioning' (Scherer & Zentner).

Second, delayed progress in understanding musical emotions partly reflects the cultural assumptions about music that have been shared by those people most well-equipped to move the field decisively forward. At the centre of these assumptions for the last 150 years or more has been the belief that music is epitomized by the canon of masterworks produced by a small circle of 'great' composers. Such works are expressions of things that the composer has grasped (which may include emotions), and the job of the listener is thus to 'learn'. What the listener 'feels' is largely irrelevant to this serious task. This task is traditionally accomplished through a grasp of formal structural descriptions of the work; that is, analysis. This activity can be undertaken at elementary level through the digestion of programme notes and other forms of commentary

designed for a wider public, in what has sometimes been called the 'music appreciation movement', or at an advanced level through participation in formal analytical activities.

Since dominant conceptions of music are almost entirely determined by what dominant groups *do* with music, it is inevitable that the belief should have developed that emotion is not core to music. Such a stance has contributed enormously to the disengagement of the general population from entire genres of music. Even among intellectuals in other spheres, the claim that 'I like classical music but don't understand it' is tragically common. Such people engage deeply and thoughtfully with music, are often excellently provocative concert companions, but since they lack a set of vocabularies and techniques which would allow them to coherently address structure independently of emotion they feel 'excluded' from the musical world. In effect, many people feel they are being told 'the thing that strikes you most immediately and powerfully about the music you hear, the thing which keeps you engaged with it and wanting to hear more of it, is really totally irrelevant to a proper appreciation and understanding of it'.

Fortunately, the winds of change are blowing at great speed through the musicological communities. In 20 years, many of the core assumptions of the field have been turned on their head, in favour of emotion (and much else that has traditionally been excluded). Some of the historical and social parameters of this are outlined in this book by Cook and Dibben. However, seeds of change were planted much earlier than 20 years ago, and, in the present context, there can be few more fertile seed-planters than Leonard Meyer. Meyer's (1956) persuasive arguments that structural engagement *requires* emotional engagement have provided music psychologists, philosophers, and musicologists with a firm launch pad to escape the gravitational inertia of traditional historical musicology. (His theory provides one fruitful solution to the problem of the formal object of musical emotions.) Yet, paradoxically, Meyer's theory initially had the opposite effect. Meyer, and later Lerdahl and Jackendoff (1983), implied that understanding musical emotion is best achieved through an understanding of musical structure. As noted by Cook and Dibben, this position could be taken to justify the almost indefinite postponement of consideration of expressive qualities in musicological practice, and a resultant separation between structural analysis and expressive interpretation.

This separation is, in many ways, still with us. The contributions of Persson, Bunt and Pavlicevic, Juslin, and Steptoe show, in their different ways, that a fully articulated discourse (and associated practice) of 'emotional development and management' within the musical profession is a long way off. The core aspects of the training of musicians have remained unchanged for 50 years, largely because those who are most important to the delivery of that training (that is, senior professional performing musicians) received their own training long before any of the ideas summarized in this book were in circulation. These ideas are slow to reach them, and those they teach, because there are few contexts in which the necessary dialogues can take place. However, there are huge changes that economic and cultural factors are forcing on training institutes for performing musicians. These changes provide a unique opportunity for rethinking both what is taught and who gets taught it. In this process, it surely must be inevitable that the world of 'academic and classical' music has to effect a far more thoroughgoing *rapprochement* with popular music, its audiences, and the industries it supports, than anything hitherto achieved.

## 20.6 Sources and components of emotion and different levels of analysis

In the foregoing, we have outlined five dichotomies that characterize this book. In doing so, we do not wish to imply that these dichotomies are valuable as such. Rather we feel that it is important to move beyond these dichotomies, and try to reconcile the different perspectives. Admittedly, this is far easier said than done. None the less, we believe it is important that we attempt to find a common ground upon which different approaches can meet. A case in point is provided by the biology–culture dichotomy discussed earlier. Biology and culture are both crucial aspects of phenotypic development, and it is therefore time to start exploring ways in which neurobiological factors *interact* with environmental inputs to produce phenomena like musical emotions. (How do sociocultural factors, such as learning and language, impact upon the neuronal patterning associated with musical emotions? Do neurological aspects constrain what is learned?) But how can different approaches to music and emotion be integrated? We believe that one key to understanding how the different approaches fit together is to be found in some of the theoretical distinctions introduced in various chapters of this book.

Sloboda and Juslin discuss two primary sources of emotion in music: *intrinsic* and *extrinsic* sources. Intrinsic sources are those that are non-arbitrarily embedded in structural characteristics of the music and that contribute to the creation, maintenance, confirmation, or disruption of schematic expectations (along the lines sketched by Meyer). Extrinsic sources include both *iconic* sources, which come about through some formal resemblance between a musical structure and some event carrying emotional tone, and *associative* sources which are premised on arbitrary and contingent relationships between the music being experienced and a range of non-musical factors, which also carry emotional messages of their own. We believe that failure to specify which source(s) of emotion one is investigating can lead to unwarranted controversies with those who investigate another source; what counts for one particular source may not count for another.

Second, it has been suggested in psychology that emotions can be broken down into various subcomponents, such as eliciting events, appraisal, subjective feeling, physiology, expressive behaviour, action tendencies, and emotion regulation (Mesquita & Fridja 1992). Such a componential approach drives Scherer and Zentner's review concerning emotional reactions to music. Different authors in this book tend to focus on different components. For instance, Gabrielsson focuses on subjective feeling, Peretz on physiological measures, Cohen on appraisal, Juslin on expressive behaviour, and Sloboda and O'Neill on emotion regulation. Again, what counts for one particular component does not necessarily hold for another component. One final distinction (already mentioned here) is that between emotion perception and emotion induction. It is likely that theoretical accounts of musical emotions will differ depending on which of these two processes a researcher is focusing on.

In addition to these different sources of emotion and subcomponents of emotion, it is clear that different authors operate at different levels of analysis, ranging from the biological and the psychological to the social and the cultural. Several commentators have pointed out that, unlike the natural sciences, the social and behavioural sciences

lack vertical integration; that is, they are not mutually consistent. However, an increasing awareness of research carried out at different levels of analysis should help to yield such consistency. An important goal for researchers of music could be that the theoretical accounts that they propose at one particular level of analysis should not be inconsistent with the accounts obtained at another level.

Given these distinctions, any research endeavour may focus on a specific combination of features. By being as clear as possible about the particular source(s) of emotion involved, the emotion components studied, the level(s) of analysis applied, and whether perception or induction (or both) is being studied, researchers may be able to integrate their approach with the work carried out by others. While it seems perfectly legitimate to conduct research on one particular level of analysis (or one component of emotion), and to bracket another domain of analysis, researchers should recognize that they are doing so, and acknowledge the legitimacy of other analytical levels.

In some cases, the apparent contradictions of different disciplinary approaches may be reconciled by observing that they focus on different sources of emotion, different components of emotion, or different processes (perception vs. induction). It is further possible that certain approaches may be more relevant for particular combinations of sources, components, and processes than others. A few examples may suffice to illustrate this point: An understanding of *intrinsic* sources of *induced* emotion requires structural analysis of the music and a model of the *appraisal* component involved. For this particular focus, a musicological–psychological level of analysis may be especially suitable. On the other hand, if we want to understand how *iconic* sources of *perceived* emotion are produced via reference to some non-verbal *expressive behaviour* component, a biological–evolutionary level of analysis may be highly fruitful. And, finally, if we want to understand how people talk about musical emotions, and interpret them within the social web of everyday life; how they use music to regulate their own emotions and influence their own self-identities, the *emotion regulation* component, a social constructionist approach, may be most suitable. (Of course, this does not preclude that other levels of analysis may have important contributions in each example.)

## 20.7 Is the field of music and emotion in good shape?

On the basis of the chapters collected here, and the work they refer to, there are grounds for considerable optimism. This is not a collection that largely rehashes old work, nor it is a collection of pious platitudes and promissory notes on work not yet undertaken. The vast majority of chapters refer to substantial research programmes yielding new data and new conceptualizations, the great majority of which have been undertaken and published in the last 10 years. A second source of optimism is the high level of theoretical and methodological incisiveness displayed within the research reported.

Emotion work is attracting the attention of some of the best-trained and most rigorous senior exponents of the various disciplines. Younger researchers are, as a result, gaining the confidence to develop focused programmes that branch outwards from PhD projects and find their ways into the most prestigious publication outlets. At least in some parts of the world, and in some disciplines, we have the conditions for what Kuhn (1962) described as 'normal science'. A third source of optimism is to be found in

the willingness of researchers from different disciplines to contribute to a book such as this (we do not know of many such resolutely multidisciplinary enterprises).

As psychologists we are particularly pleased that psychology appears to be playing such a central role in the resurgence of emotion studies in music. Psychology may be particularly well suited to such a role at the current time, because of the breadth of its scope (from the neurophysiological to the social), its methodologies (which are rigorous but at the same time also sufficiently 'neutral' to create a meeting place for researchers from different intellectual and ideological traditions), and its well-established record in advancing the understanding of a range of musical phenomena.

On the other hand, optimism should not lead to complacency. Several authors point to deficiencies, some rather major, in what has been achieved, and point out quite significant barriers to further achievement. Deficiencies include a lack of common understandings and definitions of key terms, a piecemeal approach to data gathering, and a narrowness in the domains studied. This narrowness applies to musical traditions (most work is on Western classical music), ages (there is very little work with children and older people), and contexts (a limited range of cross-cultural and subcultural contexts have been addressed). Barriers to further progress include the enormous difficulty of measuring emotions, particularly using multiple measures over extended events, such as pieces of music whose emotional expression may be constantly changing. The recent developments of increasingly non-intrusive methods of measuring brain activity may offer one important way forward.

Another barrier to progress is the relatively independent way in which research in the various disciplines develops. This may be an inevitable consequence of the structuring of the academic world. Researchers can only make substantial contributions if there is relatively secure employment. Employment is only available to those who have important things to contribute to their own discipline, as judged by opinion formers within that discipline. It matters far more to a young musicologist what other musicologists think of his or her research than what a psychologist might think. We have noticed several points within the collection of chapters here where links could have been made to work going on in another discipline but were not. More fundamentally, each discipline has its own trajectory, and issues which seem of great importance to psychologists might seem irrelevant to musicians. Therefore, what constitutes progress in this area may be a matter on which the representatives of the different disciplines will be unable to agree. Several authors end their chapters with a set of suggestions for further research. From outside the specific discipline one can certainly see that the research programmes suggested are possible, rational, and defensible. But are they necessary, or even important? Who would decide? Could the different disciplines ever agree, and would the practising musician or the untrained listener agree with them?

This is an issue that is not specific to music and emotion studies. It is a constant preoccupation of many people involved in the musical sciences. To put the question at its most blunt—does what we do help better music to be made, or for people to obtain more benefits from engagement with music? The answer is not clear. Music scientists sometimes look with envy on medical researchers. Everyone agrees that finding cures for major diseases such as cancer is worthwhile. As a result, governments and industry employ hundreds of thousands of people to research possible solutions.

Interdisciplinary teams are created. Researchers know what the goal is and can assess their success against the goal. What is the goal of music emotion research? There are probably as many answers as there are researchers. However, this book shows that researchers have common and overlapping concerns. Could these concerns ever coalesce into common agreed research agendas, which would direct the shape of the field for the next 10 years? Should they so coalesce? Would any major external body be concerned enough about these questions to provide major long-term funding? We feel that the answers to these questions are not obvious from the contents of the book. However, we hope that the book's existence may stimulate the debates that will lead to clearer answers to these questions, and that a future book on this topic might report progress in a rather more self-conscious and self-confident manner.

## References

**Blonder, L. X.** (1999). Brain and emotion relations in culturally diverse populations. In *Biocultural approaches to the emotions*, (ed. A. L. Hinton), pp. 275–96. Cambridge, UK: Cambridge University Press.

**Cosmides, L. & Tooby, J.** (2000). Evolutionary psychology and the emotions. In *Handbook of emotions*, (2nd edn) (ed. M. Lewis & J. M. Haviland-Jones), pp. 91–115. New York: Guilford Press.

**Kuhn, T. S.** (1962). *The structure of scientific revolutions.* Chicago: University of Chicago Press.

**LeDoux, J.** (1996). *The emotional brain.* New York: Simon & Schuster.

**Lerdahl, F. & Jackendoff, R.** (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.

**Mesquita, B. & Fridja, N. H.** (1992). Cultural variations in emotions: A review. *Psychological Bulletin,* **112**, 179–204.

**Meyer, L. B.** (1956). *Emotion and meaning in music.* Chicago: University of Chicago Press.

**Small, C.** (1998). *Musicking: The meanings of performing and listening.* Hanover: Wesleyan University Press.

**Solomon, R. C.** (2000). The philosophy of emotions. In *Handbook of emotions*, (2nd edn) (ed. M. Lewis & J. M. Haviland-Jones), pp. 3–15. New York: Guilford Press.

# AUTHOR INDEX


Abbate, C. 66, 66n
Abel, J. L. 297
Abeles, H. F. 97
Abell, T. D. 149
Abraham, F. D. 409
Abraham, R. H. 409
Abu-Lughod, L. 139
Adachi, M. 316
Adams, B. 114
Adams, B. L. 396
Addis, L. 31, 38n, 40n
Adelson, E. 89, 114, 323
Adler, R. H. 303
Adolphs, R. 110, 112, 124
Adorno, T. W. 49, 51, 57, 164, 165, 175
Agawu, V. K. 54, 62
Aggleton, J. 110, 124
Agras, W. S. 302
Aigen, K. 182, 187, 192
Aitken, K. J. 193
Akrich, M. 164
Alberg-Hulten, G. 301
Albersnagel, F. 258, 373
Aldridge, D. 182, 184, 187, 192
Allanbrook, W. J. 54
Alperson, P. 25
Alson, D. 193
Altshuler, I. M. 183
Alvin, J. 181
Anand, P. 283
Anastasi, A. 407
Anderson, A. 124
Anderson, E. 291
Anderson, J. D. 251
Andrew, C. 110
Ansdell, G. 182, 187, 192, 196, 197
Appadurai, A. 139
Appels, A. 301
Araya 345n
Aristotle 173
Arnheim, R. 217
Arnott, J. L. 254
Ash, R. L. 278
Ashley, J. 302
Asmus, E. P. 402
Austin, J. L. 65n
Avstreih, Z. A. K. 194

Baars, B. J. 262
Bach, C. P. E. 47, 47n, 275, 282
Baer, T. 400
Balasaraswati, T. 144
Balch, W. 373
Balkwill, L. L. 114, 233, 241, 242, 324
Banse, R. 254
Barker, W. 110
Barlow, H. 207
Barnard, M. 137
Barnes, B. 162
Baroni, M. 324
Barrett, A. 123
Barrett, L. F. 250, 257, 263
Barry, N. 286
Bartel, L. R. 293
Bartlett, D. 84, 125, 374, 376
Barton, A. 320, 378
Batson, C. D. 75
Bauer, R. 112
Baxandall, M. 137
Baylis, G. 109
Bayne, R. 426
Bechara, A. 110
Becker, A. L. 140
Becker, J. 9–10, 57n, 98, 136, 167, 282, 322, 416, 454, 455, 456
Becklen, R. 262
Beebe, B. 193
Beever, A. 33
Behne, K. E. 84, 232, 233, 324, 383
Behrens, G. A. 187, 188, 241, 313
Belleville, S. 116, 127
Bellugi, U. 127
Benamou, M. 140
Benveniste, E. 140
Beran, J. 409
Berezin, M. 163
Bergmann, G. 325
Berkowitz, L. 88, 378
Berlyne, D. E. 76, 82, 211, 212, 214, 219, 228, 401
Bermudez, P. 120
Bernieri, F. J. 194
Berntson, G. C. 375, 382
Besson, M. 119
Best, C. 112
Bharucha, J. J. 93
Biancolli, L. 205
Bigand, E. 234, 242
Birke, L. 161

Bitti, P. E. R. 227–8
Blacking, J. 140
Blair, R. 110
Blaney, P. H. 373, 426
Blasius, L. D. 65n
Bleske, A. 115
Bless, H. 75
Bloem, G. M. 374, 380
Blonder, L. X. 456
Blood, A. 120, 122, 124
Bloomfield, T. 49n
Blum, D. 309
Bohlman, P. 66
Boiten, F. A. 371, 372
Bolivar, V. J. 258, 261
Boltz, M. 256, 258, 263
Bond, A. 379
Bonfiglioli, L. 227–8
Bonnel, A.-M. 118
Bonny, H. L. 185
Borchgrevink, H. M. 367
Born, G. 167
Bornstein, M. H. 322
Borod, J. C. 321
Bouchard, B. 114, 229, 233, 330
Bouhuys, A. L. 374, 380
Bourdieu, P. 137, 163, 165
Bourne, E. J. 139, 141
Bouwsma, O. K. 32
Bower, G. H. 95, 279
Bowers, D. 112, 123
Bowler, A. 164
Box, G. E. P. 405, 408, 409
Boyd, J. 276, 283
Boyle, J. D. 394
Bradley, M. M. 383
Brammer, M. 110
Braudy, L. 277
Bregman, A. 259, 260
Breiter, H. 110
Brennan, T. 137
Bresin, R. 317
Brittin, R. V. 396, 397, 398, 403
Broca, P. 105, 110
Brochard, R. 6
Brodsky, W. 292, 303, 416
Broks, P. 110
Brook, B. S. 216
Brooks, V. 249
Brown, D. E. 138
Brown, J. D. 374, 380
Brown, J. J. 194
Brown, R. 253, 265, 266
Brown, S. 115, 192
Bruner, G. C. 93, 383
Brunswik, E. 320, 324, 325, 328
Bruscia, K. E. 182, 185, 186, 187, 191
Bryden, M. P. 119, 123, 124, 321
Bryson, B. 165, 174
Buck, R. 126, 313
Buckner, R. 110
Budai,R. 116
Budd, M. 25, 34, 58, 81, 309, 310, 329, 368, 373
Bukofzer, M. F. 282
Bullerjahn, C. 255, 258, 266
Bunt, L. 11, 79, 81, 89, 114, 182, 185, 187, 192, 258, 383, 426, 458
Burnham, S. 48
Burns, S. J. 185
Burr, V. 416
Burt, G. 265
Buss, D. 115
Butler, J. P. 65n, 66n
Byers, P. 371, 378
Byrnes, S. R. 396, 398, 407

Cabrera, D. 400
Cacioppo, J. T. 39, 74, 84, 98, 329, 375, 382
Calder, A. 108, 110, 124
Callner, A. 301
Cameron, D. 164
Camp, C. J. 250, 258, 373
Cantor, J. R. 364
Capperella-Sheldon, D. A. 396, 398, 403, 404, 405
Capra, F. 195
Carlsson, F. 329, 376, 378
Carmean 378
Carroll, B. G. 266, 267
Carroll, N. 38
Carterette, E. C. 309
Cavicchi, D. 49, 167
Cazden, N. 229
Cerulo, K. A. 216, 219
Cesarec, Z. 225, 226, 229
Ceschi, G. 385
Champion, L. 195
Chan, A. 107
Chapple, E. D. 151
Cherlow, D. 110
Chung, J. W. 97
Churchland, P. S. 151
Cicero, B. A. 321
Clapton, E. P. 309
Clark, D. B. 302
Clark, E. 114
Clark, M. 185
Clarke, E. F. 67, 263, 278, 312
Clore, G. L. 90, 93, 139
Clynes, M. 93, 151, 227, 278, 320, 378, 397
Coffey, P. 110
Cohen, A. J. 13, 249, 254, 256, 257, 258, 259, 260, 261, 263, 425, 459
Cohen, D. 106
Cohen, S. 167, 296, 297
Coker, W. 29
Colletti, J. 424
Collier, G. L. 402
Collins, A. 90, 93, 139
Collins, R. 162
Collins, S. C. 402

Condon, W. S. 151, 194
Cone, E. T. 50n
Conifer Classics 383
Contrada, R. J. 369
Conway, A. 296, 297
Cook, D. 444
Cook, N. 5, 8, 51n, 67, 97, 125, 172, 223, 229, 250, 257, 263, 267, 282, 399, 453, 456, 458
Cooke, D. 29, 54–7, 229
Cooksey, R. W. 326, 327
Cooley, C. H. 162, 171
Coomaraswamy, A. K. 143
Cooper, C. L. 300
Copland, A. 253
Cormack, J. 140
Corrigan, B. 326
Coslett, H. 123
Cosmides, L. 454
Costa, M. 227–8
Cotter, V. M. 192
Cox, H. 149
Crafts, S. 167
Craig, K. D. 293, 295, 296, 297, 303
Crain, W. C. 320, 378
Crandall, P. 110
Craske, M. G. 293, 295, 296, 297
Crofton, I. 275
Crowder, R. G. 89, 115, 229, 234
Crucian, G. 123
Csikszentmihalyi, M. 283, 417, 424, 432, 446
Cunningham, J. G. 89, 114, 323
Curtis, A. 110
Cuthbert, B. N. 383

Dalgleish, T. 76, 77
Dalla Bella, S. 115
Dalton, K. M. 369
Damasio, A. 74, 78, 106, 108, 110, 112, 140, 145
Damasio, H. 108, 110, 112
Damos, D. 401
Danly, M. 321
Darwin, C. 73, 106, 115, 320
Davidson, J. W. 329
Davidson, P. O. 303
Davidson, R. 6, 75, 89, 112, 139, 386
Davies, J. B. 94, 184
Davies, P. 296, 297
Davies, S. 7–8, 25, 29, 32, 34, 35n, 38, 39, 50, 55n, 58, 58n, 223, 309, 329, 368, 373, 457
Davis, M. 110
Davis, P. J. 427
Davis, T. 293
Dawes, R. M. 326
Day, H. I. 298, 299
De Backer, J. 182
DeGelder, B. 322
de Haan, E. 108
Deigh, J. 27
De Las Hera, V. 167
Dennett, D. 151
DeNora, T. 10–11, 97, 141, 165, 168, 169, 171–2, 175, 368, 416, 420–1, 421, 422, 423, 453, 454
Desberg, P. 302
De Sousa, R. 74
Deutsch, D. 3, 4
de Vries, B. 227, 378, 397, 407
Dewey, J. 82
Dews, C. L. B. 300
Diacoyanni-Tarlatzis, I. 107
Dibben, N. 8, 49n, 125, 172, 223, 229, 282, 399, 453, 456, 458
Dijkstra, K. 255, 258
DiMaggio, P. 165
Dimberg, U. 367
Dolan, R. 110, 124, 126
Dolder, M. 303
Dolgin, K. 89, 114, 323
Don, A. 127
Dowling, W. J. 3, 4, 82, 93, 94, 362, 443, 445
Downey, J. E. 224, 225
Dreyfus, H. L. 153
Driver, B. L. 369
Driver, J. 261
Duffy, E. 296, 374
Duke, R. A. 397
Duncan, E. 95, 96
Duncan, J. 195
Dunsmuir, W. 399, 404, 408
Dutta, S. 95

Eagle, C. T. 151
Edelman, G. 151
Edgerton, C. L. 182
Edmister, W. 109
Eerola, T. 92, 211
Eggebrecht, R. 322
Eibl-Eibesfeldt, I. 107, 322
Eich, E. 258
Einstein, A. 218
Eiseman, F. B. 146
Eisenberg, N. 370
Eisler, H. 56
Ekman, P. 6, 37, 74, 75, 76, 77, 89, 106, 107, 115, 139, 314, 322, 371, 372, 374, 375, 376, 378, 382, 454
Elder, S. T. 373
Ellsworth, P. C. 384
Eneroth, P. 301
Enserink, M. 352n
Erdonmez Grocke, D. 185
Erhan, H. M. 321
Ericsson, K. A. 265, 447
Etcoff, N. 108, 110, 124, 367
Evans, A. 120, 124
Eyerman, R. 175

Fabrigar, L. R. 244
Fairbanks, G. 312
Faith, M. 263
Farnsworth, P. R. 216, 244, 368

Featherstone, M. 161, 163
Fehr, B. 73
Feld, S. 57n, 140
Feldstein, S. 193, 194
Fentress, J. C. 258, 261
Ferstman, C. 140
Fidler, H. 293, 299, 301
Field, T. 106
Finarelli, L. 324
Fingret, A. 296, 297–8
Fish, S. 136
Fishbein, M. 293, 302, 304
Fisk, C. 64
Fisk, J. 223, 244, 311
Fitzgerald, D. F. P. 303
Flett, G. L. 298, 299
Flinn, C. 56
Flowers, P. J. 403
Fluegel, K. 111
Fodor, J. A. 92, 329
Foley, W. A. 152, 153
Folkman, S. 279
Fontaine, F. S. 116
Forgas, J. P. 75, 373
Fournier, M. 402
Fraisse, P. 88, 378
Fraser, D. 275
Frazer, E. 164
Fredrickson, W. E. 227, 396, 398, 401
Fredrikson, M. 295, 296, 297
Frego, R. J. D. 396
Friberg, A. 278, 317
Frick, R. W. 317, 321, 323
Fried, R. 88, 378
Friesen, W. 107, 371, 375
Frijda, N. H. 81, 83, 85, 87, 87–9, 92, 250, 263, 267, 371, 372, 375, 384, 459
Frith, C. 110
Frith, S. 5, 141, 166, 167
Frommer, J. 182
Fukuda, H. 123
Fussell, S. R. 266

Gabbard, J. O. 295
Gabriel, C. 229
Gabrielsson, A. 4, 12–13, 18, 19, 74, 81, 84, 87, 89, 93, 94, 95, 96, 113, 122, 145, 187, 219, 223, 224, 225, 226, 227, 229, 239, 250, 278, 280, 283, 310, 312, 313, 314, 323, 324, 325, 344n, 361, 368, 381, 394, 397, 416, 431, 434, 435, 445, 447, 455, 459
Gagnon, L. 114, 116, 118, 119, 122, 229, 233, 309, 330
Gamman, L. 137
Gardner, H. 4, 266
Gardner, W. L. 74, 98
Garfinkel, H. 168
Garpike, A. 409
Gaschke, Y. N. 379
Gaudreau, D. 118
Gaver, W. W. 3, 76, 82, 92
Gay, D. 368
Geertz, C. 139, 140, 349
Gembris, H. 89
George, M. 124
George-Warren, H. 276, 283
Gerardi, G. M. 115, 234, 241
Gergen, K. J. 136, 141, 424
Geringer, J. M. 394, 398, 403
Gerken, L. 115, 234, 240
Ghazzali, A. H. M. 146
Gibson, D. 110
Gibson, J. J. 67, 325
Giddens, A. 162
Gigerenzer, G. 77
Giles-David, J. 142, 173
Gilfillian, D. P. 97
Gillett, G. 416
Gilman, B. I. 224, 225
Gjerdingen, R. O. 54
Glasberg, B. P. 400
Glass, D. C. 369
Gnoli, R. 143
Godlovitch, S. 25
Goehr, L 25, 47–8
Goffman, E. 141, 372
Goins, W. E. 394, 397, 400
Goldhill, S. 137
Goldman, A. H. 25, 32, 35n, 38
Goldman, N. S. 299, 301
Goldsmith, H. 386
Goldstein, A. 84, 91, 125, 126, 396, 433
Goldstein, D. G. 77
Gomart, E. 168–9, 170, 175
Goodenough, W. H. 138
Goodman, N. 28n
Gorbman, C. 254, 410
Gordon, R. 26
Gorman, B. S. 320, 378
Gorsuch, R. L. 293
Gorton, T. L. 303
Gosselin, N. 115
Gotlieb, H. 234
Gottfries, C-G. 302
Gottman,J. M. 406, 408, 409
Gracyk, T. A. 25
Gray, J. 110
Green, L. 424
Green, S. B. 187, 188, 241, 313
Greenberg, R. 106
Greene, B. A. 394
Gregory, A. 115, 234, 257, 275, 322, 396, 398
Gregory, D. 393, 396, 404, 406
Griffiths, P. 27n, 38
Groll-Knapp, E. 296
Groothuis, T. G. 374, 380
Gross, J. J. 83, 84
Grossberg, S. F. 262
Grotowski, J. 145
Grunwald, I. S. 321
Grzywacz, M. 105

Guck, M. A. 64–5
Güldenring, M. 255, 258, 266
Gundlach, R. H. 224, 225, 232, 235, 240
Gunnarsson, R. 295, 296, 297
Gurney, E. 377

Hagegård, H. 313
Haider, M. 296
Hair, H. I. 394
Halgren, E. 110
Hall, D. E. 400
Hall, E. 151
Hall, S. 166
Hallam, S. 276
Halsey, R. S. 213
Hamann, S. 110
Hamilton, J. D. 405
Hammond, K. R. 326, 327
Hanaken, E. A. 424
Handel, S. 321, 325
Hanley, J. 110
Hanser, S. B. 182
Hanslick, E. 47, 48, 49–50, 58, 375
Haraway, D. 161
Harbuz, M. S. 185
Hargreaves, D. J. 82, 97, 374, 393, 422, 424
Haroutounian, J. 285
Harré, R. 416, 427
Harrer, G. 76, 88, 330, 378
Harrer, H. 76, 88, 330, 378
Harris, P. L. 89
Harwood, D. L. 3, 4, 82, 93, 94, 362, 443, 445
Haselton, M. 115
Hasselmo, M. 109
Hatfield, E. 39, 86, 329, 375, 382
Hatoh, T. 227, 394, 395
Hatten, R. 62–3, 65
Hattori, M. 313, 324
Haviland, J. M. 3, 6
Hay, D. 108
Heal, M. 182
Heaney, C. J. 182
Heaton, P. 127
Hebdige, D. 137
Heider, K. 107, 256
Heilman, K. 112, 123, 124
Heinlein, C. P. 228
Heinmiller, B. 120
Heitland, K. W. 286
Heldmann, M. 375
Hellawell, D. 110, 124
Helmerson, F. 283
Helmholtz, H. L. F. von 321, 377
Henderson, C. 320, 378
Hennion, A. 168–9, 170, 175
Hepworth, M. 161, 163
Heritage, J. 168
Hermelin, B. 127
Hess, U. 364
Hesse, F. W. 379
Hetherington, K. 163
Hevner, K. 5, 93, 224, 229, 230, 232, 235, 240, 241, 242, 243, 244, 381, 394, 398, 400, 402, 409
Hewitt, P. L. 298, 299
Hietanen, J. K. 329
Higgins, K. 25
Hill, D. S. 233, 234
Hilmersson, P. 329, 376, 378
Hilton, W. F. 369
Himle, D. P. 303
Hischer-Buhrmeister, M. 324
Hjort, M. 25
Hochberg, J. 249
Hochschild, A. 171, 372
Hodges, J. 108, 110
Hoehn-Saric, R. 303
Holbrook, M. 95, 97, 283
Holmes, T. H. 216
Holmes, T. S. 216
Holsti, O. R. 206
Hömberg, V. 108, 110
Horan, J. J. 303
Hornak, J. 123, 124
House, M. 258
Howell, D. C. 407
Huber, K. 228
Huckelbridge, P. 185
Hudson, J. A. 303
Hudson, R. 330
Hugdahl, K. 295
Hume, D. 41
Huntley, A. 114
Hurford, P. 275
Hursch, C. J. 326, 327
Hursch, J. L. 326, 327

Imaizumi, S. 123, 323
Imberty, M. 81, 226, 242
Irvine, J. T. 139, 142
Irwin, W. 112
Isaac, C. 110
Ivaldi, A. 417, 421, 427
Iwamiya, S. 260, 263
Iwarsson, J., 313
Izard, C. 76, 77, 107, 322, 374

Jackendoff, R. 29, 50, 92, 151, 398, 400, 458
Jackson, J 112, 123
Jackson, S. 164
Jacob, F. 349
Jaffe, J. 193
Jakobson, R. 59
James, I. M. 302
James, W. 73, 74, 144
Jamieson, A. 175
Japanese Society of Music Perception and Cognition 6
Jay, M. 137
Jefferson, T. 139
Jenike, M. 110

Jenkins, G. M. 405, 408, 409
Jenkins, J. M. 75, 322
Jenks, C. 137
Johnson, C. M. 407
Johnson, J. H. 65, 137
Johnson, M. 110, 124
Johnson-Laird, P. N. 74, 77, 322
Johnstone, T. 122, 366, 385
Jokinen, M. 300
Jourdain, R. 264
Jürgens, 321, 367
Juslin, P. N. 6, 8–9, 14–15, 35n, 77, 82, 83, 86, 87, 93, 94, 98, 107, 113, 114, 122, 124, 144, 161, 184, 187, 191, 194, 223, 224, 230, 232–3, 235, 240, 242, 243, 250, 257, 263, 278, 283, 285, 309, 311, 312, 313, 314, 315, 316, 317, 318–19, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 361, 368, 374, 384, 394, 401, 418, 426, 427, 431, 446, 453, 454, 455, 457, 458, 459

Kagan, J. 120, 367
Kainz, F. 362
Kalinak, K. 258, 261, 266–7
Kamenetsky, S. B. 233, 234
Kamiya, J. 282
Kammann, R. 212, 214
Kanner, A. D. 279
Kantra, S. 256, 258, 263
Kanungo, R. N. 95
Kappas, A. 364
Kappell, L. 303
Karl, G. 34, 60–1, 63, 64
Karlin, F. 264, 266
Karnat, A. 108, 110
Karno, M. 234
Kaser, H. E. 303
Kastner, M. P. 89, 115, 234
Kato, M. 227, 394, 395
Katz, R. 145
Kawashima, R. 123
Keegan, J. 296, 297–8
Keil, C. 58, 145, 167
Keith, T. A. 97
Kelly, G. A. 284
Keltner, D. 83
Kemp, A. E. 286
Kemper, T. D. 77
Kendall, R. 257, 265–6, 309
Kendrick, M. J. 303
Kenealy, P. 84, 282, 373, 379, 380
Kennedy, N. 284, 309
Kennedy, W. 110
Kenyon, N. 368
Kerman, J. 45, 51
Kern, S. 137
Ketter, T. 124
Kimbrell, T. 124
King, B. B. 310
Kingsbury, H. 275
Kintsch, W. 262
Kiritani, S. 123
Kirkpatrick, R. 50–1
Kirson, D. 314
Kivimäki, M. 299, 300
Kivy, P. 25, 29, 31, 32, 34n, 35, 35n, 38, 38n, 40, 47n, 50, 57, 59–61, 82, 105, 321, 361, 380
Kleiber, D. 424
Klein, D. 375, 382
Kleinen, G. 224, 225, 226, 229
Kleinginna, A. M. 74
Kleinginna, P. R. 74
Kneller, G. F. 283
Kneutgen, J. 378
Kogan, N. 115
Konecni, V. J. 81, 97, 234, 242
Konishi, T. 323
Kotlyar, G. M. 312, 313, 314, 324
Kozal, M. J. 369
Kracauer, S. 250, 252
Kramer, L. 51n, 66, 66n
Kratus, J. 89, 115, 323, 324
Krauss, R. M. 266
Krippner, S. 283
Krumhansl, C. L. 84, 93, 94, 113, 125, 205, 218, 219, 227, 229, 242, 263, 375, 396, 398, 399, 400, 401, 404, 408
Kubey, R. 424
Kuhn, T. S. 5, 460
Kuipers, P. 87
Kuwano, S. 227, 394, 395

Ladd, D. R. 325
Lader, M. 379
Lamont, M. 165
Landowska, W. 209
Lang, P. J. 257, 369, 383
Lange, H. 108, 110
Langeheinecke, E. J. 324
Langenberg, M. 182
Langer, S. 31, 58, 79, 81, 172, 194, 223, 228, 243
Larkin, K. T. 297
Larsen, R. 424
La Rue, J. 216
Lash, S. 168
Laski, M. 432, 433, 446
Latour, B. 164
Laukka, P. 314, 320, 322, 326, 327
Laver, S. 25
Law, J. 162, 164, 168
Lawson, D. W. 303
Lazarus, R. S. 73, 78, 81, 90, 279
Leach, E. 138
Leahey, T. H. 71
Ledden, P. 109
Lederman, R. J. 302
LeDoux, J. 27n, 85, 106, 110, 118, 121, 367, 369, 455
Lee, C. A. 182, 183, 184, 192
Lee, J. 111
Lee, M. 110
Lefevre, J. 417

Lehmann, A. C. 278, 319, 330, 396, 398, 401, 403, 407
Lehrer, P. M. 299, 301
Leming, J. 424
Lerdahl, F. 29, 50, 151, 398, 399, 400, 458
Levenson, R. 263, 371, 375
Leventhal, H. 366
Levi, D. S. 250
Levin, D. N. 369
Levin, D. T. 253
Levinson, J. 25, 26, 29, 34, 35n, 40, 87, 235, 258
Levinson, R. W. 84
Levi-Strauss, C. 138
Levitin, D. 127
Levy, J. M. 358
Levy, R. I. 139
Lewis, G. H. 424
Lewis, M. 3, 6
Ley, R. 119, 123, 124, 321
Liden, S. 302
Liebert, R. M. 72, 82
Limbacher, J. L. 250
Lindauer, M. S. 217
Lindström, E. 12–13, 94, 122, 219, 227, 233, 241, 242, 280, 283, 310, 313, 314, 319–20, 323, 325, 344n, 361, 381, 394, 397, 416, 431, 455
Lindström, S. 89, 250, 434, 435, 445, 447
Lindström Wik, S. 434, 435, 445
Linnankoski, I. 329
Lippman, E. A. 381
Lipps, Th. 370
Lipscomb, S. 257, 258, 265–6, 406
Lovejoy, A. O. 139
Lowis, M. J. 447–8
Loy, G. D 409
Lubin, B. 379
Luck, G. 6
Ludwig, A. M. 282
Lundqvist, L. G. 329, 376, 378
Lushene, R. E. 293
Lutz, C. 139, 415
Lychner, J. 396, 398
Lynxwiler, J. 368
Lyons, W. 26

McClary, S. 51, 51n, 139
McCormack, J. 424
McDaniel, J. 143, 144
MacDonald, B. 108, 109
McDonald, M. 276
Machotka, P. 212
McInerney, S. 110
Mcintosh, D. N. 371
Mack, A. 253
MacLean, A. 369
MacLean, P. D. 321
McNeill, W. H. 151m 152
McRobbie, A. 174
Madell, G. 35n, 40n
Madison, G. 263, 313, 314, 318–19, 327
Madsen, C. H. 192
Madsen, C. K. 192, 227, 257, 396, 397, 398, 401, 402, 403
Magliano, J. P. 255, 258
Maher, T. F. 227
Maisonneuve, S. 168, 175
Makarec, K. 402
Malik, F. 299, 301
Malloch, S. N. 193
Mandler, G. 3, 75, 76, 82, 92
Manfredo, M. J. 369
Mankowski, T. 374, 380
Manturzewska, M. 277
Maratos, E. 110
Marchant-Haycox, S. E. 293
Marchini, C. 116
Marin, O. 105
Marks, M. 54, 249
Markus, H. 382
Marshall, S. 256, 257, 258, 260, 263
Marshment, M. 137
Marston-Wyld, J. 185
Martin, M. A. 373
Martindale, C. 211, 212, 214, 218
Masataka, N. 114
Maslach, C. 379
Maslow, A. 283, 431–2, 433, 446, 448
Masson, J. L. 143
Matravers, D. 33, 39, 40n
Mattheson, J. 46, 375
Maturana, H. 140, 152, 153
Maughan, M. L. 303
Maus, F. E. 61–2, 62n, 63
Mayer, J. D. 74, 88, 379
Mayo, E. 162
Mays, K. 193
Mazzola, G. 409
Mazzoni, M. 116
Mazzuchi, A. 116
Mellman, T. A. 369
Melucci, A. 163
Melzack, R. 140
Merker, B. 115
Merriam, A. 145, 419
Mesquita, B. 384, 459
Metha, A. 373
Metzger, L. J. 369
Meyer, L. B. 15–16, 18, 29, 57–9, 76, 82, 91, 92, 144, 205, 210, 216, 219, 260, 264, 265, 278, 309, 311, 330, 344n, 347, 348n, 349n, 350n, 352n, 353n, 354, 357n, 358n, 380, 455, 456, 457, 458, 459
Meyer, R. G. 302, 304
Meylan, S. 379, 380, 381
Middlestadt, S. 293, 302, 304
Middleton, R. 55, 55n, 56, 172
Miller, G. 115
Miller, G. A. 369
Miller, L. 127
Mirowksy, J. 293
Mitchell, W. J. T. 137

Moles, A. 207
Monson, I. 66n
Montello, L. 182
Moog, H. 88
Moore, A. F. 56, 67
Moore, B. C. J. 400
Moore, K. 212
Moore, L. J. 164
Mor, S. 298, 299
Moretti, P. 116
Morgenstern, S. 207
Mori, K. 123
Morley, D. 137
Morozov,V. P. 312, 313, 315, 324
Morris, J. 110, 124, 126
Moss, A. 108
Motte-Haber, H. de la 229, 239, 276
Mottron, L. 127
Mulcahy, D. 296, 297–8
Mulder, S. J. 402
Mull, H. 394, 396, 402
Münsterberg, H. 217, 251, 258, 264, 266, 267
Muratorio, A. 116
Murray, I. R. 254
Myers, D. M. 373
Myers, F. R. 139

Näätänen, R. 296
Nagel, J. J. 275, 277, 303
Namba, S. 227, 394, 395, 400
Narmour, E. 91, 92, 260, 344n, 398
Neale, J. M. 72, 82
Needham, R. 140
Neftel, K. A. 303
Neher, A. 151
Neilly, L. 167
Neisser, U. 105, 262
Nettelbeck, T. 320, 378
Nettheim, N. 229, 397
Nettl, B. 140
Neumann, R. 329
Newcomb, A. 59–60, 62
Newcombe, F. 108
Newman, J. 260, 261
Nicolson, P. 426
Niedenthal, P. M. 85, 330
Nielsen, F. V. 224, 226, 227, 242, 396, 398, 399, 400, 401, 404, 409
Nielzén, S. 225, 226, 229
Niemann, B. K. 303
Niimi, S. 323
Nisbett, R. E. 385
Nordoff, P. 181, 192
North, A. 82, 92, 97, 374, 422, 424
Noyes, R. 293, 303
Nunez, R. E. 151, 152
Nuzzo, C. 322
Nyklicek, I. 78, 84, 125, 376

Oatley, K. 6, 74, 75, 76, 77, 79, 81, 95, 96, 322
Obler, L. K. 321
O'Connor, C. 314
O'Connor, N. 127
Odell-Miller, H. 184, 191
Ogston,W. D. 194
Ohgushi, K. 278, 312, 313, 323, 324
Öhman, A. 366, 367
Ohta, S. 108, 109
Oleron, G. 88, 378
Oleson, K. C. 75
Olivier, L. 294
O'Malley, A. 320, 378
O'Neill, J. 81, 83, 97, 137
O'Neill, S. A. 17–18, 127, 145, 167, 169, 183, 224, 368, 417, 421, 424, 427, 453, 454, 455, 459
Ornstein, R. 374
Orr, S. P. 369
Ortony, A. 77, 90, 93, 139
Osborne, R. 277
Osgood, C. E. 228, 257, 401
Oshinsky, J. 93, 229–30, 235, 244, 317
Oster, H. 107
Ostrom, C. W. 408, 409
O'Sullivan, M. 107

Paillard, J. 88, 378
Paisley, W. J. 206, 207
Palmer, C. 4, 250, 252, 258, 263, 264, 267, 309, 312
Panksepp, J. 77, 84, 91, 124, 126, 321, 322, 396, 433
Panzarella, R. 432, 433, 438, 446
Papez, J. 110
Papotto, C. 373
Papou_ek, H. 193, 322
Papou_ek, M. 193, 322, 323
Papsdorf, J. D. 303
Parasuraman, S. 299
Pardo, J. 110
Pardossi, L. 116
Parekh, P. 124
Pareto, W. 162
Park, A. 296, 297–8
Parma, M. 116
Parncutt, R. 278, 396, 398, 401, 403, 407
Parry, F. 108
Patel, A. D. 321
Pater, W. 48
Patton, M. Q. 434
Patwardhan, M. V. 143
Pauls, C. 375
Pavlicevic, M. 11, 79, 81, 89, 114, 182, 184, 187, 192, 194, 195, 258, 383, 426, 458
Pay, C. 299, 301
Pearson, P. 299, 301
Pecchinenda, A. 385
Pelland, R. 402
Penfield, W. 321
Penhune, V. 120, 122
Pennisi, E. 352n

Peretz, I. 9, 114, 115, 116, 118, 119, 120, 122, 127, 151, 229, 232, 233, 321, 330, 381, 453, 454, 456, 459
Perrett, D. 108, 110
Perry, D. 105
Perski, A. 301
Persson, R. S. 14, 145, 275, 276, 277, 278, 279, 282, 283, 284, 285, 309, 310, 311, 320, 330, 416, 432, 458
Peterson, R. 165, 174
Phelps, E. 124
Phillips, M. 110
Piedrahita, L. E. 218
Pignatiello, M. 250, 258, 373
Pike, A. 84
Pinker, S. 115
Pitman, R. K. 369
Plato 173
Ploog, D. 321, 322
Plutchik, R. 4, 74, 76, 77, 79, 315, 316, 317, 321, 432, 446
Pohlmann, K. C. 403
Pollock, L. R. 426
Porter, C. A. 217
Power, M. 76, 77
Pratt, C. 216, 377
Pratt, G. 275, 276, 278
Pratt, R. R. 303
Prendergast, R. M. 249
Press, A. 168
Preston, J. 258
Preston, J. M. 258
Price, C. 299, 301
Priestley, M. 181, 187, 192
Pring, L. 127
Pugliolo, M. 116
Pujol, K. K. 378
Purohit, Y. S. 299

Qian, J. 258
Queen, H. 112
Qureshi, R. B. 146, 154

Racy, A. J. 140
Radley, A. 175
Radocy, R. E. 394
Radway, J. 168
Raffman, D. 151
Rahe, R. H. 216
Rapoport, E. 278, 314, 328
Rapp, B. 108, 113
Rapson, R. L. 39, 329
Rasar, L. 250, 258, 373
Ratner, L. G. 53, 399, 400
Rauch, S. 110
Raymer, A. 123
Remington, N. A. 230, 244
Remmling, G. 162
Renshaw, P. 284
Repp, B. H. 312
Reubart, D. 304
Revonsuo, A. 260, 261
Rider, M. S. 151
Ridley, A. 34
Rigg, M. G. 93, 224, 225, 232, 235, 239, 240, 242, 243, 263, 283, 381
Robarts, J. 192
Robb, L. 193
Robbins, C. 181, 192
Roberts, L. 321
Robinson, C. R. 404
Robinson, J. 25, 34, 56, 60–1, 63, 64
Robitaille, B. 227, 241
Robson, C. 275, 276, 278, 309
Rochat, P. 372
Rock, I. 253
Rockland, C. 110
Rogers, C. R. 283
Rolls, E. 109, 123, 124
Rosaldo, M. Z. 138, 139
Rosar, W. 252, 364
Rosch, E. 152
Rosch, R. 78
Rosenthal, R. 194, 314
Rosinsky, N. 124
Ross, B. H. 322
Ross, C. E. 293
Rossi, M. 303
Rossing, T. D. 400
Rouget, G. 145
Rouleau, N. 127
Rourke, B. 127
Rousseau, J. J. 321, 330
Rousseau, L. 115
Rowland, D. 108, 110
Royer, J. M. 394
Rózsa, M. 264, 265, 266
Russ, S. W. 283
Russell, J. A. 73, 77–8, 139, 146, 230, 250, 257, 263, 315, 401, 402
Russell, P. A. 82, 89, 97
Russo, F. A. 255–6, 261, 264, 266
Ruud, E. 187, 192, 277

Saint-Saens, C. 258
Salmon, P. G. 302, 304
Salovey, P. 74, 88
Santesso, D. 119
Saperston, B. 182
Sarge, A. 115, 234
Saul, J. S. 108
Savage, I. 302
Savary, L. M. 185
Schacht, A. K. 365
Schacter, S. 86
Schechner, R. 65n, 145
Scheler, M. 162
Schellenberg, E. G. 127, 194, 367, 377
Schenck, D. L. 401
Schenker, H. 47

Scherer, K. R. 6, 16, 75, 84, 85, 90, 93, 96, 122, 125, 144, 219, 224, 229–30, 235, 244, 254, 313, 314, 317, 321, 322, 324, 328, 329, 341, 361, 362, 364, 365, 366, 367, 368, 369, 370, 372, 373, 375, 376, 379, 380, 381, 382, 384, 385, 386, 453, 454, 455, 456, 457, 459
Scherer, T. 375
Schermann, T. K. 205
Schindler, R. M. 95, 97
Schlosberg, H. 77, 230
Schmale, H. 296
Schmid 345n
Schmidt, C. P. 393, 394, 401, 406, 407
Schmidt, L. 119
Schmidtke, H. 296
Schmitt, E. 356
Schnitzier, H.-U. 324
Schögler, B. 195
Scholes, R. 124
Schorr, A. 366
Schubert, E. 16–17, 78, 94, 227, 243, 257, 319, 397, 399, 400, 402, 404, 405, 406, 407, 408, 409, 456
Schulkind, M. 256, 258, 263
Schulz, W. 293, 299
Schumacher, M. 309, 330
Schutz, A. 164
Schwartz, J. 314
Scott, S. 124
Scruton, R. 28n, 55n
Seashore, C. E. 279–80, 312
Seashore, H. G. 312
Seckel, A. 141
Seckerson, E. 309
Seddon, F. 6
Seeger, A. 140
Senior, C. 110
Senju, M. 278, 312, 323
Sergent, J. 108, 109
Sernagor, E. 105
Shackelford, T. 115
Shaffer, L. H. 309, 312
Shalev, A. Y. 369
Shammi, P. 105
Shankar, R. 142
Shapiro, B. E. 321
Shaver, P. 79, 314
Shaw, C. D. 409
Shaw, L. L. 75
Shepherd, J. 55n, 142, 172, 173
Sheridan, D 419
Shields, S. A. 314
Shore, R. E. 282
Shotter, J. 424
Showers, C. 85, 330
Shweder, R. A. 139, 140, 141
Siegwart, H. 328, 329
Sigala, F. 301
Silverman, K. E. A. 325
Simkus, A. 165, 174
Simmel, M. 256
Simon, H. 74, 447
Simons, D. J 253
Simonton, D. K. 12, 76, 206, 207, 208, 209, 210, 212, 213, 214, 215, 216, 217, 218, 219, 223, 455, 457
Sinclair, D. 255–6, 261, 264, 266
Singer, J. E. 86
Singer, P. 349
Sirius, G. 263
Slater, D. 137
Sloboda, J. A. 4, 6, 8–9, 17–18, 81, 82, 83, 84, 87, 89, 90, 91, 92, 95, 95–6, 97, 107, 124, 127, 144, 145, 161, 167, 169, 183, 184, 187, 194, 224, 229, 230, 243, 250, 257, 266, 278, 280, 303, 309, 310, 311, 319, 323, 326, 330, 342n, 355n, 364, 368, 374, 381, 384, 394, 396, 398, 399, 400, 401, 403, 404, 407, 416, 417, 418, 419, 421, 425, 426, 427, 433, 446, 448, 453, 454, 455, 459
Small, C. 5, 152, 183, 454
Small, M. 108
Smeijsters, H. 182
Smith, D. 162
Smith, J. 254, 262, 263
Smith, J. C. 140
Smith, L. D. 74, 228
Sobel, D. 374
Söderholm, M. 301
Solie, R. A. 139
Solomon, N. 52
Solomon, R. C. 26, 140, 457
Sonnemans, J. 383
Sony 383
Sorce, J. 86
Sorenson, E. 107
Spalding, T. L. 322
Sparrow, J. 296, 297–8
Speedie, L. 123
Spielberger, C. D. 293
Spies, K. 379
Spiro, M. E. 138
Spitzer, J. 65–6
Spivak, G. C. 139
Sprengelmeyer, R. 108, 110
Squire, L. 110
Stahl, G. 379
Standley, J. M. 182
Stanislavsky, K. 145, 281
Steck, L. 212
Stefanicci, L. 110
Stein, M. I. 283
Stemmler, G. 375
Steptoe, A. 14, 293, 295, 297, 299, 300, 301, 458
Sterling, R. S. 89, 114, 323
Stern, D. 79, 193, 194, 195, 243
Stern, I. 278
Sternagel, S. 377
Stone, R. 140
Storm, C. 257
Storm, T. 257
Strack, F. 329
Strayer, J. 370

Streeter, E. 192
Striano, T. 372
Strommen, E. F. 299, 301
Stuss, D. 105
Styge, B. 192
Suci, G. J. 228, 257, 401
Sudderth, J. 403
Suedfeld, P. 217, 218
Sugerman, J. 119, 321
Sugiura, M. 123
Sundberg, J. 312, 313, 321
Surakka, V. 329
Sweeney, G. A. 303
Symmes, D. 322
Symonds, E. R. C. 303

Tagg, P. 54, 55–6, 55n
Tait, M. 310
Talavage, T. 109
Tan, W. 250, 263, 264, 266, 267
Tannenbaum, P. H. 228, 257, 401
Tarrant, M. A. 369
Tasker, S. 119
Taylor, C. 141
Teasdale, J. 366
ter Schure, E. 87
Terwogt, M. M. 89, 114, 323
Thayer, J. F. 77, 84, 125, 263, 376
Theorell, T. 301
Thom, P. 25
Thomas, T. 250
Thompson, E. 152
Thompson, E. G. 293
Thompson, W. F. 114, 227, 233, 241, 242, 255–6, 261, 264, 266, 324
Tillman, B. 234, 242
Todd, N. 312, 377
Todd, P. M. 409
Tokeshu 345n
Tolbert, E. 57n
Tolkmitt, F. 325
Tomkins, S. S. 76, 376, 378
Tönnies, F. 162
Tooby, J. 454
Tormey, A. 32
Tota,A. L. 168
Trainor, L. 114, 119, 120, 194
Tranel, D. 108, 110, 112, 124
Trehub, S. E. 114, 126, 194, 233, 234, 316, 367, 377
Treitler, L. 51, 141, 143
Trevarthen, C. 182, 193, 195
Trussoni, S. J. 320, 378
Tsang, C. 119
Turino, T. 140
Turner, B. S. 161, 163
Turner, T. J. 77
Turner, V. 138, 175
Tyler, H. 192
Tyler, P. 397, 402

Uemura, A. 211, 218
Unyk, A. M. 194
Urry, J. 168, 175

Vaitl, D. 377
Valentine, E. R. 303
Van der Kolk, B. A. 369
van Doornen, L. J. P. 84, 125, 376
Van Goozen, S. 384
van Grinsven, F. 89, 114, 323
Van Heesch, N. C. A. 293
Van Kemenade, J. F. L. M. 293
van Langenhove, L. 427
Van Reekum, C. M. 385
Van Son, M. J. M. 293
Varela, F. J. 140, 152, 153
Varney, N. 275
Vaughn, K. 151
Vehrs, W. 377
Vermazen, B. 34
Visser, P. S. 244
Vista, M. 116
Viswanathan, T. 140
Vitz, P. C. 212
Von Campenhausen. H. 141
von Cramon, D. 321
Vrana, S. R. 84, 329, 376
Vroomen, J. 322

Wade, D. 123, 124
Wagner, R. 49
Wakelfield, J. 115
Walker, J. 227
Wallace, R. 48
Wallbott, H. G. 313, 384
Wallin, N. L. 115, 151
Walser, R. 56
Walter, R. 110
Walter, V. J. 151
Walter, W. G. 151
Walton, K. L. 33, 34
Waterman, M. 84, 90, 91, 95, 278, 396, 403, 448
Watson, K. B. 224, 225, 232, 394
Watson, R. 124
Watt, R. J. 278
Webb-Talmadge, N. 329, 376
Weber, M. 161, 162, 164
Wedin, L. 93, 224, 225, 226, 229, 283, 381, 402
Weick, K. E. 97
Weir, D. 402
Weld, H. P. 394
Welkowitz, J. 194, 321
Wells, A. 424
Wertheimer, M. 260
Wesikoff, R. 109
Wesner, R. 293
West, R. 182
Westermann, R. 379
Whalen, P. 110

Wheeler, B. 182
Whissell, C. M. 315, 316, 402
White, H. 356
Wicke, P. 172
Widmer, S. 296, 297
Wientjes, C. J. E. 371, 372
Wigram, T. 182
Williams, J. M. G. 426
Williams, M. S. 300
Williams, R. N. 74, 228
Williams, S. J. 163, 171
Willis, P. 165, 166–7
Wills, G. I. D. 300
Willson, R. 320, 378
Wilson, G. D. 293
Wilson, T. D. 385
Win, Z. 299, 301
Witkin, R. W. 164, 168, 175
Witvliet, C. V. 84, 329, 376
Wolfe, M. L. 299, 301
Womer, J. 112
Woods, C. 6
Woods, F. A. 206
Woodson, R. 106
Woodworth, R. S. 77
Woody, R. 309, 310, 311
Woolgar, S. 164
Worrall, L. 115, 234
Wright, C. 296, 297–8
Wundt, W. 73, 77, 362, 376

Young, A. 108, 110, 124
Young, J. O. 143

Zajonc, R. B. 81, 85, 90, 118, 121, 279, 382
Zald, D. 111
Zatorre, R. 120, 122, 124
Zbikowski, L. M. 66–7, 66n, 351n
Zentner, M. R. 16, 84, 96, 120, 125, 144, 219, 224, 365, 367, 379, 380, 381, 401, 402, 453, 454, 455, 456, 457, 459
Zillmann, D. 83, 364
Zizek, S. 137
Zuck, E. 108, 109
Zuckerman, M. 282, 379
Zwaan, R. A. 255, 258

# SUBJECT INDEX


abstract form, of classical music 205–6
abstraction, of emotional experience 343
accuracy, of emotion communication in performance 313–14, 326
achievement motivation, of performer 277
action tendencies
  effect of music 377–8
  invoked by emotion 87–9
active music therapy 182–3, 186–7
activity dimension 226, 228, 257, 315–16, 402
  and circumplex model 77–8
  physiological effects and 376
adaptive function
  of emotional reactions 82, 88
  of maternal singing 114
  of music 115
adolescents, music and 424
advertisements, classical music in 257
aesthetic emotion 81–2
aesthetic motivation, of performer 278
'aesthetic reflexivity' 168
aesthetics
  automatic evaluation of 368
  and belief 352–3
  of film 258
  formulation of 47, 48
  and melodic originality 211–15
affect, *see* feeling
age effects 422
agency, emotional, and music 174–6
Allen, Thomas, on joy of performing music 291
altered state, and performer 281–3
ambiguity
  emotional 95
  of music 185
  resolved by film music 258
amygdala
  role in perception of facial expressions 110–11
  role in perception of vocal emotion 124
analysis, of data in continuous response measurement 405–9
analytic philosophy 25
*Analyzing musical multimedia* (Cook) 257
anthropology
  'being' and music and emotion 150–5
  culturally inflected listening 135–8
  emotion as cultural construct 138–40, 151
  person as cultural construct 140–1
  and psychological approach 150–2
  universality of arousal 144–50
*Apollo 13*, Horner score for 262
appraisal
  and central route production of emotion 366–9
  in SEM 447
  theories 90–1, 368
  *see also* automaticity, of emotion; evaluation
arbitrariness, of connections 29–30, 86, 367, 433
architecture, functional, *see* distributed neural structures; specialization
Aristotle 173, 377
arousal
  definition of 75–6
  ecstasy 145–6
  in everyday listening 418
  and habituation 214
  Pentacostal 149–50
  and performance anxiety 296–8
  physiological 110, 296–8, 374–7
  rage 146–8
  relationship with loudness 404
  and SEM 438
  universality of 144–50
arousal dimension, *see* activity dimension
arousal potential 211–14
arousal theory 32–4, 36, 39–40
articulation 242, 316, 317, 319
associationist congruence, in film music 258
associative sources of emotion 94–5, 172, 364
  and music therapy 184
  *see also* conditioning; connotations; extrinsic emotion; memory; stereotypes
asynchrony of levels 92
attention
  and everyday listening 418
  *see also* 'inattentional blindness'; 'inattentional deafness'
attitude
  cultural 97–8
  towards objects 26
  *see also* belief
attribution 87
  *see also* composer; listener; persona
attunement 79
autism 127
automaticity, of emotion 27, 85, 110, 115, 118, 322, 366–8
autonomic nervous system, *see* physiological arousal
autonomy, and everyday listening 421–3

avant-garde music 352, 353, 368
awareness
  emotion elicited outside 85
  and performance anxiety 303

Bach, J. S.
  *Ach Herr, mich armen Sünder*, different modes of 234
  *Art of the Fugue, The*, late-life effects and 218
  Brandenburg Concert No. 2, motor empathy and 358
  changing structure of 234–5
  *Herzlich tut mich verlangen*, different modes of 234
  Suite for Orchestra No. 2, minor mode sounding happy 240
  Toccata, Adagio and Fugue in C, used in music therapy 185
'background' music 422
Balinese bebuten trancing 146–9
baroque music, and linguistic models of musical expression 46–7, 51–3
Barry, John, on composing film music 265
basic emotions
  applicability to music 374–5, 381
  automaticity of 118
  definition 76–7, 314–15
  facial expressions of 107, 111
  in SEM 446, 447
  *see also* uniqueness, of musical emotion
BBC Symphony Orchestra, study on performance anxiety 296, 297–8
Beethoven
  on communication of emotion 205
  effect of stress on 217
  Late Quartets, and melodic originality 214
  and melodic originality 212, 214, 217
  Piano Concerto No.5, used in music therapy 185
  Piano Sonata in E Flat, Op. 81a ('Les Adieux'), suspense in 347
  Quartet Op. 95, analysis by Maus 61–2
  Quartet Op. 130, analysis by Hatten 62–3
  sonatas, changing structure of 234
  String Quartet No. 15 in A Minor, Op. 132, emotional response to 399
  Symphony No. 3 *Eroica*, discussed by philosophers 24
  Symphony No. 3 *Eroica*, and melodic originality 215
  Symphony No. 6, and listener uncertainty 347
  Symphony No. 9 in D minor, critical interpretation by Fröhlich 48
  Symphony No. 9 in D minor, critical interpretation by McClary 51
  and texts about music 48, 49
  'Waldstein' Sonata, analysis by Simonton 208
behaviour
  emotion and 26, 31
  learned 138
  and performance anxiety 295
behavioural component, of emotion 84–5
belief
  in emotion's object 26–7
  and perception of music 97, 352–3
'bikeboys', Willis study of 166
'binding', in cognitive science 260
biofeedback, and performance anxiety 303
biographical history, of composer 215–18
  *see also* composer, as experiencing subject
biological process, habitus of listening as 152–4
biological survival, and emotional reaction 82, 88
biology, *see* evolution; neuropsychology
*Blazing Saddles* (Brooks) 253
bottom-up processing, in cognitive framework for film music 260–1
Brahms, Four Serious Songs, late-life effects and 218
brain/body approach 151
brain damage, effect on musical emotion 116–18, 120–1
brain programmes, and emotion communication in performance 321–2
brain specialization
  for facial expressions 107–9
  for music 116–18
Brendel, Alfred, on reconceptualization of music 284
brief mood introspection scale (BMIS) 379
Brooks, Mel, and film music 253
Brunswikian lens model of acoustic code 324–8
buskers 423

catastrophizing, and performance anxiety 298–9
categorical approach 76–7, 315, 322
  and continuous response measurement 400–1
  *see also* checklist response format
categorization, of emotion's object 26
*catharsis*, music as 46, 49, 427
causal explanations, in psychology 71–2
causation, vs. implication 346–7
central route production of emotion 366–71
change
  cognitive 373–4
  in course of performance 317, 440
  of mood, *see* emotion construction; mood induction
  and uncertainty 345
  *see also* temporality, of music
characteristics, of emotion 83–91
checklist response format 394–6
children
  and ascending melody 241
  and major and minor modes 234, 239
  perception of music 114–15
  physical response to music 88
  using cues to express emotion 316
  *see also* infants
choice
  effect on emotional change 421
  necessity of 348–51
Chopin

F Minor Fantasia 210
Prelude N. 4 in E Minor, and continuous response measurement 407
chromatic notes, and melodic originality 216
circumplex model, of emotion 77–8, 244
classical music
in advertisements 257
aesthetics 211–15
composer's biographical history 215–18
melodic originality of 206–20
'understanding' 458
classical period, and melodic originality 215
classical writings, about music 46
*see also* Aristotle; Descartes; Plato
classification
of emotional states 343–4
in SEM project 435, 436
*see also* categorical approach
clinical improvisation, in music therapy 191–7
closure
and film music 255–6, 264
and syntactic and native processes 346
*see also* gestalt principles, of musical structure
code
for communicating emotion in performance 314–20
*see also* lens model of acoustic code
coding
of musical stimuli 400
of responses 400–2
cognition, and performance anxiety 295
cognitive appraisal, *see* appraisal
cognitive-behavioural treatment, for performance anxiety 303
cognitive changes, and music 373–4
cognitive explanations, in psychology 72
cognitive feedback, to improve expressive skills 327
cognitive framework, for film music 258–63
cognitive neuroscience *see* neuropsychology
cognitive theory, of the emotions 26–7, 31, 38–9, 50
cognitivists 105
collective action, music and 174–6
commotion, *see* empathy
communicating emotion in performance
accuracy of 313–14, 326
code used 314–20
cues used 314–17, 324–7
emotion categories 314, 314–17
emotion dimensions 314–17
emotion induction 329–30
expressive contours 317–20
history of research 311–12
lens model 324–8
performer's role 309–11
standard paradigm 312–13
theoretical framework 320–7
*see also* expression
communication, social 421
'community of interpretation' 136
comparative investigations, continuous response measures used for 398
'compensation' tendency, in performance 320
complex emotions 77, 316
component-process model of emotion 373
composer
biographical history of 215–18
as experiencing subject 32, 38, 223
of film music 264–6
intentions of 311
computer-generated music 278
conceptual blending 66–7
conceptual classification, of emotional states 343–4
conceptual flexibility, of performer 283–4
conceptualization, of music 279–81, 312
conditioning 29–30, 86, 367, 433
conditions, listening, *see* listening conditions
conflicting emotions, and SEM 440–1
congruence, in film music 260–1
connections
exemplificatory 30–1
*see also* sources of emotion
connotations 348
*see also* associative sources of emotion; memory
consciousness, altered state and performer 282–3
'conservatoire culture', and creativity 284
consonant music, infants preference for 367
construct
emotion as cultural 138–40, 151
person as a cultural 140–1
sociological focus on cultural 164
construction, of emotion 165, 166–7, 167–73
*see also* mood induction
contagion, emotional 329, 369–71
*see also* response, mirroring
content analysis, *see* melodic originality
context 136
in everyday listening 425–6, 427–8
'fit' with music 422
of music 97
*see also* listening conditions
contextual features, of production rules for emotional effects of music 364–5
continental philosophy 25
continuous ratings of expression 319
continuous response measurement
definition 393–4
evaluation of 399–409
response formats 394–7
use of 397–9
contour theory, of musical expression 34–7
control
music weakening emotional 372
and orchestra musicians 300–1
Corigliano, composing film music 265
correlation of means, in continuous response measurement 406–7

cortical systems
and facial expressions 111–13
and musical emotions 118
creation, of mood, *see* emotion construction
creativity
and altered state of mind 283
and 'conservatoire culture' 284
cross-culturalism 37, 56, 57, 233, 323
and brain programmes 321
*see also* anthropology; universality
cross-modal congruence, in film music 260–1
crying, *see* weeping
cues, for communicating emotion in performance 314–17, 324–7
cultural construct
emotion as 138–40, 151
person as a 140–1
culturally inflected listening 135–8
cultural material, music as 415
cultural relativism 135, 138–40
culture
and aesthetics 368
and attitudes 97–8
and choice 350–1
and listener features 364
culture-nature divide 154, 161

Debussy
on mathematical abstraction of music 205
*Préludes*, in study of expression 226
*Reflections on the Water*, in continuous response study 394–5
decoding
accuracy and performance 313, 314, 319
automaticity of 322
*decoratio* 52
Delius, on importance of emotion in music 275
dependent motivation, of performer 277–8
depersonalised experience, of emotion 143
depression, music as treatment for 426–7
Descartes 24, 26, 105
development
emotion changing with 89
of perception of emotion 114–15
of performance anxiety 293–4
*see also* children
deviation, accounting for in continuous response measurement 407–8
diagesis, and film music 253–4
*see also* inference, and film music
dimension theories of emotion 77–8, 228, 244, 257, 315, 376
and continuous response measurement 401–2
future of 402
discourse, and emotion 64–8, 91
discursive analysis 427
dispositions 137
*see also* habitus of listening
dissonance appraisal 120–1
distributed neural structures 120–1
domain specificity of emotional systems 122–4
drug addicts, compared with music lovers 168
drugs, *see* pharmacotherapy, for performance anxiety
Dvorak
'New World Symphony', empathy and 371
Slavonic Dance Op. 42 No. 1, and response lag 405
Dylan, Bob, student identification with 175
dynamic form, in clinical improvisation 194–7
dynamic structure 31, 35, 50–1, 53, 58
*see also* lag structure; vitality affects

early experience, of music 425–6
ecological approach to perception 67
ecological validity, of emotion communication research 328, 328–9
ecstasy 145–6
*see also* 'renewal ecstasy'; 'transcendent ecstasy'
EEG (electroencephalogram) 119
in study of performance anxiety 296
Eisenstein, Sergej, and film music 253, 264
electroencephalogram (EEG) 119
elicitors, proximal 85–6, 90
emotion
action tendencies invoked by 87–9
automaticity of 27, 85, 110, 115, 118, 322, 366–8
behavioural component of 84–5
categorical approach to 76–7, 315, 322
changing with human development 89
characteristics of 83–91
as cultural construct 138–40, 151
definitions 75–6, 341–2, 446
depersonalised experience of 143
dimension theories, *see* dimension theories of emotion
experience of, *see* experiencing emotion
experiential component of 84–5
extrinsic 93–5
folk theories of 74
functionality of 83, 87–9
history of research into 73–4
interaction between sources of 95–6
intrinsic 30, 91–3
physiological component of 84–5
prototype approach to 78–9
proximal elicitors of 85–6, 90
recognition of 125–6, 187–91, 223–4, 267
social aspects of 86–7
sociological approach 162–4
studied in relation to music 81–2
vitality affects 79, 81, 194, 195, 196, 243
*see also* negative emotion; theories of the emotions; vocal emotion
emotional contagion 329, 369–71
*see also* response, mirroring
emotional intelligence 74, 88
'emotionality', and interpretation 319
emotional processes 344–6, 373, 384

*see also* emotional states; experiencing emotion; recognition of emotion
'emotional psychophysics', *see* lens model of acoustic code
emotional states 342–4
*see also* emotional processes; experiencing emotion; recognition of emotion
'emotional work' 171
*see also* emotion construction
emotion construction 165, 166–7, 167–73
*see also* mood induction
emotion dimensions, of communication during performance 314–17
emotion induction 313, 329–30
*see also* experiencing emotion; production rules, for emotional effects of music
*Emotion and meaning in music* (Meyer) 15, 18, 57, 59, 205, 278, 354
emotivists 105
empathy, and central route production of emotion 369–71
encoding accuracy, and performance 313
enculturation 59
*see also* cross-culturalism
enhancement, of 'inherent' emotional impact 312
Enlightenment, and human universality 139
epidemiology, of musical emotion 96
*Essay on the true art of playing keyboard instruments* (Bach) 47
ethnomusicology 135, 136, 138
evaluation
and performance anxiety 299
*see also* appraisal; preferences
evaluation dimension 228, 257, 402
of emotion, *see also* valence dimension
events, emotion as sequence of 75
everyday listening
autonomy and 421–3
content 425–6
context 425–6, 427–8
definition 415–16
functions of 419–21
and identity 421, 423–5
individuality and 421–3
as self-therapy 420–1, 426–7
studying of 416–18
evolution
and choice 348–9
and emotion 83, 115, 321
exemplificatory connections 30–1
exhibitionistic motivation, of performer 277
expectations, of listener 57, 59, 91–3, 136, 256, 263, 330, 443
*see also* melodic originality; uncertainty
experience with music, *see* early experience; expertise
experience sampling method (ESM) 417
experiencing emotion 125–6, 224, 267, 347–8
in music therapy 187–91
*see also* emotion induction; social learning
experiencing subjects, *see* composer; listener; performer; persona
experiential component, of emotion 84–5
expertise
effect of 94, 310, 313–14, 364, 398
and response lag 405
*see also* performer
*Explaining music* (Meyer) 354
expression, emotional 347–8
*see also* experiencing emotion
expression marks, of musical scores 315
expression, musical
accuracy of 313–14, 326
effects of separate musical factors 235–42
methodological approach to studying 224–35
*see also* communicating emotion in performance; linguistic models of musical expression; music therapy, improvisational; theories of musical expression
expression theory 32
expressive behaviour
as measure of emotion 74
*see also* facial expressions; mirroring response; weeping
expressive contours 317–20
*see also* dynamic structure
'expressive deviations', in performance 312
'expressiveness', perceived 317
expressive vocabulary, in music analysis 59–63
extrinsic emotion 93–5
*see also* intrinsic emotion
extrinsic motivation, and performer 278, 285

facial expressions
amygdala role in perception of 110–11
brain specialization for 107–9
cortical systems and 111–13
facilitated by music 376, 378
innateness of 106–7
response to music 84
to measure music preference 382
universality of 106–7
facilitation, of expression of pre-existing emotions 372
factors, musical
amplitude envelope 242
articulation 242, 316, 317, 319
harmony 241, 399, 443
interaction between 232, 242, 243, 400
intervals 227–8
intonation 316
loudness 240, 316, 317, 398, 399, 404, 443
manipulation of 227–35
melody 229, 240–1, 399, 443
mode 228–9, 234, 239–40, 443
musical form 234, 242
pauses 316
pitch 240
rhythm 229, 241, 371, 377, 443
in SEM 442–3

factors, musical (contd.)
tempo 229, 233–4, 235, 239, 283, 316, 317, 443
timbre 30, 241, 316, 443
timing 316, 319
tonality 241
tone attacks 316, 317
tone delays 316
tone sequences 229–30
vibrato 316
*see also* coding, of musical stimuli
familiarity, of music 92
Fauré, *Requiem*, used in music therapy 185
feeling 27, 39, 145, 446
and performance anxiety 295
sociological focus on 163
*see also* experiencing emotion; sensations
film music
cognitive framework 258–63
composer of, and emotion 264–6
definitions 249–50
and diagesis 253–4
functions of 258
historical background 250–3
and inference 254–8, 262
and non-diagesis 253–4
and semantics 56
'fit', between music and context 422
flexibility, of performer's conceptualization 283–4
flow, phenomenon of 432
fMRI (functional magnetic resonance imaging) 108–9
focus of attention, and everyday listening 418
folk theories, of emotion 74
form, musical 234, 242
formalism 49–50
frequency, of music listening 418
function
of classical music 205
of emotion 83, 87–9
of everyday listening 419–21
of expressive variations 312
of film music 258
functional explanations, in psychology 73
functional magnetic resonance imaging (fMRI) 108–9

Gadd, Steve, on choosing a musical career 276
Geminiani, Francesco, on communication of emotion in performance 311
gender differences, in adolescent music listening 424
gestalt principles
and film music 260
of musical structure 57, 92, 345
*see also* closure, and film music; expectations, of listener
GIM (guided imagery in music) 185–6
Glière, Reinhold, *Prelude* Op. 31 No. 1, study of music conceptualization 279–81
gospel music 149–50
group bonding, adaptive significance of music for 115
guided imagery in music (GIM) 185–6

habituation, and emotional arousal 214
habitus of listening 140
in Balinese bebuten trancing 148
culture and biology 152–4
Hindustani vs. Western listener 142–4
of Pentecostal listener 150
Sufi 146
*Handbook of emotions* (Lewis & Haviland) 3
happiness 145
*see also* ecstasy
harmonics, *see* timbre
harmony 241, 399, 443
Haydn
analysis by Kivy 60
Symphony No. 94 'Surprise', analysis by Simonton 208, 211
Symphony No. 94 'Surprise', prediction and 358
Symphony No. 104, in stimulus-response study 398, 399
hedonic motivation, of performer 277, 282, 285
hemispheric specialization 111–13
for music 118–20
Hermann, Bernard, composing film music 266
hermeneutics 49, 51
historical listening practice, reconstructed 62–3, 65
historiometric approach, to music 205–20
Hitchcock, and film music 264
Holst
*First Suite in E Flat*, judged musical tension in 398
*First Suite in E Flat*, and response lag 405
Horner, James, soundtrack for *Apollo 13* 262
humanist approach
and emotion 341
emotional states vs. emotional processes 342–6
evolution and choice 348–51
expression vs. experience of emotion 347–8
implication 346–8
uncertainty 351–9
hydraulic theory of the emotions 26
hyperventilation, and performance anxiety 296, 297

iconic symbol, music as 30–1, 93–4
in music therapy 185–6
*see also* dynamic structure
identification, of emotional states 342–3
identity 141–2, 169–70
and everyday listening 421, 423–5
*see also* person, as a cultural construct
ideologies, music expressing 51
Ifaluk, use of emotion words 416
imagery, and performer 279, 282
imagination
and central route production of emotion 369
and mirroring response 38–9
and music's expressiveness 33–4
*see also* guided imagery in music; imagery, and performer

immune system, music and 374
implication
  and native and syntactic processes 347
  vs. causation 346–7
implicit perception, of emotion 85–6
implicit role of music, in film 256
implicit use of expressive cues 326
improvisation, and interpretation 311
improvisational music therapy, *see* music therapy
'inattentional blindness' 253
'inattentional deafness' 254
independent motivation, of performer 277
indirect measurement 126
individual differences
  between performers 312, 323, 327
  and response lag 405
individuality, and everyday listening 421–3
induction of emotion *see* emotion induction
induction of mood *see* mood induction
inertia response 404–5
infants
  differentiation of vocal expressions 322
  musicality of and music therapy 193–4
  preference for consonance 367
  singing to 114
  *see also* children
inference, and film music 254–8, 262
innateness, of facial expressions 106–7
inspiration 283
instruction *see* music instruction
instrumental music
  and melodic originality 209
  and singing style 31
  unable to express emotion 25
instruments, musical, effect of different 242, 316
intense emotions, and SEM 435–6
intentions, of the composer 311
interaction
  between factors in SEM 445
  between musical factors 232, 242, 243, 400
  between social factors 415
  between sources of emotion 95–6
  between structure and performance 223–4, 319–20, 328
'interactional synchrony' 194
internal state, emotion perceived as 415–16, 427–8
inter-objective comparison (IOC) 55, 56
interoccular analysis, in continuous response measurement 406
interpersonality, of music therapy 194–7
interpretation 309–11
  and 'emotionality' 319
  *see also* conceptualization, of music
inter-subjective comparison 55
intervallic tension 55, 56
intervals 227–8
intimacy, and music 173–4
  *see also* emotion construction
intrinsic emotion 30, 91–3
  *see also* extrinsic emotion
IOC (inter-objective comparison) 55, 56
Ionian mode 234
'iso-principle', in music therapy 183

Jarre, Maurice, soundtrack for *Witness* 254
*King Kong*, Steiner score for 263
Korngold, Erich, composing film music 266, 267

lag structure, in continuous response measurement 404–5
language
  of film music 265
  *see also* linguistic models of musical expression; prototype approach
*language of music, The* (Cooke) 54
'languaging' 140
late-life effects, and melodic originality 217–18
learned behaviour 138
learning
  of stylistic norms 343
  *see also* music instruction; social learning
*leitmotiv*, in film music 258
lens model of acoustic code 324–8
lexicon approach, to musical expression 53–7
liking, *see* preference
limbic system 110, 366
linear regression models, of performer and listener 326–7
linguistic models of musical expression
  baroque music 46–7, 51–3
  cognitive theory 54
  lexicon approach 53–7
  semiotic theory 30, 54–7, 62
  structural analysis 57–9
  *see also* language
listener
  'community of interpretation' 136
  expectations of 57, 59, 91–3, 136, 256, 263, 330, 443
  as experiencing subject 33, 39, 64
  listening conditions established by 168–73
  and production rules for emotional effects of music 364
  qualification of 28–9, 37
  subjectivity of 141–2
  using cues in judgment of emotion 316–17
listening
  culturally inflected 135–8
  'flow' and 432
  frequency of 418
  in music therapy 184
  solitary 421–2
  *see also* context; habitus of listening
listening conditions 136
  established by listener 168–73
listening, everyday, *see* everyday listening
listening modes 137
Liszt, Faust Symphony, analysis by Simonton 208
long-term memory (LTM), and cognitive framework for film music 261, 262

loudness 240, 316, 317, 398, 399, 404, 443

'maestro' function, of performance 277, Symphony No. 10, strong emotional reaction during 440–1, 442–3
major mode, and melodic originality 210
Mancini, Henry, on composing film music 264–5
mapping vs. relating, in neurology 344
material, in music therapy 183–4
meaning, *see* musical meaning
measurement of emotion 74, 372–3
  motor expression behaviour 377–8
  physiological 374–7
  self-report 379–81
  *see also* continuous response measurement; indirect measurement
melodic originality
  aesthetic consequences of 211–15
  biographical antecedents to 215–18
  compositional impact of 211–14
  measurement of 206–11
  stylistic transformations in 214–15, 218
  *see also* expectations, of listener
melody 229, 240–1, 399, 443
memory
  and central route production of emotion 369
  and musical mood induction 373
  and performer 281–2
  *see also* associative sources of emotion; long-term memory (LTM), and cognitive framework for film music; nostalgia; short-term memory (STM), and cognitive framework for film music
meta perception 285
metaphorical expression, in music 28
Milstein, Nathan, on performance anxiety 294
*mimesis*
  and acting 145
  music as 46
mind, altered state and performer 281–3
'minor encoding habits' 206–7
minor mode
  and melodic originality 210
  and sad music 28, 210
mirroring response 37–40
mirror, music as a 423–4
  *see also* identity
mixtures of emotions, and SEM 440–1
mode 228–9, 234, 239–40, 443
  *see also* minor mode
modern music, *see* avant-garde music
modularity 108
  *see also* specialization
mood induction 419–20
  from film music 258
  gender differences 424
  measurement of 379
  music specifically for 383
  of self by performer 281–3
  using memory 373
'mood' music 383
  *see also* intimacy
moods
  definitions of 27, 75, 446
  design feature delimitation of 383
  *see also* feeling; sensations
*Moonraker*, study of soundtrack 255
mother-infant communication 323
  and music therapy 193–4
  *see also* children; group bonding
motivation, to become a performer 276–8
motor action, uncertainty and 352–3, 358
motor expressive behaviour, effect of music 377–8
'motor mimicry' 370
movement, *see* action tendencies; motor expressive behaviour
Mozart
  *Ave Verum Corpus*, late-life effects and 218
  changing musical structure of 234–5
  on compositional strategy 54
  *Die Entführing aus dem Serail* 54
  *'Dissonance' Quartet*, analysed by Kivy 60
  *'Dissonance' Quartet*, analysed by Simonton 208, 211
  on joy of performing music 291
  late-life effects 218
  Piano Sonata in A Major, K. 331, and syntactic process 347
  Piano Sonata in E Flat Major, K. 282, and tension response 399
  String Quintet in G Minor 210
  String Quintet No. 3 in C Major K. 515, emotional response to 399
  Symphony in G Minor, K. 550, analysis of 64–5
multicomponentialism, of emotions 373
multiple affect adjective checklist (MAACL(-R-)) 379
multiple emotions systems, for music 121–2
multiple regression analysis (MRA) of expressive cues 326
Münsterberg, Hugo, and film music 251–2
musemes 55, 56
*Musica Autoschediastike* (Burmeister) 52
musical factors, *see* factors, musical
musical form, *see* form, musical
musical imagery, *see* imagery, and performer
musical meaning 278, 309
'musical reality', and performer 284–6
musical score
  expression marks of 315
  as response form 403
musical structure, *see* structure
musical style, of classical composers 206–7
musical time, *see* timing
music, film, *see* film music
music instruction 425–6, 458
'musicking' 151, 183
*Music in the moment* (Levinson) 235
musicology
  classic approaches to emotion 51–9

definition 45–6
discourse and emotion 64–8, 91
expressive and structural vocabularies 59–63
history of 46–51
and sociological approach 172–3
structural and expressive vocabularies 59–63
*see also* ethnomusicology
music psychology, and study of emotion 81–2
music theory, *see* musicology
music therapy
active approach 182–3, 186–7
associative connections 184
clinical improvisation 191–7
generating emotion 187–91
iconic connections 185–6
improvisational 186–7, 191–3, 194–7
infant musicality and 193–4
interpersonality of 194–7
judging emotion 187–91
and mother-infant communication 193–4
musical material used 183–4
receptive approach 182–3, 184
*see also* depression, music as treatment for

naming, *see* classification, of emotional states
native processes 344–6
and implication 347
nature-culture divide 154, 161
*see also* culture, and choice; evolution, and choice
negative emotions 40–2
and SEM 439–40
*see also* performance anxiety
nerves, *see* performance anxiety
neural correlates
of musical emotion 126–7, 322
of vocal emotion 122–4
neurochemical correlates, of musical emotion 126–7
neuropsychology
brain specialization for music 116–18
and development of perception of emotion 114–15
distributed neural structures 120–1
domain specificity of emotional systems 122–4
evolution of musical emotion 115
experience of emotion 125–6
facial expressions and 106–13
hemispheric specialization for music 118–20
mapping vs. relating 344
recognition of emotion 125–6
single brain/body approach 151
single vs. multiple emotions systems for music 121–2
non-diagesis, and film music 253–4
non-instrumentality, of emotion 83
non-sentience, of music 26
nostalgia, and everyday listening 418, 419
'nostalgia' emotion 348
note transitions, *see* two-note transition probabilities
Nowlis mood adjective checklist 379

*On the musically beautiful* (Hanslick) 47, 49–50
open-ended response format 394
opera, representational function of 46
organismic explanations, in psychology 71–2
Organon, Bühler, production rules of 361
originality of melody, *see* melodic originality

Paganini, Nicolò, use of rubato 330
parallel enrichment procedure 314
parallel-processing 85, 121
passion, ethnographic sociology of 168
pauses 316
Pavarotti, Luciano, and performance anxiety 294
peak experience 431–2, 446–7, 447–8
Pentacostal arousal 149–50
perception
affected by belief 97, 352–3
confused with induction 380
of emotion in others 86–7
implicit 85–6
related to performance, *see* lens model of acoustic code
*see also* meta perception; recognition of emotion
perceptual grouping, in film music 260
perfectionism, and performance anxiety 299
performance
interacting with musical structure 223–4, 319–20, 328
and listener's expectations 330
and production rules for emotional effects of music 364
related to perception, *see* lens model of acoustic code
'systematic variations' in 312
*see also* communicating emotion in performance; performer
performance anxiety
cognitive factors in 298–9
coping with 301–2
'need' for 298
occurrence of 292–4
phenomenon of 294–6
and physiological arousal 296–8
stress and 299–301
treatment of 302–4
performer
conceptual flexibility of 283–4
as experiencing subject 143
and imagery 279–81, 282
individual differences 312, 323, 327
mood self-induction 281–3
motivation to become 276–8
'musical reality' and 284–6
role in communication of emotion 309–11
and social learning 323
using expressive cues 327
*see also* communicating emotion in performance; performance; performance anxiety
peripheral route production of emotion 371–2
person, as a cultural construct 140–1

persona, as experiencing subject 33, 34, 38–9
personal control, and orchestra musicians 300–1
personal experience, emotion perceived as 415–16, 427–8
personal factors, and SEM 443, 445
PET (positron emission tomography) 108–9
pharmacotherapy, for performance anxiety 302–3
philosophy
  analytic 25
  continental 25
  methods of 24–5
  personhood 141
  questions of 23–5
  theories of the emotions 26–8
  theories of musical expressiveness 29–42
*photoplay: A psychological study, The* (Münsterberg) 251
Phrygian mode 234
physical responses, in SEM 433, 441
physiological arousal 110, 296–8, 374–7
  *see also* physical responses, in SEM
physiological component, of emotion 84–5
physiological measurement, of music preference 382–3
physiological measures, of emotion 74, 125
physiological reactions, to music 84
physiological symptoms, of performance anxiety 295
pitch 240
Plato 173
plots, uncertainty and 357
  *see also* diagesis, and film music
'positioning triad' 427–8
positive emotions, and SEM 437–8
positivity-negativity scale, in everyday listening 418
positron emission tomography (PET) 108–9
potency dimension 77, 228, 257, 402
power
  structuring of 167
  uncertainty and 355–6
  prediction, *see* classification, of emotional states; expectations, of listener; inference, and film music; uncertainty
preferences
  changing over life span 89
  design feature delimitation of 382–3
  *see also* evaluation
private experience, emotion perceived as 415–16, 427–8
probabilistic functionalism, and emotion communication 320
problems
  of studying emotion and music 82
  with theories of the emotions 29–42
processes, emotional, *see* emotional processes
production rules, for emotional effects of music
  central route production 366–71
  contextual features 364–5
  evidence for 372–81
  future research 382–5
  listener features 364
  performance features 364
  peripheral route production 371–2
  structural features 362–4
  *see also* emotion induction; experiencing emotion
*Profane culture* (Willis) 166
professional musicians, *see* performer
proprioceptive feedback 371–2
prototype approach 78–9
proximal elicitors, of emotion 85–6, 90
psychological approach 71–3
*psychology of music, The* (Deutsch) 3
psychotherapy, for performance anxiety 303–4
public places, music in 422–3
Puccini, *La Boheme*, in stimulus-response study 398

qualifications, of listener 28–9, 37
  *see also* music instruction

Rachmaninoff, C Minor Piano Concerto No. 2 210
rage 146–8
rating scale response format 396–7
reaction triad 85
reactivity problem 82
reality
  and film music 254, 258, 263
  *see also* 'musical reality', and performer
receptive music therapy 182–3, 184
recognition of emotion 125–6, 223–4, 267
  in music therapy 187–91
  *see also* perception
reconstruction, of historical listening practice 62–3, 65
*Red Violin, The*, Corigliano soundtrack for 265
relating vs. mapping, in neurology 344
relationship
  between music and listener 136
  in music therapy 183
relativism, cultural 135, 138–40
relaxation training, and performance anxiety 303
'renewal ecstasy' 438
representational function, of music 46
  *see also* exemplificatory connections
represented emotion 313
response
  coding of 400–2
  continuous, *see* continuous response measurement
  habituation of 214
  mirroring 37–40
response formats, in continuous response measurement 394–7
rhythm 229, 241, 371, 377, 443
  *see also* motor action; motor expressive behaviour
rhythmic entrainment 151–2, 371–2
right hemisphere hypothesis 112, 119–20
ritual, and performance anxiety 301–2
Rodrigo, *Concierto de Aranjuez*, Schubert and Waterman studies compared 401
role playing, emotion and 88

romantic period, and melodic originality 215
Rosenman, Leonard, soundtrack for *Star Trek IV: The Voyage Home* 266
rubato, impact on listener 330
Rubinstein, Arthur, and performance anxiety 294

sad music
  attraction to 40–2
  and minor mode 28, 210
  physiological response to 375
Saint-Saens, on *leitmotiv* 258
sampling rate, for continuous response measurement 403–4
scale steps 54–5
Schoenberg 49, 165
  changing structure of 234–5
  *Pierrot Lunaire*, and melodic originality 215
Schubert
  *Impromptus* 169
  late-life effects 219
  McClary critical interpretation 51
  Second Symphony, Newcomb critical analysis 59–60
  Sonata in B Flat major, D. 960, Fisk analysis 64–5
scientific method 23–4
  in psychology 72
*Sea Hawk*, Korngold score for 267
secondary emotions 77, 316
segmental features, of musical structure 362–4
selective attention, and film music 253
self-assessment, and performance anxiety 303
self-discovery 356
self-esteem 380
self-expression 49
  *see also* composer, as experiencing subject
self-identity, *see* identity
self-report 74, 84, 379–81, 447
  *see also* continuous response measurement
self-therapy, everyday listening as 420–1, 426–7
SEM, *see* strong experiences with music (SEM)
semantics 29
  in film music 56, 258, 260
semiotic theory 30, 54–7, 62
  *see also* linguistic models of musical expression; symbol, music as a
sensations 27
  *see also* feeling; moods
'sentic' theory, of emotional expression 320
sentience, *see* non-sentience, of music
sentograph 397
sequence of events, emotion as 75
serial correlation, accounting for in continuous response measurement 408–9
short-term memory (STM), and cognitive framework for film music 261, 262
Shostakovitch
  Symphony No. 7 *Leningrad*, and melodic originality 216
  Symphony No. 10 60–1
significance, of music to listener 36, 85
sign, music as a 29–32
  *see also* linguistic models of musical expression
singing, to infants 114
singing style, and instrumental music 31
single brain/body approach 151
single emotions systems, for music 121–2
situational factors, in SEM 445
social communication, music as 421
social function
  of emotion 86–7
  of music, *see* everyday listening; group bonding; identity
social ideologies, music expressing 51
social learning, and emotion communication in performance 322–4
social motivation, of performer 277
social movements, music and 174–6
social referencing 86
social structures
  and music 164–5
  *see also* everyday listening
sociocultural context, and identity of emotion 342
sociology
  approach of 161–2
  emotional agency and music 174–6
  emotion considered within 162–4
  emotion construction 165, 166–7, 167–73
  intimacy and music 173–4
  models of music and emotion 165–7
  music and social structures 164–5
  socio-music studies 164
*sociology of rock, The* (Frith) 166
socio-music studies 164
solitary listening 421–2
'Somewhere Over the Rainbow' (Arlen), analysed by Meyer 345
*Sound effects* (Frith) 166
sources of emotion
  associative 94–5, 172, 184, 364
  extrinsic 93–5
  iconic 93–4
  interaction between 95–6
  intrinsic 30, 91–3
specialization
  of brain for music 116–18
  of hemispheres for music 118–20
specific emotions, induction of 90–1, 375, 376
stage fright *see* performance anxiety
*Star Trek IV: The Voyage Home*, Rosenman score for 266
states, emotional, *see* emotional states
statistical parameters, musical processes and 345, 346
Steiner, George
  composing film music 267
  score for *King Kong* 263
stereotypes 424–5
stimulus-response investigations, continuous response measures used for 398–9
story telling, emotion and 88

Strauss, Four Last Songs, late-life effects and 218
Stravinsky
  late-life effects 218
  *Rite of Spring* 212, 368
street musicians 423
stress
  and melodic originality 216–17
  and performance anxiety 295–6, 299–301
strong experiences with music (SEM)
  early research 431–2
  emotional aspects of SEM 435–45
  physical responses 433, 441
  SEM project 433–5
structural analysis 57–9, 92
structural congruence, in film music 258
'structural coupling' 152–3
structural vocabulary, in music analysis 59–63
structure
  changes across time 243
  dynamic 31, 35, 50–1, 53, 58
  and 'expressive shape' 312
  interacting with performance 223–4, 319–20, 328
  methodological approach to studying 224–35
  separate musical factors in 235–42
  *see also* factors, musical; intrinsic emotion; lag structure; vitality affects
structuring power
  of music 167
  *see also* emotion construction
style, *see* musical style
*Style and music* (Meyer) 347
stylistic transformations, in melodic originality 214–15, 218
subjective experience
  measurement of 379–81
  *see also* self-report
subjectivity
  of listener 141–2
  of performer 275–86
succession of states, uncertainty and 353–4
Sufis 145–6
supra-individuality 152–4
suprasegmental features, of musical structure 364
survival, and emotional reaction 82, 88
symbol
  music as a 29–32
  *see also* linguistic models of musical expression
sympathy, *see* empathy
synaesthesia 30
synchronisation
  in continuous response measurement 402–3
  and film music 261
syntactic processes 344–6
  and implication 347
synthesized performance, and emotion communication 317–19
'systematic variations', in performance 312
system-dynamics investigations, continuous response measures used for 399

Tchaikovsky
  B Minor *Pathétique* Symphony 210
  on expression of emotion 223
tears, *see* weeping response
tempo 229, 233–4, 235, 239, 283, 316, 317, 443
temporality, of music 35
  *see also* change; timing; uncertainty
tension 398–9, 401, 443
  intervallic 55, 56
texts, about music 48–9
theories of the emotions
  cognitive theory 26–7, 31, 38–9, 50
  desiderata for 27–8
  hydraulic theory 26
  problems with 29–42
  *see also* automaticity of emotion; theories of musical expression
theories of musical expression
  arousal theory 32–4, 36, 39–40
  contour theory 34–7
  expression theory 32
  music as a symbol 29–31
  *see also* expressiveness
therapy *see* music therapy
timbre 30, 241, 316, 443
timing 316, 319
  *see also* expressive contours
tonality 241
tone sequences 229–30
top-down processing, in cognitive framework for film music 261–2
tradition
  and meaning in music 56
  and music conceptualization 284
training, musical 425–6, 458
'transcendent ecstasy' 432
treatment, of performance anxiety 302–4
two-dimensional 'emotion spaces' (2DES) 397, 402
two-note transition probabilities 207–9

uncertainty
  and belief 352–3
  intensity of 358–9
  and motor action 352–3, 358
  and narratives 357
  and plots 357
  and power 355–6
  and risk-taking 356–7
  and succession of states 353–4
*Unheard melodies: Narrative film music* (Gorbman) 254
uniqueness, of musical emotion 122, 381, 384–5
universality 37, 57, 76, 345
  of arousal 144–50
  of emotion 138–40
  of facial expressions 106–7
  *see also* cross-culturalism
  use, of music, *see* emotion construction; intimacy, and music; mood induction; music therapy; self-therapy

valence dimension 77–8, 226, 230, 244, 257, 315, 316, 376
*see also* evaluation dimension; positivity-negativity scale
valence hypothesis 112, 119
validation, continuous response measures used for 397
Varèse, *Deserts* 368
variance, *see* deviation
Vaughan Williams, late life effects 218
Verdi
*Aida*, and native and syntactic processes 346
*La Traviata*, weeping reaction to 355–6
vibrato 316
vicarious functioning, of expressive cues 324–5, 330
Vienna Symphony Orchestra, study on performance anxiety 296
visual analogue scales (VAS) 379
visual primacy, and film music 261
vitality affects 79, 81, 194, 195, 196, 243
Vivaldi, *Gloria*, used in music therapy 185
vocabulary, of musical emotion 381
vocal emotion
and emotion induction 329
neural correlates 122–4
separate from musical emotion 124
voice, human, and music 321

Wagner 29, 95
*Lohengrin*, used in music therapy 185, 186
operas and physiological arousal 377
*Tristan und Isolde*, and melodic originality 215
and use of *leitmotiv* 258, 377
Wallfisch, Ernst, on choosing a musical career 276
Waxman, Franz, on composing film music 266
weeping response 399, 441
to *La Traviata* 355–6
William's syndrome 127
*Witness*, Jarre's soundtrack for 254
Wolof griots 142

Young, Victor, on film composer characteristics 264



That music has an incredible power to move us emotionally is without question. Whether performing music, listening to music, or creating music, this bond with our emotions is always there. The nature of this relationship is clearly complex however, and emotional aspects of music have received surprisingly little attention in the 45 years since the publication of Leonard Meyer's classic work 'Emotion and meaning in music.' During that time, both 'music psychology' and 'emotion' have developed as vibrant areas of research, and the time is fitting therefore to try and bring together this multidisciplinary interest and take stock of what we now know about this powerful relationship.

A new volume in the *Series in Affective Science*, *Music and emotion: theory and research* brings together leading researchers interested in both these topics to present the first integrative review of this subject. The first section reflects the various interdisciplinary perspectives, taking on board views from philosophy, psychology, musicology, biology, anthropology, and sociology. The second section addresses the role of our emotions in the composition of music, the ways that emotions can be communicated via musical structures, the use of music to express emotions within the cinema. The third section looks at the emotions of the performer - how do they communicate emotion, how does their emotional state affect their own performance. The final section looks at the ways in which our emotions are guided and influenced while listening to music, whether actively or passively.

A state of the art account of the subject, *Music and emotion* is a book long overdue, one that will fascinate psychologists, musicologists, music educators, and philosophers.

OXFORD
UNIVERSITY PRESS

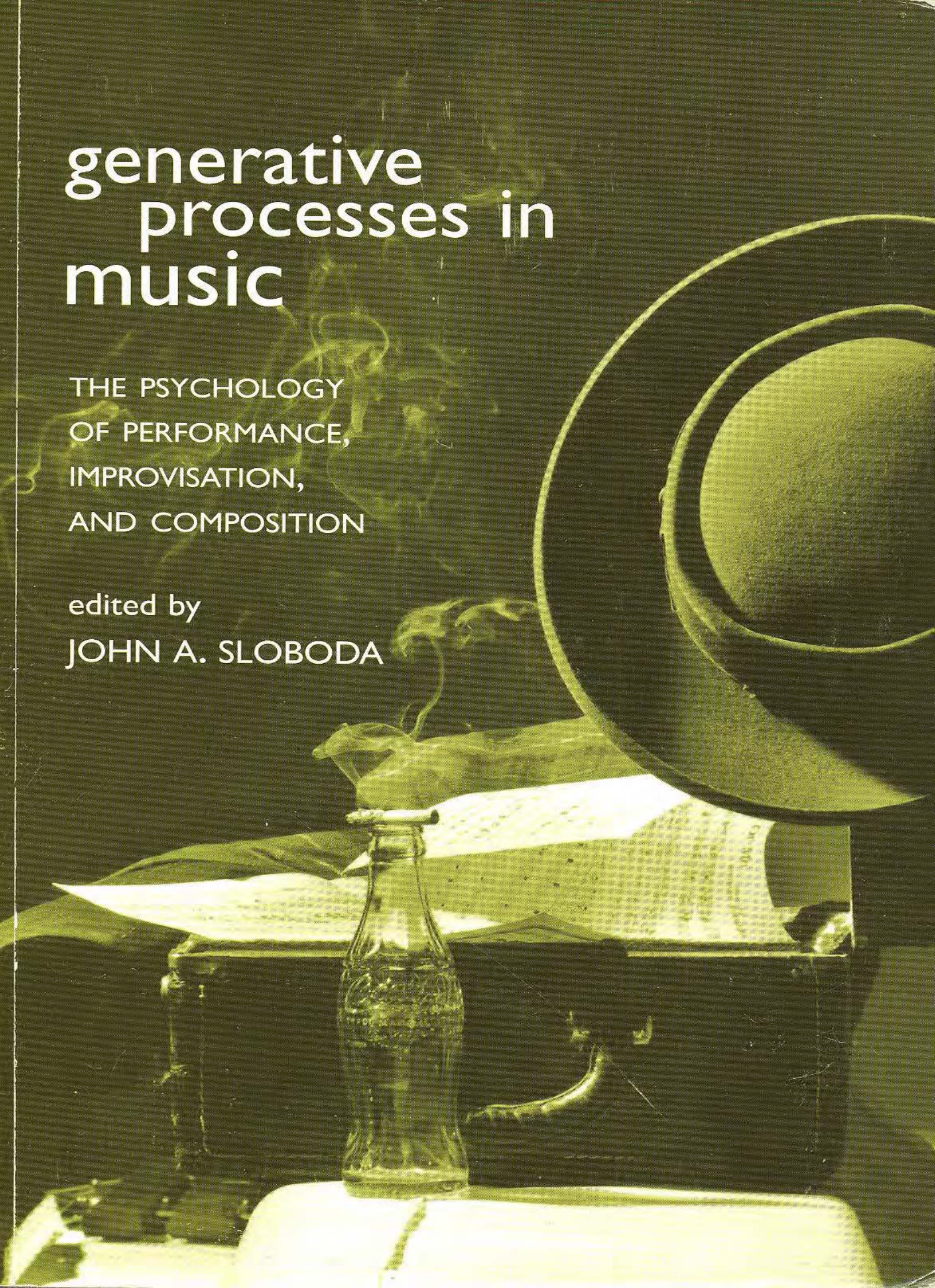
**www.oup.com**

ISBN 0-19-263188-8
9 780192 631886

# generative processes in music

## THE PSYCHOLOGY OF PERFORMANCE, IMPROVISATION, AND COMPOSITION

edited by
JOHN A. SLOBODA



# Generative Processes in Music

The Psychology of Performance, Improvisation, and Composition

---

Edited by

John A. Sloboda

*Department of Psychology*
*University of Keele*

CLARENDON PRESS · OXFORD

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP
Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in
Oxford New York
Athens Auckland Bangkok Bogotá Buenos Aires Calcutta
Cape Town Chennai Dar es Salaam Delhi Florence Hong Kong Istanbul
Karachi Kuala Lumpur Madrid Melbourne Mexico City Mumbai
Nairobi Paris São Paulo Shanghai Singapore Taipei Tokyo Toronto Warsaw
with associated companies in Berlin Ibadan
Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries
Published in the United States
by Oxford University Press Inc., New York
© John A. Sloboda and the contributors
listed on p. xvii, 1988
First published new as paperback 2000
All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above
You must not circulate this book in any other binding or cover
and you must impose this same condition on any acquirer
A catalogue record for this book is available from the British Library
British Library Cataloguing in Publication Data
Sloboda, John A.
Generative processes in music: the psychology
of performance, improvisation and composition.
1. Music—Psychology
I. Title
781'.15 ML3830
ISBN 0 19 850846 8
Library of Congress Cataloging in Publication Data
Generative processes in music.
Bibliography: p.
Includes indexes.
1. Music—Performance—Psychology.
2. Improvisation (Music)—Psychology.
3. Composition (Music)—Psychology.
4. Cognition. 5. Developmental psychology.
I. Sloboda, John A.
ML3838.G38 1988 781.6'01'9 87-12369
ISBN 0 19 850846 8
Printed in Great Britain by
St Edmundsbury Press Limited,
Bury St Edmunds, Suffolk

# Preface

JOHN SLOBODA

This book was conceived to remedy an imbalance in the Psychology of Music literature. The vast majority of serious psychological research has been devoted to an analysis and understanding of the processes involved in listening to music. This is reflected in the existence of a major journal with the title *Music Perception*. There could not be a journal entitled *Music Performance*, because there just would not be enough research to fill it.

The reasons for the relative neglect of generative processes (including performance, improvisation, and composition) are various. I wish to comment briefly on three of these: cultural bias, measurement problems, and problems of control. I shall then turn to some general comments about the contributions to this volume.

Through a long social and historical process contemporary Western art culture has become characterized by functional specialism. In the case of music, as with most other art forms, a gulf has emerged between producer and consumer. Adult producers are typically few in number, often institutionally trained. Adult consumers are typically large in number, usually untrained, and often unskilled in most forms of music production. With the advent of sound recording and other forms of mechanical transmission it has become possible for the various functions to be completely separated in both time and space. Consider, for instance, a Beethoven symphony. A group of performers construct their interpretation in necessary isolation from the composer, and also from the audience for whom the performance is intended (for example in a recording studio). A member of the audience may then 'receive' the interpretation in social isolation from both composer and performer as a disembodied aural experience. This is particularly true of the domestic listener who may, through use of headphones and darkened room, create the impression of being totally and exclusively enfolded by sound. Even at the concert there are usually strong social and geographical factors which separate listeners from performers. Performers and listeners go in and out by separate doors; they do not interact with one another. Any form of audience interruption is usually violently resisted. Such constraints tend to reinforce an 'illusion' which projects the sound of the music away from the realities of its origins

in human work, both physical and mental. As in the puppet theatre, the modes of production become veiled in mystery, and we may have no particular wish to venture behind the proscenium arch.

In such a culture it is, perhaps, inevitable that the listener's response to music should become the primary focus of the psychology of music. This focus would, however, be less than fully comprehensible in cultures (maybe the majority of the world's musical cultures so far) where functional specialism had not gone so far. In many cultures it would be strange to hear music without joining in in one way or another. The separation between performer and composer is also necessarily less clear. Especially when the music is without notation, the originator of the music must necessarily perform it to communicate it. Also, composition/improvisation is often a social process, with several individuals in a group contributing substantially to the form of the music.

One of the recurrent themes of the contributions to this book is inextricable connection of generative and receptive processes. Their separation can only ever be partial. The reality is that all music must reflect the psychological propensities and capacities of humans as composers, performers, and listeners. To an extent, each type of involvement with music includes the others. Composers temper their compositional processes against their own aural judgements (see Lerdahl, Chapter 10) and their understanding of the performance characteristics of various instruments. Performers refine their performances against both their own aural awareness and their ability to image alternatives (see Pressing, Chapter 7). Listeners grasp a work of music by attempting to sing or hum parts of it, or by engaging in some form of rhythmic movement. They may also 'compose' variants or elaborations of the music in informal (but not necessarily overt) behaviour. Sagi and Vitanyi (Chapter 8) go so far as to propose that the ability to receive music is constrained by the extent to which people are capable of generating variants within a particular genre. In a similar vein, Dowling (Chapter 6) shows that young children's ability to discriminate tonal from atonal melodies is dependent on the level of singing ability that they have achieved.

We may suppose, in fact, that the impulse to generate and perform music is as inherent a part of the human constitution as is the impulse to generate language. The research reviewed by Dowling (Chapter 6) shows that very young children produce spontaneous songs which incorporate but by no means mimic elements of adult productions. Just as in language development, song development seems to go through a set of ordered, rule-governed stages. It may be that the failure of early generativity to flower into full-scale compositional ability is a result of both the ubiquitous concern of the 8–11 year old with 'correctness' and also traditional educational styles which undervalue experimentation and self-discovery. Where opportunities are given for self-directed musical productivity (for

example Swanwick and Tillman 1986), children demonstrate an impressively varied yet ordered progression of compositional strategies right through the school ages. Davidson and Scripp (Chapter 9) support this picture by an intriguing study of children's spontaneous notations for music.

A second reason for the neglect of generative processes relates to problems of measurement in psychological science. It is always easier to collect one's data in the form of responses from a limited and pre-ordained set (for example yes–no decision, same–different decisions) than it is from relatively unconstrained and multidimensional behaviour. This concern for precision and simplicity in measurement has led to a concentration within the psychological community on rather simple laboratory tasks. Claxton (1980, p. 13) paints a humorous but not altogether inapposite picture of the typical study in cognitive psychology:

> cognitive psychology does not, after all, deal with whole people, but with a very special and bizarre—almost Frankenstinian—preparation, with consists of a brain attached to two eyes, two ears, and two index fingers. The preparation is only to be found inside small, gloomy cubicles, outside which red lights burn to warn ordinary people away. It stares fixedly at a small screen, and its fingers rest lightly and expectantly on two small squares of black plastic—microswitches. It does not feel hungry or tired or inquisitive; it does not think extraneous thoughts or try to understand what is going on. It simply processes information.

In terms of music psychology, this 'Frankenstinian' orientation has led to many studies in which subjects hear short computer-generated fragments through headphones and then are required to make some (usually binary) judgement about them, whose accuracy and reaction time may be measured relatively simply. The advantages and limitations of this approach are well known and discussed (for example Sloboda 1985; Longuet-Higgins 1987) and need not be rehearsed again here. Suffice to say that every single empirical contribution to this volume has, at one stage or another, adopted a different approach to measurement. In general, subjects have been allowed to produce responses of some complexity and length, free from excessive experimental constraints. This approach is essential is one is to make any meaningful contribution to the psychology of performance or composition.

The problems of recording and analysing these responses have been approached in a number of ways. Some researchers rely on auditory transcription into conventional (or amended) notation, and then perform formal or informal anayses on the transcripts (for example Gruson, Chapter 5; Dowling, Chapter 6; Sagi and Vitanyi, Chapter 8; Davidson and Welsh, Chapter 10). The advantages and disadvantages of the transcription method have been discussed at length by Sloboda and Parker (1985). Whilst transcription produces a record which is easily understandable by literate musicians, it runs the risk of making the performance seem more coherent than it actually was (by the transcriber's 'intelligent'

cleaning up of inherently messy data). It is also necessarily selective, and cannot reliably record information about expressive parameters of the performance. There also remains the problem of how to turn such transcriptions into quantitative form, for the application of statistical techniques. The contributors to this volume solve this problem in a number of ways, ranging from highly elaborate coding schemes (Gruson, Chapter 5) to informal stylistic analyses (Sagi and Vitanyi, Chapter 8).

Other researchers have used mechanical and electronic means of recording performance parameters in a precise quantitative fashion (for example Clarke, Chapter 1; Gabrielsson, Chapter 2; Sundberg, Chapter 3; Rasch, Chapter 4). This allows for more sophisticated statistical analysis, the testing of precise models with quantiative predictions, and the capturing of highly important expressive dimensions of performance. The disadvantages of this method arise largely from the data overload that is generated, and from the fact that it is much more difficult (although not impossible) to apply informed musical intelligence to the interpretation of data. The ubiquity of the computer has made the first problem more manageable and has, one suspects, been largely responsible for the modest upturn in the quantity of this work in the last ten years.

A third set of researchers have used tasks which require subjects to write down the end-products of their cognition in notation, traditional (Davidson and Welsh, Chapter 10) or otherwise (Davidson and Scripp, Chapter 8). This is a particularly useful method of studying composition, especially when coupled with verbal protocols. The disadvantage of the method is that schemes of quantitative measurement are hard to devise, and are usually descriptive in nature. They are, however, particularly rich sources of ideas and insights.

The third reason for the neglect of generative processes is linked to the second, and relates to the difficulties of devising suitable experimental controls over generative behaviour. Without some form of control over the conditions under which subjects exercise their generativity it becomes very difficult (but not entirely impossible) to extract any generalities about the processes involved; and without generality there is no science. The range of controls deployed in this volume varies from the very broad (in the case of improvisation and composition) to the very tight (in the case of some performance studies, where note, speed, and fingerings are specified). For instance, Sagi and Vitanyi (Chapter 8) offer the minimum constraint of a text. Subjects of varying abilities and backgrounds are required to improvise any tune they like to that text. Davidson and Welsh (Chapter 10) adopt a similar but rather more constrained tactic. They offer their subjects a rhythm, and ask them to assign pitches to the rhythms so as to produce a melody which must modulate in a certain way. Again, subjects of varying levels of ability are tested. Gruson (Chapter 5), in her study of piano

practising, simply specifies the piece to be practised, and leaves everything else up to the subjects.

Clarke (Chapter 1) is at the other extreme in his studies of expressive performance. He studies piano performance so as to control (by exclusion) the possibility of subjects using pitch and timbral variations, and then devises melodies with special qualities to test very specific hypotheses (such as that sliding the bar line along a single melodic fragment will affect expressive performance of it in systematic ways). Subjects are required to produce accurate performances of these melodies. Control is also obtained in another way by several authors. Instead of gathering data by looking at many different subjects, they obtain relatively large amounts of data from one or few subjects, looking for generalities *within* the behaviour of individual subjects (for example Clarke, Chapter 1; Gabrielsson, Chapter 2; Rasch, Chapter 4). It may be more important to show that the behaviour of an individual is systematic and rule governed, than to show that many individuals are systematic in precisely the same way. Indeed, in the study of artistic behaviour one may argue that one would neither expect, nor wish for, high degrees of uniformity between subjects.

The final part of this preface is concerned to make some general points about the contributions to the volume, and to highlight some specific aspects of individual contributions.

The first thing to say is that most of these contributions report pioneering work. There is no long tradition of research into most aspects of music generation. Thus, to my knowledge, Gruson (Chapter 5) is the first person ever to have subjected rehearsal behaviour to systematic quantitative analysis. Davidson and Welsh (Chapter 10) are the first to apply systematic analysis to a compositional task undertaken by groups chosen to represent differing ability levels. Sagi and Vitanyi (Chapter 8) have created the first large-scale corpus of improvisations by untrained people. With the pioneering comes an element of risk. It is not clear which of these approaches and ideas have a future. Some of the work is necessarily exploratory, and without the substantial backing of related research. It is not clear that all the theorizing contained here will actually be confirmed by data. Lerdahl makes a provocative and highly original claim about what constitutes a viable relationship between methods by which composers compose and how listeners listen. Will composers be able to use Lerdahl's constraints to guide them towards composing structurally rich yet comprehensible pieces? We do not yet know. Can Pressing's impressive theory of improvisation be tested against actual improvisations? The ideas are rich and plausible, but do they have predictive value? Such issues are raised but not resolved here.

The second point about the contributions is that they are truly interdisciplinary in nature. The contributors are roughly equally divided

into those working in psychology departments or other scientific setings, and those working in music or socio-cultural studies. The individaul contributions draw on ideas from linguistics, sociology, neurophysiology, acoustics, artificial intelligence and computer science, as well as psychology and music. This diversity of approach is welcome, particularly when similar ideas recur in contributions from quite different sources. It does mean, however, that several languages and perspectives rub shoulders with one another, and this sometimes makes for tensions. Psychologists may well find the contributions of the musicians unsatisfactory because of their apparent disconnection from the task of data gathering. Experimentalists may well find the rather informal analyses in some of the papers disappointing. On the other hand, musicians may well find the interpretative caution of some psychologists irritating and pedantic. Others may find the cultural narrowness of the types of music discussed less than satisfactory. Neither I nor the contributors should have to apologize for this. As a reflection of the 'state of the art' this volume represents the growth pains of a young discipline, full of ideas and enthusiasm, but not wholly co-ordinated in approach or methodological priorities.

Chapters 1 to 3 form a relatively coherent group of contributions, all focusing in slightly different ways on the question of what makes a musical performance. It is abundantly clear that a mechanical performance reproducing exactly what is contained in the notation rarely makes a satisfactory performance. Clarke (Chapter 1) emphasizes the role of performance in exemplifying structural aspects of the music through expressive gradients, discontinuities, and contrasts. He embeds several ingenious experimental studies within a theoretical framework that emphasizes the importance of hierarchical mental representations as a means of controlling performances. Gabrielsson (Chapter 2) supplements Clarke's study with some naturalistic data on rhythmic performance. This data confirms that performers make systematic and significant deviations from strict metricality, but also that it is hard to make generalizations about the nature of the deviations. The same figure may be played differently in different contexts, or by the same performer on different occasions. Sundberg (Chapter 3) describes some ingenious studies where a computer 'learns' to play expressively by being supplied with a set of rules for modulating timing and loudness in certain musical contexts. These rules are modified by playing the results to a panel of human judges and asking for judgements of appropriateness. Sundberg amplifies the interesting point that 'apt' expressive devices are very often not explicitly identified as such by judges, but simply contribute to a sense of 'rightness' about a performance.

Rasch (Chapter 4) summarizes his painstaking and elegant work on synchronization in ensemble performance. He is one of a tiny number of researchers who have examined any aspect of ensemble playing. He

reports many intriguing results, including the finding that musicians achieve better synchrony at faster speeds than at slower speeds. It also seems that a degree of asynchrony increases the perceptual clarity of individual lines, even though listeners are not consciously aware of quite large degrees of asynchronization in musical performance. Finally, Rasch provides an elegant mathematical justification for the fact that ensemble of more than 10 players tend to need a conductor while smaller groups do not.

Gruson (Chapter 5) reports a pioneering study on rehearsal strategies of musicians at different levels. She finds that people at any given level do not change their strategy much over the course of learning a given piece, but that significant strategy differences do exist between different levels of expertise. For instance, experienced players are more likely to repeat musically coherent sections of the piece. Beginners are much more likely to repeat individual notes. The general drift of her evidence is that experienced rehearsers are more aware of the structure of what they are rehearsing, and that this makes rehearsal more efficient.

Dowling (Chapter 6) explores the strong links between perception and production in the development of musical competence in young children. He reviews the literature on developmental aspects of song production and reports two new empirical studies. The first shows that children as young as three years old can reliably discriminate tonal from atonal melodies so long as their song performance is relatively advanced. The second study explores the structure of nursery songs as opposed to folk-songs and art songs. The analysis shows that nursery rhymes very rarely depart from strict diatonicity, and Dowling argues that such songs are optimal for children extracting and internalizing rules of tonality. This chapter is, in a sense, the pivot of the book, since it reminds us that children's song has elements of reproduction and improvisation inextricably intertwined. The preceding chapters deal exclusively with reproductive performance. Those to follow deal exclusively with the creation of new music through improvisation or composition.

Pressing (Chapter 7) provides a monumental survey of concepts from a number of disciplines that are relevant to an understanding of improvisation, and outlines a theory of improvisation that concentrates on the psychological bases for moment-to-moment choices within an improvisational structure. The theory is associationistic in nature, and proposes that improvisation proceeds by means of choices of elements which either *continue* or *interrupt* some aspect of the immediately preceding context. In its concern with the note-by-note structure of an improvisation it contrasts with the contribution of Sagi and Vitanyi (Chapter 8). Here, the emphasis is on global features of structure and style in the spontaneous improvisations of ordinary Hungarians. A principal finding of the study was that the majority of untrained subjects were able to use *some* coherent strategy to provide a

melody, often based on forms available in their culture, but to some extent affected by the nature of the words they were required to set.

Davidson and Scripp (Chapter 9) present intriguing data from children aged five to seven who were asked to notate various musical fragments without explicit instruction. What is particularly impressive about these data is the coherence of the schemes devised by children. Even at age five, many children devise abstract symbols to represent particularly the rhythmic structure of the music. By age seven, pitch becomes an increasing focus for notational attention. Interestingly, at age five performance skill and notational skill seem largely uncorrelated; but by age seven the two seem to have become 'yoked' together into a unified domain of musical intelligence.

The final two chapters in the book turn to the skill of composition. Lerdahl (Chapter 10) draws some logical yet provocative consequences from his influential work with Jackendoff (Lerdahl and Jackendoff 1983). If a listener understands music by constructing a hierarchical mental representation of it, then effective compositions will abide by constraints which allow listeners to do just that. Lerdahl identifies a number of specific constraints, and shows that serial music falls foul of several of them. This, he argues, is the reason why such music is so impenetrable.

Davidson and Welsh (Chapter 11) adopt a data-driven approach, in contrast to Lerdahl's theory-driven one. They look at the processes by which music students of differing expertise construct a melody given certain conditions that the melody must fulfil. Their analysis yields some tempting pointers as to the nature of compositional expertise. The more expert group worked with larger units (as shown by the size of section they would play over). They also engaged in more reflection, while the less experienced group worked in an enactive fashion (for example trying out possibilities at random until hitting on one that 'worked'). The study also highlights the dissociation of skills at this level. Even very gifted performers can be 'poor' composers without specific practice.

It is customary for the editors of multi-author volumes to draw some general overarching conclusion from the set of studies he or she has gathered together. I am reluctant to do this. My reluctance stems from the fact that the unity of a set of papers is often to be found, not in their outcomes, but in the shared assumptions that motivated them. In the case of this volume, many of the contributors attended an international conference on Psychology and the Arts held in Cardiff in 1983 (Crozier and Chapman 1984). It was primarily as a result of conversations there, and a sense of shared purpose, that the idea for this book was born. To find, therefore, that the contributions point in the same general direction is as little of a surprise as it would be to discover that the occupants of a given aeroplane are all headed for the same destination.

In the case of this volume, I would identify five core assumptions which run explicitly or implicitly through the majority of contributions:

(1) Music generative capacity is inherent in all human beings, although it may be developed to a greater or lesser extent.
(2) The capacity to generate any but the most primitive musical sequences is based on the ability to derive sound sequences from higher-order structures or rule systems.
(3) These rule systems have some universal constraints on them (arising from general facts about human cognitive capacity) but incorporate specific constraints picked up from the prevailing musical culture.
(4) Specific instruction is not necessary for skill acquisition, but practice is. Through practice, and possibly through general developmental changes, similar stages in skill acquisition can be observed in the several generative domains.
(5) Many aspects of skill become partly automated, and not open to conscious introspection. Their nature must therefore be elucidated by observation and analysis of generative behaviour rather than (or in addition to) verbal self-report.

The papers in this volume all attempt to put flesh on the bare bones of one or more of these assumptions through theorizing and/or empirical observation. Their success may be judged by the degree to which the reader is more inclined to assent to these assumptions after reading the book than he or she was before. If the reader wishes for elucidation on different but equally important issues, such as the role of motivation and emotion in sustaining and shaping generative behaviour, or the nature of effective training in these skills, then he or she must search elsewhere. If the search is successful then I would certainly like to know about it.

## References

Crozier, W. R. and Chapman, A. J. (eds) (1984). *Cognitive processes in the perception of art.* Elsevier, North-Holland.

Claxton, G. (1980). Cognitive Psychology: a suitable case for what sort of treatment. In *Cognitive psychology: new directions* (ed. G. Claxton). Routledge Kegan Paul, London.

Lerdahl, F. and Jackendoff, R. (1983). *A generative theory of tonal music.* MIT Press, Cambridge, Mass.

Longuet-Higgins, H. C. (1987). Music and the psychologists; or the poet and the peasant. *Psychology of Music* **15** (1), 3–7.

Sloboda, J. A. (1985). *The musical mind: the cognitive psychology of music.* Oxford University Press, Oxford.

Sloboda, J. A. and Parker, D. H. H. (1985). Immediate recall for melodies. In *Musical structure and cognition* (eds. P. Howell, I. Cross, and R. West), Academic Press, London.

Swanwick, K. and Tillman, J. (1986). The sequence of musical development. *British Journal of Music Education* **3**, 305–39.

# Contents


List of contributors xvii

1. Generative principles in music performance 1
   *Eric F. Clarke*
2. Timing in music performance and its relations to music experience *Alf Gabrielsson* 27
3. Computer synthesis of music performance 52
   *Johan Sundberg*
4. Timing and synchronization in ensemble performance 70
   *Rudolf A. Rasch*
5. Rehearsal skill and musical competence: does practice make perfect? *Linda M. Gruson* 91
6. Tonal structure and children's early learning of music 113
   *W. Jay Dowling*
7. Improvisation: methods and models 129
   *Jeff Pressing*
8. Experimental research into musical generative ability 179
   *Maria Sagi and Ivan Vitanyi*
9. Young children's musical representations: windows on music cognition 195
   *Lyle Davidson and Lawrence Scripp*
10. Cognitive constraints on compositional systems 231
    *Fred Lerdahl*
11. From collections to structure: the developmental path of tonal thinking 260
    *Lyle Davidson and Patricia Welsh*

Appendix: Examples 8.11–8.18 286

Author index 291

Subject index 295

# Contributors

**Eric F. Clarke**, Music Department, The City University, Northampton Square, London EC1V 0HB, UK.

**Lyle Davidson**, Project Zero, Harvard University, Longfellow Hall, Appian Way, Cambridge, MA 02138, USA.

**W. Jay Dowling**, Program in Human Development and Communication Sciences, The University of Texas at Dallas, Richardson, TX 75083–0688, USA.

**Alf Gabrielsson**, Department of Psychology, Uppsala University, Box 227, S–75104 Uppsala, Sweden.

**Linda M. Gruson**, 28 Vesta Drive, Toronto, Ontario M5P 2Z5, Canada.

**Fred Lerdahl**, School of Music, The University of Michigan, Ann Arbor, MI 48109–2085, USA.

**Jeff Pressing**, Department of Music, La Trobe University, Bundoora, Victoria 3083, Australia.

**Rudolf A. Rasch**, Institute of Musicology and Institute of Phonetics, University of Utrecht, Drift 21, 3512 BR Utrecht, The Netherlands.

**Maria Sagi**, Research Institute for Culture, H–1251 Budapest, Corvin Ter 8, Hungary.

**Lawrence Scripp**, Project Zero, Harvard University, Longfellow Hall, Appian Way, Cambridge, MA 02138, USA.

**John A. Sloboda** (editor), Department of Psychology, University of Keele, Keele, Staffordshire ST5 5BG, UK.

**Johan Sundberg**, Department of Speech Communication and Music Acoustics, Royal Institute of Technology, S–10044 Stockholm 70, Sweden.

**Ivan Vitanyi**, Research Institute for Culture, H–1251 Budapest, Corvin Ter 8, Hungary.

**Patricia Welsh**, Project Zero, Harvard University, Longfellow Hall, Appian Way, Cambridge, MA 02138, USA.

# 1

# Generative principles in music performance

ERIC F. CLARKE

## Introduction

The aim of this chapter is to give an account of generative principles involved in music performance at two levels. One level concerns the representation of musical structure in a form that gives a coherent and intelligent input into a motor system. The word generative is understood here in the same descriptive and analytic sense as in Chomsky (1957), and in the more recent music theory of Lerdahl and Jackendoff (1983). The second level at which generative principles are identifiable is in the production and control of the expressive aspects of performance, which function so as to convey a particular interpretation of a musical structure. This sense of the word generative is much closer to the idea that something is actually generated, rather than simply being describable in terms of generative principles. These two usages of the term correspond to rather different psychological orientations, and while there are certain circumstances in which the two become closely associated, one purpose of this chapter is to illustrate the distinction between them.

## The generative structure of musical knowledge

Playing music is an activity that is comparable in cognitive complexity to speaking a language, and comparable in its demands on motor control to playing a sport like tennis. It has thus been a focus of attention for psychologists interested in issues of motor control (for example Gates and Bradshaw 1974; Shaffer 1981), and for those interested primarily in musical cognition (for example Bengtsson and Gabrielsson 1983; Sloboda 1983). The motor programming perspective adopted in most recent work on motor control has, however, diminished the separation between these two lines of enquiry by emphasizing the importance of cognitive structures in the control of movement. Continuous reference to a large body of musical knowledge is required in music performance if the result is to be

fluent and intelligent, making it difficult to maintain a definite distinction between the cognitive structures of abstract musical understanding and those embodied in a motor programme for musical performance. This discussion of the knowledge structures that form the basis for musical performances therefore starts at a comparatively high level of abstraction by considering first some aspects of music theory.

Although the relationship between music theory and psychology is rather loosely defined, most music theory not only incorporates psychological principles of one sort or another (for example gestalt principles in Meyer 1973), but also takes the explanation of musical experience as one of its primary aims. It is, in other words, a theory of how people hear music, as well as a theory of how music is formally structured.* The structures described by music theory can therefore be taken as reasonable indicators of the nature of cognitive structures for music, particularly since recent work (for example Shepard 1982) has tended to confirm the perceptual significance of these structures.

The most widespread characteristic of musical structure embodied in music theory is its hierarchical nature (see Narmour 1983 for a recent review). In the parameters of both pitch and rhythm, structures are represented almost without exception as being organized in a series of levels, between which relationships of reduction or elaboration operate. Although hierarchical structures should not simply be equated with generative structures, they are nevertheless closely related, and as a number of authors have shown (for example Longuet-Higgins 1976; Sundberg and Lindblom 1976; Lerdahl and Jackendoff 1983) truly generative theories can be shown to have considerable explanatory power.

All of these studies make use of tree diagrams as representations for a generative structure, although it is not a requirement that a generative structure be represented in this way. Since a tree diagram is a useful shorthand that conveys generative relationships, I will make use of it for illustrative purposes in the following discussion. In considering the generative structures of musical knowledge, my aim is to examine their more global characteristics, rather than the precise nature of the generative principles themselves. It is the topology of generative relationships that is of interest, or put another way, the pattern of hierarchical structures that constitutes musical knowledge.

The issue can be tackled only in relation to specific performance conditions. It is inevitable that the knowledge structures that underlie a performance of a piece from memory will be different, at the highest levels at least, from those associated with a free improvisation, or a sight-read performance from notation. Starting at one extreme, performances of

* Note that this is not the case with all music theory. For example many of the structures discussed in Forte (1973) are of a level and type of abstraction that makes it extremely unlikely that they could be perceived in any direct fashion.

classical music from memory appear to offer the most deeply embedded generative structures. Though something of an idealization, we can imagine a performer who, at the start of a performance, has a complete knowledge of the generative structure of the piece, from the very highest level, where the whole piece is represented as a unity, down to the lowest level, where each individual note is represented. Figure 1.1 is a schematic representation of such a knowledge structure. Evidence for the highest level in this structure is rather sparse, and is confined to statements by a number of composers (Mozart, Beethoven, Hindemith) which indicate that they were able to hear (or imagine) their own compositions in a single 'glance'. Since these composers were also performers, the unified conceptions of which they claimed to be capable can be regarded as the basis of performances as well as compositions.

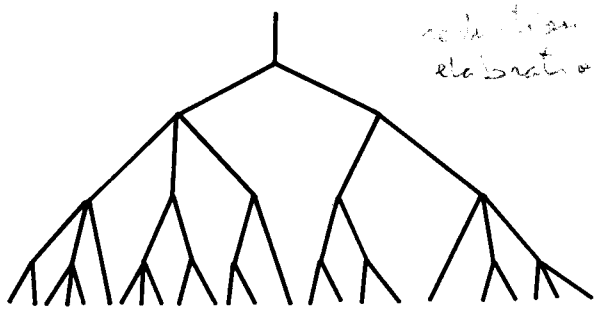


**Fig. 1.1.** Schematic representation of an idealized knowledge structure for a memorized musical performance.

The evidence for totally unified structural knowledge is a little tenuous, but does not affect the essential point that high-level musical structures acquire unified cognitive representations. It is easier to accept that a performer may have a unified conception of each of three large sections, for instance, that together make up a movement. These units connect with all the individual events of the piece through the multi-levelled branching structure that a tree diagram illustrates. It is difficult to know quite how one might experimentally test the existence of these very high levels of generative structure, but tonal structures of a fairly abstract nature (and hence at a fairly high hierarchical level) such as the concept of key, or tonal area, have been demonstrated to have psychological reality as well as theoretical value (see Krumhansl 1983).

The idea that a performer who has memorized a piece of music has a complete generative representation is an idealization not only because the structural depth of the representation is uncertain, but also because it seems implausible that a complete structural representation of the

complexity required for a piece of even moderate length could be activated. It is more likely that only a part of the entire structure is active at any time, the remainder being in a latent state, or active only in broad outline. At any particular moment the generative structure is incomplete or only partially activated, the active region shifting as the performer progresses through the music, revealing different areas and levels of the structure. For instance, in the middle of a deeply embedded musical phrase, only a region of low-level generative connections might be active, since there is little need for a performer to have access to high-level structural information; his or her main concern is the detailed structure of connections within the phrase itself. At a phrase boundary, however, it may be important for the performer to know how the previous and subsequent phrases are related to one another and to the overall structure of the piece. At these moments a small area of low-level structural connections may be active, sufficient to specify the immediate succession of events to be played, together with a section of the higher levels of generative structure specifying larger-scale relationships. Figures 1.2(a) and 1.2(b) illustrate schematically these two conditions. This suggests that in the course of a performance, a player's structural awareness constantly

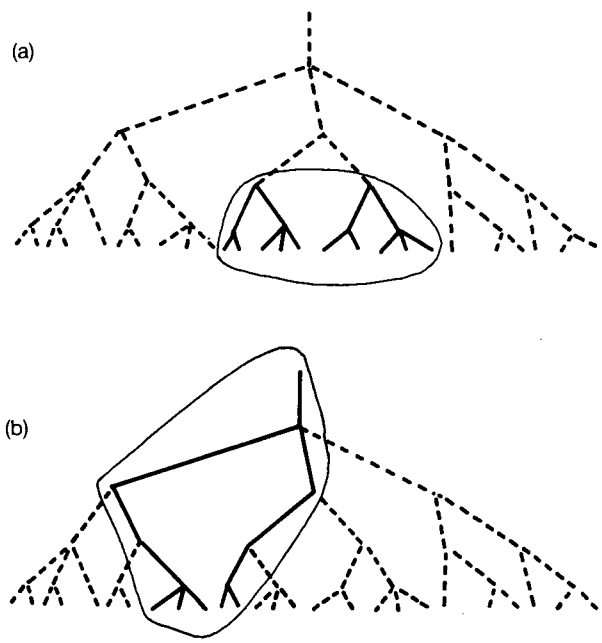



**Fig. 1.2.** Schematic representations for partially activated knowledge structures in a memorized performance: (a) illustrates active parts of the generative structure (ringed) in mid-phrase; (b) illustrates active parts approaching a phrase boundary.

shifts between regions of activated structure that vary in durational extent and generative depth. As a general rule, the depth to which the generative structure is activated is directly related to the structural significance of phrase boundaries lying close to, or at, the player's current musical location.

The second type of performance whose representational structure I wish to consider is music played from a notated source. Unless the performer already knows the music extremely well, and is simply looking at the notation for reassurance, a variable amount of uncertainty surrounds his or her anticipation of how the music may develop, and the relationship between current and future events. The future course of musical events is not, of course, completely unknown, since listeners and performers continually make projections on the basis of acquired stylistic knowledge (see Meyer 1956; Narmour 1977); but it seldom turns out that the music exactly matches these projections for a number of reasons: (1) performers and listeners may envisage a number of different possible continuations; (2) stylistic knowledge, on which projection is based, is never perfect and complete; (3) it is in the nature of music to depart from stylistic norms in the interest of creativity. Figure 1.3, which represents the knowledge structure of a piece performed from notation midway through the music, indicates these uncertainties with dashed connections and question marks.

The significance of the mismatch between projection and outcome lies in its effect upon the continuous construction of a generative representation during a performance. With memorized performance, the generative structure is known entirely (or at least to a high level) in advance, and is simply unpacked during a performance. Thus misrepresentation of the generative structure on the part of the performer is essentially restricted to memory lapses at various levels. By contrast, in an unmemorized performance it is likely that incorrect projections will result in inappropriate attributions of structual significance, failure to grasp long-term connections,

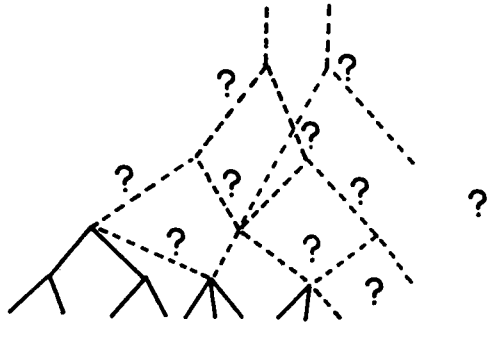


**Fig. 1.3.** Schematic representation for the knowledge structure of a musical performance from a notated score.

and so on. It is a feature of generative structures that they can be perfectly represented only when all the information has become available. In the course of assembling a representation from the raw data of notation, mistaken assumptions and oversights will inevitably result in a flawed structure that must be continually reassessed and retrospectively modified in the light of new evidence. Such re-evaluations are jeopardized by memory limitations, and may also lead to constructive rationalizations which distort the true structure of the music. In short, a performer playing unfamiliar music from notation is viewed as working from a generative representation that is incomplete in certain respects, almost certainly limited in the depth to which it extends, and subject to continuous modification in response to unfolding musical developments.

As a final example, let us consider the knowledge structures of improvised performance. Improvisation itself varies from the comparatively constrained conditions found in traditional jazz, where a harmonic sequence and metrical structure must be closely adhered to, to the almost entirely undetermined format of free improvisation. Taking the freer end of this spectrum, an improvising performer is necessarily working from a structural representation that is radically incomplete, since s/he does not yet know how the structure will turn out. A performer may start out with a more or less clear idea of the overall shape of the piece, but will typically have only the vaguest idea of how that shape will be realized. Alternatively, s/he may have no preconceived ideas whatsoever concerning the overall outline and may simply start off from a comparatively low-level structure —a motivic germ for a piece.

The questions are therefore how some sort of structure is developed (if it is), and whether that structure is generative in nature. Sudnow (1978, 1979) has provided an account of some of the processes involved in playing improvised jazz on the piano, based on an introspective analysis of his own playing. He takes the view that improvised piano playing should be viewed as the creation of 'an improvisatory hand' (Sudnow 1979, p. 14) and the logic and necessity of its movements, rather than as the outcome of abstract musical thought, with the hands acting simply as executive agents.

In accounting for the way in which these movements are structured, Sudnow questions the extent to which a hierarchical, generative process is involved, particularly at high levels of skill. Three stages of skill acquisition can be identified in Sudnow's account, differing in the extent to which they involve generative structure. At an early stage, his improvisations consist of relatively rigid routines that are executed in an invariant fashion, and which are structured together either according to the dictates of a predetermined framework, or simply by virtue of what comes most easily to the fingers. Such improvisations are therefore hierarchical up to a certain level (the level of the individual routine), and then either

predetermined (and usually hierarchical in the case of jazz 'standards') or regulated by manual dexterity at higher levels. However, the rigidity of the individual units, and the problems of accommodating them to a variety of overall frameworks, cause Sudnow considerable problems in performance, and result in a rather frantic and discontinuous style.

The second stage is to break down the preformed gestures of the first stage into their individual elements so as to recover the flexibility that his performance lacks. This is relatively easy to achieve, but presents a problem which the final stage must overcome: a method of structuring the elements together that achieves coherence without rigidity. Sudnow implies that the strategy he adopts is principally associative, or if hierarchical then only to a very limited extent. His aim is to start an improvisation with few preconceptions as to how it will develop, and to allow the hand to explore new consequences and ramifications of its current activity in an unfolding exploration of the keyboard. The impetus to continue comes not from an overall scheme into which the material must fit, but from a continual monitoring of the events just completed that generates ideas for events just ahead. The style is thus characterized by a precarious, unpredictable, associative structure.

> a single voice at the tips of the fingers, going for each next note in sayings just now and then, just this soft and just this hard, just here and just there, with definitions of aim throughout, taking my fingers to places, so to speak, and being guided, so to speak. (Sudnow 1978, p. 152)

A note of caution, however, must accompany Sudnow's account. It has been demonstrated by many authors that complex motor skills cannot achieve fluency and accuracy if they do not develop the hierarchical structures of motor programmes. Shaffer (1981) cites the example of moving a stylus to a number of targets with and without advance information, the former condition leading to hesitant, slow, and inaccurate performance when compared with the latter. Similarly, in his own work (Shaffer 1976), Shaffer showed that a copy-typist given no text preview (and thus unable to construct a motor programme) was reduced to a low level of speed and fluency when compared with her performance when she was given an eight-character preview. As far as motor control is concerned, typing is very similar to playing music, and it demonstrates that a literal interpretation of Sudnow's claims is untenable. An improviser must construct a representation for at least a short sequence of events in advance, and cannot operate at speed or with any fluency on an event-to-event level. Sudnow may have become unaware of these hierarchical structures, and the structures themselves may be no more elaborate than the eight-character preview of Shaffer's typist, but they undoubtedly exist.

The associative structure that Sudnow describes corresponds to a particular style of jazz improvisation, and in a different stylistic context a

more considered approach is often required. This can either take the form of carefully constructed and rehearsed schemes, involving complex extended melodic and harmonic strucutres—in others words hierarchical/generative structures—or can be characterized by a process of selecting patterns from a fixed repertoire. With this latter genre in mind, Sudnow comments that:

> Some musicians do no more improvising than the stand-up comic who decides only in which order to tell the same jokes the same way at each performance. (Sudnow 1979, p. 43)

The various representations underlying an improvised performance can be brought together by considering the abstract representation of an improvised performance in its very earliest stages, as shown in Fig. 1.4. All

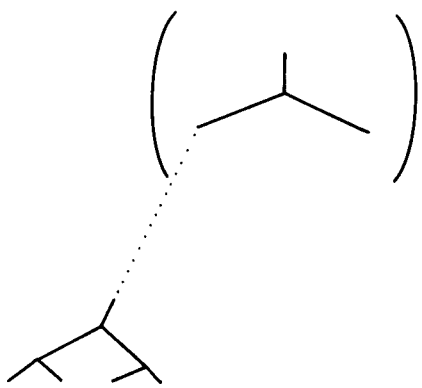


**Fig. 1.4.** Representation of the musical knowledge structure at the start of an improvised performance.

that exists is a low-level musical unit, characterized here as a small-scale, hierarchically organized event. A complete performance will consist of a large collection of such events, organized in different ways, and related to this first event according to three possible principles:

(1) The first event may be part of a hierarchical structure, to some extent worked out in advance, and to some extent constructed in the course of the improvisation. Figure 1.5 (a) is a schematic illustration of a complete piece formed according to this principle.

(2) The first event may be part of an associative chain of events, each new event derived from the previous sequence by the forward transfer of information. Figure 1.5(b) illustrates a complete piece formed on the basis of this principle.

(3) The first event may be selected from a number of events contained within the performer's repertoire, the rest of the improvisation consisting

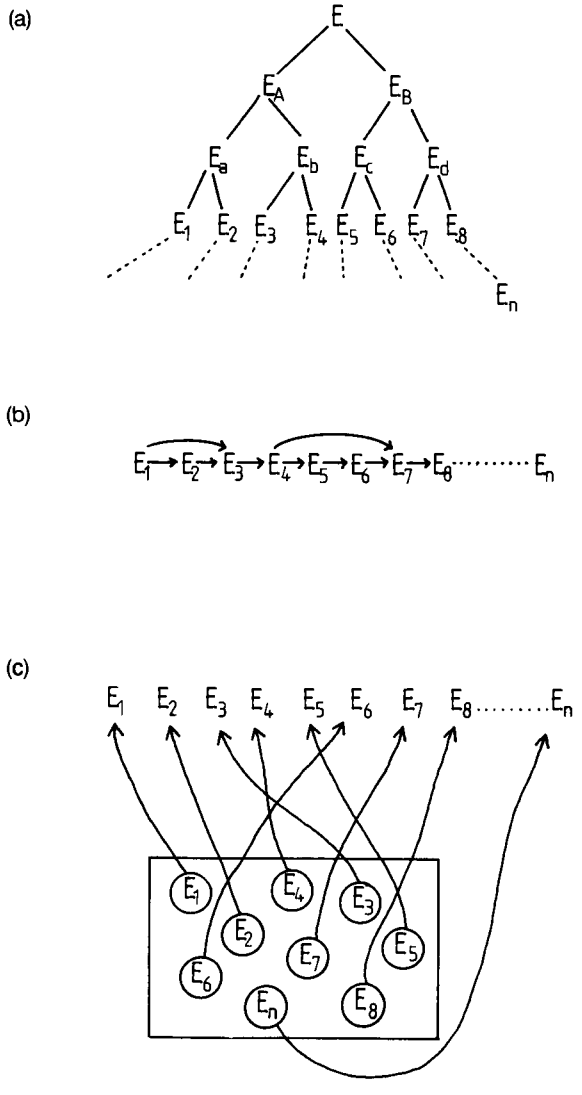



**Fig. 1.5.** The idealized knowledge structures for improvisations structured in three different ways: (a) hierarchically; (b) associatively; (c) by repertoire selection.

of further selections from this same repertoire, with a varying degree of relatedness between selections. Figure 1.5(c) illustrates this principle.

In any real performance all three principles will contribute to the improvisation. As we have seen, performance structures are never entirely associative, or perfectly hierarchical, and the selection of patterns from a

repertoire of possibilities will necessarily take place within an overall framework (associative or hierarchical) of some sort, however vague. It may be, however, that different improvising idioms can be characterized by the balance of the three principles, and the interactions between them. The improvising style known as free jazz is principally characterized by associative structure, since it eschews the constraints of a pre-planned structure, and attempts to avoid the use of recognizable 'riffs'. More traditional jazz improvisation tends towards the hierarchical principle, in its adherence to a fairly strict harmonic outline. And be-bop improvisation illustrates the selective principle in the way in which a performer may try to construct an improvisation so as to include as many 'quotes' from other sources as possible (ranging from other jazz pieces to national anthems).

To summarize the main points of this section, the representational structure of musical knowledge can be shown to have a generative component in all performance contexts. However, the strength of that generative component differs considerably when the variety of performance contexts (memorized performance, performance from notation, improvisation) is considered. Memorized performance illustrates the most pervasively generative representational structure, though it seems likely that at any moment during a performance only a part of that structure is active. Performance from notation demonstrates a less complete generative structure which, in the case of music that is poorly known or sight-read, is constructed from the projections of a player's stylistic knowledge coupled with the continuously unfolding input of detailed information derived from the score. The result is a hierarchical structure that is incomplete in certain respects at high levels, and perhaps at odds with the music's formal structure. These are the errors and omissions which are rectified through extended practising and general musical training. Finally, improvisation illustrates a performance context in which musical representations of a variably generative character are constructed on the spot, or with an element of pre-planning, according to the principles of hierarchical, associative, and selective elaboration. The balance of these three may offer a succinct way in which to characterize improvising idioms.

As stated at the outset, no radical discontinuity need be imagined between the abstract, and essentially static, representational structures discussed here and the representations for action contained within a motor program. Indeed, Shaffer (1976, 1981) has shown that in both typing and musical performance abstract properties play a vital role in the control of performance. None the less, motor programs result in the generation of real activity in a way that abstract knowledge structures do not. At a certain point declarative knowledge becomes procedural, and the term generative comes to mean not description, but production. One area of particular interest in which such a process is crucially implicated is the control and shaping of the expressive properties of a performance.

## Generative processes in performance expression

Whenever a skilled performer plays music, whether at sight or after years of practice, alone or in an ensemble, s/he plays with expression. Continuously variable modifications of the timing, dynamic, articulation, vibrato, and timbre of notes and note groups are identifiable and can be preserved with astonishing precision from one performance to another, sometimes separated by years. On the other hand dramatic changes in these expressive parameters can be observed between successive performances of the same piece of music separated by only a few minutes. These observations prompt the following questions: how is it possible for a performer to introduce expression into a performance when the music is unknown, to preserve an interpretation over long periods of time, and to alter an interpretation at a moment's notice?

The options available that account for these facts are limited. Either one must assume that a performer has simply learned a fixed expressive pattern that is applied to each piece of music, and exists in a number of fixed variants; or one must assume that the expressive profile is generated at the time of performance from information specified in the musical structure. The first option is not really a serious possibility since expressive profiles have been shown by a number of authors (Shaffer 1981, 1984; Sloboda 1983; Clarke 1984) to be closely related to musical structure. Given the structural differences between pieces of music even within the same style, it is inevitable that a mismatch between structure and expression would result from the application of a fixed expressive profile to different pieces. There is no plausible alternative, therefore, to the idea that expression is derived from structure.

### Structure and expression

Some aspects of the relationship between structure and expression can be illustrated with two different studies of piano performance. The first (Clarke 1982) is primarily concerned with the interaction between rhythm and performance tempo, and demonstrates that the changes in expressive timing structure that accompany changes in tempo relate to structural properties of the music. The performances (of a highly repetitive piece by Erik Satie called 'Vexations') show consistent deviations from strictly metrical timing that produce a profile of partially periodic timing curves. Maxima in these curves indicate boundaries which segment each performance into a number of groups. Significantly, these group boundaries correspond to positions in the musical structure at which formal breaks can be identified, using analytical principles similar to those of Lerdahl and Jackendoff (1983). At faster tempi there are relatively fewer groups in

the timing profile, and at slower tempi relatively more, the additional boundaries being formed at positions where the music indicates a structural discontinuity of some sort. Thus the expressive changes that accompany changes in performance tempo are based on structural properties of the music, and can be characterized as the transformation of latent expressive possibilities into manifest expressive features in accordance with the dictates of tempo and musical structure.

A second study investigated the expressive changes that result from modifying the relationship between a constant rhythmic structure and its metrical context. The metre of a piece of music is a regular framework of accented and weak beats, existing at a number of levels of structure, around which individual notes and groups of notes are organized. A partial representation of metrical structure is contained in the time signature, bar lines, and beams connecting notes that are used in conventional music notation. In the study a single-line melody was constructed, which was written so as to start in a variety of positions within the bar while keeping all the note values and pitches constant. The result was a complete cycle of orientations between the metrical framework and the note sequence, such that the sequence started on each of the available beats in the bar. Experienced pianists were asked to play each of these variants a number of times, in random order. Figure 1.6 shows two of the variants as examples.

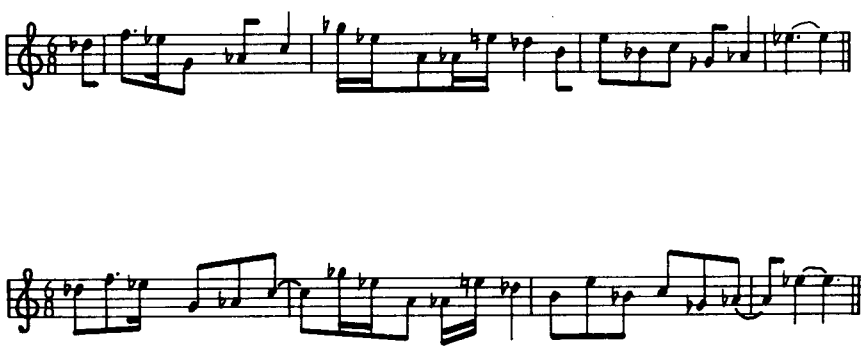


**Fig. 1.6.** Two of the metrical variants of a constant pitch/duration sequence used in the study of interactions between rhythm and metre.

The study showed that significant changes in the expressive profile of the sequence accompanied the cycle of metrical reorientations, and could be identified in all three of the expressive parameters available (timing, dynamic, and articulation). The expressive changes are, however, related to the metrical changes in a manner that is rather more complex than might be expected, for the following reasons.

(1) Changes in the notated metre of the sequence precipitate other

structural changes. These include changes in the position of group boundaries and changes in the melodic/harmonic function of notes brought about by an altered relationship to the main metrical accents. Figure 1.7 shows two of the tunes used in the study which illustrate this point. The metrical change is simply specified and is directly observable in the notation, while changes in group structure and melodic/harmonic function (see Fig. 1.7) are more concealed since they are not directly represented in the notation. Expression is related to a whole range of structural features, and will reflect these second-order structural changes no less than the metrical changes. Added to this are further structural effects brought about by what may be termed the functional resistance of notes: a short note, for instance, resists functioning as a group ending simply by virtue of its brevity, and similar resistances can be identified for long notes.

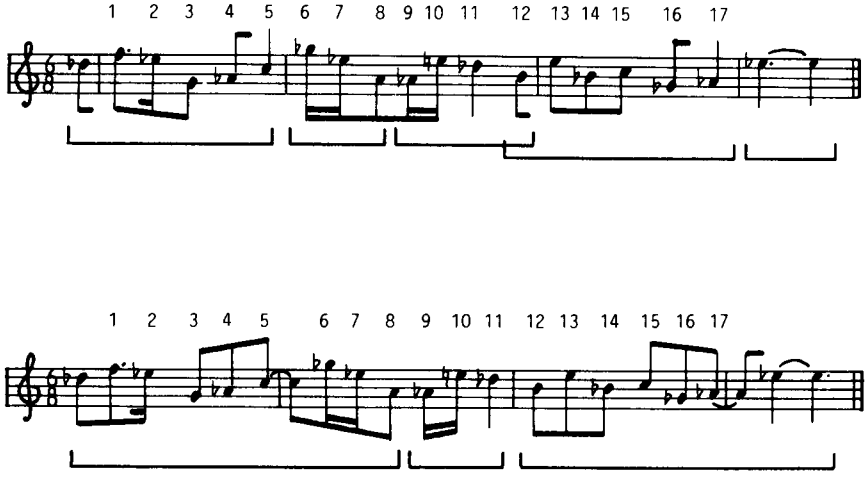



**Fig. 1.7.** Two tunes from the rhythm/metre study showing changes in group structure (indicated by brackets) and melodic/harmonic function that are triggered by metrical changes. Note 3 functions as an appoggiatura to note 4 in tune 2, but not in tune 1; and note 16 functions as an appoggiatura to note 17 in tune 1, but not in tune 2.

(2) Within each of the three expressive parameters (timing, dynamic, articulation) expressive gestures can perform a number of different functions: these include altering the relative proportions of events within a rhythmic group, indicating the position of a group boundary, marking a metrical accent, and creating an expressive gradient towards a focal point (that is, a pattern of directed motion). As a result expressive gestures are functionally ambiguous, in the sense that they specify a number of alternative interpretations. It is only through interactions with the underlying musical structure that these ambiguities are resolved.

(3) The three expressive parameters interact with one another in at least two ways: first, they may substitute for one another, so that, for instance, an accent may be conveyed by making a note longer, or louder, or more legato, or delaying its entry; and second, they may combine to form expressive complexes that possess a compound function that is not simply the sum of the individual expressive components. For example, in the appropriate context a note that is played loud and staccato will receive a particularly emphatic accent, despite the fact that staccato articulation is usually associated with metrical weakness. The combination of dynamic intensity and staccato articulation appears to convey an expressive meaning that is not explicable in terms of a summation of the individual components.

In the study described here, most expressive changes could be explained on the basis of changes in the position of metrical accents (those points being played louder, or longer, or more legato) and group boundaries (indicated by points of discontinuity in the timing, dynamic, and articulation curves), a more minor role being the emphasis of peaks in the melodic contour. With more complex musical material, the subtlety and complexity of the relationship between expressive features and the associated structure would undoubtedly increase.

What this study demonstrates, however, is that a performer generates an expressive strategy for a musical extract on the spot. This is evident from the brevity of the preparation time that the players needed for the study; from the way in which the players were able to activate an appropriate expressive strategy almost immediately when going from one variant of the sequence to another randomly selected sequence; and finally from the evidence that in many cases a player used a number of different expressive strategies for the same sequence on different repetitions, each strategy nonetheless conveying the main structural characteristics of the material. This suggests strongly that a performer is aware of a number of interpretative options in his or her approach to the music, which might emphasize different structural aspects, but which are in some sense equally adequate. In order to understand how this range of options is structured and controlled, the nature of an interpretation and the general function of expression must first be considered.

## Interpretation and the function of expression

Shaffer (1984) has described an interpretation as a compact coding of expressive forms, from which an expressive performance can be generated when required. Two qualifications, or additions, can be made to this. First, the expressive forms which constitute the interpretation are rather abstract. This is evident in the way in which performers use a variety of different expressive strategies to project essentially the same interpretation,

as mentioned in connection with rhythm and metre. This suggests that an interpretation consists of a set of abstract expressive markers that can take a material expressive form within any of the parameters available. Second, an interpretation is not only an expressive coding, but also a structural coding. A performer must form an understanding of the musical structure, or decide between structural alternatives offered by the music, and encode that in some stable and compact manner. The structural component then acts as a framework around which the expressive markers are organized.

Pieces of music are invariably open to a number of different structural interpretations, and the primary role of expression is to limit the extent of this ambiguity by emphasizing certain structural interpretations at the expense of others. This does not mean that expression necessarily operates in the service of clarity, since a performer may choose to emphasize either those structural characteristics that are most obvious, or to extract more hidden and disruptive features of the music. Similarly, although a performance must aim to be expressively coherent, that aim does not entail the resolution of all structural ambiguity, since a performance may intentionally attempt to manipulate conflicting aspects of the music.

At a detailed level, however, each expressive act operates so as to project a particular functional meaning for a given musical structure. This is achieved in a variety of ways, the most general underlying principle being the intensification of gestalt properties of the musical structure that are already evident, or the establishment of gestalt features when the music is structurally neutral. Examples of this are the establishment of boundaries in the grouping structure of the music by means of changes in dynamic, articulation, or timing; the imposition or emphasis of a sense of direction towards a structural focal point by means of dynamic, articulation, or timing gradients; or the modification of the accentual status of events (that is, changes in figure/ground relations) by means of dynamic or agogic emphasis. (Agogic emphasis is the accentuation of a note by prolonging its inter-onset duration.)

In a majority of these examples the relationship between the expressive aim and the expressive means is very direct: boundaries are indicated by relatively large parametric changes, directed motion is indicated by graduated parametric increase, and accentual strength is indicated by relative parametric intensity. Within at least two of these parameters (timing and articulation), however, the directness of this expressive function is threatened by ambiguity. The lengthening of a note can indicate that it is accented, that it finishes a structural unit at some level, or that the following (delayed) note is of structural importance. Similarly, although an underlying correlation between structural significance and increased legato articulation appears to hold, structural significance can also be indicated by means of emphatically staccato articulation.

This uncertainty concerning the meaning of an individual expressive

gesture is usually clarified by two considerations. First, the expressive sequence in which a particular gesture is incorporated will frequently resolve any ambiguity, and second, the structural context in which the gesture occurs may clarify its meaning. Figure 1.8 shows a short musical extract with a hypothetical expressive profile in three parameters.The penultimate note receives an expressive treatment (long, legato, loud) which in isolation might indicate that it was a strong beat. In reality the correct interpretation of the note as an emphasized upbeat (weak beat) is virtually guaranteed by the previous expressive profile, which helps to establish the metre, and by the structural context, which makes the note weak and unstable. Under conditions of structural clarity, and for listeners well-versed in the musical idiom, expressive characteristics function as responses to, or refinements of, the immanent properties of the music.

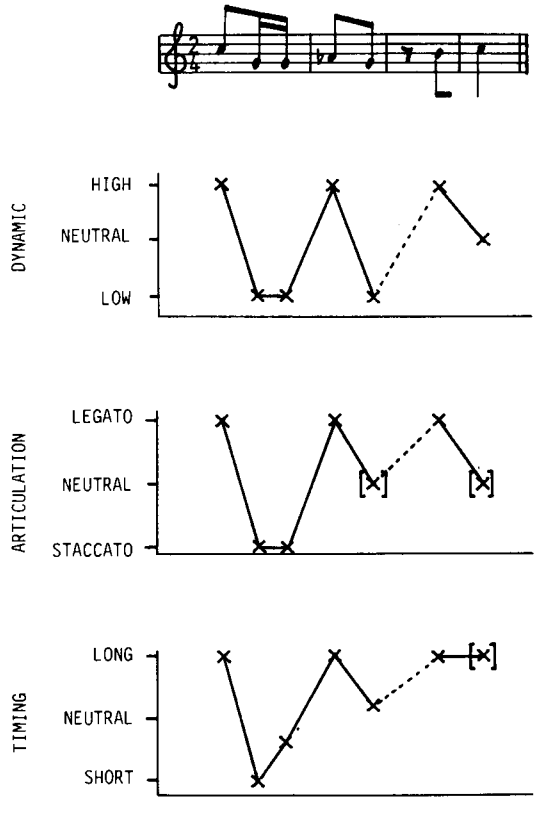



**Fig. 1.8.** A short melodic sequence with associated hypothetical expressive features, illustrating the way in which the expressive and structural context affect the meaning of an individual expressive gesture (on the penultimate note). See text for details.

When the musical structure is weak or indeterminate, however, expressive effects may function primarily to impose a particular structural interpretation onto a neutral structural base.

The distinction betweens strong and weak musical structures may also affect the way in which listeners distinguish betwen expressive deviations and mistakes. A strong structure may be sufficiently stable that if a feature of the performance conflicts radically with the structural implications of the piece, it is perceived as a mistake, or, if less prominent, simply ignored. In an ambiguous piece, on the other hand, a listener is likely to be more strongly influenced by expression and is more or less forced to take at face value anything that a performer does. Strong structures therefore act selectively to filter out inappropriate expressive features, which are then labelled as mistakes by a listener, or not noticed. Weak or unfamiliar musical structures, on the other hand, provide a listener with no clear reference against which to judge performance features, which must therefore be accepted as structural indicators, with all the ambiguities that have been discussed. In these circumstances the performer is under particular pressure to pay the most acute attention to details of expression, since an unintended deviation may jeopardize a listener's understanding of the music. It may be for this reason that the performance of contemporary music puts such great expressive demands on performers, since both listeners and performers are on the whole unfamiliar with its structural principles.

**The generative source of expression**

The close relationship between structure and expression embodied in an interpretation forms the basis for an account of the source and systematic construction of expression. Using the structural component of an interpretation as its input, a set of generative rules derives an expresive strategy as output. A range of studies (for example, Shaffer 1981, 1984; Clarke 1982, 1984; Sloboda 1983) indicate that the generative rules for expression in piano performance can be summarized as follows. Within the domain of timing three principles operate:

(1) Graduated timing changes can be used to indicate the group structure of the music, group boundaries being marked by maxima in the timing profile. This principle can be described by a quadratic function, with the position of the minimum point in the curve specified by a variable parameter, allowing a range of related curves to be generated that differ in the balance of upbeats and afterbeats. When the minimum point is displaced to the left in a group, upbeats predominate, and the internal dynamic or directed motion of the group is towards its ending. When displaced to the right, afterbeats predominate and the internal dynamic or directed motion of the group is a dissipation away from its beginning.

Figure 1.9 illustrates this principle with three schematic curves which differ in the balance of upbeats and afterbeats. Todd (1985) has formalized this function using a parabolic curve, and has shown that with only a single function at one level of structure a reasonable approximation to some of the piano data can be obtained which also reflects structural properties of the music being performed. More recent attempts at elaborating this formalization, using hierarchically nested levels of the function (Shaffer 1985), indicate that even better approximations can be achieved with only a minimal increase in formal complexity. It appears that the amount of timing modification is directly related to the structural significance of a musical segment. By displacing the minimum point of the parabola relative

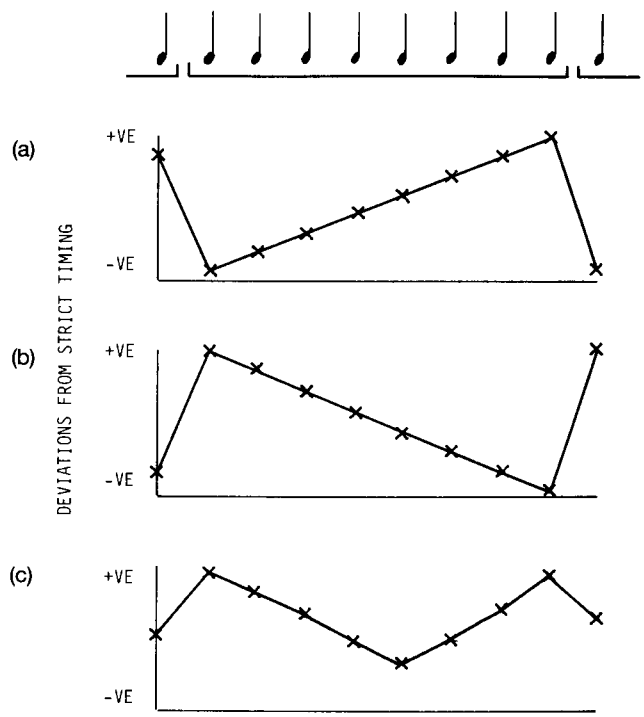



**Fig. 1.9.** Schematic diagram of three related timing curves differing in their pattern of upbeats and afterbeats, and produced by displacement of the minimum point in the curve to left or right. The top of the figure shows a succession of notes with group structure indicated by brackets. In (a) the minimum is displaced to the left, resulting in a predominance of upbeats and a sense of end-oriented direction. In (b) the minimum is displaced to the right, resulting in a predominance of afterbeats and a sense of dissipation away from the beginning of the group. In (c) the minimum is positioned roughly in the middle of the group, giving a balance of upbeats and afterbeats, and a centred focus of direction.

to the segmental boundaries according to criteria of structural significance, a timing curve of appropriate slope and degree of rubato is generated.
(2) A note may be lengthened so as to delay a following note, the function of the delay being to heighten the impact of the significant delayed note. This function is frequently associated with the first principle, since the delayed note is often the start of a new group, the previous note being lengthened both for reasons of delay and because it falls at a group boundary. Delay is therefore only observably distinct when it occurs mid-group, or when applied to isolated notes that are not part of a graduated timing curve.
(3) A note that is structurally significant may be emphasized by increasing its inter-onset interval (agogic accent). Since significant events usually occur towards the beginning of musical groups (see Lerdahl and Jackendoff 1983), the principle balances the predominantly end-effects of the first principle.

Within the domain of dynamics, three principles can be identified:
(1) Graduated changes in dynamic may be used to indicate a pattern of upbeats and afterbeats, and directed motion, within the group structure of the music. This is essentially a translation of the first timing principle into the dynamic domain and should be formalizable in the same way, namely by a quadratic function fitted to the group structure of the music, with the position of the dynamic minimum specified by a variable parameter that alters the balance of upbeats and afterbeats. Figure 1.10 illustrates the principle with performance data from the rhythm and metre study discussed earlier. The figure shows two performances of the same tune, in which the positions of the dynamic maxima remain the same (positions 1, 4, 6, 11, 13 and 16), but the positions of three minima change. This has the effect of transforming notes 3, 9, and 10 from upbeats in the first performance into afterbeats in the second, and notes 14 and 15 from afterbeats in the first performance to upbeats in the second.
(2) Dynamic contrast may be used to articulate group boundaries, employing an association between dynamic discontinuities and underlying structural discontinuity. As an illustration, Fig. 1.11 shows dynamic data from performances of another tune in the rhythm and metric study. The sharp change in dynamic level between positions 8 and 9 coincides with the first clear opportunity in this tune of establishing a group boundary.
(3) Significant structural events (for example metrical accents, melodic peaks, harmonic changes) may be dynamically intensified, a principle equivalent to the third timing principle.

Within the parameter of articulation, three principles can once again be identified, equivalent to those of dynamic usage:
(1) Graduated changes in articulation may be used to indicate a pattern of upbeats and afterbeats, and directed motion, within the group structure of the music. This appears to be rather less distinct as a principle than its

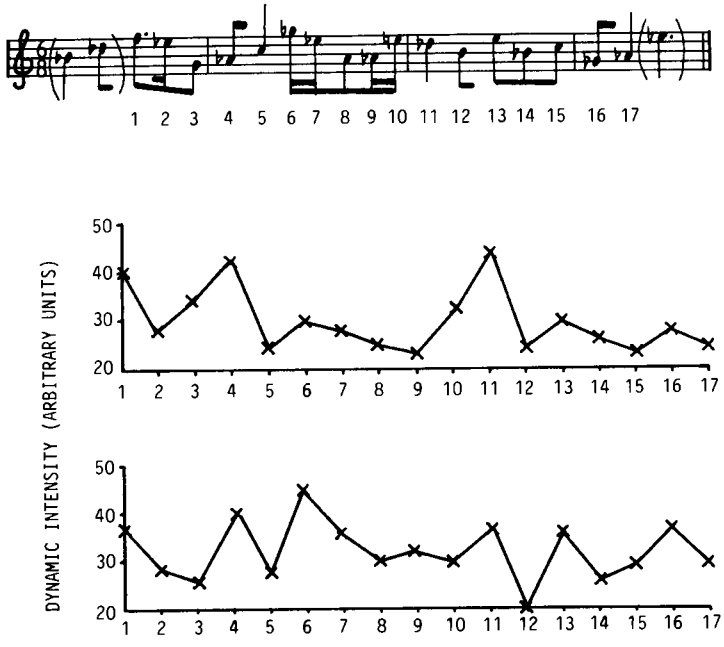



**Fig. 1.10.** Dynamic data from two performances of the same tune showing changes in the pattern of upbeats and afterbeats reflected in displacements of three dynamic minima.

dynamic equivalent, and cannot be formalized with the same quadratic function. It is more often associated with upbeats, a pattern of increasing staccato indicating the approach to a downbeat, though it can also be used to indicate afterbeats.

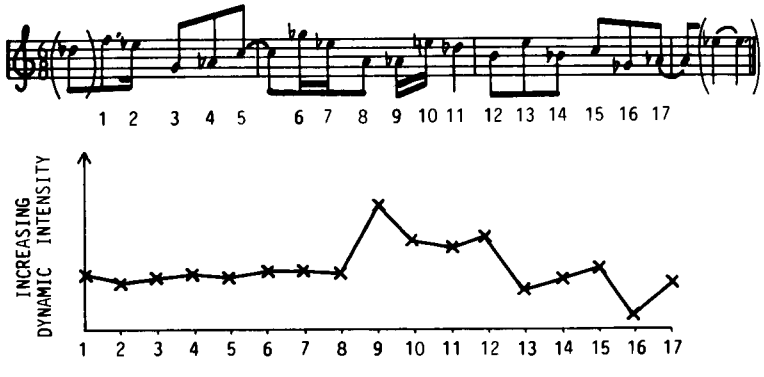



**Fig. 1.11.** Dynamic data illustrating the use of dynamic contrast to mark a group boundary between notes 8 and 9.

(2) Group boundaries may be indicted using discountinuities of articulation. This can either take the form of sustained articulatory contrasts that define different groups, or the use of an isolated staccato (or detached note) at a boundary as a separating device. This is illustrated in Fig. 1.12, where the isolated staccato articulation at note 8 serves to establish a group boundary. Conversely, an exaggerated legato articulation can be used as a device to hold together the components of a group.

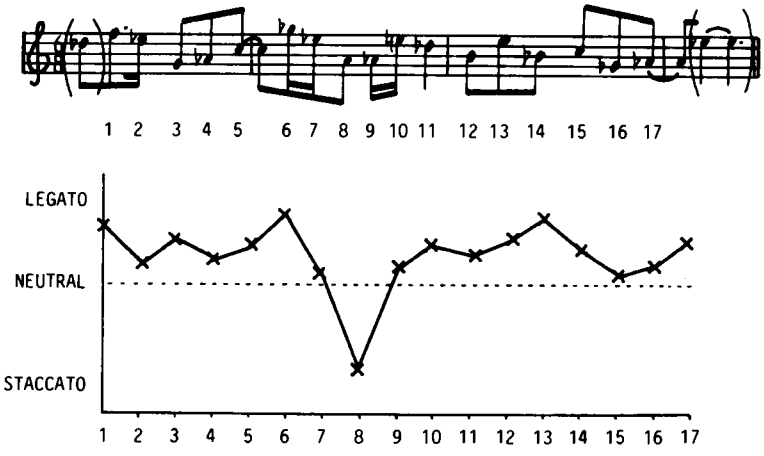



**Fig. 1.12.** Articulation data illustrating the use of an isolated staccato note (data point 8) to mark a group boundary.

(3) Individual events may be accentuated by articulation. There appears to be an underlying association between structural significance and legato articulation, although this association is rather weak, and in the correct structural context staccato articulation may serve to accentuate a note (Clarke 1984).

These, the essential principles of expression in piano performance, demonstrate three underlying expressive functions:
(1) The indication of structural direction by means of parametric gradients.
(2) The indication of group structures by means of parametric continuities and discontinuities.
(3) The accentuation of individual events by means of local parametric intensification or contrast.

A number of factors influence choices within this system of expressive options. To begin with there are the resources of the chosen instrument, which vary considerably in restrictivenesss. At one end of the spectrum only timing and articulation can be varied on a harpsichord, whereas the voice, for instance, offers expressive opportunities that include timing, dynamic, articulation, timbre, verbal content, vibrato, and the continuum

between speech and song. The piano falls towards the restricted end of this range, offering systematic control over only timing, dynamic, articulation, and pedalling. Since the same piece of music may be played on different instruments, a performer who switches from one instrument to another may have to make a considerable readjustment to the resources that the instrument offers. Quite apart from the different tone qualities of the two instruments, the differences in the use of timing in harpsichord and piano performances of Bach's keyboard music illustrate this adjustment. Performers who play both instruments confirm that they must make a conscious stylistic modification to suit the instrument. On a more refined level, changes in instrument design and construction affect performance expression: the range of dynamic contrast offered by a modern grand piano was not possible on the pianos available to Mozart or Haydn, and expressive effects are consequently obtained in different ways on the instruments of that period.

The second factor affecting expressive options is the influence of performance style. Just as instruments show historical development in their design and construction, so also do conventions of performance expression show significant historical (and geographical) variation. These may be linked to instrument technology, but are by no means simply the consequence of such changes. Writings on performance practice have a long history (see Donington 1977), but it is only in the last 30 years or so that the study of historical changes in performance practice, and the attempt to reconstruct stylistically authentic performances, have become established.

Writings on performance practice vary from prescriptive statements to rather more considered and closely argued discussions of particular problems. Whether well-informed or not, performers certainly have opinions about the manner in which different styles of music should be played, and adapt their expressive usage accordingly. The debate over the use of rubato and vibrato in performances of baroque music exemplifies the differences of opinion among performers over the need for, and nature of, historical authenticity.

We can regard a stylistic convention as a particular configuration of the expressive principles set out above, that brings about a certain type of expressive characterization. In Baroque performance practice, for example, group-final lengthening (the first of the three timing principles) is used very little, while the second and particularly the third principles (delay and agogic accent) are used more extensively. This results in an overall absence of ritardandi and a more widespread use of agogic accents, coupled with delay, particularly of the final tonic at cadences. By contrast, the nineteenth-century Romantic style placed enormous emphasis on the first timing principle, and the first dynamic principle (by contrast with Baroque and Classical layered dynamics, which correspond to the second dynamic

principle). It should be possible to charactize the performance practice of any period in this manner, as well as the personal style of a performer whose performance practice is sufficiently consistent.

Matters of stylistic convention may be broadly considered part of the performance context—in this case historical context. The details of expression are, however, also sensitive to contextual features of a far more local and immediate kind. It is this detailed context that constitutes the third factor operating on the range of expressive options. Sloboda (1983) has pointed out that it is implausible to assume that the expressive features of a performance are triggered by individual notational symbols. This means that structural information must be integrated from a collection of notational symbols before it can form the basis for an expressive strategy. It must also be supplemented by information about the current expressive context, and any logistical constraints that might apply: it is, for example, impossible to play notes legato that do not lie within the hand's span (without using the sustaining pedal).

Although performers make use of a variety of expressive strategies in the same structural context, this does not mean that expressive features can therefore be freely substituted. Each parameter has certain intrinsic qualities that make it more or less suitable for use in any given context. While articulatory differences may be good indicators of grouping properties, they are less effective than dynamics for outlining the shape and direction of large-scale melodic processes by virtue of their disruptive and divisive characteristics. On the other hand, when playing in an ensemble it may be more effective for a soloist to project an individual line by means of articulation than with dynamics, which might result in an unacceptable overall dynamic level or a muddy texture. Similarly, rubato is effective in solo performance, but in ensemble performances must be more constrained if overall synchrony is to be preserved, although Shaffer (1984) has shown that considerable rubato can be used in piano duet performances. Furthermore, even in solo performances, rubato must be used with discretion when the music involves a high level of rhythmic complexity, so as to avoid obscuring the structure.

The final factor influencing the choice of expressive options is the manner in which a player choses to perform. Seashore (1938/1967) demonstrated that a performer could play the same piece of music with an expressive profile that differed according to whether he was instructed to play artistically or metronomically. The differences were principally in the degree of rubato, rather than the pattern of rubato, but there is every indication that a performer can choose to alter the pattern as well. Sloboda (1983) showed that a professional performer who was experienced in playing exam dictations conveyed the metre of a short tune to listeners more effectively than a performer of similar calibre who did not undertake such duties, apparently by adopting a more didactic approach to the

relationship between metre and expression. This suggests that performers can choose to play in a manner that emphasizes the metre of the piece, or some other basic structural feature, or may subtly contradict those structural features for artistic effect.

Clearly a performer must not contradict the structure to such an extent that the music becomes incomprehensible to a listener, but a performer must also avoid playing music in a trite and obvious way. The choice of a 'mode' of performance is therefore influenced by the audience and occasion, as well as the performer's artistic intentions. The style of performance in a concert recital is likely to be rather different from the same performer's style at a children's concert, or during a melodic dictation exercise.

At a more subtle level, an individual interpretation can be regarded as a personally selected patterning of the expressive principles that may persist over a long period of time, or may be quite ephemeral. The expressive principles will be set in a particular configuration according to considerations such as the desire for contrast or continuity, and novelty or consistency. In a rondo movement, for instance, a pianist may choose either to retain the same expressive configuration at each return of the rondo theme, if the aim is to achieve continuity and consistency, or to make use of as wide a variety of expressive characteristics as possible, if the aim is to convey the potential novelty of the theme on each of its appearances. In a sonata form movement, on the other hand, it is usual to play the recapitulation with essentially the same expressive characteristics as the exposition, in order to make clear the significance of the return to material that has already been heard. As a third example, in a highly contrapuntal piece of music with a number of distinct motivic elements, it would be appropriate to give each an identifiable expressive profile if motivic clarity and identity are to be achieved. The structural 'archetypal schemata' that Meyer (1973) proposes may therefore have their expressive parallels.

Finally, the broad aims of a performance must be considered. Sloboda (1983) has pointed out that a performance can be considered to have primarily artistic or primarily communicative aims. If communication is the primary aim, then it is not too difficult to prescribe a successful configuration of the expressive principles outlined above. If, on the other hand, artistry is the primary aim, then prescription is both impossible and inappropriate. The essential characteristic of artistic activity (and aesthetic objects) is a radical form of ambiguity and creativity, and while the expressive resources may be outlined, their precise disposition on any occasion can be accounted for only in retrospect, not predicted. Were this not the case, and our curiosity in these ambiguities and possibilities not boundless, we might all have given up going to concerts long ago.

## References


Bengtsson, I. and Gabrielsson, A. (1983). Analysis and synthesis of musical rhythm. In *Studies of music performance* (ed. J. Sundberg). Publications issued by the Royal Swedish Academy of Music 5 No. 39.

Chomsky, N. (1957). *Syntactic structures.* Mouton, The Hague.

Clarke, E. F. (1982). Timing in the performance of Erik Satie's 'Vexations'. *Acta Psychologica* **50**, 1–19.

Clarke, E. F. (1984). Structure and expression in the rhythm of piano performance. Unpublished Ph.D. thesis. University of Exeter.

Donington, R. (1977) *The interpretation of early music.* Faber and Faber, London.

Forte, A. (1973) *The structure of atonal music.* Yale University Press, New Haven.

Gates, A. and Bradshaw, J. L. (1974). Effects of auditory feedback on a musical performance task. *Perception and Psychophysics* **16**, 105–9.

Krumhansl, C. (1983). Perceptual structures for tonal music. *Music Perception* **1**, (1), 28–63.

Lerdahl, F. and Jackendoff, R. (1983). *A generative theory of tonal music.* MIT Press, Cambridge, Mass.

Longuet-Higgins, H. C. (1976). Perception of melodies. *Nature* **263**, (5579), 646–53.

Meyer, L. B. (1956). *Emotion and meaning in music.* University of Chicago Press.

Meyer, L. B. (1973). *Explaining music: essays and explorations.* University of California Press.

Narmour, E. (1977), *Beyond Schenkerism: the need for alternatives in music analysis.* University of Chicago Press.

Narmour, E. (1983). Some major theoretical problems concerning the concept of hierarchy in the analysis of tonal music. *Music Perception* **1**, (2), 129–99.

Seashore, C. E. (1938). *Psychology of music.* McGraw-Hill, New York (Reissued 1967 by Dover, New York).

Shaffer, L. H. (1976). Intention and performance. *Psychological Review* **83**, (5), 375–93.

Shaffer, L. H. (1981). Performances of Chopin, Bach and Bartok: studies in motor programming. *Cognitive Psychology* **13**, 326–76.

Shaffer, L. H. (1984). Timing in solo and duet piano performances. *Quarterly Journal of Experimental Psychology* **A36**, 577–95.

Shaffer, L. H. (1985). The expressive component in musical performance. Paper delivered at the Third International Conference on Event Perception and Action, Uppsala, Sweden.

Shepard, R. (1982). Geometrical approximations to the structure of musical pitch. *Psychological Review* **89**, (4), 305–33.

Sloboda, J. A. (1983). The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology* **A35**, 377–96.

Sudnow, D. (1978). *Ways of the hand: the organisation of improvised conduct.* Routledge and Kegan Paul, London.

Sudnow, D. (1979). *Talks body.* Alfred Knopf, New York.

Sundberg, J. and Lindblom, B. (1976). Generative theories in language and music description. *Cognition* **4**, 99–122.

Todd, N. P. (1985). A model of expressive timing in tonal music. *Music Perception* **3**, (1), 33–58.

# 2

# Timing in music performance and its relations to music experience

ALF GABRIELSSON

## Performance and notation of music

Our experience of music and our ideas about music mainly derive from listening to performed music. The characteristics of musical performances are therefore of fundamental importance. The performance of a piece of music represents the overt, auditory manifestation of the composer's and/or the performer's intentions, ideas, feelings, and whatever is being communicated in music. It is the direct stimulus to the listener and will profoundly affect his way of perceiving and understanding this piece of music. His impression from the performance may remain for years to come, perhaps for his whole lifetime. However, another performance of the same piece may make him revise his ideas about it in different directions. Music critics excel in discussing the merits of different performances, for example, how conductors such as Toscanini or Furtwängler shape a Beethoven symphony, how pianists such as Backhaus or Kempff play a Mozart sonata, or how Glenn Gould performs music by Bach. The relations between the composer's idea, the actually performed music, and the listener's impression are complex, indeed.

In Western art music as well as in much popular music there is also a visual manifestation of the music, the score. However, the 'translation' from the score into sounding music is obviously made very differently by different performers. The notation has to be supplemented by various implicit rules, which relate to the type of music and associated performance practice (for example, church music, dance music, Baroque music, music from the Romantic era), with which different performers are more or less familiar. Of course, they are also influenced by teachers, other performers, their own earlier performances, and their ideas about the meaning of the music they are going to play. The musical score is only a starting point for the performer. It provides him with a skeleton, which he should bring to life with sound. It is one of the fascinating aspects of music that the result of this creative process may become so varied for different musicians and in different situations.

Music may thus be said to exist at different levels and in different forms: as an idea (in a wide sense) in the composer's and performer's mind, in the written score, in sounds, and in the listener's experience. The communication between the composer/ performer and the listener is meant to be realized by the *sounding* music, not by reading of a score. In fact much music (such as folk music over the whole world, and certain jazz music) never appears in notation but relies totally on performance and aural tradition to spread and survive. Nevertheless most thinkers and writers about music take musical notation as the starting point for their discussion. They even provide notations of music which has never appeared in that form (for example, transcriptions of folk music).

There are several reasons for this emphasis on notated, rather than performed, music. The musical notation is in many respects very convenient and efficient. It is also easily reproduced in print, while the handling of phonograph records or magnetic tapes to provide sound illustrations is more clumsy and requires special equipment. Certainly many writers also point out the limitations of musical notation and make due reservations as to its appropriateness. Still, I think that it is very easy to be led astray, even unconsciously, by the properties of notation; they may affect our experience and conceptualizing of music in a similar way that words may do in other areas, encouraging us to 'force' pitches into a limited number of pitch categories, to write down a rhythm using the most common note values, etc. The results from many studies of performances provide further support for this opinion.

## Analysis of performances

The characteristics of different performances may of course be studied by simply listening to them. It is often easy to hear differences and changes in tempo, in dynamics, in intonation, articulation, and phrasing. For closer study of certain points you can use simple tricks; for instance, you can play tape-recorded music at half-speed, or even slower, to have more time for listening. However, many characteristic features of a performance elude our perceptual analysis and require accurate physical registrations. Here we meet one important explanation of the relative scarcity of performance studies: to make detailed registrations of sounding music is very difficult, even with the technology available today. I have elsewhere (Gabrielsson 1985) reviewed earlier attempts in this direction, including registrations of organ playing (Sears 1902), of piano performance stored on 'player rolls' for mechanical pianos (Hartmann 1932; Vernon 1937) or studied by filming the action of the hammers (Hendersson 1937; Skinner and C. E. Seashore 1937), of singing (H. Seashore 1937), and of violin playing (Small 1937). The last-mentioned researchers belonged to a very active group gathered

around C. E. Seashore at the Iowa University. Their techniques and results, summarized by C. E. Seashore (1938), are still very impressive. Later advances in registrations of piano performances were presented by Wagner (1971, 1974) and Shaffer (1980, 1981). The methods used in the Uppsala studies of performances are described below.

From a physical point of view the variables manipulated by the performer are durations, intensities (amplitudes), frequencies ('pitch'), and the spectral composition of the sounds. On some instruments (for example, the whole violin family) and in singing one may vary all of them over a considerable range, and continuously. In other instruments, such as the piano and the organ, pitch can be varied only in discrete steps, and once the tone is struck one has no, or very limited, possibilities of affecting its properties any longer. Each instrument presents its own possibilites and limitations. In the above-mentioned investigations on singing and violin playing there were numerous examples of more or less continuous variations in pitch and intensity, for example, pitch and intensity vibrato, sharp or flat intonation before reaching the 'proper' pitch, dynamic swelling and shrinking within tones or phrases, glidings between successive tones, etc. In the piano studies the interest was focused on durations and intensities. Striking variations of tempo and different types of deviations from expected durations of various tones appeared, and it was noted that there was no consistent relation between intensity and (perceived) accent. Of course, none of these phenomena in the actual performances were indicated in the corresponding musical scores; they were truly outside the notation.

## Timing in performance

The only variable over which the performer has practically complete control, regardless of which instrument he uses, is the duration of the sound events, as well as of 'non-sound' events (rests or silences). This circumstance is one of the explanations for why the manipulation of various durations, here generally called *timing*, is often considered to be the most important tool available to the performer. If there is a score, this prescribes, of course, the timing in many respects, but there is still a great freedom within these constraints. In the case of non-notated or improvised music this freedom is even greater.

In writings and research on music, timing is usually related to rhythm. This connection is self-evident, but other factors should also be mentioned. Since there is no melody without rhythm, timing is important for melody as well; and since a melody is often harmonized, timing influences the harmonic progression. It may also affect the single chord, since tones within a chord are often played asynchronously (Hartmann 1932; Vernon

1937) or even *arpeggio*. It is known that the time course of different partials in a tone is critical for the perceived timbre (Risset and Wessel 1982). Rasch (1978, 1979, 1981; see also Chapter 4 in this volume) showed that the tone onsets from different instruments performing polyphonic music are usually asynchronized, and that this may facilitate the perception of the different parts.

These hints are enough to show that timing, in one way or another, influences practically every aspect of the music. Instead of using the traditional musical terminology, we may also refer to such categories as the *structural*, the *motional*, and the *emotional* aspects of music experience (Gabrielsson 1973*a*, *b*, *c*; 1982*a*, *b*; 1985; 1986). It is clear that timing affects the way we experience the structure of a piece of music, as a whole and at different levels (Sloboda 1983; Gabrielsson 1986; Clarke in Chapter 1 of this volume). However, little is said about the effects of timing upon the experienced motion character of the music and the related emotional qualities. These two aspects will be briefly discussed before proceeding to different examples of timing.

## Motion and emotion in music

Music is often said to be 'the language of emotions', which 'bypasses the intellect' and takes over, when words do not suffice any longer. Sweeping statements like these require much critical analysis, but this will be left for another occasion. On the other hand music is also often considered to reflect various kinds of motions or gestures. That music often elicits overt movements is obvious and sometimes explicitly intentional, as with dance music. However, one can also experience motion in music without moving oneself. A simple example of that is again dance music: there are striking differences in the experienced motion character of, say, waltz, foxtrot, rock 'n' roll, samba, and tango. The concept of 'swing', often referred to in jazz, is an example of a specific motion character. A march should certainly have a special motion character, and so should a properly performed Viennese waltz. The motion character of Baroque music is different from that of music in the Viennese Classical style. Becking (1928) took the difference in motion character between Mozart's and Beethoven's music as the starting point for his speculative but interesting attempts at using 'accompanying movements' (*Begleitbewegungen*) to catch such differences between different composers; see also Truslit (1938). Clynes (1977, 1980) and Clynes and Walker (1982) proposed to use the sentograph (a device for registering the strength and direction of finger pressure in studies of emotion and motion) as a means of finding out the 'inner pulse' in the music of different composers.

Music as expression of emotions and music as expression of motions are

not competing, but rather complementary ideas. The close psychological connection between emotion and (overt or inner) motion is already suggested by the common origin of those two words (from Latin *movere*) and amply supported by commonplace observation: to jump for joy, fight in anger, sink down in despair, etc. To find a suitable method for studying the motional-emotional aspects of music experience is difficult, however. In my earlier research (Gabrielsson 1973*b*) I mainly used conventional verbal methods and identified motion characters such as 'rapidity', 'forward motion', 'dancing', 'walking', 'rocking', 'swinging', and many others, as well as emotional qualities including 'vital vs. dull', 'excited vs. calm', 'rigid vs. flexible', etc. It is obvious, however, that words are clumsy tools for describing these phenomena in all their continuously varying shades and shapes. The increased interest in non-verbal methods (cf. Scherer and Ekman 1982) suggests other possibilities. Clynes's (1977, 1980; Clynes and Nettheim 1982) theory of emotion and its expression primarily relies on a non-verbal method, registration of finger pressure on a 'sentograph'; it has also been used for studying experienced motion in music. Nielsen (1983) used a similar device, pressing a pair of tongs to reflect experienced tension in two pieces of symphonic music. We can expect further developments of these and other methods to facilitate the study of the ongoing processes in music listening and not remain confined to retrospective verbal reports.

This excursion has been motivated by my belief that many types of timing in music performance aim at eliciting various motional-emotional reactions in the listener. Because of the lack of proper methods to study these questions in enough detail there is, of course, little empirical evidence to rely upon. Nevertheless it seems possible to make some reasonable suggestions. They will be discussed in relation to four categories of timing; tempo, different classes of duration, articulation, and deviations from mechanical regularity.

## Tempo

Tempo is a common and seemingly simple concept in connection with music. The tempo of a piece is usually indicated either by designations such as allegro, moderato, adagio, etc., or more precisely by stating a metronomic value, for example, MM ♩ = 120, that is, 120 crochets per minute. Musicians differ much in their ability to judge the absolute tempo and changes in tempo: while Madsen's (1979) and Wapnick's (1980) subjects were not especially good in this respect, Wagner (1974) found extreme accuracy in a study with the conductor Herbert von Karajan.

However, how fast the music seems to go depends not only on the metronomic tempo but on other factors as well. A piece with many 'short'

tones (say, semiquavers) seems faster than another piece with the same metronomic tempo but containing only 'long' tones (say, crotchets). The number of tones per unit of time is thus important. Sometimes the situation may be ambiguous. For instance, in certain parts of a Nocturne by Chopin the left hand may have a slow pace, while the right hand performs extremely rapid passages (see Fig. 2.1). In this case the motion may be felt as rapid, or slow, or rapid and slow at the same time (a rapid motion within a slow motion), depending on how you direct your attention.

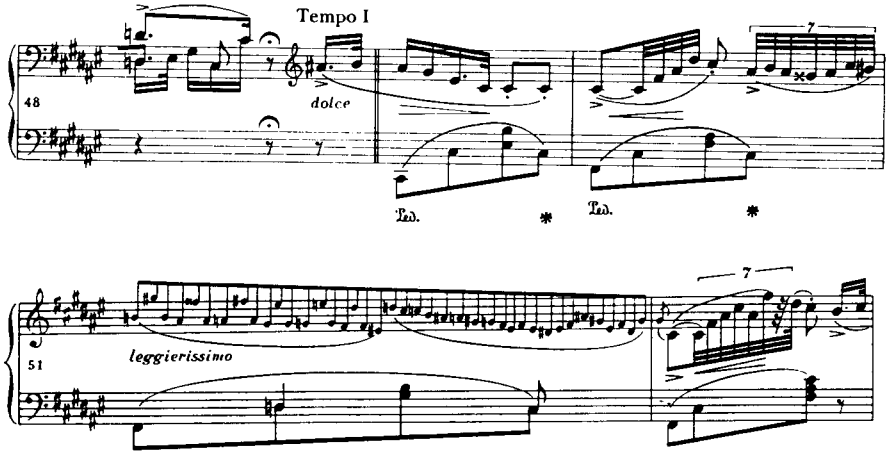



**Fig. 2.1.** Excerpt from 'Nocturne' Op. 15 No. 2 by Frédéric Chopin. Tempo designation: Larghetto. (Reproduced with permission from *Frédéric Chopin, Nocturnes and Polonaises*, Dover Publications, Inc., New York, 1983.)

With perceived tempo we usually mean the *rapidity of the beat* or pulse. This often coincides with the indicated metronomic tempo but not necessarily. Moreover, since the beat often can be alternatively felt at different rates (say, at the crotchet level *or* at the quaver level in a certain piece), the perceived tempo depends on which beat rate is meant. In either case, however, the perceived tempo is only one aspect, although a very important one, of the listener's impression of the speed of the piece. The experienced 'overall rapidity', as well as the rapidity of different parts, also depends on such factors as the number of tones per unit of time (already mentioned) and various melodic and harmonic factors. For instance, a melody with many large leaps may seem faster than a melody which proceeds stepwise (see data on coherence and fission in tone sequences, van Noorden 1975), and a piece with frequent harmonic shifts may appear more rapid than a piece with fewer shifts. There are thus many problems of definition and of influencing factors. No single work covers all these

aspects, but data and instructive discussions on various related points appear in Behne (1972), Motte-Haber (1968), Bengtsson (1974), Longuet-Higgins (1979), Longuet-Higgins and Lee (1982), and Steedman (1977).

Even if an exact metronomic tempo is given, the performer often deviates from that, consciously or unconsciously, depending on his idea of the piece, his present psychic state, acoustical conditions, etc. During the piece there is usually much more variation in tempo than is realized or consciously perceived. The tolerance for such variation is, of course, different in different music, for example, small tolerance in a march but large in a Chopin Nocturne. One common type of variation is where a phrase begins somewhat slowly, accelerates towards the middle, and then retards towards the end. This contributes, together with other means (for example, dynamic increase and decrease), to the phrasing. Numerous examples appear in the early performance studies mentioned above and in Bengtsson and Gabrielsson (1980, 1983) (see also in Fig. 2.8 below). The exact shaping and the extent of the variation differs, of course, with each single piece and each performer. If you measure the performance of a phrase and calculate its average tempo, you may find that this average tempo in fact never appears at any position in the phrase—it is an average of a continuous variation. But as a listener you find it quite natural and adequate. You would, rather, react (negatively) if the tempo *was* perfectly constant.

A detailed study of the tempo variations in a Viennese waltz (from *Die Fledermaus* by Johann Strauss, Jr.) was presented in Bengtsson and Gabrielsson (1983). Although usually not indicated in the score, a genuine performance of a Viennese waltz includes sizeable, sometimes extreme, variations of tempo, the details of which vary with the structure of different pieces. The calculated mean tempo for the 32 bars in this case was about 170 crotchets per minute, but the actual tempo varied between 148 and 196; in the initial bar of the second period it was in fact only 90.

In such a case the mean tempo is of very limited interest. (How boring the waltz sounds when the tempo is kept constant was illustrated by a synthesized example appended to the paper.) From this, and other, examples it was concluded that

> four different meanings of tempo should be distinguished: (a) the abstract *mean tempo*, calculated as the total duration of a music section divided by the number of beats in the same section, (b) the *main tempo*, being the prevailing (and intended) tempo when initial and final retardations as well as more amorphous caesurae are deleted, (c) *local tempi*, maintained only for short periods but perceptibly differing, and (d) *beat rate* . . . for describing minor fluctuations, which may not be perceptible as such.
>
> (Bengtsson and Gabrielsson 1983, p. 50)

How does the tempo affect the listener's impression of the music? If

there is a relatively stable tempo (regular beat), this provides a fundamental temporal framework for the whole piece, a structural basis for whatever else is happening. By appropriate variations of the tempo the performer can contribute to make the structural subdivisions clear to the listener: where a new motive, a new phrase, a new period, etc., begins, and where each ends. The extent of these variations often corresponds to the size or hierarchical level of the respective structural entities, for example, a certain retardation at the end of a phrase, a more pronounced retardation at the end of a period, and most pronounced of all at the very end of the piece.

However, tempo first of all relates to *experienced motion*. Although the original Latin word *tempus* means time, in music tempo designates the experienced rapidity of motion or flow. When there is a stable tempo, the listener feels a regular motion forwards. And when the tempo is varied, one feels how the motion is accelerated or retarded, how it might be brought almost to a stop now and then—but is resumed and carries on until the definite stop at the end. The performer's way of indicating the structure of the piece by variations of tempo is thus made by affecting the speed of the ongoing motion.

When one compares different performances of a piece, the first thing one notices is in fact often differences in tempo, as well as in the ways of varying the tempo. In this are also included other components of perceived rapidity, such as those discussed above in connection with Fig. 2.1 The 'total' impression of rapidity is thus a more or less complex function of various motions. In a completely homophonic texture the situation is rather unproblematic (Fig. 2.2(a)), while it is more intricate in polyphonic music (Fig. 2.2(b)), and in a piece without a regular beat at all (Fig. 2.2(c)). Whatever the case may be, there are usually considerable possibilities for the performer to manipulate various factors influencing the listener's impression of speed and motion, for the whole as well as for different parts.

With regard to the emotional effects several studies (Hevner 1937; Rigg 1940; Motte-Haber 1968; Behne 1972; Wedin 1972; Gabrielsson 1973*b*) indicated that higher tempo/rapidity was associated with judgements like 'vital', 'happy', 'exciting', 'light', 'restless', and others, while slower tempos got roughly opposite descriptions. However, the authors also point out that because of the complex structure of music, it is very difficult to ascribe such effects to differences in tempo alone—they may result from constellations of various factors, such as dynamics, pitch level, articulation, mode, etc. Although the above results certainly seem plausible and can be easily associated with various types of rapid motion, much work remains to make the picture more complete and accurate.

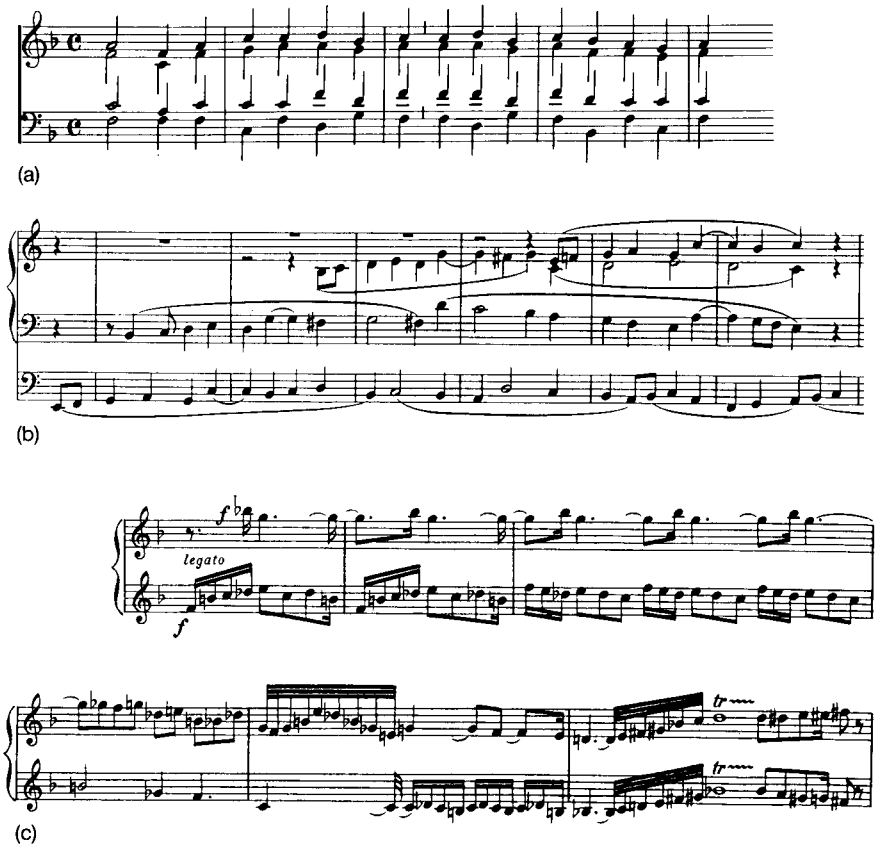



**Fig. 2.2(a).** The beginning of the German chorale 'Lob sei dem allmächtigen Gott'. *(b)* Excerpt from a choral prelude, 'Was mein Gott will, das gscheh allzeit', by Wilhelm Friedemann Bach. (From *Choralvorspiele alter Meister*, Edition Peters, No. 3048. Copyright 1951 by C. F. Peters Corporation, New York. Reproduced by permission of Peters Edition Limited, London.)
*(c)*. The beginning of 'Les Anges' by Olivier Messiaen. (From *Oliver Messiaen, La nativité du seigneur, neuf méditations pour orgue*, Alphonse Leduc, Paris, 1936. Reproduced by permission from Alphonse Leduc & Cie.)

## Different classes of duration

In the score the duration of the tones is indicated by the prescribed tempo and the note value of the respective tone. Note values (semibreve, crotchet, quaver, etc.) are defined in simple ratios to each other, such as 2:1 (for example, semibreve to crotchet) or 3:1 (dotted semibreve to crotchet). Of course, it is realized that there is a certain tolerance around

these ratios in actual performance. However, it is rarely understood how extreme and varying the deviations really are, as found in empirical investigations of performance (see p.41 onwards).

Disregarding this for the moment, it is self-evident that the specific sequence of note values given in the score provides a primary structural characteristic of the piece in question. But in the same sense that the tempo reflects experienced motion, the various tone sequences also have *motional-emotional properties*. Due to the practically infinite number of possible sequences, these qualities are almost innumerable. They provide an endless variety of means for the composer and performer to shape the ongoing flow in the music (also including the use of 'non-tones', rests). As a few, extremely simple, examples consider the short patterns in Fig. 2.3. The motional characters of these certainly vary and may here be loosely described by expressions such as 'even flow', 'rocking', more or less 'jerky', 'galloping', and the like. The last, 'galloping' pattern has a very pronounced 'moving forwards' quality, as used, for instance, in the overture to *Guillaume Tell* by Rossini.

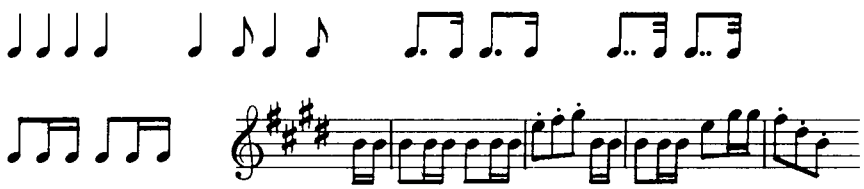


**Fig. 2.3.** Five short patterns illustrating different motion characters. The music example is taken from the overture to *Guillaume Tell* by Rossini, tempo designation Allegro vivace.

However, there have been few empirical attempts at elucidating these fundamental questions. Suggestions may be found in Motte-Haber (1968) and Gabrielsson (1973*a*, *b*). Clynes and Walker (1982) used a sentograph to study the responses to two-tone patterns, which were varied in the duration and amplitude of the tones. The multitude of different responses obtained even for these simple cases indicates something of the enormous possibilities available in composing and performing music to elicit various motional and emotional reactions in the listener—even richer when including variations of pitch, intensity, timbre, etc. as are found in most music.

In a certain contrast to the statements above, Fraisse (1956, 1982) claims that in a piece of music there are essentially only two classes of durations: 'long times' and 'short times' (*temps longs*, *temps courts*). In his excellent monograph *Les structures rythmiques* (1956), Fraisse described the results of many experiments on motor production or reproduction of rhythm patterns. His subjects' tappings of these patterns indicated the existence of

two separate duration regions, the 'long times' and the 'short times', in an approximate ratio of 2:1 or higher. The absolute duration of these classes varied in different contexts, but generally 'short times' were associated with perception of 'grouping' and 'long times' with perception of 'duration'. Rhythm perception was seen as a play between those two qualities, collection and duration.

Fraisse also finds the two classes of time in the 'long' and 'short' syllables in prosody and in the dashes and dots of the Morse alphabet. With regard to music he studied excerpts from 15 pieces of piano music and found that in each piece two note values, in a ratio of 2:1 or 3:1, were dominant. Of course, there were other note values as well, but they were seen as somehow 'occasional'. The shorter note value corresponded to a duration of 135–410 ms in different pieces, which agrees well with the 'short times' in the above-mentioned experiments; the longer note value corresponded to 270–860 ms ('long times'). Fraisse (1981) further argued that the limitation to two durations relates to our inability to 'discriminate in an absolute way more than two durations in the range of perceived durations (from 10 to 200 centiseconds) (1981, p. 220).

However interesting this may sound, the theory seems simplistic from a musical point of view. There are certainly many pieces dominated by only two note values, but to make a generalization from only 15 pieces is unwarranted. To provide only one counter-example, consider another piano piece, the beginning of Beethoven's Piano Sonata Op. 14 No. 1 (Fig. 2.4). In the exposition of the first movement at least four note values (semibreve, crotchet, quaver, and semiquaver) play an important role in the themes and in the elaborations and transitions between them. A listener should have no difficulties in recognizing them as four different categories.

To reduce timing in music to be a matter of essentially only two classes of duration is not very convincing. And no empirical evidence is given to support the statement that we are unable to discriminate in an absolute way more than two durations within 10-200 cs.

## Articulation

Legato and staccato are contrasting forms of articulation. In legato (or even legatissimo) the sounding tone goes on all the way until the beginning of the next tone with no perceptible break between the tones. In staccato (and even more in staccatissimo) the sounding part of the tone is very short and is followed by silence until the onset of the following tone. The silent part is often considerably longer than the sounding part. Between legato and staccato there are various intermediate levels such as non legato,

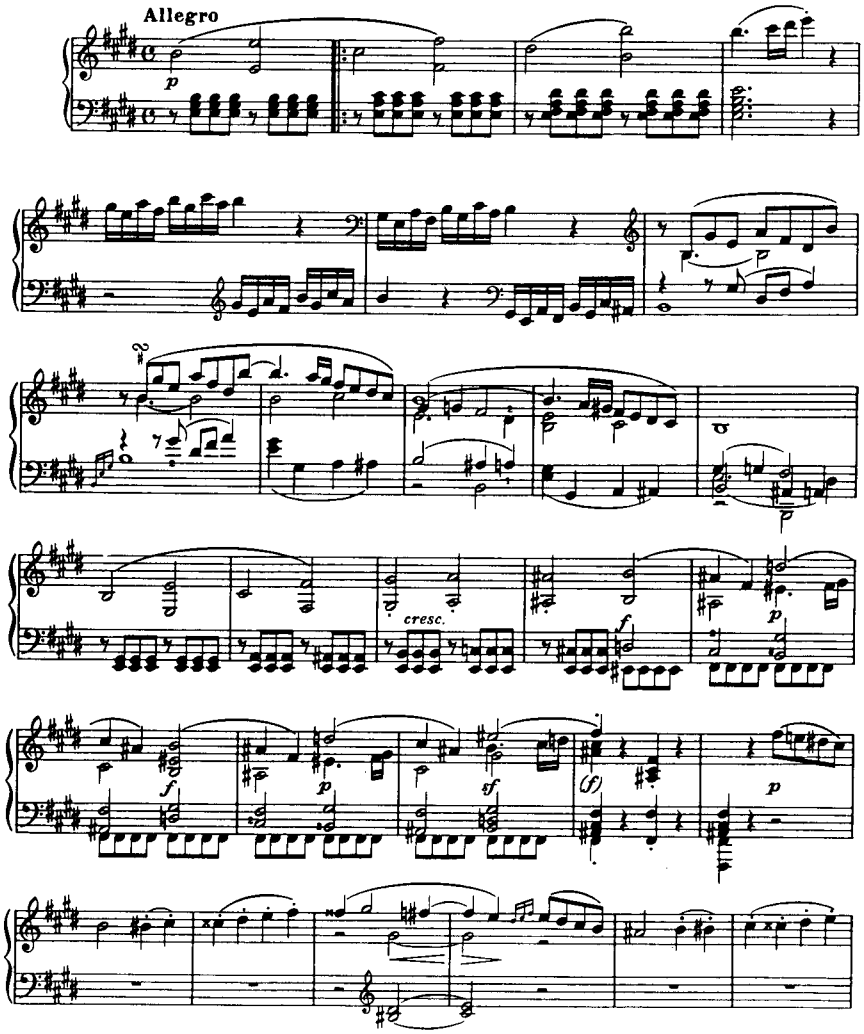



**Fig. 2.4.** The beginning of Beethoven's Piano Sonata Op. 14 No. 1. (From *Beethoven, Klaviersonaten, Band I*, G. Henle-Verlag, München-Duisburg, 1952. Reproduced by permission from G. Henle Verlag.)

portato, and marcato; for string instruments there are also other terms such as *detaché* and martellato.

In the score the type of articulation is indicated by the above labels and/or by various graphic signs at the notes, such as points (staccato), dashes, wedges, bows, and various combinations of these. The silent part of the tone may be written as a rest, for example, ♩ (staccato) might instead be written as 𝅘𝅥𝅯 𝄾 𝄾 .However, a frequent use of rests makes the

score hard to read and is usually avoided. When there is no indication at all of the articulation, this may be interpreted as legato, or that the articulation should be as is usual for this kind of music, for this type of musical ensemble etc. Sometimes the title of the piece or other designations suggest the proper articulation. Composers differ very much with regard to how accurately they prescribe various articulations.

Whatever the case may be, the articulation is achieved by timing in performance—timing to get the intended or proper proportion between the sounding and non-sounding parts of the tone(s). From an acoustical point of view the situation is very complex, since the performer's control of this timing is affected by the properties of different instruments and by the acoustics of the room. For instance, articulation on the violin is different from articulation on the organ, and playing in a small room is different from playing in a cathedral. Moreover, a perceptual distinction between the sounding and non-sounding part may not reflect acoustic reality: the tone may be sounding all the time, first loud and then soft (for example, in reverberation)— however, so soft that it is masked or otherwise perceptually silent.

The importance of the articulation is well known by performers and is given much attention in musical training. However, there have been few empirical investigations of the performance of articulations in various situations. Even if careful physical registrations are made of the performance, they have to be supplemented by judgements from listening in order to clarify the relations to the perceived articulation. One way of approaching these problems is by means of synthesized sound sequences, which are judged by listeners. An example is given in Fig. 2.5, which contains four bars of a usual form of waltz accompaniment (see the notation at the top). Eight versions were synthesized to represent (almost) legato performance in all three beats of the bar (version 1), staccato performance in all three beats (version 8), and various combinations of legato and staccato in versions 2–7 (for example, legato in the first and second beats but staccato in the third as in version 2). The example is taken from Bengtsson and Gabrielsson (1983), and the sound example appears on an appended phonograph record. Lacking this, you may perform or try to imagine the eight versions yourself.

The perceptual effects of these manipulations are striking. The basic tripartite structure of the bar is felt the same way in all cases, but there is a tremendous difference in the motional-emotional character of the eight versions. Some of them are completely 'impossible' and sound ridiculous (nos. 1 and 5), while the others represent more or less plausible alternatives, each of them with a specific character.

This example was in fact not primarily made to discuss waltz accompaniment, but to demonstrate the effects of different articulations in a systematic way. Considering the possibilities available in real music

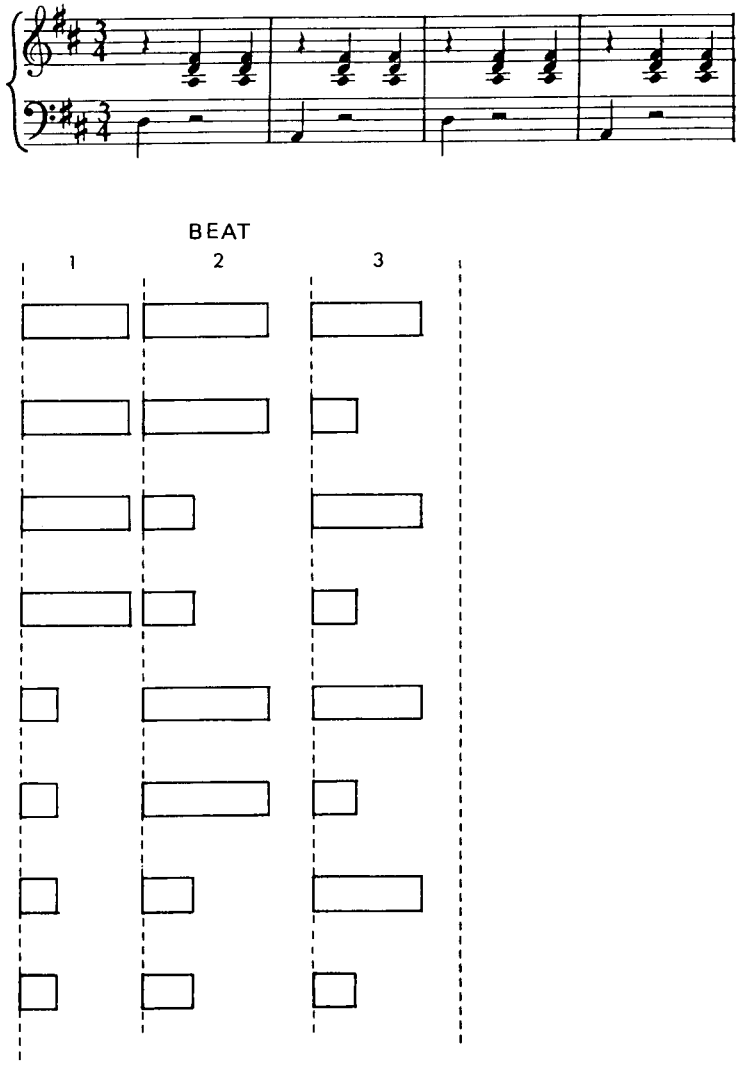



**Fig. 2.5.** Different articulations of a common accompaniment to waltzes, see further in text. (From Bengtsson and Gabrielsson 1983, p. 45. Reproduced by permission from the The Royal Swedish Academy of Music, Stockholm.)

performance—using even more legato or staccato, or intermediate levels such as non legato or portato (and still finer gradations for which we lack words), and combining them with dynamic means, etc.—we begin to realize the scope of the articulation repertoire available to the performer, and how he may use it to generate so many different and subtly varied

experiences, which we feel so well but which elude our attempts at verbalization.

## Deviations from mechanical performance

A spontaneous response to music boxes, rhythm machines, and the like is that they sound mechanical, dead, and boring. The main reason for that is the mechanical regularity in their 'performance': absolutely constant tempo, and perfect ratios (1:1, 2:1, 3:1, etc.) between successive sound events. It is generally known that human performance is not like that, and various types of deviations from mechanical regularity were described in the early investigations by Sears (1902), Hartmann (1932), and the research group around C. E. Seashore (1938); these are summarized in Gabrielsson (1985).

Extensive studies of music performance have been made in Uppsala, Sweden, in collaboration between musicologists and psychologists (Bengtsson 1974; Gabrielsson 1974, 1982*a*, 1982*b*, 1985, 1986; Bengtsson and Gabrielsson 1977, 1980, 1983; Gabrielsson *et al.* 1983; Gabrielsson and Johnson 1985). The general purpose has been to investigate timing in performance, especially with regard to deviations from mechanical regularity in different types of music and with different performers. In most cases the performances were especially for this research project, but performances stored on phonograph records or magnetic tapes are used as well. They are analysed by means of special electronic equipment (Tove *et al.* 1966), which simultaneously displays the variations of fundamental frequency ('pitch') and total amplitude (intensity). By adequate filtering and calibrations it is possible to get very accurate representations of monophonic performances and in many cases also of polyphonic music. The durations may be measured with an accuracy of $\pm 5$ ms. An example of a registration is given in Fig. 2.6.

More recently the equipment has been supplemented by a computerized synthesizer and a sampling system. This allows digital recording and storing of performances, live or recorded earlier. They can be studied in extreme detail with a time resolution of 20 μs and can be manipulated in various ways; for example, any parts(s) of the performance can be extracted for detailed listening and analysis of transitions between tones, for spectral analysis, etc. (see Fig. 2.7 for some illustrations).

Despite the high accuracy there are still problems in certain cases, especially regarding physical and perceptual tone onsets as well as decays. While Wagner (1974) and Shaffer (1980, 1981) make registrations of piano performance directly at the sound source, our techniques mainly rely on registrations by microphone, for any instrument, placed at a distance in the room. The acoustical events thus registered are often complex. Tones

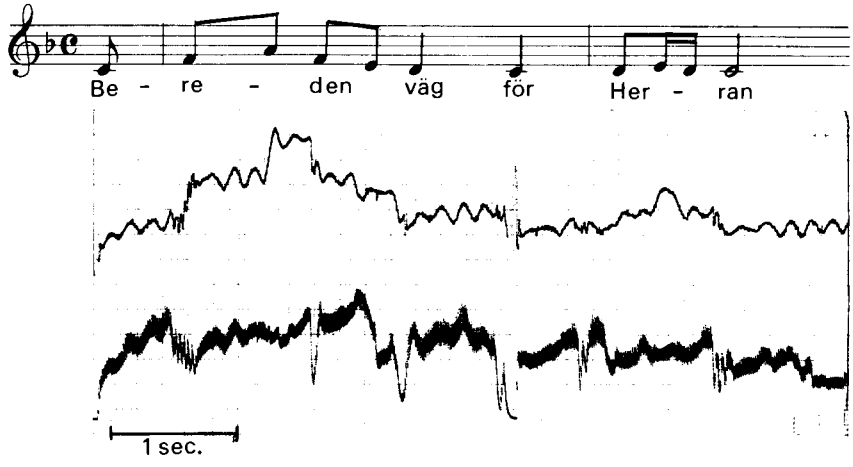



**Fig. 2.6.** Registration of a professional singer performing in the first phrase of a Swedish folk chorale. The upper curve shows the fundamental frequency, the lower curve the intensity (calibration values not given here). Note the continuous variation in both curves, including vibrato and various glidings. The notes in the notation have been put to positions approximately corresponding to the onsets of the respective tones as seen in the registrations below. For further description, see Gabrielsson and Johnson (1985).

appearing after another in the score overlap each other in the acoustical reality and make us wonder exactly when, and because of which specific feature in the ongoing stimulus, the listener detects the onset of the next tone. This is one reason why performance studies have to be supplemented by studies of the listener's experience. Other, more general, reasons for that are discussed later.

On the other hand the deviations from strict regularity in performance are usually so obvious that they are detected by any reasonable technique in spite of the above-mentioned difficulties. The results described below come from monophonic performances on the flute, the clarinet, and the piano, and from orchestral performances. The music is Western art music, popular music, and folk music; many of the pieces are of Swedish origin. The statistical analysis of the data obtained for each performance generally proceeds as follows:

(1) Calculate the total duration of the performed piece (however, often omitting the very last tone, since its duration may be difficult to measure accurately).

(2) Transform the duration of each tone (from its onset to the onset of the next tone) to per mille of the total duration.

(3) Take the difference between this per mille value and the per mille value the tone would have had in a completely mechanical performance.

The last step thus gives the difference (deviation) for each tone, between the real performance and a mechanical performance. The deviation values are tabulated and/or displayed graphically as in Fig. 2.8 below. Then follows:

(4) Perform adequate statistical operations to get measures of central

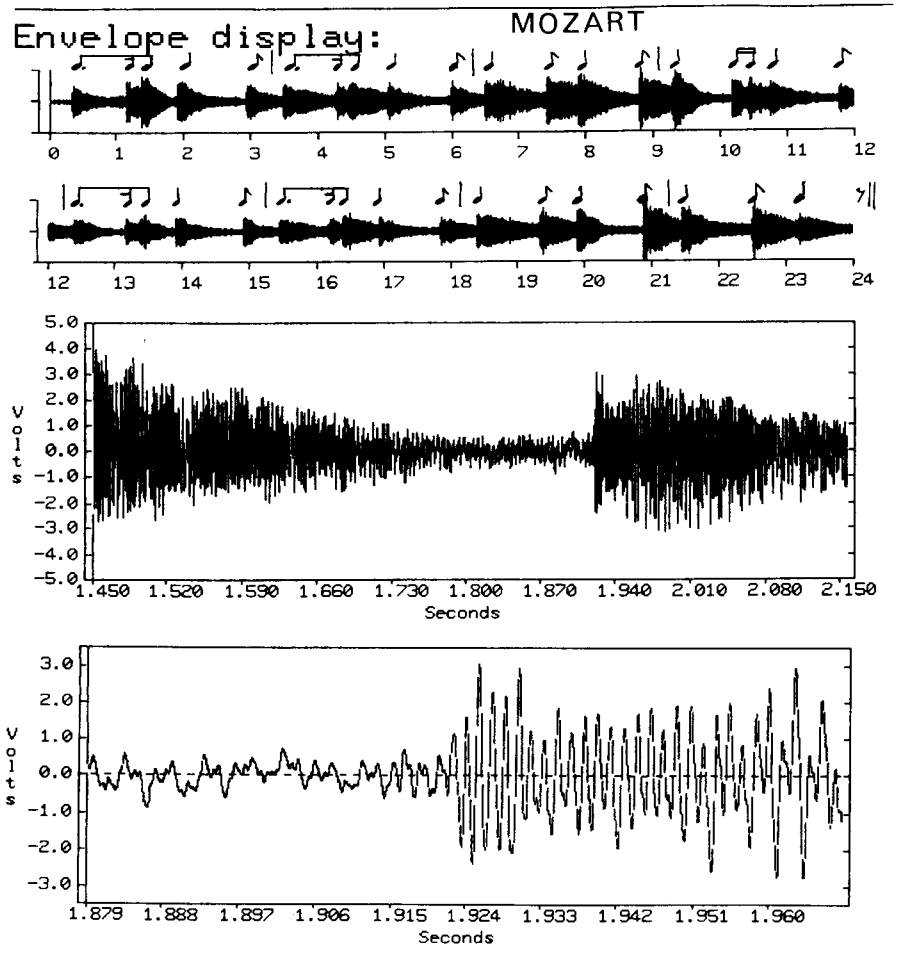



**Fig. 2.7.** Examples of registrations obtained from a sampling system (Synclavier II) for the pianist Ingrid Haebler's performance of the initial eight bars of Mozart's Piano Sonata in A major (K. 331; see the notation in Fig. 2.8 below). The *upper* figure shows the amplitude envelope; the onsets of the chords are clearly seen (the rhythmic notation is added here to facilitate the understanding of the figure). The numbers represent time in seconds. The *middle* figure shows the decay of the third chord and the onset of the fourth. The *lower* figure displays this in more detail: the very end of the third chord and the beginning of the fourth.

tendency and dispersion for various entities, calculate proportion values for neighbouring tones, etc., within the single performer and over several performers.

(5) Perform factor analysis on the deviation values from different performers or versions to see if there are some characteristically different ways of performing the piece.

This is only a very condensed description, and many other calculations are made, depending on the specific results for each piece or performer (see Bengtsson and Gabrielsson 1980 for details). However, the principle is to find out the deviations from mechanical performance, and to see whether they are the same for all performers, or if there are different types of deviations.

As expected, frequent deviations appear in all investigated cases. However, they are different in different pieces of music and also often between performers. A detailed description can thus only be made in connection with a certain piece and performer, for which the reader is referred to the original papers. Here we will only briefly describe some fairly common tendencies, proceeding from smaller to larger entities.

Sequences of equally long tones (according to the notation in the score), for example ♩ ♩ or ♫ , often show a long-short (LS) or short-long (SL) relation between the two tones; that is, the first tone is longer than the second (LS), or the converse is true (SL). Which alternative appears depends on the context as well as, sometimes, on the performer. In a sequence like ♬ the pattern is often LSSL, that is, the first and the last quavers are lengthened, the two middle ones shortened; but LSLS or SLSL patterns may also occur. In a pattern such as ♩ ♫ the quavers are often performed SL, while the relation at the crotchet level may be LS (the crotchet is longer than the two quavers together) or SL. The corresponding phenomena appear in the analogous ♬ sequence. In patterns such as 𝅗𝅥 ♩ in 3/4 metre and ♩ ♪ in 6/8 metre there is usually a much lower proportion than the mechanical 2:1 ratio between the two tones; it is often between 1.7:1 and 1.9:1 (cf. the example in Fig. 2.8) and may approach 1.5:1 in certain cases. But there are exceptions: for instance, in a march notated in 6/8 metre the ♩ ♪ sequence was performed close to the 2:1 ratio. This suggests that the general motion character of a piece is a decisive factor for the performance of various sequences within it.

The performance of dotted note values such as in ♩. ♪ or ♩. 𝅘𝅥𝅯 varies very much depending on context and performer. The mechanical ratio is 3:1, but the real ratio may be higher than ('sharp dotting'), lower than ('soft dotting'), or close to 3:1. It seems obvious that a ratio higher than 3:1 gives a more marked and 'jerky' character, while a ratio lower than 3:1 tends to sound 'softer'. If there are several sequences with the same type of dotting , it may happen that the dotting is made progressively 'softer' during the course of the phrase. A double dotting like ♩.. 𝅘𝅥𝅯 has a

mechanical ratio of 7:1 but may deviate very much from that in actual performance, sometimes down to about 4:1 or even lower.

At the beat level a well-known example of deviations is the accompaniment in Viennese waltzes. The first beat is considerably shortened, and the second beat thus comes 'too early', which (together with other factors) gives an unmistakable motion character to these waltzes. A synthesized sound example, which demonstrates the perceptual effects of the second beat coming earlier and earlier, is given in Bengtsson and Gabrielsson (1983). There are numerous other examples of deviations at the beat level, more or less specific for each piece; we found, for instance, that in the popular tune 'Oh my darling Clementine' the third beat was generally lengthened.

At levels corresponding to half-measures, measures, and to still larger entities there are deviations of various types, which have to be described in connection with the performed piece. Some results were mentioned under 'Tempo' above. In order to have at least one concrete example of several phenomena mentioned above, we may look at Fig. 2.8. It shows the performance of the first eight bars of Mozart's Piano Sonata in A Major

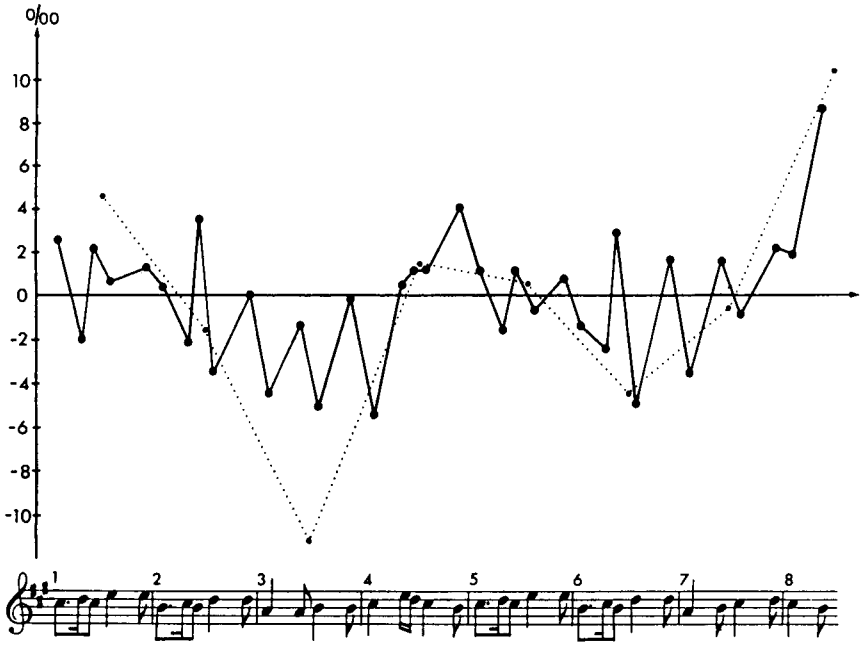



**Fig. 2.8.** Deviations from mechanical performance in the pianist Ingrid Haebler's performance of the beginning of Mozart's Piano Sonata in A Major (K. 331). The deviations are expressed in per mille of the total duration (until the last tone); one per mille corresponds to approximately 23 ms. See text for further explanations.

(K. 331) by the pianist Ingrid Haebler, renowned interpreter of Mozart. The horizontal line marked 0 represents a mechanical performance. The deviation for each tone is given by the black circles joined by the whole-drawn profile. A position above the zero line means that the corresponding tone is lengthened, a position below this line says that the tone is shortened, in relation to mechanical performance. The profile for the ♩. ♪ ♩ in the beginning of bars nos. 1, 2, 5, and 6 is V-shaped, that is, the dotted quaver and the last quaver are (relatively) lengthened, the semiquaver shortened. The sequences of ♩ ♪ are performed with the crotchet shortened or slightly lengthened; the quaver is mostly lengthened or less shortened than the crotchet. The ratio between the crotchet and the quaver is thus lower than 2:1. In average over all cases it is 1.76:1, but it varies from 1.92:1 down to 1.50:1. The dotted profile shows the deviations at the measure level. It is seen that the relation between the four bars in each half-period is LSSL, that is, the first and the fourth bars are longer, the second and third shorter. In other words, the tempo varies: somewhat slow start, accelerating in the second bar to reach the fastest tempo (= largest deviation downwards) in the third bar, then retarding towards the end of the half-period in the fourth bar. A similar profile is seen in the second half-period (the marking for the eighth bar refers to the first half of this measure). Many further observations can be made, but this may suffice to illustrate the general idea. Of course, the amount of data and interpretation grows considerably for analysis of longer examples.

## Relations to the listener's experience

Results from performance studies must be interpreted with caution for several reasons. They may be more or less specific for the individual performer. The aforementioned factor analyses usually indicated two or three different ways of playing even very short pieces, and the variation among different performers' interpretations of a musical work is well known from general musical experience. There may be intra-individual variations as well. The musician changes his performance from time to time, sometimes deliberately in order to present a 'new' interpretation, sometimes unconsciously and for obscure reasons, still other times for adapting to the acoustical conditions (for example, changing the tempo).

To find truly general results in performance data is therefore difficult. The generalities should rather be sought in the *relations between performance and* (the listener's) *experience*. The performance is the overt manifestation of the composer's and/or the performer's idea, and it is the stimulus for the listener's experience of the music. However, the performer is also a listener—he continuously checks that there is a correspondence between his intentions and the music he produces. He strives to convey his

experience of the music as truly as possible to his listeners—so that they should experience the music as he does. The performance is a link joining the performer's experience and the listener's experience.

Of course, various individual characteristics (life history, attitudes towards music, personality traits etc.) affect the experience and make it more or less different for the performer and any individual listener. But there is usually also much in common, and we may assume that this communality is based on 'natural' relations between performance and experience—in other words, that *certain properties of the performance elicit certain qualities in the experience.* Long ago H. Seashore (1937, p. 118) expressed this opinion in the following words:

> The psychophysical relations between the performer and the listener must be worked out; the data presented will contribute to such studies and also will depend for their final interpretation upon such investigations.

However, practically nothing has been done in the direction suggested by H. Seashore. We still discuss the relation between performance and experience in very general terms, as when C. E. Seashore stated that deviation from the regular is a general principle of art, or that 'deviation from the exact is . . . the medium for the creation of the beautiful—for the conveying of emotion' (quoted from H. Seashore 1937, p. 155). My repeated suggestions that various performance characteristics may affect structural and motional-emotional aspects of the experience is another example of such general propositions.

A basic difficulty is that we have no clear identification of the numerous dimensions which constitute our experience of music, nor do we have adequate methods to study them in a reliable way. It is my guess that non-verbal methods will play a greater role in future work on these questions. To study the relations between performance and experience we furthermore need a kind of iterative interplay between analyses of performed music and systematically varied syntheses of 'music'. A beginning of such a programme, with special reference to performance and experience of musical rhythm, is sketched in Gabrielsson (1985).

Meanwhile I again suggest that the frequent deviations from mechanical performance have structural as well as motional-emotional consequences. The experienced structure is certainly affected by variations of tempo, various other deviations, articulation, etc. (see Sloboda 1983). At the same time, however, these phenomena have obvious motional and emotional qualities: to accelerate or retard the experienced flow of the music, to provide a march or a Viennese waltz their proper motion characters, to make a dotting feel 'jerky' or 'soft', etc. For instance, the 'early second beat' in the Viennese waltz does not essentially change the basic tripartite structure of the bar—but it does change the experienced motion character.

Listening to synthesized, systematically varied sound sequences reveals that the ratio between the durations of adjacent tones may vary a lot without obscuring their structural relations (experiments to be reported). This indicates, among other things, that rhythm perception is categorical, and that considerable variation is available within the categories to affect various motional-emotional qualities. Similar results are obtained for synthesized melodies. Some of the variation in performance data may be accounted for in terms of such effects.

## Concluding remarks

In this chapter timing has been treated in isolation from other means used by the performer, such as dynamics (in a wide sense), intonation, and spectral variations. Moreover, different aspects of timing (tempo, classes of duration, articulation, deviations from mechanical regularity) have been discussed separately from each other. In reality, however, all these means occur together in a complex, ever-changing interaction, depending on the instrument, the performer, and the situation. It is for the purpose of understanding the ingredients of this interaction that the present analytic approach may be justified. I hope it has also made clear the importance of studying the *sounding* music and its relations to the listener's experience rather than staying content with the simplistic picture given in musical notation. Studying the relations between performance and experience means coming very close to the musical reality, and is a fascinating challenge to all of us.

## Acknowledgements

The author wants to express his gratitude to Ingmar Bengtsson, Barbro Gabrielsson, Dan Malmström, Peter Reinholdsson, and Bo von Schéele for invaluable collaboration in discussions and experiments. The investigations were supported by The Swedish Council for Research in the Humanities and Social Sciences, The Bank of Sweden Tercentenary Foundation, Knut and Alice Wallenberg's Foundation, and Uppsala University.

## References

Becking, G. (1928). *Der musikalische Rhythmus als Erkenntnisquelle.* Benno Filser, Augsburg.

Behne, K. E. (1972). *Der Einfluss des Tempos auf die Beurteilung von Musik.* Arno Volk, Köln.

Bengtsson, I. (1974). Empirische Rhythmusforschung in Uppsala. *Hamburger Jahrbuch für Musikwissenschaft* **1**, 195–220.

Bengtsson, I. and Gabrielsson, A. (1977). Rhythm research in Uppsala. In *Music, room, and acoustics*; Publications issued by the Royal Swedish Academy of Music, No. 17, pp. 19–56. Stockholm.

Bengtsson, I. and Gabrielsson, A. (1980). Methods for analyzing performance of musical rhythm. *Scandinavian Journal of Psychology* **21**, 257–68.

Bengtsson, I. and Gabrielsson, A. (1983). Analysis and synthesis of musical rhythm. In *Studies of music performance*, Publications issued by the Royal Swedish Academy of Music, No. 39 (ed. J. Sundberg), pp. 27–60. Stockholm.

Clynes, M. (1977). *Sentics, the touch of emotions.* Anchor Press/Doubleday, New York.

Clynes, M. (1980). The communication of emotion: theory of sentics. In *Emotion: theory, research, and experience* (ed. R. Plutchik and H. Kellerman), Vol. 1, *Theories of emotion.* Academic Press, New York.

Clynes, M. and Nettheim, N. (1982). The living quality of music: neurobiologic basis of communicating feeling. In *Music, mind, and brain: the neuropsychology of music* (ed. M. Clynes). Plenum, New York.

Clynes, M. and Walker, J. (1982). Neurobiologic functions of rhythm, time, and pulse in music. In *Music, mind, and brain: the neuropsychology of music* (ed. M. Clynes). Plenum, New York.

Fraisse, P. (1956). *Les structures rythmiques.* Editions Universitaires, Louvain.

Fraisse, P. (1981). Multisensory aspects of rhythm. In *Intersensory perception and sensory integration* (ed. R. D. Walk and H. L. Pick), pp. 217–48. Plenum, New York.

Fraisse, P. (1982). Rhythm and tempo. In *The psychology of music* (ed. D. Deutsch). Academic Press, New York.

Gabrielsson, A. (1973*a*). Similarity ratings and dimension analyses of auditory rhythm patterns. I and II. *Scandinavian Journal of Psychology* **14**, 138–60, 161–76.

Gabrielsson, A (1973*b*). Adjective ratings and dimension analyses of auditory rhythm patterns. *Scandinavian Journal of Psychology* **14**, 244–60.

Gabrielsson, A. (1973*c*). Studies in rhythm. *Acta Universitatis Upsaliensis: abstract of Uppsala dissertations from the faculty of social sciences*, No. 7.

Gabrielsson, A. (1974). Performance of rhythm patterns. *Scandinavian Journal of Psychology* **15**, 63–72.

Gabrielsson, A. (1982*a*). Perception and performance of musical rhythm. In *Music, mind and brain: the neuropsychology of music* (ed. M. Clynes). Plenum, New York.

Gabrielsson, A. (1982*b*). Performance and training in musical rhythm. *Psychology of Music, special issue: Proceedings of the Ninth International Seminar on Research in Music Education* , 42–6.

Gabrielsson, A. (1985). Interplay between analysis and synthesis in studies of music performance and music experience. *Music Perception* **3**, 59–86.

Gabrielsson, A. (1986). Rhythm in music. In *Rhythm in psychological, linguistic and musical processes* (ed. J. R. Evans and M. Clynes). Charles C. Thomas, Springfield, Ill.

Gabrielsson, A. and Johnson, A. (1985). Melodic motion in different vocal styles.

In *Analytica, studies in the description and analysis of music in honour of Ingmar Bengtsson*, Publications issued by the Royal Swedish Academy of Music, No. 47. Stockholm.

Gabrielsson, A., Bengtsson, I., and Gabrielsson, B. (1983). Performance of musical rhythm in 3/4 and 6/8 meter. *Scandinavian Journal of Psychology* **24**, 193–213.

Hartmann, A. (1932). Untersuchungen über metrisches Verhalten in musikalischen Interpretationsvarianten. *Archiv für die gesamte Psychologie* **84**, 103–92.

Henderson, M. T. (1937). Rhythmic organization in artistic piano performance. In *Objective analysis of musical performance: University of Iowa studies in the psychology of music* (ed. C. E. Seashore), Vol. 4, pp. 281–305.

Hevner, K. (1937). The affective value of pitch and tempo in music. *American Journal of Psychology* **49**, 621–30.

Longuet-Higgins, H. C. (1979). The perception of music. *Proceedings of the Royal Society of London* **B205**, 307–22.

Longuet-Higgins, N. C. and Lee, C. S. (1982). The perception of musical rhythms. *Perception* **11**, 115–28.

Madsen, C. K. (1979). Modulated beat discrimination among musicians and nonmusicians. *Journal of Research in Music Education* **27**, 57–67.

Motte-Haber, H. de la (1968). *Ein Beitrag zur Klassifikation musikalischer Rhythmen.* Arno Volk, Köln.

Nielsen, F. V. (1983). *Oplevelse av musikalsk spaending* (The experience of musical tension) (with summary in English). Akademisk Forlag, Copenhagen.

Noorden, L. P. A. S. van (1975). *Temporal coherence in the perception of tone sequences.* Institute for Perception Research, Eindhoven.

Rasch, R. A. (1978). The perception of simultaneous notes such as in polyphonic music. *Acustica* **40**, 21–33.

Rasch. R. A. (1979). Synchronization in performed ensemble music. *Acustica* **43**, 121–31.

Rasch, R. A. (1981). *Aspects of the perception and performance of polyphonic music.* Drukkerij Elinkwijk BV, Utrecht.

Rigg, M. G. (1940). Speed as a determiner of musical mood. *Journal of Experimental Psychology* **27**, 566–71.

Risset, J. C. and Wessel, D. L. (1982). Exploration of timbre by analysis and synthesis. In *The psychology of music* (ed. D. Deutsch). Academic Press, New York.

Scherer, K. R. and Ekman, P. (1982). *Handbook of methods in nonverbal behavior research.* Cambridge University Press, Cambridge.

Sears, C. H. (1902). A contribution to the psychology of rhythm. *American Journal of Psychology* **13**, 28–61.

Seashore, C. E. (1938). *Psychology of music.* McGraw-Hill, New York.

Seashore, H. G. (1937). An objective analysis of artistic singing. In *Objective analysis of musical performance: University of Iowa studies in the psychology of music* (ed. C. E. Seashore), Vol. 4, pp. 12–157.

Shaffer, L. H. (1980). Analysing piano performance: a study of concert pianists. In *Tutorials in motor behavior* (ed. G. E. Stelmach and J. Requin). North-Holland, Amsterdam.

Shaffer, L. H. (1981). Performances of Chopin, Bach and Bartók: studies in motor programming. *Cognitive Psychology* **13**, 326–76.

Skinner, L. and Seashore, C. E. (1937). A musical pattern score of the first movement of the Beethoven sonata, opus 27, no. 2. In *Objective analysis of musical performance: University of Iowa studies in the psychology of music* (ed. C. E. Seashore), Vol. 4, pp. 263–80.

Sloboda, J. A. (1983). The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology* **A35**, 377–96.

Small, A. M. (1937). An objective analysis of artistic violin performance. In *Objective analysis of musical performance: University of Iowa studies in the psychology of music* (ed. C. E. Seashore), Vol. 4, pp. 172–231.

Steedman, M. J. (1977). The perception of musical rhythm and metre. *Perception* **6**, 555–69.

Tove, P. A., Norman, B., Isaksson, L., and Czekajewski, J. (1966). Direct-recording frequency and amplitude meter for analysis of musical and other sonic waveforms. *Journal of the Acoustical Society of America* **39**, 362–71.

Truslit, A. (1938). *Gestaltung und Bewegung in der Musik.* Chr. Friedrich Vieweg, Berlin-Lichterfelde.

Vernon, L. N. (1937). Synchronization of chords in artistic piano music. In *Objective analysis of musical performance: University of Iowa studies in the psychology of music* (ed. C. E. Seashore), Vol. 4, pp. 306–45.

Wagner, C. (1971). The influence of the tempo of playing on the rhythmic structure studied at pianist's playing scales. In *Medicine and sport*, Vol. 6, *Biomechanics II*, pp. 129–32. Karger, Basel.

Wagner, C. (1974). Experimentelle Untersuchungen über das Tempo. *Österreichische Musikzeitschrift* **29**, 589–604.

Wapnik, J. (1980). The perception of musical and metronomic tempo change in musicians. *Psychology of Music* **8**, 3–12.

Wedin, L. (1972) A multidimensional study of perceptual-emotional qualities in music. *Scandinavian Journal of Psychology* **13**, 241–57.

# 3

# Computer synthesis of music performance

JOHAN SUNDBERG

## Introduction

### Background of the project

This chapter presents a research project which was started some years ago. Its goal has been to analyse music performance by synthesizing it using a digital computer.

The origin of this project was apparently chance, as is the case with so many other research projects. Three crucial circumstances appeared in combination at the same time and in the same place.

One circumstance was the construction of the singing synthesizer MUSSE (see Larsson 1977), which offered synthesis of sung vowels. These synthetic vowels sounded very natural, as the interested reader can hear from sound illustrations published in Sundberg (1978). When this synthesizer could be controlled from a computer, it could be programmed with melodies, which were subsequently performed by the synthesizer. Musically experienced listeners typically noted that these performances sounded like a singer who was gifted with a good voice but who unfortunately lacked the desire to communicate anything in particular with his singing. This raised the question as to exactly what it is that the performer adds to the music score.

A second circumstance was the author's contact with Lars Frydén, a musician with an extensive background as a violin soloist and quartet player, as well as an accomplished teacher of violin playing at higher levels. His long-term musical experience had resulted in a number of hypotheses regarding some basic principles for the conversion of note signs into sounding music and also in a desire to test these hypotheses.

A third circumstance was the fact that at the laboratory, two of the author's colleagues, Rolf Carlson and Björn Granström, had devoted much work and ingenuity to the conversion of text to sounding speech (Carlson and Granström 1974). This had resulted in computer facilities offering means to synthesize by rules. These rules could concern not only

pronunciation but also pitch contour and timing. It appeared that the task of converting text signs to speech sequences was in important respects similar to the task of converting note signs to tone sequences.

These three circumstances, in no way related to each other, turned out to form a good platform for attempts to improve the musical quality of computer-generated performances of music. Several reports have been published on this project over the years (Sundberg *et al* 1983*a*, *b*; Sundberg and Frydén 1985). This chapter aims both at presenting the strategy, including its advantages and limitations, and at mediating some experiences that have been gained in working with the project.

## Topic

It is a well-known fact that traditional music sounds musically unacceptable when performed by a digital computer in an exact agreement with what is nominally given in the music score. This implies that the discrepancies between the sound sequences shown in the music score and those really generated in music performance constitute an essential part of music communication.

Evidently, these discrepancies cannot be random; for instance, it takes considerable training to learn playing, and, also, there is mostly a general agreement among competent listeners as to what is an unacceptable performance of a piece of music.

The above suggests that the discrepancies between notation and sound must be in some sense meaningful, that is, they carry information which the listener needs in order to enjoy listening to the performance. From the point of view of scientific research this is an interesting fact. What is this information? How is the acoustic code chosen which is used by the musician for conveying it? Is this code used in music only or is part of it borrowed from extramusical communication? These question touch on basic aspects of music communication. It seems that important contributions to the answers could be found by analysing music performance.

Nevertheless, music performance has attracted comparatively few researchers in the past (see Sloboda 1982). After Seashore's pioneering work in the thirties (see Seashore 1938), the work of Ingmar Bengtsson and co-workers, started more than two decades ago, remained more or less singular for a long time. (A recent account of their work is given in Bengtsson and Gabrielsson 1983). Researchers in the music acoustic area have been no exception in this respect, even though music performance must be regarded as an essential topic in the science of musical instruments; for instance, it would be hard to define 'instrument quality' were there no scientific understanding of exactly what the sound characteristics are which the musicians make use of when they play their instruments. However, during the last few years, the scientific interest in

music performance has increased considerably (see for instance Clynes 1983). Perhaps the reason for this is the significance of such research to music which is performed without musicians, such as electro-acoustic music, particularly computer music.

The key questions in scientific research are *how?* and *why?* Applied to music performance they concern *how* the note signs are related to their corresponding acoustic signals in acceptable music performances; and *why* such a performance deviates from the nominal values specified in the music score. The present project attempts to contribute to the answers to these questions by formulating and interpreting rules for the conversion of note signs into sounding music.

## Method

### Set-up

Music performance can be studied by means of two different strategies, namely (1) measurements (see Bengtsson and Gabrielsson 1980) and (2) analysis-by-synthesis. *Measurements* have revealed a complex reality, to say the least. Seashore (1938) found that tones having the same nominal value in the music score occasionally may differ widely in duration. For instance, in the same piece a crotchet sometimes may be played shorter than a dotted quaver and sometimes longer than a dotted crotchet. Similar observations have been made in subsequent studies of timing in music performance (Bengtsson and Gabrielsson 1977; Sloboda 1983). It seems that the rules relating notation to sounding music are numerous and form a complex system. However, when combined with systematic experiments, measurement of music performance is a useful method.

Analysis-by-synthesis is a well-established method in natural sciences, including speech research, of testing hypotheses. When applied to music performance, hypotheses can be tested regarding rules for the conversion of note signs to tones. Thereby, even aspects of performance for which there is no generally accepted term may be investigated, so that a great variety of effects can be investigated.

The block diagram in Fig. 3.1 describes the procedure used in the present project. The overall aim is to convert a melody into tones according to a set of ordered rules. Using a program written by J. Liljencrants the one-part input melody can be written in ordinary music notation on the computer's graphic terminal. In addition to the information normally included in a music score, special signs were introduced to mark (1) chord changes and (2) boundaries between phrases and subphrases.

The rule system then converts the entire string of signs into a string of tone symbols. The rules are expressed in a special programming language,

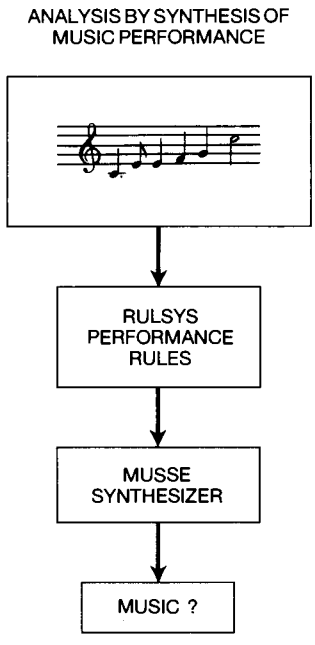



**Fig. 3.1.** Block diagram of the procedure used in the present project.

'RULSYS', written by Rolf Carlson and Björn Granström (1975) and originally devised for expressing rules regarding pronunciation, etc. in a text-to-speech conversion, as mentioned. The entire melody is processed by each rule in the order given by the rules in the system. The end result is that the tones are represented by vowel symbols which possess a number of different characteristics, such as duration, fundamental frequency, amplitude, vibrato amplitude, and other parameters, which can be manipulated by the rules. Finally, these characteristics are converted into control voltages for the MUSSE synthesizer.

In the present experiment the sound characteristics of the synthesizer were adjusted so as to be similar to the timbre of a double-reed wind instrument with a small vibrato. The fundamental frequencies were in accordance with the equally tempered scale, and could be modified in steps of 7 cents, which is close to the difference limen for frequency in normal subjects (Rakowski 1971). The amplitude was controlled in steps of 1/4 dB. The duration of the tones was computed in ms and then converted into cs. According to measurements by van Noorden (1975) a change in the perceived rhythm of a sequence of tones appeared when the durations were changed by more than 10 ms.

### Procedure

As mentioned, the basic idea of the analysis-by-synthesis strategy is that of testing hypotheses. The long-term musical experience of my co-worker Lars Frydén has resulted in a number of hypotheses regarding basic principles for converting note signs into tone sequences.

The hypotheses suggested by Frydén have been expressed in terms of rules which have been implemented in the rule system. Each hypothesis has then been tested informally by listening to several examples which clearly expose the sounding consequences of the rule. Depending on the result, the rule has then been modified. Eventually, formal listening tests were realized in which musically highly competent listeners compared two slightly different versions of the same melody; in one version, all rules were applied in generating the performance, and in the other version all rules except the one tested by that example were applied. In this way each rule can be validated to the extent that it can be decided whether or not it significantly contributes to raising the musical quality of a performance of the melody tested.

## Rules

Our present rule system includes about 20 rules. Some of these rules have been presented in detail together with sound examples in some previous reports (Sundberg *et al.* 1983*a*, *b*; Sundberg and Frydén 1985). Even though some rules have been more or less modified since then, only a few examples of the rules will be presented in this section; all of these have been approved of by a listening panel, unless otherwise stated.

The formulation of the rules is quite simple. The basic form of a rule is 'Every note possessing the characteristics A, B, . . . should be assigned the characteristics of D, E, . . . provided that is preceded by F, G, . . . and followed by H, I, . . .'. The context is optional. Also, there are means to fetch values from tones far back and far ahead in the sequence of note symbols.

There are several rules which work within a narrow time window including two or three notes only. In other words, the context is short. Some examples of such rules are the following.

(1) *The higher, the louder.* The amplitude is increased by about 3 dB per octave rise in fundamental frequency.

(2) *Heightening of durational contrasts.* This rule simply states that short note values are shortened. The shortening is very small indeed and, in terms of percentage, it is higher for quavers than for crotchets. No compensaion, such as a pause, is made for the resulting perturbation of the strict regularity of the beats.

(3) *Softening of durational contrasts.* In apparent contradiction to the previous rule this rule states that the durational contrast should be *reduced.* However, there is an important contextual condition for applying the rule, namely that the note is preceded by a note of twice its duration and followed by a tone with longer duration. This rule has not yet been tested in listening experiments.

(4) *Shortening of start tone in rising intervals.* The duration of a note initiating a rising interval is slightly decreased. In sequences of rising intervals the rule has the effect of increasing the tempo somewhat. As before, no compensation is made for the resulting perturbations of the meter.

There are also some rules which work with a wider time window. For these rules the signs representing chords and phrase and subphrase boundaries are decisive. In the input notation the chords are indicated by numbers representing the distance in semitone steps from the root of the chord to the root of the tonic. This number may be complemented by a sign for minor chords. As regards the phrasing we assume that melodies are built up by phrases, which consist of two or more subphrases. The boundaries between such phrases and subphrases are marked in the input notation.

(5) *Marking of harmonic charge.* This rule generates crescendos/decrescendos and ritardandos/accelerandos reflecting the chord progressions.The technicalities are as follows.

The rule assigns to each chord a value reflecting its *harmonic charge.* This harmonic charge is a weighted sum of the chord notes' *tonic charges**. As illustrated in Fig. 3.2, the tonic charge reflects the note's distance from the root of the tonic along the circle of fifths. The harmonic charge value is converted into an amplitude increment, which equals 6 dB for the highest harmonic charge. This amplitude increment is added to the first tone appearing over each chord. Then, the amplitudes of the intermediate tones are linearly interpolated on a dB-scale. Too-slow crescendos are avoided by letting the amplitude remain constant until about 2 s before the new chord appears. The result is crescendos and decrescendos towards harmonic changes which introduce chords having a higher or a lower harmonic charge than the preceding chord, respectively. Figure 3.3 demonstrates how the rule operates. This rule has not yet been verified by a formal listening test.

In addition to this, the tempo is also decreased and increased in synchrony with the crescendos and the diminuendos, respectively. These

* In previous reports on this project the terms 'harmonic distance' and 'tonic distance' have been used. However, as pointed out by Drs D. Butler and H. Brown, Music Department, Ohio State University, Columbus (personal communication), the word 'distance' may be misleading. The term 'charge' seems preferable, although a similar term, '*Quintspannung*', was defined slightly differently by Kurth (1917).

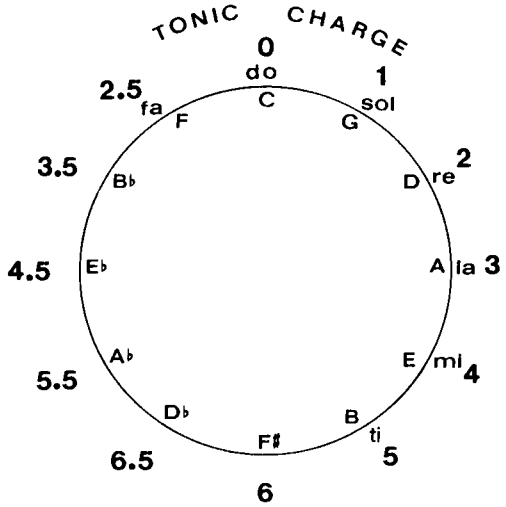



**Fig. 3.2.** Values of tonic charge for the various tones in the diatonic system. The tonic charge of a tone depends on the tone's distance from the root of the tonic along the circle of fifths. The *harmonic* charge value can be computed for any chord as a weighted sum of the chord notes' *tonic* charge values.

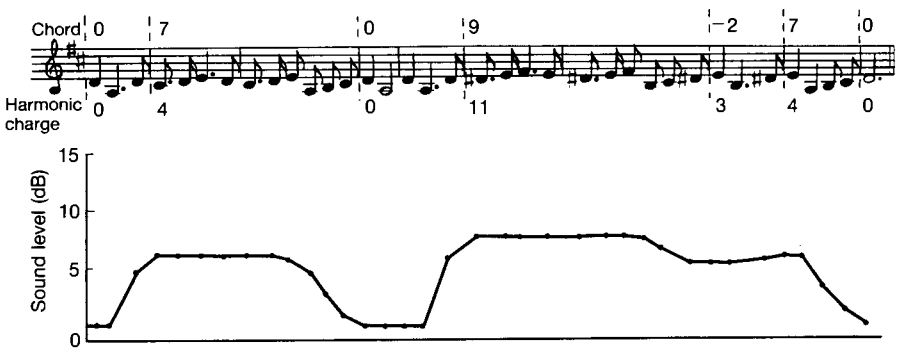



**Fig. 3.3.** Illustration of the amplitude effects of the rule reflecting changes in the chords' harmonic charge as applied to the second theme from the first movement of Schubert's Symphony No. 8 in B minor, 'Unfinished'. The chords are symbolized by numbers corresponding to the interval (in semitone steps) between the root of the chord and the root of the tonic (top line); the minus sign represents a minor triad. The numbers below the staff represent the chord's harmonic charge value. Crescendos and decrescendos accompany increases and decreases of harmonic charge, respectively.

tempo changes mostly involve note lengthenings of 0 to about 15 per cent. (6) *Phrase and subphrase marking.* The rule first adds 4 cs to the note terminating a phrase and 4 additional cs to the final note of the melody. In addition the final note in each phrase is terminated by a short pause, as the amplitude falls to zero during the last 20 per cent of the note's duration. Similarly, in notes terminating subphrases, the amplitude decays to zero during the last 80 ms of the note.

## Discussion

All these rules share an *ad hoc* character, and in some cases they may even seem mysterious. Therefore, a primary concern is to verify the rules. This can be realized in various ways. Before this, it is necessary to consider a more general question, namely the apparent contradiction between two obvious facts: on the one hand, a melody can be played in many different ways, all of which are fully acceptable from a musical point of view; on the other hand, a rule system will predict only one single performance of a given melody.

First it should be pointed out that identity with respect to the sequence of notes does not imply that the melodies are the same. This relates to an experiment recently carried out by Sloboda (1983), who had pianists play a set of melodies several times. Two melodies in this set were identical except for the location of the bar lines. The performances of these two different but 'homonymous' melodies were then analysed. The performances of the most experienced players differed consistently with respect to both timing and amplitude in certain places in the melody. The question then is whether or not different performances will be generated by our rule system in such a case. The answer is yes; the input notation may differ even in cases where the note sequence is identical, because in the input, signs are added for chord changes and phrase and subphrase boundaries, and these would differ in a case such as the one used in Sloboda's experiment. Thus, two slightly differing performances will result, if the rule system processes such 'homonymous' melodies. In many cases it is not even necessary to move the bar lines or the chord changes for obtaining different interpretations in terms of phrase and subphrase boundaries. Music is often ambiguous in this respect, and these ambiguities will cause our rule system to predict differing performances.

Apart from these cases, where the same sequences of tones represent different melodies, there is another possibility for our rule system to predict a variety of performances for the same melody. Even though this possibility has not yet been explored, it will be mentioned here. In our present rule system there are about 20 rules, and several of these affect one parameter, such as the duration of the note. This means that the duration

of one single note in a melody may be lengthened by some rules and shortened by some other rules. Still it is obvious that our system is not yet complete; there must be more rules, which would add to the musical quality of the performance. Introducing them into the rule system may cause problems, though. It seems likely that there is an upper limit on the number of rules which can be used simulaneously to affect one specific sound parameter. Perhaps there are many rules which are alternative, so that the musician may make a new selection of the alternative rules he wants to apply each time he plays the piece. This possibility should be tested in future research. In any event, it would lead to a rule system that could generate different performances of the same piece.

As suggested before, the above has important implications for the possibilities of verifying the different rules in music practice. Obviously, it is not correct to assume that eventually there will be a perfect match between the performance predicted by a complete rule system and an actual perfect performance, as there are many perfect performances. Still, similarities may occasionally occur in some actual performances, so it is still meaningful to compare generated and observed performances.

In Fig. 3.4 a performance predicted by the rule system is compared with

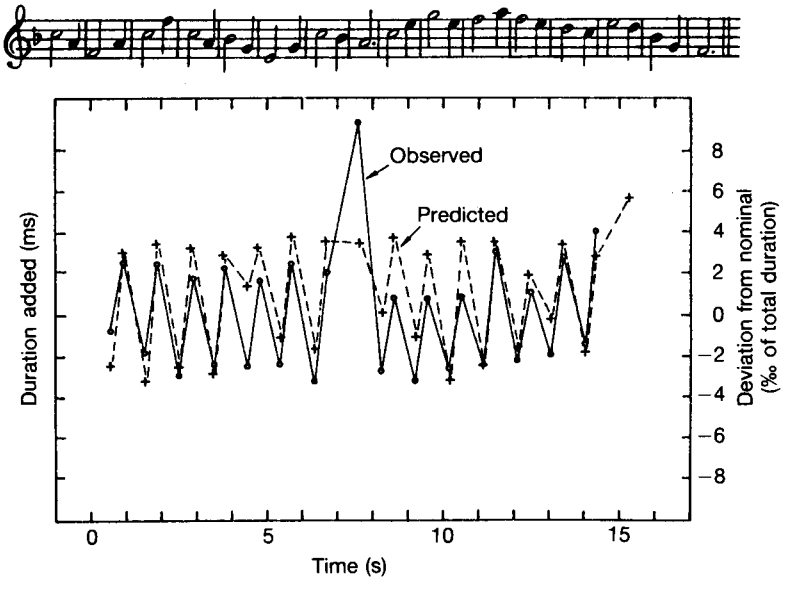



**Fig. 3.4.** Comparison of observed and predicted durations in a performance of the melody shown. The predicted values are given in cs, and the observed values, quoted from Gabrielsson and Bengtsson (1983), are given in thousandths of the summed duration of the melody.

an actual performance as measured by Gabrielsson and Bengtsson (1983). The figure compares observed average deviations from nominal note values in actual performances with the deviations predicted by the rule system. It appears that rule (3), which reduces durational contrasts in halving of duration, has had a major effect on the end result. Actually, this rule, which was not yet verified in a formal listening test, seems to be efficiently validated by the good agreement between observed and predicted data. Such a high degree of agreement between predicted and observed performances is rare. Particularly in more complex melodic structures, such as the one illustrated in Fig. 3.5, the relationship between observation and prediction is much less clear.

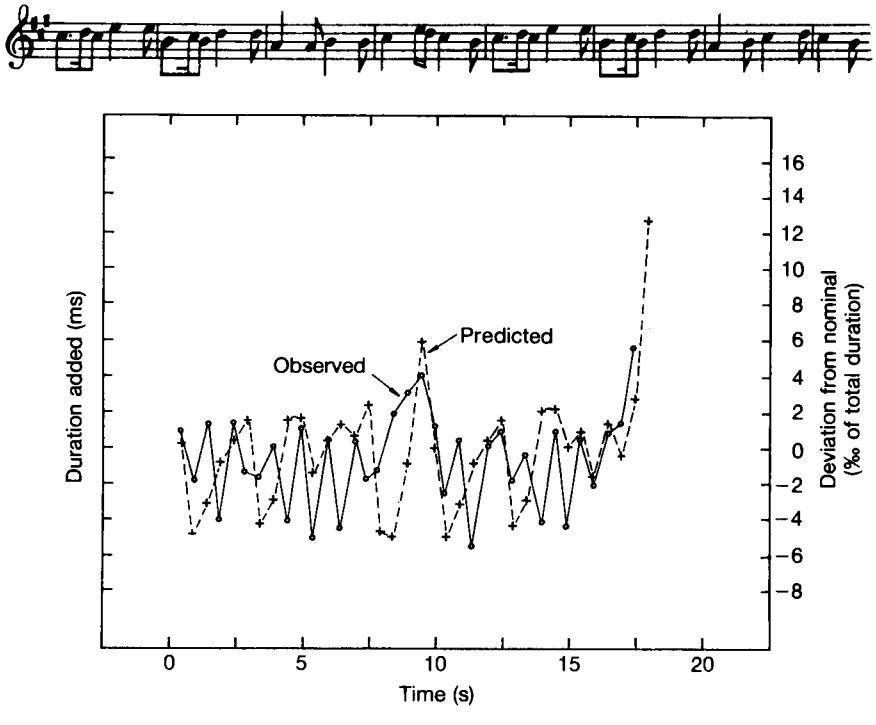



**Fig. 3.5.** Same type of comparison as in Fig. 3.4 regarding the theme of the first movement in Mozart's Piano Sonata in A (K. 331).

What do such similarities and dissimilarities between predicted and observed performances tell us? Our rule system can be regarded as a predictive model of the reality manifested in music performance. When there is an agreement between observation and prediction, it is safe to conclude that the predicting model represents a fair description of reality.

However, the reverse does not seem to be necessarily true. In other words, a disagreement cannot be used for rejecting the model. The reason for this is the fact that there are many ways of performing the same melody which, although differnt, are all musically perfectly acceptable. Rather, the criterion for an acceptable system of rules is that experienced listeners agree that the performance generated by the rule system sounds musically acceptable. This has been tested in the listening experiments mentioned previously. An evaluation of the complete rule system will be carried out in the future (see Thompson *et al.* 1986).

Let us know turn to a different aspect of the results, namely the quantity of the rules; that is, *how much* amplitude or duration, etc., is added by the various rules. The first question to consider is then to what extent random variation of duration can be tolerated in musically acceptable performances (Gabrielsson and Bengtsson 1983). Our experience is that, mostly, duration changes as small as 10 ms can be noted in motor music (where there is a rapid sequence of beats). This is in accordance with van Noorden's results (1975). It implies that random variations of duration must be smaller than 10 ms, as, obviously, noticeable random variation cannot be accepted in a perfect performance.

As was mentioned, the quantities by which these rules affect the amplitude and duration of the individual note are sometimes exceedingly small. Yet, the effects thus generated as essential to the impression we get of the performance. This is evident when listening to examples where in the input notation the phrase markers have been replaced by subphrase markers and vice versa (the interested reader may experience this by listening to the corresponding sound illustration in Sundberg and Frydén 1985). The typical reaction of music listeners is that this simple substitution results in an unacceptable performance of the melody (Friberg *et al.* 1987). This shows that our sensitivity to these minute perturbations of amplitude and duration is very high indeed. As soon as the microperturbations appear in wrong places, they are easy to hear, but when they appear in correct places, they are hard to notice.

Interference between rules has often forced us to reduce the quantity of one rule when another rule was introduced. For instance, our first crescendo simply implied an increase of the output amplitude. This is an unrealistic crescendo, however, as higher overtones typically gain more than lower overtones when amplitude is increased on traditional musical instruments. When this feature was introduced accompanying the changes in amplitude, it became necessary to reduce the amplitude changes. Similarly, the quantity of other rules has been reduced to a quantity smaller than that which was used when the rule was tested in a listening experiment with experienced musicians. As yet, we have simply added the latest rule to the previous rules which have been adjusted accordingly, if required. This means that the quantity for a rule which is just right with the

present combination of rules may turn out to be too large if more rules are added in the future.

A typical observation should be mentioned which we have made several times when trying to adjust the quantity by which a rule affects a note or a series of notes. When the quantity is too small, the effect cannot be heard at all, of course. More interestingly, when the quantity is too great, two phenomena can be noted. One is that it is easy to hear exactly what the rule is doing in a physical sense. The other is that the effect appears musically unacceptable, simply because it sounds exaggerated. The 'correct' or musically useful quantity is in between these two extremes. It is typically characterized by the situation that a listener notices the effect but is unable to analyse it correctly in physical terms; as soon as this is possible, the quantity is too great and the effect sounds exaggerated. Thus, there seems to be a difference between the quantity needed in order to make a correct aural analysis of the physical means used; what is acceptable from a musical point of view seems to be just beyond the border of what is noticeable.

The analysis-by-synthesis method has previously been applied extensively in speech research and there it has been found useful. Still, the method has certain limitations. The most important one is that, being a basically comparative method, it can merely show what is better or worse than something else. For instance, we can find out if a rule improves the musical quality of a performance, but no conclusions can be drawn as to what is the best possible performance. On the other hand, this limitation is no great concern; even with respect to 'real' music it would be wrong to pretend that this or that performance is *the* best possible performance; there are always several ideal performances of a piece.

The analysis-by-synthesis technique as applied to music performance presupposes that musical experts are capable of agreeing on what is more acceptable and what is less acceptable from a musical point of view in a musical performance. We have regarded this ability as a psychological fact, and there seem to be no serious reasons for doubting its existence among music experts. For instance, if it did not exist, it would be hard to explain why music conservatories can exist.

The analysis-by-synthesis procedure shows striking similarities with the procedure used in teaching music playing; in both cases the principle is to give instructions that strive to eliminate some obvious deficiency in the performance. A major difference is that computers always apply all rules, while this would probably be regarded as a sign of little musical talent in a human musician! However, this difference between the computer and the music student is essential for the evaluation of the 'instructions' which we give to the computer.

Summarizing this slightly rhapsodic discussion so far, our examples and listening tests prove that our rules improve the musical acceptability of a

performance at least in the piece of music that has been tested. This is an interesting fact that asks for an explanation, regardless of the generality of the rule, and regardless of whether the rule could be formulated in a better or alternative way. Thus, as soon as we have formulated a rule and found that it improves music performance, we have reasons to question why it improves the performance. The answer to this question would be interesting, because the rules must be assumed to represent some sort of reality in music communication. Next we will attempt to analyse this reality.

**Purpose of the rules**

It seems that each of the rules that we have formulated as yet serves a specific purpose or, rather, it seems to have a function in music communication. There are a number of such purposes, and the rules can be divided into different groups according to their apparent purposes.

One group of rules seems to serve the purpose of emphasizing the contrasts along a commonly used dimension. An example of this in the rules described above is rule (2), which exaggerates durational contrasts.

A second type of rule seems to serve the purpose of eliminating sound characteristics which sound queer because no traditional musical instrument possesses them. An example is rule (1), which increases the loudness as function of pitch; most traditional instruments adhere to this principle. If the synthesis does not adhere to this rule, it sounds unnatural or 'electronic'. This is logical: the listener has never heard this sound characteristic before except from electronic instruments.

A third group of rules seems to take into account the fact that a listener might make associations with previous experiences when listening to music. For instance, (a) a rapid, regular pulse would often lead the listener's thoughts to rapid locomotion such as swift walking or running, and (b) an 'upward motion' on the scale would cause many listeners to imagine something climbing upward in space. Perhaps a combination of these associations constitutes the background of rule (4), increasing the tempo with the upward motion of a melody: the slight increase in tempo can be interpreted as an indication that it is not heavy to walk, even though the road leads upward.

A fourth group of rules seems to serve the purpose of emphasizing more or less unexpected tones. It is as if the musician wanted to say that even though this note sounds a bit peculiar, it is just this note that he intends to play. The occurrence of chords having a high harmonic charge is prepared by loudness increases, while the amplitude is decreased towards chords which have a smaller harmonic charge.

A fifth group of rules apparently serves the function of marking the structure of the melody. One example is the rule lengthening the final

element of a phrase. Rule (5), which emphasizes harmonically remote chords by loudness increases, probably serves this purpose too. The background is that in traditional harmony the chord progressions are often used to mark the structure, so that a phrase ending is often marked by a cadenza, that is, a chord preceded by its own dominant—or, in other words, a decrease of the harmonic charge (see Sundberg and Lindblom 1976). For instance, one of the phrases in a melody written in C major might be terminated by the chord progression E major—A minor. Under these circumstances our rules would predict a crescendo towards the E major chord and then a decrescendo towards the A minor chord. This loudness change would then serve a double purpose. One purpose is to inform the listener that the E major is actually intended even though it is rather far from the tonic (no-mistake case), and the other purpose is to inform the listener that this is a termination of a structural unit in the melody (marking-of-structure case).

Apart from the groups of rules exemplified above there is at least one more group. The purpose of the rules in this group is to embed the performance in a sort of emotional atmosphere. Particularly in the synthesis of singing, the need for such rules is experienced as evident; in a melody with a tragic character, it is most disturbing if the performance does not sound tragic. Interesting attempts to investigate how such emotional atmospheres are signalled in music performance are being made by Clynes (see for example Clynes 1983).

If we now survey the groups of rules described above, we find that the rules apparently serve purposes which are not specific to music. The heightening of contrasts in a frequently used dimension seems to be a common psychological phenomenon in communication. Examples of the preference for phenomena that the listener/spectator is already familiar with can probably be found in most fields of human communication. Emphasis is a phenomenon that would appear in all systems of formalized communication. To embed a message in an emotional atmosphere is typical in human behaviour, including speech. Musical communication thus seems to contain several elements that are not specific to music, and many elements seem to be directly imported from speech. This is not surprising: speech and music are the two main forms of systematized interhuman communication by means of sound signals. In future research it would be interesting to see whether great differences in the speech of two cultures correspond to great differences regarding the rules used in music performances.

## Code

Above we have considered the question of what *purposes* the various rules in our performance rule system seem to serve. We can now pose another

question, namely, what the *means* are that are used to fulfil these different purposes. Thus, we will consider questions such as 'How is emphasis signalled?' and 'How is the structure of the melody marked?' It is by no means obvious how the code is chosen. It is possible that the code is imported from extramusical communication, such as speech, but it is also possible that the code is unique for music.

Before going into these questions it is improtant to point out that the acoustic parameters that we have dealt with in this work are only pitch, amplitude, timing, and vibrato. There is no doubt that these are sufficient for making music; this is evident from the fact that the organ is useful as a musical instrument, because in this instrument it is only timing that can be continuously controlled by the player. Nevertheless, it is important to remember that we can make inferences about the coding in terms of amplitude and timing only.

Much of the code seems to be taken from speech. One example is the phenomenon that a note appears emphasized when it is lengthened; in speech, emphasis can be signalled by means of an increased syllable duration as has been demonstrated by Carlson *et al.* (1974). Another example is the marking of phrase endings by means of lengthening, because many languages use 'final lengthening' for signalling the end of a phrase (Lindblom 1978). By listening to speech, music listeners have been 'programmed' to interpret lengthening as a possible sign of a termination. If so, lengthening would be impossible to use as a sign of continuation. The previously mentioned queer effect obtained when phrase and subphrase markers were switched in a melody supports this hypothesis.

However, it seems doubtful that speech is the ultimate source of the code. For instance, in some languages, such as Danish, lengthening of syllable duration is not the code used in order to flag the end of a phrase. Still, Danish and Swedish musicians can of course make music together. This lends support to the previously mentioned hypothesis that there may be several codes available for expressing the same thing in music performance; the flagging of termination by means of lengthening is perhaps only *one of many* possibilities.

When a player uses lengthening for marking termination, he exploits the fact that the listener is 'programmed' to couple the phenomenon of lengthening with the notion of termination. In the mind of a music listener there would be many other examples of such couplings between sound patterns and various phenomena which allow the listener to interpret the sound sequences. Such couplings are likely to result partly from the listener's extramusical experience, such as his acquaintance with speech, and partly from his previous musical experiences. The typical final retard is a convincing example of a coupling of the first-mentioned type. It is typically used in motor music, where there is a regular, rapid pulse. It is likely that most listeners tend to associate this pattern of pulses with

locomotion. One way of stopping locomotion abruptly is by collision, which most listeners would agree is an unpleasant experience. It may be for that reason that the final retard is often used in performance of motor music. (Kronman and Sundberg 1987).

The main point, we believe, is that the player always applies rules in his playing, because it is only by means of following rules that the player enables the listener to interpret the performance. Thus, we would speculate that all details in a music performance are interpretable as long as these details were derived from performance rules which exploit the listener's previous intra- as well as extramusical experience. We also believe that the basic requirements for musical communication can be studied efficiently by means of formulating these performance rules.

## Concluding remarks

Music performance seems to be a scientifically rewarding topic, where analysis-by-synthesis is a useful method. A number of unexpected findings and observations have emerged, and, no doubt, much work remains to be done, such as a thorough testing of all hypotheses generated by the results. Still, even before this has been done some conclusions can be drawn.

By demonstrating how the sound sequences must be modelled in order to increase the musical acceptability of a performance, explicit information is gained on what is demanded from a musically acceptable performance. The analysis-by-synthesis method offers information in terms of a precise description of sound patterns. In many case these patterns can be recognized as allusions to patterns, which must be well known to the listener because of his or her experience of speech and also of motion as manifested by, for example, footsteps. By means of such allusions, emphasis, structure, and emotional atmosphere can be expressed in a way that the listener can understand. In this way music seems to gain meaning.

In summary, it seems that analysis of music performance is beginning to reveal what exactly it is that musicans add to written music, what the code they use is and from where this code is taken. Therefore it appears fair to propose that analysis of music performance can shed light on basic aspects of music communication.

## Acknowledgements

This project is supported by the Bank of Sweden Tercentenary Fund.

This is one of many articles about this project written for different readerships during the last years. As the rule system has been continuously

modified during this period, some rules can be found in a different formulation in previous articles.

## References

Bengtsson, I. and Gabrielsson, A. (1977). Rhythm research in Uppsala. In *Music room acoustics*, Publications issued by the Royal Swedish Academy of Music, No. 17, pp. 19–56, Stockholm.

Bengtsson, I. and Gabrielsson, A. (1980). Methods for analyzing performance of musical rhythm. *Scandinavian Journal of Psychology* **21**, 257–68.

Bengtsson, I. and Gabrielsson, A. (1983). Analysis and synthesis of musical rhythm. In *Studies of music performance* (ed. J. Sundberg), Publications issued by the Royal Swedish Academy of Music, No. 39, pp. 76–181. Stockholm.

Carlson, R. and Granström, B. (1974). A phonetically oriented programming language for rule description of speech. In *Proceedings of the speech communication seminar*, Vol. 2, *Speech production and synthesis by rules*, ed. G. Fant) pp. 245–53. Almqvist and Wicksell, Stockholm.

Carlson, R., Erikson, Y., Granström, B. Lindblom, B., and Rapp, K. (1974). Neutral and emphatic stress patterns in Swedish. In *Proceedings of the speech communication seminar* Vol. 2, *Speech production and synthesis by rules*, (ed. G. Fant), pp. 209–18. Almqvist and Wicksell, Stockholm.

Clynes, M. (1983). Expressive microstructure in music, linked to living qualities. In *Studies of music performance* (ed. J. Sundberg), Publications issued by the Royal Swedish Academy of Music, No. 39, pp. 76–181. Stockholm.

Friberg, A., Sundberg, J., and Frydén, L. (1987). How to terminate a phrase. An analysis-by-synthesis experiment on a perceptual aspect of music performance. In *Action and perception in rhythm and music* (ed. A. Gabrielsson), Publications issued by the Royal Swedish Academy of Music, No.55.

Gabrielsson, A., Bengtsson, I., and Gabrielsson, B. (1983). Performance of musical rhythm in 3/4 and 6/8 meter. *Scandinavian Journal of Psychology* **24**, 193–213.

Kronman, U. and Sundberg, J. (1987). Is the musical ritard an allusion to physical motion? In *Action and perception in rhythm and music* (ed. A. Gabrielsson), Publications issued by the Royal Swedish Academy of Music, No. 55.

Larsson, B. (1977). Music and singing synthesis equipment (MUSSE). Speech Transmission Laboratory Quarterly Progress and Status Report UTH, Stockholm 1977/1, pp. 38–40.

Lindblom, B. (1978). Final lengthening in speech and music. In *Nordic prosody* (eds. E. Gårding, G. Bruce, and R. Barnett), Dept. of Linguistics, Lund University pp. 85–101.

Noorden, L. P. A. S. van (1975). Temporal coherence in the perception of tone sequences. Doctoral thesis. Druk vam Voorschooten.

Rakowski, A. (1971). Pitch discrimination at the threshold of hearing. *Proceedings of the 7th International Congress on Acoustics*, Budapest, Vol. 3, pp. 373–76.

Seashore, C. E. (1938). *Psychology of music*. McGraw-Hill, New York. (Reissued 1967 by Dover, New York).

Sloboda, J. (1982). Music performance. In *The psychology of music* (ed. D. Deutsch), pp. 479–96. Academic Press, New York.

Sloboda, J. (1983). The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology* **A35**, 377–96.

Sundberg, J. (1978). Synthesis of singing. *Swedish Journal of Musicology* **60**, (1), 107–12.

Sundberg, J. (1982). In tune or not? A study of fundamental frequency in music practice. In *Tiefenstruktur der Musik, Festschrift für Fritz Winckel* (ed. M. Krause), pp. 69–96. Universität und Akademie der Kunste, Berlin.

Sundberg, J. and Frydén, L. (1985). Teaching a computer to play melodies musically. In *Analytica: studies in the descriptions and analysis of music in honour of Ingmar Bengtsson*, Publications issued by the Royal Swedish Academy of Music, No. 47, pp. 67–76. Stockholm.

Sundberg, J. and Lindblom, B. (1976). Generative theories in language and music descriptions. *Cognition* **4**, 99–122.

Sundberg, J., Frydén L., and Askenfelt, A. (1983*a*). What tells you the player is musical? An analysis-by-synthesis study of music performance. In *Studies of music performance* (ed. J. Sundberg), Publications issued by the Royal Swedish Academy of Music, No. 39, pp. 61–7. Stockholm.

Sundberg, J. Askenfelt, A., and Frydén, L. (1983*b*). Music performance: a synthesis-by-rule approach. *Computer Music Journal* **7**, 37–43.

Thompson, W. F., Friberg, A., Frydén, L., and Sundberg, J. (1986). Evaluating rules for the synthetic performance of melodies. Speech Transmission Laboratory Quarterly Progress and Status Report, UTH, Stockholm, 2–3/1986, 27–44.

# 4

# Timing and synchronization in ensemble performance

RUDOLF A. RASCH

## Introduction

For the last six centuries Western music has been largely polyphonic. This means that a piece of music consists of two or more simultaneous parts or 'voices' (either vocal or instrumental). In order to make performance possible, the temporal relations between the various voices as well as those within a single voice have to be defined exactly and unambiguously. These relations must be clearly stated in the score of the music and appropriately realized in a performance. In its initial stage polyphonic music was often called 'measured music' (*musica mensurata*) as opposed to 'plain music' (*musica plana*), that is, plainchant in which there was no requirement of strict time structure.

Observations from such widely different fields as timbre experiments, instrumental sounds, and notation lead to the conclusion that the beginning or onset of a tone is the decisive point in time for that tone. Timbre experiments have shown that recognition of musical instrument tones is severely impaired by taking away the starting transient portion (Clark *et al.* 163; Berger 1964; Saldanha and Corso 1964). Widely used instruments such as piano, harpsichord, and guitar (in general, all plucked or struck instruments) produce tones without steady-state parts, which begin to decay immediately after the onset. Traditional music notation puts notes at the onset moment of the horizontal time scale, and places those with simultaneous onsets together irrespective of their lengths. Investigation of the temporal structure of music is therefore predominantly a study of onset moments.

In this chapter we deal with the problem of synchronization in musical performance, that is, how musicians manage to co-ordinate their own temporal (onset) structures with those of the other performers in such a way that the temporal structures actually match each other and fuse into

one common temporal structure. This means that the tones which are prescribed in the score as simultaneous must be played at the same moment of time, especially with regard to their onsets. In reality, perfect synchronization is never realized and there will always be some asynchronization. Onsets meant to be simultaneous but in performance not exactly simultaneous will be called quasi-simultaneous onsets.

In the first and second sections of this chapter, we will present a model for the description of synchronization of ensemble performances (Rasch 1979, 1981). It is based on the standard deviations of relative onset times and onset time differences. The third section is devoted to the question of how to define the onset times of musical tones in a synchronization situation (Vos and Rasch 1981). Finally, we will speculate a little about the role of the conductor in the synchronization of musical performances.

## The description of synchronization

When musicians play together in an ensemble, they will try to synchronize as much as possible the tones meant to be simultaneous. For a number of reasons, such as the restricted accuracy of human motor performance and time perception, the relative ease of tone production within or between instruments, and the time lag between the production of a player's own tones and the perception of the tones produced by others, a perfect synchronization is not possible in a live performance. There will always be some degree of asynchronization.

Onsets of tones meant to be simultaneous will, in reality, scatter a little in time. Assume an ensemble of $n$ players. Figure 4.1, illustrates some of

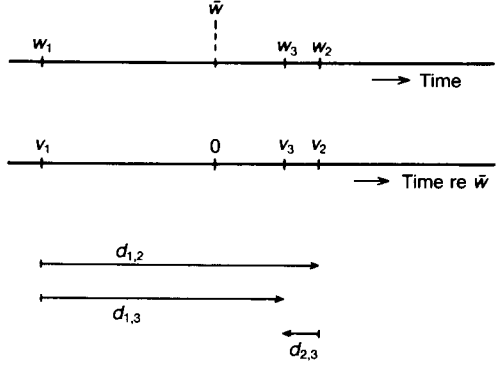



**Fig. 4.1.** Illustration of the concepts used for synchronization analysis. $w_1$, $w_2$ and $w_3$ are three onsets meant to be simultaneous. $\bar{w}$ is the mean onset time. $v_1$, $v_2$, and $v_3$ are the onset times expressed relative to $\bar{w}$. The onset time differences are denoted by $d_{12}$, $d_{23}$, $d_{31}$.

the concepts used for an ensemble of three performers. The *absolute onset times* of the tones meant to be simultaneous will be denoted by $w_1, w_2, \ldots, w_n$. For each group of such onsets there will be a mean onset time:

$$\bar{w} = \frac{w_1 + w_2 \ldots + w_n}{n}.$$

Now, the absolute onset times can be converted into *relative onset times*:

$$\begin{aligned} v_1 &= w_1 - \bar{w}, \\ v_2 &= w_2 - \bar{w}, \\ &\ldots, \text{ and} \\ v_n &= w_n - \bar{w}. \end{aligned}$$

The relative onset times are relative to the mean onset time, not to other onset times. Below, we will as a rule use the relative onset times. The absolute onset times of the tones in a piece of performed music can be measured and transformed into relative onset times. This will result in a set of $n$ distributions of relative onset times, one per voice. These distributions can be characterized by their means, which will be indicated by $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n$, and their variances, which will be denoted by $s_1^2, s_2^2, \ldots, s_n^2$. The means indicate which instruments lead or lag in their onsets relative to the other instruments in a performance.

Time differences between onsets are particularly significant for perception. The *onset time differences* will be denoted by the letter $d$:

$$\begin{aligned} d_{12} &= v_2 - v_1, \\ d_{13} &= v_3 - v_1, \\ &\ldots, \text{ and} \\ d_{n-1,n} &= v_n - v_{n-1}. \end{aligned}$$

The subscripts refer to the respective voices. Onset time difference has been termed *onset asynchrony* in other contexts. For a piece of music with $n$ voices or performers, there will be $n(n-1)/2$ distributions of onset time differences. These distributions can be characterized by their respective means, to be indicated by $\bar{d}_{12}, \bar{d}_{13}, \ldots, \bar{d}_{n-1,n}$, and variances, to be indicated by $s_{12}^2, s_{13}^2, \ldots, s_{n-1,n}^2$.

The distributions of onset time differences are related to those of the relative onset times. The means of the onset time differences can be calculated directly from the means of the relative onset times. The relation between the variances of relative onset times and onset time differences is more complex. It was found empirically that the variances of the several distributions of relative onset times for one specific piece of music do not differ very much from each other. For simplicity we will assume that they are equal, and denote their value by $s_v^2$. In addition we will assume that the covariances between distributions are equal. Their value will be denoted

by $c_v$. With these assumptions it is possible to derive a simple formula that relates the variances of the relative onset times and the onset time differences.

Each relative onset time can be written as a linear sum of all other relative onset times, for example:

$$v_1 = -(v_2 + v_2 + v_3 + \ldots + v_n).$$

The variances of such a sum equals the sum of the variances of the terms of the sum plus twice the sum of all covariances among the terms:

$$s_v^2 = \sum_{i=2}^{i=N} s_{v_i}^2 + 2 \sum_{i=2}^{i=n} \sum_{j=i+1}^{i=n} c_{v_{ij}} \; (j \neq i)$$

in which $c_{v_{ij}}$ is the covariance of $v_i$ and $v_j$.

Since, by assumption, $s_{v_1}^2 = s_{v_i}^2 = s_v^2$ and $s_{v_{ij}}^2 = c_v$, the variance of the relative onset times in general can be expressed in the following equation:

$$s_v^2 = (n-1)s_v^2 + (n-1)\,(n-2)c_v.$$

This leads to:

$$c_v = \frac{s_v^2}{1-n}.$$

Since an onset time difference can be considered to be a linear combination of the form $d_{12} = v_2 - v_1$, its variance is given by:

$$s_d^2 = 2s_v^2 - 2c_v.$$

Substituting the value of $c_v$ in the last equation leads to:

$$s_d^2 = \frac{2n}{n-1} s_v^2.$$

So the relation between the relative onset times and the onset time differences depends on the number of performers in the ensemble. For a duo ($n=2$) the relation is simply $s_d^2 = 4s_v^2$. For a large ensemble, the relation converges to the limit $s_d^2 = 2s_v^2$.

In actual practice, onset time difference is a variable with a mean not very different from zero (mostly within the range of $-5$ to $+5$ ms) and a standard deviation that is much greater (mostly within the range of 30 to 50 ms). For this reason the standard deviation of the onset time differences is a better measure for the amount of asynchronization in performed music than the mean onset time differences. We will define the *asynchronization of a pair of voices* as the standard deviation of the onset time differences of simultaneous tones of those voice parts. In formula, the asynchronization will be denoted by the capital $A$: $A_{ij} = s_{ij}$.

The standard deviations of the various distributions of onset time

differences in one piece of music do not differ very much from each other. We do not lose very much information if the asychronizations of the various voice pairs are averaged, by taking the root-mean-square. This averaging has the advantage that the asynchronization of a piece of music can be expressed as a single value. We will define the *asynchronization of a piece of performed music* as the root-mean-square of the standard deviations of the onset time differences for all pairs of voice parts. The asynchronization of a piece of music will be denoted by the capital $A$ without subscripts.

It should be noted that the asynchronization defined quantitatively in this way indicates ranges in which onset time differences will fall with certain probabilities. The range from $-A$ to $+A$ includes 68 per cent of onset time differences, the range from $-2A$ to $+2A$ 95 per cent. If we discard the sign of the onset time differences, it can be stated that the median unsigned onset time difference is $0.68A$, the lower quartile $0.32A$, and the higher quartile $1.15A$.

The synchronization of simultaneous tones is only one of the temporal tasks of performing musicians. Within one voice, the succession of tones also has to be strictly timed. We will call the timing of tones of equal duration in one voice the *isochronization* of the tones. The isochronization of a voice will be defined as the standard deviation of tone durations meant to be equal. In actual practice, there will be a number of sources that contribute to the variance of equal tone durations, like tempo trends and fluctuations. Also, the asynchronization of simultaneous tones will be mirrored in isochronization. If we exclude all other sources of the variance of 'equal' tone durations, there is a certain relation between asynchronization and isochronization (actually, a-isochronization). We consider two consecutive tones with absolute onset times $w_i$ and $w_{i+1}$ (see Fig. 4.2). The duration of the first tone is

$$h_i = w_{i+1} - w_i.$$

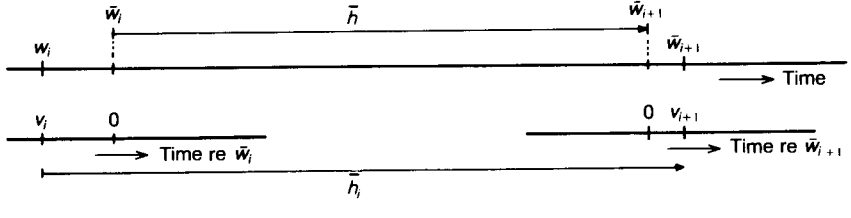



**Fig. 4.2.** Illustration of the concepts used in determining the relation between asynchronization and the standard deviation of tone durations. $\bar{w}_i$ and $\bar{w}_{i+1}$ are the mean onset times of two successive triplets of tones. $w_i$ and $w_{i+1}$ are the onsets of two successive tones in one voice part; $v_i$ and $v_{i+1}$ are the same onsets but expressed relative to the mean onset times. $\bar{h}$ and $h_i$ are the time intervals between the mean onsets and the onsets in a single voice, respectively.

The absolute onset time is the sum of the mean onset time and the relative onset time. We denote the mean onset time of tones simultaneous to $w_i$ as $\bar{w}_i$, that of tones simultaneous to $\bar{w}_{i+1}$. We assume that the successive mean onset times (. . ., $\bar{w}_i$, $\bar{w}_{i+1}$, . . .) are perfectly isochronous. Therefore, mean tone duration $\bar{h}$ is

$$\bar{h} = \bar{w}_{i+1} - \bar{w}_i.$$

The relative onset times of the two tones in succession will be denoted by $v_i$ and $v_{i+1}$. Now we can rewrite the actual tone duration as follows:

$$h_i = \bar{h} + (v_{i+1} - v_i).$$

The variance of this last expression is

$$s_h^2 = 2(1-r)s_x^2.$$

in which $r$ is the correlation between $v_i$ and $v_{i+1}$, or, the autocorrelation of relative onset times. The above equation can also be written as

$$s_h^2 = \frac{n-1}{n}(1-r)s_d^2.$$

The relation between isochronization and synchronization can be used in the comparison of the results of our measurements with the results of experiments on sequential timing behaviour which exist in the literature.

## The measurement of synchronization

The descriptive model given will now be illustrated with a set of measurements of live musical performances specifically recorded for this purpose (Rasch 1979, 1981).

In order to study synchronization in musical performances it is necessary to know the exact temporal structure of each individual part of the composition performed. Attempts have been made to analyse the compound sound of a piece into its individual parts, for instance by Tove *et al.* (1967) at the Royal Institute of Technology in Stockholm, and by Moorer (1975) at the Center for Computer Research in Music and Acoustics of Stanford University. Up to now, these attempts have been only partially successful. Several serious problems prevent application of the methods employed to larger, not specially selected samples of music. The main problem areas in this field are the storage and processing of the prohibitively large amounts of data, and the separation of fused harmonics in the cases of musical intervals with simple frequency ratios.

A more feasible approach is the one in which the acoustics signals from each instrument are picked up and recorded before they are mixed into the ensemble sound. This can be done by using either contact microphones

under normal acoustic conditions or directional microphones in an anechoic chamber. The latter method was chosen for our recordings because the recorded sound most resembled the sound in normal listening conditions. The recorded signals passed a logarithmic envelope detector with an output voltage proportional to the sound-pressure level of the temporal envelope of the signals. The sampling rate was 200 Hz, the time interval between two successive samples being 5 ms. The envelopes could be displayed on the graphic terminal of the computer. Peaks in the enveloped were matched with the notes of the score played by the musicians. The onset time of a peak was defined as the moment that the envelope reached or surpassed a certain, previously defined threshold level. This level was taken about 15 or 20 dB below the maximum levels of the signals. Figure 4.3 shows the envelopes of a fragment of 9 s of music, with the peaks numbered according to the computer output and the score of the music inserted for comparison and identification purposes.

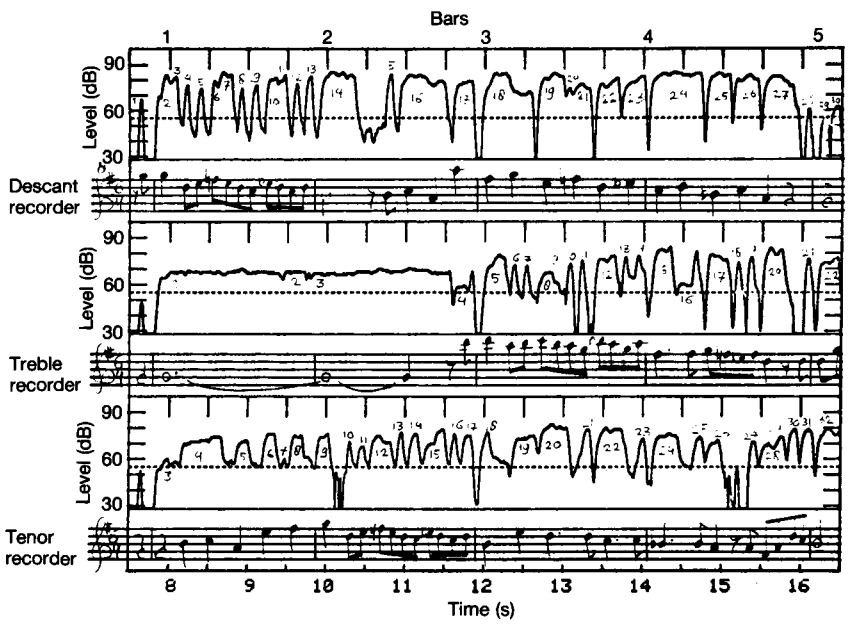



**Fig. 4.3.** The envelopes of three simultaneous voices of a fragment of 9 s. The peaks have been numbered, and the notes of the score have been inserted.

The following step was to determine which peaks in the three envelopes corresponded to a triplet of simultaneous notes in the musical score. Table 4.1 shows the simultaneous onsets from Fig. 4.3, the relative onset times, and the onset time differences. Note that the temporal resolution of the

**Table 4.1** *Peaks corresponding to simultaneous notes from the music fragment of Fig. 4.3, with absolute onset times, relative onset times, and onset time differences.*

| Instrument | | | absolute onset time (ms) | | | relative onset time (ms) | | | onset time differences (ms) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | | | | | | | | | |
| Peak no. | | | $w_1$ | $w_2$ | $w_3$ | $v_1$ | $v_2$ | $v_3$ | $d_{12}$ | $d_{23}$ | $d_{31}$ |
| 18 | 5 | 18 | 11970 | 11995 | 11955 | −3 | 21 | −18 | 25 | −40 | 15 |
| 19 | 8 | 20 | 12670 | 12675 | 12695 | −10 | −5 | 15 | 5 | 20 | −25 |
| 22 | 12 | 22 | 13405 | 13390 | 13410 | 3 | −12 | 8 | −15 | 20 | −5 |
| 24 | 15 | 24 | 14075 | 14105 | 14125 | −27 | 3 | 23 | 30 | 20 | −50 |
| 25 | 17 | 26 | 14795 | 14800 | 14795 | −2 | 2 | −2 | 5 | −5 | 0 |
| 27 | 20 | 18 | 15500 | 15510 | 15470 | 7 | 17 | −23 | 10 | −40 | 30 |

processing system was 5 ms. As a matter of fact, not all the triplets of simultaneous notes found in the score had a corresponding triplet of peaks in the envelopes. If the initial part of a tone did not correspond to a peak in the envelope, no corresponding onset time was available for the calculations.

The statistical computations were mostly done per composition, sometimes per movement of a composition. First, means, standard deviations, and root-mean-squares of standard deviations of relative onset times and onset time differences were calculated. These figures make up the basic data set of our analysis. Further computations included the autocorrelations of relative onset times, correlations with tempo measures, and conversion of synchronization into isochronization values.

Recordings were made of music played by three professional trio ensembles: a recorder trio, a reed trio (oboe, clarinet, bassoon), and a string trio. Each ensemble played two compositions that were typical for the repertoire.

Means and standard deviations of relative onset times were calculated for each recorded composition. The mean relative onset times are graphically presented in Fig. 4(a). They indicate to what extent certain instruments tend to lead or lag in simultaneous onsets relative to the other instruments. Certain tendencies can be discerned. In the string and wind trios the main melody instruments (violin, oboe) tend to lead relative to the others, with the bass instruments (cello, bassoon) in the second place, and the middle voice (viola, clarinet) in the third place. This corresponds to our intuitive notion that in most of the music for these ensembles the melody part is the leading, 'first' voice. Recorder ensemble music is more polyphonic as a rule, and the bass is the most fundamental voice. This is confirmed by the observation that the tenor recorder—the lowest voice—is the leading voice, on the average.

There are three distributions of onset time differences for each

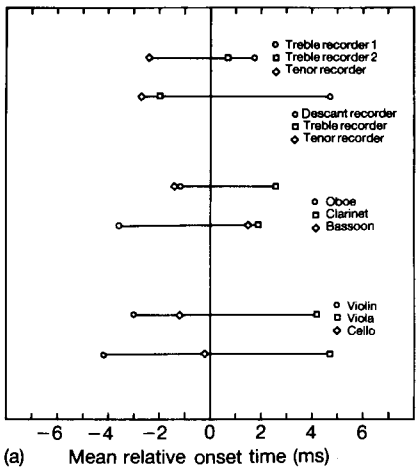



**Fig. 4.4(a).** Mean relative onset times of the instruments per composition.

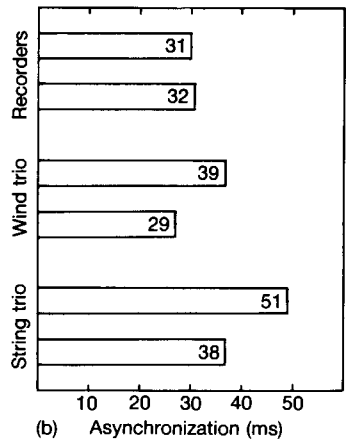



**Fig. 4.4(b).** Mean asynchronization per composition, averaged over the instruments.

composition. The root-mean-squares of the standard deviations are used as measures for asynchronization. Asynchronization per composition is graphically illustrated in Fig. 4(b). Asycnchronization values fall largely within the range of 30 to 50 ms. Mean asynchronization of this set of data is 36 ms. There are some systematic differences between the ensembles. The recorder ensemble shows a relatively small asynchronization of about 30 ms. The two examples of the wind trio have different amounts of asynchronization: 27 and 37 ms respectively. The largest asynchronizations are found in the string trio examples: 37 and 49 ms. These differences may be related to the rise times of the tones of the instruments in question. Recorders have short rise times, so that the beginnings of the tones are clearly marked. These short rise times make good synchronization both possible and necessary. The conventional woodwinds of the wind trio also have relatively short rise times. The onset of string tones is a more gradual process resulting in long rise times (30 to 100 ms). These longer rise times permit a greater leniency as to synchronization.

The relation between asynchronizatoin and tempo was also investigated (see Fig. 4.5). All data in this table concern single movements of the compositions. As tempo measure we used the time between two successive counting beats, the 'inter-beat time', determined as 60 divided by the metronome value (the number of counting beats per minute). This measure includes a subjective element in the choice of the counting unit. The correlation between the asynchronization of the movements and the tempo measure of 0.80. Evidently, the conclusion is justified that faster

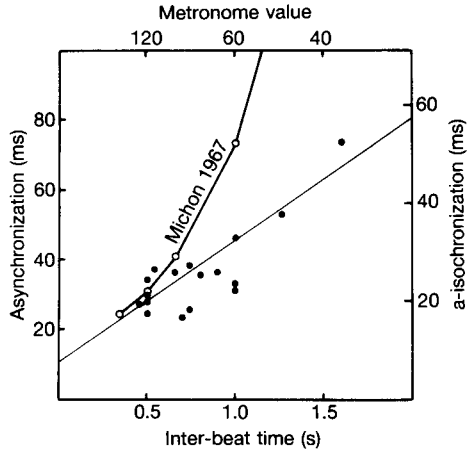



**Fig. 4.5.** Relation between tempo and asynchronization. Inter-beat time is used as a measure of tempo. The vertical scales indicate asynchronization, defined on the left as the standard deviation of onset time differences, averaged over instruments. On the right, a-isochronization is indicated. The dots are data points for movements. The regression line of tempo on asychronization is based on the correlation between inter-beat time and asynchronization. The results of Michon's experiments (1967, p. 31) are inserted for comparison.

tempo goes with more synchronization, and slower tempo goes with less synchronization.

As already noted, asynchronization of simultaneous onsets can be the cause of a-isochronization, that is, the existence of deviations from equal durations of successive tones within single voices. It is possible to make an estimate of this a-isochronization from asynchronization. For this estimate the correlation between subsequent relative onset times must be known. This correlation is the autocorrelation of a series of relative onset times. The autocorrelations were first calculated per composition and then algebraically averaged over voices and compositions. The resulting value proved to be 0.21. This indicates that there is a slight tendency of relative onset times to be preserved in the subsequent onsets. With $r = 0.21$, we find a ratio of 0.73 between a-isochronization and asynchronization. Apart from the autocorrelation, the asynchronization patterns seem to be more or less random.

Now we can compare our results with those of experiments conducted by Michon (1967). His subjects had to tap synchronously with a pulse train presented through headphones. His results (p. 31, Table 2B) are standard deviations of time intervals meant to be equal, and as such indicate a-isochronization. Tapping rates were 60, 90, 120, and 180 beats per minute. Figure 4.5 includes both the results of Michon's experiments and the

regression line for asynchronization to tempo. For high tempos the correspondence between Michon's data and ours is quite good. For lower tempos the synchronization found in our recordings is better than would be expected from Michon's data. A likely explanation of this discrepancy is that subdivisions of the beat also can contribute to synchronization.

## Perceptual effects of synchronization

In a study by Rasch (1977) the effect of onset difference times on the perception of quasi-simultaneous tones was investigated. Each trial included two stimuli in those experiments. Each stimulus contained two simultaneous tones. The lower tones of the stimuli in a trial had the same fundamental frequency, the upper ones differed in fundamental frequency by an (upward or downward) fifth (or other interval). The threshold for perceiving the upper tone of the stimulus was defined as the level resulting in a 75 per cent correct score in a 2AFC paradigm of the judgement of the direction of the pitch jump. For synchronous notes (onset difference time 0 ms) the threshold was between 0 and −20 dB depending on the frequency ratio and phase relation of the simultaneous tones. The threshold could be decreased drastically, down to about −60 dB, by introducing an onset difference time of, say, 30 ms. In the latter case, the threshold was largely independent of factors other than temporal ones. This condition can be compared directly with asynchronization in performed music. In the experiments there was no sensation of order. In supra-threshold conditions asynchronization contributes to the apparent transparency of the compound sound multi-tone stimuli.

When listening to music, onset differences mostly go unnoticed. Actually, the performances of professional ensembles give the impression of perfect synchronization. In the now classic investigation by Hirsh (1959) it was found that, roughly speaking, a time interval of 20 ms between two sounds (tones, noises, clicks) was sufficient for judging which one came first. Several reasons can be mentioned why onset difference times that are considerably longer are not perceived in a musical listening situation. Firstly, the qualities of the simultaneous tones differ, especially as to pitch and timbre (instrument); this impairs an exact judgement of the temporal structures. Secondly, rise times of musical tones are much longer (for example 20–80 ms) than the stimuli used in psychophysical experiments (often 5 or 10 ms). The smoother rise curves also impair exact judgements. Thirdly, the distribution of onset differences is practically random. Fourthly, attention of the listener is not directed to onset differences between voices ('vertical') but rather to melody lines within one voice ('horizontal'). The musicians themselves are, in general, also unaware of the amount of asynchronization, except when unwanted asynchronization

cannot be avoided in very difficult passages or in cases of 'temporal mistakes'.

Among the onset difference times there appeared to be a small number of relatively large ones, that is, from 100 to 200 ms. They were clearly audible during repeated replay. They can be a consequence of mistakes by the performer, but they are also induced by some musical characteristics, such as: the first onsets of a movement, a section with long notes in one voice and short notes in other voices; ritenuto or accelerando sections; onsets of individual part after rests; final notes following a ritardando or separated from the preceding notes by a rest; sections that are complicated with regard to rhythm and/or metre. Underlying factors for these harder-to-synchronize fragments are the absence of directly preceding tones or uncertainty concerning the temporal structure. The threshold for order of auditory stimuli sets a lower limit for synchronization. As yet, there are no experimental data for thresholds of temporal order in musical situations.

As a final remark, mention must be made of Seashore's theory that 'beauty in music largely lies in the artistic deviation from the exact or rigid' (1938, p. 249). This postulate leaves open the question of how to define artistic deviation. However, its negative formulation ('there is no beauty in music if the performers adhere exactly and rigidly to the instructions of the score') is certainly true. In every artistic musical performance there is constant deviation from what is prescribed exactly in the score, as to time, frequency, duration, level, etc. These deviations have strong statistical components without being purely random. They are of primary importance for the 'live' character of music performed by human beings. The asynchronization of simultaneous tones should be regarded as one of the vital deviations in the performance of music.

## The moment of synchronization

In the preceding section, the decisive time moment of the tone, used as a basis for the calculation of synchronization, was taken where the temporal envelope of the tones passed a threshold, quasi-arbitrarily defined as the level 15 or 20 dB below the average maximum level of the respective voice part. Since this choice is fairly central in the description of synchronization behaviour, the threshold model concerning the perceptual onset of musical tones will be considered here in some detail. The model assumes that music listeners (including music performers) base their timing on the passing of the temporal envelope through a certain threshold level, which is defined in relation to the average maximum level of the stream of musical tones (Vos and Rasch 1981). The physical temporal envelope of a musical tone can be roughly divided into three successive portions: the rise, the steady-state, and the decay portions (see Fig. 4.6).

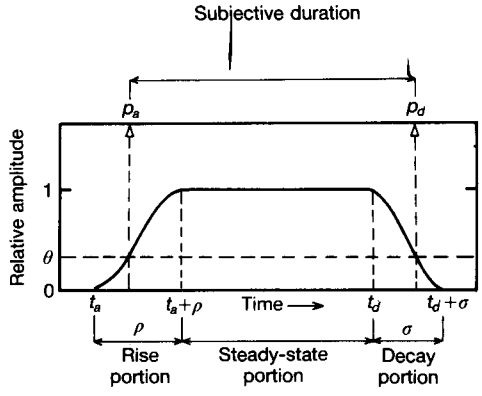



**Fig. 4.6.** Temporal envelope of a musical tone, divided into three successive portions. The perceptual onset ($p_a$) and offset ($p_d$) of the tone are defined as the moment at which the temporal envelope passes the relative threshold amplitude, $\theta$.

If the rise function, which describes the envelope during the rise, is monotonously increasing, then the inverse function is unambiguous. Throughout this chapter, we will regard rise functions as monotonously increasing functions. In our model, the perceptual onset of a tone is the moment at which the temporal envelope during the rise portion passes a certain relative threshold level. If the moment of the perceptual onset is known, the threshold level can be calculated. If the threshold level is known, the perceptual onset can be calculated. The subjective duration of a tone is defined as the time interval between the perceptual onset and offset of a tone. The following paragraphs, describing an extension of our model, will deal only with the perceptual onset.

The model can be extended to groups of tones, either simultaneous or successive ones. Two tones are called *perceptually synchronous* when their perceptual onsets coincide in time (see Fig. 4.7).

Tone sequences are defined as *perceptually isochronous* if the time intervals between successive perceptual onsets are all equal to each other. Figure 4.8(c) shows the temporal envelopes of the tones A, B, and A', which are perceptually isochronous. An experimental paradigm with successive perceptually isochronous tones with different rise functions and/or rise times can be used to determine the threshold amplitude for the perceptual onset (Vos and Rasch 1981). If tones are physically isochronous, that is, when the time intervals between successive physical onsets are all equal to each other but have different rise times and/or functions, the perceptual onsets will not, as a rule, be isochronous.

In the experiments exploring the applicability of the model a paradigm was used in which a sequence of tones had to be isochronized, that is, the tones had to be adjusted in such a way that the onsets were perceived isochronously. Each trial started with a tone sequence that was decidedly nonisochronous. The starting sequence consisted of successive pairs of

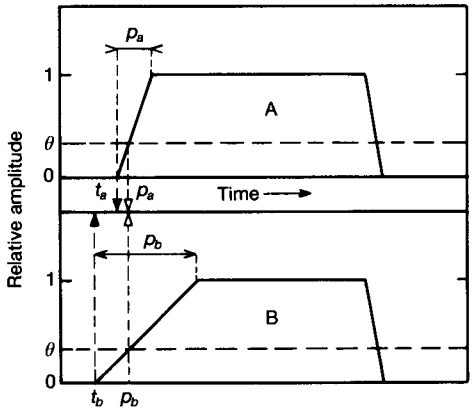



**Fig. 4.7.** Temporal structure of a perceptually synchronous stimulus, containing tones A and B with physical onsets $t_a$ and $t_b$, rise times $p_a$ and $p_b$, and maximum amplitude 1. The perceptual onsets, $p_a$ and $p_b$, coincide in time and are by definition located at the moment at which the temporal envelopes pass threshold amplitude, $\theta$.

tones, A and B, with a different interval between the physical onsets of A and the following B and between B and the following A (see Fig. 4.8(a)). The onset times of tones B, relative to those of A, could be adjusted by the subjects, by turning a knob. The experimental task was to adjust the onset times of tones B in such a way that the sequence ABABAB . . . was perceptually isochronous, that is, that the perceived onsets of the tones followed each other after strictly the same time interval. This is illustrated in Fig. 4.8(b). Because the tones A were repeated every 800 ms, the subjective repetition time $t$ of the tones in the entirely isochronized sequence is 400 ms. Rise times were varied independently. The time between the successive physical onsets was derived from the position of the turning knob at which the subject judged the tone sequence to be isochronous. Now all variables that are necessary for computing the threshold amplitude for the perceptual onset were known.

Relative threshold amplitudes were computed for a number of combinations of unequal rise times. Mean threshold level was $-15.4$ dB, and the standard deviation equalled 1.6 dB. Considering this consistent result—that is, that about the *same* relative threshold level was found in ten physically different conditions—it is justified, for the time being, to define the perceptual onset moment of a tone as the time at which its envelope passes a certain threshold level. Moreover, from these experimental data, this threshold level can be estimated as $-15$ dB, relative to the maximum level of the tones.

The perceptual onset threshold has up to now been given relative to the

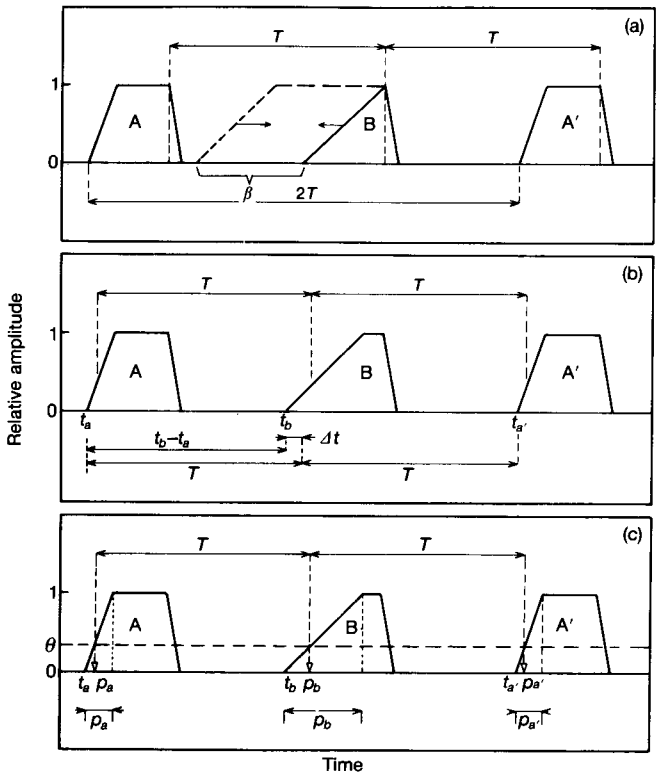



**Fig 4.8.** (a) Illustration of a perceptually nonisochronous starting sequence. The physical onsets of tones B could be adjusted by the subject within time interval $\beta$. At the start of a trial, the physical onset of tone B was either at the beginning or at the end of $\beta$. The physical offset of tones A and B, were fixed. (b) Temporal structure of a perceptually isochronous tone sequence in which the time intervals between successive perceptual onsets are all equal to each other. The physical onsets of tones A and B are $t_a$ and $t_b$, respectively. The repetition time of tone A equals $2T$, so that $T$ is the perceptual repetition time. The dependent variable is denoted by the time interval $\Delta t = T-(t_b-t_a)$. (c) Temporal structure of tones A, B, and A', which are perceptually isochronous. The time interval between successive perceptual onsets is denoted by $T$. The perceptual onsets $p_a$, $p_b$, and $p_a'$ are defined as the moment at which the temporal envelopes pass relative threshold amplitude, $\theta$.

maximum level of the tones. But it is evident that the threshold can also be defined as a level above background noise or hearing threshold. In short, the following question can be asked: Is the threshold fixed, with respect to maximum level, to background, or to some other criterion? Our first experiments were designed to test the threshold hypothesis in general, not

to discriminate between these alternatives. Further experiments were carried out to determine the reference relative to which the threshold had to be defined. This was done by varying the maximum levels of the tones. The stimuli were the same as those in the earlier experiments. The sound-pressure level of the tones, however, was 77 dB(A) in the highest level condition; in the other conditions it was 57 and 37 dB(A). Again, relative threshold amplitudes were computed. In the 37-dB condition, the mean relative threshold was −8.0 dB and the standard deviation equalled 0.9 dB. In the 57- and 77-dB conditions, the mean thresholds were −11.7 and −13.0 dB, and the standard deviations were 0.7 and 1.6 dB, respectively.

The results of these experiments show that the time difference between physical and perceptual onsets increases with decreasing tone intensity, that this increase can be described as an upward shift in the relative threshold by which the perceptual onset is determined, and that the shift in threshold is small relative to the shift in stimulus level. Therefore, the threshold can be most conveniently described relative to the maximum level of the stimulus.

From the experiments, it may be concluded that the perceptual onsets of musical tones can be defined as the times at which the envelopes pass a *relative* threshold of about 6 to 15 dB below the maximum level of the tones. In a number of experiments, we have shown that the level of the relative threshold depends on the tone level above masked or absolute threshold. The data from the present experiments seem to suggest adaptation to a certain constant stimulus level. At the time the adaptation threshold is passed by the stimulus level presented, the onset of the stimulus is perceived.

In studies on the temporal structure of performed music, our threshold model can be applied to determine the perceptual onsets of musical tones. When music is performed on instruments with very short rise times, like the piano, harpsichord, and drums (Gabrielsson 1974; Povel 1977; Sundberg and Verrillo 1980), the difference between the physical onset and the perceptual onset is very small. In these cases, level above threshold, too, does not have a great impact on this difference. However, when ensemble music is performed on instruments producing tones with relatively long rise times, such as bowed string instruments, the perceptual onset heavily depends on the relative threshold.

In such musical practice in which dynamic differences are not very large, perceptual onset is clearly affected only by the rise times and rise functions of the different instruments. This, however, is a variable with which the respective musicians can cope by adjusting their physical onset times in order to establish the appropriate timing of the perceptual onsets of their tones.

Future research should be focused on the perceptual onset of musical tones in synchronously perceived tone pairs. The sensation levels of

simultaneously presented tones, especially, are dependent on the amount of auditory masking (Zwislocki 1978). To apply our model to simultaneously produced tones, experimental results of binaural masking experiments with complex tones are needed. In addition, it would be interesting to see if our model also works in cases of complex tones consisting of partials with unequal rise times and unequal physical onsets (Freedman 1967; Grey and Moorer 1977) and of tones with substantially differing amplitude envelopes (Strong and Clark 1967).

## Synchronization and the conductor

Small ensembles usually perform without a conductor, while larger ensembles and orchestras are guided by the gestures of a conductor. Since one of the tasks of the conductor is to keep the performing musicians synchronized, it is interesting to comment upon the present or absence of a conductor in terms of our descriptive model of synchronization.

When there is no conductor, the musicians will have to synchronize upon each others' sound outputs. For this condition, we consider the standard deviation of the onset time differences as being of some fixed, constant magnitude. In that case, the standard deviation of the relative onset times varies as a function of the number of performers in such a way that it becomes larger when that number is increasing (since $s_v^2 = (\frac{1}{2}(n-1)/n)s_d^2$). If we use our mean value of the standard deviation of onset time differences (36 ms), then the standard deviation of relative onset times is 18 ms for an ensemble of two players, 20.8 ms for three players, 22 for four, 22.8 for five, 23.2 for six, 24 for nine. For ensembles with more than nine players it converges to its maximum value of $36/2\frac{1}{2}$ or 25.5 ms. Evidently, synchronization quality is less in larger ensembles.

When a conductor is leading an ensemble or an orchestra, it may be supposed that the performers will try to synchronize not with each other, but with the movements of the conductor. For that condition, it may be supposed that the standard deviation of the relative onset times will be of some fixed, constant magnitude, since all musicians have the same task in this respect. We cannot extract an estimate for this quantity from our experiments, but via a detour a guess can be made. When the standard deviation of relative onset times is a constant quantity, the standard deviation of onset time differences is diminishing when a number of performers is increasing (since $s_d^2 = (2n/(n-1))s_v^2$), or, put the other way around, the standard deviation of onset time differences is increasing when the number of performers is increasing. With this relationship in mind, it is easy to understand that a large orchestra will always need a conductor, or, at least, will do better with a conductor. If the ensemble shrinks, however,

synchronization quality is diminishing when the conductor stays in charge of the synchronization.

A survey of recorded chamber music with five to 15 performances showed that ensembles with up to and including nine performers did, more often than not, without a conductor, while ensembles with 10 or more players almost always had a conductor (see Fig. 4.9). This must mean that the standard deviation of relative onset times of an imaginary ensemble of 9.5 players marks, on the average, the maximum allowable value of that standard deviation (in our average case 24.1 ms). It may be assumed that in this imaginary situation synchronization is as good with conductor as without, which gives us at the same time the assumed constant value of the standard devation of relative onset times in a situation with conductor. When there are those 9.5 performers, the standard deviation of onset time differences (or the asynchronization) is just 36 ms. But when there were only two performers, asynchronization would have been 48 ms; with three, 42 ms; with four, 39 ms; etc. When the number of performers gets substantially larger than 10, asynchronization converges to its limit of 34.1 ms.

Our line of reasoning concerning the usefulness of a conductor when it comes to synchronization is summarized in Fig. 4.10. It amounts to the

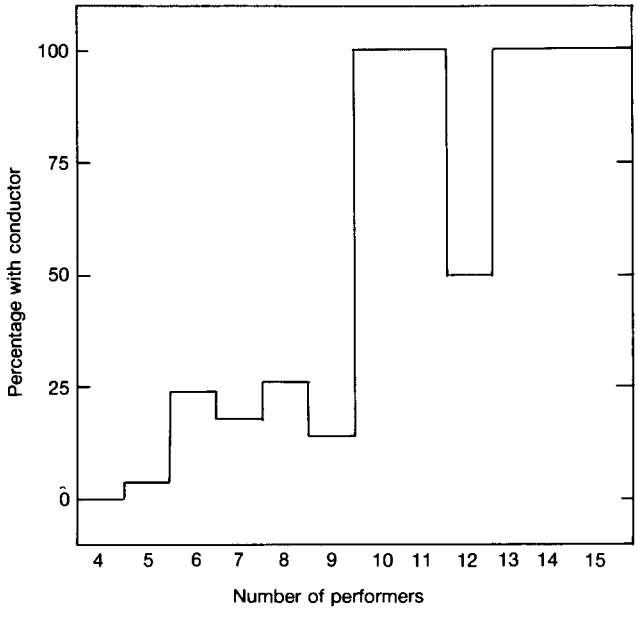



**Fig. 4.9.** Percentage of performances with a conductor, taken from a representative set of recorded quartets, quintets, sextets, etc. Only two compositions for 12 players could be located, of which one was with, one without conductor.

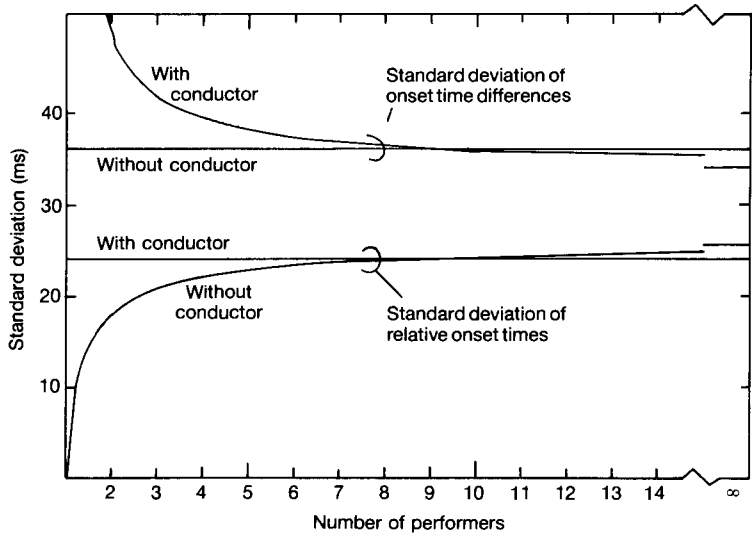



**Fig. 4.10.** Idealized standard deviations of relative onset times and onset times differences of performances with and without conductor, as a function of the number of performers. Numerical values are anchored to a mean asynchronization of 36 ms.

conclusion that below a certain cut-off value of the number of performers synchronization among performers is better without a conductor, while above that value it is better with a conductor. A caveat must be added, however. The lines marked 'with' and 'without' (conductor) cross at rather small angles, so that the crossing points are relatively strongly influenced by small up or down movements of the lines, which makes the choice of the crossing open to serious discussion. We found differences in synchronization related to the instruments of the ensemble and related to the tempo of the movement performed. The kind of music performed may also be expected to affect the value of a conductor. All these effects may cause upward or downward shifts of the lines inserted in Fig. 4.10. Indeed, the cut-off between the numbers of performers that need and that do not need a conductor is not sharply defined. Our cut-off value of 9.5 performers represents only an average value, although it is much sharper than was to be expected from the shallow curve of the standard deviation of the relative onset times as a function of the number of performers.

## Conclusion

In this chapter, we have presented some thoughts and data about timing and synchronization in ensemble performance. The position taken in this

chapter has to do with model making and acoustical measurement rather than with processes of the psychology of music. However, before any such psychological interpretation comes into play, a sound and reliable way of measuring the relevant characteristics of musical behaviour should exist, and that is what this chapter aims to provide.

Measures of musical behaviour, including musical performance, are essentially statistical measures. For this reason we developed our statistical model of synchronization, based on relations between variances of relative time moments. We showed that with this framework it is possible to describe in a sensible way synchronization in ensemble performance. The research we have done was not initiated by the wish to uncover the psychological processes that are behind synchronization behaviour. Eventually, however, the descriptive part should be supplemented by a psychological, interpretative part.

## References


Berger, K. (1964). Some factors in the recognition of timbre. *Journal of the Acoustical Society of America* **36**, 1888.

Clark, M., *et al.* (1963). Preliminary experiments on the aural significance of parts on tones of orchestral instruments and on choral tones. *Journal of the Audio Engineering Society* **11**, 45.

Freedman, M. D. (1967). Analysis of musical instrument tones. *Journal of the Acoustical Society of America* **41**, 793–806.

Gabrielsson, A. (1974). Performance of rhythm patterns. *Scandinavian Journal of Psychology* **15**, 63–72.

Grey, J. M. and Moorer, J. A. (1977). Perceptual evaluations of synthesized musical instrument tones. *Journal of the Acoustical Society of America* **62**, 454–62.

Hirsh, I. J. (1959). Auditory perception of order. *Journal of the Acoustical Society of America* **31**, 759.

Michon, J. A. (1967). *Timing in temporal tracking.* Institute for Perception TNO, Soesterberg.

Moorer, J. A. (1975). On the segmentation and analysis of continuous musical sound by digital computer. Report STAN-M-3, Center for Computer Research in Music and Acoustics, Stanford University, Stanford, Calif.

Povel, D. (1977). Temporal structure of performed music, some preliminary observations. *Acta Psychologica* **41**, 309–20.

Rasch, R. A. (1978). The perception of simultaneous notes such as in polyphonic music. *Acustica* **40**, 21.

Rasch, R. A. (1979). Synchronization in performed ensemble music. *Acustica* **43**, 121–31.

Rasch, R. A. (1981). Aspects of the perception and performance of polyphonic music. Unpublished thesis. Groningen.

Saldanha, E. and Corso, J. (1964). Timbre cues and the identification of musical instruments. *Journal of the Acoustical Society of America* **36**, 2021.

Seashore, C. E. (1938). *Psychology of music.* McGraw-Hill, New York. (Reissued in 1967 by Dover, New York).

Strong, W. and Clark, M. (1967). Synthesis of wind-instrument tones. *Journal of the Acoustical Society of America* **41**, 39–52.

Sundberg, J. and Verrillo, V. (1980). On the anatomy of the retard: a study of timing in music. *Journal of the Acoustical Society of America 68*, 772–9.

Tove, P. A., Ejdesjö, L., and Svärdström, A. (1967). Frequency and time analysis of polyphonic music. *Journal of the Acoustical Society of America* **41**, 1265.

Vos, J. and Rasch, R. (1981). The perceptual onset of musical tones. *Perception and Psychophysics* **29**, 323–35.

Zwislocki, J. J. (1978). Masking: experimental and theoretical aspects of simultaneous, forward, backward and central masking. In *Handbook of perception* (ed. E. Carterette and M. Friedman), Vol. 4. Academic Press, New York.

5

# Rehearsal skill and musical competence: does practice make perfect?*

LINDA M. GRUSON

In recent years, research on the acquisition of skills has focused upon the study of a variety of complex skills including, for example, competence in games such as chess (de Groot 1965; Chase and Simon 1973; Simon and Chase 1973), bridge (Charness 1979) and the board game, Go (Reitman 1976); in the academic domain, acquisition of skills in such subjects as mathematics (Krutetskii 1976), physics (Simon and Simon 1978; 1980; Chi *et al.* 1980; Larkin *et al.* in press), compositional writing (Flower and Hayes 1979), and engineering thermodynamics (Bhaskar and Simon 1977); and a number of cognitive-motor skills such as typewriting (Thomas and Jones 1970; Shaffer 1975, 1978) and, most pertinent to the present study, musical skills (Bean 1938; Weaver 1943; Wolf 1976; Shaffer 1980, 1981; Sloboda 1974, 1977; MacKenzie *et al.* 1983).

While the research paradigms and target behaviours studied have been quite diverse, one of the most consistent findings that emerges from the research on skills is that as an individual learns a skill, he or she acquires the ability to process increasingly larger and more complex units of meaningful, skill-related information. Chase and Simon's (1973) work with chess players illustrates this process of 'chunking'. They have demonstrated that with increasing mastery, chess players are able to reconstruct significantly more pieces of meaningful chess positions exposed briefly to them. The reconstruction of randomly placed pieces did not improve with increased mastery. Chase and Simon hypothesized that chess masters' superiority over less expert players was due to their ability to encode

* This study was carried out in partial fulfilment of the Ph.D. degree. An earlier version of this paper was presented at the eighty-ninth Annual Convention of the American Psychological Association, Los Angeles, California, August 1981

meaningful chess information more efficiently into 'chunks' in short-term memory.

The research on musical sight-reading lends support to the 'chunking' notion. Bean (1938) found that professional musicians could play almost five notes of a musical segment presented to them tachistoscopically while non-professional musicians could recall between two and three notes, and beginners, only one to two notes. He discovered that differences in perceptual span could be accounted for by differences in perceptual style; the majority of his professional pianists approached the task by reading musical patterns while most of the non-professinals read the music analytically, note-by-note. Weaver's (1943) finding that professional musicians could perceive between three and five notes at a single glance provided further evidence that proficient sight-readers can, in fact, process groups of notes as a single perceptual unit. Lannert and Ullman (1945) found that sight-reading skill was positively correlated with span of reproduction and ability to read ahead in the score, and Sloboda (1974) reported a correlation of 0.86 between sight-reading ability and the eye-hand span (the number of notes correctly played after the unexpected removal of the stimulus). In an interview study, Wolf (1976) discovered that good sight-readers report that they rely upon their recognition of familiar patterns and on their ability to organize the music into a few familiar patterns and dependable pattern cues. He also found that sight-reading appeared to be a very specific skill depending upon such a pattern-recognition style of information processing. The professional musicians who reported using an analytic strategy (note-by-note) as opposed to a pattern-recognition strategy were poor sight-readers although their ability to memorize was superior to that of expert sight-readers.

Chase and Simon (1973) attribute the skilled experts' superiority to their extensive information bank in long-term memory, developed through experience with the skill domain, which enables them to rapidly perceive and recognize familiar patterns. Thus, long-term memory is hypothesized to contain an internal representation of such familiar patterns together with information concerning appropriate associated actions. Furthermore, Chase and Chi (1980) propose that the information is organized hierarchically. A number of patterns may, therefore, be represented in a single familiar configuration in long-term memory. Experience is proposed as the process through which one's knowledge base and cognitive structure are built up. The importance of skill-related experience or practice in the acquisition of a skill was illustrated in a study by Chi (1978) who compared experienced and inexperienced chess players. In her study, there was an interesting twist; the expert players were children (mean age, 10.5 years) and the novice players were adults. As in the Chase and Simon study, players were required to reconstruct meaningful chess positions presented briefly to

them. They were also tested on a digit span task. Although adults were found to have superior recall for digits, the child experts' memory for chess positions was significantly superior. These results indicate that it is specifically skill-related practice that underlies expert performance rather than maturational factors or general, non-specific experience. It would not come as a surprise to music teachers to assert that 'practice is the major independent variable in the acquisition of skill' (Chase and Simon 1973). However, although practice is undoubtedly a crucial aspect of musical skills acquisition, the behaviour of music students as they practise remains an unexplored area.

In the domain of music, research has focused on sight-reading to investigate the acquisition of skills. While Shaffer (1981) has examined motor programming by investigating musical performance of professional pianists, little work has been directed to the acquisition process itself. The present study was designed to investigate the process of musical acquisition through an examination of the rehearsal behaviour of piano students varying in musical competence. The specific focus of the study was on practising strategies and how they change with increasing musical knowledge and expertise. In the first part of the study, the behaviour of 43 subjects from all music levels from Grade I to the Artist (Diploma) level, including three concert pianists, was compared as they practised three musical selections during a single practice session. In addition, the effect of increased familiarity with particular musical scores on practising behaviour was investigated in an attempt to discern the relative contributions to changes in rehearsal behaviour of musical competence acquired through long-term study and through short-term practice of specific materials. In the second part of the study, a subset of the players representing novice (Grade II students), intermediate (Grade VI students), and advanced (Artist; ARCT) levels of skill continued to practise their pieces for nine additional sessions (10 sessions in all). The pattern of practising behaviour across sessions was investigated in this part of the study. The study therefore used a cross-sectional approach, investigating performance across multiple trials. The behavioural domain under investigation was the subjects' natural practising of musical segments selected for their level of skill. Assessment consisted of two components, behavioural observations coded according to an observational rating scale, the Observational Scale for Piano Practising (OSPP), and self-report, tapped by means of a structured interview format. Of interest in the study were changes in the form and content of both behaviour and cognitive strategy as practising behaviour changed with increasing experience in music. Attention was specifically focused on how students chunked the music and how their chunking changed with musical experience, together with other changes that occurred concomitantly.

## Method

### Subjects

Subjects were 40 piano students taking individual instruction in piano, and three professional concert pianists. The students were graded according to the requirements of the University of Toronto's Royal Conservatory of Music, which provides grading standards for the region of South-Western Ontario. Subjects were defined as belonging to a particular level if they passed the Conservatory examinations for the previous level and were working on pieces for the designated grade. The subjects were selected so that three or four students from each musical level from Grade I to Grade X, and from the artist (ARCT) and professional levels, were included in the study. There were 33 females and 10 males ranging in age from six to 46 years and in experience from six months to 34 years, receiving instruction from one of 17 piano teachers. Details of the subjects' characteristics are presented in Table 5.1.

**Table 5.1** *Subject characteristics*

| *Grade Level* | *n* | *Sex* M | F | *Age* ($\bar{m}$ years) | *Experience* ($\bar{m}$ years) |
|---|---|---|---|---|---|
| I | 3 | 2 | 1 | 10.00 | 2.53 |
| II | 4 | 1 | 3 | 9.75 | 2.15 |
| III | 3 | 1 | 2 | 10.30 | 2.40 |
| IV | 4 | 1 | 3 | 10.50 | 3.25 |
| V | 3 | 0 | 4 | 12.33 | 6.00 |
| VI | 4 | 1 | 3 | 13.75 | 4.87 |
| VII | 4 | 1 | 3 | 14.50 | 7.65 |
| VIII | 4 | 0 | 4 | 15.25 | 8.20 |
| IX | 3 | 0 | 3 | 20.33 | 10.17 |
| X | 4 | 0 | 4 | 17.50 | 10.87 |
| ARCT | 4 | 1 | 3 | 23.75 | 17.00 |
| PROF | 3 | 2 | 1 | 39.33 | 30.67 |

### Materials

In order to present the subjects with music representing a variety of styles and challenges, music pieces were selected for each music level from three musical forms: a Baroque piece, a contemporary piece, and a sonatina. The selections were limited to approximately one page in length in order to maintain the practising demand for the students at a reasonable level.

Accurate grading of the music was important to ensure comparable levels of difficulty for the pieces across grade levels. In order to control for

famliarity, pieces were selected that would have been unlikely to have been played previously. The Conservatory syllabus and the ratings of experienced music teachers were used to select pieces appropriate to the players' grade level that were unlikely to have been practised before.

All subjects were queried concerning whether they had actually played the pieces before; this revealed that the pieces were unknown to all but three subjects. When students had played a piece before, it was excluded from the study for those subjects.

## Procedure

### *Part I*

All 43 players were instructed to practise the three experimental scores as they would usually practise new pieces, and the practice sessions were recorded on cassette audio tape and later recorded according to the Observational Scale for Piano Practising described below. The very high reliability (Pearson $r = 0.97$) between *in vivo* observations and observations taken from audio tapes obtained in a previous study (Gruson 1978) confirmed the validity of the audio-taping procedure.

### *Part II*

A subgroup of the players, four students in each of Grade II, Grade VI, and the Artist (ARCT) levels, were selected to represent novice, intermediate, and advanced levels of music proficiency. This subgroup of students was instructed to tape-record nine additional practice sessions (10 sessions in all). To represent a number of points in the acquisition process, the first, fourth, seventh, and tenth practice sessions were coded and analysed.

#### *Observational Scale for Piano Practising (OSPP)*

The Observational Scale for Piano Practising (OSPP) consists of the 20 categories defined in Table 5.2. The scale tapped length of uninterrupted playing, errors, repetitions of various unit lengths (notes, measures, larger sections), tempo changes, self-guiding speech, breaking down the piece into separate hands, giving up, playing music other than the designated piece, intervention of another person (for instance, parent), and a number of measures of practice time.

The format for recording required the rater to note, in its natural sequence, each incidence of a behavioural category as it occurred. To aid in this task, a recording sheet was constructed which divided the practice sessions into 5-s intervals. A stopwatch was used to mark the beginning of each 5-s interval. Frequency measures were employed for the majority of behaviours recorded since they tended to be discrete, momentary events. For measures that were more continuous (for example count aloud, hands

separately, and uninterrupted playing), the behaviour was scored for each 5-s interval in which it occurred.

To assess inter-rater reliability, a randomly selected sample from 17 practice sessions recorded from 15 subjects was chosen. The 15 subjects represented in the sample were randomly selected with respect to experience level, age, and sex. In addition, a random selection was made from all practice sessions using the experimental stimulus pieces, from the first to the tenth session. A sample of 3174 behavioural codes recorded over a period of 185.8 min was rated by two experienced observers from audio-tapes using the OSPP. The Kappa coefficients ($K$) for each behavioural category ranged from 0.96 to 1.00, all of which were highly significant and confirm that the OSPP may be considered to have a high level of inter-observer reliability.

**Table 5.2** *Observational Scale for Piano Practising (OSPP)*

| *Symbol* | *Name* | *Definition* |
|---|---|---|
| Unint | Playing without interruption | S plays the piece for 5-s interval during which no other behavioural categories occur. Score for each 5-s interval S plays uninterrupted. |
| Error | Error | S plays incorrect note. |
| *Repetition* | | |
| RNote | Repeat Note | S repeats a note on which an error may or may not have occurred with or without correction. |
| RMeasure | Repeat Measure | S repeats a measure, or a part of a measure, in which an error may or may not have occurred with or without correction. |
| RSection | Repeat Section | S repeats a section of the piece longer than a measure in which an error may or may not have occurred with or without correction. |
| RPiece | Repeat Piece | S repeats whole piece from the beginning after completing previous practice of entire piece. If S goes to the beginning of piece after playing only part of the piece, score as RSection. |
| *Tempo changes* | | |
| Slow | Slow down | S slows the musical tempo beyond the range permitted by the music. |

| | | |
|---|---|---|
| Pause | Pause | S stops playing for at least 2 s in a 5-s interval. Score for each 5-s interval. |
| Fast | Speed up | S increases the musical tempo from the normal tempo for a part of the piece. Score for each 5-s interval. |
| *Verbalizations* | | |
| Guide Sp | Self-guiding statements | S directs his/her playing verbally, e.g. 'Slow down', 'I made a mistake, I'd better try again'. |
| Read | Read Aloud | S reads notes aloud before playing them. |
| Count | Count Aloud | S counts aloud to aid timing. Scored for each 5-s interval. |
| Frust | Frustration | S verbalizes or emits signs of frustration, e.g. sighs, grimaces, groans, frowns, complaints, negative self-statements. |
| Hands Sep | Hands Separately | S plays a piece or a section of a piece with one hand only. The onset only is scored by H. The offset is marked by TOG (indicating coordinated hands) or by the end of the piece (/). |
| Give Up | Gives up | S fails to complete the piece. This behaviour is followed by a new piece or the end of the session. This category would also be scored if S attempts only a part of the piece and does not play piece through to completion. |
| Play other | Playing other than the designated piece. | S plays something other than the music in the score. |
| Person | Intervention by another | Other (usually a parent) intervenes in the practice, offering assistance or direction. |
| TTime | Total time | The duration of time spent in practising, calculated as the sum of time spent practising each piece. |
| Time | Mean time/piece | $\frac{\text{Total time}}{\text{No. of pieces practised}}$ |
| Tempo | Tempo | No. of notes/s of first practice in each session, averaged over pieces. This value was calculated from the first stretch of *unint* of at least 10-s duration. |

*Interview*

An additional concern of the present study was the subjects' knowledge and understanding of their own practising behaviour. In order to assess the subjects' cognitions concerning their piano practising, all subjects in Part II and the three concert pianists were given an interview concerning their rehearsal behaviour following the tenth or last practice session. The interview included the following six questions:

(1) How did you go about practising the pieces?
(2) What did you pay attention to?
(3) How did your practising change as you learned the pieces?
(4) What did you do when you made an error?
(5) What advice would you have for another student at your level learning these pieces?
(6) Is there anything that you should have done when you practised that you know that you did not do?

## Results

### Part I

The rate of OSPP behaviours for the Part I data is presented in Fig. 5.1. Across all subjects, the most frequently occurring event was uninterrupted

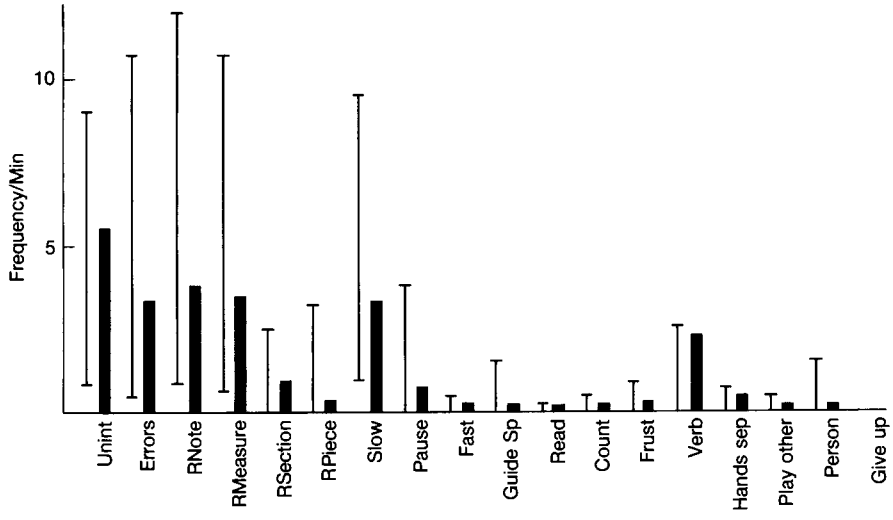



**Fig. 5.1.** Mean rate and range of practising behaviours per min: Session 1 (n = 43).

playing. The repetition of shorter units was the next most frequent type of event, repeating a single note accounting for 16.9 per cent of behaviours, and repeating measures, 16.7 per cent, while repeating sections accounted for only 2.6 per cent and repeating pieces, only 1.3 per cent. Slowing down was a frequently occurring event (16.0 per cent) while other tempo changes occurred less frequently. Errors also occurred frequently (14.3 per cent). The remaining behaviours occurred rarely, each accounting for less than 1 per cent of the total number of behaviours, although the sum of all self-verbalizations totalled 1.6 per cent of the total behaviours. In sum, students generally spend a quarter of their initial practice sessions in uninterrupted playing. When the flow of playing is broken, it is usually interrupted by errors, repetition (usually of short units), tempo reductions, pauses, and self-verbalizations.

But do all students practise alike? In order to investigate differences that may occur with increasing musical experience, the correlations between practising behaviours and skill level may be examined in Table 5.4. While it is a relatively infrequent event, repeating sections was the variable most significantly correlated with music level. Indeed, music level accounted for over 50 per cent of the variance in repeating sections. Music level

**Table 5.3** *Correlations of OSPP categories with music level:* Study 1*

| Category | *r* |
|---|---|
| Unint | 0.22 |
| Error | −0.31† |
| RNote | −0.31† |
| RMeasure | 0.17 |
| RSection | 0.72†† |
| RPiece | −0.21 |
| Slow | −0.21 |
| Fast | 0.23 |
| Pause | −0.31†† |
| Guide Sp | 0.39† |
| Read | −0.05† |
| Count | 0.20 |
| Frust | 0.14 |
| Hands Sep | 0.49†† |
| Play Other | 0.07 |
| Person | −0.26 |
| Give Up | — |
| Verb | 0.37† |
| Time | 0.40† |
| Tempo | 0.28 |

* Pearson product-moment correlation coefficient (*r*)
† $p<0.05$
†† $p<0.001$

accounted for 24 per cent of the variance in playing hands separately and between 10 and 16 per cent for each of errors, repeating notes, pauses, verbalizations, and, specifically, self-guiding speech, playing hands separately, and time. Higher music levels were also associated with reduced frequency of errors, repeating notes, and pauses.

Another way of looking at the relationship between competence and practising behaviour is to use the practising behaviour itself to discriminate between students of varying levels of competence. A discriminant functions analysis was carried out in order to determine which combination of practising behaviours could most parsimoniously and accurately discriminate students of differing skill levels. An examination of the 43 subjects suggested three groups in terms of the way music educators conceptualize music study. The three groups were divided by points that reflect certain watersheds in the career of a music student. Basic music training takes place in the first eight grades and students successfully completing Grade VIII receive a high-school credit for music. Senior music study begins in Grade IX, and the more difficult study of harmony is included in the programme at this level. Thus, the first cutpoint for the discriminant analysis was set between Grades VIII and IX. The second cutpoint was placed between the ARCT level and the concert pianist level, therefore distinguishing senior students from professional musicians. The three groups were, therefore, Grades I to VIII, Grade IX to ARCT, and professional concert pianists.

In order to arrive at the most parsimonious combination of variables that would successfully discriminate the three groups, a series of discriminant analyses was carried out, successively reducing the number of variables until the shortest, most discriminating series of variables was derived.

The analysis yielded 16 variables which in combination accurately classified all the 43 students into the three groups. A jack-knifed classification (which indicates the classification of a particular case into a group according to a function computed from all the data excluding that particular case) yielded an 88.4 per cent accuracy rate. The discriminant analysis indicated that the number of variables that emerged from the correlational analysis as relating to music level, such as repeating sections, pauses, playing hands separately, and time, were in fact important in discriminating the practising behaviour of students varying in level of skill. Repeating sections, particularly, emerged as an important discriminating variable and in order investigate the proportion of subjects that could be classified using only this single variable as discriminator, a discriminant function was computed including only repeating sections. The function yielded a 79.1 per cent success rate indicating that almost four-fifths of the subjects could be classified according to their use of this single behaviour.

In *summary*, the analysis of the data from the initial session revealed that

a number of changes in practising behaviour occurred as students gained musical skill. Errors, repeated notes, and pauses tended to decrease with competence while self-guiding speech, total verbalizations, playing hands separately, time spent on each piece, and, particularly, repeating sections (0.18 per min for Grades I to VIII, 0.75 per min for Grades IX to ARCT, and 2.40 per min for professionals) increased as music level increased.

**Part II**

Let us now take a closer look at the subset of Grade II, Grade VI, and ARCT students, how their practising differed and how it changed across practice sessions. Sessions 1, 4, 7, and 10 were selected as samples of the practising process. The descriptive data are presented in Fig. 5.2.

A correlational analysis was first carried out between practising behaviour and music level for each of the four sessions sampled. The results are presented in Table 5.6. In the initial session of Study 2, grade level accounted for the most significant proportion of variance in repeating sections (77 per cent). Repeating sections continued to increase with skill level across all sessions (0.39 per min in Session 1 to 0.62 per min in Session 4). While playing hands separately and time also increased significantly as music level increased in the first session, these variables were not significantly correlated with skill in the later sessions. A number of other behaviours were related to experience at various points in the practising process. Specifically, uninterrupted play in Sessions 4 and 10 increased as music level increased; slowing down and repeating notes in Session 4, and tempo in Session 10, decreased as skill increased.

Another way of assessing changes in the relationship between practising behaviour and music level across practice sessions is by means of discriminant analyses repeated for the data from each of the four sessions examined. The discriminant analyses were used to investigate the stability of the variables discriminating between the three levels (Grade II, Grade VI, and ARCT) across sessions.

The three groups of subjects were successfully discriminated by three variables. For sessions 1, 4, and 7, the discriminating variables were identical: repeating sections, playing hands separately and pausing. In the final (tenth) session, playing hands separately was no longer found to be discriminating and uninterrupted play entered the function as the third discriminator. In all sessions, repeating sections was the most discriminating variable. Indeed, as the sole variable in the discriminant function, repeating sections was able to correctly classify 81.8 per cent of subjects in Session 1 and 72.7 per cent in each of Sessions 4, 7, and 10.

In order to investigate more directly the relationship between practising behaviour and sessions, a correlational analysis was carried out on all

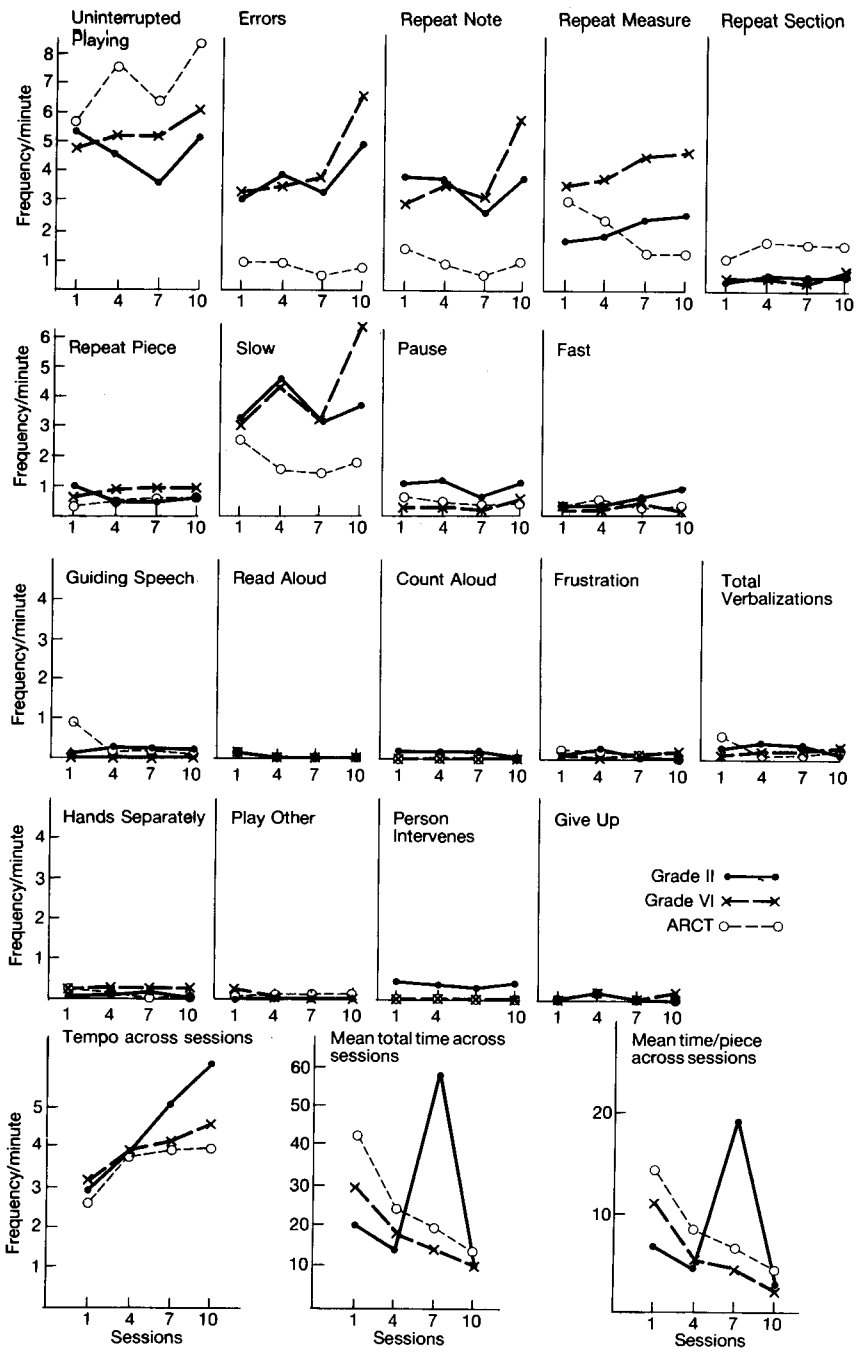



**Fig. 5.2.** Descriptive data for each level across sessions.

**Table 5.4** *Correlations of OSPP categories with music level:* Study 2, Sessions 1, 4, 7, and 10.*

| Category | Session 1 | Session 4 | Session 7 | Session 10 |
|---|---|---|---|---|
| Unint | 0.00 | 0.63† | 0.36 | 0.69† |
| Error | −0.54 | −0.72 | −0.53 | −0.35 |
| RNote | −0.44 | −0.61† | −0.41 | 0.25 |
| RMeasure | 0.54 | 0.10 | −0.16 | 0.14 |
| RSection | 0.88†† | 0.72† | 0.72† | 0.79† |
| RPiece | −0.36 | −0.02 | 0.13 | 0.03 |
| Slow | −0.21 | −0.68† | −0.48 | −0.16 |
| Fast | 0.50 | 0.30 | −0.30 | −0.34 |
| Pause | −0.30 | −0.55 | −0.47 | −0.50 |
| Guidesp | 0.43 | −0.44 | −0.35 | −0.51 |
| Read | −0.35 | — | −0.35 | — |
| Count | −0.25 | −0.35 | −0.26 | 0.35 |
| Frust | 0.28 | −0.32 | 0.01 | 0.12 |
| Hands Sep | 0.62† | 0.09 | −0.05 | 0.15 |
| Play Other | 0.01 | 0.45 | — | 0.45 |
| Person | −0.35 | −0.35 | 0.35 | −0.35 |
| Give Up | — | −0.35 | — | 0.01 |
| Verb | 0.37 | −0.23 | −0.31 | 0.04 |
| Time | 0.71† | 0.48 | −0.27 | 0.25 |
| Tempo | −0.19 | −0.13 | −0.55 | −0.67† |

* Pearson product-moment correlation coefficient ($r$)
† $p<0.05$
†† $p<0.001$

variables for each group and a one-way repeated measures analysis of variance was performed.

A number of significant correlations emerged although they were not found consistently for all three groups. Across sessions, repeating measures increased for Grade II (1.62 per min in Session 1 to 2.50 in Session 10) and Grade VI (3.55 per min in Session 1 to 4.69 in Session 10) students, repeating pieces increased for ARCT students (0.20 to 0.38), and speeding up increased only for the Grade II group (0.03 to 0.45). Repeating sections increased across sessions for all groups but not to a significant degree.

Analyses of variance were carried out involving a one-way repeated measures design for each behavioural category with music level as the independent variable and session number as the repeated variable. Significant main effects of music level were found for repeating sections, pausing, and verbalizations. These results support the findings of the correlational analyses that as music level increased, repeating sections increased. The correlational analysis for pauses and verbalizations produced no significant correlations although the discriminant functions analysis indicated that pauses could be helpful in discriminating between groups.

The descriptive data indicated a decrease in pauses with increasing skill. Looking at the data for repeated sessions, only tempo yielded a significant main effect across sessions and this variable also produced a significant interaction between music level and session number. The ANOVA and descriptive data indicate that while there were no large, consistent differences in tempo across the three subject groups, the Grade II students increased their tempo considerably more rapidly over sessions than did the more experienced groups. In addition, while the tempo maintained by the Grade II and Grade VI students increased steadily as session number increased, the tempo of the ARCT students had reached an asymptotic level before the seventh session. Thus, the less experienced students spent all 10 sessions increasing their tempo while the most experienced students mastered the pieces' tempo early in practising and were free to focus on other aspects. Self-reports indicated that the ARCT students knew the pieces fairly well after a few sessions and spent the later sessions experimenting with the pieces or memorizing them.

The impact of increased competence and practise experience may also be investigated on a more microscopic level by analysing the sequences of behaviours within practice sessions. In addition to the above analyses, a sequence analysis was carried out in which the behaviours following errors were examined. The analysis indicated that a repeated note was the most common event following an error, although for the more senior artist (ARCT) group, repeated measures increased in frequency across sessions. The category Repeating sections occurred rarely. The subjects' self-reports indicated that repeating sections was an event that would occur later in the response chain, following the repetition of shorter units.

In *summary* then, the results of Part II did not support the hypothesis that consistent changes in practising behaviours occurred with increasing practise of particular music pieces. The most salient finding was the impressive consistency across sessions of the importance of a number of variables, particularly repeating sections, as discriminators of students of different musical levels. Thus it appears to be many hours of practising a wide variety of music pieces that influences practising behaviours.

## Interview

The interview was carried out as a means of tapping students' knowledge concerning their practising and investigating differences between subjects differing in musical experience. The Grade II, Grade VI, ARCT students and concert pianists (professionals) were included in this part of the study. Their responses to the six questions were given a quantitative and qualitative score, the quantitative score being the number of strategies described in the response and the qualitative score representing the degree

of cognitive complexity and abstraction reflected in the response. The latter measure involved the rating of each response according to a four-point scale, each point representing the following types of practising strategy:

(1) *Simple undifferentiated.* This category refers to responses that do not specify behavioural strategies nor differentiate between types of music pieces played. The response is excessively general and gives no indication of any particular knowledge of specific strategies, for example, 'I try to do my best'; 'I try to work hard on the piece'.

(2) *Concrete behavioural.* This category includes responses that articulate specific behavioural responses that remain the same regardless of the piece played. The strategies are mechanically employed for all pieces and do not take into consideration the kind of material being practised nor the individual's experience with that kind of piece, for example, 'I go over the piece three times'; 'I play each piece hands separately first', 'I go over each error five times'.

(3) *General strategies.* This category includes responses that indicate the use of generalized strategies for dealing with the practising task. The strategy allows for specific behavioural elaborations that can be flexibly modified for the demands of the particular musical piece, for example, 'I divide the piece into chunks'; 'I correct the error and fit it into its phrase context'; 'I work on the fingering' (or timing, or phrasing' etc.).

(4) *Higher-order strategies.* This category refers to responses that involve higher-order chunking of the task. The strategies specify the general strategies that are carried out. They are abstract and general rather than behaviourally oriented, for example, 'I try to get a feel for the piece'; 'I work on technical problems'; 'I work on interpretation'.

The mean qualitative and quantitative scores for each music level are presented in Fig. 5.3. The graphs indicate that both the qualitative and quantitative scores increased with increasing musical skill. Analyses of variance revealed highly significant main effects for musical level for both qualitative scores ($F$=8.72; df=3,11; $p$<0.005) and quantitative scores ($F$=32.19; df=3.11; $p$<0.001). Thus, more experienced musicians report using more complex practising strategies and these strategies are more complex and abstract than those of less experienced students. One difficulty with the present results is the possible confounding of linguistic and cognitive complexity associated with age. In the present sample, age was highly correlated with music level ($r$=0.75) and the number of words produced in the subjects' interview protocols was highly correlated both with the number of strategies enumerated ($r$=0.78) and with their cognitive complexity ($r$=0.73). The effects of age on music students' self-reports merits further investigation in future research. For example, with a study following the methodology of Chi (1978) comparing the interview data of older, inexperienced players and younger, experienced musicians,

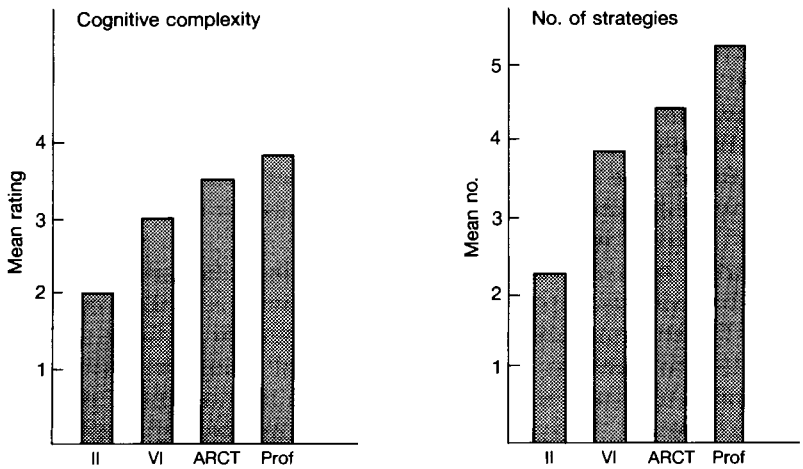



**Fig. 5.3.** Interview: descriptive statistics and ANOVA.

the confounding of age and cognitive and linguistic complexity could be evaluated.

The behavioural analyses pointed to the importance of repetition in the practising of music students. A specific analysis of Question (4) of the interview (What did you do when you made an error?) was carried out in order to investigate subjects' knowledge concerning the units of music they repeated. Fig. 5.4 presents the percentage of responses represented by each of the units of repetition examined for each group, confirmed by the performers as following errors in the interview protocols. Examination of the figure reveals that repeating notes occurred less frequently and repeating sections more frequently as music level increased. Thus the interview data corroborates the findings from the behavioural data.

## Discussion

The results of the study indicate that significant changes in practising behaviour take place as individuals acquire competence as musicians. As music level increased, there was an increase in repeating sections, verbalizing, playing hands separately, and practice time, while errors, repeating notes, and pauses decreased. There were fewer behavioural changes across sessions. Tempo increased across practising sessions for all music levels but reached an asymptote in the early sessions for the more advanced pianists. The most consistent and significant finding was the

change in the units of music repeated during practising as musical experience increased. The behavioural analyses indicated that as musical skill increased, students repeated sections of music more frequently and there was a tendency to reduce the number of notes repeated. Subjects' self-reports also indicated that the size of the unit of repetition increased with music level, corroborating the behavioural data. The interview data also revealed that more experienced students were able to conceptualize their practising behaviour in a more differentiated and abstract manner. They described more practising strategies and these strategies were found to be more cognitively complex. The changes in practising behaviour were far more significant between music levels than between practice sessions of individual pieces. Thus it appears to be the many, many hours of practising involved in mastering a new music level rather than the relatively few sessions it takes to learn an individual piece that bring about consistent and enduring changes in information processing.

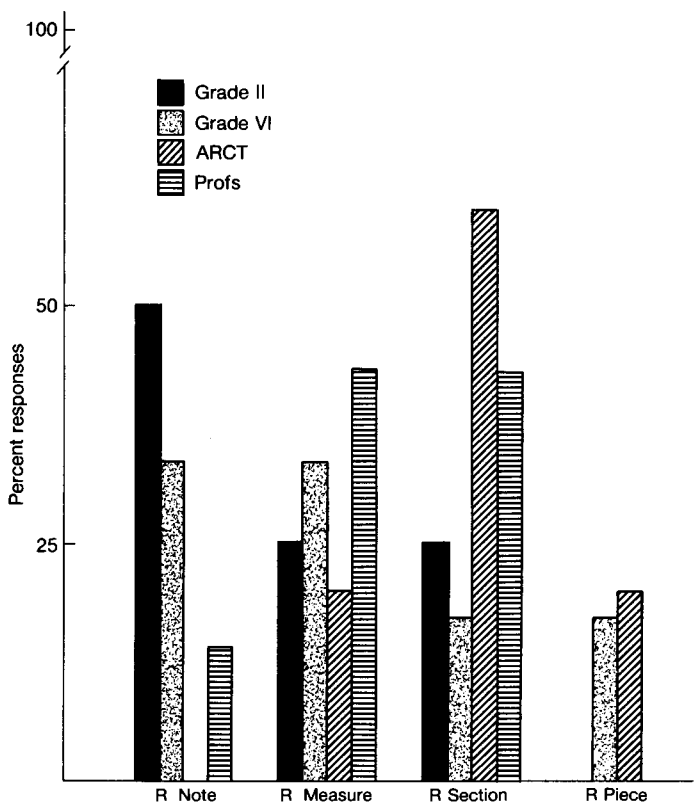



**Fig. 5.4** Interview: size of units repeated following error.

In order to understand how changes are brought about through rehearsal, that is, how chunking takes place, two models may be helpful, that of automatization (Shiffrin and Schneider 1977) and that of motor programming (Shaffer 1980).The notion of automatization as articulated by Shiffrin and Schneider proposed that learning a skill requires a shift between information processing modes from controlled to automatic processing. Controlled processing was described as a slow and effortful mode under attentional control which is subject to the limitations of short-term memory and is serial in nature. Controlled processes tend to be accurate and can be flexibly applied to unfamiliar or changing situational demands. Extended and consistent repetition, such as occurs in practising, permits the establishment of associations in the long-term memory store which, with continued practice, may be triggered by familiar stimulus configurations appearing in short-term memory. Thus, an overlearned, routinized stimulus pattern comes to trigger the performance of a highly stereotyped inflexible behavioural sequence which runs to completion automatically, without the involvement of conscious control. Automatically learned sequences may be combined to form more complex units through a process labelled 'composition' by Neves and Anderson (1981). The composition process enables the more rapid execution of the combined routine.

Musical practising may be viewed as a sequence of transitions from controlled to automatic processing in which larger and larger chunks of musical information are built up from more basic subcomponents. The novice student might be expected to focus upon associating individual notes in the printed score with the corresponding positions on the musical instrument by means of controlled processing. With practice, the associations between printed notes and manual positions become automatized and attention may be focused upon more complex musical patterns such as chords, measures, and ultimately phrases and larger units which may, in turn, come to be executed automatically from a single glance at the score. With repeated practising, the unique sequence and integration of musical patterns in the piece may be executed automatically and the practiser finds himself or herself able to play the piece from memory.

Students' repetition patterns following errors can provide us with clues regarding how they process musical information. With musical expertise, a process of structural differentiation takes place. The beginning student tends to perceive a piece as an amorphous whole consisting of a series of independent notes which, with practising, he or she is able to execute in a more automated fashion. Thus, when a mistake is made, the novice will be likely to repeat the erroneous note or measure. The acquisition process for the beginner involves construction with musical building blocks to erect an integrated whole. The expert musician, however, tends to approach a piece from the top down. He or she is more likely to break the piece down into

hierarchically organized patterns, differentiating it, for example, into a number of independent but linked sections. To correct an error, the skilled player is more likely to repeat the entire structural unit, the section, rather than a component note.

The results of the present research indicated that with increasing skill, on a cognitive level, pianists are able to describe a wider variety of practising strategies and these strategies are more complex, more abstract, and more flexible. Together with the notion of structural differentiation, the notion of motor programming may be added to the formulation to explain these changes. A motor programme, according to Shaffer (1980), may be viewed as a hierarchically organized series of plans. The acquisition of skill involves the development of higher-order plans that are more abstract and incorporate a number of simple components of the skill into more complex, superordinate units. Shaffer defines a plan as:

> an abstract homomorphism of the performance, representing its essential structure. It does not represent all the details of the performance but allows these to be generated or accessed as they are needed in the execution.

In the performance of a skill, the motor programme calls into action an intention that provides a goal and the overall structure for the performance. This intention then calls into play increasingly more specific plans that may be executed automatically and in parallel. Thus this model incorporates the notion of the hierarchical organization of cognitive structures proposed by Chase and Chi (1980). In music, such hierarchical processing is found to take place both at a behavioural and a cognitive level. The inclusion of plans in the formulation takes into account not only the execution of automatized sequences of behaviour, but also the flexibility required in the performance of a complex skill.

In the study discussed here, the articulation of higher-order, more flexible plans was found in the self-reports of the more advanced musicians. On a behavioural level, the increase in the size of the units of repetition associated with increasing expertise reflects the hierarchical nature of motor programming in music performance.

## Directions for future research

The research decribed in this paper was exploratory. There are a number of limitations and opportunities for future research efforts.

One of the major difficulties in the study was the small number of students investigated at each musical level. The reliability and generalizability of the findings need to be assessed by means of a replication incorporating a larger sample. In addition, a problem in the sample selection was the high correlation between age and music level, which made it difficult to ascertain whether differences in practising behaviour were related to

increasing musical experience, increasing age, or a combination of both. While research in other areas, such as chess (for example Chi 1978), indicates that skill depends upon domain-specific experience rather than age, a similar finding needs to be demonstrated in the musical domain by comparing older beginners to more advanced younger students and including such individuals in the general sample. In addition to age, a number of other variables such as differences in teachers' style, students' motivation, and musical aptitude could also be investigated.

While the present study focused on differences in behaviour among music levels, it would be interesting to examine whether differences in practising behaviours exist between subjects of the same level. A larger subject sample would make such an analysis feasible and would provide important information concerning whether some subjects are able to use comparable musical experience to develop more mature patterns of practising than other students. Such data could be used to predict which students might become virtuosi in the future. In addressing the issue of prediction, a longitudinal investigation would be valuable in examining more directly the development of practising technique with experience. One could compare the earlier practising behaviour of students who are later musically successful to that of those who are less successful (for example, those who fail examinations) and those who drop out of music study. It would be particularly interesting to identify those behaviours which best predict future musical involvement and expertise; aspects of the performance, knowledge assessed by the interview, or both. If it is indeed possible to isolate more mature or effective practising strategies, it would be interesting to investigate by means of a specific training programme whether practising is modifiable or whether it is purely a function of experience.

In terms of the analyses, further modification of the OSPP would be useful. While the results underscored the importance of repetition of sections in the development of musical competence, this category could vary widely in size from a few notes to several lines of music. For future work, it would be important to examine sections of varying length separately. The research of Sloboda (1974, 1977) and the self-reports of students in the present study suggest that the phrase is a fundamental musical unit and it could be investigated singly, in addition to larger units such as a thematic line, a physical line, a page, a movement, and so forth. Hence, a more discriminating investigation of the units of information processing could be carried out.

Also of interest is what happens when information processing breaks down and errors occur. Subjects' interview data indicated that they differentiate between different types of errors and respond differently depending upon the kind of error made. In order to obtain a taxonomy of errors for use in future research, a microscopic analysis of errors derived from videotaped practice sessions could be performed.

A number of problems emerged in the interview study. First, there existed a possibility of confounding linguistic fluency and complexity, since verbal productivity was found to be highly correlated with both the qualitative and quantitative scores. To control for this variable, one could compose one's sample of subjects whose age and music level are not as highly correlated, and include age-matched control subjects with no musical experience. Another direction for future research would be a more formal analysis of the content of the protocols of subjects of varying levels.

In the present study, the interview was carried out following the tenth and last practice session. In order to gain an understanding of changes in students' knowledge and affective experience regarding their practising as it progresses, their self-reports may be obtained at a number of points in the practising process. Such a procedure would also permit an assessment of the impact of subjects' cognitions and affects on their practising behaviour. In addition, one could investigate alternate ways of assessing cognitions and affects by changing the format of the interview. For example, one might investigate differences in the data derived from a structured interview such as the one used in the present study, an endorsement (forced choice) procedure, and a coaching procedure (in which a student is required to teach another student at his or her level to play the same piece).

Finally, the generalizability of the findings may be assessed through the application of the methodology to a variety of complex skill domains. Indeed, the methodology employed in this research may be seen as a model which may be applied more generally to the investigation of the development of any skill.

## Acknowledgement

The author wishes to thank Dr Donald Meichenbaum for his advice in designing the study and his helpful comments on an earlier version of this paper.

## References

Bean, K. L. (1937). An approach to the reading of music. *Psychological Monograph,* **226**, 1–80.

Bhaskar, R. and Simon, H. A. (1977). Problem solving in semantically rich domains: an example from engineering thermodynamics. *Cognitive Science* **1**, 193–215.

Charness, N. (1979). Components of skill in bridge. *Canadian Journal of Psychology* **33** (1), 1–16.

Chase, W. G. and Chi, M. T. H. (1980). Cognitivie skill: implications for spatial skill in large scale environments. In *Cognition, social behaviour and the*

*environment* (ed. T. Harvey). Lawrence Erlbaum Associates, Pontiac, Md.
Chase, W. G. and Simon, H. A. (1973). The mind's eye in chess. In *Visual Information Processing* (ed. W. G. Chase), Academic Press, New York.
Chi, M. T. H. (1978). Knowledge structures and memory development. In *Children's thinking: what develops?* (ed. R. Sieder), Thirteenth annual Carnegie Symposium on cognition, Lawrence Erlbaum Associates, Hillsdale, NJ.
Chi, M. T. H., Feltovich, P. J., and Glaser, R. (1980). Representation of physics knowledge by experts and novices. Technical Report No. 2. Learning Research and Development Center, University of Pittsburgh.
de Groot, A. D. (1965). *Thought and choice in chess.* Basic Books, New York.
Flower, L. S. and Hayes, L. R. (1979). *A process model of composition.* Technical Report No. 1. Carnegie-Mellon University, Pittsburgh.
Gruson, L. M. (1978). An observational study of the acquisition of musical skills. Unpublished M.A. thesis. University of Waterloo.
Krutetskii, V. A. (1976). *The Psychology of mathematical abilities in school children.* University of Chicago Press, Chicago.
Lannert, V. and Ullman, M. (1945). Factors in the reading of piano music. *American Journal of Psychology*, **58**, 91–9.
Larkin, J., McDermott, J., Simon, D. P., and Simon, H. A. (in press). Expert and novice performance in solving physics problems. *Science.*
Neves, D. M. and Anderson, J. R. (1981). Knowledge complication: mechanisms for the automatization of cognitive skills. In *Cognitive skills and their acquisition* (ed. J. R. Anderson). Lawrence Erlbaum, Hillsdale, NJ.
Reitman, J. S. (1976). Skilled perception in Go: deducing memory structures from inter-response times. *Cognitive Psychology* **8**, 336–56.
Shaffer, L. H. (1980). Analysing piano performance: a study of concert pianists. In *Tutorials in Motor Behaviour* (ed. G. E. Stelmach and T. Requin). North-Holland Publishing, New York.
Shaffer, L. H. (1981). Performances of Chopin, Bach and Bartok: studies in motor programming. *Cognitive Psychology* **13**, 326–76.
Shiffrin, R. M. and Schneider, W. (1977). Controlled and automatic human information processing II: Perceptual learning, automatic attending and a general theory. *Psychological Review* **84**, (2), 127–90.
Simon, D. P. and Simon, H. A. (1978). individual differences in solving physics problems. In *Children's thinking: What develops?* (ed. R. S. Siegler). Lawrence Erlbaum, Hillsdale, NJ.
Simon, H. A. and Chase, W. G. (1973). Skill in chess. *American Scientist* **61**, 394–403.
Sloboda, J. A. (1974). The eye-hand span—an approach to the study of sight-reading. *Psychology of Music* **2**, (2), 4–10.
Sloboda, J. A. (1977). Phrase units as determinants of visual processing in music reading. *British Journal of Psychology* **68**, 117–24.
Thomas, E. H. C. and Jones, R. G. (1970). A model for subjective grouping in typewriting. *Quarterly Journal of Experimental Psychology* **22**, 353–67.
Weaver, H. E. (1943). Studies of ocular behaviour in music reading. *Psychological Monographs* **55**, (1), 1–29.
Wolf, T. (1976). A cognitive model of musical sight-reading. *Journal of Psycholinguistic Research* **5**, (1), 143–72.

# 6

# Tonal structure and children's early learning of music

W. JAY DOWLING

When we listen to speech in our own language we hear words and sentences, and not mere sounds. Just so, when we listen to the music of our culture we hear melodies composed of meaningful pitches, and not uninterpreted sound. As Pick (1979) has pointed out, to hear a melody is to perceive a meaningful event extended in time—an event with structure, with constraints on its pattern organization, and in which still briefer events are embedded. Even the simplest of melodies presents a complex rhythmic organization of pitches in time. This structure is expressed in pitch and time relationships that link the notes of a melody and that remain largely invariant over the course of the event; melodies tend to move at a regular tempo and draw their pitches from a limited set of the infinite range of possibilities. Furthermore, some event patterns found within a given melody remain invariant across other melodies within the style and culture. For example, in Western music melodies usually end and often begin on the tonic (the focal pitch of a key or tonality) and share patterns of phrase repetition (such as the AABA pattern common to many popular songs and hymns). As Pick and Gibson (1969; Gibson and Levin 1975) have suggested, the perceptual learning that provides the listener with the means of understanding meaningful events such as melodies involves learning to extract relevant pattern invariants from the complex of information available.

## Invariants of tonal structure

Some invariants of melodic structure are explicitly accessible in conscious experience to the typical adult. For example, most adults know explicitly that 'Three Blind Mice' closes with a copy of its opening phrase, and notice when a child does not end it that way. Other invariants are not explicitly

accessible to adults without training, but with training can become so. For example, adults without music lessons are rarely able to describe the pattern of 1- and 2-semitone intervals in the major diatonic scale, while that pattern is one of the fundamentals learned early in formal musical training. Still other invariants remain implicit in the knowledge of even highly trained adults. For example, certain broad aspects of pattern organization described by Lerdahl and Jackendoff (1983) are unlikely to occur to musicians before reading their book (otherwise the book would not be so interesting). And the regularities of melodic structure pointed out by Rosner and Meyer (1982) had probably not occurred to the reader beforehand. Yet presumably such pattern invariants formed an implicit part of musicians' knowledge of the style described.

Even invariants that can on occasion be explicitly accessible usually remain implicit in their usual functioning in the cognition of music. In fact, effective, enjoyable listening normally relies on their remaining so, since, as Meyer (1956) has noted, focusing attention on formal properties while listening can easily interfere with meaningful experience of the music. That our knowledge of pattern invariants shapes our cognition of music can be illustrated by numerous examples. Shepard and Jordan (1984) showed that listeners' judgements of a scale formed of actually equal logarithmic intervals were distorted in ways predicted by their reliance on knowledge of the irregular diatonic pattern. As Francès (1958) noted, our implicit knowledge of the diatonic scale allows us to hear a performer's minor intonation errors as instances of scale categories. Greater departures from the pattern invariant such as out-of-key notes are glaringly noticeable, as Cuddy *et al.* (1979) showed. Furthermore, listeners often confuse exact transpositions of a melody with imitations in which the same diatonic interval pattern has been shifted along the scale they have just been hearing (Dowling 1978), and this confusion diminishes with increasing alterations in the scale; that is, with shifts in the underlying pattern invariants (Bartlett and Dowling 1980).

## Expectancy and encoding

The listener's largely implicit knowledge of invariants of musical patterns serves both to guide expectancies (or 'set') for new events and to provide ready encoding schemes for the events that actually occur. These two processes typically work hand in hand in the cognition of pieces of music. The expectancies created by initial events in a piece probably facilitate efficient encoding of ensuing events, as Haber (1966) suggests. And the ways in which those events are encoded in turn determines the expectancies that are generated for subsequent events. The readying of knowledge-based encoding schemes is very useful in the processing of the rapid and

complex stimulus sequences found in speech and music. As Haber (1966) noted,

> Stimuli for which [a person] has no code for translation into memory are notoriously difficult to remember, even though [an observer] may perform perfectly on . . . discriminations with such stimuli. . . . Most people experience this difficulty in describing colors, tastes, odors, and feelings from recollection. . . . These examples represent instances of excellent matching between the stimulus and prior experience, but very poor encoding of the stimulus into memory.

Fortunately, mere exposure to the music of one's culture induces sufficient perceptual learning to provide a fundamental encoding scheme for pitch. With further training the listener's encoding schemes become more elaborated, leading to changes in performance on memory tasks, and presumably in what is heard when listening to music. This is a process in which listeners become attuned to the perceptual invariants of the music of their culture, and consequently become more efficient perceivers of that music.

This encoding scheme for pitch develops during the first eight years of life, and the pattern of its development is quite regular. New evidence, however, suggests that children's progress in this development may be more rapid than had previously been supposed.

## Invariants and development

### Infancy

Relatively recent reviews have depicted the development of sensitivity to melodic invariants in broad outline (Dowling 1982*a*, Dowling and Harwood 1986, Chap. 5). The evidence indicates that melodic contour—the pitch pattern of ups and downs—is an important feature in melody perception from early infancy on. In a noteworthy study Trehub *et al.* (1984) used conditioning of head turning to show that infants of 8 to 11 months are sensitive to a variety of changes in immediate repetitions of six-note melodies—changes in pitches height of the whole melody (transposition and changes in pitch within a melody (either substitutions at the same general pitch height, or octave alterations of some pitches), as well as changes in melodic contour. However, when Trehub *et al.* increased the difficulty of the task by adding a 'distractor sequence' interpolated between repetitions of the melodies, the infants responded principally to changes in contour. The infants treated repetitions with altered diatonic pitches and transpositions involving small alterations of pitch height as equivalent to exact repetitions, while noticing gross changes in the pitch height of single notes (octave alterations) and contour changes. Melodic contour is a

critically important feature of melodies for infants, and remains so into adulthood (Dowling and Fujitani 1971). Contour is one of the most obvious features of a melody to remain invariant across its instances.

Contour is important to early song production as well as to perception. Around the age of one year children begin to sing spontaneous songs that are clearly differentiated from speech patterns by their rhythmic regularity and their tendency to elongate vowels on discrete pitch levels.* While the child during the first two years uses discrete pitch levels, those pitch patterns do not correspond to adult pitch categories established by the culture. That is, though single notes typically rest on well-defined pitches (as opposed to a siren effect), those pitches do not recur in invariant patterns from song to song, or even within songs. Children when they begin singing do not have pitch scales, though they have a prerequisite component of scales, namely, discrete pitches.

Singing during the early years consists largely of the control of melodic-rhythmic contour patterns.** In an analysis of 579 spontaneous songs produced by two children between the ages of one and six years, Dowling (1984*b*) found that even at the earliest ages songs consisted of patterns of repetition and alternation of brief melodic-rhythmic kernels. Around the age of two years a typical song consisted of the repetition at different pitch levels of a brief melodic phrase. These phrases remained constant in rhythmic and melodic outline throughout the song, and typically involved the repetition of a verbal phrase (such as 'Come a duck on my house', or 'Hoppy-hoppy run 'round the road'). Time within phrases was organized in a metrical beat structure on which the speech rhythms of the words of the phrase were overlaid, but that structure was often interrupted between phrases by breathing and other discontinuities.

The complexity of spontaneous songs increased over the first four years of life, both in terms of the number of phrases used and the pattern of their use. While songs with more than four distinguishable phrase contours were rare throughout that age range, there was a shift in the frequency of songs with just one repeated phrase. Songs with one phrase contour were fairly common between the ages of one and two years. By the age of three years

* Though I had noted indications in the literature that this spontaneous singing behaviour sometimes began as early as six months (for example, Ostwald 1973), I had been doubtful concerning the clarity of discrete pitch levels and rhythmic organization at such early ages. However, Ann Tubbs, a graduate student at UT/Dallas, has recently collected examples from a six-month-old boy that clearly exhibit rhythmic regularity and discrete pitches. I believe the onset of spontaneous singing can vary widely in the range of six to 18 months.

** We should not allow the emphasis in the adult experimental psychology literature on purely melodic contour (with rhythmic patterns typically held constant) to obscure the fact that in the perception and production of real melodies in the world the critically important component is an abstract contour pattern that captures both rhythmic and pitch-pattern invariants (Monahan 1984). As Dowling and Harwood (1986) pointed out, changing the rhythmic pattern of a tune while retaining the melodic contour will make the tune all but unrecognizable.

those had become rare, and songs in which three phrases occurred in varied patterns of repetition and alternation predominated. For songs where two or three phrases were involved the phrase-structure pattern changed between the ages of one and three years. A pattern that became more and more common with increasing age was one in which two, three, or four phrases would recur in some pattern of repetition, and which ended with yet another 'coda' phrase that had not yet occurred in the song. This last structure occurs in a number of nursery songs such as 'Mary Had a Little Lamb', 'Here We Go 'Round the Mulberry Bush', 'London Bridge', and 'On Top of Old Smokey' (all ABAC); 'D'ye Ken John Peel' (AABC); and 'Good King Wenceslaus' (ABABCBDE). But interestingly, this regular phrase structure pattern was more common in the children's songs that in their nursery song models. The children had latched on to a simple pattern that they used to control their songs, thereby constraining them in this respect more tightly than the song models provided them.

## Childhood

Singing during the ages of one to five years is thus largely a matter of controlling the pattern of repetition of rhythmic-melodic phrase contours. On the basis of earlier evidence I had come to believe that until the age of five years or so children remain insensitive to subtler aspects of pitch; that is, they mostly sing 'out of tune' by adult standards. However, during that period they do learn more and more nursery songs that model the adult pitch classes of the *do-re-mi* scale, and they do sing them more and more in tune. And now there is new evidence that such experience leads at least some children by the age of four years to be sensitive to invariants in the details of the pitch pattern of melodies, both in noticing contour-preserving changes in the pitch intervals of familiar melodies, and in distingushing novel tonal melodies (that conform to the scale invariants) from atonal ones (that do not).

Trehub *et al.* (1985) asked four-to-six-year-olds to identify variants of 'Happy Birthday' and 'Twinkle, Twinkle, Little Star' as either normal or 'funny'. Presumably distortions that violated the child's own internal models of the songs, and that could be detected given the precision of the child's processing capabilities, would trigger 'funny' responses. If the child were relying solely on contour in responding to the songs, then alterations that preserved contour should not sound 'funny'. However, 'funny' responses to changes in pitch intervals where contour is preserved can be taken to indicate that the child is relying on more than just contour in judging the melodies. Adults, even with no musical training, are very sensitive to changes of pitch intervals in familiar melodies. Trehub *et al.* found that the children behaved like adults in rejecting versions of both of these tunes in which pitch intervals had been changed, as well as versions

that changed contour. For 'Twinkle, Twinkle' rejection on the basis of interval change was as great as for contour change, while for 'Happy Birthday' contour violations caused more rejections than interval violations. Thus for some well-known, familiar melodies four-and five-year-olds are able to notice changes in intervals from the models, as well as changes in contour.

Trehub *et al.* (1985) incidentally tested the notion that kindergarteners are sensitive to changes of key—to shifts in the tonal center and the set of pitches that make up the 'do-re-mi' scale underlying a melody. This tendency had been observed by Bartlett and Dowling (1980) as an increased tendency to reject a comparison melody the further it departed from a standard stimulus in terms of 'key distance' (a measure of shared pitches in the diatonic scale—the greater the number of shared pitches the smaller the distance). Trehub *et al.* observed such an effect as well, obtaining more 'funny' responses to far-key comparisons than to same-key comparisons (with near-but-different-key comparisons intermediate). This adds to our evidence that children in this age range are becoming sensitive to certain invariants of the culturally determined scale patterns.

The foregoing evidence shows that preschoolers are sometimes sensitive to pitch substitutions that alter the intervals of familiar tunes, and to transpositions of familiar tunes that take them into relatively distant keys using different sets of tonal scale pitches. Adults, especially musically trained adults, can also discriminate between melodies that are tonal—that conform to the invariants of a tonal scale pattern—and those that are not (Dowling 1978). I was curious to find out whether preschoolers would display sensitivity to the tonality vs. the atonality of melodies. Sensitivity to altered pitches in familiar melodies might simply reflect specific knowledge of certain well-learned tunes, and not reflect general knowledge of invariants of scale structure. However, sensitivity to 'atonal' alterations of pitch that violate scale invariants, as compared with tonal alterations that remain within the scale structure, can be taken as an indication that the child had some knowledge of the tonal structure of melodies in general—knowledge of the cultural invariants. Therefore, we conducted the following experiment.*

## Preschoolers' sensitivity to atonal melodies

### Method

The experimenter visited preschools and kindergartens in the Dallas area, and tested the children individually in isolated rooms in the schools. Eight

* I thank James C. Bartlett for helpful suggestions in the design of this experiment, and Melinda Andrews for carrying it out.

three-year-olds (that is subjects between the ages of 3.0 and 4.0 years), 12 four-year-olds, and 12 five- and six-year-olds served in individual sessions. The experiment was replicated with a group session for eight college students at UT/Dallas (mean age = 29.6 years), who divided evenly into those with two or more years of music lessons, and those with no musical training.

The stimuli were 16 brief melodies approximately 8 s long, presented at a rate of two crotchets per s (where a crotchet represents the most common note value of the melody). The melodies were derived from eight familiar tunes in the key of C major: 'Mary Had a Little Lamb', 'Frère Jacques', 'Yankee Doodle', 'Old Macdonald', 'Oh, Susanna', 'Good King Wenceslaus', 'London Bridge', and 'Waltzing Matilda'. From each tune two new melodies were derived: one tonal and the other atonal. The tonal derivatives were created by shifting the melodic pattern along the C major scale so as to preserve the contour and diatonic intervals, but alter the interval sizes among the notes. (All the tonal derivatives still either began or ended—and usually both began and ended—on one of the pitches of the tonic triad: C–E–G.) The atonal derivatives were created by placing the melodic patterns at analogous points on an 'atonal' scale designed to produce interval patterns not contained in either the major or the minor diatonic scales in Western music. This scale consisted of the set of pitches C–D flat–E–F–G flat–A flat–B—an interval set with half-steps between the first and second, third and fourth, fourth and fifth, and seventh and eighth degrees, augmented seconds between the second and third degrees and sixth and seventh degrees, and a whole-step between the fifth and sixth degrees. The purpose of this method of generating stimuli was to preserve the average interval size and the amount of pitch repetition characteristic of melodies familiar to the children. This was especially important in the atonal derivatives, where it would not have been as satisfactory simply to substitute non-scalar pitches for scalar ones in isolated instances. The mean interval sizes were 1.66 and 1.69 semitones/interval for tonal and atonal stimuli respectively, counting unisons (note repetitions) as zero semitones.

All 16 stimuli were presented to each subject. Two different random orders of the stimuli were tape-recorded using the 'clarinet' voice of a Yamaha synthesizer, and each tape was presented to half the subjects in each age group. Thus the two derivatives—tonal and atonal—of each seed melody were presented to each subject.

The experimenter brought a tape-recorder, two stuffed animals (rabbit and bear), and a 'puppet-theatre' screen to the school. At the start of the session the experimenter conversed briefly with the child, and asked the child to sing a song, most often 'Happy Birthday'. (For the younger children the experimenter suggested that it was the bear's birthday.) All the children except for four three-year-olds sang something, which the experimenter recorded. The experimenter sat behind the puppet-theatre

screen, and the child sat in front of it. The experimenter explained that the two animals liked different sorts of songs. (The animals' preferences were alternated from subject to subject. For older children who considered them too childish the animals were dispensed with.) One of the animals liked songs that had all normal notes in them. The experimenter played an example of a tonal melody, and displayed the animal that liked that type. The other animal liked 'funny' songs that had notes in them that sounded 'odd' or 'different from the other notes'. The experimenter displayed the other animal and played an example of an atonal melody. Then the experimenter presented the task in which the curtain of the puppet theatre would be drawn and the subject would listen to a melody. Behind the screen was the animal that liked that song, and the subject had to guess which animal it was. After the subject guessed, the curtain was opened to reveal which animal it really was. (This provided feedback for the trial.)

The songs the children taped at the start of the session were rated by 12 independent judges for 'how well in tune' the children sang. The children in each age group were divided at the median for those ratings into higher- and lower-rated singing groups. The four three-year-olds who did not sing at all constituted the lower-rated group for that age.

## Results

Table 6.1 shows per cent correct judgements for the subjects with higher- and lower-rated singing in the three children's age groups, and for the adults. A 2-ratings $\times 3$ ages analysis of variance on the children's data showed only a main effect of song rating, $F\ (1,26) = 5.37$, $p<.05$. Children's performance in all the age groups was essentially at chance for those whose singing received low ratings, and was significantly better than that for the chidren with higher ratings of singing intonation.

**Table 6.1** *Per cent correct responses discriminating tonal from atonal melodies distributed according to song performance (children) or musical experience (adults)*

| | Group | |
|---|---|---|
| Age | Low Song Rating (inexperienced) | High Song Rating (experienced) |
| 3 yr | 50 | 58 |
| 4 yr | 50 | 59 |
| 5–6 yr | 46 | 61 |
| Mean for children | 49 | 59 |
| adult | 70 | 72 |

**Discussion**

The most interesting feature of these results is that children as young as three and four years, provided they were among the more competent singers, displayed the ability to respond to the degree of tonality of the stimuli. Presumably even at such early ages increased facility with simple musical materials—the songs learned at home and at school—goes hand in hand with facility in discriminating melodies on the basis of invariant properties such as conformity to standard scales. It seems plausible that nursery songs that emphasize the pitch content of the basic scales in the culture, without shifts of key or the introduction of altered pitches foreign to the key, would best facilitate perceptual learning involving the scale invariants. The next piece of evidence suggests that nursery songs are simple in just that way.

## Tonal structure of nursery tunes

The best materials from which to learn the invariants of the tonal scale patterns are tunes in which those invariants are most clearly present; namely, tunes that introduce few alterations of pitch into the basic pattern. To the extent that nursery songs fill this need for simple learning materials for the scale system, we should expect them to be strongly tonal, and to contain relatively few departures from the key in which they start. Folk-songs from the same culture, but intended for adult singing, could be expected to contain more variant pitches departing from their initial keys. 'Art' songs, intended to be sung by advanced amateurs and professionals, should contain still more variation of pitch material. To test these possibilities I counted songs in those three categories that contained, and that did not contain, alterations of pitch which respect to the starting key of the song.

**Method**

I assembled a collection of 223 nursery songs from a group of books labelled as such by their editors (Seeger 1948; Moorat 1980; Beall and Nipp 1982). For adult folk-songs I used the 317 songs in Lomax (1975). And for art songs I used the 44 songs in Schubert's (1823, 1827) *Die schöne Müllerin* and *Die Winterreise*. I chose Schubert for this comparison because his songs, especially in these song cycles, have a folk-like, singable character, and are of the same general length and form as the songs in the other categories. That is, these songs of Schubert are as close as any in the Classical–Romantic tradition of European music to the European folk-song tradition.

## Results

I categorized each song according to whether it contained alterations of pitch that departed from its initial key. The results are shown in Table 6.2. Clearly very few nursery tunes contain such departures, while the adult folk-songs contained more. All but two of the Schubert songs contained alterations. A chi-square on the 2×2 matrix involving nursery and adult folk-songs, altered or not, was signficant, chi-square (1)=45.38, $p<.01$.

**Table 6.2** *Number of nursery, folk, and Schubert songs with and without pitches altered with respect to the starting key of the song*

| | Nursery Songs | Adult Folk-songs | Schubert Songs |
|---|---|---|---|
| No Altered Notes | 208 | 220 | 2 |
| Altered Notes | 15 | 97 | 42 |
| Total | 223 | 317 | 44 |

## Discussion

Nursery tunes clearly differ, not only from Classical–Romantic art songs, but also from adult folk-songs in the culture, with respect to the simplicity of their tonal schemes. Nursery songs emphasize the repetition of strongly tonal patterns, and avoid the introduction of pitches that depart from the key established at the outset. They are thus ideal vehicles for perceptual learning of the invariants of the tonal scale pattern.

In passing, let me add a word about the differences between the adult folk-songs, one-fourth of which include altered pitches, and the Schubert songs. Generally, when a folk-song uses an altered pitch it is either as a chromatic passing tone, interpolated into the passage between two solidly tonal scale steps, or it involves playing with the differences between the major and minor (or other) modes, as with the alterations in the song 'Greensleeves'. Rarely do such alterations involve a change of key—a shift in the tonal centre and in the scale pattern. However, of the 42 songs in the Schubert sample that introduced alterations, 36 involved brief changes of key, typically modulations to the dominant. Thus the Schubert songs, simple and folk-like though they appear, contain not only more instances of altered pitches, but also a preponderance of altered pitches that signal shifts in the tonal structure of the song. Kessen (1981), in emphasizing the child's need for simple learning materials in acquiring basic knowledge of musical structure, points out the inappropriateness of the derivatives of tonally elaborate late-Romanticism for that function. From this comparison it

appears that even Schubert songs may be more complicated than is optimal for preschoolers' acquisition of the tonal system.

## Musical experience and adult encoding

I wish to turn now to some new evidence bearing on the role of perceptual learning in the development of encoding schemes that reflect invariant patterns in musical structure. I will present evidence that even musically untrained listeners will, under certain circumstances, encode pitches in terms of the culturally defined pitch categories of the equal-tempered scale, as well as evidence that listeners with more training utilize more elaborate encoding strategies that involve diatonic pitch categories, defined within key structures.

### Assimilation of quarter-steps

Even without specialized training in music, listeners use an encoding scheme for pitch arising from perceptual learning through exposure to their culture's 'tonal material'—the invariant describing the set of all possible pitch relationships in the culture (in Western music the 'chromatic' scale of all the semitones on the piano—Dowling 1978, 1982*b*, 1984*a*). This categorical encoding is apparent in the judgements of both musically trained and untrained listeners when judging pitches embedded in melodic contexts. In order to disclose this effect listeners must be put in a situation requiring rapid processing of tones, presumably via encoding guided by expectancies. As the quotation cited above from Haber (1966) suggests, listeners tend to be quite accurate when they can simply make direct comparisons between stimuli without using memory. However, following a familiar melody temporally interleaved with distractor tones presents them with a much more complicated problem. Previous results had shown that listeners can detect a familiar melody when its notes are temporally interleaved with the notes of a distractor pattern (at a rate of 8 or 10 notes/s in the same pitch range and of the same timbre and loudness, Dowling 1973). Recent results show that when listeners follow the pitch of a wandering tone in such a melody, and then compare its pitch with that of a subsequent probe tone, they make systematic errors (Dowling 1985). Those errors generally consist of judging half-step probes equal in pitch to actually presented quarter-steps, while making accurate responses to actually present half-steps. Thus quarter-steps appear to be encoded as neighbouring half-steps, while the half-steps appear to be encoded accurately.

Since listeners with no formal musical training show this assimilation effect of encoding quarter-steps as neighbouring semitones of the tonal

Lerdahl, F. and Jackendoff, R. (1983). *A generative theory of tonal music.* MIT Press, Cambridge, Mass.

Lomax, A. (1975). *The folk songs of North America.* Doubleday, Garden City, NY.

Meyer, L. B. (1956). *Emotion and meaning in music.* University of Chicago Press, Chicago.

Monahan, C. B. (1984). Parallels between pitch and time: the determinants of musical space. Unpublished Ph. D. thesis. University of California, Los Angeles.

Moorat, J. (ed.) (1980). *Thirty old-time nursery songs.* Metropolitan Museum of Art, New York.

Ostwald, P. F. (1973). Musical behaviour in early childhood. *Developmental Medicine and Child Neurology* **15**, 367–75.

Pick, A. D. (1979). Listening to melodies: perceiving events. In *Perception and its development* (ed. A. D. Pick), pp. 145–65. Erlbaum, Hillsdale, NJ.

Rosner, B. S. and Meyer, L. B. (1982). Melodic processes and the perception of music. In *The psychology of music* (ed. D. Deutsch), pp. 317–41. Academic Press, New York.

Schubert, F. (1823). *Die schöne Müllerin*, Op. 25. Kalmus, New York.

Schubert, F. (1827). *Die Winterreise*, Op. 89. Kalmus, New York.

Seeger, R. C. (1948). *American folk songs for children.* Doubleday, Garden City, NY.

Serafine, M. L. (1983). Cognition in music. *Cognition* **14**, 119–83.

Shepard, R. N. (1984). Ecological constraints on internal representation: resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review* **91**, 417–47.

Shepard, R. N. and Jordan, D. S. (1984). Auditory illusions demonstrating that tones are assimilated to an internalized musical scale. *Science* **226**, 1333–4.

Trehub, S. E., Bull, D., and Thorpe, L. A. (1984). Infants' perception of melodies: the role of melodic contour. *Child Development* **55**, 821–30.

Trehub, S. E., Morongiello, B. A., and Thorpe, L. A. (1985). Children's perception of familiar melodies: the role of intervals, contour, and key. *Psychomusicology* **5**, 39–48.

# 7

# Improvisation: methods and models

JEFF PRESSING

## Introduction

How do people improvise? How is improvisational skill learned and taught? These questions are the subject of this chapter. They are difficult questions, for behind them are long-standing philosophical quandaries such as the origins of novelty and the nature of expertise, which trouble psychologists and artificial intelligence workers today almost as much as they did Plato and Socrates in the fourth and fifth centuries BC.

In a previous article (Pressing 1984*a*) I summarized a number of general properties of the improvisation process on the basis of the diverse historical writings of artists, teachers, and musicologists. This material was integrated with precepts from cognitive psychology to sketch out the beginnings of a general theory of improvisation.

In this article a much more explicit cognitive formulation is presented, the first proper (though by no means necessarily correct) theory of improvised behaviour in music. The building of this theory has required input from many disparate fields with which the general musical reader may not be familiar. For this reason I begin with the survey of appropriate background research and its relation to improvisation. Some of these areas may initially seem distant from the topic at hand.

## A survey of pertinent research

### Some physiology and neuropsychology

Although our state of knowledge in these areas is far too meagre to have any definite repercussions for improvisation, there are a few facts which are at least strongly suggestive.

To begin with, improvisation (or any type of music performance) includes the folowing components, roughly in the following order:

(1) Complex electrochemical signals are passed between parts of the nervous system and on to endocrine and muscle systems;
(2) muscles, bones, and connective tissues execute a complex sequence of actions;
(3) rapid visual, tactile, and proprioceptive monitoring of actions takes place;
(4) music is produced by the instrument or voice;
(5) self-produced sounds, and other auditory input, are sensed;
(6) sensed sounds are set into cognitive representations and evaluated as music;
(7) further cognitive processing in the central nervous system generates the design of the next action sequence and triggers it.

– return to step (1) and repeat –

It seems apparent that the most starkly drawn distinctions between improvisation and fixed performance lie in steps (6) and (7), with possibly important differences in step (3). This chapter therefore inevitably focuses on these aspects.

The given steps are often collapsed into a three-component information-processing model of human behaviour which has ready physiological analogies: input (sense organs), processing and decision-making (central nervous system, abbreviated CNS), and motor output (muscle systems and glands).

Control of movement by the CNS is complex: the cerebral cortex sends signals to both the cerebellum and the basal ganglia, which process the information and send a new set of signals back to the motor cortex. The brainstem nuclei are also involved in details of motor co-ordination. It has been suggested that the basal ganglia and cerebellum have complementary roles, with the basal ganglia initiating and controlling slow movements while the cerebellum is active in the co-ordination of fast, ballistic movements (Sage 1977).

Motor signals from the cortex pass to the spinal cord and motor nuclei of the cranial nerves via two separate channels: the pyramidal and extra-pyramidal systems. These two nerve tracts illustrate the simultaneously hierarchical and parallel-processing aspects of CNS control, for they run in parallel but interconnect at all main levels: cortex, brainstem, and spinal cord. Hence while each tract has some separate functions there is a redundancy that can be used to facilitate error correction and motor refinement. Similar redundancy and parallel processing is found at lower levels of motor control. Alpha-gamma coactivation, for example, describes the partial redundancy of neural information sent to two distinct types of motoneurons, alpha and gamma, whose axons and collaterals terminate on the main skeletal muscles and the intrafusal muscle fibres, respectively.

The organization of behaviour has often been linked with the existence

of motor action units (or equivalent concepts), and their aggregation into long chains to develop more complex movements. The validity of the concept of motor action units can be seen mirrored physiologically in the existence of command neurons, single nerve cells in invertebrates whose activation alone suffices to elicit a recognizable fragment of behaviour. The effect is achieved by excitation and/or inhibition of a constellation of motoneurons (Bentley and Konishi 1978; Shepherd 1983). While there are no known single cells that fully trigger complex behaviour in mammals, populations of neurons in the brains of higher animals are strongly suspected of serving a similar function (Beatty 1975). It is therefore possible to speculate that skilled improvisers would, through practice, develop general patterns of neural connections specific to improvisational motor control.

Finally, it is of interest that neurological correlates have recently been discovered for a division of knowledge and memory into two separate categories: declarative and procedural. A degree of independence of these two types of memory (for facts or procedures) has been reported among amnesic and post-encephalitic patients for some time (for example Milner 1962; Brooks and Baddeley 1976). Typically, patients can not remember new facts, but are able to learn new motor skills over a period of time, yet without any awareness on successive days of having performed the tasks before. In recent studies, Cohen (1981) and Cohen and Squire (1981) have shown that declarative learning is linked to specific diencephalic and bitemporal brain structures. Unaware of this work, I drew a related distinction in a recent paper (Pressing 1984*a*) between *object* and *process* memory, based on the rehearsal strategies of improvising musicians. As Squire (1982) has pointed out, there are parallel distinctions in earlier writings: artificial intelligence (Winograd 1975), knowing how and knowing that (Ryle 1949), habit memory and pure memory (Bergson 1910), and memory with or without record (Bruner 1969). What is suggestive about these correlations is that physiological locations for some specific cognitive skills used in improvisation might very well exist.

**Motor control and skilled performance**

This area traditionally has centered around industrial skills, sport, typing, handwriting, specially designed laboratory tasks like tracking, and to a lesser degree music. It is a complex field of considerable relevance to improvisation, even though improvisation *per se* is scarcely mentioned. Therefore I first review general theories of motor control, and then delve into a number of special issues in skilled performance and skill development that are relevant here.

*Theories of motor control and skill*

The starting point for nearly all the existing theories is the three-stage information-processing model mentioned earlier, based on sensory input, cognitive processing, and motor output. To this must be added the notion of feedback (auditory, visual, tactile, or proprioceptive). Traditional 'open-loop' theories include no feedback, and hence no mechanisms for error correction. In its starkest form this theory is clearly inappropriate for improvisation; however, there is persistent evidence, dating back to the medical work of Lashley (1917), and including studies of insect behaviour and de-afferentiation techniques in monkeys that points to the existence of motor programmes that can run off actions in open-loop fashion.

In contrast stand 'closed-loop' theories, which contain feedback, and hence allow for the intuitively natural possibilities of error detection and correction. The closed-loop negative feedback (CLNF) model is one of the oldest. In this model the feedback (primarily auditory in the case of musical improvisation) is sent back to an earlier stage in the control system which compares actual output with intended output, producing a correction based on the difference between the two (see for example Bernstein 1967). Such closed-loop models have their historical roots in engineering models of servomechanisms, control theory, and cybernetics.

A wide variety of closed-loop formulations has been given. Gel'fand and Tsetlin (1962, 1971) used a mathematical minimization procedure to model the cognitive search for appropriate motor behaviour. Pew and Baron (1978) sketched out a theory of skilled performance based on optimum control (see also Kleinman *et al.* 1971). Powers (1973) proposed a hierarchy of motor control systems whereby the correction procedures of higher-order control systems constitute reference signals for lower-order systems. Another hierarchical model was given by Pew (1974), in which specific single movements are combined into sequences, and ultimately into various subroutines that make up goal-directed action. Actions are then organized and initiated by an executive programme (Fitts 1964). As is apparent, many such hierarchy theories are based on the application of computer programming principles (see Miller *et al.* 1960).

These ideas offer a more sophisticated understanding of motor behaviour, but they have serious limitations. They model motor learning either poorly or not at all, and are not based on empirical findings about human actions (Adams 1961). A closed-loop theory of motor *learning* was proposed by Adams (1971, 1976) in an attempt to rectify some of these problems. In this theory there are 'memory traces' which select and initiate movements and 'perceptual traces' which are representations of the intended movements, and are used as templates for error correction. A perceptual trace is gradually built up by repeated practice from feedback, knowledge of

results (often abbreviated KR), and error correction. Eventually the perceptual trace can function as an internalized goal, diminishing dependence on the externally based knowledge of results (Namikas 1983). Hence open-loop control characteristics are not completely excluded.

By the late 1970s the consensus was that both open- and closed-loop control must occur in skilled performance (Keele and Summers 1976; Delcomyn 1980; Paillard 1980; see Summers 1981 for a review). That is, movements are both centrally stored as motor programmes, and susceptible to tuning (adjustment) on the basis of feedback. Coupled with the well-established concept of *flexibility* characteristic of skilled (but not rote) performance (Welford 1976), this promoted approaches based on more abstract programming notions that brought the field closer to artificial intelligence (and made it more germane to improvisation).

In this spirit Schmidt (1975, 1976) proposed a theory of motor schemata that models both recall and recognition. The schema is considered to contain the general characteristics of a movement which must be organized in any given situation to satisfy environmental requirements and the goals of the performer. Context then guides the production of a series of motor commands that ultimately generate a spatiotemporal pattern of muscle actions. Feedback is based on a template-comparison idea.

Because schemata are not specific movement instructions but are 'generalized' motor programmes, this theory is capable of modelling novelty (at least in a very general way), which the others above could not (except Pew 1974, which also uses a schema notion). The possibility of novelty is also catered to by Schmidt's inclusion of degree of variability of practice conditions as one determiner of schema 'strength'. At its core, the 'novelty problem' is very close to that of improvisation.

Similar to schemata is the notion of action plan. Miller *et al* (1960) gave a general description of plans, while Clark and Clark (1977) described plans for language discourse, and Sloboda (1982) and Shaffer (1980, 1981, 1984) specified plans for playing music. As discussed by Shaffer (1980), a plan is an abstract homomorphism representing the essential structure of the performance and allowing finer details to be generated or located as they are needed during execution.

Other related theories include Allport's proposal of a system of condition-action units which are links between sensory calling patterns and categories of action (Allport 1980). Also related are adjustable control or description structures for artificial intelligence such as *frames* and *scripts* (see below).

This convergence of theory is useful in constructing a model of improvisation (see below). However, it remains rather unspecific, and has run far ahead of experiment. But as of this writing there seems only one alternative in the area of motor behaviour. This is the organizational invariant approach of Turvey, Kugler, Kelso, and others (Turvey 1977;

Kugler *et al.* 1980; see Kelso 1982 for further references). This approach draws on two sources: the ecological perspective of Gibson (1966, 1979) and the dissipative structure model of non-equilibrum thermodynamics (Prigogine 1967; Prigogine and Nicolis 1971). Essentially the theory de-emphasizes notions of cognitive process and control, replacing them with, in so far as is possible, 'organization invariants'. These organization invariants are characteristic constraint structures that allow the emergence of specific spatial relationships and dynamic processes in the behaviour of non-linear systems when the parameters controlling these systems fall in certain critical ranges. Thus if the human motor action apparatus is considered to be (as it certainly is) a non-linear system, characteristic properties of muscle groups and patterns of human limb co-ordination will naturally emerge from the constraints imposed by a given task situation (Kelso *et al.* 1981; Saltzman and Kelso 1983). The proposals are exciting, but their ultimate fate remains unclear. The theory is still being formulated, and comparable ideas from non-linear mathematics have infiltrated many fields in the last 10 years, with uneven results.

Organizational invariant theory seems also likely to apply primarily to the dynamics of motor programme execution, and not to the formulation of intentions and high-level decision-making (Wilberg 1983). Since these functions are vital elements in improvisation in any but an extreme mechanistic approach, the theory as it stands is not particularly suitable for improvisation modelling. Nevertheless, these ideas may be used in an understanding of the sources of behavioural novelty, and are discussed further below.

*Some special issues relevant to improvisation*

**Skill classification** Various dimensions of skill classification have been proposed and improvisation can be placed within these. Two possible categories are 'open' skills, which require extensive interaction with external stimuli, and 'closed' skills, which may be run off without reference to the environment (Poulton 1957). Solo improvisation is basically a closed skill, as it relies only on self-produced stimuli, whereas ensemble improvisation is more open. Other dimensions of skill classification are gross-fine, discrete-serial-continuous, complex-simple, and perceptual-motor (Holding 1981). Improvisation is a fine, complex skill, with both perceptual and motor components; continuous actions predominate, although there are also discrete and serial motor aspects. This last point varies somewhat with the nature of the instrument played.

It is important to also emphasize the contrast between unskilled and highly skilled performance. A vast majority of reported skill studies treat simple motor tasks like tracking, under an implicity reductionistic scientific methodology. It is increasingly acknowledged, however, that highly developed skills have distinctive emergent properties missed in these

earlier short-term studies, properties such as adaptability, efficiency, fluency, flexibility, and expressiveness (Welford 1976; Shaffer 1980; Sparrow 1983). These are vital components of improvisatory skill.

**Feedback and error correction** Feedback is a vital component in improvisation for it enables error correction and adaptation—a narrowing of the gap between intended and actual motor and musical effects. But feedback is also important for its motivational (Gibbs and Brown 1956) and attention-focusing effects (Pressing 1984*a*).

Feedback redundancy is an important concept for music. Aural, visual, proprioceptive, and touch feedback reinforce each other for the instrumental improviser, whereas the vocalist has only hearing and proprioception available (Pressing 1984*a*). Likewise the design of some instruments allows more precise visual feedback and more categorical kinaesthetic feedback than others. This is almost certainly why sophisticated improvisation using advanced pitch materials is more difficult on the violin than the piano, and extremely challenging for the vocalist. For every first-rate scat-singer in the world there must be 500 talented jazz saxophonists.

Feedback can also be considered to operate over different time scales. Thus short-term feedback guides ongoing movements, while longer term feedback is used in decision-making and response selection. Still longer term feedback exists in the form of knowledge of results (KR) for skills where external evaluation is present or result perception is not sufficiently precise or immediate. The importance of this for improvisation has been demonstrated by Partchey (1974), who compared the effects of feedback, models, and repetition on students' ability to improvise melodies. Feedback, in the form of playbacks of recordings of the students' own improvisations, was clearly superior to listening to pre-composed model melodies, or repetition, as an improvisation learning technique. In group improvisation, feedback loops would also operate between performers (Pressing 1980).

In view of the interconnectedness of the parts of the central nervous system, it is also clear that there exist internal feedback (and feedforward) loops not based on sensory processing (Brooks 1978). That is, if higher cognitive levels set the design of motor programmes while movement fine structure is specified in closed-loop fashion by lower levels of the CNS, notably the spinal cord, then copies of these lower level motor instructions are almost certainly sent directly back up to higher centres. In other words, there is some kind of central monitoring of efference. This would serve to increase overall processing speed and accuracy.

The role of errors in improvisation has been discussed previously (Pressing 1984*a*). It will simply be pointed out here that errors may accrue at all stages of the human information processing system: perception, movement/musical gesture selection and design, and execution. Minor

errors typically demand no compensation in following actions, whereas major errors typically do.

**Anticipation, preselection, and feedforward** These three concepts have to do with preparation for acton. Physiological recording of the Bereitschaft potential (BP) and contingent negative variation (CNV) (see Brunia 1980) now provides explicit support for the long-standing idea that higher cognitive control centres bias lower ones towards anticipated movements. This is therefore a type of feedforward, and has been described from various perspectives: spinal 'tuning' (Turvey 1977; Easton 1978), corollary discharge or efference copy (von Holst 1954), and preselection (see Kelso and Wallace 1978 for discussion).

The idea of preparation is very important for improvisation, where real-time cognitive processing is often pushed up near its attentional limits. It can be formally proved, for example, that only a control system with a model of disturbances and predictive power can become error free (Kickert *et al.* 1978). For improvised performance that aims at artistic presentation, where discrepancies between intention and result must be kept within strict bounds, practice must attempt to explore the full range of possible motor actions and musical effects, to enable both finer control and the internal modelling of discrepancies and correction procedures, including feedforward.

**Hierarchy vs. heterarchy** Because of influences of the physical sciences and control theory, an overwhelming majority of models for motor behaviour have used a hierarchical control system. However, the interconnectedness between difference locations in the CNS and the many documented types of feedack and feedforward processes mentioned above argue that this perspective is probably too narrow. Furthermore, explicit parallel-processing possibilities exist due to the separate pyramidal and extrapyramidal neural tracts, alpha-gamma coactivation, etc., as mentioned above. Hence other types of organization, referred to as heterarchical or coalition, have been proposed (McCulloch 1945; von Foerster 1960; Greene 1972; Turvey 1977). In this perspective, executive control of the system may be transferred between different 'levels' depending on the needs of the situation (Miller *et al.* 1960).

**Time scales for the control of movement** This is a subject with an enormous and complex literature. For background purposes in modelling improvisation a few points only seem sufficient.

Actual neural transmission times are on the order of tens of milliseconds. According to Davis (1957; see also Sage 1977), auditory stimulus activity reaches the cerebral cortex 8–9 ms after stimulation while visual stimulation involves a longer latency of 20–40 ms. Since the two neural pathways are of comparable length, this difference points to a greater

transmission speed for audition than vision. It should, however, be noted that the auditory system contains both ipsilateral and contralateral pathways, while the pathways of the visual system are exclusively crossed. The cortical response time for a movement stimulus appears to be on the order of 10–20 ms (Adams 1976).

Reaction time is the time taken for a sense stimulus to travel to the CNS and return to initiate and execute a largely pre-programmed motor response. Simple reaction time (RT) with only one chosen motor response typically fall in the range 100–250 ms, depending on conditions and sensory modality (Summers 1981). Auditory, kinaesthetic, and tactile reaction times have typically been found to fall in the range 100–160 ms (Chernikoff and Taylor 1952; Higgins and Angel 1970; Glencross 1977; Sage 1977), while visual reaction times have been considered longer, typically reported as at least 190 ms (Keele and Posner 1968). Reaction times for other sensory modalities seem to be in the range above 200 ms, while RTs involving choice of response are in general longer and are reasonably modelled by Hick's Law (Hick 1952). Kinaesthetic and tactile choice reactions seem also to be faster than visual (Leonard 1959; Glencross and Koreman 1979). Data on auditory choice RTs do not seem to be readily available.

Error correction (EC) times vary with sensory modality and context. EC times are important for improvisation because it may reasonably be argued that they reflect minimum times for decision-making that is expressive or compositional. Visual error correction is usually reported to be about 200 ms, whereas kinaesthetic EC can occur over intervals as short as 50–60 ms (Kerr 1982), as seen in reports on tracking tasks (Gibbs 1965; Higgins and Angel 1970). However, other recent work in the case of vision has found some instances of visual EC times down in the range near 100 ms as well (Smith and Brown 1980; Zelaznik *et al.* 1983). It seems likely that the time taken for error correction would be a function of the degree of invoked processing involvement; that is, motor programme construction would take more time than selection, while exacting criteria of discrimination or motor accuracy or a wide range of response choice would naturally increase EC time. Rabbitt and Vyas (1970) and Welford (1974) have enunciated this view, one which is well supported by the introspective reports of improvisers going back for many centuries (Ferand 1961).

Explicit information on auditory error correction times does not seem to be available, but it is possible to point out a general tendency in the above data. Namely, processing speed seems to be greatest for audition and touch/kinaesthesia, of all the possible sensory systems. These are precisely the elements involved in musical improvisation and provide a vivid psychological interpretation for the historical fact that music, of all art and sport forms, has developed improvisation to by far the greatest degree. Under this interpretation, human beings, as creative agents, have as a

matter of course drawn on the sensory systems most adapted to quick decision-making: in other words, a predilection for improvised sound manipulation might be genetically programmed. Of course, such an interpretation remains highly speculative.

Finally it should be noted that unexpected sensory changes requiring significant voluntary compensations require a minimum time of about 400–500 ms (Welford 1976). This is therefore the time scale over which improvising players in ensembles can react to each others' introduced novelties (about twice a second). Nuances in continuous improvised performance based on self-monitoring are probably limited by error correction times of about 100 ms (Welford 1976), so that speeds of approximately 10 actions per second and higher involve virtually exclusively pre-programmed actions (Pressing 1984*a*). An informal analysis of jazz solos over a variety of tempos supports this ball-park estimate of the time limits for improvisational novelty.

**Timing and movement invariants** Up to this point very little has been said about the timing of skilled performance, yet it is obviously a vital point. Considerable experimental work in the domains of fluent speech (Huggins 1978), typing (Shaffer 1978; Terzuolo and Viviani 1979), handwriting (Denier van der Gon and Thuring 1965; Viviani and Terzuolo 1980; Hollerbach 1981), generalized arm trajectories (Morasso 1983), and piano performance (Shaffer 1980, 1984) has established that invariant timing and spatial sequences, strongly suggestive of schemata, underlie skilled actions. Such performance rhythms, or 'hometetic' behaviour, as some have termed it, shows great tuneability: over wide variations in distance and overall time constraints, invariance of phasing and accelerations (equivalently, forces) can be observed (Schmidt 1983). By phasing is meant the relative timings of component parts of the entire movement sequence.

But it is also true that the relative timings of movement components can be changed intentionally, at least to a considerable degree. Hence the improviser has access to generalized action programmes (in both motor and music representation), which allow overall parametric control (time, space, force) *and* subprogram tuneability. This may well be responsible for the flexibility of conception characteristic of experienced improvisation.

**Motor memory** It has often been suggested that a distinct form of memory for action, called motor memory, exists. The subjective impression of improvisers (and other performers) is certainly that potentially separate yet often interconnected motor, symbolic, and aural forms of memory do exist. For a review of this extensive topic and its relationship to verbal memory the reader may wish to consult Laabs and Simmons (1981).

*Skill development*

All skill learning seems to share certain common features. In the early stages, a basic movement vocabulary is being assembled and fundamental perceptual distinctions needed for the use of feedback are drawn. In intermediate stages, larger action units are assembled, based on stringing together the existing movement vocabulary in accordance with the developing cognitive framework. These action units begin to enable predictive open-loop response. The ability to perceive distinctions is refined considerably, and internal models of action and error correction are developed. Expressive fluency begins to appear, characterized by a feeling of mindful 'letting go' (Schneider and Fisk 1983; Pressing 1984*a*). By the time advanced or expert stages have been reached, the performer has become highly attuned to subtle perceptual information and has available a vast array of finely timed and tuneable motor programmes. This results in the qualities of efficiency, fluency, flexibility, and expressiveness. All motor organization functions can be handled automatically (without conscious attention) and the performer attends almost exclusively to a higher level of emergent expressive control parameters.

In the case of improvised music these emergent control parameters are notions such as form, timbre, texture, articulation, gesture, activity level, pitch relationships, motoric 'feel', expressive design, emotion, note placement, and dynamics. There must also be a developed priority given to auditory monitoring over kinaesthetic and especially visual monitoring. This idea is supported by research on typists (West 1967), which showed that the dominant visual control used for optimal results in early stages of learning to type gave way later to reliance on tactile and kinaesthetic cues. It also seems likely that sensory discrimination and motor control functions make increasing use of higher order space–time relationships (velocity, acceleration) as skill learning progresses (Marteniuk and Romanow 1983).

The change from *controlled* processing to *automatic* motor processing as a result of extensive skill rehearsal is an idea of long standing (James 1890; Shiffrin and Schneider 1977), and it undoubtedly improves movement quality and integration (Eccles 1972). The accompanying feeling of automaticity, about which much metaphysical speculation exists in the improvisation literature, can be simply viewed as a natural result of considerable practice, a stage at which it has become possible to completely dispense with conscious monitoring of motor programmes, so that the hands appear to have a life of their own, driven by the musical constraints of the situation (Bartlett 1947; Welford 1976; Pressing 1984*a*). In a sense, the performer is played by the music. The same thing happens with common actions like walking and eating. As Welford (1976) has cogently pointed out, automaticity is therefore especially likely when the actions involved are always, or virtually always, accurate to within the

requirements of the task. Hence automaticity in improvisation can be frequent in both free and highly structured contexts, since task requirements are often self-chosen, but is more likely to be successful in musical terms for the less experienced player towards the free end of the specturm.

Schneider and Fisk (1983) have proposed an interesting corollary to the above, based upon a classification of tasks into those requiring *consistent* or *varied* processing: 'Practice leads to apparently resource free automatic productions for consistent processing but does not reduce (attentional) resources needed for a varied processing task.' (p. 129) This idea is appealing and perhaps widely valid, but is too simple to encompass the full complexity of improvisation. For part of the result of extensive practice of improvisation is an abstraction to greater and greater generality of motor and musical controls to the point where highly variable, often novel, *specific* results can be produced based on the automatic use of *general*, highly flexible and tuneable motor programmes. More irrevocable constraints causing attentional loading seem to be timing and interhand co-ordination (Pressing 1984*a*).

Another relevant area is the optimum distribution and nature of practice. Generalizations here are particularly hazardous (Newell 1981) and I will confine my comments specifically to improvisation.

The extremes of massed and distributed practice typically have complementary functions for the improviser. Distributed practice develops immediacy, and consistency of results under variable conditions, whereas massed rehearsal, by perhaps bringing to the player's awareness otherwise unperceived repetitive aspects of his or her music, enables the transcendence or improvement of stale musical design. One is reminded of the opinion of master trumpeter Miles Davis that his sidemen only really got loose in the last set of the night, after they had used up all their well-learned tricks (Carr 1982).

Variability of practice conditions is vital for improvisation, for obvious reasons, and this seems to be true of nearly all skilled behaviour (Schmidt 1983). Mental practice away from the instrument can be important for performers of fixed music, based on internal hearing of scores, but there seems very little record of its use in improvisation. This is presumably due to the intrinsically vital motoric link between performer and instrument for improvisation.

Techniques used by musicians to teach improvisation will be described below. However, some general principles of skill teaching are pertinent here. The successful yet contrasting approaches of the 'discovery' method and structural prescription (the use of instructions or demonstrations) may be mentioned. The basic trial-and-error idea of the discovery method probably requires little explanation; it has often been used as an industrial training procedure, where learning sessions are arranged so that trainees must make active choices which are normally correct, and which therefore

do not lead to ingrained errors (Welford 1976). Less formalized self-discovery techniques are certainly characteristic of much learning in the arts. But structural prescription is also a vital part of skill learning. For all but very simple skills, instructions seem particularly effective when kept simple, and when focusing on goals and general action principles rather than kinematic details (Hendrickson and Schroeder 1941; Holding 1965; Newell 1981). This certainly holds for improvisation. Probably too much intellectual detail both interferes with the fluid organization of action sequences, as mentioned earlier, and strains attentional resources.

**Studies and theories of musical improvisation**

A cognitive overview of much of this literature has been given earlier (Pressing 1984*a*, which includes references to dance and theatre), and will not be repeated here. Historical surveys of improvisation in Western music may be found in Ferand (1938, 1961), *The new Grove dictionary of music* (1983), and Pressing (1984*b*,*c*). These deal primarily with the period to 1900. Discussion of avant-garde improvisation since 1950 is included in Cope (1984). Non-Western musical improvisation is described by Reck (1983), Datta and Lath (1967), Wade (1973), Jairazbhoy (1971), and Lipiczky (1985) for Indian music; by Nettl and Riddle (1974), Nettl and Foltin (1972), Zonis (1973), Signell (1974, 1977), and Touma (1971) for various Middle Eastern traditions; by Béhague (1980) for Latin American musics; by Hood (1971, 1975), Sumarsam (1981) for gamelans and other stratified ensembles in Southeast Asia, and by Jones (1959) and Locke (1979) for Ewe music of Ghana. Park (1985) has described the improvisation techniques of Korean shamans, Avery (1984) structure and strategy in Azorean–Canadian folkloric song duelling, and Erlmann (1985) variational procedures in Ful'be praise song. Nettl (1974) has provided thoughtful general insights from the perspective of the ethnomusicologist.

In the twentieth century, prescriptive teaching texts on Western music improvisation are legion. Few, however, have the sorts of cognitive insights useful in model building, and almost all are concerned with the specifics of jazz (a small related number with blues and rock) or keyboard (particularly French-tradition organ) improvisation. The jazz texts are too numerous to survey fully here and are in any case mostly quite repetitious. Important perspectives are however given by Coker (1964, 1975), Schuller (1968), Baker (1969), Owens (1974), Liebman *et al.* (1978), Dobbins (1978), Howard (1978), Murphy (1982), and Radano (1985). Among the better organ and piano texts may be mentioned the works of Dupré (1925/37), Schouten (no date given), Gehring (1963), Berkowitz (1975) and Weidner (1984). Analytical and prescriptive texts which stand apart from the typical stylistic conventions above are the works of Bailey (1980),

Bresgen (1960), Sperber (1974), Stumme (1972), and Whitmer (1934). Except for Bailey, all of these take tonal music as their primary area of discourse. Discussions which emphasize free improvisation often take a more cognitive approach, but their usefulness is sometimes compromised by vagueness or subjectivity. Valuable readings in this area include Silverman (1962), Jost (1974), Parsons (1978), Bailey (1980), and special issues of *Perspectives of new music* (Fall–Winter 1982/Spring–Summer 1983, 26–111), the *Music educator's journal* (1980, **66**, (5), 36–147), *Keyboard* (1984, **10**(10)), and *The British Journal of Music Education* (1985, **2**(2)). Other works of interest are those on choir improvisation (Ehmann 1950, Ueltzen 1986), silent-film accompaniment (Hanlon 1975), dulcimer improvisation (Schickhaus 1978), and percussion gestures (Goldstein 1983).

Musical improvisation has also been considered as a vehicle for consciousness expansion and the tapping of deep intuitions. A full history of this 'transpersonal' approach would go back thousands of years to the sacred texts of many religions. Here I only survey recent Western opinion. Hamel (1979) has intelligently chronicled music of the avant-garde (for example Riley, Stockhausen) from this perspective. Laneri (1975) has developed a philosophy of improvisation based on different states of consciousness, featuring the concepts of synchronicity and introversion. The resultant music is primarily vocal, since the voice is considered the primal instrument. A powerful system of sonic meditation most applicable to local improvisation groups has been developed by Oliveros (1971). 'Sensing' compositions have been published by Gaburo (1968). An attempt to connect music, altered states of consciousness, and research in parapsychology has been given by Pressing (1980), while Galás (1981/82) has created a primal vocal music based on obsession, excessive behaviour, and trance states of severe concentration.

The approaches in the literature to the teaching of improvisation may be broadly grouped as follows. First, there is the perspective overwhelmingly found in historical Western texts, that improvisation is real-time composition and that no fundamental distinction need be drawn between the two. This philosophy was dominant in pre-Baroque times but had become rare by the eighteenth century. In practice this results in a nuts-and-bolts approach with few implications for the modelling of improvisation beyond basic ideas of variation, embellishment, and other traditional processes of musical development. A second approach, which historically took over as the first one waned, sets out patterns, models, and procedures specific to the improvisational situation, which, if followed by those possessing a solid enough level of musicianship, will produce stylistically appropriate music. In this category fall the many figured bass and melodic embellishment texts of the seventeenth and eighteenth centuries (for example Mersenne 1635; Quantz 1752/1966; Bach 1778/1949; Arnold 1965), as well as the riff

compendia and how-to-do-it books in the field of jazz (such as Coker, *et al.* 1970; Slonimsky 1975; Nelson 1966).

A third technique is the setting of a spectrum of improvisational problems or constraints. The philosophy behind this technique shows a clear contrast with the second approach above, as described by Doerschuk (1984), referring to the Dalcroze system.

> The art of improvisation rests on . . . a developed awareness of one's expressive individuality. This knowledge grows through interactive exercises with a teacher, whose function is not to present models for imitation, but to pose problems intended to provoke personal responses. (p. 52)

Jaques-Dalcroze (1921) seems to have pioneered this approach in our century with a revealing series of improvisation exercises for piano. These include composition-like problems in rhythm, melody, expressive nuance, and harmony; muscular exercises; imitation of a teacher; exercises in hand independence; the notation of improvisation just after performing it; and what may be termed an 'interrupt' technique. In this last technique the word 'hopp' is recited by the teacher, as a cue for the student to perform pre-set operations such as transposition or change of tempo during the performance. This technique is reminiscent of a much later suggestion by Roads (1979) that musical grammars used in improvisation might be 'interrupt-driven'. This idea is developed in the model below.

Parsons (1978) has made effective use of this third technique in a collection of short pieces by many different composers defined largely by improvisational instruction sets; he also presents a taxonomy of psycho-improvisational faults and recommended exercises for correcting them. A shorter multi-author collection of improvisational exercises is found in Armbruster (1984). Jazz fake books like the *Real book* (no listed authors or dates) or *The world's greatest fake book* (Sher 1983) may also be considered to act along the lines of this technique.

A fourth approach is the presentation of multiple versions of important musical entities (most commonly motives) by the teacher, leaving the student to infer completely on his or her own the ways in which improvisation or variation may occur by an appreciation of the intrinsic 'fuzziness' of the musical concept. This imitative self-discovery approach is found in the Persian *radif*, which is a repository of musical material learned in a series of increasingly complex versions by the aspiring performer (Nettl and Foltin 1972), and in Ghanaian traditions (K. Ladzekpo, personal communication), for example. A related procedure made possible by the use of recording technology in the twentieth century is for the student to directly copy a number of improvised solos by repeated listening to recordings, and from this extract common elements and variation procedures. Song-form based improvisations, in which solos consist of a number of choruses which repeat the same underlying chord

progression, are particularly suitable. This method has been widely used in jazz and blues since the end of the First World War.

A fifth approach is allied to the self-realization ideas of humanistic psychology. It is based on concepts of creativity and expressive individuality which go back in music explicitly at least to Coleman (1922), implicitly certainly to Czerny (1829/1983), and probably in a general sense at least to the Enlightenment. Important educational applications of this idea are found in the works of Carl Orff, Zoltán Kodály, Suzuki (see Mills and Murthy 1973), and particularly Jaques-Dalcroze (1976, 1930) and Shafer (1969). In the words of Jaques-Dalcroze,

> Improvisation is the study of direct relations between cerebral commands and muscular interpretations in order to express one's own musical feelings . . . Performance is propelled by developing the students' powers of sensation, imagination, and memory.
>
> (In Abramson 1980, p. 64.)

Little actual research on optimal techniques for teaching improvisation has been carried out. The important study by Partchey (1973) which showed the value of models and particularly of subsequent aural feedback in learning to improvise has already been mentioned above. Work by Hores (1977) has shown that visual and aural approaches to the teaching of jazz improvisation can be equally effective. Burnsed (1978) looked at the efficacy of design of an introductory jazz improvisation sequence for band students. Seuhs (1979) developed and assessed (by adjudication) a course of study in Baroque improvisation techniques. Bash (1983) compared the effectiveness of three different instructional methods in learning to improvise jazz. Method I was a standard technical procedure based on scales and chords. Method II supplemented this technical dimension with aural perception techniques which included rote vocal responses to blues patterns, blues vocalizations, and instrumental echo response patterns based on rote or procedures of generalization. Method III supplemented the same technical procedures of Method I with a historical–analytical treatment. All three methods gave improved results over that of a control group, and methods II and III, though no significant difference was found between them, were both superior to method I. The results show the value of specific theoretical and technical instruction, and also of its supplementation by relevant aural training or analyses of performance strategies used by virtuoso improvisers.

One final comment on improvisation teaching seems apposite. This is the fact that the optimally effective teacher is able to direct evaluative comments on several different levels. One is the technical—'Your notes don't fit the chord', 'The piano is lagging behind the bass', etc. Another is the compositional—'Try to develop that motive more before discarding it', 'Use more rhythmic variety in pacing your solo', 'Musical quotations seem

inappropriate in this free a context', etc. Yet another level is the use of organizing metaphor, a vital part of the tradition of jazz teaching—'Use more space', 'Dig in', 'Go for it', 'Play more laid-back', 'Don't force it—follow the flow', etc. Simple comments of this kind can be remarkably effective at removing improvisational blocks, when delivered at a proper time.

Pike (1974) has presented a brief but insightful phenomenology of jazz. His approach considers the projection of 'tonal imagery' to be the fundamental process in jazz improvisation. Tonal imagery is either 'reproductive' (memory-based) or 'productive' (creative). The improviser operates in a 'perceptual field' which acts as a framework in which the improviser's imagery appears and originates. This field includes not only the perception of external tonal events, but the perception of internal images, as well as the states of consciousness evoked by these images. Images in this field are combined, associated, contrasted, and otherwise organized. The phenomenological operations describing this are processes such as repetition, contrast, continuity, completion, closure, and deviation. Other aspects of improvisation defined by Pike include 'intuitive cognition', an immediate penetration into the singular and expressive nature of an image, and 'prevision', a glimpse into the developmental horizons of an embryonic jazz idea.

Although some of Pike's claims are open to question, for example his uncritical acceptance of concepts like Hodeir's 'vital drive' (Hodeir 1956), his short paper remains an important introspective analysis of the experience of improvisation. The only other extensive phenomenological treatment of improvisation seems to be Mathieu's (1984) study of musician/dancer duo performances. Other perspectives on the experiences of the improviser have been given by Milano (1984), in an interview with jazz pianist/psychiatrist Denny Zeitlin, and Sudnow (1978), who has produced a basic ethnomethodological description of learning to play jazz on the piano. Related philosophical issues have been raised by Alperson (1984) and Kleeman (1985/86).

Finally it may be proper to note that the computer age has spawned new hybrids of composition and improvisation. Fry (1980, 1982/83) has described music and dance improvisation set-ups using computer sensing and control devices. Chadabe (1984) has described a method of 'interactive composition' whereby movements of the hands in space near two proximity-sensitive antennas trigger and exert partial control over real-time computer sound generation. Interactive computer-based performance systems have also been used by trombonist George Lewis and a host of 'performance artists', including this writer. And recently available commercial software, such as the Macintosh-based *M* and *Jam Factory*, has an interactive improvisational component that seems rich with promise.

## Oral traditions and folklore

The idea that traditional folk-tales from many cultures have underlying unities, which may be interpreted as narrative grammars, is a fairly well-established one (Propp 1927; Thompson 1946; Nagler 1974). Explanations of this fact have tended towards one or the other of two viewpoints.

A common (particularly European) perspective in the study of oral tradition and folklore has been a focus on their repetitive and imitative aspects, with the frequent assumption of an *Urtext* which has undergone historical and geographic transformation. A powerful opposing view, and one which seems increasingly relevant as a description of referent-based improvisation, is found in the 'formulaic composition' proposals of Milman Parry and Albert Lord (Parry 1930, 1932; Lord 1964, 1965).

Formulaic composition was derived from Milman's intense study of the Homeric epics, particularly the *Odyssey*, and given further support by research on Yugoslav folk-epic poetry conducted by Milman and Lord. It is also considered to be applicable to other oral epics such as *Beowulf* and the *Chanson de Roland*, and has been used to analyse Latvian folk-song texts (Vīķis-Freibergs 1984). In this view epic oral poetry is created anew at each performance by the singer from a store of formulas, a store of themes, and a technique of composition. There is no 'original' version; instead the tradition is multiform. A 'formula' is a group of words regularly employed under the same metrical conditions to express a given essential idea; it has melodic, metric, syntactic, and acoustic dimensions. By choosing from a repertoire of roughly synonymous formulas of different lengths and expanding or deleting subthemes according to the needs of the performance situation, the experienced performer is able to formulaically compose (in real-time, hence improvise) a detailed and freshly compelling version of a known song epic. As a result of the composition system, instances of pleonasm and parataxis are common.

The formulas considered as a group reveal further patterns. In the words of Lord (1964): 'the really significant element in the process is . . . the setting up of various patterns that make adjustment of phrase and creation of phrases by analogy possible' (p.37). In addition, the permutation of events and formulas may occur, as well as the substitution of one theme for another.

Yet the traditional singer does not seek originality with this technique, but heightened expression. Lord speculates that formulas originally grew out of a need for intensification of meaning or evocation. 'The poet was sorcerer and seer before he became artist' (Lord 1964, p. 67).

The relevance of formulaic composition to specific types of musical improvisation has recently been discussed by several writers. Treitler (1974) has argued that Gregorian chant was composed and transmitted in an analogous process to that used in the oral epics. Smith (1983) has used

the process to describe the constraints imposed on the song-based jazz performer, and has gone on to analyse piano improvisations by Bill Evans. Kernfeld (1983) has examined how far formulas may be used to describe the music of saxophonist John Coltrane. Reck (1983) has produced the evocative idea of a musician's 'tool-kit', in a mammoth study of five performances by South Indian musician Thirugokarnam Ramachandra Iyer. The tool-kit is considered to be piece-specific and to contain both individually chosen and culturally determined formulas, musical habits, models of improvisational and compositional forms, aesthetic values, and social attitudes.

The application of Parry–Lord theory to musical improvisation is thus a clear contemporary trend. The limits of its validity and usefulness are still open questions, and are probably linked to whether a satisfactory agreement can be reached on the principles to be used to define musical 'formulas'.

**Intuition and creativity**

These are two related concepts, each with a vast literature. Their connection with improvisation is undeniable, yet explicit mention of it in either field is rare. On the other hand, 'free' musicians and many music educators commonly use the two terms, but often without a very clear notion of just what is being discussed. This section attempts to bridge that gap.

The concept of intuition is much older than creativity, and it has separate philosophical and psychological traditions. Westcott (1968) has provided an excellent general survey, enumerating three historical approaches to philosophies of intuition. First comes Classical Intuition (for example Spinoza, Croce, Bergson), which views intuition as a special kind of contact with a prime reality, a glimpse of ultimate truth unclouded by the machinations of reason or the compulsions of instinct. Knowledge gained through this kind of intuition is unique, immediate, personal, unverifiable. The second approach, called by Westcott Contemporary Intuitionism (for example Stocks 1939; Ewing 1941; Bahm 1960), takes the more restricted view that intuition is the immediate apprehension of certain basic truths (of deduction, mathematical axioms, causality, etc.). This immediate knowing stands outside logic or reason and yet is the only foundation upon which they can be built. Knowledge gained through intuition constitutes a set of 'justifiable beliefs', which are nevertheless subject to the possibility of error. A third approach is positivistic (for example Bunge 1962) in that it rejects as illusory both the notions of immediacy and ultimate truth found in some earlier views. Rather, an intuition is simply a rapid inference which produces a hypothesis.

Of all these views, it is perhaps that of French philosopher Henri

Bergson (1859–1941) which shows the greatest affinities with the common metaphors of improvisation. Bergson saw intuition as a way to attain direct contact with a prime reality ordinarily masked from human knowledge. This prime reality is an ongoing movement, an evolving dynamic flux which proceeds along a definite but unpredictable course.

> The prime reality is referred to as 'the perpetual happening' or 'duration'. The mind of man, according to Bergson, is shielded from the perpetual happening by the intellect, which imposes 'patterned immobility' on prime reality, distorting, inmobilizing, and separating it into discrete objects, events and processes. In the perpetual happening itself, all events, objects, and processes are unified'
>
> (In Westcott 1968, p. 8).

In Bergson's view, the intellect can freely interact with the fruits of intuition (special knowledge and experience) to develop an enriched personal perspective.

The notion of tapping a prime reality is very similar to the improviser's aesthetic of tapping the flow of the music, as mentioned above. The same apparent process has been eloquently described with regard to the origins of folk-tales from many cultures by English writer Richard Adams:

> I have a vision of—the world as the astronauts saw it—a shining globe, poised in space and rotating on its polar axis. Round it, enveloping it entirely, as one Chinese carved ivory ball encloses another within it, is a second . . . gossamer-like sphere . . . rotating freely and independently of the rotation of the earth.
>
> Within this outer web we live. It soaks up, transmutes and is charged with human experience, exuded from the world within like steam or an aroma from cooking food. The story-teller is he who reaches up, grasps that part of the web which happens to be above his head at the moment and draws it down—it is, of course, elastic and unbreakable—to touch the earth. When he has told his story—its story—he releases it and it springs back and continues in rotation. The web moves continually above us, so that in time every point on its interior surface passes directly above every point on the surface of the world. This is why the same stories are found all over the world, among different people who can have had little or no communication with each other.
>
> (Adams 1980, p. 12.)

There is a clear convergence of imagery in this and other descriptions that points to a likely transpersonal component to improvisation.

The psychological perspectives on intuition are many and varied, but only two seem relevant here. The first is the widely occurring idea that intuition is a special case of inference which draws on cues and associations not ordinarily used (Westcott 1968). A similarity with certain theories of skill learning mentioned above is apparent. A second and wide-ranging approach is found in the recent work by Bastick (1982), which includes a search of over 2.5 million sources for common properties underlying intuition. After the identification and detailed analysis of some 20 of these

properties, Bastick ends up describing intuition as a combinatorial process operating over pre-existing connections among elements of different 'emotional sets'. These emotional sets apparently contain encodings, often redundant, of many different life events (intellectual activities, movement, emotion, etc.). By giving strong emphasis to the role of dynamics, bodily experience, and the maximizing of redundancy in encoding, and by a series of suggestive diagrams of intuitive processing, Bastick seems to be on an important track parallel to emerging ideas of improvisation.

Research in creativity is probably more extensive than that in intuition, for intuition is most commonly considered a subcategory of creativity. Creativity research in music education has been recently surveyed by Richardson (1983). The only clear relations to improvisation she found were in specialized educational methods and a growing tendency to use improvisation tests in assessing musical creativity. Vaughan (1971), Gorder (1976), and Webster (1977) have designed and implemented such tests, but results show uneven patterns of correlation between general intelligence, creativity, musicality, composition, and improvisation, and seem to have no clear consequences for improvisation modelling.

General studies of creativity abound, and follow many divergent paths. Two alone seem relevant here. Guilford's Structure-of-Intellect (SI) model proposed a taxonomy of factors of intelligence (Guilford and Hoepfner 1971 (and earlier references mentioned therein); Guilford 1977). These intelligence factors, which number 120, are classified along three dimensions:
*thought content:* visual, auditory figural, semantic, symbolic, and behavioural information;
*kinds of operation performed on the content*: cognition, memory, convergent production, divergent production, evaluation;
*products* (the results of applying operations to content): units, classes, relations, systems, transformations, and implications.
These classifications are related to improvisation in a general way, but despite their intuitive appeal, they have so far been fairly resistant to empirical verification.

Guilford further defined a set of six aptitudes for creative thinking: fluency, flexibility, originality, elaboration, redefinition, and sensitivity to problems. Torrance (1966) used this same set in designing a more open-ended approach to the testing and definition of creativity. Some of these six aptitudes are identical to the ones found in skilled performance above; they are considered here to be further guidelines for testing the plausibility of improvisational modelling.

Finally, Guilford and Hoepfner classified techniques of evaluation (in problem-solving), which they held to be due to appeals to logical consistency, past experiences, feeling of rightness, or aesthetic principles. Such a classification also has implications for improvisation (see model below).

**Artificial intelligence**

This field is concerned with programming computers to be intelligent problem solvers. The framework of action is usually formulated in terms of a problem space which must be searched for correct solutions. Since interesting problem spaces are nearly always too large to be investigated completely, a major focus of the field is the design of better heuristic search techniques. Coupled naturally with this are many methods and frameworks for the representation of knowledge.

There is traditionally no explicit mention of improvisation in the field. In making such a link, it seems clear that the successful application of artificial intelligence concepts to improvisation rests to a large degree on the appropriateness of considering improvisation to be a kind of problem solving. There is little doubt that such an analogy can be fruitful, particularly for referent-guided improvisation. For example, the process of improvisation may be divided up into a number of time points, and viewed as a succession of small problems, each of which is the production of an appropriate chunk of musical action at the current time point, where the constraints on action are the referent, goals, and musical actions at earlier time points. Alternatively, the time-scale may be drawn much coarser, and each complete improvisation may be considered a solution to a much more generally stated problem: for instances, improvise a chorus on 'I Got Rhythm' changes, within the constraints of be-bop style.

Before surveying the fruits of this approach it may be wise to spell out its limitations. Experientially, improvisation can seem to be far removed from problem solving. This is particularly so where the goals of the music making are exploration and process, rather than the presentation of artistic product. It is also very difficult to imagine how one could ever specify the 'problems' in freer types of improvisation with sufficient detail to allow specific artificial intelligence techniques to be used in modelling. Such problem formulations, even if possible, would be very personal, open ended, and sometimes contradictory.

With these provisos, we examine how various artificial intelligence problem-solving techniques might apply to improvisation. Search techniques come in several variants, including depth-first, breadth-first, and best-first. All use a generate-and-test procedure to find solutions to a problem. Clearly there are possible connections with improvisation. Generate-and-test could be applied to learning to improvise, where generation is sound production and testing is listening to generated music; or, it could describe internal cognitive selection processes, where testing is based on internal hearing of generated possibilities, before one is chosen as the actual musical output at a given time. Unfortunately with regard to this second interpretation there is a serious limitation: the inevitable use of back-tracking in the search processes cannot be very significant in improvisation

due to the cognitive limitations of real-time processing. The need of the improviser is for a good solution, not the best, for there is probably no single 'best' solution, and even if there were, it would take too long to find it. Therefore, the number of solution paths compared at any one step is probably very strongly limited, perhaps to two or three.

Another problem-solving technique is problem reduction: that is, reducing a problem to a set of subproblems. This is a common way to look at the teaching of improvisation, but seems less likely to apply to doing it, where integration of action is required. Of course there is no proof of this; we know far too little about the workings of the brain. Constraint satisfaction, on the other hand, is a technique whose principles seem to apply to improvisation. The constraints are the referent, goals of the performer, stylistic norms, etc. Finally, means-ends analysis is a technique that is based on comparing current and goal states. Because it involves considerable back-tracking, it is unlikely to apply to the improvisation process. Yet like other methods above, it seems relevant to the process of learning improvisational skill. In general, then, learning to improvise (that is, to structure musical impulses within aesthetic guidelines) is more like problem solving than is improvising itself.

Another main branch of artificial intelligence is knowledge representation. The relevance to improvisation seems clear, for any particular mode of knowledge representation makes it efficient to do certain things and inefficient to do others. And efficiency is what the improviser needs above all.

Knowledge representation in artificial intelligence is based on many ideas, including indexing, conceptual dependency, hierarchies, semantic nets, multiple representation, blackboards (actually a type of interprocess communication), frames, scripts, stereotypes, and rule models (Rich 1983; Lenat 1984). With respect to improvisation, many of these are more suggestive than readily applicable. Indexing, for example, is too artificial, whereas conceptual dependency, in which information is represented by certain conceptual primitives, is too strongly linked with natural language structure. Hierarchies have been discussed previously. Semantic nets are perhaps more promising: information is represented as a network of nodes connected to each other by labelled arcs, each node representing an object, event, or concept, and each arc a relation between nodes. Such a graph could be drawn for musical objects and events, but parametrically tuneable processes are not easy to represent, and this is a serious drawback.

Multiple representation, however, is an important idea, and one which is implicit in parallel-processing ideas mentioned earlier. The increased flexibility and efficiency possible with multiple representation argue very strongly for its inclusion in any model of improvisation. Gelernter (1963) successfully applied the idea to problems in plane geometry by using simultaneous axiomatic and diagrammatic representations. Another interesting application is the notion of the 'blackboard', an organization of the

problem space into multiple levels of representation, typically along a dimension indicating level of abstractness. Thus a spoken sentence may be processed at levels of acoustic wave form, phonemes, syllables, words, word sequences, phrases, etc. Each part of the blackboard is triggered automatically as relevant information comes in. Multiple representation also strengthens the possibilities for analogy, and promotes synergy, by which is meant the co-operative action of parts of a complex system (Lenat 1984).

The last four ideas mentioned above, frames, scripts, stereotypes, and rule models, are considered to be various types of schemata (Rich 1983). The use of the word here is slightly different from that in the area of motor behaviour (see Adams 1976 for a survey). Frames are used to describe collections of attributes of an object. A frame consists of slots filled with attributes and associated values. Like most slot-and-filler structures, frames facilitate the drawing of analogies. Ideas equivalent to the frame are found in the improvisation model below. Scripts are simply normative event sequences and in so far as they apply to improvisation have much in common with the generalized motor schemata described above. Stereotypes have their usual meaning and are parts of the norms of musical style, but are often avoided by the best improvisers. Rule models describe the common features shared by a set of rules which form the basis for a 'production system'. If the improvising musician is the production system, the important rules will be largely heuristic and the rules about rules may be termed metaheuristics. Some of these will be culturally and historically based, while others presumably reflect intrinsic properties of the cognitive apparatus. Serafine (1983) has presented an insightful discussion of this distinction from the standpoint of the cognitive psychologist.

In principle it should be possible to integrate appropriate artificial intelligence techniques to construct an expert system which improvises. One of the very few such attempts is the unpublished work of Levitt (1981), which dealt with jazz improvisation. The idea awaits further development.

## A model of improvisation

Any theory of improvisation must explain three things: how people improvise; how people learn improvisational skill; and the origin of novel behaviour. It must also be consistent with the numerous recurring themes reviewed above. The model given here seems to satisfy these conditions.

### How people improvise

The first part of this model describes the process of improvisation. It begins with the observation that any improvisation may be partitioned into a

sequence of non-overlapping sections. By non-overlapping it is simply meant that sounds are assigned to only one section, not that the sounds themselves do not overlap. Let each of these sections contain a number of musical events and be called an event cluster $E_i$. Then the improvisation I is simply an ordered union of all these event clusters. Formally,

$$\mathrm{I} = \{E_1, E_2 \ldots E_n\} \tag{1}$$

From a naïve analytical perspective there is a large number of ways such a partitioning could be made. Our first major assumption is that every improvisation is actually generated by triggers at specific time points $t_1, t_2 \ldots t_n$ that instigate the movement patterns appropriate to effect intended musical actions. Each time point is thus the point at which decided action begins to be executed. Note that it is schemata for action that are triggered, not precise movement details, and subsequent motor fine tuning based on feedback processes goes on after each time point. Often time points will have clear musical correlates, with adjacent event clusters being set off from each other by local musical boundary criteria; pauses, phrase junctures, cadences, grouping by sequence etc.; but this need not always be the case.

With this interpretation, equation (1) is a unique specification of the timing of central decision making made by the improviser. The improvisation may then be viewed as a series of 'situations', where the $(i+1)$th situation is confined primarily to the time interval $(t_i, t_{i+1})$ and entails the generation of the cluster $E_{i+1}$ on the basis of the previous events $\{E_1, E_2, \ldots E_i\} \equiv \{E\}_i$, the referent $R$ (if one exists), a set of current goals $\mathcal{G}$, and long-term memory $M$. The referent $R$ is an underlying piece-specific guide or scheme used by the musician to facilitate the generation of improvised behaviour (Pressing 1984*a*). The process of event-cluster generation may then be written

$$(\{E\}, R, \mathcal{G}, M)_i \rightarrow E_{i+1}. \tag{2}$$

Decision-making in the $(i+1)$th situation may in principle extend well back before time $t_i$, depending on the degree of pre-selection used by the performer, and will also extend slightly into the future, in that fine details of motor control will be left to lower control centres and hence may occur after $t_{i+1}$.

Equation (2) applies strictly only to solo improvisation. The only changes with group improvisation are that, first, all performers will have their own distinct time-point sequences (even though they would often be partially correlated), and, second, players will normally interact. Equation (2) can be readily extended to apply to all $K$ members of an improvisation ensemble by writing

$$(\{E\}, C, R, \mathcal{G}, M)_{i_k} \rightarrow E_{i_k+1}, \; k=1, \ldots K, \tag{3}$$

where subscripts refer to the $k$th performer, and $C$ stands for performer $k$'s cognitive representation of all previous event clusters produced by the other performers and any expectations of their likely future actions. For simplicity, we use the formalism of equation (2) and speak primarily in terms of solo improvisation in what follows, adding in the effects of other performers in a straightforward manner as needed at certain points.

Any given event cluster $E$ has a number of simultaneously valid and partially redundant 'aspects'. Each aspect is a representation of $E$ from a certain perspective. Most important are the acoustic aspect (produced and sensed sound), the musical aspect (cognitive representation of the sounds in terms of music-technical and expressive dimensions), and the movement aspect (including timing of muscular actions, proprioception, touch, spatial perception, and central monitoring of efference). Visual and emotional aspects normally also play a role, and in principle there may be others. Furthermore each aspect exists in two forms, intended and actual. Each intended form is specified at a specific time point: the corresponding actual form is constructed from subsequent sensory feedback. The gap between these two forms is reduced by sound training in musicianship and improvisation practice, but it never dwindles completely to zero. Hence in equation (2) or (3) the variable $\{E\}_i$ represents intended and actual forms of all aspects of event clusters $E_1$ to $E_{i-1}$, the intended form of $E_i$, plus, over the course of the time interval $(t_i, t_i+1)$, increasing feedback on the actual form of $E_i$. By $t_{i+1}$, when central commands for $E_{i+1}$ are transmitted, the ongoing nature of improvisation probably demands that integration of the intended and actual forms of $E_i$ be virtually complete.

The details of the proposed model of what occurs in the $(i+1)$th situation, that is, the selection of $E_{i+1}$, are as follows:

(A) $E_i$ is triggered and executed (it may spill on briefly to times $t>t_{i+1}$).

(B) Each aspect of $E_i$ may be decomposed into three types of analytical representation: objects, features, and processes. An 'object' is a unified cognitive or perceptual entity. It may, for example, correspond to a chord, a sound, or a certain finger motion. 'Features' are parameters that describe shared properties of objects, and 'processes' are descriptions of changes of objects or features over time. At $t_i$ this decomposition is based only on intended information (efference); by $t_{i+1}$ much of the actual form of $E_i$, received through the senses and internal feedback, has been used to refine the cognitive representation of $E_i$. This may continue after $t_{i+1}$. Let this decomposition into objects, features, and processes (for each aspect) be represented by three variable-dimension arrays **O**, **F**, and **P**, and assume that they represent all information about $E_i$ needed by the improviser in decision making.

(C) The structures of the three types of arrays are as follows. The object array is a $2\times N$ array where row 1 components label the objects present and row 2 gives their associated cognitive strengths $s_k$ (explained below). The

feature and process arrays are typically non-rectangular. Their first rows consist of object and process labels respectively, and each column below that row is built up of a number of pairs of elements which give the values $v_{jk}$ of associated features or process parameters and their corresponding cognitive strengths $s_{jk}$. The arrays are non-rectangular because different objects may possess different numbers of significant features or process parameters. The feature and parameter process values $v_{jk}$ vary over ranges appropriate to their nature, whereas cognitive strengths $s_{jk}$ are normalized to vary between 0 and 1. Cognitive strength is essentially an indicator of attentional loading, that is, the importance that the given factor has in the performer's internal representation. Thus even though certain features may be objectively present, as analysed by others, if the player does not use them in his or her cognitive representation, their $s$ values would be zero. Sample object, feature, and process arrays for the following event cluster (a short trombone motive) are given by way of example (Fig. 7.1), for the musical aspect only. Considerable redundancy of representation has been set out in the process array.

(D) Production of $E_{i+1}$ occurs primarily on the basis of long-term factors (R, $\mathscr{G}$ , stylistic norms, and ongoing processes), and by evaluation of the effects and possibilities of $E_i$. There seem to be only two methods of continuation used: associative or interrupt generation. In associative generation the improviser desires to effect continuity between $E_i$ and $E_{i+1}$ and picks new arrays $\mathbf{O}_{i+1}$, $\mathbf{F}_{i+1}$, $\mathbf{P}_{i+1}$ whose set of strong cognitive components includes all or nearly all of the strong cognitive components of $\mathbf{O}_i$, $\mathbf{F}_i$, and $\mathbf{P}_i$, with the parameter values of these shared components being directly related (as described in (E) below). In other words the $E_i$ components with high $s$ values carry their information on in some way to $E_{i+1}$. These new arrays act as a set of constraints which determine, in conjunction with various generation processes, the musical actions generated for $E_{i+1}$. The relative importance of different constraints in the generation process is indicated by their respective cognitive strengths $s_k$ and $s_{jk}$. Note that the $E_{i+1}$ arrays may contain new strong components (constraints) that were previously weak or completely absent. In particular, it is possible to add a new independent musical process to a continuing one to produce an associative continuation which has a clear sense of novelty (e.g. the introduction of a new part in polyphonic music). In the case of interrupt generation the improviser has had enough of the event train ending with $E_i$ (for whatever reasons) and breaks off into a different musical direction by resetting a significant number of strong components of $\mathbf{O}_{i+1}$, $\mathbf{F}_{i+1}$, $\mathbf{P}_{i+1}$ without any relations to $E_i$ except possibly those chosen to be normative with regard to style in the piece, or intrinsic to the referent (if present) or goals. Clearly, the more strong components that are reset, the greater the sense of interruption.




Object array **O**

$$\mathbf{O}=\begin{pmatrix} \text{N} & \text{G} & \text{R} & \text{N} & \text{S} \\ \tfrac{1}{2} & \tfrac{1}{2} & \tfrac{1}{2} & 1 & 1 \end{pmatrix}$$

← Label
← Cognitive strength

N=note
G=glissando
R=rest
S=scale

Feature array **F**

$$\mathbf{F}=\begin{pmatrix} \text{N} & \text{G} & \text{R} & \text{N} & \text{S} \\ A^2(.25) & B^2(.5) & \text{♪}(.5) & A^3(.5) & 6^{\dagger}(.25) \\ \text{♩}(.5) & D^{\sharp 3}(.5) & & \text{♪}(.5) & A(.5) \\ \text{II}^*(.5) & \text{♪}(.5) & & \text{II}^*(.5) & \text{whole} \\ p(.1) & \text{VII}^*(.4) & & f(.5) & \text{tone}(1) \\ & \text{III}^*(.4) & & & \\ & <(.6) & & & \\ & p(.8) & & & \\ & f(.8) & & & \end{pmatrix}$$

← Label
← Values $v_{jk}$ and cognitive strengths ($s_{jk}$)

*Slide position of trombone
†Scale size

Process array **P**

$$\mathbf{P}=\begin{pmatrix} \text{Randomly select notes from scale} & \text{Follow contour} & \text{Follow interval sequence} & \text{Use trichords} \\ A(1)\ \text{whole tone}(1) & 1\ \text{octave}^*\ (.25) & (2,4,6)\ (.5) & (026)\ (1) \end{pmatrix}$$

← Parameters $v_{jk}$ and cognitive strengths ($s_{jk}$)

*Range

**Fig. 7.1.** Possible object, feature, and process arrays corresponding to a short trombone motive.

(E) Associative generation is based on either similarity or contrast. In the case of similarity all or nearly all important (important as determined from the vantage point of the improviser) array components stay approximately the same. In other words, for those components $v_{jk}$ with $s_{jk}$'s signicantly above zero at time $t=t_i$, $(v_{jk})_{t_i} \approx (v_{jk})t_{i+1}$. Significant object array components behave analogously. In the case of constrast-type associative generation, at least one strong component of either the feature or process arrays must either move from near one end of its possible range of values to near the opposite end, or cross some perceptually significant boundary. Meanwhile, all other strong components change either very

little or not at all. Examples are when a group of high notes is followed by a group of low notes, or an accelerando changes of decelerando, or bright timbres are replaced by dull timbres. The idea behind this classification is that the most powerful and general types of improvisational control are those that are cued to features and processes. The objects, though a crucial part of the entire procedure, are at the same time often merely the very familiar musical clothing of cognitive action space.

(F) Interrupt generation is based on the resetting of all or a significant number of the strong array components without regard to their values in the current event cluster $E_i$. A decision to interrupt brings to an end a sequence of related event clusters, say $K = \{E_{i-r}, E_{i-r+1}, \ldots E_i\}$, where the number of event clusters in this 'event-cluster class' is $r+1$. Hence interrupt decisions partition the entire improvisation into $A$ discontinuous event-cluster classes $K_\alpha$, so that the formal design of the piece becomes

$$I = \{K_1, K_2 \ldots K_A\}. \tag{4}$$

Each event-cluster class $K_\alpha$ contains at least one event cluster, and may be defined in terms of the strong components of the object, feature, and process arrays shared by all the member event clusters. If these special components are represented as $\mathbf{O}_s^\alpha$, $\mathbf{F}_s^\alpha$, and $\mathbf{P}_s^\alpha$, then $K_\alpha$ is defined by $(\mathbf{O}_s, \mathbf{F}_s, \mathbf{P}_s)^\alpha$. One of the sets $\mathbf{F}_s^\alpha$ and $\mathbf{P}_s^\alpha$ must be non-empty. If $(\mathbf{O}_s, \mathbf{F}_s, \mathbf{P}_s)^\alpha =$ or $\approx (\mathbf{O}_s, \mathbf{F}_s, \mathbf{P}_s)^\beta$, for some $\beta$ not immediately following $\alpha$, we have recursion in formal design of the improvisation. Under these assumptions the process of improvisation may be sketched diagrammatically as in Fig. 7.2.

(G) The choice between association and interrupt generation may be formally modelled by a time-dependent tolerance level for repetition, $L(t)$. An interrupt tester, whose inputs are presumably the time since the onset of the $K_\alpha$ event cluster class, $(t-t_{i-r})$, and the size and nature of $K_\alpha$, computes the degree of current repetition, $Z(t)$, and if $Z(t) \geqslant L(t)$, institutes an interrupt generation, so that $Z(t)$ jumps to a low value. Otherwise associative generation continues. Diagrammatically this is shown in Fig. 7.3 for the same improvisation as in Fig. 7.2.

(H) Once $\mathbf{O}_{i+1}$, $\mathbf{F}_{i+1}$, and $\mathbf{P}_{i+1}$ are selected for all relevant aspects, tuneable cognitive and motor subprogrammes are set in motion that generate, on the basis of these higher constraints and current motor positions, a specific action design. At this point we have reached $t_{i+1}$ and this loop of the process $(E_i \rightarrow E_{i+1})$ is complete. By iteration, then, the entire improvisation is built up. The starting point $E_1$ may be considered a situation of interrupt generation (where $E_0$ is silence) and the final event cluster $E_n$ is simply a second case of interrupt generation where $E_{n+1}$ = silence, after which the improvisation process is turned off.

These, then, are the salient features of the model in outline. They are diagramatically displayed in Fig. 7.4.

Next we look more deeply at certain critical stages of the improvisation

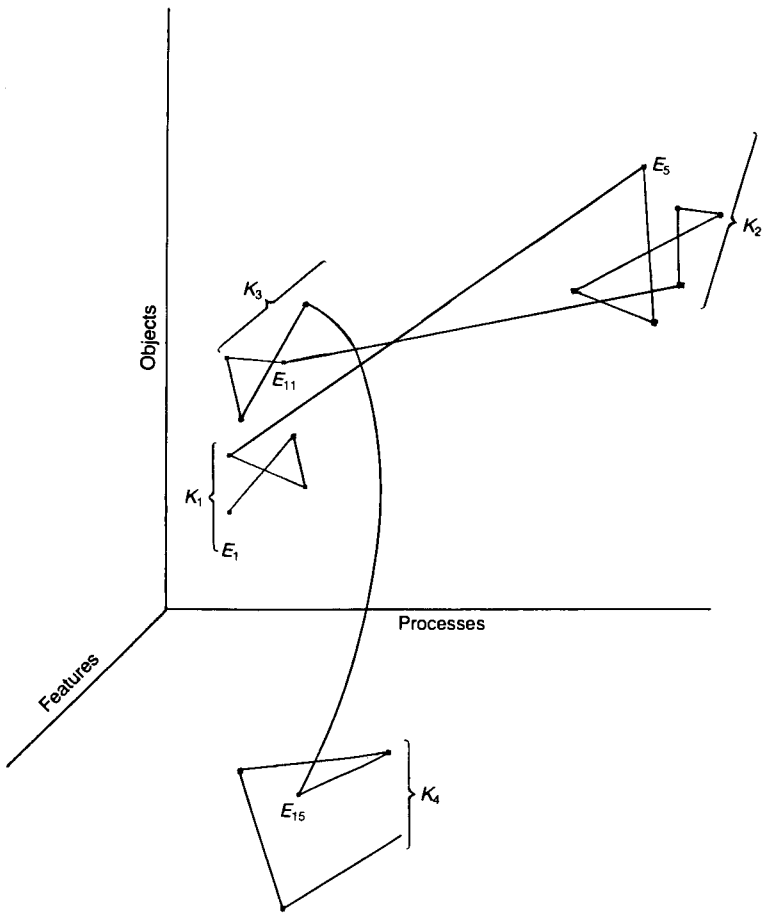



**Fig. 7.2.** An improvisation in musical action space, showing four event-cluster classes, and form ABA'C.

model. To begin with, it is characterized throughout by extensive redundancy. There is first of all redundancy between the aspects of each event cluster. The performer knows, for example, that certain motor actions involved in striking a kettle drum (motor aspect) will correspond to a particular sound (acoustic aspect), with associated musical implications (musical aspect). Furthermore, each aspect is decomposed into extensive object, feature, and process representations which contain considerable redundancy. For example, the musical motive of Example 7.1 may be pitch encoded as the objects $D^2F^2A^2B^2$, or as the object $B\phi$ diminished 7 chord in first inversion, or as a diatonic sweep to the leading tone in the key of C major, or as a ii $\phi$ diminished 7 chord in a minor, or as an ascending contour, and so forth. Its features include melodic motion by seconds or

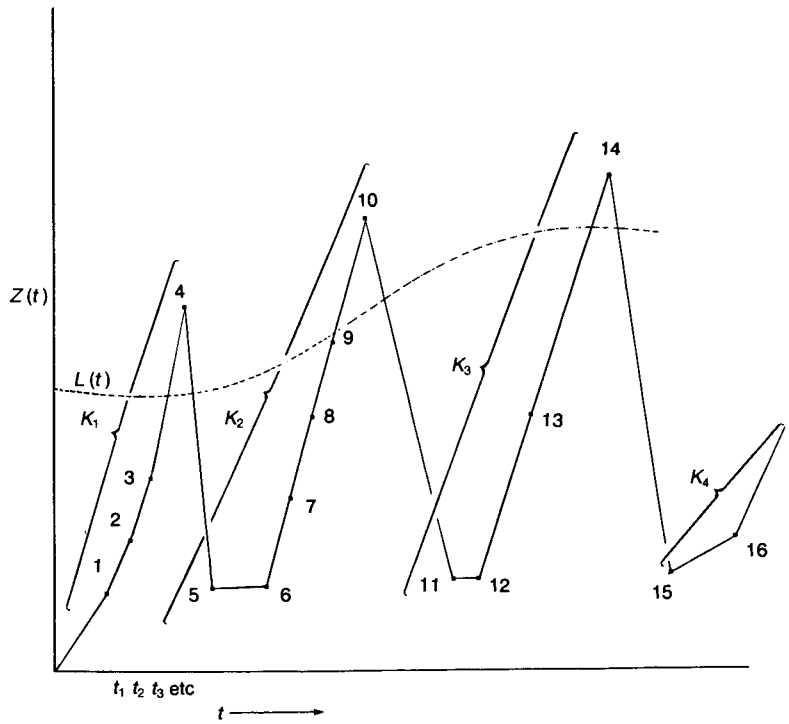
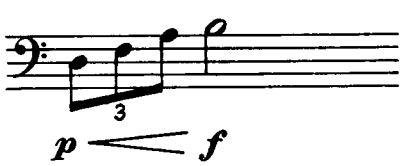



**Fig. 7.3.** Interrupt generation via the repetition functions $L$ and $Z$.

thirds, diatonic note choice, the degree and speed of crescendo, rhythmic regularity of attack, certain values of finger force and velocity used by the performer, and so forth. Many processes could be implicated to generate the given motive: arpeggiate a B$\phi$ diminished 7 chord, pick notes consistent with a triplet feel in C major, move the fingers 4321 of the left hand in such a fashion as to depress keys on the piano, and so forth. If the nature of improvisation entails the seeking out of a satisfactory trajectory in musical action space, such redundancy of description and generation allows maximal flexibility of path selection, so that whatever creative impulse presents itself as an intention, and whatever attentional loadings




**Example 7.1**

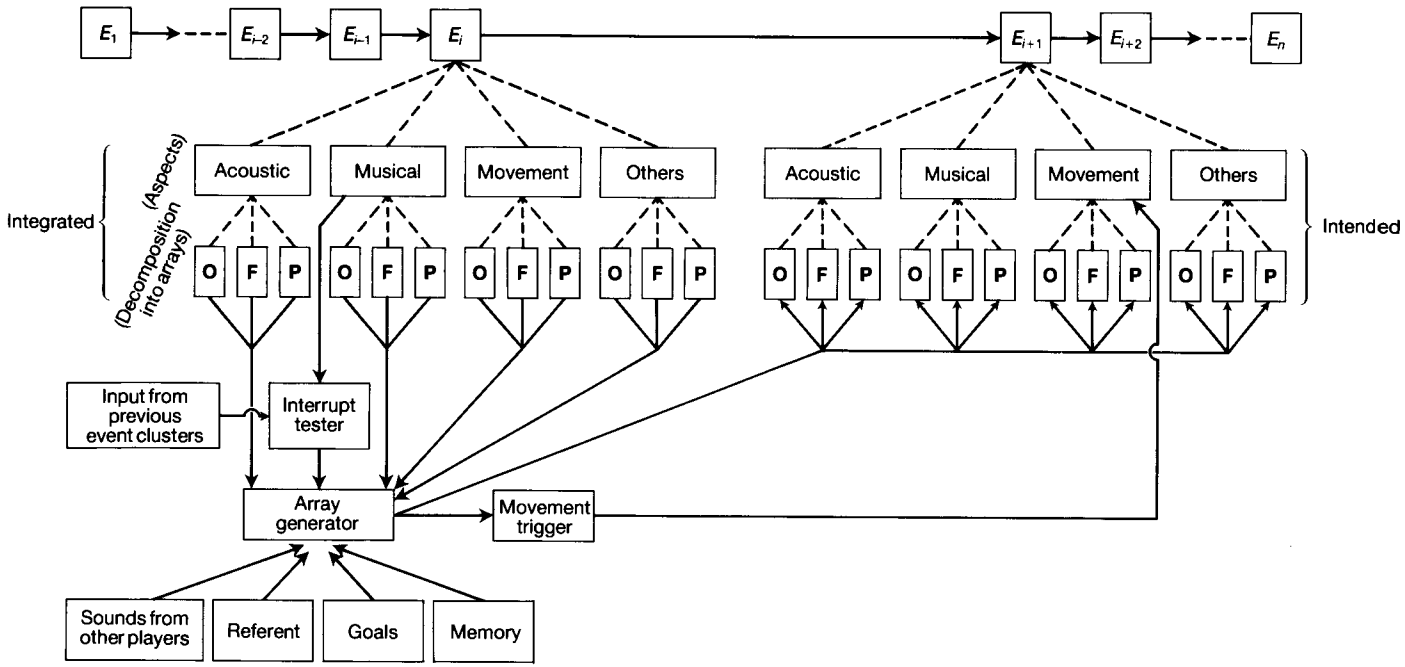



**Fig. 7.4** The improvisation model in diagrammatic form. Only the process $E_i \rightarrow E_{i+1}$ (intended) is detailed. Each event cluster $E_i$ is present in a number of partially redundant aspects, and each of these is decomposed into object, feature, and process arrays. Largely on the basis of musical representation a decision about type of continuation is made by an interrupt tester. In accordance with this decision an intended array decomposition is generated, with input from $E_i$ arrays, referent, goals, and memory. This decomposition acts as a set of constraints in the generation of musical action, and production of $E_{i+1}$ is subsequently begun by a movement trigger at $t_{i+1}$.The diagram detail shows what happens in the time interval $(t_i, t_{i+1})$, so that the indicated decomposition of $E_i$ is integrated (that is, intended plus actual forms of $E_i$ are combined), whereas the indicated decomposition of $E_{i+1}$ is intended (no feedback has been received yet). Hence **O**, **F**, and **P** at time point $t_{i+1}$ do not have indicated outputs.

may be set up, some means of cognitive organization and corresponding motor realization will be available within the limiting constraints of real-time processing.

Such extensive redundancy I take here to mean that control of event production is heterarchical, and may potentially shift rapidly from one cognitive control area to another. Indeed this must be considered the most effective strategy for improvisation. Experientially it very probably corresponds to 'letting go', or 'going with the flow' as described earlier, whereby central hierarchical control, identified here with conscious monitoring of decision making, yields to heterarchical control (and corresponding unconscious allocation of attention).

Next we look further at the object, feature, and process arrays that are critical in the representation and generation of event clusters. First of all it may well be asked how such arrays are formed. The answer given here is based on an ecological perspective, which considers that the capacity to extract or create such arrays is neurologically innate, but that they are only brought into being by interaction with the environment. More specifically, cognitive objects are inferred to exist on the basis of perceived invariance in sensory input over time, and boundedness in a space (whether physical, musical, or abstract). Features are tuneable parameters and come to be abstracted on the basis of perceived similarity or contrast in sensory input. Processes come about from perceived change in an object or along a feature dimension with time.

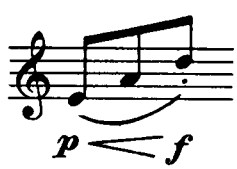
Thus over the course of one's life new arrays and array components are constantly being created by new perceptions and new perceptual groupings. During any given improvisation at most very few new features or processes will be created, and only a limited number of new objects. In general, though, this is one source of novel behaviour: the evolution of movement control structures for newly discovered objects, features, and processes. However, there seems to be another, probably more common source of behavioural novelty: the motor enactment of novel combinations of values of array components. This second possibility is shown for example by considering a child musician who has learned motor actions corresponding to the distinctions loud/soft and fast/slow separately, but without encountering soft and fast simultaneously. By combining these two dimensions an action novel to the child's experience can result. Furthermore, the results of such novel parametric combinations need not be so predictable. If we recall that the human performance system is non-linear, then, as mentioned above in the paragraphs on organizational invariant theory, novel, strikingly different behaviour may follow when controlling system parameters assume certain novel combinations of ranges. It can further be shown mathematically that behaviour described as 'chaotic' may occur under such conditions (Li and Yorke 1975; May 1976), even for simple systems. This perspective has led to a biomathematical analysis, for example, of many so-

called 'dynamical' diseases, including schizophrenia, AV heart block, epilepsy, and some haematological disorders (see Guevara *et al.* 1983 for a survey). The point with regard to improvisation is that the same sort of smooth parametric tuning can be used to generate abrupt intentional novelties in movement and musical expression. The integration of the results of novel ranges of array components is presumed to be handled by control structures of the CNS responsible for timing and smoothness of action.

During any given improvisation, when possible object, feature, and process array types are basically fixed, novel sensory input will be analysed and assigned to existing categories, or, if the fit is too poor, into existing categories plus deviations. In this model such a description is also considered to apply to the generation of action. That is, novel actions are built primarily by distorting aspects of existing ones. This sheds light on the organizing power of the metaphor, mentioned earlier, since it may be considered to be a global link across categories, one that facilitates movement integration. In other words, the image or metaphor enables the co-ordinated modification and resetting of whole classes of array components in a fashion ensuring spatial and temporal coherence.

The central core of the model is the generation of a new set of array components for $E_{i+1}$ from those preceding it. To make this process clearer, we now look at two examples.

(1) Let $E_i$ be




**Example 7.2**

played by the right hand at the piano.

Above are a number of possible improvisational continuations, based on attentional emphasis (that is, cognitive strength) given to the mentioned array components (see Fig. 7.5). Emphasis given to a particular component means that it will guide the generation of subsequent events. The type of arrays emphasized are also indicated; note that this is not uniquely determined, since the model makes a feature of redundancy. Continuations 1–8 exemplify associative continuation, with numbers 7 and 8 more abstract than the others, while number 9 is interrupt based.

| Continuation | Emphasized components used for continuation | Type of arrays |
|---|---|---|
| 1 | key of A major; quaver durations | **O,F,P** |

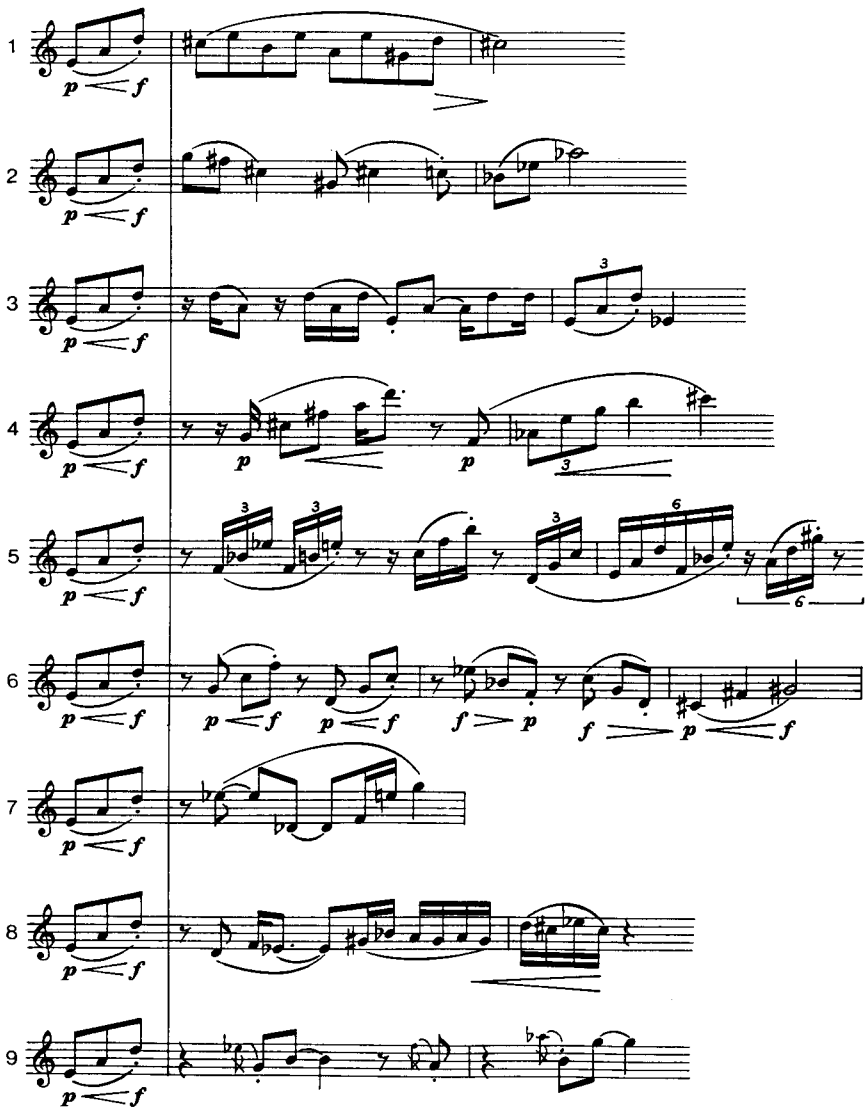


**Fig. 7.5.** Examples of continuation of an event cluster under the emphasis of selected array components.

| | | |
|---|---|---|
| 2 | perfect fourth interval | **F** |
| 3 | notes E, A, D; rhythmic displacement | **O,P** |
| 4 | melodic contour | **O,F** |
| 5 | motor generation with right-hand fingers 1, 2, and 4 | **O,F** |

| | | |
|---|---|---|
| 6 | gesture (note use of contrast), perfect fourth interval | **F,P** |
| 7 | phrase design (antecedent/consequent), interval class 2 | **F** |
| 8 | notes E, A, D: chromatic decoration | **O,P** |
| 9 | interrupt generation: new motive | **O,F,P** |

If the same line had been played on flute, the continuations might all have been very similar except for continuation number 5, which has as its constraint focus the actual movement patterns for manipulating the instrument.

(2) Here we consider an event cluster less clearly tied to structural-historical processes. Let $E_i$ be a segment of sounds produced by a single slow tilting and rotation of a tambourine one-quarter filled with a single layer of small lead shot. $E_i$ is a coloured noise sound which subjectively is reminiscent of distant ocean waves or rain. Some possible continuations are then as follows:

| Continuation | Description | Type of arrays |
|---|---|---|
| 1 | continue tilt but speed up rotation of tambourine | **F,P** |
| 2 | shake tambourine from side to side | **F,P** |
| 3 | stop motion of tambourine | **F,P** |
| 4 | toss lead shot in air and catch it | **F,P** |
| 5 | perform a drum roll on the bottom of the tambourine skin with fingers of the right hand | **O,F,P** |

Continuations 1 and 2 are of associative type, whereas 3, 4, and 5 are interrupt type. Notice that here description emphasizes the motor aspect, since there is no extensive tradition of music theory which applies to such a sound source.

If these examples succeed in illustrating how continuations may be constructed, they are mute on the details of how one continuation comes to be chosen over all other possible ones. What has been said so far is only that, in associative generation, a set of constraints is produced associatively, while in interrupt generation the set of strong constraints on action includes uncorrelated resetting. Obviously, event generation is informed by a vast panorama of culturally and cognitively based musical processes and stylistic preferences (motivic development, phrase design, historical forms, transposition, rhythmic design, etc.). But a considerable degree of residual decision-making remains, as for example the choice of array components that will be singled out to act as strong constraints or to be reset. How are such residual decisions made?

It does not seem possible to give a final answer to this question, for it has

at its ultimate root the question of volition and hence the mind-body problem, about which there is no general philosophical agreement in our culture or even among scientists. There is also no conclusive empirical evidence to support one view or another, despite the opposing claims of some positivists and phenomenologists. It seems useful therefore to characterize a number of strategies of explanation for the residual decision-making mentioned above, and subsequently explore what possibilities exist for experimentally decided among them.

It is first of all possible to take the intuitive perspective, that the individual acts best when he or she merely taps a certain powerful source that dictates the course of musical action in a naturally correct fashion, one that may not be analysable or predictable in physical or musical terms. Although this perspective is usually transpersonal and may seem romantic to some, this does not imply that it is untestable and therefore unscientific.

A second perspective is to assume that this residual decision making actually reflects the effects of individual free will. In other words, the improviser is a unique conscious entity, and residual decision making rests to some degree on internal variables not predictable even in principle from a fully detailed knowledge of the physical-state variables of the improviser and his or her environment.

A third perspective is the physicalist one. Here complex decision making is seen to be an emergent property of the fantastically complex physical system known as a human being, in interaction with a series of environments. Free will in this perspective is either illusory, or simply a somewhat misleading metaphor for certain complex characteristics of the system. There are a number of models possible within this perspective for residual decision making: interactive control with lower CNS centres, network statistical voting models, distributed memory-type models, decision making based on fuzzy logic, etc.

Fourth and last is the perspective of randomness. Here the unconstrained residual decision making is simply modelled by use of random generators. As the improviser becomes more and more expert through practice and more and more control procedures are built up, random processes need to be invoked less and less frequently and overall error levels decrease, perhaps approaching a minimum threshold.

To experimentally distinguish between these points of view a high resolution improvisation transcription system has been built here at La Trobe University. Co-worker Greg Troup and myself, as well as technical staff of the departments of psychology and music, have designed and set up the apparatus. It is a synthesizer-based system, using modified MIDI format, and enables detailed recording, to millisecond resolution, of musical actions at a keyboard. It is also possible to input sound from other (non-keyboard) instruments. Simultaneously as music is recorded a videotape of the performance can be made. The results of this investigation

are not yet complete and will be reported elsewhere, but it seems likely that the limits of validity of intuitionist and random perspectives will be determinable. There seems to be, however, no obvious experimental design that will decide between the physicalist and free-will perspectives. Hence the two may be considered co-existing formulations. The problem in deciding between the two rests with setting up the repeatable conditions which should theoretically lead to the same 'improvised' result in the hard-core physicalist model and to a different improvised result under conditions of free will. But since each event potentially affects all those that follow, all initial conditions are intrinsically unrepeatable.

## The development of improvisational skill

The modelling of this process remains in a less developed state and only a brief discussion is included here. Its starting point is the emergent results of practice found in all types of skill, as mentioned earlier: improved efficiency, fluency, flexibility, capacity for error correction, and, less universally, expressiveness. But there are at least two additional components of improvisational skill: inventiveness and the achievement of coherence. In more fixed skills these are less important, since inventiveness provides few tangible advantages, and coherence is built in by the rigidity of the task demands.*

The specific cognitive changes that allow these properties to develop in improvised musical behaviour are considered to be:

(1) an increase in the memory store of objects, features, and processes—in musical, acoustic, motor (and other) aspects;

(2) an increase in accessibility of this memory store due to the build-up of redundant relationships between its constituents and the aggregation of these constituents into larger cognitive assemblies;

(3) an increasingly refined attunement to subtle and contextually relevant perceptual information.

The build-up and improved access to memory of points (1) and (2) is presumably central to any learning process. In the language of the model of this chapter this involves the use of extensive redundancy, and also the aggregation of memory constituents (objects, features, processes) into new cognitive assembles which may be accessed autonomously. Because such a procedure can presumably be nested to arbitrary depth, very complicated interconnected knowledge structures may develop.

This last idea is not new here. It has a considerable history and has been most clearly outlined for the purposes of this paper by Hayes-Roth (1977), who generalized an earlier model of Hebb (1949). The central feature of

* It is interesting to note that these two skills push in opposite directions, for inventiveness comes from the commitment to avoid repetition as much as possible, while coherence is only achieved by some degree of structural unity, which is only possible with repetition.

aggregation of memory elements Hayes-Roth termed *unitisation*, and her knowledge-assembly theory was built up around the presence of elemental 'cognitive units'. In the terminology of this chapter these are object, feature, and process array components. In knowledge-assembly theory such cognitive units are associatively activated and may combine to form assemblies, whose 'strengths' are increasing functions of recency and frequency of activation, and decreasing functions of their own complexity. From these strengths are derived probabilities and speeds of activation. There is a level of redundancy appropriate to improvisation in this model, since for example a cognitive unit may be activated individually or as part of a larger assembly of cognitive units. In her paper, Hayes-Roth shows that knowledge-assembly learning theory is consistent with a large body of experimental results. It is also consistent with the introspective reports of improvisers and the review given above of improvisation teaching methods. But a decision on the superior applicability of this theory to improvisation over those of other related formulations must rest upon experimental work as yet undone. For this reason I give no further speculation.

The third point mentioned above may be elaborated as follows. The refinement of improvisational skill must depend partly on increasing the efficiency of perceptual processing to allow the inclusion of more and better-selected information in the improviser's decision-making procedures. The need for this efficiency is imposed by every performer's more or less limited individual capacity, per unit time, to process novel sensory input. It seems likely that practice leads to the increasingly efficient use of information in two ways: by reducing the effective amount of information by the recognition of patterns of redundancy in the sensory input, and by focusing attention increasingly on the information that is most relevant for producing a successful improvisation. The increased use of such subtle and 'higher-order' information leads to the higher-order skill characteristics mentioned earlier. The main differences in this process between fixed and improvised actions may be said to reside in the nature of the attention focus used in the two situations. The fixed-skill situation evolves towards a minimal size attention set, whereas the unpredictability of improvisation demands that the attention focus remain wide. To go beyond such insufficiently specific observations experimental work is clearly required.

## Conclusions

This chapter has attempted to illuminate the process of musical improvisation by first examining the modelling tools available from a number of different disciplines. Based on this examination, a cognitive model has then been presented for the process itself, followed by a brief discussion of its relation

to improvisational skill acquisition. The central features of the model are as follows. It is reductionist, in that cognitive structures of processing and control are considered to be broken down into aspects (acoustic, musical, movement, etc.), each of these into types of analytical representation (objects, features, processes), and each of these into characterizing elements (array components). At the same time the model is synergistic and capable of behavioural novelty, due to the extensive redundancy of the cognitive representations and the distributed and non-linear character of the outlined control processes. The extensive presence of feedback and feedforward contributes to this. The fundamental nature of the improvisation process is considered to be the stringing together of a series of 'event clusters' during each of which a continuation is chosen, based upon either the continuing of some existing stream of musical development (called here an event-cluster class) by association of array entries, or the interruption of that stream by the choosing of a new set of array entries that act as constraints in the generation of a new stream (new event-cluster class).

The model seems to be specific enough to allow its use as a basis for the design of 'improvising' computer programs. Work in this direction is in progress. At the same time some fundamental philosophical questions remain about the origin of certain kinds of decision making in any such model, and four types of answers to these have been outlined: intuition, free will, physical causation, and randomness. Some of these alternatives should be distinguishable on the basis of experimental work currently in progress at our laboratories, which also has as its aim the testing of the basic assumptions of the model. This will be described in subsequent publications.

## Acknowledgement

I am indebted to John Sloboda, Margot Prior, Geoff Cumming, Geoff Webb, Denis Glencross, and Glynda Kinsella for helpful criticism.

## References

Abramson, R. M. (1980). Dalcroze-based improvisation. *Music Educator's Journal* **66**, (5), 62–8.

Adams, J. A. (1961). Human tracking behaviour. *Psychological Bulletin* **58**, 55–79.

Adams, J. A. (1971). A closed-loop theory of motor learning. *Journal of Motor Behaviour* **3**, 111–49.

Adams, J. A. (1976). Issues for a closed-loop theory of motor learning. In *Motor control: issues and trends* (ed. G. E. Stelmach). Academic Press, New York.

Adams, R. (1980). *The iron wolf.* Penguin, Reading.

Allport, D. A. (1980). Attention and performance. In *Cognitive psychology: new directions,* (ed. G. Claxton). Routledge and Kegan Paul, London.

Alperson, P. (1984). On musical improvisation. *Journal of Aesthetics and Art Criticism* **43**, 17–29

Armbruster, G. (1984). First steps in improvisation. *Keyboard* **10** (Oct.), 37–44.

Arnold, F. T. (1965). *The art of accompaniment from a thoroughbass as practised in the XVIIth and XVIIIth centuries.* Dover, New York.

Austin, L., Oliveros, P., *et al.* (1982/83). Forum: improvisation. *Perspectives of New Music,* Fall-Winter 1982/Spring-Summer 1983, 26–111.

Avery, T. L. (1984) Structure and strategy in Azorean–Canadian song duels. Unpublished Ph.D. thesis, Indiana University.

Bach, C. P. E. (1778/1949). *Essay on the true art of playing keyboard instruments.* Norton, New York.

Bahm, A. (1960). Types of intuition. *University of New Mexico publications in social sciences and philosophy*, No. 3.

Bailey, D. (1980). *Improvisation: its nature and practice in music.* Moorland, London.

Baker, D. (1969). *Jazz improvisation.* Maher, Chicago.

Bartlett, F. C. (1947). The measurement of human skill. *British Medical Journal* **1**, 835, 877.

Bash, L. (1983). The effectiveness of three instructional methods on the acquisition of jazz improvisation skills. Unpublished Ph.D. thesis. State University of New York at Buffalo.

Bastick, T. (1982). *Intuition: how we think and act.* Wiley, Chichester.

Beatty, J. (1975). *Introduction to physiological psychology.* Brooks/Cole, Monterey, Calif.

Béhague, G. (1980). Improvisation in Latin American musics. *Music Educator's Journal* **66**, (5), 118–25.

Bentley, D. and Konishi, M. (1978). Neural control of behaviour. *Annual Review of Neurosciences* **1**, 35–59.

Bergson, H. L. (1910). *Matter and memory* (authorized trans. N. M. Paul and W. S. Palmer). Allen, London.

Berkowitz, S. (1975). *Improvisation through keyboard harmony.* Prentice-Hall, Englewood Cliffs, NJ.

Bernstein, N. (1967). *The coordination and regulation of movements.* Pergamon, London.

Bresgen, C. (1960). *Die improvisation.* Quelle and Meyer, Heidelberg.

Brooks, D. N. and Baddeley, A. (1976), What can amnesic patients learn? *Neuropsychologia* **14**, 111–22.

Brooks, V. B. (1978). Motor programs revisited. In *Posture and movement: perspectives for integrating sensory and motor research on the mammalian nervous sytem.* Raven, New York.

Bruner, J. S. (1969). Modalities of memory. In *The pathology of memory* (ed. G. A. Tallard, and N. C. Waugh), pp. 253–9. Academic Press, New York.

Brunia, C. H. M. (1980). Motor preparation, recorded on the cortical and spinal level. In *Tutorials in motor behaviour* (ed. G. E. Stelmach and J. Requin). North-Holland, Amsterdam.

Bunge, M. (1962). *Intuition and science.* Prentice-Hall, Englewood Cliffs, NJ.

Burnsed, C. V. (1978). The development and evaluation of an introductory jazz improvisation sequence for intermediate band students. *Dissertation Abstracts International*, **41A**, 1214A.

Carr, I. (1982). *Miles Davis*. Paladin, London.

Chadabe, J. (1984). Interactive composing: an overview. *Computer Music Journal* **8**, (1), 22–7.

Chernikoff, R. and Taylor, F. V. (1952). Reaction time to kinesthetic stimulation resulting from sudden arm displacement. *Journal of Experimental Psychology* **43**, 1–8.

Clark, H. H. and Clark, E. V. (1977). *Psychology and language*. Harcourt Brace Jovanovitch, New York.

Cohen, N. J. (1981). Neuropsychological evidence for a distinction between procedural and declarative knowledge in human memory and amnesia. Unpublished Ph.D. thesis. University of California at San Diego.

Cohen, N. J. and Squire, L. R. (1980). Preserved learning and retention of pattern analysing skill in amnesia: dissociation of knowing how and knowing that. *Science* **210**, 207–9.

Coker, J. (1964). *Improvising jazz*. Prentice-Hall, Englewood Cliffs, NJ.

Coker, J. (1975). *The jazz idiom*. Prentice-Hall, Englewood Cliffs, NJ.

Coker J., Casale, J., Campbell, G., and Greene, J. (1970). *Patterns for jazz*. Studio Productions, Lebanon, Ind.

Coleman, S. N. (1922). *Creative music for children*. G. P. Putnam's Sons, New York.

Cope, D. H. (1984). *New directions in music*. William Brown, Dubuque, Iowa.

Czerny, C. (1829/1983). *Systematic introduction to improvisation on the piano* (Trans. A. L. Mitchell). Longman, New York.

Datta, V. and Lath, M. (1967). Improvisation in Indian music. *World of Music*, **9**, (1), 27–34.

Davis, R. (1957). The human operator as a single channel information system. *Quarterly Journal of Experimental Psychology* **9**, 119–29.

Delcomyn, F. (1980). Neural basis of rhythmic behaviour in animals. *Science* **210**, 492–8.

Denier van der Gon, J. J. and Thuring, J. Ph. (1965). The guiding of human writing movements. *Kybernetik* **2**, 145–8.

Dobbins, B. (1978). *The contemporary jazz pianist*, 4 vols. GAMT Music Press, Jamestown, RI.

Doerschuk, B. (1984). The literature of improvisation. *Keyboard* **10** (Oct.), 48–52.

Dupré, M. (1925/37). *Cours complet d'improvisation a' l'orgue*, 2 vols. A. Leduc, Paris.

Easton, T. A. (1978). Coordinative structure—the basis for a motor program. In *Psychology of motor behaviour and sport* (ed. D. M. Landers and R. W. Christina). Human Kinetics, Champaign, Ill.

Eccles, J. C. (1972). *The understanding of the brain*. McGraw-Hill, New York.

Ehmann, Wilhelm (1950). Chorische Improvisation in der Kantorei. *Kirchenchor* **10** (Sept.–Oct.), 65–71.

Erlmann, V. (1985). Model, variation and performance. Ful'be praise song in Northern Cameroon. *Yearbook for traditional music* **17**, 88–112.

Ewing, A. (1941). Reason and intuition. *Proceedings of the British academy* **27**, 67–107.

Ferand, Ernst (1938). *Die Improvisation in der Musik.* Rhein-Verlag, Zurich.

Ferand, Ernst (1961). *Improvisation in nine centuries of western music.* Arno Volk Verlag, Hans Gerig KG, Cologne.

Fitts, P. M. (1964). Perceptual-motor skill learning. In *Categories of human learning* (ed. A. W. Melton). Academic Press, New York.

Fry, C. (1980). Computer improvisation. *Computer Music Journal* **4** (3), 48 ff.

Fry, C. (1982/83). Dancing musicians. *Perspectivies of New Music*, Fall–Winter 1982/Spring–Summer 1983, 585–9.

Gaburo, K. (1968). *Twenty sensing compositions.* Lingua Press, La Jolla, Calif.

Galás, D. (1981/82). Intravenal song. *Perspectives of New Music*, Fall–Winter 1981/Spring–Summer 1982, 59–62.

Gehring, P. K. (1963). Improvisation in contemporary organ playing. Unpublished Ph.D. thesis Syracuse University.

Gelernter, H. (1963). Realisation of a geometry-theorem-proving machine. In *Computers and thought.* (ed. E. A. Feigenbaum and J. Feldman). McGraw-Hill, New York.

Gel'fand, I. M. and Tsetlin, M.L. (1962). Some methods of control for complex systems. *Russian Mathematical Surveys* **17**, 95–116.

Gel'fand, I. M. and Tsetlin, M. L. (1971). Mathematical modelling of mechanisms of the central nervous system. In *Models of the structural–functional organisation of certain biological systems* (eds. I. M. Gel'fand, V. S. Gurfinkel, S. V. Fomin, and M. T. Tsetlin). MIT Press, Cambridge, Mass.

Gibbs, C. B. (1965). Probability learning in step-input tracking. *British Journal of Psychology* **56**, 233–42.

Gibbs, C. B. and Brown, I. C. (1956). Increased production from information incentives in an uninteresting repetitive task. *Manager* **24**, 374–9.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Houghton Mifflin, Boston.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Houghton Mifflin, Boston.

Glencross, D. J. (1977). Control of skilled movements. *Psychological Bulletin* **84**, 14–29.

Glencross, D. J. and Koreman, M. M. (1979). The processing of proprioceptive signals. *Neuropsychologia* **17**, 683–7.

Goldstein, M. (1983). The gesture of improvisation: some thoughts, reflections and questions regarding percussion music. *Percussionist* **21**, (3), 18–24.

Gorder, W. D. (1976) An investigation of divergent production abilities as constructs of musical creativity. Unpublished Ed. D. thesis. University of Illinois at Urbana–Champaign.

Greene, P. M. (1972). Problems of organisation of motor systems. In *Progress in theoretical biology*, Vol. 2 (eds. R. Rosen and F. M. Snell). Academic Press, New York.

Guevara, M., Glass, L., Mackey, M., and Shrier, A. (1983). Chaos in neurobiology. *IEEE Transactions on systems, Man, and Cybernetics* **SMC-13**, (5), 790–8.

Guildford, J. P. (1977). *Way beyond the I.Q.* Creative Education Foundation, New York.
Guildford, J. P. and Hoepfner, R. (1971). *The analysis of intelligence.* McGraw-Hill, New York.
Hamel, P. M. (1979). *Through music to the self.* Shambhala, Boulder, Col.
Hanlon, E. S. (1975). Improvisation: theory and application for theatrical music and silent film. Unpublished Ph.D. thesis. University of Cincinnati.
Hayes-Roth, B. (1977). Evolution of cognitive structures and processes. *Psychological Review* **84**, (3), 260–78.
Hebb, D. D. (1949). *Organisation of behaviour.* Wiley, New York.
Hendrickson, G. and Schroeder, W. H. (1941). Transfer of training in learning to hit a submerged target. *Journal of Educational Psychology* **32**, 205–13.
Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology* **4**, 11–26.
Higgins, J. R. and Angel, R. W. (1970). Correction of tracking errors without sensory feedback. *Journal of Experimental Psychology* **84**, 412–16.
Hodeir, Andre (1956). *Jazz: its evolution and essence* (trans. David Noakes). Grove, New York.
Holding, D. H. (1965). *Principles of training.* Pergamon, London.
Holding, D. (1981). Skills research. In *Human skills* (ed. D. Holding). Wiley, Chichester.
Hollerbach, J. M. (1981). An oscillation theory of handwriting. *Biological Cybernetics* **39**, 139–56.
Hood, M. (1971). Aspects of group improvisation in the Javanese gamelan. In *Musics of Asia* (ed. José Maceda), pp. 17–21. Manila.
Hood, M. (1975). Improvisation in the stratified ensembles of Southeast Asia. *Selected Reports in Ethnomusicology* (UCLA) **2**, (2), 25–33.
Hores, R. G. (1977). A comparative study of visual- and aural-orientated approaches to jazz improvisation with implications for instruction. Unpublished Ed. D. thesis. Indiana University.
Howard, J. (1978). Improvisational techniques of Art Tatum. Unpublished Ph.D. thesis. Case Western Reserve University.
Huggins, A. W. F. (1978). Speech timing and intelligibility. In *Attention and performance VII* (ed. J. Requin). Lawrence Erlbaum, Hillsdale, NJ.
Jairazbhoy, N. A. (1971). *The rags of North Indian music.* Faber and Faber, London.
Jaques-Dalcroze, E. (1921/1976). *Rhythm, music and education.* B. Blom, New York.
Jaques-Dalcroze, E. (1930). *Eurythmics, art, and education.* Ayer, New York.
James, W. (1890). *Principles of psychology*, Vol. 1. Holt, New York.
Jones, A. M. (1959). *Studies in African Music.* Oxford, London.
Jost, E. (1974). *Free jazz.* Universal, Graz.
Keele, S. W. and Posner, M. I. (1968). Processing of feedback in rapid movements. *Journal of Experimental Psychology* **77**, 353–63.
Keele, S. W. and Summers, J. J. (1976). The structure of motor programs. In *Motor control: issues and trends* (ed. G. E. Stelmach). Academic Press, New York.
Kelso, J. A. S. (1982). Two strategies for investigating action. In *Human motor*

*behaviour* (ed. J. A. S. Kelso). Lawrence Erlbaum, Hillsdale, NJ.

Kelso, J. A. S. and Wallace, S. A. (1978). Conscious mechanisms in movement. In *Information processing in motor control and learning* (ed. G. E. Stelmach). Academic Press, New York.

Kelso, J. A., Holt, J. G., Rubin, P., and Kugler, P. N. (1981). Patterns of human interlimb coordination emerge from the properties of non-linear, limit cycle oscillatory processes: theory and data. *Journal of Motor Behaviour* **13**, 226–61.

Kernfeld, B. (1983). Two Coltranes. *Annual Review of Jazz Studies* **2**, 7–66.

Kerr, R. (1982). *Psychomotor learning.* Saunders, Philadelphia.

Kickert, W. G., Bertrand, J. W., and Praagman, J. (1978). Some comments on cybernetics and control. *IEEE transactions on systems, man and cybernetics* **SMC-8**, 805–9.

Kleeman, J. E. (1985/86). The parameters of musical transmission. *The Journal of Musicology* **4**, 1–22.

Kleinman, D. L., Baron, S., and Levison, W. H. (1971). A control theoretic approach to manned-vehicle systems analysis. *IEEE Transactions on Automatic Control* **AC-16**, 824–32.

Kugler, P. N., Kelso, J. A. S., and Turvey, M. T. (1980). On the concept of coordinative structures as dissipative structures: I. theoretical lines of convergence. In *Tutorials in motor behaviour* (ed. G. E. Stelmach and J. Requin). North-Holland, Amsterdam.

Laabs, G. J. and Simmons, R. W. (1981). Motor memory. In *Human skills* (ed. D. Holding). Wiley, New York.

Laneri, Roberto (1975). Prima materia: an opus in progress. The natural dimension of music. Unpublished Ph.D. thesis. University of California at San Diego.

Lenat, D. B. (1984). Computer software for intelligent systems. *Scientific American* **251**, (3), 152–60.

Leonard, J. A. (1959). Tactual choice reactions. *Quarterly Journal of Experimental Psychology* **11**, 76–83.

Levitt, D. A. (1981). A melody description system for jazz improvisation. Unpublished M.S. thesis. MIT.

Li, T. Y. and Yorke, J. (1975). Period three implies chaos. *American Mathematical Monthly* **82**, 985–92.

Liebman, D., Beirach, R., Tusa, F., Williams, J.,. and Roy, B. (1978). *Lookout farm.* Almo Publications, Hollywood.

Lipiczky, T. (1985). Tihai formulas and the 'composition' and improvisation in North Indian music. *The Musical Quarterly* **71**, 157–71.

Locke, D. (1979). The music of atsiabeko. Unpublished Ph.D. thesis. Wesleyan University.

Lord, A. B. (1964). *The singer of tales.* Harvard University Press, Cambridge, Mass.

Lord, A. B. (1965). Yugoslav epic folk poetry. In *The study in folklore* (ed. A. Dundes). Prentice-Hall, Englewood Cliffs, NJ.

McCulloch, W. S. (1945). A heterarchy of values determined by the topology of nervous nets. *Bulletin of Mathematical Biophysics* **7**, 89–93.

Marteniuk, R. G. and Romanow, S. K. E. (1983). Human movement organisation and learning as revealed by variability of movement, use of kinematic

information, and Fourier analysis. In *Memory and control of action* (ed. R. A. Magill). North-Holland, Amsterdam.
Mathieu, L. (1984). A phenomenological investigation of improvisation in music and dance. Unpublished Ph.D. thesis, New York University.
May, R. M. (1976). Simple mathematical models with very complicated dynamics. *Nature* **261**, 459–67.
Mersenne, M. (1635). *Harmonie Universelle.*
Milano, D. (1984). The psychology of improvisation. *Keyboard* **10** (Oct.), 25, 30–5.
Miller, G. A., Galanter, E., and Pribram, K. (1960). *Plans and the structure of behaviour.* Holt, Rinehart and Winston, New York.
Milner, B. (1962). Les troubles de la memoire accompagnant des lesions hippocampiques bilaterales. In *Phsyiologie de l'Hippocampe.* Cent. Natl. Recherche Scientifique, Paris.
Morasso, P. (1983). Three dimensional arm trajectories. *Biological Cybernetics* **48**, 187–94.
Murphy, F. (1982). The cornet style of Leon Bix Beiderbecke (1903–1931). Unpublished Ph.D. thesis. La Trobe University, Melbourne.
Nagler, M. N. (1974). *Spontaneity and tradition: a study in the oral art of Homer.* University of California Press, Berkeley.
Namikas, G. (1983). Vertical process and motor performance. In *Memory and control of action* (ed. R. A. Magill). North-Holland, Amsterdam.
Nelson, O. (1966). *Patterns for improvisation.* Nelson Music, Hollywood.
Nettl, B. (1974). Thoughts on improvisation . *Musical Quarterly* **60**, 1–19.
Nettl, B. and Foltin, B. Jr. (1972). *Daramad of Chahargah.* Information Coordinators, Detroit.
Nettl, B. and Riddle, R. (1974). Taqsim Nahawand: a study of 16 performances by Jihad Racy. *Yearbook of the International Folk Music Council* **5**, 11–50.
Newell, K. M. (1981). Skill learning. In *Human skills* (ed. D. Holding). Wiley, Chichester.
Oliveros, P. (1971). *Sonic meditations.* Smith Publications/Sonic Art Editions, Baltimore, Md.
Owens, T. (1974). Charlie Parker: techniques of improvisation, 2 Vols. Unpublished Ph.D. thesis. University of California at Los Angeles.
Paillard, J. (1980). The multichanneling of visual cues and the organisation of visually guided response. In *Tutorials in motor behaviour* (eds. G. E. Stelmach and J. Requin). North-Holland, Amsterdam.
Park, M. (1985). Music and shamanism in Korea: a study of selected ssikkŭm-gut rituals for the dead. Unpublished Ph.D. thesis, University of California at Los Angeles.
Parry, M. (1930). Studies in the epic technique of oral versemaking: I. Homer and Homeric style. *Harvard Studies in Classical Philology* **41**, 73–147.
Parry, M. (1932). Studies in the epic technique of oral versemaking: II. the Homeric language as the language of poetry. *Harvard Studies in Classical Philology* **43**, 1–50.
Parsons, W. (1978). *Music for citizen's band.* W. Parsons, La Jolla, Calif.
Partchey, K. C. (1973). The effects of feedback, models, and repetition on the

ability to improvise melodies. Unpublished D.Ed. thesis. Pennsylvania State University.

Pew, R. W. (1974). Human perceptual-motor performance. In *Human information processing: tutorials in performance and cognition* (ed. B. H. Kantowitz). Erlbaum, NJ.

Pew, R. W. and Baron, S. (1978). The components of an information processing theory of skilled performance based on an optimal control perspective. In *Information processing motor control and learning* (ed. G. E. Stelmach). Academic Press, New York.

Pike, A. (1974). A Phenomenology of jazz. *Journal of Jazz Studies* **2**, (1), 88–94.

Poulton, E. C. (1957). On the stimulus and response in pursuit tracking. *Journal of Experimental Psychology* **53**, 57–65.

Powers, W. T. (1973). *Behaviour: the control of perception.* Aldine, Chicago.

Pressing, J. (1980). Music, altered states of consciousness, and psi. In *Proceedings of the Psychic Orientation Conference* (ed. A. Gabay). La Trobe University, Melbourne.

Pressing, J. (1984*a*). Cognitive processes in improvisation. In *Cognitive processes in the perception of art* (ed. W. R. Crozier and A. J. Chapman), pp. 345–63. North-Holland, Amsterdam.

Pressing, J. (1984*b*). A history of musical improvisation to 1600. *Keyboard* **10** (11), 64–8.

Pressing J. (1984*c*). A history of musical improvisation: 1600–1900. *Keyboard* **10**, (12), 59–67.

Prigogine, I. (1967). *Introduction to thermodynamics of irreversible processes.* Interscience, New York.

Prigogine, I. and Nicholis, G. (1971). Biological order, structure and instabilities. *Quarterly Review of Biophysics* **4**, 107–48.

Propp, V. I. (1927/1968). *Morphology of the folk tale* (trans. L. Scott). University of Texas Press, Austin.

Quantz, J. J. (1752/1966). *On playing the flute* (trans. E. Reilley). Free Press, New York.

Rabbitt, P. M. A. and Vyas, S. M. (1970). An elementary preliminary taxonomy for some errors in laboratory choice RT tasks. *Acta Psychologica* **33**, 56–76.

Radano, R. M. (1985). Anthony Braxton and his two musical traditions, the meeting of concert music and jazz (2 vols). Unpublished Ph.D. thesis, University of Michigan.

Reck, D. B. (1983). A musician's tool-kit: a study of five performances by Thirugokarnam Ramachandra Iyer, 2 Vols. Unpublished Ph.D. thesis. Wesleyan University.

Rich, E. (1983). *Artificial intelligence.* McGraw-Hill, New York.

Richardson, C. P. (1983). Creativity research in music education: a review. *Council for Research in Music Education* **74**, (Spring 1983), 1–21.

Roads, C. (1979). Grammars as representations for music. *Computer Music Journal* **3**, 48–55.

Ryle, G. (1949). *The concept of mind.* Hutchinson, London.

Sadie, S. (ed.) (1980). *The new Grove dictionary of music and musicians.* Macmillan, London. Listings under improvisation, aleatory, cadenza, continuo, division. prélude non mesuré.

Sage, G. S. (1977). *Introduction to motor behaviour: a neuropsychological approach.* Addison-Wesley, Reading, Mass.

Saltzman, E. L. and Kelso, J. A. S. (1983). Toward a dynamical account of motor memory and control. In *Memory and control of action* (ed. R. A. Magill). North-Holland, Amsterdam.

Schickhaus, K-H. (1978). *Neues Schulwerk für Hackbrett.* Josef Preissler, Munich.

Schmidt, R. A. (1983). On the underlying response structure of well-learned motor responses: a discussion of Namikas and Schneider and Fisk. In *Memory and control of action* (ed. R. A. Magill). North-Holland, Amsterdam.

Schneider, W. and Fisk, A. D. (1983). Attention theory and mechanisms for skilled performance. In *Memory and control of action* (ed. R. A. Magill). North-Holland, Amsterdam.

Schouten, H. (no date given). *Improvisation on the organ* (trans. by J. L. Warren). W. Paxton, London.

Schuller, G. (1968). *Early jazz: its roots and musical development.* Oxford University Press, New York.

Serafine, M. L. (1983). Cognition in music. *Cognition* **14**, 119–83.

Shafer, R. M. (1969). *Ear cleaning.* Universal, London.

Shaffer, L. H. (1978). Timing in the motor programming of typing. *Quarterly Journal of Experimental Psychology* **30**, 333–45.

Shaffer, L. H. (1980). Analysing piano performance: a study of concert pianists. In *Tutorials in motor behaviour* (ed. G. E. Stelmach and J. Requin). North-Holland, Amsterdam.

Shaffer, L. H. (1981). Performances of Chopin, Bach and Bartok: studies in motor programming. *Cognitive psychology* **13**, 327–76.

Shaffer, L. H. (1984). Timing in solo and duet piano performances. *Quarterly Journal of Experimental Psychology* **A36**, 577–95.

Shepherd, G. M. (1983). *Neurobiology.* Oxford University Press, New York.

Sher, C. (1983). *The world's greatest fake book.* Sher Music, San Francisco.

Shiffrin, R. M. and Schneider, W. (1977). Controlled and automatic human information processing: II. perceptual learning, automatic attending, and a general theory. *Psychological Review* **84**, 127–90.

Signell, K. (1974). Esthetics of improvisation in Turkish art music. *Asian Music* **5**, (2), 45–9.

Signell, K. (1977). *Makam: modal practice in Turkish art music.* Asian music Publications, Seattle.

Silverman, M. L. (1962) Ensemble improvisation as a creative technique in the secondary instrumental program. Unpublished Ed.D. thesis. Stanford University.

Sloboda, J. A. (1982). Music performance. In *Psychology of music* (ed. D. Deutsch). Academic Press, New York.

Slonimsky, N. (1975). *Thesaurus of scales and melodic patterns.* Duckworth, London.

Smith, G. E. (1983). Homer, Gregory, and Bill Evans? the theory of formulaic composition in the context of jazz piano improvisation. Unpublished Ph.D. thesis, Harvard University.

Smith, W. M. and Bowen, K. F. (1980). The effects of delayed and displaced visual feedback on motor control. *Journal of Motor Behaviour* **12**, 91–101.

Sparrow, W. A. (1983). The efficiency of skilled performance. *Journal of Motor Behaviour* **15**, 237–61.

Sperber, M. (1974). Improvisation in the performing arts: music, dance and theatre. Unpublished Ed.D. thesis. Columbia University.

Squire, L. R. (1982). The neuropsychology of human memory. *Annual Review of Neuroscience* **5**, 241–73.

Stocks, J. (1939). *Reason and intuition* (ed. D. M. Emmett). Oxford University Press, New York.

Stumme, W. (1972). *Über improvisation.* B. Schott, Mainz.

Sudnow, D. (1978). *Ways of the hand.* Harvard University Press, Cambridge, Mass.

Suehs, H. C. (1979). The development, implementation, and assessment of a course of study for instruction in certain improvisational techniques in the performance of Baroque music from 1679 to 1741. Unpublished Ph.D. thesis, Catholic University of America at Washington.

Sumarsam (1981). The musical practice of the gamelan sekaten. *Asian Music* **12**, (2), 54–73.

Summers, J. J. (1981). Motor programs. In *Human skills* (ed. D. Holding), Wiley, Chichester.

Suzuki, S., Mills, E. and Murphy, T. (1973). *The Suzuki concept: an introduction to a successful method for early music education.* Diablo Press, Berkeley, Calif.

Terzudo, C. and Viviani, P. (1979). About the central representation of learned motor patterns. In *Posture and movement* (eds. R. Talbot and D. R. Humphrey). Raven, New York.

Thompson, S. (1946). *The folktale.* Holt, Rinehart and Winston, New York.

Torrance, E. P. (1966) *Guiding creative talent.* Prentice-Hall, Englewood Cliffs, NJ.

Touma, H. H. (1971). The *maqam* phenomenon: an improvisation technique in the music of the middle east. *Ethnomusicology* **15**, 38–48.

Treitler, L. (1974). Homer and Gregory: the transmission of epic poetry and plainchant. *Musical Quarterly* **9**, 333–72.

Turvey, M. T. (1977). Preliminaries to a theory of action with reference to vision. In *Perceiving, acting and knowing: toward an ecological psychology* (eds. R. Shaw and J. Bransford). Lawrence Erlbaum, Hillsdale, NJ.

Ueltzen, D. R. (1986). Improvisation in Kirchengesang und mit dem Kirchenlied und am Gesangbuch für das Jahr 2000. *Gottesdienst und Kirchenmusik* **3**, 71–7.

Vaughan, M. M. (1971). Music as model and metaphor in the cultivation and measurement of creative behaviour in children. Unpublished Ed.D. thesis. University of Georgia.

Vīķis-Fribergs, V. (1984). Creativity and tradition in oral folklore or the balance of innovation and repetition in the oral poet's art. In *Cognitive processes in the perception of art* (eds. W. R. Crozier and A. J. Chapman). North-Holland, Amsterdam.

Viviani, P. and Terzudo, C. (1980). Space-time invariance in learned motor skills. In *Tutorials in motor behaviour* (eds. G. E. Stelmach and J. Requin). North-Holland, Amsterdam.

von Foerster, H. (1960). On self-organising systems and their environments. In

*Self-organising systems* (ed. M. C. Yorits and S. Cameron). Pergamon, New York.

Von Holst, E. (1954). Relations between the central nervous system and the peripheral organs. *British Journal of Animal Behaviour* **2**, 89–94.

Wade, Bonnie (1973). Chiz in Khyal: the traditional composition in the improvised performance. *Ethnomusicology* **17**, 443–59.

Webster, R. P. (1977). A factor of intellect approach to creative thinking in music. Unpublished Ph.D. thesis. University of Rochester, Eastman School of Music.

Weidner, R. F. (1984). The improvisation techniques of Charles Tournemire as extracted from his five reconstructed organ improvisations. Unpublished Ph.D. thesis, Michigan State University.

Welford, A. T. (1974). On the sequencing of action. *Brain Research* **71**, 381–92.

Welford, A. T. (1976). *Skilled performance.* Scott, Foresman, Glenview, Ill.

West, L. J. (1967). Vision and kinesthesis in the acquisition of typewriting skill. *Journal of Applied Psychology* **51** (2), 161–6.

Westcott, M. R. (1968). *Towards a contemporary psychology of intuition*, Holt, Reinhart and Winston, New York.

Whitmer, T. (1934). *The art of improvisation.* M. Witmark and Sons, New York.

Wilberg, R. B. (1983). Memory for movement: discussion of Adams and Saltzman and Kelso. In *Memory and control of action* (ed. R. A. Magill). North-Holland, Amsterdam.

Winograd, T. (1975). Frame representations and the declarative-procedural controversy. In *Representation and understanding* (ed. D. Bobrow and A. Collins). Academic Press, New York.

Zelaznik, H. N., Hawkins, B., and Kisselburgh, L. (1983). Rapid visual feedback processing in single-arming movements. *Journal of Motor Behaviour* **15**, 217–36.

Zonis, Ella (1973). *Classical Persian music: an introduction.* Harvard University Press, Cambridge, Mass.

# 8

# Experimental research into musical generative ability*

MÁRIA SÁGI and IVÁN VITÁNYI

One of the projects carried on for about 15 years in the Institute for Culture of Hungary has been concerned with analysing human creative faculties; in particular, artistic creativity. We have been concerned to discover the place, role, and modes of creativity in Hungarian culture, and we have attempted to relate different levels of creativity to economic, social, and educational factors. Our empirical studies have been extended to differnt arts, including music, fine arts, fiction, and poetry.

This chapter describes part of our research into musical creativity. Music is a particularly suitable idiom for our purposes. Its abstract character makes it easy to react to on many different levels. Also, music has deeply penetrated many different social groups, thus 'pre-conditioning' a wide range of subjects for giving relevant responses. In addition, the wide use in Hungarian education of the Kodály method has placed great importance on the practical development of musical creativity.

The music-psychological research to be described below was provoked by the view widespread in the literature both of the psychology and the sociology of art that two kinds of artistic experiences can be distinguished: that of the creator and that of the receiver. This is a view we do not share. Proponents of this view tend to argue that only a narrow group in contemporary society takes a creative part in music (for example from Hungary's population of 10 million, taking into account all the composers of popular songs and rock music, only about 2 to 3000; that is 0.02 to 0.03 per cent).

A contrasting view is provided by Weber's (1921) and Blaukopf's (1972) ideas on the role of sound systems. According to them, the systems of musical structures adopted by a composer form a language, a semiotic system to which the most ingenious composer can merely add something and which he is unable to create himself.

* A short summary of part of this study was published earlier in *Contribution to the Sociology of the Arts* (1983), ISA Research Committee 37, pp. 242–9. Research Institute for Culture, Sofia.

Of course, there are different levels of composing ability. In particular, we can distinguish between *constructive* and *generative* forms of composing ability. We do not wish to dwell at length on the general theoretical issues of creativity because our aim is to describe our empirical investigation. Let us simply note that we agree with authors like Taylor (1959) who distinguish levels of creativity. We consider creativity as consisting of two main types: constructive creativity (which is—in the stricter sense of the term—artistic creativity) and generative creativity.

We speak about constructive creative ability in music where the composer gives a final form to an original opus by means of conscious work, employing and (partly) reshaping the elements and rules known to him. By generative composing we mean a largely unconscious or intuitive variational application of the elements and rules which does not result in a final opus of unchangeable form but merely in a new variant. We have adapted this term from Chomsky's (1965) 'generative linguistic ability'. Generative linguistic ability means that with the knowledge of a definite number of elements and of a definite number of generative rules all the speakers of a language can produce an indefinite number of sentences. That is exactly how generative musical ability acts: with a definite number of elements (stock of sounds) 'and a definite number of generative rules' an indefinite number of melodies can be constructed.

Historical evidence of generative musical ability is readily available. We see this in the way that folk-songs take shape, are transformed, and are sung time and again by folk-singers. We may legitimately ask, however, if, through the destruction of traditional popular art, the generative musical ability of mankind may not have been destroyed as well. Has the pervasive music-supplying apparatus of radio, television, cinema, music-halls, and restaurants killed the ability of man to compose music for himself? This was the question to which the following investigation sought a reply.

## Procedure

### Exploration

Prior to undertaking the musical tasks subjects were engaged in a personal discussion to put them at ease. The subjects gladly discussed their lives, past and present, their general views on life, their relations to the world and the arts, and their general and artistic tastes.

### Examination of generative musical ability

Subjects were asked to make immediate sung improvisations to poems.

One was Sándor Petőfi's (mid-19th century) '*Falu végén kurta kocsma*' ('Little inn outside the village') in the style of a folk-song.

| | |
|---|---|
| *Falu végén kurta kocsma,* | At village end a small saloon, |
| *Oda rug ki a Szamosra,* | That is where Szamosra hangs out, |
| *Meg is látná magát benne,* | He would even be able to see himself in it, |
| *Ha az éj nem közeledne.* | If night was not approaching. |
| *Az éjszaka közeledik,* | Night is coming, |
| *A világ lecsendesedik,* | The world grows quiet, |
| *Pihen a komp, kikötötték,* | The ferryboat rests, they have tied it out, |
| *Benne hallgat a sötétség.* | In it the darkness remains silent. |
| | (*translation by Dr Thomas Kabdebo*) |

Two other poems were chosen to represent the Hungarian poetry of the twentieth century: Endre Ady's poem '*A Kirsztusok mártirja*' ('The Martyr of the Christ').

| | **Literal translation** |
|---|---|
| *Vad, nagyszerü rajongást oltott* | The brooklet's bank has infused in me |
| *Az Érnek partja én belém* | A worship wild and passionate, |
| *Csupa pogányság volt a lelkem,* | Of paganism my soul was brimful |
| *Gondtalan vágy és vak remény* | Carefree desire and blind of hope. |
| *Forgott körültem zagyva módon* | Turning confusedly around me |
| *Lármával, vadul a világ* | The world was clamorous and wild. |
| *És én kerestem egyre-egyre* | I was searching unfalteringly |
| *Valami nagy Harmóniát.* | In it some major Harmony. |

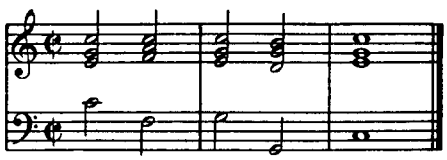
and eight lines of Attila József's '*Óda*' ('Ode').

| | |
|---|---|
| *Óh, hát miféle anyag vagyok én,* | O what kind of matter am I |
| *hogy pillantásod metsz és alakit?* | that your glance cuts and shapes me? |
| *Miféle lélek és miféle fény* | What kind of soul and what kind of light |
| *s ámulatra méltó tünemény,* | and what kind of amazing phenomenon am I |
| *hogy bejárhatom a semmiség ködén* | that in the mist of emptiness I can walk around |
| *termékeny tested lankás tájait?* | the gentle slopes of your fertile body? |
| *S mint megnyilt értelembe az ige,* | And like the word entering into an enlightened mind |
| *alászállhatok rejtelmeidbe.* | I can enter into its mysteries . . . |
| | (*translation by Dr Thomas Kabdebo*) |

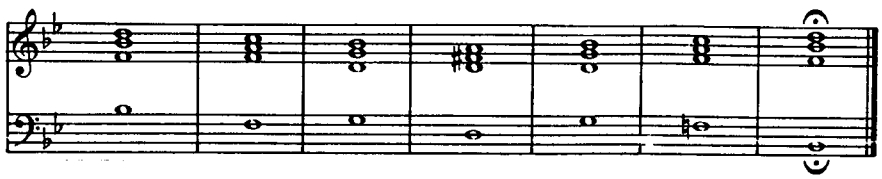
The fourth poem was the lyric of a hit song.

Subjects were also asked to make improvisations to simple harmonic progressions; first to the familiar I–IV–V–I, then to the more sophisticated progressions, including some typical of Bartók (see Examples 8. 1–7).

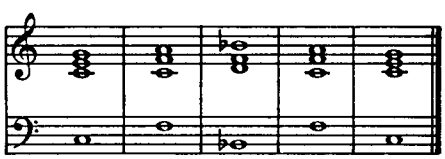
The sense of tonality of the subjects was examined by asking them to improvise completions to three melodies (see Examples 8. 8–10). Each was



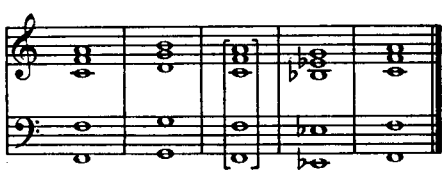
**Example 8.1**



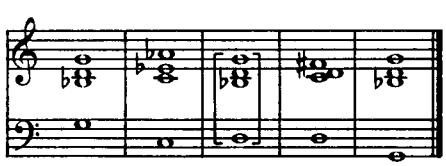
**Example 8.2**



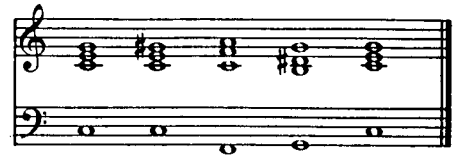
**Example 8.3**



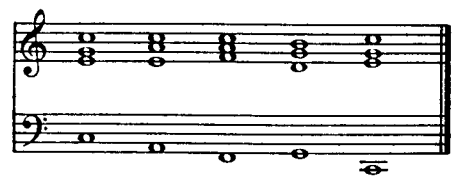
**Example 8.4**



**Example 8.5**

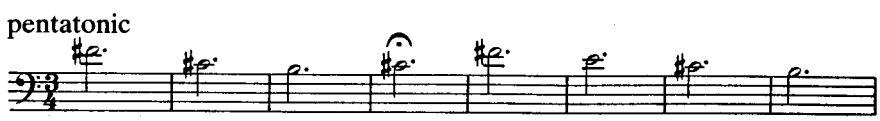


**Example 8.6**

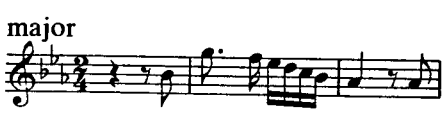


**Example 8.7**

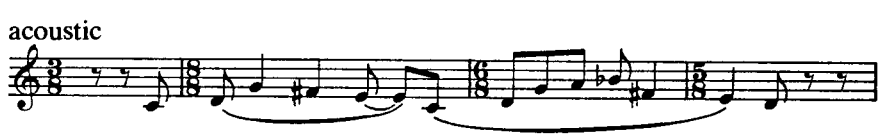
in a different mode: pentatonic, major, and acoustic (or overtone). The latter, often used by Bartók, produces scales such as C–D–E–F♯–G–A–B flat–C.




**Example 8.8.** Bartók.




**Example 8.9.** Mozart.




**Example 8.10.** Bartók.

### Tests of tonal awareness

Some additional receptive tests of tonal awareness were administered, but the results of these tests are not considered here.

## The subjects

The experiment was conceived on a large scale, with 220 subjects in 11 samples, after a pilot study with 20 peasants from a Hungarian village, Aradva'nypuszta.

Our hypothesis is that any person, whether musician or not is capable of composing music such as a song verse, using the musical patterns and structures provided by his/her daily musical environment (radio, TV, singing, etc.). Some musicologists have challenged our view. This is why our sample included untrained subjects from different age groups and social strata. A few control groups were selected from those studying or having studied music.

The impact of European musical culture arose as a separate issue which was studied by examining control groups consisting of students in Budapest of non-European (Asian, African, Latin American) origin.

The 11 samples are shown in Table 8.1.

**Table 8.1** *Subject samples used in the experiment.*

| Group No. | Age | Laymen | Age | Those having studied music |
|---|---|---|---|---|
| 1/2 | 13 | pupils | 13 | music pupils |
| 3 | 17 | middle school pupils | – | – |
| 4 | 17 | apprentices | – | – |
| 5 | 20–25 | industrial workers | – | – |
| 6/7 | 20–25 | university students | 20–25 | conservatory students |
| 8 | about 40 | industrial workers | – | – |
| 9/10 | about 60 | industrial workers | about 60 | industrial workers singing in choirs |
| 11 | 20–25 | | | non-European students |

Although the experiment involved a great deal of work (each person was tested for two to three hours), we succeeded in examining 20 persons in each of the 11 samples.

Before the beginning of the experiments a dispute over principles arose: many musicologists said we were embarking upon an impossible enterprise, that it was inconceivable that laymen might be able to perform the task. However, the result of the investigation has corroborated the correctness of our assumption.

Three thousand four hundred and twenty tunes, the length of folk-songs were recorded. If all subjects had succeeded in completing all the musical tasks, 3740 tunes could have been recorded. However, in several cases the subjects could not complete some of the tasks; for instance, they could not improvise tunes to the poems of Ady and József. This also happened when the chords proved too difficult to inspire a tune. The subjects were never forced to produce melodies since a negative result was also considered an important indicator in the experiment.

The tape-recorded music was transcribed into musical notation by 10 professional musicians whose transcriptions (both notes and bar lines) were checked by two independent judges; first by a colleague, then by a specialist, a former collaborator of Bartók. This was followed by a qualitative analysis of the tunes based on Bartók's method of folk-song classification. The analyses thus obtained were collated with the musical tastes obtained from the questionnaires and with the tests of tonal awareness. The entire work lasted eight years, from 1972 to 1980, including the analysis of the material in some 1300 pages and the writing of a 600-page book. (The book is now in press and will be published by the Hungarian Academy of Sciences.)

The scope of this chapter does not allow us to give a full description of the experiment but we can present our conclusions concerning the existence of a concrete level of generative musical creativity by means of the improvisational tests.

## Results

In our experiment we endeavoured to study the phenomenon of musical creative ability by using new methods. Our work displayed all the positive and negative aspects of a pilot study and this was why keen attention had to be paid to methodology. In such situations the first tasks to be performed include the testing of the method applied and, if possible, its validation. If we have been able to prove, to a reasonable extent, that it is possible and worthwhile to follow this path, our work would have achieved its main purpose.

We feel that certain achievements can already be seen. In the following we shall summarize these results and inferences, and in certain cases we would like to add a few reflections which are not strictly entailed by the results, but which we are encouraged to propound in view of the relatively large amount of accumulated material.

### Generative abilities in the universe of music

First of all, we have to note that the persons questioned willingly

could not really grasp the content and the message of the poems by Ady and József, their reasoning, their lyric world; this lack of understanding was indicated by frequent misreadings. The other is that, simultaneously with this, they could not grasp the structure of the poem and create an adequate form.

Thus, it may be said that the ideal form for most experimental subjects is the four-line song form; this is what they can best grasp, learn, and create. Those who know it better, repeat the first verse. The less accomplished cannot do this, but they endeavour to do so. And when they are faced with more difficult contents and poetic forms, they attempt to approach them from the angle of this ideal form; therefore they solve the task in an inadequate manner.

### Melodic features

Very much like for the structure, a clear picture was gained also for melodics: 60–70 per cent of the improvisations made for the Petőfi, Ady, and József poems were in the minor or major scales, minor character always having priority, as shown in Table 8.2.

**Table 8.2** *Use of major and minor modes*

| | Poem | | |
|---|---|---|---|
| | Petőfi | Ady | József |
| Total responses | 198 | 132 | 117 |
| Major | 38 | 18 | 17 |
| Minor | 64 | 45 | 32 |
| Major and minor mixed | 24 | 41 | 26 |
| Total major and minor | 126 | 104 | 75 |
| as % of all responses | 63 | 74 | 63 |

If we deduct the control groups (the musicians and foreign students) from the full sample, that is, if we only consider the improvisations of the Hungarian non-musicians, the proportion of the minor and major tunes is even higher (Petőfi songs 65 per cent, Ady songs 81 per cent, Attila József songs 66 per cent). Whereas non-professionals tend to stay in the same mode, young professionals tend to switch between different modes.

In the case of the Petőfi song, many subjects produced improvisations which had a folk-song character. The number of strictly pentatonic tunes (see Appendix, p. 286, Example 8.11) (10) was the highest here, which, together with the tunes having a mainly pentatonic character (25) accounted for 17 per cent of all the tunes. Modal tunes were similarly high

## The concrete level of generative musicality

It was not our purpose merely to prove musical generative ability, but also to measure its concrete form and level. The collected data may serve as a basis for drawing certain general conclusions. We shall endeavour to describe the most general features valid for almost all the groups and which, therefore, best characterize the musical disposition of the subjects of our experiment.

### Morphological features

We shall draw some general conclusions about the structure of the tunes improvised for the poems. The majority of the responses are characterized by strophic musical thinking on a level extending to the production and understanding of four-line musical forms. As soon as this limit has to be exceeded, they become uncertain.

Most subjects sang a strophic tune to the excerpt from the Petőfi poem. Forty-three subjects closely followed the strophic structure in the second verse (that is, they repeated the tune of the first verse with very little alteration); 47 endeavoured to achieve this but did not quite succeed; 54 sang two different verses of four lines each; another 30 considered the whole of the excerpt (eight lines) as a four-line tune and sang one musical line for each two lines of the poem. Thus, 174 of the 198 subjects were found to be capable of creating a more or less adequate form. The tendency to use the strophic structure is well illustrated by the fact that subjects tended to repeat the same musical line at different points in their song. Thirty per cent of the subjects used just four musical lines, 52 per cent used three, four or five lines for construction. The proportion of subjects using six, seven, or eight lines, or just one line, was 18 per cent.

The excerpt from the poem by Ady was more difficult. The test passage, of eight lines, was interpreted as such by only 44 of the 132 subjects; 63 subjects formed two four-line verses, while 20 created four two-line verses. Therefore, it seems that an eight-line strophe is more difficult to grasp, and even more difficult to structure, than a four-line verse. Many of the seemingly eight-line structures were quite loose in form and were not regular strophes.

Finally, the excerpt from the poem by Attila József was interpreted as an eight-line structure by most (70 of 117) subjects, for whom it was too difficult to establish a suitable system of interrelations between the lines. Thirty-two people created two four-line verses, while 15 created four two-line verses.

This may be due to two interrelated reasons. One is that most subjects

could not really grasp the content and the message of the poems by Ady and József, their reasoning, their lyric world; this lack of understanding was indicated by frequent misreadings. The other is that, simultaneously with this, they could not grasp the structure of the poem and create an adequate form.

Thus, it may be said that the ideal form for most experimental subjects is the four-line song form; this is what they can best grasp, learn, and create. Those who know it better, repeat the first verse. The less accomplished cannot do this, but they endeavour to do so. And when they are faced with more difficult contents and poetic forms, they attempt to approach them from the angle of this ideal form; therefore they solve the task in an inadequate manner.

### Melodic features

Very much like for the structure, a clear picture was gained also for melodics: 60–70 per cent of the improvisations made for the Petőfi, Ady, and József poems were in the minor or major scales, minor character always having priority, as shown in Table 8.2.

**Table 8.2** *Use of major and minor modes*

| | Poem | | |
|---|---|---|---|
| | Petőfi | Ady | József |
| Total responses | 198 | 132 | 117 |
| Major | 38 | 18 | 17 |
| Minor | 64 | 45 | 32 |
| Major and minor mixed | 24 | 41 | 26 |
| Total major and minor | 126 | 104 | 75 |
| as % of all responses | 63 | 74 | 63 |

If we deduct the control groups (the musicians and foreign students) from the full sample, that is, if we only consider the improvisations of the Hungarian non-musicians, the proportion of the minor and major tunes is even higher (Petőfi songs 65 per cent, Ady songs 81 per cent, Attila József songs 66 per cent). Whereas non-professionals tend to stay in the same mode, young professionals tend to switch between different modes.

In the case of the Petőfi song, many subjects produced improvisations which had a folk-song character. The number of strictly pentatonic tunes (see Appendix, p. 286, Example 8.11) (10) was the highest here, which, together with the tunes having a mainly pentatonic character (25) accounted for 17 per cent of all the tunes. Modal tunes were similarly high

in number, primarily those in the Dorian and Aeolian modes (that is, those nearest to minor), and so were those in the Phrygian and Mixolydian modes, which mainly occurred in turns and cadences. In the case of young workers we also met Lydian-acoustic turns (with raised fourth). All these modes are characteristic of folk-songs.

We obtain a similar picture when considering the stylistic characteristics of the tunes. The total of folk-song-based improvisations reached 32 per cent (see Appendix, p. 287; Examples 8.12 and 8.13). But by including those conceived in the style of the popular Hungarian art (see Appendix, p. 288, Example 8.14), the ratio reached 51 per cent. One may hesitate about how to categorize intonations typical of Hungarian popular songs; if these are attached to the two former types it may be said that 72 per cent of the subjects related their improvisations to a style that in some way originates from current Hungarian national music culture. About half of the remaining 28 per cent endeavoured to use non-Hungarian (that is, non-Oriental) intonations (see Appendix, p. 288, Example 8.15) which are related to European art music, while the other half—not being able to tackle the problem—produced a few melodic fragments based on seconds.

A different picture is obtained from improvisations to the Ady poem. Here the highest rate is represented by tunes in minor-like and major-like scales (78.7 per cent), presumably because this poem did not make a folk-song-like impression. Therefore, the number of pentatonic tunes and phrases, as well as of modal tunes and phrases, descreased considerably. However, the rate of compositions with a mixture of minor and major increased, from 18 per cent in Petőfi songs to 39 per cent in Ady songs. In the majority of the cases this reflected uncertainty which led to the dissolution of the classical major-minor polarity. The result of this seems to be a tendency to turn towards tunes that reflected a 12-note scale, not only in the case of the musicians, but also in that of the musically untrained middle school students and the apprentices.

Similar differences are found with regard to styles. Only nine of the tunes were folk-music-like in character (7 per cent) and only seven had the character of composed folk-songs (5 per cent). The impact of Hungarian popular song and operetta increased (26 per cent), as well as that of school songs, marches, pop, and religious songs (see Appendix, p. 289, Example 8.16). None of these styles is particularly appropriate from the Ady poem; they may all derive from the endeavour to present the contents of the poem in music different from that used for the Petőfi poem and for this purpose an intonation subjects happened to remember was applied. Only 24 tunes (18 per cent) show the composer actually tried to find a modern tonality.

The tunes created for the Attila József poem yield again different pictures. The rate of minor-like and major-like tunes again decreased, to approximately the level of those improvised to Petőfi songs, but presumably for other reasons. We may be justified in presuming that the József poem

had an even more remote content for the majority of the experimental subjects. They understood all the words but they could not grasp the sentences. However, there were some who obtained a feeling of the ballad, more so than in the case of the Ady poem, and this might be the reason for the repeated increase in the number of pentatonic and modal tunes and phrases, and for the unexpected appearance of other archaic structures (possibly from musical subconsciousness). This is how Lydian modes, Phrygian cadences, and whole-tone and acoustic scale fragments appeared; one experimental subject produced a pentatonic structure based on the 3/2/3/2 intervals (A–F sharp–E–C sharp–B descending structure with fourth structuring). The rate of tunes nearing dodecaphony had increased even compared to the Ady songs. The other reason for the decrease in the number of melodies using major–minor tonality was the increasing number of improvisations close to speech, almost pre-modal owing to uncertainty, based on seconds and varying one or two tones.

An even more complex picture is shown by the analysis of the styles. The rate of styles that are in no way suitable for formulating the musical material of this poem increased considerably. The rate of second-by-second tune fragments is 27 per cent, the same structure somewhat more fancifully conceived is 3.3 per cent, march tunes make up 5.6 per cent, pop 10 per cent, and all these combined constitute nearly half of the improvisations (47 per cent). This is followed by tunes of the Hungarian art song and operetta type (23 per cent) and by the composed folk-song (8 per cent) which can not be considered fit for the 'Ode' either. The rest is inspired by folk-songs, including songs reaching back to ancient tunes (8 per cent), and some initiatives resembling modern composed music (14 per cent). (See Appendix, p. 289, Example 8.17.)

On tune completion (Examples 8.8–10 above) we found that the same number of people (58 per cent) could complete both the major tune (Mozart) and the pentatonic melody (Bartók). In the case of the acoustic tune, not one single subject managed to complete the music without deviating from the tonality.

### Tonal-functional thinking on an elementary level

A general characteristic of improvisation related to minor–major tonality is a stereotyped level of tonal-functional thinking. Sixty-two per cent of the Petőfi songs, 57 per cent of the Ady songs, and 40 per cent of the Attila József songs fit a simple functional structure, either the simplest T–S–D–T cadence (I–IV–V–I), or the T–D or T–S variants of functions.

We obtained similar results from improvisations to sequences of chords (or harmonic sequences, in other terms). As shown in our following analysis, the subjects could improvise on different levels to the seven sequences of chords (Examples 8.1–7 above). The improvisations were

rated according to the degree to which they corresponded to the accompanying chords, according to traditional rules of harmony.

Based on the level of solution the seven sequences of chords could be separated into the following groups:

(1) An outstanding place is due to the harmonic sequence shown in Example 8.1, the classical I–IV–V–I cadence. Here, not only is the proportion of good solutions high (45 per cent), but the number of tunes harmonically unrelated to the chords is low (38 per cent). Also the number of variants that may be considered real improvisations is the largest; such tunes do not simply follow one or the other part, but create elaborations by arpeggiating the chord (18 per cent) and by flourishes. (See Appendix, p. 290, Example 8.18.)

(2) The second group includes harmonic sequences examples 8.2, 8.5, and 8.7. Example 8.2 comes into this group presumably because all the parts create suitable and understandable descending-ascending motifs. Example 8.7 is in this group because the addition of VI (in I–IV–V–I) extends harmonic sequence Example 8.1 in a manner that is familiar from hit songs; and Example 8.5 because the minor-character given by the Naples chord renders it accessible due to its romantic connotation. The number of good solutions in these three harmonic sequences is high. Example 8.2 inspired 41 per cent successful improvisations, 8.5 44 per cent, and 8.7 39 per cent. Despite this, we must distinguish these harmonic sequences from Example 8.1, because in these cases the proportion of unrelated and bad solutions is larger (38, 44, and 36 per cent) and even the acceptable responses include a larger number of tunes that merely follow the top part (that is, in unison) rather than being real improvisations.

(3) The third group includes harmonic sequences Examples 8.3, 8.4, and 8.6, which proved to be the most difficult ones. The rate of acceptable answers was 35 per cent for Example 8.6, 33 per cent for 8.3, and 31 per cent for 8.4. Sixty-four, 67, and 69 per cent of the experimental subjects could only sing notes completely unrelated to the chords. In the case of sequence 8.3, the problem was presumably caused by the flattened seventh degree; in Example 8.6 the source of problems (although popular in romantic music) was chromatics. A special place is occupied by sequence 8.4 which suggests Lydian-acoustic modality and therefore remains uninterpretable for most of the experimental subjects. In these three sequences of chords even the good solutions follow the parts, with only a few exceptions.

### Rhythm

The analysis of the rhythmic structure of the improvisations requires a preliminary note. All three poems or excerpts involve parlando-like rhythmic structures. None of our examples would adequately fit a giusto

(strict) rhythmic tune. Thus, we could achieve only a partial evaluation of sensitivity to rhythm, but this was enough to draw some conclusions.

The highest rate of parlando rhythm was created for the Petofi excerpt: a total of 103 (51 per cent). The vast majority of these had good prosody interpreting the 'ancient eight' in a 6+2 division. This provided for several intonations, between the explicitly old-style folk song and the style of the Hungarian popular song. Seventy songs (27 per cent) were prepared in a strict rhythm: however, the majority did not suit the poem, but were done with wrong prosody. The three-fourths rhythm was often used, though it can only forcibly be applied to the poem. The same can be said of march-like and hit song rhythms, as well as of repeated crotchets or quavers. It is, however, noteworthy that bound rhythms were primarily used by young people, and among them mainly by those attending special music elementary school and by the students of the vocational school. Finally, only 13 per cent belonged to the mixed category, who changed from free to bound rhythms and back.

In the Ady excerpt the proportion of improvisations in strict rhythm was the same (34 per cent). These were used mainly by the young, but now from all the categories. The proportion of free rhythms was 27 per cent, and most improvisations conformed to the prosodic and rhythmic requirements of the poem. The rate of mixed rhythms increased (39 per cent). Half of these (19 per cent) were still characterized by suitable prosody; these were basically parlando songs with bound scanning sections. The other half (nearly 20 per cent) endeavoured to impose some known style of Hungarian popular song (or of opera arias) on the poem, but the music turned out to be incompatible with the text due to the complexity of the latter.

Finally, in the performance of the 'Ode', the rates were again different. Here the rate of strict rhythms increased to 43 per cent, that of mixed rhythms remained unchanged (37 per cent), and that of the free-parlando rhythms decreased (25 per cent). Despite all this, the best rhythmic solutions were those in the free category and this was chosen most often by the music students. Others (peasants from Aradva'nypuszta, members of the workers' choirs) formulated parlando-recitativo hymn-like songs, or opera excerpts, selected ways with which they could best approach the poem according to their own knowledge. In this case, the mixed rhythms meant mainly infeasible tasks and a resort to scanning (the tune mostly moved in seconds here). The strict rhythm was again the 'solution' of the young in hit-like, march-like isorhythmic, or triplet-like style.

All this indicates that the subjects have a strong inclination for parlando-strict polarity which corresponds not only to the traditions of Hungarian folk-song but also has its equivalents in Hungarian popular song as well as in the style of hit music. This duality, naturally, also applies to art music but here there are far many more transitions and sepithese in genre than in

current music. However, for the experimental subjects this wealth is unattainable; thus, when they are faced with a task that cannot be solved either with a folk-song-like parlando or a march, they become uncertain. It is especially important to note that parlando is so closely linked to folk-songs (and to Hungarian popular songs) that it can hardly be applied to the Ady and József poems. The young (even pupils of music elementary school) are less acquainted with folk-song than older people, and that is why their improvisation is dominated by a march-like and a hit-like character.

## Conclusions

Altogether, these data demonstrate a certain parallel with Bernstein's (1970) sociolinguistic investigations and might lay the foundation of a kind of sociomusicological attitude. Bernstein makes a distinction between a tied-to-context and an elaborated method of use of the language. The situation in music is similar—the difference being that music as a language is less generalized or is generalized in a different way.

Our observations concerning the state or generative musical ability relate to our investigations regarding the development of popular music. In our opinion, generative ability determines also the spectrum of receptive experience. We listen in a distinct way to music that does not exceed the limits of our own generative musical ability. And so we can distinguish three degrees of musical receptive capacity:

(1) In the first degree, nearly the whole of the music listened to is interpreted in terms of structures that can be handled by the generative musical ability of the receivers themselves. This will be realized, for instance, in popular and beat music. Here the receiver can become identified with music. He feels as if he were able to make it himself.

(2) In the second degree, receptive generative musical ability does not comprise the whole of the composition received, only some of its elements. That is how for instance a listener with little formal musical education listens to, say, Beethoven's Symphony No. 5. The listener, being able to translate only a few elements—for example, the melody—to his own language, will understand the symphony only partially. The practice of everyday musical life proves, however, that he can enjoy it even in this case.

(3) Finally, in the third degree, the 'vocabulary' of the listener's generative musical ability and the composition listened to have nothing in common: in such cases no receptive experience comes about. These are the impressions of a listener brought up on Classical and Romantic music when hearing the works of, say, Webern or Bartók (not to speak of Boulez,

Stockhausen, or Ligeti). He does not meet a single familiar structure: he feels as though he were in a barren desert.

Our research not only revealed the attraction of the younger generation to pop music but also proved that this new type of investigation was a much more negotiable and applicable method in music psychology than the usual investigation of musical tastes. We can say that almost all the 11 groups had a specific musical characteristic or aspect which accorded with their different social backgrounds, and these musical attitudes definitely tally with the actual musical tastes and practices of the social strata in question.

## References

Bernstein, B. (1970). Social Class, language and socialization. In *Current Trends in Linguistics* (ed. A. S. Abramson). Mouton, The Hague.

Blaukopf, K. (1972). Musiksoziologie: eine Einführung in die Grundbegriffe mit besonderer Berücksichtigung der Soziologie der Tonsystems. VI. Auflag. Verlag Arthur Niggli, Niederteufen.

Chomsky, N. (1965). *Aspects of theory of syntax*. MIT Press, Cambridge, Mass.

Taylor, I. (1959). The nature of the creative process. In *Creativity: an examination of the creative process* (ed. Paul Smith). Hastings House, New York.

Weber, M. (1921). *Die rationalen und sociologischen Grundlagen der Musik*. Drei Masken Verlag, München.

# 9

# Young children's musical representations: windows on music cognition

LYLE DAVIDSON and LAWRENCE SCRIPP

Soon after children begin to speak, they begin to make marks on paper. At first, these marks form primitive depictions of everyday objects—flowers, houses, the family dog, and pictures of mummy and daddy. Later, marks change their meaning. For example, scribbled marks, which formerly stood for a stream, might now each represent a 'list' of things to get at the store. Gradually individual letters are formed and words begin to take shape. When children enter formal schooling, much effort is devoted to mastering the conventions of the written language, the importance of keeping the letters in a straight line, the common forms of spelling, word order, and grammar. Eventually, children's ability to make meaning from and express themselves in written language becomes a sign of understanding, an index of intellectual ability and perhaps even of station in life.

In music, the case is very different. Parents know that young children take great delight in listening to music, singing songs, and inventing songs of their own. However, they rarely suggest that their children write out songs they know. Performance ability alone is taken as the index of musical ability. Unlike language, mastery of the written form of music is neither viewed as necessary for musical ability, nor as an index of musical understanding. However, in music, as in language, the ability to represent relationships among elements in various ways is a more powerful measure of understanding than relying on performance or speech alone.

This raises many questions for those interested in young children and music. Is it possible that young children's invented marks or pictures could contain any musical meaning? Or is it possible that writing the words of a song could also convey some rhythmic grouping, melodic contour, or musical structure? Without training in standard music symbol systems, children are forced to invent symbol systems in order to indicate the musical features of a song or phrase. Thus, interpreting children's early symbolic representations of music is the initial task. It is only by finding

ways to understand these unique symbol systems that we can begin to look at the range of sophistication in their representational knowledge of music and for possible patterns of development.

While some findings have been reported about the child's representational knowledge of music during these early years (Bamberger 1980, 1982; Upitis 1985), recent researchers in developmental and cognitive psychology have shown more interest in the music perception and performance skills of young children. Drawing from Piaget's notion of discrete and ordered stages of cognitive development, researchers in music report considerable evidence for different levels of perceptual development (Gordon 1965; Petzold 1966; Zenatti 1969; Dowling 1982) and performance development (Werner 1948; Pflederer 1964; McKernon 1979; Davidson 1985) in pre-aldolescent children. It is our contention that for a more complete knowledge of children's understanding of music, their representations of music are critical 'windows' for viewing their musical cognitive development. For example, by asking children to write down a song they know 'so that someone else can sing it', we can find evidence of those musical features children think are most important for remembering 'how the tune goes'. Viewing the same children's musical representations over time, we can witness the development of their understanding of pitch, rhythm, or phrase relationships in the songs they know.

Consider the closing phrase of the common nursery tune 'Row, Row, Row Your Boat' (Fig. 9.1).

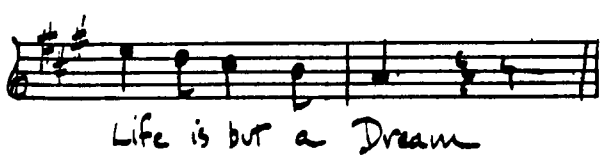



**Fig. 9.1.** Closing phrase of the song 'Row, Row, Row Your Boat'.

As a participant in our longitudinal study of young children's music notations, Janet provides remarkably clear representations of this musical phrase (Fig. 9.2). At age five Janet enactively traces only the rhythmic groupings of the 'fast and slow notes' while later (age six) she provides precise seriated lines that represent the melodic contour of the phrase ('the small lines are the higher notes'). Finally (age seven), she *simultaneously* suggests the rhythmic grouping and the melodic contour of the song (using the added dash to indicate rhythmic grouping and the layout of the words in a downward slope for the melodic contour), and additionally provides the text of the phrase.

This example illustrates three premises of this chapter. First, young

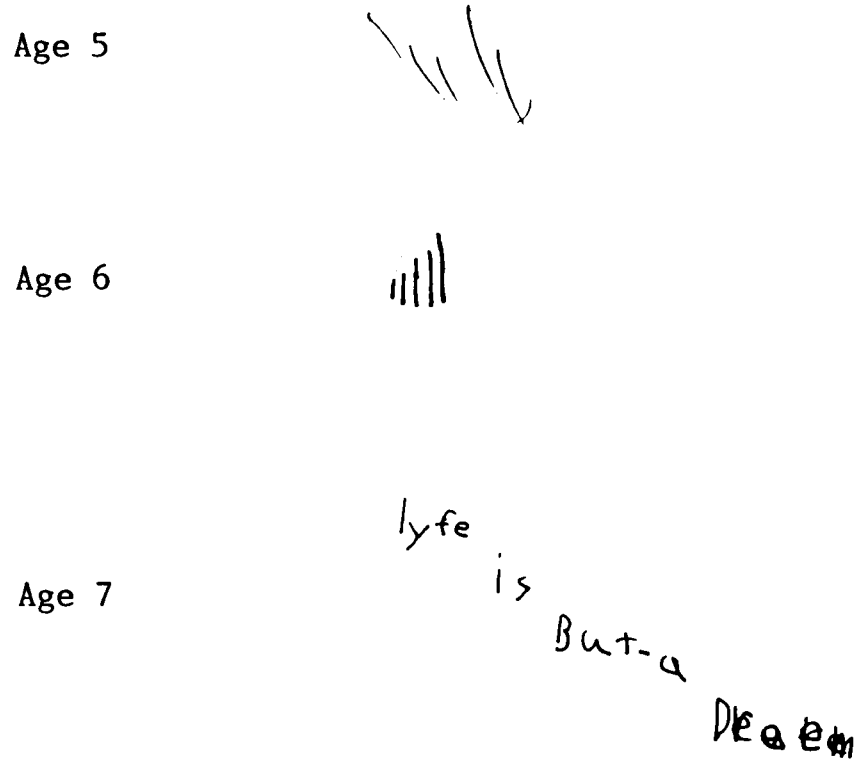



**Fig. 9.2.** Closing phrase of the song 'Row, Row, Row Your Boat' as represented by Janet at ages 5, 6, and 7.

children, untrained in music notation, can invent rich and articulate representations of the songs they know. Like many other children between the ages of five and seven, Janet effectively represents a wide range of musical features in her notations. Second, children's music notations reveal crucial insight into their understanding of music at various stages of musical cognitive development. Powerful developmental trends can be seen in Janet's and other children's notations as they increasingly specify and combine the rhythm and pitch dimensions of a single phrase. Between the ages of five and seven their notations become strikingly articulate and multi-dimensional. Third, taking both representational and performance development into account, musical pitch emerges as the primary component of children's musical cognitive development by the age of seven. Janet, typical of most young children we saw, increasingly focused on pitch relationships in her music notations and performance.

The study of young children's notations reported in this chapter offers a

wide-ranging view of their representations of music. While several recent studies have considered older children's (ages 8–12) spontaneous representations of rhythmic fragments (Bamberger 1980, 1982) and instructional effects in children's notations (Walker 1981; Upitis 1985), no previous study has explicitly looked at the early longitudinal development of younger children's spontaneous musical notations across various rhythmic, melodic, and structural contexts. Before looking at this new evidence we will first consider in more detail what music psychologists have found in music perceptual, performance, and representational development in children. Drawing on general models of cognitive development, we will piece together a 'three-dimensional view' of music cognition in which children's representations will play a vital role.

## Towards a three-dimensional view of young children's musical development

In the past, studies of musical development have been largely influenced by behaviourist, associationist, and atomist psychology. As a result, music psychologists traditionally tended to focus on extremely local stimulus-response tasks in music perception and performance. In music perception tests (Seashore 1919), for example, children were typically asked to discriminate two narrowly differentiated frequencies as 'same' or 'different' without regard to their musical temperament (within a standard scale) or their melodic context. Similarly, in music performance, research (Hattwick and Williams 1935) focused on children's ability to match individual pitches or to reproduce extremely short melodic phrases. Although this early research documented changes in children's perceptual or performance acuity with isolated pitches or melodic fragments, it offered little insight into children's ability to discriminate or produce pitch or rhythm patterns within broader, more realistic musical contexts.

More recently, studies in music perception and performance have taken a more global, cognitive approach. Current research in music perception, for example, has focused on children's developing ability to recognize melodic patterns within wider and more tonal contexts (Gordon 1965; Dowling 1982). Children were asked to recognize short melodies as 'same or different' depending on pitch transpositions or alterations of the contour (directionality) of the phrase. Results indicate that children at the age of five typically do not recognize two melodic fragments as the 'same' if the pitches are not identical in both fragments. However, by the age of eight, an important 'cognitive shift' occurs. Children now can recognize pitch relationships betwen two melodies as the 'same' even though the actual individual pitches of each melody are not shared (Pflederer 1964; Dowling 1982). Children's perception of music by the age of eight is now understood

as operating within an increasingly stable *tonal system* where the musical scale and a tonic reference are used to store melodic 'information'.

A 'cognitive shift' occurs in children's performance abilities as well. Whereas young toddlers approach the task of singing melodies by means of loosely structured pitch 'contour schemes' (Davidson 1985), six- or seven-year-olds sing familiar and invented songs with increasingly discrete and stable pitch reference (Werner 1948; McKernon 1979). The music transcriptions in Fig. 9.3 show performance development typical of most young children between the ages of four and seven. Jeannie, at the age of four, sang the melody of 'Twinkle, Twinkle, Little Star' without reference to the major scale system. By the age of six, however, her singing rivals adult performance standards.

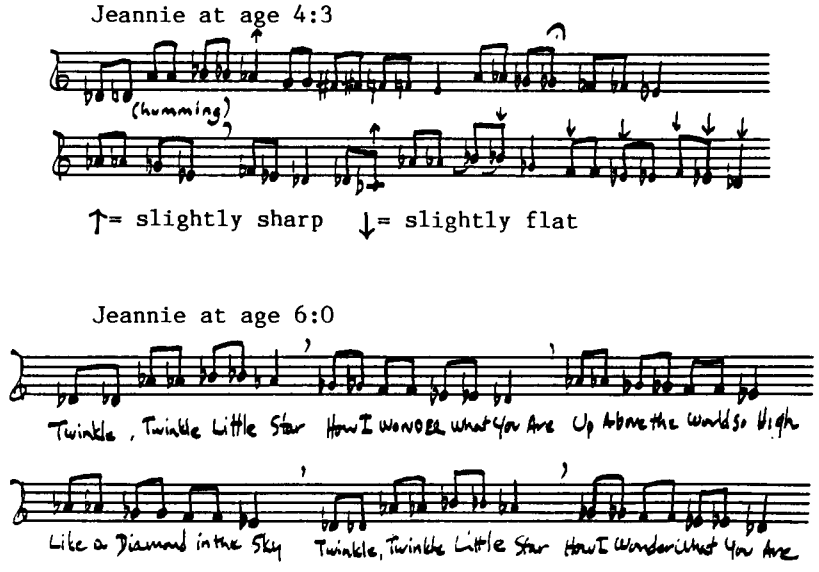



**Fig. 9.3.** Two versions of 'Twinkle, Twinkle, Little Star' by Jeannie.

Taking two dimensions together—perception and performance—we can characterize children's musical development as progressing rapidly through parallel stages of increasingly systematic knowledge. This development is measured by the emergence of a stable tonal framework. This is in contrast to rote imitation of musical fragments or perception of isolated pitches. From this 'two-dimensional' point of view we can see that untrained children between the ages of four and seven develop into surprisingly sophisticated musical listeners and performers.

Still, important questions remain regarding the depth of children's

musical understanding. For example, although children may be able to simultaneously perform the rhythms, pitches, and text of a familiar song, we know very little about which of these components children themselves regard to be most important to represent. In fact we have no idea whether young children primarily attend to more central components (rhythm and pitch) or to the relatively peripheral aspects (words or story-line) of the music they sing. Nor do we know how these patterns may change with age. Questions such as these are not easily answered with the research techniques traditionally used in studying young children's music perception and performance development. By turning to yet another dimension, children's initial representations of their performance knowledge, we may be able to better understand the roots of their developing musical comprehension.

## Music and representational development in children

Psychologists researching cognitive development in other domains have interpreted children's representations from several different points of view. Children's drawings, for example, have been used as evidence for levels of mental development (Werner 1948). From this point of view representational development occurs as a result of general mental development which Werner believes proceeds 'from a state of relative globality and lack of differentiation to a state of increasing differentiation, articulation and hierarchic integration' (1957). Piaget, particularly with his study of children's knowledge of spatial relations (Piaget and Inhelder 1956), used young children's representations to show precisely ordered levels of cognitive development. Piaget argues that children's representations are the product of what children 'know' rather than what they 'see'. To substantiate this claim, in later research, he asked young children to make drawings of a set of seriated vertical lines from memory, and reported positive development in their drawings *without perceptual intervention* six months to one year later (Piaget and Inhelder 1973). These drawings suggest that children later had access to a more intellectually advanced 'cognitive lens' which allowed them to *refocus* on their stored mental images. Other psychologists report that children's representations can also reflect stages of cognitive strategies (Bruner 1973) and levels of cognitive development within artistic domains (Goodnow 1977; Gardner 1980). Olson (1970) reports the importance of children co-ordinating pencil and papr skills with their perceptual skills in order to explore the link between cognitive skills and their representation.

Clearly children's representations are of critical value to a wide range of psychologists studying cognitive development within various domains. Similarly, children's music representations are crucial to our understanding of their musical cognitive development. Only very recently have music

psychologists begun to view this 'third dimension' of children's musical cognitive development.

## Cognitive processes and children's representation of simple rhythms

By asking young children and musically untrained adults to perform simple rhythms and then later represent them, Bamberger (1980, 1982) allows us to glimpse the 'cognitive structures' children use to express their understanding of music. Bamberger's research suggests that writing down rhythms involves a 'figural-formal transaction' (Bamberger and Schon 1977). That is, when representing these rhythmic fragments, notators may attend to the more figural properties (for example the motivic groupings formed by adjacent 'fast' and 'slow' notes) or to the more formal/metric properties (for example the organization of the surface patterns with respect to an underlying pulse) of the target phrase. Bamberger reports a developmentally ordered typology of symbols used by her subjects in respect to either side of this 'transaction'.

Bamberger reports that four- and five-year-old children typically represent simple rhythms (1980) using primitive enactive scribbling to trace the action of the rhythmic performance. Later on, children (age six or seven) use early 'figural' representations that group the surface units by duration, or notations that merely record the number of events in the phrase. Only after training is there evidence for a 'cognitive shift' towards a system where figural grouping can become subsumed by a formal/metric system classifying beat units, rhythmic patterns, and their co-ordinations. Bamberger reports:

> Results indicate that musically untrained adults and older children (9–12) seem to develop the capacity only to make figural drawings: metric/formal drawings are made almost solely by those with the trained ability to read standard music notation. (1980, p. 172)

Bamberger's work provides an intriguing first view of young children's music representational development. We can see categorically different processes in young children's notations as they discover ways of restructuring their performance knowledge of simple rhythms. Through this cognitive 'window' we begin to see how musically untrained children (without knowledge of a conventional code) represent the complex and sometimes conflicting musical dimensions of surface patterns, underlying metric pulse, and their possible co-ordinations in 'simple' rhythmic phrases.

However, Bamberger's research also raises vexing questions concerning children's musical cognitive development. Considering children's music representations within the larger framework of cognitive development, we need to ask why young children without musical training are unable to develop beyond the 'figural' representations of simple rhythms while

'formal' development apparently occurs in other domains (Flavell 1963; Furth 1969). Within the domain of music, on the other hand, we need to know why children's music representations of rhythm apparently lag behind the more developed levels of their pitch perception and performance described earlier. Because we only have evidence detailing children's representations of short rhythms, we have no way of directly comparing the development of tonality in children's perception and performance with their representational development. In sum, Banberger's work offers an enlightening point of departure for assessing children's music representations. A wider-ranging study is needed in order to get a more complete picture of children's music representational development.

In the next sections we will report findings from the Harvard Project Zero notations study (Wolf *et al.* 1986), the first longitudinal study to look at children's (ages five to seven) music notations across a wide range of musical contexts. We begin by describing in some detail the structure of the research.

## The research design: children's representations of musical content across various contexts

In this study we asked 39 musically untrained children to perform a wide variety of tasks. In the spring of each of three years, we requested the children (over two separate sessions) to reproduce short rhythmic or melodic patterns, recall a familiar song ('Row, Row, Row Your Boat'), and learn a short, unfamiliar song. At the beginning of each session it was explained that all the items were to be written down in their 'own music book'. To ensure that each child had a secure performance knowledge of the items, a pattern of phrase of the song was repeated either until the child made a correct performance or until the level of performance had stabilized. Afterwards, children were asked, 'Write the song down on your paper so that someone else who doesn't know the song can sing it back.' For those who objected to the task of writing down a song they'd known since they were 'babies', it was suggested that they should write down the song carefully for children from 'other countries' who might not know the song. They were also told they might use any mark they wished to write down the music if it helped them remember the sound of the rhythms or melodies. The children were allowed as much time as they needed to write in their 'music books'.

Immediately upon completing their notations, children were asked to perform each item while reading from their 'music books'. The experimenter explained that he wanted to 'see if the music book would work for someone who couldn't remember the clapping or the singing'. Later on in the session the experimenter asked each child to read back the notations in the 'music book' in random order to determine whether or not children were using

their own notation rather than recalling the target items by rote. At the end of the final session, each child was asked to explain or demonstrate 'how the music book works'. Each session was about 20 minutes long.

The materials for the tasks were simple. Two cassette tape-recorders were used, one to play the pre-recorded musical examples, and one to record the child's performances for future transcription and scoring. Children were each given a 12″ × 18″ piece of white drawing paper and a choice of coloured felt-tipped markers with which to notate their performances. When the task was completed (and after the child had left) the experimenter freely added comments to the notation (in pencil) or recorded remarks that might serve to clarify the child's process of making the notation.

Rather than simply focussing on one set of materials, this research design allows us to analyse children's representational as well as performance development across various musical contexts: rhythm patterns, pitch patterns, a familiar song, and a new song (see figure 9.16). These four different contexts share the same pitches or rhythms. In designing our study we structured our tasks to provide answers to the following broad questions:

(1) What is the general nature of young children's (five to seven) musical representations?

(2) What musical dimensions are young children able to extract from the melodic fragments or familiar songs they are asked to write down?

(3) What evidence is there for development in young children's music notations from the ages of five to seven?

(4) How does context affect children's notations of music?

In following sections we will analyse the notational responses to the familiar song task. First, we will show how young children use an extremely wide range of invented and borrowed graphic symbols to create increasingly sophisticated notational systems. Second, we will examine their ability to extract and code various musical dimensions without the use of standard musical symbols in their music representations. Finally, we will present evidence of representational development in children's ability to increasingly specify various musical dimensions such as phrase strucure, musical units, rhythmic pulse and grouping, and pitch contour. This allows us to appreciate the remarkably rich and articulate knowledge of music which many of these young children can express through their invented notations.

## Borrowed symbols and invented systems: young children's first notations of song

Children approached the task of writing down a familiar song with great earnestness. For five-year-olds, the task of choosing an appropriate symbol

for representing music involves a good deal of personal investment. Many openly confided they didn't know how to do it and it was not uncommon to observe children pause several minutes in silence before writing down their first mark. Careful not to prompt young children towards a particular set of graphic symbols, we witnessed an extraordinary range of symbol use. More importantly, we saw how they adapted and organized various families of graphic symbols into *symbol systems* that map various musical dimensions of their songs.

We have grouped the children's notations of song into five symbol systems: the *pictorial*, *abstract patterning*, *rebus*, *text*, and *combination/ elaboration* symbol systems. Although these systems can be identified largely by their ingredient graphic symbol types, our analysis will focus primarily on which features of the song were coded and how they were organized within the notation.

### The pictorial system

This system features a variety of pictures, icons, and images that young children use to record the song 'Row, Row, Row Your Boat'. These evocative 'picture drawers' represent the song with a drawing that can effectively render the title of the song, the look of a 'music book' or a 'person singing', or perhaps the entire story-line condensed into one salient image. Because these 'picture notations' do little more than record the song as a global event or action, this system usually remains a primitive and ineffective system of notation. Indeed, the pictorial system probably best meets the common adult expectations of the 'artistic', yet notationally untutored, child.

Emily's notation is a particularly good example of the limitations of her particular symbol system (Fig. 9.4). After observing Emily carefully drawing a rich landscape of rippling water, clouds, and a prominent boat, we asked her how 'to read the song' from this picture. Patiently explaining 'the book isn't finished yet', she turned the paper over, folded it in half, and began drawing the notes that 'belong' in her 'music book'. Although the picture of the boat may remind us of the title of the song and the notes drawn inside tell us that it is a music book, there is little here to help us remember how the tune goes. Alas, these notes appear to have all the 'look' but little of the substance of standard music notation.

### The abstract patterning system

This system usually features lines and dots which children use to represent the melodic units ('notes') of the song and record the rhythmic groupings, underlying pulse, melodic contour, or phrase structure. Due to the ambiguous nature of these symbols, it is often difficult to tell at first which

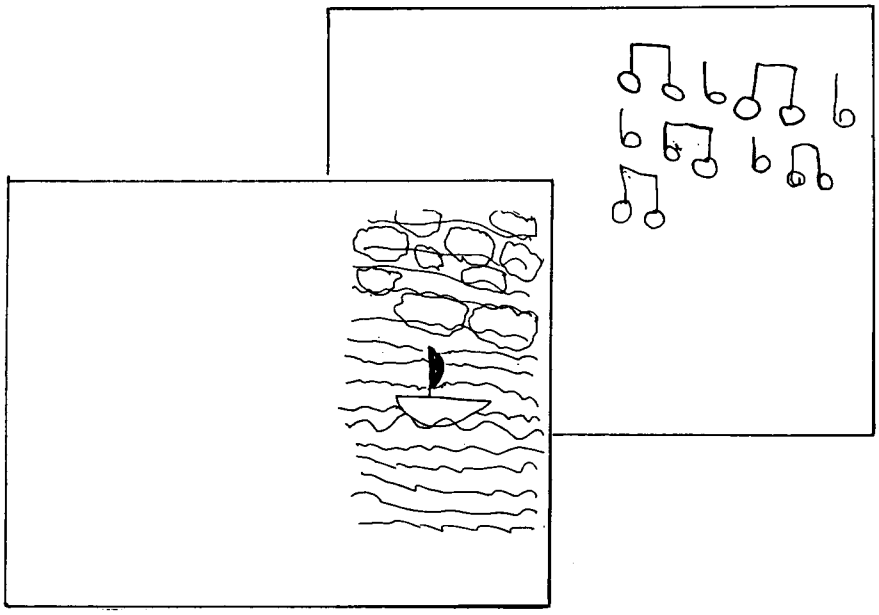


**Fig. 9.4.** Emily's (age 6) pictorial system.

musical dimensions of the song are being coded, but it is clear that these 'abstracters' are primarily concerned with the rhythmic and melodic *relationships* of the notes of the song rather than being content with merely representing the story-line or the text.

As is typical with 'abstracters', no conclusive interpretation of Janet's notation (Fig. 9.5) can be made without a transcript. What at first appears to be an arbitrary set of seriated vertical lines turns out to be an elegant system for recording pitch relationships in the song. And, once we understand that the smaller 'sticks' indicate the higher notes, we can track through the entire system with little difficulty. We see that the structural boundaries of the musical phrases are indicated not only by gaps in the notation (that is, musical 'rests'), but also by the corresponding intervallic pitch relationships recorded throughout the song (for example, the last note of phrase 1 appears to be equivalent to the first note of phrase 2 or, more remarkably, the first note of phrase 1 seems equivalent with the last note of phrase 3 and the first note of phrase 5). In fact, after Janet performed her notation while pointing at the lines, she spontaneously announced that the last phrase 'is just like the first one'!

### The rebus system

This system usually features elaborate and flexible *alternations* of icons,

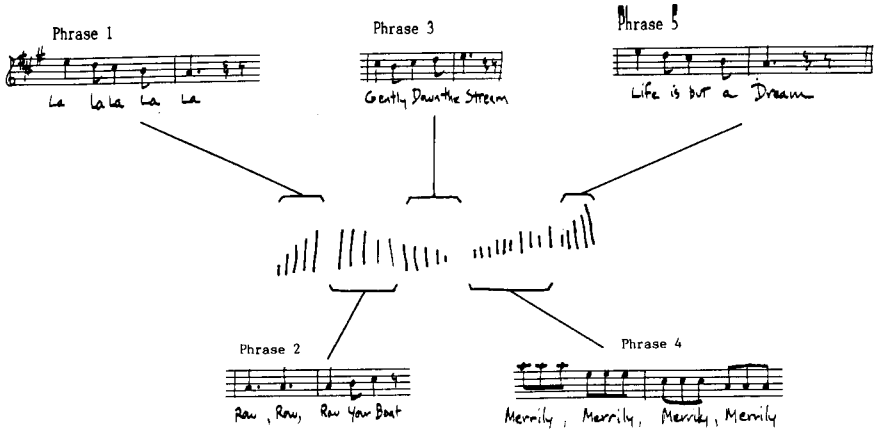



**Fig. 9.5.** Janet's (age 6) abstract patterning system with corresponding phrases of 'Row, Row, Row Your Boat' in standard notation.

conventional signs, and words children use to record the text of the song. While 'rebus makers' effectively depict the text or illustrate the lyrical content of the song, they can also occasionally reveal the underlying pulse (word stress) or pitch direction (contour) of the tune.

At first glance, it appears that Aaron (Fig. 9.6) has freely selected letters, words, and icons only to track the lyrics of the song. With closer inspection, however, we see that there are occasional omissions of the text, which suggest a focus on the rhythmic organization of the words as well. In phrase 2, for example, Aaron redundantly draws three 'rowing boats' and spells out the word 'boat' while leaving out the word 'your'. Apparently Aaron is concerned more with notating the underlying *pulse* (or stress) of the rhythm than with merely recording all the 'units' of the text. Phrase 5 shows a similar understanding of musical pulse as the stressed words are depicted and the 'weak beats' appear to be deliberately left out. In phrase one, he shows sensitivity to the longer duration of the final 'la' and explains that the last 'la' is 'different from the rest'. Typical of rebus makers, Aaron inconsistently yet imaginatively suggests various salient dimensions of the song in his notations.

### The text system

This system features the words, letters, or imitations of conventional language symbols children use to record the text of the song. Although these children seem primarily concerned with simply 'writing down the words', they also often take advantage of graphic layout devices to code

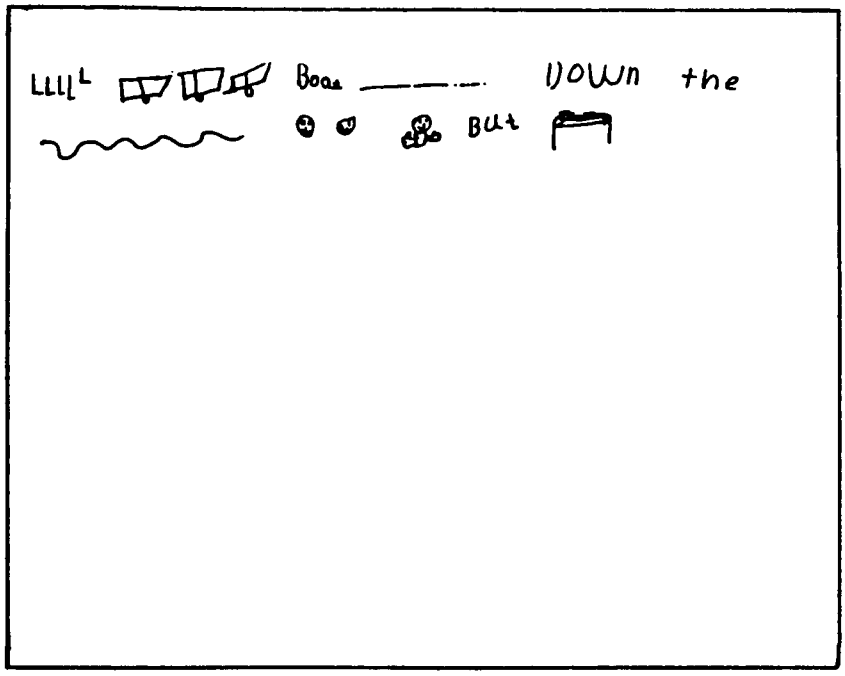



**Fig 9.6.** Aaron's (age 6) rebus system.

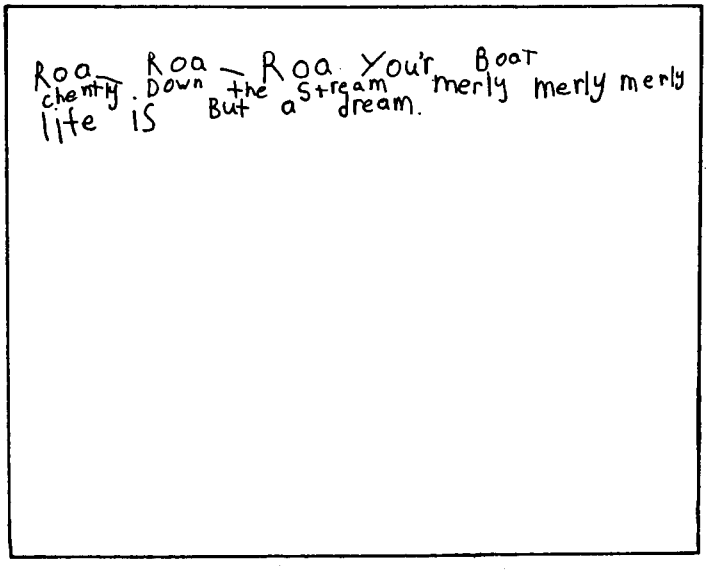



**Fig. 9.7.** Eleanor's (age 6) text system.

the verse structure, the direction of pitch, or the rhythmic grouping of the song.

Eleanor's notation (Fig. 9.7) is typical of most text notators. Eager to show off her newly learned writing skills, she quickly wrote down the words of the song. Although her transcription of the lyrics is almost impeccable, more importantly, the notation also clearly reveals her attention to the phrase structure of the song. Looking closely at the text system we see that she uses 'carriage returns' and additional space between 'stream' and 'merly' to show the individual lines of the text. Also she provides elongating marks after the first few words ('*Roa*') and demonstrates their purpose to the experimenter by gesturing the pulse and sustain while singing the opening phrase. Like most text systems, however, this notation has fallen considerably short of a consistent, systematic representation of rhythmic or pitch dimensions of song.

### The combination/elaboration system

This system features *simultaneous* use of abstract symbols and words to represent the text and musical dimensions together. Unlike the pictorial and rebus system, this system does not utilize pictures or icons, nor does it primarily focus on the lyrical content of the song. 'Elaborators' are children who consistently employ abstract symbols to modify the text to show how the words are to be sung. 'Combiners' are children who record both the text and abstract patternings of the 'notes' as two separate and equally necessary ingredients of the song.

Mary's notation (Fig. 9.8) demonstrates both approaches children used in constructing their combination/elaboration system. Like most 'elaborators' Mary records the text first and later modifies these 'notes' with marks (arrows) showing relative pitch direction. Like most 'combiners' she works through one entire system first and then fashions a separate series of marks to be read simultaneously with the first. In this case Mary uses abstract symbols to represent the 'long and short notes' of the song. As with all children using combination/elaboration notation, Mary is embarking on an ambitious path: the integration of the melodic and lyrical content.

In sum, what is more impressive in these examples of children's symbol systems is the sophistication with which they approach the task of

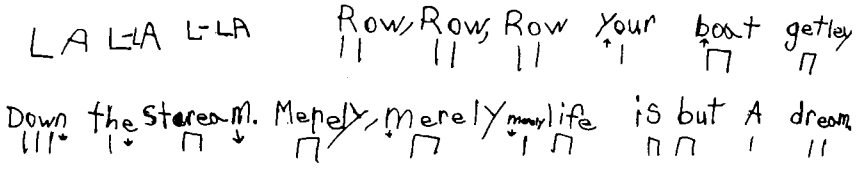



**Fig. 9.8.** Mary's (age 6) elaboration and combination system.

representing music. Using highly individual symbol systems these children can successfully code not only the song title, the story, and the text, but also the melody's more salient musical features—its underlying rhythmic pulse, pitch contour, or phrase structure. In the most gifted of notators we see children attending to several musical dimensions *at once*, constructing highly elaborated or elegantly patterned descriptions of the song. Recognizing the wide range of symbol use and the degree of sophistication with which children employ these symbols to write down a familiar song, new questions can be addressed: What changes in symbol use occur with age? Can these changes be ordered into developmental sequences?

## Signs of development: preliminary evidence for children's emerging musical cognition in their representations of song

The sections that follow present evidence for development trends in young children's representational knowledge of music between the ages of five and seven. In this section we will only look at children's developing notations of the familiar song 'Row, Row, Row Your Boat'. In later sections, we will look at children's notation of a single phrase across various musical contexts.

Our data show that children between the ages of five and seven not only change radically in their choice of symbol systems to represent songs, but also show a dramatic increase in the level of sophistication of their notations. Examples from our entire sample of children's songs can be organized in the following developmental sequence:

Age 5—children (who may not have the option of writing words) primarily use pictures and abstract symbols to show musical structure or the units of the phrase.
Age 6—children primarily use abstract or language symbols and sometimes organize the units to show either underlying pulse, rhythmic grouping, or melodic contour.
Age 7—children combine and modify language with abstract symbols and consistently show multiple features of the music with particular attention to melodic contour.

Janet's and Mary's song notations (Fig. 9.9) between the ages of five and seven provide clear examples of this developmental sequence.

Mary's initial pictorial system exemplifies the average child's notation at the age of five. Although she makes a clear distinction between the introductory phrase and the rest of the song, little attention is given to the musical units within the phrases of the song.

Janet's precocious abstract notation at age five and Mary's elaborate rebus at age six are both typical of the average six-year-old's attention to

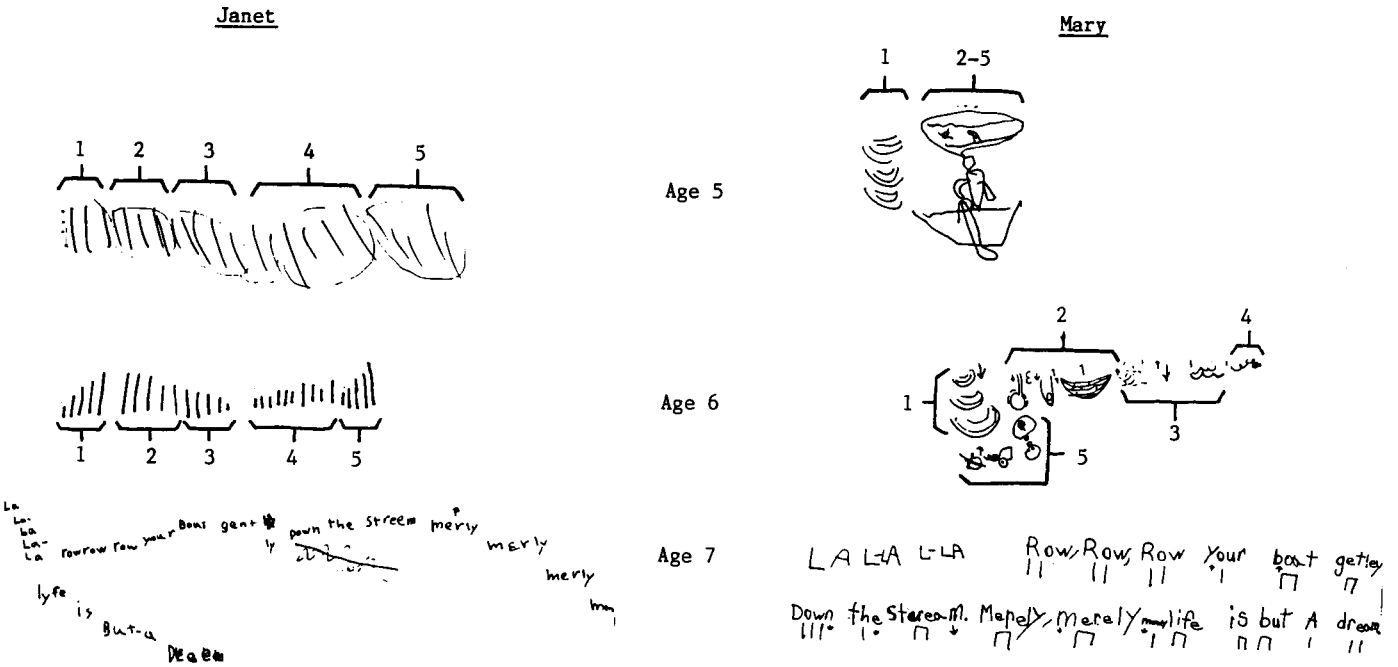



**Fig. 9.9.** Janet's and Mary's notations of the song 'Row, Row, Row Your Boat' at ages 5, 6, and 7 (phrases numbered as in standard notation).

the rhythmic organization of the units of the song. Here Janet provides an enactive tracing of the musical units with some semblance of their rhythmic pulse (phrase 1) or grouping (phrases 2 and 5); Mary clearly depicts the words of the text that receive the underlying pulse or stress (phrases 1, 3, and 5).

Both Janet and Mary anticipate the average seven-year-old child's attention to melodic contour at age six but without using language symbols. More typically, the average seven-year-old uses the combination/elaboration symbol system to record the text of the song while *simultaneously* providing indications of melodic contour.

### The symbol system sequence

Janet and Mary are both examples of particularly fluent and richly sophisticated notators. As different as their 'tracks' of development at first appear, the notations of these two children suggest that the pictures or abstract patterns used to represent songs at age five will later be superseded by words combined or modified with abstract symbols. Data from the entire sample ($n$=39) over three years provide considerable evidence for this developmental trend in children's symbol system use.

An analysis of figure 9.10 shows that most children (69 per cent) use either the *pictorial* or *abstract patterning* systems to represent songs at age

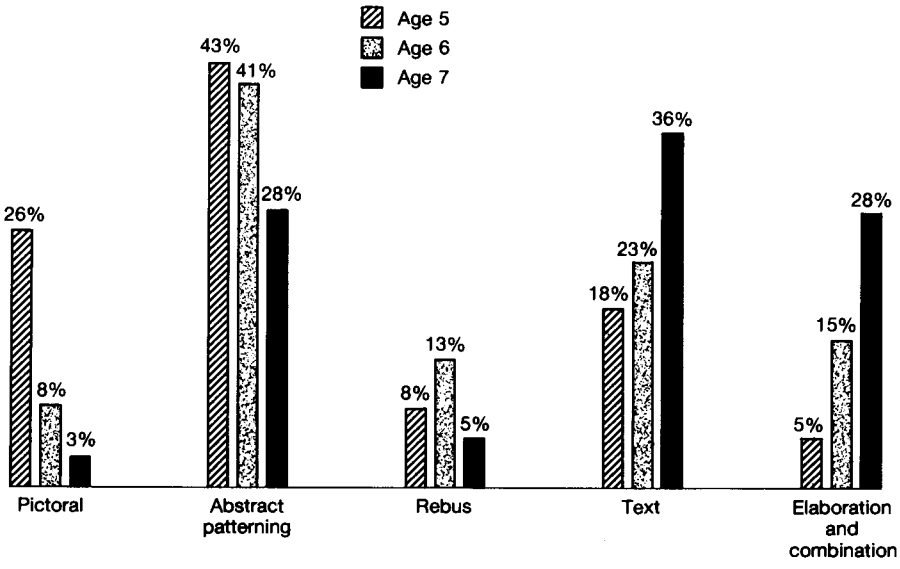



**Fig. 9.10.** Change in use of symbol systems for representing the song 'Row, Row, Row Your Boat' by children ages 5–7.

five while a majority of children (64 per cent) are using either the *text* systems or *combination/elaboration* systems by the age of seven. Statistical analysis supports the notion of these ordinal trends in symbol system choice with age [Friedman Chi-square = 9.94, 2 df, $p<0.007$, $n = 39$].

Probing further into this 'symbol system' sequence, two findings are particularly important to our understanding of the children's music representational development. First, since we assumed that most five- and six-year-old children would draw pictures to represent the song, the predominant use of abstract patterning systems throughout these early years was totally unexpected. Although young children receive considerable instruction with language literacy skills and are encouraged to draw, many feel compelled to use a totally different set of invented symbols to represent music. This suggests that the average five- or six-year-old perceives music as a separate domain, requiring its own particular symbol system uniquely suited to the task of representing musical features.

Second, although young children increasingly abandon pictorial systems and employ their newly acquired writing skills to transcribe the text of the song, many find language symbols insufficient for the task of representing the song. Surprisingly, children insist on representing important musical features as well as the words of the song. This is particularly signficant at age seven when there is a dramatic increase of elaborating marks or abstract patternings in conjunction with text.

This sequence illuminates significant trends of symbol system development. In the next section we will look at development in levels of extraction of musical dimensions in children's notations and possible correlations with symbol system or performance development.

## Representational development in song

Analysing children's representational development in song, we coded children's notations according to the musical dimensions they were able to extract and represent. In the dimension of musical pitch, we predicted a simple sequence of development. We assumed that children would at first code the *units* (notes) of the phrase (level 1), and later indicate the *melodic contour*, or directionality, of these units (level 2), with a few older children eventually mapping the *intervallic boundaries* of these melodic contours (level 3). Although we did not expect many children to code exact pitch relationships, that is, *regulated pitch* (level 4), this was the next logical step in pitch representation. This was confirmed throughout the various contexts of our pitch tasks.

In rhythm, the sequence predicted was unconfirmed by our data. Originally we assumed that children would first code *units* (level 1), followed by the *grouping* of the units (level 2), the *underlying pulse* of the phrase (level 3), and finally the rhythmic grouping regulated by pulse, that

is, *pulse plus grouping* (level 4). This sequence was based on Bamberger's (1982) research which suggested that untrained children first attend to 'figural' features of rhythm (the grouping of fast and slow notes) while only a few later attempt a more 'formal' regulation of the rhythmic figures. We were unable to confirm this finding. Surprisingly, we found many children representing the underlying (metric) pulse or stresses of the text before providing any indication of the (figural) grouping of these units. As a result, our notational sequence for rhythm allows for selecting *either* grouping or underlying pulse as an intermediate step towards simultaneously constructing both pulse and surface groupings of rhythms.

The following examples (Fig. 9.11) of the introductory phrase of our song shows how children achieved various notation levels in rhythm and in pitch.

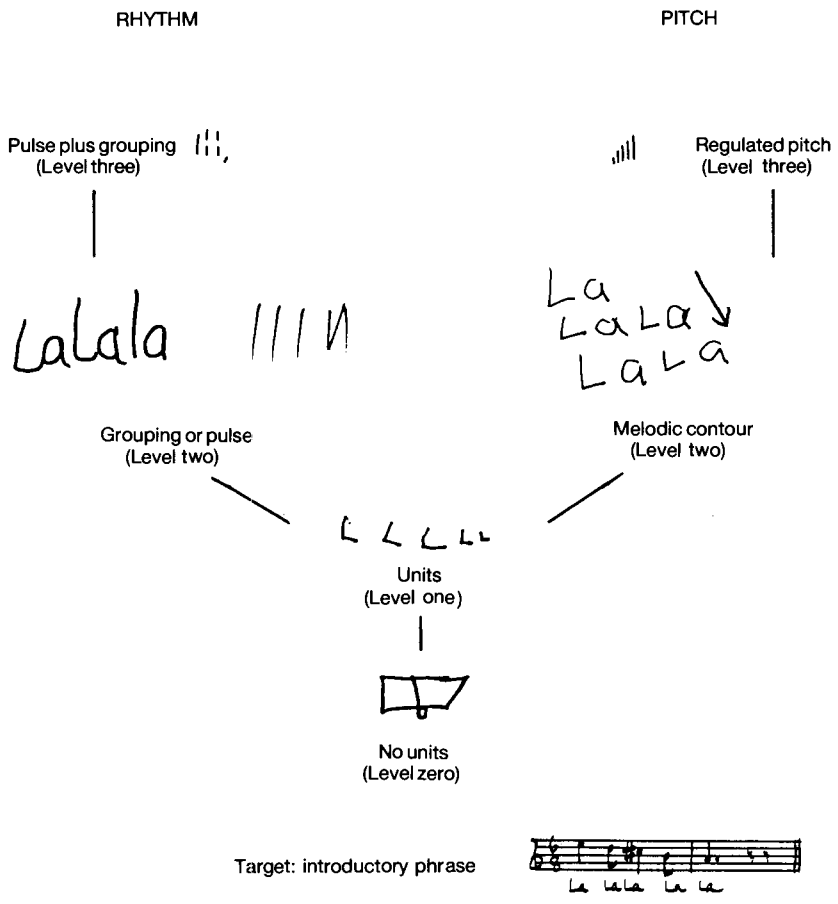



**Fig. 9.11.** The ordinal notation levels for pitch and rhythm.

Overall, we scored children's notations according to the level of the rhythm or pitch they consistently (three out of five phrases) encoded in their representations of song. A summary measure of notational consistency in rhythm and in pitch was achieved by calculating an average score produced by these levels. Looking at the mean notation scores in Fig. 9.12 we see that, on the average, children at ages five and six were able to represent the song only at the level of musical units. At age seven, however, there is a dramatic increase in representational skills as children demonstrate higher levels of rhythmic structure (that is, grouping or pulse) or pitch relationships (for instance, melodic contour or relative interval boundaries) in their notations of song.

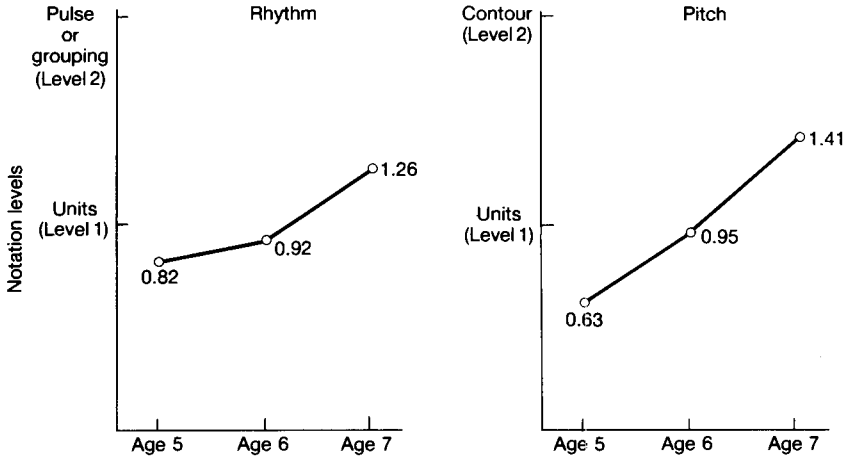



**Fig. 9.12.** Mean notation scores for rhythm and pitch in song by children ages 5–7.

Development in levels of representation of pitch can be clearly contrasted with representation of rhythm. Repeated measures one-way ANOVA suggests highly significant improvement in pitch notation levels [Friedman Chi-Square 20.55, df 2, $p<0.00005$] while improvement in rhythm notation levels appears to be relatively weaker [Friedman Chi-Square 6.62, df 2, $p<0.05$] over the three-year period. From these results it appears that young children are increasingly able to develop notations that consistently show the melodic contour rather than the surface rhythms or underlying pulse of the song.

Instead of looking at children's choice of graphic symbols as a measure of their notational sophistication, we can view children's music representational development in song independent of their choice of symbol systems. If we factor out children's symbol system choices, we are in a better

position to assess the relationship between young children's level of performance and levels of representational skills using familiar songs.

Analysis of the relationship between symbol system choice and level of notational sophistication shows decreasing correlation with age. In Fig. 9.13, we see that the more 'advanced' symbol systems (for example text or elaboration and combination systems) predict better notation scores in rhythm and pitch at age five but correlate less strongly a year later. In other words, one can expect five-year-olds to achieve higher levels of notation only if they are using a more sophisticated symbol system. In contrast, by the age of seven there are no significant correlations between children's notation scores and their choice of symbol system. In other words, the level of notation achieved and the symbol system used become increasingly independent for older children.

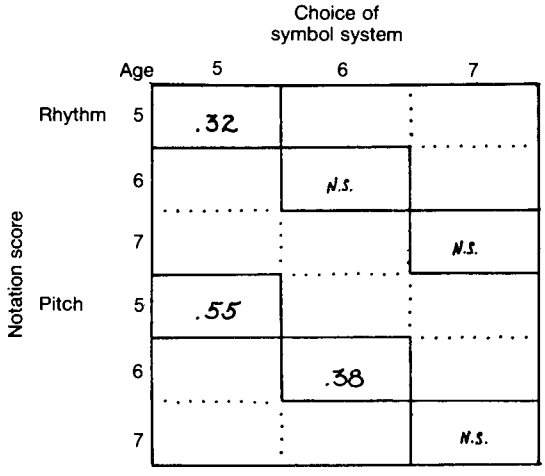



**Fig. 9.13.** Matrix showing significant ($p<0.05$) rank correlation coefficients (Kendall Tau) between symbol system choice and notation scores in children's notations of song.

Surprisingly, children's choice of symbol system in representing songs comes to have little bearing on development in notation. The more sophisticated the level of children's representations, the less of a factor the choice of specific symbol system becomes. In other words, the vehicle of expression becomes less visible with development.

## Musical performance and representational development in song

Throughout the three-year study, we observed the development of children's performance knowledge of the song 'Row, Row, Row Your

Boat'. For rhythm, children were scored according to their ability to reproduce accurately the *number of units*, steady *underlying pulse*, the *surface grouping*, or the co-ordination of *underlying pulse plus the rhythmic grouping* of a phrase of the song. With respect to pitch, the children were scored according to their ability to match the *initial pitch*, the *melodic contour*, the *interval boundary*, and the *key* (scale) of the phrase.

From Fig. 9.14 one can assess performance development in rhythm and pitch in terms of the percentage of items accurately matched by children between the ages of five and seven. Performance in both rhythm and pitch improves over the three-year period but, as we take a closer look at these overall performance scores, we see strikingly different profiles of development. In contrast to reproducing the rhythm, the task of accurately singing the pitches of a familiar song proves to be initially far more difficult. Whereas five-year-old children, on the average, reproduced the rhythm measures with 85 per cent accuracy, approximately half of the pitch measures were matched accurately. With age, however, this gap narrows considerably, suggesting that either rapid and substantial improvement in pitch accuracy occurs in contrast to rhythm, or that rhythmic development is more precocious and thus its growth is less dramatic over the course of our study.

Statistical analysis supports the notion of rapidly improving pitch in children's performance development. Repeated measures one-way ANOVA and Duncan Multiple Comparison procedures (Cody and Smith 1985) suggests highly significant [$p<0.001$] improvement in both rhythm and pitch over the three-year period. Comparisons between years show significant [$p<0.05$] improvement in rhythm between the ages of six and seven while pitch dramatically progresses in each of the three years. As in representational development, longitudinal growth in young children's performance accuracy in song between the ages of five and seven apparently is best viewed within the dimension of pitch.

At no point in the three-year study did performance skills correlate significantly with children's choice of symbol system to represent music. Apparently, the ability of children at ages five through seven to accurately sing a tune or clap a rhythm in no way predicts their choice of graphic symbols to notate the music. This finding allows us to better appreciate music research as a domain uncontaminated by conventional symbol systems that may influence young children's ability to represent their understanding of that domain. Furthermore, it underscores the initial independence of children's notational and performance abilities.

In Fig. 9.15 one can view the developing relationships between performance and representational knowledge of the song. When the dimensions of pitch and rhythm are combined, we see significant correlations between performance skills and level of representational knowledge only at the age of seven. This suggests that children, as they

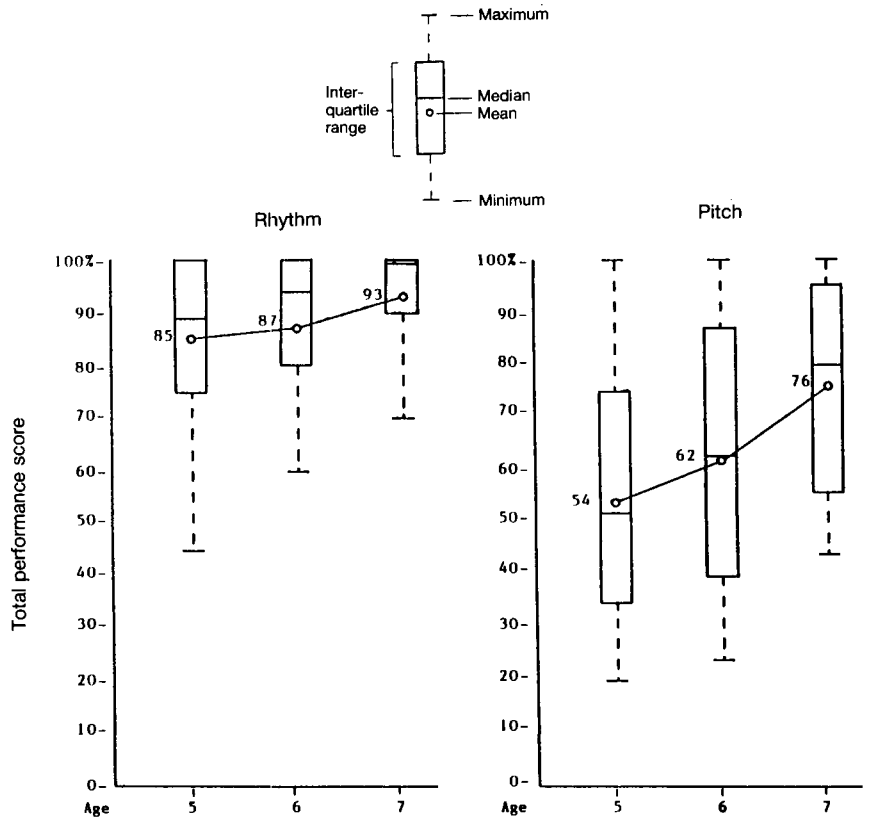



**Fig. 9.14.** Performance scores for rhythm and pitch in song by children ages 5–7.

develop, are increasingly able to draw on their performance knowledge in their representations. However, looking at pitch and rhythm separately, we see that pitch development in notation and performance increasingly correlates with age, whereas correlations are decreasing with respect to rhythm. This suggests that children's understanding of pitch is more closely linked to their developing representational knowledge of song than is their understanding of rhythm.

In sum, although there is substantial evidence for developmental trends in children's choice of symbol systems and in the level of sophistication of children's notation of songs, it appears that these are both increasingly independent 'tracks' in young children's representational development. However, as children's representations become more sophisticated, their level of representation increasingly corresponds to their performance knowledge of the song. This suggests, by the age of seven, the emergence

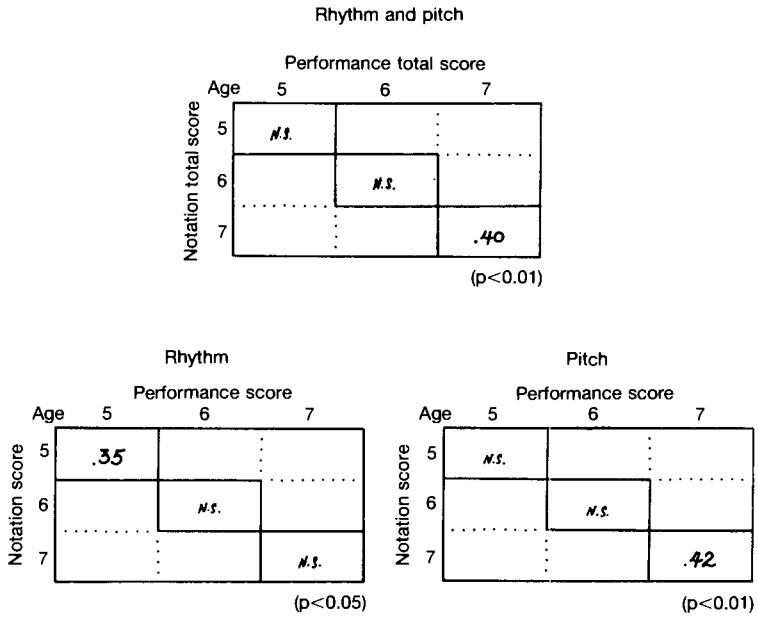



**Fig 9.15.** Matrices showing rank correlation coefficients (Spearman) between performance and notation scores in rhythm and pitch in children's song.

of a 'cognitively unified' domain of musical intelligence (Gardner 1984), where children's knowledge of musical performance and representation become increasingly yoked together. In both representational and performance knowledge of song, the dimension of pitch appears as the most stabilizing and powerfully predictive measure of children's musical development.

In the next section we look at children's notation of a single phrase across various musical contexts. Looking closely at their use of graphic symbols, we see young children's increasing sensitivity to the context of a phrase. Analysing the level of sophistication in their representations of a single phrase, we witness the increasingly powerful and consistent coding of musical features regardless of musical context. We also find more evidence establishing pitch as the primary factor in young children's developing musical cognition.

## Making the difference: a closer look at young children's representations across musical contexts

Although we have seen developmental stages in children's notations of

song, a closer view of their development is needed in respect to a wider range of musical contexts. Children's music representations may profoundly differ according to *performance attributes* (is the phrase sung or clapped?), *structural presentation* (is the phrase an isolated pattern or embedded within a song?), or *familiarity* (is the phrase familiar or newly learned?).

In our sessions we asked children to perform and notate a single five-note phrase in five distinctly different contexts : (A) as an *isolated rhythmic fragment*, (B) as an *isolated melodic fragment*, (C) as an *added introductory phrase to a familiar song*, (D) as a *concluding phrase of a familiar song*, and (E) as a *concluding phrase of an unfamiliar song*. As we can see from Fig. 9.16, these five contexts contain virtually identical pitch and rhythmic musical motives.

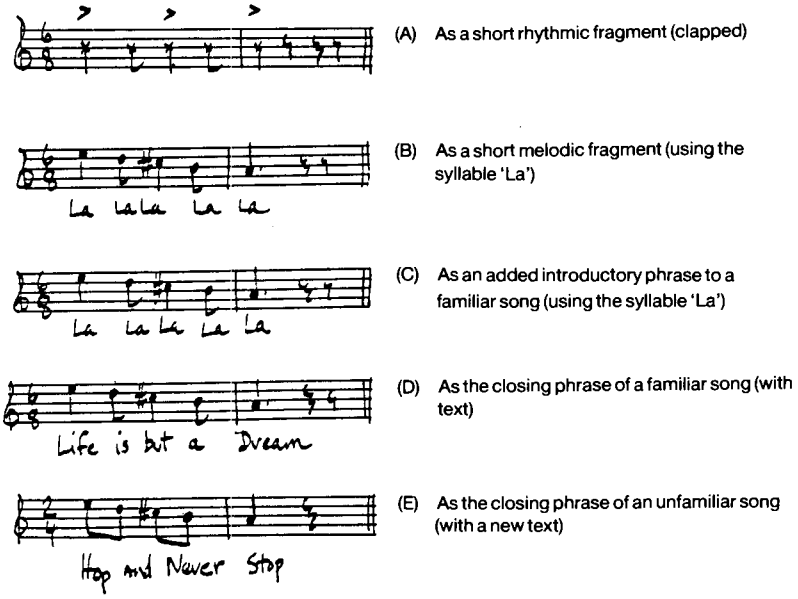



**Fig. 9.16.** The five musical contexts of a single musical phrase.

Using this single phrase as a control, we explore the ways in which young children are responsive to musical context. In particular, we contrast developmental trends in notation within the context of isolated rhythmic or pitch fragments with trends occurring within the structure of a whole song. Also, we want to know the effect of text on children's musical representations. Comparisons between no text (clapping), a generic text (the syllable 'la'), the text in a familiar song, and the text of a recently

learned song allow us to see how text may confound young children's attention to the musical features of a song. Finally, we can see which musical features children are apt to represent *regardless* of musical context. This last question may help us understand which dimensions of music are most important to (or less easily ignored by) young children.

## Choosing graphic symbols for a single phrase: evidence for the increasingly 'context sensitive' musical child

Analysing notations of a single phrase, we again see the broad developmental trends discussed in the analysis of children's notations of a familiar song. These trends include the dramatic decrease of icons and pictures, the increase of words and mixed symbols with age, and the relatively consistent use of abstract lines and dots (Fig. 9.17). More importantly, we see evidence for increasing sensitivity to musical context.

Children use *icons* less often with age, perhaps because they consider musical context when they choose these symbols. From Fig. 9.17 below we see that chldren ages five and six are just as likely to use icons to represent the single phrase in the context of isolated rhythm or pitch pattern as they are within a familiar or newly learned song. In other words, young children consider these symbols unconstrained by musical context and increasingly inappropriate for music representation.

This finding substantially contrasts with the children's increasingly 'context-sensitive' use of *abstract lines and dots* and *letters and words*. At first, children use these symbols equally often across the various musical contexts, but, in the succeeding years, there is a great deal of fluctuation of use, depending on the musical context. In particular, children use lines and dots more often for short rhythm patterns (with no pitch or text), less often for isolated or introductory melodic phrases (with generic text 'la la la'), and least often for verses (using the lyrics) of a familiar or unfamiliar song. This effect is increasingly pronounced with age. Children also use letters and words with increased context sensitivity but in opposite contexts: these symbols appear more often in phrases of the song, less often in isolated pitch patterns, and rarely in rhythm patterns.

Although language symbols appear as an obvious choice for representing a phrase within the context of a familiar song, we also see the sudden emergence of *mixed symbols* (usually combining words with abstract symbols or signs) with seven-year-olds, particularly in the context of isolated melodic patterns. Observing the relatively profuse use of mixed symbols in pitch patterns compared with relatively scant use with the identical phrase used as an introduction to the song, we see evidence for young children's increasing sensitivity to the structural aspects of song.

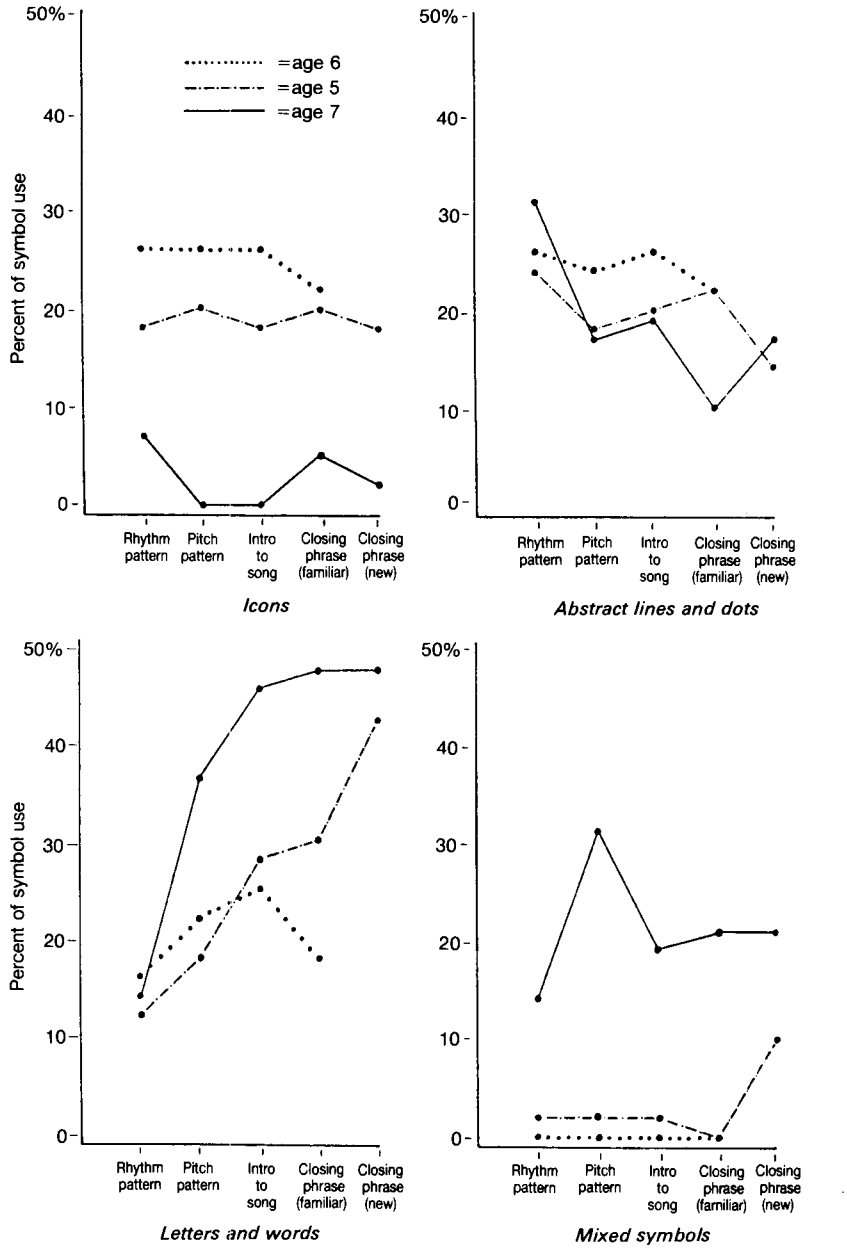



**Fig. 9.17.** Use of the four major graphic symbol types in a single phrase across five musical contexts by children ages 5–7.

Summarizing children's use of graphic symbols *by age*, we see an increasingly wider range of sensitivity to musical context. At the age of five, children's range of symbol use is surprisingly broad, yet relatively context-independent. At this age children use a large repertoire of symbol types of grasp the musical phrase. By the age of seven, however, a remarkably distinct profile of children's sensitivity to musical context emerges as children consistently choose language or abstract symbols (and their combinations) according to the context of the phrase. Looking across *context* at this age, we see children predominantly using abstract symbols for rhythm patterns, abstract and language symbols for pitch patterns, and mixed symbols for phrases within new or familiar songs.

As we trace Mary's notations of the single phrase through three years we encounter a clear example of this developmental sequence. In her notations we see a consistent use of abstract symbols for the rhythm patterns at all ages, while language and mixed symbols increasingly appear in the remaining contexts (Fig. 9.18).

Controlling for the structural context of the phrase, we see the symbols chosen for phrases as isolated patterns are markedly different from those chosen for phrases functioning within the song. We see a contrast in graphic symbols chosen for rhythmic fragments as opposed to pitch patterns. Continuing this analysis of a single phrase across the five contexts, we can take a closer look at children's ability to represent pitch and rhythmic content.

## Evidence for the increasingly divergent paths of rhythm and pitch in representational development

In our previous analysis of young children's representational development we contrasted their growth in notating the rhythmic and pitch dimensions of a familiar song. On the average, we saw the power of children's notations in both rhythm and pitch significantly expand from musical units (at age five) towards the organization of these units into rhythmic pulse or grouping and melodic contour (at age seven). We noticed also that growth in music representation and performance was particularly powerful in pitch. In our present analysis of the single phrase across five musical contexts we will see a telling clarification of this finding.

Comparing children's notations of musical dimensions across contexts we find that developmental trends in rhythm and pitch very different (Fig. 9.19). Whereas rhythmic levels are relatively unstable across contexts, pitch levels are much more stable. And, while representational development in rhythm occurs in spurts across years, growth patterns in pitch are unambiguous and increasingly substantial. Additionally, these tendencies are increasingly pronounced with age.

Although children at age five can organize the rhythmic units in the context

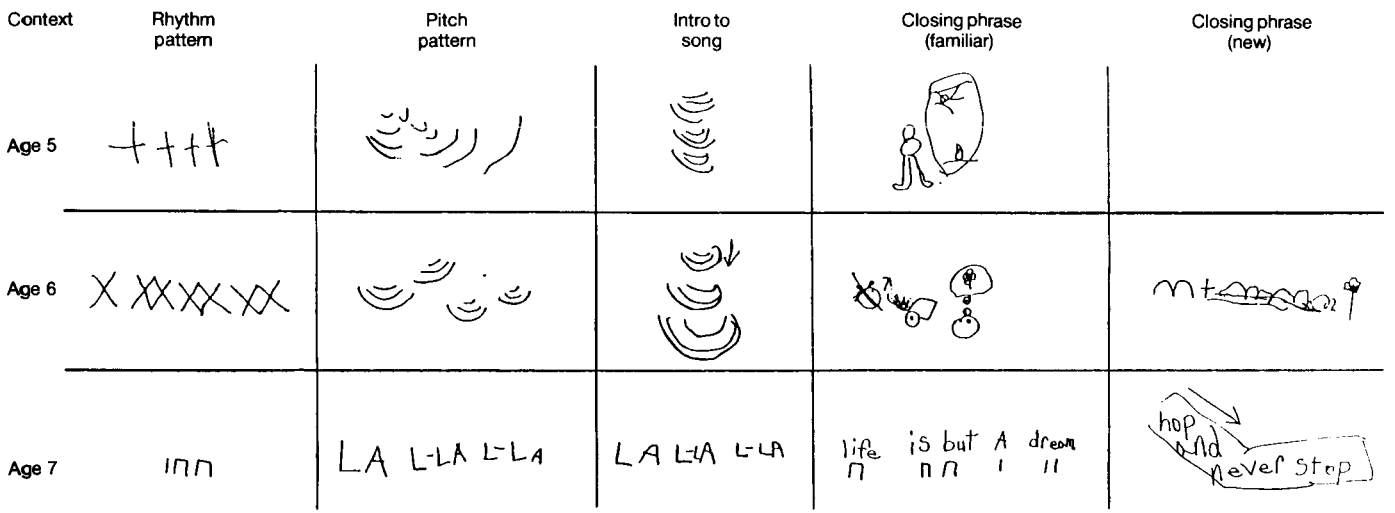



**Fig. 9.18.** Mary's notation of a single musical phrase across five contexts at ages 5–7.

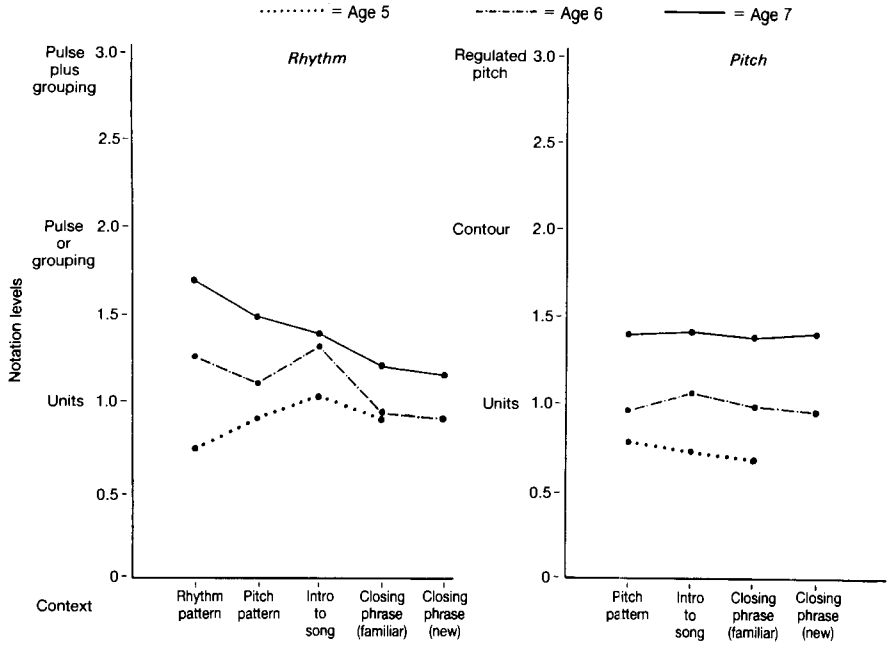



**Fig. 9.19.** Mean notation scores for rhythm and pitch in a single phrase across five musical contexts by children ages 5–7.

of a song, later they reveal their most sophisticated understanding of rhythmic features within the context of isolated rhythm patterns. Accordingly, statistical analysis of rhythmic notations using one-way repeated measures ANOVA reveals highly significant development [$p<0.0001$] over the three-year period *only* for isolated rhythm patterns. In this context it is not unusual to see seven-year-olds representing rhythmic grouping and the underlying pulse *simultaneously*. However, as soon as the dimension of pitch is introduced to the phrase, the notation of rhythmic pulse or grouping appears less often. This decline continues as the rhythmic features of the phrase competes with those features of pitch, text, and song structure. Although seven-year-olds continue to develop higher levels of rhythmic complexity in their musical representations, these only occur in the context of isolated rhythm patterns.

On the other hand, children's developing representations of pitch appear regardless of musical context. Whether the phrase occurs as an isolated melodic pattern, as an added introduction phrase with generic syllables, or as a phrase in a familiar or unfamiliar song, children attend to pitch equally well at every level of development. When compared to rhythm, pitch

development appears universally stable and increasingly robust with age. Statistical analysis in pitch notations using one-way repeated measure ANOVA shows highly significant [$p<0.002$] growth over the entire three-year period in all of the contexts.

This stable representation of pitch is most impressive at the age of seven. The average child is no longer content to code units of the phrase without providing some indication of melodic contour. Many children are now concerned with higher levels of organizing and formatting pitch as they invent ways of regulating the placement of notes so that intervallic or even scalar relationships can be represented. Furthermore, representing this level of pitch is important to the child in every musical context.

As we trace Janet's notations of the single phrase at age five and six, we see a clear example of the preference children generally give to pitch (Fig. 9.20). At age five, Janet represents the units of the phrase in the context of isolated patterns while tracking the underlying pulse (context (C)) or rhythmic grouping (context (D)) in song. At age six, she is primarily concerned with showing melodic contour except with isolated rhythm patterns. At age seven, she again consistently represents the melodic contour of the phrase, while occasionally adding indications of rhythmic grouping.

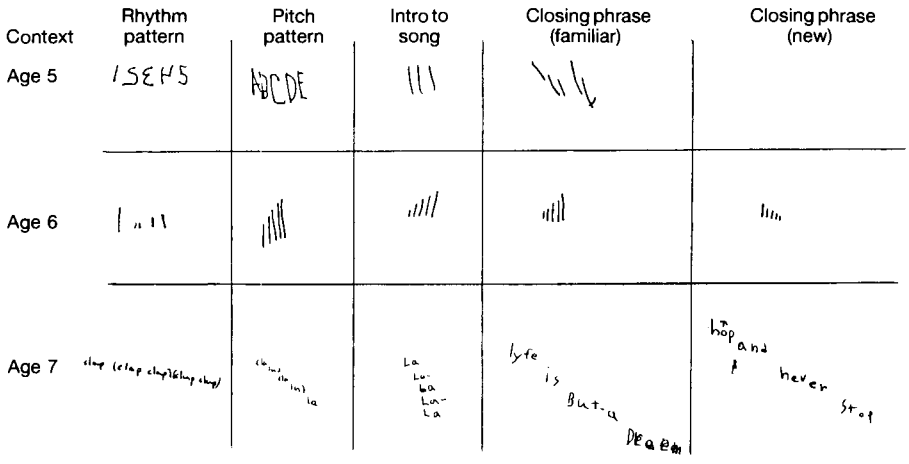



**Fig. 9.20.** Janet's notation of a single phrase across five musical contexts at ages 5–7.

Correlational analysis provides another way of viewing the diverging patterns of association between levels of rhythm and pitch representation. As seven-year-olds develop more complex representations of the phrase,

notations of pitch in all contexts are highly correlated, while rhythm notation levels are not. This strongly suggests that once children are able to record melodic contour in one context they are likely to retain this level of pitch representation throughout various contexts.

Remarkably, only children's notations of *isolated* rhythm patterns are highly correlated with the level of sophistication of pitch. Those who represent more sophisticated notational levels of isolated rhythm patterns tend to be the same ones who do well at representing pitch in their notations of short melodic fragments or phrases of songs. These findings reveal a distinctive profile of their musical understanding. Although clearly these children *can* represent the complex patterning of rhythmic relationships, when pitch is present they more often *choose* to notate the melodic contour. This switch of attention places rhythm and pitch in a figure-ground relationship, that is, the rhythmic 'figure' in isolation becomes 'ground' when pitch is introduced into the context of the phrase.

Looking briefly at the relationship between performance skills and ability to represent rhythm and pitch across contexts, we see additional evidence for the developing link between children's performance abilities and their representations of music. As children grow older and develop more sophisticated representations of pitch, they appear increasingly able to draw upon their performance knowledge. Surprisingly, this is not the case with their representation of the rhythmic content of the phrase. At first, correlations between performance scores and notation scores are practically nonexistent. By the age of seven, however, children's performance scores almost all significantly correlate with notation levels of pitch across virtually every context. Again, this is not the case with rhythm.

In sum, this 'across context' analysis of children's ability to attend to increasingly sophisticated levels of rhythm and pitch in their notations provides a rare 'composite' image of their implicit understanding of music. By looking at their notations of a single phrase across a wide range of musical contexts, we see the representation of pitch as their primary focus of attention. Although children can produce remarkably sophisticated representations of isolated rhythmic patterns during this time, we witness the dramatic emergence of pitch as the musical ingredient most consistently important to young children.

Again we witness the divergent paths of rhythm and pitch in young children's representational development. Although rhythm appears as the initial focus of focus of children's notations, pitch emerges as the increasingly cohesive underpinning of representational development. Not only is development more striking within pitch, but representational development with pitch can now be seen as extremely stable within a wide range of musical contexts and graphic symbols. We also find that representation of pitch—particularly in the context of songs—is increasingly tied to the performance knowledge of young children.

## Conclusion

Originally, in this chapter, we were concerned with the relationship of representational development to perceptual and performance development in the domain of music. For the most part, earlier studies in musical development concentrated on perception and performance. Although this research shows that young children progress considerably in their ability to recognize or discriminate melodic patterns, or in their ability to produce or imitate short melodic fragments, it may be that children's musical cognitive development is best viewed by their symbolic representations of music. Simply put, although dogs can recognize melodic calls and birds can create and reproduce tuneful ditties, human musical development is uniquely tied to symbolic 're-presentations' of the music we hear and perform. Therefore, through young children's first musical representations we can begin to view the uniquely human inception of musical cognitive development.

The age range of five to seven has been documented by many development psychologists (White 1970) as a period of considerable change in most children's learning and perception modes. Past research in musical perception and performance shows that music cognition is no exception. In our study of children's representation of musical features across a full range of contexts, we have documented dramatic developmental trends in their use of symbols, the level of sophistication of their notations, and their selective attention to musical features depending on the context. These trends appear particularly between the ages of six and seven. These findings root music cognition more firmly within the framework of general cognitive development than originally thought. Children not only reveal an increasingly systematic representation of musical features, they also increasingly draw on their performance knowledge when inventing their notations.

More importantly to psychologists of music, by studying the divergent paths of development in rhythm and in pitch we gain a more precise understanding of the course of musical development between the ages of five and seven. As these children's notations become more sophisticated, our results suggest the emergence of pitch as a prime ingredient of musical cognitive development. Pitch emerges as the most robust component of musical cognitive development, independent of musical context or symbol choice in representation, and increasingly associated with development of performance skills. We speculate from these findings that, as pitch development occurs with age, it becomes increasingly independent of the language, kinaesthetic, or number skills that confound our assessment of musical development in young children. For example, although performing the words of a well-known song may at first help control the rhythmic

surface of the performance, it does not predict the ability to sing these words with pitch. On the other hand, as children get older, the ability to reproduce 'pitch in song' increasingly predicts the ability to notate pitch in all contexts. Pitch, therefore, appears to be an essential 'core' ingredient of

In music education, we view these results as important to our understanding of the developing child. This study reveals that young children, without explicit musical training in notation, can represent an increasingly sophisticated understanding of the music they perform. While recognizing the *success* of performance training at an early age (for instance Suzuki training), we must also pay attention to the musical *understanding* of young children. For example, music teachers might be better off assessing children's sensitivity to musical features by looking at their invented representations of music rather than at their reproduction of music symbols. Also, music educators might restructure the focus and timing of musical training in accordance with these developmental findings. We would emphasize the assessment of pitch development as a critical focus in young children's musical training.

Finally we strongly urge the continued exploration of young children's understanding of music. The early development of music perception, performance, and symbolic representation can only help us better understand the early roots of human cognition. Contained in children's spontaneous notations of short nursery songs are glimpses of growing minds at work. As we better understand children's extraordinary invented symbol systems for music, we discover how they mindfully abstract and reconstruct an increasingly rich and integrated world of musical understanding.

## Acknowledgements

This research was supported in part by grants from the Carnegie Corporation, Rockefeller Foundation, and the Spencer Foundation. The work was carried out in the Cambridge, Mass, Public Schools and at the Cambridge Friends School. We are grateful to the teachers, children, and parents who helped us to pursue our interest in musical cognition. For their careful and critical readings of early drafts, we thank Howard Gardner and Jeanne Bamberger. John Willet's command of non-parametric statistical techniques aided us greatly in our data analysis. We are also indebted to Martha Davis-Perry, Joan Meyaard, Dennie Wolf, Joe Walters, and Rena Upitis for their suggestions and assistance.

## References

Bamberger, J. (1980). Cognitive structuring in the apprehension and description of simple rhythms. *Archives de Psychologie* **48** , 177–99.

Bamberger, J. (1982). Revisiting children's drawings of simple rhythms: a function for reflection in action. *U-Shaped behavioural growth* (ed. S. Strauss). Academic Press, New York.

Bamberger, J. and Schon, D. A. (1977). The figural-formal transaction: a parable of generative metaphor. Unpublished paper, Division for Study and Research in Education at the Massachusetts Institute of Technology.

Bruner, J. (1973). The growth of representational processes in childhood. *Beyond the information given*. Norton, New York.

Cody, R. and Smith, J. (1985). *Applied statistics and the SAS programming language*. North-Holland, New York.

Davidson, L. (1985). Tonal structures of children's early songs. *Music Perception* **2**, No. 3 361–74.

Dowling, W. J. (1982). Melodic information processing and its development. *The psychology of music* (ed. D. Deutsch). Academic Press, New York.

Flavell, J. H. (1963). *The developmental psychology of Jean Piaget*. Van Nostrand, New York.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association* **32**, 675–701.

Furth, H. (1969). *Piaget and knowledge*. University of Chicago Press, Chicago.

Gardner, H. (1979). Developmental psychology after Piaget: an approach in terms of symbolization. *Human development*. Academic Press, New York.

Gardner, H. (1980). *Artful scribbles: the significance of children's drawings*. Basic Books, New York.

Gardner, H. (1983). *Frames of mind: the theory of multiple intelligences*. Basic Books, New York.

Gardner, H., Howard, V., and Perkins, D. (1974). Symbol systems: a philosophical, psychological, and educational investigation. *Media and symbols: the forms of expression, communication, and education* (ed. David Olson). University of Chicago Press, Chicago.

Goodman, N. (1976). *Languages of art*. Hackett, Cambridge, Mass.

Goodman, N. (1978). *Ways of worldmaking*. Hackett, Cambridge, Mass.

Goodnow, J. (1977). *Children drawing*. Harvard University Press, Cambridge, Mass.

Gordon, E. (1965). *Musical aptitude profile*. Houghton Mifflin, Boston.

Hattwick, M., and Williams, H. (1935). Measurement of musical development II, *University of Iowa Studies in Child Welfare*, Vol. 11, No. 2, pp. 63–6. Iowa City: University of Iowa Press.

Hollander, M., and Wolfe, D. (1973). *Nonparametric statistical methods*. Wiley, New York.

McKernon, P. (1979). The development of first songs in young children. *New Directions for Child Development* **3**, 43–58.

Olson, D. (1970). *Cognitive development: the child's acquisition of diagonality*. Academic Press, New York.

Petzold, R. G. (1966). *Auditory perception of musical sounds by children in the first six grades*. Cooperative Research Project No. 1051, University of Wisconsin.

Pflederer, M. (1964). The responses of chlidren to musical tasks embodying

Piaget's principle of conservation. *Journal of Research in Music Education* **12**, (4) 251–68.

Piaget, J. and Inhelder, B. (1956). *The child's conception of space.* Norton, New York.

Piaget, J. and Inhelder, B. (1973). *Memory and intelligence.* Routledge and Kegan Paul, London.

Seashore, C. E. (1919). *Measures of musical talents.* The Psychological Corporation, New York.

Serafine, M. L. (1984). The development of cognition in music. *The Music Quarterly* **70**, 218–33.

Upitis, R. (1985). Children's understanding of rhythm: the relationship between development and musical training. Unpublished Ph.D. thesis. Harvard University. University Microfilms International, London.

Walker, R. (1981) Teaching basic musical concepts and their staff notations through cross-modal matching symbols. *Psychology of Music* **9**, (1), 31–8.

Werner, H. (1948) *The comparative psychology of mental development.* International Universities Press, New York.

Werner, H. (1957). The concept of development from a comparative and organismic point of view. In *The concept of development* (ed. D. Harris). University of Minnesota Press, Minneapolis.

White, S. H. (1970). Some general outlines of the matrix of developmental changes between five and seven years. *Bulletin of the Orton Society* **20**, 41–57.

Wolf, D., Davidson, L., Davis, M., Walters, J., Hodges, M., and Scripp, L., (1986). Beyond A, B and C: a broader and deeper view of literacy. In *The psychological bases of early education* (ed. A. Pellegrini). Wiley, London.

Zenatti, A. (1969). *Le développement génétique de la perception musicale.* Monographies Françaises de Psychologie, No. 17.

Zuckerkandl, V. (1969). *Sound and symbol: music and the external world.* Princeton University Press.

# 10

# Cognitive constraints on compositional systems

FRED LERDAHL

In this chapter I explore the relationship between composing and listening. I begin with a problematic story, draw some general conclusions, introduce relevant concepts from Lerdahl and Jackendoff (1983) and related work, propose some cognitive constraints on compositional systems, discuss 'pitch space', and explain why serial (or 12-tone) organizations are cognitively opaque. Most of these topics deserve fuller treatment than is given in these pages. My concern here is just to lay out a basic, if wide-ranging, argument.

I am not interested in passing judgement on the composers and compositions that are mentioned, particularly not on the remarkable work by Boulez that I use as a representative example. The thrust of my argument is psychological rather than aesthetic. But since aesthetic issues inevitably impinge on the discussion, I treat them briefly at the end.

## Hidden organization in *Le Marteau sans Maître*

Boulez's *Le Marteau sans Maître* (1954) was widely hailed as a masterpiece of post-war serialism. Yet nobody could figure out, much less hear, how the piece was serial. From hints in Boulez (1963), Koblyakov (1977) at last determined that it was indeed serial, though in an idiosyncratic way. In the interim listeners made what sense they could of the piece in ways unrelated to its construction. Nor has Koblyakov's decipherment subsequently changed how the piece is heard. Meanwhile most composers have discarded serialism, with the result that Koblyakov's contribution has caused barely a ripple of professional interest. The serial organization of *Le Marteau* would appear, 30 years later, to be irrelevant.

This story is, or should be, disturbing. There is a huge gap here between compositional system and cognized result. How can this be?

One might suppose that the impenetrability of *Le Marteau's* serial organization is due to insufficient exposure. After all, the piece was

innovative; listeners must become accustomed to novel stimuli. Such has been the traditional defence of new art in the face of incomprehension. One might refine this view by pointing out that there is little repetition in *Le Marteau*. The lack of redundancy perhaps overwhelms the listener's processing capacities. Comprehensibility, then, is arguably a consequence both of the degree of conditioning to the materials and of the number and variety of events per unit time. (Is it a coincidence that *Le Marteau* was composed in the heyday of behaviourism and information theory?)

But this explanation is inadequate. For one thing, competent listeners to *Le Marteau*, even after many hearings, still cannot even begin to hear its serial organization. For many passages they cannot even tell if wrong pitches or rhythms have been played. The piece is hard to learn by ear in a specific sense; its details have a somewhat statistical quality. Conditioning, in short, does not suffice. For another thing, *Le Marteau* does not feel structurally complex in the way, for example, that compositions by Beethoven or Schoenberg do. Vast numbers of nonredundant events fly by, but the effect is of a smooth sheen of pretty sounds. The listener's processing capacities, in short, are not overwhelmed.

This is not to deny the influence of exposure and redundancy on comprehensibility. But these factors do not go far enough; they do not address issues of specific organization. Cognitive psychology has shown in recent decades that humans structure stimuli in certain ways rather than others. Comprehension takes place when the perceiver is able to assign a precise mental representation to what is perceived. Not all stimuli, however, facilitate the formation of a mental representation. Comprehension requires a degree of ecological fit between the stimulus and the mental capabilities of the perceiver.

Experienced listeners do not find *Le Marteau* totally incomprehensible, but neither, I would argue, do they assign to it a detailed mental representation. This is why the details of the piece are hard to learn and why the piece does not in the end feel complex. The serial organization that Koblyakov found is opaque to such structuring.

Of course a musician of Boulez's calibre would not use a compositional system without drawing crucially upon his musical intuition and experience. Music-generating algorithms alone have always produced primitive outputs; not enough is known about musical composition and cognition for them to succeed. Boulez had the intellectually less ambitious goal of developing a system that could just produce a quantity of musical material having a certain consistency. He then shaped his materials more or less intuitively, using both his 'ear' and various unacknowledged constraints. In so doing, he listened much as another listener might. The organization deciphered by Koblyakov was just a means, and not the only one, towards an artistic end.

The degree to which *Le Marteau* is comprehensible, then, depends not

on its serial organization but on what the composer added to that organization. On the other hand, the serial procedures profoundly influenced the stimulus structure, leading to a situation in which the listener cannot form a detailed mental representation of the music. The result is a piece that sounds partly patterned and partly stochastic.*

## Compositional and listening grammars

It will be useful to put these observations in schematic form. Let me introduce the notion of a *musical grammar*, a limited set of rules that can generate indefinitely large sets of musical events and/or their structural descriptions. Such a grammar can take the form of (a) or (b) in Fig. 10.1. (The rectangles stand for sets of rules, the ellipses for inputs and outputs of rules.) In (a), the rule system generates the events of a passage or piece, and in the process provides some sort of specification (the 'input organization'). In (b), the sequence of events is taken as given, and the role of the rule system is to assign a structural description to the sequence.

Our discussion of *Le Marteau* suggests that there are two kinds of musical grammar at work here. The first is the *compositional grammar*, consciously employed by Boulez, that generated both the events of the piece and their serial organization as discovered by Koblyakov. This grammar is an instance of (a) in Fig. 10.1. The second kind is the *listening grammar*, more or less unconsciously employed by auditors, that generates mental representations of the music. This grammar is an instance of (b) in Fig. 10.1; the events themselves are present when *Le Marteau* is played.

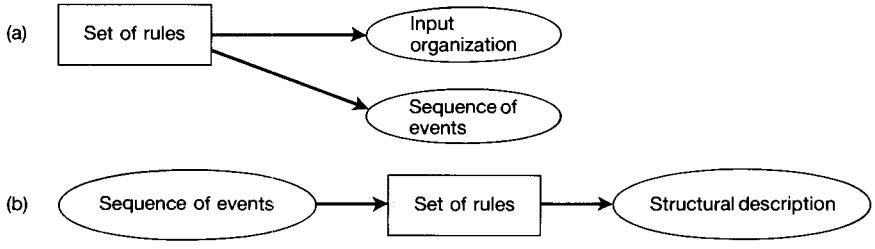



**Fig. 10.1**

* In *Structures*, Book I (1952), Boulez flirted with 'algorithmic composition', and in *Pli selon pli* (1958–61) he incorporated aleatoric elements; so *Le Marteau* lies *en route* in his dialectic of determinism and chance (see Boulez 1964). His interest in algorithmic composition continues to the extent that for a while there was a research group at IRCAM pursuing it. For six months in 1981 I was part of that group, though in my project the purpose of the computer program was not to compose but to assist composers in composing, much as a word processor assists writers (see a brief description in Lerdahl 1985). Many of the ideas expressed in this essay grew out of that experience, for which I remain deeply grateful.

Fig. 10.2 summarizes these remarks. The diagram shows that the compositional grammar generates the sequence of events and the manner in which they are specified. Only the sequence of events, however, is available as input to the listening grammar: the listener hears the acoustic signal, not its compositional specification. The listening grammar then generates the mental representation that comprises the 'heard structure' of the piece. Note that Fig. 10.2 incorporates both (a) and (b) of Fig. 10.1.

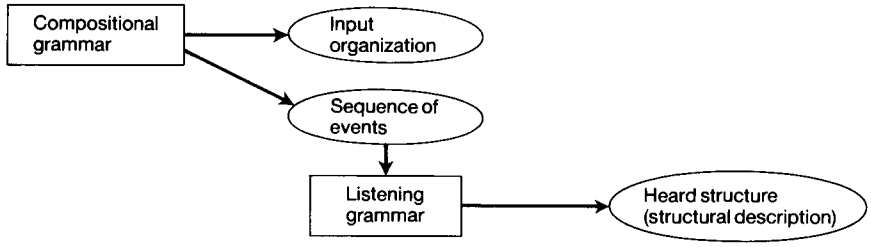



**Fig. 10.2**

This account is complicated by the fact that, as noted above, Boulez created *Le Marteau* not only through serial procedures but through his own inner listening. In the process he followed constraints that, while operating on the sequence of events produced by the compositional grammar, utilized principles from the listening grammar. This more elaborate picture is drawn in Fig. 10.3.

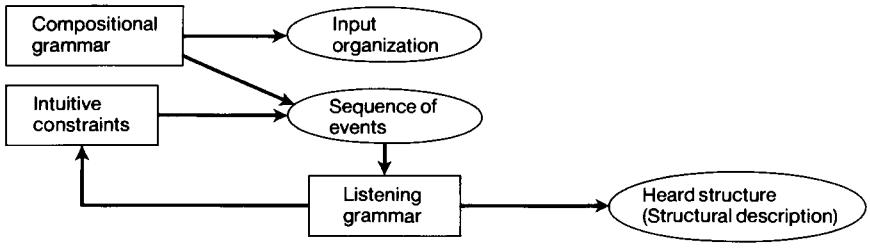



**Fig. 10.3**

Observe that none of the arrows in the diagram are inputs to the 'compositional grammar'. This component is inaccessible to the rest of the system. Hence it becomes quite possible for the 'compositional grammar' to be unrelated to the other rules, the 'listening grammar' and 'intuitive constraints'. If this happens, the 'input organization' will bear no relation to the 'heard structure'. Here, then, lies the gap between compositional system and cognized result with which we began.

This situation exists not only for *Le Marteau* but for much of

contemporary music. I could have illustrated just as well with works by Babbitt, Carter, Nono, Stockhausen, or Xenakis. This gap is a fundamental problem of contemporary music. It divorces method from intuition. Composers are faced with the unpleasant alternative of working with private codes or with no compositional grammar at all. Private codes remain idiosyncratic, competing against other private codes and creating no larger continuity—so that, for example, 30 years later the serial organization of *Le Marteau* becomes irrelevant even to other composers.

## Natural and artificial compositional grammars

Where does a compositional grammar come from? The answer varies, but a few generalizations may be helpful. Let us distinguish between a 'natural' and an 'artificial' compositional grammar. A natural grammar arises spontaneously in a musical culture. An artificial grammar is the conscious invention of an individual or group within a culture. The two mix fruitfully in a complex and long-lived musical culture such as that of Western tonality. A natural grammar will dominate in a culture emphasizing improvisation and encouraging active participation of the community in all the varieties of musical behaviour. An artificial grammar will tend to dominate in a culture that utilizes musical notation, that is self-conscious, and that separates musical activity into composer, performer, and listener.

The gap between compositional and listening grammars arises only when the compositional grammar is 'artificial', when there is a split between production and consumption. Such a gap, incidentally, cannot arise so easily in human language. People must communicate; a member of a culture must master a linguistic grammar common to both speaking and hearing. But music has primarily an aesthetic function and need not communicate its specified structure. Hidden musical organizations can and do appear.

A natural compositional grammar depends on the listening grammar as a source. Otherwise the various musical functions could not evolve in such a spontaneous and unified fashion. An artificial compositional grammar, on the other hand, can have a variety of sources—metaphysical, numerical, historical, or whatever. It can be desirable for an artificial grammar to grow out of a natural grammar; think, for example, of the salutary role that Fux (1725) played in the history of tonality. The trouble starts only when the artificial grammar loses touch with the listening grammar.

In the Western tradition the trouble began with the exhaustion of tonality at the turn of the century. Anything became possible. Faced with chaos, composers reacted by inventing their own compositional grammars. Within an avant-garde aesthetic it became possible to believe that one's own new system was the wave of the future. Boulez's generation was the

last to believe this. To a younger generation these systems have come to seem merely arbitrary. The avant-garde has withered away, and all methods and styles are available to the point of confusion.

One can react to this situation by giving up on compositional grammars and relying solely on ear and habit—on the 'intuitive constraints' of Fig. 10.3. But composing is too difficult for such a solution; there are too many possibilities. Or one can react by reverting to earlier 'natural' styles, a move that condemns one to a parasitic relationship with the past. Both reactions are common these days, and both are motivated by the desire to avoid the gap between composing procedure and what is heard.

My own reaction as a composer has been less to avoid than to confront this gap. These were my initial ground rules: (1) a compositional grammar is necessary; (2) it need not be nostalgic; (3) our musical culture is too fragmented and self-conscious for a natural grammar to emerge; but (4) an artificial grammar unresponsive to musical listening is unacceptable. It was inevitable, I concluded, that early attempts at artificial grammars—say, from the 1920s to the 1950s—were defective in their relation to listening. Not enough was known about musical cognition; the basic questions had not even been framed. Beginning around 1970, however, a new perspective became possible through the simultaneous decline of the avant-garde and rise of cognitive psychology. Contemporary music had lost its way. What other foundation was there to turn to than the nature of musical understanding itself? In other words, I decided that a compositional grammar must be based on the listening grammar. Fig. 10.4 illustrates this by adding a dashed arrow to the flow chart of Fig. 10.3. In principle this arrow integrates the compositional grammar into the overall framework, rather than leaving it isolated from input.

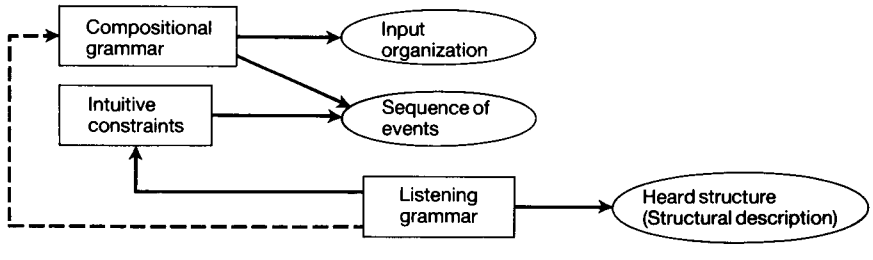



**Fig. 10.4**

But for this proposal to have substance, a great deal must be known about the listening grammar. Hence it became necessary to develop a detailed theory of musical cognition (Lerdahl and Jackendoff 1983). Such a theory, I reasoned, could provide the basis for artificial compositional grammars that could be intellectually complex yet spontaneously accessible

to mental representation. The commonality of compositional and listening grammars could produce a rich yet transparent music.

## Theoretical overview

This is not the place to describe the specifics of Jackendoff's and my theory. I propose only to outline some of its general features. The reader should bear in mind, however, that the theory is concrete and detailed.

Our theory models musical listening within the framework of some standard methodological idealizations. Assuming (1) an experienced listener, (2) the psychoacoustical organization of the physical signal into a 'musical surface' comprising a sequence of discrete events, and (3) a final-state knowledge of the sequence, the theory predicts structural descriptions from musical surfaces by means of a set of rules that ideally corresponds to the 'listening grammar' of Figs. 10.2–4. See Fig. 10.5; and note that the theory is 'generative' in the sense of (b) in Fig. 10.1. The predictions of the rules are testable by introspection and, in principle, by experiment.*

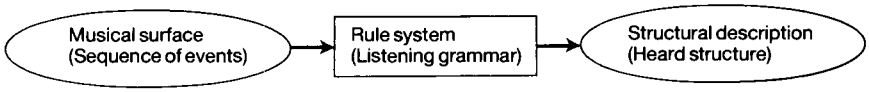



**Fig. 10.5**

As a practical matter the theory focuses on Classical tonal music, but most of the rules appear to have a more general psychological basis and therefore to be applicable to other idioms as well. The extent of the universality of the rules is an empirical question. Again for practical reasons, the rules assign structural descriptions only to the hierarchical aspects of musical structure, neglecting 'associational' dimensions such as motivic processes and timbral relations (but see Lerdahl 1987). By 'hierarchy' is meant the strict nesting of elements or regions in relation to other elements or regions. The theory claims that, if the signal permits, the listener unconsciously infers four types of hierarchical structure from a musical surface: *grouping structure*, or the segmentation of the musical flow into units such as motives, phrases, and sections; *metrical structure*, or the pattern of periodically recurring strong and weak beats associated with the surface; *time-span reduction*, or the relative structural importance of

* Deliège (1985) has experimentally verified—and, in a few cases, improved upon—the local grouping preference rules of the theory. Todd (1985) has used the time-span reductional component of the theory to model expressive timing at cadences in performance. Palmer and Krumhansl (in press) have found empirical support for predictions made by the metrical and time-span components.

events as heard within contextually established rhythmic units; and *prolongational reduction*, or the perceived pattern of tension and relaxation among events at various levels of structure. Both kinds of reduction are described by structural trees. The prolongational component incorporates aspects of Schenker's (1935) theory.

For the most part these four hierarchies interact as shown in Fig. 10.6. From the grouping and metrical structures the listener forms the rhythmic units, or *time-span segmentation*, over which the dominating-subordinating relationships of time-span reduction take place; and from the time-span reduction the listener in turn projects the tensing-relaxing hierarchy of prolongational reduction. Thus the mapping from musical surface to prolongational structure is indirect. In addition, both kinds of reduction depend for their operation on *stability conditions* among pitch configurations as considered 'out of time'. These conditions, which also can be described hierarchically (Krumhansl 1983; Bharucha 1984*b*), are internal schemata induced from previously heard musical surfaces and brought to bear on the in-time event sequences described by the reductions.

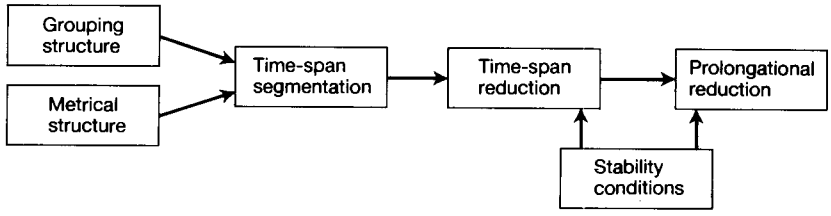



**Fig. 10.6**

For each of the four organizations there is a set of *well-formedness rules* that defines the conditions for hierarchy. Well-formed structures can be modified in limited ways by *transformational rules*. These two rules types are abstract in that they only define formal possibilities. By contrast, a third rule type, *preference rules*, registers particular aspects of presented musical surfaces and selects which well-formed or transformed structures in fact apply to those surfaces. The preference rules, which do the bulk of the work of the theory, do not generate categorically right or wrong analyses but instead predict descriptions as more or less coherent. When the preference rules reinforce one another, the analysis is stable and the passage in question is judged as stereotypical: when they conflict, the analysis is unstable and the passage is judged as ambiguous or vague.

In summary, the theory asserts that (1) the listener has at his disposal a number of hierarchical organizations possessing certain formal properties, and (2) the listener unconsciously attempts to assign the most stable overall description (or descriptions) by means of preferential principles that activate and interact in response to specific features of musical surfaces.

## Constraints on event sequences

What does Jackendoff's and my theory have to say about closing the gap between compositional system and cognized result? Let me explore this question by proposing, in this and the following two sections, a number of psychologically plausible constraints on compositional grammars. These constraints will give substance to the dashed arrow in Fig. 10.4. Then I will use the constraints to explain why serial organizations are not easily learnable.

We begin with a presupposition of the theory, namely that the musical surface breaks down into individual events. Certain recent musical developments (pioneered for instance by Ligeti) have tended to blur distinctions between events. Sensuously attractive though this blurring may be, it inhibits the inference of structure.

> *Constraint 1*: The musical surface must be capable of being parsed into a sequence of discrete events.*

It is hardly fortuitous that our theory concentrates on hierarchies. Most of human cognition relies on hierarchical structuring (Miller *et al.* 1960; Simon 1962; Neisser 1967). Studies in music psychology have indicated that the absence of perceived hierarchy substantially reduces the listener's ability to learn and remember structure from musical surfaces (Deutsch 1982). It does not suffice for the input organization to be structured hierarchically; such in fact is the case for *Le Marteau* (Koblyakov 1977). It is the relationship to the listening grammar that matters.

> *Constraint 2*: The musical surface must be available for hierarchical structuring by the listening grammar.

As suggested above, there are four ways that the listening grammar can structure event sequences hierarchically. Let us return to Fig. 10.6 and consider what is needed at the musical surface for these organizations to come into play.

Segmentation into groups is accomplished at local levels largely by detection of *distinctive transitions* in the musical flow. A distinctive transition corresponds at a less fine-grained level to the principles by which events themselves are perceived as discrete rather than continuous. It is a change in some musical dimension, such as greater distance in attack points or shift in dynamics, timbre, or register, with respect to an immediate context that is relatively invariant. From this it follows that constant change will not give rise to salient distinctive transitions; nor of course will

* This formulation skirts the issue of simultaneous sequences of discrete events, an area that has been researched as 'auditory stream segregation' by Bregman and his associates (see McAdams and Bregman 1979 for a review).

no change at all. Babbitt and Reich have something in common after all. The steady-state quality of both their musics is due to a paucity of distinctive transitions, hence of grouping boundaries.

> *Constraint 3*: The establishment of local grouping boundaries requires the presence of salient distinctive transitions at the musical surface.

The other factor creating local groups is repetition or, more generally, *parallelism* of musical units. Repetition is an obsession of minimalism. But without distinctive transitions such groups produce only 'flat' hierarchies; that is, the repetitions are not heard within larger groups to any depth of embedding. Groups of groups tend to arise rather through the reinforcing action of distinctive transitions and parallelisms, often supplemented by a third principle of *symmetry* (the approximately equal division of a larger time span). At global levels parallelism becomes the overriding grouping principle: listeners try to hear parallel passages in parallel places in the overall structure.

> *Constraint 4*: Projection of groups, especially at larger levels, depends on symmetry and on the establishment of musical parallelisms.

Much contemporary music avoids symmetry and parallelism. Indeed, in the 1950s and 1960s literal repetition was widely regarded as aesthetically inexcusable. But symmetry and especially parallelism are basic ingredients of any complex grouping structure. The reliance on these principles in Classical tonal music and in various ethnic musics is not necessarily a sign of lack of invention but is an essential technique for the creation of deeply embedded groups.

We turn now to metrical structure. A well-formed metre is a looping pattern of equidistant beats occurring at multiple hierarchical levels. A beat at any metrical level is felt to be strong if it is also a beat at the next larger level. Depending on the musical idiom, the criterion of equidistance may be loosened, but not to the extent of abandoning all sense of periodicity. A musical idiom typically has available a limited repertory of well-formed metrical possibilities.

When hearing a musical surface, the listener seeks an optimal fit between the accents perceived directly from the signal and the repertory of available metrical structures. These *phenomenal accents* arise from local stresses such as sforzandos, changing bass notes and chords, contextually long events, and so forth. Insofar as possible, the onsets of phenomenal accents are aligned with relatively strong beats. In other words, the listener looks for maximal periodicity in the musical signal.

Great regularity in phenomenal accents produces a rhythmic quality of stability and squareness. In much rhythmically vital music, however, many

phenomenal accents are forced to align with weak beats, creating cross-accents or syncopations. But if there is very little regularity to the phenomenal accents, the listener may not be able to infer any metrical structure. Most music is metrical, and the immediate location of events is established largely in relation to this or that strong or weak beat. An inability to assign a metrical grid weakens the precision of location of events, resulting in a quality of suspended rhythm.

*Constraint 5*: The establishment of a metrical structure requires a degree of regularity in the placement of phenomenal accents.

Much contemporary music, even if notated traditionally, avoids regularity of phenomenal accents and does not give rise to a sense of metre. Complicated rhythms can cancel as well as create structure. An interesting case is Carter's practice of establishing two or more simultaneous tempos. Each such tempo corresponds to a single metrical level, or pulsation, but usually does not evoke a hierarchy of strong and weak beats. Carter contracts metrical depth in order to focus on multiple speeds. However, two or three simultaneous tempos cannot be taken in as readily as the four or five metrical levels that are commonplace in a march or waltz. Musical cognition has a bias for hierarchical organization. Simultaneous tempos instead produce independent organizations competing for attention. The difficulties in attending to more than one such organization are well known (Cherry 1953) and are ameliorated only if the vertical correspondences are coherent (Sloboda 1985), a condition that only partly obtains in this case. The situation is analogous to that of harmonically controlled versus uncontrolled polyphony.

The grouping and metrical structures come together (via well-formedness rules) to form the time-span segmentation. Details aside, it should be intuitively obvious that events are heard and thought of within such units. We normally speak of an event at small levels in terms of this or that beat, at larger levels in terms of this or that phrase or section. If the grouping or metrical structure is impoverished, then so too is the time-span segmentation.

*Constraint 6*: A complex time-span segmentation depends on the projection of complex grouping and metrical structures.

Various passages of *Le Marteau*, in which no metrical structure is apparent and the groups are not deeply embedded, provide a contrary illustration. Because of the consequent lack of hierarchical time-span segmentation, the sense of these passages is of a flurry of events followed by a pause, then another flurry and another pause, like beads on a string. Each flurry tends to become one complex 'event' rather than an organized sequence of events. The phrase is replaced by a sonorous object extended in time. (This is one way of explaining why, despite its generally fast tempos, *Le Marteau* feels slow.)

The notion of an extended sonorous object filled with inner movement is a compelling one, especially in light of recent developments in computer music. Here lies an unexpected link beteween *Le Marteau* and Boulez's recent work, *Répons*. But from the perspective of musical cognition, one must ask how the inner movement of an object receives structure. We have returned to Constraint 1, the issue of musical surfaces as discrete events. If the listener can assign any rich internal structure to the sonorous object, it must be perceived as consisting of 'subobjects'. But if the subobjects are to be experienced as aspects of one object, they cannot be too discriminable or allow salient structuring.

The time-span segmentation must be highly structured for a complex event hierarchy to emerge—something not achieved, and perhaps not tried for, in *Le Marteau, Répons*, or indeed most contemporary music. Since 'events' in music are usually pitch events, event hierarchies are normally pitch hierarchies. A unified pitch hierarchy is an instance of tonality (broadly defined), the stability conditions for which will be discussed in the next two sections. For now let us assume stability conditions and review the projection of event hierarchies from musical surfaces.

The two kinds of reduction are complementary descriptions of event hierarchies. In the time-span reduction, the listener compares the relative stability of events within each time-span segment. The less stable event in a segment is judged as subordinate to, or as an elaboration of, the more stable event. This process continues recursively from local to global levels of the time-span segmentation until an entire time-span tree is built up. The relative stability of events within a span is determined partly by rhythmic position, for example whether an event occurs on a strong beat or cadences a group. But the main determinant of domination or subordination within a span is the set of stability conditions.

> *Constraint 7*: The projection of a time-span tree depends on a complex time-span segmentation in conjunction with a set of stablility conditions.

A little thought-experiment will demonstrate the necessity for both the time-span segmentation and the stability conditions for the hearing of an event hierarchy. Imagine the pitches of *Le Marteau* plugged into the rhythms of Beethoven's Fifth Symphony. The hierarchical time-span segmentation would largely be erased by the lack of any concomitant pitch articulation, and, in any case, the absence of stability conditions would prevent the inference of any pitch hierarchy. Or, conversely, imagine the pitches of the Fifth Symphony plugged into the rhythms of *Le Marteau*. The potentially hierarchical pitch structure would be garbled by the lack of supportive time-span segmentation (including the absence of any metrical structure or patterned grouping structure). In short, rhythm and pitch must work in concert if there is to be any time-span reduction.

The prolongational reduction derives from the time-span tree and, again, from the stability conditions. Although this is the most important of the ways in which events relate to one another, its description is rather technical and abstract, and so will be avoided here. Suffice it to say that in this component events connect hierarchically in terms of progression and continuity, creating *prolongational regions* that act in structural counterpoint to the time-span segmentation (see Lerdahl and Jackendoff 1983 for a detailed treatment).

> *Constraint 8*: The projection of a prolongational tree depends on a corresponding time-span tree in conjunction with a set of stability conditions.

Before continuing, it would be well to emphasize how interdependent the contraints are that have been listed so far. Most of them call on their immediate predecessors. Thus a hierarchy of events (Constraint 2) requires a musical surface that has been parsed into events (Constraint 1); the recursive segmentation of the surface into rhythmic units (Constraint 6) cannot take place without the establishment of grouping and metrical structures (Constraints 3–5); the relative structural importance of events (Constraint 7) cannot be assessed without the time-span segmentation (Constraint 6); and the representation of how events depart and return (Constraint 8) demands an evaluation of their function in the rhythmic structure (Constraint 7). The requirement of projecting event hierarchies to the listener has brought a great deal in its wake.

## Constraints on underlying materials

We turn now to the stability conditions underlying the two reductions. Generally, a stability condition says, 'Musical context aside, this structure is judged as more stable than that.' Stability conditions interact like preference rules, sometimes reinforcing and sometimes conflicting with one another. In Classical tonality, for example, a triad in root position is more stable than a triad in first inversion, and a I chord is more stable than a V chord. These conditions reinforce one another when the judgement is between a I and a $V^6$, but conflict when the judgement is between a $I^6$ and a V. In actual music such conflicts are resolved by context: on the assumption that the two events come up for comparison in the same unit (time-span segment or prolongational region), the reductional preference rules decide which event is more stable on the basis of rhythmic position, linear function, and other factors.

Stability conditions are contingent on the basic materials and properties out of which the event structure of a piece is made. In Classical tonal music, for instance, the basic materials include the diatonic scale, which

has certain internal and transpositional properties that can project judgements of relative stability (of which more below). Musical idioms vary in these respects. It is possible though rare for the basic materials not to be used in a way that gives rise to stability conditions: one can imagine a treatment of the diatonic scale where all scale degrees are functionally equivalent and the circle of fifths is not employed. Similarly, both Bach and Schoenberg used the well-tempered chromatic scale, but only the former carved it up so as to induce clear judgements of relative stability among pitch configurations; the latter instead sought a noncentric musical universe in which the 'twelve tones . . . are related only with one another' (Schoenberg 1941). However, since learning and memory depend on hierarchical structuring (Constraint 2), and since event hierarchies depend on stability conditions (Constraints 7 and 8), it follows that stability conditions are cognitively advantageous.

What are the constraints on stability conditions? First, there must be a fixed collection of sonic elements (normally pitches; but see Lerdahl 1987). The assumption of such collections might go unremarked except that in electronic music they are fairly rare. When there is a virtual infinitude of sonic possibilities, it is hard for a composer to select and stick with a limited collection. Non-selection, however, means no syntax.

> *Constraint 9*: Stability conditions must operate on a fixed collection of elements.

A fixed collection should not be thought of as completely rigid. First, an element of a collection need only be perceived categorically (Burns and Ward 1982); otherwise a slightly mistuned pitch would cause havoc. Second, as long as the categories remain the points of reference, there can be embellishing or sliding within or between categories (as in microtonal inflexions in Indian music or string glissandos in Western music).

The elements of a collection can be placed along a dimension to form a scale. The intervals of a scale have general size limitations: they must be large enough for adjacent elements to be easily discernible, but not so large as to use up excessive space along the continuum.

> *Constraint 10*: Intervals between elements of a collection arranged along a scale should fall within a certain range of magnitude.

Constraints 9 and 10 are analogous to Constraint 1 (the requirement of discrete events). Both cases provide the building blocks for further organization, Constraints 9 and 10 for stability conditions and Constraint 1 for event hierarchies. A fixed collection with moderate interval sizes allows other stability conditions to emerge. In surveying some of these, let us assume that the collections are pitch collections—as opposed, say, to drums or synthesized timbres.

Most pitch collections take advantage of the 2:1 frequency ratio of the

octave, so that in one respect pitches reduce to 'pitch classes'. There are 88 pitches on the piano but only 12 pitch classes. Thus, in addition to giving a recurring structure to the overall collection, the octave decreases to a more manageable size the memory load for elements of the collection.

> *Constraint 11:* A pitch collection should recur at the octave to produce pitch classes.

The octave can be divided up in innumerable ways. A common route has been to make a collection beginning with other small-ratio intervals and progressing gradually to intervals with larger ratios. From the present perspective, this route is advantageous because the resultant intervals provide a broad and graduated palette of *sensory* consonance and dissonance. Sensory consonance and dissonance can in turn form the basis for *musical* consonance and dissonance, where in a general sense consonance is equivalent to stability and dissonance to instability. Thus a seventh in Classical tonal music resolves to a sixth not just out of cultural convention but because the syntactical resolution is supported by sensory experience.

The claim is not that pitch configurations exhibiting relative sensory consonance must necessarily be more stable in musical contexts than those exhibiting relative sensory dissonance. The reverse can occur contextually, and shows up in the reductions where appropriate. Rather the claim is that the stability conditions will be relatively ineffectual unless they are supported by sensory consonance and dissonance. Stability conditions in which a seventh is supposed to be more stable than a sixth would contradict sensory experience and would not lead to event hierarchies of any depth of embedding.

Strictly speaking, these remarks hold for fundamentals with harmonic partials, such as produced by the voice and by all instruments in which the source is periodically excited. For intervals with very small ratios, the partials of the respective fundamentals are largely reinforcing, creating the sensation of consonance. Sensory dissonance, or 'roughness', arises when frequencies (whether of fundamentals or of partials) interfere within a critical bandwidth, which extends in most musical registers from a fraction of a semitone to a little less than a minor third (Plomp and Levelt 1965). The critical bandwidth is a consequence of the resolving power of the hearing mechanism in the peripheral auditory system. One could also invoke pattern-recognizing templates of the central auditory system, such as proposed in Terhardt (1974); these for present purposes lead to the same conclusion.

> *Constraint 12*: There must be a strong psychoacoustical basis for stability conditions. For pitch collections, this entails intervals that proceed gradually from very small to comparatively large frequency ratios.

However, as is well known, small interval ratios—whether in just or Pythagorean tuning—present problems for the transposition of intervals and chords. A reading of Partch (1949) might additionally suggest that the number of pitches and intervals in such a collection can become alarmingly large. An alternative route that avoids these problems is to divide the octave into equal intervals. The collection then permits equivalence under transposition, and the pitches and intervals to be learned are limited.

> *Constraint 13*: Division of the octave into equal parts facilitates transposition and reduces memory load.

But these advantages are offset by the fact that almost all equal divisions of the octave lead to interval ratios that are complex and that do not easily round off categorically to simple ratios. As a result the intervals give a small range of sensory consonance and dissonance, and Constraint 12 is not satisfied.

The 12-fold division of the octave (and to a lesser extent the 19-fold division; see Yasser 1932) is an exception: its intervals are sufficiently close to just intervals to be useful; that is, the deviations from just intonation are not large enough to cause serious roughness from otherwise consonant intervals. The slight loss in distinctions in sensory consonance and dissonance is compensated for by the possibility of unfettered transposition. The familiar chromatic scale has presumably survived so well because it meets the demands of both Constraints 12 and 13.

At this point two routes are available, one viewing the chromatic collection (or set) essentially as a whole, the other thinking of it as material for subsets that have various musical and psychological properties. Let us review these alternatives.

One can evoke stability conditions directly from the chromatic set through a selection and patterning of different degrees of sensory consonance and dissonance. For example, Debussy's and Bartók's harmonies often 'modulate' among chord types of varying degrees of dissonance. Schoenberg was at least tacitly occupied with such a notion both early and late in his career: early, notably in the First Chamber Symphony (1906), where the basic dissonant-to-consonant progression is from a chord built in fourths to a whole-tone chord to a triad; late, in the 12-tone works of the 1940s that include tonal features, particularly at cadential points. Hindemith's (1937) theory of 'harmonic fluctuation' is based on a similar idea, generalized on a supposedly psychoacoustic foundation to embrace all possible sonorities. It is profitable to view harmonic progression in the work of Machaut and other fourteenth-century composers in such a light. Much of my own work also fits in this category.

The other route is taken by the Classical tonal system, in which the referential sonority is always the triad. Various versions of the triad alone do not project adequate differences in sensory consonance and dissonance.

Stability conditions among triads are instead achieved through the use of subsets of the chromatic set, notably the diatonic. (Historically, of course, the diatonic scale precedes the chromatic, and until Wagner it was musically paramount as well. It can be illuminating, however, to view the diatonic in terms of the chromatic.)

Balzano (1980, 1982) shows through a group-theoretic analysis that only certain equal-interval sets—the 12-fold among them—lead to subsets incorporating the psychologically important criteria of intervallic 'uniqueness', scale-step 'coherence', and transpositional 'simplicity'. All three criteria have to do with location in 'pitch space'. The first criterion demands a unique 'vector of relations' for each pitch in a subset, so that the listener can orient himself unambiguously in relation to the other pitches. For example, each scale degree of the asymmetrical diatonic subset of the chromatic set has a non-duplicating intervallic relation to the other scale degrees. In contrast, the octatonic subset (alternating half- and whole-steps), which has been widely used in this century, is not 'unique', since its vector repeats four times in an octave (this is the 'charm of impossibilities' described in Messiaen 1944). The second criterion ensures that the intervals of a scale proceed in an orderly manner, so that, for instance, the same distance is not traversed by one step in one part of a scale but by two or more steps in another part. The harmonic minor scale, not to mention more extreme examples, would thereby be disqualified, since the augmented second between the sixth and seventh scale degrees is equal in distance to the minor thirds that elsewhere take two steps to traverse; this presumably explains why the scale is avoided melodically in much diatonic music. The third criterion ensures orderly transposition of a subset. The diatonic subset, for example, drops one pitch and adds another as it transposes along the circle of fifths, so that distance in transposition correlates with the number of pitches shared in two diatonic scales.

> *Constraint 14*: Assume pitch sets of *n*-fold equal divisions of the octave. Then subsets that satisfy uniqueness, coherence, and simplicity will facilitate location within the overall pitch space.

Balzano surprisingly demonstrates that in the 12-fold chromatic set only the pentatonic and diatonic subsets satisfy all three criteria. He advances this rather than traditional psychoacoustical explanations as the reason for the cultural ubiquity of these scales; and he recommends that the few other *n*-fold divisions of the octave that permit these features (such as the 20-fold set) be explored for their musical potential. However, none of the other candidates approximates just intonation. The real cognitive appeal of the chromatic set, I would argue, is that only here do Constraints 12, 13 and 14 all converge intact. The structural features of the pentatonic and diatonic subsets would appear to be correspondingly special. (For discussion relating to this section, see Sloboda 1985; Dowling and Harwood 1986.)

## Pitch space

I referred above to 'pitch space'. I want now to examine this notion with the purpose of extracting a few more constraints on stability conditions.

A number of cognitive psychologists, notably Longuet-Higgins (1978) and Shepard (1982), have proposed multidimensional representations for pitch relations.* The basic idea is that in a tonal framework pitches are cognized as relatively close to or far from one another, and that these distances can be represented geometrically. To take the simplest case, pitches are heard as proximate in terms of both pitch height and pitch class, so at least a two-dimensional representation is necessary. Another dimension can be added to account for fifth-relatedness; and so on. One can think of these representations as internalized maps. Just as cities are close to or far from home and can best be reached via some roads rather than others, so it is with tonal relations.

> *Constraint 15*: Any but the most primitive stability conditions must be susceptible to multidimensional representation, where spatial distance correlates with cognitive distance.

The geometric pitch spaces proposed in the psychological literature seem promising, but I have lingering doubts. First, granting the insights provided by Balzano (1982) and Shepard (1982), it remains odd that the chromatic collection is taken as the basis for the diatonic music they seek to explain. A second and related point is that, in pursuit of geometric symmetry, these models tend to come up with more regularity than diatonic music warrants. For example Longuet-Higgins's (1978) array has a major-third axis, and Shepard's (1982) double helix is made up to two strands of whole-tone scales. But only the semitone and perfect-fifth interval cycles are relevant to diatonic music, which is in many ways asymmetrical. Other interval cycles do not begin to assume musical significance until nineteenth-century chromaticism.

These questions lead me to suggest a complementary pitch space that has a reductional rather than a geometric format. This space is musically obvious yet has explanatory power. I take the lead from Deutsch and Feroe (1981), who speak of different pitch 'alphabets'—chiefly the chromatic, diatonic, and triadic—that the tonal listener 'highly overlearns' from previously heard musical surfaces. This in itself is an important point, since without repeated exposure to such regularities, listeners could not spontaneously organize new input in terms of them.

* Actually, there is a long tradition in music theory for such models, dating back to Weber (1824) and his eighteenth-century predecessors; see Werts (1983). Schoenberg's (1954) chart of regions, which has been cited in the psychological literature, can be traced through Riemann back to Weber.

*Constraint 16*: Levels of pitch space must be sufficiently available from musical surfaces to be internalized.

Deutsch and Feroe show that memory load is greatly reduced if tonal melodies are structured by means of their alphabets. As their formalism suggests, such alphabets are hierarchically related as in Fig. 10.7: the octave level *a* elaborates into the perfect-fifth level *b*, which elaborates into the triadic level *c*, which elaborates into the diatonic level *d*, which elaborates into the chromatic level *e*. In addition, pitches at higher levels in the hierarchy are more stable than pitches that do not appear until successively lower levels. This correlates with Krumhansl's (1979) experimt results. So Fig. 10.7 includes a partial representation of stability conditions.




**Fig. 10.7**

This pitch space circumvents the reservations mentioned above. First, the overall orientation is not chromatic. Non-well-tempered distinctions can be incorporated merely by expanding the content of level *e*, leaving the other levels untouched. Second, level *e* expresses an equal interval cycle and other unwanted equal divisions of the octave are absent.

As chords and diatonic collections change, so too must the hierarchy in Fig. 10.7. For example, in the framework of a IV chord in C major, levels *a*, *b* and *c* would shift respectively to F–F, F–C–F, and F–A–C–F. If the diatonic collection changes to that of F major, level *d* would also shift, with B flat replacing B. The structure of this shifting pitch space demands further representation, perhaps geometic in format, but we will not pursue this here.

Each level in Fig. 10.7 can be thought of as a scale with stepwise motion occurring between adjacent members. An octave is a step away at level *a*, as is a fifth at level *b*; an arpeggiation proceeds by step at level *c*; diatonic and chromatic progressions are by step at levels *d* and *e*. A skip is a progression between two elements that requires more than one step for traversal. This step/skip distinction is important for establishing cognitive distance, and was implicit in Constraint 14 above.

Fig. 10.7 can further be seen as expressing preferred melodic routes via the Gestalt laws of proximity and good continuation. A preferred route is one that is 'complete', where completeness is defined as a stepwise

progression at any level that begins and ends at elements represented at the next higher level. Thus a complete chromatic progression must begin and end on diatonic pitches, and so on. Anything else—a chromatic appoggiatura, for instance—is felt as 'incomplete'. Generally, incompleteness is more acceptable at the start than at the end of a progression; unstable elements must be subsequently 'anchored' (Bharucha 1984).

> *Constraint 17*; A reductionally organized pitch space is needed to express the steps and skips by which cognitive distance is measured and to express degrees of melodic completeness.

The notion of completeness ties in significantly with current music theories that are otherwise dissimilar. Meyer's (1973) and Narmour's (1977) implication-realization theory depends exactly on the space in Fig. 10.7: a stepwise motion at any level implies the next step at that level; such an implication can be realized or not, as the case may be. And Schenker's (1935) central conception of the *Zug*, a diatonic progression bounded by pitches of the prolonged harmony, falls out as a special case of completeness at levels *c* and *d* in Fig. 10.7. Other Schenkerian concepts (coupling, arpeggiation, the initial ascent, the *Urlinie*, the *Bassbrechung*) can be illuminated in related fashion. Completeness appears to be the basic voice-leading principle in Classical tonal music.

Two general psychological points may be inserted here. First, even if—as at the fifth, triadic, and diatonic levels—intervals along a level are psychoacoustically unequal, they are cognitively equal because they are steps at that level. Second, because each of these levels constitutes a category in terms of which the next level is understood, Miller's (1956) restrictions on memory load ('seven plus or minus two') are easily met, something not accomplished by the single category of the chromatic set.

Finally, Fig. 10.7 reflects certain psychoacoustic facts. In general, a descent from level *a* to level *e* brings increasing sensory dissonance. Level *a* gives the 'virtual pitch' (or root) of the collections in levels *b* and *c* (Terhardt 1978). Level *b* serves as the harmonic norm for medieval music and various ethnic musics, level *c* for Classical tonal music. Level *d* provides the melodic basis for many musical idioms, with level *e* offering inflectional possibilities. Thus between levels *c* and *d* there appears a conceptual line separating harmony and melody. This differentiation may in part stem from the critical bandwidth, since for most registers steps at level *c* fall outside the bandwidth and steps at level *d* fall inside.

Fig. 10.7 embodies all of the constraints on stability conditions enumerated in the last two sections: fixed collections with appropriate interval sizes (Constraints 9 and 10); octave equivalence at level *a* (Constraint 11); increasing sensory dissonance from level *a* to level *d* (Constraint 12); equal division of the octave at level *e* (Constraint 13); uniqueness, coherence, and simplicity at level *d* (Constraint 14—and,

incidentally, uniqueness and coherence at level *c*); multidimensional representation expressing cognitive distances (Constraint 15); levels of pitch space that are easily induced from a wide variety of musical surfaces (Constraint 16); and steps, skips, and degrees of melodic completeness at multiple reductional levels (Constraint 17). The cultural persistence of the many musical idioms relating to Fig. 10.7 may be due to this convergence.

Some of these constraints seem to me binding, others optional. Constraints 9–12 are essential for the very existence of stability conditions. Constraints 13–17, on the other hand, can be variously jettisoned. The resulting stability conditions may be weaker, but they can still lead to hierarchically rich music. For example, South Indian music approximates just intonation and does not modulate, thereby ignoring Constraint 13 and part of Constraint 14; Debussy, Bartók, and others have developed consonance-dissonance patterns directly from the total chromatic, thereby ignoring Constraints 14–17. It will be fascinating to discover how the new sound materials of computer music will be able to meet these constraints. In all probability the new materials will bring additional requirements into play. Meanwhile there remains ample leeway within the chromatic collection and within the constraints listed above for music as yet unimagined.

## Cognitive opacity of serialism

Now we are in a position to see why serial organizations are inaccessible to mental representation. Rather than explore the subject through the idiosyncratic serialism of *Le Marteau*, I will refer to elementary aspects of the Classical 12-tone system, which has been more widely influential (Schoenberg 1941). I will give three causes for serialism's cognitive opacity.*

Before proceeding, I must emphasize that the issue is not whether serial pieces are good or bad. As with tonal music, some serial pieces are good and most are bad. Nor am I claiming that listeners infer no structure at all from musical surfaces composed with serial techniques. What listeners in fact infer from such surfaces is an interesting question, one that deserves theory and experiment in its own right. But this is not the issue here. The issue is why competent listeners do not hear tone rows when they hear serial pieces.

The first reason is that serialism is a permutational rather than

* It must be said that, if only from the evidence in Schoenberg's 12-tone sketches (Hyde 1980), Schoenberg's (1941) account of the 12-tone system is partly misleading. Plainly he used the system not just to generate rows but to create certain systematic but non-serial relationships among subsets. This, however, does not affect my argument, which concerns the cognitive opacity of serial structures (tone rows) as such.

elaborational system. Pitch relations in virtually all 'natural' compositional grammars are elaborational. Take for example the pitch sequence B–E–F sharp. It is easy to imagine any number of melodic embellishments internal to the sequence. This typically happens when children sing and performers improvise. Like the syntax of a sentence (Chomsky 1957), musical elaborations can continue recursively to an indefinite level of complexity; for example, sonata form is an expansion of the Classical phrase (Schenker 1935: Rosen 1972; Lerdahl and Jackendoff 1983). This feature enables pitch relations to be described hierarchically by a tree notation.

Serialism instead depends on specific orderings of the elements of a set. Distinctions arise from permutations of the elements; for example, the inversion of a 12-tone row has the same 12 pitch classes but in a different order. The order position of the elements is therefore essential to the identity of individual set forms. From this it follows that internal elaboration of any element by other elements will undermine a set form. Consider the row of Schoenberg's Violin Concerto (1936), shown in Fig. 10.8 with order numbers given above. The sequence B–E–F sharp appears at order numbers 3–5. If one wanted to elaborate E, say, with G or with B flat–A, creating the sequences B–E–G–E–F sharp or B–E–B flat–A–E–F sharp, the integrity of the row would be destroyed, since G, B flat, and A have order positions elsewhere in the row. A tone row is not an elaborational structure. (The same point could be made with reference not to elements but to the intervals between elements.)

| Order number: | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 12-tone set: | A | B♭ | E♭ | B | E | F♯ | C | C♯ | G | A♭ | D | F |

**Fig. 10.8**

The little research that has been directed towards serialism (as in Francés 1958; Dowling 1972; Deutsch 1982; Bruner 1984) supports the contention that permutational structures are hard to learn and remember. Since other human activities are not organized in such a fashion, it is hardly surprising that the issue has in general been ignored by psychologists. The situation reminds me, though, of Chomsky's (1965) observation that many logically possible grammatical constructions never occur, such as forming interrogatives from declaratives by word reversal or by exchange of odd and even words. There is unfortunately no one on whom to test these linguistic counter-examples. Music offers a unique opportunity in this regard. Surely children can be found who have been raised on a steady diet of serial music. Do they identify tone rows? A negative answer will provide strong counter-evidence concerning the structure of musical cognition, and may suggest inherent limitations on cognitive organization in other domains as well.

Ironically, Schoenberg was much preoccupied with the issue of compre-

hensibility. I suspect this is one reason why in his 12-tone phase he adopted Classical motivic, phrasal, and formal structures. As a result, his serial music satisfies the rhythmic Constraints 3–6. But the permutational basis of his pitch organization assures a gap between the compositional and listening grammars.

It is of course possible to organize the combinations and sequences of individual rows on a hierarchical rather than permutational basis. Boulez, Babbitt, and others have done just that. But these higher-level hierarchies are extremely difficult to cognize in a hierarchical fashion because their underlying basis, the row, remains non-tree-like. There are further reasons as well, which brings us to the other causes of serialism's cognitive opacity.

The second cause is the avoidance of sensory consonance and dissonance as an input to the system (Constraint 12). This avoidance in turn stems from the historical collapse of the stability conditions underlying Classical tonal pieces. The pre-serial atonal period had established an aesthetic in which hierarchical pitch relations gave way to purely contextual, associative ones. The 12-tone system reinforced this trend through constant reshuffling of the total chromatic and through constant repetition of the intervallic patterns of the row. Concerned about the natural basis of his new aesthetic, Schoenberg (1911, 1941) invoked the doctrine of the 'emancipation of the dissonance': consonance and dissonance are not opposites but are on a continuum determined by the overtone series; dissonances are harder to comprehend than consonances; but aquaintance with the more remote overtones has made dissonant intervals just as comprehensible as consonant intervals; so now intervals may be treated as equally consonant (or dissonant). Babbitt (1965) goes a step further, debunking the overtone series as a basis for pitch systems and claiming that consonance-dissonance distinctions among intervals are entirely contextual in origin.

Neither Schoenberg nor Babbitt distinguishes between sensory and musical consonance and dissonance. Though the details of pitch perception are not fully resolved, there is no doubt about the objective existence of sensory consonance and dissonance. For intervals between pitches with harmonic partials, the degree of sensory dissonance bears an important relationship to the overtone series, by virtue of the degree of interference among partials within the critical bandwidth. Thus far Schoenberg, at least, is right. What one does with this musically is another matter. Classical tonality founds musical consonance and dissonance largely on sensory consonance and dissonance (partial exception must be made for the perfect fourth, the minor triad, and so on). Serialism, on the other hand, treats intervallic combinations in terms of the row and in effect defines musical consonance and dissonance out of existence (with the glaring exception of the octave: pitch classes are needed). But this strategy, while perfectly logical, does not neutralize sensory consonance and dissonance. The sensory dissonance of a seventh remains greater than that of a sixth,

regardless of the musical purposes to which these intervals are put.

There is an important psychological consequence: by ignoring sensory distinctions, serialism creates musical contexts that are not apprehended hierarchically. One pitch or hexachord may be associated with another pitch or hexachord, but these relationships are not easily heard in a dominating-subordinating manner. The requisite intuitions of stability and instability are missing. And because hierarchies are not inferred, tone rows and their combinations are difficult to comprehend (Constraint 2).

The third cause is that serialism does not induce a pitch space where spatial distance correlates with cognitive distance (Constraint 15). Assume for simplicity that serial pitch space looks like the 'combinatorial' space shown in Fig. 10.9 for Schoenberg's Violin Concerto. Here the first six pitch classes of the inverted set, transposed up five semitones, are identical to the second six pitch classes of the prime set, though in a different order. For a variety of reasons such a space was useful to Schoenberg (see Schoenberg 1941). A complete 12-tone pitch space would include transpositions of this and other relationships, just as the tonal space of Fig. 10.7 would have to be enlarged for other chords and regions.

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| P0: | A | B♭ | E | B | E | F♯ | C | C♯ | G | A♭ | D | F |
| I5: | D | C♯ | A♭ | C | G | F | B | B♭ | E | E♭ | A | F♯ |

**Fig. 10.9**

One might suppose that a 'step' in such highly chromatic music is a semitone. But this is not reflected at all in the pitch space of Fig. 10.9. Or, more in the spirit of the row, one might say that a 'step' occurs from any pitch class to any horizontally adjacent pitch class. But in this case step distance varies wildly from one adjacency to the next, and there is no correlation with psychoacoustic (log frequency) distance. In short, serialism does not incorporate any psychologically coherent notion of step and skip. The listener consequently has difficulty locating a pitch as close to or far from another pitch.

Further causes of serialism's cognitive opacity could be adduced, such as its failure to provide consistent exposure from piece to piece to a limited number of 'alphabets' (Constraint 16). But the main points have been addressed. We turn now to some aesthetic issues that have been lurking behind this entire discussion.

## Comprehensibility and value

There is no obvious relationship between the comprehensibility of a piece and its value. Many masterpieces are esoteric, while most ephemeral

music is all too comprehensible. On the other hand, if a piece cannot be understood, how can it be good? Most would agree that comprehensibility is a necessary if not sufficient condition for value.

Care must be taken with this formulation in three respects. First, comprehension presupposes listening competence for the music in question. This competence varies with ability and especially with exposure, but is not less real for that. Second, comprehension pertains to the listening grammar rather than to the compositional grammar. A serial piece may be understood in non-serial ways. Third, we are talking about intuitive rather than analytic comprehension. Along the lines of Fodor (1983), the mind's music module must be able spontaneously to form mental representations of musical structure from musical surfaces. This is quite different from using the all-purpose reasoning faculty to figure out the structure of a piece.

With these provisos in mind, I think the above formulation stands. Appreciation depends on cognition. I now want to go an aesthetic step further.

> *Aesthetic Claim 1*; The best music utilizes the full potential of our cognitive resources.

This seemingly innocuous statement carries weight because a great deal is becoming known about how musical cognition works. I have outlined aspects of this understanding in the constraints presented above. Following them will not guarantee quality. I maintain only that following them will lead to cognitively transparent musical surfaces, and that this is in itself a positive value; and, conversely, that not following them will lead in varying degrees to cognitively opaque surfaces, and that this is in itself a negative value.

This stance can be refined through the notion of musical 'complexity', which is to be contrasted with musical 'complicatedness' (Lerdahl 1985). A musical surface is complicated if it has numerous non-redundant events per unit time. Complexity refers not to musical surfaces but to the richness of the structures inferred from surfaces and to the richness of their (unconscious) derivation by the listener. For example, a grouping structure is complex if it is deeply embedded and reveals structural patterns within and across levels. The derivation of such a structure is complex if all the grouping preference rules come into play and if they conflict with one another to a certain degree. (Total reinforcement of the rules would produce a stereotypical grouping structure; total conflict would create intolerable ambiguity.) I take complicatedness to be a neutral value and complexity to be a positive one. Many musical surfaces meet the various constraints, but only those that lead to complexity employ 'the full potential of our cognitive resources'.

All sorts of music satisfy these criteria—for example, Indian raga,

Japanese koto, jazz, and most Western art music. Balinese gamelan falls short with respect to its primitive pitch space. Rock music fails on grounds of insufficient complexity. Much contemporary music pursues complicatedness as compensation for a lack of complexity. In short, these criteria allow for infinite variety, but only along certain lines.

I find this conclusion both exciting and—initially, at least—alarming. It is exciting because psychology really does have something substantive to say about how music might be; here is the foundation I was seeking. It is alarming because the constraints are tighter than I had bargained for. Like the old avant-gardists, I dream of the breath of other planets. Yet my argument has led from pitch hierarchies (Constraints 7–8) to an approximation of pure intervals (Constraint 12), to diatonic scales and the circle of fifths (Constraint 14), and to a pitch space that prominently includes triads (Constraint 17).

However, the constraints do not prescribe outworn styles. Rather they provide a prototype (Rosch 1975). Let me first give an uncontroversial rhythmic example. A musical surface in which the note values are multiples of 2 is intrinsically more stable and easier to cognize than one in which the note values are multiples of 7 and 11. This does not mean that the latter surface is somehow impermissible. It instead amounts to the observation that, because of the resultant ease in forming a metrical structure, note values that are multiples of smaller prime numbers are easier to process and remember, and therefore that multiples of 2 (or 3) inevitably remain a cognitive reference point for more complicated rhythms. I claim that a similar situation holds for pitch: the structure in Fig. 10.7 (with its hierarchically organized octave, fifth, triad, and diatonic scale) remains a reference point for other kinds of pitch organization, not because of its cultural ubiquity but because it incorporates all of the constraints developed above. I take this as a given, with or against which a composer can play creatively. Of course one may opt for a less constrained pitch space. But if a composer chooses the space represented by Fig. 10.7, I am sure there are innumerable and radically new ways to use and extend it. The future is open.

My second aesthetic step was discussed above; I list it here only for completeness.

> *Aesthetic Claim 2*: The best music arises from an alliance of a compositional grammar with the listening grammar.

This claim does not exclude the artifice hidden in a Bach fugue or a Brahms intermezzo. Such artifice is rooted in the bedrock of a 'natural' compositional grammar. At our present musical juncture, however, composers would do well to heed the claim.

This claim carries with it a historical implication. The avant-gardists from Wagner to Boulez thought of music in terms of a 'progressivist'

philosophy of history: a new work achieved value by its supposed role *en route* to a better (or at least more sophisticated) future. My second aesthetic claim in effect rejects this attitude in favour of the older view that music making should be based on 'nature'. For the ancients, nature may have resided in the music of the spheres, but for us it lies in the musical mind. I think the music of the future will emerge less from twentieth-century progressivist aesthetics than from newly acquired knowledge of the structure of musical perception and cognition.

## References


Babbitt, M. (1965). The structure and function of musical theory. *College Music Symposium* **5**, 10–21. (Reprinted in 1972 in *Perspectives on contemporary music theory* (ed. B. Boretz and E. T. Cone). Norton, New York.)

Balzano, G. J. (1980). The group-theoretic description of 12-fold and microtonal pitch systems. *Computer Music Journal* **4**, (4), 66–84.

Balzano, G. J. (1982). The pitch set as a level of description for studying musical pitch perception. In *Music, mind, and brain* (ed. M. Clynes). Plenum, New York.

Bharucha, J. J. (1984*a*). Anchoring effects in music: the resolution of dissonance. *Cognitive Psychology* **16**, 485–518.

Bharucha, J. J. (1984*b*). Event hierarchies, tonal hierarchies, and assimilation: a reply to Deutsch and Dowling. *Journal of Experimental Psychology: General* **113** (3), 421–5.

Boulez, P. (1954). *Le marteau sans maître.* Universal Edition, London.

Boulez, P. (1963). *Penser la musique aujourd'hui.* Schott, Mainz. (Tr. 1971 by S. Bradshaw and R. R. Bennett. Harvard University Press, Cambridge, Mass.)

Boulez, P. (1964). Alea. *Perspectives of new music* **3** (1), 42–53.

Bruner, C. L. (1984). The perception of contemporary pitch structures. *Music Perception* **2** (1), 25–40.

Burns, E. M. and Ward, W.D. (1982). Intervals, scales, and tuning. In *The psychology of music* (ed. D. Deutsch). Academic Press, New York.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and two ears. *Journal of the Acoustical Society of America* **25**, 975–9.

Chomsky, N. (1957). *Syntactic structures.* Mouton, The Hague.

Chomsky, N. (1965). *Aspects of the theory of syntax.* MIT Press, Cambridge, Mass.

Deliège, I. S. (1985). Les régles preferentielles de groupement dans la perception musicale. Unpublished Ph.D. thesis. Université de Bruxelles.

Deutsch, D. (1982). The processing of pitch combinations. In *The psychology of music* (ed. D. Deutsch). Academic Press, New York.

Deutsch, D., and Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review* **88**, 503–22.

Dowling, W. J. (1972). Recognition of melodic transformations: inversion, retrograde, and retrograde-inversion. *Perception and Psychophysics* **12**, 417–21.

Dowling W. J., and Harwood, D. L. (1986). *Music cognition.* Academic Press, New York.

Fodor, J. A. (1983). *The modularity of mind.* MIT Press, Cambridge, Mass.
Francés, R. (1958). *La perception de la musique.* Vrin, Paris.
Fux, J. J. (1725). *Gradus ad parnassum.* Joannis Petri van Gehlen, Vienna. (Tr. and ed. 1965 by A. Mann. Norton, New York).
Hindemith, P. (1937). *Unterweisung im tonsatz*, Vol. 1. Schott, Mainz. (Tr. 1942 by A. Mendel. Associated Music Publishers, New York.)
Hyde, M. (1980). The roots of form in Schoenberg's sketches. *Journal of Music Theory* **24** (1), 1–36.
Koblyakov, L. (1977). P. Boulez, 'Le Marteau sans Maître,' analysis of pitch structure. *Zeitschrift für Musiktheorie* **8**, (1), 24–39.
Krumhansl, C. L. (1979). The psychological representation of pitch in a tonal context. *Cognitive Psychology* **11**, 346–74.
Krumhansl, C. L. (1983). Perceptual structures for tonal music. *Music Perception* **1** (1), 28–62.
Lerdahl, F. (1985). Théorie générative de la musique et composition musicale. In *Quoi? quand? comment?: la recherche musicale* (ed. T. Machover). Christian Bourgois, Paris.
Lerdahl, F. (1987). Timbral hierarchies. *Contemporary Music Review* **1**, (3–4), 135–60.
Lerdahl, F. and Jackendoff, R. (1983). *A generative theory of tonal music.* MIT Press, Cambridge, Mass.
Longuet-Higgins, H. C. (1978). The perception of music. *Interdisciplinary Science Reviews* **3**, 148–56.
McAdams, S. and Bregman A. (1979). Hearing musical streams. *Computer Music Journal,* **3** (4), 26–43. (Reprinted in 1985 in C. Roads and J. Strawn, eds., *Foundations of computer music.* MIT Press, Cambridge, Mass.)
Messiaen, O. (1944). *Technique de mon langage musical.* Alphonse Leduc, Paris.
Meyer, L. B. (1973). *Explaining music.* University of California Press, Berkeley.
Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review* **63**, 81–96.
Miller, G. A., Galanter, E., and Pribram, K. (1960). *Plans and the structure of behaviour.* Holt, Rinehart and Winston, New York.
Narmour, E. (1977). *Beyond Schenkerism.* University of Chicago Press.
Neisser, Ulric. (1967). *Cognitive psychology.* Prentice-Hall, Englewood Cliffs, NJ.
Palmer, C., and Krumhansl, C. L. (in press). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance.*
Partch, H. (1949). *Genesis of a music.* University of Wisconsin Press, Madison. (Republished 1974 by Da Capo, New York.)
Plomp, R., and Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America* **38**, 548–60.
Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology* **7**, 532–47.
Rosen, C. (1972). *The classical style.* Norton, New York.
Schenker, H. (1935). *Der Freie Satz.* Universal Edition, Vienna. (Tr. 1979 by E. Oster. Longman, New York).
Schoenberg, A. (1911). *Harmonielehre.* Universal Edition, Vienna. (Tr. 1978 by R. E. Carter, University of California Press, Berkeley.)

Schoenberg, A. (1941). Composition with twelve tones. In A. Schoenberg. *Style and Idea.* St Martins, New York, new expanded edn 1975.

Schoenberg, A. (1954). *Structural functions of harmony.* Norton, New York.

Shepard, R. (1982). Structural representations of musical pitch. In *The psychology of music* (ed. D. Deutsch). Academic Press, New York.

Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society* **106**, 467–82. (Reprinted 1972 in H. A. Simon, *The sciences of the artificial.* MIT Press, Cambridge, Mass.)

Sloboda, J. A. (1985). *The musical mind.* Oxford University Press, Oxford.

Terhardt, E. (1974). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America* **55**, 1061–9.

Terhardt, E. (1978). Psychoacoustic evaluation of musical sounds. *Perception and Psychophysics* **23**, 483–92.

Todd, N. (1985). A model of expressive timing in music. *Music Perception* **3** (1), 33–57.

Weber, G. (1824). *Versuch einer Geordeneten Theorie.* Schott, Mainz.

Werts, D. (1983). A theory of scale references. Unpublished Ph.D. thesis. Princeton University.

Yasser, J. (1932). *A theory of evolving tonality.* American Library of Musicology, New York.

# 11

# From collections to structure: the developmental path of tonal thinking

LYLE DAVIDSON and PATRICIA WELSH

Imagine three musicians composing a piece. Each is seated at a piano; manuscript paper is on the music rack; pens are alongside. The first musician plays some notes on the piano, erases a few notes he has already scribbled on the manuscript paper, and replaces them with new ones. He continues in this manner until he finally completes the melody. After completing it, he plays it on the piano, immediately starts revising it, rubbing out whole sections. He stops before finishing and goes for a walk. The second musician spends some time staring at the blank manuscript paper, plays a few notes over and over on the piano, and after a few false starts writes out the complete melody on paper. During the final playing of the melody on the piano his finger slips and he hits a 'wrong' note. Immediately, he is interested in his mistake. He incorporates the 'error' into his written melody, replacing the original note. The third musician scans the page, checks the pen, thinks a moment, and writes out the melody without hesitation, as if taking down dictation.

These three working styles represent those of well-known composers. It is easy to imagine Beethoven, Stravinsky, and Mozart hard at work, creating significant musical compositions, solving complex musical problems. But at the same time, it is difficult to imagine the thought processes which supported their efforts. We know only that each composer had a very different working style and the working styles themselves tell us very little beyond the fact that there are different ways of composing.

If we replay the scene, and specify the task, 'compose a melody which starts in one key, modulates to a specific key, and returns to the original key at the end', our interpretation of these working styles changes. We expect composers of the stature of Beethoven, Stravinsky, or Mozart to be able to dash off such a melody easily. However, it is easy to imagine

inexperienced musicians having trouble with this task. Young musicians writing such a melody might show working styles very much like those described. Furthermore, the working styles now take on additional meaning. Under these circumstances, it is relatively easy to imagine the working style of the beginner, the intermediate, and the expert. Different working styles can reflect different levels of skill.

Tasks similar to this one are common in music schools. While the writing of melody is rarely studied, work in harmony and counterpoint focuses on specific aspects of composing within a tonal framework. These studies are a standard part of professional training for the classically oriented musician. However, little is known about the mental and cognitive processes which support different orientations to the musical task of writing a melody which modulates. There are two reasons for this. First, melody writing has not been systematically studied. Second, in music schools, musicians tend not to be interested in the processes involved in arriving at solutions to musical problems, but only in musically successful products.

## Perception and production of tonality

While musicians are more interested in production, psychologists take a different tack. They focus on the perception of tonality by untrained individuals. Several studies have compared the recognition of melodic shapes which form tonal patterns to those which form atonal patterns (Frances 1958; Dowling and Fujitani 1971; Zenatti 1975; Cuddy and Lyons 1981). These studies consistently report that tonal patterns are easier to process, recognize, and recall than atonal patterns. Still other work suggests the presence of tonality plays a central role in our perception and cognition of richer musical contexts. Research indicates that the components of melodic textures are easier to grasp when they are in the context of closely related keys than in more distantly related keys (Sloboda and Edworthy 1981).

Apparently, little training is needed to establish highly useful tonal schemata when perception is required. Krumhansl (1979) shows that scales are highly stable mental structures even for untrained listeners. Her work also suggests that relationships between keys parallel those of the notes of the diatonic scale (Krumhansl *et al.* 1982). Taken together, these various studies indicate that tonality and scales are powerful psychological structures which play an important role in our comprehension and enjoyment of music.

It seems clear that adults spontaneously employ the principles of tonality when they perceive melodic structures. But we should not assume that they spontaneously use that same tonal knowledge when they invent or compose even simple melodies. Indeed, except for the work of Bamberger,

we know very little empirically about how the perception of tonality is employed in the construction of tonal musical structures.

Bamberger (1986) finds that children construct tonal melodies on a set of bells in very different ways. For example, when constructing melodies from a chromatic collection of bells, some subjects physically separate the bells with pitches which do not 'fit' the tune they are attempting to play from those bells that sound notes in the tune. In doing so, they place the bells in the order in which they are played in the melody. Performance of the melody then means playing the bells in order from left to right. Another strategy these subjects use when making a tune is to select one bell for each sound of the melody. For example, the first three notes of 'Row, Row, Row Your Boat' would require three different bells using this strategy, even though all three notes are the same and could be played on a single bell.

In contast to those subjects, subjects of another group first position the bells to make a scale, then play the tune in an action path which maps the sequence of pitches of the melody on to the pitches of the scale pattern. Similarly, these same subjects simply restrike one bell to produce repetitions of pitches. Bamberger suggests that these very different strategies for finding the notes for an already known tune reflect the different levels of musical knowledge held by novice and gifted children.

## Domains as arenas for solving problems

Psychological research makes it clear that developmental studies of skill and cognition have much to offer for developing a more articulated understanding of a domain. Differences between the responses of subjects with different levels of ability provide windows on the subskills and foundations of specific skills involved in relatively simple tasks. For example, looking at the difference between novice and expert chess players has told us much about the nature of the knowledge required for the game, its representation, and how chess players use that representation to process new bits of information, make decisions, and orient themselves to subtle differences in patterns of information (DeGroot 1965, 1966). The tools of cognitive psychology and artificial intelligence have been used to look at other domains as well (Simon 1979).

The information gained in these studies is due in part to the use of protocols and protocol analysis and in part to the design of the questions guiding the interpretation of the protocols (Simon 1979). While this paradigm is not traditionally experimental, it has been extraordinarily useful in unravelling various aspects of cognition—for example, the creative process (Perkins 1981). Perkins observed poets in the process of

writing poems, and he highlights the role of everyday perception, reflection, and production in the creative process.

Combining the approach taken by Perkins, and including a developmental component, we can define and unravel still more of the mysteries of tonality. A great deal can be learned by studying musicians in the process of writing a modulating melody. Comparing musicians who have a sophisticated knowledge of tonality with those who do not can shed light on the underlying constructs of tonality. A musically and psychologically interesting task, it provides an opportunity to see how novices and experts think in a key, weaken one tonal centre, and strengthen another.

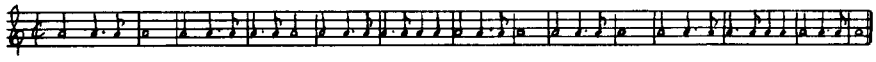
In addition to identifying developmental profiles, we hoped to be able to identify various problem-solving strategies. One might expect to find strategies in musical composition similar to those reported in studies of general problem solving (Johnson-Laird and Wason 1977), especially since the use of 'means-ends analysis' (breaking the larger composition problem into a series of subproblems) is routinely taught as a strategy in harmony and counterpoint classes. One might anticipate other heuristic procedures for compositional problem solving as well.

We also wondered about the prerequisites for the tonal understanding necessary for a simple composing task. Is the ability to play an instrument sufficient to enable one to compose melodic shapes which reflect tonal knowledge? Does the tonal knowledge acquired with advanced performance ability transfer to the composition of a tonal melody?

## Novice and expert solutions to a musical problem

We chose a composition task because it allowed us to observe the process of constructing tonal relationships. Many different ways of observing subjects composing modulating melodies were considered before the final version was designed. One approach considered was that of asking students who were studying composition to allow us to observe while they wrote melodies which modulated. This task would provide subjects who were accustomed to writing, but it would not allow us to control for individual differences which might arise between composition students over issues of metre, key, rhythm, length, phrase structure, or style. More importantly, it failed to recognize the possible biasing which personal writing habits might introduce into the task. For example, one composer might not be accustomed to writing tonal music at all while another might be acustomed to making melodies that were tonal but which moved only to certain keys and not others. For these reasons, we decided two things—first, not to include composers in the study and second, to control for variables which were not central to the issue of tonality. These variables included the length of the melody and the rhythmic durations to be used.

The rhythm of the melody was controlled by using a rhythmic structure modelled on the first section of the march by Schubert, Op. 27 No. 1. The rhythmic pattern, in 2/2 metre, was 14 measures long. The rhythmic shape allowed different interpretations of the phrase structure while providing the opportunity for constructing melodic motives. This rhythmic pattern



**Example 11.1**

created a 'shell' which the subjects filled in with notes they determined to be necessary to fulfil the requirements of the task.

It was also necessary to specify the nature of the modulation. If we left the modulation unspecified, we ran the risk of every one making a favourite and presumably well-practised key shift, or relying on their memory of a specific modulation from a specific piece as a model. Because we wanted to draw out the principles on which beginners and more experienced music students relied, we decided to ask for a relatively rare and unpractised modulation.

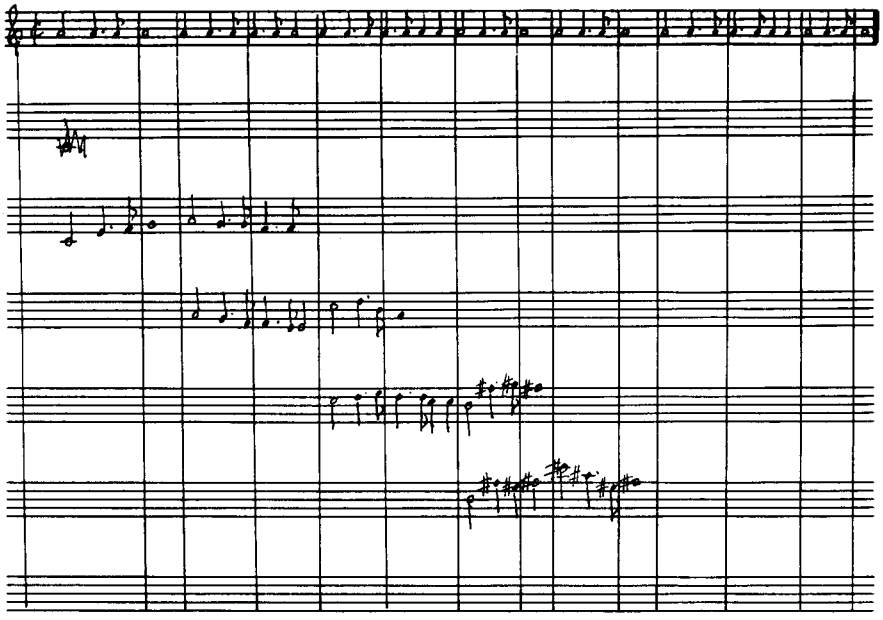
We controlled for the modulation by requesting a melody which would move from the tonic to the key a tritone away and back to the tonic. This provided three specific areas within the task where tonal thinking could be observed: the opening of the melody (to determine how the first key was established), the transition into the new key (to observe how the stability of the first key was undermined and the second key established), and finally, the ending of the melody (to determine how the return to the first key was negotiated). Interest was sharply focused on the areas in which the key was established or undermined. In this way, we could observe the ebb and flow of tonal support around specific points of reference.

## Conditions of the study

Ten music students took part in the study. They were all performers and were assigned to one of two groups. Five students were in their first year of conservatory study (we shall refer to this group as 'beginners'). Their average age was 20, and they had studied an instrument, on average, for six years. The second and more experienced group of students had an average age of 22 and all had undergone two years of training at the conservatory (we shall call these students 'experienced').

At the beginning of the session individual students were seated at a piano and given a sheet of 12-staff manuscript paper. The rhythmic pattern was written in the second space of the top staff. The remaining staves were sectioned into measures corresponding with those of the rhythm, but were

otherwise blank. All subjects were asked to 'use the rhythmic pattern at the top of the page and write a melody starting in the key of C major, passing through the key of F sharp major, and closing in C major.' A pen was provided and subjects were instructed to make all revisions in the measure immediately below their initial choice (see Example 11.2). They were requested not to cross out any of their work. The subjects were asked to talk aloud as they worked and to explain what they were considering at each step of the process; they were also instructed to use the piano whenever they wished. The task had a time limit of 30 minutes; subjects were not informed of this limit, but were stopped after 30 minutes had elapsed. The entire session was tape-recorded and an observer was present.



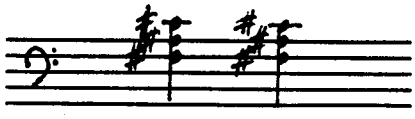
**Example 11.2**

Later, the second author made a transcription of each session which organized the student's activity in a chronological manner. The transcriptions were based upon the tape-recordings of the sessions, the notes of the observer, and the music manuscript pages on which the students notated their melodies and made their revisions.

Below is an excerpt from the transcript of a beginning student. The transcription excerpt begins just after this student has completed the first four measures of his melody. He has written those measures in C major.

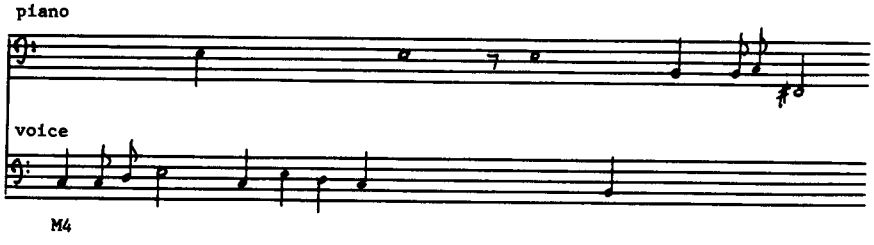
When the student replays material already written down it is designated beneath the music as M1, M2, etc., for measure 1, measure 1, and so forth. In all examples of playing or singing, the rhythmic notation is proportional and reflects only relative durations.

*Writes measure 4*
*Plays:*



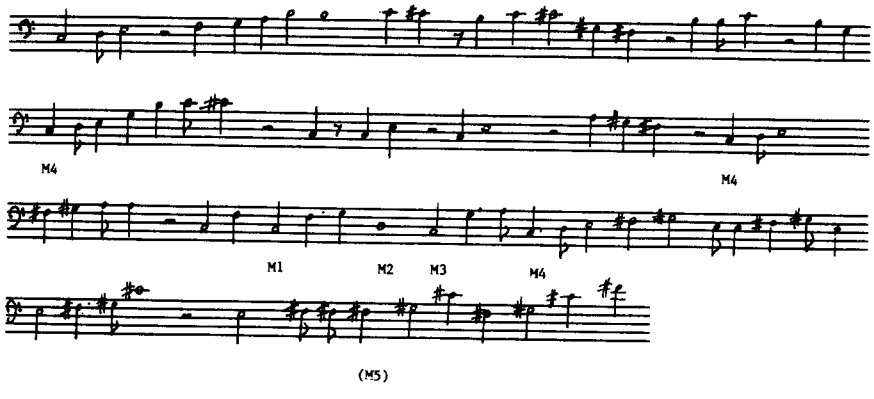
**Example 11.3**

*11 seconds silence*
*Plays and hums:*




**Example 11.4**

*37 seconds silence*
'Um, now I'm trying to get into F sharp somehow, smoothly.'
*Plays:*




**Example 11.5**

*Writes measure 5*

The sessions were analysed on the basis of categories which took into account several aspects of the composition process. These categories focused upon working procedures, rhythmic and tonal considerations, and the overall success of the activity and its final melody. The specific categories were:

(A) Size of the working unit
(B) Generation of new melodic material
(C) Integration of rhythm into the working process
(D) Use of motivic relationships
(E) Orientation to the first key
(F) Transition to the new key
(G) Return to the home key
(H) Efficiency of the modulation process
(I) How a key is defined
(J) Strategies
(K) Success of melody

The analysis was carried out by the two authors independently; a high degree of reliability was achieved. For measure (J), strategies, the transcripts were examined for evidence of strategies and both scorers had to agree on the specific strategy being used. For measure (K), success of melody, seven musicians rated the melodies independently. The definitions of each analytic category and the findings for each category are presented in the results section which follows.

## Results

In this portion of the paper, we first consider general features of the working style. These features are the size of the working unit, sequence and manner in which melodic material is generated, the role of the rhythm in the process of constructing the melody, and the degree to which motivic relationships are used in the melody. After we examine the profile of working styles, we consider the issues of tonality. The tonality measures discussed include the orientation to the first key (C major), the location of the cadences (in C and F sharp), the location of the first indication of key change (in C and F sharp), and how tonalities are defined while composing this melody. Finally, after the discussion of strategies used while composing the melody, we consider the transfer of knowledge of tonality acquired through performance to the compositional task.

### General working style

*Size of the working unit*

One of the findings from the 'novice-expert' literature suggests that the

amount of organized information at an individual's disposal is dependent, at least in part, on the level of that individual's experience (DeGroot 1965, 1966). The key word is 'organized'. The expert sees patterns of organized information where the novice sees only unrelated details. In music composition, the amount of information an individual works with at any single moment can also be an index of the level of development. Presumably, the novice will be able to work with only the immediate context or next note, while the more developmentally advanced musician will be able to chunk information into larger units and to work more efficiently—projecting further ahead and reflecting further back while at the same time holding the moment of focus in place. This makes it possible for the expert to consider immediate contextual issues and the degree to which a specific instance is forecast by earlier moments of the composition, and to develop a higher degree of motivic organization than the novice would be able to achieve. Consequently, we kept track of the size of the frame of reference within which individuals worked. We refer to this as the 'size of the working unit', which measures the area to which a subject made either verbal or behavioural reference during the task. There are five levels in this measure: (1) note-to-note, (2) measure-to-measure, (3) two-measure chunks, (4) three-measure chunks, and (5) whole phrases.

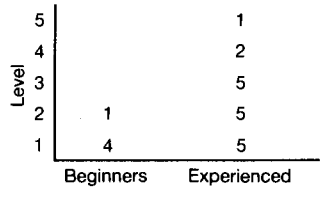



**Fig. 11.1.** Size of the working unit.

The size of the working unit proved to be a useful measure for distinguishing among the subjects. The scores reflect the magnitude of the differences between the first-year students and those in the more experienced group. Four of the five beginners used only the note as a working unit, while the third-year students all used two entire measures as a working unit at least once during the session. Furthermore, the more experienced subjects used much more variation in the size of the unit they employed when they constructed their melodies. At some time during the task, all of the third-year students used the next note, and one- and two-measure chunks. In addition, two of these subjects once used a three-measure unit, and one spoke in terms of an entire phrase. In contrast, the first-year students were less flexible in their ability to look ahead and to use units of various sizes. Specifically, four of the beginners usually looked

only to the next note. One was able to occasionally use a whole measure as a unit of focus.

*Generation of new material*

The order or sequence in which a mature composer might work on a piece is largely unexplored. Obviously, one can work in many different ways. One might work on different places at once, shuttling back and forth between one area and another throughout the compositional process; another might systematically write in a front-to-back order. There is no reason to expect that one procedure is superior to another. However, one might expect to see evidence of more back-and-forth movement (more flexibility) in the experienced students, but more beginning-to-end movement (less flexibility) in the beginning students.

When we looked at the sequence in which the new material was generated we saw two different patterns. The beginners normally worked in a left-to-right order, usually playing back only within a local context. The experienced students also worked left-to-right, but when they played back their results they were careful to connect the new material to larger chunks—the previous measure or phrase—or even to the beginning.

How composers generate new ideas or the next chunk of music is also largely unexplored. Sloboda's analysis of his own compositional process is one of the few attempts to look systematically at this important issue (Sloboda 1985). We did observe the first-year students generating new ideas, but the processes of the experienced students proved to be more elusive.

The first-year students' reliance on an enactive, sensory-motor mode of generating new melodic material was striking. They would often begin by playing or writing out a melodic shape and then get stuck. Once stuck, they would abandon their plan and start to explore ways of getting from the note they last wrote down to the next one of the rhythmic pattern. Remarkably, they often did this by letting their fingers sound the notes between two goal pitches, sometimes slowly playing all the semitones between. It is immediately striking that *all* of the first-year students, but *none* of the third-year students, let their fingers lead in this way when searching for the next note or bit of melody.

It was more difficult to determine how the experienced students were generating new material. They would more often consider the next part of the rhythmic pattern, and then write out or play a fully developed bit of melody. Clearly, they were able to consider the contextual constraints and generate pitch shapes and integrate them into the rhythmic pattern, as well as reflect on the effect without committing themselves to a solution.

*Integration of rhythm into the working process*

The ability to integrate the given rhythm into the compositional process

proved to be another useful index of difference between the two groups. When the beginners were engaged in the process of generating melody, they ignored the given rhythm entirely. On the other hand, all of the more experienced students used the given rhythm in the compositional process. In addition, when playing the completed melody at the end of the task, none of the beginners played the pitches consistently in rhythm, while the experienced students always played their melodies back in rhythm.

*Use of motivic relationships*

The rhythmic shape given as a control in this study suggests many possible motives and therefore allows the possibility of establishing and strengthening the formal relationship between the opening and the ending of the piece. The use of melodic motives is not only a powerful means of achieving structural unity and contrast, but their use also reflects the ability to think in larger structural chunks.

One of the beginners and four of the experienced students used the rhythmic pattern as a source of motivic material.

## How a tonality is defined

The centre of the study was tonality. Measures of tonal thinking included the orientation to the first key, transition to the new key, return to the home key, the relative efficiency of the modulation process, and how a key was defined. Other topics included the degree of difficulty in changing key and the way in which notes were used to define a key in the introduction as well as during the composition of the melody.

*Orientation to first key*

During the first moments of the task, students were in the process of defining the problem for themselves; trying out different ways of establishing the first key and testing the distance to the second key. Five behaviours occurred during the opening:

(1) Establishing the key by repeating the tonic note of the scale;
(2) Establishing the key by extending the duration of the tonic note of the scale;
(3) Establishing the key by using tonal formula, for example, scale degree motion from 7 to 1;
(4) Establishing the key by using scalar motion;
(5) Establishing the key by using the notes of the scale to suggest long-range tonal relationships, specifically by starting on a weaker portion of the scale, thereby preparing a smoother transition to the new key.

There was a sharp distinction between the two groups on this measure.

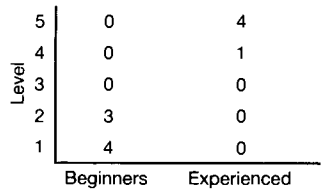



**Fig. 11.2.** Orientation to the first key.

*Entrance of the modulation*

Each student had to determine from the given rhythmic structure where critical tonal shifts had to be made in order to meet the modulation requirements of the task. The location of first indication of modulation is important because it provides an index of how the problem is interpreted, how the boundaries of key are determined, and how the moment of modulation is defined.

Six of the 10 students (three of the beginners and three of the experienced students) placed their first accidental in measure 4; no one used an accidental before measure 4. Four students placed the first accidental of their modulations after measure 4.

Most students tended to make the first move of the modulation in the fourth measure. This suggests they recognized the traditional phrase structure implicit in the rhythmic pattern.

*Location of cadence in F sharp*

The location of the cadence in F sharp is critical. If it occurs too early, the second key is too prominent. On the other hand, if it appears too late in the melody there is not enough time to get back smoothly to the home key of C major. The rhythmic structure suggests a cadence would be best placed at measure 8. All of the beginning students made a cadence in F sharp at measure 8. Three of the experienced students also made the cadence in this measure.

*Return to the home tonic and the ending*

This measurement has three levels. Level one places the moment of modulation at bar 8, and level two places the moment of modulation at bar 10, while level three indicates the modulation took place at some other point (see Fig. 11.3).

**Economy when making a modulation**

We expected that students who did not have strong and flexible representations of tonal structures would find it difficult to make a

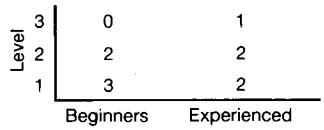



**Fig. 11.3.** Return to the home tonic.

modulation. They would not be able to easily and quickly create the tonal shift necessary to effect a new key. Consequently, we anticipated that the first-year students would use many more trial-and-error attempts to make a modulation than the more experienced subjects. We measured the students' efficiency by counting the number of notes played between the first written accidental and the cadence in the new key. There was a significant difference between the two groups of students.

The first transition was defined as the area between the first written accidental in the first section (C major) and the cadence goal on the key of F sharp major. The beginners played or hummed an average of 232 notes while composing this portion of their melody, while the experienced students averaged only 72 notes.

The second transition was defined as the area between the cadence on F sharp and the clear statement of C major at the end of the piece. Beginners sounded an average of 132 notes in the course of writing this section, in contrast to an average of 92 notes for the experienced students. The first-year students had great difficulty negotiating the shift of keys from C major to F sharp major, and somewhat less difficulty moving from F sharp major back to C major. Interestingly, the experienced students had somewhat more difficulty returning to C major than moving away from it.

### Definition of key

There are five levels in this measurement. These levels chart the distance between between treating notes as mere collections of pitches to simultaneously considering the tonal value of each note in relation to both keys required for the task.

Level (1) Collection of notes without any specific order; key is defined by an aggregate of pitches common to the scale but not necessarily used in a way that suggested a tonal framework.

Level (2) Key defined only by local emphasis on the tonic of the scale; other notes of the scale are ignored. Focus is created by the (frequency of occurrence or) repetition of the tonic note, or by the duration of the tonic triad.

Level (3) Key defined by the use of a scheme (any of several functional chunks such as 'leading tone' function, or triad shape). This level is an

extension of level (2), in which the knowledge of a triad or a scale degree function is fixed, for example, scale degree 7 is transposed as a unit with sd 1 (7–1 in C, then 7–1 in F sharp), or as a triad (1–3–5 in C [C–E–G], then in F sharp [F sharp–A sharp–C sharp]).

Level (4) Key defined by using specific transferable relationships within the context of different keys. Specifically, this includes the use of dominant-to-tonic relationships, and the use of pitches as part of small scalar fragments. This reflects an understanding of the relation of pitches to a local tonic.

Level (5) This level shows the use of single pitches to suggest different scales. A pitch may be chosen because of its relation to the local tonic *and* the target or goal tonic. A common example of this level is to start the melody on a weak scale degree, thereby heightening the likelihood of a smooth transition to the second key. Pitches are chosen which can suggest the local tonic while forecasting other tonics, scales, and scale degree functions, for example using sd 6 (A) as a starting point because it can also function as sd 3 in F sharp minor.

Figure 11.4 shows the frequency of use for each level by each group.

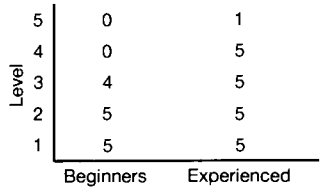



**Fig. 11.4.** Definition of key.

To summarize the results of looking at working styles, clearly, there are differences between groups. It is possible to show contrasting profiles of the two groups on the basis of the differences between them on a variety of measures. The groups differed significantly in the number of notes which they used to make the transition between the two keys. There were also differences in the ability to separate the identity of a note when it had a specific function, for example, a note as a leading tone, and the function of 'leading tone' in general. The beginning students had difficulty finding the leading tone of F sharp major because they tended to confuse it with the leading tone of the scale of C major. Similarly, none of the beginning students were able to use the note 'B' to mediate between the two scales (even though it is the single common tone between the two keys). In contrast, the members of the more experienced group were easily able to find and use the common tone as a means of entering or leaving either key.

## Strategies

The differences in working styles, defined by the size of working unit, the work sequence, and the integration of rhythm and pitches when composing the melody, were supported by a wide range of strategies. Several strategies have family resemblances which enable us to group them into categories.

The strategies can be arranged to form a three-step scale from the least sophisticated (level (1)) to the most sophisticated (level (3)). The less advanced strategies tend to rely only on local features or single, uncoordinated attributes, while more advanced strategies require the ability to consider moments of the task from the perspective of the structured whole.

*First-level strategies*

The first level contains the strategies used by the five beginning students. These are the most local of all the strategies observed, and they tend to focus on single attributes. They include the use of repetition of a single note, a heavy reliance on the use of sensory-motor or finger knowledge when finding new notes, and the use of a simple 'next-to' rule which defines the extent the beginners look ahead or back when they co-ordinate the notes on which they are currently working. That is, the beginners tended to compare what they were working on with only the immediate notes rather than phrase-length units.

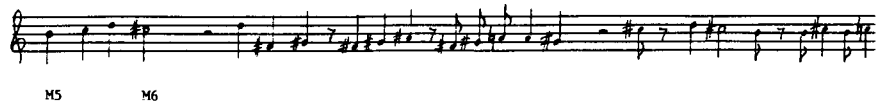
One striking characteristic of this level of strategy is the reduction of the working unit to the immediate context. The field of vision and hearing appears to be restricted to the notes on either side of the one(s) being worked on at the moment. It is especially difficult to achieve a high degree of integration when working in this bit-to-bit manner. Only by ignoring the rhythmic durations of notes is it possible to weigh each note in relative isolation. Because each note occupies the same amount of attention, the effect of shorter and longer durations on the tonal weight a pitch may receive is either not realized or not taken into account during the composition process. It is easy to observe that the beginners fail to understand the tonal implications of the given rhythmic series.

Sounding a note and repeating it is an early means of either establishing a tonally weighted note, such as the tonic note of a scale, or of filling up a space of musical time. Another beginning strategy is the reliance on whatever falls under the fingers when generating new material. The uncritical acceptance of melodic shapes which appear to grow out of finger motion alone, and which appear to be unconnected to whatever has been generated during the previous chunk, is surprising.

The transcript excerpt (Examples 11.6–12) shows many characteristic behaviours of the beginners. Most important is the reliance on the finger

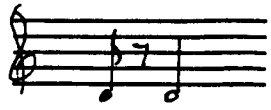
knowledge to find new notes of the melody. It also shows the small unit within which the beginners work, and the comment at the end suggests the importance of the number of occurrences of F sharp as a means of defining the F sharp major tonality. This student has just completed writing measure 6 and is starting to explore possible ways of continuing the melody into measure 7.

*Plays:*



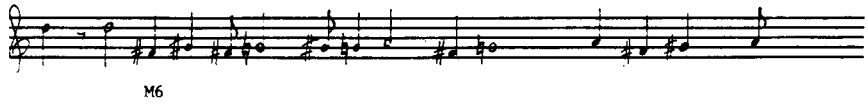

**Example 11.6**

*Hums:*



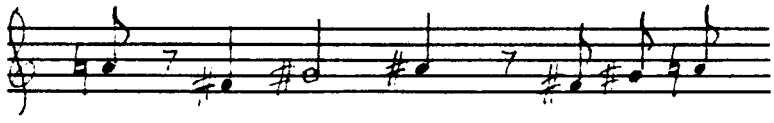
**Example 11.7**

*8 seconds silence*
*Plays:*



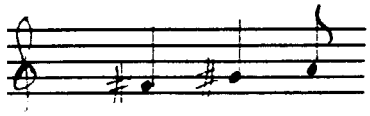

**Example 11.8**

*18 seconds silence*
*Plays:*



**Example 11.9**

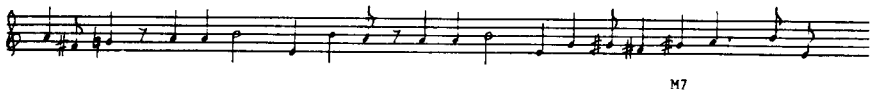
*12 seconds silence*
*Plays:*



**Example 11.10**

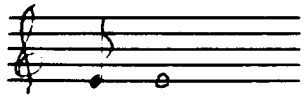
*9 seconds silence*

*Writes first part of measure 7*
*Plays as she continues to write measures 7 and 8:*




**Example 11.11**

'Got to get back to C again?' *(meaning C major)*
*Observer:* 'Yes.'
*Hums:*



**Example 11.12**

'Only two F sharps'. Gotta get more than two out of this! OK?'

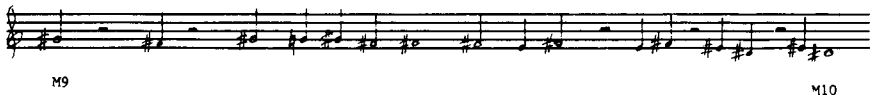
*Second-level strategies*

Three of the experienced students used second-level strategies which feature various local rules. Unlike level (1), however, this level prescribes the use of chunks of material which are larger than individual notes. These chunks consist either of portions or scales or of triads.

The least-developed aspect of this level is the use of a descending semitone motion from a goal pitch in one measure to a goal pitch in another measure. This is similar to the sensory motor exploration of the first level. The difference between the two strategies is that the sensory motor strategy appears to have no clearly directed goal, while the descending semi-tone motion has specific boundaries forming its top and bottom pitches. Thus there is a feeling of intentional design guiding the use of this strategy.

In the example below, the G sharp in measure 9 forms the top pitch and D sharp at the end of excerpt forms the final pitch of what becomes measure 10. This experienced student plays the entire completed portion of the melody before he explores the next chunk.

*Plays:*



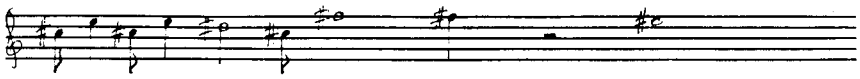

**Example 11.13**

*6 seconds silence*
*Writes measure 11*

A slightly more elaborate strategy is the use of a specific tonal function such as the leading-tone-to-tonic motion. Successful use of this step shows that the subject can easily locate the critical scale degrees (7 and 1) in C major *and* in F sharp major. These are usually co-ordinated with a specific rhythmic motive as well, knitting the whole into a coherent melodic shape.

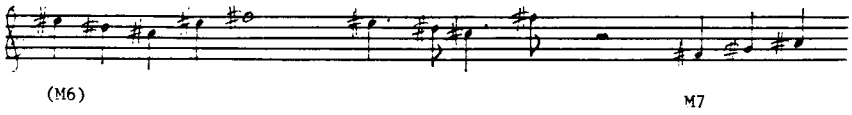
The most advanced of these local rules uses fragments of scales to introduce the modulation. The next example shows the use of a 'leading-tone' strategy to approach the new key and the use of scale fragments (the upper tetrachord of F sharp major: C sharp, D sharp, E sharp, and F sharp).

*Plays:*



**Example 11.14**

*7 seconds silence*
'E sharp is enharmonic for F natural. OK, I'll go from there maybe.'
*Plays:*




**Example 11.15**

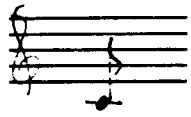
*20-seconds silence*
'OK, this is what I'll do, you see?'
*Writes measure 6 and 7*

*Third-level strategies*

Two of the experienced students used strategies which were very different from those of the other students. Students at this level easily talk about their plans and strategies, and often forecast the success of an expected solution—all before sounding a note. These students were able to balance the focal and the subsidiary aspects of the task and integrate the particular shapes into a coherent and whole melody. For example, students at this level looked at the rhythmic shape, thought about the two keys, and then stated a specific plan which would take the melody from C major to F sharp major, for example, through the intermediary key of A major.

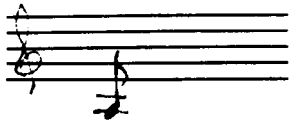
'So this is just the rhythm?' (*points to the page*)
*Observer*: 'Yes, that's the rhythm'.
'It's just the melody?'
*Observer:* 'Right.'

*14 seconds silence*
*Writes in the clef*
*Plays:*



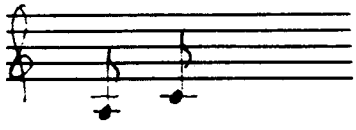
**Example 11.16**

*5 seconds silence*
*Plays:*



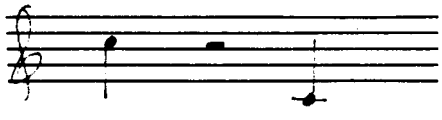
**Example 11.17**

*4 seconds silence*
*Plays:*



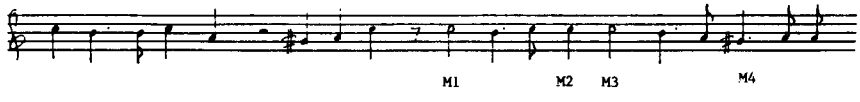
**Example 11.18**

*3 seconds silence*
*Plays:*



**Example 11.19**

*8 seconds silence*
*Plays:*




**Example 11.20**

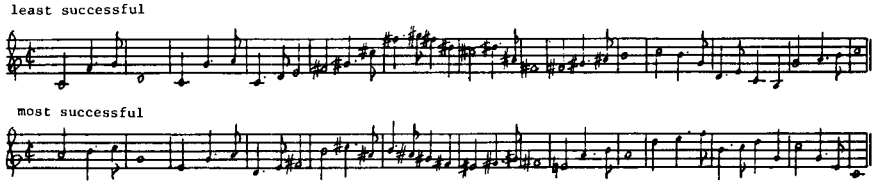
*39 seconds silence in which she writes measures 1, 2, 3, and 4*
'OK. What I'm going to do is go from C major to a minor to, um, F sharp major.'

Characteristically, individual decisions at the note-by-note level are assessed in terms of their overall effect on the design. These students' ability to maintain the most global perspective while considering the local situation is evident from the use of a hypothesis-testing model throughout the task. The model appears to have four steps, *hypotheses*, *test*, *try out in context*, and *evaluate* (accept or reject). The similarity to the TOTE (Test–Operate–Test–Exit) model described by Miller, *et al.* (1960) is obvious.

### Success of melody

Is it possible that composing a successful melody is the result of pure inspiration? If this is the case, one might expect to find little correlation among the factors of experience and the most successful of these melodies. On the other hand, to the extent that writing a successful tonal melody is dependent on rules of composition, one can expect the more experienced students to write more successful melodies. In an attempt to consider these questions, we asked a group of seven professional musicians to judge the melodies. Some of the musicians were college teachers while others were freelance musicians; all had a great deal of performing experience. We asked them to place the melodies in rank order from most to least successful. The judges were not told about the students or the study, and the melodies were presented in random order. The results are interesting because they shed more light on the role of formal knowledge in music. The resulting rank order of the melodies showed a high correlation with the experience of the student (Spearman Rho, $r = 0.86$). This means that the more sophisticated the strategies used in the process of composing the melody, the more likely judges were to consider it a successful melody. This suggests that experience, including theoretical knowledge of tonality, does indeed play a role in the creation of a more successful melody.

The melodies below were judged the least successful and the most successful respectively.




**Example 11.21**

## Summary and discussion

Looking at the working styles and the strategies employed by members of the two groups suggests that a developmental stance combined with a problem-

finding/problem-solving approach is a very useful way of investigating musical questions. Examining how a task is represented, and determining the role a representation plays in forming strategies, operations, and definitions of goals, sheds light on even such a simple task as adding notes to a rhythmic pattern.

**Representation of the problem**

The beginning and experienced music students' representations of the task were very different because of their distinctive knowledge states. The dimensions of their different knowledge states are seen in the particular ways members of the two groups defined keys and integrate rhythm into the composing process.

At a verbal level all the students appeared to have the problem well defined. For example, the beginning and experienced students understood the keys to be used as C major and F sharp major without difficulty. However, the meaning supporting the text was very different for the two groups. The beginning students defined keys with single notes (usually the tonic) or the tonic triad. This suggests their schema for key is limited to rather simple and static units. The notes of these units are not hierarchically integrated with other notes of the scale, but simply exist as part of a familiar collection of tones. In contrast, the experienced students relied on more dynamic and structured tonal schemata which supported the hierarchical relationships implicit in scales and scalar fragments. They were able to 'lift out' overlapping portions of different scales, co-ordinating and integrating them into a tonal blend which linked the two.

Similarly, the difference rhythm played in the representation of the task was striking. Although the rhythm was specified for everyone, only the third-year students used it when constructing their melodies. The beginning students appeared not to include rhythm in their representation of the task, since they focused exclusively on the pitches of the melody. The beginners even ignored the rhythm when playing back their final version, as if slowly testing each note as they played the entire melody through. Again in contrast to the beginners, the experienced students used the rhythm at every stage of the task, making it clear that their concept of melody included rhythm *as well as* pitch.

There was a striking difference between the groups in their choice of when to edit the melody. Members of the third-year group characteristically edited measures and phrases of the melody immediately, while they were composing. This reflects their ability to work on precise moments within a larger framework as evidenced by their not needing to listen when they had finished working and not wanting to make revisions once they had completed the melody. This was very different from the editing behaviour of the beginning students. Characteristically, they did not assess the effect of

their melodic accretions until the final performance of the melody. They usually played the melody the last time as though they were listening carefully to the effect of the whole shape. After that performance, they would often pick up the pen again and make 'corrections'. Usually, only a few notes were changed and, confident of their work, they usually did not perform the entire melody a second time to check their revisions.

In summary, members of the two groups see the task differently. Beginners do not have many ways to define a tonality, whereas the more sophisticated musicians employ very different ways of establishing key centres. The beginners fail to see rhythm as a central or crucial factor in constructing a tonality or designing a tonal melody. This leads them to an unsatisfactory definition of key, because they fail to realize the importance of placing more important notes of the key on longer and/or more accented moments of the rhythmic series. In sharp contrast, experienced subjects were keenly aware of the role rhythm plays in contributing to the tonal weight of a note, both during the composition process and during the playing of the final melody.

Defining the limits of the problem of making a modulation around the use of appropriate triads is very different from defining the problem as involving the use of scales. While the triad definition has the virtue of being simple, that focus tends to make melodies which are abrupt in their shifts. The scale definition does not place a limit on the nature of melodic materials (one can use triads and stepwise scalar motions), and at the same time allows abrupt as well as smooth moments of transition. Seeing the problem of establishing a key as requiring units of two or three notes of the scale (for example, scale degrees 7 and 1, or a triad) is very different from interpreting it as based on a concept of a collection of seven pitches which exist in specific hierarchical tension.

Clearly, having alternative representations of the problem is an advantage. It was possible to see the task essentially as a problem of linking tonic triads, making a harmonic progression, or locating a specific tonic note, scalar fragment, or scale. The beginners characteristically had fewer alternative representations, while the experienced subjects not only had a wider range of representations, but were able to move among them, selecting an approach which seemed appropriate for a given moment.

## Representation of the problem and strategies employed

Each representation of the problem defines and limits the range of operations used to arrive at solutions. For the problem solver, the formulation of a plan of action or strategy depends on how the problem is understood and represented. For the researcher, understanding the dimensions within which the problem solver is working makes it possible to

observe and evaluate the set of actions or operations which the problem solver hopes will lead to a solution of the problem.

The task we presented was open ended in so far as the moment of modulation was not defined. An algorithmic or rule-governed approach, even though it might lead to a solution, was not used by anyone. On the other hand, there were a wide variety of heuristic approaches to the problem. The most local strategies included the use of finger patterns and semitone chromatic pitch motion, the identification of a specific tonal relation such as the leading tone, and the use of a tone common to both keys as a pivot point. More advanced strategies included the use of melodic shapes, contours, and motives (sometimes in melodic sequence), as well as the hypothesis–test–incorporate–evaluate model. Finally, the most advanced and long-range strategies included the hypothesis-testing operation in conjunction with verbal statement of means, and the ability to plan future areas of focus.

Characteristically, the beginners were unable to design or construct a strategy except at the most local level. Consequently, they usually added notes to their melody without a long-range plan or design in mind. If they began with a plan, it was usually only temporary since they were often unable to transform it into a flexible heuristic attitude. Because of this, their attempts either looked as though there was no plan at all, or like perseveration, since they would use the same unsuccessful strategy over and over. In contrast, the most experienced students were able to use means-ends analysis, and even use a generate-test model.

### Two levels of tonal thinking: enactive vs. reflective processes

First-year students, the beginners, were generally unable to reflect on the task without enactively sounding out each step of a solution. Their reliance on their fingers when finding the next notes of their melody was particularly striking. Their knowledge of tonality appeared to have not yet been integrated into a flexible working system. They were aware of individual notes or triads but not able to use them in larger functional contexts. Their goals emerged as they worked. They were in constant conflict between an explicit local context within which they were working and the implicit global context imposed by the task definition. It was symptomatic of this conflict that they were unable to reflect on and evaluate the focal moment in terms of the whole product.

These students also tended to work from front to back in note-to-note or one-measure units and with little regard for the rhythmic shape of the melody either during construction or during the playback at the end of the session. They were unable to talk about their strategies or goals until after they had finished solving a specific problem. They were unable to assess the effect of choices they made during the process of writing the melody. They

also had trouble keeping the context in focus while they were making additional bits of melody. Similarly, they seemed to have a limited awareness of the final shape of their melody since they were frequently surprised when they played their melody at the end of the session. They used one of three strategies: using the leading-tone-to-tonic relationship as a means of entering a key; using the root-position triad to either stand for a key or as a reference point for a key; and relying on finger movement over the keyboard as a means of exploring for new notes and expanding the melody.

Third-year students, the experienced group, were able to reflect on the dimensions of the problem and solve much of it before they committed themselves to a specific version on paper. Two of them had to be requested to play their final melodies. Their tonal knowledge was integrated into a rich and flexible working system. They had an understanding of tonality which included awareness of how scale degrees functioned and how even distant key areas could be related. They were able to define their methods and their goals in advance of starting to work on them. They were able to work on a local problem while evaluating it in terms of the whole.

Like the beginning students, they also worked from front to back, but the experienced students more often worked in larger chunks (normally one-, two-, and three-measure units). They were able to talk about their strategies and goals in advance of the immediate working context. In addition, they quickly integrated the pitches into the rhythmic pattern provided at the top of the page. Since they used significantly fewer notes (and relied less on trial and error) they appeared to be relatively sure of the effect of their choices while they were constructing the melody. When they played their melody at the end of the session, they appeared to be satisfied with the result. This suggests they already knew how it was going to sound. They were able to plan an approach to the problem and adjust it to changing contexts as they worked through the task. Significantly, they were able to make sections of melody without losing the context of the whole melodic shape. Two in this group used a clear hypothesis-testing model when they worked. Specifically, they would generate a melodic fragment and test it out in isolation of its context. If it satisfied their criteria, they would integrate it into the already composed section by playing it and evaluating it. On the basis of its effect within the context of the whole shape, they would either accept it and integrate it into the melodic line, or throw it out and start again.

One final question remains. On the basis of this study, what is the relation between perceptual knowledge of tonality and the knowledge of tonality required for production? The results of this study strongly suggest that the composition of a tonal melody requires a qualitatively different understanding of tonality than does the perception of tonality. Surprisingly, it appears that even the years of practice involved in studying an instrument

fail to provide the breadth and depth of knowledge required for a tonal composition task like the one in this study. It seems that problem solving in a domain like music is fundamentally different from problem solving in a domain like chess. In chess, one is either a novice or an expert. In music, unlike chess, one can be both expert and novice at the same time, depending on whether the task is a performance task or a composition task, a perception task or a production task. The musicians in this study were all gifted performers. Only the older group approached the level of tonal knowledge which was necessary for the successful solution of the melody-writing task. That group had received two additional years of intensive training in the charged musical environment of a conservatory.

In light of this, we see there is a great deal of difference in the knowledge which supports each of the working styles described in the opening of this chapter, especially when we are considering less than fully mature composers. While a master composer may appear at times like a beginner, for example, proceeding on a note-to-note level or making revisions after a 'final' playing, on the basis of this study it is likely that the master composer has a much richer representation of the task and has strategic procedures which are more flexible than perhaps even the relatively advanced student. More than that, the results of this study indicate some of the milestones occurring along the developmental path which leads from beginning to advanced levels of tonal knowledge. We see that the definition of a task like writing a melody is very different for musicians at different levels of development. We see that these differences play an important role not only in the definition and goals of the task, but also in the strategies employed and in the quality of the results.

## Acknowledgements

This research was supported in part by grants from the Carnegie Corporation, Rockefeller Foundation, and the Spencer Foundation. The authors wish to thank members of the Thursday Night Group and especially Erin Phelps and Joan Meyaard for their assistance.

## References

Bamberger, J. (1986). Cognitive issues in the development of musically gifted children. In *Conceptions of giftedness* (ed. R. J. Sternberg and J. E. Davidson), pp. 388–413. Cambridge University Press, Cambridge.

Bartlett, F. C. (1958). *Thinking: an experimental and social study*. Basic Books, New York.

DeGroot, A. D. (1965). *Thought and choice in chess*. Mouton, The Hague.

DeGroot, A. D. (1966). Perception and memory versus thought: some old ideas and recent findings. *Problem solving: research, method and theory* (ed. B. Kleinmunte). Wiley, New York.

Cuddy, L. L. and Lyons, H. I. (1981). Musical pattern recognition: a comparison of listening to and studying tonal structures and tonal ambiguities. *Psychomusicology* **1**(2), 15–33.

Dowling, W. J. and Fujitani, D. S. (1971). Contour, interval and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America* **49**, 524–31.

Frances, R. (1958). *La perception de la musique*, pp. 81–90. Vrin, Paris.

Johnson-Laird, P. N. and Wason, P. C. (1977). *Thinking: readings in cognitive science*. Cambridge University Press, Cambridge.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology* **11**, 346–74.

Krumhansl, C. L., Bharucha, J., and Castellano, M. A. (1982). Key distance effects on perceived harmonic structure in music. *Perception and Psychophysics* **31**, 75–85.

Miller, G. A., Galanter, E., and Pribram, K. (1960). *Plans and the structure of behavior*. Holt, Rinehart and Winston, New York.

Newell, A., and Simon, H. A. (1972). *Human problem solving*. Prentice-Hall, Englewood Cliffs, NJ.

Perkins, D. N. (1981). *The mind's best work*. Harvard University Press, Cambridge, Mass.

Simon, H. A. (1979). *Models of thought*. Yale University Press, New Haven.

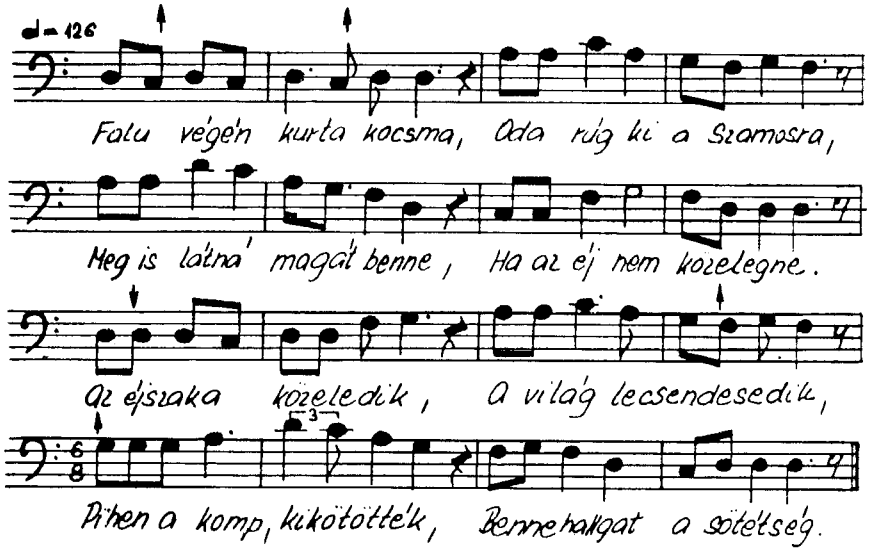
Sloboda, J. (1985). *The musical mind: the cognitive psychology of music*. Clarendon press, Oxford

Sloboda, J. and Edworthy, J. (1981). Attending to two melodies at once: the effect of key relatedness. *Psychology of Music* **9**, 39–43.

Zenatti, A. (1975). Melodic memory tests: a comparison of normal children and mental defectives. *Journal of Research in Music Education* **23**, 41–52.

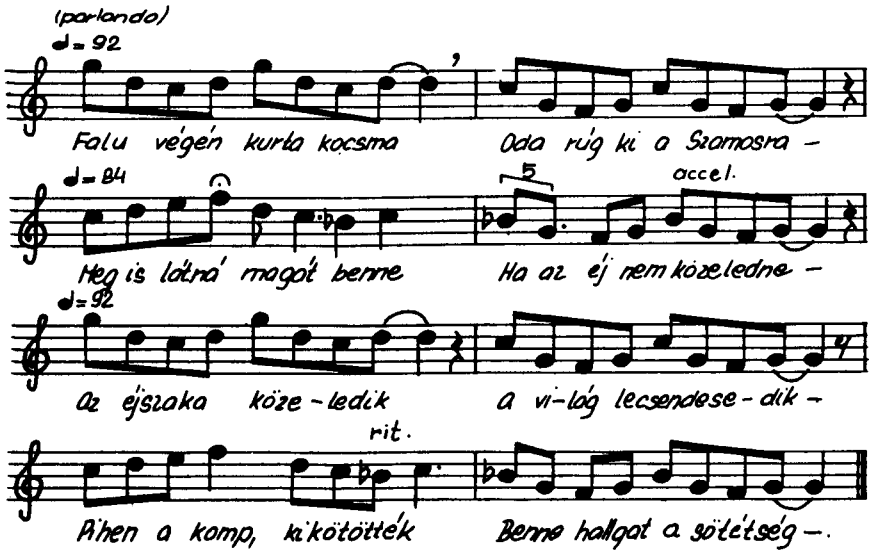
# Appendix (Chapter 8)

**Example 8.11.** Conservatory student, 24 years old.




Form: A $A_v^5$ B C
A $A_v^5$ $B_v$ $C_v$
Tonality: pentatonic

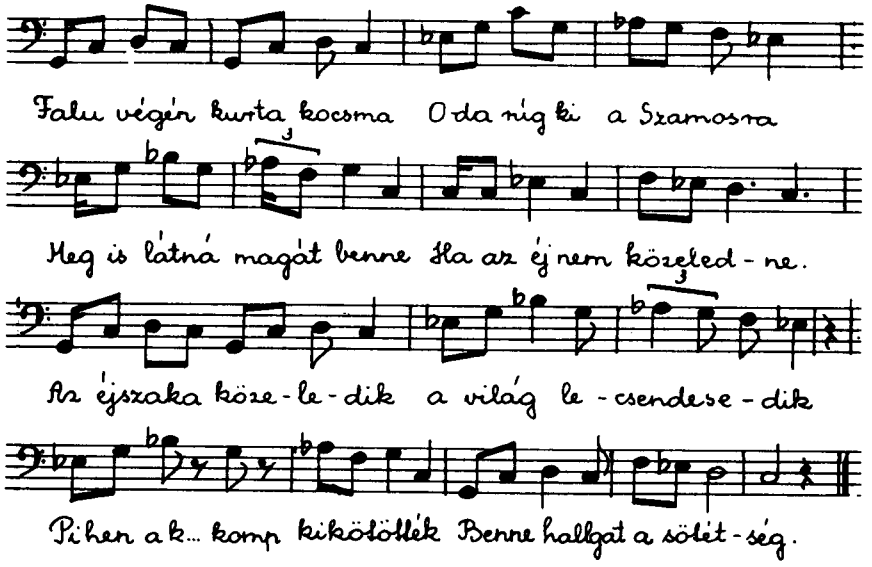
**Example 8.12.** Music pupil, 13 years old.




Form: $A^5$ A B $A_v$
$A^5$ A B $A_v$

Tonality: minor + pentaton old style, folk-song like

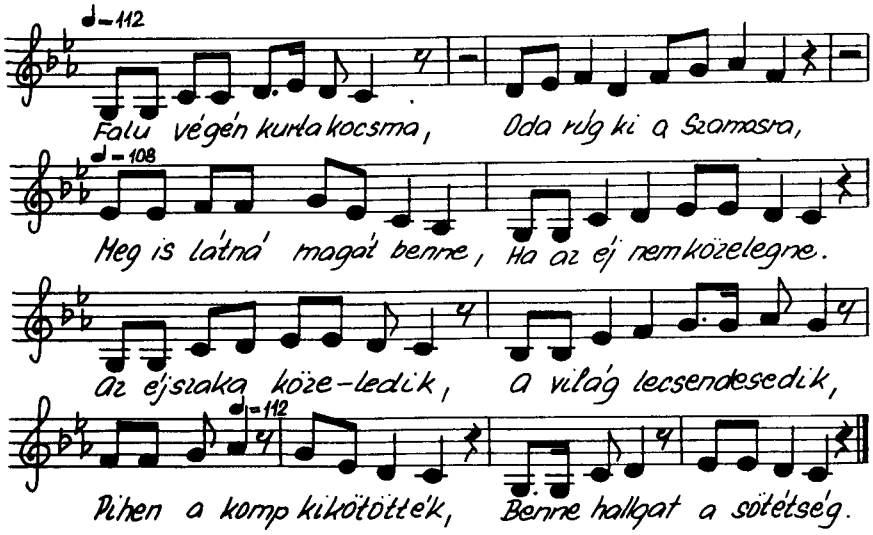
**Example 8.13.** Peasant, 47-year-old-man.




Form: A B $B_v$ C
A $B_{v_2}$ $_v$B C

Tonality: in dome-structure C-minor, E flat major, E flat major, C minor

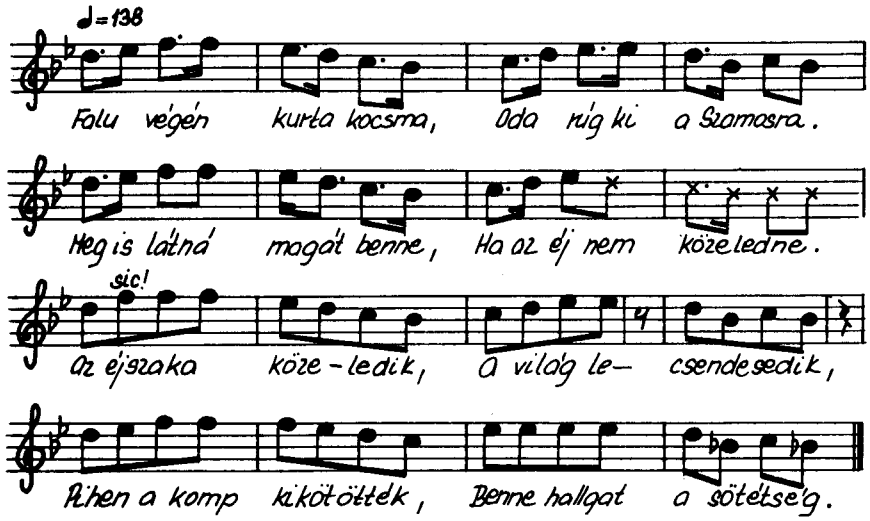
**Example 8.14.** Industrial worker, 30 years old.




Form: A B C $A_v$
$A_v$ $A_v^3$ D $A_v$

Tonality: minor in composed folk song style example for tonality minor

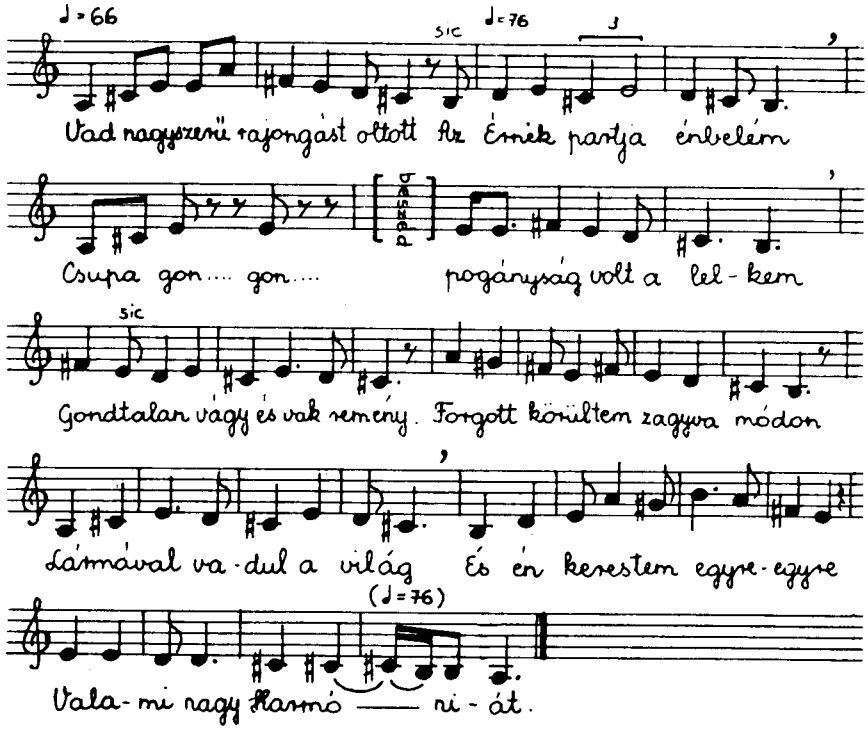
**Example 8.15.** Pupil, 13 years old.



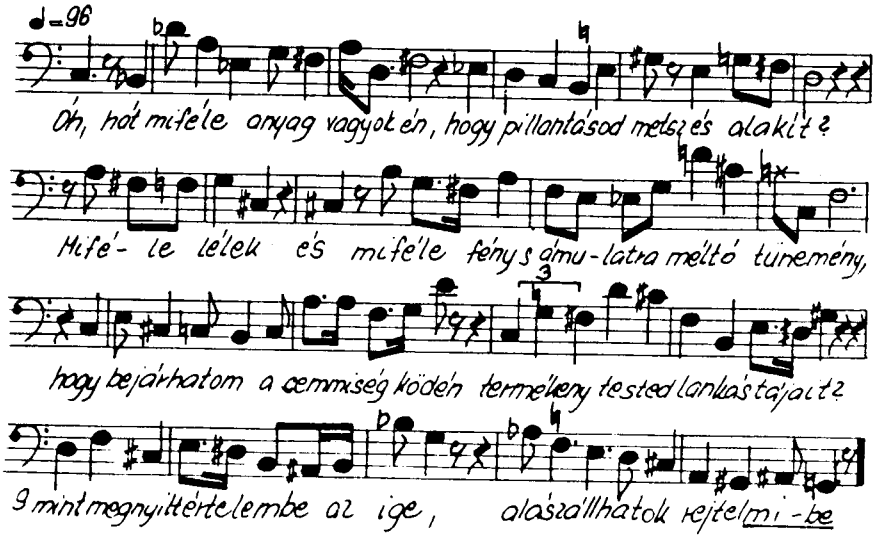

Form: A $A_{2v}$ A $A_{2v}$
$A_v$ $A_2$ A B

Tonality: major pentachord in 'non-Hungarian intonations', i.e. hit song style

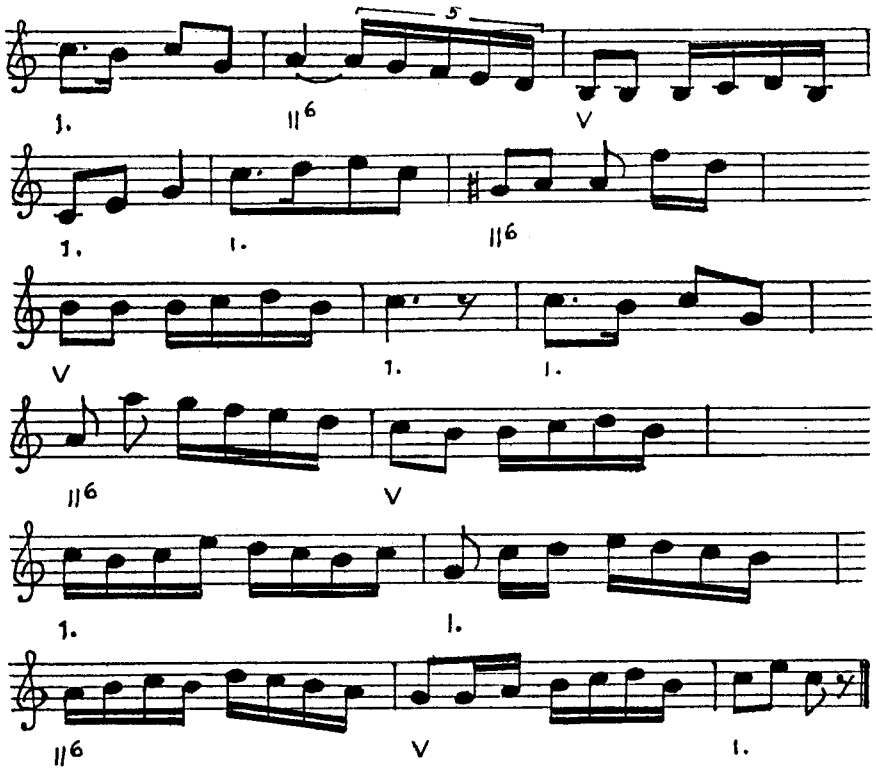
**Example 8.16.** Peasant, 52-year-old woman.




**Example 8.17.** Apprentice, 17-year-old boy.



**Example 8.18.** Grammar school pupil, 17 years old.



# Author index


Adams, J.A. 123, 137, 152
Adams, R. 148
Allport, C.A. 133
Alperson, P. 145
Armbruster, G. 143
Arnold, F.T. 142
Avery, T.L. 141

Babbitt, M. 253
Bach, C.P.E. 142
Bahm, A. 147
Bailey, D. 141, 142
Baker, D. 141
Balzano, G.J. 247, 248
Bamberger, J. 196, 198, 201, 213, 261–2
Bartlett, J.C. 114, 118, 124, 139
Bash, L. 144
Bastick, T. 148
Beall, P.C. 121
Bean, K.L. 91, 92
Beatty, J. 131
Béhague, G. 141
Behne, K.E. 34
Bengtsson, I. 1, 33, 39, 41, 44, 53, 54, 60, 61, 62
Bentley, D. 131
Berger, K. 70
Bergson, H.L. 131, 148
Berkowitz, S. 141
Bernstein, B. 193
Bernstein, N. 132
Bharucha, J.J. 238, 250
Bhaskar, R. 91
Blaukopf, K. 179
Bresgen, C. 141
Brooks, V.B. 131, 135
Bruner, C.L. 252
Bruner, J. 200
Bruner, J.S. 131
Brunia, C.H.M. 135
Bunge, M. 147
Burns, E.M. 244
Burnsed, C.V. 144

Carlson, R. 52, 55, 66
Carr, I. 140
Chadabe, J. 145
Charness, N. 91
Chase, W.G. 91, 92, 93, 109
Chernik, R. 137
Cherry, E.C. 241
Chi, M.T.H. 91, 92, 105, 109, 110
Chomsky, N. 180, 252
Clark, M. 70, 86, 133
Clarke, E.F. 11, 17, 21, 30
Claxton, G. vii, 13
Clynes, M. 30, 31, 36, 54, 65
Cody, R. 216
Cohen, N.J. 131
Coleman, S.N. 144
Coker, J. 141, 142
Cope, D.H. 141
Crozier, W.R. xii
Cuddy, L.L. 261
Czerny, C. 144

Datta, V. 141
Davidson, L. 196, 199
Davis, R. 136
DeGroot, A.D. 91, 262, 268
Delcomyn, F. 133
Denier van der Gon, J.J. 138
Deutsch, D. 239, 248, 249, 252
Deutsch, D.D. 125
Dewitt, L.A. 124
Dobbins, B. 141
Doerschuk, B. 142
Donington, R. 22
Dowling, W.J. 114, 115, 116, 118, 123, 124, 125, 196, 198, 247, 252, 261
Dupré, M. 141

Easton, T.A. 136
Eccles, J.C. 139
Ehmann, W. 142
Erlmann, V. 141
Ewing, A. 147

Ferand, E. 137, 141
Fitts, P.M. 132
Flavell, J.H. 202
Flower, L.S. 91

Fodor, J.A. 255
Foerster, H. von 136
Forte, A. 2
Fraisse, P. 36, 37
Francès, R. 114, 124, 252, 261
Freedman, M.D. 86
Fry, C. 145
Furth, H. 202

Gabrielsson, A. 1, 28, 30, 31, 33, 34, 36, 39, 40, 44, 47, 53, 54, 60, 61, 62, 85
Gaburo, K. 142
Galás, D. 142
Gardner, H. 200, 218
Gates, A. 1
Gehring, P.K. 141
Gelernter, H. 151
Gel'fand, I.M. 132
Gibbs, C.B. 135, 137
Gibson, E.J. 113, 134
Glencross, D.J. 137
Goldstein, M. 142
Goodnow, J. 200
Gorder, W.D. 149
Gordon, E. 196, 198
Greene, P.M. 136
Grey, J.M. 86
Gruson, L.M. 95
Guevara, M. 162
Guilford, J.P. 149

Haber, R.N. 114, 115, 123
Hamel, P.M. 142
Hanlon, E.S. 142
Hartmann, A. 28, 29, 41
Hattwick, M., 198
Hayes-Roth, B. 166
Hebb, D.D. 166
Hendrickson, G. 141
Hevner, K. 34
Hick, W.E. 137
Higgins, J.R. 137
Hindemith, P. 246
Hirsh, I.J. 80
Hodeir, A. 145
Holding, D. 134
Holding, D.H. 141
Hollerbach, J.M. 138
Holst, E. von 136
Hoepfner, R. 149
Hood, M. 141
Hores, R.G. 144
Howard, J. 141
Huggins, A.W.F. 138

Jairazbhoy, N.A. 141
James, W. 139
Jaques-Dalcroze, E. 143, 144

Johnson-Laird, P.N. 263
Jones, A.M. 141
Jones, R.G. 91
Jost, E. 141

Keele, S.W. 133, 137
Kelso, J.A.S. 133, 134, 136
Kernfeld, B. 147
Kessen, W. 122
Kickert, W.G. 136
Kleeman, J.E. 145
Kleinman, D.L. 132
Koblyakov, L. 231, 233, 239
Krumhansl, C.L. 3, 238, 249, 261
Krutetskii, V.A. 91
Kugler, P.N. 133, 134
Kurth, 57

Laabs, G.J. 138
Laneri, R. 142
Lannert, V. 92
Larkin, J. 91
Larsson, B. 52
Lashley, K. 132
Lenat, D.B. 151, 152
Lerdahl, F. ix, xii, 1, 2, 11, 19, 114, 231, 233, 236, 237, 243, 244, 252, 255
Levitt, D.A. 152
Li, T.Y. 161
Liebman, D. 141
Lindblom, B. 2, 65, 66
Lipiczky, T. 141
Locke, D. 141
Lomax, A. 121
Longuet-Higgins, H.C. vii, 2, 33, 248
Lord, A.B. 146, 147

McAdams, S. 239
McCulloch, W.S. 136
MacKenzie, C.L. 9
McKermon, P. 196, 199
Madsen, C.K. 31
Martenuick, R.G. 139
Mathieu, 145
May, R.M. 161
Mersenne, M. 142
Meyer, L.B. 2, 5, 24, 114, 250
Michon, J.A. 79, 80
Milano, D. 145
Miller, G.A. 132, 133, 136, 239, 250, 279
Mills, E. 144
Milner, B. 131
Monahan, C.B. 116
Moorat, J. 121
Moorer, J.A. 75, 86
Morasso, P. 138
Motte-Haber, H. de la 33, 34, 36
Murphy, F. 141

Nagler, M.N. 146
Namikas, G. 133
Narmour, E. 2, 5, 250
Neisser, V. 239
Nelson, O. 142
Nettl, B. 141, 143
Neves, D.M. 108
Newell, K.M. 140, 141
Nielson, F.V. 31
Noorden, L.P.A.S. van 55, 62

Oliveros, P. 142
Olson, D. 200
Ostwald, P.F. 116
Owens, T. 141

Paillard, J. 133
Park, M. 141
Parry, M. 146, 147
Parsons, W. 142, 143
Partch, H. 246
Partchey, K.C. 135, 144
Perkins, D.N. 262
Petzold, R.G. 196
Pew, R.W. 132, 133
Pflederer, M. 196, 198
Piåget, J. 196, 200
Pick, A.D. 113, 126
Pike, A. 145
Plomp, R. 245
Povel, D. 85
Poulton, E.C. 134
Powers, W.T. 132
Pressing, J. ix, 129, 131, 135, 138, 139, 140, 141, 142, 153
Prigogine, I. 134
Propp, V.I. 146

Quantz, J.J. 142

Rabbitt, P.M.A. 137
Radano, R.M. 141
Rakowski, P. 55
Rasch, R.A. 30, 71, 75, 80, 81, 82
Reck, D.B. 141, 147
Reitman, J.J. 91
Rich, E. 151, 152
Richardson, C.P. 149
Rigg, M.G. 34
Risset, J.C. 30
Roads, C. 143
Rosch, E. 256
Rosen, C. 252
Rosner, B.S. 114
Ryle, G. 131

Sage, G.S. 130, 136, 137
Saldanha, E. 70
Saltzman, E.L. 134
Scherer, K.R. 31
Schenker, H. 238, 250, 252
Schickhaus, K.-H. 142
Schmidt, R.A. 133, 138, 140
Schneider, W. 108, 139, 140
Schoenberg, A. 248
Schouten, H. 141
Schroeder, W.H. 141
Schuller, G. 141
Sears, C.H. 28, 41
Seashore, C.E. 23, 28, 29, 41, 47, 53, 54, 81, 200
Seashore, H.G. 28, 47
Seeger, R.C. 121
Serafine, M.L. 124, 152
Shafer, R.M. 144
Shaffer, L.H. 1, 7, 10, 11, 14, 17, 18, 23, 29, 41, 91, 93, 108, 109, 133, 135, 138
Shepard, R.N. 2, 114, 126, 248
Shepherd, G.M. 131
Sher, C. 143
Shiffrin, R.M. 108, 139
Signell, K. 141
Silverman, M.L. 141
Simon, D.P. 91
Simon, H.A. 91, 92, 93, 239, 262
Skinner, L. 28
Sloboda, J.A. vii, 1, 11, 17, 23, 24, 30, 47, 53, 54, 59, 91, 92, 110, 133, 241, 247, 261, 269
Slonimsky, N. 142
Small, A.M. 28
Smith, G.E. 137, 146
Sparrow, W.A. 135
Sperber, M. 141
Steedman, M.J. 33
Stocks, J. 147
Strong, W. 86
Stumme, W. 141
Sudnow, D. 6–8, 145
Suehs, H.C. 144
Sumarsam 141
Summers, J.J. 133, 137
Sundberg, J. 2, 52, 53, 56, 62, 65, 85
Swanwick, K. vii

Taylor, I. 180
Terhardt, E. 245
Thomas, E.H.C. 91
Thompson, S. 146
Todd, N.D. 18
Torrance, E.P., 149
Touma, H.H. 141
Tove, P.A. 75
Trehub, S.E. 115, 117, 118
Treitler, L. 146
Truslit, A. 30

Tubbs, A. 116
Turvey, M.T. 133, 136

Ueltzen, D.R. 142
Upitis, R. 198

Vaughan, M.M. 149
Vernon, L.N. 28, 29
Vikis-Freibergs, V. 146
Vos, J. 71, 81, 82

Wade, B. 141
Wagner, C. 29, 31, 41
Walker, R. 198
Wapnick, J. 31
Weaver, H.E. 9, 92
Weber, M. 179
Webster, R.P. 149
Wedin, L. 34
Weidner, R.F. 141
Welford, A.T. 133, 135, 137, 138, 139, 141
Werner, H. 196, 199, 200
Werts, D. 248
West, L.J. 139
Westcott, M.R. 147, 148
White, S.H. 227
Whitmer, T.C. 141
Wilberg, R.B. 134
Winograd, T. 131
Wolf, T. 91, 92
Webster, R.P. 148

Yorke, J. 161

Zelanik, H.N. 137
Zenatti, A. 196, 261
Zonis, E. 141
Zwislocki, J.J. 86

# Subject index


action plan theory in motor control 133
agogic accent 19
articulation 37–41
  and expression 19–21
  indication 38–9
  synthesized sequence experiments 39, 40
artificial intelligence and improvisation 150–2
  knowledge representation 151–2
  problem reduction 150–1
  search techniques 150
automaticity and skill development 139–40

Baroque style 22
Bethoven piano sonata note values 37, 38
Boulez, P. *Le Marteau sans Maître*, structure of 231–3, 241
  and grammars, musical 233, 234

children, musical representations by 195–230
  and cognitive processes 201–2
  and context, musical 218–20
    rhythm/pitch 222–6, 227–8
    symbols, choice of 220–2
  developing notations 208–18
    and developing performance 215–18
    for song 212–15
    symbol system sequence 211–12
  development of 200–1
  first notations 203–9
    abstract 204–6
    combination/elaboration 208–9
    pictorial 204
    rebus 206, 207
    text 206, 207, 208
  research design 202–3
  three-dimensional view 198–203
    cognitive shift in perception 198–9
    performance ability 199
  *see also* melody construction by children; tonal structure and early learning
compositional systems, cognitive constraints on 231–59
  comprehensibility, aesthetic value of 254–7
    complexity/complicatedness 255–6
    compositional/listening alliance 256–7
  grammar, compositional/listening 233–7
    gap between 234, 235–6
    natural/artificial 235
  listening theory 237–43
    compositional constraints 244–51
  *Le Marteau sans Maître*, structure of 231–3, 241
  serialism, cognitive capacity of 251–4
    consonance/dissonance avoidance 253–4
    permutational structure 252–3
    and pitch space 254
composition study, novice/expert 263–84
  the experiment
    conditions 264–7
    task 263–4
  key definition 272–4
  modulation, economy in 271–2
  strategy levels 274–9
  success of melodies 279
  summary and discussion 279–84
    problem representation 280–1
    strategies 281–2
    tonal thinking 282–3
    tonality, composition/perception of 283–4
  tonality definition 270–1
  working styles
    working units 267–9
    new material genertion 269
    rhythm integration 269–70
    motivic relationships 270
comprehensibility, aesthetic value of 254–7
  complexity/complicatedness 255–6
  compositional/listenign alliance 256–7
computer synthesis, *see* synthesis, computer
conductor and synchronization 86–8
consonance/dissonance
  and octave division 245–6
  and serialism 253–4
creativity studies 179–94
  composing ability 180
  conclusions 193–4
  experimental procedure
    creativity assessment 180–3
    discussion 183–4

creativity studies (*cont.*)
  experimental procedure (*cont.*)
    subjects and data treatment 184–5
  improvisations; *see also* creativity in improvisation
    melodic features 188–90
    structure 187–8
  rhythm 191–3
  tonal-functional thinking 190–1
creativity in improvisation 149–50

diatonic scale 114
  and stability 247
dynamics 19

emotion/motion and timing 30–1
error correction in improvisation 135–6
  times 137
expression in performance 11–24
  interpretation 14–17
  source 17–24
    articulation 19–21
    dynamics 19
    and instrument 21–2
    and manner 23–4
    and style 22–3
    timing 17–19
  structure and 11–14
    interpretative options 14
    rhythm 12–14
    tempo 11–12

feedback and error correction in improvisation 135–6
feedforward 136
folklore and oral tradition and improvisation 146–7
folk songs, tonal structure of 121–3
formulaic composition in oral tradition 146–7

grammars, composition/listening 233–7
  gap between 234, 235–6
  natural/artificial 235

hometetic behaviour 138

improvisation 129–76
  and artificial intelligence 150–2
    knowledge representation 151–2
    problem reduction 150–1
    search techniques 150
  creativity 149–50
  intuition 147–9
  modelling 152–67
    behaviour novelty 161–2
    continuation construction 162–4
    decision making, residual 164–6
    the process 152–7, 160
    redundancy 157–8, 161
    skill development 166–7
  motor control and skilled performance 131–41
    feedback and error correction 135–6
    hierarchy *vs* heterarchy 136
    motor memory 138
    neurology of 130–1
    preparation for action 136
    skill classification 134–5
    skill development 139–41
    theories 132–4
    time scales 136–8
    timing 138
  oral tradition and folklore 146–7
  physiology and neuropsychology 129–31
  structure 6–10
    and motor skills 7
    representations 8–10
  studies and theories 141–5
    teaching 141–5
  *see also* creativity
information processing and musical competence 108–9
instrument, choice of, and expression 21–2
interpretation and expression 14–17
intuition in improvisation 147–9

jazz and improvisation 141, 144–5
  structure 6–10
  teaching 144

knowledge-assembly theory in skill development 167
knowledge representation in artificial intelligence 151–2

legato 21, 37–8
listening, *see* compositional systems, cognitive constraints on

melody construction by children 262; *see also* composition study, novice/expert
memory
  declarative/procedural 131
  motor, and improvisation 138
motion/emotion and timing 30–1
motor control and improvisation 7
  feedback and error correction 135–6
  hierarchy *vs* heterarchy 136
  motor memory 138
  neurology of 130–1
  preparation for action 136
  and skill 131–5
    classification 134–5
    development 139–41
    theories 131–4

time scales 136–8
error correction times 137
reaction times 137
voluntary compensation 138
timing 138
Mozart piano concerto timing deviations 45–6

neuropsychology and improvisation 129–31
motor control 130–1
notation
children's
developing 208–18
first 203–9
and performance 5–6, 27–8
nursery rhymes, tonal structure of 121–3

Observational Scale for Piano Practising (OSPP) 95–7
octave division and stability 247
and consonance/dissonance 245–6
oral tradition and folklore and improvisation 146–7
organizational invariant theory in motor control 133–4

parallelism in composition 240
performance
children, development in 199, 215–18
and representational development 216–17
and notation 27–8
*see also* expression in performance; improvisation; synchronization in ensemble performance
phenomenal accents and rhythm 240–1
physiology and improvisation 129–31
motor control 130–1
pitch
assimilation 123–4
and children
representational development 222–6
sensitivity 117–18
collections and stability 244–7
consonance/dissonance and octave division 245–6
nursery rhymes/folk songs 121–3
space
and serialism, cognitive opacity 254
and stability 248–51
practice and skill development 140

reaction times in improvisation 137
rehearsal skill and musical competence 91–112
discussion 106–11
automatic/controlled processing 108
future research 109–11
information processing and competence 108–9
materials 94–5
perceptual span 92
procedure 95–8
Observational Scale for Piano practising 95–7
results 98–106
analyses 98–104
interview 104–6
study outline 93
subjects 93
rhythm
in creativity 191–3
and expression 12–14
perception and tone durtion 36–7
phenomenal accents and 240–1
representational development in children 222–6, 227–8
and performance skill 226
Romantic style 22
rubato 23

scale categories and stability 244–5
schemata theory of motor control 133
serialism
cognitive opacity 251–4
consonance/dissonance avoidance 253–4
permutational structure 252–3
and pitch space 254
and *Le Marteau sans Maître* 231–3
sight reading and perceptual span 92
singing
by children/infants 116–17
synthesizer 52
skill
classification and improvisation 134–5
development 139–41
automaticity 139–40
expressive control parameters 139
modelling 166–7
and practice 140
teaching 140–1
motor control and 131–4
theories 132–4
*see also* rehearsal skill and musical competence
staccato 14, 20–1, 37–8
structures, generative 1–10
hierarchical nature 2
improvised performance 6–10
and motor skills 7
representations 8–10
notation, performance from 5–6
partial activation and phrase boundaries 4–5
unified conception 3–4
symmetry in composition 240

synchronization in ensemble performance 70–90
  and the conductor 86–8
  description 71–5
    asynchronization and onset times 71–4
    isochronization and tone duration 74–5
  measurement 75–81
    perceptual effects 80–1
  'onset moment' 70–1
  perceptual onset threshold model 81–6
synthesis, computer 52–69
  analysis-by-synthesis 63
  homonymous melodies and performance variability 59
  note sign/acoustic signal discrepancy 53–4
  rules 56–9
    code 65–7
    interaction of 59–60
    performance, predicted/actual 60–2
    purposes 64–5
    quantity 62–3
  singing synthesizer 52
  *see also* artificial intelligence

teaching
  improvisation 141–5
  and problem solving technique 150–1
tempo 31–5
  and expression 11–12
  and motion 34
  perception 31–2
  simultaneous, in contemporary music 241
  and skill acquisition 104
  variation 33
timing 27–51
  analysis of performance 28–9
  articulation 37–41
    indication 38–9
    synthesized sequence experiments 39–40
  deviations in performance 41–6
    analytical methods 41–4
    at beat level 45
    in sequences of notes 44–5
  duration of tones 35–7
    rhythm perception 36–7
    tone sequences and motion 36
  and expression 17–19
  and improvisation 138
  motion/emotion 30–1
  performance and notation 27–8
  performer's *vs* listener's experience 46–8
  *see also* tempo
tonal structure and early learning 113–28
  adult encoding and experience 123–6
    and octave scrambling 125–6
    quarter step assimilation 123–4
    and tonal content 124–5
  children 117–21
    atonal melody sensitivity 118–21
    pitch sensitivity 117–18
  expectancy and coding 114–15
  infancy responses 115–17
    early reactions 115
    singing 116–17
  invariants, 113–14
  nursery rhymes 121–3
tonal thinking development 260–85
  perception and production 261–2
    children and bells 262
  domains and problem solving 262–3
  composition, novice/expert study
    the experiment 263–7
    key definition 272–4
    modulation, economy in 271–2
    strategy levels 274–9
    success of melodies 279
    summary and discussion 279–84
    tonality definition 270–1
    working styles 267–70

waltz
  tempo variation 33
  timing deviations in performance 45

Whereas most of the literature in the psychology of music has focused on the processes involved when listening to music, little has been written about the processes involved in making music.

Re-issued by popular demand, and for the first time in paperback, *Generative processes: the psychology of performance, improvisation, and composition* brings together leading figures in music psychology to present pioneering studies of the processes by which music is generated.

The book looks at the generation of expression in musical performance, the problems of synchrony in ensemble performance, the development of children's song, rehearsal strategies of pianists, improvisational skill in trained and untrained musicians, children's spontaneous notations for music, formal constraints on compositional systems, and compositional strategies of music students.

Edited by a leading authority on music psychology, the book will be of great interest to cognitive and developmental psychologists, as well as music educators and musicologists.

*Cover image © Herman Leonard*

OXFORD
UNIVERSITY PRESS

ISBN 0-19-850846-8
9 780198 508465

# Musical Perceptions



Edited by **Rita Aiello** with John A. Sloboda

# MUSICAL PERCEPTIONS

*edited by*

RITA AIELLO

*with*

JOHN A. SLOBODA

*New York* *Oxford*

OXFORD UNIVERSITY PRESS

1994

Oxford University Press

Oxford New York Toronto
Delhi Bombay Calcutta Madras Karachi
Kuala Lumpur Singapore Hong Kong Tokyo
Nairobi Dar es Salaam Cape Town
Melbourne Auckland Madrid

and associated companies in
Berlin Ibadan

Copyright © 1994 by Oxford University Press, Inc.

Published by Oxford University Press, Inc.,
200 Madison Avenue, New York, New York 10016

Oxford is a registered trademark of Oxford University Press

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
electronic, mechanical, photocopying, recording, or otherwise,
without the prior permission of Oxford University Press.

Library of Congress Cataloging-in-Publication Data
Musical perceptions / edited by Rita Aiello with John A. Sloboda.
p. cm. Includes bibilographical references and index.
ISBN 0-19-506475-5
ISBN 0-19-506476-3 (pbk.)
1. Music—Psychology.
2. Auditory perception.
3. Music—Philosophy and aesthetics.
I. Aiello, Rita.
II. Sloboda, John A.
ML3830.M983 1994 781'.11—dc20
93-28888

3 5 7 9 8 6 4

Printed in the United States of America
on acid-free paper

*To my daughter Daniela Caterina*

# PREFACE

*Musical Perceptions* is an introduction to the perception and cognition of music for readers in music and psychology. It presents a spectrum of the current research, and describes some of the many developments that have occurred in the perception and cognition of music over the last decades. This book includes twelve contributed chapters that discuss musical meaning, philosophical perspectives of music, comparisons beween music and language, the importance of music theory in perceptual research, research in musical development and music performance, the perception of melodies, tonality and rhythm, and the application of neural network models in music perception. The contributors' chapters give a detailed description of their research. I have prefaced each chapter with an introduction to provide a general background in which to place the significance of the contributor's research (John Sloboda contributed the introduction to his own chapter).

Because *Musical Perceptions* aims to foster a closer interaction between the artistic, the pedagogical, and the scientific aspects of music perception research, it includes contributions from both musicians and psychologists. Undoubtedly, the question "How do we perceive music?" is the main theme of this book. But there is also a second question, more difficult to answer, that underlines its intent: "How does music perception give rise to musical meaning?" Although the material discussed here offers only a small glimpse into the current vast literature on music perception and cognition, the reader will find the many references useful for more in-depth study.

I have tested the chapters that follow in my classes because they reflect the topics I teach in my perception of music courses at the Juilliard School, the Manhattan School of Music, and in the department of psychology at New York University. I am very grateful to my students for their comments and suggestions. In addition to being used in psychology of music courses (be they offered in music or in psychology departments), this volume could also be used in courses in the philosophy and the acsthetics of music, music education, and music theory.

The book opens with three chapters that address the perception of music from a philosophical perspective. Chapter 1 by Leonard Meyer, reprinted from *Emotion and Meaning in Music*, is of historical significance, and Meyer's ideas are still at the heart of the psychology of music today. Meyer explains how a listener may listen to music from several philosophical positions. Furthermore, Meyer shows how the melodic, rhythmic, and harmonic organization of music gives rise to musical meaning.

In Chapter 2, I describe comparisons between the perception of music and the perception of language at the phonetic, syntactic, and semantic levels. My

intent is not to underline the obvious similarities between the two, but to use the comparisons to point out what is intrinsically unique to music. Nicholas Cook's contribution (Chapter 3) offers a music theorist's view on music perception research, and his criticisms are thought-provoking. Cook argues that many studies in music perception are experiments in ear training strategies that do not reflect what music listeners actually do when they listen to music.

Chapters 4, 5, and 6 focus on musical development and music performance. Lyle Davidson (Chapter 4) traces the development of songsinging from the young child to the adult, and discusses the perceptual changes in tonal relationship that a student in music school undergoes. Davidson proposes that we do not yet have a comprehensive view of musical development because we fail to learn from the uncoordinated data we already have at hand. For Jeanne Bamberger (Chapter 5), listening to music in a new way enriches one's understanding of music. She elegantly describes how music students acquire new ways of listening to a piece when sharing with each other their descriptions of the piece.

In Chapter 6, John Sloboda brings new insights to research on music performance. Specifically, Sloboda examines recent studies that evaluate the ability to play notated solo compositions expressively, and the musical intentions of the performers. In addition, Sloboda discusses musical talent, practicing, and other factors that may be considered biographical precursors of musical excellence.

Aspects of the perception of melody and tonality are addressed in Chapters 7, 8, and 9. Jay Dowling (Chapter 7) defines the features that make a melody perceptually relevant. Drawing from his extensive research in this area, Dowling explains the perceptual strategies that we engage in when listening to tonal, atonal, and transposed melodies. Butler and Brown (Chapter 9) propose that the perception of tonality is influenced by the temporal organization of the pitches and not by the pitches themselves. They show experimentally that the frequency with which certain musical intervals are heard within a piece is what conveys information about the tonality of that piece.

In Chapter 9, Jamshed Bharucha demonstrates the relationship between tonality and expectation. Specifically, he explains the cognitive processes that underlie tonal expectation and their acquisition, and differentiates between the musical expectations that are related to the general style of a piece and those that are related to an actual composition. Moreover, Bharucha illustrates how musical expectations may be investigated with the application of neural networks.

Both Chapters 10 and 11 address the perception of rhythm and timing, but from different perspectives. While Sternberg and Knoll (Chapter 10) report on the ability of musicians to perceive, produce, and imitate microrhythmic patterns in the context of a laboratory setting, Shaffer and Todd (Chapter 11) describe the ways in which pianists reproduce rhythms while playing actual performances at the piano. In chapter 12, I address the importance of developing more studies in the perception of music that use as stimuli compositions from the musical repertoire rather than aesthetically impoverished laboratory sounds. Ul-

timately, research in the perception of music should help us understand the meaning and the aesthetics of music.

R. A.

New York
April 1993

# ACKNOWLEDGMENTS

My students are a very important part of my life. It is with them that I share my love for music and my interest in understanding how music is perceived. Their questions and comments have given me insights in the meanings of music, and have presented me with new perspectives. I thank all of them. It was a milestone for me to read the transcript of Leonard Bernstein's lectures *The Unanswered Question: Six Talks at Harvard.* As a musician and as a teacher, I thank the late maestro for what I learned.

Many colleagues have helped me with various aspects of this book. I acknowledge with deep gratitude the late Louis Gerstman, professor of psychology at the City College and the Graduate Center of the City University of New York, and a musician at heart. His enthusiasm and expert advice helped me envision how my lectures on the perception of music could become the book now at hand. In addition, I thank John Lewers, a dear friend, for the exchanges of ideas we have had since our student days; Edgar (Ted) Coons for his guidance and friendship; and Thomas Bever for introducing me to the literature in the perception of music. My thanks also to John Sloboda, David Butler, Leonard Meyer, David Burrows, Kendall Durell Briggs, Jerry Kerlin, Michael Palij, Valerie Bernhardt, Joan Bossert, and Paul Schlotthauer. Jane Gottlieb, the chief librarian at the Juilliard School, was very gracious in answering all my questions and in discussing many references at length.

Above all, I am grateful to my mother and to my husband, John G. J. Ritter, for their support.

# CONTENTS

Contributors, *xiii*

PHILOSOPHICAL PERSPECTIVES

1. Emotion and Meaning in Music, *3*
   Leonard B. Meyer
2. Music and Language: Parallels and Contrasts, *40*
   Rita Aiello
3. Perception: A Perspective from Music Theory, *64*
   Nicholas Cook

DEVELOPMENTAL PERSPECTIVES

4. Songsinging by Young and Old: A Developmental Approach to Music, *99*
   Lyle Davidson
5. Coming to Hear in a New Way, *131*
   Jeanne Bamberger
6. Music Performance: Expression and the Development of Excellence, *152*
   John A. Sloboda

THE PERCEPTION OF MELODY, TONALITY RHYTHM AND TIMING

7. Melodic Contour in Hearing and Remembering Melodies, *173*
   W. Jay Dowling
8. Describing the Mental Representation of Tonality in Music, *191*
   David Butler and Helen Brown
9. Tonality and Expectation, *213*
   Jamshed J. Bharucha
10. Perception, Production, and Imitation of Time Ratios by Skilled Musicians, *240*
    Saul Sternberg and Ronald L. Knoll
11. The Interpretive Component in Musical Performance, *258*
    L. H. Shaffer and Neil P. McAngus Todd

THE PERCEPTION OF MUSICAL COMPOSITIONS

12. Can Listening to Music Be Experimentally Studied?, *273*
    Rita Aiello

Author Index, *283*
Subject Index, *287*

# CONTRIBUTORS


RITA AIELLO
The Juilliard School and
Department of Psychology
New York University and
The Manhattan School of Music
New York, New York

JEANNE BAMBERGER
Department of Music
Massachusetts Institute of Technology
Cambridge, Massachusetts

JAMSHED J. BHARUCHA
Department of Psychology
Dartmouth College
Hanover, New Hampshire

HELEN BROWN
Division of Music
Purdue University
West Lafayette, Indiana

DAVID BUTLER
School of Music
The Ohio State University
Columbus, Ohio

NICHOLAS COOK
Department of Music
University of Southampton
Highfield, Southampton
England

LYLE DAVIDSON
Project Zero
Harvard University
Cambridge, Massachusetts and
The New England Conservatory
Boston, Massachusetts

W. JAY DOWLING
Program in Applied Cognition
and Neuroscience
University of Texas, Dallas
Richardson, Texas

RONALD L. KNOLL
Bell Laboratories
Murray Hill, New Jersey

LEONARD B. MEYER
Department of Music
University of Pennsylvania
Philadelphia, Pennsylvania

L. H. SHAFFER
Department of Psychology
University of Exeter
Exeter, England

JOHN A. SLOBODA
Department of Psychology
University of Keele
Keele, Staffordshire
England

SAUL STERNBERG
Department of Psychology
University of Pennsylvania
Philadelphia, Pennsylvania

NEIL P. McANGUS TODD
Department of Psychology
University of Exeter
Exeter, England and
Department of Music
City University
London, England

# Philosophical Perspectives

# 1

# Emotion and Meaning in Music

LEONARD B. MEYER

## Introduction

Leonard B. Meyer, one of the most important figures in the field of music aesthetics, has been pivotal in the growth of the psychology of music for almost four decades. The historical importance of his first book, *Emotion and Meaning in Music,* 1956, is evident in the frequency with which it is quoted, and in the impact that the theories it set forth have had on the psychology of music as a whole. *Emotion and Meaning in Music* focuses primarily on the moment-to-moment processes that take place in listening to total compositions and argues that the central meaning of music is to be found in the structural interplay of its elements and forms, rather than in associations to nonmusical events. Specifically, Meyer considers what kind of patterning underlies music, what implications are suggested by its melodic, rhythmic, and harmonic organization, and how these implications are realized or not realized in the ongoing sequence of events.

Meyer's writings address musical meaning from a psychological perspective. His theories are grounded in Gestalt psychology and information theory. The laws of Gestalt psychology (e.g., good continuation, similarity, common fate) relate to patterns in any modality, be it visual, auditory, or tactile, and provide explanations for the fact that some patterns seem particularly "good" or "complete" (the reader may wish to consult Rudolph Arnheim's *Art and Visual Perception,* 1969, for an in-depth discussion of Gestalt psychology; see also the introduction to Dowling's chapter 7 in this volume). The principles of information theory offer parameters for understanding how much information

Reprinted from L. B. Meyer, *Emotion and Meaning in Music.* Chicago: University of Chicago Press, 1956. Used by permission.

the listener can possibly derive when listening to music. (The works of Moles, 1966, Kraehenbuehl and Coons, 1959, Coons and Kraehenbuehl, 1958, and Youngblood, 1958, provide a background in information theory and music). In this vein, a style is a system of probabilities; the music, as it unfolds, sets the probabilities for what will happen next. The listener who is aware of the style may successfully predict what will follow, thus deriving a satisfactory meaning from the music. Deviations from the musical norms violate the listener's expectations and create patterns of tension and resolution.

Meyer's theories of meaning in music (1956) *preceded* a great deal of psychological research in meaning and emotional responses (see Gaver & Mandler, 1987; Mandler, 1984; Schachter & Singer, 1962). Psychologists (Mandler, 1984; Schachter & Singer, 1962) have shown that a subject's knowledge of a stimulus is of paramount importance in determining his or her emotional reactions to that stimulus. Meyer's theories of expectation and resolution in music have been found to be empirically valid (Schmuckler, 1989).

The first chapter of *Emotion and Meaning in Music* addresses music from a philosophical point of view. Music can have an intrinsic "absolute" meaning or can have an extrinsic "referential" meaning. Furthermore, the meaning of music may be interpreted from the viewpoint of the "formalist" (i.e., the listener who values the abstract patterning of music in and of itself) or from the viewpoint of the "expressionist" (i.e., the listener who views music as a means to express emotions). Meyer claims that a philosophy of music may legitimately endorse either of the opposing positions on each dimension. For example, an absolute expressionist (Meyer himself is an absolute expressionist) is a listener who believes that musical meaning is intrinsic in the music, and who derives an emotional meaning from understanding the interplay of the musical relationships.

We reprint the first chapter of *Emotion and Meaning in Music* not only because of its historical importance but to bring into focus the importance of having a philosophical perspective in addressing questions about the perception of music and to emphasize the historical importance of the book. Like Meyer, we believe that the philosophical positions of listeners may influence their perception of music. Research in the psychology of music asks what people perceive, but seldom have the subjects' philosophical views been taken into account. Studying what subjects report when hearing music should go hand in hand with asking questions about their philosophical perspective of music.

Meyer's latest book, *Style and Music: Theory, History, and Ideology,* 1989, further develops his belief that a musical style is a result of human choices that develop because of psychological, cultural, and musical constraints.

## REFERENCES

Arnheim, R. (1969). *Art and visual perception.* Berkeley: University of California Press.

Coons, E., & Kraehenbuehl, D. (1958). Information as a measure of structure in music. *Journal of Music Theory, 2,* 127–61.

Gaver, W. W., & Mandler, G. (1987). Play it again, Sam: On liking music. *Cognition and Emotion, 1*(3), 259–82.

Kraehenbuehl, D., & Coons, E. (1959). Information as a measure of experience in music. *Journal of Aesthetics and Art Criticism, XVII,* 510–22.

Mandler, G. (1984). *Mind and body: Psychology of emotion and stress.* New York: Norton.

Meyer, L. B. (1956). *Emotion and meaning in music.* University of Chicago Press.

Meyer, L. B. (1957). Meaning in music and information theory. *Journal of Aesthetics and Art Criticism, XV*(4), 412–24.

Meyer, L. B. (1961). On rehearing music. *Journal of the American Musicological Society, XIV*(2), 257–73.

Meyer, L. B. (1967). *Music, the arts and ideas: Patterns and predictions in twentieth-century music.* Chicago: University of Chicago Press.

Meyer, L. B. (1973). *Explaining music: Essays and explorations.* Chicago: University of Chicago Press.

Meyer, L. B. (1974). Concerning the sciences, the arts—AND the humanities. *Critical Inquiry, 1*(1), 163–217.

Meyer, L. B. (1975). Grammatical simplicity and relational richness: The trio of Mozart's G minor symphony. *Critical Inquiry, 2*(4), 693–761.

Meyer, L. B. (1980). Exploiting limits: Creation, archetypes, and style change. *Daedalus, 109*(2), 177–205.

Meyer, L. B. (1982). Process and morphology in the music of Mozart. *The Journal of Musicology, 1*(1), 317–41.

Meyer, L. B. (1989). *Style and music: Theory, history, and ideology.* Philadelphia: University of Pennsylvania Press.

Moles, A. (1966). *Information theory and esthetic perception.* Translated by J. E. Cohen. Urbana: University of Illinois Press.

Schachter, S., & Singer, J. (1962). Cognitive, social and physiological determinants of emotional state. *Psychological Review, 69,* 379–99.

Schmuckler, M. A. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception, 7*(2), 109–50.

Youngblood, J. E. (1958). Style as information. *Journal of Music Theory, 2,* 24–35.

## Past Positions as to the Nature of Musical Experience

Composers and performers of all cultures, theorists of diverse schools and styles, aestheticians, and critics of many different persuasions are all agreed that music has meaning and that this meaning is somehow communicated to both participants and listeners. This much, as least, we may take for granted. But what constitutes musical meaning and by what processes it is communicated has been the subject of numerous and often heated debates.

The first main difference of opinion exists between those who insist that musical meaning lies exclusively within the context of the work itself, in the perception of the relationship set forth within the musical work of art, and those who contend that, in addition to these abstract, intellectual meanings, music also communicates meanings which in some way refer to the extramusical world of concepts, actions, emotional states, and character. Let us call the former group the "absolutists" and the latter group the "referentialists."

In spite of the persistent wrangling of these two groups, it seems obvious that absolute meanings and referential meanings are not mutually exclusive: that they can and do coexist in one and the same piece of music, just as they do in a poem or a painting. In short, the arguments are the result of a tendency toward philosophical monism rather than a product of any logical opposition between types of meaning.

Because this study deals primarily with those meanings which lie within the closed context of the musical work itself, it is necessary to emphasize that the prominence given to this aspect of musical meaning does not imply that other kinds of meaning do not exist or are not important.

On the contrary, the musical theory and practice of many different cultures in many different epochs indicates that music can and does convey referential meaning. The musical cosmologies of the Orient in which tempi, pitches, rhythms, and modes are linked to and express concepts, emotions, and moral qualities; the musical symbolisms depicting actions, character, and emotion, utilized by many Western composers since the Middle Ages; and evidence furnished by testing listeners who have learned to understand Western music—all these indicate that music can communicate referential meanings.

Some of those who have doubted that referential meanings are "real" have based their skepticism upon the fact that such meanings are not "natural" and universal. Of course, such meanings depend upon learning. But so, too, do purely musical meanings—a fact that will become very clear in the course of this study.

Others have found the fact that referential meanings are not specific in their denotation a great difficulty in granting status to such meanings. Yet such precision is not a characteristic of the nonmusical arts either. The many levels of

connotation play a vital role in our understanding of the meanings communicated by the literary and plastic arts.

Let us now make a second point clear, namely, that the distinction just drawn between absolute and referential meanings is not the same as the distinction between the aesthetic positions which are commonly called "formalist" and "expressionist." Both the formalist and the expressionist may be absolutists; that is, both may see the meaning of music as being essentially intramusical (nonreferential); but the formalist would contend that the meaning of music lies in the perception and understanding of the musical relationships set forth in the work of art and that meaning in music is primarily intellectual, while the expressionist would argue that these same relationships are in some sense capable of exciting feelings and emotions in the listener.

This point is important because the expressionist position has often been confused with that of the referentialist. For although almost all referentialists are expressionists, believing that music communicates emotional meanings, not all expressionists are referentialists. Thus when formalists, such as Hanslick or Stravinsky, reacting against what they feel to be an overemphasis upon referential meaning, have denied the possibility or relevance of any emotional response to music, they have adopted an untenable position partly because they have confused expressionism and referentialism.

One might, in other words, divide expressionists into two groups: absolute expressionists and referential expressionists. The former group believe that expressive emotional meanings arise in response to music and that these exist without reference to the extramusical world of concepts, actions, and human emotional states, while the latter group would assert that emotional expression is dependent upon an understanding of the referential content of music.

### *The Present Position and Criticism of Past Assumptions*

The present study is concerned with an examination and analysis of those aspects of meaning which result from the understanding of and response to relationships inherent in the musical progress rather than with any relationships between the musical organization and the extramusical world of concepts, actions, characters, and situations. The position adopted admits both formalist and absolute expressionist viewpoints. For though the referential expressionists and the formalists are concerned with genuinely different aspects of musical experience, the absolute expressionists and the formalists are actually considering the same musical processes and similar human experiences from different, but not incompatible, viewpoints.

Broadly speaking, then, the present investigation seeks to present an analysis of musical meaning and experience in which both the expressionist and the formalist positions will be accounted for and in which the relationship between them will become clear.

Past accounts given by the proponents of each of these positions have suffered from certain important weaknesses. The chief difficulty of those who have adopted the absolutist expressionist position is that they have been unable to account for the processes by which perceived sound patterns become experienced as feelings and emotions. In fact, strange as it may seem, they have generally avoided any discussion of emotional responses whatsoever. These shortcomings have led to a general lack of precision both in their account of musical experience and in their discussions of musical perception.

But, at least, the expressionists have recognized the existence of problems in their position. The formalists, on the other hand, have either found no problems to recognize or have simply turned the other way, seeking to divert attention from their difficulties by attacking referentialism whenever possible. Yet the formalists are faced with a problem very similar to that confronting the expressionists: namely, the difficulty and necessity of explaining the manner in which an abstract, nonreferential succession of tones becomes meaningful. In failing to explain in what sense such musical patterns can be said to have meaning, they have also found themselves unable to show the relation of musical meaning to meaning in general.

Finally, this failure to explain the processes by which feelings are aroused and meanings communicated has prevented both groups from seeing that their positions should make them allies rather than opponents. For the same musical processes and similar psychological behavior give rise to both types of meaning; and both must be analyzed if the variety made possible by this aspect of musical experience is to be understood.

Readers familiar with past studies in the aesthetics and psychology of music will perhaps note that much of the earlier work in these fields is not discussed in this study and that many traditional problems are ignored. This neglect stems from the conviction that the assumptions and orientation of this literature have proved sterile and are today untenable. Since this literature has been explicitly and cogently criticized by such writers as Cazden,[1] Farnsworth,[2] and Langer,[3] only a brief comment on these earlier assumptions seems necessary here, in the hope that the position of this book will thereby be clarified.

The psychology of music has, since its beginnings, been plagued by three interrelated errors: hedonism, atomism, and universalism. Hedonism is the confusion of aesthetic experience with the sensuously pleasing. As Susanne Langer writes:

> Helmholtz, Wundt, Stumpf, and other psychologists . . . based their inquiries on the assumption that music was a form of *pleasurable sensation*. . . . This gave rise to an aesthetic based on liking and disliking, a hunt for a sensationist definition of beauty. . . . But beyond a description of tested pleasure-displeasure reactions to simple sounds or elementary sound complexes . . . this approach has not taken us. . . .[4]

The attempt to explain and understand music as a succession of separable, discrete sounds and sound complexes is the error of atomism. Even the meager achievement which Mrs. Langer allows to studies of this kind must be still further depreciated. For the tested pleasure-displeasure reactions are not what most of the psychologists tacitly assumed them to be: they are not universals (good for all times and all places) but products of learning and experience.

This is the third error, the error of universalism: the belief that the responses obtained by experiment or otherwise are universal, natural, and necessary. This universalist approach is also related to the time-honored search for a physical, quasi-acoustical explanation of musical experience—the attempt, that is, to account for musical communication in terms of vibrations, ratios of intervals, and the like.

These same errors have also plagued music theory. Attempts to explain the effect of the minor mode of Western music, to cite but one example, in terms of consonance and dissonance or in terms of the harmonic series have resulted in uncontrolled speculations and untenable theories. Even those not thus haunted by the ghost of Pythagoras have contributed little to our understanding of musical meaning and its communication. For, on the whole, music theorists have concerned themselves with the grammar and syntax of music rather than with its meaning or the affective experiences to which it gives rise.

Today we are, I think, able to take a somewhat more enlightened view of these matters. The easy access which almost all individuals have to great music makes it quite apparent that a Beethoven symphony is not a kind of musical banana split, a matter of purely sensuous enjoyment. The work of the Gestalt psychologists has shown beyond a doubt that understanding is not a matter of perceiving single stimuli, or simple sound combinations in isolation, but is rather a matter of grouping stimuli into patterns and relating these patterns to one another. And finally, the studies of comparative musicologists, bringing to our attention the music of other cultures, have made us increasingly aware that the particular organization developed in Western music is not universal, natural, or God-given.

## Evidence as to the Nature and Existence of the Emotional Response to Music

Any discussion of the emotional response to music is faced at the very outset with the fact that very little is known about this response and its relation to the stimulus. Evidence that it exists at all is based largely upon the introspective reports of listeners and the testimony of composers, performers, and critics. Other evidence of the existence of emotional responses to music is based upon the behavior of performers and audiences and upon the physiological changes that accompany musical perception. Although the volume and intercultural

character of this evidence compel us to believe that an emotional response to music does take place, it tells us almost nothing about the nature of the response or about the causal connection between the musical stimulus and the affective response it evokes in listeners.

### *Subjective Evidence*

From Plato down to the most recent discussions of aesthetics and the meaning of music, philosophers and critics have, with few exceptions, affirmed their belief in the ability of music to evoke emotional responses in listeners. Most of the treatises on musical composition and performance stress the importance of the communication of feeling and emotion. Composers have demonstrated in their writings and by the expression marks used in their musical scores their faith in the affective power of music. And finally, listeners, past and present, have reported with remarkable consistency that music does arouse feelings and emotions in them.

The first difficulty with this evidence is that, taken at its face value, without benefit of a general theory of emotions as a basis for interpretation, it yields no precise knowledge of the stimulus which created the emotional response. Because music flows through time, listeners and critics have generally been unable to pinpoint the particular musical process which evoked the affective response which they describe. They have been prone, therefore, to characterize a whole passage, section, or composition. In such cases the response must have been made to those elements of the musical organization which tend to be constant, e.g., tempo, general range, dynamic level, instrumentation, and texture. What these elements characterize are those aspects of mental life which are also relatively stable and persistent, namely, moods and associations, rather than the changing and developing affective response with which this study is concerned.

Much confusion has resulted from the failure to distinguish between emotion felt (or affect) and mood. Few psychologists dealing with music have been as accurate on this point as Weld, who notes that: "The emotional experiences which our observers reported are to be characterized rather as moods than as emotions in the ordinary sense of the term. . . . The emotion is temporary and evanescent; the mood is relatively permanent and stable."[5] As a matter of fact, most of the supposed studies of emotion in music are actually concerned with mood and association.

Taken at face value the introspective data under consideration not only fail to provide accurate knowledge of the stimulus (music) but they cannot even furnish clear and unequivocal information about the responses reported. For several reasons the verbalizations of emotions, particularly those evoked by music, are usually deceptive and misleading.

Emotions are named and distinguished from one another largely in terms of the external circumstances in which the response takes place. Since, aside from the often fortuitous associations which may be aroused, music presents no exter-

nal circumstances, descriptions of emotions felt while listening to music are usually apocryphal and misleading. If they are to be used at all, they must be analyzed and considered in the light of a general theory of the relation of musical stimuli to emotional responses.

Second, a clear distinction must be maintained between the emotions felt by the composer, listener, or critic—the emotional response itself—and the emotional states denoted by different aspects of the musical stimulus. The depiction of musical moods in conjunction with conventional melodic or harmonic formulas, perhaps specified by the presence of a text, can become signs which designate human emotional states. Motives of grief or joy, anger or despair, found in the works of baroque composers or the affective and moral qualities attributed to special modes or *rāgas* in Arabian or Indian music are examples of such conventional denotative signs. And it may well be that when a listener reports that he felt this or that emotion, he is describing the emotion which he believes the passage is supposed to indicate, not anything which he himself has experienced.

Finally, even where the report given is of a genuine emotional experience, it is liable to become garbled and perverted in the process of verbalization. For emotional states are much more subtle and varied than are the few crude and standardized words which we use to denote them.

In this connection it would seem that many of the introspections supplied by subjects in the studies made by Vernon Lee, C. S. Myers, Max Schoen, and others contain a large amount of what psychiatrists call "distortion." For example, when a subject in an experiment by Myers reports that while listening to a particular musical selection she had "a restful feeling throughout . . . like one of going downstream while swimming,"[6] she is obviously translating unspeakable feelings into symbolic form. The interpretation of such symbols is the task of the psychiatrist, not the music critic. To the music critic such introspections show only that some response, not necessarily a specifically musical one, was present. For it is always possible that the thoughts and reveries thus revealed are without any relation to musical experience. The musical stimuli may have functioned merely as a kind of catalytic agent, enabling the response to take place but playing no controlling part in shaping or determining the experience and figuring nowhere in the end result, except perhaps negatively.

### *Objective Evidence: Behavior*

The responses of listeners can also be observed and studied objectively. Two general categories of observable responses can be distinguished: (a) those responses which take the form of overt changes in behavior and (b) those responses which take the form of less readily observable physiological changes. Such objective evidence, though it undoubtedly avoids the difficulties of the verbalization of subjective feelings and emotions, presents other difficulties no less perplexing.

In the first place, emotional responses need not result in overt, observable behavior. As Henry Aiken points out,[7] one of the special characteristics of our responses to aesthetic objects is the very fact that, due to our beliefs as to the nature of aesthetic experience, we tend to suppress overt behavior. Furthermore, as an important adjunct to this point it should be noted that emotion-felt or affect is most intense precisely in those cases where feeling does not result in or take the form of overt behavior or mental fantasy. This is clear as soon as one considers the tendency of human beings to "work off" or relieve emotional tension in physical effort and bodily behavior. In short, the absence of overt emotional behavior, particularly in response to aesthetic stimuli, is no indication as to either the presence or force of emotional responses.

However, even where overt behavior is present, its interpretation is difficult and problematical. When, on the one hand, overt behavior is the product of particularly powerful emotional tensions, it tends to be diffuse, generalized, or chaotic. Extreme conflict, for example, may result in either motionless rigidity or frenzied activity; weeping may be the product of either profound grief or extreme joy. Unless we have accurate knowledge of the stimulus situation beforehand, such behavior can tell us little or nothing as to the significance of the response or of its relation to the stimulus.

On the other hand, when emotional behavior does become differentiated it tends to be standardized—to become part of more general patterns of social behavior. Thus although the philosophical aspect of the stimulus situation, the fact that an aesthetic object is being considered, tends toward the suppression of overt behavior, the social aspect of the stimulus situation permits and at times indeed encourages certain standardized types of emotional behavior. This is apparent in the conduct of performers and audiences alike. The jazz performer and his audience, for example, have one mode of socially sanctioned emotional behavior; the concert performer and his audience have another. The difference between the two is more a matter of conventionally determined behavior patterns than it is a matter of musical differences.

Such behavior must be regarded at least in part as a means of communication rather than as a set of natural, reflex reactions. It indicates and designates not only appropriate mental sets but also the proper (i.e., socially acceptable) modes of response. Once this sort of behavior becomes habitual, and it does so very early in life, then it may be activated by the social aspects of the stimulus situation alone, without regard for the stimulus itself. In short, given no theory as to the relation of musical stimuli to affective responses, observed behavior can provide little information as to either the nature of the stimulus, the significance of the response, or the relation between them. For conduct which might to an observer appear to indicate the presence of an emotional response might in point of fact be the result of the subject's day dreams, his observation and imitation of the behavior of others, or his beliefs as to the kind of behavior appropriate and expected in the given social situation.

### *Objective Evidence: Physiological Responses*

On the physiological level music evokes definite and impressive responses. It "has a marked effect on pulse, respiration and external blood pressure. . . . [It] delays the onset of muscular fatigue . . . [and] has a marked effect upon the psychogalvanic reflex. . . ."[8] In spite of the fact that these changes are the very ones which normally accompany emotional experience, the significance of these data is not completely clear. Two principal difficulties are involved.

To begin with, no relation can be found between the character or pattern of the musical selection evoking the response and the particular physiological changes which take place. These changes appear to be completely independent of any particular style, form, medium, or general character. The same responses will take place whether the music is fast or slow, exciting or soothing, instrumental or vocal, classical or jazz.

Because tonal stimulation is a constant factor of all musical stimuli, Mursell is led to conclude that the power of "tone as such" must be the cause of the physiological changes observed.[9]

There is, however, another constant involved in the perception of music; namely, the mental attitude of the audience. The listener brings to the act of perception definite beliefs in the affective power of music. Even before the first sound is heard, these beliefs activate dispositions to respond in an emotional way, bringing expectant ideomotor sets into play. And it seems more reasonable to suppose that the physiological changes observed are a response to the listener's mental set rather than to assume that tone as such can, in some mysterious and unexplained way, bring these changes about directly. For while the relationship between mental sets and physiological changes has been demonstrated beyond doubt, the effect of "tone as such" has not.

This does not imply that the presence of a physiological environment, which is a necessary condition for the arousal of emotion, is not a significant fact. The existence of this necessary condition increases the likelihood that emotional responses do take place—a fact which some critics have sought to deny. What this analysis indicates is that not only are these physiological adjustments pre-emotional, as Mursell would admit, but they are also premusical.

Furthermore, even the conclusions just reached about the significance of the physiological data are probably an exaggeration, if not from a psychological point of view, at least, from a logical one. For such adjustments not only accompany affective responses, but they are also concomitants of clearly nonemotional responses.

In the light of present knowledge it seems clear that though physiological adjustments are probably necessary adjuncts of affective responses they cannot be shown to be sufficient causes for such responses and have, in fact, been able to

throw very little light upon the relationship between affective responses and the stimuli which produce them. The situation is concisely summarized by Rapaport:

> (a) On the basis of the material surveyed nothing can be definitely stated as to the relation to "emotion felt" of physiological processes concomitant with emotions. Proof has not been offered to show that the usually described physiological processes are *always* present when emotion is felt. (b) Nothing is known about the physiological processes *underlying* emotional experience. However, sufficient proof has been adduced that neither the James-Lange theory nor the hypothalamic theory explains the origin of "emotion felt." (c) The investigations into the physiology and the neural correlates of emotional expression are of importance; their relation to the psychic process designated as "emotion felt" is the crucial point of every theory of emotions. However, the knowledge concerning this relation is so scant that investigations into the influence of emotions on other physiological processes will have to be based rather on what is known about the psychology of emotions.[10]

There is one basic problem with all the objective data discussed: namely, that even when affective experiences result in objective adjustments, whether behavioral or physiological, what can be observed is not the emotion-felt, the affect, but only its adjuncts and concomitants, which in the case of behavior tend to become standardized and in the case of physiological changes are not specific to emotion. What we wish to consider, however, is that which is most vital and essential in emotional experience: the feeling-tone accompanying emotional experience, that is, the affect.

Here we face a dilemma. On the one hand, the response with which we are concerned is profoundly and permanently subjective and hence of necessity concealed from the scrutiny of even the most scrupulous observers; and, on the other hand, we have found that the subjective data available, taken by themselves, provide no definite and unequivocal information about the musical stimulus, the affective response, or the relation between them. This difficulty can be resolved only if the subjective data available, including the responses of the readers and the author of this study, can be examined, sifted, and studied in the light of a general hypothesis as to the nature of affective experience and the processes by which musical stimuli might arouse such experience.

Such a hypothesis is provided by the psychological theory of emotions. For although much work undoubtedly remains to be done in the field of emotional theory, there appears to be general agreement among psychologists and psychiatrists at least as to the conditions under which emotional responses arise and as to the relationship between the affective stimulus and the affective response.

## The Psychological Theory of Emotions

Since the physiological changes which accompany emotional experience, whatever their importance, do not provide a basis for differentiating affective from

nonaffective states, the differentia must be sought in the realm of mental activity.

However, not all mental responses are affective. We speak of dispassionate observation, calm deliberation, and cool calculation. These are nonemotional states of mind.

If we then ask what distinguishes nonemotional states from emotional ones, it is clear that the difference does not lie in the stimulus alone. The same stimulus may excite emotion in one person but not in another. Nor does the difference lie in the responding individual. The same individual may respond emotionally to a given stimulus in one situation but not in another. The difference lies in the relationship between the stimulus and the responding individual.

This relationship must first of all be such that the stimulus produces a tendency in the organism to think or act in a particular way. An object or situation which evokes no tendency, to which the organism is indifferent, can only result in a nonemotional state of mind.

But even when a tendency is aroused, emotion may not result. If, for example, a habitual smoker wants a cigarette and, reaching into his pocket, finds one, there will be no affective response. If the tendency is satisfied without delay, no emotional response will take place. If, however, the man finds no cigarette in his pocket, discovers that there are none in the house, and then remembers that the stores are closed and he cannot purchase any, he will very likely begin to respond in an emotional way. He will feel restless, excited, then irritated, and finally angry.

This brings us to the central thesis of the psychological theory of emotions. Namely: Emotion or affect is aroused when a tendency to respond is arrested or inhibited.

## *Supporting Theories*

In 1894 John Dewey set forth what has since become known as the conflict theory of emotions.[11]

In an article entitled "The Conflict Theory of Emotions,"[12] Angier shows that this general position has been adopted, in more or less modified form, by many psychologists of widely different viewpoints. For instance, the behaviorists, who emphasize the excitement and confusion which disrupt behavior as important characteristics of emotional conduct, would seem to be describing objectively what others view as the result of inner conflict. But the difficulty with examining emotions from the point of view of behaviorism is that, as we have seen, emotion may be felt without becoming manifest as overt behavior.

MacCurdy, whose own attitude is psychoanalytical, points out that it is precisely "when instinctive reactions are stimulated that do not gain expression either in conduct, emotional expression, or fantasy, that affect is most intense. It is the prevention of the expression of instinct either in behavior or conscious thought that leads to intense affect. In other words the energy of the organism, activating an instinct process, must be blocked by repression before poignant

feeling is excited."[13] MacCurdy's analysis involves three separate phases: (a) the arousal of nervous energy in connection with the instinct or tendency;[14] (b) the propensity for this energy to become manifest as behavior or conscious thought once the tendency is blocked; and (c) the manifestation of the energy as emotion-felt or affect if behavior and conscious thought are also inhibited. Of course, if the stimulation is so powerful that the total energy cannot be absorbed by either behavior or affect alone, both will result.[15]

It is obvious that a shift of emphasis has taken place in the statement of the theory of emotions. Dewey and his followers tended to stress the conflict or opposition of tendencies as being the cause of emotional response. MacCurdy and most of the more recent workers in the field believe that it is the blocking or inhibiting of a tendency which arouses affect. Actually the concept of conflict through the opposition of simultaneously aroused conflicting tendencies may be regarded as a special and more complicated case of the arrest of tendency.

This point was made in Paulhan's brilliant work, which in 1887, almost 10 years before Dewey's formulation, set forth a highly sophisticated theory of emotions. "If we ascend in the hierarchy of human needs and deal with desires of a higher order, we still find that they only give rise to affective phenomena when the tendency awakened undergoes inhibition."[16]

However, more complex phenomena are possible as the result of "the simultaneous or almost simultaneous coming into play of systems which tend toward opposite or different actions and which cannot both culminate in action at the same time; always provided that the psychical systems brought into play do not differ too widely in intensity. . . ."[17] Such a situation results, according to Paulhan, in an emotion or affect characterized by confusion and lack of clarity.

In other words, in one case a tendency is inhibited not by another opposed tendency but simply by the fact that for some reason, whether physical or mental, it cannot reach completion. This is the situation of the inveterate smoker in the example given earlier. In the other case two tendencies which cannot both reach fruition at the same time are brought into play almost simultaneously. If they are about equal in strength, each tendency will block the completion of the other. The result is not only affect, as a product of inhibition, but doubt, confusion, and uncertainty as well.

These latter concomitants of conflict are of importance because they may themselves become the basis for further tendencies. For to the human mind such states of doubt and confusion are abhorrent; and, when confronted with them, the mind attempts to resolve them into clarity and certainty, even if this means abandoning all other previously activated tendencies.

Thus confusion and lack of clarity, growing out of conflicting tendencies, may themselves become stimuli producing further tendencies—tendencies toward clarification—which may become independent of the originally conflicting tendencies. Such tendencies need not be definite in the sense that the ultimate resolution of the doubt and confusion is specified. Some resolution of the confusion may be more important than this or that particular solution, assuming that

the final result is not in conflict with other aspects of the stimulus situation or other mental sets.

Furthermore, it should be noted that uncertainty and lack of clarity may be products not only of conflicting tendencies but also of a situation which itself is structurally confused and ambiguous. This is of capital importance because it indicates that a situation which is structurally weak and doubtful in organization may directly create tendencies toward clarification. Delay in such a generalized tendency toward clarification may also give rise to affect.

Although the main tenets of the psychological theory of emotions have been widely accepted, there have, needless to say, been criticisms of the theory. In the main these have come from those who have sought, as yet without success, to account for, describe, and distinguish emotions in purely physiological terms. The theory of emotions, it is objected, does not tell us what an emotion is; it does not tell us precisely what takes place in the body to make us feel.

This objection, though valid, is irrelevant for our purposes. For just as the physicist long defined magnetism in terms of the laws of its operation and was able to deal with the phenomena without knowing the nature of the magnetic states so, too, the psychologist can define emotion in terms of the laws governing its operation, without stipulating precisely what, in physiological terms, constitutes feeling—what makes affect felt.

## *The Differentiation of Affect*

Thus far we have considered emotion as though it were a general, undifferentiated response, a feeling whose character and quality were always more or less the same. While there is a good deal of evidence for this view, it is nevertheless clear that in common speech and everyday experience we do recognize a variety of emotional states—love, fear, anger, jealousy, and the like.[18] The whole problem of whether undifferentiated feelings, affects per se, exist, of their relation to differentiated emotional experience, and of the basis for such differentiation is of importance in the present study. For while music theorists and aestheticians have found it difficult to explain how music designates particular emotions, they have found it almost impossible to account for the existence of less specific affective experience.

Were the evidence to show that each affect or type of affect had its own peculiar physiological composition, then obviously undifferentiated feeling would be out of the question. However, Woodworth's summary of the work in this field makes it clear that this is not the case.[19]

The evidence in the case of "emotional (affective) behavior" (the term which will henceforth be used to designate the overt and observable aspects of emotional conduct) is more complex. Much emotional behavior, though habitual and hence seemingly automatic and natural, is actually learned. Because this aspect of behavior serves in the main as a means of communication, it will be called "designative (denotative) behavior." To this category belong most of the postural

sets, facial expressions, and motor responses accompanying emotional behavior. Though designative behavior is definitely and clearly differentiated, the differentiation is not a necessary one and indicates nothing as to the possible differentiation of the affect itself.

Other aspects of affective behavior, such as skeletal and muscular adjustments, have been said to be automatic, natural concomitants of the affective response. These will be called "emotional reactions." Supposing that such automatic reactions do exist, a fact that has been debated, it has not been definitely shown that they are differentiated as between types of affective experience.

However, even if it were demonstrated that emotional reactions were differentiated, this would not necessarily prove or even indicate that the affects which they accompany are also differentiated. For the reaction is a response made to the total emotion-provoking situation and not necessarily a product of affect itself. In other words, it may well be that such automatic behavior is called forth by the peculiar nature of the objective situation rather than by the operation of the law of affect itself. Were this the case, such a reaction would be independent of affect and might indeed take place, as does designative emotional behavior, in the absence of affect.

The suppositions that behavior reactions are essentially undifferentiated, becoming characteristic only in certain stimulus situations, and that affect itself is basically undifferentiated are given added plausibility when one considers the following:

a. The more intense emotional behavior is, and presumably therefore the more intense the affective stimulation, the less the control exerted by the ego over behavior and the greater the probability that the behavior is automatic and natural.

b. The more intense affective behavior is, the less differentiated such behavior tends to be. In general, the total inhibition of powerful tendencies produces diffuse and characterless activity. For example, extreme conflict may result in either complete immobility or in frenzied activity, while weeping may accompany deepest grief, tremendous joy, or probably any particularly intense emotion.

c. Thus the more automatic affective behavior is, the less differentiated it tends to be.

It seems reasonable then to conclude that automatic reflex reactions not only fail to provide reasons for believing that affect itself is differentiated but the evidence seems to point to just the opposite conclusion.

Finally, our own introspective experience and the reports of the experiences of others testify to the existence of undifferentiated emotions. It is affect as such which Cassirer is discussing when he writes: "Art gives us the motions of the human soul in all their depth and variety. But the form, the measure and rhythm, of these motions is not comparable to any single state of emotion. What we feel in art is not a simple or single emotional quality. It is the dynamic process of life itself."[20]

The conclusion that affect itself is undifferentiated does not mean that affective experience is a kind of disembodied generality. For the affective experience, as distinguished from affect per se, includes an awareness and cognition of a stimulus situation which always involves particular responding individuals and specific stimuli.

Not only do we become aware of and know our own emotions in terms of a particular stimulus situation but we interpret and characterize the behavior of others in these terms. "When an organism is in a situation which results in a disturbed or wrought-up condition, then the situation plus the reaction gives us the name or word which characterizes the whole as a specific emotion. The reaction itself is not sufficient to differentiate the emotion, the character of the situation is involved in this differentiation."[21]

Thus while affects and emotions are in themselves undifferentiated, affective experience is differentiated because it involves awareness and cognition of a stimulus situation which itself is necessarily differentiated. The affective states for which we have names are grouped and named because of similarities of the stimulus situation, not because the affects of different groups are per se different. Love and fear are not different affects, but they are different affective experiences.

Awareness of the nature of the stimulus situation also seems to be the real basis for the distinction which Hebb draws between "pleasant" and "unpleasant" emotions. According to the present analysis, there are no pleasant or unpleasant emotions. There are only pleasant or unpleasant emotional experiences. This is of importance in understanding the distinction made by Hebb.

According to Hebb, the difference between pleasant and unpleasant emotions lies in the fact that pleasant emotions (or, in our terminology, pleasant emotional experiences) are always resolved. They depend "on first arousing apprehension, then dispelling it."[22] But were this actually the case we could only know whether an emotion were pleasant or unpleasant after it was over. Yet, surely, we know more than this while we are experiencing affect. The pleasantness of an emotion seems to lie not so much in the fact of resolution itself as in the belief in resolution—the knowledge, whether true or false, that there will be a resolution. It is not, as Hebb seems to assert when he cites as pleasurable the "mildly frustrating or the mildly fear-provoking,"[23] the control actually exercised over a situation which distinguishes pleasant from unpleasant emotions. It is the control which is believed to exist over the situation.

The sensation of falling through space, unconditioned by any belief or knowledge as to the ultimate outcome, will, for instance, arouse highly unpleasant emotions. Yet a similar fall experienced as a parachute jump in an amusement park may, because of our belief in the presence of control and in the nature of the resolution, prove most pleasurable.

The foregoing analysis is of genuine importance in the present study because it explains and accounts for the existence and nature of the intangible, nonreferential affective states experienced in response to music. For insofar as the

stimulus situation, the music, is nonreferential (in the sense that it pictures, describes, or symbolizes none of the actions, persons, passions, and concepts ordinarily associated with human experience), there is no reason to expect that our emotional experience of it should be referential. The affective experience made in response to music is specific and differentiated, but it is so in terms of the musical stimulus situation rather than in terms of extramusical stimuli.[24]

In the light of this discussion it is evident that, though it is wrong to assert, as some have done, that emotions exist which are *sui generis* musical or aesthetic, it is possible to contend that there are emotional experiences which are so.[25] By the same token, however, any number of emotional experiences can be grouped together so long as their stimulus situations are in some respects similar. Musical affective experiences, for example, might be differentiated into operatic, orchestral, baroque, and so forth. But the most significant distinction would still lie in the fact that musical stimuli, and hence musical affective experiences, are nonreferential.

## *Emotional Designation*

Although emotional behavior is frequently characterless and diffuse, often it is differentiated and intelligible. Even without knowledge of the stimulus situation, motor behavior, facial expression, tone of voice, and manner of speaking can tell us not only that an individual is responding in an emotional way but also something of the character of his feelings or, more accurately, of the character of his affective experience.

Differentiated behavior, as we have seen, is not an automatic or a necessary concomitant of affect itself or even of affective experience. The more automatic behavior is, the less likely it is to be differentiated. Differentiation involves control, and control implies purpose.

The purpose of emotionally differentiated behavior is communication. The individual responding, having an affective experience or simulating one, seeks to make others aware of his experience through a series of nonverbal behavioral signs. Because the gestures and signs which differentiate such behavior are purposeful, this mode of behavior will be called "emotional designation" or "designative behavior."[26]

Such signs not only act as cues for appropriate behavior in the social situation but are probably, at least in part, aimed at making other individuals respond in an empathetic way. As the saying goes: Misery loves company. And so do other emotional states. Not only do we dislike physical isolation, but we want to share our emotional life with others. And, indeed, such sharing does take place. For an observer, recalling a situation in his own experience similar to the one signified by the behavior of another, may respond to the remembered situation in an affective way. Though designative affective behavior may, through constant use, become habitual and automatic so that it is almost invariably called up as part of

the total emotional response, it is not basically a necessary concomitant of the response but one brought into play as a result of a desire to communicate.

Designative behavior is differentiated largely by custom and tradition. It varies from culture to culture and among different groups within a single culture. This does not mean that there are no features of such behavior which are natural and widespread. In all probability there are. However, three points should be kept in mind: (1) There is no real evidence to show that there is only one single natural mode of behavior relevant to a given stimulus situation. When alternative modes of behavior are possible, cultural selection probably determines the composition of any particular pattern of affective designation. (2) Whatever natural tendencies toward a particular pattern of behavior exist may be altered or suppressed by the demands of the larger behavior patterns of the culture. (3) Even where natural behavior is retained in the pattern of emotional behavior, it inevitably becomes codified and standardized for the sake of more efficient communication.

Above all, we understand and make appropriate responses to designative behavior as a total behavior pattern, not just to some features of it, whether natural or otherwise. As a total pattern designative behavior is a cultural phenomenon, not a natural one. It is, in the final analysis, learned.

This is important as it takes most of the sting out of the criticism that music which attempts to designate emotional states depends for its effect upon the learning of conventional signs and symbols. For this fact is not peculiar to music but is characteristic of all emotional designation. If one excludes such designation as a legitimate means of musical communication, one must by the same token exclude it as a means of human communication in general.[27]

## The Theory of Emotions Related to Musical Experience

### *An Assumption*

An examination of the psychological theory of emotions was made because the evidence furnished by the introspections of musicians, aestheticians, and listeners and the objective data gathered from the observation of behavior and the study of the physiological responses to musical stimuli did not yield reliable information about the musical stimulus or the affective responses made to it. Implicit in this examination was an assumption which must now be made explicit: Though the stimulus situation may vary indefinitely, the conditions which evoke affect are general and are, therefore, applicable to music. In other words, it was assumed that the law of affect, which states that emotion is evoked when a tendency to respond is inhibited, is a general proposition relevant to human psychology in all realms of experience.

This assumption does not, however, imply or stipulate that musical affective experiences are the same as the affective experiences made in response to other stimulus situations. Musical experience differs from nonmusical or, more specifically, nonaesthetic experience in three important ways.

First, as we have seen, affective experience includes an awareness and knowledge of the stimulus situation. This being so, the affective experience of music will differ from other types of affective experience, particularly insofar as musical stimuli are nonreferential.

Second, in everyday experience the tensions created by the inhibition of tendencies often go unresolved. They are merely dissipated in the press of irrelevant events. In this sense daily experience is meaningless and accidental. In art inhibition of tendency becomes meaningful because the relationship between the tendency and its necessary resolution is made explicit and apparent. Tendencies do not simply cease to exist: they are resolved, they conclude.[28]

Third, in life the factors which keep a tendency from reaching completion may be different in kind from those which activated the tendency in the first place. The stimulus activating a tendency may, for example, be a physical or psychic need of the organism, while the inhibiting factors may simply be a series of external circumstances which keep the organism from satisfying the need. This is the situation in the case of the habitual smoker who can find no cigarette. Or the situation may be reversed; that is, a tendency activated by an external stimulus may be inhibited by the psychic processes of the organism.

Furthermore, in everyday experience the resolutions of the tensions brought into play by inhibition may be irrelevant to the tendencies themselves. Tensions arising from psychic needs may be "worked off" in sheer physical activity which is without meaningful relation to the original stimulus or to the tendency itself.

In music, on the other hand, the same stimulus, the music, activates tendencies, inhibits them, and provides meaningful and relevant resolutions.

## *Tendency and Expectation in Music*

The assumption that the same basic psychological processes underlie all affective responses, whether the stimulus be musical or of some other kind, has been implicit in much musical theory and in the speculations of many aestheticians. But this does not in itself increase our understanding of the nature of musical experience and of the musical processes which form it. It does not explain the nature of the relationships which exist between the stimulus, the listener's perceptions and mental processes, and his responses. To do this it is necessary to demonstrate precisely how musical stimuli do, in fact, arouse and inhibit tendencies and thereby give rise to emotions.

What is meant by a tendency to respond? A "tendency" or, as MacCurdy uses the term, an "instinct" "is a pattern reaction that operates, or tends to operate, when activated, in an automatic way."[29] A pattern reaction consists of a set or series of regularly coincident mental or motor responses which, once brought

into play as part of the response to a given stimulus, follow a previously ordered course, unless inhibited or blocked in some way. The order established by a pattern reaction is both temporal and structural; that is, the series involves not only the relation of the parts of the total pattern to each other but also their timing. Thus a series may be disturbed either because the succession of the parts of the pattern is upset or because the timing of the series is upset or both.

The term "tendency," as used in this study, comprises all automatic response patterns, whether natural or learned. Since habit is "a mechanism of action, physiologically ingrained, which operates spontaneously and automatically,"[30] the term "tendency" also "includes habit reactions and, inevitably, acquired concepts and meanings."[31]

The tendency to respond may be either conscious or unconscious. If the pattern reaction runs its normal course to completion, then the whole process may be completely unconscious. Countless reaction patterns, of which the responding individual is unaware, are initiated and completed each hour. The more automatic behavior becomes, the less conscious it is. The tendency to respond becomes conscious where inhibition of some sort is present, when the normal course of the reaction pattern is disturbed or its final completion is inhibited. Such conscious and self-conscious tendencies are often thought of and referred to as "expectations."

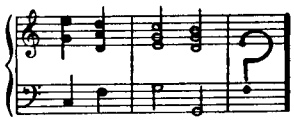
In a broader sense all tendencies, even those which never reach the level of consciousness, are expectations. For since a tendency is a kind of chain reaction in which a present stimulus leads through a series of adjustments to a more or less specified consequent, the consequent is always implied in the tendency, once the tendency has been brought into play. Thus while our conscious minds do not actively expect a consequent unless the pattern reaction is disturbed, our habits and tendencies are expectant in the sense that each seeks out or "expects" the consequents relevant and appropriate to itself. Though he may never become aware of his expectations as he reaches in his pocket for a pack of cigarettes, the behavior of the habitual smoker shows that he does expect or, perhaps more accurately, his habits expect for him.

If tendencies are pattern reactions that are expectant in the broad sense, including unconscious as well as conscious anticipations, then it is not difficult to see how music is able to evoke tendencies. For it has been generally acknowledged that music arouses expectations, some conscious and others unconscious, which may or may not be directly and immediately satisfied.

> . . . the pleasure . . . arises from the perception of the artist's play with forms and conventions which are ingrained as *habits* of perception both in the artist and his audience. Without such habits . . . there would be no awareness whatever of the artist's fulfillment of and subtle departures from established forms. . . . But the pleasure which we derive from style is not an intellectual interest in detecting similarities and differences, but an immediate aesthetic delight in perception which results from the arousal and suspension or fulfillment of expectations which are the products of many previous encounters with works of art.[32]

### *Expectation, Suspense, and the Unexpected*

Sometimes a very specific consequent is expected. In Western music of the eighteenth century, for example, we expect a specific chord, namely, the tonic (C major), to follow this sequence of harmonies (see example). Furthermore, the consequent chord is expected to arrive at a particular time, i.e., on the first beat of the next measure.



Of course, the consequent which is actually forthcoming, though it must be possible within the style, need not be the one which was specifically expected. Nor is it necessary that the consequent arrive at the expected time. It may arrive too soon or it may be delayed. But no matter which of these forms the consequent actually takes, the crucial point to be noted is that the ultimate and particular effect of the total pattern is clearly conditioned by the specificity of the original expectation.

At other times, expectation is more general; that is, though our expectations may be definite, in the sense of being marked, they are nonspecific, in that we are not sure precisely how they will be fulfilled. The antecedent stimulus situation may be such that several consequents may be almost equally probable. For instance, after a melodic fragment has been repeated several times, we begin to expect a change and also the completion of the fragment. A change is expected because we believe that the composer is not so illogical as to repeat the figure indefinitely and because we look forward to the completion of the incomplete figure. But precisely what the change will be or how the completion will be accomplished cannot perhaps be anticipated. The introductions to many movements written in the eighteenth or nineteenth centuries create expectation in this way, e.g., the opening measures of Beethoven's Ninth Symphony or the opening measures of the "March to the Gallows" from Berlioz's *Symphonie Fantastique.*

Expectation may also result because the stimulus situation is doubtful or ambiguous. If the musical patterns are less clear than expected, if there is confusion as to the relationship between melody and accompaniment, or if our expectations are continually mistaken or inhibited, then doubt and uncertainty as to the general significance, function, and outcome of the passage will result. As we have already seen, the mind rejects and reacts against such uncomfortable states and, if they are more than momentary, looks forward to and expects a return to the certainty of regularity and clarity. This is particularly striking in the responses made to works of art where, because of a firm belief in the purposeful-

ness and integrity of the artist, we expect that order will in the end triumph, and precision will replace ambiguity.

However, the manner in which clarification and order will be restored may not be predicted or envisaged. Expectation is not specific; the state is one of suspense. In fact, if doubt and uncertainty are strong enough, almost any resolution, within the realm of probability, which returns us to certainty will be acceptable, though no doubt some resolutions will, given the style, seem more natural than others.

The inclusion of suspense arising out of uncertainty may, at first sight, appear to be an extension and amplification of the concept of arrest and inhibition of a tendency. But when the matter is considered more carefully, it will be seen that every inhibition or delay creates uncertainty or suspense, if only briefly, because in the moment of delay we become aware of the possibility of alternative modes of continuation. The difference is one of scale and duration, not of kind. Both arouse uncertainties and anxieties as to coming events.

Suspense is essentially a product of ignorance as to the future course of events. This ignorance may arise either because the present course of events, though in a sense understandable in itself, presents several alternative and equally probable consequents or because the present course of events is itself so unusual and upsetting that, since it cannot be understood, no predictions as to the future can be made.

From the outset ignorance arouses strong mental tendencies toward clarification which are immediately affective. If ignorance persists in spite of all, then the individual is thrown into a state of doubt and uncertainty. He commences to sense his lack of control over the situation, has inability to act on the basis of the knowledge which he supposed that he possessed. In short, he begins to feel apprehensive, even fearful, though there is no object for his fear. Ignorance and its concomitant feelings of impotence breed apprehension and anxiety, even in music. But ignorance also gives rise to more sanguine feelings; for since the outcome cannot be envisaged, it may be pleasant. These feelings are themselves tendencies (the avoidance of painful apprehension and the expectation of a propitious conclusion) which become focused upon an expected resolution of the unpleasant stimulus situation.

The longer doubt and uncertainty persist, the greater the feeling of suspense will tend to be. The stimulus situation creating doubt and uncertainty must, of course, be progressively intensified if suspense is to be maintained or increased. For as we become accustomed to a given stimulus situation, even an unpleasant one, its effectiveness tends to diminish. Moreover, without a change in the stimulus situation in the direction of complication and uncertainty, those vital anticipatory feelings (that a break must come, that doubt and perplexity must give way to knowledge) which make us expect (both apprehensively and hopefully) would be lost.

The greater the buildup of suspense, of tension, the greater the emotional release upon resolution. This observation points up the fact that in aesthetic

experience emotional pattern must be considered not only in terms of tension itself but also in terms of the progression from tension to release. And the experience of suspense is aesthetically valueless unless it is followed by a release which is understandable in the given context.

Musical experiences of suspense are very similar to those experienced in real life. Both in life and in music the emotions thus arising have essentially the same stimulus situation: the situation of ignorance, the awareness of the individual's impotence and inability to act where the future course of events is unknown. Because these musical experiences are so very similar to those existing in the drama and in life itself, they are often felt to be particularly powerful and effective.

Musical suspense seems to have direct analogies in experience in general; it makes us feel something of the insignificance and powerlessness of man in the face of the inscrutable workings of destiny. The low, foreboding rumble of distant thunder on an oppressive summer afternoon, its growing intensity as it approaches, the crescendo of the gradually rising wind, the ominous darkening of the sky, all give rise to an emotional experience in which expectation is fraught with powerful uncertainty—the primordial and poignant uncertainty of human existence in the face of the inexorable forces of nature. With mixed feelings of hope and apprehension in the presence of the unknown, we anxiously await the breaking of the storm, the discovery of what unrelenting fate has decreed.

Similarly in music the state of suspense involves an awareness of the powerlessness of man in the face of the unknown.

What is expected in this state of suspense may not be specified, but this does not mean that any consequent is possible. Our expectations are inevitably circumscribed by the possibilities and probabilities of the style of the composition in question. The consequent must, given the circumstances, be possible within what Aiken has called "an ordering system of beliefs and attitudes."[33]

Although the consequent in any musical sequence must, in this sense, be possible, it may nevertheless be unexpected. But the unexpected should not be confused with the surprising. For when expectation is aroused, the unexpected is always considered to be a possibility, and, though it remains the less expected of several alternatives, it is not a complete surprise. Conditions of active expectation (especially general expectation and suspense) are not the most favorable to surprise. For the listener is on guard, anticipating a new and possibly unexpected consequent. Surprise is most intense where no special expectation is active, where, because there has been no inhibition of a tendency, continuity is expected.

As soon as the unexpected, or for that matter the surprising, is experienced, the listener attempts to fit it into the general system of beliefs relevant to the style of the work. This requires a very rapid re-evaluation of either the stimulus situation itself or its cause—the events antecedent to the stimulus. Or it may require a review of the whole system of beliefs that the listener supposed appropriate and relevant to the work. If this mental synthesis does not take place

immediately, three things may happen: (1) The mind may suspend judgment, so to speak, trusting that what follows will clarify the meaning of the unexpected consequent. (2) If no clarification takes place, the mind may reject the whole stimulus and irritation will set in.[34] (3) The unexpected consequent may be seen as a purposeful blunder. Whether the listener responds in the first or third manner will depend partly upon the character of the piece, its mood or designative content. The third response might well be made to music whose character was comic or satirical. Beckmesser's music in Wagner's *Die Meistersinger* would probably elicit this type of interpretive understanding.[35] In a piece whose character admitted no such purposeful blunders, the second response would probably be elicited.

### *Conscious and Unconscious Expectations*

In the light of these observations it is clear that an expectation is not a blind, unthinking conditioned reflex. Expectation frequently involves a high order of mental activity. The fulfillment of a habit response, in art as well as in daily life, requires judgment and cognition both of the stimulus itself and of the situation in which it acts. The stimulus as a physical thing becomes a stimulus in the world of behavior only insofar as the mind of the perceiver is able to relate it, on the one hand, to the habit responses which the perceiver has developed and, on the other hand, to the particular stimulus situation. This is clear as soon as one considers that the same physical stimulus may call forth different tendencies in different stylistic contexts or in different situations within one and the same stylistic context. For example, a modal cadential progression will arouse one set of expectations in the musical style of the sixteenth century and quite another in the style of the nineteenth century. Likewise the same musical progression will evoke one set of expectations at the beginning of a piece and another at the end.

Expectation then is a product of the habit responses developed in connection with particular musical styles and of the modes of human perception, cognition, and response—the psychological laws of mental life.[36]

The mental activity involved in the perception of and response to music need not, however, be conscious. ". . . the intellectual satisfaction which the listener derives from continually following and anticipating the composer's intentions—now, to see his expectations fulfilled, and now, to see himself agreeably mistaken . . . this intellectual flux and reflux, this perpetual giving and receiving takes place unconsciously, and with the rapidity of lightning flashes"[37] So long as expectations are satisfied without delay, so long as tendencies are uninhibited, though intelligence is clearly and necessarily involved in the perception and understanding of the stimulus situation, the response will probably remain unconscious.

Mental activity tends to become conscious when reflection and deliberation are involved in the completion of the response pattern, that is, when automatic behavior is disturbed because a tendency has been inhibited. "Impulsion forever

boosted on its forward way would run its course thoughtless, and dead to emotion. . . . The only way it can become aware of its nature and its goal is by obstacles surmounted and means employed."[38]

But even when a habit response is inhibited, conscious awareness of the mental activity involved in the perception of and response to the stimulus situation is by no means inevitable. Intellectual experience (the conscious awareness of one's own expectations or, objectively, of the tendencies of the music), as distinguished from intellectual activity, is largely a product of the listener's own attitude toward his responses and hence toward the stimuli and mental activities which bring them into existence. That is to say, some listeners, whether because of training or natural psychological inclination, are disposed to rationalize their responses, to make experience self-conscious; others are not so disposed. If intellectual activity is allowed to remain unconscious, then the mental tensions and the deliberations involved when a tendency is inhibited are experienced as feeling or affect rather than as conscious cognition.

Having shown that music arouses tendencies and thus fulfils the conditions necessary for the arousal of affect and having demonstrated how this is accomplished, we can now state one of the basic hypotheses of this study. Namely: Affect or emotion-felt is aroused when an expectation—a tendency to respond—activated by the musical stimulus situation, is temporarily inhibited or permanently blocked.

As noted earlier, in musical experience the same stimulus, the music, activates tendencies, inhibits them, and provides meaningful and relevant resolutions for them. This is of particular importance from a methodological standpoint. For it means that granted listeners who have developed reaction patterns appropriate to the work in question, the structure of the affective response to a piece of music can be studied by examining the music itself.

Once those sound successions common to a culture, a style, or a particular work have been ascertained, then, if the customary succession is presented and completed without delay, it can be assumed that, since no tendency would have been inhibited, the listener would not respond in an affective way. If, on the other hand, the sound succession fails to follow its customary course, or if it involves obscurity or ambiguity, then it can be assumed that the listener's tendencies would be inhibited or otherwise upset and that the tensions arising in this process would be experienced as affect, provided that they were not rationalized as conscious intellectual experience.

In other words, the customary or expected progression of sounds can be considered as a norm, which from a stylistic point of view it is; and alteration in the expected progression can be considered as a deviation. Hence deviations can be regarded as emotional or affective stimuli.

The importance of this "objective" point of view of musical experience is clear. It means that once the norms of a style have been ascertained, the study and analysis of the affective content of a particular work in that style can be made

without continual and explicit reference to the responses of the listener or critic. That is, subjective content can be discussed objectively.[39]

## The Meaning of Music

### *The Problem of Meaning in Music*

The meaning of music has of late been the subject of much confused argument and controversy. The controversy has stemmed largely from disagreements as to what music communicates, while the confusion has resulted for the most part from a lack of clarity as to the nature and definition of meaning itself.

The debates as to what music communicates have centered around the question of whether music can designate, depict, or otherwise communicate referential concepts, images, experiences, and emotional states. This is the old argument between the absolutists and the referentialists.

Because it has not appeared problematical to them, the referentialists have not as a rule explicitly considered the problem of musical meaning. Musical meaning according to the referentialists lies in the relationship between a musical symbol or sign and the extramusical thing which it designates.

Since our concern in this study is not primarily with the referential meaning of music, suffice it to say that the disagreement between the referentialists and the absolutists is, as was pointed out at the beginning of this chapter, the result of a tendency toward philosophical monism rather than the result of any logical incompatibility. Both designative and nondesignative meanings arise out of musical experience, just as they do in other types of aesthetic experience.

The absolutists have contended that the meaning of music lies specifically, and some would assert exclusively, in the musical processes themselves. For them musical meaning is nondesignative. But in what sense these processes are meaningful, in what sense a succession or sequence of nonreferential musical stimuli can be said to give rise to meaning, they have been unable to state with either clarity or precision. They have also failed to relate musical meaning to other kinds of meaning—to meaning in general. This failure has led some critics to assert that musical meaning is a thing apart, different in some unexplained way from all other kinds of meaning. This is simply an evasion of the real issue. For it is obvious that if the term "meaning" is to have any signification at all as applied to music, then it must have the same signification as when applied to other kinds of experience.

Without reviewing all the untenable positions to which writers have tenaciously adhered, it seems fair to say that much of the confusion and uncertainty as to the nature of nonreferential musical meaning has resulted from two fallacies. On the one hand, there has been a tendency to locate meaning exclusively in one aspect of the communicative process; on the other hand, there has been a

propensity to regard all meanings arising in human communication as designative, as involving symbolism of some sort.

Since these difficulties can be best resolved in the light of a general definition of meaning, let us begin with such a definition: ". . . anything requires meaning if it is connected with, or indicates, or refers to, something beyond itself, so that its full nature points to and is revealed in that connection."[40]

Meaning is thus not a property of things. It cannot be located in the stimulus alone. The same stimulus may have many different meanings. To a geologist a large rock may indicate that at one time a glacier began to recede at a given spot; to a farmer the same rock may point to the necessity of having the field cleared for plowing; and to the sculptor the rock may indicate the possibility of artistic creation. A rock, a word, or motion in and of itself, merely as a stimulus, is meaningless.

Thus it is pointless to ask what the intrinsic meaning of a single tone or a series of tones is. Purely as physical existences they are meaningless. They become meaningful only insofar as they point to, indicate, or imply something beyond themselves.

Nor can meaning be located exclusively in the objects, events, or experiences which the stimulus indicates, refers to, or implies. The meaning of the rock is the product of the relationship between the stimulus and the thing it points to or indicates.

Though the perception of a relationship can only arise as the result of some individual's mental behavior, the relationship itself is not to be located in the mind of the perceiver. The meanings observed are not subjective. Thus the relationships existing between the tones themselves or those existing between the tones and the things they designate or connote, though a product of cultural experience, are real connections existing objectively in culture.[41] They are not arbitrary connections imposed by the capricious mind of the particular listener.

Meaning, then, is not in either the stimulus, or what it points to, or the observer. Rather it arises out of what both Cohen and Mead have called the "triadic" relationship between (1) an object or stimulus; (2) that to which the stimulus points—that which is its consequent; and (3) the conscious observer.

Discussions of the meaning of music have also been muddled by the failure to state explicitly what musical stimuli indicate or point to. A stimulus may indicate events or consequences which are different from itself in kind, as when a word designates or points to an object or action which is not itself a word. Or a stimulus may indicate or imply events or consequences which are of the same kind as the stimulus itself, as when a dim light on the eastern horizon heralds the coming of day. Here both the antecedent stimulus and the consequent event are natural phenomena. The former type of meaning may be called designative, the latter embodied.

Because most of the meanings which arise in human communication are of the designative type, employing linguistic signs or the iconic signs of the plastic

arts, numerous critics have failed to realize that this is not necessarily or exclusively the case. This mistake has led even avowed absolutists to allow designation to slip in through the secret door of semantic chicanery.[42]

But even more important than designative meaning is what we have called embodied meaning. From this point of view what a musical stimulus or a series of stimuli indicate and point to are not extramusical concepts and objects but other musical events which are about to happen. That is, one musical event (be it a tone, a phrase, or a whole section) has meaning because it points to and makes us expect another musical event. This is what music means from the viewpoint of the absolutist.

## *Music and Meaning*

Embodied musical meaning is, in short, a product of expectation. If, on the basis of past experience, a present stimulus leads us to expect a more or less definite consequent musical event, then that stimulus has meaning.[43]

From this it follows that a stimulus or gesture which does not point to or arouse expectations of a subsequent musical event or consequent is meaningless. Because expectation is largely a product of stylistic experience, music in a style with which we are totally unfamiliar is meaningless.[44]

However, once the aesthetic attitude has been brought into play, very few gestures actually appear to be meaningless so long as the listener has some experience with the style of the work in question. For so long as a stimulus is possible within any known style, the listener will do his best to relate it to the style, to understand its meaning.

In and of themselves, for example, the opening chords of Beethoven's Third Symphony have no particular musical stylistic tendency. They establish no pattern of motion, arouse no tensions toward a particular fulfillment. Yet as part of the total aesthetic cultural act of attention they are meaningful. For since they are the first chords of a piece, we not only expect more music but our expectations are circumscribed by the limitations of the style which we believe the piece to be in and by the psychological demand for a more palpable pattern.

Thus the phrase "past experience," used in the definition of meaning given above, must be understood in a broad sense. It includes the immediate past of the particular stimulus or gesture, that which has already taken place in this particular work to condition the listener's opinion of the stimulus and hence his expectations as to the impending, consequent event. In the example given above, the past was silence. But this fact of the past is just as potent in conditioning expectation as a whole section of past events.[45] The phrase "past experience" also refers to the more remote, but ever present, past experience of similar musical stimuli and similar musical situations in other works. That is, it refers to those past experiences which constitute our sense and knowledge of style. The phrase also comprehends the dispositions and beliefs which the listener brings to the musical experience as well as the laws of mental behavior which govern his

organization of stimuli into patterns and the expectations aroused on the basis of those patterns.

The words "consequent musical event" must be understood to include: (1) those consequents which are envisaged or expected; (2) the events which do, in fact, follow the stimulus, whether they were the ones envisaged or not; and (3) the more distant ramifications or events which, because the total series of gestures is presumed to be causally connected, are considered as being the later consequences of the stimulus in question. Seen in this light, the meaning of the stimulus is not confined to or limited by the initial triadic relationship out of which it arises. As the later stages of the musical process establish new relationships with the stimulus, new meanings arise. These later meanings coexist in memory with the earlier ones and, combining with them, constitute the meaning of the work as a total experience.

In this development three stages of meaning may be distinguished.

"Hypothetical meanings" are those which arise during the act of expectation. Since what is envisaged is a product of the probability relationships which exist as part of style, and since these probability relationships always involve the possibility of alternative consequences, a given stimulus invariably gives rise to several alternative hypothetical meanings. One consequent may, of course, be so much more probable than any other that the listener, though aware of the possibility of less likely consequences, is really set and ready only for the most probable. In such a case hypothetical meaning is without ambiguity. In other cases several consequents may be almost equally probable, and, since the listener is in doubt as to which alternative will actually materialize, meaning is ambiguous, though not necessarily less forceful and marked.[46]

Though the consequent which is actually forthcoming must be possible within the style, it may or may not be one of those which was most probable. Or it may arrive only after a delay or a deceptive diversion through alternative consequences. But whether our expectations are confirmed or not, a new stage of meaning is reached when the consequent becomes actualized as a concrete musical event.

"Evident meanings" are those which are attributed to the antecedent gesture when the consequent becomes a physicopsychic fact and when the relationship between the antecedent and consequent is perceived. Since the consequent of a stimulus itself becomes a stimulus with consequents, evident meaning also includes the later stages of musical development which are presumed to be the products of a chain of causality. Thus in the following sequence, where a stimulus (S) leads to a consequent (C), which is also a stimulus that indicates and is actualized in further consequents,

$$S_1 \ldots\ldots C_1S_2 \ldots\ldots C_1S_3 \ldots\ldots \text{etc.}$$

evident meaning arises not only out of the relationship between $S_1$ and $C_1$ but also out of the relationships between $S_1$ and all subsequent consequences, insofar

as these are considered to issue from $S_1$. It is also important to realize that the motion $S_1 \ldots\ldots C_1$ may itself become a gesture that gives rise to envisaged and actual consequents and hence becomes a term or gesture on another level of triadic relationships. In other words, both evident and hypothetical meanings come into being and exist on several architectonic levels.

Evident meaning is colored and conditioned by hypothetical meaning. For the actual relationship between the gesture and its consequent is always considered in the light of the expected relationship. In a sense the listener even revises his opinion of the hypothetical meaning when the stimulus does not move to the expected consequent.

"Determinate meanings" are those meanings which arise out of the relationships existing between hypothetical meaning, evident meaning, and the later stages of the musical development. In other words, determinate meaning arises only after the experience of the work is timeless in memory, only when all the meanings which the stimulus has had in the particular experience are realized and their relationships to one another comprehended as fully as possible.

## *The Objectification of Meaning*

A distinction must be drawn between the understanding of musical meaning which involves the awareness of the tendencies, resistances, tensions, and fulfilments embodied in a work and the self-conscious objectification of that meaning in the mind of the individual listener. The former may be said to involve a meaningful experience, the latter involves knowing what that meaning is, considering it as an objective thing in consciousness.

The operation of intelligence in listening to music need never become self-conscious. We are continually behaving in an intelligent way, comprehending meanings and acting upon our perceptions, cognitions, and evaluations without ever making the meanings themselves the objects of our scrutiny—without ever becoming self-conscious about what experience means. What Bertrand Russell says about understanding language also applies to the understanding of music: "Understanding language is . . . like understanding cricket: it is a matter of habits acquired in oneself and rightly presumed in others."[47]

Meanings become objectified only under conditions of self-consciousness and when reflection takes place. "One attains self-consciousness only as he takes, or finds himself stimulated to take, the attitude of the other."[48] Though training may make for a generally self-conscious attitude, one is stimulated to take the attitude of the other when the normal habits of response are disturbed in some way; when one is driven to ask one's self: What does this mean, what is the intention of this passage? Reflection is likewise brought into play where some tendency is delayed, some pattern of habitual behavior disturbed. So long as behavior is automatic and habitual there is no urge for it to become self-conscious, though it may become so. If meaning is to become objectified at all, it will as a rule become so when difficulties are encountered that make normal,

automatic behavior impossible. In other words, given a mind disposed toward objectification, meaning will become the focus of attention, an object of conscious consideration, when a tendency or habit reaction is delayed or inhibited.

## *Meaning and Affect*

It thus appears that the same processes which were said to give rise to affect are now said to give rise to the objectification of embodied meaning.

But this is a dilemma only so long as the traditional dichotomy between reason and emotion and the parent polarity between mind and body are adopted. Once it is recognized that affective experience is just as dependent upon intelligent cognition as conscious intellection, that both involve perception, taking account of, envisaging, and so forth, then thinking and feeling need not be viewed as polar opposites but as different manifestations of a single psychological process.

There is no diametric opposition, no inseparable gulf, between the affective and the intellectual responses made to music. Though they are psychologically differentiated as responses, both depend upon the same perceptive processes, the same stylistic habits, the same modes of mental organization; and the same musical processes give rise to and shape both types of experience. Seen in this light, the formalist's conception of musical experience and the expressionist's conception of it appear as complementary rather than contradictory positions. They are considering not different processes but different ways of experiencing the same process.

Whether a piece of music gives rise to affective experience or to intellectual experience depends upon the disposition and training of the listener. To some minds the disembodied feeling of affective experience is uncanny and unpleasant and a process of rationalization is undertaken in which the musical processes are objectified as conscious meaning. Belief also probably plays an important role in determining the character of the response. Those who have been taught to believe that musical experience is primarily emotional and who are therefore disposed to respond affectively will probably do so. Those listeners who have learned to understand music in technical terms will tend to make musical processes an object of conscious consideration. This probably accounts for the fact that most trained critics and aestheticians favor the formalist position. Thus while the trained musician consciously waits for the expected resolution of a dominant seventh chord the untrained, but practiced, listener feels the delay as affect.

## *Music and Communication*

Meanings and affects may, however, arise without communication taking place. Individual *A* observes another individual *B* wink and interprets the wink as a friendly gesture. It has meaning for *A* who observes it. But if the wink was not

intentional—if, for instance, *B* simply has a nervous tic—then no communication has taken place, for to *B* the act had no meaning. Communication, as Mead has pointed out, takes place only where the gesture made has the same meaning for the individual who makes it that it has for the individual who responds to it.[49]

It is this internalization of gestures, what Mead calls "taking the attitude of the other"[50] (the audience), which enables the creative artist, the composer, to communicate with listeners. It is because the composer is also a listener that he is able to control his inspiration with reference to the listener.[51] For instance, the composer knows how the listener will respond to a deceptive cadence and controls the later stages of the composition with reference to that supposed response. The performer too is continually "taking the attitude of the other"—of the listener. As Leopold Mozart puts it, the performer "must play everything in such a way that he will himself be moved by it."[52]

It is precisely because he is continually taking the attitude of the listener that the composer becomes aware and conscious of his own self, his ego, in the process of creation. In this process of differentiation between himself as composer and himself as audience, the composer becomes self-conscious and objective.[53]

But though the listener participates in the musical process, assuming the role which the composer envisaged for him, and though he must, in some sense, create his own experience, yet he need not take the attitude of the composer in order to do so. He need not ask: How will someone else respond to this stimulus? Nor is he obliged to objectify his own responses, to ask, How am I responding? Unlike the composer, the listener may and frequently does "lose himself in the music"; and, in following and responding to the sound gestures made by the composer, the listener may become oblivious of his own ego, which has literally become one with that of the music.

We must, then, be wary of easy and high-sounding statements to the effect that "we cannot understand a work of art without, to a certain degree, repeating and reconstructing the creative process by which it has come into being."[54] Certainly the listener must respond to the work of art as the artist intended, and the listener's experience of the work must be similar to that which the composer envisaged for him. But this is a different thing from experiencing the "creative process which brought it into being."

However, the listener may take the attitude of the composer. He may be self-conscious in the act of listening. Those trained in music, and perhaps those trained in the other arts as well, tend, because of the critical attitudes which they have developed in connection with their own artistic efforts, to become self-conscious and objective in all their aesthetic experiences. And it is no doubt partly for this reason that, as noted above, trained musicians tend to objectify meaning, to consider it as an object of conscious cognition.

Finally, and perhaps most important of all, this analysis of communication emphasizes the absolute necessity of a common universe of discourse in art. For without a set of gestures common to the social group, and without common

habit responses to those gestures, no communication whatsoever would be possible. Communication depends upon, presupposes, and arises out of the universe of discourse which in the aesthetics of music is called style.

NOTES

1. Norman Cazden, "Musical Consonance and Dissonance: A Cultural Criterion," *Journal of Aesthetics*, IV (1945), 3–11.

2. Paul R. Farnsworth, "Sacred Cows in the Psychology of Music," *Journal of Aesthetics*, VII (1948), 48–51.

3. Susanne K. Langer, *Philosophy in a New Key* (New York: Mentor Book Co., 1951).

4. *Ibid.*, p. 171.

5. H. P. Weld, "An Experimental Study in Musical Enjoyment," *American Journal of Psychology*, XXIII (1912), 283.

6. C. S. Myers, "Individual Differences in Listening to Music," in *The Effects of Music*, ed. Max Schoen (New York: Harcourt, Brace & Co., 1927), p. 14.

7. See H. D. Aiken, "The Aesthetic Relevance of Belief," *Journal of Aesthetics*, IX (1950), 301–15.

8. James L. Mursell, *The Psychology of Music* (New York: W. W. Norton & Co., Inc., 1937), pp. 27–28.

9. *Ibid.*, p. 37.

10. David Rapaport, *Emotions and Memory* (New York: International Universities Press Inc., 1950), p. 21.

11. John Dewey, "The Theory of Emotion," *Psychological Review*, I (1894), 553–69; II (1895), 13–32.

12. R. P. Angier, "The Conflict Theory of Emotions," *American Journal of Psychology*, XXXIX (1927), 390–401.

13. J. T. MacCurdy, *The Psychology of Emotion* (New York: Harcourt, Brace & Co., 1925), p. 475.

14. For MacCurdy the term "instinct" includes learned habit responses.

15. Notice that this analysis makes apparent the great significance of Aiken's contention that our beliefs as to the nature of aesthetic experience lead to the suppression of overt responses; for such inhibiting of overt behavior tends to intensify the affective response.

16. F. Paulhan, *The Laws of Feeling*, trans. C. K. Ogden (New York: Harcourt, Brace & Co., 1930), p. 19.

17. *Ibid.*, p. 123.

18. The term affective or emotional "state" will henceforth be used to designate those aspects of emotional experience which have been given names and which are in one way or another fairly standardized in a broad sense.

19. See Robert S. Woodworth, "How Emotions Are Identified and Classified," in *Feelings and Emotions: The Wittenberg Symposium*, ed. M. L. Reymert (Worcester, Mass.: Clark University Press, 1928), p. 224.

20. Ernst Cassirer, *An Essay on Man: An Introduction to a Philosophy of Human Culture* (New York: Doubleday & Co., 1953), p. 190. This admirable statement like so many of its kind suffers at the end from an irritating vagueness in which an intangible "the

dynamic process of life itself" is substituted for a definite account of how and why the emotions of art are not comparable to any single state of emotion. It is for a solution to this problem that we are searching in the present discussion of emotional differentiation.

21. C. Landis, "Studies in Emotional Reactions: II, General Behavior and Facial Expression," *Journal of Comparative Psychology*, IV (1924), 496.

22. Donald O. Hebb, *The Organization of Behavior* (New York: John Wiley & Sons, 1952), p. 258.

23. *Ibid.*, p. 232.

24. This statement must be qualified by the reservation that insofar as it can designate or represent extramusical stimuli, music can be said to evoke such affective states as are normally connected with the situations represented.

25. It is also clear that since the world of emotions is not composed of a series of separate compartments, a given listener may feel that a purely musical emotion is comparable or analogous to affects experienced in real life.

26. The term "emotional expression" is misleading in that it implies that such behavior is the direct, necessary expression of affect.

27. Though not within the province of this study, it can, I believe, be shown that similar aspects of experience are involved in musical and other designation. Both, for example, utilize the generality of motion (fast or slow, continuous or interrupted, smooth or disjunct, intense or weak) in such designation. And musical designation, though probably in some respects natural, is, like designative behavior, in the last analysis a product of culture and learning rather than a product of nature.

28. See John Dewey, *Art as Experience* (New York: Minton, Balch & Co., 1934), pp. 35, 56.

29. MacCurdy, *op. cit.*, p. 556.

30. John Dewey, *Intelligence in the Modern World*, ed. J. Ratner ("Modern Library" [New York: Random House, 1939]), p. 733.

31. MacCurdy, *op. cit.*, p. 556.

32. Aiken, *op. cit.*, p. 313; also see Arthur D. Bissell, *The Role of Expectation in Music* (New Haven: Yale University Press, 1921), p. vii; and Hugo Riemann, *Catechism of Musical Aesthetics*, trans. H. Bewerung (London: Augener & Co., n.d.), p. 29.

33. Aiken, *op. cit.*, p. 305.

34. If this takes place, the listener may shift his attention to another aspect of the musical materials, or he may simply abandon the attempt to make sense of the music altogether.

35. Thus the designation of mood and character, whether accomplished in purely musical terms or with the aid of a program or text, is important not only for its own sake, as a source of enjoyment, but also because, as part of the stimulus situation, it is necessary for the proper understanding of the musical processes in progress.

36. Both these aspects of the process of expectation are discussed in later chapters of *Emotion and Meaning in Music*, where much of the preceding discussion is treated in more detail.

37. Eduard Hanslick, *The Beautiful in Music*, trans. E. Cohen (London: Novello, Ewer & Co., 1901), p. 135. The difficulty with this statement is that Hanslick confuses intellectual satisfaction with intellectual activity. For although intellectual activity, in the sense of mental awareness and cognition, may, as we shall see, be unconscious, intellectual satisfaction implies a self-conscious awareness of the activity taking place.

38. Dewey, *Art as Experience*, p. 59; also see *Intelligence in the Modern World*, pp. 755 ff. Robert Penn Warren writes to much the same effect: ". . . a poem, to be good, must earn itself. It is a motion toward a point of rest, but if it is not a resisted motion, it is a motion of no consequence" ("Pure and Impure Poetry," *Kenyon Review*, V [1943], 251).

39. It is clear that the terms "norm" and "deviation" are being used in a very broad and general sense. Deviation includes all delays and inhibitions which give rise to expectation within the context of the particular style in question.

40. Morris R. Cohen, *A Preface to Logic* (New York: Henry Holt & Co., 1944), p. 47.

41. See George H. Mead, *Mind, Self, and Society* (Chicago: University of Chicago Press, 1934), p. 76.

42. Thus Pratt, while maintaining that the ideas aroused in association with music "have little to do with the intrinsic nature of musical sound" (C. C. Pratt, "Music and Meaning," *Proceedings of the Music Teacher's National Association*, Series XXXVII [1942], p. 113), does contend that "music *sounds* the way emotions *feel*" (*Ibid.*, p. 117), a statement which seems to be a disguised form of referentialism.

Incidentally though Pratt's first statement is undoubtedly true, its implications are not. For while our associations may have nothing to do with the intrinsic nature of sound, whatever that may be, they do have something to do with our experience of sound.

43. The term "stimulus" as used here includes any tone or combination of tones which are marked off as a unitary event which is related to other musical events. It is, to use Mead's terminology, "a musical gesture." Or it is a "sound term." In this sense, a single tone, a phrase, or a whole composition may be considered to be a gesture, a stimulus, or a sound term. In other words, meaning must be considered as being architectonic as well as consecutive.

44. Of course, it may have designative meaning. The more difficult it is to grasp the embodied meaning of a work, the greater the tendency to search for designative meanings.

45. Notice, too, that the final tones of a piece, conditioned by all that has gone before, lead us to expect silence and that it is this expectation which makes them meaningful.

46. Note that hypothetical meanings as well as the other kinds of meaning are architectonic. If we are set to listen to a Haydn rondo, then the idea we have of Haydn rondos is, in a sense, the hypothetical meaning of that particular rondo; it is what we envisage and what points to the impinging stimulus.

47. Bertrand Russell, *Selected Papers* ("Modern Library," [New York: Random House, n.d.]), p. 358.

48. Mead, *op. cit.*, p. 194.

49. Mead, *op. cit.*, p. 42–75.

50. *Ibid.*, p. 47.

51. Of course, if the composer is developing a relatively new style, as many contemporary composers have tried to do, the imagined listener may correspond to no listener who actually exists. He is rather one whom the composer hopes to create as his style becomes part of the general style, part of the listening public's stock of habit responses.

52. Leopold Mozart, *Versuch einer gründlichen Violinschule*, quoted in *Source Readings in Music History*, ed. Oliver Strunk (New York: W. W. Norton & Co., Inc., 1950), p. 602.

53. Generally speaking this study is not concerned with the creative act but rather

with the experience which the art work brings into being. This aspect of the composer's creative life is discussed because it clearly concerns the problem of aesthetic experience. Obviously many other mental processes and attitudes which have not been touched upon are involved in the act of composition.

54. Cassirer, *op. cit.*, p. 191.

# 2

# Music and Language: Parallels and Contrasts

RITA AIELLO

## Introduction

The relationship between music and language has been addressed by musicians, philosophers, and scholars (Bernstein, 1976; Langer, 1956; see Winn, 1981, for a rich historical account on this topic). The musicologist Curt Sachs (1943) proposed that "music began with singing." Comparisons between music and language are prompted by the strong similarities between the two. Both have an inherent structure and evolve over a temporal continuum, both have a meaning for the listener, and are innate expressions of human capacities. In music and in language there is a phonetic, a syntactic, and a semantic level.

The research in psycholinguistics has influenced the research in the perception of music (Clarke, 1989; Lerdahl & Jackendoff, 1983; Sloboda, 1985). Psychologists have applied some of the experimental designs used to investigate the perception of language to research the perception of music (De Witt & Samuel, 1990; Locke & Kellar, 1973; Tan, Aiello, & Bever, 1985). Drawing on the psychological literature, this chapter discusses phonetic, syntactic, and semantic comparisons between language and music and urges that there be a greater emphasis on research in the emotional meaning of music.

REFERENCES

Bernstein, L. (1976). *The unanswered question: Six talks at Harvard.* Cambridge, MA: Harvard University Press.

Clarke, E. F. (1989). Issues in language and music. *Contemporary Music Review, 4,* 9–22.

De Witt, L., & Samuel, A. (1990). The role of knowledge-based expectations in music perception: Evidence from musical restoration. *Journal of Experimental Psychology: General, 119*(2), 123–44.

Langer, S. (1956). *Philosophy in a new key* (3rd ed.). Cambridge, MA: Harvard University Press.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.

Locke, S., & Kellar, L. (1973). Categorical perception in a non-linguistic mode. *Cortex, 9,* 355–69.

Sachs, C. (1943). *The rise of music in the ancient world.* New York: Norton.

Sloboda, J. A. (1985). *The musical mind.* Oxford: Oxford University Press.

Tan, N., Aiello, R., & Bever, T. G. (1985). Harmonic structure as a determinant of melodic organization. *Memory and Cognition, 9,* 533–39.

Winn, J. A. (1981). *Unsuspected eloquence: A history of the relation between poetry and music.* New Haven, CT: Yale University Press.

---

Music begins where words end. Johann Wolfgang von Goethe (attr.)

## Similarities Between Music and Language

Music and language are universal, innate expressions of human cognition and communication (Clarke, 1989b; McAdams, 1987; Sloboda, 1985). Their universality is evident in the fact that all cultures express themselves verbally and musically. Both are innate abilities that develop through personal interaction, as can be observed in the spontaneity with which children learn to act and interact musically and verbally from a very young age. But, specifically, how does the child acquire a language? In his book *Reflections on Language,* Noam Chomsky (1975) writes:

> [The child] quite effortlessly makes use of an intricate structure of specific rules and guiding principles to convey his thoughts and feelings to others, arousing in them novel ideas and subtle perception and judgments. . . . [I]t remains a distant goal to reconstruct and comprehend what the child has done intuitively and with minimal effort. Thus language is a mirror of the mind in a deep and significant sense. It is a product of human intelligence, created anew in each individual by operations that lie far beyond the reach of will or consciousness. (p. 4)

To learn a language, then, the child subconsciously abstracts the rules of the language structure and utilizes them to create his or her own novel sentences.

Children do not just repeat the sentences they hear, but create new sentences by applying the rules they have abstracted from their language. A similar proposal can be made for the way in which children learn and create music. Along these lines, music, too, can be viewed as "a mirror of the mind in a deep and significant sense." In order to learn and to create music, children abstract some of the rules of their musical culture and use them creatively. Just as they learn spontaneously the rules that allow them to originate an unlimited number of sentences, so they seem to learn the musical rules that allow them to create novel sequences in music. The musical growth of children occurs in stages that have been compared to stages of their cognitive and symbolic development (Gardner, 1973; Hargreaves, 1986; Serafine, 1980; see also Chapters 4 and 5 in this volume).

Although music and language vary across cultures, they share several significant cognitive characteristics (Clarke, 1989b; McAdams, 1987; Sloboda, 1985). Both evolve over time and involve a meaningful use of sound patterns. Throughout the world people use sound patterns to create words, motives, scales, ragas, and drones. The psychologist George Miller (1956) brilliantly showed that our capacity for segmenting and remembering verbal units is approximately seven items: "The magical number seven, plus or minus two;" as he called it. And to a great extent this limit also exists in music if we consider that the pentatonic scale (five pitches) and the diatonic scale (seven pitches) serve as the bases for most of the world's music. The finite length of words, of musical motives, and of verbal and musical phrases reflects our need to segment along the temporal continuum so that we may process most effectively the sounds we hear and create.

Music and language are both modes of communication, yet they have different goals. Generally speaking, while the primary aim of language is to communicate thought, one of the main goals of music is to heighten emotions and express them aesthetically. Music is born out of the need to express ourselves and to communicate aesthetically through the abstractness and the characteristics of sound.

Meter and rhythm are at the heart of music and poetry. Indeed, accents and stresses are the means used to define linguistic and musical patterns. In processing language and music as sensory data, we make use of many of the characteristics of sound—pitch, tone, and dynamics. Some Oriental languages offer striking evidence of the importance that pitch and tone can have, and we find that syllables may be pronounced with different pitch contours. For example, in the Peking dialect of the Chinese language, there are four tones: flat, rising, falling-rising, and falling. The same syllable has different meaning depending on the tone with which it is pronounced: *mā* (pronounced at the level tone) can mean "mother", *má* (rising tone) can mean "hemp", *mǎ* (falling-rising tone) can mean "horse," and *mà* (falling tone) can mean "scold" (Miller, 1981: 59). The dynamic level with which one speaks may be particularly revealing of one's emotions, and in some Western languages (e.g., Italian) the intonation alone of a sentence, not its grammatical structure, conveys whether it is a declarative sentence or a question.

The music of several African cultures shows a close bond between the pitches played by some instruments and the language itself. In west and central Africa, talking drums are used both as musical instruments and as speech surrogates. These instruments reproduce the intonation of words and the rhythms of spoken texts (Nketia, 1974). In the African languages of Kele and Mbane, each syllable possesses a fixed tone, which is high or low. The talking drum sends messages imitating the patterns of high and low that occur in the words (Carrington, 1969: 18). And the *masengo* and *endingidi* fiddles of Ethiopia and Uganda appear even to imitate vocal timbre (Wachsmann & Cooke, 1980).

Words and music are close neighbors physiologically and perceptually (Burrows, 1990). The link is so ingrained that some theories of the origin of music propose that music is a form of heightened speech (Bernstein, 1976; Revesz, 1953; Stacey, 1989). In 1974, Leonard Bernstein devoted the Norton lectures which he delivered at Harvard University to a comparison of the origins of music and language and to an exploration of the similarities in the ways in which we create and process them. These thought-provoking lectures have been recorded in the book *The Unanswered Question* (Bernstein, 1976), which borrows its title from a composition by Charles Ives. Seeking to answer the question "Whither music?" these lectures have influenced research on the similarities of music and language by both musicians and psychologists (Clarke, 1989b; Lerdahl & Jackendoff, 1983; Sloboda, 1985).

Bernstein proposed that, since sound is based on the harmonic series, it functions as a universal for all types of music. He compared the harmonic series to the monogenesis of language: just as languages can be traced to a common origin, so tones can be traced to the harmonic series. Moreover, Bernstein emphasized that the order in which the partials occur in the harmonic series is responsible for the fact that certain intervals, such of the octave, the fifth, and the minor third, are so prevalent the world over, and he compared this recurrence with that of certain syllables (e.g., *ma*) in many of the world's languages.

Many musicologists have mused over the origins of music. Curt Sachs proposed that music sprang from singing (1943: 21). Indeed, the presence and importance of singing in most of the world's cultures attest to the universal binding of music and speech. The word *melody* derives from the musical intonation of a cadenced phrase. As Stravinsky noted, "*melodia* in Greek, is the intonation of *melos*, which signifies a fragment, a part of a phrase" (1947: 42). The singing styles of *recitativo* and *sprechstimme*, for example, are particularly vivid examples of how speech and Western music can be fused.

Classical Greece gave special attention to the union of music and language. The Greeks used singing and instrumental accompaniments to heighten the significance of poetry and drama, a tradition that is at the core of opera. Not until the late Middle Ages did music emerge as an autonomous art form transcending the scope of language. In the splendid book *Unsuspected Eloquence: A History of the Relations Between Poetry and Music*, James Anderson Winn (1981) describes the birth of medieval polyphony:

> The fact that music could achieve actual simultaneity, that it could have vertical as well as horizontal events, was a revolutionary discovery. . . . In literature, one could claim that a line of poetry had four layers of meaning, but in music one could actually write four simultaneous parts . . . . Until this point in history, both the recitation of a poem and the singing or playing of a piece of music had reached the ear as a single series of sounds in time; now music had a new kind of interest, the accidental or contrived vertical combination of two or more pitches. (p. 89)

Until the late sixteenth century, a major role of instrumental accompaniments was to enhance the poignancy of the sung word and of dance. Given today's vast repertoire of instrumental music, it may be easy to forget when instrumental music began to acquire its own independence from the vocal line and to develop into an autonomous art form. Before then, instruments mostly accompanied or imitated dance and song. Igor Stravinsky (1947) in *Poetics of Music in the Form of Six Lessons* remarked: "Music of the kind that has meaning for us today is the youngest of all the arts, although its origin may be as old as man's" (p. 28).

## Musical Phonology: Some Comparisons with Language Phonology

The linguist seeks to describe the structure of language just as the music theorist seeks to describe the structure of music. The psycholinguist and the psychomusicologist investigate in their respective domains the psychological processes by which language and music are created and understood. Psycholinguists study how language is created and processed according to three fundamental components: phonetics (the speech sounds of a language), syntax (the way in which words are put together to form phrases, clauses, or sentences), and semantics (the meanings of words). To understand a language, the listener must sufficiently process all three components of the language. Given the fundamental cognitive characteristics shared by music and language, a number of researchers investigating the perception of music have utilized some of the techniques that were developed to study the perception of language. In this section I will discuss two perceptual phenomena that occur in the perception of language and music at the phonetic level: categorical perception and phonemic restoration.

The phonetics of a language are the sound units that constitute the language itself. A phoneme is the smallest phonetic unit. It distinguishes one utterance from another and is roughly the same as a written letter of the alphabet (e.g., b, c, f). The words *bat* and *cat* and *fat* begin with three different phonemes, *b*, *c*, and *f*, but share others. The analysis of speech sounds into phonemes, as well as the division of phonemes into distinctive features, is critical for any analysis of speech perception.

How do we perceive a language at the phonetic level? Partly by categorizing

the sounds that we hear. Categorical perception is the way in which we categorize what we hear. Jusczyk (1986) explains: "Generally speaking, categorical perception is said to occur when the ability to discriminate among members of the same category is very poor, whereas the ability to discriminate among members of different categories is very good. As applied to speech sounds, categorical perception implies that discrimination of sounds from two different categories such as /b/ and /p/ should be relatively easy, but that the discrimination of two different tokens of /p/ from the same speaker should be almost nonexistent" (27-11).

This ability to categorize is strikingly similar to a musical phenomenon. Musicians categorize musical sounds more accurately than nonmusicians just as native speakers categorize speech sounds that are phonemic in their language more accurately than nonspeakers do (for a review of these latter data see Repp, 1984).

A musical note can roughly be compared to a phoneme (Bernstein, 1976; Sloboda, 1985). Locke and Kellar (1973) investigated the categorical perception of musicians and nonmusicians in detecting a difference between an A minor triad and an A major triad. They asked subjects to compare pairs of sine wave triads. The outer notes of the stimuli were always 440 Hz (A) and 659 Hz (E) while the middle note varied by 2, 4, 6, or 8 Hz, ranging from 523 Hz (C) to 554 Hz (C-sharp). It was found that the musicians discriminated better across category boundaries than the nonmusicians. They were more sensitive to the differences between the chords and were able to discriminate between the triads when the middle pitch varied by 4 Hz. In general, nonmusicians require a greater difference between the triads to perceive a change. The superiority of musicians in categorizing musical sounds has been proven using different types of musical stimuli (Wapnick, Bourassa, & Sampson, 1982; Zatorre & Halpern, 1979).

The perceptual phenomenon of phonemic restoration also occurs in both language and music; it describes the ability of listeners to use their previous knowledge to supply what is missing (Warren, 1970). De Witt and Samuel (1990) explain that if a phoneme occurring within a word is replaced with noise, listeners cannot indicate which phoneme has been distorted because their knowledge of the word supplies the missing information; therefore, they perceive having heard the intact word with nothing missing. Similarly, in the musical case, restoration occurs when listeners believe they have heard an auditory signal intact when in fact noise has replaced some of the notes.

De Witt and Samuel (1990: experiment 1) asked subjects to hear a series of familiar melodies and indicate whether a noise replaced or coincided with a musical pitch. The researchers found that the subjects' specific expectations about a particular musical sequence allowed them to notice disruptions to that sequence. The musical training of the listeners affected their perception of the stimuli as well: overall, musicians showed better discrimination performance than the less trained subjects. What one expects to hear influences the nature of what one hears. As De Witt and Samuel note, "musical restoration is an example

of a more general perceptual phenomenon in which prior knowledge and expectations lead to a filling in of a missing information" (p. 142). The perception of music is not simply a passive reception of tones but an active process. The occurrence of categorical perception and of restoration effects in speech and music demonstrates a certain level of commonality of processing across these two domains.

## Musical Syntax: Some Comparisons with Language Syntax

Consciously or subconsciously, one listens to music according to some stylistic and perceptual laws. Music is composed and improvised according to a style, and even the musically untutored listeners acquire a musical syntax, a system of rules by which they make sense of music. Despite varying degrees of awareness, listeners understand the syntactic complexity of composition and the arguments of musical form (Spender, 1987: 502).

In music as in language, we are aware of the grammatical features of the stimulus. Experiments show that there are strong similarities in the way in which people perceive structure in music and in language. Deutsch (1981, 1982) asked musically trained subjects to recall sequences of 12 pitches by writing them down in musical notation. She found that the subjects' performance depended on the degree of structure in the stimuli. Sequences that were tonally structured were easier to write down than sequences that were not tonally structured. In addition, the placement of rests within the sequences also had an effect. The subjects' performance decreased when the rests were placed in locations that did not coincide with the tonal structure of the sequences.

A number of experiments have applied techniques that were developed to investigate the syntactical aspects of language perception to study the syntactical aspects of music perception. I will describe three categories of experiments that show that listeners are aware of the grammatical features of the stimuli, and whose results are rather similar in the language and the music domain: reading experiments, phrase boundary experiments, and experiments that consider the location of clicks within sequences.

### *Reading Experiments*

A number of studies (Levin & Kaplan, 1970; Sloboda, 1974, 1977) have investigated the span (i.e., the number of words or notes) that readers take in ahead of time when reading language or sight-reading music. In language reading, experienced readers tend to take in material extending up to a phrase boundary, wait until the complete processing of that phrase, and then rapidly sweep to the next phrase boundary to take in the material of the next phrase. We tend to pause the longest after reading complete major grammatical units (Sloboda, 1985).

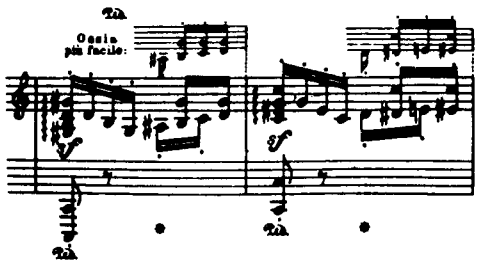


FIGURE 2.1. Brahms, Cappriccio, op. 76, no. 2, bars 76–78.

In sight-reading music, experienced readers identify significant structural units and scan them differently, according to whether the music is homophonic or contrapuntal. Experienced sight readers do not read note-by-note but shape the music as they read it by emphasizing patterns, thus creating a meaningful interpretation. They create meaningful patterns by imposing performance markers onto the music they play. Although an experienced language reader reads a text roughly in the same way as an experienced musician sight-reads music, the proficiency of the musician is evident in one additional way: in the skill with which he or she eliminates some material from the score so that the sight-reading will be heard as being both technically proficient and musically coherent. It is very interesting that in sight-reading a score musicians tend to commit "proof-readers'" errors. Pianists fail to notice the errors that may be in a score and subconsciously correct them, possibly on the basis of their tacit knowledge of the stylistic norms of the music (Sloboda, 1985: 75)

In the second measure of Figure 2.1, most pianists will play the G as if it were written G-sharp. The fact that it is actually a G-natural will easily go unnoticed. It is the *inexperienced* pianist who, lacking the musical intuition, will tend to sight-read this passage as it is printed.

## *Phrase Boundary Experiments*

We segment language and music in locations that reflect the grammatical structure of the stimuli. A phrase, whether it occurs in language or in music, is a psychological unit. Experiments indicate that, in both music and language processing, we tend to process more effectively material belonging to one phrase than material that integrates two phrases. Tan, Aiello, and Bever (1985) asked musicians and nonmusicians to listen to single-line melodies that used only equal-duration notes and that formed a two-phrase musical period. Using a psycholinguistic paradigm, the subjects were asked to recognize a two-note probe that either (1) occurred at the end of the first phrase, (2) straddled the phrase boundary, or (3) occurred at the beginning of the second phrase. All subjects, regardless of their musical training, perceived with greater difficulty the two-note probe that straddled the phrase boundary. The familiarity with the harmonic structure influenced the listener's perceptual organization of the melodies in

ways analogous to the perceptual organization of clauses or sentences. This result, however, was more accentuated for the musicians.

### *Click Experiments*

Additional evidence of the syntactical similarities between language and music comes from studies that investigate the location of clicks (Fodor & Bever, 1965; Gregory, 1978). These experiments look at the perceived location of clicks and reveal the strategies that subjects use in processing the auditory material. Click experiments have shown that subjects use basically the same strategies in processing verbal and musical sequences because the grammatical structure of the stimuli guides the listeners' perception. In click experiments, a click is presented at one or more strategic locations while the subjects hear some verbal or musical material; the subjects' task is to indicate the syllables or notes that occurred simultaneously with the click (or clicks). Frequently listeners report the location of the clicks inaccurately. They believe they heard them at the end of phrases, regardless of the location where the clicks actually occurred. It is as if the clicks *migrated* toward the direction of the phrase boundary. This result illustrates that phrase boundaries are important in the perception of language and music.

## Musical Grammars

The theories of language proposed by the linguist Noam Chomsky (1957) have been utilized by psychologists and musicians in making comparisons between the nature of music and language. Chomsky postulated that every possible sentence consists of two levels: a surface structure and a deep structure. The surface structure is closely related to the string of words that we hear or see, while the deep structure is closely related to the underlying meaning of the string of words. Let us consider an example. Although at the surface the sentences *John is easy to please* and *John is eager to please* differ in only one word, we understand that there is a substantial difference in their meaning. Indeed, they have different underlying deep structures. On the other hand, we understand that "the dog chased the cat" and "the cat was chased by the dog" are related in meaning (that is, have the same underlying deep structure) despite the dissimilarities in their surface structures.

Chomsky proposed that by applying various transformational rules to the deep structure—for example, *John* + *Love* + *Mary*—we can derive all the sentences that are related to this deep structure, such as *John loves Mary, John does not love Mary, Does John love Mary?* and so on. As I pointed out at the beginning of this chapter, we do not store the sentences that we would like to say but, rather, the rules that give rise to the sentences. If language were based exclusively on repeating what we heard, we could *only* repeat what we had heard before and say nothing new. Instead, we abstract the grammatical rules from the

previously heard sentences and use them to generate new sentences, thus expressing our own novel thoughts. And, in a general sense, the same can be said for music.

Leonard Bernstein (1976) was among the musicians seeking analogies between aspects of Chomskian theories and the structure of music (other interesting proposals have also been made by Lerdahl & Jackendoff, 1983; Sloboda, 1985; and Clarke, 1989b). Bernstein illustrated how a Chomskian transformational grammar could be compared, in some fashion, to the theories of music proposed by Heinrich Schenker (1935/1979). For Bernstein, Schenker's analyses trace what happens in a piece of music to a basic chordal structure, just as Chomskian transformational theories trace what we hear as the surface of language to the deep structure of language. In this vein, the melody (Schenker's *urline*) is equivalent to the surface structure of the language, and the harmonic structure (Schenker's *bassbrechung*) can be compared to the deep language structure.

One must keep in mind that while Schenker's musical analyses are based on long structures of music that can reflect as much as an entire piece (Schenker, 1969), Chomsky's theories of language are based on strings of words that are relatively short. Nevertheless, there is validity in the Chomsky-Schenker comparison because there is a relationship between the melody and the harmony that can be compared to the deep and the surface structure of language. But although this language-music analogy is attractive and stimulating, I believe that it fails to do justice to a major distinctive element between language grammar and music grammar: ambiguity. Ambiguity occurs rather seldom in language because language is based on communication (I will discuss ambiguity again later in this chapter with reference to the meaning of music). On the other hand, ambiguity is a major element of the grammar and aesthetics of music. Chomskian transformational theories of language explain that what is ambiguous in language is the result of two possible, but different, deep structures. For example, the sentence *they are visiting firemen* could be interpreted as *(they) ([are] [visiting firemen])* or *(they) [are visiting] [firemen])*.

In music, however, a specific harmonic progression can give rise to many melodies *and* the same melody can be harmonized by several different harmonies. Figures 2.2 and 2.3 present the Cantabile movement from Beethoven's Quartet, op. 18, no. 5, and two settings of the chorale "O Ewigkeit, du Donnerwort" by J. S. Bach which illustrate this last point.

Moreover, when listening to music there are many possible perceptions. Many elements occur simultaneously, and each musical element may have a specific syntax of its own. When listening to music we are subconsciously aware that there are *many* grammatical possibilities between the parts. The grammar of music allows for complexity and simultaneity. Overall, the syntax of music has much more latitude than that of language. Thus, in comparing the syntaxes of music and of language, we must remember that music is by far more flexible and ambiguous than language, almost regardless of style.

Let us look at an example where the syntactical ambiguity occurs because of

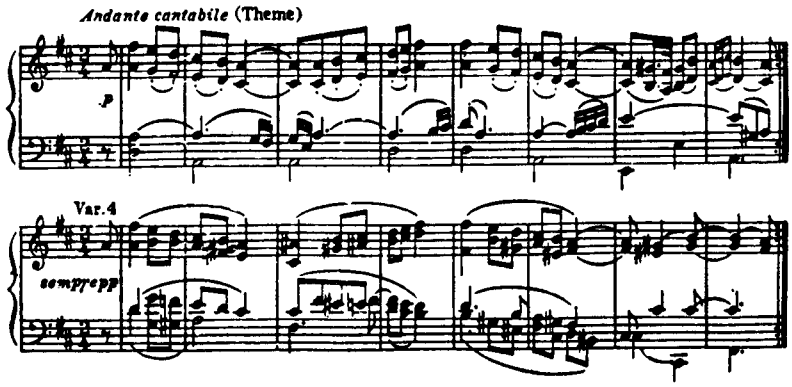



FIGURE 2.2. Beethoven, Andante Cantabile from Quartet, op. 18, no. 5, theme and variation 4.

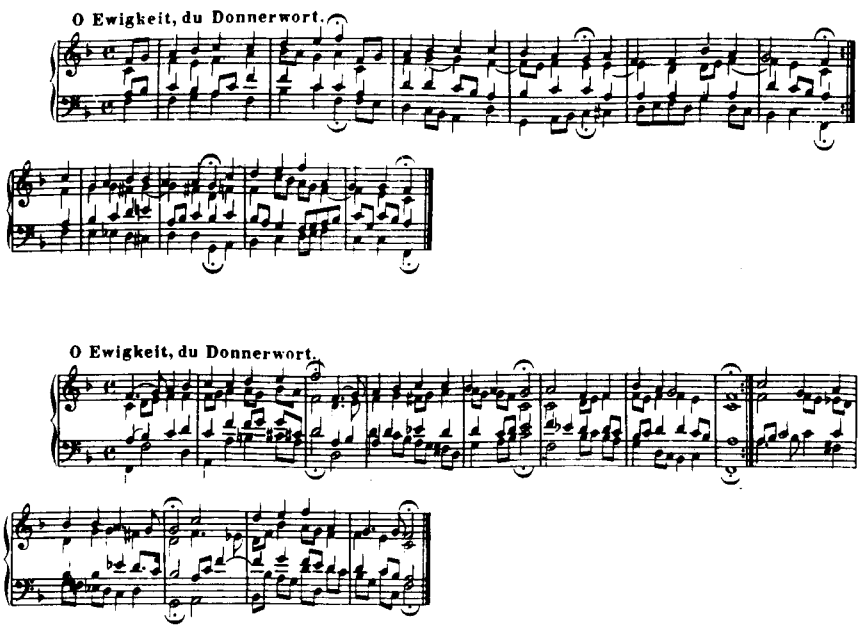



FIGURE 2.3. J. S. Bach, two settings of the chorale "O Ewigkeit, du Donnerwort."

the tonality of the piece: Chopin's Mazurka, op. 17, no. 4. What key is this mazurka in? We do not know until later in the piece when we finally realize that it is in the key of A minor. Clearly, in this case, it is the harmony that provides the element of ambiguity to the piece[1] (Figure 2.4).

In returning to Bernstein's analogy that the deep structure and the surface structure of language can be compared to the harmonic structure and the melody of music, I would like to propose that the deep structure of language be compared to a musical theme, and that the surface structure be compared to the musical variations that can be made with the theme. In this vein, by applying rhythmic, melodic, and harmonic transformations to the theme (the deep structure) the composer creates musical variations that reflect their original source but that, at the same time, have the freedom and the ambiguity which are characteristic of music.

For the psycholinguist, a grammar is a set of rules capable of generating an aspect of a structure of language. Grammars are collections of strategies for understanding and producing sentences. The developments in the psychology of music have influenced many researchers to write about musical grammars (Baroni, 1983; Roads, 1979). In recent years, grammars have been proposed for generating Swedish nursery songs (Lindblom & Sundberg, 1970), J. S. Bach's chorales (Baroni & Jacoboni, 1978), chord progressions in jazz (Steedman, 1984), and the setting of words to Gregorian chant (Chen, 1983), to mention only a few. And in *A Generative Theory of Tonal Music*, Lerdahl and Jackendoff (1983) combined principles of Chomskian transformational grammar and Schenkerian theory and proposed a theory of how listeners familiar with the tonal style listen to tonal music. This theory takes into account principles of: (1) grouping, (2) meter, (3) time-span reduction (a system that assigns the relative dominance of each surface event according to a tree structure), and (4) prolongation reduction (a system that describes the patterns of harmonic tension and relaxation). To date, however, only parts of this theory have found experimental support (Bigard, 1990; Deliège, 1987).

Although the proposal of musical grammars has been very instrumental in advancing research in the psychology of music as a whole, a psychology of music must be sensitive to features that are *unique* to music, because the application of psycholinguistic paradigms alone cannot answer the questions that are at the core of music itself (see also Clarke, 1989a, 1989b). A grammar prescribes an order, a set of rules; it does not guarantee an aesthetic reaction, it is not a prescription for it. Productive research in the psychology of music will need to address not only how the mind perceives the sounds, but the aesthetic and emotional meaning that the sounds give rise to.

Musically speaking, the style of a piece reflects its musical grammar. If the grammar of a musical composition is too taxing perceptually, will listeners relate well to the piece? David Smith and Jordan Witt (1989, see also Smith, 1987) tested how listeners react aesthetically to tonal and serial compositions. They asked subjects classified as "nonexperts who enjoyed classical music" to listen to

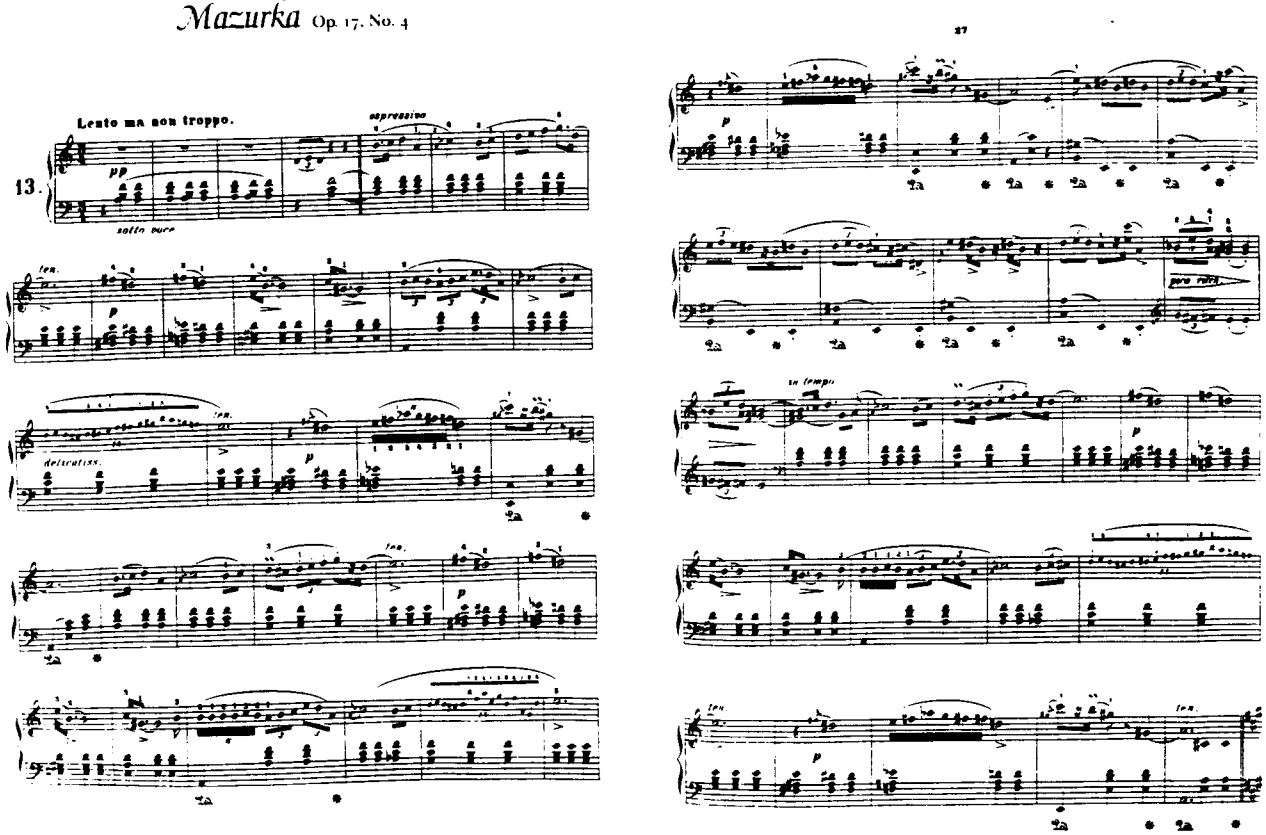



FIGURE 2.4. Chopin, Mazurka, op. 17, no. 4.

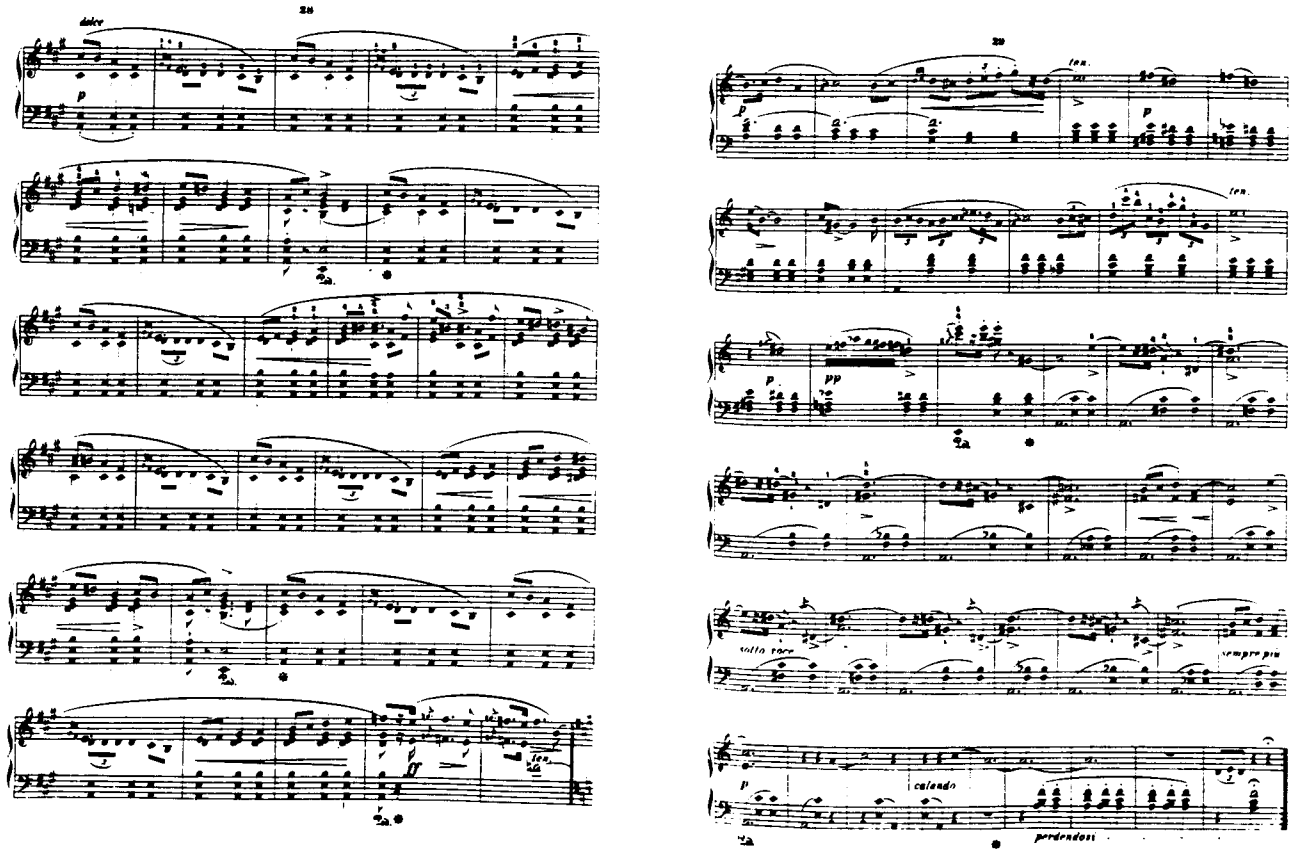



FIGURE 2.4 (*continued*)

brief excerpts from the beginning and the end of tonal and serial works by Schoenberg and Webern and to describe them using a set of 26 adjectives and a rating scale from 1 to 7. Thus, it became possible to draw an aesthetic profile of each musical excerpt. The results showed that listeners find serial music, with its atonality and rather unfamiliar grammar, to be less rich emotionally than tonal pieces written by the same composers. Smith and Witt, therefore, found that if the stylistic constraints of a piece are perceptually too demanding, listeners may not be able to understand the music and thus may not derive as much enjoyment from what they hear.

This is indeed the general reaction that many listeners will have, at least in the beginning. They will tend to perceive that the stylistic complexity of a serial piece is the result only of the composer's intellectualism. When that happens, the music loses its aesthetic and emotional message for many listeners. In this context, I am reminded of the remarks made by the writer Isaac Bashevis Singer on literature: "The very essence of literature is the war between emotion and intellect. When literature becomes too intellectual—when it begins to ignore the passions, the emotions—it becomes sterile, silly, and actually without any substance" (Page, *The New York Times*, July 26, 1991).

## Listening to Music: Where Is the Meaning?

What is musical meaning? Specifically, how does the listener derive a meaning from music? Earlier I discussed some comparisons between language and music at the phonetic and the syntactical level. But, semantically, where is the musical meaning to be found? Is the semantic level of music to be found in the emotions that music elicits? To a great extent, *Yes*. But ultimately, *why* do we listen to music? To follow a grammar? To abide by a perceptual theory? Probably not. Most of us listen to music to partake of the emotional and aesthetic experience that music offers.

Unfortunately, there is a wide gap between the progress that has been made in understanding how music is perceived and what it means to listeners. A close link clearly exists between music cognition and the emotions that music elicits, but an evaluation of this link is yet to be researched thoroughly. Meaning, in all its complexity, most likely lies hidden in the emotional content of music, and all of us, consciously or not, perceive it.

Language is based on understanding the interaction of its phonetic, syntactic, and semantic levels. And while music conveys meanings, its semantic level is not as uniformly defined as that of language. Language is the means through which verbal information is shared, and at the level of literature and poetry, it becomes the art form through which we express ourselves verbally. Music is the art form that expresses feelings and meanings through the qualities of sounds and the relationships between sounds. In language, the specificity of the semantic component provides information. Indeed, verbal communication cannot take place

if the listener does not understand the semantic meaning. But in music this is not necessarily the case. Although the meaning of music is enhanced by our knowledge of musical styles and practices, musical meaning remains both pluralistic and personal.

To understand a language, the listener combines the sounds into words that have a semantic meaning and places them within an organized structure (i.e., a grammar). To understand music we combine sounds into motives, melodies, and phrases and listen with various degrees of awareness to many musical elements occurring simultaneously: rhythm, tonality, dynamics, interpretation, all within an overall structure that reflects stylistic influences. There is not one single, uniform meaning that can be attached to a musical composition by all listeners at all times with the same certainty as there is a semantic meaning in a language composition. Broadly speaking, the lack of a specific semantic component leaves the listeners quite free to select from the meanings created by the combinations of the musical elements, and to attach their own meanings to the musical composition as a whole. Because of the lack of a uniformly defined semantic meaning and the multiplicity of elements occurring in a musical piece, music has a plasticity and a richness that language does not.

Music contains a simultaneity of meanings and a plurality of meanings. The simultaneity of meanings exists because each musical element has its own line, its own importance, yet all contribute to an overall form and significance. In music many musical elements occur, simultaneously creating musical shapes and giving rise to musical meanings. The combination of the musical elements and the effect of the performer's interpretation create an exceptionally rich and complex stimulus.

Although our tone of voice and the dynamics of our speech influence very much the nuances of meaning that we wish to convey, communication through language is possible because of the specificity of the semantic level. For example, pronounce "good morning" placing a different emphasis each time on one of the two words. Now pronounce it using a variety of speech dynamics from a bitter, sarcastic comment to an affectionate greeting. Because music is not based on a unified semantic level, we become particularly sensitive to the specific combinations of the musical elements, the nuances, and the dynamics of each performance. Lewers (1980) writes: "If inflection and nuance *enhance* the effect of spoken language, in music they *create* the meaning of the notes. Unlike words, notes and rests do not point to ideas beyond themselves; their meaning lies precisely in the quality of the sounds and silences, so that the exact rendering of the notes, the nuances, the inflection, the intensity and energy with which notes are performed *become* their musical meaning" (p. 67). Indeed, a composition may take a particular meaning just because of the interpretation of a performer.

How many musical elements contribute to create a single measure of music? One of the principal reasons we can keep listening over and over again to a brief piece of music is because of the multiplicity of patterns that potentially we can attend to, and because of the potentially great number of meanings that listening

to music can offer us. The art of listening is the art of discovery, and it motivates us to listen again. Consciously or subconsciously, listeners realize that they could discover different degrees of saliency for specific elements in the music with every hearing. The simultaneity with which lines and elements occur in music is different from that of language; this influences the way in which we listen to music perceptually and influences the way in which we retrieve a meaning from music emotionally. When several people speak different words at the same time, they produce cacophony. But when they sing different words at the same time, it may result in immense beauty if the individual parts are coordinated harmonically, melodically, and rhythmically.

The knowledge of a musical style or tradition can be of great importance in relating the intent of the music. In listening to music, the meaning can be derived from (1) the intellectual appreciation of the musical elements, (2) the emotional, aesthetic reaction that results in the appreciation of the stylistic characteristics of the music (Meyer, 1956, see also Chapter 1 in this volume), and (3) the association of a piece with a specific event or place (what John Booth Davis called the "*Darling, they are playing our tune*" theory, 1978: 69). In the book *The Open Work*, Umberto Eco (1989) writes: "the form of the work of art gains its aesthetic validity precisely in proportion to the number of different perspectives from which it can be viewed and understood. These give it a wealth of different resonances and echoes without impairing its original essence" (p. 3).

In some way, then, listening to music is like looking into a prism: to appreciate its light, we view it from different angles since we may only observe the prism's light one way at a time. This does not deprive the prism of its multiplicity of splendors, but attests to our limitation in being able to focus on one side at a time. In this sense, we are always "correct" when we listen to music because there is no exclusive meaning that can be attributed to a piece. How many elements could we potentially focus on, and how many meanings could a musical composition potentially have?

Researchers have probed extensively into the relationship between music and semiotics, the philosophical theory that deals with the interpretation of signs and symbols. A number of writers have proposed that music, too, may be interpreted semiotically (Clarke, 1989b; Eco, 1979, 1989; Nattiez, 1990; Stefani, 1974). From this perspective, musical excerpts or entire musical compositions may be interpreted by listeners as signs having an explicit meaning because of their structure or their historical association. Umberto Eco (1979) explains:

> [M]usic presents, on the one hand, the problem of a semiotic system without a semantic level . . . on the other hand, however, there are musical signs . . . with an explicit denotative value (trumpet signals in the army) and there are . . . entire "texts" possessing pre-culturalized connotative value ("pastoral" or "thrilling music," etc.). In some historical eras music was conceived as conveying precise emotional and conceptual meanings, established by codes, or at least, "repertoires." (p. 11)

## Information Theory

One reason people listen to music is that the listening confirms one's expectations. From this perspective, the listeners' awareness of the patterns of musical tensions and resolutions is what gives music its significance. The listeners' knowledge of the musical style allows them to partake of the tensions and resolutions of the music. As Meyer (1956, see also Chapter 1 in this volume) has proposed: "in music . . . the same stimulus, the music, activates tendencies, inhibits them and provides meaningful and relevant resolutions" (p. 23). In an insightful commentary on the first movement of Beethoven's Fifth Symphony, titled "Music, Mind and Meaning," Marvin Minsky (1981) asks: "Why do we like music? . . . everyone knows how it touches our emotions, but few think of how music touches other kinds of thoughts" (p. 28). He goes on to describe how listening to music engages the previously acquired personal knowledge of the listener. Specifically, Minsky asks: "Why do we like certain tunes? Because they have certain structural features. Because they resemble other tunes we like" (p. 34).

The views of Meyer and Minsky (see also Moles, 1968; Pinkerton, 1956; Youngblood, 1958) are rooted in information theory, the theory based on the mathematical probabilities of what will occur next in a chain of events. In brief, research in the applications of information theory to music examines the probabilities that a musical message will be understood by the listener considering the arousal, the satisfaction, and the frustration of the expectations of the listener. In music there are units (i.e., chords, isolated notes, etc.) that represent meaningful events. How can the listener predict what will happen in hearing the chain of musical events? And which specific predictions can the listener make? Events in music form patterns, and these occur especially at meaningful moments in the unfolding of the music (Coons & Kraehenbuehl, 1985; Kraehenbuehl & Coons, 1959; for a summary on information theory and music, see Bent, 1987). The listener, then, searches for a musical meaning by constantly interpreting what is occurring in the music. This view is consistent with research in the psychology of emotions as proposed by Mandler (1984).

In listening to music, the listener seeks an equilibrium: there must be a proportion of novelty within a background of predictability. So the meaning of music is based on the enigma: on the tension of hearing simultaneously the old and the new, in rejoicing in what we recognize, yet in being fascinated and intrigued by the new possibilities that the composer presents us. The theme from "Frère Jacques," as it appears in the beginning of the third movement of Mahler's First Symphony, is an example of how a familiar melody may present us with a completely different meaning because of the context in which it is heard, and

because of the way in which the composer has altered it. Here the French children's song is in a minor key, at a slower tempo than the usual one, and accompanied by the poignant ostinato of the timpani. Our emotions in recognizing the "Frere Jacques" theme are likely to reflect our ambiguity—on the one hand, recalling the way we heard it before and, on the other, wanting to embrace the way Mahler presents it now (Figure 2.5).

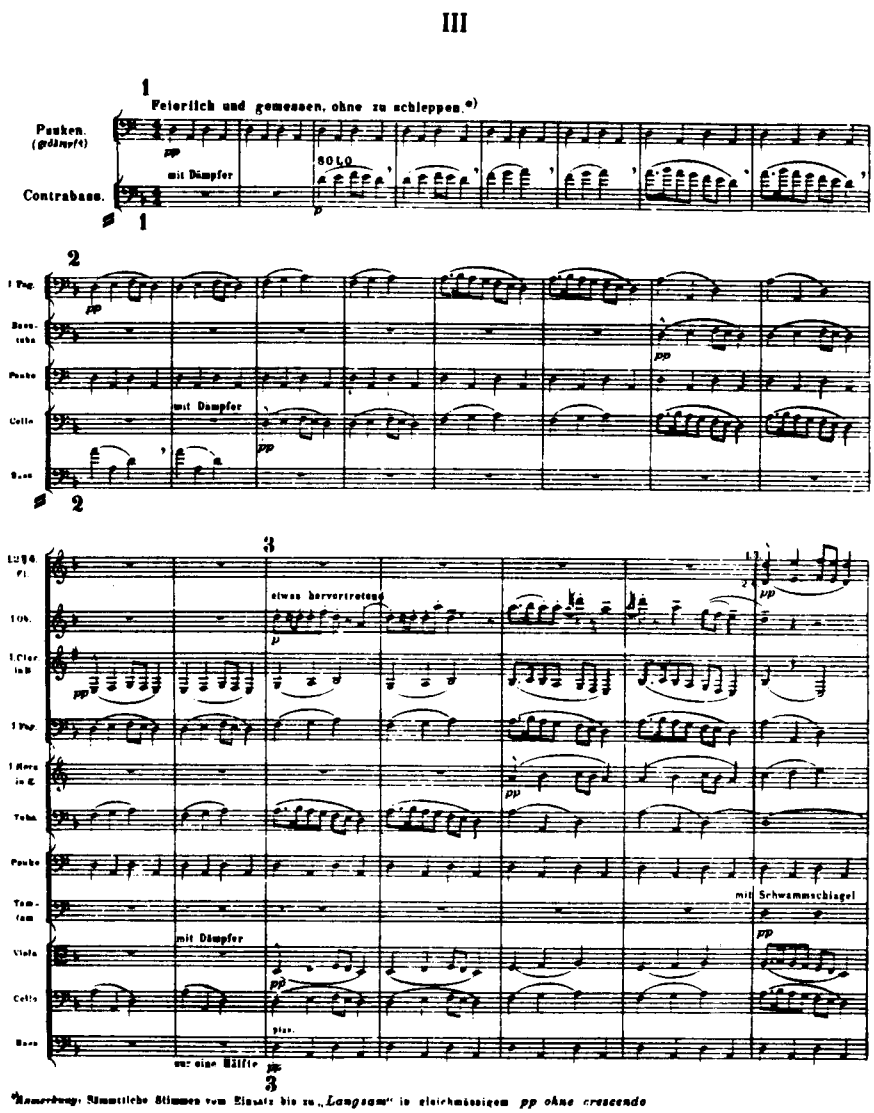



FIGURE 2.5. Mahler, Symphony No. 1, beginning of the third movement.

### *Experimental Research in the Meaning of Music*

The scarcity of research in musical meaning must be attributed in part to the difficulties of the methods that would be needed to investigate this topic, the major one being caused by the fact that there is no uniform meaning in music for all listeners at all times, and that the meaning that the listeners derive from a piece of music may change over time. But, overall, when subjects are asked to assign a title or a descriptive adjective to musical excerpts, their responses do fall within broad clusters of agreement (Hevner, 1936; Wedin, 1972).

In the book *The Language of Music*, Deryck Cooke (1959) proposed that the intervals of the diatonic scale suggest different emotional qualities and described how some basic melodic tonal combinations give rise to specific emotions. Cooke based his theory on the music of the classic and romantic periods, focusing only on the tonal structure of recurring, brief, melodic motives, without taking into account the importance of their orchestration, tempo, rhythm, dynamics, and the broader musical context in which these motivic fragments occurred.

Recently, Sloboda (1991a, 1991b) has begun to investigate whether emotional reaction may be related to specific musical events. He asked listeners to fill out a questionnaire naming the particular pieces of music to which they could recall having experienced physical manifestations commonly associated with emotions. Sloboda found that there was much agreement in the subjects' responses, and that their emotional reactions could be linked to particular structural features of the musical score. Listeners attributed feelings associated with tears to passages containing appoggiaturas and melodic or harmonic sequences (such as the opening of Albinoni's Adagio for Strings), and they associated shivers with musical passages containing new or unprepared harmonies, or sudden shifts in tonality.

Many listeners emphasized how music promoted an intensification or a release of their existing emotions (Sloboda, 1991a). The comments of these listeners link pertinently with the words of the philosopher Susanne Langer. For Langer, musical structure logically resembles the dynamic pattern of human experience. In her book *Philosophy in a New Key* (1956), she writes: "Music is not the cause or the cure of feelings, but their logical expression. It is not usually derived from affects nor intended for them; but . . . it is about them" (p. 218). Thus Langer finds that music "is not self-expression. It is the expression of feelings" (p. 221).

## Summary

Some analogies can be drawn between music and language, and a number of experiments have shown similarities between the perception of music and lan-

guage at the phonetic and syntactic levels. Unlike language, which is attached to semantic structure, music can truly be seen as a universal language in the sense that music has meaning for everyone who listens to it. This meaning is not static, but can change for each person with each listening as our focus shifts from melody to harmony to rhythm to form to the subtleties of the performer's interpretation.

## NOTE

1. Bernstein (1976) also used this Chopin mazurka as an example of ambiguity, although his discussion had a different emphasis.

## REFERENCES

Baroni, M. (1983). The concept of musical grammar. *Music Analysis*, 2(2), 175–207.

Baroni, M., & Jacoboni, C. (1978). *Proposal for a grammar of melody: The Bach chorales.* Montreal: University of Montreal Press.

Bent, I. (1987). *Analysis.* New York: Norton.

Bernstein, L. (1976). *The unanswered question: Six talks at Harvard.* Cambridge, MA: Harvard University Press.

Bigard, E. (1990). Abstraction of two forms of underlying structure in a tonal melody. *Psychology of Music, 18*(1), 45–59.

Burrows, D. (1990). *Sound, speech and music.* Amherst, MA: University of Massachusetts Press.

Carrington, J. (1969). *Talking drums of Africa.* New York: Negro University Press.

Chen, M. (1983). Toward a grammar of singing: The tune text association in Gregorian chant. *Music Perception, 1*, 84–122.

Chomsky, N. (1957). *Syntactic structures.* The Hague: Mouton.

Chomsky, N. (1975). *Reflections on language.* New York: Pantheon Books.

Clarke, E. F. (1989a). Mind the gap: Formal structures and psychological processes in music. *Contemporary Music Review, 3*, 1–13.

Clarke, E. F. (1989b). Issues in language and music. *Contemporary Music Review, 4*, 9–22.

Cooke, D. (1959). *The language of music.* Oxford: Oxford University Press.

Coons, E., & Kraehenbuehl, D. (1958). Information as a measure of structure in music. *Journal of Music Theory, 2*, 127–61.

Davis, J. B. (1978). *The psychology of music.* Stanford, CA: University of Stanford Press.

Deliège, I. (1987). Grouping conditions in listening to music: An approach to Lerdahl and Jackendoff's grouping preference rules. *Music Perception, 4*, 325–60.

Deutsch, D. (1981). The processing of structured and unstructured tonal sequences. *Perception and Psychophysics, 28*, 381–89.

Deutsch, D. (1982). The processing of pitch combinations. In D. Deutsch (Ed.), *The psychology of music.* New York: Academic Press.

De Witt, L., & Samuel, A. (1990). The role of knowledge-based expectations in music perception: Evidence from musical restoration. *Journal of Experimental Psychology: General, 119*(2), 123–44.

Eco, U. (1979). *A theory of semiotics* (1st Midland Book Ed.). Bloomington: Indiana University Press.

Eco, U. (1989). *The open work* (A. Cancogni, Trans.). Cambridge, MA: Harvard University Press.

Fodor, J. A., & Bever, T. G. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior, 4*, 414–21.

Gardner, H. (1973). *The arts and human development*. New York: Wiley.

Gregory, A. H. (1978). Perception of clicks in music. *Perception and Psychophysics, 24*, 171–74.

Hargreaves, D. (1986). *The developmental psychology of music*. Cambridge: Cambridge University Press.

Hevner, K. (1936). Experimental studies in the elements of expression in music. *American Journal of Psychology, 48*, 246–68.

Jusczyk, P. (1986). Speech perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 2. Cognitive processes and performance*. New York: Wiley.

Kraehenbuehl, D., & Coons, E. (1959). Information as a measure of experience in music. *Journal of Aesthetics and Art Criticism, 17*, 510–22.

Langer, S. (1956). *Philosophy in a new key* (3rd ed.). Cambridge, MA: Harvard University Press.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Levin, H., & Kaplan, E. (1970). Grammatical structure and reading. In H. Levin and J. Williams (Eds.), *Basic studies in reading*. New York: Basic Books.

Lewers, J. M. (1980). *Rehearsal as the search for expressiveness: Implications for music reading in the high school mixed chorus*. Unpublished doctoral dissertation, Teachers College, Columbia University, New York.

Lindblom, B., & Sundberg, J. (1970). Towards a generative theory of melody. *Swedish Journal of Music, 52*, 71–88.

Locke, S., & Kellar, L. (1973). Categorical perception in a non-linguistic mode. *Cortex, 9*, 355–69.

Mandler, G. (1984). *Mind and body: Psychology of emotion and stress*. New York: Norton.

McAdams, S. (1987). Music: A science of the mind? *Contemporary Music Review, 2*, 1–61.

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Miller, G. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

Miller, G. (1981). *Language and speech*. San Francisco: W. H. Freeman.

Minsky, M. (1981). Music, mind and meaning. *Computer Music Journal, 5*(3), 28–44.

Moles, A. (1968). *Information theory and esthetic perception* (J. E. Cohen, Trans.). Urbana: University of Illinois Press.

Nattiez, J. J. (1990). *Music and discourse: Toward a semiology of music* (C. Abbate, Trans.). Princeton, NJ: Princeton University Press.

Nketia, K. (1974). *The music of Africa*. New York: Norton.

Page, E. (1991, July 26). Isaac Bashevis Singer, Nobel Laureate for his Yiddish stories, is dead at 87. *The New York Times*, p. B 5.

Pinkerton, R. (1956). Information theory and melody. *Scientific American, 194*(2), 77–86.

Repp, B. (1984). Categorical perception: Issues, methods, findings. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 10). New York: Academic Press.

Revesz, G. (1953). *Introduction to the psychology of music* (G. I. C. De Courcy, Trans.). London: Longmans, Green.

Roads, C. (1979). Grammars as representations for music. *Computer Music Journal, 3*(1), 48–56.

Sachs, C. (1943). *The rise of music in the ancient world.* New York: Norton.

Schenker, H. (1979). *Free composition* (E. Oster, Trans.). New York: Longman. (Original work published 1935)

Schenker, H. (1969). *Five graphic music analyses.* New York: Dover Publications.

Serafine, M. L. (1980). *Music as cognition.* New York: Columbia University Press.

Sloboda, J. (1974). The eye-hand span: An approach to the study of sight-reading. *Psychology of Music, 2,* 4–10.

Sloboda, J. (1977). Phrase units as determinants of visual processing in music reading. *British Journal of Psychology, 68,* 117–24.

Sloboda, J. (1985). *The musical mind: The cognitive psychology of music.* Oxford: Oxford University Press.

Sloboda, J. (1991a). Empirical studies of emotional responses to music. In M. R. Jones and S. Halloran (Eds.), *Cognitive bases of musical communication.* Washington, DC: American Psychological Association.

Sloboda, J. (1991b). Music structure and emotional response: Some empirical findings. *Psychology of Music, 19,* 110–20.

Smith, D. (1987). Conflicting aesthetic ideals in a musical culture. *Music Perception, 4*(4), 373–92.

Smith, D., & Witt, J. (1989). Spun steel and stardust: The rejection of contemporary compositions. *Music Perception, 7*(2), 169–86.

Spender, N. (1987). Psychology of music. In R. L. Gregory (Ed.), *Oxford companion to the mind.* Oxford: Oxford University Press.

Stacey, P. (1989). Towards the analysis of the relationship of music and text in contemporary compositions. *Contemporary Music Review, 5,* 9–27.

Steedman, M. (1984). A generative grammar for jazz chord sequences. *Music Perception, 2,* 52–77.

Stefani, G. (1974). Progetto semiotico di una musicologia sistematica. *International Review of the Aesthetics and Sociology of Music, 5,* 277–89.

Stravinsky, I. (1947). *Poetics of music in the form of six lessons.* New York: Vintage Books.

Tan, N., Aiello, R., & Bever, T. G. (1985). Harmonic structure as a determinant of melodic organization. *Memory and Cognition, 9,* 533–39.

Wachsmann, K., & Cooke, P. (1980). Africa. In S. Sadie (Ed.), *The new Grove dictionary of music and musicians* (pp. 144–53). London: Macmillan.

Wapnick, J., Bourassa, G., & Sampson, J. (1982). The perception of tonal intervals in isolation and in melodic context. *Psychomusicology, 2*(1), 21–36.

Warren, R. (1970). Perceptual restoration of missing speech sounds. *Science, 167,* 392–93.

Wedin, L. (1972). A multidimensional study of perceptual-emotional qualities of music. *Scandinavian Journal of Psychology, 13,* 1–17.

Winn, J. A. (1981). *Unsuspected eloquence: A history of the relations between poetry and music*. New Haven, CT: Yale University Press.

Youngblood, J. (1958). Style as information. *Journal of Music Theory*, 2, 24–35.

Zatorre, R., & Halpern, A. (1979). Identification, discrimination and selective adaptation of simultaneous musical intervals. *Perception and Psychophysics*, 26, 384–95.

# 3

# Perception: A Perspective from Music Theory

NICHOLAS COOK

## Introduction

The study of the relationship between music theory and psychology can be traced to the work of Leonard Meyer (1956, see also Chapter 1 in this volume) and Christopher Longuet-Higgins (1976, 1978). It can also be traced to Marvin Minsky, a leading exponent of artificial intelligence, who in 1981 asked: "What might we discover if we were to study musical thinking? Have we the tools for such work?" And then went on to recommend that "music theory is not only about music, but about how people process it. To understand any art, we must look below its surface into the psychological details of its creation and absorption" (pp. 28–29).

The last decade has seen a blossoming of the research in the cognitive psychology of music (see Deutsch, 1982; Dowling & Harwood, 1986; Howell, Cross, & West, 1985; Jones & Holleran, 1992; Longuet-Higgins & Lisle, 1989; McAdams, 1987; Sloboda, 1985). Indeed, psychologists have become interested in music because the perception and cognition of music reflect mental structures. Their work has yielded models to test music theory empirically, and at the same time, some music theorists have placed musical analysis within a wide spectrum of knowledge that draws from linguistics and information theory (see Bent, 1987; and Lerdahl & Jackendoff, 1983).

Broadly speaking, musicians and music psychologists have a similar goal: to know how music is perceived and understood. The interest of music theorists and musicians is in understanding better the musical structure and musical interpretation of actual compositions; cognitive psychologists, on the other hand, have been more interested in researching mental theories of how specific musical events may be perceived. Their emphasis has been primarily to

investigate the perception of separate musical parameters rather than the musical experience.

In this chapter, Nicholas Cook acknowledges the growing interest in the psychology of music, but argues that cognitive psychology and music theory are two disciplines with essentially different and mutually incompatible aims. Cook challenges the notion that music theory ought to be considered a branch of psychology. He bases his position on the fact that psychological research has placed too much emphasis on the psychoacoustical parameters of music and has not addressed questions intrinsic to the meaning and the cultural value of music. Throughout this chapter, Cook points out that there are potential pitfalls in applying general psychological theories to music without taking into account what listeners actually hear, and why.

As an example of this research, Cook refers to *A Generative Theory of Tonal Music* by Lerdahl and Jackendoff (1983), which proposes that tonal music is based on a grammar. They suggest that in order to understand tonal music the listener needs to know its grammar. Since their book provides a model of the knowledge necessary to listen to tonal music, it is a contribution to both music theory and psychology. For Lerdahl and Jackendoff, the same grammatical principles of tonal music apply whether one listens to individual musical phrases or to entire tonal compositions.

Cook here argues with several aspects of Lerdahl and Jackendoff's theory and quotes experimental evidence demonstrating that people do not listen to music according to large-scale structures. In addition, Cook points out that studies on the recognition of intervals, chord progressions, and key centers are merely tests of ear training. They do not consider the significance of music. He laments: "[P]erceptual psychologists assume that music is made out of notes." Cook concludes by proposing that there be a more careful translation of the concepts from one discipline to the other.

## REFERENCES

Bent, I. (1987). *Analysis.* New York: Norton.

Deutsch, D. (Ed.). (1982). *The psychology of music.* Orlando, FL: Academic Press.

Dowling, W. J., & Harwood, D. (1986). *Music cognition.* Orlando, FL: Academic Press.

Howell, O., Cross, I., & West, R. (Eds.). (1985). *Musical structure and cognition.* London: Academic Press.

Jones, M. R., & Holleran, S. (Eds.). (1992). *Cognitive bases of musical communication.* Washington, DC: American Psychological Association.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.

Longuet-Higgins, C. (1976). Perception of melodies. *Nature, 263,* 646–53.

Longuet-Higgins, C. (1978). Perception of music. *Interdisciplinary Science Review, 3,* 148–56.

Longuet-Higgins, C., & Lisle, E. R. (1989). Modeling musical cognition. *Contemporary Music Review, 3,* 15–27.
McAdams, S. (1987). Music: A science of the mind? *Contemporary Music Review, 2,* 1–61.
Meyer, L. B. (1956). *Emotion and meaning in music.* Chicago: University of Chicago Press.
Minsky, M. (1981). Music, mind, and meaning. *Computer Music Journal, 5*(3), 28–44.
Sloboda, J. A. (1985). *The music mind: A cognitive psychology of music.* Oxford: Clarendon Press.

## Music and Psychology: A Mutual Regard?

There is no doubt that music theorists and perceptual psychologists have become better acquainted with one another's work in recent years. Journals have been set up, such as *Music Perception*, which welcome contributions from either camp. There have been conferences with titles such as "Music and Cognition" which are equally cross-disciplinary in thrust. And, as one might expect, there have been friendly and cooperative programs for cross-disciplinary research. David Clarke, a music theorist, recommends Schenkerian analysis to psychologists as a source of formal theories that may be transformed into empirically testable hypotheses (1989: 130). Carol Krumhansl, a psychologist, recommends empirical data to music theorists as an aid to framing theory in more psychologically relevant terms (1990a: 287). And Jay Dowling puts this picture of mutual cooperation into a nutshell: "Experiments provide tests of theories, and theories show experiments how to define phenomena" (1989: 249).

Of course, the relationship between music psychology and theory is not always so harmonious. One problem has been when musicians draw on perceptual data to justify their theoretical edifices. A classic example of this is the composer and theorist Milton Babbitt, who justified his theory of combinatorial serialism by saying that the system is built out of interval classes, and therefore anyone who can identify pitch classes can hear combinatorial structure (1962: 120). Another is the composer Karlheinz Stockhausen (1959), who justified his use of logarithmic relationships of duration on the grounds that logarithmic relationships are found in the perception of pitch, and that pitch is simply duration transferred to a higher level;[1] therefore, Stockhausen concluded, logarithmic duration structures are perceptually privileged. In each case, the argument—Eric Clarke (1989a: 11) calls it "sleight of hand"—works the same way. A perceptual principle that applies to one set of circumstances (identifying

an interval in an ear-training class in Babbitt's case, the logarithmic nature of pitch perception in Stockhausen's) is generalized to a quite different situation, without any consideration as to whether the generalization in fact has any perceptual validity. These are merely extreme and notorious examples of a quite general music-theoretical ploy, the purpose of which is essentially rhetorical: make an initial bow in the direction of the psychology of perception and then carry on with your theorizing regardless, but claiming the enduring benediction of psychology.

Psychologists of music have on occasion lost patience with theorists' cavalier attitudes toward perception. Vladimir Konečni and Heidi Gotlieb (hereafter referred to jointly as Konečni) carried out a number of tests in which well-known works from the classical canon—Bach's *Goldberg* Variations, Beethoven's sonatas—were played with their movements in the wrong order. College-level listeners showed no preference for the original ordering of the movements over the modified one (Konečni, 1984; Gotlieb & Konečni, 1985). Konečni then quoted a number of assertions by theorists to the effect that the movements *had* to be in the order the composer put them in because of considerations of higher unity. (It is worth pointing out that, as is often the case in psychological writings on music, the theoretical sources cited are not held in high regard by contemporary practitioners; genuine theories of intermovemental ordering, as opposed to the kind of old-fashioned pontificating Konečni cited, are in fact quite rare.) And from the contrast between what the theorists said and what the experiments showed, Konečni concluded sourly that "[a] greater degree of caution, moderation, and humility in the music critics' and theorists' often sweeping claims—mere speculations really—would be a welcome consequence of the type of research in the psychology of music that this article advocates" (Gotlieb & Konečni 1985: 98). Or to put it another way: let theorists submit their findings to the psychological tribunal.

Musicians have generally been less vocal in their complaints about music psychologists; they are inclined to mutter that psychologists do not understand music (or, as Eugene Narmour [1989: 97] puts it, "most music psychologists are not sufficiently well-trained to understand how hierarchical patterning emerges in music"), and leave it at that. One symptom of this is the tendency for psychologists to use test materials which are so musically impoverished that they do not really provide a context for *musical* perception at all. This has been widely criticized, and I am not going to repeat the arguments here.[2] There are other respects, too, in which psychological work can strike musicians as being unrealistic. And here a specific example may help.

The basic message of Rita Wolpert's article "Recognition of Melody, Harmonic Accompaniment, and Instrumentation: Musicians vs. Nonmusicians" (1990) is similar to Konečni's: untrained listeners do not listen to music in the same way that musicians do, and what matters to them is not the same as what matters to music theorists. Wolpert's starting point is that theorists view melody, harmony, and rhythm as the essential constituents of music, with instrumenta-

tion having a secondary function. (This is as much as to say that "Twinkle, Twinkle, Little Star" is "Twinkle, Twinkle, Little Star" regardless of whether it's played on a soprano saxophone, a sitar, or a sine wave generator.) The aim of Wolpert's research was to find out whether this corresponds to the perceptions of untrained listeners. She carried out a series of experiments, in the first of which trained and untrained subjects heard (1) a tune played on a given instrument (for instance, "Twinkle, Twinkle, Little Star" played on a celesta), then (2) the same tune played on another instrument ("Twinkle, Twinkle, Little Star" played on an oboe), and finally (3) a different tune played on the first instrument ("Hot Cross Buns" played on a celesta).[3] The subjects had to say whether (2) or (3) was "most like" (1). Whereas all the musicians chose (2)—that is, they chose the same tune on a different instrument—nearly half the nonmusicians chose (3)—that is, they chose the same instrument playing a different tune.

Wolpert then conducted two further experiments. In one of them, the same tunes were played as before, but this time they were played with simple harmonic accompaniments; the results were much the same as in the case of the melodies alone. In the other, listeners heard (1) a tune on a given instrument, (2) the same tune on a different instrument, and (3) the same tune on the original instrument, but now with the accompaniment transposed down a fifth, so that it was in a different key from the tune. This time 40 of the 42 nonmusicians chose (3); for them, the identity of the instrument outweighed any changes they noticed in the accompaniment. The musicians, on the other hand, continued to choose (2); for them, playing the accompaniment in the wrong key made a bigger difference than playing the music on a different instrument.

What does all this mean? According to Wolpert (pp. 103–4), it shows that

> Musicians and nonmusicians do not . . . listen by the same rules. Musicians consistently chose melody and correct harmonic accompaniment over instrumentation when matching an excerpt to a model, but nonmusicians did not do so. . . . This choice of instrumentation over correct harmonic accompaniment . . . has important ramifications. It suggests a profound overestimation of what it is that most listeners hear. . . . The implications of these findings are sobering. . . . Music educators may need to teach a new awareness of harmonic accompaniment and key relationships in a more systematic way than they are doing at present. Further, this research may lend empirical verification to what music therapists have long known: instrumentation is attended to and found of interest. . . . The need for empirical verification of what it is that untrained listeners actually attend to can and should inform music professionals.

But a musician might be inclined to retort that it shows nothing of the sort. What it shows is how different listeners respond to questions as to whether one musical extract is more or less like another. Wolpert anticipates objections regarding what she asked the subjects to do: "It was necessary to make the instructions vague, because it was important that the instructions not direct the subject's attention to any one aspect of the music, nor in any way make clear the musical

assumptions of the culture" (p. 102). But the result is that the two groups of listeners in Wolpert's tests may well have been engaged in different tasks.

To be a musician trained in the Western tradition means to operate within a language game organized around the concept of "things to play," with a "thing to play" being defined primarily in notational terms (that is, in terms of pitch and duration). You play "the same piece" in one key or another, in one style or another, on one instrument or another—and often to very different effect. When Western musicians, then, are asked whether "Twinkle, Twinkle, Little Star" played on a celesta is most like "Twinkle, Twinkle, Little Star" played on an oboe or "Hot Cross Buns" played on a celesta, they will generally choose "Twinkle, Twinkle, Little Star" played on an oboe because, for them, it is the same piece. This does not mean that they do not care whether the tune is played on a celesta or an oboe. It does not mean that they think other listeners do not care whether the tune is played on a celesta or an oboe. It does *not* mean that they are unaware that "instrumentation is attended to and found of interest." It just means that, when they reply to questions about what they have heard, they do so in terms of the language game of Western musicianship.[4]

But there is no corresponding language game to govern the responses of untrained listeners; that is part of what it means to say that they are untrained. Asking such listeners whether extract (1) is most like extract (2) or extract (3) is rather like asking them whether a melon is most like a banana or a football. It is a bit like both, only in different ways. In culinary and botanical terms, melons are much more like bananas than footballs. On the other hand, imagine an Englishman trying to order fruit in a Tunisian restaurant: "No, no, *not* a banana, what I want is like a *football!*" It is the situation that gives content to the comparison; posed in the abstract, the question whether a melon is more like a banana or a football is pretty meaningless and can be answered equally well either way. The same applies to Wolpert's tunes. Of the untrained listeners in the first two experiments, roughly half said that one excerpt was like another because it had the same tune, and roughly half said it was like another because it was played on the same instrument. In the third experiment, the untrained listeners may or may not have noticed that something had gone wrong with the accompaniment, but most of them said that one excerpt was like another when it had the same tune and was played on the same instrument. This all seems quite straightforward, and it is not an adequate basis for Wolpert's conclusion that "musicians and nonmusicians do not . . . listen by the same rules." In fact, it is not really an adequate basis for saying anything about how people listen to music.

What we have here, as often in psychological writings about music, is an attempt to give a purely psychological explanation for what are in part social phenomena. As I have tried to show, how people respond to psychologists' questions about music is largely a matter of language games, which are essentially social; how they listen to music (for instance, when there are no psychologists about) is quite a different matter. It is because Wolpert does not distinguish

between these two things that she arrives at an astonishing conclusion: "There is an immense gulf between the creative intent of the composer or improviser and the perception of the audience. Harmonic practice in classical music, as well as popular, rock, country, jazz, and middle of the road music, shares a similar traditional relationship between melody and harmonic accompaniment of which these listeners are unaware" (p. 104). According to Wolpert, then, even easy-listening music passes over the heads of its listeners; music educators please take note!

In saying this, Wolpert introduces an additional idea, again one that is frequently referred to in the music-psychological literature: the idea of the composer's intent being realized in perception. And this raises the question of how far music succeeds in communicating something from composer to listener.

## Do People Hear Tonal Structure?

A good way to approach this question is by focusing on cases in which there is an apparent failure of communication between composer and listener. In his article "Cognitive Constraints on Compositional Systems" (1988a), Fred Lerdahl cites Pierre Boulez's *Le Marteau sans Maître* as an example of such a failure. As he says, *Le Marteau*, which was completed in 1954, was widely hailed as a masterpiece of postwar serialism, yet for years nobody could figure out just how the serial structure worked. Only recently has the code been cracked. But nobody hears the music differently because of this; the serial structure of *Le Marteau* is so complex and abstract that you cannot hear it, even if you make a special effort to. Here, then, we have a mismatch between the structure the composer puts into the piece and the mental representation of it that the listener forms; Boulez wrote a serial composition, but all that listeners hear is more or less pretty sounds. Lerdahl refers to this as a disjunction between "compositional grammar" and "listening grammar."

As Lerdahl's terminology makes clear, the idea that music involves the encoding and decoding of some kind of message—something that passes from the composer's brain to the listener's—is based on an analogy with language, and this analogy is central to the interaction between music theory and psychology with which this chapter is concerned. Lerdahl is the co-author, with Ray Jackendoff, of *A Generative Theory of Tonal Music*,[5] which might be described (not quite accurately) as a reformulation of established theoretical thinking on pitch and rhythm within the general framework of transformational grammar. This means that they are presenting insights drawn from Schenkerian analysis, Cooper and Meyer's rhythmic theory, and other sources as contributions toward "an explicit formal musical grammar that models the listener's connection between the presented musical surface of a piece and the structure he attributes to the piece" (1983: 3). In other words, Lerdahl and Jackendoff are saying that music theory is really a branch of psychology: "A piece of music is a mentally con-

structed entity, of which scores and performances are partial representations by which the piece is transmitted. . . . The central task of music theory should be to explicate this mentally produced organization" (p. 2).

Lerdahl and Jackendoff offer their theory to psychologists as a source of predictions regarding the perception of music (1984: 230)—predictions that can be subjected to empirical investigation, along the cooperative lines that I mentioned at the beginning of the chapter. Psychologists have responded to the challenge, and indeed a cursory survey of current writing in music psychology might suggest that *GTTM* is the dominant model in tonal music theory, almost to the exclusion of any other. (The picture is rather different in the music-theoretical literature, where Lerdahl and Jackendoff's work is widely admired, but people tend to stick to the established analytical methodologies or invent their own.) The first major experiment testing Lerdahl and Jackendoff's predictions was carried out by Irène Deliège (1987), who played both musicians and nonmusicians a series of short musical extracts ranging from Bach to Stravinsky. The listeners were supplied with simplified scores of the music, in which each melody note was represented by a dot; they had to draw vertical lines on these scores to show how the music fell into separate phrases or other segments. The subjects' responses were then compared with the segmentation predicted by *GTTM*. On the whole, there was a reasonable correspondence between the two; the theory worked better with musicians than nonmusicians, but Deliège comments that "musical training does not appear to induce the emergence of a grouping behavior radically different from the one used by the naïve listener" (1987: 344).

Now it is almost always possible to raise objections to psychological tests of musical listening, for the simple reason that listeners do not generally listen to music under test conditions, and there is therefore a problem as to how far the results of tests can be generalized to normal listening. (In this case, the question would be to what extent information about how people divide up dots on paper tells us about how people listen to music in the ordinary way.) If, however, we accept Deliège's conclusions—and they seem plausible enough—then how much of *GTTM* have we validated? The answer is: the initial segmentation of the musical surface, but not the larger hierarchical structure that Lerdahl and Jackendoff derive from the initial segmentation. From the standpoint of the music theorist, it is this hierarchical structure that lies at the heart of *GTTM*, because it shows how compositions can be coherent in a specifically *tonal* manner. We can clarify this by referring to the way in which Lerdahl has extended the original *GTTM* to cover atonal music. The essential part of his argument runs as follows (1989: 73–74):

> Given that two events connect, the more *stable* is the one that is more consonant or spatially closer to the (local) tonic; the more *salient* is the one that is in a strong metrical position, at a registral extreme, or more significant motivically. If one event turns out to be more stable and more salient, it unambiguously dominates the other. But if one event is more stable and the other is more salient, there is a

> conflict in the rules. In tonal music stability almost always overrides salience. One might say that the grammatical force of tonal pitch structures can be gauged by their ability to override surface salience. Atonal music may not have stability conditions, but it does project the relative salience of events. The absence of stability conditions makes salience cognitively all the more important. I argue that listeners organize atonal surfaces by means of it. As a result, atonal music collapses the distinction between salience and structural importance.

And a few pages later Lerdahl remarks that this "amounts to an acknowledgement that atonal music is not very grammatical" (p. 94).

Several psychologists have carried out tests showing how this kind of salience influences segmentation. Deliège's (1987) experiments, mentioned above, document the role of such attributes as timbre, register, and dynamics on segmentation. Again, in an experiment where listeners were asked to match musical excerpts with middle-ground reductions, Serafine, Glassman, and Overbeeke found that "metric accent is a major determiner of the structural importance given to tones" (1989: 405). And in tests where musically trained listeners had to rate the strength of boundaries between segments in Mozart's Fantasie in C minor (K. 475), Eric Clarke and Carol Krumhansl found that segmentation was based on such salience-related criteria as change of texture, change of dynamics, and change of register; moreover, they found that these same criteria operated on different levels of structure, from the musical surface to higher-level grouping involving durations of 30–50 seconds (1990: 248–9).

But there is a problem. All these tests were based on tonal music—music, that is, in which criteria based on salience should (according to Lerdahl's theory) have been overridden by criteria based on stability. This emerges with particular clarity from Clarke and Krumhansl's experiments, which compared responses to the Mozart Fantasie with responses to a highly atonal piece: Stockhausen's *Klavierstück IX*. Listeners (and remember these were musically trained listeners) used very much the same criteria in segmenting the Stockhausen piece as the Mozart; in fact, the tables in which Clarke and Krumhansl summarize these criteria for the two pieces are strikingly similar (pp. 227, 243). As the authors comment, there are only two differences between them: "the greater emphasis on pauses and silences as grouping factors in the *Klavierstück* than in the Fantasie, which is a direct result of the more continuous nature of the music in the Mozart, and the greater importance of repetition as a boundary indicator in the Stockhausen" (p. 249). What is strikingly absent from the criteria adopted in the Mozart is any sign of specifically *tonal* features, such as modulations, cadence points, or tonal closure (the return of a tonal excursion to the chord or key with which it began). The conclusion seems inescapable: if people (musically trained people) listen to tonal and atonal music in much the same way, and if atonal music is not very grammatical, then tonal music cannot be very grammatical either.

A few years ago, I carried out a series of tests designed to address the perception of tonal closure in a more direct manner. The significance of tonal closure is that it represents the basic principle of virtually all theories of tonal music—not

just Lerdahl and Jackendoff's, but Meyer's, Schenker's, and Schoenberg's. (It is a much more important principle for music theorists than the issue of intermovemental ordering that Gotlieb and Konečni were dealing with.) Put rather too simply, the theory works like this. Tonal compositions begin and end in the same key; that is, a sonata or a symphony is in G major or F-sharp minor, even though it may well contain movements in other keys. Similarly, an individual movement will be in a given key, even though it will probably contain passages in other keys. The idea is that the "home" key, or tonic, represents a point of stability; excursions to other keys represent an increase of tension, which is resolved by the return to the tonic. So the typical sonata movement, which begins on the tonic, moves to the dominant and passes through a series of other keys before returning to the tonic, represents an archlike increase and decrease of tension. The theory, then, purports to explain the practice of tonal composers in terms of the psychological effects that their music has on listeners.

As an illustration of this, consider the third of Liszt's *Fünf kleine Klavierstücke* (Figure 3.1).[6] Bars 1–16 are in F-sharp major, with no more than slight inflections toward the dominant (bars 6–7) and subdominant (bar 11); at the beginning of bar 16 there is a cadence on a diminished seventh, which can be understood as standing for $V_7$ of F-sharp. Liszt could easily have continued in F-sharp major, for instance by repeating the opening. But he does not do this. Instead, the music shifts unexpectedly into A major. (The voice-leading in the bass—the D of bar 16 falling to C-sharp—means that, although the A major is unexpected, it is not incoherent.) However, the sense of an F-sharp major tonic does not entirely disappear during the A major episode; the B major chord over a C-sharp bass at bar 22, which functions as a modified $V_7$ of F-sharp and leads to the final tonic, creates a sense of homecoming. The return to F-sharp resolves the tension created by the move to A, and the piece concludes with a satisfying finality. That, at any rate, is the received theoretical wisdom regarding this kind of tonal structure, and similar patterns are found—and explained along the same lines—in much more extended pieces than this *Kleines Klavierstück*.

The question is: do listeners actually respond to this large-scale tonal closure? A simple way to answer this is to play listeners two versions of the piece: the original, as shown in Figure 3.1, and a recomposed version in which the music does not end in the same key as it began in. Figure 3.2 shows such a recomposed version, following on from bar 14 of Figure 3.1; here the music continues, perfectly logically, in F-sharp major. This means that it ends in a rather hard-to-decipher D-sharp major. There is no tonal closure. Tonal theory, then, predicts that the recomposed version of the music will end inconclusively, without creating the satisfying sense of completion engendered by Liszt's version. It might, in fact, lead one to expect a less rich aesthetic experience. My tests, in which the subjects were music students, did not confirm these theoretical predictions.[7] In the case of this particular piece, there was a slight, and not statistically significant, preference for the original version over the recomposed one. In the case of other test pieces there were slight, and not statistically significant, preferences for

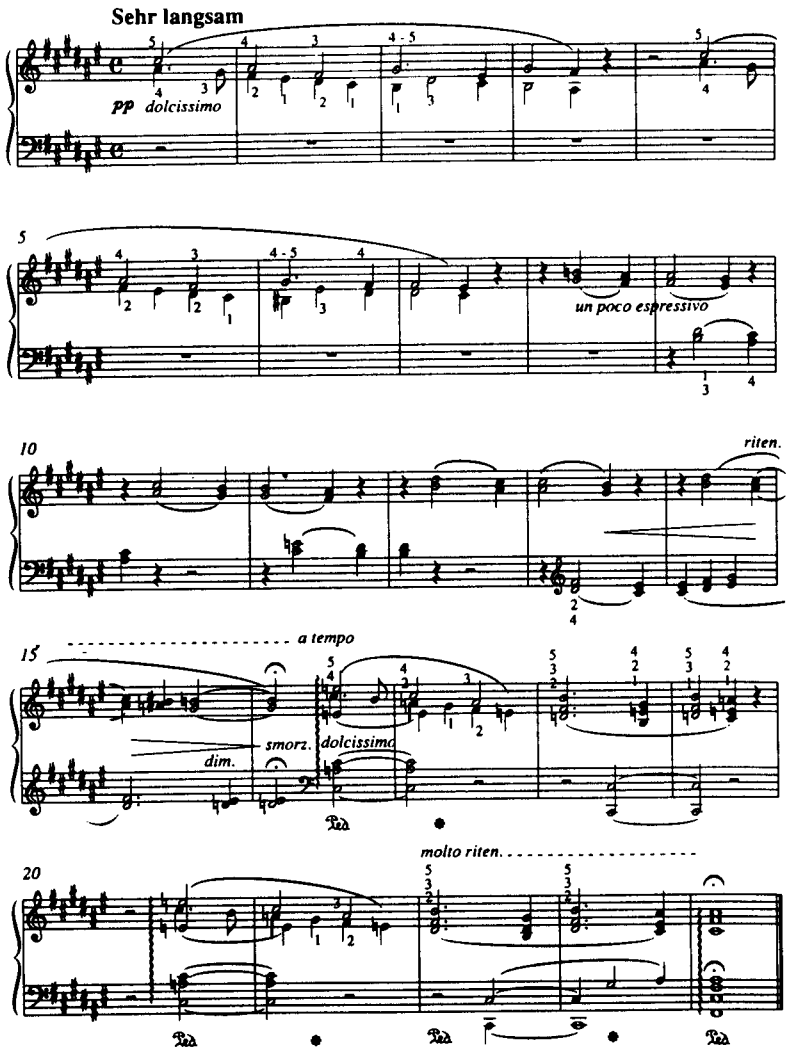



FIGURE 3.1. Liszt, *Kleines Klavierstück* no. 3.

the recomposed version over the original one. Only in the case of the two shortest pieces was there a significant correlation of tonal closure and listener preference, and even here preferences were partly the result of ordering effects (listeners tended to prefer whichever version they heard second over the one they heard first). To eliminate the effects of ordering, I carried out further tests. The listeners were divided into groups, each of which heard the two shortest pieces, but with the original and recomposed versions in a different order. This time only the very shortest piece yielded a significant correlation between tonal closure and listener response.

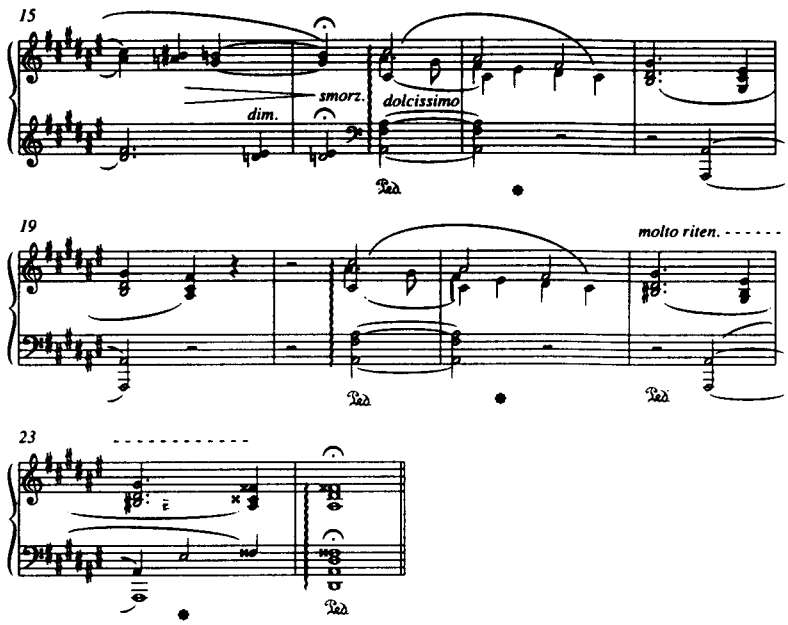



FIGURE 3.2. Liszt, *Kleines Klavierstück* no. 3, bars 15–24 (modified).

The shortest piece of music in these tests was about 30 seconds long,[8] while the next shortest was about 55 seconds. So the tests indicate that tonal closure only has a direct effect on listeners (even musically trained listeners) when the time scale involved is well under a minute. This means that, in perceptual terms, an extended composition cannot have the tonal unity that a single phrase has. Such a conclusion must be disturbing to anybody who expects tonal theory to be psychologically valid, in the sense of correctly predicting the effect of music on listeners. After all, the basic explanatory thrust of theories such as Schenker's or Lerdahl and Jackendoff's is that large-scale musical structures are to be explained along the same lines as small-scale structures. That is what it means to have a hierarchical theory of music; that is what it means to explain music by analogy with the hierarchical structures of transformational grammar.

Burton Rosner and Leonard Meyer have criticized the analogy between musical and linguistic hierarchies from just this point of view; their analysis is so penetrating that it is worth quoting in full (1986: 37).

> Analysts such as Schenker . . . and Lerdahl and Jackendoff . . . may well produce hierarchical diagrams of quite long musical passages. But if "higher" events exert increasingly tenuous perceptual influences, the psychological force of the outputs from such analytical machinery will dwindle quickly, whatever may be their value for the theory of music. In this regard, we must point out a fundamental difference between the tree structures used in linguistics and those presented by music theorists like Lerdahl and Jackendoff. The top node of a grammatical tree is an immediately observed datum: a sentence or an utterance. It represents

> some incident, occurring over time, which can be entered completely and rapidly into memory. The associated tree decomposes that uppermost node into parts at several lower levels of a strict hierarchy. The lowest nodes in music-theoretic tree structures, however, represent a datum: an actual stretch of music. Quite often, only fragments of it are held faithfully in memory. The lower nodes in the tree are not decompositions of higher ones. Instead, higher nodes are *selections* from among lower ones. We therefore cannot believe that the increasingly higher nodes, which represent ever more rarified selections, form the core of music perception.

Rosner and Meyer are saying that while the top node of a grammatical tree corresponds to something that is actually said or heard, the musical equivalent has a purely theoretical existence. A related observation—or maybe it is just another way of saying the same thing—is that, whereas the analysis of language converges on semantic content (Chomsky's "deep structure"), the analysis of music does not converge on anything; the "deepest" level of Schenkerian structure, the *Ursatz*, is simply an abstraction. It means nothing at all.

## Psychology of Music or Psychology of Ear Training?

One of the major problems in applying models drawn from linguistics to music theory, then, has to do with the significance to be attached to different levels of structure. Another set of problems can be focused around a further concept borrowed from the study of language: pragmatics. Now pragmatics comes in two flavors. What we could call the "mild" flavor means "studying the use of a language as distinct from, but complementary to, the language itself seen as a formal system" (Leech, 1983: x). Or to put it another way, pragmatics deals with what grammar leaves out; to say this is to define pragmatics in terms of grammar. We can focus what is at issue here around Chomsky's well-known distinction between "competence" and "performance." By "competence," Chomsky means the grammatical knowledge that a competent language user has internalized; as the use of the word "internalize" shows, this is considered to be a mental construct. "Performance," on the other hand, refers to the way in which this knowledge is used in actual situations. If theories of competence deal exclusively with perfect, well-formed sentences, theories of performance deal also with interruptions, syntactic breakdowns, failures of comprehension, and the goals that people have in using language in specific contexts (the last of these being the subject matter of pragmatics more narrowly defined). And it is an axiom of formal linguistics that issues of performance cannot be properly understood except on the basis of a theory of competence: as Eytan Agmon puts it, using a different terminology, "one cannot answer the question 'how' before having answered the question 'what'" (1990: 298).

This principle has been transferred intact into linguistically orientated music theory. Lerdahl and Jackendoff state it thus: "In our view it would be fruitless to

theorize about mental processing before understanding the organization to which the processing leads" (1983: 4). But there are distinct problems in applying the idea of competence to music. The classic definition of a grammar is a finite set of rules that will generate all, and only, well-formed sentences in a given language; as I said, competence means having internalized such a grammar. Now it is not at all obvious how a clear distinction is to be drawn between musical structures that are well formed and those that are not. To be sure, there are some styles so constrained that one can always say with assurance that this note is right whereas that note is wrong: species counterpoint is one. But then species counterpoint is a pedagogical construct, a demonstration of rules; Palestrina notoriously did things that counterpoint textbooks don't allow. Again, imagine misprints in a Schubert song: some would be obviously wrong (that *can't* be an F-sharp!), but others would be merely questionable (Schumann might write that F-sharp, but surely not Schubert?). To cope with the last case we would need to hypothesize a grammar for Schubert that is different from a grammar for Schumann. And obviously, the process could continue until we had different grammars for early Schubert and late Schubert, for songs and symphonies, for heroic songs and sentimental ones, and so on.

As Eric Clarke (1989b) has said, there are many factors that militate against the usefulness of explaining music in terms of strict grammars: the ways in which composers consciously mix musical styles, the coexistence of different musical styles within a single cultural community, the rate at which style change occurs in music. He explains these things in terms of the different functions served by language and music (p. 19):

> Natural language is essential to the cohesion and survival of human social organization, and hence must make use of structural constraints (grammar) which, by remaining relatively unchanging, ensure successful communication. By contrast, music in European art culture, although serving as an important cultural indicator, does not bear the same functional responsibility, and thus has less need for long-lasting structural constraints. Each piece can make use of a substantial number of principles that are specific only to that work, and which are consequently inexplicable (or at best explicable only at a very general level) in terms of a broad and general structural theory.

Music, in other words, is much more fluid than language. Before you start a conversation, you know what the grammar is going to be; in music, you can never be sure of the grammar until you have heard the piece.

Alan Marsden (1989) has outlined a model of musical listening based on the idea of discovery learning. This model suggests that, as somebody listens to music, certain parsing rules (the kind of rules that *GTTM* is based on) are dynamically strengthened or weakened, according to how well they fit the music. Moreover, the model allows for new rules to be derived. What Marsden is suggesting, then, is that listening to music involves creating a grammar to fit the piece being heard. But to say this is, in effect, to abandon the idea of grammar

altogether, at least as it is understood in formal linguistics. For it means that a musical grammar is essentially a collection of strategies for understanding music.[9] In other words, it means that the distinction between competence and performance cannot be sustained as applied to music. So it comes as no surprise that Lerdahl and Jackendoff fail to sustain it in their own work. As Uwe Seifert has pointed out (1991), *GTTM* is not really a theory of competence at all. It is a model of the way in which a structural description may be derived from a musical surface. It is an input-output mechanism. That is, it is a theory of performance, though an incomplete one.

We started with what I called the "mild" flavor of pragmatics. But by now we have advanced to the "strong" flavor, which maintains that the whole idea of a formal grammar (and therefore the distinction between performance and competence) is a theoretical artifact, an illegitimate abstraction from the real-world phenomenon of language. Roy Harris (1980, 1981) is a leading proponent of this position, and he argues that formal linguistics is a pseudoscience that has come into being by adopting the principles of traditional grammar (which mainly means Latin grammar), and then claiming what traditional grammarians never claimed, that these principles have some kind of psychological validity. In other words, pedagogical principles (for traditional grammar was a system of teaching) have been reinterpreted as psychological ones. Harris sees the competence/performance distinction as playing a crucial role in this transition, because it means that while grammatical rules are asserted to be "in the brain," it is impossible to prove or disprove their existence; it is only the *use* of these rules, that is to say performance, that can be subjected to psychological investigation. If experiments cast doubt on any given model of language, that need not affect the issue of competence; it can always be said that, by its nature, performance is determined by all sorts of factors, and the model under investigation has failed to make proper allowance for these (1981: 109). In this way the linguist's claim to be talking about psychological realities is secured against all possible counterevidence from psychologists.

I am going to argue that there has been a corresponding, if less thoroughgoing, transformation of music theory (which again is in essence a system of teaching) into a theory of perception. Perhaps the best way to approach this is to look at a key issue in Harris's account of the transformation of traditional grammar into formal linguistics and see how it applies to music. This is the issue of *scriptism*.

Harris defines scriptism (a word of his own coinage) as "the assumption that writing is a more ideal form of linguistic representation than speech" (1980: 6). Modern linguists vociferously disassociate themselves from any such idea, insisting that speech is primary and that the written word is not to be understood as anything other than a representation of the spoken word. But Harris argues persuasively that, in reality, linguists analyze speech on the model of writing. They regard *Mary had a little lamb* as a linguistic unit that can be articulated in two different ways (depending on whether it is a reply to *What did Mary have?* or

*Who had a little lamb?*), rather than as two basic linguistic units that happen to be written the same. They regard words as basic linguistic units, and as late as the 1950s were defining them by reference to "potential pauses" in speech—clearly on the model of the space on either side of the written word. They regard phonemes as linguistic units that happen to correspond rather closely to letters, and as late as the 1970s were trying to analyze speech into fixed sound units corresponding to phonemes (1981: 57). Linguistics, as Harris puts it, is cryptoscriptist.

What about music theory? Again, we have the vociferous insistence that the proper object of study is the sound of music, not the written score. Except in a few special instances (say a study of music preserved in an imperfectly understood notation), no music theorist would care—or dare—to say that she was interested in the score as anything other than a representation of the sound, the listener's experience of the sound, or both. In other words, music theorists *use* scores, but they do not *study* scores—they study the music. Or at least that is the claim. But, in fact, it is obvious that the claim is not always true. Consider a set-theoretical analysis of a piece of atonal piano music, say Schoenberg's *Sechs Kleine Klavierstücke*, op. 19, no. 3. Set theory[10] deals with the relationships between the individual notes you can see in the score. Of course, you hear the notes when you listen to the music. Or do you? Well, you hear a succession of composite sounds which, if you have a good musical training, you may be able to distinguish into their component pitches. But then again, you may not; there is no law saying that, in order to do an analysis of atonal piano music, you have to be able to "hear out" each note of the music. Of course, you might hear each note in the sense that if a wrong note was played, or if a note was left out or added, you would recognize that something was wrong. But then again, you might not; in practice it would probably depend on what the note in question was. In other words, there is, to put it mildly, a nonlinear relationship between the notes in the score and what people hear when they listen to a performance of it. Set theory takes no account of this.

At least, in the case of atonal piano pieces, the note is actually there, in the sense that either the pianist plays it or he doesn't. But there is a lot of music in which even this is doubtful proposition. An opera singer sings a rapid scale: does she actually sing every note? Is there a clear transition between each note and the next, or does one slide into the other? If so, how do you tell where one note stops and the next one starts? Or is there just a rather uneven glissando? Is the only way of answering these questions to record the sound and slow it down or print out a time/frequency graph of it? But if you have to do this to decide what notes are there, can you really say they are there in a *musical* sense? Problems of this sort become particularly pressing in the case of the florid singing characteristic of many Asian cultures. Robert Gjerdingen (1988) recently published a computer analysis of extracts, a few seconds long, from recordings made by the South Indian singer Gayathri Kassebaum. As he explains, Kassebaum's voice is constantly in motion; traditional melograms (the ethnomusicologist's word for com-

bined time/frequency and time/intensity graphs) reveal a bewildering multiplicity of events. Yet comparison between Kassebaum's different recordings of this passage, and between Kassebaum's recordings and students' imitations of her singing, reveals a much smaller number of events that remain invariant between the different performances. It is presumably these that are the notes—that is, if there *are* notes in Kassebaum's singing.

Gjerdingen offers a transcription of the passage into Western notation based, not on traditional melograms, but on graphs of the combined rates of change in frequency and intensity; these graphs throw the invariant events into relief. But he observes that such graphs can provide no more than a "good foundation" for transcription, because any musical notation involves cultural values (pp. 59–60):

> When we move from the analysis of curves, functions, and rates of change to a discussion of notes or particular rhythms, we are moving towards the consideration of cultural units of meaning, not simply pitches and intensities. Sensitivity is required when different cultures have differing ideas about the nature of such units.

Another way to put this is that the notes are not "in" the recording of Kassebaum's singing. Rather, they are constructed *out* of it through appropriate (or inappropriate) interpretation. And this is also true of Western music. There are no notes "in" the opera singer's rapid scale; if we interpret what we hear as a scale, a series of notes, and not as a rather uneven glissando, that is because in most operatic styles we expect scales rather than glissandos. (If she were singing music by Ligeti, we might hear the glissando.) Again, music students often have difficulty when they are asked to write down the inner lines of four-part chorale settings, which is part of the traditional music curriculum. It's not that they hear the lines but can't write them down; it's that they can't hear them. Later, when they have learned more harmony, they find that they can hear the inner lines quite easily. This is because they now know how to "compose" possible inner lines, so to speak, and compare these with what they hear. In other words, they have learned how to *create* notes and lines in what was previously a composite, undifferentiated mass of sound. Untrained listeners, of course, cannot do this. For them, there are no notes in the inner lines. Or rather, there are no inner lines—just, at most, inner parts.

If music theorists really wanted to analyze how music sounds, and only how it sounds, they would put away their scores and forget all about notes. Some phenomenologically minded theorists have tried to do just this. The trouble is that they almost invariably end up creating scores of their own. The only difference is that their scores tend to be relatively impoverished representations of the music, so the analytical results are generally unsatisfying from a musical point of view. This is not really surprising. In the first place, music theorists can no more forget about notes than linguists can forget about letters; they are part of our mental equipment. In the second place, any music theorist who achieves the impossible and forgets about notes "risks being left with a shapeless mass which

he does not know what to make of" (Harris, 1980: 17). Harris is talking about language, but what he says applies with no less force to music. For all music has its source in mental representations based either on notation or on kinesthetic imagery derived from singing or playing musical instruments; to compose is to manipulate images (Cook, 1990: Chapters 2, 4). It stands to reason, then, that the kind of full and satisfying comprehension of music at which the theorist aims is unlikely to be attained, except on the basis of a representation as richly determinate as that in terms of which the music was conceived. For this reason, there is nothing wrong in principle with the theorist who carries out a set-theoretical analysis of notes that he could not pick out from the musical sound. And in any case, he may well end up hearing them. It is often claimed that one of the benefits of music theory is that it enables people to hear music better. In fact, you might say that the main point of analyzing music is to enable us to respond to the excess of music over sound. I shall come back to this later.

This kind of defense against charges of scriptism is not available to psychologists, whose studies of musical perception are meant to be descriptive rather than prescriptive (that is, we expect them to tell us how people hear music, not how they might best hear it). And even a cursory examination of the music-psychological literature will show that such charges are warranted. The general picture that emerges is that, as Eric Clarke puts it, listeners "accumulate and integrate individual notes of a performance . . . into successively more abstract and global representations" (1989a: 4); we have studies of how listeners distinguish the voices of polyphonic music, discriminate intervals, follow chord progressions, and judge key centers. Now people can do all these things, at least within certain constraints. They are trained to do these things if they become music students. (They have to be trained, because otherwise they can't do them.) Such training—called ear training, aural training, or aural analysis—creates the interface between musical sound and the theoretical knowledge in terms of which musicians create, notate, and reproduce music. So we have a lot of psychological studies of ear training. The trouble is that they are not called studies of ear training. They are called studies of musical listening. And as such they are thoroughly unsatisfactory, because they begin with the premise that people hear music in terms of music-theoretical categories. They assume what music theory assumes; that music is made out of notes. This is scriptism with a vengeance. I am tempted to improve on Harris's coinage and call it *theorism*.

## Theory of Music or Theory of Mind?

Perhaps the most revealing example of theorism is tonal hierarchy theory. This theory, most conspicuously represented by the work of Carol Krumhansl, can be traced to two sources: one theoretical, and the other psychological. The theoretical source is a tradition of representing key relationships in topological terms that originated in eighteenth-century Germany, the best-known example being

the "chart of the regions" in Schoenberg's *Structural Functions of Harmony* (Figure 3.3). Schoenberg's chart shows a two-dimensional space with minor thirds on one axis and perfect fifths on the other, and its purpose is to indicate how closely different tonal regions are related to one another. For instance, the chart shows the distance from C major to F-sharp major as being much greater than the distance from C major to G major. What this means is that, in writing music in the tonal style, it is not easy to make the relationship between C major and F-sharp major sound coherent; modulating from C to F-sharp major requires a lot more care, if it is to be done convincingly, than modulating from C to G major. The purpose of the chart, then, is pedagogical. *Structural Functions of Harmony* is a textbook.

The psychological source of tonal hierarchy theory is a methodological one: the probe tone technique, introduced by Krumhansl and Roger Shepard in 1979. The purpose of this technique is to find out how the perception of individual pitches is affected by the context in which they are situated. Usually the contexts provided in experiments of this sort are very simple: subjects may hear a rising major scale, or 16 notes selected more or less at random from a given diatonic set. (Some recent work has involved more extensive contexts.) After listeners have heard these materials, they are played a final note—this is the probe tone—and asked how well this final note follows from the others "in a musical sense" (Krumhansl, 1983: 35). Such tests are carried out using a range of probe tones (usually all the notes of the chromatic set). The results are analyzed statistically, and this yields a profile showing how well each of the different probe tones has been judged to fit in that particular context; this is what Krumhansl calls a tonal hierarchy. For instance, you find that the tonic and fifth are judged to fit better than other notes in a diatonic major context. Moreover, you can compare the profiles of different diatonic sets; it turns out that, as you might expect, C major and G major are much more similar than C and F-sharp major. Krumhansl and her co-workers take this to be a measure of the distance between different keys, and if you express the results of such comparisons topologically, you end up with a chart that is essentially identical to Schoenberg's "chart of the regions."[11] So these experiments may be said to demonstrate the psychological reality of Schoenberg's topological representation of key relationships.

Krumhansl's work has come in for a good deal of criticism, some of it

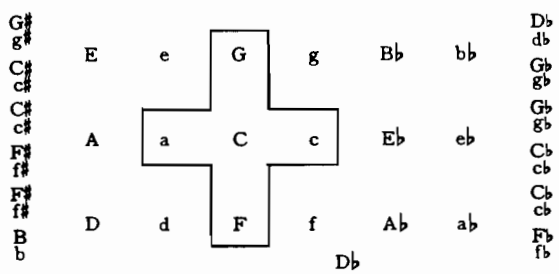



FIGURE 3.3. "Chart of the regions" from Schoenberg, 1969.

acrimonious. One of the main complaints is the one I leveled at Wolpert's experiments: the task required of the subjects is too vague. Just what does it mean to ask how well the probe tone fits "in a musical sense"? In *what* musical sense? One of Krumhansl's experiments highlights the complexities of the response situation: in probe tone tests based on serial contexts, listeners inexperienced in serial music produced profiles similar to those found in tonal music, whereas experienced listeners produced exactly the opposite profiles. Krumhansl and her associates suggest that this was because the experienced listeners adopted "a strategy, whether conscious or not, of relating the serial contexts to tonal hierarchies of major and minor keys, and simply reversing the ordering of the ratings" (Krumhansl, Sandell, & Sergeant, 1987: 73–74). In this case, since the responses were fairly consistent within each group of subjects, it can't have been too hard to see that they were adopting different strategies. In other cases, however, the divergences between response strategies might be much harder to spot. David Butler, who has been foremost among Krumhansl's critics, writes of a particular experiment that "the rating of any probe tone following the contextual series of chords might have signaled one listener's judgment of melodic continuation at the same time that it recorded another listener's judgment that the probe tone represented a well- or ill-fitting extrapolation of the roots of the chords in the contextual pattern" (1990: 329). Adding such divergent ratings together would make little sense, because the listeners are not engaged in the same task. Each is judging how well the probe tone fits, but in a *different* musical sense.[12]

Krumhansl responds to such objections by pointing out that there is considerable consistency in the results obtained by means of the probe tone technique; "Different listeners exhibit similar patterns of responses, and the same individual tends to produce the same or nearly the same response on different occasions when the identical musical stimulus is presented" (1990a: 11). Butler, however, has another explanation for this consistency: the test materials in these experiments have been consistently weighted toward the tones that end up with the highest ratings in the profiles. In other words, listeners have generally said that the tonic and the fifth fit best in a major-scale context because there have generally been more tonics and fifths in the test materials than other scale degrees; the resulting profile is "a stimulus artifact . . . rather than a reflection of mental representations of pitch relations" (1989: 230). Now this is at first sight a very damaging claim. Krumhansl, however, agrees that her tests do indeed show that "humans . . . are highly sensitive to information about frequency of occurrence" (1990b: 315). And she refers to comparative tests of Western and Indian listeners who heard materials taken from North Indian ragas. Although the Western listeners had no experience of Indian music, they produced profiles very similar to those of the Indian listeners, all of whom were trained in Indian music. Krumhansl's (1990a: 262–63) explanation of this result is that the Western listeners must have registered

> the relative salience or emphasis given the various tones in the musical contexts. In this case, this strategy produced style-appropriate tonal hierarchies. Thus, it could be used to achieve a sense of tonal organization that, when refined and solidified through experience, would serve as the basis for appreciating and remembering more complex and subtle structural features. For the Indian listeners, the high correlations between the probe tone ratings and the tone durations reflect the convergence between the surface emphasis given the various tones and the underlying tonal organization of the rāgs.

And in her response to Butler, she cites some statistical measures of the relative frequency of occurrence of different scale degrees in different works by composers like Mozart, Schubert, and Schumann, showing how these note frequencies correspond closely to the profiles for major and minor keys yielded by probe tone studies (1990b: 316, also in 1990a: 68).

By this stage in the argument it is becoming less and less clear what the probe tone studies are actually telling us. How can the same results for Western and Indian listeners in Krumhansl's experiment tell us different things in each case—that Western listeners are responding to surface features, while Indian listeners are responding to the convergence between surface features and underlying tonal organization? If perceived tonal hierarchies are a function of the frequency of occurrence of different notes, chords, or keys, then won't any probe tone experiment yield profiles that simply reflect the test materials and so cannot be generalized to any other musical context? But it isn't this simple; sometimes the attributes of the test materials and the results of probe tone experiments don't go hand in hand. For instance, tests using Balinese materials yielded different results for Western and Balinese listeners (Krumhansl, 1990a: 266–68). Krumhansl explains this in terms of the different knowledge of musical style with which the two groups of listeners approached the task. Her basic assumption is that probe tone tests reveal stable, long-term knowledge that listeners use in processing musical sounds. In an overview of her work written in 1983, she described it as "an initial attempt to understand the kinds of internal representations and processes involved in listening to tonal music. The results provide a quantitative, structural description of the kinds of knowledge listeners have about fundamental aspects of tonal organization" (1983: 59). And she referred to the tonal hierarchy as "the fundamental underlying feature of tonal music, one that is generative of interkey relations" (p. 56). The word "generative" says it all; tonal hierarchy theory was to provide something like a competence theory for tonal music. Krumhansl began her 1983 overview with a reference to Chomsky.

Since then, however, it has become increasingly evident that the static nature of tonal hierarchy theory drastically limits its explanatory scope. In Butler's words, "it is necessarily based on a rigid, time-independent conception of musical consonance in which the tonic is always considered most consonant and most stable, the dominant is next-most consonant and next-most stable, and so on" (1989: 223). But key identification, to take one of the things Krumhansl has tried

to explain in terms of tonal hierarchies, is clearly time-dependent. Whether a given passage projects a specific tonic, say C, does not simply depend on the notes in the passage: it depends on their ordering (which is to say, on their harmonic implications, contour, and so forth). A recent experimental study of key identification grew out of "a class in the psychology of music in which the . . . author had programmed a computer to sound out the notes of a diatonic scale in random succession and, after a minute or so of this continuous stream of notes, asked the class of musically highly educated listeners to hum what they thought was the major-mode tonic. To his (and the class's) surprise, the result was a cacophony" (West and Fryer, 1990: 254). The ensuing experiment, based on a version of the probe tone method in which subjects were asked "to judge the suitability of the probe tone *as a tonic*" (p. 255, my italics), confirmed the impression: responses depend on the order in which the notes are presented. This means that tonal hierarchy theory can capture, at best, only part of the phenomenon of key identification.

In her 1990 response to Butler (who for years has been arguing that key identification is time-dependent), Krumhansl takes an accommodating line over the question of key identification, suggesting that tonal hierarchies constitute just one of several relevant factors (1990b: 317). And her characterization of the general scope of tonal hierarchy theory is also more modest than it was in 1983: "tonal contexts (even simple, schematic ones) evoke listeners' knowledge about how musical entities, such as tones, chords, and keys, are generally used in the style. These abstract pitch relations provide a background that facilitates understanding the organization of a particular piece" (pp. 313–14). As a matter of fact, Lerdahl has recently incorporated tonal hierarchy theory within *GTTM* in a way that is very much in line with this. In the original version of *GTTM*, there were informally stated "stability conditions" favoring certain structural interpretations over others. (The basic idea was that connections are more likely between elements that are tonally close to one another.) In his article "Tonal Pitch Space" (1988b), however, Lerdahl presents a more formal model of these stability conditions; this is, in essence, an amplified and generalized version of Schoenberg's "chart of the regions."[13] And he invokes Krumhansl's experiments as showing the psychological reality of the model. He doesn't express a view as to exactly what the model represents; as a theorist (rather than a psychologist) he doesn't need to. He is using tonal hierarchy theory as a rule of thumb, and you can do this without knowing whether it reflects listeners' long-term knowledge of a given style, or their responses to note distributions in a given passage, or, for that matter, some other factor such as tonal consonance.

Roy Harris says of the transformation of traditional grammar into modern linguistics that "[t]he situation in which an established descriptive format devised for one particular purpose is taken over and adapted to serve some new and quite different purpose is a situation fraught with potential errors and inconsistencies of all kinds" (1981: 54). That a similar situation obtains in the study of music is

clear from statements like "the existing body of music theory is essentially a non-formalised cognitive theory" (Marsden and Pople, 1989: 30), or "Rameau's *basse fondamentale* or Schenker's *Ursatz*, and more generally, 'music analysis' of almost any familiar kind, may be convincingly construed as attempted descriptions of underlying mental realities" (Agmon, 1990: 285). The attempt to transform music theory into a theory of mind accounts for some of the most basic characteristics of current music psychology: not just the emphasis on notational categories that I mentioned earlier, but the whole guiding principle that there are specialized psychological mechanisms which process different parameters of the music more or less independently of one another. As Eric Clarke notes (1989a: 4), this approach is derived from music theory; psychologists have taken theoretical constructs and attempted to give them a psychological interpretation. Tonal hierarchy theory is an example of just this process, and the acrimonious debate over it exemplifies the difficulties of which Harris speaks.

Nothing in the practice of Western tonal music suggests that unordered collections of notes, like the C major diatonic set, function in their own right as significant musical structures. It is *ordered* structures that are significant: harmonic progressions, melodic patterns. The diatonic set is simply a theoretical abstraction, a generalization from specific harmonic and melodic contexts. It is a pedagogical construct, like Schoenberg's chart of the regions. Tonal hierarchy theory, then, does not arise from the observable phenomena of music. It arises from the attempt to give a psychological interpretation to concepts drawn from music theory. In this sense, it is theory-driven rather than data-driven. Now in saying this I am going against Krumhansl's own claims regarding her topological representations of tonal structure: they are not theoretical models, she tells David Butler in reply to his criticisms, they are simply descriptions of data. Indeed, she appears to question whether there is such a thing as tonal hierarchy "theory" at all (1990b: 309–10). But there is something disingenuous about this argument. In the first place, what Krumhansl presents as (say) a major-key profile is an extrapolation from test materials in which certain notes of the C major scale appeared in a certain order; the extrapolation is made on the assumption that unordered pitch sets constitute meaningful contexts. This is a theoretical assumption. In the second place, as Eric Clarke says, "there is a strong sense in much of this work that these descriptions or formalisms are regarded as being very close to truly mental structures—not least because they are presented as contributions to the cognitive psychology of music" (1989a: 6). The title of Krumhansl's recent book-length exposition of her work is *Cognitive Foundations of Musical Pitch*.

Why did the debate between Krumhansl and her critics become so acrimonious? The basic reason seems to be that the protagonists could not agree what they were talking about. They could not agree what tonal hierarchy theory was a theory of. They could not agree whether it was a theory at all. The story of tonal hierarchy theory is the story of a description in search of a meaning.

## Should Theory Reflect or Challenge Musical Perceptions?

Jay Dowling writes that "psychologists like myself often assume that cognitive processing (whether consciously explicit or not) proceeds in an analytic mode—that like Schumann's 'ideal listener' the brain reconstructs the score as it goes along" (1989: 250). And he adds: "This assumption feeds a theoretic position built on the experimental analysis of the brain's feature analysis." When he says "theoretic," Dowling is referring to psychological theory. But the same assumption feeds—and is fed by—music theory, which has always had a basically scriptist orientation. The result of the psychologization of music theory (I am tempted to write, the music theorization of psychology) is that it has come to seem natural for both theorists and psychologists to explain music in terms of formal principles or rules—if not actually the rules offered by current music theory, then future rules, better rules, that will be shown to be psychologically valid. Lerdahl and Jackendoff's work obviously represents a giant step in this direction and holds out the promise of a single theory that will be at the same time musical and psychological. What I want to suggest in this final section is that such total convergence between music theory and psychology may not be desirable, or even feasible, because the two disciplines have essentially different and mutually incompatible aims.

Lerdahl and Jackendoff's theory of music is both music theory and psychology because it assumes, not just that there are rules governing both composing and listening to music, but that they are essentially the same rules. Or rather, it assumes that they *ought* to be essentially the same rules, and that to the extent that they are not, there is a failure in the communicative process. Earlier I mentioned how, in his "Cognitive Constraints on Compositional Systems," Lerdahl attacks *Le Marteau sans Maître* on the grounds that there was a mismatch between its "compositional grammar" and its "listening grammar." He goes on to prescribe a number of conditions that must be fulfilled if this kind of mismatch is to be avoided. And he puts forward two general aesthetic laws: "The best music arises from an alliance of a compositional grammar with the listening grammar," and "[t]he best music utilizes the full potential of our cognitive resources" (1988a: 255–56). If the first of these laws excludes *Le Marteau* and serial music in general, the second has still more drastic effects: Ligeti's *Atmosphères* drops out because it blurs distinction between events (p. 239), Balinese gamelan music "falls short with respect to its primitive pitch space" (p. 256), and "[r]ock music fails on grounds of insufficient complexity." (Rock music attracts massive audiences, plays a major role in the personal and social development of young people, and makes fortunes, but it fails all the same.) Such conclusions will strike many people as astonishing; however, this doesn't invalidate Lerdahl's starting point, which is that a composer can't just stuff structure into a piece of music and expect listeners to be able to reconstruct it

simply because it is there. There are constraints on the kind of structures that listeners can reconstruct.

But this brings us back to the question that Dowling raised: how far *are* people engaged in reconstructing structure when they listen to music? Dowling continues, "We should not forget that the brain and mind is not always so precise and analytic, and that cognitive processing can operate effectively on a more global level." In other words: even if it is possible for people to listen analytically, they may choose to listen holistically instead. There is plenty of evidence that they do just this. If Lerdahl's account of listening grammar is correct, then the subjects in Clarke and Krumhansl's experiments could have listened to the Mozart Fantasie grammatically—that is, could have reconstructed the structure—in a manner impossible in the case of *Klavierstück IX;* but they apparently listened to both pieces in much the same way. (Similar, though less formal, experiments of my own produced comparable results.)[14] Again, Alan Smith played incomplete movements from Haydn symphonies to a group of music students, and asked them to say at what point in the form the music had stopped; in the first test "[t]he results were not better than might have been expected had the subjects all guessed. However, during two sequent selections the subjects, who now knew what was expected of them, performed well" (1973: 212–13). In other words, the students could perfectly well have listened grammatically the first time, but they *chose* not to. Introspective and anecdotal evidence suggests that this is how most people listen most of the time.[15]

The idea that music is a process of communication in which listeners decode structures that composers have encoded—the idea we saw at its most basic in Wolpert's article—is, then, based on several disputable assumptions: that people choose to listen grammatically; that there is, or ought to be, an equivalence between compositional and listening grammars; and, most fundamentally, that there is such a thing as musical grammar. It seems to me that a strong argument can be made that, if music is like language at all, then it is most like everything in language *except* grammar. In a book entitled *The Violence of Language*, Jean-Jacques Lecercle writes, "Words speak. They alliterate . . . , they proliferate in the inventories of subjects and passions, . . . they quibble . . . , and of course they endlessly pun. Homophony is the law of the remainder" (1990: 98). By "the remainder," Lecercle means the power of words to create signification of their own, to subvert intended meanings, to do violence to their users. He means the "parts of language that no grammar can ever reach" (p. 18). Lecercle, like Harris, does not believe in formal grammars. In fact, he does not believe in grammars at all. Or rather, he believes that they consist of "largely arbitrary and negative frontiers" (p. 52); he sees them as interventions in the world, historical institutions, part of the politics of language. Now Lecercle's analysis is at least equally applicable to music; what is music but the speaking, alliteration, proliferation, quibbling, and endless punning of sounds? What is its basic law, if not homophony? And it seems clear enough that the rules of music theory serve to codify,

legitimize, and institutionalize musical styles. They explain music, to be sure, but not in the sense that formal grammars explain language; more in the sense that books of etiquette explain how to behave in public.

I can clarify this by going back to the issue of tonal closure. As I said, theorists explain forms like sonata in terms of tonal excursions creating and resolving tension. But we have seen that listeners do not, in practice, respond to large-scale tonal structure. The theory of tonal closure, then, fails to predict listener responses. It is bad psychology. But then, it was never meant to be psychology. It was meant to be a way of thinking about large-scale musical structure. For the tonal plan of a sonata movement governs the disposition of themes and textures, the patterning of loud and soft and high and low, the pacing of climax and relaxation—all aspects of music that *do* directly bear on the listener's experience (Rosner & Meyer, 1986). The idea of tonal closure allows us to recapture something of the way in which composers of the eighteenth and nineteenth centuries conceived their music. Schenker's reductive graphs, or Lerdahl and Jackendoff's, show how these composers adopted the same principles in shaping the small-scale and the large-scale structure of their music. They built sections and even entire movements on the model of a single phrase, articulating each through the motion of tonic to dominant and back to tonic. They created a hierarchy of closures ranging from the perceptible to the imperceptible. And because of this, it does not really make sense to look upon hierarchical theories of music as theories of perception. At their higher levels, at least, theories such as Schenker's or Lerdahl and Jackendoff's tell us far more about what composers of tonal music did than about what listeners to their music do.

A psychological theory of music is successful to the extent that it reflects perception. Musical perception is pluralistic and fluid; listeners make use of multiple cognitive frameworks (Marsden & Pople, 1989) and shift their strategies from one moment to the next (Hantz, 1984: 255). A psychological model of perception needs to embody these characteristics. But there is no reason why music theory should be like this. When a theorist analyzes a piece of music, she is taking up an interpretive stance in relation to it. She is following through the implications of a particular way of thinking about it. She is aiming, perhaps, to change the way people experience the music. From this point of view, the psychologist's perceptual model is totally inappropriate because, as David Osmond-Smith puts it, "the resultant theoretical construct will be devoid of critical content: it cannot challenge our pre-formed responses because it sets out to reproduce them" (1989: 92). The aim of music theory, then, is to go beyond perception. In fact, you might put the point polemically and say that, from the music-theoretical point of view, an analysis can have value only to the extent that it *deviates* from perception. As an illustration of this, consider Schenker's analysis of the opening theme from Beethoven's Sonata, op. 90 (Figures 3.4, 3.5) Schenker suppresses the most perceptually salient feature of Beethoven's music—the pause on the dominant and registral displacement at bar 16. He

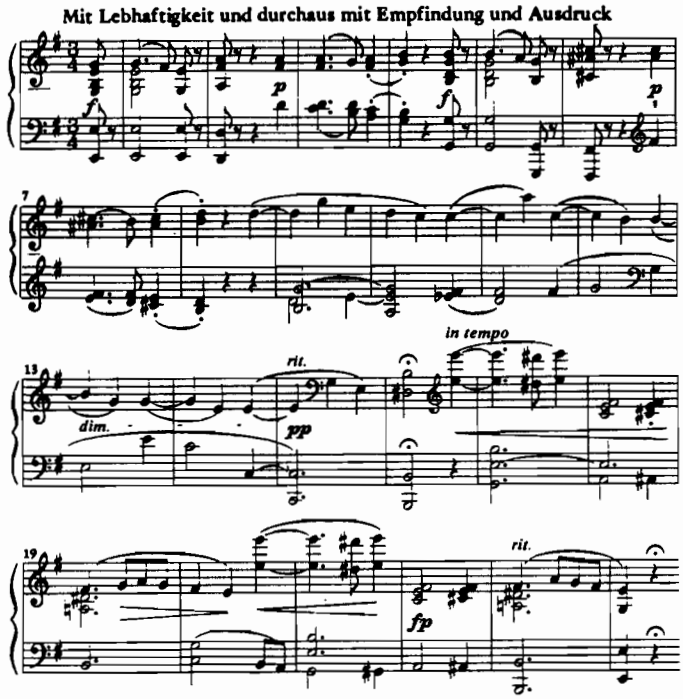



FIGURE 3.4. Beethoven, Sonata, op. 90, first movement, bars 1–24.

reads a tonic prolongation right across this passage. Critics who accuse Schenker of reducing the classics to abstract formulae could have a field day here! But such accusations are the result of looking *at* the analytical graph, instead of looking *through* it to the music. Schenker's graph expresses a stylistic norm, against which Beethoven's music stands out in all its singularity. It casts an oblique light on the music, thus throwing its characteristic features into relief. The content of the analysis, then, lies in the divergence between the graph and the music listeners hear.[16]

Another illustration of this is Charles Rosen's analysis of Haydn's String Quartet in B minor, op. 33, no. 1. This is focused round the interaction of B minor and D major. The tonal ambiguities of the opening (Figure 3.6), leading to the trenchant A-natural/A-sharp clashes of bars 3–4, create a tension that Rosen sees as generating much of the work's large-scale structure. The rest of the movement resolves the dissonant D major, he says, "first by treating it as the dominant or secondary key . . . , and then by expanding this in the 'development' and resolving everything in a recapitulation which—except for two measures at the opening—insists dramatically on the tonic" (1976: 120). But do listeners hear all this? Well, maybe they *do* after they have read Rosen's analysis, which succeeds wonderfully in sharpening perceptions. (It's an obvious fact,

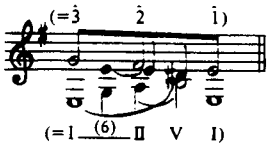



FIGURE 3.5. Schenker's reduction of Beethoven's Sonata, op. 90, first movement, bars 1–24.

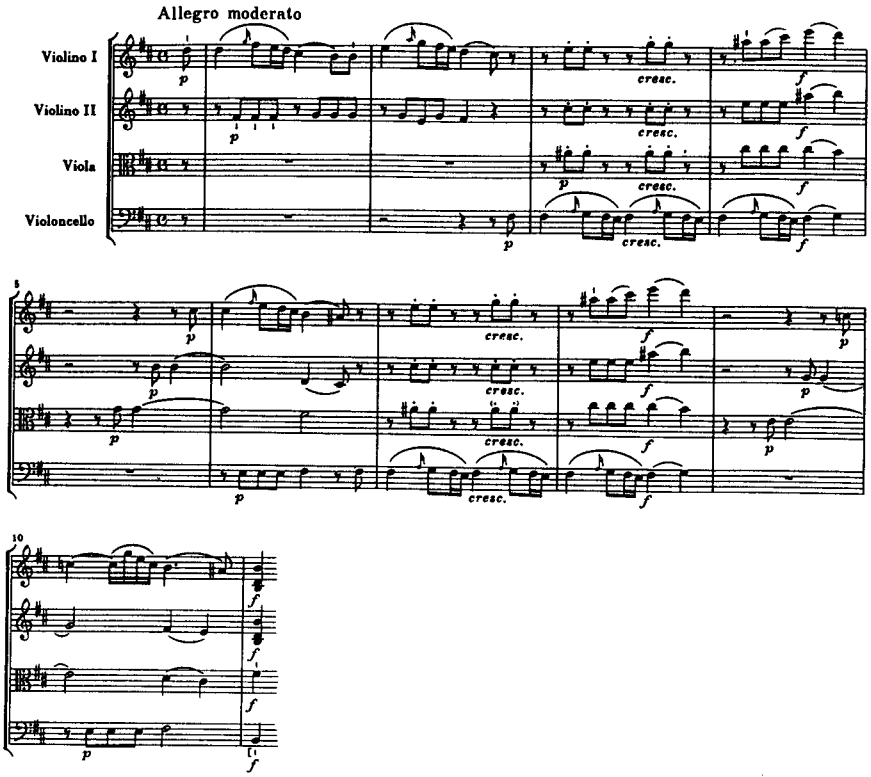



FIGURE 3.6. Haydn, Quartet, op. 33, no. 1, movement 1, bars 1–10.

though not often mentioned in the psychological literature, that people hear music differently after they have studied the score.) And maybe Rosen's analysis recaptures something of the refined sensibility of the aristocratic connoisseurs and dilettanti for whom Haydn wrote his quartets. Or maybe it doesn't; there is no way of knowing for sure. But that does not really matter. If we accept Rosen's analysis, we do not do so because we believe it to be in some absolute sense correct. We do so because we find it persuasive. We do so because it contributes to our experience of the music. The value of the analysis, to repeat what I said earlier, is that it enables us to respond to the excess of music over sound.

Only a few years ago, music theorists with an interest in psychology were inclined to complain that psychologists ought to stop reinventing the wheel and learn some theory.[17] Things have changed; psychologists know a lot of music theory nowadays. The problem, as I have argued, is that they have tended to assimilate theory, treating it as a kind of incompletely formed psychology, rather than seeing it as a related but distinct discipline with its own aims and its own methods. Of course we do not want to go back to the bad old days, when psychologists knew little or nothing about theory and theorists knew little or nothing about psychology. But what Clarke (1989a: 2) calls the "rather unstructured 'leakage' between the disciplines" is potentially damaging to both of them. Music psychology risks losing the ability to pose fundamental questions through too ready an acceptance of music-theoretical concepts and categories; it becomes unduly influenced by music theorists' preoccupation with issues of formal structure. And conversely, music theory risks losing its identity in a welter of incompatible claims regarding both procedures and objectives. What is needed is the careful translation—or maybe it would be more accurate to say the careful *mistranslation*—of concepts from one discipline to the other.

## NOTES

1. Clap your hands regularly once a second, record the sound, play it back 440 times faster, and you have the pitch A = 440.

2. See, for example, Sloboda 1985: 151–54.

3. The instruments were, in fact, synthesized.

4. This is rather like the musicians in an experiment conducted by Lucy Pollard-Gott; they judged the similarity of different excerpts from Liszt's B minor Sonata purely in thematic terms, because theme is an essential component in the language game of traditional formal analysis. "One expert commented that after the similarity task was over, he realized that there might have been other criteria for judging similarity such as relative loudness. Yet when he heard the passages, their thematic relations were compelling and seemed to be the only natural basis for comparing them" (1983: 81).

5. Often abbreviated *GTTM*. The first comprehensive statement of the theory was in the book of that name (1983); subsequent extensions to it have appeared in Lerdahl's articles on timbral hierarchies (1987), tonal pitch space (1988b), and atonal prolongational structure (1989). An outline of *GTTM* will be found in Chapter 2.

6. Also published, minus one piece, as *Vier kleine Klavierstücke*. Examples 1 and 2 are taken from Cook, 1990, where a more detailed discussion of these issues may be found.

7. The tests are fully reported in Cook, 1987. Subjects evaluated the test pieces along four dimensions: coherence, completion, pleasure, and expressiveness. There was, in general, a high level of redundancy between these measures. Bengt Edlund, of the University of Lund, reports similar results from a series of tests carried out in 1991 (unpublished paper, "Tonics and Returns").

8. This was in fact only a fragment: the theme from Brahms's *Haydn Variations*. All other pieces in the tests were complete movements.

9. This is an adaptation of G. Lakoff and H. Thompson's description of language, cited in Hantz, 1984: 247.

10. Short for pitch class set theory, a music-analytical system first described in Forte, 1973.

11. See Figures 2.8 and 7.4 in Krumhansl, 1990.

12. At least the subjects in Krumhansl's experiments probably knew what they were doing, which seems unlikely to have been the case in a cross-cultural study of tone hierarchies carried out by Kessler, Hansen, and Shepard (1984). In what must rank as one of the most farcical episodes in the history of music psychology, the team from Stanford jumped on their motorbikes and set out to find the remotest village in Bali. Eventually they discovered a village whose inhabitants had never heard the Indonesian national anthem, and who regarded the Americans as "Javanese" on the grounds that they obviously were not from Bali. The probe tone tests presented some difficulties, and the rating scale had to be reduced from 7 to 5 so that the villages could respond by holding up their fingers. When processed back at Stanford, the results indicated that the villagers were equally unfamiliar with the tonal hierarchies of Western and Balinese music. Another interpretation might be that the villagers were completely baffled by the whole undertaking and simply gave the first answers that came into their heads in order not to disappoint their visitors. What Kippen (1987: 175) refers to as the "cultural insensitivity" of the whole episode underscores the difficulty of adapting the procedures of experimental psychology to non-Western or non-Westernized cultures.

13. The chart now appears in yet another form: see Lerdahl's Figure 17.

14. See Cook, 1990: 44–45, 56–58.

15. For a particularly telling introspective account of musical listening, see Osmond-Smith, 1989: 95. An extended treatment of this whole topic may be found in Cook, 1990.

16. A fuller version of this argument may be found in Cook, 1989: 131–34. The graph is taken from Schenker, trans. 1979, Example 109a (1).

17. This is the general tenor of, for instance, Walsh, 1984.

## REFERENCES

Agmon, E. (1990). Music theory as cognitive science: Some conceptual and methodological issues. *Music Perception*, 7, 285–308.

Babbitt, M. (1962). Twelve-tone invariants as compositional determinants. In P. H. Lang (Ed.), *Problems of modern music* (pp. 108–21), New York.

Butler, D. (1989). Describing the perception of tonality in music: A critique of the tonal hierarchy theory and a proposal for a theory of intervallic rivalry. *Music Perception*, 6, 219–41.

Butler, D. (1990). Response to Carol Krumhansl. *Music Perception*, 7, 325–38.
Clarke, D. (1989). Structural, cognitive and semiotic aspects of the musical present. *Contemporary Music Review*, 3, 111–31.
Clarke, E. (1989a). Mind the gap: Formal structures and psychological processes in music. *Contemporary Music Review*, 3, 1–13.
Clarke, E. (1989b). Issues in language and music. *Contemporary Music Review*, 4, 9–22.
Clarke, E., & Krumhansl, C. (1990). Perceiving musical time. *Music Perception*, 7, 213–51.
Cook, N. (1987). The perception of large-scale tonal closure. *Music Perception*, 5, 197–205.
Cook, N. (1989). Music theory and "good comparison": A Viennese perspective. *Journal of Music Theory*, 33, 117–41.
Cook, N. (1990). *Music, Imagination, and Culture*. Oxford.
Deliège, I. (1987). Grouping conditions in listening to music: An approach to Lerdahl and Jackendoff's grouping preference rules. *Music Perception*, 4, 325–59.
Dowling, W. J. (1989). Simplicity and complexity in music and cognition. *Contemporary Music Review*, 4, 247–53.
Forte, A. (1973). *The structure of atonal music*. New Haven, CT.
Gjerdingen, R. (1988). Shape and motion in the microstructure of song. *Music Perception*, 6, 35–64.
Gotlieb, H., & Konečni, V. (1985). The effects of instrumentation, playing style and structure in the Goldberg Variations by Johann Sebastian Bach. *Music Perception*, 3, 87–101.
Hantz, E. (1984). Studies in musical cognition: Comments from a music theorist. *Music Perception*, 2, 245–64.
Harris, R. (1980). *The language-makers*. Ithaca, NY.
Harris, R. (1981). *The language myth*. London.
Kessler, E., Hansen, C., & Shepard, R. (1984). Tonal schemata in the perception of music in Bali and the West. *Music Perception*, 2, 131–65.
Kippen, J. (1987). An ethnomusicological approach to the analysis of musical cognition. *Music Perception*, 5, 173–95.
Konečni, V. (1984). Elusive effects of artists' "messages." In W. R. Crozier & A. J. Chapman (Eds.), *Cognitive processes in the perception of art* (pp. 71–93), North-Holland.
Krumhansl, C. (1983). Perceptual structures for tonal music. *Music Perception*, 1, 28–62.
Krumhansl, C. (1990a). *Cognitive foundations of musical pitch*. New York.
Krumhansl, C. (1990b). Tonal hierarchies and rare intervals in music cognition. *Music Perception*, 7, 309–24.
Krumhansl, C., Sandell, G., & Sergeant, D. (1987). The perception of tone hierarchies and mirror forms in twelve-tone serial music. *Music Perception*, 5, 31–77.
Lecercle, J.-J. (1990). *The violence of language*. London.
Leech, G. (1983). *Principles of pragmatics*. London.
Lerdahl, F. (1987). Timbral hierarchies. *Contemporary Music Review*, 2, 135–60.
Lerdahl, F. (1988a). Cognitive constraints on compositional systems. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 231–59), Oxford.

Lerdahl, F. (1988b). Tonal pitch space. *Music Perception, 5*, 315–49.
Lerdahl, F. (1989). Atonal prolongational structure. *Contemporary Music Review, 4*, 65–87.
Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA.
Lerdahl, F., & Jackendoff, R. (1984). An overview of hierarchical structure in music. *Music Perception, 1*, 229–52.
Marsden, A. (1989). Listening as discovery learning. *Contemporary Music Review, 4*, 327–40.
Marsden, A., & Pople, A. (1989). Modelling musical cognition as a community of experts. *Contemporary Music Review, 3*, 29–42.
Narmour, E. (1989). The future of theory. *Indiana Theory Review, 10*, 95–97.
Osmond-Smith, D. (1989). Between music and language: A view from the bridge. *Contemporary Music Review, 4*, 89–96.
Pollard-Gott, L. (1983). Emergence of thematic concepts in repeated listening to music. *Cognitive Psychology, 15*, 66–94.
Rosen, C. (1976). *The classical style* (rev. ed.). London.
Rosner, B., & Meyer, L. (1986). The perceptual roles of melodic process, contour, and form. *Music Perception, 4*, 1–39.
Schenker, H. (trans. 1979). *Free composition*. New York.
Schoenberg, A. (1969). *Structural functions of harmony* (2nd ed., rev. by L. Stein). London.
Seifert, U. (1991). Competence and performance in cognitive science: On the relation between music-theoretical research and the modelling of musical cognition. Abstract of paper scheduled for presentation at the Computers in music research conference, Queen's University of Belfast, April 7–10, 1991.
Serafine, M. L., Glassman, N., & Overbeeke, C. (1989). The cognitive reality of hierarchic structure in music. *Music Perception, 6*, 397–430.
Sloboda, J. (1985). *The musical mind: The cognitive psychology of music*. Oxford.
Smith, A. (1973). Feasibility of tracking musical form as a cognitive listening objective. *Journal of Research in Music Education, 21*, 200–13.
Stockhausen, K. (1959). How time passes . . . . *Die Reihe, 3*, 10–41.
Walsh, S. (1984). Musical analysis: Hearing is believing? *Music Perception, 2*, 237–44.
West, R., & Fryer, R. (1990). Ratings of suitability of probe tones as tonics after random orderings of notes of the diatonic scale. *Music Perception, 7*, 253–58.
Wolpert, R. (1990). Recognition of melody, harmonic accompaniment, and instrumentation: Musicians vs. nonmusicians. *Music Perception, 8*, 95–105.

# Developmental Perspectives

# 4

# Songsinging by Young and Old: A Developmental Approach to Music

LYLE DAVIDSON

## Introduction

There have been many investigations of the musical development and the musical abilities of children in recent years (see overviews by Hargreaves, 1986a; Shuter-Dyson & Gabriel, 1981). Research ought to benefit educational practice, but there have been difficulties in establishing an effective dialogue between the scientific and educational communities (Hargreaves, 1986b; M.E.N.C., 1981; Sloboda, 1986). In Chapter 4 and 5, Lyle Davidson and Jeanne Bamberger discuss issues related to musical development and education. Both writers stress the importance of integrating experimental results in pedagogical practice.

Scientific research in music development has examined the abilities of children from a very young age. Chang and Trehub (1977) found that five-month-old infants are already sensitive to the sequential structure of a melody. The investigations presented a series of short melodies to the babies and, by monitoring their heart rate, noticed that when a new melody was presented some of the infants showed changes in their heart rate, a reliable indicator of perceived novelty. Studies in the perception of tonality have also investigated the young child's sensitivity to tonality. Zenatti (1969) asked five-, six-, and seven-year-old children to listen to a series of three-note tonal or atonal melodies. Each melody was followed by a repetition in which one of the notes had been altered in pitch. While five-year-olds performed equally poorly on both types of melodies, children six and seven years of age performed better on the tonal melodies than on the atonal ones, thus showing that they had extracted some structural information that helped them in comparing the melodies.

The psychologist Jean Piaget (1960, 1967) believed that to study the

intellectual growth of the child we must look at growth over a period of years, and his work has led to a developmental model for researching musical growth longitudinally (Serafine, 1980). The research headed by Howard Gardner at Harvard University's Project Zero, has investigated the musical development of children, providing a naturalistic and detailed record of the development of their song output over several years. Gardner (1973) has proposed that the artistic development of children can be traced to their acquisition of symbols and their sensitivity to artistic styles.

Lyle Davidson, a senior investigator on the Project Zero research team, has extensively researched musical development, focusing on songs (Davidson, 1985; Davidson & Scripp, 1988a, 1988b). In Chapter 4 Davidson discusses the development of tonal knowledge, looking at the preschoolers' earliest songs, the singing of musically untrained adults, and the perceptual changes in tonal relationships that a student undergoes in music school. Davidson notes that we do not have a comprehensive view of the goal of musical development and instruction because we fail to learn from the uncoordinated and loosely connected chunks of knowledge that we have. He encourages educators and professional musicians to consider the psychological nature of musical knowledge. Davidson describes the musical growth of the child by showing how the child's early approximations of a song later acquire a definite melodic shape and a stable contour, and he illustrates this process with drawings of the child's musical notation (see also Chapter 7 by Jay Dowling in this volume, and Dowling, 1984, for additional research on the development of children's songs).

## REFERENCES

Chang, H., & Trehub, S. E. (1977). Auditory processing of relational information by young infants. *Journal of Experimental Child Psychology, 24,* 324–31.

Davidson, L. (1985). Tonal structures of children's early songs. *Music Perception, 2*(3), 361–74.

Davidson, L., & Scripp, L. (1988a). Young children's musical representations: Windows on music cognition. In J. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition.* Oxford: Clarendon Press.

Davidson, L., & Scripp, L. (1988b). A developmental view of sight-singing: The internalization of tonal and temporal space. *Journal of Music Theory Pedagogy, 2*(1), 10–23.

Dowling, W. J. (1984). Development of musical schemata in children's spontaneous singing. In W. R. Crozier & A. J. Chapman (Eds.), *Cognitive processes in the perception of art.* Amsterdam: Elsevier.

Gardner, H. (1973). *The arts and human development.* New York: Wiley.

Hargreaves, D. J. (1986a). *The developmental psychology of music.* Cambridge: Cambridge University Press.

Hargreaves, D. J. (1986b). Developmental psychology and music education. *Psychology of Music, 14,* 83–96.
M.E.N.C. (1981). *Documentary report of the Ann Arbor symposium.* Reston, VA: M.E.N.C.
Piaget, J. (1960). *The psychology of intelligence.* Totowa, NJ: Littlefield, Adams.
Piaget, J. (1967). *The child's conception of space.* New York: Norton.
Serafine, M. L. (1980). Piagetian research in music. *Council for Research in Music Education Bulletin, 62,* 1–21.
Shuter-Dyson, R., & Gabriel, C. (1981). *The psychology of musical ability* (2nd ed.). London: Methuen.
Sloboda, J. A. (1986). Achieving our aims in music education research. *Psychology of Music, 14*(2), 144–45.
Zenatti, A. (1969). *Le développement génétique de la perception musicale.* Monographies Françaises de psychologie No. 17.

---

In this chapter, I argue that taking a developmental perspective reveals a great deal about the nature of musical ability and the nature of effective music instruction. Based on methods of developmental psychology, the development of tonal knowledge is traced by considering preschoolers' earliest songs, the songsinging of musically untrained adults, and finally, the transformation in a student's grasp of tonal relationships in performance, which result from the intensive training that takes place in music schools. Contour schemes—the organizing mental structures of tonal materials in music typical of young children—continue to function beyond the years of childhood, unless there is considerable broad musical training like that typically found in music schools.

## Musical Talent: A Determined Ability or a Propensity for Development?

A deep-seated assumption about the nature of musical ability is the belief that musical ability is largely (if not wholly) innate and fixed. From this point of view (I will call it the "talent model"), development is assumed to occur very nearly all by itself—in the person endowed with talent. Environment, training, education, hard work: these are assumed to play little role in musical ability and knowledge. At least two difficulties follow from this assumption: only the gifted and talented are seen as worth the teacher's investment of time and energy, and the scope of

the teacher's and the student's activity is reduced. Let's consider each of these issues in turn.

First, the talent model suggests that fine teachers, good instruments, and time devoted to mastering even elementary musical skills are wasted on those not obviously talented. This line of thinking is reinforced by the assumption that an individual's natural gifts should be allowed to develop freely. Unfortunately, a model of education based on the romantic concept of the flowering of talent ultimately fails to recognize and develop the capacities for musical expression and thought in those students who are not identified as gifted.

Research suggests that *the musically untrained have surprisingly sophisticated abilities.* For example, the average child and adult, although unable to read music notation and without any performance training, can represent the musical dimensions of phrase structure, melodic contour, pitch, and rhythmic elements surprisingly well (Davidson & Scripp, 1988a; Davidson, Scripp, & Welsh, 1988) (Figure 4.1). Furthermore, research suggests that the average adult, although unable to read music notation and without any performance training, can solve musical problems *at least* at the level of undergraduate music students who have years of performance experience in their background (Scripp, Meyaard, & Davidson, 1988). *By investigating the development of musical ability,* it is possible to see degrees of difference and similarity among various groups, children and adults. From this more articulated perspective, it is possible to make better arguments in support of teaching music to the general student, and better strategies for teaching students at all levels of education (Davidson & Scripp, 1988a).

Another difficulty is the assumption that the talented student basically learns by herself. This reduces the scope of the teacher's and the student's interaction. There is little incentive to probe deeply into the nature of musical knowledge and understanding. The narrow version of the talent model assumes that instrumental proficiency alone accounts for careers and alone is powerful enough to override the need for strong interpersonal skills. It reduces the opportunity, as well as the motivation, to consider other aspects of musical life and musical thinking. For the student, *the talent model isolates later development from early experience.* The continuity of growth is cut off. Early forms of musical knowing are ignored in favor of the formal, articulated knowledge of the expert. Talent is necessary but not sufficient.

The oversimplification typical of the talent model greatly misrepresents the reality of the engaging and dynamic exchange in music, just as learning in the classroom misrepresents the knowledge required beyond school (Resnick, 1981).

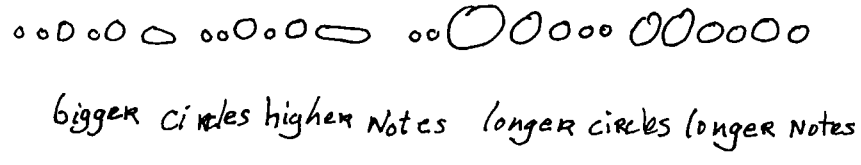



FIGURE 4.1a. Child's invented notation of "Happy Birthday to You."

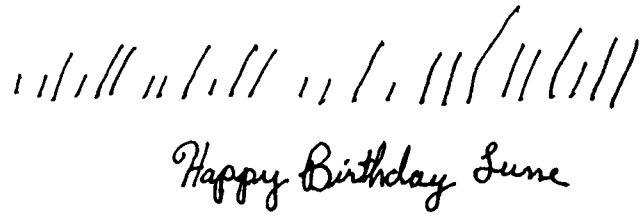



FIGURE 4.1b. Adult's invented notation of "Happy Birthday to You."

It presents a lean concept of what is involved in the development of a rich form of expression. *It results, in part, because educators and professional musicians do not consider the psychological nature of musical knowledge.*

## Development: Knowledge Under Construction

It is axiomatic in developmental psychology that our knowledge changes as it develops and matures (Wohlwill, 1973). As we acquire new degrees of mastery in a skill, our relationship to our former level of skill, knowledge, and understanding changes significantly as well. But, as musicians know if they think about it for a moment, this change is more than a simple addition to an already formed base. Its impact is deeper than adding one more element to what is already known. For example, when I learn second position on the violoncello, I do more than add another finger position to my repertoire of skills. Knowledge of the new position forces me to reinvestigate the notes I mastered in first position. My relationship to the fingerboard changes. My physical connection to the instrument is different. The implications of these changes continue to spiral beyond the moment. The change in my physical position changes my sense of what the issues are when I play a line. My choice of string on which I play a given note now reflects my understanding of the role timbre plays in my interpretation. As my options expand, the choices I make become increasingly significant, their impact increasingly wide ranging.

This dramatically transforms my perspective of the musical experience as well as the nature of performance. For example, as my ability to differentiate among possible physical movements and choices increases, my awareness of the elements of a piece also increases. When second position has become integrated into and extends the knowing represented by first position, I can begin to construct multiple views of a note, a passage, or a section. For instance, I can imagine the note A below middle C on one of two strings, and choose which one to play on the basis of the nature of the passage I am trying to play and on the basis of the musical shape I am trying to project. Finally, in this manner, my ability to consider simultaneously multiple aspects of an event, or concept, reflects considerable transformation in my mental representations of that event or concept. All of this reflects considerable developmental change.

The understanding at various points along the way to mastery is usually not noticed, except when deficient. This is not surprising, because when considering musical knowing, we normally take the most complete knowledge (the expert's) as the point of reference. When we adopt such a perspective from which to develop the musical knowledge of novices, we fail to take into account the knowledge the student brings to the task. The knowledge of an expert is very different from that of the less developed student in at least three ways: the degree of its articulateness, its stability, and the extent of its internalization (Trotter, 1986). Indeed, research shows that, contrary to what we might expect, novices and beginning musicians often confuse distinctions easily made by mature musicians (e.g., pulse and duration, contour and interval) when they make representations of each (Bamberger, 1991; Davidson & Scripp, 1988a).

We have a view of the goal of development and instruction without a corresponding view of the path that leads to the goal, because we fail to learn from the uncoordinated and loosely connected chunks of knowledge. If we don't consider the nature of the diffuse, unstable, and relatively unregulated knowledge of the musically naïve, we have little opportunity for reflecting on the knowing typical of the average nonmusician and the beginning musician as well. The study of development reveals changes in what we know, as well as in how we know it. The process of bridging the gap between the musically sophisticated and the musical novice leads to a transformation of our understanding of issues and origins, and ultimately produces new forms of knowing.

## Stages of Musical Development: The Expert's Perspective

> Domains were (previously) broken down into their principles and theories, but not studied in terms of the ways in which individuals move from encounter to engagement, from engagement to mastery of early levels, from apprentice to journeyman, and so on. Feldman, 1980: 168

Musical experts provide the models of pedagogy and set the standards for mastery. Yet unlikely as it may seem, these experts may not be the best sources of information about how knowledge of a domain develops and matures. Developmental psychologist David Feldman argues for a way to bridge the gap that separates the expert and the novice. He makes a strong case for the need to redefine a domain in terms of its psychological reality, not its logical analysis. He advocates looking at how the individual knowledge changes with experience and training, instead of defining what a person needs to know on the basis of expert *a priori* assumptions. Traditionally, developmental psychologists have been concerned with building and refining models of intellectual growth within various domains (Furth, 1969; Fischer, 1980). They seek to describe convincingly, whenever possible, invariant *and* universal stages representing changes in cogni-

tive processes and skills as they develop from novice to expert. Consider children's understanding of how to draw a person.

When children begin to develop a concept of space, what do they use as the reference point around which to organize their picture? Unlike adults, children do not always honor the horizontal baseline as an element in the organization of pictures. Instead, children first work from unit to unit, figure to figure, without an overarching reference system (Piaget & Inhelder, 1967). Thus, they first attend to and create the topological features of a drawing without reference to the whole picture. The integrating structure occurs later. Children work with the contour and more general aspects of depiction in the absence of the physical coordinates that involve the horizontal and vertical relationships which adults and experts assume to form the foundation of spatial perception and formation. So, in children's early drawing of human figures, one sees that the proximities are generally correct: arms and legs are appropriately attached to the body, the eyes are placed on the head, and they may even be placed beside one another. But the overall design is uncoordinated, proportions are unrelated to the subject as they work only from one unit to the next (Goodnow, 1977).

However, as children become able to make reference to multiple dimensions of relationships they can begin to show coordinated and geometrical relations. Thus, only when they are able to consider and honor the relationship of the horizon, the house, the angle of the roof, and the chimney in reference to one another are they able to represent the three-dimensional features necessary to make what would be an acceptable adult-style picture (Piaget & Inhelder, 1967). In music, the ability to relate individual pitches to a highly structured order (the musical scale) and the ability to coordinate the regular pulse and the varied pattern of the rhythmic surface demonstrate levels of integration necessary for an advanced knowledge of music and suggest an advanced stage of musical development which I call operational (Davidson & Scripp, 1989).

## Operational Thinking: Performance Accompanied by Awareness and Coordination

> . . . as a result of their internalized and integrated nature, concrete operations are actions accompanied by an awareness on the part of the subject of the techniques and coordinations of his own behavior. Piaget & Inhelder, 1969

Piaget's statement about the quality of actions taken in a context of awareness of techniques and coordination is a useful reference point when considering the further development of musical skills. His concept of operationality is important when considering the evolving mastery of tonal knowledge by children and adults, because it describes what takes place in the training of professional musicians. In Piaget's terms, the difference between the different grasps of tonal relationships involves contrasting aspects of thought: *figurative* and *operational*

thinking (Piaget, 1970). The figurative aspect is manifest in the initially passive aspect of knowledge as it relates to perception, imitation, and mental imagery. Mental imagery at this level is internalized imitation. For example, the finger movements that so often accompany the piano student's sightsinging, the arm movements of the trombone player, or the physical feeling of the vocal placement for the singer are examples of representation linked to an auditory image at the figurative level.

Musical training, and not only instrumental training, provides the structure for transforming the more "figurative" thinking represented by contour scheme knowledge and knowledge of tonal relationships as "tunes" into progressively operationalized thinking. This is often achieved by mixing images from different modes of representation. For example, by asking the pianist to "play" the notes while singing them with syllables or note names, the teacher is deliberately combining kinesthetic and symbolic modalities. In solfège classes, the student must initially "call up" and confirm the sounds of notes by naming them while singing melodies. Gradually, more of the sound can be imagined directly from the page, until eventually, the imagined sounds can be transformed mentally with relative ease (without externalized vocal expression). During this process, the student's knowledge of tonal space becomes increasingly articulated and internalized (Davidson & Scripp, 1988c; Scripp & Davidson, 1988; Davidson, Scripp, & Meyaard, 1988).

This provides some indication of the symptoms of what operational thinking in music might sound like. Operational thinking in music is evident in the degree to which one can make transformations and reversals of actions from one state to another (e.g., singing "Happy Birthday" in a major mode and then in a minor mode). Children and adults without musical training cannot do this simple task (except when supported by a great deal of rote imitation) because they lack the stable, yet articulated and internalized, structure on which the transformation is based. Musically untrained adults, who use either contour schemes or their memory of the specific and unique melodic shapes as whole units, can perform their songs in one way, but not both.

## Tonal Knowledge: Contour Schemes, Standard Songs, and Scales

The process of making tonal knowledge operational is as long and complex as the developmental profiles of other intelligences (Piaget, 1976). Piaget describes three distinct stages leading to the development of operational and internalized knowledge: sensorimotor, concrete operational, and formal operational (Piaget & Inhelder, 1969). In relation to tonal development, we might think of tonal knowledge as being constructed first through contour schemes, then as phrases of melodies in which the individual relationships of the pitches are not articulated,

and finally, as scales which reflect the hierarchical knowledge typical of operational thought.

Examples of the first phase clearly show the process of constructing the schemes or mental tools that precede higher levels of musical knowledge. The sound of the singing is often diffuse and unfocused. Contours of songs are typically reduced. Even if learned correctly, the range of the song usually shrinks with time to match the level of the child's current contour scheme. There is no evidence of key structure.

The second, or concrete operational, phase is different in various ways. The diffuse singing of the beginning years is replaced by discrete pitches. At this level, a singer can usually make transformations on an entire melodic set, but not on elements of the set. For example, a singer may be able to make stylistic shifts, imitate the sound of a specific performer, or sing the entire melody at a different tempo or with a different dynamic. However, other changes will be more difficult, especially changes that affect individual pitches of the melody. Singers during this stage depend on their externalized vocal production to maintain relative intonation of notes of a melody. When they imagine notes without singing them, their tonal reference slips. Their knowledge of music appears to be partitioned as performance, perceptual, or reflective ways of knowing, each one relatively independent and isolated from the other (Davidson & Torff, 1991).

We most often see the final phase of the developmental process (making tonal knowledge operational) in the context of formal music education. In music schools the instrumental lessons are linked to learning and development in several ways: through performing without the aid of one's own instrument (e.g., in the solfège class, the study of a second instrument, or chorus); through specifying one's perception (e.g., when transcribing solos, taking dictation, or making an analysis by ear), and when discussing and critiquing the music and performances of one's peers and mentors. Singers with third-stage operational knowledge do not require explicit vocal monitoring to maintain their tonal orientation or to check their notation when learning by imitation, sightsinging, or taking dictation. They can imagine the notes of a melody, transform or modify them on demand, and sing specific notes with intonation that reflects understanding of the tonal context.

To be useful, a developmental view takes into account the entire range of procedural and discursive knowledge, in a context of productions, perceptions, and reflections. These reflect the intellectual skills necessary for musical expression (Davidson & Scripp, 1992). Mapping how individuals change as they move from the "initial engagement" of their youthful years to the "early mastery" of middle teens, to become mature artists, is the business of developmental psychology. Therefore, we want to place the musical knowledge and skills of children, adolescents, and adults side by side and see what the trajectories of change really are. We ask about the antecedents of adults' musical knowledge by approaching it from a psychological as well as a musical perspective, in the hope

that with this experience we will be better able to understand and assess the skill, talent, and artistry that support mature musical expression.

Before we continue, it will be useful to review some of the arguments. First, I will argue that a logical analysis of the elements of music will not reveal the psychological knowledge on which naïve understanding is based. The psychological nature of understanding is not revealed by a logical analysis of features. It can, however, become evident by close comparison of different levels of ability. Second, a psychological analysis of musical development reveals there are plateaus of musical development. Each of the preceding levels prepares the way for the next level, but does not necessarily lead immediately to it.

## Showing What We Know: The "Felt Path" or the Formal System

The pioneering work of Jeanne Bamberger (Bamberger, 1986) provides an indication of the differences between the musically untrained and the musically trained mind. It reveals an unexpected difference between the two groups' use of the tonal reference system. Bamberger reports that untrained children and adults play melodies on Montessori bells by arranging the bells in the order of the notes of the melody as performed (a "felt path" of pitch relationships). For example, the untrained novices require one bell for each of the notes of the melody, *including the repeated tones*. Thus, for them, the first two notes of the opening of "Happy Birthday" require two different bells, one for each note (even though both pitches are the same), and the first phase, therefore, requires six bells. In the novices' understanding of melody, the notes of the tune and the order of the action of bell playing are fused together, making it unthinkable to restrike a bell in order to sound the repetition of a note.

The approach favored by the musically trained is very different. They first construct a scale of eight notes (making a formal system of pitch relationships). They then play the notes of the song, extracting them from an operational scheme of tonality. Only two notes are needed to play the first four notes of "Happy Birthday," and only four bells are required to play the first phrase. The musically sophisticated people begin the task of representing the melody with the bells by constructing an abstract scheme with which they can more efficiently represent all the notes of the melody.

The difference between these two approaches to tonal structure reflects a qualitatively different orientation to musical elements. At one level of development, the functional units of tonal space stand in such fragile relation to one another that issues like key establishment are moot. At the other level of development, the construction of the hierarchical structure of scales makes several issues relevant for consideration: the identity and stability of pitches on the one hand, and the capacity to establish and maintain keys on the other hand.

From this perspective, one index of musical sophistication is the ability to

integrate localized pitch knowledge into increasingly less contextualized tonal structures as represented by scales and keys. That is, the untrained subjects can be expected to work only within the context of the sequence of the notes of the tune, whereas the trained subjects can be expected to infer a higher-order structure: a tool with which to integrate the various notes of the tune.

Bamberger's work reveals some of the power that can be attained by comparing the performance of the same task by individuals possessing very different levels of expertise. In the next section, we will consider the songsinging abilities of young children and adults.

## Songsinging in Preschoolers: Constructing a Basic Vocabulary

Let's consider the very beginnings of performance ability, the first songs a child sings. By viewing the prototypes of musical practice as a moment on the path that leads to sophisticated musical thought, it is possible to learn a great deal about music if we watch children's first years of songsinging. Even in the beginning songs, the earliest efforts to participate in musical thinking and practice, musical elements common to all songs emerge. However, during the development of musical knowledge, the significance of various dimensions may change. For example, rhythmic structures appear to be relatively robust at first hearing, especially and when compared with children's grasp of pitch. However, careful analysis of the quality of early rhythmic structures which children use in songs appears to be first closely linked to words (Davidson & Colley, 1987). Later, independence of rhythmic development from words begins to emerge as children become able to hold a note across a beat or pulse (as with a dotted quarter note in common time). Similarly, dimensions that fail to appear early may require considerable support at later stages of development, as in the case of key stability (Davidson & Scripp, 1988c) or music students' knowledge of key use (Davidson & Welsh, 1988).

The following three songs illustrate the kinds of changes that are at issue. Despite their differences, each of these songs is a version of "Twinkle, Twinkle, Little Star," sung by the same child at three different ages (Figure 4.2).

At age three, Jeannie already knows several important things about the way songs combine words and pitches. She knows that in singing, more than in speaking, the voice is consistently modulated in a way that produces a complex contour of ups and downs, and she also knows that the pitches in the melody should be distinct from one another (i.e., not continuous slides). In addition to understanding these general characteristics of music, Jeannie has some more precise knowledge. She duplicates some of the rhythms of the standard version of the song and can divide the song into smaller units or phrases, yet she makes no attempt to measure the interval between the phrases. Where pitch relations are concerned, the contours the child sings do not match the model and her song

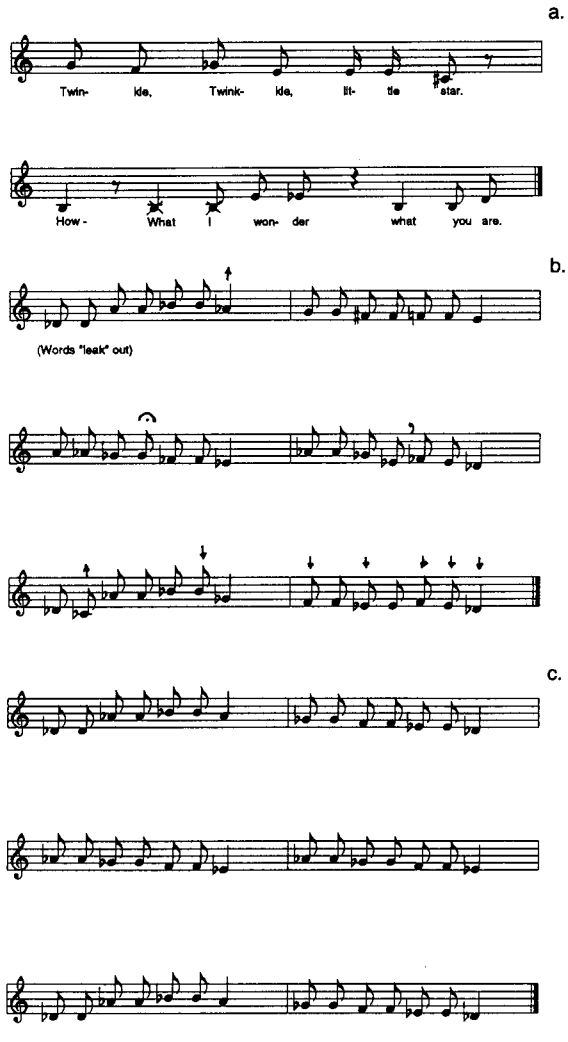



FIGURE 4.2(a,b,c). Jeannie's three versions of "Twinkle, Twinkle, Little Star."

lacks the tonal coherence that a stable scale would provide. The song structure is difficult to recognize: Jeannie sings only two of the song's four phrases, and the ends of the phrases are not marked by notes of longer duration.

In the second version of the song, when Jeannie is four, she carries the melodic contour by humming. At the same time, she forms the words with her lips closed, thus sounding all the changing vowel sounds in the text. This suggests that at four, she is still dependent on the text to keep her place in the melody. While the contour of each phrase is closer to the model, some of the notes and intervals remain inaccurate. The way these phrases approximate the model suggests a greater pitch stability. However, the fluctuating intonation and the lack of

a note-to-note match across phrases make it difficult to give her credit for having a firm sense of tonality or scale. However, the rhythmic shapes are much sharper in this version. At the end of the phrases, there is a longer note, but the interval of time between phrases is not yet measured by an underlying pulse (i.e., the steady beat of the music).

By six years of age, Jeannie's version of the song matches the model. The contours of the phrases are accurate, and the pitches and intervals suggest regulation by a sense of tonality and scale. She lends a steady pulse to each phrase and this pulse controls the duration of time between the phrases. While the words are in place, the completeness of the music suggests that the text no longer dominates the musical aspects of the performance. It is equally significant that the child now uses the constituent phrases of the song (the first and second phrases) to make a simple three-part form, which can be represented by ABA. The repetition of the melodic material of the A section at the end of the song is exact. Every aspect of the song is related to a structure that governs the whole: a single scale, a regulating pulse and surface pattern, and a specific form that binds the individual phrases into a coherent musical whole.

These three songs provide a sense of the nature of the essential questions in musical development. The streams of pitch and rhythm emerge as primary, for two reasons. First, studying these aspects of musical understanding provides explicit connections between children's songs and those of adult singers. Moreover, as the examples illustrate, each of these, pitch and rhythm, exhibits considerable development in the years between one and six.

With respect to each of these streams, there are critical questions to consider. First, we want to explore how development occurs in pitch: Is it a process of gradual approximation of adult models? Or, do children organize their songs using qualitatively different principles? If the latter is the case, can we describe these principles in terms of rules or stable structures? Is development steady or are there marked shifts in musical abilities?

The three versions of "Twinkle, Twinkle, Little Star" that Jeannie sang suggest that learning to sing songs is a more complex task than it appears at first. Learning to sing songs involves: (1) singing articulate pitches that reflect an underlying scale or tonal structure; (2) performing both surface and underlying rhythmic patterns regulated by a common pulse; (3) mastering the canons of song form that guide standard songs so they can be applied to invented compositions; (4) forming structures that make possible the use of internal reference (repetition, variation, and development).

## Development of Pitch Relationships

Three components are basic to the control and understanding of pitch relationships: (1) the capacity to grasp the figurative shape or contour of a phrase; (2) the ability to match individual pitches; and (3) sufficient memory or organizing

structure to maintain key stability across the phrases of a song. There has been considerable discussion about the manner in which these various aspects of pitch relations develop. Some earlier research suggests that knowledge of pitch relations may develop through a series of broad levels or phases in which children grasp first the contour, then the specific pitches, then intonation, and finally, key stability. For example, Moorhead and Pond (1941) and Shuter-Dyson (1968) report that contour is perhaps the first aspect of melody to be grasped by the child. Teplov (1966) finds that children sing the contours of the melody first and are only later able to match the specific intervals contained in a model melody. Still other research suggests that tonal knowledge may develop out of the ability to match specific pitches. Specifically, Wing (1963) reports that children's ability to match pitches and intervals precedes their ability to produce the general shape of the melody. However, Sargeant and Roche (1973) report that children's ability to match pitches even declines as their ability to match contours improves.

There is more widespread agreement that the ability to sing in key appears late. Michel (1973) reports that by the end of the preschool years, children are able to perform music in major and minor scales, and even to appreciate harmony. However, this rich description of young children's tonal understanding is questioned by other researchers who report that children even four or five years of age are not able to sing scales at more than chance levels (Bridges, 1965).

Beneath these descriptions lie two difficult, but highly significant, problems in analyzing children's tonal knowledge. First, many analyses have been too narrow in focus insofar as they have examined either the development of interval matching or the growth of the ability to seize the figural shape of a melody. If one studies tonal development by breaking up musical performances into their smallest units or intervals, one loses an important aspect of pitch relationships—the ongoing motion of the melodic impulse. On the other hand, if one focuses only on the motion of unfolding melodies as contours, one misses a close look at the precision, accuracy, and stability of individual intervals.

The second difficulty arises because studies of children's songsinging are often based on an inappropriate unit of analysis (Wertsch, 1985). It is evident that the tonal knowledge of musicians reflects the control of individual contours, pitches, intervals, and keys. However, and this is critically important, these musical organizers are the by-products of considerable and protracted activity constructing pitch relations. Children's early songsinging simply does not depend on linking of fixed pitches or key structures. Because of this, there is little reason to analyze their productions in those terms. Indeed, to analyze the songs of a three-year-old in these terms is akin to applying the rules of point perspective to early drawings or the rules of transformational grammar to one-word utterances. In effect, the key to understanding the development of musical ability lies in discovering a unit which captures the melodic aspect of children's songs but which does not presume that children organize their songs in terms of notes, contours, or keys from the outset. Thus, an early and major focus of our work was

to develop a way of describing children's growing understanding of pitches, intervals, and contours which did not assume that the child used fixed pitches, measured intervals in diatonic terms, or organized performances based on knowledge of keys.

So how do children, in fact, learn to control pitches of their songs? Although researchers confirm that musical pitch development follows neither course associated with the adult predictions suggested (Dowling, 1982; Ries, 1982), a systematic and empirical model of pitch development has been offered. Systematic observation of 78 children over a six-year period at Harvard Project Zero suggests that *contour schemes* rather than gradual refinements in the overall contour or interval matching account for the route of emerging tonal knowledge (Davidson, 1985). The term *contour* conveys the importance played by the figurative shape of melody—its contour; while *scheme* conveys the importance of recognizing levels of mental organization. This model, which emerged out of the Project Zero research, suggested five specific levels of pitch development in young children's singing between the ages of one and six years. These levels characterize children's knowledge of tonal relationships as expressed through their performances of songs.

When singing invented and conventional songs, preschool children first use these melodic schemes with a descending direction. Except for the first contour scheme, the phrases children sing are longer than individual notes or intervals, yet, at the same time, smaller than the individual melodic phrases of model songs. With development, the vocabulary of schemes grows from those that span roughly a third to those that span roughly a sixth. It would simplify the story if, as one might assume, development occurred by filling in the initial space with stepwise movement, and then expanding the tonal space by adding more steps. This would be efficient (on the surface of it) and would reflect some models of ear training and sightsinging currently proposed (Adler, 1979). But the path of development is not so neat and tidy. Contour schemes do not develop in ways one might logically predict from interval learning—that is, proceeding from smaller to larger fixed intervals. Although they begin as small, descending motions and expand through a fourth, fifth, and sixth, this simple line of development does not immediately account for the occurrence of stepwise motion within each tonal frame.

Expand the boundaries of the tonal space first, then go back and fill in the gaps in previously acquired schemes with steps—that appears to be the rule. For example, only after developing the next larger contour scheme does the child fill in the gap between the top and bottom of the previous contour scheme with stepwise motion. For example, not until a child has acquired the ability to reliably sing a fourth (as a leap) does she fill in the earlier leap (the third) with stepwise motion. This "backing and filling" process continues until finally the contour scheme of a sixth has been achieved. At this point the accretive process of development stops. Contour schemes do not continue to develop in the same

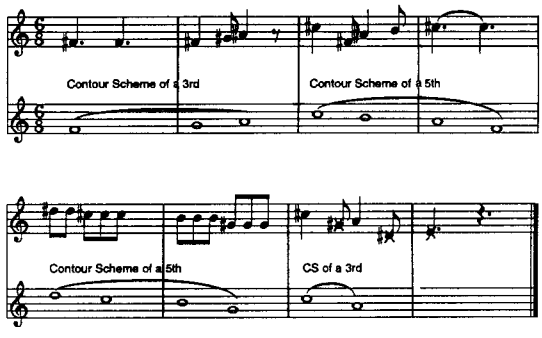



FIGURE 4.3. Transcription of a typical preschooler's performance of "Row, Row, Row Your Boat" with contour scheme analysis below.

way beyond the span of the sixth. The octave, which emerges only later, appears to arise from fundamentally different processes (perhaps having a great deal to do with memory for register, for example).

Figure 4.3 shows how these schemes function only in very local contexts (i.e., within an individual phrase) when preschool children sing "Row, Row, Row Your Boat." Substantially more developmental change is required to construct a level of tonal stability that can be sustained from phrase to phrase and in a variety of contexts.

The increasing differentiation and stability of pitches are two characteristics of pitch-singing development in the early years. At first, preschool children fail to return to established pitches or reference points in their songs as they generate only the individual contours which form the child's repertory of melodic schemes. As the repertory of contour schemes grows, the number of different reference pitches used when singing a song decreases until there is a relatively stable one-to-one correspondence between the pitches of a model song and the pitches the child sings. However, during the preschool years, the locus of this stability is restricted to pitches within the unit of the phrase. It does not yet extend across all the phrases of songs.

## Collecting Songs and Forming a Database

The setting in which musical behavior occurs is very important. Much research is conducted in laboratories in which designated variables can be controlled. However, what is learned from laboratory and clinical settings cannot easily be generalized to everyday experience. By taking a more ecologically sensitive approach to the study of music and how our musical knowledge develops, we can trace the connections between the productions of children, adults, and professional musicians. The songs on which this study is based were collected in the homes as part of a larger study of children's symbolic development. Seeing the children every other week for six years made the researchers "part of the family."

The songs collected during the study include the traditional songs American

children learn during the preschool years, as well as some songs that were composed for experimental purposes. In addition, many novel songs were invented by the children under a wide variety of circumstances—sometimes requested as a part of warm-ups to experimental tasks in other domains, sometimes as spontaneous responses to looking at "wordless books." The more than 500 songs that form the corpus were collected under a wide variety of contexts: inventing a song at the request of an observer, singing while playing with the materials of another domain (most often drawing or playing with blocks), singing favorite songs on request, and singing on private occasions when parents thoughtfully made home recordings of their children's performances.

A word about the methods. While the data from this study are rich, the corpus of songs scored includes only those songs which consisted of two or more phrases. The songs were transcribed by experienced musicians and checked by outside experts. Because every attempt was made to reflect the characteristics of the children's performances faithfully, the notations include special symbols. For pitch, in addition to conventional notation, there are symbols for speech, contour, and notes which are half spoken and half sung. The rhythm symbols used reflect free groupings, proportional durations, and more measured and regular rhythmic patterns. Analysis of each song was done from both the notation and a tape of the child's performance.

The transcription and the analysis of the songs were separate processes. Before scoring a song, we determined the number of phrases it contained and, when possible, the tempo at which it was sung. We also classified the song as either invented or standard, and the manner in which it was performed as either spoken, sung with diffuse pitches, or sung with articulated pitches. We then rated each phrase of the song for the way in which the boundaries of the phrases were marked, the type of tonal motion it contained, and finally, the type of song structure that the song exemplified.

What conditions must be present in order to determine that a musical response or behavior is stable? In short, when does a child acquire a behavior as opposed to simply showing an isolated instance of its occurrence? A conservative rule was developed. In order for a musical shape to be counted as a stable part of a child's repertoire of pitch relationships, it had to appear twice in different songs and in different sessions. Separate analysis was carried out on each group, invented and standard songs. This made it possible to distinguish between development within the two different repertoires of songs the children made up and the songs of the culture they were taught. (For more detailed description of the scoring process the reader is referred to Davidson, 1985.)

The following composite picture is based on the findings of a longitudinal study of nine firstborn children who were followed for five years. The first visits—which took place two times each month—were made shortly after the children's first birthdays. The findings of this longitudinal sample were replicated in a cross-sectional study that was made up of 69 children who were grouped by age across the same age range (from a matching socioeconomic class). Among

other tasks, the children in the cross-sectional sample were asked to sing their favorite songs and a song was taught to them during three separate visits. These songs were tape-recorded and later scored on the same measures as the songs from the longitudinal study.

## Constructing the First Boundaries of Tonal Space

### *Contour Scheme: The Interval of a Third*

There are many ways to tell the story of what is happening in Heather's song, or to put it another way, to determine the units of meaning in this song (Figure 4.4). Here, the danger is that the observer attributes an irrelevant meaning to the child's production. For example, from one perspective, one could argue the child is attempting to use many adult-like goals. One could say that Heather "decided" to invent a new version of the song "A, B, C, D" by beginning on what are "intuitively" more stable notes (the notes of a tonic, G-sharp-E), and closing on less stable notes (notes of its dominant chord, F-sharp-D-sharp). But once stated, this version seems somehow inadequate. For one thing, it at least implies that, since these tonal concepts are already visible in the songs of toddlers, appreciation and understanding of these tonal concepts (tonic and dominant functions) may be genetic givens, independent of culture, and thus musical universals. This story suggests that the child is already considering and using the materials of music in ways which music students find difficult even 17 years later (Davidson & Welsh, 1988). Thus, the unlikeliness of this precocious grasp of the harmonic structure of Western music by any but a very few makes it difficult to argue that the units of meaning for children in their songsinging are derived from the expert's formulations of musical theory and practice.

From another point of view, of the units of meaning a child uses when songsinging can be redirected to address the issue of the musical units of pitch, interval, contour, and phrase. A story based on this perspective might be about how children learn the rudiments of tonal structures by long experience, attempting to match the correct intervals or engaging in "trial and error" approximations of melodic contours, without regard for intervals. In any case, the stability and regularity of contour schemes, revealed by analysis of hundreds of children's songs, argues against such a successive approximation model of learning that controls the tonal materials of music.

The children's songs themselves suggest another possibility for units, the phrases. It is characteristic of young children's songs to *not* reflect a concern for melodic continuity beyond the level of the individual phrase. For example, at

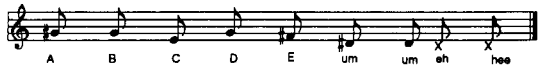



FIGURE 4.4. Heather's "A, B, C, D."

first they do not heed the importance of "reading" or acknowledging the rests and silences between phrases. Instead, they treat the space between phrases as "free" areas which do not need to be measured or regulated, quite unlike adults' performances. Children may develop the ability to control and regulate musical silences most naturally under the focused circumstances of a specific repertoire. The child distinctively segments the musical line into units, taking a breath between the units in order to begin speaking. Thus, phrases are the initial musical units of song.

There is more to say about Heather's version of "A, B, C, D." The reduced range of this song illustrates an important feature of psychological analysis: psychological boundaries in music are different from the physiological boundaries. When children first draw, they work from feature to feature—for example, from head to body to arms, and so on. They work without an overall scheme or design with which to regulate the individual units. They move from one unit to the next and the drawing is the result of a series of localized moves (Goodnow, 1977). When singing, children operate in much the same way. Early songs suggest that children operate on a principle of nearness similar to that found in drawing research: the notes of early songs and melodies are not extremely disjunct, but surprisingly contained within a small register. Regardless of the melodic model, and they appear to be used as repetitive chunks, the notes of individual phrases are relatively close together.

The constrained range children use when singing is not physiological. Children can vocalize across a wide range (Fox, 1990; Jersild & Bienstock, 1931). Yet Heather fails to imitate the melody of the song as she has heard it. This is consistent with the performances of the other children of this study. Unless excited or pushed to do so by an adult, children don't use their full vocal range when singing songs. Instead, they use a psychological unit or scheme as a means of coordinating the tonal aspects of their singing.

There is no evidence of the highly distinctive contour of the song's melodic line. The ascending perfect fifth leap of the first phrase is nowhere to be seen. Typical of children's first songs, the contour of this earliest of melodies is a small descending motion. Greatly reduced in length, in number of words, using a compressed register with a descending contour, Heather's song is representative of the first songs children sing.

There are several more things to notice about Heather's early performance of "A, B, C, D," however. The song is not only reduced in register, it is also severely reduced in length. Only two phrases of the song are sung. The words of the song are very simple, yet only a small portion of them is rendered. Nevertheless, the singer's ability to render the song appears to draw more heavily on the text than on the melody—that is, the song is more fully represented by its text than by its melody. Also typical of the first songs children sing, the text consists of single words. However, the order and sequence of phrases of early songs are not always that of the model. The salience of the text plays an important role in what phrases of the song are going to be produced.

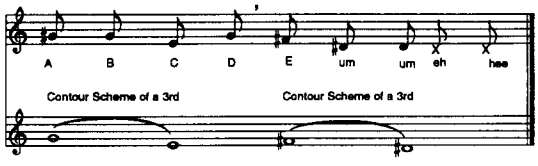



FIGURE 4.5. Heather's "A, B, C, D," showing the construction of contour schemes.

Throughout the development of songsinging there is a consistent pattern in children's grasp of pitch relationships: once a set of boundary pitches has been established and integrated as a stable mental structure, that scheme is used in conjunction with others to reflect more accurately the contour of a given standard song. Contour schemes have several characteristics. First, each contour scheme is identified by a single framing interval: a third, a fourth, a fifth, and a sixth. These "framing intervals" mark the boundaries of the contour of a phrase of song as the child sings it. Thus, the initial contour scheme delineates the interval of a third, and there is a leap between the two notes that form this interval. Later, this leap is filled in, and a more mature form of contour scheme is revealed through stepwise motion. This stepwise motion appears first in the descending form, and with the development and an increasingly operational grasp of the contour scheme, it appears in the ascending form as well. In the initial stages, the pitches of a contour scheme are relatively unstable. Often sounding like a portamento (a special manner of singing with the voice gliding gradually from one tone to the next through all the intermediate pitches; *Harvard Dictionary of Music*), these diffuse pitches become increasingly stable and articulated with development.

Contour schemes appear to represent the tonal structures for the musically untrained child as scales represent the tonal structure of the more musically trained. Children use their repertoire of contour schemes, either singularly or in tandem, to form and control the tonal materials of their songs. Children who are using contour schemes as their mental representation of the pitches of a song will reduce the range of the contour of a song that goes between the range of the most recent contour scheme acquired. In many instances, they will even inappropriately expand the range of a phrase in order to match the size of the framing interval of a new contour scheme under construction. In "A, B, C, D," Heather is using two juxtaposed contour schemes of a third. As Figure 4.5 shows, Heather is using two intervals of a third.

### *Contour Scheme: The Interval of a Fourth*

Kori's version of "Itsy Bitsy Spider" illustrates the contour scheme of a fourth. Throughout the melody the chunks of phrases are bounded by the intervals of a third or fourth. Like "A, B, C, D," the text appears to be driving the content of this song, in that the pitches of the song are not nearly as well represented as the words.

This contour scheme is characterized by a tonal frame of a fourth—that is, the

interval of a fourth forms the top and bottom of the unit of measure at this level of development. As with the previous level, the child may imitate larger intervals when they are modeled, but the child will revert to the fourth as the largest stable interval within a phrase when left alone. Leaps larger than the fourth, when they occur in standard songs, are typically compressed to the smaller size (Figure 4.6).

Children appear to construct the contour scheme of a fourth by adding an outside step to the tone of the boundary notes of the already stable contour scheme of a third. The first step of this level is characterized by either descending step-leap, or leap-step, motion, which more or less outlines an interval of a fourth (approximately in perfect or augmented form). As with the contour scheme of a third, this motion still does not reveal an underlying scale structure. The high-low dichotomy functions as it did in the contour scheme of the third, only now expanded to the slightly wider interval. There is still little stability of individual pitches at this stage. During this period of development, the framework of a third and fourth provides the stable structures for songsinging.

The average range of songs sung during this phase of development expands. It includes the notes of an octave, specifically from c′ to c″, or from middle c to c one octave higher. Each song does not use the entire range. Because the contours of standard songs are compressed to fit the framing interval of the contour scheme, however, individual songs can start higher. The contours of the melodies of this level are still sometimes diffuse and unmarked by articulated pitches, and now, both rising and falling contours are sometimes found. However, more of the pitches are clear.

An important milestone of development occurs during this level: individual pitches of songs can now be matched when they are sung in the appropriate range. Even intervals larger than a fourth can be matched if they are sufficiently supported, prompted, or modeled. These larger intervals are not used in songs, however, suggesting that songsinging uses different skills than interval matching; that representations of melody are based on contour scheme rather than intervals. Additional evidence supports this perspective. When the child is learning a

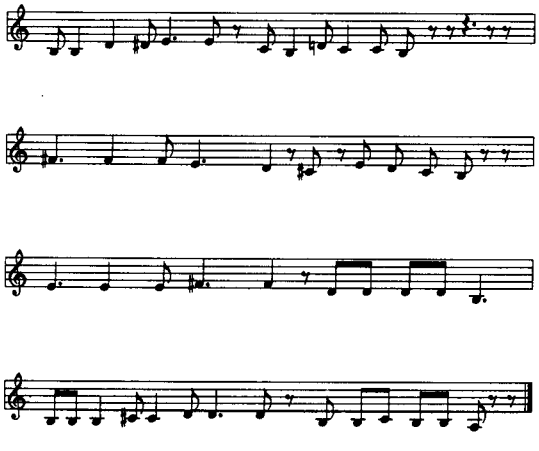


FIGURE 4.6. Kori's "Itsy Bitsy Spider."

new song, or singing a song while performing some other task (such as playing a set of bells, drawing a picture, or moving about the room), the contour scheme currently under development, when it occurs in a song, may itself be replaced by a more stable and smaller scheme—in this case, that of a third.

As noted earlier, the child at this level does not have a reliable way of representing leaps larger than a fourth, so larger leaps in standard songs are compressed. Also typical of this level, there is little evidence of key and pitch stability. Indeed, the tonal structures used at this stage are still extremely local in nature, which means the child does not always match the high or low note within the contour scheme being used.

As with the contour scheme of a third, the second step of this level involves a filling in of the framing interval. This stepwise descending fourth pattern occurs only after the boundary interval of a fourth is stable and after the contour scheme of the fifth is acquired.

### *Contour Scheme: The Interval of a Fifth*

Max's rendition of "Happy Birthday" illustrates the use of the contour scheme of a fifth (Figure 4.7). The range of the first two phrases (d′–a′) establishes the framing interval. The notes he uses in the first two phrases (a′, g′, f-sharp and d′) are typical of the first, unfilled form of the contour scheme. They represent the linking of the two forms of the contour scheme of a third (see Figure 4.8).

Like the others, this is another two-step contour scheme—the first unfilled, the second filled. In both, as in the earlier contour schemes, the descending contour initially predominates. Later, the contour can be rendered flexibly (as in Figure 4.7). The shapes that result in this contour scheme are few in number and can be derived from the two previous contour schemes. The framing interval of the fifth coordinates the notes within this register.

When children sing a song using the previous contour schemes, they appear to not return to specific pitches, but only move up or down reflecting the contour of the phrases. Thus, often there are relatively few shared pitches within a song. By contrast, when children use the contour scheme of a fifth to sing a song, they tend to return to specific pitches as they sing the contour of the piece. Thus, now there are more shared pitches across the phrases of the song. This produces songs that begin to sound as though there were a superordinate structure like a scale in place. However, while the effect of this is to create songs that may use as few as five pitches, this should probably not be taken as evidence for a scale or key structure. For one thing, as in the earlier levels, the child may imitate intervals

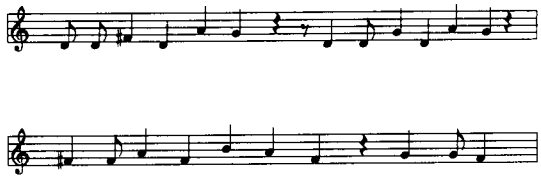


FIGURE 4.7. Max's "Happy Birthday to You."

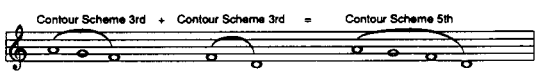



FIGURE 4.8. Constructing the contour scheme of a fifth.

larger than those of a fifth when they are modeled, but the fifth is still the largest stable interval. That is, intervals exceeding a fifth are typically compressed to a fifth. Furthermore, although the child may give an appropriate tonic (or ending note) in a completion task (especially if prepared by a stepwise descending motion), as in the last phrase of "Row, Row, Row Your Boat," the use of contour schemes in his invented songs does not show the coordination among pitches or the constraints which go with knowledge represented by scales. The child's use of the contour schemes is the same as in the earlier levels, as a high-low boundary.

When a song contains leaps larger than the interval of a fifth, they are compressed to a fifth. Not only does compression of larger intervals occur, but also smaller intervals are sometimes expanded. In these cases a smaller interval (which we know a child has acquired) may be pulled in the direction of a fifth (or whatever contour scheme is being constructed) and thus distorted in a new way. This illustrates the powerful influence that a mental structure can have on memory, perception, and production.

During this level, the average range of the singing expands very little, up a step from c″ to d″. As in the earlier levels, an individual song probably will not use the entire range. Also during this level, the nature of diffuse singing changes markedly. Diffuse singing does not occur as part of a rising and falling contour, but may sometimes occur as a sliding between two discrete pitches. This may reflect an unstable contour scheme, but the individual pitches are much clearer and more distinct than during the previous contour schemes.

There is another development. The child is now able to differentiate between the starting and ending pitches of phrases by often making the ending pitches lower. There is also evidence that during this level a child begins to use an internal model when singing. The following example shows that the child was able to correct the intervals of the first phrase during the musical repetition in the second phrase.

When learning a new song, the child characteristically reduces the contour of the melody to earlier acquired levels (e.g., contour scheme of a third). Later versions of the song show that the child has replaced those smaller intervals with the appropriate one of a fifth. There are a variety of ways of ordering the internal space of this contour scheme. All of these ways can be related to the previous two levels. For example, one way of constructing this level is to repeat the process of adding on to an already stable contour scheme. This often occurs after the contour scheme of a filled-in third has been acquired. Another way of constructing this contour scheme involves interlocking thirds, which can also be derived from the contour scheme of the third.

As with the contour scheme of a fourth, the second step of this level involves filling in the framing interval. This stepwise descending fifth pattern occurs only

after the boundary interval of a fifth is stable. Also, this level occurs relatively late, as anyone who listens carefully to a five-year-old's version of the last phrase of "Row, Row, Row Your Boat" can testify.

### *Contour Scheme: The Interval of a Sixth*

Maja's version of "Jack and Jill" provides a good example of contour schemes at this level (Figure 4.9). She is the only one to acquire the filled-in version of the contour scheme of a sixth. As in the previous levels, the many routes to this contour scheme may be derived from a limited number of strategies, based on previous contour schemes. Phrases of melody that employ a range of tonal frames of more than a fifth, but less than an octave, are characteristic of this level. Some of the ingredients of this level are the result of linking the contour schemes already mastered. Thus, the contents of this contour scheme are less narrowly defined than the earlier ones. Because there is not much evidence of a stable octave, the organization is still based on the local knowledge that contour schemes represent.

The average range of the songs at this level extends a little on both ends, spanning the notes between b and e″. Songs end on pitches that are lower than those on which they begin. As before, individual songs do not use the entire range, but only a portion of the tonal register available. The child sings pitches that are well formed and clear during this level. Diffuse singing is sometimes found around the outer limits of the register, and there may be an occasional sliding between two distinct pitches.

There is a major change in the stability of individual pitches within a phrase and across phrases. At this level, a child can remember a specific pitch and, having left it, can return to it. This makes it possible for the child to sing and resing notes and pitches consistently throughout an entire song. The result is that the songs are much less chromatic than during earlier levels, because the phrases of songs at this level tend to share the same pitches—rather than to use different pitches. Since the child still lacks a scheme for the octave, intervals larger than a sixth are usually reduced to the interval of a sixth. "Somewhere over the Rainbow" shows how this reduction occurs (Figure 4.10).

Another interesting strategy begins to emerge. It shows the role memory can now play in melodic construction. The greater memory capacity that the child now has is used to fix a specific register as a reference point. For example, when

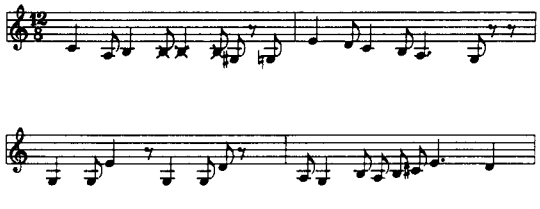


FIGURE 4.9. Maja's "Jack and Jill."

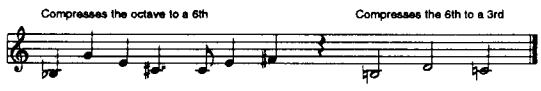



FIGURE 4.10. Jeannie (4:6) "Somewhere Over the Rainbow."

Jeannie now sings "Old MacDonald Had a Farm," she places the melody of the phrase "EIEIO" in a higher register and returns to the same register and same notes whenever the phrase is needed. She does this without regard for the appropriate tonal relationships across phrases. The song appears to exist in two distinct registers, one distinctively low and the other distinctively high. The song is sung maintaining these two registers throughout.

### *Contour Scheme: The Interval of an Octave*

While none of the children sing the interval of an octave before the age of six, there are some single instances that indicate what can be expected. As with earlier levels, the unfilled space appears first. Also, like the earlier levels, this level is constructed by combining earlier contour schemes. The song in Figure 4.11 exemplifies songs that count as instances of this level.

This song is very different from earlier ones in several ways. First, the only notes present in the song are those of an appropriate scale (i.e., they meet a musician's expectation). For example, Maja's song uses only the notes of the F major scale (or, if you prefer, a pentatonic on F since there is no B-flat in the melody). Second, these songs differ from those of earlier levels in that they tend to begin and end on the tonic of the scale, even though the range is now bound by the vast space of the octave. Another way this level differs from earlier ones is that notes on either side of leaps are discrete—there is rarely any glissandi or diffuse singing. Finally, some completions of target pitch patterns (the child is asked to finish fragments of a descending scale) show that the appropriate tonic can be inferred from the context of a given pattern of notes. It appears that these traits are characteristic of both invented and standard songs.

This ends the story of how young children construct the tonal space of their songs. It has been argued that children's grasp of tonal relationships undergoes systematic development as they expand their repertory of contour schemes and, correspondingly, as they more often return to individual referent notes of the model song. Additionally, it appears that development does not proceed neatly from steps to leaps, nor does it follow a direct sequence from simple to more complex intervals. Finally, a significant first phase leading to mastery of tonal space appears to be completed by the age of six or seven. What happens next?

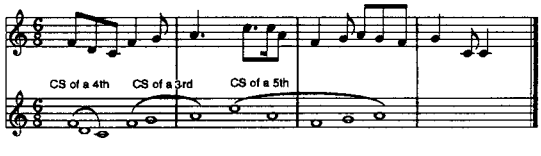



FIGURE 4.11. Maja invented song.

## Meeting the Culture: Middle Childhood and Beyond

After this phase, development appears to take another direction; knowledge of tonal relationships is linked to the individual's song repertory. In many ways this progression parallels that found in the visual arts, where, after a burst of creative brilliance, young children refocus their drawing efforts toward mastering the somewhat mundane and canonical forms of objects. The resulting flatness of their efforts is dull when compared with their earlier products.

A similar closing down and focusing on cultural forms may also occur in music. Young children, after a long period of eagerly engaging in musical tasks like inventing a song about a rabbit, become shy about singing. Now, after being asked to invent a song, they are likely to respond by throwing back a challenge, "You sing one." One may lament the loss of invention, but there is another interpretation of this phenomenon. This seeming reduction of creativity is perhaps the next important step toward mastery of the materials of the domain. Children have, by now, a rudimentary grasp of the materials of music, pitch, and rhythm. During this new phase, children are becoming aware that there is a cultural bank of standard songs, and they learn that their knowledge of those songs is limited. From this perspective, their response, "You sing one," is social in nature. It reflects a desire to take part in the culture which lies beyond that of their own making. Thus, during the years of middle childhood, the focus of activity shifts away from the materials of music toward practicing the canonic forms, thereby learning the artistic language of culture. There are, doubtless, other interpretations. Yet the parallel in visual arts strongly suggests that in music there may also be something like the "retreat" from free invention into prolonged study of cultural forms.

During the years of early adolescence music becomes extremely important to children. Children now spend an enormous amount of time listening to and learning songs that reflect the repertoire of the culture, usually the favorites of their peer group. During the middle childhood years the focus of musical activity shifts, perhaps beyond the canonical forms of the music of the culture to using music as a means of negotiating and establishing social relationships. What is the effect of so much listening and the singing it generates? Such intensive exposure should have a powerful effect on their understanding of tonal space, but does it? Does this amount of singing result in the construction of a hierarchical tonal structure to be used when performing songs? If contour schemes are characteristic of young children's song performances, at what point in development and under what conditions does a singer's command of tonal space reflect the hierarchical thinking evident in the construction and use of scales?

## Song of the Culture: Way Points to Operational Knowledge

Ongoing work at Project Zero is taking up the issue of the continued development of tonal knowledge into adulthood: How do musically untrained adults use tonal relationships in singing songs? Do adults use scales to integrate pitches when they sing songs, or do they continue to use the looser tonal structure typical of the contour schemes we observed in children (Davidson & Torff, 1991)? Put another way, how do musically untrained adults differ from or resemble the seven- or eight-year-old child's performance of children's songs? Are their performances similar to those of children in spite of years of experience listening to the music of culture?

There are some clues that indicate what we might expect. There is a strong resemblance between the invented notations of seven-year-olds and untrained adults (Davidson, Scripp, & Welsh, 1988), as Bamberger noted in her comparison of musically untrained children and adults (Bamberger, 1986). Surprisingly, there is little difference between the notations made by the two groups, in spite of the fact that adults have had nearly 15 years more experience hearing music than the children.

Turning from children's performance of songs to that of adults, we found that the study of musically untrained adults' songsinging ability has three conditions (Davidson & Torff, 1991). The first one is simple: Sing "Happy Birthday" once, then resing it. Since the second performance might be much better than or different from the first, this controls for the effect of warming up. In the second condition, subjects are asked to sing the song and then listen to a tape of someone else singing the song—this version contains the most frequent errors made by adults. After identifying the location of the errors, they sing the song again. This allows us to look at the relationship of their singing to their perception of errors. The final condition is like the second, except that the tape is of their own singing (not someone else's version). They are asked to reflect on their performance, imagine how the song goes, and compare that with the version they just performed. At the end of the discussion, they are asked to sing the song again.

Although the study is in progress, some of the results already noted have a bearing on the topic of this chapter. In the case of tonal knowledge, it is clear that a significant number of musically untrained adults sing "Happy Birthday" without employing the integrating structure of the scale. As with young children, adults typically compress the range of the third phrase (an octave) and surprisingly end the song in a different key from the opening phrases. These two errors strongly suggest the use of contour schemes, as we saw in our study of children, for these reasons: first, the octave is usually performed as a smaller

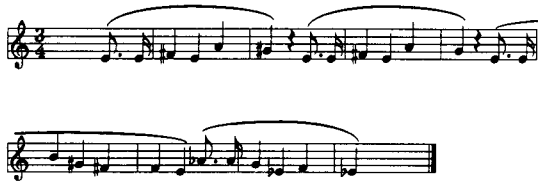


FIGURE 4.12. Typical errors in an adult's version of "Happy Birthday to You."

interval, typically compressed to a fifth or sixth, just as young children do; second, the last phrase is lower than it should be, ending in the register of the starting notes, suggesting that closure depends more on register than on hierarchical tonal relationships (Figure 4.12).

In our ongoing work with adults we are finding that musical performance knowledge is remarkably stable. And there is no significant warm-up effect; repeated singing does not change the quality or accuracy of the first version. Furthermore, even after hearing their flawed version of the song, many subjects report that it represents their "best" version of the song. Others will describe their errors as "not high enough" (for the third phrase) and "got off" (for the ending) and then resing the song without correcting the problems they have clearly identified.

The questions this work raises extend beyond the scope of this chapter, but one question is worth pursuing. What is the nature of the tonal understanding that results from the songsinging of middle childhood?

## Musical Study: Continuing the Path of Development

The course of development enters a new phase when music instruction begins around the age of 10 or 11 (for those students who take lessons on an instrument). Contour schemes begin to be integrated into the hierarchical structure provided by the scales of the culture. However, as tonal memory and the use of scalar structures expand, sensitivity to tonal functions increases and the contour schemes are gradually, yet still only somewhat, coordinated with the tonal vocabulary of the culture. Clues about the limitations of this path of development lie in the performance of major scales under a variety of conditions by students entering professional music schools. Their performances suggest they know scales as whole chunks, as melodies. These structures are difficult to pull apart and are unstable if not performed in a manner or speed which allows the student to perform, perceive, or monitor the entire shape as a single unit.

By the time music students enroll in a college-level sightsinging class, they bring with them a vocabulary of contour schemes, a repertoire of standard songs that they share with nonmusicians, the perceptual knowledge of tonal functions, a first-draft knowledge of the scale, and years of instrumental training and exposure to music literature. Given this background, it is odd that when they enter professional schools to begin formal music study, their knowledge of scale sys-

tems and tonal melodies is so incomplete and their grasp of tonal structure is so superficial.

Yet this superficial stage of tonal knowledge is easy to demonstrate. Just a few moments with a beginning solfège (sightsinging) class in the United States reveals students' relatively figurative grasp of scales. Although able to sing scales and melodies, many will be unable to easily sing specific notes of a scale on demand, for example, scale degrees four or six. Additionally, if individuals are asked to sing a one-octave span on a neutral syllable starting on scale degree two of a major mode (singing the notes of the C major scale starting on the note D), one of three things happens. Some students will sing the octave but transform the second scale degree into a tonic and sing a major scale. Many more students stop singing when they reach the tonic instead of completing the octave by continuing up one more note. In each case the students often are not aware of what they've done. Still other students use knowledge extrinsic to music—that eight steps make an octave—to determine when to stop. They solve the task by counting up eight notes as they sing.

These responses are evidence that students have not yet learned to partition musical space on demand, nor do they have an articulated and internalized representation of tonal space. They know the scale as a whole performance act, a string of notes linked into specific relationships. Singing the scale as if it were a "melody," they stop where it "feels right," not where they are asked to stop. In other words, they are unable to monitor and control their performance of the scale to show musical understanding of the parts in the context of the whole; merely counting the number of steps is not enough to engage intrinsic musical sensibilities.

Their knowledge of pitches within a scale is initially very much like the nonmusician's grasp of the notes of a song. The nonmusician's understanding of a melody is static, a rote performance without reflection, a recitation without comprehension. The song becomes a sign for music, formulaic and canonical. It is understood as a whole, not as a collection of independent units yoked into special connections. It is as if the smaller units of contour schemes are replaced by the relatively larger, yet completely context dependent, phrases of melodies.

## Development and Talent

This chapter looked at the abilities of children and adults to sing tonal songs, but our findings carry implications for the relationship of development to talent. The context for investigation has not been the laboratory, but the natural settings in which nonmusicians first begin songsinging and young musicians first begin training. Given the point of view advanced here, one can begin to appreciate the effect of musical training on the relatively naïve musical knowledge of the novice in unsuspected ways. For one thing, it suggests that, to a remarkable degree, nonmusicians and beginning musicians share the same reference system when

they are singing unaccompanied songs. Surprisingly, when deprived of the advantage of their instruments, these young talented musicians look like the nonmusicians. This, in turn, strongly suggests that instrumental training by itself does not guarantee a grasp of musical relationships, and that musical thinking in these cases is constrained to the use of the instrument.

This finding has broad implications for the study of musical talent. A developmental picture of musical knowledge suggests that natural gifts may not develop very far when allowed to grow freely or when students are not encouraged to use their knowledge and skills in ways that promote integration. Instruction focused exclusively on playing an instrument or passing on the masterpiece of the culture fails to carry the mind of the individual far enough beyond the instrument and into musical thought to make integration and operational understanding likely. A far broader program is suggested. A developmental perspective suggests that a range of musical instruments and contexts be explored rather than a restrictive focus of musical study and instruction on an instrument of choice. The differences between younger and more mature musicians suggest that integration does occur, but that it may be dependent on the extent to which knowledge and skills are used in diverse and increasingly complex relationships to one another. Studies designed from this perspective will ensure that young musicians will be able to coordinate their use of musical knowledge and skills across a range of situations.

Finally, for all of us, knowledge of the developmental path that leads to a moving musical performance can help us better understand the remarkableness of the achievement. In addition, it can help us appreciate more fully the meaning and richness of the musical knowledge that supports such a performance.

## ACKNOWLEDGMENTS

The basic research in music reported in this chapter has been generously supported by the Spencer Foundation, the Carnegie Corporation, and the Markle Foundation. The applied research is funded by the Rockefeller Foundation.

## REFERENCES

Adler, S. (1979). *Sight-singing: Pitch, interval, and rhythm.* New York: Norton.

Bamberger, J. (1986). Cognitive issues in the development of musically gifted children. In R. J. Sternberg, and J. E. Davidson (Eds.), *Conceptions of giftedness.* Cambridge: Cambridge University Press.

Bamberger, J. (1991). *The mind behind the musical ear.* Cambridge: Harvard University Press.

Bridges, V. A. (1965). *An exploratory study of the harmonic discrimination ability of children in kindergarten through grade three in two selected schools.* Unpublished doctoral dissertation, Ohio State University.

Davidson, L. (1985). Tonal structures of children's early songs. *Music Perception,* 2(3), 361–73.

Davidson, L., & Colley, B. (1987). Children's rhythmic development from age 5 to 7: Performance, notation, and reading of rhythmic patterns. In J. C. Peery, I. W. Peery, & T. W. Draper (Eds.), *Music and child development* (pp. 107–36). New York: Springer-Verlag.

Davidson, L., & Scripp, L. (1988a). Young children's musical representations: Windows on music cognition. In J. Sloboda (Ed.), *Generative processes in music* (pp. 195–230). Oxford: Oxford University Press.

Davidson, L., & Scripp, L. (1988b). Sightsinging at New England Conservatory of Music. *Journal of Music Theory Pedagogy*, 2(1), 3–9.

Davidson, L., & Scripp, L. (1988c). A developmental view of sightsinging: The internalization of tonal and temporal space. *Journal of Music Theory Pedagogy*, 2(1), 10–23.

Davidson, L., & Scripp, L. (1989). Education and development in music from a cognitive perspective. In D. J. Hargreaves (Ed.), *Children and arts: The psychology of creative development* (pp. 59–86). Leicester, UK: Open University Press.

Davidson, L., & Scripp, L. (1992). Surveying the coordinates of cognitive skills in music. In R. Colwell (Ed.), *The handbook of music research and learning* (pp. 1293–1328). New York: Macmillan.

Davidson, L., Scripp, L., & Meyaard, J. (1988). Sight-singing ability: A quantitative and a qualitative point of view. *Journal of Music Theory Pedagogy*, 2(1), 51–68.

Davidson, L., Scripp, L., & Welsh, P. (1988). "Happy Birthday": Evidence for conflicts of perceptual knowledge and conceptual understanding. *Journal for Aesthetic Education*, 22(1), 65–74.

Davidson, L., & Torff, B. (1991). *Adult's singing of children's songs*. Unpublished manuscript. Cambridge: Harvard Project Zero.

Davidson, L., & Welsh, P. (1988). From collections to structure: The developmental path of tonal thinking. In J. Sloboda (Ed.), *Generative processes in music* (pp. 260–85). Oxford: Oxford University Press.

Dowling, W. J. (1982). Melodic information processing and its development. In D. Deutsch (Ed.), *The Psychology of Music* (pp. 413–29). New York: Academic Press.

Feldman, D. (1980). *Beyond universals in cognitive development*. Norwood, NJ: Ablex.

Fischer, K. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87(6), 477–531.

Fox, D. B. (1990). An analysis of the pitch characteristics of infant vocalizations. *Psychomusicology*, 9, 21–30.

Furth, H. G. (1969). *Piaget and knowledge: Theoretical foundations*. Englewood Cliffs, NJ: Prentice-Hall.

Goodnow, J. (1977). *Children drawing*. Cambridge: Harvard University Press.

Gregg, L. W., & Steinberg, E. R. (1980). *Cognitive processes in writing*. NJ: Lawrence Erlbaum.

Jersild, A. T., & Bienstock, S. F. (1931). The influence of training on the vocal ability of three-year-old children. *Child Development*, 2, 272–91.

Michel, P. (1973). The optimum development of musical abilities in the first years of life. *Psychology of Music*, 1, 1.

Moog, H. (1976). *The musical experience of the pre-school child*. London: Schott.

Moorehead, G. E., & Pond, D. (1941). *Music of young children*. Santa Barbara: Pillsbury Foundation Study.

Piaget, J. (1970). *Genetic epistemology*. New York: Norton.

Piaget, J. (1976). Walking on all fours. In *The Grasp of Consciousness*. Cambridge: Harvard University Press.

Piaget, J., & Inhelder, B. (1967). *The child's conception of space*. New York: Norton.

Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York: Norton.

Pond, D., Shelley, S. J., and Wilson, B. D. (1978). The Pillsbury Foundation School Revisited. Paper presented to the 26th Music Educators National Conference, Chicago.

Resnick, L. (1987, December 13–20). Learning in school and out. *Educational Researcher*.

Ries, N. L. L. (1982). *An analysis of the characteristics of infant-child singing expressions*. Unpublished doctoral dissertation, Arizona State University, Tempe.

Sargeant, D. C., & Roche, S. (1973). Perceptual shifts in the auditory information processing of young children. *Psychology of Music, 1*.

Scripp, L., & Davidson, L. (1988). Framing the dimensions of sight-singing: Teaching towards musical development. *Journal of Music Theory Pedagogy*, 2(1), 24–50.

Scripp, L., Meyaard, J., & Davidson, L. (1988). Discerning musical development. *Journal of Aesthetic Education*, 22(1), 75–88.

Shuter-Dyson, R. (1968). *The psychology of musical ability*. London: Methuen.

Teplov, B. M. (1966). *Psychologie des aptitudes musicales*. Paris: Presses Universitaires de France.

Trotter, R. (1986, July). The mystery of mastery. *Psychology Today*, pp. 34–38.

Wertsch, J. V. (1985). *Vygotsky and the social formation of mind*. Cambridge: Harvard University Press.

Wing, H. D. (1963). Is musical aptitude innate? *Psychology of Music, 1*, 1.

Wohlwill, J. F. (1973). *The study of behavioral development*. New York: Academic Press.

Wolf, D., Davidson, L., Davies, M., Walters, J., Hodges, M., & Scripp, L. (1988). Beyond A, B, and C: A broader and deeper view of literacy. In A. D. Pellegrini (Ed.), *Psychological basis for early education* (pp. 123–54). New York: Wiley.

# 5

# Coming to Hear in a New Way

JEANNE BAMBERGER

## Introduction

In the second developmental chapter "Coming to Hear in a New Way," Jeanne Bamberger focuses on the relationship between the description of music and the perception of music. She describes in detail how the students' descriptions of the music represent their perception of it. Bamberger illustrates how differently students may perceive a rhythm by setting up an imaginary dialogue between two students: one who stresses the metric features of the rhythm, the other who emphasizes the functional aspects of the rhythmic structure. Through their exchange of ideas, both students gain a new dimension in the ways in which music may be heard. Bamberger urges teachers to utilize the differences in the students' descriptions as a tool for broadening the students' understanding of music.

Moreover, Bamberger encourages teachers to see *behind* what, at first, may seem wrong in the students' descriptions, and to seek instead "the logic" reflected in their descriptions. For Bamberger, hearing music is a process of instant perceptual problem solving, and hearing in a new way is learning to enrich one's own understanding of music, to perceive in a new way (see Bamberger, 1978, 1981, 1990, 1991, and Bamberger and Schön, 1991, for a description of these theories and their pedagogical applications, and Bamberger, 1986, for a discussion on the development of musically gifted children).

### REFERENCES

Bamberger, J. (1978). Intuitive and formal musical knowing. In S. Madeja (Ed.), *The arts cognition and basic skills.* St. Louis: CEMREL.

Bamberger, J. (1981). Revisiting children's descriptions of simple rhythms: A function for reflection-in-action. In S. Strauss (Ed.), *U-Shaped behavioral growth.* New York: Academic Press.

Bamberger, J. (1986). Cognitive issues in the development of musically gifted children. In R. J. Sternberg & J. E. Davidson (Eds.), *New conceptions of giftedness.* ambridge: Cambridge University Press.

Bamberger, J. (1990). The laboratory for making things: Developing multiple representations of knowledge. In D. A. Schön (Ed.), *The reflective turn.* New York: Teachers College Press.

Bamberger, J. (1991). *The mind behind the musical ear.* Cambridge, MA: Harvard University Press.

Bamberger, J., & Schön, D. A. (1991). Learning as reflective conversation with materials. In F. Steier (Ed.), *Research and reflexivity.* London: Sage Publications.

## Knowing How and Knowing About

I begin with a conversation that will most likely sound familiar to many readers and yet its very familiarity continues to puzzle me. It starts with an observation by the speaker that she enjoys listening to music and can make sense of most of what she hears. Upon questioning, she agrees that she can clap simple rhythms, recognize tunes she has heard before, even sing or whistle at least some of them. But she then hastens to add: "But of course, I don't know anything about music."

These remarks are so familiar that, like much that is commonplace, the puzzlements in them pass by unnoticed. But if, for some reason, they catch our attention and we turn back to reflect on them, we may be surprised by what we have heard. Wittgenstein puts it like this:

> The aspects of things that are most important for us are hidden because of their simplicity and familiarity. (One is unable to notice something—because it is always before one's eyes.) The real foundations of his enquiry do not strike a man at all. Unless THAT fact has at some time struck him.—And this means: we fail to be struck by what, once seen, is most striking and most powerful. (Wittgenstein, 1953: 50)

The puzzlement that catches my attention in the conversation is this: How can a person say that she can remember, enjoy, sing, and understand the music she hears everyday and still say that she "doesn't know anything about music"? Is it the difference between knowing how to do something in contrast to knowing about it? Or perhaps the distinction isn't in the "knowing" at all, but in the "it"—the music, itself. Could "music" stand for two different kinds of things—one kind that you sing, dance to, recognize, and enjoy; and another kind that you

"know about"? Or could it be both: when you "know about music," the music itself changes?

The conversation raises, in a quite natural way, questions such as the following: What do we mean by "knowledge," or more specifically, what does it mean to know, to have, or to use musical knowledge? And even if we could come to some agreement about that, we are quickly led on to other questions: How does musical knowledge develop? What, for instance, is the difference between what you or I hear in listening to the same piece of music? Why do even professional musicians so often disagree in their "interpretations" of the same piece of music? And how do we ever come to hear in new ways?

Musicians are more likely to talk to one another about their "hearings" of a piece rather than about knowledge. In rehearsing a string quartet, for instance, the violist might say to the cellist, "How do you hear that last passage?" And the cellist might answer, "Well, the second phrase begins on the downbeat of bar 19." To which the violist might answer, "No wonder we aren't together, I hear the phrase beginning with an upbeat starting in the middle of the previous bar." Or among music theorists you might hear one saying of another's analysis of a piece, "Your hearing just doesn't make sense to me." What, then, is meant by a "hearing" and how can we characterize the differences among them; or, even better, try to elicit and account for the usually tacit underlying assumptions that give rise to these differences?

But what is the evidence for another's hearing? My evidence for such hearings are usually subjects' descriptions of them, albeit in several media of which verbal descriptions is one, drawings and the spatial ordering of pitch-playing objects another. While each of these modes of description has its special qualities and constraints giving each a differing potential for what it might be trying to "say," still they all raise a very knotty problem for the observer: since a hearing is by its nature a necessarily private, internal experience, an external description in whatever mode can provide only impoverished clues to a subject's momentary organizing of a melody or rhythm—that is, to his or her hearing. Further, descriptions are influenced in interesting and often very specific ways by the terms—the "units of description"—that the hearer has available. We need to ask, then, what are the relations, often reciprocal relations since terms also influence hearings, between these units of description and what I will call the hearer's "units of perception"? And since a hearing is, perhaps paradoxically, a silent affair, how can anyone know?

I have addressed some of these questions elsewhere[1] but here I want to focus attention on an argument that must necessarily precede them. My argument is that *a hearing is a performance*; what the hearer seems simply to find in the music is actually a process of instant perceptual problem solving—an active process of sense making something like that evoked by the comments of the painter Ben Shahn: "So one must say that painting is both creative and responsive. It is an intimately communicative affair between the painter and his painting, a conversation back and forth, the painting telling the painter even as it

receives its shape and form" (Shahn, 1957: 49). Like a painting, a hearing, too, is both creative and responsive, a conversation back and forth between the music, as material, and the hearer as he or she shapes its meaning and form in some particular way.

## Making Meaning

In arguing that a hearing is a process of instant perceptual problem solving, I obviously intend to suggest that what we casually call "the mind" is actively engaged in making meaning—that is, in organizing incoming sensory material. And I also want to suggest that this is an ongoing generative process—that we are actively doing this organizing in real time as the sound/time phenomena are occurring "out there." But I do not want to suggest that by "organizing" I mean some kind of "decoding" process, as if the incoming material has already been segmented, and these entities labeled or otherwise symbolically "encoded." Rather, it is exactly because sound/time phenomena do not come already structured, but rather *hold the potential for being structured* that different hearings are possible. As Israel Rosenfield has said, "we perceive the world without labels, and we can label it only when we have decided how its features should be organized" (Rosenfield, 1988: 187).

However, I find Rosenfield's word, "decided," somewhat problematic here, since even that suggests more of an overtly intentional action than I mean to imply with the notion of instant perceptual problem solving. In fact, I suspect, and have some evidence to show, that the processes through which we actively organize incoming pitch/time phenomena are closely linked with the very basic, sentient organizing of our bodies as we move through space and time. For example, it is the *sequence of actions* that we practice and internalize in the process of carrying out familiar activities—most particularly action-paths that we internalize in learning to perform a piece on an instrument—that become our most intimate ways of knowing that piece. I call these internalized action-paths "felt paths."[2]

## Research Methodology and an Example

In searching for research questions that might effectively address some of these larger issues, I have found that the most compelling and also most productive questions emerge from puzzling events that occur in the course of my everyday work with students in the classroom, or in working on a composition in preparation for teaching or performing it. Once I recognize an issue, and when I find it impossible to turn a question or puzzle aside, I am led to the design of more formal experimental situations.

For instance, initial evidence that there are specific differences in the hear-

ings of familiar rhythms emerged first in children's drawings of rhythms, drawings that "happened" quite spontaneously in response to a practical necessity: As part of a composition project, the children in a fourth-grade music class had invented a rhythm they called their "Class Piece." The Class Piece was created during a regular class period and was to be played by all the children as a central part of their composition. But with much work still to be done on the composition, the children noticed that the class period was coming to an end. In response, one child suggested, "We had better put our rhythm on paper so we can remember it tomorrow." In the next 10 minutes, each of the children invented a way of "notating" the piece so he or she could remember it the next day. And, as is often the case with invention, there was an unexpected spin-off: While the drawings were indeed there the next day to help the children remember their Class Piece, subsequent analysis of the drawings suggested surprising but very specific differences among them.

The drawings fell into two types, which I tentatively labeled "figural" and "metric." The differences between the two types rested on the kinds of features and relations the children chose or were able to attend to: Those who made metric drawings, as the term suggests, focused their attention on *measuring* the relative duration of events. Thus, clapped events were classified with respect to "same relative duration"—for instance, all "longs" were drawn with one size shape (large) and all "shorts" were drawn with another size shape (small). Metric drawings, then, were based on a notational strategy that is similar to standard rhythm notation and thus could easily be mapped onto it. In contrast, figural drawings exposed for scrutiny functional aspects of musical structure that are not captured by standard music notation. These children focused their attention on the *grouping of performed events into phrases or figures.* The Class Piece, notated in standard rhythm notation along with prototypical examples of each type, is shown in Figure 5.1.

It was only after the distinction between figural and metric drawings had emerged from the spontaneous drawings of the fourth graders that I was able to design an experiment to test the robustness of that distinction. To do so, I asked 186 children between the ages of 4 and 12 to clap and draw five different rhythms, including the children's Class Piece. The data from this larger experiment confirmed the robustness of the initial distinction. In addition, it was possible to organize the wider range of drawings into a typology that included

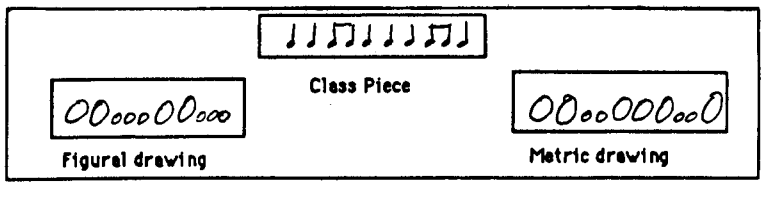



FIGURE 5.1. The Class Piece.

developmental tendencies within the figural/metric distinction (see Bamberger, 1980, 1981, 1988, 1991).

But the robustness of the distinction produced another unanticipated result: In discussing the drawings with the children and later with musically untrained adults who were asked to perform a similar task, it turned out that when subjects were asked to play back a rhythm from another's drawing, each type of drawer found it difficult to make sense of the other's. Figural drawers usually found a metric drawing to be inscrutable or they would suggest that "it must be a different rhythm"; metric drawers typically found a figural drawing simply wrong. (See Hildebrandt & Bamberger, 1979.)

Puzzling over this consistent and somewhat troubling response, it was important to remember that differences in subjects' drawings were reflecting, or were at least clues to, differences in their spontaneous hearings. It was not so surprising, then, that if a description of a hearing ran counter to one's own hearing, it was quite natural to find it difficult to grasp and more difficult to believe.

## A Paradox and a Proposal

Considering that metric drawings give priority to the same kinds of entities as standard rhythm notation (SRN), while figural drawings give priority to entities that are not represented at all by SRN, it was also not surprising that regular users of standard notation (including some, but not all, professional musicians) had difficulty making sense of these figural drawings. At the same time, it is paradoxical that figural drawings express the ability to hear just those structures that professional musicians put great value on in their actual practice, namely, the grouping together of melodic events to form meaningful gestures, motives, or what I have called figures. To do so is often associated with what we call a "musical hearing," specifically with the performer's ability to "shape a phrase." In short, figural drawings show evidence of a musical intelligence we value—the mind behind a "musical ear."

It should now be clear that while figural and metric drawings point to different *kinds* of entities, both figural and metric relations are inherent in and contribute to the structure and coherence of rhythms; indeed, it is the particular intersection of metric and figural groupings that gives a rhythm its unique coherence. (See Cooper & Meyer, 1960; Lerdahl & Jackendoff, 1983.) And yet, the evidence suggests that musicians give a different status to the two aspects of musical structure—the one, metric relations, are considered "objective," the ability to recognize and encode them being associated with what can be taught and with knowledge in the domain; the other, figural relations, are considered "subjective," associated with intuition and musicality, even with the magic of musical "talent," but not easily notated and more difficult to teach and learn.

But in the interest of coming to hear in new ways, it would seem that rather than pitting one kind of structure against the other, or giving each a specially

favored status, what is required is to help individuals learn selectively, knowingly, and at will to shift focus among the possible features and relations that give a composition its unique coherence, and to coordinate these features in new ways. In short, we need to help individuals learn to make *multiple hearings* of even simple rhythms or melodies.

But the difficulty in doing so lies in the nature of the spontaneous hearings we make: a hearing, particularly of an apparent simple rhythm, is usually experienced all at once, and at the moment it is made, it seems immutable—how could it be heard otherwise? And assuming that these ineluctable hearings, like drawings of them, differ with respect to the features one chooses or is able to attend to, then to make a hearing other than one's own, may require a shift, perhaps even a fundamental restructuring of one's focus of attention—giving priority to different features, regrouping, making new boundaries that carve out new entities, and liberating from the meld features that were previously unnoticed, even inaccessible.

## Making Multiple Hearings

How, then, can we help individuals to make multiple hearings? As a first step we need to ask: How can we help individuals *to become aware of their own hearings* of even a simple rhythm so as to recognize which of the possible aspects they are giving priority to in their instant perceptual problem solving? We can then go on to ask: How can we help them to shift their focus of attention so as to participate in and to experience a hearing that differs from their own?

In working with students in my MIT classes, I have found one way that seems to be effective in achieving these goals. These classes usually include students with varied music backgrounds and thus it is predictable that, when asked to notate a rhythm, some will make figural notations and others will make metric notations. It is also predictable that each type of student will defend his or her notation as "right," rejecting others as "wrong." My ploy is to bring the two types of students into direct confrontation with one another, the challenge being that each should try to "put on another's head"—that is, each should try to make sense of and to hear the rhythm in the way his or her counterpart does.

To illustrate this process and to help readers actually live through the sometimes confusing experience of coming to hear in a new way, I have created a dialogue between myself and two college students modeled after those that have actually taken place in these classes. I think of the two students as prototypical figural and metric types: One, whom I have called Met, hears rhythms *met*rically; the other, whom I have called Mot, hears rhythms *mot*ivically (or what I have been more generally calling a figural hearing). While the students initially hold firmly to their respective hearings, I encourage their confrontations and also try to mediate between them. Through this process, the two students eventually succeed in making sense of one another's hearings but only after struggling with

and revealing some of the fundamental assumptions that each of them holds. And since my experience strongly suggests that one or the other of these hearings is that of the reader, it is my hope that in following these conversations, the reader, too, will come to hear in a new way. However, I am also quite aware that to do so may require just as much work on the part of the reader as it does for Met and Mot.

## In Which Met and Mot Come to Understand One Another's Hearings

The dialogue begins after I have asked each of the students to clap and draw the children's Class Piece. They have already looked at one another's drawings and are now in the midst of arguing about them. As shown in Figure 5.2, Mot makes a prototypical figural drawing, while Met makes a prototypical formal drawing. Met is a computer science major who is very forthright in his opinions. Met took clarinet lessons in elementary school and currently plays the guitar. Mot is a humanities major who has had no formal music training and is rather easily intimidated by Met.

MOT (*to Met*): I don't see how you can say that my third little circle should be a big one or a long. With a big circle there, it doesn't look the way it feels when you clap it.

MET: It just seems obvious to me that your third little circle has to be a big one—at least if the circles are supposed to stand for longs and shorts.

MOT: Of course they are. . . .

MET: O.K. But then it's very difficult for me to understand the rhythm the way you've drawn it.

JEANNE: Let me try to help. What I think Mot is paying attention to, and what I see in her drawing, are the little motives or groups of claps inside each of the larger repeated rhythms.

MOT: Yes, that's right. Inside of each larger repeated pattern there are two groups of claps—you can see them quite clearly with my two big circles and then three little ones. And in between the two larger patterns I sort of take a breath and start again. I could make the grouping even more clear with a drawing (Figure 5.3).

JEANNE: That's good. Now could you clap the rhythm again and while

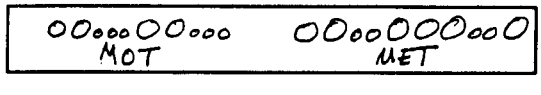



FIGURE 5.2. Mot's figural drawing and Met's metric drawing.

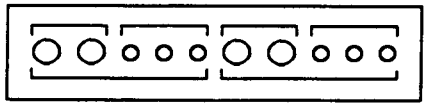


FIGURE 5.3. Mot's grouping structure for the Class Piece.

you are clapping say some numbers that seem to go along with your clapping?

MOT: Sure. It would be (*clapping the class piece*): 1 2 1-2-3, 1 2 1-2-3.

MET: But that's wrong again!

JEANNE: Please try to be a little patient, Met. Now, Mot, would you put the numbers you just said under the circles in your drawing so that we can see just what you mean? (*Mot adds numbers to her drawing; see Figure 5.4*).

JEANNE: O.K. Now try something else. Can you clap just the rhythm for the tune "Twinkle, twinkle, little star; How I wonder what you are" and, while you're clapping, say some numbers that seem to go along with your clapping?

MOT: (*clapping and counting*):

1 2 3 4 5 6 7 1 2 3 4 5 6 7
(Twin-kle twin-kle lit-tle star, How I won-der what you are.)

MET: Oh my!

JEANNE: All right, Met, so how would you count it?

MET: It's obvious. (*Met claps and counts*):

1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
(Twin-kle twin-kle lit-tle star, How I won-der what you are. )

JEANNE: Does that make any sense to you, Mot?

MOT: Not really. I especially don't understand why Met keeps counting up to 4 and then starting over again with 1.

JEANNE: Well, let's leave that problem for the moment. Could you draw a picture of the rhythm of Twinkle like you did for the Class Piece, Mot, and then put in numbers just like you said them when you were clapping?

MOT: I think so. (*She draws and numbers her graphics; Figure 5.5*).

MET: Now you're getting some place! Did you notice that you left a space between the two phrases, I mean between the words "star" and "How I"?

MOT: Of course. That's where you sort of inhale, take a breath. Like I

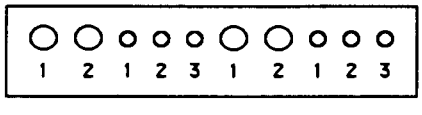



FIGURE 5.4. Mot adds numbers to her drawing of the Class Piece.

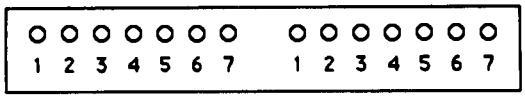



FIGURE 5.5. Mot adds numbers to her drawing of "Twinkle, Twinkle, Little Star."

said about the Class Piece, you take a breath and then start again.

MET: Exactly. But you have to *count that breath*, too—that stop before you start again.

MOT: What do you mean?

MET: Well if I count up the way you were counting up, then it would be like this: (*Met writes his numbers above Mot's drawing; Figure 5.6*).

MOT: That's interesting! You mean it's like there's a ghost beat there (*she points to the space where Met has written an* "8"). You don't actually sing or clap the ghost beat, but it's there anyhow.

MET: A ghost beat; I really like that. You see, time never stops, and the beats keep right on going, too, so you have to keep right on counting, right through the breath, even though you're not making any clap there.

MOT: Why don't we actually put the ghost beat in—the one you call "8." (*Mot draws her ghost beat; Figure 5.7*).

MET: You got it.

MOT: And now I see how you got your "8." Where I counted up to 7 and stopped, you counted up to 8 because you counted the ghost beat, too. So the ghost beat is the stop which sort of isn't a stop after all. But still, it is a stop, you know. I'm getting confused again.

JEANNE: But confusions need to be nurtured, too; they are often the necessary first step to learning anything new. Mot, would you mind clapping the Class Piece, again? And this time pay attention especially to the breath you talked about between the two larger figures.

MOT (*clapping the class piece again*): Hmm. Yes, it seems like there's a space there, too; I mean between the two larger figures. Or I guess you could call it a space-of-time. How come I didn't notice that space before?

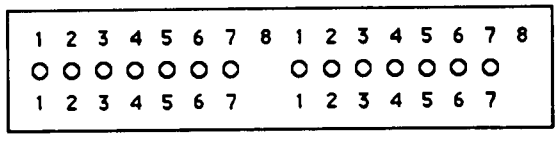



FIGURE 5.6. Met's numbering for "Twinkle, Twinkle, Little Star."

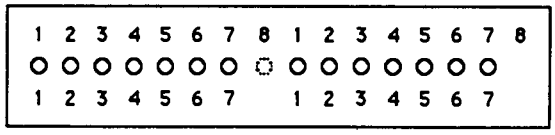



FIGURE 5.7. Mot's "ghost beat" in "Twinkle, Twinkle, Little Star."

JEANNE: Well, what do you think?

MOT: I seem to come to a boundary. I mean, when I clap the rhythm it's like I'm going along heading for a goal. For instance, I'm going along the path inside the inner fast figure till I come to the goal; and then I'm just inside the next figure, going along that one. I never paid attention to what was going on in between the two larger figures. In fact, there really wasn't any in between; I was simply in one figure and then I was in the other. The rhythm of Twinkle made me realize there is an in between. I sort of get it, but I don't.

JEANNE: That's quite understandable; it's hard to pay attention to all the different things that are going on: the claps you're making as you move along anticipating the goal or boundary; then arriving but also going across the boundary crossing. I think the way you talked about being inside one figure, moving along the path to the goal and then inside the next one, really captures what it feels like when you're actually clapping a rhythm.

MET: Sounds pretty strange to me. The rhythm is just there. . . .

MOT: Wait, I just realized something else.

MET: What's that?

MOT: I could put a ghost beat into my picture of our rhythm, too. (*Draws the Class Piece adding a "ghost beat" between the two larger figures; Figure* 5.8).

MET: That's great!

JEANNE: Do you think you could clap the Class Piece and actually clap the ghost beat, too, Mot?

MOT: I can try. (*Claps the Class Piece and includes the "ghost beat" as in Figure* 5.8) It seems like you almost expect that ghost beat to be there, but instead there's nothing.

MET: There isn't nothing. If there were nothing, you would clap it wrong. In fact, that's interesting—you clap it right, but you draw it wrong. Think about it this way: The *breath* you were talking about exactly takes up the space, I mean time, of your ghost beat.

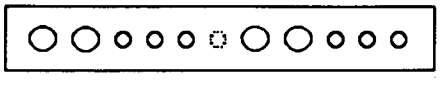


FIGURE 5.8. Mot's "ghost beat" in the Class Piece.

JEANNE: Good point, Met. And that suggests another question: Mot, how many faster claps would you say there are for each slower one?

MOT: Well, I think the little ones go twice as fast as the big ones. So I guess there are two for one. In fact, you could say there is a ghost beat inside of every big circle, too.

JEANNE: Could you explain that a little more?

MOT: It's hard to explain. . . . It seems like there's always a silence, a space, between every clap sound and I think the time of the actual clap sounds is always the same. So it must be the space-of-time between clap sounds that makes the difference between longs and shorts. Gosh this is getting so detailed!

JEANNE: But getting down into the details is useful, provided you can integrate what you've learned down there into the larger design when you climb back up again.

MET: Anyhow, what we are calling the length of a clap is really the clap sound and the gap before the next clap sound.

MOT: Yes, the length of the clap is really a package that includes both the clap and the gap. But when you're actually clapping, you don't think about that at all. With the first two claps you set a natural pace, and then you double it . . . or halve it. . . . Which is it, anyhow?

MET: When you start the faster claps, you double the speed but halve the time from one clap to the next.

MOT: This is beginning to sound like physics, not music.

MET: Not really. But now it should be obvious that if there is a clap and a ghost beat inside of every big circle like you said, then your fifth short clap, together with your ghost beat, will make up one longer clap. So the fifth clap that you've drawn as a little circle together with the gap is a slow clap so it should be a big circle—it's just the same as the slow sixth clap. (*He draws*; *Figure* 5.9).

MOT: But now that just doesn't look right. I can't see the larger repeated patterns any more. And besides, the fifth clap isn't just like the sixth clap. The fifth clap is an *ending*, the sixth clap is a *beginning*, it starts the repetition, and that's not the same at all!

MET: Well, that's all beside the point.

JEANNE: No, Met, that is the point! You see, Mot is saying something very important. It's true that both the fifth clap and the sixth clap are the same if what you're paying attention to is simply measuring the time from one clap to the next. But that's not all there is to rhythms. Did you notice, for instance, that Mot talked about "beginnings" and "endings"?

MET: I did, but . . .

JEANNE: What do you make of that?

MET: Well, I guess you can't have a beginning and an ending without

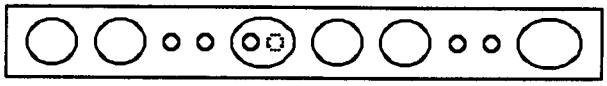


FIGURE 5.9. The fifth clap together with the "ghost beat."

something happening that makes a start and stop. So Mot's first little circle or faster clap *starts off a new thing*—what you would call a new figure; and the last of her little circles (that fifth clap, again) ends the figure, or more like stops it.

MOT: Yes, and that's exactly why your big circle for the fifth clap gets me so upset. If you draw the fifth clap as a big circle, it makes the fifth clap look like the beginning of a three-clap slower figure when really the fifth clap is the ending of the previous faster figure.

MET: But we're not talking about how things *look*. I think we're getting off the track; we're supposed to be talking about time and motion, not about how things look standing still in space. Now I'm getting confused.

JEANNE: That's a beginning, too, Met. In fact, what you just said is the beginning of a real philosophical insight, so just hang onto your confusion for a bit. What do you think makes an event function as the beginning or the ending of a figure?

MOT: I've been wondering that, too. It seems like the function of a clap—whether you hear it as beginning or ending a motive—depends on where it happens—what comes before and after it. For instance, when I start clapping the rhythm, I sort of set the normal pace with the first two claps. So when I start going faster with the third clap, or really when I go from the third to the fourth clap, it feels like the beginning of something new. That new pace sets off the new figure.

MET: And I suppose the next one, the fifth clap that we keep arguing about, stops that run of claps. You get going at this new faster pace and when you get to the fifth clap it's longer than what you expect.

MOT: Well, I'd say it's because what you expect doesn't happen.

MET: I don't get it.

MOT: I mean, you're going along faster, and then you're stopped short because nothing happens. All right, I know, Met, we've been through this before; it's the gap, the ghost beat that's there but you don't actually play it. But that silent ghost beat is more like a comma at the end of a clause. So I guess another reason I made a small circle for the fifth clap is because I was drawing the claps I actually made and not the silent punctuation mark. After all, when we're reading out loud, we don't say commas either; we

just make a little pause and that helps you to group the words so they make sense. Anyhow, just because the fifth clap has the same time as the sixth clap that comes next, they're still not the same.

JEANNE: Now, let's see if I've got it straight. What I hear you saying, Mot, is that even if two events have the same time value, they can still be different depending on where they happen. To use your example, if there's a change to a faster pace, that will set off a new little figure. Then you are moving along inside the new figure at that faster pace, expecting that pace to keep going. But when that pace changes because there is a silence where you might expect another event to occur if the faster pace continued, the new figure is brought to a halt. And that's what makes the boundary.

MOT: That's right.

JEANNE: But Met wants to make things simpler.

MET: Exactly. Since the fifth clap and the gap after it together make a long, it's a long event. What's the problem?

MOT: But if you only pay attention to how *long* an event lasts, never mind where or when, then you'll miss the boundaries altogether. And you will also miss the functional differences between beginning and ending events.

MET: That's funny. I never thought of it before: In a rhythm "where" and "when" mean the same thing.

MOT: That's true. Anyhow, a long after a run of shorts just isn't the same as a long after a long.

MET: Wait a minute, are you trying to tell me that the same two claps can be both the same and different at the same time?

MOT: Yes, depending on where they happen and what you're paying attention to.

JEANNE: Well put, Mot. And that turns out to be very important when you're performing a real piece of music. For instance, even though two notes have the same time value—let's say they both look like this in standard music notation (*she draws* ♩ ♩), you might very well play them differently exactly in order to express a beginning in contrast to an ending of a motive or a phrase. And the difference might involve, among other things, just how long you sustain a note. That is, inside the same sound-silence package, what we usually call the same notated duration, there can be a different relation between the sound and the silence. But when you're really playing a piece, you don't think about it at that level of detail; if you did, you'd feel paralyzed. It all becomes part of expressing the meaning of the piece, including its motives and phrases.

But the funny thing is, that while conventional music nota-

tion in its barest form is very good at showing you the time value of events in relation to one another, it really doesn't tell you anything about beginnings and endings of phrases or figures. That's why violinists, for instance, spend so much time and thought experimenting with bowings and fingerings, and clarinetists worry so much about where to take a breath—they're trying to find the best way to project these important differences between notes that look the same on the page. For instance, in some of Bach's suites for violin solo, practically all the note values are the same. Can you imagine what that would sound like if the violinist wasn't paying attention to the differences in the functions of events that otherwise look the same timewise? Composers, or sometimes editors in the case of Bach, try to help by adding various kinds of so-called expressive markings—slurs, or staccato marks like these (*she draws; Figure 5.10*).

But still, if you strictly follow the rule "what you see is what you get," you'll hear a pretty boring and very unmusical performance. I remember Louis Krasner, the violinist who commissioned the Berg Violin Concerto, saying to a student, "Just forget the score and play the music." But to get back to where we were, yes, two events can be both the same and different depending on the context and on what you're paying attention to.

MET: That reminds me of a picture I saw in a book on vision, once. When you look at it, you see this same shape in two different ways depending on what's next to it (Figure 5.11). At first it seems like you have no choice—you see two different shapes, a diamond and a square. But if you really pay attention to the geometric properties—just stare at the angles and the sides of one shape at a time—of course you can see that they're the same. So, yes, you could say they're both the same and different depending on what you're paying attention to.

MOT: That's a great example, Met.

MET: But you know, we keep shifting back and forth between time and space, and that makes me think of something else. Talking about performers and experimenting, let's try one. You, Mot, pull this piece of paper slowly across the table. While you're doing that, I'll move my pencil up and down in one place, tapping out the rhythm over your moving paper.

(*Mot and Met make the experiment. It leaves this trace; Figure 5.12*).

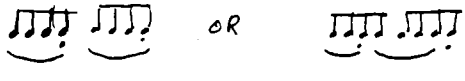



FIGURE 5.10. Slurs as expressive markings.

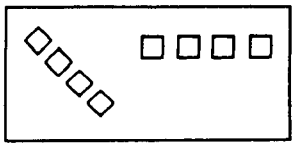


FIGURE 5.11. The same shape can look different.

MOT: That looks fine. You can see the repeated big patterns and the little inner figures, too. How come that works?

MET: Well, because we really made time become space. The trace left behind shows how continuous time or motion—the paper moving across the table—was marked off by my taps. That's the way the old piano rolls worked, and the first mechanized looms, too. First the paper was punched out as someone played—much like we just marked off the moving paper. Then, whenever the sensing mechanism moved across a hole in the paper piano roll, a piano key would be activated. So spatial relations turn into time relations or rhythms.

MOT: I guess I understand, but I keep getting confused between the dots and the spaces between the dots.

MET: Well, just reverse the process. Run your hand at a steady pace over the dots we made and every time you come to a dot make a click or something (*Mot tries it.*)

MOT: It works, all right. And I can really see the faster inner figures, too; the run of three faster claps stands out because there's a bigger space on either side of it—that's really a space-of-time, of course. In fact, what you see is the change of pace that actually generates the boundaries of figures.

MET: And putting it that way helps me to understand your drawing better, Mot, and to hear what you mean by figures, too.

MOT: Well, I'm glad of that. But what about the last clap—you could just go on forever.

MET: As a matter of fact, that's been bothering me all along. You can't know how long the last clap is because it always takes two to tango. I mean, seriously, you always need two claps to mark a time interval. It's your beginning and ending problem, in a way, but on a smaller scale: a beginning by itself, one mark or one clap, like you said, can go on forever; you need another one to stop it. But then, of course, you wouldn't know how long that one was.

JEANNE: This is getting pretty metaphysical. I really liked your experiment, Met, and I especially like the trace it left behind. It's a new notation that I think might have real possibilities.



FIGURE 5.12. The trace left by Met's experiment.

MET: By the way, you said I'd made some kind of philosophical insight back there. Aren't you going to tell me what it was?

JEANNE: Actually, the two of you sort of acted it all out just now. It was about the problems we make for ourselves when we try to describe actions moving through time. We make descriptions of moving things on paper in order to make them stand still, but then we end up with all these issues and confusions about time and space—like the ones that you've both noticed as we've been going along. When you learn how to make motion stand still like you've just been doing through your drawings of rhythms, you gain real power because then you can see, all at one time, what we necessarily experience as fleeting, evanescent, gone.

For instance, it was only after Mot had made her drawing of Twinkle that she was able to invent her ghost beat. Once the rhythm was held still in paper-space, then she could see the bigger space between the two phrases. And once seen, she could also recognize it as really a space-of-time. Only then could she catch on to what Met was saying, even if not totally agreeing with him. Making these static, spatial descriptions solves the problems of time and motion that Heraclitus might have put this way: "You can never put your foot in the same river once."

The trouble is, we are able to find ways of making actions and motion stand still on paper through the invention of a notation system, and that is very useful, even sometimes profound. But once it is made and deeply internalized through continuous use, we come to believe in the notation so thoroughly that we actually experience whatever it is—clapping, wheels turning, drummers drumming—through the lens of the notation, itself. So we go from Heraclitus's problem to taking literally what-you-see-is-what-you've-got.

MET: This has all been more interesting than I thought it was going to be at the beginning. And I think I do understand, now, what you were showing in your drawing, Mot.

MOT: Thank goodness. Things got better once you calmed down, Met. And I have to say I learned a lot, too. In spite of your unpleasantness at the beginning, and in spite of feeling like my head was swirling around some of the time, you helped me get at things I probably would never have gotten to by myself.

## Comments on our Dialogue

While Met and Mot begin their conversation in what seems like irreconcilable conflict, by the end they have clearly come to a resolution of their sometimes

contentious disputes. But interestingly, they have done so without either of them giving up the hearings expressed in their initial drawings. Instead, by actively confronting one another, each helped the other to notice features and relations of the rhythm that had been essentially inaccessible at the outset. In short, each has succeeded in making the hearing of the other but without either of them giving up his or her own.

And in the course of this educational resolution of their figural/formal disputes, Mot also answers a question that has remained quietly waiting beneath all the discussion of figural grouping: What are the temporal relationships that *generate the boundaries* of figures that figural drawings describe? Through Met's experiment—Mot pulling the paper while Met taps out the rhythm with his pencil—Mot is able to see that "the run of three faster claps stands out because there's a bigger space on either side of it—that's really a space-of-time, of course." In other words, it is a *change of pace* that generates the boundaries of figures. And in summarizing her comments, I point out that if there is a change from a slower pace to a faster pace, it will set off a new little figure; and within that faster pace if there is an event of longer duration, that will bring the new figure to a halt, thus creating a boundary.

Such contrasts in the prevailing rate of events can suffice in accounting for figural boundary making within these musically minimal examples—in particular, where events differ only with respect to their time values. But an account that may suffice for brief rhythms played on a nonpitch instrument (claps) and where there is not even any change in loudness will hardly suffice at the moment we consider "real compositions" such as a Beethoven piano sonata movement or a movement from a Bach partita with their much richer network of intersecting, sometimes conflicting, dimensions. Indeed, the issue of what generates grouping boundaries, and how hearings may differ with respect to the groupings heard, is a vexing and fascinating topic to which a number of music theorists have given their attention in recent years. (See Cooper and Meyer, 1960; Lerdahl and Jackendoff, 1983; Bamberger, forthcoming.) And yet, musically impoverished examples such as the Class Piece are useful because they can help to make clear just what kinds of questions we need to ask, what kinds of clues we need to look for, and how seemingly irreconcilable conflicts among disparate hearings might be resolved.

## Conclusions

In the dialogue between Met, Mot, and myself, I have shown an instance of just how such clues to hearings can emerge, and how, through recognizing them, hearings can develop and change as each participant brings into existence features and relations that were inaccessible at the outset. For example, Mot's initial drawing provides clues that her attention is focused on the grouping of adjacent events as these form figures. And despite Met's objections, she seems unable to

recognize, or is at least inattentive to, the shared proportional time values of some events. Met, recognizing these clues, helps Mot to hear that "time keeps on going" and that she needs to count the time "in between" clapped events, as well. This leads Mot to invent her "ghost beat," which, in turn, helps her to bring into existence, so to speak, an entity that had simply not existed for her before. And once seen, this new entity also helps Mot to understand why Met insists that the disputed fifth clap in the Class Piece should be a big circle instead of a small one—even though she doesn't like "the way it looks."

Met's initial drawing shows us clues to his singular focus on the measured time properties of events—all events that share the same event time are for him simply the same. Despite Mot's repeated attempts, she is at first unable to help Met shift his focus so as to hear her figures. But as she recognizes a critical aspect of her own hearing, Mot is able finally to help Met, too. Pointing out that even though two events may share the same time value, their *function* can change depending on the context in which they are embedded, Met is finally able to hear and appreciate the significance of the figural groupings within the rhythm. And in doing so, he too brings into existence an aspect of rhythms that had not existed for him before.

Through this process of reflective confrontation, then, new features emerge, making it possible for both students to include in their hearings new aspects of even these trivial rhythms. And it is this "liberation" of new features that eventually enables Met and Mot to resolve their disputes. With each having access to the preferred features of the other's hearing, they are able to agree that two events can be both the same and different *depending on what you are paying attention to:* Specifically, even though two events may be the same in that they *share the same event time* (e.g., claps 5 and 6), they can still be different with respect to their *figural function*—"where and when they happen" and what role they play as the rhythm unfolds in performance. Thus claps 5 and 6 are indeed the same with respect to their time values, but they still differ in their functions—one an ending, the other a beginning.

This, then, is an example of individuals coming to hear in new ways. Through the work of the two students, I have tried to illustrate that to make a new hearing we must confront those intimate, silent conversations back and forth through which we make and shape the meaning and form we seem simply and immutably to find in the musical phenomena "out there." I have also tried to show that one way to do that is to confront and perturb the hearings of others so as to become aware of *possible* aspects of a rhythm. For in that way we learn to recognize just those aspects to which we are currently giving privileged attention in making our own hearings. Experimenting with one another's hearings, then, we can learn to shift focus at will among the possible features and relations that give a rhythm, or even a moment in a complex composition, its unique coherence.

And through this process, knowing how and knowing about develop a new, *reciprocal* relation to one another in contrast to the adversary relationship I

proposed at the outset: Interrogating what it is you know how to *do* becomes the means for learning to know *about* the music you can sing, clap, or dance to. Thus, instead of knowing about being distanced in and through the privileged languages of others, knowing about becomes a useful way of informing what is most intimate and most familiar. Together, then, knowing how and knowing about, in reflective interaction with one another, help us to hear in new ways, at the same time giving us the power to choose selectively and knowingly among the multiple possible hearings that we are able to make—depending on where, when, and what we want to use them for.

ACKNOWLEDGMENTS

This chapter includes versions of material covered more completely in my book *The Mind Behind the Musical Ear*, Harvard University Press, Cambridge, 1991.

NOTES

1. See Bamberger, 1991.
2. For more on the importance of sequential actions see Bartlett (1932), Lashley (1951), Gardner (1973), and Bamberger (1991).

REFERENCES

Bamberger, J. (1980). Cognitive structuring in the apprehension and description of simple rhythms. *Archives de Psychology, XLVLII*, 171–99.

Bamberger, J. (1981). Revisiting children's descriptions of simple rhythms: A function for reflection-in-action. In S. Strauss (Ed.), *U-shaped behavioral growth*. New York: Academic Press.

Bamberger, J. (1986). Cognitive issues in the development of musically gifted children. In R. J. Sternberg & J. E. Davidson (Eds.), *New conceptions of giftedness*. Cambridge, England: Cambridge University Press.

Bamberger, J. (1988). Les structurations cognitives de l'apprehension et de la notation de rhythmes simples. In Hermine Sinclair (Ed.), *Notations*. Paris: Presses Universitaires de France.

Bamberger, J. (1991). *The mind behind the musical ear*. Cambridge: Harvard University Press.

Bamberger, J. (forthcoming). *Parables of mind and music making*. New York: Schirmer Books.

Bartlett, F. C. (1932). *Remembering*. Cambridge: Cambridge University Press.

Cooper, G., & Meyer, L. B. (1960). *The rhythmic structure of music*. Chicago: University of Chicago Press.

Gardner, H. (1973). *The arts and human development*. New York: Wiley.

Hildebrandt, C., & Bamberger, J. (1979). Claps and gaps. Unpublished manuscript.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. P. Jeffress (Ed.), *Cerebral mechanisms in behavior: The Hixon symposium*. New York: Wiley.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge: MIT Press.

Rosenfield, I. (1988). *The invention of memory.* New York: Basic Books.

Shahn, B. (1957). *The shape of content.* Cambridge: Harvard University Press.

Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, Trans.). New York: Macmillan.

# 6

# Music Performance: Expression and the Development of Excellence

JOHN A. SLOBODA

## Introduction

Music performance is a skill. As with any skill, there are two main psychological questions: What is the skilled performer like (the nature of skill) and how did the skilled performer get to be skilled (the acquisition of skill)? In the last two decades psychologists have given these questions intensive study (e.g., Anderson, 1981; Chi, Glaser, & Farr, 1988; Ericsson & Smith, 1991).

From studying a very wide range of skills (e.g., chess, sport, medical diagnosis, mathematics), a number of general conclusions have emerged. First, skill almost always depends on the ability of the performer to detect pattern and structure in the material, and to conceive the activity in terms of these patterns. Second, the level of skill is almost entirely dependent on the amount of relevant practice undertaken. In other words, learning to detect patterns and structures takes time, so the more time you have spent on a skill, the better you are at it. Third, as skills become practiced and fluent, they tend also to become "automatic"; that is, the details of how they are executed disappear from the conscious awareness of the performer. Fourth, because patterns and structures tend to be specific to a particular domain, there is little transfer between skills. A person can become very good at one skill (e.g., chess) while remaining very poor at another (piano performance). Partly because of this, some skills seem to be independent of general intelligence, and the psychological literature contains numerous examples of people with low general IQ who nonetheless excel at particular cognitive tasks (e.g., Ceci, 1990; Howe, 1989).

All these conclusions are true of music as well. The present chapter concentrates on the evidence for the first three of the above conclusions. For

discussion of the fourth conclusion, including some fascinating studies of musical skill in the absence of high general intellect, the reader is referred to Howe (1991) and Miller (1989).

The chapter concentrates on a key aspect of music performance, that is, the ability to play notated solo compositions expressively. For a discussion of wider aspects of performance, including improvisation, ensemble playing, sight-reading, and performance anxiety, the reader is referred to Sloboda (1982, 1985, 1988).

The final prefatory comment concerns the status of music performance research. Although such research has a long history, it has, in recent decades, been eclipsed by studies of musical perception. This partly reflects a fact of our culture, that many more people engage in music passively (through listening) than they do actively (through performance). But it may also reflect the fact that experimental science likes to control situations in a rather tight way. It is much easier to do this when constructing computer-generated musical sounds than when confronted with a free-performance situation. Nonetheless, in the past decade, performance studies have reestablished themselves as able to meet the rigors of experimental science while addressing questions of theoretical and practical importance in fairly realistic settings.

## REFERENCES

Anderson, J. R. (Ed.) (1981). *Cognitive skills and their acquisition.* Hillsdale, NJ: Erlbaum.

Ceci, S. J. (1990). *On intelligence . . . more or less: A bioecological treatise on intellectual development.* Englewood Cliffs, NJ: Prentice-Hall.

Chi, M. T. H., Glaser, R., & Farr, M. J. (Eds.) (1988). *The nature of expertise.* Hillsdale, NJ: Erlbaum.

Ericsson, K. A., & Smith, J. (1991). *Toward a general theory of expertise: Prospects and limits.* Cambridge: Cambridge University Press.

Howe, M. J. A. (1989). *Fragments of genius: The strange feats of idiots savants.* London: Routledge.

Howe, M. J. A. (1991). *The origins of exceptional abilities.* Oxford: Blackwell.

Miller, L. K. (1989). *Musical savants: Exceptional skill in the mentally retarded.* Hillsdale, NJ: Erlbaum.

Sloboda, J. A. (1982). Music performance. In D. Deutsch (Ed.), *The psychology of music.* New York: Academic Press.

Sloboda, J. A. (1985). *The musical mind: The cognitive psychology of music.* London: Oxford University Press.

Sloboda, J. A. (Ed.) (1988). *Generative processes in music: The psychology of performance, improvisation, and composition.* London: Oxford University Press.

## Expressive Performance

### *The Rationality of Expression*

Technical mastery of a difficult composition is undoubtedly an impressive achievement. We can admire such qualities as rapidity of playing, evenness of touch and timing, and large-scale accurate memory. Yet these things, necessary as they are to musical performance, are insufficient to justify the huge effort that goes into producing top-class performers. If that was all there was to performance, we might as well program scores into a computer, have them performed on instruments via a digital interface, record "perfect" performances once and for all, and close our conservatories!

It is well known, however, that computer renditions of familiar compositions sound dull and lifeless, no matter how fast or accurate they are. Human performances gain their lasting interest and appeal from the fact that they go beyond the information contained in the printed score. It was one of the earliest discoveries of the scientific studies of music performance (see Seashore, 1938) that musicians rarely play two apparently identical notes (from the point of view of the notated score) in the same way. Timing, loudness, tone quality, and intonation can be substantially varied, even within the same measure. Such variations together constitute what we usually call "expression." Variations in expression are what make one note-perfect performance different from another note-perfect performance. They explain why we bother to hear different people play the same piece. They explain why it is worth training successive generations of performers to repeat the same classical repertoire over and over. New and insightful interpretations are always possible.

Most musicians would agree that assigning expressive features to a performance (i.e., deciding on an interpretation of a composition) is a rational enterprise. That is to say, we agree that it matters which notes we vary and how. We can compare different interpretations and offer reasons why we feel one interpretation to be more appropriate or pleasing than another. This is not to say that there is only one acceptable interpretation of a piece of music. There may be an indefinite number. Rather, it is saying that there are criteria by means of which most musicians could come to a judgment that certain possible interpretations are not appropriate to the composition. Expression is not some kind of random or arbitrary variation that we apply to a performance. The structures in the music constrain the expression. Performance and pedagogical traditions constrain the expression. Dialogues between performers, teachers, judges, and critics formulate and make explicit these constraints and apply them to particular pieces and particular performances.

There is, however, a difference between saying, on the one hand, that an activity is rational and, on the other hand, that it is a matter of explicit conscious

awareness to the performer during performance. On any particular occasion, a performer may be unaware of the exact nature of all the expressive variations that are being used, or may be aware only of some rather high-level description of the nature of the expression (e.g., "jerky") rather than the precise means by which this expressive quality is being produced. This has been demonstrated in several experimental studies. For example, Gabrielsson (1974, 1988) asked performers to produce absolutely strict rhythms, as notated, yet they produced performances that consistently deviated from strictness by making some notes shorter than their notated values, other notes longer. It appears that some types of expressive variations have become so automatic for some performers that they are not usually aware of them. Psychological investigations of performance expression cannot, therefore, proceed purely from verbal accounts given by performers of what they do. We actually have to measure performance parameters with a high degree of accuracy. The development of sophisticated computer-instrument links (e.g., the MIDI interface) has made such studies relatively easy, and the last 15 years has seen a steady growth in performance studies.

Quite a lot of the experimental research on performance has been devoted to showing that expressive variations are, indeed, rational, related to musical structure, and have a clear effect on perception. Such research is, however, nowhere near providing a complete explanation for, or prediction of, every expressive parameter of a performance. Musicians who are afraid that this kind of study will rob them of their art have really grossly overestimated the power of contemporary scientific investigations. Attempts to synthesize expressive performance by the computer application of expressive rules (e.g., Sundberg, 1988) are still rather primitive and pose no strong threat to real performers!

### *Experimental Studies of Expression in Performance*

The first step in establishing the rationality of expressive performance is to demonstrate that expressive variations are *intentional*. There are two kinds of variation present in performance which we would not consider as intentional. The first we could call "noise." Every human performer has variations that reflect such limiting factors as muscular control or time estimation. These limits will create small random fluctuations. Hopefully, practice and experience will reduce the magnitude of these variations, but there will always be a residual "jitter." Because of the random nature of noise we would not expect it to show itself in identical ways in two performances. Therefore, the way to find out what is noise and what is not is to ask a person to repeat the same passage of music with exactly the same interpretation, several times over. Only those expressive variations which show up each time can be counted as intentional.

A number of studies of musical performances by professionals have shown a remarkable consistency in successive performances of the same piece (Seashore, 1938; Shaffer, 1981, 1984; Shaffer, Clarke, & Todd, 1985). For instance, Shaffer (1984) compared three performances by the pianist Penelope Blackie of a

Chopin étude, two of which were given a year after the first. Beat duration is plotted in Figure 6.1, which shows a remarkable consistency, not only in the location of major slowings (measures 20, 28, 53, and 57), but also in the distribution of beat durations within a measure (for instance, in measure 30 the first beat is always longer than the second).

The second kind of "unintentional" performance variation comes about be-

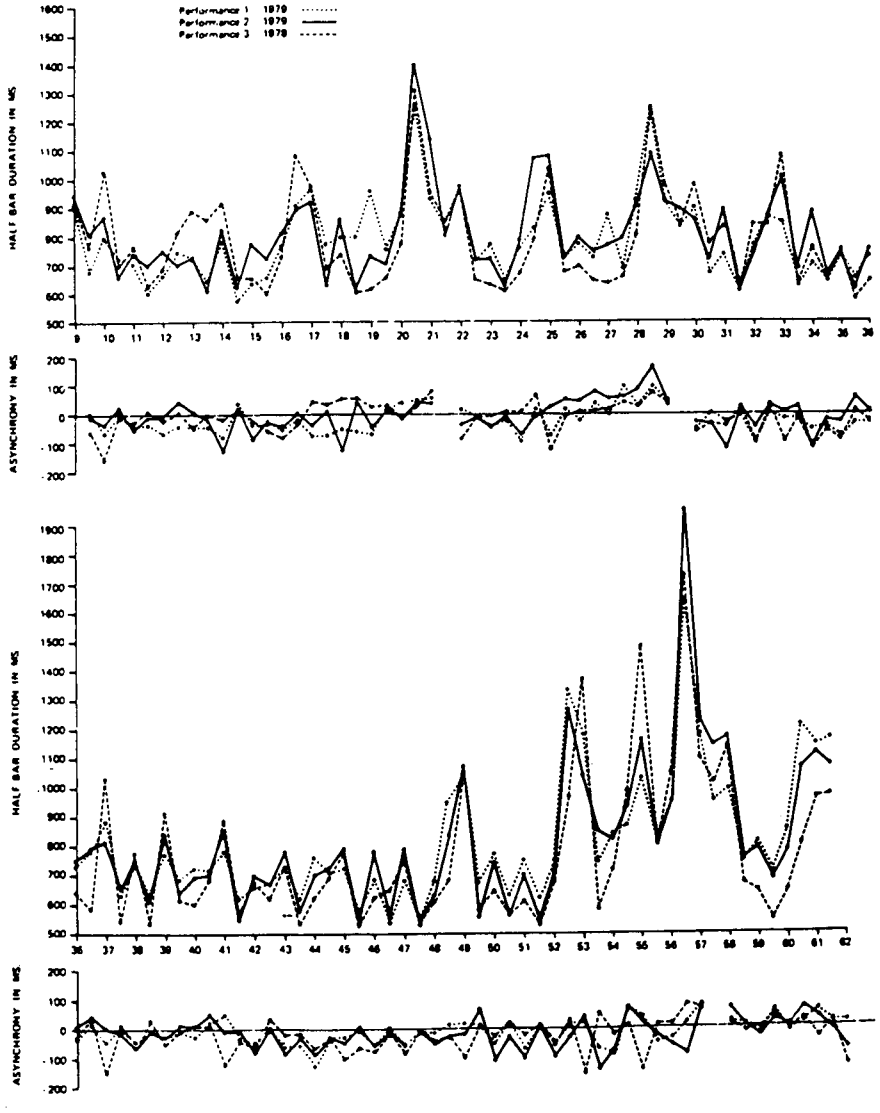



FIGURE 6.1. Beat duration in three performances of a Chopin Etude by the same pianist (Shaffer, 1984).

cause of some technical difficulty. Suppose, for example, that a pianist slowed the performance prior to every large intervallic leap. We would soon suspect that the sheer distance between two notes was the unintended cause of the delay, and would advise the pianist to go away and do some more practice. The most convincing way of showing that this kind of problem is absent is to find instances where exactly the same set of notes is played in two quite different expressive ways. If we can find such instances, then neither expressive pattern can be due to a technical difficulty. Sloboda (1983, 1985) used this method in a study where pianists were asked to provide five consecutive performances at sight of a set of simple melodies. Unknown to the performers, each melody was paired with a "twin" produced by shifting the bar-lines one note along the original melody, leaving the actual notes, their speed, and the fingers with which they were to be played exactly the same. An example of such a pair is shown in Figure 6.2.

Because the members of a pair were always separated by other melodies, no performer realized that the notes were identical. Despite being identical in all respects other than metrical location, the two melodies were played quite differently by all performers. Furthermore, these differences related very closely to the metrical structure. For instance, performers tended to make the first note of a measure significantly louder, longer, and more legato than the notes on either side. These differences were, therefore, truly intentional, even though the details of them may not have been fully available to the conscious awareness of each performer.

The results of this research help to elucidate the reason why a shared language of expressive "rules" does not lead to uniformity in performance. Each pianist's performance of a given melody in this study was quite different from every other pianist's performance although all used the same basic expressive strategies. This is because each pianist varied the extent, degree, and precise location of the variations available. Some pianists made more use of duration parameters, others of loudness. Some made large variations, others made more subtle ones. Some marked every measure, others marked only selected measures, and so on. Even within a fairly tightly defined "rational" expressive system, the potential for individual variation is still immense. There are many "free" parameters, even within a constrained system. We may suppose that the way which different performers "fill in" these free parameters is what is responsible for their unmistakable individual style that permeates all their performances (Sloboda, 1985). Connoisseurs of particular genres can often tell within a few measures of a recording who the performer is.

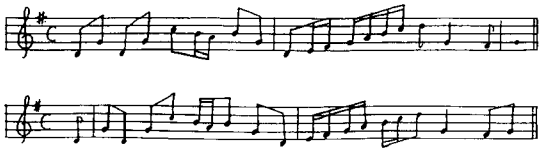


FIGURE 6.2. A sample pair of same-note melodies used in Sloboda's (1983) study.

### *The Effectiveness of Expressive Performance*

It is one thing to show that measurable and intentional expressive variations are present. It is quite another to show that these variations are *effective*. What, in the context of music performance, do we mean by effective? First, differences in expression must be *detectable* by listeners; second, they must be *interpretable*. The first condition is there to ensure that expressive variations are large enough to make a difference to listeners; the second, to ensure that the variations mean something to the listener, at either an intellectual, aesthetic, or affective level.

Psychological studies of expressive musical communication have tended to remain at the level of intellectual judgment. So, for instance, Sloboda (1983) showed that listeners had some ability to correctly judge which member of a twinned pair of melodies a particular performance was representing. Different performers, however, had differing levels of success in conveying meter. This seemed to relate to performance experience. The most experienced pianist in the sample had 40 years of professional playing, and he was the most effective in conveying meter to a panel of listeners. The least experienced player had only (!) been playing for 12 years, and was the least effective communicator of meter. The most effective communicator was both more consistent and more exaggerated in the use of expressive devices. The least effective communicator showed some inconsistencies, with some expressive variations going against the "rules" extracted from the majority of the performances.

Meter is only one small part of what an expressive performance might communicate. Palmer (1988, 1989) in an extended series of studies showed that pianists asked to exaggerate or accentuate a melody tended to make the melody notes precede other notes which were notated as simultaneous. When asked to exaggerate or highlight a particular phrasing, they tended to increase rubato at intended phrase boundaries (similar results have been obtained by Todd, 1985) and increase legato playing within the phrase. In one experiment, pianists were asked to play the same piece in three different ways—as they normally do and then two exaggerated performances which accentuated different melodies or phrasings. Listeners tended to be able to accurately judge which performance was intended. Palmer also showed that it was, indeed, the timing variations that communicated the intended interpretation to the listeners. She resynthesized the performances by computer in such a way that all timing variations were eliminated while the loudness variations of the original were retained. Listeners were poor at identifying expressive intention from these resynthesized performances.

Davidson (1991, 1993) has recently shown that not all of the expressive intent of a performance is necessarily contained within the sound parameters. Her studies demonstrated that the body movements made by performers while playing contribute to the expressivity of the performance as judged by observers.

Movements of wrist, torso, and head tend to be systematic and related to specific structural features of the music. It is as if such body movements draw the observer's attention to particular parts of the music, enhancing their salience. Even when performers are asked to play the music "deadpan" (i.e., without expression), small but detectable body movements still occur. These are of exactly the same type as the movements observed when the performer is asked to play expressively and occur at the same points in the music. Just as in the Gabrielsson (1974) study discussed earlier, these movements seem to have become automatic. Davidson's results raise some interesting questions about the context of musical performance. Do they help explain why live performances are often so much more engaging than recorded ones? Does our society's increasing dependence on audio recordings of music mean that a whole dimension of musical expressivity is being downplayed or abandoned?

### *Summary*

It has become increasingly clear from performance studies of the type described here that expressive variations in music are not simply a matter of arbitrary convention. They actually help listeners assign a structural interpretation to a passage of music that is the one intended by the performer. Many theorists (notably Lerdahl & Jackendoff, 1983; see Chapter 3) believe that an important part of what it is to have understood a piece of music is to have assigned to it a structural description in which related notes are grouped together, and in which different events are assigned different importance or prominence. Lerdahl and Jackendoff propose that this description is arrived at by assessing various cues in the music and then trying to achieve an interpretation that is consistent with them. In some musical passages, all the different cues (rhythm, harmony, articulation) are consistent, and so strongly suggest one structural description. In other cases, there are several possible descriptions, and the performer's role may well be to add expressive cues that will incline the listener toward one or another interpretation. These cues follow from general principles of musical perception. For instance, all other things being equal, notes that are longer than their neighbors will be heard as more prominent. Therefore, if a performer wishes to disambiguate a metrically ambiguous passage, he or she will tend to lengthen the first note of a metrical unit.

## The Development of Excellence in Performance

### *Talent Versus Learning*

There can be few questions which are of more intrinsic interest and pressing practical concern than the question of how music performers acquire high levels of expressive skill. Why is it that musicians differ in their level of expressivity? It

has been common to discuss differences between musicians as differences in "talent" or "musicality." Yet, on their own, these words are simply redescriptions of the phenomenon to be explained. How *is* it that one person is more "musical" than another? The implicit assumption behind such language is that these differences are inborn or innate. Yet there is remarkably little evidence to support such an assumption. Studies by behavior geneticists (see Coon & Carey, 1989) suggest that the impact of environmental variables on musical achievement is unusually high when normal methods of estimating heritability (i.e., twin studies) are used. The validity of such twin studies has, of course, been widely questioned in any case (Ceci, 1990; Rose, Lewontin, & Kamin, 1984).

An alternative perspective, and one which follows rather naturally from the view that expressive performance is a rational activity closely related to the performer's understanding of the musical structure, is that such performance is a cognitive skill that is learned. If this latter perspective is correct, then to call someone exceptionally talented is to assert no more than the fact that such a person has had a certain sort of exceptional life, in which crucial and statistically unusual experiences occurred, allowing the development of skill beyond the norm. As I have argued elsewhere (Sloboda, 1990), exceptionality cannot be observed by its inherent qualities, as it were from within the person. A person is only exceptional in a certain cultural or historical context by virtue of doing better than most other people. For instance, the normal reading ability of a contemporary high-school graduate would have been considered totally exceptional 500 years ago.

Few people would consider it plausible to explain the general increase in reading efficiency by a massive increase in innate ability for reading over the last 500 years. Rather, we would look to broad social changes which turned exceptional experiences into common ones. As intensive exposure to the products of literacy and literacy education became more and more normal, it became possible for increasing numbers of people to undergo the set of experiences that resulted in the level of attainment we consider normal today.

What are the general characteristics of the lives of those people who have attained exceptional levels of musical skill by today's standards? To answer this question, researchers have recently begun to examine elite groups of musicians at various stages of their careers: students at special music schools, conservatory students, successful professionals, and competition winners.

## *The Role of Practice*

Ericsson, Krampe, and Tesch-Romer (1991) recently conducted an extensive study of the practicing history of 40 violinists. Thirty of these were students at the Music Academy of West Berlin; 10 "best" students nominated by their professors as having the potential for careers as international soloists; 10 "good" violinists from the same department; and 10 "teachers" studying the violin in a department (education) with lower performance admission standards. All these groups were

matched for age and sex. The fourth group ("professionals") comprised 10 middle-aged violinists from two symphony orchestras in West Berlin with international reputation.

Each participant undertook extensive interviews, was asked to keep a daily diary of activities for a week, and was asked to estimate the number of hours of practice on the violin each week for each year since he or she had started to play the violin. Figure 6.3 shows the accumulated hours of practice for each group at each age up to 20. The "best" and "professional" violinists show a very similar curve, which suggests that the professors' classification was accurate. By age 20, they had accumulated an average of over 10,000 hours of practicing, compared to some 8,000 for the "good" students, and some 5,000 for the "teachers." Indeed, the "best" violinists had accumulated by the age of 15 the same amount of practice as the "teachers" had by age 20. Even as young as 10, there is a distinct difference between the "best" players and the others.

The reliability of the estimated figures was checked by comparing estimated figures for the current year to the actual figures recorded in the diaries. The concordance was quite high. At the time of the study the "best" students were averaging 30 hours per week of practice, the "good" students 25, and the "teachers" only 10.

Ericsson, Tesch-Romer, and Krampe (1990) point out that accumulated skill through practice has "many of the properties associated often attributed to inherent characteristics such as talent. If, for example, two individuals of the same age differ dramatically in accumulated practice time . . . there is no way for the less trained individual to catch up with the more experienced person, if both maintain the same weekly schedule."

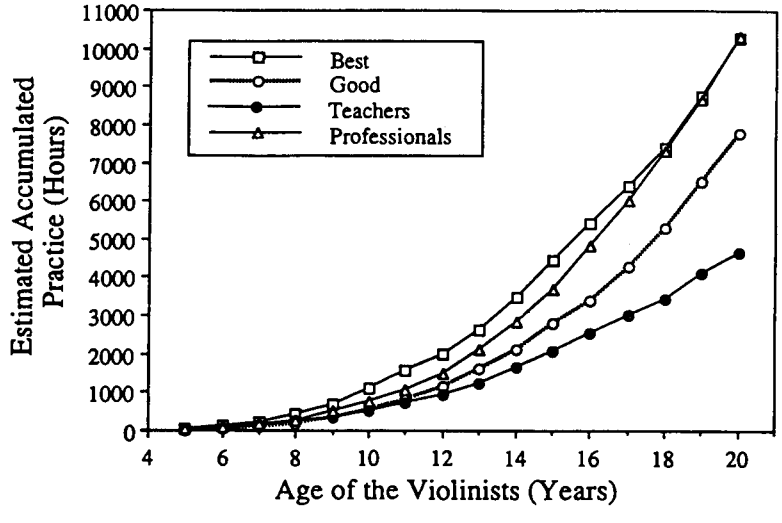



FIGURE 6.3. Accumulated hours of practice in four groups of violinists (Ericsson, Krampe, & Tesch-Romer, 1991).

Where, as in the case of music, access to the highest levels of teaching and coaching is based on competitive hurdles to be overcome in a narrow age window (typically 18 to 20), then clearly the successful musicians will be those who managed to accumulate as much practice as possible by this age. There is now considerable evidence that the highest levels of musical achievement are normally attained by individuals who began systematic study of an instrument before the age of 10 and maintained regular and increasing amounts of practice through the years of adolescence. In Ericsson et al.'s (1991) study, the mean age of the first violin lesson was 7 years. In a study of 21 international pianists in the United States, Sosniak (1985, 1990) found that the mean age of the first lesson was 6. A study by Sloboda and Howe (1991) on 42 adolescent students of a specialist English music school also revealed a mean starting age of 6 (range 3 to 10).

Of course, sheer amount of practice is a rather crude measure from a psychological point of view. The nature of learning will depend on what is done during practice as well as how much of it is done. Surprisingly, there is almost no reliable scientific information on how musicians practice. Gruson (1988) is the only researcher to date who has compared the practicing strategies of expert and novice instrumentalists. She found that experienced musicians showed much more sensitivity to structural units in the music (by, for instance, taking such units in isolation for repeated practice). Novices tended to indulge in more playing through of a whole piece, and repeating individual wrong notes, rather than repeating the sections in which they were embedded. A detailed case study of an individual professional pianist (Miklaszewski, 1989) confirms the importance of identifying structural "fragments" for detailed work. Clearly, since expressive performance is closely related to the detailed structure of a composition, practice that shows awareness of structural factors is more likely to lead to the development of an appropriate expressive repertoire.

Although amount of practice may be a conveniently simple predictor of "talent," it is far from simple to explain how large amounts of practice are initiated and sustained over a long period of a person's life. Ericsson et al. (1990: 113) expresses the problem succinctly:

> Practice activities, which involve careful analysis of one's performance, deliberate fine-tuning and acquisition of new procedures through massive repetition, are of a very different nature from spontaneously enjoyable activities within the domain. Examples of such spontaneously enjoyable activities are the play of children and the "flow state" which is defined by Csikszentmihalyi (1975) as the optimal match between the individual's skill level and the demands of the chosen field, with the individual feeling fully consumed by the activity. In contrast, practice requires deliberate control and effort, which most people, especially children, find unpleasant, particularly as it is extended over longer periods of time.

The issues here are basically ones of motivation and reward. Why do some children start music lessons and then continue effortful involvement through many years, when the "typical" child abandons effort after a few months or years?

### *Early Biographical Precursors of Excellence*

One prevalent "myth" concerning exceptional musicians is that they typically demonstrated very early signs of their exceptionality by, for instance, early accurate singing, early reproduction of melodies on an instrument, exceptional concentration on music, or spontaneous early requests to begin lessons. Evidence from neither Sosniak (1985, 1990) nor Sloboda and Howe (1991) unequivocally confirms this view. Although parents of some children reported such things, the majority of children showed no special or outstanding ability or interest in music significantly before the start of lessons. In Sloboda and Howe's study, most children began instrumental lessons for reasons other than their own choice. In some cases it was a normal part of the school routine; in others the parents had decided it was a good thing for all their children to do, regardless of "ability." Sosniak discovered that it was sometimes not until five or six years after the start of formal instruction that her top pianists began to show the signs that led teachers to believe that a professional career was possible. When they started lessons, all of her sample were perceived as normal and unexceptional.

Sloboda (1989) has provided evidence to suggest that where the motivation for beginning lessons comes from the child, it is more likely to arise from some intense and pleasurable experience of music listening rather than from performance accomplishment. Such experiences motivated their recipients to wish to be able to reproduce these experiences through performance. The study, which asked adults to recall music-related events from their childhoods, suggested that early performance-related activities were very often negatively, rather than positively, motivating, being perceived as sources of anxiety, humiliation, or even physical abuse. It is possible to account for many apparent cases of "unmusicality" through early negative reinforcement, usually administered by a schoolteacher, which inhibited further engagement or effort in music performance.

Even when instrumental learning is not inhibited by negative reinforcement, it is clear that many young people become unhelpfully focused on "playing the right notes," and are so anxious about their own performance that they cease to enjoy the music or become emotionally engaged in it for its own sake. In these contexts it may become rather difficult to learn to play expressively. The tuition in expressive performance given by a teacher will fail to generalize unless a student can experience *why* that expressive device is appropriate through his or her own listening responses. If a young person has never experienced the emotional delight of, say, an enharmonic change, then he or she won't grasp, from the inside, why a slight slowing before the change may heighten the aesthetic effect. The teacher's performance instruction will seem arbitrary, rather than something which is immediately understood and assimilated for use in analogous situations.

Recent data on emotional responses to music (Sloboda, 1991a, 1991b) demonstrate that intense physical concomitants of emotional responses to music (such as tears or shivers down the spine) are linked to specific structures in the music, such as appogiaturas or enharmonic changes. If a child rarely has the opportunity to experience music in the kind of relaxed settings that promote these intense responses, then he or she will never be able to apply appropriate performance expression to such events other than something learned mechanically. The amount of expressive detail in a single performance is so great that it would be very difficult and time consuming to learn it "from scratch" for each piece. Good performers must be able to apply general rules of expression which make intuitive sense to them. Intense, involved, and enjoyable listening experiences are the key to the acquisition of this intuitive "sensibility."

## *The Early Years of Instrumental Learning*

The general conclusion that emerges from the various biographical studies mentioned is that, for those children who later achieve performance excellence, the earliest years of instrumental learning are characterized by an atmosphere of pleasure and enjoyment. In this respect, the first teacher seems to have been crucial. In both Sosniak's (1985) and Sloboda and Howe's (1991) studies, respondents tended to describe their first teachers with such adjectives as warm, friendly, fun, loving, and encouraging. Lessons were enjoyable events to be looked forward to, and it seemed to be the quality of the teacher-pupil relationship, rather than the technical excellence of the teacher, which counted for most. In Sosniak's study first teachers were very often characterized as the "nice old lady down the road."

Two findings from Sloboda and Howe's study extend this "pleasure principle" to the home environment. Their sample contained both outstanding and average students from a specialist music school. It was discovered that the outstanding students generally had parents who were not themselves performing musicians. Average students were more likely to have come from families where one or both parents earned their living as a musician. Second, the outstanding students had accumulated significantly less practice than the average students prior to entry at the specialist. Table 6.1 shows accumulated practice for each group by instrument. It is necessary to present the results by instrument, because an unexpected finding of the study was the fact that most students had studied at least two instruments. The first instrument learned was not necessarily the one on which the student currently excelled. Table 6.1 makes it clear that the differential practice effect shows most strongly for the first instrument studied.

A tentative interpretation of these findings, supported by the more qualitative aspects of the data (see Howe & Sloboda, 1991a, 1991b), is that children of nonmusicians are more likely to be praised for their early accomplishments so that they acquire an early sense of themselves as "musician" and "special." Parents tend to be supportive but nondirective, allowing their children freedom

TABLE 6.1 Accumulated Hours of Practice as a Function of Ability

| | *Ability* | |
|---|---|---|
| *Instrument* | *Exceptional* | *Average* |
| First | 454 | 1,095 |
| Second | 320 | 422 |
| Third | 217 | 62 |
| All | 1,012 | 1,580 |

*Source*: Adapted from Sloboda and Howe, 1991.

to decide whether to do formal practice or more free improvisatory musical play. Musician parents, in contrast, tend to view their children's accomplishments in the light of known standards to be attained. They are less likely to praise early achievements and are more likely to impose a rather rigid or structured practice regime. Such children may become technically adept, but fail to find their own musical motivation and creativity. Staff at specialist music schools are only too aware of the possibility of "adolescent rebellion" among such children, who belatedly realize that they are fulfilling their parents' ambitions rather than their own.

### *The Transition to Commitment*

Practice cannot always be enjoyable, and at some stage in an instrumental career, a young person must make the decision to invest in practice activities beyond what is immediately rewarding. Transitions to new teachers or schools appear to be a vital component of this process. Almost all respondents in the various studies discussed reported a progression through a series of teachers who imposed increasingly demanding requirements on their pupils. Terms like "stretching" or "challenging" tended to be used much more when describing later teachers. Many students at music school came to see their teachers as role models or "heroes." The desire to be like the teacher, or like another eminent performer, is apparent in the motivation of many young performers.

As adolescence approaches, the existence of a reference group for whom high musical achievement is the norm becomes more important. One clear reason why many children drop out of instrumental learning is that their efforts are not valued by their peers (Howe & Sloboda, 1991c). In many environments, including schools, musical achievement is met with indifference, or even mockery. This seems to be particularly so for boys, because other boys brand musical activities as "unmasculine." Specialist music schools can provide a haven for young musicians where, for the first time in their lives, their activities are accepted as normal by their peers (Sloboda & Howe, 1990). Those who do not go

to such schools typically find similar support in specialist music groups such as orchestras.

Because there are so many pressures on children to do other things than devote time to instrumental learning, parents play a crucial role. Despite misgivings about the "ambitious parent," the studies of Sosniak (1985), Manturszewska (1990), and Sloboda and Howe (1991) all confirm the vital importance of a high level of parental support and encouragement. Almost all parents of successful musicians devoted large amounts of time, effort, and money to their children's musical activities. In Sloboda and Howe's (1991) study, the vast majority of parents had regular meetings with instrumental teachers and provided specific incentives to undertake practice. In a large minority of cases this actually involved being present throughout a practice period, helping the child to structure the practice activity. Research currently underway intends to specify the precise differences in parental behavior between successful and unsuccessful musicians.

## Conclusions

This part of the chapter has outlined a picture of exceptional achievement as resulting, not from some unique or rare characteristic of an individual, but from a series of events and experiences that stimulate a statistically rare learning trajectory. Following from, but extending, a framework devised by Ericsson et al. (1990), a three-stage developmental framework may be proposed. In the first stage musical engagement is sustained primarily through pleasure: aesthetic/sensual pleasure in the experience of musical sound; pleasure in rewarding relationships; pleasure in praised achievements. Many nonmusicians seem never to have experienced this stage at the appropriate point in childhood. Some average musicians may also have experienced this stage inadequately, leading to "proficiency without spark." In the studies examined here, this first stage typically encompasses the first three to four years of instrumental study.

In the second stage, musical learning is sustained primarily through a constructed environment in which practice levels may be increased and sustained. This involves some external regulation of activity by parents and teachers, often in the context of specialist environments with like-minded peers. Specific achievement targets (such as examinations) often become prominent motivators. This stage typically coincides with the period of secondary schooling and ends when the individual decides on his or her long-term goals for music, whether to give up serious effort, to remain at an amateur level of proficiency, or to progress to professional levels.

In the third stage, the impetus for musical learning becomes fully internalized to the maturing individual, who will sustain and direct his or her own practice activities independently. Typically, in the highest group of performers, practice activity now predominates over all other forms of life activity. The role of the teacher moves toward that of mentor and guide rather than director. This

stage typically begins in the late teens or early twenties and continues to the point where a professional career is established.

There are, of course, further stages in the life-span of a professional musician (see Manturszewska, 1990), but there is not space in this chapter to describe them. It is more important to point out that the stages described are relative to a particular culture, that of the Western classical music profession. Little serious study has been made of the development of the popular or jazz musician (although see Sloboda, 1991c, for a discussion of the untutored development of Louis Armstrong). What is clear from case studies of jazz musicians and idiots savants (see Miller, 1989; Sloboda, Hermelin, & O'Connor, 1985) is that high levels of intrinsic motivation can sustain learning in the absence of formal tuition provided certain environmental conditions appertain (Sloboda, 1991c).

Finally, it is necessary to point out that there exist many cultures, especially traditional non-Western cultures, where the spread of musical accomplishment in the population is much more even than it is in modern Western culture. In such cultures, participation in musical activity is often more rooted in the whole life and work of society, rather than being something separate (e.g., Blacking, 1976). In the long view, it may turn out to be modern Western society that is unusual in its capacity to inhibit musical development in all but a small minority. Rather than looking within a small number of individuals for signs of genetic superiority in "musicality," we may need to look more closely at the way our culture operates to inhibit the majority of people from becoming the fully expressive musical performers that they could be.

## REFERENCES

Blacking, J. (1976). *How musical is man?* London: Faber.

Ceci, S. J. (1990). *On intelligence . . . more or less: A bioecological treatise on intellectual development.* Englewood Cliffs, NJ: Prentice-Hall.

Coon, H., & Carey, G. (1989). Genetic and environmental determinants of musical ability in twins. *Behavior Genetics, 19,* 183–93.

Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety.* San Francisco: Jossey-Bass.

Davidson, J. W. (1991). *The perception of expressive movements in music performance.* Unpublished doctoral dissertation, City University, London.

Davidson, J. W. (1993). Visual perception of performance manner in the movement of solo musicians. *Psychology of Music, 21,2,* in press.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1991). *The role of deliberate practice in the acquisition of expert performance* (Tech Rep. No. 91-06). Boulder, CO: Institute of Cognitive Sciences, University of Colorado at Boulder.

Ericsson, K. A., Tesch-Romer, C., & Krampe, R. T. (1990). The role of practice and motivation in the acquisition of expert-level performance in real life. In M. J. A. Howe (Ed.), *Encouraging the development of exceptional skills and talents.* Leicester: British Psychological Society.

Gabrielsson, A. (1974). Performance of rhythm patterns. *Scandinavian Journal of Psychology, 15,* 63–72.

Gabrielsson, A. (1988). Timing in music performance and its relations to music experience. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition*. London: Oxford University Press.

Gruson, L. (1988). Rehearsal skill and musical competence: Does practice make perfect? In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition*. London: Oxford University Press.

Howe, M. J. A., & Sloboda, J. A. (1991a). Young musicians' accounts of significant influences in their early lives. 1. The family and the musical background. *British Journal of Music Education, 8*, 39–52.

Howe, M. J. A., & Sloboda, J. A. (1991b). Young musicians' accounts of significant influences in their early lives. 2. Teachers, practicing and performance. *British Journal of Music Education, 8*, 53–63.

Howe, M. J. A., & Sloboda, J. A. (1991c). Problems experienced by young musicians as a result of the failure of other children to value musical accomplishments. *Gifted Education, 8*, 102–11.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Manturszewska, M. (1990). A biographical study of the life-span development of professional musicians. *Psychology of Music, 18*, 112–39.

Miklaszewski, K. (1989). A case study of a pianist preparing a musical performance. *Psychology of Music, 17*, 95–109.

Miller, L. K. (1989). *Musical savants: Exceptional skill in the mentally retarded*. Hillsdale, NJ: Erlbaum.

Palmer, C. (1988). *Timing in skilled music performance*. Unpublished doctoral dissertation, Cornell University, Ithaca, NY.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 331–46.

Rose, S., Lewontin, R. C., & Kamin, L. J. (1984). *Not in our genes: Biology, ideology, and human nature*. London: Penguin Books.

Seashore, C. E. (1938). *The psychology of music*. New York: McGraw-Hill.

Shaffer, L. H. (1981). Performances of Chopin, Bach, and Bartok: Studies in motor programming. *Cognitive Psychology, 13*, 326–76.

Shaffer, L. H. (1984). Timing in solo and duet piano performances. *Quarterly Journal of Experimental Psychology, 36A*, 577–95.

Shaffer, L. H., Clarke, E., & Todd, N. P. (1985). Meter and rhythm in piano playing. *Cognition, 20*, 61–77.

Sloboda, J. A. (1983). The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology, 35A*, 377–96.

Sloboda, J. A. (1985). Expressive skill in two pianists: Style and effectiveness in music performance. *Canadian Journal of Psychology, 39*, 273–93.

Sloboda, J. A. (1989). Music as a language. In F. Wilson & F. Roehmann (Eds.), *Music and child development*. St. Louis: MMB.

Sloboda, J. A. (1990). Musical excellence: How does it develop? In M. J. A. Howe (Ed.), *Encouraging the development of exceptional skills and talents*. Leicester: British Psychological Society.

Sloboda, J. A. (1991a). Music structure and emotional response: Some empirical findings. *Psychology of Music, 19*, 110–20.

Sloboda, J. A. (1991b). Empirical studies of emotional response to music. In M. Riess Jones & S. Holleran (Eds.), *Cognitive bases of musical communication*. Washington, DC: American Psychological Association.

Sloboda, J. A. (1991c). Musical expertise. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits*. Cambridge: Cambridge University Press.

Sloboda, J. A., Hermelin, B., & O'Connor, N. (1985). An exceptional musical memory. *Music Perception, 3*, 155–70.

Sloboda, J. A., & Howe, M. J. A. (1990). Reasons for choosing specialist music schooling: An interview study. *Journal of the Music Masters and Mistresses Association, 25*, 17–20.

Sloboda, J.A., & Howe, M. J. A. (1991). Biographical precursors of musical excellence: An interview study. *Psychology of Music, 19*, 3–21.

Sosniak, L. A. (1985). Learning to be a concert pianist. In B. S. Bloom (Ed.), *Developing talent in young people*. New York: Ballantine Books.

Sosniak, L. A. (1990). The tortoise, the hare, and the development of talent. In M. J. A. Howe (Ed.), *Encouraging the development of exceptional skills and talents*. Leicester: British Psychological Society.

Sundberg, J. (1988). Computer synthesis of music performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition*. London: Oxford University Press.

Todd, N. P. (1985). A model of expressive timing in tonal music. *Music Perception, 3*, 33–58.

# The Perception of Melody, Tonality, Rhythm and Timing

# 7

# Melodic Contour in Hearing and Remembering Melodies

W. JAY DOWLING

## Introduction

In recent years, research on the perception of melody has received much attention. Frequently, the melody is the musical element we most easily recognize, reproduce, and remember because it is perceptually the most salient. A melody is a sequence of single pitches organized as an aesthetic musical whole. In trying to explain how some sets of notes are heard as integrated melodic wholes, psychologists have invoked Gestalt principles (see discussions in Deutsch, 1982a; Dowling & Harwood, 1986; Sloboda, 1985). These are principles of perception that were originally formulated to investigate the perception of visual stimuli (Wertheimer, 1923; see Arnheim, 1969, for application of Gestalt principles to the perception of visual art). Some of the most important of these principles are based on laws that govern proximity, good continuation, and similarity, as illustrated in Figure 1.

If we consider a melody as a line, it is easy to see how principles of perception may be successfully applied in researching melodic patterns. Indeed, in music performance and music teaching there is an emphasis on the "direction," the "shape" of the melodic line. We often speak of a melody and its accompaniment as the "foreground" and the "background." All these are visual analogies that are very pertinent musically.

A melodic contour is a sequence of single pitches that resembles the original intervallic pattern of a melody. The importance of the melodic contour, or shape, is such that it is nearly impossible to recognize a familiar tune if its pitches are played in random octaves on the keyboard: that is, if the original melodic line is broken as illustrated in Figure 2.

The more the pitches of the melody are played randomly among several octaves, the more difficult it is to recognize the melody. Such randomness violates the Gestalt principles of good continuation and proximity. Our recognition of this familiar melody increases as more pitches are played within the

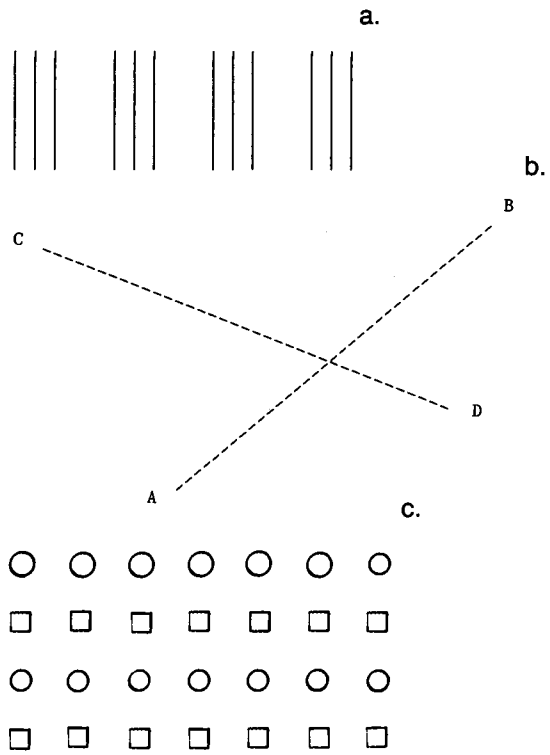



FIGURE 1. Gestalt principles. (A) *Proximity.* We tend to group nearer elements together rather than spacing them apart; therefore, we tend to perceive this diagram as four sets of three lines each rather than a set of 12 lines. (B) *Good continuation.* We tend to perceive that the set of dots starting from point A goes to point B rather than going to point C or D. (C) *Similarity.* We tend to perceive the above as rows of similar elements—namely, rows of squares or circles—rather than perceive columns combining both squares and circles.

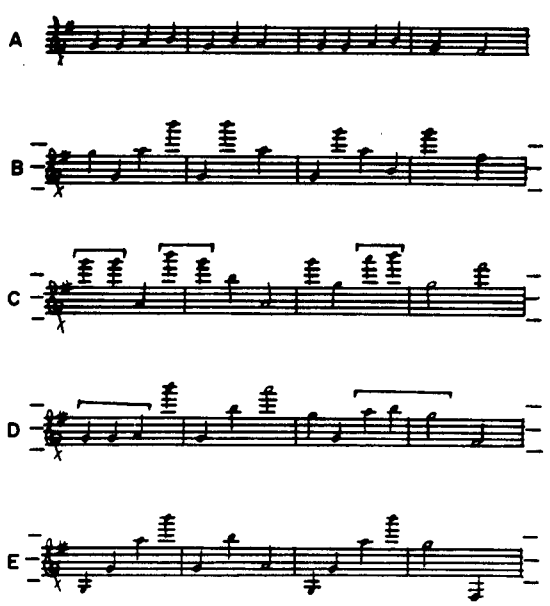



FIGURE 2. From Dowling & Hollombe (1977). Copyright © 1991 by The Psychonomic Society, Inc. Used by permission.

Bouree

*from Suite No. 3 in C for Unaccompanied Cello* J. S. Bach

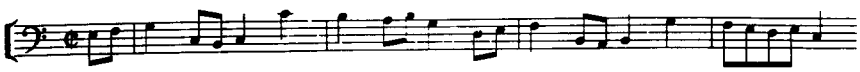


FIGURE 3. J. S. Bach, Bourée, from Suite No. 3 in C for unaccompanied cello.

same octave, or within an overall order of ascending or descending octaves—that is, as the distorted version begins to resemble more and more the original version. These findings, which are based on earlier studies by Diana Deutsch (1972, 1973), encouraged much research on the importance that pitch height (the location of a pitch on a piano keyboard) and pitch chroma (the position of a pitch within a scale) have in the perception of melodies.

To perceive a melody, we must perceive a coherence within the sequence of pitches that make up the melody. We group sounds according to various perceptual and cognitive organizational mechanisms. Some of the musical characteristics that may influence our perception of a melody are its contour, timbre, rhythm, intensity, and tempo (Deutsch, 1982a,b). Bregman (summarized in Bregman, 1990; see also Bregman & Campbell, 1971) has shown experimentally that when we hear rapid sequences of tones varying widely in frequency, we organize the sounds into perceptual streams of narrower range to create a psychological coherence. Musically speaking, in hearing the beginning of the Bourree from the Suite for Violoncello No. 3 in C major by J. S. Bach, although only one instrument is playing, we segregate the sounds perceptually as if we were hearing two lines: a higher and a lower one (Figure 3).

Polyphonic compositions are excellent examples of how perceptual principles are applied when listening to music. For instance, the practice of avoiding part-crossing in writing polyphonic compositions reflects the composer's sensitivity to avoiding potential perceptual confusion on the part of the listener (Huron, 1991).

We may not necessarily recognize, reproduce, or remember a melody exactly as we heard it. Instead, we may recognize, reproduce, or remember a melodic contour that approximates the original melody. In this chapter, W. Jay Dowling discusses the perception of the melodic contour and describes some of the features that make it perceptually relevant. He demonstrates how we perceive the difference between the overall contour of a melody and the specific intervals that make up the contour. He explains the interaction between contour and tonality and describes the results of several experiments that look at the contour of tonal and atonal melodies and transpositions to closely or remotely related keys.

REFERENCES

Arnheim, R. (1969). *Art and visual perception.* Berkeley: University of California Press.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Bregman, A. S., & Campbell, J. (1971). Primary auditory steam segregation and perceptual order in rapid sequences of tones. *Journal of Experimental Psychology, 89,* 244–49.

Deutsch, D. (1972). Octave generalization and tune recognition. *Perception and Psychophysics, 11*(6), 411–12.

Deutsch, D. (1973). Octave generalization of specific interference effects in memory for tonal pitch. *Perception and Psychophysics, 13,* 273–75.

Deutsch, D. (1982a). Grouping mechanisms in music. In D. Deutsch (Ed.), *Psychology of music.* Orlando, FL: Academic Press.

Deutsch, D. (1982b). The processing of pitch combinations. In D. Deutsch (Ed.), *Psychology of music.* Orlando, FL: Academic Press.

Dowling, J. W., & Harwood, D. L. (1986). *Music cognition.* Orlando, FL: Academic Press.

Dowling, J. W., & Hollombe, A. W. (1977). The perception of melodies distorted by splitting into several octaves: Effects of increasing proximity and melodic contour. *Perception and Psychophysics, 21*(1), 60–64.

Huron, D. (1991). The avoidance of part-crossing in polyphonic music: Perceptual evidence and musical practice. *Music Perception, 9*(1), 93–104.

Sloboda, J. A. (1985). *The musical mind.* Oxford: Oxford University Press.

Wertheimer, M. (1923). Untersuchung zur Lehre von der Gestalt II. *Psychologische Forschung, 4,* 301–50.

## What Is Melodic Contour?

Melodic contour—the overall pattern of intervals that make up a melody—is a feature of music that stands out distinctively for the listener on first hearing, and indeed composers use this feature in the construction of pieces of music, making use of the aesthetic tension that results by repeating a contour with altered intervals. A striking example is provided by the first movement of Beethoven's Fifth Symphony, which is built out of a brief melodic motif presented twice in succession at the outset (Figure 7.1). That motif and its sequences share contour (indicated under the score) but differ in the size of the interval between the two pitches: four semitones between G and E-flat in the first presentation and three semitones between F and D in the second. Throughout the movement,

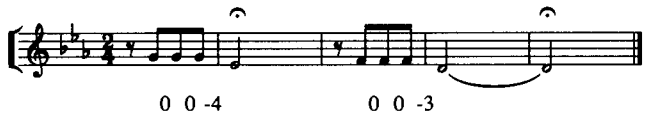



FIGURE 7.1. The opening measures of Beethoven's Fifth Symphony. Contour and interval sizes are indicated under the score by 0 for unison and – with the interval size in semitones (half steps) for a descending interval. Ascending intervals would be indicated by +.

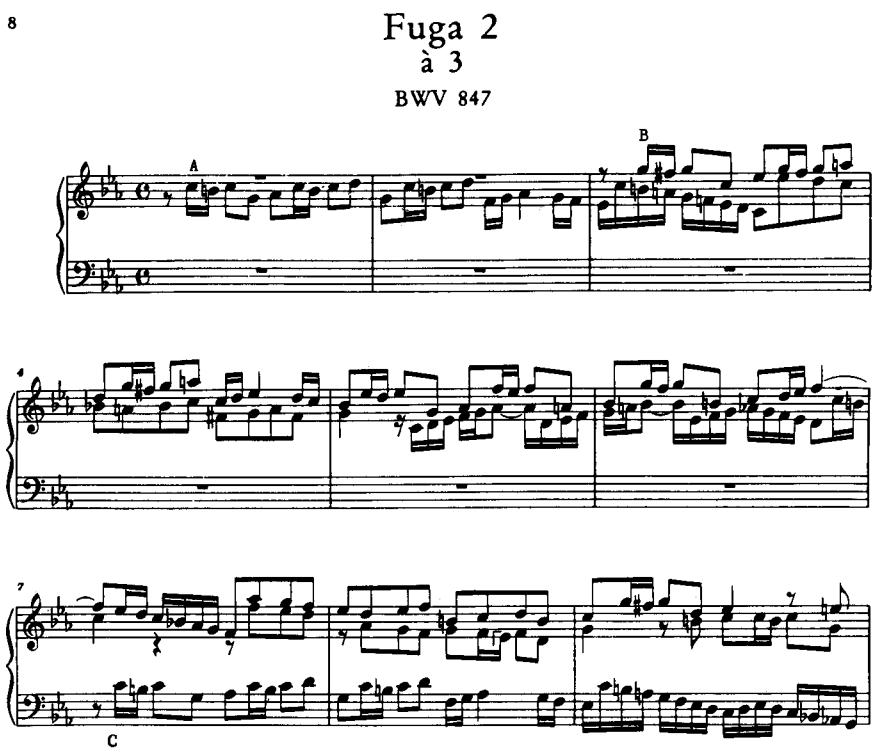



FIGURE 7.2. Bach's Fugue in C Minor from Book I of *The Well-Tempered Clavier.* Occurrences of the subject in the three voices. (Schott/Universal Edition, ed. Walther Dehnhard.)

TABLE 7.1 Interval Patterns in Semitones for Presentation of the Three Voices in Bach's Fugue in C Minor from *The Well-Tempered Clavier,* vol. 1, BWV 847

| *Occurrence/ measure* | | *Intervals* |
|---|---|---|
| 1 | 1 | −1 +1 −5 +1 +4 −1 +1 +2 −7 +5 −1 +1 +2 −9 +2 +1 −1 −2 −2 |
| b | 3 | −1 +1 −7 +3 +4 −1 +1 +2 −7 +5 −1 +1 +2 −9 +2 +1 −1 −2 −2 |
| c | 7 | like a |

Beethoven repeats this motif over and over, using virtually all the possible intervals between the two pitches, in each case preserving the contour. The repeated contour provides unity to the piece, while the small intervallic changes with each repetition provide interest.

There are many examples of the use of a repeated contour with changes in intervallic detail. One can be found in Bach's Fugue in C minor from volume 1 of *The Well-Tempered Clavier* (Figure 7.2). The initial pitch contour [ − + − + ] at a in the subject is repeated not only in recurrences of the whole subject (at b and c) but also as a simple thematic kernel throughout the fugue. If we compare the recurrences of the initial melodic contour a, we see that while b preserves the contour pattern, it does not preserve the pattern of exact interval sizes in semitones. This is because of the difference between "tonal answers" and "real" answers in a fugue—tonal answers modify the interval pattern so that they can remain in the same key, while exact answers preserve the melodic pattern at another pitch level. Table 7.1 lists the pattern of intervals for the three appearances of the initial contour. Here Bach has treated contour as a feature of the melody that can be transported into other contexts and given different intervals while still preserving its identity. In fact, melodic contours recur in fugues with each return of the subject, but only sometimes does the subject repeat the initial melodic interval pattern exactly.

## The Importance of Melodic Contour: Experimental Evidence

The importance of melodic contour in melodies we hear for the first time is borne out by experiments that test the recognition memory of subjects. In such experiments the listener hears a novel melody—never before encountered—on each trial, and then has the task of recognizing a transposition of that melody. Sample trials from a typical experiment are shown in Figure 7.3. On some trials the test melody will in fact be an exact transposition of the initial melody, in which case the listener's response should be, "Yes, that is the melody I just heard." (Figure 7.3B is a transposition of Figure 7.3A.) On other trials, however, the test melody will be different, and the listener should say, "No, that is not the

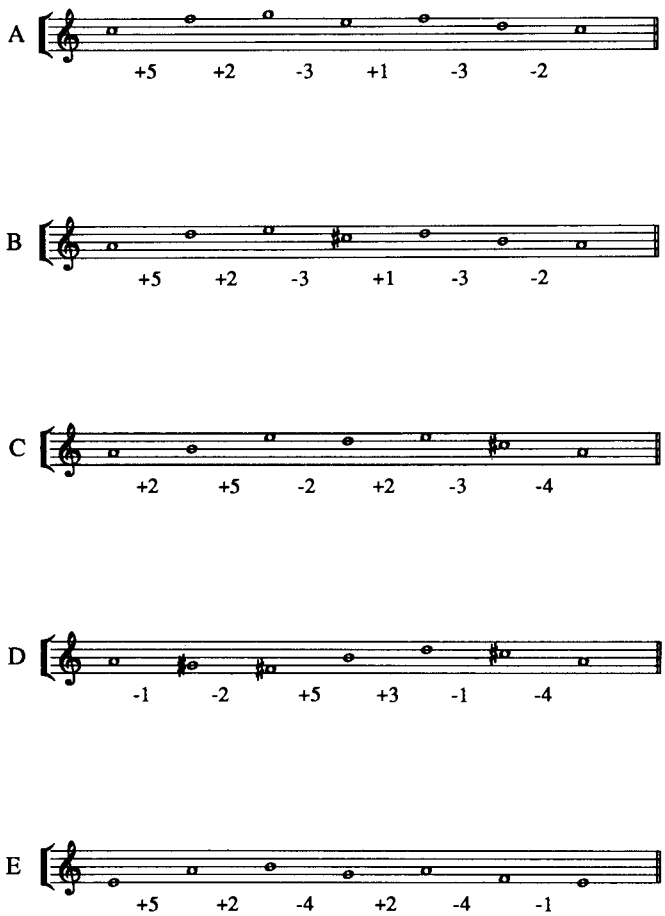



FIGURE 7.3. Typical melodies used in recognition memory experiments: (A) a novel melody in the key of C major presented on a trial; (B) an exact transposition of that melody to the key of A major; (C) a same-contour imitation of the initial melody, with different interval sizes; (D) a test melody with a different contour; (E) an imitation produced by moving the original melody along the C major scale, preserving diatonic but not semitone intervals. The contour and interval sizes are shown below each score.

same as the first melody." When the test melody is different, it might have a totally different contour (as in Figure 7.3D), in which case it will be easy for the listener to reject it. The most difficult type of melody recognition test will be an imitation of the original melody that shares its overall contour but differs in interval sizes (as in Figures 7.3C and 7.3E). Contour similarity will pull the listener in the direction of saying "Yes," while the intervallic differences, which are not always easy to notice on first hearing, prevent such a melody from being an exact transposition.

The results of a number of experiments show that listeners usually find it easy to respond positively to all comparison melodies that share contour, and respond negatively to melodies with different contours. Listeners find it much more difficult to distinguish exact transpositions from same-contour imitations. Dowling and Fujitani (1971) found this to be true using nontonal melodies; Dowling (1978) found it true using tonal melodies.

The confusion of transpositions with same-contour imitations is especially strong when the two melodies are in the same or in closely related keys (Bartlett & Dowling, 1980). That is, two melodies that share both contour and scale are easily confused, even if they differ in pattern of intervals. This confusion diminishes as the scales are made less similar by increasing key distance around the circle of fifths.

What psychological principles underlie the importance of contour in human memory? We now proceed to a description of the psychological features of melodies. The evidence suggests that contour similarity is especially important for recently encountered novel melodic phrases that are about five notes long. Contour is less important as melodies are remembered for longer periods of time, as they become more familiar, and when they are longer than five notes. Contour is more important for younger children under the age of 5 than for older ones, since before the interval patterns of tonal scales in a culture's music must be learned and incorporated into the listener's pitch, analyzing procedures and contour represents the main organizing principle for pitches in melodies (Dowling & Harwood, 1986).

## The Role of Tonality

### *Features of Melodies*

In addition to contour and interval sizes, a set of pitches also serves as a set of features to describe a melody. With that set of pitches comes the feature of its "key," or "tonality" (at least for tonal melodies). Key or tonality is the organization of the tones and chords of a piece of music in relation to a tonic. Contour and intervals remain invariant when a melody is transposed but the key changes—now it is based on a different tonality (as between C major and A major going from Figure 7.3A to 7.3B). All these features are psychologically relevant. Listeners are often led to treat an imitation as a transposition when contour and key remain similar to the original (Bartlett & Dowling, 1980). This tendency to confuse imitations and transpositions is especially strong when the key remains the same, and the imitation has simply been moved along the original scale (as in Figure 7.3E), preserving only some of the intervals (Dowling, 1978).

Key is an important feature in other ways. Watkins (1985) has shown that listeners experienced in Western music find a melody easier to recognize when

its pitch set remains within a diatonic musical scheme. Furthermore, the key is not just a static set of notes; it is a framework that assigns tonal functions to the various pitches in hierarchical relations to one another (Krumhansl, 1990). The tonic pitch is the most important, and other pitches tend to gravitate toward the tonic as a center of stability. Melodies such as those in Figure 7.3A to 7.3D begin and end with the tonic, moving from a stable point out into the rest of the scale and then returning to the stable tonic. The seventh scale degree (or "leading tone") for example, tends to pull upward toward the tonic. Francès (1958/1988) showed that when the leading tone was mistuned downward, contrary to that tendency, the mistuning was much more noticeable than when it was mistuned the same amount in the direction of the tendency.

A key defines a set of more or less expected pitches within the range of possibilities. When a melody follows those expectations, it is easier to follow and to remember. Watkins (1985) contributed additional evidence for the importance of key and the tonal hierarchy by showing that melodies whose pitches are chosen from a relatively small region around the circle of fifths are easier to recognize than melodies that involve large leaps around the circle of fifths.

A feature of melodies that should be mentioned here because of its importance is rhythm. In all the examples mentioned so far, not only melodic contour but the rhythmic pattern of the melody has been preserved. Changing the rhythm would produce a markedly different melody (Monahan, 1984). However, consideration of rhythm is beyond the scope of this chapter.

A variety of features can be used to describe a melody. In addition to contour, there are the interval pattern, pitch set, and the key. All of these have been shown to have some importance to the way in which melodies are first perceived and then remembered. But which of those patterns is most fundamental in our memory for melodies? Since interval patterns most obviously remain constant through transposition, a good case can be made that they are the most fundamental. Nevertheless, I believe that melodies are remembered as ordered sets of pitches. The pitch framework provided by the scale temporarily fixes a set of pitches as landmarks in terms of which the pitches of the melody are organized and understood.

From a psychological point of view, the scale framework itself is best understood as a set of pitches, rather than as a set of intervals, as I have argued at greater length elsewhere (Dowling, 1991a). The dynamic tendencies of pitches to gravitate toward points of stability in a tonal context, as described above in the experiment by Francès (1958/1988), are characteristic of sets of pitches, not of intervals. Furthermore, intervals, far from being psychologically more fundamental than pitches, are themselves most easily grasped with reference to pitch patterns of familiar melodies. Imagine the case of a music student faced with the task of singing the interval "ascending minor seventh" (Figure 7.4A). Abstracted from a context of melody and key, this is a very difficult task. However, if the student should happen to recall that a familiar melody (such as "Somewhere" from Bernstein's *West Side Story*; Figure 7.4B) begins with an ascending minor

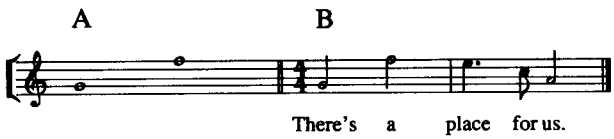



FIGURE 7.4. The interval of an ascending minor seventh (10 semitones); (B) the same interval embedded in the start of the song "Somewhere" from Leonard Bernstein's *West Side Story.*

seventh, the task is easily solved. Students learn to sing difficult intervals most easily with reference to the pitch patterns of familiar songs in which they occur.

Another line of evidence arises from multidimensional scaling studies which suggest that intervals in a melodic context are generally heard, not as intervals per se, but rather in terms of their position within the scale (Balzano & Liesch, 1982). Still more evidence comes from studies involving cued recognition, in which the listener hears a familiar tune and then is asked to discern it even when it is presented in a confusing way designed to preserve the pitch classes of its notes while destroying their interval relationships. Listeners are able to recognize tunes even when their intervals (but not their pitches) are obscured by scrambling their notes into different octaves (Dowling, 1984a) or interleaving distractor tones temporally among their notes (Dowling, 1973a). All of this strongly suggests that the psychological representation of the tonal framework is in terms of pitches, and not the logically equivalent representation in terms of intervals.

## Memory for Melodic Contour

Several factors affect listeners' memory for melodic contour relative to memory for other features of melodies such as intervallic detail. Those factors include the delay between first hearing a melody and when it is tested, the familiarity of the melody to be remembered, its length, and the age and experience of the listeners. I will treat each of these factors in turn.

### *Short versus Long Delays in Recognition Memory*

Contour is a strong determinant of transposition recognition when the comparison follows the introduction of a new melody by an unfilled delay of a few seconds. Contour becomes less important, however, after a delay of 30 or 40 seconds, while interval sizes become more important (Dowling & Bartlett, 1981; DeWitt & Crowder, 1986). One reason contour decreases in importance over time in relation to intervals is a decrease in the confusion due to scale relations, as already noted. After a short delay the scale of the initial melody is still fresh in memory, and perception of the comparison melody is affected by it. When the comparison is in a relatively close key (as in the experiments just cited), confu-

sion between exact transpositions and same-contour imitations results. Following a delay filled with other material the scale of the initial melody becomes less salient, and false recognitions of same-contour imitations decline. According to this explanation, the shift over time of the relative importance of contour and intervals should disappear with atonal melodies, since confusions due to key similarity should no longer operate. And that is exactly what happens (Dowling, 1991b). Atonal melodies are more difficult to recognize than tonal ones after both long and short delays, and (as with tonal melodies) discriminating contour differences is easier than discriminating interval differences. However, the relative difficulty of those two discriminations does not shift over time as it does for tonal melodies.

A second reason that has been put forward for the decline in importance of contour with longer delays is that contour can serve as a much better retrieval cue when the initial melody and its comparison are close together in time and there are fewer alternatives to choose from (Dowling & Bartlett, 1981). However, when the comparison comes after a delay filled with alternative possible choices lying between it and the initial melody, the relative importance of contour declines. (This would be like a series of trials with melodies such as in Figure 7.3, but with numerous items like 7.3D intervening between 7.3A and 7.3B.) We cannot rule out the possibility that more precise information than contour (namely, interval sizes) may be needed to differentiate among the numerous melodies we store in our memory. One could argue that such an effect ought to operate as well with atonal as with tonal melodies. Moreover, Dowling (1991b) failed to find any decline in the usefulness of contour information between short and long delays in that experiment.

Nevertheless, it is clear that with the passage of time the contours of tonal melodies become relatively less important to recognition and specific intervals become relatively more important. The end result of this shift can be seen in the great importance of intervals in memory for overlearned melodies in long-term memory storage.

### *Familiar versus Unfamiliar Melodies*

Contour is highly salient in memory for novel melodies. In contrast, the exact interval patterns of familiar melodies are very well remembered. Dowling and Fujitani (1971) found that listeners rarely confused same-contour imitations with familiar versions of tunes such as "Twinkle, Twinkle, Little Star" and "Yankee Doodle." In fact, when asked to respond positively to the imitations, listeners succeeded about 60% of the time. Listeners' knowledge of the intervals of familiar tunes is so precise that Attneave and Olson (1971) were able to use memory for the intervals of the call signal of NBC, an American television network (G-E-C; intervals +9 semitones, −4 semitones) as the basis for a pitch-scaling task. This accuracy with familiar melodies is true of musically inexperienced as well as experienced listeners.

### *Melody Length*

The familiar versus unfamiliar melody results were obtained with five-note-long melodies (Dowling & Fujitani, 1971; Dowling, 1978; Bartlett & Dowling, 1980) and seven-note-long melodies (Dowling & Bartlett, 1981; Dowling, 1991b). Edworthy (1985) has shown, however, that contour is not nearly so easy to remember with longer tonal melodies of 11 or 13 or 15 notes. Part of this result must be due to complexity, and Cuddy, Cohen, and Mewhort (1981) have shown that contour complexity affects memory. Part of the result may, as Edworthy suggests, also be due to the fact that a longer melody provides time for the listener to achieve a firmer grasp of the key. Edworthy's task was quite difficult, requiring the listener to notice particular details in the contour rather than make a global same-different judgment. However, it is plausible that contour should decline in importance with melody length, since inevitably the number of possible contours (hence the amount of contour information conveyed) increases with length.

The way longer melodies are remembered should be a fruitful area for research. Longer melodies must break down into phrases that should be to some extent remembered independently. We know that phrases are in fact remembered as units. Dowling (1973b) constructed melodies using four five-note phrases to form the rhythmic pattern [. . . ._ . . . ._ . . . ._ . . . ._ ]. Each phrase had a different pitch contour. After hearing such a melody, listeners were given test patterns that replicated whole phrases from the longer melody [. . . ._] or that cut across phrase boundaries [. ._ . . ]. Memory performance was much better with the whole phrases (about 70% correct) than with the ends of one phrase and the beginnings of another (about 60% correct). We also know something of the way melodies are parsed into phrases on the basis of rhythmic grouping, changes in melodic contour, melodic leaps, patterns of repetition, and so forth (Jones, 1987). The relationship between memory for phrases and memory for whole melodies poses a particularly interesting question for future research.

### *The Role of Development*

Infants are sensitive to changes in melodic contour long before they begin to speak or sing (Trehub, Bull, & Thorpe, 1984). As soon as children begin to speak they also begin to sing. That singing generally takes two forms: vocal play exploring the limits in pitch and loudness and rapidity of what their voice can do, and the creation of spontaneous songs (McKernon, 1979; Dowling, 1982, 1984b, 1988; Dowling & Harwood, 1986). Spontaneous songs at the age of one or two years are characterized by a more or less regular beat on which the speech

rhythms of the words of the song and the use of discrete levels of pitch are overlaid. The pitches are not those of the adult scale and the pitch levels are continually in flux. Nevertheless the steadiness of the metrical pattern (the beat) and the use of discrete, sustained pitch levels for vowels distinguishes singing from speech at this age. What the child controls at this age in terms of pitch is not the exact frequency or pitch of the tone, but the pattern of ups and downs—the contour.

Children sing spontaneously created songs beginning around the age of one. Such songs divide naturally into phrases that are characterized by melodic-rhythmic contours. The melodic contour patterns recur within the structure of the song. Sometimes contour patterns are copied from nursery tunes the child has heard, but typically they are made up by the child. Sometimes a contour will repeat over and over with the same words, whether nonsense ("Sippah, suppah. Sippah, suppah. . . .") or relatively meaningful ("Come a duck on my house. Come a duck on my house. . . ."). Sometimes one contour shape will alternate with another in a regular pattern throughout the song. Often the song ends with a "coda" constructed from new contour material. As the child gets older, some of the songs come to approximate nursery tune models more and more closely, but even then those songs are not literal reproductions of adult models, but are versions that have been regularized according to the child's own musical syntax (Dowling, 1984b). Children's spontaneous singing through the preschool years involves principally a manipulation of musical materials encoded as rhythmic and melodic contours.

By the age of five children's pitch levels are beginning to approximate adult categories, and some children younger than that can distinguish between tonal versions of familiar melodies that conform to adult pitch categories for the diatonic scale and atonal versions that do not (Trehub, Morongiello, & Thorpe, 1985), especially if they are rated by adult judges as being able to sing "in tune" (Dowling, 1988). Also, by five, children are able to stay within the scale of a single key in singing a familiar tune (Davidson, McKernon, & Gardner, 1981). By the age of eight, children notice changes of mode from major to minor (Imberty, 1969), are better at noticing pitch alterations in tonal (vs. atonal) melodies (Zenatti, 1969), and are beginning to acquire a sophisticated scheme for the hierarchical organization of pitches in a tonality (Krumhansl & Keil, 1982).

These shifts in relative importance of the various melodic features with changes in age and experience can be seen in the results of Andrews and Dowling (1991). Their listeners across a variety of ages were asked to identify familiar melodies in versions that either exactly reproduced their familiar pitch intervals or were changed in pitch while preserving contour. The changed versions either remained within the same key ("tonal") or did not remain within any key ("atonal"). Andrews and Dowling found a regular developmental progression of reliance on different melodic features. Five- and six-year-olds identified the melodies equally well in all forms, changed as well as original versions. Seven- and

eight-year-olds tended to identify principally tonal melodies, original and changed. Nine- and 10-year-olds performed much better with original version of the melodies than with either tonal or atonal changed melodies. With these highly familiar melodies, adults performed about equally well on all versions.

### *Contour and Gestalt Principles*

One result that has gradually emerged from experiments in which listeners have judged melodic contour in a variety of contexts is that a melody is very much an integrated whole, a Gestalt. Contour is not an entirely separable feature of a melody in Garner's (1974) sense. Contour interacts with both tonality and rhythmic pattern in perception and memory. This is not what I initially thought when I began to investigate this area. In 1978 (Dowling, 1978) I suggested that contour and scale (the embodiment of the features of tonality) might be processed and remembered independently of one another. Thus a melody could be created by generating a contour and hanging it on a scale. Scale similarity might cause confusions in short-term recognition tasks, but only with interval, not contour, recognition. That is, changes in the scale should not affect recognition of the pattern of ups and downs.

Recent studies have shown, however, that tonal context affects memory for contour. Bartlett and Dowling (1988) had listeners judge sameness of contour in pairs of seven-note melodies in which one member of the pair was tonal and the other atonal. The contour change was produced by making the lowest pitch in the melody the highest, or vice versa. There were always the same number of specific pitch differences between members of a pair. Listeners were better at discriminating contour changes when the pairs were presented in atonal-tonal order than the reverse. And in a study already cited (Dowling, 1991b), I had listeners attempt to discriminate transpositions from same-contour imitations and different-contour test melodies (like those in Figure 7.3) either immediately or after a delay of about 40 seconds filled with other melodies. The melodies were strongly tonal, weakly tonal, or atonal. Strongly tonal melodies began and ended on the tonic (like Figure 7.3A to 7.3D), and weakly tonal melodies began and ended on pitches other than the tonic (like Figure 7.3E). Listeners' performance is shown in Figure 7.5. At short delays discriminating transpositions from different-contour lures was uniformly easier than discriminating them from same-contour lures for both tonal and atonal melodies (Figure 7.5). With increasing delay this difference tended to disappear for the tonal melodies, but not for the atonal. Furthermore, the tendency to falsely recognize different-contour lures was markedly affected by tonality, being lowest for strongly tonal melodies. Thus listeners' ability to reject a different-contour test item depended on tonal context, and not just on its difference of contour.

Contours are interrelated in memory with tonal and rhythmic context. The parsing of a contour into phrases and the assignment of melodic accents depend

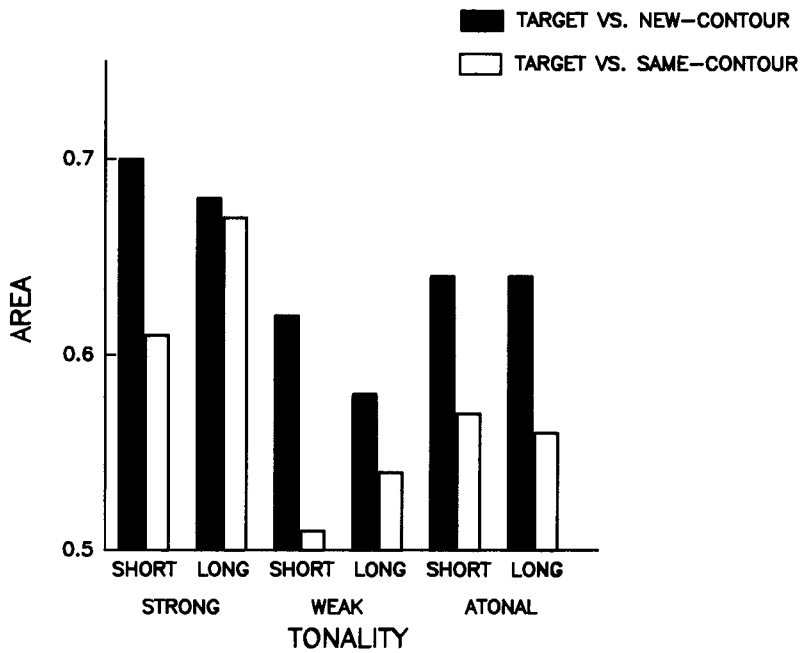



FIGURE 7.5. The interaction of tonality (Strongly Tonal, Weakly Tonal, Atonal) by delay (Short and Long) by test comparison (Targets vs. Different—or Same-Contour Lures). Performance distinguishing between exactly transposed targets and different-contour lures is shown by solid bars; that for targets versus same-contour lures is shown by open bars (after Dowling, 1991b).

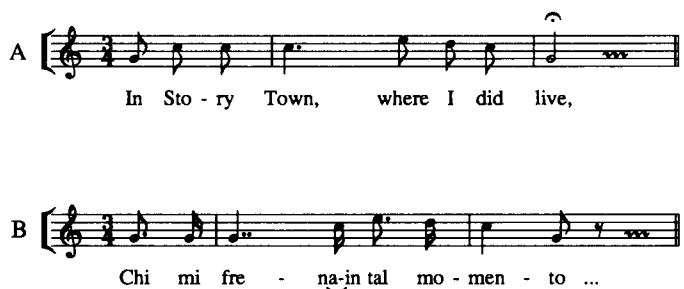



FIGURE 7.6. The opening phrases of a version of the folksong "Barbara Allen" and the sextet from Donizetti's *Lucia di Lamermoor* (after Kolinski, 1969, Fig. 17). Note the identical pattern of pitches with different rhythm, which in effect makes it a different melody.

not only on note durations, but also on accents arising from contour inflections (Jones, 1987; Drake, Dowling, & Palmer, 1991). Kolinski (1969) vividly demonstrated the interrelatedness of pitch and rhythmic patterns by showing that, if the rhythm and the number of a repeated pitch were altered, a variant of "Barbara Allen" would become the tune from the Sextet from Donizetti's *Lucia di Lammermoor* (see Figure 7.6). The very same pitch pattern (and contour) can become an entirely different tune with a change of rhythm.

It would have been simpler for psychology if melodies had divided neatly into contours, rhythms, and pitch sets, but it would probably have been a loss to music. The ways in which musical material is transformed by shifts of context provide an important source of complexity which contributes to its richness as art.

## Summary

In this chapter I discussed melodic contour in relation to other features of melodies in the pitch and time domains. An important feature of melodic contour is the distinction between pitches as interval patterns and pitches as patterns within a tonality. Evidence was cited suggesting that both representations are necessary in music cognition. Along with evidence showing the importance of contour in melody recognition and production, I presented evidence showing the limits of that importance. Contour is most important in memory for novel melodies tested after brief delays, and less so after longer delays and with familiar melodies. While contour dominates much of the preschool child's music behavior, adults are able to use contour more flexibly when the task demands it. Finally, I noted that melodic contour does not function entirely independently of the other features of melodies, but rather is part of the integrated melodic Gestalt.

### REFERENCES

Andrews, M. W., & Dowling, W. J. (1991). The development of perception of interleaved melodies and control of auditory attention. *Music Perception*, 8, 349–68.

Attneave, F., & Olson, R. K. (1971). Pitch as medium: A new approach to psychophysical scaling. *American Journal of Psychology*, 84, 147–66.

Balzano, G. J., & Liesch, B. W. (1982). The role of chroma and scalestep in the recognition of musical intervals in and out of context. *Psychomusicology*, 2, 3–31.

Bartlett, J. C., & Dowling, W. J. (1980). Recognition of transposed melodies: A key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 501–15.

Bartlett, J. C., & Dowling, W. J. (1988). Scale structure and similarity of melodies. *Music Perception*, 5, 285–314.

Cuddy, L. L., Cohen, A. J., & Mewhort, D. J. K. (1981). Perception of structure in short melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 869–83.

Davidson, L., McKernon, P., & Gardner, H. (1981). The acquisition of song: A developmental approach. In *Documentary report of the Ann Arbor symposium* (pp. 301–15). Reston, VA: Music Educators National Conference.

DeWitt, L. A., & Crowder, R. G. (1986). Recognition of novel melodies after brief delays. *Music Perception*, 3, 259–74.

Dowling, W. J. (1973a). The perception of interleaved melodies. *Cognitive Psychology*, 5, 322–37.

Dowling, W. J. (1973b). Rhythmic groups and subjective chunks in memory for melodies. *Perception and Psychophysics*, 14, 37–40.

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, 85, 341–54.

Dowling, W. J. (1982). Melodic information processing and its development. In D. Deutsch (Ed.), *The psychology of music* (pp. 413–29). New York: Academic Press.

Dowling, W. J. (1984a). Musical experience and tonal scales in the recognition of octave-scrambled melodies. *Psychomusicology*, 4, 13–32.

Dowling, W. J. (1984b). Development of musical schemata in children's spontaneous singing. In W. R. Crozier & A. J. Chapman (Eds.), *Cognitive processes in the perception of art* (pp. 145–63). Amsterdam: North-Holland.

Dowling, W. J. (1988). Tonal structure and children's early learning of music. In J. Sloboda (Ed.), *Generative processes in music* (pp. 113–28). Oxford: Oxford University Press.

Dowling, W. J. (1991a). Pitch structure. In P. Howell, R. West, & I. Cross (Eds.), *Representing musical structure* (pp. 33–57). London: Academic Press.

Dowling, W. J. (1991b). Tonal strength and melody recognition after long and short delays. *Perception and Psychophysics*, 50, 305–13.

Dowling, W. J., & Bartlett, J. C. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology*, 1, 30–49.

Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, 49, 524–31.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition*. New York: Academic Press.

Dowling, W. J., & Hollombe, A. W. (1977). The perception of melodies distorted by splitting into several octaves: Effects of increasing proximity and melodic contour. *Perception and Psychophysics*, 21, 60–64.

Drake, C., Dowling, W. J., & Palmer, C. (1991). Accent structures in the reproduction of simple tunes by children and adult pianists. *Music Perception*, 8, 315–34.

Edworthy, J. (1985). Melodic contour and musical structure. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition* (pp. 169–88). London: Academic Press.

Francès, R. (1958). *La perception de la musique*. Translated by W. J. Dowling. Hillsdale, N. J.: Earlbaum, 1988.

Garner, W. R. (1974). *The processing of information and structure*. New York: Wiley.

Imberty, M. (1969). *L'acquisition des structures tonales chez l'enfant*. Paris: Klincksieck.

Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception and Psychophysics*, 41, 621–34.

Kolinski, M. (1969). "Barbara Allen": Tonal versus melodic structure, part II. *Ethnomusicology*, 13, 1–73.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Krumhansl, C. L., & Keil, F. C. (1982). Acquisition of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 579–94.

McKernon, P. E. (1979). The development of first songs in young children. *New Directions for Child Development, 3,* 43–58.

Monahan, C. B. (1984). *Parallels between pitch and time: The determinants of musical space.* Unpublished doctoral dissertation, University of California at Los Angeles, Los Angeles.

Povel, D.-J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception, 2,* 411–40.

Trehub, S. E., Bull, D., & Thorpe, L. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development, 55,* 821–30.

Trehub, S. E., Morongiello, B. A., & Thorpe, L. A. (1985). Children's perception of familiar melodies: The role of intervals. *Psychomusicology, 5,* 39–48.

Watkins, A. J. (1985). Scale, key, and contour in the discrimination of tuned and mistuned approximations to melody. *Perception and Psychophysics, 37,* 275–85.

Zenatti, A. (1969). Le developpement génétique de la perception musicale. *Monographies Françaises de Psychologie, 17.*

# 8

# Describing the Mental Representation of Tonality in Music

DAVID BUTLER AND HELEN BROWN

## Introduction

The research of Hermann von Helmholtz (1885/1954) on the combinations of musical sounds and the qualities of musical tones has influenced most subsequent investigations of the perception of consonance and tonality. Helmholtz believed that to study auditory events carefully they should be broken down into their most elementary units, and that certain pitch relationships are musically significant because of the physical properties of the tones themselves.

During the last decade, research in the perception of tonality has figured prominently in the psychology of music. The scholarly work of Carol Krumhansl described in *Cognitive Foundations of Musical Pitch* (1990) has been very significant (see also Krumhansl, 1991; for summaries of her earlier work see Krumhansl, 1983, 1985). Krumhansl's research has offered empirical validity to aspects of music theory such as the prominence that specific tones assume within a scale, and the relationships between the keys in the circle of fifths. Often she utilized geometric models to describe her results. Most of Krumhansl's experiments investigated tonality using what she called "the probe-tone technique." This technique required subjects to listen to a musical context that could be an ascending or descending diatonic scale, a chord, a chord progression, or a brief excerpt from a musical composition. The musical context was followed by the presentation of a tone (called "the probe tone") which was randomly chosen from among the 12 pitches. Subjects are asked to rate on a scale from 1 to 7 (1 = very bad to 7 = very good) how well the probe tone fit in the previously heard musical context.

In one of her early experiments (Krumhansl & Shepard, 1979) Krumhansl investigated how subjects perceived tonality using as the musical context an

incomplete ascending or descending C major scale played without the final tonic and followed by the probe tone. She found that the results of subjects with a formal musical background conformed to the qualitative predictions of music theory: these subjects perceived the tonic, the dominant, and the mediant tone of the scale as the tones best fitting the given musical context.

Krumhansl's probe-tone ratings have also been applied in investigating the perception of North Indian (Castellano, Bharucha, & Krumhansl, 1984) and Balinese music (Kessler, Hansen, & Shepard, 1984), demonstrating that the probe-tone technique could also be successfully used to investigate the tonal hierarchies of these non-Western musical styles.

The research conducted by Butler, Brown, and Bharucha has provided valuable information that emphasizes other aspects of the perception of tonality. In this chapter, Butler and Brown describe the importance of the temporal element in perceiving tonality. They look at how single pitches and combinations of pitches are organized over time and show that it is the order in which pitches are played, not just the pitches themselves, that influences which key the listener perceives a piece to be in (see Brown, 1988; Brown & Butler, 1981; Butler, 1989, 1991). Butler and Brown believe that the scale per se does not adequately describe the relationships that listeners actually hear in the music; instead, they propose that the intervals of the scale indicate the tonality to the listeners.

Specifically, Butler and Brown demonstrate that within a major scale, the minor second and the tritone provide distinctive cues in perceiving tonality. Butler and Brown call the minor second and the tritone "rare" intervals because within a major scale the interval of a minor second occurs only twice (between the third and fourth scale degrees, and between the seventh scale degree and the tonic) and the tritone occurs only once (between the fourth and seventh scale degrees); thus they are heard less frequently than the other intervals. Butler and Brown propose that the infrequency with which these intervals occur within the major scale gives the listener distinctive cues for perceiving which tonality a piece is in.

## REFERENCES

Brown, H. (1988). The interplay of set content and temporal context in a functional theory of tonality perception. *Music Perception, 5,* 219–50.

Brown, H., & Butler, D. (1981). Diatonic trichords as minimal tonal cue-cells.:I2 In Theory Only, 5(6,7), 37–55.

Butler, D. (1989). Describing the perception of tonality in music: A proposal for a theory of intervallic revalry. *Music Perception, 6,* 219–41.

Butler, D. (1991). *The musician's guide to perception and cognition.* New York: Schirmer Books.

Castellano, M., Bharucha, J., & Krumhansl, L. C. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General, 113,* 394–412.

Helmholtz, H. L. F. (1954). *On the sensations of tone as a physiological basis for the theory of music* (A. J. Ellis, Ed. and Trans.). New York: Dover. (Original work published 1885)

Kessler, E., Hansen, C., & Shepard, R. N. (1984). Tonal schemata in the perception of music in Bali and the West. *Music Perception, 2,* 131–65.

Krumhansl, C. L. (1983). Perceptual structures for tonal music. *Music Perception, 1,* 28–62.

Krumhansl, C. L. (1985). Perceiving tonal structure in music. *American Scientist, 73,* 371–78.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch.* New York: Oxford University Press.

Krumhansl, C. L. (1991). Music psychology: Tonal structures in perception and memory. *Annual Review of Psychology, 42,* 277–303.

Krumhansl, C., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 579–94.

---

During the past three centuries, listeners in our Western musical culture have learned to expect certain underlying conventions in ways tones are combined to form meaningful patterns. A few of these conventional combinations are related to acoustical and sensory characteristics of tones, but most tonal patterns can change from one musical culture to another and from one style to another within a culture. Our purpose in this chapter is to examine these tonal patterns as perceptual objects and as cognitive processes. A goal of this discussion is to impress on the reader the importance of temporal relations in the comprehension of pitch relationships in tonal music—the importance of the manner in which pitches provide context for one another as the music unfolds before the listener.

Nature presents us with choices of smoother- or rougher-sounding combinations of tones and these combinations might have important connotations within one musical style, or throughout an entire musical culture; yet, musical cultures are sometimes very different. Our own musical culture includes highly varied approaches to combining pitches and even to contriving pitch-sets from which pitches may be combined; as the same concert, for example, you can hear a minimalistic work that might contain only five or six pitch classes, followed immediately by a microtonal work containing two dozen or more tones within the octave. Within these extremes, several well-worn, well-understood stylistic norms have been tremendously important in shaping Western music for the past

three centuries or so. Another way of saying this is that composers and listeners have formed an unspoken pact, a fairly sophisticated set of ground rules, regarding how pitches are to be organized to form normal-sounding patterns. This set of conventions is known as *tonality*, which undergirds the learned system of "keyness" that to some extent governs how melodies and chords are organized throughout the duration of a tonal musical composition. In this chapter, we describe these conventions of organization and then review experimental studies that have described listeners' recognition of these conventions.

## Pitch Relationships

It is necessary to make a few distinctions about the ways that the term "pitch" might be used.[1] Pitch may refer to nothing more than a loose registral differentiation, the difference between "high" and "low." Sounds with little or no discernible pitch—clicks, thumps—can nonetheless be described as having a higher or lower pitch. The definitions of pitch relating more directly to this discussion are two levels more specific than this, however. The first level of specificity refers to the pitch material from which tonality is built. There are 12 equal logarithmic divisions of the octave; these equal divisions produce the equal-tempered 12-tone set, variously called the *chromatic set*, the *chromatic scale* (if the tones are arranged in scalar order), or the *equal-tempered twelfth-octave collection*. Distances among tones in this 12-tone set can be described neutrally using modulo-12 (clockface) numbers: 0 stands for C, B-sharp, D double-flat, and its other enharmonic spellings in all possible octave registers; 1 stands for the next-higher semitone, and so on through 11, which stands for B and its enharmonics. The system therefore refers not to single, specific pitches, but rather to 12 *pitch classes*. The neutrality of these pitch-class numbers is appropriate for discussions of pitch relations beyond the scope of the Western tonal tradition, but discussions of tonal pitch relations cannot remain neutral. The second level of specificity pertains to tonal context. Intervals can sometimes be considered nothing more than gaps between pitches. The term "interval class" is conventionally applied to these neutral spans of pitch space: the intervals E–F and D-double-sharp–G-double-flat both belong to interval-class (ic) 1, for example, because each spans one semitone. Interval-class labels, however, are not precise enough to describe pitch relationships in tonal music. For example, the three-semitone gap between pitch classes 0 and 3 has no distinctive tonal meaning, but when 0 becomes the specific pitch middle C, and 3 becomes the E-flat a minor third above it, that particular three-semitone gap will be found in different harmonic and tonal contexts than it would if 0 were C and 3 were the D-sharp an augmented second above it. When discussions involve pitch sets derived from specific tonal contexts, spelling definitely counts.

This discussion implies that the pitch set comes close to describing a pitch pattern found in a piece of music already composed. This is not just academic

fussiness; rather, it is one of the more important musical observations in this chapter. Although some accounts describe scales as the basis of music, historically the music comes first, and the scales come later. Scales are theorists' attempts to describe the pitch materials in music that already exists.[2] The scale is a handy index of the pitch contents of some piece or pieces of music, but the scale's simplicity comes at a hefty price. The scale is a highly simplistic construction that fails to describe much of what is most important about pitch relationships in tonal music. First, there is no guarantee that all notes in the scale may be used in a composition, and there is definitely no guarantee whatever that the scale fully represents all notes that may be encountered within the music. Second, the scale says nothing at all about how the notes within the piece of music are to be arranged throughout the composition—how they are to unfold as the composition is heard.

A listener who has become familiar with the tonal idiom tacitly gains an underlying sense of *key level*, and within that a sense of *harmonic level*, in a piece of tonal music as it unfolds across time. This is one of the strongest perceptual binding forces—that is, features through which we perceive coherence—in tonal music (see, for example, Krumhansl, 1979; Brown & Butler, 1981; Deutsch, 1982; Krumhansl & Kessler, 1982; Brown, 1988; Butler, 1989). Although we may never have been formally trained to do it (see, for example, Moog, 1976; Shuter-Dyson & Gabriel, 1981; Hargreaves, 1986; Serafine, 1988), most of us seem to pick up some skill—some of us have a little, others a lot—at hearing tonal music as arranged around the context of a perceptually central pitch (e.g., Brown & Butler, 1981; Butler, 1983, 1989; Butler & Brown, 1984; Brown, 1988).

## Harmonic Context

The term "harmony" has acquired several meanings. One colloquial definition is roughly equivalent to "good-sounding accompaniment" as, for example, when two people might sing in two-part harmony. Harmony also is often equated with *chord*. On a more abstract musical level, harmony is often used to mean harmonic level, something bigger than a single chord. Music analysts may speak, for example, of a lengthy segment of a composition existing on a single harmonic level, even though it may be represented in some compositions by a sequence of several (or several dozen) chords. Although this chapter does not contain a discourse on harmony on this level of abstraction, it is important to recognize that harmony can carry any or all of these connotations. At both the simplest and the most abstract levels, we can refer to harmony without dealing directly with chords.

Chords, however, are what first come to mind when we speak of harmony. To some extent, chords are similar to the harmonic partials of a complex tone. In fact, many music theorists have argued that chords in Western harmonic practice

are (or ought always be) patterned after the harmonic series; some theorists have gone so far as to call the harmonic series the "chord of nature" (e.g., Catel, 1801). Although it has become a mainstay in some theories of tonal music, this resemblance between *harmonic* partials of complex tones and musical *harmony* is restricted in three important ways. First, the harmonic partials of complex tones are ideally positioned at exact integer relations (1:2:3: . . . :*n*) to one another, but this isn't always true in real musical sound; the partials of some "harmonic" complex tones (such as tones produced by the piano) are actually slightly inharmonic. Second, virtually all fixed-pitch instruments are tuned so that the octave is divided into 12 equal logarithmic increments, and this equal-tempered tuning system produces no intervals other than the octave that match exactly with the intervals between the partials of a complex tone. Third, chords can be built of intervals other than thirds, although chords built in thirds and containing perfect fifths are considered inherently consonant in many accounts of tonal relationships. But tastes change. Only a brief millennium ago, for example, a major triad was considered a clashing discord by many writers: the two outer tones (the fifth) would have been thought acceptable, but the middle tone (the third) was considered a clangorous dissonance. Twentieth-century composers such as Debussy and Hindemith have made consonant sonorities of chords built in fourths and seconds.

There is another aspect of consonance and dissonance beyond the sensory response that two tones may sound either smooth or rough in the presence of one another. In most accounts, dissonance (a "rough-sounding" combination such as a second, a seventh, a tritone, or any chord containing one or more of these intervals) is regarded as a point of musical instability, of musical tension: the listener is expected to respond to dissonant combinations of tones with the feeling that they do not project the sense that the music has "arrived" at a convincing point of repose, a good stopping point. Alternatively, consonant combinations of tones ("smooth-sounding" intervals such as thirds, perfect fourths, fifths, and octaves, or chords limited to them) are presented in these accounts of music as points of stability, good stopping points. Historical and speculative music theory, as well as empirical evidence (e.g., Cazden, 1980; Brown, 1988), offer convincing counterexamples to this point of view. To illustrate, you might try playing a chord containing the pitches $C_3$-$B_3$-$E_4$-$A_4$-$D_5$ (Figure 8.1a); although it contains sensory dissonances in the form of several minor and major seconds (or their octave complements or compounds), the chord forms a perfectly acceptable ending chord in some tonal styles, particularly the big-band and jazz styles of the 1940s and later. Conversely, a major triad built on the dominant scale degree may sound perfectly smooth and agreeable (Cazden used the term "euphonous"), and yet be interpreted by a listener as an unstable, unfinished point within a tonal composition. Because sensory consonance and dissonance are not necessarily equivalent to perceived stability and instability, there is growing acceptance of the use of such terms as "systemic" or "musical" consonance and dissonance to refer to perceived instability (nonfinality) and stability (finality), and "sensory" consonance or dissonance when referring to the smoothness or roughness of the

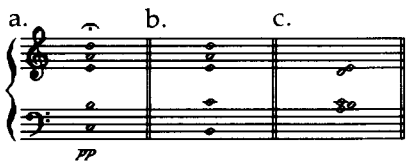



FIGURE 8.1. (a): Simultaneity containing several sensory dissonances forming a stable harmony in some styles. (b) and (c): Simultaneities discussed in note 3, p. 210 containing the same pitch classes but in varied registral arrangements.

combination's sound.[3] Moreover, dissonance means something more than tension and consonance means more than release. Meyer (1956; see esp. pp. 35–42), for example, has argued persuasively that musical events—which may be single tones, chords, entire phrases, or even entire sections of compositions—have "embodied meanings": they point to, or make us expect, other musical events. For example, an expert musical listener, upon interpreting a chord as a dominant harmony, will likely be able to know more than simply that the chord is unstable or tension-arousing; in addition, the listener will quite firmly anticipate that the dominant seventh chord will lead directly (or seldom with more than a perfunctory delay) to a tonic harmony. The essential two points in all of this are that there is a resemblance between acoustical and musical harmony, and that there is a relationship between sensory and musical consonance and dissonance, but the relationships are neither as rigid nor as extensive as they are sometimes portrayed.

## Key Context

Earlier, the term "tonality" was characterized as "keyness": a system of pitch relationships governing how we interpret music to be situated "in" a key, and how we interpret what that key is. Tonality can actually be considered a bigger and more abstract concept than key, however, and in fact even bigger and more abstract than a single *style* of music (Brody, 1985). Tonality is perhaps best considered as an idiom—similar to a dialect, or similar perhaps even to an entire grammatical system. Within the tonal idiom, one can point to a number of different Western musical styles including late Baroque, Classical, Romantic, and some twentieth-century art-music styles, and also including most of the Western popular and folk styles of the last three centuries.

One perceptual feature shared by all these styles is that music written in any of them will give the enculturated listener some sense that the music has a home base, a central and most-important pitch. This pitch stands for the *key* (sometimes called *key level*) of the piece. It is common to speak of the key as the more-general scheme of pitch relationships within a piece of music, and of the mode—major or minor—as the more specialized set of relationships within that key level. For example, Mozart's Symphony No. 40 is in the key of G minor, which is a brief way of saying that its most-important focal pitch is G and the pitch relationships in the minor scale come closer to describing the music than

do the pitch relationships in the major scale. Even though the symphony is said to be in the key of G minor, at various points the listener hears a fair amount of music in B-flat major, one movement in E-flat major (in which the key migrates to B-flat major and then returns), and other sections in even other keys (such as G major and F-sharp minor). In trying simply to describe the perceptual object, then, we must recognize that tonal music is complex enough that the concept of key exists on more than a single level. Returning to the Mozart example, G minor is a global description of the primary key level of the entire four-movement composition, but the listener hears a succession of keys within different parts of each movement of the composition. As complicated as this sort of listening activity seems, both musical and psychological theorists concur (e.g., Brown & Butler, 1981; Browne, 1981; Deutsch & Feroe, 1981; Deutsch, 1982; Boltz, 1989) that musical listeners do interpret tones according to the context of the single key governing the music.

Depending on the experience of the listener and on the specific piece, the perceptual clarity of this central pitch class will vary. Yet (as stated earlier) listeners who can successfully hear musical pitch patterns within the mental context of a key do not necessarily have to have gone through a great deal of formal training in music. Nor do they necessarily have to be able to read musical notation, or have a working knowledge of scales or chords or key relationships. It is not customary to distribute scores to audience members at concerts, nor is it common for the performer(s) to play a tone, a scale, or a chord, as a point of reference with which to prime the audience before each composition is performed. How, then, do we get a mental grip on the music—that is, how do listeners recognize that the music is tonal, and simultaneously recognize the key in that piece of tonal music, given nothing more than an aural presentation of the music itself? Obviously, the code must be located within and among the pitch patterns in the music, rather than inside some abstract construct such as a chord, a scale, or some other set of pitches.

Most accounts of this code have to do with the traditional notion that the *tonic* or keynote is the reference point in any key, it is somehow invested with the most stability of any of the tones in any key, and that all other tones in the set relate to the tonic in varying ways, and to varying degrees. Until recently there has been comparatively little experimental research conducted to explore how we recognize these tonal relationships. So how do we recognize tonic as tonic, dominant as dominant, and so on? In the next section, we examine studies undertaken in the attempt to answer these questions.

## Research on Key Identification

### *Major Mode*

One testable model of this recognition was proposed by Richmond Browne (1981). Browne began with a well-known property of the major pitch set (Bab-

bitt, 1960), which is that each interval class occurs at least once in the set and each class occurs a unique number of times. This is shown in Figure 8.2, represented by the C major set. Although the intervals in Figure 8.2 are notated specifically out of necessity, they are merely a single representation of all possible enharmonic spellings, octave complements, and octave compounds of those intervals.

It should be emphasized that there is an important difference between *scale* and *pitch set:* while pitch sets are not necessarily ordered or grouped in any certain way, scales are conventionally ordered within an octave span, and also ordered so that they form a ladder-like arrangement. The order of tones in any scale is a matter of convention and convenience, but it does not impose any necessary order on the pitches as they may be used in a piece of music.

Although the tones in Figure 8.2 are arranged as scales, the intervallic properties of the major pitch set go beyond the obvious pattern of whole steps and semitones that we see in the scale. For example, consider the intervals formed by tones not adjacent in the scale. Some intervals span three semitones, some span four, and so on. Browne (1981) recognized that the major diatonic set stands out from other pitch-class sets because every interval class is present somewhere in the set—none is omitted. Each occurs a unique number of times. There are two semitones in the set (see Figure 8.2a); no other interval is present that exact number of times. Each interval class is pointed out in each succeeding line within Figure 8.2: two members of interval class (ic) 1 (i.e., semitones, Figure 8.2a) and five members of ic2 (Figure 8.2b); four members of ic3 (Figure 8.2c) and three members of ic4 (Figure 8.2d); six members of ic5 (Figure 8.2e) and

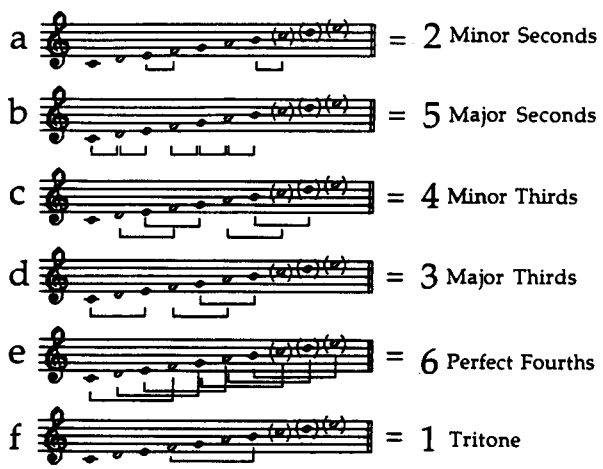



FIGURE 8.2. The various intervals comprising in the C major set (shown in scalar order in this figure). Beginning with the smallest interval other than the unison, there are (a) two intervals of one semitone (interval class or ic1); (b) five intervals comprising a two-semitone span (ic2); (c) four three-semitone spans (ic3); (d) three four-semitone spans (ic4); (e) six five-semitone spans (ic5); and (f) one six-semitone span (ic6).

only one member of ic6 (Figure 8.1f). Octave complements and compounds are assumed within interval classes.

Browne pointed out that while some intervals are abundant, others such as the semitones and the tritone occur infrequently within the major diatonic set. These rarer intervals (a term indicating relative rarity in the set, but not necessarily rare in actual tonal music), Browne reasoned, ought to serve better as perceptual cues that listeners can use to gain a sense of key in music they hear. A series of experiments has produced evidence bearing out Browne's prediction.

In the first study (Brown & Butler, 1981) listeners heard a minimal amount of pitch information, in melodic trichords, from which they made determinations of key center. Choices were limited to the major mode and excluded chromatic interpretations (these will be discussed later). After hearing only three tones in each pattern, listeners were asked to hum, whistle, sing, or otherwise produce a direct judgment of key center. Examples of these three-tone stimulus patterns are shown in Figure 8.3.

Although the pattern in Figure 8.3a looks quite similar to that in Figure 8.3e, the small change of pitch (one semitone) in the third note of the pattern produces a very different intervallic configuration, as can be verified by reviewing Figure 8.2. The rarest interval to be found among the three tones F-G-C in Figure 8.3e is the two-semitone span between F and G. The tones G and C fall within interval class 5, as shown in Figure 8.3e. This means that the intervals between the notes in Figure 8.3 are quite common in the diatonic set. One way of demonstrating this is that we can find that same pattern of ic2 combined with ic5 at five different starting points in the C major scale, as may be seen in Figure 8.4a. Another demonstration (Figure 8.4b–f) is that we can find the specific notes F, G, and C at various points in five different major scales. But the tritone (Fig. 8.3a) occurs just once in the major set; thus there is only one correct major-mode key interpretation for the pitches shown in Figure 8.3a.

When a listener is required to identify appropriate key centers for sequential combinations of tones such as those illustrated in Figure 8.3a and 8.3e, the listener may reasonably expect two things to be true. First, since a rare-interval pattern such as that in Figure 8.3a could fit into only one major pitch set, it

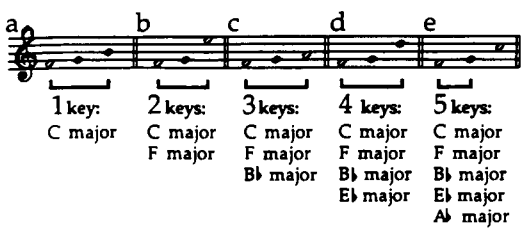



FIGURE 8.3. Transcribed examples of trichordal patterns representative of stimuli used by Brown and Butler (1981). The rarest intervals in each trichordal pattern narrowed its tonal field such that pattern (a) could fit into one major pitch set, pattern (b) into two, and so on.

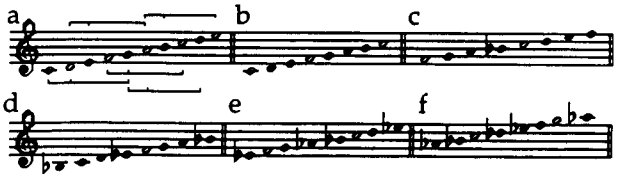



FIGURE 8.4. (a) Five places within the C Major scale where one can find the intervallic combination of ic2 and ic5. (b–f) Conversely, the five scales containing ic2 and ic5 represented by the tones F, G, and C.

ought to be much easier to identify an appropriate tonal center for a pattern such as that in Figure 8.3e, to which one could respond with five different and equally appropriate key centers. Second, although there would be more potential right answers to a pattern such as Figure 8.3e illustrates, listeners' responses might well be spread out among the several appropriate alternatives. Both expectations were borne out in the results reported by Brown and Butler (1981), who asked musicians to sing, whistle, or hum the most appropriate tonal centers for three-tone patterns such as those illustrated in Figure 8.3. Although overall accuracy in key identifications for the common-interval combinations (as represented by Figure 8.3e) were a few percentage points higher than for the rare-interval combinations (as represented by Figure 8.3a), listeners were fully twice as likely to agree on a key center for the rare-interval combinations as for the combinations containing only common intervals.

Another result of this test was that variations in the order of presentation of tones made a difference in the accuracy levels of responses to the patterns containing rare intervals. For instance, listeners concurred more on the choice of tonal center when they heard the pattern shown in Figure 8.3a than when they heard its reverse. The clearest sense of key center seemed to arise when scale step 7 was heard after scale step 4; whenever the temporal order of these two steps was reversed (regardless of the placement of the third tone in the pattern), response accuracy and confidence levels both dropped.

The stimulus patterns drawing the strongest responses (in both accuracy and confidence levels) were not necessarily the patterns adhering to any scalar order, so it was suspected that the patterns carried implied harmonic messages—in other words, that listeners mentally supplied something like a chordal accompaniment as a perceptual framework for these tiny melodic patterns. If this were true, the impact of the varying of time orders of tones on listeners' responses could be explained in terms of common harmonic practices in Western tonal music: it is much more common in typical harmonic successions for the subdominant harmony to precede the dominant than to follow it. If the fourth scale degree were interpreted as representative of the subdominant harmony, and the seventh scale degree (the leading tone) as representative of the dominant, listeners could be expected to feel that they were on firmer tonal harmonic ground hearing the four-to-seven order of presentation.

The second study (Butler, 1983) dealt directly with chordal successions. The test explored this question of temporal order by presenting, as paired dyads, four members of the major diatonic set: scale members 3, 4, 7, and 8(1). The critical variable was the time order in which the tones were presented. In the first of two stimulus categories, set members 4 and 7 were always heard simultaneously, as were scale members 8(1) and 3; these combinations are shown in Figure 8.5a. The 4-7 combination comprises the tritone, the rarest interval in the set; it also is the clearest, most unequivocal presentation of dominant harmony that can be made with a dyad. (Scale members 4 and 7 were combined to imply the dominant harmony in this test, in contrast to the earlier study in which scale degrees 4 and 7 were presented separately, perhaps implying different chords.) The 8(1)-3 combination is the clearest representation of tonic harmony that can be produced with two tones, and—as Figure 8.5a shows—the 4-7 combination can proceed to these two tones by semitone succession. Patterns such as these were presented both to expert musicians and to musical novices, and it was found that as long as the simultaneities clearly represented the dominant and tonic harmonies, as they do in Figure 8.5a, both experts and novices concurred strongly in their choices of the key that each of these patterns represented. On the other hand, both experts and novices were much less accurate in their identification of the key that patterns such as those in Figure 8.5b represented—this even though exactly the same set members were heard, and even though the intervals separating the simultaneities were still semitones. The critical difference apparently was that the simultaneities no longer clearly represented an unequivocal dominant-tonic or tonic-dominant harmonic succession. It seemed clear from this result that pitch relations alone were not sufficient to convey a sense of key to musical listeners; the clarity came when those pitch relations were presented in a temporal order that allowed listeners to interpret those tones within a meaningful harmonic context that the listener tacitly provided.

Subsequent experimental studies have added interesting detail to what we know about the effects of temporality on the perception of pitch relations in tonal (and posttonal) music. Brown (1985, 1987, 1988) demonstrated that when well-known musical excerpts are distilled by deleting repeated pitches, the source pitch set can be ordered so that listeners identify the melody's original key as the key of the new pattern, but when those same pitches are ordered another way, listeners are just as convinced that the pitches belong to a second key. When the pitches are reordered in a third way to induce tonal confusion in listeners, listeners are indeed much less likely to agree on the key for that pattern.

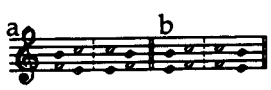



FIGURE 8.5. Transcribed examples of three-tone patterns representative of stimuli used by Butler (1983). Tones were arranged temporally either so that (a) the tritone was presented as a simultaneity, or so that (b) the tritone was presented asynchronously, within different vertical dyads.

Figure 8.6 is an example from Brown's (1988) study. Listeners were 86% accurate in identifying D as the tonal center of an excerpt (see Figure 8.6a) from the third movement of Schubert's Sonata in D Major, D.664, in which tonic and dominant seventh chords are arpeggiated over a tonic pedal. The remaining 14% mistook F-sharp, the last tone heard in the melody, as tonic. When the excerpt (which comprises the diatonic set of D major, missing only the submediant) was reduced to one instance of each within an octave of its component pitches and was ordered to promote identification of the original key (shown in Figure 8.6b), agreement on D as tonic increased to 95%. The pitches were also ordered and presented with the intent of eliciting the choice of G major as the key (Figure 8.6c). While 41% responded to the diatonic set, identifying D as the tonal center, 45% chose G, a key that no listener thought was implied by the musical excerpt or by the ordering of pitches shown in Figure 8.6b. When the sequence was arranged so that rare intervals did not imply clear harmonic progressions in any key (Figure 8.6d), choices of key were indeed more diffuse: Forty-five percent of the listeners chose D, but tonic identifications were scattered among seven other pitches, three of which were not in the tonal pattern (as indicated by whole notes). Responses to the pitch set as ordered in Figure 8.6c and 8.6d indicate that reversal of the single tritone from its optimal 4-7 ordering (the ordering that clearly suggested its major tonic in the pattern shown in Figure 8.6b) greatly diminishes the power of even the rarest interval to suggest a key.

This effect can also be observed in melodic patterns comprising the tones found in a single major triad. In a recent study by Brown, Butler, and Jones (in progress), listeners heard short contextual patterns such as those shown in Figure 8.7, and then rated all members of the chromatic set (which followed the

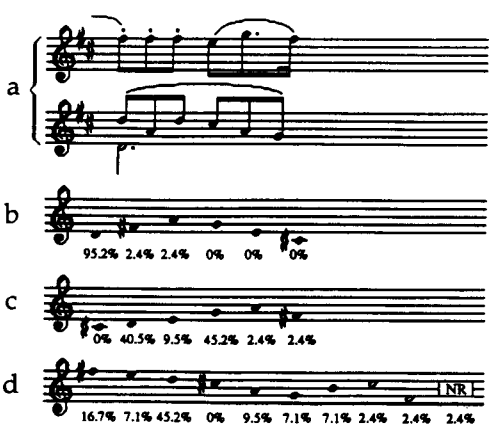



FIGURE 8.6. (a) Excerpt from Schubert, Sonata D. 664, movement III, containing at least one of each interval in the diatonic set and containing no conflicting rare intervals; (b) pitches extracted from the Schubert excerpt and ordered serially to promote identification of D as tonal center; (c) the same pitches reordered to promote G as tonic; (d) the same pitches reordered to be tonally ambiguous. Levels of identification as tonal center for each pattern are given in percentages; whole notes in (d) show out-of-pattern pitches identified as tonic. (NR means no response.) (After Brown, 1988.)

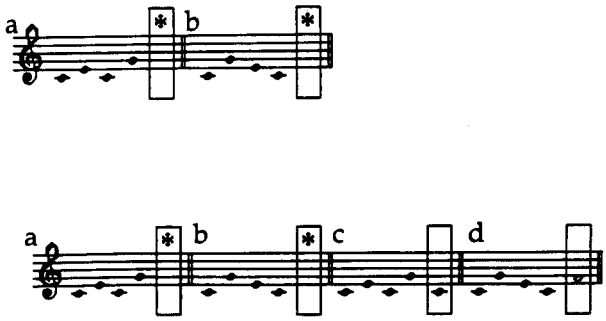



FIGURE 8.7. Tones of a C major triad arranged in two temporal orders (a,b) that preceded all members of the chromatic set (presented at the points indicated by the asterisks). (c) The tone with the highest rating for well-fittedness for the first pattern; (d) the tone rated as best-fitting for the second pattern.

patterns at the point indicated at the asterisk in Figure 8.7a,b); ratings were based on the listeners' estimates of how well each tone fit with the pattern preceding it. Listeners rated the root of the triad as the most appropriate completion when the triad was arranged to reinforce that choice (see Figure 8.7c). When the same contextual tones were rearranged (as is shown in Figure 8.7b,d) to undo the rhythmic regularity of C4 within the contextual pattern, the highest rating indicated an interpretation of the contextual pattern as a dominant chord, with the test tone implying resolution to tonic (Figure 8.7d).

### *Minor Mode*

Our discussion until this point has been limited to tonal harmony in the major mode. This is an important limitation. The perceptibility of other important aspects of tonal harmony has received relatively little experimental investigation thus far. We next consider the topics of minor mode and chromaticism and their relationships with tonality.

One likely reason that most research in tonal harmony has concentrated on the major mode is that it is so much simpler to discuss than the minor mode is. The harmonic form of the minor scale is most relevant to the topic of harmony and tonality because this form has both a leading tone below the tonic and an "upper leading tone" above the dominant. There are three semitones and one augmented second among the adjacent scale members in harmonic minor; this is shown in Figure 8.8.

As Figure 8.8 illustrates, the minor diatonic set has very different intervallic properties from the major set. There are two of the rarest interval class in this set, and only five of the most common. The tritone in major mode is likely to sound unique because it is; the two tritones in minor, however, are obviously not unique. They are inherently more ambiguous in regard to placement in the

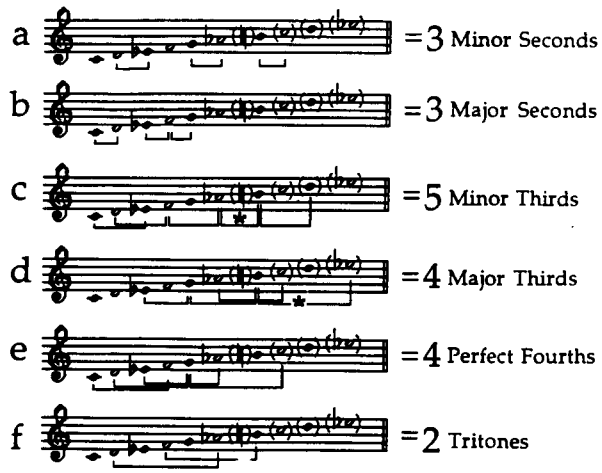



FIGURE 8.8. The various intervals found among tones in the C minor diatonic set (shown in scalar order in this figure). Beginning with the smallest interval other than the unison, there are (a) three members of ic1; (b) three also from ic2; (c) five members of ic3; (d) four members of ic4; (e) four members of ic5; and (f) two tritones (ic6).

pitch set, because any tritone we hear in a piece of music in the minor mode has two potential diatonic points of reference.

There are unique intervals within the six categories shown in Figure 8.8, however. Notice that in Figure 8.8c there are five intervals within interval class 3. Four of these are notated as minor thirds, while one, marked with an asterisk, is notated as an augmented second. Similarly in Figure 8.8d, there are four intervals within ic4: three are notated as major thirds, while the fourth interval is notated as a diminished fourth. Within a musical context (see Figure 8.9) sufficient to clarify the identities of the augmented second and the diminished fourth, it is logical that these rare intervals both assist listeners as they ascertain their tonal whereabouts in music in minor mode. Butler's (1992) data show that this is the case; listeners who heard brief melodic patterns resembling those in Figure 8.9 exhibited almost the level of agreement on minor-mode tonal center as they did for analogous patterns made up of the rarer intervals in the major diatonic set.

## *Chromaticism*

Chromaticism has been equated with randomness or atonality in some perceptual studies (e.g., Taylor, 1976; Krumhansl, 1990). Notice, however, in the second melodic sequence derived from the Schubert excerpt above (Figure 8.6c), in which pitches were arranged to elicit tonic responses of G major, that listeners who did interpret it in that key understood C-sharp to be the leading tone of D,

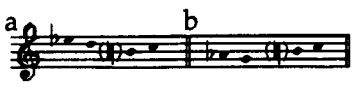



FIGURE 8.9. (a) A melodic pattern that includes a diminished fourth between notes 1 and 3; (b) a melodic pattern that includes an augmented second between notes 1 and 3.

the dominant pitch, which follows it here. Listeners in Brown's (1988) study were able to interpret this C-sharp as a raised fourth scale step that harmonically intensifies the approach to the dominant. This melodic emphasis on a note other than tonic, achieved by chromatically raising the note below it, is known as *microtonicization*.[4] Listeners' responses indicate that, in direct judgments of tonal center, chromaticism is not an element of the tonal fabric that necessarily suggests randomness or atonality. Temporal arrangements and harmonic implications in experimental stimuli can be evaluated as reflective of cognitive processes that enable listeners to incorporate chromaticism into their interpretations of tonality in music. The discussion below draws from both analytic interpretations and experimental data to support this position.

Figure 8.10 contains an example of a set of stimuli (Brown, 1985) manipulated in a way similar to the Schubert excerpt discussed above, but the pitch set of this musical excerpt from Ives's *Concord Sonata* can be reduced neither to a diatonic set nor to any familiar scale. Its interval-class content includes one more minor second than the major diatonic set and one fewer tritone (that is, none). The musical excerpt (Figure 8.10a) consists of an arpeggiated E-flat major triad, in addition to these nonchord tones: F functioning as an appoggiatura on the downbeat, and D and F-sharp functioning as lower neighbors to the root and third of the triad. F-sharp, the chromatic tone that adds two minor seconds to the set, acts as a microtonicization of G both in the excerpt and in Figure 8.10b, the melodic sequence arranged to suggest the key of E-flat to listeners. In the latter context, an arpeggiation of the dominant triad (F–B-flat–D) resolving to the tonic note by the best rare interval available for E-flat, this chromatic pattern elicited unanimous agreement among listeners. In Figure 8.10c, although there is no A-flat and no rare interval for the key of A-flat major in this set, arpeggiation of its dominant triad (B-flat–E-flat–G) ends with its leading tone to suggest resolution to A-flat to listeners. This resulted in 60% agreement on that key while choices of E-flat as tonic dropped to 12%. Responses to the last ordering, shown in Figure 8.10d, clearly illustrate the features of a tonally ambiguous melodic pattern: Responses were scattered among 10 of the 12 possible tonics, with listeners exhibiting the most reliance on the first and last pitches.

With musical excerpts and their reduced melodic sequences that included from as few as three to as many as 10 of the 12 pitches of the chromatic set, results in these studies indicated consistently that virtually any set of pitches could be manipulated temporally either to suggest various tonal centers or to be tonally confusing. Even subsets of a single diatonic pitch set could not be assumed to suggest that key on the basis of nontemporal considerations. In other words, the results indicated that critical relationships among pitches in tonal

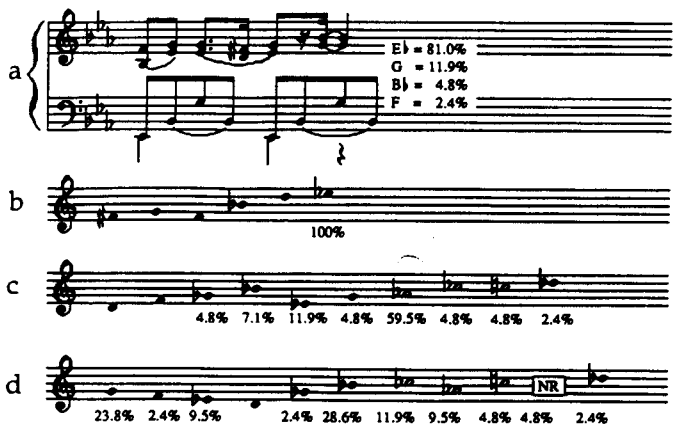



FIGURE 8.10. (a) Excerpt from Ives, Sonata No. 2 (Concord), "The Alcotts"; (b) pitches extracted from the Ives excerpt and ordered serially to promote the identification of E-flat as tonal center; (c) the same pitches reordered to promote A-flat as tonic; (d) same pitches reordered to be tonally ambiguous. Levels of identification as tonal center for each pattern are given in percentages; whole notes show out-of-pattern pitches identified as tonic. (NR means no response.) (After Brown, 1985.)

music do not arise solely from the pitch-set properties of those tones, but arise also to a great extent from their temporal context in the specific pattern being heard.

It is possible, though, for each pitch in a chromatic set to carry a unique meaning that suggests relationships to a tonal center. It takes more than the usual chromatic scale to illustrate this, because tonal music may have only 12 tones per octave, but it has more than 12 functional pitches per octave. In Mitchell's words:

> The ordering force of diatonicism becomes a dispersing force of the chromatic scale. Where each degree of the diatonic major scale is imbued with a single meaning, due to its characterizing interval set and its distance from a tonic degree, the tones of the chromatic scale have multiple meanings. (Mitchell, 1962: 8)

A representation of pitch classes called the "tonal chromatic collection," developed by Brown (1987, 1992), illustrates a distinction among the multiple meanings resulting from mixture and tonicization. According to Proctor (1977), these two processes are the sole sources of chromaticism in diatonic tonality. Although there can be abundant chromaticism in diatonic tonality, the asymmetry of the diatonic basis, with its unique pitch relationships, is preserved. This is distinguished from symmetrical structures in chromatic tonality (Proctor, 1977), where an equal-tempered 12-pitch-class collection is appropriate and where enharmonicism is a tonal, as well as a notational, phenomenon. An example of

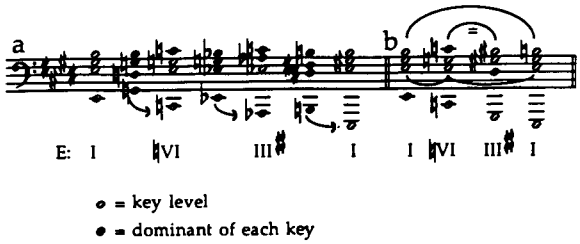



FIGURE 8.11. Synopsis of key areas in Chopin's Prelude in E Major, op. 28, no. 9. (a) Each key level in the prelude is prefaced by its dominant; (b) ties indicate pitches in common among the tonic harmonies.

this symmetrical construction is found in Chopin's E major Prelude, op. 28, no. 9, where key relationships represent a symmetrical division of the octave into major thirds: E, C, A-flat/G-sharp, E. These relationships are shown schematically in Figure 8.11.

The process of chromaticism has also received experimental attention (see, e.g., Brown, 1988, 1992; Butler, 1980, 1988), and these studies indicate that the process involves much more than simply adding nondiatonic pitches to somehow dilute the strength of the prevailing tonality. Chromatic chords can lead skilled listeners to expect that some chords will form more appropriate resolutions than others and can thus be clear signals both of the key level and of unfolding harmonic events at the local level. Until much more experimental research has been carried out, however, it is impossible to say just how unequivocal these signals are, or to say just how many listeners can reliably pick them up.

## Summary

Theoretical discussions of tonal harmony tend to begin with abstractions such as scales, idealized chords, and key signatures. It is important to remember, however, that listeners are forced to pick up information about tonal harmony from one or several tones at a time as the music unfolds perceptually across time. Leading theories of tonal harmony have been only partly successful in explaining how this information is picked up, however. There are two leading reasons for this limited success. First, some of the most influential theories assert (or imply) that musical harmony is intrinsically related to the harmonic frequency relations among partials of periodic tones (e.g., Helmholtz, 1863; Schenker, 1906; Schoenberg, 1911). One hazard of this assertion is that harmony could then be considered solely as a relationship of simultaneous sounds; musical harmony, on the other hand, comprises relationships among sounds that may be successive, or even sounds that may be temporally far removed from one another. Second, theories have emphasized abstractions such as scales and chordal prototypes (e.g., Forte, 1979; Piston, 1987), devoting comparatively little attention to the temporal aspects of harmony and tonality.

The discussion in this chapter, on the other hand, has proceeded on the assumption that the perception of tonal music is inherently temporal—that listeners discover harmonic relations within a key, and track harmonic successions hinting at or actually moving to other keys, as tonal music unfolds in time. This discussion has concentrated on the perception of two facets of musical harmony: first, how listeners recognize the key in which the harmonies reside, and second, how listeners recognize harmonic successions that may or may not be limited to the notes within a single diatonic system. These harmonic successions may be conveyed explicitly by chains of chords, or they may be implicit in a single line—in as few as a handful of tones. Riemann's theory of harmonic function includes this central assertion:

> There are only three kinds of tonal functions (significance within the key), namely, tonic, dominant, and subdominant. In the change of these functions lies the essence of modulation. (Riemann, 1895: 9)

If we are to learn how a listener can recognize any of these three fundamental harmonic functions, we must begin by asking how the listener recognizes the key forming the context for these functions. A number of studies reviewed in this chapter (e.g., Brown & Butler, 1981; Butler & Brown, 1984; Brown, 1988; Butler, 1989) have gathered evidence that listeners get their tonal bearings from the critical intervallic relationships, arranged in conventional patterns across time, that convey the harmonic message of dominant harmony progressing to tonic. If we then ask how a listener is to recognize the change of these functions wherein the essence of modulation lies, we find that the functions change when critical intervallic relationships are taken away from the primary dominant-tonic progression and transposed to some alternate progression. Riemann, in fact, stated that modulation—the change of harmonic functions—resulted from "leading-tone exchange" (*leittonwechsel*). A number of studies cited in this chapter have presented evidence that this leading-tone exchange is recognized when the rare intervals of the diatonic set are encountered in conventional temporal patterns: when one rare interval—the tritone—can indicate a dominant harmonic condition (either on the original dominant, or elsewhere in chromatic harmony), and when other rare intervals—the semitones—lead to the dominant's objective harmony (either the original tonic, or elsewhere in chromatic harmony, or in modulation). These leading-tone relations carry the essence of dominant-to-tonic harmonic progression. They can be fragmented and transposed to produce chromatic harmonic successions that are clearly tonal, or they can be subverted to deny tonality.

## NOTES

1. While we assume that the reader has a solid working acquaintance with traditional musical terminology, some of this nomenclature is problematic because some of the most basic terms have been encumbered with too many meanings. Thus, we will give narrow operational definitions to the terms "pitch," "harmony," and "key," and will discuss their relationships to traditional Western tonality.

2. It is obviously true that most instruments in our musical culture are designed such that they will most readily produce tones separated by twelfth-octave pitch increments. This certainly does not mean that composers are limited to those choices of pitch materials, as compositions by Harry Partch, Easley Blackwood, and numerous others will attest. More germane to this discussion, scales possess no intrinsic properties that force—or even necessarily guide—composers' choices on how the members of the scale are to be arranged across the time-span of the composition.

3. What if the voicing of the chord $C_3$-$B_3$-$E_4$-$A_4$-$D_5$ were changed to $B_3$-$C_4$-$E_4$-$A_4$-$D_5$ (Figure 8.1b)? —or to $A_3$-$B_3$-$C_4$-$D_4$-$E_4$ (see Figure 8.1c)? It might well be that fewer listeners would accept the second voicing as pleasant, and fewer still the third. Although some readers might argue that this undermines Cazden's distinction between sensory and systemic dissonance, other readers could argue that it strengthens Cazden's argument: if the same pitch-classes (and thus the same interval classes) can give rise to different reactions of sensory dissonance, how can sensory consonance and dissonance be a sufficient basis for a rigorous and useful musical theory of consonance and dissonance?

4. The reader is invited to compare the process of microtonicization with Bharucha's (1984) perceptual study of melodic anchoring.

## REFERENCES

Babbitt, M. (1960). Twelve-tone invariants as compositional determinants. *The Musical Quarterly, 46,* 246–59.

Bharucha, J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology, 16,* 485–518.

Boltz, M. (1989). Perceiving the end: Effects of tonal relationships on melodic completion. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 749–761.

Brody, M. (1985). Reply to Serafine and to Marantz on Serafine. *Cognition, 19,* 93–98.

Brown, H. (1985). *The effects of set content and temporal context of pitches on musicians' aural perception of tonality.* Unpublished, doctoral dissertation, Columbus, OH: The Ohio State University.

Brown, H. (1987, October). *Relationships between aural and visual analytic skills.* Paper presented to the College Music Society, New Orleans, LA.

Brown, H. (1988). The interplay of set content and temporal context in a functional theory of tonality perception. *Music Perception, 5,* 219–50.

Brown, H. (1992). Cognitive interpretations of functional chromaticism in tonal music. In M. R. Jones & S. Holleran (Eds.), *Cognitive bases of musical communication.* Washington, DC: American Psychological Association, 139–160.

Brown, H., & Butler, D. (1981). Diatonic trichords as minimal tonal cue-cells. *In Theory Only,* 5(6&7), 39–55.

Browne, R. (1981). Tonal implications of the diatonic set. *In Theory Only,* 5(6–7), 3–21.

Butler, D. (1980, August). *Some remarks on tonality cues and tonal stimuli.* Paper presented at the Third Workshop on Physical and Neuropsychological Foundations of Music, Ossiach, Austria.

Butler, D. (1983). The initial identification of tonal centers in music. In J. Sloboda & D. Rogers (Eds.), *Acquisition of symbolic skills.* New York: Plenum Press.

Butler, D. (1988). A study of event hierarchies in tonal and post-tonal music. *Psychology of Music, 18*(1), 4–17.

Butler, D. (1989). Describing the perception of tonality in music: A proposal for a theory of intervallic rivalry. *Music Perception, 6*, 219–41.

Butler, D. (1992). On pitch-set properties and perceptual attributes of the minor mode. In M. R. Jones & S. Holleran (Eds.), *Cognitive bases of musical communication*. Washington, DC: American Psychological Association, 161–69.

Butler, D., & Brown, H. (1984). Tonal function versus structure: Studies of the recognition of harmonic motion. *Music Perception, 2*(1), 6–24.

Catel, C. S. (1801). *A treatise on harmony*. Novello's Theoretical Series, No. 5. London: J. A. Novello. (Original work published 1854)

Cazden, N. (1980). The definition of consonance and dissonance. *International Review of Aesthetics and Sociology of Music, 11*, 123–68.

Deutsch, D. (1982). Grouping mechanisms in music. In D. Deutsch (Ed.), *The psychology of music*. New York: Academic Press.

Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review, 88*, 503–22.

Forte, A. (1979). *Tonal harmony in concept and practice* (3rd ed). New York: Holt, Rinehart & Winston.

Hargreaves, D. (1986). *The developmental psychology of music*. Cambridge: Cambridge University Press.

Helmholtz, H. L. F. von. (1954). *On the sensations of tone as a physiological basis for the theory of music* (A. Ellis, Trans.). New York: Dover. (Original work published 1863)

Krumhansl, C. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology, 11*, 346–74.

Krumhansl, C. (1990, Spring). Tonal hierarchies and rare intervals in music cognition. *Music Perception, 7*(3), 309–24.

Krumhansl, C., & Kessler, E. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89*, 334–68.

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Mitchell, W. (1962). The study of chromaticism. *Journal of Music Theory, 6*, 2–31.

Moog, H. (1976). *The musical experience of the pre-school child* (C. Clarke, Trans.). London: Schott.

Piston, W. (1987). *Harmony*, (5th ed., rev. by M. DeVoto). New York: Norton.

Proctor, G. (1977). *Technical bases of nineteenth-century chromatic tonality: A study in chromaticism*. Unpublished doctoral dissertation, Princeton University, Princeton, NJ.

Riemann, H. (1895). *Harmony simplified; or, the theory of the tonal functions of chords* (H. Bewerunge, Trans.). London: Augener.

Schenker, H. (1906). *Neue musikalische Theorien und Phantasien. v. 1. Harmonielehre*. Vienna: Universal-Edition A.G. *Harmony* (E. M. Borgese, Trans.; O. Jonas, Ed.). Chicago: University of Chicago Press, 1954.

Schenker, H. (1979). *Free composition* (E. Oster, Trans.) New York: Longman, 1979. (Original work published 1935)

Schoenberg, A. (1911). *Harmonielehre*. Leipzig, Wien: Verlagseigentum der Universal-Edition. *Theory of harmony* (R. E. Carter, Trans.). Berkeley: University of California Press, 1978.

Serafine, M. (1988). *Music as cognition; The development of thought in sound.* New York: Columbia University Press.

Shuter-Dyson, R., & Gabriel, C. (1981). *The psychology of musical ability* (2nd ed.). London: Methuen.

Taylor, J. (1976). Perception of tonality in short melodies. *Journal of Research in Music Education*, 24(4), 197–208.

# 9

# Tonality and Expectation

JAMSHED J. BHARUCHA

## Introduction

Jamshed Bharucha's research has been highly influential in proposing that tonality gives rise to both expectations and aesthetic experiences (Bharucha, 1984). In this chapter Bharucha discusses how musical expectations are generated experimentally. He describes the relationship between tonal expectations, consonance, stability, and memory, and questions whether musical expectations are learned, innate, or based on the structure of the sounds themselves. His interest is in understanding the cognitive processes that underline tonal expectations and their acquisition. Bharucha points out that when listening to music listeners rejoice at what they already know; therefore, he asks, are they giving up the element of surprise or the element of familiarity? For Bharucha, tonal expectations are a compelling aspect of the aesthetic experience of listening to music. He divides expectations into two kinds: schematic expectations, which are culturally generic, automatic, and related to a general schema, and veridical expectations, which are related to the actual piece of music, to a specific memory.

Bharucha explains how he has used neural networks to account for certain findings about music perception. Neural networks are computer models that simulate brain-like systems based on fundamental principles of neural organization: they have been used to model human behavior. The application of neural networks to the perception of music is a promising and fast-growing area of research (see Bharucha, 1987 and 1991; Gjerdingen, 1989; Todd & Loy, 1991).

REFERENCES

Bharucha, J. J. (1984). Anchoring effects in music: the resolution of dissonance. *Cognitive Psychology, 16,* 485–518.

Bharucha, J. J. (1987a). MUSACT: A connectionist model of musical harmony. In *Ninth Annual Conference of the Cognitive Science Society* (pp. 508–17). Hillsdale, NJ: Erlbaum.

Bharucha, J. J. (1987b). Music cognition and perceptual facilitation. A connectionist framwork. *Music Perception, 5,* 1–30.

Bharucha, J. J. (1991). Tonality and learnability. In M. R. Jones & S. Holleran (Eds.), *Cognitive bases of musical communication.* Washington, DC: American Psychological Association.

Gjerdingen, R. O. (1989). Using connectionist models to explore complex musical patterns. *Computer Music Journal, 13*(3), 67–75.

Todd, P. M., & Loy, G. D. (Eds.) (1991). *Music and connectionism.* Cambridge, MA: MIT Press.

---

Why does a segment of music ending in a dominant chord generate strong expectations for the tonic to follow? What is the relationship between tonal expectations, consonance, stability, and memory? Are expectations of this kind learned, innate, or based on the structure of sound? If they are learned, how?

Some of these and related questions about tonality and expectation have been pondered through the ages by music theorists,[1] and today we benefit from a wealth of work on which to build. In this chapter I shall sketch an approach to these questions based on psychological experiments and computational modeling designed to explore the cognitive processes underlying tonal expectations,[2] their relationship to consonance, stability, and memory, and their acquisition.[3]

Tonal expectations are important to study for several reasons, including their role in the aesthetic experience of music, their role in the mental organization of music, their influence on consonance and dissonance, and their role in learning. The advantage of viewing these disparate phenomena in terms of expectation is that there is the promise that the processes that underlie expectations underlie aspects of these other phenomena as well.

The interested reader is directed to a substantial body of important work that will not be reviewed in this chapter, particularly work on the specific musical determinants of expectation (Meyer, 1956; Carlsen, Divenyi, & Taylor, 1970; Narmour, 1977, 1990; Carlsen, 1981; Schmuckler, 1989). The focus in this chapter will be on trying to understand the perceptual processes that underlie

expectation. Toward this end, only a small set of specific musical expectations will be addressed—tonal harmonic expectations—by way of illustrating the potential of certain experimental and computational approaches. The reader is also directed to the work of Jones (1981a, 1981b) for a more overarching theoretical approach to understanding musical expectation.

## Tonal Expectation and Aesthetics

Tonal expectations are a compelling aspect of our aesthetic experience of music. The deception of a deceptive cadence, the irresolution of a half cadence, and the finality of a final cadence are only the most obvious examples of tonal expectations at work in Western tonal music. Delayed resolution of tonal expectations is often a central distinguishing characteristic ascribed to late romantic composers such as Wagner. Tonal expectations can be found in other cultures as well. Some of the most dramatic moments of a performance of Indian classical music involve an eventual resolution to the tonal center after a prolonged yearning for it.

The most influential modern exponent of the role of expectation in musical aesthetics is Leonard Meyer (1956). According to a widely accepted view based largely on the work of Meyer, a piece of music in a familiar genre generates expectations, the subtle violation of which is emotionally important. There is considerable support in psychology for the relationship between expectancy violation and emotion (e.g., Mandler, 1984).[4]

Little, if any, work has been done by cognitive psychologists on the affective aspect of experiences associated with expectation. In contrast, a fair amount of research has been done on the processing of sound that leads up to (and is a prerequisite to) this affective experience. This chapter thus focuses exclusively on trying to understand these prerequisite processes, without taking a position on the nature or interpretation of the affective experience induced thereby.

## The Puzzle of Expectations

There is a puzzle concerning the role of expectations in music that Dowling and Harwood (1986) call "Wittgenstein's paradox": The puzzle is that when we are familiar with a piece of music, there can be no more surprises. Hence if expectancy violation is aesthetically important, a piece would lose this quality as it becomes familiar.

Attempts to resolve this apparent paradox often postulate a mental structure, called a schema (Dowling & Harwood, 1986), that has assimilated the music of a genre over a lifetime of experience. The expectations produced by this system are generic in that they represent the most likely transitions in one's musical culture.[5] They are also automatic, in the sense that they cannot easily be suppressed. Even when a given piece has been heard often enough to be familiar, it

cannot completely override the generic, automatic expectations. Surprises in a new piece thus continue to have a surprising quality because they are heard as surprises relative to these irrepressible expectations.

This line of reasoning seems necessary but cannot be sufficient. If the surprises in a new piece continue to be surprises even after repeated hearing, the piece would never sound familiar. Furthermore, a performer playing from a schematic representation would be unable to predict the next event, but would instead drift into the most generic progression; the performer would, for example, always follow a dominant chord with a tonic. Using this explanation, you either give up surprise or you give up familiarity.

Yet it clearly must be the case that an atypical transition in a familiar piece is both familiar and surprising. It is familiar in a sense and surprising in another. It is thus necessary to postulate two kinds of expectations, which may be called *schematic* and *veridical* (Bharucha, 1987a). Schematic expectations are the automatic, culturally generic expectations mentioned above. They must be generated by a system that has learned to expect the events that are the most likely. Veridical expectations are for the actual next event in a familiar piece, even though this next event may be schematically unexpected. They must be generated by a system that has learned that particular piece. Since schematic expectancies are acquired from hearing many individual pieces, the two kinds of expectancies will converge more often than not for pieces typical of one's musical culture. Sometimes they will diverge, however, creating the sense of violation of which Meyer wrote.

## Expectation and Mental Representation

The existence of two kinds of expectations suggests (but doesn't require) the existence of two systems that generate them. These may be called a *schema* and a *memory*, which generate schematic and veridical expectancies, respectively. A schema is a term used by cognitive psychologists to refer to a mental representation of generic relationships. The memory, in contrast, would represent specific sequences of music in a way that enables them to be retrieved (recognized or recalled) individually. A schema enables one to recognize the style of a new piece; a memory enables one to recognize a particular piece.[6]

A schema generates expectations for events (such as chords) that *typically* follow. A schema thus embodies the information that would generate *typical* hierarchies of stability. These hierarchies, called *tonal hierarchies*, reflect the relative stability of tones or chords in given key-inducing contexts.

A memory, in contrast, generates expectations for events that *actually* follow in a particular familiar piece of music. A memory thus embodies the information that would enable a listener to recognize an error in a familiar piece, and that would enable a performer to play the correct note, however schematically unexpected.

Expected events are heard as more consonant and more stable than unex-

pected events, as will be seen below. Thus expectation, consonance, and stability covary with the musical context. If this is true, these terms refer to slightly different experiential aspects of the same underlying process.

The activation of a schema while listening means that, in addition to hearing the sounded pitch events (along with their temporal, timbral, and other qualities), schematic expectations are generated and are fulfilled or violated to varying degrees, with concomitant effects on stability. The schema thus imposes a cultural elaboration on the sounded events. Elaboration is widespread in perception and cognition (see Anderson, 1990). While reading a story, for example, inferences are drawn and familiar themes are triggered. These happen automatically, unavoidably, and constitute a mental elaboration of the text.

The representations postulated by Lerdahl and Jackendoff (1983) and by Deutsch and Feroe (1981) are essentially representations of mentally elaborated pieces of music. According to Lerdahl and Jackendoff, a piece of tonal music is represented as an inverted tree, of which the pitch-time events of the piece are at the terminal branches (see Figure 9.1). An inverted tree diagram shows levels of abstraction from the piece of music (eighth-note level, quarter-note level, etc.).

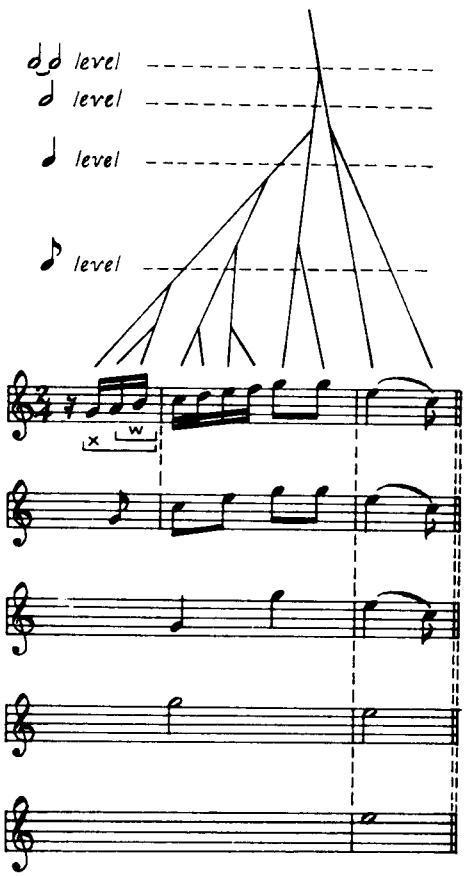



FIGURE 9.1. An inverted tree diagram shows levels of abstraction from the piece of music (eighth-note level, quarter-note level, etc.). A branch that terminates at a longer branch represents the subordination of an event to a more stable event. The most stable event (E in the last measure) is represented by the uninterrupted trunk of the tree. The increasing levels of abstraction are written out in musical notation below the score of the piece. (Reprinted from Lerdahl & Jackendoff, 1983. Copyright © 1983 by M.I.T. Press.)

to maximize the percent of trials they get correct (which is displayed at the end of the experiment).

Another reason for employing an intonation judgment is that it enables us to study how the perceived consonance of a chord is influenced on the basis of whether or not it was expected. To summarize, the priming task using an intonation judgment has two goals: (1) to measure expectation using reaction time, and (2) to study the effect of expectation on consonance. The task is designed so as to pursue these goals for musically untrained as well as trained listeners.

From trial to trial, the experimenter varies the degree to which the target chord is expected[8] given the prime. In the simplest such experiment (Bharucha & Stoeckig, 1986, experiments 2 and 3), the prime was a single chord, and the target was either most highly expected or most highly unexpected given the prime. Thus if the prime was C major, the target was either G major or F-sharp major, as shown in Figure 9.2 (all keys were used in the experiment).[9]

This yielded four types of trials: in-tune expected, out-of-tune expected, in-tune unexpected, and out-of-tune unexpected. The listeners only needed to concern themselves explicitly with whether or not the target was in-tune or out-of-tune.

The results were as follows. First, listeners were faster and more accurate, on average, to make intonation judgments for the expected chords than for the unexpected chords (see table). This demonstrates that the process underlying expectation is one of preparation or facilitation. It is as if a representation of the expected chord had been mentally activated even before the chord occurred, based on its being highly probable in the familiar tonal style.

The second central result was that listeners were faster and more likely to respond "in-tune" when the chord was expected than when it was unexpected, or, equivalently, they were faster and more likely to respond "out-of-tune" when the chord was unexpected than when it was expected. This indicates that expected chords sounded more consonant, or, equivalently, unexpected chords sounded more dissonant. This result shows starkly that there are at least two determinants of consonance. One is the relationship between simultaneous sounds. This in turn is based on the way sound interacts with the inner ear (see below) as well as learning. The second determinant is the relationship between

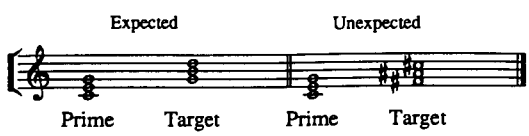



FIGURE 9.2. The design of the priming experiment (Bharucha & Stoeckig, 1986). The figure shows root position triads in close position only to illustrate the difference between expected and unexpected. The actual chords heard by the subjects were considerably more complex (see note 9).

Results of priming experiment (in milliseconds)

| | *Expected* | *Unexpected* |
|---|---|---|
| In-tune | 788 | 982 |
| Out-of-tune | 864 | 803 |
| Average | 826 | 893 |

two sequential sounds. This, it is claimed here, is primarily learned and is mediated by expectations.

Both effects described above were observed for rank novices and professionals alike. The novices were novices only insofar as they had no formal training and no experience playing an instrument. But by virtue of having grown up in a particular culture (in this case, the United States), they have presumably learned what to expect from extended passive exposure.

It is important to note that expected events aren't necessarily preferred over unexpected events. This caution about expectation holds true of consonance and stability as well. If there is any relationship between expectation and its associated concepts, on the one hand, and preference, on the other, it takes the form of an inverted-U function (see Dowling & Harwood, 1986, for a discussion): A moderate amount of violation of expectations is generally preferred over always fulfilling expectations or always violating them.

Even this trend should be viewed with some caution. People who are immersed in experimenting with new musical forms may develop preferences based on phenomena other than expectation, consonance, and stability. Although a composer may have exposure to several genres, including some that are very different from tonal-harmonic music, and would thus develop a rather diverse set of schemata, it is almost impossible to escape exposure to tonal-harmonic music if one grows up in a Western society. Tonal-harmonic music pervades television, radio, supermarkets, and the dominant popular forms of music. Some of the subjects in the above priming experiment were composers who viewed the experiment with grave suspicion, seriously doubted they would show any evidence of schematic tonal expectations, and asserted their lack of preference for expected tonal events. Nevertheless, they showed the same pattern of results as novices. A person is in a sense a prisoner of the schematic expectations driven by the inadvertent learning of musical patterns that pervade the environment.

This means that it is difficult to suppress schematic expectations. Can this be demonstrated empirically? In one experiment, 80% of the trials contained highly unexpected chords and only 20% contained highly expected chords. Even over a session of 100 trials, the chords that were processed more quickly were those that were traditionally expected even though they were in the clear minority during the session (Bharucha & Stoeckig, 1989). Distorting the musical environment in this way was not sufficient to wipe out a lifetime of exposure.

In another experiment, each prime-target pair was played twice in succession. For example, in an "unexpected" trial, the subject would hear C-F-sharp-pause-C-F-sharp. The last chord was mistuned on half the trials, and subjects had to decide whether it was in-tune or out-of-tune. The last chord, F-sharp, is schematically unexpected (since it rarely follows C) but veridically expected (since it actually follows C and the subject knows the same pair will be repeated). Nevertheless, schematically expected chords are still processed more quickly than schematically unexpected chords (Bharucha & Stoeckig, 1989). These experiments thus provide some preliminary evidence that schematic expectations are very difficult to suppress.

## Expectation and Tonal Hierarchies

There has been a more extensive study to date of tonal hierarchies than of expectations. Since one of the claims in this chapter is that the processes that underlie schematic expectations may also be responsible for tonal hierarchies, a summary of research on tonal hierarchies is relevant.

Tonal hierarchies have been studied most exhaustively by Krumhansl (1979, 1990; Krumhansl & Shepard, 1979). The term "tonal hierarchy" signifies the relative differences in stability of tonal elements. For example, when the key of C major has been strongly established during a piece of music, the diatonic tones (C, D, E, F, G, A, B) are more stable than the nondiatonic tones. If a different key is established, a different set of tones will be the stable diatonic set. Among the diatonic set in any given key, the tonic is the most stable, and the major third and perfect fifth are generally more stable than the remaining four.

Although this tonal hierarchy for tones in the context of an established key is well known from elementary music theory of the common practice period, Krumhansl sought to quantify this hierarchy. Among the many advantages of quantifying a tonal hierarchy are the following. First, although the tonal hierarchy is quite clear when a key is strongly and unambiguously established, it is less clear when a key is weakly established or when there is more than one established key (as in polytonality). A quantification of tonal hierarchies makes it possible to track the development of keys even in the latter case. For example, there has been considerable debate as to whether the opening passage of Stravinsky's *Petroushka* is simultaneously in C major and F-sharp major or, alternatively, in no key at all. By quantifying the tonal hierarchy in this passage, Krumhansl and Schmuckler (1986) were able to determine that indeed both keys, C major and F-sharp major, were established.

A second advantage (perhaps the more significant psychologically) of quantifying the tonal hierarchies is that it makes it possible to depict tonal relationships spatially. For example, the circle of fifths is a spatial depiction of the relationships between major keys (see Figure 9.3). A single spatial representation such as a circle captures, all at once, the many pairwise relationships between

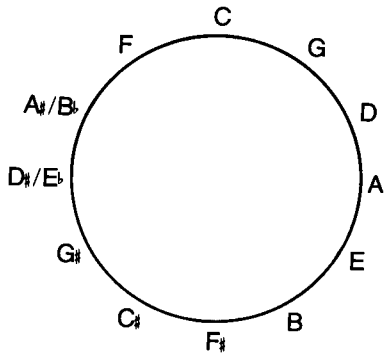



FIGURE 9.3. The circle of fifths.

the 12 possible major keys. Thus, instead of individually enumerating the relationship between all the keys (e.g., C is most closely related to G and F, less closely related to D and B-flat and so on, G is most closely related to D and C, less closely related to A and F, and so on, and so forth), a single diagram with the 12 keys labeled shows all the relationships simultaneously. The advantage is even more striking if minor keys are included. A single spatial representation that captures all these pairwise relationships at once provides a vivid insight into the psychological representation of tonality. That is not to say that the same spatial representation necessarily exists physically in the brain.[10] Rather, it means that the physical representation in the brain encodes the information that is contained in the spatial representation.[11]

At this point one might be inclined to ask why quantification of tonal hierarchies is an advance, since music theorists have a long tradition of tracking keys and employing spatial representations (the circle of fifths, in particular). The answer is that methods of quantification permit us to investigate tonal hierarchies in the minds of listeners other than theorists, who, after all, have extensive formal training and who, by the very nature of the enterprise, adopt a highly analytical stance toward listening.

The method Krumhansl has used most extensively to investigate tonal hierarchies is called the *probe tone technique*, originally developed with Roger Shepard (Krumhansl & Shepard, 1979). In this technique subjects hear a musical context followed by a probe tone. The musical context could be either a musical scale, a chord, a chord progression, or an actual sample of music. The *probe tone* is picked at random from the 12 chromatic pitch classes. The listener's task is to rate how well the probe tone fits within the musical context by rating it on a scale from 1 to 7. The context is then repeated and a different probe tone is presented. The context has been fully probed when all 12 chromatic pitch classes have been presented. This process may be repeated for the same listener and also repeated for many listeners, yielding several replications of each of the 12 probe tones. The ratings for each probe tone are averaged across the replications, yielding 12 numbers, one for each probe tone. This set of 12 numbers is called a *profile* (see Figure 9.4).

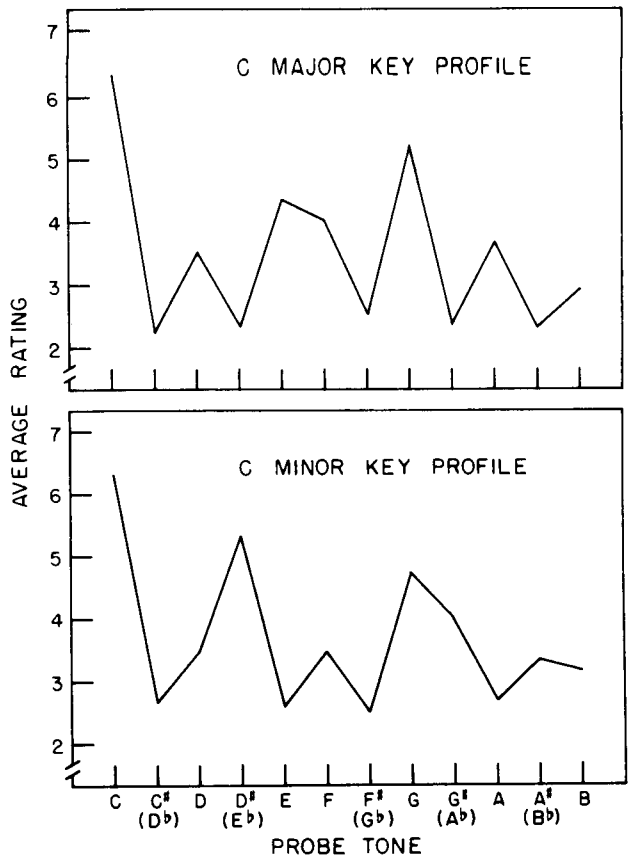



FIGURE 9.4. Profiles for major and minor keys. The rating for any given probe tone depicts the perceived stability of that tone in that key. The profiles are shown with respect to C major and minor for convenience only. (Reprinted from Krumhansl & Kessler, 1982. Copyright © 1982 by the American Psychological Association. Reprinted by permission.)

When the musical sample establishes a key strongly and unambiguously (e.g., a IV-V-I cadence), the profile shows the tonic being the most highly rated, followed by the other two tones of the tonic triad, followed by the remaining diatonic tones (Krumhansl, 1979, 1990). The profiles from samples with a strong and unambiguous key are averaged and the resulting profile is used as a standard profile for that key. The standard profiles for the 24 major and minor keys constitute the standards against which the profiles of samples with weaker or more ambiguous keys are assessed.

Krumhansl and Kessler (1982) showed how one can track the development of keys in a musical sample of interest. To do this, the profile obtained for that sample is correlated with each of the 24 standard profiles. The strength of the

sample's correlation[12] with a key standard indicates the strength of that key in the sample as of the moment at which the probe was presented. By probing a piece of music at different points in time, the development and change of keys over time can be tracked.

To represent the relationships between keys spatially, the standard profiles for the 12 major and minor keys are themselves correlated with each other. The higher the correlation between two keys, the closer they are in the spatial representation. Thus C major is highly correlated with G major and F major, as well as with A minor and C minor, and therefore needs to be located close to them. In addition, it has a range of correlations with the other keys, which must therefore be located at corresponding distances. Simultaneously satisfying all these constraints is very difficult. This is accomplished by a computerized method called multidimensional scaling, developed by Shepard (1964). Krumhansl and Kessler (1982) found that the spatial representation that best satisfies these multiple constraints is a torus. A torus requires four dimensions, but for convenience can be imagined in three dimensions as a doughnut, with the proviso that the inner and outer perimeters of the doughnut be equal (which is why three dimensions don't suffice). This solution captures the circle of fifths for major keys and for minor keys, and puts the two together in such as way as to preserve the relative and parallel major-minor relationships. The relative positions of keys on the torus can be examined by cutting it and unfolding it into a flat sheet, as shown in Figure 9.5. A torus results from joining the top and bottom edges

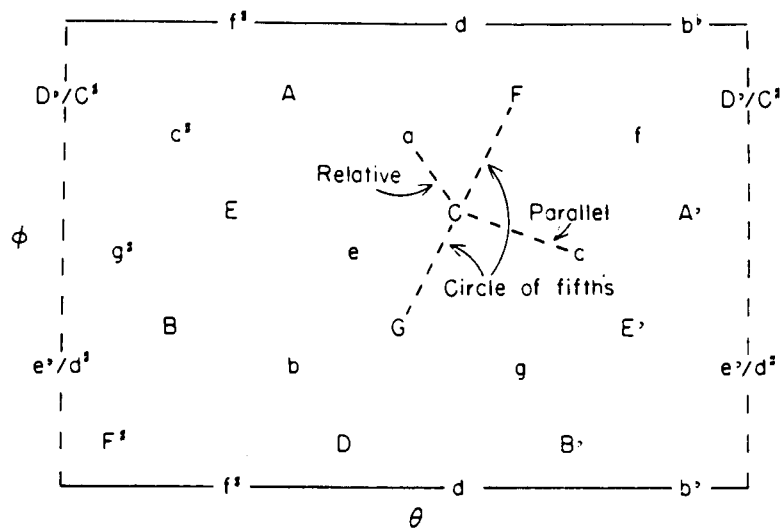



FIGURE 9.5. An equivalent two-dimensional map of the multidimensional scaling solution of the 24 major and minor keys. (Reprinted from Krumhansl & Kessler, 1982. Copyright © 1982 by the American Psychological Association. Reprinted by permission.)

together and joining the left and right edges together. If you actually try this with a piece of paper, it will either wrinkle or rip, which is why an additional dimension is needed.[13]

Tonal hierarchies usually, but not always, reflect the frequency of occurrence of tones just heard: tones heard more often are judged to be more stable. When American college students were tested with samples of North Indian music, for example, their profiles were very similar to those of Indian students (Castellano, Bharucha, & Krumhansl, 1984). The profiles of both groups of listeners matched the hierarchies described by theorists of Indian music (Jairazbhoy, 1971), but they also reflected the frequency of occurrence of tones in the samples. This should not be surprising: Tones that are heard as stable have typically occurred more often in the preceding sample. Numerous studies of note counts (summarized extensively by Krumhansl, 1990) show this to be true. In tonal music of the common practice period, tonics tend to occur most often, followed by the other tones of the tonic triad, and then the remaining diatonic tones. Tonal pieces tend to reveal their tonality in the selection of pitch classes and their relative priority. Tonal hierarchies obtained in these experiments thus show, at the very least, that listeners have a memory system that can integrate the pitch classes heard in the immediately preceding sample (called the *proximal context*). This integrated representation then leads them to hear pitch classes that immediately follow as being differentiated on the basis of stability.

The distribution of pitch classes in the proximal context is not the sole determinant of the tonal hierarchy. The profiles of the Indian students in the study by Castellano, Bharucha, and Krumhansl (1984) were found to be influenced by an additional factor: past exposure (called the *distal context*) to these musical forms. This determination was made by a regression analysis. A regression analysis enables one to determine the extent to which different factors contribute individually to an outcome. After determining that frequency of occurrence of pitch classes in the proximal context was the strongest determinant of the profiles for both American and Indian listeners, frequency of occurrence was factored out, and the residual portions of the profiles were analyzed for the influence of additional factors. The residuals of the Indian listeners correlated with profiles based on music theory (Jairazbhoy), whereas the residuals of the American listeners did not. Since the essential difference between the two groups of listeners was their background, it can be concluded that the ratings given by the Indian listeners were based in part on their familiarity with that genre and not just on the information contained in the proximal context.

It is this residual effect of the distal context—that is, prior exposure—that justifies the postulation of a schema, a mental structure that encodes the regularities of a genre as opposed to the specifics of individual pieces. There must nevertheless be enough information in the proximal context to engage the appropriate schema; after all, we have schemas for many things, possibly even many different musical genres. The information in the proximal context (such as

relative occurrence of different pitch classes) serves to activate the appropriate schema, which then generates expectancies that may compensate for the absence of tones that typically occur in those contexts.

There may be other determinants of tonality besides the distribution of pitch classes in either the proximal or distal context. Among them are the rareness of the intervals (see Chapter 8 by Butler and Brown in this volume).

## Can Schematic Expectation Be Explained by the Structure of Sound and the Physiology of the Ear?

Consider two tones an octave apart, each containing eight harmonics. Every harmonic of the higher tone is present in the lower tone, and every second harmonic of the lower tone is present in the higher tone. This can be represented by the ratio 1:2, which, for convenience, we shall call the *overlap ratio*. Now consider two tones a perfect fifth apart. Every second harmonic of the higher tone is present in the lower tone, and every third harmonic of the lower tone is present in the higher tone. This interval has an overlap ratio of 2:3. For two tones a perfect fourth apart, every third harmonic in the higher tone is present in the lower tone, and every fourth harmonic in the lower tone is present in the higher tone. This interval has an overlap ratio of 3:4. If you are familiar with how these intervals sound, you know that an octave is the most consonant interval, followed by the perfect fifth, followed by the perfect fourth. As the overlap ratio gets more complex, the intervals sound more dissonant.

Pythagoras, the ancient Greek mathematician and philosopher, arrived at the same ratios more than 2,000 years ago on the basis of the lengths of stretched strings. A string player produces an octave above an open string by stopping the string at half its length, and produces a perfect fifth by stopping the string at two-thirds of its length, and so on. Since Pythagoras couldn't have known much about harmonics, and even less about the workings of the ear, he concluded that there was something inherently special about simple ratios. Today the most compelling account of why consonance is related to ratio simplicity is in terms of how the harmonics affect the inner ear.

The inner ear serves as a prism for sound: it separates sound into its various frequency (pure tone) components. For example, a tone with eight harmonics will cause the basilar membrane in the inner ear to vibrate at eight different places along its length. The vibration of the basilar membrane causes hair cells that are lined up against it to respond. The response of these hair cells thus constitutes a (spatial) representation of the sound spectrum.

Two frequencies that are very close together cause the basilar membrane to vibrate at two points that are close together. Under these circumstances, the membrane's response to one interferes with its response to the other. This produces a sensation of roughness that is typically heard as dissonant. Thus the more

complex the overlap ratio between two tones, the greater is the extent to which the vibrations of the basilar membrane caused by one will interfere with those caused by the other, and the more dissonant the tones will sound together.

If the consonance of intervals can be explained by spectral overlap, can other relationships in musical harmony be explained in a similar way? At first glance, this approach might seem promising. Chords that are considered most closely related tend to have tones in common—e.g., tonic-dominant, parallel major and minor, and relative major and minor relationships. In contrast, chords that are considered most unrelated have no tones in common, C and F-sharp major. Could it be that schematic expectations are simply driven by common tones?

To answer this question, the priming paradigm mentioned earlier was employed. In the original priming experiments, the expected target chord was one step away from the prime chord along the circle of fifths; for example, if the prime was C major, the target was G major. Notice that these two chords share a component tone (as well as all the harmonics of this shared tone). In contrast, the unexpected target was diametrically opposite the prime chord on the circle, and these two chords share no component tones. Thus the faster and more accurate responses to the expected chord could possibly be due simply to the fact that some of its tone components had just been heard. To eliminate this problem, it was necessary for the expected target to have no shared tones with the prime. This condition is easily met by selecting a target two steps away from the prime along the circle of fifths. For example, a C major prime leads to a greater expectation for D major than for F-sharp major, yet shares tones with neither. To strengthen the case, all harmonics that were common to the prime and target were removed, thereby ensuring that the expected target wasn't favored because of its physical similarity to the prime. In spite of this, the expected chord was processed faster and more accurately and was more likely to be judged in-tune than the unexpected chord (Bharucha & Stoeckig, 1987).

There are purely theoretical reasons to reject spectral overlap as the sole determinant of schematic expectations. Consider the following pairs of major triads: C-D versus C-A. C and D do not share any tones, whereas C and A share one tone (E). Thus C and D have less spectral overlap than do C and A. Yet C and D are considered to be more closely related to each other than are C and A, because they are closer along the circle of fifths of chords (see Figure 9.3).

Before rejecting spectral overlap as the basis for harmony, however, perhaps we should question the validity of the circle of fifths for chords. The circle of fifths is typically used to describe the relationship between keys (of the same mode). The (shortest) number of steps from one key to another along the perimeter of the circle represents exactly the number of tones present in the diatonic set of one key but not in the diatonic set of the other. For keys (of the same mode), then, spectral overlap (i.e., the intersection of the diatonic sets of different keys) exactly mirrors the circle of fifths. For chords, however, it does not. Yet the circle of fifths accurately describes the harmonic distances between chords, because

chords closer on the circle are more likely to occur in the same tonal piece (they have more parent keys in common).

Since we cannot rest our case on the mere use of the circle of chords as a descriptive device, a test is needed. The appropriate test is to see whether a chord generates stronger expectations for a chord that is closer on the circle of fifths but doesn't share a tone than for a chord that is further but does share a tone. This experimental design pits convention against acoustics in determining expectations. If schematic expectations are acoustically driven, C major would generate stronger expectations for E major (with which it shares a tone) than for D major (with which it shares no tones).[14] If, on the other hand, schematic expectations are driven by a schema that has encoded the chord relationships on the basis of their conventional typicality, C major should generate stronger expectations for D major than for E major. Recent results support the latter outcome (Tekman & Bharucha, 1991).

If schematic expectations are driven by spectral overlap, then any impairment of one's ability to process auditory spectra ought to be accompanied by a corresponding diminution of schematic expectations while listening to music. A man suffered a stroke that destroyed much of the auditory cortex of his brain. Although he could detect the presence of sounds at almost normal levels, he simply could not tell the difference between the in-tune and out-of-tune chords used in the priming experiments. He was unable to detect a mistuning so severe that people with normal hearing cringe almost without exception. In the priming task, his performance was never much above 50%, which would be expected from sheer guessing. What is remarkable is that he tended to judge expected targets to be in-tune and unexpected targets to be out-of-tune (Tramo, Bharucha, & Musiek, 1990). In other words, his sense of schematic expectation was intact even though his ability to process fine-grained spectral information was severely impaired. It seems that as long as some auditory information get through, the relevant schema can be activated, thereby generating schematic expectations. This is good news for people who experience a diminished ability to process the full range of the audible spectrum, as is typical with age.[15]

Even the relative consonance of simultaneously sounded tones cannot entirely be explained by the structure of sound and the mechanical characteristics of the ear. For otherwise one should be able to synthesize artificial sounds that have very different spectra, as Pierce (1969, 1983) has done, and predict the relative consonance of intervals formed with these tones. Although there is evidence for this using tones with nonharmonic partials that are so discrepant from harmonics that the tone has no clear pitch (Geary, 1980), cases in which the tones have clear pitch reveal an overwhelming tendency to hear the intervals as we are accustomed to.

Consider a clarinet tone, which has very little energy at the even-numbered harmonics. If the fundamental frequency is 440 Hz, the next harmonic would be $440 \times 3 = 1320$ Hz. The tone that should sound most consonant with the 440-

Hz tone is one whose fundamental frequency is 1320. All the harmonics present in the higher tone would also be present in the lower tone, and every third harmonic in the lower tone would also be present in the higher tone, yielding a 1:3 ratio of spectral overlap. This yields an interval of an octave and a fifth. Two clarinet tones an octave apart have no spectral overlap at all. If spectral overlap is the sole determinant of consonance, two clarinet tones should sound more consonant when they are an octave and a fifth apart than when they are an octave apart. Indeed, many other intervals should sound more consonant than an octave. Although this has not been tested with a careful laboratory experiment, the clear sense from hearing clarinet tones is that the intervals that are the most consonant (the octave, for example) are the same as with other instruments. Once again, this must be because of the effect of prior exposure to other instruments and human speech, which do contain even-numbered harmonics. One can only speculate about how the various intervals would sound if our only auditory exposure was to clarinet-like timbres.

One is compelled to conclude that there is much more to the perception of harmony, and the schematic expectations that are such a salient part of it, than is determined by the structure of sound and the mechanical properties of the ear. But it is a mistake to discount these factors completely. Since the time of Helmholtz (1885/1954), the notion that relationships in harmony can be explained by aspects of the structure of sound and processes in early audition has been a leading hypothesis and remains so today (see Parncutt, 1989). It would be surprising if the conventions that have been established were chosen arbitrarily. Factors such as spectral overlap and the consonance and dissonance it engenders may have played a role in the early establishment of the conventions that today drive our perception. One can imagine musicians in ancient and medieval times experimenting with combinations of tones and finding certain combinations to be more consonant than others, based entirely on spectral overlap. The ensuing choices then became convention, and today we are so inundated with music that adheres to those conventions that our internalization of those conventions can compensate to some extent when spectral overlap fails.

## How Expectations May Be Learned: Neural Net Modeling

The above evidence forces us to conclude that tonal expectations must be either innate or learned. There are obvious problems with the innateness hypothesis. First, although it is axiomatic that we must be innately endowed with a mechanism capable of supporting schematic expectations, it is difficult to think of the evolutionary pressures that would have given rise to the specific schematic expectations that occur in the perception of Western tonal music by Western listeners.

Cross-cultural studies would, of course, shed some light on how different or similar schematic expectations are around the world.[16] It is important to note,

however, that universality does not imply innateness (see Bharucha & Olney, 1989). Studies with infants can only be suggestive because perceptual learning can occur very early, possibly even in utero.

Perhaps the most compelling argument against the innateness hypothesis is the existence of a plausible learning hypothesis. If schematic expectancies can be easily learned by mere exposure, there is no reason for them to have been wired innately. Much of the work in music cognition (as well as work on artificially intelligent musical systems) consists of postulating rules. This approach is most pronounced in the case of grammars. This work is invaluable because the rules serve as crystallizations of the constraints on the system. In particular, they characterize the knowledge that a musical schema must exhibit. Psychologists are often interested in also understanding the mechanisms that underlie the encoding of this knowledge. From this point of view, it is interesting to ask how the properties that are captured by rules are learned by passive exposure.

Recent advances in our understanding of how brainlike systems might learn provide an opportunity to explore how musical schemas might be learned (Rumelhart & McClelland, 1986). Models of brainlike systems are broadly referred to as neural nets, connectionist models, or parallel distributed models. These are, in turn, a special subset of a larger class of models that may simply be called associative networks.

In associative models, objects or features in the world are represented by units (or nodes). Representational units are linked together to form a network. The network as a whole represents the complex array of relationships between the objects represented by the units. The state of activation of a unit represents the degree to which the represented object is attended (suggested, implied, or expected).

If a network contains units that represent musical objects such as tones and chords, the connections between these units collectively encode the relationships that hold between these tones and chords as they are typically used. The state of activation of a unit represents the degree to which that tone or chord is currently being attended. A unit may be activated directly by a stimulus in the environment, such as a tone or chord, or indirectly through an associative network of such units. When a unit is indirectly activated, that tone or chord is suggested, implied, or expected.[17]

There are many arguments in favor of understanding music cognition in terms of associative networks. First, they easily accommodate ambiguities, with which music is rife. Many units can be active simultaneously to varying degrees, representing varying degrees to which different tones, chords, or keys are suggested. Second, they fill in missing pieces that are parts of typical patterns. For example, if the root and fifth of a triad are played, a network will suggest the major and minor third in proportion to their association with the prior context. Or if an F major chord is followed by a G major chord, the representational unit for C major will be strongly activated, representing a strong expectation for C major.

There are at least two additional arguments in favor of neural net models specifically. First, they are based on fundamental principles of neural organization in the brain. Second, and most important, they can learn.

Learning in a neural net model occurs by changing the strength of the connections between representational units. Consider a neural net that receives its input via units that represent octave-equivalent pitch classes (shown at the bottom of Figure 9.6a). Each such unit is said to be tuned to a particular pitch class. This is a valid assumption given that there are neurons in the auditory system that are tuned to specific frequency ranges and are laid out spatially according to frequency.[18]

Now suppose that this set of units is connected to another set of units that initially do not have any specialization (labeled "unspecialized" in Figure 9.6a). Each pitch class unit is connected to each unspecialized unit (the connections are not shown in the figure), such that the strengths of the connections are initially random. When a combination of tones is sounded, say an F-sharp major chord, the units representing the tones F-sharp, A-sharp, and C-sharp are activated (as shown by filled circles in Figure 9.6b). They in turn activate the unspecialized units. Since the connection strengths are initially random, the activations of the activation levels of the unspecialized units will be random.

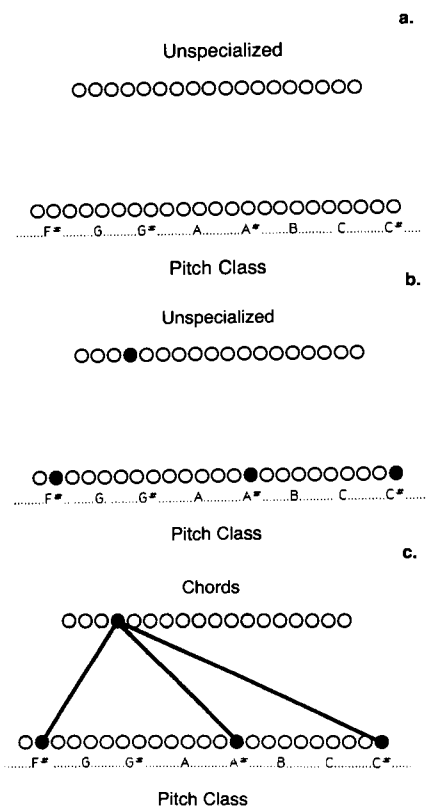



FIGURE 9.6. A neural network learning to represent typical combinations of tones. (a) Pitch class units in the bottom layer are connected to unspecialized units in the top layer via connections with random strengths. (b) In response to a combination of tones (F-sharp, A-sharp, C-sharp), the corresponding pitch class units are activated, and in turn activate the specialized units, one of which wins over the others. (c) The connections between the activated pitch class units and the winning unspecialized unit are strengthened, producing a unit that is now specialized for that combination of pitch classes.

A principle found in many parts of the brain is that units within a layer inhibit each other in proportion to their activation. The consequence of this is that only the most highly activated unspecialized unit remains active, and the others are shut down. So in response to the F-sharp major chord, some arbitrary unspecialized unit (the filled circle in Figure 9.6b) is going to win over the others in that set.

At this point, learning occurs by changing the strengths of the connections between the pitch class units and the winning unspecialized unit. The strengths are changed according to Hebb's (1949) rule: A connection becomes stronger in proportion to the product of the activations of the two units it connects. Since the F-sharp, A-sharp, and C-sharp pitch class units are the only pitch class units activated, their connections to the winning unit become strengthened (indicated by bold lines in Figure 9.6c) and the other connections to the winning unit become weakened. This winning unit has now become specialized to respond to F-sharp major chords. It will be even more likely to win the next time an F-sharp major chord occurs and will be even less likely to win in response to some other chord.

In this way, specialized representational units can form for all combinations of musical events that occur with great regularity. This process can be continued to yet another set of units to which the newly formed chord units are connected. This third set can become specialized to respond to typical combinations of chords, such as tonic, dominant, and subdominant.

Consider an idealized end-product of this learning scheme. It consists of pitch class units connected to units representing their parent chords, and chord units connected to units representing their parent keys, as shown in Figure 9.7 (see Bharucha, 1987a, 1987b, for details). The minor chord units are shown below the pitch class units for convenience only. Furthermore, the chords are shown in a systematic order for illustrative convenience only. What's important is the pattern of connectivity between units, not the layout of the units themselves.

After learning, the network can be tested to see if it has any interesting properties. When presented with the tones C, E, and G, units representing the chords that contain any of these tones are activated. These chord units in turn activate all units representing tones that are contained in these chords and keys that contain these chords. The activated key units in turn activate all units representing chords that are contained in these keys, and so on. The activation process eventually dissipates, and the network reaches an equilibrium state.

At this point, the activations of the chord and key units represent the degree to which chords and keys are suggested or expected. In response to three tones comprising a major triad, the expectations generated by the network are exactly in accord with the circle of fifths: chords and keys closest around the circle are most highly expected. In response to two major chords in succession whose roots are a whole tone apart (e.g., F major followed by G major), the network most strongly activates the key (C major) of which these chords are subdominant and

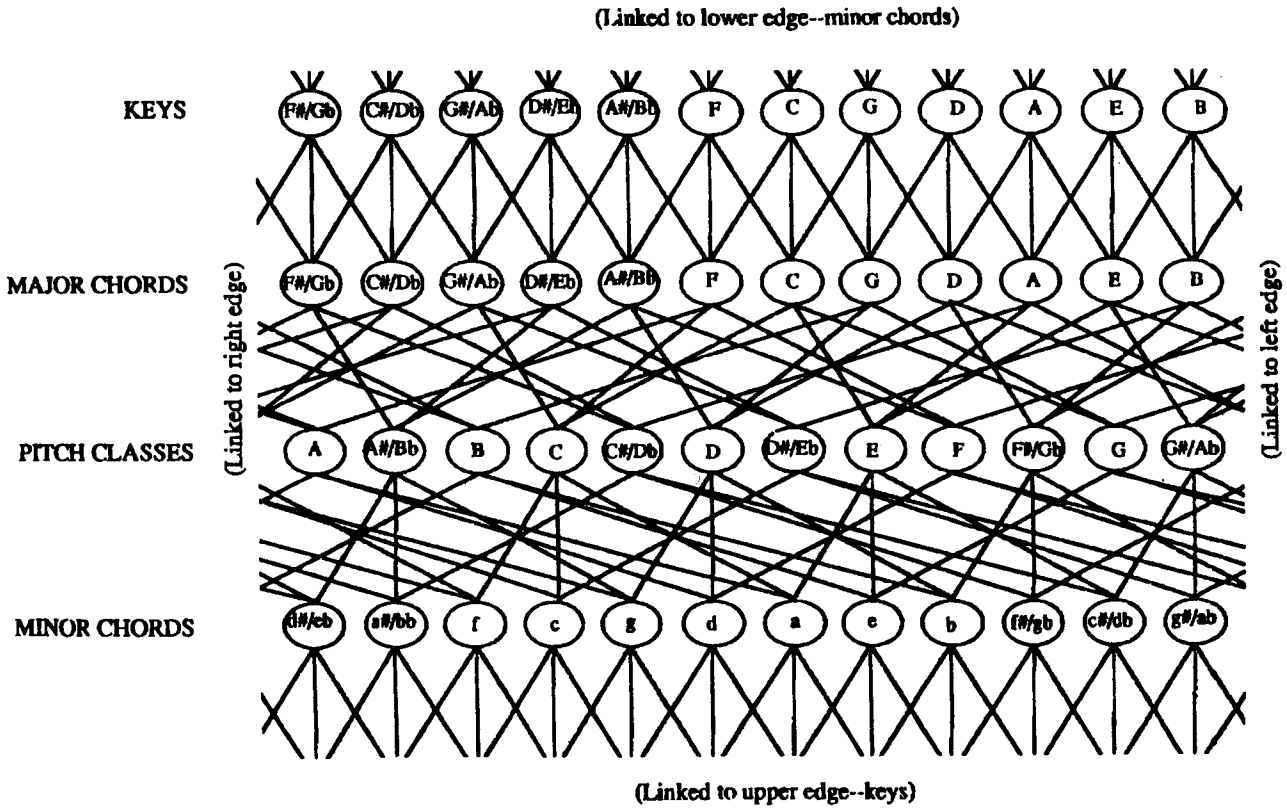



FIGURE 9.7. A network resulting from learning by exposure to typical combinations of tones as found in Western music. (Adapted from Bharucha, 1987a. Copyright by the Cognitive Science Society.)

dominant, and generates a strong expectation for the tonic chord. Finally, in response to a key-establishing context such as a major chord or a final cadence, the network generates expectations for the diatonic tones in that key. To this extent, the network can be seen to generate tonal hierarchies.

The network exhibits the patterns of expectation found in experiments on chord priming. The indirect activation of a chord unit represents an anticipatory processing of the represented object, resulting in faster and more accurate judgments about that chord, as observed in priming.

These complex behaviors of the network cannot be seen immediately by simply examining a diagram of the network such as Figure 9.7. They emerge from computer simulations of the activation process operating on the network. The network thus has emergent properties—that is, properties that are not directly evident from the initial assumptions but emerge from the way in which these assumptions interact. Emergent properties that are in accord with experimental data and music theory, as they are in this case, constitute strong support for a theory. Most theories that consist of rules (such as Lerdahl and Jackendoff's theory) postulate those rules explicitly and do not provide an account of how those rules are acquired. In a neural net, in contrast, the system behaves in accord with rules that weren't postulated explicitly but instead emerge from the complex interactions of more elementary processes that are themselves learned.

Neural nets thus offer the potential to understand how tonal schemata may be learned by passive perceptual learning, thereby giving rise to the schematic expectations that form such an integral part of our musical experience.

## NOTES

1. These include (but are by no means limited to) Zarlino (1558/1968), Fux (1725/1943), Schenker (1935/1979), Schoenberg (1954/1969), Meyer (1956), Lerdahl and Jackendoff (1983), Narmour (1990), and Butler and Brown (Chapter 8 in this volume).

2. As opposed to rhythmic or melodic expectations.

3. In keeping with the goals of this book, I shall attempt to avoid detail and excursions in favor of elucidating a clear line of reasoning. The reader is directed to other sources for details, for additional work, and for alternative explanations that bear on the question of harmonic expectation. These include Deutsch and Feroe (1981), Sloboda (1985), Dowling and Harwood (1986), Handel (1989), and, most relevant to some of the material covered in this chapter, Krumhansl (1990). For details about the work of the author, see Bharucha (1987, 1991; Bharucha & Todd, 1989).

4. It should be noted that there is some debate about whether "emotion" is the appropriate term to describe the often subtle and ineffable experiences induced by the fulfillment and violation of tonal expectations. Whether one is willing to call these experiences emotions, however, they are clearly affective.

5. There is no known reason to suppose that one cannot have more than one musical culture.

6. Many cognitive psychologists would consider a schema to be a memory system,

too. For the sake of simplicity, however, we shall restrict the term "memory" to memory for pieces of music.

It may be possible for the schematic expectations to emerge from the memory instead of requiring a completely distinct structure. One way in which this might occur has been explored by Bharucha and Todd (1989). Other methods might generate the schema by an average or composite (e.g., Metcalfe, 1982) of all individual memory traces. Whether the schema and the memory are part of the same physical structure or separate structures is still an open research question. In this chapter it will be assumed that they are separate.

7. This conclusion must await further research, however, since the empirical determination of this issue is very tricky. When a piece of music is recalled from memory in order to perform or imagine it, the schema is activated just as it would be by the sound itself, thereby generating the event hierarchy. It therefore seems redundant to store the elaboration in memory, since it is generated automatically every time the piece is recalled. This is what makes an empirical answer to the question so difficult.

8. This is based on the probability with which chords follow each other in the musical genre to which the subjects have been most extensively exposed (Western tonal music for the subjects in these experiments). C major is most often followed by G major or F major and least often by F-sharp major. Hence, for these subjects, G and F are objectively more highly expected than is F-sharp, following C.

9. The figure shows root position triads in close position as an example only. Although the original experiments employed "Shepard" tones (Shepard, 1964), which render tones more ambiguous with respect to their particular octave, subsequent experiments (Stoeckig, 1990) have shown that the results from this priming experiment are quite robust for different inversions, positions, and voice leading. Thus the effect seems to truly capture the perceptual properties of the triads as wholes.

10. In the case of the linear spatial representation of pitch height this is true! Such a *tonotopic* representation of pitch height is found not just in the inner ear, but even in the cerebral cortex.

11. The information contained in a circle can be encoded by a computer in many possible ways, including a string of symbols specifying the mathematical equation of the circle. Similarly, the brain might encode this information in terms of a particular arrangement of neurons, even though this arrangement may not itself be circular. Cognitive psychologists explore mental representations at different levels of information abstraction, spatial representations being some of the most abstract and neural representations some of the least abstract. The advantage of abstract representations is that the essence of the information can be captured without worrying about the representational medium. The advantage of less abstract representations is that the causal mechanisms involved in encoding the information can be addressed.

12. The strength of a correlation of two profiles is given by a single number, the correlation coefficient, $r$, that ranges between $-1$ and 1. A correlation coefficient of zero signifies no correlation.

13. There is nothing mystical about a four-dimensional space. Mathematically, to say that more than three dimensions are required to capture all the relationships between major and minor keys simply means that more than three factors are needed for a complete description of these relationships. The difficulty of depicting the four-dimensional key space stems only from an attempt to illustrate the space in terms of a model that can be visualized.

14. Parncutt's (1989) theory of the psychoacoustic foundations of harmony predicts this outcome.

15. This is not the same as being able to imagine music after becoming completely deaf, which only requires auditory imagery. The familiar account of Beethoven continuing to compose while deaf doesn't in and of itself discount the spectral overlap hypothesis. If sounds are stored in memory in terms of their spectra, then, while imagining music, the imagined spectra would overlap or not, just as would real spectra.

16. See Bharucha (1987b) for a report of a preliminary priming experiment using Indian music and Indian subjects.

17. See Deutsch (1969) for an early model of this kind. Unlike more recent models, it couldn't learn, since the mechanisms by which neural nets might learn were not then well understood.

18. This is called a tonotopic mapping, and can be found all the way from the inner ear to the auditory cortex.

## REFERENCES

Anderson, J. R. (1990). *Cognitive psychology and its implications.* New York: Freeman.

Bharucha, J. J. (1984). Event hierarchies, tonal hierarchies, and assimilation: A reply to Deutsch and Dowling. *Journal of Experimental Psychology: General, 113,* 421–25.

Bharucha, J. J. (1987a). MUSACT: A connectionist model of musical harmony. In *Proceedings of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Bharucha, J. J. (1987b). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception 5,* 1–30.

Bharucha, J. J. (1991). Pitch, harmony, and neural nets: A psychological perspective. In P. Todd & D. G. Loy (Eds.), *Music and connectionism.* Cambridge: MIT Press.

Bharucha, J. J., & Olney, K. L. (1989). Tonal cognition, artificial intelligence and neural nets. *Contemporary Music Review, 4,* 341–56.

Bharucha, J. J., & Stoeckig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 403–10.

Bharucha, J. J., & Stoeckig, K. (1987). Priming of chords: Spreading activation or overlapping frequency spectra? *Perception and Psychophysics, 41,* 519–24.

Bharucha, J. J., & Stoeckig, K. (1989). *Chord priming: The automaticity of schematic expectancies.* Paper presented at the meeting of the Psychonomic Society, Atlanta, GA.

Bharucha, J. J., & Todd, P. (1989). Modeling the perception of tonal structure with neural nets. *Computer Music Journal, 13*(4), 44–53.

Carlsen, J. C. (1981). Some factors which influence melodic expectancy. *Psychomusicology, 1,* 12–29.

Carlsen, J. C., Divenyi, P. L., & Taylor, J. A. (1970). A preliminary study of perceptual expectancy in melodic configurations. *Council of Research in Music Education, 22,* 4–12.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology; General, 113,* 394–412.

Deutsch, D. (1969). Music recognition. *Psychological Review, 76,* 300–7.

Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review, 88,* 503–22.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition*. San Diego: Academic Press.

Fux, J. J. (1943). *Steps to Parnassus* (A. Mann, Ed. and Trans.). New York: Norton. (Original work published 1725)

Geary, J. M. (1980). Consonance and dissonance of pairs of inharmonic sounds. *Journal of the Acoustical Society of America, 67*, 1785–89.

Handel, S. (1989). *Listening*. Cambridge: MIT Press.

Hebb, D. O. (1949). *The organization of behavior*, New York: Wiley.

Helmholtz, H. (1954). *On the sensation of tone*. New York: Dover. (Original work published 1885)

Jairazbhoy, N. (1971). *The rags of North Indian music: Their structure and evolution*. London: Faber & Faber.

Jones, M. R. (1981a). Music as a stimulus for psychological motion. Part I. Some determinates of expectancies. *Psychomusicology, 2*, 34–51.

Jones, M. R. (1981b). Music as a stimulus for psychological motion. Part II. An expectancy model. *Psychomusicology, 2*, 1–13.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology, 11*, 346–74.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. Oxford: Oxford University Press.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89*, 334–68.

Krumhansl, C. L., & Schmuckler, M. A. (1986). The *Petroushka* chord: A perceptual investigation. *Music Perception, 4*, 153–84.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 579–94.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge: MIT Press.

Mandler, G. (1984). *Mind and body*. New York: Norton.

Metcalfe Eich, J. (1982). A composite holographic associative recall model. *Psychological Review, 89*, 627–61.

Meyer, L. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Narmour, E. (1977). *Beyond Schenkerism: The need for alternatives in music analysis*. Chicago: University of Chicago Press.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures*. Chicago: University of Chicago Press.

Parncutt, R. (1989). *Harmony: A psychoacoustic approach*. Berlin: Springer-Verlag.

Pierce, J. R. (1969). Attaining consonance in arbitrary scales. *Journal of the Acoustical Society of America, 40*, 249.

Pierce, J. R. (1983). *The science of musical sound*. Scientific American Books.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vols. 1 & 2). Cambridge: MIT Press.

Schenker, H. (1979). *Free composition* (E. Oster, Ed. and Trans.). New York: Longman. (Original work published 1935)

Schmuckler, M. A. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception, 7*, 109–50.

Schoenberg, A. (1969). *Structural functions of harmony* (L. Stein, Ed.). New York: Norton. (Original work published 1954)

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America, 36,* 2346–53.

Sloboda, J. A. (1985). *The music mind.* Oxford: Oxford University Press.

Stoeckig, K. (1990). *The effects of octave transformations: Evidence from harmonic priming concerning the functional equivalence of chords and their inversions.* Unpublished doctoral dissertation, Department of Psychology, Dartmouth College, Hanover, NH.

Tekman, H., & Bharucha, J. J. (1991). *Dynamics of the development of harmonic priming: From tones to keys and back.* Paper presented at the meeting of the Eastern Psychological Association, New York.

Tramo, M. J., Bharucha, J. J., & Musiek, F. E. (1990). Music perception and cognition following bilateral lesions of auditory cortex. *Journal of Cognitive Neuroscience,* 2, 195–212.

Zarlino, G. (1968). *The art of counterpoint* (G. A. Marco & C. V. Palisca, Trans.). New Haven, CT: Yale University Press. (Original work published 1558)

# 10

# Perception, Production, and Imitation of Time Ratios by Skilled Musicians

SAUL STERNBERG AND RONALD L. KNOLL

## Introduction

The rhythmic element of music has always been a central topic for musicians (Cooper & Meyer, 1960; Yeston, 1976; Kramer, 1988; see Jaques-Dalcroze, 1972, for applications of rhythm to theories of music pedagogy). Psychologists, however, have shown increased interest in this area only in the last few years. Early research in the perception of rhythm was based more on the research in the perception of motor control and time (Lashley, 1951; Fraisse, 1963) than on rhythm within a musical context (reviewed in Fraisse, 1987).

A number of psychological studies have investigated the perception of polyrhythms, namely: sequences of rhythmic patterns occurring simultaneously (see Handel, 1984, for an overview). These experiments did not utilize musical stimuli, but rather sequences of tapping played at different pitches. They evaluated the listeners' ability to describe verbally what they heard, or to demonstrate by tapping one rhythmic pattern selected from several rhythms heard simultaneously.

Researchers have provided models of how listeners perceive rhythm, meter, and time. Steedman (1977) proposed a computer model of how listeners perceive musical rhythm and meter that utilized the occurrence of relatively long notes and the repetition of melodic phrases as cues to infer the meter of unaccompanied melodies. Steedman applied this model to analyze the metric structure of the 48 fugue subjects of *The Well-Tempered Clavier* by J. S. Bach.

Longuet-Higgins and Lee (1982) set forth a model that explained how listeners perceive the rhythms of note sequences based on the relative length

From J. Gibbon & Lorraine Allan (Eds.), *Timing and Time Perception*, Annals of the New York Academy of Sciences, vol. 423, copyright © 1984. Reprinted by permission.

of the notes. According to this model, after hearing the first two notes of a sequence, listeners already have in mind some hypothetical grouping, which they revise or confirm as they continue hearing the sequence of notes. The influential research of Mari Riess Jones has dealt with models of time perception and attention within music (Jones, 1987a, 1987b, 1992; Jones & Boltz, 1989).

The two chapters that follow illustrate how musicians perceive aspects of rhythm and timing. In "Perception, Production, and Imitation of Time Ratios by Skilled Musicians," by Sternberg and Knoll, the research is carried out in a laboratory setting, while in "The Interpretive Component in Musical Performance," by Shaffer and Todd, it is evaluated within musical performances.

"Perception, Production, and Imitation of Time Ratios by Skilled Musicians" summarizes a series of experiments that measured the abilities of musicians to reproduce, judge, and imitate (by tapping or by reproducing a note on the violin) a brief fraction of time corresponding to a nonstandard subdivision of the beat. The data showed that the subjects were inaccurate in reproducing, judging, and imitating the small fraction of time required by the experimental task.

The results demonstrate how tasks on the perception of timing that are psychologically very pertinent do not necessarily translate into tasks that are musically as pertinent. Subjects who are experts in responding to rhythm and timing within a musical framework cannot necessarily transfer their skills when making judgments or providing tapping responses to fractions of time in a context devoid of musical significance (vitality).

Regardless of their inability to be on target with the required experimental tasks, these expert musicians will nevertheless perform as skilled soloists and/or as a fine ensemble in a musical setting. In other words, they will all come in slightly early or slightly late, but they will all make their musical entrances *together,* as a group. In listening to music, listeners are not sensitive to how performers respond to a fraction of a second, but rather to their ability to *synchronize* their performances (Rasch, 1988).

## REFERENCES

Cooper, G., & Meyer, L. (1960). *The rhythmic structure of music.* Chicago: University of Chicago Press.

Fraisse, P. (1963). *Psychology of time.* New York: Harper.

Fraisse, P. (1987). A historical approach to rhythm as perception. In A. Gabrielsson (Ed.) *Action and perception in rhythm and music,* volume *55.* Stockholm: The Royal Swedish Academy of Music.

Handel, S. (1984). Using polyrhythms to study rhythm. *Music Perception, 1,* 465–84.

Jaques-Dalcroze, E. (1972). *Rhythm, music and education* (H. Rubenstein, Trans.). New York: B. Blom. (Original work published 1921)

Jones, M. R. (1987a). Perspectives on musical time. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (Vol. 55). Stockholm: Royal Swedish Academy of Music.

Jones, M. R. (1987b). Dynamic pattern structure in music: Recent theory and research. *Perception and Psychophysics, 41*(6), 621–34.

Jones, M. R. (1992). Attending to musical events. In M. R. Jones & S. Holleran (Eds.), *Cognitive bases of musical communication.* Washington, DC: American Psychological Association.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96* (3), 459–91.

Kramer, J. D. (1988). *The time of music.* New York: Schirmer Books.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. P. Jeffress (Ed.), *Cerebral mechanisms in behavior: The Hixon Symposium.* New York: Wiley.

Longuet-Higgins, H. C., & Lee, C. S. (1982). The perception of musical rhythms. *Perception, 11,* 115–28.

Rasch, R. A. (1988). Timing and synchronization in ensemble performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition.* Oxford: Oxford University Press.

Steedman, M. J. (1977). The perception of musical rhythm and metre. *Perception, 6,* 555–69.

Yeston, M. (1976). *The stratification of musical rhythm.* New Haven, CT: Yale University Press.

---

Among the human skills in which timing and time perception are critical, musical skill is distinctive: for most players a notation specifies temporal pattern explicitly and provides a criterion to which performance can be compared. Because of their years of practice in the use of this notation, the behavior of professional musicians can plausibly reveal some of the ultimate capacities and constraints of human timing mechanisms. We have examined the performance of skilled musicians in three laboratory tasks designed to capture temporal aspects of music. We focused on the short time intervals—fractions of a second—that are among the shortest durations specified by musical notation. As in Western music, these intervals occurred in the context of a train of periodic beats and were defined as fractions of the beat interval. Our three tasks—perception, production, and imitation—all appear to be required of musicians during ensemble rehearsal and performance, for example. It is plausible that because players try to "keep together," ensemble experience would cause performance in the three tasks to become at least consistent and probably correct as well (that is,

consistent with the notation). Neither of these expectations was borne out by our experiments; instead, we observed surprisingly large systematic errors and inconsistencies.

The principal subjects were three professional musicians: a flutist, a cellist, and Paul Zukofsky (PZ), violinist and conductor. We also obtained a small amount of corroborative data from Pierre Boulez, composer and conductor. A detailed report of our results is available, based on group data.[1] In the present chapter we describe only our more interesting findings, illustrated with data from PZ, who is the most musically experienced of our principal subjects, whose performance we examined in a wider variety of procedures than the other subjects, and whose data are more consistent than theirs, both within and across experiments. The picture generated by the group data is somewhat less clear, but leads to the same conclusions.

## Three Tasks of Temporal Psychophysics

We used two kinds of stimuli, shown on the left of Figure 10.1. A *time pattern stimulus* contained two or more *beat clicks* separated by a *beat interval*. The beat interval was usually 1 sec. One or more of the beat clicks was followed by a *marker click* after a *fractional interval*, *f*, that defined a fraction of the beat. A *fraction-name stimulus*, *n*, was presented both as a numerical fraction and in musical notation, where a quarter note was defined as one beat. On the right of the figure are shown the two kinds of response. In making a *fraction-name response*, *N*, the subject would select a category such as "less than 1/8 beat" or "between 1/8 and 1/7 beat." In making a *timed response*, the subject tapped his finger after a beat click, thereby producing a fractional interval, *F* between beat click and tap.

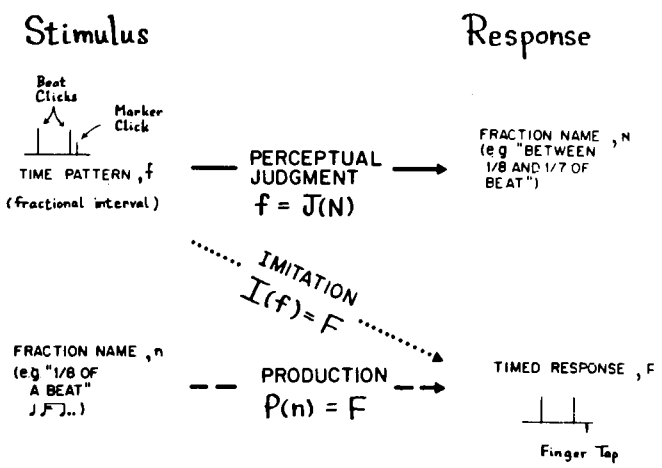



FIGURE 10.1. Stimuli, tasks, and responses.

Stimuli and responses were linked by three different tasks, as shown in the center of the figure. In *perceptual judgment* the subject assigned fraction names to time patterns. He thereby generated a *judgment function*, $f = J(N)$, that maps fraction names onto their subjectively equivalent fractional intervals. In *production* the subject made a timed response to produce a fractional interval associated with a specified fraction name. He thereby generated a production function, $P(n) = F$. In *imitation* (sometimes called the "method of reproduction") a time-pattern stimulus elicited a subjectively equivalent timed response. We thereby obtained an imitation function, $I(f) = F$. To permit performance to stabilize, conditions in all experiments remained constant for at least 25 trials.

## Perceptual Judgment of Beat Fractions

The stimulus pattern in our first perception experiment is shown in Figure 10.2. Two pairs of beat clicks separated by two beat intervals (a *rest*) were followed by a single marker click. For each of a set of fraction names we used an adaptive psychophysical procedure to determine the fractional interval that is subjectively equivalent to it. The resulting judgment function is labeled J, in Figure 10.3. The fractional interval, $f$. is plotted as a function of fraction name, $N$. Both scales are logarithmic.[1] If stimulus and response agreed, data would fall on the straight line with unit slope. Instead, our subject radically overestimated fractions less than 1/4 beat (or 250 msec). Consider the case of 1/8 beat, for example. Accurate performance would assign this fraction name to a fractional interval of 125 msec; instead, it is assigned to an interval of about 67 msec, or 1/16 beat. Thus, for small fractions, the assigned fraction name is larger than the fractional

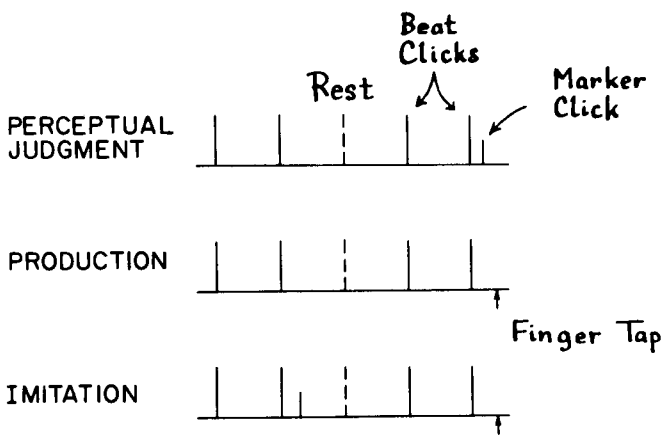



FIGURE 10.2. Sequences of beat clicks, marker clicks, and finger taps in experiments $J_1$, $P_3$, and $I_5$.

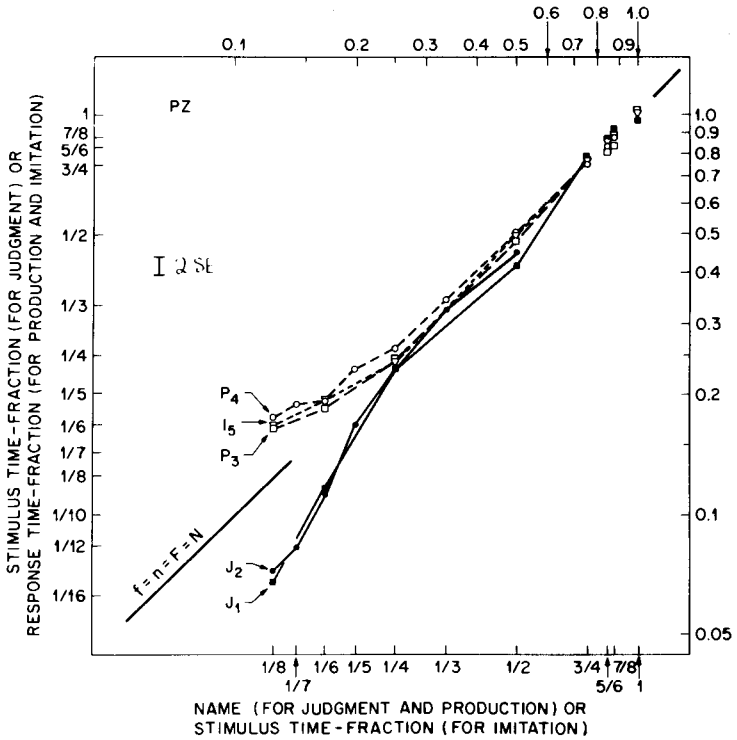



FIGURE 10.3. Results for subject PZ from five experiments: judgment (J), production (P), and imitation (I). Ordinate (*abscissa*) values are given as fractions on a logarithmic scale on the left (bottom) and as decimal values on the right (top). Error bar indicates approximately 2 S.E. based on between-session variability. If the ordinate value is regarded as a function of the abscissa value, the functions shown are estimates of $f = J^{-1}(N)$, $F = P(n)$, and $F = I(f)$.

interval.[2] The *N*-value grows more slowly than the *f*-value, however, coming into approximate agreement at about 1/4 beat. Despite the poor *accuracy* for small fractions, judgment *precision* is high: the difference threshold at 1/8 beat was only 4.3 msec (about 6%), for example.

## Production of Beat Fractions

The large errors in the *perception* of small fractions make it particularly interesting to examine musicians' accuracy in *producing* such fractions. One appealing possibility is that production is mediated by a simple feedback process in which the subject judges the fraction he produces with respect to the fraction name he is trying to produce, and adjusts later productions accordingly. Suppose also that

produced fractions are judged by means of the same perceptual mechanism used in the judgment task. The production function should then approximate the judgment function: $P(n) = J(N)$ when $n - N$. Thus, overestimation of small fractions (assigning fraction names that are too large) would lead to underproduction (producing fractional intervals that are too small).

The events on each trial in our first production experiment are shown in Figure 10.2. They are the same as in the perception experiment, except for replacement of the marker-click stimulus by a finger-tap response.[3] With each tap the subject heard the thump of his finger striking a hard surface. The production function we obtained is labeled $P_3$ in Figure 10.3. The expectation from a feedback mechanism is violated dramatically. For small fractions the intervals produced are too large rather than too small. That is, we have overproduction rather than underproduction. Whereas the subject *judged* 67 msec to be 1/8 of a 1-sec beat, for example, he *produced* a mean fractional interval of 164 msec for the same fraction name. Furthermore, he seemed satisfied with his responses, and did not experience his taps as being late.[4] For small fractions in the perception task we have seen that the fraction stimulus, *f*, that is subjectively equivalent to a fraction name, *N*, changes more rapidly than the *N*-value. In contrast, the fraction, *F*, produced changes more slowly than the *n*-value.

Several analyses and experimental variants were directed at understanding the large systematic errors we had found in judgment and production, as well as their inconsistency. In the present report we consider nine of these. Others, together with these, are considered in greater detail in the full report (Sternberg, Knoll, & Zukovsky, 1982).

## Further Analysis of Perceptual Judgment

### *Psychophysical Procedure*

One concern is whether aspects of the judgment data may depend on special features of our psychophysical procedure. In any block of trials in the adaptive procedure, subjects judged time fractions with respect to just one fraction name, and the fractional intervals were concentrated in the range where judgments were the most difficult. In a second experiment we used a method more akin to traditional psychophysical scaling in which both these features were altered. The subject categorized a wide range of time-pattern fractions into one of eight categories of fraction names, ranging from "less than 1/8 beat" to "greater than 1/2 beat," in each trial block. Also in contrast to the first experiment, two beat clicks rather than only one were followed by marker clicks, providing two observations of the fractional interval on each trial. Results are labeled $J_2$ in Figure 10.3. The two judgment functions are almost identical, despite the differences between procedures.[5]

### *Fractional Interval Defined by Subjective Onset Versus Offset of Marker*

A second attempt to discover a source of the judgment errors was based on the possibility that the internal representation of a brief click may have a longer duration than the click itself. Suppose that when beat and marker clicks are close together, the subjective duration of the fractional interval is defined by the *onset* of the internal representation of the beat and the *offset* of the internal representation of the marker. Such a mechanism could then produce overestimation of the kind we observed. To test this possibility we compared judgments of our normal stimuli, in which clicks were 5-msec tone bursts, with stimuli in which the marker duration was prolonged by about 60 msec. Contrary to the idea that offset time is important, this variation produced no change in the judgments, for either small or large fractions.[6]

## Further Analysis of Production

We next turn to five of our efforts to understand the systematic errors in the production task and the inconsistency between production and judgment performances for small fractions.

### *Opportunity for Adjustment to Feedback*

As mentioned above, our subject did not report experiencing his finger tap responses as being late for either small fractions or large. Nonetheless there may have been discrepant feedback from the perception mechanism, but too little opportunity to adjust to it, given only one production per trial. On each trial in a second production experiment the subject produced timed responses after each of 10 successive beat clicks. We found no tendency for the error to be reduced over the 10 successive responses. Mean produced fractions based on all the responses are labeled $P_4$ in Figure 10.3. Performance agrees closely with the first production experiment. If anything, the tendency toward overproduction is slightly greater.

### *Subjective Delay of Tap versus Click*

A second potential source of inaccuracies in production might be a difference between two critical subjective delays. One is the interval between the occurrence of a click and its perceptual registration. The second is the interval (possi-

bly negative) between our measurement of a tap and when the subject perceives it to have occurred. If these two delays differed, then direct comparison of the intervals between beat click and marker click in perception and between beat click and tap in production would be inappropriate. To estimate the difference between the two subjective delays we asked subjects to tap in synchrony with one or more beat clicks. The difference can be estimated by the mean asynchrony between tap and click; the asynchrony was small—about 10 or 20 msec, depending on the procedure—clearly too small to explain the observed effects. On the assumption that the difference between subjective delays does not depend on the beat fraction, the production data in Figure 10.3 have been corrected by these small amounts.

### *Improved Feedback from Finger Tap*

The disparity between performances in perception and production led us to question a feedback model of production, which suggested in turn that we scrutinize the feedback itself. The feedback from tapping the finger included tactile, proprioceptive, and auditory cues, but not the marker click used in the perception task. In additional production experiments with both single- and multiple-tap procedures, but limited to the fraction 1/8 beat, each finger tap generated a marker click. The sequence of clicks in judgment and production thereby became identical. The production performance was virtually unchanged, however; we found only a 10-msec change in the mean interval produced.

### *Musical Instrument Response*

Another potential source of the production error might have been our choice of finger-tapping as a response. (The subject was a skilled violinist but not a skilled finger-tapper, at least at the start of these experiments.) We ran the single-response production experiment again, but now the response was to play a single note on the violin after the final beat click. We measured the onset time of the note as the subject attempted to produce the fractions 1/8, 1/2, and 1 beat. The amount of overproduction of the small fraction did not decrease. (In fact, it was nonsignificantly greater by 12 ± 11 msec for PZ.)

### *Minimum Reaction Time*

The potential sources discussed above of the production error and the production-perception disparity had to be considered. However, we had little *a priori* reason to expect that even if they had been important, their effects would have depended on fraction size. One constraint that might have such differential effects is the existence of a minimum reaction time (RT). The minimum RT to auditory stimuli is between 100 and 150 msec. Furthermore, there are delays

between excitation of a musical instrument and its acoustic response. The combination of these two effects makes it virtually impossible to produce a note 125 msec after a signal to respond (such as a beat click) when the signal is the event that initiates the response timing process. If we assume that the timing of an offbeat response starts with the immediately preceding beat, it follows that the notation often calls for production of discriminably different response delays, some of which are less than the minimum RT. One solution would be to bias productions just as we have observed, so that for different small fractions the mean intervals produced are greater than the minimum RT, but still distinct.

A test of this explanation is provided by variances of the finger-tap delays, *Var*(*F*), together with an argument suggested by Snodgrass, Luce, and Galanter (1967). We assume that as its mean increases, the variance of a response delay also increases, where the delay is measured from the event that initiates the timing process. If the responses for all fractions were timed from the final beat click, we would therefore expect *Var*(*F*) to increase with fraction size. Instead, we found it to vary as a *U*-shaped function of fraction size, with a minimum between $n = 1/4$ and $n = 1/2$.[7] This pattern of variability would be expected if small fractions, but not large ones, were timed from the penultimate beat; if so, overproduction cannot be explained as compensation for a limited speed of response.

## Production of Multiple Subdivisions of the Beat

How can the existence of large production errors for small fractions be reconciled with our belief that musicians are able to fill a beat interval accurately with a sequence of equally spaced actions? Could the production error depend on our use of a single, isolated response?[8] To address this question we studied the three conditions shown in Figure 10.4 in a new production experiment. The beat interval here was 1/2 sec instead of 1 sec, incidentally testing the generality of our effects. One of the conditions required an isolated offbeat response, with a target fraction of 1/4 beat, equivalent to a fractional interval of 125 msec. There were two multiple-response conditions. In one the subject started with a tap on the beat, and alternated between index fingers to fill the beat interval with quarter-beat taps. An unusual interval between the first two taps here would reveal any general distortion of subjective time near the beat. In another condition the initial on-beat tap was withheld. If overproduction depends on response isolation, then the presence of later taps within the same beat interval should eliminate it, especially since the final tap was supposed to be made on the next beat.[9]

Consider first the results from the 5-tap condition. All tap delays, and in particular the delay of the second tap, fall close to the fitted line. There appears to be no general distortion of subjective time near the beat. Furthermore, the slope of the linear function is less than unity, giving a value of 435 rather than 500 msec for position 5. (Since the subject felt satisfied with his productions, we

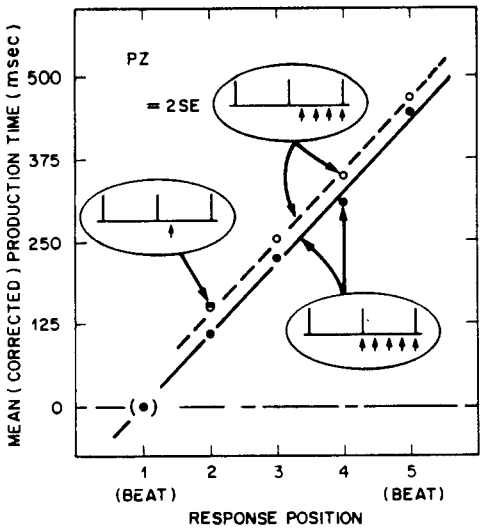



FIGURE 10.4. Stimulus and response patterns in three conditions in an experiment on multiple subdivisions of the beat, together with results. Mean response delays have all been corrected by the mean delay (−34.7 msec) between the second beat click and the first tap in the 5-response condition. Data from the two multiple-response conditions have been fitted by parallel lines. The displayed ± S.E. bar is appropriate for assessing adequacy of the fitted lines.

must assume that the subjective beat interval was shorter than the actual beat interval by about 13%.) In other words, with multiple subdivisions starting on the beat there is no evidence for overproduction of the small fraction. Consider next the 1-tap condition. Just as in our other experiments the response is delayed relative to the correct response time of 125 msec. Because of the shortening of the subjective beat interval, the amount of overproduction is even greater when referred to the second tap in the 5-tap condition. Consider finally the 4-tap condition. The first response—the tap in position 2—is delayed here as much as in the 1-tap condition. The overproduction effect therefore does not depend on the offbeat response being isolated. Instead, it appears to result from an onbeat response being withheld.

The displaced parallel lines provide a good description of performance in the two multiple-response conditions, indicating that every tap in the 4-tap condition is delayed relative to the 5-tap condition, and by about the same amount. This phenomenon seems best described as *displacement of the subjective beat*. As in some circadian phenomena, the phase of a periodic process has been changed with no alteration in its period. We find it especially remarkable that the delay of the first tap is propagated all the way through the last tap, despite the presence of a final beat click.

## Conjectures About Failure of the Feedback Model of Production

The beat-displacement effect suggests one possible source of failure of the feedback model. It is plausible that subjects judge their response delays relative to the

*subjective* train of beats. If so, the beat displacement associated with the delayed offbeat response would reduce the apparent delay of this response. Such an effect might explain a subject's inability to recognize and correct his overproduction. It would be too small an effect, however, to explain the major part of the discrepancy between overproduction and the underestimation we observed in the judgment task, which would require that the magnitude of beat displacement *exceed* the delay of the offbeat response, rather than merely equaling it.

The expectations from the feedback model of production that were violated depend on the assumption that the time-perception processes that accompany production are the same as those used in the judgment task. The beat-displacement effect leads us to question this assumption. The existence of the effect reminds us that in production the fractional interval terminated by the finger tap must not only be timed, but must also be placed in proper phase relation to the train of beats. Performance of small fractions in the production task therefore requires the timing of both a beat interval and a beat fraction within the same sequence. In contrast, there is no reason to believe that *judgment* of a beat fraction depends on concurrent judgment of a full beat interval.[10] It is possible that this difference between tasks contributes to the failure of the feedback model.

Results from a final judgment experiment provide weak evidence that favors this possibility, indicating that if successive long and short intervals must both be judged, the perception of at least one of them may be dramatically altered. We used click patterns like those shown in Figure 10.5, with a 1-sec beat interval, and asked whether judgment of the large interval between the final beat click and the pair of marker clicks would be influenced by the requirement also to judge the small interval between markers. On each trial the subject had to judge whether the interval between markers was large or small relative to 1/8 beat, and then also to judge whether the interval between the final beat click and the markers was large or small relative to a full beat. Judgments of the small beat

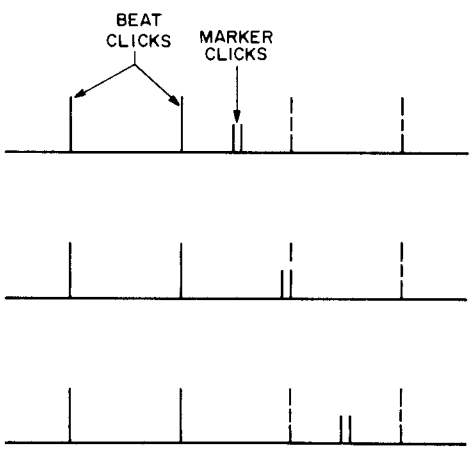



FIGURE 10.5. Examples of sequences of beat clicks and marker clicks in a dual-judgment experiment. Broken lines represent beats for which no beat click was presented.

fraction were very similar to performance in a single-judgment control condition. Judgments of the larger interval were enormously more variable than in its control condition, however: The difference threshold was increased by a factor of 10, from about 4% of the beat interval to about 40%.[11]

## Imitation of Beat Fractions

Imitation of beat fractions is of interest partly because it provides a further opportunity to determine the sources of error in perception and production. In judgment and production tasks, subjects must associate beat fractions with fraction names. In imitation (see Figure 10.1), the stimulus of the judgment task is mapped onto the response of the production task; fraction names are not explicitly involved. If the errors in judgment and production are due to the input or output of fraction names, it follows that imitation performance should be accurate.[12]

Such accurate imitation of time intervals is one possibility that has been considered in previous research (Carlson & Feinberg, 1968). A second possibility that has been advanced is that imitation is accomplished by concatenating two processes, a judgment process that covertly assigns a name to the time pattern, and a production process that converts this name into a timed response (Thomas & Brown, 1974). Given our findings of overestimation as well as overproduction of small fractions, this *full-concatenation model* implies that the response fraction in imitation will be too large by the sum of the errors in the other two tasks; we call such an outcome *strong overimitation.*

The imitation data are labeled $I_5$ in Figure 10.3. Results conform to *neither* possibility. Instead, responses to small fractional intervals in imitation were virtually identical to responses to the names of these fractions in production. The significance of this outcome is best explained in the context of the following model of performance in our three tasks.

## An Information-Flow Model of the Perception, Production, and Imitation of Beat Fractions

In the skeleton model diagrammed in Figure 10.6 we limit ourselves to accounting for the relations among performances in our three tasks. Each task involves processes that perform input, output, and possibly translation functions. Parsimony leads us to postulate the minimum number of processes consistent with our data, and hence the maximum number of processes shared between tasks.

The processes underlying perceptual judgment are shown by the two upper boxes. A time-pattern stimulus generates an internal fraction representation. This representation is converted into a fraction name to generate the required response. The processes underlying production are shown by the two lower

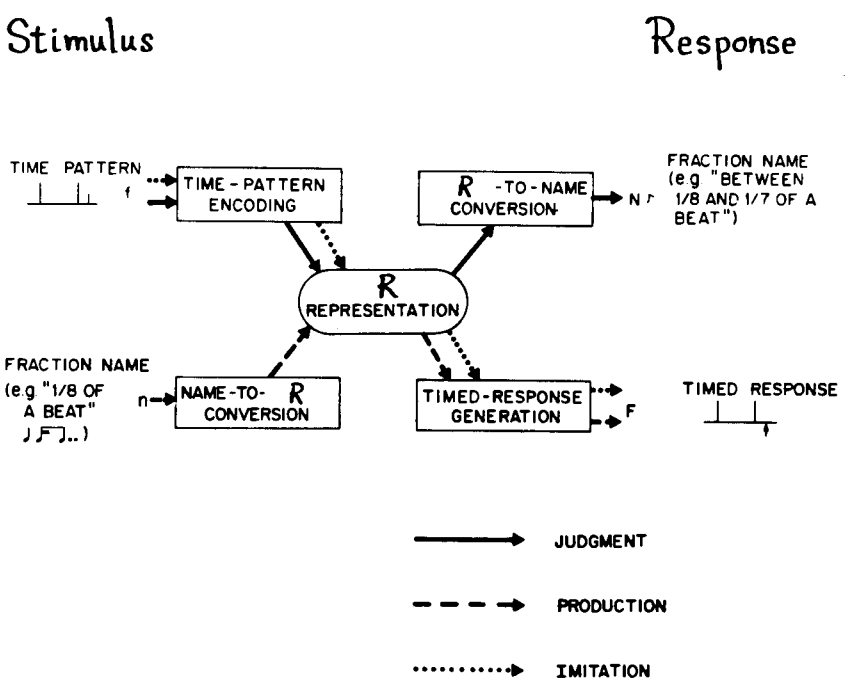



FIGURE 10.6 An information-flow model of perception, production, and imitation of beat fractions. The model incorporates four processes that convert time-pattern (*f*) or fraction-name (*n*) stimuli into time-pattern (*F*) or fraction-name (*N*) responses, and that make use of a common internal fraction-representation. Paths of information flow for the three tasks are represented by *arrows*.

boxes. A fraction name is converted into an internal representation of the same kind as in the judgment task. This representation is then used to produce the required timed response. A full-concatenation model of imitation would most naturally be represented by a mechanism in which the upper and lower pairs of processes made use of distinct internal representations. In such a mechanism information could not flow directly from time-pattern encoding to timed-response generation. Instead, to connect those two operations the covert output of the pair of processes used in judgment (a fraction name) would become the input for the pair of processes used in production. Because such a model can be rejected, we adopt a *partial-concatenation model* in which imitation makes use of a common internal representation, and shares only the encoding process with the judgment task and only the response-generation process with the production task.

Each of the four processes in the model can be thought of as a function or transformation that maps its input onto its output. Qualitative aspects of our data restrict the relations among these transformations. The most important of these relations is based on the close agreement of the data from production and

imitation. When time-pattern and fraction-name correspond, they lead to the same timed response, so that P ≅ I. This in turn means that they must have produced the same internal representation, thus the two input transformations are the same. The internal representation therefore creates a veridical mapping between the two kinds of stimuli.[13] By a complementary argument from the finding that responses to the same fractional interval in imitation and judgment are *not* the same (J ≠ I) we conclude that the two output transformations are distinct.

Both the production and judgment tasks generate psychophysical scales that describe the subject's mapping of musical notation onto time fractions. As we have already seen, neither of these scales is veridical, and, moreover, they disagree with each other. Both scales are *explicit*, in that the subject's response is identified directly with one of the terms in the psychophysical function. By combining results from imitation, $F = I(f)$, and production, $F = P(n)$, which involves the same response, we generate an *implicit* psychophysical scale that relates musical notation, $n$, to time fractions, $f$, when both are presented as stimuli: $n = P^{-1}I(f)$. Our data indicate that this implicit scale is veridical, despite the systematic errors in each of the three explicit relations, J, P, and I.[14]

## Summary

We have described our exploration of the judgment, production, and imitation of fractions of a beat by skilled musicians, illustrating our findings with data from violinist and conductor Paul Zukofsky. For small fractions we found systematic and substantial errors. In the judgment task small stimulus fractions are associ-

Summary of principal expectations and findings

| *For Small Fractions (<¼ Beat)* | |
|---|---|
| Perceptual Judgment | |
| *Results: Overestimation* | $f = J(N) < N$ |
| Production | |
| Expectation: | |
| Underproduction | $F = P(n) < n$ |
| (Feedback Model) | |
| *Results: Overproduction* | $F = P(n) > n$ |
| Imitation | |
| Alternative Expectations: | |
| a) Veridical | $F = I(f) = f$ |
| b) Strong Overimitation | $F = I(f) >> f$ |
| (Full-Concatenation Model) | |
| *Results: Overimitation* | $F = I(f) > f,$ |
| *and, for* $f = n$. | $I(f) = P(n)$ |

ated with names that are too large (overestimation). In both production and imitation tasks the fractions produced were too large (overproduction, overimitation). A summary of our findings and of the expectations they violate is provided in the table on page 254.

The temporal patterns we used are perhaps the simplest that qualify as rhythms, incorporating just a beat interval and a fraction. The phenomena we discovered in relation to these simple patterns, and their implications for underlying mechanisms, must be considered in attempts to understand the perception and production of more complex rhythms, as in actual music.

We explored and rejected several plausible explanations for the overestimation and overproduction of small fractions. Although we have as yet no satisfactory explanations of the errors themselves, relations among the errors have powerful implications for human timing mechanisms. The relation between the errors in judgment and production requires us to reject a feedback model of production, in which a subject uses the same processes as in the judgment task to evaluate and adjust his performance in the production task. An explanation of the inconsistency between judgment and production seems most likely to lie in a change in time perception induced by the production task. Together with the existence of systematic errors in judgment, the equality of the errors in production and imitation argues that imitation is not accomplished by concatenating all the processes used in judgment and production. Our results are instead consistent with a model containing four internal transformation processes, in which judgment and production share no process, but do involve the same internal-fraction representation, and in which imitation shares one process with judgment and another with production.

ACKNOWLEDGMENTS

We thank Paul Zukofsky, coauthor of the full report (Sternberg, Knoll, & Zukovsky, 1982) for significant contributions both as subject and collaborator. We are also grateful to Marilyn I. Shaw for helpful comments on the manuscript.

NOTES

1. Power functions appear as straight lines on such a plot.

2. Figure seven provides an outline of our principal findings.

3. Note that although infrequent in earlier music, the playing of a note after the beat without playing a note on the beat is not unusual in the music of the past 60 years.

4. This is an informal impression, not based on analysis of systematically collected data.

5. The procedural differences did influence judgment precision; however, difference thresholds were about twice as large in the second (multiple-fraction) experiment.

6. The measured mean change in $f = J(N)$, averaged over six fraction names, was 0.6 ± 1.9 msec. There was neither a mean effect of marker duration nor an interaction of marker duration with fraction size. This finding may reflect a general property of the perception of timing and rhythm in music, the dominance of the sequence of time intervals between successive *attacks*, and the relative unimportance of *release* times.

7. Because imitation performance is virtually identical to production, in variability as well as mean, a good estimate of the effect of fraction size on the variability of tap delay is provided by $Var(F)$ averaged over the two production experiments described above and the imitation experiment to be described [later]. Fractions that were examined in all three experiments include 1/8, 1/6, 1/4, 1/2, 3/4, 5/6, and 7/8; for PZ the corresponding RMS values of the S.D. are 21.3, 20.3, 16.7, 25.2, 32.4, 42.2, and 42.4 msec, respectively.

8. Musicians could perhaps learn to fill an interval evenly, without accurate perception or production of isolated beat fractions, by employing judgments of evenness and synchrony, together with the ability to count actions.

9. The first tap in the 5-tap condition provided the synchronization correction that we used to adjust all taps in the three conditions. Application of the correction in this way requires us to assume that the location of the initial subjective beat as well as the subjective beat rate depend only on the beat clicks, and are influenced neither by whether there is a tap on the beat not by the number of taps that follow.

10. Judgment of a fractional interval could use a stored representation of the beat interval. Alternatively, the beat interval might not be directly represented at all, but would determine the calibration of a mechanism that assessed beat fractions.

11. In further work along these lines it will be important to force high accuracy in judgments of the large interval and to search for effects on both mean and variability of judgments of small fractions.

12. The converse does not follow: if imitation were accurate, we would know only that the function that relates stimulus fractions to their internal representations must be the inverse of the function that relates internal representations to produced fractions.

13. By a mapping of psychophysical scale that is veridical we mean one that associates the beat fraction $1/n$ with the fractional interval $b/n$, where $b$ is the beat interval.

14. Given our model it is tempting to inquire whether "error" or "distortion" in any single constituent in the model can account for the performance errors in all three tasks, and their relations. Such an inquiry succeeds qualitatively: The only such single constituent that could be responsible is the internal representation, $R$, since this is the only constituent common to the three tasks. Suppose that for small fractions, $R$ is "expanded" so as to correspond to a larger fraction. Examination of Figure 10.6 reveals that such an expansion, alone, would produce the three effects we observed: overestimation, overproduction, and overimitation. Quantitatively, however, this explanation fails, because it requires that $J^{-1} = 1$, a relation we can reject reliably if we use the full range of data. The explanation might succeed, however, if we limited it to small fractions.

## REFERENCES

Carlson, V. R., & Feinberg, I. (1968). Individual variations in time judgments and the concept of an internal clock. *Journal of Experimental Psychology*, 77, 631–40.

Sternberg, S., Knoll, R. I., & Zukofsky, P. (1982). Timing by skilled musicians. In D. Deutsch (Ed.), *The psychology of music* (pp. 181–239). New York: Academic Press.

Snodgrass, J., Luce, R. D. & Galanter, F. (1967). Some experiments on simple and choice reaction time. *Journal of Experimental Psychology, 75*, 1–17.

Thomas, F., & Brown, I, Jr. (1974). Time perception and the filled-duration-illusion. *Perceptual Psychophysiology, 16*, 449–58.

# 11

# The Interpretive Component in Musical Performance

L. H. SHAFFER AND NEIL P. McANGUS TODD

## Introduction

"The Interpretive Component in Musical Performance," by L. H. Shaffer and N. P. Todd, represents a line of studies that address the perception of rhythm and timing utilizing actual performances of music (see Gabrielsson, 1987a, for a review). With the aid of computerized equipment, this research analyzes the nuances in the rhythmic patterns that pianists display in repeated performances of a composition, and traces a profile of their rhythmic and expressive interpretations. Musical performance research is very valuable to psychologists and musicians alike (see Davidson's Chapter 4 and Sloboda's Chapter 6 in this volume for a discussion of musical skill and performance. See also Clarke, 1982, 1989; Gabrielsson, 1987b; Palmer, 1989; Shaffer, 1981, 1992; Shaffer, Clarke, & Todd, 1985; Sloboda, 1983; Repp, 1990).

Todd (1985) has suggested that pianists reflect the hierarchical structure of a piece by slowing down at points of stability in the piece. In the research described in this chapter, Shaffer and Todd used repeated performances of works by Chopin, J. S. Bach, and Satie to demonstrate that pianists are consistent in reproducing the musical expressions they wish to convey. The performers repeatedly utilize timing to convey and emphasize the structure of the compositions. Psychologically, the performers' overall consistencies indicate that their motor system has a definite representation of the timing parameters. Musically, this research illustrates that nuances such as *rubato* and deviations in tempo reflect a clear musical plan on the part of the performing artists.

From A. Gabrielsson (Ed.), *Action and Perception in Rhythm and Music*. Publications issued by the Royal Swedish Academy of Music No. 55, 1987. Reprinted by permission of the Royal Swedish Academy of Music.

REFERENCES

Clarke, E. F. (1982). Timing in the performance of Erik Satie's "Vexations." *Acta Psychologica, 50,* 1–19.

Clarke, E. F. (1989). The perception of expressive timing in music. *Psychological Review, 5,* 2–9.

Gabrielsson, A. (1987a). *Action and perception in rhythm and music* (Vol. 55). Stockholm: Royal Swedish Academy of Music.

Gabrielsson, A. (1987b). Once again: The theme from Mozart's piano sonata in A major (K.331). In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (Vol. 55). Stockholm: Royal Swedish Academy of Music.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance, 15*(12), 331–46.

Repp, B. (1990). Patterns of expressive timing in performances of a Beethoven minuet by nineteen famous pianists. *Journal of the Acoustical Society of America, 88,* 622–41.

Shaffer, L. H. (1981). Performances of Chopin, Bach and Bartok: Studies in motor programming. *Cognitive Psychology, 13,* 326–76.

Shaffer, L. H. (1992). How to interpret music. In M. R. Jones & S. Holleran (Eds.), *Cognitive bases of musical communication.* Washington, DC: American Psychological Association.

Shaffer, L. H., Clarke, E. F., & Todd, N. P. (1985). Metre and rhythm in piano playing. *Cognition, 20*(1), 61–77.

Sloboda, J. A. (1983). The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology, 35*(A), 377–90.

Todd, N. P. (1985). A model of expressive timing in tonal music. *Music Perception, 3,* 33–58.

---

Musical performance can be creative in a variety of ways. A jazz or classical musician can improvise on a tune, theme, or chord sequence, to embellish it, provide a set of variations, or construct a fugue. With a little less freedom, a player can extemporize on the harmonic skeleton of a figured bass. These feats may appear to be in contrast with that of a musician playing from a score, but the freedom to improvise is different in kind rather than in degree. Scores vary in explicitness and often do little more than adumbrate a musical form, stating a metre and giving pitch and time values for a sequence of notes. Parameters of tempo, intensity, and articulation of notes—the manner of their sounding and the extent to which they are connected or detached—are usually left to the discretion of the player, and it is implicitly understood that even the written pitch

and time values of notes may be deformed in order to convey or enhance aspects of musical meaning. Thus the emphasis on creativity shifts from musical invention to the subtle shaping of sound.

There are, to be sure, differences of opinion among both composers and players, on the interpretive license given to the player. At one extreme a score is viewed as a text to be transcribed as accurately as possible. On this view there is an authentic interpretation and only this fully realizes the composer's intention. At the other extreme a score is seen as providing a series of basic forms the musical meaning of which can be re-created each time by the player. The sparse notation of a Bach or Scarlatti score suggests that these composers encouraged more license in interpreting their music than, say, Beethoven.

Interpreting a piece of music, then, requires a good working knowledge of the musical intentions of the composer in relation to the conventions and styles of his period, but it is an axiom of performance that it should clearly convey the meaning of the music. A pleasant feature of the world of music is that players are happy to take this axiom as a working brief when the concept of musical meaning is poorly defined. There is no paradox in this, since a player can have a tacit concept that is more definite than the public version. The corollary, however, is that different players can have different concepts. This raises the empirical questions of what players contribute in interpreting a piece of music, whether there is significant agreement in the interpretations of a piece by different players, and what consistent principles may players bring to playing different pieces. Some of these questions are taken up here. They are of interest, though perhaps in different ways, to students of cognitive science and music theory.

Though the concept of musical meaning is inexplicit, we can at least agree with Meyer (1956) that it is a manifold concept having major components that relate to formal structure, emotion, and physical movement. We shall not consider music such as song, theatre music, and tone poem, that is literary in origin or descriptive of the external world. In these it is the programmatic origin rather than the music itself that is likely to govern the use of expression. Putting them aside, what remains is a vast repertoire of absolute music, mainly concerned with developing musical ideas. Such music can convey nuances of mood through choices of tempo, key, modality, pitch register, and pitch contour. It can also suggest kinds of physical movement or gesture by abstractly imitating the rhythms of dancing, walking, speaking, and so on (Gabrielsson, Bengtsson, & Gabrielsson, 1983). Indeed the titles or Italian markings in a score often indicate these nonmusical references, allowing a three-beat measure to be differentiated into the swirl of a Viennese waltz or the more stately stepping of a minuet. It is further claimed that some kinds of ethnic music are based on the motor gestures that are formally required in playing them on a particular instrument (Bailey, 1985). Similar claims may be made for some of the Etudes of Chopin and Debussy, written as virtuoso pieces to capture complex patterns of movement on the keyboard.

Thus even absolute music makes reference to external states and events in conveying mood or physical movement. In these referential functions it acts as an alternative to natural language. Perhaps the major sense of musical meaning, however, lies in the structure of the music itself, and in this sense music is a closed system making reference only to its own sonic domain (Lerdahl & Jackendoff, 1983). The structures have their own meaning, and this autoreferential function of music is in contrast to natural language, in which syntax provides only a key to a more abstract semantics (Lyons, 1981). The structures of music are contained in its metre, rhythm, melody, and harmony and they can act both independently and together in patterning the music.

The different aspects of meaning must all enter into the planning of a performance. There is scope for giving prominence to one or other of these but ignoring or misjudging any of them carries the risk of leaving the musical listener baffled or unsatisfied. Altering the mood of a piece or giving a dance a song-like character (these are not independent factors) may give the music an arbitrary flavor. Failure to convey its structure may leave it without shape, a directionless concatenation of sounds. Simple monodic pieces can look after themselves, as long as they are accurately transcribed, but the more complex the texture of the music the more relevant this last point becomes.

This chapter looks at some of the ways in which concert pianists make expressive use of timing to convey musical structure. The exercise requires two things, the ability to capture sensitively the pianist's use of timing and the ability to describe musical structure in a way that is informative and relatively unequivocal. Difficulties arise on both points. Objective timing is easily measured, but subjective timing for both player and listener may depend on a subtle blend of real time, intensity, tone quality, and the articulation of successive sounds. Also, although there has been progress in the description of musical structure (Lerdahl & Jackendoff, 1983; Schenker, 1969) the methods of description still require a degree of musical intuition in their application and there are not yet accepted canons for describing even relatively conventional tonal music. Nor can it be claimed that the current alternatives are variants that can be generated by a common abstract formalism. Thus lowering our sights a little we shall concentrate on a particular expressive device, the use of slowing to indicate the boundary of a musical group. This device is ubiquitous in music (Seashore, 1938; Shaffer, 1980) and speech (Klatt, 1976), but it gains its effect by being used with discretion. Some of the factors that govern its use are examined here.

## The Data

The performances were given on a Bechstein grand piano modified to obtain computer recordings without affecting the sound or touch of the instrument (Shaffer, 1981). Photocells placed in vertical pairs opposite each of the hammers

in the action provide recordings of both the timing and intensity of every key-press. The timing gives information about note duration and the legato/staccato articulation of its succession.

### *Chopin: Prelude in F-Sharp Minor*

The pianist, PB, had recently practiced this piece for a recital. She gave two performances of the Prelude in a single recording session. The first few measures of the piece are shown in Figure 11.1. It is a four-beat measure and there is a fixed rhythmic pattern at each beat that recurs throughout the piece. This allows us to consider the performances in a hierarchy of musical units and we have chosen the beat as the basic unit of timing. The graph in Figure 11.1 shows played duration on the ordinate for each beat on the axis, for all but the last measure. Note that if there were no expressive timing the graph of a performance would be a horizontal line.

The departure from metrical timing is immediately apparent in the graph and its magnitude is quite large. The median beat duration is about 800 ms, and the duration ranges between 600 and 1600 ms, excluding the final rallentando. What is impressive, and important for a cognitive theory, is the very high agreement of timing between the two performances. The shapes of the two curves are almost identical. They differ occasionally in the degree of speeding or slowing but are almost unanimous in where these occur. Such precision has been repeatedly found in studies of concert musicians since the work of Seashore (1938). What it demonstrates is that an expressive form can have a precise mental representation and can be precisely executed. This in turn implies the presence in the brain of an excellent timekeeper and a very good way of coding expression. The musical interest lies more perhaps in the shape of the curves. This reveals a patterning of timing that is responsive to a hierarchy of musical units. There is a typical gesture of acceleration and slowing that can be seen in the beats in a measure, the measures in a four- or two-measure phrase, the phrases in a section and indeed the sections in the piece. Thus the gesture has been used recursively in each level of the hierarchy. Furthermore, there is the stronger result that at each level, the amount of slowing at a boundary tends to reflect the depth of that boundary in the hierarchy. This last result becomes clearer when assigning the parameters to a mathematical model in order to simulate the expressive form, the result of which is shown in Figure 11.2.

The model takes a structural description of the music as its input and generates an expressive form from this. Its success in simulating the data rests on the validity of two factors, the choice of a parabola to represent a basic timing gesture, and the choice of a structural description setting the parameters and hence the shape of each parabola. The assignment of structure makes use of the idea of time-span reduction discussed by Lerdahl and Jackendoff (1983), but with a piece as harmonically complex as the Chopin Prelude some intuition is in-

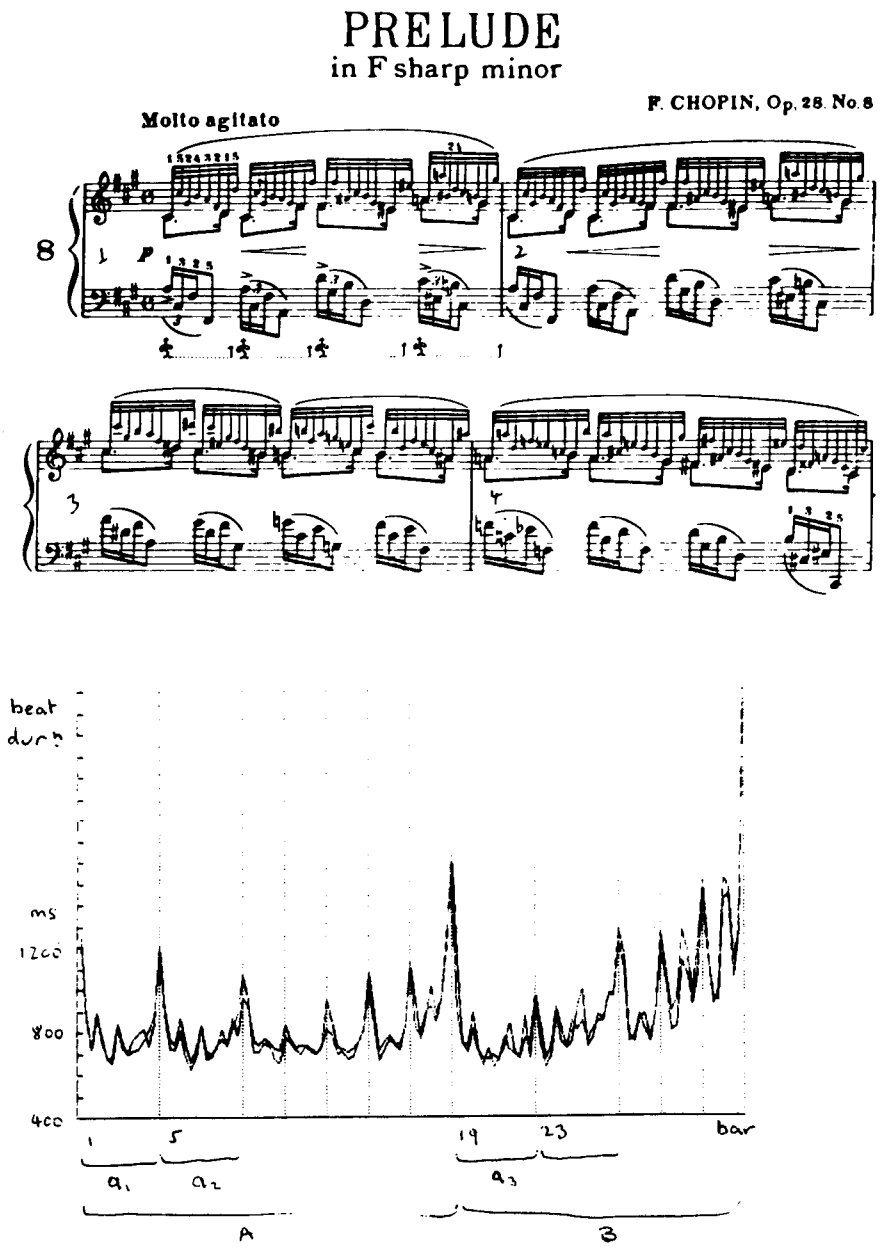



FIGURE 11.1 The first four measures of the Chopin Prelude in F sharp minor and a graph of played beat duration in two performances of the whole piece by PB. The bar, section, and phrase boundaries of the piece are indicated. (The score reproduced by permission of The Associated Board of the Royal Schools of Music, Publishing Limited.)

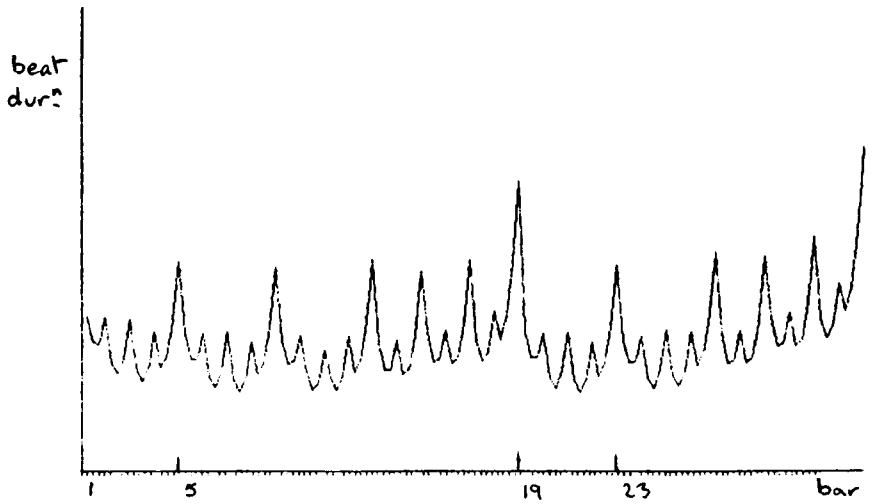



FIGURE 11.2. Beat duration for the Chopin Prelude, generated by a mathematical model.

volved in using evidence of cadence and parallel structure to assign the hierarchy.

### *Bach: Prelude in C*

The first Prelude in the first book of Bach's *Well-Tempered Clavier* is familiar to any pianist and it was given two performances here by the pianist RH with little prior practice. Like the Chopin Prelude the piece has a fixed rhythmic pattern recurring (almost) throughout. Again there is a hierarchy that contains repeating eight-note patterns within four-beat measures, and four-measure phrases, except for an elision at the beginning that makes the last measure of the first phrase also the first measure of the second. Thus there is some reason to expect that expressive patterns should be similar in performances of the two pieces. The timing of the performances of the Bach, taking the measure as the timing unit, is shown in Figure 11.3. There is less agreement of tempo and rubato between the two performances than in the Chopin. This may reflect a willingness to experiment with tempo and rubato in a piece that the pianist plays seldom, and it will later be shown that in more practiced pieces his timing is much more consistent. There is an additional factor obscured by the different timescales of the graphs in Figures 11.1 and 11.3: all the time differences shown in Figure 11.3 are small and of a lower order than in Figure 11.1. Bar duration in the second, faster, performance had a median of about 2500 ms and ranged between 2400 and 2600 ms, excluding the final rallentando. Hence in the Bach there was relatively little departure from metrical timing, apart from the final rallentando, the last measure of which is omitted from the graph.

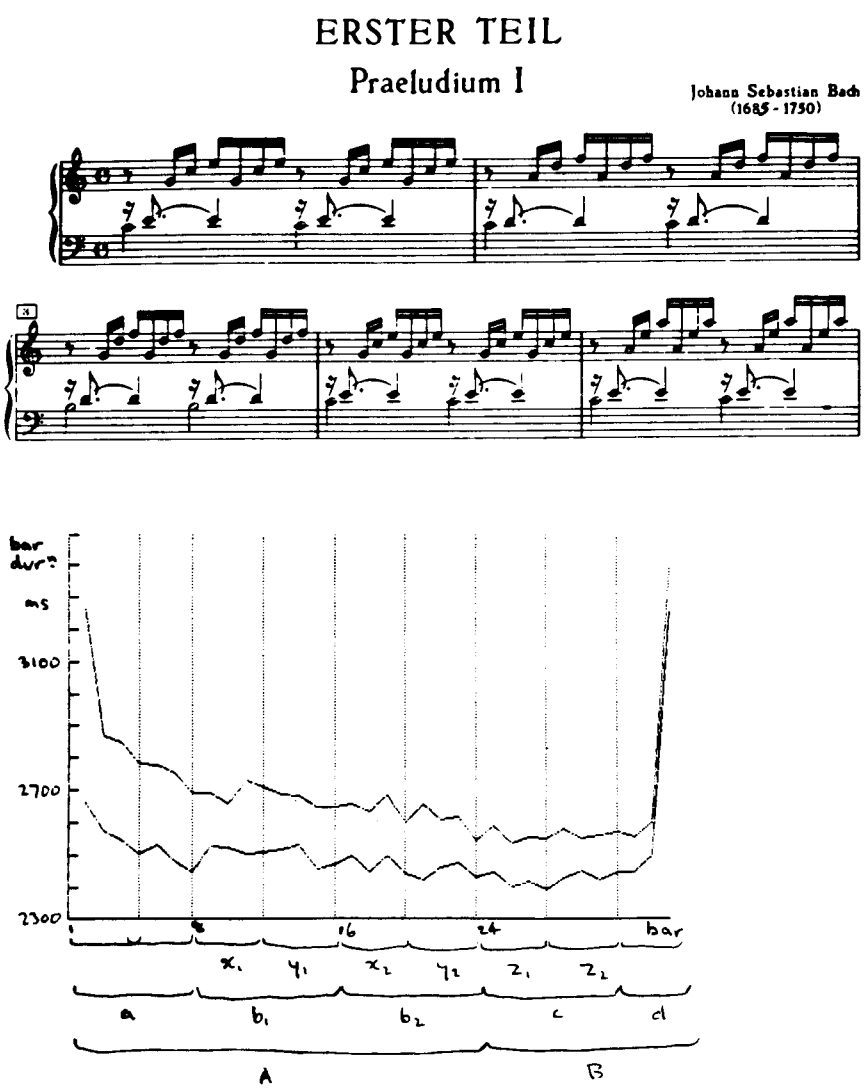



FIGURE 11.3. The first five measures of the Bach Prelude in C and a graph of played measure duration in two performances of the whole piece by RH. The boundaries of bars and phrases, and a hierarchic grouping for these are indicated.

The musically relevant result is that contrary to what was expected the timing gesture of acceleration and slowing was not used to convey hierarchy in the musical structure. Instead there was a single parabolic gesture embracing the whole piece. Yet there are cadences at the phrase boundaries marking intermediate points of rest in the music, which are not always present in the Chopin

Prelude. Unlike that piece, however, the Bach Prelude never properly leaves its home key and it has only a rudimentary melodic line. It can be interpreted, then, as a didactic exercise in exploring the varieties of sonority and progression made possible by inverting and chromatically altering chords, sometimes thereby creating secondary dominants, of the key of C. This didactic function can be underlined by playing the piece as a single tonic to dominant to tonic trajectory (cf. Schenker, 1969).

### *Satie: Gnossienne No. 5*

The pianist JM has recorded the piano music of Satie, including this piece, for the gramophone. He gave three performances here of the Gnossienne among other pieces recorded during the session (Shaffer, Clarke, & Todd, 1985).

In contrast to the other pieces described so far, the Gnossienne has a simple periodic rhythm in the left hand but the right hand has a melodic rhythm that changes continually, embracing a wide variety of rhythmic figures and note durations. Also, whereas the Bach has a clear tonal center throughout and the Chopin much of the time, the Satie appears to vacillate between G and its relative, E minor, without establishing either and so remaining tonally ambiguous, except in a brief bridging passage that is clearly in C. Nor is there a simple relationship between melodic grouping and metre. Hierarchic structure can be discerned only in terms of the recurrence of melodic groups (see Shaffer et al., 1985).

The timing in the three performances, taking the beat as the unit, is shown in Figure 11.4. Again we see the precisely reproducible timing that a concert musician brings to a practiced piece. As in the performances of Bach, there is little use of rubato, the beat duration ranging between 2000 and 2400 ms, nearly half of which is accounted for by a tempo drift described below. However, there is almost an inversion of expressive gesture: the underlining of grouping is now reserved for the lowest level of unit, motivic groups of notes, marked in the score by slurs; there is also a rallentando at the end and, tentatively, before the recapitulation. The other notable feature of timing is that the performances drift throughout to a progressively slower tempo, as if to emphasize the essential lack of harmonic direction of the music. Thus the player appears to have interpreted the piece as a concatenation of musical fragments, an interpretation consistent with a melodic rhythm that alternates groups of short notes with long notes, acting as pauses.

### *Chopin: Mazurka in A Minor*

In order to establish a more direct comparability between players and pieces, we turn to performances of the same piece by the pianists PB and RH, who played the Chopin and Bach preludes, respectively. PB gave one performance of the Mazurka and RH gave two.

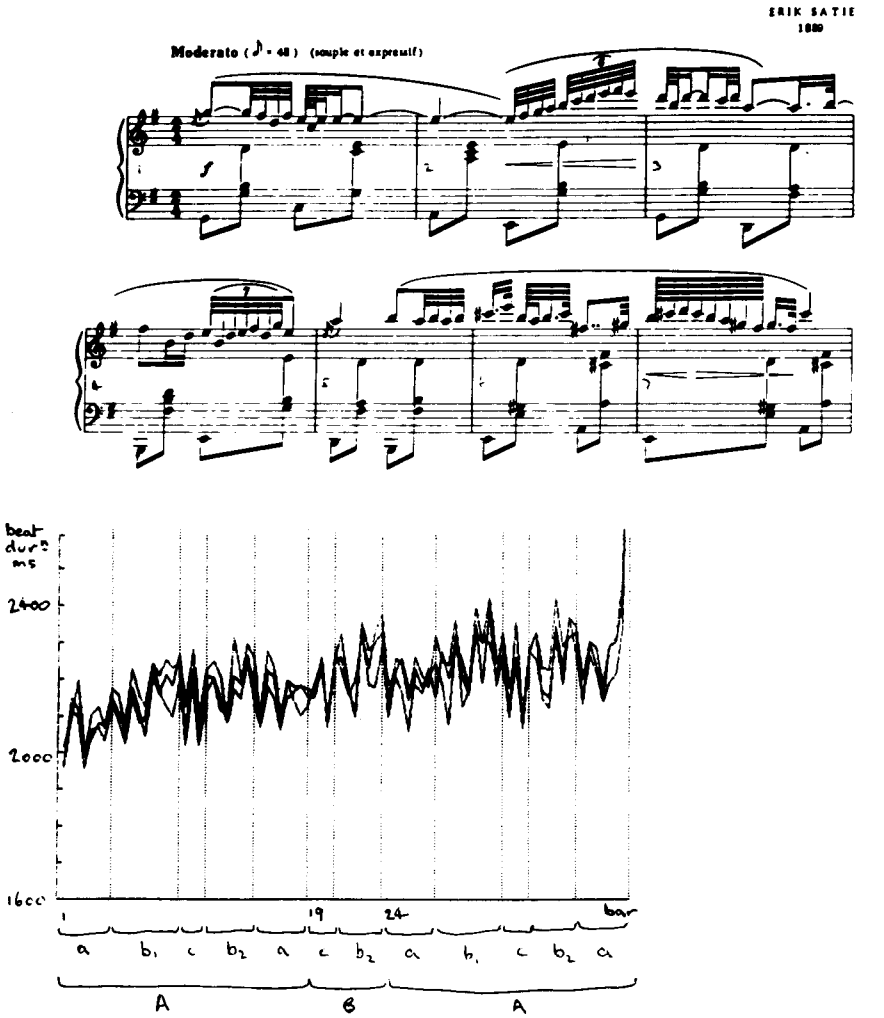



FIGURE 11.4. The opening measures of Satie's Gnossienne No. 5, together with a graph of played beat duration in three performances of the whole piece by JM. Boundaries of the melodic groupings are indicated. (The score reproduced by permission of Editions Salabert, Paris/UMP.)

The graphs in Figure 11.5 show timing based on the unit of a measure. The opening and closing phrases are omitted, because the use of rests in the music makes it difficult to estimate measure duration. Several points can be made from the graph. The first is that, in contrast to his Bach performances, RH demonstrates a very high order of timing consistency in the repeat performances of the more often practiced Mazurka, and he also shows a greater freedom in using

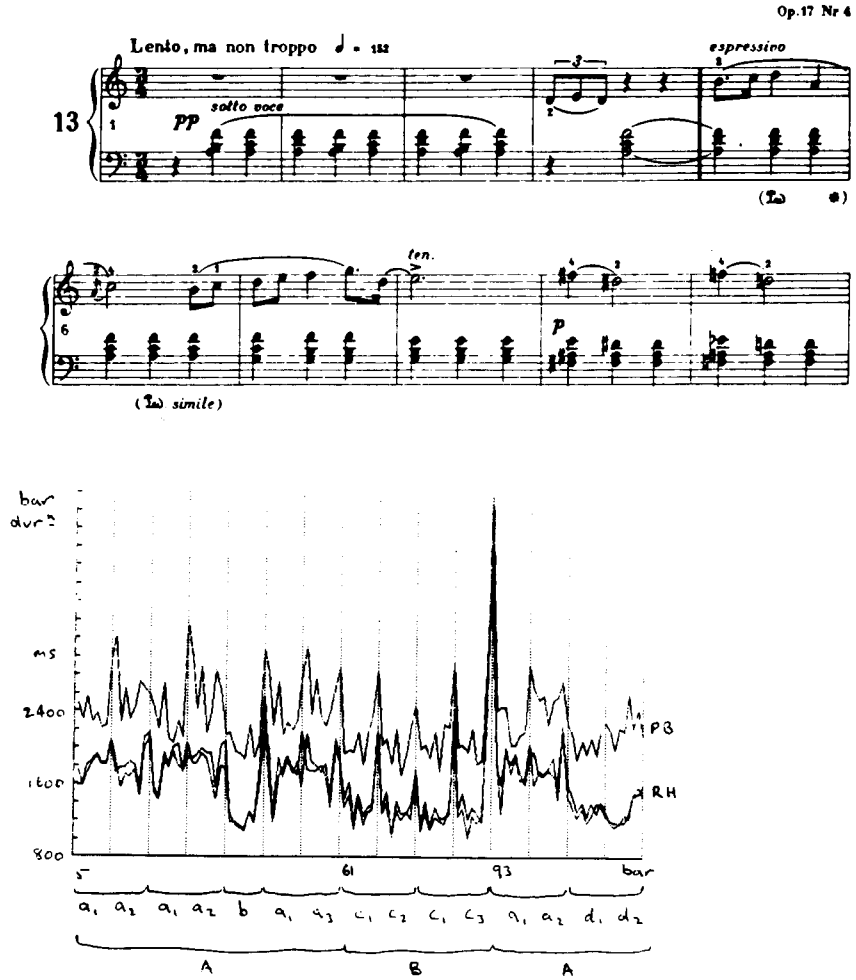



FIGURE 11.5. The opening measures of the Chopin Mazurka in A minor, together with a graph of played measure duration in two performances of the whole piece by RH and one by PB (shown as the uppermost curve). The boundaries of bars, phrases and their hierarchic groupings are indicated (excluding the opening and closing phrases). (The score reproduced by permission of PWM Editions (Alfred A. Kalmus Ltd.))

rubato, though still more restrained than PB. The two players choose different tempi, reflected in measure durations around 2300 ms for PB and 1600 ms for RH. The range of variation in using rubato is about 1200 ms for both players, excluding the rallentandi at the recapitulation and end. Yet RH reserves rubato more to mark phrase boundaries. The effect of these and other timing differences is that RH plays the piece more consistently as a dance. Despite such differences

in interpretation they agree in marking the boundaries of the eight-measure phrases and, at the sectional level, the imminence of the recapitulation. If anything it is RH who more punctiliously observes the hierarchy in the musical structure. The players also agree in playing the themes (marked b, c, and d in Figure 11.5) that alternate with the major theme (marked a) at a faster tempo. Note that in the b theme the music modulates from the minor to the major mode, and in the c and d themes there is a higher degree of restlessness in the music produced by harmonic dissonance.

## Discussion

It would be preferable to have data of each pianist playing all pieces to enable direct comparisons of players and pieces, but the elegance of experimental design is not easily arranged with concert pianists having different repertoires. Luckily PB and RH both have a special interest in the music of Chopin.

A result reiterated in this study and confirming the findings in earlier studies is that a concert musician has a very precise means of reproducing tempo and its variation in playing a practiced piece of music. At the cognitive level, this implies that the motor system has access to a stable and accurate timekeeper together with a fairly definite representation of the relevant timing parameters. At the musical level it demonstrates that tempo and rubato are not matters of current whim but are the outcome of coded decisions of interpretation. The timing profile in playing a piece typically has a complex pattern, and so we can suppose that it stands in some definite relation to properties implicit in the musical notation. The alternative would be that the player has to memorize a large number of arbitrary parameters.

We have concentrated on the way in which timing is used expressively to convey, or emphasize, musical structure. It has long been known that slowing is used to convey the boundaries of musical units, and the present results extend the picture of this, showing it to depend on a number of factors. It is too early yet to say whether there may be systematic principles governing its use, but the results suggest where to look for these.

Most music can be shown by analysis to contain hierarchic structures, and the results demonstrate that the player can convey one or more of these hierarchies by a recursive and hence graded use of slowing at a boundary. Rather than make this an invariant principle of performance, which would turn this form of expression into a somewhat pedantic exercise, the player may choose which aspects of hierarchy to reveal so as best to characterize the music. This character may in turn relate to a structural or referential component of musical meaning. Fortuitously, the Bach Prelude and the Satie Gnossienne provide almost a polar contrast in musical purpose, the one developing a linear thesis that invites a single trajectory of musical tension in its performance, the other dissipating its trajectory in a series of tonally ambiguous gestures, encouraging an

interpretation in terms of overall stasis. It is also worth noting that in these performances, in contrast to the Chopin, the selective use of boundary slowing is associated with an overall limited compass of rubato. One may argue that a decision to restrict the compass limits the scope for discriminating hierarchy, or conversely that a decision to ignore hierarchy allows a smaller use of rubato, but it is also possible that the association is accidental and neither variable constrains the other. We shall require many more performances of different pieces before we can settle on the rules that best describe, or generate, these results. What is clear from the performances of the Mazurka by different players is that the use of structural information is not arbitrary.

ACKNOWLEDGMENTS

This research is supported by a grant, No. C00232213, from the Economic and Social Research Council.

REFERENCES

Bailey, J. (1985). Musical structure and human movement. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition*. London: Academic Press.

Gabrielsson, A., Bengtsson, I., & Gabrielsson, B. (1983). Performance of musical rhythm in 3/4 and 6/8 meter. *Scandinavian Journal of Psychology, 24*, 193–213.

Klatt, D. H. (1976). Linguistic uses of segmental durations in English. *Journal of the Acoustical Society of America, 59*, 1208–21.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Lyons, J. (1981). *Language, meaning and context*. Bungay, Suffolk: Fontana.

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Schenker, H. (1969). *Five graphic music analyses*. New York: Dover Publications.

Seashore, C. E. (1938). *Psychology of music*. New York: McGraw-Hill.

Shaffer, L. H. (1980). Analysing piano performance. In G. Stelmach & J. Requin (Eds.), *Tutorials in motor behaviour*. Amsterdam: North-Holland.

Shaffer, L. H. (1981). Performances of Chopin, Bach and Bartok: Studies in motor programming. *Cognitive Psychology, 13*, 327–76.

Shaffer, L. H., Clarke, E. F., & Todd, N. P. (1985). Metre and rhythm in piano playing. *Cognition, 20*, 61–77.

Todd, N. P. (1985). A model of expressive timing in tonal music. *Music Perception, 3*, 33–58.

# The Perception of Musical Compositions

# 12

# Can Listening to Music Be Experimentally Studied?

RITA AIELLO

## Introduction

Research in the perception of music has been influenced by research in psychoacoustics and psycholinguistics (see Chapter 2 in this volume; for reviews see Deutsch, 1982; Dowling & Harwood, 1986; and Sloboda, 1985). This led to many experiments that focused on the perception of isolated musical variables and used very brief, computerized musical stimuli. In most cases, to control the variables, the musical stimuli were extremely impoverished, and therefore they did not represent the richness of music as we all experience it.

This chapter discusses some of the valid methodological reasons that originally favored the reductionistic approach to the research in the perception of music. Furthermore, it addresses some of the difficulties that must be considered if we investigate listening to actual performances from the repertoire, and reviews some recent experiments that used recorded performances of musical works as stimuli. Finally, it argues that research in the perception of music cannot remain research in the perception of isolated musical variables couched within very brief stimuli because music is more than the sum of its parts and is experienced over a temporal continuum. The perception of actual musical compositions or excerpts ought to become a priority for the research in the psychology of music.

### REFERENCES

Deutsch, D. (Ed.) (1982). *The psychology of music.* New York: Academic Press.

Dowling, J. W., & Harwood, D. (1986). *Music cognition.* New York: Academic Press.

Sloboda, J. A. (1985). *The musical mind.* Oxford: Oxford University Press.

Interested readers may ask: Does the research in the psychology of music explain how the listener perceives an actual musical composition? Which experiments describe this process, and how? Such questions are valid indeed. Musicians, musicologists, and music educators have always emphasized the value of listening and describing music (Corredor, 1956; Meyer, 1956, 1967; Copland, 1957; Sessions, 1970; Brendel, 1976; see also the first six chapters in this volume). But despite the major advances in the psychology of music in the last two decades, music psychologists have been reluctant to investigate how listeners perceive actual pieces of the repertoire. Therefore, little is known about the processes that take place while listening to a composition.

By and large, psychologists have addressed music listening theoretically. Several authors have proposed models of how a listener may listen to tonal music in general, or to certain tonal compositions specifically (Deutsch & Feroe, 1981; Lerdahl & Jackendoff, 1983; Marsden, 1987; Jackendoff, 1991). These models have been based on the hierarchical or perceptual processes that the authors theorize to be inherent in music compositions or fragments of music, and not on empirical data gathered from the listeners' experiences.

There are several reasons why researchers have not conducted in-depth investigations of listeners' perceptions of musical compositions. Let us examine some of them.

## The Influence of Psychoacoustics

First, speaking from a historical perspective, research in the psychology of music developed mainly from the research in psychoacoustics. Therefore, early studies in music psychology addressed mainly the psychoacoustical attributes of isolated pitches. Although psychoacoustical studies have proven to be most valuable for our understanding of isolated sounds, they aimed to describe exclusively the parameters of single pitches or of a few pitches heard together.

The psychoacoustical influence favored the construction of brief musical stimuli that have been used in the experiments, and allowed sound variables to be clearly manipulated. But this has happened too often at the expense of the aesthetic qualities of the musical stimuli themselves. Referring to the tradition to use brief and artificial sequences as stimuli, Clarke and Krumhansl (1990) pointed out:

> There are certain obvious advantages in this very controlled kind of approach, and it has proved extremely powerful and productive for advancing our understanding of tonal and metric hierarchies. However, it has left untouched a range of issues concerned with listeners' understanding of more extended and elaborate structures in which a considerable degree of interaction between different parameters can be expected. (p. 214)

In the past, to suit experimental goals, stimuli were often designed at the expense of musical aesthetics. Deliège and El Ahmadi (1990) write:

> The usual practice, in our field, as in any scientific discipline, is to isolate the variable that one wishes to study and to incorporate it in a series of brief and repetitive sound sequences (that are called musical), constructed by the psychologist for the needs of the experiment, in order to be able to identify it afterward, in appropriate manner, in the statistical analysis of the data. . . . Unfortunately, many studies in the field of psychology of music scarcely achieve their aims because a musical objective is being sought through the use of material that is both too simple and too trivial. (p. 22)

## The Influence of Psycholinguistics

The psychology of music has also been influenced by developments in psycholinguistics. As in psycholinguistic research, the emphasis in the psychology of music has been on investigating the phonetic and the grammatical levels of brief strings of words without an in-depth look at the perception of entire discourses, so research in the cognitive psychology of music has examined the perception of brief melodies or sequences of musical stimuli, and has not thoroughly investigated the perceptual processes that take place when listening to longer musical excerpts. The psychology of language emphasizes research on the perception of isolated sentences. Similarly, the psychology of music, for the most part, has looked at the perception of single musical phrases not occurring within a full musical context. Generally, the length of the experimental stimuli has ranged from a few notes to one or two phrases. This limitation does not replicate the actual musical experience.

Perhaps it was premature to expect that we could have built a house without foundations. Therefore, the reader who sincerely asks about the perception of genuine musical compositions can consider the last two decades of the research in the psychology of music as a foundation on which to build better investigations. Notwithstanding music psychology's good intentions, however, questions regarding how listeners perceive actual musical compositions still remain to be answered and it becomes much more pressing to ask them now. We must face that to do full justice to the laboratory, the psychology of music has not yet done justice to the concert hall or the cassette player.

Several researchers have already voiced concerns along these same lines (Sloboda, 1985; Deliège, 1989; Clarke & Krumhansl, 1990; Deliège & El Ahmadi, 1990; Butler, 1992; see also Nicholas Cook's chapter in this volume). But in order to research how one listens to a composition there are methodological difficulties to address. It is very difficult for the psychologist to design an experiment that investigates what may occur while listening to an actual musical composition. Performing, composing, and writing about music produce palpable results: a composition, an interpretation, an article, or a book—all evidence of what has actually happened between the individual and the music. But lis-

tening to music, *per se* does not yield such evidence, does not lay bare the intellectual and emotional responses of the listener, gives us no product to manipulate and analyze. Therefore, it is much more difficult to research music listening adequately (Sloboda, 1985: 151–54). Sloboda writes: "when I go to a symphony concert or listen to a gramophone record there may well be a lot of 'mental' activity, but there is not necessarily any observable physical activity. The principal end-product of my listening activity is a series of fleeting, largely incommunicable mental images, feelings, memories and anticipations" (p. 151). Nevertheless, can we say anything about how we perceive music? How can we research and measure what occurs when we listen to a musical piece?

When we listen to a composition, we hear an array of very complex temporal stimuli. Many musical variables occur simultaneously. It is a difficult task for the listeners themselves to listen attentively to a music performance. But unless there is control of variables, to what could we ultimately attribute the experimental results? "Musical material . . . cannot offer dimensions that are in all points controllable and systematically repeatable" (Deliège & El Ahmadi, 1990: 23). On the other hand, listening to an actual piece, *not* to artificially controlled laboratory stimuli, is what has provided us with the knowledge and the richness of our musical heritage.

I believe that when we listen, consciously or subconsciously, we *choose* what to focus on. This may be because of outstanding features of the music itself, certain features that are emphasized in the particular performance we are hearing, or because we just choose to focus on this or that element during that particular hearing. A piece of classical music is filled with more information than a listener can process in a single hearing. Music has a temporal, horizontal dimension (melody and rhythm), a vertical dimension (harmony), and a depth dimension (including combinations of dynamics, timbre, orchestration, and interpretation). Which among these dimensions does the listener choose to attend to from time to time? In designing an experiment, how can the freedom on the part of the subjects and the complexity on the part of the temporal stimulus be accounted for properly? How can they be tested and analyzed? If the psychologist asks the subjects to focus on this or that detail, already the task does not resemble listening as it occurs when we listen at our own discretion.

Given the complexity and the richness of the musical stimulus, listening implies choosing which elements to attend to. The artificial musical stimuli used in experiments limit the choices that a listener can make. For the reasons stated above, experimentally, the musical possibilities have been greatly impoverished in favor of reliable, controllable results. Clearly, many methodological difficulties must be considered in addressing listening to actual music, but shouldn't the music listening experience be at the very heart of the psychology of music?

In summary, listening to music in a laboratory condition does not duplicate listening to music outside of the laboratory. The experimental experience and the actual experience may differ in the complexity of the stimuli, their lack of aesthetic value, their brevity, and the very validity of the task. The research of the

last decades has provided a valuable foundation for understanding how we perceive controlled, brief, auditory sequences. What are needed are ways to develop a more sensitive understanding of the music listening experience and the mental representation of listening. The psychology of music must strive toward these goals.

Recently, there has been a growing interest in using excerpts or entire musical compositions as stimuli (Cook, 1987; Ruggieri & Sebastiani, 1987; Smith & Witt, 1989; Asada & Ohgushi, 1991). A number of studies have focused on examining some properties of live piano performances. With the help of computerized equipment, these studies investigate the similarities between the patterns of expressive characteristics that pianists display during repeated performances of a piece (Shaffer, 1981; Clarke, 1982; Gabrielsson, 1987).

In the remainder of this chapter I will review five recent studies that investigated aspects of music listening using recorded performances from the musical repertoire, all of which are approximately the same duration. All these studies investigated how listeners listen to music *as the music unfolded*. They did not ask the subjects only to respond post facto. By the nature of the required tasks, all these studies aimed to look at the subjects' perception of music as they heard the music.

Pollard-Gott (1983) researched whether repeated listening changes a listener's perception of a piece by focusing on the perception of musical themes. She asked musically trained and musically naïve subjects to rate the similarity of pairs of brief passages from the Franz Liszt Piano Sonata in B minor. This piece is in a single movement, structured around two principal themes and three minor themes. The standard division of a sonata in exposition, development, and recapitulation are all combined into a single, closely structured movement. In this study, the subjects heard the first 12 minutes (ending on measure 328) of a recording by Vladimir Horowitz.

The piece begins with the presentation of theme A (measures 8–13), a gap-filled melody, which is immediately followed by a contrasting, staccato theme B (measures 13–17). The piece continues with three variations of theme A and theme B in the following order: theme A, theme B, first variation on a2, a3, b2, b3, a4, and b4. The experimenter selected four examples of each of the two principal themes and recorded them separately.

The subjects were asked to write observations while listening, and, after listening to the first 12 minutes of the piece, they were asked to compare pairs of passages from the sonata judging their similarity on an 11-point scale (1 extremely dissimilar, 11 extremely similar). The listeners were also asked to decide whether each passage was best described as an example of theme A or theme B. It was found that, with repeated listenings, subjects improved in rating the similarity between the themes and also became more able to apply their understanding of the piece to passages that occurred much later in the same composition. Subjects with some musical background seemed to perceive the importance of the thematic material more quickly than the musically naïve subjects.

Clarke and Krumhansl (1990) have investigated how musically experienced subjects segment music, focusing on how they identify boundaries, locate excerpts within the overall piece, and estimate their duration. The researchers used performances of two stylistically different compositions as the stimuli: Karlheinz Stockhausen's *Klavierstücke IX*, and W. A. Mozart's Fantasie in C minor, K. 475. Both pieces are for solo piano, in a single movement, and contain a variety of different musical ideas. Each piece lasts approximately 10 minutes. Stockhausen's *Klavierstücke IX*, composed in 1961, is an atonal piece organized according to the principles of 12-note serialism. It is broadly divided in three regions, with the middle region being the most developed. Rhythmically and dynamically it is highly varied, alternating sections at different tempi. Mozart's Fantasie in C minor, K. 475 is also extremely varied rhythmically. There is tonal diversity and a multiplicity of different musical ideas. It may be divided in five sections, each one marked by different tempi and meter. The piece reflects the improvisatory style suggested by its title and displays many contrasts and sharply demarcated transitions from one section to the other.

To investigate the identification of boundaries, the subjects, after hearing each piece in its entirety, were asked to indicate where segmentation occurred in the music by pressing a foot pedal immediately after hearing a boundary. In later trials, they were asked to mark in the score the location where they perceived that segmentation had occurred, and were asked to describe briefly the features of the music that helped them form the boundaries. In general, the listeners agreed on the relative strength of the boundaries in both pieces, and reported this task to be quite easy. They cited similar criteria for making their judgments of where boundaries occurred in both compositions.

To investigate how listeners located brief excerpts in relation to the entire piece, the researchers first presented the piece in its entirety to the subjects, and then followed it with the presentation of brief segments. The subjects were asked to mark on a horizontal line on a piece of paper where they thought each segment occurred in the piece considering the left of the line to represent the beginning of the piece and the right of the line to represent the end.

The segments that occurred toward the beginning and toward the end of the piece appeared to be located closer to the center of the piece than was actually the case. This finding was true for both the atonal and the tonal composition. Clarke and Krumhansl propose that these results may be due to systematic changes in attention in the course of listening to music. In subsequent experiments, the subjects were asked to estimate the approximate duration of brief excerpts, being told that on average they lasted 30 seconds. The listeners were quite accurate in judging the relative duration of the excerpts in both pieces.

Deliège and El Ahmadi (1990; see also Deliège, 1989) have addressed listening and the perception of musical form. Specifically, these researchers investigated how musical cues help the listeners perceive boundaries in a musical composition, and whether there may be perceptual differences between musi-

cally trained and musically naïve subjects. The piece chosen as the stimulus was a recorded performance by Walter Trampler of *Sequenza IV*, a composition for solo viola by the contemporary composer Luciano Berio. This work is in six main sections and is characterized by contrasting structures. It fully exploits the timbre of the viola. The experiments presented the composition to the subjects three times. The first hearing was meant for the subjects to familiarize themselves with the piece, the second and third hearings were intended for the collection of the experimental data. The playing of the piece was synchronized with a timing mechanism on a microcomputer, and the subjects were asked to indicate where they perceived segmentation by depressing a key on the computer console.

The data revealed that the listeners utilized two organizing principles in their analytical hearings of this piece: "sameness" and "difference." The principle of sameness helped them constitute groups out of what they heard, while the principle of difference helped them differentiate what they listened to. Both musicians and nonmusicians showed important similarities in their perception of the composition.

Deliège also investigated how these same subjects located excerpts within this piece (reported in Deliège, 1989). In a subsequent experiment, first she explained to them that the piece was in six main sections, and then played for them 47 brief excerpts from the piece, ranging from 5 to 10 seconds. The subjects were asked to indicate after each excerpt in which of the six main sections of the piece each excerpt occurred. Significant differences were found in the accuracy with which some excerpts were localized. In general, the results were quite uniform for all subjects, regardless of their musical training. However, nonmusicians localized excerpts that occurred in the opening and the closing sections of the work better than they localized excerpts from the middle of the piece.

I am particularly interested in understanding how musicians with extensive training perceive music. Given the complexity of music, musicians are the ones who, potentially, have at their disposal the vocabulary necessary to describe what is occurring in the music. I am also interested in the perceptual freedom that one exercises when listening to music, and in understanding the differences among listeners. How do musicians capture what is happening in the music *as* the music unfolds? Given the scarcity of literature on this topic, this is a question that is as fascinating as is it difficult.

In a recent experiment (Aiello, Tanaka, & Winborne, 1990) we began to address what information musicians select from the music as it unfolds over time. Specifically, this experiment looked at which musical elements musicians report focusing on when listening to a complex musical work if they were free to report any of the elements present in the music. Graduate music students from the Juilliard School served as subjects in this study. They were asked to name on a sheet of plain notebook paper (not music paper) which musical elements they heard as they listened to a recording of the first movement of W. A. Mozart's

Symphony in G minor, K. 550, a work well known to all of them to eliminate any potential familiarity bias. The performance was by the London Symphony Orchestra, Carlo Maria Giulini conducting.

The subjects were asked to listen to this performance twice: during the first hearing, they were asked only to listen attentively. During the second hearing, *as* the music unfolded, they were to write down what they heard using any musical term that they wished. It was emphasized that they could write down any musical element that they wanted, and that this investigation addressed only the analytical aspect of listening to music.

The data were analyzed by grouping the musical elements listed by the subjects into categories. Strong individual differences emerged: some musicians' data reflected a gestalt holistic mode of perception, while the responses of other subjects showed an analytical, more segmented approach. These latter subjects focused on smaller details of the musical work. Regardless of the quantity of elements that the subjects reported listening to, they all reported attending most often to musical elements that are very salient in this movement: tonality, rhythm, and the use of specific chords. In addition, the data indicate that the listening strategies of each subject were not related to his or her major performance instrument or training in composition.

In a series of subsequent experiments (Aiello, Palij, and Gerstman, manuscript in preparation) we addressed whether the perceptual strategies of musicians remain consistent over repeated hearings of a piece, and evaluated their responses to compositions of different musical styles. The overall procedure was the same as that of our previous experiment, except that now subjects repeated the experiment listening to the very same pieces after approximately six weeks. In addition, we evaluated the subjects' responses to compositions of different musical styles presenting Mozart, Beethoven, Chopin, and a jazz improvisation by Miles Davis.

Generally, the data indicate that the subjects seem to write down less during the second experimental session. We think that this finding may be attributable to the nature of the experimental task. Perhaps, when the subjects repeat the experiment, they may wonder what they should write during the second time that is different from what they wrote during the first experimental session. In general, these subjects seem to list more elements when listening to Mozart than when listening to the other composers.

In my experiments, as well as in the other ones that I have reviewed here, the subjects listen to "real" musical stimuli, but their responses may or may not duplicate their actual experience when listening to music in a nonexperimental situation. But I look at all these data as steps being taken toward a direction, toward a kind of research in the psychology of music that addresses the ecological validity of listening to performances of the repertoire, and considers the processes that occur while listening to music.

In summary, I believe that listening to performances from the repertoire is what shapes our sensitivity to music. Many methodological difficulties must be

solved to investigate thoroughly how we perceive actual performances rather than controlled laboratory stimuli. But, unless we address these methodological difficulties we will never find potential solutions to them. Research in the psychology of music has begun to consider the importance of listening to musical performances, and there is much to discover.

## REFERENCES


Aiello, R., Palij, M., & Gerstman, L. Manuscript in preparation.

Aiello, R., Tanaka, J., & Winborne, W. (1990). Listening to Mozart: Perceptual differences among musicians. *Journal of Music Theory Pedagogy, 4*(2), 269–93.

Asada, M., & Ohgushi, K. 1991. Perceptual analyses of Ravel's *Bolero*. *Music Perception, 8*(3), 241–50.

Brendel, A. (1976). *Musical thoughts and afterthoughts*. Princeton, NJ: Princeton University Press.

Butler, D. (1992). *The musician's guide to perception and cognition*. New York: Schirmer Books.

Clarke, E. F. (1982). Timing in the performance of Erik Satie's *Vexations*. *Acta Psychologica, 50*, 1–19.

Clarke, E. F., & Krumhansl, C. (1990). Perceiving musical time. *Music Perception, 7*, 213–51.

Cook, N. (1987). The perception of large-scale tonal closure. *Music Perception, 5*(2), 197–206.

Cook, N. (1990). *Music, imagination and culture*. Oxford: Oxford University Press.

Copland, A. (1957). *What to listen for in music*. New York: McGraw-Hill.

Corredor, J. M. (1956). *Conversations with Casals* (A. Mangeot, Trans.). New York: Dutton.

Deliège, I. (1989). A perceptual approach to contemporary musical forms. *Contemporary Music Review, 4*, 213–30.

Deliège, I., & El Ahmadi, A. (1990). Mechanisms of cue extraction in musical groupings: A study of perception of *Sequenza IV* for solo viola by Luciano Berio. *Psychology of Music, 18*(1), 18–44.

Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review, 88*, 503–22.

Gabrielsson, A. (1987). Once again: The theme from Mozart's *Piano Sonata in A major (K. 331)*: A comparison of five performances. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (Publication No. 55). Stockholm: Royal Swedish Academy of Music.

Jackendoff, R. (1991). Musical parsing and musical affect. *Music Perception, 9*(2), 199–230.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Marsden, A. (1987). A study of cognitive demands in listening to Mozart's Quintet for Piano and Wind Instruments, K. 452. *Psychology of Music, 15*(1), 30–57.

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Meyer, L. B. (1967). *Music, the arts, and ideas*. Chicago: University of Chicago Press.

Pollard-Gott, L. (1983). Emergence of thematic concepts in repeated listening to music. *Cognitive Psychology, 15*, 66–94.

Ruggieri, V., & Sebastiani, M. P. (1987). New approaches to musical interpretations from a psychophysiological point of view: Analysis of some instrumental interpretations. *Musik Psychologie, 4*, 65–89.

Sessions, R. (1970). *Questions about music*. New York: Norton.

Shaffer, L. H. (1981). Performances of Chopin, Bach, and Bartok: Studies in motor programming. *Cognitive Psychology, 13*, 326–76.

Sloboda, J. A. (1985). *The musical mind: The cognitive psychology of music*. Oxford: Oxford University Press.

Smith, D., & Witt, J. (1989). Spun steel and stardust: The rejection of contemporary compositions. *Music Perception, 7*(2), 169–86.

# AUTHOR INDEX


Adler, S., 113
Agmon, E., 76, 86
Aiello, R., 40, 47, 279, 280
Aiken, H. D., 12
Anderson, J. R., 152, 217
Andrews, M. W., 185
Angier, R. P., 15
Arnheim, R., 3, 173
Asada, M., 277
Attneave, F., 183

Babbitt, M., 66, 198
Bailey, J., 260
Balzano, G. J., 182
Bamberger, J., 104, 108, 125, 131, 136
Baroni, M., 51
Bartlett, J. C., 180, 182–84, 186
Bengtsson, I., 260
Bent, I., 57, 64
Bernstein, L., 40, 43, 45, 49, 99
Bever, T. G., 40, 47–48
Bharucha, J. J., 192, 213, 216, 220–22, 226, 228–29, 233
Bienstock, S. F., 117
Bigard, E., 51
Blacking, J., 167
Boltz, M., 198, 241
Bourassa, G., 45
Bregman, A. S., 175
Brendel, A., 274
Bridges, V. A., 112
Brody, M., 197
Brown, H., 192, 195–96, 198, 200-203, 206–9, 254
Brown, I., 252
Browne, R., 198, 200
Bull, D., 184
Burrows, D., 43
Butler, D., 83, 86, 192, 195, 198, 200–202, 205, 208–9, 275

Campbell, J., 175
Carey, G., 160
Carlsen, J. C., 214, 218
Carlson, V. R., 252
Carrington, J., 43
Cassirer, E., 18
Castellano, M., 192, 226
Catel, C. S., 196
Cazden, N., 8, 196
Ceci, S. J., 152, 160
Chang, H., 99
Chen, M., 51
Chi, M. T. H., 152
Chomsky, N., 41, 48, 76
Clarke, D., 66
Clarke, E. F., 40, 41–43, 49, 51, 56, 66, 72, 77, 81, 86, 88, 92, 155, 258, 266, 274, 275, 277–78
Cohen, A. J., 184
Cohen, M. R., 30
Colley, B., 109
Cook, N., 65, 81, 277
Cooke, D., 59
Cooke, P., 43
Coon, H., 160
Coons, E., 4, 57
Cooper, G., 136, 148, 240
Copland, A., 274
Corredor, J. M., 274

Cross, I., 64
Crowder, R. G., 182
Cuddy, L. L., 184

Davidson, J. W., 158
Davidson, L., 102, 104, 107, 109, 113, 115–16, 125, 185, 258
Davis, J. B., 56
Deliège, I., 51, 71–72, 275–76, 278–79
Deutsch, D., 46, 64, 173, 175, 195, 198, 217, 273–74
Dewey, J., 15–16
DeWitt, L., 40, 45, 182
Divenyi, P. L., 214, 218
Dowling, W. J., 64, 66, 87, 88, 113, 173–74, 180–86, 188, 215, 221, 273
Drake, C., 188

Eco, U., 56
Edworthy, J., 184
El Ahmadi, A., 275–76, 278
Ericsson, K. A., 152, 160–62, 166

Farnsworth, P., 8
Farr, M. J., 152
Feinberg, I., 252
Feldman, D., 104
Feroe, J., 198, 217, 274
Fischer, K., 104
Fodor, J. A., 48
Forte, A., 208
Fox, D. B., 117
Fraisse, P., 240
Frances, R., 181
Fryer, R., 85
Fujitani, D. S., 180, 183, 184
Furth, A. G., 104

Gabriel, C., 99, 195
Gabrielsson, A., 155, 159, 258, 260, 277
Gabrielsson, B., 260
Galanter, F., 249
Gardner, H., 42, 100, 185
Garner, W. R., 186
Gaver, W. W., 4
Geary, J. M., 229
Gjerdingen, R., 79–80, 213
Glaser, R., 152
Glassman, N., 72
Goodnow, J., 105, 117
Gotlieb, K., 66
Gregory, H., 48
Gruson, L., 162

Halpern, A., 45
Handel, S., 240
Hansen, E., 192
Hanslick, E., 7
Hantz, E., 89
Hargreaves, D., 42, 99, 195
Harris, R., 78–79, 81, 85, 86
Harwood, D., 64, 173, 180, 184, 215, 221, 273
Hebb, D. O., 19, 233
Helmholtz, H., 191, 208, 230
Hermelin, B., 167
Hevner, K., 59
Hildebrandt, C., 136
Holleran, S., 64
Hollombe, A., 174
Howe, M. J. A., 152–53, 162–66
Howell, O., 64
Huron, D., 175

Imberty, M., 185
Inhelder, B., 105–6

Jackendoff, R., 40, 43, 49, 51, 64, 65, 70–71, 73, 75–76, 78, 87, 89, 136, 148, 159, 217, 218, 235, 261, 262, 274
Jacoboni, C., 51
Jairazbhoy, N., 226
Jacques-Dalcroze, E., 240
Jersild, A. T., 117
Jones, M. R., 64, 184, 188, 215, 241
Jusczyk, P., 45

Kamin, L. J., 160
Kaplan, E., 46
Keil, F. C., 185
Kellar, L., 40, 45
Kessler, E., 192, 195, 224–25
Klatt, D. H., 261
Knoll, R. L., 241, 246, 255
Kolinsky, M., 188
Konečni, V., 66
Kraehenbuehl, D., 4, 57
Kramer, J. D., 240
Krampe, R. T., 160–61
Krumhansl, C., 66, 72, 81–86, 88, 181, 185, 191, 192, 195, 205, 222–26, 274, 275, 278

Langer, S., 8–9, 40, 59
Lashley, K. S., 240
Lecercle, J., 88
Lee, C. S., 240
Lee, V., 11
Lerdahl, F., 40, 43, 49, 51, 64–65, 70–73, 75–76, 78, 85, 87–89, 136, 148, 159, 217, 218, 235, 261–62, 274
Levin, H., 46
Lewers, J. M., 55
Lewontin, R. C., 160
Liesch, B. W., 182
Lindblom, B., 51
Lisle, E. R., 64
Locke, S., 40, 45
Longuet-Higgins, C., 64, 240
Loy, G. D., 213
Luce, R. D., 249
Lyons, J., 261

Fraisse, P., 240
Francès, R., 181

McAdams, S., 41–42, 64
McClelland, J. L., 231
MacCurdy, J. T., 15–16, 22
McKern, P. E., 184–85
Mandler, G., 4, 57, 215
Manturszewska, M., 166–67
Marsden, A., 77, 86, 89, 274
Mead, G. H., 30
Mewhort, D. J. K., 184
Meyaard, J., 102, 106
Meyer, L. B., 3–4, 56–57, 64, 75–76, 89, 136, 148, 197, 214–16, 240, 260, 274
Meyers, C. S., 11
Michel, P., 112
Miklaszewski, K., 162
Miller, G., 52
Miller, L. K., 153, 167
Minsky, M., 57, 64
Mitchell, W., 207
Moles, A., 4, 57
Monahan, C. B., 181
Moog, H., 195
Moorhead, G. E., 112
Morongiello, B. A., 185
Mozart, L., 35
Mursell, J. L., 13
Musick, F. E., 229
Myers, C. S., 11

Narmour, E., 67, 214, 218
Nattiez, J. J., 56
Nketia, K., 43

O'Connor, N., 167
Ohgushi, K., 277
Olney, K. L., 231
Olson, R. K., 183
Osmond-Smith, D., 89
Overbeeke, C., 72

Palmer, C., 158, 188, 258
Parncutt, R., 230
Paulhan, F., 15
Piaget, J., 99, 105, 106
Pierce, J. R., 229
Pinkerton, R., 57
Piston, W., 208
Pollard-Gott, L., 277
Pond, D., 112
Pople, A., 86, 89
Proctor, G., 207

Rapaport, D., 14
Rasch, R. A., 241
Repp, B., 45, 258
Resnick, L., 102
Revesz, G., 43
Riemann, H., 209
Ries, N. L. L., 113
Roads, C., 51
Roche, S., 112
Rose, S., 160
Rosen, C., 90, 92
Rosenfield, I., 134
Rosner, B., 75, 76, 89
Ruggieri, V., 277
Rumelhart, D. E., 231
Russell, B., 33

Sachs, C., 40, 43
Sampson, J., 45
Samuel, A., 40, 45
Sandell, G., 83
Sargeant, D., 112
Schacter, S., 4
Schenker, H., 49, 73, 89–90, 208, 261, 266
Schmuckler, M. A., 214, 218, 221
Schoen, M., 11
Schoenberg, A., 82, 208
Schön, D. A., 131
Scripp, L., 100, 102, 104–7, 109, 125
Seashore, C. E., 154, 155, 261, 262
Sebastiani, M. D., 277
Sessions, R., 274
Serafine, M. L., 42, 72, 100, 195
Sergeant, D. C., 83, 112
Shaffer, L. H., 155, 277
Shahn, B., 133–34
Shepard, R. N., 82, 191, 192, 221, 222, 225
Shuter-Dyson, R., 99, 112, 195
Siefert, U., 78
Singer, I. B., 54
Singer, J., 4
Sloboda, J., 40–43, 45–47, 59, 64, 66, 99, 153, 157–58, 160, 162–64, 166–67, 173, 258, 273, 275–76
Smith, A., 88
Smith, D., 51, 54, 277
Smith, J., 152
Snodgrass, J., 249
Sosniak, L. A., 162–64, 166
Spender, N., 46
Stacey, P., 43
Steedman, M., 51, 240
Stefani, G., 56
Sternberg, S., 241, 246, 255
Stockhausen, K., 66
Stoeckig, K., 220–22, 228
Stravinsky, I., 7, 43, 44
Sundberg, J., 51, 155

Tan, N., 40, 47
Tanaka, J., 279
Taylor, J. A., 205, 214, 218
Tekman, H., 229
Teplov, B., 112
Tesch-Romer, C., 160, 161
Thomas, F., 252
Thorpe, L. A., 184, 185
Todd, N. P., 155, 158, 241, 258
Todd, P. M., 213
Torff, B., 107, 125

Tramo, M. J., 229
Trehub, S. E., 99, 184, 185
Trotter, R., 104

Wachsman, K., 43
Wapnick, J., 45
Warren, R., 45
Watkins, A. J., 180–81
Wedin, L., 59
Weld, H. P., 10
Welsh, P., 102, 109, 116, 125
Wertheimer, M., 173
Wertsch, J. V., 112
West, R., 64, 85
Winborne, W., 279
Wing, H. D., 112
Winn, J. A., 40, 43, 44
Witt, J., 51, 54, 277
Wittgenstein, L., 132
Wohlwill, J. P., 103
Wolpert, R., 67–70, 83
Woodworth, R. S., 17

Yeston, M., 240
Youngblood, J. E., 4, 57

Zatorre, R., 45
Zenatti, A., 99
Zukofsky, P., 246

# SUBJECT INDEX


"A, b, c, d," 116–18
Activation, 233
Aesthetic experience, 12
Aesthetic positions
  absolute expressionist, 7
  absolutist, 7, 29
  expressionist, 7, 8
  formalist, 7, 8
  referential expressionist, 7
  referentialist, 7, 29
Aesthetics of music, 8
Affect, 14, 15, 17, 19, 34. *See also* Emotions; Feeling
Affective behavior, 18
Affective experience, 19, 20, 22
  in music, 20
Albinoni, T., Adagio for Strings, 59
Ambiguity
  in language, 49
  in music, 49
Associative networks, 231
Aural analysis, 81
Aural training, 81

Bach, J. S.
  Bouree (Suite No. 3 in C major for violoncello), 175
  Fugue in C minor (Book I of *The Well-Tempered Clavier*), 177, 178
  Goldberg Variations, 67
  "O Ewigkeit, du Donnerwort," 49, 50
  Prelude in C major (Book I of *The Well-Tempered Clavier*), 264–66, 269
  *The Well-Tempered Clavier*, 240
"Barbara Allen," 187
Basilar membrane, 227
*Bassbrechung*, 49
*Basse fundamentale*, 86
Beat fraction
  displacement of, 250
  imitation of, 244, 252–54
  perceptual judgment of, 244, 246, 253, 254
  production of, 243–45, 253, 254
Beat interval, 243
Beethoven, L. van
  Fifth Symphony, 57, 177
  Ninth Symphony, 24
  Quartet, op. 18, no. 5, 49
  Sonata, op. 90, 89
Berlioz, H., "March to the Gallows" (Symphonie Fantastique), 24
Bernstein, L., "Somewhere" (*West Side Story*), 181, 182
Body movement of solo musicians, 158
Boulez, P., *Le Marteau sans Maître*, 70, 87
Boundary
  of figures, 146
  of groupings, 148
  of phrases, 47, 72, 278–79
Brahms, J., Capriccio, op. 76, no. 2, 47

Categorical perception. *See also* Phonology
  in language, 45
  in music, 45
Children's drawings of rhythms, 135
Children's singing, 113, 184–86
Children's songs compared to adults' songs, 109–26, 125
Chopin, F.
  Etude, 156
  Mazurka in A minor, op. 17, no. 4, 51–53, 266–68

Chopin, F. (*continued*)
Prelude in f-sharp minor, 262–64
Chords, effect of culture on the perception of, 195–96, 220–21
Chromaticism, 204–5, 208
Chromatic set, 194
Circle of fifths, 225, 228
Class piece, 135, 138
Click experiments
in language, 48
in music, 48
Communication
language as, 42–49, 55
music as, 42, 55
in music, 9, 34–36, 55
Competence, 76
Competition winners, 160
Computer-instrument links, 155
Connectionist model, 231
Conservatory students, 160
Consonance, 84, 196–97, 214
Context
distal, 226
proximal, 226
Contour schemes, 101, 106, 113, 116, 118, 120–23, 126. *See also* Melodic contour; Melody
Conventional music notation, 144
Counting, 139
Creativity, 165

Deep structure
in language, 48–51, 76
in music, 48–51, 76
Designative behavior, 20–21
Developmental psychology, 101
Diary of activities, 161
Dissonance, 196–97
Donizetti, G., *Lucia di Lamermoor* (sextet), 187–88

Ear training, 81
Emergent properties, 235
Emotions, 8, 10–17, 19, 54, 59. *See also* Affect; Feeling
conflict theory of, 15
emotional behavior, 12, 18–20
emotional experience, 11, 20
psychological theory of, 14–17, 21
in response to music, 9, 13, 59, 164
Exceptionality, 160
Expectation
cognitive processes in, 214
conscious, 27
in music, 22–25, 27
schematic, 216, 222, 229–31
tonal, 214–15, 218, 230
unconscious, 27
veridical, 216
Expressive markings, 145

Feedback model of production, 248, 250–51
Feeling in music, 8. *See also* Affect; Emotions
Felt-paths, 134
Figurative thinking, 105–6
First teacher, 164
Focus of attention, 137
"Frère Jacques," 57–58
Functional differences, 144

Gestalt psychology, 9, 173–74, 186
Gestures, 136
Ghost beat, 140–41, 143
Grammar
in language, 46–49, 51
listening, 87
of music, 46–49, 51
transformational, 51
Grouping
of figures, 135
of phrases, 135

Hair cells, 227
"Happy Birthday to You," 102–3, 106, 108, 120, 125–26
Harmonic functions, 209
Harmonic partials, 196
Harmonic series, as universal for music, 43
Harmonics, 227
Harmony, 49–51, 195. *See also* Key; Tonality
Haydn, F. J., Quartet in B minor, op. 33, no. 1, 90–91
Hearings of a piece, 133, 138
Heritability, 160
Hierarchy
event, 218
linguistic, 75
musical, 75
tonal, 218, 222
Home environment, 164
"Hot Cross Buns," 68–69

Imitation in timing, 242, 244
Inferences, 218
Information theory, 57
Innateness
of language, 41–42
of music, 41–42
of tonal expectations, 230
Inner ear, 227
basilar membrane in, 227
hair cells in, 227
Instrumental lessons, 163
Interpretation of music, 55, 260
Interval, 199, 205, 209
class, 194
patterns, 178, 181
"Itsy Bitsy Spider," 119
Ives, C., Concord Sonata, 206–7

"Jack and Jill," 122

Key, 180–81, 197, 200–201, 208, 225. *See also* Tonality; Tonic
  centers, 200–201
  identification, 84–85
  level, 195, 197, 208
Knowing about, 132, 150
Knowing how, 132, 150

Language, 40–41, 44, 46, 59
  deep structure of, 49
  innateness of, 41
  reading, 46
  surface structure of, 49
Languages
  African, 43
  Oriental, 42
  Western, 42
Learning
  in neural net models, 230–32
  in tonal expectations, 214
Ligeti, G., *Atmosphères*, 87
Listening
  to compositions, 274–77
  grammar, 87
Liszt, F.
  *Fünf kleine Klavierstücke*, 73–75
  Sonata in B minor, 277

Mahler, G., First Symphony, 57–58
Major mode, 198
Meaning
  absolute, 6
  designative, 31
  determinate, 32–33
  embodied, 31
  evident, 32
  hypothetical, 32
  in language, 54–56
  in music, 6, 29, 30, 33, 34, 54–56
  objectification of, 33
  referential, 6
Melodic contour, 173, 175–76, 178–84, 186, 188. *See also* Contour schemes; Melody
Melody, 43, 173, 175, 184. *See also* Contour schemes; Melodic contour
  memory for, 181–82, 184
  perception of, 173
Meter
  in music, 42
  in poetry, 42
Microtonicization, 206
Minor mode, 204
Models
  associative, 231
  connectionist, 231
  neural net, 231
  parallel distributed, 231
Motivation, 162
Motives, 136

Mozart, W. A.
  Fantasie in C minor, K. 475, 72, 278
  Symphony in G minor, K. 550, 197–98, 280
Multidimensional scaling, 225
Multiple hearings, 137–38
Music
  aesthetics of, 8
  affect in, 15–17, 19
  Balinese, 84, 87, 192
  emotions in, 9–11, 15–17, 19, 59
  feeling in, 8
  Indian, 83, 192, 226
  innateness of, 41–42
  origins of, 43
  physical manifestation when listening to, 59
  psychology of, 8, 66–67, 274
  serial, 83
  reading at sight, 46–47
  theory of, 64–65, 79, 81
  tonal, 72
Musical ability, 101
  development of, 42, 99, 102, 104, 111–12, 126–27, 184
Musical excellence, early signs of, 163
Musical experience, 9, 21–22, 28
Musical knowledge, 133
Musical motivation, 165
Musical performance, 155–59, 259–69
Musical talent, 102, 127–28, 160–62
Musical training, 106
Musical understanding, 111
Musicians, 68, 242–43, 279–80

Neural network, 231–35
  associative, 231
Neural nets, 231–35
Neurons, 232

"Old MacDonald Had A Farm," 123
Operational thinking, 105–6
Overlap ratio, 227

Parallel distributed model, 231
Parents, 164
  as support, 166
Performance, 76
Performances at sight, at the piano, 47, 157. *See also* Pianists
Phonology. *See also* Categorical perception
  in language, 44
  in music, 44
Phrases. *See also* Boundary; Segmentation
  in language, 47–48
  in music, 47–48, 135
Physiological responses to music, 13, 59
Pianists, 47, 157, 261–69
Pitch, 42, 193
  chroma, 175
  class, 194

Pitch (*continued*)
height, 175
relations, 202
set, 181, 193–94, 199
Practicing strategies, 162
of violinists, 160–61
Pragmatics, 76
Priming task, 219, 228
Probe-tone technique, 82, 191, 223
Profile, 223
Project Zero, 100, 113, 125
Psychoacoustics, 274
Psycholinguistics, 40, 275
Psychophysical scales, 254

Reaction time, 219
*Recitativo*, 43
Reflective confrontation, 149
Repeated patterns, 138
Resolution, 215
Responses
of college students, 67
of music students, 73
of musically trained, 72, 277–78
of musicians, 47, 68–69, 201, 219, 221, 243–59, 278–79
of nonmusicians, 47, 51, 68–69, 221, 277, 279
of pianists, 157–58, 261–69
Rhythm. *See also* Time; Timing
figural drawings of, 135
metric drawings of, 135
in music and poetry, 42
simple, 132
Role models, teachers as, 165
Roughness, 227
"Row, Row, Row Your Boat," 114, 121
*Rubato*, 158, 264, 266, 268–70
Rules, 235

Satie, E., *Gnossienne* No. 5, 266–67, 269
Scales, 106, 195, 199
Schema, 215–17
Schenkerian analysis, 49, 66, 70
structure of, 76
theory of, 51
Schoenberg, A.
charts of the regions, 82, 86
*Sechs Kleine*, 79
*Klavierstücke*, op. 19, no. 3, 79
*Structural Functions of Harmony*, 82
Schubert, F., Sonata in D major (D. 664), 203
Scriptism, 78
Segmentation in music, 71, 278–79
Semantics
in language, 54
in music, 54
Semiotics, 56
Set-theory analysis, 79
Shift focus, 149
Sight-reading at the piano, 47, 157
"Somewhere over the Rainbow," 123
Space, 142, 147
Space-of-time, 142
Spatial relationships, 222
Spatial representation, 222–23
Spectral overlap, 228–29
*Sprechstimme*, 43
Stability, 214
Standard rhythm notation, 136
Stockhausen, K., *Klavierstücke IX*, 72, 278
Stravinsky, I., *Petroushka*, 222
Students at special music schools, 160
Style, 36, 56
Surface structure
in language, 48–51
in music, 48–51
Suspense in music, 25, 26

Teacher-pupil relationship, 164
Teachers as role models, 165
Tendencies, 22–23
in music, 22, 28
Tensions, 22, 25
Theorism, 81
Thinking
figurative, 105
operational, 105–6
Time. *See also* Rhythm; Timing
perception of, 242, 244
ratios of, 240
Timing, 154, 241–44, 261–62. *See also* Rhythm; Time
differences in, 268
gestures, 265
of piano performances, 264–69
Tonality, 73, 180, 191, 194, 197, 214, 223. *See also* Key; Tonic
tonal center, 201
tonal closure, 72
tonal hierarchy, 86, 222
tonal structures, 70, 87
Tone, 42
Tonic, 198. *See also* Key; Tonality
Torus, 225
Transpositions, 180
Tritone, 196, 200, 209
"Twinkle, Twinkle, Little Star," 68–69, 109–11, 139–41, 183

Units of description, 133
Units of perception, 133
*Urline*, 49
*Ursatz*, 76

Violinists, 160–61

Wagner, R., *Die Meistersinger*, 27

Yankee Doodle," 183

# Advance Praise

"It is a demanding and rewarding book. I find it a unique and scholarly document, an exegetic anthology of various perspectives and perceptions of that most mysteriously abstract art—Music. It is a break-through to the clarification of once considered dead-ends. Philosophically sound and scholarly researched, this is a formidable accomplishment."

—David Diamond, *Professor and Lecturer of Composition, The Juilliard School*

"This collection should be welcomed by students and teachers of music, particularly in conservatories and in colleges and universities. Too often we overlook the psychology of hearing and listening and the relationship between modes of perception and methods of analysis. This collection brings to the reader in a concise way the widest range of issues, presented through both historic and new essays, about how we hear and understand music. This overview will be useful for practicing musicians and historians, as well as for students of psychology and philosophy."

—Leon Botstein, *President of Bard College and Music Director and Conductor of the American Symphony Orchestra*

## About the Editors

Rita Aiello, Ed.D., is a professor at the Juilliard School, an adjunct associate professor in the Department of Psychology at New York University, and an adjunct professor at the Manhattan School of Music. Dr. Aiello, a pianist, researches the perception and cognition of music.

John A. Sloboda, Ph.D., is Professor of Psychology at the University of Keele, Staffordshire, and is author of *The Musical Mind* (Oxford, 1985).

*Cover art:* The Three Musicians *by Pablo Picasso, courtesy of the Philadelphia Museum of Art*
*Cover design by Linda Roppolo*

9 780195 064766 90000

ISBN 0-19-506476-3

# La inteligencia hábil

## El desarrollo de las capacidades cognitivas

Angus Gellatly (Comp.)



Estrategias de aula

AIQUE

# LA INTELIGENCIA HÁBIL

## El desarrollo de las capacidades cognitivas

# LA INTELIGENCIA HÁBIL

## El desarrollo de las capacidades cognitivas

Angus Gellatly (compilador)

AIQUE

**Título original en inglés:** *The Skilful Mind. An Introduction to Cognitive Psychology.*
**Traducción:** Miguel Wald
**Supervisión de traducción:** Berta Braslavsky

© 1986 Open University Press, Milton Keynes, Philadelphia, U.S.A.
Esta obra se publica bajo acuerdo con Open University Press, Buckingham, totalmente en idioma inglés.

© Copyright Aique Grupo Editor S.A.
Arturo Jauretche 162 (ex Méndez de Andés) (1405) Capital Federal
Tel.-Fax: 958-0200

Hecho el depósito que previene la ley 11.723
LIBRO DE EDICIÓN ARGENTINA
I.S.B.N. 950-701-377-6
Primera edición

La reproducción total o parcial de este libro en cualquier forma que sea, idéntica o modificada, escrita a máquina por el sistema "multigraph", mimeógrafo, impreso, etc. no autorizada por los editores, viola derechos reservados.

# ÍNDICE

**Agradecimientos** ..... 9
**Prefacio** ..... 11

## *Sección 1*
## *La naturaleza de las habilidades*

**1. Cognición y psicología** ..... 17
*Angus Gellatly*
Dos ejemplos de habilidad ..... 17
Algunas características de la habilidad ..... 19
Habilidad para la navegación ..... 20
¿Qué es la cognición? ..... 21
El estudio de la cognición ..... 23
El estudio contemporáneo de la cognición ..... 25
Conocimiento psicológico ..... 29
Resumen ..... 33

**2. ¿Qué es la habilidad?** ..... 35
*John Sloboda*
Las características de la habilidad ..... 36
La estructura de las habilidades ..... 45

**3. La adquisición de la habilidad** ..... 49
*John Sloboda*
El talento y la adquisición de habilidades ..... 50
La práctica y la adquisición de habilidades ..... 52
Explicación de la ley de las mejoras decrecientes ..... 55
Mejorando los efectos de la práctica ..... 57
Un talento excepcional - un estudio de caso ..... 61

**4. La lectura: un estudio de caso de habilidades cognitivas** ..... 65
*John Sloboda*
Los movimientos oculares en la lectura ..... 66
La lectura como aplicación de conocimientos ..... 68
El papel del discurso en la lectura ..... 71
Leer con y sin sonido ..... 72
Dificultades en la lectura ..... 74
Las estrategias de lectura y el estudio ..... 76
Conclusiones ..... 77

## *Sección 2*
## *Las habilidades de la memoria*

**5. La naturaleza de la memoria** ..... 81
*Angus Gellatly*
Memorización automática y esforzada ..... 83

Más sobre la jerga de la memoria ....84
La naturaleza de la memoria de corto plazo ....85
La naturaleza de la memoria de largo plazo ....87
Codificación y recuperación ....90
Resumen ....95

**6. Adquisición de habilidades de la memoria** ....97
*Angus Gellatly*
Estrategias de ensayo y categorización ....100
Diferencias culturales en el rendimiento de la memoria ....102
El concepto de metamemoria ....104
Memoria y conocimiento ....107
Aptitudes de memoria especiales en los niños ....108
Conclusión ....110

**7. La habilidad de memoria excepcional** ....111
*Ian M. L. Hunter*
Trasfondo biográfico y social ....112
Mejora intencional y mejora casual ....120
Conclusión ....124

**8. ¿Cómo se pueden mejorar las habilidades de la memoria?** ....125
*Angus Gellatly*
La tradición mnemónica y la terapia de la memoria ....125
Mejora de la memoria de la gente normal ....127
El mejoramiento de la memoria en los niños con retardo mental leve ....131
Terapia de la memoria para aquellos que han sufrido daño cerebral ....133
La mejora del testimonio legal ....135
Ayudas para la memoria en los exámenes ....137
Resumen ....140

***Sección 3***
***Las habilidades de comprensión y comunicación***

**9. Las habilidades del hablante** ....143
*Don Rogers*
Hablar como habilidad ....144
El arte de hablar ....147
Conclusión ....157

**10. Las habilidades del oyente** ....159
*Don Rogers*
La comprensión del lenguaje ....159
La comprensión del significado social de las oraciones ....163
Más allá de la oración ....165
Mejorar las habilidades del hablante y del oyente ....168
Conclusión ....173

**11. Las habilidades sociales** ..... 175
*Helen Graham*
La competencia social como habilidad ..... 175
Las habilidades sociales ..... 178
Las investigaciones sobre interacción social ..... 185
Entrenamiento en las habilidades sociales ..... 187
Grupos de entrenamiento ..... 188

**12. ¿Cómo mejorar las habilidades de aprendizaje?** ..... 191
*James Hartley*
La psicología cognitiva y el aprendizaje ..... 191
Implicaciones para la enseñanza de habilidades de estudio ..... 201
Conclusiones ..... 205

***Sección 4***
***La habilidad y el pensamiento***

**13. La habilidad en el razonamiento** ..... 209
*Angus Gellatly*
Deducción y razonamiento ..... 211
Lógica, cultura y lenguaje ..... 219
El razonamiento como habilidad ..... 221

**14. La resolución de problemas** ..... 225
*Angus Gellatly*
La representación de problemas ..... 227
El uso de heurísticas ..... 232
El entrenamiento en las habilidades de resolución de problemas ..... 236

**15. La creatividad y las habilidades cognitivas** ..... 241
*Alan Branthwaite*
La naturaleza del pensamiento creativo ..... 241
El proceso creativo ..... 244
Los procesos creativos y el funcionamiento de los dos hemisferios ..... 249
El desarrollo de las habilidades del "hemisferio izquierdo" ..... 251
El desarrollo de las habilidades del "hemisferio derecho" ..... 253
La superación de los bloqueos en el pensamiento creativo ..... 256
Conclusión ..... 257

**16. Las computadoras y la cognición** ..... 259
*John Sloboda*
¿Qué es una computadora? ..... 260
La metáfora de la computadora para la mente ..... 262
¿Las computadoras piensan? ..... 263
Programas y simulaciones ..... 268

**Referencias bibliográficas** ..... 273

# Agradecimientos

Estoy sumamente agradecido a los muchos que me ayudaron en la realización de este libro. En particular, quisiera agradecer a E.M. Bestwick la información sobre la discriminación del sexo de los pollitos; a Jonathan Dancy, por sus comentarios sobre el capítulo 13; a John Coleman, que hizo los dibujos e ilustraciones; a Margaret Woodward y Dorothy Masters, por todas las horas de paciente y sufrido mecanografiado y, finalmente, a todos los que aportaron su esfuerzo en las interminables revisiones y con su apoyo constante.

# Prefacio

"El hombre es la medida de todas las cosas", decía el filósofo griego Protágoras; y, si dejamos de lado la política sexual de la terminología empleada (ver también más abajo), sus palabras deberían proporcionar a los psicólogos alimento para el pensamiento. Lo que Protágoras parece haber querido decir es que toda comprensión se realiza, finalmente, en términos humanos, y esto no podría ser más cierto que en el caso de la psicología. Es verdad que mucho de lo que aprendemos sobre la psicología humana lo hacemos mediante comparaciones con otras especies, estudiando las propiedades de las células nerviosas o de conjuntos de células, o buscando analogías entre los seres humanos y las máquinas. Todos esos emprendimientos contribuyen a aumentar nuestra comprensión de las formas en que funcionan los seres humanos. Sin embargo, en el análisis final, las habilidades y los logros humanos deben comprenderse solamente en términos humanos. En esta introducción a la psicología cognitiva habrá oportunidades para referirnos a estudios realizados con animales, a hallazgos de las neurociencias y a la importancia de las computadoras para comprender la mente, pero estas oportunidades serán relativamente pocas. Nuestro interés se centrará fundamentalmente en el ser humano y en sus habilidades.

Este libro ofrece un análisis de las habilidades humanas, las formas en que éstas se adquieren y las técnicas mediante las cuales se las puede mejorar. El foco estará puesto fundamentalmente en habilidades cognitivas comunes tales como hablar, recordar y razonar; como éstas son habilidades que se suelen considerar naturales, pueden aparecer, superficialmente, irrelevantes. Es importante no equivocarse en este sentido. Por ejemplo, la capacidad de leer y escribir o la de correr maratones son ejemplos extremos de habilidades que en una época se creían sólo eran accesibles a unos pocos individuos pero que, en la actualidad, muchos dominan. La frecuencia con la que se despliega una habilidad dentro de una población depende de factores ajenos a la dificultad intrínseca de esa habilidad, del mismo modo que los límites del desempeño experto se van ampliando constantemente mediante el ejercicio del ingenio humano. Las habilidades comunes pueden ser tan notables como las no comunes; en los próximos capítulos el lector podrá admirar las complejas acciones de las que son capaces tanto él mismo como otras

personas. La investigación psicológica ilumina los detalles de estas acciones y sus descubrimientos ayudan a justificar el orgullo legítimo por nuestros logros, que todos tenemos derecho a disfrutar.

El libro comprende dieciséis capítulos escritos especialmente por integrantes del Departamento de Psicología de la Universidad de Keele. Cada capítulo puede servir independientemente como introducción a un tópico en particular y por lo tanto el libro podría leerse, al menos teóricamente hablando, en cualquier orden. De todos modos, hay varios temas que se desarrollan y elaboran a través de todo el libro. Estos temas enlazan los diferentes capítulos y les proporcionan unidad y coherencia. Debido a esta organización, es conveniente que al menos los primeros tres capítulos se lean en orden, antes de encarar otras partes del libro.

El capítulo 1 ofrece una introducción a los conceptos de *cognición* y *habilidades cognitivas*, los métodos que se pueden usar para estudiarlas y la naturaleza del conocimiento que constituye la psicología cognitiva. Dos temas fundamentales del libro cobran una importancia especial en este capítulo. Primero, la cognición es una cuestión de habilidad y las habilidades cognitivas son análogas en muchos sentidos a las habilidades corporales. Segundo, el psicólogo cognitivo debe jugar roles gemelos: el de científico y el de entrenador.

En los capítulos 2 y 3 se encaran en mayor detalle las ideas sobre las habilidades en general y sobre las habilidades cognitivas en particular. Se identifican las características de la habilidad y se describen diversas formas de adquisición de las mismas. También se presentan ciertas nociones sobre cómo organizamos nuestra conducta en la persecución de los objetivos cotidianos. Más adelante, en el capítulo 4, se examinan y ejemplifican todos estos aspectos de la habilidad con respecto a la lectura. Se identifican varias habilidades de lectura, mostrando diferencias para distinguir los lectores expertos de los inexpertos; también se hacen algunas sugerencias para mejorar la lectura.

Los capítulos 5 a 8 se ocupan de la memoria. En cada uno de ellos se destaca otro tema del libro: las habilidades deben entenderse en su contexto social. En primer lugar, se ofrece al lector una introducción al tópico de la memoria, así como también al lenguaje y los métodos con que los psicólogos intentan analizarlo. A continuación, el capítulo 6 describe cómo los niños, empezando por las habilidades pasivas de la memoria que los humanos comparten con otros animales, se las arreglan para desarrollar las técnicas memorísticas con fines determinados según se practican en su propia cultura. En el capítulo 7 se analiza en detalle la destreza excepcional de una tarea, y también la destreza memorística específica. El capítulo 8 está dedica-

do a analizar cómo lo que se sabe sobre la memoria puede usarse para mejorar su rendimiento en todos, en general, en los niños con retardo mental leve (educables), en personas con lesiones cerebrales y otros.

La comunicación y la comprensión son el tópico común de los capítulos 9, 10, 11 y 12. El capítulo 9 describe los muchos niveles de habilidad que debe dominar el individuo para poder hablar de manera coherente y comprensible. El habla resulta especialmente apta para un análisis en términos de habilidad. También lo es al escuchar, tema del capítulo 10. El oyente es presentado como un receptor activo (no pasivo) de información, que ejercita habilidades complementarias a las del hablante.

Muy vinculadas a las habilidades de lenguaje de ambos tipos (del hablante y del oyente) están las habilidades sociales, que constituyen el tema del capítulo 11. Se identifican las señales inconscientes que los actores sociales transmiten y reciben constantemente, y se describen y evalúan métodos recientemente diseñados para mejorar las habilidades sociales. El aumento y la mejora de las habilidades existentes es también tema fundamental del capítulo 12. En este caso, se analizan las habilidades aplicadas al estudio; se describen investigaciones sobre la naturaleza de las habilidades de estudiar y aprender, seguidas nuevamente por una evaluación de los métodos para mejorarlas.

Los capítulos 13 a 16 se centran en el pensamiento y las habilidades para la resolución de problemas de diferentes tipos. El primero de estos capítulos examina algunas clases de razonamiento que suelen resultar difíciles, algunos que siempre se hacen automáticamente y la relación entre ambos. Una vez más, resulta evidente el carácter social de las habilidades cognitivas: lo que es "natural" para un miembro de una cultura puede resultar extraño para alguien de una cultura diferente. En el capítulo 14, el punto central es la resolución de problemas en general, sus principales características importantes y los hábitos inconscientes que pueden actuar como obstáculos con mucha facilidad. Una vez más, se plantea la cuestión de si se pueden enseñar y entrenar con éxito las habilidades cognitivas de este tipo. El mismo tema aparece en el capítulo 15, en relación con la creatividad y la resolución de problemas. Este capítulo se basa en registros históricos y en estudios de laboratorio. Se examinan varios aspectos del pensamiento creativo pensando en sugerencias para su mejoría. Finalmente, el capítulo 16 se concentra en la analogía entre las computadoras y la mente y en la cuestión de si se puede (o si alguna vez se podrá) decir que las máquinas piensan. Aunque las ideas tomadas de la computación informan gran parte del análisis y la teorización en los capítulos anteriores, el capítulo 16 ofrece la posibilidad de considerar

en profundidad la analogía con la mente. Sin embargo, aun aquí se mantiene el énfasis en la dimensión humana de la cognición.

Estos son, entonces, los temas fundamentales del libro: la cognición se entiende mejor en términos de habilidad; las habilidades deben verse en sus contextos sociales; mediante el entrenamiento pueden lograrse mejoras deliberadas en la habilidad; es función del psicólogo cognitivo tanto investigar la naturaleza de las habilidades como aplicar el conocimiento obtenido para ayudar a los que desean mejorar sus habilidades cognitivas.

# *Sección 1*

## *La naturaleza de las habilidades*

# 1

# Cognición y psicología

*Angus Gellatly*

Aunque en este libro nos concentraremos en la habilidades cognitivas y en otras que también nos resultan familiares a todos, es útil, para entender inicialmente la naturaleza de la habilidad, comenzar con dos ejemplos bastante exóticos. El primero es de una habilidad relativamente simple pero limitada a unos pocos miembros de la población. El segundo es de una habilidad extremadamente compleja que ha sido desarrollada en una cultura diferente de la nuestra y que, también, sólo realiza un número relativamente escaso de practicantes.

## Dos ejemplos de habilidad

### El sexo de los pollos

La granja es una actividad competitiva y, por varias razones, resulta económicamente conveniente determinar las identidades sexuales de los pollos tan pronto como sea posible. Como los machos y hembras pequeños son visualmente muy parecidos, incluso en la apariencia de sus órganos genitales, en Occidente se pensaba que la determinación del sexo de los pollos sólo era posible después de que los machos comenzaban a tener cresta. Pero durante el primer cuarto de este siglo los granjeros de Occidente empezaron a escuchar rumores sobre la existencia en Japón de "expertos en el sexo de los pollos", que eran capaces de discriminar machos de hembras al día siguiente de nacer. Estos rumores continuaron difundiéndose en el decenio 1930-1940. Se invitó a los expertos japoneses a exhibir sus habilidades en los Estados Unidos y en Europa y, a continuación, a dar instrucciones a sus anfitriones. Los japoneses demostraron que sosteniendo los pollos de la manera adecuada para revelar mejor la zona genital se podía aprender a distinguir machos de hembras. Más aun, los expertos podían clasificar los pollos a gran velocidad y con mucha precisión.

Como consecuencia de este encuentro entre Oriente y Occidente, se establecieron en los Estados Unidos y en Gran Bretaña escuelas para enseñar a determinar el sexo de los pollos. Lo habitual era que los alumnos se entrenaran durante un período de entre 4 y 6 semanas antes de obtener un título habilitante. En esta etapa, el nuevo profesional podía fijar el sexo de unos 200 pollos en 25 minutos, con una precisión del 95% o más (Lunn, 1948). Después de años de experiencia, la cifra podía aumentar a 1000 o 1100 pollos por hora, y se informó sobre el caso de un experto que llegó a determinar el sexo de hasta 1400 pollos en una hora, con una precisión del 98%. La tarea de determinar el sexo de los pollos es sumamente repetitiva y exige un alto nivel de atención, pero permite el desarrollo de sorprendentes niveles de habilidad.

En los años posteriores al establecimiento de las escuelas de entrenamiento, a fines del decenio 1930-1940, los graduados podían aspirar a empleos bien pagados, con la perspectiva adicional de ingresos todavía superiores en la medida en que se convertían en expertos en la tarea. Pero en los años cincuenta y sesenta el mercado laboral sufrió una aguda retracción. Esto se debió fundamentalmente al éxito de técnicas de reproducción que llevaron al desarrollo de cepas de pollos con diferencias de color según el sexo. Ahora resultaba fácil distinguir a los machos de las hembras y ya no se necesitaban los costosos servicios de los profesionales en la determinación del sexo de los pollos. La matrícula en la escuelas de entrenamiento cayó marcadamente cuando se deprimieron los ingresos y las posibilidades de empleo. Parecía que la historia del experto en pollos en Occidente iba a ser breve. Pero entonces, hacia fines del decenio 1960-1970, hubo un nuevo vuelco, otra vez a partir de las actividades de los productores. Los experimentos que se desarrollaron con apareamiento selectivo habían producido pollos con huesos más livianos y que por lo tanto requerían menos alimento para alcanzar su peso de madurez. Estas nuevas cepas resultaron populares entre los productores masivos conscientes de los costos, que criaban animales para consumo. Pero muchas de las nuevas cepas sufrían una falla particular. En la búsqueda de animales con huesos más livianos, el vínculo entre el sexo y el color de los pollos se había roto. Los expertos en la determinación del sexo de los pollos volvieron al mercado.

## Algunas características de la habilidad

Puede parecer extraño comenzar un libro sobre la psicología de la cognición con una breve historia sobre un aspecto bastante desconocido de la granja. Pero, de hecho, la determinación del sexo de los pollos es un ejemplo excelente de una habilidad humana relativamente simple, que muestra muchas de las características comunes a todas las habilidades. Una de ellas es la aparente imposibilidad, a los ojos del no iniciado, de la tarea que ha de desempeñar.

En el circo, los logros de malabaristas, acróbatas y trapecistas parecen sorprendentes, pero en la vida cotidiana también abundan las habilidades sorprendentes. Algunas de ellas, como caminar y hablar, son tan comunes que no llaman la atención, mientras que otras, aunque están confinadas a unos pocos ejecutantes, sólo se realizan fuera de la vista del público, como el trabajo de tornería de precisión. Pero hay algunas habilidades, y esto es especialmente cierto en lo que hace a las habilidades cognitivas y perceptuales en tanto opuestas a las físicas (motrices), que pasan inadvertidas porque su realización no tiene aspectos públicos. Pensar es el primer ejemplo de ello, especialmente el pensamiento que no lleva al discurso o la acción, como cuando uno ensaya en silencio un argumento o una obra musical.

Otra característica que la determinación del sexo de los pollos comparte con otras habilidades es su dependencia de la práctica intensiva. Aprender a realizar la tarea exige experiencia práctica. Aun cuando se haya logrado una destreza considerable, puede llevar años de práctica alcanzar los niveles de pericia más elevados, especialmente en lo que tiene que ver con la velocidad de realización (capítulo 2).

Finalmente, para nuestros propósitos, la historia de la discriminación del sexo de los pollos ilustra una tercera característica, absolutamente esencial, de la habilidad humana: su dependencia del contexto social. Ya hemos advertido que ciertas habilidades, como caminar y hablar, están tan ampliamente representadas que son casi características de la vida humana misma. Pero en su especificidad, incluso estas habilidades están determinadas socialmente. Un idioma puede convertirse en una lengua muerta si en la gente cesa la necesidad social de hablarlo; una forma de caminar puede ser profundamente emblemática de una sociedad o un régimen particulares. En habilidades cognitivas menos universales, su inserción social es todavía más evidente. Las exigencias perceptuales y cognitivas varían drásticamente entre diferentes hábitats. El rastreo de animales en la selva y la organización del sis-

tema de tránsito de una ciudad son habilidades configuradas por sus respectivos ambientes físicos y sociales. En el caso de la determinación del sexo de los pollos, la habilidad llegó a Occidente sólo como consecuencia del rápido crecimiento del comercio y los lazos culturales con Japón, a comienzos de este siglo. La difusión de la habilidad en Gran Bretaña y los Estados Unidos dependió luego de la existencia de una intensiva industria de aves de corral que había crecido para satisfacer las necesidades de una sociedad urbana cada vez más grande y con buen transporte y redes de comunicación. Finalmente, la suerte cambiante de los expertos en la discriminación del sexo de los pollos fue determinada por las actividades de otro grupo dentro de la misma industria: los criadores. Las habilidades humanas se elaboran en respuesta a las condiciones sociales y pueden caer en desuso y perderse rápidamente cuando esas condiciones cambian.

La determinación del sexo de los pollos es una habilidad de clasificación perceptual relativamente simple; requiere lo que en muchos sentidos son discriminaciones menos complejas que las que debe hacer quien combina diferentes tipos de té o un catador de vinos. Nuestro segundo ejemplo es de una habilidad mucho más compleja, que aquí sólo podemos esbozar brevemente.

## Habilidad para la navegación

Los isleños de Puluwat son famosos por su habilidad para la navegación sin instrumentos. Su hogar se encuentra en el océano Pacífico, a unos 1600 kilómetros al nor-nordeste de Nueva Guinea, y la navegación juega un papel esencial en la cultura Puluwat. Gladwin (1970) describe la importancia que se otorga a la construcción y mantenimiento de canoas con flotadores laterales, las funciones sociales esenciales de los viajes entre islas y el respeto especial acordado a los capitanes de canoas que han logrado destreza en su arte. Todos los habitantes de Puluwat participan en los viajes pero sólo unos pocos son capaces de pilotear una nave en 300 o 500 kilómetros de mar abierto y, más aún, de hacerlo en casi cualquier circunstancia climática y sin estar del todo sobrios. ¿Cómo lo logran?

Hay tres hechos preliminares en la descripción que Gladwin hace de las habilidades de navegación de Puluwat. En primer lugar, se necesita mucho tiempo para convertirse en un navegante experto: veinte o treinta años. En segundo lugar, la habilidad está localizada en el grupo de islas al que pertenece Puluwat, una superficie de unos 650 km de norte a sur y 1600 km de

este a oeste. En tercer lugar, el sistema de navegación es altamente conservador. Gladwin (1970) subraya que los Puluwat no corren riesgos, los navegantes utilizan muchas clases de conocimientos que se refuerzan mutuamente para asegurar la precisión de su calculada navegación.

El conocimiento es, de hecho, el rasgo característico de la navegación Puluwat. Los navegantes son, antes que nada, conocedores de la ubicación de las estrellas en el cielo, los lugares por donde salen y las trayectorias que siguen. Como está cerca del Ecuador, el cielo nocturno sobre Puluwat cambia poco con las diferentes estaciones, y las estrellas, cuando son visibles, ofrecen guías confiables. Pero los navegantes también pueden fijar cursos a seguir durante el día. Al partir de una isla con destino hacia cualquier otra, los Puluwat saben qué hitos, cuando están alineados, les darán el curso que necesitan; esta técnica es conocida como retrovisión. Incluso en la oscuridad y en días nublados, los navegantes pueden guiarse distinguiendo entre tres tipos de olas, que se originan en el este, el norte y el sur, respectivamente. Un navegante pilotea sintiendo estas olas bajo su canoa y por los intervalos característicos entre ellas.

Durante su etapa de aprendizaje, los navegantes también aprenden las posiciones y los contornos de todos los arrecifes importantes en su zona del océano, algunos de los cuales se encuentran a más de 30 metros de profundidad. Aprenden a identificar los efectos que un arrecife tiene sobre el agua y, por lo tanto, sobre los movimientos de la canoa. Allí donde un extraño sólo vería una amplia extensión oceánica, los navegantes de Puluwat ven claras señales que marcan caminos por el mar. También están familiarizados con la conducta de las aves marinas: saben cuáles vuelan directamente hacia tierra cuando termina el día y las usan para dirigirse a destinos a veces muy precisos. Finalmente, también dominan una compleja teoría de navegación que les permite calcular la distancia recorrida cuando avanzan contra el viento. La teoría utiliza islas de referencia, que a menudo están sobre el horizonte, y representa a la canoa como estacionaria mientras el océano y las olas se mueven pasando junto a ella. La teoría es el aspecto más puramente cognitivo (en tanto opuesto a lo perceptual) de la navegación Puluwat.

## ¿Qué es la cognición?

La expresión habilidad "cognitiva" se ha usado varias veces en las páginas anteriores, sin explicación. Pero, ¿qué es la cognición?

Si se reflexiona brevemente sobre algunas palabras cotidianas, se po-

drá tener cierta idea de la respuesta. Hablamos de reconocimiento (en inglés, recognition) cuando se identifica a alguien o algo que ya se conocía. O decimos que una persona viaja "de incógnito" cuando se esfuerza por ocultar su identidad. Estos ejemplos ayudan a aclarar que la palabra "cognición" se refiere a las actividades de conocer, recoger, organizar y utilizar el conocimiento. Cuando un encuentro con alguien nos da una sensación de familiaridad, decimos que es un acto de re-conocimiento (en inglés, re-cognition); conocemos al individuo de nuevo. Pero la cognición incluye muchos tipos de actividad además del reconocimiento de la identidad. Recordar que uno tiene una cita con el dentista al día siguiente, calcular el descuento en un artículo que uno está comprando o encontrar el camino a casa desde el trabajo sin perderse son formas de cognición. También lo es el jugar una partida de ajedrez, tratar de diagnosticar por qué el auto no arranca, comprender lo que otra persona está diciendo o entender qué más quiere decir esa persona además de lo que pone en palabras. En realidad, cualquier cosa que involucre percepción, memoria, aprendizaje o pensamiento es parte de la cognición, lo cual significa que casi cualquier cosa que una persona haga tiene, al menos, un componente cognitivo.

Pero uno no debe sentirse muy desilusionado al enterarse de que "cognición" es un término tan amplio. Primero, se puede dar por sentado que el alcance de los temas de este libro será selectivo. Segundo, aunque el campo de la cognición es amplio, hay un número relativamente pequeño de principios básicos. De hecho, un tema central de este libro es que la cognición en todas sus manifestaciones es una cuestión de habilidad y que, por lo tanto, todos los logros cognitivos se pueden analizar según ciertos principios involucrados en la adquisición de la habilidad (ver capítulos 2 y 3). En tercer lugar, como la cognición prácticamente entra en todo lo que hacemos, hay muchas oportunidades para usar bien esos principios, para mejorar nuestras propias habilidades cognitivas. Por ejemplo, muchas personas suelen decir que, aunque recuerdan muy bien la fisonomía de los rostros, encuentran dificultades para memorizar los nombres que les corresponden. Sin embargo, como veremos, hay ciertas cosas que se pueden hacer para mejorar esta habilidad cognitiva (capítulo 5).

# El estudio de la cognición

## El trasfondo histórico

¿Cómo podemos estudiar la cognición? Una respuesta obvia podría ser que podemos estudiarla de manera directa, observando nuestras propias mentes en funcionamiento mediante la introspección, o de manera indirecta, preguntándole a otros sobre lo que observan qué sucede en de sus mentes cuando trabajan, es decir, recogiendo informes de sus introspecciones. Ciertamente, la introspección es un método importante y muy usado para estudiar la psique. Gran parte de nuestro conocimiento de sentido común sobre la percepción, el pensamiento y la memoria se basa en los informes introspectivos de gente reflexiva, desde los antiguos griegos hasta nuestros días. Además, cuando hacia fines del siglo XIX la psicología se constituyó como ciencia de la psique independiente, separada de la filosofía o de cualquier otra rama de investigación, la introspección fue adoptada como su método natural. Sin embargo, no se suponía que cualquiera podía convertirse en psicólogo sólo haciendo introspección en sus propias cogniciones. Por el contrario, las observaciones de la experiencia consciente tenían que hacerlas individuos muy bien entrenados en técnicas introspectivas desarrolladas especialmente, y debían encarar temas teóricos significativos. Entender el funcionamiento de la mente iba a convertirse en una cosa seria.

Para ser justos con estos psicólogos tempranos, sus ideas sobre la técnica y el entrenamiento no estaban tan motivados por el deseo de marcar un territorio y cercarlo con prácticas restrictivas como por la convicción de que las observaciones introspectivas podían valer la pena sólo dentro del marco de una metodología especialmente construida. Esta convicción halló eco, unos años más tarde, en las técnicas observacionales cuidadosamente elaboradas de la etología, la disciplina que estudia la conducta animal (incluyendo la humana) en sus ámbitos naturales. Al desentrañar el secreto de la danza de las abejas y darle sentido a la conducta de los peces, los gansos y muchas otras especies, los etólogos demostraron cuántas cosas pueden saberse mediante la observación (que un observador sin guía puede perder). Pero en tanto los logros de los etólogos iban a obtener reconocimiento en 1973 mediante la entrega de un premio Nobel, los psicólogos introspectivos hicieron pocos progresos con sus observaciones sistemáticas de la conciencia. (En el capítulo 9 hay brillantes análisis, recientemente realizados, sobre las habilidades del habla, basadas fundamentalmente en observaciones de errores del discurso que ocurren naturalmente.)

## Puntos débiles y fuertes de la introspección

Hay muchos motivos por los que la introspección sola no resulta adecuada como método para revelar la naturaleza de la cognición. Probablemente, el defecto más serio es que hacer introspección en la propia conciencia nos expone a alterar el contenido de esa conciencia. Más que tomar conciencia de la actividad cognitiva corriente, uno toma conciencia del hecho de que está haciendo introspección. El problema se vuelve particularmente agudo cuando hay que traducir cogniciones no verbales a palabras, para un informe introspectivo, ya que quizá no sea posible capturar su naturaleza esencial en términos verbales.

Un segundo problema, relacionado con el anterior, es el que señala Neisser (1976). Destaca que cualquier psicología de importancia debe tener algo que decir sobre lo que de verdad la gente hace y piensa en situaciones reales y culturalmente significativas. Sin embargo, es probable que ésas sean las situaciones en las que la gente está demasiado ocupada como para realizar una introspección cuidadosa. Como resultado, la introspección puede aplicarse sobre todo a tareas artificiales en ámbitos de laboratorio. Neisser culpa fundamentalmente a esta esterilidad —la falta de lo que se suele denominar "validez ecológica"— de la decadencia del método introspectivo. Pero no es difícil detectar otras fallas. Por ejemplo, una observación común (al menos desde que Freud enfatizó especialmente el hecho) es que la conducta suele estar determinada por motivos inconscientes. A veces, un observador perspicaz puede tener mejor idea de los motivos de las acciones de una persona que la propia persona. Aunque la psicología cognitiva no se ocupa mucho de los motivos, tiene un lugar importante para las cogniciones inconscientes. Lo inconsciente no posee, en este sentido, la connotación de represión intencional asociada con la noción freudiana del "inconsciente". Más bien se refiere al simple hecho de que no tenemos conciencia de muchos de los procesos que regulan nuestras funciones físicas y psíquicas. Un ejemplo de esto puede ser la oxigenación de la sangre en nuestro músculo cardíaco o los procesos involucrados en el reconocimiento y la comprensión de las palabras de esta página podría ser otro. Como en el caso de los motivos inconscientes, el mensaje es claro. Lo que sucede fuera de la conciencia no puede revelarse mediante introspección, no importa con cuánto cuidado se realice.

Es evidente que, como método, la introspección presenta muchas deficiencias. Ya hemos enumerado varias, pero también debemos admitir que posee ciertos puntos fuertes. Es definitivamente cierto, por ejemplo, que las

descripciones introspectivas de la actividad cognitiva pueden ser muy útiles e informativas en las circunstancias adecuadas. Algunas personas son capaces de comunicar sus propias técnicas al encarar problemas cognitivos específicos. Los demás nos beneficiamos con la forma en que estas personas razonan sobre las fallas en el encendido o la estrategia adecuada para encarar un problema matemático, los mecanismos que facilitan las interacciones sociales o el mejoramiento de la memoria para retener la información importante y los detalles a los que prestan atención cuando evalúan la fuerza de un caso legal o hacen un diagnóstico basándose en radiografías. Hemos aprendido demasiado acerca del funcionamiento de la mente mediante la introspección para querer omitirlo por completo. Aunque hay un permanente debate sobre la situación precisa de varios tipos de informe introspectivo (Evans, 1982), los psicólogos cognitivos siguen confiando en ellos. Debería también destacarse que todos los "sistemas de conocimiento experto", que representan uno de los mayores logros de la investigación en inteligencia artificial, se basan en las introspecciones de los especialistas humanos. La lección que se debe aprender del fracaso de la psicología introspectiva temprana no es que el método no tiene valor sino que se lo debe ver tan solo como uno entre los muchos métodos posibles de investigación de la cognición.

## El estudio contemporáneo de la cognición

Entre las dos guerras mundiales, el conductismo dominó la psicología en los Estados Unidos y, en menor medida, en Gran Bretaña. No sólo decayó el introspeccionismo sino también todo el estudio de la cognición. Los conductistas destacaban correctamente la importancia de los datos objetivos y especificaban cuidadosamente métodos experimentales. Esto los llevó a promover explicaciones según estímulos y respuestas observables, y a negar la necesidad de referirse a los aspectos cognitivos involucrados. John Watson, padre fundador de la psicología conductista, supuso que pensar no era más que discurso subvocal, como hablarse a sí mismo. Se esperaba que medidas suficientemente delicadas de la actividad de la musculatura del habla confirmarían este supuesto, pero el conductismo no facilitó otros medios para investigar el pensamiento y nada dijo sobre el pensamiento no verbal.

Sin embargo, en la Europa continental surgía un cuadro algo diferente. El introspeccionismo no fue sucedido por el conductismo sino por la psicología de la Gestalt, preocupada especialmente por los procesos cognitivos in-

ternos y por fenómenos como el del *insight* (ver capítulo 14). Además, primero en Francia y luego en Suiza, Jean Piaget inició sus estudios, que resultarían profundamente influyentes sobre el desarrollo cognitivo de los niños. Un tema básico de la teoría de Piaget consistiría en que el pensamiento se desarrolla como resultado de la internalización. La conducta previamente manifiesta, tanto en las acciones como en el discurso, podía internalizarse plenamente, sin dejar manifestaciones externas. De modo que mientras los conductistas buscaban crear una psicología de la conducta observable que resultara metodológicamente convincente, Piaget enfatizaba, al mismo tiempo, la continuidad existente entre la conducta y la cognición.

La psicología cognitiva moderna combina el rigor metodológico de los conductistas con el énfasis piagetiano en los procesos internos que sostienen a la conducta evidente. También incluye a una tercera rama importante.

## La metáfora de la computadora para la mente

A pesar de Piaget, el interés de los psicólogos por los procesos cognitivos no revivió hasta mediados del decenio de 1950, cuando se produjo un explosivo crecimiento de la investigación, que no se ha detenido en la actualidad. Indudablemente, la causa más importante de este cambio fue la rápida evolución y difusión de la computadora digital. Una máquina que desempeña operaciones complejas sin la ayuda de partes móviles recuerda inmediatamente los procesos internos de la mente. En la entrada *(input)* a la máquina se producen operaciones no observables, que se pueden especificar precisamente y tienen efectos demostrables en la conducta de los dispositivos de salida *(output)*. Para entender de qué es capaz la computadora, no se puede depender sólo del conocimiento de lo que sucede dentro y lo que sale de ella. Es necesario saber qué le está haciendo hacer el programa a la máquina, qué transformaciones y manipulaciones de la entrada codificada se están realizando, qué objetivos se supone que debe alcanzar el programa (capítulo 16). Pero si tal comprensión es posible en el caso de una computadora, ¿por qué no también en la mente humana? Dada la computadora como modelo, repentinamente los procesos cognitivos de la percepción, el pensamiento y la memoria parecían más reales, menos etéreos. Así como las máquinas de partes móviles han ofrecido con frecuencia metáforas fértiles para las funciones corporales (por ejemplo, el corazón como una bomba), el nuevo tipo de máquina sugiere una metáfora para la mente.

Durante los últimos veinticinco años, los psicólogos cognitivos han explorado ampliamente la metáfora de la computadora para la mente, y los libros contemporáneos sobre cognición suelen escribirse en lenguaje informático. Esto hace que su lectura resulte, con frecuencia, ardua. También puede dar la impresión de que los psicólogos cognitivos están más preocupados por la naturaleza de la computadora que por la naturaleza de la cognición humana. Sólo los aspectos de la cognición que se pueden discutir plausiblemente con terminología de programación reciben atención; los puntos desemejantes entre las computadoras y la mente son por lo tanto borrados de la conciencia del lector (y de la del autor). Por este motivo, en este libro se intenta cuidadosamente evitar la metáfora y su jerga asociada. Esto no resulta fácil ya que la metáfora de la computadora ha invadido gran parte de nuestro pensamiento, especialmente en la psicología cognitiva. De vez en cuando recurriremos a la metáfora, pero sólo la retomaremos y analizaremos explícitamente en el último capítulo.

## Cognición y conducta

Si no vamos a valernos del lenguaje y las analogías de la computación, ¿cómo pensamos hablar de la cognición? ¿Cómo pueden describirse los procesos psíquicos no observables? La respuesta ya se ha sugerido antes, cuando nos referimos al énfasis piagetiano en la continuidad entre conducta y pensamiento. Muchas cogniciones pueden considerarse acciones internas. Son acciones que se han vuelto tan practicadas que podemos "llevarlas a cabo en nuestras cabezas" (por ejemplo, contar con los dedos deja paso luego a la aritmética mental); o, si se construyen sobre conocimientos y destrezas previas, podemos haber aprendido a hacer cálculos mentalmente desde el primer momento. Las metáforas de acción para el pensamiento, y más particularmente las metáforas manipulativas, que se refieren al uso de las manos, son comunes en el lenguaje cotidiano. Cuando hablamos de "aferrarse a una idea" o acusamos a alguien de "hacer una excepción", entra en juego implícitamente la noción de acción internalizada. Sin embargo, no todas las cogniciones son afines a las acciones. El lenguaje cotidiano tiene un segundo grupo importante de metáforas para describir las cogniciones, basándose en la percepción, especialmente en la de tipo visual. Las bromas como "Ya veo, dijo un ciego" se valen de estas metáforas, del mismo modo que expresiones como "Esa es una idea brillante". La prevalencia de los dos tipos de metáfo-

ra no es coincidente. Al habérselas con el mundo, un organismo tiene que percibir lo que está sucediendo a su alrededor y realizar las acciones adecuadas, como alimentarse, copular o huir. En las criaturas simples, tanto las percepciones como las respuestas tienden a ser incorporadas, y las percepciones activan automáticamente las respuestas. Pero en los animales con sistemas nerviosos más complejos aparece un elemento de elección, ya que en la mayoría de las circunstancias hay más de un curso de acción posible. La actividad nerviosa que interviene entre la percepción y la respuesta se torna elaborada, evoluciona en lo que se denomina cognición. Mediante los dos tipos de metáfora, el lenguaje común reconoce muy sensatamente que la cognición comparte características de la percepción y la acción. Las metáforas visuales y manipulativas predominan porque para los humanos los ojos son el canal principal de información perceptual y las manos ofrecen el medio más flexible y delicado para actuar sobre el mundo.

## Cognición y habilidad

Seguir la idea de Piaget sobre la relación existente entre cognición y conducta no implica necesariamente, por lo tanto, abrir nuevos caminos. Hace mucho que el lenguaje común ha reconocido la relación. Sin embargo, hay una característica aún más importante que debemos destacar aquí. El psicólogo Bartlett (1958) señaló que el pensamiento —o, más generalmente, la cognición— es en muchos sentidos similar a la conducta entrenada en el uso de habilidades. Las habilidades perceptivo-motrices, como jugar al tenis, involucran la percepción e interpretación de alguna parte del mundo (por ejemplo, la trayectoria, la velocidad o el giro de la pelota de tenis) y la realización de alguna acción motriz (por ejemplo, pegarle a la pelota hacia algún punto conveniente en el campo del adversario). Tanto la percepción como la respuesta requieren un grado de habilidad mayor. De manera similar, muchas actividades cognitivas que consideramos naturales son desempeños expertos. Aunque nuestras cogniciones pueden parecer, a veces, bastante absurdas e inconexas, en su mayoría son sumamente controladas, y como adultos casi todos hemos logrado la percepción, el recuerdo, el pensamiento y la comprensión de unos con otros. Estos logros cognitivos serán el tema central del presente libro.

De hecho, no puede trazarse un límite firme entre las habilidades perceptivo-motrices y las cognitivas. Casi todas las conductas motrices están bajo control cognitivo y muchas cogniciones sirven para actuar en el mundo.

Esta interacción queda bien clara en un informe clásico de Bryan y Harter (1899) sobre el aprendizaje de la telegrafía Morse, que fue el primer análisis completo del desempeño experto. La telegrafía involucra el aprendizaje de nuevas habilidades percepto-cognitivas para la identificación e interpretación de esquemas de puntos y rayas y la adquisición de nuevas habilidades cognitivo-motrices para manipular la llave de transmisión. Junto con Bartlett, sus sucesores, Bryan y Harter, construyeron los cimientos de una psicología de la habilidad cognitiva.

## Conocimiento psicológico

### Teoría y práctica

Francis Bacon, pensador del siglo XVI, redactó un plan para un nuevo método de obtención de conocimientos en lo que habrían de ser las ciencias naturales. Bacon se ocupó tanto del conocimiento teórico como del práctico. Como parte de su método, sostuvo que la comprensión teórica clara de un tema debe establecerse primero y que luego pueden derivarse de ella las aplicaciones prácticas. En otras palabras, la ciencia llevaría a la tecnología o, en los términos del propio Bacon, después de la "luz" vendrían los "frutos". Aunque la relación entre teoría y práctica no es tan simple como sugiere Bacon, el indudable éxito de las tecnologías basadas en la ciencia en los últimos trescientos años ha afectado notoriamente nuestras ideas sobre el conocimiento. No es exagerado decir que en la actualidad la validez del conocimiento teórico se evalúa sobre la base de las tecnologías que genera. Raramente se considera que las teorías sin aplicaciones instrumentales constituyan conocimiento.

Una consecuencia de este desarrollo es que los psicólogos cognitivos, como los investigadores de otras disciplinas, están bajo considerable presión para demostrar las aplicaciones prácticas de su trabajo. El público en general, los estudiantes y los administradores de fondos para la investigación tienden a no impresionarse por los atractivos teóricos. Quieren saber para qué sirve la psicología cognitiva. Tampoco es externa toda la presión. Los propios psicólogos se estimulan entre sí, constantemente, a hacer que su disciplina sea más significativa y utilitaria. En lo que queda de este capítulo, examinaremos dos formas posibles de responder a dicha presión para obtener resultados prácticos.

## El psicólogo cognitivo como científico

Los psicólogos pueden responder a la presión de dos formas diversas, que suponen la adopción de roles alternativos. Estos roles no necesitan ser excluyentes. Los psicólogos pueden jugar un rol y luego el otro de manera secuencial o, si son inteligentes, ambos a la vez. Los dos roles son: el de científico y el de entrenador.

Una respuesta que los psicólogos dan a veces a la demanda de resultados prácticos de la psicología cognitiva es la de representarse a sí mismos como aún en la primera de las dos etapas planteadas por Bacon; es decir, que todavía están estableciendo una comprensión teórica y necesitan un poco más de tiempo para la investigación pura. El psicólogo-científico es presentado como un especialista cuyo conocimiento pronto lo pondrá en una posición privilegiada en lo que respecta a la cognición. Pero éste no es un cuadro de situación realista. Aunque la cognición debe ser estudiada científicamente, incluso teniendo en cuenta los avances posibles en inteligencia artificial, no hay indicios de que la investigación científica esté a punto de llegar a conmovedoras innovaciones en los usos de nuestras mentes o de que alguna vez logre hacerlo (ver, por ejemplo, Joynson, 1974). La psicología sencillamente no se presta a un modelo baconiano de teoría seguida de práctica. A diferencia, por ejemplo, de la física nuclear (en la que las aplicaciones fueron imposibles hasta que se establecieron teorías precisas), la psicología se ocupa de un asunto en el cual la práctica precedió en mucho a las teorías explícitas. Los expertos, ya sean físicos nucleares o técnicos en reparación de televisores, construyen y reparan cosas que el resto de nosotros no pretendemos entender. Los psicólogos, en cambio, no se sitúan en una relación de expertos con respecto a los demás. Las personas que consultan a un psicólogo ya saben mucho sobre sus propias psiques. La situación es muy diferente de cuando esa misma persona necesita que se le repare el televisor. Es cierto que a veces la gente busca que psicólogos y psiquiatras "arreglen" su psique, pero ésa es una actitud desafortunada de su parte. Los tratamientos por drogadicción a veces pueden servir para controlar el pensamiento psicótico o las perturbaciones del carácter, pero no son curaciones. Los especialistas en salud mental no pueden hacer más que ayudar a las personas a ordenar sus mentes y sus conductas para sí mismas.

Esto no significa que la ayuda de ese tipo no sea importante. En muchos casos, puede ser vital. Lo que queremos decir es que el valor de estudiar la cognición científicamente no ha de hallarse en las nuevas tecnologías. An-

tes bien, su importancia radica en el conocimiento disponible para el psicólogo que actúa con un carácter distinto del de experto científico.

## El psicólogo cognitivo como entrenador

Cuando se trata de la aplicación del conocimiento, el papel de entrenador o maestro ofrece una alternativa al rol de experto tecnológico. Ni los entrenadores atléticos ni los maestros de piano, por ejemplo, tienen teorías formales de las que deducir su práctica, pero su capacidad para contribuir a las mejoras en el rendimiento no está en duda. Este libro sostiene que los psicólogos cognitivos tienen una capacidad similar para ayudar a otros a mejorar sus rendimientos cognitivos. No a la manera del experto sino a la del entrenador.

El papel de entrenador atlético proporciona, de hecho, una analogía útil que demuestra cómo el conocimiento científico puede tener un valor práctico que no es tecnológico en ningún sentido manifiesto. Los entrenadores atléticos se vuelcan cada vez más hacia la medicina del deporte en busca de información; toman cosas de la psicología, la anatomía y la bioquímica y aplican este conocimiento científico a los problemas relacionados con el mejoramiento del rendimiento deportivo. Estas ciencias no nos permiten usos novedosos de nuestros cuerpos pero ayudan a mejorar la eficiencia con la que se ejecutan los viejos usos (correr, nadar, saltar, lanzar). De manera similar, el estudio científico de la cognición jamás va a revelar habilidades cognitivas insospechadas, pero nos proporciona información que puede servir para mejorar la habilidad con la que se ejercitan aptitudes que hace mucho nos resultan familiares.

Para los psicólogos cognitivos, hay muchas ventajas en presentarse a sí mismos como entrenadores. Primero y principalmente, la adopción del papel de entrenador lo pondrá al psicólogo, junto con muchos de sus colegas del asesoramiento y la psicología clínica en el lugar de "socios" de un "cliente". El psicólogo puede tratar directamente con la gente de una manera no autoritaria sino constructiva, proporcionando esencialmente una forma de realimentación sobre la manera en que el cliente está desempeñando la tarea en cuestión. Una segunda ventaja es que el papel de entrenador clarifica dónde radica la responsabilidad en lo que hace al logro de mejorías cognitivas. Como en el caso del tenis, un entrenador sólo puede ayudar a aquellos que están dispuestos a esforzarse por sí mismos. La habilidad no puede inyectarse desde afuera. La tercera ventaja de jugar el papel de entrenador es precisamente que les ahorra a los psicólogos cognitivos la etiqueta de exper-

tes bien, su importancia radica en el conocimiento disponible para el psicólogo que actúa con un carácter distinto del de experto científico.

## El psicólogo cognitivo como entrenador

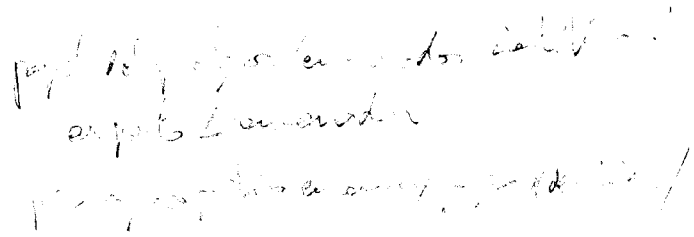
Cuando se trata de la aplicación del conocimiento, el papel de entrenador o maestro ofrece una alternativa al rol de experto tecnológico. Ni los entrenadores atléticos ni los maestros de piano, por ejemplo, tienen teorías formales de las que deducir su práctica, pero su capacidad para contribuir a las mejoras en el rendimiento no está en duda. Este libro sostiene que los psicólogos cognitivos tienen una capacidad similar para ayudar a otros a mejorar sus rendimientos cognitivos. No a la manera del experto sino a la del entrenador.

El papel de entrenador atlético proporciona, de hecho, una analogía útil que demuestra cómo el conocimiento científico puede tener un valor práctico que no es tecnológico en ningún sentido manifiesto. Los entrenadores atléticos se vuelcan cada vez más hacia la medicina del deporte en busca de información; toman cosas de la psicología, la anatomía y la bioquímica y aplican este conocimiento científico a los problemas relacionados con el mejoramiento del rendimiento deportivo. Estas ciencias no nos permiten usos novedosos de nuestros cuerpos pero ayudan a mejorar la eficiencia con la que se ejecutan los viejos usos (correr, nadar, saltar, lanzar). De manera similar, el estudio científico de la cognición jamás va a revelar habilidades cognitivas insospechadas, pero nos proporciona información que puede servir para mejorar la habilidad con la que se ejercitan aptitudes que hace mucho nos resultan familiares.

Para los psicólogos cognitivos, hay muchas ventajas en presentarse a sí mismos como entrenadores. Primero y principalmente, la adopción del papel de entrenador lo pondrá al psicólogo, junto con muchos de sus colegas del asesoramiento y la psicología clínica en el lugar de "socios" de un "cliente". El psicólogo puede tratar directamente con la gente de una manera no autoritaria sino constructiva, proporcionando esencialmente una forma de realimentación sobre la manera en que el cliente está desempeñando la tarea en cuestión. Una segunda ventaja es que el papel de entrenador clarifica dónde radica la responsabilidad en lo que hace al logro de mejorías cognitivas. Como en el caso del tenis, un entrenador sólo puede ayudar a aquellos que están dispuestos a esforzarse por sí mismos. La habilidad no puede inyectarse desde afuera. La tercera ventaja de jugar el papel de entrenador es precisamente que les ahorra a los psicólogos cognitivos la etiqueta de exper-

tos. Ya no necesitan temer las burlas familiares del tipo de: "Así que estudias el pensamiento/la memoria. ¿Y por qué tu pensamiento/memoria no es mucho mejor que el de los demás?" Después de todo, nadie pretende que los mejores entrenadores de tenis hayan sido campeones ni que los campeones de tenis vayan a ser buenos entrenadores.

En el resto de este libro se demostrarán los roles gemelos del psicólogo cognitivo. En la investigación, veremos al psicólogo trabajando como científico, proponiendo y verificando hipótesis, realizando experimentaciones, analizando datos y haciendo observaciones de los estudios de caso. Cuando se trate de aconsejar sobre el modo de mejorar los rendimientos cognitivos, veremos al psicólogo actuando como entrenador: aconsejando, guiando, dirigiendo, utilizando máximas, reglas empíricas y unos cuantos principios generales de la habilidad humana.



## Resumen

Este capítulo bosqueja varios asuntos y temas que se elaborarán en el resto del libro. La noción de *habilidad cognitiva* se introduce mediante dos ejemplos: la determinación del sexo de los pollos y la navegación sin instrumentos. Se señala que las habilidades cognitivas pueden parecer casi milagrosas a los no iniciados, que su adquisición requiere práctica intensiva y que estas habilidades están profundamente insertas en su contexto social. Se comenta brevemente de qué manera los psicólogos deben estudiar y estudian la cognición y se compara la cognición con la percepción y la conducta motriz. Se prefiere analizar las cogniciones como un modelo de habilidad perceptivo-motriz antes que hacerlo según la metáfora de la computadora para la mente, aunque se reconocen la influencia y el valor de esa metáfora.

Por fin, se considera brevemente la naturaleza del conocimiento psicológico. Se sostiene que el psicólogo cognitivo debe jugar dos roles: el del científico, cuando está involucrado en la investigación básica, y el de entrenador, cuando aplica su conocimiento a los problemas humanos.

## Lecturas adicionales

NEISSER, U. (1976) *Cognition and Reality*. San Francisco: Freeman.

Proporciona una vívida introducción al tema de la psicología cognitiva desde la perspectiva de las habilidades, enfatizando fuertemente las habilidades perceptuales.

# 2

# ¿Qué es la habilidad?

*John Sloboda*

En mi casa hay dos habitaciones recientemente pintadas: una la pintó un pintor profesional; la otra, la pinté yo. El pintor es un profesional experto; yo, no. ¿Qué detalles ayudan a detectar la diferencia? Bueno, el producto final ciertamente algo dice al respecto. Si uno mira cuidadosamente las paredes que yo pinté, encontrará una textura despareja; demasiada pintura por aquí, muy poca por allá. Encontrará también leves rayas verticales donde la pintura se ha corrido y desprolijidades en los bordes. Las paredes pintadas por el profesional tienen una textura pareja y los bordes son maravillosamente rectos y prolijos.

No obstante estas diferencias, en realidad me las arreglé para cubrir mis huellas bastante bien. A menos que uno observe detalladamente mis paredes, no encontrará ningún defecto evidente. Pero las diferencias más drásticas entre el profesional y yo habrían sido muy evidentes si se nos hubiera visto mientras trabajábamos.

Para empezar, el profesional terminó su trabajo en la mitad del tiempo que yo necesité. No sólo sus pinceladas eran más rápidas sino que se detenía menos que yo, que necesitaba parar con bastante frecuencia, no por holgazanería sino para poder evaluar lo que había hecho y determinar qué debía hacer a continuación. Tuve también otras demoras. Por ejemplo, me quedé sin pintura cuando aún me faltaba pintar la cuarta parte de la habitación y tuve que ir a comprar más.

En segundo lugar, el profesional hacía que su trabajo pareciera fácil. La pintura fluía sobre la pared en una sucesión serena y coordinada de movimientos. Mi secuencia de movimientos era forzada y espasmódica. Por ejemplo, una pincelada tenía demasiada pintura; la siguiente, muy poca; a veces empezaba con una pincelada ascendente y luego seguía con otra descendente; constantemente me detenía en minucias como derrames y goteos.

Finalmente, la secuencia del profesional era impecable. Ordenaba las cosas de tal manera que no necesitaba mover todo constantemente. Siempre terminaba en el lugar correcto en el momento preciso (por ejemplo, comen-

zaba la segunda capa de pintura justo cuando la primera se había secado lo necesario) y encaraba cada pared de una manera estrictamente sistemática. Yo trabajaba al azar, trataba cada pared de diferente manera, dejaba sectores sin pintar cuando, por ejemplo, la escalera no estaba colocada en el lugar adecuado; constantemente tenía que dejar de pintar para buscar algo que necesitaba o quitar obstáculos del camino.

## Las características de la habilidad

Durante mucho tiempo, los psicólogos han tratado de especificar con precisión las características de las actividades que requieren habilidades. Mi ejemplo sobre la pintura incluye de manera informal muchas de las características principales, pero me gustaría ocuparme de ellas de una forma más sistemática, bajo cinco encabezamientos principales: Fluidez, Rapidez, Automaticidad, Simultaneidad y Conocimiento. Para recordarlas, puede ser útil la regla mnemotécnica de la primera letra de cada palabra: FRASC (por reglas mnemotécnicas, ver capítulo 5).

### Fluidez

Una actividad es fluida si sus componentes avanzan juntos en una secuencia integrada e ininterrumpida. La palabra "fluidez" suele aplicarse a la microestructura de una tarea (los elementos que se dan en un lapso de unos pocos segundos) más que a la estructura de largo plazo. Por lo tanto, decimos que un traductor es fluido cuando puede realizar una traducción adecuada con un mínimo de pausas o vacilaciones. Un mecanógrafo fluido es el que puede mantener un ritmo de golpes parejo y continuo sobre el teclado.

Parece probable que la fluidez se origine en dos causas. Una es la superposición en el tiempo de una secuencia de movimientos; es decir, los movimientos preparatorios para la acción B se inician mientras se está desarrollando la acción A. La otra causa es la construcción de un conjunto de acciones al modo de un "agrupamiento" simple, que se puede controlar y ejecutar como si fuera sólo una unidad de conducta.

Shaffer (1976) ha demostrado elegantemente que los mecanógrafos expertos hacen estos agrupamientos. En su investigación, los mecanógrafos tenían frente a sí una pantalla de computadora en la que se veía una única línea de texto. Esta pantalla estaba conectada a un teclado de manera tal que cada

vez que los mecanógrafos presionaban una tecla, el texto se movía un espacio hacia la izquierda. Esto significaba que el carácter (o espacio) más a la izquierda "desaparecía" de la pantalla y un nuevo carácter aparecía a la derecha.

Shaffer podía variar tanto el tamaño de la ventana (es decir, la longitud de la línea exhibida) como la cantidad de visión previa (es decir, a qué distancia del extremo derecho de la pantalla se había movido un carácter antes de que fuera necesario escribirlo). Si no había visión previa, el mecanógrafo debía escribir un carácter ni bien aparecía a la derecha de la pantalla. Si había, tipeaba un carácter cuando llegaba a una posición prefijada hacia la mitad de la pantalla. Los mecanógrafos recibieron tres diferentes tipos de texto:

a) prosa normal;
b) prosa mezclada, en la que las palabras aparecían en orden azaroso.
c) palabras mezcladas, donde el orden de las letras era azaroso.

En la Figura 2.1 se ilustran estos tres tipos de texto.

*Figura 2.1. Tipos de texto usados en el estudio de Shaffer sobre mecanografía*

La línea vertical indica la posición de la letra que debe mecanografiarse, con una visión previa de 8 caracteres:

a) en Córdoba. Le dije a Juan sobre la reunión y | también

b) reunión dije le. Juan a también Córdoba en so|bre la y

c) unióren djie el. Nuaj a tmbinéa ódroCba ne br|ose al y

Shaffer halló que los mecanógrafos necesitaban ver unos 8 caracteres previos para poder alcanzar su máxima velocidad. Ver más de esos 8 caracteres no daba como resultado incremento alguno de la velocidad o la precisión. Las máximas velocidades uniformes fueron de unos 10 caracteres por segundo (100 palabras por minuto) en los ejemplos del tipo (a) y (b) pero sólo de 2 caracteres por segundo en el caso (c). Aparentemente, los mecanógrafos expertos pueden encarar las palabras familiares mucho mejor que las cadenas de letras sin sentido. Podemos lograr una mejor comprensión de por qué se tipean más rápido las palabras si observamos lo que sucedió cuando Shaffer redujo la visión previa a menos de 8 caracteres. En el caso (c), la re-

ducción hasta 0 no tuvo efecto alguno en la velocidad, que se mantuvo a un ritmo de 2 caracteres por segundo. Pero esa reducción tuvo un efecto drástico en los otros dos casos, que cayeron directamente, también, a 2 caracteres por segundo.

Aparentemente, la fluidez en el mecanografiado depende de que el mecanógrafo pueda ver la totalidad de una palabra promedio[1] antes de escribirla. Recién entonces puede ensamblarse en una única unidad de realización todo el conjunto de movimientos de los dedos necesario para escribir la palabra. Semejante unidad puede expresarse con rapidez. Cuando no hay visión previa, o cuando las letras no forman palabras familiares, el agrupamiento no es posible y el trabajo debe hacerse letra por letra.

La falta de diferencia entre los casos (a) y (b) también es interesante. Demuestra que estos mecanógrafos no podían formar agrupamientos mayores que las palabras individuales (a diferencia de los hablantes; ver capítulo 9). Si lo hubieran podido hacer, lo lógico habría sido que en la situación (a) fueran más veloces que en la (b). Aparentemente, en general, cualquier par dado de palabras adyacentes aparece con muy poca frecuencia como para agruparlo tal cual se hace con las palabras individualmente. Sin embargo, es probable que secuencias de palabras como "estimado señor" o "saludo atentamente" logren la condición de "agrupamientos" para algunas secretarias.

En el mecanografiado fluido se producen dos niveles de agrupamiento. Uno de ellos es lo que podría denominarse agrupamiento de *entrada* (el acto perceptual de agrupar letras en unidades tales como las palabras); el otro sería el agrupamiento de salida (el ensamble de secuencias de movimientos coordinados). Quizá la mejor manera de apreciar la naturaleza de la fluidez de esta salida sea experimentar su pérdida. Una forma efectiva de interrumpir la fluidez es perturbar la realimentación sensorial que normalmente recibimos de la conducta motriz. Esto interfiere con la capacidad de ensamblar entre sí los movimientos separados. Por ejemplo, supongamos que hablamos por un micrófono que alimenta unos auriculares que, al mismo tiempo, estamos usando. En esta situación es posible introducir una demora, de manera tal que en vez de oír la propia voz mientras uno está hablando, la oiga un poco después. El retraso de la realimentación auditiva de la propia voz en 0,2 segundos suele provocar graves fracturas del discurso. Lleva al tar-

---

*1. El experimento se realizó en inglés, lengua en la que las "palabras promedio" tienen unas 6 letras. (N. del T.)*

tamudeo, la repetición de sonidos y las pausas excesivas entre palabras y sílabas. La conducta fluida depende crucialmente de que la realimentación normal llegue en tiempo normal. En lo cotidiano, la pérdida de la fluidez se experimenta cuando estamos nerviosos. Tropezamos y tambaleamos con las palabras. La fluidez suele ser el último rasgo que se adquiere en una habilidad y su pérdida es el primer signo de fractura, debida a enfermedad, intoxicación o miedo.

## Rapidez

La mayoría de las habilidades incluyen la capacidad de ofrecer una respuesta adecuada rápidamente. El jugador de tenis experto no sólo debe llegar al lugar correcto y elegir el golpe adecuado sino que debe hacerlo en un lapso increíblemente corto. La capacidad de ofrecer la respuesta correcta casi inmediatamente es característica en todas las habilidades.

Uno de los estudios más importante y más citados, que demuestra la velocidad del desempeño experto, es un estudio sobre jugadores de ajedrez, realizado por Chase y Simon (1973). Tomaron personas con diferentes niveles de dominio del juego (desde principiantes hasta grandes maestros) y les mostraron tableros de ajedrez en los que se reproducía una partida real, ya empezada, jugada entre expertos. Se les permitió ver el tablero durante 5 segundos. Luego se retiraron las piezas y se les pidió que reconstruyeran las posiciones sobre el tablero vacío. Como promedio, los principiantes pudieron volver a poner 4 de cada 20 piezas correctamente. Los maestros, en cambio, pusieron 18 de cada 20 piezas.

Como todos vieron el tablero durante el mismo lapso, los resultados demuestran que un maestro puede aprovechar la misma información mucho más rápidamente que un principiante. Una segunda parte de la investigación arrojó luz sobre los mecanismos involucrados. Los tableros que se mostraron esta vez a los individuos contenían piezas colocadas al azar, de una manera que jamás podría haberse dado en una partida lógica. Enfrentados a semejantes tableros, los principiantes tuvieron el mismo rendimiento que en el primer experimento, y colocaron correctamente 4 de 20 trebejos. Los maestros, por su parte, mostraron una drástica caída en su rendimiento, de 18 a 4 piezas, de cada 20. En estos tableros azarosos el maestro no fue superior al principiante.

Esto nos muestra que la habilidad del maestro no depende simplemente de una percepción superior o de la memoria sino que está vinculada con la detección en el estímulo de esquemas familiares y relevantes para el juego.

El maestro "entiende" inmediatamente lo que ve en términos significativos, de manera que, por ejemplo, un grupo de 5 piezas en determinada posición se convierte para él en un "enroque", otro grupo en un "jaque doble", etc. Esto se confirma por la forma en que el maestro realiza la tarea de reconstrucción. Tiende a colocar conjuntamente los grupos asociados y hace pausas más largas entre los diferentes grupos. La capacidad del jugador experto para percibir esquemas significativos sobre el tablero es análoga a la capacidad de un espectador experto de fútbol para reconocer, en medio de la frenética actividad de un partido, movimientos cuidadosamente ensayados.

## Automaticidad

Una de las características más universales de la habilidad es la forma en que se vuelve "fácil" para quienes la practican. Ya no experimentamos esfuerzo alguno cuando llevamos a cabo alguna habilidad bien aprendida (como caminar). Sólo "sucede", sin que tengamos que pensar en ella. Si no fuera así, nuestra capacidad para actuar en el mundo se vería drásticamente limitada. Pasaríamos todo el tiempo ocupándonos de las cosas más elementales.

La *repentización*[2] en piano es una habilidad compleja. En un estudio sobre profesionales de la repentización, Wolf (1976) entrevistó a varios ejecutantes y les preguntó cuáles eran los principales problemas de la repentización. Un pianista contestó: "Para mí, personalmente, ninguno". Esto no es arrogancia sino una respuesta honesta sobre una habilidad altamente automatizada. Es el tipo de respuesta que la mayoría de nosotros tendría que dar si se nos preguntara sobre los problemas que tenemos para caminar. Sin embargo, caminar es una habilidad de una complejidad comparable a la de la repentización en piano, que exige a los niños un largo tiempo de aprendizaje.

Una de las formas de comprobar si una habilidad está automatizada es ver si el ejecutante puede resolver adecuadamente una situación, aun cuando no esté concentrado o no espere que ésta se presente. Así, por ejemplo, un conductor experto sería capaz de frenar rápidamente ante un riesgo potencial, aun cuando su mente hubiera estado en otra cosa en los momentos previos.

La Berge (1973) ofreció una clara demostración experimental de este tipo de fenómenos. Los sujetos debían mirar un símbolo. Después de un breve lapso, el simbolo desaparecía y otro lo reemplazaba. Los dos símbolos podían ser letras familiares (como a, b, p, d) o caracteres no familiares (como 1↓).

2. *Ejecución de una obra musical a la primera lectura.*

A uno de los sujetos se le pidió que juzgara si ambos símbolos eran iguales o diferentes y que apretara un botón (de dos posibles) para indicar su decisión tan rápidamente como pudiera, después de la presentación del segundo símbolo. Se midió el tiempo transcurrido entre la presentación del segundo símbolo y la respuesta. Este tiempo suele denominarse tiempo de reacción y es una medida usada con frecuencia en el estudio de todo tipo de habilidades. La Figura 2.2 resume los resultados de La Berge. En esta tarea, no se encontró diferencia en el tiempo de reacción entre los símbolos familiares y los no familiares. La Berge sostuvo que el primer símbolo ofrecía una pista al sujeto sobre el tipo de símbolo que podía esperar en la segunda presentación, facilitando la decisión sobre si ambos eran de la misma forma, sin importar su familiaridad.

Pero, en algunas pruebas, La Berge introdujo algo inesperado. En lugar del segundo símbolo, presentó un par de símbolos. En los casos en que hizo esto, se pidió a los sujetos que ignoraran el primer símbolo y se concentraran en el par de símbolos. Debían juzgar si los miembros de este par eran idénticos o no. En esta tarea inesperada, La Berge halló una diferencia significativa entre el tiempo de reacción entre pares familiares (por ejemplo, bd) y no familiares (por ejemplo, ⅂ ⅃). Los sujetos respondieron más rápidamente a las letras familiares. La Berge sostiene que este resultado muestra que, en los lectores experimentados, el reconocimiento experto de letras es automático mientras que el reconocimiento de símbolos nuevos y no familiares no lo es. Cuando se presenta una tarea inesperada, sólo se pueden procesar rápidamente las letras. Los símbolos no familiares se asocian con respuestas más lentas precisamente porque no se los puede reconocer automáticamente.

*Figura 2.2: Tareas y resultados de La Berge*

| | tarea esperada | | tarea inesperada | |
|---|---|---|---|---|
| | 1. | 2. | 3. | 4. |
| primer estímulo | a | ⅂ | a | a |
| segundo estímulo | a | ⅂ | bb | ⅂⅂ |
| velocidad de reacción | rápida | rápida | rápida | lenta |

Otra característica de las habilidades automáticas parece ser que, en algunos casos, son obligatorias. Es decir, un estímulo desencadena su respuesta automática sin importar si la deseamos. Cuando vemos escrita una palabra familiar, no podemos evitar el pensar en su significado. También nos re-

sulta igualmente imposible no reconocer un rostro familiar, como el de una persona que conocemos, y verlo en su lugar como un mero conjunto de rasgos individuales.

La tendencia de las habilidades automáticas a entrar en juego "a pesar de nosotros mismos" puede llevar a veces a situaciones embarazosas y graciosas, que han sido estudiadas en el contexto de la "distracción" (por ejemplo, Reason y Mycielska, 1982). La madre de una familia numerosa cuenta que una vez, en una fiesta, se encontró cortando en pedacitos la comida de su sorprendido vecino de mesa, mientras charlaba con otra persona. Si eso nos causa gracia, deberíamos reirnos con la desafortunada madre (y no de ella) ya que los estudios han demostrado que la mayoría de la gente puede recordar muchos deslices de este tipo en su vida cotidiana. Los resultados no siempre son tan embarazosos, pero el principio es el mismo. Los deslices de esta naturaleza parecen consecuencias inevitables de la automatización de la conducta experta.

## Simultaneidad

Todavía recuerdo con horror mi primera lección para aprender a conducir. Mi instructor intentaba enseñarme cómo cambiar de velocidad. Esto incluía una complicada secuencia de movimientos que involucraban al embrague, el acelerador y la palanca de cambios simultáneamente. Por si esto no era suficiente, mi instructor no dejaba de gritar: "¡Mantén la vista en el camino! ¡No sueltes el volante!". Las múltiples demandas a mi atención parecían imposibles de satisfacer. Ahora, cuando ya han pasado unos 20 años, todo parece trivialmente sencillo. Cambiar de marcha es una secuencia de movimientos fluida y coordinada que puedo realizar mientras mantengo una discusión de alto nivel intelectual con mi pasajero sobre algún tema que nada tiene que ver.

La simultaneidad es una característica de la habilidad en dos sentidos. En primer lugar, los componentes de una actividad experta pueden ejecutarse simultáneamente (como el ejemplo de los movimientos conjuntos de manos y pies para cambiar la marcha del auto). En segundo lugar, debido al alto grado de automaticidad, suele ser posible desarrollar una actividad no relacionada al mismo tiempo que se realiza una actividad experta. Una de las formas que usan los psicólogos para comprobar cuánto se ha automatizado una habilidad consiste en medir el efecto sobre el rendimiento del agregado de una segunda tarea, simultánea, para que el sujeto la realice.

Allport, Antonis y Reynolds (1972) ofrecieron una de las demostraciones antiintuitivas más sorprendentes, en un experimento en el que analizaban el rendimiento en dos habilidades: la repentización en piano y la "persecución" en prosa. Los sujetos del experimento eran repentizadores experimentados. Primero, se les pidió que ejecutaran cada tarea por separado. En la prueba de repentización, se colocó frente al sujeto una partitura para piano no familiar, y se le pidió que la ejecutara sin ensayo previo. En la prueba de "persecución", el sujeto escuchaba un fragmento en prosa a través de auriculares y debía repetirlo en voz alta mientras lo oía. Luego, se combinaron las dos tareas. Se le pedía al sujeto que realizara ambas tareas a la vez lo mejor posible.

Después de muy poca práctica, los sujetos podían realizar las dos tareas a la vez sin fallar en ninguna. Aun más sorprendente fue la comprobación de que ninguna de las dos tareas afectaba seriamente a la otra. Cada tarea fue realizada en la situación dual con la misma destreza que en la realización por separado.

Un largo debate se abrió en la psicología sobre si se puede prestar atención a dos cosas a la vez. Algunos teóricos sugirieron que la atención es como un rayo indivisible: sólo se puede apuntar hacia una cosa por vez. La realización de dos tareas debe explicarse, entonces, como el rápido cambio de la atención de una tarea a la otra (Broadbent, 1982). Otros lo han visto más como una cantidad de recursos que se pueden dedicar íntegramente a una tarea o que se pueden dividir entre varias. Los defensores de ambos puntos de vista concuerdan ampliamente en que las tareas automáticas necesitan que se les destine poca o ninguna atención. Por lo tanto, cuando al menos una de dos actividades que se desarrollan simultáneamente es automática, hay capacidad de atención suficiente como para realizar ambas. En el experimento recién comentado, parece probable que la repentización haya logrado un alto grado de automaticidad.

## Conocimiento

Cuando discuto algún tema que me resulta entrañable, no siempre gano. Mientras curo mis heridas después de la discusión, suelo encontrar la respuesta exacta que debería haber dado ante el golpe aparentemente devastador de mi oponente. Pero en el calor del debate no puedo llegar al conocimiento adecuado, al argumento que quisiera expresar. De manera similar, los alumnos en un examen suelen comprender, minutos después de salir del aula, lo que deberían haber escrito.

Uso estos ejemplos para subrayar que la habilidad no es una mera cuestión de posesión del conocimiento. Se necesita que ese conocimiento esté disponible en el momento adecuado, en respuesta a la situación que exige su uso. Por ejemplo, lo importante al conducir un automóvil es disminuir la velocidad inmediatamente al ver que la luz verde del semáforo está cambiando a amarilla. No sirve de nada tener el conocimiento conceptual de que la luz amarilla significa "alto" si no se lo puede aplicar a la hora de conducir. Recientemente, los psicólogos cognitivos se han estado ocupando del papel posible en la conducta experta de lo que ellos llaman pares asociativos de *modelo-acción*. Estas asociaciones son algo similar a reglas que un organismo puede aplicar en situaciones particulares. Tienen la siguiente forma: *Si se cumple la condición X, debe realizarse la acción Y.* Se parecen a la relación estímulo-respuesta que tiene una larga historia en las explicaciones conductistas del aprendizaje animal. Pero en estas "reglas de producción", X no necesita ser un simple estímulo externo y, por su parte, Y no tiene que ser una conducta evidente. Podemos incluir cosas tales como los estados psíquicos internos y los objetivos. De modo que la regla "Si ves un auto de policía por el espejo retrovisor, entra en pánico", es una regla de producción, aunque preferiríamos que no existiera.

Un conjunto de reglas de producción suficiente para realizar alguna tarea coherente puede convertirse en un *sistema de producción*. Una de las razones fundamentales por las que los psicólogos han comenzado a pensar en las habilidades en términos de sistemas de producción es el hecho de que semejantes sistemas pueden estimularse muy fácilmente en una computadora. Volveremos a analizar la utilidad de esto en el capítulo 16.

La teoría de los sistemas de producción ofrece una forma útil de comprender cómo puede organizarse el conocimiento al servicio de la habilidad en varios niveles diferentes. Primero, a nivel de las acciones individuales. Ciertos esquemas ambientales pueden gatillar inmediatamente conocimientos sobre lo que se debe hacer en determinada circunstancia. Al conducir, por ejemplo, una luz verde que cambia al amarillo gatilla una aplicación inmediata de los frenos. Sobresalen dos hechos significativos en semejante conocimiento. Uno es que el conocimiento puede ser completamente inaccesible (u "olvidado") hasta que se dan las circunstancias que requieren su empleo. Por ejemplo: ¿sabe usted, sin intentar hacerlo, qué músculo usaría primero si se levantara desde una posición de sentado? El segundo hecho es que nuestra capacidad para adquirir nuevos pares esquema-acción parece ilimitada.

Un estudio realizado sobre conductores de taxímetros (Chase, 1983) nos habla claramente de la accesibilidad. Se llamó al laboratorio a taxistas experimentados de Pittsburgh y se les pidió que describieran los mejores caminos entre pares de puntos de la ciudad. Se trataba de algo que podían hacer. Sin embargo, las rutas que eligieron no fueron en todos los casos las mejores ni las que tomarían los conductores realmente. Esto se demostró repitiendo el estudio en la vida real: se les pidió a los taxistas que condujeran desde A hasta B. Al hacerlo, solían recordar atajos mejores que los que habían mencionado en el laboratorio. Estos atajos parecían impulsados por la experiencia de llegar a determinados lugares. Los conductores hacían comentarios como: "Por supuesto que, si doblo a la izquierda aquí, puedo evitar aquel cuello de botella". Esto, naturalmente, es completamente adaptativo. Los taxistas necesitan encontrar buenos caminos cuando están trabajando, no cuando están pensando en ello.

Una habilidad cognitiva compleja, como jugar al ajedrez, también puede someterse al estudio de los procesos de producción. Podemos suponer que el maestro de ajedrez tiene muchos miles de reglas de producción que vinculan cada esquema concurrente o agrupamiento con una buena movida asociada con ese esquema. De esta forma, los grandes maestros sienten que las buenas jugadas "surgen del tablero". Los malos jugadores tienen que luchar contra las consecuencias de muchos malos movimientos antes de tropezar con uno bueno.

## La estructura de las habilidades

También podemos usar la idea de proceso de producción para conocer la naturaleza de la planificación y la estructura de la habilidad, y la capacidad de seguir la pista de dónde está uno. Para esto, además de las reglas de producción, debemos postular un sistema de memoria operativo en el cual se puede "apilar" un conjunto de objetivos (ver capítulo 5). Un ejemplo simple de una habilidad común es la capacidad de viajar por un país industrializado. Si, por ejemplo, estoy en un ámbito rural y necesito llegar a una gran ciudad rápidamente, puedo fijarme el objetivo de "llegar a Londres". Este objetivo puede actuar como condición X de una regla de producción que podría ser: *"Si el objetivo es llegar a Londres, fijar el objetivo de llegar a la estación de tren más cercana"*. De este modo, el objetivo de llegar a Londres es "empujado hacia abajo" en la pila y el segundo objetivo, —"llegar a la estación"—, pasa arriba de la pila como objetivo actual. Ahora, el objetivo es "lle-

gar a la estación". Esto puede requerir otra regla, que diga: *"Si el objetivo es llegar a la estación, fijar el objetivo de conseguir un taxi"*. Ahora, "conseguir un taxi" se convierte en el objetivo actual, con dos objetivos no logrados apilados debajo. Esto puede requerir otras reglas, como "buscar el número telefónico de la compañía de taxis". En algún punto habrá un objetivo que ponga en juego una regla de producción implicando a la conducta para que alcance el objetivo superior. Cuando esto ocurra, dicho objetivo podrá eliminarse de la pila y el siguiente objetivo se convertirá en el actual. Entonces podrá cumplirse este objetivo, y así con toda la pila hasta que ya no quede ninguno. La pila de objetivos da estructura y dirección a la conducta total.

Sin la estructuración de una pila de objetivos, la conducta tendería a ser incoordinada y plenamente dirigida por la situación en curso. De hecho, cuando la conservación de objetivos en la pila es inadecuada, tiene lugar una falla parcial en la conducta. Podemos observar esto mediante una clase peculiar de error de distracción. Esto nos sucede a la mayoría: entramos a un negocio y nos quedamos allí parados, tontamente, tratando de recordar para qué entramos. El objetivo de "comprar X" nos llevó al nuevo objetivo de "ir al negocio", pero en el camino empezamos a pensar en otras cosas y esta actividad mental interviniente nos llevó a perder el objetivo original de la pila.

En la vida real, la mayoría de la gente no persigue sólo un conjunto de objetivos como los que se describen arriba sino muchas objetivos dispares. A veces resulta difícil tenerlos a todos en mente, y jamás podemos trabajar de manera directa en más de un par de ellos a la vez. En consecuencia, la capacidad de recordar los múltiples objetivos que uno tiene y en qué etapa está en el logro de cada uno de ellos es una importante habilidad por sí misma. Todos conocemos individuos que siempre parece estar olvidándose de hacer cosas mientras que hay otros que siempre se manifiestan capaces de cumplir con una amplia variedad de compromisos diferentes y complejos. El papel de las ayudas externas a la memoria resulta muy importante aquí: las agendas, los nudos en los pañuelos, las listas de cosas para hacer y las notas dirigidas a uno mismo (ver capítulo 8). Podemos ver a la movilización de estas ayudas como un reconocimiento explícito del hecho de que los estímulos externos son los mejores y más confiables disparadores de la acción. Ser hábil para cumplir múltiples compromisos supone ser hábil para manejar el propio entorno, sobre todo la memoria necesaria en el lugar y el momento correctos.

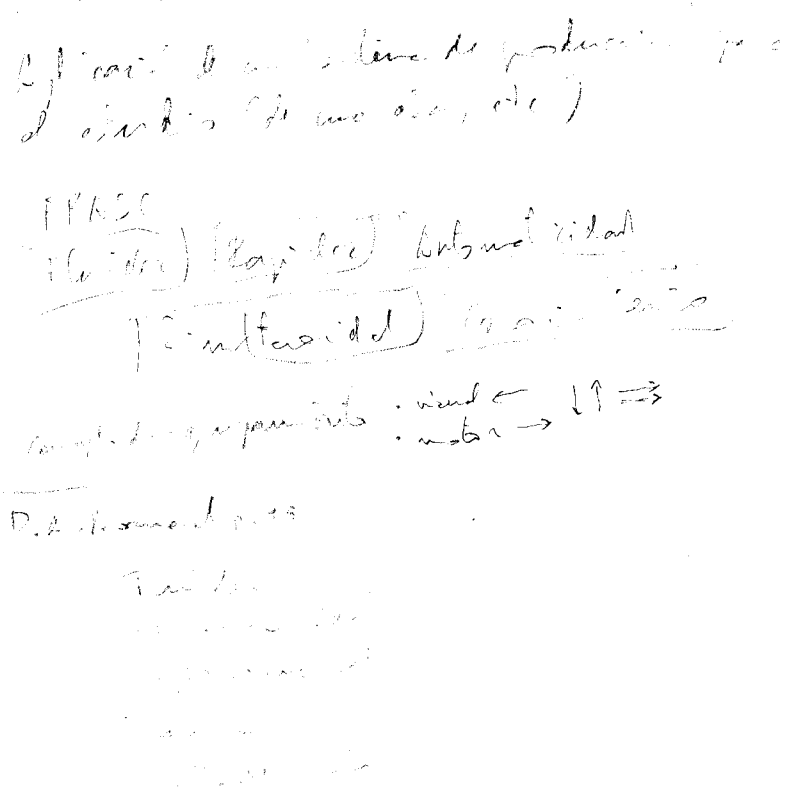
## Lecturas adicionales

REASON, J. y MYCIELSKA, K. (1982) *Absent Minded? The Psychology of Mental Lapses and Everyday Errors.* Nueva York: Prentice Hall.

Relato entretenido y de fácil lectura sobre muchas de las nociones teóricas subyacentes en el pensamiento actual acerca de la habilidad.

CHASE, W.G. y CHI, M.T.H. (1980) "Cognitive skill: implications for spatial skill in large scale environments". En J. Harvey (Comp.) *Cognition, Social Behaviour and the Environment.* Hillsdale, NJ: Erlbaum.

Útil resumen de la investigación sobre ajedrez y las habilidades relacionadas. Breve, claro y de lectura muy amena.



# 3

# La adquisición de la habilidad

*John Sloboda*

En el capítulo anterior vimos algunas de las características de la conducta experta. Nos ocuparemos ahora de una cuestión relacionada con ella: ¿qué tiene que hacer uno para convertirse en experto? El animal humano es único en cuanto a la cantidad de habilidades que adquiere mediante el aprendizaje. También somos únicos en la diversidad de nuestras habilidades. Algunos seres humanos son matemáticos expertos, otros son músicos expertos y otros, mecánicos expertos. ¿Cómo es posible que la gente sea tan diferente en su perfil de habilidades?

Muchos animales parecen extremadamente hábiles. Obsérvese, por ejemplo, la forma en que un gato caza un pájaro. Se agazapa para que el suelo disimule su presencia y se mueve lenta y silenciosamente hacia su presa; luego, cuando ya es difícil permanecer oculto, salta con las garras extendidas para dar a su presa el mínimo tiempo de reacción posible. La habilidad de mi gato en esta operación queda demostrada por la cantidad de pájaros y ratones muertos que deposita con orgullo en el suelo de la cocina de mi casa. Podría aprender mucho sobre técnicas de caza observando cuidadosamente a mi gato.

¿De dónde proviene la habilidad del gato? Hasta donde sabemos, la mayoría de las habilidades animales son heredadas y no aprendidas. Los gatitos criados en aislamiento muestran también un buen repertorio de conductas de acecho y cacería (Hinde, 1966). Parecen instintivas. Además, todos los individuos de una especie dada parecen heredar globalmente el mismo conjunto de habilidades.

¿Es posible que las habilidades humanas también sean heredadas? En la primera sección de este capítulo, examinaremos el punto de vista según el cual la adquisición de habilidades se atribuye al talento como potencial innato. Luego, contrastaremos este punto de vista con otro, según el cual el talento es menos importante que la oportunidad y la práctica. Concluiremos echando una mirada a un estudio de caso de una habilidad excepcional, que muestra que realmente necesitamos las dos perspectivas para comprender bien la adquisición de habilidades.

## El talento y la adquisición de habilidades

Es indiscutible que las personas dotadas tienden a tener hijos dotados. A veces, talentos específicos (como el talento para la música) parecen correr en las familias durante varias generaciones. Por desgracia, hace mucho que se sabe que este hallazgo, en sí, dice muy poco sobre cómo se adquirió la habilidad. Podría ser que el hijo de un músico hubiera heredado cierta propensión especial hacia la música. De la misma forma, también es posible que un padre músico proporcione a su hijo más oportunidades y aliento para que pueda adquirir habilidades musicales que un padre no músico. En la vida real, las influencias de la herencia y el medio están inseparablemente unidas, y es muy difícil establecer qué influencia en particular resultará decisiva. Gran parte de la investigación psicológica seria en este campo se ha centrado en la noción de inteligencia y ha sido objeto de grandes controversias. Una de estas controversias se refiere a si existe alguna capacidad o facultad unitaria que constituya el fundamento para la acción inteligente o si, en cambio, la inteligencia es sólo la conjunción de varias habilidades distintas e independientes. Quienes sostienen el primero de dichos puntos de vista apuntan al hecho de que aquellos que son buenos en un tipo de cosas tienden a ser buenos también en toda otra gama de cosas. La base de los tests de inteligencia reside en esto. Precisamente porque muchas habilidades se correlacionan entre sí, podemos obtener una imagen útil de la capacidad de una persona mediante el muestreo de unas pocas habilidades. Pero la correlación no es perfecta y encontramos en casi todas las personas áreas de habilidad que son mucho mejores o mucho peores de lo que se podría predecir sobre la base de las mediciones de su inteligencia. Un buen ejemplo de esto (sobre el que volveremos en el próximo capítulo) es el del disléxico, que lee mucho peor de lo que se podría predecir a partir de sus otros logros intelectuales. Gardner (1983) ofrece una visión accesible de los aspectos que rodean a esta controversia general.

Un segundo punto controvertido es la medida en que podemos hallar medios confiables para evaluar los efectos separados de la herencia y el medio. Este tópico está lleno de dificultades metodológicas y teóricas. También es un campo minado en lo político e ideológico, ya que frecuentemente se lo vincula con la cuestión de las diferencias del racismo y la herencia (Eysenck versus Kamin, 1981). Muy poca gente llega al tema sin prejuicios. Sin embargo, me arriesgaré y diré que hay probablemente tres formas en que las características heredadas podrían tener un efecto significativo sobre la adquisición de habilidades:

a) Hay algunas condiciones transmitidas genéticamente que provocan retardos mentales irreversibles, presumiblemente como consecuencia del desarrollo anormal del cerebro o el sistema nervioso (por ejemplo, el síndrome de Down). Estas condiciones pueden destruir aptitudes esenciales para algunas habilidades. Todas esas condiciones, desgraciadamente, tienden a imponer impedimentos a quienes las poseen. No hay condiciones conocidas de este tipo que confieran ventajas significativas.

b) Las características físicas heredadas son indudablemente un factor para la determinación del logro en algunas habilidades. Las posibilidades para que alguien se convierta en cantante de ópera dependerán mucho de la forma de la cavidad vocal de esa persona; las posibilidades de que una persona se dedique a la danza dependerán mucho de la conformación de su esqueleto, etc. También es probable que sean importantes las características físicas heredadas del sistema nervioso.

c) Hay todo un conjunto de lo que se podrían denominar factores de disposición y motivación que podrían tener un efecto indirecto sobre la adquisición de habilidades. Desde su más tierna infancia, los niños parecen diferir en su nivel de actividad, la cantidad de horas de sueño que necesitan, su temperamento primario, etc. Las investigaciones sobre la personalidad adulta muestran significativas diferencias individuales en factores tales como la capacidad de concentración en tareas aburridas durante períodos largos. También hay evidencias de que existen preferencias estables por ciertos campos de actividad. Algunas personas parecen tener tendencia a las actividades visuales y otra a las verbales. Estas preferencias pueden vincularse a la infancia temprana, donde se observan significativas diferencias en el perfil evolutivo (Kagan, Lapidus y Moore, 1978). Algunos niños desarrollan la habilidad del lenguaje tempranamente, pero su desarrollo físico se retrasa. Otros muestran una tendencia inversa. Todos estos factores podrían tener una fuerte influencia tanto sobre el tipo como sobre el nivel de habilidad adquiridos. Podemos suponer que es más probable que el extrovertido físicamente precoz adquiera habilidades futbolísticas mientras que el introvertido verbal podría desarrollar más fácilmente habilidad para la poesía. Esto no es una cuestión de capacidad. Más bien, representa la cantidad de esfuerzo que una persona puede estar dispuesta a destinar a tipos peculiares de actividad. Algunas actividades parecen "ir contra la naturaleza de la persona", mientras que otras parecen intrínsecamente gratificantes.

## La práctica y la adquisición de habilidades

Paul Tortelier es uno de los principales violonchelistas profesionales del mundo. Sus padres eran obreros con escasos recursos económicos y sin logro musical peculiar alguno. Antes de que Paul naciera, decidieron que su hijo debería lograr grandes cosas en el mundo y que iban a formarlo como violonchelista. Mediante el esfuerzo, los ahorros y el sacrificio personal llegaron a comprarle instrumentos, pagarle lecciones y ofrecerle todo tipo de apoyo, y su hijo logró lo que sus padres querían.

Semejante carrera de la pobreza a la riqueza toca una cuerda íntima en muchos de nosotros y sugiere la siguiente pregunta. Supongamos que al pequeño Tortelier lo hubieran arrancado de su cuna al nacer y lo hubieran reemplazado por otro niño, elegido al azar entre toda la población; supongamos, además, que sus padres no hubieran advertido el cambio y lo hubieran criado como si fuera su hijo, ¿habrían podido convertir a cualquiera en un violonchelista de nivel mundial?

Si nos inclinamos a responder que "sí" a esta pregunta, estaremos basando nuestra respuesta en el supuesto de que las habilidades no se adquieren en virtud de lo que uno es sino de lo que uno hace. Cualquiera puede adquirir una habilidad si hace lo necesario para ello.

Es imposible, en la práctica, hacer el experimento de la sustitución hipotética del niño. Pero lo que podemos hacer es observar cómo se comportan los individuos en el proceso de adquisición de habilidades y preguntarnos si hay algo que parezca ligado a la adquisición exitosa de una gran variedad de habilidades para una gran cantidad de personas. No resultará sorprendente enterarnos de que el factor simple más importante que los psicólogos conocen es la práctica. De la misma manera, la práctica es el ingrediente que más destacan los entrenadores deportivos y los maestros de piano.

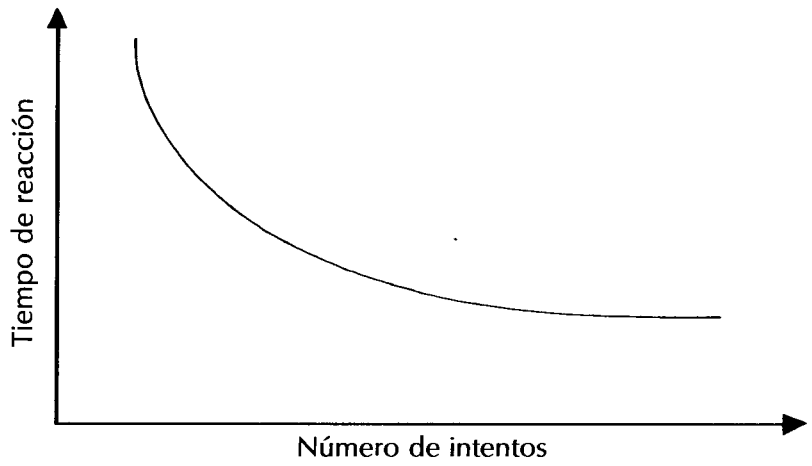
Muchas investigaciones se han ocupado de los efectos de la práctica en las habilidades perceptivo-motrices simples. Un experimento muy usual es aquel en el que se mide el tiempo de reacción para apretar un botón en respuesta a una luz que se enciende. A veces sólo hay un botón y una luz, mientras que en otros experimentos puede haber varios botones y luces y los sujetos deben responder quizá a una luz por vez o a varios esquemas de luz que aparecen juntos. Un estudio de Siebel (1963) resulta representativo: los sujetos veían 10 luces ordenadas en una fila horizontal; debajo de cada luz había un botón; los sujetos debían apoyar sus dedos sobre las teclas y mirar las luces; después de una señal, se encendían algunas de las luces, de mane-

ra azarosa, y los sujetos debían presionar las teclas correspondientes a las luces encendidas (y no otras). Esta tarea fue repetida más de 40.000 veces.

Siebel midió el tiempo de reacción medio para presionar la combinación correcta de teclas. Como vimos en el capítulo anterior, la rapidez en la respuesta es una característica crucial de la habilidad, y por eso muchos investigadores han usado el tiempo de reacción como medida primaria del nivel de habilidad. Cuanto más rápido responde uno, más hábil (o experto) es, siempre y cuando, naturalmente, las respuestas sigan siendo correctas. Siebel descubrió que el tiempo de reacción disminuía de un nivel inicial de unos 2 segundos a 0,4 segundo después de tanta práctica. Es difícil imaginar que esta tarea revistiera algún interés para los sujetos que realizaban el experimento, y sin embargo la repetición pura y persistente fue causando una continua mejoría a lo largo de los 40.000 intentos.

Se puede obtener información más detallada haciendo un gráfico en el que las variables sean el tiempo de reacción (TR) y el número de intentos. La figura 3.1 muestra el resultado que se obtiene al hacerlo. Ejemplifica lo que a veces se denomina "ley de los resultados descendentes". La práctica temprana da como resultado grandes ganancias en la velocidad, pero la misma práctica más adelante sólo ofrece pequeñas mejorías. A veces este achatamiento en la curva de adquisición se da porque se llega a los límites fisiológicos; por ejemplo, la velocidad fija a la que los nervios conducen los impulsos. Pero más frecuentemente el achatamiento comienza mucho antes de acercarnos a esos límites.

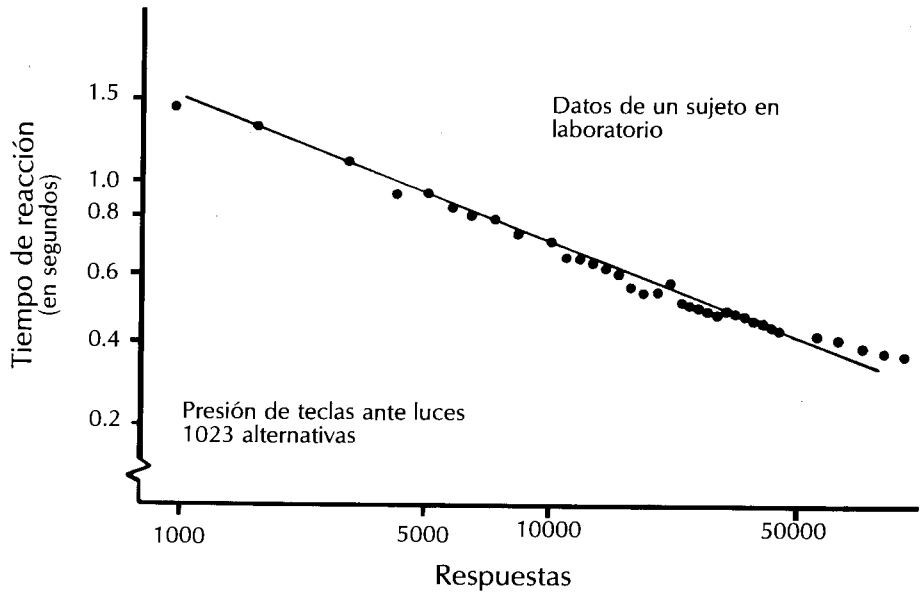
*Figura 3.1: Gráfico típico de tiempo de reacción contra cantidad de ensayos*



¿Por qué las mejoras disminuyen con la práctica? Una posibilidad sería que la gente se concentra cada vez menos mientras repite una tarea y no logra obtener beneficios de los intentos en la misma medida que al comienzo. Esta podría ser una explicación plausible para la declinación en una sesión. Pero no explicaría una caída pareja en 40.000 ensayos realizados en un lapso de muchos días. Si el aprendizaje dependiera enteramente de la concentración, deberían verse muchos altibajos, en la medida en que la propia concentración los sufriera.

Antes de considerar una explicación más probable, es importante notar que las curvas obtenidas experimentalmente, como la que se muestra en la figura 3.1, frecuentemente se aproximan a una función logarítmica. Si se las vuelve a organizar sobre ejes logarítmicos, se convierten en líneas rectas. La Figura 3.2 muestra los datos de Siebel así reorganizados.

*Figura 3.2 Datos de Siebel reorganizados sobre ejes logarítmicos. (Según Fitts y Posner, 1967.)*



## Explicación de la ley de las mejoras decrecientes

Newell y Rosenbloom (1981) ofrecieron una explicación para las mejoras decrecientes; se basa directamente en las nociones de agrupamiento y las reglas de producción modelo-acción que ya hemos explorado en el capítulo 2. Sostienen que mientras un sujeto desarrolla alguna tarea como en el experimento de Siebel, continuamente intenta agrupar las luces en esquemas. Cada agrupamiento se asocia a su vez con un conjunto de dedos determinado para la respuesta. Para empezar, el conjunto se procesa como diez luces individuales y las respuestas se dan ante cada luz por separado. A medida que pasa el tiempo, el sujeto empieza a responder, por ejemplo, a pares de luces como si fueran unidades simples, luego a grupos de tres, etc., hasta llegar a un techo teórico en donde las 10 luces se ven como un conjunto único (así como al aprender a leer los alumnos progresan desde la lectura letra por letra hasta la lectura de palabras completas — ver capítulo 4). Si suponemos, por cuestiones de simplicidad, que lleva un tiempo constante percibir y responder a un agrupamiento simple, observamos que cuanto mayores sean los agrupamientos, menos tiempo toma responder al conjunto total.

Es bastante fácil ver intuitivamente que cuantos más agrupamientos aprenda una persona, más rápida será su respuesta a cualquier conjunto de luces dado, porque más posibilidades tendrá de detectar un esquema familiar. Por lo tanto, la práctica aumentará la velocidad de la tarea. Lo que resulta menos fácil de ver es por qué esta noción también predice la ley de las mejoras decrecientes, pero en realidad lo hace. Se debe a que un agrupamiento pequeño determinado se repetirá con mayor frecuencia que uno grande, y por lo tanto se lo puede aprender y poner en práctica más rápidamente.

Para ver qué importancia tiene esto, consideremos primero qué sucede cuando agrupamos al conjunto en cinco grupos de un par de luces adyacentes cada uno. Cada grupo (es decir, cada par de luces) tiene cuatro configuraciones posibles (apagada-apagada, apagada-encendida, encendida-apagada, encendida-encendida). Hay 5 pares, de modo que hay 5 veces 4 configuraciones, lo que da un total de 20 posibilidades. Hay que aprenderlas todas junto con sus 20 esquemas de respuesta. Si en cualquier intento del experimento se encienden al azar la mitad de las luces, es probable que cualquier configuración de un par determinado aparezca, como promedio, una vez cada cuatro ensayos. De manera que como cada uno de los 20 grupos aparece frecuentemente, hay muchas posibilidades de ver y aprender a responder a cada uno de ellos. Dado nuestro supuesto de que se necesita un tiempo constante para responder a cualquier agrupamiento, una persona que

ha aprendido estos 20 esquemas de pares podrá reducir su tiempo de respuesta, viendo siempre al conjunto como 5 agrupamientos y no como 10 luces independientes. Este es el motivo por el cual las primeras etapas de la práctica pueden dar como resultado grandes ganancias en el rendimiento.

Consideremos ahora lo que incluiría el aprendizaje para responder al total como 2 conjuntos, cada uno de los cuales contuviera 5 ítem. Hay 32 configuraciones diferentes posibles de encendido y apagado para un conjunto de 5 luces. Hay 2 grupos en el total de 10, de modo que hay 64 configuraciones para aprender en total (2 veces 32). Si en cualquier ensayo al azar se encienden la mitad de las luces, cada configuración de un grupo sólo aparecerá una de cada 32 veces. Hay más esquemas para aprender y muchas menos posibilidades de observar cualquier esquema determinado. Por lo tanto, le tomará a una persona mucho más tiempo pasar de procesar el total como 5 grupos a procesarlo como 2 grupos que lo que le tomó pasar de 10 a 5 grupos. El primer movimiento reduce el tiempo de reacción a la mitad, de 2 segundos a 1 segundo. El segundo movimiento reduce el tiempo de reacción en 3/5, de 1 segundo a 0,4. No sólo habrá tomado más tiempo realizar el segundo pasaje sino que también habrá producido un aumento absoluto menor en la velocidad. Esto explica el fenómeno de las mejoras decrecientes y su aproximación a la función logarítmica.

Estudios como el de Siebel y la teorización construida en torno a él ofrecen oportunidades para la reflexión matemáticamente precisa sobre las habilidades. Muestran con especial elegancia porque hace falta tanto tiempo para llegar a ser de verdad bueno en algo. Y, sin embargo, esos estudios tienen varios defectos. En primer lugar, las habilidades estudiadas son muy simples. Requieren una respuesta simple a un estímulo y no una secuencia coordinada. En segundo lugar, los propios estímulos carecen de esquema y estructura naturales. Es muy raro en el mundo real que la materia con la que tratamos sea tan impredecible. Hay semejanzas entre las situaciones nuevas y las viejas, que ya conocemos. Podemos hacer suposiciones y planes informados. Algunos esquemas aparecen con mucha mayor frecuencia de lo que esperaríamos por casualidad. En tercer lugar, normalmente tenemos la posibilidad de decidir cómo practicaremos una habilidad. En la tarea de Siebel, los individuos debían responder a diez luces en todos los intentos. Aprender las respuestas para cada mano separadamente podría ser una buena estrategia.

Esto nos lleva a la cuarta crítica. En experimentos como el de Siebel, los individuos suelen trabajar totalmente solos; la habilidad carece por completo de contexto social. Como se observa en el capítulo 1, esto no sucede

con las habilidades de la vida real. Usualmente, tales habilidades se aprenden con la ayuda de algún tipo de entrenador. Puede ser alguien que hace mucho practica la habilidad, un libro o quizá sólo el deseo de copiar lo que una vez se vio hacer. Sin embargo, por lo general se cuenta con alguna fuente de comentarios y sugerencias. Por todos estos motivos, el grado de progreso en una tarea como la propuesta por Siebel probablemente muestre un cálculo pesimista de nuestra capacidad de adquirir habilidades en la vida real.

## Mejorando los efectos de la práctica

En esta sección, nos ocuparemos brevemente de algunas cosas que han demostrado ayudar a que muchos tipos de práctica resulten más eficaces. Son características de la práctica en torno a las que se debe construir cualquier programa de entrenamiento. La primera y principal es la retroalimentación (*feedback*).

### Retroalimentación

La retroalimentación es, sencillamente, conocimiento de lo que se ha logrado con las propias acciones. Puede ser de dos tipos, intrínseca y extrínseca. La retroalimentación intrínseca es la que se obtiene como consecuencia directa del movimiento y las sensaciones corporales. Si uno mueve un brazo, puede sentir el resultado por sinestesia y puede también ver dónde está el brazo por medio de la vista. La retroalimentación extrínseca es la proporcionada por alguna fuente o agente externo, que suele ser un entrenador o un maestro. Que a uno le digan si una respuesta que ha dado es correcta o incorrecta es una forma de retroalimentación extrínseca. La retroalimentación de una u otra clase es esencial para la adquisición de cualquier habilidad. Uno no puede mejorar a menos que tenga manera de juzgar su rendimiento actual y en qué dirección se deben producir cambios.

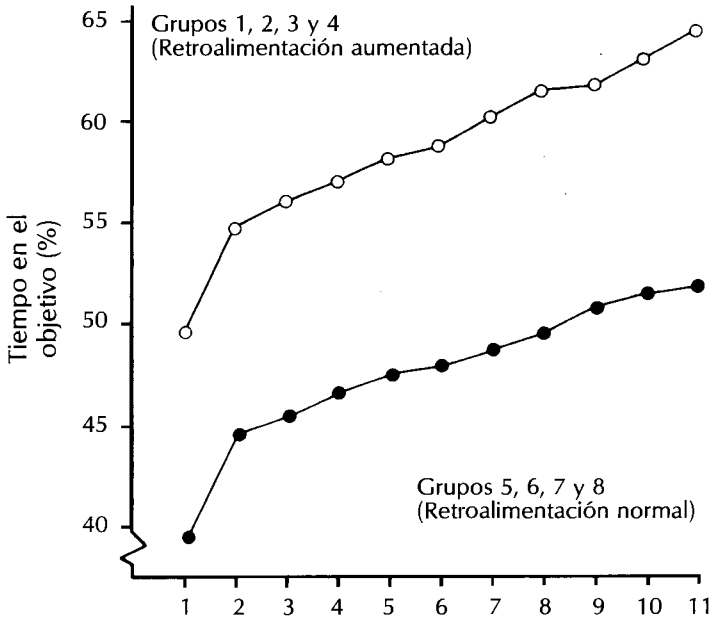
Hay varias formas de aprovechar la retroalimentación para una mejor adquisición de la habilidad. La primera es prestar verdadera atención a la retroalimentación existente. Esto puede ser algo tan simple como revisar lo que uno ha escrito cuando está haciendo una carta o una composición. Muchos estudiantes (y mucha gente, en general) no lo hacen, y sufren las consecuencias. La segunda es asegurarse de que la retroalimentación sea inmediata. Los comentarios sobre una composición devuelta después de varias semanas resultan poco útiles, porque el autor probablemente ya se ha olvida-

do de los pasos precisos que dio en la preparación del texto. Ya no es posible relacionar la retroalimentación con aspectos específicos de la conducta. En un mundo ideal, el tutor supervisaría cada paso del proceso de escritura de una composición, proporcionando los comentarios pertinentes en cada etapa. Esto significa que las dos personas tendrían que estar en contacto estrecho y prolongado, con la probabilidad de que se desarrollara una relación intensa. Y las relaciones intensas son familiares en los contextos en los que hay un maestro y un aprendiz, un entrenador y un atleta o un profesor de teatro y un alumno. Un ejemplo en la literatura es la relación entre Svengali y la cantante joven, Trilby, en la novela homónima de Du Maurier.

Más exóticas que la retroalimentación común son las formas de ayudar a la gente a darse cuenta de que tienen sensaciones internas de las que normalmente no son conscientes. Utilizando una muestra visual que controle algún aspecto de nuestro estado corporal (como, por ejemplo, el latido del corazón) se puede aprender a identificar las sensaciones corporales que acompañan a los cambios en estos estados, y alguna gente es capaz de controlar estos estados aun cuando la realimentación aumentada (conocida como bioretroalimentación) desaparece.

A veces, la retroalimentación tiene un efecto básicamente motivador. Proporciona información que no es esencial pero que parece aumentar el compromiso en la tarea. Por ejemplo, Smode (1958) realizó una investigación en la que se practicaba lo que se conoce como "tarea de localización": había que mantener una aguja de movimiento azaroso sobre un punto central, rotando un dial con movimientos compensatorios. Después de cada intento, todos los sujetos recibían retroalimentación "normal", en la que se les decía el tiempo que habían logrado mantener la aguja centrada durante ese intento. Además, a un grupo se le proporcionó retroalimentación "aumentada", con un contador en el que se acumulaba su puntaje total en la suma de intentos. Esta retroalimentación aumentada tuvo un efecto drástico en el rendimiento, como lo demuestra la figura 3.3. La ventaja se mantuvo aun después de eliminar la retroalimentación aumentada. En un segundo día de aprendizaje, en el que todos los sujetos recibieron retroalimentación normal, los que habían recibido retroalimentación aumentada el primer día siguieron obteniendo mejores resultados.

*Figura 3.3*




## Espaciamiento

Si uno puede dedicarle seis horas a la práctica de una habilidad determinada, ¿hay que dedicarle seis horas de trabajo continuo o seis horas separadas en días diferentes? En general, parece mejor espaciar la práctica en varias sesiones que amontonarla toda en una sola sesión (Welford, 1968). Hay varias razones para hacerlo. Una es que, cuando se realiza una tarea repetitiva, tienden a aparecer la fatiga y la disminución de la atención. Como consecuencia, pueden formarse malos hábitos y la práctica posterior reforzará esta conducta indeseable. Los descansos frecuentes pueden lograr que la fatiga y los malos hábitos se disipen. Por otra parte, cuando la motivación es alta, es muy posible sumergirse totalmente en una tarea durante muchas horas sin pérdida evidente de la atención, de modo que la regla de espaciar la práctica no es de nigún modo universalmente beneficiosa.

## Actitud

Un tercer tema es la cuestión de qué actitud o actividad psíquica debe acompañar a la práctica. ¿Debe uno repetir una tarea mecánicamente, con una actitud esencialmente pasiva, pretendiendo fijar la habilidad por mera repetición? ¿O debe estructurar la tarea activamente, buscando patrones y semejanzas y adivinando lo que podría seguir? Esta es una pregunta algo inquietante, y los psicólogos no han podido ofrecer hasta hoy respuestas definitivas. Es muy evidente que la repetición pura e insistente tiene efectos beneficiosos (ver el estudio de Siebel que se describió antes). También se habla mucho, en campos como el de la psicología del deporte, sobre la importancia de no ser demasiado cerebral y permitir que el cuerpo "se haga cargo". Hay muchas evidencias de que hacemos las cosas mejor si podemos llevarlas al nivel del control automático y permitir que sea el sistema el que "se ocupe"; y es posible que donde haya movimientos físicos, parte de semejante proceso sea la recuperación de esquemas de movimiento tempranos y parcialmente instintivos que hayan sido reprimidos por intentos erróneos de controlarlos (Tinbergen, 1974).

Incluso en un nivel cognitivo más complejo, gran parte del aprendizaje de habilidades resulta sólo incidental y casual. A veces parece que si uno está lo suficientemente interesado en una actividad, su compromiso dará como resultado, por sí mismo, la adquisición de cierto grado de habilidad. Cuando uno intenta adquirir una habilidad vinculada con una actividad que no le interesa intrínsecamente, hace falta un sustituto para el interés. En el capítulo 12 se presentan algunas de las formas en las que los investigadores de la educación han llegado a establecer la naturaleza de las estrategias y habilidades de estudio eficaces, así como también las estrategias de enseñanza necesarias para apoyarlas. Puede resultar útil ver estas estrategias como formas de lograr que el alumno no motivado haga aquello que (con motivación) haría naturalmente. Si alguien piensa que esto construye una falacia para que la gente estudie lo que realmente no le interesa y una acusación a nuestro sistema educativo, no estará solo (Illich, 1973). Pero este punto de vista es por sí mismo incompleto. Porque aun cuando nos interese dominar una habilidad determinada, también podemos aprovechar el consejo y la retroalimentación. En la adquisición de habilidades, la función del entrenador es proporcionarlos.

## Un talento excepcional - un estudio de caso

Me gustaría terminar este capítulo describiendo una habilidad bastante extraordinaria que he estado estudiando recientemente (Sloboda, Hermelin y O'Connor, 1985). El sujeto de este estudio, a quien llamaremos NP, es un joven autista de poco más de 20 años de edad. Como muchos autistas, tiene un retardo mental grave. Casi no tiene lenguaje espontáneo, manifiesta retracción social y otros patrones de conducta extraños y necesita el cuidado pleno de una institución especializada. Lo que lo hace diferente de la mayoría de los autistas es que en una pequeña región de su cerebro es cualquier cosa menos un retardado. Tiene una sorprendente capacidad para memorizar música para piano.

En las pruebas que realizamos, NP fue capaz de memorizar prácticamente nota por nota una pieza para piano clásica (Grieg, Op. 47, Nº 3) de más de 2 minutos de duración... en 12 minutos. Lo hizo simplemente escuchando segmentos de una grabación de la obra y repitiendo en el piano lo que oía. Al día siguiente, sin que hubiera tenido posibilidades de ensayar, todavía podía reproducir la obra casi a la perfección. En ningún momento vio la partitura o a otra persona interpretando la obra; tampoco recibió ningún otro tipo de retroalimentación externa. No se le dijo en qué clave estaba la obra, pero de inmediato fue capaz de elegir las notas correctas y tocarlas con la combinación correcta de dedos, aunque probablemente nunca la había oído antes y obviamente jamás había intentado interpretarla.

Este nivel de habilidad de memoria es muy raro, incluso entre los músicos de inteligencia superior al promedio. Probablemente sólo lo igualan los logros legendarios de unos pocos prodigios, como Mozart y Erwin Nyerghihazi (Revesz, 1925). NP es un caso de prueba ideal para algunas de las preguntas que hemos planteado en este capítulo sobre la adquisición de habilidades.

En primer lugar, podemos preguntar si la habilidad se ajusta a las nociones de talento heredado que hemos analizado al comienzo de este capítulo. Ciertamente, hay pocas evidencias de logros musicales en el medio familiar inmediato de NP, carente casi por completo de recursos. Sin embargo, NP demostró gran precocidad para la música sin tener mucho apoyo familiar. Las grabaciones que se conservan de cuando tenía 6 años muestran un nivel de habilidad muy superior al de cualquier niño de esa edad. Aunque en su casa no había piano, tuvo un buen desempeño en su primera exposición pública al instrumento. Por lo que sabemos sobre sus circunstancias, la posibilidad de entrenamiento secreto puede descartarse casi por completo. Todavía

hoy hay un sentido real en el que a NP no se le puede enseñar. Sus "lecciones" consisten en piezas que su maestro toca para que él memorice. Nadie le ha enseñado a memorizar como lo hace (ver capítulo 6). Nadie sabe cómo lo hace y él no puede explicarlo. Se le han dado posibilidades, pero NP debe ser uno de los ejemplos más puros de experto autodidacta existentes.

A primera vista, NP parece proporcionar fuertes evidencias de la existencia de un talento heredado para la música. Su carencia de logros en cualquier otro campo de habilidad lo hace parecer un don específico para la música. Pero las apariencias engañan y convendría examinar más cuidadosamente este "don". Lo primero que se observa es que la medida del don es igualada por lo que podemos llamar obsesión de NP por la música. Cuando puede elegir libremente una actividad, casi siempre quiere escuchar música (en la radio, en grabaciones, etc.) o interpretarla. También parece que, cuando no está haciendo ninguna de esas dos cosas, está pensando en la música. No sería absurdo suponer que su mente está ocupada con música, con una concentración intensa, de 5 a 6 horas por día. Como para Mozart, la música es su amor y su vida. Es difícil saber de dónde proviene esta fascinación obsesiva por la música. Podría tratarse de una disposición genuinamente heredada o provenir de alguna experiencia temprana crucial. Una característica general de la personalidad autista parece ser la obsesividad; y hay quienes tienen habilidades excepcionales similares en otras áreas, como la del cálculo mental. Quizá la mente autista esté conformada para aferrarse a una obsesión, pero la obsesión específica a la que se aferra puede estar totalmente determinada por las circunstancias.

Un segundo argumento para no tratar la habilidad de NP como un don musical específico proviene de un análisis más estrecho de la forma en que memoriza música. Además de interpretarle piezas clásicas como la de Grieg, lo probamos con una obra atonal de Bartok, que rompía muchas de las reglas de la música "normal". Aunque esta pieza de Bartok (Mikrohosuios, Escalas Tonales Totales del Libro 5) tenía muchas menos notas que la de Grieg, a NP le resultó casi imposible de memorizar. Entendemos que esto demuestra que la habilidad memorística de NP se basa en la capacidad de agrupar la música convencional en esquemas de nivel superior. Cuando estos patrones están ausentes, como en el caso de Bartok, el rendimiento se ve gravemente perjudicado. Esto es igual que en los estudios sobre ajedrez de los que hablamos en el capítulo 2, donde la memoria del gran maestro para las posiciones de las piezas en el tablero era superior a la del principiante sólo cuando el tablero estaba tomado de un juego real y no construido al azar. Es-

to lo confirma el análisis de los pocos errores que NP cometió en la obra de Grieg: casi todos fueron errores estructuralmente plausibles, en los que reemplazó un agrupamiento musicalmente adecuado por algo similar y también adecuado (de la misma forma en que uno puede recordar mal una oración, sustituyendo una palabra por un sinónimo).

Toda la evidencia que tenemos (ver también la descripción de la adquisición de la habilidad de la memoria en Steve Falloon, en el capítulo 7) es que la práctica intensa es el camino único e inevitable para la formación de agrupamientos que posibiliten habilidades tan impresionantes. Pero, como lo demuestran los datos de Siebel, este proceso de agrupamiento funciona con cualquier material y no parece muy diferente para la música que para cualquier otra materia específica. El caso de NP confirma la ecuación general *Motivación + Práctica = Habilidad.* Las raíces de las motivaciones específicas son casi por completo desconocidas, pero fácilmente podrían tener un componente hereditario. Sin embargo, la práctica puede realizarse sin motivación particularmente fuerte y, por lo menos en apariencia, nuestro sistema cognitivo responderá agrupando las entradas y las salidas de cualquier tarea que bastante suficientemente. De todos modos, la práctica necesaria para ganar experiencia es casi siempre insostenible sin una fuerte motivación. Incluso en el mejor de los ámbitos del mundo real, la mayoría de nosotros fracasaría en el intento de adquirir muchas habilidades, no porque seamos incapaces de hacerlo sino más bien porque, para nosotros, no parecen importar demasiado.

## Lecturas adicionales

VERNON, P.E., ADAMSON, G. & VERNON, D.F. (1977) *The Psychology and Education of Gifted Children*. Londres: Methuen.
Sobrio relato de la perspectiva de las habilidades como "talento" y su adquisición.

FITTS, P.M. & POSNER, M.I. (1967) *Human Performance*. Belmont: Brooks/Cole.
Pese al tiempo transcurrido desde su publicación, sigue siendo una de las mejores introducciones a la psicología de las habilidades percepto-motrices y su adquisición.

ANDERSON, J.R. (1982) "Acquisition of cognitive skills". *Psychological Review*, 89, 369-406.
Este ensayo, si bien es algo técnico, ofrece una visión global clara del pensamiento reciente.

# 4

# La lectura: un estudio de caso de habilidades cognitivas

*John Sloboda*

En los capítulos 2 y 3 nos ocupamos de algunas características generales de las habilidades, usando ejemplos tomados de diferentes habilidades específicas. Esto nos ha proporcionado una imagen global de lo que comparten la mayoría de las habilidades. Pero cada habilidad es única, en algunos sentidos, y plantea problemas psicológicos también únicos. En este capítulo sometemos una habilidad muy importante a un escrutinio más detallado que el que ha sido posible hasta ahora. Semejante escrutinio es necesario para recordarnos que las habilidades humanas son más variadas y complejas que cualquier afirmación general sobre ellas.

Hay varias buenas razones para considerar a la lectura una habilidad representativa. En primer lugar, para todos los que viven en nuestra cultura, saber leer tiene una importancia vital. Pese a ello, muchas personas tienen dificultades para aprender a leer y muchas más sienten que no lee tan bien o tan rápido como quisiera. Comprender lo que conlleva la lectura experta bien puede sugerir algunos remedios que ayuden a superar las dificultades de la lectura.

Un segundo motivo para elegir el análisis de la lectura es que se han hecho muchas investigaciones sobre ella. Esto refleja en parte su importancia cultural y educativa, pero también la facilidad relativa con la que se puede construir material textual bien controlado para trabajo experimental. La tecnología moderna nos permite ejercer un control experimental casi absoluto sobre las características relevantes del entorno de un lector y verificar hipótesis específicas acerca de la forma en que los individuos procesan estas características durante la lectura. Esto no sucede con muchas otras habilidades, como, por ejemplo, jugar al fútbol.

En tercer lugar, la lectura también presenta, en una forma particularmente aguda, un problema que domina gran parte de la investigación cognitiva: que la lectura en silencio casi no implica conducta observable alguna. Los psicólogos han tenido que diseñar métodos ingeniosos para seguir los

procesos psíquicos sin interrumpir o alterar la conducta en estudio. La dificultad se incrementa porque, para la mayoría de nosotros, gran parte de los subprocesos de la lectura no están abiertos a la inspección consciente. Los procesos de la lectura son rápidos y automáticos. Somos conscientes en primer lugar del sentido que "salta de la página" hacia nosotros, no del proceso por el cual lo comprendemos. Mientras usted, lector, lee esta oración, vea si puede responder, por sus propias observaciones, alguna de las siguientes preguntas:

a) ¿Sus ojos se mueven suavemente a lo largo del renglón mientras lee?

b) ¿Cuánto tiempo necesita desde que ve una palabra hasta que la identifica?

c) ¿Descubre usted el sonido de una palabra antes que su significado o es a la inversa?

d) ¿En qué medida usa sus conocimientos previos para que lo ayuden en la identificación de palabras?

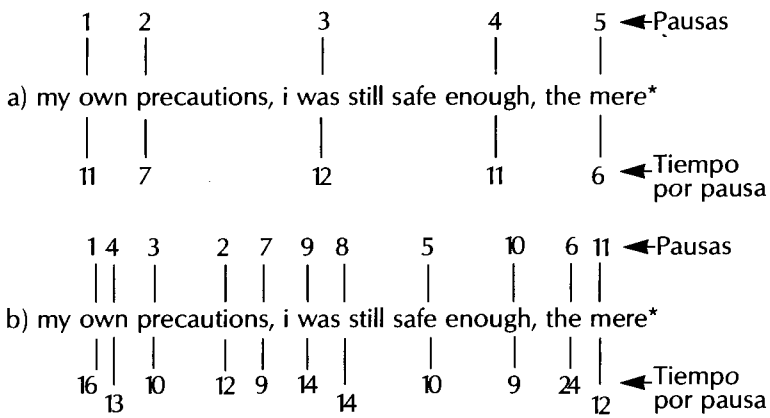
A menos que usted sea una persona muy poco común, no creo que sea capaz de usar sus propias introspecciones sobre lo que está haciendo para contestar semejantes preguntas, aunque pueda tener algunas opiniones sobre respuestas posibles. Esta inescrutabilidad del proceso de lectura hace que resulte muy difícil para cualquier lector experto ofrecer muchos consejos útiles a quien no posea la misma habilidad. Usualmente, un lector honesto dirá: "No sé cómo lo hago pero lo hago". Como psicólogos cognitivos necesitamos y, de hecho, podemos ofrecer, mejores respuestas.

## Los movimientos oculares en la lectura

El estudio de los movimientos oculares es útil para comenzar porque es la única conducta evidente que acompaña a casi toda actividad de lectura. Se puede medir la dirección de la mirada en cualquier momento, usando combinaciones de registros fotográficos y eléctricos y sin restringir los movimientos oculares normales. Semejantes estudios han demostrado que cuando vemos cualquier escena estática, contenga objetos o texto, el ojo se mueve en pequeños saltos rápidos (conocidos en inglés como *saccades*) con pausas más largas (llamadas *fijaciones*) en el medio. Cada *saccade* dura unos 50 ms (1/20 de segundo) y una *fijación* dura unos 200-250 ms (de 1/5 a 1/4 de segundo). La visión más útil tiene lugar durante las fijaciones, así que podemos pensar en la experiencia de la lectura como una serie de "instantáneas" de pequeñas porciones de texto.

La figura 4.1 muestra cómo se distribuyeron las fijaciones en un único renglón de texto leído por un experto lector y por otro que no lo era. Los números superiores muestran el orden de las fijaciones; los de abajo, sus duraciones. Esto ilustra varias características típicas: los buenos lectores hacen menos movimientos oculares regresivos, *saccades* más grandes y fijaciones más cortas. Cada una de estas características significa que un lector experto leerá el mismo pasaje más rápidamente que uno inexperto. ¿Podemos deducir, a partir de esto, que las diferencias en los patrones de los movimientos oculares son responsables de las diferencias en la habilidad de lectura? Desafortunadamente, no. Hacer menos y más cortas fijaciones sólo será posible si un lector puede manejar la información de las fijaciones con más eficacia. De lo contrario, el aumento de la velocidad de los movimientos oculares sólo llevará a la pérdida de la comprensión. Los sistemas de lectura veloz que se concentran en entrenar a las personas para que economice movimientos oculares (por ejemplo, derecho hacia abajo por el centro de la página, en lugar de hacerlo de lado a lado) raramente llevan a una mejoría duradera de la velocidad que no esté acompañada de una pérdida en la comprensión (Crowder, 1982; Gibson y Levin, 1975). Más adelante, en este mismo capítulo, nos ocuparemos de evidencias adicionales que muestran que los lectores expertos superan a los inexpertos en algunas tareas de lectura en las que no se requieren movimientos oculares. La economía de movimientos oculares es un síntoma de la habilidad de lectura y no su causa.

*Figura 4.1. Distribución de las fijaciones en un lector bueno y uno malo (según Tinker, 1965).*




*. "mis propias precauciones, yo todavía estaba bastante seguro, la mera"

## La lectura como aplicación de conocimientos

Cuando leemos, no estamos simplemente *adquiriendo* conocimientos nuevos de la página, sino que estamos utilizando muchas formas diferentes de conocimiento que ya poseemos. En los capítulos 2 y 3 vimos que todas las habilidades dependen de la movilización de los conocimientos adecuados. En la lectura, nuestro conocimiento del lenguaje nos permite reconocer palabras en las que no se han identificado todas las letras. Las "completamos" con lo que esperamos ver, como lo demuestra el bien documetando fenómeno del "error de los correctores de pruebas". Consiste en no advertir errores de copia u ortografía en el texto (como el que aparece en la oración anterior: documetando en vez de documentado). Por suerte, la redundancia del lenguaje hace que la mayoría de esos errores sean inofensivos. Obsérvese, por ejemplo, qué fácil resulta leer esta oración: lxs pxlxbrxs pxsxblxs pxxdxn xdxntxfxcxrsx fxcxlmxntx xxnqxx nx hxyx vxcxlxs prxsxntxs. En este caso, el contexto nos permite descubrir lo que el texto no ofrece.

Morton (1964) proporcionó una de las demostraciones clásicas del papel del contexto en la identificación de palabras. Mostraba a la gente oraciones a las que les faltaba una sola palabra (por ejemplo, "Entró a su casa y colgó su _____"). Luego exhibía muy brevemente la palabra faltante en un taquitoscopio (un simple equipo de laboratorio que presenta estímulos visuales para duraciones breves controladas con precisión). La duración de la exposición aumentaba gradualmente hasta que la palabra era reconocida. Morton halló que las palabras coherentes con el contexto precedente (es decir, por ejemplo, "abrigo") eran reconocidas más rápidamente que las palabras que no tenían nada que ver (por ejemplo, "barco").

Para que el contexto sea efectivo, una persona debe tener conocimiento de la situación a la que se está haciendo referencia. Su conocimiento debe llevar a esa persona a hacer deducciones sobre continuaciones posibles. Cuanto más sepa de aquello de lo que se está hablando, mejores serán sus inferencias y más eficiente su lectura. En igualdad de condiciones, es probable que el graduado en Psicología lea este libro más rápidamente que el estudiante principiante. En todas las etapas podrá usar los conocimientos que ya posee para darle sentido a lo que lee. Algunas partes del libro le resultarán familiares y le alcanzará con una lectura superficial de las mismas. Otros fragmentos usarán terminología o argumentos que ya tienen un significado bien articulado para él. Los buenos docentes de lectura saben esto y ayudan a sus alumnos eligiendo materiales de lectura iniciales que se relacio-

nen con situaciones que ya sean familiares para sus alumnos. De este modo, los alumnos pueden hacer anticipaciones inteligentes, a partir del contexto, que los ayudan a identificar las palabras que no reconocen; el éxito que obtienen les sirve para aumentar su confianza.

Una serie de estudios realizados por Sanford y Garrod (1981) ilustran muy bien el papel del conocimiento en la lectura. En estos estudios, los lectores veían historias simples, presentadas en una computadora, de a un renglón por vez. Cada lector marcaba el ritmo de su propia lectura apretando un botón cada vez que quería una nueva oración. De esta forma, los experimentadores podían medir el tiempo utilizado para leer cada oración. Veamos un ejemplo del tipo de historia que utilizaron:

Juan iba a la escuela.
El colectivo avanzaba muy lentamente.
No tenía ganas de dar lección de Matemática.
Después de todo, no era una tarea habitual para un celador.

La mayoría de los lectores experimenta una leve sensación de confusión y desorientación al leer la tercera y la cuarta frases. Esperamos que Juan sea un alumno, no un docente. Cuando hemos hecho los reajustes necesarios a la nueva información, nos enteramos de que, en realidad, es celador. Sanford y Garrod mostraron que este reajuste toma un tiempo mensurable. La tercera oración de la historia requirió más tiempo de lectura que otra ("No tenía ganas de tener clase de Matemática") coherente con la idea de que Juan era un alumno. La prosa normal bien construida no suele violar las expectativas de la forma en que decidieron hacerlo Sanford y Garrod. Sus resultados muestran que haciendo las inferencias adecuadas a partir del texto se puede luego leer oraciones con más rapidez, suponiendo que, en general, no se violen las expectativas.

## Conocimiento específico de la lectura

Las fuentes de conocimientos que acabamos de analizar no son específicas de la lectura. Tienen que ver con las formas en que operan el lenguaje (en general) y el mundo, y entran en juego cuando escuchamos hablar, cuando intentamos interpretar una escena visual y en muchas otras ocasiones. Pero también hay fuentes de conocimientos más específicas. Por ejemplo, los lectores llegan a adquirir conocimientos sobre la forma usual en que se unen las letras en su idioma, y resulta evidente que éste es un conocimien-

to específico de la alfabetización. Un lector hispanohablante sabe que es posible que TRENO sea una palabra, mientras que ONETR y OTENR no. Gibson, Pick, Osser y Hammond (1962) han demostrado que semejante conocimiento afecta la eficacia de la lectura. Presentaron estímulos hechos de cadenas de letras ante diferentes lectores y les pidieron que dijeran todas las letras que podían recordar. Los lectores podían recordar las letras de estímulos regulares (como TRENO) con una exposición a ellas más breve que la que necesitaban para los estímulos irregulares (como ONETR).

Otro campo que involucra conocimientos específicos de lectura es el uso de los espacios entre las palabras. Hochberg (1970) informa sobre un experimento en el que se pidió a niños de diverso nivel de lectura que leyeran textos normales y textos en los que los espacios entre palabras habían sido completados con "x", de la siguiente forma: Recurrióxaxdosxgruposxdexlectores,xbuenosxyxmalos.

Aunque los buenos lectores fueron en general más rápidos que los otros, su velocidad de lectura se redujo notablemente con las "x". En cambio, los malos lectores avanzaron más o menos a la misma velocidad (lenta) en ambos tipos de texto.

¿Por qué los espacios ayudan a los lectores expertos? Los datos proporcionados por O'Regan (1979) muestran que los lectores experimentados hacen movimientos oculares *sacádicos* más largos cuando la palabra situada a la derecha de la fijación es larga que cuando es corta. Tales lectores usan el espacio a la derecha de la palabra en la periferia de la visión como marcador para planificar el siguiente movimiento del ojo. Esto se hace, naturalmente, de una forma por completo automática. Un lector no es consciente de la medida constantemente fluctuante de sus movimientos *sacádicos* o de la información que rige a tales movimientos. Como sucede con muchas habilidades, nuestro sistema cognitivo parece capaz de aprender reglas complejas y adaptativas sin nuestra intervención consciente. Todo lo que tenemos que hacer es ofrecerle la posibilidad de construir estas reglas mediante la exposición frecuente a una amplia gama de situaciones relevantes.

El consejo más confiable (aunque aburrido) que se le puede dar a cualquiera que desee mejorar su lectura es que lea mucho, con frecuencia y durante mucho tiempo (igual que para cualquier otra habilidad). En un nivel de mayor detalle (y bastante menor confiabilidad) hay más que decir al respecto, especialmente si se está diseñando un programa de práctica para un individuo en particular. Quiero centrar el resto de este capítulo en tres tópicos, cada uno de los cuales proporciona información relevante para posibles

causas de problemas en la lectura: el papel del discurso en la lectura; la incapacidad de lectura y las estrategias particulares que adopta un individuo cuando lee para estudiar.

## El papel del discurso en la lectura

La difundida capacidad para leer en silencio no tiene una larga historia. En la Edad Media, mucha gente pensaba que la lectura silenciosa era un logro inexplicable y que quienes eran capaces de hacerla estaban poseídos por el demonio. Parecía increíble afirmar que las marcas negras sobre el papel podían tener significado si antes no se los convertía en los sonidos discursivos que representaban. Semejante punto de vista probablemente sería compartido por algunos lectores principiantes en la actualidad, especialmente aquellos entrenados dentro de una filosofía educativa denominada "fónica", según la cual uno aprende a leer aprendiendo los sonidos típicos de las letras y grupos de letras. Si se unen estos sonidos se producen sonidos de palabras, que luego "el oído de la mente" puede reconocer.

Otros aprenden con filosofías diferentes, a una de las cuales se la ha llamado "de mirar y decir". En este enfoque, los niños aprenden pares objeto-palabra mediante contacto visual directo, sin una enseñanza sistemática sobre correspondencias letras-sonidos. Así, las paredes de un aula "de mirar y decir" estarán cubiertas de ilustraciones de objetos familiares, cada uno junto con la palabra correspondiente. El reconocimiento rápido de palabras sigue al constante ensayo de estos pares.

La discusión entre quienes apoyan estas dos filosofías ha sido larga y feroz. Quienes defienden lo fónico sostienen que el otro método no proporciona al niño recursos para elaborar el sonido (y, consecuentemente, el significado) de palabras que no ha visto antes. Quienes defienden el "mirar y decir" afirman que la ortografía inglesa no tiene la regularidad suficiente como para que la simple traducción letra-sonido funcione de manera confiable[2]. Ambos puntos de vista son válidos. Consideremos, por ejemplo, la siguiente oración: "Lah mashoría dhe loz hautos tienem kuatro rruedaz"[3]. Aunque la oración no tiene ninguna palabra castellana, es de todos modos comprensi-

---

*2. Las vocales y los grupos vocálicos tienen, en inglés, diferentes pronunciaciones, que no los hacen fácilmente predecibles para el lector principiante.*

*3. El ejemplo usado en el original en inglés dice: "Moast xahs hav foor weels"; hemos intentado mantener la misma idea en castellano. (N. del T.)*

ble como representación de una oración en nuestra lengua. A menos que pudiéramos realizar traducciones letra-sonido, la comprensión sería imposible.

Por otra parte, sabemos que no siempre se pueden invocar reglas letra-sonido en la lectura. No hay nada en tales reglas que nos diga que *tough* y *bough* suenan de manera completamente diferente[4]. Hay otras palabras que son imposibles de leer por el camino letra-sonido, porque necesitamos establecer su significado antes de poder asignarle el sonido correcto. La palabra inglesa *tear* se pronunciará de manera diferente según se refiera a una gota (o una lágrima) o a una rotura (o desgarro).

## Leer con y sin sonido

Aunque es evidente que los lectores expertos no tienen que pronunciar en voz alta las palabras de la página, eso no significa necesariamente que la lectura basada en lo visual se apoye en la existencia interna del "sonido". De hecho, hay evidencias experimentales de lectura "sin sonido". Green y Shallice (1976) demostraron que el tiempo que se necesita para tomar una decisión sobre una palabra, basada en su significado, no depende del número de sílabas de la misma. Por ejemplo, al pedirles que presionaran un botón para demostrar que habían considerado que SALTAMONTES era un insecto, los lectores necesitaban el mismo tiempo que para la palabra MOSCA. Green y Shallice sostienen que si el sonido tuviera que ver, SALTAMONTES debería haber necesitado más tiempo por su mayor número de sílabas.

Como ocurre con muchas controversias en Psicología, la creciente complejidad de nuestro conocimiento ha demostrado que no podemos sacar una conclusión simple, ya sea en favor de la lectura basada en el sonido o en lo visual. La situación real es demasiado compleja como para poder hacerlo. En cuanto a los métodos de enseñanza, las investigaciones no han logrado demostrar ventajas claras en favor de ninguno de los dos métodos. En cambio, el factor más importante para el logro de una lectura experta es el docente (Williams, 1970). El buen docente probablemente usa una mezcla pragmática de enfoques, junto con la capacidad de simpatizar con sus alumnos y alentarlos. En lo que respecta a los lectores adultos expertos, el consenso generalizado afirma que estos disponen tanto de un proceso basado en el sonido como de uno basado en la visión directa para llegar al significado de las palabras, y que eligen el camino según la circunstancia. Además, actualmente se cree que no existe sólo un tipo de relación basado en el sonido si-

*4. En una transcripción fonética aproximada, la pronunciación de estas palabras sería: "taf" y "bau". (N. del T.)*

no que, más bien, puede haber tanto una forma abierta de compromiso discursivo asociada con movimientos de la laringe como una más encubierta de imaginación semejante a la del discurso, que operaría en un nivel totalmente subvocal. Esto se comprende mejor analizando los estudios de los efectos sobre la lectura de lo que se denomina *supresión articulatoria*.

Kleiman (1975) realizó una investigación en la que pedía a la gente que fuera repitiendo secuencias de dígitos que escuchaba a través de auriculares. Esto incluía la pronunciación en voz alta de los dígitos a medida que los iban oyendo. Pero semejante tarea restringe el sistema sonido-discurso. Mientras los sujetos repetían, se les pedía que emitieran diversos tipos de opinión sobre las palabras que se les presentaban visualmente. Una tarea consistía en que juzgaran si pares de palabras visualmente similares rimaban o no (por ejemplo: CASA-PASA, sí; ÚLTIMA-ULTIMA, no). Aunque la repetición de lo que oían disminuía la velocidad, los sujetos podían emitir opiniones correctas. Esto sugiere decididamente que no es necesario el sistema de sonido-discurso abierto para hacer juicios basados en el sonido. Parece existir algún proceso "interior" más profundamente asentado. Como observa Baddeley (1979): "Es muy evidente para mí que cuando leo, mi lectura va acompañada de algo parecido a una imagen auditiva o articulatoria de las palabras que se están procesando. Esto ocurre con o sin supresión articulatoria".

El hecho de que algunos aspectos de la lectura basados en el sonido puedan tener lugar sin el compromiso evidente del sistema discursivo no significa, sin embargo, que el sistema discursivo no tiene importancia alguna en la lectura. Veamos un estudio de Hardyck y Petrinovich (1970), que emplearon la biorretroalimentación (ver capítulo 3) para entrenar a los individuos en la supresión de la articulación. Descubrieron que la supresión articulatoria no tenía efectos adversos sobre la comprensión del material de lectura fácil pero que dificultaba la comprensión del material difícil. Podemos concluir que el discurso manifiesto (si es silencioso) ayuda a la comprensión de los pasajes difíciles. Probablemente, la mayoría de nosotros somos conscientes de esto por nuestra propia experiencia. Con textos difíciles, podemos tender a pronunciar las palabras en voz alta. Pero, ¿a qué se deberá esto? Crowder (1982) sugiere que semejante discurso resulta útil para mantener las primeras palabras de las oraciones complejas en la memoria de corto plazo, mientras se lee el resto de la oración. Esto es necesario porque uno frecuentemente necesita leer el final de una oración antes de que el significado de su primera parte quede claro (ver capítulos 5 y 10).

Este hallazgo también tiene un aspecto opuesto. Si la articulación no es particularmente útil para leer material fácil, podría resultar un obstáculo para la lectura más veloz. La tasa máxima de discurso es, quizá, de unas 4 o 5 palabras por segundo. Por lo tanto, la lectura con articulación no puede desarrollarse a una tasa mayor que unas 300 palabras por minuto. Sin embargo, hay considerables evidencias de que los lectores consumados pueden leer material simple a más del doble de esa velocidad sin pérdida de comprensión. Esto sólo se puede lograr si se suprime la articulación. Si usted tiene la impresión de que articula durante la lectura, puede convenirle practicar lectura sin vocalización. Una forma simple de obligarse a uno mismo a no vocalizar lo que lee consiste en repetir alguna palabra una y otra vez mientras se lee (por ejemplo, "doble-doble-doble", etc.).

## Dificultades en la lectura

Las dificultades de lectura pueden producirse por muchas razones. Algunos psicólogos creen que deberíamos hacer una distinción entre aquellos cuyas dificultades de lectura van acompañadas de toda otra gama de dificultades intelectuales y aquellos cuya incapacidad parece limitarse solamente a la lectura. Estos psicólogos sostienen que la palabra *disléxico* debería reservarse para la gente de este segundo tipo. En consecuencia, un disléxico es aquel individuo de capacidad intelectual promedio o superior al promedio que tiene dificultades de lectura. Curiosamente, alrededor del 80% de ellos son hombres. Otros lectores inexpertos, que tienden a no desempeñarse tan bien como el promedio del espectro intelectual, pueden ser tanto hombres como mujeres (Jorm, 1983).

Otros psicólogos desconfían de una categorización tan prolija y les resultaría difícil distinguir entre disléxicos y otro tipo de malos lectores. En todo caso, muchas investigaciones no hacen esta diferenciación al elegir sus sujetos. En este ensayo, usaré la palabra dislexia para referirme a cualquier incapacidad crónica de lectura que se resista a los métodos de enseñanza normales.

¿Hay algo específico que tienda a funcionar mal en la dislexia? Parece que a muchos disléxicos les cuesta elaborar los sonidos de las palabras que intentan leer: el camino hacia el significado basado en los sonidos. Así, por ejemplo, los disléxicos tienen un desempeño particularmente pobre en la pronunciación de palabras absurdas e inexistentes como BLEN, PORSUN y TELD. Un experimento realizado por Perfetti, Goldman y Hogaboam (1979) demostró que los disléxicos tienden a compensar esta dificultad apoyándose

mucho más en el contexto que los lectores normales. Se midió el tiempo que les llevaba a los niños pronunciar palabras impresas. En un caso, se les presentaron palabras sueltas; en un segundo caso, se les presentaron palabras incluidas en un cuento. La tabla 4.1 muestra los tiempos promedio de lectura por palabra en los dos grupos, en ambos casos. Aunque ambos grupos se beneficiaron con el contexto, los lectores inexpertos se beneficiaron más. Aparentemente, los lectores inexpertos intentan compensar su dificultad con el "camino del sonido" asignando mayor peso a la información contextual.

*Tabla 4.1. Tiempos promedio de lectura por palabra (ms)*

| | Sin contexto | Con contexto |
|---|---|---|
| Buenos lectores | 786 | 665 |
| Malos lectores | 1069 | 751 |

El déficit de lectura en el método basado en el sonido formaría parte de una dificultad mayor que la que tienen los disléxicos para utilizar la memoria de corto plazo (u operativa), que, según se sabe, es la responsable de retener información verbal de manera acústica (ver capítulo 5). Parecen ser menos capaces de recordar la redacción precisa (en oposición a la esencia) de cuentos que acaban de oír, y la duración de su memoria de corto plazo suele resultar más baja que el promedio. En la sección previa afirmamos que la capacidad de retener material verbal en la memoria de corto plazo ayuda a la comprensión del material difícil. Por lo tanto, un déficit en la capacidad de usar memoria de corto plazo bien puede llevar a dificultades en la comprensión.

¿Cuál es el origen de la dislexia? Hay ciertos indicios sobre la existencia de un componente genético. La desigual distribución por sexo es típica de ciertos rasgos que se transmiten genéticamente. La condición también tiende a darse en familias, y la incidencia conjunta de la dislexia en los gemelos es mayor que en los mellizos (Hermann, 1959). Aunque no es posible descartar las explicaciones ambientales para estos hechos, es significativo que los niños disléxicos tengan con frecuencia hermanos y hermanas que son muy buenos lectores, pese a crecer en el mismo medio.

La cuestión de importancia práctica más apremiante es: ¿puede ayudarse a los disléxicos a mejorar su lectura? Y, si la respuesta es afirmativa, ¿cómo? La respuesta a la primera parte de la pregunta parece ser un cauteloso "sí". Aunque pocos son capaces de superar sus dificultades por completo, algunos

programas de entrenamiento ayudan a los disléxicos a conseguir mejorías duraderas cuando se los compara con grupos de control que no han recibido ese entrenamiento. El enfoque más exitoso parece ser el que ofrece al lector la posibilidad de practicar precisamente aquellas habilidades en las que tiene dificultades, es decir los procesos basados en los sonidos (Naidoo, 1970). Pero algunos investigadores han sostenido que también se puede tener éxito alentando a los lectores inexpertos a utilizar mejor sus puntos fuertes (como su capacidad para aprovechar el contexto). Desafortunadamente, no hay estudios de evaluación suficientes para tener un juicio definitivo sobre este punto.

Es difícil obtener conocimiento científico confiable en el campo de la dislexia porque el docente/investigador se enfrenta a un dilema: ¿debe ofrecerse al alumno (en un juicio considerado pero subjetivo) el mejor enfoque y el más factible para el docente? ¿O debe ofrecérsele algo que no esté bien probado (o que no tenga, de hecho, "control" comparativo alguno) con el objeto de avanzar en los conocimientos científicos? No se puede culpar a quienes elijan la primera opción. Los dilemas de este tipo son comunes en muchas ramas de la Psicología. Explican en parte por qué, pese a todas las investigaciones hay, en este campo, menos respuestas definitivas que las que nos gustaría tener.

## Las estrategias de lectura y el estudio

Supongamos que uno es un lector experimentado, que lee razonablemente rápido y sin vocalización excesiva, tiene un vocabulario amplio y conocimientos generales buenos. ¿Hay algún motivo por el que podría tener dificultades con las exigencias que plantea la lectura para el estudio? Por desgracia, sí. Smith (1967) les pidió a lectores-alumnos expertos e inexpertos que leyeran textos para obtener: a) impresiones generales, b) detalles. Descubrió que los buenos lectores ajustaban sus procedimientos de acuerdo con las instrucciones. Cuando leían en busca de detalles, releían cosas del tipo de nombres y fechas. Al leer para obtener impresiones generales, se detenían a evaluar ideas al terminar los párrafos y, si releían, leían secciones enteras en vez de puntos individuales. Los lectores inexpertos, en cambio, recurrían a procedimientos mucho más parecidos en los dos casos. (Este es un ejemplo de la importancia de las habilidades metacognitivas - ver capítulo 12).

Una clave para la lectura eficiente parece ser la noción de economía. Los alumnos de bajo nivel de rendimiento pierden tiempo leyendo material que no es esencial en un nivel de detalle demasiado elevado. En cambio, es

muy probable que los lectores expertos recorran el material superficialmente, notando lo que es irrelevante para sus propósitos y aquello que deben releer para tener más detalles. Semejante conducta presupone una clara idea sobre cuál es el propósito de la lectura. Así, cuando el docente pide: "Lea el capítulo 5", el lector experto debe saber correctamente el propósito de lectura del capítulo 5 y realizar ese modo de lectura eficiente.

En lo que concierne a la lectura superficial, es conveniente hacer una advertencia. Si se pide a alguien que lea algo "sólo para saber las ideas básicas", la lectura superficial no hará el trabajo por sí sola. No es posible seguir un argumento destacando una palabra o una frase especial cada tanto. Antes bien, la lectura superficial sirve más como "análisis" preliminar, particularmente útil cuando uno tiene en mente un conjunto específico de propósitos.

## Conclusiones

Éste es, quizá, el punto adecuado para terminar nuestro análisis de la lectura. Aunque la cuestión de la evaluación del resultado de la lectura es crucial para el estudio efectivo, se lo encarará más adelante en el contexto más amplio de las habilidades de aprendizaje (capítulo 12). Apenas hemos recorrido superficialmente la vasta literatura existente sobre lectura, pero espero que por lo menos cinco características importantes de la lectura experta hayan quedado claras:

1) La lectura involucra una compleja interacción del conocimiento en muchos niveles, desde el conocimiento del mundo hasta el de las reglas ortográficas.

2) Este conocimiento está ampliamente automatizado y se aplica sin tener conciencia de ello.

3) No hay sólo un modo de leer sino una variedad de procesos y estrategias entre las que los lectores expertos pueden elegir según su propósito y las circunstancias.

d) Las dificultades de lectura pueden asociarse, con frecuencia, a deficiencias en los subcomponentes específicos del proceso total, como la traducción de símbolo a sonido.

e) La mejora de las habilidades de lectura en los adultos puede lograrse mediante la eliminación de la vocalización innecesaria y mediante una mayor conciencia de las posibilidades de adaptar la propia estrategia de lectura a objetivos de lectura formulados específicamente.

## Lecturas adicionales

CROWDER, R.G. (1982) *The Psychology of Reading: an Introduction.* Nueva York: Oxford University Press.

Clara y concisa introducción al tema; particularmente buena en lo que hace a conocimientos específicos de la lectura.

SANFORD, A.J. y GARROD, S.C. (1981) *Understanding Written Language*: Explorations in comprehension beyond the sentence. Chichester: Wiley.

Complementa al libro de Crowder, ya que se ocupa fundamentalmente del papel de los conocimientos contextuales.

BRYANT, P.E. y BRADLEY, L. (1985) *Children's Reading Problems.* Oxford: Blackwell.

Evaluación crítica actualizada de las investigaciones sobre la dislexia y problemas afines.

# *Sección 2*

## *Las habilidades de la memoria*

# 5

# La naturaleza de la memoria

*Angus Gellatly*

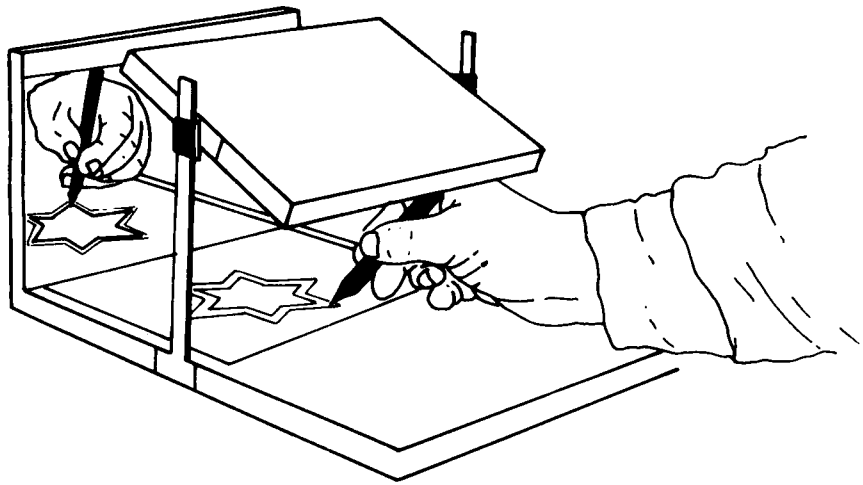
## Una "buena" memoria

Harry Lorayne, memorizador profesional, solía hacer lo siguiente: mientras las personas del público iban entrando al auditorio en el que realizaba su función, les decía sus nombres; luego, cuando ya había cientos de personas sentadas, Lorayne señalaba a cada uno y recordaba con precisión cómo se llamaba. A diferencia de la mayoría de nosotros, Lorayne no tenía problemas para asociar nombres con rostros. En realidad, su hazaña memorística es tan sorprendente que podríamos suponer la existencia de alguna facultad innata especial que explicara su desempeño y decir que estaba dotado de una muy "buena" memoria. Pero semejante explicación no es, por supuesto, explicación alguna (ver capítulo 3). El propio Lorayne insiste en que cualquiera que esté preparado para hacer el esfuerzo puede adquirir la misma habilidad. ¿Cómo se hace y qué nos dice sobre la naturaleza de la memoria y las habilidades memorísticas? En el presente capítulo ofreceremos respuestas a estas preguntas. Pero antes de que tales respuestas puedan tener sentido, es necesario saber algo sobre cómo estudian la memoria los psicólogos. Para ello, también será útil considerar un caso muy diferente del de Lorayne, en el que el rendimiento de la memoria es muy inferior al normal.

## Una "mala" memoria

Este ejemplo se ocupará del caso de un paciente típico, a quien nos referiremos como K., que sufre de una enfermedad conocida como síndrome de Korsakoff. Esta enfermedad se relaciona con el alcoholismo prolongado, que lleva a daños en el sistema límbico cerebral. Los pacientes que sufren del mal de Korsakoff muestran diversos síntomas, según su grado de deterioro cerebral y sus personalidades individuales. De todos modos, hay varios síntomas típicos que pueden atribuirse a nuestro paciente ideal, K.

*Figura 5.1. Aparato de dibujo en espejo*



Imaginemos que K. está sentado en la guardia del hospital cuando viene a buscarlo una enfermera a quien conoce bien. Sin embargo, K. no demuestra reconocerla en absoluto. Cuando la enfermera le pregunta si ha desayunado, K. responde cautelosamente que no está seguro. Luego, lo llevan a un laboratorio en otra parte del hospital y lo sientan frente a un aparato de dibujo en espejo (ver figura 5.1). K. niega estar familiarizado con el laboratorio o con el aparato, aunque ha estado en este lugar cada uno de los últimos días para practicar dibujo en espejo. Se le explica que tendrá que usar un lápiz para delinear una forma mientras ve su mano en un espejo. (En esta tarea, la persona recibe una retroalimentación invertida especular de la posición y el movimiento de sus manos, y le resulta difícil controlar lo que está haciendo; después de unas pocas sesiones de práctica, adquiere destreza y ya no experimenta problemas.) K. niega con la cabeza y afirma que le parece muy difícil; inmediatamente, empieza a delinear la figura con precisión. Se desempeña como la persona adiestrada que es, pero no recuerda haber intentado realizar esta tarea previamente. Su memoria sobre cómo hacer dibujos en espejo se ha disociado de su memoria de haberlo hecho previamente.

## Memorización automática y esforzada

Lorayne y K. se ubican en los extremos de lo que se denomina "buena" y "mala" memoria, pero la diferencia entre ellos no es meramente una diferencia de grado. Son, respectivamente, bueno y malo recordando diferentes tipos de información. La habilidad de Lorayne radica en haber dominado una técnica para fijar en la memoria cosas que la mayoría de las personas suelen olvidar, mientras que la deficiencia más soprendente de K. es con cosas que los demás recuerdan automáticamente y sin esfuerzo alguno.

Hasher y Zacks (1979) han sugerido la existencia de un continuo desde la información que se codifica en la memoria automáticamente, sin necesidad de atención consciente, hasta aquella que sólo se codifica con esfuerzo. Lo que sugieren es una aplicación al área de la memoria de las ideas generales de habilidad y automaticidad analizadas en el capítulo 2. El tipo de información que probablemente se codifique de manera automática en la memoria tiende a ser de naturaleza elusiva (cosas como olores, sabores y climas emocionales). Hasher y Zacks sugieren que el contexto espacio-temporal de los acontecimientos también puede codificarse automáticamente. Esta es la información que le permite a un organismo orientarse en el tiempo y el espacio. Sin memoria para la información espacio-temporal, el suave discurrir de la conducta natural se vería interrumpido, porque semejante conducta exige que se reconozcan los contextos familiares y se responda a ellos inconscientemente. Los errores en este procesamiento inconsciente del contexto pueden llevar al tipo de conducta distraída de la que hablaba John Sloboda en el capítulo 2. La codificación y el subsiguiente reconocimiento del contexto espacio-temporal también son importantes para ese sentido de familiaridad que no poseen, trágicamente, los enfermos de Korsakoff. Incluso los supuestamente normales no se sienten seguros de haber reconocido un rostro a menos que puedan "ubicarlo", identificando tanto el lugar como la ocasión en que lo encontraron antes.

A diferencia de la memorización y el reconocimiento automáticos, que son aspectos constantes y discernibles del mundo natural, la memoria esforzada es algo extraño. es la característica de un animal que actúa sobre su medio así como también dentro de él. La memoria esforzada requiere la invención de técnicas artificiales para complementar las habilidades innatas de la codificación automática. Esas técnicas constituyen el interés central de este capítulo y de los tres siguientes.

## Más sobre la jerga de la memoria

Harry Lorayne es un *mnemotécnico* y las técnicas que usa para lograr sus sorprendentes hazañas se denominan mnemotécnicas. La idea de mnemotecnia hace referencia a cualquier recurso para ayudar a la memoria; un ejemplo común en nuestra cultura es el famoso "Treinta días trae septiembre...", que hace uso de las propiedades mnemotécnicas de la rima y el ritmo. Todos nosotros somos mnemotécnicos en varios sentidos, pero pocos pueden alcanzar el nivel de profesionalismo de Lorayne.

K., por su parte, es un amnésico. Sufre de *amnesia*, término universal para todo tipo de pérdida de la memoria. Todos sufren, en mayor o menor medida, de *amnesia infantil*, la incapacidad de recordar los primeros años de la vida, pero la palabra amnésico se reserva para aquellos que tienen una forma anormal de pérdida de la memoria. Hay dos categorías mayores de amnesia. En la amnesia retrógrada, no se recuerdan los acontecimientos anteriores al trauma o daño relevante. En la amnesia *anterógrada*, lo que no se recuerda son los acontecimientos posteriores al trauma. Frecuentemente, ambos tipos de amnesia coexisten pero en K., como en la mayoría de los pacientes que sufren de Korsakoff, la pérdida anterógrada es mucho peor que la retrógrada (Kolb y Whishaw, 1985). K. puede recordar mucho sobre los acontecimientos previos a que sufriera el daño cerebral, pero bastante poco de lo que le ocurrió a él y al mundo desde entonces.

Esta forma típica de pérdida lleva a otras distinciones. Muchos psicólogos piensan que la memoria se da en tres etapas. La nueva información debe ser ingresada (o codificada) en la memoria, luego debe almacenarse en algún lapso de retención y, finalmente, para poder ser de utilidad, debe estar sujeta a recuperación cuando se la requiera. En los casos de amnesia es posible preguntarse en cuál de los tres estadios tiene lugar el déficit. Como K. sigue siendo capaz de retener y recuperar recuerdos de acontecimientos distantes, parecería que es en la codificación de la nueva información donde tienen problemas los enfermos de Korsakoff. La noción de codificación se irá elaborando a lo largo de este capítulo, pero debemos hacer otra diferenciación relacionada con la hipótesis de que la amnesia anterógrada es consecuencia de un problema de codificación.

Si se le lee a K. media docena de palabras, puede repetirlas normalmente, pero a los cinco minutos ya no recordará ninguna de ellas ni será capaz de recordar que se le hubiera leído algo. Su memoria de corto plazo no tiene problemas pero luego cae drásticamente. Como el protagonista de *El castillo*, (la

novela de Kafka), K. puede darle sentido al momento presente pero está aislado de la continuidad y los patrones de acontecimientos de largo plazo. Esta observación nos lleva a la distinción que hacen los psicólogos entre la *memoria de corto plazo* (MCP) y la *memoria de largo plazo* (MLP). No se puede definir con precisión ninguna de estas dos expresiones, pero la recuperación o el reconocimiento de información más de unos minutos después de que se presenta suele tomarse como prueba de almacenamiento de MLP. La hipótesis del déficit de capacidad de codificación sugiere que aunque K. puede codificar la información en la MCP, es deficiente en la codificación en la MLP. Este caso se opone al de Lorayne, que tiene un alto grado de eficiencia en la codificación en la MLP, al menos en lo que concierne a nombres.

## La naturaleza de la memoria de corto plazo

Supongamos que uno ve que dos hombres salen corriendo de un banco y se suben a un auto que inmediatamente huye a toda velocidad. Si uno es rápido, puede tomar nota mental de la patente del auto. Pero, ¿y luego? Si no se coloca ese número de patente en alguna forma de almacenamiento más duradera, tendrá que seguir repitiéndoselo constantemente o lo olvidará. El número de patente está en la memoria de corto plazo, que globalmente corresponde a aquello en lo que estamos pensando, o de lo que somos conscientes, en algún momento en particular. Como se la suele entender, la memoria de corto plazo (MCP) tiene una capacidad limitada; puede retener sólo un número fijo de ítem, o agrupamientos, por vez. Miller (1956) estimó que esta capacidad para los adultos era de siete ± dos agrupamientos, y la naturaleza de un agrupamiento depende de la familiaridad de la persona con la materia en cuestión. Para alguien que no sepa castellano, A-M-I-G-O-S podría consistir en seis agrupamientos, mientras que para un lector experto de castellano sólo sería uno (ver capítulo 2).

El método usual para medir la MCP consiste en encontrar el mayor número de palabras, letras o dígitos que una persona puede recordar con precisión en el 50% de sus intentos. Esta cifra es la duración de la memoria de una persona. Como la mayoría de los adultos tiene una cifra cercana a los siete ítem, se suele considerar a la MCP una especie de caja con una cantidad fija de ranuras. Sin embargo, los niños de 3 años tienen una capacidad medida de sólo tres ítem y el nivel adulto no se alcanza hasta los 12 años de edad, aproximadamente. ¿Significa esto que la cantidad de ranuras de la MCP au-

menta con la edad? Casi seguro que no. El aumento de la cifra con la edad es precisamente paralelo al aumento de la tasa de velocidad a la que pueden hablar los niños (Hulme, Thompson, Muir y Lawrence, 1984). Esto sugiere que un test de duración no mide el número de ranuras en un hipotético espacio de almacenamiento sino la capacidad de la persona para hacer el test. Para la mayoría de las personas, el nivel de capacidad en las pruebas de duración de la MCP está determinado por su habilidad para hablar rápido y por lo tanto por su tasa máxima de ensayo subvocal. (En el capítulo 7 se describen algunos otros medios por los que los individuos pueden lograr duraciones muy largas de su MCP.)

El punto antedicho señala un obstáculo mayor para la investigación de las aptitudes cognitivas. Todas las pruebas diseñadas para detectar y evaluar una aptitud específica desafían al sujeto a desarrollar nuevas habilidades cognitivas para mejorar el rendimiento en esa prueba, a menudo usando medios no anticipados por el experimentador. Inversamente, el bajo rendimiento puede no reflejar falta de aptitud sino, como se verá en el capítulo 6, falta de práctica en la selección y puesta en operación de la mejor estrategia para encarar la prueba.

## Memoria operativa

Seguramente ya estará claro que la naturaleza precisa de la memoria de corto plazo es problemática. Nadie sabe bien cómo se caracterizaría mejor, pero es difícil de negar la existencia de un almacenamiento de corto plazo de poca capacidad. Uno puede apreciarlo por sí mismo si piensa en cómo haría un cálculo mental, como multiplicar 36 x 36. Hay que realizar toda una serie de operaciones, incluyendo el almacenamiento y la subsiguiente utilización de los resultados intermedios. Sin un almacenamiento de corto plazo de alguna naturaleza, ¿cómo podría lograrse? Y si ese almacenamiento no tiene capacidad limitada, ¿por qué será tan difícil acordarse de todo cuando se multiplican números mayores? No es sólo en la aritmética mental en donde se encuentran estas necesidades de almacenamiento de corto plazo. Todo tipo de actividades cognitivas involucra almacenamiento de corto plazo de la información.

Debido a los tipos de dificultades recién delineados, Baddeley y Hitch (1974; Hitch, 1980) sostuvieron que la MCP debe ser más compleja de lo que previamente se creía. Más que un almacenamiento unitario, sugirieron la posible existencia de un sistema de memoria de corto plazo con varios componen-

tes. A este sistema lo denominaron memoria operativa. Sobre la base de experimentos que no tenemos espacio para detallar aquí, determinaron que la memoria operativa tiene por lo menos cuatro componentes: un sistema de salida del discurso, que juega un papel fundamental en las pruebas de duración de la memoria de corto plazo; un sistema de entrada de la audición que mantiene un registro del discurso recientemente escuchado, para que el comienzo de una oración esté aún disponible cuando se llega a su fin; un sistema de entrada visual para mantener impresiones visuales en el corto plazo; y un sistema de control que utilice los demás componentes. El modelo de memoria operativa es obviamente bastante complicado y los psicólogos aún lo están evaluando. Lo importante es que, sea cual fuere su complejidad, parece existir alguna clase de sistema para el almacenamiento de la información en el corto plazo.

## La naturaleza de la memoria de largo plazo

Ya que la MCP resulta una entidad no unitaria, no debería sorprendernos que sucediera lo mismo con la memoria de largo plazo (MLP). De hecho, ya hemos visto que para K. la capacidad de aprender, y por lo tanto de recordar, cómo ejecutar una tarea motriz está disociada del recuerdo de haber enfrentado ya antes la tarea. Nuestro análisis de las subclasificaciones de la MLP puede empezar por esta observación.

### Memoria motriz

La memoria motriz subyace a todo tipo de habilidades motrices, desde escribir a máquina hasta pintar paredes. Su naturaleza y organización se han investigado más profundamente en términos de adquisición de habilidades motrices, como las descritas en el capítulo 2. La característica más sorprendente de la memoria motriz es su durabilidad. Una vez aprendidas, habilidades como andar en bicicleta o conducir un automóvil no se olvidan. Por este motivo, a veces se ha pensado que la memoria motriz difiere fundamentalmente de otros tipos de memoria de largo plazo. Sin embargo, lo notable de la mayoría de las habilidades motrices es que se practican mucho (se sobreaprenden); éste es el motivo por el que son tan resistentes al olvido.

Cuando una persona llega finalmente a dominar una bicicleta o un automóvil, no suele dejar la actividad. Por el contrario, con frecuencia la reali-

za repetidamente, de manera que se sobreaprende. No siempre sucede lo mismo con tipos de aprendizaje más cognitivos. En este caso, se suele practicar la materia hasta algún nivel, por ejemplo la sensación de que se la puede sacar a relucir en el examen del día siguiente, y de ahí en más se la elimina de la mente. Cuando meses después se la requiere una vez más, no está disponible. Pero éste no tiene por qué ser el destino de la materia cognitiva. El sobreaprendizaje puede darse tanto en las habilidades motrices como en las cognitivas. Habilidades cognitivas sobreaprendidas como entender un discurso, reconocer la identidad de rostros humanos o multiplicar cifras largas no se olvidan.

Debido a que las habilidades cognitivas y las habilidades motrices sobreaprendidas comparten las características de durabilidad, agrupamiento, fluidez y demás, se las puede unir como ejemplos de conocimiento procesal o memoria procesal (Ericsson, 1985). Ambos son ejemplos de *recordar cómo más que de recordar qué*. En este sentido, contrastan con el siguiente tipo de almacenamiento de largo plazo.

## Memoria episódica

Tulving (1972, 1984) fue quien introdujo la expresión "memoria episódica" para referirse a la memoria para acontecimientos (o episodios) de la vida cotidiana. Se trata de una memoria autobiográfica, ya que es el registro de nuestra propia vida y de aquellas cosas de las que hemos tomado conciencia. En el caso de K., parecería que no han existido (o que han existido muy pocos) agregados a este registro desde que sufrió el daño cerebral. La memoria episódica es aquello a lo que probablemente se refiere la mayoría de la gente cuando habla de "buena/mala" memoria, o lo que piensan ante la expresión MLP (memoria de largo plazo). Cuando dijimos que Lorayne estaba especialmente entrenado en la codificación de información en la MLP, estábamos refiriéndonos a la MLP episódica.

Una característica de la memoria episódica es que la información (codificada automáticamente o con esfuerzo) se almacena según el contexto espacio-temporal. Esto ya se ha dicho antes. Otra característica de la información en la memoria episódica es que generalmente se codifica según su significado más que según su forma superficial. El registro de un episodio durante el día se hará de acuerdo con la esencia de lo que se dijo (o se dejó de decir); por lo tanto, no será equivalente a una grabación. Esto se opone a la

MCP, en la que la retención precisa se juzga según el recuerdo al pie de la letra. De hecho, la información en la memoria episódica suele ser más *reconstruida* que recordada. Los relatos de acontecimientos pasados se organizan sobre la base de cómo se los interpretó en su momento, cómo se los ha llegado a interpretar desde entonces y usando como sistema los "libretos" (o "guiones") que todos tenemos sobre la forma en que se suelen desarrollar los acontecimientos. Esto no significa afirmar que la gente no puede hacer citas textuales de poemas o discursos, pero eso sólo puede construirse sobre la práctica intensa, y son ejercicios de memoria procesal. Como lo demuestra el dibujo especular de K., las memorias procesal y episódica pueden separarse.

## Memoria semántica

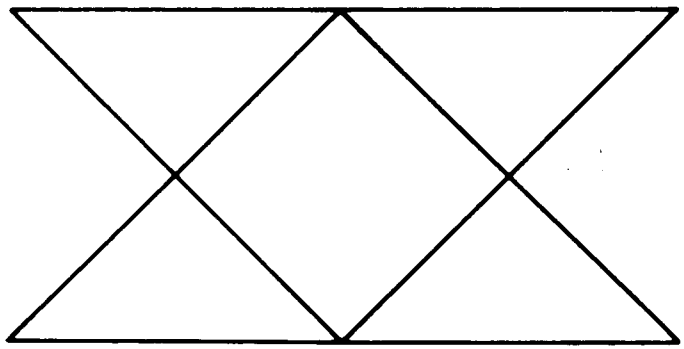
Tulving (1972) contrastó la memoria episódica con la semántica. Mientras que la memoria episódica es de naturaleza personal y contextual, la memoria semántica es conceptual y abstracta. Mantiene muchos tipos de información en diferentes clases de relaciones: el significado de las palabras, las reglas gramaticales, los hechos empíricos (como que el agua encuentra su propio nivel), los "libretos" de los que hablábamos antes, los nombres de los meses del año, etc.

Todos los usos de la memoria episódica incluyen un componente semántico, y el rendimiento en una tarea episódica está determinado por la riqueza de las estructuras semánticas del campo de conocimiento en cuestión (Ericsson, 1985). Para expresarlo con claridad, a la mayoría de las personas le resulta más fácil recordar una lista de palabras con sentido que una lista de palabras sin sentido (como MAFER), de la misma forma en que los maestros de ajedrez recuerdan configuraciones de partidas reales mejor que configuraciones azarosas (capítulo 2). Sin embargo, las memorias episódica y semántica pueden disociarse, como en la amnesia retrógrada y en la anterógrada. En el caso de K., el paciente recuerda cómo hablar y pensar, lo cual debe involucrar memoria semántica, pero su memoria episódica es gravemente deficiente.

Baddeley (1984) ha cuestionado la forma en que se hace la distinción entre memoria episódica y memoria semántica. Sostiene que la información de la memoria semántica es conocimiento *procesal* sobre el que se basan habilidades cognitivas tales como hablar, escuchar y pensar. El conocimiento procesal puede tener la forma de reglas de producción (como se describe en el capítulo 2) que el individuo no puede establecer en palabras. En contras-

te, Baddeley afirma que la información en la memoria episódica es conocimiento declarativo. El conocimiento declarativo es aquel que puede verbalizarse; es conocimiento de que la cosa es de determinada manera más que conocimiento de cómo hacer algo. Actualmente, no está claro si es conveniente pensar que la MLP tiene tres componentes (motriz, semántico y episódico) o que es dicotómica (memoria procesal y memoria declarativa)

*Figura 5.2. Una figura ambigua*



## Codificación y recuperación

La codificación de la información es la forma en que ésta se retiene. La influencia de la codificación en la memoria subsiguiente queda demostrada examinando la Figura 5.2 y luego –más adelante– la figura 5.3. El reconocimiento de la forma en 5.3A tiende a ser mucho más rápido que el de la forma de 5.3B. Aunque ambas formas estaban presentes en el original, 5.3A es más comparable con la codificación o interpretación usual de 5.2. Como se dijo antes, la codificación de palabras en la MCP tiende a ser acústica, mientras que la MLP tiende a codificar según significado, semánticamente. Craik y Lockhart (1972) teorizaron que quizá no hubiera diferencia entre la MCP y la MLP excepto en las formas de codificación. Supusieron que la durabilidad de la memoria para acontecimientos, palabras o rostros era una función de la "profundidad de procesamiento" de los ítem originalmente recibidos y que

un código acústico era menos profundo que uno semántico. Craik y Tulving (1975) hicieron preguntas sobre: a) la estructura física de una palabra ("¿Está escrita con mayúscula?"); b) el sonido de una palabra ("¿Rima con carnada?"); o c) el significado de una palabra ("¿Podría usarse en la oración: 'Encontró un _____ en la calle'?"). Luego, dieron a los sujetos un test de memoria (no anticipado) para todas las palabras que habían visto, y el rendimiento mejoró en las tres situaciones según la profundidad de procesamiento supuesta, apoyando así la hipótesis original (ver capítulo 2).

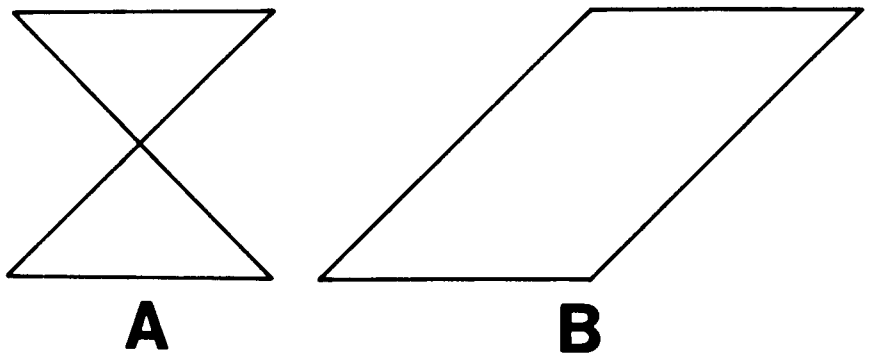
Craik y Tulving (1975) demostraron también que si todas las palabras presentadas eran codificadas semánticamente, luego el recuerdo dependía del nivel de elaboración de esa codificación. Cuando se pidió a los sujetos que juzgaran si la palabra *reloj* encajaba en una oración muy elaborada ("El anciano tropezó y se le cayó su valioso _____") la recordaron mejor que cuando se les preguntó si *reloj* encajaba en una menos elaborada ("Se le cayó el _____"). Cuanto mayor es la red de asociaciones en la que se inserta la palabra en la codificación, mejor es la memoria para ella. De modo que tanto la profundidad como la elaboración son importantes.

Pero el rendimiento de la memoria no depende sólo de la codificación. También depende del contexto de recuperación, de cuántas son las pistas de recuperación que se tienen para ayudar a la memoria. Smith (1979) presentó en una investigación listas de palabras y luego probó su recuerdo un día después, en la misma habitación y en otra diferente. El recuerdo fue mucho mejor en la misma habitación, lo cual demuestra la importancia de las ayudas externas (ver capítulo 2). Las ayudas internas también son importantes. Eich (1980) analiza evidencias para demostrar que el recuerdo es mejor cuando el humor o la situación (por ejemplo, de ebriedad) es igual al momento en que se codificó. Más aun, incluso el valor del procesamiento más profundo depende del tipo de test de memoria que se emplee. Morris, Bransford y Franks (1977) repitieron el descubrimiento de Craik y Tulving de que el procesamiento más profundo proporcionaba un mejor rendimiento en una prueba de reconocimiento directo. Sin embargo, en una prueba en la que los sujetos debían elegir palabras que rimaran con palabras previamente presentadas, se obtuvieron mejores resultados en los casos en los que la codificación se había hecho según la rima y no según el significado.

En síntesis: la forma de codificación ideal de la información dependerá de para qué se la va a usar y, por lo tanto, de las condiciones en las que se la habrá de recuperar. No obstante, es globalmente cierto que probablemente una codificación más profunda y elaborada resulte más eficiente. Po-

demos ver cómo esta generalización se sostiene en el caso de la sorprendente memoria para los nombres de Harry Lorayne.

*Figuras 5.3A y 5.3B: componentes alternativos de la figura 5.2.*




## Un sistema para recordar nombres

En uno de sus libros (Lorayne y Lucas, 1974), Lorayne detalla su sistema para recordar nombres. Como muchos sistemas mnemotécnicos, el suyo se basa en el empleo de imaginería visual y se lo puede desglosar en etapas (ver capítulo 8). La etapa inicial es muy directa y tiene que ver con asegurarse, en primer lugar, de que se ha oído el nombre que se quiere recordar. Muchas veces, no es que se olvide el nombre de alguien que a uno le acaban de presentar sino que nunca se lo registró. Así que la primera regla de oro es asegurarse de escuchar correctamente el nombre, incluso si ello implica pedir que lo repitan o repetirlo uno mismo, como cuando se dice: "Encantado de conocerlo, Sr. Acuña".

La segunda etapa es la sustitución del nombre por una palabra (o palabras) de sonido similar y fácil de recordar. Así, "Acuña" podría convertirse en "La-cuña". Luego, en la tercera etapa, es necesario examinar la cara de la persona y elegir su característica más distintiva; por ejemplo, sus grandes orejas. Finalmente, en la cuarta etapa, debe vincularse la característica elegida

con la palabra sustituta mediante una imagen tan clara como sea posible. En este caso, se podría imaginar la cara del Sr. Acuña con dos inmensas cuñas en lugar de orejas. Una vez hecho esto, la próxima vez que vea al Sr. Acuña y vuelva a notar sus orejas le vendrá la imagen a la mente y podrá recordar su nombre.

El de Lorayne es un sistema simple que cualquiera puede emplear. Pero manipular efectivamente el sistema exige más que una familiaridad superficial con las etapas que lo componen. La clave aquí (como en todo) radica en la práctica intensa. El libro de Lorayne incluye una lista de palabras sustitutas para casi seiscientos de los apellidos más comunes en los Estados Unidos. Si se recuerdan los pares, de modo que cada nombre sugiera inmediatamente su sustituto asociado, se habrá automatizado una de las etapas del proceso. Del mismo modo, la experiencia facilitará la generación de imágenes para vincular palabras sustitutas con características faciales.

El sistema de Lorayne para recordar nombres ofrece un ejemplo más que pertinente de habilidad memorística. Primero, el sistema mismo es simple y fácil de comprender. Segundo, la necesidad de recordar nombres es común y, sin embargo, pocos reflexionan sobre las exigencias de la tarea y cómo pueden satisfacerse. Tercero, recordar nombres no es sólo una habilidad cognitiva sino una habilidad cognitiva socialmente importante. Se dice que uno de los últimos presidentes de los Estados Unidos debió su triunfo electoral menos a su perspicacia política que a la lealtad que inspiraba por recordar los nombres de los trabajadores humildes de su partido. Aquí tenemos, entonces, una habilidad de memoria simple e importante en la que casi todos podemos mejorar nuestro rendimiento. Tristemente, es también una habilidad en la que nuestro paciente, K., tiene graves deficiencias.

## Amnesia anterógrada

Se han intentado diversas formas de explicar la amnesia anterógrada. Muchos incluyen la idea de que hay un déficit en las aptitudes de codificación de los pacientes amnésicos (Mayes y Meudell, 1983; Squire y Cohen, 1984). En diferentes momentos, se ha afirmado que los pacientes amnésicos directamente no pueden codificar en la memoria episódica (o declarativa), que sólo codifican en niveles superficiales y no profundos y que codifican acontecimientos pero no sus contextos espacio-temporales. Hay evidencias a favor y en contra de cada una de estas ideas. Parte de la dificultad para deci-

dir entre ellas es que el rendimiento de pacientes como K. depende en mucho de la tarea que se les encargue y también de las instrucciones sobre cómo se la debe realizar. Por ejemplo, Cermak, Butters y Moreines (1974) hallaron evidencias de que los pacientes a los que se les daban listas cortas de palabras para recordar fallaban en la codificación semántica, es decir, en un nivel profundo. Pero Winocur, Kinsbourne y Moscovitch (1981) descubrieron, con una tarea similar, que los pacientes podían codificar la información semántica cuando las instrucciones subrayaban la importancia de hacerlo. Pareciera que los límites del rendimiento amnésico no fueran fijos y no hubiera una pérdida absoluta de ninguna aptitud de codificación.

Ante este esquema de datos tan confuso, Hirst y Volpe (1984) han hipotetizado que lo que no funciona bien en amnésicos como K. es la automaticidad de la codificación. Se necesita esfuerzo para retener información que por lo común se codifica de manera automática (como el contexto espaciotemporal). Pero el consumo de esfuerzo en este sentido sólo puede hacerse a costa de otros tipos de información que normalmente podrían codificarse con esfuerzo. Lo que se codifica, y se recuerda, depende de dónde se pone el esfuerzo. Una persona normal a la que se le presentan diferentes imágenes de objetos simples puede prestar atención para recordar la identidad de los objetos, y sus ubicaciones espaciales se registrarán automáticamente. Pero para los amnésicos no es lo mismo (Hirst y Volpe, 1984). Parece que los amnésicos tienen que hacer una elección sobre cómo distribuir el esfuerzo entre los diferentes tipos de información que podrán recordar.

Esta idea de Hirst y Volpe es todavía una hipótesis que espera evaluación a la luz de futuras investigaciones. Por la misma razón que los procesos automáticos no están abiertos a la inspección consciente, puede resultar difícil para el lector imaginar cómo esos procesos podrían fracasar en una persona que estaba prestando atención al material experimental. La forma en que esto podría suceder se comprenderá mejor en el próximo capítulo, en el que examinaremos el desarrollo de las aptitudes mnemotécnicas. Algunas pueden automatizarse tanto que tal vez nos sorprenda descubrir que no todos las poseen.

## Resumen

En este capítulo nos hemos ocupado de mucha información. Hemos explorado las ideas de habilidades de memoria automática y memoria con esfuerzo, con referencia a los logros y deficiencias mnemónicos de Lorayne, el mnemotécnico, y de K., el amnésico. También presentamos parte de la terminología y las distinciones empleadas en el estudio de la memoria. Le hemos prestado especial atención a las variedades de almacenamiento de memoria de corto y largo plazo que reconocen los psicólogos, así como también a la importancia de las variables de codificación y los factores de recuperación para el rendimiento de la memoria en distintas circunstancias.

## Lecturas adicionales

WINGFIELD, A. y BYRNES, D.C. (1981) *The Psychology of Human Memory*. Nueva York: Academic Press.

Buena cobertura de las investigaciones experimentales sobre la memoria.

NEISSER, U. (comp.) *Memory observed*. San Francisco: W.H. Freeman and Company.

Interesante selección de artículos relativos a la memoria en contextos naturales.

# 6

# Adquisición de habilidades de la memoria

*Angus Gellatly*

## La memoria: ¿qué desarrolla?

Aunque a veces se atribuye a los niños pequeños una capacidad excepcional para ciertos tipos de pruebas de memoria (ver más abajo), generalmente la memoria se incrementa a medida que el niño va creciendo. Podemos preguntarnos, entonces, ¿por qué? ¿Qué es lo que tiende a mejorar con la mayoría de edad?

Esta pregunta fue planteada por el psicólogo John Flavell en 1971, cuando convocó a un simposio con el título: "¿Qué desarrolla el desarrollo de la memoria?" En síntesis, la respuesta que surgió del simposio fue que el desarrollo de la memoria es parte del desarrollo cognitivo general. Hay, por supuesto, técnicas memorísticas especializadas (estrategias), pero el desarrollo de la memoria implica mucho más que la mera adquisición de estas estrategias. La memoria no puede divorciarse del pensamiento y la comprensión, del conocimiento del mundo físico y el mundo social y del conocimiento de la propia capacidad para actuar en estos mundos. El desarrollo de la memoria depende del desarrollo de otras áreas de la cognición y se realimenta con ellas. Refleja el creciente control del niño sobre la dirección de su atención, su naciente conciencia de sí mismo como actor social y su mayor dominio de la importancia de la planificación para la memorización activa.

La pregunta de Flavell y los términos en los que fue respondida han configurado mucha de la investigación posterior sobre la memoria en los niños (Chi, 1983). En gran parte, este capítulo se ocupará de ampliar los roles jugados en el desarrollo de la memoria por las estrategias, el conocimiento y la comprensión de las propias aptitudes mnemotécnicas. Pero, antes de continuar, conviene considerar cierto proceso muy básico de la memoria al que los colegas de Flavell le han prestado muy poca atención. Su interés se centraba en las memorizaciones activas de los niños que ya han adquirido el lenguaje. Sin embargo, estos son logros relativamente sofisticados que tienen su

origen en potencialidades no lingüísticas que pueden atribuirse a otros animales, además de los humanos. Los fenómenos de aprendizaje asociativo como el *condicionamiento* y la *habituación* son de esta naturaleza. Se los puede observar incluso en casos en los que la palabra "memoria" casi no parezca siquiera aplicable. Tanto en la evolución como en el desarrollo individual, la memoria se desarrolla de lo simple a lo complejo. En sus formas más elementales, la memoria es pasiva y automática. La mayor complejidad llega con la intervención de la intención y el esfuerzo, como bien observara Flavell (1971). Pero para empezar este capítulo podemos echar una mirada a algunos de los ejemplos menos intencionales del recordar.

## Habituación

"Habituación" es la palabra usada para denotar la pérdida de respuesta de un organismo a estímulos repetidos o continuos. Por ejemplo, si ponemos un objeto extraño en la jaula de un ratón, puede provocar miedo y respuestas evasivas, pero en poco tiempo el ratón se habituará al objeto y no mostrará respuestas evidentes ante él. A nivel celular, el estímulo eléctrico repetido de una neurona puede llevar a una disminución de la respuesta ante estímulos posteriores (Kandel, 1976). Semejantes casos de habituación celular pueden explicarse en términos de bioquímica celular alterada y probablemente no atraigan la etiqueta de "memoria". Pero en el nivel de los animales completos, la habituación suele definirse como la pérdida del interés (o de la atención) ante un estímulo auditivo o visual repetido. El efecto suele explicarse por el establecimiento de alguna representación interna del estímulo, que permite el reconocimiento inconsciente en intentos posteriores; aquí parece hacer falta la palabra "memoria".

La técnica de mirada selectiva suele usarse para demostrar habituación en los niños. Primero, se le permite al niño ver un modelo, como un ojo de buey, durante algún período extenso. Luego, tras cierto tiempo, se le presenta el mismo ojo de buey junto a un nuevo modelo, como un tablero de ajedrez, y se cuenta cuánto tiempo se fija cada estímulo. En estudios de esta naturaleza se ha demostrado que niños de entre 1 y 4 días de vida demuestran una preferencia selectiva por los estímulos nuevos sobre los previamente vistos (Kail, 1984). Incluso los recién nacidos, entonces, pueden tener recuerdos que duran más de los breves lapsos que son típicos en los experimentos de mirada selectiva, y otros estudios han señalado que se pueden ob-

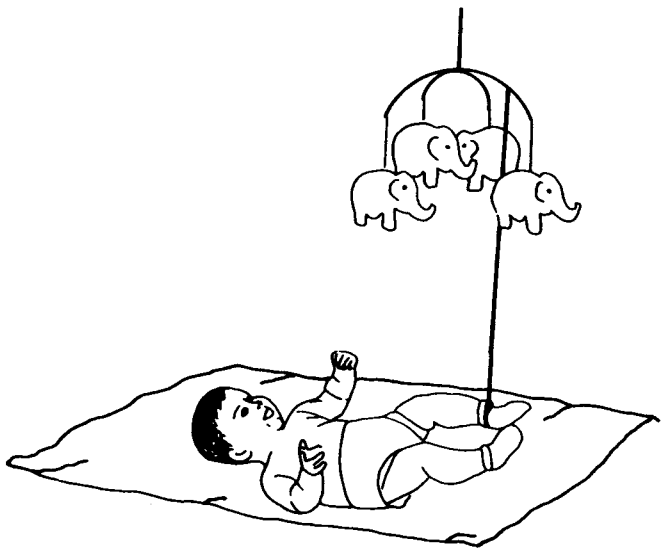
servar efectos que duran horas o días en niños de apenas dos meses de vida. Sin embargo, debe destacarse que aunque esta forma simple y pasiva de memoria puede notarse en los muy pequeños, el efecto se hace más evidente con el aumento de la edad.

## Condicionamiento

El *condicionamiento clásico* es la forma en que una respuesta natural a un estímulo puede llegar a asociarse con un segundo estímulo neutral, que predice la llegada del primero. El ejemplo más famoso es el del perro de Pavlov, que aprendió a generar saliva al oír una campana cuyo sonido precedía al momento en que le daban la comida. ¿Se pueden obtener efectos similares con los niños? A veces, se ha pensado que los niños de menos de 6 meses tenían un déficit en el condicionamiento clásico debido a inmadurez cortical. Pero en los últimos veinte años se ha demostrado que, en circunstancias controladas, el condicionamiento clásico del niño no difiere mucho del de un adulto (Rovee-Collier y Lipsitt, 1982). De modo que, una vez más, el potencial para la memoria pasiva se ve como algo inherente a los muy pequeños.

En el *condicionamiento instrumental* hay una forma más activa de aprendizaje y memoria. Esta expresión hace referencia al hecho de que si se recompensa a un animal por alguna acción, probablemente la repita. En este caso, el ejemplo más famoso está dado por las ratas de Skinner, que aprendieron a presionar una palanca para obtener alimento. El condicionamiento instrumental es la base para la mayor parte del entrenamiento animal y al menos parte del aprendizaje humano, pero ¿es discernible en los recién nacidos? Nuevamente, la respuesta es sí. Siqueland y Lipsitt (1966) pudieron condicionar el movimiento de la cabeza en niños de dos a cuatro días de vida, usando una solución de dextrosa como recompensa. Rovee-Collier y sus colegas (citado por Kail, 1984) lograron aumentar la frecuencia de movimiento de piernas de los niños, recompensándolos con el movimiento de un móvil atado a la pierna por un hilo (ver figura 6.1). Los niños de dos meses "recordaban" esta contingencia durante tres días y los de tres meses, durante ocho días.

*Figura 6.1 Preparación para el movimiento de piernas condicionado. (Según Kail, 1984.)*



Para concluir esta sección: los niños son capaces de más de un tipo de aprendizaje desde muy poco después de nacer, y por lo tanto deben venir equipados con el potencial para formas simples de memoria. Sin embargo, en los tres casos que hemos examinado (habituación, condicionamiento clásico y condicionamiento instrumental) la memorización no ha sido intencional de parte de los niños. Incluso en el caso del aprendizaje instrumental, donde parece haber una intención de ganar más comida o el movimiento del móvil, no parece haber memorización intencional. El niño no elige recordar que el movimiento de su pierna puede producir el del móvil. La memoria es un subproducto del intento de producir el movimiento del móvil. En el pasaje progresivo de la memoria pasiva y automática hacia la memoria esforzada e intencional es donde los niños, como los adultos, encuentran dificultades.

## Estrategias de ensayo y categorización

Supongamos que a una persona le muestran una lista de ítem identificables y categorizables (por ejemplo, cuatro animales, cuatro alimentos y

cuatro colores) y le dicen que en unos segundos le pedirán que los recuerde. Casi con seguridad, en ese lapso hará dos cosas: los repetirá para sí mismo y, al hacerlo, los agrupará según sus categorías. El ensayo y la categorización son estrategias mnemotécnicas tan obvias y automáticas que uno casi podría sentir que no puede considerarlas estrategias intencionales. En ese caso, puede resultar sorprendente enterarse de que los niños pequeños de nuestra cultura, y las personas mucho mayores de otras culturas, no usan de manera espontánea el ensayo y la categorización.

Flavell, Beach y Chinsky (1966) usaron la lectura de labios para hacer un índice del ensayo en una tarea que exigía el recuerdo de imágenes. Descubrieron que el 10% de los niños de 5 años, el 60% de los de 7 años y el 85% de los niños de 10 años realizaban ensayos. Estudios posteriores demostraron que el hecho de no utilizar el ensayo no implica necesariamente incapacidad para hacerlo. Los niños de 6 y 7 años que no muestran ensayo espontáneamente pueden ser entrenados para seguir la estrategia y, como consecuencia, su rendimiento en el recuerdo aumenta hasta el nivel de aquellos de su misma edad que lo realizan espontáneamente. Pero, aunque son totalmente capaces de usar la estrategia, los que no lo hacen de manera espontánea tienden a abandonarla cuando no los alienta un experimentador. Al no usar una estrategia disponible cuando es adecuada, exhiben lo que se conoce como "deficiencia de producción".

Otros experimentos demuestran que incluso cuando los niños pequeños producen el ensayo espontáneo, no lo hacen muy bien (ineficiencia de producción). Con los ítem no categorizables existe la tendencia a repetir solamente una palabra por vez antes que recitar sucesiones de palabras que podrían vincularse entre sí rítmicamente (Ornstein, Naus y Stone, 1977). También es interesante observar que los enfermos de Korsakoff evidencian además ineficiencia en este tipo de ensayo.

Cuando se usan imágenes categorizables como estímulo para el recuerdo, se observa un patrón similar de deficiencia de producción seguida de ineficiencia de producción. Los niños de menos de 10 años tienen poca tendencia a agrupar los ítem en categorías mientras los estudian para su posterior recuerdo (Moely, Olson, Halwes y Flavell, 1969). Además, cuando se les proporcionan pistas de recuperación similares a los nombres de las categorías, los niños más pequeños no las usan eficientemente para organizar su recuerdo (Kobasigawa, 1974). Mientras que los niños de 11 años intentan recordar todos los miembros de una categoría antes de pasar a la siguiente, los niños de 6 y 8 años tienden a recordar sólo uno de ellos antes de seguir ade-

lante. Los niños más pequeños saben a qué categorías pertenecen los ítem pero *no saben cómo usar su conocimiento* para mejorar el rendimiento en la tarea memorística.

Todo lo antedicho podría servir para sugerir que el hecho de no emplear estrategias mnemotécnicas se debe fundamentalmente a una cuestión técnica de motivación. Después de todo, la mayoría de nosotros sabemos que tomar notas puede ayudar a recordar los contenidos de un libro que estamos estudiando, pero no siempre nos tomamos la molestia de hacerlo. De manera similar, algunos niños pueden no sentirse motivados para esforzarse en una insignificante tarea de recuerdo propuesta por un psicólogo.

Aunque ésta es una explicación posible de las deficiencias e ineficiencias de producción, hay muchas razones por las que probablemente no sirva en la mayoría de los casos. En primer lugar, muchos niños pequeños parecen ansiosos por desempeñarse bien en las tareas que encaran. En segundo lugar, incluso los estímulos económicos producen pocas mejoras en el rendimiento. En tercer lugar, hay una explicación alternativa que dará cuenta tanto de las diferencias culturales como de las etarias en la habilidad memorística.

## Diferencias culturales en el rendimiento de la memoria

Todos los experimentos a los que hemos hecho referencia en la última sección investigan el rendimiento de niños educados en una cultura occidental. También se han hecho estudios del recuerdo de listas categorizables en gente no escolarizada y analfabeta de la tribu Kpelle, de África Occidental (Cole y Scribner, 1974). En esos estudios, se descubrió que el recuerdo era pobre y se recurría poco a la categorización. No hubo signos de mejorías drásticas con la edad, típicas de la cultura occidental, y los adultos Kpelle se desempeñaron muy poco mejor que los niños pequeños de cualquiera de las dos culturas.

El hecho de que incluso los adultos Kpelle usaran poco la categorización en el recuerdo es un descubrimiento particularmente útil. En nuestra cultura, la categorización es una forma tan obvia de encarar el recuerdo que la no agrupación automática de ítem en categorías se interpreta, a veces, como evidencia de daño cerebral (Goodglass y Kaplan, 1979). Los enfermos de Korsakoff exhiben precisamente este déficit (Hirst y Volpe, 1984). Sin embargo, se requiere una explicación muy diferente para saber por qué los Kpelle no muestran categorización espontánea en el recuerdo.

La respuesta al dilema ha sido resumida por Hunter (1985). Junto con Cole y Scribner, sostiene que la gente desarrolla habilidades mnemotécnicas que resulten útiles en su propia cultura y, por lo tanto, en su vida cotidiana. El recuerdo de un conjunto arbitrario de cosas por el sólo hecho de recordarlo es el tipo de tarea que la gente occidental educada enfrenta repetidamente en la escuela. Como consecuencia de ello, los occidentales educados desarrollan las estrategias adecuadas de ensayo y categorización hasta el punto en que se tornan automáticas. No sucede lo mismo entre los Kpelle. Por otra parte, la narración de historias es un arte muy practicado en la cultura Kpelle y es esperable que sus miembros sean narradores expertos. Cuando Cole y Scribner insertaron las palabras que se debían recordar en una historia (en lugar de presentarlas en una lista aislada), los Kpelle mostraron una alta capacidad de recuerdo y, también, recordaron los ítem por categoría. Puede verse que la gente no escolarizada puede recordar listas de palabras y utilizar categorías de información, pero en circunstancias diferentes de aquellas en las que la gente escolarizada hace lo mismo.

Otro estudio intercultural de la memoria, según el informe de Kearins (1981), fue realizado con niños aborígenes del desierto y niños blancos criados en la ciudad. En este experimento, se daban 30 segundos a un niño para que viera un conjunto de 5 por 4 cuadrados en los que había 20 objetos simples y luego se le mostraban los objetos desordenados y debía volver a ponerlos en sus posiciones originales. En todos los grupos etarios, desde los 6 hasta los 16 años, y con varios tipos de objetos, el rendimiento de los niños aborígenes fue notoriamente superior. Pero no sólo estos niños reubicaron los objetos mejor sino que, además, encararon la tarea de manera diferente. Los niños blancos intentaban recordar las posiciones de los objetos ensayando sus nombres en orden; cuando se les permitía empezar a reconstruir el conjunto, empezaban con mucho apuro, como si tuvieran miedo de olvidarse rápidamente de la información relevante, y después de colocar un objeto en el tablero solían cambiarlo de posición. En cambio, los niños aborígenes no realizaban ensayos verbales, no tenían apuro en comenzar la reconstrucción y, una vez que habían colocado un objeto en un cuadrado, muy raramente lo cambiaban de lugar. Kearins sugiere que, a diferencia de los niños blancos, que intentaban aplicar una estrategia verbal, los niños aborígenes se basaban en una estrategia visual-espacial.

Para nuestros propósitos, dos puntos del estudio de Kearins resultan particularmente relevantes. En primer lugar, como se dijo antes, la tendencia automática a adoptar una estrategia de ensayo verbal no es universal ni con-

fiere siempre ventajas. En segundo lugar, las estrategias de memoria que no entendemos son mucho más intrigantes que las que usamos. No se comprende claramente cómo actúan tan bien los niños aborígenes en la tarea propuesta por Kearins y, por lo tanto, su capacidad parace muy destacable. Sin embargo, puede suponerse que a los Kpelle también les resultaría impresionante la capacidad de los occidentales educados para recordar cantidades relativamente grandes de palabras categorizables en una lista aislada, especialmente si el ensayo se realizara en silencio y sin movimiento de labios. La conclusión es que la gente desarrolla habilidades mnemotécnicas en las que está ejercitada por la vida cotidiana en su cultura.

## El concepto de metamemoria

Las últimas dos secciones han destacado la importancia de las estrategias de memorización para adecuar el rendimiento en una serie de tareas de laboratorio. Se ha descubierto que los niños pequeños, que aún no han dominado estas habilidades, suelen tener un bajo rendimiento en las tareas. Sin embargo, sería un error exagerar las debilidades de los más pequeños. Como señala Kail (1984), los niños en edad preescolar suelen exhibir una excelente memoria para los hechos de sus propias vidas. Y muchos padres podrán confirmar que sus hijos frecuentemente recuerdan acontecimientos y comentarios en circunstancias socialmente incómodas. ¿Por qué, entonces, habrá esta discrepancia entre la memoria de los niños pequeños para los episodios naturales y para los ítem presentados en experimentos?

Kail señala que los niños de cuatro años obtienen mejores resultados en las pruebas de reconocimiento aunque su rendimiento en las pruebas de recuerdo sea asombroso. Bien puede deberse a que muchos de los ámbitos naturales en los que los niños pequeños exhiben sus aptitudes de memoria incluyen reconocimiento, como cuando el niño señala un edificio o un barrio de la ciudad en el que ya ha estado antes. Incluso cuando hay recuerdo (por ejemplo: "Sí que dijiste que la Sra. Pérez bebía demasiado", dicho en presencia de la Sra. Pérez) puede haber múltiples pistas en el contexto físico y social que promuevan el recuerdo. Estas suposiciones quieren sugerir que cuando los niños pequeños demuestran buen recuerdo, es probable que sea de naturaleza pasiva. En la mayoría de los casos, los recuerdos le ocurren al niño. En cambio, cuando se requiere que el niño almacene o recupere información intencionalmente, el rendimiento resulta pobre (Flavell, 1985).

Si se acepta la argumentación, se explica la discrepancia entre los resultados de laboratorio y las experiencias cotidianas. Pero, en ese caso, persiste la pregunta de por qué los niños pequeños tienen tan bajo rendimiento en el recuerdo intencional. Sabemos que no usan espontáneamente estrategias básicas, pero también sabemos que pueden hacerlo si se les dan las instrucciones adecuadas. ¿Por qué muestran, entonces, deficiencias de producción?

Para Flavell, los niños pequeños tienen dificultades con la metamemoria. Metamemoria es el conocimiento de cualquier cosa que tenga que ver con la memoria. Se refiere a nuestro conocimiento personal de cómo funciona el recuerdo en diversas circunstancias. Incluso los niños de 4 y 5 años tienen cierto control de la metamemoria. Comprenden la diferencia entre adivinar y recordar (Johnson y Wellman, 1980) y saben emplear ayudas memorísticas externas, como dejar en lugar visible un objeto que se quieren llevar (Kreutzer, Leonard y Flavell, 1975). Sin embargo, no resulta sorprendente que los niños pequeños carezcan de muchos aspectos de la metamemoria. Son especialmente malos en la predicción de la dificultad de hasta la más simple tarea de memoria. Como los Kpelle, los niños de 4 años tienen poca experiencia en el recuerdo de listas arbitrarias de palabras y por lo tanto no saben en qué pueden resultar beneficiosos el ensayo y la categorización. Esta carencia de metamemoria, la capacidad de diagnosticar las exigencias mnemónicas de diferentes tareas, provoca deficiencias de producción. En algunos casos, puede llevar a discrepancias extrañas. Istomina (1982) halló que los niños podían recordar más ítem si jugaban a "ir de compras" que cuando sólo se les pedía que recordaran las palabras aisladamente. Ir de compras era algo que los niños podían comprender y que automáticamente promovía intentos de recordar, pero la tarea de recuerdo por el recuerdo mismo les resultaba extraña y sin sentido. No provocaba ensayo automáticamente, y como los niños sabían poco sobre la función del ensayo y las exigencias de las tareas de recuerdo, no podían llegar a esa estrategia mediante el pensamiento consciente.

El desarrollo de la metamemoria probablemente continúe durante toda la vida, y deberíamos sentir cierta simpatía por el niño que encara una prueba de memoria. Aun si se les diera una semana entera, pocos adultos sabrían cómo encarar el aprendizaje del papel de Hamlet; sin embargo, los actores pueden lograr notables hazañas de este tipo: como consecuencia de la experiencia, han descubierto las exigencias mnemotécnicas del aprendizaje de papeles y tienen las estrategias y la metamemoria necesarias.

## La naturaleza de la metamemoria

La metamemoria es un campo de estudio complejo y aquí sólo podemos hacer un breve esbozo del mismo. Flavell (1977) distingue dos categorías principales de metamemoria. Una es la sensibilidad a las exigencias mnemónicas de las diferentes actividades. El niño debe aprender que, para recordar cosas, es necesario esforzarse por memorizarlas. La segunda categoría es el conocimiento de las propias habilidades, cómo funcionan en las diferentes tareas, qué estrategias usan con mayor eficiencia. En esta categoría se encuentra la capacidad de controlar la propia experiencia, saber que algo que por el momento no se puede recordar está igualmente disponible, o prever que aunque ahora se puedan repetir las instrucciones que se acaban de recibir, dentro de una hora no será posible hacerlo. La metamemoria incluye la capacidad de predecir el propio rendimiento futuro.

Como se dijo antes, los recuerdos pueden aparecer porque las pistas externas lo imponen. Puede haber pistas de entrada (o de codificación) de la información y pistas para recuperarla. La exigencia de juzgar si un rostro parece honesto o deshonesto actúa como pista de codificación; da como resultado un mejor reconocimiento de ese rostro más adelante, aunque no hubiera habido intención de memorizarlo al principio (Bower y Karlin, 1974). El aroma de un perfume puede actuar como pista para la recuperación no intencional de muchísmos recuerdos.

Todos estamos a merced de las pistas de la memoria en los medios físico y social, incluyendo las demandas implícitas de los roles sociales ("recordar nuestro lugar"). Ellas proporcionan un andamiaje externo a nuestra experiencia. La habilidad mnemotécnica reside en la capacidad adicional de proporcionar nuestro propio andamiaje interno y externo y ejercitar así cierto control intencional sobre la experiencia presente y futura. La adquisición de la metamemoria es parte de este proceso.

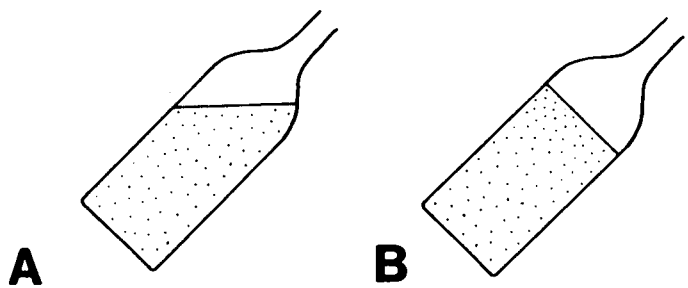
La metamemoria se encuentra estrechamente ligada a las exigencias mnemónicas que la vida cotidiana impone al individuo. En una cultura escolarizada, uno se torna sensible a las exigencias de memoria de las tareas de tipo escolar y aprende sobre las propias aptitudes al respecto (ver capítulo 12). En una sociedad de cazadores y recolectores, o en una cultura en la que la narración de historias sea una actividad significativa, se explorarán tareas muy diferentes. La gente desarrolla las habilidades memorísticas que su forma de vida necesita y la metamemoria correspondiente a esas habilidades. Actualmente, los investigadores se ocupan de cómo se puede acelerar el de-

sarrollo de la metamemoria mediante la provisión de experiencias de aprendizaje adecuadas (Flavell, 1985).

## Memoria y conocimiento

El rendimiento en las pruebas de memoria no sólo está determinado por estrategias memorísticas y la metamemoria. Como sucede con otras habilidades, el conocimiento juega aquí un papel importante (Chi, 1985). Esto es muy evidente en un experimento de Piaget en el cual se pidió a algunos niños que reprodujeran una ilustración que se les había mostrado una hora antes (ver figura 6.2A). Los niños de menos de 8 años con frecuencia no logran dibujar una reproducción precisa del original. Cometen ciertos errores típicos. Uno de los más comunes es el que se ilustra en la figura 6.2B. Aquí, el nivel del agua es dibujado paralelo al fondo de la botella y, por lo tanto, oblicuo en vez de horizontal. Al tratar de recordar la apariencia del original, el niño reconstruye sobre la base de un conocimiento inadecuado del mundo físico. Sin embargo, si al mismo niño se le vuelve a pedir que dibuje la imagen unos meses después, la nueva reproducción bien puede ser correcta. La memoria, paradójicamente, ha mejorado con el paso del tiempo. La mejora se debe a un aumento en la comprensión y esto, según Piaget e Inhelder (1973), es consecuencia del desarrollo de las estructuras cognitivas que se supone subyacen al pensamiento del niño.

*Figuras 6.2. Ilustración original y una de las reproducciones obtenidas en el experimento de Piaget.*



La prueba del nivel del agua es un buen ejemplo del papel de la reconstrucción en la memoria. A veces recordamos mal hechos pasados (o disentimos con otra gente sobre ellos) debido a la reconstrucción. Recordar es con frecuencia reordenar informaciones parciales o usar los conocimientos lógicos y empíricos para decidir cómo debe haber sido todo. Como señala Luria (1974), el niño inicialmente aprende a pensar recordando; luego, más adelante, aprende a recordar pensando.

Otra demostración experimental de la importancia del conocimiento para el rendimiento en las pruebas de memoria fue la de Chi (1978), quien utilizó una prueba de recuerdo de dígitos y la tarea de reconstrucción del ajedrez que se describe en el capítulo 2. Un grupo de adultos con alto nivel de educación pero sin conocimiento alguno sobre ajedrez y un grupo de niños entrenados en ajedrez actuaron como sujetos del experimento. En el recuerdo de dígitos, se observó la superioridad adulta usual; pero al reconstruir configuraciones significativas de piezas de ajedrez, aquellos que conocían el juego, aunque eran menores, fueron mucho más eficientes. Este resultado no es sorprendente; por el contrario, es lo esperable por la destreza de los niños en el ajedrez. Sin embargo, proporciona un ejemplo evidente de que el conocimiento suele ser el factor determinante de la capacidad mnemónica. Los adultos demuestran por lo general un mejor rendimiento que los niños en las pruebas de memoria porque tienen más práctica en ese tipo de pruebas y por lo tanto son mnemotécnicos más diestros, y también porque usualmente conocen más el campo en el que se está probando la memoria.

## Aptitudes de memoria especiales en los niños

Al comenzar este capítulo dijmos que, a veces, se atribuye a los niños pequeños una capacidad excepcional para ciertos tipos de pruebas de memoria. El análisis posterior enfatizó las deficiencias de memoria en los pequeños y las formas en que estas deficiencias son concomitantes inevitables del conocimiento y la práctica insuficientes. ¿Cómo se concilia, entonces, esta supuesta capacidad excepcional con la imagen que se ha hecho del niño como aprendiz mnemotécnico? Intentaremos responder la pregunta en esta sección, examinando dos tareas que incluyen memoria visual-espacial y de las que se dice que muestran a los niños bajo una luz más positiva.

La primera no es una tarea inventada por psicólogos curiosos sino un juego que disfrutan tanto niños como adultos. Este juego recibe varios nombres: pares, "memotest" o memory. En una versión simple, se mezcla un ma-

zo de cartas y luego se las esparce con la figura hacia abajo sobre una mesa. A su turno, cada jugador debe dar vuelta dos cartas y dejarlas a la vista de todos los demás. Si esas cartas forman un par, el jugador se las queda y tiene un turno más; si no, debe volver a ponerlas con la figura hacia abajo. El ganador es el jugador que termina llevándose más pares de cartas; evidentemente, el juego exige recordar qué cartas se encuentran en qué posiciones. Varios autores (por ejemplo, Meacham, 1972; Gardner, 1982; Wood, 1983) han escrito que los niños suelen vencer a los adultos en el juego pero en ningún caso han aportado datos que lo corroboren. Para rectificar esta omisión, se hicieron estudios en niños de 5, 7 y 9 años y en estudiantes universitarios que practicaran una versión individual del juego, usando como índice de su rendimiento el tiempo y la cantidad de intentos (Gellatly, Jones y Best, en prensa). Los resultados demostraron que los niños de 5 y 7 años tenían un rendimiento marcadamente más bajo que los demás, pero entre los otros grupos no hubo diferencias significativas. Más aún, el bajo rendimiento de los niños de 5 y 7 años fue visto como resultado de la dificultad de implementar confiablemente una estrategia típica de los niños mayores y los adultos. Naturalmente, algunos niños de 5 años son muy buenos en el juego, pero las evidencias demuestran que el rendimiento en la tarea mejora con la edad de la misma forma en que lo hace en otras tareas de memoria y debido a un tipo de desarrollo similar. No es necesario atribuir habilidad especial a los muy pequeños en este caso.

La segunda habilidad especial que se dice que tienen los niños pequeños —y que a veces se cita como explicación del alto rendimiento en este juego (Gardner, 1982)— es la de la imaginería eidética. Una imagen eidética es una imagen de memoria particularmente vívida y detallada de una ilustración o escena vista recientemente. En general, esas imágenes se experimentan como "estando allí" más que en la mente, y se las puede proyectar en un muro o en una pantalla. Se las describe en presente y no en pasado, se las puede escudriñar visualmente y se pueden superponer pares de imágenes para que formen un compuesto que contiene información que no está presente en ninguna de las dos imágenes por separado (Haber, 1979). La imaginería eidética casi no existe en los adultos pero se encuentra en alrededor del 5% de los niños de 6 a 12 años. Su naturaleza no está muy clara pero parece representar una instancia genuina de niños pequeños que demuestran mejor memoria que los adultos.

Aunque resulta interesante por sí misma, la imaginación eidética es diferente de los tipos de memoria que hemos tratado en este capítulo, ya que

constituye un registro incidental de información más que una forma activa e intencional de memorización. Aunque su desaparición antes de llegar a la edad adulta sigue resultando incomprensible, la existencia de la imaginación eidética no es incoherente con la consideración del desarrollo de la memoria como adquisición de estrategias intencionales de memorización, crecimiento de la metamemoria y acumulación de conocimientos.

## Conclusión

En este capítulo hemos analizado el desarrollo de la memorización intencional desde sus orígenes en la memoria pasiva. Se ha enfatizado el modo en que una forma de vida se asocia con la adquisición de habilidades de memoria específicas. Se ha visto su adquisición, en este sentido, como algo más o menos fortuito, ya que las habilidades reales se adquieren simplemente como consecuencia de las actividades cotidianas. En los capítulos 7 y 8 analizaremos cómo las habilidades mnemotécnicas pueden desarrollarse mediante entrenamiento y práctica intencionales.

### Lecturas adicionales

HUNTER, I.M.L. (1985) "Memory development: the cultural background". En J. A. Branthwaite y D.R. Rogers (comp.), *Children Growing Up*. Milton Keynes: Open University Press.

Breve recuento de los factores culturales que influyen en el desarrollo de la memoria.

KAIL, R. (1984) *The Development of Memory in Children* (2ª edición). San Francisco: W.H. Freeman and Company.

Entretenido análisis de los estudios experimentales sobre el desarrollo de la memoria.

# 7

# La habilidad de memoria excepcional

*Ian M. L. Hunter*

Este capítulo se ocupa de aquellos que logran notables hazañas con su memoria. El objetivo es destacar el papel que, en tales hechos, juegan el conocimiento y la habilidad previamente adquiridos.

Consideremos un ejemplo real. Un artista visita el museo del Louvre en París, observa cuidadosamente el retrato de Erasmo pintado por Holbein y luego hace una reproducción del retrato, de memoria. Este hecho está descrito por Boisbaudran (1911) en un libro que incluye fotografías del retrato original y de la reproducción. Bien podemos decir que el artista "tiene una memoria sorprendente". Si con esto sólo queremos decir que lo que hizo es poco común e irrealizable para la mayoría de la gente, nuestro comentario será correcto. Pero si queremos decir que la naturaleza lo ha dotado de una facultad especial de memoria visual-espacial, estaremos sugiriendo algo que va más allá de las evidencias e ignorando que, de hecho, el artista ha pasado por un largo entrenamiento en la observación de cuadros, lo que lo ayuda para poder copiarlos de memoria.

La idea general es la siguiente: los hechos inusuales de memoria no implican, necesariamente, que la persona que los realiza sea constitucionalmente anormal, es decir, dotada por la naturaleza de un don, una facultad o una aptitud cualitativamente especiales. Es posible que esos dones biológicos existan, aunque es difícil establecerlo con certeza. Pero hay dos cosas seguras. En primer lugar, la práctica intensiva puede llevar a habilidades de memoria verdaderamente sorprendentes. En segundo lugar, quienes realizan hazañas memorísticas inusuales son con frecuencia personas que han practicado mucho.

Para dar sustento a las afirmaciones generales que acabamos de hacer, nos centraremos en una investigación en la que un estudiante de una universidad estadounidense, Steve Faloon, se convirtió en un experto en la tarea de memorizar dígitos (ver capítulo 5). Chase y Ericsson (1981) hicieron el relato más completo de esta investigación; sus implicaciones fueron elaboradas por Chase y Ericsson (1982), Ericsson y Chase (1982) y Ericsson (1985).

### La hazaña de Steve con la memoria

Steve escucha una sucesión de dígitos leídos en voz alta a una velocidad de uno por segundo. Han sido seleccionados al azar, de manera que no puede saber cuál será el número siguiente. Sigue escuchando hasta que se le han mencionado 80 dígitos. Luego, los repite todos. No comete errores en los dígitos ni en la secuencia y no olvida ninguno.

Este hecho es evidentemente excepcional. Cuando la mayoría de los adultos normales encara la tarea, la cantidad de dígitos que puede repetir correctamente es, como promedio, siete (más o menos dos dígitos). También se sabe, en términos generales, por qué la capacidad de recordar dígitos es normalmente tan pequeña. Tenemos dos tipos de sistemas de almacenamiento en los que podemos contener información (ver capítulo 5). Uno es la memoria de corto plazo (MCP), que es temporaria y tiene una muy baja capacidad de almacenamiento. El otro es la memoria de largo plazo (MLP), que es más duradera y tiene, virtualmente, una capacidad de almacenamiento infinita. Cuando percibimos información nueva y le prestamos atención, podemos mantenerla disponible durante un breve lapso, pero luego la perdemos irremediablemente a menos que le volvamos a prestar atención, es decir, que la ensayemos de alguna forma. Por ejemplo, si buscamos un número telefónico, podemos retenerlo mediante ensayo atento hasta que llegamos al teléfono pero, si algo nos distrae, lo olvidaremos. En cambio, si aprendemos el número en la memoria de largo plazo, estará disponible durante más tiempo sin necesidad de que le prestemos atención de manera constante. En la tarea de memorización de dígitos, las personas por lo general recurren a la MCP y su capacidad muy limitada se debe fundamentalmente a la pequeña capacidad de almacenamiento de este tipo de memoria.

Para descubrir qué hay detrás de las hazañas memorísticas de alguna persona, debemos analizar dos aspectos clave. Primero, el trasfondo biográfico y social; segundo, los procesos usados durante la hazaña. Las próximas dos secciones se ocupan de estas cuestiones en relación con la hazaña de Steve.

## Trasfondo biográfico y social

Steve realizó su hazaña memorística en el ámbito de una investigación psicológica de laboratorio, realizada a fines de los años setenta en la Univer-

sidad Carnegie-Mellon, en los Estados Unidos. Los tres participantes en esa investigación fueron: William Chase y Anders Ericsson, psicólogos experimentales, y Steve Faloon, que actuó como sujeto. Steve era un estudiante de poco más de 20 años de edad. Tenía altas calificaciones y, en cambio, no demostraba aptitudes extraordinarias para un estudiante brillante. Pero era corredor de fondo, de distancias largas. Esto le daba una actitud competitiva y una ambición de ruptura de récords y también un conocimiento de los tiempos de carrera, que habría de capitalizar cuando encarara la tarea de recuerdo de dígitos. Pertenecía a los equipos de pista y campo y de maratón de su universidad. Se entrenaba corriendo entre 16 y 20 kilómetros por día y, en los últimos nueve años, había competido en muchas carreras de larga distancia. Sus mejores carreras eran la de 3 millas (aproximadamente 5 km), la de 5 millas (aproximadamente 8000 metros) y el maratón; sus mejores tiempos en estas carreras eran, respectivamente, 14'39", 25'40" y 2h 39'36".

Cuando comenzó la investigación, el objetivo era darle a Steve un par de semanas de práctica en la tarea de recuerdo de dígitos y averiguar los efectos que producía. Pero la práctica resultó tan interesante que los tres participantes quisieron seguir adelante. La investigación siguió extendiéndose y llegó a durar dos años. En ese período, Steve concurrió al laboratorio para sesiones de una hora de dos a cinco días por semana.

En cada sesión, la rutina usual consistía en probar el recuerdo de dígitos de Steve mediante el método de fluctuaciones ascendentes y descendentes. Oía una lista de dígitos al azar leídos a una velocidad de uno por segundo y luego intentaba repetirlos. Si tenía éxito, la lista siguiente contenía un dígito más; si no tenía éxito, la lista se acortaba un dígito. Este método lo mantenía trabajando cerca de los límites de su capacidad.

Inmediatamente después de tratar de recordar una lista, se le pedía que comentara los procedimientos que había utilizado. Al final de cada sesión (de una hora de duración) se le pedía que recordara todo lo que podía sobre las listas leídas ese día. A veces, se reemplazaba la rutina usual por sesiones experimentales que exploraban aspectos de su rendimiento. Todos los procedimientos de esta investigación ampliada eran grabados.

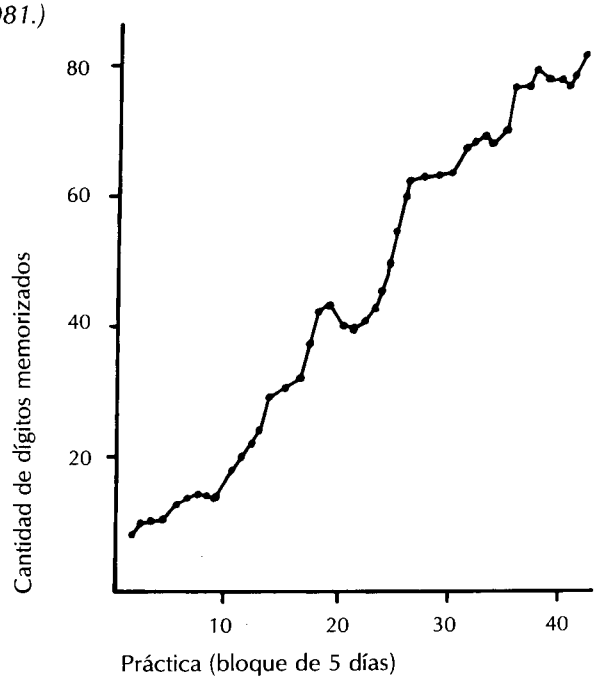
Steve no recibió entrenamiento específico en lo que debía hacer pero constantemente se esforzaba para aumentar el número de dígitos que recordaba. Como un atleta en entrenamiento, enfrentaba barreras para progresar y probaba formas de superarlas. Si una forma funcionaba, la conservaba, la dominaba y seguía adelante hasta que aparecía el siguiente obstáculo. Si no funcionaba, intentaba de otro modo. Para él, entonces, la tarea de memori-

zación de dígitos era un desafío estimulante y, a medida que ganaba experiencia, aparecían nuevas situaciones y el desafío aumentaba. Además, informaba con regularidad y minuciosamente sobre los detalles de su desempeño, y los experimentadores, a su vez, hacían comentarios sobre estos detalles y su progreso. En resumen, la investigación de laboratorio creó un ámbito social que alentaba a Steve a ganar más experiencia creativa en la tarea de memorización de dígitos.

El progreso de Steve se presenta en la tabla de la figura 7.1, que muestra que, en 215 sesiones de práctica, su recuerdo aumentó de 7 a 80 dígitos. Todo indica que, si la práctica hubiera continuado, habría hecho más progresos aún. Aparentemente, el número de dígitos que se recuerda puede aumentar indefinidamente.

El punto clave de esta sección es que la hazaña de Steve se realizó en un ámbito social y biográfico que lo alentaba a ampliar su práctica y su experiencia en algo considerado importante. Los logros de las personas suelen tener que ver con aquellas cosas que les resultan útiles para su vida cotidiana. Para la mayoría de las personas, recordar una gran cantidad de dígitos es inútil y no tienen oportunidades ni incentivos para involucrarse en actividades que los llevarían a mejorar en ello. Pero supongamos un mundo social en el que los récords en la memorización de dígitos fueran tan valorados como los récords deportivos. En semejante ámbito, indudablemente mucha gente lograría recordar largas listas de dígitos. (Chase y Ericsson trabajaron con un segundo sujeto, también corredor de larga distancia, que llegó a recordar 100 dígitos.) Pero lo más extraordinario de la memoria de Steve es que tuvo la oportunidad, la motivación y la persistencia necesarias para lograrlo. Nos recuerda drásticamente que las capacidades humanas se desarrollan a partir de las interacciones entre individuos cuyos entornos ofrecen las posibilidades y los incentivos necesarios para nutrir las aptitudes en cuestión.

*Figura 7.1. Mejoría en el recuerdo de dígitos de Steve Faloon con práctica extendida. Los 43 bloques de práctica aquí mostrados representan unas 215 horas de práctica. Intercaladas entre estas sesiones de práctica hay aproximadamente 75 horas de sesiones experimentales (no mostradas). (Según Chase y Ericsson, 1981.)*




## LOS PROCESOS INVOLUCRADOS

Durante esta investigación, Steve desarrolló nuevas formas de encarar la tarea. En esta sección bosquejaremos cuatro desarrollos fundamentales:

a) Codificación significativa.
b) Estructuración.
c) Rol más importante de la memoria de largo plazo.
d) Aumento de producción.

Obsérvese, de paso, que estos cuatro aspectos son ingredientes esenciales de todos los sistemas mnemotécnicos conocidos, de los cuales el más famoso es el *Método de loci* (ver Lorayne y Lucas, 1974; también, capítulo 8).

## Codificación significativa

Al principio, Steve encaró la tarea de la misma forma que la mayoría. Intentó retener la secuencia de dígitos fonémicamente, ensayando hasta recordarlos. Observaba patrones ocasionales, como una serie de números ascendentes, pero se basaba fundamentalmente en el ensayo atento de la memoria del corto plazo (MCP) y lograba recordar 7 dígitos. Durante las primeras cuatro sesiones no superó esa cifra. En la cuarta sesión, comentó ampliamente que había llegado a su límite y ya no había mejora posible.

Luego, en la quinta sesión logró recordar 10 dígitos y, por primera vez, se refirió a una codificación significativa. Para ejemplificarlo, consideremos la secuencia 4131946273. Trató a los primeros cuatro dígitos como un grupo y pensó en él como "cuatro-trece-coma-uno: el tiempo de una milla"; es decir, 4 minutos, 13 segundos y 1 décima, que es el tiempo que a un atleta le lleva correr una milla. Luego pensó en el segundo grupo de cuatro dígitos como "nueve-cuarenta-y-seis-coma-dos: el tiempo de dos millas". Para su regocijo, descubrió que ahora podía recordar una lista de 10 dígitos como dos grupos codificados de cuatro dígitos cada uno, más otros dos dígitos sueltos al final.

Así, en la quinta sesión, Steve empezó a considerar a los grupos de dígitos números que resultaban significativos si los pensaba como tiempos de carrera. Luego, siguió construyendo un repertorio cada vez mayor de números significativos; la mayoría de ellos eran tiempos de carrera, pero también edades de personas y fechas históricas. Usó estos números para codificar casi cualquier secuencia de tres o cuatro dígitos que se le presentaba. Su empleo de la codificación significativa cobró un propósito claro. Al escuchar los dígitos que se le leían, no esperaba hasta oír los cuatro dígitos para buscar una codificación significativa sino que, al escuchar el primero, empezaba a apuntar hacia una codificación posible y, con cada dígito que oía, estrechaba el campo de búsqueda hasta que, al oír el cuarto dígito o poco después, lograba una codificación adecuada.

## Estructuración

Para la sesión 30, Steve ya recordaba 18 dígitos codificando tres grupos de cuatro dígitos cada uno y reteniendo los últimos seis dígitos como grupo no codificado de dos triples. Pero ahora se enfrentaba a un nuevo obs-

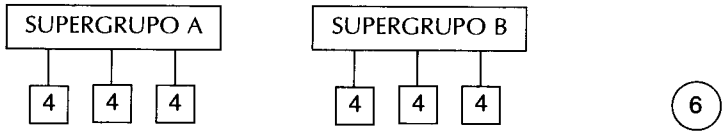
táculo. Cuando aumentaba la cantidad de grupos codificados de tres a cuatro, no podía retener el orden secuencial de los cuatro grupos. En la sesión 32 logró vencer este obstáculo cuando, por primera vez, armó los grupos. Llamemos "supergrupo" a cada agrupamiento de grupos. Veamos su nueva técnica para encarar la memorización de 20 dígitos: pensó en dos grupos, de cuatro dígitos cada uno, que conformaban un supergrupo; luego, dos grupos de tres dígitos cada uno, que integraban un segundo supergrupo; finalmente, un grupo no codificado de seis dígitos.

En las semanas siguientes, mejoró su uso de la técnica de los supergrupos. Pasó a usar supergrupos que no contenían dos sino tres grupos codificados cada uno. Su capacidad de memorización fue aumentando gradualmente hasta llegar a los 40 dígitos. Y entonces apareció otro obstáculo. Cuando intentaba usar supergrupos que contenían cuatro grupos cada uno (en vez de tres), no podía recordar el orden secuencial de los cuatro grupos. En la sesión 109, salvó este obstáculo. Usó supergrupos que sólo contenían tres grupos codificados cada uno e introdujo un nuevo agrupamiento, de nivel superior, que contenía dos supergrupos.

Para aquellos que no conozcan en detalle el rendimiento de Steve, quizá resulte difícil comprender su uso de supergrupos y de grupos de nivel superior, pero el principio general es simple. Steve estaba imponiendo una estructura jerárquica a la lista. Veamos cómo Steve, el experto, encaraba una lista, y recordemos que en todo momento era consciente de qué longitud tendría cada lista.

Incluso antes de escuchar el primero de los dígitos, Steve decidía la estructura exacta y completa que habría de imponer. Con una lista de 30 dígitos, por ejemplo, podía imponer la estructura que se presenta en la figura 7.2, que contiene seis grupos de cuatro dígitos combinados en dos supergrupos y seis dígitos sin codificar al final. A medida que iban apareciendo los dígitos, Steve los incluía en su estructura predeterminada. Cuando, más tarde, recordaba la lista, lo hacía recorriendo la estructura sistemáticamente, usándola de modo que le proporcionara pistas para la recuperación (ver también el capítulo 8).

*Figura 7.2. Representación de la estructura de una lista. Se combinan grupos de cuatro dígitos en tríos, para formar supergrupos. El círculo representa seis dígitos sin codificar.*




Al terminar la investigación, Steve estaba familiarizado con estructuras que podían incluir hasta 80 dígitos y las usaba con suma destreza. Podía hablar de ellas y dibujarlas esquemáticamente. Además, sus estructuras se reflejaban en el ritmo al que recordaba las listas. Al recordar las listas, no mantenía un ritmo de un dígito por segundo sino que los recitaba en "ráfagas" de tres o cuatro dígitos con pausas de diferente longitud entre cada ráfaga. Cada una de estas ráfagas correspondía a los grupos de dígitos, y las pausas dependían de la estructura de la lista; por ejemplo, hacía pausas más largas cuando pasaba de un supergrupo a otro que cuando pasaba entre grupos dentro de cada supergrupo.

## Papel mayor de la memoria a largo plazo

¿Steve aumentó la capacidad de almacenamiento de su MCP? No hay evidencias de que lo haya hecho. En esencia, aumentó la cantidad de dígitos que podía recordar adquiriendo recursos de la MLP que le permitieron obviar las persistentes limitaciones de la MCP. Como novato, su capacidad de recuerdo dependía fundamentalmente de su MCP porque no tenía recursos de la MLP aplicables a la tarea. Con la práctica, construyó recursos de la MLP que podían aplicarse. Como experto, su capacidad de memorización de dígitos pasó a depender fundamentalmente de la MLP; así lo demuestran las siguientes evidencias.

En primer lugar, al terminar cada sesión de una hora, Steve intentaba recordar las listas que se le habían presentado en esa sesión. Cuando el experimento recién comenzaba, no podía recordar prácticamente nada. Cuando su pericia aumentó, recordaba virtualmente todo. Era capaz de recitar las

listas que se le habían presentado. Podía volver a contar cómo había encarado las listas, las estructuras que les impusieron, los grupos que formó y las codificaciones realizadas. Si se le daba cualquier grupo de cuatro dígitos de la lista, frecuentemente podía contar cómo lo había codificado y dónde lo había colocado en la estructura de la lista; es decir, si lo había puesto al principio, en la mitad o al final de un supergrupo. Si se le pedía que recordara porciones seleccionadas de la lista, podía recordarlas sin tener que recorrer toda la lista desde el principio. Usaba la estructura de la lista para moverse selectivamente por ella.

La segunda evidencia surgió de lo que hacía inmediatamente después de que se le presentaba una lista. Cuando era novato, no perdía tiempo y empezaba a recitar los dígitos de inmediato. Como experto, esperaba entre treinta segundos y dos minutos antes de empezar a recitarlos. En esta pausa, consolidaba su comprensión de la lista, la revisaba y ordenaba sus detalles y su estructura.

Las dos evidencias que acabamos de mencionar demuestran que Steve, el experto, procesaba las listas que se le presentaban a través de su MLP. Lo hacía usando el cuerpo de conocimientos que había construido con la práctica. Este conocimiento (de códigos, estructuras y técnicas) estaba almacenado en la MLP y él lo tenía "a mano", disponible rápidamente y sin vacilaciones. Usaba este conocimiento para obviar los límites de la MCP que, de todas formas, seguía imponiendo restricciones a su rendimiento. Por ejemplo, Steve nunca codificaba un grupo de más de cinco dígitos, jamás construía un supergrupo que contuviera más de cuatro grupos, siempre trabajaba con pequeños segmentos y, con tanta rapidez como podía, incorporaba cada segmento en su MLP y se ocupaba totalmente del siguiente segmento de la lista.

Un hallazgo importante sobre los recursos de la MLP, de los que dependía la pericia de Steve, fue que eran específicos, es decir, sólo se relacionaban con el tema de los dígitos y no con otros temas. Por ejemplo, cuando se le daban tareas de memorización de letras del alfabeto elegidas al azar, su capacidad de recuerdo era de seis letras, no más que lo normal. Evidentemente, Steve no había adquirido una "memoria superpoderosa" para todo uso, aplicable a cualquier campo de manera indiscriminada.

## Aumento de producción

Con la práctica, Steve fue ganando mayor velocidad en la realización de los procesos componentes. El proceso de codificación, por ejemplo, me-

joró muchísimo y no dejó de mejorar durante toda la investigación. Semejante aumento de la producción es, por supuesto, enormemente beneficioso porque permite al individuo planificar y ejecutar maniobras más complejas (ver también capítulos 2 y 3).

## Mejora intencional y mejora casual

Al comenzar este capítulo, hice dos planteos generales. En primer lugar, sostuve que la práctica intensa puede llevar a habilidades memorísticas excepcionales. En otras palabras, si empezamos a trabajar con alguien que no demuestra aptitudes memorísticas inusuales, con la práctica podemos lograr que esa persona desarrolle habilidades memorísticas claramente fuera de lo común. Esto se hace evidente en el ejemplo de Steve Faloon. Ericsson y Chase (1982) y Ericsson (1985) presentan otros ejemplos. En segundo lugar, la gente que demuestra tener aptitudes memorísticas inusuales a menudo ha realizado mucha práctica importante para ello. En otras palabras, si trabajamos con alguien que demuestra una aptitud memorística excepcional, con frecuencia veremos que tuvo una larga historia de práctica en el campo en cuestión. Esta sección se ocupa de este segundo tema y, en particular, de la distinción entre mejora intencional y mejora casual de la memoria.

Las habilidades memorísticas de Steve se desarrollaron mientras deliberadamente intentaba aumentar su capacidad de memorización. Durante la persecución de este objetivo explícito Steve incorporó, en su MLP, un variado repertorio de códigos, estructuras y técnicas que le permitió aumentar su capacidad de memorización. Sin embargo, es posible que la gente desarrolle habilidades memorísticas de manera casual, como subproducto de la persecución de objetivos que no apuntan específicamente a la mejoría de la memoria. Semejante mejoría casual de la memoria es una de las características de lo que puede denominarse "efecto memorístico de la pericia"; es decir, que los expertos son buenos para aprender y recordar cosas pertenecientes precisamente al campo en el que son expertos.

Hay muchos campos en los que la gente puede tornarse experta, como el ajedrez, la música, la abogacía, la entomología. La gente llega a ser experta en un campo mediante largas experiencias en él: la materia del campo se torna significativamente familiar. Ahora bien, supongamos que, como psicólogos, elegimos material de cierto campo específico y lo usamos para comprobar cómo puede la gente aprender y recordar este tipo de material.

¿Qué encontramos? Los expertos en ese campo usualmente se desempeñan mejor que los que no son expertos.

Por ejemplo, cuando se les dan tareas de memorización de dígitos a expertos en cálculos mentales, alcanzan cifras dos o tres veces mayores que las comunes. Estos expertos no han practicado estas tareas deliberadamente pero, en el proceso de convertirse en expertos en cálculos mentales, han desarrollado cuerpos de conocimiento y habilidad que les permiten, como subproducto, desenvolverse bien en las tareas de memorización de dígitos. Para dar ejemplos específicos, Hatano y Osawa (1983) probaron las habilidades memorísticas de tres personas que habían ganado en la categoría División Mental en el Campeonato Nacional de Ábacos del Japón. Se les dieron tareas de memorización de letras del alfabeto: los tres campeones memorizaron 5, 8 y 5 letras, respectivamente. Cuando se les dieron palabras en lugar de letras (nombres de frutas), su rendimiento fue de 5, 9 y 5 palabras, respectivamente. Pero cuando se les dieron números, su rendimiento aumentó a 15, 16 y 16 dígitos cada uno. No hay motivos para suponer que pudieran recordar tantos dígitos antes de convertirse en expertos en cálculo mental.

El "efecto memorístico de la pericia" también puede funcionar al revés. Regresemos al artista del que hablábamos al principio de este capítulo. Era uno de los muchos alumnos entrenados por el maestro de arte francés Horace Lecoq de Boisbaudran. Los alumnos pasaban por un entrenamiento que deliberadamente buscaba lograr en los alumnos la capacidad de copiar complejas pinturas de memoria. Adquirían experiencia y práctica en la descomposición de imágenes en componentes que les resultaran familiares y en el uso de estos componentes para reconstituir las imágenes de memoria. En la persecución de este objetivo, adquirían conocimientos y habilidades que podían emplear cuando encaraban una tarea creativa en el campo pictórico. El objetivo de largo plazo de la pericia artística era enfocado a través del objetivo de corto plazo del entrenamiento memorístico en arte.

## El caso S.

El caso de S. V. Shereshevskii (conocido como S.) ofrece un ejemplo de pericia memorística de otro tipo. Las extraordinarias aptitudes memorísticas de S. fueron estudiadas por el psicólogo ruso Luria (1969). S. demuestra cuán desafiante (y a veces frustrante) resulta superar lo que implica ser alguien que llama la atención por su excepcional memoria. S. tenía casi 30

años cuando Luria lo conoció. En esa época, a mediados del decenio 1920-1930, S. era periodista de un diario.

> Cada mañana, el editor se reunía con el personal y les entregaba las tareas del día: las listas de los lugares que quería que cubrieran y la información que debían conseguir en cada lugar. La lista de direcciones e instrucciones solía ser bastante larga, pero el editor comprobó con sorpresa que S. jamás tomaba notas. Estaba por reprocharle su falta de atención cuando, en respuesta a su pedido, S. repitió la lista completa, palabra por palabra. El editor, curioso por saber más sobre la forma en que funcionaba su empleado, empezó a preguntarle por su memoria. Pero S. sólo respondió con sorpresa: ¿realmente había algo de raro en el hecho de que recordara todo lo que le decían?... El editor lo envió al laboratorio de psicología para que se hicieran estudios sobre la capacidad de su memoria.
>
> (Luria, 1969, págs. 7-8).

Luria probó a S. dándole largas secuencias de palabras, sílabas sin sentido y números, para que los aprendiera. Si los ítem se presentaban de manera lenta y clara, al final de la secuencia, S. hacía una pausa para revisar mentalmente el material y luego recordaba toda la secuencia sin cometer errores. S. también podía recordar la secuencia en el orden inverso y, si se le daba cualquier ítem de la secuencia, podía decir de inmediato cuál lo precedía y cuál lo seguía. En una sola presentación, S. aprendía la secuencia en la memoria de largo plazo.

Pero cuando Luria intentó descubrir cómo llevaba a cabo S. sus hazañas de aprendizaje y recuerdo, no logró avanzar demasiado. S. dijo que codificaba el material según *características como las de la percepción*. Una palabra, por ejemplo, podía desencadenar complejas imágenes psíquicas que no eran meramente visuales sino también auditivas, táctiles y gustativas. Sin embargo, S. no podía decir mucho sobre los procesos implicados. Tampoco podía hablar del desarrollo de sus procesos memorísticos en el pasado. Por lo que podía recordar, siempre había tenido esas habilidades. En resumen, en esta etapa de la investigación, Luria no podía determinar si las aptitudes de S. se debían a una peculiaridad innata o a una disposición adquirida para encarar los acontecimientos que experimentaba de maneras que llevaban, como subproducto casual, a recordarlos en detalle. En la vida real, hay algunas preguntas para las que no se pueden encontrar respuestas claras.

Algún tiempo después de conocer a Luria, S. decidió aprovechar su habilidad y convertirse en un mnemotécnico profesional que hacía exhibiciones públicas de su capacidad para aprender y recordar cosas que, para la ma-

yoría, no tienen sentido y son difíciles de memorizar; por ejemplo, listas de palabras o números elegidos al azar. Luria descubrió que, para satisfacer las necesidades de estas actuaciones públicas, S. desarrolló formas de operar significativamente modificadas, que curiosamente incluían los cuatro procesos mencionados cuando hablábamos de Steve Faloon. En resumen, S. desarrolló, con la práctica, cuerpos de conocimiento y habilidad que le permitieron, con más rapidez y eficiencia, codificar y estructurar el material presentado.

Básicamente, en sus actuaciones públicas, S. usaba el *Método de loci* (ver capítulo 8). Por ejemplo, se imaginaba a sí mismo caminando por una calle conocida, llena de cosas que le eran familiares (como la calle Gorki de Moscú) y visualizaba cada ítem que se le presentaba asociándolo con cada hito de la calle, sucesivamente. Uno de los procesos para lograr mayor eficiencia tenía que ver con la simplificación de la visualización realizada.

El hecho de que S. pudiera adaptar sus aptitudes de imaginación para satisfacer las necesidades de su nueva profesión nos dice algo sobre la naturaleza de sus habilidades. No podemos saber si su facilidad para las imágenes era innata o adquirida, pero es evidente que se la podía controlar y usar intencionalmente.

> Antes, para recordar algo tenía que recurrir a una imagen de toda la escena. Ahora, sólo tengo que tomar algún detalle que he elegido de antemano y representará a toda la imagen. Digamos que me dicen la palabra *jinete*. Lo único que se necesita es una imagen de un pie con una espuela... Así que mis imágenes han cambiado bastante. Antes eran más definidas, más realistas. Mis imágenes actuales no están tan bien definidas ni son tan vívidas como las de antes... Sólo trato de visualizar algún detalle que necesitaré para recordar una palabra.
>
> (Luria, 1969, pág. 42).

## Conclusión

Al terminar este capítulo, sabemos algo de la gente que demuestra tener aptitudes memorísticas excepcionales. Sabemos que esa gente existe. Sabemos sobre la importancia de la codificación y estructuración significativas rápidas. Sabemos que sus sorprendentes hazañas a veces pueden ser realizadas por gente normal que tiene mucha práctica. Pero, hasta el momento, la comprensión acabada de las habilidades excepcionales de la memoria sigue siendo un desafío para los psicólogos.

### Lecturas adicionales

ERICSSON, K.A. & CHASE, W.G. (1982) Exceptional memory. *American Scientist*, 70, 607-615.

Breve síntesis de los experimentos realizados con Steve Faloon y sus implicaciones.

LURIA, A.R. (1969) *The Mind of a Mnemonist.* Londres: Cape.

Maravilloso estudio de un caso, sobre la memoria de S.

# 8

# ¿Cómo se pueden mejorar las habilidades de la memoria?

*Angus Gellatly*

Los últimos tres capítulos han hecho referencia, de una u otra manera, a la mejoría de las habilidades de la memoria. Este capítulo, a diferencia de los anteriores, es prescriptivo en lugar de descriptivo. La atención hasta aquí estuvo puesta en la observación de los individuos cuando encaran tareas memorísticas, con el objetivo de describir el proceso cognitivo que realizan. En general, no se enseñó a los sujetos la forma de lograr la tarea y, cuando se daban instrucciones, el objetivo era teórico: explorar los límites de las capacidades del sujeto antes que lograr una mejora en la memorización por sí misma. Este capítulo tiene un objetivo muy diferente: pondremos el acento en la mejora de la memoria y no en su comprensión teórica. Nos ocuparemos menos del psicólogo como científico y más del psicólogo como entrenador. Naturalmente, sólo se trata de una cambio en el énfasis, puesto que seguimos deseando que las investigaciones teóricas y aplicadas sobre la memoria se sostengan mutuamente.

## La tradición mnemónica y la terapia de la memoria

Hace mucho ya que existen los cursos para mejorar la memoria. Actualmente, los periódicos publican avisos sobre estos cursos, pero la explotación práctica y comercial del arte de la memoria tienen una venerable historia (Yates, 1966). Los sistemas mnemónicos fueron un elemento importante en Grecia y, más adelante, en la educación en la antigua Roma. En una época en que la educación estaba reservada para unos pocos y las telecomunicaciones aún no habían sido inventadas, la importancia social de la oratoria pública era mucho mayor de lo que es en la actualidad. Los mejores oradores evitaban el empleo de notas y, por lo tanto, eran muy valiosos los métodos para aprender de memoria la información, y así siguió siendo hasta la Edad Moderna. De hecho, viviendo en una cultura de lápiz y papel, de foto-

grafías y cintas grabadas, es difícil imaginar el papel vital de las técnicas mnemónicas en las sociedades no alfabetizadas, o en aquellas con prácticas de escritura diferentes o menos difundidas que en la nuestra. Por esto, la investigación de los "efectos colaterales" cognitivos de la alfabetización y sus usos sociales ha demostrado ser una fuente de intrigantes profundizaciones en la naturaleza de la memoria y el pensamiento.

El arte tradicional de la memoria y sus descendientes modernos, los cursos publicitados en los periódicos, se han concentrado en la adquisición de técnicas mnemónicas (y lo que ahora denominamos *metamemoria*) en personas normalmente saludables. En cambio, y con la excepción de trabajos como el descrito en el capítulo 7, la investigación moderna del mejoramiento de la memoria se ha centrado en los mnemónicamente discapacitados. Los retardados, aquellos que han sufrido daños cerebrales y los ancianos han sido estudiados desde el punto de vista de la terapia mnemónica, que busca reducir los problemas de la vida que sean consecuencia de las fallas anormales de la memoria. En este trabajo, hay un reconocimiento general de los ideales del estudio controlado y los análisis teóricos, pero en la práctica se necesita una marcada dependencia de las soluciones que ofrece el sentido común. La terapia de la memoria está diseñada para ayudar a pacientes individuales con problemas que, a menudo, son esencialmente cotidianos. Las dificultades que enfrentan los discapacitados son las mismas que tienen las personas normales pero magnificadas, y las soluciones necesarias pueden ser igual de prosaicas. Sin embargo, aunque los trabajos prácticos y de laboratorio actuales no van unidos, hay mucho para aprender sobre los procesos de la memoria a partir de su disolución en quienes sufren daño orgánico. La terapia de la memoria puede demostrar su utilidad para ayudar a los teóricos tanto como a los pacientes.

## Ayudas internas y externas para la memoria

Como es poco lo que se puede hacer para mejorar la etapa de retención de la memoria, las mejoras tienen que implementarse en la codificación o en la recuperación. Algunas ayudas mnemónicas (como la repetición pura) tienen efecto en la codificación; otras técnicas (como el ajuste del contexto y las condiciones imperantes para que resulten similares a los que había en el momento de la codificación) pueden considerarse importantes en la etapa de recuperación. Pero, en la práctica, muchas de las mejores técnicas mnemónicas operan tanto en la codificación como en la recuperación. Con

frecuencia, se emplea una estructura conocida para organizar la información en la codificación y luego se recurre a la misma estructura para su recuperación (ver las técnicas de Steve Faloon descritas en el capítulo 7). Un ejemplo de ello es la categorización. Se agrupan en categorías las palabras que se quieren recordar para su ensayo y luego, al recordarlas, también se recuperan según su categoría.

La distinción entre ayudas memorísticas internas y externas (Wilson y Moffat, 1984) se relaciona con esta función de codificación/recuperación, aunque no se superpone completamente. Las ayudas internas son las mnemotécnicas, como las imágenes del sistema de Lorayne para recordar nombres. Son internas a la psique del individuo y pueden funcionar en la codificación, en la recuperación o en ambas. Por su parte, las ayudas externas existen en el medio: las notas escritas para uno mismo, los relojes con alarma, los nudos en los pañuelos y las manipulaciones del medio (como dejar un paquete junto a la puerta de calle para no olvidarse de llevarlo al salir). Aunque algunas ayudas externas pueden actuar como pistas para la recuperación, su propósito más frecuente es promover acciones más que recuperar información.

Las ayudas externas nos recuerdan algo que hay que *hacer*, quizá sugiriendo sólo indirectamente qué es ese algo. Las ayudas internas *pueden* servir como promotoras de la acción (Lorayne y Lucas, 1974) pero esa no es su función usual, al menos en nuestra cultura.

Las ayudas externas suelen emplearse en la vida cotidiana para recordar que tenemos que hacer cosas. Pero hace muy poco tiempo que los psicólogos experimentales han comenzado a estudiar el empleo de ayudas externas y los métodos que la gente usa para ayudarse a sí misma a recordar lo que tienen que hacer (Harris, 1980a; 1984). En parte, esto se debe a la importancia de estos dos tópicos para la terapia de la memoria. También es indicador de la tendencia existente hacia los estudios del rendimiento de la memoria práctica (por ejemplo, Gruneberg, Morris y Sykes, 1978).

Una vez establecida una base para el estudio del mejoramiento de la memoria, veamos ahora algo de lo que se sabe sobre el tema.

## Mejora de la memoria de la gente normal

Como se dijo antes, el mejoramiento de la memoria de la gente normal se ha basado, por lo general, en el uso de una u otra técnica mnemónicas. Pero el dominio de un conjunto de técnicas mnemónicas no es lo único

que existe para la memorización inteligente. La capacidad de emplear estrategias (o técnicas mnemónicas) es tan importante como la comprensión de las circunstancias adecuadas para ponerlas en uso, que ya ha sido llamado *metamemoria* en el capítulo 6. Mejorar el rendimiento de la memoria en algún campo de actividad implica, en primer lugar, una evaluación cuidadosa de las exigencias memorísticas que presenta la actividad, las condiciones probables de recuerdo, etc. Una vez terminado este análisis, es posible comenzar a formular un plan de acción, evaluando la viabilidad de las diferentes estrategias y las formas en que éstas podrían adaptarse fructíferamente a la tarea en cuestión. La utilización real de una estrategia mnemónica es el último paso en el proceso de memorización inteligente. Por ejemplo, aprender un idioma extranjero requiere memorización en muchos niveles, desde las palabras hasta las reglas de conjugación y las normas gramaticales. Además, los objetivos y necesidades de los alumnos variarán según sus propósitos: ¿Quieren aprender a leer en la lengua extranjera o a hablarla? ¿O las dos cosas? ¿Necesitarán conversar socialmente o están dispuestos a limitarse a hablar de ciertos temas técnicos? Sin una evaluación de esa naturaleza, las técnicas mnemónicas no pueden proporcionar más que una ayuda fragmentaria. Pero, junto con el análisis completo de las tareas, la aplicación sistemática de las técnicas puede ser una ayuda efectiva para el aprendizaje del idioma (Gruneberg, 1985).

## Mnemotécnicas verbales

La estructura impuesta al material por las técnicas mnemónicas verbales suele ser la del ritmo o la de la rima, como en la rima para recordar los días del mes (capítulo 5). El ritmo es un principio natural de organización y muchos sujetos lo adoptan de manera espontánea en los experimentos de recuerdo de dígitos, agrupando de a tres o cuatro ítem que son presentados a un ritmo constante, para su ensayo. Los agrupamientos rítmicos fueron construidos a propósito en el diseño de los códigos postales británicos, la mayoría de los cuales tiene la forma letra-letra-dígito, dígito-letra-letra.

El uso más sorprendente del ritmo y la rima para propósitos mnemónicos se encuentra en las historias y genealogías orales de las culturas sin escritura, como en la epopeya homérica (Lord, 1982; Bateson, 1982). Los bardos expertos pueden recitar muchos cientos (e incluso miles) de versos. Y no lo hacen recordando sus canciones palabra por palabra sino componiéndolas

nuevamente en cada recitado, según fórmulas y temas. Al acotar el proceso de reconstrucción, el metro del verso actúa como guía para la memoria. En las sociedades orales no hay epopeyas en prosa precisamente porque la prosa carece de las propiedades mnemotécnicas del verso.

Naturalmente, la propia trama de la historia es un elemento mnemónico importante de la narración épica. Las virtudes educativas de la narrativa siempre han sido reconocidas y en todas las culturas abundan las fábulas, parábolas y cuentos morales. Si no hay metro, la narrativa adopta una escala menor que la épica, pero hasta la estructura de la historia más simple (comienzo-medio-final) tiene características mnemónicas. Así como los Kpelle demostraban una mejor memoria para las palabras incluidas en una historia que para las mismas palabras en una lista de vocablos aislados, también los occidentales recuerdan más palabras cuando se les pide que creen historias que las vinculen que cuando sencillamente repiten las palabras (Bower, 1972).

Además del ritmo y la narración, algunas técnicas mnemónicas verbales sustituyen las palabras o consonantes por números que se deben recordar. Estos sistemas se desarrollaron especialmente para ocuparse de la memorización de cosas tales como las fechas, aunque en algunos casos se las puede generalizar a otros campos. Hunter (1964) y Norman (1976) ofrecen algunos ejemplos, además de cierto marco histórico.

## Mnemotécnicas de imágenes visuales

Los recursos mnemotécnicos con imágenes visuales son los más comunes y utilizados. Quizá el más simple sea el que Hunter (1964) denominó sistema de comparaciones sucesivas. Se lo puede usar para recordar cualquier lista de ítem (como palabras) siempre que cada uno de ellos pueda representarse mediante una imagen visual. Para empezar, los dos primeros ítem deben imaginarse juntos tan vívidamente como se pueda. Luego, se produce una nueva imagen para vincular el segundo ítem con el tercero, y así sucesivamente. Este es un método de encadenamiento muy sencillo y se puede hacer cualquier clase de asociación entre ítem. El método puede generalizarse y usarse no sólo con palabras, sino siempre que se pueda representar todo con imágenes; así, también una secuencia de ideas o de acciones puede memorizarse con este método.

El recurso mnemotécnico de comparación sucesiva es el menos estructurado de todos los que se basan en imágenes visuales, ya que sólo tiene la estructura lineal de conexiones diseñada por el individuo. Otro sistema es el recurso de la "palabra asociada", para el cual hay que aprenderse una rima como la siguiente: "Uno es un tribuno, dos tiene tos, tres está al revés..." y así hasta el diez. El primer ítem que debe memorizarse, entonces, se imagina en interacción con un tribuno. Así, si el primer ítem fuera la palabra "ventana", se podría pensar en un tribuno romano atravesando de un salto una inmensa ventana. Las imágenes producidas deben ser tan detalladas y vívidas como sea posible. Después de este paso, se imagina el segundo ítem "con tos", y así sucesivamente. Este método conlleva cierta estructura, ya que incluye una secuencia hiperaprendida de numerales. Hunter (1964) demostró que un grupo de gente a la que se le enseñara esta técnica recordaría muchas más palabras que un grupo de control al que no se le enseñara. Igual que las comparaciones sucesivas, el método es generalizable a cualquier campo que pueda entrar en la imaginería visual.

El recurso mnemotécnico visual más común es también el más estructurado. Es conocido como método de loci, por la palabra latina *locus*, que significa "lugar". El método, que fue desarrollado en gran detalle por los oradores romanos, se basa en el conocimiento del plano de un edificio en particular, o de los hitos en un camino familiar. Si hay que aprender de memoria una lista de ítem, o una sucesión de ideas para un discurso, lo único que hay que hacer es un recorrido imaginario por ese edificio, depositando (por así decirlo) una imagen de cada idea en cada habitación, sucesivamente. Si el orden en que se recorren las habitaciones es siempre el mismo, las ideas podrán recuperarse serialmente repitiendo la caminata imaginaria. Cuanto más ricas son las imágenes empleadas, más confiable se torna el método, y si se imagina una habitación en detalle, puede depositarse más de una idea dentro de ella, siempre que cada una siga inmediatamente a alguna característica prominente. La esencia de este sistema es que una secuencia conocida de lugares proporciona una estructura en la que se pueden incorporar nuevas informaciones.

El Método de *loci* tiene muchos puntos a favor. El más importante es que es altamente efectivo, incluso cuando se lo emplea por primera vez. Sin embargo, como sucede con toda técnica de memorización, con la práctica del método se obtienen mejoras radicales.

### Por qué funcionan los recursos mnemotécnicos

En esta sección nos ocuparemos brevemente de dos preguntas que tienen que ver con los recursos mnemotécnicos. En primer lugar, ¿por qué funcionan estos recursos? En segundo lugar, ¿por qué las historias y las imágenes visuales tienen una importancia mnemónica tan grande?

La respuesta a la primera pregunta es que los recursos mnemotécnicos sirven porque imponen organización al material que se debe recordar. La estructura del sistema mnemotécnico proporciona un plan de recuperación que dirige la búsqueda de la información deseada: actúa como marco.

Esto nos lleva a la segunda pregunta: ¿por qué las historias y las imágenes visuales resultan métodos tan buenos para fijar la información en un determinado contexto? Una posibilidad es que, al crear o comprender una historia, la gente automáticamente genera imágenes del material, de modo que los dos métodos son, en realidad, el mismo. Pero, sean o no distintos, es probable que resulten efectivos porque ambos incluyen un procesamiento profundo y elaborado que, como vimos en el capítulo 5, es importante para determinar la memorabilidad. Una razón para creer que ésta puede ser la función esencial de la enseñanza de imágenes es que, aunque hay grandes variaciones individuales en la capacidad aportada para crear imágenes visuales, los efectos que produce son bastante uniformes.

## El mejoramiento de la memoria en los niños con retardo mental leve

Los recientes intentos para mejorar el rendimiento de la memoria de los niños con retardo mental leve ("educables") han estado estrechamente vinculados a las investigaciones sobre el desarrollo de la memoria en los niños normales (ver capítulo 6). Antes, se pensaba que los niños retardados eran neurológicamente diferentes de los normales porque su cerebro no retenía tan bien la información ya codificada, pero el enfoque más reciente considera que el déficit aparece en el mismo proceso de codificación (Kail, 1984). Los niños levemente retardados suelen tener el mismo rendimiento (o casi) que los niños normales en las pruebas de reconocimiento de la memoria. Pero cuando se les pide que realicen la memorización más activa propia de las pruebas de recuerdo, se observa una disminución del rendimiento. Se ha descubierto que los adolescentes con retardo no tienden a involucrarse

en ensayos espontáneos para preparar el recuerdo (Belmont y Butterfield, 1971) y no suelen organizar su recuerdo en categorías (Spitz, 1966). En otras palabras, en estas pruebas los adolescentes con retardos leves demostraron la misma falta de estrategias mnemotécnicas que habían exhibido los niños normales muy pequeños.

Como en el caso de los niños normales pequeños, la siguiente pregunta era si se podía entrenar a los adolescentes con retardo para que emplearan estrategias mnemotécnicas. Una vez más, la respuesta fue afirmativa. Belmont y Butterfield pudieron entrenar a los retardados para que realizaran ensayos y su rendimiento mejoró como consecuencia de ello. Sin embargo, después de una semana el efecto del entrenamiento había desaparecido por completo. Pero cuando se realiza un entrenamiento a largo plazo, sus resultados pueden ser más duraderos y se automatiza —o "hiperaprende"— el uso de estrategias (Brown, Campione y Murphy, 1974). De modo que se puede mejorar el rendimiento de los retardados tanto en el corto como en el largo plazo. Y cuando no sólo se ejercita una estrategia sino que también se explica su objetivo a los niños retardados y se les enseñan hábitos de autocontrol que los niños normales adquieren espontáneamente, se obtienen resultados aún mejores (Wong y Jones, 1982; ver también el capítulo 12). Esos métodos también crean mayores posibilidades de transferir el entrenamiento.

Esa transferencia consiste en ver los nuevos problemas como análogos a los viejos y comprender que las habilidades que ya se poseen pueden utilizarse para resolverlos. El logro de esa transferencia es un problema perenne de toda forma de educación y entrenamiento, pero en las personas con retardo suele ser particularmente pobre. Los primeros estudios hallaron que incluso las habilidades de ensayo simples no se generalizaban espontáneamente a estrategias que apenas diferían del original. Sin embargo, cuando el entrenamiento en estrategias va acompañado de discusión, sugerencias y entrenamiento en habilidades metacognitivas como el autocontrol, se observa una mejor capacidad de transferencia (Campione y Brown, 1984; Wong y Jones, 1982). Lo mismo sucede con los niños y adultos normales, pero en el caso de las personas con retardo la transferencia tiende a ser más limitada, exige ayudas mayores y más explícitas que en los demás y puede fracasar por completo.

Para resumir esta sección: la expresión "con retardo" se aplica a las personas de la que hemos estado hablando. El curso de su desarrollo es normal pero mucho más lento y por lo general se detiene en una etapa temprana. El desarrollo de las habilidades cognitivas se ve inhibido a todo nivel y es-

pecialmente en las habilidades superiores. El aumento del rendimiento en las tareas de memoria simples es posible y se pueden lograr *insights* frágiles en su propio funcionamiento cognitivo, pero estos éxitos son escasos y sigue siendo difícil consolidarlos.

## Terapia de la memoria para aquellos que han sufrido daño cerebral

En respuesta a los hallazgos de laboratorio, se ha enseñado a pacientes que han sufrido daño cerebral a codificar la información profundamente, a recurrir a sistemas de "palabra asociada", a emplear el método de loci, a agrupar la información, a aplicar la técnica mnemónica de las primeras letras para listas de compras, a intentar recordar nombres con el sistema de Lorayne y a recurrir a técnicas de estudio desarrolladas para ayudar en el aprendizaje escolar. Todos estos métodos parecen resultar útiles para algunos pacientes en algunas tareas, y su efectividad ha sido evaluada por Powell (1981) y por Wilson y Moffat (1984).

Parte de la destreza del terapeuta radica en adecuar un método en particular a un paciente en particular. Wilson y Moffat sugieren que los pacientes con daño en uno solo de los hemisferios cerebrales, como suele suceder en las apoplejías, pueden adecuarse selectivamente a sistemas mnemónicos apropiados. Se puede alentar a quienes han sufrido daños en el hemisferio cerebral izquierdo y tienen dificultades con el habla a que adopten mnemotécnicas visuales, mientras que a aquellos que han sufrido daños en el hemisferio cerebral derecho y tienen su capacidad de imaginación reducida se les pueden enseñar mnemotécnicas verbales. También se sugiere una adaptación en los pacientes que tienen dificultades específicas para materiales presentados de manera auditiva o visual. En ambos tipos de ajuste, se aprovecha el hecho de que aquello que ayuda a la gente normal a recordar también ayudará, en conjunto, a los que han sufrido daño cerebral. Pero no siempre sucede así. Por ejemplo, varios estudios han hallado que el sistema de memorización de nombres de Lorayne resulta demasiado complejo para el aprendizaje de los pacientes con graves dificultades con los nombres.

Quienes no tengan experiencia con pacientes con daño cerebral pueden tener dificultades para imaginar la magnitud de los problemas que a veces hay que enfrentar. Un ejemplo puede ayudar. Wilson y Moffat describen el caso de un paciente que podía recordar los nombres de tres de sus fisioterapeutas sólo si le decían con qué letra empezaba. Una cuarta fisioterapeu-

ta, llamada Anita, le dijo que podría recordar su nombre fácilmente porque ella tenía *a neater way of doing her hair*[1], y sacudió su cabello. Después de esto, cada vez que ella repetía el gesto, él podía recordar su nombre. Por desgracia, hacía lo mismo con cada mujer que sacudiera su cabello frente a él. El nombre se había ligado a la acción en vez de unirse a la persona.

Cuando se trata de individuos imposibilitados, las mínimas innovaciones pueden devengar, a veces, grandes beneficios, tanto para los pacientes como para quienes los atienden. Harris (1980b) describe cómo en una guardia (más geriátrica que de pacientes con daño cerebral) se redujo la tasa de incontinencia pintando las puertas de los baños de un color distinto al de todas las demás puertas. Esta señal ambiental promovió una conducta adecuada que, de otro modo, no se hubiera producido. De manera similar, los hospitales y centros de rehabilitación suelen recurrir a líneas de tiza en las paredes y el piso para guiar a los pacientes de un sector del edificio a otro. Las señales ambientales de esta clase no son diferentes en principio de las que facilitan muchas conductas automáticas y semiautomáticas en la vida cotidiana. Actúan como ayudas externas y públicas para la memoria (ver capítulo 2).

Wilson y Moffat sugieren que las ayudas externas para la memoria son especialmente aplicables en los casos de amnesia global. Algunas ayudas "obvias", como las agendas de bolsillo, son poco útiles para estos casos porque los pacientes olvidan consultarlas. Una forma de superar este obstáculo es una alarma que suene de vez en cuando y le recuerde así al paciente que debe consultar su agenda. Fowler, Hart y Sheehan (1972) hallaron que después de varias semanas el hábito de consultar la lista se había arraigado profundamente y se podía prescindir de la alarma. En la actualidad, se venden ayudas electrónicas para la memoria, que pueden almacenar información permanente, como el nombre del paciente, su domicilio y demás, y también se pueden programar con mensajes diarios importantes y proporcionar señales recordatorias auditivas. Jones y Adam (1979) han evaluado estas ayudas. En muchos sentidos, muestran pocas ventajas sobre ayudas menos sofisticadas en lo tecnológico, pero se esperan mejoras que las acerquen a tres criterios que Harris ha establecido para lograr ayudas efectivas para la memoria. Estos tres criterios son:

a) la ayuda debe tener su efecto lo más cerca posible antes de la acción deseada;

b) debe ser un recordatorio activo y no pasivo;

---

*1. La pronunciación de "a neater" en inglés es muy similar a la del nombre Anita. "A neater way of doing her hair" significa: un peinado más prolijo. (N. del T.)*

c) debe especificar la acción requerida (a diferencia, por ejemplo, de un nudo en un pañuelo).

## La mejora del testimonio legal

Un campo importante de la investigación en memoria aplicada es el del testimonio legal. Los psicólogos ayudan al menos de dos formas en los procesos legales de evaluación de evidencias. En primer lugar, intentan influir en los procedimientos policiales y de justicia a la luz de los hallazgos de las investigaciones. En segundo lugar, los psicólogos pueden actuar como testigos expertos que aconsejan al juez y al jurado analizando los factores que pueden haber influido en un testimonio en particular. En la práctica, estas dos funciones se superponen y entre ambas cubren una enorme gama de tópicos, de los que apenas mencionaremos unos pocos aquí.

### El mejoramiento del sistema

Davies (1981) ha reseñado la historia de los equipos para recuerdo de rostros (como el *identikit*) y subrayado su bajo nivel de confiabilidad. En parte, atribuye su debilidad a que recurren a la memoria de rasgos más que del conjunto del rostro. Las técnicas más modernas, basadas en el método de "rostros enteros" y ayudadas por la flexibilidad de los gráficos de computadora, prometen mejores resultados para el futuro.

La identificación hecha a partir de "modelos" o "fotos de prontuario" suele llevar a la realización de una ronda de reconocimiento, y éste es probablemente el procedimiento legal más estudiado por los psicólogos. La probabilidad de hacer una identificación correcta está sujeta a la influencia de muchos factores (Deffenbacher y Horney, 1981). La composición del grupo puede influir en el resultado. Si un testigo busca un hombre rubio y alto, y en el grupo sólo hay una persona con esas características y los demás son mujeres pequeñas y de cabello oscuro, es muy probable que el testigo lo señale. Pero aun cuando se eviten situaciones absurdas como esa, la composición del grupo que se someterá a una ronda de reconocimiento no es tema trivial. Por ejemplo, un sospechoso que se diferencia del resto del grupo por detalles más sutiles, como su ropa, su postura o su modo de caminar, corre el riesgo de sufrir una identificación incorrecta. La cantidad de personas que inte-

gra el grupo también influye en los porcentajes de detecciones verdaderas y falsas, así como lo que piensa el testigo sobre el grupo: que el sospechoso "indudablemente está" o que "puede estar" entre el grupo. En este último caso, hay menos presiones para llegar a una identificación positiva que en el anterior, y esto puede ser muy importante cuando los testigos buscan un miembro de un grupo étnico para el que su capacidad de discriminación está menos agudizada.

Deffenbacher y Horney (1981) cuentan que en Suecia, si el hecho fue visto originalmente con una visibilidad inferior a la óptima, para el proceso de identificación se reproducen las condiciones ambientales. Además, en la ronda de reconocimiento se puede hacer que el testigo recorra una habitación llena de gente que conversa. Así, no sólo hay características faciales y posturales para ayudar a la memoria del testigo sino también peculiaridades gestuales o faciales y características de la voz. Estos dos procedimientos no tradicionales actúan como una reproducción en el momento del reconocimiento de mayor cantidad de pistas para la recuperación. Proporcionando información adicional a los testigos, es más probable reconocer a los sospechosos culpables y notar que los inocentes difieren de algún modo del criminal original.

## El testigo experto

Más que ayudar en los procesos de recolección de evidencias previas a un juicio, el psicólogo puede ayudar en el propio juicio. Se le pueden señalar al juez y al jurado ciertos factores influyentes en el testimonio de un testigo. Por ejemplo, se han hecho experimentos en los que los sujetos: (a) presencian un "crimen", (b) ven una serie de "fotos de prontuario" y (c) se les pide que elijan al criminal en una ronda de reconocimiento. Los resultados demostraron que era más probable que eligieran en la ronda a un "inocente" que habían visto entre las fotos de prontuario que a un individuo cuyo rostro no hubiera aparecido. En otras palabras, los rostros que son conocidos a partir de la sesión de "fotos de prontuario" se confunden con el rostro (o los rostros) del crimen. Así, la identificación de un acusado por un testigo que, en el lapso transcurrido desde el crimen, ha visto fotos del acusado en prontuarios o en los periódicos, es menos confiable que la identificación hecha por un testigo que no ha visto nada de eso.

El psicólogo también puede informar a la corte sobre las investigaciones que muestran que la confianza de los "testigos" experimentales en su testimonio no tiene relación con su precisión literal (Deffenbacher, Brown y Sturgill, 1978). También puede haber ocasiones, además de aquellas que podría observar cualquier abogado competente, en las que puede ser necesario señalar de qué forma la reconstrucción de los acontecimientos de parte de un testigo puede haber sido "guiada" por el tipo de interrogatorio realizado. Por ejemplo, Loftus y Palmer (1974) demostraron que el recuerdo de los sujetos de la velocidad a la que iban dos autos que chocaban, en un accidente filmado, dependía de si se les preguntaba a qué velocidad iban los autos cuando se *estrellaron* (65 km por hora), *chocaron* (63 km por hora), se *golpearon* (61 km por hora) o *entraron en contacto* (51 km por hora).

## Ayudas para la memoria en los exámenes

Para terminar este capítulo sobre el mejoramiento de las habilidades de la memoria, nos ocuparemos de un tópico que probablemente resulte importante para la mayoría de los lectores del libro: cómo mejorar el rendimiento en los exámenes.

Beard y Hartley (1984) reseñaron las investigaciones existentes sobre la preparación para los exámenes y los exámenes mismos. Para este capítulo, las ideas fundamentales de estos autores son:

a) que los alumnos deben *organizar* sus actividades de revisión;

b) que los alumnos deben revisar *activamente* en vez de hacerlo pasivamente;

c) que los alumnos necesitan *práctica* de exámenes si quieren mejorar su habilidad en este sentido.

En cuanto a la organización, se sugiere que los alumnos estudien exámenes anteriores para ver cuáles son los requisitos de cada examen en particular, tanto en lo que hace a su estructura como a sus contenidos. Beard y Hartley también sostienen que los alumnos deben pensar en horarios de revisión para poder planificar cómo espaciar su revisión sistemáticamente en vez de repasar apresuradamente todo la noche anterior al examen. (Las ventajas de esta práctica de espaciamiento fueron analizadas en el capítulo 3.) La revisión de último momento puede resultar útil (en especial si antes se ha hecho revisión espaciada) porque:

- se pueden unir los puntos sueltos;
- se pueden obtener perspectivas del curso completo;
- es más eficiente refrescar finalmente la memoria a partir de notas.

Por su parte, la revisión espaciada:

- asegura una cobertura más completa del programa;
- facilita una tarea que intimida, dividiéndola en partes manejables;
- supera el problema del cansancio, las enfermedades, etc., que puede intensificarse o aparecer hacia el final de un curso (James, 1967).

Beard y Hartley también ofrecen una lista de estrategias de revisión activas, que incluyen las bien conocidas actividades de subrayado, resumen y escritura de respuestas a preguntas de exámenes, ya sea en forma de notas o como respuestas completas. También incluyen estrategias menos frecuentes, como la explicación de tópicos a compañeros que no están familiarizados con ellos y las reuniones en grupos pequeños para repasar temas específicos, elaborar respuestas para determinadas preguntas y preparar bosquejos sobre tópicos específicos. Las ventajas del uso de estas estrategias son las que se esbozan en los capítulos 5 y 7: cuanto más profundamente se procese el material y más completamente se lo organice e interrelacione, mejor se lo recordará.

Finalmente, Beard y Hartley señalan que el simple hecho de permitir que los alumnos practiquen los exámenes puede mejorar su rendimiento, y que es posible considerar los exámenes como una habilidad. Como todas las habilidades, hay ciertas subrutinas útiles que conviene practicar, entre las cuales se cuentan la lectura exhaustiva de la pregunta y la recuperación del material adecuado. La práctica, y la automaticidad que de ella se deriva, es especialmente importante, porque los alumnos que rinden exámenes frecuentemente están ansiosos, y la ansiedad interrumpe la ejecución fluida de subrutinas (ver capítulo 2). La tabla 8.1 propone algunas técnicas valiosas para que los alumnos puedan practicar.

*Tabla 8.1. Notas sobre técnicas para exámenes (basadas en sugerencias de A.B. Hill).*

1. No llegues más de 10 minutos antes de que comience el examen, a menos que sea imprescindible.

2. Cuando llegues, no hables con tus amgios sobre el repaso que tú o ellos hayan hecho.

3. Verifica las instrucciones y luego lee *todas* las preguntas dos veces, lenta y cuidadosamente, antes de elegir las que contestarás.

4. Piensa cuánto tiempo tienes para contestar cada pregunta. Por ejemplo, si tienes que responder tres preguntas en dos horas, significa que dispones de 40 minutos para cada pregunta. *Respeta este límite estrictamente*. Si te queda tiempo después de contestar todas las preguntas, completa las que hayas dejado inconclusas por tu límite de tiempo. Si terminas todas las preguntas y te sobra tiempo, relee tus respuestas y corrígelas.

5. Asegúrate de responder las preguntas obligatorias.

6. Verifica cuidadosamente qué te exigen las preguntas elegidas. Por ejemplo, si la pregunta dice "Evaluá..." o "Realiza un análisis crítico de..." asegúrate de que tu respuesta sea una *evaluación* o *una crítica*. Es fácil dar una respuesta que no es relevante para la pregunta si uno no se asegura exactamente qué es lo que se le pide.

7. Para que tus respuestas sean relevantes y coherentes y para asegurarte de decir todo lo que puedes o todo lo que quieres, confecciona una lista que incluya los puntos importantes *antes de empezar a escribir tu respuesta*. Numéralos en el orden que quieres que tengan y táchalos cuando los vayas incluyendo en la respuesta.

8. *Asegúrate* de responder *todas* las preguntas; es decir, si tienes que responder tres preguntas, *asegúrate* de intentar responder las tres y no sólo dos o dos y media. Si respondes sólo dos de tres preguntas, tu calificación máxima posible (si tus respuestas son perfectas) será del 66,6%.

9. Si puedes elegir entre preguntas que requieran conocimientos específicos y preguntas generales, elige las preguntas específicas a menos que tengas plena confianza en tu capacidad para responder a las generales. Es más fácil responder adecuadamente en ellas.

10. Si te pones nervioso o entras en pánico, intenta calmarte haciendo dos o tres minutos de práctica de relajación.

11. Después de salir del examen, no comentes con tus amigos las preguntas que elegiste o las respuestas que diste.

## Resumen

En este capítulo analizamos cómo puede aplicarse lo que se sabe sobre la memoria humana para mejorar su rendimiento en diferentes grupos de gente, en distintas circunstancias. Los recursos mnemotécnicos usados para amplificar los recuerdos de la gente común pueden entenderse dentro del sistema de habilidades de memoria presentado en este libro, y los datos sobre el rendimiento de los que tienen retardo leve también se ajustan bien a la teoría. En el campo de la terapia de la memoria para aquellos que presentan daño cerebral, la teoría ofrece cierta comprensión de los déficit adquiridos pero las ayudas para su mejora deben pensarse en cada caso individual. De manera similar, el rendimiento de los testigos oculares se ve influido por múltiples factores, de los cuales sólo algunos tienen, hasta ahora, explicación teórica.

## Lecturas adicionales

LORAYNE, H. y LUCAS, J. (1974) *The Memory Book*. Nueva York: Stein and Day.

Fascinante análisis de las formas en que todos podemos mejorar nuestra memoria.

HARRIS, J.E. y MORRIS, P.E. (1984) *Everyday Memory, Actions and Absent-Mindedness*. Londres: Academic Press.

Compilación de capítulos sobre diversos aspectos de la memoria cotidiana, incluyendo trabajos con pacientes que han sufrido daño cerebral.

# *Sección 3*

## *Las habilidades de comprensión y comunicación*

# 9

# Las habilidades del hablante

*Don Rogers*

En muchos sentidos, hablar es la habilidad humana más característica. Es exclusivamente humana y, además, demuestra fehacientemente muchos de los rasgos de otras habilidades motrices e intelectuales que se comentan en otras partes de este libro (ver especialmente los capítulos 2 y 3).

En el mundo se hablan cientos de lenguas diferentes. Algunas sólo cuentan con unos pocos hablantes nativos en tanto que otras son habladas por millones de personas. Algunos de estos idiomas son muy diferentes de otros: por ejemplo, los que tienen una tradición escrita muy desarrollada tienen también grandes cantidades de palabras poco usadas, que la mayoría de los hablantes de ese idioma no conocen (cosa que se puede comprobar mirando cualquier diccionario grande de castellano); algunos tienen un vocabulario especializado bien desarrollado, del que otros idiomas carecen (por ejemplo, las lenguas esquimales tienen muchas palabras para los diferentes tipos de nieve, mientras que los idiomas de regiones tropicales no las necesitan); y algunos idiomas usan recursos gramaticales muy diferentes. Por ejemplo, idiomas como el inglés indican las relaciones gramaticales como sujeto y objeto mediante el orden de las palabras: *John liked Mary* es diferente de *Mary liked John*, mientras que para los idiomas como el ruso o el latín el orden de las palabras es menos importante y sujeto y objeto se indican mediante las diversas terminaciones de sus sustantivos.

Si bien los idiomas muestran diferencias entre sí, éstas son comparativamente menores. Hay ciertos puntos importantes que se aplican al lenguaje en general y que serán el tema de este análisis de las habilidades del habla. Pero antes, vale la pena hacer dos aclaraciones preliminares. En primer lugar, todos los humanos pueden aprender a usar el lenguaje, aunque con excepciones, ya que algunas formas de evolución anormal pueden asociarse con la incapacidad para usarlo; pero de todas formas es sorprendente que los niños sordos profundos puedan aprender a usar lenguaje de signos, que haya casos de niños que han aprendido a comunicarse lingüísticamente aunque son ciegos y sordos y que existan otros niños que han estado encerrados y no han

hablado los primeros cinco o seis años de su vida pero que, al entrar en contacto con un medio normal, rápidamente han aprendido a hablar de una manera relativamente normal. (Ver Clarke y Clarke, 1976, por detalles sobre algunos de esos estudios.)

En segundo lugar, no hay especie, además de la humana, que parezca capaz de usar el lenguaje. Los loros pueden imitar el habla, los perros saben expresar emociones y deseos, pero nada de esto se acerca al lenguaje. Lo más cercano que parece haber al lenguaje humano es el caso de los chimpancés y otros primates a los que se les ha enseñado a usar lenguaje de signos: el sistema de comunicación usado por los sordomudos en los Estados Unidos. Al principio, estos estudios alimentaron las esperanzas de que los animales podrían aprender a usar el lenguaje de manera plena, pero actualmente parece que, si bien los chimpancés pueden aprender unos cuantos cientos de signos (más o menos equivalentes a palabras), no pueden combinarlos rápidamente para producir "oraciones" nuevas (Terrace y otros, 1980).

Esta productividad combinatoria es lo más impresionante del lenguaje. El lenguaje es creativo. Podemos producir fácilmente una oración que nadie ha dicho antes y no nos sorprende que esa oración resulte comprensible.

El lenguaje obtiene este recurso productivo de la forma en que se estructura. Cualquier lenguaje hablado tiene unos pocos sonidos que pueden combinarse para producir miles de palabras, que a su vez pueden combinarse según la gramática de la lengua y producir infinitas oraciones. Los hablantes de una lengua pueden fácilmente producir oraciones "legítimas" desde el punto de vista de la gramática y pueden entender fácilmente esas oraciones.

Los usuarios del lenguaje pueden ser tanto hablantes como oyentes: productores y receptores de lenguaje. En este capítulo nos ocuparemos de la producción del habla y, en el próximo, de la comprensión.

## Hablar como habilidad

Habiendo visto que el lenguaje es una característica exclusivamente humana, analicemos ahora en qué sentido se puede decir que hablar es una habilidad. Recordemos que en el capítulo 2 se señaló que la persona experta es aquella que actúa con fluidez, rapidez, automáticamente (es decir, sin pensar conscientemente en todos los aspectos de la tarea), puede integrar diferentes componentes de la actividad simultáneamente y producir (habitual-

mente) el componente preciso en el momento justo. Lo mismo puede decirse del hablante experto.

## Fluidez

Cuando un niño pequeño comienza a hablar, suele empezar diciendo de a una palabra por vez. Por ejemplo, si un niño de unos 18 meses ve a un perro levantando una pelota, quizá diga algo parecido a "perro" o a "pelota". Si tiene más experiencia con el habla, podrá usar ambas palabras, pero no juntas, así que quizá diga: "Perro... pelota". Con el tiempo, podrá unir las dos palabras en una sola emisión: "Perro, pelota"; "Perro, mordió"; etc. El niño mayor puede unir varias palabras en una sola emisión, con una sola curva de entonación, para expresar un pensamiento: "El perrito malo me lamió la mano". Como veremos en este capítulo, el discurso fluido de los adultos no carece de vacilaciones y pausas —según Goldman-Eisler, "dos tercios del lenguaje hablado vienen en grupos de menos de seis palabras" (ver Beattie, 1983)— pero de todos modos, aunque no sea totalmente fluido, el lenguaje de un adulto muestra muchas más palabras entre las pausas que el discurso de un niño pequeño. Y, por supuesto, como se vio en los capítulos 2 y 3, este agrupamiento de salida es una característica típica de todo tipo de habilidad.

## Rapidez

Vacilaciones aparte, el hablante de inglés normalmente dice entre dos y tres palabras por segundo. Cada palabra tiene un promedio de cinco sonidos, lo cual significa que el hablante produce unos 12,5 sonidos por segundo. Si cada uno de estos sonidos fuera independiente de los demás, la velocidad de decisión debería ser imposiblemente rápida. Pero, naturalmente, no son independientes: los sonidos se agrupan en palabras; las palabras, en frases; las frases, en proposiciones (una oración puede estar integrada por una proposición o más, cada una de las cuales contiene un verbo). Los sonidos de una palabra y las palabras de una proposición se relacionan entre sí de manera tal que, juntos, los sonidos producen una palabra significativa o las palabras producen una cláusula significativa. Y, como veremos, en vez de tomar una decisión separada independiente sobre cada sonido, el hablante planifica por adelantado, exhibiendo la anticipación que también se observa en todas las demás actividades expertas.

## Automaticidad

Una característica del desempeño experto es que el que lo realiza es inconsciente de los detalles de su actuación: sólo le preocupa la planificación de nivel superior y los detalles se resuelven mediante un "piloto automático". Como veremos, el habla posee la misma característica jerárquica: el hablante se ocupa de contar una historia, transmitir información, convencer a alguien de su inocencia o lo que sea; no planifica conscientemente el uso del plural o del singular para el sustantivo que está usando. Esto se lo deja a una subrutina bien ejercitada; y, si no hay perturbación, intoxicación o dificultad alguna; suele ser lo adecuado.

## Simultaneidad

Igual que manejar un auto, hablar implica hacer varias cosas diferentes al mismo tiempo. Esto es evidente cuando alguien habla y, simultáneamente, mueve sus manos, su cuerpo y sus ojos y cambia su expresión facial. Se dice: "Hablamos con nuestras bocas pero conversamos con todo el cuerpo" (ver capítulo 11).

La simultaneidad también puede verse en la forma en que anticipamos lo que vamos a decir mientras decimos otra cosa. A veces esto nos lleva a cometer errores y decimos, por ejemplo, "el jran jinete" en vez de "el gran jinete", pero habitualmente no suele suceder; como la anticipación siempre está presente. Un ejemplo sencillo mostrará una forma de anticipación: diga la palabra "cruzo"; luego, diga "craso". Puede observar que sus labios ya se están cerrando para formar el sonido "u" mientras pronuncia la "r" o incluso la "c" inicial de "cruzo", mientras que se abren al decir esos mismos sonidos cuando les sigue la "a" de "craso" (Fodor, Bever y Garrett, 1974).[1]

Como uno planifica por adelantado lo que va a decir mientras está diciendo otra cosa, es evidente que existe cierta forma de simultaneidad dentro mismo de diferentes aspectos de la habilidad del lenguaje. Además, naturalmente, la gente puede hablar mientras camina, cocina, maneja o realiza casi cualquier otra actividad.

---

*1. Los ejemplos usados en castellano en este capítulo y en otros son equivalentes a los del original en inglés. (N. del T.)*

### Conocimiento

El último aspecto de la habilidad que se menciona en el capítulo 2 incluye la observación común de que el individuo experto puede traer un ítem de la memoria durante la ejecución de una serie de acciones, pero en otra ocasión bien puede no traerlo. Evidentemente, lo mismo sucede en el caso del habla. A pesar de las vacilaciones y de las raras ocasiones en que no recordamos alguna palabra, el discurso avanza con bastante fluidez y la palabra adecuada se produce en el lugar correcto de la oración en la mayoría de los casos. Pero esto no sucede cuando a uno le piden que diga una palabra exclusivamente a partir de su significado. Mientras escribo esto, mi hija intenta resolver un crucigrama y de vez en cuando me pregunta, por ejemplo: "¿Qué palabra tiene siete letras, significa cansar y empieza con *d*?" Por lo general, no puedo contestarle.

Producir la palabra correcta en el momento adecuado al decir una oración nos exige planificar por adelantado. En el capítulo 2, esto se analizó en términos de un "apilamiento de objetivos". Uno podría ocuparse de elaborar, por ejemplo, el apilamiento de objetivos involucrado en el relato de la historia de Cenicienta.

Hablar es, entonces, una conducta experta no muy distinta de las demás conductas expertas. Se necesita mucho tiempo para adquirir habilidades: se dice que se necesitan unas 10.000 horas de práctica para llegar a ser un experto, ya sea un maestro de ajedrez o un compositor. Pero, después de todo, 10.000 horas es menos de tres horas por día durante 10 años, de modo que al terminar la escuela primaria los niños ya han tenido práctica suficiente como para ser expertos en las habilidades del lenguaje.

## El arte de hablar

Hay varios problemas en el estudio del discurso y en particular en la observación de los procesos cognitivos involucrados en la decisión del hablante sobre lo que va a decir. Es evidente que el hablante tiene muchas posibilidades abiertas ante sí: elecciones sobre el tópico, elecciones sobre la selección de puntos relevantes que se quieren destacar, cómo se los ordenará y vinculará para contar, por ejemplo, una historia o una broma, elecciones sobre cómo expresar un punto en particular y cómo afectará al oyente, elecciones de estructuras gramaticales específicas, elecciones de palabras y frases

entre las muchas que pueden transmitir la misma idea, elecciones sobre el tono de voz y el acento adecuados, etc. En el próximo capítulo veremos que los hablantes deben tener en cuenta qué saben y qué creen sus oyentes, deben enmarcar su discurso según las necesidades y propósitos que le atribuyen al oyente y deben intentar hacer que su discurso sea socialmente adecuado (no contar bromas en un funeral, por ejemplo). Pero por el momento nos bastará con observar que las elecciones del hablante no sólo afectan la medida en que el oyente puede comprender lo que se dice sino también la medida en que el oyente estará interesado, le gustará el hablante, la condición social que le atribuirá, etc.

Precisamente porque hay tantas elecciones abiertas a los hablantes y tantos factores que afectan esas elecciones, es extraordinariamente difícil para el psicólogo entender los procesos subyacentes. En este capítulo analizaremos tres técnicas que se han empleado para examinar los procesos cognitivos involucrados cuando un hablante habla.

En la primera técnica, se les da a los individuos alguna tarea: describir sus departamentos, contar una historia o una película, dar instrucciones sobre el camino que se debe seguir para llegar a algún lugar, etc., y se examinan los contenidos de sus discursos. En la segunda de las técnicas, se supone que los procesos involucrados en la toma de decisiones se reflejarán en dudas y pausas en el discurso, de modo que se examinan estas dudas y pausas para descubrir con qué frecuencia y en qué circunstancias se producen. En la tercera técnica, se examinan los errores del habla, con la esperanza de que si se observan los casos en que algo sale mal en el desempeño de alguna acción experta, se pueden hacer deducciones sobre el caso normal en que las cosas salen bien. Estas tres técnicas se ajustan con bastante naturalidad a diferentes niveles de planificación del discurso: el análisis de contenidos es más útil al examinar las elecciones hechas en un nivel global de planificación estratégica, en el que se toman decisiones sobre los puntos que se plantearán y el orden en que se los mencionará; el análisis de las dudas resulta más útil en un nivel intermedio de planificación y el análisis de errores es útil para examinar la planificación puntual en el nivel de la frase o la proposición.

## Planificación de largo alcance

Un hablante que se embarca en una descripción es, en cierto sentido, como un explorador, ya que tiene que planificar su ruta por adelantado. En

particular, tiene que pensar en: a) qué puntos debe mencionar; b) en qué orden debe mencionarlos; c) cuál es el nivel de detalle adecuado. El proceso puede ser altamente deliberado y bien estructurado, como en la escritura de un plan para una composición (ver capítulo 12), pero por lo general no es así; el pensamiento involucrado parece ser extremadamente rápido.

Un estudio clásico de la planificación es el de Linde y Labov (1975), en el que se grabaron las respuestas de un grupo de neoyorquinos a la pregunta: "¿Puede contarme cómo es su departamento?". Este estudio es muy atractivo, porque es real: aunque los datos sirvieron para el análisis lingüístico, fueron grabados como si fueran parte de una encuesta sobre la vida en la ciudad de Nueva York.

Todos los que respondieron resolvieron los tres problemas antes mencionados de la misma manera. El 97% de los encuestados realizó una recorrida imaginaria por su departamento comenzando por la puerta de entrada y avanzando sistemáticamente por el lugar, recorriendo por turno cada habitación, que sólo cruzaban si había otra al otro lado. Destacaban más las habitaciones grandes que las chicas, a las que veían como accesorias: "Allí derecho hay un dormitorio con un pequeño placard".

Todos los que respondieron dieron por sentado que la pregunta requería un nivel particular de detalle para las habitaciones. Esto determinó en parte el contenido de las descripciones, pero algunas habitaciones eran consideradas más importantes que otras. Finalmente, el problema de ordenar los contenidos se resolvió imaginando una recorrida por el departamento.

Otras descripciones requieren soluciones diferentes, pero los supuestos sobre qué es lo que resulta más comprensible, útil e interesante para el oyente afectan las descripciones de manera tal que, con frecuencia, distintas personas llegan a soluciones similares. Por ejemplo, el relato de una serie de acontecimientos por lo general exige una descripción cuyo orden es el mismo de la serie de eventos (ver, por ejemplo, Levelt, 1981). Esto no es obligatorio y los relatores de cuentos pueden crear efectos peculiares si no observan esta regla.

El nivel estratégico de planificación implica: decisiones sobre los detalles que son adecuados (si alguien pregunta cómo llegar a algún lugar, no es necesario describir todos los negocios por los que pasará); decisiones sobre, por lo menos, algunos de los puntos principales que se deben incluir; decisiones sobre el orden en que se los debe mencionar. Puede haber desviaciones respecto del nivel esperable de detalle, los puntos principales o el orden, pero esas desviaciones tendrán consecuencias en la interpretación del oyen-

te. Es parte del arte del narrador de historias (o, más aún, de quien cuenta bromas) ser capaz de predecir el efecto que tendrán las alteraciones de las expectativas del oyente (ver capítulo 4).

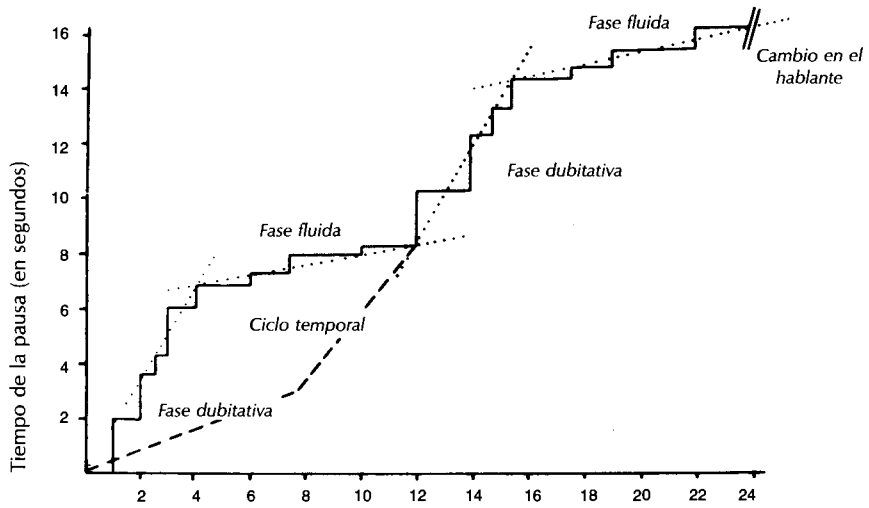
## Planificación de alcance intermedio

El desarrollo de un plan general es importante pero, como bien lo sabe cualquiera que escriba composiciones, no resuelve un montón de problemas sobre la ejecución detallada de ese plan.

Al escribir una composición, el paso siguiente después de terminar el plan general es hacer una pausa para pensar; luego, escribir algunas oraciones; luego, hacer otra pausa para pensar, etc. Según Butterworth (por ejemplo, 1980), esto también es válido para el habla, que, según descubrió, tiende a producirse en ciclos, con un período marcado por muchas vacilaciones, seguido por una frase fluida, seguido por otra frase vacilante, etc. Esto puede verse claramente en la figura 9.1. Según Butterworth, los ciclos usualmente "duran alrededor de 18 segundos, pero hay algunos de hasta 30 segundos, lo cual significa que contienen, como promedio, de 5 a 8 proporsiciones; es decir, dos oraciones o más". Sostiene que cada uno de estos ciclos coincide con una "idea" única. Para analizar esto, pidió a ciertas personas que dividieran un discurso espontáneo en "ideas" separadas y encontró que existe una buena correspondencia (aunque no absoluta) entre el comienzo de una etapa dubitativa y el comienzo de una "idea".

La planificación de la expresión de ideas parece requerir tiempo, pero una vez que se desarrolla un plan para expresar la idea, el discurso puede proceder con mayor fluidez.

*Figura 9.1. Ciclos temporales en el habla (Según Beattie, 1983).*




Las pausas aparecen especialmente entre las cláusulas y antes de las palabras de contenido (Butterworth, 1980; Beattie, 1983).

Una oración puede dividirse en proposiciones, cada una de las cuales debe contener un verbo. Las pausas que aparecen entre proposiciones probablemente tengan tres funciones. En primer lugar, le permiten respirar al hablante; en segundo lugar, pueden ayudar al oyente a dividir el discurso en unidades apropiadas; y, en tercer lugar, probablemente sean importantes en la planificación de la articulación detallada de la proposición. Como las proposiciones son las unidades mayores de una oración, necesitan planificarse como un todo, para que los acentos puedan aparecer en los lugares adecuados. Sin embargo, vale la pena observar que de ninguna manera se producen pausas antes de todas las proposiciones: menos de la mitad de las proposiciones en la fase dubitativa y alrededor de un cuarto en la fase fluida tienen pausas que las anteceden (Beattie, 1983).

Las vacilaciones también aparecen antes de las "palabras de contenido"; es decir, antes de los sustantivos, verbos, adjetivos y adverbios que llevan el sentido de la oración. Tienden a no aparecer antes de las "palabras de función" (preposiciones, conjunciones, artículos, etc.), que indican el marco

gramatical de la oración. Estas pausas a menudo se asocian con gestos que parecen representar el significado de la palabra cuando finalmente aparece (Beattie, 1983). En otras palabras, al hablar, incluso en una fase fluida, a veces tenemos una palabra en la punta de la lengua pero no podemos decirla, aunque manifestamos su significado por medio de un gesto. Fenómenos como éste han sugerido a mucha gente que, cuando planificamos el discurso, planificamos el sentido general de una idea, el marco gramatical de las oraciones y proposiciones y una especificación de las palabras involucradas. Al avanzar por el marco llegamos a un hueco que hay que llenar con una palabra que tiene una especificidad semántica peculiar. Buscamos en nuestra memoria una secuencia de sonidos determinada que se ajuste al significado, pero no podemos hallarla de inmediato, aunque es evidente que conocemos claramente su significado ya que podemos sugerirlo mediante un gesto. Tal vez parezca sorprendente que se pueda establecer el marco de una oración antes de elegir las palabras que la integran, pero en realidad eso es lo que indica la experiencia común; después de todo, es bastante fácil completar con palabras el marco de una rima jocosa:

*De la cima de aquel cerro*
*Venía bajando un .....*

Si observamos las vacilaciones, obtenemos evidencias sobre dos aspectos de la planificación del discurso. Primero, sobre la planificación de cómo expresar ideas, lo cual implica planificar unas pocas oraciones por vez; y segundo, el estudio de las vacilaciones empieza a proporcionarnos ciertas indicios sobre cómo se planifican las cláusulas en detalle. Parece que el hablante planifica una cláusula por adelantado, al menos en lo que respecta a la estructura principal, y conoce el significado de las palabras involucradas, pero no siempre sabe precisamente cómo dirá estas palabras.

Sin embargo, para analizar con más detalle algunos de los procesos involucrados en la producción de cláusulas, tenemos que observar los errores que la gente comete al hablar.

## Ejecución de corto alcance de los planes de discurso

Hay varias clases de errores de discurso, como se puede observar en la tabla 9.1. Las especulaciones sobre por qué pueden producirse estos erro-

res nos dicen mucho sobre el proceso involucrado en la producción de discurso. En primer lugar, como es esperable, el proceso de búsqueda de palabras provoca errores; a veces se necesita tiempo para encontrar las palabras, lo cual provoca vacilaciones, y a veces se encuentra una palabra equivocada, que puede tener sonido similar al de la palabra que se buscaba, por ejemplo:

*1* "No, soy *burdo" (en vez de* "zurdo").

También puede tener un *significado similar,* como:

*2* "Fui en bicicleta *a la escuela mañana" (en vez de* "ayer")

En algunas ocasiones, se extraen de la memoria dos palabras al mismo tiempo y se las combina inadvertidamente:

*3* "una madera *pulisa" (pulida + lisa)*

*4* "Es *torrible" (terrible + horrible)*

*Tabla 9.1. Algunos tipos de errores de discurso*

| | |
|---|---|
| 1 | *Sustitución* |
| (a) | De palabras con sonido similar |
| Ejemplo: | No simpatiza |
| | (en vez de "sintetiza") |
| (b) | De palabras con significado similar |
| Ejemplo: | Pregúntame si te parece posible. |
| | (en vez de "Dime...") |
| 2 | *Combinación* |
| Ejemplo: | ¿Alguna vez volejaste de noche? |
| | (en vez de "volaste" o "manejaste") |
| 3 | *Anticipación* |
| Ejemplo: | Una clección clara |
| | (en vez de "lección") |
| 4 | *Intercambio* |
| Ejemplos: | |
| | *Abre la casa de tu puerta* |
| | (en vez de "la puerta de tu casa") |
| | *un tejo rocho* |
| | (en vez de "un techo rojo") |

Estos errores nos dicen algo sobre los procesos involucrados en el ajuste de palabras en una estructura, y lo que nos dicen es coherente con lo que se desprende de los estudios sobre las vacilaciones: parece que el hablante llega a un lugar de la oración en el que tiene la intención de producir una palabra. Aparentemente, su plan contiene una especificación del significado de esa palabra, como puede observarse en el hecho de que a veces se dice por error una palabra con significado parecido al correcto, y también en el hecho de que en ocasiones se unen dos palabras alternativas.

Sin embargo, en algún punto el hablante tiene que extraer de la memoria una palabra que tenga los sonidos y el significado correctos, y parece que a veces este proceso produce deslices: es como si hubiera un proceso que cumple las especificaciones de sonidos adecuados para una palabra que tiene un significado en particular. Usualmente, esto lleva a la emisión de la palabra correcta, pero a veces otra palabra cumple con los requisitos sonoros bastante bien y se la dice en lugar de la correcta: de ahí los errores de sustitución de palabras por otras con sonido similar (como en el ejemplo 1 de la tabla).

Pero los errores de hallazgo de palabras no nos dicen cómo se planifica la proposición: para eso necesitamos analizar los errores de anticipación y, en especial, los de intercambio.

El estudio de estos errores nos hace comprender que: primero, si anticipamos los sonidos de una palabra cuando estamos diciendo otra y decimos, por ejemplo, "un tino tinto" en vez de "un vino tinto", las especificaciones para la "v" y la "t" deben estar presentes de manera simultánea en nuestras mentes. Del mismo modo, si intercambiamos dos palabras (por ejemplo: "¿Me prestas un cuento de libros?"), ambas palabras deben haber estado presentes a la vez. Por lo tanto, podemos ver una vez más que el habla no procede de a una palabra por vez sino que se planifica por adelantado.

Segundo, si observamos las palabras que se intercambian, podremos tener cierta idea sobre la duración del discurso que se planifica detalladamente. Según Garrett (1980), las palabras suelen intercambiarse dentro de una proposición y sólo un pequeño porcentaje de los errores de intercambio (20%) se produce en extensiones discursivas mayores que la proposición (recordemos que las oraciones pueden tener una o más proposiciones, cada una de las cuales contiene un verbo). Aparentemente, entonces, la proposición constituye una especie de unidad articulatoria. Es, por supuesto, mucho más corta que los ciclos de planificación a plazo intermedio que describió Butterworth, que contenían de 5 a 8 proposiciones. Sin embargo, parece pro-

bable que lo que se planifica a ese nivel sea una especificación bastante abstracta del significado de lo que se va a decir, y que no tiene suficiente nivel de detalle como para articularse realmente. El "programa articulatorio" en sí se ocupa de fragmentos de discurso más breves, y parece involucrar una especificación detallada de cómo decir una proposición en particular.

El tercer punto, y el más interesante, es que el análisis de los errores de intercambio, para varios especialistas (Fromkin, 1973; Garrett, 1980), sugiere que es posible deducir una secuencia de pasos de procesamiento en la formación de un programa articulatorio. Para analizar estas ideas, tenemos que examinar más de cerca los errores de intercambio.

Existen errores de intercambio de tres tipos diferentes:

a) Intercambios de *palabras*
"¿Me prestas un cuento de libros?"
"Abre la casa de tu puerta"

b) Intercambio de *sílabas*
"Justad y liberticia" (justicia y libertad)
"Listas y revibros" (libros y revistas)

c) Intercambio de *sonidos*
"Hay artenes y sollas" (sartenes y ollas)
"Fatos y dotos" (datos y fotos)

Hay tres aspectos de estos intercambios que proporcionan indicadores útiles para dilucidar los procesos que pueden estar involucrados en la producción de una cláusula: obsérvese primero que los ítem intercambiados tienden a ser muy similares entre sí (ya que las palabras se intercambian con otras palabras, las sílabas con otras sílabas y los sonidos con otros sonidos) y los intercambios tienden a involucrar ítem en posiciones similares (por ejemplo, las sílabas intercambiadas se encuentran en la misma posición en sus respectivas palabras).

En segundo lugar, obsérvese que muchos intercambios dejan intacta la estructura gramatical de la proposición:

"Esperar mejora." (Mejor esperar.)
"Tengo que chequear un cobre." (Tengo que cobrar un cheque.)

Observaciones como éstas han sido usadas por Fromkin y Garrett pa-

ra construir sugerencias sobre cómo una sucesión de procesos puede estar involucrada en la articulación de una proposición. Algunas de estas sugerencias pueden resumirse así:

a) *Selección de un significado.* La primera etapa es, presumiblemente, la de decidir la idea que se expresará en la cláusula: por ejemplo, que una persona o un objeto en especial están realizando una acción de alguna clase.

b) La segunda etapa parece ser la de *selección de un marco gramatical.* Como se dijo antes, parece bastante extraño elegir un marco gramatical antes de escoger las palabras que llenarán los huecos en ese marco, pero ésta parece ser la única justificación para algunas de las formas características en que funciona el intercambio: específicamente, que el marco casi siempre permanece intacto. Por ejemplo, cuando alguien quiere decir "Abre la puerta de tu casa" pero dice, en cambio, "Abre la casa de tu puerta", parece que el marco "Abre la _____ de tu _____" ya estaba establecido y luego se ubicaron las palabras en los lugares incorrectos.

c) La tercera etapa es la de *selección de las palabras de contenido* para llenar los huecos. Como vimos antes, esta etapa también puede funcionar mal ya que puede elegirse por equivocación una palabra de sonido similar a la buscada, o una palabra de significado similar, o pueden combinarse dos palabras correctas o, directamente, puede haber dificultades para encontrar la palabra correcta.

d) Finalmente, hay una etapa en la que se especifican los *detalles de pronunciación.* La prueba de que esta etapa es posterior a la que mencionamos antes es que cuando se produce un intercambio, la pronunciación de los plurales, tiempos verbales, etc., es adecuada para las palabras que *realmente* llenan los huecos y no para las que debían llenarlos.

## Conclusión

Para un ser humano (y sólo para un ser humano) los años de práctica dan como resultado un individuo experto que puede hablar sin esfuerzo, planificar por adelantado y producir oraciones que nunca ha dicho u oído antes. Se han empleado tres técnicas fundamentales para estudiar esta habilidad extraordinaria: el análisis de las elecciones que hacen los hablantes, el análisis de los puntos en los que los hablantes hacen pausas para el pensamiento, y el análisis de las (sorprendentemente raras) ocasiones en que el habla no funciona bien. Se han usado estas técnicas para empezar a bosquejar detalladamente los procesos involucrados en el habla. Sin embargo, escuchar también es una habilidad y de ella nos ocuparemos en el próximo capítulo. Detallaremos algunas de las formas en que las elecciones de los hablantes afectan a los oyentes y cómo vale la pena considerar todo esto al pensar tanto en el hablar como en el escuchar.

### Lecturas adicionales

BEATTIE, G. (1983) *Talk: An analysis of speech and non-verbal behaviour in conversation.* Milton Keynes: Open University Press.

Interesante análisis de instancias reales de conversación.

AITCHISON, J. (1983) *The Articulate Mammal.* Londres: Hutchinson.

Introducción útil, inteligente y de fácil lectura al campo de la psicolingüística.

# 10

# Las habilidades del oyente

*Don Rogers*

## La comprensión del lenguaje

Cuando pensamos en los procesos de comprensión, una teoría muy obvia sobre cómo funcionan nos viene de inmediato a la mente: la comprensión funciona "de abajo hacia arriba". Es decir, oímos y reconocemos sonidos individualmente (o, cuando leemos, reconocemos las letras individualmente); luego, reconocemos que ciertas combinaciones peculiares de esos sonidos o letras son palabras; nuestro conocimiento de la gramática nos permite comprender que ciertas combinaciones de esas palabras forman oraciones; y, finalmente, entendemos que la combinación de oraciones forma una narración, una conversación, etc.

Sin embargo, si bien esta idea es bastante acertada, también parecen existir opiniones según las cuales los procesos de comprensión funcionan "de arriba hacia abajo", ya que comprendemos lo que significan las letras, las palabras y las oraciones según el contexto y de acuerdo con nuestros conocimientos.

Analizaremos este fenómeno en relación con varios niveles lingüísticos, empezando por el más simple, es decir, los sonidos o las letras. Nos ocuparemos de la lectura más que de los sonidos, simplemente porque sus efectos son más fáciles de ilustrar en el papel y la comprensión de los sonidos presenta exactamente las mismas dificultades. Por supuesto, esto relaciona a este capítulo con el análisis de la lectura que se hace en el capítulo 4 de este mismo libro.

### El reconocimiento de letras

Si bien las letras son bastante fáciles de reconocer, todos sabemos que a veces (en especial cuando intentamos leer textos manuscritos) puede haber dificultades. Sin embargo, es sorprendente ver con cuánta frecuencia pode-

mos leer sin dificultad letras que están muy mal conformadas. Obsérvense los ejemplos de la figura 10.1. Hay muchos casos en los que una letra determinada es muy difícil de reconocer, a menos que uno pueda ver el resto de la palabra en que esa letra se inscribe. También se verán casos en los que se pueden reconocer formas bastante diferentes como ejemplos de la misma letra. En otros casos, la misma forma se ve como letras distintas en diferentes contextos, como lo ilustra la figura 10.2.

En general, entonces, hay ocasiones en las que la decisión sobre cuáles son las letras presentes en una palabra depende de la lectura global de esa palabra. Es decir, el contexto que la precede y la sigue se tiene en cuenta para reconocer las letras. Lo mismo es válido para el reconocimiento de los sonidos que conforman el discurso. Por ejemplo, la misma letra tendrá formas diferentes según la pronunciación y el acento del hablante. Donde esa pronunciación no resulte familiar, la identificación de los sonidos individuales (o, incluso, de palabras enteras) dependerá del contexto.

---

*Figura 10.1. Las letras pueden tener variadas formas físicas. Lo mismo sucede con los sonidos.*



## La comprensión de palabras y oraciones

Así como las letras son fáciles de reconocer en contexto, éste también facilita mucho la comprensión y el reconocimiento de las palabras. Se pueden hacer varios comentarios al respecto. En primer lugar, como oyentes experimentados del lenguaje, tenemos la impresión de que el discurso se segmenta en palabras, con pausas entre ellas, igual que en la página impresa. Pero esto no es así: como vimos en el capítulo anterior, el patrón de pausas en el discurso es mucho más complejo. Si uno trata de decir una o dos oraciones haciendo pausas entre cada palabra, comprobará que suena muy extraña. La impresión de que existen espacios surge porque comprendemos la oración e imaginamos esos espacios entre las palabras. Es divertido construir oraciones que se puedan dividir en palabras de diferentes maneras, de modo que se puedan cambiar las posiciones de los aparentes espacios entre ellas. Esto sucede en algunas adivinanzas infantiles. Veamos un ejemplo conocido:

A. Tengo 12 huevos metidos en una cesta. Si pongo 3 más, ¿cuántos habrá en ella?

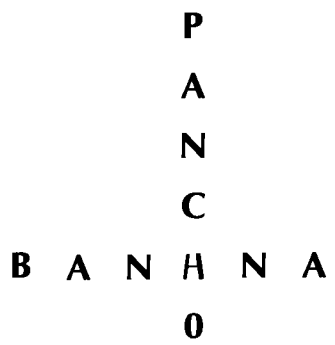
B. Quince.

A. No, cinco.

(En vez de "metidos", la respuesta de A sugiere que lo que había dicho en su primera intervención había sido "metí dos".)

---

*Figura 10.2. El mismo estímulo puede merecer diferentes interpretaciones, según el contexto.*



Una segunda función del contexto al nivel de la comprensión de palabras tiene que ver con el hecho de que muchas palabras tienen más de un significado, y en muchos casos hay palabras diferentes que suenan parecido. Para decidir cuál es el significado correcto en cada caso, y cuál es la palabra pertinente, hay que tener en cuenta el significado y la estructura gramatical de toda la oración y, frecuentemente, también de las oraciones anteriores y posteriores. En general, es muy fácil pensar ejemplos de palabras que puedan tener un significado en un contexto y otro muy distinto en otro contexto, sencillamente porque muchas palabras tienen varios significados diferentes. Veamos un caso:

* En el norte argentino hay muchas *llamas*... debido a la abundancia de incendios forestales.

* Si no *llamas* pronto por teléfono a tu hermano, se enojará mucho.

En tercer lugar, cuando escuchamos o leemos, tenemos que decidir las partes del discurso y la función gramatical de cada palabra. Éste no es un procedimiento directo o mecánico, como se puede ver examinando algunas oraciones. Por ejemplo:

*La bella llama de las montañas*

es superficialmente similar a:

*La bella llama desde las montañas.*

Pero, evidentemente, en el primer caso, *llama* es un sustantivo y *bella* un adjetivo, en tanto que en el segundo caso *llama* actúa como verbo y bella se sustantiviza.

Estos tres conceptos (el discurso no se divide en palabras separadas por pausas, las palabras pueden tener varios significados diferentes, una palabra dada puede cumplir distintas funciones gramaticales en diferentes oraciones) sugieren que tenemos que usar considerablemente el contexto para determinar la parte del discurso en que se halla incluida una palabra, el significado peculiar que tiene y, en realidad, cuáles son las palabras que están involucradas.

Esto ha llevado a algunos investigadores a pensar que el oyente espera hasta el final de la oración antes de llegar a una decisión sobre su significado. Sin embargo, parece que los procesos "de abajo hacia arriba" y "de arriba hacia abajo" operan simultáneamente. Esto ocurre en parte porque hay que tomar decisiones muy rápidamente. En el discurso normal, escuchamos

unas 2 o 3 palabras por segundo. Marslen-Wilson y Tyler (1981) demostraron que en los experimentos la gente puede repetir un discurso que se está diciendo con una demora de sólo un cuarto de segundo: se pidió a los sujetos que escucharan hablar a alguien y repitieran el mensaje mientras lo oían. Los experimentadores descubrieron que la pronunciación de algunas palabras requería hasta 375 milisegundos pero los sujetos podían empezar a repetirlas sólo 250 milisegundos después del principio. Es decir, comenzaban a repetir la palabra antes de haberla oído entera. Además, había un buen nivel de comprensión del material que se repetía; no era una mera repetición como un eco. La conclusión inevitable es que en esta situación la gente podía decidir correctamente qué palabra estaba oyendo cuando sólo había oído la mitad de ella. Esto demuestra el poderoso efecto del contexto, pero también que no es necesario esperar hasta el final de la oración para comprender las palabras que están incluidas en ella. Por este motivo es bastante fácil engañar al sistema, y en muchos estudios se han usado oraciones que (como dice Aitchison, 1983) inducen a los oyentes al error. Por ejemplo:

¡No seas tonto!... le dije a mi hermano.

Quien escucha esto piensa, al principio, que el "tonto" al que se hace referencia es él mismo; sólo al escuchar la última parte comprende que no es así y que el "tonto" en cuestión es el hermano del hablante.

## La comprensión del significado social de las oraciones

Pero aun cuando entiende las palabras y la estructura gramatical de una oración, al oyente le quedan otros problemas. En particular, necesita apreciar las intenciones del hablante para poder responder adecuadamente.

Cuando escuchamos a alguien, no sólo está emitiendo una sucesión de sonidos organizados de forma que podamos entenderlos como oración sino que también está desempeñando una acción social: quizá nos está acusando, advirtiendo o pidiendo algo; a estos se los suele llamar actos de habla. La cuestión es: ¿cómo sabemos qué quiere? (Ver, también, capítulo 13.)

Una forma de saberlo podría ser que nos dijera lo que quiere:

*"Te pido algo: sal de mi camino."*

Pero, como usuarios del lenguaje, sabemos perfectamente bien que la gente no suele hablar así y que si lo hace es más un desafío que un pedido normal.

En segundo lugar, puede haber estructuras gramaticales específicas

asociadas con actos de habla específicos. Aquí se presentan dos problemas. Primero, como ha señalado Austin (1962), se pueden desempeñar cientos de diferentes actos de habla, y el número de formas gramaticales especiales tendría que ser muy grande. Además, existe de hecho una forma especial para hacer que alguien haga algo. Es el imperativo:

"¡Apártate de mi camino!"
"¡Abre la puerta!"
"¡Préstenme atención!"

Es obvio que el imperativo aparece en contextos en los que hay una diferencia de condición social amplia (y con frecuencia de naturaleza institucional) entre el hablante y el oyente (como, por ejemplo, en el ejército) o resulta extraordinariamente grosero (precisamente porque sugiere la existencia de esa diferencia social).

Por lo tanto, los pedidos suelen hacerse de manera amable, como "pedidos indirectos", y forman parte de los *actos de habla indirectos*. Por ejemplo, podríamos decir:

"¿Podrías lavar los platos?"
"¿No te toca a ti lavar los platos?"
"Hay que lavar los platos."
Etcétera

Es evidente que la forma en que el oyente distingue un pedido indirecto de una pregunta o una afirmación comunes presenta otro problema. El filósofo John Searle tiene un análisis que resulta de mucha utilidad al respecto (Searle, 1975). Señala que para que cualquier expresión pueda ser un pedido, debe cumplir ciertas condiciones. En los ejemplos antes mencionados:

a) debe haber necesidad de la acción (por ejemplo, debe haber necesidad de que se laven los platos).

b) el oyente debe ser capaz de desempeñar la acción y estar dispuesto a hacerlo.

c) el hablante debe ser sincero y querer que el oyente realice la acción.

d) la expresión debe ser un intento de lograr que la acción se realice.

e) el hablante debe tener derecho a pedirle algo al oyente.

Sólo si se cumplen todas estas condiciones puede considerarse que una expresión es un pedido y, con frecuencia, para pedir algo se señala alguna de ellas. Por ejemplo: "¿Podrías lavar los platos?" pregunta al oyente si es-

taría dispuesto a desempeñar la acción. Si se cumplieran todas las demás condiciones (es decir, que hiciera falta que se lavaran los platos, etc.), esto se podría considerar un pedido. Por otra parte, es más factible que algo como "¿Te gustaría ir a la luna?" sea considerado como una pregunta genuina, que espera una respuesta afirmativa o negativa pero ninguna acción posterior (Stubbs, 1983, presenta un análisis más profundo sobre el tema).

Pero para responder a algunos pedidos no basta con evaluar si se nos está pidiendo algo. Con frecuencia, tenemos que ponernos en el lugar del hablante, considerar sus propósitos, sus conocimientos y las características salientes de la situación en general. Veamos un ejemplo simple, como que a alguien se le pregunte dónde queda un lugar:

"¿Sabes dónde queda Harrod's?"

La respuesta que uno dependerá de:

a) Dónde está uno en ese momento. Si uno está en Escocia, la respuesta adecuada podría ser: "En Londres". Si uno está en Londres, podría ser: "En Knightsbridge". Y si uno está en Knightsbridge: "Derecho por esta calle".

b) Su evaluación de por qué el hablante necesita la información. Por ejemplo, si el hablante estuviera mirando un mapa de Londres, la respuesta "Knightsbridge" sería más adecuada que "Londres", incluso si uno estuviera en Escocia en ese momento. Si uno sabe que el hablante es un especulador bursátil y que ha visto que uno está leyendo un informe financiero, la respuesta podría ser: "Tres puntos y medio arriba".

c) Su evaluación de los conocimientos del hablante. Con frecuencia, esto es algo establecido en la conversación antes de elegir la respuesta adecuada:

A: "¿Sabes dónde queda Harrod's?"

B: "¿Conoces Hyde Park?"

A: "No."

B: "Ah, bueno. Si sigues por aquí verás un gran arco. Al otro lado está el parque. Si lo cruzas por el sendero y al llegar al otro lado vuelves a preguntar, llegarás a Harrod's. Pensándolo mejor, ¿por qué no tomas un taxi?"

## Más allá de la oración

Las oraciones no aparecen de manera aislada: son parte de textos más amplios (relatos, conversaciones o lo que sea) y también se las puede usar en contextos físicos particulares. Naturalmente, esto ayuda considerablemente

en la interpretación de oraciones: una oración ambigua como "Están comiendo manzanas" tiene menos lecturas alternativas si uno tiene a los comensales delante o si las oraciones anteriores han tenido que ver con manzanas.

Sin embargo, la interpretación de oraciones en contexto también presenta problemas y nos proporciona más evidencias de cuán poderosos, hábiles y activos son los procesos de comprensión.

En un estudio ya clásico, Gordon Bower (1976) les leyó a sus alumnos historias como ésta:

(Nancy) fue a ver al médico. Llegó al consultorio y se presentó ante la secretaria. Fue a ver a la enfermera, que realizó los procedimientos de rutina. Luego, (Nancy) subió a la balanza y la enfermera anotó su peso. El médico entró al consultorio y examinó los resultados. Le sonrió a (Nancy) y dijo: "Bueno, parece que mis expectativas se confirman". Cuando el examen terminó, (Nancy) salió del consultorio.

Estas historias estaban precedidas por una breve descripción, que podía ser, por ejemplo:

> Nancy se despertó sintiéndose mal nuevamente y se preguntó si realmente estaría embarazada. ¿Cómo se lo diría al profesor con el que había estado saliendo? Además, estaba el problema del dinero.

Por otra parte, podía tratarse de una breve descripción de Jack, un joven que está preocupado porque no pesa lo suficiente como para integrar el equipo de luchadores. (Naturalmente, cuando ésta era la descripción que precedía a la historia, los nombres y pronombres cambiaban por "Jack" y "él".)

Veinticuatro horas después, Bower puso a prueba la memoria de los estudiantes. Se descubrió que había muchas distorsiones en sus recuerdos de la historia y que estos variaban según la descripción que la precedía. Por ejemplo, en el relato de "Nancy embarazada" los alumnos recordaron mal la afirmación del médico: "parece que mis expectativas se confirman" y dijeron "Sus temores parecen confirmarse" o simplemente "Ud. está embarazada". En cambio, en la historia de "Jack el luchador", recordaron que el médico le había dicho a Jack que estaba llegando al peso necesario.

A partir de este experimento, pueden hacerse al menos dos comentarios. En primer lugar, demuestra que comprender un relato implica un proceso de inferencia activa (ver capítulos 4 y 13). Se trata, por supuesto, de un fe-

nómeno muy generalizado en la lectura o la audición de historias, y los escritores de novelas policiales como Agatha Christie se valen mucho de esta propensión de sus lectores. Las pistas falsas que el novelista siembra nos confunden sólo porque hacemos inferencias activamente.

En segundo lugar, tanto en el relato de Bower como en las novelas policiales, parece que hacemos inferencias como consecuencia de nuestro intento de ensamblar los fragmentos de información contenidos en la historia. Parece que damos por supuesto que si se ha mencionado algo, debe ser relevante y se lo debe poder ensamblar con las demás cosas que se mencionan en la historia.

Con respecto a esto, el filósofo H.P. Grice (1975) señala que hablante y oyente actúan como si observaran un conjunto de principios co-operativos. Grice encapsula algunos de estos principios en cuatro "máximas":

Cantidad: Sea tan informativo como se necesite: no proporcione información de más ni de menos.
Calidad: Sea sincero.
Relación: Sea relevante.
Modo: Sea claro y ordenado.

Estas máximas no pretenden ser una receta para quienes aspiran a tener una conversación sino más bien un conjunto de principios intuitivos que cualquier conversador hábil y experimentado obedece de manera implícita. Es debido a esto que el lector o el oyente pueden hacer deducciones; precisamente porque damos por sentado que se obedecen las máximas de cantidad (de modo que no se proporciona más información que la necesaria) y de relación (y la información que tenemos es relevante) podemos sacar la conclusión de que la afirmación del médico tiene que ver con el embarazo de Nancy.

La violación de las máximas de Grice produce efectos interesantes. Por ejemplo, como señala él mismo, si después de haber sido tratado abominablemente por otra persona, un hombre dice: "Es un buen amigo", habrá violado la máxima de calidad, diciendo algo que no es cierto. Si su oyente lo sabe, podrá deducir que esa afirmación pretende ser irónica.

Estos temas, en especial la cuestión de la relevancia, son analizados más profundamente por Downes (1984).

Puede hacerse otro comentario sobre la historia de Bower. Damos por sentado que el malestar de Nancy se relaciona con su embarazo. Sin embar-

go, esta deducción tiene que ver con lo que podría denominarse conocimiento "no lingüístico" o "mundano": "sentirse mal" no significa estar embarazada. Entender el lenguaje, por lo tanto, implica un conocimiento no sólo del lenguaje mismo sino también del mundo. Esto puede verse claramente en un experimento bastante extraño realizado por Dooling y Lachman (1971). Este experimento incluye una historia muy difícil de entender hasta que se le da un título:

> Con joyas empeñadas para financiarlo, nuestro héroe desafió valientemente todas las burlas que habían intentado impedir sus planes. Tus ojos mienten, había dicho, un huevo y no una mesa tipifica correctamente este inexplorado planeta. Ahora, tres tenaces hermanas buscaban pruebas, abriéndose camino por los a veces serenos pero con más frecuencia turbulentos picos y valles. Los días se convirtieron en semanas y muchos incrédulos difundieron terribles rumores sobre el abismo. Finalmente, de la nada, bienvenidas criaturas aladas aparecieron, lo cual implicaba un éxito trascendental.

Mucha gente que lee o escucha esta historia puede hallarla incomprensible: entienden las palabras pero no encuentran un marco que haga que todos los datos mencionados resulten relevantes. Pero cuando se les dice que es una historia llamada "Cristóbal Colón descubre América", la relevancia de todos los puntos se hace evidente y pueden interpretarse las ambiciosas metáforas. Las palabras no han cambiado: no hay cambios en el conocimiento lingüístico pero ahora puede articularse el marco de referencia adecuado y el relato se entiende sobre la base de nuestros conocimientos históricos, del mar, los buques a vela, etc.

## Mejorar las habilidades del hablante y del oyente

Evidentemente, el oyente enfrenta un gran problema si encuentra dificultades para comprender y recordar los contenidos, por ejemplo, de una conferencia. A diferencia del lector, no puede descansar cuando su concentración flaquea ni puede volver atrás si no alcanza a seguir un argumento: la velocidad a la que tiene que incorporar la información la establece el hablante. Pero, como todos sabemos, es muy difícil mantener la concentración durante toda una conferencia de 50 minutos: del mismo modo que un operador de radar no puede mantener el mismo nivel de vigilancia constantemente, un

alumno no puede prestar la misma atención a cada gloriosa palabra que sale de los labios de su docente. De hecho, según Wilkinson, Stratton y Dudley (1974), "se ha descubierto que los alumnos que oyen conferencias sólo comprenden la mitad (o menos) de la materia básica". Esto es especialmente desafortunado ya que, como señalan más adelante, los estudiantes pasan gran parte del tiempo sólo escuchando; incluso en la escuela primaria se descubrió que los niños sólo escuchaban el 58% de lo dicho en cada día de clase.

Si tenemos en cuenta esto, es evidente que gran parte de la responsabilidad de aliviar la carga del oyente recae sobre el hablante, y hay muchos cursos y libros para ayudar a los oradores públicos a comunicarse con eficacia; por ejemplo, *Lecturing and Explaining*, de George Brown (1978). Naturalmente, hay muchas soluciones diferentes para el problema de la exposición oral eficiente, pero todas tienen que tener en cuenta al oyente. Por ejemplo, muchos de los recursos que emplea el hablante apuntan a superar la incapacidad del oyente para prestar atención todo el tiempo. Así, el docente puede señalar por adelantado los puntos principales que piensa tratar, puede resumir los aspectos principales, puede llamar la atención sobre ellos escribiéndolos en un pizarrón o proporcionar información adicional haciendo referencia al mismo punto mediante su discurso y con ayudas visuales.

Además, debería ser obvio, por lo que ya se ha dicho en este capítulo, que el hablante también tiene que tener en cuenta los conocimientos del oyente, así como sus necesidades y propósitos. Suele ser difícil explicar un tópico nuevo a un grupo de manera tal que todos puedan entenderlo, y quien habla en público tiene que aprender a evaluar correctamente qué sabe su auditorio: si sobreestima sus conocimientos, el auditorio no entenderá, no recordará y ni siquiera se interesará en lo que dice; por otra parte, si subestima los conocimientos de su auditorio, puede aburrirlo.

Si bien el peso principal de la comunicación efectiva recae sobre los hablantes, también parece haber habilidades de comprensión que los oyentes pueden ejercitar con mayor o menor pericia. Pueden discernirse tres tipos de habilidad de comprensión. El primer tipo tiene que ver con la atención y con tener la voluntad de oír al hablante. Se incluyen aquí cosas tales como la decisión del nivel en el que se quiere entender y recordar el material del hablante. Por ejemplo, un ladrón y un dueño de casa podrían extraer diferente información de una charla sobre prevención del crimen. El segundo tipo de habilidad tiene que ver con el desempeño de ciertas operaciones cognitivas, como distinguir entre ideas principales y subordinadas, mantener en la mente detalles relacionados, sacar las conclusiones adecuadas o percibir la es-

tructura lógica de un argumento. En tercer lugar, hay otro conjunto de habilidades involucrado en la identificación de las "pistas" con los que los hablantes señalan la dirección de sus pensamientos. Estos indicios pueden ser verbales ("en primer lugar", "sin embargo", "en síntesis"), pueden incluir énfasis o tonos de voz o pueden ser totalmente no verbales (como los gestos, las posturas y los movimientos corporales en general). Todos ellos pueden proporcionar información sobre la forma en que el hablante considera la organización de lo que está diciendo (ver también el capítulo 11).

En su interesante intento de mejorar las habilidades del oyente, Brown (1968) enseñó a sus alumnos a buscar las pistas verbales (ver tabla 10.1) y no verbales (ver Tabla 10.2) de los conferencistas. Después de practicar con grabaciones de charlas en audio y video, descubrió que la calidad de las notas tomadas por sus alumnos había mejorado. Aparentemente, también los alumnos sentían que se habían beneficiado con el entrenamiento al que se los había sometido.

En general, entonces, un hablante será más eficiente en su discurso si tiene en cuenta las limitaciones y posibilidades cognitivas de sus oyentes. Necesita, particularmente, tomar precauciones contra las consecuencias de las caídas en la atención de parte de sus oyentes; para ello, debe estructurar claramente su discurso, aclarando cuál es la estructura que emplea, señalando sus puntos principales, repitiendo y resumiendo, y proporcionando información adicional por otros medios. Además, tiene que recurrir a las estructuras de conocimiento que el oyente ya posee de manera que lo que diga resulte comprensible para el oyente y pueda aumentar sus conocimientos. Por otra parte, el oyente puede ayudarse a sí mismo a entender, escuchando con mayor atención y buscando las pistas verbales y no verbales que sugieren lo que el hablante pretende decir. La tarea del oyente es recíproca a la del hablante. Tiene que prestar atención a (todos) los mismos puntos que el hablante e intentar deducir su sentido y sus intenciones.

*Tabla 10.1 Tipos de discurso de los conferencistas.*
*(tomado de Brown, 1978, bajo su autorización)*

*Preámbulos.* Murmullos rituales tales como: "Eh... este... buenas tardes. Hoy... estee... vamos a... eeh... analizar... esteee... un punto... eh... muy importante".

*Orientación.* El marco de la charla. No hace falta decir que algunos conferencistas no lo hacen explícito. En el mejor de los casos, establecen de qué van a hablar y escriben los títulos. Las orientaciones pueden darse al comienzo de una charla o al comenzar cada tópico o subtópico (ver a continuación).

*Marcos.* El encuadre de los subtópicos de una charla. Suele empezar con expresiones como: "Ahora bien...", "hasta el momento...", "entonces...". A veces, el orador sencillamente hace una pausa y echa una mirada a sus notas. Las "palabras de encuadre" señalan un cambio de tópico. Lo que les sigue suele ser importante (aunque no siempre sea necesariamente así).

*Claves.* Las afirmaciones clave. Son principios, reglas, hechos, suposiciones o preguntas fundamentales. Cuidado con ellas. Las preguntas que uno ha anotado probablemente alertarán sobre ellas.

*Ejemplos.* Pueden ser ejemplos de los principios o hechos; o pueden ser metáforas, ilustraciones y analogías. Un breve comentario sobre ellos ayudará a comprender el principio.

*Resúmenes.* Pueden darse al final de los tópicos o subtópicos o al terminar la charla.

*Tabla 10.2 Tipos de conductas de los expositores (tomado de Brown, 1978, bajo su autorización)*

| | |
|---|---|
| El expositor escribe en el pizarrón | (¿siguiente sección o sólo inquietud del expositor? |
| El expositor da un paso atrás y observa la pizarra | (está verificando) |
| El expositor se quita los anteojos | (probablemente viene algo importante) |
| El expositor se quita los anteojos y hace una pausa | (probablemente sea una salvedad o una reserva) |
| El expositor mira por la ventana | (piensa en lo que va a decir; quizá está pasando una chica bonita) |
| El expositor levanta un dedo | (viene una salvedad) |
| El expositor cierra la mano y hace un gesto | (probablemente ahora viene algo importante) |
| El expositor hace la señal de "alto" | (una salvedad) |
| El expositor mira sus notas | (viene la siguiente sección o está perdido) |
| El expositor mira fijamente sus notas | (está perdido) |
| El expositor revisa sus transparencias | (está perdido) |
| El expositor levanta la vista de sus notas y deja de leer | (está por explicar un punto; usualmente, es importante) |
| El expositor se sienta sobre el escritorio y enciende un cigarrillo. | (se avecina la técnica del "¿no soy un buen tipo?" Puede reemplazar a la charla, pero no siempre) |

## Conclusión

En este capítulo, hemos demostrado que la comprensión no es de ningún modo un proceso pasivo sino que parece ser, a todo nivel, activo y experto. Si utilizamos el contexto y las expectativas podemos comprender formas de letras muy distorsionadas o ambiguas y emplear poderosas rutinas de reconocimiento de palabras; como consecuencia del conocimiento acumulado sobre el mundo social, podemos llegar más allá de las formas de las expresiones, ponernos en el lugar del otro y adivinar sus intenciones; y como consecuencia de nuestro conocimiento general del mundo, podemos encarar la tarea de comprender formas adecuadas de referencia que nos permitan sacar conclusiones que en su mayoría se cruzan correctamente con las que pretenden los hablantes.

### Lecturas adicionales

DOWNES, W. (1984) *Language and Society*. Londres: Fontana
Análisis claro y sensato sobre los aspectos sociales del lenguaje.

AITCHISON, J. (1983) *The Articulate Mammal*. Londres: Hutchinson.
Introducción útil, inteligente y de fácil lectura al campo de la psicolingüística.

# 11

# Las habilidades sociales

*Helen Graham*

En *Rosencrantz y Guildenstern han muerto*, el dramaturgo Tom Stoppard (1967) toma sus personajes de Hamlet, la tragedia de Shakespeare. A los dos héroes condenados se les encomienda la poco envidiable misión de descubrir qué aflige a Hamlet, cuyos síntomas incluyen:

> Réplicas abundantes, alusiones místicas, identidades equivocadas, afirmar que su padre es su madre... cosas así; ideas de suicidio, privaciones, pérdida de la alegría, sensación de claustrofobia (por no decir delirio de encierro), invocaciones a camellos, camaleones, capones, ballenas, comadrejas, halcones, adivinanzas, juegos de palabras y evasivas; amnesia, paranoia, miopía; sueños diurnos, alucinaciones; acuchillamiento de sus mayores, abuso de sus padres, insultos a su amada y apariciones sin la cabeza cubierta en público.

Enfrentados a conductas tan extrañas, Rosencrantz y Guildenstern se sienten terriblemente desubicados. Guildenstern, perplejo, admite: "No sabemos qué está sucediendo ni qué hacer con nosotros mismos. No sabemos cómo actuar". Sobre esta base, un psicólogo podría llegar a la razonable conclusión de que Rosencrantz y Guildenstern son socialmente incompetentes, que sus habilidades sociales son inútiles.

## La competencia social como habilidad

Podría decirse que la competencia social consiste en "cómo hacer amigos e influir sobre la gente". Es la capacidad de lograr objetivos personales y profesionales, ya sea formando y manteniendo relaciones con otros o en actividades tales como la docencia, la venta, la entrevista o la psicoterapia. Argyle (1983) sugiere que las interacciones que se realizan en estas conductas sociales utilizan una jerarquía de habilidades cuyos elementos inferiores son automáticos y habituales, mientras que las secuencias superiores se encuentran

bajo control cognitivo. Estas últimas se planifican y están sujetas a control y acción correctiva continuos, a la luz de la retroalimentación.

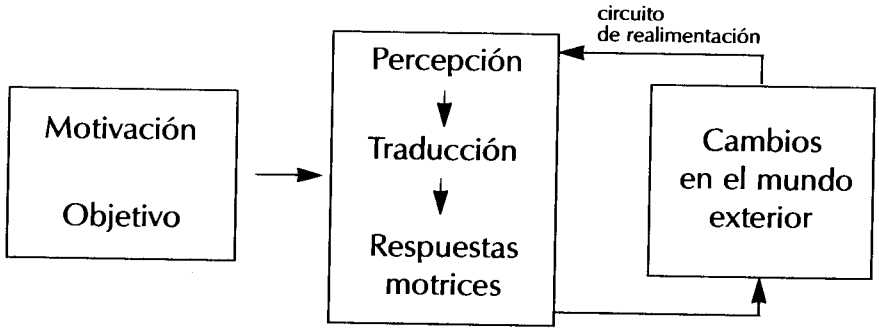
Argyle (1969) define la habilidad como una actividad organizada y coordinada en relación con un objeto o situación, que involucra toda una cadena de mecanismos perceptuales (sensoriales), cognitivos (centrales) y motrices.

> Una de sus características principales es que el desempeño, o la corriente de acción, se encuentra continuamente bajo el control de la "entrada" sensorial. Esta entrada deriva en parte del objeto o situación hacia el que puede decirse que apunta el desempeño y controla ese desempeño, ya que los resultados de las acciones se ajustan continuamente de acuerdo con cierto criterio de logro o grado de enfoque hacia un objetivo, y el desempeño se adapta consecuentemente. (pág. 180)

Sugiere que durante la interacción social un individuo se involucra en una actividad experta que incluye el ajuste y la modificación de la salida sobre la base de la información disponible. La conducta general es dirigida, adaptada y nada automática, aunque se la puede considerar como formada por elementos que en sí están automatizados.

Los elementos del modelo de habilidad social se representan como diagrama en la figura 11.1.

*Figura 11.1 El modelo de la habilidad social (según Argyle, 1969)*



El modelo de la habilidad social sigue estrechamente al modelo de las habilidades perceptivo-motrices (capítulos 2 y 3). Tiene cinco etapas:

a) *Los objetivos del desempeño experto.* Pueden incluir la obtención o transmisión de información, el cambio de actitudes o de estado emocional, la supervisión o colaboración. Pueden estar motivados profesionalmente (como sucede en la enseñanza, las entrevistas, la venta o la enfermería) o de un modo puramente social, por necesidades tales como la asociación o la autoridad.

b) *Percepción selectiva de las señales.* La percepción social es altamente selectiva. Quienes participan en las interacciones se tornan altamente sensibles a la información visual, auditiva y, posiblemente, táctil, y aprenden a cuál deben prestarle atención.

c) *Procesos de traducción centrales.* La información obtenida mediante los receptores sensoriales lleva a la formación de planes de acción adecuados. La interacción social depende de las reservas aprendidas de semejantes traducciones, que se construyen durante el proceso de socialización, cuando la persona aprende cuáles son las técnicas sociales que generan en los demás las respuestas deseadas.

d) *Respuestas motrices.* Se pone en acción el plan, pero hace falta práctica para lograr un desempeño fluido y preciso de los patrones de la actividad motriz. Cuando esas habilidades se van perfeccionando, la conciencia alerta se reduce de manera tal que las habilidades sociales complejas (como establecer afinidad o reducir la ansiedad de otra persona) pueden convertirse para algunos en automáticas.

e) *Retroalimentación y acción correctiva.* Por medio del empleo de señales perceptuales, el ejecutante realiza acciones correctivas cuando se hace necesario.

Las habilidades sociales resultan análogas a habilidades motrices como escribir a máquina o conducir un vehículo; son habilidades que, en parte, tienen que aprenderse. Por lo tanto, es esperable que algunas personas las adquieran y desarrollen mejor que otras.

La observación y la experiencia diarias sugieren que esto es efectivamente así. Mucha gente es tímida y le resulta difícil ganar amigos o mantener relaciones. Pueden ser malos conversadores, incapaces de hacerse entender o quizá tengan dificultades para encarar situaciones cotidianas, como relacionarse con miembros del sexo opuesto o con la autoridad, ir a fiestas o decir "no" a los pedidos de los demás. Está claro que la gente difiere notablemen-

te en su habilidad para manejar ese tipo de situaciones, pero las posibles consecuencias de esas diferencias no son tan obvias. Argyle (1984) informa sobre varios estudios que revelan que el ausentismo, la productividad, las quejas y las ventas de un negocio en particular difieren significativamente según los supervisores. Las habilidades sociales incorrectas también pueden tener graves consecuencias y llevar al fracaso en el establecimiento y mantenimiento de relaciones personales, a la imposibilidad de conseguir socios o empleos y, en los casos extremos, al aislamiento social, la enfermedad mental y, quizá, hasta al suicidio.

De hecho, uno de los factores fundamentales para la trágica y temprana muerte de Rosencrantz y Guildenstern en la obra de Stoppard es su incapacidad de apreciar los diferentes matices existentes en las situaciones sociales. Como comprende Guildenstern antes de su ejecución: "Debe haber habido algún momento, al principio, en el que podríamos haber dicho que no. Pero, por algún motivo, no nos dimos cuenta." (pág. 95)

## Las habilidades sociales

Las habilidades sociales pueden considerarse bajo varios encabezamientos: habilidades de comunicación, habilidades interpersonales, habilidades de desempeño. Todas ellas tienden a interrelacionarse, y por eso cualquier categorización es, necesariamente, algo arbitraria. Sin embargo, como todas incluyen en alguna medida a la conversación, se puede considerar con justicia que las habilidades conversacionales son de importancia primordial.

### Las habilidades conversacionales

La conversación es esencialmente comunicación hablada, que incluye el intercambio verbal entre personas y por lo tanto es de doble sentido, ya que el "habla" o la "conversación" por sí mismas no tendrían significación.

> La conversación es sin duda la piedra basal del mundo social: los seres humanos aprenden a hablar en la conversación, encuentran compañía con la conversación, se socializan a través de la conversación, mejoran su jerarquía social como consecuencia de la conversa-

ción y, se supone, hasta pueden enfermar mentalmente debido a la conversación.

(Beattie, 1983, p. 2)

Sin embargo, como señala Beattie, mientras que la conversación es primordialmente un hecho social, los procesos cognitivos complejos están involucrados en su planificación y generación. (Ver también el capítulo 9.)

En *Rosencrantz y Guildenstern han muerto* podemos ver una amplia planificación de los contenidos de la conversación y su ensayo. Gran parte de la obra de Stoppard trata sobre qué le dirán Rosencrantz y Guildenstern a Hamlet si pueden llegar a hablar con él. Sin embargo, como muchas personas socialmente ineptas, pierden muchas oportunidades para iniciar una conversación. La siguiente escena es un ejemplo:

GUIL.: No se me ocurre cómo podemos empezar la conversación.

(*Hamlet entra en escena y se detiene... Ros. y Guil. lo observan.*)

ROS.: No obstante, supongo que se podría decir que esta era una posibilidad... Uno bien podría abordarlo... Sí, definitivamente me parece que ésta es una oportunidad... Algo así como un acercamiento informal directo... de hombre a hombre... directo desde el hombro... Oye, mira, de qué se trata todo esto... O algo por el estilo. Sí. Sí, parece que hay que aferrarlo con ambas manos, diría yo... si me preguntaran... porque a caballo regalado, etcétera. (*Avanza hacia Hamlet pero sus nervios lo traicionan. Retrocede.*) Estamos intimidados, ése es nuestro problema. Llegado el momento, sucumbimos ante su personalidad... (pág. 55)

Además, cuando Hamlet finalmente les habla, Rosencrantz y Guildenstern son incapaces de mantener la conversación: dan respuestas breves y elementales a sus preguntas y no ofrecen reciprocidad ni ayudan a llevar adelante el diálogo. Se limitan a recibir de manera pasiva las palabras de Hamlet, sin comprometerse activamente en la conversación. Como consecuencia, sólo les quedan las pocas palabras que Hamlet les ha dirigido y lo que los demás han dicho de él; nada de ello resulta particularmente esclarecedor. A Rosencrantz le parece que "La mitad de lo que dijo quería decir otra cosa y la otra mitad no quería decir nada" (pág. 41).

## La comunicación no verbal

Gran parte del problema de Rosencrantz y Guildenstern se plantea porque, al tratar de obtener información sobre la condición psíquica de Hamlet, recurren casi exclusivamente a los contenidos verbales de su discurso y el análisis que hacen se centra en los significados precisos de las expresiones de Hamlet. Desalentado, Guildenstern declara: "Palabras, palabras, palabras... Es todo lo que tenemos" (pág. 31).

Los psicólogos, tradicionalmente, también han centrado su atención en las palabras, la estructura de la lengua, la sintaxis y la gramática. Pero, como descubren Rosencrantz y Guildenstern, hablar no es sólo cuestión de palabras. El Intérprete, líder de un grupo de actores o trágicos llamados a representar ante la Corte real, los saca de su error. Dice: "Estamos atados a un idioma que compensa con oscuridad lo que le falta en estilo", y sugiere que es la gestualidad, la "pantomima" (las conductas no verbales que acompañan al lenguaje), lo que hace que la comunicación resulte más o menos comprensible.

Al igual que el personaje del Intérprete de Stoppard, varios teóricos (especialmente Goffman, 1956, y Harre y Secord, 1972) subrayan el parecido entre el teatro y la vida cotidiana, y consideran a la conducta social como una actuación dramática bajo el control continuo de la retroalimentación con su público. El modelo de las habilidades sociales adopta una posición similar. Consecuentemente, muchos psicólogos han centrado su atención, en los últimos tiempos, en la gestualidad, en aspectos no verbales de la comunicación tales como la mirada, la expresión facial, la postura, el gesto y la distancia (la forma en que se percibe, estructura y utiliza el espacio físico), y en rasgos supralingüísticos de la comunicación, como la entonación, la forma de expresión y el tono de voz.

Una de las consecuencias de esta división del interés entre el discurso y la actuación es que, en la literatura psicológica, suelen mantenerse separados los campos de la comunicación verbal y no verbal. Por lo general, se da por sentado que hay dos lenguajes diferentes con dos funciones diferentes: el verbal, para transmitir la información y el no verbal (o "lenguaje del cuerpo") para la actitud y los sentimientos. Pero en la práctica no es posible separar estos dos "canales", y Beattie sostiene que es poco lo que se puede ganar si se lo hace en la teoría. Afirma que ambos se interrelacionan dinámicamente, como los pasos y el estilo en una rutina de baile.

## La sincronización del discurso y el movimiento

El análisis de los patrones del discurso ha revelado que existen estrechos vínculos organizativos entre el discurso y la conducta no verbal en el nivel más microscópico, y que esos vínculos son probablemente esenciales para la comunicación eficaz. Además, estos estudios han revelado que la comparación entre la conversación y la danza es especialmente apropiada. Condon y Ogston (1967) demostraron, en su detallado microanálisis de una interacción filmada de tres personas, que "el cuerpo baila al mismo ritmo que el discurso", ya que en el movimiento corporal de un hablante se producen cambios sutiles y pasajeros sincrónicamente con la organización articulada y segmental de su conversación. Además, esta autosincronía (o congruencia) es sólo una de las formas en que se coordinan el discurso y las conductas no verbales.

Condon y Ogston también descubrieron que, durante la conversación, las personas interactuantes comparten con el hablante patrones armónicos de movimientos corporales.

Su estudio sobre un esquizofrénico (1966) sugiere que estos vínculos organizativos pueden ser esenciales para la comunicación exitosa. Los esquizofrénicos suelen ser personas con dificultades de comunicación y se descubrió que, aunque entre el terapeuta y su paciente se mantenía una sincronía interactiva, la autosincronía no funcionaba, y ciertas partes del cuerpo no estaban en sincronía con otras.

Una forma de sincronía interactiva más evidente es la que Scheflen (1964) denominó "congruencia postural", en la que dos o más personas adoptan posturas idénticas o especulares durante la conversación. Por ejemplo, dos o más personas en un pequeño grupo pueden estar sentadas con las piernas cruzadas, los brazos cruzados sobre el pecho y las cabezas inclinadas hacia un mismo lado. Esto no es coincidencia, ya que cuando un miembro de un conjunto congruente cambia su posición, los otros suelen hacer lo mismo rápidamente y así se mantiene la congruencia con los reiterados cambios en la postura corporal. Scheflen sugiere que, en general, la congruencia postural señala semejanza en los puntos de vista o los roles en el grupo, o en la condición social. Cuando la condición social de un miembro de un grupo difiere marcadamente de la de los demás, puede reflejarse en una postura diferente.

El Hamlet de Shakespeare, a pesar de todas sus supuestas aflicciones, reconoce claramente la importancia de los estrechos vínculos organizativos

sultar creíbles. Freud (1905, pág. 94) admite que la información puede transmitirse tanto verbal como no verbalmente y que a veces nuestros cuerpos traicionan a nuestros discursos:

> Si sus labios callan, habla con las yemas de sus dedos; la traición le sale por todos los poros.

Parecería que las acciones hablan más fuerte que las palabras. Ekman y Friesen (1969) sugieren que los "mentirosos" no verbales convincentes (como los actores profesionales, los abogados experimentados, los diplomáticos astutos, los buenos jugadores de poquer y los estafadores exitosos) han fortalecido la conciencia de su conducta no verbal y se han entrenado con la ayuda de retroalimentación externa focalizada (de un entrenador, un director, un público o un cliente) para lograr una mayor efectividad en sus simulaciones.

Mentir es, por lo tanto, una habilidad en sí, y Rosencrantz y Guildenstern no logran dominarla, como observa el propio Hamlet:

> Lo delatáis de tal modo en vuestros semblantes, que vuestra honradez no se da traza para ocultarlos (...)
>
> (*Hamlet*, acto II, escena 2).

## Los turnos en la conversación

Otro ejemplo de la forma conjunta en que se organizan las conductas verbales y no verbales es la convención aparentemente universal de los turnos al conversar; generalmente, una persona habla y luego deja que hable otra. Duncan (1972) ha identificado diversas señales que marcan el final de una secuencia discursiva, de modo que una persona sabe cuándo comenzar a hablar. Estas señales suelen incluir cambios en la entonación, el timbre o el volumen, el arrastre de las palabras o el acento en una sílaba final, alguna clase de expresión estereotipada como "¿no es verdad?" o "Ud. me entiende", o un modo tal como la terminación con un gesto o la relajación de una posición tensa de la mano durante el propio turno de conversación.

El desarrollo de los turnos en la conversación se remonta hasta las interacciones básicas entre madres e hijos (Rogers, 1985; Newson, 1978). No obstante, aunque es evidente en los intercambios entre los niños muy pequeños y sus madres, los turnos son una habilidad adquirida y algunas personas los manejan mejor que otras. Esto tiene consecuencias sociales importantes, ya que la violación de los turnos mediante la interrupción o el habla simultánea suele considerarse una grave falta de etiqueta o hasta de respeto.

Mediante nuestra asociación con otros, aprendemos la forma en que utilizan las señales para los turnos, lo cual nos permite juzgar con precisión cuándo están a punto de dejar de hablar o de comenzar a hacerlo. Esa familiaridad se hace muy evidente en los intercambios de Rosencrantz y Guildenstern, cuyos turnos para la conversación son a veces tan ajustados y precisos que da la impresión de que cada uno está completando las oraciones del otro, como en la siguiente conversación sobre Hamlet:

ROS.: Habla solo, lo cual debe significar locura.

GUIL.: Si lo que dijera no tuviera sentido, pero lo tiene.

ROS.: Lo cual sugiere lo contrario.

INTÉRP.: ¿De qué?

GUIL.: Creo que lo tengo. Un hombre que habla cosas sensatas consigo mismo no está más loco que otro que habla sin sentido con otros.

ROS.: O está tan loco como él.

GUIL.: O está tan loco como él.

ROS.: Y él hace ambas cosas.

GUIL.: Así que ahí lo tienes.

ROS.: Sano de remate.

Beattie (1983) informa sobre una cantidad de estudios que sugieren que la conducta de turnos en la conversación está influida por variables sociales como la condición social y la confianza. Parecería que cuanto más confianza sienta una persona por otra, más probable es que la interrumpa; y cuanto más insegura sea en lo social, menos probable es que lo haga. Rosencrantz y Guildenstern interrumpen con frecuencia al Intérprete, que es socialmente inferior a ellos, pero jamás interrumpen a Hamlet que, aunque es su amigo desde la infancia, es heredero de la corona y por lo tanto superior a ellos. Del mismo modo, nunca interrumpen a ningún otro miembro de la Corte real.

Ya que la condición social y la confianza son factores significativos para los turnos en la conversación, podría suponerse que los hombres, cuya condición social suele ser superior a la de las mujeres, podrían interrumpir más que ellas. Algunos estudios sobre la conversación entre adultos (Zimmerman y West, 1975) y entre niños (Esposito, 1979) apoyan esta hipótesis, pero en un estudio sobre las tutorías universitarias, Beattie (1979) no encontró diferencias sexuales en la frecuencia o en el tipo de interrupciones. Sin embargo, la condición social dentro de las universidades puede estar menos determinada sexualmente que en muchos otros contextos, de modo que las convenciones sociales "normales" pueden resultar allí menos evidentes. Los resultados del estudio de Beattie sugieren que la interrupción se ve afectada por el contexto social así como también por otras variables.

Algunas de estas variables fueron aclaradas más en un estudio sobre entrevistas políticas (Beattie, 1982). En un análisis de dos entrevistas televisadas durante 1979, una con el entonces primer ministro británico Callaghan y otra con Margaret Thatcher, que lideraba la oposición, Beattie descubrió que el entrevistador interrumpió a Margaret Thatcher el doble de veces que ella lo interrumpió a él, mientras que Callaghan interrumpió a su entrevistador mucho más de lo que fue interrumpido por él. Sin embargo, cuando se la interrumpía, Margaret Thatcher seguía hablando hasta terminar su idea, sin importar cuánto durara la conversación simultánea. Parecería, por lo tanto, que la insistencia de Margaret Thatcher para hacerse oír es lo que ha originado la difundida idea de que es una mujer estridente y dominante en sus entrevistas, a diferencia de la serena afabilidad de Callaghan. Beattie sugiere que esto no puede explicarse según las diferencias de condición social acordadas a los dos políticos ni a la condición femenina de Margaret Thatcher sino que hay que considerarlo según su respectivo uso de las señales para los turnos en la conversación; es decir, según las diferencias en la habilidad conversa-

cional. Sugiere que Margaret Thatcher ofrece señales de turnos en los lugares incorrectos y que los entrevistadores comprensiblemente malinterpretan esas señales. Es posible que Margaret Thatcher sea interrumpida con frecuencia porque deja caer su tono al final de las oraciones y parece que con ello indica que ha terminado de hablar, cuando en realidad no ha terminado; o porque suele hacer gestos con sus manos después de que la interrupción ha comenzado y no antes.

Sin embargo, en este sentido hay que recordar que Margaret Thatcher es sumamente hábil en la oratoria y el debate públicos. Durante sus discursos, expresa sus puntos de vista de un modo convincente y enérgico. Bien puede ser que adopte una forma similar en las entrevistas: la de hablar a su entrevistador en vez de hablar con él. Ciertamente, cuando es entrevistada, no parece esperar darle la palabra a otro orador, lo cual sugiere que bien puede considerar a la entrevista como una plataforma desde la cual dirigirse a un amplio público y no como una conversación en sí.

## Las investigaciones sobre interacción social

Como señala el Intérprete en *Rosencrantz y Guildenstern han muerto*, "hay un diseño en todas las artes" (pág. 59) y le toca al psicólogo descubrir, identificar y describir el diseño existente en la interacción social. Tradicionalmente, esto se ha hecho comparando grupos de ejecutantes eficientes e ineficientes, para determinar dónde radica la diferencia en sus conductas sociales, y luego enseñando las habilidades deseables a aquellos que son deficientes en ese sentido. La psicología social ha realizado en los últimos tiempos muchos aportes al entrenamiento en habilidades sociales. Durante el decenio 1970-1980 se hicieron muchas investigaciones en el campo de la comunicación no verbal, gran parte de las cuales se relacionaba con la codificación de la información interpersonal; es decir, la forma en que las personas intentan comunicar o, en algunos casos, ocultar sus actitudes, emociones, ideas y pensamientos. Los descubrimientos de las primeras investigaciones en este campo (Ruesch & Kees, 1956; Davitz, 1964; Grant, 1969; Mehrabian, 1970) fueron divulgados por Fast (1972). Estas investigaciones se vieron estimuladas por los avances en la grabación en video, que permitían un microanálisis de las conductas faciales y corporales fugaces, y también por la necesidad de muchos políticos y figuras públicas de presentarse efectiva y persuasivamente a través de la televisión.

Fueron los etólogos (como Hinde, 1972) quienes realizaron estudios sobre las conductas faciales y corporales; muchos de ellos se habían sentido impresionados por las notables semejanzas en la conducta expresiva de hombres y primates como los chimpancés. Varios de esos estudios se centraron en niños de menos de 5 años de edad, basándose en la idea de que las conductas no verbales son más fáciles de observar antes de que el lenguaje verbal se haya desarrollado plenamente (Blurton-Jones, 1972; McGrew, 1972). Muchos trabajos se ocuparon de analizar e identificar diferencias interculturales en la expresión facial y la gesticulación (Birdwhistell, 1968; Eibl-Eibesfeldt, 1972; Ingham, 1972; Ekman, 1972; Ekman y Friesen, 1976), y Morris (1977) atrajo la atención popular hacia estas variables, con su éxito editorial *Manwatching*.

## La mirada

En los últimos quince años se han hecho muchas investigaciones sobre la mirada, que previamente no había sido tenida en cuenta. La mirada tiene una importancia considerable ya que, si uno no mira en la dirección correcta en el momento apropiado, no recibirá información perceptual. La mirada es también una señal social importante. Las primeras investigaciones en este campo fueron realizadas por Neilson, en Harvard, en 1962, y posteriormente Exline trabajó mucho sobre el tema en América, y Argyle, Kendon y Cook lo hicieron en Inglaterra. Argyle y Cook (1976) afirman que en general las investigaciones realizadas apoyan la idea de que las personas que miran más y sostienen la mirada más tiempo crean una impresión más favorable y son más apreciadas que aquellas que no lo hacen. La gente que observa más tiempo resulta más amistosa y confiada. Por el contrario, quienes miran menos y sostienen la mirada menos tiempo parecen estar a la defensiva, ser evasivos y tienden a sentirse menos cómodos socialmente.

## La decodificación del afecto facial

Para responder efectivamente a las comunicaciones de los demás, es esencial la percepción precisa de la información transmitida. Esto se puede considerar como la decodificación de mensajes interpersonales. Rosenthal (1979) ha demostrado que las personas socialmente más hábiles son deco-

dificadores más precisos. Se han hecho varios intentos en la búsqueda de un código comprensivo para la expresión facial de la emoción, o el afecto, para los humanos (Lanzetta y Kleck, 1970; Ekman, Friesen y Tomkins, 1971; y Cuceloglu, 1972). Pero el sistema más comprensivo desarrollado hasta la fecha es el Código de Acción Facial (CAF), desarrollado por Ekman y Friesen (1978). Deriva de un análisis de la base anatómica de los movimientos faciales. Mide el propio rostro en vez de la información que de él puede deducirse, que ha sido el centro de atención clásico en las investigaciones sobre las expresiones faciales desde las obras pioneras de Darwin (1872). Una vez identificadas las diferentes acciones del rostro según los grupos musculares involucrados, Ekman y Friesen proporcionan un manual de auto-enseñanza, de entrenamiento, como referencia para la decodificación de la conducta facial. Comprende material textual que describe cada acción facial de acuerdo con su base muscular y los cambios en la apariencia, información sobre combinaciones de movimientos, ejemplos de fotografías y películas y detalles del desarrollo del CAF, estudios de confiabilidad y validez, experimentos en los que se ha utilizado el código con éxito y áreas de aplicación posibles. El CAF se ha usado para entrenar a diferentes grupos de profesionales en los Estados Unidos y sus autores señalan que se lo ha utilizado para aislar signos faciales entre personas sordas que usan lenguaje de signos y para medir los cambios en niños que han sufrido daño cerebral.

## Entrenamiento en las habilidades sociales

Para permitir que la gente interactúe más eficiente y agradablemente y evite las consecuencias indeseables de habilidades sociales deficientes, se han incluido, en el entrenamiento ocupacional de muchos grupos profesionales, los resultados de las investigaciones sobre las habilidades sociales, y se los ha ofrecido en programas y cursos terapéuticos para el público en general. Aquí volvemos a encontrar al psicólogo en su rol de entrenador. Es función de los psicólogos administrar ese entrenamiento en casi todos los ámbitos concebibles, ya sea ocupacionales, clínicos, educativos, sociales o de cuidado. Por ejemplo, en la Universidad de Keele, a los médicos clínicos se les enseña a realizar entrevistas con los pacientes y, en una sociedad crecientemente multirracial, se ha difundido la enseñanza de las diferencias interculturales entre varios grupos profesionales. De este modo, se entrena a la policía esta-

dounidense para que mantenga una cierta distancia física cuando detiene gente de raza negra en la calle (Garratt y otros, 1981), y los hombres de negocios británicos se entrenan en conductas sociales árabes.

El entrenamiento en entrevistas, enseñanza, supervisión y terapia para las habilidades sociales profesionales es una de las principales áreas de interés. Sin embargo, hace tiempo que se admite que los métodos clásicos de educación tales como conferencias, debates y textos resultan insuficentes para afectar las habilidades sociales. Al igual que las habilidades motrices como andar a caballo o nadar, las habilidades sociales no pueden enseñarse de esta manera, y cada vez se enfatiza más la importancia del *role playing* y de varias formas de entrenamiento grupal. La forma más común de entrenamiento en las habilidades sociales es lo que Argyle (1969) denomina "aprendizaje en el trabajo", en el que no se da a la gente entrenamiento específico alguno además de la posibilidad de practicar una habilidad en una situación relevante. Pero, como señala Argyle, esto no garantiza necesariamente una mejora en la habilidad social, ya que, para empezar, mucha gente es socialmente incompetente y sólo persiste en el uso de técnicas sociales deficientes. Para ser eficente, el aprendizaje en el trabajo requiere oportunidades para la práctica de la habilidad, la retroalimentación y la generación de nuevas respuestas. La gente puede adquirir y adquiere habilidades sociales de esta forma pero no necesariamente aprende las formas más eficientes. El método tiene, por lo tanto, sus limitaciones, pero Argyle sugiere que estas se pueden superar parcialmente si, en las situaciones reales, alguien actúa como entrenador.

### *Role-playing* y simulación

A menudo se emplean técnicas teatrales en las que los entrenados actúan la parte que les va a tocar y reciben realimentación inmediata sobre su desempeño, de parte de un entrenador o de sus pares en el entrenamiento. Estas técnicas suelen usarse para el entrenamiento de vendedores y docentes, y han mejorado mucho con los avances en las grabaciones en video.

## Grupos de entrenamiento

La idea del entrenamiento grupal se inspiró en el psicólogo de la gestalt Kurt Lewin, que reconoció que el entrenamiento en las habilidades de re-

lación humana había sido sumamente descuidado. Promovió la formación de grupos para el entrenamiento en esas habilidades. El primero de esos grupos se reunió en Maine, en 1947, y poco después se establecieron los *National Training Laboratories* (Laboratorios Nacionales de Entrenamiento). Durante el decenio 1950-1960, muchas comisiones de empleo y escuelas de negocios de los Estados Unidos enviaron sus alumnos a los grupos coordinados por los Laboratorios Nacionales para desarrollar sus habilidades gerenciales y su sensibilidad interpersonal. Inicialmente, los grupos de hasta 12 personas se reunían una vez por semana en un período de varias semanas, para observar la naturaleza del proceso grupal, la interacción social y las relaciones humanas; gradualmente, esta experiencia se fue conociendo como entrenamiento en sensibilidad. El entrenamiento tenía objetivos comprensivos y de largo alcance, que sufrieron un marcado cambio cuando se reconoció que la participación en un grupo de entrenamiento o entrenamiento en sensibilidad solía conducir a fuertes experiencias personales: revelaciones drásticas y conmovedoras, expresiones emotivas catárticas, aumento en la sensación de bienestar y cambio personal. Los efectos parecían similares a los de la psicoterapia, y en el decenio 1960-1970 los grupos de entrenamiento perdieron rápidamente su identidad como grupos de entrenamiento en habilidades sociales y se convirtieron en el Encuentro. Argyle (1969) señala que mientras que entre el 30% y el 40% de los participantes de estos grupos se ha beneficiado con el entrenamiento, muchos otros se han visto gravemente perturbados por la experiencia.

El método de entrenamiento en habilidades sociales más utilizado en la actualidad incluye elementos de trabajo grupal, *role playing* y entrenamiento en la vida real. Estos elementos incluyen explicación y modelado, seguidos de *role playing*, retroalimentación verbal de un entrenador y mediante grabaciones en video y, posteriormente, ensayo. Sin embargo, la generalización a las situaciones reales sigue siendo problemática y por lo tanto suele darse "tarea para el hogar" entre sesiones, y se le requiere al participante que repita ejercicios en ámbitos reales e informe sobre ellos. Aun así, las sesiones de *role playing* pueden ser demasiado ajenas a ciertas situaciones de la vida real como para resultar eficientes, y por lo tanto en algunos entrenamientos profesionales, como en el policial, se ha enfatizado más la práctica en el trabajo.

Se han realizado muchos estudios de evaluación del entrenamiento en las habilidades sociales. Shepherd (1983) revisó 52 de esos estudios y llegó a la conclusión de que el entrenamiento en las habilidades sociales adquiere cierto valor para los que tienen deficiencias sociales y ha resultado bastante

exitoso en los ámbitos ocupacionales. Argyle (1984) es cuidadoso en sus afirmaciones respecto del entrenamiento en las habilidades sociales, pero sugiere que está empezando a satisfacer una inmensa necesidad humana. Aunque no les habría resultado de mucha utilidad a los desfortunados Rosencrantz y Guildenstern, que demasiado tarde reconocieron que "las ruedas se han puesto en movimiento y van a su propio ritmo, al cual nosotros estamos... condenados" (pág. 44).

## Lecturas adicionales

ARGYLE, M. (1969) *Social Interaction.* Londres: Tavistock y Methuen.

ARGYLE, M. (1983) *The Psychology of Interpersonal Behaviour.* Harmondsworth: Penguin.

Son textos que proporcionan una introducción a la psicología de la interacción social y las habilidades sociales.

ARGYLE, M. (1984) "Some new developments in social skills training". *Bulletin of the British Psychological Society,* 37, 405-410.

Este breve artículo proporciona una actualización comprensiva sobre los avances en este campo.

# 12

# ¿Cómo mejorar las habilidades de aprendizaje?

*James Hartley*

Este capítulo se divide en dos partes. En la primera, analizaremos cuatro ideas para mejorar las habilidades de aprendizaje, que surgen de las investigaciones de la psicología cognitiva. En la segunda parte, consideraremos las implicaciones de estas ideas para el entrenamiento en las habilidades de estudio.

## La psicología cognitiva y el aprendizaje

Si se le preguntara a un grupo de psicólogos cognitivos cuáles serían las consideraciones principales a tener en cuenta en un análisis del aprendizaje, habría muchas respuestas. Pero las más mencionadas serían nociones sobre:

- la organización del material;
- la profundidad de procesamiento;
- el papel del conocimiento previo;
- el conocimiento que tiene la gente sobre sus propias habilidades de aprendizaje.

Este capítulo comienza describiendo experimentos realizados por psicólogos que ilustran cada una de estas cuatro áreas. En algunos casos, se trata de material similar al que ya se ha visto en capítulos anteriores, por lo que se ofrece como una posibilidad para de consolidar las ideas.

### ¿Cómo ayuda al aprendizaje la organización del material?

Se suele decir que el sello distintivo de la buena enseñanza es su buena organización. ¿Qué significa esto? ¿Pueden demostrarse sus ventajas? Comparemos los siguientes textos:

El zorro se parece al perro. A veces se queda en un agujero. Se alimenta de pequeños animales. Es rojo. Habita en los bosques. También le gusta la fruta. Tiene un gran hocico. Caza aves para alimentarse. Vive entre los arbustos. A veces come pescado. Su cola le cuelga. Duerme en el suelo.

El zorro se parece al perro. Tiene un gran hocico. Es rojo. Su cola le cuelga. Habita en los bosques. Duerme en el suelo. Vive entre los arbustos. A veces se queda en un agujero. Se alimenta de pequeños animales. Caza aves para alimentarse. A veces come pescado. También le gusta la fruta.

El segundo texto presenta la misma información que el primero, pero supongo que nadie dudará de que está mejor organizada. La información se ha agrupado en tres temas principales: cómo es el zorro, dónde vive y qué come. Danner (1976) usó textos como estos para analizar los efectos de la organización en el recuerdo. Demostró que los niños de siete, nueve y once años invariablemente recordaban mejor el segundo fragmento. En varios otros estudios se han hecho descubrimientos similares. (Ver, por ejemplo, Bransford, 1979; Howe, Brainerd y Kingsma, 1985).

Estos estudios se ocupan de la organización de los elementos en el texto. Los autores de libros de texto tratan de ayudar a sus lectores organizando los materiales que presentan. Los libros se dividen en capítulos, los capítulos en secciones, que a su vez se subdividen en parágrafos y estos en oraciones, etc. En cada nivel, la secuencia de elementos es importante si se quiere que la estructura resulte clara para el lector. Recursos tales como los índices, los títulos de los capítulos y los subencabezamientos se utilizan para hacer más evidente la estructura, del mismo modo que los usos más sutiles de la tipografía y el diseño (Hartley, 1985).

Sin recursos de este tipo, gran parte de la organización impuesta por el autor de un libro de texto (o el docente) puede no ser clara para el alumno. Los alumnos, como veremos más adelante, tienen conocimientos previos y expectativas diferentes respecto de lo que se está enseñando, y esto puede influir en la forma en que perciben el material que se les presenta. De hecho, las evidencias existentes sugieren que el aprendizaje mejora más si los alumnos también imponen su propia estructura al material que están aprendiendo.

En un importante estudio, Tulving (1962) presentó a sus alumnos una lista integrada por dieciséis palabras sin relación entre sí, en un orden diferente cada vez, y después de cada presentación les pidió que las recordaran en

cualquier orden. Tulving analizó cada uno de los recuerdos, para ver si sus alumnos empezaban a agrupar las palabras, sin importar el hecho de que cada vez aparecieran en un orden diferente. Y esto fue lo que algunos de sus alumnos hicieron. Además, los que agrupaban ciertas palabras y las aprendían en los agrupamientos subjetivos obtenían mejores resultados que los alumnos que no lo hacían.

Los estudios sobre los alumnos que organizan sus materiales en situaciones de aprendizaje más naturales (por ejemplo, reformulando la información en diagramas y tablas de resumen o usando notas esquemáticas) también sugieren que esos procedimientos organizativos son de considerable valor (Davies y Greene, 1984; Jonassen, 1984).

## ¿Cómo influye en el aprendizaje la profundidad a la que se procesa el material?

Para considerar esta pregunta, me gustaría presentar antes el siguiente experimento (basado en Howe, 1984). Lea la siguiente lista de palabras y responda mentalmente cada pregunta por turno:

| | |
|---|---|
| discurso | ¿Es una forma de comunicación? |
| CEPILLO | ¿Está escrito con minúsculas? |
| mejilla | ¿Rima con "piso"? |
| CERCO | ¿Rima con "terco"? |
| LLAMA | ¿Es algo caliente? |
| suelo | ¿Es un tipo de casa? |
| miel | ¿Está escrito con minúsculas? |
| CUCHILLO | ¿Rima con cazador? |
| guante | ¿Está escrito con minúsculas? |
| cobre | ¿Rima con "montaña"? |
| OVEJA | ¿Es un mueble? |
| MONJE | ¿Está escrito con mayúsculas? |

Volveré a ocuparme de estas palabras en unos pocos minutos (cuando ustedes hayan tenido tiempo suficiente para olvidarlas). Mientras tanto, me ocuparé del tercer tema.

## ¿Cómo afectan los conocimientos previos al aprendizaje?

Los conocimientos previos pueden tanto ayudar como dificultar el aprendizaje. Rosemary Driver (1983), en *The Pupil as Scientist?* (¿El alumno como científico?) da ejemplos que demuestran que muchos alumnos llegan a las clases de ciencias con ideas ya formadas, que difieren de las teorías que el docente puede querer desarrollar, y que esto puede dificultar el aprendizaje. Estas ideas influyen en las observaciones que hacen los niños en sus experimentos y en las explicaciones que dan sobre ellos. Driver señala que lo que le parece obvio al docente puede no serlo para el alumno y viceversa. Por eso, por ejemplo, los docentes tienen que decirles a sus alumnos que miren por el microscopio y deben ayudarlos a dibujar "líneas de fuerza" cuando bosquejan patrones de limaduras de hierro en torno a un imán.

Un estudio de Morris y otros (1981) demuestra claramente cómo, por otra parte, los conocimientos previos pueden ayudar a la memoria. Morris y sus colegas tomaron a un grupo de alumnos una prueba de 42 puntos sobre su conocimiento del fútbol profesional inglés. (La prueba contenía preguntas como: ¿Quién es el arquero del Manchester? ¿Quiénes juegan en Brisbane Road? ¿Quién es el entrenador del Celtic? ¿Cuál fue el último equipo inglés que ganó la copa UEFA?) Los alumnos completaron la prueba un sábado por la tarde y luego, a las 5 de la tarde, escucharon los resultados de 64 partidos que se habían jugado ese día. Después, se les presentó una lista de los partidos y se les pidió que recordaran los resultados.

Las respuestas demostraron que los alumnos que habían respondido bien en la prueba sobre conocimientos de fútbol recordaban mucho mejor los resultados que aquellos a los que les había ido mal. El promedio de recuerdo para los partidos de primera división (11 partidos) fue de 7,8 para diez alumnos que habían obtenido 25 o más en la prueba sobre conocimientos de fútbol, y de sólo 3,3 para once alumnos que habían obtenido 5 o menos. En general, la correlación entre los conocimientos previos y el recuerdo exitoso fue de 0,81.

Si los conocimientos previos pueden ayudar u obstaculizar el aprendizaje, parece razonable decir que la buena enseñanza debería empezar por lo que los alumnos ya saben. Esto es más fácil de decir que de hacer, pero hay ciertas demostraciones del principio en acción. Wittrock, Marks y Doctorow (1975), por ejemplo, intentaron usar el conocimiento previo de los niños como base para el aprendizaje de vocabulario. Tomaron parte en el estudio cuatro grupos de niños de entre 10 y 11 años. El grupo 1 leyó dos ve-

ces un fragmento que contenía varias palabras difíciles y luego se verificó su comprensión de las palabras. El grupo 2 escuchó el fragmento dos veces y luego se lo sometió a la misma prueba. Los grupos 3 y 4 leyeron y escucharon, respectivamente, el mismo fragmento, y luego se sometieron a la misma prueba, pero en la primera ocasión se reemplazaron las palabras difíciles por sinónimos de las mismas, más sencillos y conocidos. Sólo en la segunda lectura se incluyeron las palabras difíciles. Con todo, los grupos 3 y 4 obtuvieron mucho mejores resultados que los grupos 1 y 2 en la prueba de vocabulario. Su familiaridad previa con las palabras más fáciles les permitió dar significado a las palabras más difíciles.

## Profundidad de procesamiento (nueva visita)

Volvamos ahora a la lista de palabras que examinamos hace unos minutos. Observe las 24 palabras que citamos a continuación y vea cuántas puede reconocer de las doce palabras originales. No mire la lista anterior.

| | | |
|---|---|---|
| BRUJA | CUCHILLO | CONDE |
| CERCO | LANA | MONJE |
| BRUTO | LLAMA | CALAVERA |
| CHARCA | MIEL | CUERPO |
| MEJILLA | ESPECIA | DISCURSO |
| ENFERMERA | PINTOR | GUANTE |
| SUELO | OVEJA | AGUA |
| ESTRELLA | CEPILLO | COBRE |

Con la lista original de doce palabras, se pidió una respuesta a una pregunta sobre cada una. Había tres tipos de preguntas: sobre apariencia visual (mayúsculas o minúsculas), sobre sonidos (qué palabras rimaban con otras) y sobre significados (por ejemplo, si una oveja era un mueble).

Experimentos como éste (con listas mucho más largas) han sugerido que uno recordará mejor algo si tiene que procesarlo o pensar en ello más profundamente (Craik y Tulving, 1975). Por lo tanto, en este experimento, es esperable que el lector recuerde la mayoría de las palabras *discurso, suelo, llama* y *oveja*, porque las preguntas sobre estas palabras se relacionaban con su significado. Es probable que recuerde menos las palabras *cepillo, guante, miel* y *monje*, porque las preguntas sobre estas palabras sólo tenían que ver

con su aspecto. Su puntaje en el grupo de palabras restante (*mejilla, cuchillo, cerco* y *cobre*) probablemente sea intermedio, ya que las preguntas sobre estas palabras implicaban cierto nivel de procesamiento pero no tanto como las preguntas sobre su significado. (No se preocupe si su puntaje individual no es así: sólo quiero sugerir el patrón promedio de resultados obtenidos cuando hay muchos participantes.) Resultados como estos sugieren que la gente recuerda mejor cuando piensa en algo o cuando lo procesa más profundamente. (Ver también el capítulo 5.)

La idea de que cuanto más profundamente se procese la información mejor se recordará es atractiva (aunque a veces es difícil de comprobar). Pero, hasta ahora, sólo hemos considerado breves presentaciones de materiales no relacionados entre sí y su recuerdo. ¿Qué sucede con las situaciones de aprendizaje más naturales? Hay conjuntos de experimentos que sugieren que varias estrategias de procesamiento de la información pueden afectar notablemente el recuerdo de largo plazo.

Estas estrategias (incluidas en la tabla 12.1) pueden considerarse de acuerdo con la profundidad de procesamiento involucrado. Por ejemplo, Jonassen (1984) sugiere que tomar notas pautadas es más eficiente que tomar notas de manera convencional o subrayar, y yo mismo he sugerido antes que la reorganización de la información es una estrategia de aprendizaje más eficaz que la mera repetición.

Otros estudios analizan la forma en que los alumnos leen los textos de enseñanza. Estos estudios distinguen entre los alumnos que adoptan lo que se denomina un "enfoque superficial" para estudiar y los que adoptan un "enfoque profundo" (Marton, Hounsell y Entwistle, 1984).

Los alumnos que adoptan un enfoque superficial para el aprendizaje probablemente aprendan las cosas "como loros", para poder escupirlas en los exámenes, y se concentren en el recuerdo de ejemplos más que en los principios organizativos. En este capítulo, por ejemplo, una persona así podría recordar que tuvo que hacer un experimento que tenía que ver con el procesamiento de palabras pero no recordar que lo que pretendo es *hacer que el lector considere los aportes de la psicología cognitiva al estudio del aprendizaje*. Los alumnos que adoptan un enfoque profundo, en cambio, intentarían buscar los principios organizativos del texto y relacionar las ideas nuevas con los conocimientos previos ganados en los capítulos anteriores. El alumno "superficial" encara la tarea con la intención de memorizar y completar las exigencias de la misma; el alumno "profundo" encara el texto con la intención de comprender.

Hay evidencias que apoyan la idea de que los alumnos "profundos" obtienen mejores resultados en los exámenes, pero es algo contradictoria. Esto puede deberse, en parte, a que:

a) algunos alumnos parecen capaces de cambiar de estrategia (según se trate de material nuevo o no y de por qué se lo esté aprendiendo)

b) muchos exámenes fomentan la memorización y por eso los alumnos pueden aprobarlos con un enfoque superficial. Marton y otros (1984) hacen un análisis muy útil de estos temas.

*Tabla 12.1. Estrategias de procesamiento de la información para el aprendizaje*

| |
|---|
| • Estrategias de práctica del material. Por ejemplo: copiar, subrayar, repasar. |
| • Estrategias de organización del material. Por ejemplo: crear encabezamientos, agrupamientos, resúmenes, mapas. |
| • Estrategias de integración del material. Por ejemplo: resumir, hacer diagramas, construir tablas de resumen integradoras. |
| • Estrategias de elaboración de la información. Por ejemplo: usar la imaginación, crear analogías, parafrasear, relacionar la nueva información con la vieja. |

*Nota: Muchas estrategias (como el repaso) pueden clasificarse de varias formas. Jonassen (1984) comenta algunas de las estrategias menos comunes (como los mapas y las redes).*

## ¿Cómo afecta el aprendizaje nuestro conocimiento sobre nuestras propias habilidades de aprendizaje?

El conocimiento que tenemos sobre nuestras propias habilidades de aprendizaje se denomina metacognición. En el capítulo 6, Angus Gellatly analiza el concepto de metamemoria según lo que los niños saben sobre sus aptitudes memorísticas. El concepto de metacognición es similar pero de alcance más amplio. La metacognición infantil (o la adulta, para el caso) puede analizarse de acuerdo con lo que los niños saben o creen saber sobre sus propios procesos cognitivos (y esto incluiría a su memoria).

El concepto de metacognición es importante por dos motivos. Primero, nos permite explorar el conocimiento que tenemos sobre nuestras propias habilidades de aprendizaje. Segundo, nos permite considerar en qué medida las estrategias que hemos analizado hasta aquí (organización del material, relación entre el material nuevo y el viejo, fomento del procesamiento más profundo) pueden volverse parte de nuestras propias actividades cotidianas cuando estamos aprendiendo.

Consideremos primero cuánto conocimiento podrían tener los adultos sobre sus propios procesos de aprendizaje. En un estudio realizado en 1979, Linda Baker presentó seis textos de unas 250 palabras cada uno a un grupo de quince alumnos. Cada texto contenía tres párrafos, y el párrafo del medio tenía alguna clase de problema, que podía ser: a) incoherencias, b) referencias incorrectas o confusas a lo anterior, c) conectores lógicos incorrectos (por ejemplo, el uso de "sin embargo" en vez de "además"). Primero, se pidió a los alumnos que leyeran los seis textos. Luego, se les pidió que recordaran el párrafo del medio de cada uno de los textos. Finalmente, se les comentaron los tres tipos de problemas y se les preguntó si los habían notado.

Los resultados mostraron que los problemas no afectaron notoriamente el libre recuerdo de los párrafos del medio, aunque los alumnos tuvieron tendencia a recordar menos que un grupo de control que había leído los mismos párrafos pero sin esos problemas. Los alumnos del grupo experimental ignoraron los problemas, los corrigieron o hicieron deducciones correctas o incorrectas al darles sentido a los párrafos.

Los alumnos notaron un 23% de los problemas cuando estaban leyendo y sólo el 38% cuando se les pidió que los buscaran en los textos. La conclusión de Baker fue que algunos alumnos pueden controlar y de hecho controlan su comprensión mientras están leyendo (al menos, en una situación experimental) y que se puede lograr que algunos lo hagan (por ejemplo, si se les pide que busquen los errores) pero que muchos otros no.

La investigación posterior de Baker y Anderson (1982), usando los mismos textos presentados oración por oración en terminales de computadora, demostró más precisamente que los alumnos pasaban más tiempo leyendo las oraciones que tenían problemas y que las revisaban con mayor frecuencia. Pero, incluso en este estudio, sólo el 62% de los alumnos detectó uno o más problemas.

Estudios como estos (con niños y adultos) han alentado a los investigadores a ver si pueden ayudar a los alumnos a evaluar su comprensión mientras leen, para aumentar su procesamiento metacognitivo. Brown y otros

(1983), por ejemplo, desarrollando el concepto de aprendizaje autocontrolado. La idea es que los alumnos controlen, verifiquen y evalúen su progreso en el aprendizaje. En el estudio de un texto, por ejemplo, sugieren que a los lectores se les deben enseñar cinco aptitudes críticas: el resumen (la autorrevisión), la interrogación, la aclaración, la predicción y el control de su comprensión.

Un ejemplo típico de uno de los muchos estudios llevados a cabo por Brown y sus colegas es el siguiente: los investigadores debían ayudar a los alumnos a resumir textos; sobre la base de varios experimentos previos, Brown y Day (1983) aislaron las estrategias que se ilustran en la tabla 12.2, usadas (con diferentes niveles de éxito de parte de alumnos y de expertos) al resumir textos.

Brown, Campione y Day (1981) describen las tendencias evolutivas en el uso correcto de estas estrategias: por ejemplo, los niños de 6 y 7 años se concentran en las estrategias de borrado y los de 15 años sólo usan bien la estrategia de "inventar una oración temática" en el 30% de los casos, a diferencia de los expertos, que lo hacen bien casi siempre. Brown y otros (1981) se interesaron en enseñar a los alumnos a resumir mejor. Para ello, diseñaron un estudio que comparaba tres métodos de enseñanza:

a) *enseñanza "ciega"*: se alentaba a los alumnos de manera general para que hicieran un buen resumen, incluyeran las ideas principales y omitieran las trivialidades y todas las palabras innecesarias, pero no se les daban reglas para lograrlo.

b) *enseñanza informada*: igual que el método anterior, pero se explicaban las reglas para hacerlo y se realizaba ejercitación.

c) *aprendizaje autocontrolado*: igual que (b), pero se proporcionaba entrenamiento explícito en cómo controlar las reglas: es decir, se mostraba a los alumnos cómo asegurarse de que se eliminaran todas las trivialidades y redundancias, de que las listas de ítem fueran reemplazadas por subordinadas, de tener una oración temática para cada párrafo y de saber si necesitaban inventar alguna oración temática.

Los resultados de este estudio demostraron que hubo beneficios para todos los alumnos en su capacidad de escoger e inventar oraciones tópicas con los tres métodos de enseñanza, pero que las técnicas de aprendizaje autocontrolado resultaban especialmente efectivas con los alumnos de menor rendimiento y con las tareas más difíciles.

Los estudios sobre aprendizaje auto-controlado de Brown y sus colegas sugieren que las mejoras en el aprendizaje pueden conseguirse: a) estructurando más la tarea, y b) haciendo más explícito el reconocimiento del éxito. Parece que en todo hay efectos relacionados con la edad y la capacidad, pero que el aprendizaje auto-controlado (donde la responsabilidad de la evaluación del éxito del aprendizaje pasa del docente al alumno) es una poderosa técnica de aprendizaje.

---

*Tabla 12.2. Ejemplos de verbalizaciones de los expertos sobre las reglas del resumen.*
*(Tomado de Brown y Day, 1983, y reproducido con permiso)*

| | |
|---|---|
| Borrado | "En un resumen de este tipo, los detalles, se eliminan. Se necesitan generalizaciones, no detalles" (trivialidades). "A este texto le sobran oraciones. Hay dos que dicen lo mismo: que los animales del desierto tienen hábitos nocturnos por el calor. Se puede omitir una." (redundancia) |
| Subordinación | "Una cosa que hice fue eliminar las diferentes clases de plantas. En vez de escribir 'margaritas, amapolas, caléndulas y lirios' sólo puse 'plantas anuales' y nuevamente omití los detalles y sólo hice generalizaciones." |
| Selección de oración temática | "Esta oración contiene los puntos esenciales del párrafo; especifica el proceso por el cual se conserva la vida de la planta. Hay que incluirla en cualquier resumen." |
| Invención de oración temática | "El párrafo habla del ciclo de las plantas anuales, que producen semillas, esperan las lluvias, florecen, vuelven a producir semillas, etc. Aunque no lo dice explícitamente, sólo hay que mencionar el ciclo y omitir lo demás." |
| Combinación entre párrafos | "En los dos primeros párrafos, la única información esencial es lo que dice del calor y la falta de agua en el desierto. Los combinaré en dos oraciones que contengan toda la información que necesito. Una oración es simple; la otra, es compuesta." |

## Implicaciones para la enseñanza de habilidades de estudio

Los cursos y textos tradicionales sobre las habilidades de estudio se concentran en dar a los alumnos información sobre diversas habilidades, como tomar notas, redactar composiciones y prepararse para exámenes, sin importar la materia (sea Inglés, Matemáticas o Psicología). Esos cursos descuidan los problemas que plantean las diferentes materias y también las diferencias individuales sobre el tema de aprendizaje en las percepciones de los alumnos y, de hecho, cómo estas percepciones pueden ir cambiando con el tiempo (ver Perry, 1981). También descuidan los problemas planteados por las ideas de aprendizaje autocontrolado.

Los cursos y textos modernos sobre habilidades de estudio tienden a centrarse más en todos estos aspectos. Por ejemplo, un curso de esa naturaleza diseñado por Gibbs (1981) considera las actividades de los alumnos dentro de disciplinas específicas. Para tomar notas, por ejemplo, Gibbs les pide primero a los alumnos que tomen notas de un texto en particular, que es parte del tema de enseñanza. Luego les pide que trabajen en equipos de dos y analicen las notas del otro e intenten comprender por qué están escritas como están. Luego, los equipos se unen en grupos de cuatro alumnos y cada uno de ellos debe explicar las notas de su compañero de equipo a los otros tres. Después, se pide a los cuatro alumnos que consideren cuáles son las características que hacen que sus notas sean "buenas", "pobres" o "inútiles" y qué señalen las estrategias útiles y las que es conveniente evitar. Se escriben estas características para después poder leerlas a toda la clase en una sesión plenaria final.

Gibbs adopta un enfoque similar con las habilidades de escritura. Sugiere que se deben proporcionar a los alumnos dos composiciones sobre el mismo tema, escritas por alumnos anteriores, con perspectivas marcadamente diferentes. Los alumnos trabajan primero de manera individual y responden a las siguientes preguntas: ¿Qué composición es mejor? ¿Por qué? ¿En qué difieren? Luego, trabajan en parejas para comparar sus comentarios y elaborar una respuesta conjunta a las mismas preguntas. Después, en grupos de cuatro, se les pide que combinen sus conclusiones y describan lo que piensan que pretendía cada uno de los dos autores de las composiciones. Finalmente, con toda la clase, cada grupo cuenta a los demás sus conclusiones y luego se realiza una discusión general.

Otros cursos modernos de habilidades de estudio se concentran en la enseñanza de habilidades generales y específicas. Habilidad general es la capacidad de elegir, para la tarea que se va a encarar, un conjunto de subhabilidades específicas adecuadas, a partir de una gran variedad de posibilidades. Se enseñan conjuntos de estrategias a los alumnos, que pueden desplegarse de manera diferente en las distintas situaciones de aprendizaje.

Para el análisis de cómo las diferentes estrategias pueden comprometer a los alumnos en un procesamiento más profundo, incluí, en la tabla 12.1, conjuntos de estrategias de procesamiento de la información. A éstas les podemos agregar ahora dos clases más: las estrategias metacognitivas y las de apoyo. Las estrategias metacognitivas, como se analizó antes, se ocupan de las técnicas de control de la propia comprensión (tales como revisión, parafrasis, imágenes, predicción y preguntas para verificar la propia comprensión). Las estrategias de apoyo tienen que ver con las formas de encarar los problemas que provocan las distracciones y la inquietud, y con la organización del propio entorno para el estudio efectivo. Estas estrategias incluyen: tener un lugar especial para estudiar habitualmente, fijar y cumplir fechas, planificar horarios de estudio y hacer ejercicios de relajación para reducir la ansiedad. Weinstein y Mayer (1985) describen cursos de habilidades de estudio en todos los niveles educativos (de primarios a terciarios) que se han centrado en el procesamiento de información particular, las estrategias metacognitivas y las de apoyo. Son escasos los estudios que combinan entrenamiento en varias estrategias de los tres grupos, pero Weinstein y Mayer comentan cuatro ejemplos. Aquí, sólo para ilustración, describiré dos de esos enfoques integradores.

*Estudio 1.* Dansereau y otros critican a gran parte de los primeros cursos sobre habilidades de estudio porque consideran que son demasiado amplios y no prestan suficiente atención a la psicología cognitiva actual. Desde mediados del decenio 1970-1980, Dansereau y sus colegas han diseñado varios programas de entrenamiento en estrategias de aprendizaje, que buscan remediar estas deficiencias (por ejemplo, Dansereau y otros, 1979; Diekhoff, Brown y Dansereau, 1982; Dansereau y otros, 1983; Dansereau, 1985).

El programa de Dansereau más conocido es el representado por el acrónimo MURDER ("asesinato", en inglés), en el cual:

M = *mood* (establecer el *humor* para el estudio)

U = *understanding* (leer para *entender* las ideas difíciles e importantes)

R = *recalling* (*recordar* el material sin referirse al texto)

D = *digesting* (*digerir* el material, corrigiendo los errores del recuerdo y amplificando y reprocesando el material)

E = *expanding* (*ampliar* el conocimiento por medio de preguntas)

R = *reviewing* (*repasar* los errores y aprender de las evaluaciones).

El acrónimo MURDER describe un conjunto de seis estrategias ejecutivas fundamentales (es decir, qué hacer). Cada estrategia ejecutiva va acompañada de varias subestrategias operativas más detalladas (es decir, cómo hacerlo). Así, establecer el humor para la enseñanza incluye el uso de técnicas de relajación y conversación constructiva con uno mismo (por ejemplo: "Ahora, lo que tengo que hacer... Bien... Listo... Perfecto"). Recordar incluye utilizar subestrategias como la paráfrasis, el uso de imágenes, mapas y el análisis de las ideas cruciales. Además, el control continuo de la propia comprensión es una subestrategia operativa de todas las estrategias ejecutivas.

Para evaluar la eficacia de este esquema, Dansereau y otros (1979) entrenaron a 38 estudiantes de Psicología en el uso de estas subestrategias ejecutivas y operativas, dos horas semanales durante quince semanas. En las comparaciones con los grupos de control, los grupos experimentales evidenciaron mejoras significativamente superiores en las subsiguientes pruebas de comprensión y recuerdo y en los autoinformes sobre las habilidades de aprendizaje.

*Estudio 2.* Paris, Cross y Lipson (1984) diseñaron un programa de entrenamiento denominado *Estrategias Informadas para el Aprendizaje* (EIA), para ayudar a los niños más pequeños de la escuela a mejorar sus habilidades de lectura. Este programa de entrenamiento en particular duraba catorce semanas e incluía varias características de enseñanza directiva, tales como: a) centrar la atención del niño en el material que se va a enseñar, b) generar altos niveles de compromiso, y c) proporcionar práctica frecuente y retroalimentación inmediata. Además, los módulos de enseñanza del programa contenían una "gran variedad" de actividades diseñadas para ayudar a la comprensión: incluían lecciones sobre la comprensión de los propósitos de la lectura, activar el conocimiento relevante de fondo, destinar atención a la idea principal de cada párrafo, detectar inferencias y controlar la propia comprensión.

El programa EIA fue probado en dos grupos de alumnos de 8-9 años y dos grupos de 10-11 años (cada uno con un grupo de control). Los alumnos de los grupos de EIA obtuvieron mayores beneficios que los de los grupos de control en las pruebas de completamiento de un texto con las palabras faltantes y en las tareas de detección de errores, pero no mostraron mejoría alguna en dos pruebas estandarizadas de comprensión de textos. Sin embargo, Paris y sus colegas sostienen que los grupos de EIA demostraron un mayor conocimiento de las estrategias de lectura.

Estos dos estudios destacan las principales preocupaciones de quienes trabajan en el campo de las estrategias de aprendizaje. Las investigaciones incluyen: a) un análisis detallado de qué implica la realización de tareas específicas, b) el desarrollo de conjuntos de estrategias dentro de un sistema total, para encarar los diferentes aspectos de las distintas tareas, y c) una evaluación del éxito del producto. Los resultados de los estudios de evaluación se usan luego para modificar y mejorar el programa de entrenamiento original.

## Conclusiones

En este capítulo, he resumido brevemente algunas de las áreas de la psicología cognitiva que se relacionan con el aprendizaje. Hemos puesto el acento más en los experimentos de laboratorio que en los estudios sobre los efectos del entrenamiento en las habilidades de estudio, fundamentalmente porque hay más experimentos que estudios sobre sus efectos, y son más precisos. Sin embargo, hay dos críticas importantes a los estudios de laboratorio en este contexto:

a) su marcado uso de materiales y condiciones experimentales simplificados.

b) su concentración en el aprendizaje verbal (a expensas de lo práctico, social y afectivo).

Estas deficiencias no han pasado inadvertidas y, sin duda, las futuras investigaciones procurarán remediarlas.

### Lecturas adicionales

BRANSFORD, J.D. (1979) Human Cognition: *Learning, Remembering and Understanding*. Belmont: Wadsworth.

Buena reseña general sobre psicología y aprendizaje cognitivos.

NISBET, J. y SHUCKSMIT J. (1986) *Learning Strategies*. Londres: Routledge and Kegan Paul.

Transfiere el trabajo sobre metacognición a prácticas para docentes.

# *Sección 4*

## *La habilidad y el pensamiento*

# 13

# La habilidad en el razonamiento

*Angus Gellatly*

Algunos lectores conocerán ya la siguiente adivinanza: "Tres indios caminan en fila por un sendero en el bosque. El primero es padre del segundo, pero el segundo es hijo del tercero. ¿Cómo puede ser?"

A muchos, si no han visto antes una adivinanza del mismo tipo, les resultará confusa, al menos por un rato. La dificultad surge del tipo de interpretación que la gente tiende a dar a las oraciones que integran la adivinanza, la tendencia a ir más allá de la información que realmente se proporciona. Naturalmente, la solución es que el tercer indio es la madre del segundo indio, pero si uno inconscientemente ha dado por supuesto que los tres indios eran hombres, puede ser difícil hallarla.

Hay al menos dos factores que pueden fomentar un supuesto, o inferencia, en cuanto al género de los tres indios. En primer lugar, las referencias masculinas de las palabras "hijo" y "padre" pueden crear un ámbito, o expectativa, tendiente a la masculinidad[1]. En segundo lugar, puede estar operando una tendencia a asignar valores naturalmente masculinos cuando no se especifica el género, como en el uso común de pronombres masculinos en los casos en que el género es indefinido. Sea cual fuere la importancia relativa de estos dos factores, la adivinanza funciona aprovechando el hecho (analizado por John Sloboda en el capítulo 4 y por Don Rogers en el capítulo 10) de que la comunicación mediante el lenguaje sólo es posible sobre la base de un conjunto de supuestos compartidos entre hablantes/escritores y oyentes/lectores. Para entender un discurso o texto, el oyente/lector debe hacer todo tipo de inferencias más allá de lo que dice el texto, y la importancia de la adivinanza sobre los indios es que demuestra cuán injustificadas pueden ser esas inferencias.

En este capítulo analizaremos la propensión de la gente a hacer inferencias al razonar sobre cuestiones cotidianas o sobre problemas lógicos, de modo que es necesario que seamos claros respecto del significado de la palabra "inferencia". Se dice que se hace una inferencia cuando, sobre la base de afirmaciones particulares (premisas), sacamos una conclusión que no está

*1. En inglés esto es más marcado aún, ya que no existen diferencias de género en los sustantivos en general: Indian es tanto masculino como femenino (N. del T.)*

explícitamente contenida en las afirmaciones originales. Para tomar un ejemplo del capítulo 4, si leemos: "Juan iba camino de la escuela; no tenía ganas de dar lección de Matemáticas", podemos inferir que Juan es un alumno al que no le va muy bien en Matemáticas, o que es un docente. Bien o mal, hemos ido más allá de las premisas hacia una conclusión. Hemos hecho inferencias.

Los filósofos distinguen entre dos tipos de inferencia: la deducción y la inducción. Como el tema central de este capítulo es el papel que juegan en el razonamiento humano estos dos tipos de inferencia, debemos analizar más detalladamente la diferencia entre ellos.

## Inducción

Además de la comprensión del lenguaje, casi todas las conductas y pensamientos se basan en supuestos e inferencias que no se pueden justificar definitivamente, idea que originalmente destacó el filósofo David Hume en el siglo XVIII. Una inferencia que sólo afirma que la conclusión es probablemente verdadera es una inducción. Por ejemplo, por mi observación de las nubes oscuras que hay hoy en el cielo y mi conocimiento de que habitualmente llueve cuando hay nubes oscuras, inductivamente infiero que hoy va a llover, tomando esta conclusión sólo como algo probable. Mi conclusión puede expresarse verbalmente o mediante una acción: en el caso del cielo nublado, tomo mi impermeable antes de salir de casa. La vida (y el lenguaje) sería imposible si no nos permitiéramos hacer inferencias que no es seguro que sean correctas. Muchos de nuestros mejores empeños se basan en inferencias de conclusiones probables, y parte de la destreza depende de juzgar con precisión los grados de probabilidad. Sin embargo, desde que Aristóteles descubrió (o, según el punto de vista, inventó) el razonamiento deductivo, el pensamiento occidental reverencia de manera especial las inferencias que se consideran concluyentes y no meramente probables.

## Deducción

Una inferencia en la que las premisas garantizan la conclusión es una deducción, y puede cobrar la forma de un silogismo. Un ejemplo famoso de silogismo válido es:

(1) Todos los hombres son mortales.
Sócrates es hombre.
Por lo tanto, Sócrates es mortal.

Al leer este silogismo uno comprende que la conclusión sigue naturalmente a las premisas. Si Sócrates no fuera mortal, podríamos asegurar con confianza que una o las dos premisas deben haber sido falsas. Esta confianza surge de nuestro sentido intuitivo de que el silogismo tiene una forma válida, y la forma de un argumento válido es aquella en la cual las premisas verdaderas deben siempre llevar a una conclusión verdadera.

Desde Aristóteles, las deducciones han sido algo impresionante para la gente. El conocimiento al que se llega deductivamente, como en las matemáticas, parece especialmente satisfactorio. Por ejemplo, Euclides pudo establecer un pequeño número de premisas (o postulados, como él los llamó) y deducir a partir de ellas toda su geometría. Al hacerlo, fijó un ideal deductivo-matemático con el que tienden a medirse los métodos de todas las otras ramas del conocimiento.

## Deducción y razonamiento

Una consecuencia de la estima en que se llegó a tener a la deducción fueron los interrogantes que se plantearon acerca de su relación con la racionalidad y el razonamiento humano. ¿La racionalidad es la capacidad de razonar de acuerdo con una lógica deductiva? Si esto es así, ¿todos los humanos pueden ser considerados racionales? ¿O sólo algunos? ¿O ninguno? Una posición que se puede adoptar en este tema (y la han adoptado, entre otros, los filósofos Kant y Boole) es que las reglas de la lógica deductiva son descriptivas del pensamiento humano, y el pensamiento cumple sus reglas. Esas reglas son como en el siguiente par:

| *Modus Ponens* | *Modus Tollens* |
|---|---|
| Si p implica q | Si p implica q |
| p | No q |
| ergo q | ergo no p |

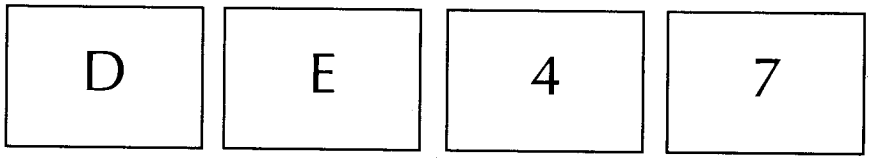
En vez de concebir al razonamiento como algo guiado por una lógica mental, se lo puede considerar una habilidad, tal como todas las otras formas de cognición de las que se ocupa este libro. Los desempeños expertos tienen que entenderse en sus propios términos. No recurrimos a lógicas subyacentes para explicarlos sino a evidencias de los rasgos característicos de la habilidad, como el agrupamiento de entrada y salida, la mejora con la práctica, la fluidez y rapidez de las respuestas, la automaticidad y demás (capítulo 2). De modo que si el razonamiento es una habilidad, debemos poder analizarlo sin recurrir a la lógica mental. De hecho, el debate sobre si el razonamiento se puede entender mejor en términos de habilidad o de lógica mental, o en términos compatibles con alguno de los dos enfoques, es complejo (Johnson-Laird, 1982, 1983; Gellatly, 1986). En el resto de este capítulo, analizaremos algunas de las evidencias a favor y en contra de ambos puntos de vista sin pretender acabar con la discusión en un sentido o en otro. Para empezar nuestra investigación sobre el razonamiento, podemos ver un problema diseñado por Peter Wason (1966).

## El problema de Wason

El problema de Wason se refiere a cuatro tarjetas, que se ilustran en la figura 13.1. Sólo se ve un lado de esas tarjetas, que muestran los símbolos D, E, 4 y 7, respectivamente. Se nos dice que cada tarjeta tiene una letra de un lado y un dígito del otro, y no se duda de que esto sea así. Luego, se nos da la siguiente regla: "Si una tarjeta tiene una D de un lado, tendrá un 4 del otro lado". El problema consiste en decir qué tarjetas hace falta dar vuelta (y cuáles, por lo tanto, no es necesario dar vuelta) para comprobar si esa regla es cierta.

No hay trampas en este problema, aunque puede ser de ayuda subrayar que la regla no es reversible, de modo que la inversa no es aplicable. Intente averiguar la respuesta usted mismo antes de seguir leyendo.

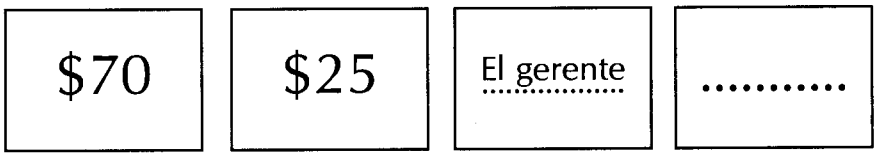
*Figura 13.1 El problema de selección de tarjetas de Wason*



## El problema de D'Andrade

Antes de revelar la solución correcta al problema de Wason, veamos un segundo problema, originalmente creado por Roy D'Andrade (ver Rumelhart y Norman, 1981). Imagine que usted es empleado de una gran tienda y supervisa los recibos de ventas. La empresa tiene una norma: "Si una compra excede los $50.-, el recibo de venta debe estar firmado al dorso por el gerente". Le muestran cuatro recibos como los de la figura 13.2: uno es por $70, otro es por $25; uno está firmado al dorso y el otro no está firmado. El problema consiste en decir qué recibos hace falta dar vuelta (y cuáles no) para comprobar si se ha respetado la norma. Una vez más, intente resolver usted mismo este problema.

*Figura 13.2.La versión de D'Andrade del problema de selección*




Los experimentos han demostrado que la mayoría de la gente da la solución correcta al problema de D'Andrade. Eligen dar vuelta el recibo de $70 y el que no está firmado. Es bastante evidente que la norma se viola sólo si el recibo que supera los $50 no está firmado (contraviniendo el *modus ponens*) o si el recibo no firmado resulta ser por una cifra superior a los $50 (contraviniendo el *modus tollens*). Fácil, ¿no?

Volvamos ahora al problema de Wason. Mucha menos gente (el 10% o menos de los estudiantes universitarios) pueden conseguir la respuesta correcta aquí. La solución consiste en dar vuelta la tarjeta que tiene una D y la que tiene un 7. El problema es decidir si la regla dada es verdadera o falsa. Si la combinación de símbolos en cualquiera de las cuatro tarjetas rompe la regla, la regla es falsa. De lo contrario, es verdadera. La regla dice que "si una tarjeta tiene una D de un lado, tendrá un 4 del otro lado". Sólo si la tarjeta con la D no tiene un 4 del otro lado se violará la regla, o también si la tarjeta que no tiene un 4 (y por lo tanto tiene un 7) tiene una D al dorso. Sin embargo, lo más frecuente es que la gente elija dar vuelta la tarjeta con la D y

la tarjeta con el 4, en vez de las tarjetas con la D y el 7. Es decir, dan vuelta las tarjetas que muestran los símbolos mencionados en la regla. En vez de resolver el problema, simplemente "adaptan" sus respuestas a los símbolos mencionados. La adaptación no suele ser un proceso consciente y a veces los sujetos experimentales sostienen acaloradamente que la solución correcta es "D y 4" y no "D y 7" (Wason, 1983).

## El razonamiento y la memoria

Lo interesante de esta diferencia en el desempeño en los dos problemas es que se trata de dos versiones del mismo problema; tienen exactamente la misma estructura lógica. De hecho, el objetivo de D'Andrade, al crear su problema, fue precisamente encontrar una versión más fácil del problema de Wason, y si a alguien le resulta difícil aceptar la solución correcta en el de Wason puede comprender claramente su forma elaborándolo en paralelo con la versión de D'Andrade. Los problemas con la misma estructura lógica pueden ser fáciles o difíciles, según el contenido que los rellene. No se trata simplemente de que los contenidos abstractos sean más difíciles que los concretos o realistas, ya que algunas versiones concretas y realistas del problema de Wason son tan difíciles como su versión abstracta.

¿Por qué, entonces, la gente responde bien a la versión de D'Andrade del problema? No hace falta haber sido empleado de una tienda; la experiencia o la comprensión de tipos de relación análogos alcanza. Griggs (1983) sostiene que la gente no resuelve en realidad la versión de D'Andrade sino que recuerda un problema análogo que alguna vez enfrentó antes y del cual ya sabe la solución. Esto sugiere que la gente tenderá a resolver un problema novedoso sólo si reconoce su semejanza con un problema familiar para el cual ya tiene la solución o un procedimiento para solucionarlo.

Para decirlo de una forma levemente distinta, pero que enfatiza igualmente el papel de la memoria: desde el capítulo 5 sabemos que los problemas que imponen una pesada carga a la MCP, o memoria a corto plazo operativa, son difíciles de resolver. Una forma de reducir la carga de la MCP es el agrupamiento de los datos que se ingresan siguiendo esquemas almacenados en la MLP. De este modo, será más fácil interpretar ("tener en mente") un problema que se presenta de una forma que se ajusta a uno o más esquemas (o modelos) de la MLP que un problema que no se ajusta a ningún esquema almacenado en la MLP. Uno de los modelos de la MLP (de los que

ya se habló en el capítulo 5) es el que se denomina libreto. La versión de D'Andrade extrae de la MLP el libreto sobre cómo actúan en las tiendas los gerentes y los vendedores, y el problema puede así interpretarse razonablemente. Pero la versión original de Wason no se ajusta a ningún esquema familiar. Hay que interpretarlo más o menos palabra por palabra en la memoria operativa. Esto supera la capacidad de la MCP y el resultado es que la mayoría de la gente no llega a comprender bien el problema.

Según lo que acabamos de ver, el papel del razonamiento es raramente (si alguna vez lo es) el de tomar la información dada sobre un problema y deducir directamente una conclusión. Es más probable que el razonamiento actúe retrospectivamente para justificar una solución a la que se ha llegado por otros medios. Ese punto de vista sobre la forma en que se resuelven los problemas es coherente con la idea de la regla de producción *"si... entonces..."* presentada en el capítulo 2. La *acción* de revisar superficialmente un procedimiento de solución sólo se realiza si el problema está configurado de una forma que satisfaga las *condiciones* para esa *acción*. La secuencia de *acción* puede dispararse automáticamente y ofrecerá una solución que luego puede justificarse mediante el razonamiento. Aprender a configurar los problemas nuevos para que se asemejen a los conocidos es el proceso por el cual se mediatiza, o transfiere, la generalización (ver capítulo 8).

Los estudios realizados sobre las diferencias entre novatos y expertos resolviendo problemas de física apoyan la idea de que la resolución deductiva de problemas depende de la activación de las producciones almacenadas en la memoria de largo plazo y que la generalización se puede hacer aprendiendo a configurar los problemas de la manera adecuada. Los físicos profesionales reconocen (configuran) tipos de problemas más rápidamente y responden a ellos con procedimientos de solución más amplios y automáticos. Sus soluciones se ejecutan en unos pocos cálculos (a veces en uno solo), en vez de en una serie de pasos más pequeños. Incluso cuando los novatos son capaces de ejecutar la acción de un procedimiento de solución, suelen tener dificultades para reconocer las condiciones en las que resulta adecuado. Larkin, McDermott, Simon y Simon (1980) sugieren que puede entenderse a la intuición matemática como este reconocimiento de esquemas altamente desarrollado sobre las condiciones a partir de las cuales una acción asociada surge automáticamente.

## El razonamiento y las acciones

Hemos visto que la familiaridad con el contenido de los problemas verbales influye en el éxito con el que se los puede enfrentar. Se ha descubierto que lo mismo sucede con los problemas no verbales que se deben resolver más con acciones que con palabras.

Kendler y Kendler (1963) hicieron una prueba con niños en un aparato que comprendía tres paneles: A, B y C. Primero, se presentaba el panel A y se enseñaba a los niños que si apretaban un botón obtenían una bolita. Luego, se reemplazaba al panel A por el B, en el que si apretaban un botón conseguían un rodamiento. A continuación, se presentaba el panel C junto con una pila de bolitas y rodamientos, cualquiera de los cuales podía entrar en una ranura del panel. Los niños aprendían que si insertaban una bolita en la ranura conseguían una recompensa, que podía ser un juguete o una golosina, pero que si insertaban un rodamiento no conseguían nada. Finalmente, se les presentaban los tres paneles a los niños y se los invitaba a obtener una recompensa. Kendler y Kendler encontraron que hasta los 10 años los niños tenían dificultades para unir las dos acciones: apretar un botón del panel A para obtener una bolita y luego insertarla en el panel C para conseguir la recompensa. Combinar las dos cosas en una "inferencia de acción" era muy difícil para los niños más pequeños.

Cole, Gay, Glick y Sharp (1971) repitieron este experimento en Liberia y hallaron que los adultos Kpelle iletrados tenían un rendimiento pobre, y sólo alrededor del 8% alcanzaba la solución. ¿Los Kpelle eran incapaces de hacer la inferencia? Cole y sus colegas consideraron que la compleja maquinaria del aparato (que funcionaba según los principios de una máquina tragamonedas) podía resultar inhibitoria para los Kpelle. Transformaron la tarea en otra, con llaves y cerraduras, objetos que resultaban conocidos para los Kpelle. Reemplazaron los paneles A y B por dos cajas de fósforos diferentes e identificables que contenían, respectivamente, una llave roja y una negra. El panel C fue reemplazado por una caja cerrada que se podía abrir con una de las llaves para obtener una recompensa. Como antes, los sujetos aprendían primero qué llave había en cada caja y luego, separadamente, qué llave encajaba en la caja cerrada. Cuando se les presentó la tarea completa, más del 70% de los Kpelle lograron obtener la recompensa. El contenido del problema (en este caso, la naturaleza del aparato) influía en la tasa de solución. Cole y sus colegas pensaron que esto se debía a que la falta de familiaridad con el aparato inhibía a los sujetos y no les permitía pensar en el procedimiento

de solución. Su hallazgo no resultará soprendente, quizá, para quienes no se lleven bien con los equipos mecánicos.

## El razonamiento y la lógica

No resulta nada novedoso el hecho de que el éxito con un problema sea una función de su contenido así como también de su forma lógica. Ya Aristóteles era consciente de que la facilidad con la que se puede juzgar la validez de un silogismo varía según su contenido. Por ejemplo, veamos el siguiente silogismo:

2) Algunos miembros del Partido Laborista son autodidactas.
Ningún valón pertenece al Partido Laborista.
Por lo tanto, al menos algunos valones no son autodidactas.

Begg y Denny (1969) hallaron que el 55% de los alumnos juzgaron erróneamente que un silogismo similar al 2) era válido. A veces la gente defiende acaloradamente la validez de un silogismo de este tipo aun cuando se les dice que no es válido. Sin embargo, un cambio en los contenidos puede revelar de inmediato que la forma del silogismo es defectuosa. Por ejemplo, 3) tiene la misma forma que 2) pero muy poca gente lo considera válido.

3) Algunos mamíferos saben nadar.
Los peces no son mamíferos.
Por lo tanto, algunos peces no saben nadar.

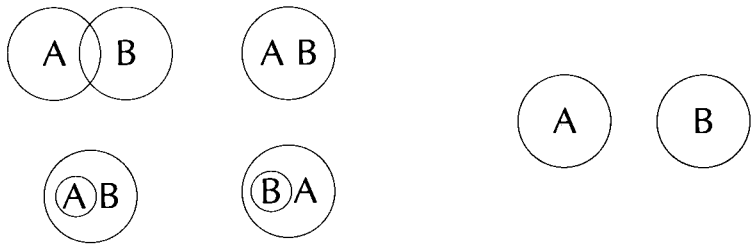
El hecho de que el contenido de los problemas tenga tanta relación con la capacidad de la gente para razonar correctamente sobre ellos podría tomarse como una negación de la doctrina de la lógica mental. Se podría afirmar que si hubiera una lógica mental, su eficiencia operativa debería depender sólo de la forma y no de los contenidos. Además, la posición alternativa (que el razonamiento es una habilidad) se ve apoyada por muchas evidencias. Sabemos que cuanto más tiempo pase alguien practicando la resolución de problemas matemáticos o de ajedrez, mejor será en ese tipo de problemas. Del mismo modo, la capacidad de evaluar categorías silogísticas puede mejorarse si se enseña la técnica de los círculos de Euler. Esta técnica emplea círculos para representar las categorías de las entidades mencionadas en el silogismo, como se muestra en la figura 13.3, lo cual también demuestra que una premisa dada puede tener más de una representación posible. Es com-

prensible que la mayoría de la gente tienda a cometer más errores cuando hay varias formas de representar un par de premisas que cuando sólo hay una representación posible (Johnson-Laird, 1983). En el caso anterior, simplemente hay más para pensar y por lo tanto hay más posibilidad de error. No obstante, sigue siendo cierto que la práctica y la adquisición de estrategias, o técnicas, lleva a mejoras en el desempeño, y éste es uno de los puntos claves de la habilidad.

Quienes defienden la lógica mental, sin embargo, no aceptan que la existencia en el razonamiento humano de errores que dependen de los contenidos haga insostenible su planteo. Una defensa común, recientemente esbozada por Henle (1962, 1978), consiste en afirmar que la razón misma nunca está en falta pero que a veces se la aplica a información distorsionada o mal percibida. En el caso de los silogismos, puede olvidarse una de las premisas, incorporarse al problema una premisa extra que deriva del conocimiento fáctico o malinterpretarse una premisa. Por ejemplo, se puede pensar que si se dice que "Todos los A son B", también significa que "Todos los B son A". Henle sugiere que cuando se realiza el esfuerzo de reconstruir el razonamiento real aplicado a un problema, se comprueba que la gente siempre razona de manera válida pero a veces es el problema lo que está equivocado. Alternativamente, se puede no intentar resolver el problema sino tratar de adivinar la respuesta, o evaluar la verdad de una conclusión en vez de la validez del argumento del cual se deriva. Sea cual fuere el caso, podría afirmarse que el razonamiento mismo no equivoca el camino.

Una de las dificultades de la propuesta de Henle es que no parece estar abierta a la refutación. Sea cual fuere la respuesta a un problema, siempre será posible hacer una reconstrucción que la haga parecer racional, siempre que se suponga un adecuado malentendido de las premisas. Por otra parte, no hay duda de que los malentendidos ofrecen una explicación plausible de muchos errores manifiestos en las tareas de razonamiento. Podemos ver esto muy claramente en relación con los estudios interculturales de la cognición.

*Figura 13.3. Algunos ejemplos del uso de los círculos de Euler*




(a)
Algunos A son B

(b)
Ningún A es B

## Lógica, cultura y lenguaje

En la cultura occidental siempre nos hemos preguntado si la gente de otras culturas tiene los mismos procesos mentales que nosotros. Frecuentemente, se ha afirmado que las culturas llamadas tradicionales se parecen a los niños occidentales en sus procesos psíquicos, y que ambos grupos difieren de los adultos occidentales (por ejemplo, Piaget, 1974). Este tipo de comparaciones demuestra, en la práctica, graves dificultades conceptuales (Gellatly, 1986). No obstante, los estudios interculturales han tenido resultados estimulantes.

Se han hecho varios estudios en los que se ha pedido a gente iletrada, sin escolarización, con frecuencia habitantes de pequeñas aldeas de diversos lugares del mundo, que razonara con silogismos. Scribner (1977) observa que cuando se les pide que repitan las premisas que se les acaban de leer, con frecuencia no pueden hacerlo. Omiten premisas o las sustituyen por conocimientos personales. De manera similar, cuando se les presentaba el siguiente silogismo:

4) Todas las mujeres que viven en Monrovia están casadas.
Kemu no está casada.
¿Kemu vive en Monrovia?

Una respuesta típica era: "Monrovia no es para una clase de personas solamente, y Kemu fue a vivir allí". Se ignora la condición hipotética del silogismo en favor de la información práctica relativa al tema de conversación

que se ha introducido; en este caso, la residencia en Monrovia. Como consecuencia de este hábito, los tipos de distorsión descritos por Henle se observan abundantemente. Y, como Henle, Scribner sostiene que la gente no alfabetizada siempre razona de manera lógica sobre sus propias interpretaciones de un problema, si es que razona sobre él.

Esta última aclaración es necesaria, porque a veces la gente iletrada rechaza por completo las tareas de razonamiento propuestas por los psicólogos. En el decenio 1930-1940, Luria (1976) presentó el siguiente problema ante un grupo de campesinos de Uzbek, en Asia Central:

5) En el norte todos los osos son blancos.
Novaya Zemlya queda en el norte.
¿De qué color son los osos allí?

Entre las respuestas que obtuvo, algunas decían: "Habría que preguntárselo a la gente que ha estado allí y los ha visto", y "Siempre hablamos exclusivamente de lo que vemos; no hablamos de lo que no vemos". Estas personas simplemente no estaban preparadas para jugar al juego de preguntas y respuestas de Luria, y su rechazo nos lleva a comprender que efectivamente se trata de un juego. Scribner (1977) dice que el juego depende de la comprensión de un tipo peculiar de discurso y lo que se hace evidente es que un juego (o un tipo de discurso) y sus reglas pueden darse por sentadas en una cultura pero resultar extraños para otra. El juego del silogismo (el tipo de discurso en el que se funda) resulta extraño en las sociedades iletradas no escolarizadas.

## Las reglas y el discurso

Ya hemos visto que la comprensión del lenguaje involucra al oyente en todo tipo de inferencias lingüísticas relativas al tema de conversación. Pero esto no es todo. Como ya se dijo en el capítulo 10, los oyentes también deben inferir las intenciones del hablante, lo que se puede denominar "sentido implícito" de una expresión (Clarke, 1977). Si alguien pregunta: "¿Sabe dónde queda la estación?", y uno responde: "Sí", uno no ha inferido correctamente las intenciones del otro o ha decidido ser hostil, irónico o algo por el estilo. Para inferir las intenciones del hablante (el sentido implícito) hace falta comprender a qué tipo de discurso pertenece su expresión. Los diferentes tipos de discurso tienen sus propios conjuntos de reglas (que se superpo-

nen) para guiar a hablantes y oyentes. Las reglas que rigen el discurso conversacional en castellano incluyen que la expresión "¿Sabe dónde queda la estación?" se interprete como: (a) un pedido de información y no un interrogatorio sobre los conocimientos geográficos del oyente, y (b) que requiere, en caso de ser afirmativa la respuesta, información sobre la forma de llegar a dicha estación. En el discurso conversacional de otras culturas, es posible que se mantenga la regla (a) pero en lugar de la regla (b) quizá los buenos modales exijan que el oyente intente ofrecer la información solicitada aun cuando no tenga idea de dónde está la estación. Esto puede ser desconcertante para el turista que le pregunta a algún residente local sin tener en cuenta la divergencia entre sus respectivas reglas de discurso.

Debería ser evidente la importancia de este análisis para el desempeño en los silogismos. Al preguntar si Kemu vive en Monrovia o de qué color son los osos de Novaya Zemlya, el psicólogo recurre a un conjunto de reglas del discurso que incluyen una regla que indica que los oyentes deben ignorar sus conocimientos personales, o su falta de esos conocimientos, y responder sólo a la información dada en las premisas, que bien puede ser falsa. En la conversación, ésta es una exigencia muy poco común (Gellatly, 1987) porque, como hemos visto, en general se espera que los oyentes aporten sus conocimientos sobre un tema en vez de omitirlos. De hecho, tan extraño resulta el pedido que los miembros de culturas iletradas no lo reconocen. Se aferran a sus reglas habituales de conversación y responden a la pregunta recurriendo a sus conocimientos personales. O, como en el caso de Luria, pueden llegar a increpar al psicólogo por violar las reglas con una pregunta que los lleva más allá de sus conocimientos personales.

## El razonamiento como habilidad

Cumplir las reglas del discurso es, de por sí, una habilidad, como lo analiza Don Rogers en el capítulo 10. Esto se plantea notoriamente cuando uno entra a un grupo social en el que se comparte cierta forma idiosincrática de sentido del humor, por ejemplo. Uno bien puede sentirse incapaz de interpretar rápidamente el sentido implícito de las expresiones de los hablantes, de distinguir los insultos en broma de los verdaderos insultos o las emociones reales de las fingidas. Pero si uno pasa en este grupo el tiempo suficiente, esos problemas se resolverán solos. Mediante la aplicación de las habilidades sociales que posee (capítulo 11), uno aprende a decodificar los

mensajes que se transmiten e ingresa en la forma de pensar y hablar del grupo. Esto nos dice, entre otras cosas, que hay más de un nivel de habilidad en el razonamiento deductivo. Ya hemos visto que la resolución de problemas deductivos es una actividad que implica habilidad, la cual puede, como en otros casos, estar representada en términos de sistemas de producción. El éxito en los problemas deductivos mejora con la práctica, que lleva a que los problemas se configuren en agrupamientos mayores, que activan automáticamente unidades más amplias de procedimientos de solución. Ahora vemos que también hay involucrada una forma previa de habilidad. Antes de poder involucrarse en la resolución de problemas deductivos, hay que dominar las reglas del discurso que rigen a esa forma de cognición. Esto consiste en hacer inferencias lingüísticas-discursivas que no son en sí deductivamente seguras; pueden resultar falsas. Así, la deducción, ese ideal del conocimiento seguro, se construye sobre los cimientos de la inferencia inductiva. Lejos de trascenderla, la deducción resulta ser dependiente de la capacidad de hacer inferencias inductivas.

Esto no significa que lo antedicho deba leerse como una desvalorización de la deducción. La configuración de argumentos de manera deductiva es, cuando resulta aplicable, un método de suma importancia para garantizar la fuerza y el rigor de nuestro pensamiento. En la ciencia, así como en muchos otros campos de la vida cotidiana, procedemos planteando una hipótesis y luego verificándola para ver qué consecuencias se pueden deducir válidamente de ella, y luego verificando esas consecuencias. También en las matemáticas el modelo euclideano de establecer postulados seguidos de deducciones sigue teniendo vigencia. En este capítulo hemos intentado demostrar que el éxito en los problemas de naturaleza deductiva es una cuestión de aprendizaje cultural, de habilidad cognitiva. La resolución de los problemas deductivos no difiere, en principio, de la resolución general de problemas, que es el tema del próximo capítulo.

## Lecturas adicionales

GILHOOLY K,J. (1981) *Thinking: Directed, Undirected and Creative*. Londres: Academic Press.

ANDERSON, J.R. (1984) *Cognitive Psychology and its implications* (2ª edición). San Francisco: Freeman.

Introducciones de fácil lectura al tema de la deducción.

JOHNSON-LAIRD P.N. (1983) *Mental Models*. Cambridge: Cambridge University Press.

Un tratamiento amplio pero más difícil sobre la deducción.

# 14

# La resolución de problemas

*Angus Gellatly*

Al empezar este capítulo, presentamos dos problemas para que el lector intente resolverlos.

> a) Una mañana, al amanecer, un monje budista empieza a escalar una montaña por un sendero estrecho que serpentea en torno a ella. Asciende a un ritmo regular de 4,5 km por hora.
>
> Después de 12 horas llega a la cima, donde hay un templo, y se queda allí meditando durante varios días hasta que, un amanecer, empieza el descenso por el mismo sendero, a un ritmo regular de 7,5 km por hora.
>
> Demuestre que el monje pasa por algún punto del camino a la misma hora del día a la ida y a la vuelta.
>
> (Basado en un problema creado por Duncker.)

> b) Usted está organizando un torneo de tenis por eliminación simple y tiene 111 inscriptos. Para cada partido se necesita un juego de pelotas nuevo. ¿Cuántos juegos de pelotas harán falta?
>
> (Basado en un problema creado por De Bono.)

Veremos las soluciones a estos problemas unas páginas más adelante. Por el momento, ya habrán cumplido su objetivo si han hecho pensar al lector en qué es lo que implica la resolución de problemas, por qué a veces es una tarea difícil y cómo se puede mejorar. Estos son los temas de los que se ocupa este capítulo. Examinaremos los procesos que constituyen la resolución de problemas y la cuestión de si se puede entrenar a la gente para que mejore su eficiencia en la resolución de problemas.

Pero, para empezar, se necesitan algunos comentarios generales sobre la naturaleza de los problemas y sus soluciones.

## La naturaleza de la resolución de problemas

Puede parecer extraño que este libro contenga un capítulo dedicado específicamente a la resolución de problemas. Después de todo, los capítulos anteriores también se han ocupado de las soluciones a los problemas relativos a la memoria, la comunicación y el estudio. ¿En qué sentido, entonces, este capítulo tiene más que ver con la resolución de problemas que los anteriores?

Eso depende de qué sentido le demos a la palabra *problema*. Newell y Simon (1972) definen un problema de la siguiente forma: "Una persona se enfrenta a un problema cuando desea algo y no sabe de inmediato qué serie de acciones puede desempeñar para conseguirlo". En esta definición, un problema se relaciona con las aptitudes del individuo. Para un niño de 4 años, recordar una lista de palabras que incluya dos animales, dos frutas y dos colores puede ser un problema, pero para el adulto alfabetizado ya no es un problema, porque da los pasos adecuados (la práctica categorizada) de manera automática. Del mismo modo, leer o conducir un automóvil plantean al novato toda clase de problemas que sencillamente no existen para el experto. Muchos de los problemas mencionados en los capítulos anteriores (aunque no todos) han sido de este tipo. Es decir, son problemas para los que se adquieren los procedimientos de solución en el curso de la evolución social y cognitiva normal, o como consecuencia del entrenamiento especial, y por lo tanto (según la definición de Newell y Simon) dejan de ser problemas.

Evidentemente, la gente puede aprender (y aprende) habilidades para eliminar todo tipo de problemas culturalmente identificados, desde el dominio de la gramática de la lengua hasta la navegación sin instrumentos, pasando por el juicio sobre la validez de un silogismo. Algunas de estas habilidades, como el control de la comprensión, tienen una amplia gama de aplicaciones, en tanto que otras sólo resultan útiles en circunstancias muy específicas. Sin embargo, pese a todas estas habilidades adquiridas, la vida cotidiana nos presenta permanentemente nuevos problemas para los que no tenemos una solución preparada. Esto nos lleva a la pregunta sobre si es posible (o no) aprender un conjunto muy general de habilidades, o principios, para resolver todo tipo de problemas. En otras palabras, además de adquirir habilidad en el uso de procedimientos de solución para problemas específicos, como en el ejemplo de los físicos planteado en el capítulo 13, ¿es posible adquirir habilidad para resolver problemas en general? El resto de este capítulo se ocupará de este tema. Analizaremos los procesos (como el de representación) involucra-

dos en la resolución de problemas y luego ofreceremos una evaluación de los intentos de mejorar la eficacia con la que se ejecutan.

## La representación de problemas

Ya hemos visto, en la selección de tareas de Wason (capítulo 13), que se puede encarar el mismo problema lógico con mayor o menor facilidad, según los contenidos que lo informan. Este es un ejemplo de la influencia de la representación. ¿Pero de qué otro modo, además del contenido, pueden diferir las representaciones? Bueno, un camino es el modo de representación.

### Las representaciones verbales y las imágenes visuales

Un buen ejemplo de diferencia en la forma de representación de los problemas es la existente entre la codificación verbal y la codificación en imágenes. En la práctica, ésta es una cuestión de grado ya que una formulación verbal puede provocar al menos algunas imágenes visuales y una imagen probablemente esté asociada con determinadas palabras. De todas maneras, un problema puede representarse predominantemente en palabras o predominantemente en imágenes, y la forma elegida puede tener consecuencias importantes. El problema del monje budista de Duncker, planteado al comienzo de este capítulo, está presentado verbalmente y de manera tal que estimula los intentos de resolverlo de manera matemático-verbal. Pero esta forma de representar el problema no resulta nada conveniente. Por el contrario, la solución se hace sencilla cuando representamos el problema en imágenes. Para hacerlo, imaginen un monje trepando una montaña durante un día. Imaginen también un segundo monje, descendiendo por la misma montaña durante un día. Superpongan ahora las dos imágenes, como si los dos viajes se hicieran el mismo día. Evidentemente, en algún momento del día los dos monjes deben cruzarse, y por lo tanto estarán en el mismo lugar al mismo tiempo. Esto es igual para el caso de un monje solo que haga los dos viajes en días separados.

En el problema de los monjes, la representación visual es superior a la verbal. Pero esto no siempre es así. El problema del campeonato de tenis que plantea De Bono es mucho más difícil si uno empieza con una hoja con el diagrama de un torneo por eliminación: una forma de pirámide caída hacia

un costado. Esto nos lleva a empezar agregando un partido final a dos semifinales, a cuatro cuartos de final, y así hasta llegar a la primera ronda, donde alguno no tendrá que jugar, ya que 111 no es múltiplo de 2, como 64 o 128. Si uno representa el problema de manera verbal la solución es mucho más sencilla: el número de participantes es 111; todos ellos perderán una vez, excepto el ganador del torneo; por lo tanto, habrá 110 partidos y ésa es la cantidad de juegos de pelotas que se necesitan.

Estos dos ejemplos demuestran la importancia de probar las diferentes representaciones de un problema y evaluar su potencial relativo para ofrecer una solución. Desafortunadamente, varios factores se unen para hacer que la representación flexible sea difícil de lograr en la práctica; nos ocuparemos ahora del análisis de algunos de esos factores.

## Configuraciones mentales

Las experiencias pasadas influyen sobre nuestra conducta presente de muchas maneras, y si no tuviéramos la capacidad de aprender de ellas estaríamos realmente muy mal. De todos modos, los viejos hábitos no siempre se adecuan a las exigencias de las nuevas situaciones. Podemos quedarnos demasiado inmóviles en algunos sentidos. Thomas Kuhn, el notable historiador de la ciencia, sostiene que los hábitos de pensamiento son un importante factor en la determinación de los compromisos teóricos en la ciencia. Cuando se propone una nueva teoría, que desafía a otra establecida, raramente hay evidencias concluyentes para decidir entre ellas. La vieja teoría gradualmente se va abandonando porque la generación que creía en ella (y había aprendido a pensar según ella) desaparece. Los científicos más viejos encuentran a las teorías novedosas, que introducen conceptos no conocidos y postulan la existencia de nuevas entidades, difíciles de aceptar. Como han aprendido a construir el mundo según una teoría previa, no hacen la transformación hacia una perspectiva moderna de inmediato. Kuhn (1970) ilustra el tema con un relato sobre el "descubrimiento" del oxígeno. En 1774, Priestley aisló una muestra de lo que ahora llamamos "oxígeno" pero que él pensó que era óxido nitroso o "aire deflogisticado". En 1777, Lavoisier propuso que el gas era uno de los dos constituyentes mayores del aire y desarrolló la base de la concepción moderna del elemento oxígeno. Priestley jamás pudo aceptar la teoría de Lavoisier y siguió conceptualizando al nuevo gas según la antigua teoría "flogística", con la que estaba familiarizado. Así, si bien Priestley fue el pri-

mero de los dos en aislar el gas ahora denominado oxígeno, no puede considerarse su descubridor. En realidad, la idea de "descubrimiento" es un concepto bastante difícil.

El ejemplo de Kuhn se refiere a los efectos de largo plazo de la experiencia, con hábitos mentales acumulados en años de aprendizaje, pero la configuración psíquica también puede solidificarse en el corto plazo. Esto queda demostrado con el famoso experimento de Luchins y Luchins (1950) con las jarras de agua, que se ilustra en la tabla 14.1. Dadas tres jarras de diversas medidas y una cantidad de agua ilimitada, hay que obtener cantidades específicas de agua. Se presentan varios problemas: el primero es bastante fácil. Los problemas 2 a 6 pueden resolverse de la misma manera (b - a - 2c). El problema 7 puede resolverse de la misma manera, pero tiene también una solución más fácil. Luchins y Luchins descubrieron que la gente que había realizado los problemas del 2 al 6 tendía a no ver la solución más sencilla del problema 7. Como se había creado una configuración mental solidificada, el problema 7 se representaba como si fuera uno más de una serie de problemas idénticos.

Las configuraciones mentales de corto plazo son fáciles de romper si uno toma conciencia sobre ellas. Pero las configuraciones mentales de largo plazo pueden resultar totalmente inevitables, o hay que estar constantemente cuidándose de ellas. Las configuraciones de largo plazo son parte del estilo de pensamiento de una persona.

*Tabla 14.1 Ilustración de los problemas de las jarras de Luchins*

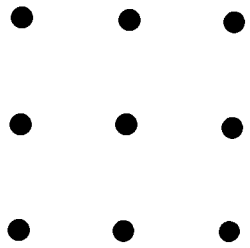
| Problemas | Capacidad jarra A | Capacidad jarra B | Capacidad jarra C | Cantidad deseada |
|---|---|---|---|---|
| 1 | 10 | 40 | 18 | 28 |
| 2 | 21 | 127 | 3 | 100 |
| 3 | 14 | 163 | 25 | 99 |
| 4 | 9 | 42 | 6 | 21 |
| 5 | 20 | 59 | 4 | 31 |
| 6 | 18 | 43 | 10 | 5 |
| 7 | 15 | 39 | 3 | 8 |

## Inmutabilidad funcional

No sólo es importante la representación general de un problema. La representación de sus elementos individuales también puede influir en la probabilidad de éxito en su resolución. Tendemos a representarnos los objetos según sus funciones, y si bien esto suele ser correcto, a veces puede ser excesivamente limitativo. La inmutabilidad funcional, como se denomina a este hábito, fue investigada por Maier (1931), quien planteó un problema en el que había que atar dos cuerdas que colgaban del techo de una habitación. Las cuerdas estaban tan separadas que no era posible sostener una y estirarse, al mismo tiempo, para alcanzar la otra. En el suelo de la habitación había varios objetos, incluyendo unas pinzas. Una solución posible consistía en atar las pinzas a la primera cuerda, hacerla balancearse, tomar la segunda cuerda y estirarse para alcanzar la primera cuando se acercara por el balanceo. Maier descubrió que la gente tendía a pensar en las pinzas sólo según su función usual. A menos que se hicieran sugerencias al respecto, era difícil percibir su potencial como plomada de un péndulo.

La historia de la tecnología es una abundante fuente de ejemplos de inmutabilidad funcional y de repentinos escapes de esa inmutabilidad. La innovación técnica con frecuencia consiste en la apreciación de cómo se puede adaptar una máquina ya existente a la resolución de problemas muy ajenos a sus actuales posibilidades de aplicación. Weizenbaum (1984) recuerda que la máquina de vapor había sido usada durante cien años para bombear agua de las minas antes de que Trevithik percibiera por primera vez su potencial como fuente de potencia de locomoción. Trevithik puso una máquina de vapor en un carruaje, que a su vez colocó sobre unas vías ya existentes, y ofreció así una novedosa solución al perenne problema del transporte. Weizenbaum señala que una inmutabilidad funcional similar dominaba también el uso inicial de las computadoras, que a menudo se consideraban meras máquinas de calcular o tabular con rapidez. Pasó bastante tiempo hasta que se reconoció su versatilidad como manipuladoras de símbolos en general (ver capítulo 16).

*Figura 14.1. El problema de los nueve puntos*



## Contrabando de supuestos

Tanto las configuraciones psíquicas solidificadas como la inmutabilidad funcional son casos especiales de un factor general que puede inhibir la flexibilidad en la representación de los problemas. Este factor es el contrabando de supuestos. En el capítulo 13 vimos que todo el pensamiento se basa en supuestos implícitos e inferencias, y que éstas no siempre son correctas. Un conocido ejemplo de contrabando de supuestos se da en el problema de los 9 puntos, en el que hay que unir los nueve puntos de la figura 14.1 usando cuatro líneas rectas o menos, y sin levantar el lápiz del papel. Adams (1979) analiza el problema detalladamente y ofrece varias soluciones posibles. A la mayoría de la gente le resulta difícil hallar alguna solución; Adams dice que esto se debe a que contrabandean en sus representaciones del problema varios supuestos limitativos. Por ejemplo, la mayoría de la gente supone que las líneas no pueden extenderse fuera del cuadrado marcado por los puntos. Esto no les permite ver la posibilidad de la solución mostrada en la figura 14.2. Mucha gente también elimina la posibilidad de dibujar una sola línea muy gruesa que abarque todos los puntos de un trazo.

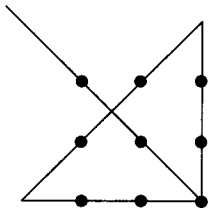
Otro ejemplo de dificultad provocada por contrabando de supuestos se da en el problema de los seis fósforos (Catona, 1940), en el que hay que formar cuatro triángulos equiláteros idénticos, todos con lados de un fósforo de largo, utilizando sólo seis fósforos. La gente tiende a suponer que el problema exige una solución en dos dimensiones. Esto no les permite ver la solución de contruir una pirámide triangular.

En los dos ejemplos que acabamos de comentar, el contrabando de supuestos nos impone una limitación innecesaria a lo que Newell y Simon (1972) denominaron "dimensión del problema del problema", metáfora lite-

ral satisfactoria para ambos casos. Lo que ya debería estar claro es que la forma en que se representa un problema influye notoriamente en la probabilidad de solución. Es necesario tener en cuenta el modo de representación y el omnipresente peligro de configuraciones psíquicas e inmutabilidad funcional. A veces, como sucede en el ejemplo del monje budista, una representación correcta del problema es equivalente a su solución. Pero, con más frecuencia, la representación es sólo la primera etapa y a ella le debe seguir la búsqueda de una solución. Ahora nos ocuparemos de la etapa de solución en la resolución de problemas.

---

*Figura 14.2. Una solución elegante al problema de los 9 puntos*



## El uso de heurísticas

Para algunos problemas existen procedimientos de solución conocidos que, si se ejecutan bien, garantizan respuestas correctas. Esos procedimientos se conocen con el nombre de *algoritmos* (ver capítulo 16). Pero la mayoría de los problemas no se puede encarar algorítmicamente o, si existe un algoritmo que ayude a resolverlos, puede ser extenso o difícil de operar. En esos casos solemos avanzar a tientas hacia la solución, recurriendo a varias reglas empíricas denominadas *heurísticas*. Las heurísticas son guías para la acción que se han adquirido con la experiencia. Sus ventajas incluyen la simplicidad y la familiaridad y su desventaja es que nunca podemos saber por anticipado si nos llevará en la dirección correcta.

## Análisis de medios-fines

John Sloboda proporciona un ejemplo de análisis de medios-fines en el capítulo 2 de este mismo libro, cuando introduce la idea de una "pila" de objetivos. El problema consistía en llegar del campo a Londres. Un análisis de medios-fines nos lleva a la pregunta: ¿cuál es la diferencia entre su situación presente y la del objetivo deseado? Es una cuestión de distancia. ¿Qué reduce la distancia? Un tren. ¿Dónde hay un tren? En la estación. ¿Cómo se llega a la estación? Busque un mapa. O pregúntele el camino a alguien del lugar. (Y ojalá que las convenciones discursivas locales se ajusten a las suyas. Ver capítulo 13.) Y así sucesivamente.

En este tipo de análisis, quien resuelve el problema oscila entre varios fines (objetivos y subobjetivos) por un lado y diferentes medios de lograrlos por el otro. La táctica predominante consiste en trabajar hacia atrás, desde la solución final, y eliminar las diferencias percibidas entre la situación presente y algún nivel de la situación del objetivo, repitiendo el proceso en cada nivel.

El análisis de medios-fines constituye la base del programa de computación tan ampliamente aplicable, el *General Problem Solver* (Newell y Simon, 1972). Es un método que básicamente utiliza el sentido común, pero su aplicación sistemática es un poderoso medio para resolver problemas, desde los más simples hasta aquellos relativos al ajedrez de alto nivel.

El uso del análisis de medios-fines fue explícito en los registros de "pensamiento en voz alta" de los sujetos del experimento de Duncker (1975), con el "problema de la radiación". Se refiere a un paciente con un tumor de estómago inoperable, y los rayos destruyen tejidos a la intensidad suficiente. El problema consiste en llegar a un procedimiento por el cual se pueda destruir el tumor sin dañar los tejidos sanos que lo rodean, como sucedería con el sencillo enfoque de dirigir un haz de rayos directamente hacia el tumor. Los sujetos tendían a trabajar hacia atrás, empezando por el objetivo general y siguiendo por subobjetivos, como irradiar el tumor mientras evitaban el tejido sano. Una solución posible era fijarse el subobjetivo de reducir la intensidad de los rayos mientras pasaban por el tejido sano. Este subobjetivo llevó a las ideas de enfocar, o hacer converger, rayos débiles en el lugar del tumor.

El análisis de medios-fines incluye una serie de procesos componentes como la reducción del problema, el trabajo hacia atrás y la evaluación de las semejanzas entre la situación presente y la del objetivo. Cada uno de estos puntos merecería un capítulo entero (ver Gilhooly, 1982).

## Disponibilidad

Al abrir juicio respecto de un problema tenemos que basarnos en la información que tenemos a nuestra disposición. Pero esa información puede no retratar con precisión la situación. Esto suele darse en casos en que nuestra exposición a la información está determinada por criterios que resultan irrelevantes para el problema que nos ocupa. Por ejemplo, si uno quiere decidir cuál es el medio de transporte de larga distancia más seguro, la información que proporcione la prensa no resultará útil. La cobertura que realizan los medios depende fundamentalmente del criterio de "noticia": destaca los desastres y no la seguridad.

Tversky y Kahneman (1973) hicieron una ilustración experimental de la influencia de la disponibilidad de la información. Le preguntaron a angloparlantes si las letras K, L, N, R y V aparecían, en inglés, con más frecuencia en la primera o la tercera posición en las palabras del idioma. Los resultados dieron que la gente pensaba que las cinco letras aparecían más en la primera posición, aunque en realidad todas son más frecuentes en la tercera. Presumiblemente, la gente comete este error porque es mucho más fácil pensar en palabras que empiecen con una letra en particular que en otras que lleven la misma letra en tercer lugar, y no se tienen en cuenta los efectos distorsionantes de esta desviación.

La disponibilidad es una heurística útil en la resolución de problemas pero sólo cuando la frecuencia objetiva de un tipo de acontecimientos se correlaciona con la disponibilidad subjetiva de ejemplos de esa clase. Pero esto no suele suceder. Además de la exposición irregular de la información en los medios de prensa, todos tendemos a buscar la información que confirma nuestros propios prejuicios. Flavell (1985) sugiere que la razón por la que es tan difícil no ver las cosas desde nuestro propio punto de vista es que nuestro propio punto de vista está inmediatamente disponible. Las interpretaciones que más practicamos son las nuestras propias, y se han automatizado. Es necesario mucho esfuerzo consciente para adquirir eficiencia en adoptar el punto de vista de otro.

## Representatividad

Lea la siguiente descripción:

> Linda tiene 31 años, es soltera, sincera y brillante. Se graduó en filosofía. Cuando era estudiante, le preocupaban los problemas de la discriminación y la justicia social, y participó en demostraciones antinucleares.

Elija ahora cuál de las siguientes alternativas le resulta más factible:

1. Linda es cajera de banco.
2. Linda es cajera de banco y militante feminista.

Tversky y Kahneman (1982) hallaron que el 87% de los entrevistados elegía la segunda alternativa como más factible que la primera. Sin embargo, un momento de reflexión nos señalará que esto no puede ser así. La probabilidad de conjunción de dos acontecimientos (o características) no puede ser mayor que la probabilidad de uno de ellos por sí mismo. Por ejemplo, si es improbable que usted gane un premio Nobel y también es improbable que gane una medalla de oro olímpica, difícilmente sea más probable que usted logre ambas cosas en vez de una sola de ellas. Sin embargo, incluso aquellos entrenados en estadísticas y teoría de las probabilidades cometen el error de considerar a la segunda alternativa como la más probable. Parece que la gente responde según lo que Tversky y Kahneman llaman "representatividad". La descripción de Linda se construyó para que fuera representativa de una feminista activa pero no de una cajera de banco, y la gente parece basar su juicio en esta representatividad, o semejanza, más que en las probabilidades objetivas.

En los problemas para los que hay poca información disponible sobre las probabilidades objetivas, lo cual es muy frecuente, la representatividad puede ser un recurso heurístico razonable. Pero en ciertas circunstancias puede desencaminar marcadamente nuestros pensamientos. Por ejemplo, unos pocos casos bien publicitados de abusos en el sistema de bienestar social pueden, si se supone que son representativos, desviar las evaluaciones del nivel de abuso entre la población total de beneficiarios del bienestar social.

## Anclaje

Tversky y Kahneman (1974) analizan otra fuente de error en el pensamiento humano: la influencia que puede tener un punto de partida en el desempeño en las tareas de evaluación. Esto quedó claramente demostrado al hacer que dos grupos evaluaran, en 5 segundos, el resultado de un cálculo numérico escrito en la pizarra. Un grupo vio el siguiente cálculo: 8 x 7 x 6 x 5 x 4 x 3 x 2 x 1, y dio una estimación media de 2.250. El otro grupo vio: 1 x 2 x 3 x 4 x 5 x 6 x 7 x 8, y dio un resultado promedio de 512. La respuesta correcta es 40.320, por lo que ambos grupos estuvieron muy por debajo de lo correcto, pero la subestimación fue mucho mayor en aquellos que, presumiblemente, tuvieron tiempo de multiplicar algunas de las cifras más pequeñas que en aquellos que pudieron multiplicar las mayores. El punto de partida actuó como anclaje.

Fiske y Taylor (1984) señalan que los efectos del anclaje pueden darse siempre que alguien tenga que evaluar la frecuencia de aparición de alguna característica en la población en general o la probabilidad de alguna forma específica de conducta. En la mayoría de los casos, la gente comienza sus deliberaciones de acuerdo con su propia actitud ante el tema, y esto les ofrece un punto de anclaje. Es probable que una persona muy honesta sobreestime el peso de la honestidad en la población ("¿Cómo podrían...?") y que una persona deshonesta sobreestime el peso de la deshonestidad ("Te apuesto a que tú, en realidad..."). De manera similar, aunque podemos dudar de lo que dice un publicista a favor de un producto, la mera exposición a avisos favorables bien puede proporcionar el anclaje a partir del cual haremos nuestra evaluación real. La consecuencia puede ser una evaluación ventajosa para el fabricante del producto.

# El entrenamiento en las habilidades de resolución de problemas

En las últimas dos secciones analizamos varios procesos relacionados con la representación de problemas y la búsqueda y ejecución de procedimientos de solución. En el curso normal de los acontecimientos, estos importantes procesos operan, fundamentalmente, de manera inconsciente. La conducta común sería imposible si no detectáramos automáticamente las seme-

janzas entre el pasado y el presente, si no simplificáramos los problemas complejos y extrapoláramos a partir de datos incorrectos. Nada de lo que se ha dicho debe tomarse como crítica de la utilidad de estos hábitos cognitivos. Sólo cuando se encuentran dificultades, cuando la conducta resulta ineficaz o los problemas resultan extraños para los enfoques habituales, nuestros procedimientos irreflexivos deben convertirse en objeto de estudio. De alguna manera hay que llevarlos a la conciencia.

## Hábitos y conciencia

En los últimos años se han hecho esfuerzos combinados, especialmente en los Estados Unidos, para desarrollar planes de entrenamiento para las habilidades de resolución de problemas generales (Hunt, 1982; Mayer, 1983). En todos estos programas hay un énfasis común en el "conocimiento consciente aumentado".

Como la vida cotidiana está llena de problemas que todos debemos enfrentar, es evidente que debemos tener mucha experiencia en la resolución de problemas. La experiencia es indudablemente buena pero puede llevar a la formación de malos hábitos y ciertamente da como resultado diversas formas de falta de atención a los procesos de resolución de problemas. El entrenamiento apunta a hacer conscientes esos procesos, a que la gente conozca sus propios hábitos de pensamiento y esto le permita ejercer mayor control sobre sus actividades de resolución de problemas. Como en los grupos de encuentro y otras formas de entrenamiento de la sensibilidad, los orígenes de este enfoque derivan de la psicología de la gestalt, surgida en la Alemania de comienzos de siglo (capítulo 11).

Mientras que la tradición conductista siempre ha subrayado la importancia del aprendizaje mediante ensayo y error para la resolución de problemas, que hace que las soluciones exitosas se vuelvan habituales, los psicólogos gestálticos destacan fundamentalmente el papel de lo que denominan *insight*. Se refiere a la comprensión de las relaciones involucradas en un problema. Los psicólogos gestálticos sostienen que la forma en que se percibe (o configura) un problema es de vital importancia para el diseño de una solución y que las soluciones suelen estar precedidas de *insights* en los que repentinamente se ve al problema según una nueva organización. Ejemplo de ello sería el problema de los 6 fósforos antes mencionado. La comprensión de que la solución puede estar en tres dimensiones (la pirámide triangular) puede ex-

perimentarse como un *insight* en la naturaleza del problema. La noción de *insight* y los procesos cognitivos de la representación de los problemas se consideran análogos a los de la percepción visual. La reorganización abrupta del problema se asemeja a la reestructuración perceptual de una figura ambigua, como cuando se percibe la figura 5.2 del capítulo 5 como dos paralelogramos en vez de como un par de soportes para huevo duro. El entrenamiento en resolución de problemas que deriva de los principios de la gestalt se centra en hacer que la gente tome conciencia de la importancia de la representación, de la necesidad de probar diferentes formas de conceptualizar un problema y de las restricciones que actúan contra la reconceptualización. Si se aumenta la conciencia en los factores involucrados en la representación de problemas, es esperable que su pensamiento se torne más flexible.

La pregunta es, entonces, ¿sirve el entrenamiento?

## Evaluación del entrenamiento

Suele afirmarse que las habilidades de resolución de problemas generales pueden mejorarse mediante el entrenamiento. Se escriben libros sobre el tema y se organizan cursos universitarios y comerciales por los que los participantes (o sus compañías auspiciantes) pagan grandes sumas. Si uno debiera juzgar exclusivamente a partir de los informes subjetivos sobre si la gente siente que se ha beneficiado con un curso en particular, tendría que llegar a la conclusión de que al menos algunas personas mejoran su habilidad en la resolución de problemas como consecuencia del entrenamiento. Hay un considerable grado de satisfacción entre los consumidores.

Pero los informes subjetivos no son necesariamente confiables. Mayer (1983) analiza varios intentos de evaluación objetiva de los cursos de entrenamiento y llega a una conclusión menos optimista. Sugiere que aunque se puede demostrar que el entrenamiento mejora la capacidad de resolución de problemas en áreas específicas tales como las matemáticas y la ingeniería, las evidencias sobre la adquisición de habilidades generales que se transfieren a otros tipos de temas siguen siendo dudosas. Los estudios menores (en los que tanto entrenadores como estudiantes pueden ser conscientes de que toman parte en un estudio de evaluación y en los que la motivación puede ser consecuentemente alta) suelen tener resultados positivos. Los estudios a gran escala, por otra parte, frecuentemente muestran que los efectos del entrenamiento son escasos (cuando tienen algún efecto).

La demostración concluyente de resultados positivos es una dificultad común a todos los intentos de entrenamiento en las habilidades cognitivas de nivel superior (ver capítulos 8, 11, 12 y 15). Es indiscutiblemente difícil encontrar evidencias firmes del éxito en el entrenamiento en habilidades de resolución de problemas generales, pero parte de la dificultad puede radicar en la naturaleza de la propia evaluación. ¿Las pruebas de laboratorio son los mejores indicadores de la mejora en una habilidad? Es posible que la gente se esfuerce más y use mejor su entrenamiento en los ámbitos naturales. Pero, quizá, el obstáculo principal para una buena evaluación del entrenamiento en estas habilidades se relacione con la práctica. Se ha dicho reiteradamente que el desarrollo de la habilidad exige mucha práctica. Naturalmente, más de la que se puede proporcionar durante un curso de resolución de problemas que dura apenas unas semanas. Se pueden enseñar técnicas y enfoques sistemáticos pero la práctica depende finalmente de cada individuo. Puede ser por esto que sólo en los estudios de evaluación de menor escala, con sus niveles de motivación concomitantemente más altos, se observen resultados positivos constantes.

De todos modos, el tema del entrenamiento para la resolución de problemas volverá a plantearse en el próximo capítulo, esta vez en el contexto del entrenamiento en creatividad.

## Lecturas adicionales

MAYER, R.E. (1983) *Thinking, Problem Solving,* Cognition. Nueva York: Freeman.

Revisión completa y amena de la literatura sobre resolución deproblemas.

ADAMS, J.L. (1979) *Conceptual Blockbusting* (2ª edición). Nueva York: Norton.

Una visión del tema entretenida y notécnica.

# 15

# La creatividad y las habilidades cognitivas

*Alan Branthwaite*

En los capítulos 13 y 14 se sostiene que los procedimientos relativos a la resolución de problemas pueden volverse rutinarios dentro de campos específicos y que se pueden aprender ciertas habilidades y enfoques para ayudar a la resolución de una amplia variedad de problemas. Pero las ideas brillantes no aparecen por un simple pedido. Algunos problemas son insolubles durante muchos años y algunas ideas son tan novedosas que quienes las tuvieron sólo alcanzan el reconocimiento póstumo. Estos son casos de pensamiento creativo; algunos de ellos se describirán en este capítulo, mientras exploramos el papel que pueden jugar las habilidades cognitivas en la creatividad.

## La naturaleza del pensamiento creativo

No siempre resulta fácil reconocer las ideas creativas, ya que no hay un estándar claro para medirlas. La creatividad, como el genio, es esquiva, y sólo el consenso social decidirá si una idea se considera creativa e innovadora. Esto es especialmente así en el campo del arte, pero también en la ciencia y la tecnología, donde algunas ideas o descubrimientos han estado dormidos durante generaciones porque su valor creativo no era reconocido en su tiempo (por ejemplo, las leyes de la herencia de Mendel).

La historia ofrece muchos ejemplos célebres de soluciones a problemas difíciles e importantes que ganaron reconocimiento y aceptación públicos como innovaciones creativas. Al observar algunos ejemplos, prestaremos especial atención a la forma en que aparecieron las soluciones para ver qué luz arroja esto sobre el proceso creativo.

### Arquímedes

Es famosa la exclamación de Arquímedes cuando encontró la solución a su problema: "¡Eureka!" La historia fue así: Arquímedes tenía que averiguar

si parte del oro de una corona que habían fabricado unos artesanos había sido sustituido por plata. La corona pesaba lo mismo que el oro que les habían dado a los artesanos para hacerla. Arquímedes razonó que, como la plata tiene más volumen que el oro para el mismo peso, una corona adulterada sería más grande que una de oro puro. El problema de Arquímedes consistía en medir el volumen de la corona, que tenía una forma irregular. Como todos sabemos, la solución se le ocurrió en un contexto diferente, ya que casualmente Arquímedes observó, mientras visitaba los baños públicos, que cuanto más se sumergía más agua desbordaba. La creatividad de Arquímedes consistió en aplicar las implicaciones de esta observación común al problema de medir el volumen de la corona. De repente, descubrió la solución: sumergiendo la corona en el agua pudo medir su volumen por la cantidad de agua que desplazaba. La leyenda termina con Arquímedes corriendo desnudo por las calles de Siracusa mientras gritaba, excitado, "¡eureka!", que significa "lo encontré". Por coincidencia, una actividad no relacionada le proporcionó la pista para resolver el problema que lo preocupaba, aunque probablemente no estaba pensando en él cuando se le ocurrió la solución. Ah... parte del oro, efectivamente, había sido reemplazado por plata.

## Darwin

Darwin tenía libretas en las que registraba sus ideas y pensamientos; estas libretas ofrecen un informe más preciso y detallado sobre su creatividad que la leyenda de Arquímedes. Las libretas de Darwin muestran cómo desarrolló su explicación para el proceso de la evolución (Gruber, 1974). Le llevó 15 meses, durante los cuales probó diversas explicaciones que resultaron insatisfactorias y, por lo tanto, tuvo que liberar a su mente de ellas. Lentamente, Darwin fue acumulando un conjunto de ideas que le parecían relevantes y en torno a las cuales fue trabajosamente amasando un considerable cuerpo de evidencia de apoyo. Un avance especialmente importante fue el de la idea de la selección natural. En su libreta, Darwin escribió: "En octubre de 1838, es decir, quince meses después de empezar mi investigación sistemática, leí casualmente, sólo por diversión, *Malthus y la población*, y al estar bien preparado para apreciar la lucha por la existencia que se da en todas partes, a partir de mi larga e ininterrumpida observación de los hábitos de plantas y animales, *de inmediato se me ocurrió* que en estas circunstancias las variaciones favorables tenderían a preservarse y las no favorables a destruirse." (Dar-

win, 1911, p. 68; las bastardillas son nuestras.) Darwin se sintió inmediatamente impresionado por esta nueva idea pero siguió trabajando en otros pensamientos, y quizá no se haya dado cuenta por entonces de que este proceso de selección natural iba a convertirse en la piedra fundamental de su explicación de la evolución. Habría que recordar, al examinar otros relatos sobre creatividad, que es muy fácil destacar un punto en particular como el momento de inspiración después de los hechos, cuando toda la solución ha quedado ya bien clara, sin tener en cuenta todo el trabajo realizado antes y después. De hecho, Darwin pasó otros 20 años analizando la teoría y buscando evidencias que la sustentaran.

## Kekule

El "descubrimiento" de que las moléculas de benceno tienen átomos ordenados en un anillo cerrado en vez de tenerlos en forma de cadena, se le ocurrió a Kekule, según su propio relato, del siguiente modo (citado por Koestler, 1964): "Volví mi sillón hacia el fuego y me eché a dormitar. Una vez más, los átomos bailaban ante mis ojos. Esta vez, los grupos más pequeños se mantenían humildemente atrás. Mi ojo mental, agudizado por reiteradas visiones de este tipo, podía ahora distinguir estructuras mayores, de conformación múltiple; largas filas, a veces más estrechamente unidas, que se retorcían en un movimiento de serpiente. Pero de pronto... ¿Qué fue eso? Una de las serpientes se atrapó su propia cola y la forma remolineó burlona ante mis ojos. Como sobresaltado por un relámpago, desperté... Aprendamos a soñar, caballeros."

## Babbage

Otro ejemplo similar de creatividad surgida en parte gracias a la lógica sistemática y en parte a debido a los sueños imaginativos fue la invención de la Máquina Analítica de Babbage, precursora de las computadoras modernas (citado por Cohen, 1970, p. 65): "La primera idea que puedo rastrear en mi propia memoria acerca de realizar cálculos aritméticos mediante máquinas surgió de esta forma: una tarde yo estaba sentado en los salones de la Sociedad Analítica de Cambridge, con mi cabeza inclinada hacia adelante sobre la mesa en una especie de ensoñación y con una tabla de logaritmos

abierta ante mis ojos. Otro miembro de la sociedad entró al salón, me vio medio dormido y dijo: '¿Qué estás soñando, Babbage?'. Y yo respondí: 'Estoy pensando que todas estas tablas (señalando los logaritmos) podrían calcularse mediante máquinas'".

Koestler, 1964, hace otras descripciones sobre la forma en que surgieron muchas ideas creativas trascendentales. Pero es importante comprender que la creatividad no se limita a los casos famosos que transformaron la comprensión del mundo. Estos ejemplos de creatividad son célebres por la seriedad de los problemas que resolvieron y la utilidad de las soluciones y no porque implicaran ideas especialmente imaginativas o inventivas. Pero la creatividad también se encuentra en la soluciones inventivas e imaginativas a los problemas cotidianos de un individuo o también en el niño que "descubre" por sí mismo el principio de Arquímedes, aunque es improbable que la historia registre tales logros. Después de todo, Darwin sólo hizo públicas sus ideas cuando otro (Alfred Wallace) concibió la misma idea independientemente, proporcionándole apoyo y aliento para que superara sus dudas y temores. Eso no hace menos creativa a la idea de Darwin o a la de Wallace, aunque la creatividad de este último es mucho menos conocida. Al analizar qué es la creatividad debemos tener en cuenta no sólo la importancia social de los resultados sino también la resistencia del problema a la solución y la dificultad que supera un individuo para alcanzar una resolución satisfactoria.

## El proceso creativo

¿Qué nos dicen estos relatos sobre algunas ideas creativas acerca de los procesos cognitivos involucrados? Los ejemplos dados ofrecen ciertos informes introspectivos en cuanto a la forma en que surgieron esas ideas. En otros capítulos hemos visto algunas de las limitaciones de los informes introspectivos como medio para aclarar los procesos psíquicos. Esto es especialmente cierto cuando el informe se escribe después de los hechos que describe (en lugar de hacerlo mientras suceden), ya que la memoria reconstructiva es selectiva y parcial (ver capítulo 5). Algunos investigadores (por ejemplo, Perkins, 1981) han combinado los informes biográficos sobre creatividad natural con observaciones sobre la creatividad en el laboratorio, cuando se pide a los individuos que vayan contando en voz alta cómo resuelven los problemas o componen poesías (procedimiento que tiene sus propias limitacio-

nes). Sobre la base de estas evidencias, parece haber cuatro procesos involucrados, que Wallas (1926) fue el primero en denominar *Preparación*, *Incubación*, *Iluminación* y *Verificación*. Esta descripción de la creatividad ha sido confirmada por autores posteriores, aunque ahora estos procesos no se consideran etapas discernibles sino parte de un interjuego dinámico que se da durante la creatividad.

## 1) Preparación

La preparación incluye la definición y exploración de las características de la tarea o problema, y es afín a la cuestión de la representación del mismo, analizada en el capítulo 14. Los procesos involucrados en la preparación se ilustran en algunos de los ejemplos de creatividad descritos: las deducciones de Arquímedes sobre el hecho de que la plata tiene más volumen que el mismo peso en oro; las observaciones de Darwin de las manifestaciones de la evolución y los componentes sobre los cuales podría operar el proceso, tales como la variabilidad dentro de cada especie y la adaptación al medio. Durante la preparación, se consideran las soluciones posibles y se las prueba lógicamente para ver si funcionan. En muchos problemas, con esto basta para llegar a una solución (ver capítulo 14). Sólo con los problemas más difíciles e intransigentes hace falta mayor creatividad. Sin embargo, este proceso preparatorio de exploración del problema es precursor esencial de las soluciones más creativas e imaginativas. En este proceso, surge una conciencia sobre las propiedades de una solución satisfactoria que es importante más tarde para poder reconocer la solución cuando aparece. La preparación también aumenta la relación con el problema, que probablemente resulte beneficiosa para aumentar la motivación y la atención a soluciones posibles cuando surgen. Pero este aumento del compromiso también puede resultar contraproducente, ya que provoca configuraciones psíquicas solidificadas y pensamientos rígidos sobre las mismas soluciones posibles que se repiten una y otra vez.

Algunos psicólogos han descrito el proceso de preparación como un sistema de pensamiento cerrado que se centra en las posibilidades lógicas y corre sistemáticamente por carriles familiares, usando enfoques que se han aprendido por la experiencia pasada en lo que parecen ser problemas similares. Pero, durante la preparación, los pensadores creativos "sondean" un problema, como si exploraran un rompecabezas para familiarizarse con la forma

de la pieza faltante, como en el caso de Kekule, que sabía muy bien cuáles eran los criterios que debía satisfacer la solución del problema del ordenamiento de los átomos en el anillo de benceno. De manera similar, durante mucho tiempo Darwin tuvo una clara noción de lo que habría que obtener mediante el mecanismo de la evolución que estaba buscando.

## 2) Incubación

La incubación se refiere a los procesos aparentemente inconscientes que tienen lugar cuando un problema no está siendo considerado de manera activa pero a partir de los cuales surge, de todos modos, una solución imaginativa y exitosa. Esto ocurrió claramente en los sueños diurnos de Kekule y Babbage. La esencia de la incubación es que durante la meditación serena (o a veces desasosegada) del problema se generan nuevas posibilidades y combinaciones de ideas. Éstas surgen del libre juego de la imaginación, de la intuición y del pensamiento abierto y fluido que trae nuevas perspectivas para encarar el problema. Bruner (1962) afirma que la intuición "produce interesantes combinaciones de ideas antes de que se sepa su validez. Precede a la prueba; de hecho, es lo que estudian y comprueban las técnicas de análisis y verificación" (pág. 102). Estas características son muy evidentes en la descripción que hace Poincaré de la forma en que estableció una prueba matemática después de un período de preparación:

> "Todos los días me sentaba ante mi mesa de trabajo, me quedaba una o dos horas, probaba varias combinaciones y no llegaba a ningún resultado. Una noche, contra mi costumbre, bebí café y no me podía dormir. Las ideas surgieron en grandes cantidades; las sentía chocar entre sí hasta que se formaban pares que, por así decirlo, constituían una combinación estable. A la mañana siguiente, ya había establecido la existencia de una clase de funciones fuchsianas, las que vienen de las series hipergeométricas; sólo tenía que escribir los resultados, lo cual no me llevó más que unas horas."
>
> (Ghiselin, 1952, pág. 37)

Durante la incubación, la mente parece volver a formas de operación más primitivas (como en los sueños) y recurrir a la fantasía, la asociación, las analogías y las metáforas. Piaget (1926) lo denominó pensamiento autista (en oposición al pensamiento inteligente lógico). No se adapta a la realidad sino

que crea por sí mismo un mundo de ensueño de la imaginación, que es estrictamente individual e incomunicable mediante el lenguaje y que funciona mediante imágenes y símbolos.

Perkins (1981) ha sugerido una visión alternativa de la incubación. Se pregunta si la incubación incluye un tipo especial de pensamiento inconsciente que genere nuevas alternativas como soluciones posibles a un problema. Subraya el papel de los procesos cognitivos más comunes: el olvido fructífero de las posibilidades apreciadas pero fracasadas; el refresco físico y psíquico; la observación de nuevas pistas en experiencias no relacionadas (como en el caso de Arquímedes), que exige un alto nivel de interés, alerta y atención no focalizada; el reconocimiento contrario, que incluye el reconocimiento de cosas no sólo por lo que son sino también por lo que podrían simbolizar (por ejemplo, ver animales en la forma de las papas). Sin embargo, el proceso de reconocimiento contrario presenta grandes semejanzas con el pensamiento autista y las ensoñaciones que se describen más arriba.

No veo razón alguna por la cual los procesos cognitivos comunes sugeridos por Perkins, así como las formas de pensamiento más primitivas, no puedan estar incluidos en la incubación (ver más abajo). La función principal de la incubación es aumentar la gama de soluciones posibles y la habilidad para generar ideas alternativas fácilmente y sin esfuerzo es característica de la gente creativa. Cohen (1970) observa que Norbert Weiner, el creador de la cibernética, "atesoraba en sí mismo lo que sentía que era una imaginación caleidoscópica de libre fluir, que por sí misma revelaba las posibilidades de una situación complicada" (pág. 57) y Picasso afirmaba, tajante: "Yo no busco, encuentro", con lo que quería decir que las ideas se le ocurrían sin necesidad de pensar en ellas deliberada o laboriosamente. Es posible que la práctica pueda llevar a un aumento en la habilidad de generar ideas; éste es un supuesto subyacente en las técnicas de entrenamiento de la creatividad que se verán más adelante. Pero, además, no debemos subestimar el papel de las ocurrencias casuales al estimular las soluciones alternativas y de la atención alerta ante tales hallazgos casuales.

## 3) Iluminación

La iluminación es el momento de *insight* o comprensión repentina de que una idea se ajusta a un problema y ofrece una solución. Queda ilustrada con claridad en la historia de Arquímedes, cuando comprendió la importan-

cia del agua desbordando en el baño, o en la experiencia de Darwin, cuando de pronto se le ocurrió la idea de la selección natural (aunque siguió trabajando en otras ideas mientras buscaba verificar esta 'iluminación'). Lo importante aquí es el "salto psíquico" por el cual algo aparentemente irrelevante se relaciona con las características del problema. Esto requiere alerta y vigilancia mientras no se presta atención consciente al problema, y se da si uno está profundamente comprometido con la tarea. Por este motivo es importante, en el proceso de preparación, una clara comprensión de la naturaleza y las características (la "forma") de una solución.

En la creación artística, la iluminación surge cuando una idea posible le da al artista la sensación de que se ha logrado una expresión más satisfactoria y valiosa, que enriquece el significado de su obra. Tiene una fuerte cualidad emocional, como la experimentada por Arquímedes, que acompaña a la convicción de que la nueva idea completará el problema. Con la iluminapensamiento lógico y sistemático para verificar si las soluciones potenciales se ajustan a las características del problema. Se trabaja y elabora la inspiración creativa. Este proceso queda claramente ejemplificado en el caso de Darwin, pero también en los de Babbage y Poincaré. Se aducen evidencias y argumentos a favor para poder comunicar la solución a los demás de una manera racional y coherente. Se utilizan los conocimientos y habilidades previamente adquiridos para comprobar la sensatez y coherencia de la idea iluminadora. Desafortunadamente, algunas de estas ideas que parecían satisfacer los requisitos del problema resultan erróneas en el análisis posterior, de modo que el proceso de verificación se convierte en apenas uno más de los procesos de preparación.

## Los procesos creativos y el funcionamiento de los dos hemisferios

Los cuatro procesos que acabamos de describir incluyen dos tipos contrastantes de habilidad psíquica:

a) el análisis consciente, racional, lógico y sistemático (importante en la preparación y la verificación).

b) la intuición consciente, las analogías simbólicas, las metáforas y las asociaciones que crean nuevas perspectivas y diferentes posibilidades de solución (en la incubación e iluminación).

La existencia de dos formas de pensamiento contrastantes ha sido reconocida en muchos aspectos de la psicología (Bogen, 1977). Por ejemplo:

| | | | |
|---|---|---|---|
| racional | versus | metafórico | (Bruner) |
| realista | versus | impulsivo | (Hilgard) |
| diferencial | versus | existencial | (William James) |
| positivo | versus | místico | (Lévi-Strauss) |
| racional | versus | intuitivo | (Maslow) |
| secuencial | versus | múltiple | (Neisser) |
| dirigido | versus | autista | (Piaget) |

Recientemente, se identificó una posible base para estos diferentes modos de pensamiento en las diferentes funciones de los dos hemisferios cerebrales. (Por un resumen de las evidencias, ver Trueman, 1985; Springer y Deutsch, 1985). Se sugiere que los dos hemisferios se especializan en diferentes formas de actividad psíquica.

| *Hemisferio izquierdo* | *Hemisferio derecho* |
|---|---|
| * pensamiento secuencial, analítico | * holístico, simultáneo, integrador |
| * verbal | * manipulación espacial, imagen y orientación corporal en el espacio |
| * lógico | * intuitivo |
| * procesamiento serial | * procesamiento en paralelo |
| * modo lineal de operación, temporal | * procesa la información difusamente |
| * intelectual (consciente) | * relacional, asociativo |
| * convergente | * divergente |

La evidencia de que ambos hemisferios se relacionan diferenciadamente con aspectos del proceso creativo no es concluyente pero es decididamente sugerente (Katz, 1978; Torrence, 1982). El hemisferio derecho parece estar involucrado en aquellos aspectos de la creatividad que dependen de la imaginación, la analogía, la flexibilidad cognitiva y el pensamiento divergente, característicos de la incubación y la iluminación. Naturalmente, resulta prácticamente demasiado simplista decir que el hemisferio derecho se especializa en la creatividad. Casi todo el pensamiento sofisticado involucra aspectos de la actividad de ambos hemisferios y la interacción entre ellos, y esto también sucede con la creatividad. En particular, hay un continuo interjuego entre la imaginación verbal y la no verbal en nuestro pensamiento. Sin embargo, esta comparación entre procesos que durante mucho tiempo se han considerado importantes en la creatividad y las formas especializadas de trabajo en los dos hemisferios ayuda a clarificar la naturaleza de los procesos involucrados. Amplía nuestra comprensión de los aspectos misteriosos, primitivos y no verbales de la incubación y la iluminación. Esta explicación sitúa la descripción de la creatividad en un contexto más amplio del actualmente conocido sobre las funciones cognitivas y explica por qué la gente creativa suele sentir la solución antes de poder articularla. Además, sugiere que se podría mejorar la creatividad aprovechando las habilidades de cada hemisferio en las diferentes etapas del proceso creativo. Pero esta idea no es nueva, como puede verse en la respuesta de Schiller a un amigo que se quejaba de su falta de creatividad:

> "Creo que el motivo de tu queja radica en las restricciones que tu intelecto impone a tu imaginación... Aparentemente, no es bueno (y, de hecho, obstaculiza la creatividad de tu mente) que el intelecto examine de muy cerca las ideas que ya están llegando en forma abundante, por así decirlo, a la salida. Considerada aisladamente, una idea puede ser insignificante y en extremo aventurada, pero puede adquirir importancia por una idea que le sigue; quizá, en determinada ubicación junto a otras ideas, que pueden parecer igualmente absurdas, pueda ser capaz de ofrecer una vinculación beneficiosa... En el caso de una mente creativa, me parece, el intelecto ha retirado a sus vigías de las puertas y las ideas fluyen tumultuosamente, y recién entonces el intelecto se ocupa de revisar e inspeccionar a la multitud. Ustedes, los críticos de valía, o como quiera que se llamen a sí mismos, se avergüenzan o le temen a la locura momentánea y pasajera que se encuentra en todos los creadores... De ahí sus quejas por la falta de frutos, ya que rechazan demasiado pronto y discriminan con excesiva severidad."
>
> (Brill, 1938, pág. 193)

El argumento de Schiller recibe nueva fuerza y apoyo cuando se lo ve de acuerdo con la operación de los dos modos de la conciencia en los hemisferios cerebrales.

El conocimiento de la existencia de dos tipos complementarios de habilidad psíquica involucrados en la creatividad sugiere dos estrategias amplias para mejorarla. Cuando revisamos los diversos métodos que se han planteado para aumentar la creatividad, encontramos que entran en dos categorías amplias que buscan mejorar lo que podemos llamar, por conveniencia, modos de pensamiento del hemisferio izquierdo y derecho. El desarrollo de habilidades creativas involucra la búsqueda de mejores estrategias para producir soluciones posibles; en el resto de este capítulo analizaremos algunas de las técnicas que se han sugerido. Estas técnicas comparten dos propiedades con muchas de las habilidades cognitivas que examinamos en el resto del libro. En primer lugar, son heurísticas (ver capítulo 16), ya que ninguna de ellas garantiza la solución de un problema en particular pero muchas pueden ayudar en la búsqueda estimulando ideas. En segundo lugar, su aplicación mejora con la práctica (ver capítulo 3). Al principio, muchas de las técnicas parecerán extrañas y difíciles de llevar a cabo. Sólo se convierten en realmente efectivas tras una aplicación persistente. No ofrecerán creatividad instantánea al primer intento.

## El desarrollo de las habilidades del "hemisferio izquierdo"

Estas técnicas fomentan un análisis más exhaustivo, sistemático, lógico y racional del problema. Algunas de las tácticas usadas para lograrlo son:

- reformule el problema en sus propias palabras;
- formule preguntas sobre el problema;
- aclare mentalmente todos los datos del problema;
- haga anotaciones;
- haga una lista incluyendo todos los objetos/aspectos importantes del problema y piense cuidadosamente en cada uno de ellos;
- busque las cosas raras o confusas de un problema y concéntrese en intentar explicarlas;
- cuando haya varias cosas confusas en un problema, intente explicarlas con una sola idea que las conecte a todas.

Algunas técnicas buscan crear nuevas ideas identificando las omisiones y creando nuevos cambios a partir de las características existentes. Las técnicas suelen incluir listas de verificación, para fomentar el análisis más completo y exhaustivo del problema. Los siguientes ejemplos proceden fundamentalmente del campo del desarrollo de productos (Whiting, 1958):

## 1) Revisión sistemática

Revise todos los aspectos de un producto existente y sus mercados de manera sistemática, usando la siguiente lista de verificación. Escriba todas las posibilidades que se le ocurrran. No sea negativo: convénzase de que puede haber mejores soluciones.

| | | |
|---|---|---|
| PRODUCTO | Considerar cambios | • características físicas<br>• envoltorio<br>• método de fabricación |
| | Analizar | • posición frente a la competencia<br>• posibilidades técnicas |
| CONSUMIDOR | Considerar cambios | • método de uso<br>• propósito del uso<br>• posición percibida<br>• actitudes |
| | Analizar | • compradores alternativos<br>• estilos y tendencias de compra |
| PROMOCIÓN | Considerar cambios | • publicidad, promoción |
| | Analizar | • precios<br>• actividad competitiva |

### 2) Análisis morfológico

Esta técnica incluye la división de un problema en sus dimensiones constituyentes. Cada dimensión se subdivide a su vez en sus posibles valores o formas y las varias combinaciones resultantes son examinadas para verificar la utilidad de las soluciones que ofrecen. Consideremos, por ejemplo, el problema del desarrollo de un nuevo tipo de envase para leche. El envase tiene tres dimensiones: tamaño, forma y material. La dimensión del tamaño se subdivide en varias unidades (1 l, 750 cl, etc.); la forma se subdivide en las formas posibles (cubo, cilindro, tetraedro, etc.) y lo mismo sucede con los materiales (vidrio, cartón, metal, etc.). Se examinan sistemáticamente las combinaciones de estas categorías para averiguar su utilidad potencial como soluciones. Éste es un procedimiento mecánico y quizá el uso de computadoras podría simplificar la tarea.

### 3) Listado de atributos

Se hace una lista con los principales atributos de un producto o de un servicio (o de otro problema). Luego, se consideran las formas de cambiar o mejorar cada uno de ellos en todas las maneras imaginables, sin abrir juicios de valor. Concentrarse en los atributos peculiares de un producto ha resultado productivo para el desarrollo de nuevos productos, ya que se trata de sus características más significativas. El lector podría tratar de diseñar un nuevo cepillo de dientes, por ejemplo, usando este enfoque.

## El desarrollo de las habilidades del "hemisferio derecho"

Estas técnicas mejoran la creatividad estimulando asociaciones novedosas y a menudo excéntricas respecto de un problema. No buscan soluciones directas sino alentar el pensamiento lateral para generar analogías y metáforas, que casi parecen agregar confusión pero que pueden producir enfoques totalmente nuevos de un problema.

La esencia de estas técnicas radica en el uso de metáforas por las cuales se transportan diferentes perspectivas de un campo de referencia a otro. Schon (1979) ha enfatizado que los conceptos y el lenguaje que usamos pa-

ra representar un problema guían nuestro pensamiento hacia ciertos tipos de solución. Bajo el hechizo de la metáfora espontánea e implícita que usamos para plantear un problema, somos propensos a cierto tipo de soluciones más que a otras. Usando diferentes metáforas podemos cambiar las connotaciones del problema y el marco de referencia para una solución. Esto se conoce como "reestructuración del marco" y es un fenómeno psíquico importante y muy difundido involucrado en la percepción, las ilusiones, la memoria e incluso en la capacidad de apreciar las bromas. Este proceso queda claro en el siguiente ejemplo, tomado de Schon.

Un grupo de investigadores de desarrollo de productos intentaba mejorar el rendimiento de un pincel de cerda sintética. Habían probado varias soluciones para hacer que las cerdas sintéticas parecieran reales; habían, incluso, llegado a cambiar el diámetro de las cerdas y les habían hecho puntas divididas. Pero su pensamiento estaba limitado por la palabra "pincel", y eso los hacía buscar una solución que incluyera cambios en las cerdas. Cuando alguien del grupo observó que un pincel era una especie de bomba que propulsaba la pintura a través del espacio entre las cerdas, se produjo un gran avance. Esta analogía cambió la perspectiva del problema y se convirtió en una nueva metáfora que alteró el contexto, y se empezó a buscar entonces otra solución, cambiando el flujo de pintura entre las cerdas.

Las siguientes técnicas pretenden desarrollar habilidades en la creación de metáforas, analogías y asociaciones. (En Stein, 1975, pueden hallarse más detalles sobre las técnicas que se describen a continuación.)

## 1) Sinéctica

La sinéctica se basa en la unión de elementos diferentes y aparentemente irrelevantes (Gordon, 1961). El objetivo es hacer familiar lo extraño y viceversa, para cambiar de perspectiva. Hay muchos recursos para hacer esto, generando analogías y metáforas:

• *Analogías personales.* Una persona trata de imaginarse a sí misma como el objeto con el que está trabajando; primero, describiendo las propiedades y características del objeto; luego, imaginando los sentimientos y emociones del objeto; finalmente, identificándose con el objeto para experimentar lo que podría sentir si fuese eso. Fácilmente puede imaginarse la analogía del pincel-bomba, que acabamos de mencionar, surgiendo de este tipo de análisis íntimo de cómo sería uno si fuera un pincel.

• *Analogías directas.* Se buscan metáforas del problema pertenecientes a otro campo. Con frecuencia, se recurre a analogías para los problemas inorgánicos haciendo referencia a la forma en que operan los seres vivos. Por ejemplo, Brunel resolvió el problema de la construcción de túneles bajo el agua mediante una analogía directa con la forma en que una polilla excavaba en la madera; y, para inventar el teléfono, Bell encontró una forma de convertir la voz en energía mecánica y eléctrica mediante comparaciones con el mecanismo del oído humano.
• *Analogías simbólicas.* Se recurre a metáforas poéticas o simbólicas para caracterizar la esencia del problema. Se exploran las asociaciones no verbales tratando de encontrar formas, colores, texturas, olores, etc., que caractericen los conceptos del problema. El objetivo es lograr que los problemas abstractos resulten más concretos y viceversa.
• *Analogías fantasiosas.* El problema se plantea pensando en formas en que podría resolverse de manera simple y rápida en un mundo ideal en el que todo fuera posible. Al diseñar un cierre hermético para trajes espaciales, un grupo que utilizaba esta técnica soñó con una línea de insectos que cerrara sus garras en respuesta a una orden, para cerrar herméticamente los trajes. Esto llevó a otras ideas que proporcionaron una solución factible (Gordon, 1951).

La técnica de usar analogías aleja al pensamiento del problema por varios desvíos. Luego, hay que regresar al problema desde el lugar al que se ha llegado.

## 2) Tormenta de ideas

La tormenta de ideas suele realizarse con un grupo de personas, pero eso no es esencial. Se hace un listado de las ideas relacionadas con el problema como cadena de asociaciones, a menudo en forma de fluir de la conciencia. Hay cuatro reglas básicas para los torbellinos de ideas:

• Se eliminan las críticas. Deben generarse ideas sin pensar en su valor, su factibilidad o su importancia.
• Se fomenta la "conversación libre", para que pueda plantearse cualquier idea, no importa cuán absurda o intrascendente parezca. La intención es otra vez la de eliminar la selección y la censura y promover la libertad en las sugerencias.

• Se busca la cantidad: se intenta generar el mayor número posible de ideas. Suele suponerse que el énfasis en la cantidad da como resultado una mejor calidad, pero esto no ha quedado siempre demostrado en los estudios de investigación. Sin embargo, el énfasis en la cantidad ayuda a suspender la crítica y el juicio de valor.
• Se recurre a la combinación y la mejora de ideas y se da prioridad a las que se basen constructivamente sobre sugerencias previas. Esta regla fomenta el desarrollo de ideas y busca eliminar, a la vez, la vergüenza o la incomodidad al usar ideas de otros.

## La superación de los bloqueos en el pensamiento creativo

Las técnicas descritas para facilitar la creatividad sólo tienen éxito si se tiene la actitud mental correcta. Los bloqueos obedecen a razones cognitivas y sociales, aunque esos dos tipos de motivos no están, de ninguna manera, separados. El clima social en el que se realiza el trabajo creativo puede servir de ayuda o de obstáculo. Cuando las recompensas o premios dependen de la creatividad y hay supervisión externa y controles extrínsecos, la creatividad se reduce (Amabile, 1983). Esas condiciones limitan los riesgos que la gente se atreve a correr en su búsqueda de alternativas imaginativas, enfatizando el valor social que se podría poner en sus excursiones creativas. Es más probable que la gente sea creativa cuando está motivada intrínsecamente, cuando un problema le importa realmente. La creatividad exige una perspectiva positiva y confiada y la seguridad de ser no conformista y de cuestionar el status quo. El proceso creativo exige individualidad y singularidad, que se consiguen mejor en un medio que sea socialmente poco amenazante, tranquilo y seguro.

Los bloqueos en la creatividad suelen ser dificultades en el paso del análisis sistemático y racional del problema (pensamiento del hemisferio izquierdo) al uso de la imaginación y la libre asociación para crear nuevas perspectivas (pensamiento del hemisferio derecho). Esto no es sorprendente, ya que muchas instituciones (la escuela, el trabajo) le otorgan mayor valor y prestigio al análisis lógico y racional que a las asociaciones libres "infantiles" y a las analogías fantasiosas. Sin embargo, debemos recordar el consejo de Schiller de no permitir que el intelecto limite la imaginación. Para lograrlo, hay ciertas ideas que vale la pena recordar:

- Elija condiciones de trabajo conducentes y apropiadas, que eviten las distracciones y alienten la satisfacción, la atención relajada y la diversión.
- Tenga a mano objetos relacionados con el problema, cámbielos de posición y juegue con ellos; preste atención a la búsqueda de los objetos y la observación de sus características; use la imaginación visual de manera similar para estudiar las propiedades de las cosas.
- No intente retener las ideas mentalmente; tome notas y haga listas; esto ayuda a evitar la sobrecarga de aptitudes cognitivas y también libera al pensamiento para permitir que surjan otras ideas.
- Intente, deliberadamente, ser original, creativo y osado.

## Conclusión

La creatividad como la de Arquímedes, Darwin o Picasso empieza con la búsqueda de una solución a un problema con el que el individuo creativo está profundamente comprometido. La búsqueda de una solución pone en actividad muchas habilidades cognitivas, incluyendo el análisis, el razonamiento, el recuerdo, la observación, el reconocimiento, la asociación y la creación de analogías y metáforas cambiantes. Hace mucho que se piensa que la creatividad tiene procesos conscientes y procesos inconscientes. Se ha dicho que la mística peculiar y la fuerza de la creatividad radica en la incubación inconsciente y el *insight* en el problema. Sin embargo, una forma alternativa de observar los procesos involucrados es que la creatividad aprovecha aptitudes complementarias en las que se especializan los hemisferios cerebrales derecho e izquierdo. La esencia de la creatividad radica en la capacidad de generar alternativas en la búsqueda de una solución, tanto a través del razonamiento lógico (hemisferio izquierdo) como mediante la asociación y las metáforas, que crean nuevas perspectivas o marcos de referencia para el problema (hemisferio derecho). Cuanto mayor sea la gama de alternativas consideradas, más probable es que surja una solución creativa a un problema nuevo y difícil. Para ser creativos, entonces, debemos concentrarnos en habilidades para generar una gama de alternativas nuevas, mejores y más amplias en nuestro pensamiento. Pero en el análisis final debemos recordar el papel de la casualidad y la suerte para tropezar con la alternativa profunda, para tener en cuenta el incidente importante pero inadvertido o para vivir la experiencia correcta en el momento preciso. Quizá lo máximo que pueden hacer las personas es fomentar la aparición de esas casualidades y volverse receptivos a las que se cruzan en su camino.

## Lecturas adicionales

BRANSFORD, J.D. y STEIN, B.S. (1984) *The Ideal Problem Solver*. San Francisco: Freeman.

Este libro quiere fomentar la mejor resolución de problemas describiendo los procesos involucrados y usando ejercicios para estimular las habilidades necesarias.

PERKINS, D.N. (1981) *The Mind's Best Work*. Cambridge, Mass.: Harvard University Press.

Profunda reevaluación de la naturaleza de la creatividad, que subraya el papel que juegan las habilidades cognitivas cotidianas normales.

# 16

# Las computadoras y la cognición

*John Sloboda*

Como dice Angus Gellatly en el capítulo 1, hemos intentado evitar el uso excesivo de la metáfora de la computadora en nuestro análisis de los diversos aspectos de la cognición. Una razón fundamental para hacerlo es que la mayoría de las personas no entiende muy bien a las computadoras y no tiene mucho sentido intentar explicar el funcionamiento de la mente según algo que en general no se conoce. Sin embargo, ignorar por completo el papel de la computadora en la psicología cognitiva moderna sería un grave perjuicio.

Este capítulo se ocupa de dos usos específicos de las computadoras, que se han convertido en el sello distintivo de una disciplina llamada ciencia cognitiva, lugar de encuentro de psicólogos, filósofos, lingüistas y científicos computacionales (Norman, 1981). El primero de estos usos es el de la metáfora de la computadora como medio para precisar nuestra comprensión de qué es o no es la mente. Esta es una empresa que involucra a los exponentes más filosófica y conceptualmente representativos de la ciencia cognitiva. La segunda, aunque estrechamente vinculada, es la escritura de programas de computación que imitan algún aspecto del funcionamiento humano. Esto se denomina *simulación por computadora* y es una rama de la disciplina conocida como inteligencia artificial. Requiere cierto grado de dominio de la programación y comprensión de las posibilidades y límites de las computadoras como herramientas prácticas. En un libro de este tipo, es más fácil disfrutar del primer aspecto, conceptual, que escribir sobre el segundo, más práctico. Aunque me referiré a ambos, dedicaré proporcionalmente más espacio al primero. Quienes deseen profundizar en los detalles de la simulación deben consultar las sugerencias de lecturas adicionales, al final del capítulo.

## ¿Qué es una computadora?

Una computadora es una máquina. Las máquinas realizan funciones físicas específicas y repetibles, según ciertas ordenes, llamadas, en inglés, *inputs*. Por ejemplo, mi máquina de escribir es una máquina. Cuando le doy una orden específica (*input*) presionando una tecla, responde moviendo un martillo determinado. El mismo martillo se mueve cada vez que presiono esa tecla. Las máquinas generalmente actúan trasnformando la energía de alguna forma. Una máquina de escribir transforma la presión hacia abajo de mis dedos en los golpes horizontales de los martillos sobre el papel. Un automóvil transforma la fuerza explosiva de la ignición de combustible y los movimientos de manos y pies de su conductor en movimientos rápidos pero controlados. Las formas de la energía involucrada pueden ser diversas: mecánicas, eléctricas, electromagnéticas o atómicas.

Una de las características principales de la mayoría de las máquinas es que sus funciones están rígidamente limitadas. La configuración física de sus partes las restringe a una tarea en particular. No se puede convertir una máquina de escribir en otra cosa sin desarmar sus partes y reconstruirlas. Podemos decir que la configuración física fija de una máquina es su *hardware*. En la mayoría de las máquinas, entonces, el *hardware* determina la función.

En cierto sentido, todo esto también es cierto respecto de las computadoras. Una computadora específica toma los impulsos eléctricos como *inputs*, los mueve por dentro de ella en sentidos totalmente determinados por su configuración física y produce salidas (*outputs*) eléctricas transformadas. Si uno alimenta la misma computadora con los mismos *inputs*, siempre obtendrá los mismos *outputs* (siempre y cuando la configuración física no funcione mal). Pero cuando observamos la función de una computadora, aparece una inmensa diferencia. No podemos definir una función específica de una computadora de la misma forma que podemos hacerlo con una máquina de escribir. Más estrictamente, podemos decir que la función de una computadora consiste en manipular símbolos. Si esto no resulta suficiente, podemos aclararlo diciendo que, en principio, se puede hacer que una computadora imite a cualquier otra máquina, sin alterar su configuración física. Esto se hace dividiendo el *input* en dos conjuntos: un conjunto de *inputs* es llamado el programa, y este determina cómo se manejará el segundo conjunto (los datos).

Regresemos un momento a mi máquina de escribir. Funciona como

funciona porque tiene un conjunto de varillas y palancas que conectan determinadas teclas con determinados martillos. Es fácil observar el funcionamiento de estas palancas levantando la tapa. Pero supongamos que tuviéramos que sellar la tapa y ya no pudiéramos mirar el interior; supongamos, además, que tuviéramos que reemplazar la conexión mecánica entre teclas y martillos por conexiones eléctricas que se dirigieran mediante una computadora. Si programáramos correctamente la computadora, no podríamos decir (con sólo escribir en la máquina) que las conexiones no son mecánicas. De hecho, la computadora hace que mi máquina de escribir sea mucho más versátil. Puedo reprogramarla para que cualquier presión sobre una tecla dé como resultado el golpe de martillo que se me ocurra. Incluso puedo hacer que la presión sobre una sola tecla se convierta en una palabra o en una oración entera.

El programa de computación hace esto básicamente informando a la computadora qué *outputs* corresponden a determinados *inputs* o combinaciones de *inputs*. Si identificamos cada *input* de las diferentes teclas de la máquina de escribir con un número arbitrario y cada *output* de los martillos con otro número, parte del programa podría decir:

SI EL *INPUT* ES 27, QUE EL *OUPUT* SEA 36

o bien

SI EL *INPUT* ES 13, QUE EL *OUPUT* SEA 11, LUEGO 3, LUEGO 15 Y LUEGO 19

Lo crucial consiste en advertir que la computadora no tiene idea de qué significan estos *inputs* y *outputs* en la realidad. No tiene forma de saber que, en este caso en particular, el *input* 27 significa que alguien está presionando la tecla W, o que el *output* 36 activa el martillo W. Para la computadora, sólo se trata de símbolos abstractos que conecta según las reglas de su programa. Estos símbolos podrían ordenar conducir un automóvil, calcular un salario o comenzar la tercera guerra mundial.

Por lo tanto, se puede pensar en un programa de computación como un conjunto de instrucciones para que una computadora manipule símbolos de determinada manera al recibir determinados *inputs*. La computadora es una *máquina de uso general* que un programa convierte, cuando está cargado, en una máquina *específica* para realizar una tarea específica.

Para nuestros propósitos, no es necesario conocer en detalle cómo la configuración física de una computadora nos permite almacenar programas y datos. Se la puede imaginar como un conjunto de muchos miles de células

eléctricas conectadas de una manera específica, donde cada célula puede encenderse o apagarse mediante la aplicación de corrientes eléctricas a un número relativamente pequeño de células de *input*. Un programa determina la secuencia o la lógica de los pasajes de impulsos a través de la configuración física. El diseño de la configuración física de las computadoras ha cambiado mucho en los últimos 30 años. Donde antes había válvulas, ahora hay chips microelectrónicos. Sin embargo, uno puede hacer funcionar programas escritos hace 30 años en las computadoras actuales y obtener los mismos resultados. Así como las computadoras pueden imitar a otras máquinas, también pueden imitarse entre sí. Los avances en las configuraciones físicas no han alterado en esencia el tipo de programas que se pueden hacer funcionar; antes bien, han hecho que las computadoras sean más baratas, más rápidas, más confiables y capaces de manejar programas mayores y grandes cantidades de datos.

## La metáfora de la computadora para la mente

Como se señala en el capítulo 1, muchos investigadores se han sentido impresionados por las semejanzas entre las computadoras y el cerebro humano, y esto ha planteado nuevas posibilidades para comprender los procesos mentales. Veamos algunas de las principales semejanzas que se han propuesto:

a) Tanto las computadoras como los cerebros pueden considerarse procesadores de información. Es decir, reciben *inputs* de diversas fuentes, los transforman y manipulan y producen *outputs*.

b) Se puede decir que ambos logran sus diversos resultados mediante la combinación de muchos elementos similares que desempeñan funciones simples. En las computadoras, las células eléctricas se encienden y apagan. En los cerebros, las células nerviosas se excitan o inhiben.

c) Tanto las computadoras como los cerebros pueden considerarse dispositivos de uso general. Pueden realizar muchas operaciones diferentes con la misma configuración física.

d) Ambos pueden almacenar y utilizar grandes cantidades de información.

e) Se podría decir que ambos siguen planes de acción organizados (programas).

Estas similitudes se relacionan con algunas conclusiones generalmente aceptadas y no controvertidas. En primer lugar, que las computadoras pueden imitar exitosamente algunos aspectos de la conducta humana. Para tomar sólo un ejemplo, Siegler (1986) ha escrito un programa de computación que simula el desempeño de los niños en los problemas de resta (por ejemplo, "seis menos cuatro"). Es decir, que cuando se lo hace funcionar en una computadora, el programa comete el mismo tipo de errores que los niños. También le lleva el mismo tiempo que a los niños resolver los problemas. Del mismo modo, mejora con la experiencia, igual que los niños. Como Siegler escribió el programa, puede describir su lógica subyacente y proponerla como modelo hipotético de la forma en que los niños realmente encaran la tarea de restar números. Si los niños usan esta lógica, tendrán exactamente el mismo patrón de errores y tiempos que, de hecho, tienen. (Sin embargo, debemos recordar que es posible que haya otros sistemas que puedan producir los mismos resultados.)

En segundo lugar, la mayoría de los científicos acepta que la disciplina intelectual necesaria para producir un programa que realmente funcione es una valiosa ayuda para teorizar mejor sobre el funcionamiento humano. Como a una computadora hay que decirle exactamente lo que debe hacer, no hay lugar para el pensamiento confuso o impreciso. La computadora también demuestra las incoherencias ocultas en el pensamiento de una forma muy clara. Si hay contradicciones internas, el programa sencillamente no funcionará o dará resultados extraños. Por esta razón, muchos educadores (por ejemplo, Papert, 1972) ven a la computadora como un valioso "gimnasio intelectual". Hay algunos aspectos del control de una computadora que pueden resultar casi necesariamente útiles (ver también Weizenbaum, 1984) y así los niños adquieren la habilidad como subproducto incidental de "hacer que el maldito aparato funcione".

Pero más allá de esos puntos generalmente aceptados, hay un gran debate en torno a lo que a veces se denomina "ardiente defensa" de la inteligencia artificial, que sostiene que cuando ciertos tipos de programas están operando en una computadora, ésta no se limita a imitar la conducta humana sino que realmente puede decirse que "piensa", "sabe" o "entiende".

## ¿Las computadoras piensan?

El advenimiento de las computadoras planteó de inmediato en la conciencia popular profundas preguntas sobre la naturaleza última de la concien-

cia y la personalidad. Muchas de estas preguntas han sido obsesivamente trajinadas en la ciencia ficción, particularmente en lo que tiene que ver con los "androides", robots de carne y piel sintética que los hacen parecer humanos. Algunas de esas preguntas tienen implicaciones morales. Por ejemplo: ¿debe tratarse a los androides como a seres humanos con derechos? Los relatos en los que las computadoras esclavas enloquecen y se vuelven contra sus amos expresan el profundo temor (ya desde la historia de Frankenstein) de que los científicos estén abriendo una caja de Pandora.

El desarrollo, en la ficción popular, de nuestra fascinación con las computadoras no debería disminuir la profundidad de los temas planteados. De hecho, la obsesión popular puede considerarse una medida de su profundidad. Estos temas siguen acosando y polarizando a la comunidad intelectual. Me gustaría esbozar las opiniones de dos claros y convincentes defensores de puntos de vista opuestos sobre las computadoras: Alan Turing y John Searle. Nunca se conocieron personalmente. Turing murió hace unos años y Searle mantiene su debate con los discípulos intelectuales de Turing. Las ideas de Turing están bella e inteligentemente expresadas en un artículo publicado en 1950 y las de Searle muestran el mismo aplomo en un ensayo de 1980. Hofstadter y Dennett (1981) han reimpreso extractos de lo más relevante de ambos autores.

## El test de Turing

Turing empieza preguntándose cuál podría ser un test adecuado para comprobar si una computadora piensa. Rechaza con justicia la idea de que debamos ocuparnos de la apariencia de la computadora (es decir, si podemos hacerla parecer humana) y nos pide que imaginemos una situación en la que un interrogador se sienta en una habitación frente a dos teletipos. Estas máquinas le permiten enviar mensajes a otras personas ubicadas en otro lugar y recibir mensajes enviados por ellas. En el test de Turing, una de las "personas" es un ser humano y la otra es una computadora. La tarea del interrogador es tratar de determinar, preguntando lo que se le ocurra, cuál de sus teletipos lo comunica con el ser humano y cuál con la computadora. Si quien interroga no puede tomar una decisión segura después de un extenso interrogatorio, podremos decir que la computadora ha pasado el test de Turing. Turing sostiene que si una computadora puede pasar una prueba así, no nos quedará más remedio que aceptar que piensa.

Esta es la formulación clásica de la "ardiente defensa" de la inteligencia artificial. En su ensayo, Turing menciona luego algunas objeciones posibles. Veamos apenas dos de las principales y las respuestas que él mismo da:

OBJECIÓN 1: Aunque una computadora actúe como si pensara, no se puede concluir que dentro de ella suceda algo afín a lo que llamaríamos *experiencia consciente*.

RESPUESTA: El único ser del cual uno puede saber con certeza que tiene esas experiencias es uno mismo. Uno no sabe si los demás seres humanos también tienen conciencia pero supone que la tienen. ¿Por qué no, entonces, las computadoras?

OBJECIÓN 2: Aunque las computadoras pueden *simular* algunos aspectos de la conducta humana, jamás podrán enamorarse, hacer algo realmente creativo, etc.

RESPUESTA: Acepto que no hay programa de computación existente que pueda aprobar todavía el test de Turing. Pero esta no es razón suficiente para suponer que no podemos seguir creando programas que cubran cada vez más aspectos de la conducta humana y que algún día no tendremos éxito. Habría que ofrecer alguna razón fundamentada que demostrara que ciertas cosas son imposibles de programar.

Nos ocuparemos ahora de Searle, quien, como Turing, nos pide que juguemos con él un juego del pensamiento, para demostrar su posición.

## El "salón chino" de Searle

El juego imaginario de Searle se denomina "Test del salón chino". Se basa en un programa de computación descrito por Schank y Abelson (1977) que, según sostienen algunos, "entiende" historias. El programa simula la operación del tipo de conocimientos que hemos visto que los lectores usan para interpretar historias o discurso (capítulos 4 y 10). Las palabras de Searle son tan claras y convincentes que dejaré que hablen por sí mismas:

En resumen, y sin mencionar los detalles, puede describirse así al programa de Schank: el objetivo del programa es simular la capacidad humana

para entender historias. Es característico de la capacidad de comprensión de historias de los seres humanos el poder responder preguntas sobre la historia, aun cuando la información pedida no esté explícitamente mencionada en la historia. Así, por ejemplo, supongamos que a uno se le cuenta la siguiente historia: "Un hombre entra a un restaurante y pide una hamburguesa. Cuando se la traen, está dura como una piedra. El hombre sale del restaurante enojado, sin pagar la hamburguesa ni dejar propina." Ahora bien, si a uno le preguntan si el hombre se comió la hamburguesa, probablemente responda que no. Del mismo modo, veamos la siguiente historia: "Un hombre entra a un restaurante y pide una hamburguesa. Cuando se la traen, el hombre piensa que tiene un aspecto apetitoso. Cuando se va del restaurante, le deja a la camarera una buena propina después de pagar su cuenta." Si esta vez nos preguntan si el hombre comió la hamburguesa, probablemente responderemos que sí. Las máquinas de Schank pueden responder preguntas sobre restaurantes de una forma parecida. Para hacerlo, tienen una "representación" del tipo de información que los seres humanos poseen sobre los restaurantes, lo cual les permite contestar preguntas como las planteadas ante historias de este tipo. Cuando a la máquina se le da la historia y luego se le hace la pregunta, imprimirá respuestas semejantes a las que serían esperables en una persona si se le contara una historia así. Los defensores de la inteligencia artificial sostienen que en esta secuencia de preguntas y respuestas, la máquina no sólo está simulando una capacidad humana sino también: 1) que se puede decir que la máquina literalmente entiende la historia y responde las preguntas, y 2) que lo que hacen la máquina y su programa *explica* la capacidad humana de entender la historia y responder preguntas sobre ella.

Me parece que el trabajo de Schank no ofrece sustento a ninguna de las dos afirmaciones, e intentaré demostrarlo a continuación.

Una forma de verificar cualquier hipótesis sobre la mente es preguntarse cómo sería si mi mente realmente funcionara según los principios que, de acuerdo con la teoría, rigen el funcionamiento de todas las mentes. Apliquemos esta prueba al problema de Schank con el siguiente *Gedankenexperiment* (experimento del pensamiento). Supongamos que estoy encerrado en una habitación con una gran carpeta con escritura china. Supongamos, además, que no sé nada de chino, oral o escrito (como sucede en la realidad), y que no creo poder reconocer la diferencia entre la escritura china y, por ejemplo, la japonesa o meros garabatos sin sentido. Para mí, la escritura china son muchos garabatos sin sentido. Supongamos también que después de

esta primera carpeta con escritura china me dan una segunda carpeta con escritura china junto con un conjunto de reglas para correlacionar la segunda carpeta con la primera. Las reglas están en inglés y yo las entiendo como las entendería cualquier otro angloparlante. Me permiten correlacionar un conjunto de símbolos formales con otro. "Formales" significa, en este caso, que puedo identificar los símbolos por sus formas. Supongamos también que me dan una tercera carpeta de símbolos chinos junto con instrucciones, nuevamente en inglés, que me permiten correlacionar elementos de esta tercera carpeta con las dos primeras, y estas reglas me enseñan a darles ciertos tipos de formas a ciertos símbolos chinos, en respuesta a ciertos tipos de formas que hay en la tercera carpeta. Quienes me dan todos estos símbolos, desconocidos para mí, llaman "libreto" a la primera carpeta, "historia" a la segunda y "preguntas" a la tercera. Además, a los símbolos que les doy en respuesta a la tercera carpeta los llaman "respuestas a la pregunta", y denominan "programa" al conjunto de reglas en inglés que me han dado. Para complicar un poco la cosa, imaginen que esta gente también me da historias en inglés, que comprendo, y luego me hacen preguntas en inglés sobre estas historias, y yo les respondo en inglés. Supongamos también que después de un tiempo llego a ser tan bueno siguiendo las instrucciones para manipular los símbolos chinos y los programadores llegan a ser tan buenos escribiendo los programas que, desde un punto de vista exterior (es decir, desde el punto de vista de alguien que esté fuera de la habitación en la que estoy encerrado), mis respuestas a las preguntas son absolutamente imposibles de diferenciar de las respuestas de los chinos nativos. Nadie puede decir, viendo mis respuestas, que no sé una sola palabra de chino. Supongamos también que mis respuestas a las preguntas en inglés son (como seguramente serían) imposibles de diferenciar de las de otros hablantes nativos de inglés, por la simple razón de que yo mismo soy un hablante nativo de inglés. Desde el punto de vista exterior (desde el punto de vista de alguien que lee mis "respuestas"), las respuestas a las preguntas en chino y a las preguntas en inglés son igualmente buenas. Pero, en el caso del chino (a diferencia del inglés), produzco las respuestas manipulando símbolos formales no interpretados. En lo que respecta al chino, me comporto simplemente como una computadora; realizo operaciones computacionales con elementos formalmente especificados. Para el chino, soy simplemente una ejemplificación del programa de computación.

Los defensores de la inteligencia artificial sostienen que la computadora programada entiende las historias y que el programa en alguna medida ex-

plica la comprensión humana. Pero ahora estamos en condiciones de examinar estos argumentos a la luz de nuestro experimento del pensamiento.

1. En lo que respecta a la primera afirmación, me parece más que evidente, por el ejemplo, que no entiendo una sola palabra de las historias chinas. Tengo *inputs* y *outputs* imposibles de distinguir de los de un hablante de chino y puedo tener cualquier programa formal, pero sigo sin entender nada. Por las mismas razones, la computadora de Schank no entiende nada de ninguna historia (escrita en chino, inglés o en cualquier otro idioma), ya que en el ejemplo chino la computadora soy yo, y cuando la computadora no soy yo, ella no tiene más de lo que yo tengo en el caso en que no entiendo nada.

2. En lo que respecta a la segunda afirmación, que el programa explica la comprensión humana, podemos ver que la computadora y su programa no proporcionan las condiciones de comprensión suficientes ya que, aunque ambos están funcionando, no hay comprensión.

Si aceptamos los argumentos de Searle (y son muchos los que han intentado demolerlos), debemos preguntarnos qué es lo que hace tan especiales a los seres humanos. En esto, Searle no es tan preciso. En síntesis, considera que atributos intencionales tales como el pensamiento, el saber y la comprensión tienen que ver con el hecho de que la computadora humana se manifiesta en nervios, carne y hueso.

Lo que importa de las operaciones cerebrales no es la representación formal de las secuencias sinápticas sino más bien las propiedades reales de las secuencias. Todos los argumentos que he visto en defensa de la inteligencia artificial insisten en trazar un perfil en torno a las imágenes que produce la cognición y luego afirman que esas imágenes son la realidad.

En este punto, dejemos descansar al debate. Si se resuelve (y cuando se resuelva) nos acercará a la esencia de lo humano. Por ahora, debemos seguir viviendo en la incertidumbre.

## Programas y simulaciones

Durante muchos años, los expertos en computación han estado creando programas de ajedrez, buscando producir alguno que ningún ser humano pueda derrotar. El ajedrez representa un desafío estimulante pára los programadores

porque sus reglas básicas son muy simples y bien definidas y sin embargo casi no hay límites para el refinamiento en el dominio del juego. En 1957, Herbert Simon predijo que en diez años habría un programa de computación al que nadie podría vencer. Treinta años después su predicción sigue sin cumplirse, aunque hay un par de sistemas de origen estadounidense (en especial el "Belle", de Bell Laboratories, Nueva Jersey) que pueden derrotar al 99,5% de los jugadores.

## Algoritmos

Como ya se comentó en el capítulo 14, hay básicamente dos enfoques para encarar un problema de cálculo. Uno es intentar descubrir un *algoritmo* eficiente para la tarea; el otro consiste en desarrollar un sistema *heurístico* funcional. Un algoritmo es un procedimiento que garantiza una solución para problemas de determinado tipo. Por ejemplo, la mayoría de nosotros aprendimos algoritmos para resolver multiplicaciones largas cuando estábamos en la escuela. Si se siguen todos los pasos en el orden correcto y no se cometen errores de cálculo, siempre se llegará a la solución correcta.

¿Existe un algoritmo para el ajedrez? En teoría, sí. Incluye el análisis de cada movimiento posible, luego el de cada respuesta posible del rival, luego el de cada respuesta propia posible, y así sucesivamente hasta haber explorado las consecuencias de todos los movimientos posibles hasta el momento en que alguno de los jugadores gana. Luego, uno elige el movimiento asociado con la mayor cantidad de resultados ganadores.

Por desgracia, un algoritmo así es impracticable. Se ha calculado que hay más partidas de ajedrez posibles que átomos en el universo. Si uno quisiera analizar apenas cinco movimientos por adelantado, tendría que considerar hasta 50.000.000 de diferentes combinaciones de movimientos. Probablemente a la computadora más veloz del mundo (que contara con una capacidad de almacenamiento de información ilimitada) le llevaría varios millones de años jugar una partida de ajedrez usando semejante algoritmo.

## Heurística

Para "reducir el campo de búsqueda" en los problemas de este tipo, hacen falta las heurísticas. Son reglas empíricas que tienen una probabilidad razonable de producir resultados aceptables. Es bastante evidente que la resolu-

ción de problemas en el ser humano normalmente se realiza con una gran provisión heurística (ver capítulos 13 y 14) y que esas heurísticas conforman el caudal con el que cuentan la mayoría de los entrenadores. Así, casi todos los manuales de ajedrez aconsejan a los jugadores elegir movimientos tendientes a ganar el control del centro del tablero. Otra heurística consiste en comer las piezas del adversario cada vez que se presenta la oportunidad. Cada heurística tiene sus propias limitaciones. En el ajedrez, un jugador puede hacer un "sacrificio" mediante el cual gana una ventaja mayor si su adversario se siente tentado por la recompensa menor de comer una de sus piezas.

Pero en general las ventajas de las heurísticas son mucho mayores que sus limitaciones. Muchos problemas humanos están demasiado mal definidos como para poder aplicar algoritmos. A menudo, no tenemos conciencia de todas las posibilidades disponibles y por lo tanto no podemos evaluar todos los cursos de acción posibles. Incluso cuando podemos hacerlo, puede requerir demasiado tiempo o exigir demasiado esfuerzo psíquico. Las heurísticas ofrecen la perspectiva de respuestas rápidas a diversas situaciones. En la vida real, uno no suele contar con tiempo ilimitado para tomar una decisión. Las heurísticas simples también han demostrado tener muy amplia aplicabilidad en computación. Los trabajos pioneros de Simon y sus colegas (Newell y Simon, 1972) dieron como resultado un programa llamado *General Problem Solver* (Solucionador General de Problemas, SGP), basado en un conjunto de heurísticas, una de las cuales ha sido denominada "heurística de medios-fines" (ver capítulo 14). En pocas palabras, dice que si uno quiere pasar de una condición inicial A a una condición final Z y hay algún estadio intermedio M que de alguna forma está más cerca de Z que A, lo que hay que hacer es intentar llegar a M. Este es el enfoque subyacente en el problema de "llegar a Londres" que se plantea en el capítulo 2.

Newell y Simon demostraron que heurísticas como ésta pueden formularse e implementarse con rigor en un programa de computación que sea capaz de encarar coherentemente una gran variedad de problemas (por ejemplo, encontrar pruebas de las afirmaciones matemáticas). También se puede considerar a las reglas de producción descritas en el capítulo 2 como corporizaciones de heurísticas (relacionadas con la heurística de medios-fines, pero de una naturaleza bastante más específica). Un sistema como el SGP puede implementarse fácilmente como sistema de producción de computación, y ya existen programas que pueden aprender, a partir de heurísticas bastante generales que se dan como supuestos, reglas específicas adaptadas a situaciones individuales (Anderson, 1982).

Los programadores de ajedrez deben recurrir a las heurísticas porque es la única forma que tienen para lograr que alguna vez las computadoras jueguen al ajedrez bien y rápido. En general, no se preocupan por imitar el desempeño humano. Su objetivo es superarlo, hacer el mejor trabajo posible. Como ellos, muchos trabajadores en inteligencia artificial están preocupados fundamentalmente por proporcionar buenas soluciones a los problemas existentes. Siempre que haya algoritmos posibles preferirán usarlos porque aumentan la confiabilidad. También aprovechan la increíble velocidad con que las computadoras pueden realizar operaciones reiterativas. La velocidad y la confiabilidad son las principales razones por las que las computadoras han reemplazado al ser humano en muchos de los campos laborales más mecánicos. Una computadora puede encargarse de los salarios del personal con mucha mayor eficiencia que una persona, si está programada para hacerlo. Pero para aquellos que están interesados fundamentalmente en usar las computadoras como vinculadores, y así entender mejor el funcionamiento humano, no es importante tener el mejor programa. Precisamente el programa que resulta de mayor interés psicológico es aquel que comete errores "humanos" y muestra la capacidad humana para ir más allá de los datos que se proporcionan (como en el programa de "comprensión del relato" de Schank).

Es posible que en niveles bastante periféricos del funcionamiento sensorial y motriz (ver, por ejemplo, Marr, 1983) el cerebro actúe según principios algorítmicos. El cálculo de la profundidad visual a partir de una imagen estereoscópica parece estar basado en un algoritmo. Pero tan pronto como entramos en pleno territorio cognitivo, las heurísticas parecen abundar. Desde la percepción de la forma visual hasta la creación de una obra de arte, la característica humana de la imperfección parece conferirnos una ventaja peculiar. Aunque las computadoras están llegando lentamente a poder imitar algunas de estas cualidades de flexibilidad e ilimitación que caracterizan a la mente humana, no es para nada evidente que el enfoque de la computadora haya explicado o vaya a explicar todos los misterios no resueltos. Como todas las metáforas, realza las semejanzas y aclara las ideas, pero la comprensión que ofrece sólo es parcial.

WEIZENBAUM, J. (1984) *Computer Power and Human Reason* (2ª edición). San Francisco: Freeman.

Evaluación provocativa y crítica de las simulaciones por computadora, que adopta un punto de vista similar al de Searle.

NEWELL, A. y SIMON, H. (1972) *Human Problem Solving*. Englewood Cliffs, Nueva Jersey: Prentice Hall.

Clásico informe de su propia obra germinal. Por momentos, de lectura ardua.

# Referencias bibliográficas


ADAMS, J. L. (1979) *Conceptual Blockbusting* (2da. edición). San Francisco: Freeman.

AITCHISON, J. (1983) *The Articulate Mammal.* Londres (2da. edición).

ALLPORT, D. A., ANTONIS, B. Y REYNOLDS, P. (1972) "On the division of attention: a disproof of the single channel hypothesis". *Quarterly Journal of Experimental Psychology, 24,* 225-35.

AMABILE, T. M. (1983) *The Social Psychology of Creativity.* Nueva York: Springer Verlag.

ANDERSON, J. R. (1982) "Acquisition of cognitive skills". *Psychological Review, 89,* 369-406.

ARGYLE, M. (1969) *Social Interaction.* Londres: Tavistock publications in association with Methuen & Co.

ARGYLE, M. (1983) *The Psychology of Interpersonal Behaviour.* Harmondsworth: Penguin.

ARGYLE, M. (1984) "Some new developments in social skills training". *Bulletin of the British Psychological Society, 37,* 405-10.

ARGYLE, M. Y COOK, M. (1976) *Gaze and Mutual Gaze.* Cambridge: Cambridge University Press.

AUSTIN, J. L. (1962) *How to do things with words.* J. O. Urmson (Comp.). Londres: Oxford University Press.

BADDELEY, A. D. (1979) "Working memory and reading". En: P. A. Kolers, M. E. Wrolstad y H. Bouma (Comps.). *Processing of Visible Language.* Vol. I. Nueva York: Plenum.

BADDELEY, A. D. (1984) "Neuropsychological evidence and the semantic episodie distinction". *The Behavioural and Brain Sciences, 7,* 238-39.

BADDELEY, A. D. Y HITCH, G. J. (1974) "Working Memory". En: G. H. Bower (Comp.). *The Psychology of Learning and Motivation* (Vol. 8). Nueva York: Academic Press.

BAKER, L. (1979) "Comprehension monitoring: identifying and coping with text confusions". *Journal of Reading Behaviour, 11,* 366-74.

BAKER, L. Y ANDERSON, R. I. (1982) "Effects of inconsistent information on text processing: evidence for comprehension monitoring". *Reading Research Quarterly, 17,* 281-94.

BARTLETT, F. C. (1958) *Thinking.* Londres: Allen & Unwin.

BATESON, G. (1982) "Totemic knowledge in New Guinea". En: U. Neisser (Comp.). *Memory Observed.* San Francisco: Freeman.

BEARD, R. M. Y HARTELEY, J. (1984) *Teaching and Learning in Higher Education* (4ta. edición). Londres: Harper & Row.

BEATTIE, G. W. (1981) "Interruption in conversational interaction, and its relation to the sex and status of the interactants". *Linguistics, 19,* 15-35.

BEATTIE, G. W. (1982) "Turntaking and interruption in political interiews - Margaret Thatcher and Jim Callaghan compared and contrasted". *Semiotica, 39,* 93-114.

BEATTIE, G. W. (1983) *Talk: An analysis of speech and non-verbal behaviour in conversation.* Milton Keynes: Open University Press.

BEGG, I. Y DENNY, J. P. (1969) Empirical reconciliation of atmosphere and conversion interpretations of syllogistic reasoning errors. *Journal of Experimental Psychology, 81,* 351-54.

BELMONT, J. M. Y BUTTERFIELD, E. C. (1971) "Learning strategies as determinantes of mental deficiencies". *Cognitive Psychology, 2,* 411-20.

BIRDWHISTELL, R. L. (1968) "Kinesics". *International Encyclopedia of the Social Sciences*, Vol. 8, 379-385. Nueva York: MacMillan.

BLURTON-JONES, N. (1972) *Ethological Studies of Child Behaviour.* Londres: Cambridge University Press.

BOGEN, J. E. (1977) "Some educational implications of hemispheric specialization". En: M. C. Wittrock (Comp.). *The Human Brain.* Nueva York: Prentice Hall.

DE BOISBAUDRAN, H. L. (1911) *The Training of the Memory in Art.* Londres: Macmillan.

BOWER, G. H.(1972) "A selective review of organizational factors in memory". En: E. Tulving y W. Donaldson (Comps.). Organization of Memory. Nueva York: Academic Press.

BOWER, G. H. (1976) "Comprehending and recalling stories". Documento presentado en el Encuentro en Washington de la Asociación Psicológica Americana.

BOWER, G. H. Y KARLIN, M. B. (1974) "Depth of processing pictures of faces and recognition memory". *Journal of Experimental Psychology, 103,* 751-57.

BRANSFORD, J. D. (1979) *Human Cognition: Learning, Remembering and Understanding.* Belmont: Wadsworth.

BRILL, A. A. (1938) *The basic writings of Sigmund Freud.* Nueva York: Random House.

BROADBENT, D. E. (1982) "Task Combination and selective intake of information". *Acta Psychologica, 50,* 253-90.

BROWN, A. L., BRANSFORD, J. D. , FERRARA, R. A. Y CAMPIONE, J. C. (1983) "Learning, remembering and understanding". En: P. H. Mussen (Comp.). *Handbook of Child Psychology,* Vol. 3. Nueva York: Wiley.

BROWN, A. L., CAMPIONE, J. C. Y DAY, J. D. (1981) "Learning to learn: on training students to learn from texts". *Education Researcher, 10,* 14-21.

BROWN, A. L., CAMPIONE, J. C. Y MURPHY, M. D. (1974) "Keeping track of changing variables: Long-term retention of a trained rehearsal strategy by retarded adolescents". *American Journal of Mental Deficiency, 78,* 446-53.

BROWN, A. L. Y DAY, J. D. (1983) "Macro-rules for summarizing texts: The development of expertise". *Journal of Verbal Learning and Verbal Behavior, 22,* 1-14.

BROWN, G. (1978) *Lecturing and Explaining.* Londres: Methuen.

BRUNER, J. S. (1962) *On Knowing - Essays for the left hand.* Cambridge, Mass.: Belknop Press.

BRYAN, W. L. Y HARTER, N. (1899) "Studies on the telegraphic language: The acquisition of a hierarchy of habits". *Psychological Reiew, 6,* 345-75.

BUTTERWORTH, B. (1980) "Evidence from pauses in speech". En: Butterworth, B. (Comp.). *Language Production,* Vol. 1: *Speech and Talk.* Londres: Academic Press.

CERMAK, K. S., BUTTERS, N. Y MOREINES, J. (1974) "Some analyses of the verbal encoding deficit of alcoholic Korsakoff pacients". *Brain and Language, 1,* 141-150.

CHASE, W. G. (1983) "Spatial representations of taxi-drivers". En: D. R. Rogers y J. A. Sloboda (Comps.). *Acquisition of Symbolic Skills.* Nueva York: Plenum.

CHASE, W. G. Y ERICSSON, K. A. (1981) "Skilled memory". En: *Cognitive Skills and their Acquisition.* Compilado por J. R. Anderson. Hillsdale, Nueva Jersey: Lawrence Erlbaum. Pp. 141-89.

CHASE, W. G. Y ERICSSON, K. A. (1982) "Skill and working memory". En: *The Psychology of Learning and Motivation,* Vol. 16. Compilado por G. H. Bower. Nueva York: Academic Press. Pp. 1-58.

CHASE, W. G. Y SIMON, H. A. (1973) "The mind's eye in chess". En: W. G. Chase (Comp.). *Visual Information Processing.* Nueva York:

CHI, M. T. H. (1978) "Knowledge structures and memory development". En: R. Seigler (Comp.). *Children's Thinking: What Develops?* Hillsdale: Lawrence Erlbaum.

CHI, M. T. H. (1983) *Trends in Memory Development Research.* Nueva York: Karger.

CHI, M. T. H. (1985) Changing conceptions of sources of memory development. *Human Development, 28,* 50-6.

CLARKE, A. M. Y CLARKE, A. D. B. (1976) *Early Experience: Myth and Evidence.* Londres: Open Books.

CLARKE, H. H. (1977) "Bridging". En: P. N. Johnson-Laird y P. C. Wason (Comps.). *Thinking: Readings in Cognitive Science.* Cambridge: Cambridge University Press.

COHEN, J. (1970) *Homo Psychologicus.* Londres: Allen & Unwin.

COLE, M., GAY, J. Y SHARP, D. (1971) *The Cultural Context of Learning and Thinking.* Nueva York: Basic Books.

COLE, M. Y SCRIBNER, S. (1974) *Culture and Thought.* Nueva York: Wiley.

CONDON, W. S. Y OGSTON, W. D. (1966) "Sound film analysis of normal and pathological behaviour patterns". *Journal of Nervous and Mental Disease, 143,* 338-47.

CONDON, W. S. Y OGSTON, W. D. (1967) "A segmentation of behaviour". *Journal of Psychiatric Research, 5,* 221-35.

CRAIK, F. I. M. Y LOCKHART, R. S. (1972) "Levels of processing: A framework for memory research". *Journal of Verbal Learning and Verbal Behaviour, 11,* 671-84.

CRAIK, F. I. M. Y TULVING, E. (1975) Depth of processing and retention of words in episodic memory. *Journal of Experimental Psychology: General, 104,* 268-94.

CROWDER, R. G. (1982) *The Psychology of Reading: an Introduction.* Nueva York: Oxford University Press.

CUCELOGLU, D. M. (1972) "Facial code in affective communication". En: B. C. Speer (Comp.). *Non-Verbal communication: Contemporary Science Issues, 10*. Beverly Hills: Sage Publications.

DANNER, F. W. (1976) "Children's understanding of intersentence organization in the recall of short descriptive passages". *Journal of Educational Psychology, 68*, 2, 174-83.

DANSEREAU, F. W. (1985) Learning strategy research. En: J. W. Segal, S. F. Chifman y R. Glaser (Comps.). *Thinking and Learning Skills*, Vol. 1, *Relating Instruction to Research*. Hillsdale, N. J.: Erlbaum.

DANSEREAU, D. F., BROOKS, L. W., HOLLEY, C. D. Y COLLINS, K. W. (1983) "Learning strategies training: effects of sequencing". *Journal of Experimental Education, 51*, 3, 102-08.

DANSEREAU, D. F., COLLINS, K. W., MCDONALD, B. A., GARLAND, J., DIEKHOFF, G. M. Y EVANS, S. H. (1979) "Development and evaluation of a learning strategy program". *Journal of Education Psychology, 71*, 64-73.

DARWIN, C. (1872) The Expression of Emotion in Man and Animals. Londres: J. Murray.

DARWIN, C. (1911) *The life and letters of Charles Darwin*. Vol. I. Nueva York: Appleton.

DAVIES, F. Y GREENE, T. (1984) *Reading for Learning in the Sciences*. Edinburgh: Oliver y Boyd.

DAVIES, G. (1981) "Face-recall systems". En: G. Davies, H. Ellis y J. Shepherd (Comps.). *Perceiving and Remembering Faces*. Londres: Academic Press.

DAVITZ, J. R. (1964) *The Communication of Emotional Meaning*. Nueva York: McGraw Hill.

DEFFENBACHER, K. A., BROWN, E. K. Y STURGILL, W. (1978) "Some predictors of eyewitness memory accuracy". En: M. M. Gruneberg, P. E. Morris y R. N. Sykes (Comps.). *Practical Aspects of memory*. Londres: Academic Press.

DEFFENBACHER, K. A. Y HORNEY, J. (1981) "Psycho-legal aspects of face identification". En: G. Davies, H. Ellis y J. Shepherd (Comps.). *Perceiving and Remembering Faces*. Londres: Academic Press.

DIEKHOFF, G. M., BROWN, P. J. Y DANSEREAU, D. F. (1982) "A prose-learning strategies training program based upon networking and depth-of-processing models". *Journal of Experimental Education, 50*, 4, 180-84.

DOOLING, D. J. Y LACHMAN, R. (1971) "Effects of comprehension on retention of prose". *Journal of Experimental Psychology, 88,* 216-22.

DOWNES, W. (1984) *Language and Society*. Londres: Fontana.

DRIVER, R. (1983) *The Pupil as Scientist?* Milton Keynes: Open University Press.

DUNCAN, S. (1972) "Some signals and rules for taking speaking turns in conversation". *Journal of Personality and Social Psychology, 23*, 283-92.

DUNCKER, K. (1945) "On problem solving". *Psychological Monographs, 58:5*, Whole Nro. 270.

EIBL-EIBESFELDT, I. (1972) "Similarities and differences between cultures in expressive movements". En: Hinde, R. A. (Comp.). *Non-verbal Communication*. Cambridge: Cambridge University Press.

EICH, J. E. (1980) "The cue-dependent nature of state-dependent retrieval". *Memory and Cognition, 8*, 157-73.

EKMAN, P. (1972) "Universal and cultural differences in facial expressions of emotion". En: J. Cole (Comp.). *Nebraska Symposium on Motivation*. Lincoln, Nebraska: University of Nebraska Press.

EKMAN, P. Y FRIESEN, W. V. (1969) Non-verbal leakage and cues to deception. *Psychiatry, 32*, 88-105.

EKMAN, P. Y FRIESEN, W. V. (1976) "Communicative Body Movement: American Emblems". *Semiótica, 15*, 4, 335-53.

EKMAN, P. Y FRIESEN, W. V. (1978) *Facial Action Coding System: investi-*

*gator's guide*. Palo Alto, California: Consulting Psychologists' Press.

EKMAN, P., FRIESEN, W. V. Y TOMKINS, S. S. (1971) "Facial Affect Scoring Technique: a first validity study". *Semiotica, 3*, 37-58.

ERICSSON, K. A. (1985) "Memory skill". *Canadian Journal of Psychology, 39*, 188-231.

EYSENCK, H. J. [illegible] York: John Wiley.

FAST, J. (1972) *Body Language*. Londres: Pan Books.

FISKE, S. T. Y TAYLOR, S. E. (1984) *Social Cognition*. Reading, Mass.: Addison-Wesley.

FITTS, P. M. Y POSNER, M. L. (1967) *Human Performance*. Belmont, California: Brooks/Cole.

FLAVELL, J. H. (1971) "What is memory development the development of?" *Human Development, 14*, 272-86.

FLAVELL, J. H. (1985) *Cognitive Development* (2da. edición). Englewood Cliffs: Prentice Hall (primera publicación 1977).

FLAVELL, J. H., BEACH, D. R. Y CHINSKY, J. M. (1966) "Spontaneous verbal rehearsal in a memory task as a function of age". *Child Development, 37*, 283-99.

FODOR, J. A., BEVER, T. G. Y GARRET, H. M. F. (1974) *The Psychology of Language*. Nueva York: McGraw Hill.

FOWLER, R., HART, J. Y SHEEHAN, M. (1972) "A prosthetic memory: an application of the prosthetic environment concept". *Rehabilitation Counselling Bulletin, 15*, 80-85.

FREUD, S. (1905) "Fragment of an analysis of a case of hysteria". *Collected Papers*, Vol. 3. N.Y. Basic Books, 1959.

FROMKIN, V. A. (1973) *Speech errors as linguistic evidence*. The Hague: Mouton.

GARDNER, H. (1982) *Developmental Psychology*. Nueva York: Little, Brown and Company.

GARDNER, H. (1943) *Frames of Mind*. Londres: Heinemann.

GARRETT, G. A., BAXTER, J. C. Y ROZELLE, R. M. (1981) "Training University Police in Black-American nonverbal behaviours: An application to police-community relations". *Journal of Social Psychology, 113*, 217-29.

GARRETT, M. F. (1980) "Levels of planning in sentence production". En: Butterworth, B. (Comp.). *Language Production*, Vol. 1: *Speech and Talk*. Londres: Academic Press.

GELLATLY, A. R. H. (1986) "Influences on conceptions of logic and mind". En: I. Hronszky, M. Feher y B. Dajka (Comps.). *Scientific Knowledge Socialized*. Dordrecht: Reidel.

GELLATLY, A. R. H. (1987) "The acquisition of a concept of logical necessity". *Human Development*. (Forthcoming.).

GELLATLY, A. R. H., JONES, S. Y BEST, A. (En prensa) "How good are young children at concentration?"

GHISELIN, B. (1952) *The Creative Process*. Berkeley: University of California Press.

GIBBS, G. (1981) *Teaching Students to Learn*. Milton Keynes: Open University Press.

GIBSON, E. J. Y LEVIN, H. (1975) *The Psychology of Reading*. Cambridge, Mass.: MIT Press.

GIBSON, E. J., PICK, A., OSSER, H. Y HAMMOND, M. (1962) "The role of grapheme-phoneme correspondence in the perception of words". *American Journal of Psychology, 75*, 554-70.

GILHOOLY, K. J. (1982) *Thinking: Directed, Undirected, and Creative*. Londres: Academic Press.

GLADWIN, T. (1970) *East is a Big Bird*. Cambridge, Mass.: Harvard University Press.

GOFFMAN, E. (1956) *The Presentation of Self in Everyday Life*. Edinburgh: University of Edinburgh Press.

GOODGLASS, H. Y KAPLAN, E. (1979) Assessment of cognitive deficiency in the brain-injured patient". En: M. Gazzaniga (Comp.). *Handbook of Behavioural Neurobiology, 2*. Nueva York: Plenum.

GORDON, W. J. J. (1961) *Synectics: the development of creative capacity*. Nueva York: Harper and Row.

GRANT, E. (1969) "Human Facial Expression". *Man, 4*, Nro. 4.

GREEN, D. W. Y SHALLICE, T. (1976) "Direct visual access in reading for meaning". *Memory and Cognition, 4*, 753-58.

GRICE, H. (1975) Logic and Conversation. En: P. Cole y J. L. Morgan (Comps.). *Syntax and Semantics*, Vol. 3: *Speech Acts*. Nueva York: Academic Press.

GRIGGS, R. A. (1983) "The role of problem content in the selection task and in the THOG problem". En: J. St B. T. Evans (Comp.). *Thinking and Reasoning*. Londres: Routledge & Kegan Paul.

GRUBER, H. E. (1974) *Darwin on Man: A psychological study of Scientific Creativity*. Nueva York: E. P. Dutton.

GRUNEBERG, M. M. (1985) "The Gruneberg-Linkword Language System for Teaching Foreign Languages". *Training and Development*, February, 1985.

GRUNEBERG, M. M., MORRIS, P. E. Y SYKES, R. N. (1978) (Comps.) *Practical Aspects of Memory*. Londres: Academic Press.

HABER, R. N. (1979) "Twenty years of haunting eidetic imagery: where's the ghost?" *The Behavioural and Brain Sciences, 2,* 583-629.

HARDYCK, C. D. Y PETRINOVICH, I. F. (1970) "Subvocal speech and comprehension level as a function of the difficulty level of the reading material". *Journal of Verbal Learning and Verbal Behaviour, 9,* 647-52.

HARRE, R. Y SECORD, P. W. (1972) *The Explanation of Social Behaviour.* Oxford: Blackwell.

HARRIS, J. E. (1980a) "Memory aids people use: two interview studies". *Memory and Cognition, 8,* 31-38.

HARRIS, J. E. (1980b) "We have ways of helping you remember". *Concord: Journal of the British Association for Service to the Elderly, 17 (Mayo),* 21-27.

HARRIS, J. E. (1984) "Remembering to do things". En: J. E. Harris y P. E. Morris (Comps.) *Everyday Memory, Actions and Absent-mindedness.* Londres: Academic Press.

HARTLEY, J. (1985) *Designing Instructional Text* (2da. edición). Londres: Kogan Page.

HASHER, L. Y ZACKS, R. T. (1979) "Automatic and effortul processes in memory". *Journal of Experimental Psychology: General, 108,* 356-88.

HATANO, G. Y OSAWA, K. (1983) "Digit memory of grand experts in abacus-derived mental calculation". *Cognition, 15,* 95-110.

HENLE, M. (1962) "On the relation between logic and thinking". *Psychological Review, 69,* 366-78.

HENLE, M. (1978) Foreword to R. Revlin and R. E. Mayer (Comps.) *Human Reasoning.* Washington, DC: Winston.

HERMANN, K. (1959) *Reading Disability.* Copenhagen: Munksgaard.

HINDE, R. A. (1966) *Animal Behaviour.* Nueva York: McGraw Hill.

HINDE, R. A. (1972) *Non-verbal communication.* Londres: Cambridge University Press.

HIRST, W. Y VOLPE, B. T. (1984) "Automatic and effortul encoding in amnesia". En: M. S. Gazzaniga (Comp.). *Handbook of Cognitive Neuroscience*, Nueva York: Plenum Press.

HITCH, G. J. (1980) "Developing the concept of working memory". En: G. Claxton (Comp.). *Cognitive Psychology: New Directions*. Londres: Routledge and Kegan Paul.

HOCHBERG, J. (1970) "Components of literacy: speculations and exploratory research". En: H. Levin y J. P. Williams (Comps.). *Basic Studies on Reading*. Nueva York: Basic Books.

HOFSTADTER, D. R. Y DENNETT, D. C. (1981) *The Mind's I*. Nueva York: Basic Books.

HOWE, N. J. A. (1984) "Questions of memory". En: J. Nicholson y M. Lucas (Comps.). *All in the Mind: Psychology in Action*. Londres: Thames Methuen.

HOWE, M. L., BRAINERD, C. J. Y JINGSMA, H. (1985). "Development of organization in recall: a stage of learning analysis". *Journal of Experimental Child Psychology, 39*, 230-51.

HULME, C., THOMSON, N., MUIR, C. Y LAWRENCE, A. (1984) "Speech rate and the development of short-term memory span". *Journal of Experimental Child Psychology, 38*, 241-53.

HUNT, M. (1982) *The Universe Within*. Hassocks, Sussex: Harvester.

HUNTER, I. M. L. (1964) *Memory*. Baltimore: Penguin.

HUNTER, I. M. L. (1985) "Memory development: the cultural background". En: J. A. Branthwaite y D. R. Rogers (Comps.). *Children Growing Up*. Milton Keynes: Open University Press.

ILLICH, I. D. (1973). *Deschooling Society*. Londres: Penguin.

INGHAM, R. J. (1972) "Cross cultural differences in social behaviour". D. Phil. Thesis, Oxford University.

ISTOMINA, Z. M. (1982) "The development of memory in children of pre-school age". En: U. Neisser (Comp.). *Memory Observed*. San Francisco: Freeman.

JAMES, D. E. (1967) *A Student's Guide to Efficient Study*. Londres: Pergamon.

JOHNSON, C. N. Y WELLMAN, H. M. (1980) "Children's developing understanding of mental verbs: remember, know and guess". *Child Development, 51*, 1.095-1.102.

JOHNSON-LAIRD, P. N. (1982) "Thinking as a skill". *Quarterly Journal of Experimental Psychology, 34A*, 1-30.

JOHNSON-LAIRD, P. N. (1983) *Mental Models*. Cambridge: Cambridge University Press.

JONASSEN, D. H. (1984) "Developing a learning strategy using pattern notes: a new technology". *Programmed Learning and Educational Technology, 21*, 3, 163-75.

JONES, G. Y ADAM, J. (1979) "Towards a prosthetic memory". *Bulletin of the British Psychological Society, 32*, 165-67.

JORM, A. F. (1983) *The Psychology of Reading and Spelling Difficulties*. Londres: Routledge and Kegan Paul.

JOYNSON, R. B. (1974) *Psychology and Common Sense*. Londres: Routledge & Kegan Paul.

KAGAN, J., LAPIDUS, D. R. Y MOORE, N. (1978) "Infant antecedents of cognitive functioning: a longitudinal study". *Child Development, 49*, 1.005-23.

KAIL, R. (1984) *The Development of Memory in Children* (2da. edición). Nueva York: Freeman.

KANDEL, E. R. (1976) *Cellular Basis of Behaviour*. Nueva York: Freeman.

KATONA, G. (1940) *Organizing and Memorizing*. Nueva York: Columbia University Press.

KATZ, A. N. (1978) "Creativity and the right cerebral hemisphere". *Journal of Creative Behaviour, 12,* 253-64.

KEARINS, J. (1981) "Visual spatial memory in Australian Aboriginal Children of desert regions". *Cognitive Psychology, 13,* 435-60.

KENDLER, H. H. Y KENDLER, T. S. (1967) "Experimental analysis of inferential behaviour in children". En: L. P. Lipsitt y C. C. Spiker (Comps.). *Advances in Child Development and Behaviour,* Vol. 3. Nueva York: Academic Press.

KENDON, A. (1967) "Some functions of gaze patterns in social interactions". *Acta Psychologica,* 26, 24-47.

KLEIMAN, G. M. (1975) "Speech recoding in reading". *Journal of Verbal Learning and Verbal Behaviour, 14,* 323-29.

KOESTLER, A. (1964) *The Act of Creation.* Nueva York: Dell.

KOBASIGAWA (1974) "Utilization of retrieval-cues by children in recall". *Child Development, 45,* 127-34.

KOLB, B. Y WHISHAW, I. Q. (1985) *The Fundamentals of Human Neuropsychology* (2da. edición). Nueva York: W. H. Freeman and Company.

KREUTZER, M. A., LEONARD, C. Y FLAVELL, J. H. (1975) "An interview study of children's knowledge about memory". *Memographs of the Society for Research in Child Development, 40* (1, Serial Nro. 159), 1-58.

KUHN, T. S. (1970) *The Structure of Scientific Revolutions* (2da. edición). Chicago: University of Chicago Press.

LABERGE, D. (1973) "Attention and the measurements of perceptual learning". *Memory and Cognition, 1,* 268-76.

LANZETTA, J. T. Y KLECK, R. (1970) "Encoding and decoding facial affect in humans". *Journal Personality and Social Psychology, 16,* 12-19.

LARKIN, J. H., MCDERMOTT, J., SIMON, D. P. Y SIMON, H. A. (1980) "Expert and novice performance in solving physics problems". *Science, 208,* 1.335-42.

LEVELT, W. J. M. (1981) "The speaker's linearization problem". En: *The Psychological Mechanisms of Language*. Londres: Royal Society and British Academy.

LINDE, C. Y LABOV, W. (1975) "Spatial networks as a site for the study of language and thought". *Language, 51*, 924-39.

LINDSAY, P. H. Y NORMAN, D. A. (1977) *Human Information Processing* (2da. edición). Nueva York: Academic Press.

LOFTUS, E. F. Y PALMER, J. C. (1974) "Reconstruction of automobile destruction: an example of the interaction between language and memory". *Journal of Verbal Learning and Verbal Behaviour, 13*, 585-89.

LORAYNE, H. Y LUCAS, J. (1974) *The Memory Book*. Nueva York: Stein and Day.

LORD, A. B. (1982) "Oral poetry in Yugoslavia". En: U. Neisser (Comp.). *Memory Observed*. San Francisco: Freeman.

LUCHINS, A. S. Y LUCHINS, E. H. (1950): "New experimental attempts at preventing mechanization in problem solving". *Journal of General Psychology, 42*, 279-97.

LUNN, J. H. (1948) "Chicken Sexing". *American Scientist, 36*, 280-87.

LURIA, A. R. (1969) *The Mind of the Mnemonist*. Londres: Cape.

LURIA, A. R. (1973) *The Working Brain*. Harmondsworth: Penguin.

LURIA, A. R. (1976) *Cognitive Development: Its Cultural and Social Foundations*. Cambridge, Mass.: Harvard University Press.

MAIER, N. R. F. (1931) "Reasoning in Humans II: The solutions of a problem and its appearance in consciousness". *Journal of Comparative Psychology, 12*, 181-94.

MARR, D. (1983) *Vision*. San Francisco: Freeman.

MARSLEN-WILSON, W. D. Y TYLER, L. J. (1981) "Central processes in speech understanding". En: *Psychological Mechanisms of Language*. Londres: Royal Society and British Academy.

MARTON, F., HOUNSELL, D. Y ENTWISTLE, N. (E) (1984) *The Experience of Learning*. Edinburgh: Scottish Academic Press.

MAYER, R. E. (1943) *Thinking, Problem Solving, Cognition*. San Francisco: Freeman.

MAYES, A. R. Y MEUDELL, P. R. (1983) "Amnesia in man and other animals". En: A. R. Mayes (Comp.). *Memory in Animals and Humans*. Wokingham: Van Nostrand Reinhold.

MCGREW, W. C. (1972) *An Ethological Study of Child Behaviour*. Cambridge: Cambridge University Press.

MEACHAM, J. (1972) "The development of memory abilities in the individual and society". *Human Development, 15*, 205-28.

MEHRABIAN, A. (1970) "Non-verbal communication". Chicago: Aldine Atherton.

MEYER, B. H. F. (1985) "Signaling the structure of text". En: D. H. Jonassen (Comp.). *The Technology of Text*, Vol. 2. Englewood Cliffs, N.J.: Educational Technology Publications.

MILLER, G. A. (1956) "The magical number seven, plus or minus two: some limits on our capacity for processing information". *Psychological Review, 63*, 81-96.

MOELY, B. E., OLSON, F. A., HALWES, T. G. Y FLAVELL, J. H. (1969) "Production deficiency in young children's clustered recall". *Developmental Psychology, I*, 26-34.

MORRIS, C. D., BRANSFORD, J. D. Y FRANKS, J. J. (1977) "Levels of processing versus transfer appropriate processing". *Journal of Verbal Learning and Verbal Behaviour, 16*, 519-33.

MORRIS, D. (1977) *Manwatching: A field guide to human behaviour*. Londres: Book Club Associates.

MORRIS, P. E., GRUNEBERG, M. M., SYKES, R. N. Y MERRICK, A. (1981) "Football knowledge and the acquisition of new results". *British Journal of Psychology, 72*, 479-84.

MORTON, H. (1964) "The effects of context on the visual duration threshold for words". *British Journal of Psychology, 55*, 165-80.

NAIDOO, S. (1970) "Remedial re-education". En: A. W. Franklin y S. Naidoo (Comps.). *Assessment and teaching of dyslexic children.* Londres: Invalid Childrens' Aid Association.

NEISSER, U. (1976) *Cognition and Reality.* San Francisco: Freeman.

NEWELL, A. Y ROSENBLOOM, P. S. (1981) "Mechanisms of skill acquisition and the law of practise". En: J. R. Anderson (Comp.) Cognitive Skills and their Acquisition. Hillsdale NJ: Erlbaum.

NEWELL, A. Y SIMON, H. (1972) Human Problem Solving. Englewood Cliffs, N.J.: Prentice Hall.

NEWON, J. (1978) *Dialogue and Development Action, Symbol and Gesture, the emergence of language.* J. Newson (Comp.) Londres: Academic Press.

NORMAN, D. A. (1976) *Memory and Attention* (2da. edición). Nueva York: Wiley.

NORMAN, D. A. (Comp.) (1981) Perspectives on Cognitive Science. Nueva York: Ablex.

O'REAGAN, K. (1979) "Moment to moment control of eye saccades as a function of textual parameters in reading". En: P. A. Kolers, M. E. Wrolstad y H. Bouma (Comps). *Processing of Visible Language*, Vol. 1. Nueva York: Plenum.

OPIE, I. Y OPIE, P. (1959) *The Lore and Language of schollchildren.* Londres: Oxford University Press.

ORNSTEIN, P. A., NAUS, M. J. Y STONE, B. P. (1977) "Rehearsal training and developmental differences in memory". *Developmental Psychology, 13*, 15-24.

PAPERT, S. (1972) "Teaching children thinking". *Mathematical Teaching, 58*, Spring, 1972.

PARIS, S. G., CROSS, D. R. Y LIPSON, M. Y. (1984) "Informed strategies for learning: a program to improve children's reading awareness and comprehension". *Journal of Educational Psychology, 51*, 3, 131-33.

PERFETTI, C. A., GOLDMAN, S. A. Y HOGABOAM, T. W. (1979) "Reading skill and the identification of words in connected discourse". *Memory and Cognition, 7*, 273-82.

PERKINS, D. N. (1981) *The Minds's Best Work*. Cambridge, Mass.: Harvard University Press.

PERRY, W. G. (1981) "Cognitive and ethical growth: the making of meaning". En: A. Chickering (Comp.). *The Modern American College*. San Francisco: Jossey Bass.

PIAGET, J. (1926) *The Language and Thought of the Child*. Londres: Routledge and Kegan Paul.

PIAGET, J. (1974) "Need and significance of cross-cultural studies in genetic psychology". En: J. W. Berry & R. R. Dasen (Comps.). *Culture and Cognition: Readings in Cross-Cultural Psychology*. Londres: Methuen.

PIAGET, J. Y INHELDER, B. (1973) *Memory and Intelligence*. Nueva York: Basic Books.

POWELL, G. E. (1981) *Brain Function Therapy*. Aldershot: Gower Press.

REASON, J. Y MYCIELSKA, K. (1982) *Absent Minded? The Psychology of Mental Lapses and Eeryday Errors*. Nueva York: Prentice Hall.

REVESZ, G. (1925) *The Psychology of a Musical Prodigy*. Londres: Kegan Paul, Trench and Trubner.

ROGERS, D. R. (1985) "Infants, mothers and intentional communication". En: J. A. Branthwaite y D. R. Rogers (Comps.). Children Growing Up. Milton Keynes: Open University Press.

ROSENTHAL, R. (Comp.) (1979) Skill in Non-verbal Communication: Inidividual Differences. Cambridge, Mass.: Oelgeschlager, Gunn and Main.

ROVEE-COLLIER, C. K. Y LIPSITT, L. P. (1982) "Learning, adaptation and memory in the newborn". En: P. Stratton (Comp.). *Psychobiology of the Human Newborn*. Chichester: Wiley.

RUESCH, J. Y KEES, W. (1956) *Non-verbal Communication: Notes on the visual perception of human relations*. Berkeley: University of California Press.

RUMELHART, D. E. Y NORMAN, D. A. (1981) "Analogical processes in learning". En: J. R. Anderson (Comp.). *Cognitive Skills and their Acquisition*. Hillsdale, N.J.: Erlbaum.

SANFORD, A. J. Y GARROD, S. C. (1981) *Understanding Written Language: Explorations in Comprehension Beyond the Sentence*. Chichester: Wiley.

SCHANK, R. C. Y ABELSON, R. (1977) *Scripts, Plans, Goals and Understanding*. Hillsdale N.J.: Erlbaum.

SCHEFREN, A. E. (1964) "The significance of posture in communication systems". *Psychiatry, 27*, 316-31.

SCHON, D. A. (1979) "Generative metaphor: a perspective on problem-setting in social policy". En: A. Ortony (Comp.). *Metaphor and Thought*. Cambridge: Cambridge University Press.

SCRIBNER, S. (1977) "Modes of thinking and ways of speaking: culture and logic reconsidered". En: P. N. Johnson-Laird y P. C. Wason (Comps.). *Thinking: Readings in Cognitive Science*. Londres: Cambridge University Press.

SEARLE, J. R. (1975) "Indirect Speech Acts". En: Cole, P. y Morgan, J. L. (Comps.). *Syntax and Semantics*, Vol. 3: *Speech Acts*. Nueva York: Academic Press.

SEARLE, J. R. (1980) "Minds, brains, and programs". *The Behavioural and Brain Sciences, 3*, 417-457.

SHAFFER, L. H. (1976) "Intention and performance". *Psychological Review, 83*, 375-93.

SHEPHERD, G. (1983) "Social-skills training with adults". En: S. Spence y G. Sheperd (Comps.). *Development in Social Skills Training*. Londres: Academic Press.

SIEBEL, R. (1963) "Discrimination reaction time for a 1023 alternative task". *Journal of Experimental Psychology*, 66, 1.005-23.

SIEGLER, R. S. (1986) "Strategy choices in substraction". En: J. A. Sloboda y D. R. Rogers (Comps.). *Cognitive Processes in Mathematics*. Londres: Oxford University Press.

SIQUELAND, E. R. Y LIPSITT, L. P. (1966) "Conditioned head-turning in human newborns". *Journal of Experimental Child Psychology*, 3, 357-76.

SLOBODA, J. A., HERMELIN, B. Y O'CONNOR, N. (1985) "An exceptional musical memory". *Music Perception*, en prensa.

SMITH, H. K. (1967) "The responses of good and poor readers when asked to read for different purposes". *Reading Research Quarterly*, 3, 53-83.

SMITH, S. M. (1979) "Remembering in and out of context". *Journal of Experimental psychology: Human Learning and Memory*, 5, 460-71.

SMODE, A. (1958) "Learning and performance in a tracking task under two levels of achievement information feedback". *Journal of Experimental Psychology*, 56, 297-304.

SPITZ, H. H. (1966) "The role of input information in the learning and memory of mental retardates". En: N. R. Ellis (Comp.). *International Review of Research in Mental Retardation*, Vol. 2. Nueva York: Academic Press.

SPRINGER, S. P. Y DEUTSCH, G. (1985) *Left Brain, Right Brain* (2da. edición). San Francisco: Freeman.

SQUIRE, L. R. Y COHEN, H. J. (1984) "Human memory and amnesia". En: G. Lynch, J. L. McGaugh y N. M. Weinberger (Comps.). *Neurobiology of Learning and Memory*. Nueva York: The Guildford Press.

STEIN, M. I. (1975) *Stimulating Creativity*. Vol. II. New York: Academic Press.

STOPPARD, T. (1967) *Rosencrantz and Guildenstern Are Dead*. Londres: Faber. Reimpreso en 1978.

STUBBS, M. (1983) *Discourse Analysis*. Oxford: Blackwell.

TERRACE, H. S., PETTITO, L. A., SANDERS, R. J. Y BEVER, T. G. (1980) "On the grammatical capacity of apes". En: Nelson, K. (Comp.). *Children's Language*, Vol. 2. Nueva York: Gardner.

THOMAS, A. Y CHESS, S. (1977) *Temperament and Development*. Nueva York: Brunner.

TINBERGEN, N. (1974) "Ethology and stress disease". *Science, 185*, 20-27.

TINKER, M. A. (1965) *Bases for Effective Reading*. Minneapolis: University of Minnesota Press.

TORRENCE, E. P. (1982) "Hemisphericity and creative functioning". *Journal of Research and Development in Education, 15*, 29-37.

TRUEMAN, M. (1985) "The different abilities of the two halves of the brain". En: J. A. Branthwaite y D. R. Rogers (Comps.). *Children Growing Up*. Milton Keynes: Open University Press.

TULVING, E. (1962) "Subjective organization in free recall of 'unrelated' words". *Psychological Bulletin, 69*, 344-54.

TULVING, E. (1972) "Episodic and Semantic memory". En: E. Tulving y W. Donaldson (Comps.). *Organization of memory*. Londres: Academic Press.

TULVING, E. (1984) Precis of "Elements of episodic memory". *The Behavioural and Brain Sciences, 7*, 223-268.

TURING, A. M. (1950) "Computing machinery and intelligence". *Mind, 59*, 433-60.

TVERSKY, A. Y KAHNEMAN, D. (1973) "A heuristic for judging frequency and probability". *Cognitive Psychology, 5*, 207-32.

TVERSKY, A. Y KAHNEMAN, D. (1974) "Judgement under uncertainty: Heuristics and biases". *Science, 185*, 1.124-31.

TVERSKY, A. Y KAHNEMAN, D. (1982) "Judgements of and by representativeness". En: D. Kahneman, P. Slovic y A. Tversky (Comps.). *Judgement under uncertainty: Heuristic and biases.* Nueva York: Cambridge University Press.

WALLAS, G. (1926) *The Art of Thought.* Nueva York: Harcourt Brace.

WASON, P. C. (1966) "Reasoning". En: B. Foss (Comp.). *New Horizons in Psychology, 1.* Harmondsworth: Penguin.

WASON, P. C. (1983) "Realism and rationality in the selection task". En: J. St B. T. Evans (Comp.). *Thinking and Reasoning.* Londres: Routledge and Kegan Paul.

WEINSTEIN, C. E. Y MAYER, R. E. (1985) "The teaching of learning strategies". En: M. C. Wittrock (Comp.). *Handbook of Research on Teaching.* Nueva York: Macmillan.

WEIZENBAUM, J. (1984) *Computer Power and Human Reason* (2da. edición). Harmondsworth: Penguin.

WELFORD, A. T. (1968) *Fundamentals of Skill.* Londres: Methuen.

WHITING, C. S. (1958) *Creative thinking.* Nueva York: Reinhold.
Wilkinson, A., Stratton, L. y Dudley, P. (1974) *The Quality of Listening.* Londres: Macmillan.

WILLIAMS, J. P. (1970) "From Basic Research on Reading to Educational Practice". En: H. Levin y J. P. Williams (Comps.). *Basic Studies on Reading.* Nueva York: Basic Books.

WILSON, B. Y MOFFAT, N. (1984) "Rehabilitation of Memory for Everyday Life". En: J. E. Harris y P. E. Morris (Comps.). *Everyday Memory, Actions and Absent-Mindedness.* Londres: Academic Press.

WINOCUR, G., KINSBOURNE, M. Y MOSCOVITCH, M. (1981) "The effect of cuing on release from productive interference in Korsakoff amnesic patients". *Journal of Experimental Psychology: Human Learning and Memory, 1*, 56-65.

WITTROCK, M. C., MARKS, C. Y DOCTOROW, M. (1975) "Reading as a generative process". *Journal of Educational Psychology, 67*, 4, 484-89.

WOLF, T. (1976) "A cognitive model of musical sight-reading". *Journal of Psycholinguistic Research, 5*, 143-71.

WONG, B. Y. Z. Y JONES, W. (1982) "Increasing metacomprehension in learning-disabled and normally-achieving students through self-questioning training". *Learning Disability Quarterly, 5*, 228-40.

WOOD, G. (1983) *Cognitive Psychology: A Skills Approach*. Monterey: Brooks Cole.

YATES, F. A. (1966) *The Art of Memory*. Chicago: University of Chicago Press.

ZIMMERMAN, D. H. Y WEST, C. (1975) "Sex roles, interruptions and silences in conversation". En: B. Thorne y N. Henley (Comps.). *Language and Sex: Difference and Dominance*. Rowley, Mass.: Newbury House.

Esta edición de 2.000 ejemplares
se terminó de imprimir
en febrero de 1997 en Indugraf S.A.
Sánchez de Loria 2251, Buenos Aires.

Estrategias de aula

Esta obra se estructura en torno al tema de las habilidades humanas, analizando sus naturalezas, las formas en que se adquieren y las técnicas que pueden contribuir a mejorarlas.

El acento está puesto fundamentalmente en habilidades cognitivas comunes como *hablar*, *recordar* y *razonar*, que suelen darse por sentadas y pueden parecer irrelevantes; pero que, como demuestra este libro, son tan extraordinarias y complejas como las habilidades no comunes.

**La inteligencia hábil** sostiene que la cognición se entiende mejor en términos de *habilidad*; que las habilidades deben ser analizadas en su contexto social, y que pueden mejorarse por medio del entrenamiento.

Angus Gellatly es profesor del Departamento de Psicología de la Universidad de Keele (U.S.A.). Las investigaciones sobre la memoria, el razonamiento y la percepción que presenta en este libro constituyen un aporte significativo a la búsqueda de estrategias para desarrollar y mejorar las habilidades de aprendizaje.

ISBN 950-701-377-6

9 789507 013775

AIQUE

Cód: 1377

JOHN A. SLOBODA

# The Musical Mind

## The Cognitive Psychology of Music



OXFORD SCIENCE PUBLICATIONS

# The Musical Mind

*The cognitive psychology of music*

JOHN A. SLOBODA

*Department of Psychology, University of Keele*

OXFORD PSYCHOLOGY SERIES NO. 5

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP
Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in
Oxford New York
Athens Auckland Bangkok Bogotá Buenos Aires Calcutta
Cape Town Chennai Dar es Salaam Delhi Florence Hong Kong Istanbul
Karachi Kuala Lumpur Madrid Melbourne Mexico City Mumbai
Nairobi Paris São Paulo Singapore Taipei Tokyo Toronto Warsaw
with associated companies in Berlin Ibadan
Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries
Published in the United States
by Oxford University Press Inc., New York
© John A. Sloboda, 1985
The moral rights of the author have been asserted
Database right Oxford University Press (maker)
First published 1985
Reprinted in paperback 1986
Reprinted 1987 (with corrections) 1988, 1989,1990, 1991, 1993, 1994,
1996, 1997, 1999 (with corrections), 2000 (twice)
All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above
You must not circulate this book in any other binding or cover
and you must impose this same condition on any acquirer
A catalogue record for this book is available from the British Library
Library of Congress Cataloging in Publication Data
(Data available)
ISBN 0 19 852128 6 (pbk)
Printed in Great Britain
on acid-free paper by
Bookcraft (Bath) Ltd., Midsomer Norton




Editors:

James L. McGaugh
Nicholas J. Mackintosh
Anne Treisman
Endel Tulving
Lawrence Weiskrantz

## OXFORD PSYCHOLOGY SERIES

1. The neuropsychology of anxiety: an inquiry into the functions of the septo-hippocampal system *Jeffrey A. Gray*
2. Elements of episodic memory *Endel Tulving*
3. Conditioning and associative learning *N. J. Mackintosh*
4. Visual masking: an integrative approach *Bruno G. Breitmeyer*
5. The musical mind: the cognitive psychology of music *John A. Sloboda*
6. Elements of psychophysical theory *Jean-Claude Falmagne*
7. Animal intelligence: *Edited by L. Weiskrantz* Proceedings of a Royal Society Discussion Meeting
8. Response times: their role in inferring elementary mental organization *R. Duncan Luce*
9. Mental representations: a dual coding approach *Allan Paivio*
10. Memory, imprinting, and the brain: an enquiry into mechanisms *Gabriel Horn*
11. Working memory *Alan Baddeley*
12. Blindsight: a case study and implications *L. Weiskrantz*
13. Profile analysis *David M. Green*
14. Spatial vision *Russell L. DeValois and Karen K. DeValois*
15. The neural and behavioural organization of goal-directed movements *Marc Jeannerod*
16. Visual pattern analyzers *Norma V. S. Graham*
17. Cognitive foundations of musical pitch *Carol L. Krumhansl*
18. Perceptual and associative learning *Geoffrey Hall*
19. Implicit learning and tacit knowledge *A. S. Reber*
20. Neuropsychology of communicative behavior *D. Kimura*
21. The frontal lobes and voluntary action *R. E. Passingham*

# Preface

When it was first suggested that I should write a book on the cognitive psychology of music, I had no clear conception of what sort of book might be the outcome. As I began to take stock of the existing literature, it became clear that there was a gap that needed to be filled. On one side of this gap could be placed the vast majority of the psychological research on music. On the other side was the experience and insight of the musician, as expressed both through his practices and through the rigorous disciplines of musicology, music theory, and analysis. It seemed, to me at least, that the psychology of music related rather little to what musicians actually did, and so was failing to tackle questions of central musical importance.

There are several reasons why this state of affairs might have arisen. One is that many psychologists studying music have not had the good fortune to receive extended musical training, and so have had a limited range of musical insights and intuitions to guide their work. The difficulties caused by this are akin to the difficulties that might be experienced by someone attempting to carry out psycholinguistic research on a language that he did not know at all well, and without a thorough grounding in linguistics. Even psychologists *with* musical expertise have often seemed to detach their scientific work from their musical knowledge.

Another reason is that theoretical developments in the psychology of music have been slow. Experimentalists have tended to construct micro-theories to account for their own results so that research from different sources cannot easily be synthesized. Only in the last few years have people well-grounded in both psychology and music begun to construct psychologically interesting theories of musical functioning which might be capable of unifying and giving direction to a wide range of research endeavours.

A third reason is that psychological research has been dominated by the view that one must understand the most peripheral and simple aspects of intellectual functioning as a prelude to the study of more central and complex aspects. Although this view is now accepted to be largely misconceived, a vast amount of so-called 'music' research has concentrated on the processes involved in the perception of single tones. Topics such as the representation of large-scale musical structure, performance, and composition, have received comparatively little attention. Yet these are the topics which are of most direct relevance to musicians.

A fourth and related reason is that psychologists have, rightly, wished to conduct their research with rigorous control and measurement. It is much

easier to construct brief auditory stimuli to exact specifications whilst constraining subjects to simple yes/no responses, than it is to to allow subjects the freedom to indulge in some complex but musically interesting behaviour in response to complex musically structured conditions. Yet the latter *can* be done with rigour, and is beginning to be done with increasing frequency.

Fifth, and finally, writers on the psychology of music have tended to address themselves exclusively either to the professional psychologist or to the music educator and educational researcher. There has been little fruitful dialogue between psychologists and practising or academic musicians. This may not be entirely the fault of psychologists, but few psychologists seem to have been really concerned to break the interdisciplinary barrier.

I have written with this barrier very much in mind, and have tried to offer something to both the psychologist and the musician. The psychologist will, I hope, find a coherent account of the important research carried out in this area. He will also be exposed to some of the questions and issues which have not yet stimulated a great deal of research. These questions arise as I critically examine the field through the eyes of a musician, but also through the eyes of a cognitive psychologist who finds psychologists of music slow to accept the wider perspectives which cognitive science has brought to other branches of psychology in the last decade. I have assumed a moderate degree of musical literacy in my readers. To have done otherwise would have made my task impossibly slow. Likewise, an acquaintance is assumed with the broad range of topics and methods embraced by scientific psychology, although I have attempted to write in such a way that detailed knowledge in specific areas is not required. I have made liberal reference to suitable introductory reading where the general reader might wish to seek further information.

Musicians will, I hope, find that cognitive psychology has something to offer in helping them understand the mental bases of their skills. I have not, however, set out to provide a 'recipe book' of psychological aids to specific musical problems. Rather, I have tried to highlight and elucidate some basic characteristics of musical skills and the cognitive mechanisms which serve them. An understanding of these underlying factors will, hopefully, allow musicians to formulate better solutions to some of their problems.

I am very much aware that the subject I present abounds with loose ends and unsolved problems. I have not tried to give the field more substantiality than it has. What I hope to have done is to describe the main achievements in this rapidly growing area up to the end of 1982 or thereabouts, and map out some important topics and questions which need more attention than they have received.

My own approach to the cognitive psychology of music has been most greatly influenced by the writings of Christopher Longuet-Higgins who, in the early 1970s, was asking and answering important questions about the psychological representation of extended tonal music at a time when many

others were hooked on peripheral questions. I gratefully acknowledge his encouragement and support, which has been directly instrumental in getting this book finished. I am also grateful to several people for putting me on the track of material that I otherwise would have discovered much later or not at all. Alan Mosley first told me about the work of Heinrich Schenker; Michael Durand first showed me the fascinating story revealed by ink changes in Mozart's operatic autographs; and Ian Hunter introduced me to David Sudnow's unique account of his own improvisatory development, as well as opening my eyes to the profound contrasts between skills of literate and non-literate people.

Several people deserve special mention in connection with the immediate preparation of this book. I am grateful to Stephen Banfield, Ian Hunter, Pamela Liebeck, George Pratt, Henry Shaffer, Ursula Sharma, my wife Judith, and anonymous reviewers for the Oxford University Press, for reading portions of the first draft and offering useful comments. My late father also made detailed comments on the earlier portion of the book which much improved its intelligibility, and I dedicate this book to his memory. I am particularly grateful to the Senate of the University of Keele for the award of a Keele Research Fellowship in 1982 which allowed me the uninterrupted time required to complete the first draft. I gratefully acknowledge the secretarial help of Dorothy Masters and Margaret Woodward, and technical assistance from John Coleman and the Keele University Computer Centre.

The psychology of music is a rapidly growing subject, and it is exciting to be directly involved in that growth. I have enjoyed writing this book, but I am also aware that the field has leapt forward even in the two years I have taken out of the laboratory to complete it. I am now anxious to return to the laboratory. My principal hope is that psychologists who enjoy reading this book will also be stimulated to repair to the laboratory and render some of my critical comments obsolete.

J. A. S.

*July 1983*

# Preface to the 1999 reprinting

The text of *The Musical Mind* was completed in 1983. Fifteen years have now passed since then, and a new printing makes it timely for me to say something about the progress of music psychology since 1983. In effect, I need to justify why any reader should turn to this book, given the substantial progress in the field since that time. Fifteen years is a long time in science, and anyone wishing to enter the field of music research as an active participant cannot do so without concentrating their attention on materials published in the last five years or so. However, it seems that *The Musical Mind* has served a variety of constituencies outside the research laboratory. Musicians, students, and interested lay-people have been kind enough to say that they have found something of more lasting value in the book. This pleases me, for it was never my intention to provide a comprehensive textbook, of the kind that could be straightforwardly updated by the addition of new sections every few years. I wanted, as I hope I expressed in the Preface to the first printing, to address the field from a critical perspective (in which I saw psychologists as insufficiently interested in making their research relevant to the interests and concerns of musicians) and highlight areas of investigation which I felt to be important, regardless of the quantity of research that had been published in those areas.

This Preface is not in any sense, an exhaustive summary of where the cognitive psychology of music has got to since 1893. That would require another book. Rather, it is a brief evaluation of the development of some of the key themes that I identified. It therefore is probably best read *after* consulting the relevant chapters, rather than before, and I have therefore referenced all my comments to the numbered sections and page numbers of the chapters. I have tried to be selective in my citation of more recent research, favouring books and review articles over specific empirical studies unless these latter are of foundational significance. In this way, interested readers will gain the fastest access to recent surveys of specific areas of research.

### Music, language, and meaning

My organizing principle for discussing links between language and music was the tripartite distinction between phonology, syntax, and semantics (2.3(h) ff. pp. 22–65). In so far as music could be mapped on to this distinction, it seemed appropriate to think of music and language as overlapping

(both functionally and conceptually). In 1983 my broad conclusion was that evidence for equivalence at the phonological and syntactic level was rather impressive, but that musical and linguistic semantics operated in rather different ways, in so far as the scanty research literature in music allowed any definite conclusions at all.

About phonology (2.4, pp. 23–32) I wish to add rather little. The ability to make categorical distinctions between musical elements is a foundational aspect of music perception and performance. My treatment concentrated on pitch and duration. More recent research has widened the range of dimensions about which we have knowledge. Of these, timbre is undoubtedly the most significant (Hajda, Kendall, Carterette, and Harchberger 1997).

The study of psychological aspects of musical syntax (2.5, pp. 32–57) has seen phenomenal growth since 1983. To me, the most interesting feature of this growth is that it has been largely driven organically by developments within experimental psychology, rather than developments within musicology. Our understanding of pitch relations has deepened largely through the application of traditional psychophysical techniques to long-established cultural conceptualizations of tonality (e.g. Krumhansl 1990) rather than in response to new theory, even when that theory is explicitly psychologically oriented (e.g. Lerdahl and Jackendoff 1983: Narmour 1989, 1992). Earlier (Sloboda 1986) I predicted that Lerdahl and Jackendoff's *Generative Theory of Tonal Music* (1983) would give the kind of new impetus to music psychology research that Chomsky (1965) gave to psycholinguistics. With a few brave exceptions (e.g. Deliege 1987) I have been proved entirely wrong about that. More fruitful growth points have been at the junctions between psychology and neuroscience, partly as a result of the rise of connectionism and neural-net modelling (see Bharucha 1991), but also as a result of a range of new techniques which allow direct measurement of the overlap of brain activity between language and music processing (e.g. Besson 1997; Patel and Peretz 1997). One possible reason for the slowness for psychology to take on new musicological theorizing is that such work attempts to grapple with sytactical issues at the global as well as the local level (i.e. pertaining to entire movements and not just single phrases). Our methods for studying global cognition are still in their infancy, and psychology is still best suited to studying our interactions with materials lasting seconds rather than minutes.

Interest in musical semantics (2.6, pp. 57–65) has increased considerably since 1983. A particularly acute philosophical contribution has been made by Raffman (1993) who pointed out that the phrase 'music has meaning'. can mean both 'music has semantics', and 'music has personal significance' (as in 'music means a lot to me'). Much conceptual confusion has been created by previous unclarity on this point. The psychology of musical semantics is concerned with the mapping and explanation of relationships,

between musical elements and their non-musical connotations, which are often emotional (e.g. Juslin 1997), but by no mans always need be so (e.g. Watt and Ash 1998). Over and above semantics, music can be used to achieve valued psychological states, including emotional states, and an explanation of these effects requires an understanding, not only of cognitive psychology, but also social psychology (Hargreaves and North 1997), therapy (Bunt 1994), and the psychology of motivation (Sloboda and Davidson 1995). Cognition remains central to this aspect, however, since there is now some evidence to support Meyer's (1973) contention (2.6, pp. 64–5) that violations of structural expectancy (describable in syntactic terms) contribute in a very direct way to the intensity of experienced emotion, if not its content (Krumhansl 1998; Sloboda 1991).

**Music performance**

My treatment of performance (3) was also based round three issues, sight-reading, rehearsal, and expert performance. A great deal of new work has been done in the latter two areas, almost nothing in the first.

The emphasis on sight-reading in 1983 (3.2, pp. 68–90) largely reflected my own research interests, in which reading was a dominant theme at that time. Despite the fact that notational competence remains central to much musical skill and development, there is little more known about it than in 1983. Waters and Underwood (1998) report a study on eye movements which summarizes the small recent literature.

The study of rehearsal (3.3, pp. 90–3) has seen a recent renaissance as a result of the rise of a branch of cognitive psychology known as 'expertise studies' (see Ericsson and Smith 1991). This area was foreshadowed by seminal work on problem-solving (1 p. 8) begun by Simon and colleagues, and has led to detailed investigations of the conditions under which skilled individuals improve their skills. Not only do we understand much more about the details of effective musical practice (e.g. Ericsson, Krampe, and Tesch-Romer 1993; Jorgensen and Lehmann 1997), but we have a much clearer idea of the social, motivational, and emotional factors which allow unusually high levels of musical achievement to develop (Manturszewska 1990; Davidson, Howe, and Sloboda 1997). This work is entirely in keeping with the 'demystification' of talent and genius, which characterized my approach in 1983 (see 1 pp, 3–6), and which has been a central theme in the cognitive psychology of creativity and achievement ever since (e.g. Howe, Davidson, and Sloboda 1998; Weisberg 1993)

The section on expert musical performance was, in 1983, largely conceptual (3.4, pp. 93–101), and driven by the hope that psychologists would begin to study it. This hope has been fulfilled beyond all expectations. Work begun in Europe by Shaffer (pp. 98–9) and Gabrielsson (p. 30), built on by their students (Clarke 1995; Juslin 1997; Todd 1985), and taken up enthusi-

astically in the US (e.g. Palmer 1996, 1997, Repp 1990, 1997) has confirmed my hypotheses that expert performance depends upon the flexible deployment of structural knowledge. One reason why this area of work progressed much faster than I expected was the development of MIDI (Musical Instrument Digital Interface) technology, which allowed for easy storage and manipulation of performances through commercially available industry-standard equipment. This now allows the achievement in days of what it took Shaffer years to develop. A notable feature of much of this research is its use of entire pieces of music performed in naturalistic contexts, thus meeting demands of ecological validity, and forming one of the really productive meeting points between psychologists and practising musicians, for whom 'performance studies' is now an established subdiscipline, sharing many of the technologies and methodologies developed in psychology (e.g. Rink 1995). These developments recently allowed the opportunity to revisit the question of fingering strategies (Sloboda, Clarke, Parncutt, and Raekallio 1998), and do the research suggested in 1983 (3.4, pp. 96–7), a source of some personal satisfaction.

**Composition and improvization**

Despite a considerable general interest in creativity and the products of creative individuals (e.g. Simonton 1996), the amount of new work on cognitive processes in composition and improvisation since 1983 has been small. A collection of preliminary studies in this area (Sloboda 1998) has not led to major subsequent developments. In general, what little work there is focuses on detailed analyses of improvisations (e.g. Jarvinen 1995). The only study of composition published in a major music psychology journal this decade is a small expert–novice comparison on a classroom composition exercise (Colley, Banton, Down, and Pither, 1992). It is hard to understand why so little attention has been paid to these areas. Admittedly, they pose methodological problems; but it is possible that their neglect also reflects the general neglect of musical creativity in the arena of the high art or 'classical' tradition which dominates schools, colleges, and universities. Whatever the reason, it means that my 1983 chapter remains as pertinent today as it did then.

**Listening to music**

The notion that music listening draws on general auditory mechanisms designed for a range of detection tasks in the natural world (5.2, pp. 154–66) has now motivated an active research area called 'Auditory Scene Analysis' of which Bregman's (1990) contribution is the defining work (see also McAdams and Bigand 1993). The most impressive application of this work to music is undoubtedly the series of studies by Huron (e.g. 1991, 1993)

which show, through computer-based analyses of very large corpuses of classical compositions, that composers have intuitively conformed to many of the perceptual principles that allow, for instance, separate musical voices to be heard as separate perceptual objects. A natural consequence of this line of work is the demonstration that, at many levels, what experts and novices extract from their hearing of music is remarkably similar, despite the effects of training. Significant evidence of these commonalities has been provided by in a series of studies on cue-abstraction while listening to complete compositions (Deliege and Melen 1997). This work has also been supported by the strong theoretical claim (Lerdahl 1998) that certain structural features (intended by, for instance, some avant-garde composers) are simply incapable of being detected by natural hearing mechanisms, and this accounts for their difficulty and relative unpopularity among listeners.

Given that memorizing music is such a central and ecologically pertinent aspect of musical cognition, I tried to give significant coverage to memory for extended musical sequences in my treatment of listening (5.4, pp. 174–93). A number of books treating music perception and cognition have been published since (chronologically: Howell, Cross, and West 1995; Dowling and Bartlett 1996; Riess-Jones and Holleran 1990; Butler 1992; Aiello 1994; Deliege and Sloboda 1997). It is surprising how little explicit attention is given to memorization (although the role of memory is implied in studies of recognition and other forms of judgment). The most comprehensive coverage of memory is given by Dowling and Harwood (1986) although it reviews almost the same materials as I did. Only two other chapters focus primarily on memory and recall (Sloboda and Parker 1985; Krumhansl 1990). Curiously, the most extended examination of the memorization process since 1983 has been provided through the study of anomalous cognition displayed by autistic musical savants (Miller 1989). It would be fair to say that the predominant strand in recent research on music listening has been to further our understanding of the experience of pitch, within the context of western tonal music (e.g. Cuddy 1997; Cross 1997; Krumhansl 1991; Sloboda 1992). Since memorization requires integration of information across all relevant domains, pitch, rhythm, melody, and large-scale structure, it is perhaps unsurprising that our understanding of it has not progressed as fast as one would have liked. More integrative syntheses in other areas of the kind attempted by Parncutt (1994) for rhythm and metre are required.

### Learning and development

Since 1983, substantial progress has been made in our understanding of the development of musical awareness from birth to school age (Deliege and Sloboda 1995; Trehub, Schellenberg, and Hill 1997). One of the most interesting new findings relating to the first year of life (6.2.1, pp. 198–202) is that

infants have musical preferences which seem to be informed by an appreciation of such high-level musical features as phrase structure (e.g. Jusczyk and Krumhansl 1993). The scaffolding for early musical development provided by parent–child interaction is now much better understood (Papousek 1995).

In middle childhood two research strands have developed. On the one side is research into autonomous or naturalistic musical behaviour, such as song singing (Davidson 1994), or spontaneous notations (Bamberger 1991), following in the tradition of Moog and Gardner (6.2.2, pp. 202–5). On the other are experimental studies, using paridigms developed for the study of adult cognition and which allow direct comparison between different age groups on the same materials, (e.g. Drake 1993; Lamont 1998). However, it would be fair to say that adult cognition still accounts for the vast majority of work in music cognition, and there is still much to be understood about the development of musical cognition.

Perhaps as a result of the slowness of the developmental literature to progress, there have been a few successful applications of cognitive psychology to musical education and training. My enthusiasm for production systems as a theoretical tool (6.3, pp. 215–29) proved to be overstated. The rise of connectionism as an alternative theoretical framework both rendered production-system theorizing unfashionable, and at the same time seemed to disable or inhibit experimental work which might apply this theorizing to concrete musical learning situations (Morrongiello 1992). Instead, the predominant interest of the last few years has been in ways in which involvement with music might aid developmental progress in other areas of cognition (Sharp, Benefield, and Kendall 1998). This may reflect pressures on music within the school curriculum in many countries.

Work on the assessment of musical ability (6.4, pp. 233–8) has not been a major concern of recent research within the cognitive tradition. By and large there is too little clarity concerning what precise cognitive skills are being tapped by many ability tests, and predictive validity is generally low. Individual researchers seem to devise their own tests to suit their particular research question (e.g. Macpherson 1995), with the consequent loss of generality.

### Culture and biology

Cross-cultural issues have continued to be a neglected aspect of the study of musical cognition. Comparisons of oral and literate cultures are almost non-existent, notable and pioneering exceptions being the work of Arom and colleagues on perception of scales and tuning in Central Africa and Indonesia (Arom, Leothaud, and Voisin 1997) and that of Kessler and colleagues on perception of tonality in Bali (Kessler, Hansen, and Shepard 1984). Indeed, even studies of western music show a worrying narrowness,

concentrating on perception and performance of high-art classical music to the near exclusion of other forms.

In relation to biology, on the other hand, new techniques such as Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) have revolutionized the amount and quality of information available concerning brain activity during music engagement. Most of the work in music is in the somewhat better-established area of Event-Related Potentials (ERPs) (see Besson 1997). Much of this has focused on the measuring of musical expectancy as a function of manipulations of structure. There are also some excellent case studies on individuals with brain damage which add to our knowledge of functional dissociation (see Patel and Peretz 1997).

**The scientific milieu**

It is not possible to conclude this brief update without reference to the very changed intellectual climate for the study of music psychology. In 1983, researchers in this topic had only just begun to co-ordinate their activities intentionally. Fifteen years later the activities of the field are co-ordinated by major national and international scholarly bodies, including the Society for Music Perception and Cognition (SMPC), the European Society for the Cognitive Sciences of Music (ESCOM), as well as the Deutsche Gesellscahft fur Musikpsychologie (DGM), Japanese Society for Music Perception and Cognition (JSMPC), and the Society for Research in Psychology of Music and Music Education (SRPMME). A major international conference, the International Conference on Music Perception and Cognition (ICMPC) is held every two years in a different part of the world. When the conference was last held in Europe, in 1994, there were over 200 papers presented. The major journals in the field, *Music perception* and *Psychology of music* go from strength to strength.

Despite the great advances made, challenges remain. The relationship between music cognition and more traditional academic disciplines remains problematical. I see music cognition as primarily a science, with most of its roots in psychology, but with important inputs from linguistics, computer science, and philosophy. However, the subject matter, music, is traditionally studied from within faculties of fine arts. Attempts to locate music cognition teachers and researchers within arts contexts has never proved easy. The assumptions of scientists and artists about how to approach the subject matter are usually too different. This means that the dialogue between scientists and musicians has been and remains problematical.

Another challenge relates to the lack of cohesion within the discipline of music cognition itself. Understanding the phenomenon of music requires inputs from many disciplines, and even from many parts of the discipline of psychology. For instance, the areas covered in this book touch social psychology, developmental psychology, educational psychology, neuro-

psychology, as well as cognitive psychology. Each of these has its own theories, nethdologies, and preoccupations. To become truly expert in any one of them could take a lifetime. Music psychology is, thus, not a coherent discipline, but a loose confederation of disciplines converging around the same object of study. It is hard, therefore, to agree on what the important problems are, or how they should be addressed. International funded research collaborations, such as those found in medicine or physics, are absent from music psychology, with the result that the discipline grows by accretion, each researcher or group making their own idiosyncratic contribution. Governments and other funding agencies are, in general, not interested enough to impose discipline through national or international research policies.

Finally, both psychology and music are undergoing rapid changes as disciplines. Both are entering 'post-modern' phases where some of the earlier certainties are evaporating. In psychology, the hope for a unified theory of mind, possibly in the form of some computational model has faded. In music, activities such as historical musicology, which were once considered the unshakable foundations of the discipline, are being supplanted by a range of interdisciplinary approaches within what is sometimes characterized as 'the new musicology' that relativize, and to a certain extend, undermine the central focus of tranditional musicology, on the historical canon of 'masterworks' which universities study, and music academies prepare young musicians to perform.

These interesting problems form, in themselves, part of my justification for not attempting a thoroughgoing revision of this book. The *Musical mind* was written from within a point of some stability and self confidence in the historical development of cognitive psychology. To reflect the feel of today's intellectual climate, and to foresee some important questions for the next millenium, would require an altogether different approach from a refocused perspective. And two things are very clear to me about that approach: (1) it would need to take much more account of the messy world of emotion and motivation, to balance the precisions and tidinesses of cognition; (2) it would need to encompass a far greater range of the world's musics, cultures, and sub-cultures than the classical art tradition that has moulded the discipline up to this point.

## References

Aiello, R. (Ed.) (1994). *Musical perceptions*. Oxford University Press, New York.

Arom, S., Leothaud, G. and Voisin, F. (1997). Experimental ethnomusicology: an interactive approach to the study of musical scales. In I. Deliege and J. A. Sloboda (Eds) *Perception and cognition of music*. Psychology Press, Hove.

Bamberger, J. (1991). *The mind behind the musical ear*. Harvard University Press, Cambridge, Mass.

Besson, M. (1997). Electrophysiological studies of music processing. In I. Deliege and J. A. Sloboda (Eds.) *Perception and cognition of music.* Psychology Press, Hove.

Bharucha, J. J. (1991). Pitch, harmony and neural nets: a psychological perspective. In P. Todd and D. G. Loy (Eds), *Music and connectionism.* MIT Press, Cambridge, Mass.

Brehman, A. S. (1990). *Auditory scene analysis: the perceptual organisation of sound.* MIT Press, Cambridge, Mass.

Bunt, L. (1994). *Music therapy: an art beyond words.* Routledge, London.

Butler, D. (1992). *The musicians' guide to perception and cognition.* Schirmer, New York,

Chomsky, N. (1965). *Aspects of the theory of syntax.* MIT Press, Cambridge, Mass.

Clarke, E. F. (1995). Expression in performance: generativity, perception and semiosis. In J. Rink, (Ed.) *The practice of performance: studies in musical interpretation.* Cambridge University Press, Cambridge.

Colley, A., Banton, L., Down, J., and Pither, A. (1992). An expert–novice comparison in musical composition. *Psychology of Music*, **20.2**, 124–38.

Cross, I. (1997). Pitch schemata. In I. Deliege and J. A. Sloboda (Eds.) *Percetion and cognition of music.* Psychology Press, Hove.

Cuddy, L. (1997). Tonal relations. In I. Deliege and J. A. Sloboda (Eds.) *Perception and cognition of music.* Psychology Press, Hove.

Davidson, L. (1994). Songsinging by young and old: a developmental approach. In R. Aiello (Ed.) *Musical perceptions.* Oxford University Press, New York.

Davidson, J. W., Howe, M. J. A., and Sloboda, J. A. (1997). Environmental factors in the development of musical performance skill in the first twenty years of life. In D. Hargreaves and A. C. North (Eds.) *The social psychology of music.* Oxford University Press, Oxford.

Deliege, I. (1987). Grouping conditions in listening to music: an approach to Lerdahl and Jackendoff's grouping preference rules. *Music Perception*, **4.4**, 325–60.

Deliege, I. and Melen, M. (1997). Cue abstraction in the representation of musical form. In I. Deliege and J. A. Sloboda (Eds.) *Perception and Cognition of Music.* Psychology Press, Hove.

Deliege, I. and Sloboda, J. A. (Eds.) (1995). *Musical beginnings: the origins and development of musical competence* Oxford University Press, Oxford.

Deliege, I. and Sloboda, J. A. (Eds.) (1997). *Perception and cognition of music.* Psychology Press, Hove.

Dowling, W. J. and Harwood, D. L. (1996). *Music cognition.* Academic Press, New York.

Drake, C. (1993). Reproduction of musical rhythms by children, adult musicians, and non-musicians. *Perception and Psychophysics*, **53**, 25–33.

Ericsson, K. A., Krampe, R. T., and Tesch-Romer, C. (1993). The role of deliberate practice is the acquisition of expert performance. *Psychological review*, **100**, 363–406.

Ericsson, K. A. and Smith, J. (1991). *Towards a general theory of expertise: prospects and limits.* Cambridge University Press, Cambridge.

Hajda, J. M., Kendall, R. A., Careterette, E. C., and Harshberger, M. L. (1997). Methodological issues in timbre research. In I. Deliege and J. A. Sloboda (Eds.) *Perception and cognition of music.* Psychology Press, Hove.

Hargreaves, D. J. and North, A. C. (1997). *The social psychology of music.* Oxford University Press, Oxford.

Howe, M. J. A., Davidson, J. W., and Sloboda, J. A. (1998). Innate talent: reality or myth? *Behavioural and Brain Sciences*, **2**, 399–442.

Howell, P., Cross, I., and West, R. (Eds.) (1995). *Musical structure and cognition.* Academic Press, London.

Huron, D. (1991). The avoidance of part-crossing in polyphonic music: perceptual evidence and musical practice. *Music Perception*, **9.1**, 93–104.

Huron, D. (1993). Note onset asynchrony in J. S. Bach's Two-Part Inventions. *Music Perception*, **10.4**, 435–44.

Jarvinen, T. (1995). Tonal hierarchies in jazz improvisation. *Music Perception*, **12.4**, 415–37.

Jorgensen, H. and Lehmann, A. (Eds.) (1997). *Does practice make perfect? Current theory and research on instrumental music practice.* Norges Musikkhogskole, Oslo.

Jusczyk, P. and Krumhansl, C. L. (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception and Performance*, **19**, 1–14.

Juslin, P. (1997). Emotional communication in music performance: a functionalist perspective and some data. *Music Perception*, **14**, 383–418.

Kessler, E. J., Hansen, C., and Shepard, R. N. (1984). Tonal schemata in the perception of music in Bali and in the West, *Music Perception*, **2**, 131–65.

Krumhansl, C. L. (1990). *The cognitive foundations of musical pitch.* Oxford University Press, Oxford.

Krumhansl, C. L. (1990). Internal representations for musical perception and performance. In M. Riess-Jones, and S. Holleran (1990) (Eds.). *Cognitive Bases of Musical communication.* American Psychological Association, Washington D.C.

Krumhansl, C. L. (1991). Music psychology: tonal structures in perception and memory. *Annual Review of Psychology*, **42**, 277–303.

Krumhansl, C. L. (1998). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Psychology*, **51**, 36–52.

Lamont, A. (1998). Music, education, and the development of pitch perception: the role of context, age, and musical experience. *Psychology of Music*, **26.1**, 7–25.

Lerdahl, F. (1988). Cognitive constraints on compositional systems. In J. Sloboda (Ed.) *Generative processes in music: the psychology of composition, performance, and improvisation.* Oxford University Press, London.

Lerdahl, F. and Jackendoff, R. (1983). *A Generative theory of tonal music.* MIT Press, Cambridge, Mass.

McAdams, S. and Bigand, E. (Eds.) (1993). *Thinking in sound: The cognitive psychology of human audition.* Oxford University Press, Oxford.

MacPherson, G. (1995). The assessment of musical performance: development and validation of five new measures. *Psychology of Music*, **23.2**, 142–61.

Miller, L. K. (1989). *Musical Savants: exceptional skill in the mentally retarded.* Erlbaum, Hillsdale, NJ.

Morrongiello, B. (1992). Effects of training on children's perception of music: a review. *Psychology of Music*, **20.1**, 1992.

Narmour, E. (1989). *The analysis and cognition of basic melodic structures.* University of Chicago Press, London.

Narmour, E. (1992). *The analysis and cognition of melodic complexity.* University of Chicago Press, London.

Palmer, C. (1996). On the assignment of structure in music performance. *Music Perception*, **14**, 21–54.

Palmer, C. (1997). Music performance. *Annual Review of Psychology*, **48**, 115–38.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, **11.4**, 409–64.

Papousek, M. (1995). Intuitive parenting: a hidden source of musical stimulation in infancy. In I. Deliege, and J. A. Sloboda (1995) (Eds.) *Musical beginnings: the origins and development of musical competence.* Oxford University Press, Oxford.

Patel, A., and Peretz, I. (1997). Is music autonomous from language? A neuropsychological approach. In I. Deliege and J. A. Sloboda (Eds.) *Perception and cognition of music.* Psychology Press, Hove.

Raffman, D. (1993). *Language, music, and mind.* MIT Press, Cambridge, Mass.

Repp, B. (1990). Patterns of expressive timing in performances of a Beethoven Minuet by nineteen famous pianists. *Journal of the Acoustical Society of America*, **88**, 622–41.

Repp, B. (1997). The aesthetic quality of a quantitatively average music performance: two preliminary experiments. *Music Perception*, **14.4**, 419–44.

Riess-Jones, M. and Holleran, S. (Eds.) (1990). *Cognitive bases of musical communication.* American Psychological Association, Washington, D.C.

J. Rink, (Ed.) *The practice of performance: studies in musical interpretation.* Cambridge University Press, Cambridge.

Sharpe, C., Benenfield, P., and Kendall, L. (1998). *The effects of teaching and learning in the arts.* Qualifications and Curriculum Authority, London.

Simonton, D. K. (1996). Creative expertise: a life-span developmental perspective. In K. A. Ericsson (Ed.) *The road to excellence: the acquisition of expert performance in the arts and sciences, sports and games.* Erlbaum, Mahway, NJ.

Sloboda, J. A. (1986). Cognitive psychology and real music: the psychology of music comes of age. *Psychologica Belgica*, **26.2**, R.

Sloboda, J. A. (Ed.) (1988). *Generative processes in music: the psychology of composition, performance, and improvisation.* Oxford University Press, London.

Sloboda, J. A. (1991). Music structure and emotional response: some empirical findings. *Psychology of Music*, **19**, 110–20.

Sloboda, J. A. (1992). Psychological structures in music: core research 1980–1990. In. J. Paynter, T. Howell, R. Orton, and P. Seymour (Eds.). *Companion to contemporary musical thought.* Routledge and Kegan Paul, London.

Sloboda, J. A. and Davidson, J. W. (1995). The young performing musician. In I. Deliege and J. A. Sloboda (Eds.) *The origins and development of musical competence.* Oxford University Press, London.

Sloboda, J. A., Clarke, E. F., Parncutt, R., and Raekallio, M. (1998). Determinants of finger choice in piano sight reading. *Journal of Experimental Psychology: Human Perception and Performance*, **24.1**, 185–203.

Sloboda, J. A., and Parker, D. H. H. (1985). Immediate recall of melodies. In P.

Howell, I. Cross, and R. West (1995) (Eds.) *Musical structure and cognition.* Academic Press, London.

Todd, N. P. M. (1985). A model of expressive timing in music. *Music Perception*, **3**, 33–58.

Trehub, S., Schellenberg, G., and Hill, D. (1997). The origins of music perception and cognition: a developmental perspective. In I. Deliege and J. A. Sloboda (Eds.) *Perception and cognition of music.* Psychology Press, Hove.

Waters, A. J. and Underwood, G. (1998). Eye movements in a simple music-reading task: a study of expert and novice musicians. *Psychology of Music*, **26.1**, 46–60.

Watt, R. and Ash, R. (1997). A psychological investigation of meaning in music. *Musicae Scientiae*, **2.1**, 33–54.

Weisberg, R. W. (1993). *Creativity: Beyond the Myth of Genius.* Freeman, San Francisco.

# Contents


1. MUSIC AS A COGNITIVE SKILL 1

2. MUSIC, LANGUAGE, AND MEANING 11

2.1 Introduction 11
2.2 Chomsky and Schenker 11
2.3 Other comparisons between language and music 17
2.4 Musical phonology 23
2.5 Musical syntax 32
2.6 Musical semantics 57
2.7 Conclusions 65

3. THE PERFORMANCE OF MUSIC 67

3.1 Introduction 67
3.2 Sight reading 68
3.3 Rehearsal 90
3.4 Expert performance 93

4. COMPOSITION AND IMPROVISATION 102

4.1 Introduction 102
4.2 Manuscript evidence 104
4.3 Composers' writings on their own compositional process 115
4.4 Observations of composers at work 123
4.5 Improvisation 138

5. LISTENING TO MUSIC

5.1 Introduction 151
5.2 'Natural hearing': primitive grouping mechanisms in music 154
5.3 Attention in music listening 166
5.4 Memory in music listening 174

6. MUSICAL LEARNING AND DEVELOPMENT

6.1 Introduction 194
6.2 Musical enculturation 198

6.3 Training and skill acquisition 215
6.4 Assessing musical ability 233

7. THE MUSICAL MIND IN CONTEXT: CULTURE AND BIOLOGY

7.1 Introduction 239
7.2 Culture and musical thinking 240
7.3 Biology and musical thinking 260

REFERENCES 269

INDEXES 283

# 1. Music as a cognitive skill

The reason that most of us take part in musical activity, be it composing, performing, or listening, is that music is capable of arousing in us deep and significant emotions. These emotions can range from the 'pure' aesthetic delight in a sound construction, through emotions like joy or sorrow which music sometimes evokes or enhances, to the simple relief from monotony, boredom or depression which everyday musical experiences can provide.

If someone from a civilization without music were to ask us why our civilisation supported so much musical activity, our answer would surely point to this capacity of music to heighten emotional life. Of course there are other reasons for individuals or societies to make use of music. Because much musical activity is also social activity it can come to have many social meanings, offering a variety of social rewards to those who participate in it. For instance, knowledge about certain types of music is a prerequisite for full membership of particular subcultures. The adoption of music for use in formally defined social situations allows people to earn a living through music. Aspects of musical training, such as the need for discipline and co-operation, may be held to be generally worth while. And so on. However, we may label these social motivations as secondary, if only because they are so closely tied to particular cultures. The emotional factor is, however, transcultural. It seems unlikely that music could have penetrated to the core of so many different cultures unless there were some fundamental human attraction to organized sound which transcended cultural boundaries.

If emotional factors are fundamental to the existence of music, then the fundamental question for a psychological investigation into music is *how* music is able to affect people. Seen with the cold eye of physics a musical event is just a collection of sounds with various pitches, durations, and other measurable qualities. Somehow the human mind endows these sounds with significance. They become symbols for something other than pure sound, something which enables us to laugh or cry, like or dislike, be moved or be indifferent.

There are two reasons why this issue leads us straight into the realms of cognitive psychology. The first is that most of our responses to music are learned. This is not to deny the possibility of some primitive responses to music shared by the whole species. For instance, loud fast music is arousing, whilst soft, slow music is soothing. Certain pitch ranges and timbres seem particularly attractive to infants, as do simple repetitive rhythms. But these primitive tendencies cannot account for the subtle and multidimensional

nature of the adult response to music, nor can they account for the many significant cultural differences in response. I can still recall the surprise I felt when, after hearing a rather jolly sounding Greek folk song, my Greek companion indicated that this was a desperately sad piece of music. No doubt, an understanding of the words was partly responsible for her different view of the music; but I realized that my own response was determined by the major key, the simple and open harmonies, and the general rhythmic impetus. These qualities are, in fact, common to most Greek folk music. More subtle differences, undetected by me, are responsible for the communication of affect.

A second pointer to the cognitive component of our emotional responses is the fact that they are not to be explained simply in terms of 'conditioning'. Conditioning theory (amusingly called the 'Darling, They're Playing Our Tune' theory by Davies 1978) supposes that a piece of music acquires the emotional significance of the circumstances in which it happened to be heard. On this theory, the form and content of the music is irrelevant to its acquired emotional character. Only its context is important. Although this type of conditioning may occur in certain circumstances, and may help explain particular idiosyncrasies of musical taste, there are several observations which show that it is inadequate as a complete explanation for emotional responses:

1. Listeners within a musical culture generally agree on the emotional character of a given piece of music, even though they may have never heard it before (Hevner, 1936). Conditioning theory, in contrast, would predict wide individual differences according to circumstances of hearing.

2. The emotional character of a piece of music is not unitary and unchanging. On the contrary, experienced listeners to some types of music are able to identify a web of differing emotions which are evoked by the detailed sequence of events, and which become more finely differentiatied the better the music is known. Conditioning theory, in contrast, would lead one to expect that a piece of music would always be dominated by a single generalized emotional tone acquired from the conditioning context.

3. Our emotional response to the very same music can vary considerably from hearing to hearing. For instance, there have been occasions when the last movement of Tchaikovsky's Sixth Symphony has moved me to tears of grief, and others on which it has left me completely unmoved. What remained the same about my mental state on the different occasions was my *knowledge* that this was music *expressing* extreme grief, regardless of my emotional response to it.

An analogy which I find helpful is that between music and humour. When hearing a joke the listener must first understand it; he must perceive and identify the constituent words, recognise them as sentences, form a mental representation of the propositions that the sentences assert, and then determine the nature of the incongruity, double meaning, or whatever that makes the joke. Thus, to 'get the joke' involves a large set of cognitive

processes, drawing upon knowledge of language and the world. After this the listener may, depending on his mood and taste, experience an emotional reaction involving laughter. So there is both a cognitive and an affective stage to his response. The cognitive stage is a necessary precursor of the affective stage; a listener cannot find a joke funny unless he understands it. However, the affective stage does not *necessarily* follow the cognitive stage. A person may understand a joke perfectly well without being moved to laughter by it. So it is with music. A person may understand the music he hears without being moved by it. If he *is* moved by it then he must have passed through the cognitive stage, which involves forming an abstract or symbolic *internal representation* of the music. The nature of this internal representation, and the things it allows a person to do with music, is the central subject matter of the cognitive psychology of music.

The way in which people represent music to themselves determines how well they can remember and perform it. Composition and improvisation require the generation of such representations, and perception involves the listener constructing them. These representations, and the processes which create them, are not directly observable. We have to infer their existence and nature from observations of the way in which people listen to, memorize, perform, create, and react to music.

Because such activities are learned, they may be viewed as skilled behaviours. Although composing and performing are universally acknowledged to be skills of some complexity, one must also remember that more commonplace activities, such as the ability to whistle a familiar melody, or to detect a 'wrong sounding' note in an unfamiliar melody, are also complex skills which are capable of shedding light on the nature of internal representations in music. What makes the composer or performer special is his rarity rather than anything *fundamentally* different about his mental equipment.

To illustrate this point, let us take two contrasting examples of musical skill and see what a cognitive psychologist might make of them. The first is the much quoted incident in the life of the young Mozart who, denied access to the score of Allegri's *Miserere*, listened to two church performances and then wrote it out from memory. The primary source for the authenticity of this incident is a letter from Mozart's father dated 14 April 1770 (Anderson, 1966). Such an accomplishment, although by no means unique, is well beyond the capacity of most people, and has become one of a handful of legends about the apparently superhuman powers of great musicians. In discussing this example the psychologist Farnsworth (1969) claims that Mozart's 'eidetic' memory was responsible for his ability. Eidetic memory is a peculiarly vivid, almost hallucinatory, recall which is alleged to be present in some children. On this view, the unusual accomplishment is the result of possessing unusual mental powers, since most people do not possess eidetic memory.

An alternative explanation is that Mozart was more adept, through

experience, at doing something we all do when trying to memorize complex material; that is, to identify patterns in the material, thereby remembering groups as single units or 'chunks'. It is difficult to decide between these two alternatives with respect to Mozart's memory, since we cannot test him further. There is, however, contemporary evidence which tends to favour the latter explanation. For instance, there has been much recent interest in the abilities of game players with various levels of expertise to remember game positions. Chase and Simon (1973) asked chess players of varying standards to examine boards taken from actual games for five seconds and then attempt to reproduce the boards from memory. Players in the category of chess master were able to reproduce the boards almost perfectly, whereas novices did very poorly. This is, perhaps, unsurprising, and similar results have been obtained when musicians and non-musicians are asked to reproduce the positions of a few notes displayed briefly on a stave (Halpern and Bower 1982; Sloboda 1976a). The really important result is that when chess players were asked to reproduce random boards which could not have occured in any rational game, the masters and novices performed equally, and very poorly. This shows that masters do not have better visual (eidetic) memory than novices. They are superior *only* when the material to be remembered makes some sense, and can be grouped into strategic 'chunks' (e.g. a pawn chain, a fork, a concealed check). This is confirmed by detailed observation of the way masters carry out the reproduction task. They tend to recall all the members of a chunk together, with pauses between separate chunks.

For a chess master, the mental representation of a chess position is not a copy of the physical board. It is a more abstract *structural* description of the meaningful relationships between groups of pieces. Through many years of experience he has acquired automatic perceptual mechanisms which rapidly pick out frequently occuring strategic patterns from the input. It has been estimated that the master must be able to recognise several thousands of patterns in this way.

Most musical material has just as much patterning and structure as a chess position. Within particular idioms there are many regularly occuring patterns (e.g. chords, scales and arpeggios) and someone exposed to an idiom will rapidly become familiar with them. Patterning also occurs on a larger scale. Thematic repetition is a cornerstone of most music, certain harmonic progressions are ubiquitous, and so on. The master can make use of all this patterning when building up a representation of a piece of music.

My second example is commonplace. Nearly everybody is capable of identifying well known melodies, or of recalling them if given the title (e.g. 'Yankee Doodle', 'Frère Jacques', 'God Save the Queen'). What makes this observation particularly important in the current context is that a precise pitch, speed, or medium is not crucial for identification, nor will a person necessarily reproduce a melody at a pitch he has ever heard before. For instance, Attneave and Olson (1971) asked people to sing the NBC chimes

from various given starting notes. The chimes are broadcast several times each day as a station identifier for the American National Broadcasting Corporation, but they are always played at the same pitch. The experimental subjects were able to transpose the chimes to any starting note within their range instantly and effortlessly. This is one of several demonstrations that people do not remember simple melodies in terms of precise pitches and durations but in terms of patterns and relationships. The NBC chimes remain identifiably the same so long as the pitch and duration *ratios* between the notes remain constant. This simple demonstration of musical memory already demonstrates how it is a *abstraction* from the physical stimulus. Sequential structure also influences untrained musicians. For instance, music which conforms to the rules of diatonic tonal harmony is much easier for Western listeners to remember than is music which breaks these rules. Some of the most important demonstrations of this are reviewed in Chapter 5, which examines the processes involved in listening to music. Here I wish to stress the closeness of the analogy to the chess experiments. Music which does not contain familiar patterns and structures cannot easily be represented in a listener's memory.

If expert and commonplace musical achievements are to be explained in the same sort of way, how are we to account for the *differences* between experts and novices? The most important factor is probably the *number* and *complexity* of the structural features in terms of which the listener is capable of representing music. The 'ordinary' listener is probably aware only of the musical 'foreground', noticing small patterns which are made up from a few adjacent notes. The Mozarts of the world probably share this level of representation, but have overlaid it with higher-order grouping processes which form single 'background' units from groups of 'foreground' units. The same process could be repeated on the new units to form even larger units, building up a hierarchical schema which could span a piece of music lasting several minutes.

A second difference between trained and untrained musical memories may lie in the degree of awareness that the memorizer has of the structures he is using. Commonly, musical training involves acquiring the vocabulary in terms of which to describe the structure of music. Elements of this vocabulary (e.g. cadential sequence, tonic chord, passing note) often function as means for extending memory capacity. A musician can be consciously looking out for instances of such structures. This does not imply that such structures are not available to untrained people. The evidence suggests, rather, that untrained musicians have implicit knowledge of that which musicians can talk about explicitly. In this respect, music is similar to language. Ordinary people speak their natural language according to the same grammatical rules as professional linguists even though they may have very limited conscious knowledge of those rules.

The 'representation of musical structure' is a unifying concept for the

various topics treated in this book; and it may be helpful at this point to give a brief sketch of the topics to be covered, and the approach I shall be taking to them.

One of the key features of any cognitive skill is speed. A process such as reading or listening is useless if it cannot keep up with the inflow of new material. Mozart had not only to notice structural relationships in the piece he memorized, he also had to perform all the necessary mental operations quickly enough to keep abreast of the music. Skilled performance must also be rapid and fluent. This has led psychologists to think that performance is controlled by the same sort of hierarchical structures as seem to account for input skill. Here, highly automated procedures control detailed performance while the performer is free to attend to higher-order aspects. For instance, the experienced sight reader confronted with a familiar scale passage will not need to make conscious decisions about which fingers to use for which notes. His hand will automatically take up the right configuration, while his attention may be on expressive factors, or on mentally preparing the following phrase. When we hear the phenomenal speed of a virtuoso performance, we can render it explicable, although no less of an achievement, by recalling that what strikes the ear as 20 separate notes may well be, for the performer, a single integrated and automated unit. What makes the achievement impressive is the many hundreds of hours' practice which may be needed to develop the degree of automation required. In examining psychological investigations of music performance (Chapter 3) we shall find considerable evidence for abstract high-level structures governing all aspects of skilled performance.

When we turn to the creation of music through composition or improvisation (Chapter 4), we shall find that the same notions of structured representation and rapid deployment of structural units are helpful in understanding these skills. Succesful composition relies on the ability to be sensitive to very large-scale structures, so that detailed working out can be governed by a conception of the overall framework. This applies to any large creative endeavour, such as writing a book. Without some unifying conception the author will find it hard to co-ordinate and shape the many sub-sections. Success in improvisation relies on structural grasp too, but also on the ability to rapidly choose appropriate, yet novel, elements to fill the slots in a basic script. Improvisers achieve their best results when working within a highly constrained form (e.g. jazz blues, fugue), whereas composers are more likely to achieve successful innovation of forms.

Since a basic premise of this book is that we *learn* the structures that we use to represent music, an examination of the learning process itself is central to our concerns (Chapter 6). This process can be seen as having two phases. The first is developmental enculturation. This is the learning that takes place as a result of our exposure to the normal musical products of our culture in childhood, together with the acquisition of simple skills like the ability to

reproduce short songs. The knowledge acquired at this stage is not, on the whole, the result of self-conscious learning or teaching effort. Rather, children simply acquire the knowledge through their everyday social experiences. In consequence such knowledge tends to be universal in a culture, and is the ground on which more specialized skills may be built. The second phase is, then, the acquisition of specialized skills through training. These skills are not universal within a culture, and are what turn ordinary citizens into 'musicians'. Our treatment of enculturation can be only descriptive. The treatment of specialized skills should arguably be prescriptive as well. Musicians have some right to expect that cognitive psychologists should be particularly well qualified to advise teachers and students on the self-conscious acquisition of skill.

Because musical skills are culturally transmitted, it follows that they may vary significantly between cultures. Many writers on the psychology of music talk about Western tonal music as if it were the only type of music. Whilst ignorance prevents me from dwelling on music from other cultures at any length, Chapter 7 contains an attempt to evaluate just how far some of the cognitive principles which underlie Western musical skill are relevant to other forms of music.

Most contemporary thought in cognitive psychology is functional, in that it attempts to discover the *principles* of structure and operation which would allow an information handling system to behave in the way we observe humans to behave. We have to keep in mind, however, that these principles are *in fact*, in our case, realized in the structures and functions of the *brain*, and that it is worthwhile finding out what we can about the brain in relation to the skills we examine. Studies reviewed in Chapter 7 reveal that, as with other cognitive skills, components of musical skill appear to have specific brain localization. The functional organization of the brain is modular, with specific portions doing specific jobs.

I began this chapter with reference to what is, humanly, the most important psychological fact about music; that it carries emotional significance or meaning for us. To say this prompts the thought that music is, in some respects, like a language. Although many writers on music have used this analogy in a loosely metaphorical way, there is now a growing body of theory and data which make a more rigorous assessment possible. Chapter 2 attempts such an assessment. It is possibly the most technical of the chapters, but is central to the book in that it provides a source of ideas about the nature of cognitive representations in music. These ideas are put to work elsewhere in attempting to understand specific musical skills. I hope, though, that the selective reader will be able to read the other chapters in isolation, and still obtain a broad idea of how music is represented in the mind, and why an understanding of these representations is so central to an understanding of all the skills discussed.

Before going further, I must say something about the term 'cognitive

psychology' itself and its role in this book. Cognitive psychology is a fairly new branch of psychology. It is also a rapidly changing discipline, and a generally acceptable definition of its aims and boundaries would be hard to produce. The emergence of an identifiable discipline coincided approximately with the publication of Broadbent's seminal *Perception and Communication* (1958). It marked the advent of a cautious experimental mentalism, which showed that it was possible to subject processes such as attention and recognition to rigorous laboratory investigation. Subjects were not required to introspect on these processes but to perform tasks from whose accuracy or speed of execution could be inferred something about the internal structure of the system controlling these tasks. It was another nine years before Neisser's influential *Cognitive Psychology* (1967) gave the discipline its name, but during the years spanned by these two books the preoccupations of researchers in this field remained quite narrowly defined. The research concentrated on 'input' skills, including the recognition of letters, words, and other simple stimuli, in artificially simplified laboratory surroundings. Major theoretical disputes centred on questions of capacity: how many items could be held in memory at once? were cognitive operations on incoming stimuli performed in series or in parallel? The 1970s saw a broadening of the field in several directions. One development was a greater interest in higher cognitive processes and the control of complex behaviour, and an allied interest in the organization of knowledge and its deployment in cognitive skills. Several influences were important here. One was a renewal of interest in problem-solving behaviour, much neglected since the 1930s. Newell and Simon's *Human Problem Solving* (1972) is arguably the most important contribution to this branch of the subject, and is also a prime example of another very important influence on cognitive psychology—the growth of the science of artificial intelligence. The opportunity to embody theories in working computer programs has stimulated concern with theories of cognitive functioning which are specific enough to predict actual instances of behaviour, but which are broad enough to account for a wide range of cognitive achievements.

The second major development of recent years has been the attempt to study cognitive skills in situations more closely resembling those in which people would normally employ them outside the laboratory. Thus, concern has focused on how people deal with extended and meaningful material rather than on fragmented, meaningless stimuli. This movement has brought in its train an increased interest in developmental aspects of cognition, and cultural variance; and a greater commitment to finding practical applications for cognitive research. Neisser's second book, *Cognition and Reality* (1976) has been a rallying point for this approach, and some of the first fruits of research in this vein are exemplified by collections such as Gruneberg, Morris, and Sykes' *Practical Aspects of Memory* (1978).

Luckily for the beginning student there are now many excellent introductions to the subject available (e.g. Anderson 1980; Solso 1980), although

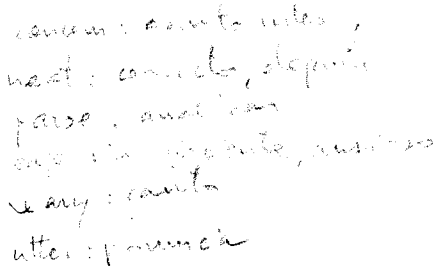
they sometimes tend to exaggerate the stability and unity of the subject. Claxton (1980) provides some more candid and critical assessments, although some of the arguments will be difficult for beginners to follow. In this book I identify, although fairly loosely, with what I perceive as the mainstream of current cognitive research outlined above. The focus is on empirical study of what people *do* with music, rather than what they *say* they do, and, where possible, I address myself to the real-life behaviour of musicians, rather than to their behaviour in artificially constricted laboratory situations. I try to give as much prominence to output skill (e.g. performing) and higher cognitive functioning (e.g. composition) as I do to input skill (e.g. listening). Finally, my criterion for selecting empirical studies to describe is that they should have a particular theoretical or practical importance. I make no claim to provide a complete survey of the literature.

The cognitive psychology of music is still a very young discipline and there are several areas of enquiry which have not yet received detailed empirical attention. I have felt it important to identify some of these areas, and say something about the kinds of questions which need to be asked. In this respect the book looks forward to the future, and is not simply a statement of the current position. I have freely used my experiences and intuitions as a practising musician to augment and complement experimental findings; indeed the section on composition would hardly have been possible without them. Here, as elsewhere, my aim has been to stimulate scientific thinking and experiment. If some of the ideas in this book are proved wrong through new research no-one will be more pleased than I.

The last purpose of this chapter is to place this book in the context of other writing on the psychology of music, and to suggest some further general reading in the area. Some aspects of the psychology of music fall outside the concerns of this book. I make little mention of musical taste or preference, and the reader is referred to Farnsworth (1969) for an introduction to this area. I do not deal with social applications of music (e.g. to industry or therapy). Farnsworth is quite useful here, and Priestley (1975) and Alvin (1975) are two of several introductions to music therapy. There is also a *Journal of Music Therapy*. Purely educational research falls outside the scope of this book. Two main journals in this area are the *Journal of Research into Music Education* and *Psychology of Music*. Sensory psychology and psychoacoustics are not covered in any detail. The *Journal of the Acoustical Society of America* is a prime source of research in this area. Plomp (1977) offers a recent synthesis, and Moore (1982) provides a sensible introduction to the area. Research on music reported in the main perceptual journals tends to be of this sort. My last area of exclusion concerns psychometric approaches to musical ability. Shuter-Dyson and Gabriel (1981) provide a useful summary of research in this tradition.

General collections of papers by Critchley and Henson (1977), Clynes (1982) and Deutsch (1982a) are useful source books for the whole subject. A

recent book by Davies (1978) approaches nearest to the orientation of this book, although his emphasis, and that of a new journal *Music Perception*, is on the processes involved in listening. Finally, the journal *Psychomusicology* has recently been founded with an editorial policy which closely fits the aims of this book. The name, although new and unwieldy, is perhaps the best single-word description for the subject which this book explores.



# 2. Music, language, and meaning

## 2.1. Introduction

This chapter begins by looking at two influential theorists, the linguist Chomsky and the musicologist Schenker. Their theories have some striking similarities. They both argue, for their own subject matter, that human behaviour *must* be supported by the ability to form abstract underlying representations. Two of man's highest and most complex products seem to display something central about his intellect.

Chomsky and Schenker gained their principal insights from examining the *structure* of language and music rather than from examining linguistic and musical *behaviour*. Recent empirical work has shown that music and language share behavioural as well as formal features; and so the chapter moves on to review this work. The major portion of the chapter is organized around the subdivision of language and music into three components: phonology, syntax, and semantics. Phonology concerns the way in which a potentially infinite variety of sounds are 'parcelled up' into a finite number of discrete sound categories which constitute the basic communicative units. Syntax concerns the way in which these units are combined into sequences. A major concern of those studying syntax has been the discovery of rules which reliably produce legal sequences and eliminate illegal ones. Semantics concerns the way in which meaning is carried by the sequences so constructed. Particular fixed combinations of phonological units can have fixed meanings, but it is a characteristic of both language and music that meaning is also carried by the ordering and combination of elements in longer sequences.

The concern of this chapter is not only to describe linguistic and musical rules or categories, but also to examine the degree to which psychological evidence confirms their reality. A linguist searches for the neatest and most economical way of describing a language's structure. A *psycholinguist* seeks the description which most closely resembles the psychological processes by which language is produced and understood. And so it must be for the psychomusicologist.

## 2.2. Chomsky and Schenker

The cognitive psychology of language almost owes its very existence to the linguist Noam Chomsky. His radical approach to the problem of describing language structure (Chomsky 1957, 1965, 1968) was the main impetus to the formation of the discipline of psycholinguistics, and the impact of his work

still reverberates around the psychological laboratories of the world. As is the case with all seminal thinkers his work has attracted as much cogent opposition as it has support, but it is doubtful if any current approach to language owes nothing to him.

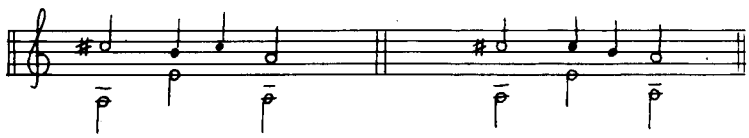
One of Chomsky's main claims has been that, at a deep level, all natural languages have the same structure, and that this structure tells us something universal about the human intellect. Many writers on music have been attracted to the idea that there are musical universals; maybe that there is something particularly natural about tonality. Could it be that Chomsky's theses receive independent support from music? A few writers seem to think so. For instance, Leonard Bernstein has recently celebrated his discovery of the 'Chomsky connection' with characteristic exuberance (Bernstein 1976). Sundberg and Lindblom (1976) have written a 'Chomskyan' grammar for a set of Swedish folk-songs; and researchers in artificial intelligence have written programs for parsing music which share formal properties with language programs. For the most part, however, psychologists investigating music have not been eager to pursue this path, preferring to treat music as a totally independent phenomenon.

There are, of course, good scientific reasons for approaching the study of music without too many presuppositions, but my own attitude to this issue has been profoundly affected by the writings of the German musicologist Heinrich Schenker (1868–1935) which have been gaining influence in contemporary analytic thought (e.g. Schenker, 1935). There are some striking parallels between Schenker's views on music and Chomsky's on language. For instance, Schenker would wish to claim that, at a deep level, all good musical compositions have the same type of structure, and that this structure reveals to us something about the nature of musical intuition. There is no evidence to suggest that Chomsky knew of Schenker's work at the time of formulating his theories of language. It appears, therefore, that similar ways of viewing language and music arose quite independently in two creative intellects steeped in their own subject matter.

My view is that the linguistic analogy in music deserves serious attention; although I would make three qualifying remarks. Firstly, it would be foolish to claim that music is simply another natural language. There are many fundamental differences which cannot be overlooked, the most obvious being that we use language to make assertions or ask questions about the real world and the objects and relationships in it. If music has any subject matter at all, then it is certainly not the same as that of normal language. Secondly, this analogy can clearly be exploited in metaphorical, poetic ways, of which scientists are right to be wary (e.g. 'music is the language of the emotions'). However, this does not exclude the possibility of a more rigorous use of the analogy. Indeed, it is arguable that the most productive scientific analogies are those which fire the imagination as much as they fuel empirical and theoretical endeavours. Thirdly, it is an open question just how far the

analogy can be pursued, and its application is subject to the weighing of empirical evidence and argument that characterizes any scientific endeavour. In other words, the analogy is something to be evaluated, not assumed.

What, then, are the main similarities between Chomsky and Schenker? Perhaps the most fundamental similarity concerns the differentiation between *surface* and *deep* structure. Surface structure is, roughly, the form taken by a linguistic (or musical) sequence as it is uttered (or written). The two sentences 'John phoned up Mary' and 'John phoned Mary up' have different surface structures; and so do the two musical sequences given in Example 2.1. Nevertheless, in both cases we would wish to say that there is a particular closeness between the two sequences, which is not simply a result of their containing the same elements in a different order (e.g. the closeness would not be shared by 'Phoned John up Mary'). In the language example we would wish to say, at the least, that the two sentences *mean* the same thing. Chomsky proposed that we could capture the closeness of such sentences within a grammar, if we assigned them to the same deep structure. This deep structure is an abstract entity, not itself a sentence, from which both surface structures can be derived by the application of *transformation* rules.



Example 2.1.

In order to represent the deep structure of a sentence it is useful to employ an 'immediate constituent tree' notation as shown in Figs. 2.1–2.3. Each tree has a single *node* at its head; S for sentence. This node can be decomposed into two or more nodes at the next lower level. The rules governing the decomposition of nodes are called the generative (or rewrite) rules of the grammar. One generative rule allows S to be rewritten as NP + VP (noun phrase + verb phrase). Other rules allow further rewriting (e.g. VP goes to V + NP). At the lowest level, individual words can be specified. For many simple sentences, immediate constituent analysis of this kind gives a satisfactory account of their structure because words connected to a single node are next to one another. Grammatical relation is signalled by sequential adjacency. In Fig. 2.1 'The' has a closer grammatical relation to 'girl' than any other word in the sentence. An informal way of testing the grammatical relation of two words is to see whether they could be replaced by a single word to form a new meaningful sentence. Thus, 'The girl' could be replaced by 'Mary', and 'hit the ball' by 'sang'; but 'girl hit the' cannot be so replaced.

There are, however, many sentences where an immediate constituent analysis does not capture our intuitions about the grammatical relation of words. In Fig. 2.2 we find that the constituent tree links 'up' to 'Mary', whereas one

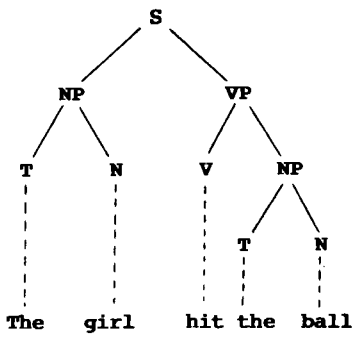



Fig. 2.1

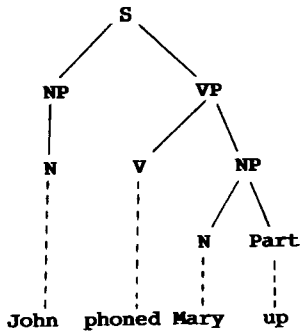



Fig. 2.2

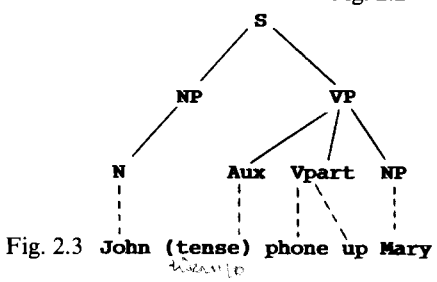


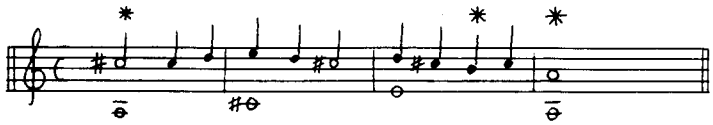

Fig. 2.3

Fig. 2.1–2.3. Examples of immediate constituent trees for English sentences. Key: S = sentence, NP = noun phrase, VP = verb phrase, T = article, N = noun, V = verb.

may feel that 'up' is, in fact, more related to 'phoned'. What is more, we can introduce longer sequences between 'phoned' and 'up' without destroying this relation. For instance, 'John phoned the brown eyed girl he met for the first time yesterday up' is an acceptable sentence. Chomsky's solution to this problem is to propose that the deep structure is a tree in which words having close grammatical relationship are governed by the same node. So 'John phoned up Mary' and 'John phoned Mary up' are given the *same* deep structure, shown in Fig. 2.3. Note that the symbols in the bottom row (the terminal string) do not constitute a sentence. There are some obligatory transformations that must be performed (in this case the inflection of the verb form to indicate tense). Other transformations are optional, such as the moving of 'up' to the end of the noun phrase governed by the verb phrase. These transformations generate the various different surface structures that a single deep structure may exhibit.

Turning now to the music sequences in Example 2.1, although Schenker analysis would rarely be applied to such trivial excerpts, we can note some interesting parallels to the language examples. Schenker would give both excerpts the same deep structure (*Ursatz*). This deep structure always has two components: a fundamental melodic line (*Urlinie*) which, in this case, is the progression 3–2–1 in the scale of A major, and a bass arpeggiation

(*Bassbrechung*) which, in this case, is the progression I–V–I in the same key. In each excerpt the melodic C receives no explicit representation in the deep structure. Rather, it is an interpolation between the notes of the *Urlinie*which results in expansion or *prolongation* of the material. However, no amount of appropriate prolongation destroys the integrity of the *Ursatz* or the fundamental interdependence of its elements. For instance, Example 2.2 possesses precisely the same deep structure as both sequences in Example 2.1. The asterisked notes show the *Ursatz*. The relation of the melodic B to the preceding C is not weakened by the intervening notes, just as the relationship of 'phoned' to 'up' is not weakened by intervening words.



Example 2.2.

The source of the *Ursatz* (the top node in a hypothetical constituent structure tree) is the chord of the tonic triad. It must, however, be admitted that neither Schenker nor any of his followers have produced anything approaching a formal generative grammar for deriving *Ursatze* from triads, or for transforming an *Ursatz* into an acceptable composition through prolongation. What they have shown is that *Ursatze* are discoverable in a very large body of tonal compositions. One aim of modern linguists is to produce a set of generative and transformational rules which generates all and *only* acceptable sentences of a language like English. While Schenkerians might claim that all good compositions are built upon *Ursatze*, they would not claim to be in possession of rules which would generate *only* good music. Schenker's method is, in essence, analytic, not generative.

There are three significant contrasts between Schenkerian and Chomskyan structures which we should note before proceeding. Firstly, the *Ursatz* is, in itself, a legal although trivial piece of music. It obeys rules of counterpoint and harmonic progression just as do the 'surface' elements. Schenker analysis proceeds by a recursive 'reduction' of a finished composition, eliminating subsidiary prolongations to reveal a harmonic and contrapuntal 'skeleton'. This skeleton is then further reduced by the very same methods over and again until the simplest and most fundamental structure is revealed. In contrast, as we have already noticed, Chomskyan deep structures are not, in themselves, acceptable sentences.

Secondly, the transformations applied to an *Ursatz* result in a lengthening and complexifying of the music. Very long movements containing hundreds of notes may be subsumed under a single *Ursatz*. To demonstrate this was one of Schenker's major achievments. In contrast, Chomskyan deep structures govern only single sentences, which are typically short, and a common effect

of applying transformations is condensation rather than expansion. For instance the sentence 'John likes Mary and John likes Jane' mirrors rather closely the deep structure which would normally be transformed and shortened to 'John likes Mary and Jane.'

Thirdly, differentiation between pieces of music is achieved at the surface level rather than at the level of deep structure. According to Schenker there are very few different types of *Ursatz*. If the *Ursatz* determined the significance of music, then there would be very little for composers to say. Much of the 'meaning' of any piece of music is given to it through the actual surface details. In language it is different. The significance of a sentence, in practical discourse, lies almost entirely in its deep structure. There are an indefinite number of deep structures, corresponding to all the different propositions thata language can express. Transformations of these deep structures often *reduce* differentiation because, for instance, different deep structures can be transformed into identical surface structures—as in 'Bill shot the man with a gun'. This is ambiguous: one possible deep structure would have Bill doing the shooting with a gun; the other would have the man possessing a gun.

The demonstration that events quite far apart in both linguistic and musical sequences can have a close structural relationship has two major consequences. The first is that whatever grammars are eventually shown to be fully adequate for language and music, they must be more powerful than 'finite state' grammars. Chomsky (1957) provided the classic proof that this type of grammar is inadequate for a natural langauge. In such a grammar, words are generated one at a time, each word determining the set from which the next word may be chosen. The rules are thus 'context independent'. It does not matter in which context a word is found; exactly the same set of consecutive words is permissible in each case. In a finite state grammar, for instance, the word 'washes' must allow both 'himself' and 'herself' to follow, since 'the boy washes himself' and 'the girl washes herself' are both correct English sentences. Yet such a grammar would allow 'the boy washes herself', which is unacceptable. To improve on a finite-state grammar we must introduce 'context sensitive' rules which take into account more than the immediately preceding word. The same arguments apply by analogy to music.

The second consequence is more directly psychological. If music and language have properties which demand grammars of a certain complexity, then humans must have the psychological resources to represent such grammars. This claim is hardly controversial nowadays, but it is incompatible with the radical behaviourism of Skinner (1957) and others which dominated many branches of psychology at the time of Chomsky's early work. Their basic claim was that behavioural sequences were brought about by simple causal chains in which a given response was triggered by an immediately preceding stimulus (which could, in verbal behaviour, be internal feedback from the preceding word uttered). Today, Lashley's (1951) classic argument for hierarchical models of sequential behaviour, in which the whole sequence is,

in crucial respects, mapped out in the mind before any behaviour is initiated, is the basis of standard theory (see, for instance, Martin 1972, and Shaffer 1976). A speaker is able to produce grammatical sentences precisely because he is able to represent each sentence as a unified structure in which the parts have the kinds of relationships between them which are exemplified in a deep-structure tree. Similarly, as Schenker would claim, a composer is able to produce a masterwork precisely because he has an intuition of the *Ursatz* underlying it, which guides and unifies the process of generating the individual notes.

As this chapter progresses I shall return to Chomsky and Schenker on several occasions. I will not, however, be able to provide a detailed analysis or critique of either transformational grammar or Schenker analysis. To do so would require lengthy excursions into linguistics and musicology. For a useful introduction to Chomsky's work the reader is referred to Lyons (1970). Greene (1972) and Fodor, Bever, and Garrett (1974) provide psychological assessments of transformational grammar. A criticism of Chomskyan linguistics may be found in Robinson (1975).

A difficulty with approaching the work of Schenker is that good English translations of all his major writings are not yet available. The major English-speaking musicologists to employ Schenkerian concepts are Salzer (1952) and Forte (1962). Yeston (1977) has made a collection of articles on Schenker, and a recent critical approach is provided by Narmour (1977). It is worth emphasizing that Schenker's work does not occupy the central position in musicological thought that Chomsky's work does in linguistics. Schenker did not present his ideas with the formal precision of Chomsky, and their application to anything other than mainstream classical compositions is problematical.

### 2.3. Other comparisons between language and music

Although the Chomsky–Schenker parallel seems one of the most striking links between language and music, it is not, of course, the only one. Before reviewing selected research in detail I would like to list, and comment upon, some of the other major similarities between language and music.

(*a*) Both language and music are characteristics of the human species that seem to be *universal* to all humans and *specific* to humans. To say that language and music are universal is to say that humans have a general capacity to acquire linguistic and musical competence. There may, of course, be special reasons why particular individuals do not acquire competence (e.g. brain damage, lack of appropriate experience). The claim that language is *specific* to humans has been the subject of psychological controversy. Clearly, there seems to be no other species whose natural communication shares all the features of human language, but attempts to teach chimpanzees a modified human language (Gardner and Gardner 1969) have met with enough

success to sustain lengthy argument in the psychological literature. On balance, most psychologists would conclude that such studies, whilst possibly forcing us to reassess our ideas about chimpanzees' mental capacities, do not seriously challenge the specificity hypothesis. Although chimpanzees have learned modest vocabularies of sign-language words, they use and combine them in ways which do not go significantly beyond the capacities of a normal two-year-old human child, and the human effort required to produce and sustain this behaviour is prodigious.

The specificity hypothesis for music has not been subject to the degree of psychological assessment that has been devoted to language. However, it is clear that the functions of music for man find no parallel in the animal world; many of the most highly patterned sound behaviours (e.g. bird song) are relatively rigid intra-specific signals of territory, aggression, warning, etc. I shall discuss more fully in Chapter 7 the question of whether non-human primates display capacities in virtue of which man's musical origins could be explained. Certainly, no serious scientific effort has been made to exploit such capacities. For now, a quote from Williams (1980) will suffice to typify the generally accepted conclusion about non-human music. 'In the vocalising of apes we do not find the elements, nor even the beginnings of the chant and dance ritual of primitive man.'

(*b*) Both language and music are capable of generating an unlimited number of novel sequences. People can produce sentences they have never heard before, and composers can write melodies which nobody has ever produced before. Animal 'languages' in contrast usually contain a finite (and small) set of utterances which exhaust the communicative repertoire.

(*c*) Children seem to have a natural ability to learn the rules of language and music through exposure to examples. Spontaneous speech and spontaneous singing are first exhibited at around the same age (between one and two years). Language development proceeds through intermediate forms to model adult grammar at about the age of five. The studies reviewed in Chapter 6 show an analogous progression for music.

(*d*) The natural medium for both language and music is auditory-vocal. That is, both language and music are primarily received as sequences of sounds and produced as sequences of vocal movements which create sounds. Thus, many of the neural mechanisms for analysing input and producing output must be shared. The most universal of all musical forms is the song, where words and music are intimately combined.

(*e*) Although the auditory-vocal mode is primary, many cultures have developed ways of writing down music. In such notational systems the writer expresses the sequence he wishes to communicate as a set of written symbols, and the receiver retrieves the message by decoding the visual symbols. Although most people acquire the ability to understand and use spoken language without specific tuition, reading and writing has usually to be taught *after* a person is a competent user of spoken language. Many cultures have

no notational systems for their language or music, and many individuals in literate cultures have difficulty in acquiring literacy. Adult illiteracy is acknowledged to be a major social problem in nearly all literate cultures, and inability to sight read bedevils a great number of music performers (see Chapter 3). The inability to deal with notation is all the more striking when, as is the case with standard Western notations, the grammatical structure is actually made more explicit in the notated form than in the spoken form. For instance, English orthography marks each word with a gap, and each major segment with a punctuation mark; whereas in speech words and clauses run together with no explicit marker. In standard musical notation key, metre, phrasing, and other structural information is explicitly signalled by notational elements; whereas in the auditory mode the listener must frequently work this out from the note sequence itself (see this chapter, section 5, and Chapter 5).

Although literacy is hard to acquire, its acquisition can profoundly alter cognitive functioning. Activities such as the accurate memorisation of long passages, or extended composition, would be impossible without it. In Chapter 7, I shall be exploring some of the ways in which literate and non-literate musical activity differ.

(*f*) Receptive skills precede productive skills in development. Children can understand sentences using certain constructions before they can invent sentences using those same constructions. In music, children are able to respond to musical devices before they can use these devices to make their own music. Generally, most adults retain a severe production deficiency in music. Although they may be able to learn to reproduce specific musical sequences, and although they may develop complex analytic listening abilities, they can be totally incapable of producing novel musical sequences of a style and complexity similar to those they are competent to deal with receptively. So, for instance, a listener may have a deep knowledge of Mozart's symphonies, as evidenced by his ability to recall, describe and classify them, or to analyse their structure, yet he may be totally unable to generate a theme, let alone a whole work, which is original yet 'Mozartian'.

(*g*) The forms taken by natural language and natural music differ across cultures, but some universal features constrain these forms. The fact that there are different forms has the consequence that people familiar with a particular form are often unable to deal appropriately with other forms. Thus, a native English speaker is unable to understand Chinese. Similarly, someone whose musical culture has consisted of English nursery rhymes is unlikely to be able to understand Tibetan chant. In both language and music, there are varying degrees of formal difference. For instance, spoken Slovakian is quite similar to spoken Polish, and so a native Pole will have partial success in understanding a Slovak speaker. Similarly, Mozart has more in common with nursery rhymes than Tibetan chant, so someone familiar with nursery rhymes will have partial success in understanding

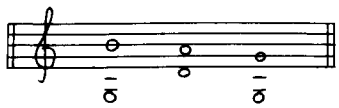
Mozart. Formal differences arise because there is a certain degree of arbitrariness about the precise sounds used to realize linguistic and musical forms. In language, any pronounceable set of sounds will do as a word, provided that all users know its function. Similarly, there are many ways of signalling grammatical dependency, from word ordering to inflection.

Given the diversity in linguistic and musical forms, it seems initially puzzling that one should wish to claim that there are universal features nonetheless. One form of the argument for linguistic universals runs something as follows. The role of language is to express thought. The form of human thought is innate and common to all humans. The deep structure of an utterance is closely related to the thought it represents (whether or not is is identical *to* the thought need not concern us here). In order to represent a thought faithfully, the deep structure is constrained in certain ways. In other words, because all human prelinguistic thoughts have the same type of form, then all linguistic deep structures which represent them must have the same kind of form. In order to turn the deep structure into a sentence, the speaker must transform it into a linear sound sequence which supplies the necessary information to a listener. These transformations are thus doubly constrained; by the nature of the vocal apparatus shared by all humans (which determines the general surface appearance of an utterance); and by the deep structure which they must carry through to the surface. This argument draws in many complex issues in philosophy, linguistics, and cognitive psychology. For instance, if language expresses propositions, does this imply that human thought is propositional? If so, what role does imagery play in thought (Pylyshyn 1973, Kosslyn and Pomerantz 1977)? Rather than pursue these issues, our purpose here is to see what a similar argument might look like for music.

The first question we must ask is whether there is any entity which bears the same relationship to a musical sequence as a thought bears to a linguistic sequence. A thought is not, in itself, a linguistic sequence on the argument we have outlined. It exists independently of language and could be entertained by a non-linguistic or pre-linguistic human. Is there any form of mental activity which could take place in a mind without musical knowledge that could be somehow *expressed* by a musical sequence? Such activity would be, precisely, one which could find musical expression in such natural but diverse forms as a Tibetan chant or a nursery rhyme. One suggestion is that the mental substrate of music is something like that which underlies certain types of story. In these stories a starting position of equilibrium or rest is specified. Then some disturbance is introduced into the situation, producing various problems and tensions which must be resolved. The story ends with a return to equilibrium. The underlying representation for music could be seen as a highly abstracted blueprint for such stories, retaining only the features they all have in common. The learning of a musical language could then be seen as the acquisition of a way of representing these features in sound. Maybe,

therefore, we should look more closely at Schenker's *Ursatz* for insight into the possible nature of universals; for, as a deep structure, it is likely to have a close resemblance to the underlying thought representation of music.

The *Ursatz* is derived from a tonic triad. The tonic triad is a fundamental musical structure built on the natural harmonic series. Leaving aside octaves, which we may hold to be harmonically equivalent, the triad is the chord which results from combining the three lowest distinct harmonics of a note (the second, third, and fifth harmonics). If the fundamental note is C then these harmonics are C, G, and E, respectively (the fourth harmonic is a further C). For Schenker, this tonic triad represents the ultimate resting place in music, the place from which it starts and to which it returns. It has this function precisely because it is present in many natural periodic sounds (*Klang in die Natur*). When an elastic medium of any sort is set in periodic motion, it will produce sound waves corresponding to twice, three times, four times, the lowest (fundamental) frequency of vibration, and so on (see Jeans 1937, for an introduction to the physics of sound in relation to music). These sounds constitute the harmonic series, although higher harmonics are usually very faint compared to the fundamental and lower harmonics. The tonic triad is 'natural' in a very real sense because it is audibly present in most natural sounds.



Example 2.3. An *Ursatz*, the fundamental Schenkerian structure.

If we examine an *Ursatz* such as that given in Example 2.3, we find that all its notes are contained in the tonic triad (of G major in this case) except the middle note of the upper line (the *Urlinie*). At the midpoint of the *Urlinie* we thus find a departure from the resting position which is established at either end of the *Ursatz*. Tension and discord have been introduced; but it is *motivated* tension. One may argue that in good stories neither the tensions nor their resolutions are arbitrary. We find it unsatisfactory when the author introduces some *deus ex machina* to extricate the hero from a seemingly impossible situation. We prefer it when the kernel of the solution is somehow implicit in what has gone before. For instance, the villain's evil designs have within them them the seeds of their own destruction; the internal dynamics of a relationship lead the partners to the brink of breakdown and also provide the final resources to save it; and so on. Similarly, the *Ursatz* satisfies because it is not just *any* note (say F) which introduces tension. It is, in this case, an A which has two highly important pivotal functions. Firstly, it creates a linear progression in the *Urlinie*, B–A–G. The line has its own logic or pattern (two consecutive linear descents of one scale step) so that, in one sense, the A becomes an inevitable consequence of travelling from the B to the G.

Secondly, it creates, together with its accompanying bass note, the elements of a new triad based on D (A is the third harmonic of D). The tension-inducing element thus operates by attempting to set up a 'rival' triad. In the final chord of the *Ursatz* we witness the 'defeat' of this rival system. Let us, then, hypothesize that one appropriate 'deep' universal for musical thought is to be summarized in the phrases 'creation and resolution of motivated tension'. This notion has a family resemblance to the 'implicative' theory of L. B. Meyer (see this chapter, section 2.5).

What, if anything, can we say about the universal features of music brought about by 'surface' constraints, i.e. the constraints dictated by the fact that music is primarily vocal, and thus melodic? Firstly, since the deep structure has discrete phases, then the surface constituents must display discreteness or discontinuity, whether at the level of 'episode' or 'note'. At the very least, music must contain two discrete aspects of a sound event which begin 'in unity' with one another, move to a state of 'disunity' and then return to unity.

Second, there must be sound dimensions along which distance can be measured. The tension-inducing elements must be more distant from the initial elements along one of these dimension than the final elements. So, clearly, linear time cannot be this dimension, because the middle tension-inducing elements will always be closer in time to the beginning of the music than will the resolving terminal elements.

Third, sounds must be capable of being classified along a *second* independent dimension of sound representing distance. This is so that *motivated* tension can be introduced. As we have seen, motivated tension is achieved when an element is discrepant on one dimension (e.g. harmony) but unified on another (e.g. scale), thus displaying ambiguity. The system of tonality may be seen as one particularly satisfactory solution to the problem of representation that is consistent with these constraints. Within the tonal system there are two ways of representing the closeness of two notes with definite pitch. One is scalar: notes are close if their pitches are as near to one another as the scale allows. The other is harmonic: notes are close when their pitch *ratios* are those between low harmonics in the harmonic series. Thus, a given note can be close to another on the scalar dimension whilst being distant on the harmonic dimension, and *vice versa*. It is possible that one reason for tonality's success is that its dimensions of closeness are ones which it is easy for people to work with. We return to this issue in greater detail in Chapter 7.

(*h*) It is common to consider a human language as comprising three components: phonology—a way of characterising the basic sound units of a language; syntax—the rules governing the way in which sound units are combined; and semantics—the way in which meaning is assigned to sound sequences. Music seems to break down quite naturally into the same three

components, and the following sections take a more detailed look at research on these three aspects of language and music.

## 2.4. Musical phonology

All utterances of a natural human language like English can be analysed into a small set of basic speech-sound classes called 'phonemes'. The characterization of a phoneme is somewhat technical, but many phonemes correspond directly to letters. Thus, the word 'cat' has three phonemes, one for each letter. Each phoneme can be thought of as a certain sound pattern, having certain frequency and duration parameters, typically produced by a characteristic combination of lip, tongue, and vocal chord motion. The continuing debate about the extent to which phoneme characteristics are invariant in different sound environments need not concern us here (Cole and Scott 1974; Liberman, Cooper, Shankweiler, and Studdert-Kennedy 1967). Two points about phonemes are of special significance. One is that most phonemes are characterized by a *range* of sounds along particular dimensions rather than totally specific values. The other is that different languages divide up the sound continuum in different ways. Thus, two sounds which are distinct phonemes in one language may be heard as the same phoneme in a different language. Hence the well documented difficulty encountered by people trying to learn a language having substantially different phonology to their own. They simply fail to hear important distinctions. The human tendency to perceive speech *categorically* provides much of the explanation for these phenomena.

Categorical perception is best described by means of a specific example. It appears that the principal determiner of whether one hears a consonant as voiced (e.g., /d/) or unvoiced (e.g., /t/) is the time relationship between the onset of two components of the sound known as the first and second formants ($f_1$ and $f_2$). Formants are bursts of sound occupying discrete frequency bands. A typical speech sound will have four such bands. The first formant is the name for the lowest frequency band present, the second for the next highest, and so on. Such sounds can be synthesized by machine, and it appears that the optimal /d/ is heard when the onsets of $f_1$ and $f_2$ are simultaneous. In contrast, the optimal /t/ is heard when $f_1$ commences about 60 ms after $f_2$. With a machine it is possible to synthesize sounds with onset delays intermediate between 0 and 60 ms, and then discover how listeners classify them. If we generalize from other perceptual studies where one dimension is varied between two extreme values, our expectation would be that as onset delay increased from 0 to 30 ms, listeners would hear a sound less and less like /d/. When the midpoint was reached we would expect to hear an ambiguous sound, and then, as the delay approached 60 ms, the sound would become more and more like /t/.

The actual result of such an experiment is very different (Liberman, Harris, Kinney, and Lane 1961). Up to about 20 ms delay there is almost no change in the perceived sound. Subjects identified nearly all the sounds as absolutely clear /d/s, indiscriminable from the case where the delay was zero. Then over a very small range of delay (about 10 ms) perception 'flipped over' to /t/. Further increases in delay produced no perceptible change, all sounds beyond this boundary being indiscriminable from the case where the delay was a full 60 ms.

It has been found that this type of effect is demonstrable for nearly all phonemic distinctions, and this shows that the level of analysis of which we are habitually aware has discarded much of the specific information present in the sound. From sounds which vary continuously on a number of dimensions we extract a few categories into which all normal speech sounds are assigned. These are the categories which feed onto higher level syntactic and semantic analyses.

It appears that we acquire the categories of our native language very early in life. A set of studies by Eimas and his colleagues (Eimas 1975; Eimas, Siqueland, Jusczyck, and Vigorito 1971) has demonstrated that three-day-old infants already categorize sounds in the same way as adults. This precocious ability strongly suggests the existence of special learning mechanisms for speech patterns, since—of course—infants categorize differently according to the language they are exposed to in the first few days of life.

These features of language seem to have some rather close musical parallels. The basic 'phoneme' of music is a 'note'. Like a phoneme a note is characterized by frequency and duration parameters. Within a particular musical culture, all music is composed from a small set of these notes, chosen from an indefinitely large set of possibilities. Different cultures, however, choose different subsets of possible notes for their music. The selection takes place along two dimensions of sound: frequency and duration; these merit separate discussion.

### 2.4.1. *Categorical perception of frequency*

The perceived pitch of a musical sound is, in most naturally occurring instances, determined by its fundamental frequency of vibration. The exceptions to this will not concern us here. Furthermore, we will assume that octaves are musically equivalent in the context of this discussion. The 'phonology' of pitch thus concerns the way in which the octave is subdivided. The subset(s) of pitches chosen by a musical culture constitutes the musical scale(s) of that culture.

There is some evidence that people brought up within one musical culture have difficulty in perceiving and reproducing scales of another culture. It is hard to find convincing contemporary evidence for this since the Western diatonic scale has infiltrated most cultures through the broadcasting media. There is, however, some ethnomusicological data which shows how tunes

which have migrated from one culture to another have been adapted to the new scale (Cohen and Cohen 1973). Hodeir (1956) has put forward the interesting hypothesis that the 'blue notes' in the jazz scale came about through the efforts of African musicians (from a culture using a pentatonic scale) to assimilate the diatonic scale of the North American culture. If we represent the diatonic scale as a sequence of note names (e.g. C, D, E, F, G, A, B) then the proposed pentatonic scale would use only C, D, F, G, and A. Users of the pentatonic scale would have no normal representation for the diatonic E and B of this example. When attempting to reproduce a Western melody using these notes, Africans would have heard them as 'mistuned' notes falling somewhere between D and F or A and C. Accordingly they would 'fish' for these notes when playing or singing, producing something which would sound unstable and mistuned to Western ears. Gradually this accidental feature would become formalized as part of the jazz style—the flattened third and seventh. One must emphasize that this is an unproven hypothesis and does not receive universal acceptance (Virden and Wishart 1977).

The psychological principle underlying such examples appears to be the categorical nature of pitch perception. It has been demonstrated in laboratory conditions by Locke and Kellar (1973). They presented listeners with triads in which the tuning of the middle note (the third of the chord) was systematically varied from major to minor, passing through several intermediate pitches. The outside pitches were always 440 Hz (A) and 659 Hz (E). The middle note varied between 554 Hz (C sharp) and 523 Hz (C natural). There were two experimental tests. In one, subjects heard a single chord (the standard) and then, after a short delay, were required to say whether each of a sequence of chords was the same as, or different from, the standard. In the other test, subjects heard pairs of chords in which the middle notes differed by a few Hz at most. They were required to say whether the chords were the same or different. In both tests, musician subjects showed clear evidence of categorical perception. Fig. 2.4 shows the mean responses for 15 musicians. The hatched line shows the result of the first test when the standard was 554 Hz (A major chord). Almost all chords with middle notes above 546 Hz were heard as A major. Almost all chords with middle notes belowo 540 Hz were heard as A minor. The evidence suggests a category boundary at about 542 Hz.

The data from the second test (full lines) corroborates this picture. The four lines are for different degrees of separation of the test pitches. In each case there is a clear peak in discrimination at the category boundary. That is, when members of the pair of chords straddle the boundary, discrimination is good. When they are both on the same side of the boundary, discrimination is poor.

Interestingly, non-musicians did not show such clear evidence of categorical perception. The discrimination functions have peaks at around

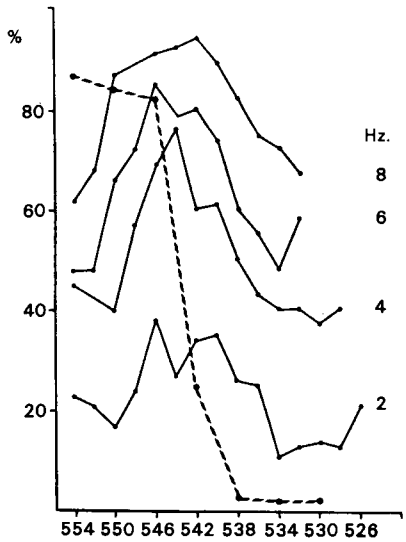



Fig. 2.4. Categorical perception in musicians (from Locke and Kellar, 1973).

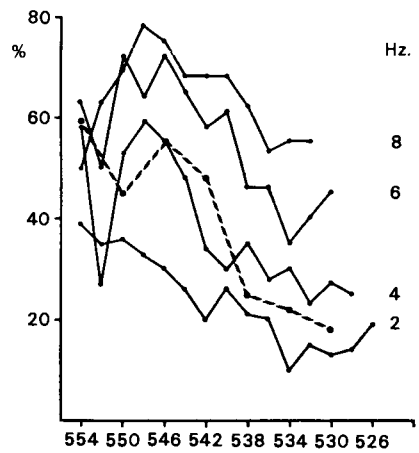



Fig. 2.5. Categorical perception in non-musicians (from Locke and Kellar, 1973).

Figs. 2.4. and 2.5. Key: the dashed line gives the data from the categorization test; the full lines give the data from the discrimination test at various pitch separations; the abscissa represents the frequency of the middle note of the chord; the ordinate represents percent correct judgement.

546 Hz, but the comparisons to a standard show a more smoothly declining function: the greater the pitch difference, the less likely a 'same' judgement (see Fig. 2.5).

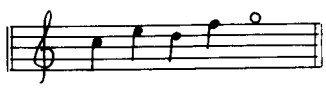
This study raises several issues. Firstly, there is the question of the accessibility to listeners of the *uncategorized* frequency information. In the speech studies it would appear that this information is not normally available to conscious perception. In music, however, this information certainly *can* be made available to consciousness. If it were not so, then no chord could ever sound badly tuned—the assimilation to categories would 'complete' the perceptual experience. The music listener, we must conclude, has some ability to operate both within and outside the categorical mode.

A second question is, therefore, what perceptual contexts encourage categorical perception? For most listeners the likely answer is that the context must be one which provides a framework for the construction of a musical scale. This framework will supply, at the very least, two simultaneous (or closely consequent) notes, such that one can act as a 'tonic' for the other. The notes of a musical scale are defined not absolutely but relatively, so that a scale can be constructed on any pitch using a characteristic set of frequency ratios to generate the notes. Thus, what music listeners carry in their memories are not the absolute pitches of any particular scale, but procedures for generating a scale from any given tonic. In Locke and Kellar's study this framework was supplied by the invariant outer notes of the chord which maintained a 3:2 frequency ratio throughout. This identified them, within diatonic tonality, as the tonic (first step) and dominant (fifth step) of a diatonic scale. This framework imposed a categorization on the middle note as the submediant (third step) of either a major or a minor diatonic scale. Had the experiment been carried out using only a single variable note without chordal context, then it is unlikely that any categorization would have taken place. Data from normal psychophysical studies support this assertion. There is no evidence of discontinuity in the discrimination functions for single frequencies.

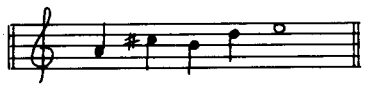
A third question concerns the genesis of categorization within the individual. The music data suggest that, unlike language, mere exposure to tonal music is insufficient to bring about categorical perception. Some aspect of musical training heightens the tendency to categorize. For instance, most musical training involves learning note names and scale terminology. It is possible that possessing verbal labels increases the likelihood of categorical information being extracted and stored. Another possibility is that, for some musicians, each category becomes associated with a prototypical 'absolute' frequency band. This is possible, at least in Western cultures, because there are generally agreed conventions about the precise frequencies to which instruments should be tuned. Concert A is defined as 440 Hz. Such frequencies could come to represent the 'central' positions of scale categories for listeners, with deviant pitches being assigned to the nearest prototype.

There is considerable evidence for an ability that could support such behaviour. It is called 'absolute pitch' or 'perfect pitch' and is possessed by a significant minority of trained musicians. This is an accurate long-term memory for prototypical pitches and their associated scale names; and I shall discuss it more fully in Chapter 5. We do not know, unfortunately, how many of Locke and Kellar's 'musician' subjects possessed absolute pitch.

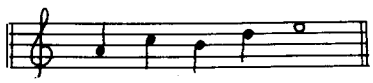
A different type of evidence for assimilation of music to scale categories is provided by Dowling (1978). In this study subjects were required to judge pairs of brief melodies as the same or different. In some cases the second melody (which was always at a different pitch from the first) was an exact transposition of the first melody. In other cases, the second melody was a 'tonal answer' to the first; that is, the melodic contour was transposed up or down *within the same key*. This has the consequence that the exact intervals of the original melody are not preserved. Example 2.4 is an illustration of the type of melodic stimulus used by Dowling. Example 2.5 is an exact transposition of it. Example 2.6 is a tonal answer starting on the same note as 2.5. Dowling found that his subjects could not consistently discriminate exact transpositions from tonal answers. One plausible explanation for this finding is that, at least for unfamiliar melodies, subjects code melodies as contours in which the *number* of scale steps between adjacent notes is represented, but not the precise pitch distance.



Ex. 2.4.



Ex. 2.5.



Ex. 2.6.

Examples 2.4–2.6. Examples of stimuli used by Dowling (1978).

### 2.4.2. *Categorical perception of duration*

One of the most well researched areas in this field concerns the perception of plucked and bowed string notes (Blechner, Day, and Cutting 1976; Cutting and Rosner 1974; Cutting, Rosner, and Foard 1976). In a carefully designed series of experiments, Cutting and his colleagues have demonstrated that variations in rise time (the time from sound onset until the time when waveform amplitude reaches its peak) are responsible for the perception of this quality, rise times of 30 ms or less giving rise to 'plucked' sounds, and those of 60 ms or more producing 'bowed' sounds. The discrimination function shows a peak at the category boundary (around 40 ms) and troughs within each category. The interest of this phenomenon is threefold. First, it exactly matches the functions for the categorical perception of a phonetic distinction

in speech; the one displayed between the words 'chop' and 'shop'. This also depends on rise times and shows a category boundary at about 40 msec (Cutting and Rosner 1974). Secondly, adaptation to a sound well within one category shifts the category boundary towards the unadapted category. This exactly matches speech perception data (Eimas and Corbit 1973). Thirdly, infants as young as two months demonstrate categorical perception for these sounds (Jusczyck, Rosner, Cutting, Foard, and Smith 1977), just as they do for many speech sounds (Eimas 1974; Eimas *et al.* 1971).

These findings have been used to dispute the view that speech perception involves unique psychological processes and mechanisms. Cutting *et al.* (1976) state that 'it is evident that the arsenal of empirical findings which once distinguished speech perception as a unique type of auditory perception is being steadily depleted.' It seems, in fact, that we possess some perceptual mechanisms which are present from an early age and are deployed in both speech and music perception to produce categorical perception. The reason why the same physical attribute (rise time) gives rise to different perceptual experience in speech and music is probably that the acoustic contexts are different. There is evidence that the nature of the immediately following vowel/note affects the way in which the rise portion is heard. For instance, the sound patterns for stop consonants (like /t/ and /p/) are heard as chirps when presented in isolation (Mattingley, Liberman, Syrdal, and Halwes 1971). They require a subsequent vowel in order to be heard as speech sounds. Similarly, Cutting *et al.* (1976) show that synthetic musical sounds must persist for at least 250 ms after the initial rise in order to be heard as plucked or bowed.

The issue of context also assumes central importance in the second major aspect of musical duration that I wish to discuss. This is the perception of the duration of a note. The results of standard psychophysical tests tell us that when two successive isolated tones are presented for discrimination of duration (i.e. subjects must say which is longer), then the longer the sounds are, the greater must be the difference between them if it is to be reliably detected (Woodrow 1951). In music, however, absolute durations are less important than the rhythmic implications which notes acquire through their immediate context. Fundamental to most Western music is the concept of a *beat*, a musical pulse which underlies any melody. In general, notes will either begin on the beat or at some simple subdivision of the beat (half, third, and quarter being common). This fact is reflected in musical terminology and notation, where there exists a limited set of categories for describing durations of notes. These categories are, for the most part, divisions of a longer category into two equal halves. Thus, there are two crotchets (quarter notes) to every minim (half note); two quavers (eighth notes) to every crotchet, and so on. In a particular performing situation, one of these symbols may be defined as having a particular duration (for instance, 'crotchet = 120' is a standard way of indicating that there should be 120 crotchet beats per minute). Thus,

all the other symbols acquire a defined duration which is exactly double, or half that of the standard. None the less, in any human performance it is certain that there will be significant deviations from these defined durations.

In all empirical studies which have measured times between the beginnings of successive notes, such deviations have been found in abundance (Michon 1974; Seashore 1938; Shaffer 1981b; Sloboda 1983), even when the performer has been asked to produce absolutely strict rhythms (Gabrielsson 1974). None the less, identification of intended rhythm is a commonplace accomplishment for listeners, who are continually faced with the potentially confusing phenomena of *rubato* and gradual changes in speed. In contrast, accurate perception of deviations from metricality is difficult, and requires much specific training. It is almost impossible for one performer to imitate another *exactly*. All this strongly suggests that listeners achieve a categorization of the durations of the notes they hear into crotchets, quavers, etc.

In the light of these considerations it becomes a little easier to understand some puzzling data provided by Sternberg, Knoll, and Zukofsky (1982) who showed that three highly trained professional musicians, including Pierre Boulez, were unable to reproduce non-standard subdivisions of a beat accurately. In the experiments subjects heard a series of regular beats at one second intervals, one of which was followed by a click at a delay ranging from one eighth to seven eighths of a beat. They were then required to place an equivalent click after a subsequent beat. The reproduction of the shorter subdivisions (less than one third of a beat) were systematically in error, all being overestimates. In contrast, reproduction was very accurate when the subdivision was half, three quarters, five sixths, or seven eighths of a beat. A similar pattern of results was obtained when the subjects were asked to *estimate* the delay of a click by giving a verbal categorization (e.g. 'between one eighth and one seventh of a beat'), except that in this situation the errors were *under*estimates. Why were these subjects so poor? Maybe accurately reproduced and estimated delays correspond to frequently occuring rhythmic patterns in music, which are readily categorized. In Fig. 2.6 I have written out the experimental stimuli in musical notation. The first, second, and fourth notes in each example correspond to the regular beats, and the third note corresponds to the click. The lower half dozen notational forms will be familiar to most performing musicians. It is hard to think of well-known examples of the upper four. These are precisely the ones that caused difficulty for the subjects used by Sternberg *et al.* Perhaps they are particularly hard to perceive as examples of distinct rhythmic categories.

Just as in the discussion of categorical perception of pitch, one would not wish to claim that categorical perception of duration makes finer temporal discriminations impossible. We *can* hear rhythmic imprecision and *rubato* with appropriate training, but fine differences in timing are more often experienced not as such, but as differences in the quality (the 'life' or 'swing')

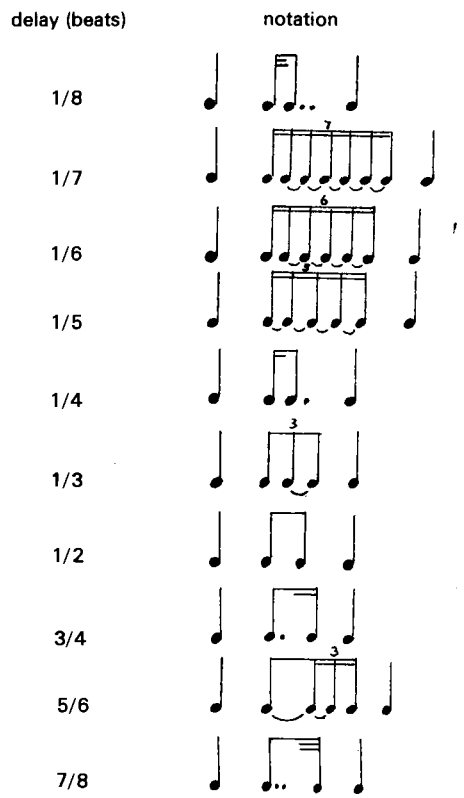



Fig. 2.6. Stimuli used by Sternberg, Knoll, and Zukofsky (1982), expressed in conventional rhythmic notation.

of a performance. This accounts in part for the intangibility of performance style (see Chapter 3). Differences in style are due to real differences in performance, but many of them cannot be consciously categorized by most listeners. Most of us have only global and imprecise means of capturing our experience of style difference.

My discussion of musical phonology has been designed to illustrate one fundamental feature of music behaviour. That is, we tend to categorize our musical experience along the available dimensions of sound, giving importance to differences between categories at the expense of differences within categories. The notions of scale and metre are the fundamental concepts underlying musical phonology (although timbre and intensity are arguably additional dimensions). The similarities between language and music in this respect are striking, although categorical perception in music is neither as complete nor as universal as it appears to be in language. To understand the musical significance of categorization we must now turn to the way notes are combined with one another. This is the subject matter of musical syntax.

## 2.5. Musical syntax

### 2.5.1. *The psychological reality of grammar*

The attempt to construct formal descriptions of musical structures is a central concern of musicology (see, for example, Bent 1980). Schenker is only one of many people who have attempted to characterize the sequential regularities of music in terms of rule systems. Since music, like language, is a human product, we can legitimately suppose that the observed structure of music tells us something about the nature of the human mind that produces it. But what? An extreme view would be that the rules of a musical (or linguistic) grammar are the very psychological procedures that someone uses to generate music (or language). This view has been strongly criticized on a number of grounds, both for music (Laske 1975) and for language (Olson and Clark 1976). One of the strongest arguments is that there is no such thing as a unique grammar for any body of data. Different rules, applied in different orders, can produce exactly the same set of sentences or musical sequences. The mere discovery of a single workable grammar is, therefore, no guarantee that *this* grammar is the one that best describes psychological processes of generation. Another argument derives from the observation that humans can (and deliberately do) violate the rules which seem to account for some of their behaviour. Thus, people *do* speak ungrammatically, and composers *do* violate the formal conventions within which they work. A formal syntax or grammar is a closed system whose sole purpose is to generate the set of sequences stipulated as 'grammatical'. The psychological generation of language is, in contrast, a process whose purposes are wider. The intention to communicate a particular meaningful proposition is the psychological motivation behind most utterances, and if a speaker has good reasons to suppose that he will be understood even without correct grammar, then he may disregard grammatical conventions. In music, the intention to violate a listener's expectancies may lead to a comparable freedom with the grammar.

The modest view supported by more psychologists is that some features of grammars have implications for the general way in which we think about psychological processes. One implication of the generative approach exemplified by Chomsky's work is that we decide what to say by the application of a finite set of rules (or procedures) to some unspoken proposition we wish to utter, rather than by having to recall, by association, the correct utterance for each separate proposition. On an associationist view, it is impossible to know how to express the proposition that, say, John loves Mary until one has learned the verbal formula 'John loves Mary' and is able to recall it when required. On the generative view, a speaker may never have heard or uttered the sentence 'John loves Mary' yet still be capable of producing it to express the appropriate proposition. This is done in the following way. A person's knowledge that John loves Mary is to be characterized, roughly, as know-

ledge that a certain person, John, is the agent of some activity, loving, this activity being directed towards a recipient, Mary. The person then *constructs* the appropriate sentence by reference to rules which stipulate how general relationships between agents and objects are expressed. One form of expression is given by the formula, agent + active verb + recipient. In simple three element propositions the conditions for this formula to be applied are satisfied, so, with appropriate knowledge of the vocabulary items 'loves', 'Mary', and 'John', the speaker can utter them in the correct order. Because the number of acceptable sentences in a language like English is indefinitely large, and because we are all capable of generating acceptable sentences that we may not have heard before, it seems inevitable to conclude that we store rules for constructing sentences rather than (or at least in addition to) storing sentences themselves.

A second, and related, implication is that listeners will be able to sort utterances they hear into acceptable and unacceptable categories. Thus, most people can tell that a novel utterance is ungrammatical, even though they judge another equally novel utterance to be acceptable. Since neither utterance has prior associative meaning, it seems incontrovertible that the utterances are being assessed in virtue of their conformation to a set of rules of construction. Hard questions arise, however, when we try to specify exactly what rules are invoked. Is a generative grammar, as such, the principal mental tool? If so, then a listener would have to judge a quasi-sentence like 'John the apple ate' ungrammatical by somehow running the grammar 'in reverse' and failing to condense the sentence to a single node. Alternatively, he might try generating sentences according to the grammar and fail to come up with a match. The trouble is that either of these processes may involve an indefinite number of steps. If one used the second process one could not, strictly speaking, be certain that a sentence was ungrammatical until one had generated all possible sentences, and noted the absence of the sentence in question. Since we may argue that the number of English sentences is indefinitely large, this is clearly a very implausible psychological mechanism. More plausibly, people have 'metaknowledge' about the grammatical systems they work with. For instance, someone may know that, in English, a sentence never starts with a verb in the indicative mood. This is not itself a rule of the generative grammar (although statements of this sort are sometimes confusingly referred to as grammatical rules). It is a generalization about the *products* of the grammar. Or again, it seems likely that people use non-syntactic considerations in judging sentences. Given 'John the apple ate', a listener might attempt to work out a plausible proposition including all the elements mentioned (which would presumably entail the apple being eaten by John) and then discovering how he or she would naturally express the proposition, rejecting the given utterance if it did not match his or her own.

A third implication, and one that I wish to examine in some detail, since it relates to comparable experimental data in language and music, is that,

regardless of our conclusions about the equivalence of generative grammars to psychological processes, the *structures* embodied in such grammars have psychological reality. Briefly, this claim is that elements of an utterance which are closely linked to one another in the grammar are closely linked psychologically, in whatever form of cognitive representation a speaker or listener may make of a sentence or melody. If the structure of a sequence is best represented in a grammar as a tree of the sort shown in Figs. 2.1–2.3, then the claim would be that adjacent items governed by the same node are closer in 'psychological space' than are adjacent items governed by different nodes. Figure 2.7 shows an idealized constituent tree which might equally well be a proposed structural tree for a linguistic or a musical sequence containing eight elements. The nodes are indicated by letters, the terminal elements that constitute the actual sequence are indicated with numbers. Taking, as an example, elements 3 and 4, we see that they are very closely linked in the structure. They are both subsumed under node E, and more remotely by node B. In constrast, elements 4 and 5, whilst also being adjacent members of the sequence, are not closely linked in the structure. Element 4 is 'governed' by node B, whilst element 5 is 'governed' by node C.

An experimental technique developed to investigate this claim involves subjects listening to short well-formed sequences of language or music over which an audible click has been superimposed, at a particular point. Subjects are asked to indicate exactly when they hear the click to occur by reporting the syllable or note which seemed to be simultaneous with it. The initial results of applying this technique have been reported for language by Fodor and Bever (1965) and Ladefoged and Broadbent (1960); and for music by Gregory (1978) and Sloboda and Gregory (1980). The basic findings in both domains are easily summarized. Firstly, click location is quite inaccurate. Secondly, clicks appear to migrate perceptually to boundaries between structural groupings. For instance, in the sentence 'That he was happy was evident from the way he smiled', a click occurring during the fourth word tended to

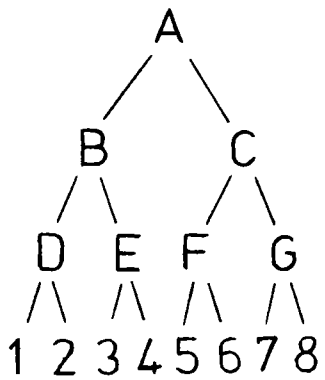



Fig. 2.7. An idealized constituent tree structure.

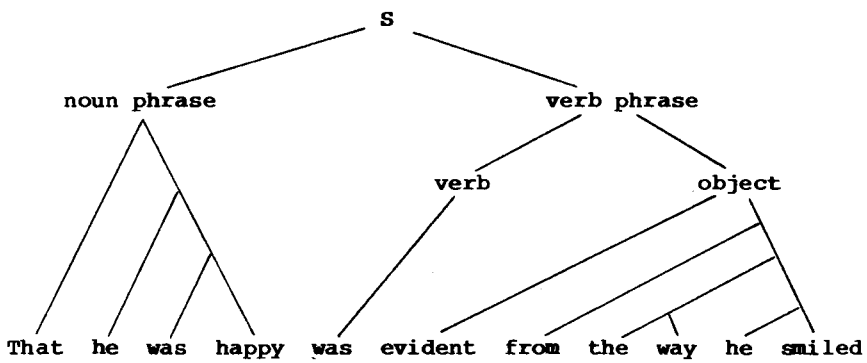


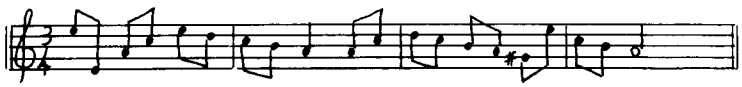

Fig. 2.8. Constituent tree analysis of a sentence used in Fodor and Bever's (1965) study.

be heard as coming later than its actual time of occurrence. If the click occurred during the fifth word, it was heard early. A constituent analysis of the sentence (Fig. 2.8) assigns the first four words to the major node NP (noun phrase) while the rest of the sentence makes up VP (the verb phrase). In structural terms, therefore, words 3 and 4 are close, as are words 5 and 6. Words 4 and 5, though, are not structurally close. The experimental results show that these structural relations *are* represented psychologically. It seems that there is a tendency to represent closely related elements as 'uninterrupted', so that an extraneous event (the click) gets pushed towards a point where the structural links are weak.

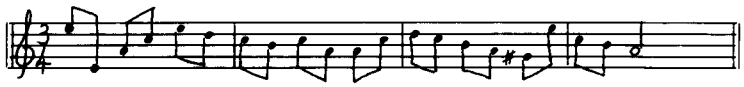
A pictorial analogy of the process may be obtained if one imagines the words (or notes) of a sequence to be laid out in a line on a surface. Each major structural grouping is covered by a smooth hemispheric dome. The click can then be envisaged as a small pellet dropped on to the surface from a position vertically above its real location in the sentence. When it hits the surface of the dome its final position will be determined by the slope of the dome at that point. Pellets falling near the apex of the dome are likely to be displaced rather little; pellets falling near the edge are likely to roll off into the space between.

The music results are to be explained in a similar way to the language results. Here, the main structural unit is the musical phrase. In Sloboda and Gregory's (1980) experiment, each melody contained two metrically equal phrases, and, in one condition, equal durations for all the notes. Example 2.8 shows one such melody. Musicians generally agree that the first phrase ends on the fourth note of the second bar. In this experiment the notes were computer-generated and exactly equal in all respects other than pitch. In physical terms, therefore, notes 4 and 5 of the second bar are as close to one another as notes 5 and 6 (arguably even closer since they are at the same pitch). Structurally, however, notes 4 and 5 are distant. A click placed on note 5 had a higher tendency to 'migrate' backwards than any other placement of

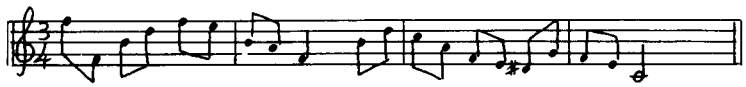
the click, confirming that listeners were representing the phrase boundary in a special way. The migration tendency was enhanced by melodies which marked the phrase boundary with a long note (see Example 2.7) but reduced by melodies which did not have a 'normal' harmonic structure (Examples 2.9 and 2.10).



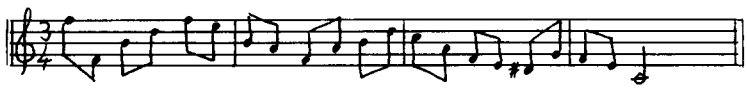
Ex. 2.7.



Ex. 2.8.



Ex. 2.9.



Ex. 2.10.

Examples 2.7–2.10. Examples of stimuli used by Sloboda and Gregory (1980).

The precise interpretation of such 'click' experiments has been the subject of some controversy in the literature (Olson and Clark 1976). For instance, Reber and Anderson (1970) presented subjects with sentences over which *no* click had been superimposed, but told their subjects that very faint clicks were present. Subjects were asked to detect the location of the 'clicks'. They tended to locate clicks at phrase boundaries. Thus, argue Reber and Anderson, the earlier results are explicable by a bias subjects display to place their responses at phrase boundaries rather than by a perceptual migration. However, it may be argued that a tendency to organise *any* aspect of cognitive behaviour, be it perceptual or response-orientated, on the basis of phrase structure is good evidence for the psychological reality of grammatical structures. More worrying, perhaps, are studies like those of Bond (1971) and Lehiste (1972) who found that non-grammatical cues, like pauses and stresses, were more important than syntax in determining the nature of the migration. In Sloboda and Gregory's study pauses (long notes) were also found to elicit click migration. However, in both these cases, no harm is done to the general thrust of our argument if it is acknowledge that pausal and stress patterns in speech and music can be generated by a grammar along with the word ordering (see Chomsky and Halle 1968, for langauge; and Martin 1972, and Sundberg and Lindblom 1976, for music). In other words,

it may be *because* they occur in a systematic syntax-related fashion that stresses and pauses become psychological 'magnets' for clicks in individuals with long experience of a particular grammar.

'Clickology' is not the only way that the psychological reality of grammatical structures can be demonstrated. For instance, if phrases are psychological units, then one might expect processing within a single phrase to proceed more rapidly or effectively than processing which requires integration of information from two phrases. This prediction has been confirmed by Green and Mitchell (1978) who showed that language readers' pauses are longest after fixations which complete major grammatical units. Also, Levin and Kaplan (1970) have shown that the eye-voice span in language reading tends to extend to a phrase boundary, and Sloboda (1974, 1977a) has shown the same for music reading. Span, here, is defined as the number of words (or notes) that a reader can produce after the text has suddenly and unexpectedly been removed. It is a measure of how much has been taken in ahead of what is actually being said (or played). The findings suggest that readers take in material up to the boundary of the current phrase quite soon after starting to read it, but then 'mark time' until the execution of that phrase is almost complete before sweeping rapidly on to the next boundary.

In the discussion so far I have been deliberately vague about musical syntax and its nature. I wished, for purposes of exposition, to let it be assumed that grammars of roughly equivalent power and sophistication had been discovered for both language and music, so that some central psychological findings could be introduced. I have now to confess that, to my knowledge, there have been very few attempts to produce full-blown generative grammars for any significant corpus of music, even in outline. In what follows I shall discuss the most significant psychologically oriented work in this area in order to help us understand what kind of structures we use to mentally represent music. This will bring to light some similarities with language, but significant differences too.

### 2.5.2. *Syntactical structures in music*

In attempting to produce a musical syntax, an important question to be answered is '*what* body of music do we wish the syntax to account for'? The answer to this question is not as clear cut as the same questions asked of language. Most people have only one language at their disposal, whose grammar is fairly stable over long periods of time. The English language is delimited as containing all utterances which the majority of English inhabitants judge to be acceptable and meaningful. In contrast, the forms and styles of music with which a given person may be familiar can be very diverse. Musical forms change rapidly, and there is heated discussion within our culture about what does and does not count as acceptable music. It is interesting to speculate on the reasons for this difference. One possible reason

is that the communicative functions of language favour unity and stability. Syntax is a vehicle for communicating knowledge about the world, and given that the world remains the same sort of place and human concerns remain the same, there seems little to gain and much to lose by diversity and rapid evolution of syntax. Art music, in contrast, has no such clearly defined function. Syntax becomes, in itself, an object of aesthetic awareness, and the pressures for novelty invite diversity and change. It is arguable that music shows the greatest stability in cultural contexts where it is the vehicle for some non-musical activity (the recounting of verse, ritual, etc.).

The most cautious approach to musical syntax would be to start from a particular small sub-area in music, say the works of a single composer, and attempt to discover a grammar adequate for this corpus. The next step might be to extend the domain of the grammar by discovering if other compositions or other composers conformed to it. In this one might be guided by historical and social factors (looking at composers who lived at the same time and were subject to the same sort of influences) and also by the subjective judgements of experienced listeners (judgements of which types of music sound most similar). On the other hand, it is commonly claimed that very large bodies of music share certain aspects of syntax. If this is so, such aspects should be given the greatest prominence. A piecemeal approach to small areas of music can, arguably, obscure wider patterns (e.g. basic features of harmonic, melodic, and rhythmic construction which are common to the vast bulk of tonal music). I shall begin with some work which takes the former, cautious approach, and then move to wider consideration about tonal music as a whole.

Sundberg and Lindblom (1976) have written a complete generative grammar for eight-bar melodies which all bear a family resemblance to a set of Swedish nursery tunes written in the late nineteenth century by Alice Tegner. These tunes are familiar to, and sung by, most Swedish children today. The outline of this grammar is given in Fig. 2.9. It resembles the generative phonology of Chomsky and Halle (1968) in two respects. Firstly, the basic grammatical structure is an hierarchical tree. Secondly, integral to the tree is a 'prominence contour'. In speech, such a contour is used to assign stress, timing, and intonation to a sentence. In this music grammar, the contour is used to generate appropriate chords and durations.

The structure of the constituent tree is given in Fig. 2.10. It is completely symmetrical and composed by binary subdivisions of each node. So, each phrase divides into two equal sub-phrases; each sub-phrase divides into two equal bars; and so on. It is noteworthy that the structure is *metrically* based (in contrast to Schenker analysis where the fundamental structure is *harmonically* based).

The prominence contour is derived from the tree by assigning left elements priority within each foot, but right elements priority within larger sub-units. The contour derived by Sundberg and Lindblom is given at the bottom of

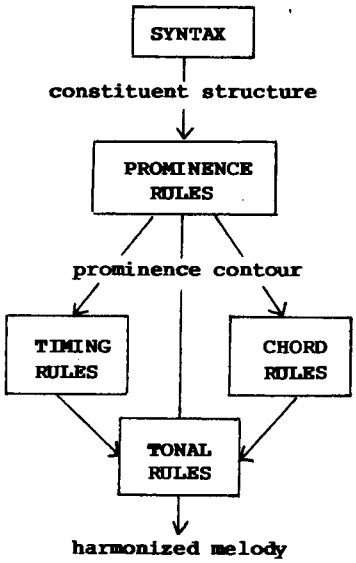



Fig. 2.9. Outline of Sundberg and Lindblom's (1976) grammar for nursery tunes.

Fig. 2.10. The numeral 1 indicates an element with the highest prominence. Elements with equal numbers have equal prominence.

The next step is to generate separately a rhythmic pattern and a chordal progression from the prominence contour. Rhythm is generated by a set of obligatory and optional rules. The most important of the obligatory rules is that no note may begin on an element *following* an element of the highest prominence. In this example, the very last element cannot contain a new note; and the penultimate element will carry the last note of the melody. Then there are other optional deletion and addition rules, with certain repetition restrictions (e.g. the second four bars must repeat the rhythm of the first four bars).

The chordal derivation rests on two main principles. Only notes with prominence of 4 or greater (i.e. 4, 3, 2, or 1) may introduce new chords; and

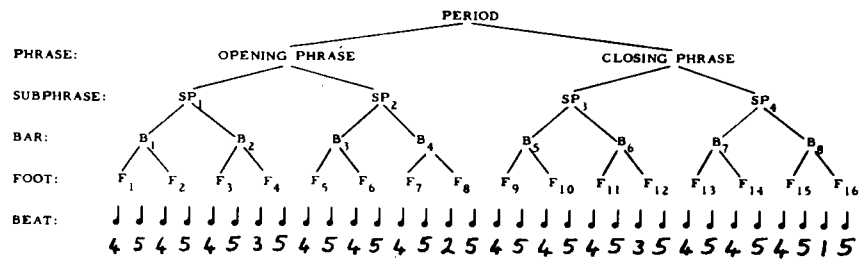
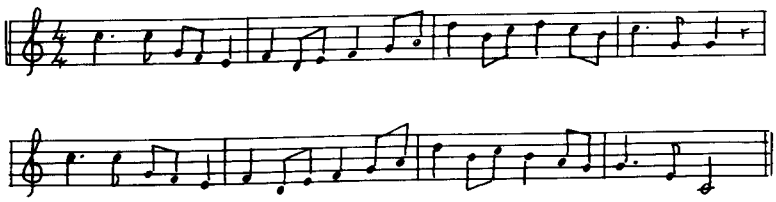


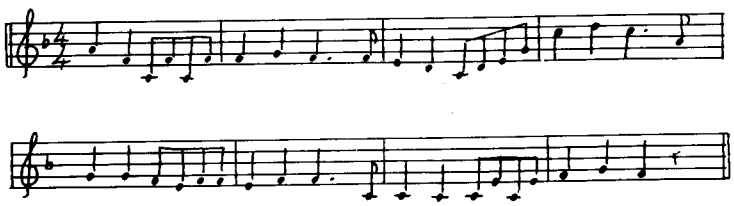

Fig 2.10. Constituent structure of tunes generated by Sundberg and Lindblom's (1976) grammar.

certain sequences of chords are obligatory at certain points (e.g. the final chord must be a tonic and must be preceded by a dominant).

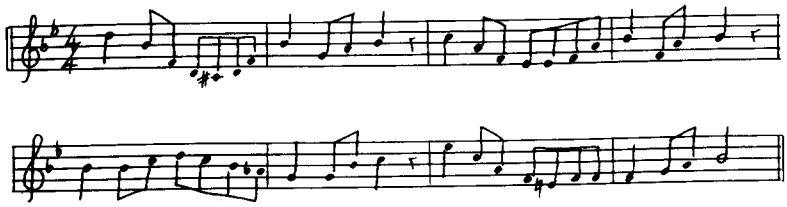
The actual notes of the melody are then generated. Notes at points of high prominence must be chosen from the chord specified at that point. Elsewhere, passing notes and suspensions may be introduced. Examples 2.11–2.16 show six melodies. One was written by Tegner. The other five were produced by the grammar, the choice among the optional alternatives being made at random on each occassion. The reader is invited to try and spot the 'real' Tegner melody before reading on.



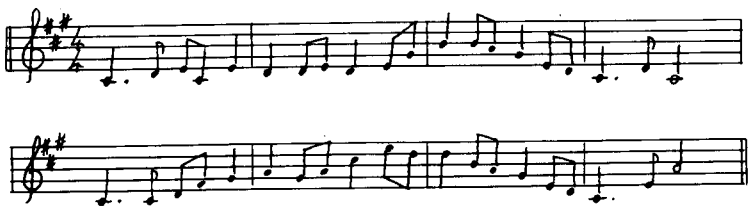
Ex. 2.11.



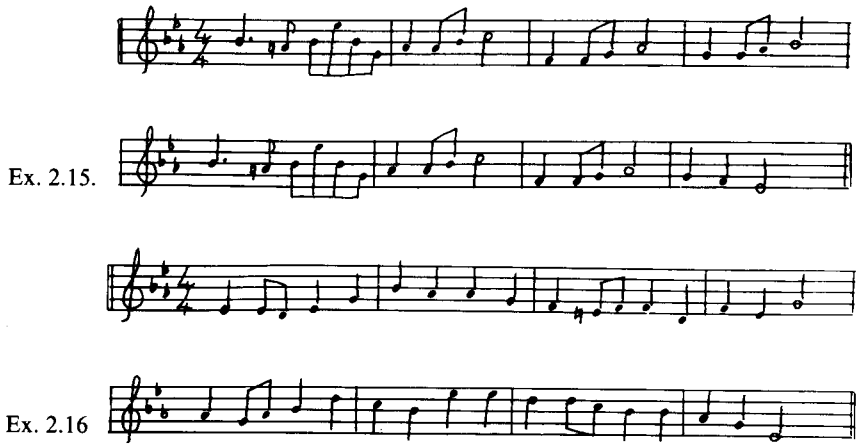
Ex. 2.12.



Ex. 2.13.



Ex. 2.14.




Examples 2.11–2.16.
Tunes produced by Alice Tegner or Sundberg and Lindblom's grammar (1976).

It is Sundberg and Lindblom's claim for the melodies generated by their grammar that 'the general reaction of Swedes listening to these melodies informally is that they are similar in style to those by Tegner.' Whilst this is hardly a rigorous proof of the adequacy of the grammar, the work demonstrates that music with a definite style can be generated from a small set of completely definite grammatical rules. The rules could, no doubt, be improved. For instance, I find it hard to believe that the repetitions of the same pitch in Example 2.12, or the intervallic leap between the second and third bars of Example 2.11, are characteristic of the idiom. However, there seems no principled reason why further restrictive rules could not eliminate such occurrences. The important feature of any such grammar is that major decisions are context-determined. They are governed by an overall structure represented in the constituent tree and the prominence contour derived from it. The psychological implication is *not* that Tegner composed her tunes by running through these precise steps in her mind. Rather we would wish to say that her compositional procedures incorporated the structures and restrictions inherent in this grammar *in some way*.

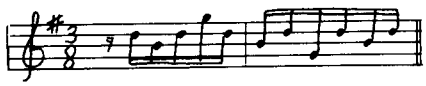
Of the examples given only 2.15 is, in fact, genuine Tegner. The structure of Sundberg and Lindblom's grammar affords no easy answer to the question of which rules characterize Tegner's music alone, and which are characteristic of wider bodies of music. Perhaps the best way to tackle this question is to consider each of the three traditional features of music—harmony, rhythm, and melody, in turn.

### 2.5.3. *Harmony and tonality*

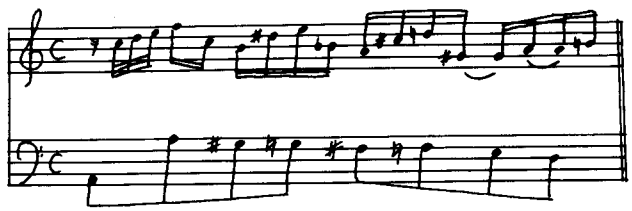
If one consults any standard textbook on harmony, one finds that it contains a large set of prescriptions for the sequences of chords that are or are not

permissible in tonal music. These are not generative rules; rather they allow one to check whether a given sequence conforms to them. The rules specify, not actual notes, but chords within a key. Thus, for instance, it might be stated that a piece of music must terminate with a tonic chord. In order to check out whether this type of rule is being obeyed by a particular sequence it is, therefore, necessary to determine what key the music is in. A key is specified by a particular scale (e.g. G major; containing the notes G, A, B, C, D, E, and F sharp) and an assignation of functional names (tonic, dominant, etc.) to the members of that scale. The notion of a key also contains the idea that the *tonic* is the fundamental or central note in relation to which all other notes are characterized.

The concept of a key raises two psychological questions. First, in virtue of what is the note C in the set (C D E F G A B) heard or understood as central? Why is it not E or F that has this centrality? Although there exist simple rules for assigning a tonic to a scale (e.g. it is the note preceded by three whole tone and one half tone scale steps when the notes of the scale are arranged in asending order) such rules do not *explain* the psychological centrality of the tonic. Secondly, how does a listener determine which note is the tonic of a musical sequence on first hearing? One possible procedure would be to examine the opening measures to enumerate all the different notes contained there, putting these notes in scalar order to discover what scale was present, and so determining the tonic by the application of a rule like the one given above. Unfortunately there is much music where we can easily hear a tonic, but where there are too few different notes in the opening measures to specify a single scale uniquely, or where the measures contain a set of notes that does not come from any single scale. Two examples from Bach's *Well-Tempered Clavier* demonstrate this point well. Example 2.17 shows the opening two bars of Fugue 15 (Book 2). Only three notes are present (excluding octaves), but most listeners will easily identify the fourth note (G) as the tonic (and not the D or B). Example 2.18 shows the opening bar of Prelude 20 (Book 2). It contains all 12 notes of the chromatic scale yet listeners will identify the key as A minor without difficulty.



Example 2.17. From J. S. Bach, *Well-Tempered Clavier*, Fugue 15, book 2.



Example 2.18. From J. S. Bach, *Well-Tempered Clavier*, Prelude 20, Book 2.

A theoretical framework in the context of which such problems may be elucidated has been provided by Longuet-Higgins (1972, 1976, 1978). In this theory every note of a composition is assigned a coordinate in a three-dimensional psychological space by a listener. In the first dimension ($x$) a move of $+1$ raises the pitch of a note by a perfect fifth (pitch ratio 3:2). In the second dimension ($y$) such a move raises the pitch by a major third (5:4). In the third dimension the octave (2:1) is the unit. For expositional purposes we may ignore the third dimension under the principle of octave equivalence and represent tonal space in a two dimensional form (see Fig. 2.11). The third dimension can be thought of as extending out of and below the page, all notes with the same name in the various octaves being 'stacked' above or below the note shown. Within this space a key is a bounded sub-space. In Fig. 2.11 the space of C major is bounded by the solid line.

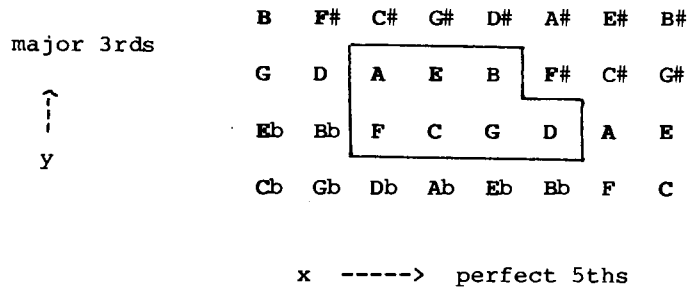


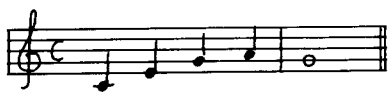

Fig. 2.11. A representational matrix for tonal music. From Longuet-Higgins (1978).

It is of note that, taking C as the origin, all other notes in the scale can be reached in a maximum of two moves (either one move of a fifth and one of a major third, or, in the case of C–D, two moves of a fifth. If we take any other note as the origin the number of moves increases (e.g. it takes four moves to get from A to D). So, a mental representation of this sort accounts in a direct way for the *centrality* of the tonic. It has the closest average distance from its fellow scale members than any other note.

Another feature of note is that the fundamental chord of tonal music, the major triad, can be obtained by taking any note together with its next highest neighbours in the $x$ and $y$ dimensions. The triad thus comes out as a compact L-shaped representation in the space. Whenever chords of this kind are encountered they are strongly suggestive of particular tonalities. The C–E–G triad is to be found within three keys only, and these keys can be visualized by sliding the key 'box' to the left or the right by one step (giving either F major or G major). Some chords are even less ambiguous than this. The dominant seventh chord (G–B–D–F) can be accommodated only within C major. If the key frame is shifted in any direction, one or more of the notes falls outside it. This reflects the fact that, in traditional tonal harmony, the dominant seventh has very strong key implications. In contrast there are

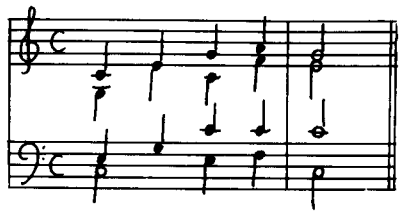
some chords which cannot be fitted into any key 'box' (e.g. C–E–G sharp, the augmented fifth; or C–E flat–F sharp–A, the diminished seventh). This is reflected in the fact that these chords do not have key implications, at least in isolation.

A further feature is that notes get multiple representation in the space. For instance, as well as the A in the C key-frame, there is another A just to the right of it. This feature reflects the musical insight that the same note can fulfil different tonal functions in different contexts. Consider the phrase in Example 2.19. Here, given a starting key of C, the A could be *either* part of an implied subdominant chord (F–A–C) in the key of C leading, via a plagal cadence, back to the tonic C, *or* part of an implied dominant chord of G major (D–F sharp–A) leading to a modulatory perfect cadence into the key of G. Longuet-Higgins' claim is that, in listening to an excerpt like this, we have to decide which tonal interpretation to give the A before we can represent the excerpt in tonal memory (or at least acknowledge the ambiguity and hold both readings until further information allows us to choose one of them). The use of equal temperament tuning on many instruments increases the possibility for ambiguity, for then C natural, D double flat, B sharp etc. are all represented by the same sound. Any note played on the piano has multiple ambiguity (in isolation). Listeners generally decide on a single tonal interpretation for a note by taking into account the context in which it appears, i.e. the ordering of the musical elements. This is equivalent to saying that the listener is sensitive to syntactical considerations.

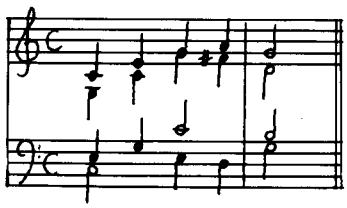


Example 2.19.

Some of these considerations may be quite simple. For instance, all other things being equal, a listener will choose a tonal interpretation for a note which places it close to preceding notes in tonal space. In other words, he operates within a restricted 'window' of tonality which changes position only slowly. Another principle could be to choose an intepretation which preserves the current key frame rather than one which shifts it. Most listeners will hear Example 2.19 as remaining in C major, rather than modulating to G. They choose the intepretation of A which preserves the initial key. If, however, we write out two alternative chordal accompaniments (Examples 2.20 and 2.21) we can disambiguate the phrase. The appearance of an F sharp in Example 2.21 'forces' the listener to shift the key frame to G major. In general, the fuller the accompaniment to a melody, the less possibility for tonal ambiguity there is. This does not mean, however, that composers are advised to 'flesh out' implied harmonies all the time. Ambiguity is, in itself, a very important compositional device. I cannot systematically describe the myriad types of



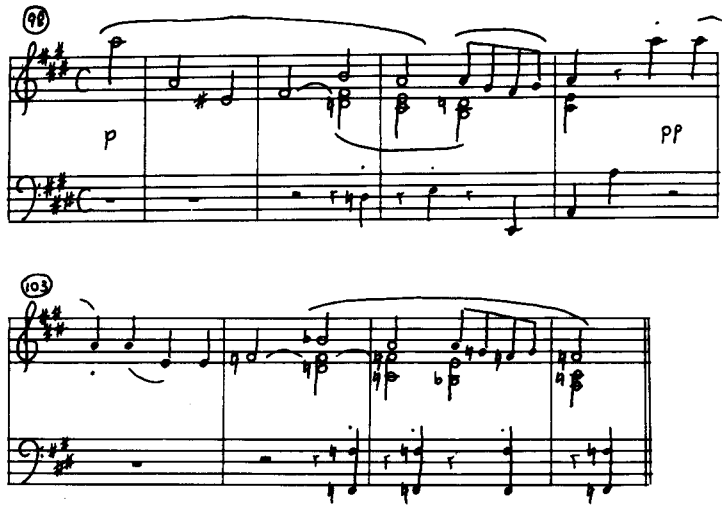
Example 2.20.



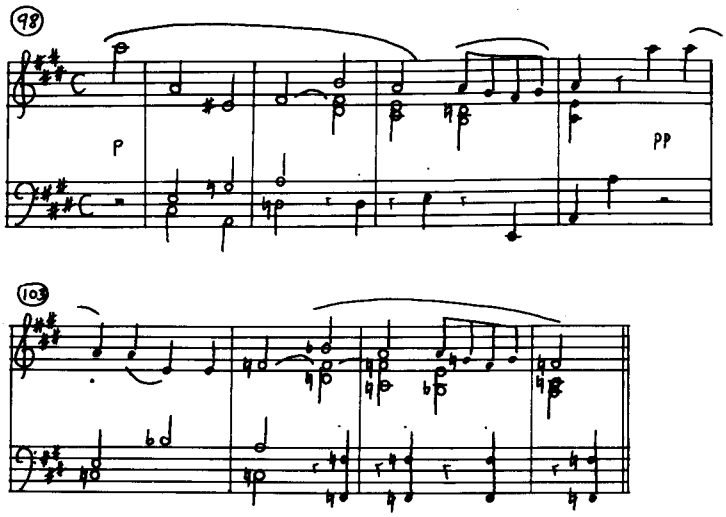
Example 2.21.

harmonic ambiguity found even in classical compositions, but here is one simple and beautiful example from Beethoven's Piano Sonata (Op. 14 No. 1). Example 2.22 gives bars 98 to 106 of the last movement. The music is firmly in A major at this point. Taken in isolation, the four-note figure which opens the excerpt (and which recurs in repeated crotchets four bars later) is harmonically ambiguous. The context will, however, suggest to many listeners a move from the chord of A to the chord of D (via an augmented fifth on A). This interpretation in confirmed by the chord in bar 100. This creates a strong expectation that the harmonic pattern will be repeated in bars 103 to 104. The first note of 104 will be heard as an E sharp awaiting resolution upwards to F sharp as before. But then comes the chord of B flat (!), taking the music firmly into the key of F major and forcing interpretation of the preceding two notes as intiating a movement into F. Both the original tendency to interpret as E sharp and the surprise which causes the reinterpretation as F are part of Beethoven's intention. The passage would lose all its dramatic power if an accompaniment such as that in Example 2.23 had been provided. This would have prepared the listener for a transition to F major, and the crucial moment of surprise would have been lost.

In this, and all other examples of musical ambiguity, there is no single syntactical structure. Rather, we have two possible structures, each of which the composer intends to be construed by the listener at different times. The rules for the introduction of ambiguity, if rules there are, lie outside the set of rules which describe legal moves within tonality. When psychologists and others talk of a musical syntax they are usually talking of rules for the construction of music which is intended to be *un*ambiguous (such as Tegner's nursery rhymes). We return, therefore, with caution to our earlier question—how to determine the key of compositions from their opening measures (Examples 2.17 and 2.18). A composer may not always wish for us to solve



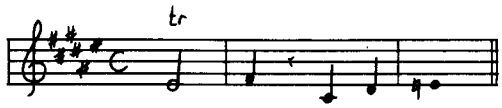

Example 2.22. From Beethoven, Piano Sonata Op. 14 No. 2, last movement.




Example 2.23. Alternative harmonization for Example 2.22.

this particular problem easily. Although it would be generally true to say that tonal melodies tend to start on a note of the tonic chord, and therefore a listener would be trying to 'fit' a tonic harmony at the soonest opportunity, there is no guarantee that melodies will not be written which start on other notes. In Example 2.17 the rapid establishment and maintenance of a G major chord strongly suggests G major, and in Example 2.18 the emphatic bass line establishes A as the tonic immediately, although the chromatic writing then does its best to dissuade the listener from his certainty. But what

of the F sharp major Fugue of Book 2 (Example 2.24)? This starts resolutely on an E sharp, which is no part of the F sharp major chord. Arguably, what allows such a strong sense of F sharp major here is the trill on the first note. This evokes the leading-note to tonic trill which characterizes the final cadence of many pieces of the period. So Bach may have been relying on listeners' ability to recognise this very distinctive figure to determine the tonality.



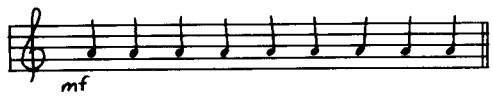

Example 2.24. From J. S. Bach, *Well-Tempered Clavier*, Fugue 13, Book 2.

How may we sum up this brief discussion of harmony in tonality? Generative grammars, such as Sundberg and Lindblom's (1976), written for a small corpus of tunes, allow further tunes in the same idiom to be produced. What they lack is reference to any principles in virtue of which the rules may be understood. The purpose of theorists such as Longuet-Higgins is to make general postulates about the way in which listeners represent large bodies of music. They can explain why it is that much music seems to operate within certain constraints, but they also help us understand the effect of music which strikes the ear as tonally ambiguous.

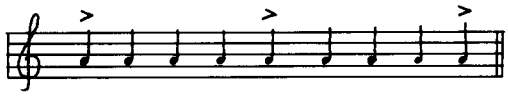
### 2.5.4. *Rhythm and metre*

The structure underlying the rhythm of a piece of music is often a metrical tree of the sort depicted in Fig. 2.10. The simplest type of structure involves a regular pulse, so marked that every *n*th occurrence (usually 2, 3, 4, or some simple multiple of these) is stressed. A stress marks the beginning of a metrical unit, which is notated in much music by the use of bar lines. Sometimes the tree has several nested levels (as in Fig. 2.10) but I will concentrate here on the rules which prescribe possible rhythmic structures *within* a single bar. As with harmony, the most fruitful discussions of rhythmic syntax have occurred in the context of the problems facing a listener trying to identify rhythmic stress from a sound pattern. How does he work out which notes are principal stresses? If a listener is presented with a set of equally spaced, equally loud notes of equal pitch (Example 2.25) there is no way in which any rhythm may be said to be present (although the listener may *impose* some rhythmic grouping of his own (Woodrow 1951). To communicate an intended rhythmic interpretation, the primary stresses must be marked in some way. One method of doing this involves variations in intensity. In the absence of other information a stressed note should be louder than its neighbours. Thus, Example 2.26 gives a strong indication of a quadruple metre, each loud note marking the start of a new bar. On instruments where dynamic change from note to note is not possible (e.g. organ) analogous ways of indicating stress

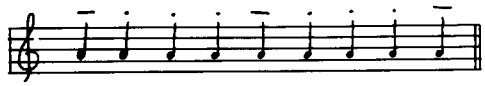
are available through articulation and phrasing. For instance, the first note may be held for the full length of its duration (tenuto) whereas the other notes are released early (staccato), as in Example 2.27.




Example 2.25.



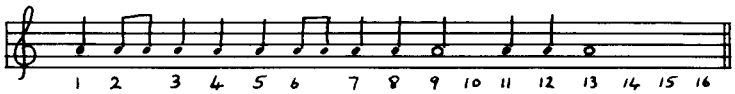
Example 2.26.



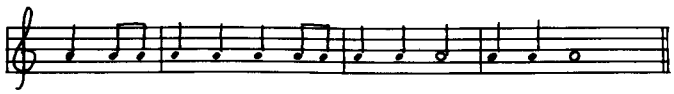
Example 2.27.

This general class of marking procedures I shall call 'performance markers' because they rely on the way the notes are played rather than on their specified pitches and durations. I shall discuss them further in Chapter 3 (particularly section 3.2.3). They are analogous to stress and intonation in speech.

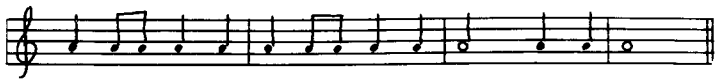
A different type of marker is illustrated in Example 2.28. Here the note durations are varied by subdividing the beat or by making a note last for more than one beat. Observation of much tonal music supports certain generalizations about the distribution of note values. For instance, the stressed beat is not usually subdivided, and does not occur within a long note other than at its beginning. Thus, in Example 2.28 we can eliminate beats 2, 6, 10, 14, 15, and 16 as primary stresses, leaving a quadruple metre stressing notes 1, 5, and 9 as the only regular interpretation. In this example, however, some possible barrings cannot be eliminated without reference to the 'empty' beats 14 to 16. This seems rather a long time for a listener to wait before the rhythm can be determined, to say nothing of the memory load imposed by holding so many notes whilst various rhythmic interpretations are tried out. Are there other features of the structure which could be used to make an earlier judgement? One possible strategy is to look for repeating patterns, on the assumption that a rhythmic pattern will fall at the same position within a bar on its two occurrences. The distance between two patterns will then be a whole bar or a simple multiple of one bar. On this principle, the repetition of the crotchet + quaver + quaver pattern in Example 2.28 limits us to a duple or quadruple metre because a triple metre would inevitably involve the pattern falling at a different position within the bar on its two consecutive occurrences. Even this strategy, however, does not help the listener decide




Example 2.28.



Example 2.29.



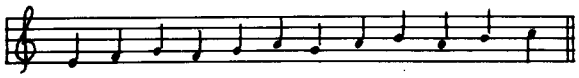
Example 2.30.

between the two quadruple barrings given in Examples 2.29 and 2.30. Here, it is only the absence of a new note at beat 14 which is decisive in favour of 2.29.

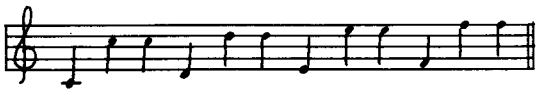
Another strategy which allows an earlier decision involves the search for dactylic groupings. A dactyl is a figure where one note is followed by two of half its duration (e.g. crotchet followed by two quavers). A recursive parsing based on this grouping can sometimes yield an appropriate barring quite early. Applied to Example 2.28 the procedure would search for the first dactyl to be encountered, duly identifying the first three notes as such. This dactyl would then be considered as a unit (effectively by replacing the three notes with a single note of their summed durations, a minim). The sequence is thene searched again from the beginning for a dactyl, and the procedure identifies the initial minim together with the following two crotchets as a superordinate dactyl. Once again, the dactyl is replaced by a single note, a semibreve. At this point, the procedure can find no further dactyl involving the first beat, and moves on to consider the fifth beat. Acting on the assumption that the beginning of a dactyl marks a stressed beat, this procedure identifies beats 1 and 5 as primary stresses.

Of course, very little music contains passages which remain at a single pitch for any length of time, and a further possibility for determining rhythm rests on the tendency of much music to contain repeating pitch patterns. Although the notes of Example 2.31 are all of the same duration the listener experiences a sense of triple metre because the pitch pattern repeats itself in transposition after every three notes. This effect can be heightened by emphasising the pitch separation of the first note in each group as in Example 2.32. For a fuller discussion of these and other strategies of rhythm recognition, see Steedman (1977) and Longuet-Higgins and Lee (1982).

The difficulty with proceeding from plausible strategies like these to a grammar of rhythm is that composers arguably use their knowledge of listeners' likely strategies to support music which is neutral with respect to,

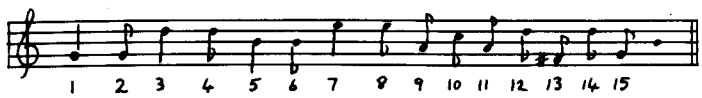


Example 2.31.



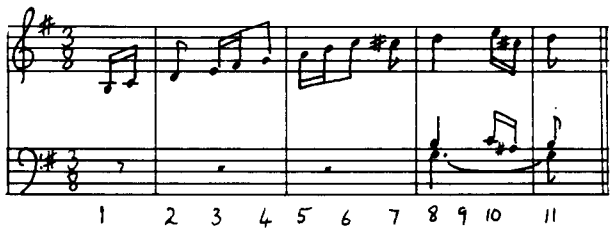
Example 2.32.

or even violates, the rules which the strategies imply. For instance, one may postulate a principle of consistency with regard to the way in which listeners ascribe metre to a piece of music. If a sequence begins with strong cues to a certain metre, the listener is likely to maintain that metre as an interpretational aid for some time, even in the absence of further strong cues. Thus, the composer may write music which would be 'without metre' if heard in isolation, but which acquires a particular interpretation in a listener's mind by extrapolation of the preceding metre. A good example of this is provided by Bach's Organ Fugue commonly known as the 'Gigue Fugue'. It gets its nickname from the unmistakeable triple metre strongly implied by the first eight notes of the theme (Example 2.33). Then follows a string of quavers. There are certainly no durational cues to metre here, and indeed, the repetition of note 9 by note 11 might be taken as a counter-suggestion of duple metre. Nonetheless one experiences a strong sense of triplet grouping 'carried over' from the opening. Elsewhere, the use of syncopation takes this idea one step further by placing an apparent stress in a place where the previous metre would not predict it. If the syncopation is to be heard as a syncopation (and not as a change of metre) then we must look to the listener's reluctance to alter the inferred metre (and the composer's exploitation of this reluctance) rather than to rules for generating sequences from a rhythmic 'tree.'

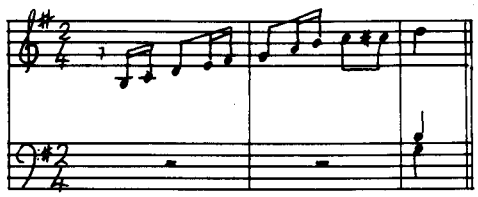



Example 2.33. Adapted from J. S. Bach, Gigue Fugue for organ.

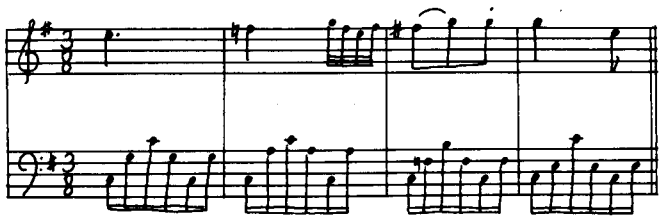
As with harmony, one must acknowledge that it is often the composer's intention to thwart or puzzle the listener with rhythmic ambiguities. Take, for instance, the opening of the last movement of Beethoven's Piano Sonata in G (Op. 14 No. 2) given in Example 2.34. Despite the barring, the effect of this sequence, up to beat 9, is to suggest the duple metre shown in Example 2.35. This grouping is suggested by the dactyl groups, particularly the superordinate dactyl formed by beats 4, 6, and 7, but the interpretation runs into trouble at beat 10 (where a subdivision of the beat is not expected) and is strongly challenged by the absence of a new note on the supposed primary




Example 2.34. From Beethoven, Piano Sonata in G, Op. 14 No. 2, Last movement.



Example 2.35.



Example 2.36. From Beethoven, Piano Sonata in G, Op. 14 No. 2, Last movement.

stress at beat 12. In this case the listener has no real way of settling for one particular interpretation of the rhythm, but it would destroy the point of the music for the pianist to thump out an accent on the first beat of each bar. The uncertainty is all part of a mischievous 'catch me if you can' first subject, full of abrupt stops and sudden flurries of activity. All manner of uncertainties are only resolved finally in the serene and rhythmically unambiguous second subject (Example 2.36).

These considerations lead us to some fundamental restrictions on the way we postulate a grammar of music to be *used*. In language we can usually assert that both speaker and listener *abide* by the grammar, and that the speaker's intention is normally to provide unambiguous utterances which the listener will interpret as the speaker intends. In music the composer abides by grammar in a much looser sense. He may certainly know of its existence, and how to write unambiguous music using it, but he also stands outside the grammar in a way that the listener may not. He anticipates the strategies that a listener in his culture will use to structure his or her experience, and seeks to thwart these strategies in interesting ways. Thus, any shared grammar does not *generate* the compositions. Rather, its existence is a major

consideration in determining the nature and degree of freedom the composer can exercise to transform and extend musical style. Put simply, if a composition is totally generated by a grammar it is likely to be dull; if it breaks rules in an unmotivated way it is likely to be unintelligible.

The best linguistic analogy for musical composition is not so much the utterance of a novel sentence as the utterance of a non-sentence (as in a poem) which, by virtue of the utterer's appreciation of his listener's linguistic apparatus, has significance for the listener.

### 2.5.5. *Melody*

Many of the constraints on melody are dictated by the requirement that the melody be capable of communicating harmonic and rhythmic structure to a listener. This is the reason why Sundberg and Lindblom's grammar generates the melody *after* the harmony and rhythm have been determined. It would be oversimplistic, however, to suppose that composers conceived harmony and rhythm prior to melody in all cases. Although a composer may sometimes be constrained by harmony and rhythm (e.g. 'I need to get from D major to A major in four bars') there are times when he may well allow harmony and rhythm to be derived from a melodic idea. The claim that rhythm and harmony are syntactically prior to melody does not rest on observations about the composing process (of which more in Chapter 4). It rests rather on observations about the musical *product* and its effect on listeners. Firstly, the number of different melodies within an idiom is much greater than the number of harmonic and rhythmic sequences common within the idiom. There are many cases where very different melodies are supported by identical harmonic progressions, but rather few cases where very different harmonies support identical melodies. Secondly, it makes some sense to talk of harmony without melody, or rhythm without melody, but very little sense, at least within the tonal idiom, of talking about melody without harmony or rhythm. When listeners hear a melody, their processing of it normally involves the attempt to retrieve implicit harmonic and rhythmic structure. Representing the melody to themselves *means* recovering this structure. In contrast, a chord sequence can be heard without any attempt to recover an 'implicit melody'.

Melody, thus, represents the level of the greatest differentiation in music, the level at which our evaluative and critical faculties are most immediately engaged. It is the aspect of music which is nearest to the 'surface', and that which, for most listeners, most immediately characterizes the music. Our appreciation of a 'great' melody rests partly on our appreciation that a composer has a great many options when generating a melody even if harmony and rhythm are determined. His choice becomes memorable when it is sufficiently different from anything we might have thought of, or heard before, and when we appreciate some particular rightness or fittingness about the choice. A syntax for melody should not aspire to formalize greatness.

What it can do, perhaps, is specify the permissible options in a way which makes clear just how formidable the composer's task is to select from among them.

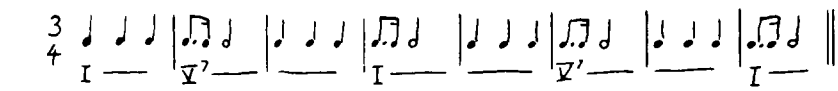
Let us, then, take a very simple harmonic and rhythmic framework as an example, and examine the kind of rules which would allow us to derive acceptable melodies from it. Example 2.37 gives the rhythm and harmony that our melody must 'carry'. We may imagine this to have been generated by the application of a context-sensitive set of rules such as those of Sundberg and Lindblom discussed in section 5.2 of this chapter. A first rule of melodic construction which we might wish to consider is:

1. A note may be chosen from the chord of the implied harmony at that point.

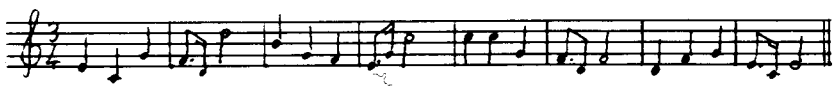
If we suppose that we are in the key of C major, then this rule allows any note in the first and fourth bars to be C, E, or G, and any note in the second and third bars to be G, B, D, or F (dominant seventh chord). Even if the choice of notes is confined to a single octave this rule allows 2 985 984 distinct melodies. Examples 2.38 and 2.39 are just two of such melodies. Examination of published melodies rapidly demonstrates, however, that Rule 1 is too restrictive for many melodies, because they contain notes outside the harmony at a given point. Consider, for example, Verdi's famous 'La donna è mobile' from his opera *Rigoletto* in which the precise pattern of Example 2.37 is repeated twice (see Example 2.40). There are five notes (asterisked) which Rule 1 does not allow. To accommodate these notes we need at least two more rules:

2. A note may be inserted between two adjacent notes of a given chord if it forms a scale step with one or both notes (passing note).
3. Two occurrences of the same note may be replaced by a pair comprising that note preceded by one a single scale step higher (*appogiatura*).

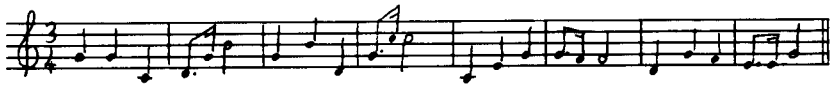
The derivation of the melody from these rules can be illustrated in stages. Example 2.41 shows a melody derived by Rule 1 alone. Example 2.42 shows the modifications introduced through Rule 2, and Example 2.43 the form achieved after selective application of Rule 3. Notice that there is still one note (asterisked) which does not conform to the Verdi melody. Is there yet another rule which we should introduce to account for this? Maybe there is; but we should also allow for the possibility that Verdi deliberately breaks formal rules for some particular reason. In this case such a reason is not hard to find. The melody breaks down quite naturally into two-bar segments (each with identical rhythm). The second segment (bars 3–4) is an exact repetition of the first segment transposed down a scale step. Clearly Verdi has it in mind that the fourth segment (bars 7–8) should be a similarly transposed version of bars 5–6. If. however, the transposition were to be *exact*, we would get Example 2.44 which leaves the three final notes



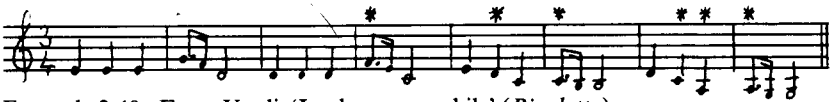

Example 2.37.



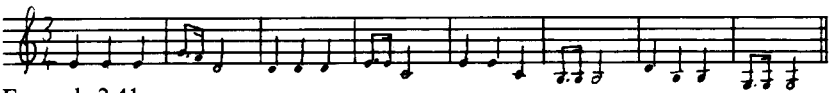
Example 2.38.



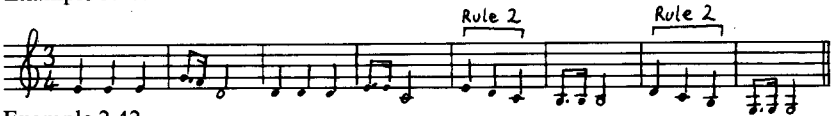
Example 2.39.



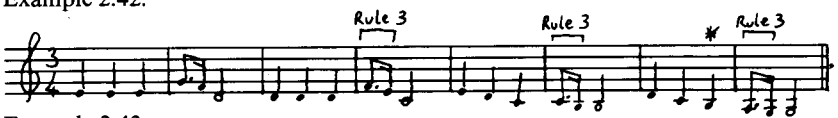
Example 2.40. From Verdi, 'La donna e mobile' (*Rigoletto*).



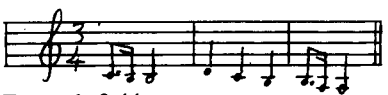
Example 2.41.



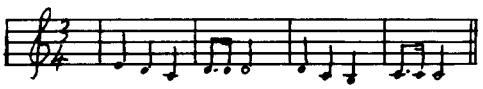

Example 2.42.




Example 2.43.



Example 2.44.



Example 2.45.

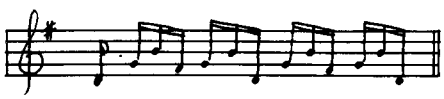
unaccounted for under the three rules we have so far. Verdi's 'compromise' is to slip the last four notes down by a tone. This preserves the contour (i.e. the 'up', 'down', or 'same' relationship of adjacent notes) of the previous segment, and, additionally, imitates in bar 7 the melodic shape of bars 2 and 4. This mixture of imitation and slight deviation from exact repetition, together with the mild discord of the 'discrepant' last note of bar 7, gives the melody

interest. A less imaginative composer might have provided a stricter, more 'correct' solution (e.g. Example 2.45).

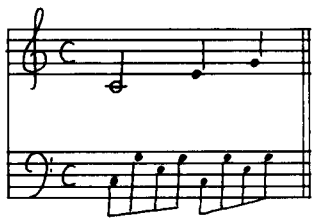
If there are simple rules of melodic construction of the sort outlined above, then I believe we must conclude two things about them. Firstly, we will find composers breaking these rules from time to time when they consider some other organizational principle to take precedence. Secondly, these rules allow such a vast number of options that the art of composition consists in developing motivated ways of restricting the options. Because the 'search space' is so large, and because there are so many relevant musical properties which could influence compositional choice, composition seems normally to involve the establishment of provisional goals and the application of heuristic procedures for attempting to achieve them. Such procedures do not formally resemble the steps of a generative grammar, and we shall examine their nature in Chapter 4.

### 2.5.6. *Pattern and structure*

Although rules of harmony, rhythm, and melody provide much in the way of detailed moment-to-moment restrictions on musical sequences, they do not account for many larger-scale aspects of musical structure which arguably fall within the realm of musical syntax. The most general of such aspects is probably *repetition*. Exact repetition can occur on many levels. The same melodic fragment can be repeated (Bach's French Suite No. 5 in G, last movement, Example 2.46), as can an accompanying figure (Alberti bass as in Mozart's Piano Sonata in C, K. 545, 1st movement, Example 2.47). On a



Example 2.46. From J. S. Bach, French Suite No. 5, last movement.



Example 2.47. From Mozart, Piano Sonata in C (K. 545), first movement.

larger scale, the same phrase may be repeated (e.g. the typical A–A–B–A song form) or a whole large-scale portion (e.g. the exposition of a classical sonata form movement). Equally often, however, the repetition is not exact. Pitch transposition of melodic material is common. Other forms of repetition with change include inversion, augmentation, contraction, segmentation, and

retrogression. Sonata form and many later compositional forms make multiple use of such transformations to extend and develop thematic material. Simon and Sumner (1968) have shown that many important characteristics of musical patterning can be captured by simple formalisms involving the operators SAME and NEXT applied to the major dimensions of musical sound. They take, as their example, the Gavotte from Bach's French Suite No. 5, and show how the first four bars can be almost completely described in terms of repetitions of elements contained within the first bar.

Whilst accepting that such patterning is integral to music, I believe that it is necessary to claim that it cannot, on its own, give a full account of the structure of a musical composition. I can, perhaps, illustrate this point by two analogies. The first is an architectural one. A curved masonry arch, such as might be found in an old church, contains several examples of patterning. For instance, the curvature of each brick may be slightly greater than that of its neighbour, whilst retaining enough similarity of shape to appear a fitting next member of the sequence. Repetition and transformation are there, but when we ask why *that* transformation is chosen, and how much transformation is required, we cannot answer the question by recourse to a formalism which describes each transformed element in terms of its predecessors. Rather, we must see what the overall function of the *set* of transformations is: to provide a self-supporting bridge of a certain shape. There is, of course, the possibility that some aspects of pattern (say, decorative carving on the masonry) are irrelevant to this level of structure, but typically, such decoration will, in some way, be *governed* by the structure (e.g. the decoration maintains a constant relation to the line of the arch).

A second analogy comes from graphic art,as exemplified by the work of M. C. Escher. In many of his drawings, a basic technique of patterning is easily discernible. Escher uses ordered series of shapes, each one of which is slightly different from its neighbour, so that as one's eye moves across the picture, subtle and disturbing metamorphoses are experienced. And yet, what makes these pictures compelling is not so much the transformational technique as the end-points which Escher uses these transformations to travel between. A drawing describes a journey with a beginning and an end, and structurally the outer destinations *control* the inner transformations. We are engaged at two levels; the technique fascinates, but a major cause of the fascination is the way the technique is used to arrive at the various significant forms.

And so it is, I believe, with music. Music uses patterning to achieve structural goals; it also uses patterning to engage the listener with decorative detail, but the structural coherence of a piece of music is not to be found solely *within* the principles of patterning. These principles do not have starting points and goals built into them. They can be applied over and over again to any musical idea whatever, so that an indefinitely long piece of music could be produced (just as, using Escher's techniques, an indefinitely large drawing

could be produced). Schenkerian concepts offer one possible basis for goal-directed structure, with the concept of a harmonic journey from the tonic which returns back there in the end, but clearly, any fully adequate characterization of goal structure is going to be complex and multidimensional. The fact that it is relatively easy to understand and characterize the small-scale techniques of repetition and transformation which operate within short musical segments, whilst it is extremely difficult to characterize the factors which give a work an integrated and directed form, probably explains Schenker's belief that it takes a genius to produce an extended piece of music with a unified and motivated structure.

A common implication of several aspects of musical syntax we have dealt with is that they are best understood as relating to propensities of *listeners* to infer particular underlying structures from note sequences rather than to the generative processes of *composers*. Recently, Lerdahl and Jackendoff (1983; also Jackendoff and Lerdahl 1982) have extensively followed through the implications of this observation by offering a set of grammatical rules which attempt to account systematically for the way in which listeners assign structural prominence to the various elements in a musical composition. They have, in effect, proposed a formal system for *analysing* a given musical sequence. Their rules can 'cope' with ambiguity because they incorporate a set of independent strategies for grouping musical elements. The final interpretation assigned by a listener will depend on the weight he or she gives to the various strategies, and on whether their effects are to reinforce or contradict one another. Their theory is too recently published to evaluate in detail here, but it promises to be a major contribution to this topic.

### 2.6. Musical semantics

It is a significant fact about language behaviour that people almost never remember what they read or hear in *verbatim* fashion. Not only does recall paraphrase the original, but it also contains many inferences about information not present in the original. Take, for instance, the following short story:

> Jill was invited to Jack's birthday party.
> She wondered if he would like a kite.
> She went to her room and shook her piggy bank.
> It made no sound.

A typical recall might produce something like the following:

> Jill was going to Jack's birthday party.
> She wanted to buy him a kite.
> She went up to her room and shook her piggy bank.
> There was no money in it.

Such a recall, although we might wish to call it accurate, is heavily laden with inference. For instance, it is possible that the reason why Jill's piggy bank made no sound was that it was full of bank notes. It is also possible that Jill's wondering about Jack's wants was a passing thought, and that she had no intention of buying him anything. These alternatives do not normally occur to us, because we bring knowledge to bear on what we hear; knowledge about birthday parties, the financial circumstances of little girls, and much else. The thrust of recent evidence (see, for instance, Sandford and Garrod 1980) is to suggest that people make these inferences during reception of the message, and that their aim in doing so is to build up an internal model of the portion of the world referred to in the message. Although it is possible that given information can be marked as such and held separate from inferred information, the evidence suggests that it is remarkably difficult to do this. When it is possible to check the verbatim record of some real-life conversation, it is found that even people claiming to have excellent memories for what other people say, and testifying under oath, deviate grossly but systematically from what was actually said (see Neisser's analysis of John Dean's 'Watergate' testimony, 1981).

The implication of this type of evidence is that people recode linguistic information into semantic information. The 'meaning' of a sentence for a listener is, roughly, what he adds to his model of the situation being described as a result of hearing the sentence. This addition allows him to do a number of things. One of these *may* be to reproduce what he received, but this is by no means a necessary, or even important, consequence (we can imitate speech that has *no* meaning for us). Other things may include taking some appropriate action (as in response to 'you have left your car headlights on'), representing the information presented in some other mode (as in drawing a picture of 'the red brick is beneath the green brick'), or experiencing some significant emotion (as to 'your grandmother died today'). From a psychological point of view language is, in some sense, a substitute for sensory experience. We respond to being told that our headlights are on in the same kind of way as we would if we saw them to be on. Contemporary theorists account for this equivalence by saying that the two inputs have equivalent effects on our mental representation of the world.

The principal issue in musical semantics is whether we can establish analogous equivalences between music and some essentially non-musical phenomenon. Does music cause psychological effects which can be caused by other means, and does it do this in a systematic way?

The first position we might consider is that music simply has no semantics. On this view, music is psychologically self-contained, a separate species of psychological activity for which unique modes of representation have been developed. The suggestion, discussed earlier, that people represent music in a tonal space such that notes are assigned harmonic functions within keys, is

certainly different from any suggestion psychologists have made about cognitive representation in other spheres.

Clearly, there is much in musical behaviour which can be accounted for by considering musical representation to be a closed sub-system with no essential links to other cognitive domains. This system merits study, is the principal topic of discussion in this book, and is the subject of the most prominent research initiatives in the area. However, I believe that the available evidence forces us to accept that there is some 'leakage'. Musical experience *is* translated into other representational modes. Consider, for example, the case of someone who has just listened to a performance of a long and complex symphonic work. It is quite possible that he or she cannot recall a single theme from the work (I have often been in this situation myself), yet he or she certainly remembers *something* about the work, and can make some appropriate response to it. When this response is expressed in words it characteristically contains remarks about the substance of the music which are neither descriptive ('it was loud') nor reactive ('I liked it') but embody an attempt to *characterize* the music through metaphor ('It had the feeling of a long heroic struggle triumphantly resolved'). It seems less significant that people often disagree about their characterizations than that they nearly always have *some* comment to offer. This is not an arbitrary reaction, but a genuine attempt to describe some real thing or experience. For a great number of us, then, music has extra-musical meaning, however intangible.

What is it that makes music have meaning for us? One possibility is that it *mimics* the sounds which occur in extra-musical contexts. There is a great deal of music which employs mimicry to considerable effect (woodwind 'birdsong' to suggest a pastoral scene, glissandi violins to suggest the howling wind of a storm, and so on). Our recognition of these meanings requires only the knowledge of the appropriate extra-musical sounds. Other music employs *symbolic* reference to an extra-musical event. In opera, for instance, if a particular theme is associated with each appearance of the hero, then it might be used to signify that, in his absence, the heroine is thinking of him. Sometimes music can be *programmatic*, aiming to illustrate a given story. In Dukas' *Sorcerer's Apprentice* there is a bizarre moment when the music suddenly stops; then out of the silence bass instruments take up the first few notes of the main theme one after another, haltingly at first, but gradually increasing in speed and intensity. This corresponds to the moment in the story when the desperate apprentice, unable to undo his spell, chops the relentless water-carrying broom into small fragments. For a moment he thinks he has solved his problem, but then, to his horror, each fragment rises, one after another, to continue its task, compounding his folly.

In all the above cases the composer has set out to depict a definite set of world events through music, and it is easy to understand how it is done.

Unfortunately, very little music has such explicit extra-musical reference, and even in that which does, its reference does not exhaust its significance. If the sole aim of such music were reference, one might as well have a sound recording of the event referred to. Musicial reference is special because the music 'makes sense' even if the reference is *not* appreciated by a listener. A referential passage nearly always serves a double function: it points to an external event, but it also forms part of the thematic structure of the piece as a whole. It is a well-formed segment of the music quite independently of its reference, and one of art of composition consists precisely in finding passages which are capable of serving this double function.

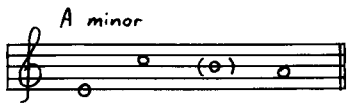
What, then, of the vast bulk of music, which does not have meaning in the specific sense outlined above? One of the most popular suggestions is that musical sequences somehow denote, or stand for, certain emotional states. Deryck Cooke (1959) has presented one of the most fully worked out versions of this thesis. He argues that the intervals of the diatonic scale suggest different emotional qualities (e.g. major third—pleasure, fifths and octaves—neutral) and goes on to isolate several basic melodic combinations of these intervals which seem to recur throughout tonal music. For instance, there are many melodies which start on the dominant of a minor scale, then rise over the tonic to the third of the scale, before falling, possibly via the super-

TABLE 2.1

| Composer | Approximate date | Words (English translation) |
|---|---|---|
| Dowland | 1612 | When the poor cripple by the pool did lie |
| Bach | 1722 | Flow my tears in floods of weeping |
| Bach | 1729 | Have mercy on me, O Lord |
| Bach | 1731 | I am weary of this world |
| Mozart | 1791 | Day of bitter lamentation |
| Verdi | 1873 | Day of bitter lamentation |
| Schubert | 1823 | Life flows away like the waters of a river |
| Schubert | 1827 | (The jilted lover bids farewell to his faithless sweetheart) |
| Verdi | 1852 | Farewell to the past, to the fair smiling dreams |
| Wagner | 1853 | (He who wishes for the power of the Ring must renounce love) |
| Mahler | 1883 | The unhappiest day of my life |
| Mahler | 1884 | Goodbye to all I ever loved |
| Britten | 1944 | We planned that their lives should have a good start |

tonic, to rest on the tonic (Example 2.48). Of this motif Cooke says: it 'conveys the feeling of a passionate outburst of painful emotion, which does not protest further, but falls back into acceptance—a flow and ebb grief. Being neither complete protest nor complete acceptance, it has an effect of restless sorrow'. In support of this, and his other analysis, Cooke collects together an impressive array of supportive examples from all periods of tonal music. The most telling of these are, perhaps, taken from vocal music, where we have some independent evidence from the words concerning the emotional mood. For Example 2.48 Cooke gives fourteen examples, all from vocal music, summarized in Table 2.1.

The ubiquity and stability of such motifs is impressive, even if it is all too easy to accuse Cooke of picking his examples to support his case. Counter-examples are not immediately damaging to the thesis, because melody is only one of the dimensions along which composers may signify emotion. Rhythm, tempo, orchestration, and dynamics are all capable of modifying the meaning. It is quite clear that Example 2.48 would have different emotional connotation when entoned fortissimo by the entire brass section of a symphony orchestra than when played liltingly by a solo violin. It is not coincidental that most of Cooke's examples of this theme are quiet and slow. I believe that Cooke has identified a real and important component of musical meaning with his melodic analyses. As psychologists, however, we need to ask what psychological mechanisms allows these meanings to be comprehended by a listener.




Example 2.48. One of Cooke's (1959) thematic motifs.

One possibility is that the motifs are intrinsically neutral with respect to the emotions and that people acquire their meanings through association with exactly the kind of words that seem to habitually accompany the motifs. Whilst this does not explain why composers should have begun to pair emotions with melodic motifs in the first place, it does seem the simplest explanation for the listener brought up within the existing musical culture. On this argument a listener would have no added difficulty in learning some different pairing of melody and emotion. For instance, it might have been that the emotion paired with Example 2.48 in our culture were paired with, say, a rising major triad in some alternative tonal culture. This is by analogy with the arbitrariness of word-meaning pairings in natural language. We believe that a language like English, where the meanings of 'cat' and 'dog' had been exchanged, would be just as easy to learn and use as the language we actually have. There is nothing about the sound 'cat' as such which makes it particularly suitable to stand for the small feline domestic mammal as opposed to anything else.

Cooke, however, invites us to reject the notion of instrinsic neutrality for music. He argues that certain motifs are more suitable than others for suggesting particular emotions by virtue of the tonal relationships that hold between their members. For instance, positive emotions are suggested by major sequences, especially those involving triads, because such sequences are directly derived from the early members of the harmonic series. Similarly, movements away from the tonic, particularly upwards, are suitable for expressing outgoing emotions, whilst movements towards the tonic signify rest or repose. This is because the tonic occupies a central or 'home' position in tonal harmony. What is suggested, therefore, is that the tonal system, as such, offers analogies for the way in which people represent emotions in some 'semantic space', and that these analogies allow a partial mapping of tonal relations onto emotions (Makeig 1982).

The psychological implications of this view are twofold. Firstly, the discovery of analogical relationships between music and emotion should be discoverable, in principle, by anyone who is capable of appreciating the tonal relationships in music and who has a normal understanding of the emotions. Thus, an explicit word–music association is not necessary to come to appreciate emotional meanings in music (although it may, of course, assist the process of acquisition). This helps to explain how so many composers, working in different times and places, seem to have hit upon the same types of melodic forms for the explicit expression of textual emotion. Secondly, it follows that pairings of melodies with emotions *other* than the ones we now have would not be as easy to learn, since they would be working against the 'ready made' analogical relationships inherent in tonality.

Although Cooke's thesis is empirical, it is not easy to know how it could be practically demonstrated. The 'ideal' demonstration would be given by someone who had never been exposed to any word–music associations (through song, etc.) showing culturally 'normal' appreciation of the emotional connotations of music. Such a person would, of course, have to have received otherwise normal musical and linguistic experience. It is unlikely that the conditions for this demonstration could ever be set up.

The research of Gabriel (1978) illustrates the considerable difficulties besetting the psychologist wishing to conduct empirical investigations into the mechanisms whereby emotional meanings in music are transmitted. He played each of Cooke's basic motifs to musically untrained subjects twice. On one occasion the motif was paired with Cooke's emotional characterization of it, on the other by a randomly chosen description from Cooke's set. Subjects were asked to rate the appropriateness of each motif–description pair. Gabriel argued that if subjects were aware of the supposed emotional connotation of each motif, then they should rate Cooke's designated description as more appropriate than the randomly chosen one. The results showed that subjects did not, in fact, distinguish between descriptions in this way. The 'random' description was equally often chosen as more appropriate. Here,

then, is an apparent failure to demonstrate any relationship between tonal pattern and emotional connotation. Gabriel's tentative conclusion is that more primitive features of melodies (such as contour) are the principal carriers of musical meaning. Before accepting this conclusion, however, it is worth questioning whether Gabriel's experiment was maximally *conducive* to the discovery of Cooke's melody–emotion links. There are three issues which seem to require consideration.

First comes the question of subjects. It is natural to suppose that, as one becomes increasingly sophisticated musically, one becomes attuned to finer emotional nuance. A hierarchy of emotional cues seems likely, with primitive cues (such as speed and loudness) available at all levels of musical sophistication, and with more subtle cues (such as aspects of tonal relations) available only to those with deeper analytic powers with respect to the music. A common experience among many music lovers (including myself) is a belated appreciation of the emotional diversity and subtlety in the music of a composer such as Mozart. The inexperienced listener may find Mozart pale, insipid, and all vaguely 'jolly', especially when set beside the kalaeidosocopic turbulence of the romantic composers. Closer knowledge of Mozart (and maybe of the emotional world) results in his music becoming richly and sublimely expressive. The ability to 'read' the emotional language of music is an acquired skill, and we should, perhaps, not be too surprised that a group of 'ordinary' subjects do not show much awareness of the finer details of this language.

Second is the consideration of the nature of the experimental stimuli. In Gabriel's study the motifs were presented in musically 'disembodied' form. Each stimulus was a series of equally loud and equally long sine waves following the minimal tonal context of a tonic chord. The justification for this is the wish to hold all factors other than melody constant. There is no real music where melody is the *only* dimension on which two excerpts differ, and so it is impossible to tell which of several differences between two pieces of music is responsible for a difference in emotional connotation. In Gabriel's experiment, any difference in emotional meaning *must* be due to a difference in melodic shape, since nothing else varies. The major criticism of Gabriel's design is that it may have destroyed the conditions under which it is possible for a listener to activate the cognitive mechanisms which mediate between music and emotion. For instance, one may argue that the ascription of emotional meaning depends upon exposure to an extended context in which the crucial melody is embedded. By analogy, in recognition of facial expression we use information from the area of the eyes, but if we see *only* the eyes without the rest of the face, our ability to use this information is diminished. We may need to know where an excerpt comes in the structure of the music before we can interpret it emotionally; or we may need to give it some metrical reading. It would be difficult even for Cooke to hold that *every* occurence of a melodic motif has emotional significance. More likely, it is

the portion of a sequence that is highlighted through its position in the musical structure, or its scoring, that carries the brunt of the emotional burden. For these, and other reasons, Gabriel's task is an unreal one. The future task is to find more realistic listening situations which nevertheless allow for the degree of control over the various dimensions of sound which is an undoubted prerequisite of the experimental investigation of response to music.

Third we must consider the nature of the response that subjects are asked to make. The judgement of the emotional meaning of a piece of music is not necessarily an instantaneous thing. It may be only gradually that one comes to the best characterization of the extra-musical themes that seem to be given metaphorical expression through the music. The analogy from language, where we hear a word or phrase and then can instantly access some central aspect of its meaning, is not necessarily the best one for understanding what goes on in music listening. A better analogy would perhaps be the ill-defined process by which we come to appreciate that a particular turn of phrase in a poem has implications which contribute to the meaning of the poem in a way that some putative 'synonymous expression' would not do. To come to this appreciation may require extended mental work during which we give the object of attention detailed and repeated scrutiny. Gabriel's subjects were required to come to a rapid unmeditated judgement about the meaning of short musical excerpts. Although music can have immediate emotional impact, there is nothing in Cooke's thesis which demands that it should be the *tonal* element of music which alone carries this impact.

I have chosen to discuss Cooke's ideas at some length because they seem particularly clear and simple. Once we accept the general idea of analogical relationships between music and extra-musical domains, however, the possibility for other ways of expressing meaning are immense. I have already outlined one possible 'meaning' of Schenkerian structures in section 3 of this chapter. Many other possibilties suggest themselves. For instance, one may wish to find in symphonic form some metaphor for human life. The transformation or metamorphosis of theme parallels the changes we experience in ourselves and our ideas. The unification of diverse themes within a coherent structure parallels our attempts to endow our chaotic lives with meaning. And so on. This is, surely, part of the power of music, and this power is possibly enhanced by the manifest lack of practical meaning of the type supported by language. Musical semantics is of a similar type to poetic semantics. This does not mean that either subject is necessarily opaque to scientific understanding, but that we may be mistaken to seek for musical meanings in the same way as psychologists have so far attempted to elucidate the semantics of normal speech.

I would like to end this section by commenting briefly on the work of the noted American musicologist, L. B. Meyer (1956, 1973). He has made a distinction between two forms of meaning in music, *embodied* meaning, and

*designative* meaning. Designative meaning is that which refers outside the music to objects or events in the non-musical domain. We have been primarily concerned with designative meaning in this section. Embodied meaning, in contrast, is the significance that a musical passage can have for a listener in terms of its own structure and the interaction of that unfolding structure with the listener's musical knowledge and expectations. The musical structure can create certain expectations (or implications) which it can then either fulfil or thwart. Among those which are fulfilled, some are fulfilled at once, others after a delay. This creates a dynamic flux of tensions and resolutions which can influence the emotional and aesthetic responsese of a listener.

Like Cooke's theory, Meyer's requires that a listener represents the music he hears in terms of the tonal and other formal categories of the musical culture. It is, however, a broader theory than Cooke's, in that it gives an explicit semantic role to large-scale musical structures, whereas Cooke concentrates on small melodic patterns. Meyer's approach also shows how the same melodic pattern could have quite different meanings in different contexts according to whether or not it fulfilled implications of earlier material. Because the theory is broader it is also more difficult to bring into the laboratory, and I know of no empirical research which bears directly on Meyer's central ideas. This is a shame because he is clearly one of the most influential and psychologically well-informed of contemporary musicologists. I shall give some examples of Meyer's notions of melodic implication in Chapter 5, but would recommend Meyer's (1973) book to anyone who is seriously interested in musical semantics.

## 2.7. Conclusions

I hope that the diverse topics reviewed in this chapter have shown that it is fruitful to consider music in relation to language. It allows us to notice fundamental similarities (e.g. the Chomsky–Schenker parallel) but also highlights differences in an illuminating way. For instance, the fact that categorical perception in music is not so complete and automatic as it is in language raises productive questions for future consideration. What is it about the two systems which makes for the difference? The fact that a generative grammar for music seems to be a less realizable goal than for language forces us to consider the fundamental differences in aim of linguistic and musical communication. In the process, our perception of both domains is sharpened.

The linguistic analogy is neither 'true' nor 'false'. Like all analogies, it achieves a partial fit with its subject. The 'true' element which I would like to emphasize most strongly is the notion that we represent sequences of individual elements by assigning theme roles in abstract underlying structures, some of which, particularly those with hierarchical organization, have strong family resemblance to one another. It is the relationship of elements

to one another within these structures, rather than their temporal or spatial proximity that determines whether or not they are psychologically close. In the following chapters we shall not explicitly pursue the analogy further, but we shall find that the aspects of musical cognition emphasized by the analogy have profound and detailed implications for every musical activity we examine.

# 3. The performance of music

## 3.1. Introduction

In its broadest sense, performance covers the whole range of overt musical behaviour. A little child's improvised play-song, the humming of a popular melody, participation in corporate ritual such as hymn or folk singing, and dancing to music, are just some of the diverse modes of performance which merit psychological study. In a narrower sense, though, a musical performance is one in which a performer, or a group of performers, self-consciously enacts music for an audience. In our Western culture, such music is often written by someone not directly involved in the performance. The performers *realize* a pre-existent composition.

In this chapter I shall concentrate on performance in this latter narrow sense. There are several reasons why I do this. Firstly, most of the contemporary research is about this type of performance. Secondly, I shall wish to treat instances of performance where the performer is also the originator of the music (improvisation) as a special case of composition, and so reserve discussion of it for Chapter 4. Music of non-literate cultures is discussed in Chapter 7, and some aspects of children's musical behaviour are examined in Chapter 6. Thirdly, in performance of a 'composition', there is an objective record of the composer's intentions, usually in the form of a score. Without such a record we have no easy way of characterizing a piece of music independently of a particular performance of it, and so questions such as how *well* a performer realizes the music become very hard to answer.

There are three main stages of involvement with a musical score at which we can study performance. First is the unpremeditated performance which a musician is capable of providing at first sight of a score. This is commonly known as sight reading. Second are the performances generated over a period of repeated exposure to a score. Usually the aim of such performances is to improve performance until it reaches some criterion of adequacy, because sight reading does not often produce a totally satisfactory performance. This type of performance is known as rehearsal or practice. Finally comes the more-or-less finished product of rehearsal, a polished performance which may well involve total memorization of a score. This we will call expert performance, although it must be pointed out that some musicians can perform much music well without extensive rehearsal. Good sight reading (and indeed, good rehearsal) are forms of expertise in their own right.

### 3.2. Sight reading

Almost all of our reading of language is sight reading. Unless we are trying to memorize a text or are scrutinizing it for some professional purpose, we tend to read a text just once, and we expect that to suffice for our purposes. In contrast, many musicians find fluent sight reading very difficult, and brave indeed is the performer who would step on to the concert platform with no preview of his material.

Part of the explanation for this difference may be that whilst the fluent sight reading of language is necessary for a full life in our culture, the same is not true for music. Enormous educational effort is expended on getting children to a state in which their language reading is reasonably fluent. On the other hand, it is not necessary for children to be able to sight read well for them to succeed at conventional musical tasks, most of which allow for, even assume, intensive rehearsal. Thus, music sight reading is taught neither so rigorously nor at such an early age as language reading.

There are other factors too. Children are already fluent language users before they learn to read. They are rarely such fluent music performers. Most children learn a new music performance skill, such as playing an instrument, *alongside* learning to read music. This double task can be an intolerable burden which many solve by memorizing each new piece at the soonest opportunity so that their performance does not depend upon their reading ability. The result of this is that some musicians give themselves very little practice at continuous reading. Some modern educational techniques attempt to circumvent this problem by giving opportunity to acquire some performing skill before introducing music notation (e.g. the Suzuki method of violin teaching).

Finally, and perhaps most importantly, there is a sense in which the *criteria* for successful language reading are more lax than those for music reading. For most purposes it is sufficient, in language reading, to determine the practical meaning of the text. Much reading is silent, and even when spoken, often little more is required than the ability to utter the correct words in the correct sequential order. In contrast, music reading requires the execution of a complex response where there is very little latitude for deviations in timing and quality. The only linguistic task whose performance demands would approach the musical task is the recitation, at sight, of some piece of fine literature, incorporating all the nuance of voice, expression, and timing which one would expect from a first class actor.

The question which many musicians ask is 'what makes a good sight reader?' This question includes two separate concerns—firstly, what can be said about the characteristics of the fluent sight reader—and secondly, what must a poor reader do to become fluent? It is an unfortunate but unavoidable fact that a reasonably comprehensive answer to the first question does not necessarily lead to prescriptions for the second. For instance, the finding

that fluent readers typically look further ahead than poor readers does not automatically yield the prescription that poor readers should practise looking further ahead. It may well be that increased ability for preview is the result of some other skill, such as the ability to detect pattern or structure in the score, and that simply trying to look ahead will not improve *this* skill. Therefore, whilst it is important that psychologists should try to to make prescriptions about the techniques for acquiring a skill, the enterprise involves many separate and complex issues. I shall discuss some of them in Chapter 6. Here my primary concern will be with the characteristics of performing skill, regardless of how the skill may be acquired.

### 3.2.1. *Eye movements in reading*

The first overt behaviour in any reading situation is the movement of the eyes over a page to expose succesive portions of the material to central vision. It is only centrally that we have the acuity to register precise details of visual material (although work on language reading, e.g. Rayner 1978, demonstrates the importance of peripheral information in the guidance of eye movements). At normal viewing distances the area of clear vision is a circle of approximately one inch diameter on the page. The mechanisms of the eye movement system operate to give the reader a series of brief 'snapshots' (known as fixations) of such circles. In fluent reading the duration of each fixation is about 250 ms (a quarter of a second) and the eye moves from one fixation to another in a rapid sweep known as a saccade, lasting about 50 ms. The available evidence suggests that we take in visual information only during the static fixations and not during saccadic movements.

The question of primary psychological interest is how the reader controls the sequence and location of fixations. The simplest hypothesis we could entertain is that eye movements are controlled by low-level features of the input. For instance, fixations might occur in regular left-to-right order, each fixation occurring a constant distance to the right of the previous fixation so that the complete text is fixated over a complete sequence. Such neat regularity is, however, hardly ever observed, either in language or in music. Instead we find vertical as well as horizontal displacements, leaps which omit significant areas, and varying amounts of backtracking.

It is possible that some of these deviations are simply random movements of a system over which we have imperfect control. However, a large body of recent research on eye movements in language reading leads us to reject this view. Instead, it appears that the irregularities in eye movements are under immediate cognitive control. Take, for instance, the eye movement pattern in the vicinity of a very common word such as 'the' (O'Regan 1979). Typically readers do not fixate a 'the' directly and their saccadic movements over this word are larger than average. It seems as though linguistic knowledge interacts with peripheral information to allow an identification in the absence of a direct fixation. The eye-movement control system immediately responds

by 'jumping' over the word it does not require to fixate. This is just one piece of evidence which contributes to the accepted view that moment-to-moment variations in the fixation sequence are motivated by the precise cognitive requirements of a reader at a particular time. The pattern of eye movements thus tells us something about these requirements.

Eye movements in music reading have not received the intensive study that has been given to eye movements in language reading. Nonetheless, the evidence that exists is consistent with the language data. For instance, studies by Weaver (Van Nuys and Weaver 1943; Weaver 1943) on piano readers show that the sequence of fixations is determined by the nature of the music. Piano music is written on two staves, so that it is impossible to view all the notes to be played together in a single fixation; one must fixate first upon one stave, then upon the other. One might imagine that the optimum strategy for reading would be a vertical sweep down (or up) the system, followed by a shift to the right and a further sweep [Fig. 3.1 (*a*)] Weaver found that this pattern was indeed used when the music was homophonic and chordal in nature. When the music was contrapuntal, however, he found fixation sequences which were grouped in horizontal sweeps along a single line, with a return to another line afterwards [Fig. 3.1 (*b*)]. Figures. 3.1 (*c*) and 3.1 (*d*) show examples of fixation patterns obtained by Weaver. Figure 3.1 (*c*) is from a chordal passage, whilst 3.1 (*d*) is contrapuntal.

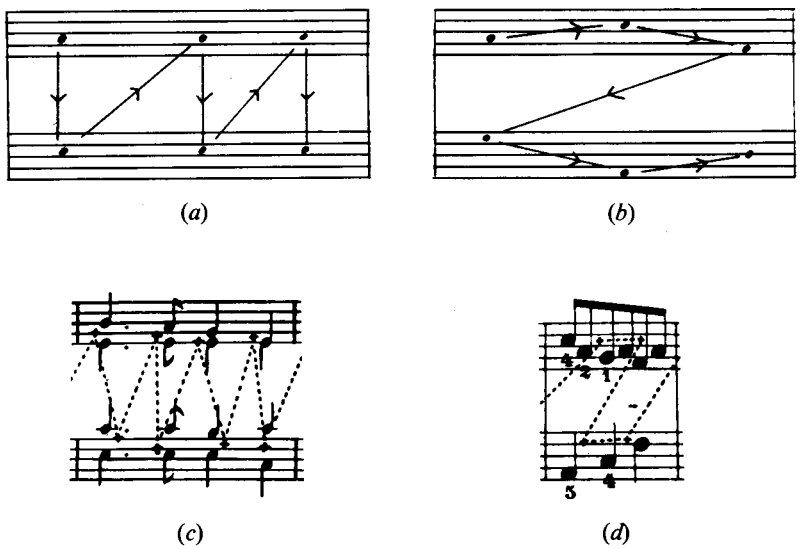



Fig. 3.1. (*a*) Vertical fixation sequence in piano reading. (*b*) Horizontal fixation sequence in piano reading. (*c*) Example of observed fixation sequence in chordal piano music (from Weaver, 1943). (*d*) Example of observed fixation sequence in contrapuntal piano music (from Weaver, 1943).

The general strategy appears to be to identify significant structural units in successive fixations. In homophonic music these units are chords, and so it is necessary to sample both staves in successive fixations. In contrapuntal music the significant units are melodic fragments which extend horizontally along a single stave.

How does such a fixation pattern help the reader? The available evidence suggests that readers require *preview* of structural units in a text if they are to organize fluent and rapid performance. Studies summarized by Shaffer (1976) illustrate this rather well. He examined the behaviour of experienced copy-typists under conditions of controlled preview. Texts were displayed on a computer console. Each time the typist input a character on a keyboard three things happened: all the characters displayed were shifted one space to the left; the leftmost letter disappeared; and a new character appeared at the right of the display. The crucial independent variables were the number of characters displayed at one time, and the number of characters on the screen that had not yet been typed. This latter variable could take any value between a single character and a whole line. It was discovered that if a typist was able to see at least eight characters ahead, performance was indistinguishable from the normal situation with unlimited preview. As preview was decreased, however, performance slowed and became less regular. With preview of eight characters the typing speed was about ten characters per second. With preview of a single letter this dropped to about two characters per second. Shaffer found that it was also possible to slow speed by leaving preview in normal range but manipulating the text. First he randomized word order; this had virtually no effect on performance. Then he randomized letter order within words; this had a highly significant effect, causing speed to drop almost as low as with single letter preview. We can draw quite a firm conclusion from this elegant study. Typists require to preview a whole word, but no more, in order to achieve fluency. It appears that seeing a word in advance allows them to plan a set of partially overlapping hand and finger movements into a single co-ordinated unit.

No-one has yet published experiments with controlled preview in music, although the rapid development of computer music systems (e.g. Tucker, Bates, Frykberg, Howarth, Kennedy, Lamb and Vaughan 1977) makes such studies increasingly more feasible. However, it has been possible to estimate the extent of preview in normal performance by the modification of a technique developed for measuring what is known as the *eye–voice span* in language reading. The technique (as used, for instance, by Levin and Addis 1980, and Levin and Kaplan 1970) involves asking subjects to continue reading out a prose passage *after* the text has been unexpectedly removed. The distance between the word being uttered at the time of removal and the final correct word uttered provides an estimate of the extent of preview. In normal reading situations the eye-voice span of an experienced reader is four to six words. I have modified the technique for instrumental music reading, by

asking various instrumentalists to read single-line melodies at sight (Sloboda 1974, 1977a). When the score was removed, proficient readers could produce up to seven further correct notes. If these performers were looking only so far ahead as required to organize fluent performance, then we may conclude, with some confidence, that the units in terms of which fluent performance is organized are not typically larger than seven consecutive notes of a melody. This seems to be true for all instruments tested, and is not a result of constraints imposed by particular instruments.

Of some theoretical importance is the finding that measures of this *eye-hand span* do not always yield precisely the same result. Depending upon the particular note in the melody where the score is removed, the span can be larger or smaller than average. It appears that there is a tendency for span to coincide with a musically defined phrase boundary, so that a boundary just beyond average span 'stretches' the span, and a boundary just before average 'contracts' it. This strongly suggests that phrase boundaries mark the boundaries between effective performance units which tend to be assembled in an 'all or none' fashion. It also suggests that the typical musical phrase is not, in itself, this unit, because most phrases contain more than seven notes. We must suppose that there exists some principled way of dividing a phrase up into smaller units, analogous to words.

Just as in the typing study, manipulating the 'text' gives further clues to the nature of the performance units. When melodies lacked normal harmonic progressions, span decreased significantly. Similarly, obscuring the rhythmic divisions between phrases by the addition of passing notes also decreased span. It seems that the reader subdivides a phrase into performance units in ways dependent on the harmonic and rhythmic structure of the phrase, and that when cues to these structures are obscured, preview is not so useful, and cannot, indeed, be sustained at normal levels.

Comparison of good and poor music readers showed that poor readers have lower spans than good readers (three or four notes) even when the melodies were structurally straightforward. Poor readers seem to behave with 'normal' music rather like good readers with 'obscure' music. If the ability to maintain preview is dependent on the ability to isolate structurally defined units, then poor readers are less capable of doing this, even when the structures are clearly present.

Before leaving eye movements, it is worth commenting that performance units in music, whilst analogous to words in some respects, cannot be very like them. The English language contains a finite number of words, each defined by a set of letters in a particular order. It is quite feasible for a typist to know every one of these words, storing the unique typing pattern for each word in memory. There is no similar agreed 'dictionary' of musical patterns. It is quite possible for a melody to contain a novel combination of notes. On the other hand, there *are* certainly familiar recurrent configurations of pitch and rhythm in many melodies (such as scale and *arpeggio* pitch movements,

dactyllic rhythms, etc.) which a reader could learn to recognise as units. Thus, music reading probably involves *some* recognition of frequently occurring patterns, but this is likely to be supported by some more general strategy for organizing groupings of notes which can function even when the patterns are novel. One such strategy would be a grouping according to metre. A metrical unit is like a frame, which can be used to hold and describe a very large number of note sequences. The same frame can be applied repetitively to any metrically regular piece of music, bar by bar. Such units are, of course, easy to find in a score. Bars are marked by bar lines; smaller units are often marked by beams. An efficient 'default' grouping strategy would therefore seem to be—take the bar as the performance unit, unless it contains more than seven notes, in which case take the half-bar. Since musical phrases are normally defined metrically (as well as harmonically and melodically) then the observed 'phrase-boundary' effect would be consistent with such a strategy. Furthermore, a metrical organizer has one extremely useful feature—it incorporates a regular pulse by reference to which finger movements can be targetted, and by reference to which the next performance unit can be 'set up'.

In other words, such an organizer is an aid to fluency and pacing. Suppose, for instance, that the organizer being used is a four beat bar with the internal structure signified by Fig. 3.2 (*a*). Level A represents the whole bar, and, if we set a constant tempo, then the onset of each bar can be timed from a clock operating at this level. Organization of the performance of the subsequent bar can be paced against a countdown of the previous bar. An eye movement sequence must be terminated in time to provide the information for the next unit before the countdown reaches zero. Levels B to D represent a possible way of internally structuring the bar, using a hierarchic binary branching tree down to the quaver level. The bar marker generates timing for half-bar units; the half-bar markers generate timings for beats; and so on. Of course, the reader need not use such a deeply structured organizer as this; he could use a simpler two-deep hierarchy such as in Fig. 3.2 (*b*). Data reviewed later in this chapter (sections 3.2.3 and 3.2.4) favours the notion that experienced performers use an organizer more like 3.2 (*a*) than 3.2 (*b*). However, the general principle shared by any such organizers is that they specify *when* each note in a group should occur, so that the performer can initiate ballistic finger movements by reference to the countdown for each note.

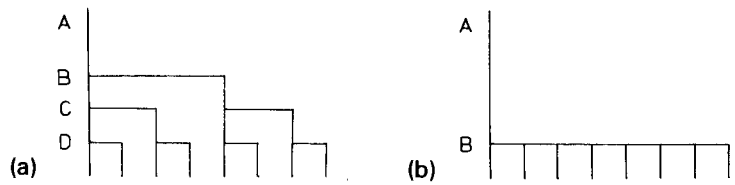



Fig. 3.2. Possible metrical organizers for performance.

### 3.2.2. *Errors in reading*

Eye movements give useful information about the organization of music reading, but it is the performance itself which provides the most direct evidence of reading behaviour. A crude but nonetheless informative measure of reading performance is what one might call 'notational accuracy'—the extent to which a performer plays the correct notes at their approximate notated values.

In a classic study Pillsbury (1897) demonstrated that experienced language readers are prone to make misreadings which transform wrongly printed words into their correct form. For instance 'internalional' might be seen as 'international'. Under many circumstances these perceptual transformations are unconscious. Readers are not aware of the original stimulus; they believe they are seeing the correctly printed word. The received explanation for this phenomenon is that readers are able to identify words on less than complete information, given their knowledge of the spellings of words. For instance, if a reader of English identifies three letters of a four letter word as W–RK, he should be able to identify the word as WORK. There is no other word in the language that fits the available information.

When people read continuous prose they are even more susceptible to this kind of error than when viewing single words. This is because they now have an additional source of information—context. For instance, the sequence 'She raised her wine glass to her . . .' strongly suggests 'lips' as a continuation. It may not be necessary for someone to actually 'read' the word in order to 'see' it. In normally constructed text, such 'seeings' are usually correct. It is when text incorporates the unexpected that trouble can occur. The phenomenon of *proof-reader's error* is well known to psychologists (Vernon 1931). This is the failure to detect spelling or typing errors when scanning through a typescript. The proof-reader makes unconscious inferences which restore the incorrect words to their correct form, and so the errors are not detected.

By deliberately manipulating texts and introducing errors we can learn something about the nature of the inferences which expert readers make. For instance, readers are more likely to spot misprints if they occur at the beginnings of words than elsewhere (Haber and Schindler, 1981; Sloboda, 1976b). People are more likely to overlook spelling errors which are visually similar to the correct form (Healy 1980).

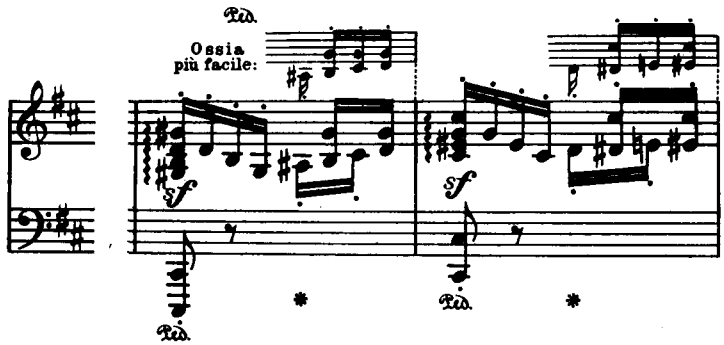
Proof-readers' error also occurs in music. Wolf (1976) quotes a striking example of such error from the experience of a professional pianist/teacher, Boris Goldovsky:

> A student whom Dr Goldovsky describes as 'technically competent but a poor reader' prepared a Brahms *Capriccio* (Op. 76 No. 2) which she brought to her lesson. She began to play the piece through but when she arrived at the C sharp major chord on the first beat of the bar 42 measures from the end, she played a G natural instead of the G sharp which would normally occur in the C sharp major triad. Goldovsky

told her to stop and correct her mistake. The student looked confused and said that she had played what was written. To Goldovsky's surprise, the girl had played the printed notes correctly—there was an apparent misprint in the music.

This misprint occurs in most standard editions of Brahms' piano music. It appears, therefore, that many hundreds of musicians had never noticed the misprint. Goldovsky proceeded to test skilled readers by telling them that the piece contained a misprint which they were to find. He allowed them to play the piece as many times as they liked and in any way that they liked. None of them ever found the mistake. Only when he specified the bar in which the misprint occured were most readers able to spot it.

It is easy to understand why this misprint was so hard to detect. But for the misprint, the bar in which it occurs is an almost exact transposition of the previous bar. The underlying harmonic movement is the very common V–I (from G sharp to C sharp) over a C sharp pedal. In addition, the notation in the previous bar has already 'set' the G as a sharp. There are thus multiple and powerful cues for a knowledgeable reader to interpret the crucial G as a sharp (see Example 3.1).




Example 3.1. Bars 76–78 from Brahms, Capricio, Op. 76, No. 2. Reprinted by kind permission of Peters Edition Limited, London

What is important to note about this story is that it was a relatively *poor* reader who was the first to uncover the error. Because she did not have the expectations of more accomplished players she required more information from the score to determine her performance, and so, paradoxically, read more accurately than more accomplished players.

This demonstration, although striking, is anecdotal, and one of my own experiments (Sloboda 1976b) was designed to verify the phenomenon under laboratory conditions. Example 3.2 shows one of the stimulus pieces used in the experiment. It is by the eighteenth-century composer J. Dussek. The musically literate reader is invited to attempt a sight performance of the piece (preferably tape-recording it for subsequent examination) according to the following instructions, which approximate to those given to my experimental

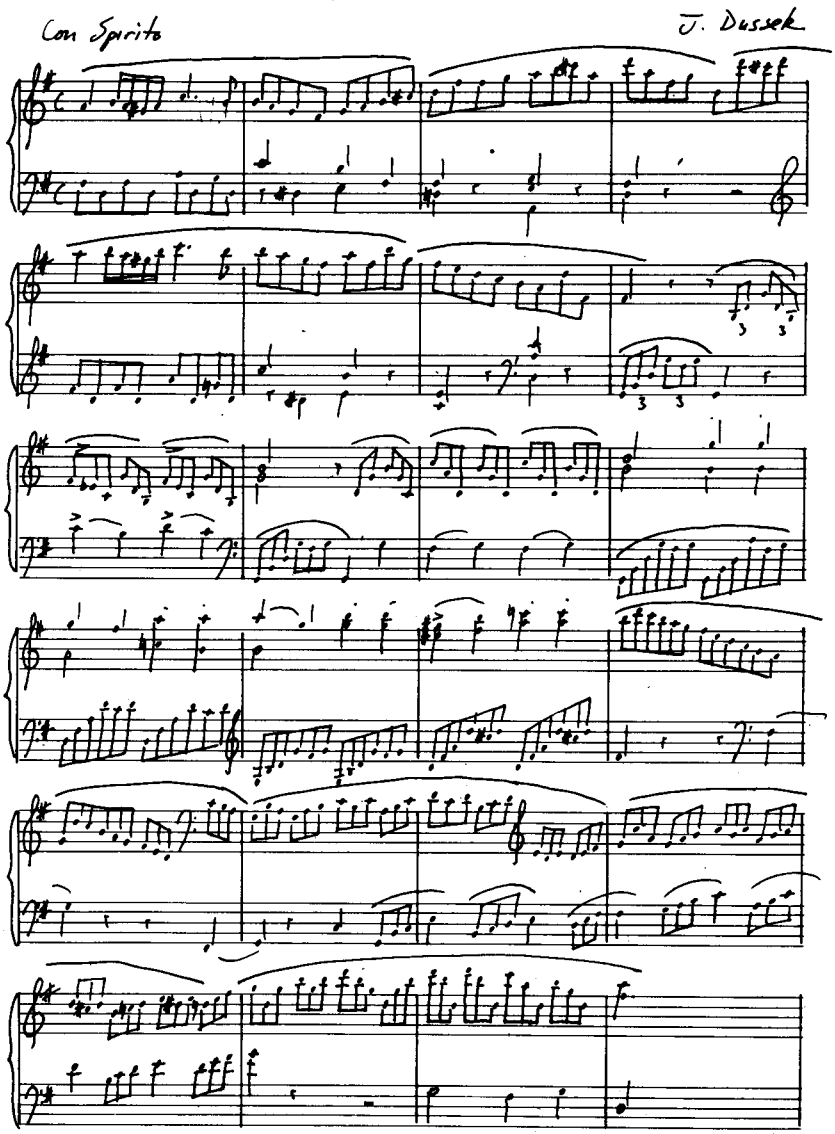
subjects: 'Please sight read this short piano piece at your own pace. Attempt to play exactly what is written, but do not stop if you make an error. Do not extensively preview the piece but start playing within five seconds of looking at the page. You may begin now.'

The stimulus materials were prepared by taking little known keyboard pieces by minor classical composers and rewriting them with deliberate notational errors. The errors had two characteristics. Firstly, they involved a shift in the notated pitch of one step on the stave, leaving all other features (accidentals, rhythm, etc.) unchanged. Thus a notated quaver on B flat might become either a quaver on C flat or on A flat. This had the consequence that the visual appearance of the error was very similar to that of the original. Secondly, they were chosen to be musically inappropriate in their contexts. This could be either by creating a dissonant harmony or by violating normal melodic progression. The errors were judged as 'blatant' by musicians. This is particularly important since there are no definite dividing lines between what is acceptable in music and what is not. There is always a grey middle area where the unexpected may in fact be a legitimate violation of norms which extends and transforms style rather than simply being a mistake. The errors chosen for this experiment were well outside such grey areas. They were downright wrong.

The results of the experiment were quite clear. All subjects tested showed 'proof-reader's error'; that is, they played some altered notes as they should have been rather than as notated. This was even though no subject was familiar with any of the test pieces, and in spite of the fact that every subject 'tumbled' the point of the experiment by noticing *some* of the deliberate errors.

Example 3.3 identifies the 'misprints'. The notes which I altered are ringed. It can be seen that the errors are distributed equally between upper and lower staves. They are also distributed within musical phrases. One third of the errors occur at the beginning of a phrase, one third in the middle, and one third at the end. Table 3.1 shows the proportion of altered notes which were restored to their originals by the subjects, for each stave, and for each position within the phrase. The data for the upper stave (containing the primary melodic material) are strikingly similar to those obtained with words. Alterations in the middle of phrases are more likely to be restored than those at beginnings or ends. We may, therefore, argue that phrases form psychologically effective units such that the reader's attention is concentrated on the material occurring at phrase boundaries, and that some form of inference allows the reader to economize on input from intermediate positions.

A further striking feature of the results is the pattern of errors in a second performance of the test pieces that subjects made. As one might expect, their performances improved, as shown by a significant drop in the total number of performance errors. This shows that they had learned something about the pieces from their first performance. However, the proportion of proof-




Example 3.2. Material from Sloboda's (1976b) sight reading study.

readers' errors did *not* decrease. On the contrary, it *increased* slightly. This suggests that, as they gained a firmer hold on the structure of the pieces, they became *more* likely to infer the expected notes. Clearly, at this early stage in learning, subjects were beginning to store the music, not only in terms of individual notes, but also in terms of higher-order structures. This ability is clearly a hallmark of skilled reading.

It is relevant to ask whether musical proof-reader's error occurs for other performers than pianists. Keyboard music presents particular difficulties precisely because it is on more than one stave, and because several notes are often to be performed simultaneously. Most other instruments are capable of playing only one note at a time, and the music is presented on a single

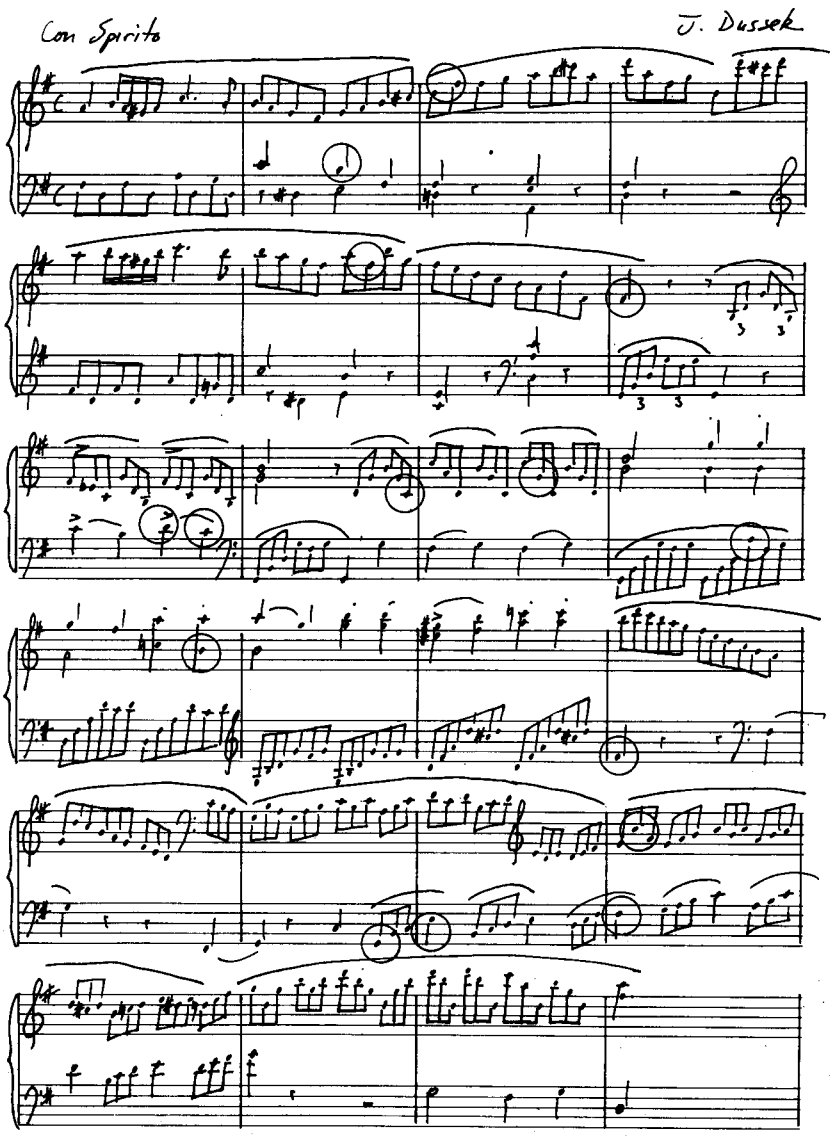
TABLE 3.1
*Number of errors on altered notes at each position in phrase*

| Subject | Beginning | Middle | End |
|---|---|---|---|
| | Upper stave | | |
| A | 7 | 19 | 12 |
| B | 5 | 10 | 4 |
| C | 2 | 16 | 6 |
| D | 4 | 10 | 6 |
| E | 7 | 10 | 9 |
| F | 5 | 8 | 7 |
| G | 7 | 20 | 7 |
| Mean | 5.3 | 13.3 | 7.3 |
| % | 23 | 62 | 34 |
| | Lower stave | | |
| A | 7 | 7 | 6 |
| B | 5 | 7 | 4 |
| C | 12 | 11 | 9 |
| D | 8 | 8 | 5 |
| E | 8 | 10 | 6 |
| F | 9 | 10 | 9 |
| G | 14 | 12 | 7 |
| Mean | 9.0 | 9.3 | 6.6 |
| % | 42 | 43 | 31 |

stave. The 'real time' demands are, therefore, reduced, and inference may be less necessary. This issue has not yet received experimental investigation.

Because it is relatively difficult to induce errors in normal skilled reading, many language investigators have degraded the visual stimulus in some way to make perception more difficult. The most fruitful way of doing this has been to expose portions of text for very brief durations, less than a single fixation. By varying the exposure duration we can learn something about the time-course of information pickup.

I have applied this approach to music in a series of experiments (Sloboda 1976a, 1978) where varying numbers of pitch symbols were displayed briefly on staves for subjects to write down. The advantage of this technique is that




Example 3.3. The material of Example 3.2 with altered notes ringed.

it allows comparison between musicians and non-musicians, since the pitch transcription task is one which non-musicians can easily carry out as straightforward visual copying. Figure 3.3 gives examples of the kind of stimulus material used. Subjects were supplied with blank staves on which to write their responses, which they were allowed to do immediately after the end of the brief exposure. Exposure durations varied from 20 ms up to 2 s.

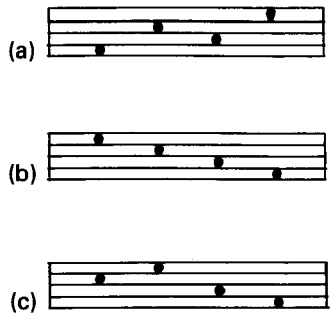



Fig. 3.3. Examples of stimulus materials used by Sloboda (1978).

At the briefer exposures (up to 100 ms) both musicians and non-musicians were very poor at recording the position of notes accurately. They were rarely able to record more than one note correctly. However, when forced to guess all the notes in a display, musicians were significantly better than non-musicians at retaining the approximate *contour* of the original. If one note was higher than its left-neighbour in the stimulus then it tended to be recorded as higher in the response, even though it may not have been placed on the correct stave-line. The ability to retain contour was particularly good when the contour of the original was a straight ascending or descending line [Fig. 3.3 (*b*)] or contained only one change of direction [Fig. 3.3 (*c*)].

We may suppose that durations below 100 ms are just too short for the visual processing mechanisms to gain accurate positional information about the notes, but that more global information about the shape of the note contour is available. This accords with data from other visual experiments (e.g. Massaro and Klitzke 1977; Navon 1977). What is of particular interest here is that the ability to use contour information seems to depend on musical experience. The music reader seems primed to store information about scale or *arpeggio*-like patterns in particular, and this suggests that such information is used in normal reading. This may be because such patterns are usually associated with a stereotyped finger and hand-movement pattern, so that rapid early identification allows advance motor planning. if one can accurately identify just the first note of such a pattern, and one has some contextually supplied knowledge of the harmonic structure at that point, then identification of the contour may be enough to uniquely specify the required action.

As the exposure duration was increased, both musicians and non-musicians increased in their ability to record the notes accurately. However, non-musicians' performance peaked at about three notes, whereas musicians were able to record six notes from a single exposure. This difference remained even at exposure durations of 2 s when there could be no *visual* problems for either group. The superior performance of the musicians is due to their superior *retention* of the stimulus material. They have means of coding and

storing the notes (as music) not available to the non-musicians, who have to remember the stimuli as purely visual patterns.

A partial clue as to the nature of the musicians' coding comes from an experiment by Halpern and Bower (1982). Using a similar technique to mine they varied the harmonic structure of the stimuli. Half of their stimuli were 'good' melodies, following normal rules of harmonic and melodic progression. The other half were 'bad', designed to violate as many musical conventions as possible. They found that musicians performed better with 'good' melodies than with 'bad', but that non-musicians showed no difference between the two types of melody. Clearly, music-specific knowledge was responsible for the advantage enjoyed by 'good' melodies with the musicians. These melodies could be coded into higher-order structures, thereby reducing memory load. However, even on the 'bad' melodies musicians retained a large superiority over the non-musicians. So structure is not the whole story. In my own experiments, many of the stimuli were randomly generated, and so 'bad'. This did not prevent musicians from showing massive superiority over non-musicians. This is rather different to the outcome in Chase and Simon's (1973) classic chess experiments. Here, masters reproduced 'bad' chess configurations as poorly as novices.

One possible reason for this difference is that the music stimuli are far simpler than the chess stimuli. They contain fewer elements distributed over fewer possible positions. Each pair of adjacent notes will have some musical meaning *whatever* their positions. A random sequence of six notes can always be represented as three pairs of notes, each encompassing a common musical interval, so that it is always possible to impose *some* musical structure on the sequence. Another reason concerns differences in experience at the specific task. Although chess masters have undoubtedly spent longer looking at chess positions than have novices, they probably have not had more practice at recording chess positions from memory. In contrast, skilled music reading involves the continuous performance of short sequences of music from memory, and musicians have developed specialized mechanisms precisely for this task that novices will not possess. One could predict that if a chess master devoted considerable effort to developing memory reproduction skills for chess positions, he would soon develop superiority over novices even for 'random' boards.

### 3.2.3. *Expressive aspects of sight performance*

Although errors of reading can reveal something of underlying psychological processes it is often difficult to induce skilled performers to make *any* errors without experimental intervention. By error, in this context, we mean a performer playing the wrong note or making gross errors in timing and rhythm. If skilled performance is error-free, how can we get behind such performance to understand the psychological mechanisms which control it?

One possibility is to ask skilled sight readers what they do. Wolf (1976)

interviewed four expert sight-readers, and although they produced some insightful remarks it is hard to know whether these remarks are more than informed rationalizations about largely unconscious and automatic processes. For instance, they displayed uncertainty about how much music they took in at a single fixation, or how far ahead they looked. One performer, when asked what problems he experienced in sight reading, simply replied 'for me personally, there are none.'

A different approach, pioneered by Seashore (1938) but first applied to sight reading by Shaffer (1980) is to record accurately the microstructure of a performance; the minute variations in touch, timing, and intensity that characterize any human performance. In some cases, these variations may be relatively uninteresting; random fluctuations due to noise in the system (such as limits in the acuity of feedback mechanisms, muscular tremor, lapses of attention). In other cases, however, the variations can be shown to bear systematic relation to structural features of the music.

This is particularly important to our understanding of sight reading in situations where there are no explicit instructions in the score to make variations in touch and timing. Consider a Bach fugue, for instance. Bach generally supplied no dynamic markings, touch, or phrasing instructions. The only explicit information given, in addition to the notes themselves, concerns the tonality (signalled by the key signature) and the metre (signalled by the time signature and the barring). It would be quite inappropriate, however, for a performer to consider that each note should be played in the same way. Such a 'dead-pan' performance, whilst technically error-free on our prior definition, would not be considered musically effective.

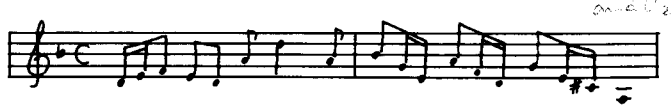
Even inexperienced performers soon become aware that their performance must be expressive. Too often, though, this can mean a haphazard application of timing and intensity changes plastered at will over the music. One simple test of how principled a performer's expressive playing is is to ask him to provide two consecutive performances of a relatively unfamiliar piece. If he is able to provide very similar expressive changes on the two occasions, in the absence of explicit score instructions, then we can infer the existence of some rule-based system for assigning variation to his performance. We can rule out the possibility that such performance details simply demonstrates good memory for the details of the first performance. As Shaffer (1981a) argues,

> the pianist could not memorize this amount of detail: musicians are no better than anyone else at remembering arbitrary analogue information, and Eric Clarke, at Exeter, has demonstrated this by asking pianists to reproduce short melodies they hear, containing timing deformation in a single pair of notes. Almost inevitably the pianist bends the deformation in the direction of a notatable rhythm. Hence the pianist playing Bach could only make consistent use of expressive variation across performance by having an expressive grammar with which to construct the variation according to features in the music.

Shaffer has provided evidence from two consecutive performances at sight of a Bach fugue (*Well-tempered clavier, Book 2*, Fugue 7) by an experienced pianist. A grand piano was linked to a computer in such a way that information about the precise duration and intensity of each note played could be stored (for details see Shaffer 1980). The subject had not played this particular fugue for many years; and following the first performance he was asked to play the fugue again with no instruction to play it the same or different. Shaffer found similarities in expressive performance at many levels. For instance, *within* bars the ranked intensity of individual notes tended to be the same for the two performances. Some notes were given accents on each performance even though the score did not specify this. On a larger scale, increases and decreases in the duration of bars showed a remarkably similar pattern.

What type of rules govern the assignment of expressive variation? Long-term variations, such as gradual changes in speed or intensity, are directed by relatively large-scale aspects of musical structure which often require the performer to appreciate and remember thematic landmarks in the music. For instance, a sequence repeated at increasingly higher pitch can signal the approach of a climax, and an appropriate response is often to increase the intensity for each repetition of the sequential pattern. This requires the reader to spot the pattern, and then visually recognize that what is about to be played is a further continuation of the pattern. At the other extreme, variations occuring within a bar (especially at the beginning of a piece) are unlikely to be determined by the large-scale structure of the piece, and so do not rely on an appreciation of this structure. The variations are linked to the microstructure of the music, the relative importance of the various notes within a bar. One fundamental aspect of this microstructure is the metrical organization, and it is on this aspect that I wish to concentrate.

As discussed in Chapter 2, section 5.4, there are two ways in which a listener can determine the rhythmic parsing for a section of music. On some occasions the patterning of pitches and durations suggest a particular grouping, so that if the performer plays the correct notes for the correct approximate durations, the metre can be determined. On other occasions, however, the patterning of the notes will not suggest a single interpretation, or may even be misleading. An example which caused me some perplexity on first hearing is the opening of Bach's *Clavier Concerto* in D minor (Example 3.4). I took the first two notes to be an upbeat, giving notes 3 and 7 as the major stresses. This intepretation is consistent with the rule that long notes generally denote major stresses. Of course, the correct parsing puts the

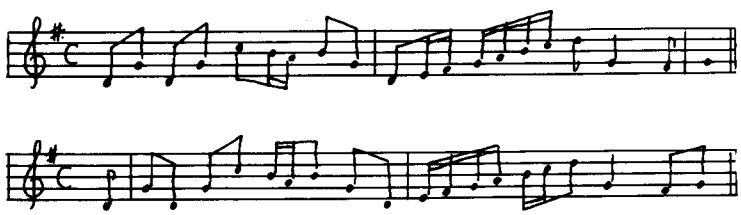


Example 3.4. From J. S. Bach, Clavier Concerto in D minor, First movement.

major stresses on notes 1 and 6. I believe that the only way to guarantee that a listener parses this correctly on first hearing is for the performer to make appropriate expressive variation of a magnitude sufficient to overcome the tendency of a listener to mis-parse.

Using Shaffer's laboratory, I have recently attempted to provide evidence about the use of metrically related variation in piano sight-performance. The basic technique was to present a metrically ambiguous melody under two notational conditions in which bar lines and beams were shifted along the melody, keeping note values and fingerings the same (see Examples 3.5. and 3.6). This is a crucial control, because observing the performance of a *single* notated melody would not be enough to demonstrate that variation was related to notated metre. It might also relate to the precise patterning of pitches and durations, and the fact that different notes use different fingers, etc. In this paradigm we have two melodies whose note and finger patterning are precisely the same. All that differentiates them is the positioning of metrical stress. Therefore, systematic differences in expressive variation *must* be a result of the shift in metre.

Pianists were asked to provide five faultless and similar performances of each of Examples 3.5 and 3.6 at sight. These examples were embedded in a larger number of short melodies, and no subject noticed the fact that the two melodies were, in fact, identical. Each set of five performances represented the spontaneous response to each melody in isolation.



Examples 3.5. and 3.6. The melodies used in Sloboda's (1983) study.

A note-by-note analysis of the differences in timing and intensity (see Sloboda 1983, for full details) showed there to be many significant differences between the performances of the two melodies. For instance, five out of the six pianists played the first bar of Example 3.5 in such a way that note 5 was more legato than its predecessors. In Example 3.6, however, note 5 received no special treatment. This is one example of what emerges from the data as a 'half-bar' effect. The subjects used a variety of devices to differentiate the note immediately following the half-bar (and so falling on the third beat) from its predecessors. These were essentially either an increase in intensity, or an increase in legato touch, or a slight delaying of the onset of the following note, or a combination of these. Individual differences were apparent in the nature of the choices made from these options.

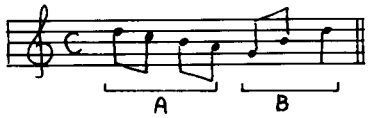
The devices discovered will be of no surprise to performing musicians,

corresponding roughly as they do to notions of *accent*, *slurring*, and *tenuto*, respectively. However, the fact that pianists can deploy these devices at sight shows that they are cued rapidly and fairly automatically from a first inspection of the score. Since there are no explicit notational cues which mark the half-bar, the performance variations must arise through some analysis of the notes which assigns them to structures in which the third beat of a four-beat bar achieves special prominence. Of course, the first beat of each bar is also given special treatment in performance, but there is an explicit notational cue here, the bar line.

These preliminary data leave us a long way from understanding the total process by which metrically related variations are assigned to a performance. The main problem is that a performer does not mark every stress, and when he does mark it he makes a choice from the available options, as well as deciding on the *magnitude* of the variations he chooses. We may suppose that individual performers develop 'styles' so that one makes more use of touch variation, another of intensity variations, and so on. Highly proficient performers may even be able to alter performance style at will within certain limits.

There is also the question of relating performance to perception. Expressive features of performance are useless unless they can be detected by listeners. Similarly, it is pointless to expend energy on achieving levels of performance accuracy (for instance, in timing or tuning) that cannot possibly be appreciated by listeners, because the performance resolution is much finer than the discriminatory powers of listeners. This area is one in which little research has been carried out. The principal problem is that one cannot simply use data from standard psychophysical measurements to make prescriptions for performers. Under 'ideal' conditions a listener may, indeed, be able to detect certain small differences in timing or tuning, but in the welter of a complex musical event discrimination may be much poorer. Experienced performers soon come to realize just how much they can 'get away with' in live performance. I have often been amazed, when listening to a recording of my own performance, just how unnoticeable were errors which, at the time of performance, struck me as catastrophic. Indeed, part of the art of sight reading is knowing which parts of the music will not be salient for a listener. One learns how to create an impression of accuracy in a performance that is actually far from faithful to the score.

We may view the process of assigning expressive variation at sight to comprise three major stages. First is the formation of a mental representation of the music, from examination of the score, which identifies features to be marked expressively in performance. For instance, a four-beat bar such as that depicted in Example 3.7 may be represented as two equal sub-units nested within the bar. To respond appropriately to a rising sequence such as that depicted in Example 3.8 more complex representations based on pitch and tonality need to be computed. The three-note pattern which repeats itself




Example 3.7.

three times at higher pitch suggests increasing intensity through A1, A2, to A3. A tonal representation of the scalar and cadential approach to the tonic will place B as the culmination of the sequence, and so suggest an intensity climax at B. The elements A1–3 and B also have internal metrical structure which, if represented, can provide the basis of expressive variation *within* each element. Since the first note of each element falls on a major stress within the bar it could be marked by a greater intensity than its neighbours. A possible intensity contour that could be derived from computing these multiple representations (scalar, tonal, and metrical) might be:

3 1 2 | 4 2 3 | 5 3 4 | 64445

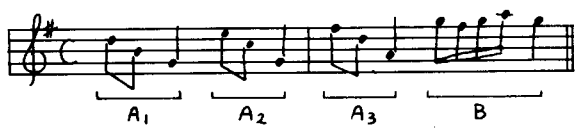
where the numbers represent positions on an ordinal intensity scale. It is, of course, possible to represent a melody on only one of these dimensions and still produce *some* expressive variation. The dimensions can be computed independently. For instance, someone who computes only a tonal representation might assign the following intensity contour:

3 3 3 | 4 4 4 | 5 5 5 | 66666

whereas someone computing only a metrical representation might assign this intensity contour:

4 2 3 | 4 2 3 | 4 2 3 | 42223

There are, of course, further levels of representation which would take us beyond a single phrase such as Example 3.8 to place it in relation to other thematic material in a longer piece. Still confining ourselves to intensity variation, we may want to set the overall level for the phrase with respect to what has gone before or what is to follow. Let us suppose, for instance, that Example 3.8 was preceded by an identical two bar phrase in a longer piece.




Example 3.8.

We might, then, wish to represent Example 3.8 as an 'echo', in which case the entire four bar section might be assigned the following intensity contour:

6 4 5 | 7 5 6 | 8 6 7 | 97778 / 3 1 2 | 4 2 3 | 5 3 4 | 64445

in which the calibration for the second phrase is reset three intensity steps lower than the first.

We may hold it as a working hypothesis that the smaller the amount of music a reader has to preview to discover the cues for a particular representation, the more likely it is that the reader will form that representation. In much music, preview of one bar is the maximum required to form an appropriate metrical representation. The important metrical groupings fall within a bar's span. On the other hand, tonal progressions can span many bars, so that preview becomes less possible, and accurate prediction becomes the only way of making the right expressive decisions. As a performer becomes increasingly familiar with an idiom such prediction is increasingly possible, but all performers will experience times when they take a 'wrong view' of a piece of music peformed at sight. For instance, one may believe that a particular cadence is the resolution of a section, and so emphasize it through *crescendo* and *allargando*, only to discover that it is the penultimate cadence of the section—by which time one has 'spent one's last ammunition.'

The second stage in the expressive process involves gaining access to a 'dictionary' of expressive variations which are accepted as effective in communicating the structural markings which have been identified. In other words, we must assume a mutually agreed 'language' of expression between performers and listeners. It is possible that part of this language is based on relatively primitive mechanisms of auditory grouping in temporal sequences which all people share. For instance, a note which is louder than its neighbours will, all other things being equal, be heard as separate from them. Some of these mechanisms are discussed further in Chapter 5.2. Other aspects of the language are culturally determined, and depend on the shared musical experience of performer and listener (for instance, the use, in classical music, of an *allargando* to signal the final cadence of a movement). Paradoxically, it may be this second type of variation which is easiest for a performer to learn and apply, because it *is* a self-conscious product of a culture to be explicitly identified and talked about.

On the other hand, the small-scale moment-to-moment variations which make performances 'alive' are not so easy to describe. In listening to a master performer we experience his grasp of the *structure* of the music (i.e. we are enabled to form a representation of the music resembling that which dictates his performance) but we tend to lose the information about the precise means by which the form is conveyed. As we noticed earlier, people are very bad at remembering analogue information. In terms of metrical information, we may experience a very keen sense of metre through a good performance but be quite unable to say *why* the performance made the structure so transparent. This phenomenon is not confined to music. We find it anywhere that performing expertise is involved. A master can reveal to us the dramatic impact of a difficult monologue that is just dull and wordy in the hands of a lesser actor; but the general listener appreciates only the power of the

interpretation, not the precise means by which the power is achieved.

Performing expertise first requires analytic listening powers of a developed kind so as to be able to 'latch on to' the minute timing and intensity variations that make a master performance and then *imitate* them. In my own experience, and that of many musicians, there is no really satisfactory way of *describing* expressive variations in a way which allows one to incorporate them into one's own performance. Expressive techniques are passed from one musician to another by demonstration. That is why master musicians are nearly always passionately interested in hearing the performances of other masters. It is why good teachers *have* to be, in my view, good performers. I can recall occasions when my own teacher taught me more by *playing* the passage under discussion than by talking about it for many minutes.

If, however, performing expertise requires listening expertise, we are left with the problem of what allows a person to listen to, and store the kind of detailed information that most of us find so difficult to retain. As an instrumental teacher I have noticed very great differences between even quite young children in the ability to notice expressive variation in the performance of others. For many learners, teaching by demonstration is useless simply because they do not appear to pick up what is being demonstrated. Others are capable of immediate and accurate imitation and retention. I view this ability as one of the best predictors of high levels of musical achievement. It is unfortunate that the cognitive bases for it are not well understood.

The third stage in the process is what has often been referred to as 'motor programming'. This is the setting up of a sequence of commands to the performing muscles which will reveal in sound the expressive devices selected from the 'dictionary'. There are two important features of fluent behaviour which a theory of motor programming must account for. One is the fact that elements in a motor sequence overlap in time. When, for instance, a performer executes a scale passage, the second finger must already be moving towards its position before the first finger is released. Because the degree of temporal overlap is variable (depending, among other things, on the speed of performance), the second movement cannot be initiated on the basis of some cue from the first movement. This is one of the classic problems of serial behaviour outlined in Lashley's seminal (1951) paper, which has greatly influenced current thinking. The second, related, feature is that fluent motor behaviour results in specified *goals* being achieved rather than specified movements being carried out (Shaffer 1982). In terms of piano playing the goal of motor programming is to ensure that a particular note is begun at a particular time and played with a particular loudness. This will be achieved by different *movements* depending upon the preceding hand position and also by what must follow. Expert pianists may use different fingerings on two consecutive performances of a passage, yet achieve precisely the same sound each time.

These features emphasise the fact that skilled human performance is seldom a rigid movement sequence where each movement is triggered in an

inflexible way by the preceding movement, and forces us to realize that performance is the result of an interaction between a mental *plan* which specifies features of the intended output and a flexible programming system 'which has learned through experience to compute the patterns of muscle contraction that will achieve the goal (of producing the specified output) over a broad range of starting conditions' (Shaffer 1981b, p. 331).

This conceptualization accounts for many aspects of musical performance. For instance, it makes sense of the phenomenon of knowing exactly what a passage of music played at sight *should* sound like, but being unable to negotiate the correct notes. In this case the performer has formed an adequate plan of the music he is playing. In terms of my three-stage division of the process he has successfully carried out the first two stages; forming a structurally based representation of the music; and assigning expressive variation to elements of the representation by reference to an expressive 'dictionary'. The failure comes in getting the hands to do what he knows they should do. A common cause of this in my own experience of piano playing is miscalculation of an unfamiliar leap, or 'running out of fingers' in rapid passagework. At such points the expert sight reader can invent a new plan, consistent with the structural context, which he *is* able to program, so 'covering' his error with a plausible alternative.

The dissociation between plan and program also allows us to understand something which many teachers identify in their (otherwise musical) pupils as a 'failure to listen to themselves'. This typically occurs when a teacher criticizes a pupil for failing to make an expressive variation in a case where the pupil is convinced that he *did* make the variation. What could be happening here is that the pupil has formed a plan of the music where a particular note sequence is marked for expressive variation, but that the motor programming does not take account of the marking (or takes insufficient account for reliable auditory detection). The pupil bases his judgement that the expressive variation has occured on the basis of an internal monitoring of his *plan*, whereas the teacher is monitoring the *output*. The inability to monitor one's performance at the level of expressive detail is quite common, and not confined to music. Many people are unpleasantly surprised when they experience recorded versions of their own verbal or motor behaviour.

Finally, the dissociation helps to explain why two quite separate activities are necessary for the acquisition of high degrees of musical skill. On the one hand, a performer needs to listen to, analyse, and discuss, much music. These activities allow him to be able to represent music more fully; and provide him with richer plans. On the other hand, he also needs to indulge in many hundreds of hours of practice at his instrument. To a certain extent, the precise nature of what he plays is not important so long as it presents to the motor system the whole range of ballistic problems which will be encountered in the musical repertoire for his instrument. Only such practice will expose the motor system to the 'broad range of starting conditions' that will allow the

system to build up effective computation routines for solving programming problems rapidly and accurately in new situations. The pedagogic use of scales and studies has, therefore, a very sound psychological basis. They ensure that the learner is exposed to the whole range of programming problems in a systematic way.

The two types of activity described here can, of course, be carried out entirely separately, and this accounts for the existence of two common 'types' of musician. The first type is the musician who can play relatively simple music with the utmost sensitivity and has a profound critical appreciation of other people's performances, but falters when high levels of speed or fluency are required. Such a person tends to spend a lot of time involved with music but neglects systematic practice. He has a highly developed skill for representing musical structure, but his motor programming is inadequate to convert representations into sound. The second type is the musician who can tackle the most demanding pieces in the repertoire from the technical point of view, but often performs them insensitively. Such a person tends to spend hours each day at his instrument, diligently attending to scales and other technical exercises, but neglects to deepen his understanding of music through analysis and critical listening. He has a finely tuned motor programming system but possesses impoverished peformance plans on which the system can act.

The master musician, of course, combines excellence in both these separate skills. Fluent sight reading is, it can be argued, a prime example of a skill which demands both these sub-skills. One needs a high degree of representational skill to be able to rapidly construct an appropriate performance plan on the basis of visual information where relevant structural cues are often only implicit, and one needs a high degree of motor programming skill to be able to assemble a fluent motor sequence at an appropriate speed. If a selection panel for some instrumental award had only five minutes in which to examine each candidate, then I believe there are good arguments for devoting the whole of that time to an observation of a candidate's sight reading performance. There can be few tasks which demand as complete a musicianship.

### 3.3. Rehearsal

Most musicians probably expend the majority of their musical time and effort on rehearsal. The nature and quantity of rehearsal carried out is, therefore, likely to be the most important determiner of performing skill. However, we know almost nothing about the precise ways in which musicians of differing skill go about their rehearsal. The undoubtedly useful and insightful pedagogical literature (e.g. Buck 1944; Hughes 1915) is largely built upon informal observation and personal experience. From this wide body of experience we can glean certain generalizations. Firstly, many novices find it difficult to

rehearse in a productive manner (i.e. one that significantly improves performance). Their principal strategy seems to be to simply play through a passage again and again, hoping, as it were, that the mistakes will be 'ironed out'. Teachers and expert performers insist that more is required. Troublesome sections must be 'taken to pieces'. Secondly, experts insist that secure knowledge of a piece of music involves forming multiple representations of it. One of the principal problems facing moderate performers is the dissociation of the performance sequence from full conscious control. Each short section is cued by association from the preceding section so that when something goes wrong the performance cannot be rescued; memory simply 'dries up'. As a teacher, I have noticed that pupils often can gain entry to a memorized piece only at certain points. If they are stopped for discussion, they have to go back to one of the entry points to continue the performance. In contrast, experts insist that full knowledge of a piece of music entails the ability to start in at any point at will. The conscious mind should know what is coming next independently of what the fingers are doing. Achievements which experts advocate as a guarantee that a piece is properly learned include the ability to think a piece through away from the instrument, or to write it out from memory. Thirdly, most teachers acknowledge that rehearsal is transferable, so that each piece rehearsed lessens the rehearsal requirements for future pieces. Conversely, there are times when intensive rehearsal of a single piece fails to yield significant improvements. Some limit or bottleneck is reached which can only be overcome by moving on to other pieces and returning to the troublesome piece some months, or even years, later.

I wish to concentrate the substantive portion of this section around a study by Gruson (1981) which is, to my knowledge, the only recent attempt to look at musicians' rehearsal behaviour in a detailed way. Her subjects were forty piano students taking individual lessons and three professional concert pianists. Their experience was roughly assessed by noting the level of achievement on the local examination board's grading system. There are ten grades in this system which would bring a performer up to the level of proficient amateur, and each grade represents roughly one year's work on the instrument. The level of experience of the professionals was, of course, way off the top of this grading scale. Each subject was given three novel pieces, consistent with his or her grade level, to practise for 30 minutes. An audio-tape recording was made of each practice session; and then each recording was scored according to a detailed observational scale (devised by Joseph 1978). The choice of pianists as subjects was arbitrary. We have every reason to suppose that their general practising strategies are typical of musicians.

Taking all subjects together, Gruson found that uninterrupted playing accounted for about one quarter of the total practising time. The other most frequent occurrences were, repeating a single note (16.9 per cent), repeating a bar (16.7 per cent), slowing down (16.0 per cent), and errors (14.3 per cent). Of these behaviours, two (repeating a single note, and making errors) are

specifically warned against by many teachers as counterproductive. Repeating a single note is unhelpful, it is argued, because the aim of rehearsal is to build up fluent integrated performance units of several notes. Errors are unhelpful because they tend to reinforce inappropriate motor patterns; the ideal is to start by playing so slowly that errors are never made. At the other end of the scale, certain advocated rehearsal techniques almost never occurred among this group of subjects. Playing with hands separately occurred only 0.97 per cent of the time, and repetitions of sections larger than the bar occured only 2.6 per cent of the time.

To examine the differences in rehearsal behaviour between pianists of different levels Gruson performed two analyses. First she calculated correlation coefficients between the incidence of each behaviour category and skill level. There were four significant positive correlations. These were behaviours which increased in frequency with increasing skill. They were: repeating a section larger than a bar ($r = 0.72$), playing hands separately ($r = 0.49$), verbalizations ($r = 0.37$), and time spent practising each piece (as opposed to playing something other than the music, stopping, etc., $r = 0.40$). The verbalizations included particularly self-guiding statements like 'slow down', 'I made a mistake, I'd better try again'. Three behaviours showed significant negative correlations; that is, they decreased with skill. They were: making errors ($r = -0.31$), repeating single notes ($r = -0.31$), and pausing for more than two seconds ($r = -0.31$). These results show that conventionally advocated rehearsal techniques increase in frequency with increasing skill, whilst non-advocated techniques become less frequent.

Gruson's second analysis used discriminant function analyses to determine which combination of practising behaviours could most parsimoniously discriminate students into three groups corresponding to apprentice (grades 1 to 8), senior student (grade 9 to diploma), and professional. This gave broadly similar results to the correlational study, and yielded repeating sections as the most reliable discriminator. Seventy-nine per cent of the subjects could be accurately classified on the basis of their use of this single behaviour. Gruson's hypothesis is that the increased use of repeating sections indicates an increased awareness of the structure of the music to be learned, where the sections repeated are isolated as significant units. The breaking down of the music into appropriate units becomes an aid to memorization and the construction of a fluent performance. For this hypothesis to be fully corroborated, we would need to see whether the sections repeated could, in fact, be isolated on *a priori* musical grounds.

Gruson's study had a second stage in which she asked a subset of her subjects (grades 2 and 6, and diploma students) to continue practising the pieces over ten sessions. This was in order to see whether continuing experience with a piece of music altered the way in which it was rehearsed. On the whole, she found few changes at any level of skill. The subjects seemed very consistent in their rehearsal behaviour, so that those behaviours which

discriminated between them on the first session (particularly repeating sections) still discriminated between them on the tenth session. It looks as though skilled pianists have a repertoire of rehearsal strategies which serve them well right through the acquisition of a new piece; the less skilled are equally consistent in adopting strategies which may not be so advantageous. Gruson concludes that the major influence on rehearsal behaviour must be the experience gained from 'many hours of practising a wide variety of music pieces'. On the most cautious calculation 'many' must mean 'thousands' at the very least.

Gruson's study clearly leaves many questions unanswered. For instance, we do not know whether the differences in rehearsal strategy actually make the rehearsal more effective. It would be interesting to elicit trial performances after the first and last practice sessions. If the strategies of the experienced players are more effective, then we would expect them to show a greater improvement between the first and last sessions than the beginners. Another difficulty with the study is that age is confounded with experience. The beginners tended to be under the age of eleven, whilst the professionals were all over twenty-five. Some of the observed differences might be due to age changes in general cognitive strategies rather than specific musical experience. A study is needed where the different skill groups are matched for age. None the less, Gruson's study is a pioneering example of the kind of work which urgently needs to be done if we are to gain further insights into the acquisition of performing skills.

### 3.4. Expert performance

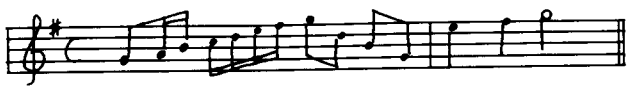
Expert performance demands a large number of sub-skills. Each of these sub-skills merits psychological investigation, and some of them have received considerable psychological study. There are specific technical skills such as vibrato in string playing and singing (Clarkson and Deutsch 1978; Seashore 1938), hand independence in piano playing (Shaffer 1981b), dynamic control of wind players (Patterson 1974), synchronization in ensemble performance (Rasch 1979), and so on. But clearly the expert is more than someone who possesses these several skills. He achieves his expertise through the simultaneous deployment of all these skills in a sustained fashion, and in a manner subservient to the overall structure of a musical composition.

As a teacher I find that one of the greatest stumbling blocks for many pupils is the requirement to attend to several dimensions of musical experience simultaneously. A pupil may be quite capable of playing a passage with appropriate dynamics, but when asked to add some phrasing then loses dynamic control, and when asked to make some third adjustment loses the other two features. Each skill is present, but to utilize it precludes significant attention to other aspects. We may express this in traditional theoretical terms by saying that each skill requires allocation of attention from a limited

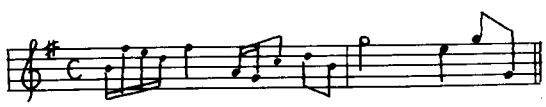
'pool' of attentional resources. Somehow, the expert has managed to escape these limits. It is usually held that the expert has no more attentional resources than anyone else; rather, each skill requires little or no attention for its execution. It has become 'automatic'. In the following section I shall try to unpack some of the implications of this for music performance.

If a performer is able to perform a particular piece of music expertly, then we would be most surprised to discover this to be the *only* piece he could play. This is because we rightly assume that some of the skills displayed in the particular performance are *general* skills, learned through the performance of previous pieces of music, and applicable to future works. These skills are *used* in the learning of a new piece, but they are not significantly altered through the learning of a single new piece. The expert performance of a given piece of music is the result of the interaction of specific knowledge of this piece alone with general knowledge acquired over a wide range of musical experience. This point is worth making since for most performers it will be the *specific* problems of a particular piece which will dominate their concerns. It is therefore easy to forget that specific activities arise from, and build upon, a vast storehouse of general musical knowledge. What distinguishes the expert from the novice is the extent and availability of this knowledge.

Example 3.9 is a simple excerpt around which we can base a discussion of some of the elements of expert performance. What would be involved in arriving at an expert performance of this? Perhaps the first thing to discuss is the role of memorization, since many expert performances are indeed from memory. It is clear that memorization of a piece is a performing advantage since it frees the performer's eyes from the score and guarantees the availability of information about what is to come next. Although the job of memorization is to hold the specific sequence of notes which makes *this* piece unique, there are clearly general skills involved too. These general skills can be demonstrated by considering the additional problems posed by trying to memorize Example 3.10. Although the two excerpts have the same key, metre, and number of notes, Example 3.10 presents more difficulty because its parts are less familiar, and because the principles of construction or patterning are less straightforward. It is easy to give a concise description of



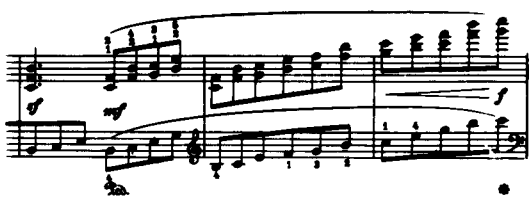
Example 3.9.



Example 3.10.

Example 3.9 such as 'a rising major scale of one octave, tonic to tonic, followed by a descending arpeggio in the same octave', etc. In contrast, it is hard to think of a description of Example 3.10 which would be more concise than an enumeration of the notes.

An important skill of memorization is the ability to encode music in terms of familiar groupings and structures. Sometimes the structure is 'hidden' and adequate performance can be achieved only when the pattern is discovered. An example from recent personal experience comes in the piano part of Samuel Barber's song 'Sea-snatch' (Op. 29 No. 6). Example 3.11 shows bars 7–9 of the piano score. I made no headway with this at all until I realised that the right hand repeats the same four-chord sequence four times (at different octaves, and minus the very last chord). Having then applied the same fingering pattern to each of the four sequences, practice rapidly improved performance. The principal reason why I did not see this pattern at the outset was the irregular beaming which visually grouped the chords within each bar, breaking up the repeated groups whose discovery was essential to my performance.



Example 3.11. From Samuel Barber, 'Sea Snatch', Op. 29, No. 6. Copyright © 1954 by G. Schirmer, Inc. Used by permission.

On other occasions the need is to find a familiar structure to which the current music can be *related.* We may remember an unusual chord as a more familiar one with something added or taken away, or an unusual melody as a scale passage with an interruption. In music of any length it is possible to use earlier events within the music to assist memory of later ones. Sometimes the same music is repeated exactly, but even when transposed or transformed, the earlier music provides a framework for economical coding. Teachers will often have had the experience of a pupil who despairs of learning a long movement until it is pointed out to him or her how fingerings worked out for one passage can be adapted for other similar passages. Sometimes learners fail to notice that two distant passages contain identical notes and so need only be learned once.

Considerations like these suggest that effective musical memory depends upon the ability to represent music in terms of groupings of notes which can be related to familiar stylistic patterns and structures, and also to other sequences within the same piece. This does not necessarily mean that these groupings will be at the forefront of consciousness during a highly polished performance from memory. Aspects of small-scale structure can become

subsumed into larger performing units which are not 'analysed out' on every occasion. Taking the excerpt from the Barber song again, I found that as my fluency improved I no longer had to consciously *think* of the music in groups of four in order for my fingers to do what was required of them. Gradually the whole section became a single fluent statement, although I could impose the subdivisions on it if I wished. Most performers find that there comes a time when they can play through a well-learned piece completely on 'auto-pilot', devoting their conscious attention to something else. This experience, which some musicians describe as 'floating', and in which the hands seem to have a life of their own, is certainly an exhilarating one. It may, however, have its dangers in public performance. Unless one is supremely confident it is hard to refrain from asking oneself 'do I know what is coming next?' and panicking if the answer appears to be 'no'.

More modest examples of 'floating' underlie practically every example of skilled performance. Take Example 3.9 again. It should be within the capacity of any skilled musician to look at this for a few moments and then give a fluent performance on an instrument. A pianist, for instance, will choose an appropriate fingering. This is not a trivial matter since the options for fingering are many. The art of fingering consists in choosing a sequence which will allow legato where the music requires it and which will leave the hand in such a position that notes to follow are within its compass. There are several legitimate solutions to Example 3.9, but a very common one will involve starting on the right hand thumb (call it 1), following with 2 and 3, and then tucking the thumb under the hand for the C, continuing 2, 3 ,4 ,5, up to the G, and returning down the arpeggio with 3, 2, 1.

This is a good fingering for several reasons. Firstly, it allows the first bar to be played legato. In allegro vivace a semiquaver passage is unlikely to be executed with precision unlesss adjacent fingers are used wherever possible. Continuous legato playing up or down a scale is possible because of the opposition of the thumb to the other fingers. In this example, the thumb can be sliding rightwards under fingers 2 and 3 so that it is in position to play the fourth note on time. Secondly, the fingering allows top G to be played with the fifth finger, so allowing an easy reverse of movement down the octave without change of hand position. Finally, it allows the thumb to fall on the second beat of bar 1, so providing the opportunity for using its extra strength to accent the note if required.

Although there exists this explicit rationale for a particular fingering, most experienced pianists will find themselves playing these fingers without conscious choice. Their hands just seem to 'fall into place'. When pianists talk of music which 'lies well under the hands' they are, in part, reporting that many fingering problems are soluble at sight in this effortless way. They tend to agree about the music of which this is true, reflecting their common experience with a musical/pedagogic literature. Most performers in our culture receive intensive exposure to the classical and romantic repertoire where

scales, arpeggios, and diatonic triads form a very large part of the substance of most compositions. The fingering problems associated with these features will have been solved many times in a performer's career, and it is not difficult to see how a relatively simple set of heuristic rules could provide adequate solutions to a large number of these problems. The observation of musicians' fingering sequences and the modelling of the heuristic strategies which could underlie them seems a particularly tractable and worthwhile area of study in which there is, at present, to my knowledge, no significant work.

Because ease of fingering does relate to past experience, most musicians will have had the experience of coming to a new composer or idiom whoses writing seems at first difficult to perform, but after moderate exposure becomes easier. The piano music of Olivier Messiaen, for instance, often seems unperformable to classically trained pianists who encounter it for the first time. It contains none of the traditional scales and arpeggios, and most of the chords are unfamiliar. However, after learning several pieces it becomes possible to recognize typical chord structures and sequences whose fingering problems are closely related to one another. Gradually, solutions become available without conscious effort.

A substantial objection can be made to this line of reasoning. Surely, one may argue, an expert performer who plays something from memory will have solved his fingering problems during the rehearsal stage and will have settled upon a definite fingering which he memorizes along with the notes. Whilst agreeing that this is possible, and may sometimes happen, I would wish to argue that expert performance is often characterized by the fresh reconstruction of performance parameters such as fingering on every occasion. I have several reasons for believing this. One is that expert performers are often prolific, of their nature and of necessity. By this I mean that an expert performer can learn a new piece to concert standard very quickly, and can have many pieces in repertoire concurrently. The argument from economy suggests that he will only learn what new information is needed to perform the new piece, using his general musicianship to supply the rest. This will, perhaps, be particularly true of orchestral players, accompanists, and exponents of new music. There is just not the time to learn each piece to the level of complete 'floating'.

A second reason concerns the existence of variability between different performances of the same piece by the same performer. The reason that it is worth going to live music performances is that each performance will be different. Regular concert-goers will talk of a performer as being 'particularly inspired' on one day, 'detached' on another, and so on. Doubtless, some of this talk is not due to anything the musician does, but the phenomenon is too omnipresent for it to be all in the mind of the listener. Musicians themselves are aware of times when their performances 'reach new heights' or 'fail to come alive'. If all aspects of a performance were predetermined in a memorized plan, such variability would be inexplicable. If, however, many

performance decisions are made anew at each performance, then it becomes easier to understand performance variability. Of course, such a process does not necessitate variability. A performer could make the *same* decision each time. Variability, therefore, also implies some *optionality* in whatever system it is that controls the construction of a performance. Such optionality is indeed highly probable when we are dealing with complex patterned material like music. Such material often displays structural ambiguity; that is, it can be grouped in different ways if different principles are invoked (rhythmic, harmonic, melodic, etc.,) or a given element can be serving a double function with respect to surrounding elements (see Chapter 2, section 5.4). Performance, on the other hand, is unambiguous—one has to assign definite values of loudness, timbre, and duration to each note. On one occasion a performer may choose performance parameters which emphasize a note in virtue of its rhythmic importance. On another occasion he may choose a less emphatic performance in virtue of the subsidiary harmonic function of the note.

Expert performance is also characterized by what commentators often call grasp of the large-scale structure or 'architecture' of a composition. One factor which may contribute to this grasp is the internal consistency of a performance. In much music we would expect to find the same theme performed in the same way on each of its occurrences through a movement. For instance, if we choose a given articulation for the first entry of a fugue, subsequent entries should have similar articulation (unless there is some principled reason for change). Inexpert performance is often characterized by inconsistency. The way a theme is played on subsequent occasions is likely to be determined by local contingencies. In a fugue, for instance, it is relatively easy for a keyboard player to give attention to the articulation of the first entry, because it is unaccompanied. A subsequent entry may comprise one line among two or three to be performed at the same time, with all the fingering problems this poses; and articulation often suffers as a result.

Over a longer time scale expert performance is characterized by control over the progression of speed and dynamic. It is not a trivial achievement to keep a piece at the same speed throughout, or to return to a starting speed after variation. The marshalling of dynamic resources to arrive at a climax at precisely the right musical moment is similarly hard to achieve. Unintentional 'drift' in speed, mismanagement of *crescendi*, reaching a maximum too soon or too late: these are commonplace faults of inexpert performers, and well known to any teacher.

There is almost no scientific work on the features of large-scale performance which contribute towards a listener's judgement of expertise. A notable exception is the work of Shaffer (1981b) on timing, and we may use his study as an example of the direction that such work could take. He recorded piano performances of classical music in such a way that the time intervals between successive notes could be accurately measured. It was then

possible to examine the variability of timing in some detail. Of particular interest was the *covariance* between note durations. Two quantities covary positively if they tend to increase or decrease together. In a smooth *accelerando*, for instance, the durations of notes near to one another in time should covary positively. Notes at one point in the sequence tend *all* to be faster or slower than notes at other points. In contrast, two quantities covary negatively if one decreases as the other increases. Shaffer's data show that in expert metrical performances of Bach and Bartok the durations of adjacent notes tended to covary negatively. This suggests the operation of a compensatory mechanism such that if one duration is shorter than an envisaged ideal, the next note is given a longer duration, so that the *pair* of notes has a duration close to the intended ideal for that pair. This means that some monitoring and timing is taking place at a level above that of the individual note. An oversimplified, but generally correct, conclusion to be drawn from Shaffer's rather complex data is that timing in expert performers operates hierarchically. The performer generates a pulse which specifies the start of a *group* of notes (be it a bar, beat, or whatever) and then uses learned procedures for timing notes within the group. Clearly, the likelihood of timing 'drift' will be reduced if the performer can operate with such a pulse. One may hypothesize that inexpert performance is characterized by an inability to maintain performance with reference to a pulse. Rather, timing is determined locally by procedures for halving or doubling preceding durations, or for playing common rhythmic patterns. Unfortunately, Shaffer did not compare his expert performers with non-experts; but one informal observation supporting the notion of a deficiency in pulse control is the inability of many inexpert musicians to count out a steady beat whilst playing music where the beat is variously subdivided. This is similar to Ibbotson and Morton's (1981) finding that non-musicians are less adept at a particular tapping task than are musicians. The task entails tapping a steady beat with one limb whilst tapping a rhythm with the other. Curiously, the ability to do this seems to depend partly upon the combination of limbs used. Most people can tap a beat with the left hand against a right hand rhythm, but not vice-versa.

The reader will have noticed a concentration on piano performance in this chapter. This is partly because I am a pianist myself, and have the greatest awareness of performance issues on this instrument. It is also because most of the available research uses pianists. To what extent are the principles isolated for piano performance applicable across the whole range of musical instruments? At the highest level I would wish to claim a degree of generality for the observations made in this chapter. The influence of structure and patterning on the construction of performance plans and motor programs arises from the nature of tonal music and its psychological representation, and not from the nature of the instrument which realizes it. At a lower level, however, the technical and expressive problems presented by the various instruments are very different. Pianists escape a whole set of problems con-

cerning pitch, intonation, and timbral modulation, which face string players, singers, and wind players. The means whereby performers achieve control over pitch and timbre are not well understood, although there is some evidence for the existence of a homeostatic feedback loop which allows performers to make micro-adjustments to pitch so as to keep it in line with an intended pitch (Deutsch and Clarkson 1959).

Modulations of pitch and timbre are not simply 'problems' to be overcome. They are part of the expressive repertoire of music. Spectral analysis of natural sounds produced by, for instance, professional horn players show that, during a single long note, there is not just one typical 'horn spectrum' but as many as 30 or 40 different spectra at different times, which give rise to the aural impression of a 'living' sound. If a computer-synthesized horn uses only one spectrum for the whole duration of a note (leaving aside the initial attack portion as a separate problem), the resulting sound is completely flat and 'un-horn-like'. Furthermore, performers can *control* pitch and timbre modulation to achieve expressive variation related to the musical structure, just as pianists use timing and loudness variations. Researchers since Seashore (1938) have made inroads into the measurement of the dimensions of these variations, but little attempt has been made to account for the control structures which *assign* variations in the service of musical communication. Sundberg's (1978, 1982) preliminary work with singers is a notable exception.

The research of Rasch (1979) is one of the few recent examples of work with instrumental performers other than pianists. He was concerned with *synchronization* between three independent performers in string and wind trios. His professional subjects played classical pieces from their repertoire, and the relations between the onset times of notionally simultaneous sounds were examined. One finding was that the violin lead the other instruments in the string trio by some 4–8 ms on average. This may have implied an intentional strategy which 'highlighted' the main melodic instrument. However, the 'scatter' of asynchrony was quite high, and the mean asynchrony between any two instruments was between 30 and 50 ms. This suggests an error component which would be consistent with a homeostatic timing mechanism which kept on target by corrective overshooting and undershooting adjustments. Interestingly enough, these asynchronies mostly went unnoticed by listeners. The performances of the professional ensembles gave the impression of perfect synchronization. This is somewhat surprising, since in ideal experimental conditions (e.g. Hirsh 1959) subjects can easily judge which of two sounds separated by 20 ms came first. This is one piece of evidence which confirms my earlier remark that we need to examine perceptual acuity in real music contexts rather than generalize from the result of 'ideal' psychophysical studies. There are several possible reasons for Rasch's finding. One has to do with the fact that the different instruments occupied different 'pitch streams' (see Chapter 5.2).

Rasch has also shown that perfect synchronization actually impairs the

ability of listeners to identify the component notes of a chord, so that a small degree of asynchronization contributes to the perceptual clarity of the separate lines. This is a reason why some computer-synthesized polyphonic music lacks the contrapuntal clarity of live performances. Advanced programs incorporate a timing 'jigger' to better simulate live sound.

Finally, Rasch discovered a small number of relatively large onset difference times in his data (i.e. from 100 to 200 ms). These *were* clearly audible to listeners. It is possible that these were technical errors, but we must be open to the possibility that some of them were *intentional* timing deformations related to the musical structure. We cannot, however, know this for certain without controls of the sort imposed by Shaffer (1981b) or Sloboda (1983). For instance, intentional variation should be repeated in a second performance.

Although the literature on expert performance is sparse, it is possible to state with some certainty what the principles underlying such performance are, albeit at a level of some generality. Firstly, we see the existence of knowledge of large-scale groupings or patterns within the music which control performance. In many instances this control is hierarchical; parameters which apply to a larger grouping determine the settings for elements within the group. In contrast, inexpert performance is typically controlled by the superficial characteristics of the musical 'foreground'. Secondly, hierarchical control is supported by highly flexible procedures for solving local problems. These procedures operate rapidly, and often without the need for conscious monitoring. The inexpert performer cannot exercise higher-level control because his resources are fully committed to managing the solution of these immediate local problems. Thirdly and finally, the expert has the means to adequately monitor his own performance and take corrective action before he deviates too grossly from plan. Adequate monitoring is not so much a question of listening as of knowing what to listen for. There must be a clear representation of the intended sound if appropriate adjustments are to be made to the actual sound. It is probable, though, that what allows expert performers to make feedback adjustments so effectively is that they are not dependent on consciously monitored feedback at all times and levels. An important feature of skilled motor programming is the ability to operate relatively independently of conscious feedback. Thus, a performer can choose how and when to monitor his performance, knowing which aspects can be safely left to learned programming procedures. For a fuller discussion of experimental work on feedback in performance the reader is referred to Sloboda (1982), which also contains a brief account of social and motivational factors in music performance, which fall outside the scope of this book.

# 4. Composition and improvisation

## 4.1. Introduction

In the previous chapter we examined how performers go about the task of realizing in sound a pre-existing composition, either held in memory or specified by a score. We now turn to the creative process itself, whereby new pieces of music are generated.

There is a vast body of literature on the musical compositions which figure prominently in our art culture, but most of this deals with the product of composition, not the process. Even critical analysis, which, according to Meyer (1973), 'attempts to understand and explain the choices made by a composer in a particular work' (p. 18), is concerned more with the musical relationships evident in the finished product than with the moment-to-moment psychological history of the genesis of a theme or passage. The analyst may assert that a theme has the structure it does because it was the composer's intention to create a particular balance, or a particular contrast. He cannot, however, say whether this intention ever figured *directly* as a goal in the compositional process. The theme may have arisen as the result of some unrelated strategy, retained because of the subsequent *discovery* by the composer that it did, in fact, possess a quality of balance or contrast that recommended it. It is also possible that certain features of a composition highlighted by critical analysis were no part of the composer's intention, conscious or otherwise. In a system of inter-related sounds as large as, say, a symphony, there are bound to be some relationships discoverable by analysis which were neither noticed nor designed by the composer.

How, then, are we to gain psychological insight into this central and fundamental musical process? It seems to me that there are *four* possible methods of enquiry. First is the examination of the history of a particular composition as displayed in the composer's written manuscripts. Sketches and notebooks, if datable, can show how a composition grew and changed over the time during which it exercised the composer's mind. Some composers. such as Beethoven, made sketches which were returned to and remoulded over periods of years. Even where sketches do not exist we can sometimes discover something about the compositional sequence from the final manuscript. For instance, we shall see that differences of ink in Mozart's orchestral manuscripts show that he often wrote out the melody and bass

line of entire sections before returning at some later date to 'fill in' the harmony lines.

Second is the examination of what composers say about their own compositional processes. Many composers have been reluctant to do this. However, a few have tried to say in a general and retrospective fashion what happens when they compose.

Third is the 'live' observation of composers during a session of composition. This requires a rare degree of co-operation from a composer and, to my knowledge, there is only one published example in the psychological literature (Reitman 1965). During such a session one can observe the notes written down, their sequence and grouping in time. In addition, one may obtain concurrent comment from the subject about what he conceives himself to be doing. It has become usual when observing problem solving behaviour of many sorts to ask the subject to 'think aloud', supplying what is known as a 'protocol' for subsequent analysis.

Fourth and last is the observation and description of improvisatory performance. This is the special case where the composer is also the performer, producing a novel musical statement without premeditation in a public context. One may argue that improvisation reveals more of psychological interest than does a score. The improviser has fewer opportunities to cover his tracks. However, the constraints of improvisation—immediacy and fluency—make it likely that there are processes which improvisation and composition do not share. Indeed, it would be very unwise to assume that all composers or all improvisers go about their art in the same way. The most superficial reading of the musical literature provides abundant evidence that different practitioners habitually go through very different mental steps to arrive at a finished composition. As Neisser (1983) has succinctly expressed it, 'practise does not make all practitioners alike, it just makes them all effective'.

My intention in this chapter is to say something about what each of these four methods can contribute to psychological understanding, with illustrations, where possible, from the musical and psychological literature. I shall be highlighting two facets of the psychological activities involved. One is the persistent occurence of superordinate structures or plans which seem to guide and determine the detailed note-by-note working out. The other is the degree to which these plans can, particularly in composition, be rather provisional. They can, for instance, be changed in the light of the way a particular passage 'turns out'. The reader should be warned that composition is the least studied and least well understood of all musical processes, and that there is no substantial psychological literature to review. Whilst I hope to retain an empirical spirit in what I say, it is more probable here than elsewhere that my story will be inadequate, or simply wrong. I am emboldened to proceed only because of the absolute centrality of this aspect of musical cognition.

## 4.2. Manuscript evidence

### 4.2.1. *Sketches*

A sketch is an incomplete or preliminary draft of a work (or part of it) which was written prior to the final manuscript. Composers have differed in the extent to which they made use of sketches, particularly in the extent to which they were prepared to commit early ideas to paper. The sketches of many composers are nearly complete drafts of the final version, requiring little or no radical revision. Many of the extant sketches by Mozart, Haydn, and Schumann, for instance, display this 'finished' quality (Forte 1961). Beethoven, on the other hand, was a prolific sketch writer, filling many notebooks with fragments and sections which show considerable difference from the finished work. He often worked on a particular composition over a period of years. For instance, a first sketch of the opening theme for Op. 111 is found in a notebook which predates the final version by almost 20 years.

This feature of Beethoven's compositional life has led some commentators to an over-romanticized view of Beethoven as a tormented genius struggling with intractable raw material to mould it into the service of higher art. Early commentators on the sketches offered them primarily as *moral* comment on the character of the composer. Johnson (1980) sees the classic transcriptions and commentaries of Nottebohm (1887) as concerned with 'portraying the demonic opposition [. . . of the raw musical material] and leaving us to marvel at the spiritual power which eventually subdued it' (p. 4).

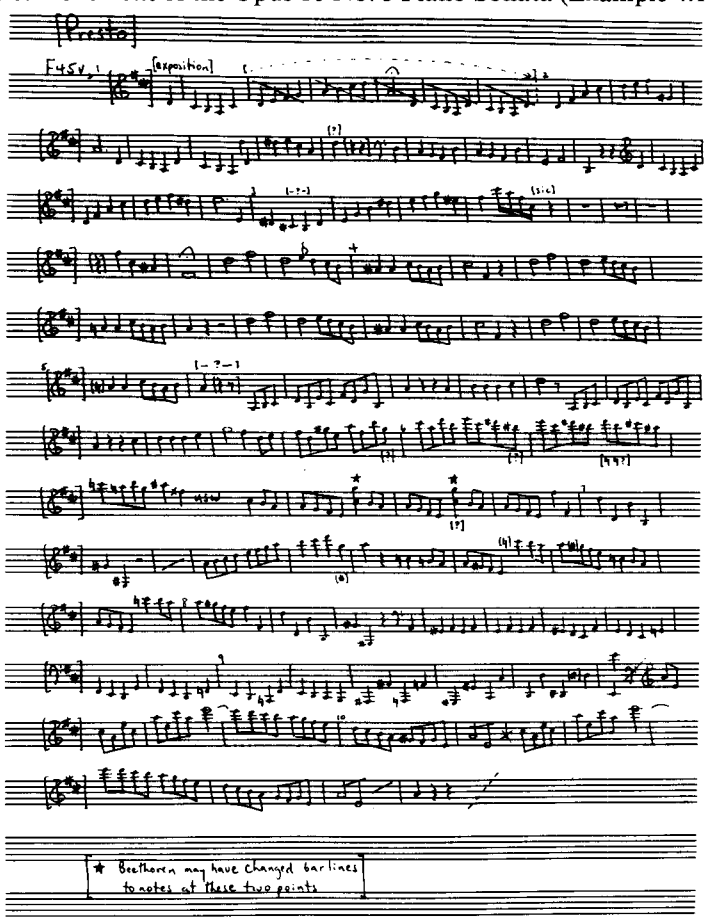
Our concerns are rather different. We must ask whether the sketches are capable of revealing anything about the cognitive processes which led to the genesis of particular works or passages. How did an initial idea come to the composer; what types of transformations and combinations were applied to the original ideas; what goals were being worked towards; in virtue of what were initial sketches rejected as unsatisfactory?

Johnson (1980) argues that musical analysis of the sketches has not yielded any fresh insights into the finished compositions. He identifies two principal reasons for this. One is the overdependence by modern musicologists on prior analysis of the finished work, and the assumption that it alone has the well-formed structure prescribed, for instance, on a harmonic level by Schenker (see Chapter 2). The sketches then tend to be branded as failed experiments—themes which go places that Beethoven didn't want. For us it would, perhaps, be more fruitful to consider the sketches as signs of competence, necessary and enabling resources for the compositional process. As psychologists we seek not so much fresh insights into the finished composition as fresh insights into the *means* by which a composition is finished.

The second reason for lack of progress is identified by Johnson as a deep pessimism among musicologists about the possibility of understanding the compositional act. He quotes Kramer (1973) as symptomatic of this attitude:

'The creative act is mysterious . . . It is an act so complex, motivated by so many impulses—as remote and impersonal as the entire web of knowable history, and as remote and intensely personal as the sum of one man's experience—that the material evidence (records of the act) are little more than occasional memos of a deeper, continual process' (pp. 516–17). Whilst Kramer is undoubtedly right, he describes conditions which appertain to many, if not most, processes of psychological interest. Psychologists, and indeed most other scientists, are used to operating in such situations. Progress is possible because there are ways of achieving a partial understanding of a system on quite restricted data.

Let us, then, take a small example from Beethoven's sketch books which illustrates the potentials and problems of sketch study. I choose the sketch of the first movement of the Opus 10 No. 3 Piano Sonata (Example 4.1) from




Example 4.1. From sketch of Beethoven, Piano Sonata, Op. 10, No. 3. First movement. Transcribed by Johnson (1980).

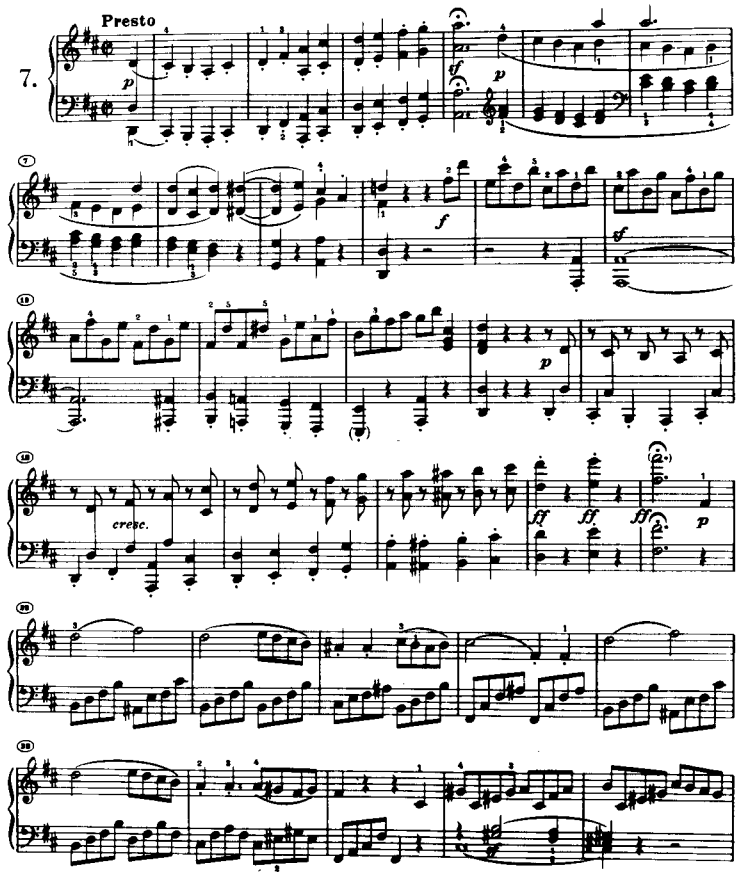
the Fischof Miscellany (transcribed by Johnson 1980; F45V, 1) more or less at random, but because it is reasonably extended, and shows many typical features of his sketches.

The first problem is to know the time scale of the composition of the sketch. Changes in paper, ink, or writing style, could signify a long break in the compositional process; but shorter breaks are of psychological interest too. For instance, what happened in this sketch on line 1 where five bars are crossed out? Were these bars the result of a compositional 'slip of the pen' noticed immediately and corrected, or do they represent a genuine first attempt which was then perceived as unsatisfactory, to be followed by more compositional thought? I believe we have some reason to prefer the latter explanation. The sketch as a whole reveals several clear musical features: first the repeated use of the descending four-note scale from the tonic D as a thematic motif; second the use of the four-bar phrase as the basic musical statement; third, the use of classical sonata form as a large-scale structure.

We can take these three features as constraining what was written, either before pen was set to paper, or very shortly after writing began. Both versions of the first phrase lead the music upwards from a tonic D harmony to the dominant A. In the first (deleted) attempt a simple ascending scale is used. In the second attempt the first half of the ascent is arpeggiated, taking the melody up to F sharp before descending by way of G sharp to the A. The two versions are not similar at a superficial level, but they do honour the constraints outlined above, as does the final version of the theme (Example 4.2) which is different from both sketches. We will never know precisely what motivated the rejection of the two sketched themes in favour of the final version, but the sketches provide compelling evidence that successive solutions were generated within the limits set by initial constraints or plans. Whatever alternative was generated, it had to leave the initial thematic motif intact, and had to move from a D harmony to an A harmony in four bars. We see in this sketch a persistent harmonic and rhythmic plan, albeit provisional, for the *structure* of the section, within which melodic alterations could be made.

Some aspects of such a plan are, however, provisional. This is clear when we examine the final version (Example 4.2). Here we discover that, after the first phrase, Beethoven abandons the strict four-bar phrase scheme. The next three phrases each have *six* bars, so as to produce an asymmetric first subject section of $4 + 6 + 6 + 6 = 22$ bars. This unconventional scheme makes the first section musically novel, although stretching rather than abandoning the original conception. The harmonic movement of the first subject section and the phrase structure of the B minor bridge passage are preserved.

In sum, then, structural similarities and contrasts between the sketches and the final version suggest a hierarchy of constraints within which Beethoven was operating. He was making successive attempts at 'filling in' a structure which, in certain respects, was already specified in his mind. We




Example 4.2. From Beethoven, Piano Sonata, Op. 10, No. 3, First movement. Copyright © G. Henle Verlag, Munchen. Reproduced with permission.

see, however, that some initial constraints (such as the four bar phrase structure) were abandoned during the compositional process in the service of some newly discovered goal.

A second problem for us to consider is the extent to which significant compositional thought takes place at times when the composer writes nothing. When comparing two consecutive sketches it is unwise to assume that they are close neighbours in the compositional process. They may be separated by extensive unwritten compositional thought. In Beethoven's case this is particularly likely in view of one of his own remarks from a written conversations with Louis Schlosser in 1822 (Hamburger, 1952; reprinted in Morgenstern 1956):

I carry my thoughts about with me for a very long time, before writing them down . . . I change many things, discard others, and try again until I am satisfied; then, in my

head, I begin to elaborate the work in its breadth, it narrowness, its height, its depth and, since I am aware of what I want to do, the underlying idea never deserts me. It rises, it grows, I hear and see the image in front of me from every angle, as if it had been cast, like sculpture, and only the labour of writing it down remains, a labour which need not take long, but varies according to the time at my disposal, since I very often work on several things at the same time.

This passage confirms what we deduce from the sketches, that a distinction can be made between the basic idea and the detailed note-by-note working out. The former is present in the mind over a longer time span; the latter is subject to repeated experiment.

A third problem concerns the extent to which a sketch means for an outsider what it means for the composer. It can be argued that the purpose of writing a sketch is to be able to remind oneself of a musical idea that one must lay aside for some time. It is an *aide memoire*, not a full record. A composer may use cryptic or skeletal signs in the knowledge that these will be sufficient to retrieve the whole idea. In Beethoven's case, sketches are often confined to a single melody line without harmony or counterpoint. It would be extremely unwise to imagine that Beethoven thus *conceived* of a melody before assigning it a harmonic or contrapuntal context. Rather, we should conclude that Beethoven wrote this much only, knowing that he would be able to retrieve or generate afresh the remainder from examination of what he had written. A teacher's lecture notes are personalized in a similar way. They allow him or her to retrieve what was in mind at the time of preparation, but there is no guarantee that a colleague would reconstruct the same lecture from the notes.

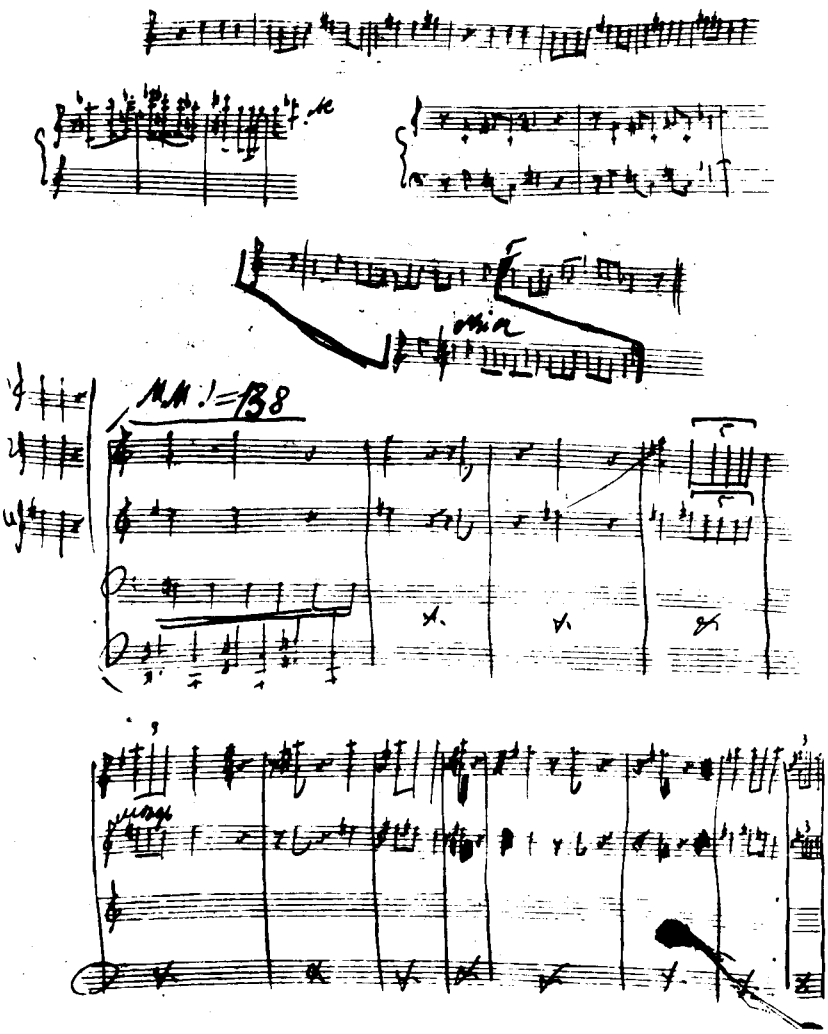
There is neither time nor space here to enter into an exhaustive study of sketching processes in different composers. It is a lengthy task in which significant musicological and psychological effort is still required. However, it is never wise to confine even preliminary remarks to one case. As a comparison I now turn to something from a completely different period and style—the sketches that Stravinsky made for his *Rite of Spring*. A collection which comprises about four-fifths of all the sketches for this work is now reproduced, with commentary by Robert Craft (Stravinsky 1969). In Stravinsky's case there is little need for transcription since the original is highly legible in most places.

There are some instructive comparisons we may make with Beethoven's sketches. Like Beethoven, Stravinsky very often records a mixture of ideas on the same page—sometimes successive workings out of the same material, sometimes different material. Secondly, also like Beethoven, he does not stick to the chronology of the final work, but jumps forwards or backwards in sequence. Thirdly, like Beethoven, many of the sketches are skeletal—a single melody line, a chord sequence, a rhythmic pattern: rarely does one find a full orchestral score. Once again the sketches seem to act as *aides-memoires*,

recording the themes and ideas from which the rest can be reconstructed at a later date.

Stravinsky differs from Beethoven as well. Whereas Beethoven's sketches use predominantly the single melodic line, a great number of Stravinsky's sketches depict several simultaneous strands. The chord structure is often fully specified, together with the main contrapuntal elements, and a preliminary indication of how the parts should be disposed among the orchestral instruments. These differences are not surprising in view of the profound differences in style and content between the music of the two composers. In his early and middle periods, at least, Beethoven worked within a conventional harmonic and formal framework. The 'rules' of diatonic harmony and sonata-form were well known. Within this framework, the melodic element became both the main individuating factor of a particular work, and also the main determining factor for the other features. Someone conversant with the classical style could, without too much difficulty, reconstruct a plausible sonata movement from the melody line alone (indeed, this kind of task is often used as a pedagogic exercise). Stravinsky, however, was working within a set of constraints whose bounds were much less well defined. His use of tonality was novel, both in terms of the immediate harmonic effects he created and also in terms of harmonic movement. For a full analysis of Stravinsky's musical style the reader is referred to Forte (1978). Suffice here to say that in this work Stravinsky was forging a new style with respect to the prevailing musical culture and also extending his *own* style. In such a situation it becomes likely that Stravinsky *needed* to capture more of the textural detail in his sketches. Such detail was sufficiently novel to constitute the *essence* of the composition, not something which could be reconstructed from a single line.

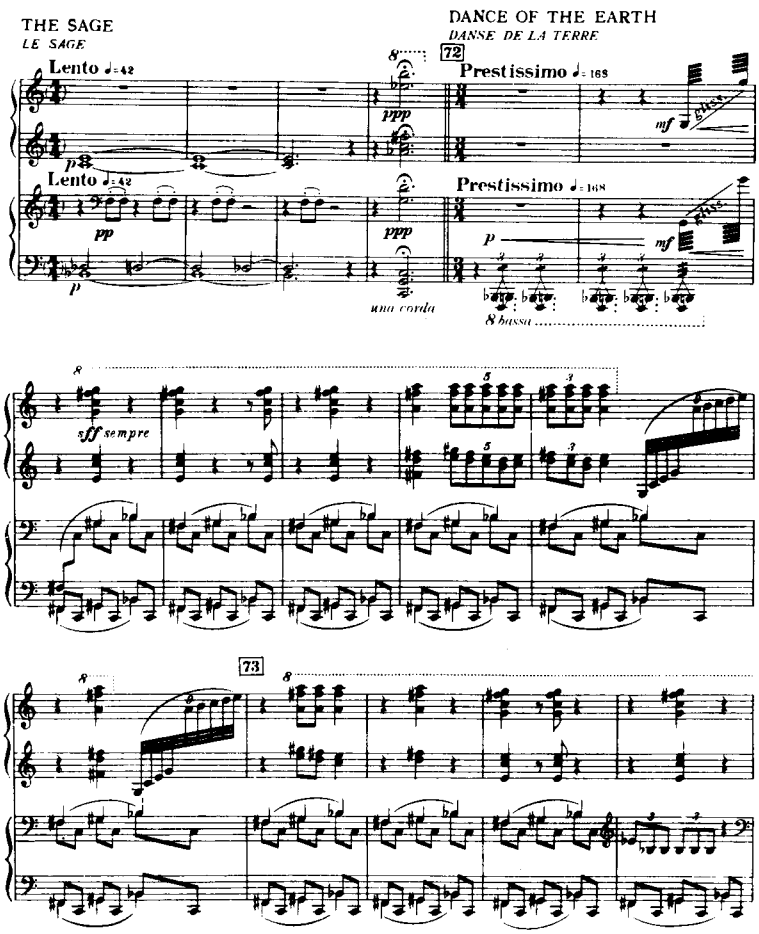
Another difference between Stravinsky and Beethoven concerns the number of preliminary sketches made before something approaching a final version was produced. Stravinsky made, on the whole, considerably fewer sketches than Beethoven. We may often observe single sketches of a few ideas which are then combined into a contrapuntal whole on the same page. The ideas themselves are often unchanged from the first sketches. Subsequent sketches show their combination and extension rather than their transformation. When transformations do occur, they are often radical and bold, leaping to the final version without any notated intermediaries. Craft draws attention to one striking example of this on p. 35 of the sketches (Example 4.3). The page is headed 'Dance of the Earth' and at the bottom we find a sketch which is essentially the final version of the first statement of the main theme of this section. Above it are four melodic fragments. Three of these arer quite clearly incorporated into the final version; but the very first sketch appears, at first, to be quite anomalous. It is a simple folk-like chorale in consecutive thirds which does not appear in the section, or anywhere else in




Example 4.3. Page 35 of Stravinsky's sketches for *The Rite of Spring* (1969). Reproduced by permission of M. F. Meyer.

the work. Despite its unfamiliarity, Craft convincingly argues that it is, in fact, the point of departure of the genesis of this section (given in the final version piano duet arrangement in Example 4.4). He writes:

> We discover, first, that the composer translates the melody from the top to a middle voice (second stave of the piano score in Example 4.4); second, that he forms harmonic aggregates from it, superimposing the notes as if they were *appogiature*; third, that he exploits its whole tone context—the harmonization in major thirds—in an ostinato bass-figure with the F sharp (rather than the C) as the root tone; and fourth, that he renovates the rhythm (p. xxiii).




Example 4.4. From Stravinsky, *The Rite of Spring*, Dance of the earth. Piano duet version. Copyright © Boosey and Hawkes Limited, London. Reproduced with permission.

This is not an isolated example, and Craft uses the sketches to postulate a basic compositional technique for Stravinsky.'Many of the composer's first notations were simple folk-like melodies which he transformed by a number of face-lifting devices including changes in note order, the extracting of essences, and the grafting of new rhythms and characters of tempo.' It is only the existence of these apparently anomalous folk melodies side by side with their transformed versions in the sketches that allows us this unique insight into Stravinsky's compositional processes.

We may ask what allowed Stravinsky to jump from first idea to radically transformed final version in one step? It is unwise to assume that Stravinsky had a surer technique, or that the transformations he used could be applied

with less extensive cogitation than Beethoven. The probable difference between the two composers is the time scale over which each was prepared to let an idea develop. Beethoven kept themes with him for several years. He needed to notate intermediate versions while he moved on to other things. Stravinsky, however, moved from initial conception to final theme over a much shorter time span, during which he worked on few other sections, and certainly on no other compositions. We know, in addition, that he composed at the piano (Stravinsky 1936, p. 14). It is plausible to suppose that ideas intermediate between initial and final versions were tried out on the piano, modified, and discarded, without any need for notation. The time scale was short enough for the whole process to be kept in mind until the final version. The structure of the work—short, contrasting episodes—would help such a strategy. At one level, each small section of 20 or 30 bars could be composed as a self-contained entity.

It is worth commenting, as Craft does, that many of the initial folk-like melodic fragments share common characteristics; such as single step movement up and down the scale, a slightly obsessive character brought about by repetitions of single notes or short motifs, harmonization in parallel thirds, etc. Such similarities lead one to suppose that there is a central core of thematic ideas to which Stravinsky returned again and again for his basic compositional material. It is this thematic unity at the heart of diversity which helps to make the work one of the most powerful compositional achievements of the twentieth century.

### 4.2.2. *The final manuscript*

It is clear that the study of sketches can, despite all the difficulties, provide some valuable insights into the compositional process. As mentioned earlier, however, not all composers have left sketches. Some composers write very little prior to the final draft. This, together with remarks attributed to such composers as Handel and Mozart has led, in some quarters, to a view of their composition as a form of dictation from inspiration, where the sounds of a work parade themselves fully formed in the imagination—the composer's task being simply to capture them with the pen. And so, composers of fluent inspiration, such as Mozart, are contrasted with composers such as Beethoven who had to do conscious work on their material, as their sketches show.

This division of composers into two classes is, I believe, a gross oversimplification based on an erroneous equation of sketching with conscious compositional effort, and on inadequate attention to the details of what composers say about their own compositional processes (see section 3 of this chapter, to follow). Beethoven's comment already quoted on p. 107 asserts that there was much compositional effort before anything was committed to paper, even though he was a prolific sketcher. Also, the evidence we have of

how composers such as Mozart wrote out their final manuscripts is inconsistent with a 'dictation' hypothesis.

It is a little unclear what the constraints of 'dictation' are, but we may start by assuming the constraints that exist in 'real' dictation, where a musician is required to notate an unknown passage played or sung to him by someone else. In this case, the length, speed, and complexity of the passage will be crucial to success. The subject will be engaged in a struggle to 'keep up' with the music, missing out fast passagework and complex harmony whilst making some attempt to get down the more straightforward and slow-moving sections. Clearly, Mozart scores show no evidence of this type of struggle. Perhaps, however, we should loosen the constraints a little. Maybe the composer has control over the speed of 'playback' so that he is able to stem the flow of inspiration for long enough to adequately record all the details. In this case we would expect to see a score written out 'vertically' with each section completely transcribed before going on to the next section. However, recently published facsimiles of Mozart's autographs (for instance, Mozart, 1979) tell a quite different story. It is fortunate for us that Mozart changed ink during the writing out of *Die Zauberflote* ('The Magic Flute'). The second ink used has faded more quickly than the first, so that the passage of time has revealed the writing history. In many of the movements, Mozart's first task was to write out the entire melody and bass lines on the outer staves, leaving gaps for the inner parts. At a later date, using a different ink, he then went back over the score to write out the inner parts.

In order to hold to a 'dictation' hypothesis, this data forces us to postulate some form of storage of the received composition such that Mozart could return to it at will and obtain repeated 'playbacks'. In other words, there would have to be *two* dictation processes: an initial one when the inspirational stream was 'fixed' in memory—a mental process requiring no overt activity; then a series of playback dictations whereby what was fixed in memory became fixed on paper.

Even this story, however, leaves unsolved puzzles. Why, for instance, should Mozart write out melody and bass first? Surely any order of writing out would be as efficient as any other. Suppose, for example, that a playback allowed a single line to be copied. Why not, then, work down the score from top to bottom? Then there is a further puzzling feature of the Mozart manuscripts. There are occasions when, in the first ink, Mozart has filled out the whole score for a few bars. At other points, isolated passages in a few interior parts are written in the first ink. Such points are not musically arbitrary. They consist of linking passages between sections, closing orchestral cadences, important counter-subjects, and other points where interior parts have particular prominence or importance. It seems as if Mozart was anxious to 'fix' these particular passages on paper, perhaps in the fear that they would be less easy to retrieve on another occasion.

We move, then, towards the notion that whatever was held in Mozart's memory was not a simple serial concatenation of notes—some mental tape-recorder which could be turned on or off at will. Rather, it was a structured and sectionalized representation which identified crucial parts within a super-ordinate plan. The first writing allowed Mozart to fix the crucial individuating elements of the composition on paper. These elements would allow the reconstruction of the rest even if the overall plan was lost from memory. For Mozart, writing out was not so much dictation as the negotiation, in a planned way, of a highly structured internal representation.

There is, of course, no manuscript evidence which could possibly be pertinent to the question of a prior 'mental dictation', at least for those composers who did not leave sketches. It is entirely consistent with the manuscript evidence that Mozart 'heard' his compositions as once-for-all hallucinations, played for him on an imaginary orchestra by his unconscious mind. The representation of such 'performances' in a structured memory form could well have been the result of a conscious strategy for retaining the details of the hallucination, just as we argued earlier that such strategies were responsible for success in remembering Allegri's *Miserere* (Chapter 1).

Such a picture is, however, profoundly unsatisfactory for a cognitive psychologist, because it leaves unexplained the unconscious processes by which a highly structured composition could come to appear, fully formed, in the composer's consciousness. If any composer was, or is, such a 'recipient' of unconsciously formed compositions, then the prospect for analysing and understanding the compositional process are not good.

In fact, when we move on to consider verbal evidence, we will find very few instances where compositional processes are totally inscrutable to the individual concerned. The most convincing case of 'unconscious composition' on a large scale is that of Rosemary Brown, who claims to receive compositions from dead composers by dictation. The works so produced, mainly for piano, are certainly coherent, and characteristic of the various composers she names. Although she is a reasonably competent pianist who is acquainted with the classical and romantic piano repertoire, she has no formal compositional training. Her verbal reports and her writing technique are consistent with a literal note-by-note dictation process, and, unless she is the perpertrator of an elaborate hoax spanning many years, her compositions provide incontrovertible evidence for the possibility of unconscious composition on an extended scale. Parrott (1978) reports a first hand observation of Rosemary Brown at work:

> I have watched the process many times and have often filmed it. You will know better than I how normal composition actually occurs: I have certainly never seen anything like the process Mrs Brown uses. The music literally flows onto the paper in a continuous stream—sometimes both clefs together, sometimes one first and then the other. And all the time Mrs Brown chats away: 'Not so fast; did you say natural

or flat? This G or the octave higher?', etc. And as fast as she can write, so it's taken down.

### 4.3. Composers' writings on their own compositional processes

To give a flavour of the kind of insight we can gain from examination of composers' writings, I wish to begin this section with four short extracts from the writings of various composers, without comment or identification.

(*a*) It has been my own experience in creative activity that a motive or a two to four measure melodic phrase occurs to me suddenly. I put it down on paper and immediately extend it to an eight, sixteen, or thirty-two bar phrase, which naturally does not remain unaltered, but after a shorter or longer 'maturing' is gradually worked out in definitive form. . . . This work now proceeds at a rate which depends primarily on my awaiting the moment at which my imagination is capable and ready to serve me further.

(*b*) You may ask me where I obtain my ideas (for a theme). I cannot answer this with certainty: they come unbidden.'

(*c*) Whence and how (my ideas) come, I know not; nor can I force them. Those ideas that please me I retain in memory, and am accustomed, as I have been told, to hum them to myself. If I continue in this way, it soon occurs to me how I may turn this or that morsel to account . . . agreeably to the rules of counterpoint, to the peculiarities of the various instruments.

(*d*) The first stage in the composer's work is . . . 'inspiration.' The composer . . . 'has an idea' . . . consisting of definite musical notes and rhythms which will engender for him the momentum with which his musical thoughts proceed. The inspiration may come in a flash, or as sometimes happens it may grow and develop gradually. [In this latter case] the inspiration takes the form . . . not of a sudden flash of music, but a clearly envisaged impulse towards a certain goal for which the composer was obliged to strive . . .

After inspiration and conception comes execution. The process of execution is first of all that of listening inwardly to the music as it shapes itself; of allowing the music to grow; of following both inspiration and conception wherever they may lead. A phrase, a motif, a rhythm, even a chord, may contain within itself, in the composer's imagination, the energy which produces movement. It will lead the composer on, through the forces of its own momentum or tension, to other phrases, other motifs, other chords. . . .

[The composer] is not so much conscious of his ideas as possessed by them. Very often he is unaware of his exact processes of thought till he is through with them; extremely often the completed work is incomprehensible to him immediately after it is finished. Why? Because his experience in creating the work is incalculably more intense than any later experience he can have from it; because the finished product is, so to speak, the goal of that experience, and not in any sense a repetition of it.

These quotes show, I believe, that composers have been capable of providing

coherent and articulate accounts of their inner experiences. They span about 150 years, yet contain so much that coheres across the accounts that one has some confidence in taking them to isolate some of the fundamental features of the compositional process and the creative experience. Before we examine the psychological implications of these quotations in detail it may be as well to identify them:

(*a*) Richard Strauss (1949); translated Morgenstern (1956).

(*b*) Beethoven (1822); from Hamburger (1952). This should be read in conjunction with the quote on p. 107, which comes from the same letter.

(*c*) Attributed to Mozart (1789?); from Anderson (1938).

(*d*) Roger Sessions (1941).

I would like to focus on the agreement across accounts about the existence of two stages in composition: the first called 'inspiration' by Sessions, where a skeletal idea or theme appears in consciousness; the second called 'execution', where the idea is subject to a series of more conscious and deliberate processes of extension and transformation. This dichotomy is by no means confined to musical composition, and is to be found in every branch of creative activity, as collections such as those of Ghiselin (1952) and Vernon (1970) show. It is as if the creative artist has a consciously known repertoire of things he can do with basic material, but has no such repertoire for generating the first germs on which he exercise his craft. The inspiration seems almost externally 'given', much as a sculptor may be given a lump of stone of particular shape and texture to work upon. Compositional skill entails the existence of a repertoire of ways of extending and building from the given, by discovering and using its inherent properties in principled ways. It is much easier to discover what composers do with themes, through examination of sketches and writings, than to discover where first themes come from. Analytic writing on musical compositions shows how extended works are built from kernel themes by means of such devices as transposition, augmentation, subdivision and recombination of elements, changes of rhythm, etc. Much of what we identify as compositional style inheres in the habitual ways in which composers modify initial thematic material.

Yet even if a composer is selective in the types of transformation he uses, there is still too much freedom. Anyone can learn 'tricks of the trade' and concatenate them together, but the result can be weak and lacking in integrity. It is clear that skilled composers are very hard on themselves, rejecting many permissible developments as unsatisfactory; but just as they are unable to articulate where inspiration comes from, so they are very unclear about the criteria by which they choose among the available ways of developing a theme.

It seems to me that the choice of the 'right' development may be governed by considerations of large-scale structure and balance, as discussed in Chapter 2. The art of composition lies, in part, in choosing extensions of initial thematic ideas that honour superordinate constraints, often to be formalized

in terms of hierarchical structures governing sections or movements. Schenkerian notions of melodic and harmonic movement constitute one way in which such superordinate constraints may be expressed, but doubtless there are others too. We may parsimoniously account for much of the sequence of compositional behaviour if we suppose that skilled composers are often unable to use superordinate constraints as starting points for *generative* processes, but are able to use them to *recognize* satisfactory musical material generated in some other way.

Newell, Shaw, and Simon (1962) have postulated a general theory of problem solving in which they distinguish two classes of process, the first of which they call *solution-generating* processes, and the second *verifying* processes. Typically, solution-generating processes involve the use of heuristics (strategies) for producing candidate solutions. Development of skilled problem solving involves finding better heuristics, i.e. ones which produce fewer bad solutions, and take less time. Heuristics are always less than perfect. They do not guarantee success; and it is arguable that in ill-defined problem solving situations such as musical composition, where the composer is at liberty to change the nature of the problem as he proceeds, algorithms (foolproof solution generators) are of limited value, even if discoverable. Because heuristics are not perfect, there has to be a process of verification whereby trial solutions are tested against criteria for success.

What we are wont to call compositional processes, those events over which composers have greatest conscious control, can possibly be identified as solution-generators in the theory of Newell *et al.* These are the repertoires of ways of 'turning themes to account' by exploiting and transforming their properties in principled ways. The process of judgement, whereby a particular development is accepted or rejected as achieving compositional goals, can be identified as the verification processes postulated by Newell *et al.* It is here where, arguably, testing of a trial solution against higher-order constraints takes place. This process also changes with increasing skill, as a composer develops more comprehensive and large-scale goals. Initally, a young composer may rely on a superficial understanding of the forms of his culture to determine his goals. For instance, he may notice surface features of a sonata form movement, such as the introduction of a second subject in the dominant key. As he grows in knowledge of his medium he may begin to extract the more fundamental principles of harmonic and melodic movement that underlie particular forms. He will understand why the standard features of sonata form usually 'work' musically; but he will also see how and when such features may be dispensed with to produce new forms which also 'work'. The verification is then based on 'intuition' that what has been composed is, in some sense, effective.

Composers throughout the whole of the last two centuries have insisted that the best route to the development of compositional skill is detailed study of Bach and the contrapuntal masters. Composers as varied as Schumann,

Kodaly, and Stravinsky, emphasize, in their different ways, the common principles underlying tonal music of the last 500 years in virtue of which it is fruitful to immerse oneself in the 'old masters'. From this immersion one gains the essential familiarity with the potentials and possibilities of the tonal medium from which one can forge one's own style. Verdi puts the point as succinctly as any (Werfel and Stefan 1942):

> *No study of the moderns!* Many people will think this strange. But today, when I hear and see so many works put together in the way a bad tailor puts clothes together on a standard model, I cannot budge in my opinions. . . . When a young man has gone through a severe course of training, when he has achieved his own style, then, if he sees fit, he can study these works, and he will no longer be in danger of turning into a mere imitator. (From a letter to Giuseppe Piroli, 1871.)

From a psychological point of view we may explain the benefit of such study as forcing the student to concentrate on basic underlying properties of the tonal system rather than on the particular compositional forms which are fashionable in the culture, and which constitute only a small subset of the ways of writing acceptable tonal music.

We may summarize our discussion so far by reference to a diagram of a 'typical' composer's compositional resources and processes (Fig. 4.1). There appears to be a distinction between those processes on which a composer is able to report fairly easily and those on which he is not. For convenience these have been labelled 'conscious' and 'unconscious'. Square-edged boxes depict knowledge or structures that are stored in long term memory. The

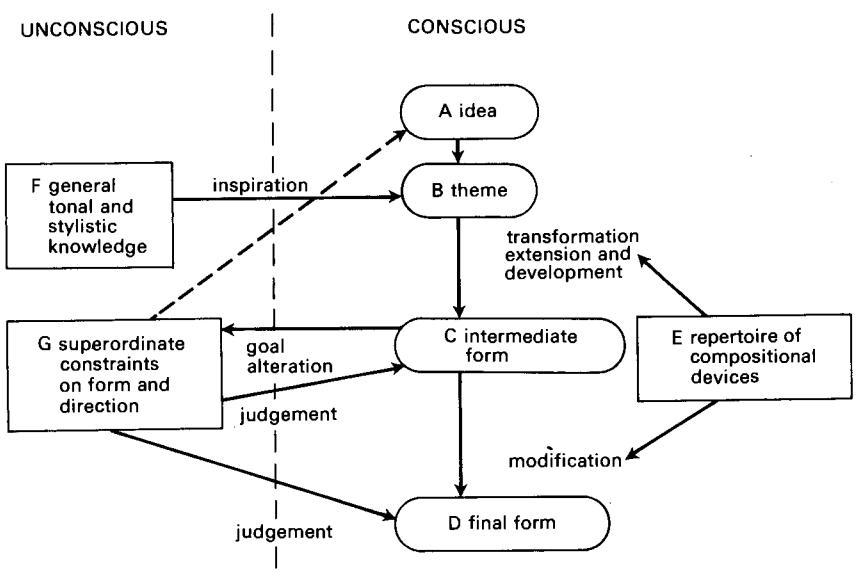



Fig. 4.1. Diagram of typical compositional resources and processes.

curved boxes contain the transitory materials that constitute successive versions of a composition as it grows in the composer's mind. The square boxes represent items of long-term knowledge which a composer has built up over the years, and which can be applied to new compositional problems. The lines joining boxes represent processes which transform or use the contents of the various boxes.

Box B represents the thematic kernel that springs 'unbidden' to mind out of the storehouse of thematic knowledge (F). Box A is optionally present in view of the comments (e.g. Sessions 1941) that sometimes a more or less specific idea of the *kind* of music required precedes an actual theme in awareness.

Box C represents the results of applying compositional techniques of transformation and modification (E) to the original theme. Its contents are then judged against criteria of 'rightness' (G) and, if found wanting, are modified until a satisfactory final form (D) is reached. The pathway 'goal alteration' acknowledges the fact that discovered properties of intermediate themes can actually overwrite originally held goals, so that the composition can appear to the composer to generate its own momentum or 'life', almost independently of his will.

It should be emphasized that Fig. 4.1 is not a 'theory' or an 'explanation' of the compositional process, but simply an economical way of describing some of the elements present in composers' accounts of their activities which makes clear the possible relationships between them. A rigorous psychological theory of composition would need to specify the contents of the various boxes and the nature of the transformations operating upon them in some detail.

The most rigorous test of the efficacy of theories in modern cognitive science is the production of a working computer program whose external behaviour mimics that to be explained. No such program has, to date, been written. The few documented programs (e.g. Sundberg and Lindblom 1976) do not attempt to model actual compositional sequences of the type portrayed in Fig. 4.1. In particular, there is no interplay between the generative process and processes of verification. The computer 'gets it right' at the first attempt.

It is also worth noting that in an effective theory of composition, the conscious/unconscious distinction is irrelevant. Both types of process must be fully specified, and it is the internal logic of a process which is of importance. We may suppose that, for different composers, the line dividing conscious from unconscious processes may be drawn in different places. Indeed, the whole issue of individual differences is an important topic of study. It would be oversimplistic to assume that there is only one route to an acceptable composition. A full analysis of such differences is beyond the scope of this book, but a single example illustrates the kind of issue which arises here.

Tchaikovsky, in his letters (Newmarch 1906; reprinted in Vernon 1970), frankly identifies a particular point of failure for him—form:

> I cannot complain of poverty of imagination, or lack of inventive power; but, on the other hand, I have always suffered from my want of skills in the management of form. Only after strenuous labour have I at last succeeded in making the form of my compositions correspond, more or less, with their contents. Formerly I was careless and did not give sufficient attention to the critical overhauling of my sketches. Consequently my *seams* showed, and there was no organic union between my individual episodes. This was a very serious defect, and I only improved gradually as time went on; but the form of my works will never be *exemplary*, because although I can modify, I cannot radically alter the essential qualities of my musical temperament.

Translating this into the formalisms embodied in Fig. 4.1, we may say that the route F–B–C–G is strong, but that the G–C and G–A–B routes are not. It is more likely that properties of a theme control the eventual structure rather than a strong conception of the structure exerting control over the thematic process.

Tchaikovsky clearly sees his own compositional process as inadequate, and others such as Hindemith (1952) stress the point that there may be as much separating the master composer from the 'ordinary' composer as there is separating the composer from the novice. Of the themes and motives which form the substance of inspiration, Hindemith says: 'they are common to all people, professionals and laymen alike; but while in the layman's mind they die away unused in their earliest infancy . . . the creative musician knows how to catch them and subject them to further treatment.' This treatment is precisely the moulding, transforming process that we have identified with the B–C–D route in Fig. 4.1. However, Hindemith goes on to claim that 'if only the work involved in reaching this goal really counted, there would be many a genius. The petty composer could do the same, technically, as the real genius did, and he would almost be justified in feeling godlike . . . because he was able to turn his bubbling inner singing and ringing into music, which [the layman] could never do.'

Hindemith then asks what separates the genius from the workaday composer and identifies the mark of the master as 'vision', which concept he explains in the following way:

> A genuine creator . . . will . . . have the gift of seeing—illuminated in the mind's eye as if by a flash of lightning—a complete musical form (though its subsequent realization in a performance may take three hours or more); [and] he will have the energy, persistence, and skill to bring this envisioned form into existence, so that even after months of work not one of its details will be lost or fail to fit into his photographic picture. This does not mean that any F-sharp in the 612th measure of the final piece would have been determined in the very first flash of cognition . . . but in working out his material he will always have before his mental eye the entire picture. In writing melodies or harmonic progressions he does not have to select them arbitrarily, he merely has to fulfil what the conceived totality demands . . . The

man of average talent may have visions too; but instead of seeing them in the clarity of lightening, he perceives dark contours which he has not the divination to fill out appropriately . . . [He] may manage the few basic rules of construction with all their combinative possibilities pretty well, and yet the highest degree of subtlety, in which each technical item is in congruence with the respective part of the vision, again may be attained only by the genius. There are relatively few masterworks in which this ultimate congruence can be felt.

Hindemith sees the acquisition of compositional mastery as one in which the demands of box G in Fig. 4.1 gain greater control over the compositional process so that the conception (A) exercises greater constraints over the compositional process from the outset. The subservience of detailed working out to a controlling plan or 'scheme' has been proposed as a partial explanation of introspective remarks attributed to Mozart and other composers (Gardner 1980), although it is clear that by 'vision' Hindemith means considerably more than the ability to plan a composition according to some stock structural formula (such as minuet and trio, rondo, or sonata form). He is referring to some underlying thematic and harmonic unity which transcends particular forms.

In the light of Hindemith's comments it becomes easier to understand the achievements of someone such as Rosemary Brown. We may see that she has grasped some of the rules of construction used in the music of classical and romantic masters. The compositions she produces are, however, mainly of simple episodic form, and lack the organic mastery sometimes achieved by the composers she claims to transmit. She is a good, albeit, unconscious imitator of surface styles. But for the total insulation of her compositional processes from her awareness she would be as little worthy of note as the countless unmemorable imitators in every sphere of creative endeavour. There is no evidence of the 'vision' that her composers had in such abundance in their lives.

The general issue of the psychological status of verbal reports has been a central one in cognitive debate. It is obviously naive to suppose that verbal introspections can be taken as true descriptions of mental processes in every respect. Nisbet and Wilson (1977) draw on a wide body of evidence in support of the view that, in many cases, people attempting to report on their cognitive processes do so not on the basis of true instrospection, but on the basis of *a priori* theories about plausible causes and effects. They reconstruct, in a more-or-less rational fashion, what they infer *should have* gone on to produce the results of cognition. Nisbet and Wilson's evaluation of the literature on creative processes leads them to the conclusion that

creative workers describe themselves almost universally as bystanders, differing from other observers only in that they are the first to witness the fruits of a process that is almost completely hidden from view. The reports of these workers are characterized by an insistence that (*a*) the influential stimuli are usually completely obscure—the individual has no idea what factors prompted the solution; and (*b*) even the fact that

a process is taking place is sometimes unknown to the individual prior to the point that a solution appears in consciousness.

Whilst it it true that *some* compositional processes seem obscure to consciousness, the material reviewed here suggests that Nisbet and Wilson's conclusions may be too sweeping. There are many aspects of the process which are open to conscious inspection and control. Furthermore, the insistence by many composers that there are processes that they cannot report upon encourages us that they are *not* 'telling more than they know'; that they are really attempting to describe the phenomenal experience *without* inference.

Ericsson and Simon (1980) argue that Nisbet and Wilson are overly pessimistic about the degree to which we can treat verbal reports as reliable indicators of cognitive processing. They point out that reporting verbally on introspective data characteristically means *remembering* a past sequence of events. The distinction between what can be reported and what cannot is not necessarily one between conscious and unconscious processes, but may be one between what can be readily recalled and what cannot.

They cite evidence which suggests two conditions under which material that has been the object of focal attention may not be recalled: cases where a solution is reached in a single step without intermediate or partial solutions; and cases where a process or a product is transient, being held in focal attention for only a brief moment before being superseded by new material. They argue that these factors are 'adequate to account for the phenomena of sudden insight that are the subject of so many anecdotes in the literature of creativity.' The studies cited in support of sudden insight are based on 'retrospective accounts of purportedly real creative acts, often reported many years after the event'. Ericsson and Simon's view is supported by evidence from studies where subjects were asked to introspect *whilst* solving the problem (e.g. Bulbrook 1932; Durkin 1937; Henry 1934). They argue that studies such as these show that the progress to solution was either gradual or determined by trial and error. No additional unconscious processes were required to be postulated.

A related cause of memory failure is concerned with the fact that the thought processes preceding 'inspiration' are often diffuse, unplanned, and undirected, and may be interspersed with other cognitive tasks which form the prime focus of attention. Although such unstructured thought is difficult to recall, there is evidence (e.g. Klinger 1971) that it is not unconscious. If people are asked to verbalize their processes *at the time* they are able to do so; but memory for the events fades very quickly.

No-one would wish to deny the immense complexity of the processes involved in musical composition. The arguments of Ericsson and Simon are simply a reminder that we may be making life too difficult for ourselves if we succumb fully to the prejudice of the contemporary psychologist against

using verbal reports as data. With a clear understanding of the conditions under which they tend to be unreliable, we can proceed to a relatively free use of this vast storehouse of psychologically relevant data. The thrust of Ericsson and Simon's argument is that we will gain the most detailed and accurate picture by examining verbalizations made *concurrent with* the act. It is to such data that we now turn.

## 4.4. Observations of composers at work

This section is based on two observations. One is reported by Reitman (1965) and involved obtaining a verbal protocol from an unnamed composer who was set the problem of writing a piano fugue (Sanchez and Reitman 1960). This yields a number of insights, but has inadequacies too. The complete protocol is unpublished, and there is no record of the concurrent musical protocol, the notes tried out at the piano and written down. Accordingly, the second observation is one made on myself as I composed a section of a work for choir and organ. As I worked on this section I wrote down as many of the steps of my thought as I could capture, alongside the relevant musical notation.

### 4.4.1. *Reitman's protocol of fugue composition*

To specify a fugue is to specify certain features which the final composition must possess. A strict fugue contains two or more independent contrapuntal voices which enter, one after another, with the statement of a single theme. The entry of the second voice is accompanied by new, contrasting material in the first voice which constitutes a 'counter-subject'. When each voice has entered, the music develops through the segmentation and recombination of the thematic elements in various ways; towards the end of the fugue one normally finds some dramatic intensification of the texture, through harmonic and dynamic climax, through the overlapping of thematic entries in the various voices (stretto), and other conventional devices. The prime constraints on a fugue are that each voice should possess some continuity and individual coherence, and that there should be a minimum of free writing where the voice material is not derived from the theme and counter-subject material (see Bullivant 1971).

A composer's problem, when working in any medium, is to set additional constraints over and above those loosely specified by the compositional type. There are an infinite number of possible fugue themes and possible fugues. The composer clearly cannot search through all possible themes in an unprincipled way. He must, in technical terms, 'reduce his problem space'. We find, in Reitman's protocol, evidence of external constraints operating, such as the need to find something which fits under a pianist's hands. When attempting to compose the countersubject the composer says: (erasing) 'I

think we'll abandon the idea of the enlarged syncopation to get away from repetition of the theme . . . and try to do something more pianistic.'

Probably more important than such constraints, however, are those *internal* constraints generated by what has already been composed, following some general principle of consistency or balance. Throughout the protocol, observations such as 'I'm playing through measure six to try and get my ear attuned to that kind of material' are common. It seems that the composer is attempting to discover initially unnoticed properties of what he has already composed to help him generate ideas for continuing. He is indulging in what Reitman calls 'after-the-fact adoption of a convention which is consistent with what has already been done'.

There comes a point in some compositions where the constraints, external or self-imposed are so many that the element of search or choice seems to be reduced or removed altogether. Towards the end of his protocol the composer calls out '*Voilà!* From here the ending is quite obvious, and there's no problem. This is the solution . . . The structure of the piece is solved now . . . I don't think it can miss.' It is a universal phenomenon of expert mental life that points occur in a problem-solving process where one 'turns onto a familiar road' traversed so many times before that the novelty and challenge of the particular problem disappears. Such points depend, of course, on the expert knowledge of the problem solver, and may be undetectable when reached by a novice working on the same problem.

I am a moderately poor Bridge player, and, sitting at the tables of the mighty, have often been astonished when, three or four tricks into a hand, the declarer lays down his hand with the claim that 'the rest is obvious'. This is followed by some explanation of his intentions of play, to which the others at the table immediately acquiesce; and I am left trying confusedly to follow what seems so obvious to the others. My self-esteem is only somewhat restored by the observation that sometimes such a declarer is wrong. A keen opponent says, 'I'm sorry, but do you mind actually playing that out?' and the unlucky declarer discovers some unforeseen twist of card distribution or tactic which undoes his strategy. The same is true of music composition, and Reitman's composer had the insight to add 'of course, you never can tell until you get to the end; it's always dangerous . . .' Some unforeseen conflict between the various constraints adopted may prevent any satisfactory ending from being found.

When multiple constraints build up over the course of a composition, it can become impossible to find ways of progressing which honour all the constraints. In such situations the protocol reveals the composer temporarily dropping some constraints, producing continuations which honour only some of the features defined earlier. He then readjusts in one of two ways: he either modifies the new material so as to bring it back into line with earlier constraints; or he may change the constraints. Thus, at one point the composer ignores all attributes of a pair of melodies except the arrangements

of harmonic intervals they form, ignoring the rhythmic aspect to generate a continuation.

What we are unable to do from the brief set of selected fragments presented by Reitman is to get any idea of the detailed *musical* working out of the general principles of problem solving which emerge with clarity. There is no opportunity to experience the music that the words are about. Accordingly, the next protocol gives some of the crucial musical material in addition to the words.

### 4.4.2. *A self-generated protocol of choral composition*

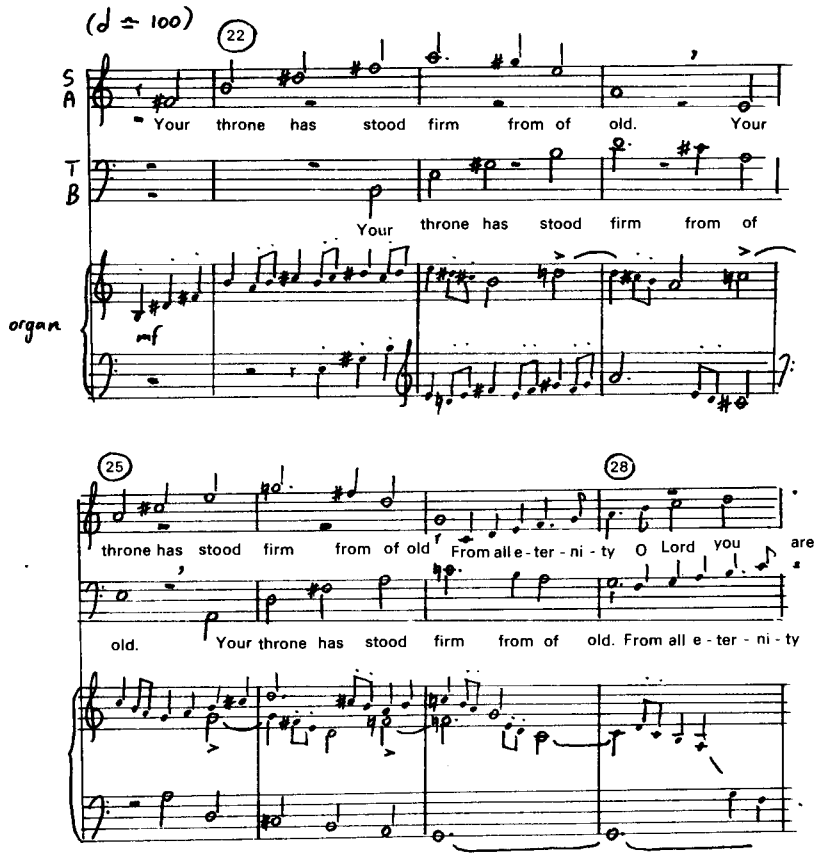
A few autobiographical details are necessary to set the context. I conduct a small amateur choir, and some years ago I decided to try and write something for the choir, taking up again a compositional interest that had absorbed me in early adulthood. My compositional style might best be described as 'twentieth-century English church' as exemplified by such composers as Herbert Howells. The style is traditional, placing great emphasis on melody, counterpoint, and a strongly tonal harmonic framework. I came to this work with a cluster of general constraints: the work must be learnable by an amateur choir; it should be for four voice parts plus organ accompaniment; it should conform to my general style—I was not deliberately intending to transform or extend my repertoire of compositional devices.

I chose to write a setting of Psalm 92, in English translation. The psalm has five verses, and the general mood is vigorous, joyful, and triumphant ('The Lord is king, with majesty enrobed'). The demands of contrast suggested to me that the third verse could take a more serene setting, and so a vague three-section notion was born; an initial vigorous setting of verses 1 and 2; a restful slow section for verse 3; returning to a triumphant close for the last two verses.

My initial compositional effort got me to the beginning of the last line of verse 2, and then I became blocked. My initial idea for the words 'From all eternity, O Lord, you are' seemed to lead nowhere, and I abandoned the project for at least a year.

It was some months after this abandonment that I first developed an explicit interest in the process of composition. When I returned to the task of completion in the summer of 1981 it occurred to me that it would be of interest to capture some of my compositional steps in a protocol. At this stage I was relatively naïve with respect to the literature reviewed in this chapter, and I had few preconceptions about the kind of thing such a protocol would show.

The protocol begins at bar 27 of the composition, where the abandoned theme was sketched in. Example 4.5 shows the score from bar 21 through 29 as it stood when I began. Since the protocol is fairly short, it can be given in full, with commentary and musical examples. The protocol is in italic type; comments and explications follow each segment of protocol in normal type.

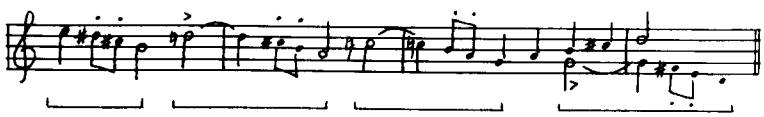



Example 4.5.

I should add that the original written protocol is highly cryptic, using personalized abbreviations and deviations from grammatical usage. The version presented here is as literal a translation into grammatical English as possible.

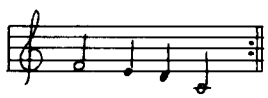
A. *How am I to continue at bar 29? The next verse is going to be quieter and slower. A theme is needed for 'From all eternity'. The tentative one where composition broke down before didn't seem right, partly because it was over too quickly.* This initial passage can be seen as a statement of the problem to be solved, and the setting of two immediate constraints. There was a general need to 'wind down' from a rather extrovert and 'busy' texture into something more easeful for the next verse; and a particular need to find a theme for the last line of verse 2. Lengthening the vocal phrases, which up till now had been short and separated, could be one way of introducing the required 'easement'. These constraints, however, were not yet specific enough to suggest actual notes.

B. *I decide to progress the accompaniment so that it slows down. It occurs to me that the word 'eternity' can be represented by a repetitive circular motif. We already have a circular four-note fragment in bars 23–27 (Example 4.6). So I try to use it to slow down (Example 4.7). The writing of bars 27–33 then becomes routine. A B flat is added at bar 33 on the intuition that harmonic movement is needed by now.* The vocal theme for 'Your throne has stood firm' is a rather self-contained 'closed' phrase. It did not have implications for any particular continuation to me. The accompaniment is, however, much more flowing and open ended at this point. To overcome the 'block' in the vocal material I decided to develop the accompaniment in the hope that this would, in its turn, suggest what the voices might do. The word 'eternity' suggested the possibility of a repetitive cyclical motif, and the accompaniment happened to contain such a motif, introduced at bar 23 for reasons I have now forgotten. It consists of a falling four-note scale, which returns to a higher note to fall again and again (see Example 4.6).) I seem to have decided



Example 4.6.

that the passage from bar 27 to 33 should be harmonically static (over a dominant G pedal). The task, then, was simply to repeat the falling motif over and over, at the same pitch, but with a spacing out and slowing so that from bar 31 there is just one cycle per bar (Example 4.7). Once the problem

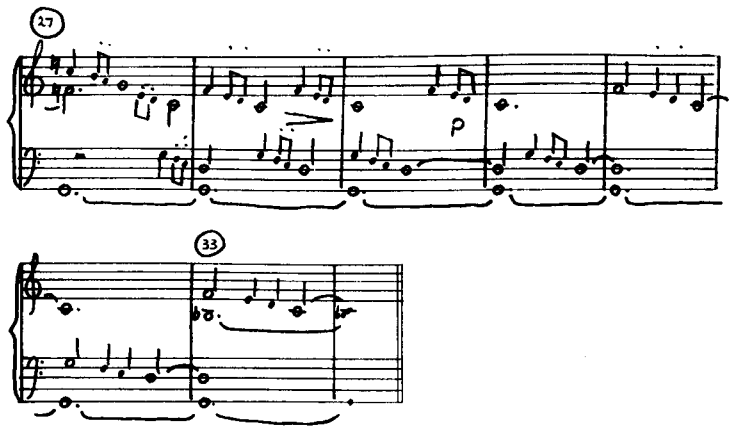


Example 4.7.

had been narrowed this far, I was able to compose fairly mechanically, thus the comment that this section is 'routine'. The full organ part for bars 27 to 33 is given in Example 4.8.

At bar 33 the problem was perceived as one of maintaining interest whilst continuing with the repetitive melodic material. The solution that occurred to me was to alter the tonal implication of the G pedal by adding a B flat in the harmony, suggesting a reinterpretation of the key as tending towards F major.

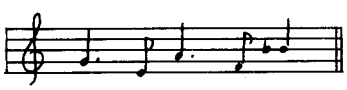
C. *Now I leave the accompaniment and return to the voices. The basic rhythmic pattern of the earlier, unsatisfactory theme (Example 4.9, bars 27–28) seems right, but it needs slowing down. The rise in pitch is right too, but it needs to be less extreme. If we incorporate alternate rising fourths and falling thirds (Example 4.10) that produces a coincidence. Starting where I feel it to*




Example 4.8.



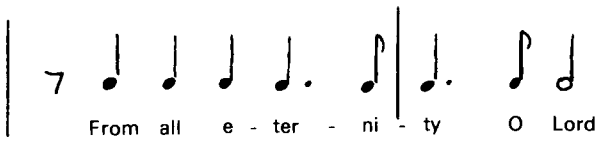

Example 4.9.



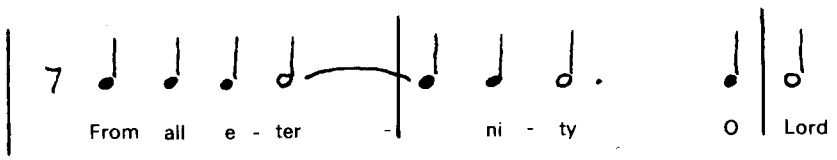
Example 4.10.

*be right (bar 31) it comes out on B flat at the same time as the accompaniment (Example 4.16, bar 33).* In this rather condensed section of protocol the accompaniment is now left. It provides a suitable bed over which new voice material can be added. The first, unrecorded step was to delete the abandoned voice entries in bars 27 and 28 (Example 4.5) so that by bar 29 only the organ was sounding. Clearly, I liked the basic rhythm of the abandoned theme, and its ascending pitch direction: but the speed of the rhythm and the rate of pitch rise were both too fast for the intended 'easing up' role that this theme had to fulfil. The first modification that occurred was to double the duration of the dactyl pairs (from Example 4.11 to Example 4.12). The second modification was to allow *alternate* notes to carry the pitch rise, with intermediate notes dropping back below (changing from Example 4.13 to Example 4.14). I cannot be sure why I chose the interval of a rising fourth at this point. It is possible that I was looking back to earlier thematic material to find some grounds for consistency. The very opening theme of the work, in fact, contains two germinal elements; ascent by rising fourths, and a falling scale of four notes (Example 4.15).

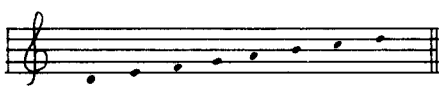
The 'coincidence' refers to the happy discovery that when I started the



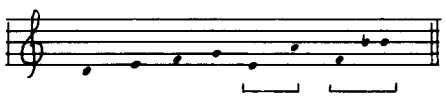

Example 4.11.



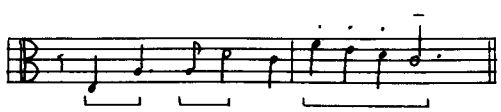

Example 4.12.



Example 4.13.



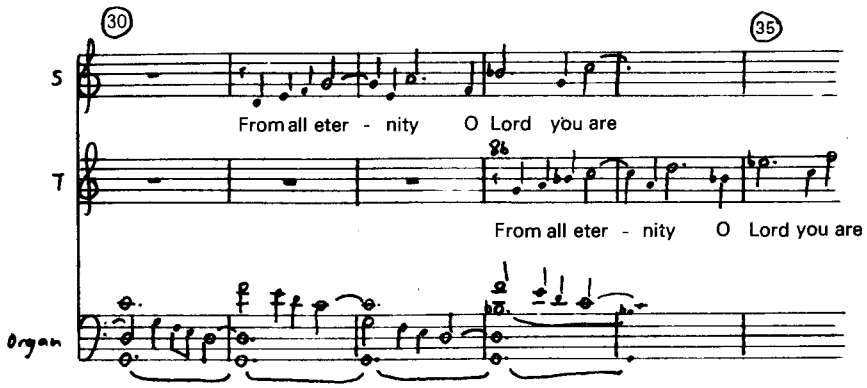
Example 4.14.



Example 4.15.

first vocal entry on D in the soprano line at bar 31, which, for reasons now lost to me, seemed the right time and note to start, the theme's already determined structure led it to B flat at precisely the point at which I had already decided to introduce a B flat in the accompaniment (see Example 4.16). This seemed to act as a confirmation that I was 'on the right track' at this point.

D. *The soprano theme finishes on C (bar 33), but I want that repetitive upward movement to continue. I will have to give the theme to another voice an octave lower. My first thought is to follow on, starting in bar 34. My second thought is to start in earlier (bar 33) and grow on out of the soprano ending. That works for the tenor voice.* The organ harmony at bar 33 is unresolved (see Example 4.8) and so this creates the feeling that the music is still moving towards a point of resolution. I experienced a strong sense that the upward movement started by the soprano voice should continue, but in a way that brought in another voice. The first thought was to start a lower voice in an imitation of the soprano in bar 34. This was discarded almost immediately in favour of an overlapping join in which the lower voice starts before the




Example 4.16.

soprano finishes, so that the upward cycle of ascending fourths and descending thirds is unbroken. Giving the tenor voice a C at bar 34 determines a start note of G in bar 33, if it is to be a strict imitation (a constraint which I did not explicitly formulate, but one that seems to be a stylistic feature of the whole work). Such imitation allowed the tenor theme to be copied out to the F at the end of bar 35. Example 4.16 shows bars 30 to 35 as they stood after this stage.

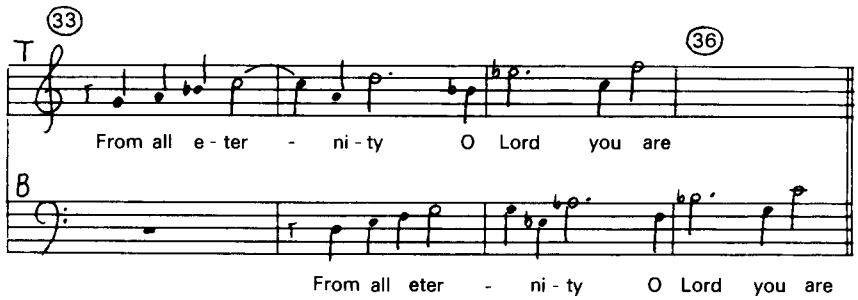
E. *It would be boring to do that again for a third vocal entry. Would an entry fit in the rhythmic spaces, to make a sort of 'stretto'? I experiment with different interval separations. A bass part entering a fourth lower than the tenor seems to work quite well, so fit that in up to bar 36* (*Example 4.18*). It seems clear that I was constraining myself by the relatively strict contrapuntal requirement that each of the four voices should enter with the theme in sequence. My problem was, thus, how and when to bring in the next voice? I decided that it would be uninteresting to generate entry three in relation to entry two by direct analogy with the relation of entry two to entry one. Instead, I conceived of an earlier entry which allows for rhythmic interplay (Example 4.17). Once this was decided, it remained only to choose a pitch




Example 4.17.

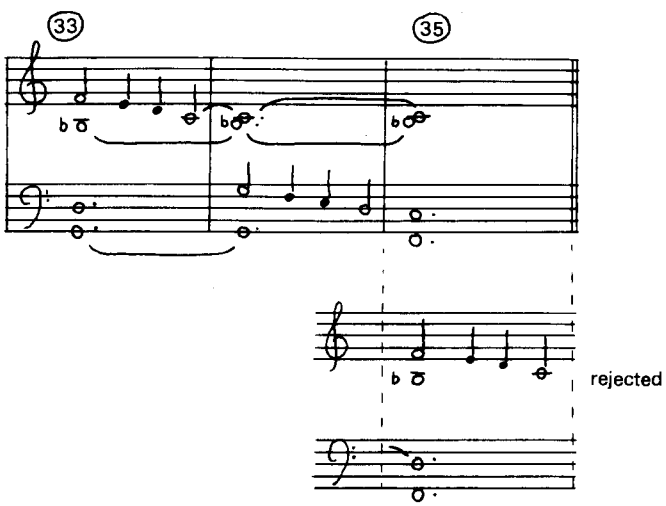
for entry three. The protocol suggests extended trial-and-error experimentation at the piano here (I always compose at the piano). Eventually I chose a bass entry on D as harmonically appealing, and wrote that out routinely to bar 36 (see Example 4.18).

F. *What happens to the accompaniment now? In experimentation I discover*




Example 4.18.

*that a drop to F in the bass fits with the voice parts quite well at bar 35. I carry through a tied B flat from bar 33 to 34, repeat bar 32's left hand descending scale in bar 34, and tie the middle C in bar 33 through to 34. By rule, the tenor D in 34 should carry through to 35 as well as the tied treble notes. Does this sound a bit messy? In bar 35 we certainly cannot repeat the FEDC figure, because the voice parts have shifted to an E flat key space. Perhaps this is the time to stop the pattern, and let the tenor D go down to C. I try it. Now we have a totally static accompaniment (Example 4.19, bar 35), so we must look to the voices to see how this might develop.*




Example 4.19.

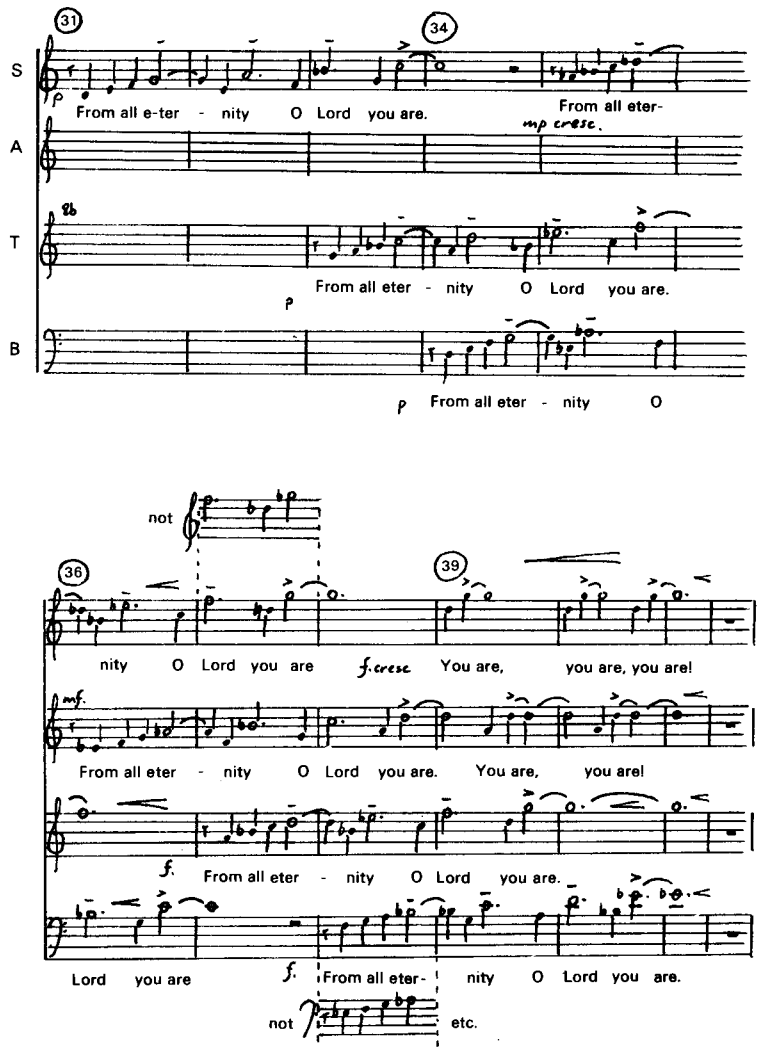
This section of protocol offers the most detailed comment so far, relating, as it does, to the composition of just over one bar of material for the organ. It was clearly a difficult and crucial moment for the direction of the music. The main problem was that there seemed no way of stopping the accompaniment just going on and on with the pattern established in bars 31

and 32. However, the voice parts were moving into remote harmonic regions, and so the accompaniment had to change somehow. The first decision was, then, to abandon the G pedal and drop to F at bar 35. The second decision was to abandon the repeated scale passage. The pattern established during bars 30 to 34 was that the last note of the scale passage should be tied over to the first beat of the next bar (see Example 4.8). If the left-hand passage at bar 34 is written according to this rule (repeating bars 32 to 33) this results in a tied D in bar 35. The chord produced by this (organ F, D, B flat, C; bass G, tenor E flat) was heard as unsatisfactory. It contained too many dissonances. By breaking the pattern and dropping the D to a C the chord was simplified and suggested as partial resting point. This suggestion was confirmed by the 'test' of putting a continuation of the right-hand scale pattern in bar 35 (F, E, D, C, as in bar 33). This created an unpleasant dissonance between the E flat in the vocal line and the organ E which was instantly rejected as a possibility. So, I was left with a simple unmoving chord at bar 35 (see Example 4.19) with no obvious implications for continuation. I returned to the voice parts to find the key to further development.

G. *I must bring the soprano voice in with the theme. I imitate the bass entry in relation to the tenor, so start on A flat in bar 35 (Example 4.20). If I follow strictly, this gets to D flat and G flat at bar 37. This is too harmonically remote, so I cheat and give D and G. Now that's four entries, but I have a feeling that this 'stretto' must be maintained for a while without working higher in pitch. How about taking the organ off at bar 35 or 36, then bringing it crashing in with some sort of perfect cadence later? Key? Who knows?* Here we find the completion (in principle) of the choral entries. I suspect that the fourth entry was given to the soprano rather than the alto because of its high tessitura. The soprano voice enters at bar 35 with the same pitch and time relations to the bass voice (at a suitable octave) as the bass has to the tenor. Once again, this pattern of successive entry can be repeated indefinitely. However, unlike the accompaniment at bars 30 to 35, there is a continual upward motion of pitch, and a movement round the circle of fifths in key, one step for each new entry. I seem to have wanted the continuing contrapuntal texture, but without further harmonic movement, so I was led to break my self-imposed constraint that the entries should be exact imitations. In bar 37, the soprano theme was shifted down a semitone from what it 'should be' so as to break the continuing movement round the circle of fifths (see Example 4.20).

At this point, an idea of the form of the next few bars presented itself in vague form. The organ stops, leaving the voices alone, intensifying the 'stretto' up to some climactic cadence, which is marked by the re-entry of the organ, fortissimo. Now there was a goal to control the next few bars, but, as yet, no idea of the final tonality.

H. *I notice that the vocal theme is four bars long, so it should be possible to bring it in each bar in a new voice* ad infinitum. *I bring the alto in at bar 36 on an E flat (same relationship to preceding voice as before). Then the tenor comes*



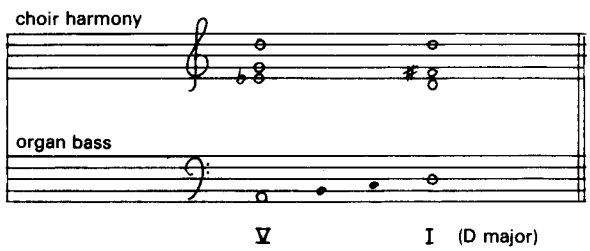

Example 4.20.

*in on A in bar 37, the bass on E flat (no—we get an unwanted A flat). I try entering on F—this seems to work. I write out the bass part to bar 40. How do we fill in the soprano and alto whilst the bass is finishing? Repeat 'you are' more insistently, off beat with other parts (bars 39 to 40).* It is not until this fairly late stage in working with the vocal material that I noticed the theme to be four bars long, deducing that it could, therefore, fit into an indefinitely long pattern where a new entry comes in once per bar in a different one of the four voices. This allowed the routine writing out of three more entries; alto at 36, tenor at 37, and bass at 38 (see Example 4.20). The first attempt

at the bass entry (following the strict pitch relationship to the previous entry already established) yielded an unsatisfactory harmonic result. The scale E flat, F, G, A flat moved too far away from the prevailing tonality, so, once again, I relaxed my constraints and allowed the theme to start on F, remaining in consonance with the other parts.

For some reason, I decided that there should be no further full entry after the bass, so the problem became what to do with the upper voices during bars 39 and 40. Anticipating an imminent climax, I hit upon the idea of taking the last two notes of the theme (ascending fourth) and repeating them with increasing insistence. This led to a chord at bar 41 which, in context, is harmonically unstable, and needs resolution (E flat, G, D). Even at this point, however, I was not clear where it should resolve.

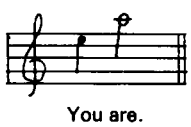
I. *Now the accompaniment comes in at bar 41. I try various bass notes. A sounds well, suggesting a perfect cadence to D major. It is too direct just to go from A to D. I try A, B, C, D ascending. That sounds good.* At this point I was 'fishing' for the right organ harmony by piano experimentation. A bass A was recognized as the right note for bar 41, and I soon found the cadence required (Example 4.21). An elaboration (small intermediate notes B and C)




Example 4.21.

soon suggested itself. I now interpret my need for these notes as a prolongation of the unresolved dissonance, screwing up the tension until the very last moment. At this point, only the bass line of the organ part was plotted out.

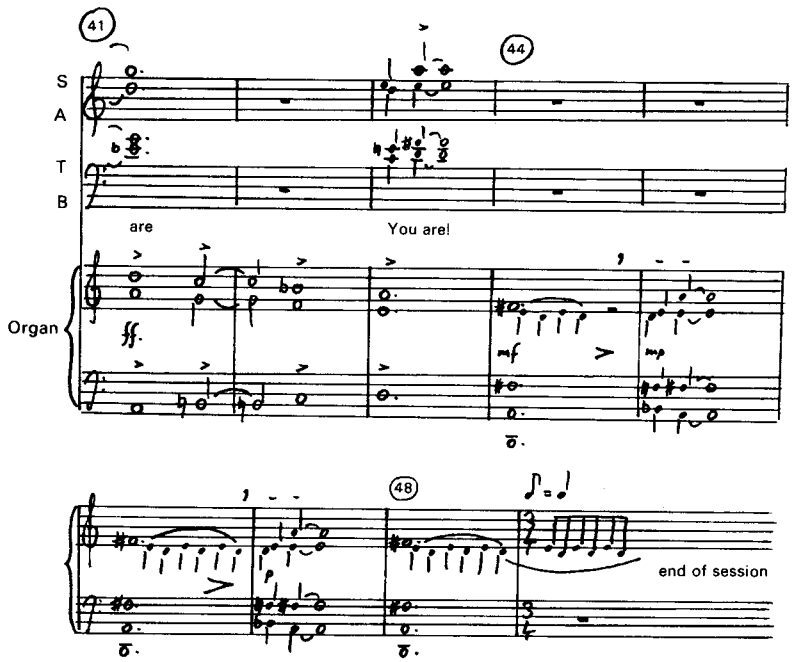
J. *The choir has to have something on the D major chord (silence sounds wrong). I tighten the screw once more in the soprano with E, A (Example 4.22).*




Example 4.22.

*I now find a choir harmony which fits, especially bearing in mind the E flat in the bass at bar 41 (Example 4.23).* I think what sounded wrong here was the unresolved choir chord at bar 41. It needed to be resolved by the choir as

well as by the organ. I thus returned to the immediate choral context to find something which could be used, and arrived at the idea of moving the ascending soprano figure of bar 40 up a whole tone, to bring the sopranos to top A (Example 4.23, bar 43), the highest note in the whole piece, and an




Example 4.23.

integral part of the D major chord planned for the organ. In this way, the chord of bar 43 becomes a climax, culmination, and summing up of the whole first section. I notice now, in retrospect, that the total composition has moved at this point from an opening G major to D major, the traditional tonic-dominant move of sonata form exposition; and that bar 43 contains the most emphatic statement of the ascending fourth motif which opens the piece. All this escaped conscious attention at the time, but, arguably, contributed towards my sense that this was a 'right' development. The filling in of the three outer choral parts at bar 43 was a piece of routine harmonization. Example 4.23 gives the final version of the choir parts in bars 41 to 43.

K. *The organ writing is a bit bare in bars 41 to 43. I add a right-hand part in descending fourths. Bar 44 is a straight D major chord, then the choir motif of bar 43 is repeated in a more luxuriant way (bar 45). I repeat bars 44 and 45 and this begins to form a flowing accompaniment out of which the next choral theme begins (Example 4.23).* Here the protocol ends.

The organ cadence, of such climactic importance, required more substance than a bare bass line. The notion of introducing a descending contrary motion line in the right hand must have come easily. As I now write, I notice that the melodic progression D, C, B, A is a slow version of the descending four-note scale which dominated the earlier accompaniment part (bars 22 to 33). The juxtaposition of this motif with the rising bass scale (derived from the opening of the vocal theme of the section at bar 31) makes a kind of grand, final statement of the thematic core of the whole section. In retrospect, the whole compositional process described here can be seen as a search for this crucial cadence. Through the fog of my compositional mediocrity loomed the dim shape of the 'inevitable' Schenkerian I–V movement, begging to be realized, but, at the outset, totally inarticulated. Climactic pivotal moment over, it then became easy to set the scene for my slower middle section. Bars 44 to 50 came quickly, allowing the organ to settle into a gentle rocking accompaniment for what was to come.

Something happened next which is of considerable importance in coming to understand the strengths and weaknesses of a protocol like this. Shortly after I arrived at bar 50, I began on the next section of the composition; but became immediately aware of a difference in the quality of the compositional experience. It was as if I had more urgent sense of direction. There came a wish to capture an unfolding shape before I lost it. As a result, I found myself incapable of continuing with the protocol, and embarked on an intense period of composition which produced the next 35 bars in an uninterrupted period of work.

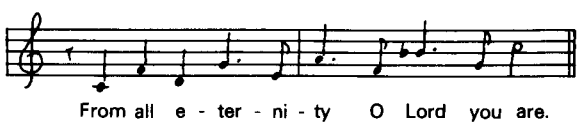
This is not to say that I experienced a pure 'flow of inspiration'. Rather, it was as if I feared to lose some momentum if I stopped to record my thoughts in words. In retrospective examination of my manuscript, I sense a lyrical, overarching quality about the next section, giving it a unity and repose quite unlike the kaleidoscopic first section. It is in D major, and moves, with few intermediate steps, along the well-trodden harmonic path I–V–I, with a single melodic idea unfolding in the voice parts. I would hesitate to say that this section is *better* than the section for which a protocol exists, but I am quite certain that the compositional strategy was different. There was not the same sense of indecision and search. I was lucky enough to start with a set of constraints which, within the bounds of my compositional style, almost allowed the section to 'compose itself'. It seems probable to me that verbal protocols reveal most of the genuine compositional process precisely when the composer is rather unsure where he is going, and casts around among his consciously held stock of transforming and extending devices to come up with a possible solution. When things go smoothly, the requirement to provide a protocol may well alter, even disrupt entirely, the process one is trying to describe.

Tchaikovsky has expressed well the concentration that is often required when creative activity seems to flow particularly well:

I forget everything and behave like a madman. Everything within me starts pulsing and quivering; hardly have I begun the sketch ere one thought follows another. In the midst of this magic process it frequently happens that some external interruption wakes me from my somnambulistic state: a ring at the bell, the entrance of my servant . . . Dreadful are such interruptions. Sometimes they break the thread of inspiration for a considerable time, so that I have to seek it again, often in vain

(From Newmarch 1906; reprinted in Vernon 1970.)

What general psychological conclusions can be drawn from this protocol? The first is that the protocol and its subsequent explication leave much unsaid, detailed as they are. They do not come near to documenting the rationale for the choice of each and every note. For instance, there are many possible melodic patterns consistent with the constraints I identify for the vocal theme starting at bar 31. Why did I choose the one I did? In retrospect, I can formulate reasons why the theme I chose is better than an alternative such as Example 4.24. The chosen theme (Example 4.15) has more melodic




Example 4.24.

interest (this latter alternative is too repetitive), and it has more harmonic movement (the alternative spans an octave, and, as such, is harmonically rather static). I doubt if I would have consciously entertained this alternative at all. It is as if I was able to choose the *kind* of theme which had musical potential almost automatically. This is reminiscent of observations on other forms of skilled choice (e.g. chess playing; Simon 1979). In choosing moves, skilled players seem to consider fewer unpromising moves than do beginners. My own protocol is littered with unexplained choices. Something is *felt* to be right, or chosen without any justification at all. I take this to be a general characteristic of skilled problem-solving.

A second conclusion is that, at least in my case, the interplay between the discovered properties of previously composed material and some broader conception is paramount. All the way through the protocol I find myself deliberately using some feature of melody or harmony in previously written material as the starting point for a continuation. On the other hand, there are many cases in which superordinate considerations of harmony and structure prescribe changes which modify or overwrite the essentially imitative strategies which generate the first continuations. Just like Reitman's composer, I find myself engaged in trying to characterize what I have already got, so that continuations will be, in some sense, fitting or 'in style'; but I also have shifting, and often dimly perceived notions about the direction of a section which allow me to discern when strict imitation must stop.

Third, because characteristics of previously written material are often dis-

covered some time *after* the writing, but can still be important in shaping future events, I suspect that an important component of compositional skill is a degree of 'trust' in one's medium—a certainty that the habitual processes of generation will yield material which is richer than one first sees, and which, even if initially unsatisfactory, usually contains within it discoverable properties which can be used to profit. This trust is partly engendered by the sheer fact of previously solved problems; but it also has something to do with increasing awareness of the richness of a medium such as the tonal system. There are so many different kinds of relationship which hold between a set of notes that even the simplest theme can become an almost inexhaustible source of discovery. The consequent delight that any creative worker experiences in the sheer fecundity of his medium is, arguably, not simply a pleasant by-product of creative activity, but the essential source of motivation which keeps him involved with his own material through 'blocks' and blind alleys in the creative process. Such delight is amply documented by the practitioners of all sorts of skill; from mathematicians to poets (for example, Poincaré 1924, and Spender 1946; both reprinted in Ghiselin 1952, and Vernon 1970).

In discussing my protocol with fellow psychologists and musicians, the question has often arisen of how one might use such protocols to generate fruitful scientific research into composition. Protocol gathering is a species of natural history, and although it is an essential first step in understanding any complex mental process, it does not, on its own, allow detailed theory building and testing. For this, two lines of attack seem possible. One would be to focus on a simple compositional task, such as the harmonization of a short melody, attempting to specify *all* the cognitive steps which underlie the choice of particular chords by a particular composer. This would naturally lead to some form of computer simulation to test the adequacy of one's theory. Another related possibility would be to take some circumscribed compositional task, maybe the same one as before, and ask a range of musicians of varying ability and experience to carry out the task. By a comparative analysis of the behaviour of the different subjects it might be possible to specify some of the ways in which compositional skills change with experience.

### 4.5. Improvisation

The keynote of the compositional process seems to be the moulding and perfecting of musical ideas. Although an idea may come spontaneously, unbidden, and instantaneously, its subsequent development may take years. In improvisation, the composer has no opportunity to mould and perfect his material. His first idea *must* work. Is, then, the improviser a different sort of animal to the composer? Is he the musician who, like Rosemary Brown, seems to exert no conscious creative effort in the production of coherent

music? The record tells us that this cannot be so; we know that composers such as Beethoven were also master improvisers. The same man that could chisel away at a theme for 20 years could also produce live compositions of extreme accomplishment.

I wish to argue in this section that what distinguishes improvisation from composition is primarily the pre-existence of a large set of formal constraints which comprise a 'blueprint' or 'skeleton' for the improvisation. The improviser can, therefore, dispense with much of the composer's habitual decision making concerning structure and direction. He uses a model which is, in most cases, externally supplied by the culture, and which he embellishes and 'fills in' in various ways. Such models frequently have a recursive sectional form, so that the performer becomes very familiar with the structure of the basic section, which can, to a large extent, be considered independently from other sections. This means that the improviser does not have to be constantly referring back to the detailed working out of earlier sections as, it may be argued, a composer must. He can rely on the given constraints of the form together with his own 'style' to give the music unity.

If we look at one of the forms in which improvisation flourishes, we can see how it helps the improviser to construct his performance. 'Theme and Variations', for instance, is a common form in classical, folk, and jazz idioms. A typical structure is

| | | | | |
|---|---|---|---|---|
| phrase: | A A B A | $A_1$ $A_1$ $B_1$ $A_1$ | $A_2$ $A_2$ $B_2$ $A_2$ | etc. |
| section: | theme | first variation | second variation | etc. |

In this structure, the theme is composed of four phrases (typically each eight bars in length). The second and fourth phrases are repeats of the first phrase; the third phrase is different, but stylistically and thematically related to the first. The theme will be characterized by a certain underlying harmonic progression, and this progression will be repeated in each new variation. The improviser maintains interest and momentum by embellishing and transforming the melody and the texture within this repeating harmonic framework. Once the character of the A phrase variant has been determined, the remainder of the variation becomes highly constrained. An improviser can, for instance, 'run off' the last phrase of a variation as a rough copy of the first phrase, whilst giving some advance thought to the construction of the next section.

It will help us to see musical improvisation in perspective if we compare it to a verbal skill which has roughly the same sort of demands, such as the recounting of a story by a skilled story-teller. First, the story-teller has knowledge of a particular set of episodes which constitute the 'plot' of his story. This is similar to the knowledge of a particular melody or chord sequence in music.

Second, each element of the plot contains reference to well-known objects, situations, people, etc., whose detailed characteristics are not necessarily

specified in the plot, but which may be used to elaborate and keep the tale fresh on any particular occasion of telling. For example, the plot may specify a character such as a farmer, who, without damage to the plot, can be further specified as old, lonely, jolly, tall, etc. on particular occasions. Minsky (1977) has used the term 'frame' to describe the cognitive structure that is accessed when a word like 'farmer' is introduced. This frame specifies attributes appropriate to the object or situation that can be specified, together with details of limits and constraints on those specifications. For instance, in a folk tale, a farmer will be an adult human being, so it is possible, although not necessary, to specify an age between, say, 15 and 70. Tellers and listeners frequently operate with 'default' values for attributes which are not actually specified. Such values may be chosen to be consistent with earlier information, or may arise from 'stereotypes'. For instance, when a farmer is mentioned, many listeners will assume a middle-aged, rather rough-cut male; and so a storyteller may choose to leave some details unspecified, knowing that the listener will still be able to make sense of the story through shared frames and stereotypes.

In music, we may equate 'frames' with characteristic harmonic or melodic progressions that underlie many different types of music. For instance, cadences are harmonic frames; particular types of melodic movement such as scales, arpeggios, gap-fill or changing note patterns (Rosner and Meyer 1982) are melodic frames. As in stories, many of these frames can be filled quite simply, with a couple of notes or chords, but can be fleshed out with various types of embellishment; passing notes and passing chords, for instance. There are many options for the way in which a musical frame may be filled. Performers and listeners within a culture will know about stereotypic ways of filling in such details. The composer and improviser achieve aesthetic impact by finding novel or surprising ways of embellishing a frame without changing its basic character.

A third feature of stories is their tendency to conform to certain structural specifications. One common type of story is the 'problem–effort–resolution' structure. An initial situation is specified. This situation presents some problem for a principal character or group of characters. They attempt to solve the problem; and this often involves the resolution of intermediate sub-problems. Finally, the goal is reached. Such structures allow both listeners and teller to keep track of the complex internal details of a story by labelling them according to their roles in the familiar structure. When stories do not conform to familiar structural models they are very hard to understand and remember (Bower 1976). The teller can use such a shared structure to pace his storytelling. If he senses his audience losing grip of the plot, he can add explanatory sentences to refresh memory of the position he is at within the structure (e.g. 'now the prince needed to find just one more thing before he would be able to slay the wicked dragon'). Similarly, if the plot seems too simple or bare for his audience, he can go down an irrelevant digressory path, provided that

he is able to leave a 'marker' which tells where he is within the main structure.

We can roughly equate story structure with 'form' in music, as exemplified by 'theme and variation' or 'fugue', etc. When such structures are known by listeners, they are able to appreciate the function of present events, and through their anticipation of future constraints they can appreciate the art that produces novel sequences to fulfil the relevant function in the correct time. In jazz ensemble music, for instance, there is often a distinction between a fairly invariant 'chorus' played by the whole group, and a 'solo' where an individual performer is given the opportunity to come to the fore in an improvisatory fashion for a fixed number of bars before leading back into the next chorus. A listener who knows this structure can realize that the solo is about to end, and wonder how on earth the soloist is going to return satisfactorily from his harmonic and melodic excursions in a way that appropriately leads into the chorus. When such a performer does 'get home' in an appropriate but novel way, the listener can derive satisfaction which would be denied to someone who did not appreciate the demands of the structure. Similarly, the performer's own knowledge of the structure allows him to plan excursions whilst retaining markers which tell him exactly when he must arrive at a particular musical event.

In a sense, every act of story-telling that is not a verbatim recall is an act of improvisation. However, such improvisation reaches a peak in feats such as the singing of epic poetry in the Homeric tradition, preserved into the twentieth century among the folk singers of Yugoslavia (Lord 1960; Parry 1971). The singers are often illiterate, yet each singer has a large repertoire of poems (as many as 100), some of which are several thousand lines long and take more than an hour to sing. Parry's observations and recordings show that epic singers composed their songs anew on each occasion. No two performances were the same.

Neisser (1982) summarizes the main features of this epic poetry:

> (the songs) show repetition of units at two levels of analysis, known as *themes* and *formulas*. Themes are types of events that occur in many different songs; a council of war, the arming of a warrior, the return of the hero in disguise. The formula, a much smaller unit, is 'a group of words which is regularly employed under the same metrical conditions to express a given essential idea' (Lord 1960, p. 30). The singer may have a dozen formulaic ways to describe daybreak: 'when dawn put forth its wings', 'when it was dawn and white day', or 'when the sun had warmed the earth', for example. When he comes to a part of the action that takes place at dawn he uses the one that best fits the metre and his mood. He knows dozens of themes and thousands of formulas, not from having memorized them but from having heard and used them, as we know the words and phrases of our own language.

This poetry is of special interest to us because it is *sung* and so involves musical improvisation too. The composer and musicologist Béla Bartók had a particular interest in Eastern European folk music, and collaborated with Lord in transcribing and analysing some of the musical recordings made by

Parry (Bartók and Lord 1951). The epic poems were always sung by a man, accompanying himself on the gusle, a one-stringed variety of lute which could provide a few different single notes for a drone-like accompaniment. The sung music was a version of chant. As far as one can undertand the comments of Bartók and Parry, who did not transcribe many of the epic poems, these chants were short simple melodies which were repeated with minor variations in the successive lines of the poems. Bartók was more interested in the shorter 'women's songs' where the degree of verbal improvisation was much less, but where the music was of considerably more complexity. His analysis identified various 'principal tones' which constituted the basic melodic frame for successive verses. The painstaking transcriptions show that a singer would embellish each verse in a different way, and would not use the same embellishments on two performances of the same song. There was much use of interpolations; glides, grace notes, ornaments, and passing notes.

In general, although vocal music is no stranger to improvisatory treatment, purely instrumental music seems to have encouraged more ambitious musical improvisation through the ages. One reason for this is that words constrain vocal music in many respects. A sung melody can be embellished only at the expense of directness of word communication. The 'basic' folk melody usually has only one note for each syllable. If more notes are added, then words must be repeated, or new syllables be created, or existing syllables lengthened and sung to several notes. The second of these devices is apparent in the Bartók transcriptions, where nonsense syllables (e.g. 'Aj' or 'Ahaj') are interpolated between words. Similarly, English folk songs are often embellished with 'fa la' or 'fol de rol', etc. The third device is known as melisma, and is common in many types of vocal music. Excessive use of melisma does, however, destroy the natural metre of the words, and can lead to severe difficulties in understanding them. In Gregorian chant, for instance, it can sometimes take five minutes to sing one sentence; a distortion which can only be tolerated if the words are well known to the listeners.

Although improvisation has flourished in various cultural and historical settings, I shall base the remainder of this section around jazz improvisation. Jazz is a contemporary Western form in which improvisation is paramount. It is a convenient form to study because (*a*) the idiom of most of its styles is derived from popular Western music (in its widest sense), and so is straightforwardly tonal; and (*b*) being contemporary, it can be experienced by any listener with access to a radio or a record player. We can only guess, from imprecise reports of observers, exactly how an improvisation by Mozart or Beethoven sounded. (For some observations on contemporary non-Western improvisatory forms, see Chapter 7).

There are two primary jazz forms; the 12-bar blues and the 32-bar song. The melody is often a 'standard' which is played in a fairly straightforward way to begin with. Repeated improvisations retain the basic harmonic sequence (although with the possibility of varying the harmonic detail); but

the *essence* of improvisation, both here and in the 12-bar blues, is in the melodic line. This is partly due to the fact that the prototypical and original jazz medium is the brass ensemble where, because each player can produce only one note at a time, solo performance is necessarily melodic. Jazz piano came later, but it copies features of brass ensemble playing. The left hand typically provides chords and rhythm; the right hand 'plays the solo', carrying a single melodic line which is the focus of improvisatory effort.

All this, though, does not uniquely specify *jazz*. It would be quite possible to produce something perfectly 'Mozartean' using the 12-bar blues frame. What gives a performance its jazzy character are the particular types of melodic, rhythmic, and harmonic devices which form the basic building blocks for larger sequences. Such devices are almost exactly analogous to the *formulas* of epic poetry. They constitute the vocabulary of jazz. However, just as Yugoslav singers do not memorize a dictionary of formulas, so jazz improvisation is not normally learned by arming oneself with a list of characteristic chords and progressions. Jazz is learned through listening intently to other musicians and performing with them. Books and notation are not necessary; many great jazz performers could not read a note of music.

The pianist David Sudnow has performed a valuable service by writing a book which records his progress at learning jazz improvisation (Sudnow 1978). He provides a detailed exposition of his experiences at various stages of his development. In pulling out the principal features of his account, we can reinforce and expand the characterization of improvisation already given in outline.

When Sudnow approached the task of learning jazz he was already a competent classical pianist. The first new thing he learned was the 'anatomy' of typical left-hand jazz chords—practising placing single chords in the various keys of the chromatic scale. This was so that his hand would come to take up the appointed finger configurations rapidly and accurately without hitting adjacent wrong notes.

Second came the task of combining chords into the sequences typically found in songs. This entailed the co-ordinated ballistic movement of arm, hand, and fingers, to negotiate the sequence in a smooth but strict rhythm. At first, much visual monitoring was required, but eventually, common song sequences could be negotiated independently of visual control.

Next, standard (non-improvised) song melodies were combined with their appropriate chord sequences, so that the first statement of the 32-bar song could be given. All this was preliminary to the main task, that of melodic improvisation over a standard left-hand chord sequence.

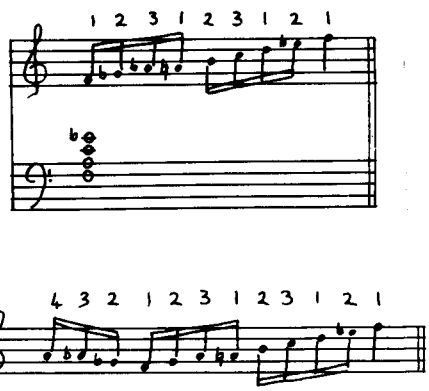
At the outset, Sudnow did not know how to start nor once started, where to go next. The hand

> might as well have gone anywhere, but once it did, there was nothing for it to do, and I discovered from the outset that if you don't know where you are going, you can't go anywhere except incorrectly. The hand has to be motivated to particular

next keys to depress, and when there was nowhere for it to go it became totally immobilised, stumbled around . . . (p. 18).

He had no *formulas* to put into play. This was despite much experience of listening to jazz improvisations. He found the improvised passagework of his favourite performers so fast and complex that he simply could not get a detailed, note-by-note grasp of the structure. He reports that 'a three second stretch of play, within the course of an ongoing jazz improvisation (on record) I had listened to for many years, demanded several hours to catch the details sufficiently . . .' (p. 16).

Sudnow's teacher dealt with this problem by giving him some specific formulas. He would say 'with this particular chord you can get a characteristic sound by playing this particular scale.' An example of a chord-scale combination that Sudnow learned in this way is given in Example 4.25. He learned enough of these scales to have one available for each of the twelve chromatic dominant-seventh chords, and became fluent at playing them rapidly, both ascending and descending. At first, he could begin these scales in only one or two places. He reports that 'over the first years of play, nearly every time I played this particular scale, I either started on the F with the thumb (Example 4.25) or on the fourth finger (A) then moving down to the F (Example 4.26). This rigidity in a scale that could, in principle, be started on




Example 4.25. and 4.26. Jazz formulas learned by Sudnow (1978).

any note, was attributed to lack of useable knowledge about what finger to use if starting on other than the familiar note. To start, for instance, on the B natural and move upwards would require a second finger. But Sudnow knew the sale only as a single 'handful'. He knew a start note and the rest 'just happened'. So, in his early improvisation, he would simply find the appropriate scale for a particular chord, varying only the octave in which it was started or the number of notes used.

Gradually a repertoire of characteristic jazz formulas was built up in this

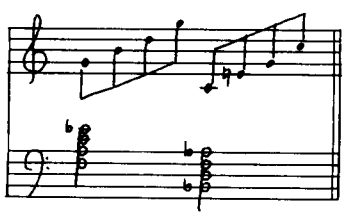
way: 'arpeggios to be taken and scales to be linearly played, various "licks" constituted by certain intervallic relationships.' It was at this point that Sudnow tried his chances with a jazz trio at a club. His graphic description makes it clear just how far from true improvisation he yet was:

> . . . and now a run for chord A is played, starting near the middle of the keyboard, and rises up, while the pathway I know best for chord B is played starting at the middle also. So that I started going up with a fast, spluttering, and nervous scale course, and the next chord came up and I had to shoot back down to the middle of the keyboard, to get the thing I knew how to do well done for it, and then there was the next chord. . . . An upward-moving line would more or less end when the next chord's sustaining had to be terminated, no matter where it was. Or, in order to get to the next starting place, I would end it a bit sooner, to give myself time to relocate, feeling the upcoming chord as an encroaching presence whose necessity was fixed by adherence to the chord chart of the song . . . so that what the left hand was doing in its pre-set ways was guiding what the right hand was then obliged to do. (p. 37.)

Sudnow continued for a few more years on essentially the same course, adding more and more formulas to the repertoire. Then, one day in the third year of his play, he began to have a different experience with his playing, which he calls 'going for the sounds'. This experience was essentially one of selecting and anticipating *particular* notes in the melodic improvisation. He explains that, up to that point

> there was, with respect to the finely textured note-to-note nature of my play, an order that was guaranteed by the path's formal theoretical construction . . . (e.g. a scale) . . . I was not doing note-to-note selectional work. I decided where to start out each run, which run to choose, how fast to play it. But, we may say, no intentionality of aim was given to particular pitch achievments—only these broad 'parameters' received a motivated determination. (p. 39.)

Now, in contrast, Sudnow started to plan particular note effects. He did this primarily by attempting to relate a particular melodic line 'backwardly' to what had been done before, on the previous chord. So for instance, two adjacent chords might be given the realization shown in Example 4.27. Here,



Example 4.27. Imitatory improvisation. From Sudnow (1978).

the second group of four notes is a melodic imitation of the first group.

Previously, Sudnow had only the capacity to respond melodically to each chord as it came up by 'pulling something out of the bag' in respect of that

chord alone (Fig. 4.2, stage 1). Now he was able to choose a continuation which both fitted the chord and was similar to the previous melodic fragment (Fig. 4.2, stage 2). This new way of playing led at first to a loss of fluency.

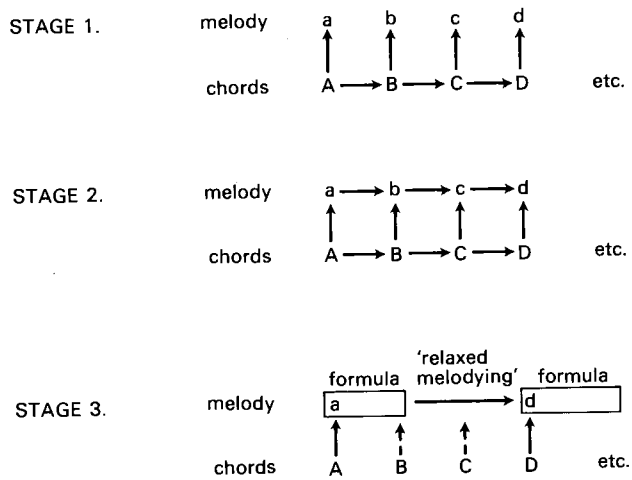


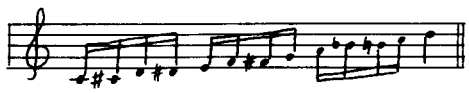

Fig. 4.2. Three stages in the development of David Sudnow's jazz improvisation.

Because the hand was attempting more ambitious things, it more often went wrong. For instance, to transpose a melodic pattern up or down a fixed number of semitones requires different fingering and hand configuration according to the particular black and white note pattern that the new key has. The melodic fragment in Example 4.27, if transposed to start on E flat, has a black-white-black-black key sequence, in contrast to the original all white sequence. This requires a shift in hand shape and, if uncertainly known, entails the possibility of hitting a wrong note. The well learned scale formulas were, however, played impeccably by this stage.

At this stage a typical experience is described as follows:

> I would play a figure, go for its repetition, get some way into it, and stumble. To fill what I felt to be the remaining empty space, keeping this jazz going, I would do something quite unrelated to the explicit continuities I had partially achieved. I would accomplish the beginning of a reiteration (transposition, inversion . . . exact duplication, etc.), and then, for example, use up a remaining allotted chord time by taking on any notes that were thereabouts to take. (p. 56.)

Gradually, Sudnow worked towards something more assured and principled to do when the imitation attempt failed. He calls this the 'chromatic way'—and it essentially involves a bunched hand which plays a close set of notes which, although not necessarily composed totally of semitones, would have a chromatic character (e.g. Example 4.28). Such devices 'were afforded potentials for keeping action going in a sector, for making streams of notes,



Example 4.28. Chromatic improvisation. From Sudnow (1978).

for doing a fair amount of unthoughtful, weakly geared aiming for sounds. They allowed sustaining, at the same time, a connective tissue of action within attempts to tie up courses taken with one chord to those taken with the next' (p. 60).

Still Sudnow knew he was not really playing jazz the way he wanted it. It was too 'lungeful', composed of sporadic 'bits' of motivated improvisation joined together by *ad hoc* connections. It was after listening to the jazz pianist Jimmy Rowles (and watching him) that a breakthrough came. Rowles 'sat rather low down and stretched back, almost lazy with the piano like a competent driver is nonchalant behind the wheel on an open road. Still, there was a taking care with the melody, a caressing of it, a giving each place its due. He was never in a hurry . . .' (p. 81).

Rowles' relaxed swaying way at the piano contrasted with Sudnow's own 'frantic' attempts to negotiate fast passagework. Rowles played more slowly, with more attention to the basic melody. In trying to imitate Rowles, Sudnow began to find that

> within my practice sessions little spates of that jazz on the records unpredictably showed themselves to me, and then disappeared. No sooner did I try to latch on to a piece of good-sounding jazz that would come out in the midst of my improvisations, than it would be undermined. . . . But there is no question that the hang of it had been glimpsed . . . All prior ways of being seem thoroughly lacking, and the new way is encountered with a 'this is it' feeling, almost like a revelation. . . . For a brief course of time while playing rapidly along, a line of melody would be generated, interweavingly flowing over the duration of several chords, fluently winding about in ways I had never seen my hands wind about before. (pp. 83–4.)

At first, Sudnow was unable to characterize exactly what he was doing in these moments, but gradually he came to realize that it was a sort of 're-leasing' from the rigid formulas which sent him scampering up and down the keyboard, on the basis of a new understanding that 'note choices could be made anywhere, that there was no need to lunge, that usable notes for any chord lay just at hand, that there was no need to find a path, image one up ahead to get ready for a blurting out . . . Good notes were everywhere at hand, right beneath the fingers.' He found

> for example, that where it before seemed required to reach for a big path for a chord, forseeing its location and organizational requirements . . . a single note would do to make melody over the course of several chords' durations. I could take my time in going for a long run, could linger, finding right beneath a non-venturing hand all sorts of melodying possibilities. (pp. 94–5.)

This allowed him to escape the tyranny of a rigid time schedule which had so hampered his earlier attempts. In the old way of playing, each chord was like a starting gun for some new pattern. So, the previous pattern had to get finished in time, maybe breaking off in an unplanned place. Sometimes he would fail to get to the new start in time, so the hands moved unintentionally out of phase, and the next figure would be hurried to get back on schedule. The new style allowed

> mobile ways with the terrain such that commitments to arrival times could be continuously altered and shifted about in the course of the negotiations, a steady beat to the song as a whole sustained all the while. . . . It was to permit nonstuttering and nontripping disengagements from the terrain when a saying was not at hand, disengagements that would not make the music stop but would be silences of the music. (pp. 103–4.)

So now, the relation between the chords and the melody was looser; chords would still suggest typical formulas, but maybe a formula could run over into the next chord, then leaving the hand in a position to do some of the new relaxed melodying, taking time to introduce the next formula (Fig. 4.2, stage 3).

Very often a 'wrong note' in an attempt to play a formula could be used to good purpose. Instead of stumbling to correct the mistake, the note could be exploited in a jazz-like way. We see, therefore, an increasing understanding of the potentialities of almost *any* note to be used to effect in jazz. This presumably comes about through the years of experience of exploring the sounds that a keyboard makes. The improviser is relaxed and unhurried because he knows that, wherever he lands up, there are a dozen different ways of getting from there to the next place.

This confidence in the availability of plenty of 'escape routes' is, surely, a hallmark of skill in any impromptu accomplishment. To take a non-musical analogy from personal experience; my profession entails the fairly frequent giving of lectures and seminars to other members of the profession. When I began doing this, I was particularly terrified of the questions session which followed my 'set piece'. I was afraid of many things: not knowing the answer to a question, having some devastating criticism pointed at my work, not being able to put together a fluent reply without preparation. These fears began to subside as (*a*) I realized that I had a natural advantage over my audience; that of knowing the details of my own material, through many months of involvement with it, far better than anyone in the audience, and (*b*) I developed a repertoire of ways of handling awkward or unpleasant questions without faltering or becoming confused. The most important thing in situations like this is not to provide the *best* answer, but to provide *an* answer of some respectability with the fluency and immediacy that signals competence. Now, I usually face a question session in a relaxed and confident frame of mind because I know that I will always be able to find *something* appropriate to say. The jazz improviser can enjoy his improvisation because

he knows that there will always be *something* he can play. Within this competence he, like all improvisers, knows that for much of the time he will produce quite ordinary improvisations; but from time to time, according to circumstance of mood and choice, something superlative will come up.

Here, then, is a fundamental difference between improvisation and composition. The composer rejects possible solutions until he finds one which seems to be the best for his purposes. The improviser must accept the first solution that comes to hand. In both cases the originator must have a repertoire of patterns and things to do with them that he can call up at will; but in the case of improvisation the crucial factor is the speed at which the stream of invention can be sustained, and the availability of things to do which do not overtax the available resources. In composition, fluency becomes less important; but it is much more important to keep long-term stuctural goals in sight, and to unify present material with what has gone before.

As we said at the beginning of this section, what absolves the improviser from the task of evaluation and long term planning is the relatively rigid formal 'frame' within which his improvisation takes place, and which dictates the large scale structure of his performances. Because this frame persists over many improvisations, the performer builds up a repertoire of 'things that have worked well in the past'. Commentators on jazz have emphasized that there is often less improvisation on the concert platform than one might imagine. The musician is often 'playing safe' by using improvisatory devices which have worked well in other circumstances, so as to create the best effects he knows how. A so-called improvisation may, in fact, be a carefully planned and rehearsed performance; although there is nothing about the performance as such that would allow us to know this. To evaluate the degree of true improvisation, one would need to ask for two or more performances of the same piece for comparison. This observation is not to denigrate the musicians concerned. Improvisation is still at the heart of their art, but the social and commercial pressures of the concert platform do not encourage the risk-taking that improvisation inevitably involves. For real improvisatory jazz at its best one may have to seek out the late-night backroom informal sessions where, amidst a dedicated and sympathetic audience, experiments that fail, as well as those that reach new heights, may be observed.

As when discussing composition, we may ask how these informal observations on improvisation may lead to more rigorous scientific study of the processes involved. Sudnow's commentaries, for all their merits, were formulated some time after the improvisational acts in question, and at a level of some generality. There is a need for detailed and controlled study of *in vivo* improvisations. Although concurrent protocols could not easily be obtained from an improviser, the next best thing would be an immediate retrospection. By recording an improvisation and then playing it back to the performer, with as many pauses and backtracks as required, we could hope to obtain a detailed record of the conscious decisions involved in constructing

the improvisation. Experimental studies also suggest themselves. For instance, one could pre-record simple repeated chord sequences and ask performers of varying ability to provide as many different melodic variations as they can. Effects of the 'stereotypicality' of the chord sequence, or its speed, could then be examined. There is a rich untapped vein of data here which urgently awaits psychological attention.

# 5. Listening to music

## 5.1. Introduction

The two musical activities we have considered in detail so far, performing and composition, each entail the production of an externally observable end-product. Performance involves making a sequence of sound-producing movements, and composition involves making a score or some other communicable record. In contrast, listening to music is, in many situations, a passive affair. When I go to a symphony concert or listen to a gramophone record there may well be a lot of *mental* activity, but there is not necessarily any observable physical activity. The principal end-product of my listening activity is a series of fleeting, largely uncommunicable mental images, feelings, memories, and anticipations. When trying to understand what happens during music listening, the psychologist, therefore, is at a considerable disadvantage. In the previous two chapters we used concrete evidence from performances, scores, sketches, etc. to infer something about the cognitive processes that produced them. Such types of evidence provide what we may call 'records' of the music concerned, because they have a detailed event-by-event correspondence with the music. Many, if not most, instances of music listening do not provide 'records' in this sense.

Sometimes there is a specific behavioural activity attached to listening. For instance, one may dance, clap one's hands, or tap one's feet in response to the music. From such behaviour the psychologists may be able to infer something about the listener's ability to derive rhythm and metre from music (see the discussion of rhythm in Chapter 2, section 5.4); but such a response is hardly a *record* of the music. Rather, it is a generalized response to rhythmic characteristics which can be shared by very many pieces of music. By simply examining the response, one would not know *which* piece of music had been heard.

We approach closest to obtaining a musical record when we ask a listener to *recall* what he has heard, as we may do when teaching someone a new song. However, unless the material to be recalled is very short we are in danger of seriously underestimating the amount of musically related mental activity that has taken place. I have frequently been unable to recall details of a long piece of music after a first hearing, even though I was totally engaged with the music while it took place. As when listening to a lecture, or reading a novel, the details of argument and plot may fully engage one's intellect *whilst* one experiences the unfolding of ideas, but they may leave few recoverable traces at the end of the session.

The principal problem facing the student of listening processes is to find a valid way of tapping the moment-to-moment history of mental involvement with the music. As we examine research in the field of perception, attention, and memory in music we shall find that this problem has not really been solved satisfactorily. Most research evades this crucial issue by examining responses to very brief segments of music, made up of between two and twenty notes. Such segments hardly present listeners with the range of patterns and relationships which they must deal with in even the simplest short song.

In defence of such research one may, of course, argue that music is made up of a large number of small fragments chained together, and that music perception is simply a concatenation of a series of perceptual acts on such fragments. If so, however, then all the art of classical composition must go for nothing, because, as has been argued earlier, events quite far apart in time can be intimately connected structurally. Intelligent listening, we may argue, picks up such relationships; it spans large numbers of notes so as to be 'structural hearing' (to use the phrase coined by Salzer 1952). Composers write for listeners, not analysts, and the testimony of many listeners is that they *can* discern the large-scale relationships which analysts characterize.

What we may, more modestly, claim for studies on short musical fragments is that they shed light on elementary listening processes whose results are then used by higher processes about which little is known. In a similar way, students of language perception have discovered a great deal about the way in which listeners deal with isolated sentences (e.g. Fodor, Bever, and Garrett 1974), but these findings do not allow a very clear picture to be formed of how listeners or readers deal with whole paragraphs or books. For that, one has to find ways of observing behaviour in response to the larger units (Bransford 1979; Clark and Clark 1977; Sandford and Garrod 1981).

A further problem hinders our attempts to appreciate the significance of music listening studies. We must distinguish effects which are due to real and abiding features of normal listening from those which are special effects of the experimental task, which may involve activity quite unlike that of normal listening. Let me give a common example. Many studies require a listener to hear two short extracts of music and then to judge whether they are the same or different in some respect. The normal inference drawn from such studies is that if a listener can reliably tell when a particular type of difference is present, then he extracts this information in normal listening, and conversely, if he cannot reliably detect a difference, then he does not extract such information in normal listening. In fact, neither inference is necessarily valid. *Success* at discrimination in the experimental task may be a result of the subject's being able to focus on the dimension concerned for a brief period. This does not imply that he will choose to, or be able to, focus on this dimension in continuous listening where other dimensions of sound also demand his attention. For instance, fine differences in timing or tuning may

be noticeable in very short sequences, but be completely unrecoverable in longer ones. *Failures* of discrimination in the experimental task may be due to lack of cues that would be supplied by a more extended context. For instance, we shall see that listeners are often unable to tell short sequences apart when they have slightly different pitch intervals but identical up–down contour. It would be wrong to conclude from this that listeners cannot use pitch interval information in extended listening. It seems that in more realistic listening situations pitch information becomes crucial in discrimination (Dowling and Bartlett 1981).

Another example concerns the degree of familiarity that a listener has with the materials. In most music perception experiments this familiarity is low—the listener is exposed to sequences that he has never heard before. In contrast, real-life listening often involves the repeated hearing of the same material so that its internal structure becomes better and better known. Most serious listeners to music will have experienced times when they suddenly 'see' some new relationship between elements of a well-known piece of music—music that they may have lived with for many years. The experimental literature has told us very little about such processes, although recent work by Pollard-Gott (1983) has shown that subjects change their coding of segments from Lizst's Sonata in B minor after only a few opportunities to listen to the complete work. Briefly, she found that repeated exposure increased the likelihood that subjects would notice 'deep' harmonic or thematic similarities between segments, as opposed to 'surface' features, such as speed or tessitura.

I do not wish to dwell at any further length on the hazards and shortcomings of 'music perception' research. I have put these cautionary comments at the beginning of this chapter because it is a source of great disappointment to me that the one area of music psychology in which research is flourishing should be characterized by a relative insensitivity to the problems of relating research findings to normal music listening. This disappointment is shared by others (e.g. Davies 1979) but rarely voiced in print.

Another factor which prompts cautionary remarks at this point is the 'halo effect' which can surround a body of research which has the credentials of a *paradigm* as typified by Kuhn (1962). A paradigm is signalled by the existence of a set of agreed research problems, a methodology capable of solving them and generating new, but similar, problems, and a group of scientists who interact around these problems. When, as often happens, such a group 'captures' prestigious communication channels, in journals and elsewhere, generating a continuous series of co-ordinated and logically connected research articles, it is tempting to accord the research a significance which is not necessarily justified by its long-term achievments. The history of psychology is littered with forgotten paradigms. Many of them were initially motivated by important and substantial psychological problems, but in the end sub-

sidiary problems raised by the research became the main focus of potentially indefinite elaboration which contributed nothing to the original substantive question. I speak as one who has been seduced by the attractions of paradigm-driven research; and I have dabbled at the edges of a paradigm (Sloboda 1976c, 1977b) which is, I believe and hope, already being consigned to obscurity. I would not wish to go so far as to say that music perception research is unproductive, or to argue that science should proceed in any other way. I would argue, however, that we need to be constantly checking that questions of real musical importance are being asked and answered. It is all too easy to take currently researched issues to be *the* issues of music perception, just because a substantial body of research is *there* in all its solidity, interconnectedness, ingenuity, and undoubted methodological soundness.

If this causes me to stand a little lightly to the current research, mentioning some, but by no means all, of the experimental approaches to music perception, then I must say that I feel happy to do this, not least because there are several very good and fairly exhaustive reviews of experimental work in this area which absolve me from duplicating the effort here. I would advise the reader who is seriously interested in music perception to read one or all of the following: articles by Deutsch, Dowling, Shepard and Ward; all in Deutsch (1982a); and the authoritative survey and bibliography provided by Spender (1980).

### 5.2. 'Natural hearing': primitive grouping mechanisms in music

The first area of study that I wish to examine relates to a fundamental question about music. This question is: what natural mechanisms and propensities of the human auditory system can determine the way in which we hear musical sounds to be grouped? The principal characteristic of music is that sounds stand in significant relation to one another, not in isolation. For music perception to 'get off the ground' listeners must start to notice relationships and identify significant groupings. At the highest level such groupings will be large and complex, corresponding to formal structures within a work. At the lowest level they will be relatively simple, features of the sound pattern which characterize small groups of notes.

Studies of visual perception have revealed a number of basic and universal grouping tendencies which go under the name of 'Gestalt principles of perception' because they were first given systematic demonstration and formulation by the Gestalt school of psychologists (e.g. Wertheimer 1923). Although visual examples are most commonly cited, these principles can be seen as having application to any perceptual domain, representing as they are held to do, primitive and innate grouping tendencies. It seems that very young children and animals are subject to the operation of these principles as well as human adults (Gibson 1969).

If these principles are innate and universal, then it is possible that they

have been incorporated into the brain's way of operating in order to solve perennial perceptual problems faced by living organisms that interact with a natural environment. Some of these problems have been identified (see Sutherland 1973, Hochberg 1974, and Marr 1982, for vision; Bregman 1981, and Kubovy 1981, for hearing). For instance, the principle of 'common fate', which states that elements that move in the same direction are perceived together, allows a solution to the problem of how to identify the elements of a visual scene which belong to a single moving object that may be partially obscured or occluded by other visual elements. This principle explains why an animal which may be perfectly camouflaged against a background whilst stationary becomes instantly visible once it moves. Striking demonstrations by Johanssen (1976) show that conjoint motion of lights fitted to the joints of otherwise invisible humans is sufficient for an observer to reconstruct a percept of them which can not only identify them as humans, but allow a normal observer to make accurate estimates of their age, sex, and actions as well.

Most of the natural mechanisms of visual grouping are motivated by 'action' requirements. They help the organism to orient and move about the environment effectively, to locate and track objects in that environment, and so on. We should expect auditory grouping mechanisms to have their roots in similar requirements. Deutsch (1982b) suggests that a primary function of auditory grouping mechanisms is to assist in the enumeration of sound-producing objects in the environment. Such objects are likely to be of importance because they are other animals (friendly or unfriendly) or because they signal significant change in the environment (e.g. fruit falling from trees).

We may imagine the sounds of a forest at night, a potentially bewildering confusion of overlapping sounds. Yet it is possible to separate and group the sounds. We hear the complex sound-pattern created by the wind rustling in the trees as unified, yet separate from the chirps of insects. And we hear the calls of three birds of the same species *as* three sounds. To each source is assigned the correct 'package' of sounds. Gestalt grouping mechanisms underlie our ability to 'parse' the acoustic environment effortlessly. Thus, Deutsch argues that 'similar sounds are likely to be emanating from the same source, and different sounds from different sources. A sound sequence that changes smoothly in frequency is likely to be emanating from a single source. Components of a complex sound spectrum that rise and fall in synchrony are also likely to be emanating from the same source' (Deutsch 1982b, p. 101). These likelihoods govern the way we group complex sound 'scenes'. The existence of such grouping tendencies is, arguably, one of the happy chances that allows us to hear *musical* groupings spontaneously. Let us look at the operation of some of these tendencies in more detail.

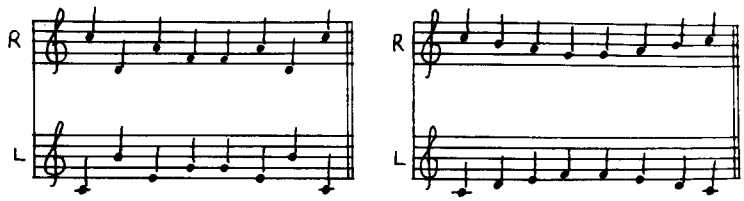
Sound sources usually have definite physical locations, and a possible way of sorting out a complex set of sounds is to group them according to perceived

location. One way of locating the direction of an unseen sound source is through the comparison of the sound signals that arrive at each ear. When a sound to the left of a listener is turned on, the first impulses arrive at the left ear slightly before they arrive at the right. The technical term for this time differences is 'onset asynchrony'. Humans, and many animals, have precocious ability to locate an unseen sound source through onset asynchrony. Quite young babies will turn their heads to look in the direction from which a sound has come.

A listener can also locate a steady tone through phase differences between the two ears. Because sound is transmitted through air by a series of air molecules, and because points of maximum compression move outwards from the source at a finite speed (taking about 680 $\mu$s to traverse the 23 cm between the two ears), there is a time difference between the energy peak at the two ears which can be detected. In this case, though, the energy differences are small compared to sounds which have sharp onset or offset bursts (known as transients). Therefore, sounds with transients, such as clicks and chirps, are more accurately localized than steady tones. The wise hi-fi salesman uses this fact when demonstrating the power of his stereo systems. Music with percussive elements will give rise to a more compelling 'stereophonic' effect than will the 'smooth' characteristics of voice or string sounds.

There are other problems with grouping by localization. Echoes and reverberations can cause sounds to arrive at the ear from other directions than the source. Also, the auditory system has no way of distinguishing between a single sound coming from directly in front of the face and two identical sounds coming from equidistant sources on either side. For these and other reasons it has been argued that location is a less important grouping cue than *pitch* (see Bregman 1981; Bregman and Campbell 1971; Deutsch 1982b; Kubovy 1981). In nature, sounds coming from the same source over a brief duration will also tend to have the same, or similar pitch. If pitch *does* change it is more likely to be smoothly, or in small steps, than discontinuously or in large steps. Additionally, in nature, perceived pitch and perceived location are often mutually reinforcing; confirming the *same* grouping of sounds.

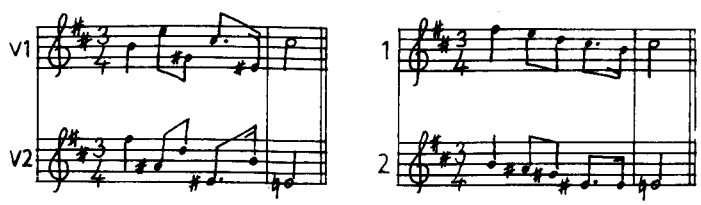
Man-made sounds, in music and in the laboratory, can deliberately set grouping cues in opposition, creating illusory or ambiguous percepts which demonstrate the relative strengths of the grouping tendencies. A good example of this work has been provided by Deutsch (1975) who played listeners two simultaneous tone sequences through headphones, one to each ear. The two sequences are shown in Example 5.1. The commonest perceptual experience which arose from these stimuli is shown in Example 5.2. Listeners heard all the high tones emanating from the right headphone and all the low tones emanating from the left headphone. Deutsch calls this phenomenon the *scale illusion* because listeners hear the stimulus as two smooth scale passages rather than as the two angular melodic contours which were actually present.




Example 5.1. and 5.2. The scale illusion. From Deutsch (1975).

Butler (1979) has demonstrated that this pitch grouping phenomenon is highly robust. Deutsch used pure tones with headphones, but Butler used real instrumental sounds (e.g. piano) with spatially separated loudspeakers. Even when the notes from one speaker had a distinctive timbre, listeners still predominantly grouped by pitch.

Composers seem to have discovered this phenomenon long before it was brought into the laboratory. The best known example occurs at the beginning of the last movement of Tchaikovsky's Sixth Symphony. Example 5.3 gives the first and second violin parts as played. Example 5.4 shows what most

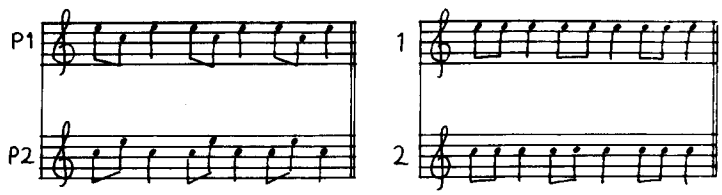



Examples 5.3. and 5.4. An analogue of the scale illusion. From Tchaikovsky, sixth symphony, last movement.

listeners hear. Those who know the work by ear before they see the score sometimes cannot believe that there has not been a typographical error! The 'illusion' does not require the sounds of the two violin sections to come from the same direction. The perception is still the same when the two sections are on opposite sides of the conductor's rostrum.

It is not totally clear *why* Tchaikovsky chose to write the parts in this way if they are really indistinguishable from the case in which the violins had actually played Example 5.4. It is possible that typical violinists produce subtly different sounds when negotiating an angular rather than a smooth melodic sequence, which would be heard as a difference in texture or timbre.

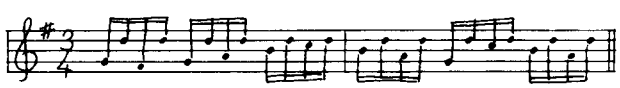
A less well-known example is to be found in the second movement of Rachmaninov's Second Suite for Two Pianos, Op. 17. Example 5.5 shows a transposed version of the notes played by first and second piano towards the end of the movement. I was recently involved in a performance of this work, playing first piano, and was subject to the compelling illusion that Example 5.6 was being played. This was even though I could see my hand playing two




Examples 5.5. and 5.6. A perceptual illusion experienced in Rachmaninov's suite for two pianos, Op. 17, second movement.

notes, and not just repeating one. It was as if someone had suddenly retuned the lower note I was playing to the same pitch as the upper one. We may suppose that Rachmaninov wrote the parts in this way because it would be difficult for a pianist to repeat the same note in the figuration of Example 5.6 with the kind of speed and dynamic control which the music requires (about six quavers per second with a controlled decrescendo to pianissimo).

In the examples we have considered so far there are two *really* distinct sound sources producing simultaneous sounds. Pitch grouping causes some of the sounds to be mislocated. There are other cases when grouping by pitch (or *pitch streaming* as it is often called) allows a single sound source to be heard *as if* it were two independent sources. Thus, a solo instrument can create an illusion of polyphony by interleaving the notes of two contrapuntal lines separated in pitch. This device has frequently been used by composers, especially in the baroque period. Example 5.7 shows this device at work in the right hand of J. S. Bach's Prelude in G from book 2 of the Well-Tempered Clavier.



Example 5.7. From J. S. Bach, *Well-Tempered Clavier*, Prelude 8, Book 2.

Whether or not the single melodic line is heard as one or two pitch streams depends crucially on the pitch separation and the speed of the component notes. Miller and Heise (1950) presented listeners with a sequence containing two notes that alternated at a rate of ten per second. They varied the pitch separation of the two notes and found that intervals of less than about one seventh of an octave were heard as a *single* stream, whereas intervals greater than this caused the notes to separate out into two independent streams. In terms of the chromatic scale, this means that intervals of a minor or major second produce *fused* percepts which correspond to musical trills. Notes a minor third apart or greater tend to be heard as separate, producing stream *fission*. Van Noorden (1975) has shown that if the rate of alternation is slowed beyond about six notes per second, it becomes possible to hear wider intervals up to an octave as fused to form a single stream. The effect is, within limits,

open to conscious control. It is possible to switch one's perception of an intermediate case so that at one time it is heard as a trill or single line, at another time as two independent lines. There seem to be limits, however, beyond which the percept is not open to such switches. Intervals of a tone or less, when repeatedly alternated, tend not to give rise to fission. No matter how slow the speed of performance, they are heard as a single line.

None the less, Example 5.7 suggests a case in which intervals of a tone or less *may* be heard in separate streams. Here, the C at the end of the first bar is heard as belonging to the lower stream, and does *not* fuse with the upper stream of repeated Ds. We may suppose that larger patterning context overrides the immediate factors, so that the repeated Ds are heard as one line, and the rising and falling scale as another. This tendency is accounted for under another principle of Gestalt grouping, the principle of *good continuation*. We may illustrate this by an almost exact visual analogy. In Fig. 5.1 (*a*) the lower dot is seen as forming a group with the two adjacent upper dots. In Fig 5.1 (*b*) it is 'captured' by the lower dots and becomes part of a curved line distinct from the upper straight line.

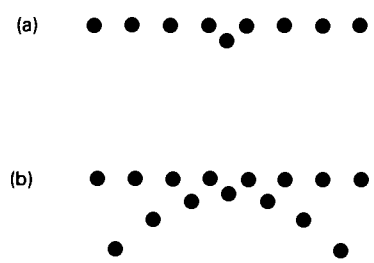



Fig. 5.1. Example of gestalt principles of perceptual organization.
(*a*) The lower circle appears joined to its neighbours (principle of proximity).
(*b*) The circle is 'captured' by the curved sequence (principle of good continuation overrides proximity).

The basic pitch streaming phenomenon demonstrated by Miller and Heise also has a fairly direct visual analogy (as noticed by Kubovy 1981). If two nearby point light sources are switched on and off in alternation in a darkened room, the impression is often of a *single* light moving backwards and forwards between the two locations. This well-known visual illusion is called the *Phi phenomenon*. The likelihood of the two lights being seen as a single moving light is decreased if the speed of alternation is increased, or if the sources are moved further apart. Under these latter conditions one is more likely to see two independent 'flashing' light sources. When tones are substituted for lights, and pitch distance for spatial distance, the Phi phenomenon becomes the pitch streaming phenomenon.

We can suppose that the striking parallel between these phenomena has

its origin in the general properties of natural objects which the perceptual system has developed to deal with. Two sounds very close to one another in time but widely separated in pitch are unlikely to come from the same object. Likewise, two visual events very close to one another in time but widely separated in space are unlikely to come from the same object. As the time separation between events increases, however, the possibility of them belonging to the same object increases. A high sound may come from the same animal as a low sound once it has had time to adjust voice register. A light on the left may come from the same object as a light on the right if it has had time to move between the two locations, unseen.

The work of researchers such as Miller and Heise (1950) and Van Noorden (1975) has demonstrated pitch streaming in very simplified perceptual situations. Dowling (1973) extended these studies to more musically realistic stimuli. He took familiar melodies (such as 'Frère Jacques', 'Happy Birthday', etc.) and interleaved pairs of them so that listeners heard note 1 of tune A, then note 1 of tune B, then note 2 of tune A, then note 2 of tune B, and so on. These composite sequences were played at a rate of eight notes per second (four of each tune). Listeners were required to identify the component melodies. Dowling found that recognition of the melodies was almost impossible when they overlapped in pitch. They merged into a single, unrecognizable sequence. However, when the melodies were moved apart in pitch, so that their notes did not overlap, the melodies could easily be recognized.

Dowling found one condition in which overlapping melodies could be recognized. This was when listeners were asked to search for particular named melodies. Melodies perceived in this way were done so with effort, however. They had none of the ease and spontaneity of recognition enjoyed by the pitch-separated melodies. The process was one of active search rather than passive awareness. A visual analogy may again be helpful. Figure 5.2 will be seen spontaneously as two triangles. You do not naturally see the

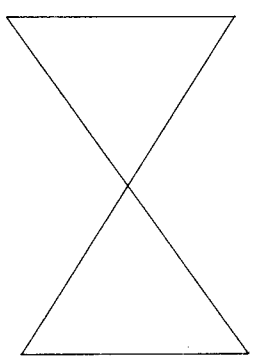


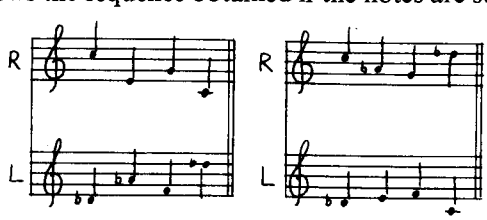
Fig. 5.2.

numeral 7; but if I ask you to *find* a 7 in the figure, you can do so. In the former case, your perception was governed by the spontaneous and largely automatic Gestalt principles of grouping (the 'laws' of proximity, closure, and symmetry being operative here). In the latter case you use prior knowledge about the shape of a figure 7 to overcome your natural grouping tendencies and achieve a more difficult, less spontaneous, way of seeing the figure.

Dowling also found that unfamiliar melodic patterns (learned by the listeners during the experiment) could not be recognized so easily when interleaved with other melodies, even when there was pitch separation of as much as an octave. This shows that the use that can be made of pitch streaming is not invariant, but can be helped or hindered by acquired knowledge about music.

This conclusion raises a general question about *all* the demonstrations we have examined in this section. Do they *really* show the natural auditory system at work, unaffected by the musical experience that most of us share? Many of the experiments use materials which make most *musical* sense when grouped by pitch. The grouping could be a result of application of knowledge about likely melodic patterns in music. This argument applies, for instance, to Deutsch's scale illusion. The melodies produced by each sound location are angular, with atypical pitch contours, whereas those arising through the supposed operation of pitch streaming conform to one of the commonest of all melodic patterns, the diatonic scale. Musical knowledge would, therefore, contribute to the tendency to group by pitch. There is, of course, a curious circularity here. It may be precisely *because* pitch grouping is a a fundamental auditory phenomenon that many common musical sequences move within restricted pitch ranges.

None the less, it should be possible to tease apart the contributions of fundamental auditory groupings and musical knowledge. One way of doing this would be to modify Deutsch's (1975) stimuli so that pairs of simultaneous melodies made more musical sense when grouped by location than when grouped by pitch. Example 5.8 shows such a pair. Here, the sequence played to the right ear forms a C major arpeggio/chord whilst that played to the left ear forms a D flat major arpeggio/chord. Each pattern is commonly found in tonal music, and has a strong and simple harmonic structure. Example 5.9 shows the sequence obtained if the notes are streamed by pitch,




Examples 5.8. and 5.9.

taking the upper note in each pair for the first stream, and the lower note for the second stream. Both streams contain unusual pitch leaps (diminished fifth, augmented second) rather than the smooth scales of Deutsch's illusion, and do not have straightforward harmonic implications. If the tendency to group by pitch is lessened in such sequences, then musical knowledge is implicated in streaming mechanisms.

A second way of determining the part played by musical knowledge in such effects is to look for differences between experienced musicians and novices. If differences exist, then musical knowledge is implicated. This approach has been taken by Smith, Hausfield, Power, and Gorta (1982) who compared musicians and non-musicians on several varieties of the scale illusion. They found that musicians were more likely than non-musicians to stream by pitch, even when there were conflicting cues (such as timbre) suggesting a streaming by ear. They argue for the conclusion that musicians are more likely to expect 'tunelike' stimuli than non-musicians.

The most convincing demonstrations that pitch streaming is a primary auditory grouping mechanism, regardless of musical knowledge or structure are, perhaps, those that require perceptual judgements that have nothing to do with music as such, and point to the existence of pitch streams by implication, rather than because listeners are asked to directly *identify* streams. I will give just one example of such a demonstration. Fitzgibbons, Pollatsek, and Thomas (1974) asked subjects to detect a small temporal gap in a sequence of four tones, the first two of which were high in pitch (2093 and 2394 Hz) and the last two of which were low in pitch (440 and 494 Hz). This sequence (HHLL) was presented with the gap (G) in three possible positions: HGHLL, HHGLL, and HHLGL. In two cases the gap was between notes close in pitch, in the other case (HHGLL) it fell between two notes widely different in pitch. In the former two cases gap detection was nearly perfect, but in the latter case, detection of the same gap was significantly poorer. The interpretation given to these results is that listeners are able to notice time relationships *within* pitch streams more easily than time relationships between different pitch streams. Such relationships are not preserved in the automatic 'parsing' mechanisms which split such stimuli into two independent parts on the basis of pitch.

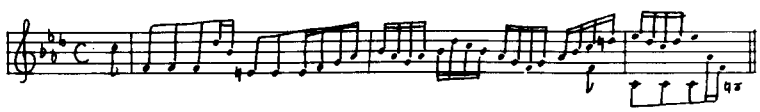
Such demonstrations provide strong evidence that pitch streaming is a real 'pre-musical' phenomenon, although musical knowledge may interact with and modify its effects. Pitch is not the only 'primitive' cue to grouping. Other cues, such as timbre and amplitude are described by Deutsch (1982b). I do not intend to deal with them in a detailed way here. Rather I would like to conclude this section by looking at some more complex and musically sophisticated ways in which composers can use pitch relationships as the basis for creating dynamic impetus in extended melodies.

This discussion owes much to the theoretical approach to melody outlined by Meyer (1973). According to Meyer, one of the functions of a melodic line

is to create *implications* for future events. In earlier writings (Meyer 1956) he used concepts of 'creating expectancies' in listeners, but this is a slightly misleading way of talking. The implications are inherent in the melody itself, by virtue of its structure; such implications may or may not be picked up by a listener and used to form expectancies on a given hearing.

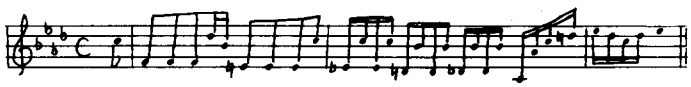
For Meyer, implications are objective features of a melody which create some patterning or direction which admits of continuation. One implication, for instance, of an ascending five-note scale passage on its own is for a sixth note to be added to continue the sequence. An implication of the sequence CDCDCD is that CD should be repeated again; and so on. These are very simple implications. Any interesting melody has multiple (and sometimes conflicting) implications. A composer may not necessarily realize all of the implications of a melodic opening. Some he may not notice, others he may deliberately go against for musical effect. Even those that he does realize in a continuation may not be noticed as such, even by an acute listener. Rather, they contribute to an undifferentiated awareness that this is a good or interesting melody.

There are many devices for creating implications which rely on tonal or rhythmic factors. Here, I want to consider the use of pitch separation alone as an implicative device. Consider the melody which J. S. Bach uses for the F minor Fugue from book 2 of the Well-Tempered Clavier (Example 5.10).



Example 5.10. From J. S. Bach, *Well-Tempered Clavier*, Fugue 12, Book 2.

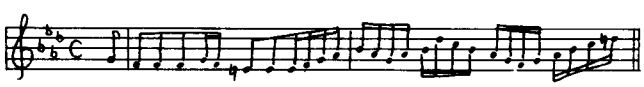
The opening nine notes contain some quite large pitch separations between adjacent notes. Thereafter, the melody moves in scalar fashion. The ideal performing speed for this fugue is about four quavers per second. This places the melody in the range established by Van Noorden (1975) where intervals of up to an octave between successive notes can be heard either as separating the two notes into distinct streams, or merging them into one, depending upon one's 'set'. So the opening of Bach's melody can be heard as a single unified line; but it has ambiguity—it can also separate out into two lines, an upper one created by notes 1, 5, and 6 (C, D flat, and B flat), and a lower one created by notes 2, 3, 4, and 7, 8, 9 (F and E). On the other hand, the material after note 9 proceeds by scalar steps of major and minor seconds, each note therefore being necessarily perceived as part of a single line with its neighbours. So these semiquaver runs serve as a sort of retrospective 'glue', bringing together the potentially separable lines of the opening notes. A different continuation could, in contrast, have sustained and accentuated the line separation (Example 5.11). Bach's intention seems to have been to create a tension between two potentially separable melodic lines which is



Example 5.11.

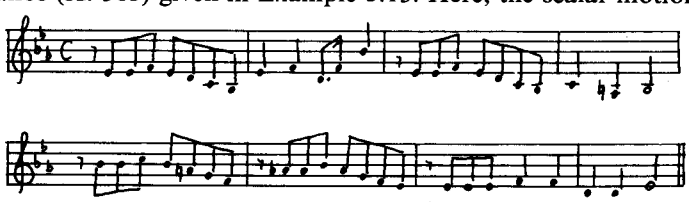
then resolved by 'knitting' them together. In Rosner and Meyer's (1982) terminology, the melody has features of a type labelled 'gap-fill' where an unfilled interval is created, then filled through scalar movement.

This is not the whole story, however. Each melodic line of the opening has its own implications which are not realized in the immediately following notes. Each creates a *direction* of pitch movement. The lower line on its own creates an implication for further downward movement (F→ E→). Conversely, the upper line has an implication for further upward movement. The topmost note is C, then D flat; an even higher note is implied. Both these implications are left unfulfilled in bar 2, but they are both realized on the first beat of bar 3: the lower movement through the second fugal entry which drops to middle C; and the upper movement through the simultaneous achievment of top E flat in the first voice, the highest note in the piece so far. This resolution of unfulfilled implications contributes to the sense of closure experienced as the second theme enters, and gives the melody a dynamic forward movement that makes it vital and satisfying. Consider how lifeless Example 5.12 is in comparison, where there is no upper pitch line, and therefore no sense that the final E flat resolves an earlier implication.



Example 5.12.

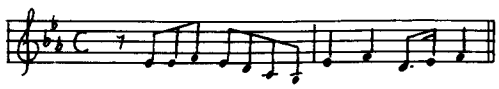
A second example of the implicative use of pitch streaming is provided by the principal theme of the last movement of Mozart's Concerto in E flat for Two Pianos (K. 365) given in Example 5.13. Here, the scalar motion of the



Example 5.13. From Mozart, Concerto for two pianos (K.365), last movement.

opening creates a unified melodic line spanning the F above middle C to the B flat below it. The last note of bar 2, an upper B flat, stands isolated from this line by a perfect fourth. In bars 3 and 4, the melody continues in the initial lower register. Arguably, the single upper note creates an implication for further notes in that pitch range. Such implication is only realized in bar

5, where the melody leaps up to this B flat before descending in scalar fashion to bring the melody back to the opening E flat. Bar 5 is, therefore, motivated by what has gone before in a way that it would not be without the bar 2 that Mozart supplied. If we rewrite the first two bars as in Example 5.14, then the pitch leap at bar 5 comes as arbitrary and surprising.



Example 5.14.

My treatment of these examples has been discursive and analytic in the musical sense. However, I have imputed psychological consequences of listening to them which have not been strictly demonstrated. What kind of evidence could we supply that a supposed implication of a melody has actually been detected by a listener? One, rather naïve, approach would be to play listeners the opening parts of melodies, and ask them to provide fitting continuations. For instance, we could compare listener's responses to the first four bars of Example 5.13 with the same four bars as modified in Example 5.14. If they pick up the implication of the high B flat, then Example 5.13 should result in significantly more continuations in a pitch range around that B flat.

I call this approach naïve because it assumes that implication is likely to be detected on a first hearing in a way that could guide conscious choice. It is equally possible that, for many listeners, the implications of a good melody are revealed gradually over many hearings, and that, even when revealed, can be used only to *recognize* fitting continuations rather than to *generate* one's own continuations. Such considerations make it quite likely that no positive results would be obtained with this technique. None the less, something of the sort is clearly worth trying. The only experiments I know to have used a similar technique (Carlsen 1981) involved presenting subjects with simple two-note sequences and asking them to provide plausible continuations. The types of implications that can be contained in two-note sequences are very limited, and hardly begin to tap the range of devices used by real composers. As far as they go, however, Carlsen's results show some evidence for pitch range sensitivity. His subjects showed a marked tendency to choose a third note close in pitch to the second note given. Sixty per cent of all continuations began within a major second of the second note given. It would be interesting to know whether subsequent continuations employed the 'gap-fill' technique, moving back to the first note through scale-like movement, but Carlsen analysed only the first note of each subject's response.

The mechanisms of melodic perception required to notice implications of the type discussed above involve more than can be supplied by the involuntary and immediate processes of auditory grouping. They require *attention* to the salient aspects of a complex stimulus, and *memory* of past

events which stand in important relationships to present material. The next two sections deal with these aspects of music listening.

### 5.3. Attention in music listening

We begin our consideration of attention by returning to Dowling's (1973) experiments with interleaved melodies. We saw earlier that subjects could recognize both melodies provided that they did not overlap in pitch, but formed separate perceptual streams. However, Dowling reports that subjects did not identify both melodies at the same time. More commonly they would identify one of the melodies on a given hearing, and the other on the next hearing. They seemed to be *selecting* either one or the other stream to monitor; but were not able to monitor both streams at once. This result points to the existence of a serious problem in music listening. A great amount of music is polyphonic; that is, it has more than one line of music proceeding at the same time. If we can attend to only one line at a time, how do we ever achieve a full experience of the polyphonic structure?

Pioneering studies on attention to speech (Cherry 1953) have shown that when two speech messages are presented simultaneously it is possible to shadow (verbally report) one of them, but that almost nothing is known about the characteristics of the other message. Thus, for instance, the non-shadowed message may change language, or repeat the same phrase over and over again without the subject being aware of it. These studies, and others like them, have been used to support the notion that our perceptual system incorporates a single, limited capacity attentional 'channel', through which only a small part of our sensory experience can pass at any one time. This mechanism operates like a 'filter' which lets through only material which is defined by some distinguishing sensory feature (like pitch, or ear of entry) at any one time. All other material is lost before it can reach those higher mechanisms which recognize and classify input. Broadbent (1958) provides the classic formulation of such a position.

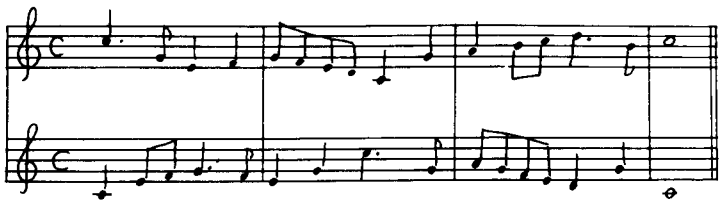
More recent studies have challenged the 'filter' theory in various ways (Allport, Antonis, and Reynolds 1972; Deutsch and Deutsch 1963; Shiffrin and Schneider 1977; Triesman 1964). For instance, Allport, Antonis, and Reynolds asked pianists to sight-read piano music at the same time as shadowing a prose message played over headphones. They found that performance on both tasks was just as good when carried out together as when carried out alone. Such a finding is incompatible with the filter theory, which would predict that subjects could attend either to the sight-reading input (visual) or to the speech input (auditory) but not to both. An alternative way of explaining attentional phenomena is to propose that processes *may* take place simultaneously provided that they do not use the same kinds of cognitive mechanisms. In Cherry's and other speech experiments, both messages were verbal. In the study of Allport *et al.*, one message was verbal, the other

was musical. Each message required a different *type* of processing.

If we examine everyday situations where we are able to do two things at once, we see that most of them have similar properties to the demonstration of Allport *et al.* We can hold a conversation whilst driving a car; listen to the radio whilst doing housework, and so on; but we *cannot* listen to a radio talk whilst reading a book.

What is the relevance of this to music? It suggests that our difficulties in attending to two concurrent melodies are not so much due to an incapacity to take them in as to an incapacity to subject them to the same kind of *analysis* simultaneously. In situations where nothing is required of an observer other than the detection of some simple event, he is capable of monitoring very many simultaneous 'channels', even when only a single sense modality is involved. For instance, Shiffrin and Schneider (1977) report demonstrations that people find it as easy to monitor 49 separate spatial locations for the occurrence of a brief visual event as they do to monitor only one location. It is when some integrative response is required to the material in each channel, such as detecting sequential structure (as in understanding a sentence or recognizing a tune) that problems begin to occur. As Neisser (1983) puts it, 'there is no way to avoid assigning the verb in speaker B's sentence to the noun in speaker A's, with resulting organizational chaos.' When we hear two people speaking at once, what each of them says is an independent message, bearing no necessary relation to the other. Trying to construct the two independent but simultaneous structures yields misleading and unintended results as elements are assigned to the wrong structure. It is hardly surprising that the conventions of normal human linguistic communication are designed to ensure that only one person speaks at a time.

In Dowling's (1973) experiment, the two interleaved melodies could be just as unrelated to one another as two randomly chosen verbal messages. In most contrapuntal music, however, the various lines are *not* independent. They are constructed with skill and care to be related to one another. Each line possesses not only its own 'horizontal' melodic identity (enabling it to stand on its own as a valid musical sequence) but also a 'vertical' harmonic function with relation to the other concurrent lines. At any instant, the simultaneously sounding notes make up a chord which has its own musical identity. The progression of chords found in the piece constitutes the harmonic structure of the music, and this structure acts as a single underlying framework in which the individual lines can be placed. We may hypothesize that this underlying harmonic unity makes it easier to attend to several simultaneous melodic lines. This hypothesis was tested in an experiment (Sloboda and Edworthy 1981) where subjects learned two melodies which cohered harmonically when played in the same key (Example 5.15). After each melody was learned on its own, subjects heard trials where the two melodies were played together. The melodies were always separated in pitch by at least an octave, so that they were easily heard as separate streams. On



Example 5.15. Melodies used by Sloboda and Edworthy (1981).

each trial, one of the melodies contained a pitch error on one of the notes. Subjects were required to say which melody contained the error, and, if they were musically literate, to mark the error on a score.

The crucial experimental variable was the key relation of the two melodies. In condition A the melodies were in the same key. In condition B they were in keys a perfect fifth apart (e.g. C and G.) In condition C they were in keys an augmented fourth apart (e.g. C and F sharp). We predicted that, if harmonic coherence aids simultaneous attention, subjects should correctly locate most errors in condition A and least in condition C. The results exactly upheld these predictions. Subjects were 80 per cent correct in condition A, but only 67 per cent correct in condition C. Condition B gave intermediate results (74 per cent).

How does harmonic coherence aid attention? Our current hypothesis is that polyphonic music is perceived as an ambiguous pattern capable of 'figure-ground reversal.' There are many well-known examples of this. Rubin's 'faces-vase' figure is one of the best known. It oscillates between being seen as a white vase on a black background, and as two black faces looking at each other across a white background (Fig. 5.3). Shape and contour is

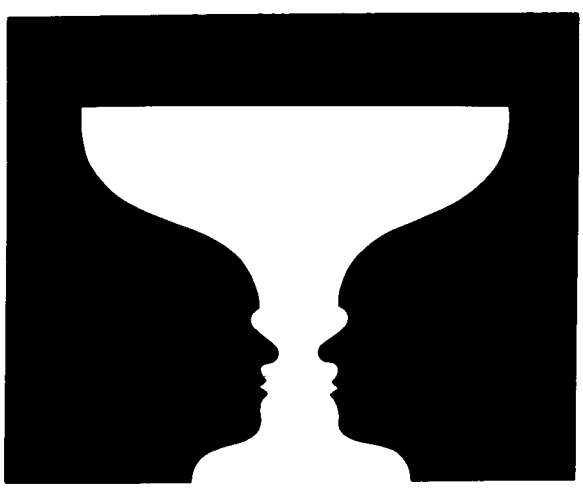


Fig. 5.3.

seen as being possessed by the elements which form the figure at a particular time. The background is seen as a contourless occluded surface.

In music we propose that only one melodic line can be treated as 'figure' at any one time. When so treated we may say that this line is being given 'focal attention'. Focal attention allows the noticing of relationships within the melodic line, so that the melody may be recognized, related to previous material, and so on. The other line, or lines, form the background. They are registered, but are not processed focally. Instead, they are fragmented into a series of individual notes which are heard 'vertically' as chords which support or accompany the focal melody. Thus, there are two types of process taking place, melodic processing of the focal line, and harmonic processing of the other parts. Furthermore, each note of the melody being focally processed has a harmonic function which is confirmed by the notes in the other parts, so both the melodic and the harmonic processes contribute to the building up of a unified structural representation for the whole piece.

A diagram may help to make clear the relationships being proposed. Consider a piece of music containing three contrapuntal lines. Fig. 5.4(*a*) represents the primary relationships noticed when line A is the subject of focal attention. The listener extracts the 'horizontal' melodic relationships between notes in A (represented by double lines) but he does not hear B and C as melodies. Rather, their individual notes are heard as harmonizations for the concurrent note of A, assisting the understanding of the structure of A. Fig. 5.4(*b*) represents the case where a listener is hearing precisely the same music, but this time giving line B focal attention. Now A and C become background harmony.

How does this formulation explain our experimental results? We suppose that listeners will succeed in incorporating most contrapuntal material into

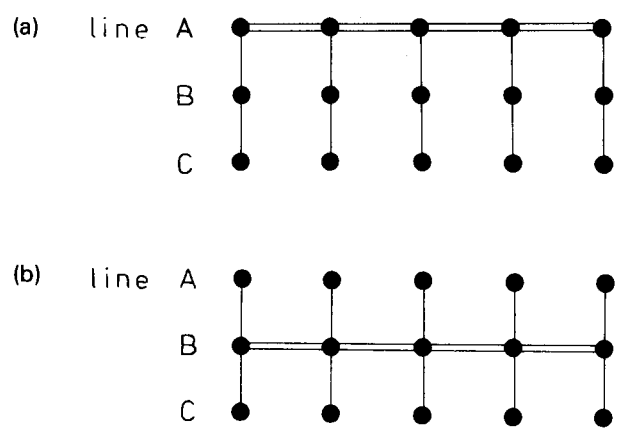



Fig. 5.4. Two different possibilties for focal processing in a contrapuntal sequence.

their representation of the music when the chords supplied by the background parts are familiar ones which support a plausible tonal harmonization for the focal melody. When the two melodies of Example 5.15 are played in the same key, they mutually reinforce a simple and common harmonic structure. Bars 1 and 2 are subsumed under the tonic chord (I), bar 3 subdominant to dominant (IV–V), and bar 4 tonic again (I). With such a conventional progression, the experience of most listeners is that a wrong note 'sticks out like a sore thumb' because it creates an odd and unexpected harmony. It does not matter what line is being given focal attention; the 'queer' harmony is heard. If the wrong note is in the focal melody it will be detected as a deviation from the known melodic pattern. If, however, the focal melody 'matches' the listener's memory of it, then the inference is that the wrong note must be in the other, non-focal melody.

If our explanation is the correct one, then there is a simple prediction which may be tested. This is that when there are *more* than two simultaneous lines, location of wrong notes will become less efficient, even when all lines cohere harmonically. This is because in the case where the mistake is in a non-focally processed line, the listener has to guess at random between these lines. All he will know is that there was a wrong note, creating an odd harmony, which was *not* in the focal melody. By simple application of probability theory we would expect the 80 per cent performance with two lines to drop to 53 per cent with three lines, and 40 per cent with four lines. This prediction is yet to be tested.

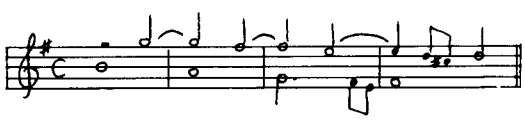
Informal support for this prediction comes from a common experience of ensemble conducting. When a conductor is rehearsing a group he will often hear 'something go wrong' but be unable to say with certainty which part went wrong without going back and hearing the section again. It appears that he was focally processing a part which did not go wrong, and needs to redirect his focal attention to locate the mistake.

If focal processing is limited to one part, it becomes pertinent to ask how, if at all, it is possible to justify the claim made by some listeners that they *can* follow all the simultaneous parts of a contrapuntal piece. One factor which may allow things to seem that way is a high degree of familiarity with the individual parts, through repeated hearing. Each part is then so well known that it does not need to be continuously monitored to refresh the experience of its being 'there'. This kind of listening is rather like the experience we achieve when casting our eyes round a familiar visual scene, such as our own room. Although the eye can only focus on one thing at a time, the fact that each glance fulfils an expectancy of what *should* be there allows a real experience that the whole scene is present in one's perception. In familiar music, the 'mind's ear' can dart around between the parts, always knowing what it is going to find at any given moment, and finding it, thus confirming to the knowing mind that all the parts *are* present.

Even if each contrapuntal line is not well known as such, it is also possible

to know a piece very well *as* a melody plus accompaniment (in the case where the same contrapuntal line has always been the object of focal attention). In such a case, each chord may be so well known that the listener has the possibility of *constructing* a non-focal line by picking out a given note from each chord and integrating a new line. In such cases all the contrapuntal lines may be *implicit* in what a listener knows, simply awaiting a reorganization of his knowledge to be made explicit. In such a way may a listener suddenly realize that the 'accompaniment' of a well-known passage actually has its own melodic integrity. It is not that there are notes present which he has not noticed before. It is simply that he has not heard them *as* melody before.

Another way of achieving the impression of total awareness in polyphony has more to do with the composer's art than with the listener's knowledge. In the cases discussed so far we have assumed that a note in any one part is associated with a simultaneously played note in every other part. Whilst this is true of some chordal counterpoint, such as hymn harmonizations, it is by no means true of all counterpoint. It is possible to lighten the listening load considerably by limiting significant events to one or two lines at any one time. Consider, for example, the counterpoint of the top two lines in no. 18 of J. S. Bach's Goldberg Variations (Example 5.16). These lines (which form



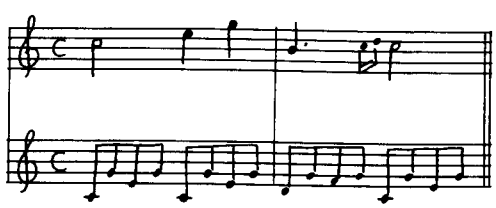
Example 5.16. From J. S. Bach, Goldberg Variations.

a canon) are distinguished by the fact that there is not a single occasion when notes begin simultaneously in the two parts, although the notes are sustained in such a way that the two-part texture is continuous. We may argue that listeners can focally process each melody, by switching from line to line as each new note is played. Secondly, once one has realized the canonic relationship it is possible to predict the next event in the upper line by reference to what has just happened in the lower one. This is a quite different situation to that of Dowling's 'interleaved melodies' experiment. First, the rate of switching between parts is much less (about one per second rather than four per second). Second, because of the musical relationship between the two parts, understanding one part can directly help a listener understand the other part.

In this Bach example both contrapuntal lines are equally important. In some cases, however, the composer may quite clearly mark one line as containing the most significant material, with another part relegated to a secondary role where texture may be more important than the precise melodic sequence. For instance, in the orchestral music of Rachmaninov one often finds subsidiary countermelodies of a 'busy' scalar nature whose function is to provide a restless undercurrent to some more salient theme. Although a

connoisseur may come to know every note of such countermelodies, there may be a sense in which this is knowing more than Rachmaninov intended. It is enough for adequate listening that one can detect 'that kind of thing going on' which can be done, perhaps, without too much focal attention, in the 'spaces' between significant events in the primary melody.

If, indeed, focal attention can only be in one place at a time we may move to consider a compositional problem. Is it possible for composers to write music which will tend to direct the focal attention of first-time listeners to the intended part? We ask this question on the grounds that it is not correct to say, for most music, that all parts are equally important for a first-time listener; that he would 'hear the music rightly' whatever part he happened to focus on. This may possibly be true for a fugue; it is not true for a symphony or a sonata. In such works there are nearly always primary themes which must be apprehended, because later transformations and extensions of these themes form the substance of the compositional structure. It would hardly do if a listener focally processed the lower line of Example 5.17 (Mozart's Piano Sonata in C, K. 545) at the expense of the upper line.



Example 5.17. From Mozart, Piano Sonata in C (K. 545), first movement.

Almost no experimental work relates directly to this question, but it is possible to point out some of the most easily observable 'attention grabbing' devices. First, it appears that, all other things being equal, it is the line with the highest pitch range that tends to be focally processed. When listening to, say, a hymn tune harmonization, many listeners are simply unaware that the lower parts comprise melodic lines. In the vast majority of classical music, principal themes are stated by high-pitched instruments, with lower ones playing subsidiary material. For instance, in string music it is nearly always violins, rather than violas or cellos, which present the principal themes; and in piano music it is nearly always the right hand. Why should this be? One possible contributory factor is that outside lines are heard better because they are *on* the outside and so are not 'surrounded' by other lines. Thus, the highest and lowest lines of a polyphonic texture should be most salient, with inner lines being less salient. We still need a further reason for the topmost line 'winning out' over the bottom line. One possible reason is that high notes in normal musical textures mask one another less than do low notes. 'Masking' is a technical term of psychoacoustics which refers to the detrimental effect the presence of one tone has on the ability of a listener to hear

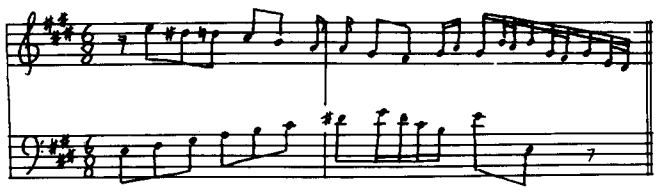
a second tone close in pitch or time to the masking tone (see Moore 1982, for a good introduction to this topic).) The masking power of a particular note may be assessed by varying the intensity of nearby notes to the point when they can only just be heard. This intensity is called the auditory threshold. The auditory thresholds of notes close in pitch to the masking tone are considerably raised by the simultaneous presence of the masking tone. However, as one moves the test tone further away from the masking tone in pitch there comes a point where the masking tone no longer has any effect on the threshold. One hears the test tone just as easily with the masking tone as without it. For any particular tone it becomes possible to estimate its 'critical bandwidth'. This is the frequency range within which it has a masking effect on other tones. Estimates of the critical bandwidth at various frequencies have been provided by Scharf (1970) and others. Table 5.1 shows the kind of results obtained, translating frequency values into approximate musical intervals.

TABLE 5.1
*Sample masking ranges of musical tones (adapted from Scharf 1970)*

| Approximate frequency of masking tone (Hz) | Musical note name of masking tone | Average musical interval from masking tone within which threshold of another tone is raised |
|---|---|---|
| 110 | A2 | octave |
| 220 | A3 | perfect fifth |
| 440 | A4 (concert A) | major third |
| 880 | A5 | major second |
| 1660 | A6 | minor second |

These rough estimates show that within the range of normal music, low tones mask over wider musical intervals than do high tones. In the range of a double bass, two octaves below middle C, tones as much as an octave apart will mask one another. In the range of a piccolo, two octaves above middle C, tones as close as a major second apart will not mask one another. In a thick contrapuntal texture there is real reason, then, to suppose that upper notes are more audible than lower ones. In a reverberant environment, the perception of low notes is also hindered by their greater tendency to persist in the environment after being played, creating a 'muddy' low-pitch noise, which makes lower parts less discriminable. Finally, we have to consider the fact that crucial acoustic information may be carried, not by the fundamental frequency of a musical note, but by its weak upper partials or harmonics. For low-pitched instruments these partials are likely to be masked by the higher pitched instruments, so making them less distinctive. All of this cannot, however, be the whole story, since I at least experience a strong 'set' for the upper line in simple two-part 'staggered' counterpoint where the effects

of masking and acoustic environment are not significant (e.g. Example 5.18, Two-part invention in E by J. S. Bach). It seems likely that experience with normal music, where the upper parts are known to be more important, on the whole, has resulted in a learned disposition to attend to the upper voice when there are no strong counter-signals.



Example 5.18. From J. S. Bach, Two part Invention in E.

A second type of characteristic of melodic lines which capture focal attention is a feature which discriminates them from the other lines in some way. Thus, a line may be louder than other lines, or have a distinctive 'solo' timbre. It may introduce change against an unchanging background. For instance, in Example 5.7 we tend to focally attend to the lower pitch stream which moves, rather than to the upper one, which is stationary. For the same sort of reason, the upper line of Example 5.17 invites attention because the lower line simply repeats the same figuration over and over again.

A third and related device is *change* of quality or texture in the focal line. This exploits a natural tendency, sometimes called the 'orienting response', to attend to a new event in a complex environment rather than to familiar ones. One principal way of introducing change is to add a new voice in a distinct pitch stream to a polyphonic texture. Successive entries of a fugue fulfil this condition.

Fourth, we may postulate a principle of 'attentional conservatism' such that a listener will tend to remain with a particular line once drawn to it unless there are strong enticements for him to shift attention. Thus, the composer can increase the likelihood of keeping a listener on the right line by avoiding sudden change in other parts.

Of course, composers rarely use such devices in isolation. An important part may be topmost in pitch, and loudest, and the most novel of the parts. The scientific evaluation of the relative effectiveness of these devices and the precise conditions under which they apply awaits detailed study of controlled musical examples where each attentional device can be manipulated separately.

## 5.4 Memory in music listening

The way one hears music is crucially dependent upon what one can remember of past events in the music. A modulation to a new key is heard only if one

remembers the previous key. A theme is heard as transformed only if one can remember the original version of which it is a transformation. And so on. A note or chord has no *musical* significance other than in relation to preceding or following events. To perceive an event musically (that is, to recognize at least part of its musical function) is to relate it to past events. It is, therefore, important for us to know how good we are at remembering past musical events, and to know what factors assist our memory.

We may begin by considering a very thorough series of experimental studies by Deutsch on memory for individual notes (Deutsch 1970, 1972, 1973; Deutsch and Feroe 1975; and others summarized in Deutsch 1982c). The basic experimental task is simple. Listeners hear two notes, separated by a five-second interval. They are required to judge whether or not the notes have the same pitch. In fact, the notes have the same pitch on half the trials On the other half they differ by a semitone. The interest of these studies lies in the ways in which Deutsch filled the five-second interval. In the basic control condition the interval was silent. Most listeners were then able to perform the pitch judgement with 100 per cent accuracy. In another condition, the interval was filled with spoken numbers, either to be recalled or ignored. In both cases listeners were still able to make the pitch judgement with 100 per cent accuracy. Another, crucial, condition interpolated randomly chosen notes between the two test notes. These interpolated notes were drawn from the same octave as the test notes. In this condition performance dropped to 68 per cent *even though* listeners were instructed to ignore the interpolated notes.

This result shows that intervening notes have a highly disruptive effect on memory for pitch of an earlier note. This is not some *general* deficiency in perceiving or remembering sequences, since the performance was not affected by intervening numbers, even when the numbers had to be recalled. It seems that the effect is specific to pitch.

In further conditions Deutsch discovered that the disruptive effect was greatest when one of the intervening notes was close in pitch to the test notes. Maximum disruption occurred when an intervening note was two-thirds of a tone above or below the first test note. The disruptive effect also generalized across the octave. The same pattern of results was obtained if the intervening tone sequence was shifted up or down an octave.

At first sight, Deutsch's results suggest a very gloomy conclusion about musical memory. Memory for individual pitches seems incredibly poor, if it cannot survive a few succeeding notes. How is it possible to remember notes across structures of symphonic proportions, containing tens of thousands of notes? The general answer to this problem would seem to lie in the opportunities which most music affords for listeners to classify and organize what they hear. Deutsch's sequences were atypical in two respects. They did not confine themselves to the intervals of a common scale (using fractions of a semitone in some instances), and their notes were randomly chosen so that

they were not designed to form common musical patterns within the scale framework. To see how musical memory is possible we will have to consider the two aspects of organization that Deutsch's materials violated—scale and sequential pattern—and show how they may be used by listeners to construct representations of music which do not rely on the preservation of precise pitch information.

### 5.4.1 *Scale*

Within a musical culture such as our own, there is general agreement about the scales from which music is constructed. In our culture the predominant scales are the diatonic major and minor scales. Furthermore, there is general agreement about the absolute pitches which scales should be based upon. This is so that a correctly tuned instrument can be played in ensemble with any other instrument. Thus A4, the traditional 'concert A' used for orchestral tuning is set at 440 Hz. With equal temperament instruments, such as piano, it is possible to set each of the twelve chromatic notes to definite pitches which allow all twelve major and minor diatonic scales to sound 'in tune'.

Given this 'fixity' of musical pitches, the possibility arises that music listeners could learn to associate particular pitches with musical note names. If they can do this, then it is possible that memory could be improved through the recoding of sensory pitch information into verbal information—remembering note names rather than raw sounds.

People with the ability to name individual heard pitches, or accurately sing named pitches are said to possess 'absolute' or 'perfect' pitch (see Ward 1963a, 1963b; Ward and Burns 1982, for reviews). This ability seems to be learnable by anyone prepared to undergo lengthy and systematic training (Brady 1970; Cuddy 1970). However, by no means all musicians have absolute pitch (AP), and a study by Sergeant (1969) shows that, among a sample of 1156 professional musicians, there was a high inverse correlation between age of commencement of musical training and the possession of AP. Almost all musicians who began training before the age of six had AP, but almost none of those who began after eleven did.

An elegant study by Siegel (1974) demonstrates the kind of memory advantage that absolute pitch confers. Her subjects were presented with two notes, separated by a silent five-second interval. These notes differed in pitch by either one-tenth of a tone or by three-quarters of a tone. Subjects were asked to judge whether the second note was higher or lower than the first. Siegel used two types of subject: music students with AP; and music students of an equivalent standard without AP. Siegel found that whilst the two types of subject did not differ in performance on the one-tenth tone intervals, the AP subjects did much better than the non-AP subjects when it came to the three-quarter tone intervals. The AP subjects were able to assign different verbal labels (note names) to the two notes. When the notes were only

one-tenth tone apart such verbal coding no longer worked, because it provided the *same* name for both notes. Thus they had to fall back on a 'sensory' coding which the other subjects reported using all the time.

Verbal coding aids retention over longer time spans, as shown in a second experiment where the time interval between the two test notes was varied. The interval was now filled with other random notes at the rate of four per second, making the task very similar to the one in which Deutsch found such poor performance. For the one-tenth tone test interval, performance of both types of subjects dropped very rapidly as the time interval was increased. However, for test intervals separated by a semitone, only the performance of the non-AP subjects dropped. The AP group were just as good after 15 seconds as they were after 5 seconds.

These results not only demonstrate the superior memory performance exhibited by subjects with AP. They also demonstrate the categorical nature of absolute pitch (see also Chapter 2.4). It is not the case that people with AP have finer pitch discrimination than anyone else. They cannot remember pitch within one-tenth of a tone over a five-second silent interval with any degree of accuracy. What they *can* do is assign the pitch to a *class* or *range* of pitches which are given the same name. Within a name category (spanning perhaps 30 Hz in the range of A4) they cannot remember pitch differences reliably. Furthermore, the results show how subjects with this categorizing ability can overcome the disruptive effect of intervening notes demonstrated in Deutsch's subjects, and by the non-AP subjects in this study.

Do these results suggest that people with absolute pitch have the best equipment for listening to and remembering music? If they do, then all of us without AP should, maybe, strive to obtain it. In fact, there are other types of ability which can be even more important in most tonal music. These include the ability to identify musical intervals and key. Of these, more shortly. AP would seem to be most advantageous in situations where music, whilst using the notes of the chromatic scale, violates the normal rules of harmonic and melodic construction, producing unfamiliar intervals and sequences where no key or tonal centre is implied. Much music in contemporary 'atonal' idioms would seem to fulfil these criteria; and those musicians particularly singers, who need to work with such music, have their task made much easier if they possess AP. There are situations, however, in which AP can be a nuisance. Listening to, or performing key-transposed music can be very difficult. The sounds produced are constantly violating expectancies created by one's own prior knowledge of the music or by the printed score. When scales or tunings are non-standard then it is possible to devise experimental tasks in which possessors of AP are actually at a disadvantage, because their categorizations are inappropriate to the task (e.g. Cuddy 1977, reported in Ward and Burns 1982).

Many musicians possess an ability which they prize much more than absolute pitch, and this is often called 'relative pitch'. This is the ability to

say what musical interval separates two simultaneously or successively played notes, or to be able to sing a named interval above or below a given starting note. This ability is important because melodic or harmonic identity is conferred on a set of notes through the interval relationships they hold to one another, and not their absolute pitches. A melody can start on any pitch within audible range and still be recognized *as* the same melody provided its pitch intervals (frequency ratios) are the same. This feature of music offers a possible explanation for why it is that relatively few people with the appropriate opportunities learn absolute pitch. When a child is taught a song, he is reinforced for producing the correct intervals, not a particular set of absolute pitches. He hears songs sung at various pitches (e.g. father sings the same song in a lower register than mother) and in general has every incentive to record interval information at the expense of absolute pitch (Pick 1979).

Siegel and Siegel (1977a) have demonstrated that relative pitch improves with training, and that highly experienced musicians can identify intervals reliably, accurately, and independently of stimulus context. Furthermore, musicians report hearing intervals as sounding qualitatively different from one another. Thus, a major third comes to sound not so much 'larger' than a minor third as a different *type* of sound. Each interval acquires its own unique character or colour that makes it unmistakable and not confusable with any other interval. Just as absolute pitch is 'categorical' so is relative pitch. Two notes do not have to be separated by the *exact* correct pitch interval to be heard as a major third, or whatever. Rather, their separations must fall within a *class* of separations about a central point. Categorical perception of intervals was demonstrated by Locke and Kellar (1973; see Chapter 2.4). It has also been demonstrated by Siegel and Siegel (1977b) who asked subjects to judge the magnitude of various standard and non-standard intervals. Although only 23 per cent of the intervals were actually 'in tune' (on just temperament tuning), subjects judged 63 per cent of them to be in tune, judging 'sharp' and 'flat' intervals to be equal in magnitude to correctly tuned intervals. There were even cases where subjects judged a sharp interval to be flat, and vice-versa. In this case it appears that they noticed a mistuning, but were unable to retain the precise pitch information which would allow them to tell whether the interval was sharp or flat. This loss of precise pitch information has many musical uses. For one, it allows the existence of the equal-temperament chromatic scale, which slightly distorts many of the 'perfect' small integer ratios of the harmonic series. These distortions are not noticeable under normal conditions, and provide musicians with the advantage of being able to play music which modulates round the circle of fifths without sounding more and more out of tune; unlike just intonation, where intervals in keys remote from the originally tuned key sound grossly mistuned. A technical discussion of tuning and temperament is beyond the scope of this book, and the interested reader is referred to Barbour (1951).

Just as absolute pitch aids memory under certain conditions, so does

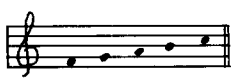
relative pitch. Idson and Massaro (1976) have provided one demonstration of this. Subjects were trained to recognize three-note sequences derived by permuting the notes A, C, and D sharp in different orders. The notes were given arbitrary labels (A, B, and C respectively) and subjects were required to identify a sequence as one of six possibilities, ABC, ACB, BAC, CAB, CBA. When they had learned to do this task accurately a 'masking' condition was introduced, where each note of the sequence was followed by an irrelevant note (e.g. AMBMCM, where M is the masking note). This masking note could be either in the same octave as the test sequence or in another octave. Idson and Massaro found that a same-octave mask disrupted identification of the test sequence, but that identification performance remained high if the mask was in another octave. In a second condition, subjects learned to identify not sequences but single notes. There were three notes, each given an arbitrary name (A, B, and C). Subjects learned to say which of the three notes had been presented on a given trial. Then they had to continue the task with various masking notes introduced, as in the first condition. It was found that *any* masking note, and not just one close in pitch to the test note, was highly disruptive of identification.

It is important to note that the first 'sequential' test could be done on the basis of relative pitch information, whilst the second task requires absolute pitch memory. We may interpret the results by saying that interference is more likely to disrupt comparison of present with past notes on the basis of absolute pitch than on the basis of relative pitch. When subjects are looking out for learned pitch intervals or contours they can perform the necessary perceptual and memory operations provided that interfering material does not fuse with the test material into a single pitch stream (see this chapter, section 2). When, however, subjects are trying to hold a particular absolute pitch in memory for comparison with later material, they are disrupted by *any* interference.

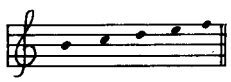
Idson and Massaro's experiment involved recognizing a stimulus which, although containing interval information, also retained the absolute pitch of the original. A common occurrence in music is the *transposition* of a melodic theme to a new key within the same piece of music. If we do, indeed, remember sequences of notes in terms of pitch interval, we should be able to recognize transpositions *as* repetitions of the same melodic stimulus. Experiments by Cuddy and Cohen (1976) and Cuddy, Cohen, and Miller (1979) have shown that both musically trained and untrained listeners can discriminate true transpositions of three-note sequences from those where one of the notes has been raised or lowered by a semitone relative to the other notes. This ability is, however, greatly affected by the precise nature of the sequence to be transposed. For instance, Dowling (1978) has shown that it is difficult for listeners to discriminate between 'real' and 'tonal' transpositions (see Chapter 2.4). Bartlett and Dowling (1980) have shown that discrimination of real from false transpositions is best when the transposition is

to a key closely related to the original. Cuddy, Cohen, and Miller (1979) have shown that if the sequence to be transposed is embedded in a tonal context, it is easier to recognize than when in a non-tonal context. All these findings suggest that memory for pitch intervals alone does not account for the way in which listeners deal with tonal sequence.

Much of the evidence points to the importance of establishing a key, or tonal centre, for memorization of melodic sequences. This provides an economical way of representing a melodic sequence in terms of scale steps within a key. An example will help make this clear. Consider the sequence in Example 5.19. Viewed as a set of successive intervals this could be seen as



Example 5.19.



Example 5.20.



Example 5.21.

the ascending sequence tone–tone–tone–semitone. If we consider intervallic relations between non-adjacent notes, we would add minor third (A–C), major third (F–A and G–B), perfect fourth (G–C), augmented fourth (F–B) ,and perfect fifth (F–C). Viewed as components of a scale, however (in this case C major), we could see it as four ascending adjacent intervals of one scale step starting on the fourth degree of the scale. The intervals between non-adjacent notes are then simple multiples of a single scale step. If we hear melodies in this second way, then we can argue that shifts of a melodic contour up or down *in the same key* are likely to be heard as 'the same melody' because a most important aspect of the melody—scale-step interval—is retained. If Example 5.19 is to be described compactly by the notation:

Key: C major
Start: 4
Sequence: +1, +1, +1, +1

then Example 5.20 is described in exactly the same way, except that the start of the sequence is now note 7 of the scale (B):

Key: C major
Start: 7
Sequence: +1, +1, +1, +1

Describing Example 5.20 in terms of its pitch interval relations, however, gives a very different pattern to that of Example 5.19. The intervals between successive adjacent notes are now semitone–tone–tone–semitone. Among the non-adjacent intervals there is no longer a perfect fifth or an augmented fourth, but a diminished fifth is added. The fact that people find such sequences (called 'tonal answers') as similar to the original as 'real' transpositions (which preserve the actual pitch intervals) is strong evidence for

viewing listeners' representations of tonal sequences as embodying the kind of notation given above. Such notations allow the economical coding of real transpositions too. Here, instead of changing the starting note, we change the key. Example 5.21 is a real transposition of Example 5.19, and in our notation would be represented:

Key: F major
Start: 4
Sequence: +1, +1, +1, +1

The foregoing offers a possible explanation for why real and tonal transpositions are often confused. Can it also explain why transpositions to related keys are easier to recognize than transpositions to remote keys? I think it can, if we look more closely at the nature of the experimental tasks which gave rise to this finding. A listener hears one short melodic sequence, and then, without any intervening preparation, hears another transposed sequence. Unless the key of the transposed sequence is close to the initial key in the circle of fifths it is possible that a listener is simply unable to decide a key (or tonic) for the new sequence. If he cannot do this, then he will not be able to assign the correct scale-step intervals to the new sequence. Once a listener has established a key, there is an inherent conservatism which makes him reluctant to shift key (see Chapter 2, section 5.3). Key modulation in much music is effected through movement around the 'circle of fifths', from C to G, G to D, D to A, and so on (or in reverse). A feature of such modulation is that the scales of the two keys share all their notes except one, the leading note, or seventh of the new scale. Thus G major contains all the notes of the C major scale except F which it replaces with F sharp. This has two consequences for transposition. One is that many transposed sequences can be successfully represented in two ways, either as resulting from a change of key, or as resulting from a contour shift within the same key For instance, the sequence CEGAG may be represented as:

Key: C major
Start: 1
Sequence: +2, +2, +1, −1

When the transposed sequence GBDED is heard, it can then either be represented:

Key: G major
Start: 1
Sequence: +2, +2, +1, −1

*or*

Key: C major
Start: 5
Sequence: +2, +2, +1, −1

This double coding means that a listener has *two* chances of hearing the transposition: crucially, it does not rely on him shifting his key frame.

The second consequence of the close relationship between keys like C and G is that the introduction of the one note *not* shared by the two keys is a strong and unambiguous cue for a key shift to the key which requires the least change of component notes to accommodate the new note. Thus CE DBC, becoming G B A Fsharp G, is strongly suggestive of a modulation from C major to G major. The first three notes of the second sequence (GBA) perform a vital pivotal function. They link the two sequences in a 'tonal space'. Probably most listeners hear these three notes as continuing in the original key of C. Then the F sharp forces a motivated key-shift one step up the circle of fifths. Consider in contrast the impression of CEDBC followed by F sharp A sharp G sharp E sharp F sharp. Here, the first notes of the transposed sequence cannot fulfil a linking function, since none of them are in the scale of the previous sequence. This has the effect of casting the listener adrift from his tonal bearings. He has to forget all about the first scale, and try to construct a plausible scale for the second sequence from scratch. His immediate memory of the first scale may interfere with this attempt, and his efforts to interpret the second sequence may cause him to lose details of the first. It is not surprising that recognition of such transpositions is hard. In much real music, two statements of a theme in distant keys will be separated by the kind of preparatory modulations that allow a listener to keep his tonal bearings and so know in advance what scales are likely to be encountered. In such cases, the transposed theme is recognized with ease.

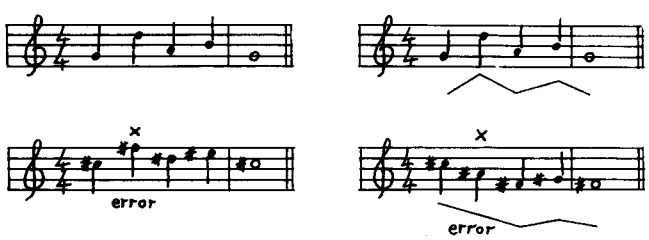
There is now a small body of research on the mechanisms by which key can most readily be detected by a listener. Some suggestions have already been discussed in Chapter 2, section 5.3. Butler (1983) reports experiments showing that one cue signalling key is the presence of intervallic patterns which are unique to that key, or which can be found in the smallest number of other keys. The interval C–D, for instance, can be found in *five* major scales—C, G, F, B flat, and E flat. On the other hand the tritone F(E sharp)–B is only found in two scales, C major and F sharp major. In Butler's experiments sequences which made this tritone relation most salient were those which gave listeners the strongest sense of a key centre. A third note added to the tritone specifies a unique key. For instance, the sequence FBC can occur only in C major (Brown and Butler 1981).

Once a key is established, each note of the implied scale has a unique configuration of intervallic relationships with the other notes of the scale (because of the uneven distribution of tone-semitone intervals up the scale—see Balzano 1980). Thus, the interval F–C *heard as* in C major has a different tonal quality to F–C *heard as* in F major. In the former case the notes have an implied scalar relationship to a B *natural*, whereas in the latter case they relate to a B *flat*. To hear F–C as in C major is to appreciate the relationship of these notes to an unheard but 'implied' B natural.

These considerations help us to understand why context is of such crucial importance in the recognition of transpositions. In Cuddy, Cohen, and Miller's experiments (1979), the diatonic context helped listeners identify the key of each passage. When the context was non-diatonic it was impossible to assign a key to each passage, and thus impossible to code the sequences in terms of scale steps. Probably the best that could be done would be a rough estimate of interval contour (e.g. 'a large interval up followed by a smaller one down'). Such information would be too crude to be useful in distinguishing exact transpositions from 'false alarms' with similar up-down contours.

Perhaps, however, this 'crude' up-down contour should not be dismissed as playing no role in musical memory. Davies (1978) and Dowling (1978) among others have proposed that such contour information *does* form the basis of our musical memory in some real listening situations, and can be useful even when more exact scalar or intervallic representations are not achieved. Such information could certainly not allow us to *recall* a melody other than very approximately, but it might indeed be sufficient, in some contexts, for a melody to be *recognized* as a repetition or transposition of an original (especially if combined with other information, such as, for example, rhythm).

Edworthy (1983) has attempted to assess the importance of contour memory by means of an experiment where listeners heard two melodies, the second of which was a transposition of the first at the tritone (e.g. C major–F sharp major). The transposition preserved the exact intervals of the original except at one point where a deliberate alteration was introduced. Listeners were required to locate the alteration. There were two types of melody, short (5 notes) and long (15 notes). Example 5.22 gives an example of a typical five-note melody. There were also two listening conditions. In the 'pitch' condition, the altered note was another note chosen from the key of the melody which did not change the original up-down contour (see Example 5.23). Subjects were asked to monitor the pitches of the notes to detect the alteration. In the 'contour' condition, the melody was altered within its key so that the up-down contour was different at one place (see Example 5.25). Subjects were asked to monitor the up-down contour, *not* the precise pitches, so as to locate the contour alteration.

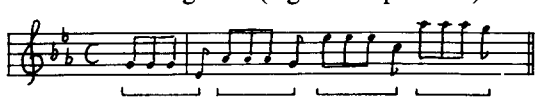



Examples 5.22.–5.25. Examples of stimuli in Edworthy's (1983) study.

Unsurprisingly, the short melodies gave more accurate results than the long ones. The important result, however, is that although the contour condition gave the best results for short melodies, the *pitch* condition gave the best results for long melodies. It seems that as soon as melodies of any length are heard, up-down contour coding becomes inefficient. Subjects must switch to a more difficult form of coding which retains pitch interval information in some form. In the light of the preceding discussion it seems that this processing involves the coding of the melody as a sequence of tonal elements in a particular scale or key.

Dowling and Bartlett (1981) provide similar conclusions to Edworthy in a study where short extracts from Beethoven string quartets were played. Subjects had to classify other extracts from the same work, heard after a delay, as similar or dissimilar on a rating scale. Some of the extracts were 'lures' chosen to resemble the originals in melodic contour but to differ in exact pitch interval, others were dissimilar in both contour and pitch interval. Subjects judged both these types of extract as 'dissimilar' to the original; with the same-contour lures being heard as only marginally more similar than the different contour lures. In further experimentation, Dowling and Bartlett found that shared contour was a better predictor of perceived similarity only when the delay between original and comparison was short. When the delay was longer, an exact pitch-interval match was required for high similarity judgements.

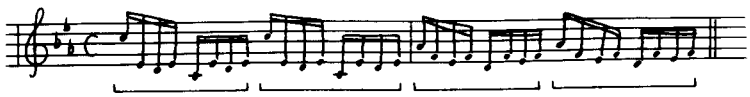
Taken together, these results suggest that contour is a characteristic of music which is most important for listening over short, rather than long spans; and that if a composer wishes to direct a listener's attention to the *contour* similarity of two sequences they must both be rather short (nearer 5 notes than 15) and be close together in the composition. The cases that spring to my mind certainly fulfil this condition. In the first movement of Beethoven's Fifth Symphony, the crucial melodic element is four notes long, and when pitch interval is altered so as to leave only contour identity then the repetitions are often contiguous (e.g. Example 5.26). In Bach's Prelude



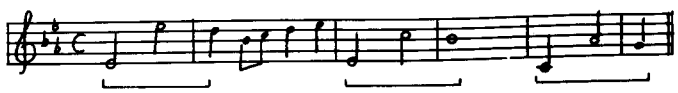
Example 5.26. Adapted from Beethoven, Fifth Symphony, first movement.

in C minor (Well-Tempered Clavier, book 2), the repeating motif is eight notes long, and because virtually the whole piece is based on contiguous repetitions of the motif which preserve contour but not precise interval relations, the contour is highly salient (Example 5.27).

Consider, though, how even a small amount of intervening material can camouflage a contour identity. The popular song 'Over the Rainbow' has a three-note motif which is subject to contour repetition (Example 5.28). It was not, however, until I was encouraged to view this very familiar melody



Example 5.27. From J. S. Bach, *Well-Tempered Clavier*, Prelude 2, Book 1.

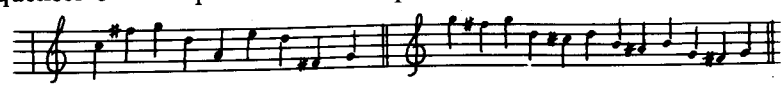


Example 5.28. Opening melody bars of 'Over the rainbow', Music by Harold Arlen, lyric by E. Y. Harburg. Copyright © 1938, 1939 (renewed 1966, 1967) Metro-Goldwyn-Mayer, Inc. All rights controlled by Leo Feist, Inc., a catalogue of CBS, a division of CBS Inc. All rights reserved. Used by permission.

analytically (by Rosner and Meyer 1982) that I noticed the contour repetition of bar 1 by bar 3. My 'parsing' of the melody had grouped the third note (D) with the other notes in bar 2, undermining its relationship to the previous two notes.

### 5.4.2. *Sequential pattern*

It seems incontrovertible that musical sequences which can be represented as occurring within a key or scale are easier to remember than those which cannot. However, this generalization needs qualification. Consider the two sequences of Example 5.29. If one spends a few seconds studying each of



Example 5.29.

these with a view to reproduction it will soon become apparent that (*b*) is much easier to remember than (*a*). This is despite the following facts:

1. (*a*) begins with a three-note sequence which uniquely specifies the key as G major, whereas (*b*) does not.
2. (*a*) contains only notes from within the G major scale, whereas (*b*) contains two chromatic notes (C sharp and A sharp).
3. (*b*) is longer than (*a*).

What makes (*b*) easier to remember is that it seems somehow simpler. One hears it as a three-note pattern which repeats four times down a G major arpeggio. Deutsch and Feroe (1981) have provided a detailed formalism which captures some features of music which appear to accompany ease of memorization. They see simple tonal melodies as operating within *alphabets* of notes. An alphabet is a culturally specified pitch-ordered set of intervals which embodies some tonal principle. The primary alphabets of our culture are the diatonic and chromatic scales and arpeggios. Taking any one of these alphabets one can construct an ordered ascending or descending sequence from any note. Applying the alphabet of the G major scale to the note G would give the ascending sequence G A B C D E, etc. Applying the alphabet

of the chromatic scale to the same note would give G G sharp A A sharp etc. Applying the alphabet of the G major arpeggio would give G B D G etc. It is proposed that within each of these alphabets, moves of a single step are the simplest and most easily manageable type of change, and that sequences which can be shown to be constructed out of a combination of such moves in the various alphabets are perceived as simple and easily memorable. For instance, Example 5.29 (*b*) can be expressed as a simple hierarchical structure based upon single step movements in the alphabets of the G major arpeggio and the chromatic scale (see Fig. 5.5). At the top of the hierarchy we start

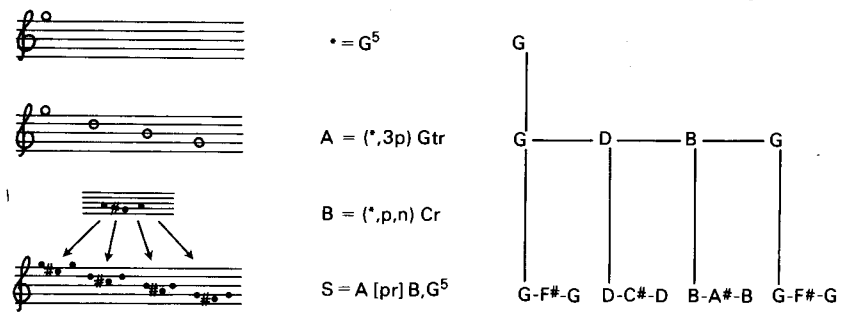



Fig. 5.5. Formal representation of an hierarchically organized tone sequence. Adapted from Deutsch (1981) and Deutsch and Feroe (1980). Key: * = reference note (i.e. the note on which the sequence is constructed); p = downward movement of one step within an alphabet; n — upward movement of one step within an alphabet; Gtr = alphabet of the G major triad; Cr = alphabet of the chromatic scale; A[pr]B = the application of pattern B to each element of pattern A.

with a single note, G4. We than apply the G major arpeggio alphabet to this note, descending in consecutive steps for four notes. We then construct a subsidiary pattern in the chromatic scale which is obtained by taking a note, moving one scale step down, and then returning to the original note.

We now apply this pattern to each note of the arpeggio in turn to produce the final melody. Figure 5.5 shows the consecutive steps in the construction, and also shows the symbolic notation that Deutsch and Feroe have developed to describe such sequences economically.

These theoretical formulations build directly on work by Simon and Sumner (1968) and Restle (1970) concerned with temporal patterning. However, Deutsch and Feroe point out that similar hierarchic structures seem to be implicated in the way we represent other complex inputs such as visual scenes (Palmer 1977) and large-scale environments (Chase and Chi 1981) as well as language (see Chapter 2); and thus represent a general feature of human cognitive functioning.

Deutsch (1980) has provided first-hand evidence for the theory by asking musicians to recall melodies by dictation. Some melodies could be economically described as hierarchic structures [such as Example 5.29 (*b*)]. Others [like Example 5.29 (*a*)] could not. Recall for the former type was 94 per cent

correct; that for the latter was only 52 per cent. Furthermore, if the structured melodies were temporally segmented into groups of three (by the introduction of a brief pause after each third note) corresponding to the major hierarchical subdivisions, performance rose to 99 per cent. If, however, the segmenting was in groups of *four*, thus 'breaking up' the melodic groupings by threes, performance dropped to 69 per cent. On the other hand, temporal grouping of *any* sort improved performance on the unstructured sequences by about 10 per cent.

These results introduce an important factor which we have not so far considered in our treatment of music listening; that is the role of timing and rhythm. It is well known that introducing temporal gaps into sequences of any sort increases the tendency for sequences to be perceived as groups or chunks between each gap (Bower and Winzenz 1969; Restle 1972). This is, in part, to be explained by Gestalt grouping principles. The 'law of proximity' states that elements grouped closely together on a particular dimension tend to be perceived as single unit, separate from other, more distant elements. A visual analogy to the auditory phenomenon is provided by the following pattern:

.. .... .... .... .....

This is seen as five unequally sized groups.

However, as we have already seen (Chapter 2, section 5.4) such simple groupings can be supplemented by a more sophisticated type of grouping based on metre, and displaying much of the same hierarchical structuring as Deutsch and Feroe have proposed for pitch grouping. The operation of both types of grouping is well demonstrated in a study by Smith (1983) who examined memory for rhythmic patterns by asking musicians and non-musicians to listen to rhythmic sequences performed on a synthesized drum, and then attempt to reproduce them. Figure 5.6 gives an example of the type of rhythmic pattern used, together with two alternative ways of grouping the pattern. Above the pattern is a representation of the sound amplitude over



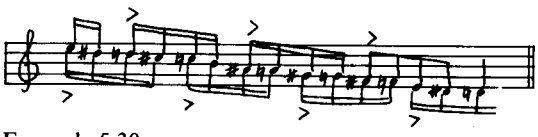
Fig. 5.6. Adapted from Smith (1983).

time. Gestalt grouping by temporal proximity yields the four bracketed groups given. Below the pattern is a representation of its hierarchical metrical structure. This assigns the sequence a 3-beat metre which groups notes *across* the boundaries of temporal groups. Each metrical group begins with a stressed note at the beginning of a bar. When Smith analysed her subjects' errors she found that musicians tended to group by metre, and that their responses tended to maintain the regular metre of the original. The errors of non-musicians tended not to retain the metre, but their responses preserved the more primitive temporal-proximity grouping. Smith found that

> Many subjects reported using a 'counting' strategy, tallying the number of sounds that formed a group, and the number of such groups in the sequence. This meant that although they might tap out the correct number of sounds, the intervals within and between groups of sounds were distorted.

We see, therefore, in Smith's study, not only the importance of relatively abstract underlying patterning in determining memory for rhythm, but also the fact that the use made of such patterning depends upon musical experience. This reminds us of the generally applicable, but sometimes overlooked, fact that many aspects of the ability to deal with music are crucially dependent on musical experience. Such experience allows the learning of the characteristics of common structuring principles in music, and mechanisms for detecting them. We shall return to a more thorough examination of the changes that take place in musical behaviour as a result of experience in the next chapter.

Several other studies confirm the importance of hierarchical organization in the perception of rhythmic patterns (e.g. Handel and Todd 1981; Povel 1981); and Jones, Kidd, and Wetzel (1981) have provided a preliminary demonstration that rhythm can act as a cue to attentional streaming in an analogous way to pitch. In general, rhythm has been 'tonality's poor relation' (Davies 1978) in studies of psychological response to musical structure. We are now beginning to see, however, that not only is rhythm just as important an organizing principle as tonality, but that the two systems are mutually interactive. Thus, in much tonal music, knowledge of the tonal structure can help determine the rhythmic structure, and *vice versa*. We have mentioned before (Chapter 2) how repetition of tonal patterns can act as a cue to metre (Steedman 1977). Thus Example 5.29(*b*) acquires a 3/4 metre as a result of the repeating three-note pattern. Conversely, rhythmic stress (as communicated through accent or timing variations—see Chapter 3, section 2.3) provides a cue for determining tonal structure. In much music, principal harmonic notes fall on the main beats, with subsidiary or passing notes at other points. Thus, the chromatic sequence of Example 5.30 is most likely to be heard in D minor if the upper accenting scheme is used, but in E minor if the lower scheme is used. In each case the accents specify the primary harmonic notes.



Example 5.30.

5.4.3. *Memory in extended music*

The literature I have reviewed in this chapter shows that we are beginning to build up a detailed, objectively verified, picture of the way listeners represent and memorize short melodic sequences. The question we return to, raised at the beginning of the chapter is, how does this help us to understand what goes on in listening to extended passages of music? One, rather naïve, view would be that listeners form representations in memory of each short sequence in the sort of way outlined here, and then simply retain these sequences in the correct serial order; in much the same way as one might try to remember a list of unrelated sentences. This view raises two problems: firstly, how does a listener segment or 'break up' a continuous musical texture into short sequences of the appropriate sort; secondly, how does he keep what may be many hundreds of such sequences distinct and correctly ordered in memory? We will examine these questions in turn.

To solve the segmentation problem, listeners must be capable of detecting 'closure' at various levels. Sometimes physical characteristics of the music, like a pause, or a change of instrumentation will suggest a suitable segmentation. But there is much music where such cues would not be sufficient. There is some evidence that listeners perceive more abstract types of closure, specified by harmonic or rhythmic considerations. In Chapter 2 we discussed the 'click migration' experiment of Sloboda and Gregory (1980) as providing evidence for perceptual segmentation of melodies into phrases separated by cadences. A study by Tan, Aiello, and Bever (1981) is more particularly relevant to the present discussion because it tested *memory* for parts of a sequence. In this study sequences of equal-duration notes were constructed which contained two melodic phrases, each phrase ending with a melodic cadence. Example 5.31 shows one of the sequences used. When played or sung it will give Western listeners the impression of breaking into two halves, a partial 'close' coming after the tenth note. The sequence leading up to this note (D G E C) is most readily interpreted in this context as a perfect cadence, moving from the harmony of G to that of C. Notes 11–14 form part of a new chord (D minor) leading to a final perfect cadence in C. Tan *et al.* played such melodies to subjects and then asked them to judge whether a particular two-note 'probe' had or had not been present in the melody. Of particular importance were three types of 'true' probe; the pair of notes ending the first phrase (E C in Example 5.31), the pair of notes beginning the second phrase (F A in Example 5.31), and a pair of notes which 'straddled' the phrase boundary (C F in Example 5.31). They found that subjects were less likely



Example 5.31. Example of stimulus used by Tan, Aiello, and Bever (1981).

to recognize the latter type of probe as having occurred than the other two. It appears that they were more likely to form accessible memory representations of intervals *within* phrases defined by cadences than they were for notes equally close in time but coming from two different phrases.

Closure can also be signalled by rhythmic or metrical features of music. For instance, there is a tendency for much melodic music to be constructed in multiples of two-bar units. Thus, a common phrase length is four bars: the 16-bar song contains four four-bar phrases. Meyer (1973) has proposed that a musical form achieves the double ends of articulation or segmentation together with unification and forward-movingness by putting the various cues to closure 'out of phase' with one another. Thus a melody may achieve rhythmic closure at a particular point, but its lack of harmonic closure implies the need for a continuation. On this argument, the end of a well-constructed piece of music is not simply the point where the sounds stop; it is the point where closure is simultaneously achieved at all the various levels. Let us then assume that there are cues which listeners can use to segment music into manageable and 'meaningful' short units.

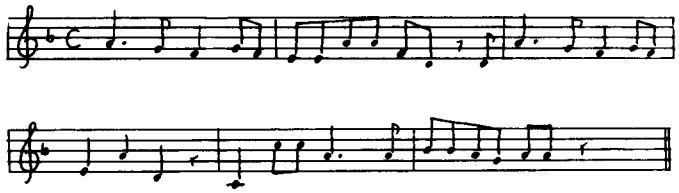
Our second question relates to the problem of holding a large number of such segments in memory. Estimates of memory capacity for a set of items which are not related in any principled way suggest a limit within single figures (e.g. Miller 1956). As more items are added, memory for other items is lost. The only way to overcome such limitations appears to be to find some way of linking or relating, the items together. In some spheres of memory this is done by 'importing' a structure which can act as a mnemonic to link otherwise unrelated elements together. In music, such relations are, to a large extent, already present in the patterning and structure of a composition. The composer deliberately writes so that the individual segments have a similarity and connectedness that links them together into larger units. It is precisely by discovering such similarities and connections that the limits on memory can be overcome.

First, for instance, economy of coding is achieved if repetitions can be identified and noted. Then, rather than making two representations of the sequence in question, it can be represented once in memory, and 'called up' on its various occurrences through the piece by some sort of 'marker' which locates the repetitions in the longer sequence. Second, inexact similarities could be noticed, and sections coded as *variants* of earlier sections. Maybe on a first hearing, the listener would not have the resources to code all the differences, but would be able to remember it as 'almost like' something that had gone before. Third, the listener could extract some underlying

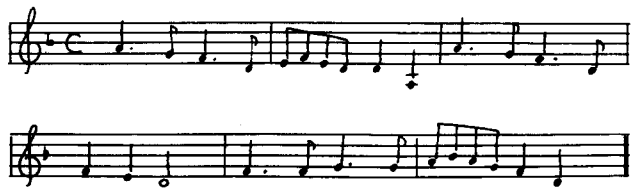
progression or pattern in a sequence of musical statements; such as traversing a familiar tonal path, or moving towards a climax, or conforming to some known form such as theme or variations. This would help him to reconstruct the correct ordering for the sequences, which would no longer be arbitrary, but guided by some rule or pattern. Fourth, the listener may be able to construct an emotional or representational 'plot' or 'drama' from the music, which he can use to help recall and order the various components. For instance, the requirements of the 'plot' at a particular point tells one that the next theme is 'cheeky' or 'resigned'. Delis, Fleer, and Kerr (1978) have provided evidence for this kind of process in an experiment where they found that people remembered musical extracts best when they were labelled with concrete representational titles as opposed to abstract conceptual ones. They argue that the concrete titles allow some kind of story to be constructed, to which the various segments of the music can be associated. In much music, there is, of course, no representational title, and so the listener must construct his own 'story' from the character of the music; but the principle is the same.

The general picture I would like to propose on the basis of such suggestions is that the listener is engaged in building up a multidimensional representation of the music he hears, and that depending upon his knowledge and cognitive style, his early memory for the music will select different dimensions from the many available ones. What is crucially important is that these memories will not necessarily be composed of isolated episodes in the music. It is equally possible that he will have extracted quite global information which specifies some parameters of the overall structure (such as metrical construction or harmonic framework, recurrent melodic types, or emotional 'argument').

How can we demonstrate such processes at work? I can, at present, think of no satisfactory way of tapping what goes on when music of really extended proportions is heard; but one intermediate approach, which David Parker and I have recently begun to explore, is to ask listeners to attempt to reproduce, by singing, moderately extended short musical forms which have a definite melodic phrase structure, but which have too many notes for a perfect first recall to be likely. We can then examine the errors made in reproduction, and also how a representation is built up over repeated hearings, and repeated opportunities to recall the melodies. Our preliminary findings are encouraging. Even people with quite little musical training are capable of providing competent responses; and we find that often the very first recall preserves much high-level information contained in the original whilst losing specific details. For instance, Example 5.33 shows a transcription of a subject's first attempt at recalling Example 5.32. Whilst many of the individual notes do not match the original, it has the correct metrical structure (three two-bar phrases in common time, the correct harmonic sequence (two I–V–I phrases in D minor followed by an implied move to the relative major key of F), and several of the characteristic melodic and rhythmic patterns of the



Example 5.32.



Example 5.33.

original. It seems very much as if this subject was reconstructing some of the details according to constraints laid down by his knowledge of the higher order structure and style of the melody, using general musical knowledge to 'fill in the gaps' in a plausible manner. Further discussion of this study maybe found in Sloboda and Parker (1985).

I hope that the preceding discussion will make it more apparent why I wish to claim that a 'feat of memory' such as Mozart's memorization of Allegri's *Misèrere* (see Chapter 1) does not involve inexplicable processes which set him apart from ordinary musicians. Rather, it distinguishes him as someone whose superior knowledge and skill allows him to accomplish something rapidly and supremely confidently which most of us can do, albeit less efficiently, and on a smaller scale. If we examine the structure of the Allegri *Miserere* we find several features which could assist memory processes of the kind we have been discussing. The work has a simple episodic structure in which a polyphonic 'chorus' is repeated several times, separated by a repeated simple and homophonic chant-like verse passage. Mozart would have had prior access to the words of this choral piece, and possibly to the reports of other listeners, which would have given him a pretty clear idea of the type of structure to expect. The primary problem for him would be to 'fix' the part writing of the 'chorus'. In this, he would be assisted by the several exact repetitions during the course of the performance. Within this section there is a reasonably straightforward harmonic sequence which supports a particularly salient and ornate soprano melody. It was the alleged beauty of this melody, heard in its harmonic and acoustic context, that was responsible for the refusal of the Vatican to publish the music. The fact that this music could only be heard in the Sistine Chapel no doubt added to its fame.

There are several strategies that Mozart could have used to memorize the crucial section. One would be to focus his attention on a different vocal part on each of the hearings (see this chapter, section 2). I suspect, however, that a priority would have been the exact replication of the celebrated soprano line together with the correct harmony. It would not have been so important that the disposition of the harmony notes among the inner voices should be absolutely correct. After all, there was no possibility of the 'illicit' copy being checked against the original score. With the primary melodic line and some of the more salient inner parts fixed, it is quite likely that Mozart could have used his stylistic knowledge to 'fill in' what he did not hear for certain. Furthermore, he would probably have had that 'instant' recognition of chords and intervals which we discussed in the context of 'relative pitch'. Thus he would know that he had heard, say, a dominant seventh chord, without necessarily hearing out the individual notes. Yet his knowledge of the contents of a dominant seventh would allow him to infer the notes 'in the mind's ear'. After a first hearing, he would have identified troublesome sections which needed more attention on a subsequent hearing, because they did not contain instantly recognizable chords or patterns. Finally, for someone with as much musical knowledge as Mozart, there exists the possibility that he recognized sequences as identical to, or similar to, sequences in other pieces already well-known to him. Once recognized, such a sequence can be disposed of as something already available in the memory representation of another work.

These comments, whilst speculative with respect to Mozart, are clearly testable with respect to contemporary musicians who can produce both written evidence of their memorization attempts and verbal protocols of their strategies. Let us hope that such research is undertaken soon. Research on other cognitive tasks (such as copy drawing from memory, e.g. Pratt 1983) offers promising models, as well as confirming the general kinds of explanatory proposal put forward here for music.

# 6. Musical learning and development

## 6.1. Introduction

Musical skill is acquired through interaction with a musical environment. It consists in the execution of some culturally specific action with respect to musical sounds. However, musical skill is constructed from a base of innate abilities and tendencies. Every human advance involves building on what is already present. Because all human brains are similar, and because certain aspects of experience are common to all humans (e.g. the experience of the physical world and its properties) it has been possible to isolate certain early aspects of skill acquisition which seem to be shared by all humans. For instance, we all acquire the basic skill of manipulating objects in the first year of life; we acquire competence in our natural language by the age of three or so. Further than this, it has been proposed by some psychologists, notably Piaget (e.g. 1950, 1952) that the *type* of learning we are capable of at any age is determined by the general features of our intellectual equipment at that age. Thus, the reason why children seem not to be able to reliably master mathematical concepts until the age of seven or so is that until that age they do not have the type of cognitive resources that would allow them to grasp necessary notions such as transitivity or conservation.

On this view cognitive development is to be partly explained in terms of the ordered acquisition of new 'general' cognitive abilities and structures ('general' because they are involved in a whole range of specific skills and not just one). Piaget's well-supported view is that there is a universally shared order of passage through various cognitive 'stages' and that each stage is characterized by a fairly rapid advance in skill acquisition as the new capacity is applied to the whole range of specific skills the child is engaged with. These stages should not be seen as mysteriously emerging in the child simply as a result of some biologically determined 'release' of new capacities at various ages. Rather, they are brought about by adjustments or 'accommodations' that occur as the child learns particular skills. When the time is right, these adjustments 'colonize' other areas of endeavour.

If we look at musical development through Piagetian eyes, we should be alert to the possibility of discovering invariant sequences of musical development, possibly linked to general changes in other cognitive domains. These sequences would account not for the precise aspects of musical behaviour found—this would depend upon culture, motivation and op-

portunity, but for the *types* of musical activities to be found at particular ages in virtue of the general cognitive capacities they require.

There is, however, another powerfully argued view of human cognitive capacity, associated particularly with the work of the linguist Chomsky (see also Chapter 2). This states that the human organism is biologically predisposed to excellence in a few *specific* cognitive skills, language being one such, and that there exist special mechanisms for acquiring these skills which do not necessarily form part of a 'general' cognitive capacity, but are confined (at least initially) to the skill in question. Taking this approach we should be alert to the possibility of discovering early aspects of musical development which point to capacities specific to music, and not mirrored in other domains.

These two views are not mutually incompatible, and Gardner and Wolf (1983) have pointed out that elements of both may be required to fully explain the course of cognitive development. They argue that human development is characterized by both separate *streams* of specific skill acquisition, possibly supported by special biologically determined mechanisms, and common *waves of symbolization*, where a new achievement in one stream 'spills over' into apparently unrelated streams. The streams are largely linked to, and supported by specific cultural roles—orator, singer, artist, craftsman, etc. Within each stream, explicit instruction or tuition is common. The waves, on the other hand, seem not to be culturally salient. No-one, until very recently, has explored the notion that general cognitive style could be changed through instruction independently of particular skills. Rather, the 'spillage' across skills seems to be a spontaneous human propensity. It might even be a defining characteristic of human thought that there exists some mental medium abstract enough to provided the means of passing general skill between two different specific skills.

It is not my intention to give an exhaustive account of developmental theories as they relate to cognitive psychology. General introductions to the work of Piaget are given by Flavell (1963) and Boden (1979); and a readable critical approach is provided by Donaldson (1978). Reviews of developmental theories in relation to music are provided by Funk and Whiteside (1981) and Shuter-Dyson and Gabriel (1981); the latter also being a comprehensive source reference for research into musical development. What I would like to do here is to use the broad theoretical framework I have outlined above to motivate the division of this chapter into two separate, but complementary, halves.

The first half follows the spontaneous acquisition of musical skill by Western children from birth up to the middle years of childhood. The second looks at the later development of specialized musical skills, which typically take place in a self-consciously educational milieu. These two processes may be roughly labelled by the terms enculturation and training respectively.

The main elements of *enculturation* seem to be these. First, we find a shared

set of primitive capacities which are present at birth or soon after. Second, there is a shared set of experiences which the culture provides as children grow up. Third, there is the impact of a rapidly changing general cognitive system as the many other skills supported by the culture are learned. These elements combine to yield a roughly similar sequence of achievements for the majority of children in a culture, and a set of roughly similar ages at which the various achievements occur. Enculturation is also typified by a lack of self-conscious effort and a lack of explicit instruction. Young children do not *aspire* to improve their ability to pick up songs, although they do improve. Adults do not *instruct* young children in the art of song memorization, yet children come to be able to memorize songs.

When we turn to *training* we take up a rather different set of concerns. We now concentrate on specific experiences which are not shared by all members of a culture. Rather, they are specific to the sub-culture where aspiration to excellence in a particular skill is encouraged. Such experiences allow a person to build on the general foundation of enculturation to achieve what we may call expertise. Very broadly, we may say that, in our Western culture, musical enculturation is the dominant process up to the age of about 10; thereafter musical training plays an increasingly important part. In general, it appears that training is more likely to contribute to depth of knowledge and accomplishment *within* a particular skill than it is to have broad implications for the whole cognitive system. Gardner's 'waves of symbolization' seem to be particularly potent during the first years of childhood when learning on all fronts is most rapid. At this time the cognitive system seems particularly susceptible to broad upheavals. In contrast the typical adult displays a highly 'skewed' profile of cognitive attainments. Someone may be a master jazz improviser whilst retaining the reasoning powers of a 10-year-old. Someone else may have a profound grasp of mathematics whilst retaining the musical accomplishments of a child. And so on. We see, therefore, a cutting of deeper channels within individual streams rather than general 'waves' of development.

Training also involves self-conscious effort on the part of the person concerned with the specific aim of becoming more accomplished. Typically he seeks, or is given, methods for increasing accomplishment. These methods are conveyed through instruction. As well as seeking to describe and understand the changes that take place during a particular type of training, the psychologist is also interested in evaluating the efficiency of different training methods, and, on the basis of such evaluation, offering prescriptions about the best way to train a particular skill. This evaluative/prescriptive approach is the basis of a branch of psychology known as educational psychology.

Over the past thirty years, relations between educational psychologists and cognitive psychologists have not always been cordial. Cognitive psychologists have viewed educationalists as unconcerned with understanding the psychological mechanisms which account for the success or failure of particular

training methods. They have caricatured the typical educational study as randomly allocating children to a number of different types of instructional regime and simply prescribing the instruction which gave the best results. On the other hand educational psychologists have seen cognitive psychologists as unconcerned with solving practical problems, and caricatured the typical cognitive study as dealing with some microscopic portion of a real task and hedging any possible prescription with so many qualifications as to be useless in a real training situation. Both these views have more than an element of truth in them, and it remains as difficult as ever to make instructional prescriptions which are soundly based on real understanding of the psychological processes involved. None the less, I believe that both sides are coming to acknowledge the importance of trying to make links between theory and practice; and writers of cognitive texts now find it almost obligatory to derive some training prescriptions from the results of pure research (e.g. Anderson 1980; Bransford 1979).) Particularly, it has been argued in various places (e.g. Levin and Williams 1970) that an intelligent teacher is more likely to benefit from a greater understanding of the psychological processes underlying a skill than he or she is from being given a particular training prescription. The increased understanding will allow the teacher to devise his or her own methods to suit particular pupils and situations and to modify them in principled ways.

This view has been supported from my own experience with teachers. For instance, I once asked an internationally known performer and teacher to take part in a sight-reading experiment. This was the one described in Chapter 3, section 2.1, where eye-hand span was estimated for 'tonal' and 'atonal' melodies. After the experiment had taken place, I discussed with the teacher my prediction that span would decrease on the atonal melodies. He was sceptical, on the grounds that he specialized in the performance of atonal music. He felt that his immersion in atonal music would counteract the possible effect of my mildly atonal stimuli. Moreover, he had not *experienced* the atonal melodies as any more difficult to read than the others. Later, I analysed the experimental data and found this player to provide results indistinguishable from other musicians whose staple diet was Bach and Beethoven. Like them, his span decreased for the atonal melodies. When I wrote to him with his results he wrote back saying that the results had forced him to change his attitudes to the teaching of sight-reading. How he put his changed attitudes into practice I do not know; but clearly there are any number of educational implications in this result, showing as it does that tonal relationships have an inescapably privileged status in the mental apparatus of a musician (presumably through their early and continued presence in the casual experience of all Westerners). However, I would not be happy to derive a single prescription for teachers from this information. I would expect the competent teacher to be able to make the best use of such information to suit his or her particular circumstances.

Educational psychology has a second major concern besides evaluating training methods, and this is the devising of *assessment* procedures which allow teachers to discover where a particular individual stands with respect to the average achievement for his age on some standardized test. In music such tests have often been labelled 'tests of musical aptitude'. One use of these tests is diagnostic, to isolate particular weaknesses which need extra training. Another use is predictive. High scores on particular tests seem to correlate with later achievement, and so, for instance, music schools use such tests to select candidates for entry to specialist training courses. Such tests are often designed with the explicit or implicit aim of detecting potential excellence in the absence of specific training. We shall examine to what extent they are capable of doing this.

## 6.2. Musical enculturation

### 6.2.1. *The first year of life*

The first evidence of musical awareness is some form of differentiation of musical sequences from one another or from non-musical sound sequences. It is not adequate to say, as many writers do, that heightened attention to particular sounds, as such, is evidence of musical awareness. It is well known that infants are particularly responsive to change in the environment (e.g. Kagan and Lewis 1965). It follows that any new or unusual sound will capture a young baby's attention. Parents who are delighted when their babies become quiet and attentive to the sound of a nursery rhyme should be wary of concluding too much about their abilities. It is possible that a child is simply responding to a change in auditory experience from the rapid pitch and amplitude modulations found in speech to the steadier parameters of song, or that he is particularly responsive to definite types of waveform (e.g. Hutt, Hutt, Lenard, Bernuth, and Muntjewerft 1968).

It is, indeed, a necessary precursor of musical awareness that a child should be able to notice differences in crucial dimensions of sound. For instance, he will be unable to notice the defining features of a particular melody if he cannot detect differences in pitch or time. None the less, real musical awareness begins only when the child is able to notice sequential relations between different sounds. How good, then, are young babies at detecting sequence? Studies by Chang and Trehub (1977a, 1977b) suggest that babies as young as five months are already sensitive to sequential structure. In one study (1977a) they repeatedly presented a six-note atonal melody to babies. They measured the perceived novelty of the melody by monitoring heart-rate. It is now well established that changes in heart rate are reliable indicators of perceived novelty, and that as the same stimulus is repeatedly presented, infants 'habituate to' (get used to) the stimulus, with a corresponding stabilization of heart rate (e.g. Bower 1971). When the babies had habituated

to the first melody, a second one was played. This could either be a melody with the same starting note as the first but a different up–down contour of pitches, or it could be the same melody transposed up or down to a different pitch level. Chang and Trehub found that heart-rate destabilized to the new melody with a different contour but *not* to the transposition of the old melody. By five months, therefore, it appears that changes in sequential pattern are already salient, whereas simple changes in pitch level are not. In another study (1977b) the same authors showed that five-month-old babies were also sensitive to changes in rhythmic patterns.

It is instructive to compare Chang and Trehub's precise but limited study to a much broader but less controlled study by Moog (1976). Moog carried out a large cross-sectional study of children's responses to a set of prepared tapes. The tape contained six 'tests'. Test 1 consisted of three children's nursery songs, sung by children. Test 2 consisted of words spoken to definite rhythms but without precise pitch. Test 3 consisted of pure rhythms played on various combinations of percussion instruments. Test 4 consisted of instrumental music. Test 5 took one of the tonally consonant items from test 4 and subjected it to various rewritings which introduced a high degree of harmonic dissonance. Test 6 consisted of non-musical sounds such as the sound of a vacuum cleaner and traffic noise. These tests were played to about 500 children at varying ages between three months and five years. Moog observed the nature of response to the various tests, and recorded details of any musical behaviour in response to them.

He reports that six-month-old children typically stop what they are doing and turn towards the source of the sound with facial expressions of astonishment, remaining motionless and attentive at first, then looking at their mothers and smiling. They will even stop feeding to look at the sound source. The response is not equally marked to all the tests. Tests 1, 4, and 5 (songs and instrumental music) attract much more attention than the others. Moog reports that 'if a subject responded to one series only, that one was always either the nursery songs or the instrumental music; if he responded to two series, then they were these same two. If more than two series of tests attracted the baby's attention, then songs and instrumental music were, in every case, among the tests to which he responded.'

It seems that babies are here selecting *quality* of sound as the criterion for attention—smooth treble-register pitched sounds. Test 3 (rhythms) attracted hardly any attention, even though it was much louder than any other test. Valuable as Moog's data are, they do not tell us whether the babies were noticing sequential aspects of the sounds. All the results could be explained on a simple preference for what Moog characterizes as 'sensuously beautiful sound'.

Up to the age of six months, babies show hardly any overt behaviour which could be called musical. It is, of course, necessary to make a distinction between behaviours which can be heard *as* music by adults and those which

reveal musical awareness in the child. When Ostwald (1973) talks of the newborn baby's cry as having musical qualities, he is clearly referring to the former interpretation. It is nonsensical to suppose that the baby *intends* anything musical by crying.

The earliest sign of intentional music-like behaviour seems to be the ability of some babies to imitate sung pitches. Kessen, Levine, and Wendrich (1979) investigated this in a study where mothers were taught a technique for training their three- to six-month-old babies to match a pitch sung by them on a pitch pipe. They found that after 40 days of practice, all infants were matching sung pitches more often than failing to match; and this was regardless of the musical ability of the parents. There is, however, no evidence that babies below the age of one can imitate melodic sequences, even as short as two notes. Furthermore, 40 days of intensive training hardly forms a part of the normal enculturation process of most babies.

A much more common behaviour which could possibly be musically relevant is the ability of infants to mimic the intonational contour of speech. This is part of the pre-speech vocal exploration often called 'babbling'. Such intonational babbling is primarily composed of microtonal pitch glides, smooth movements across a range of pitches. Gardner (1981) reports that children do not produce *discrete* pitch intervals until the age of 18 months or so. This conclusion is drawn on the basis of an extensive longitudinal study on symbolic development carried out by Gardner and colleagues (Gardner, Davies, and McKernon 1981). In this study they observed the development of a whole range of abilities, including musical ones, in a group of nine children during the first five years of life. This involved regular visits to the homes of all the children, recording their spontaneous musical behaviour, and also attempting to teach them simple songs. This study provides valuable fine-grained data on musical development, and we shall return to it on several occasions in this chapter.

Overt rhythmic behaviour is not noticeably present in the first year of life; although one must again take care to distinguish that which can be interpreted as rhythmical by an adult from that which is intended to be rhythmical by the child. My eight-month-old daughter repeatedly bangs her spoon on the desk of her chair, or repeats a nonsense syllable over and over in her babbling, but I would not wish to ascribe to either of these approximately regular occurrences a rhythmic intention. Rather, she repeats the action as soon as she can after finishing the previous one, and, because it takes roughly the same time to perform each repetition, a roughly regular pulse is set up. To ascribe rhythmic *intention* I would look for some of the following behaviours:

(*a*) subdivision of a beat, so that sometimes there are two or more events within a regular superordinate pulse;

(*b*) omission of a beat, with the picking up of the pulse at the correct time after a pause;

(*c*) imitation of a given rhythmic pattern;
(*d*) moving or beating in time to music.

None of these behaviours normally occur within the first year of life.

In Moog's (1976) study it was found that nearly all babies made some sort of movement response to his tests. The most common responses were swaying from side to side in a sitting position or bouncing up and down. These responses were most commonly observed to song and instrumental music rather than pure rhythms, and they were not rhythmically co-ordinated with the music. Between the ages of nine months and one year these movements increased in frequency, duration, and intensity. Moog also found that around nine months babies began to make distinctive vocalizations to the music. He calls these vocalizations 'song babbling' in that they tend to be varied in pitch but lacking in phonemic variety, often being 'sung' to an open vowel. Song babbling bears no relation to the pitch or rhythm of the music being played; it seems to be a characteristic response of pleasurable expression to the sounds. This is confirmed by the fact that movements and vocalizations occurred primarily to the song and instrumental tapes, and did not occur so frequently to the rhythm or 'noises' tapes. Indeed, babies over nine months began to make quite clear signs of displeasure at these latter tasks, turning away, pulling 'dissatisfied' faces, even showing signs of fear. The increase in vocalization to music over the first year is quite dramatic. At three months only 5 per cent of children made vocalizations. At six months this proportion had risen to 30 per cent and by nine months 100 per cent of the children tested vocalized during the music. It is not clear, however, from Moog's report whether he checked that vocalizing occurred more during music than at other times, or whether the results simply demonstrate an increasing propensity for babies to vocalize at any time.

Do babies of under one year *recognize* familiar pieces of music? The initial possibility of this occurring rests on the fact that parents very often sing the same nursery tune to their children over and over again. My daughter has 'Baa baa black sheep' from someone in the house at least a dozen times a day. At a later age, signature tunes to familiar TV programmes, often played records, and others, can be added to the list of repeating musical experiences. Moog (1976) reports that children towards the end of the first year are often able to make the appropriate movement when they hear 'action songs' which are associated with particular gestures. However, it is not necessary to have the melody. Children still make the appropriate movement if the words are *spoken* rhythmically. Dowling (1982) reports that 'my daughter at 18 months would run to the TV set when she heard the "Sesame Street" theme come on, but not for other melodies'. Clearly then, young children are capable of recognizing some aspect of familiar music. However, we must not conclude from this that they 'know the tune'. Their recognition may be based on the words, or on some distinctive texture or timbre. It is disappointing for me to discover that I can sing 'Baa baa black sheep' to almost any set of pitches

provided I retain the words and rhythm. My daughter will still stop struggling and look pleased.

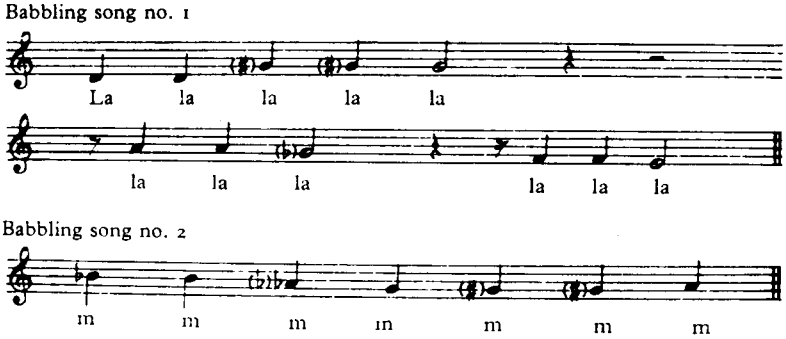
What, then, can we conclude about musical enculturation in the first year of life? It seems that children come to distinguish musical sounds from non-musical ones, as shown by greater attention, movement and vocalization. The primary qualities that mark the sounds they prefer seem to be those of the singing voice or instrument. It is possible that these responses arise from a biological propensity to respond in a special way to certain classes of sound. We also find that children can imitate individual sung pitches, and can discriminate short sequences with differing pitch or rhythmic contours if heard several times in a row. However, there is little or no evidence that children retain much precise musical information about songs they hear often; and their pitch and rhythm-like productions do not share the organizational features of the melodies that evoke them. It should be said, however, that these broad conclusions urgently require refining through the type of controlled experimental study that other aspects of infancy have now been enjoying for 20 years or more.

### 6.2.2. *The pre-school child (ages one to five)*

The first striking change in overt behaviour after the first birthday comes at about the age of 18 months when spontaneous singing begins to occur. The main characteristic of spontaneous singing is the use of discrete stable pitches (rather than the microtonal glides of the earlier 'song babbling'). These pitches are joined in sequence to form simple intervallic patterns. Although the child at this age has usually begun to speak, words are not usually used in spontaneous singing. This suggests that musical development at this age is proceeding along a genuinely separate 'stream' to speech. Moog (1976) has observed that 'single words or parts of words may occur, scattered around in a string of nonsense syllables or at the beginning of a babbling song which after beginning with the word continues with the repetition of a single syllable'. At this stage there is no evidence that children are attempting to imitate heard songs; rather it seems as though they are experimenting with melodic interval construction. Gardner *et al.* (1981) noticed that the most frequent intervals initially used approximated to seconds, and major and minor thirds. As the second birthday approaches, children begin to experiment with larger intervals such as fourths or fifths. In all cases the intervals are only approximate, and there is little evidence that children are producing the exact intervals of the diatonic scale. Several authors have attempted to transcribe these spontaneous songs using conventional notation. Whilst this provides a useful impression of the general characteristics of child song, one must be careful not to ascribe more tonal and rhythmic coherence than is actually present. The temptation for an adult listener is always to assimilate what a child sings to the adult categories.

With this cautionary statement in mind, we may examine some examples

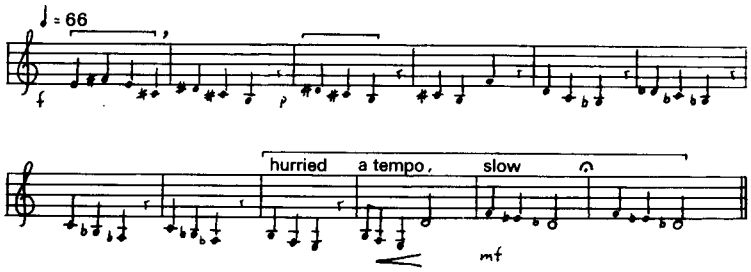
of spontaneous song. Moog (1976) provides two typical examples of songs produced in the second year of life (Examples 6.1 and 6.2). A sharp or flat




Examples 6.1 and 6.2. Spontaneous songs of one-year-old children. From Moog (1976).

sign enclosed by brackets indicates an approximate quarter-tone interval above or below the given note. We may see that the songs are dominated by simple repetitions of the same pitch, or small movements up and down in pitch. They are also rhythmically primitive, being dominated by simple repetition of a single note length. When pauses occur, they seem to be related principally to the need of the child to breathe, rather than any developed notion of rhythmic differentiation. It is, however, of interest that the end of Example 6.1 seems to involve a melodic and rhythmic imitation of a three-note figure. Clearly, though, these songs have little sense of adult tonality or harmony.

Between the ages of two and three, spontaneous songs become longer and begin to display greater signs of internal organization. Example 6.3 gives an




Example 6.3. Spontaneous song of a 32-month-old child. From Dowling (1982).

example of a song provided by a 32-month-old child (from Dowling 1982). Each note was sung to the syllable 'yeah' and brackets indicate regions of relatively accurate intonation. Elsewhere the notated pitch is only approximate. We can now see a quite clear and deliberate use of repetition. A descending major 3-2-1 is repeated nine times at different pitches with the

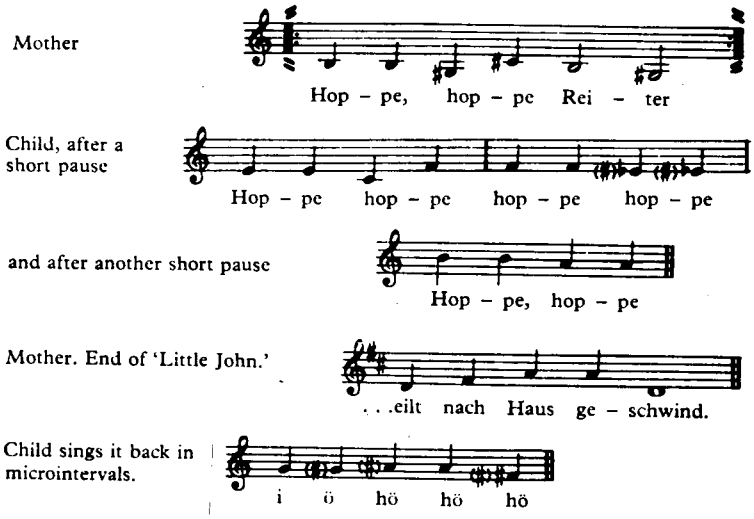
same rhythm. Furthermore, the intervals used often correspond exactly to the diatonic intervals of the scale. Although the pitch of the whole song wanders, and there is no overall sense of a stable tonal centre, each phrase tends to be tonally coherent. By two-and-a-half, the child seems to have assimilated the notions that music is constructed around a small fixed set of pitch intervals, and that repetition of intervallic and rhythmic patterns is a cornerstone of music. What he does not grasp is that there is any hierarchical structure governing groups of patterns which prescribe direction and closure. In general, child song at this age has an 'aimless' quality—it can go on and on without any sense of finishing. The decision when to stop is largely arbitrary.

By the age of two to two-and-a-half a new milestone is reached. Children begin to attempt to *imitate* parts of the songs they hear around them. The first aspects of songs that children imitate are the words—not the complete words, but particularly salient or often repeated sections. Moog reports children often imitating the 'ding ding dong dong' from one of the songs in his test series. Gardner (1981) reports observing 'an oink oink here, an oink oink there' from 'Old MacDonald'. Often these same few words will be repeated over and over. At the beginning there is little attempt to reproduce either the rhythm or the melody of the song concerned. Rather, children seem to couple such word patterns with the type of melodic fragment they have been producing in spontaneous song.

The next stage is the extraction of characteristic rhythmic and pitch patterns from the songs of the culture. In the study by Garner *et al.* (1981) study, this tended to occur towards the age of three. In Moog's (1976) study 50 per cent of the children had reached this stage by two years. This discrepancy is not something that should concern us unduly. Not only were the samples of different nationalities (Gardner—USA; Moog—W. Germany) but the sample sizes were very different. Gardner's generalizations are based primarily on in-depth studies of nine particular children. Moog's generalizations are based on a representative cross-section of 500 children. Both studies agree on the *order* in which the various milestones are reached; and we can see that within Moogs' study there were wide divergences between children at particular ages.

Moog provides an example of the kind of imitation provided by one two-year-old child in interaction with his mother (Example 6.4). We see here an exact imitation of the first four notes of 'Hoppe, Hoppe, Reiter' transposed to a new pitch. Then the child abandons strict imitation and begins spontaneous repetitive song-play on the word 'Hoppe'. In the second example, the mother sings the song 'Little John', and when she has finished, the child repeats the melodic contour of the last five notes (with much diminished intervals).

In general, melodic contour was imitated more often than exact pitch, and in no case did Moog observe *exact* imitation for more than a bar at a time,




Example 6.4. Song interaction between a two-year-old child and his mother. From Moog (1976).

and this only when the child was actually singing simultaneously with the model.

During the third and fourth years of life the child develops imitative capacity to the point where whole songs can be repeated. Generally the rhythm and the pitch contour is mastered before the ability to reproduce precise intervals and maintain the same tonality throughout a song. Most children can accurately reproduce the familiar songs and nursery rhymes of their culture by the age of five.

What happens to spontaneous songs during the child's increasing mastery of imitation? Firstly they get longer. Moog reports that three-year-olds will often produce continuous songs of several minutes' length. They do, however, occur less frequently as the preoccupation with imitation grows. By four years of age, about 30 per cent of Moog's sample were producing what he calls 'pot-pourri' songs. In these songs 'children make up new ones by putting together pieces of several songs which they already know. Words, melodic lines, and rhythms are mixed up, altered, taken apart and put together again in a different way and then fitted in between stretches of 'original ideas". Usually these 'pot-pourris' were free episodic structures with little sense of overall organization. However, in a few isolated cases figures or phrases from a learned song were altered according to a definite formal principle. One girl of three-and-a-half sang the whole of the Christmas carol 'Ihr Kinderlein kommet' in triple time instead of quadruple time. She did this by altering the dactyl pattern to a triplet wherever it occurred. Such grasp of formal structure at this age is, however, most unusual.

By the age of five, spontaneous song has declined dramatically in frequency

of occurrence. The child is now more self-conscious and is concerned with avoiding 'error' and achieving precision of imitation. This concurs with a general trend in children of this age for mastery of detail. Gardner and Wolf (1983) characterize this as a 'wave of symbolization' which moves children from a stage of 'topological mapping' to one of 'digital mapping'. In the earlier stage, children operate mainly with approximate relations of shape and size. In drawing, for instance, they will not be too concerned with getting the right number of fingers on a drawing of a human hand. In the later, digital, stage children become almost obsessed with exact quantification and classification. It is the age at which they will painstakingly put exactly the right number of fingers on each hand even at the expense of stylistic demands of the drawing. For instance, running figures often appear 'frozen' and lifeless because the child is more concerned to try for correct anatomical details rather than convey a sense of motion in which, at a younger age, maybe an extra arm or leg is added.

The same tendency may also account for the love of exact repetition. Moog reports that 'a dozen musically minded parents agreed that small children go on asking to sing the same song for weeks and months on end, just as they go on asking for the same stories over and over again . . . When they look at picture books they look for the same picture, and when they themselves draw they go on drawing the same things for a long time.'

There are a number of possible consequences of this concern with precision and repetition at age four to five. One is that the child is focusing on and using, maybe for the first time, the characteristics of song which are determined by superordinate structures of tonality and rhythm. Although the behaviour of the child may seem externally static, it is very likely that he is using his increasing knowledge of, and memory for, exact pitch and timing relationships to build up knowledge about higher-order structures in music, extracting a new level of knowledge about scales and keys, and about rhythm and metre. We see this knowledge at work in a study by Davidson and McKernon (reported in Gardner 1981) where they taught a new folk song to children aged four and five. They found significant changes between four and five in the imitative products of learning. The five-year-olds were able to maintain a single key throughout, starting on and returning to the tonic, even if some of the individual notes were wrongly remembered. In contrast, the key of the songs produced by four-year-olds tended to 'drift' as the song progressed. Secondly, five-year-olds were able to organize their reproductions in terms of an underlying pulse or beat. Their songs conformed to the beat even if individual rhythms were misremembered. In contrast, four-year-olds tended to organize rhythm 'locally', largely from placements of accents in the words, and they did not maintain a steady pulse.

Another consequence of the concentration on precision is that spontaneous musical experimentation may well cease to play any further real part in a person's life after the age of five unless it is specifically encouraged. Our

Western culture does not provide many opportunities for people to improvise. Much more importance is placed on the shared reproduction of well-known music (e.g. hymn singing, ritual singing among football supporters, the classical concert sub-culture, etc.) In our culture, composers and improvisers do not arise from normal enculturation; their skills are nurtured by specific, and culturally atypical support and training.

We have examined the development of singing between the ages of one and five in some detail. There are, of course, other aspects of musical behaviour and awareness which develop over these years as well; so we should now return and examine some of these.

We first consider the development of movement responses to music. In the second year of life, most children learn to walk. This allows a much increased range of movements, an increase which is represented in movements to music. Moog (1976) found that more than half of his 18-month-old subjects 'danced' in response to music, making circular or turning movements in conjunction with stepping. He also found a large increase in the frequency of 'conducting', that is, a waving of the arms in response to music. These movements are not rhythmically co-ordinated with the music. They seem to reflect a general motor enthusiasm and exuberance which music somehow elicits.

Around two years of age 'about 10 per cent of children begin, for short stretches of time, to match their movements to the rhythm of the music. . . . The few children who are able to do this at all are only able to keep up the co-ordination for a few bars at a time.' Up until the age of five there does not seem to be any marked change in the ability to move in synchrony with music (confirmed by Rainbow and Owen 1979; and Frega 1979). However, two trends are noticeable. First, the variety of movements increases. Second, and perhaps more significant, the number of spontaneous movements decreases very markedly. As children grow older they are more prone to a concentrated, still, attentive listening. By the age of five many children showed no spontaneous movements at all to Moog's tests. This may be part of the same developmental trend to imitation and precision that we noted earlier. Because there is no definite model for specific movements to be made, children prefer to make none.

To elicit movement responses in children of five years it seems necessary to *ask* them to make particular movements, such as clapping or tapping, to see whether they relate to the music. Moog did this, and found that very often the children would make 'any sort of clapping movement, sometimes not even rhythmically regular ones.' Only about two-thirds of the children kept in time with the music for even a short time. On the other hand, when asked to clap whilst singing a song they knew, they could clap in synchrony with the successive notes of the song. It is not clear, however, that children of this age are capable of maintaining a steady clapping *beat* whilst singing a song which variously subdivides the beat and contains silent beats. Rather they are able to use the same mechanisms that time the successive notes of the song to time successive hand-claps.

Here we see a type of dissociation that we shall have cause to consider again. This is the dissociation between implicit knowledge (in this case of metre) that children display in their enactive repertoire (such as in singing the songs they know) and the explicit ability to isolate and use such knowledge in tasks of perception and judgement. In this case, children of five have implicit knowledge of metre but seem not to be able to use this knowledge explicitly.

So much for movement to music. We now consider the development of discrimination. It appears, from Moog's extensive sample, that very few children differentiated between the 'normal' instrumental music and the 'dissonant' manipulations of it. Children attended, made rhythmic and vocal movements just as readily to both tests. 'Not a single child showed the least sign of displeasure' at the dissonances. In contrast *many* children made no response (or a negative one) to the pure rhythms (test 3). At age three, 37 per cent of the children simply ignored the test. Another 24 per cent showed signs of being uneasy; and one little boy asked his mother 'when are we going to have music again?' Another said 'that isn't music, is it, mummy?' We conclude from these observations that harmony is not particularly salient to the pre-school child, but that he has already identified music with some form of melodic content.

In respect of familiar music Moog discovered a gradual increase in true tune recognition up to the age of five. When a familiar nursery tune was played, without words, on an instrument only 40 per cent of four-year-olds recognized it. This proportion rose to 75 per cent of five-year-olds. By this age children are just beginning to reliably catch hold of the notion that melodic identity is carried by a pitch and rhythm pattern, regardless of timbre or words.

Finally, Moog examined the effect of environment on development. He summarizes his findings thus:

> Up till the age of about three we could not observe in response to music any significant differences determined by the environment; children from poorer homes reacted no differently to our tests from children in the upper income groups. Children who, according to their parents' reports, were inundated with music from morning till night showed little difference, in either the quantity or quality of their response, from children who were only allowed to hear carefully weighed amounts of carefully chosen musical stimuli. . . . But, between the ages of three and four, differences in home environment begin to show their effect in the field of music. Girls and boys who are taught songs and games by their parents, brothers and sisters, or in nursery schools, have a clear advantage over other children.

Moog also found quite marked differences between children from the *same* kind of background. It is impossible to rule out the possibility of innate differences in musical receptivity, although one should, perhaps, not move to that conclusion as readily as Moog does. What the results do suggest is that after the age of three a child does become receptive to certain types of environmental 'enrichment'. Before that age, immersing one's baby in music may not help him or her become a second Mozart.

How do we sum up the path of musical development between one and five? We see, I think, four main strands. First, the child's increasing propensity to *imitate* words, then melodic fragments, then whole songs. Second, a subordinating of a free improvisatory and unstructured note-play to the forms of the musical culture, by the incorporation of diatonic intervals, then phrases from well-known songs, and finally its disappearance in the service of exact imitation. Third, the increasing ability to organize song behaviour in accordance with tonal and metrical rules. Fourth, a concurrent lack of ability to extract metrical or harmonic information in a situation where some abstracting or evaluative response is sought. Children in this age do not reliably beat time to music, nor do they seem to notice gross dissonance. It seems as if knowledge becomes incorporated in action before it can form the basis of judgement.

This last point can be demonstrated through other aspects of cognitive development. For instance, children find it easier to recall a list of objects when they can be grouped under the headings of two or more common categories (e.g. 'fruits', 'items of clothing'); but at the same age they may be incapable of *judging* a categorized list to be more memorable than an uncategorized one. It is only at a later age that they can accurately judge which lists are likely to be easy and which difficult without actually doing the task of memorization. This reflective grasp of the situation is most generally observed to first occur between the ages of eight and ten (Hunter 1976; Moynahan 1973).

### 6.2.3. *From age five to 10*

Piaget's investigations of intellectual development led him to propose a profound change in general cognitive ability at around the age of seven or eight. He labels this a change from 'pre-operational' to 'operational' thought. The most celebrated 'symptom' of this change is the child's ability to perform well on tasks involving the notion of conservation of quantity.

Let us take a common example. A child is shown a rolled-up ball of plasticine. Before his eyes, the experimenter then rolls it out into a long thin sausage. The pre-operational child tends to believe that there is now *more* plasticine (because in one crucial respect it *looks* more). The operational child knows that the amount is the same. More than this, he knows it *must* be the same. Other examples of conservation tasks involve pouring liquids from short squat beakers to tall thin ones, or spreading out a line of initially closely bunched counters. In each case the operational child is not fooled. He knows that the amount remains the same.

Several music researchers have attempted to construct musical analogues of these tasks. Pflederer (1964), for instance, played children the same melody at two different speeds and asked them if they were the same. Only 50 per cent of five-year-olds thought the two were the same. By eight years of age this proportion had risen to 94 per cent.

Although this task has a superficial resemblance to Piaget's conservation tasks, it differs in crucial respects. First, children do not observe the melody under continuous transformation between one speed and the other, and so have no independent way of knowing that the two melodies are of the same 'stuff' (Wohlwill 1981). Second, it follows that there is no logical *necessity* that such transformations should conserve musical quantities or qualities. That is to say, one could present two melodies at two different speeds which did *not* have melodic identity. In contrast, there is no way of deforming a piece of plasticine in such a way as to change the quantity. Children become competent at Piagetian conservation tasks precisely when they realize that transformations of configuration are *always* irrelevant to judgements of quantity.

In music the notion of a melody undergoing a continuous observable transformation is problematic, if not incoherent. Each occurrence of music is a separate event in time, totally discrete from the next event. Therefore, there is no necessary relation between one musical event and the next. When Pflederer and others claim to show that children around the age of eight come to conserve musical quantities, it would perhaps be better to say that children of this age are capable of perceiving and remembering invariant aspects of otherwise different patterns. Whilst this ability is of profound psychological interest and importance, it is *not* conservation. Nonetheless, we may suppose that it shares with conservation the increasing awareness by children of the possibility of going behind surface perceptual features in their search for underlying patterns and structures.

If one were to sum up the main developmental trend in music between five and 10 it would seem to be the increasing reflective awareness of the structures and patterns that characterize music and which are already implicit in the child's enactive repertoire. For instance, at age five a child is already capable of singing the same song at different speeds or pitches. In one sense, therefore, this behaviour demonstrates the existence of knowledge that a melody is still the same after transformations. What the child of five lacks is the ability to translate that knowledge into a procedure for making a considered judgement in a perceptual situation where a change, such as in speed or pitch, is highlighted through temporal juxtaposition. The progress from enactive to reflective knowledge does not take place all at once, and we find various aspects of musical awareness undergoing this change in order over the middle years of childhood.

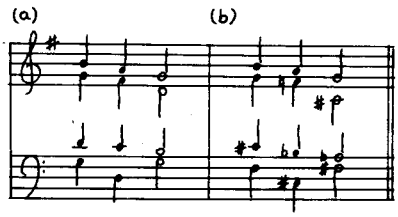
An early aspect of this change is illustrated by the work of Zenatti (1969). She presented children with three-note melodies, each followed by a second melody in which one of the notes had been altered in pitch. Some of the melodies were atonal. The others were tonal. Children were asked to locate the changed note. Zenatti found that five-year-olds performed equally badly on both types of melody, but by six or seven the tonal melodies were producing superior performance. The tonal melodies allowed the extraction of

some structural information which could be carried over in memory to aid the comparison with another melody. Although this shows the beginnings of a dissociation of structural knowledge from a purely enactive context, it is still linked to a performance task which does not demand *direct* awareness of structure. We may ask when children are able to form *judgements* about musical sequences when asked to make such judgements in a direct way.

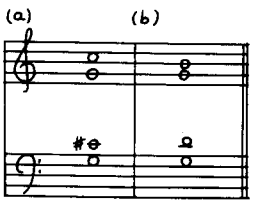
One way of examining this is to ask children to classify musical examples as'good', 'bad', 'complete', incomplete', etc. For instance, Imberty (1969) played excerpts from Bach chorales to children of various ages and asked them to judge whether excerpts were complete or incomplete. Eight-year-old children tended to hear melodies as complete when they ended on the tonic in the context of a perfect cadence; but even at age 10 they tended not to accept as complete perfect cadences with melodic resolution to the mediant.

I have examined the development of judgement in a slightly different context. In my study, children were presented with a pair of musical stimuli and were told that one was played 'correctly' the other one had 'mistakes'. Each member of the pair came from a different sound source, well separated in space and time. Children were asked to indicate (by writing or pointing depending on age) which sound source had produced the 'correct' version. There were four tests. In the first test the 'correct' item resembled a cadential sequence from a typical four-part hymn harmonization; the 'incorrect' item was derived from the correct item by modifying each chord so that it was dissonant. To the adult Western ear such dissonances are extreme (Example 6.5 gives one example of such a pair: all items in this study were played on the piano).

In the second test, the items were single chords, one being highly consonant, the other breaking one or more rule of conventional chord construction (see Example 6.6). In the third test, the items were sequences of

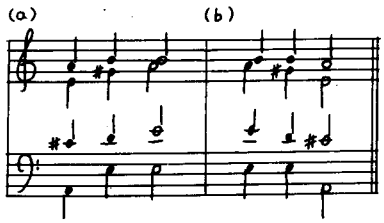



Example 6.5. A stimulus pair from test 1.

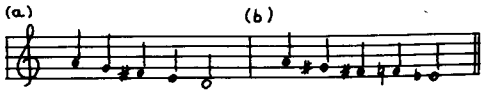



Example 6.6. A stimulus pair from test 2.

consonant chords. In one sequence of each pair the chords were in a 'musical' order that led to a conventional cadence. In the other sequence, the same chords were present but in 'scrambled' order (see Example 6.7). In the fourth and final test, the items were unaccompanied melodic sequences. One item was a diatonic melody which remained in a single key. The other item was a sequence with similar contour but containing notes from several distant keys so that the tonality became obscure (see Example 6.8).




Example 6.7. A stimulus pair from test 3.




Example 6.8. A stimulus pair from test 4.

In each test there were twelve pairs, so that with a score of 1 for the 'correct' item, the maximum score for each test was 12. A subject guessing at random would, on average, score 6 out of 12. I ran the experiment on five different age groups, five-, seven-, nine-, and 11-year-olds, and an adult control group. The average scores of each group on each test are given in Table 6.1. Circled scores are not significantly different from random guessing.

TABLE 6.1

*Mean scores (out of twelve) on each test for each age group*

| Age | 5 | 7 | 9 | 11 | Adult |
|---|---|---|---|---|---|
| Test 1 | 7.3 | 9.1 | 11.4 | 11.8 | 12.0 |
| Test 2 | (5.5) | 7.7 | 9.5 | 9.3 | 10.6 |
| Test 3 | (5.6) | 7.3 | 9.3 | 10.5 | 11.6 |
| Test 4 | (6.4) | 7.3 | 9.7 | 9.6 | 11.6 |

Scores joined by an underline are not significantly different from one another.

We may first observe that the five-year-olds performed very poorly. Only on Test 1 did they perform slightly above chance. In many cases I observed the children to be making choices on non-musical bases. For instance, one child always chose the same sound source, no matter what music came from it. Another child alternated between sound sources.

By age seven, performance significantly improved on all tests except Test 4 (unaccompanied melodies), although performance was still well below adult levels. At age nine, score on Tests 1 and 2 became indistinguishable from adults. It is pertinent to notice that around this age children began to show strong aesthetic reactions to the 'incorrect' items (dissonant chords)—screwing up their faces in disgust, or laughing and giggling. These tests require least 'work' because the incorrectness arises from simultaneous sounds. Each event is 'wrong' by itself and requires no memory of previous events in the same sequence.

It was not until the age of 11 that scores on Test 3 (chord sequences) approached those of adults. In this test, each chord on its own is acceptable; to detect incorrectness one has to notice the ordering of chords. Performance on Test 4 (unaccompanied melodies) did not improve between ages nine and eleven, and was still significantly below adult performance. Here, the 'work' required is arguably the greatest. Not only do incongruous events not happen simultaneously, but the cues to an underlying key or harmonic framework are much more sparse than in the case of fully harmonized chords.

This study shows, therefore, a progression from age five to adulthood, in which the listener becomes capable of reflective judgement on the musical quality of successively more 'difficult' aspects of music; starting with the ability to reject gross dissonance, and moving on the the ability to detect violations of normal sequential structure.

Two subsidiary findings are of note. Girls performed better than boys on most tests and at most ages. This result has been obtained in quite a lot of musical developmental research (Shuter-Dyson and Gabriel 1981). Secondly, musical training made no difference to the scores. Those children who were receiving regular instrumental tuition fared no better than those who were not. This is a strong indication that the skills we are dealing with here are true products of enculturation, and do not rely on specific training. They arise from the normal child's intellectual encounters with the music of his culture.

The final topic I wish to discuss in this section is the development of the ability to make finer classifications within music which is broadly 'acceptable'. This ability is often given the name 'style sensitivity' and relates to a person's ability to tell whether two extracts of music come from the same composition, composer, or period. Gardner (1973) investigated style sensitivity in an experiment where he played pairs of musical extracts to children, some of which came from the same composition and some of which did not. The children were required to say whether or not they thought the two extracts were from the same piece of music.

There was a significant age improvement in the ability to do this. The youngest children (aged six) had a high tendency to judge extracts 'different' and Gardner suggests that 'pieces had to sound identical or directly continuous in order to be judged as from the same composition'. Six-year-olds

had a very limited ability to explain their choices, and those that could offer comments based their discriminations largely on dimensions such as high/low, loud/soft, or fast/slow. Of the 16 pairs presented, six-year-olds judged only 9.7 of them correctly on average (chance score would be 8).

By eight years of age, average scores had risen to 10.5 on 16, and children were able to give 'same' judgements even when the two segments did not sound continuous. Subjects commonly described their perceptions using adjectives or metaphors drawn from outside music: 'peppy', 'dull', 'churchy', 'grown up', 'like a horse race'.

At 11 years of age the average score was 12.4, and subjects began to explicitly refer to aspects of instrumentation, rhythmic character, and texture in making their judgements. Although average scores did not increase above age eleven, subjects aged 14 began to justify their responses in terms of general style, using terms such as 'jazzy' or 'baroque'. Coincident with this was a greater ability to correctly differentiate 'different' pairs when they came from widely different stylistic eras. For instance, they were better at discriminating Boulez from Bach (300 years between them) than Schumann from Brahms (contemporaries). Thus, in this study, we see a progression from judgements based on simple physical features at age six to judgements made on complex multidimensional aspects of style and underlying 'language' by the age of fourteen.

We may summarize the main strands that emerge from our review of musical enculturation as follows. There is some evidence of special propensities which support the early establishment of a 'stream' of specifically musical development. Very young children seem particularly responsive to musically pitched sounds, and are also responsive to changes in pitch and rhythmic sequence. During the second and third years of life we see the emergence of improvisatory song which does not directly mimic the song of the culture (just as children's early speech is not purely imitative). By the age of five children are able to use underlying tonal and metrical structures to guide their song performance, even though they seem to have no reflective awareness of such structures, being almost totally unable, for instance, to identify gross dissonance in simultaneously sounding notes.

We also see the possible influence of wider 'waves' of intellectual development having their effect in the musical sphere. For instance, the increasing concern with precision and imitation around age five mirrors developments in other symbolic domains such as drawing (Gardner and Wolf 1983). This change also corresponds to a change from 'the stage of romance' to 'the stage of precision' suggested many years ago by A. N. Whitehead (1917).

Changes in musical awareness between the ages of five and 10 seem to reflect a general intellectual change from enactive competence, which is displayed only within the bounds of specific and directed activities, to a reflective awareness of the structures and principles which underlie such

competence. This change is characterized by Piaget as a change from pre-operational to operational thought. In music it is marked by an increased ability to explicitly classify music as conforming to rule or style, and an increasing advantage in memory and perceptual tasks for those sequences which conform to rule.

### 6.3. Training and skill acquisition

Musical training is the means to acquisition of specific skills which build on the base of enculturation. We have seen in the previous section that the normally enculturated child already has a set of musical abilities—such as the ability to recall songs and learn new ones; the ability to tell different types of music apart; the ability to make use of underlying features such as metre and tonality in organizing performance. In the 'training' phase, which may overlap with enculturation, and which may extend indefinitely into adulthood, we see the growth of a variety of skills—performance on instruments and voice, composition, aural analysis, conducting, etc. No single individual necessarily achieves expertise in all these skills, although a culture generally agrees on the skills required by an 'all round musician', and this set of skills becomes institutionalized in the training and examination procedures of music schools and colleges.

Each musical skill poses its own training problems and has related with it a long pedagogic tradition. In a book of this sort it would be impossible to examine all aspects of skill in the detail they deserve. We have already examined some aspects of the acquisition of performance skill (Chapter 3) and improvisatory skill (Chapter 4). What I want to do in this section is draw out some general principles of skill acquisition and training with illustrative examples drawn from some specific skills. My concern with principles is such that I do not pretend to be overly concerned with providing 'advice for teachers and learners'. There are several books of this latter sort, among which Buck's *Psychology for musicians* (1944) is outstanding. Despite its age I would still recommend it as essential reading for any performing musician. Its author has the main prerequisite for writing a book of this sort—a lifetime of high-quality teaching experience at all levels.

Psychology provides many, sometimes apparently conflicting, interpretative frameworks for understanding learning. I would like to organize this section by introducing just one detailed theoretical proposal, and showing, in discursive fashion, how it can be used to support broad generalizations about the business of learning and teaching. The theory is *Production System theory* as applied to skill acquisition by Anderson (1981, 1982). The basic attraction of Production System theory is that it is possible to construct machines (embodied in computer programmes) that operate according to the postulates of the theory. Some typical aspects of human thought and behaviour can be mimicked by such a machine. The basic elements of a Pro-

duction System are very simple and precise, yet behaviour of some complexity can be obtained when many of these elements are put together with a few simple rules for the way they interact. It is most important to bear in mind that production systems are not in themselves part of the human mind. Rather, they are formal analogies for mental processes. However, to simplify my task, I will talk *as if* production systems are part of the mind. This will save obscuring the argument by repeated qualifications.

Before embarking on detailed exposition, I would like to introduce a few general concepts associated with skill acquisition which I shall be filling out with the aid of theory and example. These are quite fundamental concepts, and turn up, in one form or another, in most writing on learning. First is the concept that skill learning involves acquiring *habits*. The principal feature of a habit is that it is *automatic*, and that it uses up little or no mental capacity to execute. The precursors of habits are conscious, deliberate, and effortful behaviours which commonly involve verbal control. Second is the notion that skill learning involves passing from *factual* knowledge (knowing that) to *procedural* knowledge (knowing how). Knowing what a skill entails is very different to actually executing it, and a theory of learning should be able to refine our understanding of exactly what changes when factual knowledge becomes procedural knowledge. One part of the change seems to be that knowledge comes under direct and more intimate control by *goals*. For most of us, to formulate the goal of performing some familiar task, such as saying a common word, is enough for the task to be accomplished. There seems to be no psychological gap between wanting (or intending) and doing.

Many of our goals are as small and short-lived as saying a particular word. Other goals are larger and more long-term, as, for instance, the goal of being able to play a musical instrument. The ability to form and sustain goals seems to be an essential condition of learning. Such ability is often called *motivation*. Other general conditions essential for most skill learning are *repetition* and *feedback*. People generally become skilled at some task by being presented with repeated opportunities to engage in elements of the task. The sheer amount of time that a person has spent actually doing an activity is one of the best predictors of level of skill. For any complex skill, such as writing a computer program, playing chess, or playing a musical instrument, moderate levels of ability are achieved by those who have devoted hundreds of hours to it. For real expertise, thousands of hours are required. Feedback (or reinforcement as it is called in some contexts) is essential so that only succesful procedures are learned; any procedure which leads to repeated failure is discarded.

We begin our detailed treatment with reference to a proposal by Fitts (1964) that the process of skill acqusition can be broken down into three phases or stages: the *cognitive stage*, the *associative stage*, and the *autonomous stage*. Anderson (1982) succinctly summarizes these stages as follows:

> . . . the *cognitive* stage involves an initial encoding of the skill into a form sufficient

to permit the learner to generate the desired behaviour to at least some crude approximation. In this stage it is common to observe verbal mediation in which the learner rehearses information required for the execution of the skill. The . . . *associative* stage involves the smoothing out of the skill performance. Errors in the initial understanding of the skill are gradually detected and eliminated. Concomitantly there is a drop out of verbal mediation. The . . . *autonomous* stage is one of the gradual continued improvement in the performance of the skill. The improvements in this stage often continue indefinitely.

Let us now consider what an aspect of the *cognitive* stage might look like for someone who is already competent to read and perform on one instrument beginning to learn a new instrument. I found myself in this situation when, after several years of piano playing, I took up the clarinet. When attempting to play my first simple melodies on the clarinet I had constantly to make recourse to verbal information, sometimes recalled with difficulty from memory, more often retrievable only from my instruction manual. The information was primarily in the form 'the fingering for note X is combination Y.' When the right set of fingers were engaged with the right set of keys, there was still the problem of producing a sound by blowing. This was not straightforward, and to produce any recognizable sound at all required constant adjustment of mouth, tongue, and lips, in accordance with another set of verbal prescriptions. There seemed to be so many different things to attend to at once that fluent performance seemed an impossible dream.

This sense of the crowding in of demands which cannot all be fulfilled is a characteristic of early learning in any endeavour, be it driving a car or learning a new language. There is just too much to think of, too much to remember. Such a sense can be profoundly discouraging, and may cause a learner to give up before any progress is made. The fortunate learner will have available a strategy, possibly imparted by a good teacher or a manual, which may help to circumvent such a block.This is the breaking down of the skill to be learned into a set of components which can be acquired stepwise. In each step the amount of new verbal information is small, and the learner is able to progress beyond the cognitive stage with this small package before proceeding further. In this way the learner is able to shield himself from the ultimate (and daunting oal), setting his sights instead on the very next stage.

On the whole, it seems fair to say that the younger the child, the less able he will be to devise and impose his own schedule. Even adults may find such a task difficult, and will benefit from the advance organization of material to be learned by an experienced teacher. A teacher has oversight which the novice lacks. He knows which aspects of a skill typically cause most difficulty, and which aspects, learned early on, help the course of later learning. On the clarinet, for instance, there is much to be said for acquiring some fluency in the lowest register before attempting music which 'crosses the break'. Each register re-uses the same fingering sequence, uncovering successive holes as the register is ascended. When one arrives at the top of the lowest register,

one must depress a register-shift key, re-engage all the fingers, and modify the blow in order to get the next highest note. On a B flat clarinet the top note of the bottom register is the B flat above middle C. A self-directed adult learner might, without guidance, decide to learn the C major scale, one octave up from middle C as his first task. This would involve crossing the break, with all its complications. The experienced teacher (or a good manual) will advise that a better scale to start on would be the F or G major scale that can be totally encompassed within the lower register.

The application of this principle can be extended indefinitely. For instance, in piano learning it makes sense to acquire patterns which do not require shifts in hand position before those that do. In violin learning it makes sense to master open-string positions before learning fingered positions; and so on. Only the experienced practitioner of the particular skill is fully qualified to devise generally useful training schemes by virtue of his oversight of the whole skill. It is, however, worth emphasizing the point that there are very many possible ways of segmenting a learning task which adequately fulfil the requirements of a learner; and it is perhaps more important that the *principle* of segmentation is adopted, than that the learner should have the belief that *one* of the many reputable schemes available is somehow best.

What is happening in a learner as he passes from the cognitive stage to the associative stage with respect to a particular section of skill knowledge? Our overview of concepts in skill acquisition suggested that he is acquiring some *procedural* knowledge in addition to the factual (or declarative) knowledge that was acquired from the printed or spoken word. To continue with the clarinet example, a piece of declarative knowledge might be the following:

D1. The clarinet fingering for Middle C is the left-hand fingers covering the top three holes.

A simple part of the equivalent procedural knowledge might, in Production System theory, look like Table 6.2.

Each of the statements P1 to P5 is a procedural step called a production rule. Each rule has the form IF condition(s) THEN action(s). It tells one, in simple and unambiguous terms, what to do when a particular state of affairs occurs. Although the language of a production rule is plain, it can often cause perplexity to a first-time reader. One basis for perplexity is that each rule, in itself, seems so trivial and lacking in substance. The reason for this is that by writing out a rule in English for conscious scrutiny one is subjecting it to far more attention than it would ever receive by the individual who possessed such a rule in his procedural repertoire. The whole point about a production rule is that it is so simple that the subconscious mind is capable of executing many dozens of them, without attention and effort, in a few seconds. What appear as single steps in thought to our *conscious* awareness, are actually composed of many subconscious but logically primitive steps which we have run through so many times that they have become completely

TABLE 6.2

*Example of Production System*

| | | | |
|---|---|---|---|
| P1. | IF | the goal is to play middle C | |
| | and | the fingers are not in the configuration of having the left-hand fingers covering the top three holes (configuration L3) | |
| | THEN | the sub-goal is to achieve configuration L3 | |
| P2. | IF | the goal is to achieve configuration L3 | |
| | and | no configuration is presently in force | |
| | THEN | place the fingers in configuration L3 | |
| | and | the goal is achieved | |
| P3. | IF | the goal is to achieve configuration L3 | |
| | and | a configuration other than L3 is presently in force | |
| | THEN | the sub-goal is to achieve the state of having no configuration in force | |
| P4. | IF | the goal is to achieve the state of having no configuration presently in force | |
| | and | some configuration is presently in force | |
| | THEN | remove all fingers from the keys | |
| | and | the goal is achieved | |
| P5. | IF | the goal is to play middle C | |
| | and | the fingers are in configuration L3 | |
| | THEN | blow | |
| | and | the goal is achieved | |

automated. Thus, a clarinet player using the procedures P1-P5 would actually run through the cognitive steps contained therein in a fraction of a second. This is in contrast to the laborious minutes it took me to write out the productions, or you to comprehend them. I chose a very simple task, that of playing middle C, precisely so that the number of production rules would be few. If I had chosen a more complex and interesting task, it would have taken me many weeks to write out the production rules required, and it would take a reader many hours to comprehend them.

Let us tease out some of the characteristics of production systems, using this example. A crucial feature is the way that the rules embody a hierarchy of goals and sub-goals. In this example, the main goal is to play middle C. The two rules explicitly mentioning this goal as a condition (P1 and P5) are left-aligned to help make the goal structure clear. If the player is lucky enough to find his fingers already in the required position, then all he has to do is blow, and his goal is achieved (P5).) If, however, his fingers are not correctly positioned, then he must get them correctly positioned. So he for the moment sets aside his main goal of playing C and embarks on achieving a subsidiary goal, that of getting his fingers into the right configuration (P1). Rules P2

and P3 (indented one step) relate to this new goal. In this set of rules I have attempted to model the procedures of a fairly raw beginner. One commonly observes beginners completely removing the hand from the keys prior to each note, and then choosing the new configuration 'from cold'. This is, of course, wasteful of effort, and more experienced players learn procedures whereby the existing hand position is used to find the next position. So, for instance, if a sub-set of the required fingers were already in position, they would not be removed.

With our hypothetical performer we assume that if he finds his hands taking up the configuration for playing some other note than middle C, he must set aside the sub-goal of putting the hands in the middle C configuration and pursue a *third* subsidiary goal, that of 'clearing' the keyboard (P3). P4 is the production rule pertaining to this third goal (indented two steps) and its execution allows the goal to be achieved. The system then passes 'control' back to the second goal and now P2 pertains. This allows goal achievment, and finally control is handed back to the original goal, whereupon P5 can come into force and the task is completed. In a 'realistic' system, the goal of playing middle C would itself be a sub-goal of a larger process, and so when it had been achieved, control would pass up to the next higher, and so on. We can represent the goal structure inherent in our little example by reference to a diagram (Fig. 6.1). The goals are enclosed by boxes, and the numbered arrows show how the procedural steps pass control between one goal and another.

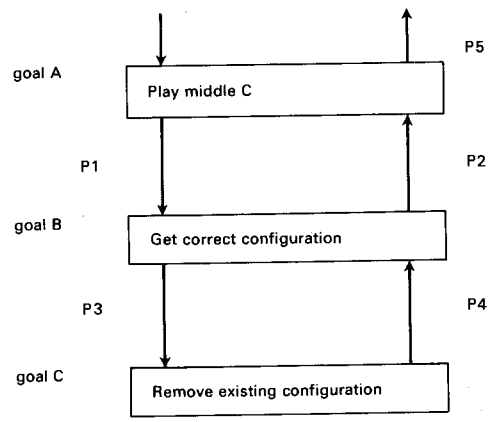



Fig. 6.1. Diagram of goal structure of production system for playing middle C.

It can be seen that a crucial aspect of a Production System is some means of remembering higher goals while sub-goals are being pursued. For instance, it would not be of much use to our clarinet player if he got to the point of executing rule P4 (clearing the keys) only to discover that he had forgotten what he had cleared the keys in order to do. What efficient performance

requires is some form of *goal stack*. The idea of a goal stack is very simple, and based on the mechanical analogy of the spring-loaded plate holders that one finds in canteens and restaurants. The design is such that as more plates are put into the holder the base drops so that the top plate is always just flush with the lip of the holder. When a plate is removed, the next highest plate 'pops' back into view. In this analogy the primary goal corresponds to the first plate put into the holder. The next plate is the first sub-goal, which 'pushes' down the original goal. The second sub-goal pushes down the previous ones and so on. When a particular goal is achieved it is jettisoned from the stack and forgotten, and the preceding goal 'pops' back into control.

When executing a complex skill it should be apparent that the goal stack can easily become very large. Capacity for holding goals in memory is not unlimited, and it often occurs, especially in early learning that the goal stack becomes overloaded, and some higher goals are lost. Some varieties of 'absent-minded error' can be attributed to goal forgetting (Reason 1977), such as going into a shop and then not knowing why one went in. In music performance, goal-forgetting can help explain why learners sometimes find it difficult to 'manage' large-scale form by appropriate long-term changes in dynamic and other performance qualities. They have so many subsidiary goals to fulfil in the execution of individual notes that the goal stack overloads, and the overall 'direction' of the performance is lost. We shall return later to the issue of how increased experience allows such difficulties to be overcome.

Another crucial aspect of a production system is that each rule operates entirely automatically (with certain restrictions which need not concern us here) if it finds evidence that its *conditions* pertain. When it does find such evidence, then the actions that it prescribes automatically take place. It does not seem to matter how many *other* production rules the system contains; as soon as the conditions for a particular rule pertain, then its actions are immediately and automatically carried out. The capacity of the human organism for skill learning is apparently unlimited.

The place where production rules 'look' to see if their conditions pertain is called *working memory* by Production System theorists. A close lay approximation to this term would be 'consciousness'. Working memory contains details of aspects of the external environment that one is currently observing. It contains items of declarative (factual) information that have recently been received or retrieved from memory; the current goal stack; details of one's own internal or bodily state; and like things. It specifically does *not* contain all the facts and memories that are stored in the mind and retrievable *somehow*. For instance, most readers will know the name of the city that is the capital of France. It is, of course, Paris. However, until the moment that the topic was raised, the fact that 'Paris is the capital of France' was not part of working memory—it was not in conscious awareness. Rather it had to be retrieved from 'long-term' memory, and the retrieval process

took time. In the case of this example, the time taken was probably not very long; but for less well-known facts, the time could be considerable, and indeed the search could fail on a particular instance.

The conditions of production rules are, therefore, contents of working memory. They can be facts about the observable world. They can be items of factual knowledge retrieved from long-term memory. They can be current goals (i.e. goals that are at the top of the goal stack). The actions of production rules can involve overt behaviours; but they need not do so. What they must do is alter the contents of working memory in some way. An overt behaviour will normally do this because it will result in an observable change in the environment or bodily state. However, other kinds of action are just as common. Sometimes some mental operation is demanded. For instance, in a production system for mental arithmetic, a rule might specify that two digits, both in working memory be added together, and their sum recorded in working memory. A third type of action involves either adding a new goal to the goal stack, or noting that a goal is achieved and jettisoning it from the stack. In our simple example, the rule P1 involves no overt behaviour. Its only action is to set a new goal.

Thus, although production rules have several similarities to the 'stimulus-response' link beloved of learning theorists in the behaviourist tradition, they are more sophisticated and versatile in many ways. They directly incorporate inner mental states and goals rather than trying to explain them away.

The production rules possessed by our hypothetical clarinet player incorporated specific knowledge about several things, including the finger configuration needed to play middle C. At one time he did not have such specific knowledge. It follows that these rules must somehow have been *constructed* as a result of his ongoing experience with clarinet playing. How were these rules added to his knowledge?

A primary postulate of Production System theory is that knowledge can only affect behaviour if there is a production rule that can act on it. Another way of putting this is to say that all behaviour is procedural, and that a fact cannot influence the course of behaviour unless there is a procedure that can use it. By analogy, although a book contains many ideas, those ideas can have no influence on the course of history unless someone reads those ideas and acts on them. Earlier on, we said that the passage from the cognitive to the associative stage involved acquiring procedural knowledge from declarative knowledge. At the outset of learning, there were no production rules which *directly* incorporated knowledge about the fingering for middle C. It would be wrong, therefore, to imagine that the fact 'the fingering for middle C is . . . etc.' could directly evoke the required behaviour. Rather, there must have been some more general procedures *already available* which could pick up and use the new knowledge. In Production System terms, there must have existed a set of production rules with 'empty slots' into which specific pieces of information could be fitted.

Perhaps the easiest way to get a grasp on this notion is to write out part of the set of production rules which, hypothetically, could have been present at the initial cognitive stage of performance. The terms (VA) and (VB) represent the empty slots which can be filled by a variety of specific values. In formal language, VA and VB are *variables*.

| | | |
|---|---|---|
| P6. | IF | the goal is to play a note (VA) |
| | and | information of the form 'the fingering for note (VA) is configuration (VB)' is not in working memory |
| | THEN | the goal is to place information of the form 'the fingering for note (VA) is configuration (VB)' in working memory |
| P7. | IF | the goal is to play a note (VA) |
| | and | information of the form 'the fingering for note (VA) is configuration (VB)' is in working memory |
| | and | the fingers are not in configuration (VB) |
| | THEN | the sub-goal is to achieve configuration (VB). |
| P8. | IF | the goal is to play a note (VA) |
| | and | information of the form 'the fingering for note (VA) is configuration (VB)' is in working memory |
| | and | the fingers are in configuration (VB) |
| | THEN | blow |
| | and | the goal is achieved. |

P6 is an expression of the knowledge that if one wants to play any particular note and one does not know the fingering for that note, one must discover the fingering before one can proceed. The complete sete of production rules would thus incorporate a set of procedures for getting the required knowledge into working memory (by scanning long-term memory, consulting the manual, attempting to work it out from first principles, etc.). Only when this is done can P7 come into force. It is worth pointing out that P6 is a general rule which would help *any* instrumental learning to get off the ground. It is not specific to clarinet. It is, however, specific to music in its application, but is related to an even more general class of rules which can be obtained by turning *all* the specific values in the rule ('play', 'note', 'fingering', etc.) into variables. Roughly, such a rule would state that if you want to achieve a goal and you do not know what state of affairs would bring about that goal, then you must undertake to discover what that state of affairs is. Such a sentiment may seem trivial, but it remains true that without a procedure that embodied it, people would never be motivated to search for solutions that were not immediately obvious. In such a situation the problem would simply be abandoned.

We see, therefore, in most specific examples of learning behaviour the application (through specification of variables) of a set of very general procedures or strategies for setting about learning. With these strategies, it is

possible to make some kind of coherent attempt to master any task. Newell and Simon (1972) have constructed a computer system embodying these principles which they call 'General Problem Solver'. We cannot pursue the characteristics and capabilities of the General Problem Solver in any depth here. Suffice to say that it incorporates the general feature of goal structuring; breaking down a larger goal into a set of sub-goals to be achieved on the way.

It is pertinent to ask where these general strategies come from. Are we born with them, or do we have to acquire them? Although no-one has constructed a detailed answer to this question, it appears that we are born with a set of very specific production rules which allow us to respond instinctively to certain states of affairs. A young baby does not, for instance, realize what is trivially obvious to you or me, that if you lose sight of something and want to find it again, you increase your changes of doing so by searching for it. Rules of this generality are acquired gradually through childhood by some kind of process of abstraction from specific successes. The mechanisms for this abstraction process are little understood, although it is clear that humans have a particular gift for doing this, and that language is an important mediator of the process. What we can say with certainty is that when a child of 10 or so tackles a new learning problem, he or she has already acquired a large set of general abstract procedures which can be brought into play.

If we return to our specific production rules, it will be easily seen that P7 is a general version of P1, and that P8 is a general version of P5. By a similar process of replacing specific notes and fingerings with variables the corresponding general versions of P2, P3 and P4 may be constructed. This *general* set would, in some ways, seem much more economical that a set *specific* to middle C. This is because, for clarinet playing to be possible, one would need a different specific set for *every* note on the clarinet. In contrast the one *general* set will do for the whole instrument. This economy is, however, a false one precisely because the general rules will not work until specific fingering information has been found. If such information is not available, then the set of procedures which a rule like P6 would initiate might be very complicated, and uneconomical of time or effort. Under the right conditions, therefore, which we shall specify in more detail shortly, there is every reason for a learner to construct the large number of specific production rules which will absolve him from the need to search for particular information in the service of more general productions.

What, then, are the conditions under which a learner adds production rules like P1–P5 to his repertoire, where before only general rules like P6–P8 existed? Production System theorists are only just beginning to formulate a detailed answer to this question (Anderson 1982) but we can already see the general outlines of an answer. Such additions to knowledge occur when a general production rule is repeatedly and successfully used in conjunction

with specific factual information. The cornerstones of any procedural learning are thus *repetition* and *feedback*.

A naïve view of the role of repetition is that it somehow 'stamps in' new learning, as repeated swings of the axe gradually bite further into a piece of wood. The cognitive view espoused here is a little more sophisticated. We suppose that an adaptive system is necessarily conservative, and will not add a new *specific* production rule to its repertoire unless there is some evidence that it is going to be generally useful. This is because, once inserted into the system, a production rule is immensely powerful. When its conditions are fulfilled it automatically 'takes control'. The system has to be quite sure that it wants to give so much control to a single rule before incorporating it. The best evidence a system can have that a rule is going to be useful is that the situation which demands its application has occurred very often. In many 'natural' learning situations the repetitions which suggest the usefulness of a new procedure to the organism are supplied by the natural environment as the organism interacts with it in the pursuance of its fundamental goals. On the other hand, it is possible to self-consciously construct an environment in which one's cognitive system is presented with many more repetitions of the same circumstances than would otherwise occur. In this way one can 'force feed' the learning system. This is not to entirely distort the normal patterns of human learning. Children are often spontaneously repetitive in many aspects of behaviour, and the urge for repetition may well be part of a natural tendency which has survived in part because it is so beneficial to learning.

Pedagogic techniques of rehearsal, exercise, etc. are extensions of the natural repetitive urge rather than totally alien impositions. None the less, many experts will testify that rapid progress is often achieved only by degrees of repetitive practise far in excess of what is pleasant or intrinsically rewarding. To achieve culturally valued goals a learner must often find ways of making intrinsically disagreeable effort enjoyable, or at least bearable. Much of teaching and learning technique is concerned with this problem.

*Feedback* is also essential to an adaptive sytem in order to prevent unsuccessful or potentially damaging productions being formed. The learner must have a way of discovering the success of the application of a specific piece of declarative knowledge in a general production system. If it is repeatedly successful, only then will it become proceduralized. As Anderson (1982) puts it:

> if a new piece of declarative knowledge proves to be faulty it can be tagged as such and so disregarded. It is much more difficult to correct a faulty procedure. As a gross example, suppose I told a gullible child If you want something, then you can assume it has happened. Translated into a production it would take on the following form:
>
> IF the goal is to achieve X
> THEN the goal X is achieved

This would lead to a perhaps blissful but deluded child who never bothered to try and achieve anything because he or she believed it was already achieved. As a useful cognitive system he or she would come to an immediate halt. However, even if the child were gullible enough to encode this in declarative form at face value and perhaps even act upon it, he or she would quickly identify it as a lie (by contradiction procedures he or she has), tag it as such, and so prevent it from having further impact on behaviour and continue on a normal life of goal achievement.

Feedback comes in several ways. In some instances failure is indicated rather directly by a complete breakdown of behaviour or a manifest failure to achieve a goal. In our clarinet example, if a particular configuration is physically impossible to take up then this is dramatic evidence of something very wrong with the 'knowledge' with which one is operating. At the other extreme, a learner has no way of knowing whether he has succeeded other than by being told so. A good example of this would be the early stages of 'ear training' when a particular chord or interval is to be identified. The learner hears the chord, makes an attempt to identify it, and then receives the 'correct' answer from a teacher or manual. Most situations are intermediate between these two extremes, in that there *is* 'internal' feedback available to the learner, if he can notice it. A teacher can help by directing a learner's attention to such feedback.

When a production rule of a general sort repeatedly calls up a piece of specific knowledge and is repeatedly successful such specific declarative knowledge becomes incorporated into a specific version of the production rule. There is some evidence, however, that the cognitive system has a tendency to incorporate into a production rule *any* part of the contents of working memory that has appeared there on every occasion of the rule's operation, even if it is, in fact, irrelevant to the rule's success. This can be offered as an explanation of why it is often claimed that *distributed* practice is better than *massed* practice. To distribute practice is to spread it over several separate periods rather than massing it all together in a single session. Thus, if one has seven hours to devote to practice in a particular week, it may be better to do one hour on each of seven days rather than seven hours on one day.

In Production System terms, massed practice encourages the formation of production rules containing spurious conditions associated with the particular practice period in question. These conditions have to be fulfilled before the production operates. To take a ludicrously exaggerated example, we might imagine someone who only ever practised at night forming the production rule:

IF the goal is to play C
and it is night-time
and the fingers are not in the configuration Z
THEN the sub-goal is to achieve Z

This rule would work only at night! We call learning of this sort 'context'

or 'state' dependent. Music teachers and others often observe the effects of such learning. A pupil who is shown a fingering pattern or phrasing suggestion for one passage may fail to transfer it to an exactly similar sequence elsewhere.

What distributed practice seems to do is to allow more opportunity for context to be varied so that only essential and not spurious conditions get incorporated into a new production. In fact, one of the important consequences of persistently practising a skill for a very long time is that it becomes more and more 'decoupled' from particular contexts.

After relevant production rules have been added in the associative stage of learning, there follows a period of gradual improvement of the skill which Fitts calls the 'autonomous stage' in which performance gradually becomes faster and more fluent. There are several reasons why this might be. One is that several simple productions become merged into a single production sharing a common goal. Therefore, instead of a sequence of productions applying in turn, each taking some finite time to be found and used, only one production is needed.

For instance, if the goal is to play an arpeggio pattern (say C E G C in C major) then one could imagine an early stage of learning where there were four separate rules, one for each note of the arpeggio. Thus the first rule would apply if and only if the first note had not yet been played, and so on. If such rules were constantly and successfully applied in sequence, the cognitive system would eventually 'couple' them together so that the condition of having the goal of playing the arpeggio would, within one rule, entail the action 'play note 1 *then* note 2 *then* note 3 *then* note 4'. In traditional terms, the arpeggio would have become an integrated and automated performance pattern.

Such a 'compilation' could not take place, of course, until the performer were capable of *having* a goal such as 'perform a C major arpeggio'. This would in its turn depend upon having already understood what a C major arpeggio is, either by demonstration or by inference from its repeated occurrence in music. In performance from a score, a performance production rule such as the above would need a corresponding *perceptual* production rule which could *recognize* a C major arpeggio as being present on the page. Early instrumental learning is often hard because one is learning the perceptual and motor 'halves' of the skill at the same time.

It should be emphasized that the forming of a production rule for a performance pattern such as a C major arpeggio is not to specify every performance parameter. The performance will *not* be a rigid 'fixed action pattern' which comes out with the same timing and dynamic parameters each time. Rather, the rule specifies a series of sub-goals (of playing the various notes in turn) each of which may call up new procedures which determine the precise parameters for the notes according to transient information in working memory.

Another aspect of the 'autonomous' stage is the acquisition of further specific production rules which relieve the load on the goal stack. Consider once again the little production system for playing middle C (P1–P5). We observed that it had three levels of goal stacking. The third goal (removing all the fingers) was set by P3. We could eliminate the need for P3 (and thus P4) if we were able to incorporate a rule for going from each specific other finger position to the desired one. One such rule might be:

P9. IF the goal is to achieve the configuration of having the left-hand fingers covering the top three holes and the left-hand fingers are presently covering the top two holes
THEN cover the third hole and the goal is achieved.

To get to middle C from any other note would require about 30 such rules. To get from any note to any other note would require about 900. It is not unreasonable to suppose that the expert clarinet player has all 900 of these rules in his production system. The time and effort required to get them all there is great, but it is worth it for the elimination of one goal from the stack, and the consequent lightening of the load on working memory. Typically, such new productions will be acquired piecemeal, so that common transitions, such as moving from one note to the next member of a scale or arpeggio, are acquired early on, at a stage where the learner still reverts to P3 and P4 for more unusual transitions.

The task of writing a production system for a non-trivial aspect of musical skill has not yet been undertaken; nor yet has the even more crucial task of modelling some aspect of musical learning in detail. It is not clear, at this stage, whether the theoretical notions available are fully up to the job of music learning, and I do not wish to pretend otherwise. It does, however, seem to me that there is no other theoretical framework currently available which allows such detailed formalization of cognitive processes but at the same time supports many general but more vague conceptualizations and observations about learning. We thus move away from detailed analysis of production systems to some more general remarks which such an approach pulls into focus.

It has long been clear to insightful teachers that one does not create learning simply by offering people facts or prescriptions. Such offerings are of no use to people unless they can incorporate them in procedures they already have. Otherwise the facts remain insulated from any real influence on skill. Teachers in all spheres of activity, including music, often experience the sense that they are not 'getting through' to a learner. Facts are acquired and reproduced, yet there is some sense that the learner has not really grasped the essence of what is being said. For instance, I have had piano pupils whose playing has somehow been wooden and lifeless. I have talked to them about tonal variety, planning *crescendos* to points of accent or climax, rubato, etc.,

and at a 'declarative' level it has all been understood, in that they can provide definitions of what terms like *rubato* and *accent* mean. Furthermore, they have been able, under instruction, to obey the letter of these prescriptions in particular contexts. A pupil took my advice that a particular cadential chord should be accompanied by a holding back of the tempo; but the application of the advice achieved a grotesque result, an arbitrary 'hiccup' in the music, rather than a measured pause which emphasized the importance of the chord, but allowed the music to continue. It was clear that my pupil had no sense of the 'rightness' of my advice in virtue of which it would be possible to tailor it to the demands of the specific context. I do not know whether or how, such sensitivity can be taught. All a teacher can do is watch diligently for signs of it, and when it appears, encourage it and build on it. With all learners and at all levels there comes a point when a tactful teacher must just pass over a failing, knowing that, at this juncture, further comment will not improve it.

The greatest determiner of the progress of any piece of learning is the learner himself, the procedures he has available, and the motivations he has. To use an agricultural analogy, a teacher is like a gardener; tending, watering, pruning, and 'training' his plant. Without his attentions the plant would still grow, possibly not so straight and tall; but it remains true that the gardener cannot alter the essential form and nature of the growth. This is also, to a certain extent, true of the learner himself. Even in a self-consciously 'engineered' learning environment the learner cannot totally control his progress. He finds some things easier to grasp than others, he finds he has just not got the motivation or concentration to cross a particular hurdle, his recalcitrant fingers will just not take up the patterning he wants. And so, in many respects, the learner *observes* his own progress, sometimes with delight and sometimes with despair, rather as an outsider would.

In Production System terms we may say that a learner's conscious efforts tend to alter the contents of the 'declarative' part of his knowledge, but cannot actually get 'down among the production rules' where it really counts. He must patiently wait for the production system, operating according to its own laws and pace, to pull down from declarative knowledge that which it can make its own.

The engineering of learning is, thus, a hit-and-miss affair. No-one not even the learner himself, can be fully aware of the automatic thought procedures which form the basis of his competence. It is impossible, therefore, for anyone to plan the 'ideal' training diet in a way which is exactly tailored to the competences of the individual concerned. Rather, teachers and learners evolve broad strategies which, on the whole, seem to produce results. This helps to explain the huge gap between cognitive theory and learning/teaching practice. It is impossible to derive foolproof prescriptions for a situation in which, almost by definition, the crucial determining factors (the existing set of production rules within an individual) cannot be easily known. What good

teachers (and good learners) seem to do is build up a stock of commonly observed facts about learning environments and schedules which, on the whole, work for them. As a learner becomes more expert, the more likely are such conditions to be specific to that individual, because the shape of his knowledge is different from anyone else's. Our best hope of deriving generalizations would seem to be at the beginning of training, when one may hope that learners share a reasonably common heritage of procedures and motivations supplied by the interaction of biological propensities with the experiences of enculturation.

Thus, for instance, it seems true for our culture that most well-motivated children learn to read if they are supplied with some means of acquiring information about the sounds of letters and words together with a stock of suitably attractive books. The motivation to be able to read their *own* stories, when they want, seems strong enough for children to effectively teach themselves. Teachers assist the process by providing information and feedback when it is needed, and by some elementary structuring of sequence, so that, for instance, the words that a child first meets relate to well-known objects and ideas in his experience. The main conclusion to be drawn from a massive amount of research into teaching methods seems to be that, for the majority of children, it does not much matter what official 'method' is adopted, they learn to read just as well. The main factor determining success seems to be the individual teacher (Williams 1970). It has been observed that two teachers from the same school, with the same type of children, can produce consistently different results, although using the same 'method'. It is not unreasonable to conclude that good teachers are those who provide, reliably and efficiently, the information, feedback and encouragement that children seek as they go about the business of becoming skilled at reading. This can often boil down to such basic things as providing the kind of environment in which a child who wishes to pursue some learning is not likely to be interrupted by other children.

In music an early motivation is often the wish to be able to reproduce the musical sounds of the culture. In more cases than one might expect, the mere presence of a musical instrument in a home can allow a child the opportunity to learn some idiosyncratic music making—picking out combinations and sequences which sound familiar or interesting. In some cultures, where improvisation dominates, such idiosyncratic beginnings can, under their own momentum, produce an expert performer of some distinction, whose technique and style may be unusual, but effective for the music he chooses to play. In our own culture, where the emphasis is on reproductive performance of significant 'art' compositions, teachers feel a duty to steer their pupils towards the kind of technical and formal grasp which will eventually, enable a culturally acceptable performance of, say, a Mozart sonata, or a Chopin Etude. In this case, it is necessary to introduce learners to ways of holding their hands and moving their fingers which will at first, seem very unnatural,

but which are necessary if they are to execute fast passagework with the kind of fluency and control that the reproductive goal requires.

This contrast is exemplified by the difference between a largely self-taught jazz pianist and a graduate from a music conservatoire where a 'rigorous' technical approach is taken, and where learners are coached on the same relatively small concert repertoire that will enable them to participate in international competitions. The former will often have a technique which would be perceived as 'hopelessly awkward' by a trained concert pianist, yet it serves its purposes well. The latter will usually have a technique which is very similar to other pianists trained at the same school, often to the point where individuality becomes submerged. A 'balanced' relationship between teacher and learner will probably steer a middle course between these two extremes. The teacher will try to shape and mould what is already present without 'going back to square one', and the learner will incorporate, and make his own, what the teacher is offering.

If one accepts the way of looking at learning that I have attempted to outline in this section then it seems to me necessary to accept that the training (of oneself or others) is more an art than a science. However much information scientific research and theory can supply it will never allow a total knowledge of the learner's cognitive system, because each learner is unique. No matter how much scientific information is available, teachers and learners must still make decisions on a largely intuitive basis. I confess to experiencing some relief that this conclusion may be drawn on theoretical grounds, because if one *had* to rely on research in music education it would be hard to find a consistent and universally applicable set of findings.

Consider, for instance, the question of how to correct habitual errors in performance (such as the execution of a wrong note or a wrong fingering in a passage of music). Buck (1944) has characteristically forthright advice on this point:

> If you work at a passage where the difficulty is the fingering, your only chance is to decide on how you will finger it, and work on it with concentrated thought until it is automatic. It may mean a thousand repetitions. But if, after fifty of them you are careless and and stumble, you have not merely done it wrong for once; you have wiped out all the benefit derived from probably twenty-five of your trials (p. 20).

On this view, the only way to correct error is to play the correct fingering a very large number of times. This is intuitively plausible. However, an extraordinary study by Reitmeyer casts serious doubt on this conclusion (Reitmeyer 1972). In this study, some subjects were asked to indulge in *negative* practice. This involved deliberately practising the error, being aware all the while of the correct fingering. Other subjects practised in the 'conventional' way, playing the correct fingering. The two groups of subjects were then asked to provide repeated correct performances as a measure of the extent to which the error had been corrected and eliminated. Reitmeyer

found that there was no significant difference between the effectiveness of the two techniques.

Without a theoretical framework, such a result is practically uninterpretable, and leads to such weak conclusions as those given by the author, e.g. 'instrumental students and teachers might consider negative practice as an occasional alternative to positive practice'. On the face of it, the result renders Buck's statements false. Yet I would be loath to alter Buck's advice. It is clear that positive practice is *a* road to success, and one which is supported by teaching tradition and intuition. Even if negative practice had turned out to be *better* than positive practice in Reitmeyer's study, I would still be loath to recommend it, since I would imagine that learners would find it very hard to live with, and believe in, such a counterintuitive strategy.

As a cognitive psychologist, I am, of course, fascinated to know how it is possible for negative practice to work. It is possible that repeated attentive *thinking* about the right fingering is more important than actually *playing* the right fingering. This is supported by Neisser's (1983) studies of dart playing. He found that his subjects improved if asked simply to *imagine* themselves throwing darts. In Production System theory, it is what is going on in working memory (or consciousness) that determines the course of learning. Because Reitmeyer's subjects were actively labelling what they played as 'wrong' and another, unplayed, fingering as 'right', we may conclude that production systems picked up only that information labelled as 'right'. We may contrast this case with another outwardly similar case where someone plays the wrong note, but does *not* simultaneously think of the right note in a focused manner. In such a case, it would be *really* extraordinary if there were any improvement. Unfortunately, a teacher cannot easily know whether a learner is or is not thinking of the correct note whilst practising the wrong note. Negative practice is, therefore, an educationally disastrous technique, precisely because an observer cannot check whether it is being carried out effectively. Positive practice has the advantage that the learner *must* be thinking of the correct note in order to play it (unless, of course, he already knows the passage so well that it is automated; in which case there is no learning problem).

The problems that arise when one makes any attempt to derive specific training prescriptions from educational research have been illustrated through this (not atypical) example. This is not to suggest that such research is without value. However, it seems to me that the most useful research is that which offers the musician a general resource which he can use to suit his own purposes rather than a detailed method. One area in which this condition is fulfilled is the development of new technology which offers a musician the opportunity of enhancing the feedback available to him.

For instance, Tucker, Bates, Frykberg, Howarth, Kennedy, Lamb, and Vaughan (1977) describe a computer-based interactive aid for musicians that allows a passage played on the keyboard of an electronic organ to be dis-

played visually in a modified notation that reflects the exact duration of the notes played. With the aid of this feedback, a pianist of concert standard improved his performance of a triple trill quite dramatically within minutes. Similarly, Basmajian and Newton (1974) showed that with visual electromyographic feedback, wind players learned within minutes to suppress or activate specific parts of the buccinator muscle in the cheek. Singers can also be helped in a similar way (Fourcin and Abberton, 1971).

### 6.4. Assessing musical ability

We finish this chapter by briefly considering 'tests of musical ability'. Within musical cultures where there is an organized pedagogic tradition, achievement in specific musical skills such as performance and theory are measured directly through examination. Apart from the problems caused by 'nerves' and 'stage fright' such examinations seem to fulfil their purpose in a direct and unexceptionable manner. We shall not consider them further here. The purpose of 'tests of musical ability' are rather different. Whilst examinations presuppose intensive preparation of specific materials, tests of ability involve no foreknowledge of test content. Indeed, such tests are invalidated by extensive practice on the particular tasks they contain. This is because of the rationale which underlies their construction.

The argument is that all children within a culture experience a relatively common exposure to, and involvement in, music, at least to begin with. However, because of differences in innate potential, motivation, or experiences at certain critical periods, the 'take-up' of musical knowledge varies from child to child. Thus, even before some children are selected for specialist training, they have come to differ from one another quite widely in the stock of underlying skills and sensitivities they possess. Sometimes these differences are not immediately apparent in the overt musical behaviour of the children. For instance, there may be differences in the ability to listen to and appreciate music. Or again, a child with relatively few musical skills could have received instrumental coaching which provided him with an overt accomplishment not displayed by a child without such coaching who, nevertheless, had many more 'unnoticed' skills rather than the ones which are on the official training agenda. A child who possesses these skills in a measure above those of his age-mates is more likely to make rapid progress under formal training and is more likely to achieve excellence. A child who lacks these skills is unlikely to make such rapid progress, will require more training effort to achieve excellence, and possibly has passed some critical period of development after which it is unlikely that excellence can be achieved.

Without questioning this line of argument, let us look at a few typical situations where such tests might be of use. One example might be the case of a school with a very good choir that gives frequent public performances and broadcasts (such as many English choir schools). Members of the choir

spend many hours a week rehearsing at the expense of other school activities, and the majority of the children selected for training must reach the required level within a relatively short time if the numbers in the choir are to be maintained to take account of the older children who leave the choir every year. Very few potential entrants have had intensive singing experience. After ensuring that candidates can sing simple melodies pleasingly and with good intonation, the choirmaster makes the final selection on the basis of a test of musical ability.

A second example might be the case of an older child, in early or mid-teens, who becomes seriously interested in music after years of relative neglect. He wonders whether he would have a chance of succeeding in a musical career. He attends vocational counselling and is given a test of musical ability. His score is well below the average for his age, and he is advised that he will be unlikely to cross the selective and competitive hurdles that would gain him a place at a music college.

These two examples have in common that they are being used to assist some kind of educational selection. Although the principle of selection can have unfortunate and undesirable consequences, there are benign aspects too. It does a child no good to allow him to embark on a course of training for which he is temperamentally or intellectually unsuited, whatever his current 'wants' might be. Given that some selection is inevitable and possibly helpful, we may ask if and how 'tests' can be used to make such selection as fair and accurate as possible.

First, a test should only be used when there are no more direct signs of achievement to examine. Instrumental or vocal performance which shows technical and expressive mastery provides better evidence of musical ability than any test can do. In both our examples, the sensible teacher or counsellor would have ensured that there was no direct clinching evidence of accomplishment before embarking on testing.

Second, test results should be taken in conjunction with other evidence, formal and informal, where possible. For instance, a child who performs well on a test but does not seem particularly interested in, or excited by, music may still be unsuited to an intensive musical education. It is always possible, for instance, that over-keen parents have discovered the general nature of the tests used, and obtained special coaching for their child on the types of task to be tested.

Third, testing should only be carried out with respect to a particular educational question to be answered at a particular time, not to provide a once-for-all statement about capacity or potential for achievement. In the choir example, the tests used to select those children who, within a few months of testing, would be likely to be able to 'hold their own' in a fast-moving training programme. For the teenager the question was whether he would be able to raise his achievement level to competitive examination standard within a few years. In neither case should relatively poor results

lead to the conclusion that a person is 'unmusical' or 'incapable of musical achievement'. The person concerned is simply judged unlikely to meet *particular* achievement targets within *particular* limited time-spans.

Fourth, the test should be a *valid* measure of musical ability. Validity may be assessed in two ways. A test has 'face' or 'content' validity if it manifestly demands use of the ability it is purporting to test. A test which involves a subject hearing, and then singing back, a short melody has high content validity as a test of musical ability. A test which requires a subject to remember a list of numbers has low content validity. A test may be valid in a second way, however, even when content validity is low. This second type of validity is called 'associative validity' and arises when there are good correlations between a test and some other measure of achievement. One way of assessing associative validity is to see how well one test result correlates with some *other* test of ability. This is often a little circular, because this still leaves one with the problem of assessing the validity of the other test. A better method is to see how test performance correlates with future achievement. If there is a high positive correlation, the test is said to have good 'predictive validity'. Thus, if ability to remember numbers correlated highly with later musical achievement, it would be a valid test of musical ability. The problem with tests that have only predictive validity is that one has to place one's whole faith in the adequacy of the studies which purported to establish the validity. There may be any number of reasons why the conditions which pertained in the validation study do not pertain in a particular testing situation. For this reason, it would seem advisable to use tests which, where possible, have high content validity. If evidence exists of high associative validity, so much the better. However, it is worth saying that it is often quite difficult to obtain really convincing estimates of predictive validity. To obtain such estimates requires a non-selective use of the test with a large population, and a subsequent follow-up study to assess achievement. Ideally the validation study should be carried out in an educational context similar to the one in which the test is being proposed for use. Thus, in our choir school example, the 'ideal' validation would involve the school selecting children for entry to the choir on some basis other than a musical test. Meanwhile, the specific test of musical ability in question would be administered to all the children by a third party, and the results kept secret. All the children would then undergo normal training, and at the end of the training period their achievement assessed by conventional means. Then the test results would be unlocked and their predictive power assessed. In reality, few schools could afford this sort of experiment.

Fifth, the test should be reasonably *reliable*. A test is reliable if the same subject achieves the same score when tested more than once. In practice no test is 100 per cent reliable, and the scores of an individual will fluctuate from testing to testing. Much of this has to do with concentration and other personal factors. A test will tend to be unreliable for a given individual if a

lot of the items are just at the limits of his ability. On a 'good day' he will just succeed at all these items. On a 'bad day' he will just fail. Good test preparation can reduce this problem. A standard technique is to have a series of items of graded difficulty, starting with very easy ones, and moving through to very difficult ones, so that all subjects will find several items that they can always get right, and several that they always get wrong, good day or bad day alike. What will vary from test to test are the scores on the few items which lie around his 'cut-off' point. Because no test is completely reliable, an experienced tester will resist the temptation to make distinctions between people on the basis of small differences in test result. Tests should only be used to place people in broad categories (such as 'top quarter' or 'bottom third'). Even this does not make a test completely fair. Suppose, for instance, one can only take 25 per cent of a given population, and one uses a test as the sole criterion for selection. There is no coherent alternative to simply selecting the highest scoring quarter. This may well mean accepting someone with a score of, say, 62, and rejecting someone who scores 61, even though it is entirely possible that, on another testing, their positions would have been reversed. There is no entirely satisfactory solution to this vexing problem. One flexible and humane response is to abandon rigid quotas and examine the distribution of test scores. Quite often one can find that the scores 'cluster' around certain values, with definite gaps between the scores of different groups. In such a case it would considerably increase the fairness and reliability of the selection to place the 'cut-off' point in the largest convenient gap, even if this meant taking rather more, or rather less, than quota. Another response would involve using test scores as one among several selection criteria, so that 'borderline' cases could be distinguished on some more discriminating criterion.

There are some 24 tests of musical ability documented in the recent literature. Shuter-Dyson and Gabriel (1981) provide a comprehensive description and discussion of all these tests. All of the tests have basic content validity. That is, they all require children to listen to, and make judgements about, short extracts of music. None of the tests require overt musical performance such as singing or beating time. The most common type of test item is a pair of notes, chords, or sequences, which may or may not differ in some respect. The subject is required to say whether the members of a pair are the same or different, or to make a simple judgement about the difference ('which is louder?', 'which is better?', etc.). All of the tests comprise two or more sub-tests, examining different aspects of music. The same types of sub-test occur in many of the tests. One of the most common sub-tests, for example, involves the subject hearing two short melodies of identical length. In the second melody, one note may be changed in pitch from the first melody. The subject has to judge whether the melodies are the same or different, and if different, where the different note lies in the sequence.

Differences between the tests lie in the selection of sub-tests used, the

age-group for which the test has been designed, and the efficiency of the standardization procedures and reliability and validity studies. Several of the tests are not commercially published, and others have gone out of print. The two tests most readily available and most widely used (at least in Britain) seem to be the Seashore Measures of Musical Talents (Seashore, Lewis, and Saetvit 1960) and the Wing Standardized Tests of Musical Intelligence (Wing 1962). For a detailed evaluation of these tests see Sloboda (1984).

Both the Wing and Seashore tests have been designed for use with children of age eight and upwards. As we saw earlier in the chapter most eight-year-olds have an enactive repertoire of songs they have learned from the culture, and which they are capable of reproducing with rhythmic and tonal accuracy. They are beginning to enhance their reflective awareness of the structures and styles that the music of their culture contains. They are beginning to differentiate consonance from dissonance, and to gain a sense of harmonic direction and closure. They also have started to detect high-level 'style' similarities. The period up to the age of 12 or so is, in general, one of rapid and significant changes in musical awareness; and most of the observational and experimental studies reported in the first part of this chapter found wide variation in ability at each age. Unfortunately, both the Wing and Seashore tests were designed long before much of this research was published (Wing 1948; Seashore 1919). If one were designing a new test today it would be only rational to test the very abilities that recent investigators have shown to be developing in middle childhood. Thus the test might include versions of Gardner's (1973) style-sensitivity test, Zenatti's (1969) tonal memory tests, Imberty's (1969) harmonic closure tests, and so on.

Some of the Wing and Seashore sub-tests approximate quite well to these research tests. Both, for instance contain a test of melodic memory. The Wing tests contain a test for discriminating conventional from unconventional harmonizations, and metrical from unmetrical accenting. Other tests seem more hard to justify. The Seashore tests, for instance, contain a test of pitch discrimination with pitch differences going down to as little as 2 Hz (about one-tenth of a semitone in the pitch range tested). Research we have discussed in Chapter 5.4 shows that many experienced adult musicians are incapable of such fine discrimination, and that assimilation to pitch/scale categories is a more characteristic (and useful) feature of the musical mind. Manor (1950) found a zero correlation between scores on the Seashore pitch test and success at clarinet or trombone playing. A small positive correlation *was* obtained for success at violin playing; a skill where accurate pitch perception is arguably more important than most.

Similarly, the Seashore tests include a loudness discrimination test which requires finer discriminations than many professional bassoonists seem capable of (Patterson 1974).

It is certainly true that a total failure to make discriminations along sound

dimensions such as a pitch and loudness would render musical achievement impossible. However, we are entitled to question the value of extremely fine discriminations below a generous minimum. The thrust of most of the evidence presented in this book is that musical expertise consists in grasping the various levels of structure that exist in musical sequencing. A test of musical ability or aptitude should be demonstrably concerned with such structural skills. On these grounds the Wing test is to be preferred to the Seashore, and seems to produce better validity estimates than the Seashore (Shuter-Dyson and Gabriel 1981). Even so, there are serious technical deficiencies in the presentation of the Wing tests which makes them hard to recommend wholeheartedly (Sloboda 1984). Both tests are somewhat dated, and need re-standardizing for contemporary populations.

Tests of musical ability have never made the same kind of impact on the music teaching profession as have tests in 'core curriculum' areas such as reading or mathematics. It would be interesting to pursue the reasons for this, but here is not the place. It does seem likely, however, that lack of strong and consistent teacher demand is a major factor behind the relative unsatisfactoriness of the music test market. Without such demand it is unlikely that the funds and research impetus will be found to devise a test which serves the musical community really well. This is not to say that an individual teacher or institution will not find a test of musical ability useful in a particular context. However, it would be wise to treat test results with considerable caution. In particular, it would be foolish, and possibly unfair, to make major educational decisions on test scores alone.

# 7. The musical mind in context: culture and biology

## 7.1. Introduction

The discipline of psychology occupies a curious and rather precarious position in the world of science. It stands mid-way between the 'physical' sciences such as biology and physiology, and the 'social' sciences such as sociology and anthropology. Within its own boundaries it experiences the tension created by the opposite pulls which the 'physical' and 'social' sciences exert. Thus, there exists a sub-discipline of physiological psychology and one of social psychology. Cognitive psychology occupies a middle position between these sub-disciplines rather as psychology is between the biological and social sciences.

The biological approach to psychology attempts to account for human behaviour in terms of the operation of the brain and nervous system, which is, in its turn, influenced by the genetic constitution of the human organism. It is often 'reductionist' in that it seeks to replace psychological entities such as 'consciousness', 'intentions', 'memories' with combinations of physiological entities and events such as activities in certain brain areas. This approach tends to stress the innate basis of much human mental equipment, the universality of human psychological makeup as a result of shared genetic material, and proposes that many of our basic social conventions and institutions have arisen because they further the survival of individuals within the species, and so have been selected through Darwinian evolution. Critics of this approach point out that cultural and historical variation in human behaviour are very wide; and philosphers have argued with some cogency that the attempt to reduce mental events to brain events fails on conceptual grounds.

The social approach to psychology attempts to account for human behaviour in terms of the social institutions and conventions within which a person develops. On this account, significant behaviour thought patterns are learned through interaction with the culture, and they are specific to the particular culture. This approach stresses the *relative* nature of most human abilities, and their dependence on particular social contexts for their execution. Critics of this approach point out that all humans share a genetic heritage, a common physical environment, and certain cultural constants such as upbringing by adults, language learning, and so on. These factors converge to give rise to psychological universals.

As with most crude dichotomies, each approach states something important about man, but fails to encompass the whole truth. Cognitive psychology occupies a vital pivotal position between the extremes because both 'biologists' and 'sociologists' identify individual human thought and action as the 'given' which requires explanation. The 'neutral' cognitive psychologist attempts to articulate the structure of human thought and action in a way that leaves aside the question of cause—biological or social. For this reason, the typical cognitive psychologist is not concerned to examine a wide range of developmental and cultural contexts. Instead he takes a *particular* stage of behavioural development of an activity within a *particular* culture, and asks the question: how can we best characterize the system that is operating at that time and place? On the whole, I have tried to take a relatively 'neutral' line in this book; although authors such as Sampson (1981) argue that such a stance cannot help but offer implicit support for strong assumptions about the nature of man and society.

Whether or not one accepts such criticism, it is clear that neutrality can be, at best, a temporary strategy which leaves vitally important questions untouched. Biological and social influences must be included in a full explanation of human conduct. Accordingly, this chapter explores a few of the social and biological considerations relevant to musical thought and behaviour.

On the social/cultural side we look at some of the factors which may be responsible for cultural differences in music, and ask whether, none the less, there is any evidence of underlying universal features of music. On the biological side we examine some studies which shed a little light on the neural organization of musical function. We also ask whether there is a biological base for the origins of music in our species. Does music fulfil a biological function, and has it got evolutionary precursors in our immediate ancestors?

## 7.2. Culture and musical thinking

The musics of the world contain enormous variety. Just what variety has not, perhaps, been fully appreciated until the present century when Western scholars have begun systematically to study and record music of other cultures than their own.

One important approach to the study of the world's music is essentially taxonomic or classificatory. It is a first step in understanding the relation of music to culture. In this approach the musical sounds of a culture are classified in terms of, for instance, the instruments they use, the typical forms found, the scales and tuning systems used, the social contexts in which the music occurs, and so on. Using such information it is possible to build up a comparative 'world map' of music, grouping together those cultures whose musics seem most similar.

To take one example, it seems possible to divide the continent of Africa

into two main musical sub-continents, with the Southern boundary of the Sahara Desert as the approximate dividing line between the two. In Northern Africa, where Muslim influence has been particularly strong, one can discern features of the indigenous music which links it to the music of the Near East. Here the voice is paramount. There is relatively little instrumental music, and instruments are more often used to provide an accompanying drone or rhythm for a vocal line. The music is mainly monophonic and soloistic, and the solo singer uses a particular type of voice production which gives it (to Western ears) a forced nasal quality. Melodies are based on pentatonic (five-note) or heptatonic (seven-note) scales, but are decorated with microtonal shakes. Social contexts for music making include religious ceremonies and festivals, and the recounting of epic poems. The music of Northern Africa is generally seen as part of a larger 'Pan Islamic' tradition which extends as far up as Turkey and Eastern Europe. The Yugoslavian epic songs discussed in Chapter 4 fall into this Pan-Islamic tradition.

The music of Southern Africa has distinctly different charcteristics from Pan-Islamic music. Here music is rarely soloistic but involves groups of people. A fundamental form is the call-response, in which a solo leader sings some short phrase, followed by an answering phrase from a chorus. The music is usually polyphonic, which is to say that there are often at least two independent, but equally important and co-ordinated, things going on at once. Instruments, particularly drums, are very important, and much music has a rhythmic strength and vitality brought about by the repetition of short, highly complex, multi-layered rhythmic patterns, produced by different drummers simultaneously executing their own rhythmic devices. The tonal content of the music seems more accessible to Western ears than that of Pan-Islamic music. Singing is full throated and less prone to microtonal variation. Harmony is common, and typical chords are built from intervals of major and minor thirds, perfect fourths, and perfect fifths. The social contexts for music are many and varied. Practically any social gathering, whether for work, celebration, or religion, is the occasion for group music making. Nearly all music is text-based, with poetic or song structures lying at the core of musical forms. Even purely instrumental music usually derives from a vocal model.

This much simplified contrast between the music of Northern and Southern Africa gives a flavour of the kind of descriptions that ethnomusicologists have been able to give of the world's musics. Of course, there are many more fine distinctions within small regions than this sketch suggests, and Africa contains no less diversity than any comparable area of the world. I have neither the competence nor the space to offer a survey of the world's musics, even at the most general level. There is a large and detailed literature on all aspects of world music. Much of the fundamental research is published in the journal *Ethnomusicology*, but helpful summaries are provided by Kunst (1959), Malm (1977), and Nettl (1964, 1973). Up to date summaries and

bibliographies are provided in *The new Grove's dictionary of music and musicians* (ed. Sadie 1980).

What I wish to do here is go beyond taxonomy to questions of cause. What factors, cultural or social, can help account for the wide differences between musical cultures in the world? I shall focus on a single factor which, although not the only significant one, seem to me to have far-reaching consequences. It is particularly important because it seems to relate particularly strongly to differences between Western Art music (about which this book has mainly been) and other forms. I shall be suggesting that the form of Western music, and the nature of the cognitive skills it supports, are closely linked to the existence of a developed system for notating music. In doing this I draw particularly strongly on the ideas of Blacking (1976) and Shepherd, Virden, Vulliamy, and Wishart (1977). After this I shall turn to the question of the evidence for 'universals' underlying the surface diversity of music.

My aim is to raise a few of the issues which seem to be important for a broader understanding of musical cognition than can be gained from examining only one type of music. Psychologists have not been noticeably interested in music outside their own culture, and there has been little empirical work carried out within the framework of modern psychological thought. In this section I shall be offering some broad, even provocative, general hypotheses, which clearly need refinement in the light of psychologically oriented research. A psychologically informed evaluation of the vast ethnomusicological literature is also needed. It is only the absence of such major achievments that justifies my rashness in dealing with this vast and complex issue at all.

### 7.2.1. *Music notation as a cultural force*

Many world cultures are, or have recently been, essentially non-literate (or *oral*). That is, they have not used visual notational means of recording details of human transactions. There are many contrasts to be made between literate and oral cultures. I wish to concentrate on four broad proposals about the effects of the availability of notation:

1. The existence of written notation allows lengthy verbatim recall of complex meaningful material.
2. Notation allows proliferation and migration of material so that it exceeds the capacity of any one individual to know it all.
3. Notation encourages the separation of the content of an utterance from its context, and makes it easier for an utterance to be treated as a 'thing in itself'.
4. Notation selects certain aspects of sound for preservation, and, in doing so, both embodies current theory and also tends to restrict the future development of music in certain ways.

First, I will enlarge on these points with reference to language, and then suggest how they may also apply to music. Before doing these things, it is necessary to make some qualifying remarks. Contemporary cultures can be ranged on a continuum from those where literacy is non-existent to those where literacy pervades all aspects of social activity. There are many 'semi-literate' cultures, where literacy exists within restricted classes or situations, and where this has effects which permeate to the oral members of the culture. Even in a predominantly literate culture such as ours there may be individuals who are not literate, but indulge in activity which is supported, directly or indirectly, by the literacy of others. Conversely, individuals who are fully literate can, at the same time, possess significant areas of activity which are essentially oral. Many Westerners are linguistically literate and musically oral. Some musically literate people maintain areas of musical activity, such as jazz or folk music, which are essentially oral. So the literate/oral distinction is not as culturally clear-cut as I sometimes assume for argument's sake.

There is also another complicating factor in the form of new modes of preserving the past, particularly electrophonic and photographic recording. These fulfil many of the functions of notation, but have new properties too. Because these modes of recording are not symbolic, people can use them without acquiring literacy. This tends to increasingly blur some of the distinctions I make, at least when applied to the contemporary scene. With these reservations in mind, we may proceed.

For many people, notation is so important that reality becomes, in many ways, mediated *by* their notations. What can be written down and preserved is *correct* and definitive: man's performance or memory is judged against the written record. In oral culture, present knowledge and memory are the only guides. Disagreements are settled by negotiation, or by recourse to the judgement of respected individuals (D'Azevedo 1962). When members of a literate and an oral culture interact these differing 'world views' can cause mutual incomprehension. Goody and Watt (1963) give a clear example of this (cited in Shepherd 1977):

> Early British adminstrators among the Tiv of Nigeria were aware of the great importance attached to . . . genealogies which were continually discussed in court cases where the rights and duties of one man towards another were in dispute. Consequently they took the trouble to write down the long list of names and preserve them for posterity, so that future administrators might refer to them in giving judgement. Forty years later . . . their successors were still using the same genealogies. . . . However, these written pedigrees now gave rise to many disagreements, the Tiv maintained that they were incorrect, while the officials regarded them as statements of fact, a record of what had actually happened, and could not agree that the unlettered indigenes could be better informed about the past than their literate predecessors. What neither party recognized was that in any society of this kind changes take place which require a constant readjustment in the genealogies if they are to continue to carry out their functions as mnemonics of social relationships.

It is common for many literate people to assume that the life and knowledge of a literate culture is somehow *superior* to that of a oral culture; that literacy adds to man's resources and takes nothing away. It would be more correct to say that literate and oral cultures are *different*, and that there are both gains and losses involved in becoming literate. The advantages are so clear to us that they hardly need spelling out: the increased availability and durability of knowledge, allowing complex and geographically widespread social and knowledge structures to develop. The disadvantages sometimes escape us and need more emphasis. In a completely oral society, gaining knowledge cannot be divorced from the fundamental human interactions of that society. Knowledge is preserved in the very customs and rituals which bind the society together. In so far as the social group is stable, the knowledge it transmits tends to be stabilizing, dependable, and sufficient for life. There is little possibility of a person falling foul of society through failure to know something of importance. The limitation of knowledge to that which is important enough to become enshrined within oral traditions guarantees that each individual knows what he needs.

Literacy, in contrast, involves the possibility that knowledge becomes esoteric and unmanageable. There is so much stored knowledge in a literate culture that it is impossible for any individual to interact with anything other than a small fraction of it. Thus man becomes dependent on other 'experts' to mediate and use knowledge for him. Because written knowledge can be disseminated far from the social context that generated it, the amount of 'useless' knowledge (with respect to a particular individual or group) becomes very large. At best this can mean that a person will waste a lot of time and energy before discovering useful knowledge. At worst it can seriously mislead by supplying knowledge which is simply inappropriate to a person's own situation.

Written notation also encourages us to *distance* ourselves from our words and those of others. When they are captured in permanent form on paper it is easier to analyse, dissect, and treat them as lifeless *objects* distinct from us and our thoughts. In this way, they lose some of their power and immediacy. This has some good consequences. It allows us, for instance, to resist the power of words used for persuasion, seduction, or demagoguery, at least to some extent. But it also encourages the formation of an image of ourselves as *separate* from our words and actions. We see ourselves as observing and acting on the world rather than being part of it.

Finally, by *selecting* the aspects of an utterance which are preserved, notation can lead to impoverished communication. Our own alphabetic notation fails to preserve significant information about timing, intonation, voice, gesture—and selects phonetic information for preservation. This can have a profound influence on some forms of spoken language, legitimizing speech where the phonetic element carries more and more of the burden of the communication. In some contexts it may be implied that educated, lit-

erate, speech is grammatically correct, phonetically well-articulated, decontextualized, unsupported by gesture, and smooth in pace and tone, compared to the speech of the oral or the child. The temptation is to make spoken language more and more like its written counterpart, so that the un-notated currency of speech, in all its richness of expression, becomes debased and devalued. The epitome of this trend is to be found, perhaps, in the radio newscaster of whom (at least in Britain) it can truly be said that he supplies no more than what a listener would have learned from reading the text himself. War atrocities and Stock-Exchange fluctuations are reported in the same measured tones.

Having outlined some general aspects of the effects of literacy, let us now see how these arguments might be applied to music. In oral cultures, music is handed on from person to person without notation and so it is, like verbal knowledge, subject to mutation over time. Only the scribe or the tape-recorder can record that change, however. Within the oral culture there is no way of checking that a particular performance is 'the same music' as before, other than through consensus. There is no evidence that oral cultures support any lengthy musical recall that is exact (i.e. *note-for-note* the same as a model). The music is 'recreated' anew at each performance. What we observed to be true of Yugoslav song characterizes most of the lengthy musical products of oral cultures. Although a basic pattern or kernel may be retained, successive performances demonstrate significant differences of detail and elaboration. This is not to say that the members of the culture view the music as fluid or unfixed. There is real and detailed learning of particular pieces of music; but what is learned is inevitably a relatively high-level abstraction from a succession of performances which *are* different in notatable detail. The kind of exact knowledge which we can gain of particular pieces of music through repeated examination of scores or through repeated hearings of the same record is literally impossible in an oral culture.

There are several things to be borne in mind when evaluating this assertion. First, it is not denying the possibility of *all* verbatim recall. Short melodies *can* be easily memorized, and we have some reason to believe that the orally transmitted nursery rhymes and folk songs of our own and many other cultures have changed remarkably little over the years. These usually have a repeating verse structure, and the musical material is rarely longer than 50 or 60 notes. Second, the assertion does not deny the existence of 'prodigious' feats of memory. In oral settings, people do remember very long pieces of music. My point is that it is unlikely that the memory will be a note-for-note replica of a previous performance. In literate cultures we find many accounts of lengthy verbatim recall in cases where the memorizer is not necessarily making use of notation. This does not contradict the general assertion because it seems fairly clear that such achievements rest indirectly on the existence of notation. We may argue that lengthy verbatim recall depends upon the culture being able to provide several exact repetitions of

the same piece of music. Before the advent of modern recording techniques notation would be the only means of providing performers with the opportunity to produce note-for-note repetitions of lengthy works. When such repetitons are available to a listener, then he may begin to develop the representational skills required to achieve verbatim recall. Eventually an individual may get to the point where he can memorize a long work after one or two hearings (cf. Mozart and Allegri's *Miserere*, discussed in Chapters 1 and 5). There is, however, no psychologically plausible account of lengthy verbatim recall which does not rest on a prior history of 'training up' recall ability on shorter pieces which are repeatedly heard.

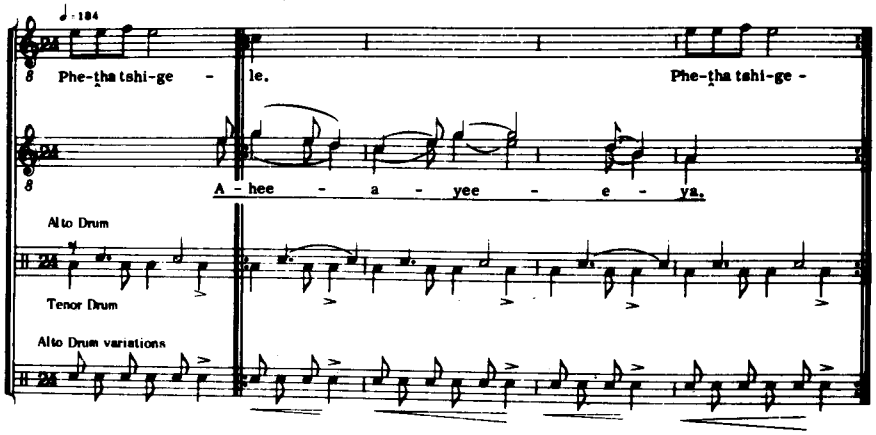
Third, the assertion is not to claim totally different memory mechanisms for people in oral and literate contexts. The fundamental basis for musical memory, as has been argued throughout this book, is the ability to extract higher-order structure from sequences of notes. In an oral context, the musician uses a stored structure to generate different, but structurally linked, note sequences on different occasions. This ability becomes labelled as 'improvisation' in a literate context. When notation (or recording technology) allows for several hearings of the identical note sequence, then a musician can elaborate his structural memory to provide for exact note-for-note recall. The memory is still *structural* as the experimental evidence reviewed in Chapter 5 shows.

A different, but related, effect of the existence of notation is on the *forms* of music that exist within a culture. One may argue that notation enables the construction of complex and lengthy forms where relationships spanning many hundreds of notes are planned in detail. It also enables the use of devices such as inversion or retrogression, and strict counterpoint. This is because a composer can make detailed records of early events and construct later events by repeated reference to them. In a purely oral context, thematic relations are unlikely to have the almost mathematical exactitude of some notated forms, but to have a looser, although still musically effective, association.

In general the forms of oral music do not match those of literate music in 'architectural' complexity, where multiple hierarchical embeddings can produce long, varied, yet integrated compositions. Rather, we find a preponderance of 'chain' structures where the same kind of short element is repeated, with variations, over and over again. The 'theme and variation' structure that we discussed in Chapter 4 with respect to Western improvisatory performance seems to incorporate fundamental characteristics of oral music, of whatever culture.

It is worth noting that even if notation is necessary for the development of such forms as sonata form, fugue, etc., it certainly cannot be a sufficient condition. There are several cultures with develped notational systems which use very different musical forms to those of the Western classical tradition. We must see notation as only one of a range of interacting influences on the development of form.

In addition, to say that oral music lacks the formal complexity of some Western forms is not to say that it lacks all complexity. Oral music provides all sorts of opportunity for elaboration and detailed craftsmanship. However such complexity tends to be due to elaborations *within* sections rather than relations *between* sections. Consider, for instance, the Venda song (South Africa) given by Blacking (1976) in Example 7.1. It repeats the same short




Example 7.1. South African Venda song. From Blacking (1976) Used by permission of Faber and Faber Publishers, London.

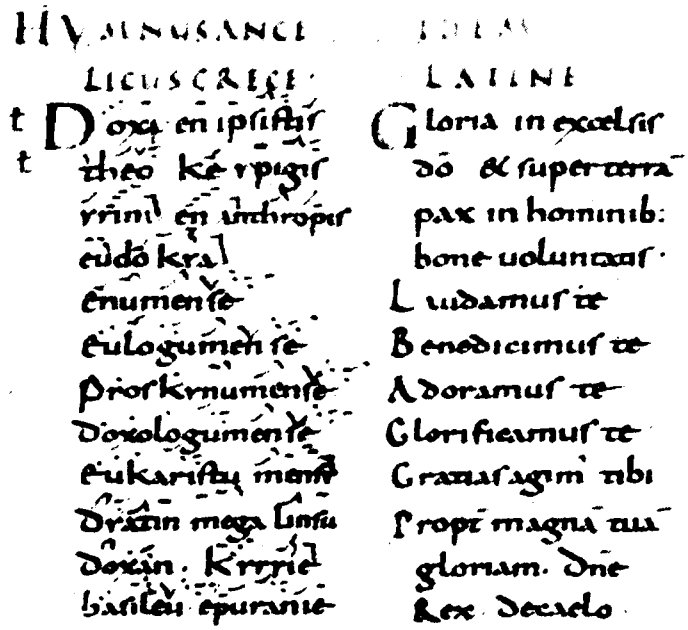
phrase over and over again, yet the rhythmic asymmetries would defeat many a Western performer. The time signature is, in effect, a very fast 5/8 + 7/8 + 5/8 + 7/8, a rhythm which the drummers sustain through a number of possible polyphonic permutations, two of which are given.

There are at least two different, but complementary, ways of understanding a piece of music. One is through its context, the other is through its content. Both ways of looking at music are available in all situations, but it is possible to argue that notation tips the balance in favour of the content approach, and tends to devalue the context. This is precisely because notation (and sound recording) allow us to examine musical content detached from particular performing contexts. Oral music cannot be detached from its context in this way, and a degree of its 'meaning' is supplied by the context. Without such context we are encouraged to view the music as a 'thing in itself', and to search for what Goodman (1976) has described as 'repleteness', where every noticeable aspect of content is taken to be significant, and *designed* to be significant. We must be careful, however, to make sure that the originator of the music intends it to be replete in this sense. We can 'read too much' into a notated composition, but the danger is exacerbated when we approach oral music with this set, for the oral musician is less likely to have turned every element over in his mind, searching for the *exact* notes to capture his *exact* meaning. There is, of course, nothing inherently wrong in listening to

any music in an analytic way. Intellectual disciplines such as musicology would be impossible without it. There is, however, the danger that such analysis will take one further away from, rather than nearer to, the mind of a musician engaged in oral music, for whom such an analysis would seem inappropriate, even meaningless.

A final important aspect of music notation to observe is that, like language, it *selects* some aspects of sound for preservation and discards others. This has important influences on theoretic conceptions of music and, over the long term, on the nature of music itself. To illustrate this we may briefly survey the development of standard Western notation over the past 1000 years or so. The serious scholar should consult Apel (1953), Cole (1974), Karkoschka (1972), and Read (1974) for full details on this topic. I have written elsewhere at more length on psychological aspects of music notation (Sloboda 1981).

Modern Western notation arises, through a complex chain of developments, from *neumes*, which are first found with frequency in religious manuscripts of the eighth century AD. These were used to notate chants to which fixed prayers were sung in religious communities. Example 7.2 shows



Example 7.2. Ninth century neumatic notation. Photograph from Bibliotheque Nationale, Paris.

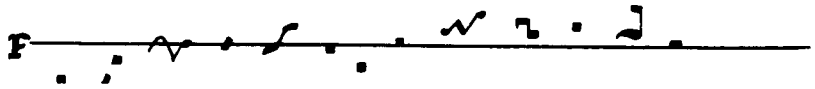
one of the earliest examples of neumes from a French manuscript *c.*871. The neumes are written above the Greek text and indicate the approximate shape of the melodic contour (the rises and falls in pitch). Thus, ⟋ denoted a

rise, ∖ a fall and ∨ a fall and a rise. Wishart (1977) comments on neumes as follows:

> The neume did not attempt to mark out what we now have come to regard as individual pitches and units of rhythm, but only shapes and contours of melodic lines customary in current practice, and hence also requiring a complete familiarity with current melodic practice, and an adherence to it before becoming usable.

Neumes were designed without any intention of recording music for posterity or dissemination to other places. Rather they served as mnemonics for recall to people already familiar with local performing conventions. A person faced with neumatic notation alone could not possibly reconstruct the melody notated, even to a first approximation.

By the twelth and thirteenth centuries, music notation had undergone a series of changes which significantly widened and decontextualized its use. First, symbols became linked to conceptually static entities, 'notes' rather than mobile 'contours'. A note is a sound which remains notionally at the same pitch for the whole of its duration. Second, a means of calibrating the pitch distance between notes was devised by the introduction of stave lines. In the earliest examples, only one line was supplied, identifying a single fixed pitch. In Example 7.3 from the tenth century, the fixed pitch is F below




Example 7.3. Tenth century neumatic notation. From Read (1974).

middle C. It requires some guesswork to decide what the other notes might be. Eventually, however, multiple lines became the rule, as we know them today, with lines and spaces indicating principal melodic pitch levels. Third, aspects of timing began to be notated through the shape of the note symbols. At first, such timing distinctions were fairly crude (e.g. long, short, very short) but gradually the binary subdivision of time units into halves, quarters, eighths, etc., became codified. Whilst notation was linked to chanted texts, timing information was arguably not so important, since the performers were all singing the same text, whose words regulated the timing. The accurate notation of time was, however, crucial to the development of polyphony and instrumental music. An accomplishment such as the performance of a string quartet would be totally impossible were there not accurate means of ensuring synchrony and coordination among the instruments (see Sloboda 1981). Although notation is not the only possible means of promoting synchrony, it has undoubtedly contributed to the rapid development of complex polyphony in Western culture.

For any notational system to work beyond the immediate realm of particular social groups it must become systematic and reliable. This is accomplished mainly by the existence of a small number of finite and discrete

constituents out of which all music can be constructed by combining the constituents. But natural music, particularly that of the voice, is not necessarily made up of discrete elements. Pitch, timing, timbre, loudness can all vary continuously along the various dimensions of sound. Fitting vocal music into the procrustean bed of discrete notation can thus be very unreliable, it would seem. Unreliable, that is, unless the voice actually conforms to the requirements of the notational system by tending towards discrete 'jumps' along the dimensions under its control, particularly the pitch dimension. Within contemporary Western culture it is certainly true that much vocal music aspires to the discreteness that our notational system embodies. Early school training involves emphasis on singing 'in tune' and 'in time', avoiding pitch glides and gross deviations from small whole-number subdivisions in timing. Performed music is thus 'like' notated music in crucial respects. Explications of pitch and time within conventional music theory embody this discreteness through elaborations of concepts of scale, metre, etc. (see Chapter 2). Thus theory, notation, and performing practice form a mutually supporting closed system within which music can be described, notated, transmitted, and understood.

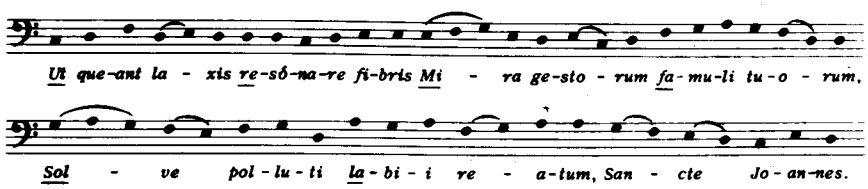
Although notation helped cement theory and practice over the ages, it would, I believe be wrong to assume that pre-notated vocal music had no qualities of discreteness in the pitch and other dimensions. At any stage of history, a notational system used in performance would have to be perceived as providing a relatively good fit to current practice before it would be accepted by performers. We know that discreteness exists in contemporary oral cultures, although it is often blurred through a greater freedom in the use of microtonal and microrhythmic embellishment. Above all, we know that, since prehistory, musical *instruments* have possessed the precise qualities of discreteness that conventional notation codifies, at least in the pitch domain. A plucked string issues a single unchanging pitch. A set of pipes or strings offers a set of relatively fixed pitches, and music played on such instruments must, of necessity, be confined to this set. To notate instrumental music with discrete pitch symbols is not a procrustean operation. It very closely mirrors the sonic state of affairs. In so far as we can assume that the sounds of fixed pitch instruments would influence the vocal practice of oral musicians, we have good reason to suppose that instruments have provided a continuous force towards discreteness in vocal practice.

We may also note that early theories of harmony were based on discrete pitch intervals, and that theoretical notations employing alphabetic conventions were used by Ancient Greek and Chinese writers. However, we should not, perhaps, make too much of this, because until the late Middle-Ages music theory based on the properties of vibrating strings and air columns, as codified in the writings and notations of philosophers, had almost nothing to do with the actual practice of making music. It would not have occurred to philosophers to use their notations and theories to record con-

temporary musical *practice*. We know almost nothing about what Ancient Greek singers and instrumentalists actually did from the writings of these theorists. This is also true of early medieval writers such as Boethius (*c*.480–524) who taught that performers and composers were 'separated from the intellect of musical science'.

What makes the later development of notation in medieval Europe so different is that it was rooted in *practice*, drawing theory, so to speak, in its wake. Furthermore, it arose in a religious context where instruments were frowned upon (although such instruments as the organ became gradually acceptable from the eighth century on: Caldwell 1978). Instrumental music was almost entirely secular—the province of the essentially oral laity—and it did not begin to be notated until long after stave notation was widely accepted.

The acceptance of stave notation seems to rest largely upon the work of Guido of Arrezzo (*c*.1000–50). It is he who is generally credited with the completion and systematization of the four-line stave still used for plainchant notation today (Hoppin 1978). Guido was fully aware of the importance of the need that his own system filled. He was able to 'produce a perfect singer in the space of one year, or at the most in two' whereas previously 10 years of study had yielded 'only an imperfect knowledge'. With his system, singers could now read and perform a melody they had never heard before. As well as perfecting the four-line stave system, Guido was also responsible for devising the 'sol-fa' mnemonic for singers. This is a way of remembering the pitch relations between notes on the stave by referring back to notes in a well-known melody The melody he chose was the hymn to St John the Baptist, '*Ut queant laxis*' (Example 7.4). This hymn 'spells out' a six-note




Example 7.4. Plainsong hymn to St. John the Baptist. From Hoppin (1978).

scale, *ut–re–mi–fa–so–la*. The system, with minor modifications is still in use today, although the archaic '*Ut queant laxis*' has been replaced with more appropriate contemporary songs (as immortalized in Rogers and Hammerstein's *Sound of Music*).

Guido's system rapidly spread through Europe, and soon, all chants of the Roman liturgy had to be notated within his system. That not every existing chant fitted into this system easily is suggested by the fact that some chants underwent significant rewritings and transpositions over the years which can be seen as attempts to 'bring them into line' (Hoppin 1978).

The notational system perfected by Guido paved the way for the emergence of a new breed of musician, the sight-reader (see Chapter 3, section 2). This is the performer who uses notation to give a reasonably polished performance of music without prior rehearsal. He does this through his intense familiarity with the notational system. With advances in printing techniques and resultant increases in legibility, sight-reading reached its apotheosis in the present century.

One should emphasize that the selectivity of music notation does not *force* performance practice to be restricted to what is notated. Even with standard Western notation, performers still have considerable interpretational latitude (see Chapter 3), and one requires much more than the ability to read music to give a culturally acceptable peformance of a Beethoven sonata. What our notation does is to fix a level of detail at which we are prepared to say that anyone providing a performance in accord with the notation has provided a performance of *that* piece, and no other. Thus, the most idiosyncratic and ill-informed performance of a Beethoven sonata is still *that* sonata. This would not be true of someone attempting to provide a performance of Gregorian chant from early neumatic notation. The notation is just not specific enough. Modern notation is normative in a way that early neumatic notation could not be. This does not, however, prevent modern notation being used for purposes where it is *not* normative, and where it significantly underdetermines performance (e.g. the baroque ground bass).

Although notation has supported positive developments, it has its negative aspects too. Wishart (1977) argues that the sterile 'dead end' of twelve-tone music was due to a constriction of the musical imagination by the categories implicit in the notational system:

> In its constant search for new modes of expression, the Western classical music tradition was . . . constrained by its very concentration upon relationships of a limited set of thus notatable 'pitches', to extend the notatable field of harmonic relationships to the limit. The final step into a 12-tone and thence 'integral', serial technique, rather than being a 'liberation' from this restricted set tonality, must be seen in historical perspective as the final total capitulation to the finitistic permutational dictates of a rationalized analytic notation system, and the gateway to much sterile rational formalism . . . More subtle uses of pitch, such as sliding inflections, which can vary over an infinitude of possible speed, interval of sliding and curvature of the slide . . . were not amenable to this finitistic approach.

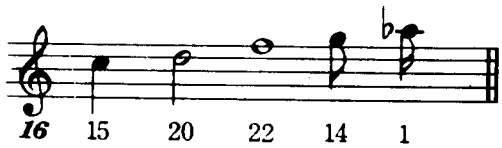
The freedom from sterile and formal music has been made possible, not so much through new notational systems, as through the opportunity now afforded to composers to bypass notation entirely and compose directly into sound with the assistance of electronic machines. This is partly because it is difficult to see how *any* notational system could cope with things like sliding inflections reliably; and partly because the fixed 12-tone system has become enshrined in conventional instrument technology. One cannot do pitch glides on a piano.

I have suggested that the selectivity of music notation reinforces the tendency to view musical elements as falling into discrete categories, a tendency which is, in any case, encouraged by the existence of discrete pitch instruments. We may still be entitled to ask, however, *why* it is that man has favoured fixed pitch instruments for his music. It has recently been argued (e.g. Balzano 1980; Pressing 1983) that discreteness in both the pitch and time dimensions are essential if music is to serve certain communicative functions. We shall see, in the following discussion, that much music in both oral and literate traditions shares basic characteristics, in view of which we must conclude that even such potent cultural forces as notation are not all-important in determining the nature of music.

### 7.2.2. *Musical universals*

We have examined one factor which can contribute towards cultural variation in music. There are, of course, many other such factors. For instance, not all cultures support the existence of musician 'specialists' such as composers, instrumentalists, etc. This can have a profound influence on the type of musical activity that takes place. Is it possible that, despite all such factors, there are some underlying features which typify most music? If there are, then these might be related to some universal cognitive basis for music which transcended individual cultures. We have already suggested the possibility of primitive auditory grouping strategies playing a part in the way that music is heard (Chapter 5.2). Another possible set of universals arises in connection with the use of the pitch dimension in different cultures.

A very large number of cultures contain, both in theory and practice, the notion that music takes place with respect to fixed *reference* pitches. These pitches need not be fixed for all time, but are usually fixed for the duration of a single piece of music. In many cultures, the principal reference pitch (or pitches) are maintained throughout the music in the form of a (usually instrumental) 'drone'. Even where drones are absent, we can usually see that certain pitches are 'privileged', in that the music often returns to them, and circles round them. Ethnomusicologists sometimes represent this feature by deriving a *weighted scale* for a given piece of music. This orders the principal pitches in ascending sequence, and associates with each pitch a notated duration which roughly reflects the number of occurrences of that note in the piece. Example 7.5 for instance, gives the weighted scale associated with a particular Aboriginal song (Malm 1977). It shows that there are five principal




Example 7.5. Weighted scale notation. From Malm (1977).

notes, with F occurring most frequently, and A flat occurring least frequently.

Thus, although tonality, as we know it, is by no means universal, the notions of scale and tonic have formal analogies in most cultures. One of the influences on the development of fixed discrete points along the continuum is, as we have argued earlier, the existence of instruments which are either of permanently fixed pitch (e.g. the *metallophones* such as gongs, bells, and xylophones) or else which can be tuned, but which offer a fixed set for the duration of a single performance (e.g. the stringed instruments such as harp, sitar, or guitar). Moreover, within most of these scales the *octave* appears to be a particularly privileged interval. The disposition of pitches repeats within each octave, and the octave frequently appears as an interval in polyphonic music. It is particularly common to find the principal reference pitch being reinforced by voice or instruments at several different octaves. In addition, intervals close to our perfect fifth and perfect fourth appear in the polyphony of most cultures.

The similarities do not end here, however, because it seems that the subdivisions of the octave into scale steps follow common principles in most cultures. First, the number of subdivisions is always moderately small. Miller's (1956) 'magic number seven plus or minus two' would encompass most of the world's scales. Particularly common are five- and seven-note scales. In fact, the term *octave* derives from the fact that, in a seven-note scale, the eighth note marks the start of the repetition of the scale pattern at a pitch with a frequency of approximately double that of the first note. A second similarity is that practically no scale divides the octave up into equal ratio steps. In other words, scales are almost never found where the pitch intervals between adjacent notes are the same for each and every pair. Rather, scales are composed by selecting from a number of smaller (often notionally equal) steps, but in such a way that there are different numbers of steps between some of the adjacent scale members. Thus, our own diatonic major scale selects seven notes from a nominal equal-interval 12-note chromatic scale (Fig. 7.1). The Indian *Sa-grama* scale is derived from a theoretical 22-interval *sruti* scale (Fig. 7.2).

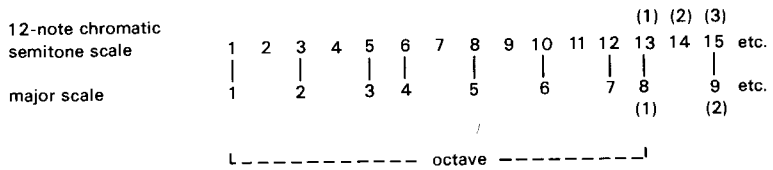



Fig. 7.1. Octave subdivision of the diatonic major scale.

The ubiquity of this unequal interval principle leads one to ask whether it serves some fundamental psychological purpose. It has been suggested earlier (Chapters 2 and 5) that one consequence of this unequal distribution is to allow a listener to 'get tonal bearings' on hearing a subset of the notes of a

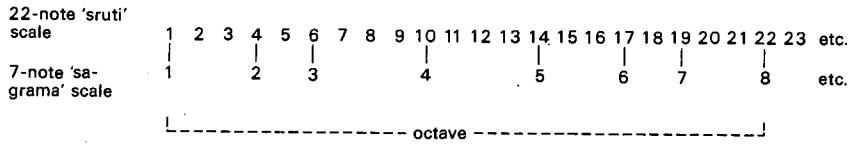



Fig. 7.2. Octave subdivision of the Indian sa-grama scale.

scale, without any necessary explicit marking of the tonic by repetition or priveleged performance treatment. For instance, in the case of a major scale, the interval separating notes four and seven (six semitones) is unique. No other pair of notes is separated by this interval (the *tritone*); and so a listener hearing a melodic sequence such as 4–7–8 in a major scale can identify the last note played as the tonic. In the Indian *Sa-grama* scale, the interval between notes three and five is similarly unique (eight *sruti*), at least in theory.

Shepard (1982) argues that it is this property of uneven spacing which

> enables the listener to have, at every moment, a clear sense of where the music is with respect to such a framework. Only with respect to such a framework can there be things such as motion or rest, tension and resolution, or, in short, the underlying dynamisms of tonal music. By contrast, the complete symmetry and regularity of the chromatic and whole-tone scales means that every tone has the same status as every other. The fact that for such scales there can be no clear sense of location, and hence of motion is, I believe, the reason that such scales have never enjoyed wide or sustained popularity as a basis for music.

It has been further argued (e.g. Balzano 1980) that the burden of the musical significance of a scale rests, not in its precise tuning, or in the degree to which perfect whole number frequency ratios are represented between its members (e.g. 3:2, 4:3, etc.), but in the configuration that the scale takes up by virtue of selecting its members from a larger set of notionally equal steps. Although 'perfect' intervals have been *historically* important in the development of scale systems, scales achieve much of their psychological significance through configurational properties which Balzano has captured in terms of group-theoretic notions borrowed from mathematics. As he points out,

> When we look at the Just, Pythagorean, and Equal Temperament tunings of our familiar major scales, it is evident that there is an important sense in which they all work in substantially the same way: we don't have to learn to compose or hear separately for each one. The ratios, which are different in all three tuning schemes, do not really address this fundamental commonality.

The evidence we have already reviewed on categorical perception (Chapters 2 and 5) shows that precise tuning of intervals is without psychological consequence in many situations. What Balzano has done is to formalize the level of description at which we can see that all the various tunings of the major scale are doing the same (group theoretic) job. It is only possible for a very limited number of other scales to perform the same functions.

One function of a scale, as we have argued, is to give a listener a sense of *location* by virtue of the different intervallic relations that hold between the various members. If this was the only function of a scale, then the 'best' kind of scale would be of the type illustrated in Fig. 7.3 where intervals between

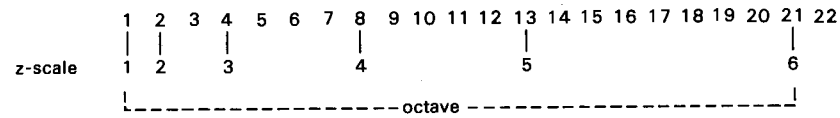



Fig. 7.3. Octave subdivision of the 'Z' scale.

adjacent notes are all different, and most non-adjacent intervals are different too. We can arbitrarily name this scale, based on a 20-fold division of the octave, the Z-scale. There are other properties of scales like the major scale, however, which the Z-scale does not share. Perhaps the most important of these is the special relationship that holds between the twelve major scales generated by starting on each note of the chromatic scale in turn. Each scale shares all but one of its notes with the scale that starts on its fifth degree. The fourth note of the old scale is sharpened to form the seventh note of the new scale; and by recursively performing this operation round the 'circle of fifths' one can generate all the major scales. This allows concepts of scale- or key-relatedness to enter music. G major is *close* to C major because it shares all but one note. F sharp major is *remote* from C major because it shares *only* one note. The Z-scale does not have this property. There is nowhere that one can start a Z-scale which causes an overlap of all notes but one with another Z-scale. One can see, intuitively, that the large number of different intervals makes that impossible. The major scale contains only two types of interval between adjacent notes, tone and semitone; and it seems that this property is crucial.

Balzano (1980) has demonstrated that there is a restricted set of artificial scales which share crucial group-theoretic properties with the diatonic major scale. These scales require the octave to be subdivided into 12, 20, 30, or 42 equal intervals, or any larger number given by the formula $n(n-1)$. Each of these subdivisions yields one, and only one, scale type with the desired properties. For the 12-fold division, that scale is the diatonic major scale. For the 20-fold division, that scale is the nine-note scale given in Figure 7.4. Its closest relation is generated by starting on the fifth degree and sharpening

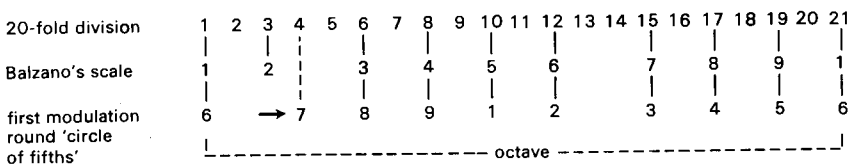



Fig. 7.4. Octave subdivision of Balzano's (1980) nine-note scale.

just one note (the second note of the old scale). Applying this operation recursively yields all 20 possible scales of this mode. Note that there are only two different intervals between adjacent notes, 2/20 or 3/20 of an octave. It is a general feature of Balzano's scale set that the two intervals which are found between adjacent notes are separated in size by just one subdivision of the octave. So, the eleven-note scale derived from a 30-fold subdivision of the octave has adjacent intervals of either 2/30 or 3/30 of an octave. The 13-note scale derived from a 42-fold subdivision of the octave has adjacent intervals of either 3/42 or 4/42 of an octave. As the size of the subdivisions becomes smaller, the absolute and relative difference between these intervals becomes smaller. It is much easier to discriminate between a tone and a semitone than between the analogous intervals in the other scales of the set. It can be argued, therefore, that the 12-fold division is psychologically optimal; and although Balzano advocates the experimental use of the other microtonal scales in computer music, it is perhaps not surprising that there is little evidence of these other scales having been developed in the world's cultures.

The considerations on which Balzano bases his formulations are essentially algebraic. Traditional formulations of scale structure have been essentially acoustic, pointing to the natural occurrence of some scale intervals in the harmonic series. The 12-note chromatic scale can be seen as the single system which satisfies both algebraic and acoustic constraints, and to which music therefore, in some sense, tends. Three observations support the centrality of the diatonic scale:

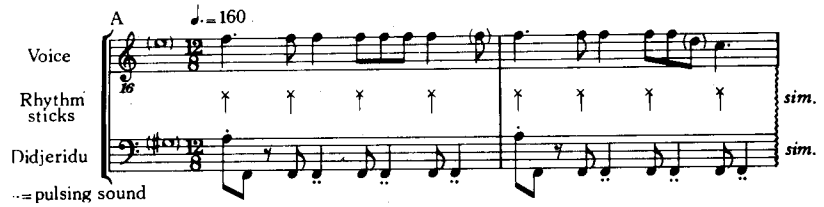
1. Many pentatonic scales are based on a subset of the full diatonic scale, and so can be converted to a diatonic scale by the addition of two further notes.

2. Far from being a recent Western invention, the diatonic scale can be traced back to the most ancient tuning systems so far deciphered from archaeological records (Kilmer, Crocker, and Brown 1976; cited in Shepard 1982).

3. Even in cultures where the octave is *theoretically* divided into other than 12 parts, practice often diverges from theory in the direction of a 12-fold system. Thus, for instance, although we have seen that Indian music theory postulates a 22-fold division of the octave in its *sruti* scale, contemporary music practice seems to be based on a 12-fold subdivision of the octave though each has a fairly wide range of tolerance as to the actual pitch (Jairazbhoy 1971). It is important to remember that many excellent Indian performers have no overt knowledge of, or interest in, music theory (Malm 1977).

We now move from the pitch domain to a consideration of possible musical universals in the temporal domain. Just as reference pitches seem universally important, so do reference *times*, important moments in relation to which other sounds are organized; and, just as many cultures maintain the reference

pitch in awareness through the use of a drone, so many cultures maintain reference times in awareness through the use of instruments marking out a regular pulse or metre. The simplest pulse involves a series of equally spaced sounds, such as is provided by a clock ticking, but in nearly every culture, such a basic pulse becomes elaborated through intermediate detail. Example 7.6 shows the first two 'bars' of an Australian Aboriginal song given by




Example 7.6. Opening of Aboriginal song. From Malm (1977).

Malm (1977). Here we see that the rhythm sticks provide a basic unadorned pulse. The didjeridu is, however, doing something more complex. It synchronizes with each beat of the rhythm sticks, but adds a second 'pulsing' sound after each beat, which has the effect of subdividing the beat, marking its passage by an asymmetric disposition of sounds (the long 'pulsing' sound marks the end of the beat, the shorter sound the beginning). Through such asymmetry, a sense of *location* within the beat is created. Such ascribing of location is given the general name of *accenting*, and it is crucial to the music of nearly all cultures. Differentiation through accenting can be achieved in a variety of ways. For instance, the time subdivision can be symetrical, with asymmetry in terms of pitch or dynamic pattern. This can also be seen in the didjeridu part of Example 7.6. Every fourth pulse of the rhythm sticks is marked by a pitch change in the didjeridu, thus creating a superordinate pulse, which justifies the positioning of the bar-lines in the transcription.

Thus, a hierarchy of reference times can be created, since a superordinate pulse can itself form the basis of a yet higher grouping. In Javanese Gamelan music, for instance, the fundamental time unit is the *gongan*. Each *gongan* is marked by the sounding of the largest gong in the ensemble. The *gongan* is divided into two equal halves by the sounding of the *kenong*. Each *kenong* is divided into two equal halves by the sounding of the *kempul*. Each *kempul* is divided into two equal halves by the sounding of the *kethuk*. Each *kethuk* may then be further subdivided (a *kethuk* corresponds very roughly to a bar in Western music, so a *gongan* is an eight-bar unit).

When a pulse unit at some level is subdivided asymmetrically in time, we usually talk of the subdivisions as forming a rhythm, or rhythmic pattern. Just as scales are usually made up by selecting pitch intervals asymmetrically from a larger set of small equal intervals, so rhythmic patterns are made up by selecting time intervals asymmetrically from a larger set of small equal time intervals; and in the same way that scales tend to repeat at the octave,

so rhythmic patterns tend to repeat in such a way as to mark off time into equal segments.

One of the simplest rhythms of many musical cultures is the *dactyl* (a long interval followed by two short ones). In Western music, this is notated as a minim followed by two crotchets, and signifies the selection of the first, third, and fourth of four equal underlying beats for marking. When this figure is repeated over and over in an *ostinato*, most listeners hear the long note as strongly marking (*accenting*) the beginning of the next four-beat grouping. On the whole, in Western music, the underlying and repeating sequence of equal time units (metre) from which rhythmic patterns are derived tends to contain small even numbers of units. The dactyl requires a four-unit metre, but six and eight are also very common. Furthermore, rhythmic asymmetries are often constructed on a principle of binary subdivision of the metrical unit. Thus, we can conceive of the dactyl as being constructed by subdividing the four-unit metre into two equal halves containing two beats each. The first half is then 'filled' by a single note, whilst the second half is further subdivided into two equal 'quarter notes'.

In some other cultures (e.g. Southern Africa and India) we find more complex rhythms. First, the number of beats contained within the metrical unit can be large and uneven. In Indian music, *tala* of between seven and sixteen beats are common. Second, subdivisions are frequently not made on a principle of binary subdivision. When the number of beats is odd, binary subdivision is impossible; but even when there are an even number of beats we often find non-binary subdivisions. Consider, for instance, the rhythmic pattern of the drum parts in Example 7.1 (South African Venda song). The pattern repeats every twelve beats, but the effect of the two long notes (crotchets on the lowest stave) is to create an unequal subdivision of five plus seven.

The ability of rhythm to give a listener a fine-grained sense of temporal location within the metrical unit through asymmetric subdivision of time units suggests that, in some sense, complex rhythmic structures perform the same kind of cognitive functions that are supported by diatonic tonality in our culture (Pressing 1983). Regardless of how far this analogy can be pursued, it is certainly true that scale and rhythm perform the same essential function, that of dividing up the pitch and time continua into discrete and re-identifiable locations, on which backdrop all the essential dialectical activities (tension–resolution, motion–rest) can flourish.

Finally, these fixed points provide cues for synchrony, so that musicians can organize their behaviour with respect to what others are doing. A universal feature of polyphonic music is the temporal and pitch co-ordination of the parts. Without reference points, it would be impossible for people to make the necessary anticipatory and planned adjustments to bring their behaviour into co-ordination with others, and thus make musical behaviour the structured *social* phenomenon that it is the world over.

## 7.3. Biology and musical thinking

### 7.3.1. *The neural organization of musical function*

On 9 October 1959 the distinguished Russian composer V. G. Shebalin (1902–63) suffered a severe stroke. When he recovered consciousness he was suffering from a partial paralysis which affected the right side of his body. He also suffered severe disturbance in speech function, which remained with him to the end of his life. He found it hard to understand, produce, or reproduce speech. About a year after his stroke, for instance, when asked to repeat three short sentences: 'The moon is shining. The house is burning. The dog is barking', he replied: 'The moon is shining . . . and the house . . . no . . . I really don't understand.' In trying to describe his compositional history he offered:

> An oratorio is descended . . . it was a long time ago . . . that is a kind of spiritual . . . oh . . . gradually it became . . . it was extended . . . and then . . . and now . . . gradually . . . ah . . . it became very different . . . how may I say . . . I have to recall . . . oh . . . nothing . . . I have thoughts . . . no . . . it is too difficult . . .

Reading and writing were also disturbed, although not so badly.

Despite these severe difficulties (as Luria, Tsvetkova, and Futer 1965, report) he continued to work as a composer.

> He worked hard with his pupils, listening to their compositions, analyzing and correcting them. He spent a considerable time over his own creative work. While aphasic, he finished compositions which he had started to write before he was taken ill, and he created a series of new compositions which other musicians considered to be up to standard, and which did not significantly differ from the compositions of his earlier years.

Of his Fifth Symphony, composed in 1962, Dimitri Shostakovitch said, 'it is a brilliant creative work, filled with highest emotions, optimistic and full of life. This symphony composed during his illness is a creation of a great master.'

How is it possible that a man's verbal ability could disintegrate, leaving his musical intellect unchanged? A large body of evidence has now been amassed to support the view that intellectual functions are localized in different areas of the brain, so that damage to one area may only disrupt a proportion of normal intellectual functions. Post-mortem examination of Professor Shebalin revealed massive damage in the left temporal and parietal regions of the brain, but not elsewhere. Such damage is often associated with language disorder, and sensorimotor defects on the right side of the body.

In order to discuss this issue in more detail, it is necessary to outline a few basic neuroanatomical facts. The human brain (and body) is divided into two roughly symmetrical halves. Each half of the brain receives nervous input from, and issues nervous commands to, one side of the body. However, for reasons which are not well understood, there is a 'crossover' of nerves

going in and out of the brain, such that nerves belonging to the right side of the body connect with the left brain, and nerves belonging to the left side of the body connect with the right brain. We say that nervous organization is essentially *contralateral* (literally 'opposite sided').

The part of the brain which seems most deeply implicated in learned cognitive skills is the cerebral cortex which, in man, is the largest part of the brain. The cortex surrounds the rest of the brain like a thick cap, but it is divided down the midline into two cerebral *hemispheres*. Although the two hemispheres are to a certain extent capable of independent operation, they are connected by a fat bundle of nerve fibres called the *corpus callosum*. When this is intact, information can be transferred from one side of the cortex to the other. In most animals, each cerebral hemisphere seems to do roughly the same job. In some higher mammals, however, especially man, there is now incontrovertible evidence that each hemisphere is, or becomes, specialized to some degree. For most people (the 90 per cent or so who are unambiguously right-handed) the mechanisms controlling language behaviour seem to be primarily concentrated in the *left* hemisphere, and those controlling spatial orientation and other non-verbal skills seem to be concentrated in the *right* hemisphere.

There are several converging sources of evidence which support this conclusion:

1. Studies of people with brain damage show that left-hemisphere damage is often associated with language disorder (as in the case of Professor Shebalin), whereas right-hemisphere damage rarely causes significant language disruption.
2. By temporarily anaesthetizing one or other hemisphere, one can selectively disrupt different intellectual functions. For instance, Gordon and Bogen (1974) found that left-hemisphere intervention caused speech disturbance whilst right-hemisphere disturbances caused disruption of singing.
3. One can observe asymmetries of various sorts in the gross electrical activity of the brain which alter according to the task being carried out. For instance, McKee, Humphrey, and McAdam (1973) took electroencephalograph (EEG) recordings from points on the scalp above the temporo-parietal regions of both hemispheres. They took recordings whilst subjects carried out various musical and linguistic tasks. The left hemisphere was more active during the linguistic tasks than during the musical ones.
4. When linguistic or non-linguistic input is restricted to one ear, then linguistic material is better processed if it is presented to the right ear (left hemisphere), and non-linguistic material such as music is better processed if presented to the left ear (right hemisphere). A well-used technique is that of 'dichotic presentation' in which two different stimulus sequences are simultaneously presented to the two ears (e.g. Kimura 1964, 1967). When the two messages are speech sequences then the message reaching the right ear

tends to be better perceived. When the two messages are musical sequences then the pattern tends to be reversed, with the left ear doing better.

The literature on hemispheric specialization is very large, and is growing all the time. Recent general introductions to the area are provided by Springer and Deutsch (1981), and Kinsbourne and Smith (1974). Reviews of the literature with particular respect to music have been provided by Benton (1977), Damasio and Damasio (1977), Gates and Bradshaw (1977), Marin (1982), Shuter-Dyson and Gabriel (1981), and Wyke (1977). Rather than attempt to review this large, and not altogether homogenous, body of research, I will confine myself to a critical examination of a few of the most noteworthy findings in the area, those that seem to have particular relevance for the arguments that we have been pursuing in the rest of the book.

The first point to make is that music is not a single, monolithic, capacity which either exists in an individual or does not. Musical skill has many, logically independent, sub-skills which may, therefore, be anatomically independent as well. A region of a cerebral hemisphere is a very big place in neural terms, and even if it were possible to assert that music was 'in' a part of the right hemisphere, it would still be pertinent to ask whether different components of the skill were in different anatomical locations *within* that region. Psychologists have been particularly alerted to the possibility that very specific sub-skills can be considered as functionally and anatomically distinct by finely analytic work carried out over the last decade or so on people with language disorders caused by brain damage (see Coltheart, Patterson, and Marshall 1980). It has been known for many years that a person may suffer receptive disorders of language without production being affected, and *vice versa*. The recent research goes beyond this in demonstrating that disruption may be confined to a specific *part* of a single skill like reading. Thus, for instance, a person may be able to read familiar concrete nouns, but have the greatest difficulty with connectives and other function words (e.g. 'and', 'which', 'that', etc.). This is important psychologically, because it shows that the mechanisms by which content and function words are understood must be, to some extent, functionally distinct.

Using such an analytic approach it has been discovered that language function is not entirely confined to the left hemisphere, even in fully right-handed people. For instance, it seems that the right hemisphere can understand words and simple sentences provided that such understanding can be signalled by non-verbal actions, such as obeying a command, pointing to, or picking up a named object.

One of the major problems in studying the effect of brain damage on musical skill is that the level of skill prior to damage can be very varied, and is often unknown; whereas most people have a high, and verifiable, level of linguistic skill prior to damage. The most informative musical cases are those where the patient was a professional musician before receiving brain damage. Here we can be fairly sure what skills were present prior to damage. Thus, a

case like that reported by Wertheim and Botez (1961) is of particular interest. The patient they studied was a professional violinist known to have perfect pitch before brain damage. He developed right sensorimotor disturbances consistent with left hemisphere damage, and experienced difficulties in receptive aspects of language functioning. Some aspects of music function remained undisturbed. These included:

1. Reproduction by singing or violin playing of single pitched notes.
2. Pitch discrimination down to intervals smaller than a semitone.
3. Noticing differences between pairs of melodies or chords.
4. Detection of tonal violations in melodies.
5. Naming of single written notes.
6. Identification of major and minor scales.

Other aspects were disturbed. These included:

1. Loss of perfect pitch (i.e. the ability to name pitches).
2. Inability to recognize familiar pieces of music, or to identify composer and/or style.
3. Inability to identify melodic intervals.
4. Inability to reproduce rhythmic patterns.
5. Inability to read note sequences.
6. Errors in singing familiar melodies.
7. Errors in copying and transcribing music.

The pattern of disorders does not admit of easy explanation, but Marin (1982) suggests that the pattern is compatible with preservation of simple perceptual abilities (pitch discrimination, pitch reproduction) along with some highly overlearned and automated knowledge structures (the major and minor scales, basic tonal progressions) together with the loss of recognition and identification of musical sequences, particularly where some lexical operation (such as naming) is required.

This patient shows a very different profile of disorders to Professor Shebalin, who seems to have suffered no musical loss at all. It turns out that Shebalin is rather unusual in the clinical literature. There exist only about twelve reported cases of language disorder in musicians which is unaccompanied by musical disorder. More commonly, both language and music are affected, especially when damage occurs to the left hemisphere. There is also a small group of documented cases where musical disorder (primarily of a receptive nature) is unaccompanied by noticeable language disruption. In these cases, brain damage is usually in the right hemisphere.

We see, therefore, that although right hemisphere damage nearly always disrupts musical functioning in musicians, left hemisphere damage often disrupts it too. It is, therefore, oversimplistic to say that music is 'in' the right hemisphere. What the complex, and somewhat unsatisfactory literature on

brain damage does at least show is that musical skills are just as dissociable and separately disruptable as are language skills.

For further insight into the neural organization of musical skill we turn from studies on brain-damaged patients to a series of studies on normal intact subjects. The line of enquiry I wish to focus on was initiated by Bever and Chiarello (1974). In their experiment they used two groups of subjects. 'Musicians' were those who had been attending music lessons for at least five years, and were currently playing or singing. 'Non-musicians' were those who had received less than three years of music lessons and had not been musically active in the last five years. The subjects were all right-handed. They listened to a tape of specially prepared melodies, and were asked to say whether each melody had occurred before in the sequence. Half the subjects had the melodies presented to the right ear (left hemisphere) and the other half had them presented to the left ear (right hemisphere). Bever and Chiarello found that non-musicians performed best when the melodies went to the left ear. This confirmed the 'traditional' view that music is processed in the right hemisphere. The musicians, however, showed exactly the reverse pattern of results; they performed best when the melodies went to the *right* ear.

Bever and Chiarello attempted to account for these results by recourse to the widely held view (first proposed by Hughlings-Jackson, see Taylor 1932) that the left hemisphere is specialized for 'analytic' and 'sequential' processing, whilst the right hemisphere is specialized for 'holistic' or 'global' processing. On this view, the 'crossover' of laterality in music recognition occurs because the musicians adopt a analytic strategy whereas the non-musicians adopt a holistic strategy. The terms 'analytic' and 'holistic' are rather broad in their connotation, but their application to music has been taken to relate to such distinctions as that between processing the contour of a melody and processing the precise pitch intervals of which the melody is composed (see Chapter 5.3), the former being a holistic process, the latter being analytic.

The work of Bever and Chiarello has stimulated a spate of experimental studies looking at the way in which musical ability and experience, processing strategy, and stimulus material, interact to give different patterns of hemispheric differences. It seems that the left hemisphere advantage is not confined to musicians. People who score highly on a test of musical aptitude tend to show left hemisphere advantage, regardless of training (Gaede, Parsons, and Bertera 1978), as do non-musicians who consciously adopt an 'analytic' strategy in a music perception task (Peretz and Morais 1980). It would be fair to say that the literature does not yet provide a fully coherent picture of hemispheric asymmetry in musical function. The differences between left and right ear scores, although statistically significant, are often very small, and they are highly sensitive to the nature of the task, the stimulus material, and subject strategy. In addition, experimenters have by no means tapped the whole range of musical activities, concentrating on the recognition of small

changes in chords or brief melodic sequences. For all these reasons, although the analytic–holistic distinction seems a current 'best bet' as an explanation of left–right differences in normal subjects, the picture is by no means complete.

Marin (1982) argues cogently that the study of the neurological bases for music will not advance substantially until the categories and distinctions between musical activities made on psychological and music-theoretic grounds are taken seriously by researchers. Too often tests are devised on the basis of methodological convenience rather than with precise and well-grounded musical considerations in mind. Traditional distinctions between, for example, receptive and productive skills, are just not fine-grained enough to tell us, for instance, whether the phenomenon under study operates at the phonological, syntactic, or semantic level (Chapter 2).

What we can see clearly, even in this rather incomplete picture, is that various sub-skills of music have a certain degree of neural independence. There is little evidence for a single 'music centre' in the brain. Furthermore, the brain areas responsible for music seem to have a partial, but incomplete, overlap with those responsible for language. We should, therefore, not be surprised if the behavioural and cognitive analogy between music and language (Chapter 2) is, whilst compelling, not complete. Music, if not a totally distinctive neural function, almost certainly employs a distinctive configuration of neural resources.

### 7.3.2. *The origins and functions of music*

The principles of evolution and natural selection propounded by Darwin continue to play a central role in theories of biological development (Dawkins 1976). Darwin proposed that man's music evolved from the vocalizations of sub-human primates like the apes (Darwin 1874, Chapter XIX) which served as 'emotional' signals. Correct though Darwin's general theory may be, his explanation of musical evolution is far from satisfactory. For one thing, the vocalizations of apes seem to be involuntary patterns, having more in common with human behaviours like laughing, crying, and screaming, which survive very distinctively in humans as something quite different from music. For another, ape vocalizations do not share any of the organizational qualities of human music. They are not based on scales or rhythmic patterns, as are all human forms, and they do not show any of the the planning or purposiveness that characterize other areas of ape behaviour such as locomotion and manipulation of objects. If apes have an intellect, there is little evidence that their vocalizations are at its service. It is for this reason that attempts to teach chimpanzees a vocal language have been such a dismal failure. Only when chimps were taught a manual sign-language was any modest progress made (e.g. Gardner and Gardner 1969).

Darwin's notion was that the most musical (i.e. pitched) sounds of apes were produced by the male during courtship, and that correspondingly man's earliest music was in the form of love songs. Modern ethnological studies force

us to reject both parts of this notion. The call cries of most primates are signal calls expressing 'all is well', 'I am here', 'danger', and variants of these (Marler 1965; Petter and Charles-Dominique 1979; Williams 1980). Williams reports that 'monkeys and apes in fact have no specific sex-calls, and there is no reason for assuming that the ancestors of primitive man serenaded their prospective mates.' Secondly, the most primitive songs about which we know are not concerned with personal love or sexual relations. The poetic songs of the Aborigines, for instance, express the feelings of men and women as members of the community rather than as individuals. Even when fertility is celebrated, it is seen in its wider role of sustaining the natural and social order.

A major function of primate calls seems to be 'group cohesive' (Petter and Charles-Dominique 1979). Calls are used so that members of a social group can locate one another, and indicate their motivational state. It is possible to argue that primitive human music has analogous group cohesive functions. For instance, Blacking (1976) reports that a Venda person can often know exactly what a distant group is doing by the song they are singing. Songs and dance rituals serve to unite and define social groupings. This analogy between apes and men is, however, by no means exact, and fails by itself to account for the temporally organized nature of human music. It seems as if some specifically human tendency to create and notice organized patterns, hierarchies, and sequences, has overrun and overhauled almost every type of behaviour 'inherited' from non-human primates.

Comparing man to his nearest biological neighbours, it has seemed to many that, despite observable similarities, there is an unbridged 'quantum leap' between the highest achievements of apes and what every normal human being can achieve as a matter of routine. Of course, in standard theory, evolution does proceed by discrete leaps, each leap being occasioned by a genetic mutation in an individual. However, simply to state this theory does not provide a satisfactory explanation of any given case. We have no idea if it is possible that a single mutation (or a manageably small number of them) could turn an ape-like creature into something resembling a man. Furthermore, in the case of music, it is not at once clear how musical behaviour makes for better adapted individuals that are more likely to survive; and this, of course, is a precondition for any mutation to be selected for transmission to future generations. This puts music in a special case apart from language and other 'practical' skills whose survival values are numerous and obvious.

Does man *need* music? If he does, then deprivation should be harmful in some way. We know, for instance, that sleep is necessary to humans, since continued deprivation of sleep causes physical and psychological ill effects. On this count music is very unlike sleep. People can go without music for very long periods without suffering noticeable ill effects. Perhaps, however, this is to take a too naïve approach. There are several activities (such as sex)

which are vital for the continuance of the species, but which can be refrained from by individuals without any apparent personal harm. Music may be like this, necessary to the species, but not to any particular individual.

If this is an appropriate line of argument, then the 'unit' for a deprivation experiment might need to be a whole culture. Cultures without music do not, in fact exist, but it seems hard to imagine why lack of music should be harmful to a civilization like our own. We may argue, however, that modern cultures have outgrown the *need* for music, and it is in primitive non-literate cultures that music is of major survival value. Society requires organization for its survival. In our own society we have many complex artefacts which help us to externalize and objectify the organizations we need and value. Primitive cultures have few artefacts, and the organization of the society must be expressed to a greater extent through transient actions and the way people interact with each other. Music, perhaps, provides a unique mnemonic framework within which humans can express, by the temporal organization of sound and gesture, the structure of their knowledge and of social relations. Songs and rhythmically organized poems and sayings form the major repository of human knowledge in non-literate cultures. This seems to be because such organized sequences are much easier to remember than the type of prose which literate societies use in books. It would, perhaps, be an exaggeration to say that there was no logical alternative to music as a mnemonic device. It does seem to me, however, that there are very few things which 'natural' peoples could congenially do with their speech that did not tend towards what we understand as music. They can move their bodies and modulate the timing of their speech. When this is done in an organized way, *rhythm* is created. They may modulate the pitch of the voice. When this is done in a organized way, *melody* and some form of *tonal structure* is created.

We began this book with the observation that music is emotionally appealing to almost all people, and ventured the view that its appeal came *through* its structure, so that there was an inescapable cognitive component to the act of musical appreciation. We have explored the nature of these cognitive structures and processes in some detail through the seven chapters of the book. I have tried to present as complete as possible a picture of what is known about human musical cognition, without too much speculation, but with an eye to the considerable gaps in our knowledge. Perhaps I may be allowed to end on a note of pure speculation with respect to the fundamental question which, in many ways, must motivate all scientific enquiry into music. This question is, of course, *why* does music pervade human culture?

Even though humans have an intellect which, in some respects is very powerful, they do not find it easy to evaluate the long-term effects of their actions, or to predict which course of actions will lead to the most desired result in a complex situation. This is why the development of scientific thinking has been one of the most important steps in human development. It allows a method for arriving at such evalutions by the application of logical

thinking and controlled observations of the world. However, long before the development of scientific method, people were notably successful at mastering their environment. They achieved success precisely through the kinds of mental processes which we can now formally defend as being effective ways of modelling and conceptualizing events in the world. They used these processes, not primarily because they had emerged through a rational process of calculation, or through trial and error, but because it was *appealing* to behave in that way. Evolution had supplied a natural propensity to behave in adaptive ways. This included a propensity to use language and music. Music is indeed of immense benefit as a mnemonic aid, but evolution did not implant a rational awareness of this fact in the human mind. Rather, it supplied a *motivation* for music, so that it became enjoyable and 'natural' for people to indulge in it.

Society has now changed, so that we have many more powerful mnemonics than music available to us. Our instincts and motivations, however, change more slowly. We still enjoy music, and thank goodness we do. Detached from its mundane origins, it can now serve as a vehicle for a whole range of aesthetic and transcendent experiences. Two words of caution are necessary, however. Because our instincts for music are rooted in the conditions that prevailed in the infancy of humanity, the *forms* that were available to early people have a primal and inescapable influence. In particular, it is the *voice* and the human body in rhythmic movement which form the motivational mainspring of music. If music departs too far from this mainspring, it will cease to have deep meaning and power for us. Contemporary musicians have, I believe, already come to realize that the unfettered development of electronic music leads to sterility and lifelessness. Electronic instruments must always be constrained by the parameters of 'human' music making, enhancing and enriching these parameters rather than striking off in arbitrary directions.

The second, final, point is more sombre. For all its achievements, our Western society has a degree of precariousness about it. It is all too easy to construct scenarios which would entail the destruction of the delicate balances which preserve our complex social arrangements. In such a situation, those of us who survived would find ourselves in a world where the artefacts of our present society would, by and large, have disappeared. The resources that we carried around in our heads would, once again, form the mainstay of our attempts to survive. Songs and poems would become vital mnemonic and cohesive tools for the construction of a new society, and musical skill would, indeed be a skill for survival. It is not, therefore, simply a task of disinterested scientific curiosity to come to a better understanding of musical skill. Music is a fundamental human resource which has played, and may well play again, a vital role in the survival and development of humanity.

# References


Allport, D. A., Antonis, B., and Reynolds, P. (1972).On the division of attention: a disproof of the single channel hypothesis. *Quart. Jl. Exp. Psychol.* **24,** 225–35.

Alvin, J. (1975). *Music therapy.* Hutchinson, London.

Anderson, E.,(ed) (1966). *The letters of Mozart and his family.* Macmillan London.

Anderson, J. R. (1980). *Cognitive psychology and its implications.*Freeman, San Francisco.

Anderson, J. R. (1981). *Cognitive skills and their acquisition.* Erlbaum, Hillsdale, New Jersey.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychol. Rev.* **89,** 369–406.

Apel, W. (1953). *The notation of polyphonic music 900–1600* (4th edn.). The Mediaeval Academy of America, Cambridge, Massachusets.

Attneave, F. and Olson, R. K. (1971). Pitch as a medium: a new approach to psychophysical scaling. *Am. J. Psychol.* **84,** 147–66.

Balzano, G. J. (1980). The group-theoretic description of twelvefold and microtonal pitch systems. *Comput. Mus. J.* **4,** 66–84.

Barbour, J. M. (1951). *Tuning and temperament.* Michegan State College Press, East Lansing.

Bartlett, J. C. and Dowling, W. J. (1980). The recognition of transposed melodies: a key-distance effect in developmental perspective. *J. Exp. Psychol. Hum. Percept Perform.* **6,** 501–15.

Bartók, B. and Lord, A. B. (1951). *Serbo-Croatian folk songs.* Columbia University Press, New York.

Basmajian, J. V. and Newton, W. J. (1974). Feedback training of parts of buccinator muscle in man. *Psychophysiol.* **11,** 92.

Bent, I. (1980). Analysis. In *The new Grove's dictionary of music and musicians* (ed. S. Sadie). Macmillan, London.

Benton, A. L. (1977). The amusias. In *Music and the brain* (eds. M. Critchley and R. A. Henson). Heinemann, London.

Bernstein, L. (1976). *The unanswered question.* Harvard University Press, Cambridge, Massachusetts.

Bever, T. G. and Chiarello, R. J. (1974). Cerebral dominance in musicians and non-musicians. *Science* **185,** 537–9.

Blacking, J. (1976). *How musical is man?* Faber, London.

Blechner, M. J., Day, R. S., and Cutting, J. E. (1976). The processing of two dimensions in nonspeech stimuli: the auditory–phonetic distinction reconsidered. *J. Exp. Psychol. Hum. Percept. Perform.* **2,** 257–66.

Boden, M. A. (1979). *Piaget.* Fontana, London.

Bond, Z. S. (1971). Units in speech perception. *Working papers in linguistics.* No. 9. Computer and Information Science Research Center, Ohio State University.

Bower, G. H. (1976). Experiments in story under standing and recall. *Quart. Jl. Exp. Psychol.* **28,** 511–34.

Bower, G. H. and Winzenz, D. (1969). Group structure, coding, and memory for digit series. *J. Exp. Psychol. Monog.* 80, 1–17.

Bower, T. G. R. (1971). The object in the world of the infant. *Scient. Am.* **225,** 30–8.
Brady, P. T. (1970). Fixed-scale mechanism of absolute pitch. *J. Acoust. Soc. Am.* **48,** 883–7.
Bransford, J. D. (1979). *Human cognition.*Wadsworth, Belmont.
Bregman, A. S. (1981). Asking the 'what for' question in auditory perception. In *Perceptual organization* (eds. M. Kubovy and J. R. Pomerantz). Erlbaum, Hillsdale, New Jersey.
Bregman, A. S. and Campbell, J. (1971). Primary auditory stream segregation and perceptual order in rapid sequences of tones. *J. Exp.Psychol.* **89,** 244–9.
Broadbent, D. E. (1958). *Perception and communication.* Pergamon, London.
Brown, R. and Butler, D. (1981). Diatonic trichords as minimal tonal cue-cells. *In Theory Only* **5,** 39.
Buck, P. C. (1944). *Psychology for musicians.* Oxford University Press, London.
Bulbrook, M. E. (1932). An experimental inquiry into the existence and nature of 'insight'. *Am. J. Psychol.* **44,** 409–53.
Bullivant, R. (1971). *Fugue.* Hutchinson, London.
Butler, D. (1979). A further study of melodic channelling. *Perc. and Psychophys* **25,** 264–8.
Butler, D. (1983). The initial identification of tonal centres in music. In *Acquisition of symbolic skills* (eds. D. Rogers and J. A. Sloboda). Plenum, New York.
Caldwell, J. (1978). *Medieval music.* Hutchinson, London.
Carlsen, J. C. (1981). Some factors which influence melodic expectancy. *Psychomusicology.* **1,**12–29.
Chang, H. and Trehub, S. E. (1977a). Auditory processing of relational information by young infants. *J. Exp. Child Psychol.* **24,** 324–31.
Chang, H. and Trehub, S. E. (1977b). Infant's perception of temporal grouping in auditory patterns. *Child Dev.* **48,** 1666–70.
Chase, W. G. and Chi, M. T. H. (1981). Cognitive skill: implications for spatial skill in large-scale environments. In *Cognition, social behaviour and the environment* (ed. J. Harvey). Erlbaum, Potomac, Maryland.
Chase, W. G. and Simon, H. A. (1973). The mind's eye in chess. In *Visual information processing* (ed. W. G. Chase). Academic Press, New York.
Cherry, E. C., (1953). Some experiments on the recognition of speech, with one and two ears. *J. Acoust. Soc. Am.* **25,**975–9.
Chomsky, N. (1957). *Syntactic structures.* Mouton, The Hague.
Chomsky, N. (1965). *Aspects of the theory of syntax.* MIT Press, Cambridge, Massachusetts.
Chomsky, N. (1968). *Language and mind.* Harcourt Brace Jovanovitch, New York.
Chomsky, N. and Halle, M. (1968). *The sound pattern of English.* Harper and Row, New York.
Clark, H. H. and Clark, E. V. (1977). *Psychology and language.* Harcourt Brace Jovanovitch, New York.
Clarkson, J. K. and Deutsch, J. A. (1978). Pitch control in the human voice: a reply to Rostron. *Quart. J. Exp. Psychol.* **30,** 167–9.
Claxton, G. (Ed). (1980). *Cognitive psychology: new directions.* Routledge and Kegan Paul, London.
Clynes, M., (ed) (1982). *Music, mind and brain: the neuropsychology of music.* Plenum, New York.

Cohen, A. and Cohen, N. (1973). Tune evolution as an indicator of traditional musical norms. *J. Am. Folklore.* **86** 37–47.

Cole, H. (1974). *Sounds and signs: aspects of musical notation.* Oxford University Press, London.

Cole, R. A. and Scott, B. (1974). Towards a theory of speech perception. *Psychol Rev.*. **15,** 348–74.

Coltheart, M., Patterson, K., and Marshall, J. C.,(eds) (1980). *Deep dyslexia.* Routledge and Kegan Paul, London.

Cooke, D. (1959). *The language of music.* Oxford University Press, London.

Critchley, M. and Henson, R. A.,(eds) (1977). *Music and the brain.* Heinemann, London.

Cuddy, L. L. (1970). Training the absolute identification of pitch.*Percept. Psychophys.* **8,** 265–9.

Cuddy, L. L. (1977). Perception of structured melodic sequences. Paper given at conference on musical perception, Paris, France, July 10–13.

Cuddy, L. L. and Cohen, A. J. (1976). Recognition of transposed melodic sequences. *Q. Jl. Exp. Psychol.* **28,** 255–70.

Cuddy, L. L., Cohen, A. J., and Miller, J. (1979). Melody recognition: the experimental application of musical rules. *Can. J. Psychol.* **33,** 148–57.

Cutting, J. E. and Rosner, B. S. (1974). Categories and boundaries in
speech and music. *Percept. Psychophys.* **16,** 564–70.

Cutting, J. E., Rosner, B. S., and Foard, C.F. (1976). Perceptual categories for musiclike sounds: implications for theories of speech perception. *Quart. Jl. Exp. Psychol.* **28,** 361–78.

D'Azevedo, W. L. (1962). Uses of the past in Gola discourse. *J. Afr. Hist.***3,** 11–34.

Damasio, A. R., and Damasio, H. (1977). Musical faculty and cerebral dominance. In *Music and the brain* (eds M. Critchley and R. A. Henson). Heinemann, London.

Darwin, C. (1874). *The descent of man* (2nd edn.). John Murray, London.

Davies, J. B. (1978). *The psychology of music.* Hutchinson, London.

Davies, J. B. (1979). Memory for melodies and tonal sequences: a theoretical note. *Brit. J. Psychol.* **70,** 205–210.

Dawkins, R. (1976). *The selfish gene.* Oxford University Press, Oxford.

Delis, D., Fleer, J., and Kerr, P. (1978). Memory for music. *Percept. and Psychophys.* **23,** 215–8.

Deutsch, D. (1970). Tones and numbers: specificity of interference in short-term memory. *Science* **168,** 1604–5.

Deutsch, D. (1972). Mapping of interactions in the pitch memory store. *Science* **175,** 1020–2.

Deutsch, D. (1973) Interference in memory between tones adjacent in the musical scale. *J. Exp. Psychol.* **100,** 228–31.

Deutsch, D. (1975). Two-channel listening to musical scales. *J. Acoust. Soc. Am.* **57,** 1156–60.

Deutsch, D. (1980) The processing of structured and unstructured tonal sequences. *Percept. and Psychophys.* **28,** 381–9.

Deutsch, D. (ed). (1982a). *The psychology of music.* Academic Press, New York.

Deutsch, D. (1982b). Grouping mechanisms in music. In *The psychology of music* (ed. D. Deutsch). Academic Press, New York.

Deutsch, D. (1982c). The processing of pitch combinations. In *The psychology of music* (ed. D. Deutsch). Academic Press, New York.

Deutsch, D. and Feroe, J. (1975). Disinhibition in pitch memory. *Percept Psychophys.* **17,** 320–4.

Deutsch, D. and Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychol. Rev.* **86,** 503–22.

Deutsch, J. A. and Clarkson, J. K. (1959). Nature of the vibrato and the control loop in singing. *Nature (Lond.).* **183,** 167–8.

Deutsch, J. A., and Deutsch, D. (1963). Attention; some theoretical considerations. *Psychol. Rev.* **70,** 80–90.

Donaldson, M. (1978). *Children's minds.* Fontana, London.

Dowling, W. J. (1973). The perception of interleaved melodies. *Cog. Psychol.* **5,** 322–37.

Dowling, W. J. (1978). Scale and contour: two components of a theory of memory for melodies. *Psychol. Rev.* **85,** 341–54.

Dowling, W. J. (1982). Melodic information processing and its development. In *The psychology of music* (ed. D. Deutsch). Academic Press, New York.

Dowling, W. J., and Bartlett, J. C. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology* **1,**30–49.

Durkin, H. E. (1937). Trial-and-error, gradual analysis and sudden reorganization: an experimental study of problem solving. *Arch. Psychol.* **210.**

Edworthy, J. (1983). Towards a pitch-contour continuum theory of memory for melodies. In *Acquisition of symbolic skills* (eds D. Rogers and J. A. Sloboda). Plenum, New York.

Eimas, P. D. (1975). Speech perception in early infancy.In *Infant perception: from sensation to cognition* (ed. L. B. Cohen and P. Salapatek) Vol.2. Academic Press, New York.

Eimas, P. D. and Corbit, J. D. (1973). Selective adaptation of linguistic feature detectors. *Cog. Psychol.* **4,** 99–109.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. and Vigorito, J.M. (1971). Speech perception in infants. *Science* **171,** 303–6.

Ericsson, K. A. and Simon, H. A. (1980). Verbal reports as data. *Psychol. Rev.* **87,** 215–51.

Farnsworth, P. R. (1968) *The social psychology of music* (2nd edn). Iowa State University Press, Iowa.

Fitts, P. M. (1964). Perceptual-motor skill learning. In *Categories of human learning* (ed. A. W. Melton). Academic Press, New York.

Fitzgibbons, J., Pollatsek, A., and Thomas, I. B., (1974). Detection of temporal gaps within and between perceptual tonal groups. *Percept. and Psychophys.* **16,** 522–8.

Flavell, J. (1963). *The developmental psychology of Jean Piaget.* Van Nostrand, Princeton, New Jersey.

Fodor, J. A. and Bever, T. G. (1965). The psychological reality of linguistic segments. *J. Verb. Learn. Verb. Behav.* **4,** 414–21.

Fodor, J. A., Bever, T. G., and Garrett, M. F. (1974). *The psychology of language.* McGraw Hill, New York.

Forte, A. (1961). *The compositional matrix.* Baldwin, New York.

Forte, A. (1962). *Tonal harmony in concept and practice.* Holt Rinehart and Winston, New York.

Forte, A. (1978). *The harmonic organization of the Rite of Spring*. Yale Univerity Press, New Haven.

Fourcin, A. J. and Abberton, E. (1971). First applications of a new laryngograph. *Med. Biol. Illust.* **21,** 172.

Frega, A. L. (1979). Rhythmic tasks with 3-, 4-, and 5-year old children. A study made in the Argentine Republic. *Bull. Council Res. Mus. Educ.* **59,** 32–4.

Funk, J. and Whiteside, J. (1981). Developmental theory and the psychology of music. *Psychol. Mus.* **9.2,** 44–53.

Gabriel, C. (1978). An experimental study of Deryck Cooke's theory of music and meaning. *Psychol. Mus.* **6,** 13–20.

Gabrielsson, A. (1974). Performance of rhythm patterns. *Scand. J. Psychol.* **15,** 63–72.

Gaede, S. E., Parsons, O. A., and Bertera, J. H. (1978). Hemispheric differences in music perception: aptitude *versus* experience. *Neuropsychologia* **16,** 369–73.

Gardner, H. (1973). Children's sensitivity to musical styles. *Merrill-Palmer Q. Behav. Dev*. **19,** 67–77.

Gardner, H. (1980). Composing symphonies and dinner parties. *Psychol. Today*. **14.** 4, 18–27.

Gardner, H. (1981). Do babies sing a universal song? *Psychol. Today*. **15.** 12, 70–6.

Gardner, H. and Wolf, D. (1983). The waves and streams of symbolization. In *Acquisition of symbolic skills* (eds D. Rogers and J. A. Sloboda). Plenum, New York.

Gardner, H., Davidson, L., and McKernon, P. (1981). The acquisition of song: a developmental approach. In *Documentary report of the Ann Arbor Symposium*. Music Educators' National Conference, Reston, Virginia.

Gardner, R. A. and Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science* **165,** 664–72.

Gates, A. and Bradshaw, J. L. (1977). The role of the cerebral hemispheres in music. *Brain Language*. **4,** 403–31.

Ghiselin, B. (1952). *The creative process*. Mentor, New York.

Gibson, E. J. (1969). *Principles of perceptual learning and development*. Appleton-Century-Crofts, New York.

Goodman, N. (1976). *Languages of art*. Hackett, Indianapolis.

Goody, J. and Watt, I. (1963). The consequences of literacy. *Comp. Stud. Soc. Hist.* **5,** 27–68.

Gordon, H. W. and Bogen, J. E. (1974). Hemispheric lateralization of singing after intracarotid sodium amylobarbitone. *J. Neurol. Neurosurg. and Psychiat.* **37,** 727–38.

Green, D. W. and Mitchell, D. C. (1978). The effects of context and content on immediate processing in reading. *Q. Jl. Exp. Psychol.* **30,** 608–36.

Greene, J. (1972). *Psycholinguistics*. Penguin, Harmondsworth.

Gregory, A. H. (1978). Perception of clicks in music. *Percept. Psychophys.***24,** 171–4.

Gruneberg, M. M., Morris, P. E., and Sykes, R., (eds) (1978). *Practical aspects of memory*. Academic Press, London.

Gruson, L. M. (1981). Investigating competence: a study of piano practising. Paper presented at the 89th Annual convention of the American Psychological Association, Los Angeles.

Haber, R. N. and Schindler, R. M. (1981). Error in proofreading: evidence of

syntactic control of letter processing? *J. Exp. Psychol. Hum. Percept. and Perform.* **7,** 573–579.

Halpern, A. R. and Bower, G. H. (1982). Musical expertise and melodic structure in memory for musical notation. *Am. J. Psychol.* **95,** 31–50.

Hamburger, M. (1952). *Beethoven: letters, journals and conversations.* Pantheon Books, New York.

Handel, S. and Todd, P. (1981). Segmentation of sequential patterns. *J. Exp. Psychol. Hum. Percept. and Perform.* **7,** 41–55.

Healy, A. F. (1980). Proofreading errors on the word *the*; new evidence on reading units. *J. Exp. Psychol. Hum. Percept. Perform.* **6,** 45–57.

Henry, L. K. (1934). The role of insight in the analytic thinking of adolescents. *Stud. Educ. Univ. Iowa Stud.* **9,** 65–102.

Hevner, K. (1936). Experimental studies of the elements of expression in music. *Am. J. Psychol.* **48,** 246–68.

Hindemith, P. (1952). *A composer's world: horizons and limitations.* Harvard University Press, Cambridge, Massachusetts.

Hirsh, I. J. (1959). Auditory perception of temporal order. *J. Acoust. Soc. Am.* **31,** 759–67.

Hochberg, J. (1974). Organization and the Gestalt tradition. In *Handbook of perception* (eds E. C. Carterette and M. P. Friedman) Vol. 1. Academic Press, New York.

Hodeir, A. (1956). *Jazz: its evolution and essence.* Secker and Warburg, London.

Hoppin, R. H. (1978). *Medieval music.* W. W. Norton, New York.

Hughes, E. (1915). Musical memory in piano playing and piano study. *Mus. Q.* **1,** 592–603.

Hunter, I. M. L. (1976). Memory: theory and application. In *Piaget, psychology and education* (eds V. P. Varma and P. Williams). Hodder and Stoghton, London.

Hutt, S. J., Hutt, C., Lenard, H. G., Bernuth, H. V., and Muntjewerff, W. J. (1968). Auditory responsivity in the human neonate. *Nature* **218,** 888–90.

Ibbotson, N. R. and Morton, J. (1981). Rhythm and dominance. *Cognition.* **9,** 125–38.

Idson, W. L. and Massaro, D. W. (1976). Cross-octave masking of single tones and musical sequences: the effects of structure of auditory recognition. *Percept. Psychophys.* **19,** 155–75.

Imberty, M. (1969). *L'acquisition des structures tonales chez l'enfant.* Klinksieck, Paris.

Jackendoff, R., and Lerdahl, F. (1982). A grammatical parallel between music and language. In *Music, mind, and brain: the neuropsychology ofmusic* (ed. M. Clynes). Plenum Press, New York.

Jairazbhoy, N. (1971). *The ragas of North Indian music.* Wesleyan University Press, Middletown.

Jeans, J. (1937). *Science and music.* Cambridge University Press, London.

Johansson, G. (1976). Visual motion perception. In *Recent progress in perception: readings from Scientific American* (eds R. Held and W. Richards). Freeman, San Francisco.

Johnson, D. P. (1980). *Beethoven's early sketches in the 'Fischof Miscellany' Berlin Autograph 28.* Vols 1 and 2. UMI Research Press, Ann Arbor, Michigan.

Jones, M R.,Kidd, G., and Wetzel, R. (1981). Evidence for rhythmic attention. *J. Exp. Psychol.: Hum. Percept. Perf.* **7,** 1059–73.

Joseph, L. G. (1978). An observational study on the acquisition of musical skills. Unpublished M. A. thesis, University of Waterloo, Canada.

Jusczyk, P. W., Rosner, B. S., Cutting, J. E. Foard, C. F., and Smith L. B. (1977). Categorical perception of non-speech sound by 2-month-old infants. *Percept. Psychophys.* **21**, 50–4.

Kagan, J. and Lewis, M. (1965). Studies of attention in the human infant. *Merrill-Palmer Q. Behav. Dev.* **11**, 95–127.

Karkoschka, E. (1972). *Notation in new music.* Universal Edition, London.

Kessen, W., Levine, J., and Wendrick, K. A. (1979). The imitation of pitch in infants. *Infant Behav. Dev.* **2**, 93–9.

Kilmer, A. D., Crocker, R. L., and Brown, R. R. (1976). *Sounds from silence: recent discoveries in ancient Near East music.* Bit Enki Publications, Berkeley.

Kimura, D. (1964). Left-right differences in the perception of melodies. *Q. Jl. Exp. Psychol.* **16**, 355–8.

Kimura, D. (1967). Functional asymmetry of the brain in dichotic listening. *Cortex.* **3**, 167–78.

Kinsbourne, M. and Smith, W. L., (eds) (1974). *Hemispheric disconnection and cerebral function.* Charles C. Thomas, Springfield, Illinois.

Klinger, E. (1971). *Structure and functions of fantasy.* Wiley, New York.

Kosslyn S. M. and Pomerantz, J. R. (1977). Imagery, propositions, and the form of internal representations. *Cog. Psychol.* **9**, 52–76.

Kramer, R. (1973). *The sketches for Beethoven's violin sonatas opus 30; history, transcription, analysis.*Princeton University, Unpublished Ph.D. dissertation.

Kubovy, M. (1981). Concurrent-pitch segregation and the theory of indispensable attributes. In *Perceptual organization* (eds M. Kubovy and J. R. Pomerantz). Erlbaum, Hillsdale, New Jersey.

Kuhn, T. S. (1962). *The structure of scientific revolutions.* University of Chicago Press, Chicago.

Kunst, J. (1959). *Ethnomusicology.* Nijhoff, The Hague.

Ladefoged, P. and Broadbent, D. E. (1960). Perception of sequence in auditory events. *Q. Jl. Exp. Psychol.* **12**, 162–70.

Lashley, K. (1951). The problem of serial order in behaviour. In *Cerebral mechanisms in behaviour* (ed. L. A. Jeffries). Wiley, New York.

Laske, O. E. (1975). On psychomusicology. *Int. Rev. Aesth. Sociol. Mus.* **6**, 269–81.

Lehiste, I. (1972). The units of speech perception. *Working papers in linguistics* No. 12. Computer and information science research center, Ohio State University.

Lerdahl, F. and Jackendoff, R. (1983). *A generative theory of tonal music.* MIT Press, Cambridge, Massachusetts.

Levin, H. and Addis, A. B. (1980). *The eye-voice span.* MIT Press, Cambridge, Massachusetts.

Levin, H. and Kaplan, E. A. (1970). Grammatical structure and reading. In *Basic studies on reading* (eds H. Levin and J. P. Williams). Basic Books, New York.

Levin, H. and Williams, J. P. (eds) (1970). *Basic studies in reading.* Basic Books, New York.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. (1967). Peception of the speech code. *Psychol. Rev.* **74**, 431–61.

Liberman, A. M., Harris, K. S., Kinney, J. A., and Lane, H. (1961). The

discrimination of the relative onset time of the components of certain speech and non-speech patterns. *J. Exp. Psychol.* **61,** 379–88.

Locke, S. and Kellar, L. (1973). Categorical perception in a non-linguistic mode. *Cortex.* **9,** 355–69.

Longuet-Higgins, H. C. (1972). Making sense of music. *Proc. R. Inst. G. Br.* **45,** 87–105.

Longuet-Higgins, H. C. (1976). Perception of melodies. *Nature* **263,** 646–53.

Longuet-Higgins, H. C. (1978). The perception of music. *Interdisc. Sci. Rev.* **3,** 148–56.

Longuet-Higgins, H. C. and Lee, C. S. (1982). The perception of musical rhythms. *Perception* **11,** 115–28.

Lord, A. B. (1960). *The singer of tales.* Harvard University Press, Cambridge, Massachusetts.

Luria, A. R., Tsvetkova, L. S., and Futer, D. S. (1965). Aphasia in a composer. *J. Neurol. Sci.* **2,** 288–92.

Lyons, J. (1970). *Chomsky.* Fontana, London.

McKee, G., Humphrey, B., and McAdam, D. W. (1973). Scaled lateralization of alpha activity during linguistic and musical tasks. *Psychophysiology.* **10** 441–3.

Makeig, S. (1982). Affective versus analytic perception of musical intervals. In *Music, mind and brain: the neuropsychology of music* (ed. M. Clynes). Plenum Press, New York.

Malm, W. P. (1977). *Music cultures of the Pacific, the Near East and Asia* (2nd edn). Prentice Hall, Englewood Cliffs, New Jersey.

Manor, H. C. (1950). A study of prognosis. *J. Educ. Psychol.* **41,** 31–50.

Marin, O. S. M. (1982). Neurological aspects of music perception and performance. In *The psychology of music.* Academic Press, New York.

Marler, P. (1965). Communication in monkeys and apes. In *Primate behaviour* (ed. I. DeVore). Holt Rinehart and Winston, New York.

Marr, D. (1982). *Vision.* Freeman, San Francisco.

Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behaviour. *Psychol. Rev.* **79,** 487–509.

Massaro, D. W. and Klitzke, D. (1977). Letters are functional in word identification. *Memory and Cognition* **5,** 292–8.

Mattingley, I. G., Liberman, A. M., Syrdal, A. K., and Halwes, T. (1971). Discrimination in speech and non-speech modes. *Cog. Psychol.* **2,** 131–57.

Meyer, L. B. (1956). *Emotion and meaning in music.* University of Chicago Press, Chicago.

Meyer, L. B. (1973). *Explaining music.* University of California Press, Berkeley.

Michon, J. A. (1974). Programs and 'programs' for sequential patterns in motor behaviour. *Brain Res.* **71,** 413–24.

Miller, G. A. (1956). The magical number seven plus or minues two: some limitations on our capacity for prcessing information. *Psychol. Rev.* **63,** 81–97.

Miller, G. A. and Heise, G. A. (1950). The trill threshold. *J. Acoust. Soc. Am.* **22,** 167–73.

Minsky, M. (1977). Frame-system theory. In *Thinking: readings in cognitive science* (eds. P. N Johnson-Laird and P. C. Wason). Cambridge University Press, London.

Moog, H. (1976). *The musical experience of the pre-school child* (translated C. Clarke). Schott, London.

Moore, B. C. J. (1982). *An introduction to the psychology of hearing* (2nd edn). Academic Press, London.

Morgenstern, S. (1956). *Composers on music: an anthology of composers' writings from Palestrina to Copland.* Pantheon Books, New York.

Moynahan, E. D. (1973). The development of knowledge concerning the effect of categorization upon free recall. *Child Dev.* **44,** 238–46.

Mozart, W. A. (1979). *Die Zauberflote.* Documenta Musicologica. Zweite Reihe: Handschriften-Faksimiles. Barenreiter, Kassel.

Narmour, E. (1977). *Beyond Schenkerism: the need for alternatives in music analysis.* University of Chicago Press, Chicago.

Navon, D. (1977). Forest before trees: the precedence of global features in word identification. *Cog. Psychol.* **9,** 353–83.

Neisser, U. (1967). *Cognitive psychology.* Appleton Century Crofts, New York.

Neisser, U. (1976). *Cognition and reality: principles and implications of cognitive psychology.* Freeman, San Francisco.

Neisser, U. (1981). John Dean's memory: a case study. *Cognition.* **9,** 1–22.

Neisser, U. (1982). *Memory observed: remembering in natural contexts.* Freeman, San Francisco.

Neisser, U. (1983). Toward a skilful psychology. In *The acquistion of symbolic skills* (ed. D. R. Rogers and J. A. Sloboda). Plenum, New York.

Nettl, B. (1964). *Theory and method of ethnomusicology.* Free Press of Glencoe, New York.

Nettl, B. (1973). *Folk and traditional music of the Western continents* (2nd edn). Prentice Hall, Englewood Cliffs, New Jersey.

Newell, A. and Simon, H. A. (1972). *Human problem solving.* Prentice Hall, Englewood Cliffs, New Jersey.

Newell, A., Shaw, J. C., and Simon, H. A. (1962). The process of creative thinking. In *Contemporary approaches to creative thinking* (eds H. Gruber, G. Terrell, and M. Wertheimer). Prentice Hall, New York.

Newmarch, R. (1906). *Life and letters of Peter Ilich Tchaikovsky.* John Lane, London.

Nisbet, R. E. and Wilson, J. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychol. Rev.* **84,** 231–59.

Nottebohm, G. (1887). *Zweite Beethoveniana* (ed. E. Mandyczewski). J. Reiter-Biedermann, Leipzig.

O'Regan, K. (1979). Moment to moment control of eye saccades as a function of textual parameters in reading. In *Processing of Visible Language* (ed. P. A. Kolers, M. E. Wrolstad, and H. Bouma) Vol. 1. Plenum Press, New York.

Olson, G. M. and Clark, H. H. (1976). Research methods in psycholinguistics. In *Handbook of perception* (eds E. C. Carterette and M. P. Friedman) Vol. 7, *Language and speech.* Academic Press, New York.

Ostwald, P. F. (1973). Musical behaviour in early childhood. *Dev Med. Child Neurol.* **15,** 367–75.

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cog. Psychol.* **9,** 441–474.

Parrott, I. (1978). *The music of Rosemary Brown.* Regency Press, London.

Parry, M. (1971). *The making of Homeric verse.* Oxford University Press, London.

Patterson, B. (1974). Musical dynamics. *Sci. Am.* **233,** 78–95.

Peretz, I. and Morais, J. (1980). Modes of processing melodies and ear asymmetry in non-musicians. *Neuropsychologia.* **18,** 477–89.

Petter, J. J. and Charles-Dominique, P. (1979). Vocal communication in prosimians. In *The study of prosimian behaviour* (eds G. A. Doyle and R. D. Martin). Academic Press, New York.

Pflederer, M. (1964). The responses of children to musical tasks embodying Piaget's principles of conservation. *J. Res. Mus. Educ.* **12,** 251–68.

Piaget, J. (1950). *The psychology of intelligence.* Harcourt Brace Jovanovitch, New York.

Piaget, J. (1952). *The origins of intelligence in the child.* Routledge and Kegan Paul, London.

Pick, A. D. (1979). Listening to melodies: perceiving events. In *Perception and its development* (ed. A. D. Pick). Erlbaum, Hillsdale, New Jersey.

Pillsbury, W. B. (1897). A study in apperception. *Am. J. Psychol.* **8,** 315–93.

Plomp, R. (1977). *Aspects of tone sensation.* Academic Press, New York.

Poincaré, H. (1924). Mathematical creation. From *The foundations of science* (translated G. B. Halstead). Science Press, New York.

Pollard-Gott, L. (1983). Emergence of thematic concepts in repeated listening to music. *Cog. Psychol.* **15,** 66–94.

Povel, D. J. (1981). Internal representation of simple temporal patterns. *J. Exp. Psychol Hum. Percept. Perform.* **7,** 3–18.

Pratt, F. (1983). Intellectual realism in adults' and children's drawings of cubes and straight lines. In *Acquisition of symbolic skills* (ed. D. R. Rogers and J. A. Sloboda). Plenum, New York.

Pressing, J. (1983). Cognitive isomorphisms in pitch and rhythm in world musics: West Africa, the Balkans, Thailand and Western Tonality. *Studies in Mus.* **17.**

Priestley, M. (1975) *Music therapy in action.* Constable, London.

Pylyshyn, Z. (1973). What the mind's eye tells the mind's brain: a critique of mental imagery. *Psychol. Bull.* **80,** 1–24.

Rainbow, E. L. and Owen, D. (1979). A progress report on a three year investigation of the rhythmic ability of pre-school aged children. *Bull. Council Res. Mus. Educ.* **59,** 32–4.

Rasch, R. A. (1979). Synchronization in performed ensemble music.*Acustica.* **43,** 121–31.

Rayner, K. (1978). Foveal and parafoveal cues in reading. In *Attention and Performance VII* (ed. J. Requin). Erlbaum, Hillsdale, New Jersey.

Read, G. (1974). *Music notation.* Gollancz, London.

Reason, J. (1977). Skill and error in everyday life. In *Adult learning* (ed. M. Howe). Wiley, London.

Reber, A. S. and Anderson, J. R. (1970). The perception of clicks in linguistic and nonlinguistic messages. *Percept. Psychophys.* **8,** 81–9.

Reitman, W. R. (1965). *Cognition and thought.* Wiley, New York.

Reitmeyer, J. W. (1972). The application of negative practice to the correction of habitual fingering errors in clarinet performance. Unpublished thesis, The Pennsylvania State University (Abstract in *Diss. Abstr. Int.* 1973, **33,** 3403).

Restle, F. (1970). Theory of serial pattern learning: structural trees. *Pschol.Rev* **79,** 487–509.

Restle, F. (1972). Serial patterns: the role of phrasing. *J. Exp. Psychol.* **92,** 385–90.

Robinson, I. (1975). *The new grammarians' funeral.* Cambridge University Press, London.

Rosner, B. S. and Meyer, L. B. (1982). Melodic processes and the perception of music. In *The psychology of music* (ed. D. Deutsch). Academic Press, New York.

Sadie, S. (ed.) (1980). *The new Grove's dictionary of music and musicians.* MacMillan, London.

Salzer, F. (1952). *Structural hearing.* Charles Boni, New York.

Sampson, E. E. (1981). Cognitive psychology as ideology. *Am. Psychologist.* **36,** 730–43.

Sanchez, M. and Reitman, W. R. (1960). The composition of a fugue: protocol and comments. Unpublished report. Carnegie Institute of Technology, Pittsburgh.

Sandford, A. J. and Garrod, S. C. (1980). *Understanding written language: explorations in comprehension beyond the sentence.* Wiley, London.

Scharf, B. (1970). Critical bands. In *Foundations of modern auditory theory* (ed. J. V. Tobias) Vol. 1. Academic Press, New York.

Schenker, H. (1935). *Der Freie Satz* (translated E. Oster, 1979). Longman, New York.

Seashore, C. E. (1938). *The psychology of music.* McGraw Hill, New York.

Seashore, C. E., Lewis, D., and Saetvit, J. G. (1960). *Seashore measures of musical talents* (2nd revision of 1939 revision). The Psychological Corporation of New York, New York.

Sergeant, D. (1969) Experimental investigation of absolute pitch. *J. Res. Mus. Educ.* **17,** 135–43.

Sessions, R. (1941). The composer and his message. In *The intent of the artist* (ed. A. Centeno). Princeton University Press, Princeton.

Shaffer, L. H. (1976). Intention and performance. *Psychol. Rev.* **83,** 375–93.

Shaffer, L. H. (1980). Analysing piano performance: a study of concert pianists. In *Tutorials in motor behaviour* (eds G. E. Stelmach and J. Requin). North Holland, Amsterdam.

Shaffer, L. H. (1981a). Creativity in skilled performance. Paper presented at NATO conference on adaptive control of ill-defined systems.

Shaffer, L. H. (1981b). Performance of Chopin, Bach, and Bartok: studies in motor programming. *Cog. Psychol.* **13,** 326–76.

Shepherd, J. (1977). Media, social process and music. In *Whose music? sociology of musical language* (by J. Shepherd, P. Virden, G. Vulliamy, and T. Wishart). Latimer, London.

Shepherd, J., Virden, P., Vulliamy, G., and Wishart, T. (1977). *Whose music? A sociology of musical language.* Latimer, London.

Shepherd, R. N. (1982). Structural representation of musical pitch. In *The psychology of music* (ed. D. Deutsch). Academic Press, New York.

Shiffrin, R. M. and Schneider, W. (1977). Towards a unitary model for selective attention, memory scanning and visual search. In *Attention and performance 6* (ed. S. Dornic). Erlbaum, Hillsdale, New Jersey.

Shuter-Dyson, R. and Gabriel, C. (1981). *The psychology of musical ability* (2nd edn). Methuen, London.

Siegel, J. A. (1974). Sensory and verbal coding strategies in subjects with absolute pitch. *J. Exp. Psychol.* **103,** 37–44.

Siegel, J. A. and Siegel, W. (1977a). Absolute identification of notes and intervals by musicians. *Perc. and Psychophys.* **21,** 143–52.

Simon, H. A. (1979). *Models of thought*. Yale University Press, New Haven.
Simon, H. A. and Sumner, R. K. (1968). Pattern in music. In *Formal representation of human judgement* (ed. B. Kleinmuntz). Wiley, New York.
Skinner, B. F. (1957). *Verbal behaviour*. Appleton Century Crofts, New York.
Sloboda, J. A. (1974). The eye-hand span: an approach to the study of sight-reading. *Psychol. Mus.* **2,** 4–10.
Sloboda, J. A. (1976a). Visual perception of musical notation: registering pitch symbols in memory. *Q. J. Exp. Psychol.* **28,** 1–16.
Sloboda, J. A. (1976b). The effect of item position on the likelihood of identification by inference in prose reading and music reading. *Can. J. Psychol.* **30,** 228–36.
Sloboda, J. A. (1976c). Decision times for word and letter search: a wholistic word identification model examined. *J. Verb. Learn. Verb. Behav.* **15,** 93–101.
Sloboda, J. A. (1977a). Phrase units as determinants of visual processing in music reading. *Br. J. Psychol.* **68,** 117–24.
Sloboda, J. A. (1977b). The locus of the word-priority effect in a target detection task. *Memory Cognition.* **5,** 371–6.
Sloboda, J. A. (1978). Perception of contour in music reading. *Perception* **6,** 323–31.
Sloboda, J. A. (1981). The uses of space in music notation. *Visible Lang.* **25,** 86–110.
Sloboda, J. A. (1982). Music performance. In *The psychology of music* (ed. D. Deutsch). Academic Press, New York.
Sloboda, J. A. (1983). The communication of musical metre in piano performance. *Q. Jl. Exp. Psychol.* **35,** 377–396.
Sloboda, J. A. (1984). Review of Wing and Seashore tests. In *Tests in education: a book of critical reviews* (eds H. Goldstein and P. Levy). Academic Press, London.
Sloboda, J. A. and Edworthy, J. (1981). Attending to two melodies at once: the effect of key relatedness. *Psychol. Mus.* **9,** 39–43.
Sloboda, J. A. and Gregory, A. H. (1980). The psychological reality of musical segments. *Can. J. Psychol.* **34,** 274–80.
Sloboda, J. A. and Parker, D. H. H. (1985). Immediate recall of melodies. In *Musical structure and cognition* (eds. P. Howell, I. Cross, and R. West). Academic Press, London.
Smith, J. (1983). Reproduction and representation of musical rhythms: the effects of musical skill. In *Acquisition of symbolic skills* (eds D. Rogers and J. A. Sloboda). Plenum, New York.
Smith, J., Hausfield, S., Power, R. P., and Gorta, A. (1982). Ambiguous musical figures and auditory streaming. *Percept. Psychophys.* **32,** 454–64.
Solso, R. L. (1980). *Cognitive psychology*. Harcourt Brace Jovanovitch, New York.
Spender, N. (1980). Psychology of music. In *The new Grove's dictionary of music and musicians* (ed. S. Sadie). MacMillan, London.
Spender, S. (1946). The making of a poem. *Partisan Rev.* Summer issue.
Springer, S. P. and Deutsch, G. (1981). *Left brain, right brain.* Freeman, San Francisco.
Steedman, M. J. (1977). The perception of musical rhythm and metre. *Perception* . **6,** 555–70.
Sternberg, S., Knoll, R. L., and Zukofsky, P. (1982). Timing by skilled musicians. In *The psychology of music* (ed. D. Deutsch). Academic Press, New York.
Strauss, R. (1949) *Betrachtungen und Erinenrungen.* Atlantis, Zurich.

Stravinsky, I. (1936). *Chronicle of my life.* Gollancz, London.
Stravinsky, I. (1969). *The Rite of Spring: sketches 1911–1913.* Facsimile reproductions with commentary by R. Craft. Boosey and Hawkes, London.
Sudnow, D. (1978). *Ways of the hand: the organization of improvised conduct.* Routledge and Kegan Paul, London.
Sundberg, J. (1978). Synthesis of singing. *Swed. J. Musicol.* **60,** 107–12.
Sundberg, J. (1982). Perception of singing. In *The psychology of music* (ed. D. Deutsch). Academic Press, New York.
Sundberg, J. and Lindblom, B. (1976). Generative theories in language and music descriptions. *Cognition.* **4,** 99–122.
Sutherland, N. S. (1973). Object recognition. In *Handbook of perception* (eds E. C. Carterette and M. P. Friedman) Vol. 3. Academic Press, New York.
Tan, N., Aiello, R., and Bever, T. G. (1981). Harmonic structure as a determinant of melodic organization. *Memory and Cognition.* **9,** 533–9.
Taylor, J. (ed.) (1932). *Selected writings of John Hughlings Jackson.* Hodder and Stoughton, London.
Triesman, A. M. (1964). Selective attention in man. *Br. Med. Bull.* **20,** 12–16.
Tucker, W. H., Bates, R. H. T., Frykberg, S. D., Howarth, R. J., Kennedy, W. K., Lamb, M. R., and Vaughan, R. G. (1977). An interactive aid for musicians. *Int. J. Man–Machine Stud.* **9,** 635–51.
Van Noorden, L. P. A. S. (1975). Temporal coherence in the perception n of tone sequences. Unpublished doctoral dissertation, Technische Hogeschool Eindhoven, The Netherlands.
Van Nuys, K., and Weaver, H. E. (1943). Memory span and visual pauses in reading rhythms and melodies. *Psychol. Monog.* **55,** 33–50.
Vernon, M. D. (1931). Characteristics of proof-reading. *Br. J. Psychol.* **21,** 368.
Vernon, P. E. (1970). *Creativity.* Penguin, London.
Virden, P. and Wishart, T. (1977). Some observations on the social stratification of twentieth century music. In *Whose music? A sociology of musical languages* (by J. Shepherd, P. Virden, G. Vulliamy, and T. Wishart). Latimer, London.
Ward, W. D. (1963a). Absolute pitch. Part I. *Sound.* **2.3,** 14–21.
Ward, W. D. (1963b). Absolute pitch. Part II. *Sound.* **2.4,** 33–41.
Ward, W. D. and Burns, E. M. (1982). Absolute pitch. In *The psychology of music* (ed. D. Deutsch). Academic Press, New York.
Weaver, H. E. (1943). A study of visual processes in reading differently constructed musical selections. *Psychol. Monog.* **55,** 1–30.
Werfel, F. and Stefan, P. (1942). *Verdi: the man and his letters.* L. B. Fischer, New York.
Wertheim, N. and Botez, M. I. (1961). Receptive amusia: a clinical analysis. *Brain.* **84,** 19–30.
Wertheimer, M. (1923). Untersuchung zur Lehre von der Gestalt. II. *Psychol. Forsch.* **4,** 301–50.
Whitehead, A. N. (1917). *The organization of thought.* Williams and Northgate, London.
Williams, J. P. (1970). From basic research on reading to educational practice. In *Basic studies in reading* (eds H. Levin and J. P. Williams). Basic Books, New York.
Williams, L. (1980). *The dancing chimpanzee* (2nd edn.). Allison and Busby, London.

Wing, H. D. (1962). *Wing standardized tests of musical intelligence* (1957 revision). National Foundation for Educational Research, Windsor.

Wishart, T. (1977). Musical writing, musical speaking. In *Whose music? A sociology of musical language* (by J. Shepherd, P. Virden, G. Vulliamy, and T. Wishart). Latimer, London.

Wohwill, J. F. (1981). Music and Piaget: spinning a slender thread. Paper presented at a meeting of the American Psychological Association, Los Angeles, August 1981.

Wolf, T. (1976). A cognitive model of musical sight-reading. *J. Psychol. Res.* **5**, 143–71.

Woodrow, H. (1951). The perception of time. In *Handbook of experimental psychology* (ed. S. S. Stevens). Wiley, New York.

Wyke, M. A. (1977). Musical ability: a neuropsychological interpretation. In *Music and the brain* (eds M. Critchley and R. A. Henson). Heinemann, London.

Yeston, M. (ed). (1977). *Readings in Schenker analysis*. Yale University Press, New Haven.

Zenatti, A. (1969). Le developpement genetique de la perception musicale. *Monog. Francais Psychol.* No. 17.

# Author index


Abberton, E. 233
Addis, A. B. 71
Aiello, R. 189
Allport, D. A. 166
Alvin, J. 9
Anderson, E. 3, 116
Anderson, J. R. 8, 36, 197, 215, 216, 224, 225
Antonis, B. 166
Apel, W. 248
Attneave, F. 4

Balzano, G. 182, 253, 255, 256, 257
Barbour, J. M. 178
Bartlett, J. C. 153, 179, 184
Bartok, B. 142
Basmajian, J. V. 233
Bates, R. H. 71, 232
Beethoven, L. van 107, 108, 116
Bent, I. 32
Benton, A. L. 262
Bernstein, L. 12
Bernuth, H. V. 198
Bertera, J. H. 264
Bever, T. G. 17, 34, 189, 264
Blacking, J. 242, 247, 266
Blechner, M. J. 28
Boden, M. 195
Bond, Z. S. 36
Botez, M. I. 263
Bower, G. H. 4, 81, 140, 187
Bower, T. G. R. 198
Bradshaw, J. L. 262
Brady, P. T. 177
Bransford, J. D. 152, 197
Bregman, A. S. 155, 156
Broadbent, D. E. 8, 34, 166
Brown, R. 182
Brown, R. R. 257
Buck, P. C. 90, 215, 231, 232
Bulbrook, M. E. 122
Bullivant, R. 123
Burns, E. M. 176, 177
Butler, D. 157, 182

Caldwell, J. 251
Campbell, J. 156
Carlsen, J. C. 165
Chang, H. 198, 199
Charles-Dominique, P. 266
Chase, W. G. 4, 81, 186
Cherry, E. C. 166
Chi, M. T. H. 186
Chiarello, R. J. 264
Chomsky, N. 11, 12, 13, 14, 15, 16, 17, 32, 36, 38, 65, 195
Clark, E. 152
Clark, H. H. 32, 36, 152
Clarke, E. 82
Clarkson, J. K. 93, 100
Claxton, G. 9
Clynes, M. 9
Cohen, A. 25
Cohen, A. J. 179, 180, 183
Cohen, N. 25
Cole, H. 248
Cole, R. A. 23
Coltheart, M. 262
Cooke, D. 60, 61, 62, 63, 64, 65
Cooper, F. S. 23
Corbit, J. D. 29
Craft, R. 108, 109, 110, 111, 112
Critchley, M. 9
Crocker, R. L. 257
Cuddy, L. L. 176, 177, 179, 180, 183
Cutting, J. E. 28, 29

Damasio, A. R. 262
Damasio, H. 262
Darwin, C. 265
Davidson, L. 200. 206
Davies, J. B. 2. 10, 153, 183, 188
Dawkins, R. 265
Day, R. S. 28
d'Azevedo, W. L. 243
Delis, D. 191
Deutsch, D. 9, 154, 155, 156, 157, 161, 162, 166, 175, 176, 177, 185, 186
Deutsch, G. 262
Deutsch, J. A. 93, 100, 166
Donaldson, M. 195
Dowling, W. J. 28, 153, 154, 160, 161, 166, 167, 179, 183, 184, 201, 203
Durkin, H. E. 122

Edworthy, J. 167, 183
Eimas, P. D. 29
Ericsson, K. A. 122, 123

Farnsworth, P. 3, 9
Feroe, J. 175, 185
Fitts, P. 216
Fitzgibbons, J. 162
Flavell, J. 195
Fleer, J. 191
Foard, C. F. 28
Fodor, J. A. 17, 34
Forte, A. 17, 104, 109
Fourcin, A. J. 233
Frykberg, S. D. 71, 232
Funk, J. 195
Futer, D. S. 260

Gabriel, C. 9, 62, 63, 64, 195, 213, 236, 238, 262
Gabrielsson, A. 30
Gaede, S. E. 264
Gardner, B. T. 17, 265
Gardner, H. 121, 195, 200, 202, 204, 206, 213, 214, 237
Gardner, R. A. 17, 265
Garrett, M. F. 17
Garrod, S. C. 58, 152
Gates, A. 262
Ghiselin, B. 116, 138
Gibson, E. J. 154
Goodman, N. 247
Goody, J. 243
Gorta, A. 162
Green, D. W. 37
Greene, J. 17
Gregory, A. H. 34, 35, 36, 189
Gruneberg, M. M. 8
Gruson, L. M. 61, 62, 63

Haber, R. N. 74
Halle, M. 36, 38
Halpern, A. R. 4, 81
Halwes, T. 29
Hamburger, M. 107, 116
Handel, S. 188
Harris, K. S. 24
Hausfield, S. 162
Healy, A. F. 74
Heise, G. A. 158, 160
Henry, L. K. 122
Henson, R. A. 9
Hevner, K. 2
Hindemith, P. 120, 121
Hirsh, I. J. 100
Hochberg, J. 155
Hodeir, A. 25
Hoppin, R. H. 251
Howarth, R. J. 71, 232
Hughes, E. 90
Hughlings-Jackson, J. 264
Humphrey, B. 261

Hunter, I. M. L. 209
Hutt, C. 198
Hutt, S. J. 198

Ibbotson, N. R. 99
Idson, W. L. 179
Imberty, M. 211, 237

Jackendoff, R. 57
Jairazbhoy, N. 257
Johanssen, G. 155
Johnson, D. P. 104, 106
Jones, M. R. 188
Joseph, L. G. 61
Jusczyck, P. 29

Kagan, J. 198
Kaplan, E. A. 37, 71
Karkoschka, E. 248
Kellar, L. 25, 26, 27, 28, 178
Kennedy, W. K. 71, 232
Kerr, P. 191
Kessen, W. 200
Kidd, G. 188
Kilmer, A. D. 257
Kimura, D. 261
Kinney, J. A. 24
Kinsbourne, M. 262
Klinger, E. 122
Klitzke, D. 80
Knoll, R. L. 30
Kosslyn, S. M. 20
Kramer, R. 104, 105
Kubovy, M. 155, 156, 159
Kuhn, T. S. 153
Kunst, J. 241

Ladefoged, P. 34
Lamb, M. R. 71, 232
Lane, H. 24
Lashley, K. 16, 88
Laske, O. 32
Lee, C. S. 49
Lehiste, I. 36
Lenard, H. G. 198
Lerdahl, F. 57
Levin, H. 37, 71, 197
Levine, J. 200
Lewis, D. 237
Lewis, R. 198
Liberman, A. M. 23, 24, 29
Lindblom, B. 12, 36, 38, 41, 47, 52, 119
Locke, S. 25, 26, 27, 28, 178
Longuet-Higgins, H. C. 43, 47, 49
Lord, A. B. 141, 142
Luria, A. R. 260
Lyons, J. 17

McAdam, D. W. 261
McKee, G. 261
McKernon, P. 200, 206
Makeig, S. 62
Malm, W. P. 241, 253, 257, 258
Manor, H. C. 237
Marin, O. S. M. 262, 263, 265
Marler, P. 266
Marr, D. 155
Marshall, J. C. 262
Martin, J. G. 17, 36
Massaro, D. W. 80, 179
Mattingley, I. G. 29
Meyer, L. B. 22, 64, 65, 102, 140, 162, 163, 164, 185, 190
Michon, J. A. 30
Miller, G. A. 158, 160, 190, 254
Miller, J. 179, 180, 183
Minsky, M. 140
Mitchell, D. C. 37
Moog, H. 199, 201, 202, 203, 204, 205, 206, 207, 208
Moore, B. C. J. 9
Morais, J. 264
Morgenstern, S. 107, 116
Morton, J. 99
Moynahan, E. D. 209
Mozart, W. A. 113, 116
Muntjewerft, W. J. 198

Narmour, E. 17
Navon, D. 80
Neisser, U. 8, 58, 103, 141, 167, 232
Nettl, B. 241
Newell, A. 8, 117, 224
Newmarch, R. 120, 137
Newton, W. J. 233
Nisbet, R. E. 121, 122
Nottebohm, G. 104

Olson, G. M. 32, 36
Olson, R. K. 4
O'Regan, K. 69
Ostwald, P. 200
Owen, D. 207

Palmer, S. E. 186
Parker, D. H. H. 191, 192
Parrott, I. 114
Parry, M. 141
Parsons, O. A. 264
Patterson, B. 93, 237
Patterson, K. 262
Peretz, I. 264
Petter, J. J. 266
Pflederer, M. 209, 210
Piaget, J. 194, 195, 209, 210
Pick, A. D. 178
Pillsbury, W. B. 74
Plomp, R. 9
Poincaré, H. 138
Pollard-Gott, L. 153
Pollatsek, A. 162
Pomerantz, J. R. 20
Povel, D. J. 188
Power, R. P. 162
Pratt, F. 193
Pressing, J. 253, 259
Priestley, M. 9
Pylyshyn, Z. 20

Rainbow, E. L. 207
Rasch, R. A. 93, 100
Rayner, K. 69
Read, G. 248
Reason, J. 221
Reber, A. S. 36
Reitman, W. 103, 123, 124, 125, 137
Reitmeyer, J. W. 231, 232
Restle, F. 186, 187
Reynolds, P. 166
Robinson, I. 17
Rosner, B. 28, 29, 140, 164, 185

Sadie, S. 242
Saetvit, J. G. 237
Salzer, F. 17, 152
Sampson, E. E. 240
Sanchez, M. 123
Sandford, A. J. 58, 152
Scharf, B. 173
Schenker, H. 11, 12, 13, 14, 15, 16, 17, 21, 65
Schindler, R. M. 74
Schlosser, L. 107
Schneider, W. 166, 167
Scott, B. 23
Seashore, C. E. 30, 82, 93, 100, 237
Sergeant, D. 176
Sessions, R. 116, 119
Shaffer, L. H. 17, 30, 71, 82, 83, 84, 88, 89, 93, 98, 101
Shankweiler, D. F. 23
Shaw, J. C. 117
Shepard, R. N. 154, 255, 257
Shepherd, J. 242, 243
Shiffrin, R. M. 166, 167
Shuter-Dyson, R. 9, 195, 213, 236, 238, 262
Siegel, J. A. 176, 178
Siegel, W. 178
Simon, H. A. 4, 8, 56, 81, 117, 122, 123, 179, 186, 224
Skinner, B. F. 16
Sloboda, J. A. 4, 30, 34, 35, 36, 37, 72, 74, 75, 78, 84, 101, 154, 167, 189, 192, 237, 238, 248, 249

Smith, J. 162, 187
Smith, W. L. 262
Solso, R. L. 8
Spender, N. 154
Spender, S. 138
Springer, S. P. 262
Steedman, M. 49, 188
Stefan, P. 118
Sternberg, S. 30
Strauss, R. 116
Stravinsky, I. 112
Studdert-Kennedy, M. 23
Sudnow, D. 143–148
Sumner, R. K. 56, 186
Sundberg, J. 12, 36, 38, 41, 47, 52, 100, 119
Sutherland, N. S. 155
Syrdal, A. K. 29

Tan, R. 189
Taylor, J. 264
Tchaikovsky, P. I. 120, 136
Thomas, I. B. 162
Todd, P. 188
Trehub, S. E. 198, 199
Tsvetkova, L. S. 260
Tucker, W. H. 71, 232

Van Noorden, L. P. A. S. 158, 160, 163
Van Nuys, K. 70
Vaughan, R. G. 71, 232
Verdi, G. 118
Vernon, M. D. 74
Vernon, P. E. 116, 120, 137, 138
Virden, P. 25, 242
Vulliamy, G. 242

Ward, W. D. 154, 176, 177
Watt, I. 243
Weaver, H. E. 70
Wendrich, K. A. 200
Werfel, F. 118
Wertheim, N. 263
Wertheimer, M. 154
Wetzel, R. 188
Whiteside, J. ???
Williams, J. P. 197, 230
Williams, L. 18
Wilson, J. D. 121, 122
Wing, H. D. ???
Winzenz, D. 187
Wishart, T. 25, 242, 249, 252
Wohlwill, J. 210
Wolf, D. 195, 206, 214
Wolf, T. 74, 81
Woodrow, H. 29, 47
Wyke, M. 262

Yeston, M. 17

Zenatti, A. 210, 237
Zukofsky, P. 30

# Subject index


Aboriginal song 258, 266
absolute pitch 28, 176–8
accenting 49, 258–9; *see also* metre and rhythm
Africa, music of 240–1, 247
Alberti bass 55
Allegri, G., *Miserere* 3, 114, 192
alphabets, in tonality 185–6
ambiguity
  harmonic 44–7
  rhythmic 50–1
analysis, of musical compositions 102
analytic processing 264–5
ape vocalizations 18, 265–6
*appogiatura* 53
aptitude, musical 198, 233–8
artificial intelligence 12
assessment, of musical aptitude 198, 233–8
associative stage, of skill acquisition 216–18
attention
  in listening 166–74
  in performance 93
auditory grouping 155–66
auditory threshold, for pitch 173
augmented fifth chord 43
Australia, Aboriginal music of 258, 266
autonomous stage, in skill acquisition 216–17

barbling 200
babies, musical awareness of 198–202
Bach, J. S. 82, 118
  Clavier Concerto in D minor 83
  French Suite no. 5 in G 55
  Gigue Fugue for organ 50
  'Goldberg' Variations no. 18 171
  two part invention in E 174
  Well-tempered Clavier
    Book 1 Prelude 2 184
    Book 2 Fugue 7 83
      Fugue 12 163
      Fugue 13 47
      Prelude 15 158
      Fugue 15 42
      Prelude 20 42
Barber, Samuel, 'Sea Snatch' Op. 29, no. 6 95
*Bassbrechung* 14
Beethoven, L. van 102, 112
  compositional processes of 116
  improvisation of 139
  Piano Sonata Op. 10, no. 3 104–12
    Op. 14, no. 1 45–6
    Op. 14, no. 2 50–1
  sketches in composition 104–12
  Symphony no. 5 184
bird song 18
bowed sounds, perception of 28
Brahms, J., Capriccio Op. 76, no. 2 74–5
brain, localization of musical function in 260–5
Brown, Rosemary, compositional processes 114–15, 121, 138

canon 171
categorical perception
  of duration (rhythm) 28–31
  effects of musical training on 25–7
  of pitch (frequency) 24–8, 176–8
  of speech 23–4
cerebral hemispheres 261–5
chess, memory for 4, 81
children, development of musical skills in 198–213
children's song 18, 202–7
choral composition 125
chord 43–5
clarinet, teaching of 217–18
click location, in music and language 34–7
closure 190
cognitive psychology 7–9, 239–40
  and education 196–8
cognitive stage, of skill acquisition 216–24
composition
  compared to improvisation 149
  execution in 116
  inspiration in 116
  as problem solving 117–20
  protocols of 123–35
  role of sketches in 104–12
  unconscious processes in 114, 119
computer synthesis
  of horn sounds 100
  of plucked and bowed sounds 28–9
  of speech sounds 23–4
conditioning 2
conservation 209–10
constituent structure trees 13–15, 34–5

contour
  babies' awareness of 198–9
  melodic 54
  role in memory 183–5
  in reading music 80
creative activity 138
critical band 173
culture, and music 240–59

dactyl, in musical rhythm 49, 259
Darwinian evolution 265–6
deep structure, in language and music 13–22
designative meaning 64–5
dichotic presentation 261
digital mapping 206
diminished seventh chord 43
discreteness, in musical practice 249–50
discrimination, of pitch 177
distributed practice 226–7
dominant seventh chord 43
Dukas, P., *The Sorcerer's Apprentice* 59
Dussek, J. 75
dynamic control, in wind performance 93

educational psychology 196–8
embodied meaning 64–5
emotion 1–3
  meaning in tonal relations 60–5
enculturation 195–6, 198–215
ensemble performance
  synchronization in 93, 100
  violin in 100
epic poetry, of Yugoslavia 141–2
Escher, M. C. 56
ethnomusicology 240–2
execution, in composition 116
expectation, in music 65
expert performance 93–101
expressive performance 81–91
  perception of 85–8
eye movements in reading
  of language 69–70
  of music 70–3
  of piano music 70–1

factual knowledge, in skill acquisition 216–23
feedback
  in performance 101
  in skill acquisition 216, 223–4
figure–ground reversal, in musical attention 168–71
fingering
  in piano performance 96–7
  in skill acquisition 231–2
finite state grammar 16
fission, in pitch streaming 158
form, effects of literacy on 246–8
fugue 123

functions, of music 266–8
fundamental frequency 173
fusion, in pitch streaming 158

general problem solver 224
generative grammar 13–17, 32–41
  for Swedish Nursery Tunes 38–41
gestalt principles, of perception 154–6
goal stack, in production systems 221
Goldovsky, B. 74–5
good continuation, principle of 159
grammar
  generative 32–41
  in language and music 32–57
  rules of 32–3
Gregorian chant 142, 248–9
group cohesive role, for music 266
group theory, applied to musical scales 255–7
Guido of Arrezzo 251–2

habits, in skill acquisition 216
hand independence, in piano performance 93
Handel, G. F. 112
harmonic series 21–2
harmonic structure
  effects on memory for notation 81
  role of in reading 72
harmony
  ambiguity in 44–7
  perception of 42–7
  role of in reading 72
hierarchical structure 186
holistic processing 264–5
horizontal structure, in polyphony 167–9

implication, in music 65, 163–5
improvisation 138–50
  compared to composition 149
  in jazz 142–50
Indian music
  Sa-Grama scale of 254–5, 257
  metrical schemes in 259
inference, in reading 74
inspiration, in composition 116

Java, Gamelan music of 258
jazz
  blue notes in 24–5
  improvisation 142–50

key, musical
  modulation of 44–5
  perception of 42–7, 179–82
key relation, effect on attention 167–9
Kodaly, Z. 118

language
  reading of 68–7(
  relationship of music to 11–66
literacy 18–19
  and culture 242–53
localization, of sounds 155–6

masking, of pitch 172–3, 179
massed practice 226
meaning, in music
  designative 64–5
  embodied 64–5
  *see also* semantics
melisma 142
melody 52–5, 63
  babies' awareness of 198–202
  children's recognition of 208
  construction of 52–5
  recall of 191–2
memory
  for chess positions 4
  for expressive variations 82
  for individual pitches 175–9
  for melody 191–2, 201–6
  for musical notation 4, 78–81
  in oral and literate cultures 242–6
  in performance 94–8
  for rhythms 187–8
Messiaen, O. 97
metre
  communication of 47–52, 82–5
  perception of 47–52
  role of in reading 73
  universal functions of 257–9
mimicry, in music 59
mnemonic aid, music as 268
modulation of key 44–5
motor programming, in music performance 88
movement to music, in children 207
Mozart, W. A. 19–20, 63, 102, 112, 121
  autograph of *Die Zauberflöte* 113
  compositional processes of 113–14, 116
  Concerto for Two Pianos, K. 365 164
  manuscripts of 102
  memory capacity of 3, 192–3
  Piano Sonata in C, K. 545 55, 172
musical judgement, development of 211–14

negative practice 231–2
neumes 248–9
neural organization, for music 260–5
notation 18–19
  and culture 242–53
  western, history of 248–52

onset asynchrony, of sounds 156
oral cultures 242–8
origin, of music 265–8

paradigms, in research 153
passing note 53
pattern, role of in memory 185–8
perfect pitch, *see* absolute pitch
performance of music 67–101
phi phenomenon 159
phoneme 23–4
phrase, perception of boundaries 36–7, 72
piano performance
  communication of metre in 82–5
  fingering in 96–7
  hand independence in 93
pitch, categorical perception of 24–8
pitch streaming 100, 158–65
plucked sounds, perception of 28
poetry 64
polyphonic structure
  perception of 166–74
  relation to notation 249
practice, *see* rehearsal
preview, in reading 69, 71, 87
problem solving, composition as 117–20
problem space 123
procedural knowledge, in skill acquisition 216–23
production deficiency 19
production system theory 215–29
programmatic music 59
proof-reader's error
  in language 74
  in music 74–9
protocol
  of choral composition 125–35
  of fugue composition 123–5
  strengths and weaknesses of 136–8

Rachmaninov, S.
  orchestral music of 171
  Suite for Two Pianos Op. 17 157–8
reading 37
  effects of brain damage on 262–3
  eye movements in 69–73
  of language 68–70
  of piano music 70–1
  of sight 67–90, 252
reference, in music 59
recall
  of melodies 191–2
  in oral and literate cultures 242, 245–6
  of rhythms 187–8
rehearsal 90–3, 226–7, 231–2
relative pitch 177–8
reliability, in testing 235–6
repetition, in skill acquisition 216, 223, 231–2

rhythm
  ambiguity in 50–1
  in babies 199–201
  categorical perception of 28–31
  communication of 47–52
  in different cultures 257–9
  perception of 47–52
  reproduction of 30–1, 187–8
  role of in reading 72
rise time, in musical sounds 28–9
Rowles, Jimmy, jazz playing of 147

scale, musical 24–5
  blue notes in jazz 24–5
  diatonic, centrality of 256–7
  in different cultures 254–5
  pentatonic 25, 257
  relation to memory 176–85
  weighted 253
scale illusion 156, 161
Schenker analysis 14–17, 21–2, 136
Schumann, R. 118
Seashore measures of musical talents 237–8
segmentation, in musical memory 189
semantics, musical 57–66
  and tonal relations 60–4
Sessions, R., compositional processes of 116
Shebalin, V. G. 260
sight reading 67–90, 166
  teaching of 197
singing, of babies 200
sketches
  of Beethoven 104–12
  in composition 104–12
  of Stravinsky 108–12
skill acquisition 215–36
sol-fa 251
song
  development of in children 202–7
  social functions of 266–8
span
  eye–hand in music reading 37, 71–2
  eye–voice in language reading 37
specialization, of cerebral hemispheres 261–5
specificity hypothesis, for language and music 17–18
spectral analysis, of horn sounds 100
spontaneous songs, in children 202–4
streams, in skill acquisition 193
story telling, relation to musical improvization 139–41
strategies, of teaching and learning 220
Strauss, R., compositional processes of 116
Stravinsky, I. 118
  sketches in composition 108–12
  *The Rite of Spring* 108–12
stress, *see* metre, rhythm
style sensitivity, in children 213–14
surface structure, in language and music 13–22
Suzuki method of violin teaching 68
synchronization, in ensemble performance 93, 100
syncopation 50
syntax, in language and music 32–57

tapping, rhythmic 99
Tchaikowsky, P. I.
  compositional processes of 120
  6th Symphony 2, 157
teaching
  of instrumental skill 228–32
  of performance 88
  of rehearsal techniques 90–3
Tegner, Alice 38–41
temperament 178
tension, in music 21–2
tests, of musical aptitude 198, 233–8
theme and variation, in improvisation 139
timing
  in different cultures 257–9
  in performance 73, 98–9
tonality 12, 42–7
  and emotional meaning 60–4
tonic
  chord 63
  in different cultures 254
  importance for memory 180–2
  in musical scale or key 42–7, 254
topological mapping 206
training, effects on
  categorical perception 25–7
  hemispheric specialization 264
  music reading 80–1
  performance 98–9
  rehearsal behaviour 72–3
  reproduction of rhythms 188
  rhythmic tapping 99
transcription, of pitch symbols 78–81
transformational grammar 13–17
transposition
  in musical recall 4–5
  recognition of 179–82
tritone 182, 255
tuning 178
  of major scales 255
  perception of 27
twelve-tone music 252
typing skill 71

unconscious processes, in composition 114, 119
universals
  in language 19–20
  in music 12, 19–22, 253–9
upper partials, of complex sounds 173

*Urlinie* 14, 21
*Ursatz* 14–17, 21–2

validity, in testing 233
variability, in expert performance 97
verbal reports, of compositional processes 114–23
Verdi, G. 118
  *La donna è mobile* 53–5
vertical structure, in polyphony 167
vibrato, in performance 93
violin
  in ensemble performance 100
  teaching of 68, 218

waves of symbolization 193, 196, 206
weighted scale 253
Wing standardized tests of musical intelligence 237–8
working memory, in production systems 221

Yugoslavia, epic poetry of 141–2

What are the mental processes involved in listening to, performing, and composing music? What is involved in 'understanding' a piece of music? How are such skills acquired? Questions such as these form the basis of the cognitive psychology of music. The author addresses these questions by surveying the growing experimental literature on the subject. The topics covered will be of interest to psychologists as windows on to a human cognitive skill of some complexity that is only now beginning to receive the attention that has been given to such skills as language. The book will also be of relevance to musicians who are seeking to understand the psychological bases of their skills. The author does not simply review existing research, but takes a critical look at what has been achieved in the subject, introducing such topics as composition and musical skill in non-literate cultures. He draws freely on his own knowledge and experience as a practising musician as well as a psychologist to provide an overview that is scholarly and also accessible to the general reader.

**Dr Sloboda** is a Professor in psychology at the University of Keele; he is also a pianist and the director of a chamber choir.

. . . the clarity of Sloboda's writing and his numerous suggestions for further research will make the book essential reading for anyone, student or researcher, interested in how minds and music interact.
*Nature*

Viewed overall, Sloboda's book deserves loud applause, for its content, for its frankness in dealing with speculative issues largely ignored by other workers, and especially because it brings with it a new kind of approach. *New Scientist*

Sloboda's book must be an essential purchase for all those with an interest in the psychology of music . . . *British Journal of Psychology*

A wholly admirable book which may yet re-establish the place of music psychology in civilized society. *Music and Letters*

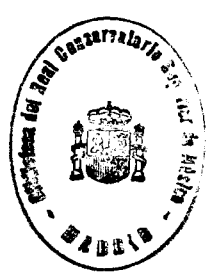
ISBN 0-19-852128-6




OXFORD UNIVERSITY PRESS

ACADEMIC PRESS SERIES IN
COGNITION AND PERCEPTION

# The Psychology of Music



This is a volume in

**ACADEMIC PRESS**
**SERIES IN COGNITION AND PERCEPTION**

A Series of Monographs and Treatises

A complete list of titles in this series appears at the end of this volume.

# The Psychology of Music

EDITED BY

Diana Deutsch

*Department of Psychology*
*University of California, San Diego*
*La Jolla, California*

R.º 63.073
8-IV-1994



**ACADEMIC PRESS, INC.**

(Harcourt Brace Jovanovich, Publishers)

**Orlando San Diego New York London**

**Toronto Montreal Sydney Tokyo**



COPYRIGHT © 1982, BY ACADEMIC PRESS, INC.
ALL RIGHTS RESERVED.
NO PART OF THIS PUBLICATION MAY BE REPRODUCED OR TRANSMITTED IN ANY FORM OR BY ANY MEANS, ELECTRONIC OR MECHANICAL, INCLUDING PHOTOCOPY, RECORDING, OR ANY INFORMATION STORAGE AND RETRIEVAL SYSTEM, WITHOUT PERMISSION IN WRITING FROM THE PUBLISHER.

ACADEMIC PRESS, INC.
Orlando, Florida 32887

*United Kingdom Edition published by*
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1 7DX

Library of Congress Cataloging in Publication Data

Main entry under title:

The Psychology of music.

(AP series in cognition and perception)
Includes index.
1. Music--Psychology. I. Deutsch, Diana.
II. Series.
ML3830.P9 781'.15 82-1646
ISBN 0-12-213560-1 AACR2

PRINTED IN THE UNITED STATES OF AMERICA

85 86 87 88 9 8 7 6 5 4 3 2

# Contents

List of Contributors xi
Preface xiii

1. The Perception of Musical Tones
*R. A. Rasch and R. Plomp*

I. The Psychoacoustics of Music 1
II. Perceptual Attributes of Single Tones 6
III. Perceptual Attributes of Simultaneous Tones 14
IV. Conclusion 21
References 21

2. Exploration of Timbre by Analysis and Synthesis
*Jean-Claude Risset and David L. Wessel*

I. Timbre 26
II. Timbre and the Fourier Spectrum: The Classical View 26
III. The Shortcomings of the Classical Conception 29
IV. Attack Transients 29
V. Complexity of Sounds: Importance of Characteristic Features 30
VI. Instrumental and Vocal Timbres: Additive Synthesis 30
VII. Cross Synthesis and Voice Synthesis 34
VIII. Additive Synthesis: Percussion Instruments 37
IX. Substractive Synthesis 39
X. Acoustic Modeling as a Synthesis Technique 41
XI. The Importance of Context 42
XII. Analysis-Synthesis as Fitting Acoustic and Perceptual Models to Data 44
XIII. The Use of Analysis-Synthesis Models of Timbre 45
XIV. Timbral Space 47
XV. Conclusion 49
Appendix: Signal Representations and Analysis-Synthesis Processes 50
References 54

3. Perception of Singing
*Johan Sundberg*

I. Introduction 59
II. Function of the Voice 60
III. Resonatory Aspects 62
IV. Phonation 76
V. Vibrato 82
VI. Pitch Accuracy in Singing Practice 89
VII. Phrasing and Emotion 91
VIII. Concluding Remarks 94
References 95

4. Grouping Mechanisms in Music
*Diana Deutsch*

I. Introduction 99
II. Grouping Principles 100
III. Two-Channel Listening to Melodic Sequences 101
IV. Channeling of Rapid Sequences of Single Tones 118
V. Voluntary Attention 127
VI. Conclusion 130
References 130

5. The Listener and the Acoustic Environment
*R. A. Rasch and R. Plomp*

I. Introduction 135
II. Methodology 137
III. Level Effects of Indirect Sound: Loudness 140
IV. Temporal Effects of Indirect Sound: Definition 141
V. Spatial Effects of Indirect Sound: Spaciousness 142
VI. The Compromise between Definition and Spaciousness 145
VII. Conclusion 146
References 146

6. Rhythm and Tempo
*Paul Fraisse*

I. Definitions 149
II. Rhythm and Spontaneous Tempo 151
III. Rhythmic Forms 157
IV. The Perception of Musical Rhythms 170
V. Conclusion 175
References 177

7. Timing by Skilled Musicians
*Saul Sternberg, Ronald L. Knoll, and Paul Zukofsky*

I. Perception, Production, and Imitation of Fractions of the Beat 182
II. Perceptual Judgment of Beat Fractions 187
III. Production of Beat Fractions 198
IV. Imitation of Beat Fractions 207
V. A Shared-Process Model of the Perception, Production, and Imitation of Beat Fractions 212
VI. Further Analysis of Perceptual Judgment 215
VII. Further Analysis of Production 224
VIII. Summary 229
Glossary 231
Appendix 231
References 237

8. Intervals, Scales, and Tuning
*Edward M. Burns and W. Dixon Ward*

I. Introduction 241
II. Are Scales Necessary? 243
III. Musical Interval Perception 246
IV. Natural Intervals and Scales 255
V. Conclusions and Caveats 264
References 265

9. The Processing of Pitch Combinations
*Diana Deutsch*

I. Introduction 271
II. Feature Abstraction 272
III. Higher Order Abstractions 282
IV. Alphabets and Hierarchies 287
V. Memory Systems 291
VI. Conclusions 311
References 312

10. Melodic Processes and the Perception of Music
*Burton S. Rosner and Leonard B. Meyer*

I. The Perception and Classification of Two Archetypal Melodic Processes 317
II. Experimental Findings 326
III. Implications 339
References 340



11. Structural Representations of Musical Pitch
*Roger N. Shepard*
I. Introduction 344
II. Unidimensional Approaches to Pitch 344
III. Potentially Multidimensional Approaches to Pitch 347
IV. The Spatial Representation of Pitch 350
V. Illustrative Analyses of Empirical Data 365
VI. Discussion 369
References 385

12. Musical Ability
*Rosamund Shuter-Dyson*
I. Concepts of Musical Ability 391
II. Correlational and Factorial Studies of Musical Ability 393
III. Musical Ability and Other Intellectual Abilities 404
References 408

13. Melodic Information Processing and Its Development
*W. Jay Dowling*
I. Introduction 413
II. Development 415
III. Adult Memory 421
IV. Contour versus Interval 427
V. Summary 427
References 428

14. Absolute Pitch
*W. Dixon Ward and Edward M. Burns*
I. Introduction 431
II. Genesis of AP 434
III. Measurement of AP 436
IV. Stability of the Internal Standard 444
V. Learning AP 445
VI. The Value of AP 447
References 449

15. Neurological Aspects of Music Perception and Performance
*Oscar S. M. Marin*
I. Introduction 453

II. Amusia 454
III. Auditory Agnosia and Verbal Deafness 462
IV. General Comments 466
References 473

16. Music Performance
*John A. Sloboda*
I. Introduction 479
II. The Nature of Performance Plans 480
III. Acquisition of Performance Plans 483
IV. The Role of Feedback in Performance 488
V. Social Factors in Performance 491
VI. Summary 494
References 494

17. Social Interaction and Musical Preference
*Vladimir J. Konečni*
I. Introduction 497
II. Effects of Social Stimulation on Aesthetic Choice 502
III. Effects of Information Load and Arousing Nonsocial Stimulation on Aesthetic Choice 505
IV. Effects of Listening to Melodies Differing in Complexity on Emotional States and Social Behavior 507
V. Listeners' Sequencing and "Chunking" of Musical Materials and the Use of Music for Mood Optimization 511
References 515

18. New Music and Psychology
*Robert Erickson*
I. Introduction 517
II. Music Theory and Music 519
III. Understanding Tonality 520
IV. Music and Perceptual Streaming 523
V. Fused Sounds in Music 529
VI. Music Theory and Experimental Science 534
References 535

Index 537

# List of Contributors

*Numbers in parentheses indicate the pages on which the authors' contributions begin.*

Edward M. Burns (241, 431), Department of Audiology and Speech Sciences, Purdue University, West Lafayette, Indiana 47907

Diana Deutsch (99, 271), Department of Psychology, University of California, San Diego, La Jolla, California 92093

W. Jay Dowling (413), Program in Psychology and Human Development, University of Texas at Dallas, Richardson, Texas 75080

Robert Erickson (517), Department of Music, University of California, San Diego, La Jolla, California 92093

Paul Fraisse (149), Laboratoire de Psychologie Expérimentale et Comparée, Université René Descartes, Paris, France

Ronald L. Knoll (181), Bell Laboratories, Murray Hill, New Jersey 07974

Vladimir J. Konečni (497), Department of Psychology, University of California, San Diego, La Jolla, California 92093

Oscar S. M. Marin (453), Department of Neurology, Good Samaritan Hospital and Medical Center, Portland, Oregon 97210

Leonard B. Meyer (317), Department of Music, University of Pennsylvania, Philadelphia, Pennsylvania 19104

R. Plomp (1, 135), Institute for Perception TNO, Soesterberg, The Netherlands, and Faculty of Medicine, Free University, Amsterdam, The Netherlands

R. A. Rasch (1, 135), Institute of Musicology, University of Utrecht, Utrecht, The Netherlands

Jean-Claude Risset (25), Faculté des Sciences de Luminy et Laboratoire de Mécanique et d'Acoustique, Marseille, France

Burton S. Rosner (317), Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Roger N. Shepard (343), Department of Psychology, Stanford University, Stanford, California 94305

Rosamund Shuter-Dyson[1] (391), Senior Lecturer in Psychology, Hatfield Polytechnic, Hertfordshire, England

John A. Sloboda (479), Department of Psychology, University of Keele, Keele, Staffordshire, England

Saul Sternberg (181), Bell Laboratories, Murray Hill, New Jersey 07974

Johan Sundberg (59), Department of Speech Communication and Music Acoustics, Royal Institute of Technology, S-100 44 Stockholm, Sweden

W. Dixon Ward (241, 431), Hearing Research Laboratory, University of Minnesota, Minneapolis, Minnesota 55455

David L. Wessel (25), Institute de Recherche et Coordination, Acoustique-Musique, Paris, France

Paul Zukofsky[2] (181), Bell Laboratories, Murray Hill, New Jersey 07974

[1]Present address: 6 Rectory Close, Tadley, Basingstoke, Hampshire RG26 6PH, England.

[2]Present address: Musical Observations, Inc., Port Jefferson, New York 11777.

# Preface

*We shall advance to our conclusions by strict demonstration.*
Aristoxenus: *Harmonic Elements*

The above words were written by Aristoxenus of Tarentum (ca. 320 B.C.) in the course of arguing that musical phenomena can be understood only through scientific investigation of human processing mechanisms. Reasonable as this view may seem to modern psychologists, it ran counter to the prevailing theoretical stance, based on Pythagorean doctrine, which has profoundly influenced the study of music throughout history. Most characteristic of this stance is a strong distrust of the evidence of the senses, together with the belief that music ought to be investigated by contemplation of numerical relationships alone. As Anaxagoras (ca. 499–428 B.C.) put it: "Through the weakness of our sense perceptions we cannot judge truth." And later as Boethius, the leading music theorist of the Middle Ages and a strong follower of Pythagoras, wrote in *De Institutione Musicae:* "For what need is there of speaking further concerning the error of the senses when this same faculty of sensing is neither equal in all men, nor at all times equal within the same man? Therefore anyone vainly puts his trust in a changing judgment since he aspires to seek the truth."

There are several good reasons that such a rationalistic position should have been adopted at the time and should have persisted for so long. One reason was a lack of understanding of the nature of sound. One can see that the inability to characterize a physical stimulus should have inhibited development of theories concerning how this stimulus is processed. A related problem was the lack of stimulus control, which made experimentation difficult. Another problem was the lack of mathematical techniques appropriate for the study of probabilistic phenomena.

These factors considered, it appears as no coincidence that the last decade has seen a sudden flowering of interest in the empirical study of music on the part of both psychologists and musicians. The nature of sound is now well understood, and the required mathematical techniques have been developed. Recent advances in computer

technology have enabled investigators for the first time to generate complex sound stimuli with versatility and precision. It has thus become possible to explore such issues as auditory shape analysis, attentional mechanisms in music, the organization of memory for musical information, and so on, with the stimulus control required for tight experimentation. Researchers have been able to build on conceptual frameworks recently developed in other branches of psychology and in music theory. This same technological development has led composers to experiment with the computer as a compositional tool. In so doing, numerous questions in perceptual and cognitive psychology that are of both practical interest to composers and theoretical interest to psychologists have been raised. As a result of this developing interest on the part of both scientists and musicians, we are experiencing a rapid expansion of work in the area of musical processing; and perhaps more importantly, collaboration between musicians and scientists is a rapidly growing phenomenon.

The purpose of this volume is to draw together the diverse and scattered literature on musical processing that has accumulated particularly over the last decade. The work is intended as a reference source for all those interested in music and the way it is processed by the listener and the performer. It is also intended as a reference source for perceptual and cognitive psychologists, who will find in the study of music much that is of general interest to their fields. The volume may also be used as a text for courses in the psychology of music.

The volume opens with a chapter on *The Perception of Musical Tones*, by Rasch and Plomp, which surveys the classical psychoacoustical literature on tone perception, focusing on characteristics of particular relevance to music. The attributes of pitch, loudness, and timbre are examined, as are psychoacoustical phenomena that occur when tones are presented simultaneously, such as beats and roughness, combination tones, and consonance and dissonance. The authors also provide a useful summary of research methods in psychoacoustics.

Timbre perception is examined in depth by Risset and Wessel in *Exploration of Timbre by Analysis and Synthesis* (Chapter 2). The authors address several issues that are of importance both to perceptual psychology and to contemporary music. For example, How is it that we can identify the sound of an instrument regardless of its pitch or loudness? What types of information can be discarded in the synthetic replication of a sound without distorting perceived timbre? How are the timbres of different instruments perceptually related? By investigating such questions it has proved possible to modify timbres in intriguing ways, for example, to extrapolate a timbre beyond its instrument register or to create a sequence of timbres that progress regularly from one instrument type to another. As the authors note, this field of research is likely to have a profound impact on the development of new music.

Johan Sundberg's chapter on *The Perception of Singing* (Chapter 3) examines some closely related issues. For example, How are we able to hear a singer's voice against a loud or orchestral background? How do we manage to identify sung vowels even though these differ markedly from those in speech? Why do we perceive singing as a set of discrete pitches even though the fundamental frequency events are not discrete? These and other questions are expertly evaluated.

Consider what happens when we listen to a live performance by an orchestra. The mixture of sounds that reaches our ears is produced by many instruments playing in parallel. Somehow our auditory system is able to sort out this mixture of sounds, so that we may choose to listen to a particular instrument or to a particular melodic configuration. What are the mechanisms whereby such groupings are achieved? To what extent are they under voluntary control? These questions are explored in *Grouping Mechanisms in Music* (Deutsch, Chapter 4). It is also shown that under certain circumstances the listener may perceptually reorganize what he or she hears, so that striking illusions may occur. When this happens, the listener's perceptions may not accord with the intentions of the composer.

As implied in Berlioz's statement "There is no such thing as music in the open air," the enclosed space of the concert hall contributes much to the aesthetic quality of music, through the complex sound reflections to which it gives rise. Experiments on the subjective effects of different sound fields are reviewed by Rasch and Plomp in *The Listener and the Acoustic Environment* (Chapter 5). The results of such studies have important practical implications for the design of concert halls.

The next two chapters in the volume are concerned with temporal aspects of music. In Chapter 6 on *Rhythm and Tempo* Paul Fraisse thoroughly surveys a field which he has so ably pioneered, including a review of his own classical contributions. This chapter is complemented by an in-depth study on *Timing by Skilled Musicians*, by Sternberg, Knoll, and Zukofsky (Chapter 7). As the authors write, "During ensemble rehearsal and performance, players must judge, produce and imitate beat fractions. It is plausible that because of the requirement that players 'keep together,' performing experience would cause the three functions to become at least consistent with each other and probably accurate as well. Neither of these expectations was borne out by our experiments." The results of this unique study are particularly valuable; such eminent musicians as the composer and conductor Pierre Boulez and the violinist and conductor Paul Zukofsky were employed as subjects. Based on their findings, the authors propose an information-flow model of timing which also has general implications.

The next four chapters are concerned with abstract structures formed by pitch relationships in music. The large majority of musical cultures employ scales. In Chapter 8, *Intervals, Scales, and Tuning*, Burns and Ward explore the perception of musical scales in detail, focusing on the interval as the basic building block. The authors argue that the use of a small set of discrete pitch relationships in music is probably dictated by inherent limitations in our ability to transmit sensory information. Such limitations probably also account for the phenomenon of categorical perception of musical intervals, the evidence for which is thoroughly reviewed. Arguments for different tuning systems are explored, and the authors conclude that patterns of intonation used by musicians are primarily a function of their acquired ability to reproduce learned interval categories.

The *Processing of Pitch Combinations* (Deutsch, Chapter 9) examines how pitch information is internally represented at different levels of abstraction. At the lowest level, local features such as intervals and pitch classes are represented, as are global

features such as contour. At the next higher level, such features are combined so as to give rise to perceptual equivalences and similarities between larger sets of pitches. At a yet-higher level, pitch information is mapped onto a set of highly overlearned alphabets and is retained in the form of hierarchies. A distinction is drawn between musical abstractions that result from passive "bottom-up" processes and those that are the result of hypothesis testing or are "top-down" in nature. The final section of the chapter is concerned with memory for pitch at these different levels of abstraction.

The internal representation of musical hierarchies is explored in an elegant collaborative study by Rosner and Meyer on *Melodic Processes and the Perception of Music* (Chapter 10). The authors argue, following Meyer's earlier theoretical work, that melodies often appear to be hierarchically structured in such a way that the type of patterning exhibited by a given melody differs from one hierarchical level to the next. The authors further hypothesize that "melodic patterns are classified by listeners, as well as music theorists, in terms of the organization of the highest level on which significant closure is created by the parameters that shape musical relationships." Using a concept identification task, the authors found that musically untrained listeners did indeed classify melodies in accordance with theoretical expectations.

The question of how pitch relationships within a musical system can best be represented has traditionally been the main focus of music theory. In his provocative *Structural Representations of Musical Pitch* (Chapter 11) Roger Shepard proposes that pitch in our tonal system be represented as a double helical structure requiring an embedding space of five dimensions. In arguing for this model, Shepard draws on recent empirical findings from psychology and psychoacoustics, as well as on considerations from tonal music.

The next four chapters consider musical ability from different points of view. In *Musical Ability* (Chapter 12) Rosamund Shuter-Dyson describes and critically evaluates different tests of musical ability. Various questions are then explored. For example, How does musical ability correlate with other intellectual abilities? How does musical ability develop with age? To what extent is musical ability a learned phenomenon and to what extent does it depend on training? Cases of exceptional musical ability are also described. This chapter should serve as a particularly useful reference for music educators.

The question of how the ability to process melodic information develops with age is explored in detail by Dowling in *Melodic Information Processing and Its Development* (Chapter 13). Dowling argues for a developmental sequence that begins with the infant's ability to match pitches and distinguish such gross features as contour, progresses to the young child's ability to recognize tonal scales and discriminate key changes, and arrives finally at the adult's ability to detect small differences in interval size. Dowling proposes that the same hierarchy of melodic features is manifest in adult musical behavior.

One intriguing example of unusual musical ability is the possession of absolute pitch. Ward and Burns discuss this phenomenon in depth in Chapter 14 (*Absolute Pitch*). Why is it that a few people possess absolute pitch, whereas most people do not? Is it innate, or is it acquired through experience, perhaps at a critical age? Is it an ability that tends to

disappear through a process of unlearning? A detailed study of the processing capacities of one listener with absolute pitch is also presented.

Chapter 15 (*Neurological Aspects of Music Perception and Performance*) was written by Oscar Marin with two aims: first, to characterize the types of deficit in music perception and performance that are associated with damage to the nervous system and second, to use this information to elucidate the neuroanatomical basis of musical function in the normal case. Marin emphasizes that music is a highly complex function, so that simplistic views of how it is represented are bound to prove incorrect. This chapter also includes a discussion of the relationship between the neuroanatomical substrates of music and those of speech.

Instrumentalists will find Sloboda's *Music Performance* (Chapter 16) of particular interest. Sloboda emphasizes the importance of abstract structural representation to understanding how music is performed. Deviations from accuracy in expert performers can often be explained as serving to communicate musical structure to the listener. Experiments involving sightreading, memorization, and improvization are discussed with this theoretical framework.

Konečni's contribution on *Social Interaction and Musical Preference* (Chapter 17) stresses the important role played by music in everyday life. As the author writes, "a thorough understanding of aesthetic behavior cannot be achieved without examining how it changes as a function of its immediate social and nonsocial antecedents, concurrent cognitive and resultant emotional states." Furthermore, "both subtle and major changes in social behavior may occur as a function of listening to music." A set of laboratory experiments documents these arguments.

In the final chapter of the volume, *New Music and Psychology*, the composer and music theorist Robert Erickson presents a thoughtful evaluation of the relationship between the two disciplines. Erickson also outlines several questions raised by contemporary music that can usefully be approached using the methods of experimental psychology.

I am grateful to Michael Jordan for compiling the index, and I should like particularly to thank E. C. Carterette for his invaluable contribution in organizing the volume.

Diana Deutsch



# 1

# The Perception of Musical Tones

*R. A. Rasch and R. Plomp*


I. The Psychoacoustics of Music ..... 1
A. Introduction ..... 1
B. Methodology ..... 2
C. The Ear as a Frequency Analyzer ..... 4
II. Perceptual Attributes of Single Tones ..... 6
A. Pitch ..... 6
B. Loudness ..... 10
C. Timbre ..... 12
III. Perceptual Attributes of Simultaneous Tones ..... 14
A. Beats and Roughness ..... 14
B. Combination Tones ..... 17
C. Consonance and Dissonance ..... 19
IV. Conclusion ..... 21
References ..... 21


## I. THE PSYCHOACOUSTICS OF MUSIC

### A. Introduction

The aim of research in music perception is to explain how we respond subjectively to musical sound signals. In this respect it is a part of psychophysics, the general denomination for scientific fields concerned with the relationship between the objective, physical properties of sensory stimuli in our environment and the subjective, psychological responses evoked by them. If the stimuli are of an acoustic nature, we speak of *psychoacoustics*. Psychoacoustics can be of a general, theoretical nature; it can


Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

also be applied to a certain class of auditory stimuli, such as music and speech. This chapter is devoted to *musical psychoacoustics*.

The most important topics of musical psychoacoustics are the subjective properties of musical tones (pitch, loudness, timbre) and the phenomena that occur when several tones are presented simultaneously, which is what usually happens in music (beats and roughness, combination tones, consonance and dissonance). We will focus our discussion on these topics. However, before we deal more extensively with them, some attention must be given to the methodology of psychoacoustics and to the frequency-analyzing power of the ear, a capacity that is fundamental to its perceptual functioning.

## B. Methodology

Psychoacoustics is an empirical or, rather, experimental science. Observations from daily life and informal tryouts may be starting points for psychoacoustical knowledge, but the core of the scientific content is the result of laboratory investigations. In this respect it is an interdisciplinary field of research. Contributions have been made both by experimental psychologists and by physicists and acousticians.

A *psychoacoustical experiment* can be described most simply in a stimulus–response scheme. The *stimulus* is the sound presented to the subject. The experimenter requires the subject to give a *response*. The experimenter tries to discover the relationship between stimulus and response characteristics. Both stimulus and response are observable events. The subject is considered a "black box" that cannot be entered by the experimenter. Psychoacoustical research is often carried out without an attempt to explain the experimental results functionally in terms of sensory processes. Such attempts are made in research that is labeled *physiological acoustics*, a part of sensory and neurophysiology.

Our ears are very sensitive organs. Because of this, very accurate control of the stimulus variables is required in psychoacoustical experiments. Sound pressure level differences of less than 1 dB, time differences of a few msec, and frequency differences of less than 1 Hz can have a profound effect on the subjective response to a stimulus. It is impossible to obtain well-controlled psychoacoustic stimuli by manual means, like playing tones or chords on a musical instrument. The precision of the ear in distinguishing fine nuances is much greater than our ability to produce these nuances. As a rule, psychoacoustics makes use of electronic audio equipment that can produce sound stimuli according to any specification. In recent years it has become feasible to run the experiments under computer control. The computer can also be used for storage and analysis of stimuli and response data. Most problems concerning the production of the stimuli in psychoacoustical experiments may be considered solved. After the sound stimulus has been produced, it must reach the subject's eardrum with the least possible distortion. Usually high-quality headphones are used unless the spatial effect of the listening environment is involved. Background noises should be reduced, if not eliminated.

It is possible to have the subject describe his perception verbally. However, this response is often insufficient because our sensations allow much finer distinctions than our vocabulary does. Moreover, the use of words may differ from subject to subject. Because of this, in psychoacoustics most results are derived from responses made on the basis of a certain perception without direct reference to the perception itself. For example, if we have to indicate in which of two time intervals a sound has occurred, the response is a time indication based on an auditory sensation. A great deal of inventiveness is often required of the experimenter in designing his experimental paradigms.

The procedures used most often in psychoacoustical experiments are choice methods and adjustment methods. A single presentation of a sound event (one or more stimuli) to which a response must be made is called a *trial*. Using *choice methods*, the subject has to make, for each trial, a choice from a limited set of well-defined alternatives. The simplest case is the one with two alternatives, the *two-alternative-forced-choice* (2AFC). The insertion of the word "forced" is essential: The subject is obliged to choose. He must guess when he is incapable of making a meaningful choice.

For example, let us assume that the investigator is studying under what conditions a probe tone can be heard simultaneously with another, or masking sound. Each trial contains two successive time periods marked by visual signals. The masking sound is continuously present; the probe tone occurs in one of two time periods, randomly determined. If the probe tone is clearly detectable, the subject indicates whether it was presented in the first or in the second period. If the tone is not perceived at all, the subject must guess, resulting in an expectation of 50% correct responses. The transition from clearly detectable to not detectable tones is gradual. It is reflected by a gradual slope of the so-called *psychometric curve* that represents the percentage of correct responses plotted as a function of the sound pressure level of the target tone. The sound pressure level that corresponds to a score of 75% correct responses is usually adopted as the threshold for detection.

In order to arrive at an accurate estimate of the threshold, the experimenter varies sound pressure level of the tone for the successive trials. In the *constant stimuli* method the experimenter presents the tones according to a fixed procedure. The method of constant stimuli is time consuming because a number of trials are definitely supra- or infra-threshold and, therefore, do not give much information. Another class of choice methods, called *adaptive methods*, makes a more efficient use of trials. The experimental series is started with a certain initial value of the stimulus variable. One or more correct responses, depending upon the experimental strategy adopted, result in a change in the stimulus variable that makes it harder for the subject to make a correct choice. If the subject makes one or more false responses, the experimental task is facilitated. In this way, the value of the stimulus variable fluctuates around a certain value, which can be defined to be the threshold for perception.

Besides choice methods there is the *adjustment method*. The subject controls the stimulus variable himself, and he uses this control to find an optimal value. This method is not always feasible. The adjustment method is suitable for stimulus var-

iables that allow an optimal quality in perception: the best pitch for a tone in a musical interval, the most comfortable loudness, the greatest similarity or dissimilarity, etc. The optimal adjustment behaves like a stable equilibrium between lower and higher, both suboptimal, adjustments. Adjustment methods have the advantage that the results can be derived directly from the adjusted value, and do not have to be derived indirectly from the psychometric curve.

## C. The Ear as a Frequency Analyzer

Only by the ear's capacity to analyze complex sounds are we able to discriminate simultaneous tones in music. *Frequency analysis* may be considered the most characteristic property of the peripheral ear. The cochlea is divided over its entire length into two parts by the basilar membrane. In 1942 Von Békésy was the first to observe, with ingenious experimentation, that at every point along its length this membrane vibrates with maximum amplitude for a specific frequency. This finding confirmed the hypothesis, launched 80 years earlier by Helmholtz, that the cochlea performs a frequency analysis. Sound components with high frequencies are represented close to the base; components with low frequencies are represented near the apex of the cochlea. The frequency scale of the sound is converted into a spatial scale along the basilar membrane.

This capacity of the ear means that any periodic sound wave or *complex tone* is resolved into its frequency components, also called *partials* or *harmonics* (see Fig. 1). In mathematics the analogous procedure of determining the sinusoidal components of a periodic function is called *Fourier analysis*. In contrast with the theoretically perfect

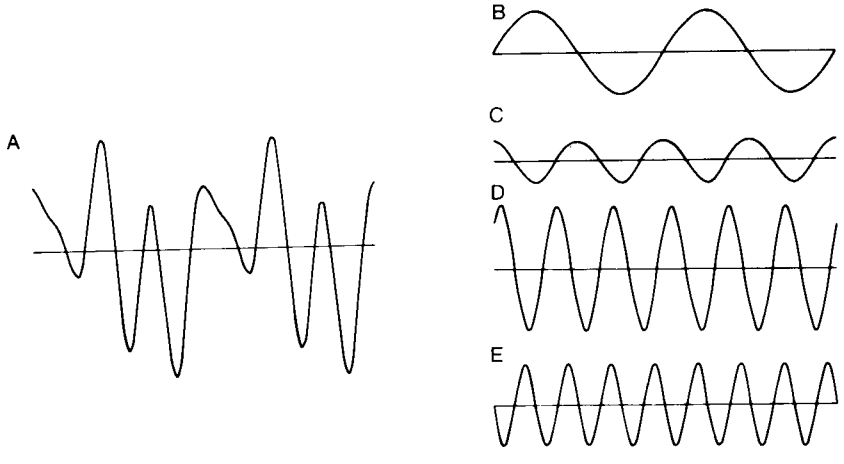



**Fig. 1.** The wave form A, a complex tone, is in fact the sum of the simple tones B-E. This is an illustration of Fourier's theorem that every vibration of frequency $f$ can be analyzed mathematically into a series of sinusoidal vibrations with frequencies $f$, $2f$, $3f$, etc. These sinusoidal vibrations are called the harmonics.

Fourier analysis, the frequency-analyzing power of the ear is limited: Only the lower harmonics can be analyzed individually.

There are many ways of studying the extent to which the ear can separate simultaneous tones. Only two approaches will be considered here. The first method investigates how many harmonics (with frequencies $nf$, $n = 1, 2, 3, 4$, etc.) can be distinguished in a complex tone. This can be done by using the 2AFC procedure: The listener has to decide which of two simple (sinusoidal) tones—one with frequency $nf$, the other with frequency $(n \pm \frac{1}{2})f$—is also present in the complex tone. The percentage of correct responses varies from 100 for low values of $n$ to about 50 for high values of $n$. Experiments along these lines have shown (Plomp, 1964) that, on the average, listeners are able to distinguish the first five to seven harmonics.

A quite different approach involves measuring the minimum sound pressure level necessary for a probe tone to be audible when presented with a complex tone. This is the so-called masked threshold; by varying the probe-tone frequency, we obtain the "masking pattern" of the complex tone. In Fig. 2 such a pattern is reproduced. The masking pattern of a complex tone of 500 Hz reveals individual peaks corresponding to the first five harmonics, nicely demonstrating the limited frequency-analyzing power of the ear.

The usual measure indicating how well a system is able to analyze complex signals is its bandwidth. The finding that the fifth harmonic can be distinguished from the fourth and the sixth means that the mutual distance should be a minor third or more. This distance constitutes a rough, general estimate of the bandwidth of the hearing mechanism, known in the psychophysical literature as the *critical bandwidth* (Fig. 3). A detailed review (Plomp, 1976) revealed that the bandwidth found experimentally is dependent on the experimental conditions. The values may differ by a factor of two.

In the lower frequency region (below 500 Hz) critical bandwidth is more or less constant if expressed in Hz. That means that musical intervals (frequency ratios) larger than the critical bandwidth at high frequencies may fall within the critical bandwidth at lower frequencies.

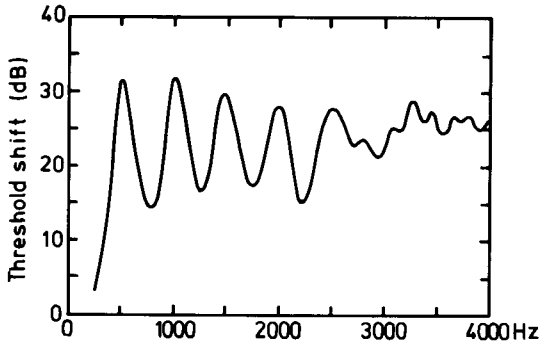






**Fig. 2.** Masking pattern of a complex tone consisting of the first 12 harmonics of 500 Hz (based on Plomp, 1964). Only the first five harmonics are analyzed by the ear.

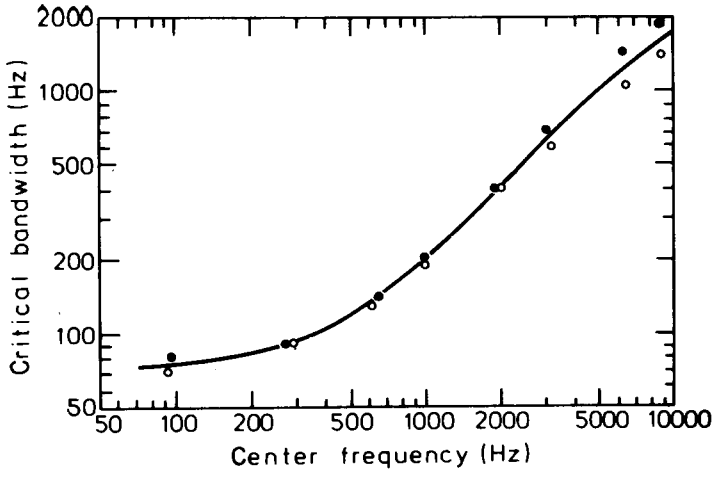



**Fig. 3.** Critical bandwidth as a function of frequency. This curve is a rough approximation only.

## II. PERCEPTUAL ATTRIBUTES OF SINGLE TONES

### A. Pitch

*Pitch* is the most characteristic property of tones, both simple (sinusoidal) and complex. Pitch systems (like the diatonic-chromatic and the 12-tone system) are among the most elaborate and intricate ever developed in Western and non-Western music. Pitch is related to the frequency of a simple tone and to the fundamental frequency of a complex tone. The frequency of a tone is a property that can usually be controlled in production and is well preserved during its propagation to the listener's ears.

For our purposes, pitch may be characterized as a *one-dimensional* attribute, i.e., all tones can be ordered along a single scale with respect to pitch (but see Chapters 8, 9, and 11, this volume). The extremes of this scale are *low* (tones with low frequencies) and *high* (tones with high frequencies). Sometimes tones with different spectral compositions (timbres) are not easily comparable as to pitch. It is possible that the clearness of pitch varies, for example, as a result of important noise components or inharmonic partials, or that the subjective character of the pitch varies, for example, when comparing the pitch of simple and complex tones. There are a number of subjective pitch scales:

1. The *mel* scale (see Stevens, Volkmann, & Newman, 1937). A simple tone of 1000 Hz has a defined pitch of 1000 mel. The pitch in mels of other tones with another frequency must be determined by comparative scaling experiments. A sound with a pitch subjectively twice that of a 1000 Hz tone is 2000 mel; "half pitch" is 500

mel, etc. Since there is no unambiguous subjective meaning of "a pitch half as high" or "double as high," the mel scale is a rather unreliable scale. It is not used very often.

2. The *musical pitch* scale (i.e., the ordinary indications C1, D1, . . . , C4, . . . , A4, etc.). These indications are only usable in musical situations.

3. The physical *frequency* scale in Hz. In psychoacoustical literature the pitch of a tone is often indicated by its frequency or, in the case of complex tones, by its fundamental frequency. Since the correspondence between frequency and pitch is monotonic, frequency is a rough indication of our pitch sensation. It must be realized however, that our perception operates more or less on the basis of a logarithmic frequency scale.

Pitch in its musical sense has a range of about 20 to 5000 Hz, roughly the range of the fundamental frequencies of piano strings and organ pipes. Tones with higher frequencies are audible but without definite pitch sensation. Low tones in the range of 10 to 50 Hz can have the character of a rattling sound. The transition from the perception of single pulses to a real pitch sensation is gradual. Pitch can be perceived after very few periods of the sound wave have been presented to the ear.

Simple tones have unambiguous pitches that can be indicated by means of their frequencies. These frequencies may serve as reference frequencies for the pitches of complex tones. The pitch sensation of complex tones is much more difficult to understand than the pitch of simple tones. As was discussed, the first five to seven harmonics of a complex tone can be distinguished individually if the listener's attention is drawn to their possible presence. However, a complex tone, as heard in practice, is characterized by a single pitch, the pitch of the fundamental component. This pitch will be referred to as *low pitch* here. In psychoacoustical literature this pitch is also known under a variety of other terms, such as periodicity pitch, repetition pitch, residue pitch, and virtual pitch. Experiments (Terhardt, 1971) have shown that the pitch of a complex tone with fundamental frequency $f$ is somewhat lower than that of a sinusoidal tone with frequency $f$. The existence of low pitch of a complex tone raises two questions. First, why are all components of the complex tones perceived as a perceptual unit; that is, why do all partials fuse into one percept? Second, why is the pitch of this perceptual tone the pitch of the fundamental component?

The first question can be answered with reference to the Gestalt theory of perception. The "Gestalt explanation" may be formulated as follows. The various components of a complex tone are always present simultaneously. We become familiar with the complex tones of speech signals (both of our own speech and of other speakers) from an early age. It would not be efficient to perceive them all separately. All components point to a single source and meaning so that perception of them as a unit gives a simpler view of the environment than separate perception. This mode of perception must be seen as a perceptual learning process. Gestalt psychology has formulated a number of laws that describe the perception of complex sensory stimuli. The perception of low pitch of complex tones can be classed under the heading of the "law of common fate." The harmonics of a complex tone exhibit "common fate."

The second question can also be answered with the help of a learning process

directed toward perceptual efficiency. The periodicity of a complex tone is the most constant feature in its composition. The amplitudes of the partials are subjected to much variation, caused by selective reflection, absorption, passing of objects, etc. Masking can also obscure certain partials. The periodicity, however, is a very stable and constant factor in a complex tone. This is reflected in the wave form built up from harmonics. The periodicity of a complex tone is at the same time the periodicity of the fundamental component of the tone. The perception of complex tones can be seen as a pattern recognition process. The presence of a complete series of harmonics is not a necessary condition for the pitch recognition process to succeed. It is sufficient that at least a few pairs of adjacent harmonics are present so that the periodicity can be determined. It is conceivable that there is a perceptual learning process that makes possible the recognition of fundamental periodicity from a limited number of harmonic partials. This learning process is based on the same experiences as those that led to singular pitch perception. Pattern recognition theories of the perception of low pitch are of relatively recent origin. Several times they have been worked out in detailed mathematical models that simulate the perception of complex tones (Goldstein, 1973; Wightman, 1973; Terhardt, 1974a; see also de Boer, 1976, 1977; Patterson & Wightman, 1976; Gerson & Goldstein, 1978; Houtsma, 1979; Piszczalski & Galler, 1979). It will probably take some time before the questions about the low singular pitch of complex tones are completely solved.

The classical literature on tone perception abounds with theories based on von Helmholtz's (1863) idea that the low pitch of a complex tone is based on the relative strength of the fundamental component. The higher harmonics are thought only to influence the timbre of the tones but not to be strong enough to affect pitch. However, low pitch perception also occurs when the fundamental component is not present in the sound stimulus. This was already observed by Seebeck (1841) and brought to the attention of the modern psychoacousticians by Schouten (1938). These observations led Schouten to the formulation of a *periodicity pitch theory*. In this theory pitch is derived from the waveform periodicity of the unresolved higher harmonics of the stimulus, the *residue*. This periodicity does not change if a component (e.g., the fundamental one) is removed. With this theory the observations of Seebeck and Schouten concerning tones without fundamental components could be explained. An attempt has also been made to explain the low pitch of a tone without fundamental ("the missing fundamental") as the result of the occurrence of combination tones, which provide a fundamental component in the inner ear. However, when these combination tones are effectively masked by low-pass noise, the sensation of low pitch remains (Licklider, 1954).

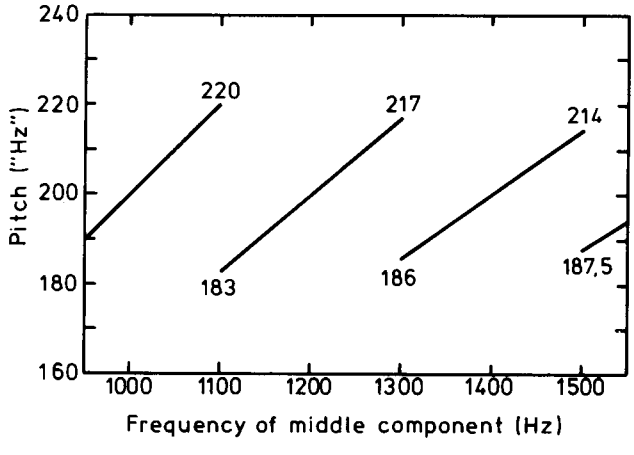
In musical practice complex tones with weak or absent fundamentals are very common. Moreover, musical tones are often partially masked by other tones. These tones can, however, possess very clear low pitches. Effective musical sound stimuli are often incomplete when compared to the sound produced by the source (instrument, voice).

Experiments in tone perception have pointed to a *dominance region* for pitch perception, roughly from 500 to 2000 Hz (Plomp, 1967; Ritsma, 1967). Partials falling in the

dominance region are most influential with regard to pitch. One way of showing this is to work with tones with inharmonic partials. Assume a tone with partials of 204, 408, 612, 800, 1000, and 1200 Hz. The first three partials in isolation would give a pitch of "204 Hz." All six together give a pitch of "200 Hz" because of the relative weight of the higher partials, which lie in the dominance region. The low pitch of complex tones with low fundamental frequencies (under 500 Hz) depends on the higher partials. The low pitch of tones with high fundamental frequencies is determined by the fundamental because it lies in the dominance region.

Tones with inharmonic components have been used quite frequently in tone perception research. An approximation of the pitch evoked by them is the fundamental of the least-deviating harmonic series. Assume a tone with components of 850, 1050, 1250, 1450, 1650 Hz. The least-deviating harmonic series is 833, 1042, 1250, 1458, and 1667 Hz, which contains the fourth, fifth, sixth, seventh, and eighth harmonics of a complex tone with a fundamental of 208.3 Hz. This fundamental can be used as an approximation of the pitch sensation of the inharmonic complex (Fig. 4). Let us consider an inharmonic tone with frequency components of 900, 1100, 1300, 1500, 1700 Hz. This tone has an ambiguous pitch, since two approximations by harmonic series are possible, namely one with a fundamental of 216.6 Hz (the component of 1300 Hz being the sixth harmonic in this case) and one with a fundamental of 185.9 Hz (1300 Hz being the seventh harmonic).

If not all partials of a complex tone are necessary for low pitch perception, how few of them are sufficient? The following series of experimental investigations show a progressively decreasing number (see Fig. 5). De Boer (1956) worked with five harmonics in the dominant region; Schouten, Ritsma, and Cardozo (1962), with three;




Biblioteca del Real Conservatorio - MADRID

**Fig. 4.** Schematic diagram of the low pitch of a complex tone consisting of five (inharmonic) components 200-Hz apart. The horizontal axis represents the frequency of the middle component. This component is taken as the fifth, sixth, seventh, or eighth pseudo-harmonic partial of a complex tone with low pitch, which is indicated along the vertical axis. The figure in the graph indicate the "pitches" of the stimuli with ambiguous pitch.

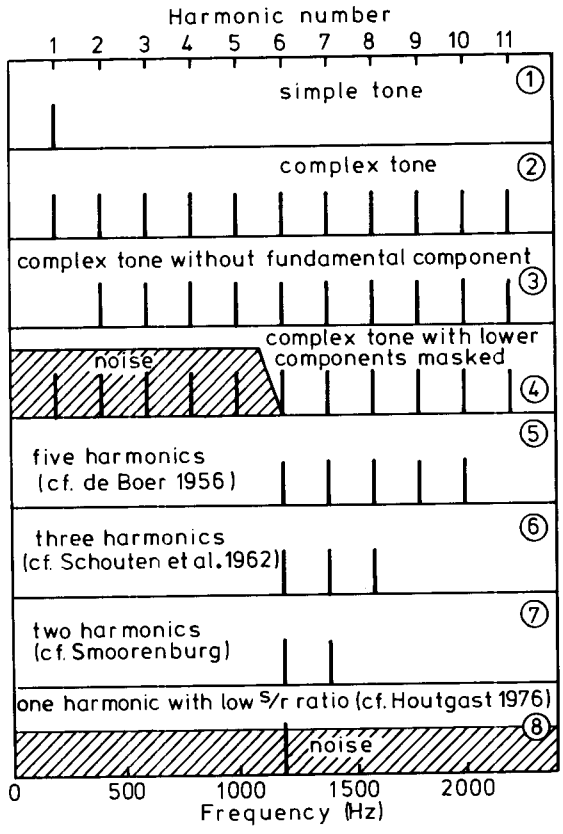



**Fig. 5.** Schematic diagram representing eight signals with the same low pitch.

Smoorenburg (1970), with two; Houtsma and Goldstein (1972), with one plus one—that is, one partial presented to each ear. In the latter case it is also possible to elicit low pitch perception. The authors concluded that low pitch was a central neural process not brought about by the peripheral sense organ (the ears). The last step in the series should be a low pitch perception evoked by one partial. That this is also possible has been shown by Houtgast (1976). The following conditions have to be fulfilled: The frequency region of the low pitch has to be filled with noise, the single partial must have a low signal-to-noise ratio, and attention has to be directed to the fundamental frequency region by prior stimuli. These conditions create a perceptual situation in which it is not certain that the fundamental is not there so that we are brought to the idea that it should be there by inference from earlier stimuli.

## B. Loudness

The physical correlate that underlies the *loudness* of a tone is intensity, usually expressed as sound pressure level (*SPL*) in dB. Sound pressure level is a relative

measure, expressed either relative to a zero level defined in the experimental situation or relative to a general reference sound pressure of $2 \times 10^{-5}$ N/m$^2$. Sound pressure levels of performed music vary roughly from 40 dB for a pianissimo to about 90 dB for a full orchestral forte-tutti (Winckel, 1962). By means of electronic amplification higher levels are reached in pop concerts. These levels, sometimes beyond 100 dB, are potentially damaging to the ear in case of prolonged presentation (Flugrath, 1969; Rintelman, Lindberg, & Smitley, 1972; Wood & Lipscomb, 1972; Fearn, 1975a,b).

The subjective assessment of loudness is more complicated than the physical measurement of the sound pressure level. Several *loudness scales* have been proposed during the last decades. None of them, however, can be applied fully satisfactorily in all conditions. We give the following summary review:

1. The *sone* scale, a purely psychophysical loudness scale (Stevens, 1936). The loudness of a simple (sinusoidal) tone of 1000 Hz with a sound pressure level of 40 dB is defined to be 1 sone; a tone with double loudness is assigned the loudness of 2 sones, etc. In general, a sound of $X$ sones is $n$ times louder than a sound of $X/n$ sones. The experimental determination of the relationship between the physical sound level and the psychophysical loudness is not very reliable because of the uncertainty of what is actually meant by "$X$ times louder."

2. The *phone* scale, a mixed physical–psychophysical loudness scale with scale values expressed in dB and, therefore, termed *loudness level* (*LL*). The loudness level of a sound in phones is equal to the sound pressure level of a 1000 Hz tone with the same loudness. For tones of 1000 Hz the identity relation $SPL = LL$ holds. The loudness level of simple tones with other tones with other frequencies and of complex tones or other sounds (noises, etc.) is found by comparison experiments, which can be done with acceptable reliability. These comparisons may be used to draw contours of equal loudness as a function of, for example, frequency.

3. The *sensation-level* scale, also a mixed scale. Sensation level is defined as the sound pressure level relative to threshold level and, as such, is also expressed in dB. It may differ as a function of frequency or other characteristics of a sound but also from subject to subject.

4. In many papers on psychoacoustics no loudness indications are given. Instead, physical levels are mentioned. For the investigator this is the most precise reference and at the same time a rough indication of subjective loudness.

In the description of the relation between sound pressure level and loudness, a clear distinction must be made between sounds with all spectral energy within one critical band and sounds with spectral energy spread over more than one critical band. If all sound energy is limited to one critical band, the loudness $L$ in sones increases monotonically with intensity $I$. The relation is often approached by the equation

$$L = kI^n$$

in which $k$ and $n$ are empirically chosen constants. A consequence of this relation is the rule that equal intensity ratios result in equal loudness ratios. Now, an intensity ratio is a fixed level difference (dB) so that the rule can also be formulated as follows: A certain loudness ratio corresponds to a certain level difference. Psychophysicists

have been much interested in the level difference that results in doubling or halving loudness, and many experiments have been carried out to establish this. The outcomes of these experiments are disappointingly dissimilar. Stevens (1955) summarized all experiments known to him with the median value of 10 dB for doubling loudness, later (1972) modified to 9 dB. These values correspond to values of $n = 0.3$ and $n = 0.33$ for the exponent in the formula. It is also possible to interpret the subjective loudness judgment as an imaginary judgment of the distance to the sound source. In this theory (Warren, 1977) half loudness must correspond to double distance, which gives, in free field conditions, a decrease of 6 dB sound pressure level. Warren conducted experiments in which this value is indeed found.

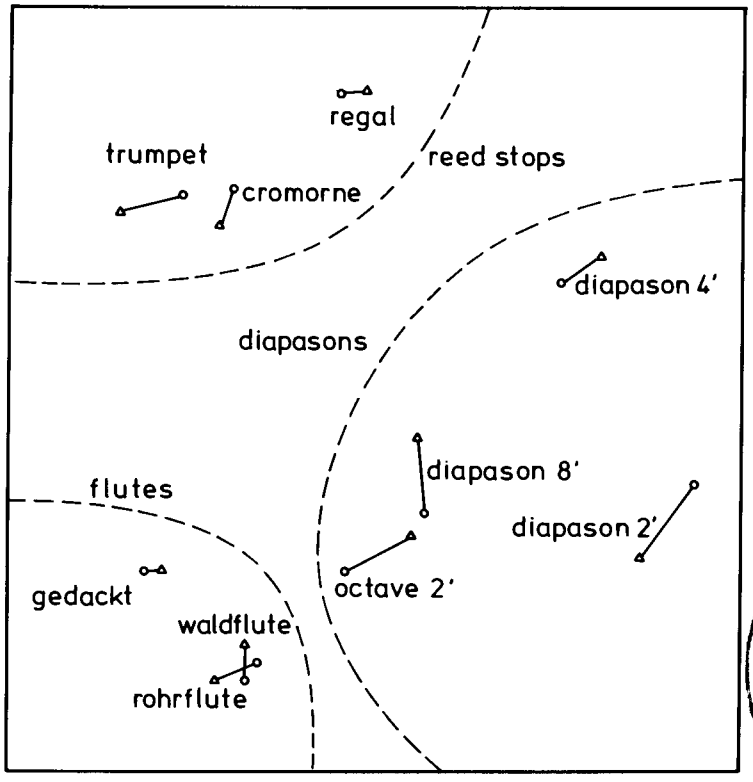
The assessment of loudness is a complicated matter if sound energy is present in more than one critical band. This situation is the common one for musical tones, especially for chords, and music played by ensembles, choirs, and orchestras. Total loudness is greater than when the same amount of sound energy is concentrated within one critical band. A number of models have been proposed that intend to be simulations of the perceptual processes involved and the parameters of which have been assigned values in accordance with psychophysical experiments. Well known are the models by Stevens (1955), Zwicker, Flottorp, & Stevens (1957), Zwicker and Scharf (1965), and Stevens (1972). These models have also been applied to musical sounds, especially to organ tones (Churcher, 1962; Pollard, 1978a,b).

Although loudness variations play an important role in music, they are less important than pitch variations. The number of assignable loudness degrees in music is limited to about five, coded musically from soft to loud as pianissimo, piano, mezzoforte, forte, and fortissimo. The definition of these loudness degrees is rather imprecise (Clark & Milner, 1964; Clark & Luce, 1965; Patterson, 1974). Judgment of musical loudness cannot have the degree of reliability and preciseness that is possible with the judgment of (relative) pitch, duration, tempo, etc. This is a consequence of the fact that the underlying physical dimension, intensity, is hard to control precisely. Sources of variation are encountered in sound production, in the fixed acoustic conditions of a room (absorption and thus attenuation by walls, floor, ceiling, etc.), in variable acoustic conditions (like the presence or the absence of an audience, the relative positions of sound source and listener, disturbing external noises), and in the audiograms of the listeners. In all the stages on the road from sound production to sound perception, sound pressure level is liable to be altered whereas frequency is not.

### C. Timbre

Timbre is, after pitch and loudness, the third attribute of the subjective experience of musical tones. Subjectively, timbre is often coded as the function of the sound source or of the meaning of the sound. We talk about the timbre of certain musical instruments, of vowels, and of sounds that signify certain events in our environment (apparatus, sounds from nature, footsteps, the slapping of a door, etc.).

What are the physical parameters that contribute to the perception of a certain timbre? In a restricted sense timbre may be considered the subjective counterpart of the spectral composition of tones. Especially important is the relative amplitude of the harmonics. This view was first stated by Helmholtz over a century ago and is reflected by the definition of timbre according to the American Standards Association (Acoust. Terminology S1.1., 1960): "Timbre is that attribute of auditory sensation in terms of which a listener can judge that two steady-state complex tones having the same loudness and pitch are dissimilar." Recent research has shown that temporal characteristics of the tones may have a profound influence on timbre as well, which has led to a broadening of the concept of timbre (Schouten, 1968). Both onset effects (rise time, presence of noise or inharmonic partials during onset, unequal rise of partials, characteristic shape of rise curve, etc.) and steady state effects (vibrato, amplitude modulation, gradual swelling, pitch instability, etc.) are important factors




Biblioteca del Real Conservatorio Superior de Música - Madrid

**Fig. 6.** Result of matching a two-dimensional perceptual timbre representation (circular symbols) of organ stops with the corresponding physical (spectral) representation (triangular) of the same sounds. The timbre scaling is the result of mulidimensional scaling applied to triadic comparison data. The physical scaling is based on a factor analysis of the spectral composition of the tones. The vertical dimension can be labeled as few versus many strong higher harmonics (based on Plomp, 1979).

in the recognition and, therefore, in the timbre of tones. Experiments (Clark, Robertson, & Luce, 1964; Berger, 1964; Saldanha & Corso, 1964) have shown that the identification of instrumental sounds is impaired when temporally characteristic parts of tones (especially the onsets) are removed.

Sounds cannot be ordered on a single scale with respect to timbre. Timbre is a *multidimensional attribute* of the perception of sounds. Dimensional research is highly time-consuming and is therefore always done with a restricted set of sound stimuli. The dimensions found in such an investigation are of course determined by the stimulus set.

Dimensional research of timbre leads to the ordering of sound stimuli on the dimensions of a timbre space. An example of such research is that by Von Bismarck (1974a,b). His stimulus set contained a large number (35) of tone and noise stimuli. The most important factors found by him can be characterized as follows: (a) *sharpness*, determined by a distribution of spectral energy that has its gravity point in the higher frequency region and (b) *compactness*, a factor that distinguishes between tonal (compact) and noise (not compact) aspects of sound.

In some investigations sound stimuli have been submitted to multidimensional scaling, both perceptual and physical. The physical scaling can be based on the spectral composition of the sounds, as was done in Plomp's (1979) experiments with tones from a number of organ stops. Figure 6 gives the two-dimensional representation of 10 sounds, both perceptual and physical. The representations correspond rather well, leading to the conclusion that in this set of stimuli the sound spectrum is the most important factor in the perception of timbre.

Other examples of dimensional research on timbre are the investigations by Plomp (1970), Wedin and Goude (1972), Plomp and Steeneken (1973), Miller and Carterette (1975), Grey (1977), and de Bruijn (1978).

## III. PERCEPTUAL ATTRIBUTES OF SIMULTANEOUS TONES

### A. Beats and Roughness

In this and the following sections we will discuss perceptual phenomena that occur as the result of two *simultaneous tones*. We will call the simultaneously sounding tones the *primary tones*.

We consider first the case of two simultaneous simple tones. Several conditions can be distinguished, depending on frequency difference (Fig. 7). If the two primary tones have equal frequencies, they fuse into one tone, in which the intensity depends on the phase relation between the two primary tones. If the tones differ somewhat in frequency, the result is a signal with periodic amplitude and frequency variations with a frequency equal to the frequency difference. The frequency variations are only slight and will not be considered here. The *amplitude variations*, however, can be

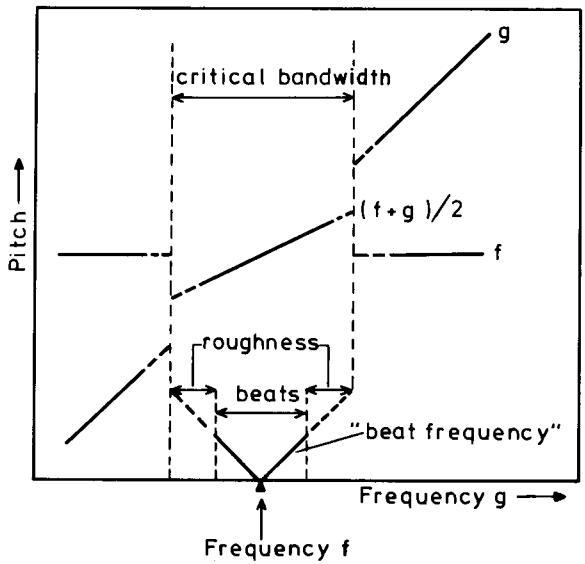



**Fig. 7.** Schematic diagram representing perceptual phenomena that may occur when two simple tones with a small frequency difference sound simultaneously. The frequency of one tone is set constant ($f$); the frequency of the other tone ($g$) varies along the horizontal axis (based on Roederer, 1975).

considerable and result in a fluctuating intensity and perceived loudness. These loudness fluctuations are called *beats*, if they can be discerned individually by the ear, which occurs if their frequency is less than about 20 Hz. A stimulus equal to the sum of two simple tones with equal amplitudes and frequencies $f$ and $g$

$$p(t) = \sin 2\pi ft + \sin 2\pi gt$$

can be described as

$$p(t) = 2 \cos 2\pi \tfrac{1}{2}(g - f)t \times \sin 2\pi \tfrac{1}{2}(f + g)t$$

This is a signal with a frequency that is the average of the original primary frequencies, and an amplitude that fluctuates slowly with a *beat frequency* of $g - f$ Hz (Fig. 8). Amplitude variation is less strong if the two primary tones have different amplitudes.

When the frequency difference is larger than about 20 Hz, the ear is no longer able to follow the rapid amplitude fluctuations individually. Instead of the sensation of fluctuating loudness, there is a rattle-like sensation called *roughness*. Beats and roughness can only occur if the two primary tones are not resolved by the ear (that means, not processed separately but combined). If the frequency difference is larger than the critical band, the tones are perceived individually with no interference phenomena.

In musical sounds beats can occur with just noncoinciding harmonics of mistuned consonant intervals of complex tones. If the fundamental frequencies of the tones of

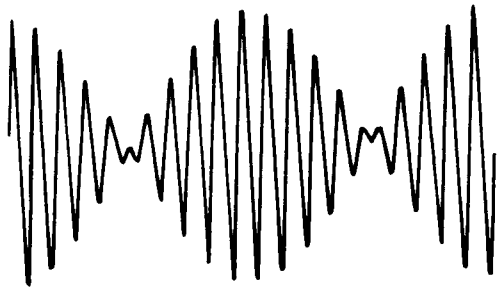


**Fig. 8.** Wave form arising from the superposition of two simple tones with a small frequency difference.

an octave (theoretically 1:2) or fifth (2:3) differ a little from the theoretical ratio, there will be harmonics that differ slightly in frequency and will cause beats. These beats play an important role when tuning musical instruments.

No psychophysical research has been done on mistuned intervals of complex tones, but to a certain extent psychophysical results found with two beating simple tones and with amplitude-modulated simple tones (see Fig. 9) can be applied to the perception of beating mistuned intervals of complex tones (Zwicker, 1952; Terhardt, 1968a,b, 1974b). The following relations can be stated. Thresholds vary with beat frequency. There appears to be a minimum at about 5 to 10 Hz. The threshold decreases when the sound pressure level increases. It is possible to define perceptual quantities called *beating strength* and *roughness strength* and to determine their values as a function of stimulus characteristics. Research following this line has shown that such a quantity increases with modulation depth and with sound pressure level. Moreover, there seems to be a modulation frequency giving maximal roughness (about 50 to 70 Hz).

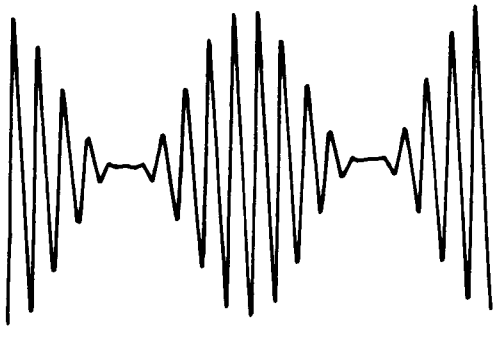


**Fig. 9.** Wave form that arises when a simple tone has been modulated in amplitude sinusoidally. This wave form is an example of complete modulation.

## B. Combination Tones

Two simple tones at a relatively high sound pressure level and with a frequency difference that is not too large can give rise to the perception of so-called combination tones. These combination tones arise in the ear as a product of nonlinear transmission characteristics. The combination tones are not present in the acoustic signal. However, they are perceived as if they were present. The ear cannot distinguish between perceived components that are "real" (in the stimulus) and those that are not (combination tones). The combination tones are simple tones that may be cancelled effectively by adding a real simple tone with the same frequency and amplitude but opposite phase. This cancellation tone can be used to investigate combination tones.

The possible frequencies of combination tones can be derived from a general transmission function. Assume a stimulus with two simple tones:

$$p(t) = \cos 2\pi ft + \cos 2\pi gt$$

$f$ and $g$ being the two frequencies. Linear transmission is described by

$$d = ap + c$$

($a$ and $c$ being constants). If transmission is not linear, higher order components are introduced:

$$d = a_1 p + a_2 p^2 + a_3 p^3 + \cdots$$

The quadratic term can be developed as follows:

$$\begin{aligned} p^2 &= (\cos 2\pi ft + \cos 2\pi gt)^2 \\ &= 1 + \tfrac{1}{2}\cos 2\pi 2ft + \tfrac{1}{2}\cos 2\pi 2gt \\ &\quad + \cos 2\pi(f+g)t + \cos 2\pi(f-g)t \end{aligned}$$

It can be seen that components with frequencies $2f$, $2g$, $f + g$, and $f - g$ are introduced in this way. Similarly, the cubic term can be developed:

$$\begin{aligned} p^3 &= (\cos 2\pi ft + \cos 2gt)^3 \\ &= \tfrac{3}{4}\cos 2\pi ft + \tfrac{3}{4}\cos 2\pi gt + \tfrac{1}{4}\cos 2\pi 3ft \\ &\quad + \tfrac{1}{4}\cos 2\pi 3gt + \tfrac{3}{4}\cos 2\pi(2f+g)t + \tfrac{3}{4}\cos 2\pi(2g+f)t \\ &\quad + \tfrac{3}{4}\cos 2\pi(2f-g)t + \tfrac{3}{4}\cos 2\pi(2g-f)t \end{aligned}$$

This term is responsible for components with frequencies $3f$, $3g$, $2f + g$, $2g + f$, $2f - g$, $2g - f$. The higher terms of the nonlinear transmission formula can be worked out analogously. The factors just preceding the cosine terms indicate the relative amplitudes of the components in their groups. Psychoacoustical research on combination tones has shown that the pitches of the combination tones agree with the frequencies predicted by nonlinear transmission (Plomp, 1965; Smoorenburg, 1972a,b; Hall, 1975; Weber & Mellert, 1975; Schroeder, 1975b; Zurek & Leskowitz, 1976). However, the correspondence between the relative amplitude predicted and the subjective loudness measured is far from perfect. Clearly, the phenomenon of combination tones

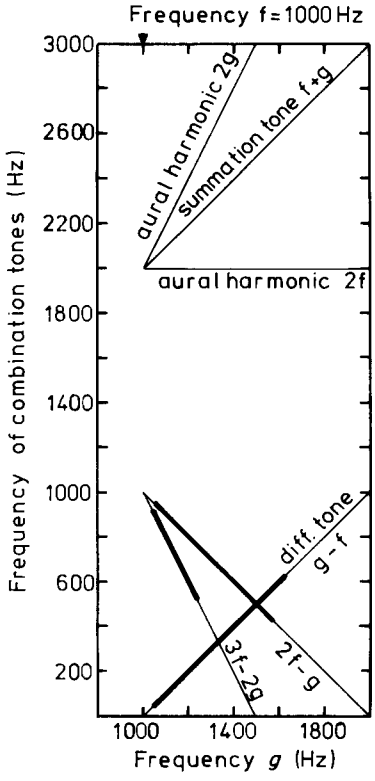



**Fig. 10.** Frequency ranges over which particular combination tones can occur. The frequency $f$ of the lower primary tone is 1000 Hz. The frequency of the higher primary tone ($g$) varies from 1000 Hz up to 2000 Hz. The thin lines correspond to the various theoretically possible combination-tone frequencies; the solid lines represent the combination tones found most often.

is more complicated than can be described in a simple formula. Moreover, there are individual differences, which should be expected since this is a distortion process. Experiments have shown (see Fig. 10) that the following combination tone frequencies are the most important: the so-called *difference tone* with frequency $g - f$ Hz, the *second-order difference tone* with frequency $2f - g$ Hz, and the *third-order difference tone* with frequency $3f - 2g$ Hz. The diagram illustrates that the combination tones are stronger for small frequency differences of the primary tones than for large differences; this indicates that the origin of combination tones is tightly connected with the frequency-analyzing process in the inner ear. It should be noted that the importance of *summation tones* (with frequency $f + g$) and the so-called *aural harmonics* (with frequencies $2f$, $3f$, etc., and $2g$, $3g$, etc.) is questionable. Although combination tones were discovered by musicians in musical contexts (Tartini and Sorge in the eighteenth century), their significance for music is not very high. They can be easily evoked by playing loud tones in the high register on two flutes or recorders or double stops on the violin. In a normal listening situation, however, their levels are usually too low to

attract attention. Moreover, they will be masked by the tones of other (lower) instruments. Some violin teachers (following Tartini) advise the use of combination tones as a tool for controlling the intonation of double-stop intervals. Because audible combination tones behave more as simple tones in lower frequency regions than the complex tones to be intonated, a pitch comparison of combination tones and played tones should not be given too much weight.

### C. Consonance and Dissonance

The simultaneous sounding of several tones may be pleasant or "euphonious" to varying degrees. The pleasant sound is called *consonant;* the unpleasant or rough one, *dissonant*. The terms consonance and dissonance have been used here in a perceptual or sensory sense. This aspect has been labeled *tonal consonance* (Plomp & Levelt, 1965) or *sensory consonance* (Terhardt, 1976), to be distinguished from consonance in a musical situation. Musical consonance has its roots in perceptual consonance, of course, but is dependent on the rules of music theory, which, to a certain extent, can operate independently from perception.

The perceptual consonance of an interval consisting of two simple tones depends directly upon the frequency difference between the tones, not upon the frequency ratio (or musical interval). If the frequency separation is very small or large (more than critical bandwidth—the tones not interfering with each other), the two tones together sound consonant. Dissonance occurs if the frequency separation is less than a critical bandwidth (see Fig. 11). The most dissonant interval arises with a frequency separation of about a quarter of the critical bandwidth: about 20 Hz in low-frequency regions, about 4% (a little less than a semitone) in the higher regions (Fig. 12). The frequency separation of the minor third (20%), major third (25%), fourth (33%), fifth (50%), and so on, is usually enough to give consonant combination of simple tones. However, if the frequencies are low, the frequency separation of thirds (and eventually also fifths) is less than critical bandwidth so that even these intervals cause a

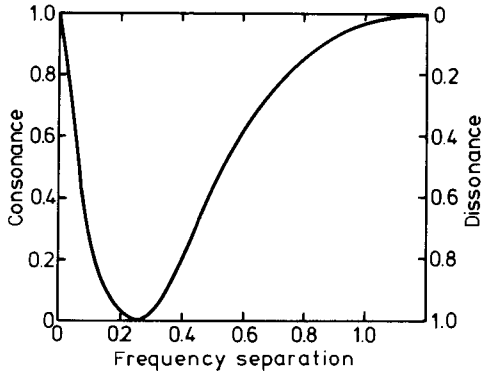





**Fig. 11.** Consonance of an interval consisting of two simple tones as a function of frequency separation, measured relative to critical bandwidth (based on Plomp & Levelt, 1965).

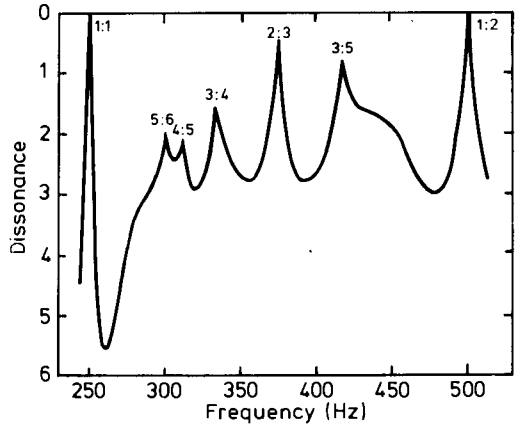



**Fig. 12.** Consonance of an interval consisting of two complex tones (with six harmonics). The lower tone ($f$) has a fundamental frequency of 250 Hz; the fundamental frequency of the higher tone is the variable along the horizontal axis. The consonance/dissonance values are predictions from the model of Plomp & Levelt (1965).

dissonant beating. For this reason, these consonant intervals are not used in the bass register in musical compositions.

The consonance of intervals of complex tones can be derived from the consonances of the simple-tone combinations comprised in them. In this case the dissonance is the additive element. The dissonance of all combinations of neighboring partials can be determined and added to give the total dissonance and, inversely, the total consonance of the sound. Sounds with widely spaced partials, such as clarinet tones (with only the odd harmonics) are more consonant than sounds with narrowly spaced partials. The composition of the plenum of an organ is such that the partials are widely spaced throughout the spectrum. Some mathematical models have been worked out that describe the dissonance of a pair of simple tones and the way in which the dissonances of partial pairs in tone complexes have to be added (Plomp & Levelt, 1965; Kameoka & Kuriyagawa, 1969a,b; Hutchinson, 1978). As far as can be decided, these models give a good picture of consonance perception.

The consonance of a musical interval, defined as the sum of two complex tones with a certain ratio in fundamental frequency, is highly dependent on the simplicity of the frequency ratio. Intervals with frequency ratios that can be expressed in small integer numbers (say, less than 6) are relatively consonant because the lower, most important components of the two tones are either widely apart or coincide. If the frequency ratio is less simple, there will be a number of partials from the two tones that differ only a little in frequency, and these partial pairs give rise to dissonance. It seems that intervals with the number 7 in their frequency proportions (7/4, 7/5, . . .) are about on the borderline between consonance and dissonance.

Experiments with inharmonic partials (Slaymaker, 1970; Pierce, 1966) have shown that consonance or dissonance is indeed dependent on the coincidence of partials and

not necessarily on the simple frequency ratio between the fundamental frequencies (which is usually the cause of the coincidence).

If the number of partials in a complex tone increases or if the strengths of the higher harmonics (with narrow spacing) increase, the tone is perceived as more dissonant (compare the trumpet with the flute, for instance). However, the $n$th partial is required in order to make an interval with frequency ratio $n:m$ or $m:n$ relatively consonant. For example, if the fifth harmonic is absent, the usual beating (dissonance) of a mistuned major third (4:5) will be absent (see also Fig. 12).

Musical consonance in Western polyphonic and harmonic music is clearly based on perceptual consonance of complex (harmonic) tones. Intervals with simple frequency ratios are consonant. Intervals with nonsimple frequency ratios are dissonant. The way in which consonance and dissonance are used in music theory and composition varies considerably from one historical period to another.

## IV. CONCLUSION

More than a century ago von Helmholtz published his classic volume *On the Sensations of Tone* (1863). The subtitle specifically indicates the intention of this study: "As a Physiological Basis for the Theory of Music." For Helmholtz the theory of music (as a compendium of rules that control composition and as such the musical sound stimulus) could only be understood fully if it could be shown that its elements had their origin in the perceptual characteristics of our hearing organ. Helmholtz's working hypothesis has been put aside by later investigators, both those who worked in music and those who worked in psychoacoustics. Several reasons for this can be given. First, before the introduction of electroacoustic means of tone production and control in the 1920s, it was not possible to carry out the necessary psychoacoustical experiments, while Helmholtz's observations proved to be insufficient in many ways. Second, it turned out that music theory has its own rules apart from the perceptual relevance of the characteristics of the sounds that it creates. Therefore, it is not clear, neither for the music theorist nor for the psychoacoustician, which aspects of music theory should be subjected to psychoacoustical research and which should not. Fortunately, in recent years much research has been initiated that is aimed at the investigation of the relationship between musical-theoretical and perceptual entities. For the time being, no complete view can be given, but there may come a time in which Helmholtz's ideas on the relation between the properties of our perceptual processes and the elements of musical composition can receive new, more complete and exact formulations than was possible a century ago.

## REFERENCES

Berger, K. W. Some factors in the recognition of timbre. *Journal of the Acoustical Society of America*, 1964, *36*, 1888–1891.

Bismarck, G. von. Timbre of steady sounds: A factorial investigation of its verbal attributes. *Acustica*, 1974, *30*, 146–159.

Bismarck, G. von. Sharpness as an attribute of the timbre of steady sounds. *Acustica*, 1974, *30*, 159–172.

Boer, E. de. On the 'residue' in hearing. Dissertation, Amsterdam, 1956.

Boer, E. de. On the 'residue' and auditory pitch perception. In W. D. Keidel & W. D. Neff (Eds.), *Handbook of sensory physiology*. (Volume V, Auditory system, Part 3, Clinical and special topics) Berlin: Springer-Verlag, 1976. Pp. 479–583.

Boer, E. de. Pitch theories unified. In E. F. Evans & J. P. Wilson (Eds.), *Psychophysics and physiology of hearing*. New York: Academic Press, 1977. Pp. 323–335.

Bruijn, A. de. Timbre-classification of complex tones. *Acustica*, 1978, *40*, 108–114.

Churcher, B. G. Calculation of loudness levels for musical sounds. *Journal of the Acoustical Society of America*, 1962, *34*, 1634–1642.

Clark, M., & Luce, D. Intensities of orchestral instrument scales played at prescribed dynamic markings. *Journal of the Audio Engineering Society*, 1965, *13*, 151–157.

Clark, M. Jr., & Milner, P. Dependence of timbre on the tonal loudness produces by musical instruments. *Journal of the Audio Engineering Society*, 1964, *12*, 28–31.

Clark, M. Jr., Robertson, P., & Luce, D. A preliminary experiment on the perceptual basis for musical instrument families. *Journal of the Audio Engineering Society*, 1964, *12*, 199–203.

Evans, E. F., & Wilson, J. P. (Eds.), *Psychophysics and physiology of hearing*. New York: Academic Press, 1977.

Fearn, R. W. Level limits on pop music. *Journal of Sound and Vibration*, 1975, *38*, 591–592. (a)

Fearn, R. W. Level measurements of music. *Journal of Sound and Vibration*, 1975, *43*, 588–591. (b)

Flugrath, J. M. Modern-day rock-and-roll music and damage-risk criteria. *Journal of the Acoustical Society of America*, 1969, *45*, 704–711.

Gerson, A., & Goldstein, J. L. Evidence for a general template in central optimal processing for pitch of complex tones. *Journal of the Acoustical Society of America*, 1978, *63*, 498–510.

Goldstein, J. L. An optimum processor theory for the central formation of the pitch of complex tones. *Journal of the Acoustical Society of America*, 1973, *54*, 1496–1516.

Green, D. M. *An introduction to hearing*. Hillsdale, New York: Lawrence Erlbaum, 1976.

Grey, J. M. Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America*, 1977, *61*, 1270–1277.

Hall, J. L. Nonmonotonic behavior of distortion product $2f_1$-$f_2$: Psychophysical observations. *Journal of the Acoustical Society of America*, 1975, *58*, 1046–1050.

Helmholtz, H. von. *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik (Sechste Ausg.)*. Braunschweig: Vieweg, 1913 (1st ed., 1863). Translated by A. J. Ellis as: *On the sensations of tone as a physiological basis for the theory of music*. London: Longmans, Green, 1885. (1st ed., 1875; reprint of the 1885 ed., New Y ork: Dover, 1954).

Houtgast, T. Subharmonic pitches of a pure tone at low S/N ratio. *Journal of the Acoustical Society of America*, 1976, *60*, 405–409.

Houtsma, A.J.M. Musical pitch of two tone complexes and predictions by modern pitch theories. *Journal of the Acoustical Society of America*, 1979, *66*, 87–99.

Houtsma, A.J.M., & Goldstein, J. L. The central origin of the pitch of complex tones: Evidence from musical interval recognition. *Journal of the Acoustical Society of America*, 1972, *51*, 520–529.

Hutchinson, W., & Knopoff, L. The acoustic component of Western consonance. *Interface*, 1978, 7, 1–29.

Kameoka, A., & Kuriyagawa, M. Consonance theory, Part I: Consonance of dyads. *Journal of the Acoustical Society of America*, 1969, *45*, 1451–1459. (a)

Kameoka, A., & Kuriyagawa, M. Consonance theory, Part II: Consonance of complex tones and its calculation method. *Journal of the Acoustical Society of America*, 1969, *45*, 1460–1469. (b)

Licklider, J.C.R. 'Periodicity' pitch and 'place' pitch. *Journal of the Acoustical Society of America*, 1954, *26*, 945.

Miller, J. R., & Carterette, E. C. Perceptual space for musical structures. *Journal of the Acoustical Society of America*, 1975, *58*, 711–720.

Patterson, B. Musical dynamics. *Scientific American*, 1974, *31*, 78–95.

Patterson, R. D., & Wrightman, F. L. Residue pitch as a function of component spacing. *Journal of the Acoustical Society of America*, 1976, *59*, 1450–1459.

Pierce, J. R. Attaining consonance in arbitrary scales. *Journal of the Acoustical Society of America*, 1966, *40*, 249.

Piszczalski, M., & Galler, B. A. Predicting musical pitch from component frequency ratios. *Journal of the Acoustical Society of America*, 1979, *66*, 710–720.

Plomp, R. The ear as a frequency analyzer. *Journal of the Acoustical Society of America*, 1964, *36*, 1628–1636.

Plomp, R. Detectability threshold for combination tones. *Journal of the Acoustical Society of America*, 1965, *37*, 1110–1123.

Plomp, R. Pitch of complex tones. *Journal of the Acoustical Society of America*, 1967, *41*, 1526–1533.

Plomp, R. Timbre as a multidimensional attribute of complex tones. In R. Plomp & G. F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing*. Leiden: Sijthoff, 1970. Pp. 397–414.

Plomp, R. Auditory psychophysics. *Annual Review of Psychology*, 1975, *26*, 207–232.

Plomp, R. *Aspects of tone sensation*. New York: Academic Press, 1976.

Plomp, R. Fysikaliska motsvarigheter till klanfärg hos stationära ljud. In *Vår hörsel och musiken*. Stockholm: Kungl. Musikaliska Akademien, 1979.

Plomp, R., & Levelt, W.J.M. Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, 1965, *38*, 548–560.

Plomp, R., & Smoorenburg, G. F. (Eds.), *Frequency analysis and periodicity detection in hearing*. Leiden: Sijthoff, 1970.

Plomp, R., & Steeneken, H.J.M. Place dependence of timbre in reverberant sound fields. *Acustica*, 1973, *28*, 49–59.

Pollard, H. F. Loudness of pipe organ sounds. I. Plenum combinations. *Acustica*, 1978, *41*, 65–74. (a)

Pollard, H. F. Loudness of pipe organ sounds. II. Single notes. *Acustica*, 1978, *41*, 75–85. (b)

Rintelmann, W. F., Lindberg, R. F., & Smitley, E. K. Temporary threshold shift and recovery patterns from two types of rock-and-roll music presentation. *Journal of the Acoustical Society of America*, 1972, *51*, 1249–1255.

Ritsma, R. J. Frequencies dominant in the perception of the pitch of complex sounds. *Journal of the Acoustical Society of America*, 1967, *42*, 191–198.

Roederer, J. G. *Introduction to the physics and psychophysics of music*. New York and Berlin: Springer, 1974 (2nd ed., 1975).

Saldanha, E. L., & Corso, J. F. Timbre cues and the identification of musical instruments. *Journal of the Acoustical Society of America*, 1964, *36*, 2021–2026.

Schouten, J. F. The perception of subjective tones. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen*, 1938, *41*, 1083–1093.

Schouten, J. R., Ritsma, R. J., & Cardozo, B. L. Pitch of the residue. *Journal of the Acoustical Society of America*, 1962, *34*, 1418–1424.

Schouten, J. F. The perception of timbre. In *Report of the Sixth International Congress on Acoustics, Tokyo*, Paper GP-6-2, 1968.

Schroeder, M. R. Models of hearing. *Proceedings of the IEEE*, 1975, *63*, 1332–1350. (a)

Schroeder, M. R. Amplitude behavior of the cubic difference tone. *Journal of the Acoustical Society of America*, 1975, *58*, 728–732. (b)

Schubert, E. D. (Ed.) *Psychological acoustics*. Stroudsburg, Pennsylvania: Dowden, 1979 (Benchmark Papers in Acoustics 13).

Seashore, C. E. *Psychology of music*. New York: McGraw-Hill, 1938 (Reprint New York: Dover, 1967).

Seebeck, A. Beobachtungen über einige Bedingungen der Entstehung von Tönen. *Annalen der Physik und Chemie*, 1841, *53*, 417–436.

Slaymaker, F. H. Chords from tones having stretched partials. *Journal of the Acoustical Society of America*, 1970, *47*, 1569–1571.

Smoorenburg, G. F. Pitch perception of two-frequency stimuli. *Journal of the Acoustical Society of America*, 1970, *48*, 924–942.

Smoorenburg, G. F. Audibility region of combination tones. *Journal of the Acoustical Society of America*, 1972, *52*, 603–614. (a)

Smoorenburg, G. F. Combination tones and their origin. *Journal of the Acoustical Society of America*, 1972, *52*, 615–632. (b)

Stevens, S. S. A scale for the measurement of a psychological magnitude: Loudness. *Psychological Review*, 1936, *43*, 405–416.

Stevens, S. S. The measurement of loudness. *Journal of the Acoustical Society of America*, 1955, 27, 815–829.

Stevens, S. S. Perceived level of noise by Mark VII and decibels (E). *Journal of the Acoustical Society of America*, 1972, *51*, 575–601.

Stevens, S. S., Volkmann, J., & Newman, E. B. A scale for the measurement of the psychological magnitude pitch. *Journal of the Acoustical Society of America*, 1937, *8*, 185–190.

Terhardt, E. Über die durch amplitudenmodulierte Sinustöne hervorgerufene Hörempfindung. *Acustica*, 1968, *20*, 210–214. (a)

Terhardt, E. Über akustische Rauhigkeit und Schwankungsstärke. *Acustica*, 1968, *20*, 215–224. (b)

Terhardt, E. Die Tonhöhe harmonischer Klänge und das Oktavintervall. *Acustica*, 1971, *24*, 126–136.

Terhardt, E. Pitch, consonance, and harmony. *Journal of the Acoustical Society of America*, 1974, *55*, 1061–1960. (a)

Terhardt, E. On the perception of periodic sound fluctuations (roughness). *Acustica*, 1974, *30*, 201–203. (b)

Terhardt, E. Ein psychoakustisch begründetes Konzept der musikalischen Konsonanz. *Acustica*, 1976, *36*, 121–137.

Terhardt, E. Psychoacoustic evaluation of musical sounds. *Perception & Psychophysics*, 1978, *23*, 483–492.

Tobias, J. V. (Ed.). *Foundations of modern auditory theory.* (Volumes 1 & 2) New York: Academic Press, 1970, 1972.

Warren, R. M. Subjective loudness and its physical correlate. *Acustica*, 1977, *37*, 334–346.

Weber, R., & Mellert, V. On the nonmonotonic behavior of cubic distortion products in the human ear. *Journal of the Acoustical Society of America*, 1975, *57*, 207–214.

Wedin, L., & Goude, G. Dimension analysis of the perception of instrumental timbre. *Scandinavian Journal of Psychology*, 1972, *13*, 228–240.

Wightman, F. L. The pattern-transformation model of pitch. *Journal of the Acoustical Society of America*, 1973, *54*, 407–416.

Winckel, F. W. Optimum acoustic criteria of concert halls for the performance of classical music. *Journal of Acoustical Society of America*, 1962, *34*, 81–86.

Wood, W. S. III & Lipscomb, D. M. Maximum available sound-pressure levels from stereo components. *Journal of the Acoustical Society of America*, 1972, *52*, 484–487.

Zurek, P. M., & Leshowitz, B. Measurements of the combination tones $f_2 - f_1$ and $2f_1 - f_2$. *Journal of the Acoustical Society of America*, 1976, *60*, 155–168.

Zwicker, E. Die Grenzen der Hörbarkeit der Amplitudenmodualtion und der Frequenzmodulation eines Tones. *Acustica*, 1952, *2*, Beihefte 125–135.

Zwicker, E., & Feldtkeller, R. *Das Ohr als Nachrichtenempfänger* (2nd Ausg.). Stuttgart: Hirzel, 1967.

Zwicker, E., Flottorp, C., & Stevens, S. S. Critical bandwidth in loudness summation. *Journal of the Acoustical Society of America*, 1957, *29*, 548–557.

Zwicker, E., & Scharf, B. A model of loudness summation. *Psychological Review*, 1965, 72, 3–26.

Zwicker, E., & Terhardt, E. (Eds.), *Facts and models in hearing.* Berlin: Springer-Verlag, 1974.

# 2

# Exploration of Timbre by Analysis and Synthesis

*Jean-Claude Risset and David L. Wessel*


| | | |
|---|---|---|
| I. | Timbre | 26 |
| II. | Timbre and the Fourier Spectrum: The Classical View | 26 |
| III. | The Shortcomings of the Classical Conception | 29 |
| IV. | Attack Transients | 30 |
| V. | Complexity of Sounds: Importance of Characteristic Features | 30 |
| VI. | Instrumental and Vocal Timbres: Additive Syntheses | 30 |
| VII. | Cross Synthesis and Voice Synthesis | 36 |
| VIII. | Additive Synthesis: Percussion Instruments | 37 |
| IX. | Substractive Synthesis | 39 |
| X. | Acoustic Modeling as a Synthesis Technique | 41 |
| XI. | The Importance of Context | 42 |
| XII. | Analysis-Synthesis as Fitting Acoustic and Perceptual Models to Data | 44 |
| XIII. | The Use of Analysis-Synthesis Models of Timbre | 45 |
| | A. Insight | 45 |
| | B. Information Reduction | 46 |
| | C. Possibility of Producing Variants | 46 |
| XIV. | Timbral Space | 47 |
| XV. | Conclusion | 49 |
| | Appendices | 50 |
| | A. Signal Representations and Analysis-Synthesis Processes | 50 |
| | B. A Synthesis Model Based on Perceptual Principles | 53 |
| | References | 54 |







Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

## I. TIMBRE

Timbre refers to the quality of sound. It is the perceptual attribute that enables us to distinguish among orchestral instruments that are playing the same pitch and are equally loud. But, unlike loudness and pitch, timbre is not a well-defined perceptual attribute. Definitions tend to indicate what timbre is not rather than what it is. Take as an example the following enigmatic definition provided by the American Standards Association (1960, p. 45): "Timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar."

The notion of timbral constancy or invariance is even vaguer than that suggested in the definitions of timbre as a basis for discrimination. It would seem that a form of timbral constancy is implied by the common observation that a sound source can be reliably identified over a wide variety of circumstances. For example, a saxophone is readily identified as such regardless of the pitch or dynamic it is playing. Furthermore, the saxophone remains a saxophone whether it is heard over a distortion-ridden pocket-sized transistor radio or directly in a concert hall. Thus, the question arises as to the physical correlates of this constancy. Is there a physical invariant or a characteristic feature mediating a given timbre?

The issue is not only academic: it has musical relevance, since the electronic and computer technology promises access to an unlimited world of timbres. One must, however, know how to evoke a given timbre; that is, how to describe it in terms of the physical structure of sound.

## II. TIMBRE AND THE FOURIER SPECTRUM: THE CLASSICAL VIEW

Physicists have been performing analyses of musical instrument tones for some time. The goal of many of these acoustical analyses is to determine the physical correlates of tone quality.

Many results of such analyses have been published (Miller, 1926; Richardson, 1954; Meyer & Buchmann, 1931; Culver, 1956; Olson, 1967). The general conclusion of such studies was that musical sounds are periodic and that the tone quality is associated solely with the waveshape, more precisely with the Fourier spectrum of the waveshape. These early analyses were strongly motivated by the theorem of Fourier, which states that a periodic waveshape is completely defined by the amplitudes and phases of a harmonic series of frequency components (see Feynman, Leighton, & Sands, 1963, Chapters 21–25; Jenkins & Watts, 1968). But the claim, often known as Ohm's acoustical law, is that the ear is phase deaf. Put more precisely, Ohm's acoustical law states that if the Fourier representation of two sounds have the same pattern of harmonic amplitudes but have different patterns of phase relationships, a listener will be unable to perceive a difference between the two sounds, even though they may have very different waveforms (see Fig. 1)

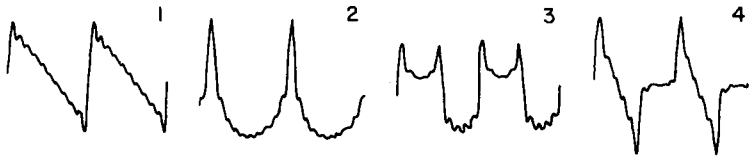



**Fig. 1.** The waves 1 to 4 correspond to tones generated with the same spectrum but with different phase relations between the components, these tones with quite different waveforms sound very similar (Plomp, 1976).

It has been argued that the ear is not actually phase deaf. It is indeed true that under certain conditions, changing the phase relationship between the harmonics of a periodic tone can alter the timbre (Mathes & Miller, 1947; Plomp & Steeneken, 1969); however, this effect is quite weak, and it is generally inaudible in a normally reverberant room where phase relations are smeared (Cabot, Mino, Dorans, Tockel, & Breed, 1976; Schroeder, 1975). One must remember, though, that this remarkable insensitivity to phase, illustrated by Fig. 1, only holds for the phase relationship between the harmonics of periodic tones.[1]

Thus, it would appear that timbre depends solely on the Fourier spectrum of the sound wave. The most authoritative proponent of this conception has been Helmholtz (Helmholtz, 1954). Helmholtz was aware that "certain characteristic particularities of the tones of several instruments depend on the mode in which they begin and end": yet he studied only "the peculiarities of the musical tones which continue uniformly," considering that they determined the "musical quality of the tone." The temporal characteristics of the instruments were averaged out by the early analyses (Hall, 1937); but since different instruments had different average spectra, it was believed that this difference in average spectrum was utterly responsible for timbre differences. This view is still widely accepted: a reputed and recent treatise like the Feynmann Lectures on Physics gives no hint that they may be factors of tone quality other than "the relative amount of the various harmonics."

Actually, even a sine wave changes quality from the low to the high end of the musical range (Köhler, 1915, Stumpf, 1926). In order to keep the timbre of a periodic tone approximately invariant when the frequency is changed, should the spectrum be transposed so as to keep the same amplitude relationship between the harmonics or should the absolute position of the spectral envelope be kept invariant? This question produced a debate between Helmholtz and Herman (cf. Winckel, 1967, p. 13). In speech, a vowel corresponds approximately to a spectrum with a given formant structure. A formant is a peak in the spectral envelope that occurs at a certain frequency, and which is often associated with a resonance in the sound source. This is the case for speech, and the formants can be related to resonances in the vocal tract.

[1] A varying phase can also be interpreted as a varying frequency. Also, dispersive media (for which the speed of propagation is frequency-dependent) cause inaudible phase distortion for periodic tones and objectionable delay distortion for nonperiodic signals (e.g., the high frequencies can be shifted by several seconds with respect to the low ones in a long cable, which makes speech quite incomprehensible).

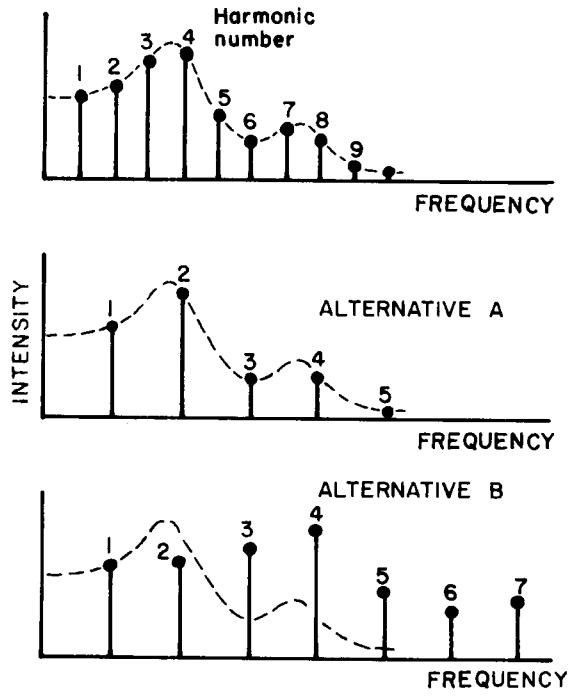



**Fig. 2.** This figure refers to an experiment by Slawson (1968) comparing alternative predictions of invariance in timbre under octave increases in fundamental frequency. The experiment rules out alternative *B*, that of the relative pitch or overtone theory, in favor of alternative *A*, that of the fixed-frequency or formant theory.

Indeed, in many cases a fixed formant structure gives a timbre that varies less with frequency than a fixed spectrum (Slawson, 1968; Plomp, 1976, pp. 107–110; Plomp & Steeneken, 1971).

Certain characteristics of the spectrum seem to induce certain timbral qualities. Brightness (or sharpness) relates to the position of the spectral envelope along the frequency axis. Presence appears to relate to strong components around 2000 Hz.

The concept of critical bandwidth, linked to the spectral resolution of the ear (Plomp, 1966), may permit a better understanding of the correlation between spectrum and timbre. In particular, if many high-order harmonics lie close together, that is, within the same critical bandwidth,[2] the sound becomes very harsh. Hence, for instance, antiresonances in the frequency response of string instruments play an important part to diminish the roughness of the tones. It may be more significant to

[2]The critical bandwidth around a certain frequency roughly measures the range within which this frequency interacts with others. The width of a critical band is about one third of an octave above 500 Hz and approximately 100 Hz below 500 Hz (cf. Zwicker & Scharf, 1965). It is an important parameter of hearing, which relates to spectral resolution (Plomp, 1964).

display spectra modified so as to take critical bands into account. This was done in some studies: the frequency axis is converted into so-called Bark units (1 Bark corresponds to the width of one critical band over the whole frequency range (Cf. Zwicker, 1961; Zwicker & Scharf, 1965: Grey & Gordon, 1978).

## III. THE SHORTCOMINGS OF THE CLASSICAL CONCEPTION

So, for periodic tones, timbre depends upon spectrum. It has long been thought that musical tones were periodic, at least for most of their duration. Musical tones are often thought of as comprising three sections: attack, steady state, and decay. Note that Helmholtz and his followers considered that timbre is determined by the spectrum of the steady state. However, this conception suffers from serious difficulties. As we noted at the beginning of this article, musical instruments can be recognized even from a very poor recording, despite the fact that their spectra are radically changed by such distortion (Eagleson & Eagleson, 1947).

In fact, a normally reverberant room has an incredibly jagged frequency response, with fluctuations up to 20 dB, and this frequency response is different at every point in the room (Wente, 1935). Hence, spectra are completely changed in ways that depend on the specific location. However, when one moves in the room, the corresponding timbres are not completely upset as one would expect them to be if they depended only on the precise structure of the frequency spectrum.

Also, various methods of sound manipulation show that temporal changes bear strongly on tone quality. Removing the initial segment of notes played by various instruments impairs the recognition of these instruments, as noted by Stumpf as early as 1910 (Stumpf, 1926). Subsequently, tape-recorder manipulation (George, 1954; Schaeffer, 1966) has made it easy to demonstrate the influence of time factors on tone quality. For instance, playing a piano tone backwards gives a nonpiano-like quality, although the original and the reversed sound have the same spectra. However, temporal factors were not taken into account in most early analyses (cf. Hall, 1937): the analysis process could not follow fast temporal evolutions.

Recently, computer sound synthesis (Mathews, 1963, 1969) has made it possible to synthesize virtually any sound from a physical description of that sound. Efforts have been made to use the results of analyses of musical instrument tones that are to be found in treatises on musical acoustics as input data for computer sound synthesis. In most cases, the sounds thus obtained bear very little resemblance to the actual tones produced by the instrument chosen; the tones thus produced are dull, lacking identity as well as liveliness (Risset & Mathews, 1969). Hence, the available descriptions of musical instrument tones must be considered inadequate, since they fail to pass the foolproof synthesis test. This failure points to the need for more detailed, relevant analyses and for a more valid conception of the physical correlates of timbre. Clearly, one must perform some kind of "running" analysis that follows the temporal evolution of the tones.

## IV. ATTACK TRANSIENTS

A few attempts have been made since 1930 to analyze the attack transients of instrument tones (Backhaus, 1932; Richardson, 1954). These transients constitute an important part of the tones—in fact, many tones like those from the piano or percussion instruments have no steady state. Yet their analysis has not produced much progress. The transients are intrinsically complex, and they are not reproducible from one tone to another, even for tones that sound very similar (Schaeffer, 1966). Most analyses have been restricted to a limited set of tones, and the researchers have tended to make generalizations that may be inappropriate even for different samples collected from the same instruments. These shortcomings have produced many discrepancies in the literature and cast doubt on the entire body of acoustic data.

## V. COMPLEXITY OF SOUNDS: IMPORTANCE OF CHARACTERISTIC FEATURES

Sounds are often intrinsically complex. Musical instruments have a complex physical behavior (Benade, 1976); often the damping is low, and transients are long compared with note duration. Also, the tones are not generated by a standardized mechanical player, but by human musicians who introduce intricacies both intentionally and unintentionally. Even if a human player wanted to, he could not repeat a note as rigorously as a machine does. If he has good control of his instrument, he should be able to play two tones sounding nearly identical, but these tones can differ substantially in their physical structure. More often the performer will not want to play all notes the same way, and his interpretation of some markings depends upon his sense of style as well as upon his technique. All these considerations, which involve different disciplines—physics, physiology, psychology, physiology, esthetics—certainly make it difficult to isolate characteristic invariants in musical instrument sounds.

This points out the need to extract significant features from a complex physical structure. Also, one must be able to control through synthesis the aural relevance of the features extracted in the analysis. Only recently has this been possible.

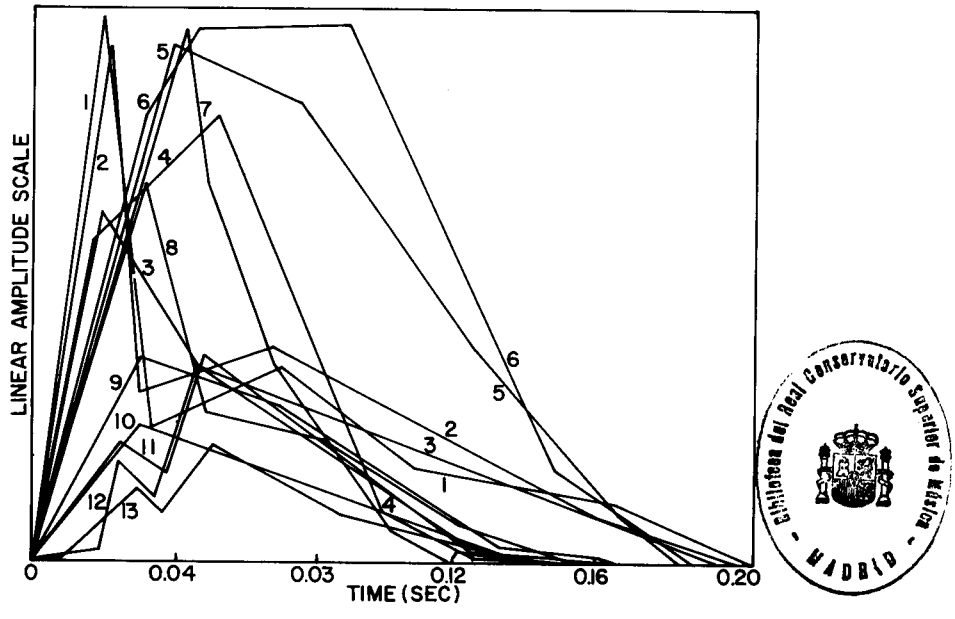
We shall now give a brief review of recent work on exploration of timbre by analysis and synthesis.

## VI. INSTRUMENTAL AND VOCAL TIMBRES: ADDITIVE SYNTHESIS

The study of trumpet tones performed in the mid-1960s by one the authors (Risset, 1966; Risset & Mathews, 1969) illustrates some of the points made above. We chose trumpet tones because we were experiencing some difficulties in synthesizing brass-like sounds with the computer. The tones synthesized with fixed spectra derived from the analysis of trumpet tones were unconvincing.

To obtain more data, we recorded musical fragments played by a professional trumpet player in an anechoic chamber. Sound spectrograms suggested that, for a given intensity, the spectrum has a formant structure; that is, it varies with frequency so as to keep a roughly invariant spectral envelope. The spectrograms gave useful information, although it was not precise enough. Thus, selected tones were converted to digital form and analyzed by computer, using a pitch-synchronous analysis (PISA program, Mathews, Miller, & David, 1961). Pitch-synchronous analysis assumes that the sound is quasi periodic; it yields displays of the amplitude of each harmonic as a function of time (one point per fundamental pitch period). The curved functions resulting from the analysis program were approximated with linear segments (see Fig. 3). These functions were then supplied to the MUSIC IV sound-synthesis program and the resulting synthetic tones were indistinguishable from the originals even when compared by musically skilled listeners. Hence, the additive synthesis model, with harmonic components controlled by piece-wise linear functions, captures the aurally important features of the sound.

Conceptually, the model is simple. The pitch-synchronous analysis yields a kind of time-variant harmonic analysis that is further reduced by fitting the linear segments to the amplitude envelope of each component. However, computationally this model is not very economical. Figure 3 shows that the functions can be quite complex and an estimation of the parameters must be performed for every tone. So further simplifica-




Biblioteca del Real Conservatorio Superior de Música - MADRID

**Fig. 3.** This figure shows line-segment functions that approximate the evolution in time of 13 harmonics of a D4 trumpet tone lasting 0.2 sec. Functions like these, obtained by analysis of real tones, have been used to control the harmonic amplitudes of synthetic tones (Risset & Mathews, 1969).

tions of the model were sought. By systematic variation of the various parameters—one at a time—the relative importance of the parameters were evaluated. Whereas some parameters were dismissed as aurally irrelevant—for example, short-term amplitude fluctuations—a few physical features were found to be of utmost importance. These include the following: the attack time, with faster build-up of the low-order harmonics than the high-order ones; for certain tones, a quasi-random frequency fluctuation; and, most importantly, a peak in the frequency spectrum between 1000 and 1500 Hz and an increase in the proportion of high-order harmonics with intensity.

In fact, the latter property permitted us to abstract a simplified model of brasslike tones. Here only the amplitude function for the first harmonic was provided and the amplitude functions for the other harmonics were deduced as fixed functions of this first harmonic amplitude so that they increased at a faster rate. The specification was much more economical than the previous one and did not need to be precisely adjusted to yield the brasslike quality. Hence this property of an increase in spectral content with amplitude seems to be the most salient physical correlate of brass-tone quality. Beauchamp (1975) studied these nonlinear interharmonic relationships in cornet tones and ascribed the brass-like character to the type of nonlinear relationship between the different harmonics, which are all functions of the first one regardless of the general level. This relationship has been found later to have an acoustical basis (Benade, 1976, pp. 439–447: Backus & Hundley, 1971). This nonlinear property has been used to produce brass-like sounds with synthesizers, in which a voltage-controlled low-pass filter's cutoff frequency is increased with amplitude. This characteristic has also been implemented in a very simple, satisfying way, using Chowning's powerful technique of spectral generation by frequency modulation (Chowning, 1973; Morrill, 1977).

It was found in the trumpet-tone study that some factors may be important in some conditions and inaudible in others. For instance, details of the attack were more audible in long sustained tones than in brief tones. Also, it appeared that some listeners, when comparing real and synthetic tones, made their decision about whether a tone was real or synthetic on the basis of some particular property. For instance, they often assumed that the real tones should be rougher, more complex than the synthetic ones. This suggests that by emphasizing roughness in a synthetic tone, one could cause the listeners to believe it was a real tone. In his striking syntheses of brassy tones, Morrill (1977) has simulated intonation slips that greatly enhance the realistic human character of the tones. Similarly, in their study of string tones, Mathews, Miller, Pierce, and Tenney (1965, 1966) had included an initial random-frequency component, which corresponds to erratic vibration when the string is first set in motion by the bow. When exaggerated, this gives a scratchy sound strikingly characteristic of a beginning string player. Such idiomatic details, imperfections, or accidents (Schaeffer, 1966) are characteristic of the sound source, and the hearing sense seems to be quite sensitive to them. Taking this into account might help to give stronger identity and interest to synthetic sounds. Indeed, a frequency skew imposed on even a simple synthetic tone can help strongly endow it with subjective naturalness

and identity. The pattern of pitch at the onset of each note is often a characteristic feature of a given instrument: the subtle differences between such patterns (e.g., a violin, a trombone, a singing voice) act for the ear as signatures of the source of sound.

The paradigm for the exploration of timbre by analysis and synthesis followed in the latter study has been much more thoroughly pursued by Grey and Moorer (1977a,b) in their perceptual evaluation of synthesized musical instrument tones. Grey and Moorer selected 16 instrumental notes of short duration played near E♭ above middle C. This pitch was selected because it was within the range of many instruments (bass clarinet, oboe, flute, saxophone, cello, violin, etc.); thus, the tones represented a variety of timbres taken from the brass, string, and woodwind families of instruments. The tones were digitally analyzed with a heterodyne filter technique, providing a set of time-varying amplitude and frequency functions for each partial of the instrumental tone. Digital additive synthesis was used to produce a synthetic tone consisting of the superposition of partials, each controlled in amplitude and frequency by functions sampled in time. Each of the 16 instrumental notes could appear in at least four of the five following conditions: (1) original tone; (2) complex resynthesized tone, using the functions abstracted from the analysis: (3) tone resynthesized with a line-segment approximation to the functions (4 to 8 line segments); (4) cut-attack approximation for some of the sounds; and (5) constant-frequencies approximation. In order to evaluate the audibility of these types of data reduction, systematic listening tests were performed with musically sophisticated listeners. The tones were first equalized in duration, pitch, and loudness. An AA AB discrimination paradigm was used. On each trial four tones were played, three of them identical and the fourth one different; the listeners had to detect whether one note was different from the others, to tell in which pair it was located and to estimate the subjective difference between this note and the others. The judgments were processed by multidimensional scaling techniques.

The results demonstrated the perceptual closeness of the original and directly resynthesized tones. The major cue helping the listeners to make a better than chance discrimination was the tape hiss accompanying the recording of the original tones and not the synthetic ones. The results also showed that the line-segment approximation to the time-varying amplitude and frequency functions for the partials constituted a successful simplification, leading to a considerable information reduction while retaining most of the characteristic subjectivity. This suggests that the highly complex microstructure in the time-varying amplitude and frequency functions is not essential to the timbre and that drastic data reduction can be performed with little harm to the timbre. The constant frequencies approximation (for tones without vibrato) was good for some tones but dramatically altered other ones. The importance of the onset pattern of the tones was confirmed by the cut-attack case.

A recent study by Charbonneau (1979) has demonstrated that the simplification can go even further for most of the tones studied by Grey and Moorer (namely short tones of nonpercussive instruments). The various envelopes controlling each harmonic are replaced by a single averaged envelope; for each harmonic, this curve is weighted in order to preserve the maximum amplitude for this harmonic; it is also warped in time

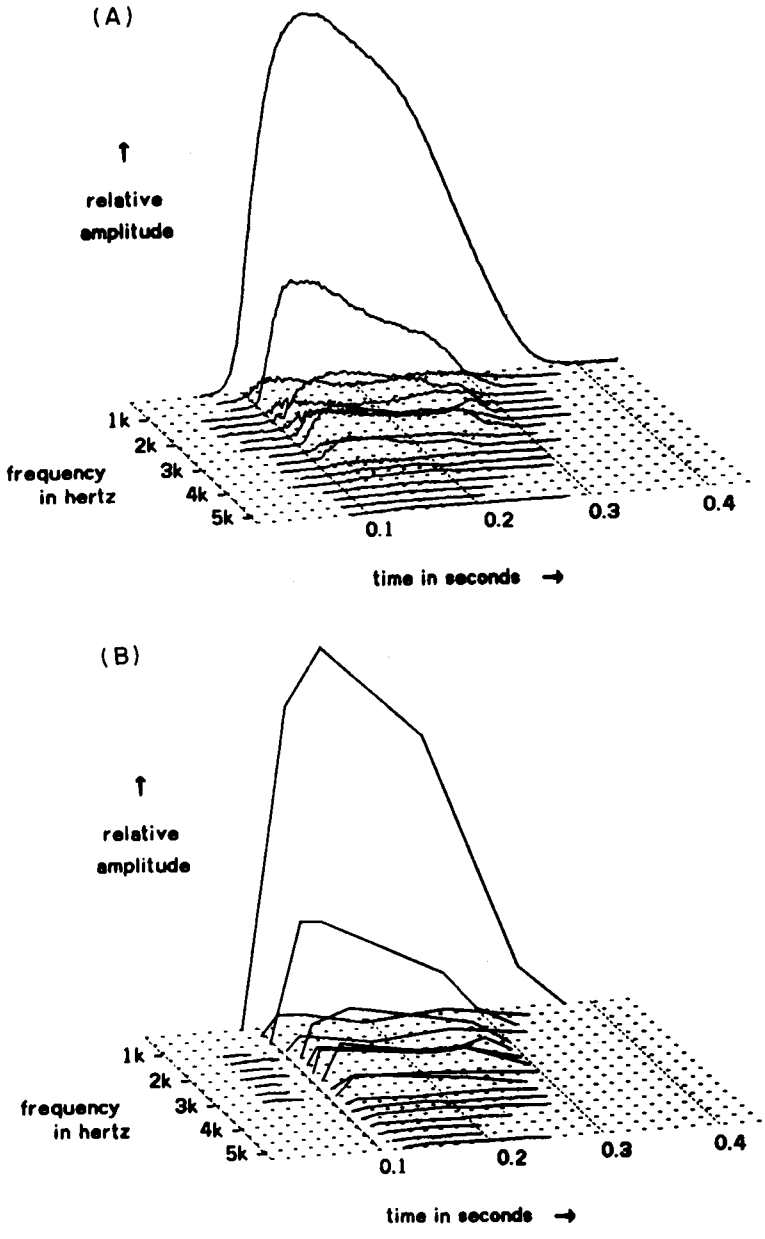



**Fig. 4.** (A): Time-varying amplitude functions derived from heterodyne analysis for a bass clarinet tone, shown in a three-dimensional perspective plot. (B): Line-segment approximation to the functions plotted in A. Both of these functions have been used to resynthesize the tone. Form B gives a considerable information reduction (Grey & Moorer, 1977).

in order to preserve the times of appearance and extinction of the various harmonics. While this is not a proper model for flute tones, it permits a good imitation for most of the other instruments.

Fletcher and his collaborators (Fletcher, Blackham, & Stratton, 1962; Fletcher & Bassett, 1978; Fletcher, Blackham, & Christensen, 1963; Fletcher & Sanders, 1967) studied the timbre of several instruments by analysis and synthesis, using an additive synthesis model. (The earlier of these studies did not use a computer but *ad hoc* analysis and synthesis devices.) A study of the quality of piano tones (1962) indicated that the attack time must be less than 0.01 sec, whereas the decay time can vary from 20 sec for the lowest notes to less than 1 sec for the very high ones. The variation of partial level versus time during the decay was highly complex and not always monotonic—the partials at time increase in intensity rather than decrease. However, the complexities of the decay pattern did not appear to be very relevant to the ear since the much simplified syntheses could sound similar to the original sounds. The study provided a major insight. It ascribed subjective warmth to the inharmonicity of the partials. The frequencies of the successive partials of a low piano tone are close to, but higher than, the frequencies of the harmonic series (to the extent that the 15th partial frequency can be 16 times that of the lowest one (Young, 1952). Now this slightly inharmonic pattern gives rise to a complex pattern of beats that induces a peculiar lively and warm quality. This is an important feature for low piano tones (and also for organ tones; cf. Fletcher *et al.*, 1963). Actually, many analyses have been performed on piano sounds (Martin, 1947). They have been used to devise electronic pianos (Dijksterhuis & Verhey, 1969) whose tone quality (while not fully satisfying) depends upon the simplified model abstracted from the analyses.

In a study of violin tones, Fletcher and Sanders (1967) investigated the slow frequency modulation (around 6 Hz) known as vibrato, showing that it also modulates the spectrum of the tone. They also pointed to two features that enhance naturalness if they are simulated in the synthetic tones: the bowing noise at the onset of the tone and the sympathetic vibrations coming from the open strings (the latter occur substantially only when certain frequencies are played).

Clark, Luce and Strong have also performed significant research on wind instrument tones by analysis and synthesis. In a first study (Strong & Clark, 1967a) wind instrument tones were synthesized as the sum of harmonics controlled by one spectral envelope (invariant with note frequency) and three temporal envelopes. (A more specific model was also sought for brass instruments, cf. Luce & Clark, 1967). Listeners were tested for their capacity to identify the source of the tones. Their identification was nearly as good as for real instrument tones, which indicates that this model grasps the elements responsible for the difference between the sounds of the different instruments. Incidentally, the probability of confusion between the tones of two instruments gives an indication of the subjective similarity between these tones; it has been used to ascertain the perceptual basis of the conventional instrument families (cf. Clark, Robertson, & Luce, 1964). The results suggest that some conventional families represent fairly well the subjective differentiations, especially the string and the brass

family. A double reed family also emerged, comprising a tight subfamily (oboe and English horn) and a more remote member (the bassoon).

## VII. CROSS-SYNTHESIS AND VOICE SYNTHESIS

In another study, Strong and Clark (1967b), in order to evaluate the relative significance of spectral and temporal envelopes, resorted to an interesting process: they exchanged the spectral and temporal envelopes among the wind instruments and asked listeners to attempt to identify these hybrid tones. The results indicated that the spectral envelope was dominant if it existed in a unique way for the instrument (as in the oboe, clarinet, bassoon, tuba, and trumpet); otherwise (as in the flute, trombone, and French horn), the temporal envelope was at least as important.

It should be noted that the above conclusions apply to wind instruments, which can have different temporal characteristics, although not very drastic ones. On the other hand, it is easy to verify by synthesis that a sharp attack followed by an exponential decay gives a plucked or percussive quality to any waveform. In this case, temporal cues tend to dominate over spectral ones.

One often speaks of cross synthesis to characterize the production of a sound that compounds certain aspects of a sound A and other aspects of a sound B. There are interesting possibilities for cross synthesis when sound production can be modeled as the combination of two relatively independent processes. In particular, a sound source can often be thought of as comprising an excitation that is transformed in ways that can be characterized in terms of a stable response (Huggins, 1952)—think of someone hitting a gong or blowing into a tube. The temporal properties of the sound are often largely attributable to the excitation insofar as the response depends on the structural properties of a relatively stable physical system; the spectral aspects result from a combination of those of the excitation and those of the response. (Huggins suggests that the hearing mechanism is well equipped to separate the structural and temporal factors of a sound wave.) A good instance is that of voice production (cf. Fant, 1960): the quasi-periodic excitation by the vocal cords is fairly independent of the vocal tract response, which is varied through articulation. Thus, the speech waveform can be characterized by the *formant* frequencies (i.e., the frequencies of the vocal tract resonances) and by the fundamental frequency (*pitch*) of the excitation—except when the excitation is noise-like (in unvoiced sounds like *s* or *f*).

A considerable amount of research on speech synthesis has demonstrated the validity of this physical model. It is possible to synthesize speech that sounds very natural. It remains difficult, however, to mimic the transitions in spectrum and frequency that occur in speech with enough accuracy and suppleness. In fact, while one can faithfully imitate a given utterance by analysis and synthesis, it is still difficult to achieve a satisfactory "synthesis by rule," whereby the phonetic elements (phonemes or dyads) would be stored in terms of their physical description and concatenated as needed to form any sentence, with the proper adjustments in the physical parameters performed automatically according to a set of generative rules. We cannot dwell at length here on

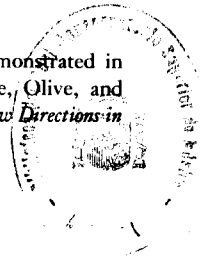
this important problem; we can notice that the correlates of speaker's identity are multiple and the spectral quality of the voice as well as the rhythmic and intonation patterns are significant. At this time one cannot reliably identify speakers from their voiceprints as one can from their fingerprints (cf. Bolt, Cooper, David, Denes, Pickett, & Stevens, 1969, 1978).

The notion of independence between the vocal tract and the vocal cords is supported by an experiment by Plomp and Steeneken (1971); however, it has to be qualified for the singing voice. For high notes, sopranos raise the first formant frequency to match that of the fundamental in order to increase the amplitude (Sundberg, 1977). Specific features detected in the singing voice have been recently confirmed by synthesis in the work of Sundberg, Chowning, Rodet, and Bennett. Through certain processes of analysis (like inverse filtering or linear predictive coding—cf. Flanagan, 1972), one can decompose a speech signal to separate out the contributions of the vocal cords and the vocal tract. These processes made it possible for Joan Miller to synthesize a voice as though it were produced with the glottis of one person and the vocal tract of another one (cf. Mathews *et al.*, 1961).[3] Actually, the source signal—due to the vocal cords—can be replaced by a different signal, provided this signal has enough frequency components to excite the vocal tract resonances (between, say, 500 and 3000 Hz). It is thus possible to give the impression of a talking (or singing?) cello or organ. Composers are often interested in less conspicuous effects, for instance in producing timbres from the combination of two specific tone qualities, using processes other than mere mixing or blending. This can be achieved through processes of analysis and synthesis—like the phase vocoder or the predictive coding process, or also through the reconstitution of the sounds through a certain model, like frequency modulation or additive synthesis. By physically interpolating the envelopes of the harmonics, Grey and Moorer (1977) have been able to gradually transform one instrumental tone into another one (e.g., a violin into an oboe) through monodic intermediary stages that do not sound like the mere superposition of a violin and a oboe.

## VIII. ADDITIVE SYNTHESIS: PERCUSSION INSTRUMENTS

The above studies of timbre resorted to models of additive synthesis, whereby the sound was reconstituted as the superposition of a number of frequency components, each of which can be controlled separately. Such models require much information specifying in detail the way each component varies in time: hence, they are not very economical in terms of the amount of specification or the quantity of computations

[3]Impressive examples of voice synthesis and processing for musical uses have been demonstrated in particular by Bennett, Chowning, Moorer, Olive and Petersen. Compositions by Dodge, Olive, and Petersen using synthesized voices can be heard on record CRI SD 348, and on the album *New Directions in Music*, Tulsa studios.



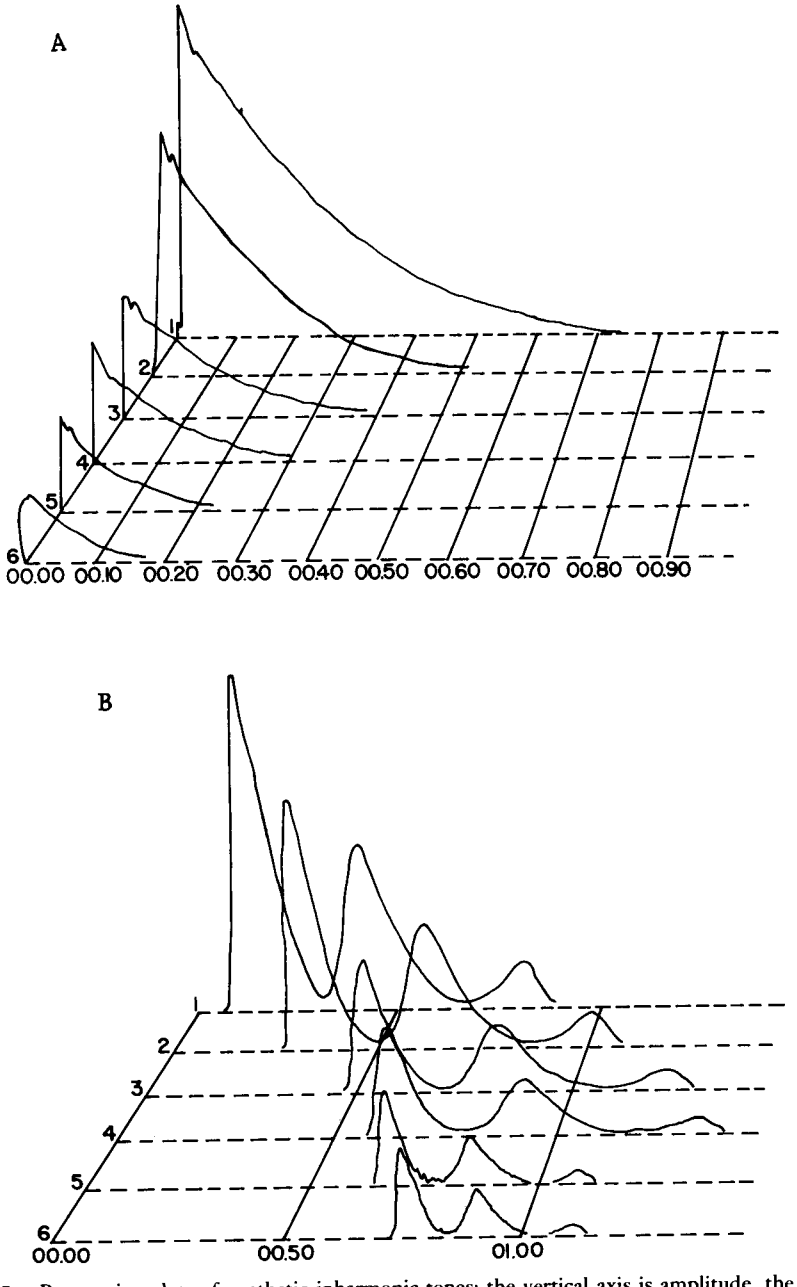



**Fig. 5.** Perspective plots of synthetic inharmonic tones: the vertical axis is amplitude, the horizontal axis is time, and the depth axis is frequency. In *A*, the sharp attack followed by a decay yields a bell-like tone. In *B*, the time-varying amplitude function yields a fluid nonpercussive tone, in which the components can be heard much better than in the fused bell-like tone. (In spite of the appearance of equal spacing, these tones have nonharmonic components.)

they require. However, as was stated, the information on the temporal behavior of the components can often be simplified. In addition, the development of the digital technology has made it possible to build special processors with considerable processing power, for instance, digital synthesizers that can yield in real time dozens of separate voices with different envelopes (Alles & Di Giugno, 1977); so additive synthesis is a process of practical interest, considering its power and generality. It is not restricted to quasi-periodic tones; in fact, it can be used to simulate the piano and percussion instruments (Risset, 1969; Fletcher & Bassett, 1978).

In percussion instruments, the partials are no longer harmonics: their frequencies, found from the analysis, are those of the modes of vibration excited by the percussion and can sometimes be predicted from consideration of theoretical acoustics. The synthesis can correspond to a considerably simplified model and still be realistic, provided it takes into account the aurally salient features. Fletcher and Bassett (1978) have simulated bass drum tones by summing the contribution of the most important components detected in the analysis—these were sine waves decaying exponentially, with a frequency shift downwards throughout the tone. The simulation was as realistic as the recorded bass drum tones. The authors noted, however, that the loudspeakers could not render the bass drum tones in a completely satisfactory way.

Timbre can often be evoked by a synthesis that crudely takes into account some salient properties of the sound. Bell-like tones can be synthesized by adding together a few sine waves of properly chosen frequencies that decay exponentially at different rates—in general, the higher the frequency, the shorter the decay time. Realism is increased by introducing slow amplitude modulation for certain components of the spectrum. Such modulations exist for real bells; they can be ascribed to beats between closely spaced modes because the bell does not have perfectly cylindrical symmetry.) Snare drums can also be imitated because the decays are much faster than for bells and the effect of the snares can be evoked by adding a high-pitched noise band (Risset, 1969). Bell-like or drum-like sounds synthesized this way can also be transformed morphologically by changing the envelopes controlling the temporal evolution of the components. Thus, for instance, bells can be changed into fluid textures with the some harmonic (or rather inharmonic)[4] content yet with a quite different tone quality (see Fig. 5).

## IX. SUBSTRACTIVE SYNTHESIS

Whereas additive synthesis, used in most of the previous examples, builds up the tone as the sum of elementary components, subtractive synthesis consists of submitting a spectrally rich wave to a specific type of filtering, thus arriving at the desired tone by eliminating unwanted elements rather than by assembling wanted ones. Subtractive synthesis is better adapted to certain types of sounds. As was mentioned, the process of speech articulation consists of shaping the vocal tract so that it filters in a specific way the spectrally rich source signal produced by the vocal cords. In fact,

[4]cf. Inharmonique, in the record *Risset-Mutations* (INA-GRM AM 546 09).

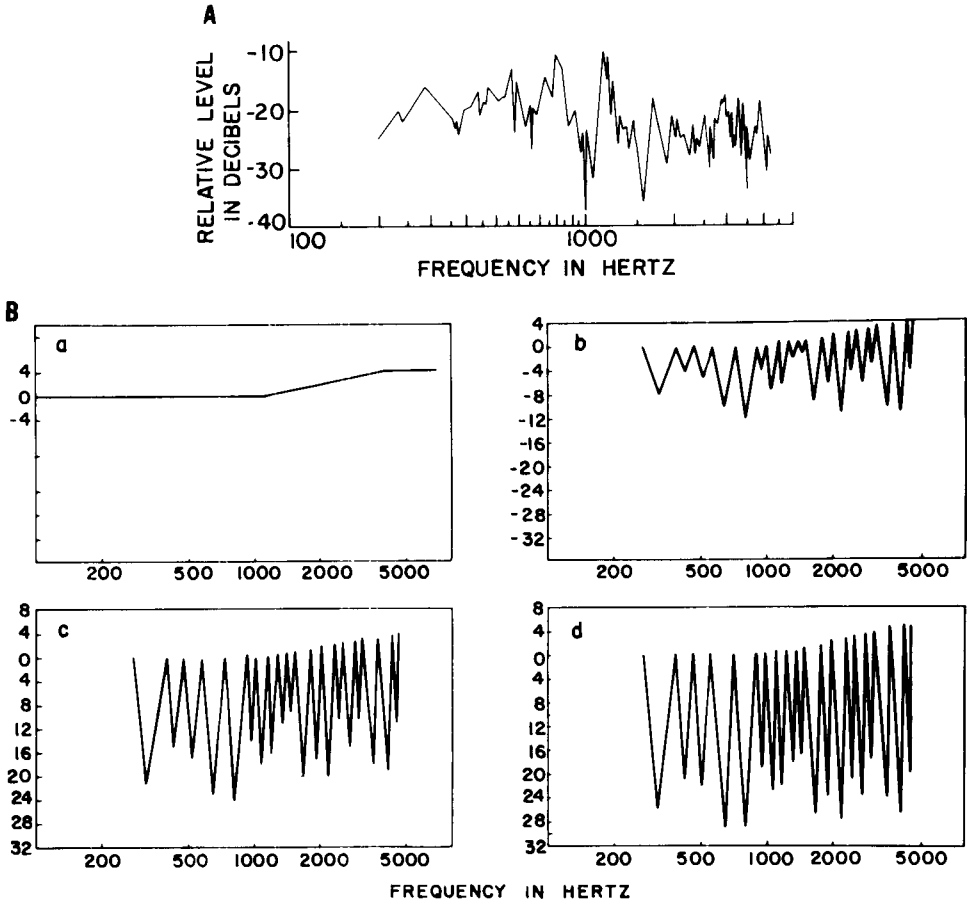



**Fig. 6.** Relative frequency response: *A*, as measured in a real violin from sinewave excitation; *B*, as simulated in the electronic replication of the violin tone: from (a) to (d), the Q of the simulated resonances increases from a too low to a too high value (Mathews & Kohut, 1973).

linear prediction coding consists of adjusting the parameters of a recursive filter so as to minimize the difference between the original speech signal and the signal obtained by filtering a single, quasi-periodic pulse wave by this time-variant recursive filter (see below).

Another instance in which subtractive synthesis has proven most useful is the case of violin tones, as demonstrated by Mathews' electronic violin. Mathews and Kohut (1973) have studied the aural effect of the resonances of the violin box through electronic simulation. They have approximated the complex frequency response of a violin (which exhibits many peaks and minima—as many as 20 or more in the audible frequency range) with a set of electrical resonant filters (between 17 and 37). In this experiment, the vibration of the violin string near the bridge was converted into an

electric signal by a magnetic pickup. This signal was approximately a triangular wave, as predicted by Helmholtz (Kohut & Mathews, 1971); hence, it consisted of a number of significant harmonic components whose amplitude decays regularly with the rank. This signal was then subjected to the complex filtering approximating the response of the box. It was possible to change the characteristics of that filtering by changing both the damping of the resonances and their distribution along the frequency axis. It was found that a violin-like tone could be achieved with 20 or 30 resonances distributed in the frequency range of 200–5000 Hz, either randomly or at equal musical intervals (see Fig. 6). The best tone was obtained with intermediate values of damping, corresponding to a peak-to-valley ratio of about 10 dB in the response curve of the resonant filter. With too small a damping, the sound was even but dull; with too great a damping, the sound was hollow and uneven in intensity for various pitches.

The experimental equipment constitutes an electronic violin, which has been used musically to obtain either violin-like tones (e.g., in a quartet by M. Sahl) or sounds of very different qualities, by changing the filter settings (e.g., in pieces by V. Globokar or M. Urbaniak).

This experiment has suggested to the authors that the specific quality of violin vibrato could be due to the interaction of the frequency modulation with the resonant peaks, producing a complex pattern of spectral modulation. Different harmonics are modulated in different ways, depending on the slope of the frequency response at the position of these harmonics. This can be verified by synthesis: the effect is not very sensitive to the parameters of a jagged frequency response. Imitative synthesis of violin tones (Schottstaedt, 1977) indicates that a good vibrato quality may be obtained in apparently simpler ways, not ensuring a truly fixed spectral envelope, but modulating the spectrum in a complex way, through a variant of Chowning's frequency modulation technique.

## X. ACOUSTIC MODELING AS A SYNTHESIS TECHNIQUE

An ambitious approach for analysis and synthesis has been tried by Hiller and Ruiz (1971) in their work on the use of physical models for string sounds. The analytical study of the acoustic behavior of an instrument can lead to differential equations governing the motion of the vibrating elements. One can try to synthesize the sound by solving these differential equations. This approach is in a way the reciprocal of the one used in analog computers, in which one assembles a physical system with parameters governed by the equations to be solved. In the latter case the measurement of these parameters gives solutions to these equations. In the study of Hiller and Ruiz the resolution of the differential equations gives an approximation to the sound of the instrument. This approximation may be a good one if the differential equations embody a good physical model.

Now this method makes it easy to change at will the physical parameters of the model—to the extent of choosing completely unrealistic values (like negative stiffness)

or transforming step-by-step a vibrating object into a different one—and to produce the corresponding sound. Unfortunately, in most cases the equations of motion provide only crude approximations to systems as complex as musical instruments, and even then the computations necessary to solve them are often overwhelming in that they require great numerical precision and a considerable amount of computing power. The development of faster computers and special fast processors may eventually make this method more conveniently usable, but at the moment it still appears difficult to apply (except for demonstration purposes). In fact, Ruiz has produced convincing demonstrations of the behavior of the violin string through the solution of the equations for motion. One may also mention here the use by Freedman (1967) of a model somewhat based on physical behavior, and also the relevance of some physical processes that have been used, if only in a schematic way, as the basis of a synthesis technique. We have already mentionned some of these processes, like nonlinear distorsion (Risset, 1969, #150, Beauchamp, 1975; Arfib, 1979; Le Brun, 1979). Weinreich (1977, 1979) has shown the contribution to the tone of the piano of the coupling between strings which are not exactly tuned to the same frequencies (this ensures the prolongation of the tone as well as a specific quality): he is currently applying this model successfully to the synthesis of piano-like tones.

## XI. THE IMPORTANCE OF CONTEXT

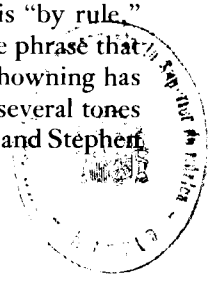
The importance of a given cue depends on context. For instance, details of the attack of trumpet-like tones (especially the rate at which various partials rise) are more significant in long sustained tones than in brief or evolving tones (Risset, 1965, 1966). In the case of a very short rise time (as in the piano), the subjective impression of the attack is actually more determined by the shape of the beginning of the amplitude decay (Schaeffer, 1966). The acoustics of the room may also play an important role (Benade, 1976; Schroeder, 1966; Leipp, 1971). The sound of an organ, for instance, depends considerably upon the hall or church where it is located.

Most of the exploration of timbre by analysis and synthesis has focused on isolated tones, but music usually involves musical phrases. Throughout these phrases, the physical parameters of the tones evolve, and this evolution can obscure the importance of certain parameters which are essential for the imitation of isolated tones. Similarly, in the case of speech the parameters of isolated acoustic elements (e.g., phonemes) undergo a considerable rearrangement when the elements are concatenated to form sentences. The specification of simple and valid models of this rearrangement is the problem of speech synthesis by rule. The importance of prosodic variations throughout the sentence is obvious in speech; pitch bends and glides even subtle ones are also essential in music. In a musical context the evolution of various parameters throughout a phrase can be significant. The prosodic variation of one parameter may subjectively dominate other parameters in a musical phrase. So, it is essential to study musical prosody by analysis and synthesis. Actually, this appears to

be the new frontier for exploration of analysis and synthesis. (One had to first understand the parameters of isolated tones to be able to describe how they evolve in a musical phrase.)

Currently, musical prosodic studies appear difficult since the phrasing is likely to depend upon the musical style. Its importance seems greater, for instance, in Japanese shakahachi flute playing than in Western instrumental playing. In the latter the music is built from fairly well-defined and relatively stable notes from which the composer can make up timbres by blending, whereas in the former the state of the instrument is constantly disrupted. Hence, a prosodic study on the shakahachi is interesting, even necessary, since the sound can only be described properly at the level of the phrase (A. Gutzwiller, private communication). Mathews has used the GROOVE hybrid synthesis system (Mathews & Moore, 1970), which permits the introduction of performance nuances in real time, to explore certain correlates of phrase, for instance the role of overlap and frequency transition between notes in achieving a slurred, legato effect. Using his algorithms for trumpet synthesis, Morrill is studying the correlate of phrasing in the trumpet. Grey (1978) has studied the capacity of listeners to distinguish between recorded instrumental tones and simplified synthetic copies when the tones were presented either in isolation or in a musical context (single- or multivoiced). He found that while multivoice patterns made discrimination more difficult, single-voice patterns seemed to enhance spectral differences between timbres, while isolated presentation made temporal details more apparent. This finding may relate to the phenomenon of *stream segregation* (Bregman & Campbell, 1971; McAdams & Bregman, 1979; see also Chapter 4, this volume), an important perceptual effect which can be described as follows: if a melodic line is made up of rapidly alternating tones belonging to two sets that are sufficiently separated, the single stream of sensory input splits perceptually into segregated lines. (Baroque composers, such as Bach, resorted to this interleaving of lines to write polyphonic structures for instruments capable of playing only one note at a time.) This segregation is helped by increasing the frequency separation between the lines. Recently, studies by van Noorden (1975) and by Wessel (1979) indicates that the influence of frequency separation on melodic fission has more to do with brightness—that is, with spectral differences—than with musical pitch per se, which appears to be linked with Grey's finding on single-voice patterns.

Chowning has recently performed synthesis of sung musical phrases that sound supple and musical. In addition to carefully tuning the tone parameters for each note, he has given due care to the change of musical parameters throughout the phrase. He has found that the parameters had to vary in ways that are to some extent systematic and to some extent unpredictable. These changes seem to be essential cues for naturalness. In fact, the musical ear may be "turned off" by a lack of variability in the parameters, which points to an unnatural sound for which even complex details may be aurally dismissed. One can hope to develop musical phrase synthesis "by rule," that is, to find rules for the change of musical parameters throughout the phrase that would yield musically acceptable phrases. In the course of this study Chowning has given strong evidence that the addition of the same vibrato and jitter to several tones enhances the fusion of these tones, a fact investigated by Michael McNabb and Stephen



McAdams (1982). Chowning's syntheses strongly suggest that the ear relies on such micromodulations to isolate voices among a complex aural mixture such as an orchestral sound.

In a complex auditory situation it often appears that one dominant feature can eradicate more subtle differences. The most striking aspect according to which the stimuli differ will be taken into consideration rather than the accumulation of various differences between a number of cues. Lashley (1942) has proposed a model of such behavior in which the dominant feature masks the less prominent features. This often seems to hold for perception in a complex environment. Certainly, in the case of musical timbre, which can depend on many different cues, context plays an essential role in assessing whether or not a given cue is significant.

## XII. ANALYSIS-SYNTHESIS AS FITTING ACOUSTIC AND PERCEPTUAL MODELS TO DATA

Having described a number of significant studies of timbre by analysis and synthesis, we shall pause here to put these studies in a conceptual framework that will help us to understand possible applications of the analysis-synthesis approach.

A general scheme that we have found useful is shown in Fig. 7. The analysis-synthesis process begins with a sound that is to be modeled. In these general terms the analysis of a sound involves estimating the parameters of a model (for example, in the Fourier analysis model the frequencies, amplitudes, and phases of a set of sine wave components must be estimated). Once the parameters of the model have been estimated, the model can be driven with them to generate a synthetic version of the

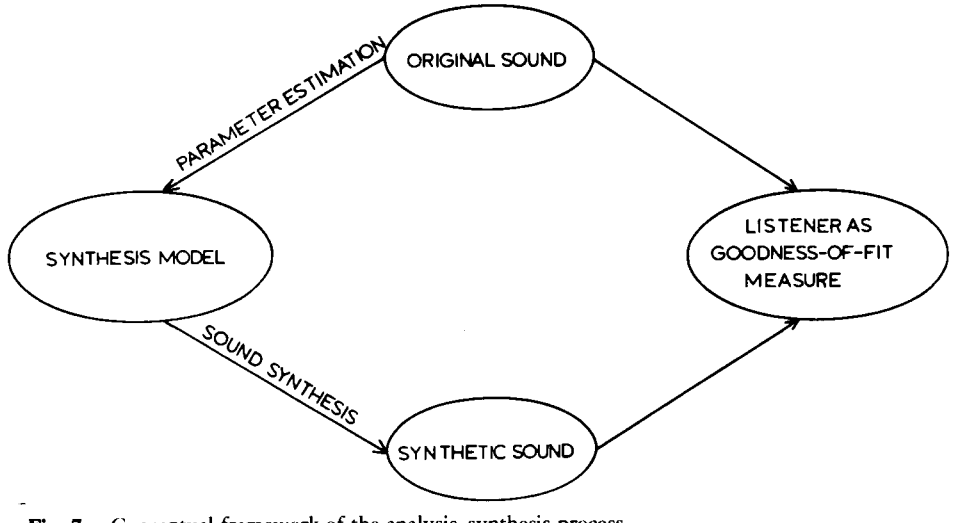



**Fig. 7.** Conceptual framework of the analysis-synthesis process.

original sound. For our purposes the appropriate goodness-of-fit evaluation technique is to make auditory comparisons between the original sound and its synthetic replica. If the analysis–synthesis model captures the essential perceptual features of the sound in a thorough way, then the listener should be unable to distinguish the difference between the original and the synthetic version.

The above criterion of validity characterizes what we call a *perceptual model*, as opposed to an *acoustic model*: the latter would mimic the physical mechanisms that give rise to the sound whereas the former simulates the sound through processes that may well not reflect the way the sound is really produced, provided the aural result comes close enough to the original. As we have seen, a good acoustic model can also be a good perceptual model; but the physical behavior of the sound-emitting bodies is very complex, and acoustic simulations require simplifications such that they can rarely sound faithful to the ear. While hearing is very demanding in some respects, it is also very tolerant in other respects: perceptual models can concentrate on those features to which the ear is most sensitive.

Acoustical and perceptual models often represent waveshapes in terms of certain mathematical functions. In the appendix the reader will find some general notions about representations of signals and their utility. A workable and reasonably general perceptual model is also described.

## XIII. THE USE OF ANALYSIS–SYNTHESIS MODELS OF TIMBRE

The models drawn from analysis–synthesis of timbre can be useful for several purposes: (1) providing of insight and understanding. (2) information reduction; and the (3) possibility of producing variants or modifications.

### A. Insight

Analysis–synthesis provides insight into the perception of timbre, which displays highly specific features. Many of these features can perhaps be better understood from an evolutionary perspective, considering the ways in which hearing has adapted to provide useful information about the environment. For instance, hearing is very sensitive to changes: it is well equipped to be on the alert, which makes sense since sounds propagate far and around obstacles. Perhaps this is why the musical ear tends to reject steady sounds as dull and uninteresting. Hearing is very sensitive to frequency aspects, which are only rarely modified between the sound source and the listener. On the other hand, the ear is quite insensitive to the phase relations between the components of a complex sound, which is fortunate since these relations are smeared in a reverberant environment. Timbre is related to rather elaborate patterns that resist distortion (e.g., the relationship between spectrum and intensity in the brass). From these elaborate patterns, hearing has intricate ways to extract informa-

tion about loudness and distance. Models of timbre shed light on our capacity to assign different sounds to the same source, for instance, recognition of a note as such regardless of the register in which it is playing. The models help us to to understand what properties form the basis of such categorization.[5] This understanding can be of importance in the fields of experimental music; a composer may want to confer some distinctive identity to certain artificial sounds.

### B. Information Reduction

Usually we require that there should be many fewer parameters in the analysis-synthesis model than there are degrees of freedom in the data of the original signal. This is a form of data reduction. For example, consider a digitally sampled sound one second duration. If the sampling rate is 40,000 samples per second and if we wish to account for all these sample values in our model, then we could trivially simulate this signal with a model containing 40,000 parameters: however, a model with a reduced amount of information would be more practical. In fact, much research on speech analysis-synthesis (e.g., the channel vocoders) has been performed to try to find a coding of speech that would reduce the bandwidth necessary to transmit the speech signal (Flanagan, 1972). Such a coding would in fact be an analysis-synthesis model because the speech would be analyzed before transmission and resynthesized at the other end (see Appendix). Such systems have only occasionally been put into practical use since it is difficult to preserve a good speech quality and since the price of the transmission bandwidth has gone down substantially, so that the devices implementing analysis and synthesis at the ends of the transmission line would be more costly that the economized bandwidth. However, information reduction can work very well for certain types of sound, as we have already seen above (Grey & Moorer, 1977): linear predictive coding is an economical way to store speech and is now used in portable speaking machines.

### C. Possibility of Producing Variants

If one manipulates the parameters before resynthesis, one will obtain modifications of the original sound and such modifications can be very useful. For instance, starting with a recording of a spoken sentence, one can change the speed by playing it on a variable-speed tape recorder; however, the pitch and the formant frequencies will also be changed, completely distorting the original speech. Now if one analyzes this sentence according to an analysis-synthesis process which separates glottal excitation and vocal tract response [e.g., channel vocoder, phase vocoder, linear predictive

[5]It seems clear that the identity of the timbre of an instrument such as the clarinet, whose high notes and low notes are physically very different, must be acquired through a learning process. It has been proposed that this learning process involves senses other than hearing; in particular, the experiments of Cadoz *et al.* (1981) aim at better understanding "motor" aspects of timbre perception, in particular how the gestural experience of producing a sound interacts with its perception.

coding (Flanagan, 1972; Moorer, 1978)], one can then alter the tempo of articulation independently of the pitch. Moorer has thus been able to slow down speech excerpts by a factor of 20 or more without loosing quality or intelligibility. This example shows the usefulness of analysis-synthesis in obtaining variants of the original sounds.

We shall distinguish here between two uses of sound modification: classical musical processing and expanding timbral resources. In classical musical processing the goal is to transform the sound so as to maintain timbral identity while changing pitch and/or duration (also possibly articulation and loudness). For instance, as mentioned above, linear predictive coding or phase vocoder analysis-synthesis permits the changing of pitch and speed independently. Also, as was discussed at the beginning of this chapter (see Fig. 2), it is often improper to keep the same spectrum when one changes pitch. It may also be necessary to change the spectrum as one changes loudness. Such changes are essential if one wants to use digitally processed real sounds (e.g., instrumental sounds) for music. Without resorting to analysis-synthesis processes, one can only perform rather superficial and often unsatisfying modifications of the sound. On the other hand, one should be aware that these processes are complex and difficult to implement, especially in real time. Even a fast digital processor can have difficulty in coping with the demands of real time if it has to perform analysis-synthesis processes.

In expanding timbral resources, the goal is different: to change certain aspects of the tone so as to modify the timbre while preserving the richness of the original model. Here again, analysis-synthesis processes are essential for allowing interesting timbral transformations (like cross-synthesis), interpolation between timbres (Grey & Moorer, 1977), extrapolation beyond an instrument register,[6] "perversion" of additive synthesis to produce sound paradoxes and illusions (Shepard, 1964; Risset, 1971, 1978a,b,c; Deutsch, 1975; Wessel & Risset, 1979), or transformation of percussive sounds into fluid textures while preserving their frequency content (see p. 39 above). The extension of the models can thus lead to the synthesis of interesting nonconventional timbres, which is a fascinating area open to musicians.

## XIV. TIMBRAL SPACE

We have discussed perceptual models; we have also said that analysis-synthesis is useful in modifying timbres. In this respect it would be useful to have a good notion of the structure of the perceptual relationship between different timbres. This can be greatly eased by geometrical models provided by multidimensional techniques, which in effect provide displays of this structure. As was stated earlier by one of the authors: "A timbre space that adequately represented the perceptual dissimilarities could conceivably serve as a kind of map that would provide navigational advice to the composer interested in structuring aspects of timbre (Wessel, 1973)."

[6]This can be heard in *Studies for Trumpet and Computer* by Morrill (recording by Marice Stith, Golden Crest, recital series RE-7068).

One can indeed propose geometric models of subjective timbral space such that individual sounds are represented as points in this space: sounds judged very dissimilar are distant, and sounds judged similar are close. The models are not constructed arbitrarily, but by asking subjects to rate for many pairs of sounds the dissimilarities between the sounds of each pair and by submitting the dissimilarity data to multidimensional scaling programs. These programs—strictly devoid of preconceptions about the data—provide a geometrical model that best fits these data. The dimensions of the model can then be interpreted (e.g., by investigating the stimuli that are least, or most, differentiated along these dimensions).[7] Wessel (1973, 1978) and Grey (1975) have thus already provided models of timbral space for string and wind instrument tones. These models unveil two dimensions—one that differs within the instruments of a same family (e.g., cello, viola, violin) and which appears to relate to the spectral distribution of energy in the sound (cf. von Bismarck, 1974); and one that is the same within a family of instruments and which seems to be linked to temporal features like the details of the attack (Wessel, 1973; Grey, 1977).

The corresponding representations of timbral space tempt one to fill the space, to draw trajectories through it, like the timbral interpolations mentioned above. According to Grey (1975), "The scaling for sets of naturalistic tones suggest a hybrid space, where some dimensions are based on low-level perceptual distinctions made with respect to obvious physical properties of tones, while other dimensions can be explained only on the basis of a higher level distinction, like musical instrument families." The intervention of cognitive facets, such as familiarity and recognition, indicates that a fully continuous timbre space may not be obtainable. Nevertheless, subjective space models can propose new paths and new intriguing concepts, such as that of analogies between timbral transitions (Wessel, 1979), which may permit one to do with timbres something similar to melodic transposition with pitches. Resolving timbre, that "attribute" defined as neither pitch nor loudness, into dimensions may uncover new features or parameters susceptible to precise differentiation and appropriate for articulating musical structures.

For instance, multidimensional scaling of timbre often unveils a dimension correlated with the spectral energy distribution, hence with that aspect of timbre termed *brightness*. As Wessel (1979) has shown, this dimension is the one that can best articulate stream segregation (McAdams & Bregman, 1979). Here, isolating dimensions of timbres permits one to make predictions about the behavior of these timbres in context.

The timbre-space representation suggests relatively straightforward schemes for controlling timbre. The basic idea is that by specifying coordinates in a particular timbre space, one could hear the timbre represented by those coordinates. If these coordinates should fall between existing tones in the space, we would want this interpolated timbre to relate to the other sounds in a manner consistent with the

[7]From only quantitative judgments of dissimilarities between sounds, multidimensional scaling in effect unveils in what ways these sounds differ. Schaeffer failed to realize that in his criticism of the process as described by Babbitt (1965) [cf. Music and Technology (1971), pp. 77–78].

structure of the space. Evidence that such interpolated sounds are consistent with the geometry of the space has been provided by Grey (1975). Grey used selected pairs of sounds from his timbre space and formed sequences of interpolated sounds by modifying the envelope break points of the two sounds with a simple linear interpolation scheme. These interpolated sequences of sounds were perceptually smooth and did not exhibit abrupt changes in timbre. Members of the original set of sounds and the newly created interpolated timbres were then used in a dissimilarity judgment experiment to determine a new timbre space. This new space had essentially the same structure as the original space with the interpolated tones appropriately located between the sounds used to construct them. It would appear from these results that the regions between the existing sounds in the space can be filled out, and that smooth, finely graded timbral transitions can be formed.

The most natural way to move about in timbral space would be to attach the handles of control directly to the dimensions of the space. One of the authors examined such a control scheme in a real-time context (Wessel, 1979). A two-dimensional timbre space was represented on the graphics terminal of the computer that controlled the Di Giugno oscillator bank at I.R.C.A.M. One dimension of this space was used to manipulate the shape of the spectral energy distribution. This was accomplished by appropriately scaling the line-segment amplitude envelopes according to a shaping function. The other axis of the space was used to control either the attack rate or the extent of synchronicity among the various components. Overall, the timbral trajectories in these spaces were smooth and otherwise perceptually well behaved. To facilitate more complex forms of control, we need an efficient computer language for dealing with envelopes. The basic idea behind such a language is to provide a flexible control structure that permits specification, sequencing, and combination of various procedures that create and modify envelopes. These procedures would include operations like stretching or shortening duration, changing pitch, reshaping spectrum, synchronizing or desynchronizing spectral components, and so forth. With such a language it will be possible to tie the operations on the envelope collections directly to the properties of the perceptual representations of the material.

## XV. CONCLUSION

As was explained above, the exploration of timbre by analysis and synthesis can serve several purposes: it provides insight into the physical parameters of the sound and the relevance of these parameters to the resulting timbre; it leads to simplified models that permit data reduction in the synthetic replication of the sound; and it uses models to perform transformations on the original sound, either from the point of view of classical musical processing (for instance, by independently changing pitch, duration, articulation, and loudness) or by expanding timbral resources (rearranging at will the complex variations abstracted from the analysis to obtain new and rich sounds).

Exploration of timbre by analysis and synthesis is difficult but rewarding. Since the

development of analysis and synthesis devices, in particular the digital computer and its descendants, it has brought a better understanding of the physical correlates of timbre as well as recipes for new musical resources.

Although much remains to be done, these new possibilities available to musicians will probably increase the musical role of timbre. In classical Western music timbres were used mostly to differentiate musical lines. Later this linear organization was disrupted by Debussy, Varése, and others. Schoenberg's *Klangfarbenmelodie* associated different timbres to successive notes; Boulez submitted the succession of timbres to serial organization. With the control of timbre now made possible through analysis and synthesis, composers can compose not only with timbres, but they can also compose timbres: they can articulate musical compositions on the basis of timbral rather on pitch variations. It has been argued that timbre perception is too vague to form the basis of elaborate musical communication; however, as Mathews has remarked, there already exists an instance of a sophisticated communication system based on timbral differentiation; namely human speech.[8] Hence, it is conceivable that proper timbral control might lead to quite new musical architectures, and analysis and synthesis processes will be of much help in providing a great musical potential. This will require the acute ear of the musician, provided with some psychoacoustic know-how and with a good interactive environment that permits him or her to achieve fine timbre tunings and manipulations. Hence, the exploration of timbre by analysis and synthesis may become a basic musical endeavor.

## APPENDICES

### A. Signal Representations and Analysis–Synthesis Processes

Analysis–synthesis according to a given process implies estimating the parameters of a model of the sound. This model may or may not be adequate; it may or may not lend itself to a good imitation of the sound. For instance, Fourier series expansion is a useful tool for periodic tones, and Fourier synthesis, using the data of Fourier analysis, indeed permits one to synthesize a faithful copy of a periodic sound. However, as was explained above, most sounds of interest are not periodic; hence, Fourier series expansion is inadequate to replicate, for instance, a sound whose spectrum varies with time.

A sound can be mathematically described by the waveshape function $p(t)$, giving the acoustic pressure as a function of time. Mathematics tells us that reasonably regular functions can be analyzed in a number of ways, that is, in terms of one or another set of basic functions. This set is said to be complete if an arbitrary function can indeed be obtained as the proper linear combination of these basic functions. (This proper combination is unveiled by the analysis process that consists of estimat-

[8]As Moorer demonstrated by analysis and synthesis, speech can remain intelligible under certain conditions after removal of pitch and rhythmic information.

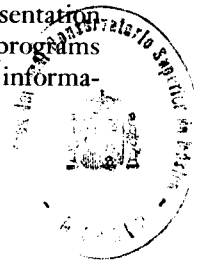
ing the parameters of the corresponding model.) For instance, Fourier's theorem states that any periodic function (of frequency *f*) can be expanded as a linear combination of the sine and cosine functions of frequencies *f, 2f, 3f,* . . ., so that this linear combination can be arbitrarily close to the periodic function. Hence, the set of sine and cosine functions of frequencies *f, 2f, 3f, etc.* is "complete" over the space of periodic functions of frequency *f* (cf. Panter, 1965; Rosenblatt, 1963).

Actually, the representation of nonperiodic signals in terms of basic functions usually requires an infinite number of basic functions so that the series expansion turns into a transformation. For instance, nonperiodic signals can be represented in terms of the so-called Fourier transform or Fourier integral, in which the discrete spectral components are replaced by a continuous amplitude spectrum; the discrete phases are also replaced by a phase spectrum. There are other transformations used for analysis-synthesis (e.g., the Walsh-Hadamard and the Karhunen-Loève transformations). Such linear expansion in terms of a basic set of signals is similar to the expansion of a vector in terms of a set of basic vectors; it is practical to use orthogonal transforms—that is, to use functions that form an orthonormal (and complete) set (cf. Harmuth, 1972).

The application of a given transform to a sound signal provides a representation of the signal that may be revealing and should make it possible to restore the signal by means of the inverse transform. Hence, representation of signals is closely linked to analysis-synthesis processes. Actually, the representation of signals purports both to characterize the information (bearing elements in the signal) and to describe in a simple way the effect of modifications of the signals (like those introduced by an imperfect transmission system or by a deliberate simplification of the signal).

Although we cannot go into much detail here, we would like to make several points:

1. Some analysis-synthesis processes and the corresponding representation are intrinsically limited to certain classes of signals. Others can be transparent if they are complete in the above sense—for instance, the Fourier or the Hadamard transform, the phase vocoder, the linear predictive coding scheme. However, the two latter schemes will permit one to reproduce the original signal only at the expense of a considerably detailed analysis, an information explosion instead of an information reduction. This can only be substantially simplified for certain classes of signals (quasi-periodic signals with relatively independent excitation and response mechanisms, like speech; for instance, linear predictive coding is efficient in simulating oboe sounds but poor for low clarinet sounds because eliminating the even harmonics is taxing for the filter). Indeed, much work on analysis-synthesis and signal transformation was originally directed toward efficient coding of speech information for economical transmission over technical channels (Flanagan, 1972; Campanella & Robinson, 1971; Schafer & Rabiner, 1975). It is also for certain types of signals that the representation of the signal will be most enlightening [but, for instance, phase vocoders' programs implemented by Moorer (1978) have permitted Castellengo to obtain useful information on nonharmonic "multiphonic" tones].



Similarly, Gabor's expansion of a signal into Gaussian elementary signals has been proven to be complete (Bastiaans, 1980). Hence, it can in principle produce exactly what Fourier or other types of synthesis can produce (cf. Gabor, 1947; Xenakis, 1971; Roads, 1978). The idiosyncrasies of different complete analysis-synthesis methods only appear in what they permit—or suggest—in a simplified, archetypal use.

2. The Walsh-Hadamard transform seems promising because it leads to operations that are easy to implement with digital circuits. However, from a psychoacoustical standpoint this transform is quite inappropriate. The basic functions do not sound elemental to the ear; they are spectrally very rich, and an approximated representation in those terms would lead to aurally unsatisfying results. The analysis-synthesis process does not deteriorate gracefully for the ear, and it has great difficulty in producing timbres that are not rich and harsh (for instance, it has trouble approaching a sine wave).

3. Fourier-type analysis (and synthesis) has been much criticized, often in a poorly documented way. Whereas Fourier series expansion is indeed inadequate for non-periodic sounds, there are more elaborate variants of Fourier analysis of great utility. The Fourier transform provides complete information of an amplitude spectrum and a phase spectrum; however, the latter characterizes the evolution of the signal in time in a way that is unintuitive and very hard to use. Since this evolution in time is very significant to the ear, one needs some kind of running analysis. This is obtained by calculating, as a function of time, the spectrum of the signal viewed through a specified *time-window* (also called *weighting function*), which at any time only shows the most recent part of the past values of the signal. Such representations are very useful: they have been used in several of the studies previously described. The sound spectrograph (Koenig, Dunn, & Lacey, 1946) implements this type of running analysis: Its windows are appropriate for a useful portrayal of speech sounds, but it often displays significant features of music as well (Leipp, 1971), even though the analysis is often too crude to provide data for a proper synthesis.

The significance of Fourier analysis has a multiple basis. There is clear evidence that the peripheral stages of hearing, through the mechanical filtering action of the basilar membrane, perform a crude frequency analysis with a resolution linked to the critical bandwidth (Plomp, 1964; Flanagan, 1972). The distribution of activity along the basilar membrane relates simply to the Fourier spectrum. Also, when the sound is quasi-periodic, the phase deafness of the ear (Fig. 1) permits a substantial reduction of information. One can also in this case take advantage of the concentration of energy at the harmonic frequencies to describe the sounds by the evolution in time of the amplitude of few harmonics. We have seen that such additive synthesis was a very useful model (cf. Risset & Mathews, 1969; Keeler, 1972; Grey & Moorer, 1977).

4. Global methods like frequency-modulation (Chowning, 1973) and nonlinear distortion or waveshaping (Arfib, 1979; Le Brun, 1979) are appealing because they provide powerful control over salient features of the spectrum in terms of few parameters: the amount of specification and processing is much reduced as compared with additive synthesis. However, strength is at the expense of generality. It is difficult with the global methods to achieve certain results unless one uses them in refined ways that can quickly become complex (cf. Schottstaedt, 1977).

## B. Synthesis Model Based on Perceptual Principles

We now give a brief account of how synthesis procedures can provide for direct control over some essential perceptual attributes of timbre. The essential principle underlying these synthesis schemes is the decomposition of a musical signal into perceptual attributes that are, for the most part, perceptually independent of each other. The motivation is to provide a reasonably general but simple control scheme for additive synthesis as this form of synthesis is becoming more and more practical with advances in the development of high-speed digital-synthesis hardware.

### *1. Pitch Versus the Global Impression of the Spectral Envelope*

Several studies (Plomp & Steeneken, 1971; Risset, 1978b,c) suggest that musical pitch and the global spectral energy distribution as perceptual attributes are reasonably independent of each other. This is true to a large extent for harmonic tones that tend to produce clear pitch percepts, but it is not true for inharmonic spectra whose ambiguous and otherwise multiple-pitch content depends on the spectral balance of the components. What we mean by independence is that it is possible to manipulate, for example, the placement and shape of formants without influencing the perceived pitch and conversely manipulate the pitch while keeping the perceived shape of the spectrum constant. The voice provides an example of such an independent control scheme that operates over a reasonably wide range of pitches and spectral shapes. A singer can sing the same pitch with a large variety of vowel qualities and can likewise maintain a constant vowel quality over a substantial range of pitches.

### *2. Roughness and Other Spectral Line-Widening Effects*

Terhardt (1978) has provided evidence that our impression of roughness in sounds depends on an additive combination of independent spectrally distributed amplitude fluctuations. Consider the following example using a tone consisting of the three components: 400, 800, and 1600 Hz. Here the components are widely distributed (i.e., more than a critical bandwidth between them) and amplitude fluctuations of say 10% of the component amplitude at frequencies between 10 and 35 Hz contribute independently to the overall impression of roughness. The implication for synthesis is to provide for independent control of the amplitude fluctations in different regions of the spectrum.

By *spectral line widening* we mean the spreading or smearing of energy around a spectral line. Such spectral line widening can be obtained by amplitude and/or frequency modulation of a sinusoid. Many instrumental timbres have noise-like effects in their attack transients and most often their spectral placement is essential to the timbre. For example, in the synthesis of string-like attacks the middle to upper spectral regions require more noise than the lower regions. It is to the synthesis model's advantage to allow for the independent placement of noise-like effects in separate spectral regions, which can be accomplished by widening the spectral lines in those regions.

### *3. Vibrato and Frequency Glides*

Our impression of timbre is often strongly dependent on the presence of a vibrato or frequency glide and the synthesis procedure should provide for an easy application of these effects without disrupting the global spectral energy distribution. A frequency glide of an oscillator with a fixed spectrum results as well in a glide of the spectral energy distribution and thus violates the desired independence. Such independence has been accomplished in the glissando version of Shepard's illusion (Shepard, 1964) produced by Risset (1978b,c).

In our additive synthesis procedure we should be able to provide an overall spectral envelope that remains constant in spite of changes in the specific frequencies of the components. In addition, the model should provide for the independent placement of roughness and noise-like effects in separate regions of the spectrum again without violating the overall spectral envelope. These kinds of control can be accomplished fairly easily in most sound synthesis languages by the use of table-look-up generators, such as the VFMULT of the MUSIC V language. These generators allow one to store a spectral envelope function that is used to determine the sample-by-sample amplitude of a given component that could be executing a frequency glide. This technique works similarly for control of the spectral distribution of roughness or other line-widening effects. To obtain time-variant effects with these attributes, the spectral envelopes and roughness distributions are defined at successive and often closely spaced points in time, and interpolation is carried out between successive pairs of these functions.

## REFERENCES


Allen, J. B., & Rabiner, L. R. A unified approach to short-time Fourier analysis and synthesis. *Proceedings of the IEEE*, 1977, *65*, 1558–1564.

Alles, H. G., & Di Giugno, P. A one-card 64-channel digital synthesizer. *Computer Music Journal*, 1977, *1* (n°4), 7–9.

American Standard Association. American Standard Acoustical Terminology, New York, 1960.

Appleton, J. H., & Perera, R. C., editors, *The development and practice of electronic music*. Prentice Hall, Englewood Cliffs, New Jersey, 1975.

Arfib, D. Digital synthesis of complex spectra by means of multiplication of non-linear distorted sine waves, *Journal of the Audio Engineering Society*, 1979, *27*, 757–768.

Babbitt, M. The use of computers in musicological research. *Perspectives of new music*, 1965, *3*, n° 2.

Backhaus, W. Einschwingvorgänge, *Zeitschrift für Technische Physik*, 1932, *13*, 31.

Backus, J. *The acoustical foundations of music*. Norton, New York, 1969.

Backus, J., & Hundley, J. C. Harmonic generation in the trumpet, *Journal of the Acoustical Society of America*, 1971, *49*, 509–519.

Bastiaans, M. J. Gabor's expansion of a signal into Gaussian elementary signals, *Proceedings of the IEEE*, 1980, *68*, 538–539.

Beauchamp, J. W. Analysis and synthesis of cornet tones using non linear interharmonic relationships. *Journal of the Audio Engineering Society*, 1975, *23*, 778–795.

Benade, A. H. *Fundamentals of musical acoustics*. London and New York: Oxford University Press, 1976.

Berger, K. W. Some factors in the recognition of timbre, *Journal of the Acoustical Society of America*, 1964, *36*, 1888–1891.

Blackman, R. B., & Tukey, J. W. *The measurement of power spectra from the point of view of communications engineering*. Dover, New York, 1958.

Bolt, R. H., Cooper, F. S., David, E. E. Jr., Denes, P. B., Pickett, J. M., & Stevens, K. N. Identification of a speaker by speech spectrogram. *Science*, 1969, *166*, 338-343.

Bolt, R. H., Cooper, F. S., David, E. E. Jr., Denes, P. B., Pickett, J. M., & Stevens, K. N. *On the theory and practice of voice identification*. National Research Council Washington, D.C., 1978

Boomsliter, P. C., & Creel, W. Hearing with ears instead of instruments. *Journal of the Audio Engineering Society*, 1970, *18*, 407-412.

Bregman, A. S., & Campbell, J. Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 1971, *89*, 244-249.

Bregman, A. S., & Pinker, S. Auditory streaming and the building of timbre. *Canadian Journal of Psychology*, 1978, *32*, 19-31.

Cabot, R. C., Mino, M. G., Dorans, D. A., Tackel, I. S., & Breed, H. E. Detection of phase shifts in harmonically related tones. *Journal of the Audio Engineering Society*, 1976, *24*, 568-571.

Cadoz, C., Luciani, A., Florens, S. L. Synthèse musicale par simulation des méchanismes instrumentaux et transducteurs gesteuls retroactifs pour l'étude du jeu instrumental. *Revue d'Acoustique*, 1981, *59*, 279-292.

Campanella, S. J., & Robinson, G. S. A comparison of orthogonal transformations for digital speech processing. *IEEE Transactions on Communication Technology*, 1971, COM-19, 1045-1050.

Charbonneau, G. Effets perceptifs de la réduction des données dans la perception du timbre. *C. R. Acad. Sci., Paris*, 1979, 289 B, 147-149.

Chowning, J. The synthesis of complex audio spectra by means of frequency modulation. *Journal of the Audio Engineering Society*, 1973, *21*, 526-534.

Clark, M. Robertson, P., & Luce, D. A preliminary experiment on the perceptual basis for musical instrument families. *Journal of the Audio Engineering Society*, 1964, *12*, 194-203.

Culver, C. A. *Musical Acoustics*. New York: McGraw Hill, 1956.

Deutsch, D. Musical illusions, *Scientific American*, 1975, *233*, 92-104.

Dijksterhuis, P. R., & Verhey, T. An electronic piano. *Journal of the Audio Engineering Society*, 1969, *17*, 266-271.

Eagleson, H. W., & Eagleson, O. W. Identification of musical instruments when heard directly and over a public address system. *Journal of the Acoustical Society of America*, 1947, *19*, 338-342.

Erickson, R. *Sound structure in music*. Berkeley, California: University of California Press, 1975.

Fant, G. *Acoustic theory of speech production*. Gravenhage: Mouton, 1960.

Feynman, R. B., Leighton, R. B., & Sands, M. *The Feynman lectures on physics*, Reading, Massachusetts: Addison-Wesley, 1963

Flanagan, J. L. *Speech analysis, synthesis and perception*. New York: Academic Press, 1972.

Fletcher, H., & Bassett, I. G. Some experiments with the bass drum. *Journal of the Acoustical Society of America*, 1978, *64*, 1570-1576.

Fletcher, H., Blackham, E. D., & Christensen, D. A. Quality of organ tones. *Journal of the Acoustical Society of America*, 1963, *35*, 314-325.

Fletcher, H., Blackham, E. D., & Stratton, R. Quality of piano tones. *Journal of the Acoustical Society of America*, 1962, *34*, 749-761.

Fletcher, H., & Sanders, L. C. Quality of violin vibrato tones. *Journal of the Acoustical Society of America*, 1967, *41*, 1534-1544.

Freedman, M. D. Analysis of musical instrument tones. *Journal of the Acoustical Society of America*, 1967, *41*, 793-806.

Gabor, D. Acoustical quanta and the nature of hearing. *Nature (London)*, 1947, *159*, No. 4.

George, W. H. A sound reversal technique applied to the study of tone quality. *Acustica*, 1954, *4*, 224-225.

Grey, J. M. An exploration of musical timbre. Thesis, Stanford University, 1975.

Grey, J. M. Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America*, 1977, *61*, 1270-1277.

Grey, J. M. Timbre discrimination in musical patterns. *Journal of the Acoustical Society of America*, 1978, *64*, 467-472.

Grey, J. M., & Gordon, J. W. Perceptual effect of spectral modifications in musical timbres. *Journal of the Acoustical Society of America*, 1978, *63*, 1493-1500.

Grey, J. M., & Moorer, J. A. Perceptual evaluation of synthesized musical instrument tones. *Journal of the Acoustical Society of America*, 1977, *62*, 454-462.

Hall, H. H. Sound analysis. *Journal of the Acoustical Society of America*, 1937, *8*, 257-262.

Harmuth, H. *Transmission of information by orthogonal functions*. New-York. Springer, 1972.

Hiller, L. & Ruiz, P. Synthesizing musical sounds by solving the wave equation for vibrating objects—Part I. *Journal of the Audio Engineering Society*, 1971, *19*, 463-470.

Huggins, W. H. A phase principle for complex frequency analysis and its implication in auditory theory. *Journal of the Acoustical Society of America*, 1952, *24*, 582-589.

Jansson, E., & Sundberg, J. Long-term average spectra applied to analysis of music. *Acustica*, 1975/1976, *34*, 15-19, 269-274.

Jenkins, G. M., & Watts, D. G. *Spectral analysis and its applications*. San Francisco, California: Holden-Day, 1968.

Keeler, J. S. Piecewise-periodic analysis of almost-periodic sounds and musical transients. *IEEE Transactions on Audio & Electroacoustics*, 1972, *AU-ZO*, 338-344.

Koenig, W., Dunn, H. K., & Lacey, L. Y. The sound spectrograph. *Journal of the Acoustical Society of America*, 1946, *18*, 19-49.

Köhler, W. Akustische Untersuchungen. *Zeitschrift für Psychologie*, 1915, 72, 159.

Kohut, J., & Mathews, M. V. Study of motion of a bowed string. *Journal of the Acoustical Society of America*, 1971, *49*, 532-537.

Lashley, K. S. An examination of the "continuity theory" as applied to discriminative learning. *Journal of General Psychology*. 1942, *26*, 241-265.

Le Brun, M. Digital waveshaping synthesis. *Journal of the Audio Engineering Society*, 1979, 27, 250-266.

Leipp, E. *Acoustique et musique*. Paris: Masson, 1971.

Luce, D., & Clark, M. Jr., Physical correlates of brass-instrument tones. *Journal of the Acoustical Society of America*, 1967, *42*, 1232-1243.

Martin, D. W. Decay rates of piano tones. *Journal of the Acoustical Society of America*, 1947, *19*, 535.

Mathes, R. C., & Miller, R. L. Phase effects in monaural perception. *Journal of the Acoustical Society of America*, 1947, *19*, 780-797.

Mathews, M. V. The digital computer as a musical instrument. *Science*, 1963, *142*, 553-557.

Mathews, M. V. *The technology of computer music*. Cambridge, Massachusetts: MIT Press, 1969.

Mathews, M. V., & Kohut, J. Electronic simulation of violin resonances. *Journal of the Acoustical Society of America*, 1973, *53*, 1620-1626.

Mathews, M. V., Miller, J. E., & David, E. E. Jr., Pitch synchronous analysis of voiced sounds, *Journal of the Acoustical Society of America*, 1961, *33*, 179-186.

Mathews, M. V., Miller, J. E., Pierce, J. R., & Tenney, J. *Journal of the Acoustical Society of America*, 1965, *38*, 912. (abstract only)

Mathews, M. V., Miller, J. E., Pierce, J. R., & Tenney, J. Computer study of violin tones. Murray Hill, New Jersey: Bell Laboratories, 1966.

Mathews, M. V., & Moore, F. R. Groove—a program to compose, store and edit functions of time. *Communications of the ACM*, 1970, *13*, 715-721.

Mathews, M. V., Moore, F. R., & Risset, J. C. Computers and future music. *Science*, 1974, *183*, 263-268.

McAdams, S. Spectral fusion and the creation of auditory images. In M. Clynes (Ed.) *Music, mind and brain: The neuropsychology of music*. Plenum: New York, 1982.

McAdams, S., & Bregman, A. Hearing musical streams. *Computer Music Journal*, 1979, *3*, No. 4, 26-43.

Meyer, E., & Buchmann, G. *Die Klangspektren der Musikinstrumente*. Berlin: 1931

Miller, D. C. *The Science of musical sounds*. New York: Mac Millan, 1926.

Moorer, J. A. Signal processing aspects of computer music: a survey. *Proceedings of the IEEE*, 1977, *65*, 1108-1137.

Moorer, J. A. The use of the phase vocoder in computer music applications. *Journal of the Audio Engineering Society*, 1978, *26*, 42-45.

Moorer, J. A., & Grey, J. Lexicon of analyzed tones.-Part I: A violin tone. *Computer Music Journal*, 1977, *1*, No. 2, 39-45. (a)

Moorer, J. A., & Grey, J. Part II: Clarinet and oboe tones. *Computer Music Journal*, 1977, *1*, No. 3, 12-29. (b)

Moorer, J. A., & Grey, J. Part III: The trumpet, *Computer Music Journal*, 1978, *2*, No. 2, 23-31.

Morrill, D. Trumpet algorithms for music composition. *Computer Music Journal*, 1977, *1*, No. 1, 46-52.

Music and Technology. UNESCO and *Revue Musicale*. Paris, 1971.

Olson, H. F. *Music, physics and engineering*. New York: Dover, 1967.

Panter, P. F. *Modulation, noise and spectral analysis*. New York: McGraw Hill, 1965.

Plomp, R. The ear as a frequency analyzer. *Journal of the Acoustical Society of America*, 1964, *36*, 1628-1636.

Plomp, R. *Aspects of tone sensation*. New York: Academic Press, 1976.

Plomp, R. Timbre as a multidimensional attribute of complex tones. In R. Plomp & F. G. Smoorenburg (Eds.) *Frequency analysis and periodicity detection in hearing*. Leiden: Suithoff, 1966.

Plomp, R., & Steeneken, J. M. Effect of phase on the timbre of complex tones. *Journal of the Acoustical Society of America*, 1969, *46*, 409-421.

Plomp, R., & Steeneken, J. M. Pitch versus timbre. *Proceeding of the 7th International Congress of Acoustics, Budapest*, 1971, *3*, 377-380.

Richardson, E. G. The transient tones of wind instruments. *Journal of the Acoustical Society of America*, 1954, *26*, 960-962.

Risset, J. C. *Journal of the Acoustical Society of America*, 1965, *38*, 912 (abstract only).

Risset, J. C. Computer study of trumpet tones, Bell Laboratories, Murray Hill, New Jersey, 1966.

Risset, J. C. *An introductory catalog of computer-synthesized sounds*, Bell Laboratories, Murray Hill, New Jersey, 1969.

Risset, J. C. Paradoxes de hauteur, *Proceedings of the 7th International Congress of Acoustics, Budapest*, 1971, 20 S 10.

Risset, J. C. Musical acoustics. In E. C. Carterette & M. P. Friedman, *Handbook of perception. (Volume IV, Hearing)* New York: Academic Press, 1978. Pp. 521-564. (a)

Risset, J. C. Paradoxes de hauteur. IRCAM Report No. 10, Paris, 1978 (with a cassette of sound examples). (b)

Risset, J. C. Hauteur et timbre. IRCAM Report No. 11, Paris, 1978. (with a cassette of sound examples). (c)

Risset, J. C., & Mathews, M. V. Analysis of musical instrument tones. *Physics Today*, 1969, *22*, No. 2, 23-30.

Roads, C. Automated granular synthesis of sound. *Computer Music Journal* 1978, *2*, No. 2, 61-62.

Rodet, X., & Bennett, G. Synthèse de la voix chanteé pour ordinateur. *Conferences des Journées d'Etudes du Festival du son, Paris*, 1980. Pp. 73-91.

Rodet, X. Time-domain formant-wave-function synthesis. *Proceedings of the NATO-ASI Meeting Bonas*, July 1979.

Roederer, J. G. *Introduction to the Physics and Psychophysics of Music*. London: The English Universities Press, 1974.

Rosenblatt, M. (Ed.), *Time series analysis*. New York: Wiley, 1963.

Saldanha, E. L., & Corso, J. F. Timbre cues and the identification of musical instruments. *Journal of the Acoustical Society of America*, 1964, *36*, 2021-2026.

Schafer, R. W., & Rabiner, L. R. Digital representations of speech signals. *Proceedings of the IEEE*, 1975, *63*, 662-677.

Schaeffer, P. *Traité des objets musicaux*. Paris: Ed. du Seuil, 1966. (with three records of sound examples)

Schottstaedt, W. The simulation of natural instrument tones using frequency modulation with a complex modulating wave. *Computer Music Journal*, 1977, *1*, No. 4, 46-50.

Schroeder, M. R. Complementarity of sound buildup and decay. *Journal of the Acoustical Society of America*, 1966, *40*, 549-551.

Schroeder, M. R. Models of hearing. *Proceedings of the Institute of Electrical and Electronics Engineers*, 1975, *63*, 1332-1350.

Conservatorio Superiore

Shepard, R. N. Circularity of relative pitch. *Journal of the Acoustical Society of America*, 1964, *36*, 2346–2353.

Slawson, A. W. Vowel quality and musical timbre as functions of spectrum envelope and fundamental frequency. *Journal of the Acoustical Society of America*, 1968, *43*, 87–101.

Strong, W., & Clark, M. Jr., Synthesis of wind-instrument tones. *Journal of the Acoustical Society of America*, 1967, *41*, 39–52. (a)

Strong, W., & Clark, M. Jr., Perturbations of synthetic orchestral wind instrument tones. *Journal of the Acoustical Society of America*, 1967, *41*, 277–285. (b)

Stumpf, C. *Die sprachlante*. Berlin and New York: Springer-Verlag, 1926.

Sundberg, J. The acoustics of the singing voice. *Scientific American*, 1977, *236*, 82–91.

Tenney, J. C. The physical correlates of timbre. *Gravesaner Blätter*, 7, Heft 26, 106–109.

Terhardt, E. Psychoacoustic evaluation of musical sounds. *Perception & Psychophysics*, 1978, *23*, 483–492.

Van Noorden, L. *Temporal coherence in the perception of tone sequences*. Eindhoven, Holland: Instituut voor Perceptie Onderzoek, 1975.

von Bismarck, G. Sharpness as an attribute of the timbre of steady sounds *Acustica*, 1974, *30*, 159–172.

von Helmholtz, H. *Sensations of tone*. 1877. English translation with notes and appendix by E. J. Ellis, New York: Dover, 1954.

Weinreich, G. Coupled piano strings. *Journal of the Acoustical Society of America*, 1977, *62*, 1474–1484.

Weinreich, G. The coupled motions of piano strings, *Scientific American*, 1979, *240*, No. 1, 118–127.

Wente, E. C. Characteristics of sound transmission in rooms. *Journal of the Acoustical Society of America*, 1935, 7, 123.

Wessel, D. L. Psychoacoustics and music: A report from Michigan State University. *Bulletin of the Computer Arts Society*, 1973, *30*.

Wessel D. L. Low dimensional control of musical timbre, IRCAM Report Paris No. 12, 1978 (with a cassette of sound examples).

Wessel, D. I. Timbre space as a musical control structure. *Computer Music Journal*, 1979, *3*, No. 2, 45–52.

Wessel, D. L., & Risset, J. C. *Les illusions auditives*. Universalia: Encyclopedia Universalis, 1979. Pp. 167–171.

Winckel, F. *Music, sound and sensation*. New York: Dover, 1967.

Xenakis, I. *Formalized music*. Bloomington, Indiana: Indiana University Press, 1971.

Young, R. W. Modes, nodes and antinodes. *Journal of the Acoustical Society of America*, 1952, *24*, 267–273.

Zwicker, E. Subdivision of the audible frequency range into critical bands. *Journal of the Acoustical Society of America*, 1961, *33*, 248.

Zwicker, E., & Scharf, B. A model of loudness summation. *Psychological Review*, 1965, *72*, 3–26.

# 3

# Perception of Singing

*Johan Sundberg*


| | | |
|---|---|---|
| I. | Introduction | 59 |
| II. | Function of the Voice | 60 |
| III. | Resonatory Aspects | 62 |
| | A. Female Singing | 62 |
| | B. Male Singing | 69 |
| | C. Voice Classification | 73 |
| IV. | Phonation | 76 |
| | A. Vocal Effort and Pitch | 77 |
| | B. Register | 78 |
| V. | Vibrato | 82 |
| | A. Background | 82 |
| | B. Perceptual Aspects | 83 |
| VI. | Pitch Accuracy in Singing Practice | 89 |
| VII. | Phrasing and Emotion | 91 |
| VIII. | Concluding Remarks | 94 |
| | References | 95 |


## I. INTRODUCTION

An understanding of singing perception may be developed from two different types of investigation. One type considers the acoustic property of singing, varies it systematically, and explores the perception of these variations. Such investigations are rare in singing research. Another type of investigation concerns the acoustic correlates of certain types of voices or phonations. The underlying typology of voices and phonations must be based mainly on aural perception. Consequently, even such investigations have perceptual relevance. Most of the investigations of singing have this type of perceptual relevance. However, research on the perception of singing is


Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

not as developed as is the closely related field of speech research. Therefore, the reader will not find an exhaustive presentation in this chapter. Rather he or she will find a presentation of different investigations only partly related to each other.

A confession on the part of the author of this chapter is in place here. Writing this chapter has been a bit embarassing, because so many references to the author's own work seemed motivated. As the reader might notice, it is difficult to be fully objective in the presentation of one's own investigations. However, as was just mentioned, there are comparatively few studies published on the perceptual aspects of singing.

When we listen to a singer there are a number of remarkable perceptual facts. For instance: How is it that we can hear the voice even when the orchestra is loud? How is it that we generally identify the singer's vowels correctly even though vowel quality in singing differs considerably from that which we are used to in speech? How is it that we can identify the individual singer's sex, register, and voice timbre when the pitch of the vowel lies within a range common to all singers and several registers? How is it that we perceive singing as a sequence of discrete pitches even though the fundamental frequency events do not form a pattern of discrete fundamental frequencies? How is it that we hear a phrase rather than a sequence of isolated tones? These are some of the main questions that will be discussed in the present chapter. In order to understand the questions as well as the answers, it is necessary to have a basic knowledge of the acoustics of singing. We will therefore first briefly present what is known about this.

## II. FUNCTION OF THE VOICE

The vocal organ consists of three basic components: (1) the respiratory system that provides an excess pressure of air in the lungs, (2) the vocal folds that chop the air stream from the lungs into a sequence of quasi-periodic air pulses, and (3) the vocal tract that gives each sound its final characteristic spectral shape and thus its timbral identity. These three components are referred to as respiration, phonation, and resonance or or articulation, respectively.

The chopped air stream (i.e., the *voice source*) is the raw material of all voiced sounds. It can be described as a complex tone composed of a number of harmonic partials. This implies that the frequency of the $n$th partial equals $n$ times the frequency of the first partial, which is called the *fundamental.* The frequency of the fundamental (i.e., the fundamental frequency) is identical to the number of air pulses occurring in one second or, in other words, to the frequency of vibration of the vocal folds. The fundamental frequency determines the pitch we perceive in the sense that the pitch would remain essentially the same even if the fundamental sounded alone. The amplitudes of the voice-source partials decrease monotonically with rising frequency. As a rule of thumb, a given partial is 12 dB stronger than a partial located one octave higher. However, for low degrees of vocal effort the slope of this source spectrum is steeper than 12 dB/octave. On the other hand, the slope of the voice-source spectrum is generally not dependent on which voiced sound is produced.

Spectral differences between various voiced sounds arise when the sound from the voice source is transferred through the vocal tract (i.e., from the vocal folds to the lip opening). The reason for this is that the ability of the vocal tract to transfer sound is highly dependent on the frequency of the sound being transferred. This ability culminates at certain frequencies, called the *formant frequencies.* In consequence, those voice-source partials that lie closest to the formant frequencies are radiated from the lip opening at greater amplitudes than the other neighboring partials. Hence, the formant frequencies are manifest as peaks in the spectrum of the radiated sound.

The formant frequencies vary within rather wide limits in response to changing the position of the articulators (i.e., lips, tongue body, tongue tip, lower jaw, velum, and larynx). We can change the two lowest formant frequencies by two octaves or more by changing the position of the articulators. The frequencies of these two formants determine the identity of most vowels. The higher formant frequencies cannot be varied as much. They seem to be more relevant as personal voice characteristics. Thus, properties of vowel sounds that are of great importance to vowel identity can be described in a chart showing the frequencies of the two lowest formants, as is done in Fig. 1. Note that each vowel is represented by a small area rather than by a point in the chart. In other words, these formant frequencies may vary within certain limits without changing the identity of the vowel. This reflects the fact that a given vowel is normally observed to possess higher formant frequencies in a child or in a woman

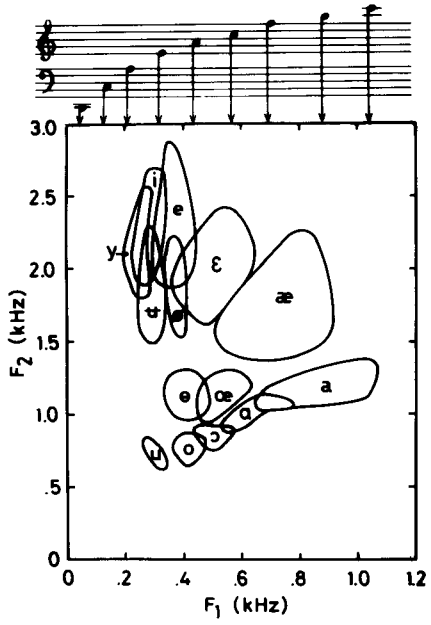






**Fig. 1.** Ranges of the two lowest formant frequencies for different vowels represented by their symbols in the International Phonetic Alphabet (IPA). Above, the scale of the first formant frequency is translated into musical notation.

than in a male adult. The reason for such differences lies in differing vocal tract dimensions, as will be shown later.

In singing, more or less substantial deviations are observed from the vowel ranges shown in Fig. 1. Indeed, a male opera singer may change the formant frequencies so much that they enter the area of a different vowel. For instance, in the vowel [i:][1] as sung by a male opera singer the two lowest formant frequencies may be those of the vowel [y:] according to Fig. 1. And in female high-pitched opera singing the formant frequencies may be totally different from those found in normal speech. Yet we tend to identify such vowels correctly. This shows that the frequencies of the two lowest formants do not determine vowel identity entirely. Next we will see how and why these deviations from normal speech are made in singing.

## III. RESONATORY ASPECTS

### A. Female Singing

#### *1. Formant Frequencies*

A soprano singer is required to sing at fundamental frequencies as high as 1000 or 1400 Hz. In normal female speech the fundamental frequency rarely exceeds about 350 Hz. The normal value of the first (and in some vowels even the second) formant frequency is far below 1000 Hz, as can be seen in Fig. 1. If the soprano were to use the same articulation in singing a high-pitched tone as in normal speech, the situation illustrated in the upper part of Fig. 2 would occur. The lowest partial in the spectrum (i.e., the fundamental) would appear at a frequency far above that of the first formant. In other words, the capability of the vocal tract to transfer sound would be optimal at a frequency where there is no sound to transfer. It seems that singers tend to avoid this situation. Instead they abandon the formant frequencies of normal speech and move the frequency of the first formant close to that of the fundamental. The main articulatory gesture used to achieve this tuning of the first formant is a change of the jaw opening, which is particularly effective for changing the first formant frequency (cf. Lindblom & Sundberg, 1971). This explains why female singers tend to change their jaw opening in a pitch-dependent manner rather than in a vowel-dependent manner, as in normal speech. The acoustic result of this maneuver is illustrated in the lower part of the same Fig. 2: The amplitude of the fundamental and hence the sound power of the vowel increases considerably. Note that this gain in sound power results from a resonatory phenomenon. It is obtained without increasing vocal effort.

Figure 3 shows formant frequencies measured in a soprano singing various vowels at varying pitches (Sundberg, 1975). As can be seen from the figure, the vowels maintain the formant frequencies of normal speech up to that pitch where the fundamental comes close to the first formant. Above that frequency the first formant is

[1]All letters appearing within [] are symbols in the International Phonetic Alphabet.

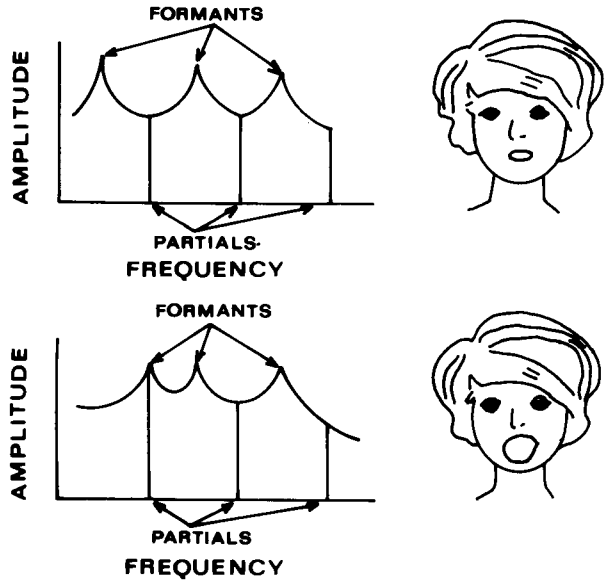



**Fig. 2.** Schematical illustration of the formant strategy in female singing at high pitches. In the upper case the singer has a small jaw opening. The first formant appears at a frequency far below the frequency of the lowest partial of the vowel spectrum. The result is a low amplitude of that partial. In the lower case the jaw opening is widened so that the first formant matches the frequency of the fundamental. The result is a considerable gain in amplitude of that partial (reprinted from Sundberg, 1977b).

raised to a frequency in the near vicinity of the fundamental. If the jaw opening is changed, the main effect observed is in the first formant frequency, but the higher formant frequencies also change to some extent. This is illustrated in Fig. 3: All formant frequencies change when the first formant starts to match the fundamental frequency.

### 2. *Sound Intensity and Masking*

As was mentioned above, the amplitude of the fundamental increases when the first formant is tuned to that frequency. This results in a gain in overall sound pressure level (SPL). The magnitude of the gain can be seen in Fig. 4, which shows the increase in SPL associated with the formant frequencies plotted in Fig. 3. We can see that the pitch-dependent choice of formant frequencies results in an amplitude gain of almost 30 dB in extreme cases. This corresponds to a thousandfold increase of sound power. A perceptually important conclusion is that the female singer will gain in loudness to a corresponding extent.

The singer's need for exceptionally high degrees of loudness is of course a consequence of the fact that opera singers are generally accompanied by an orchestra. The average SPL of an orchestra playing loudly in a concert hall is about 90 to 100 dB. This is much more than we can expect from a human speaker. The masking effect that

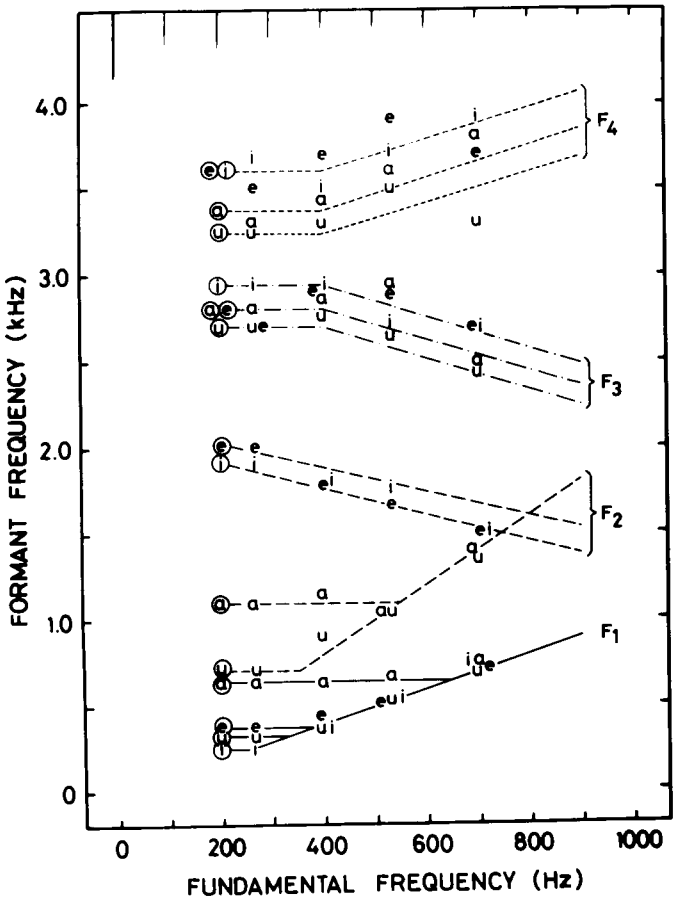



**Fig. 3.** Formant frequencies in the vowels indicated (IPA symbols) measured in a professional soprano singer singing at different fundamental frequencies. The lines show schematically how the formant frequencies are changed with fundamental frequency for the vowels indicated by the circled symbols (adapted from Sundberg, 1978b).

the orchestral sound will exert on a singer's voice is determined by the distribution of sound energy along the frequency scale. A long-time average spectrum of orchestral music shows the average of this distribution. Such a spectrum is shown in Fig. 5. It was obtained from the *Vorspiel* to the first act of Wagner's *Meistersinger* opera. The frequency scale is based on the mel unit, which is preferable when masking and spectral pitch are considered (cf. Zwicker & Feldtkeller, 1967). The graph shows that the strongest spectral components are found in the region of 400–500 Hz. The average spectrum level falls off more steeply toward higher frequencies than toward lower frequencies (Sundberg, 1972a).

The masking effect of a noise with the spectrum shown in Fig. 5 can be estimated from hearing theory (see Zwicker & Feldtkeller, 1967). Avoiding details, we may say that the masking effect will be greatest at those frequencies where the masking sound

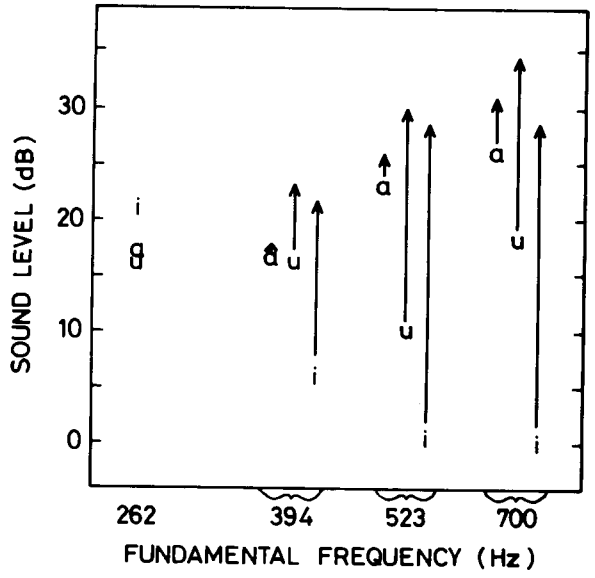



**Fig. 4.** The overall sound level of vowels indicated (IPA symbols) that would result at different fundamental frequencies if the formant frequencies were kept constant at the values observed for the fundamental frequency of 260 Hz in Fig. 3. The arrows show how much these sound levels increase when the formant frequencies are changed with the fundamental frequency in the way indicated by the noncircled vowel symbols in Fig. 3.

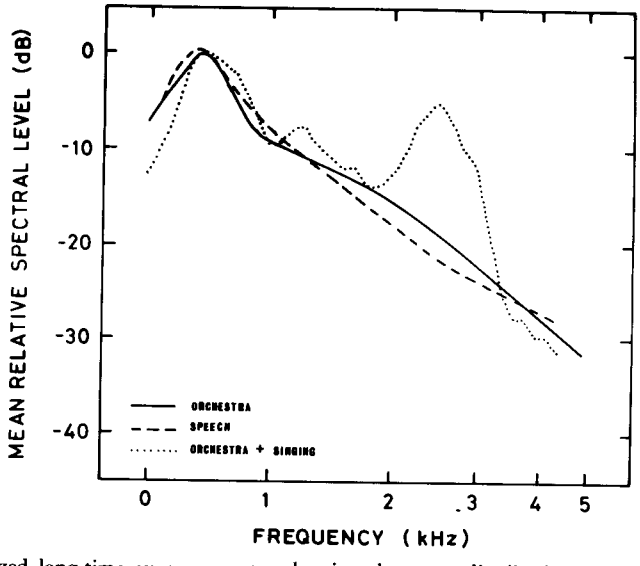






**Fig. 5.** Idealized long-time average spectra showing the mean distribution of sound energy in the "Vorspiel" of act 1 in Wagner's opera "Meistersinger" (solid line) and in normal speech (dashed line). The dotted line pertains to an opera singer singing with orchestra accompaniment (from Sundberg, 1977b).

is loudest and it will decrease as the amplitude of the masker decreases toward higher and lower frequencies. Thus, on the average, the masking effect of the sound of the orchestra will be greatest at 400–500 Hz and decrease toward higher and lower frequencies.

What types of spectra does the human voice produce, then? From Fig. 5 we can see that the long-time average spectrum of normal speech is very similar to that of the orchestra. This suggests that the combination of the sound of an orchestra with that of the human voice during normal speech is probably the most unfortunate one possible. If the sound level of the orchestra is considerably higher than that of the voice, the voice is likely to be completely masked. And, inversely, if the sound of the voice were much stronger (which is very unlikely), the orchestra may be masked. From this we can conclude that the acoustic characteristics of the human voice as observed in normal speech are not very useful for solo parts when combined with the sound of an orchestra. Therefore, these characteristics would need to be modified if both the singer's voice and the orchestral accompaniment are to be both loud and independently audible.

Let us now return to the case of female singing. The spectrum will be dominated by the fundamental if the first formant is tuned to the frequency of the fundamental. We would expect this to occur as soon as the fundamental frequency is higher than the normal frequency value of the first formant. This value is 300 to 800 Hz, depending on the vowel (see Fig. 1). From what was said about masking, we see that all vowels are likely to be masked by the orchestra as long as their first formant is below 500 Hz, approximately. This will be the case for all vowels except [*a:*, a:, æ:] sung at fundamental frequencies lower than about 500 Hz, which is close to the pitch B4. As soon as the fundamental frequency exceeds this value, it will be strong, and its frequency will be higher than that of the partial which is likely to be the strongest in the accompaniment. Summarizing, we can say that a female singer's voice can be expected to be masked by a strong orchestral accompaniment as soon as the vowel is not [*a:*, a; æ:] and the pitch is below B4. This seems to agree with the general experience of female voices in opera singing. They are rarely difficult to hear when they sing at high pitches, even when the orchestral accompaniment is loud.

### *3. Vowel Intelligibility*

We have seen that female singers gain considerably in loudness by abandoning the formant frequencies typical of normal speech when they sing at high pitches. On the other hand, the formant frequencies are extremely important to vowel intelligibility. This poses the question of how vowel intelligibility is affected by high pitches in female singing.

One of the first to study this problem was the phonetician Stumpf (1926), although he probably was not aware of its acoustic background. Stumpf used one professional opera singer and two amateur singers. Each singer sang various vowels at different pitches, turning their backs to a group of listeners who tried to identify the vowels. The identifications were found to be better when the vowels were sung by the

professional singer. These results are illustrated in Fig. 6A. The percentages of correct identifications dropped as low as 50% for several vowels sung at the pitch of G 5. The identification was far better for most vowels when the vowel was preceded by a consonant, particularly [t]. This shows that vowels are much easier to identify when they contain some transitions. Incidentally, this seems to be a perceptual universal: Changing stimuli are more easy to process than quasi-stationary stimuli.

Morozov (1965) studied intelligibility of syllables sung by professional singers (even males) as a function of fundamental frequency. According to his results, intelligibility drops below 80% correct identification above the pitch of E4 and B4 in male and female singers, respectively. At the pitches of C5 and C6 intelligibility has decreased to 50% and 10% correct identification for male and female singing, respectively. At the very highest pitches in female singing, all vowels tend to be perceived as an [a:] according to Howie and Delattre (1962). This appears to agree with results already mentioned on the formant frequencies in female high-pitched singing: The highest pitches would be sung with almost the same (i.e., maximum) jaw opening for all vowels; under such conditions the formant frequency pattern would be similar regardless of which vowel is intended by the singer. Nelson and Tiffany (1968), Scotto di Carlo (1972) and Smith and Scott (1980) all found that vowel intelligibility differed for different vowels sung at the same pitch. For instance, Nelson and Tiffany found that open vowels were harder to interpret correctly than closed vowels and diphthongs.

All these results on intelligibility of vowels and syllables sung at high pitches should be seen against the background of two different effects. One is that singers systematically deviate from the formant frequency patterns of normal speech. This deviation is likely to generate intelligibility problems, at least under certain experi-

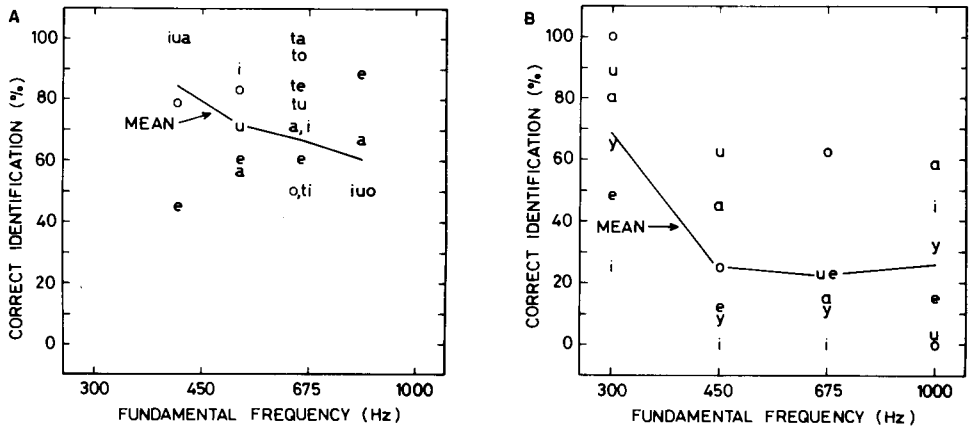



**Fig. 6.** (A) Percentages of correct identification of vowels (IPA symbols) sung by a professional singer according to Stumpf (1926). The solid line represents the average. Note that intelligibility increased when the vowels vowels were preceded by a [t]. (B) Corresponding values obtained by Sundberg (1977a) in an experiment with synthesized vibrato vowels, each of which had the same formant frequencies regardless of the fundamental frequency. The solid line represents the average.

mental conditions. The other effect is that in high-pitched vowels few partials are distributed over the frequency band that normally contains the information we needed to identify a sound as a specific vowel. Thus, a lack of information about vowel identity seems to arise when a vowel is sung at a very high pitch because of the low number of spectrum partials. This effect will certainly add to the score of failing intelligibility in tests like those we have discussed.

The question of how the female singer's deviations from the formant frequencies of normal speech affect vowel intelligibility was studied by Sundberg (1977a). A set of six vowels were synthesized (with vibrato) at different fundamental frequencies ranging from 300 to 1000 Hz. The formant frequencies were kept constant in each of the vowels. The sounds were presented to a group of phonetically trained listeners who tried to identify each of them as one of 12 given vowels. The results are shown on Fig. 6B. It can be seen in this figure that, on the average, vowel intelligibility decreased monotonically as pitch rose, although there were exceptions and minor variations. More important, though, is that the percentages of correct identification were much lower than those reported by Stumpf, using nonsynthetic vowels. A major difference between the synthetic vowels and the vowels used by Stumpf is that the first formant was presumably never lower than the fundamental in Stumpf's case. This being so, we may conclude that the pitch-dependent articulation in high-pitched female singing improves vowel intelligibility when compared to the case in which the formant frequencies are kept constant regardless of the pitch.

Smith and Scott (1980) studied the effect on vowel intelligibility of larynx height and consonantal environment as a function of pitch. Their results, which were based on material from a female singer, confirm that vowel identification is much simpler if the vowel is surrounded by consonants than when it is rendered in isolation. The results also showed that vowels sung with a raised larynx position are more intelligible than vowels sung with the lower larynx position, which their subject normally used in singing. A raised larynx shortens the vocal tract and, by consequence, raises the formant frequencies. Thus, their results suggest that high-pitched vowels are more intelligible if produced with somewhat higher formant frequencies than in normal singing. This assumption is in accordance with the findings of the Sundberg (1977a) investigation. The positive effect of transitions characterizing consonant-vowel-consonant sequences offers formal evidence for Stumpf's observation of the decisive importance to perception of changes in stimuli.

An important point in this connection is the fact that a rise in pitch *must* be accompanied by a rise in formant frequencies if vowel quality is to be preserved. Slawson (1968) found that maximum similarity in vowel quality was obtained when the formant frequencies were increased by 10% on the average for each octave increase in fundamental frequency. It should be noted that Slawson worked with speech-like sounds with a fundamental that never exceeded 270 Hz. In any case, our ears seem to expect a certain increase in the formant frequencies when the fundamental frequency is increased.

The difference in the percentage of correct identifications between Stumpf's and Sundberg's investigations may not necessarily depend solely on a difference in the

formant frequencies. Other differences between synthetic and real vowels may very well have contributed. As was just mentioned, the beginning and ending of a sound are probably very revealing, and presumably the vowels in these two investigations differed in this respect also. Therefore, a direct comparison using well-defined synthetic stimuli is needed before we can draw safe conclusions as to whether or not the pitch-dependent choice of formant frequencies in high-pitched female singing really is a positive factor in vowel identification.

## B. Male Singing

### *1. The "Singer's Formant"*

The audibility problem appears rather different for a male singer than for a female singer. The reason for this lies in the difference in the fundamental frequency ranges. In normal speech the male voice centers around approximately 110 Hz whereas the female voice is about one octave higher. The top pitch for a bass, a baritone, and a tenor is generally E4 (330 Hz), G4 (392 Hz), and C5 (523 Hz), respectively. Consulting Fig. 1 once more, we find that most vowels have a first formant frequency that is higher than these top fundamental frequencies, at least in the cases of bass and baritone voices. The case in which the fundamental frequency is higher than the normal values of the first formant frequency will occur only in the upper part of the tenor and baritone ranges. Therefore, in male singing a pitch-dependent choice of the two lowest formant frequencies is not to be expected except in vowels with a low formant frequency sung at high pitches by tenors and baritones. Measurements by Sundberg (1973) and Cleveland (1977) support this.

The consequence of this seems to be that the male singers produce spectra that are on the average similar to the average spectrum of the orchestral accompaniment (see Fig. 5). Previously, we found that such a similarity in spectrum leads to maximum masking. On the other hand, we know that male voices can be heard readily even when the orchestral accompaniment is loud.

If vowel spectra of normal speech are compared with those produced by male opera and concert singers, at least one difference can be almost invariably observed. Sung vowels contain more sound energy than spoken vowels in the partials falling in the frequency region of 2.5–3 kHz, approximately. Thus, the spectrum envelope exhibits a more or less prominent peak in the high-frequency region. This peak is generally referred to as the "singer's formant" and it has been observed in most acoustic studies of male singing (see, e.g., Bartholomew, 1934; Winckel, 1953; Rzevkin, 1956; Sundberg, 1974; Hollien, Keister, & Hollien, 1978). Figure 7 provides a typical example.

The "singer's formant" has been studied from acoustical and perceptual points of view by Sundberg (1974). There are strong reasons for assuming that the "singer's formant" is an acoustic consequence of a clustering of the third, fourth, and fifth formant frequencies. If formants approach each other in frequency, the ability of the vocal tract to transfer sound increases in the corresponding frequency region. Hence,

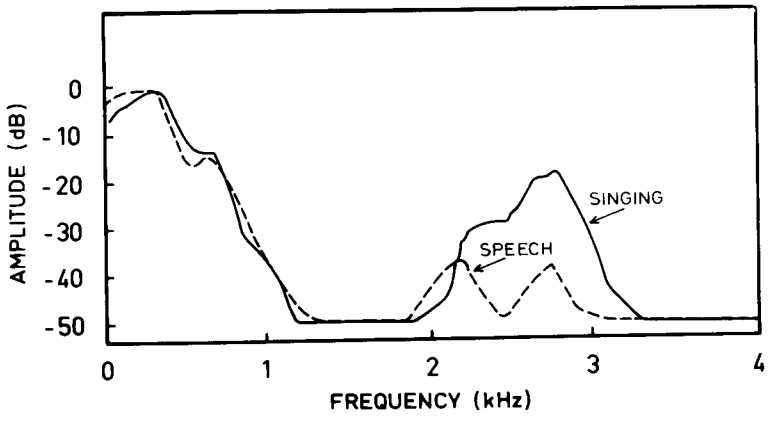



**Fig. 7.** Spectrum contours (envelopes) of the vowel [u] spoken (dashed curve) and sung (solid curve) by a professional opera singer. The amplitudes of the harmonics between 2 and 3 kHz give a marked peak in singing as compared with speech. This peak is called the "singer's formant." It is typical for all voiced sounds in male professional opera singing (adapted from Sundberg, 1978a).

the spectrum envelope peak called the the "singer's formant" seems to be primarily a resonatory phenomenon. However, the amplitude of this peak does not seem to depend on one but on several formants and also, of course, on the source spectrum characteristics.

Formant frequencies are determined by the dimensions of the vocal tract, (i.e., by articulation). According to Sundberg (1974), an articulatory configuration that clusters the higher formants in such a way that a "singer's formant" is generated involves a wide pharynx that appears to result from lowering the larynx. Such lowering of the larynx is typically observed in male singers. Thus, the "singer's formant" can be interpreted acoustically and articulatorily. It should be mentioned that other articulatory interpretations have also been suggested (Hollien et al., 1978).

### 2. *Audibility*

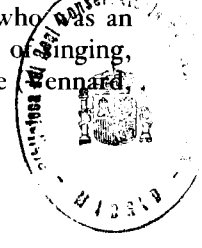
Another question is why male opera singers add a "singer's formant" to their voiced sounds in singing. Probably the reason is perceptual. In a sound illustration contained in Sundberg (1977b), it is demonstrated that a singer's voice is much easier to discern against the background of a noise with the same average spectrum as the sound of an orchestra when the voice has a prominent "singer's formant". This effect is certainly associated with masking. The average spectrum of an orchestra culminates around 400–500 Hz and then decreases toward the higher frequencies (see Fig. 5). The mean spectral level at 2.5–3 kHz is about 20 dB below the level at 400–500 Hz. It seems to be an extremely good idea to enhance the spectrum partials in this frequency range. These partials are likely to be perceived without difficulty by the audience because the concurrence from the orchestra's partials is moderate at these high frequencies.

Another perceptual advantage of producing vowels containing a "singer's formant"

has been suggested by Winckel (1953, and personal communication). It relates to the sound-radiation characteristics of the lip opening. It can be shown theoretically and has also been demonstrated by measurements (Flanagan, 1965) that low-frequency components scatter spherically from the lip opening while the radiation of the high-frequency components is more concentrated along the length axis of the mouth cavity. In other words, high spectrum partials are radiated sagitally with greater efficiency than lower partials. For a singer facing the audience, the sound radiated behind and above his head is probably lost on an opera stage because of the high sound absorption in the backstage area. The high-frequency components contained in the "singer's formant" are lost to a lesser extent as their radiation is more limited to the sagittal direction. Hence, the relative amplitude of the "singer's formant" would be greater compared to the lower partials when the sound reaches the audience. This would help the audience to discern the singer's voice against the background of the orchestral accompaniment. As yet, however, no attempts have been made to assess the magnitude of this effect. It may be mentioned here that many singers and singing teachers speak about "projecting" or "focusing" the voice as a requirement for the audibility of the voice in a large audience. These expressions appear to agree with the above reasoning that the singer's voice is radiated more efficiently in the sagittal direction if the voice contains a "singer's formant."

Before we leave masking problems, one more fact should be mentioned. There are two exceptions to the principle that we cannot hear sounds that fall below the masked threshold resulting from a louder, simultaneous sound. One exception is when the softer sound starts some fraction of a second earlier than the masker sound (cf. Rasch, 1978). The other exception applies to the situation when the masker sound is time varying. Plomp (1977) has demonstrated that we can hear a sound below the masked threshold provided that the masker signal is interrupted regularly. Under these conditions we can hear the signal continuously even in the presence of the masker. Both these cases might apply to the singer–orchestra combination. The orchestral sound, of course, varies in intensity. Given Plomp's results, it should not be necessary for the "singer's formant" to be heard all the time. It would be sufficient for it to be audible during the moments when the amplitude of the orchestral sound in this frequency region is low, provided that such moments are separated by sufficiently short time intervals.

As the "singer's formant" is a perceptually apparent characteristic of male singer voices, it is not surprising that it is recognized as an important timbral attribute among singers and singing teachers. In general, we invent names for such attributes. There seems to be a number of different names for the "singer's formant". Gibian (1972) synthesized vowels in which he varied the frequency of the fourth formant while the remaining formants were kept constant. An expert on singing found that the "placement in the head" of the tone was most "forward" when the fourth formant was 2.7 kHz, which was only .2 kHz above the third formant. Vennard, who was an eminent singing teacher and had a thorough knowledge of the acoustics of singing, simply speaks about "the 2800 Hz" that produces the "ring" of the voice (Vennard, 1967).



### 3. *Modification of Vowel Quality*

Just as in the case of female singing, male singing involves modifications of the vowel qualities characteristic of normal speech. The main articulatory background of these modifications is probably the lowering of the larynx and the widening of the pharynx required for the generation of the "singer's formant". These articulatory characteristics affect not only the third and higher formant frequencies but also the two lowest formant frequencies, which are critical to vowel quality, as was mentioned. Sundberg (1970) measured formant frequencies in vowels sung by four singers and compared these frequencies with those reported by Fant (1973) for nonsingers. As shown in Fig. 8, there are considerable differences here. For instance, the second

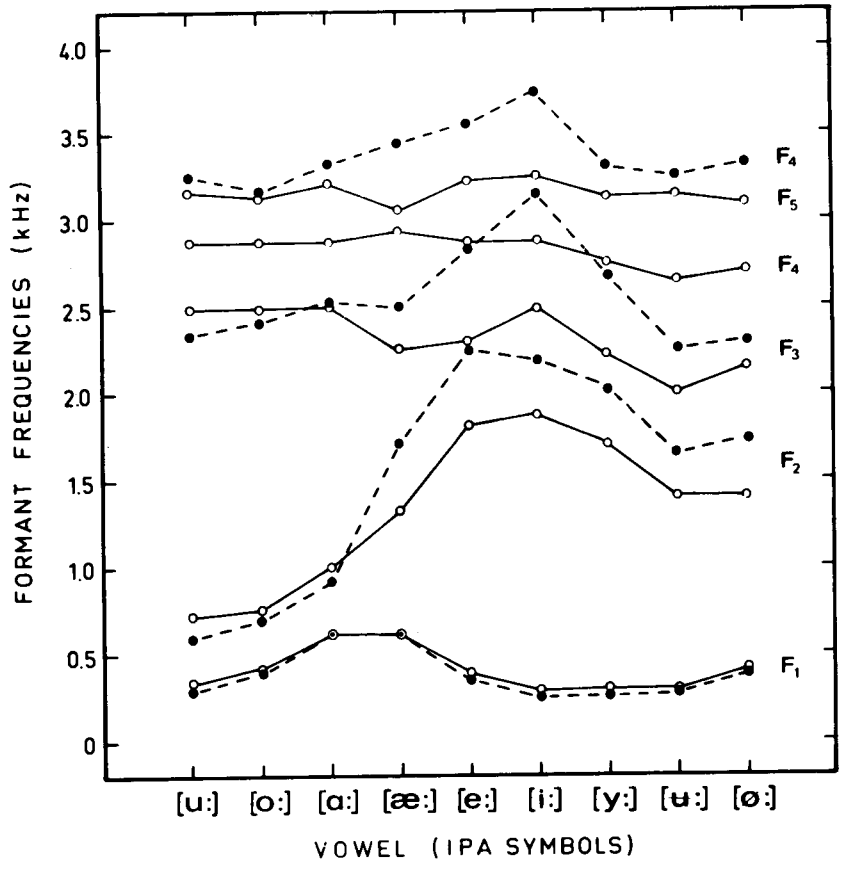



**Fig. 8.** Average formant frequencies in different vowels as produced by nonsingers (dashed curves) according to Fant (1973) and four male singers (solid curves) according to Sundberg (1970). Note that the fourth formant ($F_4$) in nonsingers is slightly higher in frequency than the fifth formant ($F_5$) in singing (reprinted from Sundberg, 1974).

formant does not reach as high a frequency in sung vowels as in spoken vowels. This is the acoustic consequence of a wide pharynx and a low larynx. As a result, some vowels do in fact assume formant frequencies typical of a different vowel in singing. This poses the same question as was posed for female singing: Can we really identify the sung vowels correctly?

Unfortunately, there is no formal evidence available to supply an answer to this question. (It will be recalled that the Morozov (1965) study concerns syllables, not isolated vowels.) On the other hand, the differences in quality between spoken and sung vowels are well known, at least to singers and singing teachers. Many singing teachers instruct their students to modify an [i:] toward a [y:], and [e:] toward an [œ], an [a:] toward an [*a:*], etc. (see, for example, Appleman, 1967). It is considered important that a vowel should not be *replaced by* but only *modified toward* another vowel. This must mean that the sung vowels do retain their vowel identity, although the two lowest formant frequencies are clearly "wrong." It is likely that a low value of the second formant frequency in front vowels can be compensated for by the presence of the "singer's formant."

In summary, we can say that the departures from the formant frequencies typical of normal speech lead to a modification of vowel quality. This modification is probably not sufficiently great to shift the vowel identity. With front vowels, part of the reason for this might be that the "singer's formant" compensates the effect from the too low frequency of the second formant. It seems likely that transitions associated with consonants are the most important factors for vowel identification.

Before we leave this subject, reference should be made to a study by Simon, Lips, and Brock (1972). It concerns the spectra of a vowel sung with differing timbres by a professional singer. These measurements show how properties of the spectrum vary when the singer mimics different types of singing with labels such as *Knödel*. It seems that formant frequencies explain many of these differences.

## C. Voice Classification

### *1. Bass, Baritone, and Tenor Timbre*

As we all know, singing voices are classified in terms of soprano, mezzosoprano, alto, tenor, baritone, bass. The main criterion for such classification is the pitch range available to the singer. If a singer's range is C3 to C5 (131–523 Hz), his classification is tenor. Pitch ranges of different voice classifications overlap to some extent. In fact, the range C4 to E4 (262–330 Hz) is common to all voices. Still, we rarely have any difficulty in deciding whether a tone in this range is sung by a male or a female singer, and often we can even judge the voice classification correctly.

Cleveland (1977) studied the acoustic background of this discrimination ability in the case of male singing. He presented five vowels sung at four pitches by eight singers classified as basses, baritones, or tenors to singing teachers who were asked to decide on the voice classification. The natural beginnings and endings of the tones were spliced out. The results revealed that the major acoustic cue in voice classifica-

tion is the fundamental frequency. Incidentally, the same result was found by Coleman (1976) in a study of maleness and femaleness in voice timbre. The result is not very surprising if we assume that we rely mainly on the most apparent acoustic characteristic in this classification task. By comparing vowels sung at the same pitches, Cleveland found that the formant frequencies serve as a secondary cue. The trend was that the lower the formant frequencies, the lower the pitch range the singer is assumed to possess. In other words, low formant frequencies seem to be associated with bass singers and high formant frequencies with tenors. In a subsequent listening test Cleveland verified these results by presenting the same singing teachers with vowels synthesized with formant frequencies that were varied systematically in accordance with his results obtained from real vowel sounds.

Cleveland also speculated about the morphological background of these findings. As has been described, formant frequencies are determined by the dimensions of the vocal tract. These dimensions are smaller in children and females than in male adults, and the formant frequencies differ accordingly. As a longer tube resonator has lower resonance frequencies than a shorter tube, the formant frequencies produced by a male tend to be lower than those produced by a female for a given vowel. The female vocal tract is not simply a small-scale copy of the male vocal tract (Fant, 1973). The pharynx-to-mouth length ratio is smaller in females than in males. The acoustic consequence is that certain formant frequencies in certain vowels exhibit greater differences between sexes than others, as can be seen in Fig. 9 (see also Nordström, 1977). The greatest variations are found in the two lowest formant frequencies. In the same figure are shown the corresponding values that Cleveland found when he compared a tenor voice with a bass voice. There is a clear similarity suggesting a similar morphologic background in the tenor/bass case as in the female/male case. This

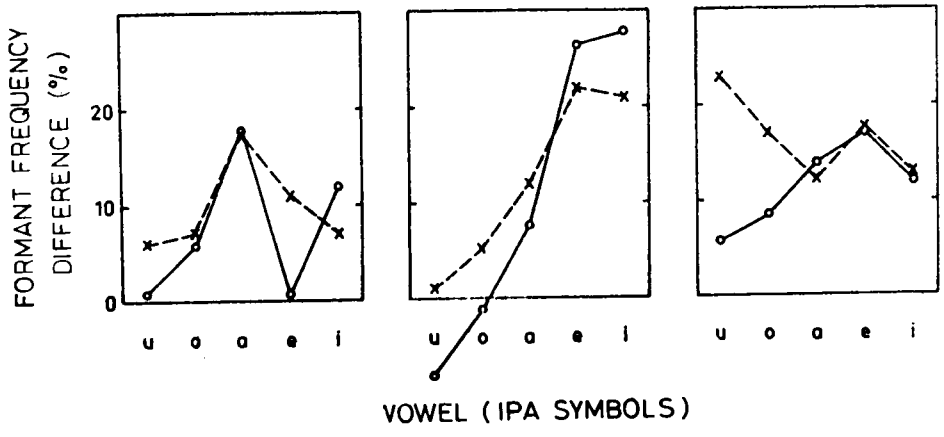



**Fig. 9.** Percentage differences between various voices in the first (left), second (middle), and third (right) formant frequency in the vowels indicated. Solid curves compare a tenor with a bass singer, according to Cleveland (1977). Dashed curves show the average over six languages of female nonsingers compared with male nonsingers according to Fant (adapted from Fant, 1975).

finding should be corroborated by X-ray measurements on a number of singers of differing voice classification. As yet, we can only hypothesize that tenors tend to have smaller pharynx-to-mouth ratios than basses.

An investigation by Dmitriev and Kiselev (1979) is of interest in this connection. It shows a clear correlation between the center frequencies of two peaks in a long-term average spectrum of voices of different classifications and the associated vocal tract length, particularly with regard to the center frequency of the higher of these two peaks, which would reflect the "singer's formant." The lower of the peaks, which occurs in the frequency range of about .4 to .9 Hz, would relate to the average frequency of the first formant.

In summary, experimental support has been found for the following conclusions. In voice classification the fundamental frequency seems to be the main acoustic cue. However, formant frequencies typically differ between bass, baritone, and tenor voices. These differences, which probably reflect differences in vocal tract dimensions as well as in the pharynx-to-mouth length ratios, serve as secondary cues in voice classification.

### 2. *Alto and Tenor Timbre*

Generally, there is a clear difference in timbre between alto and tenor voices. Since their pitch ranges overlap to a great extent, the fundamental frequency cannot always explain this difference. We have seen that tenor and bass voices differ with respect to formant frequencies in a way similar to that in which female and male voices differ. This suggests that with respect to formant frequencies a tenor voice is more similar to a female voice than a bass voice is. What, then, are the acoustic differences that account for the timbral differences between alto and tenor voices?

Ågren and Sundberg (1978) compared two alto and two tenor voices singing the same six vowels at the same pitches. Although the subjects in this study were too few to allow for general conclusions, and in spite of the fact that no perceptual evaluation of the results was attempted, the results of this investigation have perceptual relevance. Only the fourth formant frequency showed a consistent difference that could account for the considerable difference in timbre between the two voice types. This formant was observed to have a higher frequency in the alto voices than in the tenor voices. This means that the frequency distance between the third and fourth formants was smaller in the tenor voices. There was also a clear difference in the source spectrum: The amplitude of the fundamental was higher in the alto voices. As we shall see later, this might reflect the principle that the amplitude of the fundamental decreases toward the upper limit of a singer's pitch range. Obviously, the tenors sang in the upper part of their pitch range when singing at the same fundamental frequencies as the altos.

The smaller distance between the third and fourth formants in the tenor voices is not surprising. As was mentioned, it implies a stronger "singer's formant" in the tenor voices. This is in agreement with the finding of Hollien *et al.* (1978) that the "singer's formant" is more developed in male voices than in female voices.

The perceptually interesting point is the relationship between the "singer's formant" and roughness in timbre. First some words about roughness may be in order. With more extreme degrees of roughness we perceive a series of rhythmic pulsations, as in the case of a very low-pitched tone from a reed instrument, such as the bassoon. At the other extreme (i.e., with complete absence of roughness) the sound is completely smooth. A good example of this is the timbre of a sine wave. Against this background it seems intuitively reasonable to assume that male and female voices differ with respect to roughness.

Terhardt (1974) has studied the acoustical correlate of roughness. He revealed that roughness appears as soon as a spectrum contains at least two partials that (a) have high and reasonably equal amplitudes and (b) excite the same critical band of hearing. This critical band is a sort of analyzing bandwidth of the ear. It is about 100 Hz for center frequencies up to 450 Hz and approximately 20% of the center frequency for higher frequencies. Let us consider a harmonic spectrum with a fundamental frequency lower than 100 Hz. In this case, all adjacent partials are closer than 100 Hz. Therefore, any pair of adjacent partials will excite one critical band. As a consequence, any pair of partials may contribute to roughness. Let us next consider a harmonic spectrum with a higher fundamental frequency. In this case, all of the lowest five partials excite different critical bands because they are more than 20% apart in frequency. Hence, they cannot give rise to roughness. Roughness can occur in such spectra only if one or more pairs of higher partials have high and reasonably equal amplitude.

Let us now return to the alto/tenor case. In the pitch region of relevance, only pairs of partials above the fourth partial can give rise to roughness. If we take into account the fundamental frequency ranges of alto and tenor voices, this leads us to consider partials in the vicinity of the third formant, which is generally located around 2500 Hz. If the frequency distance between the third and fourth formant is on the same order of magnitude as the fundamental frequency, it is likely that these formants will enhance two adjacent partials and thus give rise to roughness. In the Ågren and Sundberg study (1978) the mean frequency distance between these formants in the six vowel sounds analyzed was found to be 785 Hz, (SD = 212 Hz) in the case of the two altos and 439 Hz (SD = 189 Hz) in the case of the two tenors. Thus, we find that this distance is of the same order of magnitude as the frequency separation between the partials only in the case of the tenor voices. Therefore, we would expect roughness from the tenor voices but not from the alto voices. It seems reasonably safe to assume that alto and tenor voices differ with respect to roughness because of the difference in the frequency distance between the third and fourth formants.

## IV. PHONATION

Up to this point we have focused primarily upon resonatory phenomena (i.e., on characteristics associated with formant frequencies). In the present section some as-

pects on phonation will be presented (i.e., the behavior of the vibrating vocal folds and the acoustic properties of the resulting voice source).

## A. Vocal Effort and Pitch

Voice-source characteristics change with vocal effort and with pitch as related to the pitch range of the individual voice. In normal speech the amplitudes of the higher overtones increase at a faster rate than the amplitude of the fundamental when vocal effort is increased while the reverse is true when pitch is raised (cf, e.g., Fant, 1960). Sundberg (1973) studied the voice source in two professional singers and found the amplitudes of the overtones above 1 kHz to increase at a faster rate than the lower overtones. This occurred not only when the vocal effort was increased, but also when pitch was raised.

In a later study Sundberg and Gauffin (1978) measured both the waveform and spectrum of the voice source in singers. They used an inverse filter technique, such as that of Rothenberg (1972), which allowed them to study the partials up to 1.5 kHz, approximately. The results showed that in this low-frequency part of the source spectrum, the amplitude relationship between the fundamental and the overtones changed with pitch rather than with vocal effort. When pitch was raised, the amplitudes of the overtones increased more than the amplitude of the fundamental. When vocal effort was increased, the amplitude of the fundamental was observed to increase at approximately the same rate as the SPL. As the SPL is mainly determined by the amplitude of the partial underlying the first formant, which was an overtone, the amplitude of the fundamental increased at about the same rate as the amplitudes of the overtones. However, the amplitudes of the source spectrum partials above 1.5 kHz would be expected to increase more rapidly than the amplitudes of the lowest source spectrum partials when vocal effort is increased. This can be inferred from the observations by Sundberg (1973) and Hollien *et al.* (1978) that the amplitude of the "singer's formant" increases faster than the SPL when vocal effort is raised. In addition to these findings Sundberg and Gauffin (1978) also found that "pressed" phonation is characterized by strong overtones as compared with the fundamental.

The above findings may explain why Ålgren and Sundberg (1978) found a stronger fundamental in the alto than in the tenor voices. In that investigation all subjects sang vowels at identical fundamental frequencies. Hence, the tenors sang in the upper part of their pitch range while the altos sang in the lower part of their pitch range. A similar reasoning can be applied whenever vowel sounds of voices with differing pitch ranges are compared under conditions of identity with respect to fundamental frequency (for instance, in Cleveland's (1977) investigation). It is likely that voice experts can hear if an individual phonates in the upper, middle, or lower part of his/her pitch range by listening to the voice timbre characteristics associated with the voice-source spectrum.

In summary, the dominance of the source spectrum fundamental is promoted by



low and medium pitch. At high pitch and in pressed phonation the dominance of the fundamental is decreased. The amplitudes of the overtones above 1.5 kHz increase more rapidly than the overall SPL when vocal effort and pitch are raised.

## B. Register

One phonatory aspect of singing which has been subject to a considerable amount of scientific effort is register (see Large, 1972). Unfortunately, register terminology is rather chaotic. On the other hand, there is general agreement that a register is a series of adjacent tones on the scale which (a) sound equal in timbre and (b) are felt to be produced in a similar way. Also, it is generally agreed that differences in register reflect differences in the mode of vibration of the vocal folds (see Hollien, 1974).

Several objections can be raised against this definition since it relies so heavily on subjective impression. Nevertheless, lacking a definition based on physiological facts, we accept it for the time. In trying to understand it, it is helpful to contrast two registers, namely the modal (normal) and the falsetto register of the male voice. These are two clear examples of different registers. In the female voice there are three main registers: chest, middle, and head. They cover the lowest, the middle, and the top part of the pitch range, respectively. However, many voice experts speak about modal and falsetto register both in male and female voices.

### *2. Female Chest and Middle Register*

Large, mostly with various co-authors, has published a series of investigations concerning the acoustic characteristics of different registers. With respect to the physiological background of registers, Large, Iwata, and Von Leden (1970) found that tones sung in the chest register consume more air than those sung in the middle register. They conclude that the conversion of air stream to sound is more efficient in the chest registers.

Large and Shipp (1969) studied the influence of various parts of the spectrum on the ability to discriminate between the chest and middle registers. The material included the vowel [*a:*] sung by 12 singers at the pitch E4 (330 Hz). The quality of the vowel (but obviously not its timbre) and its acoustic intensity were kept approximately constant by the singers. A test tape was made in which the natural beginnings and endings of each tone were spliced out. The vowel sounds were presented with and without low-pass filtering at 1400 Hz to a jury of voice experts who were asked to classify them with respect to register. The results revealed that generally the registers were correctly identified when the vowels were unfiltered. When they were low-pass filtered, identification of register became more difficult, but it never dropped as far as the level of mere guessing. The authors concluded that the higher spectrum partials merely contribute to register differences. Large (1974) returned to this question in a later study. His results agreed with those of the previous investigation, but this time he studied the spectrum of the vowels more closely. The experiment showed typical

differences between the registers in the amplitudes of the lower spectrum partials. By and large, the chest-register vowels were found to possess stronger high partials than the middle-register vowels. However, the differences were all very small. Large found that the results support the assumption that register differences reflect differences in the vocal fold vibrations.

Sundberg (1977c) studied the voice-source and the formant frequency characteristics underlying timbre differences between the chest and middle register in one soprano singer. The subject sang a vowel in both registers at the same pitches. The intensity was left to the subject to decide. The results revealed a considerable source spectrum difference in that the relative amplitude of the fundamental was more than 10 dB stronger in the middle register. This is much more than the small differences reported by Large (1974). Probably, the register difference was less pronounced in Large's subjects.

Sundberg (1977c) also found formant frequency differences between the registers, suggesting that the timbre differences between the registers may depend not only on voice source, but also on articulatory differences. In order to test this hypothesis, he synthesized pairs of vowels differing in either formant frequencies or source spectrum. A group of singing teachers were asked to identify the registers in these pairs of vowel sounds. The results confirmed that both formant frequencies and source spectrum may contribute to register identification. Thus, some of the spectral differences reported in the previously mentioned studies may have been due to formant frequency differences. We will return to this question later.

### 2. *Male Modal and Falsetto Registers*

A number of investigations into the differences between the modal and falsetto registers have been published. Although falsetto is rarely used in traditional Western singing—except, perhaps, in counter-tenor singing—the research in this field will be reviewed.

It has been shown that physiologically the vocal folds are longer, stiffer, and thinner in falsetto than in modal register. As a rule, the glottis is never completely closed in falsetto. This is in agreement with the finding of Large, Iwata, and von Leden (1972) that falsetto tones consume more air than comparable tones sung in modal register. On the other hand, complete glottal closure may occur in falsetto (see Fig. 35, frame F on page 71 in Vennard, 1967); and, inversely, incomplete glottal closure is sometimes observed in modal register phonation.

Part of the literature on falsetto and modal register focuses on the question of whether or not listeners can identify these registers from sustained, isolated vowel sounds. Even though difficulties sometimes arise, particularly when the vowels are sung by professional singers, the answer is generally found to be in the affirmative (see, for instance, Lerman & Duffy, 1970). A dependence on the subjects' voice training was also found, which is not surprising since singers are generally trained to blend registers (i.e., to reduce timbral differences between registers). An experiment by Colton and Hollien (1973) allowed for more detailed conclusions. They found vocal

registers to be a multidimensional phenomenon: "Under normal conditions it is the combination of pitch, loudness, and quality that an observer utilizes to distinguish two vocal registers. When pitch and loudness are equalized, register discrimination becomes more difficult."

The study by Large *et al.* (1972) used vowels recorded under conditions of equality in pitch and acoustic intensity. Under these conditions the falsetto was found to produce weaker high overtones than the modal register. This agrees with the observation made by the same authors that more air is consumed in falsetto singing; the conversion of air stream into sound is less efficient in falsetto than in modal register.

Again equalizing pitch and acoustic intensity, Russo and Large (1978) compared the two registers perceptually and acoustically. Twelve expert listeners judged the similiarity of pairs of tones sung in the different registers. The pairs considered most dissimilar in timbre differed mainly in (a) the amplitudes of the higher spectrum partials, which were lower in falsetto, and (b) the amplitude of the fundamental, which tended to be slightly greater in falsetto. Both these observations agree with spectral evidence collected from singers and nonsingers that Colton had published earlier (1972).

These studies have dealt with the amplitudes of spectrum partials. As we have seen, such amplitudes depend not only on the amplitudes that the partials have in the source spectrum, but also on the frequency separation between the partials and the formants. Thus, the relationships between amplitudes of individual partials and identification of registers are strongly influenced by the formant frequency differences between the spectra compared. Against this background it seems interesting to explore the properties of the voice source that characterize the registers.

Monsen and Engebretson (1977) studied the voice source in various types of phonation. To eliminate the formants, they used a reflectionless tube into which the subjects phonated. The resulting voice-source waveforms are open to question, probably because of phase distortion in the system they used. Such distortion does not, however, affect the amplitudes of spectrum partials. Hence, their results with regard to voice-source spectrum differences would represent reliable information.

They found that the slope of the spectrum envelope was much steeper in falsetto. In other words, the falsetto voice source was more dominated by the lower source spectrum partials. The data shown in Fig. 10 have been selected from the study of Sundberg and Gauffin (1978) mentioned earlier. The waveform is smoother in the falsetto register than in the modal register, and the amplitude of the source spectrum fundamental is much greater in falsetto. These results obviously agree qualitatively with those of Monsen and Engebretson.

On the other hand, Large and his co-authors mostly found very small differences with respect to the fundamental. There may be several reasons for this difference. One is the fact that all tones were sung with vibrato in the Large studies. As will be shown, this implies that the frequency of each partial varies, and a variation in the frequency of a partial leads to an amplitude variation. If spectra are compared, which have not been sampled at identical vibrato phases, errors can be expected. Such errors will be greatest for partials with frequencies close to the formants. In most of

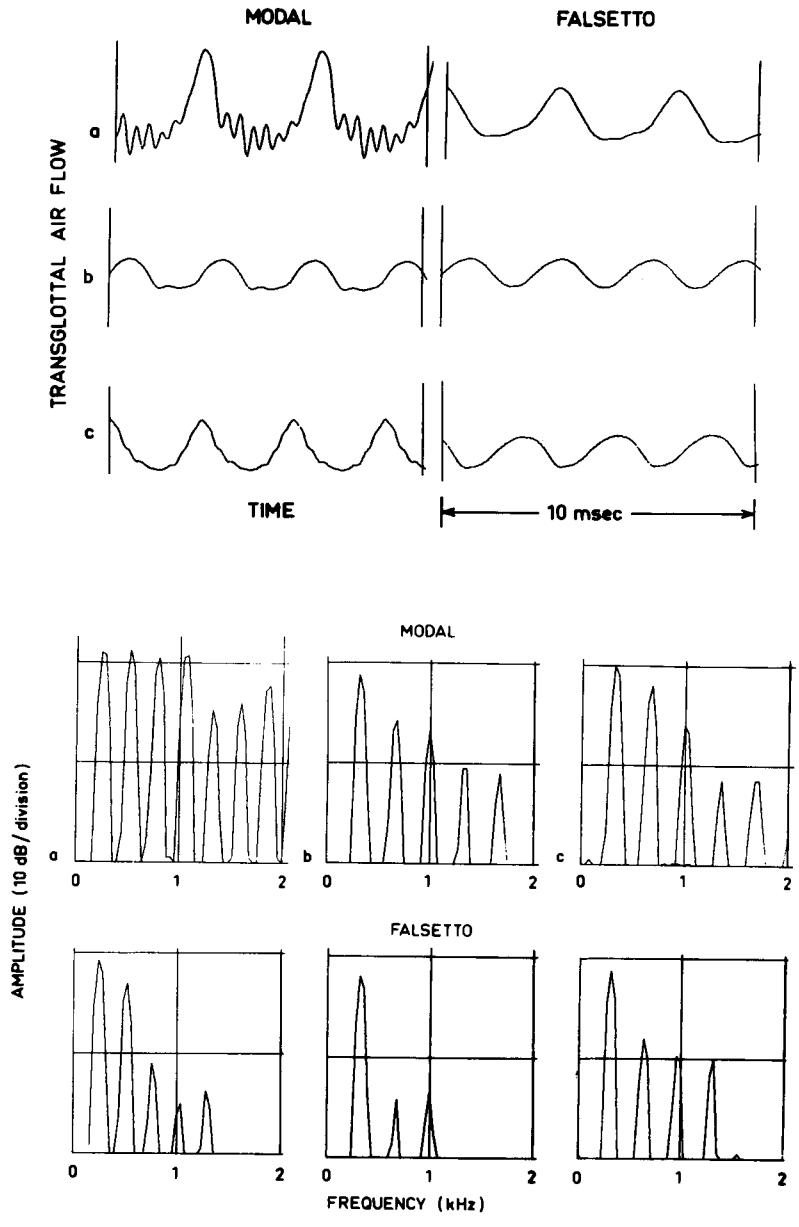



**Fig. 10.** Voice source characteristics in modal and falsetto register in three singers as determined by inverse filtering technique ad modum Rothenberg (1972). The upper series of curves shows the waveform and the lower series shows the corresponding spectrum boosted by 6 dB/octave. The ripple in the modal register waveforms is an artifact due to the particular inverse filter setup used, which could cancel the influence from the two lowest formants only. This caused the "singer's formant" to appear as a ripple. Note that the amplitude difference between the first and second partial is much greater in falsetto than in modal register.

the Large experiments the vowel was [*a:*] and the fundamental frequency was 330 Hz. Then, the amplitudes of the four lowest partials will be rather dependent on the vibrato phase.

However, the main reason that the results differ between the spectral analyses of vowels and the analyses of the voice source probably is that acoustic intensity was not equated in the two types of studies. If the vowel is [*a:*] and the pitch is E4, the second partial will be close to the first formant. The amplitude of the partial underlying this formant is normally quite decisive to the overall acoustic intensity of the vowel. Thus, if two vowels are produced in different registers at the same acoustic intensity, the amplitude of the second source spectrum partial is presumably almost the same in both cases. How, then, should a singer phonate if he is required to produce an [*a:*] at this pitch at the same acoustic intensity in falsetto as in modal register? Probably by raising the amplitude of the second source spectrum partial in the case of falsetto. This can be achieved by increasing vocal effort. Thus, we arrive at the strange conclusion that equality in acoustic intensity may require a difference in vocal effort. If the above reasoning is correct, the difference in the results is a consequence of the fact that acoustic intensity was kept constant in one case while vocal effort was probably constant in the other case.

## V. VIBRATO

### A. Background

Vibrato occurs in most Western opera and concert singing and often in popular singing. Generally, it develops more or less automatically during voice training (Bjørklund, 1961). Acoustically, it corresponds to an undulation of the fundamental frequency. As the spectra of voiced sounds are harmonic, the frequencies of all partials vary in synchrony with the fundamental. The amplitude of a partial depends on how far it is from a formant, and the formant frequencies do not seem to vary appreciably with the vibrato. Therefore, each partial varies in amplitude synchronously with the vibrato.

As the vibrato is a rhythmic undulation (or modulation) of frequency it can be described by two parameters. One is the *rate* of vibrato (i.e., the number of undulations occurring during one second); the other is the *extent* of vibrato (the depth of the modulation expressed in a percentage of the average frequency). More often, however, the vibrato extent is given in number of cents. (One cent is the interval between two tones having the frequency ratio of $1:2^{1/1200}$.)

The physiological background of vibrato is unclear. In electromyographic measurements on laryngeal muscles, pulsations in synchrony with vibrato are generally observed (Vennard, Hirano, Ohala, & Frizell, 1970–1971). Moreover, the subglottic pressure and the transglottal air flow often undulate in synchrony with vibrato as can be seen in recordings published by Rubin, Le Cover, and Vennard (1967). An observation that may prove relevant has been reported by Weait and Shea (1977) who

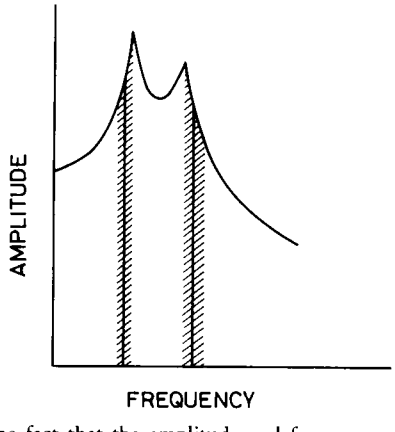
studied the glottal behavior in a bassoon player. They found that the glottal area varied in synchrony with vibrato. This can be interpreted as support for the hypothesis that vibrato originates in the laryngeal muscles while undulations in airflow and subglottic pressure are secondary effects.

Several aspects of vibrato have been studied. As early as the 1930s Seashore (1938) summarized, among other things, a series of investigations that he and his co-workers had made on vibrato. He found the vibrato rate to be rather constant for any given singer but slightly different between singers. The mean for 29 singers was 6.6 undulations per second (extremes 7.8 and 5.9). The average extent was ±48 cents (extremes ±98 and ±31).

## B. Perceptual Aspects

### *1. Vowel Intelligibility*

As mentioned before, the identification of vowels is assumed to be related to the detection of peaks in the spectrum envelope. These peaks signal the frequencies of the formants, and the formant frequencies characterize the vowel. If the number of partials is low compared to the number of formants (i.e., if the fundamental frequency is very high), the peaks in the spectral envelope signaling the formant frequencies would be impossible to detect because there may not be a partial in the neighborhood of every formant frequency. It is not unreasonable to assume that vibrato plays a role here. If the frequency of a partial is slightly lower than that of a formant, an increase in fundamental frequency will *raise* the amplitude of that partial. If the partial is slightly higher in frequency than the formant, a *decrease* of the amplitude will result from the same situation, as is illustrated in Fig. 11. Thus, the phase relationship be-




Biblioteca del Real Conservatorio Superior de Música - MADRID -

**Fig. 11.** Illustration of the fact that the amplitude and frequency of a partial in the spectrum of a vibrato tone vary in phase or in counter phase depending on whether the partial is slightly lower or higher than the closest formant. The hatched area represents the width of the frequency modulation. The frequency scale is linear.

tween the undulations in frequency and amplitude in a vibrato tone actually gives information about the frequency locations of the formants. The question, then, is whether the ear can detect and use this information. If so, vibrato would facilitate vowel identification for high-pitch vowels.

This question was studied in the experiment mentioned earlier concerning vowel identification in the soprano pitch range (Sundberg, 1977a). Each vowel in the test was presented both with and without vibrato. The interpretations made by phonetically trained subjects differed considerably. The degree of agreement between the interpretations was measured in the following manner. Each response vowel was ascribed a set of three formant frequencies. Then all responses obtained for a given stimulus vowel could be regarded as a cloud of points in a three-dimensional space, in which each dimension corresponds to a formant. The center of this cloud was determined. The mean distance between the individual points and the center was next computed using a formula for perceptual distance between vowels suggested by Plomp (1970). It was assumed that this average distance reflected the difficulty with which a vowel stimulus was identified as a specific vowel. The average distance between responses is shown in Fig. 12. As can be seen in the figure, there are no consistent differences between the values pertaining to vibrato tones and those obtained for vibrato-free tones. Therefore, it is reasonable to conclude that vibrato does not facilitate vowel identification. On the other hand, the results may have been rather different if the stimuli had been more like natural vowels sung by sopranos. It is often hard to predict how our ability to identify stimuli is affected when the stimuli do not resemble anything familiar.

### 2. *Singleness in Pitch*

It is a well-established fact that fundamental frequency generally determines pitch. In the case of vibrato tones, however, this is not quite true. While the fundamental frequency varies regularly in such tones, the pitch we perceive is perfectly constant as long as the vibrato rate and extent are kept within certain limits.

What are these limits? Ramsdell studied this question at Harvard University in a thesis work that unfortunately was never published. Ramsdell varied the vibrato rate and extent systematically and had listeners decide when the resulting tone possessed an optimum "singleness in pitch." His results for a 500 Hz tone are shown in Fig. 13. Later Gibian (1972) studied vibrato in synthetic vowels. He varied the vibrato rate and extent and had subjects assess the similarity of this vibrato with human voice vibrato. His results agree closely with Ramsdell's data as can be seen in the same figure. In addition to asking the listeners for the optimum singleness in pitch, Ramsdell also asked for an evaluation of the "richness" in the timbre. His data showed that the optimum in regard to singleness in pitch as well as timbral richness corresponds to the values of rate and extent typically observed in singers.

It is interesting that Ramsdell's curve approaches a straight line in the neighborhood of seven undulations per second. This implies that the extent is not very critical for singleness in pitch at this rate. In contrast to this, there is a strong opinion among

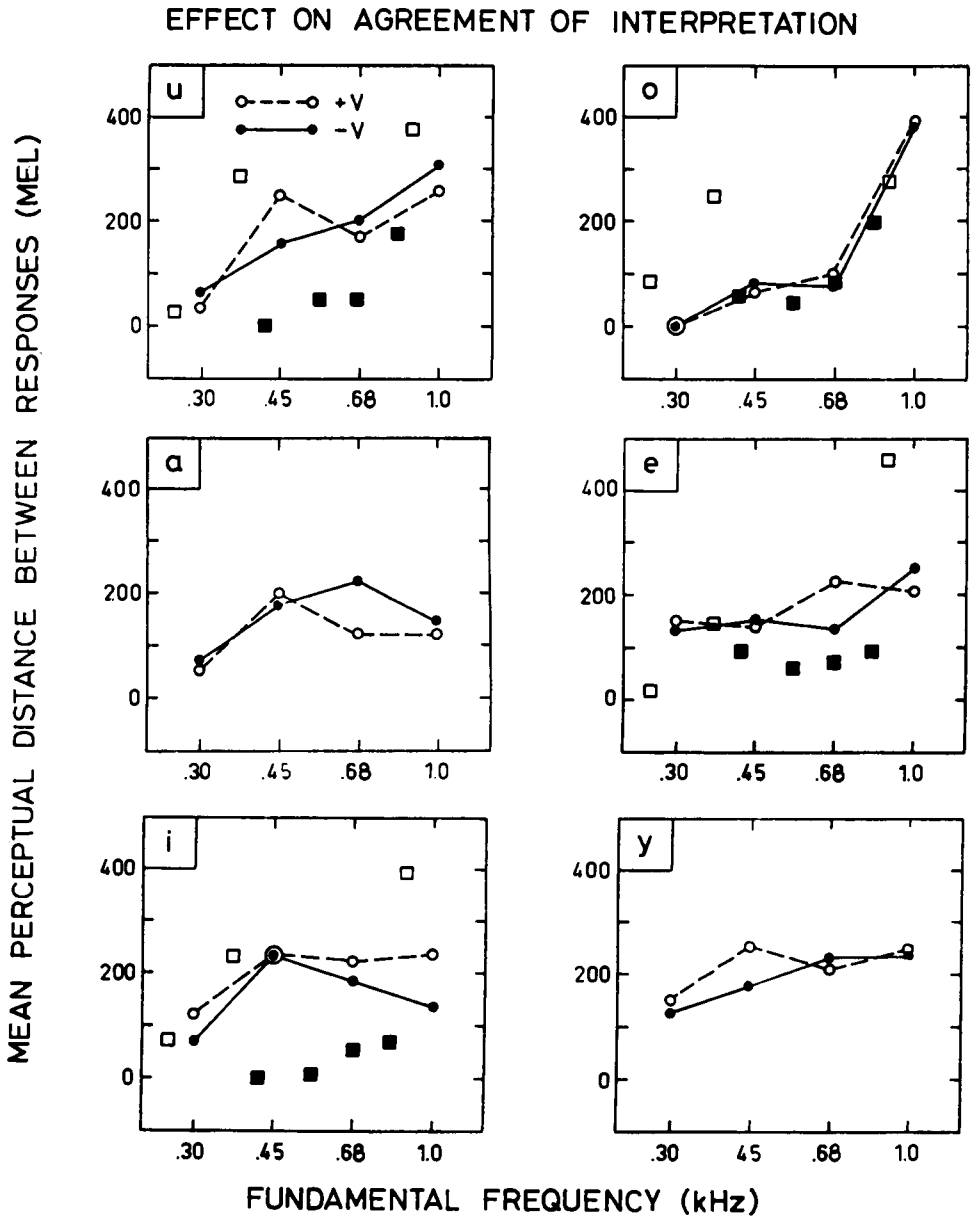



**Fig. 12.** Scatter of responses obtained when subjects attempted to identify synthetic vowels with high fundamental frequencies. The formant frequencies were those of the vowels given in the top left corner of each graph. Dashed lines pertain to vibrato vowels and solid lines to vibrato-free vowels. The squares give values reported by Stumpf (1926). They were observed when subjects identified vowels sung by two untrained singers (open squares) and one professional soprano (filled squares). The procedure for deriving the measure of the scatter is described in the text (reprinted from Sundberg (1977a).

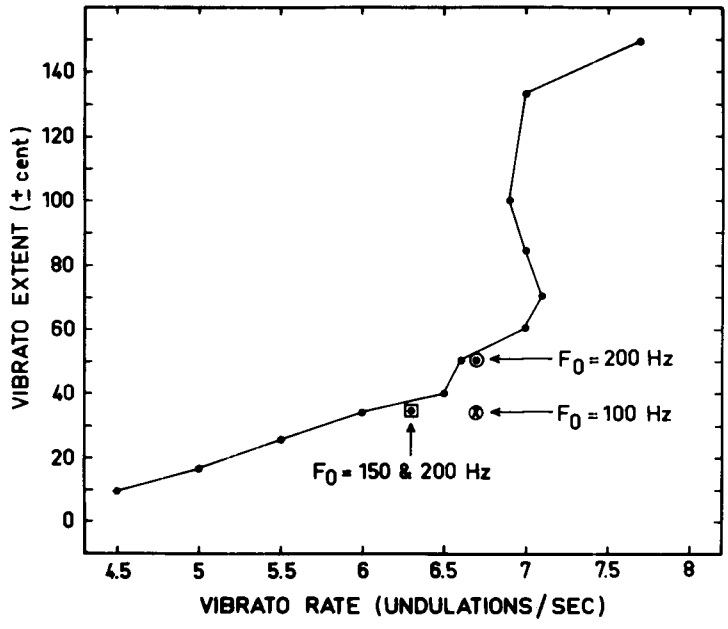



**Fig. 13.** Values of the vibrato extent at different vibrato rates which gives the impression of best singleness in pitch according to Ramsdell. The circled symbols show maximum values obtained by Gibian (1972) in a test where subjects judged the rate and extent closest to human singing. Ramsdell's data pertains to a frequency of 500 Hz while Gibian's data were obtained with the fundamental frequencies ($F_0$) indicated.

some singing teachers that not only slow but also fast vibrato rates are tolerable only if the extent is small. It would be interesting to repeat Ramsdell's experiment with modern equipment.

### *3. Pitch and Fundamental Frequency*

Another perceptual aspect of vibrato is perceived pitch. Provided that the rate and extent are kept within acceptable limits, what is the pitch we perceive? This question was studied independently by Shonle (1975) and Sundberg (1972b, 1978b). Sundberg had musically trained subjects match the pitch of a vibrato tone by adjusting the fundamental frequency of a following vibrato-free tone. The two tones, synthetic sung vowels, were identical except for the vibrato. They were presented repeatedly until the adjustment was completed. The vibrato rate was 6.5 undulations per second, and the extent was ±30 cents. Figure 14 shows the results. The ear seems to compute the average of the undulating frequency, and perceived pitch corresponds closely to this average. Shonle worked with sinewave stimuli and arrived at practically the same conclusion. He was also able to show that it is the geometric mean, not the arithmetic mean that Sundberg worked with, that determines the pitch at least in the case of

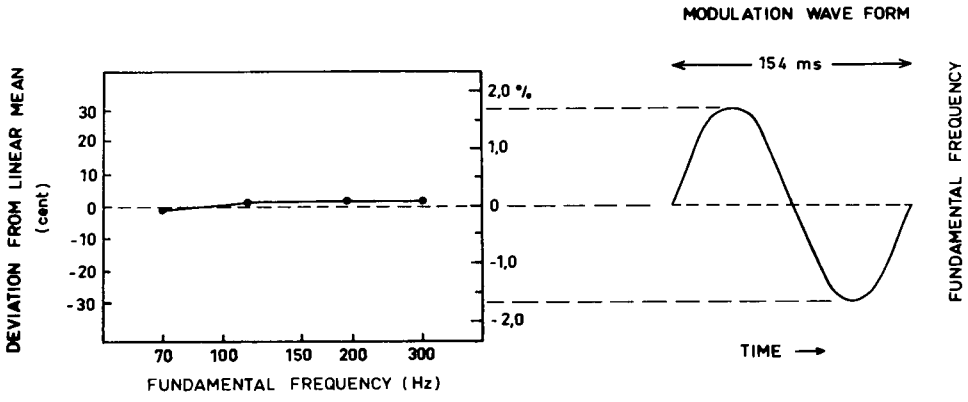



**Fig. 14.** Left graph: average for six musically trained subjects of the fundamental frequency of a vibrato-free synthetic vowel that gives the same pitch as a vibrato vowel according to Sundberg (1978b). The right graph shows the waveform, rate, and extent of vibrato used. Tbe pitch of a vibrato vowel is seen to coincide almost perfectly with the frequency corresponding to the arithmetic mean of the undulating frequency.

sinewave signals. However, the difference between these two means is insignificant in musically acceptable vibratos.

It is frequently assumed that the vibrato is useful in musical practice because it reduces the demands on accuracy of fundamental frequency (see, e.g., Stevens & Davis, 1938; Winckel, 1967). One possible interpretation of this assumption is that the pitch of a vibrato tone is less accurately perceived than the pitch of a vibrato-free tone. Another interpretation is that the pitch interval between two tones that sound simultaneously can be determined with less accuracy when they have vibrato than when they are vibrato-free.

The first interpretation was tested by Sundberg (1972b, 1978b). The standard deviations obtained when subjects matched the pitch of a vibrato tone with that of a vibrato-free tone were compared with the standard deviations obtained from similar matchings in which both tones lacked vibrato. As can be seen in Fig. 15, the differences between the standard deviations were extremely small and dropped slightly with rising fundamental frequency. This implies that the vibrato reduces pitch-perception accuracy slightly for low frequencies. On the other hand, the effects are too small to explain any measurable effects in musical practice.

The second interpretation has not yet been tested, but it is tempting to speculate about it. If two simultaneous complex tones with harmonic spectra constitute a perfectly tuned consonant interval, some partials of one tone will coincide with some partials of the other tone. Let us consider two tones with fundamental frequencies of 200 and 300 Hz (i.e., producing an interval of a perfect fifth). In this case, every third partial of the lower tone (frequencies: 600, 1200, 1800...Hz) will coincide with every second partial of the upper tone. Let us now mistune the interval by raising the frequency of the upper tone to 300.5Hz. This frequency shift equals 2.9 cents, which

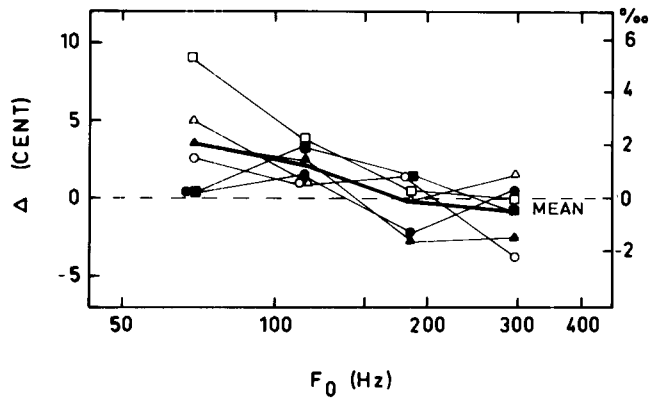



**Fig. 15.** Effect of a vibrato on pitch perception accuracy at different fundamental frequencies ($F_0$). Musically trained subjects first matched the pitch of a vibrato-free stimulus tone by adjusting the fundamental frequency of a subsequent response tone which also lacked vibrato. Then, the same experiment was repeated, except that a vibrato was added to the stimulus tone. Δ is the shift of standard deviation thereby obtained. The individual differences are given by the symbols whereas the heavy solid line shows the group average (reprinted from Sundberg, 1978b).

is impossible for almost any listener to detect under any experimental conditions. (The difference limen for frequency is at least 6 cents but may be considerably higher depending on the experimental method, see Rakowski, 1971.) On the other hand, the partials from the two tones will not coincide any longer. For instance, the fourth partial of the upper tone has a frequency of 4·300.5 = 1202 Hz. This partial will give two beats per second with the sixth partial of the lower tone, which has a frequency of 1200 Hz. There are no difficulties in detecting such beats, provided that both partials have similar and sufficiently high amplitudes. The point is that these beats will not occur if both tones have vibrato. Thus, if two voices sing perfectly "straight" (i.e., without vibrato), the demands on accuracy with respect to the fundamental frequency are higher than if they sing with vibrato. However, this advantage seems to be small. In an unpublished thesis work done at the Department of Speech Communication, Royal Institute of Technology in Stockholm, Ågren (1976) had musically trained subjects match different intervals formed by two simultaneous vibrato tones. The intervals were a major second, a major third, a pure fifth, and a pure octave. The tones were synthetic sung vowels. Some of the subjects managed to obtain a standard deviation as low as six cents in repeated matchings of a given interval. If we may believe that mistunings of this small magnitude can be detected even in musical practice, it would seem that the demands on pitch accuracy are extremely high even when the singers use vibrato. It is likely that the vibrato is accepted and used in singing for other reasons, as will be shown later.

Our conclusions are that the pitch of a vibrato tone is practically identical with the pitch of a vibrato-free tone with a fundamental frequency equal to the geometric mean of the fundamental frequency of the vibrato tone. Moreover, the accuracy with

which the pitch of a vibrato tone is perceived is not affected to any appreciable extent by the vibrato.

## VI. PITCH ACCURACY IN SINGING PRACTICE

A couple of investigations on the perceived pitch of vibrato tones were reviewed earlier. These investigations were made under well-controlled experimental conditions. Do the results obtained that way apply also to musical practice? A study of the accuracy of fundamental frequency in musical practice is likely to answer that question.

In a review of a number of investigations, Seashore (1938) included a wealthy documentation of fundamental frequency recordings of professional performances of various songs. The trend is that long notes are sung with an average fundamental frequency that coincides with the theoretically correct value. This is in agreement with the experimental findings reported previously. On the other hand, they often "begin slightly flat (about 90 cent on the average) and are gradually corrected during the initial 200 msec of the tone." Moreover, a great many of the long tones were observed to change their average frequency in various ways during the course of the tone. Bjørklund (1961) found that such deviations were typical for professional singers as opposed to nonprofessional singers. One possible interpretation of this is that pitch is used as a means of musical expression.

With regard to short tones, the relationship between fundamental frequency and pitch seems to be considerably more complicated. The case is illustrated in Fig. 16 showing the fundamental frequency during a coloratura passage as sung by a male singer. The singer judged this performance to be acceptable. The registration reveals a careful coordination of amplitude, vibrato, and fundamental. Each note takes exactly one vibrato period, and most of the vibrato periods seem to center around the target frequency. However, if we try to apply what has been shown about pitch perception for vibrato tones, we run into trouble. The average fundamental frequency in a coloratura passage does not change stepwise between the target frequencies corresponding to the pitches we perceive; rather the average rises and falls monotonically at an approximately constant rate. Thus, we cannot explain why the passage is perceived as a rapid sequence of discrete pitches. A possible explanation is that the average computation process is interrupted and started again each time there is a minimum in the amplitude and/or frequency curve. However, this is a clear case of an *ad hoc* hypothesis, and no experiments have been performed to support it.

An investigation of interest in this connection should be mentioned here. It has been shown that a glide is perceived as a pitch corresponding to the geometric mean of the extremes of the glide, provided that the product of the frequency change and the time for the change is not greater than five (Nábělek, Nábělek, & Hirsch, 1970). This case will certainly apply to some cases of short notes in singing, but it does not seem to apply to coloratura cases. For instance, the geometric mean of the upward glide does not agree with the geometric mean of the following downward glide in the same tone

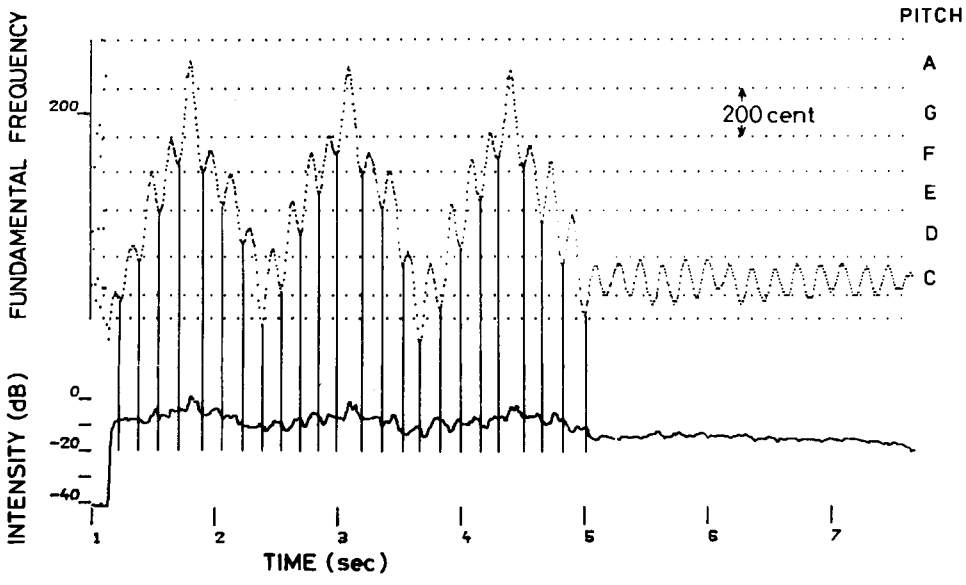



**Fig. 16.** Synchronous recording of fundamental frequency (upper graph) and overall intensity (lower graph) as measured in a professional singer performing a coloratura passage (C3, D3, E3, F3, G3, F3, E3, D3, C3, D3 . . . ). The horizontal dotted lines in the upper graph show the frequencies midway between (i.e., on the geometrical mean of) the scale-tone frequencies calculated from the average frequency of the last note by means of the equally tempered scale.

(see Fig. 16). Moreover, difficulties seem to occur when the pitch is very high. In this case, the pitch changes between the scale tones are wide in terms of absolute frequency. Thus, at high pitches the condition of the product of the change and the time being less than five can hardly be fulfilled any longer. We have to conclude that at present we cannot explain how a coloratura passage can be perceived as a sequence of discrete pitches.

From what was just said, it seems that Seashore (1938) was right in saying that the musical ear is extremely generous and operates in the interpretive mood when it listens to singing. On the other hand, there are certainly limits for this generosity: There is generally an agreement among experts as to what is in tune and what is off pitch. This would lead us to assume that the analyzing properties of the ear are more important to the pitches we perceive from singing than Seashore assumed.

In a thesis work at the Department of Musicology, Stockholm University, Lindgren and Sundberg (1972) studied what musically experienced listeners considered to be off pitch. A tape was prepared including excerpts of phonograph recordings representing different degrees of singing off pitch along with several cases of apparently perfect intonation. A chart with the notation of the excerpts was given to the listeners, who were asked to circle each note they perceived to be off pitch. The fundamental frequency was analyzed by measuring the frequencies of the high overtones in sound spectrograms (sonagrams). The results showed that tones with an

average frequency matching the theoretically correct value were mostly accepted as perfect intonation. However, several tones that did not meet this demand were also accepted as correct. Theoretically mistuned tones were accepted remarkably often (1) when they occurred on an unstressed position in the bar, (2) when they were a little sharp, and (3) when they occurred on emotionally prominent places in the text.

This last point again suggests that deviations from the theoretically perfect pitch is used as an expressive means in singing. Support for this assumption can be found in measurements of clarinet playing (Sirker, 1973). Also, it seems typical of music that the composer and the performer build up expectations in listener as to what might follow. Occasionally, minor deviations from what was expected are made. It is the author's belief that such deviations contribute to the excitement we can perceive when we listen to a good performance.

If it is true that deviations from theoretically correct frequencies are used as an expressive means in singing, an important conclusion regarding the benefit of the vibrato can be made. We have seen that vibrato-free representation of mistuned consonant intervals give rise to beats and beats seem to be avoided in most types of music. By adding a vibrato, the singer escapes the beats. Consequently, the vibrato allows him or her more freedom in using deviations from theoretically correct frequencies.

This point is illustrated in Fig. 17. It shows the distribution of fundamental frequencies averaged over approximately the duration of one vibrato cycle. The data pertain to a part of a song performed by a first-rate opera singer. For comparison, a registration of the same song performed by a singing synthesizer is shown (Sundberg, 1978b). The vibrato rate and extent were the same in the synthesis as in the real performance. The scale-tone frequencies are represented by peaks. These peaks are seen to be considerably wider in the case of the real singer than in the case of the synthesizer. This agrees with the assumption that deliberate deviations from expected pitches are used in singing. In the same figure a third distribution is shown. It pertains to the members of a distinguished barbershop quartet. The vibrato is not used in barbershop singing. Hence, the chords must be perfectly tuned to avoid beats, so the singers have very little freedom as regards fundamental frequency. The scale tones are seen to correspond to very narrow peaks. This means that the frequency value corresponding to a given scale tone varies extremely little in barbershop singing. It is likely that this is a consequence of the absence of vibrato. Although we need more measurements on vibrato-free singing, we may hypothesize that the vibrato offers the singer a freedom in the choice of fundamental frequencies and that this freedom is used for purposes of musical expression.

del Real Conservatorio Superior de Música

## VII. PHRASING AND EMOTION

One of the most essential things about singing (and music in general) is the truly remarkable fact that a performer is able to use sound signals in such a way that phrases rather than sequences of unrelated tones are communicated to the listener. This

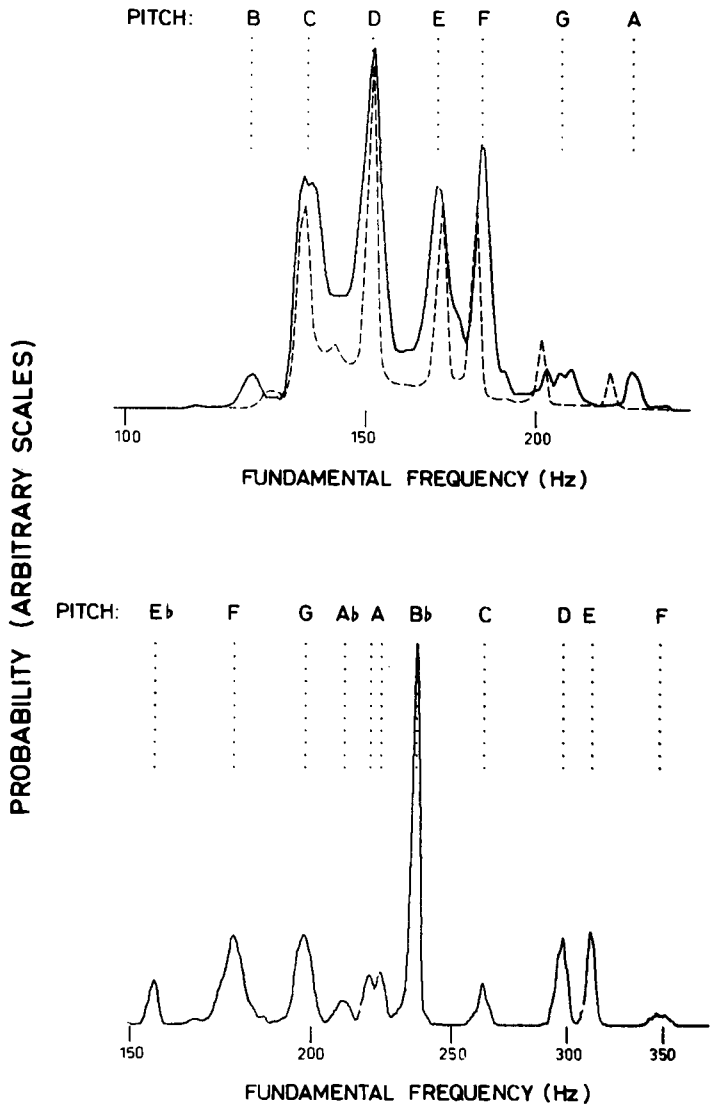



**Fig. 17.** Distribution of fundamental frequencies in singing. The upper graph pertain to a professional singer (solid curve) and a singing synthesizer (dashed curve) performing the same song. In both cases the fundamental frequency was averaged with a running time window corresponding to one vibrato cycle, approximately. Thus, the distributions should be identical if the singer was as accurate with respect to fundamental frequency as the synthesizer. The lower graph was obtained from a distinguished barbershop quartet (Happiness Emporium, Minneapolis) singing a chord progression. Note that the widths of the scale-tone peaks are generally much narrower in the barbershop singers who lack vibrato, than in the opera singer, who has vibrato. Note also that the pitch A is represented by two peaks in the case of the barbershop quartet presumably because it appeared several times in the song with different harmonic functions.

would be the perceptual effect of phrasing. Moreover, the performer can add something to these signals that evokes an emotional experience in the listener. Although this must be considered the essence of singing and music, we know little about how this effect is achieved. Two studies will be mentioned even though they merely touch on the problem.

Seashore (1938) regarded phrasing as a temporal phenomenon. He compared the durations of beats, bars, and phrases in a couple of performances of the same song. It turned out to be difficult to detect any consistent pattern in the deviations from the average durations of musical segments such as bars. Seashore's comment is "The singer's interpretation is accomplished by artistic deviations from this average. . . . In none of these singers is there the slightest approach to an even time for a measure." Thus, Seashore was convinced that these deviations are by no means random.

Sundberg (1978b) started from the hypothesis that the structure of the music is mirrored in a performance perceived as musically acceptable. Such a mirroring must be accomplished by means of a set of rules that operate on the structure. In other words, the application of the rules requires information about the structure, which means that the performer must be consciously or unconsciously aware of this structure. One question is how the singer announces that a particular sequence of notes belongs together and constitutes a structural unit, such as a phrase. A singer's performance of a song was analyzed, and certain regularities were observed. These regularities were formulated as tentative rules. The perceptual relevance of these rules was tested by implementing them one by one in a rule system used to control a singing synthesizer's performance of the same song. In sound examples, published with the investigation, one can experience how the different rules affect the performance. An apparent effect was that the tones sounded unrelated to each other as long as they did not form a long-term amplitude event. Thus, by introducing the same amplitude pattern that comprises an entire phrase, the tones in the phrase sound as if they belong together. This is a clear case in which systematic deviations from an average (namely the amplitude average) result in a specific musical effect. We can assume that systematic deviations from other averages, such as theoretically correct fundamental frequency or duration, are used analogously in singing and that the result is bringing the performance to an artistic agreement with the structure. However, a considerable amount of research is needed to corroborate this assumption.

The effects of the emotional content of the text on vocal performance has been formally investigated by Kotlyar and Morozov (1976). Eleven professional singers repeatedly performed 16 excerpts from various songs so as to represent different emotional atmospheres: joy, sorrow, fear, anger, and neutral. By means of a listening test, it was confirmed that the singers managed to convey the desired emotional information in performing these excerpts. A series of acoustic aspects of their performances were studied. It was found that each of the five emotions studied was manifested acoustically in different ways. For instance, the average syllable duration was shortest in fear and longest in sorrow. Fear and neutral were associated with the longest and the shortest unvoiced intersyllabic segments. Anger was performed with the loudest sound levels and fear with the lowest sound levels. Sorrow had slow-tone

onsets, and anger and fear had rapid-tone onsets. Thus, each of the five emotions studied showed a specific pattern of acoustic characteristics.

The authors also tried to discover to what extent these characteristics were not only necessary but also sufficient to convey the emotional information. Thereby, they used an "artificial signal comprising a tone modulated by the amplitude of the envelope (observed in the performances), i.e. a signal deprived of such informative attributes as intonational and spectral characteristics." This amplitude-modulated signal was presented to 11 listeners who tried to identify the "possible emotional content." The results suggested that the amplitude-modulation characteristics of the voice are particularly efficient in communicating fear, sorrow, and anger (80, 66, and 60% correct identification) while neutral and joy seem to be signaled mainly along some other acoustic dimensions (24 and 8% correct identification).

It does not appear farfetched to assume that the acoustic characteristics used in singing in order to communicate an emotional atmosphere are the same as those which are used for the same purpose in speech. This assumption is supported by studies of emotional speech published by Fónagy (e.g., 1962, 1976). Moreover, it seems reasonable to hypothesize, as Fónagy does, that such characteristics are merely the acoustical reflections of general patterns of body movements. For instance, we expect no rapid body movements from a truly sad person, and, of course, we would not expect any rapid movements in those laryngeal structures that regulate voice fundamental frequency in that person. There seems to be reason to believe that vocal behavior under the influence of emotions is the mere translation into the acoustical domain of such general patterns of body movements. If this is correct, it seems to follow that our ability to appreciate singing should partly rely on our ability to infer the emotional state of another person from the person's way of speaking.

## VIII. CONCLUDING REMARKS

In the present chapter two types of fact about singing have been considered. One is the choice of acoustic characteristics of vowel sounds that singers learn to adopt and that represent typical deviations from normal speech. Three examples of those characteristics have been discussed: (1) the pitch-dependent choice of formant frequencies that concerns high-pitched singing primarily in female voices, (2) the "singer's formant" that typically occurs in all voiced sounds in the male singing voice; and (3) the vibrato that occurs in both male and female singing. In all three examples we have strong reasons to assume that they serve a specific purpose. The pitch-dependent formant frequencies as well as the singer's formant are both resonatory phenomena that increase the audibility of the singer's voice when the orchestral accompaniment is loud. As resonatory phenomena occur independently of vocal effort, the increase in audibility is gained without expenses in terms of vocal effort; hence, a likely purpose in both these cases is vocal economy. The vibrato serves the purpose of allowing the singer a greater freedom in the choice of fundamental frequency as it eliminates beats with the sound of the accompaniment. Thus, in these three cases we see that *singing*

*differs from speech in a highly adequate manner.* It is tempting to speculate that such characteristics have developed as a result of evolution; the singers who developed them became successful, and hence their technique was copied by other singers.

A second kind of fact about singing discussed in this chapter is the acoustic correlates of various voice classifications that can be assumed to be based on perception. Such classifications are not only tenor, baritone, and bass, etc., but also vocal effort (e.g., piano, mezzopiano, etc.) and register. We have seen that in most of these cases it was hard to find a common acoustic denominator, because the acoustic characteristics of the categories vary with vowel and fundamental frequency. Rather, the common denominator exists within the voice organ. In the case of the male voice classification—tenor, baritone, and bass—the characteristic differences in formant frequency could be assumed to result from morphological differences in the vocal tract. The same is true for vocal effort and register, because they reflect differences in the control and operation of the vocal folds. Therefore, we may say that these examples of voice classification seem to rely on the function of the voice organ rather than on the acoustic properties of voice sounds. This is probably revealing as to the way in which we perceive singing voices. We seem to interpret the sounds in terms of how the voice organ was used in producing the sounds.

With regard to artistic interpretation, it seems that it contains at least two different components. One is the marking of structural constituents of the music. This requirement on the performance of songs thus seems to be common to both speech and music, and probably, it has a perceptual background. The other component is the signaling of the emotional atmosphere underlying the text and the music. Even in this respect perception of singing seems closely related to perception of speech. The coding of emotions in speech and singing would be similar and probably founded on a "body language" for communication of emotions. If this is true, our acquaintance with human emotional behavior and particularly speech serves as a reference in our decoding of the emotional information in the sounds from a singer's voice.

## ACKNOWLEDGMENTS

Si Felicetti of the Department of Speech Communication, KTH (Royal Institute of Technology) Stockholm, is acknowledged for her expert assistance in typing and editing this chapter. Its preparation was in part supported by the following funds: the Swedish Council for Planning and Coordination of Research, the Swedish Council for Research in the Humanities and Social Sciences, and the Swedish Natural Science Research Council.

## REFERENCES

Ågren, K. Alt- och tenorröst och harmoniska intervall mellan dem. Thesis work in speech communication. Department of Speech Communication, KTH, Stockholm, 1976.

Ågren, K., & Sundberg, J. An acoustic comparison of alto and tenor voices. *Journal of Research in Singing,* 1978, *1* (3), 26–32.

Appelman, D. R. *The science of vocal pedagogy.* London: Indiana University Press, 1967.

Bartholomew, W. T. A physical definition of 'good voice quality' in the male voice. *Journal of the Acoustical Society of America*, 1934, *6*, 25–33.

Bjørklund, A. Analyses of soprano voices. *Journal of the Acoustical Society of America*, 1961, 33, 575–582.

Carlson, R., Fant, G., & Granström, B. Two-formant models, pitch, and vowel perception. In G. Fant & M. A. A. Tatham (Eds.), *Auditory analysis and perception of speech*. New York: Academic Press, 1975, pp. 55–82.

Cleveland, T. F. Acoustic properties of voice timbre types and their influence on voice classification. *Journal of the Acoustical Society of America*, 1977, *61*, 1622–1629.

Coleman, R. O. A comparison of the contributions of two voice quality characteristics to the perception of maleness and femaleness in the voice. *Journal of Speech and Hearing Research*, 1976, *19*, 168–180.

Colton, R. H. Spectral characteristics of the modal and falsetto registers, *Folia Phoniatrica*, 1972, *24*, 337–344.

Colton, R. H., & Hollien, H. Perceptual differentiation of the modal and falsetto registers. *Folia Phoniatrica*, 1973, *25*, 270–280.

Dmitriev, L., & Kiselev, A. Relationship between the formant structure of different types of singing voices and the dimensions of the supraglottal cavities. *Folia Phoniatrica*, 1979, *31*, 238–241.

Fant, G. *Acoustic theory of speech production*. The Hague: Mouton, 1960.

Fant, G. *Speech sounds and features*. Cambridge, Massachusetts: MIT Press, 1973.

Fant, G. Non-uniform vowel normalization. *Speech Transmission Laboratory, Quarterly Progress and Status Report*, 1975, No. 2–3, 1–19.

Flanagan, J. L. *Speech analysis, synthesis and perception*. Berlin and New York: Springer-Verlag, 1965.

Fónagy, I. Mimik auf glottaler Ebene. *Phonetica*, 1962, *8*, 209–219.

Fónagy, I. La mimique buccale. *Phonetica*, 1976, *33*, 31–44.

Gibian, G. L. Synthesis of sung vowels. *Quarterly Progress Report, Massachusetts Institute of Technology*, 1972, No. 104, 243–247.

Hollien, H. On vocal registers. *Journal of Phonetics*, 1974, *2*, 125–143.

Hollien, H., Keister, E., & Hollien, P. A. Experimental data on 'singer's formant.' *Journal of the Acoustical Society of America, Supplement 1*, 1978, *64*, S171 (Abstract).

Howie, J., & Delattre, P. An experimental study of the effect of pitch on the intelligibility of vowels. *The National Association of Teachers of Singing Bulletin*, 1962, *18:4*, 6–9.

Kotlyar, G. M., & Morozov, V. P. Acoustical correlates of the emotional content of vocalized speech. *Soviet Physics Acoustics*, 1976, *22*, 208–211.

Large, J. Towards an integrated physiologic-acoustic theory of vocal registers. *National Association of Teachers of Singing Bulletin*, February-March 1972, 18–36.

Large, J. Acoustic-perceptual evaluation of register equalization. *National Association of Teachers of Singing Bulletin*, October 1974, 20–41.

Large, J., Iwata, S., & von Leden, H. The primary register transition in singing. *Folia Phoniatrica*, 1970, *22*, 385–396.

Large, J., Iwata, S., & von Leden, H. The male operatic head register versus falsetto. *Folia Phoniatrica*, 1972, *24*, 19–29.

Large, J. & Shipp, T. The effect of certain parameters on the perception of vocal registers, *National Association of Teachers of Singing Bulletin*, October 1969, 12–15.

Lerman, J. W., & Duffy, R. J. Recognition of falsetto voice quality. *Folia Phoniatrica*, 1970, *22*, 21–27.

Lindblom, B., & Sundberg, J. Acoustical consequences of lip, tongue, jaw, and larynx movement. *Journal of the Acoustical Society of America*, 1971, *50*, 1166–1179.

Lindgren, H., & Sundberg, A., Grundfrekvensförlopp och falsksång. Thesis work, Department of Musicology, Stockholm University, 1972 (stencil).

Monsen, R. B., & Engebretson, A. M. Study of vibrations in the male and female glottal wave. *Journal of the Acoustical Society of America*, 1977, *62*, 981–993.

Morozov, V. P. Intellegibility in singing as a function of fundamental voice pitch. *Soviet Physics Acoustics*, 1965, *10*, 279–283.

Nábělek, I. V., Nábělek, A. K., & Hirsh, I. J. Pitch of tone bursts of changing frequency. *Journal of the Acoustical Society of America*, 1970, *48*, 536–553.

Nelson, H. D., & Tiffany, W. R. The intellegibility of song. *The National Association of Teachers of Singing Bulletin*, December 1968, 22–33.

Nordström, P-E. Female and infant vocal tracts simulated from male area functions. *Journal of Phonetics*, 1977, *5*, 81–92.

Plomp, R. Timbre as a multidimensional attribute of complex tones. In R. Plomp & G. F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing*. Leiden: Sijthoff, 1970. Pp. 397–414.

Plomp, R. Continuity effects in the perception of sounds with interfering noise bursts. Paper given at the *Symposium sur la Psychoacoustique musicale, IRCAM, Paris*, July 1977.

Rakowski, A. Pitch discrimination at the threshold of hearing. *Proceedings of the 7th International Congress on Acoustics, Budapest*, 1971, *3*, 373–376.

Ramsdell, D. A. The psycho-physics of frequency modulation. Thesis, Harvard University (not seen, reported by F. Winckel, In *Music, sound, and sensation*. New York: Dover, 1967).

Rasch, R. A. The perception of simultaneous notes such as in polyphonic music. *Acustica*, 1978, *40*, 21–33.

Rothenberg, M. The glottal volume velocity waveform during loose and tight voiced glottal adjustments. In A. Rigault & R. Charbonneau (Eds.), *Proceedings of the 7th International Congress of Phonetic Sciences, Montreal, 1971*. The Hague: Mouton, 1972. Pp. 380–388.

Rubin, H. J., Le Cover, M., & Vennard, W. Vocal intensity, subglottic pressure and airflow relationship in singers, *Folia Phoniatrica*, 1967, *19*, 393–413.

Russo, V., & Large, J. Psychoacoustic study of the Bel Canto model for register equalization: male chest and falsetto. *Journal of Research in Singing*, 1978, *1* (3), 1–25.

Rzhevkin, S. N. Certain results of the analysis of a singer's voice. *Soviet Physics Acoustics*, 1956, *2*, 215–220.

Scotto di Carlo, N. Etude acoustique et auditive des facteurs d'intelligibilité de la voix chantée. In A. Rigault & R. Charbonneau (Eds.), *Proceedings of the 7th International Congress of Phonetic Sciences, Montreal, 1971*. The Hague, Paris: Mouton, 1972. Pp. 1017–1023.

Scotto di Carlo, N. Influence de l'articulation sur la musicalité de la phrase chantée. *Travaux de l'Institut de Phonétique d'Aix*, 1976, No. 3, 117–146.

Seashore, C. E. *Psychology of music*. New York: McGraw-Hill, 1938 & New York: Dover, 1967.

Shonle, J. I. Perceived pitch of vibrato tones. *Journal of the Acoustical Society of America*, 1975, *58*, S132 (Abstract).

Simon, P., Lips, H., & Brock, G. Etude sonagraphique et analyse acoustique en temps reel de la voix chantée a partir de differentes techniques vocales. *Travaux de l'Institut de Phonétique de Strasbourg*, 1972, No. 4, 219–276.

Sirker, U. Objektive Frequenzmessung und subjektive Tonhöhenempfinung bei Musikinstrumentklängen. *Swedish Journal of Musicology*, 1973, *55*, 47–58.

Slawson, A. W. Vowel quality and musical timbre as functions of spectrum envelope and fundamental frequency. *Journal of the Acoustical Society of America*, 1968, *43*, 87–101.

Smith, L. A., & Scott, B. L. Increasing the intellegibility of sung vowels. *Journal of the Acoustical Society of America*, 1980, *67*, 1795–1797.

Stevens, S. S., & Davis, H. *Hearing, its psychology and physiology*. New York: Wiley, 1938.

Stumpf, C. *Die Sprachlaute*. Berlin and New York: Springer-Verlag, 1926.

Sundberg, J. Formant structure and articulation of spoken and sung vowels. *Folia Phoniatrica*, 1970, *22*, 28–48.

Sundberg, J. Production and function of the 'singing formant'. In H. Glahn, S. Sørensen, & P. Ryom (Eds), *Report of the 11th Congress of the International Musicological Society*. Copenhagen: Editor Wilhelm Hansen, 1972, 679–686. (a)

Sundberg, J. Pitch of synthetic sung vowels. *Speech Transmission Laboratory, Quarterly Progress and Status Report*, 1972, No. 1, 34–44. (b) Later revised and published as Effects of the vibrato and the 'singing formant' on pitch. *Musicologica Slovaca* (in Memoriam M. Filip) 1978, *6*, 51–69. (b)

Sundberg, J. The source spectrum in professional singing. *Folia Phoniatrica*, 1973, *25*, 71–90.

Biblioteca del Real Conservatorio Superior de Música - MADRID -

Sundberg, J. Articulatory interpretation of the 'singing formant'. *Journal of the Acoustical Society of America*, 1974, *55*, 838–844.

Sundberg, J. Formant technique in a professional female singer. *Acustica*, 1975, *32*, 89–96.

Sundberg, J. Vibrato and vowel identification. *Archives of Acoustics (Polish Academy of Sciences)*, 1977, *2*, 257–266. (a)

Sundberg, J. Singing and timbre. In *Music room acoustics*. Stockholm: Royal Swedish Academy of Music Publications (Volume 17), 1977. Pp. 57–81. (b)

Sundberg, J. Studies of the soprano voice. *Journal of Research in Singing*, 1977, *1*, (1), 25–35. (c)

Sundberg, J. *Musikens ljudlära*. (2nd ed.) Stockholm: Proprius Förlag, 1978. (a)

Sundberg, J. Synthesis of singing. *Swedish Journal of Musicology*, 1978, *60*. (b)

Sundberg, J. & Gauffin, J. Waveform and spectrum of the glottal voice source. *Speech Transmission Laboratory, Quarterly Progress and Status Report*, 1978, No. 2–3, 35–50.

Terhardt, E., On the perception of periodic sound fluctuations (roughness). *Acustica*, 1974, *30*, 201–213.

Weait, C., & Shea, J. B. Vibrato: an audio-video-fluorographic investigation of a bassoonist. *Applied Radiology*, January-February 1977, *6:1*, 89–90, 148.

Vennard, W. *Singing, the mechanism and the technic*. (2nd ed.) New York: Fischer, 1967.

Vennard, W., Hirano, M., Ohala, J., & Fritzell, B. A series of four electromyographic studies. *National Association of Teachers of Singing Bulletin*, October 1970, 16–21; December 1970, 30–37; February-March 1971, 26–32; May-June 1971, 22–30.

Winckel, F. Physikalische Kriterien für objektive Stimmbeurteilung. *Folia Phoniatrica*, 1953, *5 Separatum*, 231–252.

Winckel, F. *Music, Sound and sensation. A modern exposition*. New York: Dover Publications, 1967.

Zwicker, E., & Feldtkeller, R. *Das Ohr als Nachrichtenempfänger*. (2nd ed.) Stuttgart: Hirzel Verlag, 1967.

# 4

# Grouping Mechanisms in Music

*Diana Deutsch*


I. Introduction ..... 99
II. Grouping Principles ..... 100
III. Two-Channel Listening to Melodic Sequences ..... 101
A. The Scale Illusion ..... 102
B. Temporal Relationships as Determinants of Grouping ..... 104
C. The Octave Illusion ..... 108
D. Handedness Correlates ..... 114
E. Melody Perception from Phase-Shifted Tones ..... 116
F. Discussion ..... 117
IV. Channeling of Rapid Sequences of Single Tones ..... 118
A. Grouping by Frequency Proximity ..... 118
B. Temporal Coherence as a Function of Frequency Proximity and Tempo ..... 119
C. Grouping by Frequency Proximity Builds with Repetition ..... 120
D. Grouping by Frequency Proximity and the Perception of Temporal Relationships ..... 122
E. Grouping by Good Continuation ..... 124
F. Grouping by Timbre ..... 124
G. Grouping by Amplitude ..... 126
H. Grouping by Temporal Proximity ..... 126
I. Perceptual Replacement of Missing Sounds ..... 127
V. Voluntary Attention ..... 127
VI. Conclusion ..... 130
References ..... 130


## I. INTRODUCTION

Music presents us with a complex, rapidly changing acoustic spectrum, often resulting from the superposition of sounds from many different sources. The primary task that our auditory system has to perform is to interpret this spectrum in terms of the behavior of external objects. This is analogous to the task performed by the visual


Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

system when it interprets the mosaic of light patterns impinging on the retina in terms of objects producing them (Gregory, 1970; Sutherland, 1973). Such a view of perception as a process of "unconscious inference" was proposed in the last century by Helmholtz (see Hemholtz, 1925), and we shall see that many phenomena of music perception are readily interpretable in this light.

Issues concerning organizational processes in music divide themselves basically into two. The first is the following. Given that we are presented with a set of first-order acoustic elements, how are these combined so as to form separate groupings? If all first-order elements were indiscriminately linked together, auditory shape-recognition operations could not be performed. There must, therefore, exist a set of mechanisms that permits the formation of simultaneous and sequential linkages between certain elements, and inhibits the formation of such linkages between others. Simple mechanisms underlying such linkages are explored in the present chapter. Second, we may enquire into the ways in which higher order abstractions are derived from combinations of first-order elements so as to lead to perceptual equivalences and similarities. This issue is explored in the next chapter, and it is assumed that such abstractions also form bases for grouping.

## II. GROUPING PRINCIPLES

There are two basic questions involved in considering the mechanisms involved in grouping musical stimuli into configurations. The first concerns the stimulus *attributes* along which grouping principles operate. When presented with a complex sequence, our auditory system may group stimuli according to some rule based on the frequencies of its components, on their amplitudes, on the spatial locations from which they emanate, or on the basis of some complex attribute such as timbre. As we shall see, all these attributes can function as bases for organization and grouping in music. Furthermore, the principles determining what attribute is followed for any given sequence are both complex and rigid. We shall see, for example, that with one type of sequence, organization on the basis of frequency uniformly occurs; yet given a slight change in this sequence, organization on the basis of spatial location may occur instead. Such differences in organization can be interpreted in terms of strategies most likely to lead to the correct conclusions in interpreting our environment.

Second, we may pose the following question: Assuming that organization takes place on the basis of some dimension such as frequency, what are the *principles* governing grouping along this dimension? The Gestalt psychologists proposed that we group stimuli into configurations on the basis of various simple principles (Wertheimer, 1923). One of these is the principle of Proximity, which states that nearer elements are grouped together in preference to those that are spaced farther apart. An example of this principle is shown on Figure 1A, where the closer dots appear to be grouped together in pairs. Another is the principle of Similarity, which is illustrated on Figure 1B. Here, configurations are formed out of like elements, so that we perceive one set of vertical rows formed by the filled circles and another set formed by the unfilled circles. A third is the principle of Good Continuation, which states that elements that

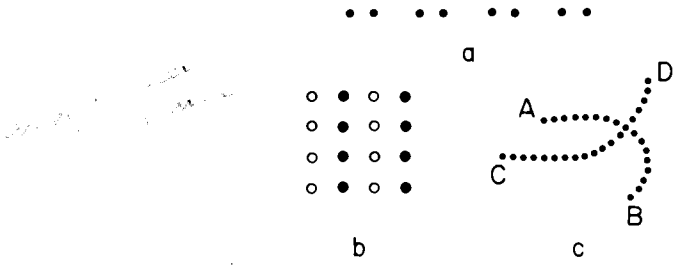



**Fig. 1.** Illustrations of the Gestalt principles of Proximity, Similarity, and Good Continuation.

follow each other in a given direction are perceived together. For instance, we perceptually group the dots in Figure 1C so as to form the two lines AB and CD. A fourth principle, known as Common Fate, states that elements which move in the same direction are perceived together.

It has been demonstrated that these principles are important determinants of grouping in visual arrays, and we shall see that this is true of music also. It seems reasonable to suppose—as argued by Gregory (1970), Sutherland (1973), Hochberg (1974), and Bregman (1978)—that grouping in conformity with such principles enables us to interpret our environment most effectively. To give some examples, in the case of vision, proximal elements are most likely to belong to the same object than elements that are spaced farther apart. The same line of reasoning holds for similar elements compared with those that are dissimilar. In the case of hearing, similar sounds are likely to be emanating from the same source and different sounds from different sources. A sound sequence that changes smoothly in frequency is likely to be emanating from a single source. Components of a complex sound spectrum that rise and fall in synchrony are also likely to be emanating from the same source.

One more point should be made before reviewing the experimental evidence. When we hear a tone, we attribute a fundamental pitch, a loudness, a timbre; and we hear the tone at a given location. Each tonal percept may therefore be described as a bundle of attribute values. If our perception is veridical, this bundle reflects the location and characteristics of the sound emitted. We shall see, however, that in situations where more than one tone is presented at a time, these bundles of attribute values may fragment and recombine in other ways, so that illusory percepts result. Perceptual grouping in music is therefore not simply a matter of linking different sets of stimuli together; rather it involves a process whereby these stimuli are fragmented into their separate attributes, followed by a process of perceptual synthesis in which the different attribute values are recombined.

Biblioteca del Real Conservatorio Superior de Música - MADRID -

## III. TWO-CHANNEL LISTENING TO MELODIC SEQUENCES

The two-channel listening technique is particularly useful for studying organizational processes in music, since it enables different attributes to be set in opposition to

each other as bases for grouping. For example, grouping by spatial location may be set in opposition to grouping by frequency or by amplitude. Similarly, different principles governing grouping along a given dimension may be set in opposition to each other; for example, the principle of Proximity may be opposed to the principle of Good Continuation. The experiments to be described show that the nature of the stimulus configuration critically determines what grouping principle is adopted, and indicate that there are complex and rigid rules of precedence for these principles.

### A. The Scale Illusion

The configuration that produced the scale illusion is shown in Fig. 2A. It can be seen that this consisted of a major scale, presented simultaneously in both ascending and descending form. When a tone from the ascending scale was delivered to one ear, a tone from the descending scale was simultaneously delivered to the other ear, and successive tones in each scale alternated from ear to ear (Deutsch, 1975b).

This sequence was found to give rise to various illusory percepts. The majority of listeners perceived two melodic lines, a higher one and a lower one, that moved in contrary motion. Further, the higher tones all appeared to be emanating from one earphone and the lower tones from the other (Fig. 2B). When the earphone positions were reversed, there was often no corresponding change in the percept. So, it appeared to the listener that the earphone that had been producing the higher tones was now producing the lower tones, and that the earphone that had been producing the lower tones was now producing the higher tones. A minority of listeners heard instead only a single stream of four tones, corresponding to the higher tones in the sequence, and little or nothing of the rest of the sequence was perceived.

So in considering what stimulus *attribute* was here used as a basic for grouping, we find that organization by spatial location never occurred; rather, organization was

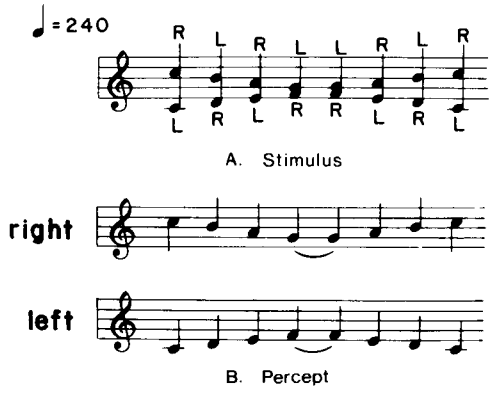



**Fig. 2.** (A) Representation of the configuration producing the scale illusion. This basic pattern was repetitively presented 10 times without pause. (B) Representation of the illusory percept most commonly obtained (from Deutsch, 1975b).

always on the basis of frequency. Second, in considering what grouping *principle* was adopted, we find that organization was always on the basis of frequency priximity. Listeners heard either two melodic lines, one corresponding to the higher tones and the other to the lower, or they heard the higher tones alone. No listener reported a full ascending or descending scale as a component of the sequence; so that grouping by Good Continuation never occurred.

Butler (1979a) has demonstrated that these findings may be extended to a broad range of musical situations. He presented the configuration shown in Fig. 2A to music students through spatially-separated loudspeakers, rather than earphones, in a free sound-field environment. The listeners notated separately the sequence that they heard as emanating from the speaker on their right and the sequence that they heard as emanating from the speaker on their left. In some conditions the stimuli were generated on a piano. Further, timbral and loudness differences were sometimes introduced between the stimuli presented through the different speakers. Butler found that despite these variations, virtually all responses reflected channeling by frequency proximity, so that higher and lower melodic lines were perceived, each apparently emanating from a different speaker. A further interesting finding was that when timbral differences were introduced between the tones presented through the two speakers, a new tone quality was perceived, but it seemed to be emanating simultaneously from both speakers. So, not only were the spatial locations of the tones perceptually rearranged in accordance with frequency proximity, but their timbres were rearranged also.

To determine whether these findings generalize to other melodic configurations, Butler presented listeners with the two-part contrapuntal patterns shown in Figs. 3a and 3b. Virtually all responses again reflected grouping by frequency range. For both configurations a perceptual reorganization occurred, so that a melody corresponding to the higher tones appeared to be emanating from one earphone or speaker and a melody corresponding to the lower tones from the other (Figs. 3c and 3d).

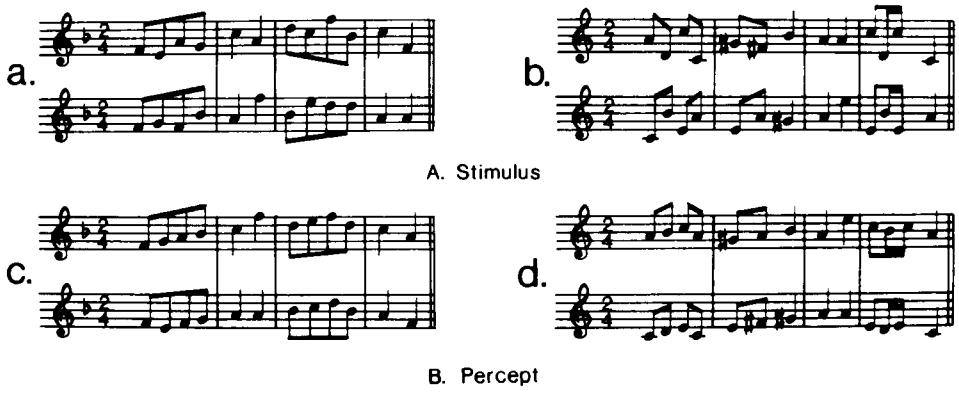



**Fig. 3.** (A) Two-part melodic patterns as presented to subjects through left and right earphones or speakers. (B) The patterns as most commonly notated by the subjects (from Butler, 1979a).

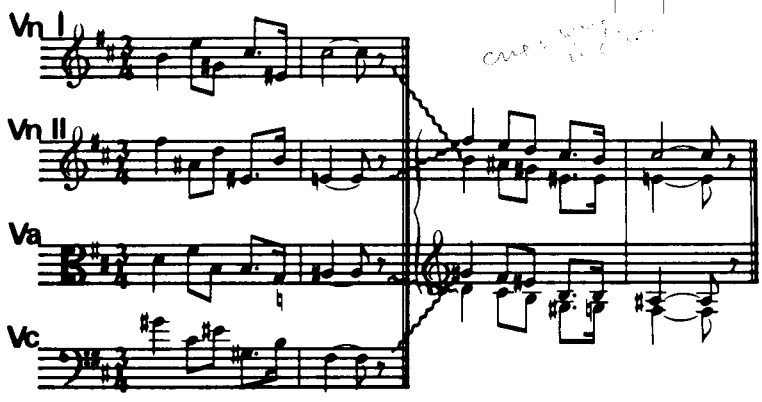



**Fig. 4.** Passage from the final movement of Tschaikowsky's Sixth (*Pathetique*) Symphony. The combination of the Violin I and Violin II melodies produces the percept shown on the upper right. The combination of the viola and violincello melodies produces the percept shown on the lower right (from Butler, 1979b).

Butler (1979b) further drew attention to an interesting passage from the final movement of Tschaikowsky's Sixth (*Pathetique*) Symphony. As shown in Fig. 4, the theme and accompaniment are each distributed between the two violin parts. However, the theme is heard as coming from one set of instruments and the accompaniment as from the other. Whether it was Tschaikovsky's intention to produce a perceptual illusion here, or whether he expected the listener to hear the theme and accompaniment waft from one set of instruments to the other, we may never know!

How should such gross mislocalization effects be explained? Our acoustic environment is very complex, and the assignment of sounds to their sources is made difficult by the presence of echoes and reverberation (Benade, 1976). So, when a sound mixture is presented such that both ears are stimulated simultaneously, it is unclear from first-order localization cues alone which components of the total spectrum should be assigned to which source. Other factors must also operate to provide cues concerning the sources of these different sounds. One such factor is similarity of frequency spectrum. Similar sounds are likely to be emanating from the same source and different sounds from different sources. Thus, with these musical examples it becomes reasonable for the listener to assume that tones in one frequency range are emanating from one source, and tones in another frequency range from a different source. We therefore reorganize the tones perceptually on the basis of this interpretation (Deutsch, 1975a).

### B. Temporal Relationships as Determinants of Grouping

Given the above line of reasoning, we should expect perceptual grouping of simultaneous sequences to be strongly influenced by the salience of first-order localization cues. Under the conditions we have been considering such localization cues were

weak, since input was always to both ears simultaneously. However, under conditions where such cues are strong and unambiguous, organization by spatial location should be expected to take precedence over organization by frequency proximity. This should be the case, for instance, where the signals to the two ears are clearly separated in time.

An experiment was therefore performed to examine perceptual grouping as a function of the temporal relationships between the signals arriving at the two ears (Deutsch, 1978a, 1979). Listeners identified melodic patterns in which the component tones switched between the ears. Conditions were compared where input was to one ear at a time and where input was to the two ears simultaneously. Such simultaneity of input was achieved by presenting a drone to the ear opposite the ear receiving the component of the melody. In order to control for the effect of the drone apart from its providing a simultaneous input to the opposite ear, a further condition was included in which the drone and the melody component were presented to the same ear. In a fourth condition the melody was presented binaurally.

It was predicted that in the conditions in which input was to one ear at a time, identification of the melodic patterns should be difficult, reflecting perceptual grouping by spatial location. However, in the conditions where both ears received input simultaneously, identification of the melodic patterns should be much easier, and should reflect organization by frequency proximity in the presence of the contralateral drone.

The experiment employed the two melodic patterns shown in Fig. 5. On each trial, one of these patterns was presented ten times without pause, and listeners identified on forced choice which of these had been presented.

The four conditions of the experiment, together with the error rates in each, are shown in Fig. 6. It can be seen that when the melody was presented to both ears at the same time, identification performance was excellent. However, when the component tones of the melody switched from ear to ear with no accompanying drone, a severe performance decrement occurred. Yet when the drone was presented to the ear *opposite* the ear receiving the melody component, the performance level was again very high, even though the melody was still switching from ear to ear. This result cannot be attributed to processing the harmonic relationships between the drone and the melody

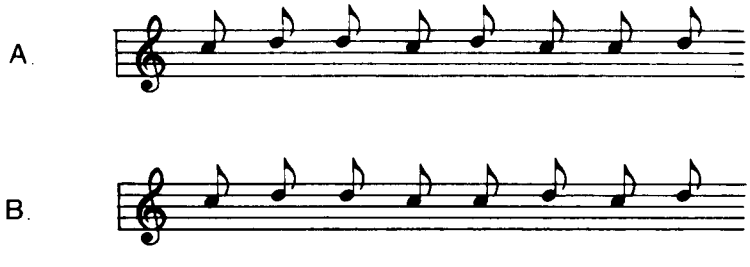






**Fig. 5.** Basic melodic patterns employed in experiment to study the effects on melody identification of rapid switching between ears. All tones were 30 msec in duration, and tones within a sequence were separated by 100-msec pauses (Deutsch, 1979).

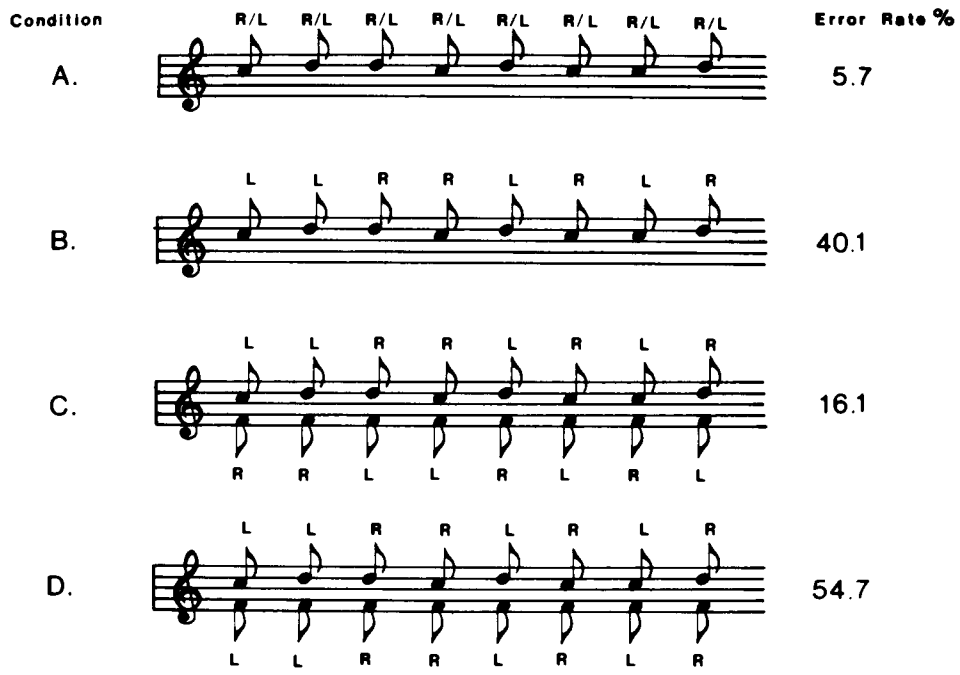



**Fig. 6.** Examples of distributions between ears of melodic pattern and drone in different conditions of the experiment. See text for details (from Deutsch, 1979).

components because when the drone was presented to the *same* ear as the one receiving the component of the melody, performance was below chance.

This experiment therefore demonstrates that with tones coming from different spatial locations, temporal relationships between them are important determinants of grouping. When signals are coming from two locations simultaneously, it is easy to integrate the information arriving at the two ears into a single perceptual stream. But when the signals coming from the two locations are clearly separated in time, subjective grouping by spatial location is so powerful as to prevent the listener from combining the signals to produce an integrated percept.

A related experiment comparing the effects of simultaneity with nonsimultaneity of input to the two ears was performed by Judd (1979). In this experiment listeners were presented with four-tone melodic patterns whose components alternated from ear to ear. Judd found that presenting noise to the ear contralateral to the ear receiving the melody component resulted in enhanced recognition performance. He also proposed an interpretation in terms of competing channeling mechanisms, reasoning that the strong localization cues present in the no-noise condition induced channeling by spatial location, and that the weaker localization cues due to the noise resulted in channeling by frequency proximity instead.

In the study by Deutsch (1979) the effects of onset–offset asynchrony between the

tones arriving at the two ears were also examined. Such temporal overlaps between signals commonly occur in normal listening, and it was predicted that results here should be intermediate between those where the input to the two ears was strictly simultaneous and those where these inputs were clearly separated in time. Such findings were indeed obtained: performance levels under conditions of asynchrony were significantly worse than where the melody components and the contralateral drone were strictly simultaneous, yet significantly better than were the melody switched between ears without an accompanying drone. This is as expected from the present line of reasoning. Temporal similarities in the waveform envelopes of two simultaneous signals are important indicators that these signals are emanating from the same source (following the "principle of Common Fate"), and discrepancies would indicate that the signals are emanating from different sources (Tobias, 1972). We should therefore expect that asynchronies between the signals arriving at the two ears would increase the tendency to treat these signals as emanating from different sources, and so permit less integration of the pattern distributed between the two ears.

Two related experiments on the effects of asynchrony should here be cited. Rasch (1978) investigated the threshold for perception of a high tone when it was accompanied by a low tone. He found that delaying the onset of the low tone relative to the high tone produced a substantial lowering of threshold. Further, under conditions of asynchrony the subjective percept was strikingly altered, so that the two tones stood apart clearly rather than being fused into a single percept. Rasch advanced an interpretation of his findings along lines very similar to those presented here.

Rasch also noted that although temporal asynchrony had strong perceptual effects, it was not recognized as such by the listeners. The same observation was made in the present experiment. In the asynchronous conditions, one obtained the subjective impression of a "plopping" sound at the onset and offset of the tones, but it was difficult to describe this percept further. The strong effect found here due to temporal asynchrony was therefore not based on conscious inference.

Another related experiment is that of Bregman and Pinker (1978). These authors presented a simultaneous two-tone complex in alternation with a third tone, and introduced various conditions of onset–offset asynchrony between the simultaneous tones. They found that with increasing asynchrony there was an increased likelihood that one of the simultaneous tones would form a melodic stream with the third tone. Bregman and Pinker reasoned that asynchrony between the simultaneous tones resulted in a decreased tendency for these tones to be treated as emanating from the same source, and so facilitated a sequential organization by frequency proximity between one of these simultaneous tones and the alternating tone.

These various experiments on the effects of asynchrony bear on an issue that was raised a century ago by von Helmholtz. In his book *On the Sensations of Tone* (1885), he posed the question of how, given the complex, rapidly changing spectrum produced by several instruments playing simultaneously, we are able to reconstruct our musical environment so that some components fuse to produce a single sound impression, while others are heard as separate melodic lines which may be simultaneously per-

ceived. For the latter instance, he posed the further question as to the basis on which such simultaneous melodic lines are constructed. Thus he wrote:

> Now there are many circumstances which assist us first in separating the musical tones arising from different sources, and secondly, in keeping together the partial tones of each separate source. Thus when one musical tone is heard for some time before being joined by the second, and then the second continues after the first has ceased, the separation in sound is facilitated by the succession of time. We have already heard the first musical tone by itself, and hence know immediately what we have to deduct from the compound effect for the effect of this first tone. Even when several parts proceed in the same rhythm in polyphonic music, the mode in which the tones of different instruments and voices commence, the nature of their increase in force, the certainty with which they are held, and the manner in which they die off, are generally slightly different for each . . . but besides all this, in good part music, especial care is taken to facilitate the separation of the parts by the ear. In polyphonic music proper, where each part has its own distinct melody, a principal means of clearly separating the progression of each part has always consisted in making them proceed in different rhythms and on different divisions of the bars (p. 59).

And later:

> All these helps fail in the resolution of musical tones into their constituent partials. When a compound tone commences to sound, all its partial tones commence with the same comparative strength; when it swells, all of them generally swell uniformly; when it ceases, all cease simultaneously. Hence no opportunity is generally given for hearing them separately and independently (p. 60).

### C. The Octave Illusion

In the experiments so far described, channeling by frequency proximity was the rule when information was presented to both ears simultaneously. Channeling by spatial location occurred only with temporal separations between the stimuli presented to the two ears. We now examine conditions where channeling by spatial location occurred even though the input to the ears was strictly simultaneous. We shall see that this principle was adopted under special conditions of frequency relationship between the tones as they were presented in sequence at the two ears.

One stimulus configuration that induced melodic channeling by spatial location is shown in Fig. 7A. It can be seen that this consisted of two tones that were spaced an octave apart and repeatedly presented in alternation. The identical sequence was delivered to the two ears simultaneously; however, when the right ear received the high tone the left ear received the low tone, and vice versa. So, essentially the configuration was that of a two-tone chord, where the ear of input for each component switched repeatedly (Duetsch, 1974a,b, 1975a).

This sequence was found to produce various illusions, the most common of which is shown on Fig. 7B. It can be seen that this consisted of a single tone that alternated from ear to ear, and whose pitch also alternated from one octave to the other in synchrony with the localization shift. When the earphones were placed in reverse position, most listeners found that the apparent locations of the high and low tones remained fixed. So it seemed to these listeners that the earphone that had been

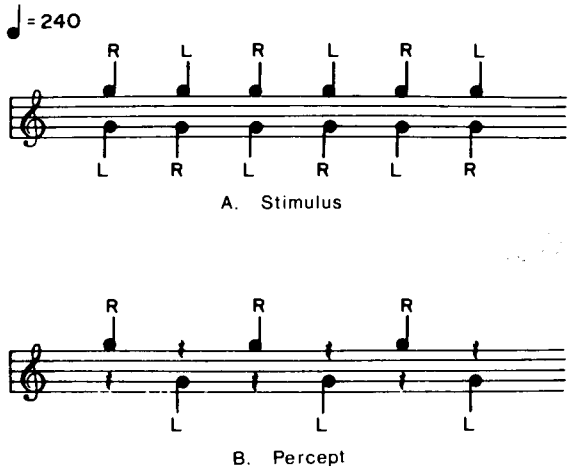



**Fig. 7.** (A) Representation of the configuration producing the octave illusion. (B) Representation of the illusory percept most commonly obtained (from Deutsch, 1974b).

producing the high tones was now producing the low tones, and that the earphone that had been producing the low tones was now producing the high tones.

It was hypothesized that this illusion results from the operation of two different selection mechanisms underlying the pitch and the localization percepts. To provide the perceived sequence of pitches (i.e., the melodic line) the frequencies arriving at one ear are followed, and those arriving at the other ear are suppressed. However, each tone is localized at the ear receiving the higher frequency signal, regardless of whether the higher or the lower frequency is in fact perceived (Deutsch, 1975a). This model was confirmed in later experiments (Deutsch, 1976, 1978b; Deutsch & Roll, 1976).

We can next ask whether the interactions giving rise to the octave illustion take place between pathways specific to information from the two ears, or whether instead pathways conveying information from different regions of auditory space are involved. In order to investigate this question, the stimuli were presented through spatially separated loudspeakers rather than earphones. It was found that the analogous illusion was obtained: A high tone apparently emanating from one speaker was perceived as alternating with a low tone apparently emanating from the other speaker (Deutsch, 1975a).

Setting aside the issue of where the tones appeared to be located in the octave illusion, and considering only what sequence of pitches was perceived, we note that here channeling was always on the basis of spatial location. This stands in sharp contrast to findings on the scale illusion, in which channeling was always on the basis of frequency proximity instead. Yet the stimuli producing these illusions were remarkably similar. In both cases listeners were presented with repeating sequences of sine wave tones at equal amplitudes and durations. In both cases the stimuli were continuous, and the frequencies presented to one ear always differed from the fre-

quencies simultaneously presented to the other ear. Yet these two configurations gave rise to radically different channeling strategies. It is especially noteworthy that when two tones that were separated by an octave were simultaneously presented in the scale illusion, both tones were generally perceived (Fig. 2B). Yet when two tones that were separated by an octave were simultaneously presented in the octave illusion, only one of these was generally perceived (Fig. 7B). These differences in channeling strategy must therefore have resulted from differences in the patterns of frequency relationship between successive tones.

Considering the octave illusion further, we find that here the frequency emanating from one side of space was always identical to the frequency that had just emanated from the opposite side. It was therefore hypothesized that this factor was responsible for inducing melodic channeling by spatial location for this configuration. A further set of experiments was performed to test this hypothesis (Deutsch, 1980, 1981).

In the first experiment listeners were presented with sequences consisting of 20 dichotic chords. Two conditions were compared, which employed the basic configurations shown on Fig. 8A. The configuration in Condition 1 consisted of the repetitive presentation of a single chord, whose components stood in octave relation and alternated from ear to ear such that when the high tone was in the right ear the low tone was in the left ear, and vice versa. It can be noted that here the two ears received the same frequencies in succession. On half the trials the sequence presented to the right ear began with the high tone and ended with the low tone, and on the other half this order was reversed. On each trial, subjects judged whether the sequence began with the high tone and ended with the low tone, or whether it began with the low tone and ended with the high tone; and from these judgments it was inferred which ear was being followed for pitch.

The basic configuration in Condition 2 consisted of the repetitive presentation of two dichotic chords in alternation, the first forming an octave and the second a minor third, so that the entire four-tone combination constituted a major triad. It will be noted that here the two ears did not receive the same frequencies in succession. On half the trials the right ear received the upper component of the first chord and the lower component of the last chord, and on the other half this order was reversed.

The amplitude relationships between the tones presented to the two ears were systematically varied across trials, and the extent to which each location was followed was plotted as a function of these amplitude relationships. The results are shown on Fig. 8B. It can be seen that in Condition 1 the frequencies presented to one location were followed until a critical level of amplitude relationship was reached, and the other location was followed beyond this level. However, in Condition 2 there was no following on the basis of spatial location, even when the signals presented to the two locations differed substantially in amplitude. However, hypothesizing instead that the subjects were following this sequence on the basis of frequency proximity a very consistent result emerged: Three of the subjects consistently followed the low tones, and one consistently followed the high tones. This result is as expected if the critical factor responsible for channeling by spatial location here is that the same frequencies emanate in succession from different regions of auditory space.

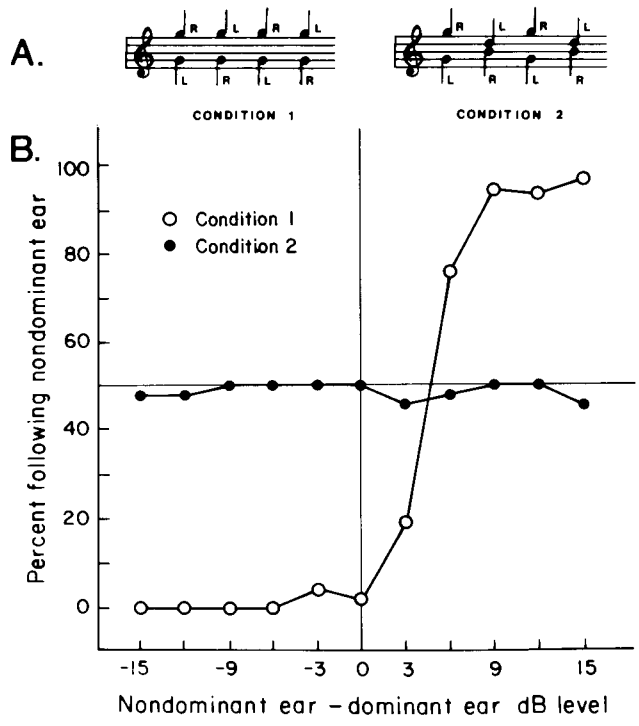



**Fig. 8.** (A) Configurations used in first experiment investigating the factors producing following on the basis of spatial location. (B) Percent following of nondominant ear in the two conditions of the experiment (from Deutsch, 1980a).

In a second experiment only two dichotic chords were presented on each trial. Two conditions were again compared, which employed the basic configurations shown on Fig. 9A. The configuration in Condition 1 consisted of two presentations of the identical chord, whose components formed an octave, such that one ear received first the high tone and then the low tone, while simultaneously the other ear received first the low tone and then the high tone. The identical frequencies were used throughout this condition. The basic configuration in Condition 2 consisted of two dichotic chords, each of which formed an octave, but which were composed of different frequencies. Trials employing chords composed of C and F♯ and of A and D♯ occurred in strict alternation, so that any given chord was repeated only after a substantial time period during which several other chords were interpolated.

Figure 9B shows the extent to which each location was followed in these two conditions, as a function of the amplitude relationships between the signals at the two ears. It can be seen that in Condition 1 following was clearly on the basis of spatial location. However, in Condition 2 such following did not occur, even when there were substantial amplitude differences between the signals at the two ears. Instead the subjects consistently followed these sequences on the basis of overall contour: Their

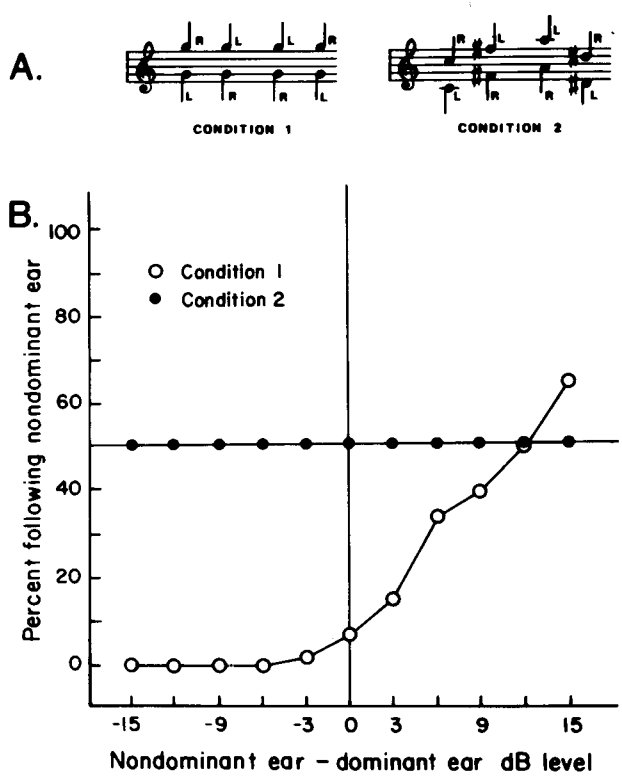



**Fig. 9.** (A) Configurations used in second experiment investigating the factors producing following by spatial location. (B) Percent following of nondominant ear in the two conditions of the experiment (from Deutsch, 1980a).

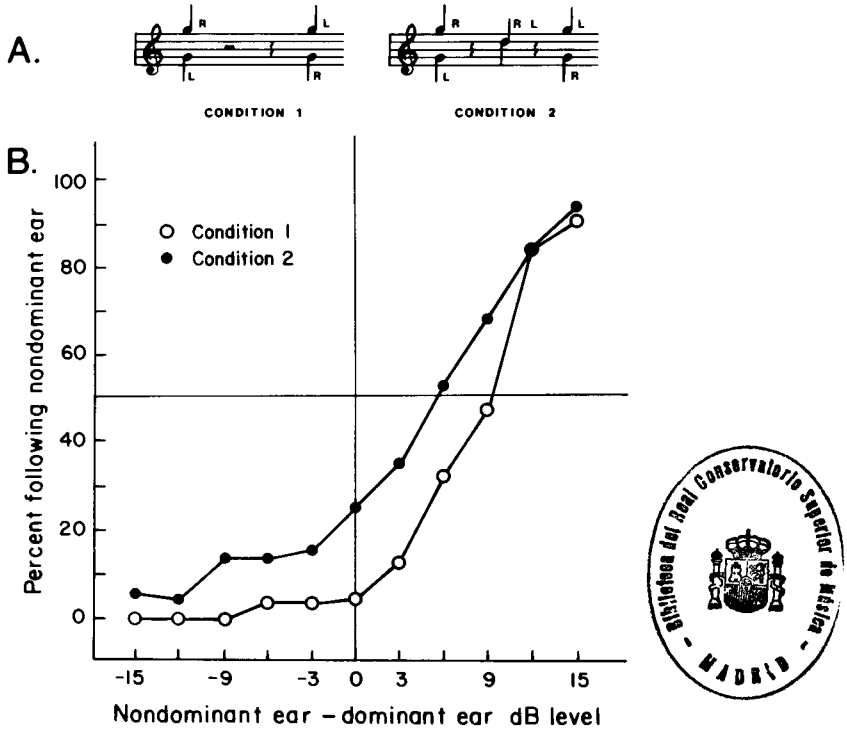
patterns of response indicated an ascending sequence when the second chord was higher than the first, and a descending sequence when the second chord was lower than the first. This was true even when the signals at the two ears differed substantially in amplitude.

It should be noted that following by contour here was consistent with following by frequency proximity, but that following the middle tones of the sequence (for example, the high C and the low F♯ in Fig. 9A) would have been equally consistent, since in either case the melodic interval formed by the two tones was a diminished fifth. So, in these sequences, overall contour was the factor that determined which tones were followed. Whether this served simply to "break a tie" or whether contour would win out in competition with frequency proximity in other configurations remains to be determined.

At all events, the second experiment showed, as did the first, that following by spatial location occurs in sequences where the same frequencies emanate in succession from two different regions of auditory space, and it occurs on other lines when this

relationship does not hold. It is particularly interesting to note that relative amplitude was found not to be an important factor in either experiment. When following was by frequency proximity or by contour, this occurred in the face of substantial amplitude differences between the signals arriving at the two ears. When following was by spatial location, the switch from following one side of space to the other did not occur at the point where the amplitude balance shifted from one side of space to the other, but at a different level of amplitude relationship. Thus, amplitude here appeared to *set the scene* for following on the basis of spatial location rather than serving as a primary following principle. [It will be recalled that in the experiment by Butler (1979a) following by frequency proximity also occurred in the face of amplitude differences between signals emanating from different spatial locations.]

We can next ask whether the lack of following by spatial location in the second conditions of these two experiments was due to the delay between successive presentations of the same frequencies to the two locations or to the interpolation of tones of different frequencies. To examine the effect of interpolated information, we studied performance under two further conditions. As shown on Fig. 10A, the configuration




Biblioteca del Real Conservatorio Superior de Música - MADRID -

**Fig. 10.** (A) Configurations used in third experiment investigating the factors producing following by spatial location. (B) Percent following of nondominant ear in the two conditions of the experiment (from Deutsch, 1980a).

in these two conditions were identical, except that in Condition 2 a single tone was interpolated between the dichotic chords and the listeners were instructed to ignore this tone. As shown on Fig. 10B, there was a less pronounced following of the preferred spatial location in the condition in which the extra tone was interpolated. To investigate the effects of temporal dealy, we varied the time interval between the onsets of successive tones at the two ears. This was achieved either by changing the durations of the tones or by interpolating gaps between them. It was found that the strength of this effect decreased with increasing time between onsets of the identical frequencies at the two locations. It did not matter whether this increase was produced by lengthening the durations of the tones, or by interpolating gaps between them. Thus, both interpolated information and temporal delay were found to reduce channeling be preferred spatial location.

These experiments raise the question of why such a channeling mechanism should have developed. We may hypothesize that this mechanism enables us to follow new, ongoing auditory information with a minimum of interference from echoes or reverberation. In everyday listening, when the same frequency emanates successively from two different regions of auditory space, the second occurrence may well be due to an echo. This is made more probable as the delay between the onsets of these two occurrences is shortened. However, if different frequencies are interpolated between two occurrences of the identical frequency, other interpretations of the second occurrence are made more likely. We may therefore hypothesize that this falls into the class of mechanisms that act to counteract misleading effects of echoes and reverberation. Such an explanation has been advanced, for instance, for the precedence effect. Wallach, Newman and Rosenweig (1949) have reported that in listening to music a single image may be obtained with the waveform presented at two different spatial locations separated by 45–70 msec intervals. The second sound is, under these conditions, attributed to the same location as the first. Analogous findings have been reported by Haas (1951).

### D. Handedness Correlates

Strong handedness correlates have been obtained for both the octave and the scale illusions. In the case of the octave illusion, there was a strong tendency among right-handers to hear the high tone on the right and the low tone on the left. This was not, however, found among left-handers (Deutsch, 1974). From further studies it was concluded that these findings reflected a tendency to perceive the pattern of frequencies presented to the dominant side of auditory space rather than the nondominant (Deutsch, 1975a, 1981; Deutsch & Roll, 1976). In the case of the scale illusion, there was also a strong tendency among right-handers to hear the higher tones on the right and the lower tones on the left; again this was not true of left-handers. Here the mislocalization of the higher tones to one spatial position and the lower tones to another cannot be interpreted in terms of a following of the input from one side of space rather than the other, since the higher and lower melodic lines were each

composed of tones that emanated from both spatial locations. One may, however, interpret this handedness correlate as reflecting relatively more activity in the dominant hemisphere on the part of neural units underlying the higher tones, and relatively more activity in the nondominant hemisphere on the part of neural units underlying the lower tones. Justification for this view comes in part from neurological studies showing that patients who experience palinacousis tend to perceive the illusory sound as located on the side of auditory space contralateral to the lesion (Jacobs, Feldman, Diamond, & Bender, 1973). Further, when patients obtain auditory sensations upon stimulation of the temporal lobe, these sensations are also generally referred to contralateral auditory space (Penfield & Perot, 1963).

A prominence of dominant over nondominant pathways is therefore implicated in both the octave and the scale illusions. These findings may be related to those of other investigators who explored patterns of ear advantage in the processing of melodies or tonal sequences. Very mixed results have been obtained in these studies. Some have found left ear superiorities (Kimura, 1964, 1967; King & Kimura, 1972; Bartholomeus, 1974; Darwin, 1969; Zatorre, 1979; Spellacy, 1970; Spreen, Spellacy and Reid, 1970). Others have found no ear differences (Gordon, 1970; Bartholomeus, Doehring, & Freygood, 1973; Berlin, 1972; Doehring, 1971, 1972). Under yet other conditions right ear superiorities have been obtained instead (Halperin, Nachshon, & Carmon, 1973; Robinson & Solomon, 1974; Papcun, Krashen, Terbeek, Remington, & Harshmann, 1974; Natale, 1977). Bever and Chiarello (1974), and Johnson (1977) obtained a right ear superiority for musicians and a left ear superiority for nonmusicians in melody recognition.

Such inconsistencies are probably due to a variety of factors. With sequences generated by voice or by natural instruments, recognition may be achieved in part by spectral cues. Loudness and temporal cues may also serve as bases for judgment, and so on. It is not unreasonable to suppose that specific attributes of a melodic segment might be processed in different parts of the nervous system. For example, Gordon (1970) obtained a left ear advantage in processing dichotically presented chords that were generated by an electronic organ, yet in this same study he failed to find an ear difference in the processing of melodies. Gaede, Parsons and Bertera (1978), using monaural presentation, found a left ear advantage in processing chords, and yet a right ear advantage in melody recognition. Further, Charbonneau and Risset (1975) studied the processing of dichotic sound sequences that varied either in fundamental frequency or in spectral envelope. When fundamental frequency was varied, a right ear advantage was obtained. Yet when spectral envelope was varied, a left ear advantage was obtained instead. If the relative involvement of the two hemispheres differs, depending on the specific musical attribute being processed, this could explain such results. Furthermore, different categories of listener might utilize specific musical attributes to varying extents. This could explain the discrepancies in performance found between musicians and nonmusicians in sone studies.

Ear differences have been interpreted by some investigators in terms of a simple dichotomy in processing strategy: the left or dominant hemisphere is assumed to specialize in "analytic" processing; the right or nondominant hemisphere in "Gestalt"

or "holistic" processing. However the meaning of such a dichotomy is far from clear. For instance, melody perception is held to be a "Gestalt" phenomenon. Indeed, Von Ehrenfels (1890) originally gave melody as an example of a Gestalt, because it retains its perceptual identity under transposition. However, in order to produce invariance under transposition, a set of specific intervals must be abstracted and their orders preserved. This requires a set of highly specific analyses. It would seem more useful, rather than invoking a nebulous "analytic-holistic" distinction, to attempt to pinpoint the types of processing responsible for different patterns of ear advantage.

There is a further factor that should be considered. This arises from findings on the scale illusion. When two melodies are simultaneously presented, one to each ear, the listener may not perceive these same melodies but may instead perceptually synthesize two different melodies, as shown in Figs. 2 and 3. When one of these melodies is later presented for recognition, accuracy may then be determined in part by the perceptual reorganization that had occurred during the dichotic presentation. For example, the typical right-handed listener, on perceiving the dichotic scale sequence, perceptually displaces the high tones from his left ear to his right, and perceptually displaces the low tones from his right ear to his left. A recent study has demonstrated that this phenomenon forms the basis of the apparent left ear advantage in dichotic listening to simultaneous sequences of tones (Deutsch, in preparation).[1]

### E. Melody Perception from Phase-Shifted Tones

Another technique relating to melodic channeling on the basis of spatial location was employed by Kubovy and co-workers. Kubovy, Cutting and McGuire (1974) presented a set of simultaneous and continuous sine wave tones to both ears. One of these tones in one ear was phase shifted relative to its counterpart in the opposite ear. When these tones were phase shifted in sequence a melody that corresponded to the shifted tones was clearly heard. However, the melody was undetectable when the stimulus was presented to either ear alone. Subjectively, the dichotically presented melody was heard as occurring inside the head but displaced to one side of the midline, while a background noise was heard as localized to the opposite side. So, it was as though a source in one spatial position was producing the melody, and a different source in another spatial position was producing the noise.

Kubovy (1981) pointed out that there are two potential interpretations of this effect. First, the segregation of the melody from the noise could have been based on *concurrent difference cues;* i.e., the target tone may have been segregated because at that time its interaural disparity—or apparent spatial location—differed from that of the background tones. Alternatively, the effect could have been based on *successive difference cues*, i.e., the target tone may have been segregated because it had moved its apparent location. Two further configurations were therefore devised to determine which of

[1]For a review of the neurological substrates of music perception see Chapter 15.

these factors was responsible. In the first, the target tones moved while the locations of the background tones remained constant, producing a successive difference cue. In the second, the target tones themselves did not move, but the background tones did, so that the target tones were segregated from the others, producing a concurrent difference cue. Kubovy found that although both types of cue were effective in producing segregation, the successive difference cue was considerably more effective than the concurrent difference cue.

In another experiment, Kubovy and Howard (1976) presented six tones simultaneously, in such a way that each occupied a different apparent position in space. They then displaced each tone in turn to a new apparent position, and so produced a melody by successive difference cues. They studied the effect of interpolating temporal gaps of different durations between successive tone bursts, and found that the melody could still be heard through such gaps. Although there was considerable intersubject variability in sensitivity to the effect of the gaps, one subject performed perfectly with gaps of 9.7 seconds (the longest duration employed). Thus, this effect of a successive difference cue was found to be capable of acting over surprisingly long silent intervals.

It is interesting to note that in Kubovy's paradigm, configurations were formed from a fixed set of tonal frequencies that simply shifted their apparent positions in space. Melodic channeling resulted from these movements of spatial position. This is analogous to the situation in the octave illusion where two continuous tones interchange their positions in space, resulting in melodic channeling by spatial location. As we have seen, when successive configurations are formed from different frequencies, rather than identical frequencies, channeling on other lines occurs instead. The issue of how differences in the frequencies of successive chords would affect channeling in Kubovy's paradigm remains to be explored.

### F. Discussion

We have found that the issue of how melodic channels are formed in two-channel listening situations is a complex one. Given certain stimulus configurations, channeling occurs on the basis of spatial location. Yet given other configurations, channeling occurs instead on the basis of frequency proximity or contour. In the conditions we have examined, amplitude plays a remarkably small role as a basis for channeling.

The radical differences in channeling strategy demonstrated here bear on certain apparent inconsistencies in the literature on divided attention. Certain investigators have found that the requirement to distribute attention across ears produced performance decrements on various tasks (e.g., Cherry & Taylor, 1954; Broadbent, 1954, 1958; Moray, 1959; Treisman, 1971). It was hypothesized that such decrements were due to an inability to switch attention between ears rapidly enough for the task demands. However, other investigators have found evidence against this view (e.g., Sorkin, Pastore, & Pohlman, 1972; Sorkin, Pohlman, & Gilliom, 1973; Moray, 1975;

Shiffrin, Pisoni, & Casteneda-Mendez, 1974; Pollack, 1978). It would appear from the work reviewed here that deficits in monitoring information simultaneously from two spatial locations should occur with certain types of stimulus configuration but not with others. It is probable that the configurations that give rise to such deficits are such as to induce the strong inference that the inputs to the two ears are emanating from separate sources rather than a single source. Integrating the information from such two sources would, in normal listening situations, lead to confusion in monitoring the environment. However, with configurations where there is an ambiguity of interpretation in terms of sources, integration of the information from the two ears could be the most useful strategy.

## IV. CHANNELING OF RAPID SEQUENCES OF SINGLE TONES

### A. Grouping by Frequency Proximity

Melodic channeling has also been studied with the use of rapid sequences of single tones. When these tones are in more than one frequency range, they tend to split apart perceptually, with the result that two or more melodic lines are heard in parallel. Composers often take advantage of this perceptual phenomenon with their use of pseudopolyphony, or compound melodic line. Here one instrument plays a rapid sequence of single tones which are drawn from different pitch ranges, so that two simultaneous melodic streams are clearly perceived. Figure 11a shows a segment of music that exploits this principle. Figure 11b shows the same segment with log frequency and time mapped into two dimensions of visual space. It can be seen that the principle of Proximity clearly emerges in the visual representation. At lower speeds, the tendency to group by pitch proximity still persists, but is subjectively less compelling.

One of the early experiments on this phenomenon was that of Miller and Heise (1950), who presented listeners with a sequence consisting of two tones that alternated at a rate of 10 per second. They found that when the frequencies of these tones differed by less than 15%, the sequence was heard as a trill (i.e., as a single string of related tones). However, when the frequency disparity between the alternating tones increased, the sequence was heard instead as two interrupted and unrelated tones. This phenomenon has been termed "fission" by several investigators (Van Noorden, 1975). Heise and Miller (1951) examined this phenomenon further, using sequences of tones that were composed of several different frequencies. They found that if one of the tones in a rapid repetitive sequence differed sufficiently in frequency from the others it was heard as in isolation from them.

Dowling (1967, 1973) has demonstrated the importance of this principle in a long term memory situation. He presented two well-known melodies with their component tones alternating at a rate of eight per second. Recognition of these melodies was

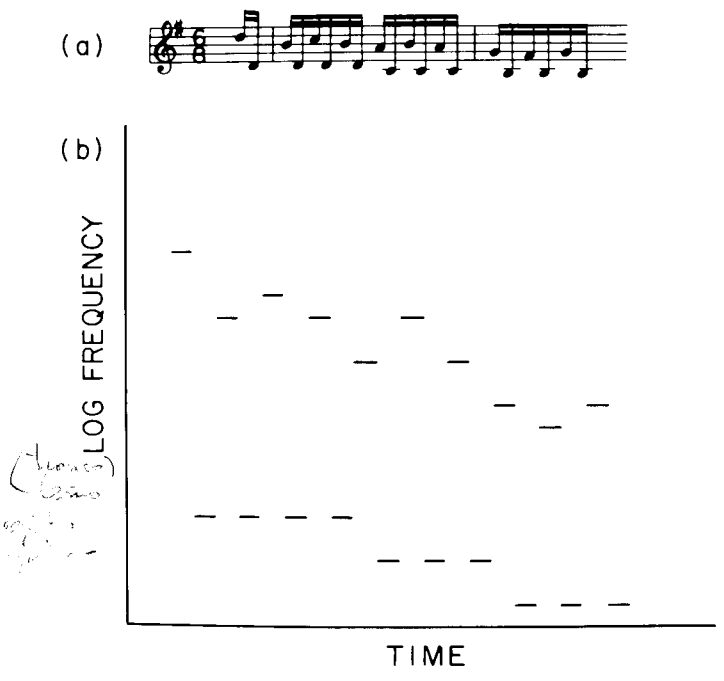



**Fig. 11.** Grouping of melodic stimuli on the basis of frequency proximity. Two parallel lines are perceived, each in a different frequency range (from Beethoven's *Six Variations on the Duet* "Nel cor piu non mi sento" from Paisiello's *La Molinara*).

found to be very difficult when they were in overlapping pitch ranges since their components were perceptually combined into a single stream. However, as one of the alternating melodies was gradually transposed so that their pitch ranges diverged, recognition became increasingly more easy.

## B. Temporal Coherence as a Function of Frequency Proximity and Tempo

*Temporal coherence* is a term used to describe the subjective impression that a sequence of tones froms a connected series. Schouten (1962) studied the conditions giving rise to the perception of temporal coherence. He varied both the frequency relationships between successive tones in a sequence and also their presentation rate. As the frequency separation between successive tones increased, the tempo of the sequence had to be reduced in order to maintain the impression of temporal coherence between these tones.

Van Noorden (1975) investigated this phenomenon in detail. Listeners were presented with sequences of alternating tones, and were instructed either to try to hear temporal coherence or to try to hear fission. Two boundaries were determined by this

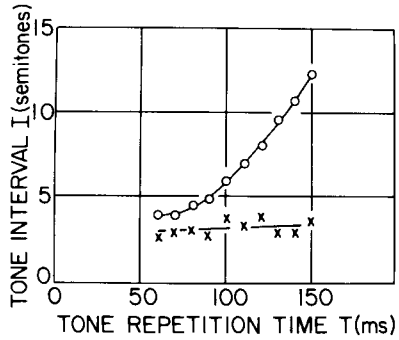



**Fig. 12.** Temporal coherence boundary (o) and fission boundary (x) as a function of frequency relationship between alternating tones and presentation rate (from Van Noorden, 1975).

method. The first, termed the *temporal coherence boundary*, established the threshold frequency separation as a function of tempo required for the listener to hear the sequence as coherent. The second, termed the *fission boundary*, established these values when the listener was attempting to hear fission. As shown in Fig. 12, when listeners were trying to hear coherence, decreasing the tempo from 50 to 150 msec per tone increased the frequency separation within which coherence could be heard from 4 to 13 semitones. However, when the listeners were trying to hear fission, decreasing the tempo had little effect on performance. Between these two boundaries there was a large region where the listener could alter his listening strategy at will, and so hear either fission or coherence. So within this region, attentional set was important in determining how the sequence was perceived; however outside this region, attentional set was not effective.

Bregman and Bernstein (quoted in Bregman, 1978) confirmed the finding of an interaction between frequency separation and tempo in judgments of coherence. They found that as two alternating tones converged in frequency, a higher rate of alternation was required for the sequence to split into two streams. This effect was found to hold throughout a substantial frequency range.

### C. Grouping by Frequency Proximity Builds with Repetition

Several experiments have shown that the splitting of tonal sequences into streams on the basis of frequency proximity builds up with repetitive presentation. For instance, Van Noorden (1975) compared the temporal coherence boundary for two-tone, three-tone, and long repetitive sequences. With three-tone sequences the frequency change was either unidirectional or bidirectional. It was found that for unidirectional three-tone sequences, temporal coherence was observed at rates that were equal to or even higher than those for two-tone sequences. (This follows the principle of Good Continuation, as described below.) But with bidirectional three-tone sequences, the rate of frequency change had to be set much lower than for two-tone

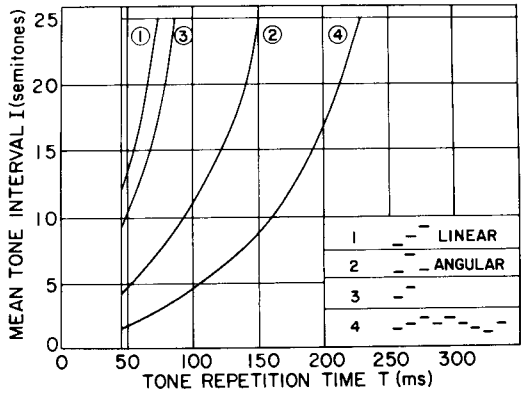



**Fig. 13.** Temporal coherence boundary for two-tone, three-tone undirectional, three-tone bidirectional, and continuous sequences (from Van Noorden, 1975).

sequences before coherence could be perceived. With long repetitive sequences the rate of frequency change had to be set lower still (Fig. 13).

In a related experiment, Bregman (1978) presented listeners with sequences consisting of two "high" tones (748 and 831 Hz) and one "low" tone (330 Hz). When this sequence split into two streams, the upper stream was perceived as an alternation of two high tones, and the lower stream as the steady repetition of a single tone. The experiment varied the number of tones packaged between four-second periods of silence. On each trial listeners adjusted the speed of the sequence until the point of splitting was determined. As shown on Fig. 14, as the package size increased, the speed required for segregation decreased.

Bregman interpreted these findings along the following lines. Stream segregation may be viewed as the product of a mechanism that acts to "parse" the auditory environment (i.e., to group together components of the acoustic spectrum in such a way as to recover the original sources). Such a mechanism would be expected to accumulate evidence over time, so that the segregation of acoustic components into groups should build up with repeated presentation.

Further evidence for the view that stream segregation results from a "parsing" mechanism was provided by Bregman and Rudnicky (1975). Listeners judged the order of two tones that were embedded in a four-tone pattern flanked by two "distractor" tones. The presence of the distractor tones made judgment of the order of the test tones difficult. However, when another stream of tones, called "captor" tones, was moved close to the "distractor" tones, this caused the "distractors" to combine with the "captors" to form a single stream, leaving the test tones in a stream of their own. This had the consequence that the order of the test tones was now easy to judge. Bregman and Rudnicky argue that this situation presents the listener with two simultaneously structured streams, of which the "distractor" tones can belong to either one, but not to both simultaneously. This is as expected on an interpretation in terms of an

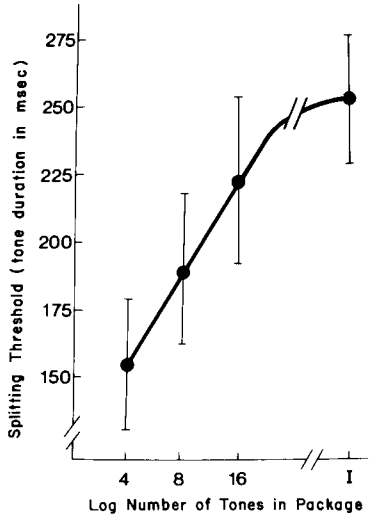



**Fig. 14.** Threshold for stream segregation as a function of number of tones per package. Two "high" tones were presented in alternation with a single "low" tone (from Bregman, 1978).

auditory parsing mechanism: any given tone is likely to be emanating from only one source; not from two sources simultaneously.

It should be noted that the cumulation of effect over time found by Bregman (1978) is analogous to cumulation effects found in the octave illusion, where the strength of tendency to follow the frequency presented to one side of auditory space rather than the other also builds up with repeated presentation, and builds up more rapidly as repetition rate increases. Analogous findings were obtained for the strength of tendency to localize toward the higher frequency signal in this illusion (Duetsch, 1976, 1978). Such a built-up of effect is also well interpreted in terms of evidence accumulation.

## D. Grouping by Frequency Proximity and the Perception of Temporal Relationships

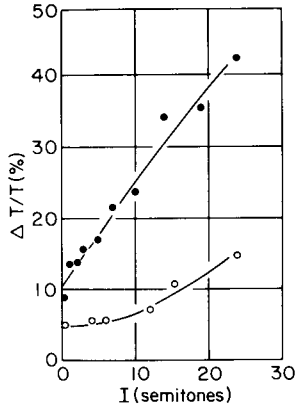
One striking consequence of the formation of separate streams out of rapidly presented sequences is that temporal relationships between the elements of different streams become difficult to process. This has been shown in several ways. Bregman and Campbell (1971) presented a repetitive sequence consisting of six tones: three from a high frequency range and three from a low frequency range. When these tones were presented at a rate of 10 per second, it was difficult for listeners to perceive a pattern of high and low tones that was embedded in the sequence.

Dannenbring and Bregman (1976) demonstrated a further perceptual consequence of this breakdown of temporal processing. They found that when two tones alternate

at high speeds so that they produce separate perceptual streams, the tones in the two streams appear to be perceptually overlapping in time. A related study was performed by Fitzgibbon, Pollatsek, and Thomas (1974) who explored the perception of temporal gaps between tones occurring in rapid sequence. When a 20-msec gap was interpolated between tones in the same frequency range, detection of this gap was easy. However, when the gap was interpolated between tones in different frequency ranges, detection performance dropped considerably.

A further reflection of this breakdown of temporal processing was found by Van Noorden (1975). He studied detection of the temporal displacement of a tone that alternated continuously with another tone of different frequency and found that as the tempo of the sequence increased, the mean just noticeable displacement also increased. This increase was substantial for sequences where the tones were widely separated in frequency, but only slight for sequences where the frequencies were contiguous. These results paralleled those found for judgments of temporal coherence.

Such deterioration of temporal processing as a result of frequency disparity occurs with two-tone sequences also. Divenyi and Hirsh (1972) found that discrimination of the size of a temporal gap between a tone pair deteriorates with increasing frequency disparity between members of the pair. Further, Williams and Perrott (1972) measured the minimum detectable gap between tone pairs. They found that for tones of 100 and 30 msec duration, the threshold rose with increasing frequency disparity between members of the pair. However, Van Noorden (1975) showed that this deterioration of temporal processing is considerably greater for long repetitive sequences than for two-tone sequences; so that it develops as a consequence of stream formation (Fig. 15). This conclusion also follows from consideration of Bregman and Campbell's results (1971).



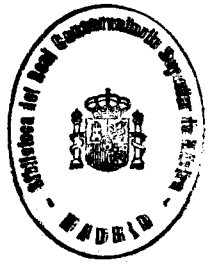



**Fig. 15.** ○ The just noticeable displacement ΔT/T of the second tone of a two-tone sequence as a function of tone interval I. ● The just noticeable displacement ΔT/T of one tone in a continuous sequence of alternating tones as a function of tone interval I (from Van Noorden, 1975).

## E. Grouping by Good Continuation

Another principle found to be effective in producing grouping is that of Good Continuation. Bregman and Dannenbring (1973) found that when a repeating cycle consisting of a high tone alternating with a low tone tended to segregate into two streams, this splitting tendency was reduced when the high and low tones were connected by frequency glides. Similarly Nabelek, Nabelek, and Hirsh (1973) reported that for complex tone bursts, frequency glides between the initial and final tones resulted in more pitch fusion than when these tones were juxtaposed with no transitions.

Related experiments have involved the perception of rapid sequences of three or more tones. Divenyi and Hirsh (1974) studied order identification for three-tone sequences, and found that sequences with unidirectional frequency changes were easier to order than sequences with bidirectional frequency changes. Analogous results were obtained by Nickerson and Freeman (1974), Warren and Byrnes (1975), and McNally and Handel (1977) for four-tone sequences. Furthermore, Van Noorden (1975) found that a sequence of three tones was more likely to be judged as coherent if these tones formed a unidirectional rather than a bidirectional frequency change.[2]

## F. Grouping by Timbre

The grouping of complex tones on the basis of sound type or timbre is an example of grouping by the principle of Similarity. (A visual example of such grouping is shown in Fig. 1b, where the open and closed circles each combine perceptually to form vertical rows.) Grouping on the basis of timbre is clearly apparent in natural musical situations (Erickson, 1975). Adjacent phrases are often played by different instruments to enhance their perceptual separation. Further, overlaps in pitch range are far more common where more than one instrument type is involved, reflecting the greater perceptual separation provided by the timbral difference.

A striking demonstration of grouping by timbre was produced by Warren, Obusek, Farmer and Warren (1969). They constructed repeating sequences consisting of four unrelated sounds, a high tone (1000 Hz), a hiss (2000 Hz octave band noise), a low tone (796 Hz), and a buzz (400 Hz square wave). Each sound was 200 msec in duration, and the sounds followed each other without pause. Listeners were quite unable to name the orders of the sounds in such repeating sequences. For correct naming to be achieved, the duration of each sound had to be increased to over half a second.

It appears that two separate factors are involved in this effect. The first factor is the organization of the elements of a sequence into separate streams on the basis of sound

[2]Further issues concerning the grouping of rapid sequences of tones involve the effect of average frequency difference between the tones. An extended discussion of these issues is beyond the scope of the present chapter, and the reader is referred to Warren and Byrnes (1975), Nickerson and Freeman (1974), and Divenyi and Hirsh (1978).

type, analogous to organization on the basis of frequency proximity. The second factor involves the lack of familiarity with such sound sequences. It has been shown that when verbal items are combined to form repeating sequences of this nature, correct ordering occurs at considerably faster rates (Warren & Warren, 1970; Thomas, Cetti, & Chase, 1971; Thomas, Hill, Carroll, & Garcia, 1970; Dorman, Cutting, & Raphael, 1975). It is likely that sequences composed of familiar musical sounds would also be more easily ordered. Although this has not been formally investigated, an observation by the author is of relevance here. It was found that a trained percussionist specializing in avant-garde music had little difficulty in discriminating sequences such as those created by Warren and his colleagues. This musician frequently produced such sequences in musical performance.

The question then arises as to the nature of the process that enables the rapid reconstruction of the order of components of complex yet familiar sound sequences such as in speech and music. Wickelgren (1969, 1976) has proposed that the correct ordering of speech components is based on an encoding of a set of context-sensitive elements that need not themselves be ordered. For instance, he proposed that the word "struck" is encoded not as the ordered set of phonemes /s/, /t/, /v/, /u/, /k/, but as the unordered set of context-sensitive allophones /♯$^{s}$t/, /s$^{t}$r/, /t$^{r}$u/, /r$^{u}$k/, /u$^{k}$♯/. Thus each of these context-sensitive elements contains some local information concerning how this element is ordered in relation to the other elements in the set. From such an unordered set of elements the information concerning their order can be derived.

This theory can readily be applied to auditory perception in general, and one may hypothesize that the easy identification of familiar sound sequences is mediated by an acquired set of such context-sensitive elements. For familiar sounds presented in unfamiliar order, these context-sensitive elements may not be encoded firmly enough to achieve correct identification.

An alternative proposal, suggested by Warren (1974), is that the ready identification of familiar sequences is mediated by a two-stage process. In the first stage the sequence is recognized in global fashion: as a "temporal compound" which can be distinguished from other compounds without being analyzed into its components. Other factors in addition to the perception of relationships between strictly adjacent items could be involved in such global processing. In the second stage there takes place an item-by-item analysis of the components of this compound and their orders.

Judgments of temporal order for only two disparate sounds are easier than for continuous repetitive sequences. Hirsh (1959) and Hirsh and Sherrick (1961) found that the threshold for ordering two disparate events was around 20 msec for highly trained listeners, though somewhat higher for untrained listeners (Hirsh, 1976). This superior performance is probably based on several factors. First, items that are preceded or followed by silence are more readily identified than those that are not (Warren, 1974). Second, there are fewer relationships to be judged between two events. And, third, an active process which organizes elements according to sound type probably acts to inhibit the perception of relationships between disparate elements, in a fashion analogous to the process that organizes elements by frequency proximity. This effect should be expected to cumulate with repetition.

### G. Grouping by Amplitude

Amplitude has been shown to be an effective grouping principle in the perception of rapid sequences of single tones. Dowling (1973) in his experiment on perception of interleaved melodies found that loudness differences between the melodies resulted in an enhanced ability to hear them as separate. Van Noorden (1975) studied perception of sequences where the tones were of identical frequency but alternated in amplitude. He found that with amplitude differences of less than 5 dB a single coherent stream was heard, even though loudness differences were clear. However, with larger amplitude differences two separate streams of different loudnesses were heard instead. Under these conditions attention could be directed to the softer stream as well as to the louder one. With even larger amplitude differences between the alternating tones, the auditory continuity effect was produced, and the softer tone was heard as though continuing through the louder tone (see below).

### H. Grouping by Temporal Proximity

When we attend to one melodic configuration rather than to another, we are forming figure-ground relationships analogous to those in vision (Gregory, 1970). Perception of sequences of tones that are interleaved in time may then be likened to visual perception of embedded figures. Divenyi and Hirsh (1978) drew this analogy, and argued that melodic configurations may be represented in two dimensions, with frequency providing one dimension and time the other (see also Julesz & Hirsh, 1972). Just as visual configurations can be more readily identified when these are spatially separated from background stimuli, so should melodic configurations be more readily identified when these are separated either in time or in frequency from background tonal stimuli.

As a test of this notion, Divenyi and Hirsh presented rapid three-tone patterns that could occur in any of six permutations and required subjects to identify on each trial which permutation had been presented. These three-tone patterns were embedded in sequences of seven or eight tones, but were not interleaved with them. Identification performance was superior when the irrelevant tones and the target tones were in different frequency ranges. Furthermore, performance levels varied considerably depending on the temporal position of the target pattern within the full sequence. Best performance was obtained when the target pattern occurred at the end of the sequence; performance was also relatively good when the target was located at the beginning, but it was close to chance when the target occurred in the middle of the sequence. Both temporal and frequency separation were therefore found to reduce interference from the background tones. Previously Ortmann (1926) had found that a single tone was more salient when it was the highest or lowest in a sequence, or when it was in the first or last position. Similar conclusions were drawn recently by Watson, Kelly, and Wroton (1975) and Watson, Wroton, Kelly, and Benbasset (1976). Divenyi and Hirsh's results, therefore, extended such findings to the case of melodic

configurations. (Further issues involving grouping by temporal position are discussed in Chapter 9.)

### I. Perceptual Replacement of Missing Sounds

So far we have examined several instances where our perceptual system reorganizes sound sequences in accordance with expectations derived from both the sequences themselves and our knowledge of the auditory environment. It has also been found that sounds which are not actually present in the stimulus may be perceptually synthesized in accordance with such expectations.

Various studies have shown that when two sounds are presented in alternation, the fainter sound may be heard as continuing through the louder one (Miller & Licklider, 1950; Thurlow, 1957; Vicario, 1960). More recently, Warren (1970) and Warren, Obusek, and Ackroff (1972) showed that if a phoneme in a sentence is replaced by a louder noise, the missing phoneme may be perceptually synthesized. Analogous results were obtained with nonverbal sounds. In a set of parametric studies, Warren and his colleagues have demonstrated that this "auditory induction effect" occurs only under stimulus conditions where it would be reasonable to assume that the substituted sound had masked the missing one.

Dannenbring (1976) produced another version of this effect. He presented a sine wave tone that repeatedly glided up and down in frequency. When a loud noise burst was substituted for a portion of this sound, it still appeared to glide through the noise. However, if the tone changed in amplitude just before the noise burst, producing evidence that something had happened to the tone itself, rather than its simply being masked, the tendency to hear the tone as continuing through the noise was reduced (Bregman & Dannenbring, 1977).

## V. VOLUNTARY ATTENTION

We now turn to a consideration of the effects of voluntary attention on channeling phenomena. In listening to music outside the laboratory, we have the impression that we can direct our attention at will; listening now to a melodic line, now to its accompaniment, now to a chosen instrument, and so on. Yet the conditions under which such attention focusing is indeed under voluntary control remain to be determined. We are dealing with two issues here. First, we may examine the role of voluntary attention in the initial division of the configuration into groupings. Second, we may examine the role of voluntary factors in determining which grouping is attended to, once such a division is established. Concerning the first issue, we have described several configurations where a particular grouping principle is so strong that listeners are generally unaware of alternative organizations. For example, most people on hearing the scale illusion form groupings so strongly on the basis of frequency proximity that they hear tones in one frequency range as emanating from one source,

and tones in another frequency range as emanating from a different source. They therefore believe that they are attending to one spatial location rather than to another; yet in reality they are synthesizing information from two different locations (Deutsch, 1975b). The same is true for the two-part contrapuntal patterns devised by Butler (1979a). Similarly, on listening to the octave illusion, many people believe that a single high tone is being delivered intermittently to one ear, and a single low tone intermittently to the other ear. Yet in fact they are being presented with a continuous two-tone chord. So, here again involuntary organizational mechanisms are so strong that the listener is unaware of the nature of the stimulus configuration. The sequence of Kubovy *et al.* (1974) provides another example. Here one hears a melody as in one spatial location and a background noise as in another, yet in reality a continuous chord is being delivered to both ears.

However, when we consider the role of voluntary attention in determining which of two channels is attended to, once these have been formed, we find that in all these examples voluntary attention plays a prominent role. For example, in the scale illusion listeners who hear two melodic lines in parallel can choose at will to attend to either the higher or the lower one. Even those listeners who initially hear only the higher melodic line may after repeated presentations focus their attention on the lower one. Again in Butler's contrapuntal patterns we can choose at will to listen to the higher or the lower of the two melodies that we have perceptually synthesized. In the case of the octave illusion, those listeners who unambiguously hear a high tone in one ear alternating with a low tone in the other ear can focus their attention on either the high tone or the low one. Similarly, with the configuration of Kubovy *et al.* (1974), listeners can direct their attention to either the melody or to the noise.

When we consider channeling of rapid sequences of tones, we also find that strong involuntary factors are involved in the formation of initial groupings. Thus, the inability to form order relationships across streams based on frequency proximity (Bregman & Campbell, 1971), sound type (Warren *et al.*, 1969), or spatial location (Deutsch, 1979) cannot readily be overcome by voluntary attention focusing.

However, other examples have been given where voluntary attention does play a role. In exploring the temporal coherence boundary, Van Noorden (1975) found that within a given range of tempos and of frequency relationships, listeners may direct their attention at will, hearing either fission or temporal coherence (Fig. 12). An ambiguous situation where channeling by timbre was set in competition with channeling by pitch was created by Erickson (1974) in a compositon called *LOOPS*. Here, a repeating melodic pattern was performed by five instruments, with each instrument playing a different note in the manner of a hockett, so that each pitch was eventually played by every instrument. Under these conditions listeners can often choose to follow the sequence on the basis of either timbre or pitch. It therefore appears that although there are strong involuntary components in the formation of groupings, ambiguous stimulus situations may be set up where voluntary attention can be the determining factor.

Considering the issue of voluntary factors in determining which stream is attended to, once a set of alternatives have been formed, we find that voluntary attention

focusing is easily achieved with rapid sequences also. For instance, Van Noorden (1975) reports that in cases where two streams were formed on the basis of frequency proximity, the listener was able to direct his attention at will and concentrate on either the upper stream or the lower one. However, he noted that an involuntary component was also present: The listener's percept would sometimes switch spontaneously to the stream he was attempting to ignore. This was true even when the unattended stream was less salient. Similar observations were made by present author using streaming by spatial location in patterns such as on Fig. 6.

In summary, it appears that the initial division of the stimulus configuration into groupings is often outside the listener's voluntary control, though ambiguous situations may be generated where attention focusing can be effective. In contrast, once a set of groupings is established, voluntary attention focusing plays a prominent role in determining which of these is attended to. This division of the attentional process into two stages corresponds in many respects to the stages identified as preattentive and postattentive by Neisser (1967) and Kahneman (1973) among others. These terms, however, have often been taken to imply different depths of analysis at these two stages, yet the issue of depth of analysis remains unsettled (Deutsch & Deutsch, 1963; Keele and Neill, 1979).

We may next consider the consequences of the selective attention process for the unattended material (i.e., for the component of music that serves as "ground" when attention is focused on another component which serves as "figure".[3] Considering the analogous issue for streams of speech, Cherry (1953) and Cherry and Taylor (1954) presented subjects with two messages, one to each ear, and required them to shadow one of these. They found that the subjects were able to report virtually nothing about the message presented to the nonattended ear, not even what language it was in. Other studies have produced similar findings (Kahneman, 1973). Recently, the present author set up an analogous situation for musical stimuli. Two familiar folk songs were recorded by piano, and were simultaneously presented, one to each ear. Listeners were required to shadow one of the melodies by singing, and were later asked to describe what had been presented to the other ear. Analogous to Cherry's finding, no listener was able to name the unattended melody, and none was able to describe much about the stimulus. Thus, voluntary attention focusing on one channel of a musical configuration can have the effect of suppressing the other channel from conscious perception.[4]

The question then arises as to what extent the unattended signal is processed under these conditions. This is a controversial issue in the literature on speech materials. Broadbent (1958) proposed that in selective listening a filter sorts out simultaneously presented stimuli on the basis of gross physical characteristics, such as spatial location or frequency range. Stimuli that share a characteristic that defines the relevant "chan-

[3]The author is indebted to R. Erickson for raising this question.

[4]Channeling by spatial location was here facilitated by the fact that the messages delivered to the two ears were asynchronous. Had they been synchronous, as with the stimuli used by Deutsch (1975b) and Butler (1979a), it would not have been possible for the listener to focus attention on one ear rather than the other.

nel" are then perceptually analyzed further, whereas the other stimuli are simply filtered out. This theory was found to be unable to account for certain findings—for instance, that the meaning of words may be important determinants of selective attention (Grey & Wedderburn, 1960; Treisman, 1960). Consequently, Treisman (1960, 1964) suggested a modification of filter theory to accommodate these findings. She proposed that the unattended message is not totally rejected, as Broadbent had suggested, but rather attenuated. An alternative view was taken by Deutsch and Deutsch (1963) who proposed that all input, whether attended to or not, is completely analyzed by the nervous system. The information thus analyzed is then weighted for importance or pertinence. Such weightings are determined both by long-term factors (for instance, there is a long-term predisposition to attend to one's own name) and by factors determined by the current situation. The information with the highest weighting of importance then controls awareness. Recent studies (e.g., Lewis, 1970; Corteen & Wood, 1972; Shiffrin & Schneider, 1977) have provided strong evidence for this view, but the issue remains controversial. In the case of music, this has not yet been the subject of experimental investigation.

## VI. CONCLUSION

In this chapter we have focused on musical channeling phenomena in two types of situation. First, we have explored the perceptual consequences of presenting two simultaneous sequences of tones in different spatial locations. Second, we have investigated channeling when rapid sequences of single tones were presented. In general, relatively simple stimulus configurations were examined, and grouping or channeling on the basis of higher order abstractions was not considered. The formation of such abstractions is the subject of Chapter 9 and we shall assume that these can also serve as bases for grouping.[5]

## REFERENCES

Bartholomeus, B. Effects of task requirements on ear superiority for sung speech. *Cortex*, 1974, *10*, 215–223.

Bartholomeus, B. N., Doehring, D. G., & Freygood, S. D. Absence of stimulus effects in dichotic singing. *Bulletin of the Psychonomic Society*, 1973, *1*, 171–172.

Benade, A. H. *Fundamentals of musical acoustics*. London and New York: Oxford University Press, 1976.

Berlin, C. I. Critical review of the literature on dichotic effects—1970. In 1971 reviews of scientific literature on hearing. *American Academy Ophtology Otology* 1972, 80–90.

Bever, T. G., & Chiarello, R. J. Cerebral dominance in musicians and nonmusicians. *Science*, 1974, *185*, 537–539.

Bregman, A. S. The formation of auditory streams. In J. Requin (Ed.), *Attention and Performance*. (Volume VII) Hillsdale, New Jersey: Erlbaum, 1978. Pp. 63–76.

Bregman, A. S., & Campbell, J. Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 1971, *89*, 244–249.

[5]This work was supported by United States Public Health Service Grant MH-21001.

Bregman, A. S., & Dannenbring, G. L. The effect of continuity on auditory stream segregation. *Perception & Psychophysics*, 1973, *13*, 308-312.

Bregman, A. S., & Dannenbring, G. L. Auditory continuity and amplitude edges. *Canadian Journal of Psychology*, 1977, *31*, 151-159.

Bregman, A. S., & Pinker, S. Auditory streaming and the building of timbre. *Canadian Journal of Psychology*, 1978, *32*, 20-31.

Bregman, A. S., & Rudnicky, A. I. Auditory segregation: Stream or streams? *Journal of Experimental Psychology: Human Perception and Performance*, 1975, *1*, 263-267.

Broadbent, D. E. The role of auditory localization in attention and memory span. *Journal of Experimental Psychology*, 1954, *47*, 191-196.

Broadbent, D. *Perception and communication*. Oxford: Pergamon, 1958.

Butler, D. A further study of melodic channeling. *Perception & Psychophysics*, 1979, *25*, 264-268. (a)

Butler, D. Melodic channeling in a musical environment. *Research Symposium on the Psychology and Acoustics of Music, Kansas*, 1979. (b)

Charbonneau, G., and Risset, J-C. Differences entre oreille droite et oreille gauche pour la perception de la hauteur des sons. *Comptes Rendus, Académie des Sciences, Paris*, 1975, *281*, 163-166.

Cherry, E. C. Some experiments on the recognition of speech, with one and two ears. *Journal of the Acoustical Society of America*, 1953, *25*, 975-979.

Cherry, E. C., & Taylor, W. K. Some further experiments upon the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, 1954, *26*, 554-559.

Corteen, R. S., & Wood, B. Autonomic responses to shock-associated words in an unattended channel. *Journal of Experimental Psychology*, 1972, *94*, 308-313.

Dannenbring, G. L. Perceived auditory continuity with alternately rising and falling frequency transitions. *Canadian Journal of Psychology*, 1976, *30*, 99-114.

Danndenbring, G. L., & Bregman, A. S. Stream segregation and the illusion of overlap. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, *2*, 544-555.

Darwin, C. J. Auditory Perception and Cerebral Dominance. Doctoral dissertation, University of Cambridge, 1969.

Deutsch, D. An auditory illusion. *Journal of the Acoustical Society of America*, 1974, *55*, S18-S19. (a)

Deutsch, D. An auditory illusion. *Nature (London)*, 1974, *251*, 307-309. (b)

Deutsch, D. Musical illusions. *Scientific American*, 1975, *233*, 92-104. (a)

Deutsch, D. Two-channel listening to musical scales. *Journal of the Acoustical Society of America*, 1975, *57*, 1156-1160 (b)

Deutsch, D. Lateralization by frequency in dichotic tonal sequences as a function of interaural amplitude and time differences. *Journal of the Acoustical Society of America*, 1976, *60*, S50.

Deutsch, D. Binaural integration of tonal patterns. *Journal of the Acoustical Society of America*, 1978, *64*, S146. (a)

Deutsch, D. Lateralization by frequency for repeating sequences of dichotic 400-Hz and 800-Hz tones. *Journal of the Acoustical Society of America*, 1978, *63*, 184-186 (b)

Deutsch, D. Binaural integration of melodic patterns. *Perception & Psychophysics*, 1979, *25*, 399-405.

Deutsch, D. Two-channel listening to tonal sequences. In R. S. Nickerson and R. W. Pew (Eds.), *Attention and performance*. (Volume VIII) Hillsdale, New Jersey: Erlbaum, 1980.

Deutsch, D. The octave illusion and auditory perceptual integration. In J. V. Tobias and E. D. Schubert (Eds.), *Hearing research and theory*. (Volume I). Academic Press: New York, 1981.

Deutsch, D., *Left ear advantage for dichotic tonal sequences: an artifact of the scale illusion*. In preparation.

Deutsch, D., & Roll, P. L. Separate 'what' and 'where' decision mechanisms in processing a dichotic tonal sequence. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, *2*, 23-29.

Deutsch, J. A., & Deutsch, D. Attention: Some theoretical considerations. *Psychological Review*, 1963*1* 70, 80-90.

Divenyi, P. L., & Hirsh, I. J. Discrimination of the silent gap in two-tone sequences of different frequencies. *Journal of the Acoustical Society of America*, 1972, *52*, 166S.

Divenyi, P. L., & Hirsh, I. J. Identification of temporal order in three-tone sequences. *Journal of the Acoustical Society of America*, 1974, *56*, 144-151.

Divenyi, P. L., & Hirsh, I. J. Some figural properties of auditory patterns. *Journal of the Acoustical Society of America*, 1978, *64*, 1369-1386.

Doehring, D. G. Discrimination of simultaneous and successive pure tones by musical and nonmusical subjects. *Psychonomic Science*, 1971, *22*, 209-210.

Doehring, D. G. Ear asymmetry in the discrimination of monaural tonal sequences. *Canadian Journal of Psychology*, 1972, *26*, 106-110.

Dorman, M. F., Cutting, J. E., & Raphael, L. J. Perception of temporal order in vowel sequences with and without formant transitions. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, *104*, 121-129.

Dowling, W. J. Rhythmic Fission and the Perceptual Organization of Tone Sequences. Unpublished doctoral dissertation. Harvard University, Cambridge, Massachusetts, 1967.

Dowling, W. J. The perception of interleaved melodies. *Cognitive Psychology*, 1973, *5*, 322-337.

Ehrenfels, C. Von. *Uber Gestaltqualitäten Vierteljahrschrift fur Wissenschaftliche Philosophie*, 1890, *14*, 249-292.

Erickson, R. *Sound structure in music*. Berkeley, California: University of California Press, 1975.

Erickson, R. LOOPS, an informal timbre experiment, Center for Music Experiment, University of California, San Diego, 1974.

Fitzgibbon, P. J., Pollatsek, A., & Thomas, I. B. Detection of temporal gaps within and between perceptual tonal groups. *Perception & Psychophysics*, 1974, *16*, 522-528.

Gaede, S. E., Parsons, O. A., and Bertera, J. H. Hemispheric differences in music preparation: aptitude vs. experience. *Neurophychologia*, 1978, *16*, 369-373.

Gordon, H. W. Hemispheric asymmetries in the perception of musical chords. *Cortex*, 1970, *6*, 387-398.

Gray, J. A., & Wedderburn, A.A.I. Grouping strategies with simultaneous stimuli. *Quarterly Journal of Experimental Psychology*, 1960, *12*, 180-184.

Gregory, R. L. *The intelligent eye*. New York: McGraw-Hill, 1970.

Hass, H. Über den einfluss eines Einfachechos auf die Hörsamkeit von Sprache. *Acustica*, 1951, *1*, 49-52.

Halperin, Y., Nachshon, I., & Carmon, A. Shift of ear superiority in dichotic listening to temporally patterned nonverbal stimuli. *Journal of the Acoustical Society of America*, 1973, *53*, 46-50.

Heise, G. A., & Miller, G. A. An experimental study of auditory patterns. *American Journal of Psychology*, 1951, *64*, 68-77.

Hirsh, I. J. Auditory perception of temporal order. *Journal of the Acoustical Society of America*, 1959, *31*, 759-767.

Hirsh, I. J. Order of events in three sensory modalities. In S. K. Hirsh, D. H. Eldridge, I. J. Hirsh, & S. R. Silverman (Eds.), *Essays honoring Hallowell Davis*, St. Louis, Missouri: Washington University Press, 1976.

Hirsh, I. J., & Sherrick, C. E. Perceived order in different sense modalities. *Journal of Experimental Psychology*, 1961, *62*, 423-432.

Hochberg, J. Organization and the Gestalt Tradition. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception*. (Volume 1) New York: Academic Press. Pp. 180-211.

Jacobs, L., Feldman, M., Diamond, S. P., & Bender, M. B. Palinacousis: Persistent or recurring auditory sensations. *Cortex*, 1973, *9*, 275-287.

Johnson, P. R. Dichotically-stimulated ear differences in musicians and nonmusicians. *Cortex*, 1977, *13*, 385-389.

Judd, T. Comments on Deutsch's musical scale illusion. *Perception and Psychophysics*, 1979. *26*, 85-92.

Julesz, B., & Hirsh, I. J. Visual and auditory perception—An essay of comparison. In E. E. David and P. B. Denes (Eds.), *Human communication: A unified view*. New York: McGraw-Hill, 1972. Pp. 283-340.

Kahneman, D. *Attention and effort*. Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

Keele, S. W., and Neill, W. T. Mechanisms of attention. In E. C. Carterette and M. P. Friedman (Eds.), *Handbook of Perception* (Vol. 9) New York: Academic Press, 1979.

Kimura, D. Left-right differences in the perception of melodies. *Quarterly Journal of Experimental Psychology*, 1964, *16*, 355-358.

Kimura, D. Functional asymmetry of the brain in dichotic listening. *Cortex*, 1967, *3*, 163-178.

King, F. D., & Kimura, D. Left-ear superiority in dichotic perception of vocal nonverbal sounds. *Canadian Journal of Psychology*, 1972, *26*, 111-116.

Kubovy, M. Concurrent pitch-segregation and the theory of indispensable attributes. In M. Kubovy and J. Pomerantz (Eds.), *Perceptual organization*. Hillsdale: Erlbaum, New Jersey, 1981.

Kubovy, M., Cutting, J. E., & McGuire, R. M. Hearing with the third ear: Dichotic perception of a melody without monaural familiarity cues. *Science*, 1974, *186*, 272–274.

Kubovy, M., & Howard, F. P. Persistence of a pitch-segregating echoic memory. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, *2*, 531–537.

Lewis, J. L. Semantic processing of unattended messages using dichotic listening. *Journal of Experimental Psychology*, 1970, *85*, 225–228.

McNally, K. A., & Handel, S. Effect of element composition on streaming and the ordering of repeating sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, *3*, 451–460.

Miller, G. A., & Heise, G. A. The trill threshold. *Journal of the Acoustical Society of America*, 1950, *22*, 637–638.

Miller, G. A., & Licklider, J.C.R. The intelligibility of interrupted speech. *Journal of the Acoustical Society of America*, 1950, *22*, 167–173.

Moray, N. Attention in dichotic listening; Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, 1959, *11*, 56–60.

Moray, N. A date base for theories of selective listening. In P.M.A. Rabbitt and S. Dornic (Eds.), *Attention and performance*. (Volumn V) New York: Academic Press, 1975.

Nabelek, I. V., Nabelek, A. K., & Hirsh, I. J. Pitch of sound bursts with continuous or discontinuous change of frequency. *Journal of the Acoustical Society of America*, 1973, *53*, 1305–1312.

Natale, M. Perception of neurolinguistic auditory rhythms by the speech hemisphere. *Brain and Language*, 1977, *4*, 32–44.

Neisser, U. *Cognitive Psychology*. New York: Appleton, 1967.

Nickerson, R. S., & Freeman, B. Discrimination of the order of the components of repeating tone sequences: Effects of frequency separation and extensive practice. *Perception & Psychophysics*, 1974, *16*, 471–477.

Ortmann, O. On the melodic relatively of tones. *Psychological Monographs*, 1926, *35* (whole No. 162).

Papcun, G., Krashen, S., Terbeek, D., Remington, R., & Harshman, R. Is the left hemisphere specialized for speech, language and/or something else? *Journal of the Acoustical Society of America*, 1974, *55*, 319–327.

Penfield, W., & Perot, P. The brain's record of auditory and visual experience. *Brain*, 1963, *86*, 595–696.

Pollack, I. Temporal switching between binaural information sources. *Journal of the Acoustical Society of America*, 1978, *63*, 550–558.

Rasch, R. A. The perception of simultaneous notes such as in polyphonic music. *Acustica*, 1978, *40*, 1–72.

Robinson, G. M., & Solomon, D. J. Rhythm is processed by the speech hemisphere. *Journal of Experimental Psychology,1974, 102*, 508–511.

Schouten, J. F. On the perception of sound and speech; Subjective time analysis. *Fourth International Congress on Acoustics, Copenhagen Congress Report II*, 1962, 201–203.

Shiffrin, R. M., Pisoni, D. B., & Castaneda-Mendez, K. Is attention shared between the ears? *Cognitive Psychology*, 1974, *6*, 190–215.

Shiffrin, R. M., & Schneider, W. Toward a unitary model for selective attention, memory scanning and visual search. In S. Dornic (Ed.), *Attention and performance*. (Volume VI) Hillsdale: Earlbaum, 1977. Pp. 413–440.

Sorkin, R. D., Pastore, R. E., & Pohlmann, L. D. Simultaneous two-channel signal detection. II. Correlated and uncorrelated signals. *Journal of the Acoustical Society of America*, 1972, *51*, 1960–1965.

Sorkin, R. D., Pohlmann, L. D., & Gilliom, J. D. Simultaneous two-channel signal detection. III. 630- and 1400-Hz signals. *Journal of the Acoustical Society of America*, 1973, *14*, 101–109.

Spellacy, F. Lateral preferences in the identification of patterned stimuli. *Journal of the Acoustical Society of America*, 1970, *47*, 574–578.

Spreen, O., Spellacy, F., and Reid, J. R. The effect of interstimulus interval and intensity on ear asymmetry for nonverbal stimuli in dichotic listening. *Neuropbychologia*, 1970, *8*, 245–250.

Sutherland, N. S. Object recognition. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of Perception*. (Volume III) New York: Academic Press, 1973. Pp. 157–186.

Thomas, I. B., Cetti, R. P., & Chase, P. W. Effect of silent intervals on the perception of temporal order for vowels. *Journal of the Acoustical Society of America*, 1971, *49*, 584.

Thomas, I. B., Hill, P. B., Carroll, F. S., & Garcia, B. Temporal order in the perception of vowels. *Journal of the Acoustical Society of America*, 1970, *48*, 1010-1013.

Thurlow, W. An auditory figure-ground effect. *American Journal of Psychology*, 1957, *70*, 653-654.

Tobias, J. V. Curious binaural phenomena. In J. V. Tobias (Ed.), *Foundations of modern auditory theory*. (Volume II) New York: Academic Press, 1972.

Treisman, A. M. Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 1960, *12*, 242-248.

Treisman, A. M. Selective attention in man. *British Medical Bulletin*, 1964, *20*, 12-16.

Treisman, A. M. Shifting attention between the ears. *Quarterly Journal of Experimental Psychology*, 1971, *23*, 157-167.

Van Noorden, L.P.A.S. Temporal Coherence in the Perception of Tone Sequences. Unpublished doctoral dissertation. Technische Hogeschoel Eindhoven, The Netherlands, 1975.

Vicario, G. L'effetto tunnel acustico. *Revista di Psyicologia*, 1960, *54*, 41-52.

Von Helmholtz, H. *On the sensations of tone as a physiological basis for the theory of music*. (2nd English ed.) New York: Dover, 1954. (Originally published 1859)

Von Helmholtz, H. *Helmholtz's physiological optics*. (Translated from the 3rd German ed.) (1909-1911 byJ.P.C. Southall, ed.) Rochester, New York: Optical Society of America, 1925.

Wallach, H., Newman, E. B., & Rosenzweig, M. R. The precedence effect in sound localization. *American Journal of Psychology*, 1949, *62*, 315-336.

Warren, R. M. Perceptual restoration of missing speech sounds. *Science*, 1970, *167*, 392-393.

Warren, R. M. Auditory temporal discrimination by trained listeners. *Cognitive Psychology*, 1974, *6*, 237-256.

Warren, R. M., & Byrnes, D. L. Temporal discrimination of recycled tonal sequences: Pattern matching and naming of order by untrained listeners. *Journal of the Acoustical Society of America*, 1975, *18*, 273-280.

Warren, R. M., Obusek, C. J., & Ackroff, J. M. Auditory induction: Perceptual synthesis of absent sounds. *Science*, 1972, *176*, 1149-1151.

Warren, R. M., Obusek, C. J., Farmer, R. M., & Warren, R. P. Auditory sequence: Confusions of patterns other than speech or music. *Science*, 1969, *164*, 586-587.

Warren, R. M., & Warren, R. P. Auditory illusions and confusions. *Scientific American*, 1970, *223*, 30-36.

Watson, C. S., Kelly, W. J., & Wroton, H. W. Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of uncertainty. *Journal of the Acoustical Society of America*, 1976, *60*, 1176-1186.

Watson, C. S., Wroton, H. W., Kelly, W. J., & Benbasset, C. A. Factors in the discrimination of tonal patterns. I. Component frequency, temporal position and silent intervals. *Journal of the Acoustical Society of America*, 1975, *75*, 1175-1185.

Wertheimer, M. Untersuchung zur Lehre von der Gestalt II. *Psychologische Forschung*, 1923, *4*, 301-350.

Wickelgren, W. A. Context-sensitive coding, associative memory and serial order in (speech) behavior. *Psychological Review*, 1969, *76*, 1-15.

Wickelgren, W. A. Phonetic coding and serial order. In E. C. Carterette and M. P. Friedman (Eds.), *Handbook of Perception*. (Volume VII) New York: Academic Press, 1976. Pp. 227-264.

Williams, K. N., and Perrott, D. R. Temporal resolution of tonal pulses. *Journal of the Acoustical Society of America*, 1972, *51*, 644-647.

Zatorre, R. J. Recognition of dichotic melodies by musicians and nonmusicians. *Neuropbychologia*, 1979, *17*, 607-617.

# 5

# The Listener and the Acoustic Environment

*R. A. Rasch and R. Plomp*


I. Introduction ..... 135
II. Methodology ..... 137
III. Level Effects of Indirect Sound: Loudness ..... 140
IV. Temporal Effects of Indirect Sound: Definition ..... 141
V. Spatial Effects of Indirect Sound: Spaciousness ..... 142
VI. The Compromise between Definition and Spaciousness ..... 145
VII. Conclusion ..... 146
References ..... 146


## I. INTRODUCTION

If a sound source and a listener are situated in an open field without any sound-reflecting surfaces in the neighborhood, the emitted sound will reach the ears of the listener only via the straight line that connects source and listener. The sound image that the listener receives will roughly correspond to the sound emitted by the source. However, that is not the usual situation in listening to music. Producing musical sounds and listening to them is almost always done in rooms or halls—technically speaking, in enclosed spaces. These enclosed spaces have bounding surfaces (walls, floor, ceiling) that reflect the incident sound. Because of these reflections the emitted sound does not only reach the ears of the listener via the straight line from source to listener, but also via numerous other paths. The sound that reaches the listener without any reflection is called the *direct sound;* the sound that arrives after one or more reflections is called the *indirect sound* or *reverberation* (see Fig. 1). The presence of an indirect sound field has a profound influence on the sound image that the



Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

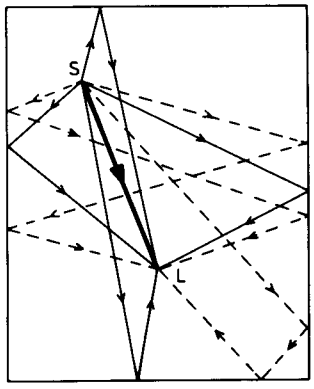



**Fig. 1.** Some sound paths from a sound source S to a listener L. The direct sound has been indicated with a heavy solid line. The four sound paths, including single reflections, have been drawn as thin solid lines. Slashed lines indicate sound paths with multiple reflections of which only a small selection has been included. As a matter of fact, the radiated sound is attenuated proportional to the square of the length of the path and by every reflection.

listener receives. The subjective effects of the indirect sound field make up what is loosely called the *acoustics* of a room or hall.

Whereas the physical aspects of sound in an enclosed space have been studied for almost a century (see Beranek, 1954; Kuttruff, 1973; Meyer, 1970), the subjective effects cannot claim a long history of research. Research in subjective room acoustics begins after World War II, and its results up to now are mainly tentative. This chapter gives a summary of empirical evidence from the experimental literature on subjective musical room acoustics, which centers in the United States, Great Britain, Germany (West and East), and Japan (Rasch 1977).

We will first briefly examine the physical aspects of indirect sound. Fig. 1 shows several paths by which the sound of a source can reach a listener. Since all sound paths work equally well in both directions, source and listener can always be interchanged. The differences between a situation with and one without indirect sound may be summarized in three points:

1. The indirect sound *adds sound energy* at the position of the listener, resulting in a higher intensity than there would be without indirect sound. The gain can be substantial and depends, of course, on the sound absorption (and reflection) of the boundaries. It can be up to 10 or 15 dB.

2. The indirect sound *arrives later* than the direct sound because its path is always longer. If the velocity of sound is approximately 340 m/sec, it can be stated that every additional meter in a sound path causes a delay of 3.4 msec. Roughly, the time delays of indirect sound can be up to 100 msec per reflection. If the indirect sound includes some strong single late reflections with delays of more than 50 msec, these are called *echoes*.

3. The indirect sound *arrives from other directions* than the direct sound. Usually, it

is possible to distinguish some discretely traceable reflections from the walls and ceiling that arrive first after the direct sound and a mass of diffuse later reflections coming from all directions.

The corresponding subjective effects may be described as follows:

1. The increase in sound intensity is perceived as an increase in *loudness*.
2. The later arrival of indirect sound has the effect that the source seems to sound a little longer than it really does. The direct sound is followed by a "cloud" of indirect sound. This gives continuity to a stream of notes that may have small discontinuities, such as staccato notes. But the indirect sound may also coincide with or even mask the direct sound of the succeeding notes, which may confuse the sound image to a lesser or greater extent. The temporal aspects of indirect sound correspond to the subjective attribute *definition*, the ability to distinguish and to recognize sounds.
3. The incidence of sound from all directions results in an impression of *spaciousness*. This is usually considered a positive quality, although it seems necessary that the position of the sound source should be recognizable in the sound field.

These three effects of indirect sound, both physical and subjective, can be quantified in scales, as will be shown later in this chapter. One of the aims of subjective musical acoustics is to relate subjective to objective scales. The objective and subjective effects are both based on one physical phenomenon (the indirect sound), and, therefore, the scales are not independent. Very often the values on different scales can be predicted from each other. Some scales have a positive extreme corresponding to a condition with a lot of indirect sound; some other scales have a positive extreme corresponding to a condition with no indirect sound.

## II. METHODOLOGY

The subjective effects of a room on the perceived sound can be separated in theoretical description. However, in practical situations and even under laboratory conditions, they can never be separated because they are all dependent on one physical source, the indirect sound. For this reason, and some other ones, the methodology of subjective acoustics research is a rather complicated affair. There are some methodological problems that are specific to subjective room acoustics. They give good insight into what kinds of research and results may be expected in this field. We will deal with three such problems.

First, there is the problem of subjective response. The most direct way to measure this response is to ask the listener to report verbally his or her subjective impression of the acoustics of a room or hall. This method was used by Beranek (1962), who based his work on interviews with musicians (mainly conductors and soloists) and music critics. Hawkes and Douglas (1971) and Wilkens (1977) made extensive use of semantic differentials. They opened the way for statistical analyses, including correlation and factor analyses. The subjective factors found can be, more or less successfully,

related to measured physical factors (Yoshida, 1965). An objection to this method is that one cannot be sure how the subjects interpret the verbal scales, especially with terms that are not applied to acoustical aspects in normal use. Better in this respect are nonverbal multidimentional scaling techniques, in which the similarity or dissimilarity between various acoustical conditions has be to compared (Yamaguchi, 1972; Edwards, 1974), or only a preference has to be reported (Schroeder *et al.*, 1974). These methods lead to unlabeled factors that can be filled in by comparing them with physical factors. Some researchers, such as Reichardt and co-workers in Dresden, made use of methods borrowed from psychophysics, such as detectability thresholds and difference limens. These methods permit only indirect conclusions concerning subjective aspects, but they are more reliable and more reproducible, both intra- and inter-individually, than subjective methods, such as interviews and semantic differentials.

A second problem lies in the method of presentation of the acoustical situation for subjective evaluation. In a number of studies audiences at live performances have been interviewed. An advantage of this method is its directness. However, it is difficult to compare performances in different halls or different performances in the same hall. In order to cope with these difficulties, researchers have performed experiments with synthetic sound fields, with the aid of which the acoustics of a hall are simulated in the laboratory (see also Wilkens & Kotterba, 1978).

These synthetic sound fields are constructed with loudspeakers in an anechoic chamber. The positions of the loudspeakers determine the angles of incident sound, thus simulating direct sound and indirect sound coming from several directions. All loudspeakers, except for the one for direct sound, are connected to time-delay and attenuation circuits. The music reproduced in such a synthetic sound field must have been recorded without reverberation. A typical setup for a synthetic sound field is illustrated in Fig. 2. It contains the following elements:

1. A loudspeaker in front of the subject, simulating the direct sound. Sometimes two loudspeakers at short distances are used.
2. Two loudspeakers, obliquely placed to the right and to the left. They simulate the first reflections from the walls. These reflection arrive 10 to 50 msec after the direct sound.
3. A loudspeaker mounted above the subject, simulating a ceiling reflection.
4. Several loudspeakers placed on all sides, simulating the later diffuse reverberation (see Fig. 2).

The time and intensity patterns of a sound field can be represented by a echogram or reflectogram. This is a diagram indicating the time delay and intensity of the various components of a sound field, determined relative to the direct sound. The reflectogram of a room or hall can be measured, for instance, by recording the acoustic response to an impulse (Fig. 3); the reflectogram of a synthetic sound field can be derived from its construction, as a matter of fact (Fig. 4). The reflection patterns used in synthetic sound fields are inspired by possible live reflectograms. A recent development in synthetic sound fields is to use headphones and to add the

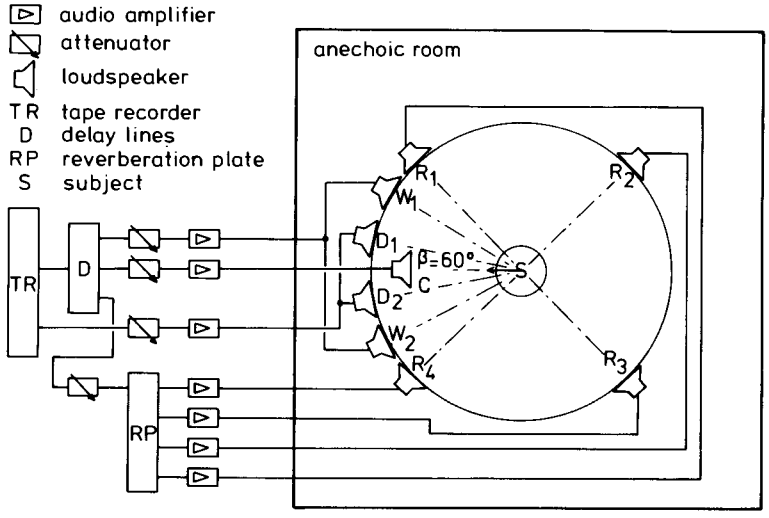



**Fig. 2.** Example of a synthetic sound field in an anechoic room. The depicted loudspeakers simulate the direct sound ($D_1$ and $D_2$), two wall reflections ($W_1$ and $W_2$), one ceiling reflection (C, with angle of incidence 60°), and diffuse reverberation (with four loudspeakers $R_1$, $R_2$, $R_3$, and $R_4$). The arrow indicates the viewing direction of the subject.

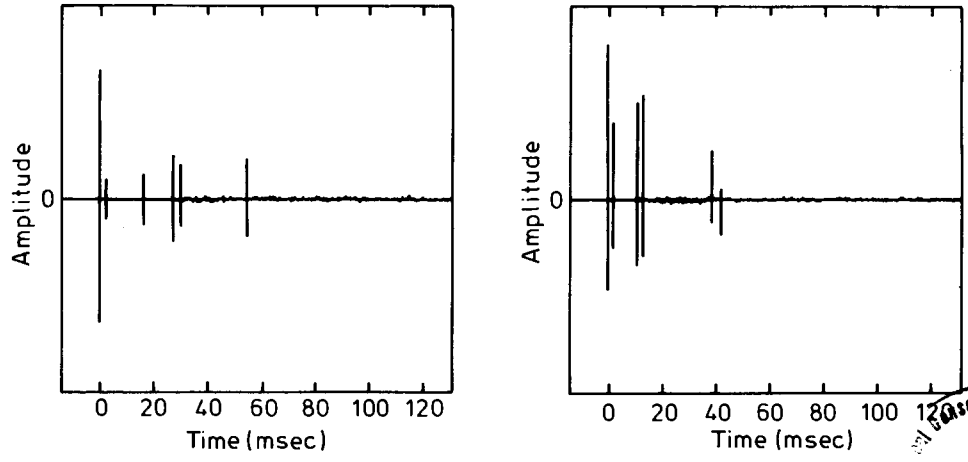



**Fig. 3.** Two examples of acoustic impulse responses recorded in rooms. The single peaks correspond to the direct sound and the early discrete reflections from walls and ceiling. These early reflections are followed by much weaker diffuse reverberation (based on Thiele, 1953, Figs. 15k and 16c.

Conservatorio Superior de Música - Madrid

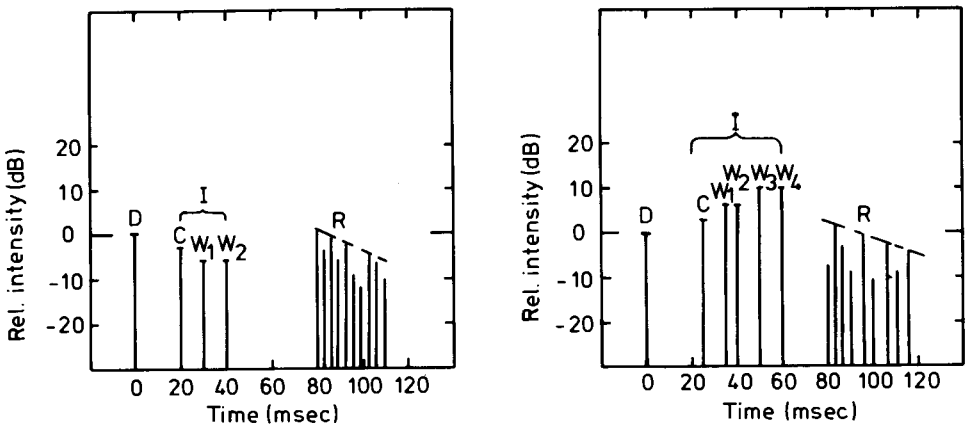



**Fig. 4.** Two examples of acoustic impulse responses of synthetic sound fields. Actually, the synthetic sound field is the result of the choice of a certain impulse response. The impulse response dictates the attenuation and the time delay of the various components of the sound field. The impulse responses depicted include direct sound (*D*), indirect sound (*I*) with wall reflections (*W*), ceiling reflections (*C*), and reverberation (*R*). The vertical scale is a logarithmic scale (based on Reichardt *et al.*, 1975, Figs. 3-2 and 4a).

acoustic response of a hall to "dead" recorded music with the help of filtering techniques (Schroeder, 1975, 1979).

A third problem lies in the choice of subjects. Who gives the best judgment? Average concert goers are often not aware of the acoustical properties of the hall in which they listen to musical performances. Acousticians and sound technicians do have such an awareness, but they may be less sensitive to the relevant musical criteria. Musicians have their own place in this respect, both literally and metaphorically. As a rule, the investigator tries to find subjects belonging to several of these categories mentioned, such as music critics, acousticians and technicians with a musical background or interest, composers, conductors, and so on.

## III. LEVEL EFFECTS OF INDIRECT SOUND: LOUDNESS

The increase in intensity caused by indirect sound is traditionally expressed in sound pressure levels (dB). However, our hearing system is not very sensitive to absolute levels unless these are outside the range of normal listening conditions. Much more important is the ratio between the intensities of the direct and the indirect sound fields. For this ratio we coin the term *indirect/direct ratio* (abbreviated as *i/d*) given by the formula

$$R = L_i - L_d \tag{1}$$

where $R$ is the indirect/direct ratio (in dB), $L_i$ the intensity of indirect sound (in dB), and $L_d$ the intensity of direct sound (in dB).

Wilkens's (1977) first subjective factor is clearly a level factor. It is characterized by the variables "large," "sounding," "loud," "brilliant," "strong," and "penetrating." Typical sound pressure levels of classical music performed in concert halls are within the range of 60 to 90 dB (Winckel, 1962).

## IV. TEMPORAL EFFECTS OF INDIRECT SOUND: DEFINITION

The subjective aspects of the temporal effects of indirect sound will be called *definition*. It is a negative scale in the sense that good definition implies little or no indirect sound. Other terms found in the literature are "clarity" (Beranek, 1962, pp. 36–40) and "clearness" (Mafune & Yoshida, 1968). The German term is *Durchsichtigkeit* (Reichardt, Abdel Alim, & Schmidt, 1975; Reichardt & Lehmann, 1976). Definition may be described subjectively as that which enables the listener to distinguish temporal details in the musical sound, and, as such, it is a necessary condition for listening to music. Beranek (1962) distinguishes between horizontal definition (holding for successive sounds) and vertical definition (holding for simultaneous sounds). The same distinction is made by Reichardt (1975) when he refers to temporal and register definition. The physical counterpart of vertical or register definition is not very clear, however. Hawkes and Douglas (1971) do not make this distinction. Their definition factor correlates with the variables "good definition," "clear," and "brilliant." Wilkens's (1977) second factor is a definition factor. Variables strongly loading this factor are "clear," "concentrated," and "definite." It is remarkeable that some evaluative variables like "pleasant," "liked," and "nice" had their highest loadings on this factor.

Mafune and Yoshida (1968) found a high correlation between definition and the intelligibility score for speech. There was also a high correlation with a subjective measure for listening comfort. Yamaguchi (1972) found two subjective definition factors, both correlated with intelligibility: speech definition (a $d/i$ ratio taking the indirect sound within 50 msec after the direct sound as direct sound) and sound pressure level. The physical factors supporting a good musical definition are evidently closely related to the factors ensuring a good understanding of speech.

Several physical scales can be given that quantify the temporal effects of indirect sound. It must be mentioned, however, that the relationship between the physical and subjective aspects have not yet been worked out in detail for all scales. The classical scale is the *reverberation time* $T$, the time required for the sound intensity to decay by 60 dB after abruptly stopping the sound source. This can be extimated from the physical characteristics of a room or hall by the following formula:

$$T = \frac{V}{6\alpha S} \tag{2}$$

where $T$ is the reverberation time in sec, $V$, the volume in $m^3$, $S$ the surface of bounding areas in $m^2$, and $\alpha$ the mean absorption coefficient of the boundaries (fraction of sound energy not reflected).

Optimal reverberation times have been reported in the literature. Kuhl (1954) mentions: 1.5 sec for classical and contemporary music and 2.1 sec for romantic music. Beranek's figures (1962, pp. 425–431) differ only slightly: 1.5 sec for baroque music and Italian opera, 1.7 sec for classical music and Wagnerian opera, and 2.1 sec for romantic music. Optimal reverberation times for music are higher than for speech.

A second physical measure of definition is the *modulation transfer function* or MTF

(Houtgast & Steeneken, 1973; Steeneken & Houtgast, 1980). If the intensity of a sound source is modulated, the modulation depth at a distance decreases to a greater or lesser extent because of the indirect sound. With a lot of indirect sound the valleys in the temporal envelope will be filled, the more so for higher modulation frequencies. The degree of modulation retained at the listener's position may be used as a measure of the influence of the indirect sound. The modulation transfer depends on the modulation frequency so that the actual measure is a curve, the modulation transfer function. For speech the relevant modulation frequency range is from 0.4 to 20 Hz. By weighting the modulation frequencies, the information in the curve can be condensed to a single measure. Up to now, the modulation transfer approach has only been applied to speech communication problems, but application to musical acoustics, both objective and subjective, seems worthwhile. Macfadyen's (1970) confusion index, the minimum perceivable modulation depth expressed in dB of an amplitude-modulated white noise (modulation frequency 10 Hz), is related to the modulation-transfer approach. Macfadyen used his measure to assess the subjective definition of different seating positions in a hall under various acoustical conditions.

A third physical measure for definition is an adaptation of the *i/d* ratio, called *clarity* by Reichardt *et al.* (1975). It is actually a *d/i* ratio, in which the direct sound has been expended to include in addition the indirect sound coming within 80 msec after the direct sound:

$$C = L'_d - L'_i \tag{3}$$

where $C$ is the clarity (physical definition) (in dB), $L'_d$ is the intensity of direct sound plus indirect sound within 80 msec in (dB), $L'_i$ is the intensity of indirect sound, arriving more than 80 msec after the direct sound (in dB). With this measure a good prediction of subjective definition is possible. Definition seems to be optimal when there is no indirect sound. This is not realistic in practical situations because another indispensable, positive aspect, spaciousness, depends on indirect sound. Reichardt *et al.* (1975) state that clarity should be at least 1.6 dB.

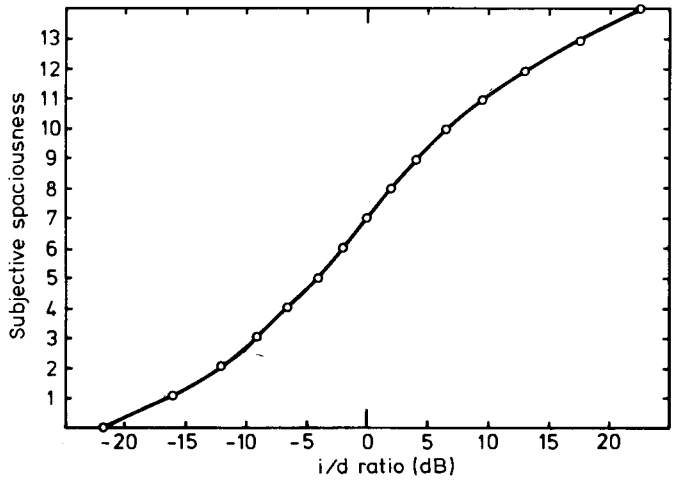
## V. SPATIAL EFFECTS OF INDIRECT SOUND: SPACIOUSNESS

The subjective aspects of the spatial effects of indirect sound are indicated here by the term *spaciousness*. In the literature no prevailing term has come up yet. One finds terms such as "liveness" (Maxfield & Albersheim, 1947; Beranek, 1962), "richness" (Mafune & Yoshida, 1968; Yoshida, 1965), "ambience" (Lochner & DeVilliers Keet, 1960), "fullness of tone" (Beranek, 1962), "spatial responsiveness" (Marshall, 1967), "spatial impression" (Barron, 1971), "resonance" (Hawkes & Douglas, 1971), and "reverberance" (Hawkes & Douglas, 1971). In the German literature the list of terms is restricted to *Raumeindruck*, *Räumlichkeit* and *Halligkeit* (room impression, spaciousness, and reverberance, respectively, in papers by Reichardt and co-workers; and Kuhl 1977, 1978). The German authors treat *Raumeindruck* as a generic term, with *Räumlichkeit* and *Halligkeit* as special aspects.

The subjective aspects of spaciousness have been described by Maxfield and Albersheim (1947) as follows: (1) a change in the general tone quality, stated by musicians to be improved "resonance" or "roundness;" (2) the blending of the sound from the various instruments of an orchestra into a single coordinated sound; (3) the sense of acoustic perspective; and (4) the realization on the part of the listener of the approximate size of the auditorium. Beranek (1962, pp. 22-24) mentions the following aspects of "liveness" or a "live room": more uniform loudness, enhancement of bass and treble, fullness of tone, range of crescendo, sound diffusion, intimacy and texture. In this list level and temporal effects are also included. Reichardt *et al.* (1974), Kuhl (1977), and Reichardt and Lehmann (1978a) give lists that do not differ essentially from the items mentioned. Hawkes and Douglas (1971) describe a factor resonance/reverberance, characterized by the variables "resonant," "reverberant," "responsive," and "large dynamic range." Wilkens' (1977) third factor may be regarded as spaciousness factor, with the corresponding variables "weak," "round," "blunt," "dark," and "not-treble."

Maxfield and Albersheim (1947) first connected "liveness" with the *i/d ratio.* This approach has in particular been elaborated by Reichardt and co-workers at the Technological University of Dresden. They used the *d/i* ratio, referred to as *Hallabstand* (Schmidt & Lehmann, 1974). Their first publication (Reichardt & Schmidt, 1966) describes a spaciousness scale with 15 subjectively just-distinguishable points. The relation between the *i/d* ratio and subjective spaciousness is represented in Fig. 5. This scale was based on measurements with synthetic sound fields consisting of direct sound and diffuse reverberation with $T = 2$ sec.

In later research it became evident that not only the amount of indirect sound but




Biblioteca del Real Conservatorio Superior de Música - MADRID

**Fig. 5.** Subjective spaciousness as a function of the *i/d* ratio (based on Reichardt & Schmidt, 1966, Fig. 5).

also its temporal spread affects spaciousness. Sound fields with equal *i/d* ratios but different reverberation times may have slightly different spaciousnesses. Actually, the early discrete reflections that come before the diffuse reverberation act subjectively as direct sound, not as indirect sound. Also, the angle of incidence influences the subjective spaciousness. Indirect sound that comes from the frontal direction strengthens the direct sound subjectively.

Reichardt *et al.* (1974, 1978) and Reichardt and Lehmann (1978a,b) summarize the spaciousness effects of the various indirect components of a sound field as follows:

1. Indirect sound arriving within 25 msec after the direct sound counts as direct sound.
2. Sound arriving between 25 to 80 msec after the direct sound must be divided into two components: (2a) the sound arriving with an angle up to 40° relative to the direct sound must be counted as direct sound, and (2b) the sound arriving from side and rear directions must be counted as indirect sound.
3. Sound arriving later than 80 msec after the direct sound must be counted as indirect sound.

With these rules a corrected *i/d* ratio can be constructed to predict spaciousness:

$$iR = L''_i - L''_d \tag{4}$$

where $R$ equals spaciousness, $L''_d$ the intensity of direct sound, plus indirect sound, within 25 msec from all directions and within 80 msec from front directions (in dB), and $L''_i$ the intensity of all other components of indirect sound (in dB).

Reichard and Lehmann (1978b) found the correlation between their *i/d* ratio and the subjectively judged spaciousness as determined in two concert halls to be 0.64 and 0.65, respectively. *I/d* ratios in these concert halls differed with seat position but were mostly within the range of 2 to 4 dB (Reichardt & Sarkov, 1972).

It is well known that traditional rectangular concert halls—like the Boston Symphony Hall, the Grosser Musikvereinsaal in Vienna, and the Concertgebouw in Amsterdam—have excellent acoustics, very often better than modern halls, which are wide and low. Marshall (1967) related this observation to the relatively strong early reflections. In high, rectangular halls these reflections come from the side walls; in wide, low halls, from the ceiling. Since our ears are located in the horizontal plane, lateral reflections introduce interaural differences that are important in contributing to the perception of spaciousness. For this reason, reflection panels hanging from the ceiling may affect speech intelligibility and musical definition positively, but their spaciousness effects are doubtful.

Barron (1971) conducted detailed research concerning early lateral reflections. He used synthetic sound fields consisting of direct sound and a single side reflection at 40° of varying intensity and delay time. His results may be summarized as follows:

1. Reflections within 10 msec after the direct sound are too early; they result in a subjective sideward shift of the sound source.

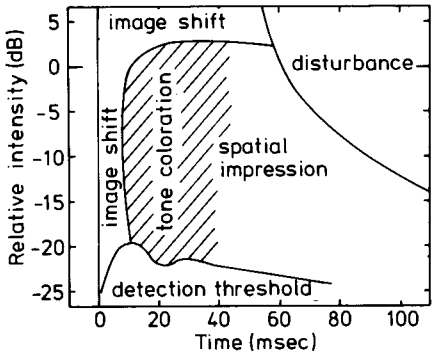



**Fig. 6.** Subjective effects of a sound field, consisting of direct sound and a lateral reflection with an angle of incidence of 40° with variable attenuation and time delay. The "tone coloration" is caused by the addition of a sound signal and its delayed repetition. The disturbances are usually called echoes (from Barron, 1971, Fig. 5).

2. Relatively strong reflections later than 50 msec after the direct sound disturb the sound image. They are perceived as echoes, distinct from the direct and early indirect sound.

3. Reflections of 20 to 25 dB weaker than the direct sound are below threshold.

4. In between the aforementioned effects there is a region of *spatial impression*. Most important for this are the reflections arriving between 40 and 100 msec after the direct sound. Reflections between 10 and 40 msec can give rise to a distortion of the timbre as an effect of the addition of the direct sound with its delayed repetition (see also Lochner & DeVilliers Keet, 1960 and Bilsen, 1968). Barron's results are illustrated in Fig. 6. The importance of lateral reflections was confirmed by Kuhl's (1978) experiments.

## VI. THE COMPROMISE BETWEEN DEFINITION AND SPACIOUSNESS

The right amount of definition and spaciousness are decisive for good subjective room acoustics. However, definition is negatively correlated with indirect sound while spaciousness is positively correlated. So, it will not come as a surprise that definition and spaciousness have a high negative intercorrelation. This means that in practice a compromise between requirements for definition and spaciousness is always necessary.

However, comparing the formulas for clarity (objective definition) and objective spaciousness, one will notice that one component of the sound field affects both definition and spaciousness positively. It is the sound coming from the sides and from the rear later than 40 msec and earlier than 80 msec after the direct sound. Coming earlier than 80 msec, it functions as direct sound with respect to definition. Coming

from nonfront directions, it functions as indirect sound with respect to spaciousness. It may be concluded that these reflections are of great importance for good acoustics of a room or hall.

## VII. CONCLUSION

Subjective musical room acoustics is a relatively new field of scientific enquiry. It does not possess time-honored concepts, methods, or basic results. It is not yet a standard component of handbooks, textbooks, university curricula, or scientific institutions. Its nature is to a large extent interdisciplinary. Methods and concepts have been derived from physical acoustics as well as from psychology and from musicology. However, the rapid growth of the literature concerning its problem areas during the last decade shows that it will soon become an undispensable part of both room acoustics and psychophysics.

Out of the recent literature two important subjective features of room acoustics have emerged: definition and spaciousness. A number of definitions and measurement procedures for both the subjective attributes and its objective counterparts have been proposed, which include a wide range of ways of thinking. Much research will still be needed before the interrelation between the two basic concepts as well as their connection with physical-acoustical properties are fully understood. However, there are several methods available that promise good progress, such as simulation techniques for room acoustics, the indirect/direct ratio, and the modulation-transfer function. These methods have a mainly physical background. They should be applied along with methods such as factor analysis and multidimensional scaling, which have a predominantly psychological origin. The recognition of the hybrid nature of subjective musical room acoustics is essential for the solving of its questions and problems.

## REFERENCES

Barron, M. The subjective effects of first reflections in concert halls—the need for lateral reflections. *Journal of Sound & Vibration*, 1971, *15*, 475–494.

Beranek, L. L. *Acoustics*. New York: McGraw-Hill, 1954.

Beranek, L. L. *Music, acoustics and architecture*. New York: Wiley, 1962.

Bilsen, F. A. On the Interaction of Sound With its Repetitions. Dissertation, Technological University Delft, 1968.

Edwards, R. M. A subjective assessment of concert hall acoustics. *Acustica*, 1974, *39*, 183–195.

Hawkes, R. J., & Douglas, H. Subjective acoustic experience in concert auditoria. *Acustica*, 1971, *24*, 135–150.

Houtgast, T., & Steenken, H.J.M. The modulation transfer function in room acoustics as a predictor of speech intellegibility. *Acustica*, 1973, *28*, 66–73.

Kuhl, W. Über Versuche zur Ermittlung der günstigsten Nachhallzeit grosser Musikstudios. *Acustica*, 1954, *4*, 618–634.

Kuhl, W. In der Raumakustik benützte hörakustische Termini. *Acustica*, 1977, *39*, 57–58.

Kuhl, W. Räumlichkeit als Komponente des Raumeindrucks. *Acustica*, 1978, *40*, 167–181.

Kuttruff, H. *Room acoustics*. London: Applied Science Publishers, 1973.

Lochner, J. P., & De Villiers Keet, W. Stereophonic and quasi-stereophonic reproduction. *Journal of the Acoustical Society of America*, 1960, *32*, 393–401.

Macfadyen, K. A. A method of assessing musical definition in an auditorium. *Applied Acoustics*, 1970, *3*, 181–190.

Mafune, Y., & Yoshida, T. Psychometric approach to the room acoustics, Report II: Distribution of subjective evaluations in five auditoriums and discussion. *Sixth International Congress on Acoustics, Tokyo*, Paper E-2-9, 1968.

Marshall, A. H. A note on the importance of room cross-section in concert halls. *Journal of Sound and Vibration*, 1967, *5*, 100–112.

Maxfield, J. P., & Albersheim, W. J. An acoustic constant of enclosed spaces with their apparent liveness. *Journal of the Acoustical Society of America*, 1947, *19*, 71–79.

Meyer, J. *Acoustics and performance of music.* Frankfurt am Main: Das Musikinstrument, 1978.

Rasch, R. A. Subjective muzikale zaalakoestiek: Een overzicht. *Mens en Melodie*, 1977, *32*, 70–81.

Reichardt, W. Vergleich der objectiven raumakustischen Kriterien für Musik. *Hochfrequenztechnik und Elektroakustik*, 1970, *79*, 121–128.

Reichardt, W., Abdel Alim, O., & Schmidt, W. Definition und Messgrundlage eines objektiven Masses zur Ermittlung der Grenze zwischen brauchbarer und unbrauchbarer Durchsichtigkeit bei Musikdarbietung. *Acustica*, 1975, *32*, 126–137.

Reichardt, W., & Lehmann, U. Sind Raumeindruck und Durchsichtigkeit des Hörerlebnisses im Konzertsaal Gegensätze? *Applied Acoustics*, 1976, *9*, 139–150.

Reichardt, W., & Lehmann, U. Raumeindruck als Oberbegriff von Räumlichkeit und Halligkeit, Erläuterungen des Raumeindrucksmasses R. *Acustica*, 1978, *40*, 277–290. (a)

Reichardt, W., & Lehmann, U. Definition eines Raumeindrucksmasses R zur Bestimmung des Raumeindrucks bei Musikdarbietungen auf der Grundlage subjektiver Untersuchungen. *Applied Acoustics*, 1978, 99–127. (b)

Reichardt, W., & Sarkov, N. Bestimmung der optimalen Räumlichkeit für eine Musikprobe im Bezugsschallfeld. *Zeitschrift für elektronische Informations- und Energietechnik*, 1972, *2*, 49–52.

Reichardt, W., & Schmidt, W. Die hörbaren Stufen des Raumeindrucks bei Musik. *Acustica*, 1966, *17*, 175–179.

Reichardt, W., Schmidt, W., Lehmann, U., & Ahnert, W. Definition und Messgrundlagen eines "wirksamen Hallabstand" als Mass für den Raumeindruck bei Musikdarbietungen. *Zeitschrift für elektronische Informations- und Energietechnik*, 1974, *4*, 225–233.

Schmidt, W., & Lehmann, U. Eignung von Hallabstand oder Hallmass zur objectiven Bestimmung des Raumeindrucks. *Zeitschrift für elektronische Informations- und Energietechnik*, 1974, *4*, 161–168.

Schroeder, M. F. Models of hearing. *Proceedings of the IEEE*, 1975, *63*, 1332–1350.

Schroeder, M. F. *Music perception in concert halls.* Stockholm: Royal Swedish Academy of Music, 1979.

Schroeder, M. F., Gottlob, D., & Siebrasse, K. F. Comparative study of European concert halls: correlation of subjective preference with geometric and acoustic parameters. *Journal of the Acoustical Society of America*, 1974, *56*, 1195–1201.

Steeneken, H.J.M., & Houtgast, T. A physical method for measuring speech-transmission quality. *Journal of the Acoustical Society of America*, 1980, *67*, 318–326.

Thiele, R. Richtungsverteilung und Zeitfolge der Schallrückwürfe in Räumen. *Acustica*, 1953, *3*, 291–302.

Wilkens, H. Mehrdimensionale Beschreibung subjektiver Beurteilungen der Akustik von Konzertsälen. *Acustica*, 1977, *38*, 10–23.

Wilkens, H., & Kotterba, B. Vergleich der Beurteilung verschiedener raumakustischer Situationen bei Anregung eines Raumes mit einem Orchester oder Lautsprechern. *Acustica*, 1978, *40*, 291–297.

Winkel, F. Optimum acoustic criteria of concert halls for the performance of classical music. *Journal of the Acoustical Society of America*, 1962, *34*, 81–86.

Yamaguchi, K. Multivariate analyses of subjective and physical measures of hall acoustics. *Journal of the Acoustical Society of America*, 1972, *52*, 1271–1279.

Yoshida, T. Psychometric approach to the room acoustics. *Fifth Internation Congress on Acoustics, Liège*, Paper B 18 (1965).

# 6

# Rhythm and Tempo

*Paul Fraisse*


I. Definitions ..... 149
II. Rhythm and Spontaneous Tempo ..... 151
A. Spontaneous Rhythmic Movements ..... 151
B. Spontaneous Tempo ..... 153
C. Motor Induction and Synchronization ..... 154
D. Subjective Rhythmization ..... 155
III. Rhythmic Forms ..... 157
A. Regular Groupings ..... 157
B. Patterns in Time ..... 162
C. Patterns of Time ..... 165
IV. The Perception of Musical Rhythms ..... 170
V. Conclusion ..... 175
References ..... 177


## I. DEFINITIONS

The task of those who study rhythm is a difficult one, because a precise, generally accepted definition of rhythm does not exist. This difficulty derives from the fact that rhythm refers to a complex reality in which several variables are fused. Our aim will be to distinguish these variables successively. However, since this work is devoted to music, it is necessary to emphasize that the problem has been complicated by music theorists who have often chosen, due to their personal aesthetic preferences, to recognize only one of the several aspects of rhythm.

Can etymology help us? Rhythm comes from the Greek words ϱυθμος (rhythm) and ϱέω (to flow). However, as Benveniste (1951) has shown, the semantic connection between rhythm and flow does not occur through the intervention of the regular movement of waves, as was often believed. In Greek one never uses *rheo* and *rhythmos* when referring to the sea. *Rhythmos* appears as one of the key words in Ionian philoso-



Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

phy, generally meaning "form," but an improvised, momentary, and modifiable form. *Rhythmos* literally signifies a "particular way of flowing." Plato essentially applied this term to bodily movements, which, like musical sounds, may be described in terms of numbers. He wrote in *The Banquet* "The system is the result of rapidity and of slowness, at first opposed, then harmonized." In *The Laws* he arrived at the fundamental definition that rhythm is "the order in the movement." We will adopt this definition, which, even in its generality, conveys different aspects of rhythm.

However, an essential distinction asserts itself. Rhythm is the ordered characteristic of succession. This order may be conceived or perceived. We speak of the rhythm of the days and of the nights, of the seasons, and of rapid or of very slow physical phenomena (such as that of light frequencies or of the planets, respectively). If by direct or by indirect observation, we ascertain the successive phases of these phenomena, in none of these cases do we directly perceive the order—that is to say, the succession of the phases itself. The rhythm is thereby inferred from a mental construction.

However, there exist cases in which there is, properly speaking, the perception of rhythm, such as in dance, song, music, and poetry. We then find precisely the connection that Plato made between order and human movement: All of the rhythms that we perceive are rhythms which originally resulted from human activity. The first psychologists of the nineteenth century felt this relationship. Mach (1865) placed motor activity at the center of our experience of rhythm, and Vierordt (1868), several years later, began to record rhythmic movements.

Let us, nevertheless, insist on the fact that rhythm is a perceptual quality specifically linked to certain successions, a *Gestaltqualität*, according to Von Ehrenfel's definition. At this point, in order to clarify the rest of our discussion, it is necessary to specify the characteristic traits of this rhythmic perception.

Most generally, we say that there is rhythm when we can predict on the basis of what is perceived, or, in other words, when we can anticipate what will follow. In this guise, we return to the idea of order found in Plato and in the most modern definitions, such as Martin's (1972): "Inherent in the rhythmic concept is that the perception of early events in a sequence generates expectancies concerning later events in real time (p. 503)."

This characteristic appears in its true form if we compare rhythm with arrhythmia. All sequences of random stimulations will be considered arrhythmic (see Section III,C,1). Nevertheless, one can more or less anticipate what is to follow, and from this the difficulties arise. At one extreme, we have the isochronous repetition of the same stimulus: the pulse, the march, the tick-tock of a clock. This repetition can be that of a pattern of stimuli having analogous structures, as in a waltz or an alexandrine. At the other extreme, we have a succession of relatively different patterns, as in free verse or in certain modern music.

The anticipation can only be temporal; that is to say, linked to the organization within the duration. But what is organized in this way? At this point, there is often a misunderstanding. Is rhythm born out of a series of stimuli, whose temporal characteristics are fundamental, as one could be led to believe by a description of a poetic

sequence, in terms of breves and longs, or by the reading of a musical score where each note is of a precise length? Or is rhythm born out of the ordering of the temporal intervals among the elements marked by a difference in intensity, of pitch or of the timbre? These two propositions are opposed by all theorists of rhythm. The problem is, without doubt, as old as is music. Plato in *The Republic* already made fun of a critic of his era: "I vaguely remember" he wrote, "that he spoke of anapaestic verse. . . ; I don't know how he arranged them and established the equality of the up beat and of the fall by a combination of long and breve. . . ." The problem remains: Is rhythm the arrangement of durable elements, or is it the succession of more or less intense elements, the upbeat and the fall, the *arsis* and the *thesis* of the Greeks being the most simple example? We will see that both forms of organization exist, one type of relation prevailing over the other. Moreover, they are most often linked and interdependent. Rhythm is the perception of an order.

One of the perceptual aspects of rhythmic organization is *tempo*. It can be lively or slow. It corresponds to the number of perceived elements per unit time, or to the absolute duration of the different values of the durations. Evidently, one passed from a definition based on frequency to a definition based on duration. We will use both of them. The possibility of rhythmic perception depends on tempo, because the organization of succession into perceptible patterns is largely determined by the law of proximity. When the tempo slows down too much, the rhythm and also the melody disappear.

We have chosen to begin this chapter with the most simple perceptions and to end up with the most complex ones, which are evidently those of artistic rhythms. The reason is not because the simple explains the complex, but because simple configurations can be more easily analyzed (see also Fraisse, 1956, 1974).

## II. RHYTHM AND SPONTANEOUS TEMPO

### A. Spontaneous Rhythmic Movements

The most easily perceived rhythm is one that is produced by the simple repetition of the same stimulus at a constant frequency. In the rest of this article, we will call this a *cadence*. The simplest examples are the beating of a clock or of a metronome. But the most important fact is that these rhythms are characteristic of some very fundamental activities such as walking, swimming, and flying. Both animals and people move about with rhythmic movements characteristic of their species. The first rhythmic movement found in the human new-born is sucking, with periods that follow at intervals of from 600 to 1200 msec. This regularity is interrupted by spontaneous pauses, but sucking movements occur at a cadence that seems to be characteristic for each infant. Later on, walking appears. While one of the limbs supports the body weight, the other swings forward, before serving, in turn, as support. In the adult there is also a brief period (100 msec) of double support. The duration of the step is about 550 msec, and corresponds to a frequency of 110–112 per minute (Mishima,

1965). This frequency depends a little on anthropometric differences between individuals, age, and environmental conditions. This spontaneous activity, which is similar to a reflex, is a fundamental element of human motor activity. It plays an important role in all of the rhythmic arts.

Spontaneous activities reveal that physiological settings exist in the human organism, and, more generally, in all living things, which are regulated by peripheral afferents and, above all, by nervous centers situated at different levels. The tempo of walking, for example, seems to be determined by the medulla. From these centers and from their activity, we have other manifestations, among which it is necessary to cite the heartbeat (with an average of 72 beats per minute) and the electrical oscillations of the cerebral cortex with frequencies varying from 1 to 3 per second for delta waves, from 14 to 30 for beta waves, and most characteristically, from 8 to 13 for alpha waves.

We still do not know where the different biological clocks that assure the regularity of these phenomena are located. However, one can think that an autorhythmicity is characteristic of certain nervous tissues. Even though Sherrington has shown that a nerve follows the rhythm of an excitation whatever its frequency, it is not the same when the excitation crosses a nervous center. The frequency of the response is then different from that of the excitation. Also, the myotatic reflex follows the cadence of a mechanical excitation up to a frequency of 4 to 5 per second. For values higher than this, a halving of the frequency of the response occurs every second shock, and for even higher values, every third shock is effective (Paillard, 1955). This fact, reminiscent of the general properties of oscillating circuits, is suggestive.

Among spontaneous movements, it is necessary to cite rocking, which clearly intervenes in games or in dances, but which is manifested from the most tender age on by the beating of the foot of the newborn lying on his back (average age of appearance 2.7 months). As soon as the child can remain seated, the rocking of the trunk appears (toward 6 months). This rocking of the trunk can be considered essentially as a movement of the head, of which one observes different modalities: rocking while on all fours (forward-backward), standing, or on the knees. In most children these movements are transitory, but in others they can last for months, sometimes until 2–3 years and even until 5 years of age. One encounters, moreover, other forms of rocking in the older child (movement of the legs, for example, while seated). Also, the use of the rocking-chair is not without a relation to this type of behavior. There is little precise information regarding the frequency of these rockings. They occur within the range of spontaneous tempos (.5–2 per second), and this frequency depends on the muscular mass concerned.

It is necessary to note that rocking is related neither to vegetative functions (as is the heart) or to relational functions (sucking, walking); it appears when the child is idle or at the moment of falling asleep. In the adult it also translates into an absence of voluntary control or a state of distraction. These movements seem to correspond to a regulation of nervous tension. The postural activity with its tonic effects then takes preponderance over the relational activity (Wallon, 1949). It above all appears when the possibility of communicating with the environment is reduced, as when the movements of the child are restricted by the intervention of an adult or by illness. It is

frequent in mental deficiency, neurosis or dementia. In all of these cases, rocking seems to aim at the maintenance of a state of excitation, and it has a heavy affective connotation.

## B. Spontaneous Tempo

The periodic activities that we have just mentioned have their own spontaneous tempo. Stern (1900) thus thought that a psychic tempo characteristic of voluntary activity exists. In order to determine this tempo, he proposed a simple motor activity: tapping a spontaneous tempo on a table. History has shown the fecundity of this proposed test, but has not confirmed the existence of a psychic tempo characteristic of all an individual's activities. There are only weak correlations between the different repetitive tests executed at a spontaneous tempo. Factorial analysis always reveals a plurality of factors (Allport and Vernon, 1933; Rimoldi, 1951).

Spontaneous tempo, also called *personal tempo* (Frischeisen-Köhler, 1933a) or *mental tempo* (Mishima, 1951–1952) and measured by the natural speed of tapping, is of great interest. The length of the interval between two taps varies, according to the authors, from 380 to 880 msec. One can assert that a duration of 600 msec is the most representative. All of the research underscores the great interindividual variability of this tempo (from 200 to 1400 msec, Fraisse, Pichot, & Clairouin, 1949). By contrast, individual variability is slight. One can verify this within a trial: the variability of intervals is from 3 to 5%, which is in the range of the differential threshold for durations of this type. Also, there is great reliability from one trial to another: the correlations are of the order of .75 to .95 (Harrel, 1937; Rimoldi, 1951). This reliability indicates that spontaneous tempo is characteristic of the individual, a statement reinforced by twin reserach. Differences in tempo between two identical twins (homozygous) are no larger than between two executions of spontaneous tempo by the same subject; however, the differences between two heterozygous twins are as great as between two individuals chosen at random (Frischeisen-Köhler, 1933a; Lehtovaara, Saarinen, & Järvinen, 1966).

Spontaneous tempo of the forefinger has a good correlation with that of the palm of the hand, with the swinging of the leg of a seated subject, and with the swinging of the arm when the subject is standing (Mishima, 1965).

It is necessary to distinguish between spontaneous tempo and *preferred tempo*. The latter corresponds to the speed of a succession of sounds or of lights that appears to be the most natural—that is to say, to a regular succession judged as being neither too slow nor too fast. Since the nineteenth century a number of German scientists have sought the interval which appeard to be neither too short nor too long. The most frequent determination has been about 600 msec. In this regard, Wundt examined the natural duration of associations between two perceptions, and he proposed a value of 720 msec.

Since Wallin (1911), the preferred tempo has most frequently been measured using a metronome. The results found are fairly close to 500 msec (Wallin, 1911;

Frischeisen-Köhler, 1933b; Mishima, 1956). Possibly, this value is in part determined by the scale of tempos which the metronome offers. The preferred tempo of an individual is an constant as is spontaneous tempo, but the correlations between the two tests are not higher than .40 (Mishima, 1965).

It is striking that the rhythm of the heart, of walking, of spontaneous and of preferred tempo are of the same order of magnitude (intervals of from 500 to 700 msec). It has been tempting to study whether one of these rhythms serves in some way as a sort of *pacemaker* for the others. The rhythm of the heart, the most often invoked, is not correlated with spontaneous tempo (Tisserand & Guillot, 1949–1950). Moreover, it has been verified several times that an acceleration of the hearbeat does not correspond to an acceleration of spontaneous tempo. By contrast, one finds a noteworthy correlation between the rhythm of walking and of spontaneous tempo (.28, Harrison, 1941; .68, Mishima, 1965). However, we cannot assume that one phenomenon can be explained by the other. There is only a narrow range of frequencies of natural or voluntary rhythms and of preferred tempo.

### C. Motor Induction and Synchronization

Spontaneous motor tempo and preferred tempo do not only have comparable frequencies, but observations and experiments show that they are also often associated. People fairly easily accompany with a motor act a regular series of sounds. This phenomenon spontaneously appears in certain children toward one year of age, sometimes even earlier. Parents are surprised to see their child sitting or standing, rocking in one way or another while listening to rhythmic music. From the age of 3–4 years on, the child is capable of accompanying, when requested, the beating of a metronome (Fraisse *et al.*, 1949). This accompaniment tends to be a synchronization between the sound and a tap—that is to say, that the stimulus and the response occur simultaneously.

This behavior is all the more remarkable, as it constitutes an exception in the field of our behaviors. As a general rule, our reactions succeed the stimuli. In synchronization the response is produced at the same time as the appearance of the stimulus. A similar behavior is possible only if the motor command is anticipated in regard to the moment when the stimulus is produced. More precisely, the signal for the response is not the sound stimulus but the temporal interval between successive signals. Synchronization is only possible when there is anticipation—that is, when the succession of signals is periodic. Thus, the most simple rhythm is evidently the isochronal production of identical stimuli. However, synchronization is also possible in cases of more complex rhythms. What is important is not the regularity but the anticipation. The subjects can, for example, synchronize their tapping with some series of accelerated or decelerated sounds, the interval between the successive sounds being modified by a fixed duration (10, 20, 50, 80, or 100 msec). Synchronization, in these cases, remains possible, but its precision diminishes with the gradient of acceleration or of deceleration (Ehrlich, 1958).

The spontaneity of this behavior is attested to by its appearance early in life and also by the fact that the so-called evolved adult has to learn how to inhibit his involuntary movements of accompaniment to music. Experiments confirm these observations. When subjects were presented with a regular series of sounds and asked to tap for each sound, they spontaneously synchronized sound and tap. When asked not to synchronize but to respond *after* each sound, as in a reaction-time experiment, all of the subjects found this task difficult, the more so the higher the frequency of sounds (Fraisse, 1966). The same difficulty arose when the subjects were asked to syncopate—that is, to interpolate the series of taps between the series of sounds. The subjects habitually succeeded only when the intervals between the sounds were longer than one second (Fraisse & Ehrlich, 1955). Conversely, it has been shown (Fraisse, 1966) that synchronization is established very rapidly, and that it is acquired from the third sound on. Let us add, in anticipation of what is to follow, that the synchronization of repetitive patterns is also realized from the third pattern on.

Not only is synchronization possible at the frequency of preferred tempo, but it is also possible in the whole range of frequencies of spontaneous tempos. More precisely, one observes that synchronization is most regular for intervals of 400 to 800 msec. If the frequencies are faster or slower, the separation between taps and sounds is more variable. For rapid cadences it is, above all, a perceptual problem: the interval between two sounds is not perceived exactly enough to permit precise synchronization (Michon, 1964). The subject oscillates between exaggerated anticipations and delays when the tap follows the sound as in a reaction time situation (Fraisse, 1966). In conclusion, the range in which synchronization is possible is at sound intervals of 200 to 1800 msec between the sounds.[1]

The synchronization that we have considered in its most elementary forms plays a fundamental role in music, not only in dance but also in all instances in which several musicians play together. The unity of their playing is possible only when they are capable of anticipation. One of the roles of the conductor of an orchestra is to furnish the signals that will result in synchronization between musicians.

## D. Subjective Rhythmization

If one listens to identical sounds that follow each other at equal intervals, that is to say, a cadence, these sounds seem to be grouped by twos or by threes. Since nothing

[1]In speaking of synchronization, it is necessary to specify what is synchronized with what. In effect, if one measures the temporal separation between a tap of the forefinger and the sound, one finds that the tap slightly anticipates the sound by about 30 msec. The subject does not perceive this error systematically. This was pointed out as early as 1902 by Miyake. Moreover, this error is greater if the sound is synchronized with the foot. The difference between hand and foot permits us to think that the subject's criterion for synchronization is the coincidence of the auditory and of the tactile-kinesthetic information at the cortical level. For this coincidence to be as precise as possible, the movement of tapping should slightly precede the sound in order to make allowance for the length of the transmission of peripheral information. This length is all the greater when the distance is longer (Fraisse, 1980, pp. 252–257).

objectively suggests this grouping, this phenomenon has been termed *subjective rhythmization*. This expression, which appeared at the end of the nineteenth century (Meumann, 1894; Bolton, 1894), must today be considered inadequate, because all perceived rhythm is the result of an activity by the subject since, physically, there are only successions. The observations made using sound series were later confirmed by using visual series (Koffka, 1909). When one thus listens to a cadence, introspection reveals that grouping seems to correspond to the lengthening of one of the intervals. If one continues to be attentive, it seems that one of the elements of the group, the first in general, also appears to be more intense than the others.

These introspective notions were confirmed by authors who asked their subjects to accompany each of the sounds by a tap (Miyake, 1902; Miner, 1903; MacDougall, 1903; Temperley, 1963). The recording of the taps corresponded to introspective observation. There were temporal differentiations and corresponding accentuations. This phenomenon, today seemingly banal, which preceded the work of the Gestalts, was considered extraordinary. Its significance remains important at the present time since it underscores the perceptual and spontaneous character of rhythmic grouping.

In his pioneering study, Bolton (1894) worked on the problem of the limits of the frequencies at which subjective rhythmization could appear. He gave as the lower limit an interval of 115 msec, and as the upper limit, 1580 msec. These limits should command our attention, since they are approximately those of the durations on which all of our perceptions of rhythm are based. The lower limit (about 120 msec) corresponds to the psychophysiological conditions of the distinction between two successive stimuli. It corresponds to the minimal interval between two rapid motor taps. The upper limit has a very important perceptual significance, revealed at the phenomenological level. It corresponds to the value at which two stimuli (or two groups of stimuli) are no longer perceptually linked.

In order to understand this, let us take the example of the tick-tock of a clock. The sounds are linked together in groups of two. Let us suppose that one can slow down this tick-tock indefinitely. There comes a moment when the tick and the tock are no longer linked perceptually. They appear as independent events. This upper limit is also that where all melody disappears, and is substituted by isolated notes. The limit proposed by Bolton (1580 msec) is without doubt too precise. MacDougall rightly situated it between 1500 and 2000 msec. We propose retaining a value of about 1800 msec. Beyond this duration subjective rhythmization becomes impossible. If one asks the subjects to continue isochronous sequences by tapping, one finds (Fraisse, 1956, pp. 13–14) that the variability of these sequences is minimal around an interval of 600 msec, that it increases a little for shorter intervals, and increases substantially beyond 1200 msec. The linkage is no longer susceptible to precise perceptual control.

On the motor level, we also find an optimum of about 600 msec for perceptual organization. This length is also that which is perceived with the greatest precision (Fraisse, 1963, p. 119; Michon, 1964).

The importance of all of these parameters will appear when we discuss more complex rhythms.

## III. RHYTHMIC FORMS

### A. Regular Groupings

As soon as a difference is introduced into an isochronous sequence of elements, this difference produces a grouping of the elements included between two repetitions of the difference. One then speaks of objective rhythmization. This difference can be a lengthening of a sound, an increase in its intensity, a change in pitch or in timbre, or simply a lengthening of an interval between two elements. This fact suggests two types of question: (a) the possible durations of rhythmic groups and (b) the nature of the effects principally produced by modifications in intensity or in duration.

#### *1. The Duration of Groups*

If one accentuates or lengthens one sound out of two, three, or four, this produces the perception of a repetitive group of two, three, or four elements. We already know that the interval between the elements is important and that the perception of the rhythm, objective or subjective, disappears if the intervals are either too short or too long. By asking subjects to produce groups of three or four taps, we found that there was, on the average, an interval of 420 msec between the taps of groups of three, and of 370 msec between groups of four (Fraisse, 1956, p. 15).

Within these limits of succession one can perceive groups of from two to six sounds that correspond to the boundaries of our immediate memory or of our capacity of apprehension. In order to obtain a good rhythmization, it is necessary when the number of elements increases to increase the frequency of the successive sounds. MacDougall (1903), while employing a method of production, found that the longer the groups were the faster the frequencies of the sounds. Thus, a group of four is only 1.8 times as long as a group of two, while a group of six is 2.2 times as long. Everything happens as if the subject was trying to strengthen the unity of the group when the number of elements to be perceived is larger.

By employing the method of reproduction of auditory series (while preventing the subject from counting by a concomitant verbalization), we found that for an interval of 17 msec, 5.7 elements were accurately perceived (total duration 800 msec); for an interval of 630 msec, 5.4 elements (total duration 2770 msec); and for an interval of 1800 msec, 3.3 elements (total duration 4140 msec). Thus, there is an interaction between the number of elements and their frequency. The total length of possible groupings depends on both. However, more complex groupings of sounds can be perceived (such as those that we will study in Section III), if subunits analogous to those that are called "chunks" can be created. Thus, one can come to perceive about 25 sounds as a unity (Dietze, 1885; Fraisse and Fraisse, 1937) if they form five subgroups of five sounds following each other at a rapid frequency (180 msec). However, the total length of the groups, in this extreme case, cannot be more than 5 sec.

This limit is found in the rhythmic arts. The slowest adagio in a 9/4 bar is no longer than 5 sec, and the longest lines of poetry have from 13 to 17 syllables, the time necessary to recite them being no longer than from 4 to 5 sec (Wallin, 1901). This length of from 4 to 5 sec is, however, an extreme limit that allows only unstable groupings. For the groups of sounds produced by subjects, MacDougall (1903) gives 3 sec as a practical limit. According to Sears (1902), the average length of a musical bar in religious hymns is 3.4 sec. According to Wallin (1901), the average duration of lines of poetry is 2.7 sec.

This duration limit corresponds to what has been called the *psychological present*. We know that we can perceive, relatively simultaneously, a series of successive events (for example, a telephone number or the elements of a sentence). This phenomenon is also called *short-term storage* or even *precategorical acoustic storage* (Crowder & Morton, 1969). We prefer, however, in the case of rhythm to speak of the psychological present. This term expresses well the organization of a sequence of events into a perceptual unity. It corresponds to our limit in organizing a succession. A similar unity introduces a perceptual discontinuity in the physical continuum into the psychological present. One should not repeat the mistake that James (1891) made when he thought that there was a continuous sliding of the present into the past. He cited as an example the recitation of the alphabet. If one's present is at moment $t$: C D E F G, at moment $t + 1$ it will be D E F G H, C having disappeared and having been replaced by H. This analysis is inexact. Language, as well as rhythm, shows that one group of stimuli succeeds another group.

Today, it is easier to accept as true the principle of the temporal *Gestalt*. At the beginning of the twentieth century, psychologists were very much preoccupied with associative links as the basis of the unification of rhythmic groups. Two hypotheses were dominant about 1900. One, introduced for the first time by Bolton (1894), referred to the pulse acts of attention; these acts follow each other. This is only a description of the same sort as that of the psychological present, although perhaps, it is less subtle. The other hypothesis looked for the unification of successive sensations through the unity of a kinesthetic response. Miner (1903) wrote: "In feeling the groups to be units is an illusion due to the presence of movement or strain sensations along with the sensations that are grouped." In this form this hypothesis goes beyond associationism. However, it is true that the perception of successive groups is facilitated if there is an accompaniment of sounds by synchronized movements that can, according to their frequency, correspond either to each of the sounds or to a single sound per group. This fact is observed in objective and subjective rhythmization. Temperley (1963) observed a negative correlation between the frequency of motor responses and the frequency of a cadence, the subject tapping only once for every two or three sounds. Miner (1903) had previously observed the same phenomenon when a person taps his or her foot to a piece of music. Movement does not create the grouping; however, the more the rhythmic experience is sensory-motor, the more it is complete (see Section III).

## 2. *Factors in Grouping*

Any differentiation in an isochronous series of identical elements serves as a basis for grouping. However, they do not all have the same effects as far as the organization of a temporal series is concerned. In general, a noteworthy lengthening of the duration of a sound or of the interval between two sounds determines the end of a group; this longer duration allows one to distinguish between two successive patterns. It imposes itself in subjective rhythmization. This lengthening creates a rupture between two groups. We call it *pause*. Its duration is not random and cannot be assimilated to a gap or to a ground according to Rubin's terminology. In effect, the perception of rhythm is not only that of a grouping but also that of a linking of groups called *Gestaltverkettung* by Werner (1919) and *Fugengestalten* by Sander (1928). In his 1909 dissertation, Koffka noted the following striking behavior: If a subject was given three lights—a, b, c—and if he was asked to continue the rhythm, he not only reproduced the intervals between a and b and c but also linked the groups of lights as though the interval between the final c of the first group and the initial a of a second one had been proposed to him. Out of seven subjects five did not even see that there was a problem.

When asked to tap regularly in groups of three or of four, subjects spontaneously separated the groups by a pause that was from about 600 to 700 msec. With more complex patterns the duration of the pause was at least equal to that of the duration of the longest interval inside the pattern. Otherwise, there was a reorganization of the pattern, so that the longest interval played the role of the pause. However, the pause was never longer than 1800 msec; since, if such were the case, there would no longer be the perception of a chain of patterns but only the perception of isolated patterns. Wallin (1901) found that the pause at the end of a line of poetry was, on the average, 680 msec. Evidently, pauses in the strict sense of the word do not exist in a musical sequence; still, one exists at the end of each pattern in the form of a slight interval.

Most often, it has been assumed that the structuring of patterns was based on the accentuation of one or of several elements. This accentuation already appears, as we have seen, in subjective rhythmization. It is important in music where the pause—*stricto sensu*—does not play a role. The accented element, when it determines the length of a group, also determines the nature of the grouping. The objective accent is situated most spontaneously at the beginning of the pattern. This fact has already appeared in subjective rhythmization.

A regular succession of a strong and of a weak sound of equal duration is perceived in 60% of the cases as a succession of trochees (strong-weak) and in 40% of the cases as a series of iambuses (weak-strong) (Fraisse, 1956, p. 95). Other discussions continued—and are continuing—concerning the relative role of accents and of pauses. In reality, there is an interaction between the two factors producing segregation. An important lengthening of a sound leads it to play the role of a pause. A slight lengthening of the duration of a sound makes it appear more intense and confers upon it the role of an accent. It then, most often, becomes the first element of a pattern. Recip-

rocally, the accenting of an element slightly modifies its duration or, if one prefers, the interval that follows it. Thus, while synchronizing the taps with a regular series of strong and weak sounds, we found, as did initially Miyake (1902), that the intervals between the taps depended on the perceived structure (Fraisse, 1956, pp. 95–96).

| | |
|---|---|
| trochee | 484–452 msec |
| iambus | 432–520 msec |

A general fact is observed: the most intense element is lengthened. But it is lengthened more when it terminates the pattern, as the effect of the properly so-called accent adds itself to the effect of the pause.

The most intense sound is spontaneously lengthened even by musicians (Stetson, 1905; Vos, 1973). In prosody when the structure is fundamentally in breves and longs, the accent is always placed on the long element. There is between the lengthening of the duration and the accent, a certain functional and perceptual equivalence. If the more intense sounds are perceived as being longer, the longer sounds are perceived as being more intense.

Evidently, by modifying the intervals between sounds of different intensities, one can make the trochees or the iambuses, for example, more frequent. This experiment was performed by Woodrow (1909) who was able to establish a point of indifference where the iambus had as many chances of being perceived as the trochee, the dactyl as the anapaest. It suffices to lengthen relatively the interval that follows the weak sound or that precedes the strong sound. Since grouping can be obtained by modifying durations or accents, and since there is an interaction between these two factors, simplistic conclusions can be discounted such as those which affirm that the perception of rhythm is based only on the perception of durations or of accents. But the roles of durations and of accents are not the same. The duration of the elements or of the intervals which separate them (in the rhythmic arts these are barely distinguishable) is always a precise quantity. Experiments show that in performances durations are less variable than are accents. By making subjects tap repetitive forms, Brown (1911) found that the relative variability of accents was of the order of 10 to 12% and those of durations of 3 to 5%. This result was confirmed by Schmidt (1939). One finds the same results in vocal performances.

One can then modify in the repetition of patterns the strength—but not the place—of some accents without modifying the nature of the perceived rhythm. One can also modify the duration of the elements, but to a lesser degree. Variations of about 6% do not in any way alter the nature of the rhythm. They are still acceptable at 12% but not beyond (Wallin, 1911). Modifications of accents can be very much more important, and artists use them a great deal.

We have reasoned until now as though accentuation signified an increase in intensity. However accents, as we said, can be produced by a slight increase in duration. They can also be obtained by changes in pitch or in timbre. A change in pitch brings with it rhythmic segregation. However, according to the best study (Woodrow, 1911), the highest sound can be spontaneously placed at the beginning as well as at the

end of a group. There is also an interaction of some sort between the intensities and the pitches of sounds. In a series of sounds the highest appears subjectively as the most intense, and vice versa.

If differences in duration, in intensity, and in pitch can organize rhythmic groups, the intensity of the accent has a specific role that we have already noticed in subjective rhythmization. The periodic repetition of accents more or less induces motor reactions that repeat themselves regularly and reinforce the salience of the perceived patterns.

We have described the modifications, which in a series of isochronous sounds, produce rhythmic structures (Fraisse, 1975). We have not explained them. We can state that these perceptual laws are identical to those pointed out by the Gestaltists, and in particular by Wertheimer (1922, 1923). The pause underscores the importance of proximity, the accent, that of repetition of identical elements, or of good continuation.

This comparison still does not explain anything, as Gestalt laws are themselves unexplained. However, current research on the perceptions of the new-born have at least shown the great precocity of these perceptual laws in the spatial domain (Vurpillot, Ruel, & Castrec, 1977). It has recently been demonstrated that the very young child is also sensitive to differences in rhythms. Demany, McKenzie, & Vurpillot (1977), by using a habituation paradigm with an operant response consisting of the fixation of a visual target, showed that the new-born child (71 ± 12 days) discriminates a series of isochronous sounds (duration 40 msec with intervals of 194 msec) from a series of patterns of four sounds separated by intervals of 194, 97, 194, 297 msec. They can also discriminate a pattern of the type 97, 291, 582 msec from another pattern 291, 97, 582 msec. These sequences are here described in an arbitrary manner, as we remain ignorant as to how the child groups sounds.

However, a child of this age does not perceive a difference in tempo between a series composed of sounds of 500 msec followed by intervals of 500 msec and another series composed of sounds of 1000 msec followed by intervals of 1000 msec (Clifton & Meyers, 1969). Does the technique of this experiment have any flaws? Were the chosen tempos too slow? One does not know at present, but Berg (1974) and then Leavitt, Brown, Morse, & Graham (1976) found that a simple change of tempo between two simple structures (séries of sounds of 400 msec followed by an interval of 600 msec compared with a series of sounds of 800 msec followed by an interval of 1200 msec) was discriminated. According to Chang and Trehib (1977), children at 5 months are capable of discriminating groups of two from groups of four sounds (children of this age are also capable of discriminating between identical groups composed of different sounds). Allen, Walker, Symonds, and Marcell (1977) also found that children can at 7 months distinguish an isochronous succession from an iambic type of grouping.

Rhythmic grouping thus appears very early in life. In consequence, hypotheses that consider it as a voluntary activity, such as the pulse of attention or a motor accompaniment, are invalid. Furthermore, the law of proximity seems to be very primitive in



time, as it is in space [Brunswick (1956) to the contrary]. Beyond early childhood the child becomes capable of perceiving or of reproducing more and more complex forms and, in particular, rhythmic structures that are not repetitive (Fraisse *et al.*, 1949; Zenatti, 1976).

In order to summarize the effects of pause and of accent, we can say

1. Any noteworthy lengthening of a sound or of an interval between sounds plays the role of a pause between two successive groups.
2. Any sound, qualitatively different from the others, especially as to its intensity, plays the role of an accent that begins the group. When these two principles act simultaneously, one can as well say that the pause preceded the accent (Vos, 1977) as the converse.
3. More intense sounds are perceived as relatively lengthened and longer sounds as relatively more intense.

### B. Patterns in Time

We will now attempt to understand the laws underlying the organization of groups when several elements are different from each other. In other words, we are going to study patterns in time (Handel, 1974). One preliminary remark is important. In a complete series of sounds the first perceived pattern tends to impose its structure on the later patterns. It becomes a privileged form of grouping (Preusser, Garner, & Gottwald, 1970a), and this fact confirms the importance of predictability as the basis of rhythmic perception. In order to avoid this effect, it is necessary to use artifices so that no pattern imposes itself due to its initial position. One can, for example, increase the intensity or the frequency of the sounds little by little, use long presentations in order to allow reorganizations, present random series before ordered ones, etc.

1. What happens if in a potential pattern there are several intense sounds? We mixed loud sounds (L) of 100 dB with softer sounds (S) of 75 dB (Fraisse & Oléron, 1954) in patterns of four or five sounds. The sounds were brief and between them the intervals were equal (475 msec for four sounds, 380 msec for five sounds). The subjects had to listen to a long series and then consecutively reproduce three times the perceived patterns by tapping on a key which thereby enabled the measurement of the force of the tap. The subjects grouped, as much as possible, sounds of the same intensity, all of which resulted in the construction of runs. Thus, beginning with a sequence L S S L L S S L S S, etc., one does not perceive the pattern L S S L but the pattern L L S S and, less often, S S L L. The pattern includes the smallest possible number of runs and is at the same time the simplest.
2. What happens if the number of loud sounds is greater than that of soft sounds? The subjects in their reproduction can invert the relative force of the elements and reproduce, for example, the series L L L S in the form S S S L or L S S S. These inversions are reminiscent of those which one can perceive in figure-ground reversals of spatial forms and allow us to think that, in these cases, the differentiating element is the least frequent one.

We have found the same phenomena by using sounds of different pitches (for example, sounds of 1040 Hz in combination with sounds of 760 or 520 Hz) (Ehrlich, Oléron, & Fraisse, 1956). The subjects tend to regroup so that the sharpest sounds tend to begin the group. With the same technique of reproduction, we have also found that the least numerous elements, high or low, were accented.

This type of research was developed by Garner and his colleagues using longer series. The first study was by Royer and Garner (1966). The patterns had eight sounds of two different types (two buzzers) and were repeated until the subject was capable of finding a pattern and of tapping it on two different keys. The authors of this research were, above all, concerned with estimating the effect of response uncertainty, evaluated in bits, on the identification of a pattern in a series. The first observation they made was that the subjects did not proceed by trial and error; they only began to respond when they had identified a pattern, and at that moment, the pattern would be responded to in complete synchrony and with little difficulty. Thus, construction does not proceed element by element, but wholistically. As for the rest, their hypothesis was partially confirmed. The simple patterns were organized quickly and the complex ones more slowly, which confirms Garner's (1962) thesis according to which perceptually good patterns should have few alternatives. This research showed that the most often chosen organizations were those in which the number of changes was minimal—that is, where the sounds of the same quality grouped themselves to the maximum extent. Thus, the pattern X X X O X X O O was the most often perceived (61 times out of 128), and from the same elements, the pattern next most often perceived was O X X O O X X X (31 times out of 128); however, the pattern X O X X O O X X was practically never perceived (once out of 128). As one also notices by this example, the longest run tended to begin the pattern. It could sometimes end the pattern, but it was practically never in the middle.

Later research (Royer & Garner, 1970; Preusser, Garner, & Gottwald, 1970a), most often using sounds presenting little difference in pitch, have confirmed this result. The longest run was placed at the beginning or, more frequently, at the end of the pattern. The solution evidently depended on the structure of the whole and on the relative length of the longest run. Thus, in the example mentioned above, the longest run more often began the pattern than finished it; another pattern X X X X O O I O was perceived only in 36 cases out of 128 whereas the pattern O O I O X X X X was perceived 61 times out of 128.

The place and the role of the run were the main determinants of grouping. Others could also play a role, all of which corresponded to making simplifications prevail. Thus, when possible, the subject chose a directional simplicity with run lengths either increasing or decreasing in regular order (Preusser, Garner, & Gottwald, 1970b). The most redundant and/or symmetric forms (for example, X O X O, or still, X X O O X O O O—where the first, third, and fifth elements were conspicuous) were more easily perceived than the pattern X X O X O X O O that did not have a simple structure nor a longer run than the others (Sturger & Martin, 1974).

Preusser *et al.* (1970b) and also Handel (1974) have analyzed these results in terms of figure-ground relations by claiming that one of the elements plays the role of figure,

and the other that of ground. In particular, they rely on the fact that if one of the elements is replaced by an empty interval, the laws of organization are the same. In the case in which the longest run is at the beginning of a pattern, it plays the role of figure and the pattern obeys the *run principle*. When it is at the end, it plays the role of ground, and the authors then speak of the *gap principle*.

The distinction between figure and ground, however, does not appear to be relevant. In a rhythmic structure there is no ground. Even the empty intrapattern intervals are part of the structure. As for the interpattern intervals, even though they have a different status, they nevertheless form links between the successive subpatterns. In the types of structures used in this research, there is no pause, *stricto sensu*, between the patterns. We also think that when the longest run is at the end of a pattern, it plays the role of accent more than that of a gap. Reference to poetry or to music, moreover, helps us to understand that all of the elements that structure a succession play a role. There is, between all of them, a relation that is not that of all-or-none (figure-ground) but of a hierarchy of salience.

What is the influence of tempo on patterns in time? All of the presented results were obtained using frequencies of two to three per second. What is the result when one increases or decreases this frequency? A frequency of two to three per second appears as an optimum. For more rapid frequencies (eight per second), more time is needed in order to discover the pattern in the presented sequence. However, the structuring phenomenon still occurs as an "integrated, immediate, compelling, and passive" process. In contrast, at the lowest frequency, .8 (which still is not very low), the subject constructs the pattern that is learned little by little according to an "integrated, derived, intellectualized and active" process (Garner & Gottwald, 1968).

Garner and Gottwald have also found that, at the lowest frequency, the structuring of patterns was all the more difficult if they deviated more from patterns constructed according to the run principle. Preusser (1972) systematically stated the problem of the interaction between the frequency of the elements and the structuring of the patterns. With two sounds of 238 and 275 Hz, the frequency being rapid (four per second), the subjects tended to place the longer run at the end of the pattern, making it play the role of a gap according to Preusser, of an accent plus pause according to us. At the slowest frequency (one per second), the longest run tended to begin the pattern. Why, at the most rapid frequency, was the run at the end? This solution seems to be characteristic of *perceived* rhythm whereas the initial run would be more characteristic of *constructed* rhythm, if we use Garner and Gottwald's distinction. Moreover, Preusser used two criteria in order to detect a pattern. One was to reproduce the pattern on one or two keys. The other consisted of asking the subject to describe the perceived rhythm by means of symbols. The delay necessary to describe the pattern is at least twice that which is necessary to reproduce it. This fact, previously found by Oléron (1959), confirms the wholistic character of rhythmic perception and also the compatibility between perceived patterns and motor patterns. In order to describe rhythms, it is necessary to analyze their structure, but this analysis is not necessary in order to reproduce them.

## C. Patterns of Time

Rhythm, understood as "order of movement" is evidently based on an order which is primarily temporal. Until now we have envisaged only the most simple temporal situation: the isochronous repetition of sounds. What are the more complex temporal situations that permit perception of rhythm, and following Handel's expression (1974), what are the characteristic patterns of time?

### *1. Rhythm and Arrhythmia*

If rhythm is order, arrhythmia is disorder (i.e., it is *a priori*, a sequence of continuous sounds where no temporal organization is perceptible). A computer can create this type of sequence. Can man? We asked subjects to produce an uninterrupted series of taps as irregularly as possible. We also asked them, in contrast, to produce patterns of five or six sounds having an internal structure of their choice, while trying to avoid reproducing known tunes (Fraisse, 1946–1947). While subjects found the task of producing a series of patterns easy, they nevertheless found it difficult to produce an irregular sequence. In order to study temporal structure, we have calculated the successive ratios between durations by computing the ratio of the shorter of the two intervals to the longer.

The first characteristic fact, in rhythm as well as in arrhythmia, is that a ratio of near-equality between two successive intervals predominates (40% of the ratios are less than 1.2). It is as though every sequence were based on a tendency to produce an interval equal to the preceding one, which is evidently the easiest and the most economical activity.

Rhythmic and arrhythmic sequences are constructed on the basis of this regularity. However, the way of breaking regularity is different in the two cases. In arrhythmia, the higher the ratio the less frequent it is. The rupture with equality then happens by a lengthening (or by a decrease) of the preceding interval: small differences become numerous, large ones become rare. In rhythm, on the contrary, small differences are rare. When the subject has broken the regularity, he or she produces a new interval of a noticeable duration. The difference forms about a ratio of one to two.

If one considers the absolute durations of the intervals, one finds the following results:

| *Intervals less than (msec)* | *Rhythm (%)* | *Arrhythmia (%)* |
|---|---|---|
| 400 | 56.2 | 35.2 |
| 1000 | 92 | 75.8 |
| 1800 | 98 | 93.8 |

Biblioteca del Real Conservatorio Superior de Música

First, these numbers indicate that in order to *perceive* regularities or irregularities, we use few intervals larger than 1800 msec. These would break the succession of sounds into independent sequences. It is also necessary to note the high proportion of short intervals in the rhythmic patterns. Moreover, ratios of the order of one to two inter-

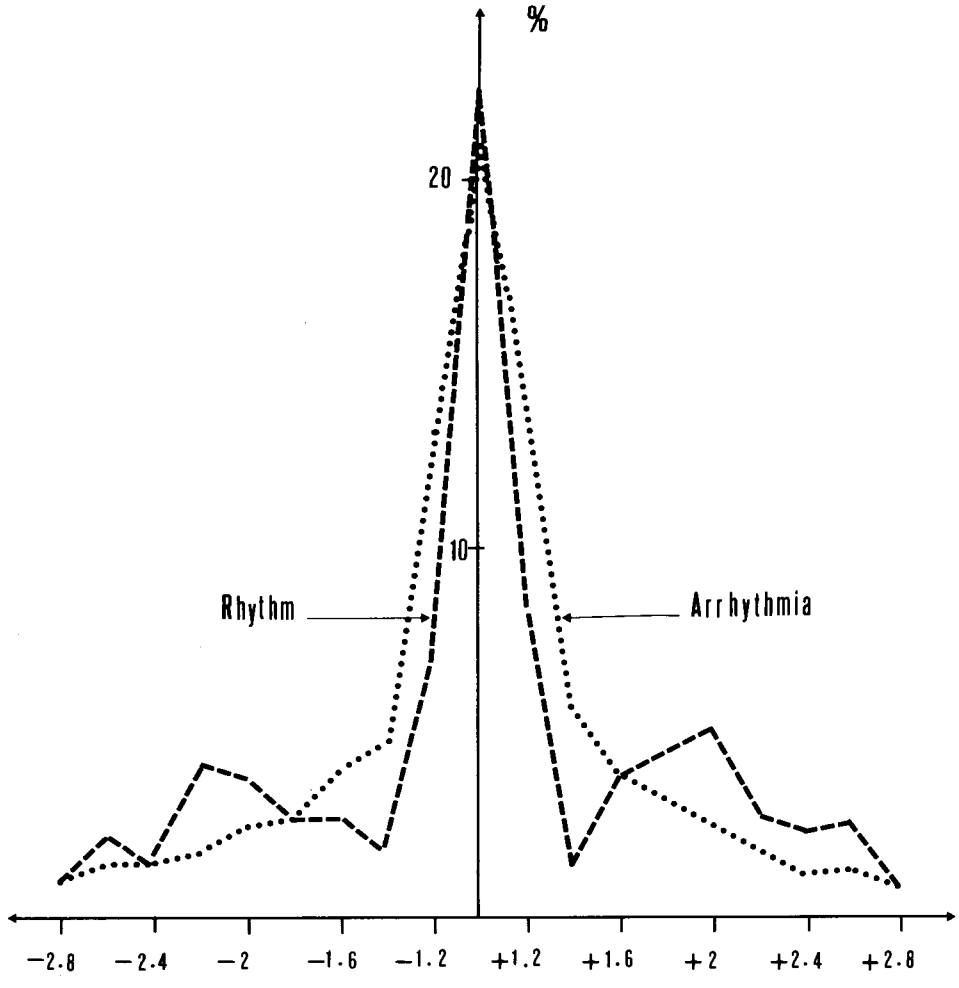



**Fig. 1.** Frequency of the ratios between successive intervals for rhythmic and arrhythmic sequences. 1 indicates equality of the intervals. Negative values indicate that the second interval is shorter than the first; positive values indicate that the second interval is longer than the first. Class interval equals .2 (i.e. class 1.2 includes ratios between −1.09 and +1.09). Only ratios inferior to 2.9 are represented here; they correspond to 85% of the ratios with rhythmic sequences and 86.19% with arrhythmic sequences (from Fraisse, 1946-47, *47-48*, 11-21 by courtesy of *Année Psychologique*).

vene most often only when the time that we call short (less than 400 msec) follows or precedes the time that we call long.[2]

A more complete analysis reveals that the relative equalization of durations in rhythmic patterns is not only produced between adjacent intervals. One thus finds patterns 680-260-630-280 (in msec) and patterns 280-300-850-290-850 (in msec) with

[2]Here the word "time" is used as synonymous with duration or with interval until Section IV, where we use time according to common usage.

equalization of the short times on the one hand and of the long times on the other hand. This phenomenon is found in patterns of three or four taps. Here are some examples (average of 10 subjects): 210-480-490 msec; 470-190-430 msec; etc. (Fraisse, 1956). The hypothesis that we formulated above—that is, of a simple tendency to repeat equal intervals—is only partially exact since we find the phenomenon of equalization between nonadjacent intervals. Briefly, patterns are characterized by a composition of basically two sorts and only two sorts of time: *short times* of 200 to 300 msec and *long times* of 450 to 900 msec. If one looks not only at the averages but at the individual performances, one finds that between short times, adjacent or not, 84% of the ratios are less than 1.15 and 97% are less than 1.55. Between long times, 54% of the ratios are less than 1.15 and 94% are less than 1.55. The modal value of the ratios' long times—short times is 2.4 of which 95% are less than 1.55. This ratio 1.55 seems to be the dividing point between two sorts of time.

If two durations belong to the same category, there is a tendency to equalize these durations. We prefer to say that there is assimilation since this equalization is not absolute. Among durations of differing categories, there is a sharp distinction. Assimilation and distinction bring us back to the classical perceptual laws which correspond to a principle of economy in perceptual organization (Fraisse, 1947).

## 2. *Temporal Rhythms as Structure*

Our previous analyses already confirm that temporal intervals in rhythmic structures are interdependent. However, one can go further and show that the basic pattern described previously corresponds to "good form." Are we capable of producing or of reproducing any other patterns? One can demonstrate, in several ways, the salience of good form.

First, by a conflict between space and time: if one lays out before the subject 4, 5, or 6 targets at different distances while asking him to tap them successively as quickly as possible without stopping, he establishes a veritable rhythmic pattern of taps. This temporal pattern is simpler than the spatial pattern. Unequal spaces are gone through either in equal times or in very distinct times (ratios of two to three) (Montpellier, 1935). When the subject is auditorily presented with temporal patterns that differ from a basic temporal model and asked to reproduce these patterns, one observes deformations of the model by a systematic reorganization that simplifies it. If there are, for example, in the model three unequal intervals, the subject tends toward a reproduction of two intervals. Let us take the very simple case of a structure with a first interval of 190 msec, a third interval of 450 msec, and a second interval which has, depending on the model, a duration of 210 to 690 msec. At first there is assimilation of the first to the second duration and then there is assimilation of the second to the third duration.

Good form then, is not only a spontaneous form but a dynamic organization that imposes itself in production or reproduction.

We can find, in very different contexts, examples of this type of structure based on ratios of two durations only. When Samuel Morse attempted to create an alphabet

based on the play of durations, the Morse code, composed of two durations called dots and dashes, was invented. Greco-Latin prosody was based on the opposition of two durations: breves and longs. In music, there is, at any given moment, a play of two notes that are in a ratio of one to two or one to three (double quaver and quaver, quaver and crotchet, quaver and pointed crotchet). These two notes represent 85 to 95% of the movement (Fraisse, 1956, p. 107). The first theorist of rhythm, Aristoxenus of Tarentum, distinguished two sorts of beats corresponding to the upbeat and the fall. One was the first beat upon which only one syllable or one note could fall; the other was worth two or three first beats. Aristoxenus claimed that only ratios corresponding to whole numbers are rational.

This generality, regarding the use of only two durations, corresponds, according to us, to a perceptual requirement revealed by psychophysics. Research done on information theory in order to measure channel capacity has shown that the channel is always limited by our ability to distinguish in an absolute way, several levels of stimulation. This capacity, which is about five, varies with the nature of the sensation. In the case of duration, the studies by Hawkes (1961), Murphy (1966), and Bovet (1974) have shown that even trained subjects could differentiate only two or, at the most, three durations in the range of perceived durations (below 2 sec). If the durations were more numerous, confusion arose.

However, these laws do not apply to the time interval between two patterns that we have called pause. Phenomenally, a pattern ends with the last element. But between one pattern and the next there is, as was revealed by subjective rhythmization, a pause that corresponds to the length of the last note in the case of music and that is an empty time in the case of taps.

Let us take one more step in the analysis of temporal patterns. When they are quite long, they often split up into several subunits. A pattern of six sound-taps is often decomposed into two subunits of 3 + 3, of 4 + 2, or of 2 + 2 + 2 as the case may be. In this case, the interval between two subunits has the characteristic of a pause: it is at least equal to the longest duration but it is not necessarily equal to it, while being more integrated with the pattern than with the pause, *stricto sensu*, between two patterns. This type of analysis explains, we think, certain groupings that intervene when models have eight or ten sounds, as in research such as Garner's

If a subject taps a pattern at his spontaneous tempo and if he is asked to continue to tap the same pattern more quickly or more slowly, it is seen that the ratio long time—short time is maximal at the spontaneous tempo. When the tempo slows down a great deal, there is no longer a sharp distinction between long time and short time. At the limit the durations are almost equal. We have seen rhythm born from a rupture with regular movement; we see it disappear by a return to this movement.

The previous analyses were based on methods of production and of reproduction of fairly short patterns. Preusser (1972) has produced new data. He not only had patterns reproduced with two types of elements (see Section III,B), but he also constructed similar patterns that presented only one type of element by replacing the other by an empty temporal interval. Two organizational principles were obvious

from this work (1) the run principle: the longest run begins the pattern (for example, 3″1′)[3] and (2) the gap principle: the longest interval terminates the pattern (for example, 1′3″).

If these two principles are compatible, as in 3′1″, the pattern is correctly identified in 90% of the cases. If they are incompatible, as in the first two examples cited the gap principle is, on the average, the decisive factor in 68% of the cases and the run principle in 32% of the cases. When there are three runs, which we consider as three subunits, a third principle, which we have already detected in spontaneous rhythms, is added to the two earlier principles: The sequence of run lengths produces an upward progression (for example, 1′2′3′). This principle evidently gives rise to an organization differing from one starting with the longest run. Preusser has compared these results with those found in the case of sequences of eight elements composed of two sounds of a different nature. By comparing the results of the two studies he concludes that the gap principle plays a more important role when there is only one element. This confirms our previous conclusions. Of two elements, one is not figure and the other ground since the empty intervals have a stronger structuring effect than the element considered as the ground in their analyses.

Handel's research (1974) brought along a supplementary piece of information. He had the duration of sounds varied (ratio of 1 to 5), and he found that, most often, the short durations began the pattern and that the long durations ended it. This effect is all the more marked when the run of the short and/or of the long durations is longer.

More recent research by Vos (1977) produced comparable results obtained by another method. The subjects had to judge, for a sequence of two durations which were in a ratio of 1 to 4, whether it was an iambus or a trochee; for a series of three durations, whether it was a dactyl, an anapaest, or an amphibrach. He used three principles in order to explain the obtained results: (1) Tones that are separated by short intervals are perceptually grouped together; (2) the first tone of a perceptual group is a tone that is immediately preceded by a long interval (which is another example of the role of the pause, or of the gap principle); and (3) long tones are perceived as accented and short tones as nonaccented.

All authors are in agreement that complexity is important among the factors which intervene to produce greater or lesser salience of rhythmic patterns. This is difficult to evaluate. It seems that one can draw several conclusions from research done on Morse code signals: the relative difficulty in learning each signal can be considered to be an index of its complexity. If using Plotkin's (1943) results, we divide the signals into three categories of 12—easy, average, and difficult to learn—we can calculate three indices, keeping in mind that the number of dots and dashes varies from three to five elements: (a) the number of elements in a signal (N), (b) the number of signals having only one category of elements, dots or dashes (E), and (c) the number of signals in which there is an interleaving of elements (for example: - - . . -) in contrast to those in

[3]The digits represent the number of elements which follow each other, the hyphens, the intervals which are equal in duration to the elements.

which there are only two runs (for example: - - -..) (R). One finds the following results:

| | N | E | R |
|---|---|---|---|
| Easy signals | 2.9 | 8 | 0 |
| Average signals | 4.2 | 1 | 3 |
| Difficult signals | 3.9 | 0 | 6 |

Complexity increases a little with the length of the signals and, above all, with the multiplication of runs, as was stated by Preusser (1970; Preusser *et al.*, 1972). Signals with only one type of element are always easy.

Generally, one can say that, the more a temporal form is brief and simple, the easier it is to perceive. Vos (1973) attempted to calculate the indices of complexity by taking into account the indices mentioned above as well as the ratio between the length of subunits.

One can, moreover, more closely approach music by studying how syncopated auditory rhythms are perceived. Polyrhythms are defined as the simultaneous presentation of two pulse trains such that the rates are not integral multiples of each other (for example, three against four). Each pulse train is a series of regularly recurrent stimuli (Oshinsky & Handel, 1978). How will this ambiguous pattern be perceived, the criterion being the choice made by the subject asked to tap in synchrony with the pattern in question? Will he follow the pattern of three or that of four elements? The most remarkable result was that the subjects most often preferred to accompany the pattern of three rather than that of four elements but that this tendency was not the same for all tempos. In this research the pattern had a duration that varied from .96 to 2.4 sec. There was a reversal of the tendency for durations of 1.2 or of 1.6 sec depending on the pitch of the sounds. These tendencies were about the same when the two patterns of three and of four consisted of identical sounds or when they consisted of sounds of differing pitch.

Is it also necessary to underscore the fact that synchronizations are very rapidly established? The subjects began to tap in a stable way after about 3 sec, which proves that the two trains of stimuli were not analyzed. The majority of subjects, moreover, did not detect that there was an ambiguity in the polyrhythms.

## IV. THE PERCEPTION OF MUSICAL RHYTHMS

The above analyses have permitted us to extract the laws characteristic of rhythm perception. However, the stimuli used were far from musical, since these researches used only taps, identical sounds, or at best, two types of sound of different duration, intensity, or pitch. Musical rules, however, do not escape the fundamental laws that we have demonstrated. Without doubt, these laws do not explain music any more than gravity explains the art of architecture. But there is not an architect who ignores gravity any more than there is a musical rhythm that does not respect perceptual laws.

First, it is necessary to underscore the difficult problem of vocabulary that we have

evoked by distinguishing rhythm, which is the perception of a pattern, and meter, which allows the description of a musical composition. We will use this distinction as we consider the perceived rhythm and meter used by the composer.

A musical composition is a synthesis of very different stimuli that are perceptually unified much as forms and colors are unified in a painting. We distinguish melody, harmony, timbre, and a rhythmic organization consisting of the succession of rhythmic patterns, at the same time identical to themselves and also varying continuously. The unity assures the characteristic of anticipation, which seems to us to be fundamental, and that Steedman (1977) finds, for example, when he tries to discover in a fugue by Bach the algorithms that allow one to give an account, if not of the rhythm, at least of its meter. What appears fundamental to him is the "principle of consistency" that corresponds to the fact that there is, with the passage of time, a constancy of predictable forms from the first bars on.

These patterns are composed of subunits that metrically correspond to times and, in performance and in perception, to a succession of beats. The metrics tell us that there are bars at two, three, four, and even nine times, but perceptually the bars at four times are often reduced to binary rhythm and the others to combinations of substructures. The longest bars have hardly more than nine times and are generally understood as a triple ternary rhythm. Reciprocally, the simplest bars can group themselves into periods as do the lines of poetry into stanzas. A famous example is that of the scherzo of Beethoven's Ninth Symphony, written in 3/4 bars, in which Beethoven indicated *ritmo a tre battute* in order to indicate that it is necessary to regroup three bars into one rhythmic unity.

Musical rhythm is based on a precise temporal organization of bars and the play of accents. Let us summarize two points. Rhythmic pattern has a duration of from 2 to 5 sec. In recent research Vos (1976) measured the duration of different bars in the *Well-Tempered Clavier* by Bach using a recorded performance. Their duration was 1.75 sec for a 2/4 bar, 3.0 sec for a 3/4 bar, and 4.8 sec for a 6/4 bar. These are only examples that permit us to say that the durations of the longest bars correspond to the possibilities of unifying successive elements into a pattern.

The bars repeated themselves and were metrically identical in their structure but varied in particular by syncopations. Sears (1902) found a variability in the duration of 3% between successive patterns of religious chants. Weaver (1939) studied the variability of a piece of music where there was a succession of normal bars and syncopated bars. The subjects, conservatory students, tapped rhythms with their right hand while they kept the tempo by moving a foot. The presence of syncopated bars does not prevent the very great temporal regularity of the executed rhythms (variability from 2% to 4% between successive patterns). The bars are composed of notes whose duration is specified by the composer who indicates the tempo. We have already mentioned the important statistical result that we found by studying the notes used in pieces having a metric unity.

Table I gives the proportion of each note in each piece studied. It is immediately evident that the compositions are based on two notes that represent more than 80% of the notes used. They are in a ratio of 1:2, sometimes 1:3. The briefest among them is

TABLE I
Frequency of Notes

| Composer and tempo | Frequency of notes (sample of 100 notes) | | | | | | | | | Frequency of the two main notes |
|---|---|---|---|---|---|---|---|---|---|---|
| Chopin I ♩ = 132 | | | 1 | | 29 | 3 | 52 | 15 | | 81 |
| Chopin II Tempo Giusto | | | 1 | 6 | 10 | | 83 | | | 93 |
| Beethoven I Adagio | | | | | 14 | 5 | 23 | 56 | 2 | 79 |
| Beethoven II Allegro Vivace | | 13 | 1 | | 54 | | 32 | | | 86 |
| Beethoven III Adagio | | | | 8 | | 29 | 31 | 32 | | 63 |
| Debussy I ♩ = 66 | | 1 | 4 | 1 | 10 | | 29 | 55 | | 84 |
| Debussy II ♩ = 184 | | | 1 | | | | 44 | 55 | | 99 |
| Fauré I ♩ = 52 | | | | | 4 | 22 | | 74 | | 96 |
| Fauré II ♩ = 96 | | | 1 | 1 | 23 | | 61 | 14 | | 84 |
| Stravinsky I ♩ = 112 | | | 5 | | 21 | | 74 | | | 95 |
| Stravinsky II ♩ = 112 | | | | | 4 | | 43 | 45 | 8 | 88 |
| Bela Bartok I ♩ = 100 | | | | 8 | 24 | | 68 | | | 92 |
| Bela Bartok II ♩ = 72 | | 3 | 9 | 11 | 12 | | 65 | | | 77 |
| Bela Bartok III ♩ = 60 | 1 | | | | 15 | | 13 | 71 | | 84 |
| Weill K. ♩ = 84 | | | 2 | | 29 | | 61 | 8 | | 90 |

also the most frequent. The duration of notes in our examples varies—taking into account the tempos—from 150 to 290 msec. These values are again found in performances recorded by Gabrielsson (1973). We found, in spontaneous rhythms, durations going from 180 to 280 msec, and we have already stated that the shortest time was also the most frequent. These comparisons are striking. The massive use of two different notes is explained, we said, by the difficulty in identifying more than two durations. (However, let us not forget that composers use other, longer durations that have an important aesthetic role and which permit syncopations and pauses.)

The two basic notes, moreover, do not have the same perceptual status (Table I). The brief note, or the interval that we have called *short time*, does not last. One was able to speak of point in time and, for the rapid succession of two sounds, of a collective perception (Schultze, 1908). The other note on the contrary, which is double or triple the first one, corresponds to the perception of a *duration*. These notes vary between 300 and 900 msec—that is, the range of durations which appear as neither too brief nor too long and which are centered around an optimal duration of 600 msec, which is also that of spontaneous tempo.

Comparisons with the results of our analysis also lead us to consider the ratio of 1.5 between two successive notes (or two intervals) that appeared to us as ambiguous from the perceptual viewpoint. It corresponds to the case of the pointed note. In partitions, one rarely finds a succession of two notes in a ratio of 1.5, and musicians are acquainted with the difficulty of realizing such a succession. The pointed note is most often differentiated from another note, in a ratio of 1 to 3, as between a ♪ and a ♩.

Even in the case where a pointed quaver follows a double quaver (ratio of 1:3), the ratio is distorted in the performances recorded by Gabrielsson (1973a); it is closer to 1:4 than 1:3. The ratio of 1:2 between the quavers and the double quavers is often slightly increased, especially when there is syncopation. The ratio of equality between two ♪. or two ♩ is not strictly respected but the difference does not attain 1.2, which remains within the limits found in the spontaneous production of tapped rhythms.

However, it is equally necessary to remember that, at the level of composition, and also of perception, notes are grouped into what in metric one calls a *time*. Two ideas are fundamental in order to define a bar, the number and duration of each time. The bar, in principle, has a unity that provides the relative accentuations of each of the times, the first being, in classical music, the most accentuated. Two questions are important: What do musicians do? And what does the audience perceive?

The accentuation of a note slightly lengthens its duration, a fact that we have already found in nonmusical rhythms. The difference in accentuation of diverse notes, in a dance rhythm, can vary from 10 dB in piano performances to 20 dB in percussion performances (Gabrielsson, 1973).

In general, the first time is accentuated. But how does the subject perceive succession and grouping of times? Recent research by Vos (1978) brings along a first answer. On commercial versions of Bach's preludes, subjects familiar with classical music but not particularly acquainted with the pieces chosen were asked to tap in synchrony with the beginning of each bar—that is, with the beginning of the perceived rhythmic pattern. The subject did not tap in all cases on the first beat of the bar. Let us take the example of a 2/4 bar that lasted 1.75 sec. Forty percent of the subjects tapped in synchrony with the first beat | 1̂ 2 |;[4] 45% tapped on the second beat of the measure thus linked in a rhythmic pattern with the first beat of the following measure | 1 2̂ | 1 2̂ |; 10% tapped each beat. For a 3/4 bar, 20% tapped each beat. Eighty percent tapped as if it were a bar with two beats, grouping two of the three beats, the last beat of a bar being grouped one time out of two with the first beat of the following bar | 1̂ 2 3̂ | 1 2̂ 3 |. This result shows that arguments regarding the place of accent are irrelevant since for the same musical performance, some perceive, as one can predict, the accent on the first beat, whereas others do so on the last. In the same ternary sequence, a binary perception corresponds to an accent successively placed on the first and last beats. As a general rule, the intervals between the subjects' taps were shorter than the length of the bars. These varied between 1.75 sec and 4.8 sec in the examples studied and the intervals between the taps varied from .8 sec to 2.4 sec. Vos wondered, without being able to answer, whether it was the melodic saliency that led to these cuts or whether it was the difficulty of storing too long a series in short-term memory? Without ignoring the importance of melodic structure, we think that the generality of the observed phenomenon is particularly explained by limits in storage capacity and by the necessity of maintaining a perceived succession between successive taps. From this point of view, tempo plays a decisive role.

[4]The sign ">" indicates the place of the tap.

Biblioteca del Real Conservatorio Superior de Música - MADRID -

The ease of the task of motor accompaniment to musical rhythm brings us back to the perceptual and motor aspects of rhythm that we have found in the most simple rhythmic forms. Musical time, like the foot in poetry, recalls the origins of rhythm or chant, dance and music linked by the beat of the foot made by a succession of *arsis* and of *thesis*. In the time of the Greeks, theorists already debated whether the accent corresponded to *arsis* or to *thesis*. Aristoxenus of Tarentum, moreover, did not speak of upbeat and of fall but of high times and of low times.

This link between rhythmic perception and movement also appeared in research requiring judgments and not performances. Gatewood (1927) asked subjects to listen to diverse pieces of music and report which of the four following qualities: rhythm, melody, harmony, or pitch appeared to be dominant. For each of the pieces, the subjects had to indicate impressions as evoked by these pieces. Impression of movement was perceived in 64% of the pieces in which rhythm was dominant, 25% of the pieces in which melody was dominant, 15% of the pieces in which harmony was dominant, and 12% of the pieces in which pitch was dominant. The movement thus appeared associated with rhythm and was, moreover, also present in the pieces where melody, harmony, or pitch were judged to be predominant.

More recent work has dealt with the problem of the dimensions of the rhythmic experience. Gabrielsson (1973a) set out from three sorts of musical samples: (a) monophonic structures played on a piano or on a drum, (b) polyphonic structures of diverse dances, in which the rhythm arises from the play of the pitch of percussion instruments and of their duration without melodic intervention, and (c) real music (most often dances).

By using diverse methods (judgments of similarity of two sequences, estimates from adjectives recalling the semantic differential, performance with monophonic rhythms, free verbal descriptions) and by using the methods of factorial analysis and of multidimensional analysis, Gabrielsson (1973b) could sum up his results by distinguishing three groups of dimensions:

1. *Structural properties of rhythm.* One distinguishes here what we call the relationship between perceived rhythms and the bar: coincidence or not, place of accent(s), degrees of perceptual prominence of a basic pattern—accentuation versus clearness, simplicity versus complexity, or uniformity versus variation in the pattern. "The more different note values used, the more duration of duration patterns, the more syncopation, the more variation in instrumentation, the more leaps in a melody, the more changes in harmonic functions, etc. the more varied and in most cases, the more complex the rhythm will be judged to be (Gabrielsson, 1973b p. 10)."

2. *Movement properties.* This defines rapidity and tempo forward movement (depending on the fact that in a pattern the movement seems to accelerate or decelerate) and movement characteristics (i.e., different aspects of experienced movements in relation to rhythmic experience: dancing–walking; floating–stuttering; solemn–swinging, and others).

3. *Emotional aspects.* Gabrielsson thinks that these are characterized by the dimensions of vital–dull, excited–calm, rigid–flexible, and solemn–playful.

## V. CONCLUSION

All the above approaches attest to some links that exist between rhythmic perception and movement. This has led us several times to speak, as did Ruckmick (1927), of rhythmic experience, thereby enlarging the perceptual aspect as it is most often accompanied by motor stimuli and by emotional reactions.

Perception is first of a temporal order—that is, of a regularity. It is, moreover, striking that in music the temporal data—at least that of the notes—are always explicit. This necessary condition is not always sufficient. The ordered elements can, in effect, be varied as to duration, intensity, pitch, and contain silences. The rhythmic structures are always perceptually complex.

However, rhythm, differing from melody, is made up, above all, of temporal and intensive patterns. One aspect can predominate over the other. The Gregorian chant is the best example of a temporal structuring without intensity and without periodicity. The march and most dances represent the other extreme where the pattern of accents imposes itself with its regularity and where isochronous patterns repeat each other.

All classical music falls between these two extremes. In order to understand the play of regularity and variety, it is perhaps necessary to mention the most simple studies. Brown (1911) has shown that while tapping successions of rhythmic patterns, the variability of the durations is half as strong as that of the intensities. Let us complete what we said above. If in music the durations are always explicit, the indications of accents are much more vague. The rule that states that the first beat of the bar is always accented is only a convention. We have seen that bars, even those of a classical musician such as Bach, and even when played by the same orchestra, can be perceived in multiple ways.

In all polyphonic musical performances, it is necessary however, that there be a regularity in order to permit the anticipation of playing and the synchrony of the artists. In reality, all musical performances consist of isochronous repetition and, simultaneously, of more varied patterns, but which, in a complex way, fit into the play of isochronous repetitions. Perhaps Chopin's remark can be generalized: "Let your left hand be an intransigent and rigorous orchestra conductor and your right hand do what it wants." In the same spirit one can cite this passage from a letter written by Wagner to Liszt (Dumesnil, 1949) that singers respect the duration of notes "by staying within the indicated bar... if they leave it, in order to go further, let them do so with an intelligent liberty and instil fire rather than caution, thereby entirely making the continuity which the bar imposes disappear. If only they produce the impression of an animated and poetic style, we will have won everything."

It is necessary not to dissociate the motor *behaviors* linked to rhythms from these complex perceptions. Still, two aspects are to be considered here. The play of music is always based on movements. Some are very voluntary and lengthily learned in order to go beyond simple determinisms. Others, on the contrary, give way to these determinisms. They appear when the repetitions of accented patterns impose themselves

on the player and on the listener. This subtle or rough motor component has the effect that rhythmic perception is plurisensorial. This is why we speak of rhythmic experience. The movement brings along more particularly affective reactions which are also a component of this experience. The affective aspect is all the more important, in part, because the anticipation of successive patterns facilitates synchrony between individuals. They are spontaneously realized in marches and in dances. All socialization of behavior, as is well known, reinforces their affective impact.

This said, in spite of the plurality of rhythmic components and of artistic realizations, is there a general sense of rhythm? We have explored this aspect (Hiriarborde & Fraisse, 1968) by using factorial analysis on a series of tests—some predominantly perceptual (discrimination, reproduction, transcription of temporal or intensive structures), others predominantly motor (adaptation to a change in pattern, polyrhythmic hand-foot, synchronization with good or with ambiguous forms)—and by using the batteries of musical aptitude tests developed by Seashore and Wing. All of these tests were in positive correlation with each other, except for the tests of intensity discrimination and pitch discrimination from the battery by Seashore, and the tests of choice of the best harmonization and judgment of rhythmic accent by Wing.

Using a centroid analysis and by looking for the orthogonal factors, we found three main factors:

1. *Perceptual structuration.* The tests that are the most saturated in this factor are based on the discrimination of temporal structures.
2. *Rhythmic anticipation.* The tests that are the most saturated in this factor are those of synchronization and several of those by Seashore and by Wing that imply the memorization of a pattern in order to compare it with a following one.
3. *Practo-rhythmic.* The tests that are the most saturated in this factor are those of coordination with alternate movements of the hand and foot and those of adaptation to changes in rhythm, both of which require a voluntary control of rhythmic movement. Moreover, we found a musical factor and a factor of discrimination.

Thackray (1969) carried out a similar study based on his experience as a music teacher. He started from a battery of tests which he classified into three categories:

1. *Rhythmic perception:* Perceiving the number of sounds in a pattern, distinguishing two cadences, comparing durations, identifying the place of an accent in a pattern.
2. *Rhythmic performance:* reproducing temporal and intensive patterns of sounds and of a short melody, conserving a tempo.
3. *Rhythmic movement:* rhythmic quality of movements that put the whole body to play—sequences of (rhythmic) movements, following music in which the tempo, the bar, and so on are varied.

The correlations between all these tests are quite high, especially within the same group of tests. Using these results and applying the method of Spearman-Burt, Thackray, in effect, extracted a general factor. We also could have extracted one had we chosen this method of analysis. We think, however, that a plurality of factors better accounts for the plurality of rhythmic aptitudes that correspond, moreover, to

the plurality of the aspects of rhythm, as Thackray hypothesized himself when constructing his tests.

## REFERENCES


Allen, T. W., Walker, K., Symonds, L., & Marcell, M. Intrasensory and intersensory perception of temporal sequences during infancy. *Developmental Psychology*, 1977, *13*, 225-229.

Allport, G. W., & Vernon, P. E. *Studies in expressive movement*. New York: Macmillan, 1933.

Benveniste, E. La notion de "rythme" dans son expression linguistique. *Journal de Psychologie Normale et Pathologique*, 1951, *44*, 401-411.

Berg, W. K. Cardiac orienting responses of 6- and 16-weeks-old infants. *Journal of Experimental Child Psychology*, 1974, *17*, 303-312.

Bolton, T. L. Rhythm. *American Journal of Psychology*, 1894, *6*, 145-238.

Bovet, P. *Quantité d'information transmise dans la perception des durées brèves*. Thèse de 3e cycle, Université René Descartes, Paris, 1974 (unpublished).

Brown, W. Temporal and accentual rhythm. *Psychological Review*, 1911, *18*, 336-346.

Brunswik, E. *Perception and the representative design of psychological experiments*. Berkeley and Los Angeles, California: University of California Press, 1956.

Chang, M. W., & Trehib, S. E. Infant's perception of temporal grouping in auditory patterns. *Child Development*, 1977, *48*, 1666-1670.

Clifton, R. K., & Meyers, W. J. The heart rate response of fourth-month-old infants to auditory stimuli, *Journal of Experimental Child Psychology*, 1969, 7, 122-135.

Crowder, R. G., & Morton, J. Precategorical acoustic storage (P A S). *Perception & Psychophysics*, 1969, *5*, 365-373.

Demany, L., Mc Kenzie, B., & Vurpillot, E. Rhythm perception in early infancy, *Nature (London)*, 1977, *266*, 718-719.

Dietze, G. Untersuchungen über den Umfang des Bewusstseins bei regelmässig auf einander folgenden Schalleindrücken. *Philosophische Studien*, 1885, *2*, 362-393.

Dumesnil, R. *Le rythme musical*. Paris: La Colombe, 1949.

Ehrlich, S. Le mécanisme de la synchronisation sensori-motrice. *L'Année Psychologique*, 1958, *58*, 7-23.

Ehrlich, S., Oléron, G., & Fraisse, P. La structuration tonale des rythmes. *L'Année Psychologique*, 1956, *56*, 27-45.

Fraisse, P. Mouvements rythmiques et arythmiques. *L'Année Psychologique*, 1946-1947, *47-48*, 11-21.

Fraisse, P. De l'assimilation et de la distinction comme processus fondamentaux de la connaissance. In *Miscellanea Psychologica Albert Michotte*. Louvain: Institut Supérieur de Philosophie, 1947. Pp. 181-195.

Fraisse, P. *Les structures rythmiques*. Louvain: Editions Universitaires, 1956.

Fraisse, P. *Psychology of time*. New York: Harper, 1963.

Fraisse, P. L'anticipation de stimulus rythmiques. Vitesse d'établissement et précision de la synchronisation. *L'Année Psychologique*, 1966, *66*, 15-36.

Fraisse, P. *Psychologie du rythme*. Paris: Presses Universitaires de France, 1974.

Fraisse, P. Is rhythm a gestalt? In S. Ertel, L. Kemmler, & Stadler (Eds.), *Gestalttheorie in der modernen Psychologie*. Darmstadt: Steinkopff, 1975. Pp. 227-232.

Fraisse, P. Les synchronisations sensori-motrices aux rythmes. In J. Requin (Ed.), *Anticipation et comportement*. Paris: Editions C.N.R.S. 1980.

Fraisse, P., & Ehrlich, S. Note sur la possibilité de syncoper en fonction du tempo d'une cadence. *L'Année Psychologique*, 1955, *55*, 61-65.

Fraisse, P., & Fraisse, R. Etudes sur la mémoire immédiate. I. L'appréhension des sons. *L'Année Psychologique*, 1937, *38*, 415-423.

Fraisse, P., & Oléron, G. La structuration intensive des rythmes. *L'Année Psychologique*, 1954, *54*, 35-52.

Fraisse, P., Pichot, P., & Clairouin, G. Les aptitudes rythmiques. Etude comparée des oligophrènes et des enfants normaux. *Journal de Psychologie Normale et Pathologique*, 1949, *42*, 309-330.

Frischeisen-Köhler, I. *Das Persönliche Tempo. Eine erbliologische Untersuchung*. Leipzig: Thieme, 1933. (a)
Frischeisen-Köhler, I. Festsellung des weder langsamen noch schnellen (mittelmässigen) Tempos. *Psychologische Forschung*, 1933, *18*, 291–298. (b)
Gabrielsson, A. Similarity ratings and dimension analyses of auditory rhythm patterns. I. *Scandinavian Journal of Psychology*, 1973, *14*, 138–160. (a)
Gabrielsson, A. Studies in rhythm. Acta Universitatis Upsaliensis 7, Uppsala 1973. (b)
Garner, W. R. *Uncertainty and structure as psychological concepts*. New York: Wiley, 1962.
Garner, W. R., & Gottwald, R. L. The perception and learning of temporal pattenrs. *Quarterly Journal of Experimental Psychology*, 1968, *20*, 97–109.
Gatewood, L. L. An experimental study of the nature of musical enjoyment. In M. Shoen (Ed.). *The effects of music*. New York: Harcourt, 1927.
Handel, S. Perceiving melodic and rhythmic auditory patterns. *Journal of Experimental Psychology*, 1974, *103*, 922–933.
Harrel, T. W. Factors influencing preference and memory for auditory rhythm. *Journal of General Psychology*, 1937, *17*, 63–104.
Harrison, R. Personal tempo and the interrelationships of voluntary and maximal rates of movement. *Journal of General Psychology*, 1941, *24–25*, 343–379.
Hawkes, G. R. Information transmitted via electrical cutaneous stimulus duration. *Journal of Psychology*, 1961, *51*, 293–298.
Hiriartborde, E., & Fraisse, P. *Les aptitudes rythmiques*, Monographies françaises de Psychologie, Paris. CNRS, 1968.
James, W. *The principles of psychology*. London: Macmillan, 1891. (2 Vols.)
Koffka, K. Experimentelle Untersuchungen zur Lehre von Rhythmus. *Zeitschrift für Psychologie*, 1909, *52*, 1–109.
Leavitt, L. A., Brown, J. W., Morse, P. A., & Graham, F. K. Cardiac orienting and auditory discrimination in 6 week-old-infants. *Developmental Psychology*, 1976, *12*, 514–523.
Lehtovaara, A., Saarinen, P., & Järvinen, J. Psychological studies on twins. II. The psychomotor rhythm: environmental versus hereditary determination. *Reports from the Psychological Institute, University of Helsinki*, 1966, No. 3.
Mac Dougall, R. The structure of simple rhythm forms. *Psychological Review, Monograph Supplements*, 1903, *4*, 309–416.
Mach, E. Untersuchungen über den Zeitsinn des Ohres. *Sitzungeberichte der Wiener Akademie der Wissenschaften*, 1865, K*ℓ*, 51.
Martin, J. G. Rhytmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, 1972, *79*, 487–509.
Meumann, E. Untersuchungen zur Psychologie und Aesthetik des Rhythmus, *Philosophische Studien*, 1894, *10*, 249–322, 393–430.
Michon, J. A. Studies on subjective duration. I. Differential sensitivity in the perception of repeated temporal intervals. *Acta Psychologica*, 1964, *22*, 441–450.
Miner, J. B. Motor, visual and applied rhythms. *Psychological Review, Monograph Supplements*, 1903, *5*, 1–106.
Mishima, J. Fundamental research on the constancy of "mental tempo". *Japanese Journal of Psychology*, 1951–1952, *22*, 27–28.
Mishima, J. On the factors of the mental tempo. *Japanese Psychological Research*, 1956, *4*, 27–38.
Mishima, J. *Introduction to the morphology of human behavior. The experimental study of mental tempo*. Tokyo: Tokyo Publishing, 1965.
Miyake, I. Researches on rhythmic activity. *Studies from the Yale Psychological Laboratory*, 1902, *10*, 1–48.
Montpellier, G. de. *Les altérations morphologiques des mouvements rapides*. Louvain: Institut Supérieur de Philosophie, 1935.
Murphy, L. E. Absolute judgments of duration. *Journal of Experimental Psychology*, 1966, *71*, 260–263.
Oléron, G. Etude de la "perception" des structures rythmiques, *Psychologie Française*, 1959, No. 4, 176–189.
Oshinsky, J. S., & Handel, S. Syncopated auditory polyrhythms: Discontinuous reversals in meter

interpretation. *Journal of the Acoustical Society of America*, 1978, *63*, 936-939.
Paillard, J. *Réflexes et régulations d'origine proprioceptive chez l'homme*. Paris: Arnette, 1955.
Preusser, D. The effect of structure and rate on the recognition and description of auditory temporal patterns. *Perception & Psychophysics*, 1972, *11*, 233-240.
Preusser, D., Garner, W. R., & Gottwald, R. L. The effect of starting pattern on descriptions of perceived temporal patterns. *Psychonomic Science*, 1970, *21*, 219-220. (a)
Preusser, D., Garner, W. R., & Gottwald, R. L. Perceptual organization of two-element temporal patterns as a function of their component one-element patterns. *American Journal of Psychology*, 1970, *83*, 151-170. (b)
Plotkin, L. Stimulus generalization in Morse code learning. *Archives of Psychology*, 1943, No. 287.
Rimoldi, H.J.A. Personal tempo. *Journal of Abnormal and Social Psychology*, 1951, *46*, 283-303.
Royer, F. L., & Garner, W. R. Response uncertainty and perceptual difficulty of auditory temporal patterns. *Perception & Psychophysics*, 1966, *1*, 41-47.
Royer, F. L., & Garner, W. R. Perceptual organization of nine element auditory temporal patterns, *Perception & Psychophysics*, 1970, 7, 115-120.
Ruckmick, C. A. The rhythmical experience from the systematic point of view. *American Journal of Psychology*, 1927, *39*, 355-366.
Sander, F. Experimentelle Ergebnisse der Gestaltpsychologie. *10 Kongres für Experimentelle Psychologie, Iena*, 1928, p. 23.
Schmidt, E. M. Uber den Aufbau rhythmischer Gestalten. *Neue Psychologische Studien*, 1939, *14*, 1-98.
Schultze, F. E. Beiträg zur Psychologie des Zeitbewusstseins. *Archiv für die Gesamte Psychologie*, 1908, *13*, 275-351.
Sears, G. H. A contribution to the psychology of rhythm. *American Journal of Psychology*, 1902, *13*, 28-61.
Steedman, M. J. The perception of musical rhythm and metre. *Perception*, 1977, *6*, 555-569.
Stern, W. Das psychisch Tempo. In *Uber psychologie der individuellen differenzen*. Leipzig: Barth, 1900.
Stetson, R. H. A motor theory of rhythm and discrete succession. *Psychological Review*, 1905, *12*, 250-270, 292-350.
Sturger, P. T., & Martin, J. G. Rhythmic structure in auditory temporal pattern perception and immediate memory. *Journal of Experimental Psychology*, 1974, *102*, 377-383.
Temperley, N. M. Personal tempo and subjective accentuation. *Journal of General Psychology*, 1963, *68*, 267-287.
Thackray, R. *An investigation into rhythmic abilities*. Londres: Noveleo, 1969.
Tisserand, M., & Guilhot, J. Etude du tempo de 335 sujets masculins dans la région parisienne. *Biotypologie*, 1949-1950, *10,11*, 89-94.
Vierordt, K. *Der Zeitsinn nach Versuchen*. Tübingen: Laupp, 1868.
Vos, P. G. Pattern perception in metrical tone sequences. Unpublished thesis, University of Nijmegen, 1973.
Vos, P. G. Identification of metre in music. Report 76 ON 06. University of Nijmegen, 1976.
Vos, P. G. Temporal duration factors in the perception of auditory rhythmic patterns. *Scientific Aesthetics*, 1977, *1*, 183-199.
Vos, P. G., Leeuwenberg, E. L., & Collard, R. F. What melody tells about meter in music. Report 78 FU 03. University of Nijmegen, 1978.
Vurpillot, E., Ruel, J., & Castrec, A. Y. L'organisation perceptive chez le nourrisson: réponse au tout et à ses éléments. *Bulletin de Psychologie*, 1977, *30*, 396-405.
Wallin, J.E.W. Researches on the rhythm of speech. *Studies from the Yale Psychological Laboratory*, 1901, *9*, 1-142.
Wallin, J.E.W. Experimental studies of rhythm and time. *Psychological Review*, 1911, *18*, 100-131, 202-222.
Wallon, H. *Les origines du caractère chez l'enfant*. Paris: Presses Universitaires de France, 1949.
Weaver, H. E. Syncopation: a study of musical rhythms. *Journal of General Psychology*, 1939, *20*, 409-429.
Werner, H. Rhythmik, eine mehrwertige Gestaltenverkettung. *Zeitschrift für Psychologie*, 1919, *82*, 198-218.
Wertheimer, M. Untersuchungen zur Lehre von der Gestalt. *Psychologische Forschung*, 1922, *1*, 47-58.

Wertheimer, M. *Psychologische Forschung*, 1923, *4*, 301-350.
Woodrow, H. A quantitative study of rhythm. *Archives of Psychology*, 1909, *18*, No. 1.
Woodrow, H. The role of pitch in rhythm. *Psychological Review*, 1911, *18*, 54-71.
Zenatti, A. Jugement esthétique et perception de l'enfant entre 4 et 10 ans dans des épreuves rythmiques. *L'Année Psychologique*, 1976, *76*, 93-115.

# 7

# Timing by Skilled Musicians

*Saul Sternberg, Ronald L. Knoll, and Paul Zukofsky*


I. Introduction: Perception, Production, and Imitation of Fractions of the Beat ..... 182
A. Previous Studies of the Psychophysics of Time ..... 183
B. Procedures and Notation ..... 184
C. Subjects ..... 185
D. Caveats ..... 186
E. Principal Findings ..... 186
II. Perceptual Judgment of Beat Fractions ..... 187
A. Single-Fraction Perceptual Judgment (Experiment 1) ..... 188
B. Multiple-Fraction Perceptual Judgment (Experiment 2) ..... 194
C. Judgment Precision ..... 196
III. Production of Beat Fractions ..... 198
A. Use of Tap-Click Synchronization to Correct for Differential Subjective Delays ..... 198
B. Expectations from a Feedback Model of Production ..... 200
C. One-Response Production (Experiment 3) ..... 200
D. Rejection of the Feedback Model ..... 201
E. Repeated-Response Production (Experiment 4) ..... 202
F. Production Precision: Evidence Against a Reaction-Time Explanation of the Production Error ..... 203
G. Implications of Other Analyses of Psychophysical Scaling for the Production–Perception Disparity ..... 205
IV. Imitation of Beat Fractions ..... 207
A. Four Simple Alternatives for Imitation ..... 207
B. One-Response Imitation (Experiment 5) ..... 209
C. Choice Among Alternative Imitation Functions: Rejection of Accuracy and Full-Concatenation Models ..... 209
D. Implicit Scaling of Beat Fractions from Imitation and Production Data ..... 211


[1]The research described in this chapter was conducted in the Human Information-Processing Research Department of Bell Laboratories, Murray Hill, N.J. Much of it was described in a doctoral dissertation in psychology submitted by R. L. Knoll to Princeton University.

Biblioteca del Conservatorio Superior



Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

V. A Shared-Process Model of the Perception, Production, and Imitation of Beat Fractions .. 212
A. Definition of the Model ..... 212
B. Restrictions on the Four Processes ..... 214
VI. Further Analysis of Perceptual Judgment ..... 215
A. Attention Shifts and Delays: Effect of Marker-Click Pitch (Experiment 6) ..... 215
B. Time Marking by Onset versus Offset: Invariance of Judgment with Prolonged Markers (Experiment 7) ..... 216
C. A Test of the Rate Constancy of Subjective Time Between Beats: Effect on Fraction Perception of Fraction Location Relative to the Beat (Experiment 8) ..... 217
D. A Constraint on the Precision of Dual Time Judgments and Its Implications for Timing Models and the Use of Feedback ..... 218
E. The Dependence of Perceptual Judgment on Duration versus Fraction: Effects of the Beat Interval (Experiment 9) ..... 220
VII. Further Analysis of Production ..... 224
A. Production Errors with Musical Instruments (Experiment 10) ..... 224
B. Production with Marker-Click Feedback (Experiment 11) ..... 225
C. Production of Single versus Multiple Subdivisions of the Beat (Experiment 12) ..... 226
VIII. Summary ..... 229
Glossary ..... 231
Appendices ..... 231
A. Staircase and Constant Stimulus Methods ..... 231
B. Additional Details of Design and Procedure ..... 233
C. Measures of Location of the Psychometric Function ..... 234
D. Measures of Spread of the Psychometric Function ..... 236
E. Stimulus-Averaging versus Response-Averaging Methods for Deriving a Psychophysical Scale ..... 236
References ..... 237

## I. INTRODUCTION: PERCEPTION, PRODUCTION, AND IMITATION OF FRACTIONS OF THE BEAT

In this chapter we report 12 experiments that explore how skilled musicians perceive time and time performance in contexts similar to those in music. We discuss some of the perception and performance constraints we found and their implications for underlying timing mechanisms. Our emphasis is on the short time intervals—fractions of a second—that are among the shortest durations specified by musical notation. In our experiments, as in Western music, these intervals occur in the context of a train of periodic beats and are defined as fractions of the beat interval, or duration ratios. Because the beats are provided externally, our experimental tasks are probably most directly analogous to aspects of ensemble playing or of solo playing with a metronome or conductor. Musician subjects are among the best for answering questions about the relation between notation and performance, the constraints on the precision of ensemble playing, and temporal illusions in listening. We hope, however, that the constraints and mechanisms we uncover are relevant to human timing in general; our choice of musicians as subjects in timing experiments is based on our belief that a fruitful approach to the understanding of any human function is the study of skilled practitioners of that function.

We have examined the performance of skilled musicians in three laboratory tasks designed to capture aspects of musical practice: perception, production, and imitation of fractions of the beat interval. All three functions are required of musicians during ensemble rehearsal and performance, for example. It is plausible that because of the requirement that players "keep together," performing experience would cause the three functions to become at least consistent with each other and probably "correct" (consistent with the notation) as well. Neither of these expectations was borne out by our experiments; instead, we observed surprisingly large systematic errors and inconsistencies across these laboratory tasks. In agreement with our observations, the one study of the temporal coordination of ensemble playing of which we know (Rasch, 1979) reveals considerable inaccuracy even in unison attacks.[2] Our findings suggest that other studies of timing in actual musical performance would be of great interest, but unfortunately too little is known at present for us to comment on the extent to which our findings apply outside the laboratory.

The study of performance in our three different but related tasks, together with the analysis of relations among performances, permits some surprisingly strong inferences about these timing mechanisms. Some of our experiments may also be regarded as first steps in establishing a relation (a psychophysical function) between traditional Western musical notation—a notation that specifies time ratios—and the corresponding perceived and produced time ratios among people highly trained in use of the notation.

## A. Previous Studies of the Psychophysics of Time

There is a vast, conflicting literature on the psychology of time that we shall only touch on here; we refer the interested reader to Fraisse (1963, 1978), Poppel (1978), Sternberg & Knoll (1973), Woodrow (1951), and Zelkind & Sprug (1974) for reviews and references.

For time intervals greater than a second, there is a long history of experiments focused on determining the psychophysical function relating subjective to actual duration. Most investigators have assumed that estimation and production are consistent, in the sense that they reflect the same psychophysical function. (We shall see this assumption fail dramatically in our experiments.) Investigators have agreed less well on how to account for reproduction (imitation) performance.[3] Based on his review of the data from over 100 studies of the reproduction, production, and estimation of

[2]The limitation in Rasch's analysis of trio playing to measurement of unison attacks probably implies that note sequences representing small fractions of a beat—where we found the largest errors and inconsistencies—were underrepresented.

[3]For example, according to Eisler's (1976) analysis, the subject in a reproduction task produces an interval whose subjective duration is half of the subjective duration of the sum of itself and the presented interval. This leads to a nonlinear relation between the presented and reproduced intervals. According to other models (e.g., Carlson & Feinberg, 1968), the subject produces an interval that is subjectively equal to the presented interval; this implies a linear reproduction function of unit slope, regardless of the underlying psychophysical function.

time intervals, Eisler (1976) concluded that subjective duration $D$ is a power function of objective duration $d$ with an exponent of about .9: $D = ad^{.9}$. (An exponent less than unity implies that the subjective ratio of two time intervals is smaller than the corresponding objective ratio; for example, 2 sec would appear less than twice as long as 1 sec.)

For intervals smaller than one second, the data are both more sparse and less consistent. Both Michon (1967) and Svenson (1973) report magnitude-estimation data supporting a change in the exponent of the power function at about .5 sec; whereas the exponent is approximately 1.0 for intervals greater than .5 sec, its value decreases to about .5 for smaller intervals. However, other investigators report either no such change in the function (e.g., Steiner, 1968) or changes at different points (Nakajima, Shimojo, & Sugita, 1980; see also Zwicker, 1969, and Fastl, 1977).[4] According to one view of human timing that has attracted interest, there exists a central timing process (or "clock") that functions similarly in the judgment, production, and reproduction of duration; results from experiments with time intervals greater than a second provide some evidence favoring such a common central process.[5] The discontinuities in judgment revealed in the experiments of Michon and Svenson suggest that a more sensitive test of a common central timing process might be obtained with time intervals less than a second; this suggestion provides one framework for the present investigation.

Despite the importance of time ratios in music, we know of no substantial studies in which judgments or productions of a range of *ratios* have been systematically examined.[6]

## B. Procedures and Notation

We have examined performance in three basic tasks (illustrated on the left side of Fig. 1) and variations thereof. All three involve a train of beats specified by *beat clicks;* in most experiments the time from one beat click to the next (*beat interval b*) was 1.0 sec. In describing these tasks we denote stimuli by lower-case letters and responses by upper-case letters.

In the *perceptual judgment* task, one (or more) of the beat clicks was followed by a *marker click,* to form a *time-pattern stimulus* (see Fig. 1A). The time interval between beat and marker clicks (called the *fractional interval*) was to be judged in relation to the beat interval. (For example, if the *stimulus fraction* is $f = 1/8$ and the beat interval is $b = 1000$ msec, the fractional interval is $bf = 125$ msec.) In musical terms our fractions corresponded to note values between a 32nd note (𝅘𝅥𝅯) and a quarter note (♩), in which the quarter note equals one beat and the rate (in most experiments) was 60 beats

[4]Data from the three subjects in the study by Nakajima *et al.* (1980) are too inconsistent to justify firm conclusions.

[5]See Treisman (1963), .25–6 sec; Carlson & Feinberg (1968, 1970), 1–10 sec; and Adam, Castro, & Clark (1974), 3–40 sec.

[6]Except in one small study (Richards, 1964), the only ratios studied in production or perception tasks have been 1:2 (halving) and 2:1 (doubling). Povel (1981) has studied a range of ratios, but only in an imitation task.

per minute. The subject's response in this task was made in terms of *fraction names*, $N$. (For example, he or she might be asked to decide whether the stimulus fraction appeared to be less than or greater than $N = 1/8$th of a beat.) The fraction name was specified both in musical notation and as a numerical fraction; for example, the subject in this case would be asked whether the marker click was early or late relative to the pattern ♩ ♪♩.. . (The *correct* or *target* fractional interval for $N = 1/8$, given $b = 1000$ msec is, of course, $bN = 125$ msec.) The outcome of this procedure is the determination, for each of a set of fraction names, which stimulus fraction $f$ corresponds to it. The relation between $f$ and $N$ defines a *judgment function*, $f = \mathbf{J}(N)$.[7]

In the *production* task a train of beat clicks was presented, but no marker click (see Fig. 1C). The "stimulus" here was a *fraction name* $n$. The subject made a *timed response* by tapping his or her finger after a beat click, with the aim of producing a fractional interval between click and tap whose duration was appropriate for the specified fraction name. The ratio of fractional interval to beat interval gives the *produced fraction* $F$; the fractional interval is $bF$. The outcome of this procedure is the determination, for each of a set of fraction names $n$ the average value of the fraction $F$ produced to correspond to it. The relation between $F$ and $n$ determines a *production function* $F = \mathbf{P}(n)$.

In the *imitation* task the stimulus was a *time-pattern stimulus*, as in the judgment task (defined by the beat interval $b$ and the fractional interval $bF$) and the response was a timed response $F$, as in the production task (see Fig. 1E). The subject attempted to equate the produced fraction $F$ to the stimulus fraction $f$. The outcome of this procedure is the determination—for each of a set of stimulus fractions $f$—the average value of the fraction $F$ produced to correspond to it. The relation between $F$ and $f$ determines an *imitation function*, $F = \mathbf{I}(f)$.[8]

Note that the marker click and the timed response of our basic procedures were single offbeat events. Although not frequent in earlier music, the playing of a note after the beat without playing a note on the beat is not unusual in the music of the past 60 years.

## C. Subjects

Our principal subjects were three professional musicians: Susan Bush, flutist (SB); Pamela Frame, 'cellist (PF); and Paul Zukofsky, violinist and conductor (PZ). PZ, who had substantially more musical experience than SB or PF, produced data that were more consistent, both within and across experiments. For this reason (and because results from other subjects usually agreed with his), we tend to weight his data more heavily. We obtained a small amount of corroborative data from Pierre Boulez (PB),

[7]Note that we have expressed the stimulus $f$ as a function $\mathbf{J}$ of the response $N$ for convenience in later discussion. We shall use the opposite convention for production and imitation.

[8]In some instances in which there is no ambiguity and the beat interval is 1 sec, we shall use the fraction symbols ($n$, $N$, $f$, and $F$) also to denote fractional intervals ($bn$, $bN$, $bf$, and $bF$). Note also that we do not distinguish notationally between quantities such as $n$, $N$, $f$, and $F$, and their means.

composer and conductor. In Experiment 12 we used two experienced amateur players (JM and SS) as well as PZ.

We carefully avoided informing subjects (including PZ and SS, coauthors) in any way about their performance until after the series of experiments in which they participated was complete, and deliberately provided no trial-to-trial feedback.

### D. Caveats

Most of our 12 experiments made use of three subjects, a relatively small number, expecially given instances of inconsistency. We are confident of our major conclusions, however, especially because we found the same trends in more than one experiment. Nonetheless, we suggest caution in generalizing from our findings. We used only two subjects in Experiments 6 and 8, and only one subject in Experiments 7 and 11; results from these experiments should therefore be treated with special caution. Our most stable subject (PZ) served in all 12 experiments, permitting useful comparisons. Some of our findings are clear as well as surprising; nonetheless, they should be regarded as starting points to be confirmed and extended.

### E. Principal Findings

In Experiments 1–5 we employed two variants each of the judgment and production tasks, and one variant of the imitation task. All three procedures resulted in proportional errors that are large (20–50%) for small fractions.

As described above, we define beat fractions (and values of $N$, $n$, $F$, and $f$) in terms of the interval between a tap or marker click and the *previous* beat click: we call these *forward fractions.* It should be noted, however, that a large forward fraction (such as 7/8) corresponds to a small *reverse fraction* (1/8) measured from tap or marker click to the *next* beat. Data from all three procedures hint at systematic errors associated with small reverse fractions that are qualitatively similar to the errors we observe for small (forward) fractions. In this report we emphasize performance for small values of $n$ and $f$, however, because it is more reliable in all three tasks, and we have more data in that region.

Subjects tend to "overestimate" small fractions (Section II). The fraction names $N$ associated with stimulus fractions $f$ are too large: $f = \mathbf{J}(N) < N$. From this result one might expect that the fraction $F = \mathbf{P}(n)$ produced in response to a stimulus name $n$ would be too small. Instead, produced fractions are too large: $F = \mathbf{P}(n) > n$ (Section III). For example, though a stimulus fraction had to be shorter than 1/8 to be called "1/8," when subjects tried to produce a fraction of "1/8" they produced an interval greater than 1/8.

This inconsistency between judgment and production performance for small fractions requires us to reject feedback models of production (Sections III,B and III,D), in which produced fractions are adjusted by judging them.

Imitation performance (Section IV) is very similar to production performance: for

small fractions, $F = \mathbf{I}(f) > f$. The existence of systematic errors in imitation argues against models in which the same transformations (psychophysical functions) relate stimulus fractions and produced fractions to their internal representations. Quantitative comparison of the imitation error to the judgment and production errors argues against a concatenation model of imitation, in which a fraction name produced by a covert judgment then serves as input to the production process.

Taken together, results from the three tasks suggest an information-flow model containing four processes (Section V) with an input process shared by judgment and imitation, and an output process shared by production and imitation. Nothing quantitative is assumed about the four processes, yet properties of the data permit some surprisingly strong inferences about them. The model is outlined in Fig. 7; readers may find it helpful to examine this figure before reading further.

In Section VI we explore and dismiss three potential sources of errors in judgment: the time to shift attention from beat to marker (Section VI,A), the possible importance of stimulus offsets ("releases") as well as onsets ("attacks") (Section VI,B), and the possibility that the rate at which subjective time elapses varies with location within the beat interval (Section VI,C).

We report evidence of a special difficulty associated with concurrent time judgments (Section VI,D), and by varying the beat interval we demonstrate that the judgment error can be described neither in terms of the fraction $f$ alone or the fractional interval $bf$ alone (Section VI,E).

In Section VII we explore and dismiss five potential sources of the errors in production: the use of finger-tap responses rather than notes played on musical instruments (Section VII,A), the absence of adequate response feedback (Section VII,B), the possibility of a distortion of subjective time near the beat (Section VII,C), and the use of single isolated responses that do not fill the beat interval and of off-beat responses not accompanied by any on-beat response (Section VII,C). We also note a tendency for errors in production to be accompanied by displacement (phase shift) of the subjective beat.

Details of experimental method and analysis are given in five appendices. We recommend that readers not interested in technical details omit these appendices, as well as the footnotes and Sections II,C, III,F, III,G, and IV,D.

## II. PERCEPTUAL JUDGMENT OF BEAT FRACTIONS

Our principal aim in Experiments 1 and 2 was to determine the stimulus fractions $f$ that were judged to be equivalent to various fraction names $N$; we would thereby have a psychophysical scale $f = \mathbf{J}(N)$ for fractions of a beat. (Note that in this chapter the term "scale" never denotes a musical scale.) A secondary aim was to measure the *precision* of expert judgments of beat fractions—the sensitivity of judgment probabilities to changes in $f$. We explored two different methods that permitted us to determine, for each of a set of fraction names, the stimulus fraction $f$ that was subjectively equivalent to it; values of $N$ ranged from 1/8 of a beat to 1 (a full beat). In

Experiment 1 (single-fraction judgment), a fraction name $N$ was specified and the subject then judged, for each of a set of stimulus fractions, whether it was larger or smaller than $N$. In Experiment 2 (multiple-fraction judgment), the subject selected a response from a set of eight categories (such as "between 1/8 and 1/7 of a beat") whose boundaries were defined by fraction names. Our use of both single- and multiple-fraction procedures was motivated partly by a desire to assess the invariance over experimental methods of the systematic perceptual errors we discovered. Other differences between the procedures are discussed below.

### A. Single-Fraction Perceptual Judgment (Experiment 1)

The stimulus patterns in Experiments 1, 3, and 5 are represented on the left side of Fig. 1; the pattern of beat clicks was held constant across these experiments to minimize stimulus differences among the three procedures. Two preliminary beat clicks were followed by a pause ("rest") of one beat (symbolized by a broken line) and then by two more beat clicks. (We used the pause so as to separate the stimulus and response components of each trial in the imitation task.) On each trial in Experiment 1 (Fig. 1A) the final beat click was followed by a marker click. Subjects judged whether the beat fraction appeared too large or too small relative to a specified fraction name $N$.[9]

Subjects judged fractions in relation to the fraction names 1/8, 1/6, 1/4, 1/2, 3/4, 5/6, 7/8, and 1. The name stayed the same for 75 consecutive trials as the stimulus fraction was varied by an "up-and-down" or "staircase" procedure. (The effect of the staircase procedure is to concentrate the stimuli close to the fraction that is judged to be neither too large nor too small relative to the specified name—i.e., subjectively equivalent to it. See Appendix A for more information about our use of this procedure.)

For each fraction name the resulting data permitted us to estimate the stimulus fraction $f$ subjectively equal to it, which we call the "PMF mean". They also provided a measure of judgment variability discussed in Section II,C, which we call the "SD (standard deviation) of the PMF." Readers not interested in details of method need not understand how these estimates are determined. For each fraction name the method starts with the estimated psychometric function (PMF) provided by our data: a function, usually S-shaped, that associates with the value of each stimulus fraction the proportion of trials on which that fraction appeared "too large." The *location* of the PMF on the $f$-axis for a specified fraction name is, roughly, the stimulus value where judgment probabilities change most rapidly as the stimulus fraction is changed. This location separates two intervals on the $f$-axis: a "small-$f$" region where $f$ tends to be judged too small relative to the name $N$ and a "large-$f$" region where $f$ tends to be judged too large. The location therefore corresponds to a fraction $f$ that appears

[9]Subjects actually selected responses from six alternatives, representing three degrees of confidence for "larger" and three for "smaller." For the present report, however, we have pooled responses from each of the two sets of three to produce two response classes.

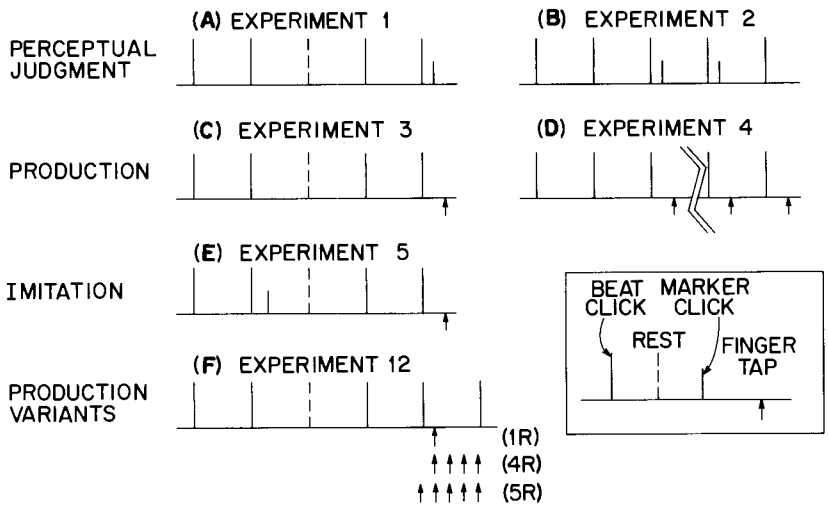



**Fig. 1.** Time-pattern stimuli and finger responses in six experiments. Beat and marker clicks were 5-msec tone bursts of 3000 and 2500 Hz, respectively, presented over headphones at approximately 30 dB above threshold. The tone bursts started at a zero-crossing and were gated by a voltage controlled amplifier so as to have a 1-msec rise time, a 3-msec steady level, and a 1-msec fall time. Judgment responses were made with a multiple response keyboard. Finger-tap responses were made by tapping on a plate and thereby completing an electric circuit. The interval between successive beat clicks was 1 sec in Experiments 1–5 and .5 sec in Experiment 12. Presented fractions were defined by the time interval between beat click and marker click. Produced fractions were defined by the time interval between beat click and tap, corrected for differential delays (see Section III,A). The train of events on each trial in Experiment 4 included ten finger-tap responses.

subjectively equal to $N$. A conventional measure of location is the estimated 50% point, or median, of the PMF (i.e., the $f$-value for which the judgments are equally divided between "smaller than $N$" and "larger than N"). Instead we report the estimated *means* of PMFs as location measures; our preference for the mean over the median—which differ little in these experiments—is explained in Appendix C, together with our estimation method.

A set of such PMF means establishes a *judgment function*, $f = \mathbf{J}(N)$, a psychophysical scale that associates with each fraction name $N$, its subjectively equal fraction.[10] (The inverse function, $N = \mathbf{J}^{-1}(f)$, to be used in Section IV,A, therefore gives the value on the name scale associated with a specified stimulus fraction.)

The results are shown in Table I; column labels give the fraction name and its equivalent fractional interval in msec, and row 1 shows the mean stimulus fraction for the three subjects (SB, PF, and PZ). If judgments were free of systematic error, entries in this row would equal the column headings. Instead, as shown by the

[10]Note that this procedure for establishing a psychophysical function, in which averaging is performed in the *stimulus* domain, differs from more common procedures in magnitude and category scaling in which *responses* are averaged. In Appendix E we discuss a comparison of the two methods applied to data from Experiment 2 (multiple-fraction judgment) in which, unlike Experiment 1, both methods can be applied.

TABLE I

Results from Experiments 1–5 and Six Critical Contrasts[a,b]

| Experiment or contrast | Fraction name $N$, $n$ (in judgment, production) or fraction $f$ (in imitation) (and corresponding fractional interval) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1/8 (125) | 1/7 (143) | 1/6 (167) | 1/5 (200) | 1/4 (250) | 1/3 (333) | 1/2 (500) | 3/4 (750) | 5/6 (833) | 7/8 (875) | 1 (1000) | SE (df) |
| 1. Single-fraction judgment, $\mathbf{J}_1$ | 85.3 (z) | — | 125.5 (fz) | — | 265.2 | — | 455.3 (z) | 787.3 | 835.4 (b) | 883.9 (b) | 977.8 | 23.9 (14) |
| 2. Multiple-fraction judgment, $\mathbf{J}_2$ | 59.3 (*bfz) | 79.7 (*bfz) | 105.4 (*bfz) | 154.4 (*bfz) | 207.3 (*bf) | 303.6 | 451.7 (*bz) | — | — | — | — | 10.1 (12) |
| 3. One-response production, $\mathbf{P}_3$ | 139.1 | — | 197.8 (*b) | — | 259.3 | — | 486.3 | 759.5 | 812.3 | 865.7 (z) | 1023.3 | 11.7 (14) |
| 4. Repeated-response production, $\mathbf{P}_4$ | 156.8 (bz) | 181.4 (bz) | 190.4 (bfz) | — | 256.7 (z) | — | 500.1 | 743.9 | 814.2 (b) | 853.7 (b) | — | 18.3 (14) |
| 5. Imitation, $\mathbf{I}_5$ | 158.2 (*z) | — | 185.2 (z) | — | 254.5 | — | 491.0 | 773.0 | 800.8 (*f) | 853.3 | 986.6 | 14.6 (14) |
| 6. $\mathbf{P}_3 - \mathbf{J}_1$ | 53.8 (z) | — | 72.3 (*fz) | — | −5.9 | — | 31.0 (z) | −27.8 | −23.1 (z) | −18.2 (bz) | 45.5 (z) | 28.2 (14) |
| 7. $\mathbf{P}_4 - \mathbf{J}_2$ | 97.5 (*bfz) | 101.7 (*bfz) | 85.0 (*bfz) | — | 49.4 (*b) | — | 48.4 (*bz) | — | — | — | — | 10.0 ( 8) |
| 8. $\mathbf{P} - \mathbf{J}$ (mean) | 75.7 (*bz) | — | 78.7 (*bz) | — | 21.8 (bz) | — | 39.7 (*bz) | — | — | — | — | 10.0 ( 6) |
| 9. $\mathbf{J}_1^{-1} - \mathbf{I}_5$ (Alternative 2)[c] | 34.2 (f) | — | 17.0 | — | 11.0 (z) | — | 43.2 (z) | −57.3 (fz) | 7.3 (fz) | −4.4 (z) | — | 26.0 ( 6) |
| 10. $\mathbf{P}_3 - \mathbf{I}_5$ (Alternative 3) | −19.1 | — | 12.6 | — | 4.8 | — | −4.7 | −13.5 | 11.5 (z) | 12.4 (z) | 33.7 (b) | 17.2 (14) |
| 11. $\mathbf{P}_3\mathbf{J}_1^{-1} - \mathbf{I}_5$ (Alternative 4)[c] | 52.5 (f) | — | 29.6 | — | 10.8 | — | 36.0 (z) | −62.0 (fz) | −1.7 (fz) | 35.9 (z) | — | 29.3 ( 6) |

[a]Entries in rows 1 and 2 (3 and 4) are mean presented (produced) fractional intervals associated with the fraction names specified by column headings. Entries in row 5 are mean produced fractional intervals associated with presented fractions specified by column headings. Entries in rows 6–11 are the mean differences indicated. All entries are in msec.

[b]The letters b, f, z are the last initials of our three principal subjects SB, PF, and PZ. A letter is placed next to an entry if the corresponding subject's value of that entry differs significantly ($p < .05$) from the column heading (rows 1–5) or from zero (rows 6–11); the tests for individual subjects were based on the between-replications variance pooled over fractions. The asterisk indicates that the mean over subjects is significantly ($p < .05$) different from the column heading (rows 1–5) or from zero (rows 6–11). Each condition in Experiments 1, 3, and 5 had two replications of the procedure; each condition in Experiments 2 and 4 had from one to three replications. Questions about characteristics of the "population" of subjects from which our "sample" of three was drawn depend on tests of the mean cell entries over subjects. For Experiments 1, 3, and 5 such tests were based on fractions × subjects × replications analyses of variance in which replications was regarded as a fixed effect. For Experiments 2 and 4, results from different replications were averaged and subjected to fractions × subjects analyses of variance. SE estimates used for t-tests of cell entries are based on subjects × fractions interaction mean squares, whose *df*s are also indicated.

[c]Data from subject SB were excluded from the means in rows 9 and 11 because the between-replications variances for these contrasts are greater than those of PF and PZ by a factor of about 19. (This exclusion has the effect of producing means that are heavily weighted by degree of precision.) None of SBs values of the contrasts associated with alternatives 2 and 4 differed significantly from zero.

Biblioteca del Real Conservatorio Superior de Música - MADRID

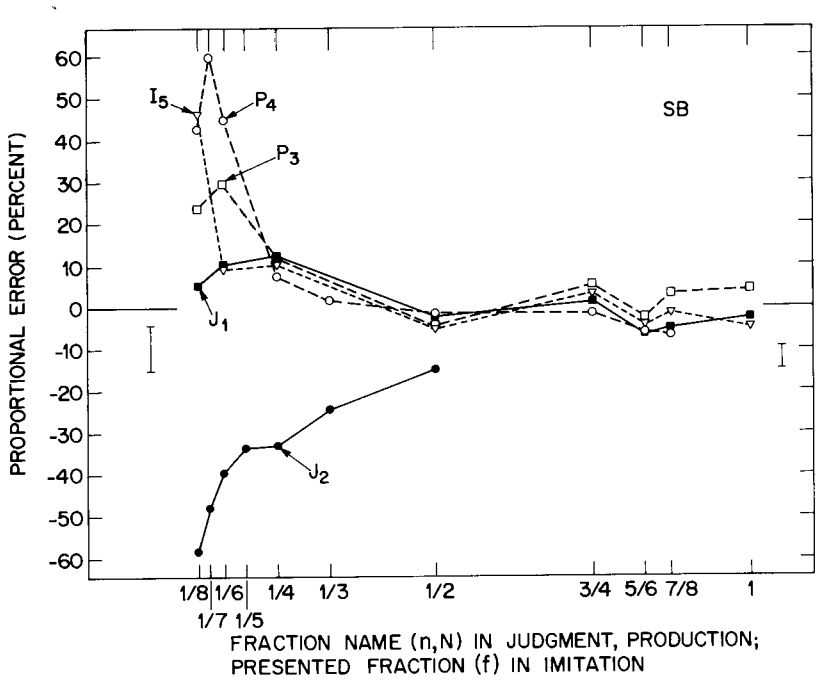



**Fig. 2.** Results for subject SB from five judgment (J), production (P), and imitation (I) experiments, Ordinate values denote signed proportional error: $(f-N)/N$ for judgment, $(F-n)/n$ for production, and $(F-f)/f$ for imitation, in percentage units. Corresponding abscissa values are $N$, $n$, and $f$, respectively, expressed as fractions. Subscripts are experiment numbers.

accompanying letters, six of the 24 tests of individual data indicated significant departures from equality. Judgment variability tends to increase with fraction size. (See Section II,C, and Getty, 1975, for examples.) This is one reason for our representing the data for individual subjects in Figs. 2, 3, and 4 as signed proportional error (in percent) versus $N$-value. Our second reason is the importance in music of time ratios and of the rates at which notes occur. The value of $\mathbf{J}(1/8)$ from a brief session with our fourth subject, PB, is included in Fig. 4.[11]

A numerical example of the proportional error measure for $N = 1/8$ may be helpful.

[11] As noted in Section I,E, a large fraction (such as 7/8) defined from the *previous* beat (as described by the instructions to subjects) corresponds to a small fraction (1/8) measured relative to the *next* beat. If such small reverse fractions were overestimated we would expect that just as we tend to find $\mathbf{J}(N) < N$ for $N < 1/2$, so we would find the symmetric relation $[1 - \mathbf{J}(N)] < 1 - N$, equivalent to $\mathbf{J}(N) > N$, for $N > 1/2$. For PZ and PF both relations tend to obtain, indicating symmetry, although the effect for large fractions is not significant; for SB both relations tend to be reversed, also indicating symmetry. Since the beat following the response was not represented by a click in this procedure, as was the previous beat, any such symmetry suggests that the beat click itself may not be an important determinant of performance and that there is indeed an internal event associated with the final beat. One deficiency of our proportional error plots, of course, is that they obscure systematic irregularities associated with large fractions.

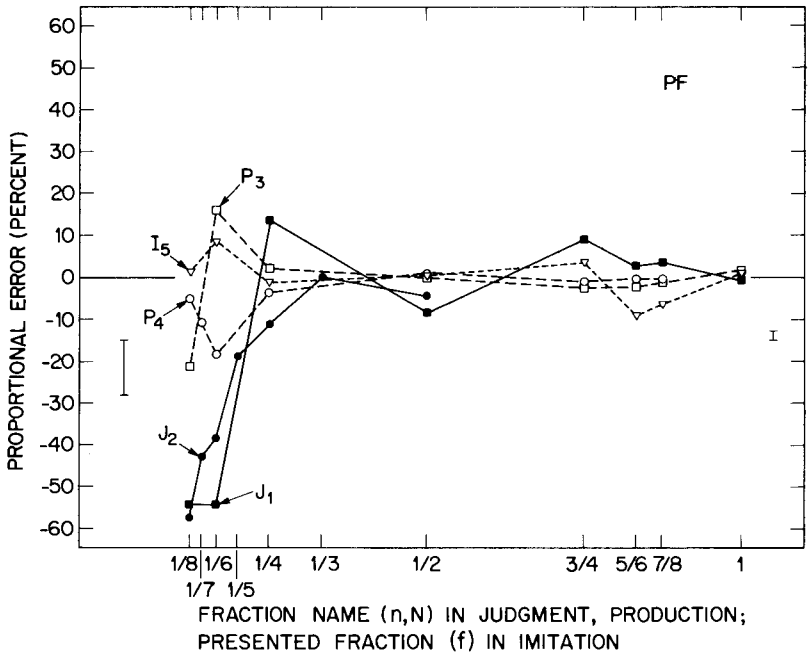



**Fig. 3.** Results for subject PF from five experiments. See caption of Fig. 2.

PB's judgment data implied that relative to a beat interval of 1 sec, 62 msec was subjectively equal to 1/8 of a beat (125 msec): $\mathbf{J}(1/8) = 62$. Since $[\mathbf{J}(1/8) - 125]/125 = (62-125)/125 = -.50$, there was a $-50\%$ error: the fraction that was judged subjectively equal to 1/8 of a beat was 50% too small. Put another way, the name (1/8) assigned to $f = 62$ was 100% too large: an instance of surprisingly radical overestimation.

Figure 5 shows a different representation of the data from PZ, the most stable of our three principal subjects. Here, $\ln(bf)$ is plotted against $\ln(N)$. If $f = \mathbf{J}(N)$ were a power function, the judgment data in this figure could be well fitted by a straight line. (The slope of such a line is the exponent of the corresponding power function.) Clearly, no one power function can describe these data; if separate linear segments were fitted to small-$N$ and large-$N$ ranges, exponents would be about 1.81 for small fractions ($N \leq 1/4$), and close to unity for larger fractions.[12] We defer further discussion of the judgment data to the next section.

[12]Note that the more conventional judgment function derived from "magnitude estimation" procedures, in which the experimenter specifies $f$ and the subject provides $N$, would give an average $N$ as a function of $f$; $N = \mathrm{M}(f)$. M can be regarded as the inverse of **J**; if they were power functions, their exponents would be reciprocals. (The exponent of M for small fractions is about .55 for the data in Fig. 5.) The change in the exponent of **J** as $n$ is increased therefore conforms approximately to the findings for magnitude scaling of subjective duration reported by Michon (1967) and Svenson (1973), mentioned in Section I.

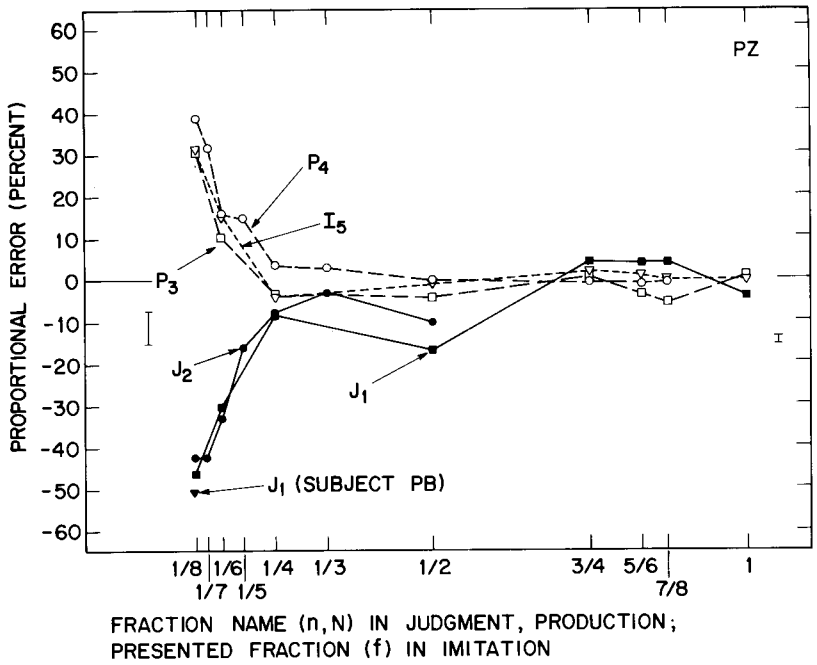



**Fig. 4.** Results for subject PZ from five experiments and one datum from subject PB. See caption of Fig. 2.

## B. Multiple-Fraction Perceptual Judgment (Experiment 2)

The stimulus pattern used in Experiment 2 is represented in Fig. 1B. On each trial the subject heard five beat clicks, with marker clicks following the third and fourth; subjects therefore had two opportunities to observe the beat-marker interval before each judgment. This interval was varied from trial to trial over a wide range (from a minimum of 43 msec to a maximum of 891 msec) by a constant-stimulus method (see Appendix A). The subject selected a response from a set of eight categories, each denoting a range of fraction names and bounded by "simple" fractions (involving small integers): "less than 1/8 of a beat," "between 1/8 and 1/7," "between 1/7 and 1/6,". . . , "between 1/3 and 1/2," and "greater than 1/2." The eight ordered categories define seven between-category boundaries on a hypothetical response continuum. For each boundary and each stimulus fraction $f$ we determined the proportion of responses in all categories above that boundary. Regarded as a function of $f$, this proportion for a specified boundary defines an estimated PMF; this procedure produces seven such PMFs. (Consider, for example, the boundary $N = 1/7$ between the second and third category. Responses to a fraction $f$ in categories above that boundary—categories 3 through 8—are all associated with judgments that $f$ appears

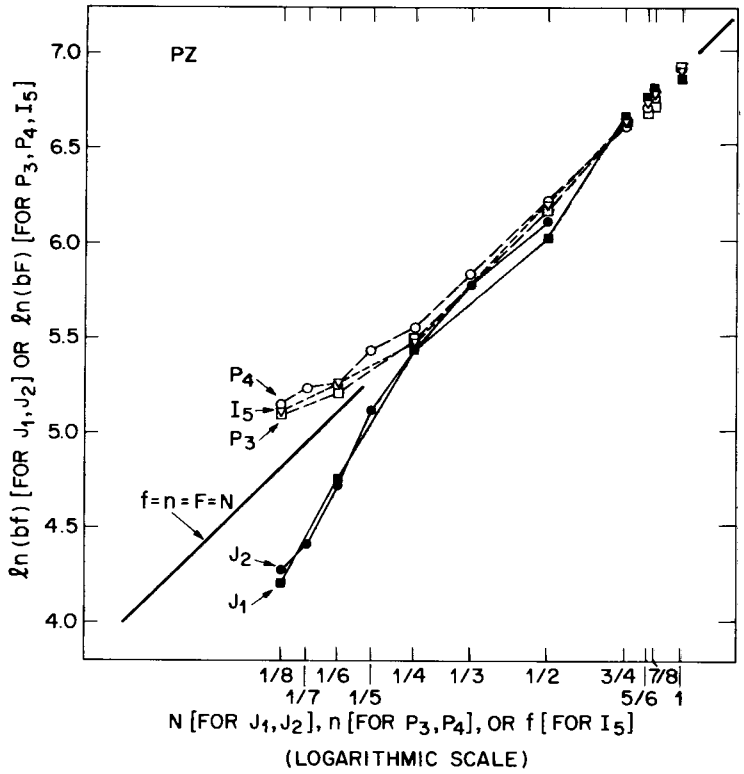



**Fig. 5.** Results for subject PZ from five judgment (**J**), production (**P**), and imitation (**I**) experiments. Ordinate values are natural logarithms of the fractional intervals *bf* (for judgment) and *bF* (for production and imitation) in msec. Abscissa values represent *N* (for judgment), *n* (for production), and *f* (for imitation), also on a logarithmic scale. Subscripts are experiment numbers. To make units the same on abscissa and ordinate, multiply abscissa values by the beat interval (1000 msec). Power functions are represented as straight lines on this kind of graph.

greater than 1/7 of a beat. As $f$ is increased, the proportion of judgments in this "supercategory" increases, defining an estimated PMF associated with $N = 1/7$.)

As in Experiment 1, means of the resulting set of PMFs were used to establish a judgment function, $f = \mathbf{J}(N)$, for each subject.[13] The mean function for three subjects

[13]This judgment function can be regarded as associating an average stimulus value with each of a set of values (category boundaries) on the response continuum. For each partitioning of the eight categories into the pair of "supercategories" defined by a particular boundary, we treated the data in the same way as in Experiment 1 (see Appendix C). Experiment 2 can also be treated by the more conventional procedure in which an average *response* value is associated with each of a set of *stimuli*. To permit such averaging in the present data, the value of a category response could be taken to be the geometric mean of the values of its two boundaries, for example. That the judgment functions from the two procedures are similar is shown in Appendix E, which also explains our preference for the PMF method.

is presented in row 2 of Table I, individual proportional error data are shown in Figs. 2, 3, and 4, and $\ln(f)$ versus $\ln(N)$ for PZ is shown in Fig. 5.

In Experiment 1, PF and PZ (corroborated by PB) showed large and systematic overestimation of small stimulus fractions, while SB did not. In Experiment 2, however, all three principal subjects showed this effect; quantitative agreement between Experiments 1 and 2 was excellent for PZ, good for PF, but poor for SB. Of the 21 tests based on individual data in Experiment 2, 16 indicated significant departures from equality of $N$ and $\mathbf{J}(N)$, and each of the mean differences also proved significant. Taken together, our data show that musicians radically overestimate small fractions of a beat: $N > f = \mathbf{J}(N)$.

Although we cannot explain the anomalous results from SB in Experiment 1, we are more impressed by the consistency of PZs mean data between experiments than by the inconsistency of SBs: PZ is the most experienced musician among our principal subjects, and his data *within* experiments are by far the most consistent. In Experiment 1 each trial included only one presentation of the fraction to be judged; the fraction size was varied by a staircase procedure over a narrow range of equally spaced values, and subjects had to judge stimulus fractions relative to only a single "target" fraction name during a group of trials. In Experiment 2, on the other hand, each trial included two successive presentations of the fraction to be judged; the fraction size was varied over a wide range and over unequally spaced values by a constant stimulus method; and subjects had to judge stimulus fractions relative to an array of seven fraction names (category boundaries). We conclude that the mean values of the criteria that subjects employ in making these perceptual judgments are affected little by either the number of observations per trial, the range of fractions to be judged, or the number of fractions with respect to which each judgment was made. We shall see in the next section, however, that the choice of procedure does influence the precision of these judgments.[14]

### C. Judgment Precision

Insofar as a subject is more sensitive to the stimulus fraction, his judgment probabilities will change more rapidly as the fraction is changed, and the PMF will rise more steeply, or have less *spread*. It is convenient to regard the PMF as a (cumulative) distribution function characterized by a standard deviation (SD) as well as a mean. The SD of the PMF is one measure of its spread, and therefore of the imprecision or variability of judgments.[15]

[14]On each trial in Experiment 1, both $N$ and $f$ were provided to the subject; there is some question whether the value of $N$ associated with the stimulus $f$ by the function $f = \mathbf{J}(N)$ should be regarded as a response to $f$. In Experiment 2, where the subject explicitly selected an interval on a continuum of $N$-values, there is less uncertainty in identifying $N$ as the response. Because of this, together with similarity of the two judgment functions for each of two subjects, we shall regard $f$ as "input" and $N$ as "output" for both experiments.

[15]An alternative and more traditional measure of precision is the difference threshold (DL), which is defined as half of the interquartile range or, roughly, the change in the stimulus fraction required to change

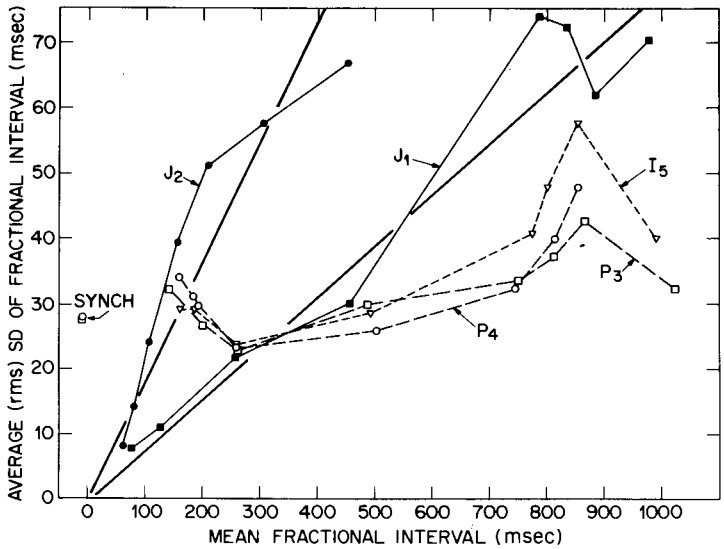



**Fig. 6.** Performance variability in five experiments. Root mean squared SDs (in msec) over replications and subjects. For judgment experiments $\mathbf{J}_1$ and $\mathbf{J}_2$, SDs were estimated from PMFs regarded as cumulative distributions of the fractional interval *bf*. (See Appendix D.) For production ($\mathbf{P}_3$ and $\mathbf{P}_4$) and imitation ($\mathbf{I}_5$) experiments, SDs were calculated from the distributions of the fractional interval *bF*. Also shown is the rms SD from the synchronization condition (Section III,A). In all cases, the SD($x$) of a quantity $x$ is plotted as a function of the mean M($x$) of that quantity. Linear functions passing through the origin—which represent Weber laws SD(*bf*) = $k$M(*bf*) with $k$=.08 and $k$=.19 for judgment Experiments 1 and 2, respectively—were fitted by eye.

Fig. 6 includes average SDs of the PMFs in Experiments 1 and 2.[16] In both experiments the SD increases approximately in proportion to *f*, consistent with a Weber law; proportionality constants are about .08 and .19 for Experiments 1 and 2, respectively. (Sampling error in these data preclude a powerful test of the Weber law, however.) Getty (1975) discusses related data and some of the implications of Weber's law for timing models. Insofar as a fraction has been better learned than others or is easier to "compute" given a beat interval, one might expect it to show greater precision (smaller SD) and therefore fall below the Weber law line; there is no dramatic evidence supporting this conjecture, however. Judgment precision is systematically greater (smaller SDs) in the one-fraction than the multiple-fraction procedure, despite

the proportion of "larger than" judgments from .50 to .75. See Appendix D for a discussion of our preference for the SD and our method of estimating it.

[16]The figure actually shows square roots of mean variances (root mean square SDs). Throughout this paper we have chosen to average variances rather than SDs because different sources of variability (such as fraction and subject, in perceptual judgments, or timing and response mechanisms, in production) are more likely to be additive in variance units. We have shown SD rather than SD$^2$ in the figure, however, because the expression of Weber's law, SD = $kf$, where $k$ is a constant, is then simply a straight line through the origin.

the similarity of judgment means discussed above. (This finding suggests that the range of stimuli $f$ or of fractions $N$ with respect to which judgments are made has a large effect on the variability of subjects' criteria, but only a small effect on their means.)

Despite the systematic errors shown by the judgment function, Fig. 6 reveals the precision of the judgments to be high. For example, in Experiment 2, the $f$-value that was subjectively equal to $N = 1/8$ (125 msec) was 59.3 msec, and the SD was 7.9 msec. This implies that an $f$-value of only 72 msec would be judged "larger than 1/8" on 95% of the trials,even though this $f$-value is 53 msec (or 42%) smaller than the "correct" value.

## III. PRODUCTION OF BEAT FRACTIONS

The systematic errors in perceptual judgment discussed above, which are proportionately very large for small fractions of a beat, make it particularly interesting to examine musicians' accuracy in *producing* brief time intervals defined as beat fractions. We used a method of *timed response*. The subjects' task in Experiments 3 and 4 was to use a finger tap to terminate a time interval that started with a beat click and thereby produce a beat fraction $F$ that corresponded to a specified fraction name $n$. (Subjects could hear as well as feel themselves tapping, since the earphones that delivered the clicks provided negligible attenuation of other sounds. A reader who taps the hard surface of a desk top will hear a "thump" similar to what our subjects heard.)

Our aim was to determine the relation between a set of fraction names and the set of corresponding fractions, and thereby establish a production function, $F = \mathbf{P}(n)$. A secondary aim was to measure the precision of such expert timed responses—the variability of the time intervals they defined. Again we used two methods, to assess the invariance of the systematic timing errors we discovered. One method required a single timed response on each trial; the other required a repeated series of responses corresponding to a fixed-beat fraction, thereby permitting more immediate adjustment to perceptual feedback.

### A. Use of Tap-Click Synchronization to Correct for Differential Subjective Delays

We wished to compare the timing mechanisms used in perceptual judgment with those used in production and imitation, partly to test the idea that they are the same. It is possible, however, that the subjective delays associated with events that mark the ends of the relevant intervals—beat click, marker click, or finger tap—are different. By correcting for any such differences we can examine the timing mechanisms more directly.

This need for correction seems especially acute for finger taps. Even abrupt taps are extended in time, which makes it unclear how to associate a single time point with a

response. Our equipment measured the time at which the finger first contacted a metal plate, but the subjective time of the response might be equally well described by the time when the "command" to make the response is issued, the time when maximum pressure is achieved, the time when the finger breaks contact with the plate, or some other feature of the response, possibly adjusted by perceptual delays.

Suppose that the beginning and end of an interval are marked by events b and e, respectively, $T_b$ and $T_e$ are their physical occurrence times, and $D_b$ and $D_e$ are the delays in registering the events internally. (For a tap, the mean "delay" may be negative.) Then the registration times are $T_b + D_b$ and $T_e + D_e$; and whereas the objective interval between events is $T_e - T_b$, the interval between registration times is $(T_e - T_b) + (D_e - D_b)$. Thus, to correct the measured interval for internal delays, we must estimate the delay difference, $D_e - D_b$.

If the two delays are equal, the difference is zero, and no correction is needed. For the perceptual judgment experiments, in which both ends of the relevant interval were marked by the same class of events (clicks), we felt that equality of delays was a plausible starting assumption. (In Section VI we report some findings favorable to this assumption.)

Our solution to this problem in the production experiments, where the critical interval begins with a beat click and ends with a finger tap, was to measure the difference between the two subjective delays by using a special condition in which subjects were asked to synchronize their responses with beat clicks. This synchronization task can be thought of as a production task with $n=0$, except for the occurrence of a beat click at the time when the response should occur. Since we shall be using $\mathbf{P}'(n)$ to denote the raw (uncorrected) mean production time for fraction $n$, we use $\mathbf{S}'(0)$ to denote the mean measured response "delay," $\mathbf{S}'(0) = T_t - T_c$, where $t$ and $c$ denote tap and click, respectively. Suppose that the subjects succeed in locating the mean of their distributions of subjective occurrence times of responses coincident with the mean of the subjective beat times. The registration-time difference defined above is then zero: $(T_t - T_c) + (D_t - D_c) = 0$. It follows that $-\mathbf{S}'(0) = D_t - D_c = T_c - T_t$ provides the desired estimator of the delay difference.

The raw (uncorrected) mean production times $\mathbf{P}'(n)$ were corrected by subtraction: $\mathbf{P}(n) = \mathbf{P}'(n) - \mathbf{S}'(0)$. In the synchronization conditions, subjects responded slightly before the beat click [$\mathbf{S}'(0)<0$], implying that the subjective delay associated with the tap response was greater than the perceptual delay associated with the beat click. The corrections, then, slightly increase the measured values of $\mathbf{P}(n)$. Application of the synchronization correction depends on the assumption that the difference between these beat-click and tap delays in the production task is the same as in the synchronization task and is independent of the interval between beat click and tap (which in turn depends, in production, on the specified fraction name).[17]

[17]One objection to the use of the synchronization correction is based on the possibility that this assumption will be violated. For example, constraints on attending simultaneously or in close succession to click and tap might produce special differential delays ("prior-entry" effects). See Sternberg and Knoll (1973), Section VI.

## B. Expectations from a Feedback Model of Production

What relation might we expect between performances in the judgment and production tasks? One appealing hypothesis is a simple feedback process in which a subject (1) judges the size of each produced fraction $F$ with respect to the target-fraction name $n$ using the same perceptual mechanisms in this judgment as in the judgment task itself, and (2) adjusts subsequent productions accordingly.[18] One would then expect systematic errors in judgment and production tasks to be equal to that $\mathbf{P}(n) = \mathbf{J}(N)$ when $n = N$. For example, if an interval of 62 msec is *judged* to correspond to 1/8 of a 1-sec beat [$\mathbf{J}(1/8) = 62$], the same interval (after correction for differential subjective delays) should be *produced* for 1/8 of a beat [$\mathbf{P}(1/8) = 62$].

## C. One-Response Production (Experiment 3)

On each trial in Experiment 3 (Fig. 1C), the subject attempted to respond with a single finger tap following the last beat click so as to produce a fraction $F$ that corresponded to a specified fraction name $n$. The name remained the same for 25 consecutive trials. We used the same fraction names as used in Experiment 1 to define a set of experimental conditions. For each condition we calculated a mean raw response time, $\mathbf{P}'(n)$.

In a synchronization condition, the subject attempted to synchronize the finger tap with a beat click added at the end of the normal stimulus pattern. The resulting values of $\mathbf{S}'(0)$ were $-13.8$ msec, $-16.3$ msec, and $-9.0$ msec, for subjects SB, PF, and PZ, respectively.

The mean production function, $\mathbf{P}(n) = \mathbf{P}'(n) - \mathbf{S}'(0)$ for the three subjects is presented in row 3 of Table I, individual proportional error data are shown in Figs. 2, 3, and 4, and $\ln(F)$ versus $\ln(n)$ for PZ is shown in Fig. 5. PZ and SB show large positive proportional errors (overproduction) for fractions 1/8 and 1/6. Because PF does not show this effect, the error shown by the mean production time for small fractions is somewhat smaller; for only one of the eight means (for $n = 1/6$) is the inaccuracy significant. Proportional errors for larger fractions are small and of varying sign, with a slight tendency for produced intervals associated with the largest fractions (5/6, 7/8) to be too small (underproduction).[19]

[18]Carlson and Feinberg (1968) and Adam, Castro, and Clark (1974) present data that favor an "internal clock" or counter model that applies to both judgment and production of time intervals in a range from 1 to 40 seconds. According to Carlson and Feinberg's model, systematic errors in production and judgment result from changes in the counter's rate between learning the count to be associated with a time interval specified by name and performing the experimental tasks. The resulting relation between errors in production and judgment is consistent with a feedback model. Furthermore, because the counter rate is assumed not to change systematically between stimulus and response in an imitation (reproduction) task (Section IV), this kind of model requires imitation to be accurate. Given that fractions are being judged in our experiments, however, and the beat interval (which can "calibrate" the counting rate) is presented on each trial, the counter mechanism does not easily lend itself to explaining any systematic errors in our tasks.

[19]This tendency can also be described as overproduction of the small "reverse fractions" 1/6 and 1/8 that are defined by the intervals between tap and subsequent beat. A tendency toward such symmetry in

### D. Rejection of the Feedback Model

All three subjects show large discrepancies between $\mathbf{P}(n)$ and $\mathbf{J}(N)$ for the two smallest fractions. In row 6 of Table I are shown values of the difference $\mathbf{P}(n) - \mathbf{J}(N)$ for the matched procedures of Experiments 1 and 3. The feedback model requires this difference to be zero for all values of $n=N$; for the two smallest fractions it is substantially greater than zero, especially in relation to the precision (Fig. 6) of perceptual judgments. The two largest fractions show smaller but directionally symmetric discrepancies.[20]

A numerical example based on mean data from Experiments 1 and 3 (rows 1 and 3 of Table I) may help to clarify the failure of the feedback model. The mean interval *produced* to correspond to $n = 1/8$ was $\mathbf{P}(1/8) = 139.1$ msec. To what fraction name $N$ would this interval correspond if the same perceptual judgment mechanisms were used as in Experiment 1? Since $\mathbf{J}(1/8) = 85.3$ msec and $\mathbf{J}(1/6) = 125.5$ msec, $N$ is clearly greater than 1/6; since $\mathbf{J}(1/4) = 265.2$ msec, it is clearly less than 1/4. To determine $N$ exactly, we must solve $\mathbf{J}(N) = 139.1$ msec for $N$. If we regard fraction names as lying on a continuum, and use linear interpolation, we find the corresponding fractional interval to be 174.8 msec or slightly greater than 1/6 of a beat. Thus, if subjects perceived intervals between click and tap in the same way as intervals between two clicks, the average subject would *perceive* the interval he or she *produced* for 1/8 as larger than 1/6.

Differences between judgment and production procedures could, in principle, result from "constant errors" based on differential subjective delays for which we have not corrected. (We have not attempted to estimate or correct for any delay difference between beat click and marker click in the judgment task, partly because of findings discussed in Section VI. The synchronization-based correction that we did apply to the production data might be inappropriate.) But on the most straightforward view, the absolute effects of differential delays should be of the same size for all fraction values. Hence, insofar as the inconsistency between perception and production depends on fraction size (which it does reliably, as shown below), we have to search elsewhere for an explanation.

One possibility is that although discrepant feedback was available, the time interval between successive trials and the number of stimulus events between one response and the next prevented subjects from making appropriate corrections. Because our subjects said neither that they were dissatisfied with their productions nor that they "came in late," this possibility seems unlikely. Nonetheless, it was tested in Experiment 4, which called for timed responses to 10 successive beat clicks on each trial.

---

judgment performance was noted in footnote 10. Note, however, that when the *precision* of performance (Fig. 6) is considered, rather than its *mean*, performance at the two ends of the beat interval appears far from symmetric.

[20] $\mathbf{J}(1)$ and $\mathbf{P}(1)$ can be regarded as alternative measures of the *subjective* beat interval. That they differ (significantly for PZ) suggests that this interval might be task dependent (also suggested by findings in Experiment 12). Deviations of $\mathbf{P}(1)$ and $\mathbf{J}(1)$ from 1000 msec are sufficiently small, however, so that our use of the actual rather than subjective beat interval to define beat fractions makes little difference, especially for small fractions.

### E. Repeated-Response Production (Experiment 4)

On each trial in Experiment 4, we presented 12 beat clicks. As shown in Fig. 1D, subjects attempted to make 10 consecutive finger-tap productions so as to produce a fraction that had been specified by name. The first response was produced after the third beat click. We used the fraction names 1/8, 1/7, 1/6, 1/4, 1/2, 3/4, 5/6, and 7/8; the name remained the same for 25 consecutive trials. For each fraction name $n$ and each position $k$ in the sequence of 10 responses, we determined the mean raw response time, $\mathbf{P}_k'(n)$.

In a synchronization condition, subjects attempted to make tap responses that coincided with each of the last 10 beat clicks, thereby generating values of $\mathbf{S}_k'(0)$ for $k = 1,\ldots, 10$. The corrected response time was obtained by subtraction: $\mathbf{P}_k(n) = \mathbf{P}_k'(n) - \mathbf{S}_k'(0)$.

#### *Trend over Repeated Responses*

Subjects produced large positive errors ("too late") for the smallest three fractions (1/8, 1/7, 1/6). Values of $\mathbf{P}_k(n)$ averaged over these fractions and the three subjects, for $k = 1,\ldots, 10$, were 168.4, 172.9, 176.9, 180.0, 175.4, 181.4, 176.3, 184.5, 176.6, and 169.8 msec, respectively. There is no evidence of a decrease in the size of the error over repetitions. (The slope of a line fitted to these values is .43 msec/repetition.) It is also instructive to compare the time of the first response $\mathbf{P}_1(n)$ (with synchronization correction $\mathbf{S}_1'(0) = -13.4$ msec) with the time of the mean response $\mathbf{P}(n)$ (with $\mathbf{S}'(0) = -24.9$ msec), separately for the three fractions. For 1/8, 1/7, and 1/6, values for the first response (mean response) are, respectively, 151.9 (156.8), 169.5 (181.4), and 183.9 (190.4) msec.

There is no evidence that subjects used feedback from one response to the next in the repeated response procedure to reduce the size of their production errors.

#### *Means over Repeated Responses*

Further analyses were based on mean response times over the 10 positions. These were adjusted by the mean synchronization corrections; $\mathbf{S}'(0)$ was $-22.6$, $-35.2$, and $-16.9$ msec for SB, PF, and PZ, respectively.

The mean of the production functions, $F = \mathbf{P}(n)$, for the three subjects is presented in row 4 of Table I, individual proportional error data are shown in Figs. 2, 3, and 4, and ln ($F$) versus ln($n$) for PZ is shown in Fig. 5.[21] As in Experiment 3, PZ and SB show large positive proportional errors for small fractions, whereas PF does not; in addition, PZ and SB show negative errors for large fractions, whereas PF does not, providing further evidence for the symmetry in performance at the two ends of the

[21]Like the judgment function, ln($F$) versus ln($n$) cannot be fitted by a single power function (straight line). A two-limb fit gives an exponent of about .57 for the range $1/8 \leq n \leq 1/4$ and about 1.0 for $n > 1/4$.

beat interval mentioned in Sections II,A and III,C. Agreement between Experiments 3 and 4 is best for PZ. Again, performance for large fractions is relatively accurate.

Values of the difference $\mathbf{P}(n)-\mathbf{J}(N)$ appropriate for another test of the feedback model are provided in row 7 of Table I. We used judgment data from Experiment 2 because it provides values that are independent of Experiment 1, because it was run in close temporal proximity to Experiment 4, and because the stimulus conditions (Figs. 1B and 1D) were similar. All three subjects show large discrepancies between judgment and production of small fractions. Each of the five mean values of the difference in row 7 of Table I (values for all the fractions $n = N$ that were common to the two experiments) is significantly and substantially greater than zero; the difference diminishes reliably as fraction size increases. (This implies, of course, that the proportional difference decreases to an even greater extent.) The failure of the feedback model is even more dramatic for these data than for the initial test (row 6 of Table I).[22]

Our best estimates of production-perception discrepancies are obtained by combining the two tests; means for the fraction values that are common to the four experiments are given in row 8 of Table I. The mean contrast is significantly greater than zero ($t_7 = 4.41$; $p < .005$), and the difference depends significantly on fraction size ($F_{3,6} = 7.56$; $p < .025$).

### F. Production Precision: Evidence against a Reaction-Time Explanation of the Production Error

Average SDs of production and synchronization times are included in Fig. 6 for Experiments 3 and 4. Data from the two experiments are in good agreement; data from each of the three subjects averaged over the two experiments produced a U-shaped function with a minimum SD (maximum precision) between $n = 1/4$ and $n = 1/2$.[23] The precision of synchronization performance, with mean $\mathbf{S}'(0)$, is slightly but not reliably greater than the precision of production of a whole beat, with mean $\mathbf{P}'(1)$.

For fractions $n \geq 1/4$, production precision exceeds judgment precision, but this

[22]It remains a puzzle how production performance might be calibrated (or learned) by players without feedback about the timing of individual responses. One possibility is that the association of produced fractions with fraction names is learned through attempts to produce extended sequences of temporally regular notes that fully occupy the interval between one beat and the next. Counting, rather than the timing of individual responses, could then provide a measure by which to adjust the rate. Given this possibility (and also to assess the generality of our findings), it is interesting to ask whether the first response by a subject attempting to produce a temporally regular sequence displays the same pattern of errors as that obtained with a single response. This question is considered in Section VII,C, in a study of multiple divisions of the beat.

[23]A U-shaped function with a minimum in this range has also been obtained from one subject in a time production task by A. B. Kristofferson (1976; Experiment II). Experimental conditions differed considerably from ours, with feedback provided on each trial, and the one subject had about 10 times as much practice in the experimental task as ours did. Nonetheless, his minimum SD (about 12.7 msec) did not differ significantly from the minimum (22.2 msec) of the mean SD produced by our subjects.

Biblioteca del Real Conservatorio Superior de Música

relation is reversed for small fractions. Insofar as subjects do not experience their productions of small fractions as highly variable, we therefore have further evidence against the feedback model.

One explanation of the production errors for small fractions (but not the similar errors for small "reverse fractions") attributes them to a combination of musical training and the existence of a minimum reaction time (RT). The minimum voluntary RT to auditory stimuli is between 100 and 150 msec. Furthermore, there are delays (which differ across instruments) between excitation and acoustic response. The combination of these two effects makes it virtually impossible to produce a note 125 msec after a signal to respond such as a beat, when the event that triggers the response is the beat itself. Players could try to "anticipate" the beat (by timing their responses from the penultimate beat), but since the occurrence of the beat in musical performance is variable, this might be risky. Players therefore appear to be in a situation in which they are musically required to produce discriminably different response delays, some of which are less than the minimum RT.

One possible solution would be to time responses from the final beat, but bias the productions so that the intervals for small fractions are both greater than the minimum RT and distinctive. For example, if to respond later than the minimum reaction time a subject produces an interval of 150 msec (rather than 125 msec) for 1/8 of a beat, then an interval greater than 150 msec must be produced for 1/7 of a beat, etc.

However, the variability of the production-time distributions suggests that responses associated with small fractions may be timed from the penultimate beat. (Synchronization responses presumably *must* be timed from this beat.) Hence, the production errors may *not* be due to a constraint imposed by a minimum RT. The argument (whose impetus and conceptual framework is provided by Snodgrass, Luce, and Galanter, 1967) is as follows.

We start by assuming that the variance of a distribution of response delays increases monotonically (or at least does not decrease) as the mean of the distribution increases, where the mean is measured relative to the reference signal from which subjects time their responses.[24] Therefore, if all responses were timed from the final beat click, we would expect the variance to increase monotonically with fraction size. As we have seen, this expectation is violated by our data. One interpretation of the increased variability for small fractions is that subjects were timing their responses from the penultimate beat click in these conditions. The most salient errors in production are then not the result of a constraint imposed by the minimum RT.

If subjects timed small fractions from the penultimate beat click but large fractions from the final beat click, the argument above, in its simplest form, implies that no large-$n$ productions should have variances greater than the small-$n$ productions. One difficulty is the suggestion in the data (Fig. 6) of a peak in the SD-function when $n$ = 7/8. (Because this difference between the SD at $n$ = 7/8 and $n$ = 1/8 is shown only by

[24]See, e.g., Snodgrass, Luce, and Galanter (1967) for intervals $\geq$ .6 sec, Treisman (1963) for intervals $\geq$ .25 sec, and references cited therein.

SB and PZ, the mean difference is not reliable, however.) A slight elaboration of our account can deal with this difficulty. Suppose that the response for small fractions can be triggered either by perception of the final beat click *or* by a timing process initiated by the penultimate beat click, whichever occurs first. This could shorten production delays that would otherwise be exceptionally long, thereby reducing the variance. (See Kornblum, 1973.) Evidence favorable to such a facilitation effect of the final beat click is provided by a comparison we made with PZ between one-response production of $n = 1/8$ and the same procedure with the final beat click replaced by a "rest" (or an "imaginary beat"). Omission of the click increased the mean response delay by about 17% (from 149.5 to 175.5 msec), thereby increasing the mean error, but more than doubled the SD (from 26.0 to 70.2 msec).[25]

However, our argument for the idea that subjects timed their productions of small fractions from the penultimate beat click depends on the assumption, introduced above, that the variance of a produced time interval cannot decrease as its mean increases. Even if true for time intervals defined in isolation, this assumption may be false for intervals that are defined as different fractions (or multiples) of a standard interval, as in our production task. Perhaps fractions that are "simpler," or more practiced, or that require less "computation" (such as $n = 1$) are produced more reliably. In models in which timing is accomplished by counting a stream of internal events until a criterion is reached, the reliability with which the criterion count is *set* may have to be considered as well as the variability of the inter-event intervals. This possibility is supported in our data by the (statistically significant) reduction in variability from **P**(7/8) to **P**(1).

## G. Implications of Other Analyses of Psychophysical Scaling for the Production–Perception Disparity

Readers with a special interest in psychophysical scaling may find it interesting to consider our perception and production experiments in relation to the "magnitude estimation" and "magnitude production" methods used to investigate many perceptual domains. These methods often produce power-function relations between stimulus and response values; such psychophysical scales are often summarized by power-function exponents. The exponent $\beta$ determines how ratios of stimuli ($\phi_1 < \phi_2$) are mapped onto ratios of the numerical magnitudes or names ($\psi_1 < \psi_2$) associated with them: $\psi_2/\psi_1 = (\phi_2/\phi_1)^\beta$. An exponent larger (smaller) than 1.0 implies that the name ratio is larger (smaller) than the stimulus ratio.

Partly because there is a "correct" relation between $n$ and $F$ that musicians are presumably trained to achieve, we are interested in the relation between $n$-values and $F$-values and not merely in the relation between *ratios* of pairs of $n$-values and pairs of $F$-values. (The latter relation, but not the former, is captured by the exponent of a fitted power function.) Sizes of the exponents are nonetheless useful to consider. We

[25]This comparison also suggests that the existence of a positive production error for small fractions does not depend on the presence of an actual beat click.

have already noted that both $\mathbf{J}(N)$ and $\mathbf{P}(n)$ deviate dramatically from power functions if the full range of fractions is considered. However, power functions fitted only to the data for small fractions fit reasonably well and do capture one aspect of the discrepancy between perception and production. A power function fitted to the data for small fractions in our experiments has an exponent greater than 1.0 for production but less than 1.0 for perception.

Under some conditions with other perceptual continua, magnitude production exponents are larger than those obtained in magnitude estimation. There are at least three reasons why our finding may not be an instance of the same phenomenon. First, Teghtsoonian and Teghtsoonian (1978) have shown that the difference between exponents depends on the stimulus range and, indeed, is reversed for narrow ranges. Stimulus ranges in our two perception experiments differed greatly; their endpoints were about 50 and 850 msec in Experiment 2, but only about 90 and 140 msec (for PZ and $n = 1/6$), for example, in Experiment 1. Nonetheless, we obtained good agreement between experiments for PZ and PF. In both production experiments, the stimulus (fraction name) was fixed for a long series of responses (25 and 250 responses in Experiments 3 and 4, respectively). The best description to assign to the stimulus range in this case would therefore appear to be "narrow."

Second, the Teghtsoonians argued that the effect depends on the avoidance by subjects of extreme response ratios (i.e., either much larger or much smaller than unity). This analysis seems inapplicable in a straightforward way to Experiment 1, where the overt responses were "larger than $N$" and "smaller than $N$." (To make it applicable, one could assume that subjects produce covert responses of particular $N$-values that they then categorize in terms of the specified target $N$-value to determine the overt response.)

Third, if mechanisms of the kind discussed by the Teghtsoonians were responsible for the difference we observed between $\mathbf{J}(N)$ and $\mathbf{P}(n)$ for small fractions, we would expect a difference in the same direction for large fractions, contrary to what we observed.

In the discussion above we have assumed that time *ratios* $F$ or $f$ (between produced or stimulus intervals and the beat interval) are the objects to be produced or judged. A less obvious alternative is to consider these objects to be time *intervals* $bF$ or $bf$. In that case, a procedural difference between our experiments and many others becomes important. The Teghtsoonians argue persuasively that in choosing a response on one continuum to associate with a stimulus on another, subjects refer to the prior stimulus and prior response and choose a response that generates a response ratio equal to the subjective stimulus ratio. With traditional methods the prior stimulus and response are those from the previous trial and usually vary from one trial to the next. In contrast, with our methods the (large) beat interval $b$—corresponding to fixed prior values ("standards") $f = 1$ and $N = 1$ on the two continua—is presented on each trial, becomes a prior stimulus, and can perhaps be regarded as generating a prior response. Suppose that we accept this alternative analysis (together with the idea that subjects produce covert $N$-responses in Experiment 1). Then, if subjects tend to avoid extreme response ratios, they would both overproduce and overestimate small fractions but

not large ones, as observed. The Teghtsoonians' analysis can therefore provide one viable account of the perception-production difference, if we combine it with rejection of the feedback model.[26] [27]

Two findings incline us against the interval alternative. The first, by Michon (1967), is the absence of effects of stimulus range in an experiment on magnitude estimation of time intervals. Given this finding, one would have to argue that the psychophysics of beat fractions and of time intervals differ, with the former more like other sensory domains, or that although the standard interval (like a beat interval) was not presented on each trial in Michon's experiment, subjects presented a fixed standard to themselves.

The second finding is our own, emerging from a comparison of perceptual judgments of beat fractions across different beat intervals (described in Section VI,E). This experiment permitted us to remove the confounding of the duration of the interval being judged with its fraction value. A straightforward application of the interval alternative, based on avoidance of extreme response ratios, implies that the judgment error depends only on the response (the $N$ value) and therefore on the fraction $f$ rather than the interval $bf$; in contrast, we found that the error depends on both fraction value and interval duration.

## IV. IMITATION OF BEAT FRACTIONS

In both the judgment and production tasks, subjects must associate beat fractions and their names. In the imitation task (sometimes called the method of reproduction), this association is not called for, at least not explicitly: the "input" is the same stimulus $f$, as in the judgment task, and the "output" is a timed response $F$, as in the production task. The relation, $F = \mathbf{I}(f)$, between $f$ and $F$ in the imitation task may therefore tell us whether the systematic errors found in the other tasks depend on the requirement to associate names with beat fractions. More generally, by exploring the imitation task and its relation to the judgment and production tasks, we hoped to explain or describe the production-perception disparity in terms of characteristics of internal transformations, some of which may be shared by pairs of the three tasks and some of which may be task-specific.

### A. Four Simple Alternatives for Imitation

There are four simple and interesting possible outcomes of the imitation experiment, two of which have been explicitly considered in studies of time-interval perception.

[26]The possible inconsistency in conventional procedures between magnitude estimates (of stimuli controlled by the experimenter) and feedback from magnitude productions (controlled by the subject) is also a puzzle, of course.

[27]Insofar as symmetric errors are found for large fractions—as discussed in Sections II,A; III,C; and III,E—this account would, of course, have to be elaborated.

*Alternative 1: Accurate Imitation*

The first possibility is that despite errors in the other tasks, imitation (after the synchronization correction) will be accurate: $F = \mathbf{I}(f) = f$. The imitation function is thus the identity transformation for which output equals input and which we denote "**E**", and we have $\mathbf{I} = \mathbf{E}$. This outcome would follow from any model in which the errors found in the judgment and production tasks result from the processing of fraction names: Carlson and Feinberg discuss an example of such a model based on a "clock" or event-counting process.[28]

*Alternative 2: Imitation Consistent with Perceptual Judgment*

A second possibility is that the relation between $f$ and $F$ in imitation is the same as the relation between $f$ and $N$ in the perceptual judgment task so that $\mathbf{I}(f) = \mathbf{J}^{-1}(f)$, or $\mathbf{I} = \mathbf{J}^{-1}$. This consistency could arise if the stimulus pattern, which is common to the two tasks, is processed in the same way, leading to the same internal representation, and if the further transformations of this representation (that lead to $F$ in imitation and $N$ in judgment) are equivalent.

*Alternative 3: Imitation Consistent with Production*

A third possibility is that the relation between $f$ and $F$ in imitation is the same as the relation between $n$ and $F$ in production so that $\mathbf{I}(f) = \mathbf{P}(n)$ when $f = n$. The relation $\mathbf{I} = \mathbf{P}$ could arise, for example, if $f$ and $n$ generate a common internal representation when $f = n$, which is then transformed by the same processes in the two tasks to produce $F$.

*Alternative 4: Imitation Combines the Errors of Judgment and Production*

The fourth possibility is easiest to motivate by considering a model of imitation performance that would generate it. We describe this as a *full-concatenation* model because it calls for the application of all of the internal transformations used in the other two tasks. According to this model, the subject covertly assigns to $f$ a value on a continuum containing fraction names $N = \mathbf{J}^{-1}(f)$ (just as in the judgment task) and then produces the timed response $F = \mathbf{P}(N)$ that corresponds to that $N$-value (just as in the production task). We ignore, for the present, the possibility that special difficulties would be introduced by N-values that did not correspond to simple fractions. The result is $F = \mathbf{I}(f) = \mathbf{P}[\mathbf{J}^{-1}(f)]$. (We represent this by $\mathbf{I}(f) = \mathbf{P}\mathbf{J}^{-1}(f)$ or $\mathbf{I} = \mathbf{P}\mathbf{J}^{-1}$.) Note that since $\mathbf{J}^{-1}(x) > x$ and $\mathbf{P}(x) > x$ for small $x$ (on average), the two combined errors are in the same direction. Thomas and Brown (1974, Section V)

[28]See footnote 18 for a description of the model.

assumed a full-concatenation model in their study of the filled-duration illusion in the perception of time intervals.[29,30]

### B. One-Response Imitation (Experiment 5)

In Experiment 5 (Fig. 1E) the first pair of beat clicks was followed by a marker click. (In the corresponding judgment task of Experiment 1, the marker followed the second pair of beat clicks, so the contexts for the time-pattern stimuli in the two experiments were not precisely the same.) The subject attempted to respond with a single finger tap after the final beat click (as in the corresponding production task of Experiment 3) to imitate the presented fraction defined by the marker click. The fractions to be reproduced were the objectively correct fractions that correspond to the fraction names used in Experiments 1 and 3.[31] In the imitation task, however, no name was specified to the subjects. The fraction to be imitated remained the same for 25 consecutive trials. We determined the raw mean response time, $\mathbf{I}'(f)$ for each fraction; we then corrected this value by subtraction: $\mathbf{I}(f) = \mathbf{I}'(f) - \mathbf{S}'(0)$, with the same synchronization correction used in Experiment 3.

Mean values of $\mathbf{I}(f)$ are given in row 5 of Table I, proportional error curves for individual subjects are shown in Figs. 2, 3, and 4, $\ln(F)$ versus $\ln(f)$ is plotted for PZ in Fig. 5, and mean SDs are shown in Fig. 6. As in the production task, imitations of the small fractions 1/8 and 1/6 tend to be too large, and imitations of the complementary large fractions 7/8 and 5/6 tend to be too small. Symmetric distortion in the direction of 1/2 has also been described by Fraisse (1956, Chapter IV) and Povel (1981). Note, however, that the effect is absent in our data for the fractions (1/4, 3/4) closest to 1/2.

### C. Choice among Alternative Imitation Functions: Rejection of Accuracy and Full-Concatenation Models

To test the four alternative possibilities for imitation performance outlined in Section IV,A, we calculated deviations between the observed imitation function and tbe function expected from that alternative for each replication within each subject's

[29]If we add to this model the assumption that the component operations are stochastically independent, it follows that the variance of $F$ in imitation must be at least as great as the variance of $F$ in production. It must also be at least as great as the variance that would be induced in the production of $F$ by virtue of variability in the $N$-values on which responses are based in the judgment task. Given plausible assumptions, the SD that measures this induced variability can be estimated by multiplying the SD of the appropriate PMF from the judgment task by the derivative of the production function $\mathbf{P}(n)$ at the appropriate $n$-value. [Since $\mathbf{P}(n) \cong n$ for $n \geq 1/4$, this derivative is close to 1.0 for $n \geq 1/4$.]

[30]Imitation would also combine the errors of judgment and production if, for example, it shared just an input process with the former and just an output process with the latter *and* if these two processes were fully responsible for the errors in their respective tasks.

[31]It is a limitation of the experiment that other fraction values, such as those *judged* to be equivalent to simple fraction names, were not used as stimuli for imitation.


Biblioteca del Real

data, based on results from the matched procedures of Experiments 1, 3, and 5. For example, for Alternative 2 we calculated the contrast $\mathbf{J}_1^{-1} - \mathbf{I}_5$ for each fraction; insofar as this alternative is valid, these contrasts (whose means over subjects are displayed in row 9 of Table I) should be close to zero.[32]

Numerical examples for Alternatives 2 and 4 may be helpful; we shall base them on second replication data from PZ, for $f = 1/8$ ($bf = 125$). These examples are clearer when the argument of **J** is expressed in fractional intervals (in msec) rather than fractions. Thus, for Alternative 2, $\mathbf{J}(1/6) = \mathbf{J}(166.7) = 96.6$, and $\mathbf{J}(1/4) = \mathbf{J}(250) = 206.2$. Linear interpolation gives $\mathbf{J}(188.6) = 125$, or $\mathbf{J}^{-1}(1/8) = \mathbf{J}^{-1}(125) = 188.6$. According to Alternative 2, this value should be equal to $\mathbf{I}(1/8) = 161.1$. The contrast is $\mathbf{J}^{-1}(1/8) - \mathbf{I}(1/8) = 27.5$ msec.

For Alternative 4 we need $\mathbf{P}[\mathbf{J}^{-1}(1/8)]$, and from above we have $\mathbf{J}^{-1}(1/8) = 188.6$. We therefore need $\mathbf{P}(188.6)$. Second replication data from PZ give $\mathbf{P}(1/6) = \mathbf{P}(166.7) = 170.3$ and $\mathbf{P}(1/4) = \mathbf{P}(250) = 242.0$. Linear interpolation gives $\mathbf{P}(188.6) = 189.2$. According to Alternative 4, this value should be equal to $\mathbf{I}(1/8) = 161.1$. The contrast is $\mathbf{P}[\mathbf{J}^{-1}(1/8)] - \mathbf{I}(1/8) = 28.1$ msec.

We have used three methods to compare the relative goodness of fit of the four alternatives to our data. Since none of these methods is ideal, but taken together they point clearly in one direction, we mention results from all three. We restrict our attention to the seven fractions $1/8 \leq f \leq 7/8$ for which we were able to calculate contrasts for all four alternatives. For Alternatives 1 through 4, respectively, the numbers of individual subject contrasts (of 21 possible) that reach significance are 3, 8, 2, and 7, respectively, favoring Alternatives 1 and 3. The numbers of tests of means over subjects that reach significance are 2, 0, 0, and 0, however, indicating more consistency over subjects in the failures of Alternative 1 and thereby favoring Alternative 3. The mean squared deviations (contrasts) for the four alternatives are 514, 972, 148, and 1454, respectively, clearly favoring Atlernative 3; the same ordering is observed for the mean squared deviations associated with the three smallest fractions, which fall within the range of our most interesting and surprising findings.

Taken together then, our results favor Alternative 3 (imitation consistent with production) for the range of fractions we examined and permit us to reject the two alternatives considered in Section IV,A (accurate imitation, and the full-concatenation model) that have been previously considered for longer durations.

Further evidence bearing on the choice among the four alternatives can be found in relations among the variabilities of performance in the three tasks (Fig. 6). First, the SD functions for $\mathbf{I}_5$ and $\mathbf{P}_3$ are strikingly similar in form, again favoring Alternative 3. The increasing divergence of the two functions with size of the produced fraction is statistically significant, however ($t_2 = 8.4$; $p < .02$). In the context of the mechanism proposed in Section IV,A, for Alternative 3, this divergence could arise if the value of the common internal representation is more variable when it is derived from $f$ than from $n$

[32]Note that nonlinearities in the computation for Alternatives 2 and 4 result in discrepancies between the *mean contrasts* displayed in rows 9 and 11 of Table I and *contrasts of the means* of the components in rows 1, 3, and 5.

and if this variability difference grows with $f$. Second, although the SD of imitation is no smaller than the SD of production (as required by the full-concatenation model together with a stochastic independence assumption; see footnote 29), the SD of imitation *is* smaller than the corresponding variability measure associated with the judgment task, which violates an additional requirement of that model. This observation argues further against Alternative 4.

### D. Implicit Scaling of Beat Fractions from Imitation and Production Data

Performances in the judgment, production, and imitation tasks interest us primarily because of the light they and the relations among them can shed on underlying timing mechanisms. As already discussed, however, the judgment and production tasks can also be regarded as two different methods for establishing a psychophysical scale—a function that relates the beat fraction $f$ and its subjective magnitude. In each method the scale is established by identifying the subjective magnitude with an $n$-value. Both methods are explicit in that the subject's response is identified directly with one of the terms in the function. If we believe there is one "true" scale, then the fact that the two methods disagree implies that the scale derived from at least one of them is incorrect. As discussed in Section III,G, this difficulty also arises in other perceptual domains and has been attributed to effects on response-generation processes that distort the inferred association between stimuli and their internal representations. In Section III,F, we considered one such explanation (the existence of a minimum RT) for the systematic error we found in production.

The assumption required to use judgment and production tasks as explicit scaling methods—that responses accurately reflect magnitudes of the internal representations that are the objects of interest in psychophysics—is therefore subject to question. We can replace this assumption by a much weaker one if we use an *implicit scaling procedure* in which the scale is derived by combining data from production and imitation. The weaker assumption permits response biases or other distorting effects associated with responses to exist and requires only that those that operate in the generation of the timed response ($F$) in production also operate in the generation of the timed response in imitation. Under this assumption (which we use in developing the model described in Fig. 7) if a stimulus $n_0$ in production leads to the same timed response as does a stimulus $f_0$ in imitation so that $\mathbf{P}(n_0) = \mathbf{I}(f_0)$, then the internal representations of $n_0$ and $f_0$ that are used by the (common) response generation processes must have equal magnitude. The equation $\mathbf{P}(n) = \mathbf{I}(f)$ therefore establishes a scale relating $n$ and $f$ that is free of response effects; it may be written $n = \mathbf{P}^{-1}\mathbf{I}(f)$.

Let us consider what the four simple alternatives for imitation (Section IV,A) imply about the resulting implicit scale:

1. *Accurate imitation.* Since $\mathbf{I}$ is the identity transformation in this case, $n = \mathbf{P}^{-1}\mathbf{I}(f) = \mathbf{P}^{-1}(f)$ or $f = \mathbf{P}(n)$, so that the implicit scale is the same as the scale based on production.

2. *Imitation consistent with judgment.* Here $\mathbf{I} = \mathbf{J}^{-1}$. Hence $n = \mathbf{P}^{-1}\mathbf{I}(f) = \mathbf{P}^{-1}\mathbf{J}^{-1}(f)$; or $f = \mathbf{JP}(n)$. Since $\mathbf{P}(x) > x$ and $\mathbf{J}(x) < x$ for small $x$, the relation between $f$ and $n$ specified by the implicit scale depends on relative magnitudes of the errors in $\mathbf{J}$ and $\mathbf{P}$.

3. *Imitation consistent with production.* Here $\mathbf{I} = \mathbf{P}$. Hence $n = \mathbf{P}^{-1}\mathbf{I}(f) = \mathbf{I}^{-1}\mathbf{I}(f) = f$, so the implicit scale is free of systematic error (veridical).

4. *Imitation combines errors.* Here $\mathbf{I} = \mathbf{PJ}^{-1}$. Hence, $n = \mathbf{P}^{-1}\mathbf{I}(f) = \mathbf{P}^{-1}\mathbf{PJ}^{-1}(f) = \mathbf{J}^{-1}(f)$, so that the implicit scale is the same as the scale based on perceptual judgment.

We have seen above that results of the imitation experiment favor Alternative 3. One implication is that despite the inaccuracy (and inconsistency) of the explicit scales based on judgment and production data, the implicit scale based on combining results from the two tasks with a common response (imitation and production) is free of systematic error.

## V. A SHARED-PROCESS MODEL OF THE PERCEPTION, PRODUCTION, AND IMITATION OF BEAT FRACTIONS

In this section we present an information-flow model of performance in our three tasks. It is a minimal model in that we make as few assumptions as we can and limit ourselves to accounting for major features of the data. We think of each task as involving processes that perform input, translation, and output functions, and a principle of parsimony leads us to assume that different tasks share whatever processes they can. Given this starting point, the model incorporates the minimum possible number of constituent processes.

### A. Definition of the Model

The processes in the model responsible for judgment performance are represented by the two upper boxes in Fig. 7. A time-pattern stimulus $f$ generates an internal uncategorized or "analog" representation by a transformation $\mathbf{T}_{fa}$. (We call the representation "analog" only to indicate that it does not reflect a categorization of the stimulus that maps 1-1 onto fraction names.)[33] This representation must then be converted by a transformation $\mathbf{T}_{aN}$ into an internal "name" representation to generate the required fraction-name response $N$. The resulting compound transformation is denoted $\mathbf{T}_{aN}\,\mathbf{T}_{fa}$.

The processes in the model responsible for production performance are represented by the two lower boxes in the figure. A fraction-name stimulus $n$ is converted into an internal analog representation by a transformation $\mathbf{T}_{na}$, which is then used to generate

[33]One possibility is that this representation encodes both fraction (normalized marker interval) and beat interval. An alternative is that the beat interval is reflected by the rate of an internal clock or accumulator so that the analog representation has to encode only the marker interval in terms of the count or value accumulated.

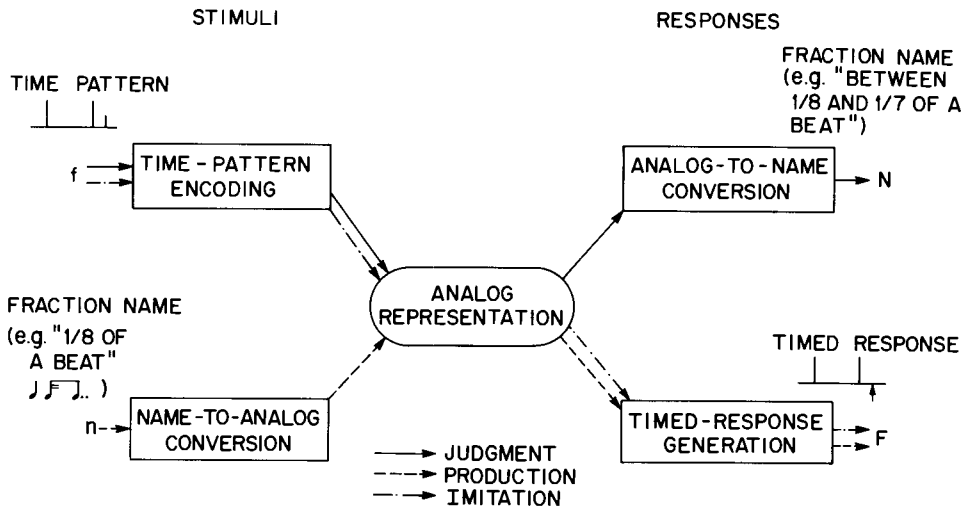



**Fig. 7.** A shared-process information-flow model of the perception, production, and imitation of beat fractions. The model incorporates four processes that convert time-pattern (*f*) or fraction-name (*n*) stimuli into time-pattern (*F*) or fraction-name (*N*) responses and that make use of a common intervening representation. Transformations carried out by the four processes are symbolized by $\mathbf{T}_{input,output}$. Paths of information flow for the three tasks are represented by unbroken, broken, and dotted arrows, respectively.

the required timed response by a transformation $\mathbf{T}_{aF}$.[34] Since no feedback process has been incorporated in this account and judgment and production share no common processes, the inconsistency between perception and production is not paradoxical.

A full-concatenation model of imitation (Alternative 4) would most naturally be represented by a system in which the upper and lower pair of processes had separate intervening representations, instead of the common analog representation shown in Fig. 7. Information could then not flow directly from time-pattern encoding to timed-response generation. Instead, a covert response output of the pair of processes used in judgment would become the input for the pair of processes used in production; the resulting compound transformation converting *f* to *F* would be $\mathbf{T}_{aF}\,\mathbf{T}_{na}\,\mathbf{T}_{aN}\,\mathbf{T}_{fa}$. Because such a model can be rejected, we adopt a *partial-concatenation model* of imitation, which shares only the encoding process of the judgment task and the response-generation process of the production task and makes use of an internal representation that is common to the two tasks. The resulting compound transformation converting *f* to *F* is $\mathbf{T}_{aF}\,\mathbf{T}_{fa}$.[35,36]

[34]Again there are several ways in which the (subjective) beat interval might be represented to provide the information that must be incorporated with the fraction name to define the response.

[35]Note that it is only because production and perception errors are not compensatory (i.e., do not conform to the feedback model) that we can discriminate a partial- from a full-concatenation model of imitation.

[36]An alternative two-process model of imitation in the same spirit would separate the information flow in judgment and production into phases that precede and follow the establishment of internal *name* repre-

### B. Restrictions on the Four Processes

Experiments to be described in Sections VI and VII help further to elucidate performance in the judgment and production tasks, and will eventually help to flesh out the skeleton shown in Fig. 7. Even with the results presented thus far, however, if we assume the structure of the model some interesting and surprising inferences can be made about the relations among the transformations $\mathbf{T}_{fa}$, $\mathbf{T}_{aN}$, $\mathbf{T}_{na}$, and $\mathbf{T}_{aF}$ carried out by its four component processes. Given these four transformations, there are six transformation pairs; our data permit an inference about the relation between the members of each pair.

The starting points for these inferences are idealizations of four of the properties that appear to characterize performance in the three tasks. The four properties are as follows:

a. $\mathbf{J} \neq \mathbf{E}$. (There are systematic errors in judgment.)
b. $\mathbf{P} \neq \mathbf{E}$. (There are systematic errors in production.)
c. $\mathbf{P} = \mathbf{I}$ (When $n = f$, response times in imitation and production are the same.)
d. $\mathbf{I} \neq \mathbf{J}^{-1}$. (The response fraction $F$ in imitation is not equal to the name $N$ associated with the same $f$ in judgment.)

To make the inferences, we start by using the model to write each of the functions **J**, **P**, and **I** in terms of the pair of transformations they reflect: $\mathbf{J}^{-1} = \mathbf{T}_{aN}\mathbf{T}_{fa}$, $\mathbf{P} = \mathbf{T}_{aF}\mathbf{T}_{na}$, and $\mathbf{I} = \mathbf{T}_{aF}\mathbf{T}_{fa}$. The inferences are as follows:

1. From property (c) we have $\mathbf{T}_{aF}\mathbf{T}_{na} = \mathbf{T}_{aF}\mathbf{T}_{fa}$, or $\mathbf{T}_{na} = \mathbf{T}_{fa}$. The two input transformations are therefore the same, and hence the internal (analog) representations of stimuli $n$ and $f$ have the same magnitude when $n = f$. This corresponds to the observation that the implicit scale relating $f$ to $n$ (Section IV,D) is free of error. Identity of the input transformations of $n$ and $f$ suggest that performance is not an accidental property of input processes; changes in details of the time-pattern stimulus should therefore not have major effects on performance. Evidence favoring this suggestion is presented in Sections VI,A, and VI,B.
2. From property (a) we have $\mathbf{J}^{-1} \neq \mathbf{E}$ or $\mathbf{T}_{aN}\mathbf{T}_{fa} \neq \mathbf{E}$ and hence $\mathbf{T}_{aN} \neq \mathbf{T}_{fa}^{-1}$. (Not surprisingly, given errors in judgment, its input and output transformations are not inverses.)
3. Combining (1) and (2), we find $\mathbf{T}_{aN} \neq \mathbf{T}_{na}^{-1}$. Thus, the transformations analog to name (in judgment) and name to analog (in production) are not inverses.
4. From property (b) we have $\mathbf{T}_{aF}\mathbf{T}_{na} \neq \mathbf{E}$, or $\mathbf{T}_{aF} \neq \mathbf{T}_{na}^{-1}$. (Not surprisingly, given errors in production, its input and output transformations are not inverses.)
5. Combining (1) and (4), we find $\mathbf{T}_{fa} \neq \mathbf{T}_{aF}^{-1}$. Thus, the transformations time-pattern stimulus to analog (in judgment and imitation) and analog to timed response

---

sentations rather than (the earlier) internal *analog* representations. We prefer our alternative because it seems less likely to us that an interesting or plausible transformation (other than the identity transformation) would relate stimulus or response *names* to their internal representations than that such a transformation would relate stimulus or response *times* to their internal representations.

(in production and imitation) are not inverses. This could also have been inferred from $\mathbf{I} \neq \mathbf{E}$.

6. From property (d) we have $\mathbf{T}_{aF}\mathbf{T}_{fa} \neq \mathbf{T}_{aN}\mathbf{T}_{fa}$, or $\mathbf{T}_{aF} \neq \mathbf{T}_{aN}$. In other words, the two output transformations are distinct (unlike the two input transformations): values of the $N$ and $F$ derived from the same internal (analog) representation are distinct. A difference between the output transformations for $N$ and $F$ makes it plausible that changes in response details might influence performance in production and imitation; some tests of this possibility are presented in Section VII.

## VI. FURTHER ANALYSIS OF PERCEPTUAL JUDGMENT

In Sections VI and VII we report results of our search for explanations of the errors associated with small fractions in the judgment and production tasks; we describe four variations of the judgment task and three variations of the production procedure. Our aim in most of these experiments was to determine not whether there was *any* effect of changes in experimental conditions, but whether there were any effects large enough to suggest major sources of the performance errors.

### A. Attention Shifts and Delays: Effect of Marker-Click Pitch (Experiment 6)

The presented fraction is defined by the difference between the onset times of the beat and marker clicks. The *subjective* occurrence time of a click, however, may differ from its objective time by an amount that depends on perceptual delay (possibly influenced by the amount of processing required to mark its occurrence). To the extent that the perceptual delays of the beat and marker clicks differ, the presented fraction that is judged subjectively equal to a fraction name will differ from the objective fraction that corresponds to that name, even in the absence of other perceptual distortions. As mentioned in Section III,D, we felt that equality of perceptual delays was a plausible starting assumption for beat and marker clicks. In this section and the next we report results that bear on its validity.

For the perceptual judgment data described thus far, the beat and marker clicks had different pitches. In one possible explanation of the judgment errors, perception of the marker click is assumed to be delayed by the shift of attention from the pitch of the beat clicks to the pitch of the marker click. (For example, findings by Van Noorden, 1975, suggest that the delay might increase with the pitch difference by about 100 msec/octave.) Suppose that this attention shift can be initiated, and possibly completed, after the beat click but before the marker click, if there is enough time between them. (The marker pitch could be learned from earlier trials.) Suppose further that if the shift has not been completed before the marker click, a time interval is required for the marker to attract attention, whose duration decreases with time after the beat click; perception of the marker is delayed until the attention shifts. This

Conservatorio Superior de Música

would explain both the judgment error for small fractions and its decrease in magnitude for larger fractions.

This hypothesis implies that the errors associated with small fractions should be influenced by any manipulation that alters the time to shift attention, such as variation of the pitch difference between the beat and marker clicks.

A second reason to suspect that the pitch difference may be implicated in the judgment errors is based on its possible influence on perceptual organization of the series of clicks into sequential groups (Woodrow, 1909, 1951) or simultaneous streams (Bregman, 1978).

To investigate the effect of pitch differences, we had two subjects (PZ and SB) perform in the procedure of Experiment 1 with $N = 1/8$, one second beat intervals, and marker clicks of 1700, 2500, and 3000 Hz; the beat-click frequency was always 3000 Hz. The frequency of the marker click remained the same for 75 consecutive trials, and **J**(1/8) was derived from the last 50 trials of the staircase procedure. According to the attention-shift hypothesis, **J**(1/8), the $f$-value associated with $N = 1/8$ should be greater (and closer to 1/8) when marker and beat clicks are closer in pitch.

An analysis of variance failed to show a significant effect on **J**(1/8) due to the frequency variation: for marker-click frequencies of 1700, 2500, and 3000 Hz, **J**(1/8) had mean values of $f = 79.2$, 69.6, and 69.9 msec, respectively, with a standard error (based on 2 *df*) of 2.1 msec (a nonsignificant effect in the wrong direction). These results make unlikely an explanation of the estimation errors in terms of the time to shift attention along the pitch continuum.

### B. Time Marking by Onset versus Offset: Invariance of Judgment with Prolonged Markers (Experiment 7)

In general, one might expect the internal response to any brief stimulus to differ from the stimulus itself in both shape and duration (see Sternberg & Knoll, 1973, Sec. IV; Fastl, 1977). Furthermore, the subjective occurrence time of a stimulus should depend on the particular feature of the internal response used to mark it. If the internal responses produced by the beat and marker clicks were different or if different features were used to mark their occurrence times, these differences by themselves could produce the observed judgment errors. For example, if the subjective duration of a time interval delimited by a pair of clicks corresponded to an interval delimited by the *onset* of the internal response produced by the first and the *offset* of the internal response produced by the second, the subjective duration would be greater than the objective duration, defined as the difference between click onset times.[37]

To test this possibility we conducted a small perceptual judgment experiment

[37]If auditory signals are presented in close temporal proximity, the internal representation of one (especially the second) is probably affected by the presence of the other (e.g., Duifhuis, 1973; Fastl, 1977; Penner, 1974). Thus, forward masking causes the first of two clicks to elevate the detection threshold of the second. Such effects become negligible with delays of at most 100 msec, however, and are therefore unlikely to be important in determining the judgment error.

(with PZ as the only subject). We used the procedure and stimulus values of Experiment 2 and compared the normal time-pattern stimuli (with all clicks 5 msec in duration) with stimuli in which the marker duration was 62 msec. Let us assume that relative to its onset, the perceived offset time of a tone burst is delayed by about the same amount as its duration is increased. Given the hypothesis, then, we are led to expect an increase in $f = \mathbf{J}(N)$ of about 57 msec in the prolonged marker condition. Instead, we obtained no change: over six fraction names the measured mean increase was a negligible .6 ± 1.9 msec. (The SE is based on variability among the effects on PMF means for $N$ = 1/7, 1/6, 1/5, 1/4, 1/3 and 1/2. For $N$ = 1/8 there were insufficient data to generate a PMF.) There was neither a main effect of marker duration nor an interaction of marker duration with fraction size.

It is reasonable to suppose that any feature of an internal response whose occurrence time is invariant with changes in stimulus duration is located at or near the beginning of that response. The absence of an effect of marker duration therefore argues that the onset rather than the offset of the marker response is the critical feature that determines its subjective occurrence time, and suggests that the systematic judgment errors cannot be attributed to different features of the internal response being used to define the occurrence times of beat click and marker click. (If the offset rather than onset of the *beat* click were used by subjects, the resulting judgment error would be in the wrong direction.)[38]

### C. A Test of the Rate Constancy of Subjective Time between Beats: Effect on Fraction Perception of Fraction Location Relative to the Beat (Experiment 8)

In the perceptual judgment experiments reported thus far, the conditions for which systematic errors are largest have two features in common: first, the interval to be judged is small, and, second, it occurs in close temporal proximity to (indeed, is bounded by) the beat click. Suppose that subjective time during the beat interval was inhomogenous in the sense that relative to physical time it elapsed faster near the beat and more slowly elsewhere in the beat interval. Then small fractions defined by intervals near the beat would be overestimated, as observed, but the same small fractions elsewhere in the beat interval, and large fractions initiated by the beat, might not be. To determine how proximity to the beat of the interval being judged affects perceptual judgment, we instructed subjects PZ and PF to judge whether the interval between a pair of marker clicks was larger or smaller than 1/8 of a beat for marker pairs at six different locations within the beat interval.

The beat and marker clicks were 5-msec tone bursts of 3000 and 2500 Hz, respectively. We used four intervals between markers (50, 60, 70, and 80 msec) chosen based on earlier results to permit us to estimate PMFs for judgments relative to 1/8 of

[38]We conjecture that this finding reflects a general property of the perception of musical timing and rhythm: the dominance of the sequence of time intervals between the *onsets* of successive notes (attacks) and the relative unimportance of offset times, which probably serve articulative rather than timing functions.

a 1-sec beat interval. (If effects of proximity to the beat in this initial experiment had been large, of course, this set of intervals might not have provided sufficiently complete functions at all proximities.) On each trial we presented clicks for three beats—$B_1$, $B_2$, and $B_3$; subjects were asked to imagine the next two beats, $B_4$ and $B_5$. A marker-click pair, $M_1$ followed by $M_2$, defined one of the four intervals, and was located at one of six positions within the sequence of beats, as follows: (1) symmetric about the midpoint of $B_3$ and $B_4$, (2) such that $M_2$ preceded $B_4$ by 100 msec, (3) such that $M_2$ was simultaneous with $B_4$, (4) such that $M_1$ was simultaneous with $B_4$, (5) such that $M_1$ followed $B_4$ by 100 msec, (6) symmetric about the midpoint of $B_4$ and $B_5$.

Position 4 is, of course, the arrangement that had been used in our previous experiments in which the interval to be judged is initiated by the beat, except that the event that marks the beat is a marker click rather than a beat click. (For this position, then, beat $B_4$ is signaled by a click and is not imaginary, unlike the other positions.) At position 3 the interval to be judged is terminated by the beat. At positions 1 and 6 the interval to be judged is as far as possible from any beat. The six positions combined with four marker intervals defined 24 stimuli that were presented in random sequence (method of constant stimuli).

We found no systematic effects of proximity to the beat on either PMF means [values of $\mathbf{J}(1/8)$] or SDs. Over the two subjects, average PMF means for the six positions are 57.5, 61.7, 55.9, 60.0, 58.3, and 56.7 msec, respectively; (rms) average PMF SDs are 9.4, 10.9, 7.0, 5.5, 12.5, and 4.4 msec, respectively. Over the six positions the average PMF mean is 58.3 msec, and the (rms) average SD is 8.8 msec.[39] Again, the judgment error is surprisingly large: 58.3 msec is about 53% smaller than the correct value of 125 msec.[40] These results show that the judgment error depends neither on the judged interval being bounded by a beat click nor on the proximity of the judged interval to the beat. The rate at which subjective time elapses during the beat interval appears to be constant.

### D. A Constraint on the Precision of Dual Time Judgments and Its Implications for Timing Models and the Use of Feedback

A further variation of the perception task revealed an interesting and unexpected limitation in the judgment of time intervals. In an extension of Experiment 8, we

[39]It is instructive to compare these results to findings for the same two subjects in Experiments 1 and 2. Like Experiment 2 the present experiment involved a method of constant stimuli rather than a staircase procedure; like Experiment 1 the present experiment called for a narrow range of $f$-values. Since results of Experiments 1 and 2 at $N = 1/8$ for PZ and PF were similar, we have combined them to obtain a PMF mean of 62.5 msec and an SD of 7.7 msec. The present experiment produced almost identical values, suggesting that uncertainty from trial to trial about the position of the interval to be judged (which was much greater in the present experiment) is an unimportant factor in judgment performance.

[40]Because the two marker clicks had the same frequency (2500 Hz), these results also provide further evidence against the notion that the estimation errors result from a pitch difference between the clicks bounding the interval to be judged.

instructed one subject (PZ) first to estimate the duration of the brief interval bounded by the marker clicks (relative to 1/8 of a beat), as in the main experiment, and then also to judge whether the longer interval between the last beat click and the marker pair was less than or greater than one beat interval (i.e., marker clicks before or after $B_4$). The subject was instructed to perform accurately in judging the brief interval (primary task) and, having done so, to judge the long interval (secondary task) as accurately as he could.

The need to make the long-interval judgment did not substantially alter either the mean or the SD of the duration PMF: without the secondary task these parameters were 54.4 and 6.0 msec, respectively; with the added task they were 58.1 and 7.9 msec, respectively. On the other hand, the subject's precision in judging the long interval appears to have been greatly impaired by having also to judge the brief interval. One measure of the loss in precision is obtained by comparing performance in the secondary task to earlier performance (Experiment 1, $N = 1$), judging only the position of a single marker click relative to the beat; this comparison reveals that the SD of the PMF from the secondary task is more than *10 times* the SD obtained in the single-judgment, single marker-click procedure.[41] (In Experiment 1 the mean and SD of the PMF were 962.5 and 48.8 msec, respectively, versus 1114.5 and 502.6 msec, respectively, in the secondary task.)[42]

It is helpful to consider this observation in relation to a particular class of mechanisms that may underlie the timing process. One candidate for the analog representation in the information-flow model of Section V is the value attained by an internal clock or accumulator. (See Creelman, 1962; Treisman, 1963; Wing, 1973; Eisler, 1975; and Getty, 1976, for particular realizations of this idea.) In the judgment task, for example, the hypothesized clock starts with the initial event defining the interval and stops with the terminal event. (Alternatively, the current value of the clock is "saved" when the terminal event is detected.) Results of the dual-task variation of Experiment 8 require elaboration of such clock models to explain why the two successive intervals (the long interval from the beat to the first marker and the short interval from the first to the second marker) could not both be accurately judged.

One possibility is that the timing process permits only intervals that are similar in duration to be accurately classified in quick succession or concurrently. For example, the clock might have an adjustable rate: a slow rate for accurately judging large intervals and a fast rate for accurately judging small intervals.[43]

[41]Strictly speaking, to control for the possibility that the poor performance in the secondary task might be due to the physical nature of the stimuli (the end of the long interval was defined by two marker clicks rather than one), performance in the secondary task should be compared for conditions that have identical physical stimuli and that differ only by the presence or absence of the primary duration judgment. Although this control is logically necessary, and therefore should be used in further investigation of the phenomenon, it would surprise us if such enormous changes in performance could be explained by such minor variation of the physical stimuli.

[42]Thus, a marker pair in position 1 (500 msec before $B_4$) was judged to have occurred *after* $B_4$ with probability about .11; a pair in position 6 (500 msec after $B_4$) was judged to have occurred *before* $B_4$ with probability about .22.

[43]This limitation could explain our finding that the multiple-fraction procedure (Experiment 2) elicited

According to a second possibility, the clock cannot be started and stopped rapidly, permitting precise timing of only one of two adjacent intervals and requiring use of an alternative, less precise mechanism for timing the interval from the beat click to the first marker click.[44]

Either of the above possibilities could also account for failure of the feedback model of production (Section III,D), if we assume that the same timer is used for production as for perception. To produce a fraction appropriate for a specified name, subjects must time an interval from the beat click to the *initiation* of the response. Since there is a delay between the start of a response and its actual occurrence, it is possible that having accurately timed when to initiate the response, subjects cannot also accurately judge when the response occurs. Alternatively, suppose that timing for the production of small fractions is initiated by the penultimate beat (Section III,F) but that judgment of fractions of all sizes depends on timing from the final beat. Again, the constraints on timing discussed above would prevent the perceptual mechanisms used in the judgment task from being used to evaluate feedback for small fractions in the production task.

### E. The Dependence of Perceptual Judgment on Duration versus Fraction: Effects of the Beat Interval (Experiment 9)

In this section we examine two simple alternative ways to characterize perceptual judgment performance and the mechanisms responsible for judgment errors. One is a *duration model*, according to which the fundamental variable is the duration of the fractional interval. For a specified fraction name $N$ the correct value of this interval can be represented as $bN$, where $b$ is the beat interval and $N$ is the fraction name; the obtained (matched) value is then $bf$, the absolute error is $bf - bN$, and the relative error is $(bf - bN)/bN$. According to the duration model, judgment error depends only on the correct duration $bN$; once that is specified there is no further effect of beat interval on either the absolute or relative error. For example, the mean judgment error for $N = 1/8$ at $b = 1000$ msec should be the same as the mean error for $N = 1/6$ at $b = 750$ msec, since in both cases the correct duration is $bN = 125$ msec. The attention-shifting mechanism considered in Section VI,A exemplifies such a model.

The second alternative is a *fraction model*, according to which the fundamental variable is the fraction, or duration *ratio*. For a specified fraction name $N$ the correct fraction is $N$ itself, the obtained value is $f$, the absolute error is $f - N$, and the relative error is $(f - N)/N$. According to the fraction model, judgment error depends only on the correct fraction; once that is specified, there is no further effect of beat interval on either the absolute or relative error. For example, the mean judgment error (expressed

---

lower judgment precision than the single-fraction procedure (Experiment 1). It could also explain the finding by Vorberg and Hambuch (1978) that subjects attempting to produce precisely timed rhythmic patterns control the timing with a set of *chained* "timers" that produce approximately equal durations rather than hierarchically *nested* (concurrent) "timers" that produce highly disparate durations.

[44]In developing this possibility, the accuracy with which subjects judge the regularity (equality) of trains of successive intervals (Schulze, 1978) would have to be considered.

TABLE II
Experiment 9: Design and Mean Data[a]

| Beat interval ($b$) (msec) | Fraction name ($N$) 1/8 | | 1/6 | | 1/4 | | 1/2 | |
|---|---|---|---|---|---|---|---|---|
| 750 | .078 (.125) | 58 ( 94) | .108 (.167) | $\alpha$ 81 (125) | .203 (.250) | $\beta$ 152 (188) | .475 (.500) | $\delta$ 356 (375) |
| 1000 | .075 (.125) | $\alpha$ 75 (125) | .098 (.167) | 98 (167) | .238 (.250) | $\gamma$ 238 (250) | .450 (.500) | 450 (500) |
| 1500 | .102 (.125) | $\beta$ 153 (188) | .158 (.167) | $\gamma$ 236 (250) | .250 (.250) | $\delta$ 374 (375) | .456 (.500) | 683 (750) |

[a]Left-hand cell entries are relevant to the fraction model. Upper left entry is the obtained mean fraction $f$; lower left entry (in parentheses) is the target fraction $N$ (equal within each column). Right-hand cell entries are relevant to the duration model. Upper right entry gives the mean obtained duration $bf$ in msec. Lower right entry (in parentheses) gives target duration $bN$. Pairs of cells marked with the same Greek letter have equal target durations.

as a fraction) for $N = 1/8$ should be the same for all beat intervals. A mechanism in which the beat interval has its effect by controlling the rate of an internal clock exemplifies such a model.

On the basis of the judgment experiments considered above, the two models cannot be distinguished because the beat interval was constant (1 sec). In Experiment 9 we used the procedure of Experiment 1 to compare judgment performance by SB, PF, and PZ for each of four fractions ($N = 1/8$, 1/6, 1/4, and 1/2) at three different beat intervals ($b = 750$, 1000, and 1500 msec). The fractions were chosen so that we could examine performance with the same target *fraction* $N$ at each of three beat rates (permitting a test of the fraction model) and also with the same target *duration* $bN$ at two beat rates each (permitting a test of the duration model). These two possibilities are most easily seen by examining cell entries in Table II. The design is orthogonal with respect to $N$ and $b$; each of the four columns represents the same target *fraction* (left value in parentheses) for different beat intervals. Greek letters indicate those cells that represent the same target *duration* (right value in parentheses) at different beat intervals. For example, for both of the cells marked $\beta$ ($N = 1/4$, $b = 750$, and $N = 1/8$, $b = 1500$) the target duration is 188 msec.

Tests of both models involved the examination of row (beat-interval) effects in an appropriate two-way table. Let us consider the fraction model first. Here the two-way table has three rows (beat interval) and four columns (fraction name). If the fraction model is correct there should be neither a row effect nor an interaction of rows with columns: error measures associated with the three cells in each column should be equal. (Means over subjects of the obtained fraction values, shown at the

upper left to each cell in Table II, appear, in contrast, to change systematically with beat interval.)

For the duration model the full design is not orthogonal; tests are made possible by reducing the design. Four fractional intervals—with durations 125, 188, 250, and 375 msec—each appear at a "smaller" and "larger" beat interval in Table II in cells designated by Greek letters. The reduced two-way table therefore has two rows ("smaller" and "larger" beat interval) and four columns (one per fractional interval). If the duration model is correct there should be neither a row effect nor an interaction of rows with columns in the reduced table: Error measures associated with the two cells in each "column" (marked by the same Greek letters in Table II) should be equal. (That pairs of beat intervals differ from column to column does not invalidate the test of this null hypothesis.) The pairs of mean obtained interval values for the same target interval, shown in Table II, do appear to depend little on beat interval. For example, for the two cells marked $\alpha$, for which the correct duration is 125 msec, we obtained 81 msec at the smaller (750 msec) beat interval and a similar 75 msec at the larger (1000 msec) beat interval. (The analysis will show this independence of beat interval to be an artifact due to averaging over subjects, however.)

We were able to use the same error measure in tests of both models, based on analysis of variance. This was possible, first, because for each model absence of row effects and interactions for its absolute error would imply their absence for its relative error[45] and, second, because the two relative error measures [$(bf-bN)/bN$ for duration and $(f-N)/N$ for fraction] are equal.

Results of the analyses of variance for the two models are shown in the upper and lower halves of Table III. For the group analyses neither test shows a significant main effect of rows (b) nor a significant row-column interaction. Both analyses, however, reveal significant interactions of beat interval and subjects, indicating that there are row effects for individual subjects and thereby violating the models; differences among these effects for individuals are apparently large enough so that they cancel each other or otherwise render the main effects insignificant.

Results of the group analyses are consistent with the possibility that each subject's behavior conforms with one of the two models but that the same model does not apply to all three subjects. We tested this possibility by performing the same analyses for each subject separately; results are shown in the right-hand section of Table III. Both models are rejected by these individual analyses, with all three subjects providing evidence against the fraction model and two subjects providing (somewhat weaker) evidence against the duration model. For PZ, duration accounted for a larger percentage of variance than did fraction (76 versus 27%, respectively). For PF, the ordering was the same but the difference was small (80 versus 75%, respectively). Thus,

[45]Relative error can be obtained from absolute error in each case by dividing by the value of the column factor (correct duration or correct fraction). Suppose a two-way table of absolute errors has no row effects or row-column interactions. Then transforming its cell entries in a way that depends only on the column factor produces a new two-way table that also can have no row effects or row-column interactions. (If there is no row effect within any column before the transformation then there can be none after the transformation.)

TABLE III
Results of Analyses of Variance for Duration and Fraction Models

| Model | Factor | $p$-value from group analysis | $p$-values from individual analyses | | |
|---|---|---|---|---|---|
| | | | PZ | SB | PF |
| Fraction | Beat interval ($b$) | n.s.[a] | <.001 | <.005 | <.01 |
| | Fraction ($N$) | n.s. | <.001 | n.s. | <.001 |
| | Subjects ($s$) | <.001 | — | — | — |
| | $b \times s$ | <.001 | — | — | — |
| | $N \times s$ | <.001 | — | — | — |
| | $b \times N$ | n.s. | <.001 | n.s. | <.025 |
| | $b \times N \times s$ | <.005 | — | — | — |
| Duration | Beat interval ($b$) | n.s. | <.005 | n.s. | <.05 |
| | Duration ($bN$) | n.s. | <.001 | n.s. | <.001 |
| | Subjects ($s$) | <.001 | — | — | — |
| | $b \times s$ | <.001 | — | — | — |
| | $bN \times s$ | <.001 | — | — | — |
| | $b \times bN$ | n.s. | n.s. | n.s. | n.s. |
| | $b \times bN \times s$ | <.01 | — | — | — |

[a]n.s. = not statistically significant ($p > .05$).

although both simple models can be rejected convincingly, the data favor the duration over the fraction model.

Experiments 1 and 2 had a fixed beat interval of 1 sec and are therefore insensitive to the distinction between fraction and duration. To generalize from the results of Experiment 9, we must demonstrate that the findings do not depend on its special conditions, in which the beat interval was changed within sessions. One test of the invariance of performance is to compare the results of Experiments 1, 2, and 9 for those conditions common to all of them (i.e., a beat interval of 1 sec and fraction names of 1/8, 1/6, 1/4, and 1/2). An analysis of variance with the factors experiment, fraction name, and subject resulted in a significant effect of fraction name $F_{3,5} = 7.64$, $p < .025$, but a nonsignificant effect of experiment $F_{2,4} = .73$. Thus, performance (the size of mean judgment errors) at one beat interval appears not to be influenced by the subject having recently worked with other beat intervals. (This finding is consistent with a common belief about musical performance.)

In summary, although performance is better described in terms of the duration being judged than the fraction, the consistent effects of beat interval require us to reject a model in which duration is the sole determinant: beat interval as well as fractional interval influences the size of the judgment error.[46]

[46]If one considers the beat interval to be a standard against which each fractional interval is compared, this conclusion bears on the typical magnitude estimation paradigm in research on time perception. The

## VII. FURTHER ANALYSIS OF PRODUCTION

In this section we describe results from three variations of the production task, aimed at testing hypotheses about major sources of the systematic error for small fractions.

### A. Production Errors with Musical Instruments (Experiment 10)

One potential source of the production errors we observed is our choice of finger tapping as a response; our subjects are skilled players but not necessarily skilled finger tappers. It was possible that the errors would disappear if subjects used their instruments instead. We tested them in the one-response production task of Experiment 3 (Fig. 1C) using their instruments: flute (SB), cello (PF), and violin (PZ). On each trial a subject played a single note after the final beat click, attempting to produce an interval that corresponded to a specified fraction between the click and the beginning of the note. We measured the occurrence time of the note with an acoustic energy detector; its threshold was adjusted to be relatively low so that it would be exceeded near the start of the acoustic signal. The fraction name ($n$ = 1/8, 1/2, or 1) remained the same for 25 consecutive trials. In Experiment 3 mean response times (corrected by synchronization) for these three fractions were 139.1, 486.3, and 1023.3 msec, respectively. Corresponding response times in Experiment 10 were 195.0, 480.0, and 1018.3 msec, respectively.[47] One reason the mean difference between experiments is so large at $n$ = 1/8 is that whereas PF differed from the other two subjects in Experiments 3 and 4 in not showing a positive production error at $n$ = 1/8, she did produce a substantial positive error in Experiment 10, bringing her into conformity with the other subjects. SDs of the response delays did not differ systematically between experiments; (rms) average SDs over fractions and subjects were 31.5 msec for finger taps and 34.0 msec for instrument notes. Mean values of $\mathbf{P}(1/2)$ and $\mathbf{P}(1)$ were almost identical between experiments, but $\mathbf{P}(1/8)$ was substantially larger in Experiment 10. Each of the three subjects shows this interaction between $n$-value and response mode, and an analysis of variance shows it to be significant.[48] That the mean response delay at $n$ = 1/8

existence of a beat-interval effect suggests that results of magnitude estimation tasks that employ an explicit standard interval may depend on the size of the standard. In the absence of an explicit standard, subjects may use an implicit standard. This could increase variability within or between experiments if, for example, the implicit standard varied between subjects or depended on experimental manipulations such as the distribution or range of intervals to be judged.

[47]Values from Experiment 10 are raw response times with no synchronization correction applied. Unfortunately, we collected synchronization data with musical instruments only for PZ; for him, $\mathbf{S}'(0)$ was 4.2 msec (as compared to −9.0 msec for one-response finger taps). Results from the present experiment are sufficiently clearcut so that the absence of these small corrections is unlikely to affect our conclusions. Note that a synchronization correction would change all three response times by the same amount and would therefore have no effect on differences among $\mathbf{P}(n)$ values.

[48]The differences between absolute errors at $n$ = 1/8 and $n$ = 1, $[\mathbf{P}(1/8)-1/8]-[\mathbf{P}(1)-1]$, for SB, PF, and PZ are 47.6, 49.8, and 57.8 msec, respectively.

associated with the musical instrument production was much larger than that of the tap response suggests that the errors measured with the tap response may *underestimate* the size of errors that occur in a more natural context.

### B. Production with Marker-Click Feedback (Experiment 11)

The disparity between performances for small fractions in the perceptual judgment and production tasks led us to reject the feedback model of production, but we did not scrutinize the feedback itself. The feedback available from tapping the finger in the production task probably includes tactile, proprioceptive, and auditory cues, but did not include the marker click we used in the judgment task.[49] Could this difference between the events that terminated the critical intervals in the two tasks explain the perception-production disparity?

We asked this question in a small production experiment with one subject (PZ), in which each finger tap produced a marker click identical to the markers in Experiments 3 and 4, thereby providing augmented feedback.[50] With this procedural change the sequences of clicks in the judgment and production tasks (but not necessarily their timing) become identical. The critical difference between tasks is limited to how the timing of the final (marker) click is controlled: in one case the subject controls it directly by choosing when to tap his finger, attempting to make the click time correspond to a specified fraction; in the other case the experimenter controls the click time, and the subject judges it relative to that fraction. We used both one-response and repeated-response procedures.

For this subject, the mean value of **J**(1/8) over perceptual judgment Experiments 1 and 2 together with the corresponding conditions in Experiments 6 and 9 is 64.3 ± 3.5 msec (SE based on between-experiment variation). For production without marker-click feedback, the mean value of **P**′(1/8) over Experiments 3 and 4 is 158.5 ± 2.9 msec. For production under the new condition with augmented feedback, the mean **P**′(1/8) was 149.0 ± 2.9 msec.[51] (With click feedback the one-response and repeated-response procedures produced values that differed by only 4 msec.) Thus, the effect on **P**′(1/8) of adding marker-click feedback is small and not statistically significant, whereas **P**′(1/8) values from experiments both with and without augmented feedback differ reliably from **J**(1/8). Again, despite the large disparity between production and judgment, the subject appeared satisfied with his performance.

That the disparity was maintained—even when the cues in production that were

[49]The cues comprising finger-tap feedback may have been less salient and punctate than marker clicks, and might therefore have limited the precision of subjects' knowledge of the occurrence times of their responses. That the production error generalizes to played notes (Section VII,A) seems to argue against this as a major basis for the error, however.

[50]We actually ran SB as well, but her data were too few and too variable to be conclusive.

[51]We report uncorrected response times because the validity of any correction based on the tap-click synchronization procedure is suspect under these conditions; with an additional tap-produced click, special perceptual mechanisms may be available on which to base simultaneity judgments of marker click and beat click— mechanisms not available for judgments of the timing of taps versus clicks.

available for use as feedback included the same marker click as in perceptual judgment—supports our suggestion (Section VI,D) that subjects process stimulus information differently in the judgment task from the way they process feedback in the production task.

### C. Production of Single versus Multiple Subdivisions of the Beat (Experiment 12)

The existence of large production errors for small fractions suggests a paradox: How can this finding be reconciled with our belief that musicians are able to produce response sequences that fill the beat interval accurately and evenly? Could the production error depend on our use of a single, isolated response? Suppose that beats are salient and relatively precise periodic time references. Then the task of accurately producing a sequence of evenly spaced responses, one or more of which are coincident with beats (certainly a more frequent pattern in music than the production of an isolated offbeat response) might be easier because better information might be available for error correction.

Experiment 12 was designed partly to explore this possibility. The three conditions are shown in Fig. 1F and are designated 1R, 4R, and 5R in accordance with the number of tap responses required. For descriptive purposes it is convenient to define potential responses in five positions, $R_1$, $R_2$,...,$R_5$. Accurate performance would require $R_1$ and $R_5$ to be coincident with beat clicks and would require the five tap times to be evenly spaced, dividing the beat interval into quarters. In Condition 1R only one of the responses ($R_2$) was executed; this is the condition studied in our other production experiments. In Condition 4R the first executed response is $R_2$, as in 1R, but the three following responses are executed as well; accurate performance would require the last response to be coincident with the beat. In Condition 5R all five responses are executed.

We used a beat interval of .5 sec instead of the 1.0 sec interval used in our other production experiments. The fractional interval associated with $n = 1/4$ therefore corresponds to 1/8 in previous experiments. Pilot work had demonstrated the existence of a production error in the 1R condition with the shorter beat interval; the experiment incidentally tests the generality of the production error at a different beat interval.

If the production error depends on planning and performing an isolated response shortly after the beat, then $R_2$ should be delayed only in Condition 1R. An alternative explanation of the error is similar to one we considered (and rejected) for the judgment error in Section VI,C: the possible inhomogeneity of subjective time during the beat interval. Because judgment and production appear to depend on timing mechanisms that are at least partially distinct, our rejection of such a possibility for the judgment task does not preclude it as an explanation of the production error. If subjective time elapsed relatively slowly near the beat, then $R_2$ would be delayed in all three conditions of Experiment 12.

The subjects were PZ and two experienced amateur players, JM and SS. Conditions were held constant for sequences of 25 trials, providing opportunities for detected errors to be corrected on later trials. Subjects tapped with alternate index fingers, chosen so that they used the same finger to produce $R_2$ in all conditions. We treated $R_1$ (in Condition 5R) as synchronization data and adjusted all response times by subtraction (as discussed in Section III,A). $R_1$ occurred early, on average, with the mean $\mathbf{S}'(0) = -29.5$ msec.[52] Results are displayed in Fig. 8.

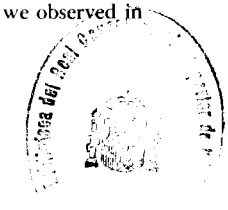
Condition 1R produced a large and reliable positive mean error of 59.1 ± 16.6 msec, or 47%, thereby generalizing our finding to additional subjects and a shorter beat interval.[53] When $R_2$ was the initial response in the 4R condition, it was produced with as much delay as in the 1R condition. This result shows clearly that the production error is not a consequence of our requiring an isolated response. The mean occurrence time of $R_2$ in the 5R condition was 119.7 msec, which corresponds to a negligible negative error. Thus, it is the withholding of $R_1$ that causes the error in $R_2$. The accuracy of the interval between $R_1$ and $R_2$ in Condition 5R argues against any explanation that depends on a distortion of subjective time near the beat during the production task. Rather than being associated with the first subdivision *after* the beat, the productionerror for small forward fractions is associated with the absence of a response *on* the beat (i.e., with "coming in" shortly after the beat).[54]

Let us consider the occurrence times of the remaining responses in Conditions 4R and 5R. The displaced parallel lines in Fig. 8 fit well, indicating that the mean times

[52]The use of this correction is supported, for the average data, by the fact that it places the mean time of $R_5$ in Condition 5R within 3 msec of the final beat click. (Indeed, the time between $R_1$ and $R_5$ can be regarded as a measure of the subjective beat interval.) As we shall see, however, the remarkable accuracy of the mean response rate (or subjective beat interval) implied by this result is an artifact of averaging and does not apply to the data from individual subjects. Using the notation $\mathbf{P}(n;b)$, for PZ we have $\mathbf{P}(1;500) = 437 \pm 20$ msec, indicating a subjective beat interval that is too short. In contrast, from Experiment 3 for PZ we have $[\mathbf{P}(1;1000)]/2 = 506 \pm 2$ msec. Since the new condition is distinguished by "filling" the interval with repeated taps, the reliable difference may be an instance in production of a "filled-duration illusion" (Michon, 1965; Ornstein, 1969) that has frequently been observed in judgment tasks. Since the synchronization correction depends on a different finger from the finger used for $R_2$, in evaluating the timing of $R_2$ we must consider the possibility of a timing difference between fingers that would produce a sawtooth pattern of production times. Figure 8 shows that any such difference is small.

[53]It is important to determine whether the production error depends exclusively on either the target interval $bn$ (duration model) or the target fraction $n$ (fraction model). (We asked this question about perception in Section VI,E.) Comparison of these results to production performance in Experiment 3, in which the beat interval was twice as long, provides a small amount of evidence (from PZ only) bearing on this question. Using the notation $\mathbf{P}(n;b)$, we have for PZ, $\mathbf{P}(1/4;500) = 151 \pm 8$ msec, $\mathbf{P}(1/8;1000) = 164 \pm 6$ msec, and $\mathbf{P}(1/4;1000) = 242 \pm 6$ msec. Whereas the interval error difference between the first two (equal $bn$) is small and not significant, the fraction-error difference between the first and third (equal $n$) is large and significant. These results argue against a fraction model for the production error and provide (weak) evidence favoring a duration model.

[54]One possibility is that withholding a response on the beat ($R_1$) requires the establishment of an inhibitory state that takes time to dissipate and thereby delays $R_2$. This mechanism could not also lead to the enlargement of short intervals *initiated* by finger taps before the beat, however, that we observed in Experiments 3 and 4. (See Sections III,C and E.)



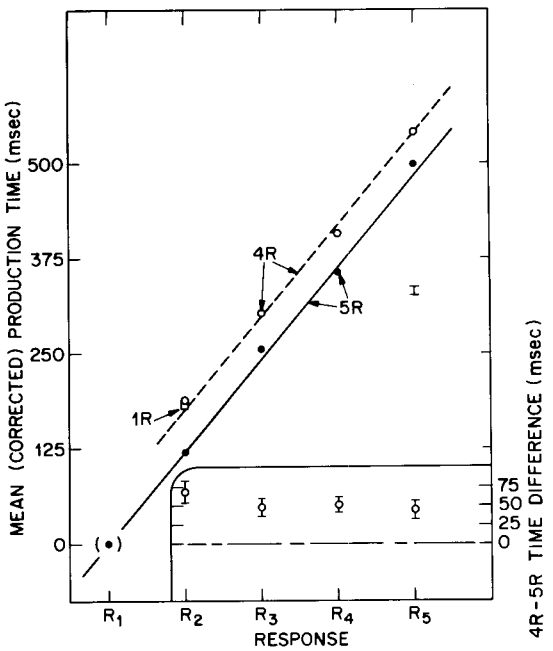



**Fig. 8.** Mean production times and production-time differences in Experiment 12 (multiple divisions of the beat). The main panel (left-hand ordinate) shows mean production time for each response in the three conditions, after a synchronization correction (−29.5 msec) based on $R_1$ is applied. (This forces production time for $R_1$ to be zero.) Parallel lines were fitted by least squares and then adjusted such that the 5R line passes through the $R_1$ point. The displayed ± 1 SE bar is appropriate for assessing adequacy of the fitted lines. The inset (right-hand ordinate) shows differences between production times of corresponding responses in 4R and 5R conditions together with SEs appropriate for assessing deviations from zero.

between successive responses are close to being equal, both within and between conditions.[55] The remarkable implication for Condition 4R is that the production error in $R_2$ is propagated through the three succeeding responses. This phenomenon seems best described as *displacement of the subjective beat*. Having produced $R_2$ with a delay, subjects do not "catch up" because their perception of the train of beats has also been shifted.[56] The lower panel of Fig. 8 displays the occurrence-time differences

[55]If subjects' response rates were accurate, the slope of the fitted lines (mean interresponse time) would be 125 msec/response. The actual mean interresponse time for Condition 5R alone is remarkably close to this value: 123.0 msec. This is an accident of averaging three very different values, however: for JM, SS, and PZ, the mean interresponse times in Condition 5R were 142.6, 117.2, and 109.3 msec, respectively. Such rate inaccuracies imply that even in Condition 5R the final response is far from coincident with the final beat click, again suggesting a surprising insensitivity to the time relation between a response and an external stimulus. (For example, in Condition 5R, PZ's mean $R_1$ occurred 34.7 msec before its beat click, and his mean $R_5$ occurred 88.3 msec before its beat click. If $R_1$ was subjectively coincident with its beat click, then $R_5$ should have subjectively anticipated the next beat by 53.6 msec, on average.)

[56]That there is a final *actual* beat click that is *not* shifted (and that should be coincident with the final response) raises the same question about beat displacement in 4R as it does about the incorrect response rate in 5R.

of corresponding responses in Conditions 4R and 5R; they range from 67.8 ± 13.9 msec for $R_2$ to 44.1 ± 12.3 msec for $R_5$. All three subjects show a decline in this difference with response number, suggesting that the amount of beat displacement is smaller than the full production error. The decline is not statistically significant, however; it remains for another experiment to determine whether all or only some of the production error is transformed into displacement of the beat. In either case, if subjects judge their response delays relative to the subjective beat, then the beat-displacement effect may help to explain why subjects seem satisfied with their delayed responses in the production task, and why the feedback model fails.

## VIII. SUMMARY

In a series of experiments we have explored the judgment, production, and imitation (reproduction) of time ratios by three professional musicians. In five initial experiments we found that for small fractions in all three procedures our subjects exhibited systematic and substantial errors. In the judgment task they associated small stimulus fractions with names that were too large (*overestimation*). In both the production task (targets specified by fraction name) and the imitation task (targets specified by fractional interval) our subjects produced intervals that were too large (*overproduction*).

The relation between judgment and production errors requires us to reject feedback models of production, in which a subject uses judgment of the time interval from perceived beat click to perceived response (response feedback) to adjust produced fractions. The approximate equality of production and imitation errors, together with the existence of systematic errors in judgment, argues that imitation is not accomplished simply by concatenating the processes used in judgment and production. Instead, we propose a model containing four internal transformations, in which judgment and production share no transformations and imitation shares one transformation with judgment and another with production. Our data permit us to infer relations among the four transformations. Since production and imitation share the same response, data from these two tasks implicitly define a psychophysical scale for fractions of a beat (a function that relates stimulus fractions to their names) that is independent of potential distortions due to response generation; our results indicate that this implicit scale is free of systematic error, unlike the scales for fractions of a beat that are defined (explicitly) by our judgment and production data.

Results from seven additional experiments increase the generality of our findings and help discriminate among alternative explanations of the judgment and production errors. (1) In the judgment task, changing the pitch of the marker click had little effect, indicating that the overestimation we observed is not a consequence of delays in shifting attention from beat click to marker click. (2) Because performance remained approximately invariant as we altered marker duration, we conclude that subjective onsets rather than offsets of marker and beat clicks were used to mark their occurrence times, discrediting an explanation based on the use of marker offsets. (3) When we varied proximity to the beat of the fractional interval being judged and found no effect on the judgment error, we rejected the possibility of a distortion of

subjective time near the beat. (4) The judgment error depends more on the absolute size (duration) of the judged interval than on its size relative to the beat interval (fraction), but both factors have systematic effects; this finding argues against any model in which either interval alone or fraction alone determines the error. (5) The unexpectedly poor judgment precision we encountered when a subject had to judge two disparate intervals within the same stimulus pattern suggests reasons for failure of the feedback model of production; they are based on the possibility that the timing of two such intervals may be required in the production task if a subject attempts both to time the initiation of a response and to judge when it actually occurs.

In the production task, we found that (a) enriching the potential perceptual cues (feedback) that mark responses had little effect on performance, and that (b) when subjects used their musical instruments to perform the task they produced even larger mean errors for small fractions than in productions with finger taps. (c) Evidence from production variability suggests that the overproduction we observed is not a consequence of the existence of a minimum reaction time. In an experiment requiring multiple subdivisions of the beat interval, we found that (d) the production delay did not depend on the number of responses produced within a beat interval, but did depend on withholding a response on the beat that initiates that interval (so the first response is required to occur after the beat). (e) Responses that followed such an initial production, including one that was supposed to be coincident with the following beat, were delayed almost as much as the initial response, suggesting that displacement of the subjective beat accompanies the production error. If subjects judged their response delays relative to the subjective beat, this would provide another explanation for failure of the feedback model of production.

Among the issues that should be addressed in future research is the effect of musical training on performance in these tasks, and the basis of the inconsistencies we observed within and among skilled performers. To help understand the extent to which our results reflect properties of human timing in general rather than musical training it would be desirable to modify our paradigms to examine the performance of subjects unskilled in the use of musical notation. Studies of additional musicians would perhaps illuminate the differences we observed among our three subjects. More measurements are needed that would permit comparison of production and perception of small fractions versus large fractions (small "reverse fractions"). Our results raise the question whether similar errors are manifested in contexts that are more musical and in performance of actual music. If they are, we need to know how musicians reconcile the perception-production conflict suggested by our experiments, and exactly what role is played by response feedback in human performance requiring precise timing.

## ACKNOWLEDGMENTS

We are especially indebted to Max V. Mathews for encouraging our collaboration, to Stephen Monsell, Marilyn L. Shaw, and Robert Teghtsoonian for helpful discussions and advice, and to A.S. Coriell for technical support. We also thank R.D. Luce, D.L. Noreen, R.T. Ollman, S. Roberts, D.A. Rosenbaum, D.L. Scarborough, M.S. Schoeffler, and J. Sloboda for useful comments on earlier drafts, and W.J. Kropfl and M.J. Melchner for technical help.

## GLOSSARY

Here we provide a list of the main symbols used in order of appearance in the text, with brief definitions and numbers of the sections in which they are introduced.

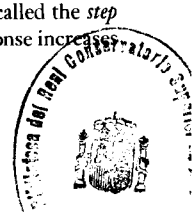
| | |
|---|---|
| $b$ | Beat interval: time interval between one beat and the next. I,B |
| $f$ | Stimulus fraction: ratio of stimulus interval duration to beat interval duration. I,B |
| $n$ | Fraction name specified in numerical and musical notation. I,B |
| $N$ | Fraction-name response in the judgment task. I,B. |
| $\mathbf{J}(N)$ | Judgment function: defines relation between stimulus fraction and fraction-name response, $f = \mathbf{J}(N)$. I,B |
| $F$ | Produced fraction in production and imitation tasks: ratio of delay of timed response to beat interval. I,B |
| $\mathbf{P}(n)$ | Production function: defines relation between fraction-name stimulus and produced fraction, $F = \mathbf{P}(n)$. I,B |
| $\mathbf{I}(f)$ | Imitation function: defines relation between stimulus fraction and produced fraction, $F = \mathbf{I}(f)$. I,B |
| PMF | Abbreviation of the term "psychometric function," a function that associates with each stimulus fraction the proportion of "too large" judgment. II,A |
| $\mathbf{G}^{-1}$ | If $y = \mathbf{G}(x)$ denotes an operation that converts $x$ to $y$, the operation that converts $y$ to $x$ is denoted $\mathbf{G}^{-1}$: $x = \mathbf{G}^{-1}(y)$. (Requires that there is only one value of $x$ for each $y$.) II,A |
| $\ln(x)$ | Natural logarithm of $x$. II,A |
| SD | Standard deviation. II,C |
| DL | Difference threshold: half the change in stimulus fraction required to change the proportion of "larger than" responses from .25 to .75. II,C |
| $\mathbf{S}'(0)$ | Produced fraction in synchronization task. III,A |
| $\mathbf{P}'(n)$ | Production function uncorrected for response delays: $\mathbf{P}(n) = \mathbf{P}'(n) - \mathbf{S}'(0)$. III,A |
| $\mathbf{I}'(f)$ | Imitation function uncorrected for response delays: $\mathbf{I}(f) = \mathbf{I}'(f) - \mathbf{S}'(0)$. IV,C |
| RT | Reaction time. III,F |
| $\mathbf{T}_{fa}$ | Transformation of stimulus fraction to analog representation. IV,B |
| $\mathbf{T}_{aN}$ | Transformation of analog representation to fraction-name response. IV,B |
| $\mathbf{T}_{na}$ | Transformation of fraction-name stimulus to analog representation. IV,B |
| $\mathbf{T}_{aF}$ | Transformation of analog representation to produced fraction. IV,B |
| $\mathbf{E}$ | Identity (or equality) transformation, $\mathbf{E}(x) = x$. IV,B |

## APPENDICES

### A. Staircase and Constant Stimulus Methods

#### *1. Collection of Staircase Data*

The "staircase" or "up-and-down" procedure used in Experiments 1, 6, and 9 can be regarded as a method of collecting observations in which the stimulus presented on each trial depends on both the stimulus and the response of the immediately previous trial. On the first trial in this procedure a fractional interval ($bf$) either longer or shorter than the "correct" or "target" interval ($bN$) is presented. If the subject presses the key indicating that he or she judged the presented fraction to be "too small" relative to the fraction-name target, a longer interval is presented on the next trial. If the subject presses the "too large" key, a shorter interval is presented on the next trial. The subject is required to make one of these two responses; no "equal" response is available. The amount by which the interval is changed is called the *step size*. Use of the procedure depends on the assumption that the probability of a "too large" response increases monotonically with interval duration.



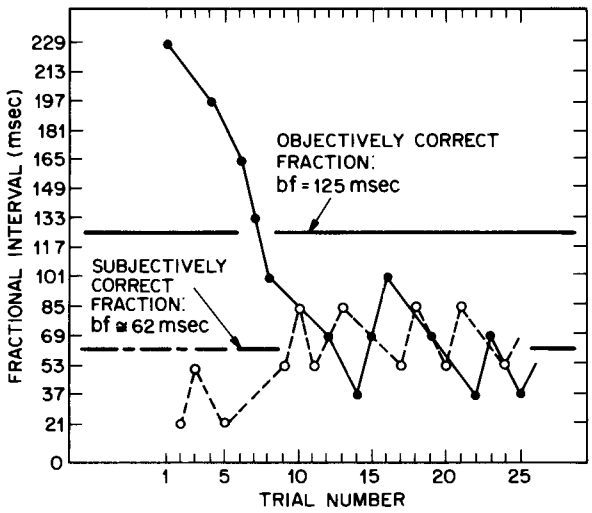



**Fig. 9.** Example of staircase procedure used in Experiment 1. The subject was PB. The fractional interval presented on a trial is shown versus trial number. Two staircases with regularly interleaved stimulus values were randomly interleaved across trials. Value of the stimulus presented on each trial depended on the history of the staircase selected on that trial. The line leaving a point rises if the response was "too small" and falls otherwise.

The immediate result of applying this stimulus selection rule is to choose an interval for each trial so as to reduce the likelihood of the previous response; the long-term result is to increase the number of presented intervals in the region where the subject is maximally uncertain (proportions of the two judgments approximately equal to each other and to .50), and where the proportion of "too large" judgments increases from being less than .50 to more than .50. We assume that in this region the presented fraction is *subjectively* close in value to the target fraction. In practice more reliable data are obtained if two or more independent staircases are randomly interleaved over trials and contain interleaved arrays of stimulus values. Under these conditions a large step size (e.g., $2.5\sigma$, where $\sigma$ is the SD of the PMF) has the virtue of producing rapid convergence with minimal bias. (See Kappauf 1967, 1969; Levitt, 1971; and references therein.)

Figure 9 illustrates 25 trials of the initial staircase procedure of Experiment 1 with data collected from PB when he was judging fractions relative to $N = 1/8$. On each trial the subject had to judge whether the presented fraction was too small or too large relative to 1/8 of a beat. The fractional interval ($bf$) is represented on the ordinate; a horizontal line marks the correct value ($bf = 125$ msec) for $N = 1/8$ and $b = 1000$. Trials are represented on the abscissa. Two staircases started at intervals much smaller and larger than 125 msec. Staircase 1 (broken line) started at 21 msec, and staircase 2 (solid line) started at 229 msec. For each staircase the step size was 32 msec, and the arrays of stimulus values were interleaved, giving a pooled array of stimulus values with 16-msec spacing. The selection of which staircase to use on a trial was random. If an interval was judged too small (large) relative to the target fraction, the next time that staircase was selected a larger (smaller) fraction was presented. In this manner, the staircases converged to a region where the subject was maximally uncertain. In the case shown in Fig. 9, the interval subjectively equivalent to $N = 1/8$, estimated by the mean of the PMF, was 62 msec.

We used two successive staircase procedures in Experiments 1, 6, and 9. First we ran 25 trials with only two interleaved staircases (as exemplified in Fig. 9) to obtain rapid convergence. The starting values (one smaller and one larger) were chosen to be symmetric about $bN$; 50 additional trials were then run with four interleaved staircases to provide the data we used to estimate the PMF. The new array of stimulus values and the four starting points were determined from the estimated mean and standard deviation ($\hat{\sigma}$) of the PMF for that fraction, based on the first 25 trials. The step size of each staircase was adjusted to be $2.5\hat{\sigma}$,

the pooled array of stimulus values had a spacing of $2.5\hat{\sigma}/4$, and the four starting points were separated from the mean by −9, −2, +4, and +9 step values. By widely distributing the starting points of the 50-trial series about the mean based on the initial 25-trial series we hoped to minimize bias. The proportion of "too large" responses was calculated for each fractional interval that had been presented,to obtain an empirical PMF.[57]

### 2. *Psychometric Function*

The plotted points at each stimulus (ordinate) value in Fig. 9 can be used to estimate a PMF. Thus, the fractional interval *bf* = 69 msec was presented on four occasions (trials 12, 15, 19, and 23). The response was "too large" on all but one (trial 15). (Only on that trial is the next *bf* value in that staircase increased.) The estimated PMF value at *bf* = 69 msec is therefore the proportion .25.

Although the sample based only on the illustrative data shown in Fig. 9 is very small (with only 21 judgments in the range from $21 \leq bf \leq 101$ msec) the estimated PMF shows no reversals: stimulus values *bf* (in msec), together with their associated sample sizes *k* and proportions *p* of "too large" responses, in the form (*bf*, *k*, *p*), are (21, 2, .00), (37, 3, .00), (53, 6, .17), (69, 4, .75), (85, 4, 1.00), and (101, 2, 1.00).

### 3. *Constant-Stimulus Method*

In the method of constant stimuli of Experiments 2, 7, and 8, the fractional interval on each trial was selected randomly without regard to the subject's prior responses from a set of intervals chosen in advance. In Experiments 2 and 7, for example, each set contained 24 different intervals, covering a wide range and with a spacing that increased with interval size.

## B. Additional Details of Design and Procedure

In this section we mention some details of design and procedure of our 12 experiments not discussed elsewhere.

*Experiment 1.* The order of the judgment, production, and imitation procedures in Experiments 1, 3, and 5 was balanced in a 3×3 Latin square design within each of the two replications: each procedure was studied first, second, and third for some subject. For each subject the order of the three procedures was reversed between the first and second replications. (The Latin squares in the two replications were therefore mirror images.) Three distinct orders of fraction names were paired with the procedures so that the complete design for each replication was a Graeco-Latin square. In each replication for each fraction name, an initial 25 practice trials with only two interleaved staircases provided approximate estimates for the mean and variance of the PMF. Following a brief pause, 50 additional trials were run with four interleaved staircases.

*Experiment 2.* Fractional intervals were varied by the method of constant stimuli. On each trial the interval between beat click and marker click was determined by random selection without replacement from a set of 24 intervals. The subject was given a brief rest after each cycle through the 24 intervals; 7 such cycles defined a session. To reduce the impact of any effects associated with specific intervals, four different sets of 24 intervals were used in the course of the experiment. (The smallest and largest interval differed from set to set; over all sets intervals ranged from 43 to 891 msec.) The spacing between successively larger intervals increased approximately in the same way (harmonically) as the spacing between successively larger fraction names (1/8, 1/7, ..., 1/2). The first cycle of each session was considered

[57] A staircase procedure is typically designed to produce data that estimate a particular quantile of the PMF, such as the 50% point (the median), rather than the entire PMF. However, results from a Monte Carlo study (Sternberg & Knoll, unpublished) show that the empirical PMF, derived from the application of the staircase procedure to a known PMF, shows little distortion when independent staircases contain interleaved grids of step values, the step size of individual staircases is large, and the underlying PMF is symmetric.

practice and was excluded from the analysis; a session therefore contributed 6 test trials for each of 24 intervals to the analysis. The number of sessions run for SB, PF, and PZ was 4, 3, and 5, respectively.

*Experiment 3.* Each of two replications of each of the eight production conditions involved 25 trials and was run as part of the balanced Latin square design described for Experiment 1. Eight replications of the synchronization condition and four additional replications of the production condition with $n = 1$ were run after the main experiment; performance in the production condition indicated no systematic change relative to the main experiment.

*Experiment 4.* Principal conditions included synchronization and eight $n$-values. Each replication of a condition included 25 trials, or 250 responses. Number of replications per condition per subject varied from 1 to 4 and averaged approximately 2; conditions were run in an irregular order. Plots of results for individual subjects (Figs. 2–4) also show results for conditions ($n$ = 1/5, 1/3) run on only a subset of the subjects.

*Experiment 5.* Each of two replications of each of the eight imitation conditions involved 25 trials and was run as part of the balanced Latin square design described for Experiment 1.

*Experiment 6.* The staircase procedure was run as in Experiment 1. Either one or two replications were run per condition; the order of conditions differed across subjects.

*Experiment 7.* The constant stimulus method of Experiment 2 was used with one session run for each of the two conditions. Intensity of the prolonged marker was reduced from 30 dB above threshold to about 21 dB to make the loudness of brief and prolonged markers more similar.

*Experiment 8.* In each cycle of 24 trials each stimulus was presented once. Seven cycles defined a session; one session was run per subject. The first cycle of a session was considered practice; each stimulus is therefore represented six times in the data for an individual subject.

*Experiment 9.* The order of the three beat intervals was balanced across subjects in a 3×3 Latin square design. In the first replication the order of the four $N$-values differed across subjects, but for an individual subject the order was the same for all three beat intervals. The four $N$-values within each beat interval were run consecutively before the beat interval changed. In the second replication the order of the 12 conditions was reversed for each subject. Use of the staircase procedure was similar to that in Experiment 1 except that in the second replication the initial 25-trial series was omitted and starting points were based on performance in the first replication.

*Experiment 10.* We asked subjects to choose a note that they felt they could produce consistently: PZ bowed an open A string on the violin, SB played C above middle C on the flute, and PF bowed D below middle C on the cello. All subjects started with a single replication of 25 trials with $n$ = 1/4 for practice and then completed two replications each for $n$ = 1/8, 1/2, and 1; the order of fraction names was varied between subjects and was reversed from the first to second replication within each subject.

*Experiment 11.* Two replications (of 25 trials) were run, first in the repeated-response procedure and then in the one-response procedure.

*Experiment 12.* Each condition was studied in three replications (of 25 trials). Within each subject the order of conditions was balanced. All three subjects are right-handed; SS and PZ performed response $R_2$ with their right index fingers, and JM performed $R_2$ with her left index finger.

## C. Measures of Location of the Psychometric Function

Both staircase and constant-stimulus methods in our perceptual judgment experiments gave rise to PMFs. Among our considerations in choosing a location measure were (1) our desire to reduce nonadditive distortions in estimates of the effects of interest on underlying processes, (2) our desire to compare the location measure to corresponding measures in production and imitation with the idea that a common underlying process might contribute to all three, (3) our reluctance to assume a functional form for the PMF (such as cumulative Gaussian), (4) the smallness of sample sizes at each stimulus value for individual subjects, and (5) our interest in a spread measure as well.

It is convenient to regard the "true" PMF as a cumulative distribution function and the observed PMF as

an estimate of this function. The most common location measure used in psychophysics is the *median* (the 50% point) of the PMF. Instead, we used an estimate of the *mean;* comparisons of the two revealed extremely good agreement.

It may be helpful first to consider why we favor the mean over the median for the production and imitation experiments. For production, suppose that the time between beat click and timed response is the sum of a timing component $X_t(n)$ that interests us and that varies with fraction name $n$ and other components such as input and output delays, $X_s$ and $X_r$, that vary from trial to trial but do not depend on target fraction: $X(n) = X_t(n) + X_s + X_r$; Var $(X_s + X_r) > 0$. Then changes with $n$ in the *mean* of $X(n)$ accurately reflect changes in the mean of the process of primary interest $X_t$, whatever the distributions of $X_t$ and $X_s + X_r$. (The mean of a sum of random variables equals the sum of the means.) This property does not characterize the median, however. It follows, for example, that if an estimate of the mean of $X_s + X_r$ is available, it can be used to "correct" the observed $X(n)$; this is not possible, in general, for the median. The same argument, with $f$ substituted for $n$, applies to imitation. Similarly, the PMF may reflect processes such as $X_s$ and $X_r$ as well as a timing process that interests us; given plausible assumptions the mean is preferable there also. Furthermore, if we use the mean for production and imitation experiments and wish to compare results across experiments, the mean becomes the favored statistic for judgment experiments.

We used the Spearman-Kärber (S-K) method to estimate PMF means (Spearman, 1908; Epstein & Churchman, 1944; Church & Cobb, 1973). Let stimulus fractions be $f_i$, $i = u, u+1, \ldots, v$, and let $p_i$ be the proportion of "too large" responses for stimulus $f_i$, which estimates a corresponding response probability $Pr\{L;f_i\}$. Let $f_u$ be a stimulus such that we can assume $Pr\{L;f\} = 0$ for $f < f_u$, and let $f_v$ be a stimulus such that we can assume $Pr\{L;f\} = 1.0$ for $f \geq f_v$. Then the S-K estimate of the $r^{th}$ raw moment of the PMF is

$$m_r' = \sum_{i=u}^{v} (p_i - p_{i-1}) \left( \frac{f_1^{r+1} = f_{i-1}^{r+1}}{(r+1)(f_i - f_{i-1})} \right) ;$$

the estimated mean $m_1'$ is obtained by setting $r = 1$.[58,59] The PMF mean can be regarded as a weighted stimulus average, where the weighting function is the derivative of the PMF, a measure of the sensitivity of the response probability to changes in the stimulus. Stimulus values in a region where the PMF rises more steeply contribute more to its mean.

We prefer to avoid strong assumptions about the form of the true PMF, or even about its symmetry. (If the PMF were symmetric then the sample mean and median would estimate the same quantity, of course.) We do feel justified in assuming that the true PMF is nondecreasing. Because our sample sizes are small, however, the *empirical* PMFs are occasionally nonmonotone. In such instances, we have used monotone regression to estimate the best-fitting (least squares) set of nondecreasing proportions $\{p_i^*\}$ from the empirical PMF $\{p_i\}$ before estimating parameters. (See Ayer, Brunk, Ewing, Reid, & Silverman, 1955; Kruskal, 1964; and de Leeuw, 1977.)[60,61]

For the data from Experiment 1, we compared and found excellent agreement between conventional median estimates obtained by applying linear interpolation to the $\{p_i^*\}$, and S-K mean estimates. The two replications for each of three subjects provided six sets of PMFs, with 8 PMFs per set. For the six sets of data, linear correlations of means versus medians ranged from .9995 to .9998, and slopes of the linear regression of means on medians ranged from .999 to 1.018.

[58]This is actually a modified S-K estimator, appropriate for a continuous distribution function (a piecewise linear integrated histogram) rather than a discrete distribution function.

[59]In our application of this method, as shown by the estimation equation, the proportions $p_i$ enter with equal weights. An alternative would be to weight them by considering differences in sample size and binomial variability. Fortunately, for PMFs derived from staircase data, the distribution of observations is approximately triangular and is centered where $p_i \cong .5$; this compensates approximately for differences in binomial variability and justifies the use of equal weights.

[60]This "monotonizing" procedure can probably be improved upon by regarding the $\{p_i\}$ as estimates of values of an underlying continuous rather than a discrete distribution function.

[61]Note that at least one advantage of applying monotone regression is that without it the more conventional quantile measures of location and spread may uot be uniquely defined. If ouly the S-K estimate of location is desired, the monotonizing transformation is not necessary, since it does not alter the estimated mean. We have used it because it influences the spread measure (SD estimate) based on the S-K method and is necessary to permit comparison of quantile with moment estimators.



## D. Measures of Spread of the Psychometric Function

The conventional measure of spread (or precision) of the PMF is the DL, defined as half of the interquartile range. We obtained DLs for the PMFs in Experiments 1 and 2 by applying linear interpolation to the $\{p_i^*\}$. We compared these quantile measures to a set of corresponding SD estimates; these were based on $m_1'$ and $m_2'$ values obtained by applying the S-K method to the same $\{p_i^*\}$. (See Epstein & Churchman, 1944; Chmiel, 1976.)

Six sets of PMFs were used for the comparison, obtained from the three subjects in each of the two experiments. For each set of PMFs we determined a constant $k$ such that $\overline{SD} = k\overline{DL}$, where $\overline{SD}$ and $\overline{DL}$ are means over the set of PMFs. (The relation between SD and DL depends on the shape of the PMF. For a Gaussian PMF, for example, $k = 1.48$. The mean of the six $k$-values we obtained is 1.53, suggesting that our PMFs are approximately Gaussian.) Our aim was to compare the variabilities of the two spread measures. This requires first adjusting them to have the same mean; we did so by multiplying the DLs in each set by the $k$-value obtained for that set. For each set of PMFs, we determined the variance of each spread measure SD and kDL across replications of the same condition (fraction name) and pooled these variances over fractions. The ratios of Var(kDL) to Var(SD) ranged from 1.5 to 11.0 over the six sets of PMFs, with a mean of $5.7 \pm 1.7$. This surprising finding indicates convincingly that, at least for our small-sample PMFs, the conventional quantile measure of spread (the DL) is far less reliable than the S-K estimator of the SD, and supports our use of the SD in the present report.

## E. Stimulus-Averaging versus Response-Averaging Methods for Deriving a Psychophysical Scale

Our method of determining the judgment function, $f = \mathbf{J}(N)$, was to associate with each value of $N$ the mean of the corresponding PMF. For each of a set of response ($N$) values this method can be regarded as providing a weighted mean stimulus, where the weighting function is given by the steepness (derivative) of the PMF. Stimulus averaging has at least three virtues relative to the more conventional response-averaging method (to be discussed below): (1) it can be applied to data from single-fraction as well as multiple-fraction experiments; (2) it does not require us to treat response values as being defined on an interval scale, as do response-averaging methods; and (3) each PMF mean presumably reflects only one subjective criterion because it is associated with only one $N$-category boundary. A response-averaging method depends, in general, on associating a set of more than two $N$-values with each $f$-value; the response average therefore reflects more than one criterion on the subjective $N$-scale. Such a method seems less comparable with the straightforward analysis of production and imitation, in which performance in any condition presumably depends on only one criterion, as in the single-fraction judgment experiment. Oyama (1969) described a similar stimulus-averaging method applied to magnitude estimates of loudness, and argued in its favor by noting virtue (2) above.

Despite these advantages it seemed important to check whether the results we obtained in the judgment experiment could be attributed to the unorthodox method we used to construct the judgment function. Data from the multiple-fraction judgment procedure of Experiment 2 permitted comparison of the two methods. Each of the six central response alternatives in that experiment were defined in terms of two category boundaries, such as "between 1/8 and 1/7." To average responses, we defined the value of each response to be the geometric mean of its two boundaries; the set of such response values associated with each stimulus fraction $f$ was then averaged, again using the geometric mean.[62]

Seven stimulus intervals were defined, containing approximately equal numbers of fractions; the set of average responses within each interval were themselves averaged (using geometric means) and associated with the geometric mean stimulus fraction. The $\mathbf{J}(N)$ functions are truncated in the response-averaging

[62]Use of the geometric mean is common in response-averaging methods; it is equivalent to applying the arithmetic mean to logarithms of response values and is thought to be appropriate when the psychophysical scale is linear on logarithmic coordinates, as in the case of a power function.

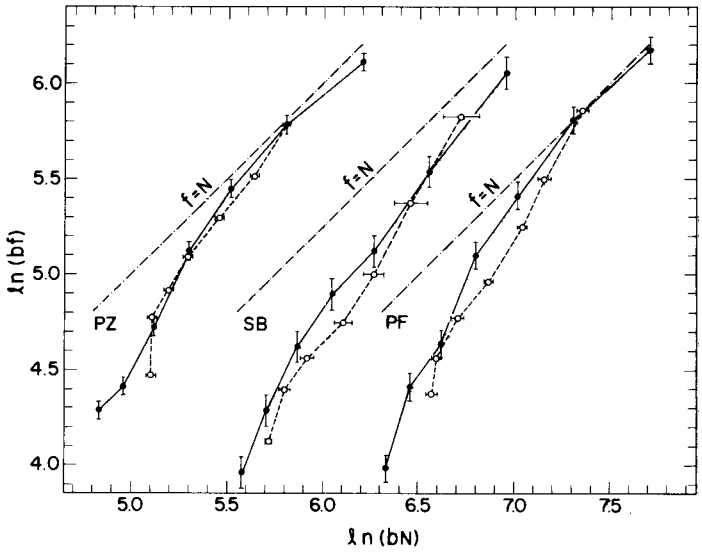



**Fig. 10.** Comparison of psychophysical scales derived by two different averaging methods from the multiple-fraction judgment experiment. The scale relates fractional intervals (*bf*) specified by time-pattern stimuli to fraction names expressed as intervals (*bN*). Data for SB (PF) are displaced .75 (1.5) log units to the right. Values derived from PMFs (weighted stimulus averaging for specified *N*) are shown by filled circles and unbroken lines. Values derived by response averaging (for specified *f*) are shown by open circles and broken lines. Lines representing $f=N$ are included for reference. Bars represent ±1 SE; SEs are based on variability across replications in the quantity averaged.

method because a stimulus must be excluded if any response to that stimulus falls in one of the end categories ("less than 1/8" or "greater than 1/2"), for which an acceptable response value cannot be defined.

Results from this procedure are shown in Fig. 10, together with results from the stimulus-averaging (PMF) method. For the PMF method, the eight response categories provide seven category boundaries and thus seven values of **J**(*N*).

Estimates shown of ±SE are based on between-replications SDs. These were pooled over fractions for the PMF method. For the response-averaging method, the SDs were smoothed by linear regression of SDs on means. The difference in SE estimates between the two methods can be regarded primarily as a result of the **J**(*N*) slopes being considerably greater than unity.

The scale based on the PMF method for each subject can be seen to be very similar to the scale based on response averaging. Oyama (1968) drew the same conclusion when he compared stimulus- and response-averaging methods for deriving a scale of loudness.

## REFERENCES

Adam, N., Castro, A., & Clark, D. Production, estimation, and reproduction of time intervals during inhalation of a general anesthetic in man. *Journal of Experimental Psychology*, 1974, *102*, 609–614.

Ayer, M., Brunk, H. D., Ewing, G. M., Reid, W. T., & Silverman, E. An empirical distribution function for sampling with incomplete information. *Annals of Mathematical Statistics*, 1955, *26*, 641–647.

Bregman, A. S. The formation of auditory streams. In J. Requin (Ed.), *Attention and performance VII*. Hillsdale, N.J.: L. Erlbaum Associates, 1978. pp. 63–75.

Carlson, V. R., & Feinberg, I. Individual variations in time judgments and the concept of an internal clock. *Journal of Experimental Psychology*, 1968, 77, 631–640.

Carlson, V. R., & Feinberg, I. Time judgment as a function of method, practice, and sex. *Journal of Experimental Psychology*, 1970, *85*, 171–180.

Chmiel, J. J. Some properties of Spearman-type estimators of the variance and percentiles in bioassay. *Biometrika*, 1976, *63*, 621–626.

Church, J. D., & Cobb, E. B. On the equivalence of Spearman-Kärber and maximum likelihood estimates of the mean. *Journal of the American Statistical Association*, 1973, *68*, 201–202.

Creelman, C. Human discrimination of auditory duration. *Journal of the Acoustical Society of America*, 1962, *34*, 582–593.

de Leeuw, J. Correctness of Kruskal's algorithms for monotone regression with ties. *Psychometrika*, 1977 *42*, 141–144.

Duifhuis, H. Consequences of peripheral frequency selectivity for nonsimultaneous masking. *Journal of the Acoustical Society of America*, 1973, *54*, 1471–1489.

Eisler, H. Subjective duration and psychophysics. *Psychological Review*, 1975, *82*, 429–450.

Eisler, H. Experiments on subjective duration 1868–1975: A collection of power function exponents. *Psychological Bulletin*, 1976, *83*, 1154–1171.

Epstein, B., & Churchman, C. W. On the statistics of sensitivity data. *Annals of Mathematical Statistics*, 1944, *15*, 90–96.

Fastl, H. Subjective duration and temporal masking patterns of broadband noise impulses. *Journal of the Acoustical Society of America*, 1977, *61*, 162–168.

Fraisse, P. *Les structures rythmiques*. Louvain: Publications Universitaires de Louvain, 1956.

Fraisse, P. *The psychology of time*. New York: Harper & Row, 1963.

Fraisse, P. Time and rhythm perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception, vol. VIII: Perceptual coding*. New York: Academic Press, 1978. pp. 203–254.

Getty, D. Discrimination of short temporal intervals: A comparison of two models. *Perception & Psychophysics*, 1975, *18*, 1–8.

Getty, D. Counting processes in human timing. *Perception & Psychophysics*, 1976, *20*, 191–197.

Kappauf, W. E. *Empirical modifications in the up-and-down method and its estimates of* $\mu$ and $\sigma$. Report No. 2, National Institutes of Health Research Grant NB-105576-01, University of Illinois, 1967.

Kappauf, W. E. An empirical sampling study of the Dixon and Mood Statistics for the up-and-down method of sensitivity testing. *American Journal of Psychology*, 1969, *82*, 40–55.

Kornblum, S. Simple reaction time as a race between signal detection and time estimation: A paradigm and model. *Perception & Psychophysics*, 1973, *13*, 108–112.

Kristofferson, A. B. Low-variance stimulus-response latencies: Deterministic internal delays? *Perception & Psychophysics*, 1976, *20*, 89–100.

Kruskal, J. B. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 1964, *29*, 115–129.

Levitt, H. Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 1971, *49*, 467–477.

Michon, J. A. Studies on subjective duration II: Subjective time measurement during tasks with different information content. *Acta Psychologica*, 1965, *24*, 205–212.

Michon, J. A. Magnitude scaling of short durations with closely spaced stimuli. *Psychonomic Science*, 1967, *9*, 359–360.

Miller, R. G., & Halpern, J. W. Robust estimators for quantal bioassay. *Biometrika*, 1980, *67*, 103–110.

Nakajima, Y., Shimojo, S., & Sugita, Y. On the perception of two successive sound bursts. *Psychological Research*, 1980, *41*, 335–344.

Ornstein, R. E. *On the experience of time*. Harmondsworth: Penguin Books, 1969.

Oyama, T. A behavioristic analysis of Stevens' magnitude estimation method. *Perception & Psychophysics*, 1968, *3*, 317–320.

Penner, M. J. Effect of masker duration and masker level on forward and backward masking. *Journal of the Acoustical Society of America*, 1974, *56*, 179–182.

Pöppel, E. Time Perception. In R. Held, H. Leibowitz, & H. Teuber (Eds.), *Handbook of sensory physiology, volume VIII: Perception*. Berlin: Springer-Verlag, 1978. pp. 713–729.

Povel, D. J. Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 1981, 7, 3–18.

Rasch, R. A. Synchronization in performed ensemble music. *Acustica*, 1979, *43*, 121–131.

Richards, W. Time estimates measured by reproduction. *Perceptual and Motor Skills*, 1964, *18*, 929–943.

Schulze, H. The detectability of local and global displacements in regular rhythmic patterns. *Psychological Research*, 1978, *40*, 173–181.

Snodgrass, J., Luce, R. & Galanter, E. Some experiments on simple and choice reaction time. *Journal of Experimental Psychology*, 1967, *75*, 1–17.

Spearman, C. The method of 'right and wrong cases' ('constant stimuli') without Gauss's formulae. *British Journal of Psychology*, 1908, *2*, 227–242.

Steiner, S. Apparent duration of auditory stimuli. *Journal of Auditory Research*, 1968, *8*, 195–205.

Sternberg, S., & Knoll, R. L. The perception of temporal order: Fundamental issues and a general model. In S. Kornblum (Ed.), *Attention and performance IV*. New York: Academic Press, 1973. pp. 629–685.

Svenson, O. Magnitude estimation of time intervals delimited by light flashes of different spatial separation. *Reports from the Psychological Laboratories (Report No. 394)*. Stockholm: University of Stockholm, 1973.

Teghtsoonian, R., & Teghtsoonian, M. Range and regression effects in magnitude scaling. *Perception & Psychophysics*, 1978, *24*, 305–314.

Thomas, E., & Brown, Jr., I. Time perception and the filled-duration illusion *Perception & Psychophysics*, 1974, *16*, 449–458.

Treisman, M. Temporal discrimination and the indifference interval: Implications for a model of the "internal clock." *Psychological Monographs: General and Applied*, 1963, 77 (Whole No. 576).

Van Noorden, L.P.A.S. *Temporal coherence in the perception of tone sequences*. Doctoral thesis, Eindhoven University of Technology, The Netherlands, 1975.

Vorberg, D., & Hambuch, R. On the temporal control of rhythmic performance. In J. Requin (Ed.), *Attention and performance VII*. New Jersey: Lawrence Erlbaum Associates, 1978. pp. 535–555.

Wing, A., & Kristofferson, A. B. The timing of interresponse intervals. *Perception & Psychophysics*, 1973, *13*, 455–460.

Woodrow, H. A quantitative study of rhythm: The effect of variations in intensity, rate and duration. *Archives of Psychology*, 1909, *14*, 1–66.

Woodrow, H. Time perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. pp. 1224–1236.

Zelkind, I., & Sprug, J. (Eds.) *Time research: 1172 studies*. Metuchen, N.J.: The Scarecrow Press, Inc., 1974.

Zwicker, E. Subjektive und objektive Dauer von Schallimpulsen und Schallpausen. *Acustica*, 1969, *22*, 214–218.



# 8

# Intervals, Scales, and Tuning

*Edward M. Burns and W. Dixon Ward*


I. Introduction ..... 241
II. Are Scales Necessary? ..... 243
III. Musical Interval Perception ..... 246
A. Identification of Isolated Musical Intervals ..... 246
B. Adjustment of Isolated Musical Intervals ..... 249
C. Musical Interval Discrimination and Categorical Perception ..... 250
IV. Natural Intervals and Scales ..... 255
A. Natural Intervals and Their Possible Basis ..... 255
B. Natural Scales and Temperament ..... 256
C. Experimental Evidence Relevant to Natural Intervals and Scales ..... 257
D. Octave Generalization and Chroma ..... 262
V. Conclusions and Caveats ..... 264
A. Conclusions ..... 264
B. Caveats ..... 264
References ..... 265


## I. INTRODUCTION

In the vast majority of musical cultures, collections of discrete pitch relationships—musical scales—are utilized as a framework for composition and improvisation. In this chapter we will explore the perception of musical scales concentrating on the basic building block of scales, the musical interval. In addition, we will speculate on possible perceptual and/or physiological bases for scales. Among the questions to be addressed are the following: (1) Are scales necessary (i.e., why are only a relatively small subset of all possible pitch relationships utilized in music?)? (2) What evidence is there for the existence of "natural" intervals and scales? (3) What is the basis for the concept of octave generalization? and (4) What, if any, are the relationships between music perception and other facets of auditory perception, especially speech perception?


Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

TABLE I

Interval Comparison in Different Mathematical Tuning Systems[a]

| Interval name | Solfeggio | Letter notation | Pythagorean tuning (PT) | | | Just intonation (JI) | | | Equal temperament (ET) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Numerical origin | Frequency ratio | Cents | Numerical origin | Frequency ratio | Cents | Frequency ratio | Cents |
| Unison | DO | C | 1 : 1 | 1.000 | 0.0 | 1 : 1 | 1.000 | 0.0 | 1.000 | 0 |
| Minor second | | D♭ | $2^8 : 3^5$ | 1.053 | 90.2 | 16 : 15 | 1.067 | | 1.059 | 100 |
| | | C♯ | $3^7 : 2^{11}$ | 1.068 | 113.7 | 16 : 15 | 1.067 | 111.7 | 1.059 | 100 |
| Major second | RE | D | $3^2 : 2^3$ | 1.125 | 203.9 | 10 : 9[b] | 1.111 | 182.4 | 1.122 | 200 |
| | | | | | | 9 : 8[c] | 1.125 | 203.9 | | |
| Minor third | | E♭ | $2^5 : 3^3$ | 1.186 | 294.1 | 6 : 5 | 1.200 | 315.6 | 1.189 | 300 |
| | | D♯ | $3^9 : 2^{14}$ | 1.201 | 317.6 | 6 : 5 | 1.200 | 315.6 | 1.189 | 300 |
| Major third | MI | E | $3^4 : 2^6$ | 1.265 | 407.8 | 5 : 4 | 1.250 | 386.3 | 1.260 | 400 |
| Fourth | FA | F | $2^2 : 3$ | 1.333 | 498.1 | 4 : 3 | 1.333 | 498.1 | 1.335 | 500 |
| Tritone | | G♭ | $2^{10} : 3^6$ | 1.407 | 588.3 | 45 : 32 | 1.406 | 590.2 | 1.414 | 600 |
| | | F♯ | $3^6 : 2^9$ | 1.424 | 611.7 | 64 : 45 | 1.422 | 609.8 | 1.414 | 600 |
| Fifth | SO | G | 3 : 2 | 1.500 | 702.0 | 3 : 2 | 1.500 | 702.0 | 1.498 | 700 |
| Minor sixth | | A♭ | $2^7 : 3^4$ | 1.580 | 792.2 | 8 : 5 | 1.600 | 813.7 | 1.587 | 800 |
| | | G♯ | $3^8 : 2^{12}$ | 1.602 | 815.6 | 8 : 5 | 1.600 | 813.7 | 1.587 | 800 |
| Major sixth | LA | A | $3^3 : 2^4$ | 1.688 | 905.0 | 5 : 3 | 1.667 | 884.4 | 1.682 | 900 |
| Minor seventh | | | | | | 7 : 4[d] | 1.750 | 968.8 | | |
| | | B♭ | $2^4 : 3^2$ | 1.788 | 996.1 | 16 : 9[e] | 1.777 | 996.1 | 1,782 | 1000 |
| | | A♯ | $3^{10} : 2^{15}$ | 1.802 | 1019.1 | 9 : 5 | 1.800 | 1017.6 | 1.782 | 1000 |
| Major seventh | TI | B | $3^5 : 2^7$ | 1.900 | 1109.8 | 15 : 8 | 1.875 | 1088.3 | 1.888 | 1180 |
| Octave | DO | C | 2 : 1 | 2.000 | 1200.0 | 2 : 1 | 2.000 | 1200.0 | 2.000 | 1200 |

[a] Adapted from Martin, 1962.
[b] Lesser.
[c] Greater.
[d] Harmonic
[e] Grave.

Since many readers may not be conversant with music theory and terminology, some of the relevant concepts of traditional Western music will be reviewed. A concept inherent in scale construction is that of the tonic (or keynote) of the scale: that tone to which the pitches of the other tones of the scale are referenced. The basic unit of the scale is the musical interval, which corresponds physically to the frequency ratio between two tones and perceptually to the pitch relationship between the tones. The adjective "harmonic" when describing musical intervals denotes that the two tones comprising the interval are presented simultaneously (regardless of whether or not the frequencies of the two tones are exact harmonics of a common fundamental frequency); the adjective "melodic" indicates that the two tones are presented sequentially. Also basic to Western theory is the concept of octave generalization (or tone chroma), which dictates that tones separated by an octave (frequency ratio of 2 : 1) are musically similar (have the same chroma), and, hence, a scale may be completely defined by specifying the intervals within an octave.

The present standard inclusive scale (i.e., the scale that includes all of the intervals used in Western music) is the 12-tone chromatic scale of equal temperament. This scale divides the octave into 12 equally spaced intervals. The basic interval unit (i.e., the smallest interval and the spacing between adjacent intervals) is called a *tempered semit* (or *semitone*) and corresponds to a frequency ratio of $2^{1/12}$ : 1. (A useful unit for measuring frequency ratio is the cent: 1 cent = $2^{1/1200}$ : 1; hence, a semit contains 100 cents and an octave contains 1200 cents.) Table I gives the name, solfeggio (vocal mnemonic), letter notation (assuming a tonic of C), frequency ratio, and cent value of the 12 intervals comprising the equally tempered scale. (The Just and Pythagorean scales are contenders for the title of "natural scale" and will be discussed in Section IV,A.) In Western music the absolute frequencies of the tones are also specified by defining a standard reference frequency (presently, "middle A" = 440 Hz) and using one of various schemes to notate octave ranges (Olson, 1952). Most compositions in Western music do not employ all of the intervals of the chromatic scale but are based on seven-interval subsets of the chromatic scale called *diatonic scales*, in which the interval between adjacent tones is either one or two semits (e.g., the diatonic major scale consists of the intervals, 200, 400, 500, 700, 900, 1100, and 1200 cents).

## II. ARE SCALES NECESSARY?

Given that present Western music utilizes a relatively small set of discrete pitch relationships, an obvious question occurs: Is this use of discrete intervals universal? That is, are there musical cultures that employ continuously variable pitches, analogous, for example, to the rising or falling tones of some tone languages? The evidence from ethnomusicological studies indicates that the use of discrete pitch relationships is essentially universal. The only exceptions appear to be certain primitive musical styles—for example, "tumbling strains" (Sachs, 1961) or "indeterminate-pitch chants" (Malm, 1967)—which are found in a few tribal cultures. Of course, pitch glides—glissandos, portamentos, trills, etc,—are used as embellishment and ornamentation in

most musical cultures. However, they are not an essential part of the structure of these musics. The concept of octave similarity, although far from universal in early and primitive music (e.g., Nettl, 1956; Sachs, 1962), also seems to be common to more advanced musical systems.

A related question follows: Does the 12-note chromatic scale represent a norm or a limit to the number of usable pitch relationships per octave? Several Western composers have written compositions in quartertone (i.e., 24 approximately equal intervals per octave) and other microtonal scales, but none of these scales have gained wide acceptance. There are numerous cultures whose scales contain less than 12 notes per octave. There are, however, apparently only two musical cultures that, in theory, utilize more than 12 intervals per octave: the Indian and the Arab–Persian. Both musical systems of India (Hindustani and Karnatic) are, according to tradition, based on 22 possible intervals per octave. They are not equal (or approximately equal) intervals, however. The basic structure of the inclusive scale is essentially the same as that of the Western 12-interval chromatic scale, and the microtones ("shrutis") are (theoretical) small variations of certain intervals, the exact values of which are dependent on the individual melodic framework ("raga") being played. There is evidence that in actual musical practice these theoretical variations are not played as discrete intervals but are denoted by a purposefully induced variability in intonation (e.g., a slow vibrato). (Jhairazbhoy & Stone, 1963; Callow & Shepherd, 1972).

The one system that may utilize true quartertones (i.e., intervals that bisect the distance between the Western chromatic intervals) is the Arab–Persian system. In this system there are various claims as to the number of possible intervals (ranging from 15 to 24) and some controversy as to whether they are true quartertones or, as in the Indian system, merely microtonal variations of certain intervals (e.g., Zonis, 1973). The limited data on measured intonation in this system are ambiguous as to how accurately these quartertones are produced in practice (e.g., Caron & Safvate, 1966; Spector, 1966). It is clear, however, that neither the Indian nor Arab–Persian scales are chromatically microtonal. That is, the quartertones (or microtones) are never played contiguously but only as alternative versions of certain larger intervals.

Thus, the evidence indicates that the 12-interval chromatic scale may, indeed, reflect some sort of limit on the number of intervals per octave that are of practical use in music or, at least, that the semit is probably the smallest usable separation between successively played tones. If we assume this limitation exists, is there any obvious perceptual basis for it or for the use of discrete intervals in general? Authors of books on musical perception often contrast the number of just noticeable differences (JNDs) in frequency to the step-size in the chromatic scale—from 20 to 300 JNDs per semit, depending on the frequency range and experimental paradigm—and point out the apparent discrepancy between the number of tones that can be "distinguished" and the number actually used in music. There are several reasons why this is not a meaningful contrast.

First, it is evident from music theory, from everyday musical experience, and from a number of experiments (e.g., White, 1960; Attneave & Olson, 1971) that frequency ratio, rather than frequency per se, is the primary mediator of melodic information.

Hence, frequency ratio JNDs, rather than frequency JNDs, are the more relevant comparison with scale step-size. Frequency-ratio JNDs for highly trained observers are an order of magnitude greater than frequency JNDs obtained in the same frequency region, using equivalent psychophysical paradigms (see Section III,C). Trotter (1967) has also pointed out that there is little correlation between the inability to perceive interval or melodic relationships ("tone deafness") and frequency discrimination ability.

The second, and probably more important, reason for the apparent discrepancy between JND magnitude (whether frequency or frequency ratio) and scale step-size is the limitation on the processing of information by sensory systems that is imposed by higher level processing (e.g., involving attention and memory). One useful concept in this regard (although certainly not the complete explanation for a complex experience such as music with its cultural and experiential overlays) is that of information theory, in which information is related to the uncertainty and complexity of the stimuli, and information transmitted is related to the reduction in uncertainty. The JNDs mentioned above were determined in minimal-uncertainty (i.e., low information) psychophysical procedures and, as such, probably reflect (at least for the case of frequency JNDs) the resolving power of the peripheral auditory system rather than any limitations imposed by higher level processing. The effect of stimulus uncertainty is illustrated in experiments by Watson, Kelly, and Wroton (1976), in which frequency JNDs were increased by a factor of 30 by paradigm manipulations that presumably increased stimulus uncertainty. Numerous studies have indicated (e.g., Estes, 1972) that when faced with high information signals and/or high information rates, observers tend to encode the information into categories as a means of reducing the information load. It is probable that such factors have dictated both the use of discrete pitch relationships (rather than continuous glides and sweeps, for example) and the number of practical discrete categories.

In a classic paper Miller (1956) reviewed experiments that tested the ability of observers to categorize stimuli in several sensory modalities. He concluded that for stimuli varying on a single physical dimension, the information transmitted to observers (i.e., their information channel capacity) was on the order of 2.8 bits in all modalities. That is, they were only able to place the stimuli, without error, into a maximum of about $2^{2.8}$ categories ("the magic number $7 \pm 2$"). This was contrasted with the ability of observers to discriminate several thousand stimuli along the same continuum in forced-choice discrimination tasks. The discrepancy between the magnitude of frequency ratio JNDs and the minimum interval size in musical scales thus appears to be just another example of this classic discrepancy between resolution in identification and discrimination tasks.

A number of investigators have formulated perceptual models based on differential modes of memory operation that predict this discrepancy. One example is the model of auditory intensity resolution proposed by Durlach and Braida (1969). This model, which is basically a version of Thurstone's "law of categorical judgment," divides the noise that limits resolution into three factors: sensation noise (dependent only on the properties of the stimulus) and two paradigm-dependent components of memory

noise (corresponding to two modes of memory operation). In the "sensory-trace" mode the observer is assumed to attempt to retain an image of the stimulus, and the noise in this mode is related to (a primarily time-dependent) interference with this memory trace. In the "context-coding" mode the subject attempts to remember an imprecise verbal code of the relationship of the individual stimulus to the context of stimuli in the experiment (e.g., the use the stimuli at the extremes of the range as perceptual anchors). The noise associated with this mode is assumed to be dependent only on the range of stimuli. For single- (temporal) interval paradigms (e.g., absolute identification) the subject is assumed to operate in the context-coding mode. The model assumes that in a two-interval discrimination paradigm the subject operates in the mode (or combination of modes) that optimizes performance. The improved resolution in the typical two-interval, two-alternative discrimination task relative to a one-interval, multi-alternative identification task is predicted by this model on the basis of the relative dominance of context-coding noise for the usual experimental parameter values.

## III. MUSICAL INTERVAL PERCEPTION

### A. Identification of Isolated Musical Intervals

#### *1. Absolute Identification*

One obvious exception to Miller's $7 \pm 2$ rule for categorization is the performance of "possessors" of absolute pitch (i.e., persons who have the ability to identify the pitch of a single tone—usually in terms of musical scale categories or keys on a piano). The best subjects (see Chapter 14 by Ward and Burns) are able to identify perfectly about 75 categories (roughly 6.2 bits of information) over the entire auditory range compared to about 5 categories (2.3 bits) for nonpossessors (Pollack, 1952).

It would also appear that musicians who have relative pitch—the ability to identify frequency ratios in terms of musical interval categories—are also exceptions to the $7 \pm 2$ rule. Clearly, most competent musicians can recognize perfectly the 12 intervals of the chromatic scale in either harmonic or melodic modes (Plomp, Wagemarr, & Mimpen, 1973; Killam, Lorton, & Schubert, 1975). In informal experiments we have found that many musicians can identify perfectly the ascending and descending melodic intervals from unison to major tenth (32 categories). However, in light of octave generalization, it is not clear how these results should be interpreted. For example, Pollack (1953) has shown that adding independent dimensions increases the information transfer. Therefore, if chroma and tone height are indeed independent dimensions (see Section IV,D,1), information transfer for identification of frequency ratios exceeding an octave would not be expected to follow the $7 \pm 2$ rule for identification or unidimensionally varying stimuli. A similar problem of interpretation exists for identification of ascending and descending melodic intervals.

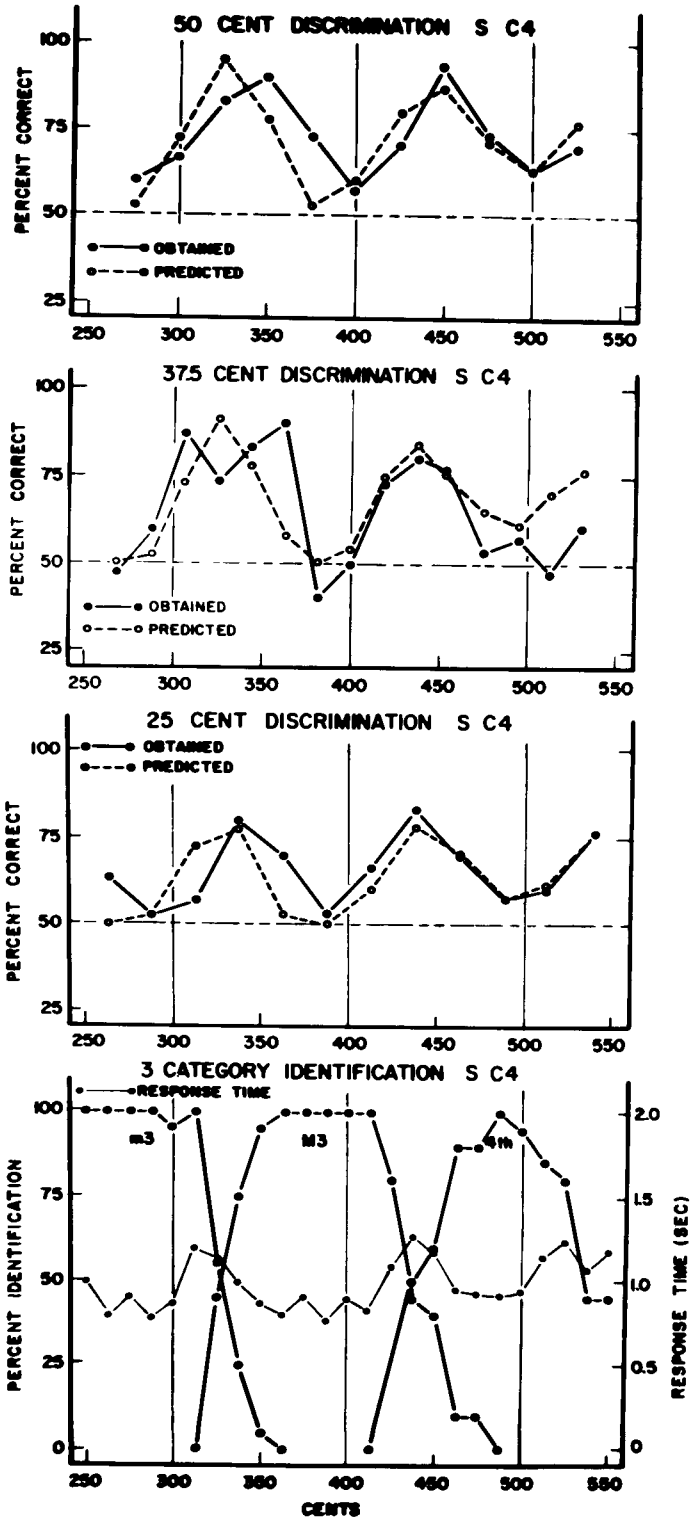
### 2. *Category Scaling*

Another apparent exception to the type of identification performance reported by Miller is performance in category-scaling paradigms using certain kinds of stimuli, primarily stimuli for which the observers possess highly-overlearned familiar labels. The most widely reported experiments of this type are those in which stimuli varying along a single acoustic dimension are categorized by observers in terms of phonemic labels from their native language. Observers in these experiments show very sharp category boundaries and high test-retest reliability (see Studdert-Kennedy, Liberman, Harris, & Cooper, 1970). More importantly, their perceptual categories are much less susceptible to contextual effects characteristic of stimuli that obey the $7 \pm 2$ rule (see Parducci, 1974; Lim, Berlinder, & Braida, 1977), although they are not completely immune from such effects (see Brady & Darwin, 1978; Diehl, Elman, & McCusker, 1978).

A number of experimenters have obtained category-scaling identification functions for melodic intervals spaced in increments of 10 to 20 cents, over ranges of 2 to 5 semits, where the labels are the relevant intervals from the chromatic scale (Burns & Ward, 1974, 1978; Rakowski, 1976; Siegel and Siegel, 1977a,b). Results for a typical musically trained subject (Burns & Ward, 1978) over the range from 250 to 550 cents (category labels: minor third, major third, and fourth) are shown in the lower graph of Fig. 1 (see p. 248). The form of the identification functions shows the sharp category boundaries that are also characteristic of the results of speech-token category-scaling experiments.

In contrast to the type of identification functions shown in Fig. 1, Siegel and Siegel (1977a) report data obtained from musically untrained observers in a similar task that show inconsistent, often multimodal, identification functions with large category overlap and poor test-retest reliability. The only stimuli categorized consistently are those at the extremes of the stimulus range. Siegel and Siegel also measured the effect on categorization of a contextual change (shift in the stimulus range); they found a large effect for the nonmusicians and essentially no effect for the musicians. This contextual independence was also evident in some data reported by Burns (1974).

Some general findings evident in the data of all experimenters are: (1) a tendency for observers to hear narrow intervals (intervals less than a fourth) as perceptually wider (i.e., a compression of the scale relative to equal temperament) and to hear wide intervals (greater than a fourth) as perceptually narrower (i.e., a stretch of the scale), and (2) large but reliable idiosyncratic differences between observers in their perception of the relative width and placement of interval categories. These effects appear for both ascending and descending intervals. Observers are not, in general, able to categorize reliably the stimuli to a finer degree than chromatic semits. For example, they are not able to identify consistently quartertones between chromatic semits (Burns, 1977; Burns & Ward, 1978) or to label consistently stimuli as "low," "pure," or "high" tokens of a single interval (Burns & Ward, 1974; Rakowski, 1976) even when the stimuli are limited to a very narrow range (Szende, 1977). (Bimodal distributions


50 CENT DISCRIMINATION S C4
PERCENT CORRECT
OBTAINED
PREDICTED
37.5 CENT DISCRIMINATION S C4
PERCENT CORRECT
OBTAINED
PREDICTED
25 CENT DISCRIMINATION S C4
OBTAINED
PREDICTED
PERCENT CORRECT
3 CATEGORY IDENTIFICATION S C4
RESPONSE TIME
m3
M3
Mm
PERCENT IDENTIFICATION
RESPONSE TIME (SEC)
CENTS

for "low" and "high" are not uncommon.) This inability to categorize the isolated intervals more precisely is also true of Indian musicians (Burns, 1974a, 1977), whose scales theoretically include microtonal variations of certain intervals. The above-mentioned category-scaling results were all obtained for melodic intervals composed of sinusoids. Similar identification functions have been obtained for harmonic intervals (Halpern & Zatorre, 1979).

### *3. Magnitude Estimation*

Magnitude estimation procedures (i.e., identification with unlimited categories) have also been used to assess musical interval perception (Siegel & Siegel, 1977b). The magnitude estimation functions of musicians typically are step-like. That is, frequency ratios over a small range are estimated to have the same magnitude; then there is an abrupt transition to another estimate. The ranges over which the estimate is constant correspond roughly to semits. In addition, the function relating the standard deviation of repeated magnitude estimates to frequency ratio has a multimodal character, in which the modal peaks correspond to the regions between the plateaus of the magnitude estimation function. These functions are unlike those found for magnitude estimation of stimuli obeying the $7 \pm 2$ rule, which are typically smooth, monotonically increasing functions of stimulus magnitude both for magnitude estimates and for standard deviation of repeated estimates (see, e.g., Stevens, 1976). The results do, however, resemble those found for certain speech stimuli (Vinegrad, 1972) and will be further discussed in Section III,C.

## B. Adjustment of Isolated Musical Intervals

Adjustment procedures have also been extensively used to study the perception of musical intervals, primarily the octave (Ward, 1953, 1954; Walliser, 1969; Terhardt, 1969; Sundberg & Lindquist, 1973; Burns, 1974b) but also other intervals (Moran & Pratt, 1926; Rakowski, 1976). In the typical paradigm the subject is presented with pairs of tones (either sequential or simultaneous), one of which is fixed in frequency and the other under the control of the subject. The subject is instructed to adjust the frequency of the variable tone so that the pitch relationship of the two tones corresponds to a specific musical interval. It should be noted that this is a one- (temporal) interval procedure because the subject is adjusting to some internal standard. Thus, these experiments are akin to the identification experiments discussed in Section

---

**Fig. 1.** Lower graph: Identification functions obtained from a musically trained subject for category scaling of isolated melodic musical intervals. Upper graphs: Discrimination functions (solid lines) obtained from the same subject in a roving-level melodic-interval discrimination experiment for interval separations of 25, 37.5, and 50 cents. (Percent correct discrimination is plotted at the mean value of the two intervals in a diserimination trial). Also shown are the discrimination functions (dashed lines) predicted from the identification functions assuming categorical perception (see text) (Burns & Ward, 1978).

Biblioteca del Real Conservatorio Superior de Música

III,A, not to the usual psychophysical adjustment experiment—in which the subject adjusts the variable stimulus to equal some physically presented standard and which is essentially a two- (temporal) interval discrimination experiment.

For repeated adjustments of intervals individual subjects typically show quite small variability relative to that obtained in magnitude production experiments for 7 ± 2-type stimuli (Stevens, 1976). The average standard deviation of repeated adjustments of sequential or simultaneous octaves composed of sinusoids is on the order of 10 cents (Ward, 1953, 1954; Terhardt, 1969). It is slightly less for octaves composed of complex tones (Walliser, 1969; Terhardt, 1969; Sundberg & Lindquist, 1973). A range of average deviations from 14 to 22 cents for adjustments of the other intervals of the chromatic scale (simultaneous presentation) has been reported by Moran and Pratt (1926). Rakowski (1976) reports variability—in interquartile ranges—of 20 to 40 cents for both ascending and descending melodic versions of the 12-chromatic intervals. Other general trends evident from the results of adjustment experiments are (1) a small but significant day-to-day variability in intrasubject judgments, (2) significant intersubject variability, and (3) a tendency to "compress" smaller intervals (adjust narrower than equal-tempered intervals) and "stretch" wider intervals (adjust wider). All of these trends have counterparts in the category-scaling results (see Section III,A,2).

### C. Musical Interval Discrimination and Categorical Perception

There have been only a few attempts to measure frequency-ratio resolution using the two- (temporal) interval, two-alternative forced-choice (2I-2AFC) discrimination paradigms typically employed to determine resolution for other auditory stimulus parameters, such as frequency and intensity. In one such experiment Houtsma (1968) obtained estimates of frequency ratio JNDs for melodic intervals composed of sinusoids using a 2I-2AFC task (i.e., subjects were asked to judge which of two melodic intervals was larger). The average JND for three subjects (based on the 75% correct point of the psychometric function) at the physical octave was 16 cents. The JNDs for other ratios in the immediate vicinity of the octave were not significantly different. The JNDs at the ratios corresponding to the (just-tempered) intervals of the chromatic scale were also determined for one subject. They ranged from 13 to 26 cents. The ordering of JND magnitude with interval magnitude indicates a general tendency for smaller JNDs for narrower intervals—a tendency also shown in the dispersion data for Rakowski's (1976) adjustment experiment. However, this ordering also roughly corresponds to smaller JNDs (or less variability) for the intervals that show the highest frequency of occurrence in melodies (e.g., Zipf, 1949).

The other experiments in which 2I-2AFC discrimination data for melodic musical intervals were obtained were those reported by Burns and Ward (1974, 1978). Before these data are reviewed, however, it is necessary to discuss the context in which they were obtained. The series of experiments comprised a study of *categorical perception*. The term "categorical perception" was coined by the speech-perception group at

Haskins Laboratories (see e.g., Studdert-Kennedy *et al.*, 1970) to describe the results of experiments in the perception of certain speech tokens that varied along a single acoustic continuum. The concept is defined by Studdert-Kennedy *et al.* according to two criteria: (1) the subjects' identification functions show sharp and reliable category boundaries when stimuli along the continuum are categorized in terms of phonetic labels (this result was discussed in Section III,A,2) and (2) the subjects' discrimination performance for stimuli equally spaced along the continuum can be correctly predicted from their identification functions if it is assumed that they can discriminate two stimuli only to the extent to which they can differentially identify them. This was contrasted by Studdert-Kennedy *et al.* with the usual situation in psychophysics (discussed in Section II in terms of the $7 \pm 2$ rule) in which subjects' resolution in 2I-2AFC discrimination tasks is much better than their resolution in single- (temporal) interval, multi-alternative identification tasks—a situation that Studdert-Kennedy *et al.* call *continuous perception.*

Results indicative of categorical perception were at first obtained only with speech stimuli and formed the basis of a portion of the rationale for speech-specific models of speech perception (Studdert-Kennedy *et al.*, 1970). However, a number of experiments showing results indicative of categorical perception for nonspeech stimuli were reported later (e.g., Locke & Kellar, 1973; Miller, Wier, Pastore, Kelly, & Dooling, 1976), including the original experiments of Burns and Ward (1974) on musical interval perception. These results forced an extension of the concept of categorical perception to signals other than speech.

One experiment in the Burns and Ward (1978) study was a musical-interval-discrimination experiment designed to duplicate closely the conditions of the typical speech-perception experiment. The basic paradigm was similar to that used by Houtsma (1968) except that the two melodic intervals to be discriminated on a given trial were adjacent ones chosen at random from a set of melodic intervals separated by an equal step-size in cents over a 300 cent range (i.e., roving-level discrimination) rather than two intervals that were fixed over a block of trials (i.e., fixed-level discrimination). The results of one subject for step-sizes of 25, 37.5, and 50 cents, over the range from 250 to 550 cents, are shown in the upper graphs in Fig. 1. Also shown are the discrimination functions predicted from the identification functions (discussed in Section III,A,2; see lower graph). The agreement between the obtained and predicted discrimination functions in Fig. 1 is comparable to that shown by categorically perceived speech stimuli. Thus, according to the criteria stated by Studdert-Kennedy *et al.*, musical intervals were perceived categorically in this experiment.

Siegel & Siegel (1977b) have interpreted the correlation between category boundaries in the category-scaling identification functions and the modal peaks in the magnitude-estimation, standard-deviation functions of their subjects as indicative of categorical perception of melodic intervals. Halpern and Zatorre (1979) have obtained results for the perception of harmonic musical intervals similar to those found by Burns and Ward (1978) for melodic intervals. Locke and Kellar (1973) also obtained results indicative of categorical perception for a harmonic triad varying from minor to major.

To obtain results indicative of categorical perception for musical intervals, one must use musically trained subjects who are able to reliably label the stimuli. Nonmusicians have significantly poorer performance and show no evidence of categorical perception (Burns & Ward, 1978; Siegel & Siegel, 1977b; Locke & Kellar, 1973). For example, the lower graph in Fig. 2 shows the average performance of four musicians and four nonmusicians in a 50-cent roving-level discrimination experiment (Burns &

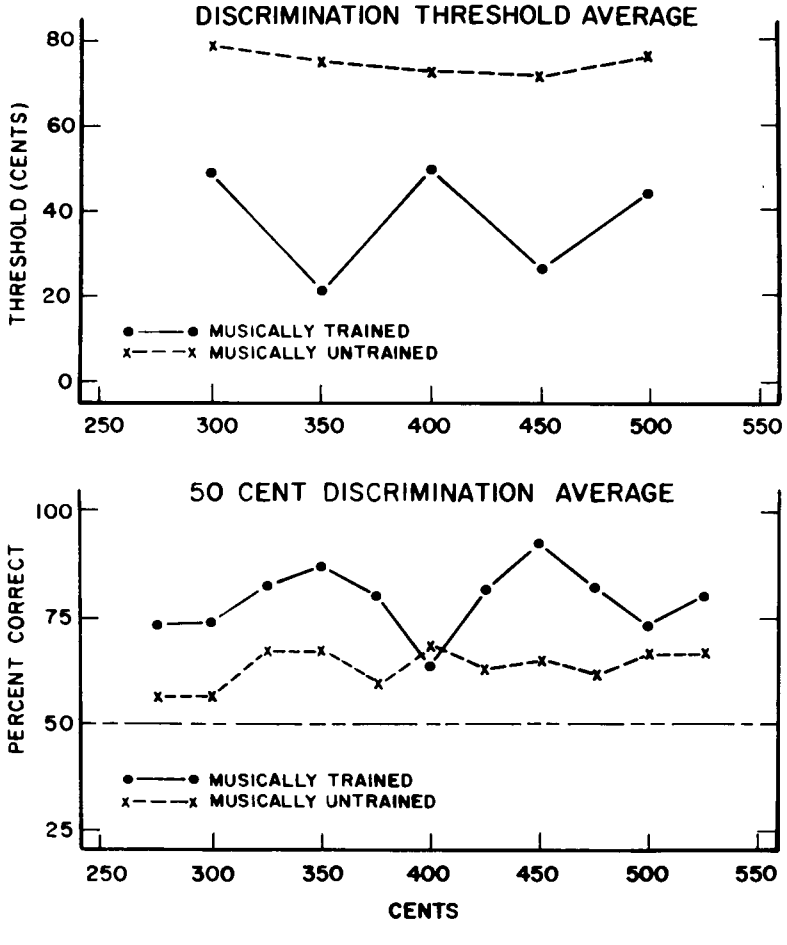



**Fig. 2.** Upper graph: The average melodic-interval JND estimates of four musically trained (points connected by solid lines) and four musically untrained subjects (points connected by dashed lines) obtained in a fixed-level discrimination experiment. Lower graph: Average discrimination functions (50 cent interval separation) of four musically trained (solid lines) and four musically untrained (dashed lines) subjects obtained in a roving-level discrimination experiment (Burns & Ward, 1978).

Ward, 1978). The discrimination function of the musicians shows the typical "peak-trough" form associated with categorical perception; the discrimination function of the nonmusicians shows significantly poorer performance with no evidence of categorical perception.

This evidence indicates that when equivalent paradigms are used, musical intervals are perceived as "categorically" as speech tokens. However, the usefulness and validity of the concept of categorical perception have become the focus of considerable controversy. Categorical perception almost certainly does not define the limits of perceptual ability, as most speech perception literature implies. This is clearly shown by another experiment (Burns & Ward, 1978) in which frequency-ratio-JND estimates were obtained for melodic intervals corresponding to "within category" and "between category" points on the identification functions, using a fixed-level 2I-2AFC adaptive paradigm. The results for one subject are shown in the upper graph of Fig. 3. JND estimates obtained from the initial adaptive run show the form that would be expected for categorically perceived stimuli (i.e., larger JNDs for "within-category" stimuli compared to "between-cateogry" stimuli). However, after subjects had reached asymptotic performance (the numbers by the data points indicate the number of adaptive runs necessary to reach asymptotic performance), the JND estimates for all ratios were roughly equal and comparable to those of Houtsma's (1968) subject for the same range. Furthermore, roving-level discrimination functions obtained after subjects had reached asymptotic performance in the adaptive paradigm still show the typical peak-trough form associated with categorical perception (see lower graph Fig. 3). Similar results for a speech continuum have been obtained by Carney, Widin, & Viemeister (1977). Thus, whether or not certain stimuli are perceived categorically depends to a large extent on the experimental paradigm. The increased stimulus uncertainty induced by a roving-level procedure is apparently a major factor in eliciting categorical perception.

These discoveries suggest that categorical perception and magic number $7 \pm 2$ results are not dichotomous phenomena but different manifestations of the processing of high information signals. This processing can probably be modeled by a slightly more general form of the Durlach-Braida model discussed in Section II. For example, using the criteria that discrimination resulution be equivalent to identification resolution, the model already predicts categorical perception for two special cases: (1) somewhat trivially, when both context-coding noise and trace noise are small and sensation noise limits performance and (2) when both context-coding noise and trace noise are high, relative to sensation noise, but trace noise is dominant (for example, for very long-time delays in a 2I-2AFC task) and, hence, optimum discrimination is accomplished by comparing context-coding representations. If it is assumed that there is a more precise form of the context-coding mode for highly overlearned, complex stimuli (such as speech tokens or musical intervals) that is not dependent on the immediate context of the stimuli in the experiment but is based on long-term experience, and if it is further assumed that sensory-trace-mode noise is inherently high due to the more complex nature of the stimuli, then sensory-trace noise would be expected

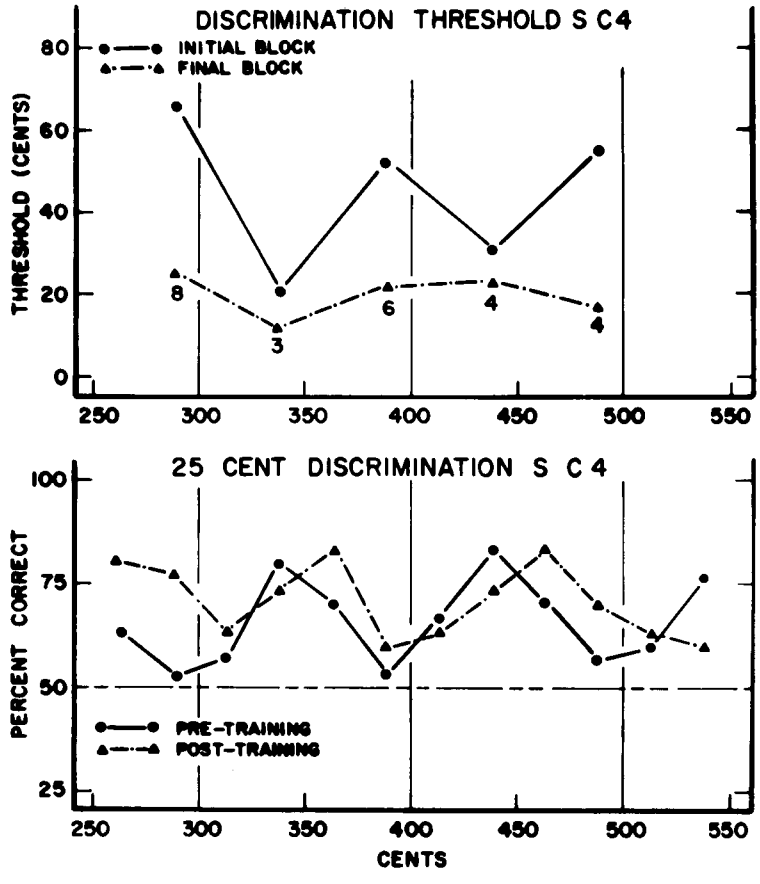



**Fig. 3.** Upper graph: Melodic-interval JND estimates obtained from a musically trained subject using a fixed-level adaptive discrimination paradigm. The solid lines connect estimates obtained in the initial adaptive run at each condition. The dash-dot lines connect estimates obtained after asymptotic performance was reached. Lower graph: Discrimination functions (25 cent interval separation) obtained from the same subject using a roving-level discrimination task. The functions plotted as solid lines were obtained prior to, and the functions shown as dash-dot lines were obtained after, the fixed-level discrimination experiment (Burns & Ward, 1978).

to dominate except for minimal uncertainty procedures. Under these assumptions optimum performance would be obtained in the context-coding mode for most paradigms, and categorical perception (i.e., equivalence of identification and discrimination) would be the norm. The fact that the standard deviation of repeated adjustments of musical intervals (an identification procedure) is of the same order of magnitude as the frequency ratio JND determined in a minimal-uncertainty 2I-2AFC discrimination procedure supports the idea of a precise context-coding mode in the case of musical intervals.

## IV. NATURAL INTERVALS AND SCALES

### A. Natural Intervals and Their Possible Basis

Given the assumption, based on the discussion in Part II, that practical music is limited to a relatively small set of discrete pitch relationships, how are the specific values of these relationships chosen? That is, are there "natural" frequency ratios inherent in the manner in which the auditory system processes tonal stimuli, which are perceptually salient or unique and, as such, define the intervals of the scale? According to traditional Western music theory, such natural intervals do exist. They are associated with the concept of consonance (and its inverse, dissonance) and are defined by small-integer frequency ratios. For example, the ratio 2:1 (octave) is the most consonant interval, the ratio 3:2 (fifth) is the next most consonant, and so on—consonance decreasing with increasing ratio complexity. It should be noted that the concept of consonance discussed in this section is defined, somewhat circularly, by music theory as the sensation associated with simultaneously presented small-integer frequency ratios. This (see Section IV,C) may or may not be synonymous with consonance as measured in psychophysical procedures that are designed to minimize any bias due to musical training. The origin of the concept of consonance in terms of small-integer ratios is usually attributed to the Greek scholar Pythagoras. However, the preference of the ancient Greeks for small-integer ratios was probably based more on metaphysics than on psychophysics.

There are essentially three current explanations of the concept of consonance and its association with small-integer frequency ratios [for an excellent review of the historical development of theories of consonance, see Plomp and Levelt (1965)] and, hence, three theories of the basis of natural intervals. The first is based on the fact that most tones in music (and in voiced speech) are complex periodic tones whose partials are (approximately, at least) harmonically related to the fundamental. This explanation states that we learn to recognize the relationships between the harmonic partials of complex tones and to consider those relationships consonant or natural. Terhardt (1974a, 1977, 1978) is a current proponent of this explanation of natural-interval categories (although he calls this concept "harmony" and divorces it from "sensory consonance", see Section IV,C,2).

The second explanation is also based on the harmonic structure of complex tones. For simultaneously presented complex tones, relatively more of the harmonics of the tones, as well as the primary nonlinear distortion products resulting from interaction between harmonics, will coincide in frequency to the extent that the tones are related by small-integer ratios. For example, if two tones are related by an octave, all of the harmonics of the higher frequency tone coincide with harmonics of the lower frequency tone resulting in a smooth (consonant) sound. Conversely, if complex tones are slightly mistuned from small-integer ratios, the interaction between nearly coinciding harmonics will create a sensation of beating or roughness, which is presumed to be related to dissonance. This explanation is usually attributed to Helmholtz (1954).

The third explanation is based on the assumption that the brain prefers combinations of frequencies whose neural-firing patterns contain a common periodicity. This extension of an old idea (e.g., Lipps, 1883; Meyer, 1898) is currently championed by Boomsliter and Creel (1961) and Roederer (1973). It essentially predicts the existence of small-integer-frequency-ratio "detectors."

The latter two explanations were both originally based on simultaneously presented tones. The argument is often made that since scales containing these natural intervals predate harmony and polyphony, explanations of scale development based on simultaneous presentation of tones are inappropriate. This argument is somewhat questionable as accompaniment in parallel octaves and fifths, as well as other intervals, is prevalent in many tribal music systems (see Sachs, 1961; Nettl, 1956). In any case, both explanations have been extended to melodic intervals. The "beats" explanation is extended via what one might term the "Flintstone hypothesis" (e.g., Wood, 1961)—that early music was played in highly reverberant caves that provided pseudosimultaneous presentation. The modern version of this hypothesis is currently promoted by Benade (1976), who assumes that cues provided by the (reverberant) interference between mistuned harmonics are the primary intonation criteria used by musicians in actual performance situations. Similarly, the neural-pattern explanation has been extended to melodic intervals by assuming some sort of "neural reverberation" (Boomsliter & Creel, 1971; Roederer, 1973). The explanation based on learning the harmonic relationships of complex tones is obviously equally applicable to melodic intervals. In Section IV,C some experimental evidence relevant to these explanations for the propensity for small-integer frequency ratios will be examined.

### B. Natural Scales and Temperament

There are numerous versions of the natural scale based on these natural intervals, however, two general forms are usually distinguished. One is the scale of just intonation (JI), in which the intervals within an octave are determined, in essence, by choosing the smallest possible whole-number frequency ratios relative to the tonic [i.e., the most consonant intervals, for example: 3:2 (fifth), 4:3 (fourth), 5:4 (major third), etc.]. This process is fairly straightforward for choosing the intervals of the diatonic major scale (do, re, me, fa, so, la, ti, do; see Table I). However, attempts to "fill in" the remainder of the octave to give 12 approximately equally spaced intervals result in fairly complex frequency ratios, and there are a number of alternative candidates for some of the intervals. Although JI is the scale most often cited as being "the" natural scale, its importance is a direct result of the development of harmony in Western music, especially the prominence of the major triad in harmony. One of the characteristics of JI is that the tonic, dominant, and subdominant triads (do, mi, so; so, ti, re; and fa, la, do—respectively) are all tuned in the exact ratios of 4:5:6.

The other prominent form of the natural scale, and the historically precedent one (allegedly devised by Pythagoras himself), is Pythagorean tuning (PT). PT is an attempt to construct a scale using only the so-called perfect consonances (2:1, octave;

3:2, fifth; and 4:3, fourth). This is accomplished by cyclic application of the fifth (and fourth) and the reduction of the resultant intervals to within an octave: for example, $(3/2 \times 3/2) = 9/4 = (2/1 \times 9/8)$ (major second); $(9/8 \times 3/2) = 27/16$ (major sixth); etc. If this process is repeated 12 times, the approximate original starting point is reached. However, since powers of three can never be an exact multiple of two, there is a discrepancy. If the original starting note was C, this resultant is the enharmonic equivalent of C, B♯, which is sharp by a Pythagorean comma (24 cents). One way to avoid this is by proceeding downward in fifths (or equivalently, upward in fourths) for half of the intervals, thereby introducing the discrepancy at the tritone (F♯ = 612 cents or G♭ = 588 cents.). It can be seen in Table I that the main differences between JI and PT are in the major and minor thirds and major and minor sixths.

If the ratios between adjacent intervals in the diatonic JI or PT scales are calculated (see Table I), it will be seen that there are three values for JI (two whole tones 10/9 and 9/8 and a semit 16/15) and two values for PT (a whole tone of 9/8 and a semit 256/242, which is not half the whole tone). This is one manifestation of the major problem with either of the natural tuning systems: key modulations on fixed-tuning instruments require an inordinate number of intervals per octave (at least 30 in the case of JI) since the same note will have a number of different values depending on the key in which it is derived. (For example, if D in the diatonic just scale of C major is used as a keynote to derive a new diatonic major scale, the notes E and A in the new scale will have slightly different frequencies than the notes E and A in the old scale.) In an effort to reduce all of these discrepancies, various temperament systems (i.e., methods of "adjusting" the natural intervals, such as "meantone tuning") were devised. With the exception of equal-temperament, these systems are of only historical interest and will not be discussed here. (For a review of various temperaments and tunings, see Barbour, 1951.)

The scale of equal temperament (ET), discussed in Section I, which divides the octave into 12 equal intervals, was devised as a compromise that would permit modulation in any key but keep fifths, fourths, and thirds as close as possible to small-integer ratio values. Many musicians claim that ET has destroyed the "inherent beauty" of the natural scales and that performers unencumbered by fixed tuning will tend to play in one of the natural scales. We will review some of the evidence regarding the perception and production of natural intervals and scales in the next section.




## C. Experimental Evidence Relevant to Natural Intervals and Scales

### *1. Measured Intonation*

**a. Non-Western Scales.** Three of the major non-Western musical systems (Indian, Chinese, and Arab-Persian) have inclusive scales approximately equivalent to the Western 12-interval scales and, hence, have the same propensity for the "perfect" con-

sonances (octaves, fourths, and fifths). There are, however, a number of musical cultures that apparently employ approximately equally tempered 5- and 7-interval scales (i.e., 240 and 171 cent step-sizes, respectively) in which the fourths and fifths are significantly mistuned from their natural values. Seven-interval scales are usually associated with Southeast Asian cultures (Malm, 1967). For example, Morton (1974) reports measurements (with a Stroboconn) of the tuning of a Thai xylophone that "varied only ±5 cents" from an equally tempered 7-interval tuning. (In ethnomusicological studies measurement variability, if reported at all, is generally reported without definition.) Haddon reported (1952) another example of a xylophone tuned in 171-cent steps from the Chopi tribe in Uganda. The 240-cent step-size, 5-interval scales are typically associated with the "gamelan" (tuned gongs and xylophone-type instruments) orchestras of Java and Bali (e.g., Kunst, 1949). However, measurements of gamelan tuning by Hood (1966) and McPhee (1966) show extremely large variations, so much so that McPhee states: "Deviations in what is considered the same scale are so large that one might with reason state that there are as many scales as there are gamelans." Another example of a 5-interval, 240-cent step tuning (measured by a Stroboconn, "variations" of 15 cents) was reported by Wachsmann (1950) for a Ugandan harp. Other examples of equally tempered scales are often reported for pre-instrumental cultures (although in these cases, the concept of scales may be of doubtful validity). For example, Boiles (1969) reports measurements (with a Stroboconn, "±5 cents accuracy") of a South American Indian scale with equal intervals of 175 cents, which results in a progressive octave stretch. Ellis (1965), in extensive measurements of melodies in Australian aboriginal pre-instrumental cultures, reports pitch distributions that apparently follow arithmetic scales (i.e., equal separation in Hz).

Thus, there seems to be a propensity for scales that do not utilize perfect consonances and that are in many cases highly variable, in cultures that either are pre-instrumental or whose main instruments are of the xylophone type. Instruments of this type produce tones whose partials are largely inharmonic (see Rossing, 1976) and whose pitches are often ambiguous (see De Boer, 1976).

**b. Intonation in Performance.** A number of measurements have been made of the intonation of musicians playing variable-tuning instruments under actual performance conditions (e.g., Greene, 1937; Nickerson, 1948; Mason, 1960; Shackford, 1961, 1962a,b). The results of these measurements have been summarized by Ward (1970). They show a fairly large variability for the tuning of a given interval in a given performance—ranges of up to 78 cents, interquartile values of up to 38 cents. The mean values of interval tunings, in general, show no consistent tendency to conform to either JI or PT in either melodic or harmonic situations. The general tendency seems to be to contract the semit and slightly expand all other intervals relative to ET. There is also some evidence of context-dependent effects [e.g., to play F♯ sharper than G♭ (Shackford, 1962a,b)]. Those results mirror, to a certain extent, the results of the adjustment and identification experiments using isolated intervals (discussed in Sections III,A and III,B), which showed a tendency to compress the scale for small intervals and stretch the scale for large intervals, in both ascending and descending modes of presentation.

The above measurements were obtained for Western classical music, but the same general tendencies are evident in measurements of intonation from a military band (Stauffer, 1954), Swedish folk musicians (Fransson, Sundberg, & Tjernland, 1970), and jazz saxophonists (Owens, 1974). Measurements of intonation in performance for Indian (Hindustani) classical music (Jhairazbhoy & Stone, 1963; Callow & Shepard, 1972) show similar variability. There are even large variations in the intonation (ranges of up to 38 cents) of a given interval in the single performance of a composition by one musician. Callow and Shepard analyzed these variations in terms of melodic context and found no significant correlations. Large variability (±50 cents) was also found in the intonation of a Thai vocalist whose frame of reference was presumably an equally tempered 7-interval scale (Morton, 1974).

There is no evidence from any of these studies that suggests that the performers tend to play intervals corresponding to exact small-integer ratios, either with reference to the tonic or to preceding notes, for either melodic or harmonic situations.

### 2. *Sensory Consonance and Dissonance*

As mentioned in Section IV,A, the degree of consonance of musical intervals is defined by music theory in terms of the simplicity of frequency ratios; hence, any attempt to rate consonance or dissonance of intervals by musicians who are able to identify the intervals is obviously subject to bias. Therefore, recent attempts to determine the physical parameters corresponding to the sensations of consonance and dissonance have employed musically naive observers. For example, van de Geer, Levelt, and Plomp (1962) showed that for naive observers the term consonance is synonymous with beautiful or euphonious, and, conversely, the term dissonance is synonymous with ugly or noneuphonious. Using such synonyms, when necessary, Plomp and Levelt (1965) and Kameoka and Kariyagawa (1969a,b) had musically naive observers scale the consonance of simultaneous stimulus pairs composed of both pure and complex tones. Both groups of experimenters found that for pure tones consonance first decreases as a function of frequency-ratio magnitude, reaching a minimum at a value corresponding roughly to one quarter of a critical band (a measure of the frequency resolving power of the ear) and then increases as a function of increasing frequency ratio. For complex-tone ratios, however, consonance shows maxima at small-integer frequency ratios (i.e., at those ratios dictated as consonant by music theory). Similarly, the results of a nonverbal triadic-comparison technique (Levelt, Van de Geer, & Plomp, 1966) implied that subjects order both pure- and complex-tone ratios in terms of ratio magnitude but that complex tones are additionally ordered along a dimension corresponding to simplicity of frequency ratio.

Other experimenters have shown that for pure-tone ratios the term "dissonance" is essentially synonymous with "roughness" (Plomp & Steeneken, 1968; Terhardt, 1974a, 1977, 1978), a quality which is highly correlated with the degree of amplitude modulation of the stimulus. The roughness of simultaneous pure-tone pairs as a function of frequency separation is limited at low frequencies by the frequen

Biblioteca del Real Conservatorio Superior de M

lution of the ear and at high frequencies by the limits of the ear's ability to follow rapid amplitude modulations (Terhardt, 1974b). Plomp and Levelt (1965) and Kameoka and Kariyagawa (1969b) have shown that the dissonance of a complex-tone interval can be accurately predicted by adding the dissonances associated with the interaction of individual partials of the two tones. Slaymaker (1970) and Pierce (1966) have demonstrated that these considerations also predict the relative dissonance of inharmonic complex tones. These results are in essential agreement with Helmholtz's original hypothesis regarding the basis of consonance and dissonance.

### *3. Musical Interval Identification and Discrimination*

The results of three harmonic-musical-interval absolute identification experiments have been analyzed in light of presumed small-integer-ratio natural categories. The rationale involved in the interpretation of these results is that the similarity of intervals is related to their confusability in absolute identification tasks (Plomp *et al.*, 1973; Killam *et al.*, 1975)—or to the response time in deciding whether a given interval is the same as an expected token (Balzano, 1977)—and the existence of small-integer-ratio detectors should result in a similarity based on ratio simplicity. The results of all three experiments show, however, that interval similarity is based primarily on interval width. That is, maximum confusions were made between adjacent intervals, or, equivalently, response time for "different" determinations was inversely proportional to difference in interval width. There were slight additional tendencies in all three experiments to consider equivalent name categories (e.g., major and minor third) and inversions (e.g., minor second, major seventh) as being more similar, but no general tendency for confusion between small-integer ratios. This, surprisingly, was true even for the case in which the intervals were composed of complex tones: Plomp *et al.* hypothesize that the short durations of the intervals used in their experiment prevented observers from using roughness cues for complex-tone pairs.

Contrary to the opinion of some authors (e.g., Husmann, 1953; cited in Schügerl, 1970), absolute identification of harmonic intervals does not require interaction of the two tones on the basilar membrane. Identification performance is essentially equivalent whether the two tones are presented in the same or in opposite ears (Burns & Ward, 1976).

These absolute identification results are consistent with the results of the category-scaling and interval-adjustment experiments (Sections III,A,2 and III,B, respectively), which show, in general, unimodal distributions of interval categories along the dimension of frequency-ratio magnitude. In addition, the large idiosyncratic differences, day-to-day variability, tendencies to stretch and compress the scale, and the apparent inability of musically untrained observers to categorize frequency ratios, shown in the category-scaling and adjustment experiments also argue against the existence of frequency-ratio detectors.

The results of melodic-musical-interval discrimination experiments (Section III,C) also show little evidence for small-integer-ratio detectors. Houtsma (1968) found no

difference between frequency ratio JNDs at the exact octave (2:1) and at nearby ratios. Burns and Ward (1978) found essentially equal frequency ratio JNDs at five different ratios over a range of 250 cents with musically untrained observers showing JNDs an order of magnitude larger than those shown by musicians (see upper graphs in Figs. 2 and 3).

### *4. Conclusions: Learned versus Innate Categories*

The evidence presented thus far implies that musical-interval categories are learned rather than are the direct result of characteristics of the auditory system. This evidence includes: (1) the variability found in measured scales and intonation, even when possible contextual effects are taken into account; (2) the intrasubject variability, large intersubject variability, and consistent deviations from small-integer-ratio categories found in category-scaling and adjustment experiments; (3) the absence of small-integer-ratio singularities in frequency-ratio-JND functions and absence of small-integer-ratio confusions in absolute-identification experiments; and (4) the relative inability of musically untrained subjects to perform musical interval identification or discrimination experiments. Although temporally coded information exists that might be utilized to discriminate exact from nonexact small-integer ratios of low-frequency simultaneous pure tones in a minimal-uncertainty psychophysical task (e.g., Plomp, 1967), this information is apparently of little use, or is ignored, as an intonation cue in actual musical situations. Similarly, there is no indication that possible roughness cues provided by the interaction of reverberant and ongoing complex tones are used as intonation cues in most musical situations.

Assuming, then, that the intonation of individual musicians is based on their ability to reproduce learned categories, there are basically three alternative hypotheses for the origin of these categories: (1) the categories are learned from the scales of a given culture, the intervals of which were originally chosen at random; (2) the categories are learned from the scales of a culture, the intervals of which were originally derived from considerations of sensory consonance (e.g., Plomp & Levelt, 1965); or (3) the categories are based on the early unconscious learning of relationships between the partials of environmental sounds—primarily voiced speech (Terhardt, 1974a, 1977, 1978).

Given the fairly universal occurrence in scales of perfect consonances, especially the octave (Section IV,C,1), hypothesis 1 seems untenable. The variability of scales in pre-instrumental music or music whose major instruments produce inharmonic tones is consistent with either hypothesis 2 or 3. One problem with Terhardt's "unconscious learning" hypothesis, however, is that this hypothesis predicts that even observers without musical training possess a sense of the basic musical intervals (e.g., octave, fifth). The evidence from a number of the experiments discussed in Sections II,C and III,D (e.g., Siegel & Siegel, 1977b; Locke & Kellar, 1973; Burns & Ward, 1978) indicates that this is not so. In fact, even the concept of octave-unison similarity

appears to be a function of musical training (Allen, 1967; Thurlow & Erchul, 1977). There also appears to be evidence that young children do not possess an innate interval sense (Frances, 1968; Zenatti, 1969; cited in Risset, 1978).

Thus, on the basis of the available evidence, the most likely possibility seems to be that natural intervals, as dictated by sensory consonance, have influenced the determination of the scales of most cultures, but that the intonation of individual musicians is primarily a function of their acquired ability to reproduce the learned interval categories of these scales.

## D. Octave Generalization and Chroma

### *1. Introduction*

As was mentioned in Section I, the scales of Western music are predicated, in part, on the concept of octave generalization (i.e., that tones separated by an octave are in some sense musically equivalent, and thus scales are uniquely defined by specifying the intervals within an octave). Octave generalization appears to be farily universal for advanced musical cultures (see Section IV,C). Octave circularity of relative pitch is inherent in the conceptualizations of pitch as having two aspects: (1) pitch height, which is correlated with absolute frequency and (2) chroma, which is correlated with relative position within an octave. This idea has often been graphically represented by a helix or torus (see Chapters 9, 11, and 14, this volume).

### *2. Possible Explanations for Octave Generalization*

There are several possible explanations for the uniqueness of the octave as the basis for an assumed circularity of relative pitch. In the discussion of sensory consonance in Section IV,C,2, it was pointed out that for simultaneous musical intervals composed of complex tones whose partials are harmonically related, the exact octave is unique in that all of the partials of the tones will coincide exactly. Therefore, the octave interval will be no more dissonant than the lower-frequency complex tone alone.

Another explanation is a consequence of the perception of the pitch of complex tones. Current models of complex-tone-pitch perception (e.g., Gerson & Goldstein, 1978; Terhardt, 1974a) assume that complex-tone-pitch perception is a pattern recognition process in which a "central pitch processor" attempts to match the partials of the complex tone with the best-fitting harmonic series. A consequence of this type of operation will be a certain amount of octave ambiguity in the model predictions of fundamental pitch. For example, a complex tone consisting of all harmonics of 200 Hz might also be estimated as the even harmonics of 100 Hz. Such ambiguity has, indeed, been found in complex-tone-pitch-perception experiments (see, e.g., Gerson & Goldstein, 1978; Houtsma, 1979).

### *3. Psychophysical Evidence Regarding Octave Generalizstion*

The bidimensional representations of pitch imply that manifestations of octave equivalence should be found in experiments for which musical training is not a prerequisite. Two experiments often cited as providing such evidence for octave generalization are the conditioning experiments of Blackwell and Schlosberg (1943) using rats and Humphreys (1939) using humans. Both experiments, however, have methodological problems, primarily in the form of stimulus distortion, and have apparently never been replicated.

Other evidence appears to derive from the experiments of Shepard (1964), employing complex tones whose partials consist of octaves of the fundamental. These complex tones are, in essence, passed through a bandpass filter that serves to keep average tone height constant regardless of fundamental frequency. The objective experiments—which have recently been expanded upon by Pollack (1978)—show that judgments of the pitch relation between tones of this type with differing fundamental frequencies are based on the relative proximity of harmonics rather than on absolute fundamental frequency differences. These results—which can be explained either by relative pitch judgments between proximal harmonics or between the dominant "residue" pitches of the complex tones (see DeBoer, 1976)—essentially separate relative pitch from pitch height and are not unexpected. The surprising result is when a set of tones of this type, whose fundamental frequencies cover the range of an octave in semit steps, are played cyclically—with a sufficient time delay between individual tones—the impression is one of constantly rising (or falling) pitch without the octave jumps one would expect (and which are indeed heard if sufficient delay between the tones is not provided). This illusion is often cited as evidence for circularity of relative pitch based on octave equivalence. However, recent experiments (Burns, 1981) show that octave separation of partials is not a prerequisite for eliciting the illusion. An equally salient illusion obtains from cyclically repetitive sequences of tones composed of inharmonic partials (e.g., stretched or compressed octaves). In addition, experiments in which observers judge the similarity of both pure or complex sequential tones provide conflicting evidence (Allen, 1967; Thurlow & Erchul, 1977); in general, only musically trained observers give results indicative of octave similarity.

Another aspect of octave perception is the perceptual stretch of sequential octaves (which seems to reflect a general tendency to stretch all of the wider intervals; see Section III) in which exact octaves are perceived as being too small. This also seems to be fairly universal across musical cultures (Ward, 1954; Hood, 1966; Burns, 1974b). Terhardt's (1974a) hypothesis (see Section IV,C,4) of unconscious learning of the relationship between harmonics provides an explanation for this phenomenon. It is based on the fact that the pitches of the individual components of a complex tone are slightly altered by the presence of other components—in general a downward pitch shift for the lower components and an upward shift for the higher components—resulting in a wider subjective interval between components (Terhardt, 1971). However, this argument is based on purely circumstantial evidence and does not appear to

be a completely satisfactory explanation. It does not, for example, explain the relative amount of stretch or compression of intervals other than the octave.

Our conclusion is that octave generalization is a learned concept that has its origins in the octave's unique position in the range of sensory consonance of complex-tone intervals. However, resolution of the question of the basis for octave generalization awaits further research, particularly on the Shepard illusion.

## V. CONCLUSIONS AND CAVEATS

### A. Conclusions

On the basis of the evidence reviewed, the following conclusions regarding the perception of musical intervals and scales seem justified.

1. The use of a relatively small number of discrete pitch relationships in music is probably dictated by inherent limitations on the processing of high information-load stimuli by human sensory systems.
2. Natural intervals, in the sense of intervals that show minimal sensory dissonance (roughness) for simultaneous presentation of complex tones, have probably influenced the evolution of the scales of many musical cultures, but the standards of intonation for a given culture are the learned interval categories of the scales of that culture. A corollary of this is that the intonation performance of a given musician is primarily determined by his or her ability to reproduce these learned categories and is little influenced, in most situations, by any psychophysical cues (e.g., roughness, beats of mistuned consonances, etc.).
3. The concept of categorical perception, also related to the limitations on processing of high information-load stimuli, is probably a reasonable description of the way in which intervals are perceived in all but minimal-uncertainty situations, an analogous situation to the perception of phonemes in speech.
4. Quartertone music might be theoretically feasible given sufficient exposure to it, but the present 12-interval Western scale is probably a practical limit. Any division of the octave into intervals smaller than quartertones is perceptually irrelevant for melodic information.
5. Octave generalization is probably a learned concept with its roots in the unique position of the octave in the spectrum of the sensory consonance of complex-tone intervals.

### B. Caveats

The perception of isolated melodic musical intervals may have little to do with the perception of melody. As several of the chapters in this volume will indicate, there is considerable evidence that melodies are perceived as Gestalts or patterns, rather than as a succession of individual intervals, and that interval magnitude is only a small

factor in the total percept. There is also an analogy here with speech perception, but the perception of individual musical intervals may be even less relevant to the perception of music than the perception of individual phonemes is to the perception of speech. The ability to label individual intervals is certainly not crucial to the perception of, or even the production of, music. Many amateur musicians who learn and accurately reproduce melodies "by ear" cannot identify isolated intervals. Categorical perception, at least as it is usually defined (i.e., in terms of the relationship between identification and discrimination), may only be relevant when musicians are listening "analytically"—for example, in order to transcribe a melody. There is obviously a need for experiments in the perception of the intonation of individual notes in familiar and unfamiliar melodic phrases.

## ACKNOWLEDGMENTS

The authors would like to thank the following colleagues for commenting on a previous version of this manuscript: Lawrence Feth, Adrian Houtsma, Ernst Terhardt, and David Woods.

## REFERENCES

Allen, D. Octave discriminability of musical and non-musical subjects. *Psychonomic Science*, 1967, 7, 421-422.

Attneave, F., & Olson, R. Pitch as a medium: A new approach to psychophysical scaling. *American Journal of Psychology*, 1971, *84*, 147-166.

Balzano, G. On the bases of similarity of musical intervals: A chronometric analysis. *Journal of the Acoustical Society of America*, 1977, *61*, S51 (A).

Barbour, J. M. *Tuning and temperament*. East Lansing, Michigan: Michigan State College, 1951.

Benade, A. *Fundamentals of musical acoustics*. London and New York: Oxford University Press, 1976.

Blackwell, H. R., & Schlosberg, H. Octave generalization, pitch discrimination, and loudness thresholds in the white rat. *Journal of Experimental Psychology*, 1943, *33*, 407-419.

deBoer, E. G. On the "residue" and auditory pitch perception. In W. Keidel and W. Neff (Eds.), *Handbook of sensory physiology*. (Volume V/3) Berlin and New York: Springer-Verlag, 1976.

Boiles, J. Terpehua thought-song. *Ethnomusicology*, 1969, *13*, 42-47.

Boomsliter, P., & Creel, W. The long pattern hypothesis in harmony and hearing. *Journal of Music Theory*, 1961, *5*, 2-31.

Boomsliter, P., & Creel, W. Toward a theory of melody. Paper presented to *Symposium on Musical Perception. Convention of the AAS, December 28, Philadelphia, Pennsylvania*, 1971.

Brady, S. A., & Darwin, C. J. Range effect in the perception of voicing. *Journal of the Acoustical Society of America*, 1978, *63*, 1556-1558.

Burns, E. M. In search of the Shruti. *Journal of the Acoustical Society of America*, *56*, Supplement, 1974, S26 (A). (a)

Burns, E. M. Octave adjustment by non-Western musicians. *Journal of the Acoustical Soceity of America, Supplement, Fall*, *56*, 1974, S25-26 (A). (b)

Burns, E. M. The perception of musical intervals (frequency ratios). Unpublished doctoral thesis, University of Minnesota, 1977.

Burns, E. M. Circularity in relative pitch judgments for inharmonic complex tones: The Shepard demonstration revisited, again. *Perception & Psychophysics*, 1981, *30*, 467-472.

Burns, E. M., & Ward, W. D. Categorical perception of musical intervals. *Journal of the Acoustical Society of America*, 1974, *55*, 456 (A).

Burns, E. M., & Ward, W. D. Perception of monotic and dichotic musical intervals. *Journal of the Acoustical Society of America*, 1976, *59*, 552 (A).

Burns. E. M., and Ward, W. D. Categorical perception—phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America*, 1978, *63*, 456–468.

Callow, G., & Shepherd, E. Intonation in the performance of North Indian classical music. Paper presented at the *17th Annual Meeting of the Society for Ethnomusicology, November 30–December 3, Toronto, Canada*, 1972.

Carney, A. E., Widin, G., & Viemeister, N. Noncategorical perception of stop consonants differing in VOT. *Journal of the Acoustical Society of America*, 1977, *62*, 961–970.

Caron, N., & Safvate, D. *Les traditions musicales, Iran*. Correa, France: Buchet/Chastel, 1966.

Diehl, R. L., Elman, J. L., & McCusker, S. B. Contrast effects on stop consonant identification. *Journal of Experimental Psychology*, 1978, *4*, 599–609.

Durlach, N. I., & Braida, L. D. Intensity perception I. Preliminary theory of intensity resolution. *Journal of the Acoustical Soceity of America*, 1969, *46*, 372–383.

Ellis, C. Pre-instrumental scales. *Ethnomusicology*, 1965, *9*, 126–144.

Estes, W. K. An associative basis for coding and organization in memory. In A. W. Melton and E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C.: Winston, 1972. Pp. 161–190.

Frances, R. *La perception de la musique*. Paris: Vrin, 1958.

Fransson, F., Sundberg, J., & Tjernlund, P. Statistical Computer Measurements of the Tone-Scale in Played Music. STL-QPSR 2-3/1970. Department of Speech Communication, KTH, Stockholm.

Gerson, A., & Goldstein, J. L. Evidence for a general template in central optimal processing for pitch of complex tones. *Journal of the Acoustical Society of America*, 1978, *63*, 498–510.

Greene, P. C. Violin intonation. *Journal of the Acoustical Society of America*, 1937, *9*, 43–44.

Haddon, E. Possible origin of the Chopi Timbila xylophone. *African Music Society Newsletter*, 1952, *1*, 61–67.

Halpern, A. R., & Zatorre, R. J. Identification, discrimination, and selective adaptation of simultaneous musical intervals. *Journal of the Acoustical Society of America*, 1979, *65*, 540 (A).

Helmholtz, H. *The sensations of tone*. Translated from the 1877 German edition. New York: Dover, 1954.

Hood, M. Slendro and Pelog redefined. *Selected Reports in Ethnomusicology, Institute of Ethnomusicology*, UCLA, 1966, *1*, 36–48.

Houtsma, A.J.M. Discrimination of frequency ratios. *Journal of the Acoustical Society of America*, 1968, *44*, 383 (A).

Houtsma, A.J.M. Musical pitch of two-tone complexes and predictions by modern pitch theories. *Journal of the Acoustical Society of America*, 1979, *66*, 87–99.

Hymphreys, L. F. Generalization as a function of method of reinforcement. *Journal of Experimental Psychology*, 1939, *25*, 361–372.

Husmann, H. *Vom Wesen der Konsonanz*. Heidelberg: Verlag Müller-Thiergarten, 1953.

Jhairazbhoy, N., & Stone, A. Intonation in present day North Indian classical music. *Bulletin of the School of Oriental and African studies, University of London*, 1963, *26*, 118–132.

Kameoka, A., & Kuriyagawa, M. Consonance theory part I: Consonance of dyads. *Journal of the Acoustical Society of America*, 1969, *45*, 1451–1459. (a)

Kameoka, A., & Kuriyagawa, M. Consonance theory part II: Consonance of complex tones and its calculation method. *Journal of the Acoustical Society of America*, 1969, *45*, 1460–1469. (b)

Killam, R. N., Lorton, P. V., & Schubert, E. D. Interval recognition: Identification of harmonic and melodic intervals. *Journal of Music Theory*, 1975, *19.2*, 212–234.

Kunst, J. *Music in java*. (Volume II) The Hague: Martinus Nyhoff, 1949.

Levelt, W., Van de Geer, J., & Plomp, R. Triadic comparisons of musical intervals. *British Journal of Mathematical & Statistical Psychology*, 1966, *19*, 163–179.

Lim, J. S., Berliner, J. E., & Braida, L. D. Revised perceptual-anchor model for context coding in intensity perception. *Journal of the Acoustical Society of America*, 1977, *61*, S89 (A).

Lipps, T. *Psychologische Studien*. Heidelberg: Weiss, 1883. Pp. 92-161.

Locke, S., & Kellar, L. Categorical perception in a non-linguistic mode. *Cortex*, 1973, *9*, 355-368.

Malm, W. P. *Music cultures of the Pacific, the Near East, and Asia*. Englewood Cliffs, New Jersey: Prentice-Hall, 1967.

Martin, D. W. Musical scales since Pythagoras. *Sound* 1962, *1*, 22-24.

Mason, J. A. Comparison of solo and ensemble performances with reference to Pythagorean, Just, and equi-tempered intonations. *Journal of Research of Music Education*, 1960, *8*, 31-38.

McPhee, C. *Music in Bali*. New Haven, Connecticut: Yale University Press, 1966.

Meyer, M. Zur Theorie der Differenztöne and der Gehorseempfindungen überhaupt. *Beitr. Akust. Musikwiss*, 1898, *2*, 25-65.

Miller, G. A. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 1956, *63*, 81-96.

Miller, J., Wier, C., Pastore, R., Kelly, W., & Dooling, R. Discrimination and labeling of noise-buzz sequences with varying noise-level times: An example of categorical perception. *Journal of the Acoustical Society of America*, 1976, *60*, 410-417.

Moran, H., & Pratt, C. C. Variability of judgements of musical intervals. *Journal of Experimental Psychology*, 1926, *9*, 492-500.

Morton, D. Vocal tones in traditional Thai music. *Selected reports in ethnomusicology*. (Volume 2) Los Angeles, California: Institute for Ethnomusicology, UCLA, 1974. Pp. 88-99.

Nettl, B. *Music in primitive culture*. Cambridge, Massachusetts: Harvard University Press, 1956.

Nickerson, J. F. A Comparison of Performances of the Same Melody Played in Solo and Ensemble with Reference to Equi-tempered, Just and Pythagorean Intonation. Unpublished doctoral thesis, University of Minnesota, 1948.

Olson, H. *Musical engineering*. New York: McGraw-Hill, 1952.

Owens, T. Applying the melograph to 'Parkers Mood.' *Selected reports in ethnomusicology*. (Volume 2) Los Angeles, California: Institute for Ethnomusicology, UCLA, 1974. Pp. 166-175.

Parducci, A. Contextual effects; a range-frequency analysis. In E. C. Carterette and M. P. Friedman (Eds.) *Handbook of Perception*. (Volume II) New York: Academic Press, 1974.

Pierce, J. C. Attaining consonance in arbitrary scales. *Journal of the Acoustical Society of America*, 1966, *40*, 249 (L).

Plomp, R. Beats of mistuned consonances. *Journal of the Acoustical Society of America*, 1967, *42*, 462-474.

Plomp, R., & Levelt, W.J.M. Tonal consonance and critical band-width. *Journal of the Acoustical Society of America*, 1965, *35*, 548-560.

Plomp, R., & Steeneken, H.J.M. Interference between two simple tones. *Journal of the Acoustical Society of America*, 1968, *43*, 883-884.

Plomp, R., Wagemarr, W., & Mimpen, A. Musical interval recognition with simultaneous tones. *Acustica*, 1973, *29*, 101-109.

Pollack, I. The information in elementary auditory displays. *Journal of the Acoustical Society of America*, 1952, *24*, 745-749.

Pollack, I. The information of elementary auditory displays, II. *Journal of the Acoustical Society of America*, 1953, *25*, 765-769.

Pollack, I. Decoupling of auditory pitch and stimulus frequency: The Shepard demonstration revisited. *Journal of the Acoustical Society of America*, 1978, *63*, 202-206.

Rakowski, A. Tuning of isolated musical intervals. *Journal of the Acoustical Society of America*, 1976, *59*, S50 (A).

Risset, J. Musical acoustics. In E. C. Carterette and M. D. Friedman (Eds.), *Handbook of perception*. (Volume IV) New York: Academic Press, 1978.

Roederer, J. *Introduction to the physics and psychophysics of music*. Berlin and New York: Springer-Verlag, 1973.

Biblioteca del Real Conservatorio Superior de Música

Rossing, T. D. Acoustics of percussion instruments—part I. *The Physics Teacher*, 1976, *14*, 546-556.
Sachs, C. In J. Kunst (Ed.), *The wellsprings of music*. The Hague: Martinus Nyhofif, 1962.
Schügerl, K. On the perception of concords. In R. Plomp and G. Smoorenburg (Eds.) *Frequency analysis and periodicity detection in hearing*. Leiden: Sijthoff, 1970.
Shackford, C. Some aspects of perception. Part I. *Journal of Music Theory*, 1961, *5*, 162-202.
Shackford, C. Some aspects of perception. Part II. *Journal of Music Theory*, 1962, *6*, 66-90. (a).
Shackford, C. Some aspects of perception. Part III. *Journal of Music Theory*, 1962, *6*, 295-303. (b)
Studdert-Kennedy, M., Liberman, A. M., Harris, K., & Cooper, F. S. The motor theory of speech perception: A reply to Lane's critical review. *Psychological Review*, 1970, 77, 234-249.
Sundberg, J., & Lindquist, J. Musical octaves and pitch. *Journal of the Acoustical Society of America*, 1973, *54*, 922-927.
Szende, O. *Intervallic hearing: Its nature and pedogogy*. Budapest: Akadémia Kiadó, 1977.
Terhardt, E. Oktavspreizung und Tonhöhen der Schiefbung bei Sinustonen. *Acustica*, 1969, *22*, 348-351.
Terhardt, E. Pitch shifts of harmonics, an explanation of the octave enlargement phenomenon. *Proceedings of the 7th International Congress on Acoustics*, 1971, *3*, 621-624.
Terhardt, E. Pitch, consonance and harmony. *Journal of the Acoustical Society of America*, 1974, *55*, 1061-1069. (a)
Terhardt, E. On the perception of periodic sound fluctuations (roughness). *Acustica*, 1974, *30*, 201. (b)
Terhardt, E. The two-component theory of musical consanance. In E. F. Evans & E. P. Wilson (Eds.), *Psychophysics and Physiology of Hearing* London, Academic Press, 1977.
Terhardt, E. Psychoacoustic evaluation of musical sounds. *Perception & Psychophysics*, 1978, *23*, 483-492.
Thurlow, W. R., & Erchul, W. P. Judged similarity in pitch of octave multiples. *Perception & Psychophysics*, 1977, *22*, 177-182.
Trotter, J. R. The psychophysics of musical intervals: Definitions, techniques, theory and problems. *Australian Journal of Psychology*, 1967, *19*, 13-25.
van de Geer, J. P., Levelt, W., & Plomp, R. The connotation of musical consonance. *Acta Psychologica*, 1962, *20*, 308-319.
Vinegrad, M. D. A direct magnitude scaling method to investigate categorical vs. continuous modes of speech perception. *Language & Speech*, 1972, *15*, 114-121.
Wachsmann, K. An equal-stepped tuning in a Ganda harp. *Nature* (*London*), 1950, *165*, 40.
Walliser, U. Über die Spreizung von empfundenen Intervallen gegenüber mathematisch harmonischer Intervallen bei Sinustones. *Frequenzy*, 1969, *23*, 139-143.
Ward, W. D. The Subjective Octave and the Pitch of Pure Tones. Unpublished doctoral thesis, Harvard University, Cambridge, Massachusetts, 1953.
Ward, W. D. Subjective musical pitch. *Journal of the Acoustical Society of America*, 1954, *26*, 369-380.
Ward, W. D. Musical perception. In J. Tobias (Ed.), *Foundations of modern auditory theory*. New York: Academic Press, 1970.

Watson, C., Kelly, W., & Wroton, H. Factors in the discrimination of tonal patterns II. Selective attention and learning under various levels of stimulus uncertainty. *Journal of the Acoustical Society of America*, 1976, *60*, 1176–1186.

White, B. W. Recognition of distorted melodies. *American Journal of Psychology*, 1960, *73*, 100–107.

Wood, A. *The physics of music*. New York: Dover, 1961.

Zenatti, A. La développement génétique de la perception musicale. (Monographies Francaises de Psychologie No. 17). Centre National de la Recherche Scientifique, 1969.

Zipf, G. K. *Human behavior and the principle of least effort*. Reading, Massachusetts: Addison-Wesley, 1949.

Zonis, E. *Classical Persian music: An introduction*. Cambridge, Massachusetts: Harvard University Press, 1973.

Biblioteca del Real Conservatorio Superior de M... MADRID

# 9

# The Processing of Pitch Combinations

*Diana Deutsch*


I. Introduction ..... 271
II. Feature Abstraction ..... 272
A. Octave Equivalence ..... 272
B. Interval and Chord Equivalence ..... 273
C. Proposed Physiological Substrates ..... 273
D. Contour ..... 277
E. Interval Class ..... 278
III. Higher Order Abstractions ..... 282
IV. Alphabets and Hierarchies ..... 287
V. Memory Systems ..... 291
A. The System Retaining Absolute Pitch Information ..... 292
B. The System Retaining Interval Information ..... 300
C. Interactions between These Systems ..... 301
D. Memory for Hierarchically Structured Sequences ..... 304
VI. Conclusion ..... 311
References ..... 312


## I. INTRODUCTION

In this chapter the processing of pitch combinations is examined at several levels. First, we inquire into the types of abstraction that give rise to the perception of local features. Such features may be considered analogous to those of orientation or angle size in vision. We have developed sufficient understanding of sensory physiology to justify speculation concerning how such abstractions are achieved by the nervous system. Other low-level abstractions result in the perception of global features, such


Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

as contour. Next, we consider how combinations of such features are abstracted so as to give rise to perceptual equivalences and similarities. We then examine processing where these higher level abstractions are themselves combined according to various rules.

Investigations into mechanisms of visual shape perception have led to a distinction between an early process, where many low-level abstractions are passively carried out in parallel, and a later process, where questions are asked of these low-level abstractions based on hypotheses about the scene to be analyzed (Hanson & Riseman, 1978). This distinction between abstractions that are formed passively from "bottom-up" and those that result from a "top-down" process is important in music also. As we shall see, much of musical shape analysis occurs only when the context is such as to allow for the confirmation of expectations.

The final sections of the chapter are concerned with memory. It is clear that a musical sequence is retained simultaneously at different levels of abstraction and that the information at these different levels combines to influence memory judgments. Interactions occurring within the different memory systems are examined, as are the ways the outputs of these systems interact during retrieval.

## II. FEATURE ABSTRACTION

### A. Octave Equivalence

It is clear that a strong perceptual similarity exists between tones separated by octaves—that is, whose fundamental frequencies stand in the ratio of 2:1 (or a power of 2:1). In the Western musical scale tones that stand in octave relation are given the same name so that a tone is specified first by its position within the abstracted octave and then by the octave in which it is placed ($G_3$, $F_4$, etc.). Octave duplications also occur in the scales of other cultures (Nettl, 1956).

Various experimental observations related to octave equivalence have been re- suggested that pitch be treated as a bidimensional attribute: the first dimension representing overall pitch level or *tone height*, and the second defining the position of a tone within the octave, or *tone chroma* (Meyer, 1904, 1914; Révész, 1913; Ruckmick, 1929; Bachem, 1948; Shepard, 1964). Contemporary music theorists make an analogous distinction between *pitch* and *pitch class* (Babbitt, 1961, 1965; Forte, 1973).

## B. Interval and Chord Equivalence

When two tones are presented either simultaneously or in succession, there results the perception of a musical interval, and intervals are perceived as being the same size when the fundamental frequencies of their components stand in the same ratio. This principle forms an important basis for the traditional musical scale. The smallest unit of this scale is the semitone, which corresponds to a frequency ratio of approximately 1:1.06. Tone pairs separated by the same number of semitones are given the same name, such as major third, minor sixth, and so on. Contemporary music theory also treats tone pairs separated by the same number of semitones as perceptually equivalent.

Chords consisting of three or more simultaneous tones are also classified in part on the basis of the frequency ratios of their components. However, a simple listing of these ratios is not sufficient to define a chord. For example, a major triad and a minor triad both have as their components a major third (five semitones) and a minor third (four semitones) and a fifth (seven semitones). Thus, the fact that the minor third lies above the major third in the one instance and below it in the other is of perceptual importance, and needs to be taken into account in considering how chord abstraction might be achieved by the nervous system.

Given the principles of octave equivalence and interval equivalence, one might hypothesize that intervals whose components are placed in different octaves are also perceptually equivalent. This assumption is frequently made by contemporary music theorists who refer to such intervals as in the same *interval class*. However, traditional music theory assumes this equivalence for simultaneous but not for successive intervals. Simultaneous intervals whose components have reversed position by being placed in different octaves are termed inversions (Piston, 1948). Thus, a simultaneous interval of $n$ semitones is considered perceptually equivalent to a simultaneous interval of 12-n semitones.

Laboratory evidence for the perceptual similarity of inverted intervals has been obtained. Plomp, Wagenaar, and Mimpen (1973) required subjects to identify intervals formed by simultaneous tone pairs, and they found that confusions occurred between intervals that were inversions of each other. Further evidence was provided by Deutsch and Roll (1974) and is discussed below.

For the case of intervals formed by successive tone pairs, the experimental evidence is complicated. As will be discussed below, it appears that interval class perception here occurs only indirectly through a process of hypothesis confirmation, where the features directly apprehended are pitch class and interval.

## C. Proposed Physiological Substrates

Following Drobisch (1846, 1855), various psychologists have suggested that the phenomenon of the tone chroma (or pitch class) be accommodated by representing

353



pitch as a helix, with tones separated by octaves lying most proximal within each turn of the helix. One can take this suggestion literally and propose that such a mapping of pitch exists in the auditory system, so that columns are formed of neural units that respond to tones spaced at octave intervals. Unfortunately, there is no physiological evidence for such columnar organization. On the other hand, units have been found that exhibit peaks of sensitivity at octave intervals. Such units could mediate the perceptual equivalence of tones standing in octave relation, and this hypothesis is described in detail below.

Various models for the perceptual equivalence of intervals and chords have been advanced. Pitts and McCulloch (1947) proposed that the auditory cortex is composed of layers, each layer containing a topographic projection of frequency-specific units. In each projection, units responding to frequencies related by equal intervals are spaced equal distances apart. These layers are arranged so as to produce columns of units that respond to the same frequencies. They further hypothesized the existence of fibers that traverse this columnar mass parallel to each other in a slantwise direction. Three such slanting pathways would therefore define a three-note chord. Such a mechanism could mediate transposition of simultaneous intervals and chords but could not mediate transposition of successive intervals nor the perceptual similarity of intervals and chords related by inversion.

Another mechanism, suggested by Boomsliter and Creel (1961), was based on the volley theory of pitch perception (Wever & Bray, 1930). They pointed out that when the components of two frequency combinations stand in the same ratio, these combinations should generate the same pattern of firing. That is, one pattern would be produced by the ratio 2:3, another by the ratio 4:5, and so on. They therefore proposed that the perceptual equivalence of simultaneous intervals is mediated by recognition of these patterns. This model would require an ability on the part of the nervous system to follow frequencies at much higher rates than has been established (Deutsch & Deutsch, 1973). A further difficulty for the model is that it cannot account for the perceptual equivalence of successive intervals and chords nor for the similarity of inverted intervals and chords.

Deutsch (1969) proposed a mechanism for the abstraction of first-order pitch relationships that accommodates both octave equivalence and also the equivalence of intervals and chords under inversion. This hypothesis was modeled on findings concerning the abstraction of low-order specific features by the visual system, such as orientation and angle size. It appears that such abstractions are accomplished in several stages (Hubel & Wiesel, 1962). Units with circular receptive fields[1] appear to project onto higher order units in such a way that units whose receptive fields taken together form straight lines converge onto the same higher order units. These higher order units respond to lines of specific orientation presented in a specific position in the visual field. It further appears that these higher order units project onto yet higher order units in such a way that those units responding to lines presented in a given

[1]The receptive field of a unit is that region which, when stimulated, causes a change in the activity of this unit.

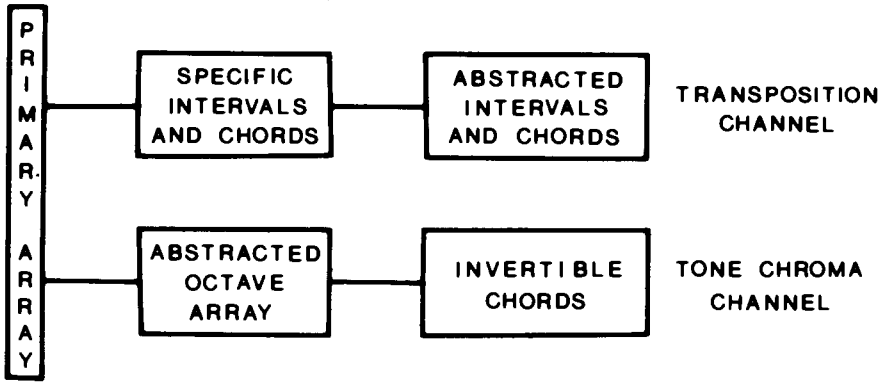



**Fig. 1.** Flow diagram for abstraction of pitch relationships. It is assumed that pitch information is simultaneously abstracted along two parallel channels, one mediating transposition phenomena and the other mediating octave equivalence effects (from Deutsch, 1969).

orientation but in different positions in the visual field converge onto the same unit. These units therefore respond to lines of specific orientation presented in different positions in the visual field.

The proposed mechanism for abstraction of first-order pitch relationships consists of two parallel channels along each of which information is abstracted in two stages (Fig. 1). Channel A mediates the perceptual equivalence of intervals and chords under transposition. In the first stage of abstraction along this channel, first-order units responding to tones of specific pitch are linked in groups of two and three to second-order units. These units therefore respond to specific intervals and chords (such as the combination of $C_4$, $E_4$ and $G_4$, or of $D_5$ and $G_5$). It is assumed that such linkages occur only between units underlying pitches that are separated by an octave or less. In the second stage of abstraction these second-order units are linked to third-order units in such a way that all units that are activated by tones standing in the same relationship are linked together. So, for example, all units activated by an ascending interval of four semitones (a major third) converge onto one unit, all units activated by a descending interval of seven semitones (a perfect fifth) converge onto another unit, all units activated by a major triad onto another unit, and so on (Fig. 2).[2]

Channel B mediates the equivalence of tones separated by octaves (tone chroma or pitch class). In the first stage of abstraction along this channel, first-order units responding to tones of specific pitch are linked in such a way that units underlying tones separated by octaves converge onto the same second-order unit. These units therefore respond to tones in a given pitch class regardless of actual pitch. In the second stage of abstraction, these second-order units converge in groups of two and

[2]Only intervals and chords formed out of elements of the 12-tone chromatic scale are described here, for the purpose of clarity. However, it is assumed that first-order units responding to tones that are not elements of the 12-tone scale are also linked to higher order units in this fashion. This theory therefore makes no assumptions about temperament.

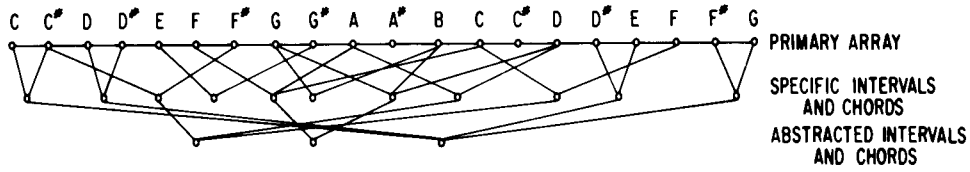



**Fig. 2.** Two stages of abstraction along the transposition channel (from Deutsch, 1969).

three onto third-order units, which therefore respond to pitch class combinations (Fig. 3). Such units would mediate the perceptual similarity of inverted intervals and chords. Since it is assumed that interval class is directly apprehended only where simultaneous intervals are concerned, this level of convergence is assumed to occur only for units that respond to simultaneously presented tones.

Although no attempt has been made to confirm this model at the neurophysiological level, some relevant findings may be cited. Suga, O'Neill, and Manabe (1979) describe neurons in the auditory cortex of the bat that showed facilitation when the first harmonic of a tone was delivered simultaneously with the second harmonic so that the combination formed a perfect fifth. Other units showed facilitation when the first and third harmonics were simultaneously presented, so that the combination formed an octave; yet others showed facilitation when the second and third harmonics were simultaneously presented, so that the combination formed a fourth. Such units often responded poorly to single tones in isolation but strongly and consistently when the appropriate tonal combination was presented. On the above model, units with such characteristics are hypothesized to occur at the first stage of abstraction along Channel A, i.e., the channel mediating interval and chord perception.

Evans (1974) reports the existence of neurons in the auditory cortex of the cat that exhibit peaks of sensitivity at more than one band of frequencies. Peaks spaced at octave intervals were commonly found. Also Suga and Jen (1976) note the presence of

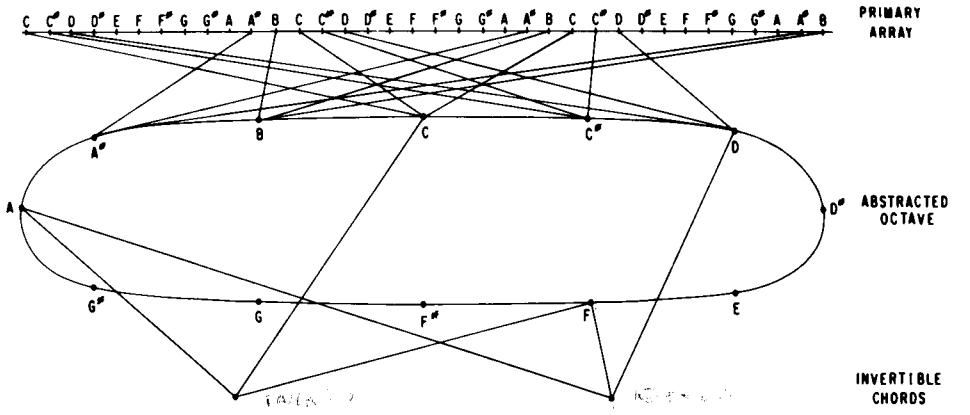


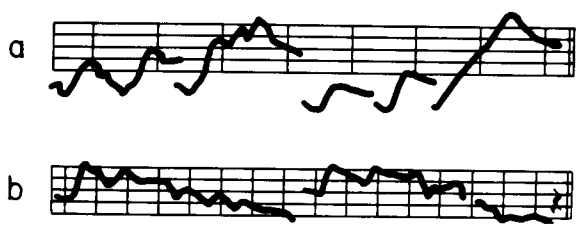

**Fig. 3.** Two stages of abstraction along the channel mediating octave equivalence effects (from Deutsch, 1969).

neurons in the auditory cortex of the bat that showed two peaks of sensitivity that were approximately harmonically related. They write: "since no neurons with a double peaked tuning curve were found at the periphery . . . it is evident that harmonically related components in acoustic signals are converging on some single neurons at higher levels in the auditory system." In the above model, units with these characteristics were hypothesized to occur at the first level of abstraction on Channel B, i.e., as mediating octave equivalence for single tones.

Units with other patterns of facilitation were also found in these neurophysiological studies. This is hardly surprising since the cat and the bat would not be expected to have the same mechanisms for abstraction of pitch relationships as humans. However, it is interesting that the two classes of neuron hypothesized on the model have been shown to exist, i.e., those with multiple peaks for single tones, and those that require the presentation of specific tones in combination to be securely activated. Furthermore, examples of neurons with characteristics as specifically hypothesized on the model were uncovered in these studies.

## D. Contour

In recognizing a segment of music, we employ global as well as specific cues. These include, for example, overall pitch range, the distribution of interval sizes, the proportion of ascending versus descending intervals, and so on. The use of global cues has been best documented for the case of contour in the processing of linear sequence. As shown in the examples in Fig. 4 (taken from Schoenberg, 1967) melodies can be represented by their distinctive contours. We appear to be very sensitive to such information. For example, Werner (1925) found that listeners were able to recognize familiar melodies when these were transformed onto very small scales, so that the intervals were grossly distorted in size. Later, White (1960) found that listeners were able to recognize melodies to some extent when all the intervals were arbitrarily set to one semitone, so that the interval information was entirely removed, apart from the directions of pitch change. When the relative sizes of the intervals were retained, even though their absolute sizes were altered, performance was considerably enhanced. Recent studies by Dowling (1978), Dowling and Fujitani (1971), Idson and Massaro (1978), and Kallman and Massaro (1979) have confirmed and extended such findings.



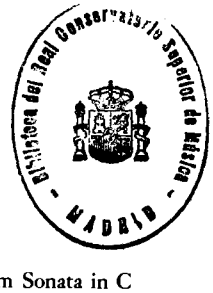




**Fig. 4.** Contours from Beethoven piano sonatas as represented by Schoenberg. a. from Sonata in C minor, Op. 10/1-III, m 1-8. b. from Sonata in D, Op. 10/3-III, m 1-16 (from Schoenberg, 1967).

### E. Interval Class

If different two-tone combinations form the same interval by appropriate octave displacement, these combinations are held to be in the same interval class. For example $C_3$ and $D_5$ in combination form the same interval class as $G_2$ and $F_6$. Whether interval class identity gives rise to perceptual equivalence is a matter for debate. As mentioned above, experimental evidence for such equivalence has been found where simultaneous intervals are concerned (Plomp *et al.*, 1973; Deutsch & Roll, 1974). Further compelling evidence is provided by the fact that we easily recognize root progressions of chords as abstractions.

Where successive intervals are concerned, however, the issue is complicated. If interval class were indeed a perceptual invariant, we should experience no difficulty in recognizing a melody when its component tones are placed in different octaves. This issue was examined experimentally by Deutsch (1972a). The first half of the tune "Yankee Doodle" was generated under various conditions. First, it was produced without transformation in each of three adjacent octaves. Then, it was generated such that each tone was in its correct position within the octave, but the choice of octave placement varied randomly between these same three octaves. And finally, the tune was generated as a series of clicks so that the pitch information was entirely removed but the rhythmic information retained. This was to provide a measure of identification performance on the basis of rhythm alone.

The different versions of the tune were played to separate groups of subjects, who were given no clues to its identity besides being assured that it was well known. Although the untransformed versions were universally recognized, recognition of the randomized octaves version was no better than for the version where the pitch information was removed entirely. However, when the subjects were later informed of the identity of the tune, and were again presented with the randomized octaves version, they now found they were able to follow the tune to a large extent. They were thus able to use octave generalization to *confirm* the identity of the tune, even though they had been unable to *recognize* it in the absence of prior information. It was concluded that this confirmation was achieved by the listeners' imagining the tune simultaneously with hearing the randomized octaves version. In this way they could match each note as it arrived with their auditory image and so confirm that the two were indeed in the same pitch class.

It would appear from this experiment that interval class can be perceived, but not as a first-order abstraction. Rather, perception occurs indirectly through a process of hypothesis testing in which the listener uses pitch class to transpose each tone to the appropriate octave, followed by perception of interval which enables the hypothesis to be confirmed or disconfirmed. By this line of reasoning, interval class, where successive intervals are concerned, is perceived through an active top-down process, in contrast with interval and pitch-class perception, which result from a passive bottom-up process. The extent to which interval class is perceived would then depend critically on the expectations of the listener.

Deutsch examined this issue again (1976, 1979), using a short-term recognition

paradigm. Listeners were presented with a standard six-tone melody, followed by a comparison melody. The comparison was always transposed up four semitones from the standard. On half of the trials, this transposition was exact so that the set of intervals and their orders were preserved. On the other half, two of the tones in the transposed melody were permuted. The permuted tones were always a semitone apart in pitch, and no tones were permuted that were either at the beginning or the end of the melody or that were adjacent to each other. Thus, in the permuted sequences, four out of the five successive intervals were changed in size by a semitone; however, the contour of the melody (as defined by the sequence of directions of pitch change) was unaltered regardless of whether or not the comparison melody was an exact transposition of the standard. (This invariance of contour was necessary to insure that contour could not be used as a basis for recognition judgments.)

There were four conditions in the experiment. In the first condition the standard melody was presented once, followed by the comparison melody. In the second condition the standard melody was repeated six times before presentation of the comparison melody. Here, all repetitions were exact. In the third condition the standard melody was again repeated six times, but now on half of the repetitions the melody was transposed in its entirety an octave higher and on the other half it was transposed an octave lower. In the fourth condition the standard melody was again repeated six times, but now on each repetition the individual tones in the melody were placed alternately in the higher and the lower octaves.

The results of the experiment are shown on Fig. 5. It can be seen that exact repetition of the melody produced a substantial improvement in comparison perfor-

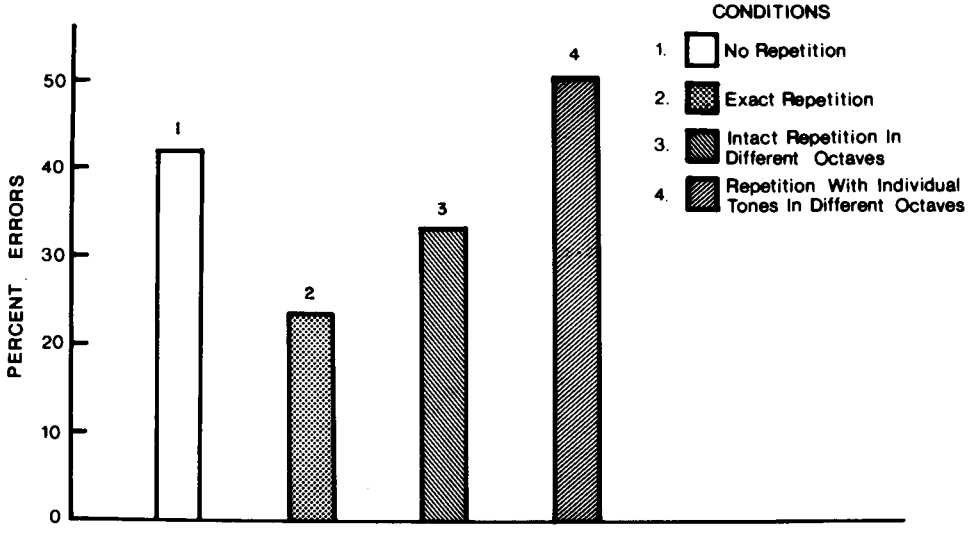



**Fig. 5.** Percent errors in different conditions of experiment studying the effects of octave displacement on consolidation of melodic information (from Deutsch, 1979).

mance, and an improvement also occurred when the melody was repeated intact in the higher and lower octaves. However, when the melody was repeated such that its component tones alternated between the higher and lower octaves, performance was significantly poorer than when the melody was not repeated at all. This experiment again strongly indicates that interval class cannot be treated as a first-order perceptual feature. Repetition of a set of successive intervals resulted in consolidation of memory for these intervals; however, repetition of a set of interval classes did not do so.

These findings are as expected on the model of Deutsch (1969), since this model assumes that the linkages that give rise to abstraction of successive intervals occur between units underlying specific pitches, and not between higher order units underlying pitch classes. Idson and Massaro (1978) proposed an alternative explanation. They argued that the placement of the tones in different octaves resulted in an alteration of melodic contour, and that this in turn acted to disrupt recognition performance. As evidence for this they point out that when the component tones of melodies were placed in different octaves but contour was preserved, recognition of these melodies was at a higher level than when contour was not preserved (Dowling & Hollombe, 1977; Idson & Massaro, 1978). However, preservation of contour would be expected to improve performance regardless of how interval calss is perceived, since contour alone has been shown to be a powerful cue in melody recognition (Werner, 1925; White, 1960, Dowling, 1978). Once the listener has guessed the identity of the melody on the basis of contour (or any other cue for that matter), he can then confirm or disconfirm his hypothesis by a process of matching each tone as it arrives with its octave equivalent (Deutsch, 1972a, 1978a). The findings on the preservation of contour are not, therefore, evidence against an explanation in terms of Deutsch's (1969) model.

Idson and Massaro (1978) proposed instead that melody recognition is mediated by two processes: first, recognition of the succession of pitch classes, and second, recognition of contour (the sequence of directions of pitch change). If their hypothesis were correct, then recognition of melodies where pitch class and contour are preserved but the component tones are in different octaves should be at as high a level as recognition of untransformed melodies. However, Kallman and Massaro (1979) found that this transformation resulted in a significant decrement in recognition performance. This poses a severe difficulty for Idson and Massaro's theory.

Kallman and Massaro (1979) also investigated the effect on recognition performance of preserving contour under octave displacement but altering pitch class by randomly raising or lowering each tone by one or two semitones. They found that recognition performance under this transformation was considerably poorer than where pitch class was preserved. However, this result would also be expected on the hypothesis-testing argument, for just as a perception of the correct set of pitch classes would tend to confirm a hypothesis, so would perception of an incorrect set of pitch classes tend to disconfirm it.

A further point should here be made. In the experiment by Idson and Massaro (1978), subjects were given the names of a small set of test melodies, and they were presented with these melodies under various transformations for hundreds of trials,

allowing ample opportunity for hypothesis-testing. On the other hand Deutsch (1972a) and Kallman and Massaro (1979) presented subjects with each melody only once and did not inform them of the identities of these melodies so that hypothesis testing was much more difficult. In comparing the results from these two types of paradigm, we find that recognition performance where pitch class was preserved but octave placement was randomized was considerably better in the study of Idson and Massaro than in the other two studies. This difference would be expected on the theory that listeners recognize interval class through the mediation of hypothesis-testing, but provides a further difficulty for Idson and Massaro's model.

In this regard it is instructive to consider the use of octave jumps in traditional music. If the present line of reasoning is correct, such jumps can be made with impunity provided the musical setting is such that the displaced tone is expected by the listener. We should therefore suppose octave jumps to be limited to such situations. Indeed, this appears to be the case. In one such situation a melodic line is presented several times without transformation. A clear set of expectations having thus been established, a jump to a different octave occurs. The melodic line on Fig. 6a, for instance, occurs after this line has been presented several times with no octave jumps. Another such situation is where the harmonic structure is clear and unambiguous, so that again the displaced tones are highly probable to the listener. This is illustrated in the segment shown on Fig. 6b. It should also be observed that when octave jumps occur in traditional music, often the identical pitch class is quickly repeated in the new octave. This also occurs in the segment in Fig. 6b. According to the model of Deutsch (1969), this pitch class identity should be recognized directly. The repeated pitch class then provides a means of placing the tones in the new octave in correct relationship with the tones in the previous octave.

The technique of 12-tone composition employs very frequent octave jumps (see below). This raises the question of whether the listener does in fact recognize as perceptually equivalent two presentations of the same tone row under octave displacement. Given the evidence and arguments outlined above, such recognition

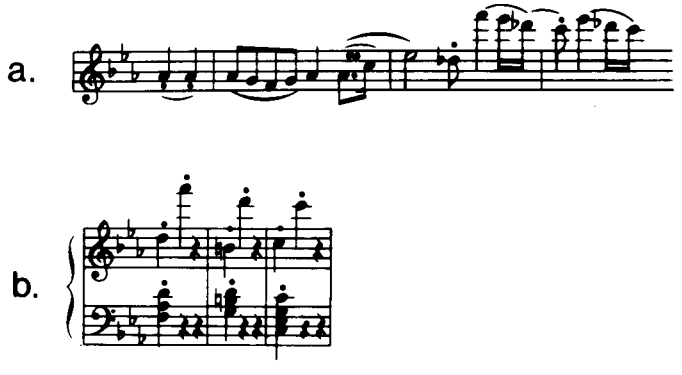






**Fig. 6.** Two examples of the use of octave jumps. In these instances the jumps are readily processed. a. from Beethoven, *Rondo in C, Op. 5, No 1.* b. from Beethoven, *Sonata in C minor, Op. 10, No. 1.*

should be possible in principle, but only if the listener is very familiar with the material, or if its structure is such as to arouse strong expectations (see also Meyer, 1973).

## III. HIGHER ORDER ABSTRACTIONS

Given that linkages are formed between first-order pitch elements, we may next inquire how higher order abstractions are further derived so as to lead to perceptual equivalences and similarities. We recognize visual shapes as equivalent when these differ in size, orientation, or position in the visual field. What transformations result in analogous equivalences in music?

Theorists have long drawn an analogy between perception of pitch relationships and of relationships in visual space (Helmholtz, 1859; Mach, 1906; Koffka, 1935). In contrast with visual space, however, pitch was conceived as represented along one dimension only. As Mach (1906) wrote:

> A tonal series occurs in something which is an analogue of space, but is a space of one dimension limited in both directions and exhibiting no symmetry like that, for instance of a straight line running from right to left in a direction perpendicular to the median plane. It more resembles a vertical right line. . . .

More recently, several investigators have shown that auditory analogues of visual grouping phenomena may be created by mapping one dimension of visual space into log frequency and the other into time (Van Noorden, 1975; Deutsch, 1975a; Bregman, 1978; Divenyi & Hirsh, 1978). In the visual representation of the sequence shown on Fig. 11 of Chapter 4, for example, the principle of proximity emerges clearly. We may therefore inquire whether auditory analogues also exist for perceptual equivalences found in vision (Julesz & Hirsh, 1972).

Von Ehrenfels (1890) in his influential paper on form perception, pointed out that a melody when transposed retains its essential form, the *Gestaltqualität*, provided that the relations among individual tones are unaltered. In this respect, he argued, melodies are similar to visual shapes. On our present intermodal analogy, transposing a melody would be like translating a shape to a different location in the visual field. Shapes may be moved in this way without destroying their perceptual identities (Deese & Grindley, 1947). Similarly, transposing melodies to different pitch levels may leave identification of these melodies unimpaired.[3]

We may next inquire whether further equivalences can be demonstrated for musical shapes that are analogous to their visuospatial counterparts. Schoenberg (1951)

[3]This is clear from everyday experience where long-term memory is concerned (Deutsch, 1969; Attneave & Olson, 1971) and also occurs in some short-term situations (Divenyi & Hirsh, 1978). However, short-term recognition of transposed melodies may be difficult (Attneave & Olson, 1971; Cuddy & Cohen, 1976; Dowling, 1978). This probably reflects the projection of interval information onto highly overlearned unequal-interval scales (Deutsch and Feroe, 1981).

argued that transformations similar to rotation and reflection in vision result in perceptual equivalences in music also. He wrote:

> *The unity of musical space demands an absolute and unitary perception.* In this space . . . there is no absolute down, no right or left, forward or backward. . . . Just as our mind always recognizes, for instance, a knife, a bottle or a watch, regardless of its position, and can reproduce it in the imagination in every possible position, even so a musical creator's mind can operate subconciously with a row of tones, regardless of their direction, regardless of the way in which a mirror might show the mutual relations, which remain a given quantity.

This statement may be compared with Helmholtz's (1844) description of imagined visuospatial transformations:

> Equipped with an awareness of the physical form of an object, we can clearly imagine all the perspective images which we may expect upon viewing it from this or that side (see Warren & Warren, 1968, p. 252).

On this basis, Schoenberg proposed that a row of tones may be recognized as equivalent wben it is transformed such that all ascending intervals become descending intervals and vice versa ("inversion")[4], when it is presented in reverse order ("retrogression"), or when it is transformed by both these operations ("retrograde-inversion"). Figure 7 illustrates Schoenberg's use of his theory in compositional practice. As he wrote:

> The employment of these mirror forms corresponds to the principle of the absolute and unitary perception of musical space.

Schoenberg did not conceive of the vertical dimension of musical space simply as pitch, but rather as pitch class. That is, he assumed that octave displacement would not destroy the perceptual identity of a musical configuration. His assumptions of perceptual equivalence under transposition, retrogression, inversion, and octave displacement are fundamental to the theory of 12-tone composition (Babbitt, 1961, 1965). This compositional technique employs the following procedure. A given ordering of the 12 tones within the octave is adopted. This tone row is repeatedly presented throughout the piece; however, the above transformations are allowed on each presentation. It is assumed that the row as an abstraction is perceived in its different manifestations.

Whether such transformations indeed result in perceptual equivalences is debatable. In the visual case we must have evolved perceptual mechanisms that preserve the identities of objects regardless of their orientation relation to the observer. An analogous argument cannot be made for the case of inversion and retrogression of sound sequences. A second doubt is based on general experience. Sound sequences may be unrecognizable when reversed in time, as the reader can determine by attempting to decode a segment of speech played backward. Furthermore, many inverted three-note combinations are perceptually very dissimilar. For example, the major and minor

[4]The use of the term "inversion" as defined by contemporary music theorists should not be confused with the traditional use of the term defined earlier.

Biblioteca del ... de Música - MADRID

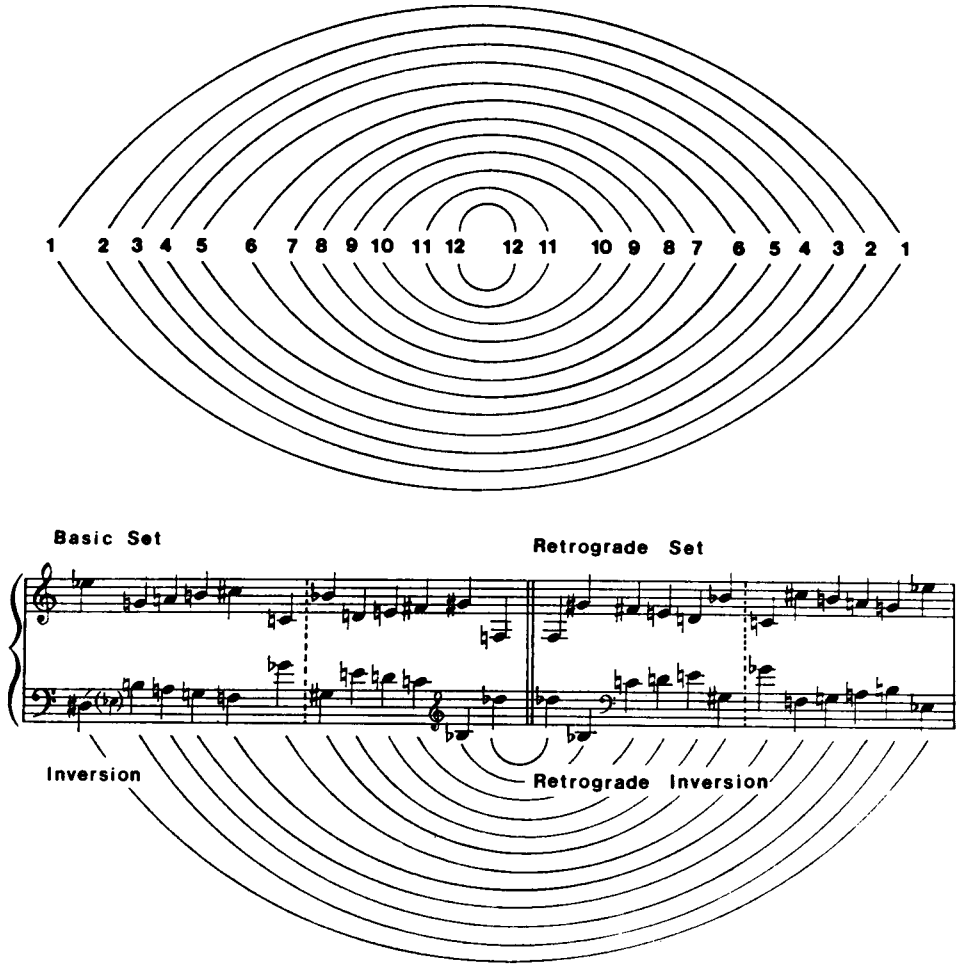



**Fig. 7.** Schoenberg's illustration of his theory of equivalence relations between pitch structures—taken from his *Wind Quartet, Op. 26* (from Schoenberg, 1951).

triads are inversions of each other; however, it appears quite implausible to regard them as perceptually equivalent. Rather it would seem that recognition of inverted and retrograde patterns is generally accomplished at a level of abstract coding equivalent to that which allows us to recite a segment of the alphabet backward, or to invert a sequence of numbers (Deutsch and Feroe, 1981).

At all events, a substantial body of music theory, aimed at defining equivalence and similarity relations between sets of pitches, is based on these assumptions of equivalence under retrogression and inversion, as well as the assumptions of pitch class and interval class identity. Although these theories are essentially concerned with stimulus structure, they do have implications about psychological representation.

A good example of this type of theory was provided by Chrisman (1971). He first defined the term "pitch-set" to refer to the set of elements derived from any collection of pitches, such that all members of a pitch class in the collection are represented by a unique element corresponding to that pitch class. In order to determine the intervallic relationships in a pitch-set, the elements of the pitch-set are given in ascending order, beginning with any pitch in the set. All members of the pitch-set are then contained within the octave above this initial pitch.

From any pitch-set whose elements have been placed in ascending order and within a single octave, one can construct an array which describes the intervallic structure of the set. Such an array consists of a linear succession of intervals and is termed a "successive-interval array." Each element of the array corresponds to the number of semitones between successive pitch-class representatives in the pitch-set. The array also includes the interval between the last note in the pitch-set and the first note an octave higher.

Cyclic reorderings of elements in each successive interval array are useful for determining intervallic relationships between pitch-sets with differing pitch contents. To take Chrisman's example, the pitch-sets S, T and V have different pitch contents and produce three different interval arrays: A, B, and C.

| Pitch-Set | Interval-Array |
|---|---|
| S = C, C♯, E, F♯, G, A♯, B | A = 1-3-2-1-3-1-1 |
| T = C, D♯, E, F, F♯, A, B | B = 3-1-1-1-3-2-1 |
| V = C, D, D♯, F♯, G, G♯, A | C = 2-1-3-1-1-1-3 |

When the elements in these arrays are cyclically reordered, the arrays are shown to be equivalent.

$$A = 1\text{-}3\text{-}2\text{-}1\text{-}3\text{-}1\text{-}1$$

$$B = 3\text{-}1\text{-}1\text{-}1\text{-}3\text{-}2\text{-}1 \qquad P_3(\underline{B}) = 1\text{-}3\text{-}2\text{-}1\text{-}3\text{-}1\text{-}1$$

$$C = 2\text{-}1\text{-}3\text{-}1\text{-}1\text{-}1\text{-}3 \qquad P_5(\underline{\overline{C}}) = 1\text{-}3\text{-}2\text{-}1\text{-}3\text{-}1\text{-}1$$

The sets T and V are, thus by this definition, shown to be transpositions of the set S.

Another, much more elaborate formulation was proposed by Forte (1973). Forte was concerned not only with the conditions under which sets of pitches should be considered equivalent, but also with defining similarity relationships between pitch-sets. He used two measures of similarity, one based on pitch class intersection and the other on interval class intersection. For other work on pitch-sets see Howe (1965), Lewin (1960, 1962), Perle (1972, 1977), and Teitelbaum (1965).

The extent to which the structures defined by such theories are processed by the listener remains to be determined. As noted by Garner (1974), some structures that exist in a stimulus configuration are perceived readily, others with difficulty, and yet others not at all. A fundamental problem with this body of theory concerns the basic equivalence assumptions on which it rests. The issue of interval class is a thorny one,

and the assumptions of equivalence under retrogression and inversion are also debatable. Reservations about these equivalence assumptions have also been raised recently by music theorists (Browne, 1974; Howe, 1974; Benjamin, 1974).

Other theorists have attempted to represent pitch relationships in terms of distances in a multidimensional space. Drobisch (1846, 1855) proposed that pitch be represented in three dimensions as a helix, with the vertical axis corresponding to pitch height, and tones separated by octaves lying closest within each turn of the helix. This representation reflects the perceptual closeness of the octave relationship. Shepard (Chapter 11) provides a detailed theoretical formulation that elaborates on this model.

Longuet-Higgins (1962a,b) has proposed that "tonal space" be represented as a three-dimensional array. Tones adjacent along the first dimension are separated by fifths, those adjacent along the second dimension by major thirds, and those adjacent along the third dimension by octaves. The intervals of tonal music then appear as vectors in this tonal space. If tones that are separated by octaves are treated as equivalent, an array is obtained such as shown on Fig. 8. Note that a closely related set of tones, such as comprise the C-major scale, forms a compact group, so that a key can be defined as a neighborhood in this space. Longuet-Higgins proposed that when presented with a segment of music, the listener initially selects a given region of space, thus attributing a key. However, if his choice forces him to engage in large jumps

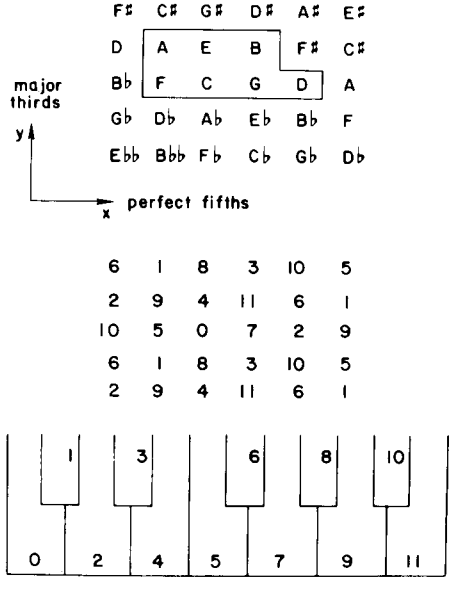



**Fig. 8.** Array hypothesized by Longuet-Higgins for the representation of "tonal space." See text for details (from Longuet-Higgins, 1978).

within this region, the listener abandons it and selects instead a region where the tones are more compactly represented, thus attributing a new key.

Another approach to the mapping of tonal space was taken by Krumhansl (1979). She performed an experiment in which subjects were presented with a set of context tones which were followed by two tones played in succession. The context tones were either the chord of the C-major triad or the C-major scale. Subjects judged on each trial how similar the first tone was to the second in the tonal system suggested by this context. Multidimensional scaling of the similarity ratings produced a three-dimensional conical structure around which tones were ordered according to pitch height. The components of the major triad formed a closely related cluster near the vertex of the cone; the remaining tones of the diatonic scale formed a less closely related subset that was further from the vertex; and the nondiatonic tones were widely dispersed, still further from the vertex. It is expected that different contexts would give rise to different patterns of similarity relationships using this technique. In the traditional music of our culture, the minor scale should produce a different configuration, and entirely different patterns would be expected from those familiar with other types of music. However, the study is important in demonstrating that pitch relationships are represented in a complex and well-defined fashion in a highly overlearned tonal context.

## IV. ALPHABETS AND HIERARCHIES

We next consider a further level of abstraction, in which pitch information is mapped onto a relatively small set of highly overlearned alphabets. Although these differ from one culture to another, the use of such alphabets appears to occur cross-culturally. The invocation of a relatively small number of alphabets, each of which consists of a relatively small number of steps, allows for music of considerable complexity without the penalty of heavy processing load (Miller, 1956; Garner, 1974). In the tonal music of our tradition, the 12-tone chromatic scale forms a parent alphabet from which a family of subalphabets is derived, such as major scale, minor scale, major triad, and so on. Each of these subalphabets is itself a family of subalphabets that are related by transposition. Other alphabets involve chord progressions, such as progression along the cycle of fifths.

The ready mapping of musical sequences onto pitch alphabets is reflected in the finding that short-term transposition often tends to occur along such alphabets rather than in terms of exact interval sizes (Fig. 9). The consequent alterations in interval size do not produce an impression of musical "incorrectness." This contrasts with transposition in long-term situations, where exact intervals are generally preserved instead (Deutsch, 1969; Attneave & Olson, 1971). Short-term perceptual equivalences, therefore, appear to be more heavily influenced by pitch content than are long-term perceptual equivalences.

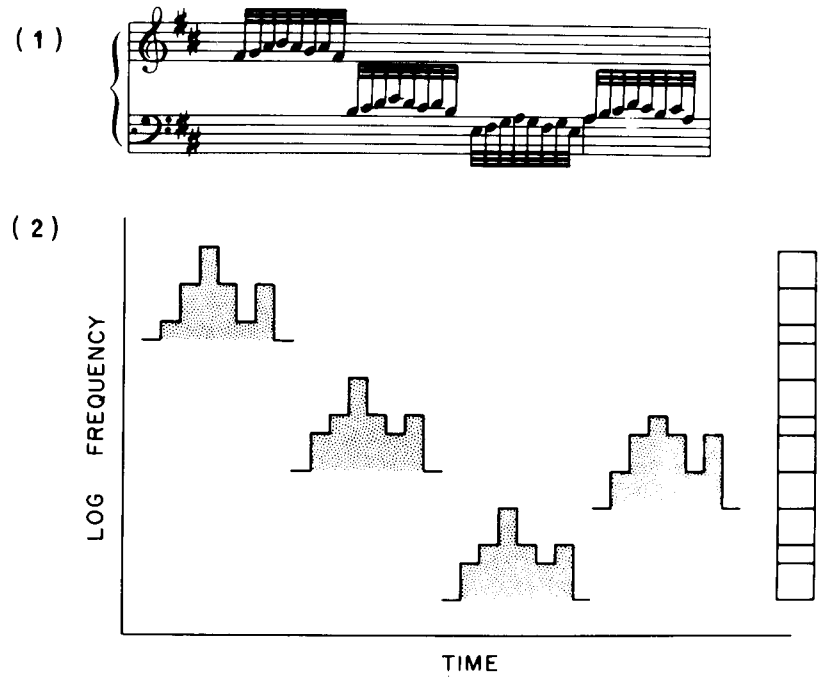



**Fig. 9.** Transposition along the alphabet of the scale. A given configuration is presented four times in succession at different positions along the scale. Since the scale has unequal intervals, there results a variation in the set of intervals involved. Ladder at right displays scale. From Bach, J. S. *The Well-Tempered Clavier, Book 1, Fugue V* (from Deutsch, 1977).

The invocation of scalar alphabets in short-term situations has been studied by several investigators. Francès (1958) found that listeners were better able to detect alterations in tonal than in atonal melodies. Similar conclusions were reached by Dewar (1974) and Dewar, Cuddy, and Mewhort (1977). Dewar also observed that for tonal melodies, when the altered tone was in the same diatonic scale as the other tones, discrimination accuracy was poorer than when the altered tone departed from this scale. Dowling (1978) further observed that listeners had considerable difficulty in distinguishing between exact transpositions of melodies to new keys and shifts along the same diatonic scale where intervals were not preserved.

The tonal music of our tradition is also composed of small segments that are systematically organized in hierarchical fashion. It is reasonable to suppose that such hierarchical organization reflects the ways in which musical information is abstracted and retained. As Greeno and Simon (1974) point out, many different types of information appear to be retained as hierarchies. In some instances, the information stored is in the form of concepts that refer to classes (Collins & Quillian, 1972). We also appear to retain hierarchies of rules (Gagné, 1962; Scandura, 1970), hierarchies of programs (Miller, Galanter, & Pribram, 1960), and hierarchies of goals in problem

solving (Ernst & Newell, 1969). Visual scenes appear to be represented as hierarchies of subscenes (Winston, 1973; Palmer, 1975; Navon, 1977; Hanson & Riseman, 1978). The phrase structure of a sentence lends itself readily to hierarchical interpretations (Yngve, 1960; Miller & Chomsky, 1963).

Experiments by Restle and Brown (1970 and Restle (1970) have demonstrated that we readily acquire serial patterns as hierarchies that reflect the structures of these patterns. In their experiments, subjects were presented with a row of six lights, which came on and off in repetitive sequence, and their task was to predict which light would come on next. The sequences were structured as hierarchies of operators. For example, if the basic subsequence is X = (1 2), then the operation R ("repeat of X") produces the sequence 1 2 1 2; the operation M ("mirror-image of X") produces 1 2 6 5; and the operation T ("transposition +1 of X") produces 1 2 2 3. By recursive application of such operations, long sequences can be generated that have compact structural descriptions. For example, the sequence 1 2 1 2 2 3 2 3 6 5 6 5 5 4 5 4 can be described as M(T(R(T(1)))) and corresponds to the *structural tree* shown on Fig. 10.

Using sequences constructed in this fashion, Restle and Brown (1970) showed that the probability of error in prediction increased monotonically with the level of transformation along such a structural tree. For example, the highest probability of error in a sequence such as on Fig. 10 occurred at Locations 1 and 9, the next highest at Locations 5 and 13, and so on. From this and other evidence, it was concluded that the observer organizes information in accordance with such structures. Parallel theoretical developments by Simon and his colleagues (Simon & Kotovsky, 1963; Simon & Sumner, 1968; Simon, 1972) and by others (Leewenberg, 1971; Jones, 1974, 1978; Vitz & Todd, 1967, 1969) also utilized hierarchies of operators.

Deutsch and Feroe (1981) argue that sequences in tonal music have characteristics not reflected in the above formalisms, and they advance a model for the encoding of pitch sequences as hierarchies of operators, which takes these characteristics into account. In essence, the model may be characterized as a hierarchical network at each level of which, structural units are represented as organized sets of elements. Elements that are present at any one level are elaborated by further elements so as to form structural units at the next-lower level, until the lowest level is reached. It is also

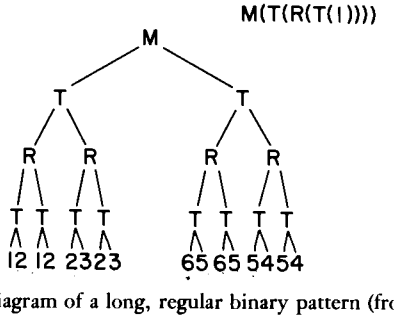





**Fig. 10.** Tree diagram of a long, regular binary pattern (from Restle, 1970).

assumed that Gestalt principles such as Proximity and Good Continuation contribute to organization at each hierarchical level.

The following are the rules for a simplified version of the system; however, the reader is referred to Deutsch and Feroe (1981) for a description of the full system:

1. A *structure* is notated by $(A_1, A_2, \ldots A_{l-2}, A_{l-1}, *, A_{l+1}, A_{l+2}, \ldots, A_n)$, where $A_j$ is one of the operators n, p, s, $n^i$, or $p^i$. (A string of length k of an operator A is abbreviated kA.)

2. Each structure $(A_1, A_2, \ldots, *, \ldots, A_n)$ has associated with it an alphabet, $\alpha$. The combination of a structure and an alphabet is called a *sequence* (or *subsequence*). This, together with the reference element r, produces a *sequence of notes*.

3. The effect of each operator in a structure is determined by that of the operator closest to it, but on the same side as the asterisk. Thus, the operator n refers to traversing one step up the alphabet associated with the structure. The operator p refers to traversing one step down this alphabet. The operator s refers to remaining in the same position. The two operators $n^i$ and $p^i$ refer to traversing up or down i steps along the alphabet, respectively.

4. The values of the sequence of notes $(A_1, A_2, \ldots, *, \ldots, A_n)$, $\alpha$, r, where $\alpha$ is the alphabet and r the reference element, are obtained by taking the value of the asterisk to be that of r.

5. To produce another sequence from the two sequences $A = (A_1, A_2, \ldots, *, \ldots, A_m)$, $\alpha$, and $B = (B_1, B_2, \ldots, *, \ldots, B_n)$, $\beta$, where $\alpha$ and $\beta$ are two alphabets, we define the compound operator pr (prime). $A$[pr]$B$;r, where r is the reference element, refers to assigning values to the notes produced from $(B_1, B_2, \ldots, *, \ldots, B_n)$, such that the value of * is the same as the value of $A_1$, when the sequence $A$ is applied to the reference element r. Values are then assigned to the notes produced from $(B_1, B_2, \ldots, *, \ldots, B_n)$, such that the value of * is the same as the value of $A_2$, and so on. This gives a sequence of length m × n. Other compound operators such as inv (inversion) and ret (retrograde) are analogously defined.

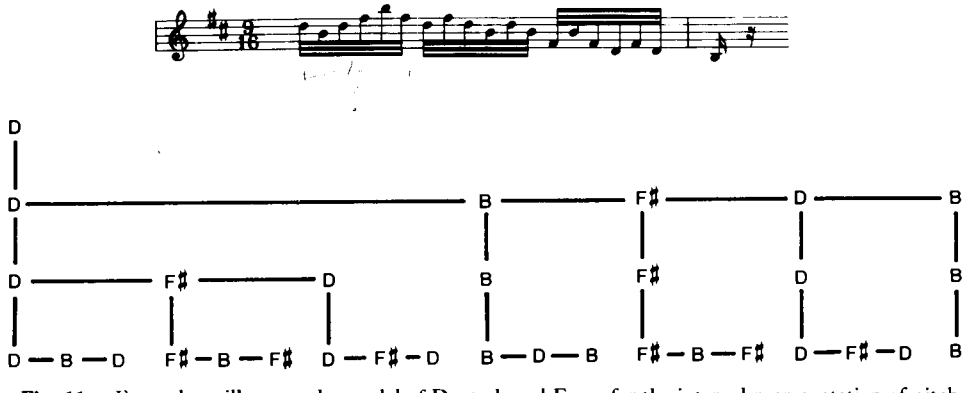



**Fig. 11.** Example to illustrate the model of Deutsch and Feroe for the internal representation of pitch sequences (see text for details). Passage is from Bach's *Sinfonia 15*, BWV 801 (from Deutsch & Feroe, 1981).

To give a simple example of the use of the formalism, Sequence 1 of Fig. 16 is represented as

$$A = (*,3p)G_{tr}$$
$$B = (*,p,n)Cr$$
$$S = A[pr]B,G_5$$

where $G_{tr}$ represents the G-major triad, Cr the chromatic scale, and $G_5$ the reference element.

It may be observed that this form of representation results in considerable parsimony of encoding. Further, the employment of distinct alphabets at different structural levels enables an encoding in terms of proximal relationships, so that the principle of proximity is upheld (see Chapter 4).

A more complex musical example is shown on Fig. 11 and is represented as on three structural levels as follows:

$$A = (*,4p)b_{tr}$$
$$B = (*,n,p)b_{tr}$$
$$S = A[pr](B,4(*))[inv,5pr](B,(*))D_5$$

So far we have considered the processing of a single melodic line. However, tonal music generally involves several such lines, and even where only one is presented a harmonic progression is often implied. It is assumed that such progressions may also be encoded in hierarchical fashion. The use of parallel linear sequences, which must themselves combine to form an acceptable harmonic sequence, places constraints on the choice of elements in each sequence, and this in turn serves to reduce processing load.

The present model may be related to characterizations of hierarchical structure in tonal music advanced by music theorists. The most important development in this field is that of Heinrich Schenker (1868–1935), who proposed a hierarchical system for tonal music that has points of similarity with the system proposed by Chomsky for linguistics (Chomsky, 1963). On Schenker's system music is considered as a hierarchy in which pitch events at any one level are considered "prolonged" by sequences of pitch events at the next-lower level. Three basic levels are distinguished in the system. First, there is the surface representation or *foreground*; second, there is the *middleground*; and third, there is the *background* or *Ursatz*. The *Ursatz* is itself considered a prolongation of the triad. For literature on Schenkerian analysis and related theoretical formulations, see particularly Lerdahl and Jackendorff (1977), Meyer (1973), Narmour (1977), Salzer (1962), and Schenker (1956,1973) and Yeston (1977).

## V. MEMORY SYSTEMS

We assume that separate memory systems exist for retaining information at different levels of abstraction. Craik and Lockhart (1972) have made the general argument that the higher the level of abstraction, or "depth of processing" of information,

the longer its persistence in memory. This may indeed be true of music. It is clear from general experience that memory for melodic and harmonic intervals persists for considerably longer than memory for absolute pitch values (Deutsch, 1969; Attneave & Olson, 1971); memory for higher order abstractions may well persist for longer still. Such differences in the persistence of memory would have the consequence that when retention of a musical sequence is examined following different time periods, it is likely to reflect different forms of encoding.

In previous sections we have examined the question of encoding in detail, and we now turn to a consideration of the influences acting on information in storage. This question has been explored in depth for the system retaining absolute pitch values, and a little is also known about the system retaining interval information, and the system retaining higher level abstractions.

### A. The System Retaining Absolute Pitch Information

When listeners make pitch comparison judgments between tones that are separated by a silent retention interval, accuracy declines gradually as the retention interval is lengthened (Koester, 1945; Harris, 1952; Bachem, 1954; Wickelgren, 1966, 1969). However, the rate of memory deterioration here is very slow. For example, Harris (1952) found that following a retention interval of 15 sec, a difference of a small fraction of a semitone could still be reliably discriminated. This stands in sharp contrast to comparison performance when a sequence of extra tones is interpolated during the retention interval. In an experiment by Deutsch (1970a), listeners were selected for obtaining a score of 100% correct in comparing tone pairs that were separated by a six-second retention interval. The test tones were either identical in pitch or they differed by a semitone. When eight extra tones were then interpolated during this interval, the error rate rose to 40%, even though the listeners were instructed to ignore the interpolated tones. In a further experiment by Deutsch (1970b), this performance decrement was shown not to be due to attention distraction nor to a general memory overload, since only a minimal decrement occurred when spoken numbers were interpolated instead. This was true even when the subjects were simultaneously required to recall the interpolated numbers. It was concluded that pitch memory is subject to interference caused specifically by other tones.

Further experiments demonstrated that tones interact with each other in memory in an orderly and systematic fashion. Such interactions were shown to be a function of both the pitch relationships between the interacting tones and their closeness in serial position.

In one experiment, the effect of a tone that formed part of a sequence interpolated between two test tones was studied as a function of its pitch relationship to the first test tone (Deutsch, 1972b). Subjects compared the pitches of two tones that were separated by a five-second retention interval during which a sequence of six extra tones was interpolated. The test tones were taken from the 12-tone chromatic scale, and ranged from Middle C to the B above. In half of the sequences the test tones were identical in pitch; in the other half they differed by a semitone. The pitches of the

intervening tones were also taken from the 12-tone chromatic scale and ranged from the F♯ below Middle C to the F an octave and a half above. The tones were chosen at random from this range with certain restrictions.[5]

The experiment consisted of eight conditions. In all conditions but the last, a tone whose pitch bore a critical relationship to the pitch of the first test tone was placed in the second serial position of the intervening sequence. The relationship between the critical intervening tone and the first test tone varied between identity to a whole-tone separation. A unique value of pitch separation was incorporated in each of the seven conditions, these values being placed at equal intervals of 1/6 tone within this whole tone range. However, in the eighth condition, the pitch of the tone in the second serial position was chosen in the same way as were the other pitches in the intervening sequence. This last ("null") condition therefore provided a baseline against which the effects of the critical interpolated tone could be evaluated.

It was found that the effect of the critical interpolated tone on memory for the test tone varied systematically as a function of their pitch relationship. When the critical interpolated tone was identical in pitch to the first test tone, memory facilitation occurred. With increasing pitch separation between the first test tone and the critical interpolated tone, the error rate rose progressively; it peaked at a separation of 2/3 tone and declined to baseline at roughly a whole tone separation.

Further experiments have reinforced and elaborated on these findings of specific interactions within the pitch memory system. One study investigated the effects of including in an intervening sequence a tone that was a semitone removed from the first test tone (Deutsch, 1973b). It was found that when the test tones were identical in pitch, including in the intervening sequence a tone that was either a semitone higher or a semitone lower, an increase in errors resulted. When both a tone a semitone higher and also a tone a semitone lower were included in the intervening sequence, a substantially greater increase in errors occurred than when only one of these was included. Further, when the test tones differed in pitch by a semitone, including in the intervening sequence a tone that was identical in pitch to the second test tone; there was a substantial increase in errors. An increase in errors that was however significantly smaller also occurred when the critical intervening tone was a semitone removed from the first test tone, but on the opposite side of the pitch continuum to the second test tone. When both of these tones were included in the same intervening sequence, a significantly greater increase in errors was produced than when only one of these was included.

This experiment demonstrated at least two separable disruptive effects in pitch memory. First, the inclusion of a tone that was a semitone removed from the tone to be remembered produced a small but significant disruptive effect which cumulated in size when two such tones, one higher and the other lower than this tone, were both included. Second, a significantly larger disruptive effect occurred when the test tones

[5]No sequences contained two tones of the same pitch class. Further, all tones were excluded from any sequence that lay within, and including, a whole tone range in either direction from the first test tone, or that were displaced by an octave from this range. This gap was necessary to prevent the random inclusion of tones in the critical range under study.

differed in pitch and the critical interpolated tone was identical in pitch to the second test tone.

To explain the second (and larger) of these two disruptive effects, the following hypothesis was advanced (Deutsch, 1972c). Memory for the pitch of a tone is laid down simultaneously both on a pitch continuum and also on a temporal or order continuum. As time proceeds, this memory distribution spreads in both directions, but particularly along the temporal continuum. As a result of this spread, when a tone of the same pitch as the second test tone is presented in the interpolated sequence, the subject sometimes concludes that this had been the first test tone. That is, errors of misrecognition result from the subject correctly recognizing that a tone of identical pitch to the second test tone had occurred, but being uncertain when it had occurred.

This hypothesis, which is described in detail in Deutsch (1972c), gives rise to various predictions. First, we should expect errors of misrecognition to be more numerous when the critical interpolated tone is placed early in the intervening sequence rather than late, since here there would be a greater chance of temporal or order confusion. That is, the closer the critical interpolated tone is to the first test tone, the more difficult it should be to discriminate their two positions along a temporal or order continuum and so the greater the number of misrecognition errors.

So, in a further experiment subjects made pitch recognition judgments between tone pairs that were separated by a sequence of six interpolated tones. In sequences where the test tones differed in pitch, a tone that was identical in pitch to the second test tone was sometimes placed in the second serial position of the intervening sequence, sometimes in the fifth serial position, and sometimes no such tone was included. It was found that the critical interpolated tone produced an increase in errors at both serial positions compared with sequences containing no such tone; however, the effect was much more pronounced when the critical tone was placed early in the intervening sequence rather than late (Deutsch, 1975b). This finding therefore corroborated the hypothesis.

Also in this experiment, in sequences where the test tones were identical in pitch, the effect was studied of including a tone that was a semitone removed from the test tone, as a function of its serial position in the interpolated sequence. No serial position effect was here found. So it seems that the first source of disruption, in contrast, does not depend on the serial position of the critical interpolated tone.

Given this large serial position effect based on interpolating a tone of identical pitch to the second test tone, we can ask whether this is due to a deterioration of information along a temporal continuum, or whether this effect is a function of order rather than time. An experiment was performed to investigate this question. Subjects compared the pitches of two tones that were separated by a sequence of six interpolated tones. In sequences where the test tones differed in pitch, a tone identical to the second test tone was placed either in the second serial position of the intervening sequence, or in the fifth serial position, or no such tone was interpolated. The experiment studied the effect of lengthening the pause following the first test tone, and also of lengthening the pause preceding the second test tone. The different temporal conditions are shown on Fig. 12. Now if this serial position effect were due simply to

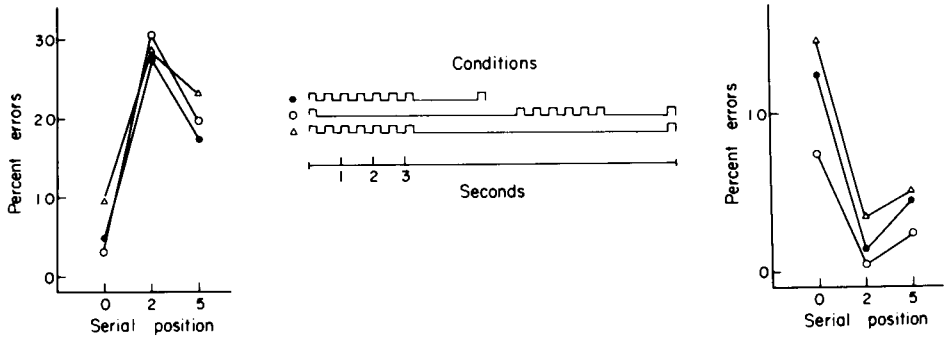



**Fig. 12.** Temporal parameters used in two experiments investigating the effects of unfilled delays on pitch recognition. Graph on left plots error rates in sequences where the test tones differed and a tone whose pitch was identical to the second test tone was placed in the intervening sequence. Graph on right plots error rates in sequences where a tone of identical pitch to the first test tone was interpolated. Numbers 2 and 5 indicate that the critical tone was placed in the second and fifth serial positions of the intervening sequence, respectively; "0" indicates that no such tone was interpolated.

interactions along a temporal continuum, we should expect substantial changes in the shape of the effect to result from such temporal manipulations. However, as shown on Fig. 14, the serial effect occurred under all temporal conditions. It would appear, therefore, that this deterioration of information takes place along an order continuum, which is not sensitive to temporal variations, at least in the range investigated here (Deutsch, in preparation).

There is a second prediction from the hypothesis of loss of temporal or order information. Suppose that we plotted errors precisely as a function of the pitch relationship between the first test tone and a critical interpolated tone, in 1/6 tone steps as described earlier, and we also varied the pitch difference between the two test tones. Then, in sequences where the critical interpolated tone and the second test tone are placed on the same side of the pitch continuum relative to the first test tone, the peak of errors should occur where the critical interpolated tone is identical in pitch to the second test tone. That is, a shift in the pitch of the second test tone, when the pitch of the first test tone is held constant, should result in a parallel shift in the peak of errors produced by the critical interpolated tone.

This prediction was tested in another experiment (Deutsch, 1975b). Here, when the test tones differed in pitch, this difference was either 1/3 tone, 1/2 tone, or 2/3 tone. Errors were plotted as a function of the pitch of a tone placed in the second serial position of an interpolated sequence, whose relationship to the pitch of the first test tone varied in 1/6 tone steps between identity and a whole tone separation. Whenever the second test tone was higher in pitch than the first, the critical interpolated tone was also higher; whenever the second test tone was lower, the critical interpolated tone was also lower. So, when the second test tone and the critical interpolated tone were separated from the first test tone by the same pitch distance, they were also identical in pitch to each other.

Biblioteca del Real Conservatorio Superior

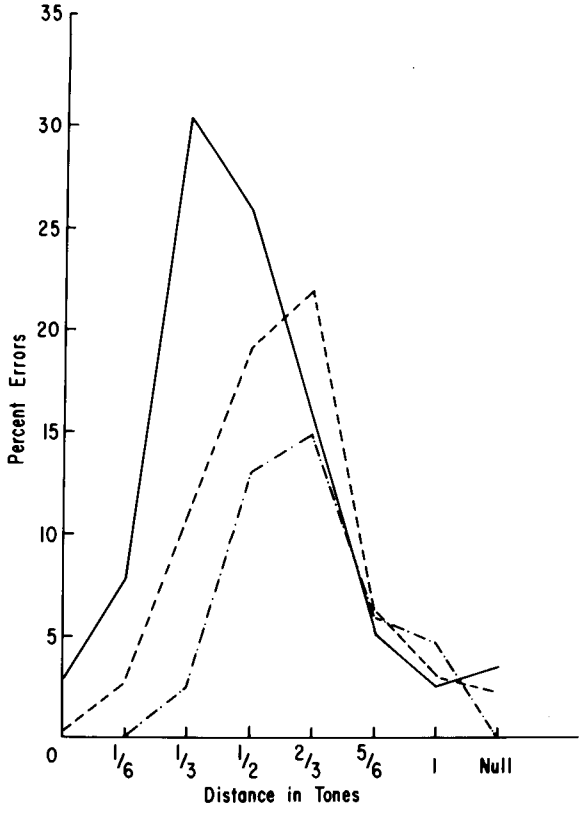



**Fig. 13.** Percent errors in pitch comparisons in sequences where the test tones differed in pitch, and the critical interpolated tone was placed on the same side of the first test tone along the pitch continuum. Errors were plotted as a function of the pitch relationship between the first test tone and the critical interpolated tone. (—) Test tones separated by 1/3 tone. (---) test tones separated by 1/2 tone. (-·-) test tones separated by 2/3 tone (from Deutsch, 1975b).

The results of this experiment are shown in Fig. 13. It can be seen that when the test tones were 1/3 tone apart, errors peaked when the critical interpolated tone was 1/3 tone removed from the first test tone, and so when it was identical in pitch to the second test tone. Similarly, when the test tones were 2/3 tone apart, errors peaked when the critical interpolated tone was 2/3 tone removed from the first test tone, and so again when it was identical in pitch to the second test tone. So, here a shift in the pitch of the second test tone relative to the first did indeed result in the predicted shift in the peak of errors produced by the critical interpolated tone. But note also that when the first and second test tones were a semitone apart, errors peaked not at a semitone but at 2/3 tone. This can be explained by assuming that the first source of disruption that peaks at 2/3 tone is superimposed on the present source of disruption produced by a relationship of identity or near identity between the second test tone and the critical interpolated tone.

In considering possible bases for this first source of disruption, two points should be noted. The first is that the relative frequency range over which this occurs corresponds well with the range over which centrally acting lateral inhibition has been found in physiological studies of the auditory system (Klinke, Boerger, & Gruber, 1969, 1970). Second, we can remember that the error rate cumulates when two such disruptive tones are interpolated, placed one on either side of the test tone along the pitch continuum (Deutsch, 1973b). There is an analogy here with lateral inhibitory interactions found in systems handling sensory information at the incoming level, where there is a cumulation of inhibition from stimuli placed on either side of the test stimulus (Ratliff, 1965). Evidence for lateral inhibition has been found in the system that handles pitch information at the incoming level (Carterette, Friedman, & Lovell, 1969, 1970; Houtgast, 1972). It was therefore theorized that elements of the pitch memory system are arranged as a recurrent lateral inhibitory network, analogous to those found in systems handling incoming sensory information. Elements of this system are activated by tones of specific pitch, and are organized tonotopically on a log frequency continuum (Deutsch & Feroe, 1975.)

Now, if this were so, we might hope to obtain an effect which would not be expected on other grounds. It has been found in physiological studies of peripheral receptors that when a unit that is inhibiting a neighboring unit is itself inhibited by a third unit, this releases the originally inhibited unit from inhibition. This phenomenon is known as disinhibition. Applying this to our present situation, we might expect that if a tone that was inhibiting memory for another tone were itself inhibited by a third tone, this could cause memory for the first tone to return. More specifically, in sequences where the test tones are identical in pitch, if two critical tones were interpolated—one always 2/3 tone removed from the first test tone and the other further removed along the pitch continuum—then the error rate should vary systematically as a function of the pitch relationship between the two critical interpolated tones. The error rate should be highest when these two tones are identical in pitch, decline as the second critical tone moves away from the first, dip maximally at a 2/3 tone separation, and then return to baseline. The curve produced should therefore be roughly the inverse of the curve plotting the original disruptive effect.

Accordingly, the following experiment was performed. Subjects compared the pitches of two test tones, which were separated by a sequence of six interpolated tones. A tone which was 2/3 tone removed from the first test tone was always placed in the second serial position of the interpolated sequence. Errors were then plotted as a function of the pitch of a further tone, which was placed in the fourth serial position, whose relationship to the tone in the second serial position varied in 1/6 tone steps between identity and a whole-tone separation. As can be seen in Fig. 14, a systematic return of memory was indeed obtained. The error rate in sequences where the second critical interpolated tone was identical in pitch to the first was significantly higher than baseline; further, the error rate where the two critical interpolated tones were separated by 2/3 tone was significantly lower than baseline. A first-order inhibitory function was also obtained experimentally, using subjects selected on the same criterion as for the disinhibition study. This function was then used to calculate the

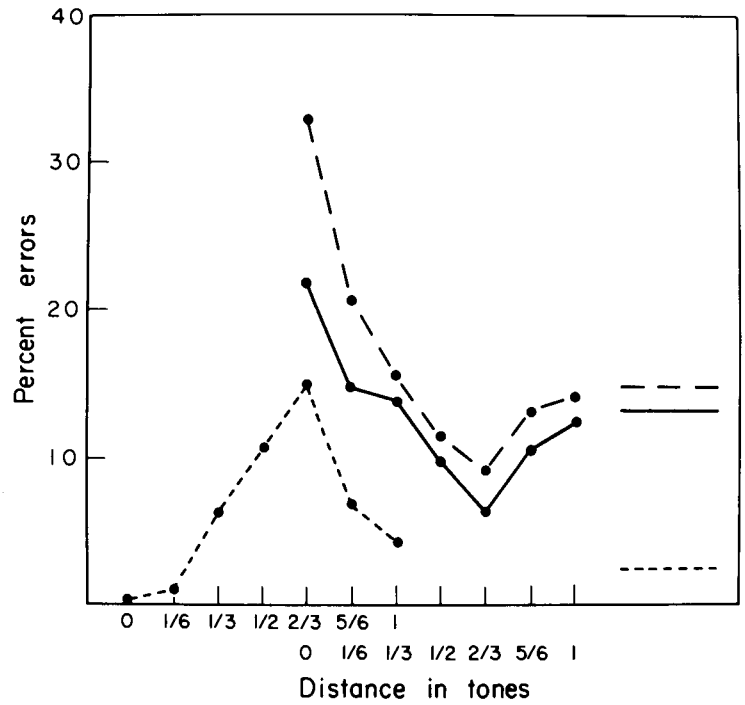



**Fig. 14.** Percent errors in pitch recognition obtained experimentally and predicted theoretically. Open triangles display percent errors in a baseline experiment that varied the pitch relationship between a test tone and a critical interpolated tone. (The open triangle at the right displays percent errors where no tones were interpolated in the critical range under study.) Filled circles display percent errors in an experiment where a tone that was 2/3 tone removed from the test tone was always interpolated. Errors are plotted as a function of the pitch relationship between this tone and a second critical interpolated tone that was further removed along the pitch continuum. Open circles display percent errors for the same experimental conditions predicted theoretically from the lateral inhibition model. (Filled and open circles at the right display percent errors obtained experimentally and assumed theoretically where no further critical tone was interpolated (from Deutsch & Feroe, 1975).

theoretical disinhibition function. As shown in Fig. 13, there is a good correspondence between the disinhibition function derived experimentally and that derived theoretically. This experiment provides strong evidence that pitch memory elements are indeed arranged as a lateral inhibitory network, analogous to those handling incoming sensory information.

### *1. Octave Generalization Effects*

So far we have been considering interactions in memory between tones that are separated by less than an octave. We can now ask whether these effects take place along an array that is organized simply in terms of pitch, or whether an abstracted octave array, such as hypothesized on pp. 275–276, is also involved.

In one experiment (Deutsch, 1973a), subjects compared the pitches of two tones that were separated by a sequence of six interpolated tones. This experiment studied the effects of including in the interpolated sequence tones that bore the same relationships to the test tones as had been found earlier to produce disruption, but that were further displaced by an octave. In sequences where the test tones were identical in pitch, the effects were studied of interpolating two tones, one a semitone higher than the test tone and the other a semitone lower, except that the critical tones were also displaced an octave up or down. In sequences where the test tones differed, the effects were investigated of interpolating a tone of the same pitch as the second test tone, but that was also displaced an octave up or down. It was found that a substantial generalization of disruptive effect occurred from tones placed in the higher octave, and a weaker effect occurred from tones placed in the lower octave. However, the error rate was greatest when the critical tones were placed in the middle octave. From this pattern of results it was concluded that these disruptive effects take place along both a monotonic pitch continuum and also along an abstracted octave array.

### 2. *Memory Consolidation*

When a tone of identical pitch to the first test tone is included in an intervening sequence, the effect on memory is facilitatory rather than disruptive. In one experiment, subjects compared the pitches of two tones that were separated by a sequence of interpolated tones. In one condition, four tones were interpolated; in another, six tones were interpolated; in yet another, six tones were again interpolated, with a tone of identical pitch to the first test tone placed in the second serial position of the interpolated sequence. It was found that the error rate was lowest in sequences where the pitch of the first test tone was repeated, even compared with sequences containing fewer interpolated tones. We can conclude that the pitch memory system is subject to consolidation through repetition (Deutsch, 1975c).

However, in another experiment it was found that this consolidation effect is very sensitive to the serial position of the repeated tone. Here, subjects made pitch comparison judgments between tone pairs that were separated by a sequence of six interpolated tones. In some sequences a tone of the same pitch as the first test tone was placed in the second serial position of the interpolated sequence; in other sequences such a tone was placed in the fifth serial position; and in yet other sequences no such tone was interpolated. It was found that the facilitation effect was much more pronounced when the repeated tone was in the second serial position than in the fifth; indeed, this effect was statistically significant only for sequences where the repeated tone was in the second serial position (Deutsch, 1975c).

It was hypothesized that this consolidation effect results from the same process as produces the errors of misrecognition discussed above—namely, the spread of memory distribution along a temporal or order continuum. When two such distributions overlap, the overlapping portions sum, resulting in a stronger memory trace (Deutsch, 1972c). This again raises the question of whether this serial position effect is based on temporal factors, or whether it is simply a function of order. One might,

for example, hypothesize that temporal proximity determines the amount of consolidation in pitch memory. If this were the case, then interpolating a long pause between the first test tone and the first interpolated tone should substantially reduce the consolidation effect. Alternatively, one might hypothesize that some mechanism tags the incoming stimuli in terms of order, and that consolidation takes place along such an order continuum, independent of temporal variations.

An experiment was therefore carried out to evaluate the effects of varying temporal parameters on the memory consolidation effect. Subjects compared the pitches of two tones that were separated by a sequence of six interpolated tones. A tone that was identical in pitch to the first test tone was placed either in the second serial position of the intervening sequence, or in the fifth serial position, or no such tone was included. The experiment studied the effect of lengthening the pause following the first test tone, and also of lengthening the pause preceding the second test tone. The different temporal conditions together with their error rates are shown on Fig. 12. These temporal manipulations did not significantly alter the consolidation effect as a function of serial position. It appears, therefore, that within this temporal range, consolidation takes place along a continuum that is organized in terms of order independent of time (Deutsch, in preparation).

### B. The System Retaining Interval Information

We now briefly consider the influences acting on interval information in storage. Deutsch (1975d) proposed that memory for such information is based on a continuum whose elements are activated by the simultaneous or successive presentation of pairs of tones. Tone pairs that stand in the same ratio project onto the same elements and so onto the same point along the continuum; tone pairs standing in closely similar ratios project onto adjacent points; and so on. It was further hypothesized that interactive effects take place along this continuum that are analogous to those that occur within the system retaining absolute pitch values. Such effects would include consolidation through repetition and similarity-based interference.

An experiment was performed to test this hypothesis (Deutsch, 1978b). Subjects compared the pitches of two test tones that were both accompanied by tones of lower pitch. The test tones were either identical in pitch or they differed by a semitone. However, the tone accompanying the first test tone was always identical in pitch to the tone accompanying the second test tone. Thus, when the test tones were identical, the intervals formed by the test tone combinations were also identical. And when the test tones differed, the intervals formed by the test tone combinations also differed.

The test tone combinations were separated by a sequence of six interpolated tones. The tones in the second and fourth serial positions of the interpolated sequence were also accompanied by tones of lower pitch. It was found that when the intervals formed by the interpolated combinations were identical in size to the interval formed by the first test tone combination, the error rate was lower than when the sizes of the intervals formed by the interpolated combinations were chosen at random. Further,

when the intervals formed by the interpolated combinations differed in size by a semitone from the interval formed by the first test tone combination, the error rate was higher than when the sizes of the intervals formed by the interpolated combinations were chosen at random.

This experiment, therefore, demonstrates the presence of both consolidation through repetition and also similarity-based interference in memory for harmonic intervals. This indicates that the system retaining such information is similar in organization to the system retaining absolute pitch values.

### C. Interactions between These Systems

So far we have examined interactions that take place within a given memory system. We now turn to a consideration of how the outputs of the different systems interact in determining memory judgments. First, we examine ways in which abstracted information influences judgments of absolute pitch values. Then, we consider how information in the pitch memory system might in turn influence memory for pitch abstractions.

Since relational information is retained in parallel with pitch information, we might expect that judgments of sameness or difference in the pitches of two test tones would be biased by a sameness or difference in the relational context in which these test tones are placed. This possibility was investigated by Deutsch and Roll (1974). Subjects compared the pitches of two tones that were both accompanied by tones of lower pitch. The test tone combinations were separated by a retention interval during which six extra tones were interpolated. In some conditions the harmonic intervals formed by the test tone combinations were identical, and in others they differed; and these patterns of relationship were present both when the test tones were identical in pitch and also when these differed. It was found that relational context had a strong influence on pitch recognition judgments. When the test tones were identical but were placed in different relational contexts, there resulted an increased tendency for their pitches to be judged as different. Further, when the test tones differed in pitch, but were placed in an identical relational context, there resulted an increased tendency for their pitches to be judged as identical. It was further found that when the test tones differed, and the test tone combinations formed intervals that were inversions of each other, there also resulted an increase in errors of misrecognition. It was concluded that this misrecognition effect was based on the perceptual equivalence of the inverted intervals.

Deutsch (in press) performed an analogous experiment to study the effect of melodic relational context. Subjects compared the pitches of two tones that were each preceded by tones of lower pitch. The test-tone combinations were again separated by a retention interval during which six extra tones were interpolated. A strong effect of melodic context was demonstrated, analogous to that found for harmonic relational context.

Melodic relationships have been shown to influence pitch recognition judgments in


Conservatorio Superior de Música - Oviedo

another way. In listening to sequences, we process not only the individual tones, but also the melodic intervals between them. These intervals then provide a framework of pitch relationships to which the test tones can be anchored. Interpolated sequences forming melodic configurations that are more easily processed should then be associated with enhanced performance.

As described in Chapter 4, there is considerable evidence that melodic sequences are processed more effectively when these are composed of intervals of smaller size rather than larger (reflecting the operation of the principle of Proximity). One would therefore expect that interpolated sequences composed of smaller melodic intervals should be associated with higher performance levels than interpolated sequences composed of larger intervals. In an experiment by Deutsch (1978a), subjects compared the pitches of two tones that were separated by a sequence of six interpolated tones. Performance was studied under four conditions. In Condition 1 the interpolated tones were chosen at random from within a one-octave range, and they were ordered at random. Condition 2 was identical to Condition 1 except that the interpolated tones were arranged in monotonically ascending or descending order, with the result that the average size of the melodic intervals in the sequence was reduced. In Condition 3 the interpolated tones were chosen at random from within a two-octave range and were also ordered at random. Condition 4 was identical to Condition 3 except that the interpolated tones were arranged in monotonically ascending or descending order.

As shown on Fig. 15, the error rate was found to increase with an increase in the average size of the melodic intervals comprising the sequence. There was no evidence that monotonic ordering of the interpolated tones had any effect (apart from producing a smaller average interval size), though this might have been hypothesized from the principle of Good Continuation (see Chapter 4).

It has been shown that there is a striking cross-cultural tendency for the frequency of occurrence of a melodic interval to be inversely correlated with its size (Ortmann, 1926; Merriam, 1964; Fucks, 1962; Dowling, 1967; Jeffries, 1974; Deutsch, 1978d).

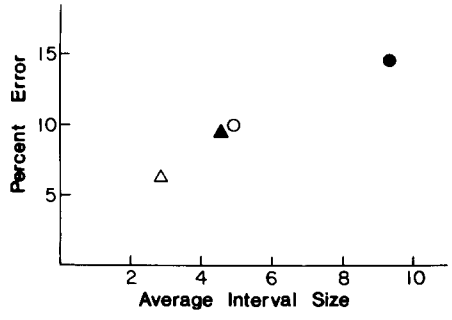



**Fig. 15.** Percent errors in pitch comparisons as a function of the average size of melodic interval in the sequence. Interpolated tones span a one-octave range: ▲ tones ordered at random; ● tones ordered monotonically. Interpolated tones span a two-octave range: ● tones ordered at random; ○ tones ordered monotonically (from Deutsch, 1978c).

One might hypothesize that this tendency is based on an increasing difficulty in the processing of melodic intervals as the sizes of these intervals increase. As indicated from the present experiment, this should in turn result in decreased accuracy in pitch recognition.

Two related studies should here be cited. Deutsch (1974) had subjects compare the pitches of two test tones that were separated by a sequence of eight interpolated tones. In one condition the interpolated tones were all drawn from the same octave as the test tones. In a second condition the interpolated tones were all drawn from the octave higher. In a third condition they were all drawn from the octave lower. In a fourth condition half of the interpolated tones were drawn from the octave higher than the test tones and the other half from the octave lower, the order of octave placement being random. It was found that when the interpolated tones were drawn from a single adjacent octave, the error rate was lower than when they were all drawn from the same octave. However, when the interpolated tones were drawn from both the octave higher and also the octave lower, the error rate was highest of all. In this last condition the average size of the intervals formed by successive tones in the sequence was considerably larger than in the other conditions, and it was concluded that these large jumps made it very difficult to make use of the melodic information in the interpolated sequence.

Another related study is that of Olson and Hanson (1977). They found, using several different paradigms, that increased errors in pitch recognition were associated with an increased pitch distance between the test tones and the interpolated tones. Since only three interpolated tones were used in this study, this increased pitch distance resulted in a significant increase in average interval size. These authors suggest an interpretation of their results that is very similar to the present line of reasoning.

Evidence for interactions based on more complex encodings has also been obtained. Krumhansl (1979) required subjects to compare the pitches of two test tones that were separated by a sequence of interpolated tones. Some of the sequences were constructed so as to suggest a tonality. It was found that for such sequences, when the first test tone was in the same scale as the interpolated tones, recognition performance was better than when it was outside this scale.

We should expect that interactions taking place within the pitch memory system would in turn affect memory for abstracted information. For example, given the memory consolidation effect produced by a repeated tone (Deutsch, 1975c), we should expect that sequences containing repeated tones would be better remembered, even in abstracted form, than sequences not containing repeated tones. Similarly, given the large disruptive effect produced by interpolating two tones, one a semitone higher than the tone to be remembered and the other a semitone lower (Deutsch, 1973b), we should expect that music composed in scales consisting mostly of semitonal steps would be more difficult to remember than music composed in scales consisting mostly of whole tone steps.

Erickson (1978) has argued that low-level effects, such as described here, have probably exerted a strong influence on the evolution of music systems. For example,

it appears true cross-culturally that musical sequences tend to contain one or two tones that are repeated considerably more often than others. In our tonal system, for example, the first tone of the scale (the tonic) has this property. Given the finding of consolidation through repetition, we should expect that such repeated tones should be particularly well remembered. Second, it appears true cross-culturally that small melodic intervals occur considerably more often than large ones. We have also seen that recognition of the pitch of a tone is better in the context of a sequence consisting of small melodic intervals rather than of large ones (Deutsch, 1978c). When we consider these two low-order effects together, we can see that a considerable processing advantage is to be gained from a system in which there are a limited number of anchor tones—which are well remembered through repetition—surrounded by satellite tones which are linked to these anchor tones by pitch proximity. Thus, examination of these low-order effects enables us to develop an understanding of why musical systems have evolved as they did. Such systems should therefore not be considered as arbitrary sets of rules but rather as reflecting constraints imposed by our processing mechanisms.

### D. Memory for Hierarchically Structured Tonal Sequences

We now turn to a consideration of memory for tonal information which is projected onto highly overlearned pitch alphabets, and which is organized in hierarchical fashion. Deutsch and Feroe (1981) have proposed that, at this level, tonal sequences are coded and retained as hierarchies of structures, each of which is associated with a given pitch alphabet. It is further proposed that such structures are encoded and retained as chunks, along with their associated alphabets and rules of combination. If this model is correct, then sequences that may be parsimoniously represented according to its rules should be better retained than those where such parsimonious representation is not possible.

An experiment was performed as a test of this hypothesis (Deutsch, 1980, 1981). In this experiment a second factor was also considered. It has been found in studies using strings of verbal materials that we tend to recall such strings in accordance with their temporal grouping (Mueller & Schumann, 1894; McLean & Gregg, 1967; Bower & Winzenz, 1969). This effect can be so powerful as to mask grouping by meaning, and so to obliterate the advantage incurred by such grouping (Bower & Springston, 1970). Analogous results have been obtained using nonverbal materials (Restle, 1972; Handel, 1973; Dowling, 1973). It was therefore expected that temporal grouping would have a strong effect on recall of these sequences also. Grouping in accordance with tonal structure was expected to result in enhanced performance, and grouping in conflict with tonal structure to give rise to performance decrements.

The experiment employed the following paradigm. Musically trained subjects were presented with sequences which they recalled in musical notation. The sequences employed are shown in Fig. 16. It can be seen that those in Fig. 16A each consisted of a higher level subsequence of four elements that acted on a lower level subsequence of

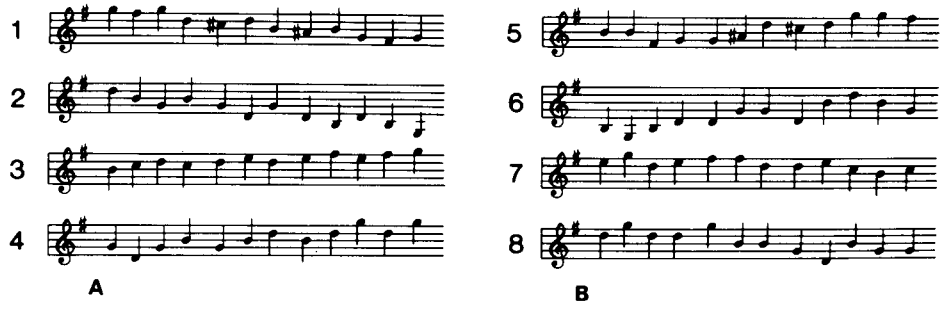



**Fig. 16.** Sequences used in first experiment to study utilization of structure in recall. See text for details (from Deutsch, 1981).

three elements. The sequences in Fig. 16B cannot be represented in this parsimonious fashion. These eight sequences were each presented in three temporal configurations, which are illustrated in Fig. 17. In the first configuration (a), the tones were spaced at equal intervals; in the second (b), they occurred in four groups of three, so that segmentation was in accordance with tonal structure; in the third (c), they occurred in three groups of four so that segmentation was in conflict with tonal structure.

Table I shows the percentages of tones correctly recalled in their correct serial positions in the different conditions of the experiment. Large effects of tonal structure and temporal segmentation are apparent. For structured sequences that were segmented in accordance with tonal structure, the performance level was extremely high. For structured sequences that were unsegmented, the performance level was slightly lower, though still very high. However, for structured sequences that were segmented in conflict with tonal structure, the performance level was much lower. For unstructured sequences the performance level was considerably lower than for structured sequences that were segmented in accordance with tonal structure or that were

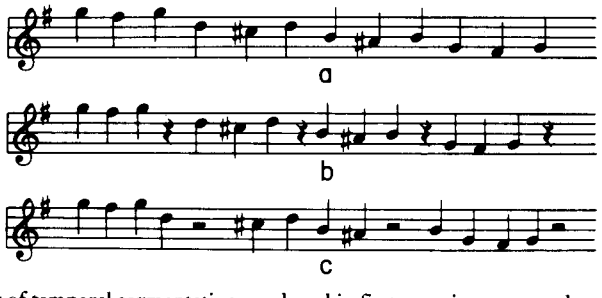






**Fig. 17.** Types of temporal segmentation employed in first experiment to study utilization of structure in recall. a. Sequence unsegmented. b. Sequence segmented in groups of three, so that segmentation was in accordance with structure. c. Sequence segmented in groups of four, so that segmentation was in conflict with structure.

TABLE I

Utilization of Structure in Recall: Percent Correct Recall of Tones in Correct Serial Positions in Experiment 1[a]

| Condition | Percent correct |
|---|---|
| Sequences structured in groups of 3 | |
| OS. Not temporally segmented | 93.5 |
| 3S. Temporally segmented in groups of 3 | 99.3 |
| 4S. Temporally segmented in groups of 4 | 69.2 |
| Sequences unstructured | |
| OU. Not temporally segmented | 52.0 |
| 3U. Temporally segmented in groups of 3 | 63.2 |
| 4U. Temporally segmented in groups of 4 | 62.3 |

[a]From Deutsch (1981).

unsegmented, but in the same range as for structured sequences that were segmented in conflict with tonal structure.

The serial position curves for the different conditions of the experiment are shown on Fig. 18. Typical bow-shaped curves are apparent, and in addition, discontinuities can be seen at the boundaries between temporal groups. This provides further evidence that temporal groups tend to be encoded as chunks and to be retained or lost independently. This type of configuration is very similar to that obtained by others with the use of verbal materials (Bower & Winzenz, 1969).

The transition shift probability (TSP) provides a further measure of interitem association. This is defined as the joint probability of either an error following a correct response on the previous item or of a correct response following an error on the previous item (Bower & Springston, 1970). If groups of elements tend to be retained or lost as chunks, we should expect the TSP values to be smaller for transitions within a chunk and to be larger for the transition into the first element of a chunk. The TSP values for sequences segmented in temporal groups of three and four are shown in Fig. 19A and B, respectively. It can be seen that the TSPs are larger on the first element of each temporal group than on the other elements. This is as expected on the hypothesis that temporal groups serve to define subjective chunks that are retained or lost independently of each other.

Finally, a very strong sensitivity to musical alphabet was demonstrated in this experiment. As shown in Fig. 16, four sequences employed only a triadic alphabet, two employed a diatonic alphabet, and two employed other tones from the 12-tone chromatic scale. Of the 12 subjects who participated in the experiment, 6 remained entirely within the alphabet of the sequence they were notating. (For example, in notating sequences 2, 4, 6, and 8, all their incorrect responses, as well as their correct ones, were G, B, or D.) Of the remaining six subjects, five made between them a total of only 15 responses that departed from the alphabet presented. This implies that the subjects were retaining information concerning alphabet independently of structure.

Now in this experiment, the structured sequences all consisted of a higher level

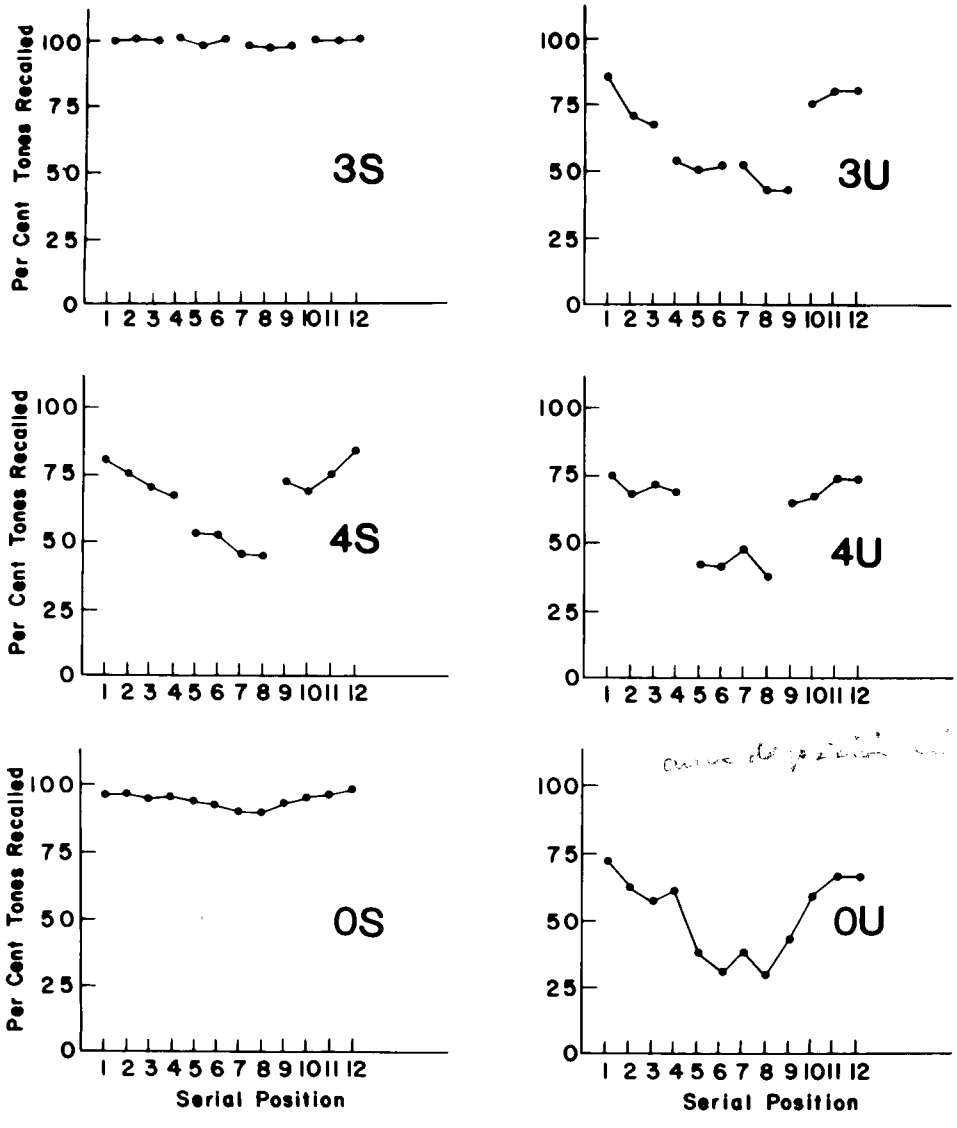



**Fig. 18.** Serial position curves for the different conditions of the first experiment to study utilization of structure in recall (from Deutsch, 1981).

subsequence of four elements that acted on a lower level subsequence of three elements. So compatible segmentation was always in groups of three and incompatible segmentation in groups of four. One might therefore argue that the enhanced performance found for structured sequences was due simply to an advantage conferred by the size of temporal group. A second experiment was designed to control for this.

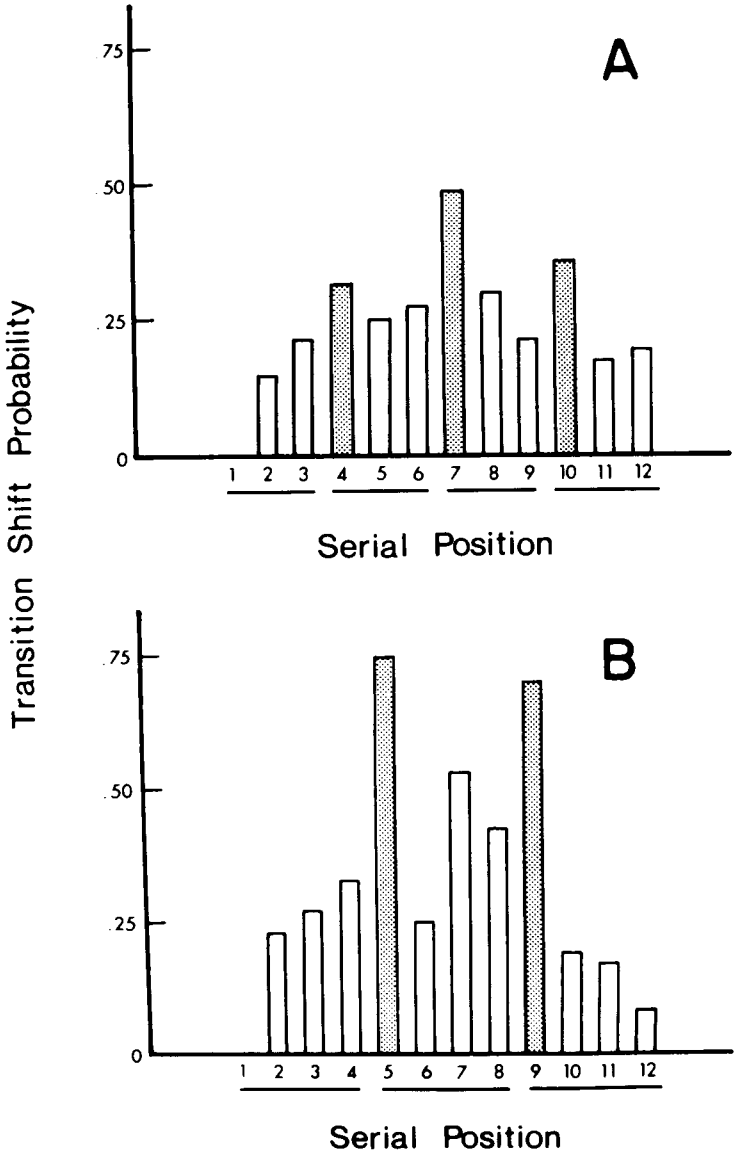



**Fig. 19.** (A) Transition shift probabilities for sequences segmented in temporal groups of three. (B) Transition shift probabilities for sequences segmented in temporal groups of four (from Deutsch, 1981).

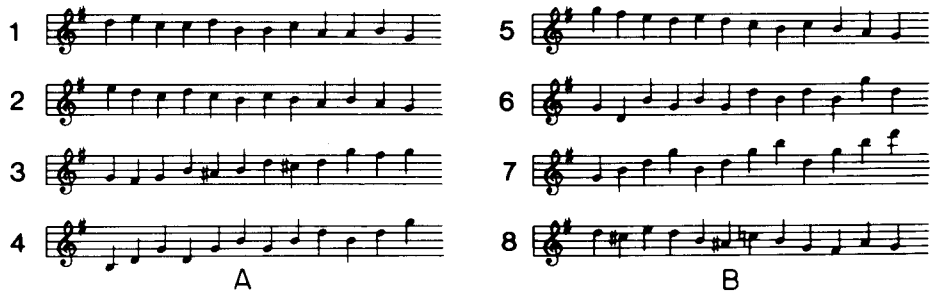



**Fig. 20.** Sequences used in second experiment to study utilization of structure in recall. See text for details (from Deutsch, 1981).

Two types of tonal structure were employed, and these are shown in Fig. 20. In the first, shown in Fig. 20A, a higher level subsequence of four elements acted on a lower level subsequence of three elements. In the second, shown in Fig. 20B, a higher level subsequence of three elements acted on a lower level subsequence of four elements. Enhanced performance was expected where the number of elements in the lower level subsequence corresponded to the number within a temporal group, compared with sequences where these numbers did not correspond. In addition, segmentation in groups of two was examined. It was expected that where the lower level subsequence consisted of three elements, such segmentation would result in considerable performance decrements since it would conflict with tonal structure. However, where the lower level subsequence consisted of four elements, segmentation in group of two would be less disruptive since pauses would still be placed between tonal groups.

There were, therefore, six conditions in the experiment, shown in Table II. Also shown are the percentages of tones correctly recalled in their correct serial positions in these different conditions. It can be seen that performance levels were considerably lower when temporal segmentation was in conflict with structure. When pauses were

TABLE II

Utilization of Structure in Recall:
Percent Correct Recall of Tones in Correct Serial Positions in Experiment 2[a]

| Condition | Percent correct |
|---|---|
| Sequences structured in groups of 3 | |
| 3-2. Temporally segmented in groups of 2 | 45.4 |
| 3-3. Temporally segmented in groups of 3 | 93.1 |
| 3-4. Temporally segmented in groups of 4 | 50.6 |
| Sequences structured in groups of 4 | |
| 4-2. Temporally segmented in groups of 2 | 80.8 |
| 4-3. Temporally segmented in groups of 3 | 52.9 |
| 4-4. Temporally segmented in groups of 4 | 85.4 |

[a]From Deutsch (1981).

Biblioteca del Real Conservatorio Superior de Música - MADRID -

placed both between and within tonal groups (that is Condition 4-2), the performance level was slightly lower than when the pauses were placed only between groups, but considerably higher than when the pauses conflicted with tonal structure. The interaction between size of tonal unit and size of temporal unit was highly significant, reflecting the damaging effect of incompatible segmentation.

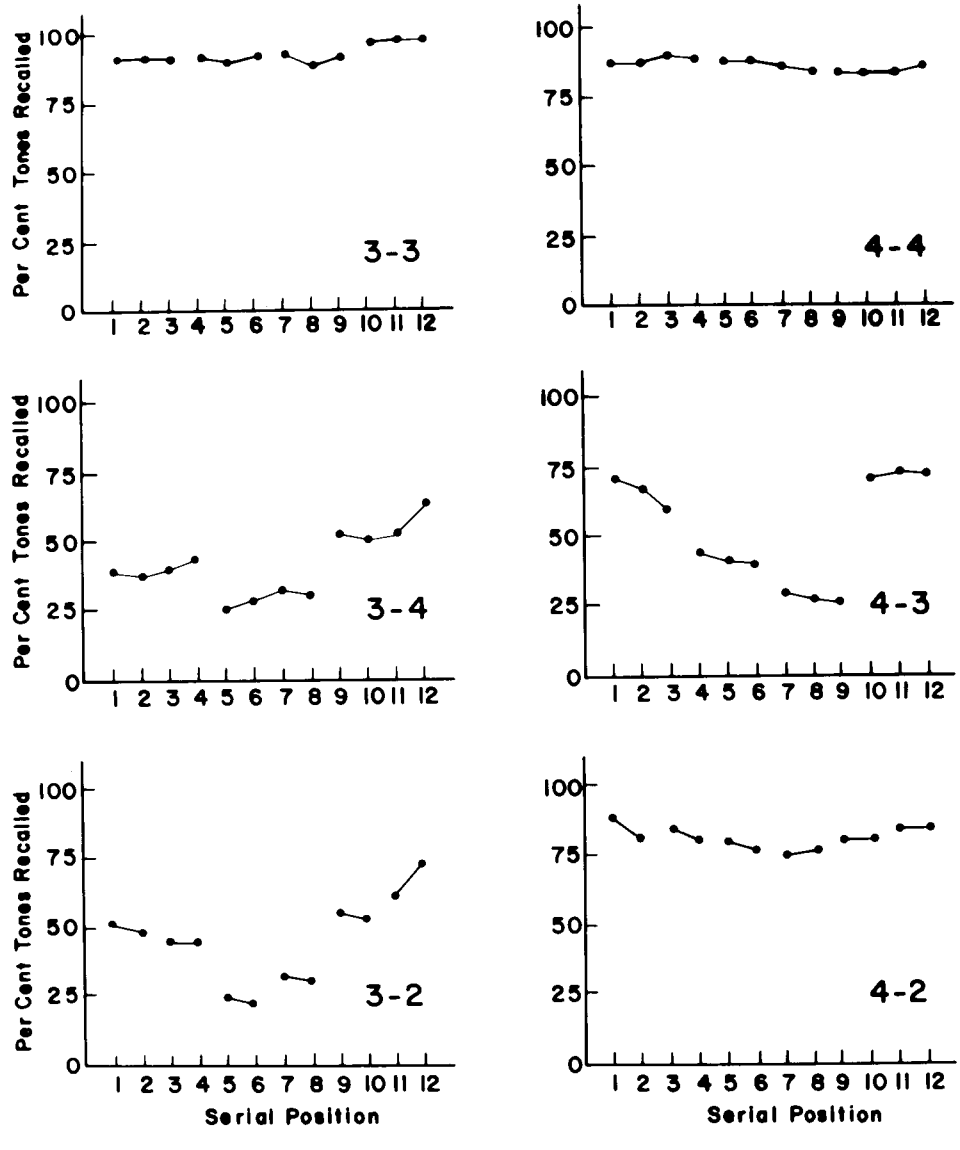



**Fig. 21.** Serial position curves for the different conditions of the second experiment to study utilization of structure in recall (from Deutsch, 1981).

Figure 21 displays the percentages of tones correctly recalled at each serial position in the different conditions of the experiment. It can be seen that again, discontinuities appear at temporal group boundaries, reflecting the formation of subjective chunks on the basis of temporal proximity.

These two experiments lead to several conclusions. First, they demonstrate that listeners perceive hierarchical structures that are present in tonal sequences and can utilize these structures in recall. For the structured sequences employed in this study, the listener need only retain two chunks of three or four items each; but for the unstructured sequences no such parsimonious encoding was possible. The unstructured sequences therefore imposed a much heavier memory load, with resultant performance decrements. Second, the experiments demonstrate that temporal segmentation has a profound effect on perceived structure, as has been shown by others with different materials. On our present line of reasoning, when segmentation is in conflict with structure, there results a less parsimonious representation, which in turn leads to decrements in recall.

For example, Sequence 2 in Fig. 20 would be encoded in the absence of temporal segmentation as

$$A = (3n,*)G$$
$$B = (2n,*)G$$
$$S = A[pr]B,G_4$$

However, with temporal segmentation in groups of four, the same sequence would be encoded as

$$A = (2n,*)G_{tr}$$
$$B = (2n,p,*)G$$
$$C = (n,p,n,*)G$$
$$D = (p,2n,*)G$$
$$S = A[pr](B,C,D),G_4$$

Thus, four structures would need to be encoded and retained, together with their alphabets and rules of combination. A much heavier memory load would therefore be imposed.

## VI. CONCLUSION

In the foregoing pages, we have considered the rules whereby abstractions based on pitch are formed and also how pitch information is retained at the different levels of abstraction. Where appropriate, we have considered underlying neurophysiological mechanisms, and we have also attempted to draw on insights provided by music theorists. The system that we are dealing with is clearly very complex, but an understanding of its operation is slowly developing.[6]

[6]This work was supported by United States Public Health Service Grant MH-21001.

## REFERENCES


Attneave, F., & Olson, R. K. Pitch as a medium: A new approach to psychophysical scaling. *American Journal of Psychology*, 1971, *84*, 147–165.

Babbitt, M. Twelve-tone invariants as compositional determinants. *The Musical Quarterly*, 1960, *46*, 246–259.

Babbitt, M. The structure and function of musical theory. *College Music Symposium 5*, 1965, 10–21.

Bachem, A. Note on Neu's review of the literature on absolute pitch. *Psychological Bulletin*, 1948, *45*, 161–162.

Bachem, A. Time factors in relative and absolute pitch determination. *Journal of the Acoustical Society of America*, 1954, *26*, 751–753.

Baird, J. W. Memory for absolute pitch; Studies in psychology. In *Titchener commemorative volume*. Worcester, 1917. Pp. 69.

Benjamin, W. E. Review of Forte's *The structure of atonal music*. *Perspectives of New Music*, 1974, *2*, 170–211.

Blackwell, H. R., & Schlosberg, H. Octave generalization, pitch discrimination, and loudness thresholds in the white rat. *Journal of Experimental Psychology*, 1943, *33*, 407–419.

Boomslitter, P., & Creel, W. The long pattern hypothesis in harmony and hearing. *Journal of Music Theory*, 1961, *5*, No. 2, 2–30.

Bower, G. H., & Springston, F. Pauses as recoding points in letter series. *Journal of Experimental Psychology*, 1970, *83*, 421–430.

Bower, G., & Winzenz, D. Group structure, coding and memory for digit series. *Journal of Experimental Psychology Monographs*, 1969, *80*, No. 2, Part 2, 1–17.

Bregman, A. S. The formation of auditory streams. In J. Requin (Ed.), *Attention and performance*. (Volume VII) Hillsdale, New Jersey: Erlbaum, 1978.

Browne, R. Review of *The structure of atonal music* by A. Forte. *Journal of Music Theory*, 1974, *18*, 390–415.

Carterette, E. C., Friedman, M. P., & Lovell, J. D. Mach bands in hearing. *Journal of the Acoustical Society of America*, 1969, *45*, 986–998.

Carterette, E. C., Friedman, M. P., & Lovell, J. D. Mach bands in auditory perception. In R. Plomp & G. F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing*. Sijthoff: Leiden, 1970.

Chrisman, R. Identification and correlation of pitch-sets. *Journal of Music Theory*, 1971, *15*, 58–83.

Collins, A. M., & Quillian, M. R. How to make a language user. In E. Tulving and W. Donaldson (Eds.). *Organization of memory*. New York: Academic Press, 1972.

Craik, F. I. M., & Lockhart, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 671–684.

Cuddy, L. L., & Cohen, A. J. Recognition of transposed melodic sequences. *Quarterly Journal of Experimental Psychology*, 1976, *28*, 255–270.

Deese, V., & Grindley, G. C. The transposition of visual patterns. *British Journal of Psychology*, 1947, *37*, 152–163.

Deutsch, D. Music recognition. *Psychological Review*, 1969, *76*, 300–307.

Deutsch, D. The deterioration of pitch information in memory. Unpublished doctoral dissertation, University of California at San Diego, 1970. (a)

Deutsch, D. Tones and Numbers: Specificity of interference in short-term memory. *Science*, 1970, *168*, 1604–1605. (b)

Deutsch, D. Octave generalization and tune recognition. *Perception & Psychophysics*, 1972, *11*, 411–412. (a)

Deutsch, D. Mapping of interactions in the pitch memory store. *Science*, 1972, *175*, 1020–1022. (b)

Deutsch, D. Effect of repetition of standard and comparison tones on recognition memory for pitch. *Journal of Experimental Psychology*, 1972, *93*, 156–162. (c)

Deutsch, D. Octave generalization of specific interference effects in memory for tonal pitch. *Perception & Psychophysics*, 1973, *13*, 271–175. (a)

Deutsch, D. Interference in memory between tones adjacent in the musical scale. *Journal of Experimental Psychology*, 1973, *100*, 228–231. (b)

Deutsch, D. Generality of interference by tonal stimuli in recognition memory for pitch. *Quarterly Journal of Experimental Psychology*, 1974, *26*, 229–234.

Deutsch, D. Musical illusions. *Scientific American*, 1975, *233*, 92–104. (a)
Deutsch, D. The organization of short-term memory for a single acoustic attribute. In D. Deutsch & J. A. Deutsch (Eds.), *Short term memory*. New York: Academic Press, 1975. Pp. 107–151. (b)
Deutsch, D. Facilitation by repetition in recognition memory for tonal pitch. *Memory & Cognition*, 1975, *3*, 263–266. (c)
Deutsch, D. Auditory memory. *Canadian Journal of Psychology*, 1975, *29*, 87–105. (d)
Deutsch, D. Octave equivalence and the processing of melodic sequences. *Journal of the Acoustical Society of America*, 1976, *60*, S94.
Deutsch, D. Memory and attention in music. In M. Critchley & R. A. Henson (Eds.). *Music and the brain*, London: Heinemann, 1977. Pp. 95–130.
Deutsch, D. Octave generalization and melody identification. *Perception & Psychophysics*, 1978, *23*, 91–92. (a)
Deutsch, D. Interactive effects in memory for harmonic intervals. *Perception & Psychophysics*, 1978, *24*, 7–10. (b)
Deutsch, D. Delayed pitch comparisons and the principle of proximity. *Perception & Psychophysics*, 1978, *23*, 227–230. (c)
Deutsch, D. The psychology of music. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception*. (Volume X) New York: Academic Press, 1978. Pp. 191–218. (d)
Deutsch, D. Octave generalization and the consolidation of melodic information. *Canadian Journal of Psychology*, 1979, *33*, 201–204.
Deutsch, D. Recall of hierarchically structured melodic sequences. *Journal of the Acoustical Society of America*, 1980, *67*, 57.
Deutsch, D. The processing of structured and unstructured tonal sequences. *Perception & Psychophysics*, 1981, *28*, 381–389.
Deutsch, D. Facilitation and misrecognition effects in pitch memory as a function of serial position and temporal proximity. In preparation.
Deutsch, D. Effects of melodic context on pitch recognition judgment. *Perception and Psychophysics*, in press.
Deutsch, D., & Feroe, J. Disinhibition in pitch memory. *Perception & Psychophysics*, 1975, *17*, 320–324.
Deutsch, D., and Feroe, J. The internal representation of pitch sequences in tonal music. *Psychological Review*, 1981, *88* 503–522.
Deutsch, D., & Roll, P. L. Error patterns in delayed pitch comparison as a function of relational context. *Journal of Experimental Psychology*, 1974, *103*, 1027–1034.
Deutsch, J. A., & Deutsch, D. *Physiological psychology*. (2nd Ed.) Homewood: Dorsey, 1973.
Dewar, K. M. Context Effects in Recognition Memory for Tones. Unpublished doctoral dissertation. Queens University, Kingston, Ontario, 1974.
Dewar, K. M., Cuddy, C. L., & Mewhort, D. J. K. Recognition memory for single tones with and without context. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, *3*, 60–67.
Divenyi, P. L., & Hirsh, I. J. Some figural properties of auditory patterns. *Journal of the Acoustical Society of America*, 1978, *64*, 1369–1385.
Dowling, W. J. Rhythmic Fission and the Perceptual Organization of Tone Sequences. Unpublished doctoral dissertation, Harvard University, 1967.
Dowling, W. J. Rhythmic groups and subjective chunks in memory for melodies. *Perception & Psychophysics*, 1973, *4*, 37–40.
Dowling, W. J. Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, 1978, *85*, 342–354.
Dowling, W. J., & Fujitani, D. S. Contour, interval and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, 1971, *49*, 524–531.
Dowling, W. J., & Hollombe, A. W. The perception of melodies distorted by splitting into several octaves: Effects of increasing proximity and melodic contour. *Perception & Psychophysics*, 1977, *21*, 60–64.
Drobisch, M. W. *Uber die mathematische Bestimmung der musikalischen*. Intervalle, 1846.
Drobisch, M. W. *Abhandlungen der Sachsischen Wissenchaften math-phys*. Cl.B.II. 1855, 35.
Erickson, R. Paper presented to Department of Music, University of California, San Diego, California, 1978.

Ernst, G. W., & Newell, A. *GPS: A Case Study in Generality and Problem Solving.* New York: Academic Press, 1969.

Evans, E. F. Neural processes for the detection of acoustic patterns and for sound localization. In F. O. Schmitt and F. T. Worden (Eds.), *The neurosciences, third study program.* Cambridge, Massachusetts: MIT Press, 1974. Pp. 131–147.

Forte, A. *The structure of atonal music,* New Haven, Connecticut: Yale University Press, 1973.

Francès, R. *La perception de de musique.* Paris, Vrin, 1958.

Fucks, W. Mathematical analysis of the formal structure of music. *Institute of Radio Engineers Transactions on Information Theory,* 1962, *8,* 225–228.

Gagné, R. M. The acquisition of knowledge. *Psychological Review,* 1962, *67,* 355–365.

Garner, W. R. *The processing of information and structure.* Hillsdale, New Jersey: Erlbaum, 1974.

Greeno, J. G., & Simon, H. A. Processes for sequence production. *Psychological Review,* 1974, *81,* 187–196.

Handel, S. Temporal segmentation of repeating auditory patterns. *Journal of Experimental Psychology,* 1973, *101,* 46–54.

Hanson, A. R., & Riseman, E. M. (Eds.) *Computer vision systems.* New York: Academic Press, 1978.

Harris, J. D. The decline of pitch discrimination with time. *Journal of Experimental Psychology,* 1952, *43,* 96–99.

Helmholtz, H. Von. *On the sensations of tone as a physiological basis for the theory of music.* (2nd English ed.) New York: Dover, 1954. (Originally published 1859).

Helmholtz, H. Von. The Origin of the Correct Interpretations of our Sensory impressions. *Zeitchrift für Psychologie und Physiologie der Sinnesorgane* 1894, 7, 81–96.

Houtgast, T. Psychophysical evidence for lateral inhibition in hearing. *Journal of the Acoustical Society of America,* 1972, *51,* 1885–1894.

Howe, H. S. Some combinatorial properties of pitch structures. *Journal of Music Theory,* 1965, *4,* 45–61.

Howe, H. S. Review of Forte's *The structure of atonal music. Collectanea,* 1974, 118–124.

Hubel, D. H., & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology,* 1962, *160,* 106–154.

Humphreys, L. F. Generalization as a function of method of reinforcement. *Journal of Experimental Psychology,* 1939, *25,* 361–372.

Idson, W. L., & Massaro, D. W. A bidimensional model of pitch in the recognition of melodies. *Perception & Psychophysics,* 1978, *24,* 551–565.

Jeffries, T. B. Relationship of interval frequency count to ratings of melodic intervals. *Journal of Experimental Psychology,* 1974, *102,* 903–905.

Jones, M. R. Cognitive representations of serial patterns. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition.* Hillsdale, New Jersey: Erlbaum, 1974.

Jones, M. R. Auditory patterns: Studies in the perception of structure. In E. C. Carterette and M. P. Friedman (Eds.), *Handbook of perception.* (Vol. VIII, Perceptual coding) New York: Academic Press, 1978.

Julesz, B., & Hirsh, I. J. Visual and auditory perception—An essay of comparison. In E. E. David and P. B. Denes (Eds.), *Human communication: A unified view.* New York: McGraw-Hill, 1972.

Kallman, H. J., & Massaro, D. W. Tone chroma is functional in melody recognition. *Perception & Psychophysics,* 1979, *26,* 32–36.

Klinke, R., Boerger, G., & Gruber, J. Alteration of afferent, tone-evoked activity of neurons of the cochlear nucleus following acoustic stimulation of the contralateral ear. *Journal of the Acoustical Society of America,* 1969, *45,* 788–789.

Klinke, R., Boerger, G., & Gruber, J. The influence of the frequency relation in dichotic stimulation upon the cochlear nucleus activity. In R. Plomp & G. F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing.* Sijthoff: Leiden, 1970.

Koester, T. The time error in pitch and loudness discrimination as a function of time interval and stimulus level. *Archives of Psychology,* 1945, *297,* whole issue.

Koffka, K. *Principles of Gestalt psychology.* New York, Harcourt, 1935.

Krumhansl, C. L. The psychological representation of musical pitch in a tonal context. *Cognitive Psychology,* 1979, *11,* 346–374.

Leewenberg, E. L. A perceptual coding language for visual and auditory patterns. *American Journal of Psychology*, 1971, *84*, 307–349.

Lerdahl, F., & Jockendorff, R. Toward a formal theory of music. *Journal of Music Theory* 1977, *21*, No. 2, 111–172.

Lewin, D. The intervallic content of a collection of notes. *Journal of Music Theory*, 1960, *4*, 98–101.

Lewin, D. A theory of segmental association in twelve-tone music. *Perspectives of New Music*, 1962, *1*, 89–116.

Longuet-Higgins, H. C. Letter to a musical friend. *Music Review*, 1962, *23*, 244–248. (a)

Longuet-Higgins, H. C. Second letter to a musical friend. *Music Review*, 1962, *23*, 271–280. (b)

Longuet-Higgins, H. C. The perception of music. *Interdisciplinary Science Reviews*, 1978, *3*, 148–156.

Mach, E. *The analysis of sensations and the relation of the physical to the psychical.* (C. M. Williams, translator; W. Waterlow, review and supplement) New York: Dover, 1959. (Originally published in German, 1906)

McLean, R. S., and Gregg, L. W. Effects of induced chunking on temporal aspects of serial retention. *Journal of Experimental Psychology*, 1967, *74*, 455–459.

Merriam, A. P. *The anthropology of music.* Evanston, Illinois: Northwestern University Press, 1964.

Meyer, L. B. *Emotion and meaning in music.* Chicago, Illinois: University of Chicago Press, 1956.

Meyer, L. B. *Explaining music: Essays and explorations.* Berkeley, California: University of California Press, 1973.

Meyer, M. On the attributes of the sensations. *Psychological Review*, 1904, *11*, 83–103.

Meyer, M. Review of G. Révész, "Zur Grundleguncy der Tonpsychologie." *Psychological Bulletin*, 1914, *11*, 349–352.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, *63*, 81–97.

Miller, G. A., & Chomsky, N. Finitary models of language users. *Handbook of Mathematical Psychology*, 1963, *2*, 419–493.

Miller, G. A., Galanter, E. H., & Pribram, K. H. *Plans and the structure of behavior.* New York: Holt, 1960.

Mueller, G. E., & Schumann, F. Experimentelle Beitrage zur Untersuchung des Gedächtnisses. *Zeitschrift fur Psychologie und Physiologie der Sinnesorgane*, 1894, *6*, 81–190, 257–339.

Narmour, E. *Beyond Schenkerism.* Chicago, Illinois: University of Chicago Press, 1977.

Navon, D. Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 1977, *9*, 353–383.

Nettl, B. *Music in primitive culture.* Cambridge, Massachusetts: Harvard University Press, 1956.

Olson, R. K., & Hanson, V. Interference effects in tone memory. *Memory & Cognition*, 1977, *5*, 32–40.

Ortmann, O. On the melodic relativity of tones. *Psychological Monographs*, 1926, *35*, Whole No. 162.

Palmer, S. E. The effects of contextual scenes on the identification of objects. *Memory and Cognition*, 1975, *3*, 519–526.

Perle, G. *Serial composition and atonality.* (3rd ed.) Berkeley, California: University of California Press, 1972.

Perle, G. *Twelve-tone tonality.* Berkeley, California: University of California Press, 1977.

Piston, W. *Harmony.* (2nd ed.) London: Norton, 1948.

Pitts, W., & McCulloch, W. S. How we know universals. The perception of auditory and visual forms. *Bulletin of Mathematical Biophysics*, 1947, *9*, 127–147.

Plomp, R., Wagenaar, W. A., & Mimpen, A. M. Musical interval recognition with simultaneous tones, *Acustica*, 1973, *29*, 101–109.

Ratliff, F. *Mach bands: Quantitative studies of neural networks in the retina.* San Francisco, California: Holden Day, 1965.

Restle, F. Theory of serial pattern learning: Structural trees. *Psychological Review*, 1970, 77, 481–495.

Restle, F. Serial patterns: The role of phrasing. *Journal of Experimental Psychology*, 1972, *92*, 385–390.

Restle, F., & Brown, E. Organization of serial pattern learning. In G. H. Bower (Ed.), *The psychology of learning and motivation.* (Volume 4) New York: Academic Press, 1970. Pp. 249–331.

Révész, G. *Zur grundleguncy der tonpsychologie.* Leipzig: Feit, 1913.

Ruckmick, C. A. A new classification of tonal qualities. *Psychological Review*, 1929, *36*, 172–180.

Salzer, F. *Structural hearing.* New York: Dover, 1962.

Biblioteca del Real Conservatorio Superior de Música

Scandura, J. M. Role of rules in behavior: Toward an operational definition of what (rule) is learned. *Psychological Review*, 1970, 77, 516–533.

Schenker, H. *Neue musikalische theorien und phantasien: Der Freie Satz*. Vienna: Universal Edition, 1956.

Schenker, H. In O. Jonas (Ed. & Annotating), *Harmony*. (E. M. Borgese, Translator). Cambridge, Massachusetts: MIT Press, 1973.

Schoenberg, A. *Style and idea*. London: Williams & Norgate, 1951.

Schoenberg, A. In G. Strong (Ed.), *Fundamentals of musical composition*. New York: St. Martin's Press, 1967.

Shepard, R. N. Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America*, 1964, *36*, 2345–2353.

Simon, H. A. Complexity and the representation of patterned sequences of symbols. *Psychological Review*, 1972, *79*, 369–382.

Simon, H. A., & Kotovsky, K. Human acquisition of concepts for sequential patterns. *Psychological Review*, 1963, *70*, 534–546.

Simon, H. A., & Sumner, R. K. Pattern in music. In B. Kleinmuntz (Ed.), *Formal representation of human judgment*. New York: Wiley, 1968.

Suga, N., O'Neill, W. E., & Manabe, T. Harmonic-sensitive neurons in the auditory cortex of the mustache bat. *Science*, 1979, *203*, 270–274.

Suga, N., & P. H-S. Jen. Disproportionate tonotopic representation for processing CF-FM sonar signals in the mustache bat auditory cortex. *Science*, 1976, *194*, 542–544.

Teitelbaum, R. Intervallic relations in atonal music. *Journal of Music Theory*, 1965, *9*, 72–127.

Van Noorden, L. P. A. S. Temporal coherence in the perception of tone sequences. Unpublished doctoral thesis. Technische Hogeschool, Eindhoven, Holland, 1975.

Vitz, P. C., & Todd, T. C. A model of learning for simple repeating binary patterns. *Journal of Experimental Psychology*, 1967, *75*, 108–117.

Vitz, P. C., & Todd, T. C. A coded element model of the perceptual processing of sequential stimuli. *Psychological Review*, 1969, *76*, 433–449.

Von Ehrenfels, C. *Uber Gestaltqualitäten Vierteljahrschrift fur Wissenschaftliche Philosophie*, 1890, *14*, 249–292.

Warren, R. M., & Warren, R. P. *Helmholtz on perception: Its physiology and development*. New York: Wiley, 1968.

Werner, H. Uber Mikromelodik und Mikroharmonik. *Zeitschrift fur Psychologie*, 1925, *98*, 74–89.

Wever, E. G., & Bray, C. W. The nature of the acoustic response; the relation between sound frequency and frequency of impulses in the auditory nerve. *Journal of Experimental Psychology*, 1930, *13*, 373–387.

White, B. Recognition of distorted melodies. *American Journal of Psychology*, 1960, *73*, 100–107.

Wickelgren, W. A. Consolidation and retroactive interference in short-term recognition memory for pitch. *Journal of Experimental Psychology*, 1966, *72*, 250–259.

Wickelgren, W. A. Associative strength theory of recognition memory for pitch. *Journal of Mathematical Psychology*, 1969, *6*, 13–61.

Winston, P. H. Learning to identify toy block structures. In R. L. Solo (Ed.), *Contemporary issues in cognitive psychology: the Loyola Symposium*. Washington, D. C.: Winston, 1973.

Yeston, M. (Ed.) *Readings in Schenker analysis and other approaches*. New Haven, Connecticut: Yale University Press, 1977.

Yngve, V. A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society*, 1960, *104*, 444–466.

# 10

# Melodic Processes and the Perception of Music

*Burton S. Rosner and Leonard B. Meyer*

I. The Perception and Classification of Two Archetypal Melodic Processes ..... 317
A. Gap-Fill Melodies ..... 323
B. Changing-Note Melodies ..... 325
II. Experimental Findings ..... 326
A. Purpose ..... 326
B. Method ..... 327
C. Results ..... 330
D. Discussion ..... 335
III. Implications ..... 339
References ..... 340

## I. THE PERCEPTION AND CLASSIFICATION OF TWO ARCHETYPAL MELODIC PROCESSES

It is a familiar fact that we tend to understand and remember experience in terms of categories and classes. And this is as true of aesthetic experience as of other realms of human activity. In literature, for instance, narrative plots—that is, goal-directed *processes* implying probable outcomes—seem to form readily recognizable types: "boy-meets-girl" stories, "who-done-it" mysteries, "revenge" dramas, and so on. Such narrative processes are often complemented by commonly replicated *formal* plans—for example cyclic schemes, episodic successions, flashback structures, and the like. When they are considered to have transcended the confines of a particular cultural group or historical period, replicated kinds of processes and forms have frequently been referred to as *archetypes* (see, for instance, Frye, 1957, pp. 99–105 and *passim*)—a term we adopt here.


Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

Archetypes are important for a number of reasons. They establish fundamental frameworks in terms of which culturally competent audiences (not only members of the general public but also creators, performers, critics, and scholars) perceive, comprehend, and respond to works of art. For what audiences enjoy and appreciate are neither the successions of stimuli *per se*, nor general principles *per se*, but the relationship between them as actualized in a specific work of art. Just as we can delight in the play of a particular football game only if we understand the constraints governing the action down on the field (the rules, strategies, physical conditions, etc.), so we can enjoy and appreciate the playful ingenuity and expressive power of works of art only if we know—and such knowledge may be tacit: a matter of ingrained habits and dispositions—the constraints that governed the choices made by the artist and, consequently, shaped the process and form of the particular work of art. Thus, archetypes are, in a sense, embodiments of fundamental stylistic constraints. As such, they connect understanding to other aspects of aesthetic experience.

For our appreciation and evaluation of a work of art involves not only understanding what is actually presented, but also what *might have been* presented given the constraints of the style. Put differently, a full appreciation of the significance of an actual patterning includes an awareness of possible alternatives that remain unrealized—what Eugene Narmour has called "implied structure" (Narmour, 1977, p. 212). For the sake of brevity, consider an example discussed at some length elsewhere (Meyer, 1979, pp. 33–38). The slow movement of Haydn's *Military* Symphony (No. 100) opens with a two-measure pattern that can without the slightest exaggeration be termed "archetypal" in the music of the eighteenth century (Fig. 1, first part). But the pattern belongs not to the class of *beginning* figures but to that of *closing* ones—as in the second and third parts of Fig. 1. One consequence of this anomalous usage is that the movement might have ended with the very same figure that it began with. (For other instances of movements that begin with a closing figure, see Meyer, 1973, pp. 212–217; and Levy, 1981, pp. 355–362). And, though Haydn did not choose this alternative, the competent listener's appreciation of the close of the movement includes some (perhaps unconscious) awareness that this option was a very real possibility. That this was indeed a possible alternative ending can, in this case, be convincingly documented. For the second movement of the *Military* Symphony is based almost in its entirety upon the slow movement of an earlier work by Haydn, his Concerto in G Major for Two Hurdy Gurdies and Orchestra. And in the Concerto, the movement closes with the same figure that it began with.

Similarly, our evaluation of works of art involves an awareness of the possibilities available to the composer in a specific set of stylistic cultural circumstances. Thus, to praise a work as being "perfect," "inevitable," or "exciting" is implicitly to assert that, given our knowledge of the constraints governing the composer's choices, the alternatives actually used were the most felicitous imaginable; while to deprecate a passage as being "awkward," "trite," or "bombastic" is to suggest that, considering existing constraints, more graceful, original, or convincing alternatives were available to the composer.

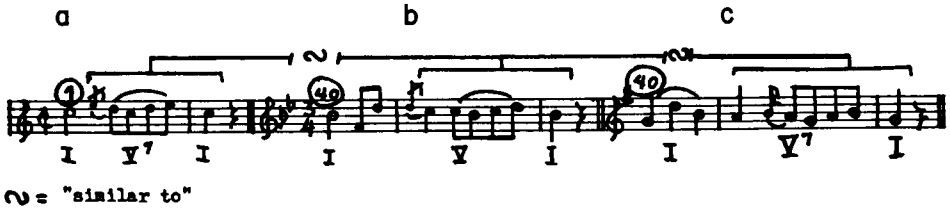



**Fig. 1.** Excerpts from (*a*) Haydn's Symphony no. 100 in G Major (*Military*), ii, mm 1–2; (*b*) Haydn's String Quartet in B flat, op. 64 no. 3, iii, mm 40–42; and (*c*) Mozart's Symphony no. 40 in G Minor (K. 550), iii, mm 40–42. Notice a similar melodic process in all three.

Because they are coherent, orderly, and simple, archetypes are easily learned and tend to be stable over time. For these reasons, they not only facilitate the appreciation and evaluation of particular works of art, but they also constitute an important means of cultural continuity. That is, they function as cognitive-mnemonic schemata—the replicated entities that Dawkins (1976) called "memes" because he considered them the cultural counterparts of the units of biological trait transmission. As such they may be an important basis for studying the histories of style of art.

As suggested earlier, archetypes may play a significant role in shaping aesthetic experience and fostering cultural continuity in the absence of any conscious conceptualization about their existence, nature, or kinds. Rather, they may be and usually are internalized as habits of perception and cognition operating within a set of cultural constraints. One of us, who has known the opera for many years, only recently recognized that the plot of Mozart's *Magic Flute* belongs to the archetypal class of quest-trial stories of which Dante's *Divine Comedy*, Bunyan's *Pilgrim's Progress*, and Cervantes' *Don Quixote* are celebrated members. And, as will soon be apparent, very familiar melodies may be members of archetypal classes without our being conscious of the fact.

Music, too, tends to be understood and remembered in terms of types and classes, some of which can, as Fig. 1 indicates, be thought of as archetypal. In the domain of musical form, this is abundantly clear. For though not usually so characterized, many of the forms described and labeled in books about musical analysis can be considered archetypal: for instance, strophic patterns (AA′A″ . . . $A^n$), exemplified in the theme and variation form; binary patterns (AB), the basis for many dance movements; bar forms (AA′B), the structural basis for countless melodies; ternary patterns (ABA), exemplified in *da capo* arias, many dance movements, and marches. Though formal organization is chiefly delineated by the syntactic processes of the primary parameters of music (melody, rhythm, and harmony), form is usually made particularly patent through the patterning of the secondary parameters (i.e., dynamics, tempi, register, instrumentation, sonority, and so on). The prevelance of such formal organizations has been amply documented both by theorists and historians of Western music and by ethnomusicologists.

The analysis and classification of the *processes* that complement such formal schemes

have proved more problematic. As textbooks on form and analysis show, most of the achievements have been in the parameter of harmony. A number of the processes typical of Western tonal music have been identified: for example, the progression through the circle of fifths and characteristic cadences—authentic, plagal, and deceptive. On a more extended hierarchic level, an archetypal harmonic-contrapuntal patterning, the so-called *Ursatz*, has been posited in the theory of Heinrich Schenker (Schenker, 1956; Yeston, 1977). In the areas of rhythm and melody, there has been less success. Though there has been increasing interest in rhythm and meter in music and some typical groupings and patterns have been tentatively distinguished (Cooper & Meyer, 1960), little agreement has been reached (Komar, 1968; Yeston, 1976).

The difficulties encountered in the analysis of melodic processes, the concern of this study, result from a combination of conceptual, methodological and systemic problems. In the realm of human behavior, we tend to conceptualize and classify processes in terms of goals—and of the strategies devised for reaching them. Thus, reading a novel or seeing a play, we understand the succession of represented events in terms of the purposes of the protagonists: the union of boy and girl—despite her father's objections; the discovery of the criminal—despite his best efforts to avoid detection; the realization of revenge—despite obstacles of wealth or power. Because music is not well-suited to, or essentially concerned with, the representation of specific human actions, successive events in a composition cannot readily be related to one another in terms of such goals. Partly for this reason, melodic processes have resisted theoretical formulation and analysis.

A second difficulty is that while the formal plans, mentioned above, seem to be cross-cultural and atemporal, the syntactic processes generated by the primary parameters of melody, rhythm, and harmony are not. They are conventional: bound to a particular cultural context and limited to a specific historical epoch. The constraints governing melodic relationships and processes are not the same in Mozart's music as in Machaut's; nor are they the same in Western tonal music as in the music of Java. Put the other way around: it seems probable that we have a reasonably workable account of harmonic process precisely because what we have is *not* a fully general theory but one restricted to the practice of a particular culture and period—that of Western culture since the end of the Renaissance. If this observation has merit, then the search for "universal" principles may hinder the development of theories (*plural*) of melodic process.

The third difficulty, and perhaps the most important, is systemic. It arises because melodies are frequently hierarchically structured. As a result, the kind of patterning exhibited by a particular melody tends to change from one hierarchic level to the next; and so, of course, does the class of archetype involved. Consider, for instance, the melody of the opening measures of Mozart's Oboe Quartet (Fig. 2). As graph *a* of Fig. 2 shows, the foreground (note-to-note) level is characterized by rising and falling linear (scale) patterns. On the next level (graph *b*), rising and falling thirds create pairs of complementary diads. An archetypal changing-note pattern (F-G-E-F) arises on the third hierarchic level (graph *c*). And on the highest level, that resulting from the relationship created by phrase beginnings, the patterning is again linear—the first

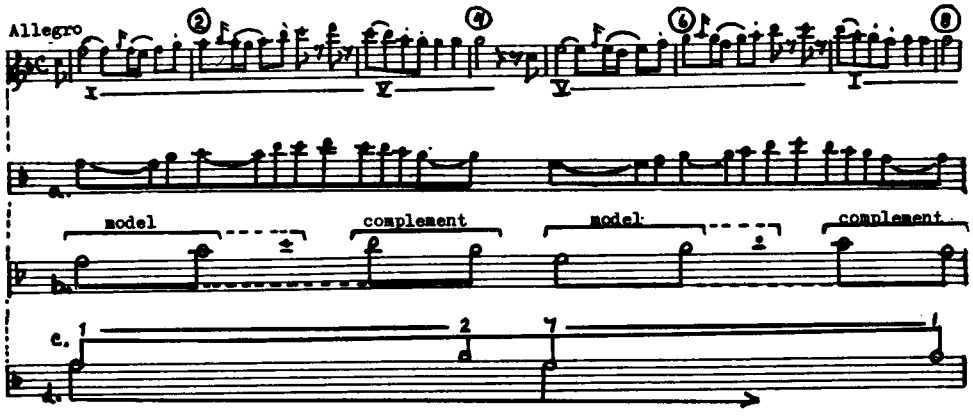



**Fig. 2.** Mozart's Oboe Quartet in F Major (K. 370), i. mm 1–8. Graphs a, b, c, and d show analysis at increasingly higher levels of hierarchy. See text for explanation.

notes of a descending F-major scale, which continues beyond the music given in this figure.

If the kind of patterning changes from one hierarchic level to another within the same melody, which of the possible patterns should be the basis for classification? For instance, should the melody given in Fig. 2 be classed as linear on the basis of its foreground organization (graph *a*)? As two sets of complementary diads (graph *b*)? As a changing-note melody (graph *c*)? Or on the highest level, as again linear (graph *d*)?

The view adopted here is that melodic patterns are classified by listeners, as well as music theorists, in terms of the organization of the highest level on which significant closure is created by the parameters that shape musical relationships. Thus, Fig. 2 would be classed as a changing-note melody because that is the organization of the highest structural level when pattern closure occurs in measure 8. Had the closure that defines the limits of the whole melody (which lasts for 20 measures) been the basis for typology, then the pattern would have been classed as linear (graph *d*). For analysis of the whole melody, see Meyer (1973, pp. 192–195). In like manner, though the first half (AA′) of Fig. 3 consists of disjunct, triadic motion and the second (B)

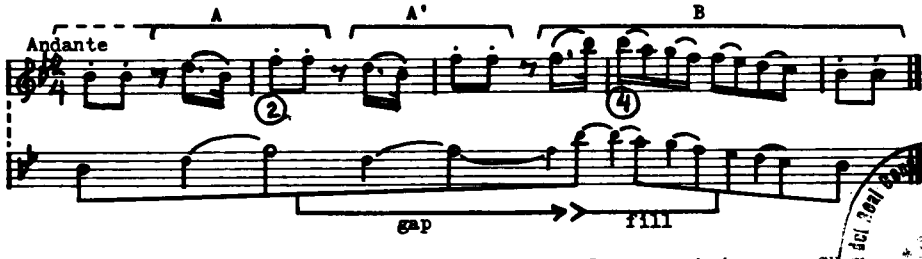



**Fig. 3.** Mozart's Symphony in F Major (K. 112), ii, mm 1–4. Lower graph shows gap-fill process.

Biblioteca del Real Conservatorio Superior de Música - MADRID

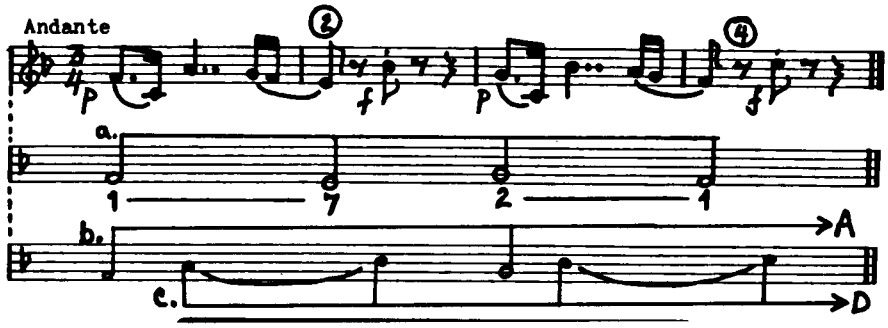



**Fig. 4.** Mozart's Symphony no. 41 in C Major (K. 551), ii, mm 1–4. Lower graphs show analysis of alternative melodic structures. See text for explanation.

involves conjunct linear descent, the patterning of the highest level of closure creates what will be called a gap-fill melody.

Ease or difficulty of classification of melodic processes may be related to another sort of melodic complexity. Namely, the possibility that a single melody may contain two different strands of patterning on the *same* hierarchic level. For instance, the main pattern of the melody of the slow movement of Mozart's *Jupiter* Symphony is a changing-note figure (Fig. 4, graph *a*). But coordinate with this pattern is a subsidiary linear organization that is presented on weakbeats (graph *c*). To classify this melody, the listener must distinguish the two strands and discern which is the main one. As graph *b* indicates, the next level of this strand transcends the melodic and harmonic closure reached in measure 4. It is linear and implies continuation to a structural A (not shown in Fig. 4). Again, the basis for classification is the degree of closure. In measure 4 both the F (graph *a*) and the C (graph *c*) occur over stable, closed tonic harmony. But the F is more closed, both because it occurs on a stable, accented beat while the C comes on a mobile weakbeat, and because F is the stable tonic of the scale while C is the less stable fifth of the scale. As a result, the more closed changing-note pattern acts as the basis for melodic classification.

Of the primary parameters of music, melody has unquestionably been the most difficult to analyze and classify in terms of process. The paucity of practical and theoretical texts on the subject are evidence of this. So is the generally antitheoretical attitude expressed in the familiar maxim that since melody is a matter of native gift and inspiration, melodic processes cannot be "reduced" to principles or "forced" into artificial classes, and, consequently, melodic writing cannot be taught. Nevertheless, one of us (Meyer, 1973) has attempted to show that in the tonal music of Western culture, melodic processes can be separated into distinguishable classes on a particular hierarchic level—usually the level on which the pattern exhibits clear closure. Since two of these classes—gap-fill melodies and changing-note melodies—were investigated in an experiment described below, the characteristics of each class merit brief discussion.

## A. Gap-Fill Melodies

Gap-fill melodies consist of two elements: (1) a disjunct interval or a succession of disjunct intervals moving in the same direction—the *gap* (or skip) and (2) a succession of conjunct (stepwise) intervals that "fill" the gap by presenting all or most of the notes previously skipped over. The relationship is one in which the incompleteness created by a gap implies subsequent completeness through a fill; and, generally speaking, the larger the gap, the stronger the sense of incompleteness and the implication of fill.

The subject of the fugue from Geminiani's Concerto Grosso in E Minor, Op. 3 No. 3, is an uncomplicated instance of a gap-fill melody (Fig. 5). As the graph (*a*) beneath the figure shows, the main gap consists of the octave from a low to a high E, and a subsidiary gap occurs from B up to E. This gap structure is immediately followed by conjunct motion—at first chromatic and then diatonic—that descends through the harmonic minor scale down to the tonic (E).

The same fundamental process, elaborated somewhat, forms the basis for the beginning of the chorus of the very familiar tune, "Over the Rainbow," by Harold Arlen (Fig. 6). Observe that whereas Geminiani's fugue subject has essentially only one skip (gap), followed by a fill that is slightly embellished by passing chromaticism, the generating gap of "Over the Rainbow" is reinforced by "auxiliary" gaps in measures 3 and 5; and most of the tones of the fill are prolonged through melodic-rhythmic elaboration. These differences affect formal structure. For instead of being a basically one-part form, as Geminiani's fugue subject is, "Over the Rainbow" is a bar form (AA'B), as indicated over Fig. 5. However, because the initial octave gap, the prime generating event, is immediately followed by the beginning of the fill, one thinks of the gap as being much shorter than the fill. Although most gaps are upward skips, downward skips may also function as gaps (see Narmour, 1977, pp. 76–77 for an example). And though upward skips are usually followed by descending fills and downward skips by rising fills, this is not necessarily the case: for instance, an ascending gap may be "filled" (completed) by a rising conjunct pattern. Moreover, the fill may be partly descending and partly ascending. For instance, in Fig. 6 the third gap (m. 5) is partly filled by the subsidiary ascending line shown in graph *b*.

But the gap part of a gap-fill process may be prolonged, so that the components of

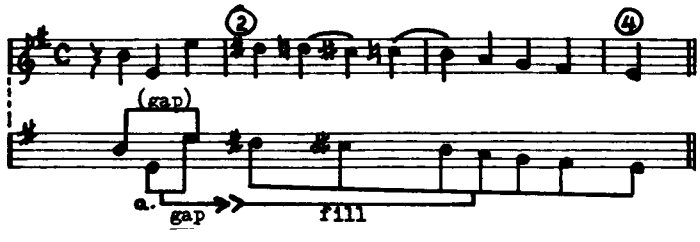



**Fig. 5.** Geminiani's Concerto Grosso in E Minor, op. 3 no. 3, mm 1–4. Lower graph shows gap-fill process.

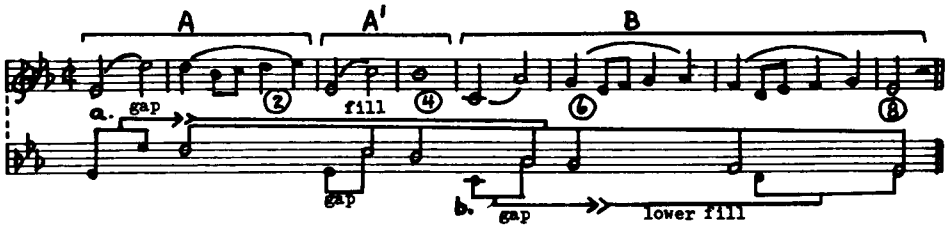



**Fig. 6.** "Over the Rainbow" by Harold Arlen and E. Y. Harburg. Copyright © 1938, 1939 (renewed 1966, 1967) Metro-Goldwyn-Mayer, Inc. Rights throughout the world controlled by Leo Feist, Inc. All rights reserved. Used by permission. Lower graph shows gap-fill processes. AA'B form indicated above. Chorus, mm 1–8.

the whole melodic entity are roughly equal in length. This is true of the opening measures of the slow movement of Mozart's Symphony in F Major (K. 112) (Fig. 3). Thus, although the form (AA'B) is like that of Fig. 6, the first half of Mozart's melody is concerned with creating the gap; the fill takes place only in the second half. It should also be noted that, as in many cases, the pitches that define the full extent of the gap (the octave from low to high B♭) are not directly connected but are the end points of a triad.

Gap-fill patterns may also occur in conjunction with the quasi-strophic form AA', called an antecedent-consequent period. The beginning of the Menuetto of Mozart's Flute Quartet in A Major (K. 298), illustrates this kind of process/form coupling (Fig. 7). In the antecedent phrase, a brief, anacrustic gap moves rapidly through the tonic triad from a lower to an upper fifth. The following fill descends conjunctly from the high A. However, before it is completed, a half cadence (IV–V) creates partial closure on the second degree of the scale, E (m. 4). The consequent phrase begins like the antecedent, reiterating the gap and the beginning of the fill. But the end of the fill is modified so that a full cadence occurs on the tonic D (m. 8). Though the melody never descends to the lower A, the gap is understood as being satisfactorily filled because the tonal/rhythmic closure at the end of the period (m. 8) is so decisive that whatever incompleteness might have been experienced is for the moment eclipsed.

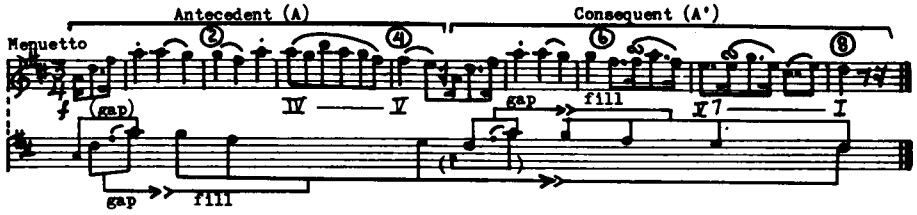



**Fig. 7.** Mozart's Flute Quartet in A Major (K. 298), ii, mm 1–8. Lower graph shows gap-fill processes. Antecedent-consequent form indicated above.

## B. Changing-Note Melodies

A changing-note melody is one in which the main structural tones of the pattern consist of the tonic (1), the seventh or leading tone of the scale (7), the second degree of the scale (2), and then the tonic again. As we have seen (Fig. 2), the order of the middle pitches may be reversed, resulting in the succession 1–2, 7–1. A variant of the changing-note process may occur beginning on the third degree of the scale, producing the succession 3–2, 4–3 (or 3–4, 2–3): an example is discussed in Meyer (1976). The pattern is always harmonized by a progression that moves from a tonic chord (I) to dominant harmony (V), and then from the dominant back to the tonic: that is, the progression is always I–V, V–I. Surrogates for these harmonies are possible: for example, vii for V, or vi for I. The form complementing the changing-note pattern is quasi-strophic, AA′. Thus, the basic plan can be diagramed as:

$$\frac{A}{I - V} : \frac{A'}{V - I}$$

It might be thought that the underlying regularity of a changing-note melody is really harmonic rather than melodic. However, since the same harmonic, and even formal, pattern may accompany other melodic processes—for instance, one that might be called a "sequential changing note pattern" (see Meyer, 1980)—the parameters are at least partially independent.

Any or all of the structural tones in a changing-note pattern may be extended—though ornamentation, triadic prolongation, etc. Consider, for instance, the beginning of the last movement of Beethoven's String Quartet in F Major, Op. 18 No. 1 (Fig. 8). In the first measure the tonic F is prolonged first through a turn involving lower and upper neighbor-notes and then through a linear descent. The leading tone E on the first beat of measure 2 is clearly structural despite the afterbeat G. In measure 3 a structural G, comparable in every way to the earlier F, is also prolonged, and then moves down by octave transfer to the tonic F followed by an afterbeat (A).

The cultural "potency" of this archetype seems evident in the fact that it acts as the basis for "Hinky Dinky, Parlee-voo," a tune that was enormously popular during the

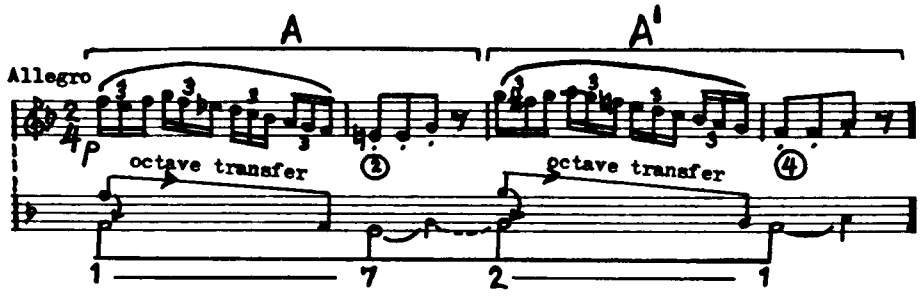



**Fig. 8.** Beethoven's String Quartet, op. 18 no. 1, iv, mm 1–4. Lower graph shows changing-note process and octave transfers. AA′ form indicated above.

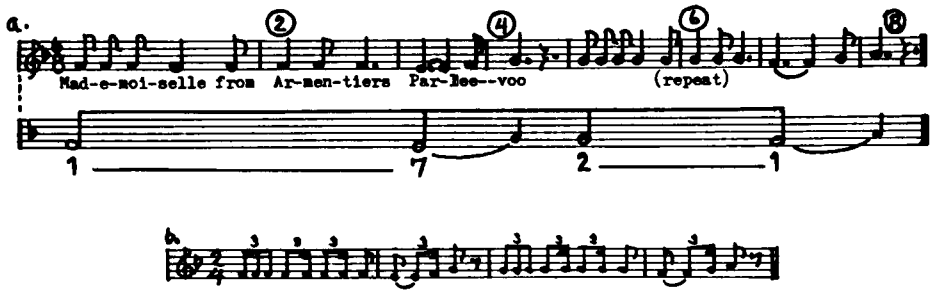



**Fig. 9.** (*a*) "Hinky Dinky, Parlee-voo," folk tune, mm 1–8, with analysis of changing-note process underlying it. (*b*) Notation of meter of (*a*) divided to conform to the meter of Fig. 8.

World War I (Fig. 9*a*). The similarity between the melodic process of this commonplace ditty and that of the last movement of Beethoven's Op. 18 No. 1 is evident enough. If the notation of its durational pattern is changed to conform to that of Fig. 8, then as Fig. 9*b* shows, the metric schemes are similar as well.

## II. EXPERIMENTAL FINDINGS

### A. Purpose

The preceding analysis of melodic processes raises an obvious psychological question. Do the different types of processes identified by musical theory play any role in the perception of music? If they do, a subject untutored in formal musical analysis should at least be able to place melodies into the classes which the analysis specifies. This argument can be tested by a concept identification experiment in which different types of melodic processes define the different classes of stimuli. The processes should be represented at the highest hierarchical level for each stimulus. A subject should be able to learn to associate different responses with the different classes. The subject then should successfully generalize the responses to a new sample of melodies.

We carried out such a concept identification experiment.[1] In this experiment, subjects had to learn to place gap-fill melodies into one class and changing-note melodies into another. The subjects were to respond to gap-fill instances by saying "A" and to changing-note instances by saying "B," without any explicit education in melodic analysis and without any didactic training about these two types of melodic processes. Training was conducted by merely informing the subject after each response what the correct response should have been. We used recordings of fully instrumented passages from the musical literature for this experiment, rather than

[1]The experiment grew out of a graduate seminar which we taught together in 1977. The students participated in the design of the experiment, selection of passages, and running subjects. We are grateful for their unusual help to James Copp, Erika Ellenberger, Christopher Foard, Robert Hopkins, Lief Laudamus, and Katherine Hirsch-Pasek. Scott Balthazar helped us to collect further necessary data and to organize the results.

melodies of our own devising. This choice rested on two grounds. First, we wanted to keep style as homogeneous as possible. We found that we could select sufficient material from works of the Classical period (ca. 1750–1827). Second, we feared that we could all too easily make the concept identification task trivial by composing our own melodies. We also rejected the procedure of abstracting melodies from the musical literature and having them played on a single instrument, such as a clarinet. We wanted to see whether classification of different types of melodic processes could emerge despite the presence of the many other parameters which are characteristic of music. Therefore, we used short passages of four to 15 measures' length chosen from works by Haydn, Mozart, Beethoven, and Schubert. Each passage reached significant closure and was an instance of either gap-fill or changing-note processes.

The results of this initial experiment suggested that subjects could indeed learn to place each melody into its appropriate class and could then generalize to new melodies. Inspection of the data, however, indicated that changing-note melodies generated fewer erroneous responses than did gap-fill ones. Subjects apparently carried out the task by learning to respond correctly first to the changing-note passages and then using the alternative response to all other passages. We could not conclude safely that the two types of melodic processes played more or less equal roles in the perception of music. Therfore, we redesigned the experiment as a two-part procedure. In one part, subjects heard gap-fill melodies along with some which represented a variety of different melodic processes other than gap-fill or changing-note. The subjects had to respond by saying "Type A" to the former and "Not type A" to the latter. A generalization test followed training. In another part of the experiment, subjects learned to respond to changing-note melodies by saying "Type B" and to melodies representing various processes other than changing-note or gap-fill by saying "Not type B." Again, a generalization test followed training. No passage used in one part of the experiment ever appeared in the other.

## B. Method

### *1. Selection of Passages*

Fifty-nine passages were selected for this experiment from the works of Haydn, Mozart, Beethoven, and Schubert. They represented parts of symphonies, chamber music, concertos for solo instrument and orchestra, and sonatas for piano or for piano and violin. Tables III and IV provide lists of the passages selected for the gap-fill and changing-note parts of the experiment, respectively. The tables show composer, work, movement, and measures. On request, we will provide details about the particular recordings used.

Real Conservatorio Superior

### *2. Gap-Fill versus Non-Gap-Fill (Part A)*

Thirty passages were used in this part of the experiment. Eight gap-fill and eight non-gap-fill passages were used for training. Four of each type were of form AA'B

and the other four were antecedent-consequent forms. In an AA'B form, the passage contains an initial motif, a repetition or a variation of it, and then a longer concluding motif (see Fig. 3). The antecedent-consequent form contains two related statements, the first of which does not resolve to the tonic while the second does (e.g., Fig. 7). There were two gap-fill passages of the form AA'B that ended on dominant harmony, and these were matched by a non-gap-fill passage of the same form that also ended on the dominant. In addition to the 16 training stimuli, two more gap-fill passages were chosen for introductory examples at the start of training. One was of form AA'B and the other was an antecedent-consequent passage. Since the non-gap-fill passages employed various melodic processes, no introductory examples of these selections were given. For the generalization test we chose six gap-fill and six non-gap-fill passages different from those used for training. Three gap-fill and three non-gap-fill passages had form AA'B. One of these ended on the dominant, and this was matched by a non-gap-fill passage that ended on the dominant. The other three gap-fill and the other three non-gap-fill passages were antecedent-consequent in form.

*3. Changing-Note versus Non-Changing-Note (Part B)*

Twenty-nine passages were used in this part of the experiment. All had form AA' (see Fig. 8). Eight changing-note and eight non-changing-note passages served as training stimuli while an additional changing-note melody provided an introductory example at the start of training. A non-changing-note introductory example was not offered since passages in this category used a variety of melodic processes. The generalization test employed six changing-note and six non-changing-note passages different from those used in training.

*4. Preparation of Tapes*

Each passage was recorded on a master selection tape, from which it was rerecorded on separate segments of tape as many times as necessary. The segments then were spliced into the experimental tapes. There was a training and a generalization tape for each part (A and B) of the experiment.

For gap-fill training, the subject responded by saying "Type A" to all instances of gap-fill melodies and by saying "Not type A" to all non-gap-fill passages. The tape started with an announcer's voice stating, "This is an experiment in music perception. You will now hear an example of type A." The AA'B gap-fill introductory example then was played once. After a 10-sec pause, the announcer's voice stated, "This is another example of type A" and the antecedent-consequent gap-fill introductory example was then played. The two introductory examples never reappeared in the rest of the experiment. The announcer then said, "Now start responding." After a 2-sec delay, the first training passage was played. Ten sec after its termination, the announcer identified it by saying, "That was type A" or "That was not type A," as appropriate. The subject had to respond within the 10-sec interval between the end of the passage and the announcer's identification of it. Five sec after the identification,

the announcer stated, "Next example," and another passage was begun 2 sec later and followed 10 sec after its conclusion by an identifying statement. The training tape continued, following this cycle of a warning of "Next example" 5 sec after termination of the previous identification, a 2-sec pause after the warning, a passage, a 10-sec pause for the subject to respond, and then identification of the type of selection just heard.

A single training trial consisted of one playing of each of the eight gap-fill passages and each of the eight non-gap-fill ones. The order of passages was randomized with the constraint that no more than three gap-fill or no more than three non-gap-fill selections could occur in sequence. One trial took about eight minutes. A different random order of the passages was used for different trials. The gap-fill training tape contained 12 trials, with no indication to the subject of where one trial ended and the next began. Our initial experiment had indicated that this would permit most subjects to reach a criterion of two successive trials with two errors or less on each.

The gap-fill generalization tape consisted of six gap-fill and six non-gap-fill passages in random order, with the constraint that no more than two of either type could occur in succession. The tape began with the message, "Now start responding." Two seconds later the first generalization passage began. It was repeated after a 2-sec pause. Then a 10-sec pause occurred, during which the subject was to respond. No information about correct identification was given. Instead, the announcer stated, "Next example," and after a 2-sec pause, two exposures to the next passage occurred. Ten seconds were then allowed for responding. The cycle of a warning, a 2-sec pause, two playings of a passage, and a 10-sec pause for responding continued throughout the gap-fill generalization tape. The entire tape ran for about 10 minutes.

The changing-note training tape was constructed exactly parallel to that for the gap-fill part of the experiment, with one minor exception. Since all passages were of form AA′, only one example of a changing-note melody was played at the start of training. Subjects were supposed to respond with "Type B" to changing-note and with "Not type B" to non-changing-note passages. The introductory example on the training tape was preceded by the announcement, "This is an experiment in music perception. You will now hear an example of type B." Identifications after training passages were either, "That was type B" or "That was not type B." The changing-note generalization tape followed exactly the pattern of the gap-fill generalization tape.

### 5. *Subjects*

Twenty subjects were recruited individually by one of the experimenters from students taking music courses at the University of Pennsylvania. They included undergraduate and graduate students. Among the former were majors in music or in other subjects. The graduate students were all studying composition. No subject had received any exposure to analysis of melodic processes, although some were versed in more traditional aspects of musical theory, such as harmony, counterpoint, and musical form. Each subject received $6 at the end of two 1½ hour sessions.

*6. Procedure*

Each subject underwent gap-fill training and generalization (part A) in one session and changing-note training and generalization (part B) in the other. Half the subject took part A first, and the other half took part B first. As they were recruited, subjects were assigned in alternation to one or the other initial condition. At the start of the first session, each subject filled out a questionnaire and a consent form. The questionnaire asked about academic status, previous training in musical performance, previous academic training in music, and current performance activities and listening habits.

The experimenter then read instructions to the subject and answered any questions. The instructions specified that subjects were to respond by saying "Type A" or "Not type A" for part A of the experiment (gap-fill) or "Type B" or "Not type B" for part B (changing-note). They also informed the subject that 10 sec were allowed for responding after the end of a passage and each failure to meet this requirement would result in losing 10 cents. The experimenter then started the appropriate training tape, which was reproduced at a comfortable listening level over an Akai model 400DS Mark II tape deck. The subject and the experimenter each listened to the tape over stereophonic earphones. After each passage had been played, the subject responded verbally. The experimenter marked correct responses on a specially prepared data sheet. At the end of each trial, the experimenter added up the number of errors and decided whether the subject had met the criterion for training: two successive trials with no more than two errors each. If criterion had not been met, the next training trial proceeded. If it had been met, the experimenter stopped the training tape, rewound it, removed it from the tape deck, and put the generalization tape on the deck. If the subject ran through 12 training trials without meeting criterion, the experimenter moved on to the generalization test nevertheless. The subject therefore had a brief rest between hearing the training and generalization tapes. The generalization tape was played and responses were recorded just as for the training tape. At the end of each subject's second session, the experimenter explained the purpose of the study to the subject and asked the subject not to discuss the experiment with any other participant.

## C. Results

*1. Gap-Fill (Part A)*

During training, one subject failed to respond within the 10-sec limit to a single selection. Table I summarizes the results of part A of the experiment. Due to various exigencies, three Os did not complete part A of the experiment. Fourteen of the remaining 17 subjects met criterion in a median of 8.8 trials. They made a median of 37.0 errors in training and 4.7 errors out of a possible 12 in generalization. Three Os went 12 trials without meeting criterion. They had a median of 58.0 errors during

TABLE I
Results of Concept Identification Experiment

| | Gap-fill (A) | Changing-note (B) |
|---|---|---|
| Os meeting criterion | 14 | 12 |
| Median trials | 8.8 | 9.3 |
| Median errors | 37.0 | 43.0 |
| Median errors in generalization | 4.7 | 4.6 |
| Os going 12 trials Without meeting criterion | 3 | 4 |
| Median errors | 58.0 | 69.0 |
| Median errors on last 2 trials | 5.0 | 8.0 |
| Median errors in generalization | 4.0 | 4.3 |
| Median errors in generalization for all Os | 4.5 | 4.4 |

training, a median of only 5.0 errors out of a possible 32 on their final two training trials, and a median of 4.0 errors out of a possible 12 in generalization. All 17 Os together made a median of 4.5 errors during the generalization test. During generalization no O ever failed to respond within the 10-sec limit after the end of a passage.

We used the Kolmogorov-Smirnov test (Siegel, 1956) to determine whether subjects responded better than chance in the generalization test. If the subjects were just guessing throughout generalization, their errors would follow a binomial distribution. Table II shows the theoretical and the observed cumulative probabilities of error scores of at least a certain size. The observed cumulative probabilities are given for the 14 Os who met criterion and for all 17 who completed the gap-fill part of the experiment. The last two lines of the table show the maximum deviation of the observed from the theoretical probabilities and the associated significance level from the Kolmogorov-Smirnov text. The 17 subjects who completed part A of the experiment had fewer generalization errors than expected by chance; this difference was significant ($p = .05$). Therefore, these subjects acquired generalizable knowledge as a result of training. The 14 subjects who reached criterion showed the same trend in their generalization error scores, but the difference fell just short of significance. The results indicate that subjects who did not reach criterion during training still learned to classify the passages for the most part correctly and could generalize their knowledge to a new sample of selections.

Inspection of the data indicated that passages differed in their difficulty during training and possibly during generalization. Table III lists the individual selections used in part A of the experiment along with the errors made on them. Results are given separately for the Os who met criterion during training and for those who did not, along with total errors for all Os. The passages are separated into gap-fill and

TABLE II
Analysis of Number of Generalization Errors

| Errors | Binomial cumulative probabilities | Gap-fill cumulative probabilities | | Changing-note cumulative probabilities | |
|---|---|---|---|---|---|
| | | Os meeting criterion ($n$ = 14) | Os completing experiment ($n$ = 17) | Os meeting criterion ($n$ = 12) | Os completing experiment (n = 16) |
| 12 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 11 | .9998 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 10 | .9968 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 9 | .9807 | .9286 | .9412 | 1.0000 | 1.0000 |
| 8 | .9270 | .9286 | .9412 | 1.0000 | 1.0000 |
| 7 | .8061 | .9286 | .9412 | .9167 | .9375 |
| 6 | .6128 | .7857 | .8235 | .9167 | .9375 |
| 5 | .3872 | .7143 | .7059 | .9167 | .9375 |
| 4 | .1938 | .5000 | .5294 | .5000 | .5625 |
| 3 | .0730 | .3571 | .3529 | .4167 | .3125 |
| 2 | .0192 | .1428 | .1765 | .3333 | .2500 |
| 1 | .0032 | .0714 | .0588 | .2500 | .1875 |
| 0 | .0002 | .0714 | .0588 | .0833 | .0625 |
| $D_{max}$ | | .3271 | .3356 | .5295 | .5503 |
| $p<$ | | .10 | .05 | .01 | .01 |

non-gap-fill sets, for both training and generalization. If all passages within a set were equally difficult, the same number of errors would be made on each. We performed Friedman two-way analyses of variance (Siegel, 1956) on errors made during training in order to see whether passages differed in difficulty. The individual subjects were replicates and the different passages were the conditions. Tests were done on the gap-fill passages for subjects who met criterion, for those who did not, and for all Os combined. The results appear in Table III, which gives the values of $\chi_r^2$ and their associated levels of significance. The results show that gap-fill training passages varied in difficulty. Three similar analyses on non-gap-fill training passages yielded no significant results, as Table III shows. Finally, chi square tests were made when possible on error data for individual selections during generalization. Tests were done on total errors per passage across all Os for gap-fill and non-gap-fill passages and for total errors per non-gap-fill passage across Os who met criterion. None of the results proved significant.

The data in Table III also suggest that subjects made more errors on non-gap-fill than on gap-fill passages. To test this hypothesis, we used the Wilcoxon signed-ranks test for matched samples (Siegel, 1956). The test was done on differences across subjects between total errors on gap-fill and total errors on non-gap-fill passages. Table III shows the results. Subjects who reached criterion and all Os combined who finished part A of the experiment made significantly more errors on non-gap-fill than

TABLE III
Errors on Gap-Fill Selections

| Selection | Total errors | | |
|---|---|---|---|
| | Criterion Os ($n$ = 14) | Noncriterion Os ($n$ = 3) | All Os ($n$ = 17) |
| Training gap-fill | | | |
| Mozart: Piano Quintet in E-flat, K. 452 (III) mm. 1-8 | 19 | 16 | 35 |
| Mozart: Flute Quartet in A, K. 298 (II) mm. 1-8 | 9 | 8 | 17 |
| Beethoven: Violin Concerto (III) mm. 1-8 | 24 | 4 | 28 |
| Mozart: Eine Kleine Nachtmusik, K. 525 (IV) mm. 1-8 | 13 | 9 | 22 |
| Schubert: String Quartet in E-flat, op. 125 (D. 87) (II) mm. 1-8 | 41 | 16 | 57 |
| Mozart: Symphony no. 18 in F, K. 130 (III) nm. 1-8 | 8 | 5 | 13 |
| Beethoven: String Quartet in B-flat, op. 18 no. 4 (I) mm. 42-49 | 31 | 22 | 53 |
| Beethoven: Piano Sonata in A, op. 2 no. 2 (III) Trio, mm. 1-8 | 36 | 10 | 46 |
| Friedman $\chi_r^2$ | 31.63 | 8.13 | 31.51 |
| $p \leqslant$ | .001 | — | .001 |
| Non-gap-fill | | | |
| Mozart: String Quartet in E-flat, K. 428 (IV) mm. 1-8 | 35 | 10 | 45 |
| Mozart: String Quartet in D, K. 499 (II) Trio, mm. 1-8 | 52 | 4 | 56 |
| Mozart: Symphony no. 41 in C, K. 551 (III) Trio, mm. 1-8 | 25 | 8 | 33 |
| Beethoven: String Quartet in B-flat, op. 130 (II) mm. 1-8 | 35 | 11 | 46 |
| Haydn: String Quartet in E-flat, op. 64 no. 6 (IV) mm. 1-4 | 35 | 14 | 49 |
| Mozart: Piano Concerto in E-flat, K. 482 (III) mm. 1-8 | 44 | 11 | 55 |
| Haydn: String Quartet, op. 33 no. 6 (III) mm. 1-8 | 50 | 17 | 67 |
| Haydn: String Quartet in G, op. 64 no. 4 (II) mm. 1-4 | 19 | 10 | 29 |
| Friedman $\chi_r^2$ | 4.53 | 6.46 | 5.03 |
| $p \leqslant$ | — | — | — |
| A-NA Mdn | −7.5 | — | 6.5 |
| T: $p \leqslant$ | .01 | | .05 |
| Generalization Gap-fill | | | |
| Haydn: Symphony no. 77 in B-flat (IV) mm. 72-80 | 6 | 0 | 6 |
| Haydn: Symphony no. 79 (II) mm. 61-68 | 5 | 1 | 6 |

(Continued)

Conservatorio Superior de Música del Real Gobierno de Madrid

TABLE III (*Continued*)

| Selection | Total errors | | |
|---|---|---|---|
| | Criterion Os ($n$ = 14) | Noncriterion Os ($n$ = 3) | All Os ($n$ = 17) |
| Mozart: Symphony no. 13 in F, K. 112 (II) mm. 1-5 | 2 | 1 | 3 |
| Mozart: Symphony no. 24 in B-flat, K. 182 (II) mm. 1-8 | 3 | 2 | 5 |
| Mozart: Sonata for Piano and Violin in G, K. 301 (II) mm. 1-6 | 2 | 0 | 2 |
| Beethoven: String Quartet in B-flat, op. 18 no. 6 (IV) mm. 1-8 | 9 | 2 | 11 |
| $\chi^2$ | — | — | 9.0 |
| $p \leq$ | | | — |
| Non-Gap-fill | | | |
| Mozart: String Quartet in B-flat, K. 458 (III) mm. 1-8 | 8 | 0 | 8 |
| Haydn: Symphony no. 104 in D (IV) mm. 3-10 | 6 | 1 | 7 |
| Mozart: Symphony no. 35 in D, K. 385 (III) Trio, mm. 1-8 | 9 | 1 | 10 |
| Haydn: Quartet op. 33 no. 3 in C, (IV) mm. 1-8 | 4 | 2 | 6 |
| Mozart: Symphony in G, K. 45a (III) mm. 1-8 | 5 | 1 | 6 |
| Haydn: String Quartet in G, op. 64 no. 4 (IV) mm. 1-4 | 4 | 1 | 5 |
| $\chi^2$ | 3.67 | — | 2.29 |
| $p \leq$ | — | | — |
| GA-GNA Mdn | .4 | — | .5 |
| $T$: $p \leq$ | — | | — |

on gap-fill training passages. No similar difference in difficulty between gap-fill and non-gap-fill selections occurred during generalization.

In summary, subjects learned to put gap-fill and non-gap-fill melodies into different classes, without explicit academic training in musical analysis. They subsequently proved able to classify new instances of each type of melody in a generalization test. During training, gap-fill passages differed among themselves in difficulty, but non-gap-fill ones did not. The gap-fill training passages were easier to identify correctly than were the non-gap-fill ones. Generalization passages showed no variation in difficulty either within or between types of passages.

### 2. *Changing-Note (Part B)*

Table I shows the results of part B of the experiment. During training, no subject ever failed to respond within the 10-sec limit. Due to various exigencies, four Os did not finish this part of the experiment. Twelve subjects met criterion in a median of 9.3

trials. They made a median of 43 errors during training and a median of 4.6 errors during generalization. Another four subjects underwent 12 training trials without meeting criterion, generating a median of 69 errors. They made a median of 8 errors on the last two training trials and a median of 4.3 errors during generalization. All 16 Os made a median of 4.4 errors during the generalization test. No O ever failed to respond within the 10-sec limit during generalization.

Table II shows the cumulative probabilities that subjects made error scores of at least a certain size. Results are given separately for Os who reached criterion and for all Os who completed part B of the experiment. Kolmogorov-Smirnov tests showed that subjects who met criterion made significantly fewer errors than expected by chance, as did all Os who completed part B. Subjects who met criterion and those who did not seemed to have acquired generalizable knowledge about changing-note melodies as a result of training.

Table IV shows the individual passages used in part B and the errors made on them. Results are given separately for the Os who met criterion, for those who did not, and for all Os combined who completed part B. The passages are separated into changing-note and non-changing-note sets, for both training and generalization. Friedman tests indicate that both the changing-note and non-changing-note passages varied among themselves in difficulty during training. The non-changing-note passages generated more errors than did the changing-note ones. Wilcoxon tests showed that this difference was statistically significant for Os who met criterion and for all Os combined. Chi square tests were made when possible on generalization error scores and gave no indication that changing-note or non-changing-note passages varied in difficulty. Nor was there evidence from Wilcoxon tests that the changing-note passages differed in difficulty from non-changing-note ones.

### *3. Gap-Fill versus Changing-Note*

Of the 20 subjects who began the experiment, 10 managed to meet criterion in both parts A and B. Data from the latter subjects permit comparisons between the two parts of the experiment. During training, these subjects made a median of 6.0 more errors on gap-fill than on changing-note passages but took a median of .7 more trials to reach criterion on changing-note than on gap-fill passages. Wilcoxon tests showed that neither difference is significant. The 10 subjects also made a median of .8 more errors in generalization on changing-note than on gap-fill selections; again, the difference is not significant. In brief, we have no evidence of any difference in difficulty between the gap-fill and the changing-note parts of the experiment.

## D. Discussion

Our experimental results indicate that at least two classes of melodic processes obtained from theoretical analysis can act as a basis for perceptual differentiation between melodic types. The melodic types, therefore, meet the minimal requirement

TABLE IV
Errors on Changing-Note Selections

| Selection | Total errors | | |
|---|---|---|---|
| | Criterion Os ($n$ = 12) | Noncriterion Os ($n$ = 4) | All Os ($n$ = 16) |
| Training changing-note | | | |
| Mozart: Piano Sonata in E-flat, K. 282 (II) mm. 1–4 | 30 | 18 | 48 |
| Mozart: Fantasy in D, K. 397 mm. 12–15 | 10 | 16 | 26 |
| Mozart: Piano Quintet in E-flat, K. 452 (III) mm. 17–20 | 33 | 19 | 52 |
| Beethoven: Piano Sonata in C, op. 10 no. 1 (I) mm. 1–8 | 14 | 17 | 31 |
| Mozart: Oboe Quartet in F, K. 370 (I) mm. 1–8 | 26 | 24 | 50 |
| Beethoven: String Quartet in F, op. 18 no. 1 (IV) mm. 1–4 | 22 | 10 | 32 |
| Mozart: Symphony no. 39 in E-flat, K. 543 (III) Trio, mm. 1–8 | 24 | 10 | 34 |
| Mozart: Piano Sonata K. 311 (I) mm. 7–10 | 10 | 16 | 26 |
| Friedman $\chi_r^2$ | 26.60 | 2.52 | 21.75 |
| $p \leqslant$ | .001 | — | .001 |
| Non-Changing-note | | | |
| Beethoven: Symphony no. 2 in D, op. 36 (III) mm. 85–92 | 27 | 13 | 40 |
| Beethoven: Piano Sonata in G, op. 49 no. 2 (I) mm. 20–23 | 45 | 24 | 69 |
| Mozart: Piano Quartet in G Minor, K. 478 (III) mm. 1–4 | 47 | 22 | 69 |
| Mozart: String Quartet in A, K. 464 (I) mm. 1–8 | 49 | 25 | 74 |
| Haydn: Symphony no. 94 in G (III) mm. 1–8 | 33 | 23 | 56 |
| Beethoven: String Quartet, op. 18 no. 2, (III) mm. 1–8 | 29 | 16 | 45 |
| Mozart: String Quartet in D Minor, K. 421 (IV) mm. 1–8 | 31 | 19 | 50 |
| Haydn: String Quartet in D, op. 50 no. 6 (III) mm. 30–40 | 33 | 16 | 49 |
| $\chi_r^2$ | 13.83 | 5.92 | 16.97 |

for attributing psychological significance to them. Explicit training in musical analysis is not necessary for the types to form the basis of perceptual distinctions. Our findings, although encouraging, are merely an initial step. Further studies are necessary in order to specify just what sort of psychological reality different types of melodic processes may be said to possess, to determine the levels at which this reality is

TABLE IV (*Continued*)

| Selection | Total errors | | |
|---|---|---|---|
| | Criterion Os ($n$ = 14) | Noncriterion Os ($n$ = 3) | All Os ($n$ = 17) |
| $p \leqslant$ | — | — | .02 |
| B-NB Mdn | −10. | — | −10.8 |
| $T$: $p \leqslant$ | .01 | | .01 |
| Generalization changing-note | | | |
| Mozart: Piano Quartet in G Minor, K. 478 (I) mm. 1–8 | 6 | 0 | 6 |
| Mozart: String Quintet in C, K. 515 (I) mm. 1–10 | 5 | 1 | 6 |
| Mozart: String Trio in E-flat, K. 563 (III) mm. 1–8 | 2 | 1 | 3 |
| Haydn: String Quartet, op. 64 no. 5 (I) mm. 1–8 | 3 | 2 | 5 |
| Haydn: Symphony no. 46 in B (II) mm. 1–4 | 2 | 0 | 2 |
| Mozart: Symphony no. 41 in C, K. 551 (II) mm. 1–4 | 9 | 2 | 11 |
| $\chi^2$ | — | — | — |
| $p \leqslant$ | | | |
| Non-changing-note | | | |
| Mozart: Sextet in F, K. 522 (III) mm. 1–4 | 2 | 1 | 3 |
| Beethoven: String Quartet, op. 18 no. 2 (II) mm. 1–6 | 3 | 0 | 3 |
| Haydn: Symphony no. 101 in D (IV) mm. 1–8 | 4 | 0 | 4 |
| Haydn: Quartet, op. 33 no. 2 (II) Trio, mm. 1–8 | 3 | 1 | 4 |
| Mozart: String Quintet in E-flat, K. 614 (I) mm. 1–8 | 9 | 2 | 11 |
| Beethoven: Symphony no. 7 in A, op. 92 (III) mm. 149–163 | 5 | 2 | 7 |
| $\chi^2$ | — | — | 9.5 |
| $p \leqslant$ | | | — |
| B-NB Mdn | .0 | | .4 |
| $p \leqslant$ | — | | — |

represented, and to show the manner in which it operates. These tests will involve techniques such as direct judgments of similarity.

Training and generalization passages of every type in both parts A and B of the experiment differed among themselves in difficulty. One of the most important reasons for this variability is probably the hierarchic nature of the melodies themselves.



As observed earlier, hierarchically complex melodies exhibit different kinds of patterning on different levels. And the more hierarchic levels there are, the more likely that classification will be based upon the patterning of some level other than that of the highest level of closure. This may help to explain why our subjects found the melody of the first movement of Mozart's Oboe Quartet (Fig. 2) particularly difficult to classify. Not only are there at least four hierarchic levels, but the forceful linear motion of the highest level (Fig. 2, graph D) may well mask the changing-note melody that patterns the highest closural level. The whole melody is analyzed as being linear in Meyer (1973, pp. 192–195). The fourth movement of Beethoven's String Quartet in F Major (Fig. 8) is also a changing-note melody. But it is hierarchically simple and, despite octave transfers and potentially problematic afterbeats, our subjects found it relatively easy to classify. The preemptive force of a linear patterning that transcends the closure of a passage may also have made it difficult for subjects to recognize that the opening measures of Schubert's String Quartet in E-Flat create a gap-fill pattern (for an analysis of this passage see Meyer, 1973, pp. 234–238). But in both these cases, errors were probably a result of other problems as well. In the Mozart example, for instance, the changing-note pattern was unusual: instead of moving 1-7-2-1 as in the other instances of this archetype, it moves 1-2-7-1. And in the Schubert example, the gap is created by a downward skip rather than by the more usual upward one.

Inspection (as well as common sense) suggests more difficulty in learning to classify bilinear melodies, such as Fig. 4, compared to those consisting of just a single melodic strand. The degree of internal redundancy also seems to affect ease of classification. For instance, both the Andante from Mozart's Symphony in F Major (Fig. 3) and the last movement of Beethoven's String Quartet in B-Flat, op. 18 no. 6, are gap-fill melodies of the form AA'B. But the former, which is characterized by a high degree of internal redundancy (particularly, the first half involves repetition of a stable tonic triad), was generally classified correctly; the latter, which involves considerably less pitch and pattern repetition and is relatively unstable harmonically, proved much more difficult to classify. In this connection, it should be noted that while changing-note melodies occur in only one form (AA'), gap-fill melodies occur either in antecedent-consequent form (AA') or in so-called bar form (AA'B). The possibility of alternative forms may in part explain why subjects generally found it more difficult to classify gap-fill patterns than changing-note ones, although the difference was not statistically significant.

A review of passages in the light of differential error scores also suggests that secondary parameters—texture and tempo, dynamics and timbre—may significantly influence ease of classification. All else being equal, it seems probable that very rapid tempi and surprising changes in dynamics make melodies more difficult to classify. The organization of texture also seems to play a role in the ease with which the class of a melody is recognized. As one might expect, clear figure-ground division (as in a homophonic texture) facilitates the classification of melodies while textural intricacy makes it more difficult. For instance, though the Adagio melody from Mozart's Fantasy for Piano in D Minor is quite complex (for a discussion, see Narmour, 1977, pp. 89–95), our subjects seem to have learned its class quite readily. And perhaps they

did so because the figure-ground relationship is very clear and the tempo is slow. It is possible, too, that familiarity facilitated classification in this case—and perhaps in some others as well. For Mozart's D-Minor Fantasy is often studied by beginning piano students, and it might have been known (and hence readily remembered) by some of our subjects.

Finally, in at least one case, the high number of errors can be attributed to the use of a passage that was not, properly speaking, a member of the class it was included to represent. The melody of the third movement of Mozart's Piano Quintet in E-Flat is indeed a changing-note melody; but instead of moving around the tonic (1-7-2-1), it moves around the third of the scale (3-4-2-3).

As the preceding discussion indicates, the ease or difficulty with which subjects classify a particularly melody is a result of the intricate interaction among *all* the features and parameters that make the patterning what it is. It cannot be sufficiently emphasized that the passages that our subjects were asked to classify were not invented abstractions or contrived simplifications but real music as it was really performed. Not only melodic organization on all hierarchic levels, but rhythmic and harmonic structure, texture, tempo and dynamics, at once shaped complex patterning and made the task of classification formidable. And because for each passage the relative importance of the various factors is different, general rules explaining how the resulting idiosyncratic relationships help or hinder classification are difficult, perhaps impossible, to discover.

## III. IMPLICATIONS

Melodic archetypes, such as those considered in this experiment, make possible an almost infinite variety of particular instantiations. Our analysis suggests that unless the essential structural tones of a melodic process are kept intact, changing individual notes in a melody by one or several octaves at random should make recognition very difficult. Such octave shifts, for example, might transform a changing-note melody into something like an abortive gap-fill one. Deutsch (1972) found that octave displacements did interfere with recognition of a familiar tune. Dowling and Hollombe (1977) and Idson and Massaro (1978) have produced evidence that maintenance of contour can partly overcome the effects that Deutsch first reported. Having the undistorted tune available for comparison also facilitates recognition (see House, 1977 and Deutsch, 1978; Idson and Massaro also played undistorted tunes in their series). Therefore, maintenance of contour is only one factor that can oppose the effects of octave displacement. Other experimental work bears out the importance of contour in recognition of distorted melodies (Francès, 1958; White, 1960) and in perception of transpositions (Dowling & Fujitani, 1971).

We suspect that one of the crucial considerations for recognition in the face of variation is the presence or absence of hierarchic structuring. If hierarchic organization is present, pattern recognition is possible even when there is significant alteration *as long as* the structural tones that shape the essential process maintain their functions.

When these are changed, recognition should be increasingly difficult. And if the melodies used in an experiment are without hierarchic structure, it should make less difference which tones are altered—recognition should be more or less equally affected.

The idea of different types of melodic processes also illuminates other experimental findings on recognition of melodies. White (1960) found that retrograde variations on a familiar melody were particularly destructive of recognition. Dowling (1971, 1972) has examined the effects of inversion, retrograde, and retrograde inversion on recognition of short atonal sequences in an immediate memory paradigm. A sequence was presented on each trial, followed immediately by a variation on it or by some different sequence. Subjects had to say whether or not the second sequence was a transformation of the first. Dowling found that retrograde inversion made recognition most difficult. Because they are syntactic and create a sense of goal-directed motion, tonal melodies should be more difficult to recognize in inversion, retrograde and retrograde inversion than nontonal ones. The analysis of melodic process, however, suggests that the difficulty of recognizing a transformation will depend on the type of melody used. Changing-note melodies, for example, are more open to retrograde variation than are gap-fill ones. Different types of melodic processes will make one or another sort of transformation easier or harder to use while maintaining recognition. Generally, however, retrograde transformations are very likely to alter severely the process that underlies a particular melody and to render recognition harder. This fact gives some account of the findings of White and of Dowling.

Finally, some attention has been directed during the last 10 years to effects of interleaving two melodies, binaurally or dichotically (Deutsch, 1975; Dowling, 1973; Butler, 1979). Melodic interleaving is essential in the practice of polyphony. The analysis presented here suggests that some melodic types are probably better suited for polyphonic combination than others. For instance, because they quickly make clear the basic melodic process involved and employ motivically as well as functionally contrasting parts, gap-fill melodies—especially those with a patent initial skip and without complex hierarchic organization (see, for instance, Fig. 3)—provide a particularly good basis for polyphonic combination, as even a casual glance at the fugue subjects in Bach's *Well-Tempered Clavier* makes evident. Because they are quite different in these respects, changing-note melodies are not specially well-suited to polyphonic usage.

## REFERENCES

Butler, D. A further study of melodic channeling. *Perception & Psychophysics*, 1979, *25*, 254-268.

Cooper, G. W., & Meyer, L. B. *The rhythmic structure of music*. Chicago, Illinois: University of Chicago Press, 1960.

Dawkins, R. *The selfish gene*. London and New York: Oxford University Press, 1976.

Deutsch, D. Octave generalization and tune recognition. *Perception & Psychophysics*, 1972, *11*, 411-412.

Deutsch, D. Two-channel listening to musical scales. *Journal of the Acoustical Society of America*, 1975, *57*, 1156-1160.

Deutsch, D. Octave generalization and melody identification. *Perception & Psychophysics*, 1978, *23*, 91-92.
Dowling, W. J. Recognition of inversions of melodies and melodic contours. *Perception & Psychophysics*, 1971, *9*, 348-349.
Dowling, W. J. Recognition of melodic transformations: Inversion, retrograde, and retrograde inversion. *Perception & Psychophysics*, 1972, *12*, 417-421.
Dowling, W. J. The perception of interleaved melodies. *Cognitive Psychology*, 1973, 5, 322-337.
Dowling, W. J., & Fujitani, D. A. Contour, interval, and pitch recognition in memory for melodies. *Perception & Psychophysics*, 1971, *9*, 524-531.
Dowling, W. J., & Hollombe, A. W. The perception of melodies distorted by splitting into several octaves: Effects of increasing proximity and melodic contour. *Perception & Psychophysics*, 1977, *21*, 60-64.
Francès, R. *La perception de la musique*. Paris: Vrin, 1958.
Frye, N. *Anatomy of criticism*. Princeton, New Jersey: Princeton University Press, 1957.
House, W. J. Octave generalization and the identification of distorted melodies. *Perception & Psychophysics*, 1977, *21*, 586-589.
Idson, W. L., & Massaro, D. W. A bidimensional model of pitch in the recognition of melodies. *Perception & Psychophysics*, 1978, *24*, 551-565.
Komar, A. *Theory of suspensions*. New York: Norton, 1968.
Levy, J. M. Gesture, form and syntax in Haydn's music. In J. P. Larsen, H. Serwer, & J. Webster (Eds.), *Haydn Studies*. New York: Norton, 1981. Pp. 355-362.
Meyer, L. B. *Explaining music: Essays and explorations*. Berkeley, California: University of California Press, 1973.
Meyer, L. B. Grammatical simplicity and relational richness: The trio of Mozart's G Minor symphony. *Critical Inquiry*, 1976, *2*, 693-761.
Meyer, L. B. Toward a theory of style. In B. Lang (Ed.), *The concept of style*. Philadelphia, Pennsylvania: University of Pennsylvania Press, 1979.
Meyer, L. B. Exploiting limits: Creation, archetypes, and change. *Daedalus*, 1980, *109*, 177-205.
Narmour, E. *Beyond Schenkerism*. Chicago, Illinois: University of Chicago Press, 1977.
Schenker, H. *Der Freie Satz*. Vienna: Universal Edition, 1956.
Siegel, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.
White, B. Recognition for distorted melodies. *American Journal of Psychology*, 1960, *73*, 100-107.
Yeston, M. *The stratification of musical rhythm*. New Haven, Connecticut: Yale University Press, 1976.
Yeston, M. (Ed.) *Reading in Schenker analysis*. New Haven, Connecticut: Yale University Press, 1977.



# 11

# Structural Representations of Musical Pitch

*Roger N. Shepard*


I. Introduction ..... 344
II. Unidimensional Approaches to Pitch ..... 344
A. The Purely Physical Approach ..... 344
B. The Psychophysical Approach ..... 345
C. The Cognitive-Structural Approach ..... 346
III. Potentially Multidimensional Approaches to Pitch ..... 347
A. A Purely Physical Approach ..... 347
B. A Psychophysical Approach ..... 347
C. A Cognitive-Structural Approach ..... 349
IV. The Spatial Representation of Pitch ..... 350
A. Foundations for Unidimensional Psychoacoustic Scales ..... 350
B. Foundations for Multidimensional Cognitive-Structural Representations ..... 351
C. The Form of the Data Required ..... 354
D. Constraints Inherent in Any Spatial Representation ..... 355
E. The Construction of the Implied Structure by Multidimensional Scaling ..... 356
F. The Construction of the Implied Structure by Euclidean Composition ..... 357
G. Considerations in Choosing Structural Components ..... 357
H. A Proposed Set of Component Structures ..... 359
I. Some Derivable Compound Structures ..... 361
V. Illustrative Analyses of Empirical Data ..... 365
A. Multidimensional Scaling (Using Individual Difference Scaling) ..... 365
B. Euclidean Composition (Using Linear Regression) ..... 367
VI. Discussion ..... 369
A. Is Pitch Really Multidimensional? ..... 369
B. Multidimensional Variation of the Pitch of Synthesized Tones ..... 370
C. Further Comments on Psychoacoustic versus Cognitive-Structural Representations ..... 372
D. Alternative Cognitive-Structural Representations of Musical Pitch ..... 373
E. The Two-dimensional Melodic Map ..... 374




Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

F. The Two-dimensional Harmonic Map ..... 376
G. Structural Properties of Musical Scales ..... 378
H. Pitch as a Morphophoric Medium ..... 380
I. The Problem of the Representation of Pitch within a Tonal Context ..... 381
VII. Conclusions ..... 384
References ..... 385

## I. INTRODUCTION

In this chapter I am concerned with the perceived relations between musical tones and with schemes for representing such tones as points in a geometrical structure (psychophysical scale, musical scale, or higher dimensional spatial model) in such a way that the perceptually important relations between the tones have direct counterparts in the geometrical structure (as relations of spatial proximity or, possibly, collinearity). I argue that the traditional psychoacoustic representations are inadequate for the representation of the relations between pitches that are perceived musically, and I propose some more complex structures that appear to capture more of these musically significant relations.

Throughout, I am concerned with relative pitch only, that is, with the perceived relations between two or more pitches. The ability to identify the absolute pitch of a single tone with accuracy, though of considerable interest in itself (Bachem, 1955; Ward, 1963), is quite rare even among musicians and may be essentially independent of the ability to discriminate differences in pitch (Oakes, 1955). For most listeners absolute identification of individual pitches—just as the absolute identification of individual loudnesses, brightnesses, durations, etc.—is comparatively crude and structurally impoverished. For ecological reasons we are quite generally more responsive to the relations and, particularly, to the ratios between physical quantities than we are to those physical quantities themselves (Shepard, 1978c). This appears to be true for the attributes of musical tones as well.

## II. UNIDIMENSIONAL APPROACHES TO PITCH

### A. The Purely Physical Approach

For the purposes purely of physical specification of an acoustic vibration, there is little reason to distinguish the pitch of the vibration from its physical correlate, frequency, if the vibration is a simple, sinusoidal vibration. For anyone concerned with the response of a listener, however, a distinction between psychological pitch and physical frequency becomes essential. For, whether the relationship between tones *a* and *b* is equivalent to that between two other tones *c* and *d*, as these tones are perceived by a human listener, is an empirical question. It can only be answered by determining such things as whether the listener will in fact accept the interval *c*–*d* as

equivalent to the interval *a–b*, whether actual frequencies of error or latencies of response will in fact be equivalent in discriminating *c* and *d* as in discriminating *a* and *b*, and so on. As a result of internal transformations in the process of perceptual registration, psychological equivalence of these kinds is not ensured merely by equating physical differences in frequencies. Perhaps we should equate, instead, differences in log frequency or differences in some other nonlinear function of frequency.

An identification of pitch with physical frequency runs into further difficulties when more complex tones or combinations of tones are considered. A complex sound may be heard to have a well-defined pitch, and listeners may be able to adjust the frequency of a variable sinusoid to match that perceived pitch. Yet the sinusoidal frequency produced by such a match may not in some cases (such as that variously referred to as subjective pitch, residue pitch, or the missing fundamental) correspond to any physically measurable frequency in the originally presented sound (Schouten, 1938; see Plomp, 1976, pp. 112–114; Ward, 1954).

## B. The Psychophysical Approach

The classical psychophysical approach, as originally proposed by Fechner (1860–1966) and most vigorously advocated in recent years by Stevens (1975), attempted to determine the mathematical form of the monotonic psychophysical function that would convert the physical value of a stimulus (its frequency, say, in the case of sinusoidal tone) into a corresponding psychological value (pitch). Pitch was thus treated in much the same way as other subjective variables, such as the loudness of a tone or brightness of a light, which were also regarded as monotonic functions of their corresponding physical variables (amplitude and luminance, respectively). By discovering the appropriate psychophysical function, it was hoped, the physical scale (in our case, frequency) would in effect be relatively stretched and compressed in different regions with the result that if tones *a* and *b* were separated by the same distance as tones *c* and *d* on the resulting transformed scale, the interval *a–b* would be perceptually equivalent to the interval *c–d*. One example of such a scale for pitch is the *mel* scale developed by Stevens and his associates (Stevens & Volkman, 1940; Stevens, Volkman, & Newman, 1937; or, for a more recent variant, see Beck & Shaw, 1961, 1963).

At this point the question arises, however, whether a unique psychological scale of pitch results regardless of the particular choice of listeners, tones, tasks, or types of data collected. Not surprisingly, in view of the complexity and flexibility of the human organism and the multidimensional variety of tones that can give rise to an impression of pitch, the answer is negative. For example, if listeners who have not been especially selected for musical ability are presented with pure sinusoidal tones in the absence of any musical context, the resulting scale of pitch will tend to be compressed at the low-frequency end relative to a log frequency scale—reflecting the fact that the ear is less sensitive to differences between tones standing in the same fre-

quency ratio when those tones are both at the low end of the continuum of audible frequencies. Indeed, successive octaves, which are equally spaced in log frequency (and on a piano keyboard), are for this reason more and more compressed as we move toward the low end of the mel scale of pitch. From the standpoint of the musician, however, such a representation is quite unsatisfactory. It fails to preserve the fact that in a musical context there is something unique about a particular musical interval, such as an octave, regardless of where the two tones separated by that interval fall—at least within a range for which musical intervals are well perceived as such, say between 100 and 4000 Hz (Attneave & Olson, 1971; Dowling, in press; Guttman & Pruzansky, 1962; Watt, 1917, p. 63).

### C. The Cognitive-Structural Approach

In order to obtain data that are more reflective of the relations between tones as they are perceived musically, one needs to depart from traditional psychoacoustic practice. In particular, one needs (a) to present the tones in a more musical context in which, for example, the intervals between the tones are those of a standard musical scale and perhaps the tones themselves possess some of the upper harmonics whose absence renders pure sinusoids so lacking in musical interest and (b) to ensure that the listeners possess at least the minimum level of musical ability necessary to be responsive to musically significant intervals.

The point was forcefully demonstrated by Attneave and Olson (1971). They showed that when the task was one of judging whether the musical intervals in a short melody were or were not the same under various overall shifts of the melody in pitch height (i.e., were or were not correct transpositions), the psychologically equivalent intervals were those on the logarithmic scale rather than on previous psychophysical scales of pitch such as the mel scale. This leads to a fundamental distinction, which I develop more fully later in this chapter—a distinction between what I refer to as the psychoacoustic and the cognitive-structural approach to pitch.

Quite apart from this distinction there is, however, a fundamental limitation inherent in any purely unidimensional representation of pitch. Even if one chooses a scale such that a musically significant interval has the same size at all locations along that scale—that is, even if one chooses the log scale—there is no way in which such a unidimensional scale can represent the fact that under appropriately musical conditions, two tones separated by an especially significant interval, such as the octave (e.g., see Allen, 1967; Krumhansl & Shepard, 1979), are preceived to be more closely related than two tones separated by a slightly smaller interval, such as the major seventh. In order to accomodate an increse in similarity between all tones separated by any particular interval, the rectilinear scale of pitch must be deformed into some more complex structure requiring, as we shall later see, a higher-dimensional embedding space. (It is for this reason that the following approaches are characterized as "potentially multidimensional." However, an explicit consideration of the nature of the multidimensional geometry to which they lead will be deferred to a later section in this chapter.)

## III. POTENTIALLY MULTIDIMENSIONAL APPROACHES TO PITCH

### A. A Purely Physical Approach

The motion that tones separated by certain intervals may actually have more in common than tones separated by somewhat smaller differences in log frequency has, in fact, a long history. Even for pure, sinusoidal tones the degree of musical significance of the relation between a pair of such tones has been recognized since the time of Pythagoras to be related to the simplicity of the numerical ratio between the two frequencies (or, more precisely, in the case of Pythagoras, between the lengths of the two vibrating strings)—with the octave first (with a ratio of 2:1), followed by the perfect fifth (3:2), the perfect fourth (4:3), the major third (5:4), and so on. However, it cannot be the simplicity of the numerical ratio of the frequencies, in any strict sense, that is the determining factor:

1. As Risset (1978) has remarked, ratios that are very close to simple ratios (e.g., 29,998:20,000) are highly complex and yet indiscriminable from the corresponding simple ratios (in this case 3:2).
2. In fact, the interval that listeners accept as the best octave is not the interval with a precisely two-to-one frequency ratio but one with a slightly larger and thus numerically much more complex ratio (Ward, 1954; also see Burns, 1974; Dowling, 1973b, in press; Elfner, 1964; Sundberg & Lindqvist, 1973; Ward & Martin, 1961 Chapter 10, this volume). And the frequencies of the tones actually produced by musicians or preferred by listeners as representative of other musical intervals, as well, often fail to conform to the predicted most simple ratios (Seashore, 1938; Ward & Martin, 1961; Van Esbroeck & Montfort—cited in Risset, 1978).
3. Preference ratings that listeners give for pairs of tones that do stand in the various simple frequency ratios tend to depart systematically from the orderings of those ratios predicted on the basis of what is usually taken to be their numerical simplicities—with, for example, the major third (5:4) often preferred to the numerically simpler perfect fourth (4:3) (e.g., Butler & Daston, 1968; Krumhansl & Shepard, 1979; see also Davies, 1978, p. 158; Fyda, 1975; Van de Geer, Levelt, & Plomp, 1962).
4. Despite numerological theories going back at least to Leibnitz (see Revesz, 1954, p. 50), to my knowledge no psychologically plausible mechanism has been offered to explain how a listener determines that two tones achieve or approximate a simple frequency ratio—particularly when the tones are pure sinusoids and are presented only successively.

del Real Conservatorio Superior de Música

### B. A Psychophysical Approach

von Helmholtz (1862–1954) held that for musical and hence harmonically rich tones, it is the extent of actual coincidence between the upper harmonics rather than

merely the simplicity of the frequency ratio of the fundamentals that primarily determines the degree to which two tones are perceived to have a harmonious relation. The rationale for this view is that failures of coincidence between harmonics of nearly the same frequencies tend to produce rapid, temporal pulsations or beats that, to the extent that they are audible, contribute to the perceived roughness, discord, or dissonance between simultaneously sounded tones. Although departures from strict coincidences between harmonics can be defined purely in terms of the physical stimuli, the importance of the audibility of the beats in producing perceptual roughness or discord ultimately leads to considerations of the response of the ear to small differences in relative frequency at various absolute levels of frequency, amplitude, and timbre—hence my characterization of this approach as psychophysical rather than as purely physical. This approach has more recently been developed in considerable quantitative detail by Plomp and Levelt (1965), Kameoka and Kuriyagawa (1969), and others. Although the degree of consonance predicted for two tones depends in general on the harmonic content of the tones (Kameoka & Kuriyagawa, 1969), for most musical tones the predicted ordering of the standard musical intervals is in approximate agreement with that predicted on the basis of the simplicity of the frequency ratios of the fundamentals—with the octave predicted to be the most harmonious (after the unison), followed by the perfect fifth, and so on.

However, it cannot be that the actual audibility of beats directly determines the perceived musical relation between tones in general. When the tones are pure sinusoids or are presented successively (as melodic rather than harmonic intervals), the same musical intervals are recognized or, under some conditions, are preferred (Krumhansl & Shepard, 1979), although audible beats will be largely or (in the case of successive presentation) completely absent. Moreover, even for simultaneously presented, harmonically rich tones, owing to the properties of the ear, the audibility of the beats for any given interval will depend, as I noted, on the timbre and the placement of the tones within the range of audible frequencies (Risset, 1978). But the significant musical intervals, as I have also noted, can be recognized and appreciated for what they are regardless of their transpositional locations within an appropriate musical range and across musical instruments of very different timbres.

Despite these and other difficulties (shortly to be considered), the psychophysical approach continues to be developed in attempts to deal (in various ways that are beyond the scope of this chapter) with perceived pitch relations between more complex sounds, inharmonic tones, combinations of tones, filtered noises, pulse trains, and the like (e.g., de Boer, 1976; Goldstein, 1973; Terhardt, 1974; Wightman, 1973). In general, the objective of the psychophysical approach to pitch is to formulate quantitative rules according to which perceived pitches of sounds are related to physically specifiable properties of those sounds. However, from the cognitive-structural standpoint taken here, any purely psychophysical approach has two related shortcomings: first, it does not by itself make sufficient provision for the apparently important role in the interpretation of musical tones of cognitive structures that the listener brings to the situation or that have been induced in the listener by a preceding musical context; second, although it may—in discriminating, for example, between "spectral"

and "periodicity" pitch—imply that perceived pitch is multidimensional, in focusing on the relation between physical variables and perceptual responses, the psychophysical approach does not generally concern itself with the precise multidimensional form of the interpretative cognitive structures themselves (although these internal structures may constrain the perceptual response quite as much as the physically specified external stimuli).

### C. A Cognitive-Structural Approach

There is in fact a growing body of empirical evidence that in the perception of tones in a musical context, what each externally presented tone gives rise to within the listener is not a percept that varies continuously with the physical parameters of that tone according to a fixed, though perhaps multidimensional, psychophysical function. Rather, a more apt description seems to be that by a kind of "categorical perception," each continuously variable tone activates the closest node in some discrete cognitive structure corresponding, for example, to a tonal framework or diatonic scale (Balzano, 1978, 1980; Dowling, 1978; Francès, 1958, p. 49; Ward, 1970), which may be explicitly established by a preceding musical context (Krumhansl & Shepard, 1979) or which may be induced by the musical relations between the test tones themselves (Blechner, 1977; Burns & Ward, 1978; Locke & Kellar, 1973; Siegel & Siegel, 1977a,b; Zatorre & Halpern, 1979). (For further indications of the importance of a tonal interpretive framework, see Cohen, 1978; Cuddy, Cohen, & Miller, 1979; Dewar, Cuddy, & Mewhort, 1977; Francès, 1958, Experiments III & IX.) Indeed, in agreement with expectations based on music theory (e.g., see Risset, 1978, p. 526), Krumhansl (1979) has obtained evidence, to which I shall later return, that the perception of the relation between tones separated by any given physical interval depends on the relation of the pair of tones to a contextually established tonal center (or "tonic") and, hence, can never be explained solely in terms of physical properties of the two presented test tones themselves—no matter how complex the specification of those properties may be.

It is understandable that psychophysical investigators, who have traditionally focused on the operating characteristic of sensory transducers, have tended to prefer simple physical stimuli—especially, in the case of the ear, sinusoidal tones varying on easily specified physical variables of frequency and amplitude—and have eschewed the cognitive complications that might arise in a musical context. It is only from the very different standpoint of the cognitive scientist—who is accustomed to working toward the characterization of the abstract mental structures that underlie our higher level perceptual, conceptual, and linguistic competencies—that an investigation of the abstract structural complexities that make possible the appreciation of a Bach fugue or a Beethoven quartet appears at all enticing.

The analogy with the study of language may be helpful here. The investigation of how the ear responds to carefully controlled acoustic events in the time-frequency domain is not irrelevant to the problem of the recognition of speech. But in order to

understand how the spoken sentences of a natural language are perceptually segmented, syntactically parsed, and semantically interpreted, we must concern ourselves even more with highly developed syntactic and semantic structures and rules, which operate at a very different and physiologically less accessible level and, perhaps, in a very different, more discrete and combinatorial mode (Chomsky, 1965; Winograd, 1972).

More specifically, the "competence-performance" distinction that Chomsky enunciated for language (see Chomsky, 1965) may have an instructive analogue for musical cognition. Chomsky argued that our linguistic competence presupposes an implicit knowledge of general algebraic rules of syntactic formation and transformation, even though our actual linguistic performance in any concrete situation may be more or less faulty, distorted, or incomplete—particularly in dealing with long, complex, or self-embedded constructions—owing to limitations of short-term memory, failures of lexical retrieval, difficulties of articulation, external distraction, and so on. Similarly, our musical competence may presuppose an implicit knowledge of underlying structures (e.g., diatonic scales, hierarchies of tonal functions) and transformations (e.g., transpositions, inversions, and the like), even though for reasons of sensory, memory, or motor limitations, we may fail on occassions to comprehend or to execute complex musical passages.

In particular, I suggest that considerations of the sensory limitations of the input transducer—such as the reduced efficiency of the ear in discriminating nearby amusical pitches at the low end of the continuum of audible frequencies, which such psychophysical scales of pitch as the mel scale had been designed to represent—are essentially irrelevant to the problem of the representation of the cognitive structures that underlie the interpretation of musical sequences. In order to preserve the musically important structural relations between tones (the musical intervals of the octave, fifth, and so on) and in order to preserve invariance under the musically important transformations (e.g., transposition), the cognitive representation of musical pitch must have properties of great regularity, symmetry, and transformational invariance.

The question of the representation of musical pitch then comes down to this: How can we represent musical tones as points in a regular structure so that (a) all pairs of tones separated by the same musical interval are separated by the same distance in this structure and (b) tones separated by especially significant musical intervals (such as the octave and perfect fifth) have corresponding special relations (e.g., of spatial proximity) within this structure?

## IV. THE SPATIAL REPRESENTATION OF PITCH

### A. Foundations for Unidimensional Psychoacoustic Scales

Although the various unidimensional scales that psychophysicists have constructed for the representation of pitch differ in detail, they all appear to presuppose certain

basic conditions. I suggest that the most essential of these can be set down roughly as follows:

1. *Unique correspondence.* Each pitch corresponds to a particular point in a metric space.
2. *Preservation of equivalence.* Pitches perceived to stand in the same psychological relation correspond to points separated by the same distance in the space.
3. *Monotonicity.* Pitches perceived to be more similar to a given pitch correspond to points that are closer in the space to the point corresponding to that given pitch.
4. *Unidimensionality.* The metric space is the one-dimensional Euclidean line.

The difference between the individual scales of pitch that have been constructed (such as those of Attneave & Olson, 1971; Beck & Shaw, 1961; Stevens & Volkman, 1940; Stevens, Volkman, & Newman, 1937) arise primarily from differences in how the investigators set about determining (in accordance with the second of these four conditions) which pitches "stand in the same psychological relation." Thus, if the determination was based on the discriminability of sinusoidal tones (just as much as if it were based on the discriminability of narrow bands of white noise) without any musical context, the points representing pitches an octave apart tended to be separated by smaller distances at the low end of the scale. Whereas if the determination was made on the basis of judgments of equivalence of intervals between musical tones under transposition (as by Attneave & Olson, 1971), the points for pitches an octave apart tended, over a suitable musical range, to be equally spaced (also see Dowling, 1978; Null, 1974; Ward, 1954; Burns and Ward, this volume, Chapter 8).

As I have already implied, however, the simultaneous satisfaction of all four of these conditions is not always possible. Particularly in the more musical context the fact that tones separated by special musical intervals (such as the octave) tend to be perceived as having more in common than tones separated by somewhat smaller intervals cannot simultaneously be reconciled with the conditions both of monotonicity (3) and of unidimensionality (4).

## B. Foundations for Multidimensional Cognitive-Structural Representations

If the richer structural relations that apparently hold between the pitches of tones that are interpreted musically are to be fully captured in the spatial representation itself (apart, that is, from the mapping rules that relate that spatial representation to the set of possible pitches), it is the condition of unidimensionality rather than the condition of monotonicity that we should abandon. (Later, however, I consider the possibility of weakening in addition the monotonicity condition and of representing some special relations of pitch by colinearity rather than exclusively by proximity of points in the multidimensional space.) For the present, then, I take as the basis for a cognitive-structural representation of musical pitch, only the first three of the four stated conditions, and I suppose that the representational space may in general be

multidimensional. If, now, tones separated by standard musical intervals are presented to musical listeners under appropriate conditions, all tones separated by the same interval should, I claim, be perceived as musically equivalent—except for tones that are so low or so high in pitch as to exceed the performance limitations of the sensory transducer (cf., again, Attneave & Olson, 1971; Dowling, in press; Guttman & Pruzansky, 1962; Watt, 1917, p. 63; and my present invocation of the competence-performance distinction). By the second of the three retained conditions, the points corresponding to such tones should be equally spaced in the geometrical structure. What, then, will be the form of this structure?

The equidistant, regular spacing of the points implies that the structure maps into itself under all distance-preserving transformations, including translations, reflections, and rotations. In fact, according to a geometrical theorem of classical kinematics, the most general rigid motion of three-dimensional space into itself is a rotation together with a translation along the axis of the rotation—that is, a helical or screwlike motion (Coxeter, 1961, pp. 101, 321; Goldstein, 1950, p. 124; Greenwood, 1965, p. 318; Hilbert & Cohn-Vossen, 1932/1952, pp. 82, 285). This theorem, properly generalized to higher-dimensional spaces and also to include reflections, implies that the structure we are looking for must be a generalized helical structure or some degenerate variant of such a structure—including, in the simplest cases, the straight line, if there is no rotational component, or the circle, if there is no translational component.

For example, as illustrated in Fig. 1, we can deform the usual rectilinear scale of pitch into a simple helix having one complete turn per octave and thus represent any increased similarity at the octave by a reduced distance of separation (e.g., by Path b versus Path a between C and C′ in the figure). Indeed by varying the ratio of the rotational component (the circumference of the helix) to the translational component (the rectilinear height of one complete turn), we can move anywhere between the two degenerate cases of the helix: (1) that in which it is stretched back into a straight line and there is, therefore, no longer any increased similarity at the octave, and (2) that in which it is compressed all the way down into a circle and there is complete perceptual identity of all tones in the octave relation (such as C, C′, C″, C‴, etc.). It was in fact through this line of thought that I was originally led to the computer synthesis of a special set of tones as a concrete instantiation of the latter possibility of complete octave equivalence and, hence, complete circularity of pitch (Shepard, 1964b, 1970; Shepard & Zajac, 1965).

In the process of my earlier explorations of the possibilities of this simple helical representation for pitch, I soon discovered that in similar efforts to account for the increased similarity at the octave, variants of this helix had long before been proposed by other theorists (for example, by Drobish, 1855; Revesz, 1954; Ruckmick, 1929; also see Pickler, 1966, concerning Drobish's proposal and the related 1874 proposal of a planar spiral by Donkin). By now the implied rectilinear and circular components of the helical representation of pitch have become widely enough accepted to be generally referred to as *tone height* and *tone chroma*, respectively (Bachem, 1950; Revesz, 1954; Risset, 1978; also see Kallman & Massaro, 1979).

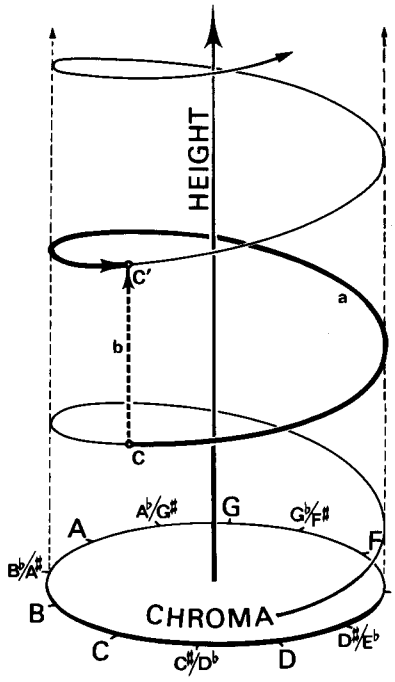



**Fig. 1.** The simple helix of pitch (relabeled from Shepard, 1965).

The distinction that I have urged between the cognitive–structural and the psychoacoustic approaches to pitch applies just as much to the multidimensional or helical as to the unidimensional or rectilinear representations. Thus, my arguments for equality of intervals and invariance under transposition in the perception of musical pitch imply that the helix must be a completely regular helix, falling on the surface of a right cylinder, with successive turns equally spaced with respect to the axis of the cylinder (just as I have previously portrayed the helix, e.g., see Shepard, 1965, p. 105). Only in this way will the helical structure map into itself (by a rigid helical motion consisting of a translation and/or a rotation) under musical transposition. This is in contrast to the helical representations proposed earlier from a more psychophysical standpoint by Drobish (1855) and Ruckmick (1929), in which the helix was intentionally distorted so that intervals (and hence successive turns of the helix) were, once again, relatively compressed toward the low end of the range of audible frequencies.

Of course, this simple helix in three-dimensional space, whether regular or distorted, cannot account for an increased similarity between tones separated by intervals other than the octave. Yet there are strong indications that there is indeed something special about other musical intervals, such as, for example, the perfect fifth or the major third—indications from music theory (e.g., Goldman, 1965; Schenker, 1954/1973), from psychoacoustic considerations (e.g., von Helmholtz, 1862/1954; Plomp & Levelt, 1965; Kameoka & Kuriyagawa, 1969), from group-theoretic argu-

ments (see, especially, Balzano, 1978, 1980); and from recent empirical evidence (Balzano, 1977, in press; Butler & Daston, 1968; Krumhansl & Shepard, 1979). In order to generalize the helical model so that it will accommodate other special relations, in addition to the octave we have to move into a still higher dimensional embedding space. Before proceeding, however, we need first to consider a little more carefully the form of the data that might force us to adopt such a higher-dimensional representation and, then, the constraints inherent in any spatial representation of such data.

### C. The Form of the Data Required

In the case of the musical pitch, we can simplify our problem by restricting consideration to the set of discrete musical tones corresponding to the equally tempered chromatic (12-tone or semitone) scale, which is obtained by dividing each octave into 12 steps that are equal in log frequency. (Later in this chapter I argue that the restriction to the scale of equal temperament, as opposed to some scale approximating "just intonation," with its implication of a distinction between the sharp of one scale degree and the flat of the degree above (e.g., C♯ versus D♭) as well as the choice of whether the octave is taken to correspond to a frequency ratio of exactly two-to-one or somewhat more, may be of little consequence from the present, cognitive standpoint.) Moreover, I propose to assume that the overall structure is determined by the perceived relations between tones at all intervals of an octave or less; that is, that the perceived relations between tones separated by more than an octave can be inferred from the relations between tones that are no more than an octave apart. As a first approximation this assumption seems reasonable on musical grounds, and it has the further virtue of simplifying geometrical considerations. However, it is an assumption that is susceptible to empirical evaluation and can subsequently be modified if necessary.

By my assumptions, then, the generalized helical structure of the set of chromatic pitches is to be determined from just 12 numbers, representing the perceived degrees of musical relations between tones separated by the 12 musical intervals within an octave—minor second, major second, minor third, major third, perfect fourth, tritone, perfect fifth, minor sixth, major sixth, minor seventh, major seventh, and octave (which more or less following others, e.g., Balzano, 1981; Balzano & Liesch, in press, I abbreviate as m2, M2, m3, M3, P4, T, P5, m6, M6, m7, M7, and O, respectively). Such data might be obtained from various sources, including subjective ratings of successive (melodic) intervals presented in a tonal context (Krumhansl, 1979; Krumhansl & Shepard, 1979) or of simultaneous (harmonic) intervals (e.g., Butler & Daston, 1968; also see Davies, 1978, p. 158), or under appropriate experimental conditions, from observed frequencies of confusion or latencies of discrimination between different intervals (e.g., Balzano, 1977; Balzano & Liesch, in press). For comparison, theoretical numbers might also be calculated on the basis of acoustic or

psychoacoustic considerations (such as those presented by Kameoka & Kuriyagawa, 1969).

## D. Constraints Inherent in Any Spatial Representation

The proposal to represent relations of pitch as relations of spatial proximity entails certain constraints on those pitch relations corresponding to constraints on distances in a metric space. Even in the case of the most general type of metric space, the distances are subject to the three standard axioms of distance, which for present purposes can be set down as follows (see Shepard, 1980):

$$d_{ij} \geq d_{ii} = 0 \qquad \text{(positivity)}$$
$$d_{ij} = d_{ji} \qquad \text{(symmetry)}$$
$$d_{ik} \leq d_{ij} + d_{jk} \qquad \text{(triangle inequality)}$$

And for more specific types of spaces, still stronger constraints hold. In the case of $n$-dimensional Euclidean space—in particular, the three-point triangle inequality is expanded in effect into a related $(n + 2)$-point condition (see Blumenthal, 1938, pp. 55 ff.; Shepard, 1974, 1980a).

In the application to relations of pitch, the triangle inequality implies that the closer two tones are to any third tone, the closer those two tones must be to each other. As an example, since the two tones a major third away on either side of any given tone are necessarily a minor sixth from each other, we cannot represent major thirds by small distances without representing minor sixths by at least moderately small distances. Moreover, from the triangle inequality together with symmetry, it follows that the closer any two tones are to each other, the more nearly equal must be the distances from those two tones to any third tone. In particular, this implies that if tones separated by an octave are close together, any two intervals related to each other by an inversion in the octave must have a comparable (close or distant) relation. For example, closeness of tones separated by a perfect fifth would then entail moderate closeness of tones separated by a perfect fourth, remoteness of tones separated by a major seventh would entail moderate remoteness of tones separated by a minor second, and so on.

These geometrical constraints are not without some musical plausibility. For example, there are music-theoretic reasons and even some empirical data (Balzano, 1977, 1981; Balzano & Liesch, in press) supporting the expectation that pairs of intervals related by an inversion in the octave (e.g., P4 and P5, M3 and m6, m2 and M7, etc.) are to some extent related to each other—especially to the extent that octaves are treated as equivalent. Indeed, the fact that a musical listener can hear a sequence of normal musical tones ascending by major sevenths—either as rising by large, slightly less than octave steps in pitch height or as falling by small, minor second steps in chroma (to which I have previously called attention—Shepard, 1964b, p. 2346) is attributable to this phenomenon of partial octave equivalence and the consequent



partial equivalence of inversions. Nevertheless, empirical data on the perceived relations between tones may systematically depart, as we shall see, from even these quite general geometical constraints. Accordingly, I shall consider in a later section how we might accommodate such departures without abandoning a spatial representation, by elaborating the mapping assumed to hold between that representation and the particular relations perceived between the pitches presented in a particular musical context.

### E. The Construction of the Implied Structure by Multidimensional Scaling

Generally, data on the degree of perceived relations between tones separated by different intervals cannot be directly interpreted as distances; they are measures merely of "proximity"—that is, numbers related only monotonically to distances. (This is an agreement with the monotonicity condition already stated.) However, as I demonstrated in 1962, such measures of proximity are often sufficient to determine the corresponding, unknown distances with considerable precision (Shepard, 1962, 1966). Accordingly, one who accepts the assumption that all tones separated by the same musical interval are equivalent can appropriately enter copies of the empirically obtained numbers for the 12 different intervals until all off-diagonal cells are filled in the $13 \times 13$ square matrix giving the estimated proximity between each tone and every other within one complete octave. (Each row or column of the completed matrix will be identical to the preceding except for a circular shift by one cell.) Application of nonmetric multidimensional scaling or "analysis of proximities" (Kruskal, 1964; Shepard, 1962) should then yield the implied spatial structure for one octave in the Euclidean space for the minimum number of dimensions needed to accommodate this structure. In a later section I describe such an analysis using, however, a metric variant of multidimensional scaling subsequently developed by Carroll and Chang (1970) to take account of differences in the patterns of judgments made by different individuals (in this case, listeners with different musical backgrounds).

One limitation of multidimensional scaling from the standpoint advocated here is that there is no guarantee that the structure obtained in any particular number of dimensions will realize the perfect symmetry and regularity that I take to be implied by our musical competence. In particular, as we shall see, if the tones from just one octave are scaled, the two ends of the octave-long structure may be somewhat distorted with the result that it cannot be extended into other octaves in such a way that the entire structure is strictly invariant under transposition. One could collect data for tones spanning two or more octaves, but this would require collecting data for intervals greater than an octave and analyzing a much larger matrix—that is $(12n + 1) \times (12n + 1)$ for $n$ octaves. Moreover, the structure, though considerably more regular, might still be subject to some distortion—again, especially toward its two ends. Fortunately, for certain kinds of data the desired structure can be synthesized in another way that ensures complete regularity and transpositional invariance.

## F. The Construction of the Implied Structure by Euclidean Composition

Just as a helical motion can be generated as a combination of two simple motions (a rotation and a translation along the axis of the rotation), a complex helical structure can be constructed out of simpler, circular and rectilinear components. According to a Euclidean rule of composition (Blackett, 1967, p. 213; Shepard, 1978a), the squared distances between $n$ points in another structural component (a one-dimensional line, say) when added to the corresponding squared distances between $n$ points in another structural component (a circle in a two-dimensional plane, say) yield the squared distances between $n$ points in a resulting combined structure (in this case, a helix in three-dimensional space). In this type of "Euclidean composition" of complex structures out of simpler components, the number of points in the resultant structure is the same as the number in each of its structural components, but the number of dimensions needed to embed the resultant structure is the sum of the dimensions needed to embed each of the components (in the example given, $1 + 2 = 3$).

The additivity of the squared distances in Euclidean composition means that under the appropriate condition we can use methods of linear regression to estimate a set of weights—one weight for each of the proposed structural components—that will provide a best fit to the matrix of data. The appropriate condition is, of course, that we must be able to put or to transform the data into a form in which those numbers are, to a sufficient approximation, linearly related to squared distances. We have in effect thus traded the absence, in multidimensional scaling, of constraints to ensure the desired strict regularity and symmetry of the obtained representation for the imposition of stronger, metric assumptions about the data. In principle, extensions of existing methods should enable one either (a) to impose the desired regularity in the process of obtaining a representation by nonmetric multidimensional scaling (cf. Bentler & Weeks, 1978; de Leeuw & Heiser, 1979) or, equivalently, (b) to carry out the linear regression without using more than the ordinal properties of the data (cf. Kruskal, 1965; Kruskal & Shepard, 1974). But in practice the same overdetermination of the solution that makes nonmetric analyses possible should make the results of a standard linear regression relatively insensitive to departures from the assumed linearity (e.g., see Abelson & Tukey, 1963; Weeks & Bentler, 1979; Shepard, 1974).

## G. Considerations in Choosing Structural Components

Precisely what set of structural components ought to be taken as the initial basis set for the proposed regression analysis is not immediately clear because the several conditions that such a set should satisfy must be balanced against each other. Principle among these conditions are, it seems to me, the following four:

1. There should be at least one component, analogous to the rectilinear dimension of pitch height in the previously proposed simple helix, that can provide for both (a)

the possibility of arbitrarily large distances between tones that are separated by arbitrarily large numbers of semitone steps and (b) the interpolation, when necessary, of additional tones (such as quartertones) within the standard semitone intervals of the chromatic scale.

2. For each musical interval, from the minor second to the octave, that might have a special perceptual relation (that is, a perceptual proximity not adequately provided for by a purely rectilinear component of pitch height), there should be a corresponding component in which the distance between all tones separated by that interval is as small as possible relative to the other distances in that component. An appropriately weighted combination of such components would then come as close as possible to realizing whatever pattern of proximities is indicated by a given set of data.

3. The components needed to fit the data should, however, be as parsimonious as possible—both with regard to the number of these components and with regard to the number of dimensions required for each component.

4. The assignment of positive weights to the components for intervals other than the octave, such as the perfect fifth, should not entail a positive increment in the distance corresponding to the interval of an octave in the overall structure. The motivation for this condition is that it is possible with musical listeners (Krumhansl & Shepard, 1979) or with specially contrived tones (Shepard, 1964b) to approach complete equivalence between all tones standing in the octave relation while at the same time maintaining an augmented proximity for certain other intervals such as the fifth.

If we had to satisfy only the first two conditions, we could choose as the structural component corresponding to the interval consisting of $n$ semitone steps, an $n$-cornered regular simplex in $(n - 1)$-dimensional space. All tones differing from any given tone by the specified interval would then collapse into the same vertex of the simplex and hence would be separated by a distance of zero while any two tones that do not differ either by that interval or by any integer multiple of it would fall on different vertices and hence would be separated by some fixed, larger distance, say unity. However, the triangle inequality constrains us, in thus making the distance arbitrarily small for any specified interval, to make the distances comparably small for all multiples of that interval as well. One consequence is that for the interval of the minor second, the components as just described would collapse into a single point in which all intervals (being alike, integer multiples of the minor second) would be represented by a distance of zero. Hence, any differential proximities of minor seconds could not be achieved by the weighted addition of such a component but would, instead, have to be provided for by the rectilinear pitch-height component (in which, however, tones differing by a minor second are adjacent rather than coincident).

More seriously, the proposal to erect the generalized helical structure by combining the rectilinear component with simplexes of various dimensionalities fails to satisfy Conditions 3 and 4. Thus, in violation of Condition 3, the components for the large intervals—including, particularly, the all-important octave—are notably unparsimonious. While the minor and major third would require, respectively, the equilateral triangle and regular tetrahedron and hence embedding spaces of just two and three

dimensions, the octave alone would require the 12-cornered regular simplex and hence an 11-dimensional embedding space. And in violation of Condition 4, since the octave is not an integer multiple of the perfect fifth, differential emphasis of the perfect fifth would entail a departure from octave equivalence. Yet the evidence indicates that it is just those listeners manifesting the strongest octave equivalence who also perceive the relation between tones separated by the perfect fifth to be very pronounced as well (Krumhansl & Shepard, 1979).

Another more parsimonious type of component is the simple helix, already proposed as a way of representing the octave relation. Thus, for each other interval we could take, as the corresponding structural component, the simple helix that completes one full turn in exactly that interval. The rectilinear subcomponent common to all of the component helices could then be factored out to yield (a) a single pitch-height component as required by Condition 1 and (b) a separate component in the form of a circle (or regular polygon) for each musically significant interval with the component for the interval consisting of $n$ semitone steps represented by a circular $n$-gon with its $n$ equally spaced vertices embedded in just two dimensions. (The representation for the minor second again becomes degenerate and, so, must be entirely provided by the common, rectilinear component.)

Condition 4 still remains unsatisfied, however. Neither of the so-called "perfect" intervals—the musically important fifth or the fourth (which is the inversion of the fifth in the octave)—can be evenly divided into the octave. Consequently, the positive weighting of the circular $n$-gons for either of these two intervals entails a departure from octave equivalence (even when the weight of the rectilinear component corresponding to pitch height is reduced to zero).

### H. A Proposed Set of Component Structures

Taking as essential the ability to represent the limiting case of full octave equivalence, I propose simply to substitute, for the five- and seven-sided polygons already considered for the perfect fourth and fifth, the well-known (and 12-sided) circle of fifths—which, under octave equivalence and hence inversion in the octave, can as well be interpreted as a circle of fourths traversed in the opposite direction. Although tones differing by a fifth (or a fourth) are not strictly coincident in this structure, they are at least adjacent (i.e., nearest neighbors) around the circle—just as tones differing by a second, while never coincident, are nearest neighbors both along the rectilinear pitch-height component and around the chroma circle. Moreover, octaves do become coincident in this structure. Hence, they no longer prevent the simultaneous achievement of octave equivalence and a strong relation at the fifth.

Finally, a further major simplification immediately becomes possible. As is to be expected under octave equivalence, each structural component, except the rectilinear one for pitch height (which is itself inherently incompatible with octave equivalence), is invariant under inversion of intervals. That is, within any component the distance is the same for any interval as for its inversion in the octave. Therefore, we can reduce

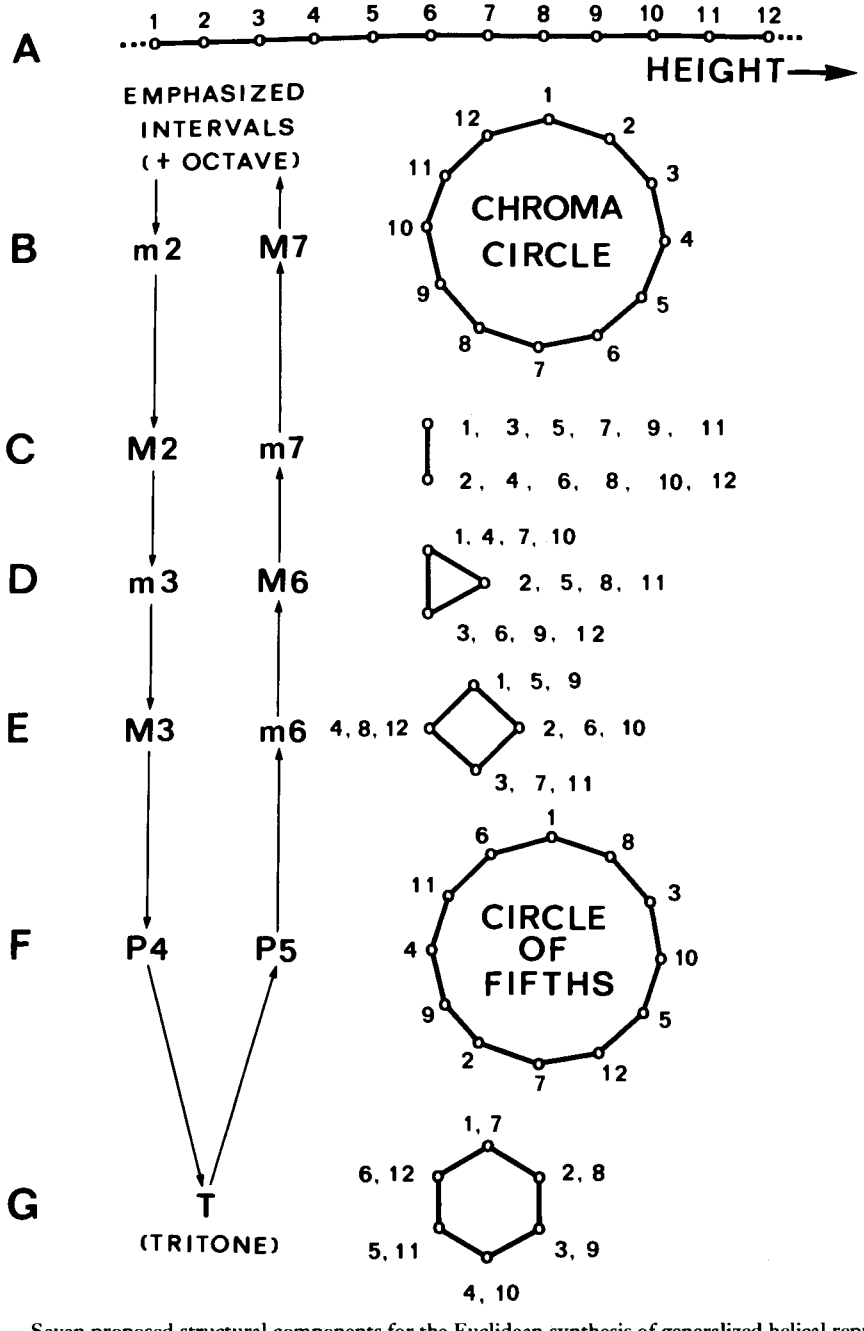



**Fig. 2.** Seven proposed structural components for the Euclidean synthesis of generalized helical representations for musical pitch.

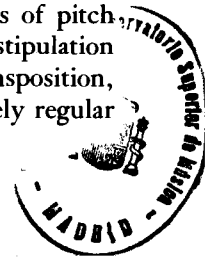
our set of structural components to just the seven exhibited in Fig. 2: (A) a rectilinear component, which emphasizes both proximity of the minor second and large differences in pitch height; (B) the chroma circle, which emphasizes octave equivalence and proximities of both the minor second and its inversion, the major seventh; (C) the degenerate polygon, which emphasizes the major second and its multiples including its inversion, the minor seventh; (D) the triangle, which emphasizes the minor third and its multiples including its inversion, the major sixth; (E) the square, which emphasizes the major third and its multiples including its inversion, the minor sixth; (F) the circle of fifths, which emphasizes the perfect fifth and its inversion the perfect fourth; (G) the hexagon, which emphasizes the tritone, which being exactly half an octave is its own inversion. (The small numerals—1, 2, 3, . . . , 12—attached to the nodes of each of the components A through G represent successively higher semitones within any one octave. For example, if we begin on C, these numerals represent the tones C, C♯, D, . . . , B, respectively, and the next tone, 13 if it were included, would represent a return to C but one octave higher.)

The maximum number of dimensions that could be required using this basic set is given by the sum of the dimensions of the seven components. Since five of the components require two-dimensional embedding spaces, the overall number is $5(2) + 2(1) = 12$ (which is, incidentally, exactly the same as the number of distinct intervals in the octave). However, if some intervals (such as the tritone and perhaps the major second and minor third) turn out to have negligible weights, their corresponding components can be eliminated, yielding a considerable reduction in complexity and dimensionality of the overall structure.

The structural components displayed in the figure already contain some of the representations that have previously been discussed in connection with musical pitch—such as the rectilinear representation of pitch height in a one-dimensional embedding space (A), which corresponds to the log frequency scale of pitch advocated by Attneave and Olson (1971); the chroma circle in two-dimensional space (B), which has been proposed by a number of workers and was concretely instantiated in my computer-generated circular tones (Shepard, 1964b); and the circle (or "cycle") of fifths in two-dimensional space (F), which has played a central role in music theory (e.g., see Goldman, 1965; Schenker, 1954/1973).

## I. Some Derivable Compound Structures

Weighted combinations of two or more of these structural components can yield a variety of more complex structures that show promise for the fuller representation of musical pitch. By combining A and B of Fig. 2, we obtain, as I have noted, a simple helix in three-dimensional space (Fig. 1) that is akin to the various helices that have previously been proposed to account simultaneously for the two factors of pitch height and similarity at the octave. Here, however, in accordance with my stipulation that the resulting structure should map strictly into itself under musical transposition, the helix generated by the proposed rule of composition will be completely regular



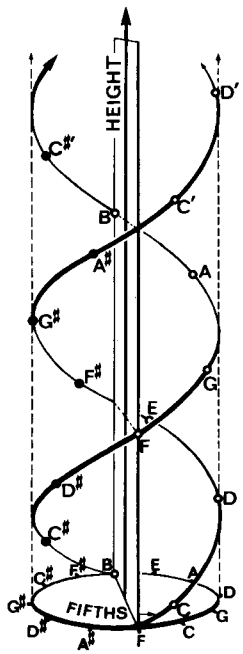



**Fig. 3.** The double helix of musical pitch in three dimensions (obtained by combining Components A and F of Fig. 2). As illustrated for C major, a vertical plane divides the tones in a key (white circles to right) from those not in that key (black circles to left).

and symmetrical rather than, for example, compressed at the low-frequency end. And, by combining other subsets of components, we can obtain other structures that I have more recently proposed for the still fuller representation of musical pitch (Shepard, 1978b, 1981a, 1982) as follows.

If we combine F (rather than B) with A, we obtain the "double helix of musical pitch" (Fig. 3) that I initially derived on the basis of rather different assumptions from those used here—including an assumption that successive steps of the diatonic scale can under some circumstances be heard as equivalent but not including any assumption about the importance of the perfect fifth itself (Shepard, 1978b). As a consequence of the fact that the circle of fifths is nevertheless present as one component, this three-dimensional structure has two related properties of some musical significance:

1. The tones included in any particular major diatonic key can be divided from the tones not in that key simply by passing a plane through the central axis of the double helix. Thus, as is illustrated, the plane passing through the diagonally opposite B and F in the base circle in the figure divides the tones (on the right) that are in the key of C major from the tones (on the left) that are not—indicated (in an analogy with a piano keyboard) by circles that are white and black, respectively.

2. Transpositions into the most closely related keys are achieved by the smallest

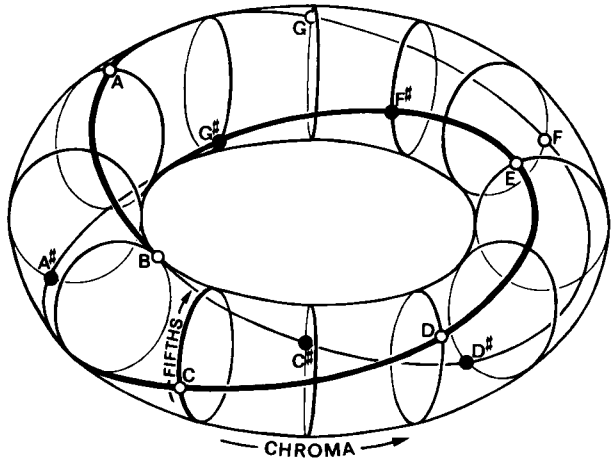



**Fig. 4.** The double helix wrapped around a torus in four dimensions (obtained by combining components B and F of Fig. 2).

angles of rotation of the dividing plane about the central axis. Thus, the two keys most closely related to the key of C are those of F and G—adjacent to C around the circle of fifths. (For convenience, in this and subsequent figures, the chromatic tones between adjacent C-major tones are indicated by sharps rather than by the equivalent flats—e.g., between C and D by C♯ rather than D♭.)

If we combine F with B rather than with A, we obtain the double helix wound around a torus. This structure, generated as it is from a pair of two-dimensional components (the chroma circle and the circle of fifths), can only be isometrically embedded in four-dimensional space, where it realizes the full degree of symmetry required for transpositional invariance. A somewhat crude and distorted idea of its form can be gained from the sketch of a three-dimensional "projection" of this toroidal structure shown in Fig. 4. Since the points on the double helix of the earlier Fig. 3 fall on the surface of a right circular cylinder (the so-called "topological," "Cartesian," or "direct product" of the circle and the straight line—Blackett, 1967), one can think of the toroidal structure in Fig. 4 as having been obtained from that earlier structure by cutting a one-octave segment out of the cylinder and then bending the cylindrical segment around until its two free ends join to form the closed torus (the direct product of two circles). The move is exactly analogous to that of bending a one-octave segment out of the rectilinear scale of pitch height until the two free ends of that segment join to form the closed chroma circle in two-dimensional space. In both cases such a move would be required for the representation of tones, like my computer-generated circular tones (Shepard, 1964b), that achieve complete octave equivalence. Again, either of the two extremes can be approached in which either component completely dominates the other.

In order to obtain a more general representation, we can combine all three of these components, namely, the rectilinear pitch-height component (A), the chroma circle

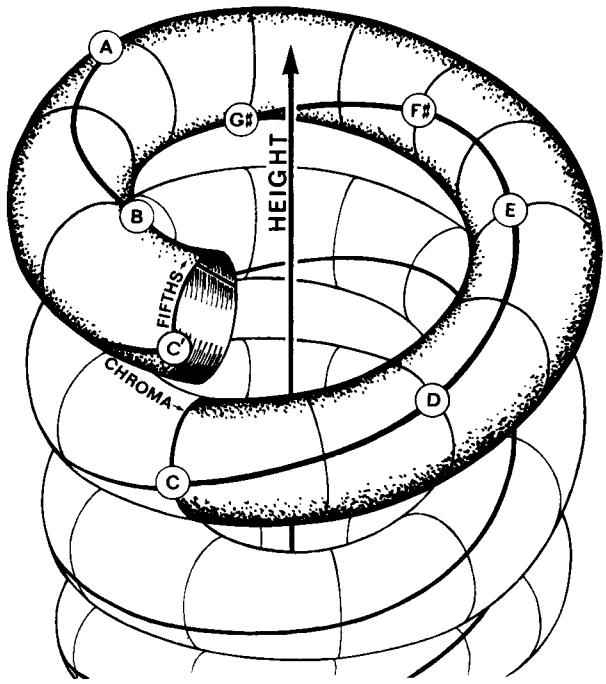



**Fig. 5.** The double helix wrapped around a helical cylinder in five dimensions (obtained by combining Components A, B, and F of Fig. 2).

(B), and the circle of fifths (F). The number of dimensions required to embed the resulting structure is five (1 + 2 + 2), and, again, the intrinsic form of the structure can be no more than crudely suggested by the two-dimensional portrayal of a three-dimensional "projection" (as in Fig. 5). This structure, which is roughly describable as a double helix wound around a helical cylinder, should, however, go farther than any other that has previously been proposed to account for perceived relations of musical pitch. By varying the relative weights of the three components, one can account for (a) proximity in pitch height—that is, perceived similarity of tones separated by small intervals such as the major and particularly the minor second (as well, of course, as the even greater perceived similarities of any tones interpolated at microtonal steps within the semitone interval)—(b) heightened similarity at the octave, including the limiting case of complete octave equivalence, and (c) heightened similarity at the perfect fifth and (to the extent that octave equivalence is more heavily weighted than pitch height) its inversion, the perfect fourth. Moreover, much as I already explained for the double helix, tones in any major diatonic key remain linearly separable from tones not in that key with transpositions between the most closely related keys achieved by the smallest rotations of the hyperplane that separates them in five-dimensional space. (For some recent evidence for the psychological importance of closeness of relations between keys see Bartlett & Dowling, 1980.)

The most prominent musical interval that is not specifically provided for by the five-dimensional generalized helical structure of Fig. 5 is the major third. One can of course add the square component, E, and hence two more dimensions, if this proves necessary in fitting data. Note, also, that the previously considered alternative component for major thirds—namely, the three-dimensional regular tetrahedron—could also be introduced, even within the scheme shown in Fig. 2, by forming an appropriately weighted combination of components E and C. In the interest of parsimony, however, I hope that special components (C, D, and G) for the major second, minor third, and tritone (and their inversons) will prove to be dispensable.

## V. ILLUSTRATIVE ANALYSES OF EMPIRICAL DATA

### A. Multidimensional Scaling (Using Individual Difference Scaling)

As indicated earlier, many different types of data on the relations between musical tones differing in pitch might usefully be subjected to multidimensional analysis—including subjective judgments of degree of musical relation, actual frequencies of confusion or latencies of discriminative reaction, or (for comparison) acoustic data (whether empirically measured or theoretically calculated) concerning, for example, alignment of harmonics. Whatever the type of data, if the predominant effects include an overall dependence on pitch height (with tones separated by a very small number of semitones generally more closely related than tones separated by a very large number of semitones) and a strengthened relation between tones separated by the octave, the perfect fifth, and/or the perfect fourth, the pattern of those data should be approximated by the five-dimensional structure schematized in Fig. 5 or, if some of these effects are relatively weak, by one of its simpler limiting cases (depicted in Figs. 1, 3, or 4).

Purely for purposes of concrete illustration, I here present the results that we obtained by applying the multidimensional scaling method INDSCAL (Carroll & Chang, 1970) to the data that Krumhansl and I collected on the judged relatedness to the tonic of tones differing by each of the 12 intervals from the tonic within the octave (see Krumhansl & Shepard, 1979). INDSCAL, which is designed to capitalize on differences between individuals in their weightings of different dimensions (or, in this case, different components), seemed particularly suitable since we found by other analyses that the individual listeners differed widely. (The rows of the matrix were circularly displaced duplicates of each other, as described earlier.)

Figure 6 shows the four-dimensional solution projected onto the plane of Dimensions 1 and 2 and onto the orthogonal plane of Dimensions 3 and 4. As suggested by the added broken-line circle, the first projection, shown on the left, is essentially the chroma circle, going clockwise from C (through C♯, D, D♯, etc.) around to C′ one octave above. The configuration departs from the chroma circle, however, in that the spacing is wider near C and C′ and, particularly, in that C and C′ are separated from

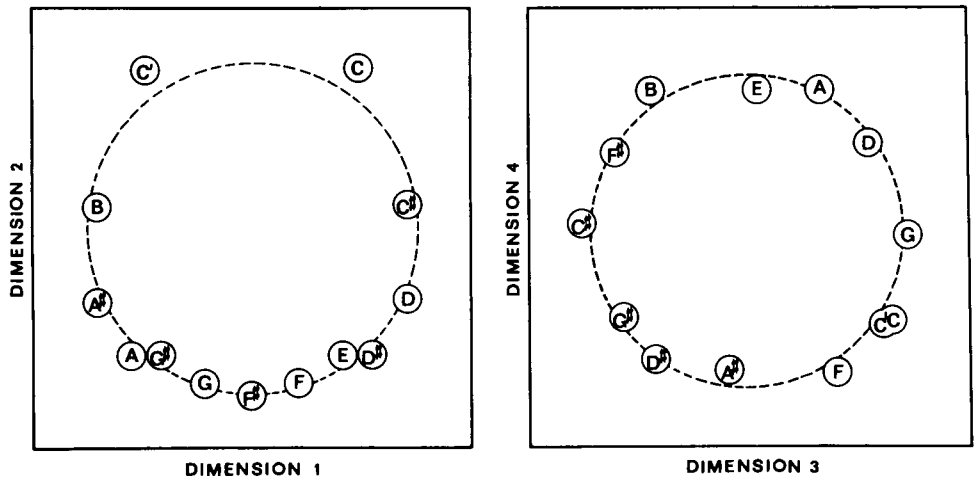



**Fig. 6.** Two orthogonal projections of the four-dimensional INDSCAL solution obtained from an analysis of Krumhansl and Shepard's (1979) data.

each other by a large gap, indicating nonequivalence of octaves on Dimension 1. Dimension 1 thus seems to combine one dimension of the chroma circle with the dimension of pitch height. The projection shown on the right, however, is an almost perfect circle of fifths, with C and C′ nearly superimposed, indicating complete octave equivalence. Except for the gap between C and C′ on Dimension 1, the four-dimensional configuration is precisely the double helix on the torus depicted in Fig. 4, which was generated by Euclidean composition from the chroma circle and the circle of fifths (B and F in Fig. 2).

Figure 7 displays the INDSCAL weights for each of the listeners on each of the four dimensions. The plot on the left shows that the listeners with the least extensive musical backgrounds (represented by the triangles) had the heaviest weights on Dimension 1, which separated the tones with respect to pitch height, while the plot on the right shows that the listeners with the most extensive musical backgrounds (represented by the circles) had the heaviest weights on Dimensions 3 and 4, which contained the circle of fifths and implied complete equivalence between octaves. Moreover, the fact that the points for all listeners fall on a 45-degree line in that right panel means that the circle of fifths emerges, to whatever extent that it does for any one listener, as an integrated whole and never one dimension at a time. That the points for Group 1 listeners also fall close to the (broken) 45-degree line in the left panel indicates that under the complete octave-equivalence characteristic of the most-musical listeners, the chroma circle, too, comes and goes as an integrated unit.

My interpretation of the obtained four-dimensional structure in Fig. 6 is as follows: It is essentially a one-octave piece of the endless five-dimensional theoretical structure portrayed in Fig. 5. But because it includes only one octave, the gap between the two ends, which should have been represented by a displacement in a separate fifth

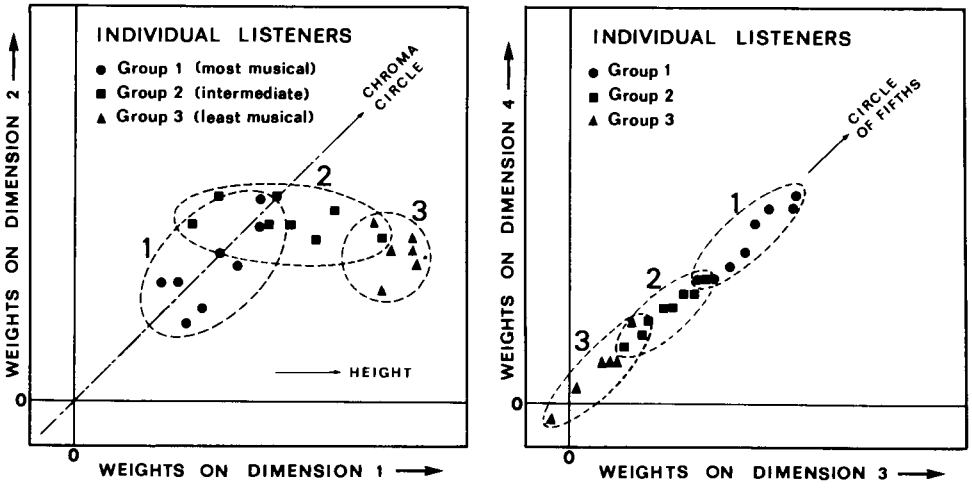



**Fig. 7.** Weights for individual listeners in Krumhansl and Shepard's experiment plotted against the four dimensions of Fig. 6. (Each broken curve encloses all points for listeners within one of the three groups.)

dimension, has (with a small distortion) been accommodated in the four dimensions of the embedding space of the torus generated from the circle of chroma and of musical fifths. In other words, the separation in pitch height between C and C′ an octave above has been achieved by cutting through the torus in Fig. 4 at C and springing it slightly apart in that same four-dimensional space rather than, as in Fig. 5, in an orthogonal fifth dimension. My prediction is that if similar data were collected and analyzed for a range of tones spanning two or three octaves, the data could no longer be fit by a small distortion of this sort in the four-dimensional space, and thus the truer, five-dimensional structure would emerge.

Although this analysis is only intended to be illustrative and should be tried with other, more complete sets of data before definite conclusions are drawn, the results do suggest that the kind of higher-dimensional structures that I have been proposing here (and in Shepard, 1978b, 1981a,b, 1982) are not merely fanciful construction but are, to some degree, actually implied by empirical data.

### B. Euclidean Composition (Using Linear Regression)

In a very preliminary exploration of the feasibility of accounting for the data, alternatively, in terms of Euclidean combinations of regular structural components, Shelley Hurwitz and I carried out a multiple linear regression of the data considered in the preceding section onto the squared interpoint distances in each of the four one-octave component configurations—A, B, F, and E of Fig. 2—that I thought to be most important for accounting for perceived relations in pitch—namely, those of height, chroma, fifths, and major thirds, respectively. (As I have already implied,

Escuela Superior de Música - MADRID

further work may succeed in improving the fits reported here by estimating the transformation that converts the data into numbers that most nearly behave as squared distances. However, I doubt that such refinements will change the general conclusions reached here concerning the relative importances of the component structures.)

Previous cluster analyses had revealed that the listeners fell into three internally homogeneous subgroups having qualitatively different patterns of judgments and, correlated with those, considerable differences in extents of musical background (Krumhansl & Shepard, 1979; also see Fig. 7). For this reason we carried out the multiple regression for each of these three subgroups separately. The results, summarized in Table I, reveal a rather simple pattern. For almost all measures of relative importance of each of the four components, the circle of fifths was dominant for the most-musical listeners (Group 1), the chroma circle was dominant for the intermediate listeners (Group 2), and the rectilinear component of pitch height tended to become dominant for the least-musical listeners (Group 3). (The significant correlations in Table I are all negative because the data were measures of proximity rather than measures of distance—squared or otherwise.)

These results are quite compatible with the multidimensional scaling results discussed in the preceding section. Presumably by taking advantage of the greater con-

TABLE I

Krumhansl and Shepard's (1979) Data on the Relatedness of Pitches to a Tonal Center, Regressed on Proposed Components of Height, Chroma, Perfect Fifths, and Major Thirds—for Listeners with Three Levels of Musical Background[a]

| Component | Height (A)[c] | Chroma (B)[c] | Fifths (F)[c] | Thirds (E)[c] |
|---|---|---|---|---|
| Group 1 (most musical) | | | | |
| Variance explained[b] | .004 | .205 | [.425] | .194 |
| Simple correlation | .061 | −.454 | [−.705] | −.566 |
| Beta weight | −.098 | −.338 | [−.606] | −.448 |
| F ratio | .417 | 4.847 | [16.551] | 9.036 |
| Group 2 (intermediate) | | | | |
| Variance explained[b] | .004 | [.359] | .206 | .157 |
| Simple correlation | −.067 | [−.563] | −.518 | −.507 |
| Beta weight | −.255 | [−.519] | −.412 | −.403 |
| F ratio | 1.774 | [7.178] | 4.973 | 4.586 |
| Group 3 (least musical) | | | | |
| Variance explained[b] | .310 | [.389] | .081 | .088 |
| Simple correlation | [−.557] | −.463 | −.321 | −.354 |
| Beta weight | [−.742] | −.576 | −.252 | −.301 |
| F ratio | [31.095] | 18.315 | 3.715 | 5.300 |

[a] The boxes enclose the largest entry in each row.

[b] "Variance explained" means the increment in multiple R-squared obtained by adding the indicated component to the components listed to its left.

[c] Designation in Fig. 2.

straints imposed here, however, the new results also go beyond the earlier results in providing support for some theoretically derived expectations that were not directly confirmed by multidimensional scaling. First, although from the generally high weighting of the two circular components of chroma (B) and fifths (F) we again see the emergence of the four-dimensional toroidal structure in Fig. 4, we now see the fifth dimension of pitch height (A) also emerging as an orthogonal component for the less-musical listeners (rather than manifesting itself merely as a distortional stretch-plus-gap in the four-dimensional torus). And, second, we see that the component emphasizing the harmonically important major thirds (E), although never the most important factor for any group of listeners, tends to be more important than the component of pitch height for the more-musical listeners (Groups 1 and 2).

Apart from the further dimensions that would be required by the addition of this further component for major thirds, the overall structure defined by the first three components—A, B, and F—is, of course, the five-dimensional structure schematically portrayed in Fig. 5. As expected, this structure systematically changes in its linear proportions between the different groups of listeners: For the most-musical Group 1, the two circular components are the largest components with the circle of fifths clearly dominant. For the intermediate Group 2, the same two circular components remain the largest, but it is the chroma circle that now dominates. And for the least-musical Group 3, pitch height and the chroma circle are the largest components with height generally dominant and with the circle of fifths (like the square of thirds) reduced essentially to naught.

## VI. DISCUSSION

### A. Is Pitch Really Multidimensional?

Psychoacoustic investigators who are accustomed to thinking of pitch as a function of a single physical attribute such as frequency may still be inclined to think that pitch itself must surely be unidimensional. Except for the dimension that I have referred to as pitch height, all the dimensions of pitch that I have been discussing, including those underlying the circles of chroma and of fifths, might from that standpoint be dismissed as existing (if at all) only at some higher level of cognitive interpretation that need not concern the psychophysical investigator and, perhaps, that should not even be subsumed under the term "pitch."

Such a view is encouraged, moreover, by the usual finding in psychoacoustic experiments that the principal determinant of the perceived relation between two tones separated in frequency is simply that separation (measured, for example, as a difference in log frequency). Seldom is there much indication of an increase in perceived similarity at any particular musical interval, except for the occasional finding of a relatively slight effect at the octave. Indeed, I myself have argued that the multidimensional scaling results of at least two psychophysical studies (including the one by Levelt, Van de Geer, & Plomp, 1966) that seemed to provide evidence for a degree

of octave equivalence and even for the existence of two or more dimensions of pitch were artifacts of the scaling method used and that the data were entirely consistent with a purely one-dimensional interpretation (Shepard, 1974, pp. 387–388).

As I have noted, however, the failure to find evidence for a heightened perceptual relation at special musical intervals in most psychoacoustic experiments appears to have been a consequence of the typical avoidance in those experiments of conditions that might ensure an appropriately musical set in the listeners—that is, an avoidance of (a) a preceding musical context, (b) specifically musical intervals between the test tones, (c) harmonically rich, and hence musical, tones, and (d) the selection of listeners with at least some degree of musical background or ability. When instead of avoiding such conditions we intentionally arrange for them, the results become quite different and, indeed, the factor of the musical significance of each interval comes to dominate the factor of mere separation in pitch height (Krumhansl, 1979; Krumhansl & Shepard, 1979).

And, incidentally, the fact that individual listeners can be shown to differ not only widely (Seashore, 1938; Shepard, 1964b) but also in accordance with qualitatively different patterns (Krumhansl & Shepard, 1979; Shepard, 1981a; also see the present Fig. 7) implies that pitch must be multidimensional. If pitch were only a matter of separation in pitch height or log frequency, listeners would be expected to differ only in their fineness of discrimination of pitches that are more or less close to each other in log frequency. In fact, however, they also differ markedly in the extent to which they are sensitive to special musical relations such as the octave and the perfect fifth (Allen, 1967; Krumhansl & Shepard, 1979).

As I argue in the following section, even from the purely psychoacoustic standpoint the attempt to dismiss the existence of more than one dimension of pitch is untenable in the face of the following three psychoacoustic facts—which, by means of digital techniques for the additive synthesis of arbitrarily specified tones, can be demonstrated even with nonmusical listeners and in the absence of a musical context: First, tones separated by musically significant intervals such as the octave and fifth, whose augmented perceptual proximities must entail the deformation of a geometrical representation of pitch into a higher dimensional embedding space, can be made arbitrarily similar and, ultimately, indiscriminable. Second, the implied component dimensions of pitch such as height and chroma, which normally vary together in perfect correlation, can be made to vary independently and even in opposite directions simultaneously. And, third, between two fixed tones differing in pitch, a variable tone can be made to traverse different paths in pitch space—paths corresponding, for example, to a shift through height, a shift through chroma, or a shift through the circle of fifths.

## B. Multidimensional Variation of the Pitch of Synthesized Tones

The first of the three just listed facts is most simply illustrated by my original computer synthesized tones, in which all tones standing in the octave relation to each

other were made identical, thereby collapsing the (helical) continuum of pitch into the endlessly repeating circle of chroma (Shepard, 1964b). More recently, I have synthesized an analogous continuous variation around the circle of fifths, which (as portrayed in Figs. 4 and 6) may be another component of pitch. Each tone consisted of 12 equally spaced sinusoidal components per octave extending throughout the audible range with the amplitudes adjusted so that for a given tone, say C, all components in the octave relation to any C (considered as the fundamental or first harmonic for that chroma) had the highest amplitudes; the components closest to the chroma of the next harmonic with a different chroma (G) had the next highest amplitude; the components closest to the chroma of the next harmonic with yet a different chroma (E) had the next highest amplitude; and so on, according to an essentially exponential decay for succeeding harmonics that did not correspond to previously used chromas. In addition, these amplitudes were weighted, as in my original circular tones, by a fixed, bell-shaped spectral envelope that fell away to below-threshold amplitudes for low and high frequencies. Taking advantage of the great overlap, already noted, between the harmonics of tones differing by a perfect fifth, I effected the shift through a fifth simply by linear interpolation between the amplitudes of components of the same frequencies (Shepard, 1978b).

The second of the listed facts is also illustrated by my original circular tones and subsequently developed variants of them. For, by rigidly displacing, up or down in log frequency, either the bell-shaped spectral envelope with the frequencies of all the sinusoidal components held fixed or the frequencies of all the sinusoidal components with the spectral envelope held fixed, we can produce a variation either in pitch height or around the chroma circle, respectively. It appears arbitrary to maintain that only one of these two orthogonal variations is a variation in pitch when listeners are quite ready to characterize both variations in this way. Particularly arbitrary would be the claim that only the pitch-height variation of the spectral envelope is a variation in pitch, because the other, circular variation in chroma is generally more salient for a small displacement of a given size in log frequency. The possibilities inherent in this scheme for independent variation of height and chroma have been most fully demonstrated by Risset (1971, 1978) who was, for example, the first to vary these two components simultaneously in opposite directions—producing a tone that is heard as continuously ascending in pitch but is paradoxically heard, at the end, to be much lower in pitch than it was at the beginning.

The third of the listed facts is also illustrated by these sorts of computer-generated tones. In my original circular tones, the very same sequence consisting of one tone followed by a second that is on the opposite side of the chroma circle (for example, C followed by F♯) can be heard in either of two distinctly different ways—as a tritone shift up or a tritone shift down—but never as something between or as both at once (Shepard, 1964b). This implies the existence of at least two alternative paths for moving between these two tones differing in pitch. In this case the paths, though in opposite directions, are both around the chroma circle. However, the paths traversed between two tones can also be carried out over quite different dimensions of pitch. Thus, with the aid of my more recent scheme for effecting shifts around the circle of

fifths, we can continuously shift between the very same C and G, for example, in either of two ways: by passing around the chroma circle (and hence through the intermediate chromas C♯, D, D♯, E, and F) or directly between these two adjacent tones on the circle of fifths without passing through any of those intermediate chromas.

### C. Further Comments on Psychoacoustic versus Cognitive-Structural Representations

If pitch does have more than a single degree of freedom even when approached from the psychoacoustic standpoint, psychoacoustic pitch as well as musical pitch cannot be represented by a one-dimensional line without losing information about the relations and possible paths of transition between tones. (Not even topological structure is preserved under dimensional reduction.) This raises the question of what multidimensional representation would be appropriate from the standpoint of the psychoacoustic investigator, who is more interested in modeling the transducing properties of the sensory system than in modeling the cognitive-structural properties underlying musical interpretation. From what has been said about the possibility of achieving arbitrary degrees of similarity between synthesized tones separated by special musical intervals such as the octave or fifth, it appears that similar, generalized helical representations are still appropriate. The extreme cases of a one-dimensional rectilinear scale or some sort of collapsed circular or toroidal structure would then be approached, even in the absence of a musical context, as the extreme situations are arranged, in which, on the one hand, the tones are pure sinusoids and are presented successively or in which, on the other, the tones either are especially synthesized to maximize equivalence at octaves or fifths or are harmonically rich and presented simultaneously.

The purposes of a psychophysical endeavor might be best served by representations of this kind that allow for systematic distortion in which, for example, the structure is relatively more compressed at the low end of the pitch-height continuum. For these purposes, standard multidimensional scaling (whether metric or nonmetric) might be the most appropriate type of data-analytic method since, as I noted, that method allows for deviations of this kind from a completely regular structure. By contrast, for the fundamentally different purposes of representing the higher-level structure presumed to underlie musical competence, I have argued that we want the regularity that will ensure invariance under transposition. And, in this case, data-analytic methods based on some variant of Euclidean composition seem more suitable.

From the psychoacoustic standpoint, what is of interest about musical tones is how the representation of those tones changes with continuous variation in the parameters of the physical stimuli. From this standpoint one is concerned with such questions as (a) whether what is perceived as the best octave corresponds to a strict 2-to-1 frequency ratio or to a somewhat larger ratio, (b) whether the favored ratio systematically changes at low or at high frequencies, and (c) whether other standard musical intervals

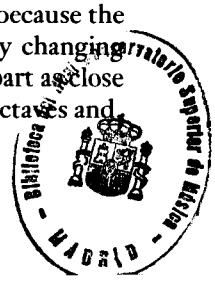
are most preferred when they correspond to steps of equal temperament (and, so, are designed to preserve equal steps in log frequency) or some scheme of so-called just tuning (and, thus, are designed to maintain numerically simple frequency ratios). However, from the cognitive–structural standpoint these psychophysical concerns are of little importance. Once the tones are categorically mapped into the discrete nodes of an internal representation that is functionally regular, it is the structural properties inherent in that regular representation that are important. Thus, the fact that all tones in octave relations to a given tone are separated by multiples of a fixed distance (and, perhaps, are colinear as well) is essential; but whether the frequency ratio that corresponds, physically, to that fixed psychological unit of octave distance is equal to or greater than 2-to-1 or, indeed, whether it changes somewhat toward low or high frequencies is no longer relevant. For the same reason, from this cognitive–structural standpoint, the long-standing psychophysical concern over the fact that it is mathematically impossible to construct a scale in which all the intervals exactly realize simple frequency ratios (e.g., see the appendix in Westergaard, 1975) is quite irrelevant also.

## D. Alternative Cognitive-Structural Representations for Musical Pitch

Even for the purposes of representing the cognitive structure of musical pitch, alternative representations are possible. This is because different representations that preserve the topological structure can be adopted as long as we compensate for any information-preserving change in the representation by an appropriate change in the rules taken to relate the representation to the data (cf., Shepard & Carroll, 1966; and for a related tradeoff between representation and process in cognitive theory, see Anderson, 1978; Shepard, 1981b).

In the present connection we can use the simple helix (Fig. 1) and the double helix (Fig. 3) to illustrate two alternative ways in which an increased similarity at any particular interval—say the octave—might be accounted for in terms of the spatial representation. In the simple helix, all tones standing in the octave relation to each other have two special geometrical relationships within the spatial representation: such tones are closer to each other than they would be without the deformation into the regular helix, and they are also colinear. In the double helix of Fig. 3, however, the tones standing in the octave relation to each other have only one special geometrical property in common: they are colinear. This is why, in attempting to represent the octave relation solely in terms of proximity, we had to bend the double helix (Fig. 3) around to form a higher order helix (Fig. 5) or, in the limiting case of complete octave equivalence, a toroidal structure (Fig. 4).

So far I have made use only of the relation of proximity. In part this is because the relation of proximity has the advantage of being continuously variable; by changing the degree of the compression of the helix, we can bring tones an octave apart as close together as we choose (until, in the limit, we achieve complete identity of octaves and


Biblioteca - Patrimonio Superior de Música - MADRID

hence, complete degeneration of the helix into the chroma circle or torus). By contrast, the relation of colinearity is all or none: either the tones in a particular set fall on a certain line or they do not.

Still there is evidence to support the notion of Revesz (1954) and Bachem (1950) that (at least within, again, a suitable musical range) musical listeners hear tones differing by multiple octaves as having some discrete feature (or chroma) in common regardless of the number of intervening octaves (e.g., see Dowling & Hollomb, 1977; Idson & Massaro, 1978; Kallman & Massaro, 1979; Thurlow & Erchul, 1977; and, for comparison, Deutsch, 1972). This is quite analogous to the fact, in the visual domain, that (for example) bars at a certain angle share the analyzable property of that certain orientation, regardless of large differences in their sizes (Shepard, 1964a). In the present, auditory case this suggests that chroma may act as an attribute of tones that is "analyzable" (Shepard, 1964a) or "separable" (Garner, 1970). If so, there is good reason for representing the perceptual equivalence of all tones in the multiple-octave relation to a particular tone by a coincidence of the corresponding points with a certain chroma line regardless of their distance of separation in the spatial model. Such coincidence could also provide for a sharper tuning to the octave relation than could plausibly be provided purely by proximity in the helix (where increased closeness at the octave must imply increased closeness at the neighboring major seventh and minor ninth as well). Only further study will reveal whether all special pitch relations can adequately be represented solely by spatial proximity or whether the extent and sharpness of tuning to other special relations will require the use of other geometrical relations such as colinearity as well.

Possibly, different forms of representation will prove to be more suited for bringing out the relations between pitches that are most salient in different situations—such as those of simultaneous presentation (harmony), immediate succession (melody), or longer-range shifts of tonality (modulation of key). Moreover, even when the alternative forms of representation are essentially isomorphic, some forms may be more accessible to visual grasp by the human mind. As a case in point, the difficulty of picturing structures in spaces of more than three dimensions, such as the structures depicted in Figs. 4 and 5, motivates the search for lower dimensional "maps" of those structures that preserve, if not all the relations, at least the relations that are most important for particular purposes.

### E. The Two-dimensional Melodic Map

In each of the structures portrayed in Figs. 3, 4, and 5, one octave is represented by a double helix wound around a two-dimensional surface that is, alternatively, a straight cylinder in three dimensions (Fig. 3), a cylinder curved into a closed torus in four dimensions (Fig. 4), or the same torus cut through again and twisted apart into an orthogonal, fifth dimension (Fig. 5). The surface in which the double helix is embedded in each case, being itself two-dimensional, has a flat "topological map" (Blackett, 1967) that is obtained by (a) cutting out a one-octave segment of the tube, (b)

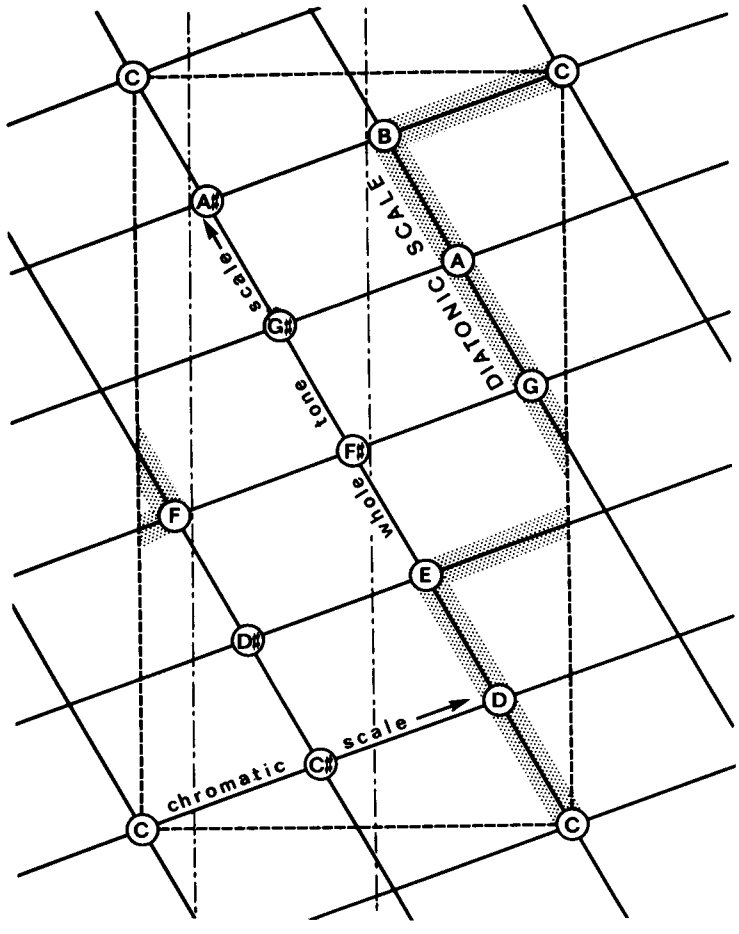



**Fig. 8.** The two-dimensional melodic map obtained by unwrapping the double helical structure of Figs. 3, 4, and 5. (The structure of a diatonic scale is illustrated for the key of C major.)

cutting along the length of the straightened one-octave tube (whether it was initially straight, curved into itself, or twisted), and (b) opening the cylindrical surface out into a flat rectangle, like that illustrated in Fig. 8. The fact that the original surface was unbounded in each case can be represented in the topological map by replicating the rectangle endlessly in all directions within the plane, as indicated in the figure. The two-dimensional map contains an infinite lattice in which adjacent nodes, representing the tones, are related by minor seconds (semitones) running in one diagonal direction, and by major seconds (whole tones) running in the other diagonal direction. (The whole-tone scales, running in the latter direction, correspond to the two opposite edges of the original double helix.)

The fact that in the original curved representations (of Figs. 3, 4, and 5), tones in a

particular diatonic key are linearly separable by a hyperplane with closely related keys obtainable by small rotations of the hyperplane has a direct analog in this two-dimensional map: Tones in a particular key fall within a narrow vertical band (demarcated for the key of C major by the vertical broken lines in Fig. 8), and the most closely related keys are obtained by the smallest horizontal shifts of this band. Notice that the tones falling within such a band always form a characteristic zig-zag pattern (illustrated by the stippled strip for the key of C major). Indeed any rigid translational motion of this zig-zag pattern into another set of nodes—whether horizontally, vertically, or diagonally—will again fall in such a band and will therefore yield another diatonic key. Moreover, the relative minor and all the so-called church modes are also represented by this same pattern and differ from each other only with regard to which of the seven distinct nodes included within each octave of the key is designated as the primary focal tone or tonic. Finally, the most common pentatonic scales are represented by the very similar five-tone zig-zag patterns that form the complements of these seven-tone diatonic sets. In the following subsection, I shall return to a consideration of the musical significance of the fact that the most common musical scales have these particular zig-zag structures. For the moment, it is sufficient to notice that all such scales, as well as the chromatic and whole-tone scales, have compact, orderly representations in this space.

Not surprisingly, since they are constructed within the frameworks of such scales, tunes or melodies also have generally compact representations within this lattice. Indeed, quite apart from any consideration of the underlying scales, statistical analyses of the melodies occurring in popular, folk, and classical music have substantiated the generalization that most intervals between successive tones in such melodies are minor or major seconds with larger intervals relatively less common (Dowling, 1978, p. 352; Fucks, 1962; Merriam, 1964; Philippot, 1970, p. 86). The reason for the preponderance of small melodic intervals seems to be that the perceptual integration of successive tones into a coherent unit depends especially strongly on proximity in pitch height (Bregman & Campbell, 1971; Deutsch, 1978; Dowling, 1973a; Jones, 1976; McAdams & Bregman, 1979; Van Noorden, 1975)—a dependence that can be seen as a manifestation of a much more general principle of proportionality between time and distance known, in connection with visual apparent motion, as "Korte's Third Law" (see Shepard, 1981b). In any case, because adjacent nodes in the two-dimensional lattice of Fig. 8 are separated by the predominant intervals in melodies (minor and major seconds), I have proposed to refer to this two-dimensional topological map of the double helix as *melodic space* (Shepard, 1978b, 1980a, 1981a,b).

### F. The Two-dimensional Harmonic Map

Although tones that differ by minor or major seconds typically follow each other in melodic sequences, they are dissonant when sounded simultaneously. Accordingly, the harmonic intervals that make up the most common chords are not the dissonant minor and major seconds but the larger, consonant intervals of the minor and major

thirds. As a consequence, whereas melodic sequences are compact in the two-dimensional lattice of Fig. 8, chords tend to be spread out in that representation. The major and minor triads, for example, form two triangles with a common longer side (corresponding to the perfect fifth) and hence together form an elongated parallelogram in Fig. 8. From the C in the lower left corner of the rectangular map for one octave in that figure, this parallelogram extends diagonally up to the G, the perfect fifth that is to the right and above that C. The other two vertices of that parallelogram are formed by E, for the major third, and E♭ (labeled D♯), for the minor third.

However, because no other standard musical tone falls within this parallelogram, all such elongated parallelograms can be compressed into compact, more nearly square configurations by a single affine transformation consisting of a linear compression of the whole two-dimensional space along the long (C-to-G) axis of the parallelogram considered—with, if desired, a compensating expansion in some other, more-or-less orthogonal direction. This affine transformation of the two-dimensional map corresponds, in the torus, to the transformation achieved (a) by cutting through the torus in the direction orthogonal to the cut needed to open the torus back into a one-octave cylinder, and (b) by giving the resulting cylinder two complete torsional twists, and (c) by reattaching the torsionally twisted ends to form an altered torus. The resulting, affinely transformed two-dimensional map is displayed in Fig. 9. As is illustrated for the key of C, the set of other tones that are consonant with the tonic of any key (represented here by the double-circled C) does now have compact representations. The same is true for any of the most common chords, whether major or minor, augmented or diminished. Thus, although none of the structures incorporating the double helix (Figs. 3, 4, 5, and 8) were explicitly based on musical thirds, those structures are nevertheless related by a simple transformation to a structure that

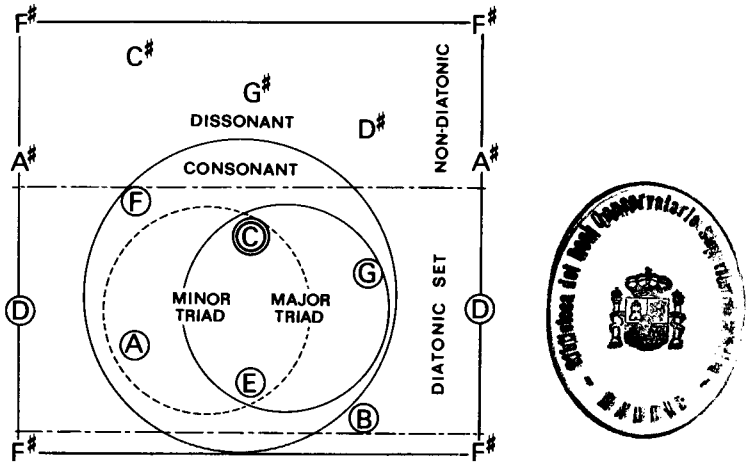



**Fig. 9.** The two-dimensional harmonic map obtained by an affine transformation of the melodic map. (Relations of consonance and a diatonic set are illustrated for an arbitrarily selected tone, C.)

does emphasize thirds—including the major third, which according to Table I does play a significant role in the perception of musical pitch.

Spaces that are essentially isomorphic to this harmonic space have previously (and often independently) been proposed by many theorists, including Balzano (1978, 1980), Hall (1974), Lakner (1960), Longuet-Higgins (1962, 1979), and O'Connell (1962). What I presume to be new, here, is the demonstration of a simple affinity (in the literal mathematical sense) between this space and the melodic space (Fig. 8) that I have derived from my double helical representation for pitch. In an especially promising development, Krumhansl and Kessler (1982), extending the probe tone technique of Krumhansl and Shepard (1979), have now shown that essentially this toroidal structure can also represent the set of major and minor tonalities, with chord progressions moving the listener's momentary tonal center around on the two-dimensional surface.

I am not claiming that the different components or alternative representations for pitch that I have proposed here or the different musical intervals that each of these emphasizes, all operate on the same level of the hierarchically organized perceptual system. On the contrary, the ways in which individual listeners differ in their relative sensitivities to these components (see Fig. 7 and Table I) suggest that the smallest intervals, the thirds, and the perfect fifths may operate at three different, increasingly cognitively structured levels. By analogical extension of my earlier remarks about proportionality of time and distance, we might conjecture that the time constants should be greater for successively higher, more complexly structured levels of the system (cf. Powers, Clark, & McFarland, 1960). Such a conjecture appears to be consonant with the observation that I, as well as some of my colleagues (Balzano, 1980; Krumhansl, Bharucha, & Kessler, 1981), have recently made; namely, that in tonal music, shifts from one tone to the next of a melody are usually the most rapid, followed by somewhat slower shifts in harmonic accompaniment with modulations of key or tonal center occurring, when they do, at a still slower rate.

### G. Structural Properties of Musical Scales

The asymmetrical zig-zag character that the principal musical scales have in the two-dimensional melodic space of Fig. 8 is fundamental in tonal music. The alternating pattern of three and four whole-tone steps between semitone shifts (3-4-3-4-3-4- . . . ) that characterizes the diatonic scales and the similar alternating pattern of two and three whole-tone steps between one-and-a-half-tone shifts (2-3-2-3-2-3- . . . ) that characterizes the pentatonic scales have the essential property that each scale degree has a unique relation (in terms of numbers of intervening semitone steps) to all other tones in that scale. As has been emphasized particularly by Zuckerkandl (1956, 1972) and, recently, by Balzano (1978, 1980), it is this property that confers on each scale degree its unique "dynamic quality" (Zuckerkandl, 1956, 1972) or "tonal function" (Meyer, 1956; Piston, 1941; Ratner, 1962) and, hence, that enables the listener to have, at every moment, a clear sense of where the music is with respect to a fixed tonal framework. Only with respect to such a framework can there be such things as motion

or rest, tension and resolution, or, in short, the underlying dynamisms of tonal music. By contrast, the complete symmetry and regularity of the chromatic and whole-tone scales means that every tone has the same status as every other. The fact that for such scales there can be no clear sense of location and, hence, of motion is, I believe, the reason that such scales have never enjoyed wide or sustained popularity as a basis for music.

The idea is often expressed, whether explicitly or by implication, that the diatonic as well, perhaps, as the complementary pentatonic scales, which are so pervasive in the music of the Western world are arbitrary and probably rather recent cultural conventions and that in Eastern cultures scales have developed with entirely different structures and even with octaves divided into quite different and sometimes much larger numbers of steps than the traditional 12 of Western music. However, this idea should be evaluated in view of the following points.

First, as Balzano (1978, 1980) in particular has shown, quite apart from the psychoacoustic considerations usually invoked to justify the diatonic scale (such as that it permits an especially good approximation to the simplest frequency ratio for the perfect fifth), there are deep cognitive-structural justifications of a group-theoretic character both for the division of the octave into 12 parts and for the selection from the resulting set of tones of the diatonic (or the pentatonic) subsets of seven (or five) tones to constitute a particular scale. One such justification is the already noted achievement, in these particular asymmetric subsets of the unique tonal function of each tone corresponding to its unique structural relation to the contextually established scale (also see Rothenberg, 1978a,b).

Second, far from being a recent Western invention, the diatonic scale has in fact been traced back to the most ancient tuning systems so far deciphered from archaeological records (Kilmer, Crocker, & Brown, 1976).

Third, although different tuning systems continue to be used in other cultures and although some (e.g., Indian and Indonesian) scales require the division of each octave into as many as 22 parts, a subset of five or seven focal tones, corresponding to something more or less akin to a Western pentatonic or diatonic scale, usually plays the central role in a particular piece of music—with the additional, microtonal steps used more for subtle melodic decorations or ornamentations around these focal tones (e.g., see Dowling, 1978, in press). Indeed, just as may be the case for different natural languages (Chomsky, 1965), what is more significant about the scales used in different cultures may be their deep structural similarities rather than their superficial differences:

1. Virtually all such scales are based on the octave (whether or not it corresponds to an exactly 2-to-1 frequency ratio) with every octave divided up in the same way (Dowling, 1978, in press).
2. Nearly all assign a central role to the so-called perfect intervals, the fifth and its inversion in the octave, the fourth (Sachs, 1943).
3. And as I have already noted, most select a subset of either five or seven tones from each octave as the principal or focal tones, and these focal tones quite generally

have a definite, irregular or asymmetric spacing much like that in the diatonic or pentatonic scales in order, I believe, to confer a unique status on each tone of the scale.

Finally, and presumably for the same deep group-theoretic reasons, in some (e.g., West African) cultures in which rhythmic structure has evolved a complexity comparable to that of Western tonal structure, the basis of rhythmic variations has recently been seen to be isomorphic, in the time domain, to the diatonic system in the pitch domain (Pressing, 1979).

Within Western tonal music, incidentally, the hierarchical organization of the levels of musical interpretation noted in the preceding section—in addition to having different associated time constants—contributes different and sometimes recursive levels of dynamism to the music. Thus, even after a melody has itself come to a stable resting point, the harmony may maintain tension by delaying resolution to the tonic; and even after the harmony and the melody have both subsequently resolved to the locally defined tonic, an implicit tension may remain at a deeper level until that locally established tonic finally modulates back to the initially instated global tonic of the piece as a whole. The forward momentum of music (as of life, generally) is thus maintained by a hierarchy of subgoals within subgoals (cf. Hofstadter, 1979, p. 612).

### H. Pitch as a Morphophoric Medium

I believe that there is a fundamental reason why pitch and time, among the potentially variable attributes of tones (including such others as loudness, timbre, and perceived spatial location of source), have always played the most essential roles in music. The two attributes of pitch and time appear to be unique in the auditory modality in the extent to which they are capable of bearing form or, to adopt Attneave's term, are "morphophoric" (see Attneave, 1972; Attneave & Olson, 1971). The essential property of the continua of pitch and time is that musical objects—namely, melodies and chords—preserve their perceived structural identities under rigid transformations of transposition in pitch or displacement in time—just as visual objects preserve their structural identities of shape under rigid transformations of translation and rotation. At present, there is little evidence that perceived patterns of loudnesses, timbres, or spatial locations (any more than perceived patterns of colors in the visual modality) will preserve their structural identities to the same extent under transformation in those continua.

The distinction appears to be intimately connected with the distinction that Kubovy (1981) has recently drawn between "dispensable" and "indispensable" attributes. According to this distinction, time is an indispensable attribute in both the visual and the auditory modalities; but whereas space is a second indispensable attribute in the case of vision, it is pitch and not literal space that is the second indispensable attribute in the case of audition. To use a variant of the example by means of which Kubovy illustrates these concepts: if two projected spots of light, one red and one blue, are superimposed on each other in the same spatial location, they are perceived

to merge into a single spot, purple in color; hence, the original difference in spatial location of the two spots was indispensable to their separate identities. If, however, the two spatially separated spots are made the same color, purple, they will still be perceived to be two in number; hence, the difference in color was dispensable in Kubovy's sense. The situation is the other way around in the auditory domain. If two sound sources, one emitting middle C and the other emitting E, are superimposed in space, the two tones will still retain their individual identities as two tones separated by a major third; hence, the difference in spatial location was dispensable. If, however, the two spatially separated tones are brought to the same pitch, say the intermediate D, they will be heard to be a single tone of that pitch centered between the two sources; hence, the difference in pitch was indispensable.

## I. The Problem of the Representation of Pitch within a Tonal Context

I have claimed that the context of a particular musical key (whether explicitly presented or only implied by a musical passage) induces an internal cognitive framework or hierarchy of tonal functions within a musical listener and that this cognitive framework plays, in turn, an essential role in the perceptual interpretation of each succeeding tone or chord. But how is the dependence of this contextually induced cognitive framework on a particular key to be represented with respect to the underlying geometrical structures that I have proposed for musical pitch? Throughout this chapter I have argued for an invariant underlying representation in which all tones separated by the same musical interval correspond to points separated by the same spatial distance and, hence, a representation that maps into itself under rigid transformations. Principally, the contrast that I have emphasized in this connection has been with alternative psychoacoustic proposals that the perceptual relations corresponding to particular intervals might systematically change toward the low or the high ends of the continuum of audible frequencies. The departures from a strictly regular representation that are suggested by considerations of cognitive interpretation within a tonal context are quite different. Departures of this latter sort would not accumulate over global shifts in pitch height; rather, they would operate differentially at the local level in exactly the same way within each octave.

Such departures, unlike those of the former, psychoacoustic kind, necessarily imply that the psychoacoustic goal, pursued since the time of Helmholtz, of accounting for the perceived musical relations between tones entirely in terms of the physical properties of the two tones making up that interval can never be achieved. The perception of musical intervals would then depend, as well, on the cognitive framework within which those physical stimuli are categorically interpreted by a musical listener.

At the same time, such departures appear to be inconsistent with the kind of invariant cognitive-structural representation proposed here—that is, one in which all tones separated by the same interval have the same distance of separation in the

underlying structure. Accordingly, we need to consider some of the kinds of evidence for—and implications of—these apparent departures from strict structural regularity and symmetry.

First, the most easily disposed of, is one phenomenon that I already claimed to hold, at least for some musical listeners; I mean that in which the successive steps of a diatonic scale—do, re, mi, etc.—can be heard as in some sense equal, or at least more equal than one would expect from the fact that two of these steps (namely, mi-fa and ti-do), being semitone intervals, are just half the physical size of the remaining (whole-tone) steps (Fig. 8). For as I have shown elsewhere (Shepard, 1978b, 1982), there exists a relative weighting of the circle of fifths (Component F in Fig. 2) such that the distances corresponding both to minor and to major seconds are all equal to each other in the resulting combined structure. Listeners who do not claim to hear the successive steps of a diatonic scale as equal, then, would presumably be those characterized by a different weighting of the circle of fifths relative to the chroma circle and to height. (See the individual variation in such weights indicated in Fig. 7.)

Of greater concern is another cognitive phenomenon, which is already well known to students of music theory (see, for example, illustrations discussed by Rosen, 1971). Here, I expand on the example given by Risset (1978, p. 526); the same physical interval, whether it is called F-B or F♯-B, may be differently perceived in different contexts. In C-major it is the upper tone in the interval that is interpreted as the leading tone, and, hence, the interval is perceived as tending to resolve to the interval E-C. In F♯-major, however, it is the lower tone that is interpreted as the leading tone, and, hence, the interval is in this case perceived as tending to resolve to the different interval F♯-A♯. Correspondingly, musicians may actually execute such an interval differently in the two cases (e.g., see Shackford, 1961, 1962; Small, 1936). Indeed, in experiments with listeners in the psychological laboratory, Krumhansl (1979) has now provided evidence that the perceived relation between tones separated by the very same interval depends systematically on the relation between that interval and the contextually established tonic. Her listeners judged the relation between two tones separated by a particular interval to be greatest when the two tones were selected from the tonality-defining major triad, next greatest when the two tones were selected from other tones that still belonged to the associated key, and least when the two tones were selected from the remaining tones not belonging to that key.

Clearly, Krumhansl's findings pose a major challenge to the psychoacoustic goal of accounting for perceived relations between tones solely in terms of physical properties of those tones. From the present standpoint, however, we should perhaps be cognizant of the possibility that the judged relatedness of tones in a contextually established key might be a reflection of the differential saliencies of the individual tones relative to that key rather than a reflection of a true alteration of the perceived distances between those tones. As Tversky (1977) has shown, in comparison with one pair of stimuli, a pair of perceptually richer or more salient stimuli may be judged to be more similar if the judgment is in terms of similarity, but also more different if the judgment is in terms of difference. (Also see Krumhansl, 1978.) Thus, in order to substantiate the

claim that the minor third is actually represented by a smaller cognitive distance when that minor third is a part of the contextually established major triad (that is, when it is E-G in the key of C major) than when it falls entirely outside that key (e.g., when it is D♯-F♯), it may be necessary to show that this same result is obtained whether the judgments are of similarity or of dissimilarity.

I presume, in any case, that the multidimensional scaling results obtained by Krumhansl (1979; also see Krumhansl *et al.*, 1980) are subject to the end-distortion that I have already said is to be expected when data are scaled for a single octave only. Neither the two- nor the three-dimensional solutions obtained by Krumhansl, as they stand, can be strictly prolonged into additional octaves. Just as in the case of the left panel of the present Fig. 6, however, a slight torsional distortion into an additional dimension will permit the structures to be extended in this way, and, indeed, they then become helical, as we should expect. The precise form that such a helix might ultimately take and, in particular, the extent to which it might correspond to any of the completely regular forms illustrated in the present Figs. 3, 4, and 5 would depend, in part, on the resolution of the similarity–dissimilarity issue just raised.

No matter how that particular issue is eventually resolved, a prevailing tonal framework nevertheless does appear to exert a powerful influence on the interpretation of tones within a musical context. Thus, even the data of Krumhansl and Shepard (1979), which the present Figs. 6 and 7 and Table I have already shown to provide support for the proposed geometrical representations of pitch at the global level, manifest systematic deviations from those structures at the local level. The most striking and consistent pattern in these deviations is this: As the condition of complete octave equivalence is approached—as it is for the most musical listeners (see Fig. 7)—pairs of intervals that are related to each other by an inversion in the octave (namely, m2 and M7, M2 and m7, m3 and M6, and M3 and m6 in Fig. 2) do not become equivalent as predicted by the geometrical model. Instead, some of these intervals—particularly M2 and M3—strongly dominate their inversions even in the limit of complete octave equivalence (see Fig. 3 in Krumhansl & Shepard, 1979).

As of this writing I can think of three possible ways in which these various tonal asymmetries may eventually prove to be reconcilable with the kinds of regular representations for pitch proposed here.

First, as I have already suggested, some of the asymmetries might be handled, without altering the underlying distances between the tones, simply by assigning to each individual tone a weight or salience representing how central that tone is in the prevailing tonal context. Perceived relations between tones, including asymmetries of these relations, would then be explained in terms of these weights operating in conjunction with the fixed underlying distances—perhaps in a manner similar to that proposed by Tversky (1977).

Second, one could go further and propose that a tonal context does invoke a fixed, asymmetric internal representation of the corresponding diatonic scale and its associated hierarchy of tonal functions quite independent of the underlying regular representation of pitch considered in abstraction from any particular musical key.



Under the prevailing tonal context, then, each succeeding tone would be mapped from a node in the underlying regular structure into the closest node in the fixed, asymmetric representation for that diatonic key. Hence, although all tones related by the same interval would be separated by the same distance in the former, helical structure, such tones might be separated by different and even asymmetric distances in the latter structure—depending on the tonal function of each tone in that tonal structure. Very generally, some analogy between melodies and sentences might even be discerned here. That is, somewhat as each noun phrase in a sentence is represented, at a deep level, as having a certain function of grammar (subject, direct object, indirect object, etc.) or of case (agentive, instrumental, locative, etc.—see Fillmore 1970, 1977), each tone in a melody may be represented as having a definite tonal function (tonic, dominant, subdominant, leading tone, etc.).

Third, the mapping of the underlying abstract, invariant structure into the particular asymmetric structure for a particular tonality might in some way be regarded, still more specifically as a kind of "projection." Thus, much as the concrete visual representation of a three-dimensional object must always be from a particular spatial point of view (Metzler & Shepard, 1974; Shepard, 1981b), the concrete musical representation of the underlying abstract helical structure of pitch must always be from a particular tonal point of view. Although the precise rules that might govern such projections remain unspecified, consideration of the structures shown in Figs. 3, 4, and 5 has indicated that change of viewpoint (i.e., change of tonality) corresponds—just as in the visual case—to a rotation (see Shepard & Cooper, 1982, Ch. 16).

## VI. CONCLUSIONS

1. Pitch is multidimensional. A one-dimensional rectilinear representation of pitch is incompatible with the facts (a) that under appropriately musical conditions, tones separated by certain intervals (especially the octave and perfect fifth) are perceived to have heightened musical relations and (b) that between two tones differing in perceived pitch a synthesized tone can be continuously varied from one to the other over distinctly different paths (in pitch height, in chroma, or even around the circle of fifths). The heightened perceptual relations between tones separated by these special musical intervals implies a generalized helical structure for pitch. In order to fit empirical data for more than a single octave, an embedding space of at least five dimensions will be required for this structure (including one dimension for a rectilinear component of pitch height and two dimensions for each of two circular components: the chroma circle and the circle of fifths).

2. For psychophysical purposes of characterizing the relations between musical tones as they are transduced by the sensory system, this generalized helical structure may be allowed to assume a globally distorted form in which, for example, the component of height is relatively compressed toward the low end of the range of audible pitch. In this way we can represent such facts about the ear as that it is less effective in discriminating differences in pitch at low frequencies. In order to extract

such a globally distorted helical structure from empirical data, multidimensional scaling appears to provide a potentially useful tool.

3. For purposes of characterizing the cognitive structure of pitch as interpreted within a musical setting, however, the helical structure should have exactly the same conformation within each octave. Only within such a regular representation will musical objects—principally, melodies and chords—retain their unique structural identities under transposition. In order to extract such a regular structure from empirical data, a different method of Euclidean composition appears to be more promising.

4. Still remaining to be fully elucidated are the rules of projection that map such an abstract, symmetrical structure into the more concrete and perhaps asymmetrical structure representing the perceived relations between tones within the context of a particular musical key—between, that is, the unique tonal functions (tonic, dominant, subdominant, leading tone, etc.) relative to that key.

## ACKNOWLEDGMENTS

This chapter is based on research supported by National Science Foundation Grants BNS 75-02806 and BNS 80-05517 and by a Sabbatical Award from the James McKeen Cattell Fund. The chapter represents in part a recasting of ideas developed in a previously written paper (Shepard, 1982). I undertook this recasting in June 1980, while I was a visiting scientist at the *Institute de Recherche et Coordination Acoustic/Musique* (*IRCAM*) of the *Centre Georges Pompidou*, Paris. My interest in the possibility of regular geometrical representations for the structure of musical pitch, dormant since 1964 (when I had explored the single helix for pitch), was reawakened by a demonstration by Gerald Balzano, while he was still a graduate student at Stanford, of the group-theoretic basis of the diatonic system. My own geometrical developments, presented here, have benefited from the ideas and information contributed by a number of colleagues and associates both in the United States and in France, including (in addition to Balzano) Carol Krumhansl, Michael Kubovy, Stephen McAdams, Jean-Claude Risset, and David Wessel. In addition, I am indebted to Shelley Hurwitz for technical assistance in the analyses of the data and to Teresa Putnam for her part in the preparation of the final manuscript.

## REFERENCES

Abelson, R. P., & Tukey, J. W. Efficient utilization of nonnumerical information in quantitative analysis: General theory and the case of simple order. *Annals of Mathematical Statistics*, 1963, *34*, 1347–1369.

Allen, D. Octave discriminability of musical and non-musical subjects. *Psychonomic Science*, 1967, 7, 421–422.

Anderson, J. R. Arguments concerning representations for mental imagery. *Psychological Review*, 1978, *85*, 249–277.

Attneave, F. The representation of physical space. In A. W. Melton and E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C.: Winston, 1972. Pp. 283–306.

Attneave, F., & Olson, R. K. Pitch as a medium: A new approach to psychophysical scaling. *American Journal of Psychology*, 1971, *84*, 147–165.

Bachem, A. Tone height and tone chroma as two different pitch qualities. *Acta Psychologica*, 1950, 7, 80–88.

Bachem, A. Absolute pitch. *Journal of the Acoustical Society of America*, 1955, *27*, 1180–1185.

Balzano, G. J. Chronometric studies of the musical interval sense. Doctoral dissertation, Stanford University, 1977. *Dissertation Abstracts International*, 1977, *38*, 2898B (University Microfilms No. 77-25, 643).

Balzano, G. J. The structural uniqueness of the diatonic order. In R. N. Shepard (Chair), *Cognitive*

*structure of musical pitch*. Symposium presented at the meeting of the Western Psychological Association, San Francisco, California, April 1978.

Balzano, G. J. The group-theoretic description of twelvefold and microtonal pitch systems. *Computer Music Journal*, 1980, *4*, 66-84.

Balzano, G. J. Musical versus psychoacoustical variables and their influence on the perception of musical intervals. *Bulletin of the Council for Research in Music Education*, 1981. (In press).

Balzano, G. J., & Liesch, B. W. The role of chroma and scalestep in the recognition of musical intervals in and out of context. *Psychomusicology*. (In press).

Bartlett, J. C., & Dowling, W. J. Recognition of transposed melodies: A key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance*. 1980, *6*, 501-515.

Beck, J., & Shaw, W. A. The scaling of pitch by the method of magnitude estimation. *American Journal of Psychology*, 1961, *74*, 242-251.

Beck, J., & Shaw, W. A. Single estimates of pitch magnitude. *Journal of the Acoustical Society of America*, 1963, *35*, 1722-1724.

Bentler, P. M., & Weeks, D. G. Restricted multidimensional scaling models. *Journal of Mathematical Psychology*, 1978, *17*, 138-151.

Blackett, D. W. *Elementary topology*. New York: Academic Press, 1967.

Blechner, M. J. Musical skill and the categorical perception of harmonic mode. Status Report on Speech Perception SR-51/52. New Haven, Connecticut: Haskins Laboratories, 1977. Pp. 139-174.

Blumenthal, L. M. Distance geometries. *University of Missouri Studies*, 1938, *13*, No. 2.

Bregman, A. S., & Campbell, J. Primary auditory stream segregation and the perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 1971, *89*, 244-249.

Burns, E. M. Octave adjustment by non-Western musicians. *Journal of the Acoustical Society of America*. 1974, *56*, 525. (Abstract)

Burns, E. M., & Ward, W. I. Categorical perception—phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America*, 1978, *63*, 456-468.

Butler, J. W., & Daston, P. G. Musical consonance as musical preference: A cross cultural study. *Journal of General Psychology*, 1968, *79*, 129-142.

Carroll, J. D., & Chang, J.-J. Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart-Young decomposition. *Psychometrika*, 1970, *35*, 283-319.

Chomsky, N. *Aspects of the theory of syntax*. Cambridge, Massachusetts: MIT Press, 1965.

Cohen, A. J. Inferred sets of pitches in melodic perception. In R. Shepard (Chairman), *Cognitive structure of musical pitch*. Symposium presented at the meeting of the Western Psychological Association, San Francisco, California, April 1978.

Coxeter, H.S.M. *Introduction to geometry*. New York: Wiley, 1961.

Cuddy, L. L., Cohen, A. J., & Miller, J. Melody recognition: The experimental application of musical rules. *Canadian Journal of Psychology*, 1979, *33*, 148-157.

Davies, J. B. *The psychology of music*. London: Hutchinson, 1978.

de Boer, E. On the "residue" and auditory pitch perception. In W. D. Keidel & W. D. Neff (Eds.), *Handbook of sensory physiology* (Vol. V/3, *Auditory system: Clinical and special topics*). New York: Springer-Verlag, 1976. Pp. 479-583.

de Leeuw, J., & Heiser, W. Multidimensional scaling with restrictions on the configuration. In P. R. Krishnaiah (Ed.), *Multivariate Analysis* (Volume 5). Amsterdam: North Holland Publ., 1979.

Deutsch, D. Octave generalization and tune recognition. *Perception & Psychophysics*, 1972, *11*, 411-412.

Deutsch, D. Delayed pitch comparisons and the principle of proximity. *Perception & Psychophysics*, 1978, *23*, 227-230.

Dewar, K. M., Cuddy, L. L., & Mewhort, D.J.K. Recognition memory for single tones with and without context. *Journal of Experimental Psychology: Human Learning and Memory*. 1977, *3*, 60-67.

Dowling, W. J. The perception of interleaved melodies. *Cognitive Psychology*, 1973, *5*, 322-337. (a)

Dowling, W. J. The 1215-cent octave: Convergence of Western and Nonwestern data on pitch-scaling. *Journal of the Acoustical Society of America*, 1973, *53*, 373A. (Abstract) (b)

Dowling, W. J. Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, 1978, *85*, 341-354.

Dowling, W. J. Musical scales and psychophysical scales: Their psychological reality. In T. Rice & R. Falck (Eds.) *Crosscultural approaches to music*. Toronto: University of Toronto Press. (In press)

Dowling, W. J., & Hollombe, A. W. The perception of melodies distorted by splitting into several octaves: Effects of increasing proximity and melodic contour. *Perception & Psychophysics*, 1977, *21*, 60-64.

Drobish, M. W. *Uber musikalische Tonbestimmung und Temperatur. Abhandl Math. Phys. Kl. Konigl. Sachs, Ges. Wiss.* 1855, *4*, 1-120.

Elfner, L. Systematic shifts in the judgment of octaves of high frequencies. *Journal of the Acoustical Society of America*, 1964, *36*, 270-276.

Fechner, G. *Elements of psychophysics*. New York: Holt, 1966. (Original publication in German, 1860)

Fillmore, C. J. The case for case. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory*. New York: Holt, 1970. Pp. 1-90.

Fillmore, C. J. The case for case reopened. In P. Cole & J. Sadock (Eds.), *Syntax and Semantics*. (Volume 8, Grammatical relations) New York: Academic Press, 1977. Pp. 59-81.

Francès, R. *La perception de la musique*. Paris: Vrin, 1958.

Fucks, W. Mathematical analysis of formal structure of music. *IRE Transactions on Information Theory*, 1962, *8*, 225-228.

Fyda, M. C. Perception and Internal Representation of Musical Intervals. M. A. Thesis, Michigan State University, East Lansing, Michigan, 1975.

Garner, W. R. The stimulus in information processing. *American Psychologist*, 1970, *25*, 350-358.

Goldman, R. F. *Harmony in Western music*. New York: Norton, 1965.

Goldstein, H. *Classical mechanics*. Reading, Massachusetts: Addison-Wesley, 1950.

Goldstein, J. L. An optimum processor theory for the central formation of the pitch of complex tones. *Journal of the Acoustical Society of America*, 1973, *54*, 1496-1516.

Greenwood, G. D. *Principles of dynamics*. Englewood Cliffs, New Jersey: Prentice-Hall, 1965.

Guttman, N., & Pruzansky, S. Lower limits of pitch and musical pitch. *Journal of Speech and Hearing Research*, 1962, *5*, 207-214.

Hall, D. E. Quantitative evaluation of musical scale tunings. *American Journal of Physics*, 1974, *42*, 543-552.

Hilbert, D., & Cohen-Vossen, S. *Geometry and the Imagination*. New York: Chelsea, 1952. (Original publication in German, 1932)

Hofstadter, D. R. *Gödel, Escher, Bach: An eternal golden braid*. New York: Basic Books, 1979.

Idson, W. L., & Massaro, D. W. A Bidimensional model of pitch in the recognition of melodies. *Perception & Psychophysics*, 1978, *24*, 551-565.

Jones, M. R. Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 1976, *83*, 323-355.

Kallman, H. J., & Massaro, D. W. Tone chroma is functional in melody recognition. *Perception & Psychophysics*, 1979, *26*, 32-36.

Kameoka, A., & Kuriyagawa, M. Consonance theory Part II: Consonance of complex tones and its calculation method. *Journal of the Acoustical Society of America*, 1969, *45*, 1460-1469.

Kilmer, A. D., Crocker, R. L., & Brown, R. R. *Sounds from silence: Recent discoveries in ancient Near Eastern music*. Berkeley, California: Bit Enki Publication, 1976.

Krumhansl, C. L. Concerning the applicability of geometrical models to similarity data: The interrelationship between similarity and spatial density. *Psychological Review*, 1978, *85*, 445-463.

Krumhansl, C. L. The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, 1979, *11*, 346-374.

Krumhansl, C. L., Bharucha, J. J., & Kessler, E. J. Perceived harmonic structure of chords in three related musical keys. *Journal of Experimental Psychology: Human Perception and Performance*, 1981, in press.

Krumhansl, C. L., & Kessler, E. J. Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, 1982, in press.

Krumhansl, C. L., & Shepard, R. N. Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception & Performance*, 1979, *5*, 579-594.

Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 1964, *29*, 1-27.

Kruskal, J. B. Analysis of factorial experiments by estimating monotone transformations of the data. *Journal of the Royal Statistical Society, Series B*, 1965, *27*, 251-263.

Kruskal, J. B., & Shepard, R. N. A nonmetric variety of linear factor analysis. *Psychometrika*, 1974, *39*, 123-157.

Kubovy, M. Concurrent pitch-segregation and the theory of indispensable attributes. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, New Jersey: Lawrence Erlbaum, 1981.

Lakner, Y. A new method of representing tonal relations. *Journal of Music Theory*, 1960, *4*, 194-209.

Levelt, W.J.M., Van de Geer, J. P., & Plomp, R. Triadic comparisons of musical intervals. *British Journal of Mathematical and Statistical Psychology*, 1966, *19*, 163-179.

Locke, S., & Kellar, L. Categorical perception in a nonlinguistic mode. *Cortex*, 1973, *9*, 355-369.

Longuet-Higgins, H. C. Letter to a musical friend. *Music Review*, 1962, *23*, 244-248.

Longuet-Higgins, H. C. The perception of music (Review Lecture). *Proceedings of the Royal Society, London, Series B*, 1979, *205*, 307-332.

Merriam, A. P. *The anthropology of music*. Evanston, Illinois: Northwestern University, 1964.

McAdams, S., & Bregman, A. Hearing musical streams. *Computer Music Journal*, 1979, *3*, 26-44.

Metzler, J., & Shepard, R. N. Transformation studies of the internal representation of three-dimensional objects. In R. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium*. Potomac, Maryland: Lawrence Erlbaum, 1974. Pp. 147-201.

Meyer, L. B. *Emotion and meaning in music*. Chicago, Illinois: University of Chicago Press, 1956.

Null, C. Symmetry in Judgments of Musical Pitch. Unpublished doctoral dissertation. Michigan State University, 1974.

Oakes, W. F. An experimental study of pitch naming and pitch discrimination reactions. *Journal of Genetic Psychology*, 1955, *86*, 237-259.

O'Connell, W. Tone spaces. *Die Reihe*, 1962, *8*, 34-67.

Philippot, M. *L' Arc, Beethoven*, No. 40, 1970.

Pickler, A. G. Logarithmic frequency systems. *Journal of the Acoustical Society of America*, 1966, *39*, 1102-1110.

Piston, W. *Harmony*. New York: Norton, 1941.

Plomp, R. *Aspects of tone sensation: A psychophysical study*. New York: Academic Press, 1976.

Plomp, R., & Levelt, W.J.M. Tonal consonance and critical band width. *Journal of the Acoustical Society of America*, 1965, *38*, 548-560.

Powers, W. T., Clark, R. K., & McFarland, R. L. A general feedback theory of human behavior: Parts I & II. *Perceptual & Motor Skills*, 1960, *11*, 71-88, 309-323.

Pressing, J. Cognitive isomorphisms in pitch and rhythm in world musics: West African, the Balkans, Thailand and Western tonality. Unpublished manuscript, Department of Music, La Trobe University, Bundoora Victoria, Australia 3083, 1979.

Ratner, L. G. *Harmony: Structure and style*. New York: McGraw-Hill, 1962.

Revesz, G. *Introduction to the psychology of music*. Norman, Oklahoma: University of Oklahoma Press, 1954.

Risset, J. C. *Paradoxes de hauteur: le concept de hauteur sonore n'est pas le meme pour tout le monde. Proceedings of the 7th International Congress on Acoustics. Budapest*, 1971, Paper 20510.

Risset, J. C. Musical acoustics. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception*. (Volume 4, Hearing) New York: Academic Press, 1978. Pp. 521-564.

Rosen, C. *The classical style: Haydn, Mozart, Beethoven*. New York, Viking Press, 1971.

Rothenberg, D. A model for pattern perception with musical applications. Part I: Pitch structures as order-preserving maps. *Mathematical Systems Theory*, 1978, *11*, 199-234. (a)

Rothenberg, D. A model for pattern perception with musical applications. Part II: The information content of pitch structures. *Mathematical Systems Theory*, 1978, *11*, 353-372. (b)

Ruckmick, C. A. A new classification of tonal qualities. *Psychological Review*, 1929, *36*, 172-180.

Sachs, C. *The rise of music in the ancient world, East and West*. New York: Norton, 1943.

Schenker, H. *Harmony*. Cambridge, Massachusetts: MIT Press, 1973. (Translated from original 1954 Edition)

Schouten, J. F. The perception of subjective tones. *Proceedings Kon. Ned. Akad. Wetensch.*, 1938, *41*, 1086-1093.

Seashore, C. E. *Psychology of music*. New York: McGraw-Hill, 1938.

Shackford, C. Some aspects of perception. I. *Journal of Music Theory*, 1961, *5*, 162-202.

Shackford, C. Some aspects of perception. II. *Journal of Music Theory*, 1962, *6*, 66-90.

Shepard, R. N. The analysis of proximities: Multidimensional scaling with an unknown distance function, I & II. *Psychometrika*, 1962, *27*, 125-140, 219-246.

Shepard, R. N. Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, 1964, *1*, 54-87. (a)

Shepard, R. N. Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America*, 1964, *36*, 2346-2353. (b)

Shepard, R. N. Approximation to uniform gradients of generalization by monotone transformations of scale. In D. I. Mostofsky (Ed.), *Stimulus generalization*. Stanford, California: Stanford University Press, 1965, Pp. 94-110.

Shepard, R. N. Metric structures in ordinal data. *Journal of Mathematical Psychology*, 1966, *3*, 287-315.

Shepard, R. N. Shepard's tones. In M. V. Mathews, J. C. Risset *et al.*, *The voice of the computer*. Decca Record DL 710180, 1970.

Shepard, R. N. Representation of structure in similarity data: Problems and prospects. *Psychometrika*, 1974, *39*, 373-421.

Shepard, R. N. The circumplex and related topological manifolds in the study of perception. In S. Shye (Ed.), *Theory construction and data analysis in the behavioral sciences*. San Francisco: Jossey-Bass, 1978. Pp. 29-80. (a)

Shepard, R. N. The double helix of musical pitch. In R. N. Shepard (Chair), *Cognitive structure of musical pitch*. Symposium presented at the meeting of the Western Psychological Association, San Francisco, California, April 1978. (b)

Shepard, R. N. On the status of "direct" psychophysical measurement. In C. W. Savage (Ed.), *Minnesota studies in the philosophy of science*. (Volume IX) Minneapolis, Minnesota: University of Minnesota Press, 1978. Pp. 441-490. (c)

Shepard, R. N. Multidimensional scaling, tree-fitting, and clustering. *Science*, 1980, *210*, 390-398.

Shepard, R. N. Individual differences in the perception of musical pitch. In *Documentary report of the Ann Arbor symposium: Applications of psychology to the teaching and learning of music*. Reston, Virginia: Music Educators National Conference, 1981. Pp. 152-174. (a)

Shepard, R. N. Psychophysical complementarity. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1981. Pp. 279-341. (b)

Shepard, R. N. Geometrical approximations to the structure of musical pitch. *Psychological Review*, 1982. (In press).

Shepard, R. N., & Carroll, J. D. Parametric representation of nonlinear data structure. In P. R. Krishnaiah (Ed.), *Multivariate analysis: Proceedings of an international symposium*. New York: Academic Press, 1966. Pp. 561-592.

Shepard, R. N., & Cooper, L. A. *Mental images and their transformations*. Cambridge, Massachusetts: The MIT Press/Bradford Books, 1982.

Shepard, R. N., & Zajac, E. *A pair of paradoxes*. Murray Hill, New Jersey: Bell Telephone Laboratories, 1965. (A computer-generated 16-mm sound film.)

Siegel, J. A., & Siegel, W. Absolute identification of notes and intervals by musicians. *Perception & Psychophysics*, 1977, *21*, 143-152. (a)

Siegel, J. A., & Siegel, W. Categorical perception of tonal intervals: Musicians can't tell *sharp* from *flat*. *Perception & Psychophysics*, 1977, *21*, 399-407. (b)

Small, A. M. An objective analysis of violin performance. *University of Iowa Studies in the Psychology of Music*, 1936, *4*, 172-231.



Stevens, S. S. *Psychophysics: Introduction to its perceptual, neural and social prospects.* New York: Wiley, 1975.

Stevens, S. S., & Volkman, J. The relation of pitch to frequency: A revised scale. *American Journal of Psychology*, 1940, *53*, 329-353.

Stevens, S. S., Volkman, J., & Newman, E. B. A scale for the measurement of the psychological magnitude of pitch. *Journal of the Acoustical Society of America*, 1937, *8*, 185-190.

Sundberg, J.E.F., & Lindqvist, J. Musical octaves and pitch. *Journal of the Acoustical Society of America*, 1973, *54*, 922-929.

Terhardt, E. Pitch, consonance, and harmony. *Journal of the Acoustical Society of America*, 1974, *55*, 1061-1069.

Thurlow, W. R., & Erchul, W. P. Judged similarity in pitch of octave multiples. *Perception & Psychophysics*, 1977, *22*, 177-182.

Tversky, A. Features of similarity. *Psychological Review*, 1977, *84*, 327-352.

Van de Geer, J. P., Levelt, W.J.M., & Plomp, R. The connotation of musical intervals. *Acta Psychologica*, 1962, *20*, 308-319.

Van Noorden, L.P.A.S. Temporal Coherence in the Perception of Tone Sequences. Unpublished doctoral dissertation. Technishe Hogeschool, Eindhoven, Holland, 1975.

von Helmholtz, H. *On the sensations of tone as a physiological basis for the theory of music.* New York: Dover, 1954. (Original publication in German, 1862)

Ward, W. D. Subjective musical pitch. *Journal of the Acoustical Society of America*, 1954, *26*, 369-380.

Ward, W. D. Absolute pitch (Parts I & II). *Sound*, 1963, *2*, 14-21, 33-41.

Ward, W. D. Musical perception. In J. V. Tobias & E. D. Hubert (Eds.), *Foundations of modern auditory theory*. (Volume 1) New York: Academic Press, 1970. Pp. 407-447.

Ward, W. D., & Martin, D. W. Psychophysical comparison of just tuning and equal temperment in sequences of individual tones. *Journal of the Acoustical Society of America*, 1961, *33*, 586-588.

Watt, H. J. *The psychology of sound.* London & New York: The University of Cambridge Press, 1917.

Weeks, D. G., & Bentler, P. M. A comparison of linear and monotone multidimensional scaling models. *Psychological Bulletin*, 1979, *86*, 349-354.

Westergaard, P. *An introduction to tonal theory.*. New York: Norton, 1975.

Wightman, F. L. The pattern-transformation model of pitch. *Journal of the Acoustical Society of America*, 1973, *54*, 407-416.

Winograd, T. Understanding natural language, *Cognitive Psychology*, 1972, *3*, 1-191.

Zatorre, R. S., & Halpern, A. R. Identification, discrimination, and selective adaptation of simultaneous musical intervals. *Perception & Psychophysics*, 1979, *26*, 384-395.

Zuckerkandl, V. *Sound and symbol.* Princeton, New Jersey: Princeton University Press, 1956.

Zuckerkandl, V. *Man the musician.* Princeton, New Jersey: Princeton University Press, 1972.

# 12

# Musical Ability

*Rosamund Shuter-Dyson*


I. Concepts of Musical Ability ..... 391
II. Correlational and Factorial Studies of Musical Ability ..... 393
A. General Considerations ..... 393
B. Pitch Perception ..... 396
C. Tonal Memory ..... 397
D. "Feeling for Tonal Center" ..... 399
E. Harmony and Polyphony ..... 400
F. Appreciation Tests ..... 401
G. Rhythmic Abilities ..... 402
H. Kinesthetic Perception ..... 403
I. Summary ..... 404
III. Musical Ability and Other Intellectual Abilities ..... 404
References ..... 408


## I. CONCEPTS OF MUSICAL ABILITY

The term *musical ability* is "the broadest and safest" in that it suggests the power to act but indicates "nothing about the heritability or congenitalness of inferred potentiality" (Farnsworth, 1969, p. 151). However, Gordon (1979) points out the words *ability* and *talent* include, and thereby confuse, aptitude and achievement. Certainly a broad distinction should, where possible, be drawn between *aptitude* (potential to learn) and *achievement* (what has been learned). The earlier view that it is possible to test innate capacity needs modification by the recognition that all aptitude tests are to some extent achievement tests just as all achievement tests necessarily reflect the initial aptitude that the individual can bring to the learning situation. However, attainment depends not only on aptitude but also on the teaching received and the child's interest in music and willingness to learn. Colwell (1970, pp. 72–73) suggests that it is useful to regard musical aptitude as referring to those qualities that develop


Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

over a long period and may cease to improve beyond a certain level in spite of further training. For example, tonal memory improves slowly and is possessed in widely varying degrees that do not seem to be directly related to length of experience with music. Differentiating between the major and minor modes may be learned at any period of life and within a relatively short time and hence may appropriately be measured in an achievement test.

A useful analogy may be drawn with intelligence. As Vernon (1968) noted in a paper on "Potential Ability," we need to recognize the existence of genetic differences between individuals in their capacities for building up neural and mental schemata, what Hebb (1949) called "Intelligence A." The psychologist can, however, only observe the effectiveness of present behavior or thinking through the interaction of the genes and the stimulation offered by the environment (i.e., Intelligence B). This is culturally conditioned by the environment provided by the ethnic group into which the child is born, especially by the child's home and his or her leisure pursuits. Vernon also notes the effects of physiological conditions before and shortly after birth, Cattell's (1965) *Constitutional Intelligence*. The results of intelligence tests, Intelligence C, are of considerable predictive value because they sample useful mental skills.

The term *musicality* is used by Revesz (1953) to denote "the ability to enjoy music aesthetically." The ultimate criterion is the depth to which a person, listening and comprehending, can penetrate the artistic structure of a composition. As Zuckerkandl (1973, Chapter 1) points out, music is equated with the compositions that have evolved in the Western world over the past thousand years, characterized by their tonal polyphony. Composing and interpreting have become the business of experts, whose works seem to demand special sensitivity to comprehend them. Hence, a distinction between "musical" and "unmusical" comes to be drawn, even though conceived as a continuum rather than a dichotomy.

A different concept of musicality implies that all men *qua* men are music makers. Blacking (1971), for example, proposes that human beings possess musical competence similar to the linguistic competence postulated by Chomsky and that "average musical ability" is as universal as average linguistic competence.

For example, the Venda in Africa assume that everyone is capable of music making, even deaf persons having the ability to dance. However, the Venda recognize that some persons perform better than others do. Exceptional musical ability is indeed expected of children born into certain families. Only a few may actually become exceptional musicians; if they do, what is considered to set them apart is not that they can do what others cannot but that they devote more time and effort to music.

Important evidence on the inherent music-making potential of humans comes from exploratory studies conducted at the Bioacoustic Laboratory in Pennsylvania. In neonates, the effects of multidimensional signals, whether tonal sequences or synthetic speech sounds, produce distinctive overt reactions whatever the state of prior arousal (Eisenberg, 1976, p. 137). Such responses occurred, for example, to descending versus ascending sequences (2000 down to 500 Hz versus 500 up to 2000 Hz). Moog (1976, p. 113) noted that up to the age of about 3 the development of a response to

music is determined by innate ability and by the musical stimuli presented to the child more or less haphazardly. But he reported from his investigation of some 500 children that "by the end of the third year all children are capable of imitative singing" (p. 97), even if only some groups of notes are sung correctly.

Certain of Moog's results are of special interest: 16% of 2-year-olds did not sing words but hummed or sang something similar to the rhythm and pitch of Moog's tests (p. 83). Singing before speaking seems to be one characteristic of children with conspicuous musical talent. For example, Erwin Nyiregyhazy, a prodigy whose talents were studied by Revesz (1925), could reproduce correctly melodies sung to him in his second year; at the beginning of his fourth year, he began to play on the piano everything he heard; and by seven could analyze complicated chords with great accuracy. Many such children are born into families of musicians, even if not distinguished ones. Others are born into families with only modest musical abilities. Conversely, highly talented parents may produce children with little aptitude or interest in music. Such "anomalies" suggest more strongly than do family resemblances the likely action of hereditary factors (Shuter, 1968, p. 112).

In one sense, then, the ability to make and apprehend music is inherent in all humans, but in another sense aptitude varies among individuals and may set a limit on ultimate achievement, no matter how favorable the environmental influences nor how highly motivated the person.

## II. CORRELATIONAL AND FACTORIAL STUDIES OF MUSICAL ABILITY

### A. General Considerations

A fair variety of material from which tests might be constructed has been examined over the last 70 years. The best of the published tests, in spite of various imperfections, have proved successful in helping to identify the relative status of individuals to an extent that makes them useful in music education. Detailed information on such tests is readily available elsewhere (Lundin, 1967; Colwell, 1970; Whybrew, 1971). All the more reputable batteries were subjected to several years of experimentation before they were published, thus throwing light on what subtests are the most efficient. Allowance has to be made for the limitations imposed by the group testing situation: tests of vocal or instrumental performance have lagged behind perceptual ones since they need to be administered individually.

It is as research tools that tests are to be considered in this paper. In assessing the meaningfulness of correlation and factor analytic studies, we must bear in mind the reliability and the validity of the tests used and their suitability for the groups tested.

Tests naturally reflect differences in their authors' concepts of musical ability. Seashore (1938, p. 3) believed that the sensory capacities measured in his laboratory and eventually presented in his *Measures of Musical Talent* (1919) were basic to musical aptitude, could be sharply defined, and could be present or absent in the individual to

varying degrees. Correlation studies of the *Measures* rarely produce zero correlations. Teplov (1966, pp. 62–64) tabulated the intercorrelations from 14 early studies of the *Measures*. The average of these ranged from .22 (intensity-rhythm) to .52 (pitch-tonal memory). Some of the agreement is probably due to intelligence. But after partialing out intelligence, Drake (1939) found a common factor and two group factors in a study of 163 13-year-old boys. He concluded that even when a special attempt is made to measure independent capacities, it is seldom entirely successful. While the revised version of the *Measures* (1939 and 1960) appears to have improved reliability (rising to .8 at least in the case of pitch, intensity, and tonal memory), the relevance of sensory capacities to functional music activities has been repeatedly questioned (Mursell, 1937; Wing, 1948; Lundin, 1967).

Wing (1948) experimented with 25 tests that were eventually reduced to seven: three of ear acuity (chord analysis, pitch, and melodic memory) and four of appreciation. They became the Wing Standardised Tests of Musical Intelligence (1961). One criterion for including a test in the short series was a satisfactory, but not too high, correlation with the total for the whole battery. This was justified, he believed, since his factorial studies had shown a general factor existed in a wide variety of tests. Though these studies were carried out on small groups by present standards, the results produced a general factor having 40% of the variance with positive saturations throughout. Minor factors separated tests and persons into "analytic" and "synthetic" types, and those depending mainly on harmony as opposed to melody. How far any common factor is evidence of the unitariness of musical ability obviously depends on how successfully the tests used cover all aspects of the ability. In Wing's case this depends essentially on how valid is his claim that "no vital test is missing from the short series" (Wing, 1948, p. 49). The battery is certainly comprehensive for one limited to an hour; its reliability and validity, at least taken as a whole as Wing intended, are good.

The Gordon *Musical Aptitude Profile* (*MAP*) (Gordon, 1965) consists of three parts: tonal (melody and harmony); rhythm (tempo and meter), and musical sensitivity (phrasing, balance, and style). The intercorrelations of the tonal with the rhythm subtests range from .45 to .65 (Gordon, 1965, pp. 56–57) and from .35 to .70 (Tarrell, 1965). Since these values were lower than the reliabilities of the subtests, Gordon (1965, p. 51) claims that this "constitutes evidence although somewhat indirect, of the multidimensional nature of musical aptitude." However, he also states (p. 9) that in music "rhythm and melody interact in an inseparate way." Considerable evidence has been reported on the reliability and validity of the MAP and its subtests (Gordon, 1967). His new *Primary Measures of Music Audiation* (*PMMA*) (Gordon, 1979) has separate tonal and rhythm subtests, since he considers that a sense of meter to be on a par with that of tonality. Intended for children kindergarten through grade 3, reliability and validity are most promising.

The first battery devised primarily for children aged 7/8 to 14 was the *Measures of Musical Abilities* (*MMA*) (Bentley, 1966a). Bentley (1966b, pp. 89–92) stated that, though the functions measured by his tests may overlap and seem to be working

together, they are in fact separate. Certainly the test of rhythmic memory correlates rather weakly with the rest of the tests in both his and Young's (1973) research.

The results of factor analyses depend in part on the method used and on whether and how the factors are rotated. Unlike Wing (1941) and McLeish (1950), who were content with unrotated solutions, Karlin (1941, 1942), Franklin (1956) and Holmstrom (1963) rotated their factors. Computer programs are now producing objectivity in rotations and giving several "primary" factors of more nearly equal importance.

Whellams (1971) reanalyzed many previous correlation and factorial studies by principal components, followed by varimax rotation, to investigate how far the various diverse studies might confirm or deny similar results. In the reanalyses of 16 studies of the Seashore *Measures*, three factors recurred, all seeming to involve tonal memory. Three rather similar factors emerged with the Kwalwasser-Dykema (K-D) (1930) tests, plus a fourth factor of tonal movement and rhythm imagery. Factorizing five correlation matrices of the MAP produced only two significant factors: tempo/meter and phrasing/style. Ten studies of the Wing tests produced five significant factors. In all, 15 factors of some significance were extracted. The tentative interpretations of these were summarized by Whellams thus:

| Significant[a] Factors | | Tests loaded | Interpretation | P.A.M.A. factor |
|---|---|---|---|---|
| Perceptual | | | | |
| I, VI, XI | (23) | Pitch, tonal memory | Factor for pitch imagery | pi |
| II, V | (11) | Rhythm, tonal memory | Dynamic or kinesthetic factor for the development of pitch perception | kp |
| III, XV | (12) | Consonance, chords, harmony, memory | Factor for harmonic ability | h |
| X | (10) | Chord analysis, harmony | Tonal separation factor | t |
| VIII | (4) | Tempo, meter | Rhythmic ability | r |
| IV | (3) | Tonal movement, memory | Experience factor | xp1 |
| VII | (2) | Rhythm imagery, tonal movement | Experience factor | xp2 |
| Appreciation or judgment | | | | |
| IX | (3) | Phrasing, style | Musical judgment factor | j |
| XII | (7) | Intensity | Wing test factor | j1 |
| XIII | (6) | Phrasing | Wing test factor | j2 |
| XIV | (4) | Rhythm | Wing test factor | j3 |

[a]Numbers in parentheses are the number of studies in which the factors appeared.

Biblioteca del Real Conservatorio Superior de Música - Madrid

From nonorthogonal primary factors second-order factors may be extracted by further analysis. For example, Cattell (1963) proposed two second-order factors of intelligence—one representing the crystallization of intellectual potential through acculturation (Gc) and one representing the capacities of intellect that have remained relatively "fluid" immune to acculturation influences (Gf). So far, only Horn and Stankhov seem to have tried to extract second-order factors from auditory tests. They carried out an extensive exploration of auditory and visual functions, using the Seashore battery, Drake's Rhythm tests, and the Wing tests 1–3, as well as visual tests and auditory tests devised to correspond to the visual ones. Their subjects were 241 men, aged between 16 and 54, inmates of a state penitentiary. Results indicated six primary auditory abilities: auditory reasoning (Ra), temporal reordering (Tr), nonsymbolic recognition memory (Mr), detection of distorted speech (DDS), masked speech comprehension (MSC), and rhythm (Ry) (Horn, 1973). Later analyses of the data (Stankov & Horn, 1980; Horn & Stankov, in preparation) include the extraction of a second order factor of general auditory functioning (Ga) and one of auditory acuity (Ac).

### B. Pitch Perception

Pitch discrimination is widely acknowledged to be an essential part of musical perception. It appears in some form in most test batteries and in many experimental studies of the development of musical ability (e.g., Andrews & Deihl, 1967). Pitch tasks are among the most difficult for younger children, though it is sometimes not clear how far this is due to a lack of understanding of "higher" and "lower" as applied to music.

Quite strong intercorrelations are typically found between tests that explore discrimination of fine pitch differences by means of oscillator-produced tones (Seashore and Bentley) and also between these and Wing's test where the pitch change is masked within two piano chords. Examples are

| | |
|---|---|
| Seashore–Wing | .63 (McLeish, 1966) |
| | .49 .67 .62 (Franklin, 1956, pp. 184–185) |
| | .53 (Faulds, 1959) |
| Bentley–Wing | .46 (Donat, 1964) |
| | .26 (Whellams, 1971) |

Closely allied to pitch discrimination are interval tests. Lundin (1944, 1949) required his subjects to judge whether a second melodic interval moved up or down. Faulds (1959) slightly modified the test and found that the Seashore pitch test correlated .58 and the Wing pitch test .76 with it.

The work of the Siegels (Siegel & Siegel, 1977b) provides a different perspective on the perception of pitch by musicians, who appear to categorize pitches in a way similar to the categorization of consonants in language. One group of six musicians, whose participation in music averaged 53 hours a week, identified in-tune intervals

with a 95% accuracy. In one experiment they were asked to identify 13 tonal intervals ranging from 20 cents below a perfect fourth by 20-cent steps to 20 cents above a perfect fifth by pressing buttons labeled perfect fourth, augmented fourth, and perfect fifth. All examples of a musical category were perceived as equivalent, variability being least at the center of the category and most at each boundary. Intermediate stimuli were heard to "pop" one way or the other, not as halfway between two musical intervals. Siegel and Siegel had previously (1977a) found that subjects chosen for their lack of musical experience made inconsistent judgments largely influenced by context in this type of experiment. For a further experiment (J. Siegel, 1976), the six musicians were joined by a performer of Indian music who was also experienced in Western music; his score on in-tune intervals was 100%. Subjects were now required to judge whether intervals (minor third, fifth, and octave) were in-tune, sharp, or flat. In one type of stimuli, all were out of tune, with sharp and flat examples occurring equally often; in another, half were in tune, 25% flat, and 25% sharp; and in a third all the intervals were in tune. The flat/sharp intervals deviated by .25 of a semitone from accuracy. As the percentage of in-tune intervals increased from 0 to 100%, the percentage of in-tune judgments increased only from 47 to 62%. Flat intervals were detected more accurately than sharp or in-tune ones. In spite of their many years of musical training, no subject approached perfect performance, two reaching only chance level. The performer of Indian music fell in the middle of the group. In a musical context he might have scored better. Citing Seashore's (1938, p. 269) conclusions from his acoustic measurements of the deviations from exact pitch in the performances of well-known artists, Siegel and Siegel (1977b) concluded that "categorical perception allows one to recognize the melody, even when the notes are out of tune, and to be blissfully unaware of the poor intonation that is characteristic of good musical performance." When, however, a musician was given an untimed production task, emulating the tuning of an instrument, he was very accurate at adjusting the frequency of an oscillator to a value corresponding to musical notes (W. Siegel, 1976).

### C. Tonal Memory

Drake (1939) believed that memory acts to knit together specific capacities such as discrimination of pitch, time, and rhythm. Bugg and Herpel (1946) investigated whether the intercorrelations of the various Seashore tests would drop substantially when tonal memory was partialed out. Partialing out the results of testing 181 subjects with the Seashore *Measures* did lead to reduced intercorrelations, but the differences were not great.

The Seashore tonal memory test contains 3-, 4-, and 5-tone items, which are repeated with one tone changed, the task being to detect which tone has been altered. Wing's test ranges from 3 to 10 notes and includes a rhythmic content but is based on the same principle. The Drake Musical Memory Test (Drake, 1957) is rather different. Two-bar melodies have to be remembered and compared for possible changes of time, key, or notes. The testee has to detect what element (if any) has been

changed. As the test progresses, up to 7 variations follow the original, considerably increasing the memory load. Gaston (1957) included in his Test of Musicality seven items built on the same principle as the Drake except that the testee must identify change of note or rhythm from a repetition. This was the most discriminating of all the tests used by Bentley (1955) in an investigation contrasting the scores of 110 noninstrument-playing high-school students with 110 instrument-playing music students, matched on the bases of sex, IQ, grade placement, and socioeconomic status on music tests, including the Wing battery.

The intercorrelations of the various tests of tonal memory are quite strong:

| | |
|---|---|
| Seashore-Wing | .74 (McLeish, 1966) |
| | .64 and .75 (Franklin, 1956, pp. 184–185) |
| Seashore-Drake | .56 (Lundin, 1958) |
| | .55 (Farnsworth, 1969) |
| | .61 (Rainbow, 1965) |

Davies (1971) sought to devise a new type of test material that would place emphasis on memory for entire sequences rather than for single elements. First, a tonal sequence of three or four tones was presented, then a longer sequence (from 4 to 8 tones). Sometimes the short sequence was contained intact in the longer one. The testees were asked whether or not the long sequence contained the short one and also in what part of the longer sequence it was located. The sequences were derived from statistical approximations to music upon the analogy with Miller and Selfridge's (1950) experiments with verbal material. The Davies test had reliabilities as high as .7 even for children as young as 7. The items were based on the equal-tempered scale; sequences based on a scale not common to any culture might also be produced.

The above mentioned tests could be subsumed under "Short-Term Memory." Long-term memory is obviously crucially involved in musical ability, as was recognized by Seashore: "Hearing is not a mere registering of sounds. It is a positive active process of reconstruction in the mind of the listener (1938, p. 168)." He also speaks of auditory imagery "the capacity to hear music in recall, in creative work, and to supplement the actual physical sounds" as perhaps the most outstanding mark of the musical mind (p. 161). Gordon (1977, p. 2) uses the term "audiation" to designate hearing music through recall or creation, in the absence of the physical sound, and thereby deriving musical meaning, as opposed to musical perception/conception when the music is actually being performed by others.

At least in the early stages of learning, auditory imagery seems to need to be supported by kinesthetic imagery (see Section II,H). Gordon recognizes this when he speaks of the person who is tonally audiating as singing silently and of rhythmic audiation requiring covert, if not overt, eurhythmic functioning. Agnew (1922) found most of 200 musicians rated their auditory image of "America" as played on the piano, "as clear as the actual hearing" or "very clear." Eighty-nine psychologists rated their auditory imagery as low and needing kinesthetic support. Curiously, a study of self-rated musical imagery produced with naive subjects a sizable association with a pitch identification test but none with Drake's memory test (Bergan, 1967).

### D. "Feeling for Tonal Center"

Even now that much contemporary music is atonal or polytonal, the ability to feel the presence or absence of a tonal center is to be considered important (Colwell, 1969, p. 112; Gordon, 1977, p. 1). Yet the melodies children spontaneously sing do not seem to relate to any observable tonal center, at least not one found in Western music (Moorhead and Pond, 1942, p. 14). Gordon (1979) suggests that a sense of tonality develops simply out of the child audiating the resting tone. A notable feature of the tonal part of his *PMMA* test is that the tonic is found in at least one of the patterns of the pair (which the child is required to judge as "same" or "different") and is included in both in strategically located items. All the patterns used had in an extensive previous research been found to be "easy," differences in difficulty in the test being attained by the pairings of the patterns for comparison. Gordon considers the sense of tonality to be basic to the development of musical understanding, for example, in the ability to follow modulations.

Just as one criterion of linguistic ability is the child's acquisition of a command of his or her native language, it is reasonable to expect musical aptitude to manifest itself in acquisition of feelings for tonality. Zenatti's experiments (1973, 1975) were concerned with comparisons between the perception and discrimination of tonal (or pentatonic) melodies as compared with atonal ones. Not until the age of 7 or 8 were tonal melodies perceived better than atonal ones above a chance level. Success with tonal as opposed to atonal melodies improved with age, though in the case of three-note tasks, when the children were either very good or very bad at discrimination, it made little difference whether the phrases were tonal or atonal. In the case of four-tone phrases, the better the discrimination score, the greater was the difference between tonal and atonal perception (Zenatti, 1969, Chapter III). Even adults perceived a change of note on a second playing more surely with tonal as opposed to atonal phrases (Francès, 1958, pp. 81–90).

One of the most valid of the K–D tests is on tonal movement, in which the testee hears an incomplete pattern of four tones and is asked to indicate whether the tone needed to complete the sequence should move "up" or "down" from the fourth tone. Drake (1931) found that few could succeed in this test but not on tonal memory, but that many could make good scores on tonal memory but not on tonal movement (i.e., "judgement, an entirely different operation from memory, is yet impossible without the latter"). Franklin (1956) based his Tonal Musical Talent (TMT) test on the principle that if a short two-part melody is interrupted immediately before the final tone, the subject who could sing the tonic would thereby demonstrate his sense of tonality. The TMT correlated .43 and .55 with Wing's pitch test, .55 with Wing's memory test, and .40 with Wing's harmony test, and .60 and .55 with Seashore's memory test but only .18 and .40 with Seashore's pitch test.

From his factor analyses of his test, the Wing battery and Seashore's pitch and memory tests, Franklin found two factors of special interest to him: (1) "mechanical-acoustic" pitch discrimination (of fine differences, as required by Seashore) and (2) "judicious-musical" pitch discrimination, the ability to deal with pitch changes in a

musical context, as in his TMT and the Wing tests. Franklin argued that judicious pitch would be the most important aspect of any general music factor that might be found. Holmstrom (1963) pointed out that fine pitch discrimination is likely to help rather than hinder pitch tasks in a musical context. Holmstrom's own analyses yielded two pitch factors: "alpha" had a loading on intensity, had possibly a physiological basis, and might be only slightly influenced by experience with music; "beta" had greater loadings on tonal memory and seemed to be more salient among older children with varied experience of music. In Whellams' more comprehensive studies (see Section II,A), the pi factor does not overshadow all others. His interpretation of pi seems more comparable to Holmstrom's beta. Pi seemed to be influential in sight-singing and to improve with length of training.

### E. Harmony and Polyphony

Franklin (1956) concluded from the researches of Rupp and of Brehmer that the establishment of an ear for tonality is preceded by a stage in which an ear for pitch and an ear for consonance and dissonance may exist but not for simultaneous horizontal and vertical listening. However, an acquisition of a sense of tonality is no guarantee of an appreciation of suitable harmonizations (on the criterion of conventional Western harmony) (Teplov, 1966, Chapter VI). Teplov cites the interesting case of a subject who found the "good" harmonization more pleasant but the discordant one less distracting from perceiving the melody. Though Zenatti (1969, Ch. II) found children between 8 and 10 began to be able to recognize a well-known tune presented fugually, even children up to 12 had difficulty perceiving the bass part.

The *MAP* harmony test is concerned with "like/different" judgments about the lower of two-part melodies. The Wing test is concerned with the judgment of the better harmonization. This test tends to be linked with the Wing pitch and memory tests, but a harmony (h) factor as distinct from a tonal separation factor was revealed in Whellams' reanalyses of the Wing and Seashore batteries. In his own study of children around 9, in which both the Wing and the Bentley chord analysis tests were included but not the Wing harmony test, a combined h:t factor emerged. In McLeish's analysis (1966) the Wing chord analysis and the harmony test loaded similarly (.66 versus .65) on the "musical cognition" factor. With professional students of the Eastman School of Music, the harmony test proved too easy, but the chord analysis test successfully separated the good from the very good student (Shuter, 1968, pp. 35–36).

Experiments with cadences have provided valuable data on the growth of musical ability. Imberty (1969, p. 216) showed that 10-year-old subjects understood both perfect and imperfect cadences, but if the third degree of the scale was used for the tonic chord, a feeling of surprise was created (i.e., judgment was still essentially melodic). Yet, a cadence test devised by Wing (1948, p. 16, 48), in which the melodic line of two cadences to be compared for their feeling of finality was kept as similar as possible, proved too easy for his subjects. It would seem to be true that at certain

stages of musical development (using the criterion of age or of level of aptitude) an appreciation of appropriate harmonization is salient; at others, the analysis of chords appears to be more discriminative.

### F. Appreciation Tests

Among tests of appreciation of music that are commercially available are the Indiana-Oregon Music Discrimination Test, the last four tests of the Wing battery, and the third part of the *MAP*. Indiana-Oregon, developed by Long, was based on the original Hevner-Landbury test. The principle remains the same: an original excerpt is paired with a "distorted" version. The testee has to state which version he or she prefers and which element—rhythm, harmony, or melody—has been changed. Wing presents an original along with an impoverished version but uses these to make separate tests of rhythm, harmony, intensity, and phrasing. A criticism of such a procedure is that testees who recognize the music are at an advantage. But, as Wing (1948, Chapter VI) makes clear, scores made on known tunes were no higher than unrecognized tunes. When Wing was devising his tests (in the late 1930s), the music of Wagner and Ireland proved too strange to be useful as test material. In Martin's (1976) research and in McLeish's (1950), the Wing appreciation tests do not correlate conspicuously with the Oregon.

Gordon's approach to establishing the "better" version was to begin by asking professional musicians which of two interpretations of a specially composed item they preferred. Only when high consensus (9 out of 10) was reached, and this was confirmed by experimental field trials, were the items incorporated into the final version of the test. [Hoffren (1964) also used a "consensus" judgment on which to base a test of expressive performance of music that would resemble as closely as possible the judgment required in an actual musical situation.] Consensus is obviously based on value judgments that may change with time—such tests may need to be updated.

An early statement of the relationship between Seashore's tonal memory test and the Oregon was that they correlated .65 (Bugg & Herpel, 1946). This finding was confirmed by McLeish (1950). The Oregon test proved to be highly related to the Wing battery, especially to the memory, pitch and harmony tests, and to Seashore's tests of pitch and tonal memory. In fact, the judgment of "Which element changed?" received the highest loading (.86) on McLeish's musical cognition factor. The correlations between Gordon's sensitivity tests and the Seashore *Measures* are much lower (Gordon, 1969). Gordon himself takes the view that musical appreciation depends on musical understanding, which depends on awareness of basic aural elements (Gordon, 1977, p. 1).

Swanick (1973), too, believes that much cognitive activity is involved in esthetic response to music and that the intensity and quality of any emotional response depends on this. His experiments were based on setting up an experimental subculture, a pair of alternating notes in a metric scheme as a basic unit as a repeated norm against which a deviation could be located. (See Fig. 1). The various musical events included



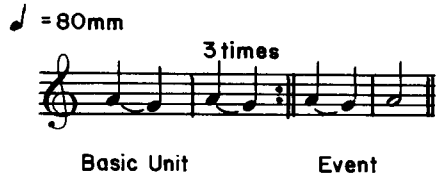



**Fig. 1.** From Swanick (1973).

change of tonality, of tempo, of direction of melody. The subjects were simply asked, "What happens in the music just before the end?" and to record their responses on semantic differential scales such as active/passive, light/heavy, and happy/sad. Even small changes carried for all subjects, even for those as young as 7, a clear and specific meaning, and there was a strong consensus of opinion. He concluded that, in addition to perceptual and memory abilities, three more musical abilities are central to the processes of understanding music: (1) the ability to locate piece and style norms, (2) the ability to understand the flux of "meaning" embodied in ongoing music, and (3) the ability to make predictions as to what may follow so that deviations arouse surprise and excitement.

Scherer and Oshinsky (1977) manipulated acoustic cues with electronically synthesized tone sequences; these produced in their naive judges systematic emotional effects—results which seem to confirm Swanick's.

### G. Rhythmic Abilities

Wing (1948, p. 47) justified the inclusion of only one rhythm ability in his short series on the grounds that the correlation between time-pattern dictation and appreciation of rhythm was very high, hence one could be regarded as redundant. Drake's rhythm tests are concerned with the ability to maintain a steady beat, in the case of Form B against a faster or slower distracting beat. Tanner and Loess (1967) reported intercorrelations of .37 between Seashore's time and rhythm tests, .56 between Seashore's rhythm test and Drake A and .38 with Drake B, .38 between Seashore's time test and Drake A and .43 with Drake B. Very low correlations were found between the Wing rhythm test and all these tests. Horn (1973) reported a correlation of .69 between the two versions of the Drake tests and a loading of the Seashore rhythm test on a rhythm factor. However, the Seashore rhythm test had a higher correlation with a factor largely involving other Seashore tests than it had with the rhythm factor. It therefore seemed more a function of nonsymbolic recognition memory (Mr) than a function of capacity to continue an established rhythm—Whellams might have interpreted this as a kp factor (See Section II,H). Hiriartborde (1964) found a factor that had high loadings on the Wing and Seashore memory tests, the Wing intensity and phrasing test, as well as on a test of synchronization. The factor that showed the highest variance, 28%, involved motor coordination and control of

movement as well as tonal memory and pitch. A factor concerned with the ability to structure rhythmic groups had a loading on the Seashore rhythm test.

For his study of the rhythmic abilities of adults, Thackray (1969) devised tests of rhythmic perception, rhythmic performance, and rhythmic movement. He later adapted the perception and performance tests for use with children (Thackray, 1972). He noted that inability to memorize some of the longer items was an important reason for low scores (1969, p. 29)—again suggesting a kp linkage between rhythm and memory. His own factor analyses (unrotated) suggested that there was a general factor, though this was complex. With adult physical education students, the following intercorrelations were found: perception and performance .63, perception and movement .55, and performance and movement .59. With child groups performance tests seemed superior to tests of perception as a means of testing general rhythmic ability. Ability to maintain a steady tempo appeared to be a highly specific rhythmic ability. Igaga (1974) compared the rhythmic abilities of Ugandan with those of English children. The Ugandan children tended to excel on the Thackray performance tests. In particular, the Ugandan girls were superior to the boys by 10 points on perception and 8 on performance—perhaps because Ugandan girls participate from their earliest years in domestic activities, accompanied by rhythmic music.

## H. Kinesthetic Perception

Whellams' (1971, p. 240) kp factor seems to span the tonal memory and rhythmic aspects of music, to have a large genetic component, and to suggest the kind of overall musical talent envisaged by Wing. It is interesting in drawing attention to the muscular component in auditory perception that is in line with contemporary feedback models of skilled behavior. From validation experiments on the Seashore tests, Whellams (p. 239) suggests that kp has a positive influence on harmonic attainments, that its connection with rhythmic ability goes beyond mere recognition of meter or time, that it is indeed an important central activity influencing the development of various kinds of overt musical behavior and of the higher level aural imagery required for tasks such as written harmony and dictation. A factorizing of Gordon's (1969) data comparing the *MAP* with the Seashore produced a first factor heavily loaded on *MAP* melody (.80), *MAP* harmony (.77) which also had sizable loadings on *MAP* tempo (.46) and meter (.41) and on Seashore pitch (.59) and memory (.30). This, too, Whellams reasonably considered a kp factor.

Mainwaring (1933) found that children and Training College students tended to translate auditory into kinesthetic processes in order to recall tunes they had listened to well enough to answer questions about them. He also cites the disturbance to recall in an experienced singer who had been forbidden to sing a tune to herself while it was being played to her. Vernon (1931, p. 126) noted the importance of the kinesthetic aspects of music in making perception of music definite but believed that at the level of the highly trained musician, many can listen in purely auditory and intellectual terms. Sergeant too (1969, p. 434) agrees that childhood perception of pitch is first

facilitated by a vocal scheme but believes that kinesthetics is not an important factor in adult perception and reports that subjects who do not possess absolute pitch are more likely to have recourse to kinesthetic sensations than those who do. Joyner (1969) and Gould (1969) both worked on the assumption that improving pitch perception and the singing voice requires a kinesthetic approach. [See Shuter (1974) for a review of such studies.] An extensive investigation by Kyme (1967) demonstrated the value of performing in an orchestra or a choir as the prime means of developing musicality (capacity to grasp in its completeness and detail a musical idea that has been heard).

The disruptive effects of DAF seem to operate on a music performance task as with verbal tasks (Gates, Bradshaw, & Nettleton, 1974).

Some evidence on the role of gross bodily movements on rhythmic perception was obtained by Moog (1978). He tested children, aged 10 to 11 (1) with serious physical handicaps, (2) with mental handicaps, and (3) normal children (IQ 80–118). Limitation of movement since earliest childhood reduces rhythmic perception nearly as much as does low intelligence. Children with physical handicaps tend to use pitch rather than rhythmic perception to discriminate rhythmn patterns that may (or may not) be presented in a melodic context. The question of whether some individuals are attentive to pitch structure and others to rhythmic structure, Moog believes, needs further research.

### I. Summary

Davies (1978, pp. 154–155) is right to say that a more open-ended approach to testing is now needed. Some older tests do not reach current recording standards. The greater exposure to music of all kinds may require an updating of norms (Tomita and Kurosu, 1976). New tests, such as Davies' (1971), seem to show up the possibilities of using "nonmusical" material that might minimize the effects of experience of music. Attempts like Vaughan's to parallel in music "creativity" tests, such as those of Torrance, need to be given due regard (Vaughan, 1977). Karma's (1975, 1978) test based on the ability to structure acoustic material would seem just the kind of culture-fair test that might throw light on "fluid musical ability."

## III. MUSICAL ABILITY AND OTHER INTELLECTUAL ABILITIES

The correlations between the results of testing musical ability with other intellectual abilities in the case of some 16,000 subjects, as tabulated by Shuter (1968), were almost without exception positive but low. Both Wing (1954) and Kwalwasser (1955) refer to .30 as being about the correlation to be expected with ordinary, unselected subjects. Wing observed that there was usually good agreement between low intelligence and low scores on his test, but that disagreement occurred when a high IQ was accompanied by a low musical ability score. Edmunds (1960) studied 118 chil-

dren, aged 12 and 13, in a mixed-ability school. He concluded that once a certain level of intelligence is reached (approximately IQ 90), intelligence no longer plays a significant part in level of musical ability. A large proportion of the correlations tabulated by Shuter refer to college students for whom intelligence would not be likely to be important in success with musical ability tests. Whellams (1971 p. 299) reported that correlations between the Wing tests 1–3 and the Bentley tests as compared with nonmusical tests frequently exceeded .5 and some were greater than .6 with his young subjects whose test scores on nonverbal tests were about average, but whose reading ages were about two years lower than the norms. If one applies Edmunds' formula to children of 10, the equivalent mental age would produce an IQ of 108. Zenatti (1975) found that the acuity of perceptive discrimination and tonal acculturation among 396 subnormal children approximated to that of normal children of the same *mental age.*

However, the correlations between the *PMMA* and intellectual tests are low. For example at kindergarten level, with Metropolitan Readiness Tests Level 2, tonal correlated .29 with language and .24 with quantitative; rhythm .16 with language and .15 with quantitative. At grade 1, correlations with the Stanford Reading and Mathematics Achievement Test range from .24 and .37; at grade 3, correlations with the Lorge–Thorndike Intelligence Tests (verbal and nonverbal) range from .19 to .30.

In the case of those Achievement in Music Tests that are largely tonally based, correlations with intellectual tests are also low. For 1400 fourth-grade students, the Musical Achievement Test (MAT) (Colwell, 1969) correlated .23 with the Iowa Tests of Basic Skills (Colwell, 1969, p. 80). Intercorrelations among scores on the Iowa Tests of Musical Literacy (ITML) (Gordon, 1970) and the verbal Lorge–Thorndike Intelligence tests fell around .3 (Gordon, 1970, p. 116). On the other hand, grades on theory of harmony and musical history can be predicted by verbal intelligence tests (Holmstrom, 1969). Mursell (1937, pp. 336–339) pointed up the close association between educational attainments and musical ability when the criteria employed are teachers' estimates of student ability. It is hard to judge how far the student's incentive to work hard and the "halo" effect of success in other school work may contaminate such results.

The statistic used in the above research was usually the correlation coefficient. Sergeant and Thatcher (1974), however, consider that analysis of variance provides a more useful result. They investigated the interrelationships among the Cattell Culture Free Scale B, ratings of the sociocultural and -economic characteristics of the family, and the abilities of 75 children, aged 10 to 11, to perform a rhythmic and a melodic task. Highly significant relationships were found among all the variables, the trends being linear. In the case of the rhythmic task quadratic and cubic trends were also significant. Phillips (1976) carried out a careful study of the scores made on Wing tests 1–3 and five of the Thackray Rhythmic Perception tests by 194 children from four contrasting schools. Highly significant differences were found among the variables whether classified by music scores, musical background or the Thorndike–Hagen Cognitive Abilities Test.

Pedagogically, it is of importance that children from disadvantaged homes (and probably with relatively low IQs) can, given suitable opportunities to learn a musical

instrument, equal or surpass the achievements of a heterogenous group both when the former's *MAP* scores were superior and when they were average. A potent motive was probably a desire to compensate for their limited academic attainments (Gordon, 1975).

That certain persons of low mental ability can show capacities for music far superior to their general intellectual functioning is exemplified by the so-called idiot-savants. Typical musical accomplishments show some similarities to those of infant prodigies, they can, for example, often play a tune by ear after hearing it only once or twice (Shuter, 1968, pp. 98–102). Feats of memory are certainly commonly reported. Blind Tom was able to memorize a piece after one hearing (Drake, 1940); a blind imbecile girl could play a new and difficult piece after hearing it only once (Rife & Snyder, 1931); *XY* had good pitch discrimination and a good tonal memory, being able to play by sight or by ear, though he produced no original compositions (Minogue, 1923; Rife & Snyder, 1931). One idiot-savant described by Rife and Snyder could play on the piano any tune sung to him, using harmonically correct chords, though he had never received any training in music. *L* showed remarkable interest and ability in music, rhythm, and counting in his third year and was apt to acquire long-lasting obsessive interest in certain music—like an aria from *Otello*, the words of which he learned to sing in Italian phonetically (Scheerer, Rothmann, & Goldstein, 1945). Current theorists on hemisphere bilaterity might postulate left-hemisphere malfunctioning. A subject (*S*) with a Binet IQ of only 67 but a WAIS verbal score of 92 was eventually taught to speak through the medium of lyrics; he became good at sight-reading and played the piano at rehearsal for a chamber music orchestra. He too had an outstanding rote memory, being able to reproduce a two-and-a-half-page article verbatim after a single silent reading, but without being able to grasp the social significance of the passage (Anastasi & Levee, 1960).

As Horn (1973) points out one of the problems of investigating the relationship between auditory abilities and intelligence is that existing intelligence tests fail to do justice to input through the auditory modality.

Musical ability must involve temporal integration, "perhaps," according to Hearnshaw (1951) "the most basic intellectual skill." In Stankov & Horn's researches on the temporal integration and reasoning aspects of auditory functions, they found two factors: Tr and Ra (Horn, 1973). A test that particularly defined Tr was one which required a subject to retain awareness of a sound while listening to determine when the same sound occurred again in a series. Seashore's time test correlated .27 with this factor. Stankov (1980) has investigated this factor, now called Tc (Temporal Tracking), further. To study the reasoning processes involved in tasks of auditory input, Stankov and Horn devised tasks parallel to visual reasoning tasks such as analogies and classifications. For example, a chord of three notes was followed by four sets of three notes played individually, the subject having to select the one that contained the same three notes that were in the original chord. Such tests came together in the Ra factor. This was also defined by variables selected to indicate closure, in their attempt to test the hypothesis that auditory and visual flexibility and speed of closure might represent similar mental functions. A test designed to imitate

the Gottschaldt figures test loaded .49 on the Ra factor and the Wing chord analysis test, .42. The Gottschaldt tonal test required the subject to choose one of four pairs of piano notes that had been embedded in a previously present chord.

Perhaps the test most successful in producing a clear-cross modality measurement was called "hidden tunes" (a test similar to Davies' described in Section II,D). In White's (1954) study it correlated .66 with "embedded figures" (its visual counterpart) and sizably with three other visual tests. Fleishman, Roberts, and Freidman (1958) found that "hidden tunes" had a major loading (.62) on a factor that involved tests of ability to detect the dots and dashes of the Morse code. Since it also loaded .56 on the Seashore rhythm test, they interpreted it as a factor of auditory rhythm perception.

Horn's (1973) factorization of the intercorrelations among the auditory and visual primary factors produced four second-order ones. In the first, Mr, Tr, and Ra appeared with five visual primaries. Horn interpreted this as a factor indicating fluid intelligence. Factor 2 reflected a crystallized form of intelligence also manifested in visual primaries and in Ra, Tr, and DDS. A third factor was concerned with visualization. The fourth was almost wholly defined by the auditory primaries. Mr had the most prominent correlation with this broad auditory Ga factor, followed by MSC (comprehension of speech distorted—e.g., by masking) with a correlation of .74. The loadings of Ra, DDS, and Tr, each of which had prominent loadings on the Gf or Gc factors, were .59, .49, and .43, respectively.

It seemed that the distinction between Gf and Gc was manifested in the auditory as well as in the visual realm. However, the auditory tests contained a systematic variance of their own. The perception keenness and retention represented by MSC and Mr suggested to Horn these were more important than temporal ordering in his Ga factor. However, his musically unselected adults may well have been operating at a rather basic "melodic" rather than "harmonic" level. It would be interesting to investigate whether different results would be found among trained musicians. For a more detailed discussion of researches of Horn and Stankov and their implications for music, see Shuter-Dyson and Gabriel (in press).

Two areas to which musical ability might seem to be especially associated are mathematical/spatial abilities and ability in the other arts. Shuter (1968) discussed the difficulties of proving or disproving the traditional view of a connection between musical and mathematical abilities. More recently, the view that some musical abilities and some mathematical or spatial abilities are more efficiently processed in the right cerebral hemisphere might give renewed credence to their association. In his earlier research Karma's test did seem to be more highly correlated with spatial abilities than with verbal (Karma, 1978). He hypothesized that some persons (perhaps females) might process musical stimuli verbally and others (perhaps males) spatially. His later research with a revised form of his test, however, pointed to the tentative conclusion that the essential difference between "verbal" and "spatial" types of musical processing is the amount of musical training. For a review of studies of musical training and lateralization, see Shuter-Dyson and Gabriel (in press, Chapter 18).

The parents of great musicians have often been distinguished in the other arts (Feis, 1910). However, attempts to obtain empiric evidence of a relationship through testing

have been disappointing. Guilford (1957)—while considering that the ability to produce and express ideas in the graphic arts was distinct from the ability to produce and express ideas in music—suspected that there might be something in common among the parallel factors in the different arts.

It is possible that connections between the arts exist on a personality rather than on a cognitive level. At least one of the Cattell 16 PF factors, apart from "B" (intelligence) has been consistently found in a number of studies of personality and musical abilities (Cattell, Eber, & Tatsuoka, 1970; Shuter, 1975). This is the "I" (tender-mindedness) factor. This factor was shown both by musical children and professional music students, but it did not clearly distinguish them from nonmusical groups; rather it was common to all whose interests lay in the arts (Martin, 1977). Beldoch (1964) found highly significant correlations among sensitivity to emotional expressions in speech (recognition of emotion expressed by the voice, given a neutral verbal content), in graphic art (abstract representations of the same 10 emotions), and music (short compositions depicting the same emotions). Though verbal intelligence was related to all three measures, significant intercorrelations remained after intelligence had been partialed out. This result Beldoch interpreted with reference to Langer's (1942, 1953) theory of discursive and nondiscursive symbols. On the general question of the personality structure of the musician, Kemp (1981) reports the results of an extensive investigation.

In summary, however valuable intelligence may be in the development of musical ability, mere intellectual efficiency, however highly oriented to music, will not make a musician. During the later years of the twentieth century, music is likely to become an important leisure interest, one for which education should be provided. Perhaps test producers may be induced to accord auditory/musical abilities a more prominent place in their batteries. If a second-order factor (or factors) were to be extracted from the results of testing, it would account nicely for the existence of a number of musical abilities in varying degrees in individuals and also for the need for an overall, *coordinated* level of efficiency for success in music whether as listener, performer, or composer.

## REFERENCES

Agnew, Marie. A comparison of auditory images of musicians, psychologists and children. *University of Iowa Studies in Psychology*, 1922, *8*, 268–278.

Anastasi, Anna, & Levee, R. E. Intellectual defect and musical talent. *American Journal of Mental Deficiency*, 1960, *64*, 695–703.

Andrews, Frances M., & Deihl, N. C. *Development of a technique for identifying elementary school children's musical concepts*. Washington, D.C.: U.S. Office of Education Project 5-0233, 1967.

Beldoch, M. Sensitivity to expression of emotional meaning in three modes of communication. In J. R. Davitz (Ed.), *The communication of emotional meaning*. New York: McGraw-Hill, 1964.

Bentley, A. *Measures of musical ability*. London: Harrap, 1966. (a)

Bentley, A. *Musical ability in children and its measurement*. London: Harrap, 1966. (b)

Bentley, R. R. *A critical comparison of certain aspects of musical aptitude test*. Doctoral thesis, Univ. Southern California. 1955.

Bergan, J. R. The relationship among pitch identification, imagery for music sounds, and music memory. *Journal of Research in Music Education*, 1967, *15*, 99-109.
Blacking, J.A.R. Towards a theory of musical competence. In E. DeJager (Ed.), *Man: Anthropological essays in honour of O. F. Raum*. Cape Town: Struik, 1971.
Bugg, E., & Herpel, L. The significance of tonal memory for musicality. *Journal of General Psychology*, 1946, *35*, 3-15.
Cattell, R. B. Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 1963, *54*, 1-22.
Cattell, R. B. *The Scientific Analysis of Personality*. Baltimore: Penguin. 1965.
Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. *Handbook for the Sixteen Personality Factor Questionnaire (16 PF)*. Champaign, Illinois: Institute for Personality & Ability Testing, 1970.
Colwell, R. *MAT Music achievement tests 1 and 2, Interpretive Manual*. Chicago, Illinois: Follett, 1969.
Colwell, R. *The evaluation of music teaching and learning*. Englewood Cliffs, New Jersey: Prentice Hall, 1970.
Davies, J. B. New tests of musical aptitude. *British Journal of Psychology*, 1971, *62*, 557-565.
Davies, J. B. *The Psychology of music*. London: Hutchinson, 1978.
Donat, Kinga. *Testing Musical Ability*. Unpublished paper, University of London, Institute of Education, 1964.
Drake, R. M. *Tests of musical talent*. Ph.D. thesis, University of London, 1931.
Drake, R. M. Factorial analysis of music tests by the Spearman-Tetrad-Difference technique. *Journal of Musicology*, 1939, *1* (1), 6-16.
Drake, R. M. The relation of musical talent to intelligence and success at school. *Journal of Musicology*, 1940, *2*, No 1.
Drake, R. M. *Manual for the Drake Musical Aptitude Tests*. (2nd ed.) Chicago, Illinois: Science Research Associates, 1957.
Edmunds, C. B. *Musical ability, intelligence and attainment of secondary modern and E.S.N. children*. Thesis, University of Leeds, 1960.
Eisenberg, Rita B. *Auditory competence in early life*. Baltimore, Maryland: University Park Press, 1976.
Farnsworth, P. R. *The social psychology of music*. (2nd ed.) Ames: Iowa State University Press, 1969.
Faulds, B. D. *The perception of pitch in music*. Princeton, New Jersey: Educational Testing Service, 1959.
Feis, O. *Studien über die Genealogie und Psychologie der Musikers*. Wiesbaden: Bergmann, 1910.
Fleishman, E. A., Roberts, N. H., & Freidman, M. P. Factor analysis of aptitude and proficiency measures in radio-telegraphy. *Journal of Applied Psychology*, 1958, *42*, 127-137.
Francès, R. *La perception de la musique*. Paris: Vrin, 1958.
Franklin, E. *Tonality as a basis for the study of musical talent*. Göteborg: Gumperts Förlag, 1956.
Gaston, E. T. *A test of musicality*. Lawrence, Kansas: Odell's Instrumental Service, 1957.
Gates, Anne, Bradshaw, J. L., & Nettleton, N. C. Effect of different delayed auditory feedback intervals on a music performance task. *Perception & Psychophysics*, 1974, *15*, 21-25.
Gordon, E. *Musical aptitude profile*. Boston, Massachusetts: Houghton Mifflin, 1965.
Gordon, E. *A three-year longitudinal predictive validity study of the musical aptitude profile*. Iowa City, Iowa: University of Iowa Press, 1967.
Gordon, E. An investigation of the intercorrelation among musical aptitude profile and Seashore measures of musical talent subtests. *Journal of Research in Music Education*, 1969, *17*, 263-271.
Gordon, E. *Iowa tests of music literacy*. Iowa City, Iowa: Bureau of Educational Research and Service, 1970.
Gordon, E. Fourth-year and fifth-year final results of a longitudinal study of the musical achievement of culturally-disadvantaged students, *Experimental Research in the Psychology of Music: Studies in the Psychology of Music*, 1975, *10*, 24-52.
Gordon, E. E. *Learning sequence and patterns in music*. Chicago, Illinois: G.I.A., 1977.
Gordon, E. E. *Primary measures of music audiation*. Chicago, Illinois: G.I.A., 1979.
Gould, A. O. Developing specialized programs for singing in the elementary schools. *Bulletin of the Council for Research in Music Education*, 1969, No 17, 9-22.
Guilford, J. P. Creative abilities in the arts. *Psychological Review*, 1957, *64*, 110-118.
Hearnshaw, L. S. Exploring the intellect. *British Journal of Psychology*, 1951, *42*, 315-321.
Hebb, D. O. *The organization of behavior*. New York: Wiley, 1949.

Biblioteca del Real Conservatorio Superior de Música

Hiriartborde, E. *Les aptitudes rythmiques.* Paris: Dactylo-Sorbonne, 1964.
Hoffren, J. The construction and validation of a test of expressive phrasing in music. *Journal of Research in Music Education*, 1964, *12*, 159–164.
Holmstrom, L-G. *Musicality and prognosis.* Uppsala: Almqvist & Wiksells, 1963.
Holmstrom, L-G. Intelligence vs. progress in music education. *Journal of Research in Music Education*, 1969, *17*, 76–81.
Horn, J. L. Theory of functions represented among auditory and visual test performances. In J. R. Royce (Ed.), *Multivariate analysis and psychological theory.* London: Academic Press, 1973. Pp. 203–239.
Horn, J. L. & Stankov, L. Auditory and visual factors in intelligence. In preparation, 1981.
Igaga, J. M. *A comparative developmental study of the rhythmic sensitivity of Ugandan and English schoolchildren.* Ph.D. thesis, University of London, 1974.
Imberty, M. *L'acquisition des structures tonales chez l'enfant.* Paris: Klincksieck, 1969.
Joyner, D. R. The monotone problem. *Journal of Research in Music Education*, 1969, *17*, 115–124.
Karlin, J. E. Musical ability. *Psychometrika*, 1941, *6*, 61–65.
Karlin, J. E. A factorial study of auditory function. *Psychometrika*, 1942, 7, 251–279.
Karma, K. *The ability to structure acoustic material as a measure of musical aptitude.* Helsinki: Institute of Education, Research Bulletin, No 43, 1975.
Karma, K. Musical, spatial and verbal abilities. Paper to the *7th International Seminar on Research in Music Education, Bloomington, Indiana*, 1978.
Kemp, A. The personality structure of the musician. I. Identifying a profile of traits for the performer. *Psychology of Music*, 1981, *9(1)*, 3–14. II. Identifying a profile of traits for the composer. *Psychology of Music*, 1981, *9(2)*, (in press).
Kwalwasser, J. *Exploring the musical mind.* New York: Coleman-Ross, 1955.
Kwalwasser, J., & Dykema, P. W. *Kwalwasser-Dykema Music tests.* New York: Fischer, 1930.
Kyme, G. H. A study of the development of musicality in the junior high school and the contributions of musical composition to this development. *Bulletin of the Council for Research in Music Education*, 1967, No 10, 15–23, No 11, 36–45.
Langer, Susanne K. *Philosophy in a new key.* Cambridge, Massachusetts: Harvard University Press, 1942.
Langer, Susanne K. *Feeling and form.* New York: Scribners, 1953.
Long, N. H. *Indiana-Oregon Music Discrimination Test.* Bloomington, Indiana: Midwest Music Tests, 1978.
Lundin, R. W. A preliminary report on some new tests of musical ability. *Journal of Applied Psychology*, 1944, *28*, 393–396.
Lundin, R. W. The development and validation of a set of musical ability tests. *Psychological Monograph*, 1949, *63(305)*, 1–20.
Lundin, R. W. What next in the psychology of musical measurement? *Psychological Record*, 1958, *8*, 1–6.
Lundin, R. W. *An objective psychology of music.* (2nd ed.) New York: Ronald Press, 1967.
McLeish, J. The validation of Seashore's measures of musical talent by factorial methods. *British Journal of Psychology (Statistical section)*, 1950, *3*, 129–140.
McLeish, J. The factor of musical cognition in Wing's and Seashore's Tests. Paper represented to *Music Education Research Conference, Reading University*, 1966.
Mainwaring, J. Kinaesthetic factors in the recall of musical experience. *British Journal of Psychology*, 1933, *23*, 284–307.
Martin, P. J. *Appreciation of music in relation to personality factors.* Ph.D. thesis, University of Glasgow, 1976.
Martin, P. J. The personalities of musicians. Paper to the *9th Conference of the Society for Research in Psychology of Music and Music Education, Cambridge*, 1977.
Miller, G. A., & Selfridge, J. A. Verbal content and recall of meaningful material. *American Journal of Psychology*, 1950, *63*, 176–185.
Minogue, B. A case of secondary mental deficiency with musical talent. *Journal of Applied Psychology*, 1923, 7, 349–352.
Moog, H. *The musical experience of the pre-school child.* London: Schotts, 1976.
Moog, H. The perception of rhythmic forms by physically handicapped children and those of low intelligence in comparison with non-handicapped children. Paper to the *7th International Seminar on Research in Music Education, Bloomington, Indiana*, 1978.

Moorhead, G. E., & Pond, D. *Music of Young Children: II General observations.* Santa Barbara, California: Pillsbury Foundation Studies, 1942.

Mursell, J. L. *The psychology of music.* New York: Norton, 1937.

Phillips, D. An investigation of the relationship between musicality and intelligence. *Psychology of Music.* 1976, *4*(2), 16-31.

Rainbow, E. L. A pilot study to investigate the constructs of musical aptitude. *Journal of Research in Music Education,* 1965, *13,* 3-14.

Revesz, G. *The psychology of a musical prodigy.* New York: Harcourt, 1925.

Revesz, G. *Introduction to the psychology of music.* London: Longmans, Green, 1953.

Rife, D. C., & Snyder, L. H. A generic refutation of the principles of behavioristic psychology. *Human Biology,* 1931, *3,* 547-559.

Scheerer, M., Rothmann, E., & Goldstein, K. A case of "idiot savant": an experimental study of personality organization. *Psychological Monographs,* 1945, *58,* No 4.

Scherer, K. R., & Oshinsky, J. Cue utilization in emotion attribution from auditory stimuli. *Motivation & Emotion,* 1977, *1,* 331-346.

Seashore, C. E. *Seasbore Measures of Musical Talent.* Chicago, Illinois: Stoelting, 1919. Revised and retitled *Measures of Musical Talents* 1939 and 1960. *Revised Manual* New York: The Psychological Corporation, 1960.

Seashore, C. E. *The psychology of music.* New York: McGraw Hill, 1938.

Sergeant, D. C. *Pitch perception and absolute pitch—a study of some aspects of musical development.* Ph.D. thesis, University of Reading, 1969.

Sergeant, D. C., & Thatcher, G. Intelligence, social status and musical abilities. *Psychology of Music,* 1974, *2*(2), 32-57.

Shuter, Rosamund. *The psychology of musical ability.* London: Methuen, 1968.

Shuter, Rosamund. Singing out of tune: A review of recent research on this problem and its educational implications. *Scientific Aesthetics,* 1974, *9,* 115-120.

Shuter, Rosamund. The relationship between musical abilities and personality characteristics in young children. In New Zealand Council for Educational Research, Studies in Education No 25, *Music education for the very young Child,* Wellington, 1975. P. 30.

Shuter-Dyson, Rosamund, & Gabriel, C. *The psychology of musical ability.* 2nd edition. London: Methuen (in press).

Siegel, Jane A. Judgment of intonation by musicians: Further evidence for categorical perception. London: University of Western Ontario Research Bulletin, No. 375, 1976.

Siegel, Jane A., & Siegel, W. Absolute identification of notes and intervals by musicians. *Perception & Psychophysics,* 1977, *21,* 143-152. (a)

Siegel, Jane A., & Siegel, W., Categorical perception of tonal intervals: musicians can't tell *sharp* from *flat. Perception & Psychophysics,* 1977, *21,* 399-407. (b)

Siegel, W. The paradox of pitch categories. London: University of Western Ontario Research Bulletin No. 373, 1976.

Stankov, L. Ear differences and implied cerebral lateralization in some intellective auditory factors. *Applied Psychological Measurement,* 1980, *4,* 21-38.

Stankov, L., & Horn, J. L. Human abilities revealed through auditory tests. *Journal of Educational Psychology,* 1980, *72,* 19-42.

Swanick, K. Musical cognition and aesthetic response. *Psychology of Music,* 1973, *1* (2), 7-13.

Tanner, J., & Loess, H. Correlations among rhythm tests. *Perceptual & Motor Skills,* 1967, *25,* 721-726.

Tarrell, V. V. An investigation of the validity of the musical aptitude profile. *Journal of Research in Music Education,* 1965, *13,* 195-206.

Teplov, B. M. *Psychologie des aptitudes musicales.* Paris: Presses Universitaires de France, 1966.

Thackray, R. *An investigation into rhythmic abilities,* Music Education Research Papers No 4. London: Novello, 1969.

Thackray, R. *Rhythmic abilities in children,* Music Education Research Papers No 5. London: Novello, 1972.

Tomita, M., & Kurosu, M. A preliminary study on the assessment of musical aptitude. Tokyo: Waseda University Psychological Report 8, 1976.

Vaughan, M. M. Musical creativity: Its cultivation and measurement. *Bulletin of the Council for Research in Music Education*, 1977, No 50, 72-77.

Vernon, P. E. The psychology of music with especial reference to its appreciation. Ph.D. dissertation, University of Cambridge, 1931.

Vernon, P. E. What is potential ability? *Bulletin of the British Psychological Society*, 1968, *21*, 211-219.

Whellams, F. The aural musical abilities of junior school children: A factorial investigation. Ph.D. thesis, University of London, 1971.

White, B. W. Visual and auditory closure. *Journal of Experimental Psychology*, 1954, *48*, 234-240.

Whybrew, W. E. *Measurement and evaluation in music* (2nd ed.) Dubuque: Brown, 1971.

Wing, H. D. A factorial study of musical tests. *British Journal of Psychology*, 1941, *31*, 341-355.

Wing, H. D. Tests of musical ability and appreciation. *British Journal of Psychology, Monograph Supplement*, No 27, 1948.

Wing, H. D. Some applications of test results to education in music. *British Journal of Educational Psychology*, 1954, *24*, 161-170.

Wing, H. D. *Standardised tests of musical intelligence*. Windsor, England: N.F.E.R., 1961.

Young, W. T. The Bentley Measures of Musical Abilities: A congruent validity report. *Journal of Research in Music Education*, 1973, *21*, 74-79.

Zenatti, Arlette. *Le développement génétique de la perception musicale*, Monographies Françaises de Psychologie 17, Paris: CNRS, 1969.

Zenatti, Arlette. Etude de l'acculturation musicale chez l'enfant dans une épreuve d'identification mélodique. *Journal de Psychologie normale et pathologique*, 1973, *4*, 453-464.

Zenatti, Arlette. Melodic memory tests: A comparison of normal children and mental defectives. *Journal of Research in Music Education*, 1975, *23*, 41-52.

Zuckerkandl, V. *Man the musician*. Princeton, New Jersey: Princeton University Press, 1973.

# 13

# Melodic Information Processing and Its Development

*W. Jay Dowling*


I. Introduction ..... 413
II. Development ..... 415
A. Infancy ..... 415
B. Childhood ..... 416
C. Training ..... 419
D. School-Age Years ..... 420
III. Adult Memory ..... 421
A. Melodic Features ..... 421
B. Immediate Memory ..... 423
C. Tonality ..... 424
D. Long-Term Memory ..... 426
IV. Contour versus Interval ..... 427
V. Summary ..... 427
References ..... 428


## I. INTRODUCTION

In this chapter I review what we presently know about humans' processing of pitch information in melodies. The time has come for such a summary because in the 25 years since Francès' (1958) classic work an increasingly clear picture of how adults hear and remember melodies has been emerging, along with a clearer picture of how those abilities develop. I will start by following the developmental pattern from infancy to adulthood, and then I will analyze the adult behaviors in more detail.

Melodies have properties that musicians and psychologists can vary more or less independently of one another. The pitches of a melody may be preserved when it is



Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

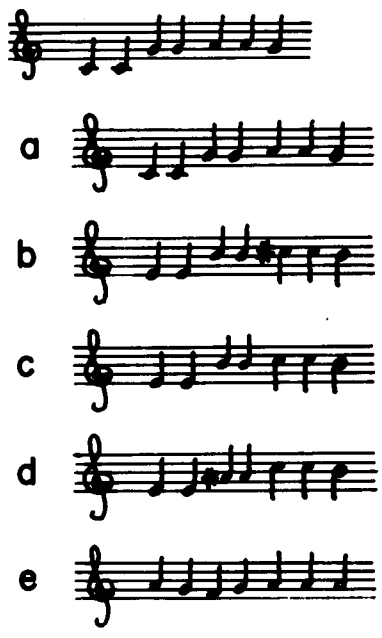



**Fig. 1.** Examples of types of stimuli described in the text. At the top is the first phrase of a familiar melody, "Twinkle, Twinkle, Little Star." Following it are (a) an exact repetition; (b) a transposition to another key; (c) a tonal imitation in the key of the original; (d) an atonal imitation; and (e) a melody with a different contour.

repeated exactly (Fig. 1A), but it is more usual for melodies to be repeated at a variety of pitch levels. In such cases the identity of the melody is preserved as long as the pitch *intervals* between notes have been preserved through transposition to a new *key* along a logarithmic scale of frequency (Fig. 1B). A further change, one that does not preserve the identity of the melody though it does result in something similar, is to preserve the melodic *contour* (the pattern of ups and downs of pitch) while allowing the interval sizes to change (a tonal imitation, Fig. 1C). Examples 1A and 1C share melodic contour and key (tonality) but not interval sizes. Fig. 1D shows the result of taking the same contour but changing the intervals even more radically so that the result lies outside any tonal scale and is *atonal*. Finally, Fig. 1E displays the result of altering the contour, leaving only the rhythmic pattern of the original song and whatever nonspecific features it might have, such as general pitch range, tempo, and instrumental timbre. These latter features are nonspecific in that they might be properties of whole families of similar songs.

The kinds of features preserved through the transformations of Fig. 1 evoke different patterns of response in adults asked to recognize similarities and differences among them. And in the developmental sequence the child's abilities to utilize these different sorts of stimulus invariance stabilize at different ages. In early infancy the child can reproduce single pitches and can notice a change of melodic contour (but not

of key). It is not until the age of 5 or 6, however, that the child develops a stable sense of tonality; that is, until he or she can detect when the key of a melody has been changed and distinguish tonal from atonal melodies. It is still later that the child becomes sensitive to small changes of interval sizes in familiar songs—that is, to the differences between examples 1A and 1C. The developmental sequence is one of going from the baby's ability to distinguish gross features such as contour and pitch level, to the 5-year-old's grasp of tonal scales and ability to discriminate key changes, to the adult's ability to detect small changes of interval size. There seems to be a hierarchy of features, and the sequence of human development follows that hierarchy closely (Pick, 1979). I will now review the evidence in more detail.

## II. DEVELOPMENT

### A. Infancy

Infants 5 months old notice changes in melodic contour while treating transpositions of melodies as equivalent. Melson and McCall (1970) demonstrated the first of these results, finding dishabituation of the baby's heartrate response when the temporal positions of all eight notes of an eight-note pattern were changed. And Kinney and Kagan (1976) found the same effect with 7-month-olds when the contour *and rhythmic pattern* of a melody were altered. Chang and Trehub (1977a) demonstrated both aspects with 5-month-olds. They presented babies with six-note melodies played at a rate of 2.5 notes/sec. The melodies were atonal and had rather large pitch intervals between notes. The pitches lay in a two-octave range covering the treble staff from the C below to the C above (262 to 1048 Hz). One melody was played 30 times over a 5.5 min period, during which the baby's heart rate habituated to the melody. (Only babies showing initial heartrate deceleration followed by habituation were used—about half the babies tested). After the 30 repetitions the pattern was switched to either an exact transposition of the first melody (up or down a minor third) or a permutation of the notes of the transpostion that destroyed the contour of the original. (Note that the transposition left the test melody in more or less the same pitch range as the original except that all the particular pitches had been changed.) Chang and Trehub found that the babies showed dishabituation—a reappearance of the heartrate deceleration "startle" response—to the melody with the altered contour, but not to the transposed melody. That is, for infants as for adults the transposition sounds like the same old melody again while the different-contour melody sounds new. (In a related study Chang and Trehub (1977b) also found dishabituation to a change in rhythmic pattern.) Thus, Chang and Trehub demonstrated both the noticing of melodic discrepancy (contour changes) and the acceptance of melodic similarity (transposition).

Vocal play in which infants explore the pitch range accessible to their voices has often been observed (Ostwald, 1973), and some infants even tried to match pitches they heard (Révész, 1954, p. 171; Shuter, 1968, p. 67). In a more controlled study,



Kessen, Levine, and Wendrich (1979) have demonstrated that with a few brief training sessions babies under 6 months of age can learn to match pitches, singing back the pitches sung to them. Using test pitches from the minor triad D-F-A, the experimenters found that the babies sang successful matches two-thirds of the time, and a large proportion of these hits were very nearly in tune. According to Kessen *et al.*, the babies enjoyed the task and worked hard at it. Apparently the singing of specific pitches is something that the 6-month-old is ready and willing to learn. In both noticing contour changes and matching pitches, the infant's behavior is already like that of the adult. But in still another way it is very adultlike: in Chang and Trehub's task in which absolute pitch level was irrelevant to the identity of the melody the infants ignored changes in pitch. Yet they were quite capable of discriminating pitches when the task required it, as Kessen *et al.* demonstrated. Of course, the infant is insensitive to many dimensions accessible to the adult, including tonality and interval size, and so I will continue the story.

### B. Childhood

During their second year children begin to recognize certain melodies as stable entities in their environments. My daughter at 18 months would run to the TV set when she heard the "Sesame Street" theme come on but not for other melodies. At 20 months, after a week or so of going around the house singing "uh-oh" rather loudly to a descending minor third, she responded with the spoken label "uh-oh" when I played the pattern on the piano. Especially after 18 months the child begins to generate recognizable, repeatable songs. As Ostwald (1973) says, "Song patterns become more coherent." Songs for a child around the age of 2 years often consist of brief phrases repeated over and over. The rhythm of these phrases is coherent, and their contour is replicable, but the pitch wanders. The same contour is repeated at different pitch levels, usually with different intervals between the notes. An example from a 24-month-old[1] consisted of a descending phrase with the words "Duck on my house" repeated 10 or 12 times at different pitch levels with small intervals within the phrase and a leap up to the start of a new phrase. The song was repeated by the child over a period of two weeks at which point it disappeared. (The contour was probably borrowed from the "E-I-E-I-O" phrase of "Old MacDonald," which the child also sang.) Such spontaneous songs have a systematic form and display two essential features of adult singing: they use discrete pitch levels, and they use repetition of rhythms and melodic contours as a formal device. They are unlike adult songs, however, because they lack a framework of stable pitches (a scale) and use a very limited

[1]This observation and following ones are drawn from tape recordings of the singing of 21 children: four of age 2; six of age 3; five of age 4; and six of age 5; with between 5 and 32 "songs" collected from each. Such sampling does not lend itself to statistical analysis, and the results reported are intended as illustrative. I have usually used examples that find substantiation elsewhere in the literature. And as with all such observations the ages mentioned should not be taken as normative in a rigid sense since children vary immensely in the age at which they display these behaviors.

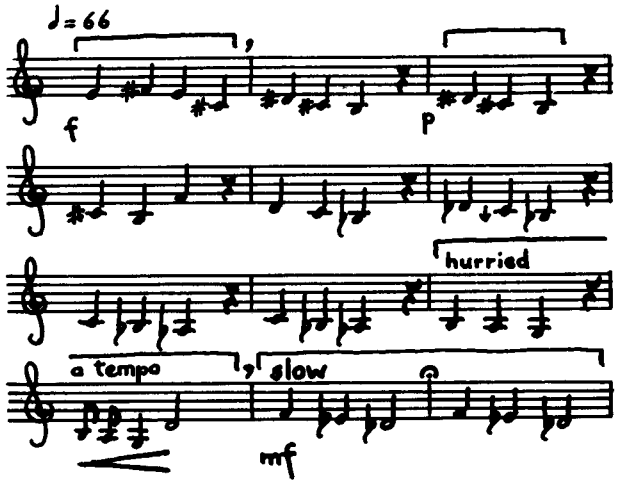



**Fig. 2.** A child's spontaneous song at 32 months. Each note was vocalized to the syllable "Yeah." Brackets indicate regions of relatively accurate intonation. Elsewhere intonation wandered.

set of contours in one song—usually just one or two. A more sophisticated construction by the same child at 32 months can be seen in Fig. 2. The pitch still wanders but is locally more coherent within phrases. Here three replicable phrases are built into a song structure that begins to be meaningful for the adult listener in a way that the "Duck on my house" song is not. The child has moved from Frank Zappa somewhat in the direction of the Beatles.

The above observations are in general agreement with those of Gardner, Davidson, and McKernon (1981) on spontaneous singing by 2-year-olds. Gardner *et al.* extended their observations by teaching a simple song to children across the preschool age range. Two- and three-year-olds generally succeeded in reproducing the contours of isolated phrases. As they got older they were able to concatenate more phrases in closer approximation of the model. It was only very gradually that the interval relationships of the major scale began to stabilize. Four-year-olds could stick to a stable scale pattern within a phrase but would often slip to a new key for the next phrase. It was not until after 5 that the children could hold onto a stable tonality throughout the song. Further, with a little practice the 5-year-olds were able to produce easily recognizable versions of the model. My own observations suggest that the typical 5-year-old has a fairly large repertoire of nursery songs of his or her culture. This phenomenon appears when children are asked to sing a song and can respond with a great variety of instances as well as when they are asked to perform short-term recognition-memory tasks and do better with familiar materials than with novel, randomly generated melodies.

A converging line of evidence on the 5-year-old's acquisition of a stable scale structure, of tonality, is provided by a study by Bartlett and Dowling (1980, Experiment 4). That study explored subjects' ability to distinguish between transpositions of

familiar melodies (Fig. 1B) and tonal imitations retaining the same contour but with changes of interval size (Fig. 1C). On each trial of the experiment the first phrase of a familiar melody would be presented, followed shortly afterwards by a comparison stimulus that was either a transposition or a tonal imitation. The comparison was either in a nearly related (or the same) key as the standard or in a distant key. Near keys share many overlapping pitches in their scales; distant keys share few. For example, starting with the key of C major (CDEFGABC), we could move to the hear key of G (GABCDEF♯G) that shares every note with C but the F♯ or to the far key of B major (BC♯D♯EF♯G♯A♯B) that shares only two pitches with C: the B and E. Table I summarizes the results in terms of the tendency to say "Same" when presented with the various types of stimulus pair. The responses of nonmusician adults are presented for comparison. Note that the adults are highly accurate in saying "Same" to transpositions (over 90%) and not saying "Same" to imitations (under 10%). The pattern for the 5-year-olds is very different: they tend to say "Same" to near key comparisons whether transpositions or imitations and to avoid "Same" responses to far-key items. That is, the 5-year-olds have one component of the adult behavior pattern—the ability to distinguish near from far keys—but not the other component—the ability to detect changes of interval sizes in the tonal imitations. Note that both components are reflected in the adults' better performance in rejecting far-key (as opposed to near-key) imitations, a result that Bartlett and I obtained in a series of experiments using a variety of materials and that we call a *key-distance* effect (Bartlett & Dowling, 1980). Note also that as tbe child grows older, the pattern of response changes markedly in the adult direction so that the 8-year-old child accepts near-key *imitations* less often than far-key *transpositions*. Both cognitive strategy components are clearly visible in the 8-year-olds' data: they can use both key distance and interval changes to reject a comparison stimulus.

As with all of these ages, 5 years of age should not be taken as rigidly normative but as an approximate guideline for when stable tonal structures appear, depending on how the experimenter tests for them. Zenatti (1969, Chap. 3), for example, found that with the rather difficult task of saying *which* note of a three-note sequence has been altered in pitch, 5-year-olds perform at about chance level with both tonal and atonal

TABLE I

Probabilities of Responding "Same" to Transpositions and Tonal Imitations in Near and Far Keys[a]

| | Transpositions | | Imitations | |
|---|---|---|---|---|
| Age group (yr) | Near | Far | Near | Far |
| 5.6 | .67 | .58 | .63 | .55 |
| 6.9 | .58 | .54 | .60 | .36 |
| 8.6 | .70 | .62 | .50 | .35 |
| 29.4 | .91 | .92 | .11 | .02 |

[a]After Bartlett and Dowling, 1980.

melodies, and only with 6- and 7-year-olds does superior performance with tonal melodies appear. And in keeping with the result of Bartlett and Dowling (1980), Imberty (1969, ch. 4) found that 7-year-olds could notice sudden changes of key in the middle of familiar tunes and that 8-year-olds could tell when a melody had been switched from the major mode to the minor, an effect very similar to that produced by creating a tonal imitation in the above sense.

The 5- or 6-year-old's grasp of stable tonal centers fits other results in the literature. For example, in a series of studies Riley and McKee (1963; Riley, McKee, & Hadley, 1964; Riley, McKee, Bell, & Schwartz, 1967) found that first graders have an overwhelming tendency to respond to pitch-interval comparison tasks by choosing the same pitch rather than the same interval. This tendency to respond to the pitch tasks in terms of a stable frame of reference contrasted with the same children's ability to respond to loudness-comparison tasks in terms of *relative* (not absolute) loudness. And the emergence of tonal scale relationships among the child's cognitive structures has implications for the conduct of research. The fact that Chang and Trehub (1977a) used atonal materials had little impact on their results since babies do not respond to tonal scale structures as such. But Wohlwill's (1971) use of atonal and (to the adult ear) rather strange sounding melodies very probably led to his result that first graders could distinguish targets from different-contour lures at a level barely better tban chance (in a long-term memory task). When first graders are asked to perform a short-term memory task with tonal materials in which they have to distinguish between repetitions of the target and lures containing the same pitches but in permuted order (that is, with different contour), they get about 65% correct compared with a chance level of 50% (Dowling & Goedecke, in preparation). At any rate, Wohlwill's (1971, p. 228) conclusion that "The establishment of pitch as a directional dimension is a relatively late phenomenon" could not possibly be true in the light of Chang and Trehub's result. What is true is that first graders have trouble with their language descriptions of pitch direction, as Zimmerman and Sechrest (1970; Pflederer & Sechrest, 1968) and Hair (1977) have documented. Words like "up" and "down" are very difficult for children of 5 or 6 to learn to apply to pitch in the way we adults do.

### C. Training

As with the development of any information processing skill, it is useful to know about the impact of training at various ages. Dowling and Goedecke (in preparation) studied first and third graders who had been in an instrumental music program for 6 to 8 months, as compared with children who were on the waiting list for the program but could not participate. The music program was designed to emphasize auditory information processing skills (as distinct from other skills such as reading music). Each child performed two short-term recognition memory tasks: one in which repetitions of a brief, novel, tonal melody were contrasted with melodies different in both their contours and in the pitches they contained; and another in which repetitions were contrasted with different-contour melodies containing the same pitches but in dif-

ferent order. This second task, which was harder than the first, was designed as a test of the child's ability to conserve the temporal order of a series of auditory events. On the first task, untrained first graders got about 75% correct (chance was 50%). After training, however, performance was around 90% correct. (This task was quite easy for the third graders—with training their performance went from somewhat under to somewhat over 90%.) Thus, first graders seem quite ready to learn more skills involving the discrimination of melodic contour. With the more difficult task in which lures contained the same pitches as targets, training had less effect on the first graders with their performances going from about 65% to about 70%. Unlike the first task this task did not benefit from simple maturation: untrained third graders performed in the same range as the first graders. However, with training, the third graders improved dramatically to around 80% correct. As with other materials it appears that there are periods in the child's life when he or she can benefit more than at other times from training that focuses attention on relevant aspects of auditory materials.

### D. School-Age Years

The period between early school years and adulthood has not been studied much. One study that helps fill this gap is Zenatti's (1969), mentioned above. Zenatti studied short-term memory for sequences of three, four, and six notes with subjects ranging in age from 5 years to adulthood. On each trial a standard melody was presented at 1.5 notes/sec, followed by a comparison melody in which one of the notes of the standard had been changed by one or two semitones. The subject had to say which of the notes had been changed—as I said before, a very difficult task. With four-note atonal sequences adults got only 75% correct (versus chance of somewhere between 20 and 40%).[2] In comparison, subjects in Dowling and Fujitani's (1971, Experiment 1) study performed above 90% (with chance at 50%) on five-note, atonal melodies at a rate of 6.0 notes/sec, where they had only to say "same" or "different," but not specify the difference. What makes this task difficult is in part an effect observed by Guilford and Nelson (1937), that when one note of a melody has been altered in pitch, it sounds as though temporally adjacent notes had been changed as well. For this reason Zenatti's results cannot be compared directly with the results of studies in which essentially easier tasks were used. However, she obtained a qualitative result of great interest. For the three-note sequences, 5-year-olds perform at about chance level with both tonal and atonal stimuli. From ages 6 through 10 the results for tonal and atonal se-

[2]The appropriate baseline comparison rate is difficult to arrive at for adults. The *a priori* chance level is .20, and this is empirically validated in the younger children's data. However, for adults I believe the lowest score that we are likely to obtain really lies somewhere between 20 and 40%. This is because the task has two components: the relatively easy task of discriminating one fifth of the trials in which the melodies are the same from four fifths the trials in which they are different; and the more difficult task of saying which of the four notes was changed. Dowling and Fujitani (1971) obtained performance better than 90% on the former task. At the extremes, pure guessing on both tasks gives 20% correct while perfect performance on the former task and pure guessing on the latter would lead to 40% correct. Unfortunately, Zenatti does not provide a further breakdown of her data beyond the percent correct.

quences diverge, with the children doing better on the tonal sequences. Then around ages 12 or 13 processing of the atonal sequences catches up with the tonal. For four- and six-note sequences the same pattern appears, but the tonal and atonal difference remains across all age groups except for the most proficient adults. This result is qualitatively very similar to the "word superiority effect" in reading, in which the youngest readers take just as long to analyze nonword letter strings as they do to analyze real words while experienced readers process real words faster (Juola, Schadler, Chabot, & McCaughey, 1978). Experience with language and orthography leads the reader to recognize more efficiently stimuli that follow the well-learned patterns, processing them on a more holistic level (LaBerge, 1976). Similarly, experience with the tonal scale system leads people to improve on recognition of tonal melodies but not atonal. If the stimuli are simple, in the three-note melodies, atonal performance catches up; but the tonal system still confers an advantage on the longer sequences. The superiority of recognition performance with tonal materials has been observed several times with adults (Francès, 1958; Dowling, 1978). It is clear from Zenatti's study that the effect can be observed early in development and can be used as an index of the internalization of the scale structures of the culture.

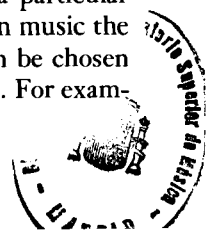
In summary, the development of melody-processing skills can be seen as a progression from the use of gross, obvious features to the use of more and more subtle features. Babies can distinguish pitch contours and produce single pitches. Around the age of 5 the child can organize songs around stable tonal centers (keys) but does not yet have a stable tonal scale system that can be used to transpose melodies accuratley to new keys. The scale system develops during the elementary school years and confers on tonal materials an advantage in memory that remains into adulthood.

## III. ADULT MEMORY

### A. Melodic Features

The same features of melodies determine adult behavior that we have seen gradually emerging in the developmental sequence. These features arrange themselves into the more global types that pertain to whole phrases and melodies and the local types that describe individual notes and note pairs (LaBerge, 1981). The more global features include the melodic contour, the rhythmic pattern, and the tonal scale of a melody. Local features include pitches and durations of individual notes and the intervals between them. The basic questions of melodic information processing concern what features will be perceived, what features will be stored in memory, and what features can be retrieved from memory. The answers to these questions vary with the context in which the features appear and with the task we give the listener.

From the point of view of music theory, the choices of features in a particular melody are not independent of each other. At the global level, in Western music the contour and the mode of the scale (for example, major versus minor) can be chosen independently, though this is not true of numerous other musical cultures. For exam-



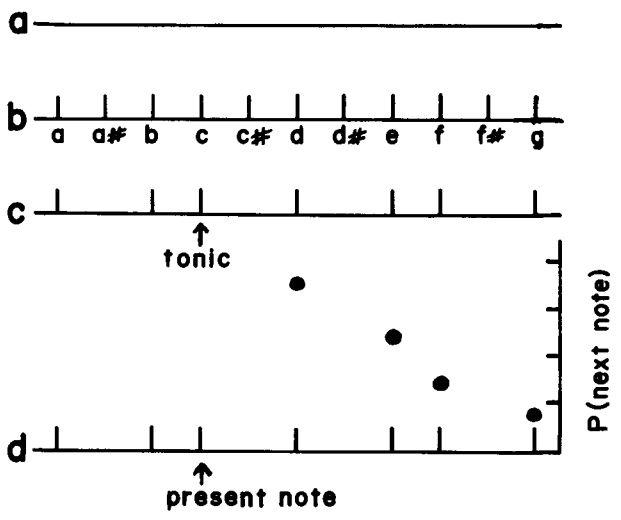



**Fig. 3.** A schematic outline of levels of pitch analysis relevant to melodic information processing: (a) the psychophysical pitch continuum, (b) tonal material used in a culture's music, (c) a tuning system and modal scale, and (d) the imposition of a melodic contour and probability distribution of interval sizes.

ple, in the music of India choosing a particular *rāg* means choosing a modal scale and at the same time limiting the choice of characteristic melodic patterns (Jairazbhoy, 1971). But even when they are independent of each other, scale and contour taken together limit the choice of individual pitches in much the same way as the global features of a word limit the choice of individual letters. The process by which this happens may be visualized in Fig. 3. All the pitches people can perceive fall on a psychophysical continuum as in Fig. 3A. This continuum is not divided into discrete steps—all the pitches of a fire siren or a baby's cry are represented.[3] The tonal material (Fig. 3B) a culture uses in its music constitutes a subset of all the possible pitches of the psychophysical continuum. In Western music the tonal material consists of the chromatic scale represented by all the notes on the piano. Choosing a tuning system and modal scale (Fig. 3C) involves selecting a subset of the pitches of the tonal material (for example, the white notes on the piano) and establishing a tonal hierarcy (for example, by selecting the scale of C major with C as tonic). When a contour is chosen, it limits further the choice of each succeeding pitch (Fig. 3D). Given the pitch of the present note of the melody, the direction of the particular

[3]This analysis draws on a more detailed description of levels of pitch coding elsewhere (Dowling, in press). The present model differs from the one Deutsch (1975) applies to similar phenomena. The main differences are as follows: (1) I include two levels of feature abstraction (namely, raw waveform data and psychophysical pitch) more primitive than her "primary array"; or, more precisely, the primary array collapses into one level what I prefer to divide into three. (2) For Deutsch, intervals are extracted directly from the primary array in a level corresponding to my "tonal material." I believe intervals *can* be so obtained, but that it is difficult. I think intervals are much more easily abstracted from the encoding level of the modal scale.

interval in the contour eliminates half the possible pitches of the next note. And the distribution of interval sizes in Western melodies (as in other cultures) favors small intervals over large, further restricting the choice (Dowling, 1968, 1978). The interdependence of global and local stimulus features could also have been described from the other direction since the choice of notes in a melody determines scale and contour. But this interdependence of features in the stimulus does not necessarily carry over to perception. It is an empirical question whether melodic contour, for example, might be stored and remembered under circumstances in which pitches and intervals are forgotten, or vice versa.

## B. Immediate Memory

I will begin my description of the listener's processing of the above sorts of features from the local end and proceed to the more global. When a person hears a sound, his or her auditory system registers an enormous amount of raw data concerning its waveform. Figure 3A represents sounds as already coded into perceived pitches, but a wealth of information is available in the stimulus that may or may not be encoded into perceived pitches, as Kubovy and Howard (1976) have demonstrated. Kubovy and Howard presented listeners with tone complexes of ambiguous pitch. Certain pitches could be disambiguated by presentation of a second-tone complex, and the disambiguating information was hidden in the fine structure of the waveform—namely in the interaural time disparity of the waveform of the target tone versus the other tones in the complex. (Interaural time disparity translates perceptually into spatial position.) Kubovy and Howard found that they could delay something like 1 sec before presenting and disambiguating tone complex, and more than half their listeners could detect the pitch relationship of the two. (One subject could wait almost 10 sec.) This means that information from the first complex remained in some sort of immediate storage without being encoded into a "perceived" pitch (of the kind represented in Fig. 3A) until the second complex arrived, a length of time that might vary considerably depending on task conditions and the listener's prior experience. It would be interesting to see how much of this information can be transformed into the perceptual features of pitch, timbre, and rhythm without mutual interference. The psychomusicological question here is, can the listener attend to the beauty of the soloist's tone and at the same time make accurate judgments about her intonation?

When pitches are coded at the psychophysical level or the level of tonal material outside of a tonal context, listeners find it difficult to integrate them into higher order units. The pitches can be remembered in and of themselves very accurately, but higher order features such as the intervals between them are difficult to extract. When Attneave and Olson (1971) presented isolated intervals to subjects and asked them to transpose them to other pitch ranges, subjects were not very successful. This contrasts with subjects' performance in the same study when asked to transpose a familiar tonal melody, which they did very well. Dowling and Fujitani (1971) found that when subjects were asked to distinguish between exact repetitions of brief atonal melodies

and imitations in which the pitches had been changed, they achieved better than 90% correct. But when subjects tried to distinguish between exact transpositions and imitations at a new pitch level having the same contour but different intervals, their performance was little better than chance. The intervals of atonal melodies—melodies in which the set of pitches is restricted to those of Fig. 3B but not to those of Fig. 3C—are hard for subjects to encode accurately. Deutsch (1979) replicated Dowling and Fujitani's result and went on to find that repeating the target melody several times at the same pitch level helps in the recognition of a transposition (as opposed to a same-contour imitation), boosting the proportion correct to 75%. Repetitions transposed to other octaves did not help as much while repetitions that destroyed melodic continuity by assigning successive pitches of the melodies to different octaves ("octave scrambling") actually hurt performance. Taken together, these results suggest that as a rule, only very specific information about pitch is extracted from atonal melodies encoded at the tonal material level. Intervals between notes are hard to encode, though exact repetition helps. The pitches and intervals of atonal melodies seem especially hard to learn. When they are remembered, it is in a relatively specific way and not easily generalized to a new context.

### C. Tonality

Intervals are easier to abstract from melodies that can be encoded in terms of a tonal scale. In an experiment similar in design to the preceding, Dowling (1978) asked subjects to distinguish between exact transpositions of novel tonal melodies and atonal, same-contour imitations (that is, between Figs. 1B and 1D). Performance was much better than that obtained by Dowling and Fujitani with atonal melodies. Moderately experienced adults achieved almost 80% correct on the task. Further, Bartlett and Dowling (1980, Experiment 2) found subjects could distinguish transpositions from far-key tonal imitations at the 75% level. This suggests that melodic intervals, difficult to encode with atonal materials, are encoded relatively well with tonal melodies so that they are recognized in transposition. Performance was still better with familiar melodies (of the sort actually shown in Fig. 1). Bartlett and Dowling (1980) found adults' performance to be better than 90% with such materials. (Some of the data are in Table I.) This agrees with the result of Attneave and Olson (1971) who found that subjects found it much easier to transpose a familiar tonal melody (the NBC chimes) than to transpose isolated intervals.

In considering the accuracy of people's memory for interval sizes, we should consider alternative ways of looking at the problem. A formal description of the "Twinkle, Twinkle" standard melody in Fig. 1 could, in terms of intervals, read:

unison, up perfect fifth, unison, up major second, unison, down major second (1)

which could be written in semitones as follows:

0, +7, 0, +2, 0, −2 (2)

Alternatively, "Twinkle, Twinkle" could be written in terms of degrees of the major scale as

$$\text{Do, Do, Sol, Sol, La, La, Sol} \tag{3}$$

or using numbers for the diatonic scale steps as

$$1, 1, 5, 5, 6, 6, 5 \tag{4}$$

Expressions (2) and (4) are formally equivalent at the level of music theory. But it remains an empirical question which of them might be closer to what people actually store in their memory when they learn a new melody. In an earlier article (Dowling, 1972) I argued on the basis of the relative difficulty subjects have with performing pitch-time transformations (such as inversion and retrograde) that their memory representations for brief melodies were more like (4) than (2). It may well be that representations of the quasi-solfeggio sort, like (4), are typical of most memory storage of tunes.

One way of summarizing the preceding series of results would be to say that exact interval sizes are remembered to the degree that the diatonic scale values of the notes are remembered. The key distance effect referred to above is best seen from this point of view. That is, in evaluating accuracy of transposition of a comparison melody, the listener typically attempts to base his or her judgment on a match to the contour of the standard and on the compatibility of the pitches he or she hears with those of the scale of the standard. Suppose the comparison were the tonal imitation of Fig. 1C. In intervals it would be represented as

$$0, +7, 0, +1, 0, -1 \tag{5}$$

In scale steps it would be

$$3, 3, 7, 7, 1, 1, 7 \tag{6}$$

retaining the scale framework of the standard in (4). With novel, tonal melodies (not "Twinkle, Twinkle"), listeners accept tonal imitations of this type as accurate transpositions provided they are in the same (or nearly the same) key as the standard (Bartlett & Dowling, 1980). Though the listener is able to match the sequence of signs of expressions (2) and (5) (the contour), he or she does not succeed well in matching the quantities (the interval sizes). Further, the listener seems to evaluate the acceptability of (6) on the basis of its sharing of pitches with the 1, 2, 3, 4, 5, 6, 7, 1 scale of (4). When the comparison melody is in a more distant key from the standard, fewer pitches are shared between the two scales and the listener is not misled by this second source of similarity. Presumably the listener is then forced to reinterpret (6) and (5) as beginning on 1 of a new mode. Bartlett and Dowling (1980, Experiment 1) found that listeners could discriminate transpositions from far-key tonal imitations with about 60% accuracy as contrasted with approximately chance performance with near-key imitations. This indicates that subjects had *some* information regarding interval sizes available in their memories but were easily misled by comparison trials that did not force them to change their key and scale frame of reference.

## D. Long-Term Memory

Contour information is easily encoded in memory, and subjects perform very well when tested immediately for contour information (Dowling & Fujitani, 1971; Dowling, 1978; Bartlett & Dowling, 1980). Interval information is difficult to encode, and subjects are easily confused about it in immediate tests, as the above results show. However, it seems that interval (or scale step) information, once stored in memory, fades slowly. Dowling and Bartlett (1981, Experiment 4) contrasted immediate tests with delayed tests for the same set of brief, novel, tonal melodies. They found that though both interval information and contour information were more difficult to retrieve with delayed testing, discrimination of true transpositions from tonal imitations did not decline with delay. Moreover, when stimuli were more "natural" brief excerpts of Beethoven quartets and subjects were instructed to respond positively *only* on the basis of contour, they were unable to do so. With delayed testing subjects could hardly discriminate same-contour imitations from different-contour lures (Fig. 1C versus 1E) while they could recognize exact repetitions with about 75% accuracy (Dowling & Bartlett, 1981, Experiments 1 & 2). It may be that the additional musical context of the "natural" stimuli facilitated interval encoding. Nevertheless, contour information is generally the easiest to encode.

A different aspect of the key distance effect appears under different testing conditions. If the listener's task is to detect the alteration in pitch of one of seven notes of a novel tonal melody, performance is better with transposition to a near (versus far) key. Cuddy, Cohen, and Miller (1979) found performance in detecting the altered note to be especially good when the alteration departed from the tonal scale of the comparison stimulus and when the stimuli provided a firm tonal-scale context. This shows that when subjects must shift keys to evaluate a comparison stimulus differing in only one pitch from the standard, they do so more easily to a near than to a far key.

One reason I have for prefering representations like (4) over those like (2) as descriptions of what listeners store in memory is that listeners are able to recognize melodies in cases in which the intervals of (2) but not the scale steps of (4) have been destroyed. One way to destroy the interval sizes is to interleave foreign pitches between the notes of a familiar melody (Dowling, 1973, Experiment 3). Thus, every other note the subject hears is from a target melody he or she is trying to detect, and every other note is irrelevant. When such sequences are presented rapidly (8 to 10 notes/sec), subjects perform almost perfectly in detecting the target melodies. It is as though the listener were able to hold a template like (4) up to the sequence consisting of every other note of the stimulus and evaluate a match. Another situation in which the interval sizes from note to note are destroyed while the list of scale steps is preserved (given octave-equivalence) is when melodies are scrambled so the successive pitches fall in different octaves. Deutsch (1972) demonstrated that such scrambling makes familiar tunes difficult to recognize. Dowling and Hollombe (1977) and Idson and Massaro (1978) showed that where both contour and tone-chroma (scale-step) information are available, recognition performance improves. However, this last result seems to depend critically on the presence of scale-step information. Kallman and Massaro (1979)

found that if the chromas of notes in octave-scrambled melodies were altered slightly, recognition performance was very poor even when contour was preserved. It seems as though contour can be used to narrow the set of possibilities in recognition, but that final verification of a match is done with reference to the chroma (scale-step) information of representations like (4).

## IV. CONTOUR VERSUS INTERVAL

From the above review it appears that contour information is very important to melody recognition under certain circumstances—especially when tonal context is weak (as with atonal melodies) or confusing (as with tonal imitations). Contour is less important with familiar melodies stored in long-term memory or even novel melodies remembered over periods of minutes. Meaningful musical context seems to aid memory for interval and/or scale-step information and not contour information. Contour is easy to extract from a melody but no easier to remember than intervals. Finally, contour information seems useful as an idexical device to access melodies in long-term storage, but recognition of such melodies seems critically dependent upon scale-step information.

One further aspect of adult behavior remains to be explored—namely, the effects of training on the processing of the various features of melodies. Even persons with little or no musical training perform at near-perfect levels when reproducing (Attneave & Olson, 1971) or recognizing (Bartlett & Dowling, 1981, Experiment 3) the intervals of familiar melodies. It is when dealing with novel materials that differences between trained and untrained people appear. ("Training" here refers to minimal training—usually two or more years of music lessons.) Training facilitates discrimination between tonal and atonal comparison melodies (Dowling, 1978). Untrained subjects do not find contour recognition much more difficult than trained subjects (Dowling, 1978) but do find interval recognition more difficult (Cuddy & Cohen, 1976; Bartlett & Dowling, 1980, Experiment 1). Trained listeners are much more sensitive to the tonal scale structure than untrained listeners when making judgments about how well various pitches fit a tonal context (Krumhansl & Shepard, 1979). It is safe to conclude from this evidence that one effect of training is to enhance the importance of the tonal scale system in information processing of melodies. The intervals of the scale system are firmly embedded in the minds of even untrained listeners. Training facilitates the application of that system to new materials.

## V. SUMMARY

The same hierarchy of melodic features—pitch, contour, tonality, and interval size—appears in adult behavior as in the developmental sequence. The psychological reality of these feature types is attested by their differential importance at different developmental stages, and the ways they can be manipulated in context to produce independent effects on the behavior of adults.



## ACKNOWLEDGMENTS

I wish to thank James Bartlett, Kirk Blackburn, and Darlene Smith for their help in developing this article.

## REFERENCES

Attneave, F., & Olson, R. K. Pitch as a medium: A new approach to psychophysical scaling. *American Journal of Psychology*, 1971, *84*, 147–166.

Bartlett, J. C., & Dowling, W. J. The recognition of transposed melodies: A key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception & Performance*, 1980, *6*, 501–515.

Chang, H-W., & Trehub, S. Auditory processing of relational information by young infants. *Journal of Experimental Child Psychology*, 1977, *24*, 324–331. (a)

Chang, H-W., & Trehub, S. Infants' perception of temporal grouping in auditory patterns. *Child Development*, 1977, *48*, 1666–1670. (b)

Cuddy, L., & Cohen, A. J. Recognition of transposed melodic sequences. *Quarterly Journal of Experimental Psychology*, 1976, *28*, 255–270.

Cuddy, L., Cohen, A. J., & Miller, J. Melody recognition: The experimental application of musical rules. *Canadian Journal of Psychology*, 1979, *33*, 148–157.

Deutsch, D. Octave generalization and tune recognition. *Perception & Psychophysics*, 1972, *11*, 411–412.

Deutsch, D. The organization of short-term memory for a single acoustic attribute. In J. A. Deutsch & D. Deutsch (Eds.), *Short-term memory*. New York: Academic Press, 1975. Pp. 107–151.

Deutsch, D. Octave generalization and the consolidation of melodic information. *Canadian Journal of Psychology*, 1979, *33*, 201–205.

Dowling, W. J. Rhythmic fission and the perceptual organization of tone sequences. Unpublished Ph. D. Thesis, Harvard University, 1968.

Dowling, W. J. Recognition of melodic transformations: Inversion, retrograde, and retrograde inversion. *Perception & Psychophysics*, 1972, *12*, 417–421.

Dowling, W. J. The perception of interleaved melodies. *Cognitive Psychology*, 1973, *5*, 322–337.

Dowling, W. J. Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, 1978, *85*, 341–354.

Dowling, W. J. Musical scales and psychophysical scales: Their psychological reality. In T. Rice & R. Falck (Eds.), *Cross-cultural perspectives on music*. Toronto: University of Toronto Press, in press.

Dowling, W. J., & Bartlett, J. C. The importance of interval information in long-term memory for melodies. *Psychomusicology*, 1981, *1*, 30–49.

Dowling, W. J., & Fujitani, D. S. Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, 1971, *49*, 524–531.

Dowling, W. J., & Goedecke, M. The impact of a Suzuki-based instrumental music program on auditory information-processing skills of inner-city school children. (in preparation)

Dowling, W. J., & Hollombe, A. W. The perception of melodies distorted by splitting into several octaves: Effects of increasing proximity and melodic contour. *Perception & Psychophysics*, 1977, *21*, 60–64.

Francès, R. *La perception de la musique*. Paris: Vrin, 1958.

Gardner, H., Davidson, L., & McKernon, P. The acquisition of song: A developmental approach. In *Documentary Report of the Ann Arbor Symposium*, Reston, Virginia: Music Educators National Conference, 1981.

Guilford, J. P., & Nelson, H. M. The pitch of tones in melodies as compared with single tones. *Journal of Experimental Psychology*, 1937, *20*, 309–335.

Hair, H. I. Discrimination of tonal direction on verbal and nonverbal tasks by first-grade children. *Journal of Research in Music Education*, 1977, *25*, 197–210.

Idson, W. L., & Massaro, D. W. A bidimensional model of pitch in the recognition of melodies. *Perception & Psychophysics*, 1978, *24*, 551–565.

Imberty, M. *L'acquisition des structures tonales chez l'enfant.* Paris: Klincksieck, 1969.

Jairazbhoy, N. A. *The rāgs of North Indian music.* Middletown, Connecticut: Wesleyan University Press, 1971.

Juola, J. F., Schadler, M., Chabot, R. J., & McCaughey, M. W. The development of visual information processing skills related to reading. *Journal of Experimental Child Psychology,* 1978, *25,* 459–476.

Kallman, H. J., & Massaro, D. W. Tone chroma is functional in melody recognition. *Perception & Psychophysics,* 1979, *26,* 32–36.

Kessen, W., Levine, J., & Wendrich, K. A. The imitation of pitch in infants. *Infant Behavior & Development,* 1979, *2,* 93–99.

Kinney, D. K., & Kagan, J. Infant attention to auditory discrepancy. *Child Development,* 1976, *47,* 155–164.

Krumhansl, C. L., & Shepard, R. N. Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception & Performance,* 1979, *5,* 579–594.

Kubovy, M., & Howard, F. P. Persistence of a pitch-segregating echoic memory. *Journal of Experimental Psychology: Human Perception & Performance,* 1976, *2,* 531–537.

LaBerge, D. Perceptual learning and attention. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 4). Hillsdale, New Jersey: Erlbaum, 1976. Pp. 237–273.

LaBerge, D. Perceptual and motor schemas in the performance of music. In *Documentary Report of the Ann Arbor Symposium,* Reston, Virginia: Music Educators National Conference, 1981.

Melson, W. H., & McCall, R. B. Attentional responses of five-month girls to discrepant auditory stimuli. *Child Development,* 1970, *41,* 1159–1171.

Ostwald, P. F. Musical behavior in early childhood. *Developmental Medicine & Child Neurology,* 1973, *15,* 367–375.

Pflederer, M., & Sechrest, L. Conservation-type responses of children to musical stimuli. *Bulletin of the Council for Research in Music Education,* 1968, No. 13, 19–36.

Pick, A. D. Listening to melodies: Perceiving events. In A. D. Pick (Ed.), *Perception and its development: A tribute to Eleanor J. Gibson.* Hillsdale, New Jersey: Erlbaum, 1979. Pp. 145–165.

Révész, G. *Introduction to the psychology of music.* Norman, Oklahoma: University of Oklahoma Press, 1954.

Riley, D. A., & McKee, J. P. Pitch and loudness transposition in children and adults. *Child Development,* 1963, *34,* 471–482.

Riley, D. A., McKee, J. P., Bell, D. D., & Schwartz, C. R. Auditory discrimination in children: The effect of relative and absolute instructions on retention and transfer. *Journal of Experimental Psychology,* 1967, *73,* 581–588.

Riley, D. A., McKee, J. P., & Hadley, R. W. Prediction of auditory discrimination learning and transposition from children's auditory ordering ability. *Journal of Experimental Psychology,* 1964, *67,* 324–329.

Shuter, R. *The psychology of musical ability.* London: Methuen, 1968.

Wohlwill, J. F. Effect of correlated visual and tactual feedback on auditory pattern learning at different age levels. *Journal of Experimental Child Psychology,* 1971, *11,* 213–228.

Zenatti, A. Le développement génétique de la perception musicale. *Monographies Françaises de Psychologie,* 1969, whole no. 17.

Zimmerman, M. P., & Sechrest, L. Brief focused instruction and musical concepts. *Journal of Research in Music Education,* 1970, *18,* 25–36.



# 14

# Absolute Pitch

*W. Dixon Ward and Edward M. Burns*


I. Introduction ..... 431
The Pitch Helix ..... 432
II. Genesis of AP ..... 434
III. Measurement of AP ..... 436
A. Extraneous Cues ..... 436
B. Absolute Piano ..... 436
C. Relative Pitch ..... 437
D. Accuracy of AP ..... 438
E. Information Transfer in an Expert Possessor ..... 440
IV. Stability of the Internal Standard ..... 444
V. Learning AP ..... 445
VI. The Value of AP ..... 447
References ..... 449


## I. INTRODUCTION

The ultimate in musical endowment is commonly regarded by musicians to be the possession of "absolute pitch" (AP), also called "perfect pitch" or "positive pitch": the ability to identify the frequency or musical name of a specific tone, or, conversely, the ability to produce some designated frequency, frequency level, or pitch without comparing the tone to any objective reference tone (i.e., without using relative pitch).

Suppose we present to an individual the following sequence of frequencies: 260, 260, 290, 330, 260, 330, 290 Hz and ask "What was that?" Individuals who are "tone-deaf" [or better, "melody-deaf" (Trotter, 1967)] are apt to answer with something no more specific than "a bunch of tones." The median American nonmusician will probably answer "Yankee Doodle" (identification of tune), although many may remember enough from grade school to add that it was "do, do, re, mi, do, mi, re" (identification of solfeggio). The typical performing musician can do all of that and


Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

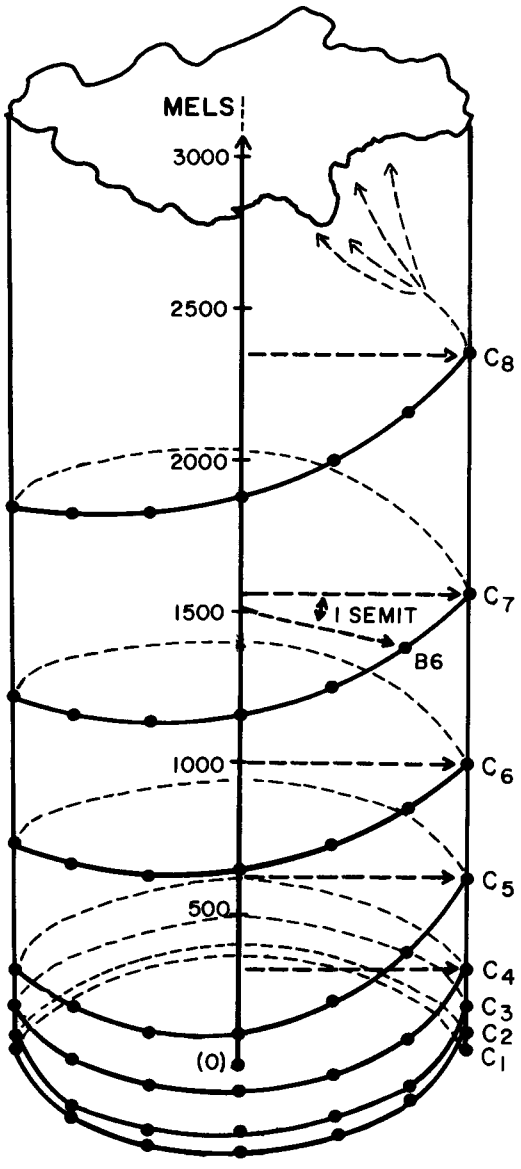
may add that "the sequence of successive intervals was unison, ascending major second, another ascending major second, descending major third, ascending major third, and descending major second" (identification of intervals). But only the person with AP is able to answer "Middle C, C, D, E, C, E, D" (identification of the musical designation of individual components).

## The Pitch Helix

Musically educated persons with good relative pitch behave as though they have developed an internal scale of pitch, a movable conceptual template that is permanently calibrated in terms of the pitch relations among the notes in our musical scale (i.e., in subjective octaves, each of which is further subdivided into 12 equal parts called subjective semitones or semits). Because corresponding semit markers within the octaves are given the same name, this template can be represented in the form of a pitch helix: a spiral that ascends the shell of an invisible vertical cylinder (Fig. 1). This schema allows several aspects of pitch to be simultaneously represented: the projection of a point on this pitch spiral on the vertical axis determines its "tone-height" in mels, the angle involved in its projection on the horizontal plane indicates its "chroma" (the scale note—A, A♯, B, etc.—or its solfeggio), and its octave designation depends on which coil of the helix it lies on, the coils being numbered in succession from low to high frequency. If the individual with good relative pitch is presented a reference tone X and told what its chroma and octave number are, the pitch engendered serves to tie one point on this pitch helix to all other pitches. The pitch helix is mentally rotated until the corresponding subjective semit marker coincides with this "anchor" pitch, and the musician can now auralize all other pitches on the scale. The musician is now prepared to make judgments of the musical interval (a concept that corresponds to distance along the spiral) that separates the anchor from some other frequency. Such judgments have been shown in Chapter 8 to be categorical in nature, involving a labeling process that is analogous to that used in the discrimination of speech sounds. So, when next given tone Y, this listener will categorize the interval between X and Y, and by using knowledge of the number of semits in that interval, the listener can indicate the musical pitch (chroma plus octave number) that tone Y must have.

However, in the foregoing case the helix is, so to speak, free-floating. That is, there are no *permanent* labels attached that serve to tie this subjective template to the objective world; there are only temporary ones. Thus if we were to present the person with relative pitch a tone of 440 Hz followed by one of 525 Hz, there might be a rapid recognition of the interval as a minor third, but only if our musician was told that the first one was A4 would he or she call the second C5. Indeed, we could say that the *first* one was C5, and the individual would then indicate that the second was D♯5 or E♭5.

Such deception would not succeed if the person has AP, however. Here the helix apparently has permanent labels attached; given the sequence 440 Hz, 525 Hz, the




Biblioteca del Real Conservatorio Superior de Música - MADRID

**Fig. 1.** The pitch spiral. The chroma of any point on the surface of the cylinder is determined by its projection on the XZ plane and its tone-height by its projection on the Y axis. Thus, for example, all Cs are shown to lie at 0° azimuth; since C6 is shown as having a tone-height of 1000 mels (which is, by definition, the pitch of a 1000-Hz sinusoid), this orientation of the pitch spiral assumes a physical A4 of about 420 Hz rather than 440 Hz. Chroma tends to become impalpable for tone-heights above 2500 mels or so (sinusoids above 5000 Hz), so the upper boundaries of both the spiral and the axis are left vague and irregular.

musician would immediately recognize A4 and C5, and if told that the first was C5, would merely snap, "Nonsense."

This ability has been subjected to scientific scrutiny for about a century (Stumpf, 1883), and we now know quite a bit about the characteristics of AP. Still an enigma, however, is the genesis of the ability: why do some people identify pitches with no apparent effort while others must engage in strenuous training to develop AP? Indeed, there was doubt that it was even possible to "learn" AP until Paul Brady did so in 1970. For this reason, its possession was for many years regarded with pride, as if it indicated that its owner were mysteriously gifted.

Yet a bit of reflection should convince one that, viewed in comparison to other sensory modalities, AP is not so strange after all. We learn labels for colors, smells, and tastes—indeed, for speech sounds, voices, and instrumental timbres. Why not also pitches? Stimuli that fall along any metathetic continuum should be labelable, one would think, and not require the comparison between stimuli either present or in short-term memory that is involved in relative-pitch judgments. One does not need to look at a rainbow in order to see that a rooster's comb is red nor take a whiff of camphor to identify a nearby skunk. Viewed in this light, the real question is why everyone does not have AP: Why cannot some people put labels on pitches?

## II. GENESIS OF AP

There are four major theories of AP: heredity, on the one hand, and learning, unlearning, and imprinting on the other. The *heredity* viewpoint, espoused most vigorously by possessors such as Revesz (1913) and Bachem (1937), contends that AP is a special innate ability that one either inherits or not, that those who do inherit the trait will demonstrate pitch-naming ability as soon as an appropriate situation arises, regardless of their early musical training, and that those who are not so genetically blessed can never attain the degree of excellence in identifying pitch displayed by the chosen few, no matter how much instruction they are given or how diligently they practice naming tones.

The *learning* theory, in its most extreme Skinnerian form, is the exact antithesis of the heredity position, asserting that heredity has nothing to do with the matter. Instead, the development of AP depends on some more or less fortuitous set of circumstances whereby the individual is reinforced for trying to put labels on pitches. Oakes (1951) postulates, for example, that "an examination of the total history of the interactions involving the organism and tonal stimuli will show a complex series of events, some gross and some extremely subtle, from which pitch-naming reactions evolved or did not evolve—depending on factors in the history itself." Just to make sure that his position cannot be disproved, he adds: "In explaining whether or not pitch-naming reactions did develop, it is necessary that we take into account every contact of the organism and tonal stimuli, and we also must consider setting and situational factors in each of the interactions." The implication, in short, is that anyone can develop AP under the "right"—but, alas, unknown—circumstances.

A variant of the learning theory is the *unlearning* viewpoint expressed by Abraham (1901) in the first extensive monograph on AP, who pointed out that most musical experience is not conducive to the development of AP. For example, a given tune may be heard in many different keys. How can a child develop absolute recognition of a particular frequency, say 261 Hz, if it is called "do" today and "sol" tomorrow or if it is heard when he presses "the white key just left of the two black keys in the middle of the piano" at home but a completely different key (perhaps even a black one) at grandfather's house? Considering all the factors that conspire to enhance the development of relative pitch at the expense of AP, Abraham suggested that perhaps an inborn potential for developing AP was relatively widespread, but that it is simply trained out of most of us. Watt (1917) pursued Abraham's line of reasoning to its ultimate and proposed that perhaps AP is initially universal: "In some favoured persons it is acquired early and more or less unwittingly and never lost. Perhaps these persons have some special refinement of hearing.... Or perhaps a highly favoured auditory disposition gives them the power to maintain their absoluteness of ear in spite of the universality of musical relativity. In that case we should all naturally possess absolute ear and then proceed to lose it or to lose the power to convert it into absolute nomenclature."

Abraham had also commented that pitch-naming ability was relatively easy to develop in children. Copp (1916) pursued this idea and on the basis of her own experience suggested that something like the modern concept of "imprinting" may be involved. Claiming that 80% of all children can be taught to produce middle C when asked and to recognize it when played or sung by others, she insisted that this is so only if they begin musical training at an early age. The notion here that AP can be developed only in children, possibly as the result of some crucial incident, may be related to the comparative ease with which children develop accent-free speech in foreign tongues.

The nature–nurture debate in this particular arena essentially ended with the death in 1957 of Bachem, who had been the most eloquent exponent of inheritance despite the lack of any convincing supportive scientific evidence. On the other hand, advocates of learning have made only slight headway either. Although Meyer (1899) indicated that he and a colleague brought themselves up to "60 and 64% terminal proficiency" after a heroic regimen of training, this improvement soon disappeared when practice was discontinued. Other later attempts to train AP [Gough (1922); Mull (1925); Wedell (1934)—and no doubt several that remained unreported] were equally unsuccessful. Wellek (1938), in a study of 27 of the best possessors in Vienna, Prague, Dresden, Leipzig, and Hamburg, noted a correlation of .80 between the age at which AP behavior appeared and number of errors on an identification test, and Sergeant (1969) has recently reported a similar result. So there is little doubt that *early* learning is important, though not all-important: Brady (1970), after months of practice, was able as an adult to achieve a degree of pitch-naming ability that was indistinguishable, in terms of error score or reaction time, from four possessors who had had the ability from childhood (Carroll, 1975). The evidence, then, favors the learning theory, although learning is clearly easier in childhood. And, of course, a genetic

component can never be ruled out completely unless some technique for teaching AP is developed that will succeed with everyone, or at least all children.

## III. MEASUREMENT OF AP

### A. Extraneous Cues

If we accept AP as the ability to attach labels to isolated auditory stimuli on the basis of pitch alone, tests for AP should not involve extraneous cues such as loudness, timbre, duration, or any other attribute. As an extreme example, nobody would take seriously as an AP test one in which stimulus #1 was a taped record of someone singing "Number One" on C3, stimulus #2 was someone else singing "Number Two" on C♯3, and so on.

There are two ways to ensure the nonuse (or at least ineffective use) of these other attributes. One is to try to hold constant all attributes except pitch. In this case, the set of tones to be judged would have to be balanced in advance by each listener to give a constant loudness, timbre, and duration. This is a tedious process, however, and even after it was finished, the tones would still differ in density and volume.

The second alternative is therefore better, to wit: vary the extraneous attributes randomly over a small range, presenting a given frequency now with one intensity and duration, next time with different ones. Under these conditions, although many parameters are varying, the only one that will provide the correct cue is pitch. Theoretically, of course, one should determine equal-pitch contours over the range of intensities to be used so that all stimuli that are supposed to be labeled A4, for example, could be adjusted in frequency as intensity is changed so that they would actually have the same pitch for that listener. However, the change of pitch with intensity over a moderate range is ordinarily negligible (Ward, 1954; Cohen, 1961), so this factor can generally be ignored. Timbre and tonal envelope could also be varied randomly, but it is practicable to hold these particular parameters constant; indeed, if anything but pure tones (sinusoids) are used, one will be in the position of presenting the listener with *several* pitches and then asking him what *one* pitch he hears (a question that has been asked all too often, despite its patent absurdity, in the field of musical perception).

### B. Absolute Piano

From the very beginning of the study of AP, it was abundantly clear to the more astute experimenters that piano tones are extraordinarily poor stimuli from the point of view of extra cues (von Kries, 1892). Abraham (1899) discusses the effect of timbre differences, nonmusical elements such as strike noises, and inharmonic partials at some length, concluding that of all instruments, tones from the piano are probably the easiest to identify because of the myriad extraneous cues that exist. On the other

hand, if the piano is struck with great force, the relative intensity of the partials may change considerably so that note identification becomes more difficult, particularly in the lowest octave or two. Thus not all of the characteristics of the piano make identification easier. The net effect is, however, that we are never quite certain just what role pitch alone plays in the identification of piano tones.

Nevertheless, a major portion of the literature on AP is based on studies using piano tones, simply because pianos are abundant. Furthermore, despite the shortcomings of the instrument, some aspects of the results are applicable to the more general picture. So, in order to avoid throwing the baby out with the bath, studies using the piano need not be completely ignored, although it must be kept in mind that "absolute piano" may not be quite the same as "absolute pitch."

### C. Relative Pitch

Much more difficult than extraneous cues to eliminate from the AP testing situation is the relative-pitch ability of a good musician. If the tones to be identified are all members of an ordinary musical scale (i.e., are separated by whole numbers of semits), it is not much of a challenge for such a listener, knowing what any one of a series is, to compare the next with it and make the second judgment agree with the estimated interval between them. Obviously, such relative-pitch judgments are likely to increase if feedback is provided on each item (e.g., Terman, 1965; Fullard, Snelbecker, & Wolk, 1972; Fulgosi & Zaja, 1975).

Various procedures have been used in attempts to destroy the short-term memory trace of preceding items on which such judgments are based. Stumpf (1883) used conversation interjected between successive stimuli, Abraham (1899) "unusual modulations" on the piano, Mull (1925) a "short period of auditory distraction," Petran (1932) reading aloud, Hartman (1954) a burst of white noise, and Hurni-Schlegel and Lang (1978) an unrelated interval-comparison task. Strangely enough, no one has yet conducted a systematic study of some of these "pitch erasers" in which subjects are *encouraged* to "cheat" and use relative pitch so that the relative efficiency of the distractors could be assessed, although a few studies have addressed the question of the effect of intervening tonal sequences on short-term memory per se—i.e., the ability to tell whether a preceding tone and a subsequent tone were the same or different (Deutsch, 1973; Siegel, 1974). Indeed, some experimenters merely rely on a blank interval of a minute or as little as 10 sec to destroy memory of a tone (Lundin & Allen, 1962; Heller & Auerbach, 1972; Carroll, 1975), apparently on the basis of a study of two possessors and two nonpossessors by Bachem (1954) in which he reported that comparison judgments deteriorated in the nonpossessor groups after as short an interval as 15 sec. However, no details of procedure or results were given by Bachem, and a recent study by Rakowski (1972) indicates that even nonpossessors can hold pitches in short-term memory over periods as long as 5 minutes.

Fortunately, there is some evidence that relative pitch is seldom used in AP experiments. Petran (1932), after a thorough review of the literature on AP at that time, did

an experiment in which each of 16 subjects was asked to identify a single piano tone as soon as he or she awoke on each of 50 days; then at the end of that time, the same 50 tones were tested at a single session. There were no significant differences between the two tests in either the number or degree of errors, which certainly suggests that in the latter case no judgments were made on the basis of relative pitch. Perhaps those who think they "have" AP feel no need for additional cues, and those who do not are seldom confident enough of any particular judgment (in the absence of feedback) to make estimating the interval between that stimulus and the next one worthwhile.

Given that there is no proven "pitch eraser," the best way of testing whether or not relative pitch was used in any given experiment is to examine the pattern of each subject's responses. If an error of say +2 categories is followed by an error of the same magnitude on the next two or three stimuli, it is likely that relative pitch was being utilized (again, a no-feedback condition is assumed). Appropriate statistical tests will allow one to determine the probability that the particular pattern of errors observed is a chance one. However, even this is not completely infallible. As Petran (1932) points out, subjects with poor relative pitch may be trying to use it but failing so that "even though there may be no trace of correct interval judgments in the results of a series test for absolute pitch, yet incorrect interval judgments may be there in numbers." This dilemma seems to have no solution.

Relative pitch may also enter into AP studies in other ways. If listeners are permitted to hum and whistle at will, many of them can come quite close to the correct pitch from knowledge of the highest or lowest note in their range (although the stability of either of these is not outstanding). Others, even without making any sound, perform as if they have AP for a single tone. That is, some violinists are apparently able to auralize A4 at will and can, given time, compare any pitch to this single internal standard. Bachem (1937) calls this type of AP "quasi-absolute pitch."

A special type of quasi-AP exists in those individuals who are afflicted with a permanent tinnitus of fixed pitch. Stumpf (1901) was one such person: he had, in essence, a built-in tuning fork whose pitch level was very nearly that of a 1500-Hz tone, so it was not necessary for him to auralize some internal standard—it was always there for the listening. There would seem to be no way to discriminate individuals with "true" AP from those with quasi-AP on the basis of error scores, although one might search for differences in the time required to make judgments, which would be expected to be greater for those with quasi-AP because they must make a relative pitch estimate (except, of course, when the stimulus is the same as their internal standard). No one has recently studied persons with quasi-AP in depth.

### D. Accuracy of AP

From this discussion of some of the pitfalls of procedure and caveats of interpretation, it appears that for the least equivocal results, one should use for the study of AP only pure-tone stimuli whose intensity and duration are varied randomly over a

narrow range. Let us turn, then, to specific procedures that have been used to measure AP. As the original definition implies, AP is manifested either by accurate production of a designated note or by correct categorization of a presented tone. The problem of course is deciding on how to define "accurate" and "correct" in the two cases, respectively. "Accuracy" in production is perhaps the easier to define, as only one psychophysical method is applicable—the method of adjustment. Even then, though, one can argue over the relative merits of "absolute accuracy" and "relative accuracy." Absolute accuracy would be measured by calculating the difference between frequencies based on A4 = 440 Hz and those produced by the subject. The subject's usefulness as an animate tuning fork in setting the pitch for a chorus would depend on absolute accuracy. However, such a procedure is in a sense "unfair" to an individual who grew up with a piano tuned a semit or so flat or who has suffered "paracusis": a change in the pitch aroused by a specific frequency (Ward, 1954), presumably due to a more or less localized disturbance on the basilar membrane. So if our interest is not so much in the practical aspects of AP as in the theoretical basis of absolute identification, the important statistic would be relative variability, as manifested in the distribution of repeated adjustments. From this point of view constant errors should be ignored; the "best" absolute pitcher is the individual with the lowest variance.

There is, however, yet another problem: what to do about "octave errors." Suppose that the subject, told to adjust an oscillator to A4, gives successive values of 444, 432, 449, 438 and 882 Hz. To say unthinkingly that the mean of these judgments is 529 Hz or that the SD is 177 Hz would clearly be misleading since the last judgment, in terms of the pitch helix, was within a few Hz of the correct chroma but was one octave off in tone-height. Although we may be in the position of trying to average apples with oranges, the most accepted solution to the problem here has been to consider only chroma. In the example given, the 882 Hz would be dropped an octave to 441 Hz, making the mean now also 441 Hz with a standard deviation of 5.7 Hz. (A more rigorous procedure would be to determine the frequency that actually appeared to be one octave lower than 882 Hz and use this value in calculating the mean, but ordinarily the difference between the subjective octave and the physical octave will be small enough—Ward, 1954—that its determination would not be worth the considerable effort involved.)

Identification techniques, although greater in variety, have the same problems as pitch production, plus a few more. Not only must one deal with octave errors and constant errors, but now the categorization process also confuses the issue. It becomes difficult to test the ability of subjects to identify quartertones because half of the stimuli will have no "name" in our diatonic scale. This problem is attacked by asking the subject to learn a new set of labels—that is, arbitrary numbers assigned to specific frequencies. One can then apply information-transfer analysis to an experiment in which a subject attempts to identify a series of stimuli consisting of some number of items from this fixed set. In theory, such a procedure, when used not only for testing but also for training, might also be a method of discovering persons with "latent



AP"—that is, individuals who can make absolute judgments but have never learned the names of the notes of the scale. However, to our knowledge there is in the literature no instance of such a person being "discovered" in this fashion.

### E. Information Transfer in an Expert Possessor

In 1952 Pollack published the first study of the information transmitted by pitch in average (unselected) listeners, finding that the maximum information that could be transmitted by pitch was only about 2.7 bits (i.e., $2^{2.7} = 7$ different pitches spread over the entire frequency range could just be named correctly by his best listener). However, it was clear that these subjects used only tone-height in their judgments: none had AP. For this reason an intensive study of one particular listener JL was undertaken (Ward, 1953). Although these data are nearly 30 years old, as far as we know no one has demonstrated greater facility at pitch-naming than JL, so the experiments will be reported here in detail.

In each of the experiments of this study, each of a set of 10 (or, for some tests, 20) frequencies was presented to the left ear of JL one at a time together with its number in that series, and JL was instructed to write on a card with the appropriate numbers anything that would help her identify number 1, number 2, and so on. After the entire list had been presented twice, a set of 100 items, 10 of each category, was judged. The intensity was varied randomly over a 20-dB range around 50 phons. JL responded vocally to each test item with a number; no feedback was given. The order of stimuli was semirandom: each subset of 20 stimuli contained two of each of the categories, the only rule of succession being that there could not be three of the same category in a row, even at the subset boundary. A confusion matrix was constructed for each test, and by means of standard formulas the information transmitted was determined. From this, the number of categories over this range that could have been correctly distinguished on a consistent basis was calculated. No attempt was made to erase the short-term memory effect via interfering tones, conversation, or noise; as will be seen, however, this was not a serious mistake because her pattern of errors indicated little if any serial dependence in her responses. JL was seated in an anechoic chamber while the experimenter was outside; communication took place via intercom.

The sequence of tests with the experimental outcome is shown in Fig. 2. The first test employed 10 stimuli from C4 (262 Hz) to A7 (3520 Hz). Successive stimuli were separated by a musical fourth (i.e., C4, F4, B♭4, etc.) so that there was no duplication of chroma (a move designed to minimize octave errors) yet with large degrees of difference in tone height. JL made no errors on this test, although the ordinary musician without absolute pitch will indicate transfer of only about 2.4 bits of information, implying that five categories could have been distinguished nearly perfectly (we routinely use this test to screen for possessors of AP in groups of listeners).

In Test 2, again the stimuli had different names, but this time they were separated only by one semit (A5 to F♯6, or 880 Hz to 1480 Hz). Again JL made no errors.

Next, a situation that would maximize the possibility of octave confusions was

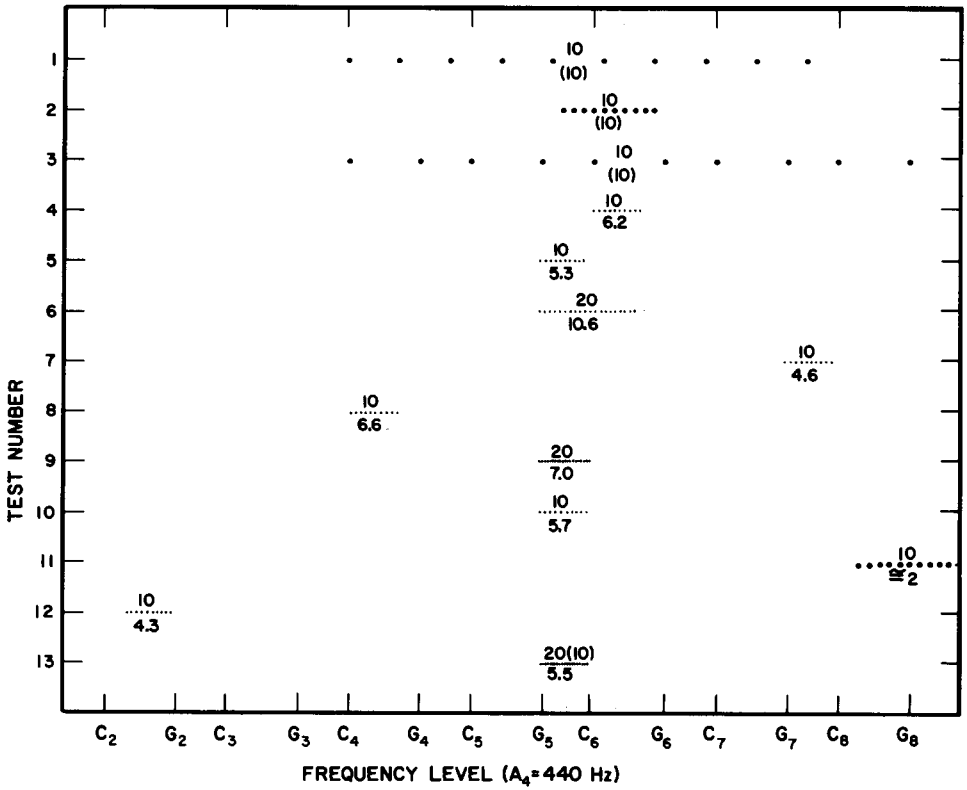



**Fig. 2.** Graphic representation of pitch identification tests on subject JL. For each test the upper number indicates the number of different stimuli involved, the dots show their frequency level, and the lower number represents the information transmitted in the form of the number of stimuli in the range concerned that could have been identified without error.

developed. In Test 3, five Cs and five Gs were used: from C4 (262 Hz) to G8 (6270 Hz). In this test JL made three mistakes: number 7 (C7) was called number 5 (C6) twice and 5 was called 7 once. However, this is still very close to perfect transmission of information; besides, the errors all occurred in the first half of the test battery, which suggests that a slight amount of learning was necessary. This result reinforces the hypothesis that octave errors are largely an artifact of using piano tones or other complex stimuli that of course *do not have* a single pitch.

Because earlier experiments on persons with AP had indicated that the best performers could distinguish quartertones with 95% accuracy (Abraham, 1901; Petran, 1932; van Krevelen, 1951; for a detailed summary of this early literature see Ward, 1963a,b), Test 4 involved 10 stimuli spaced by 50 cents from C6 to E6+50 (1046 to 1357 Hz). With this set of tones JL finally made some errors, dropping to 80% correct. The pattern of these errors is shown in Table I, which is a horizontal version

TABLE I
Responses Made to Each Stimulus in Test 4 (C6 to E6+50 cents)

| Stimulus minus response | Stimulus Number | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| +1 | | 6 | 1 | 1 | | 2 | | | | |
| 0 | 10 | 4 | 7 | 9 | 8 | 7 | 8 | 9 | 9 | 9 |
| −1 | | | 2 | | 2 | 1 | 2 | 1 | 1 | 1 |

of the diagonal of the confusion matrix. In this case the information transmitted $I_T$ was 2.64 bits, implying a potential error-free discriminability of 6.2 categories, a value slightly greater than the 5.5 semits spanned by the range employed.

One might conclude from this that JL could just identify perfectly tones 1 semit apart over the entire auditory range. However, it was next necessary to demonstrate that if one range of frequencies contained $X$ discriminable categories and an adjacent range contained $Y$, the results of a test involving both ranges would indicate a transmission of information of $X + Y$ categories. Therefore, in Test 5, 10 quartertones from G5 − 25 to B5 + 25 were used, and Test 6 involved all 20 quartertones from G5 − 25 to E6 + 25. The reason for using G5 − 25, G5 + 25, etc., instead of G5, G5 + 50, etc., was that a test for AP using the method of adjustment had shown that JL's internal template was about 25 cents flat re A4 = 440 Hz (Ward, 1954); her pitch helix was apparently tuned much closer to the old "physical" standard pitch based on C4 = 256 Hz.

Test 5 gave an $I_T$ of 2.39 bits (5.2 categories), and Test 6 indicated 3.41 bits (10.6 categories). Based on Tests 4 and 5, the expected value of distinguishable categories in Test 6 was 6.2 + 5.3, or 11.5, so one category was lost, as it were, in the process of doubling the number of alternatives. Perhaps this effect is due to elimination of a boundary.

The implication that error-free performance was limited to conventional semit categories was checked by two more 10-quartertone tests: Test 7 showed 2.20 bits transmitted (4.6 categories) in the range of G7 − 25 to B7 + 25, and Test 8 gave 2.71 bits (6.6 categories) from C♯4 + 25 to F♯4 − 25. In the latter test, stimuli 8, 9, and 10 (F4 − 25, F4 + 25, and F♯4 − 25) were correctly identified all 10 times.

In order to make sure that the last fragment of information was being extracted from JL, Test 9 used categories separated by only 25 cents. The range was the same as for Test 5 (G5 − 25 to B5 + 25) but there were 20 stimuli at 25-cent intervals. Results showed a slight improvement over Test 5, as $I_T$ rose to 2.8 bits (7.0 categories). That this improvement was more than a learning effect was shown by Test 10, which was simply a repetition of Test 5 and which gave nearly the same result: 2.5 bits, 5.7 categories.

In order to determine the limits of JL's identification range, Test 11 was designed to investigate the area from D8 to B8 in 100-cent (semit) steps. However, the test was terminated after 30 trials because JL became upset at being unable to perform accurately; her last nine responses were all either 8, 9, or 10, even though the stimuli were actually 3, 7, 6, 2, 5, 8, 2, 5, 3. JL thus displayed the "chroma fixation" reported by Bachem (1948): an inability to name notes much above 4000 Hz. This was somewhat surprising because JL, who was also serving as a subject in a study of relative pitch (Ward, 1954), had learned to make octave judgments in this range; that is, given a reference tone of A7, she would consistently set a variable tone to about A8 + 50. However, her experience with these high-frequency tones was apparently too limited to affect her ability to identify them on an absolute basis.

Performance is somewhat degraded at extremely low frequencies also. Test 12 (D2 + 25 to G2 − 25, 75 to 91 Hz, in 50-cent steps) gave $I_T = 2.1$ bits, 4.3 categories. On the final test JL was told that it was to be a repetition of Test 9 (20 stimuli at 25-cent intervals); however, only stimuli 1, 3, 5, . . . , 19 were actually presented. The results were essentially the same as for Tests 5 and 10: $I_T$ was 2.45 bits, implying 5.5 discriminable categories. JL was unaware that half of the possible categories had never been presented.

In all these tests, when errors were made, they usually occurred at random. That is, there was never the long run of errors in the same direction that would be expected if JL were attempting to use relative pitch (i.e., if she were comparing stimulus $N$ with the memory of stimulus $N - 1$). So in her case the use of a short-term-memory eraser was unnecessary. It may also be mentioned that errors occurred as often on the weaker stimuli as on the stronger (recall that the range of intensities was 20 dB).

One can infer from these data that JL should be able to identify without error some 70 to 75 pitches in the auditory range, which is about the number of semits from 60 to 4000 Hz, and that quartertones can be discriminated well above chance. Indeed, if one uses the criterion for AP proposed by Bachem—that is, ignoring errors in which the correct category was only missed by one semit—JL would be considered able to name quartertones accurately (e.g., Table I). However, that seems to be about the limit, as 25-cent intervals transfer only slightly more information than 50-cent intervals.

A test of "absolute loudness" on JL (1000-Hz tone, 10 intensities from 10 to 100 dB SPL in 9-dB steps) showed $I_T$ to be 2.69 bits (6.5 categories), so one could contend that there are, for JL, about 500 pure tones that can be completely distinguished. Unfortunately, this extrapolation was not tested directly (or perhaps fortunately, from the point of view of the sanity of JL). Even now, apparently only one experiment has attempted to have subjects judge pitch and loudness categories simultaneously (Fulgosi *et al.*, 1975), and since none of the subjects in that study had AP, the total information transmitted by both loudness and pitch was only 3.85 bits rather than 9, implying something on the order of only 14 separable pure tones—and this in a situation in which the subjects were given feedback, scored their own results, and revealed their scores to their classmates. Truly, possessors of AP *are* different.

## IV. STABILITY OF THE INTERNAL STANDARD

All pitch-frequency relations, both in possessors and in nonpossessors of AP, are apparently established early in life and cannot be changed. If something happens to disturb the normal hydromechanical or neurophysiological processes at one particular area of an ear so that a given frequency no longer affects exactly the same receptors and associated neural elements, the pitch in this region is shifted and the listener has "musical paracusis." (And, if the two ears are not affected equally, "binaural diplacusis" will exist.) However, no learning takes place; that is, musical paracusis does not gradually disappear as one "relearns" that a 3500-Hz tone arouses a percept formerly associated with a 3700-Hz tone.

Only a possessor of AP, however, will become aware of a change in the *entire* tuning of the auditory system. If, for some reason, *all* pitches were shifted by the same percentage, the individual with only relative pitch ability would hear nothing amiss. The possessor of AP, though, would complain that everything is in the "wrong" key—that all music sounds as though it has been transposed.

Interestingly enough, several possessors do indeed make this specific complaint. After age 50 or so, music was heard one or even two semits sharp from what it "ought" to be. Triepel (1934) reported that this occurred in himself, his father, and his brother. Vernon (1977) noticed about age 52 that keys were shifted about a semit. This was particularly distressing because, as a result, he heard the overture to *Die Meistersinger* in C♯ instead of C and for him C is "strong and masculine" while C♯ is "lascivious and effeminate." Now, at age 71, he hears everything *two* semits high (which at least brings Wagner away from the sauna parlor, although Vernon fails to indicate the nature of "D-ness" for him). Beck (1978) experienced a 1-semit sharping at 40 years of age that progressed to 2 semits at age 58; although he finds it somewhat disconcerting to watch a trumpet player performing in B♭ but to hear it in C, he indicates that if he can watch the score, "the eyes and ears lock in synchrony" and everything sounds all right.

It is possible that some possessors have experienced this shift without being consciously aware of it. For example, Corliss (1973) reports that she was surprised to find that when she plays Chopin's Prelude in A Major (op. 28, no. 7) from memory (pitch memory, not motor memory), she performs it in G♯. Although she attributes this to the fact that she originally learned the piece as a child on an old piano that was more than a quartertone flat, it may be that she, too, has a hearing mechanism that has aged by 1 semit. Apparently, at any rate, one aspect of presbyacusis (the change in hearing with age) may be a gradual shift of the excitation on the basilar membrane in the direction of the oval window. However, not everyone experiences this change, it would seem. Wynn (1973), for example, tells us that Carpenter's (1951) subject claims that his A4 has not changed from 435 Hz in 50 years.

Shorter term changes in tuning have also been reported in tests using the method of adjustment. Abraham (1901) indicates that his A4 varied, over a three-month period, from 451.3 to 442.8 Hz in a random manner, though consistent on a single day. Wynn (1971, 1972) tested his wife's ability to sing A4 on demand over a period of three

months, during which time the frequency produced rose from about 440 Hz in the first two weeks to 462 Hz in the last month (and has remained there six months later). Wynn sees in his data a slow cyclical change in the mean that is associated with the menstrual cycle; however, that particular evidence is far from compelling. More significant, though inexplicable, is the fact that his wife's internal standard seems to have gone *flat* by nearly a semit (since she had to raise the frequency to sound "right").

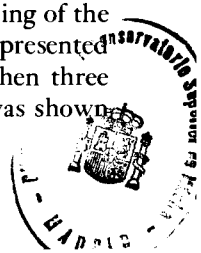
## V. LEARNING AP

Various methods have been proposed for improving proficiency at pitch naming up to the level displayed by JL. The first attempt was made by Max Meyer and Heyfelder in 1895 (Meyer, 1898). Beginning with 10 different pitches (both piano tones and tuning forks were used), they gradually increased the number to 39. Although details of procedure are not given, Meyer states that by the time they abandoned the experiment, they had achieved 64 and 60% proficiency. He also indicates that at the time he was writing the article (three years later), he had lost the increased facility.

Gough (1922) reports that the average error of nine graduate students dropped from 5 semits to about 1.5 after a year of practice on identification of piano tones. A study by Mull (1925) also produced a slight improvement in the ability to recognize one particular tone after nine months of practice (one hour weekly). Her technique, apparently designed to develop quasi-AP, consisted of presenting middle C steadily for 15 minutes (!) and then having the listener try to pick middle C from a series of tones within a two-octave range. As soon as a subject responded, the experimenter told him whether or not he was correct; then middle C was again presented (happily, only 5 sec this time), followed by another random series of tones, etc. At the end of this training, a series of tests was performed in which no correction was given. In this series the nine stimuli used were 232 to 296 Hz in 8-Hz steps (i.e., somewhere around quartertone separation); in a long series of tones, the subject was asked to indicate the occurrence of each middle C (256 Hz at this time). The percent correct identification in four successive sessions was 43, 47, 64, and 57, respectively, which is hardly evidence that AP has developed.

Wedell's (1934) study was the first to use response categories other than musical-scale values. The range of frequencies from 50 to 7500 Hz was divided in order to provide 5, 9, 13, 17, 25, or 49 stimuli separated by approximately equal values of pitch extent (tone height)—actually, by equal numbers of just-noticeable differences: 333, 167, 111, 84, 55, and 28, respectively. All were presented at a loudness level of 40 phons. The subject had a chart on which was printed a list of the frequencies used in that particular test.

In the first part of his experiment, he measured in four subjects the learning of the 25-category series. In each daily experimental session, each tone was first presented once, and the subject was asked to identify each, without correction. Then three complete series were presented with feedback. Significant improvement was shown



from day 1 to day 4 on session 1, but little thereafter (a total of 20 sessions were run). The average error dropped from about 2.5 categories (125 DLs) to just over 1 category (60 DLs). After a three-month rest, he retrained the subjects on the 25-category series; then he ran two sessions using the 49-item series. The average error remained at 60 DLs (about 3 semits) although this was now more than a two-category separation. No subjects came close to 100% correct identification. A second experiment in which listeners were first trained to identify 5, then 9, then 13 different stimuli also produced only slight improvement; none could get 100% correct on the 13-item series.

Lundin and Allen (1962) report improvement of performance in both possessors and nonpossessors using a 24-button voting board that provided feedback automatically. Lundin (1963) indicates that with this apparatus, learning was even more facilitated by starting with a few tones and gradually increasing their number but gives no details.

A variant of Mull's "learn a single reference" method was proposed by Cuddy in 1968. Here, listeners were required to respond either "A" or "not A" to a series of semits ranging from E4 to D♯5 in which A4 occurred very often during the early training, gradually dropping in later sessions; all six subjects showed improvement with practice. Brady (1970) employed this procedure in the unique development of his ability to name pitches. Using a computer, he presented himself various tones from 117 to 880 Hz in which there was a high proportion of Cs; as training progressed, the proportion gradually dropped to the "chance" 1/12. Although he admits using relative pitch at the beginning of the experiment, he reports that gradually (i.e., no sudden "revelation") he became able to recognize any C immediately. In tests using Petran's (1932) technique for confirming his ability (that is, having his wife play a single note chosen at random every morning for 57 days), he made two 2-semit errors, 18 1-semit errors, and was correct 37 times. As indicated earlier, Carroll (1975) tested Brady's ability, finding that he responded as accurately and as swiftly as four individuals who claimed AP without any formal training. However, Brady himself points out that, unlike them, he does not identify the key of a piece instantly.

Cuddy (1970, 1971) has recently extended the single-standard method to three standards. In the first study, which can be called the nonmusical experiment, she tried to get subjects to classify stimuli as low, medium, or high in pitch, and then to further subclassify each of the three areas into three specific pitches: L−, L, L+; M−, M, M+; H−, H, H+. Two types of spacing were compared: equal mels (nine stimuli from 290 to 3000 Hz spaced at 200-mel intervals), and grouped mels (500, 600, 700; 1100, 1200, 1300; 1700, 1800, 1900 mels). Half of the subjects were given training using a gradually decreasing proportion of L, M, and H (the "standards") beginning at 14/16 and progressing to 8/13, while the other half were given series in which all nine stimuli were equally probable. Spacing proved to be irrelevant (the grouping did not help); although musicians improved more with "standards" training, nonmusicians succeeded better with the equal-probability series. Performance, however, was not noteworthy—the best average $I_T$ was 1.91 bits (in musicians, after three-standard training).

Cuddy's other study (1971) was an improvement in that she used musical stimuli

and exploited musical skills. The stimuli in the main experimental group were all the Fs, As, and Cs from F3 to C7 (175 to 2093 Hz); this is called *triad spacing*. Ability to learn to identify these 12 stimuli was compared to learning to identify (1) 12 tones over the same range separated by 3 or 4 semits with no repetition of chroma or (2) 12 tones with essentially equal arithmetic spacing. All series included A4. Training led to great improvement in the musicians using triad spacing; indeed, three listeners finally achieved 100% scores (3.59 bits). Of course, this is a fairly unique "kind" of AP; the listener had only to identify the octave in which the tone occurred. Because all tones were members of the ordinary F triad, she had only to remember this triad in order to decide whether it was the tonic (F), the third (A), or the fifth of the chord (C).

Heller and Auerbach (1972) examined the importance of the type of feedback during training. Using semit stimuli from F4 to D5, they developed a series of 4 training tapes in which the proportion of A4s was successively 40, 30, 20 and 10%. One group of subjects was told whether each tone was A or not A while the other group received full feedback. Improvement in ability to identify all tones in a no-feedback trial was the same for both groups. They conclude that the single-standard procedure is not crucial. However, the improvement was only from 27% correct to 37%, and this may merely represent learning to utilize relative pitch more efficiently. In view of this equivocal result, the issue can hardly be regarded as closed.

## VI. THE VALUE OF AP

Thousands of hours have been spent trying to develop AP, yet only one person has been able to achieve as an adult the proficiency that seems to come so naturally to most possessors. One may well ask whether or not AP is worth that much effort. What are the advantages of AP? The most obvious, of course, is that if the pitch pipe is lost, an *a cappella* performance can still be started by a possessor on the correct pitch—an absolute necessity in some songs such as the American national anthem that require an octave and a half of vocal range. Beck (1978) writes me that he can tell his speed on a particular road from the pitch of his tires, McCleve (1979) judges his car's RPM from the motor whine, and Corliss (1973) indicates that identification of elements in a chord is easier for possessors because each element is recognized as such. However, outside of these few instances in which the frequency of a tone is important, the alleged advantages are mostly incapable of confirmation. For example, Bachem (1955) enthuses that "Particular characteristics of certain keys, e.g., the brilliancy of A major, the softness of D-flat major, can only be appreciated fully through absolute pitch." Although he apparently would agree with Vernon on pieces played in C♯ (unless he would maintain that C♯ is different from D♭), we can only take his word that that is how things sounded to him. As recently as 1976 nonsensical claims such as the following are still appearing despite a lack of corroborative evidence: "Absolute pitch is generally an asset to a musician because it helps in playing an instrument in tune, in sight-singing easily and accurately, and in knowing what a piece of music will sound like simply by reading the musical score" (Eaton & Siegel,

1976). These are all skills displayed by nonpossessors of AP as well as by possessors. There is also little evidence that there is much of a correlation between AP and other musical traits such as the ability to improvise (Weinert, 1929) or to make judgments of relative pitch (Ward, 1954; Baggaley, 1974).

Indeed, one would predict that if persons with AP cannot help categorizing notes, they might well be at a disadvantage in certain relative-pitch tasks. For example, in experiments dealing with categorization of intervals such as those described by Burns and Ward (Chapter 8), four tones are presented, and a subject is asked to compare the musical-pitch distance between the first and second tones to that between the third and fourth. The results (Burns and Ward, 1974, 1978) imply that in this task, as in all categorization tasks, some information is sacrificed when categorization occurs. The effect is as if the memory trace of the interval defined by the first pair of tones shifts toward the "best" interval in that category. If an analogous process occurs in AP, it would seem that the possessor might be subject to double jeopardy in this particular situation because he may be making not two but four categorical judgments. Consider the sequence A + 40, C − 40,; D − 40, E + 40 as an extreme example. The person using relative pitch will easily recognize the first interval of 220 cents (i.e., from A + 40 to B + 60) as a major second and the second interval of 280 cents as a minor third and will conclude that the second interval was the larger. On the other hand, if a possessor of AP were unable to avoid categorizing each note and had an internal pitch helix tuned to A4 = 440 Hz, he would perceive the sequence as being A, C; D, E and would therefore say that the *first* interval was the larger. The question, obviously, is how well the possessor can suppress the tendency to categorize individual tones.

Perhaps when the tones are in rapid succession even the most zealous possessor can attend to the intervals as such—can shift from the reference tonality established by the first tone (for judging the distance to the second) to the new reference tonality associated with the third tone, which of course is necessary if the pitch distance between the third and fourth tones is to be accurately evaluated. With longer interstimulus intervals, however, possessors of AP might find their memory trace of A + 40 gradually slipping to A, etc., leading to the confusion described above for certain pairings of frequencies. At the moment nobody has yet studied the ability of possessors to make relative interval judgments using these "off-frequency" tones (i.e., mistuned relative to the possessor's internal scale), although Siegel and Siegel (1977) were on the brink of such an experiment: they studied interval categorization in nonpossessors and single-stimulus pitch categorization in possessors but not interval discrimination in possessors. So how much of a handicap AP provides in such a situation (admittedly only peripherally related to musical experience) remains to be seen.

One situation that puts possessors at a disadvantage has recently been devised by Cuddy (1977). The task was to identify whether two 7-tone sequences (melodies) were identical or differed on one of the tones by half a semit (3%). Performance of nonpossessors was independent of whether the sequence was tonal or atonal (i.e., used only the 7 major tones of the scale or, instead, all 12) or whether the tones used were separated by equal ratios or by equal steps of linear frequency (log and linear scales,

respectively). Possessors were like nonpossessors in two of the tasks but showed significantly poorer performance when judging atonal sequences in the linear scale while simultaneously displaying significantly superior skill at distinguishing tonal sequences in the log scale. That is, possessors were better at picking out a single mistuned tone from a sequence, but when many of the tones were already mistuned (from the viewpoint of standard scales), they were unable to determine whether or not an additional 50-cent change had occurred. They are, it appears, less flexible—less able to learn a new scale.

In any event, AP is not an unmixed blessing. It remains a fascinating phenomenon, although its fascination, for many of us, lies in the question of why so many people do not have the ability rather than why a few do. Any situation in which one particular frequency is to be discriminated from all others, as in testing for "frequency generalization" in experimental animals, is a type of AP experiment; and although the literature is not completely unequivocal, there is evidence that a high degree of discrimination of tones separated by long intervals of time is indeed found in lower organisms such as the dog (Andreyev, 1934) or rat (Blackwell & Schlossberg, 1943), though perhaps it is not highly developed in the cat (Thompson, 1959).

Interestingly enough, mentally retarded teenagers apparently have AP. Paulson, Orlando, and Schoelkopf (1967) trained three of these retardates to tap rapidly on a lever when a 1455-Hz tone was present in order to receive bits of candy, and then the experimenters tried to measure "generalization" to tones of 1100 and 1855 Hz. To the surprise of the experimenters, they got hardly any responses to the new tones in two of their subjects. One of these was therefore tested further with tones spaced at 100-Hz intervals, and even with such small separations, only 1455 Hz was able to elicit a response. Clearly, this illustrates AP in these subjects (assuming, of course, that there was not something unique about the acoustic effects of the 1455-Hz tone in that particular situation that provided an unwanted cue, such as some object that resonated at that frequency). Paulson *et al.*, however, were not overjoyed to find this; rather than rejoicing that the children had AP, they bemoan the results as being "indicative of a broad deficit in generalization processes" of retardates. It appears that AP is typical of either musical precociousness or mental retardation. So if you have it, be sure it is for the right reason.

## REFERENCES

Abraham, O. Das absolute Tonbewusstsein. *Sammelbde. International Musikges*, 1901, *3*, 1–86.

Andreyev, L. A. Extreme limits of pitch discrimination with higher tones. *Journal of Comparative Psychology*, 1934, *18*, 315–332.

Bachem, A. Various types of absolute pitch. *Journal of the Acoustical Society of America*, 1937, *9*, 146–151.

Bachem, A. Chroma fixation at the ends of the musical frequency scale. *Journal of the Acoustical Society of America*, 1948, *20*, 704–705.

Bachem, A. Time factors in relative and absolute pitch determination. *Journal of the Acoustical Society of America*, 1954, *26*, 751–753.

Bachem, A. Absolute pitch. *Journal of the Acoustical Society of America*, 1955, *27*, 1180–1185.

Baggaley, J. Measurement of absolute pitch: A confused field. *Psychology of Music*, 1974, *2-2*, 11-17.
Beck, J. F. Personal communication, May 1978.
Blackwell, H. R., & Schlosberg, H. Octave generalization, pitch discrimination, and loudness thresholds in the white rat. *Journal of Experimental Psychology*, 1943, *33*, 407-419.
Brady, P. T. Fixed-scale mechanism of absolute pitch. *Journal of the Acoustical Society of America*, 1970, *48*, 883-887.
Burns, E. M., & W. D. Ward. Categorical perception of musical intervals. *Journal of the Acoustical Society of America*, 1974, *55*, 456.
Burns, E. M., & W. D. Ward. Categorical perception—phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America*, 1978, *63*, 456-468.
Carpenter, A. A case of absolute pitch. *Quarterly Journal of Experimental Psychology*, 1951, *3*, 92-93.
Carroll, J. B. Speed and Accuracy of absolute pitch judgments: Some latter-day results. Research Bulletin (preprint). Princeton, New Jersey: Educational Testing Service, Ocotber 1975.
Cohen, A. Further investigation of the effects of intensity upon the pitch of pure tones. *Journal of the Acoustical Society of America*, 1961,*133*, 1363-1376.
Copp, E. F. Musical ability. *Journal of Heredity*, 1916, 7, 297-305.
Corliss, E. L. Remark on "fixed-scale mechanism of absolute pitch." *Journal of the Acoustical Society of America*, 1973, *53*, 1737-1739.
Cuddy, L. L. Practice effects in the absolute judgment of pitch. *Journal of the Acoustical Society of America*, 1968, *43*, 1069-1076.
Cuddy, L. L. Training the absolute identification of pitch. *Perception & Psychophysics*, 1970, *8*, 265-269.
Cuddy, L. L. Absolute judgement of musically-related pure tones. *Canadian Journal of Psychology*, 1971, *25*, 42-55.
Cuddy, L. L. Comment on "Practice effects in the absolute judgment of frequency" by Heller and Auerbach. *Psychonomic Science*, 1972, *28*, 68.
Cuddy, L. L. Perception of structured melodic sequences. Paper given at *Conference on Musical perception in Paris, France, July 10-13, 1977*.
Deutsch, D. Octave generalization of specific interference effects in memory for tonal pitch. *Perception & Psychophysics*, 1973, *13*, 271-275.
Eaton, K. E., & M. H. Siegel. Strategies of absolute pitch possessors in the learning of an unfamiliar scale. *Bulletin of the Psychonomic Society*, 1976, *8*, 289-291.
Fulgosi, A., Bacun, D., & Zaja, B. Absolute identification of two-dimensional tones. *Bulletin of the Psychonomic Society*, 1975, *6*, 484-486.
Fullard, W., Snelbecker, G. E., & Wolk, S. Absolute judgments as a function of stimulus uncertainty and temporal effects: Methodological note. *Perception & Motor Skills*, 1972, *34*, 379-382.
Gough, E. The effects of practice on judgments of absolute pitch. *Archives of Psychology, New York*, 1922, 7, No. 47, 93.
Hartman, E. B. The influence of practice and pitch distance between tones on the absolute identification of pitch. *American Journal of Psychology*, 1954, *67*, 1-14.
Heller, M. A., & Auerbach, C. Practice effects in the absolute judgment of frequency. *Psychonomic Science*, 1972, *26*, 222-224.
Hurni-Schlegel, L., & Lang, A. Verteilung, Korrelate und Veränderbarkeit der Tonhöhen-Identifikation (sog. absolutes Musikgehör). *Schweizerische Zeitschrift fur Psychologie & Anwendungen*, 1978, *37*, 265-292.
von Kries, J. Über das absolute Gehör. *Zeitschrift fur Psychologie & Physiology Sinnesorg*, 1892, *3*, 257-279.
van Krevelen, A. The ability to make absolute judgments of pitch. *Journal of Experimental Psychology*, 1951, *42*, 207-215.
Lundin, R. W. Can perfect pitch be learned? *Music Education Journal*, 1963, 49-51.
Lundin, R. W., & Allen, J. D. A technique for training perfect pitch. *Psychology Record*, 1962, *12*, 139-146.
McCleve, D. E. Personal communication, March 1979.
Meyer, M. Is the memory of absolute pitch capable of development by training? *Psychological Review*, 1899, *6*, 514-516.

Mull, H. K. The acquisition of absolute pitch. *American Journal of Psychology*, 1925, *36*, 469–493.
Oakes, W. F. An alternative interpretation of "absolute pitch." *Transactions of the Kansas Academy of Sciences*, 1951, *54*, 396–406.
Paulson, D. G., Orlando, R., & Schoelkopf, A. M. Experimental analysis and manipulation of auditory generalization in three developmental retardates by discriminated-operant procedures. IMRID Papers and Reports 4, No. 13, 1967. (Institute on Mental Retardation and Intellectual Development, George Peabody College for Teachers, Nashville, Tennessee 37203.)
Petran, L. A. An experimental study of pitch recognition. *Psychological Monograph*, 1932, *42*, No. 6.
Pollack, I. The information of elementary auditory displays. *Journal of the Acoustical Society of America*, 1952, *24*, 745–749.
Rakowski, A. Direct comparison of absolute and relative pitch. *Symposium on Hearing Theory, Eindhoven, Netherlands*, sponsored by IPO, June 22–23, 1972.
Revesz, G. *Zur Grundlegung der Tonpsychologie*. Veit, Leipzig, 1913.
Riker, B. L. The ability to judge pitch. *Journal of Experimental Psychology*, 1946, *36*, 331–346.
Sergeant, D. Experimental investigation of absolute pitch. *Journal of Research in Music Education*, 1969, *17*, 135–143.
Siegel, J. A. Sensory and verbal coding strategies in subjects with absolute pitch. *Journal of Experimental Psychology*, 1974, *102*, 37–44.
Siegel, J. A., & W. Siegel. Absolute identification of notes and intervals by musicians. *Perception & Psychophysics*, 1977, *21*, 143–152.
Stumpf, C. *Tonpsychologie*. Herzel, Leipzig, 1883.
Stumpf, C. Beobachtungen über subjective Töne und über Doppelthören. *Beitrage zur Akustik & Musik*, 1901, *3*, 30–51.
Terman, M. Improvement of absolute pitch naming. *Psychonomic Science*, 1965, *3*, 243–244.
Thompson, R. F. The effect of training procedure upon auditory frequency discrimination in the cat. *Journal of Comparative & Physiology Psychology*, 1959, *52*, 186–190.
Triepel, H. Zur Frage des absoluten Gehörs. *Archiv für die Gesamte Psychologie*, 1934, *90*, 373–379.
Trotter, J. R. The psychophysics of melodic interval: Definition, techniques, theory and problems. *Australian Journal of Psychology*, 1967, *19*, 13–25.
Vernon, P. E. Absolute pitch: A case study. *British Journal of Psychology*, 1977, *68*, 485–489.
Ward, W. D. Information and absolute pitch. *Journal of the Acoustical Society of America*, 1953, *25*, 833.
Ward, W. D. Subjective musical pitch. *Journal of the Acoustical Society of America*, 1954, *26*, 369–380.
Ward, W. D. Absolute pitch. Part I. *Sound*, 1963, *2* (3), 14–21. (a)
Ward, W. D. Absolute pitch. Part II. *Sound*, 1963, *2* (4), 33–41. (b)
Watt, H. J. *The psychology of sound*. London and New York: Cambridge Univ. Press, 1917.
Wedell, C. H. The nature of the absolute judgment of pitch. *Journal of Experimental Psychology*, 1934, *17*, 485–503.
Weinert, L. Untersuchungen über das absolute Gehör. *Archiv. für die Gesamte Psychologie*, 1929, *73*, 1–128.
Wellek, A. Das absolute Gehör und seine Typen. *Zeitschrift für Angewandte Psychologie & Charakterkunde-Beihefte*, 1938, *83*, 1–368.
Wynn, V. T. "Absolute" pitch—a bimensual rhythm. *Nature (London)*, 1971, *230*, 337.
Wynn, V. T. Measurements of small variations in "absolute" pitch. *Journal of Physiology*, 1972, *220*, 627–637.
Wynn, V. T. Absolute pitch in humans, its variations and possible connections with other known rhythmic phenomena. In G. A. Kerkut and J. W. Phillis (Eds.), *Progress in neurobiology*, Vol. 1, Part 2. Oxford: Pergamon, 1973. Pp. 111–149.



# 15

# Neurological Aspects of Music Perception and Performance

*Oscar S. M. Marin*


I. Introduction ..... 453
II. Amusia ..... 454
A. Amusia with Aphasia ..... 455
B. Aphasia without Amusia ..... 461
C. Amusia without Aphasia ..... 462
III. Auditory Agnosia and Verbal Deafness ..... 462
IV. General Comments ..... 466
A. Neuroanatomical Basis of Musical Functions ..... 466
B. Perspectives for a Neuropsychological Study of Disorders of Musical Function ..... 469
References ..... 473


## I. INTRODUCTION

An article dealing with the neurology of music should include two levels of inquiry: (a) description of the dissociations and clinical deficits in music perception or performance found in patients with localized or diffuse damage to the nervous system and, on the basis of this data, (b) analysis of normal and abnormal psychological, physiological and anatomical functions, in order to determine the principles and modes by which the brain stores, codifies, or processes music.

Clinical descriptions of abnormalities of musical perception or performance abound in the neurological literature, particularly that of the second half of the last century and the early part of this one. It is significant, however, that in the last 20 years the confident enthusiasm with which clinical deficits were related to functional localization in the brain have become increasingly restrained. This restraint is largely due to a realization that in dealing with brain processes, one can no longer accept the simplistic


Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

concept that the central nervous system can be represented merely as a mosaic of complex mental faculties. One should reject the idea that it is possible to give an adequate account of verbal, musical, or any other complex behavior by anatomically describing a handful of fiber interconnections among a few perceptual or executive functional components. Of even more fundamental importance has been the realization that one cannot study the neuropsychology of complex behavior without a prior understanding of the cognitive structures of the systems involved.

Despite this awareness, progress in the field of neurology of music has been slow. Until recently, a similar simplistic approach controlled the study of the neurology of speech and language; and it still prevails to a large degree in understanding the agnosias and disorders of motor behavior. Disillusionment with the whole mental faculty approach caused a shift to the opposite extreme and emphasized instead the minute analysis of the hardware. This is as if in order to understand the plot of a television program, we analyze the details of the television circuit or the trajectory of the electronic beam projected onto the screen. But behavioral reductionism also had its hour, and most would now agree that from the cognitive standpoint, music is a hierarchical function that requires processing at many interacting levels.

Because some useful information on auditory perceptual processing has begun to emerge from the study of auditory agnosias and verbal deafness, these clinical syndromes will be analyzed here despite the fact that there is still little known concerning the way in which disorders of music perception relate to such cases. Because neuropsychological studies of language and speech have proven their value in the analysis of cognitive processes and because these functions are related to music, we shall frequently use neurolinguistic and psycholinguistic experience as points of reference.

The clinical characteristics of disorders of musical function will be described, but a summary of these clinical experiences will not yet provide a neuropsychologically coherent interpretation in terms of information processing, or localization in the nervous system. Rather, these clinical descriptions will be used as points of departure for critical comments on issues that await preliminary study before fruitful research in this area can proceed.

## II. AMUSIA

*Amusia* is a generic term used to designate acquired clinical disorders of music perception, performance, and reading or writing that are secondary to organic brain damage. These disorders, like other syndromes of higher cortical function—such as agnosias, aphasias, or apraxias—should represent fairly well-circumscribed disabilities and not be the result of a diffuse decay of overall mental capacity as observed in dementia, psychosis, or mental retardation.

The classical literature describing amusia has been reviewed many times in recent years. Of special value are the works of Benton (1977), Jellinek (1956), and Wertheim (1963, 1969, 1977) in English, and of Barbizet (1972), Dorgeuille (1966), and Grison (1972) in French. The German speaking authors have sustained interest in the subject

with works ranging from highly theoretical speculative essays to clinical or anatomoclinical reviews. Notable are the works of Edgren (1895), Feuchwanger (1930), Henschen (1920, 1926), Jellinek (1933), Kleist (1928, 1962), Ustveldt (1937), and Walthard (1927).

While the agnosias (i.e., auditory agnosia, visual agnosia) are clinical nosological terms used to designate modality specific disorders of perception, the amusias are terms referring to attempts to group together a variety of disabilities related to musical function. Some amusias are specifically perceptual in nature; some involve symbolic systems of reading and writing or are based on previously acquired knowledge, while still others comprise complex executive vocal or manual motor activities. We should also note that the degree of proficiency and automatization with which each of these tasks was performed before illness varies considerably in individual patients. It can be anticipated that such a heterogeneous assortment of dysfunctions will not enable us to determine whether basic musical mechanisms are related to consistent lesional topographies. Consequently, as also noted by others (Ustveldt, 1937; Wertheim, 1977), the nosological term amusia is too general; and is thus essentially meaningless for describing, and useless for explaining, the brain mechanisms that may be involved in music functions.

Efforts to improve and clarify the main issues relating to disorders of music perception or performance have utilized four main approaches:

1. Nosological studies attempting subclassifications of the amusias into clinical syndromes that are more closely related to the impairment of basic mechanisms of musical functions (Dorgeuille, 1966; Henschen, 1920; Kleist, 1934; & Wertheim, 1963, 1969).
2. Attempts to define the relationships between disturbances of speech and language (aphasias) and those of music by comparing the clinical syndromes, documenting the clinical coexistence or segregation of these syndromes in individual patients, and correlating the anatomoclinical findings of both disorders.
3. Improvements in methods of assessing premorbid material skill and knowledge in amusic patients that would serve as a reliable base line for the evaluation of deficits (Grison, 1972).
4. Improvement and systematization of the methods of clinical investigation of musical disorders (Dorgeuille, 1966; Jellinek, 1956; Wertheim & Botez, 1959; Wertheim, 1969).

### A. Amusia with Aphasia

It has been stressed that music is a "language" and, consequently, that amusia and aphasia are likely to be closely linked in their neuroanatomical substrates (Feuchtwanger, 1930). Thus, the *criteria* required for their documentation and even their nosological subclassification would be expected to be closely related (Weissenburg & McBride, 1935). Musical disabilities that are predominantly receptive (receptive or sensory amusias) would somehow parallel the receptive and sensory aphasias of

Wernicke; disorders that are predominantly expressed in terms of an inability to vocalize or sing would be somewhat equivalent to Broca's aphasia; disorders of musical writing would parallel agraphias; disorders of musical reading would correspond to alexias; while the inability to remember a familiar tune would be similar to an anomia. Nonparalytic difficulties in playing an instrument would constitute a musical apraxia.

A strong clinical association between aphasias and amusias would have important implications for the anatomical localization of musical functions, since speech and language are strongly lateralized to the dominant cerebral hemisphere (for distribution of speech and language functions, see Marin, Schwartz, & Saffran, 1979). If, on the other hand, the association is only partial or inconsistent, a conclusion that language and music share only some processes or may even be totally independent functions, would be suggested. In this respect, one should remember that neuropsychological similarities between these two functions do not necessarily mean that they must share a common physiological or anatomical substratum. The evolutionary tendency of the central nervous system is toward the assignment of task-oriented neural subsets (Simon, 1967). This subspecialization for speech and music would result in independent neural operators of greater efficiency.

It is not possible to evaluate with certainty the frequency with which aphasias and amusias occur in association. Most of the published clinical cases were preselected, examined only from the standpoint of either aphasia or amusia. Cases of amusia have been analyzed for aphasia more often than cases of aphasia have been analyzed for amusia. Feuchtwanger (1930) found various degrees of aphasia in each of his observations of amusia, and on the basis of his findings suggested that the neurolexical substrates of music and language were intimately related. Since then, a growing number of cases have shown dissociations between the two functions. Dorgeuille (1966), in his study of 26 observations, found 11 cases of amusia with aphasia and two examples of isolated amusia. In reviewing the observations of a large number of clinical articles on amusia, we were able to collect 87 case reports with fairly complete clinical descriptions. Aphasia and amusia coexisted in 33 of these cases. Well-documented examples of amusia without aphasia were found in 19 cases. Amusia and general auditory agnosia were reasonably well-documented in four cases, and there were at least five cases of amusia associated with verbal deafness with only minor or nonexistent signs of auditory agnosia. There were one or possibly two cases of auditory agnosia with a possible musical perceptual disorder but without verbal deafness or aphasia (Nielsen, 1962; Spreen, Benton, & Fincham, 1965).

No attempt will be made here to review the large number of cases already recorded by others. Examples will be selected because they represent dissociations that are clinically characteristic, or because they offer an opportunity to discuss special aspects of the neuropsychological structure of music or cognitive function in general.

The existence of fairly distinct disorders of music perception (receptive amusia) are well recognized. Their association with similar deficits in language perception and comprehension is not exceptional (Feuchtwanger, 1930), but the two functions may often be dissociated (Ustvedt, 1937). Kohl and Tschabitscher (1952) describe a patient

with left hemisphere damage, expressive aphasia, normal comprehension, and predominantly receptive amusia. Wertheim and Botez (1961) also describe the case of a 40-year-old professional violinist who, after developing a sudden right-sided sensorimotor hemiparesis and a mixed but predominantly receptive aphasia, showed a number of difficulties in musical processing. While the patient previously had absolute pitch, he now consistently makes errors in identifying and naming isolated sounds. However, his pitch discrimination was normal even for intervals smaller than a semitone, and vocal and instrumental reproduction were correct. In contrast, recognition and identification of work, composer, and even style of well-known musical themes were incorrect. He was capable of identifying differences in melodies and chords, and to recognize faulty performances of a melody (including errors in tonal modulation); however, he was unable to identify and name intervals in melodic or harmonic contexts. Identification of rhythmic measures, motor reproduction, and dictation of rhythms were also poorly performed. Music reading was impossible in sequence, but the patient was able to name notes in isolation. Reading of musical symbols was normal but rendition of their meaning was often faulty. Singing of isolated sounds or familiar melodies was often incorrect, while evaluation of major and minor scales was usually correct. Some errors were made in copying or transcribing melodies.

This case, summarized in some detail, shows the clinical complexity of some of the deficits and, at the same time, shows the variances in the cognitive nature of the errors made by the patient, even given unvarying exploration of pitch, rhythm, melody, or chord perception. It is interesting to observe that most of the tasks that involved simple perceptual discriminations were apparently performed more successfully than similar tasks that required identification or naming. In this respect, it is of significance to note that the residual aphasic disorder of this patient was also predominantly a lexical difficulty in the form of anomia. The deficits exhibited by this patient could well have been due to a basic lexical disorder common for language and music.

This case illustrates the need to distinguish between specifically perceptual defects, and disorders related to the "linguistic" or symbolic lexical level of the organization of the musical system. A further distinction between musical perceptual and auditory sensory levels will be made later in this paper when observations of auditory agnosias and verbal deafness are discussed. Souques and Baruk (1930) describe a patient with severe Wernicke's aphasia due to extensive left temproal damage. The patient was unable to recognize familiar melodies but was able to play spontaneously pieces of music at the piano, as well as to play major and minor scales. This patient was able to detect minute errors in the instrumental execution of others, and his music reading and writing was preserved. Here again we observe disabilities in the higher level functions of identification and naming, while other musical perceptual functions were intact.

Bernard (1881), cited by Dorgeuille (1966), describes the case of a professional piano teacher who was affected by a severe mixed expressive-receptive aphasia and alexia and was unable to read or recognize music. Despite this, she was able to sing the melody and lyrics of a famous song, and she did so repetitively and automatically.



Benton (1977) cites a similar description by Dalin (1745) of an aphasic patient who was nevertheless still capable of singing with words. Other such cases were reported by Gowers (1875), Knoblauch (1888), and Proust (1866). This contrast between normal and automatic performance of overlearned familiar tasks and an inability to execute new sentences or new motor acts or to sing new melodies is a well-known aspect of the neurological disintegration of behavior.

A dissociation between what Jackson called automatic and propositional behavior can be observed in practically all functions (speech and language, movements, singing). It reminds us of the critical problem that any distinct motor or language task, or any ostensive behavior, may correspond to varying levels of processing depending on an individual's previous experience and the context in which it appears. While a patient may be unable to name an object voluntarily (i.e., a drinking "glass"), he or she may a few minutes later request, "Give me the glass. I am thirsty." When one of our demented patients, whom we have studied over a prolonged period (Schwartz, Marin, & Saffran, 1979) was recently reexamined, we found that her spontaneous speech had totally disappeared. Her spontaneous vocalizations consisted only of iterative sounds that she said in sentence-like sequences with appropriate sound contours to express her emotions and intentions. She was now totally unable to name spontaneously any object, and only exceptionally did she vaguely approximate the phonology of certain words. However, when asked to count aloud, after some prompting by the examiner, she could count perfectly normally, in sequence and in pronounciation, up to 32. This woman, who is suffering from a diffuse bilateral cortical atrophy, was also able to sing, without the lyrics, the entire national anthem with perfect rhythm, intonation, and prosody. The usual explanation given in such cases is that abnormal functions reflect the effect of the damaged left cortical areas while those that are preserved are based on normal right anatomical substrata. But in this case of diffuse cortical damage such an explanation is not determinative, since there is no good reason to believe that her right cerebral cortex was functioning any better than her left language sustaining hemisphere. The researcher must accept that, depending on the automatism acquired in the performance of different tasks or skills, a patient can develop alternative representations at multiple cortical or subcortical loci or perhaps that representations become resistant to anatomical pathological simplification resulting from cortical atrophy by virtue of their widespread anatomical substrates.

Ethological literature is rich with examples of centrally programmed and complex behaviors of long duration that are stored and performed as wholes. Neurophysiology has only recently found similar examples of behavior in higher vertebrates whereas clinical neurology hardly recognizes their importance in clinical situations.

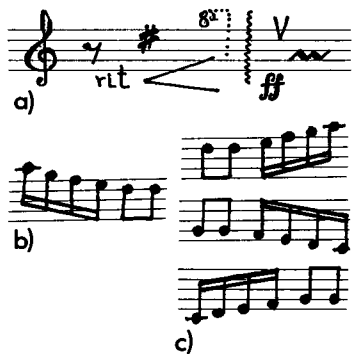
The hypothesis that under special circumstances of automatism or overlearning, a cognitive task may escape the complexity of its original function and may operate in relative independence provides another explanation for the evolutionary specialization of neural subsets. Such a hypothesis should have importance for neuropsychological theory as well as for theories of cognition and cognitive development. Although speech is a learned function in both its receptive and productive aspects, there is evidence that in adult life this phonological-articulatory loop can function with incred-

ible precision in isolation from the rest of the linguistic cognitive organization. In fact, lesions that isolate the speech areas of the left hemisphere may originate a clinical syndrome of echolalia characterized by automatic ability to repeat faithfully any heard speech (Geschwind, Quadfasel, & Segarra, 1965). In such patients there is no evidence of understanding of language, and, in fact, the semantic–linguistic disconnection seems to be partly responsible for the phonological-articulatory automatic independence (Marin *et al.*, 1979; Marin & Gordon, 1979a,b). A similar, but less frequent, phenomenon can be observed in some patients who continuously repeat gestures or movements made by others (echopraxia). Although these examples are taken from pathology, they show how functionally specialized subsets can detach themselves from the original cognitive roots that nurtured their development.

An aphasic patient described by Finkelnburg (1870) was able to play melodies on his violin, but was unable to do so on the piano, which was a less familiar instrument. In music, as in language, such automatisms are of enormous importance and may explain the sparing of some complex musical tasks while others that appear to be more basic and simple are lost. Perhaps the best example of this is seen in the motor skills of instrumentalists or in instrumental musical improvisation. In both cases, success depends to a large extent on the ability to create central programs of long duration based on tactual-motor loops. Performance, thus, demands a minimal requirement of higher cognitive control whereby the motor system performs with almost total independence, allowing the musician's mind to occupy itself with the higher level of musical aesthetic controls. Sudnow (1978), discussing his experience in jazz improvisation, wrote: "My hands have come to develop an intimate knowledge of the piano keyboard, ways of exploratory engagement with routing through its spaces, modalities of reaching and articulatory, and now I choose to go in the course of moving from place to place as a handful choosing."

A frequent and perhaps predictable association is that of verbal and musical alexia: Case 9 of Bouillaud (1865) was agraphic, alexic, and musically alexic, as were the cases of Brazier (1892), Dejerine (1892a,b), Dupre and Nathan's Case 2 (1911) and those of Jellinek (1953) and Wertheim and Botez (1961). Dorgeuille (1966) found disturbances of music reading in six of his collected cases, but only three of these were associated with verbal alexia. Jossmann's case (1927) was unable to read music or to sing but was not aphasic, and his lesion was in the right hemisphere.

Here again, we are confronted with a function, reading, which in no way may be considered simple, and a deficit, which does not correspond to a single altered mechanism or to a well-defined lesional topography. Music reading does not depend only on the ability to recognize the symbolic lexical nature of specific graphic designs such as in Fig. 1A and the assignment to each of them of a referential meaning (symbolic-semantic level). Thus, a study of music reading demands consideration of other processes involved in the perception of the lexical elements of music. Perceptual defects (visual agnosia) may account for some cases of musical alexia. Ballet (cited by Dorgeuille, 1966) tells us, for example, of one of Charcot's patients who suddenly developed an inability to read music because he became incapable of "deciphering" the notes that he saw but could not "understand" any longer.




An element of fundamental importance in reading music is the fact that the script is displayed in a highly bisymmetrical space. For music reading we have to deal not only with the perceptual recognition of elements as in Fig. 1B as such but also with the spatial problems of discriminating them from elements such as in Fig. 1C. Consequently, music reading might be simpler than verbal reading in terms of the number of lexical-semantic items represented in the musical "dictionary." However, verbal reading is more ambiguous in terms of perceptual features of object identification and more complex in terms of its visuospatial organization. The discrimination in music reading of perceptual orientation and symbolic-referential aspects may have important neuropsychological implications for the varieties of musical alexias and their possible anatomoclinical correlates.

There is an important neuropsychological, physiological, and comparative anatomical literature which indicates that form perception, and spatial orientation are two different aspects of vision (Schneider, 1969; Weiskrantz, Warrington, Sanders, & Marshall, 1974). This literature will not be reviewed here, but its implications are obvious: It cannot be expected that music reading depends on a single anatomical substrate or even that its various components will overlap entirely with those of verbal reading whether this be syllabic or ideographic. The greater role of the parietal cortex of the minor hemisphere, so significant in visuospatial functions, seems to be clear in the case of music reading as opposed to verbal reading. If reversals and other disturbances of spatial orientation are prominent abnormalities in cases of developmental dyslexias, one can easily imagine the importance of these factors in the case of music reading. There is, in fact, some evidence that relatives of dyslexic children, who are not dyslexic themselves, often have great difficulties in learning to read music (E. Saffran, personal communication).

One of the earliest reports of amusia is Case 9 of Bouillaud (1865). This was a woman who suffered a sudden right hemiplegia, associated with aphasia. Her articulatory defects ("sort d'ataxie verbale . . ."), possibly a word order syntactic defect, her agraphia and acalculia, were accompanied by a peculiar inability to read music. Occasionally, she was able to play some passages correctly, but she frequently made mistakes in the sequential ordering of the music. The case is remarkable because this

disorder allows for various alternative interpretations. On the one hand, it may be due to the association of a language and a musical disorder that is the manifestation of a common cognitive defect: abnormality in the language syntactic word order and abnormality in the sequence of the musical execution of a tune. On the other hand, the same music reading defect could be explained by a visual spatial disorganization and attentional disorder, similar to the type of verbal reading abnormality observed by Shallice and Warrington (1977). In such circumstances, there is a perceptual disorder in relation to selective visual attention that results in an attenuation of the perceived order of a sequence of letters or words.

Proust (1872) described the case of another woman who became aphasic and who was also unable to vocalize music. The patient could, however, name notes, recognize familiar tunes, and sign musical scales. Although the author did not give details of the patient's language disorder, one can be reasonably sure that comprehension was relatively well preserved and that language production was predominantly affected. In this case, the speech disability was accompanied by a corresponding defect in vocalization, an association not at all uncommon. Similar cases have been reported by Botez and Wertheim (1959), Brazier (1892), Cramer (1891), and Jellinek (1933).

It is not unusual to find cases of severely aphasic patients that are able to sing with normal intonation and rhythm but without lyrics. One of our patients with Broca's aphasia was a very well-educated woman who, despite her severe expressive disorder, was able to sing a number of themes of the classical symphonic repertoire and to sing songs without the lyrics. The dissociation in performance between verbal and melodic tasks gives strong support to the contralateral localization of these expressive functions (Barbizet, 1972).

### B. Aphasia without Amusia

There are at least 12 documented cases of aphasia in professional or amateur musicians in whom musical abilities were not noticeably affected. These cases of total dissociation are of particular interest because they decisively contradict the hypothesis that language and music must share common neural substrates.

Perhaps the most remarkable case is the one reported by Luria, Tsvetkova, and Futer (1963). The patient was a well-known composer and professor of music at the Moscow Conservatory who suffered a cerebrovascular accident that damaged the left temporal and temporal-parietal regions leaving him with a severe sensory aphasia. This patient, despite the aphasia, was able to resume his work as a composer successfully, creating works that were considered by other professional musicians as excellent in quality. Other cases of aphasia without amusia have been reported by Assal (1973), Bouillaud, (1865), Brazier (1892), Charcot (1886, cited by Dorgeuille 1966), Henschen (1926), Lasegue (1872, cited by Dorgeuille, 1966), and Proust (1866). These cases indicate that the neurological substrates of the left hemisphere that are responsible for aphasia do not sustain, to any significant degree, the musical functions of these patients.



### C. Amusia without Aphasia

Cases of amusia unaccompanied by language impairment support the hypothesis of independence of musical and verbal processes. At least 19 such clinical examples have been recorded. A significant number of cases in which there is satisfactory information concerning the site of lesion show damage to the right hemisphere, notably affecting temporal regions. The clinical symptoms are diverse, but the majority of them, as expected, are related to defects in music perception, discrimination, or memory. Dorgeuille (1966) found in his series of 26 patients, eight examples of alterations in memory for melodies, four of disorders of rhythm, and seven musical perceptual defects. In two of the four cases of specific melodic perceptual defects, there was also avocalia. The rest of his music patients had complex combinations of perceptual deficits, disorders of instrumental sound recognition, or disturbances of music reading or writing. Despite Dorgeuille's careful review, no clear nosological picture emerges from his efforts that could be interpreted in terms of basic neuropsychological mechanisms. The correlation between clinical deficits and morbid anatomy is inconclusive. The involvement of the temporal lobes on both sides, including the anterior portions, were, however, prominent anatomical features.

## III. AUDITORY AGNOSIA AND VERBAL DEAFNESS

Auditory agnosia constitutes the auditory perceptual disorder par excellence: in such instances, the association of abnormal perception of noises, speech, and animal sounds with that of music is not at all surprising. Unfortunately, it is common to find that reported cases of auditory agnosia have been only cursorily explored for their musical perceptual abilities. The syndrome usually results from bilateral cortical damage to the temporal cortex of the primary acoustic area of Heschl and surrounding fields (Albert, Sparks, von Stockert, & Sax, 1972; Jeger, Weikers, Shorbrough, & Jeger, 1969; Kogerer, 1924; Laignel-Lavastine and Alajouanine, 1921; Misch, 1928; Pötzl, 1939; Pötzl and Uiberall, 1937; Reinhold, 1948). There are recorded cases in which the syndrome resulted from a unilateral lesion in the right hemisphere. Here, the abnormal perception of nonverbal sounds contrasted with the preservation of normal speech perception and absence of aphasia (Nielsen, 1962; Spreen *et al.*, 1965).

Disorganization of speech perception (verbal deafness or word agnosia) with preservation of normal nonverbal discrimination is well known and is usually the result of localized damage to the acoustic and paracoustic cortical fields of the dominant speech hemisphere.

Typically, patients with auditory agnosia are not truly deaf, but in fact continuously experience chaotic sounds from the environment as well as auditory acoustic illusions and hallucinations. Many patients complain that noises, voices and sounds have become too loud and unpleasant (Arnold, 1943); that sounds are like crackling noises running against each other and creating indistinct, blurred noises; that sounds are incomprehensible and do not "register" in their minds; that they resound in

unpleasant echoes, vibrating or oscillating in intensity, even becoming painful to hear. For example, one of our patients constantly complained that the hinges of the doors needed oiling because they emanated constant unpleasant squeaky noises. Others heard human voices or their own thoughts. Such patients have great difficulty in distinguishing between real and illusory noises, sounds or voices; and like patients with other cortical cognitive disorders (i.e., hemiplegia, cortical blindness, somatoagnosic defects), they have difficulty in gaining full awareness of the extent of their perceptual deficits. One of our patients with visual agnosia insisted on driving his car. Another patient with an auditory agnosia that had lasted for years still insisted on calling and answering the telephone, thus renewing each time his frustration at not being able to comprehend the voices of speakers or the noises produced by the apparatus. The essential deficit here therefore does not seem to be one of threshold perception but rather one of discrimination and pattern organization.

Systematic and thorough studies of auditory deficits are practically nonexistent. This is due not only to extreme difficulties in achieving meaningful communication with such patients but also to their unusual irritability and anxiety. Communication by gestures is usually unsatisfactory and full of ambiguity, and laborious written communication is their only option. In typical cases, reading and writing are normal with only occasional paralexic errors. In some cases, spontaneous speech may show paraphasic errors. Occasionally, there is a tendency to talk too loudly and in a rather forced high pitch. One of our patients developed, together with his auditory agnosia, a high pitch falsetto voice. Another patient had completely incomprehensible utterances mostly consisting of iterative fast repetition of single syllables (i.e., da-da-da, or sh-sh-sh) or combinations that in no way resembled anarthric or typical paraphasic speech. The same patient was unable to sing, imitate animal sounds, or repeat vocal rhythms with syllables. By contrast, his written communication was perfect.

It is difficult to be sure what the relationship is between this abnormality of speech production and the auditory defect, but one cannot fail to remember the importance of acoustic templates for the production of songs in birds (Marler, 1976; Nottebohm, 1975), the effect of the delayed auditory feedback on speech perception (Yates, 1963), or the studies of jargon aphasia (Alajouanine and Lhermitte, 1961, 1964). Auditory agnosics have great difficulty with a number of sound discriminations. Although many of them have absolute thresholds for tones that are only slightly higher than normal, intensity discrimination thresholds are quite poor. Often, noises or sounds are either not heard at all, or they appear at a loud and uncomfortable level. In general, patients have no difficulty in determining whether or not a stimulus has been turned on. Very often, however, they are unable to detect the off-change, and they declare that the stimulus has terminated before its actual end. In successive on–off trials, patients soon lose track of the changes and the direction in which the stimulus operates. This on–off confusion has also been observed by us in intermittent visual stimulations of patients with visual agnosia. Temporal distribution of stimuli and temporal clusters (rhythm) are very poorly discriminated. Patients exhibit total inability to distinguish, describe, or reproduce rhythmic patterns. Binaural sound localization is usually quite abnormal. Recognition of timbre and of complex sounds is

usually hopelessly defective. One of our patients thought that the barking of a dog was a "kind of music," while a locomotive noise was interpreted as human singing, and typewriter tapping as "children playing." Patients often say that they can hear, but that they cannot make sense of the sound.

One patient, who was extensively studied by Albert *et al.* (1972) commented after hearing a bird singing: "I hear the sound but I can't seem to associate it with any of these pictures" (pictures that represented possible sources of the stimuli). Whether these patients have some clear internal auditory representation of the stimulus to which they are exposed is an important but yet incompletely answered question. One patient was exposed to the ticking noise of an alarm clock. We observed that despite the fact that she went so far as to declare that the object made a kind of a "tic tac, tic tac" noise, she was unable to identify the object from among a group of objects that were presented visually or to choose it from a list of alternative names. No other type of naming disorder was present in this patient. Another of our patients could not determine with certainty if a spoken sound "ai" corresponded to the picture of an "eye" or that of a "tie." However, he always pronounced the name of the picture of these or other objects correctly. This same patient was totally unable to provide words that rhymed with a given word or with the name of an object or picture of an object. In this task, despite our efforts to give written instruction and multiple examples, the enormous difficulty encountered by the patient raised doubts as to whether the concept of rhyming itself was still present in the patient's mind.

Many authors have attempted to differentiate levels of perceptual disorganization. From a broad perspective, Vignolo (1969) studied auditory perception in cases of right or left hemisphere damage. He devised tests that would in some cases discriminate the acoustic properties of a stimulus among distractors and in other cases identify the source or the nature of the stimulus. This allowed the author to separate auditory-perceptual from semantico-referential aspects; levels of processing previously suggested by Kleist (1928). Vignolo found that patients with damage to the left hemisphere performed well in the discriminatory tasks and worse in the identification tasks, whereas patients with damage to the right hemisphere performed poorly in the discrimination of meaningless sounds. Albert *et al.* (1972) were impressed that their patient was always correct in selecting the title of a popular tune just heard among multiple choices, but was only 30% correct in producing spontaneously the name of the tune. It is therefore an open question whether Vignolo's left hemisphere damaged patients failed in the identification task because of a semantic referential disorder or because of difficulties in retrieving the verbal lexical tags of the stimuli.

Studies reflecting this dichotomy between identification and discrimination have been performed in the area of language perception by a number of authors (Basso, Casati, & Vignolo, 1977; Blumstein, 1973, 1977). The studies suggest the existence of multiple stages in the auditory perceptual process, a notion that may have important implications for the neurological aspects of music percpetion. Therefore, it seems useful to review briefly what has been observed in the field of verbal deafness—a perceptual auditory disorder that specifically involves a hierarchically organized sys-

tem. Today this entity is well recognized as resulting from a unilateral lesion of the dominant hemisphere that isolates the auditory association areas from the primary auditory projections (Lichtheim, 1885). This syndrome has been described and analyzed in a long series of reports by Albert and Bear (1974), Barrett (1910), Hemphill and Stengel (1940), Klein and Harper (1956), Richter (1957), and Saffran, Marin, and Yeni-Komshian (1976). Saffran *et al.*, made a detailed study of a right-handed, 37-year-old patient who very likely suffered an embolic infarction in the left temporal lobe. As is usually the case with such patients, his illness began with a clincial picture that corresponded to a Wernicke's aphasia including defective comprehension and paraphasic speech. Later on, his speech production improved to normal levels, but the abnormalities of his auditory comprehension remained unchanged. The findings can be summarized as follows:

1. Monaural presentation of previously rehearsed or familiar names resulted in correct perception through either ear.
2. Dichotic presentations of similar words resulted in perception of the stimuli presented in the left ear with total unawareness of the stimuli presented in the right. An increase in intensity of an order of 20 or more dB to the right ear was needed to make the patient aware of the presence of the stimulus. Dichotic presentation of vowel sounds and consonant-vowel combinations gave a strong extinction effect in the right ear. When presented with natural and synthetic speech sounds (vowels and stop consonants), there was severe and complex disorganization of the expected discrimination and identification of these sounds. The picture that emerged was one of a considerable reduction in the number of items that could be identified with some certainty. This reduction involved overextension of some categories relative to others. There were parallelisms between these identification difficulties and errors made in purely discriminative tasks, suggesting that these two phenomena are, at least, closely interdependent. The patient was able to make discriminations and identifications of nonverbal sounds without difficulty. He was also well able to distinguish intonation contours, changes in types of voices, and the number of voices speaking simultaneously. He was quite alert in detecting when a speaker changed without transition from one language to another (English–German–Spanish) despite the fact that he was unable to understand any of these utterances, and all sounded to him as "kind of Greek." A number of tests demonstrated that the patient's speech comprehension was very sensitive to semantic, syntactic, and other contextual factors with occasional impressive improvements in understanding.

The analysis of this case emphasizes aspects of perceptual processing that are currently subject to active research and discussion in the realm of speech and language but that could be equally valid for our study of music:

1. There is strong evidence that speech sounds are decoded by a specialized perceptual apparatus. This specialization relates to the peculiar acoustical composition of speech sounds.


Biblioteca del Real Conservatorio

2. Certain parts of speech (consonants) are categorically identified and perceived. There is evidence that perceptual categorical boundaries may be present before babies begin to hear significant amounts of spoken language and well before they start uttering speech (Eimas, 1971, 1974; Eimas & Corbit, 1973). Categorical learning, however, is not restricted to speech sounds but is applicable to other acoustic events as well: bow-pluck sounds (Cutting & Rosner, 1974), chords (Locke & Kellar, 1973), and musical intervals (Burns & Ward, 1973; Siegel & Sopo, 1975, also see Chapter 8, this volume). Further, categorical perceptual boundaries are present in other modalities (Miller & Morse, 1976), and in the acoustic mode they are present for speech sounds in nonspeaking animals (Morse & Snowden, 1975; Sinnott, Beecher, Moody, & Stebbins, 1976, Walter & Wilson, 1976).

All this suggests that categorical perception is a property of the perceptual cognitive apparatus that is extensively utilized in the codification of information originating in the surrounding world. Its prevalence in speech is a consequence of the consistency and universality of human speech sounds, all of which have developed in accordance with pre-established neural mechanisms.

Perception of speech sounds seems to be hierarchically organized, as many other types of perception (Palmer, 1977). In the case of speech, one can distinguish initially between two levels of analysis: auditory and phonetic (Studdert-Kennedy, 1974). Auditory analysis extracts basic psychophysical parameters of pitch and intensity. For this, no specialized lateralization of perceptual apparatus is necessary. Phonetic analysis, closely lateralized, identifies and analyzes information in terms of phonetic features. Verbal deafness represents a disorganization of this later level with at least a relative preservation of the auditory analysis. Still higher perceptual units and processing levels are operative in speech analysis. Semantic, lexical, and syntactical contexts facilitate perception by limiting the number of possible alternatives and by allowing the listener to predict likely outcomes. Perception in this sense becomes largely a problem of recognition, by matching the incoming stimuli with internal templates (analysis by synthesis) (Neisser, 1967). One can easily see how similar phenomena could be operative in music at levels of motives, thematic, rhythmic configurations, harmonic patterns, musical forms, the stylistic characteristics.

## IV. GENERAL COMMENTS

### A. Neuroanatomical Basis of Musical Functions

With regard to the localization of lesions in cases of disorders of music perception and performance, the following generalizations may be proposed: sensory and perceptual disorders seem to be related to lesions in the temporal lobe of either hemisphere (for a recent review, see Ullrich, 1978). Pure cases of sensory amusia without aphasia are likely to correspond to right-sided lesions. The neuropsychological structure for sensory and perceptual deficits varies. Damage to the left side often generates combi-

nations of aphasia and amusia in which amusia shows a neuropsychological structure that varies from perceptual to symbolic or lexical musical disorders. Musical abnormalities that are usually called expressive have a neuropsychological structure of greater complexity and, as is the case with language production, involve complex perceptual, mnemonic, lexical, and programmative tasks of greater variety and temporal duration. Correlation with locus of lesion is uncertain, and lesions may invo;ve temporal as well as frontal and even parietal regions. To the extent that musical disorders affect temporally organized programs and the nature of the task is related to linguistic functions or vocalization, lesions tend to predominate in the left hemisphere.

In many professional and even amateur musicians, overlearned and automated tasks are frequently spared. These are tasks performed by the average individual only through careful analysis or orderly sequential programming, which in musicians are partially automatisms. This may have important implications in terms of the neural substrates involved.

Some authors have suggested that deficits in music perception are related to right hemisphere lesions whereas deficits in musical expression are more often due to left-sided damage (Wertheim, 1969). This view needs to be examined further: lateralization of functions has developed not only because they are primarily expressive or receptive but rather because required operations need neural substrata "wired" or "programmed" so specifically as to make them impossible to be shared with other functions. Lateralization, as well as physiological or anatomical localization, is the *consequence of specialization*, and this in turn becomes gradually less compatible with sharing. In some cases, specialization and functional segregation result from an expansion in the amount of data to be processed, while in other cases these may be the result of the peculiar nature of the computation that needs to be performed (i.e., temporal order as opposed to spatial distribution). Nonverbal auditory perception may not require detailed exploration in terms of serial order processing. Instead, great effort may be required to extract the components of the stimulus and to decipher its configuration and cognitive *content*.

Despite the fact that all music displays itself "in time," the temporal perceptual task of the listener who hears the recurring rhythmic clusters of a popular tune is quite simple: it consists in merely discovering and identifying a musical segment of very short duration that consists of repetition of a rhythmic, melodic or timbral pattern with hardly any harmonic, instrumental, or contrapuntal complexity. Hardly any analysis of the information is needed, and usually the perceptual task is performed by analysis-by-synthesis as well as by comparing the incoming information against previously stored and familiar abstract models (Deutsch, 1978). Perceptual tasks of this degree of complexity and variability can be performed by units of configurations that are performed well by the right hemisphere. However, perception of more elaborate music requires a greater awareness of the temporal articulation and relation of musical components. One must be aware of thematic elements required to be stored in memory; of melodic, rhythmic or harmonic developments; of contrapuntal relation of

voices; of tonal modulations that shift the tonal resting point; of subtle changes in instrumentation; and of the harmonic and formal structure of the total work of music. As suggested by dichotic listening studies several perceptual operations of this higher degree of complexity in music are preferentially performed by the right ear, left hemisphere processes. Bever (1980) has reached similar general conclusions with regard to the factors that determine lateralization of functions. Studying the operations involved in language and music perception and discriminating holistic from operational ones, he concludes that these correspond better with their right or left preferential processing than by assignment of the total function to either hemisphere.

Executive functions (i.e., singing, complex rhythmic tasks) are fundamentally dependent on serial temporal and *sequential* organization and may depend on processing systems that were lateralized very early (Divenyi & Efron, 1975).

Trained musicians tend to regard music as an ongoing series of interrelated events over longer periods of time (Bever & Chiarello, 1974; Gordon, 1978; Johnson, 1977; Locke & Keller, 1973). This perception requires short- and long-term memory (Deutsch, 1977), internal perceptual representations (Warren, 1974), and strategies of temporal organization similar to that that we observe in motor skills or verbal syntax. When this indepth analysis is required, the evidence indicates that trained musicians tend to process music with their dominant hemisphere (Bever & Chiarello, 1974; Locke & Kellar, 1973; but also Zatorre, 1979).

Thus, it is plausible to hypothesize that in the process of learning musical tasks that involve sequential programs, analogous to most of language, speech or praxis, there would be an initial tendency to share specialized neural processors with most of the other functions. Thus, the preferential left hemisphere *lateralization* found in musicians would be due both to the cognitive "linguistic" structure of classical music, the need for *their* processing along *syntactical algorithms* and in general to cognitive operations that imply related computations and prelearned internal representations.

However, it is also plausible to hypothesize that when such musical programmatic activities are to be developed to a still greater degree (as could be the case in professional composers), new conflicts in processing would arise, and the neurobiological tendency would then be for the development of specialized musical processors that become totally independent from other functions. This often happens to each of us with respect to speech and language. Although in this case there are important genetic, functional, and even anatomical preadaptations (Lenneberg, 1967, 1975; Geschwind & Levitsky, 1968; Marin *et al.*, 1979), it is clear—as demonstrated by the vicarious function of the minor hemisphere following lesions of the dominant hemisphere in infancy and early childhood—that the anatomical substrates to be assigned are not fixed until the function begins to develop. The variability in the degree of separation of the various specialized musical processors in individuals with different degrees of musical sophistication may explain the equal variability in the *lateralization* of musical functions as reflected by cases of severe aphasia unaccompanied by any noticeable alteration of musical abilities. (For a more extensive discussion of neurobiological issues in relation to lateralization of language and cognition, see Marin, 1976; Marin and Gordon, 1979a,b.)

## B. Perspectives for a Neuropsychological Study of Disorders of Musical Function

The optimistic view of classical neurology that musical functions are clearly localized has not been verified. Not only are fixed anatomical representations unattainable for the various aspects of musical perception, memory, or performance, but even basic musical operations—such as chord or melody perception and identification, music reading, or vocalization—seem to be fairly complex processes that are manifested with considerable individual variability (Barbizet, Duizabo, Enos, & Fuchs, 1969; Deutsch, 1978; Dowling, 1972, 1978; Gordon, 1970, 1974). The uncertainty surrounding the field of neuropsychology of music has been attributed to various factors. Many authors comment that because of the lack of standards for musical ability in the normal population as compared with those for speech and language, writing, or reading, one cannot anticipate whether a particular musical discrimination or performance should be expected in an otherwise normal individual. Not only is it difficult to evaluate the capacities of the musically unsophisticated, but in the amateur musician, or even in the professional, it is difficult to obtain strict criteria for normal performance. Experience with professional musicians reveals the uneven distribution of musical talents, acquired skills, or even understanding of the rules of music as a combinational system of perceptual forms. It is not at all clear that the same musical problem is processed by all musicians in the same way with equal depth or even that the problem would be solved by different musicians following similar and consistent cognitive strategies.

Some authors hope that the neuropsychological study of music will advance when studies employ more systematic and sophisticated methods of examination. Improvements have been proposed by many authors (Dorgeuille, 1966; Grison, 1972; Jellinek, 1956; Ustveldt, 1937; Wertheim and Botez, 1959). No essential argument against this approach can be made; however, it should be noted that when a similar approach was adopted in early studies of aphasia, no real progress was achieved merely by sharpening the details of examination. Real progress in the study of aphasia was not forthcoming until we began to uncover the basic structure of language (see for instance, Head, 1926; Weisenburg & McBride, 1935). Just as an elaborate study of phonetic discrimination would in itself fail to provide a better understanding of speech and language production, it is also unlikely that exhaustive psychophysical analysis of rhythmic interval, perception, or intensity discrimination will tell us anything meaningful about musical processing. Substantial headway in the neuropsychology of music can be achieved only with the study of those levels of perception and memory that are truly musical and could serve as a basis for musical information processing schemas. Recent reviews (Deutsch, 1978) and the results of research reported in the present volume indicate that a sufficient basis exists for undertaking such neuropsychological studies. Even so, this progress does not necessarily indicate that a better understanding of musical neuropsychology will delineate "musical centers" or even less well-delineated anatomical substrates for music. Music can be conceived as a communication between composer and listener through an interpreter.

The musical message contains both a *cognitive* and an *emotional* component, and the full effect of such communication is achieved only if what was conceived by the composer is reproduced in the listener's mind. The *medium* is a construction of acoustic configurations of diverse complexity and duration. The basic *cognitive structures* are abstract mental auditory forms that the composer creates and combines, following rules imposed by stylistic traditions. The *acoustic realization* is a performance that is sometimes modified by the idiomatic constraints of various modes of musical expression. As is the case in other branches of aesthetics, the degree of emotional arousal achieved by different individuals has little to do with the formal complexity or subtlety of the music heard. For the unsophisticated person, cognitive complexity of music has no significant correlation with its enjoyment and even less with the emotional changes that the music may induce. People tend to enjoy music more when it is familiar and when its rules of organization are balanced between that which can be predicted with some degree of accurate anticipation and some modest ingredient of novelty (Pribram & Melges, 1969).

Correlation between enjoyment of music and its cognitive complexity is functional for the musically initiated individual. For such an individual, musical language is fully discriminated and is allowed to interact with the individual's own prelearned internal models, thus extending the individual's enjoyment to a gamut of emotions that involve the whole of his or her cognitive experience (Meyer, 1956; Berlyne, 1971). But more often, all that is required for musical familiarity and enjoyment is the simple repetition of easy-to-recognize melodies or rhythmic patterns. Rules of organization and cognitive structure are important ingredients of music, but this degree of complexity is a matter that solely concerns and satisfies the composer, the intellectual, or sophisticated interpreter, or the analytical musicologist who sees in the musical medium an increasingly complex linguistic-like system. The universality of music is a function of its ability to arouse emotions through some organized system of sound patterns, but the complexity of a musical structure may or may not add to its function in communicating emotion. Esthetic pleasure may depend on perceptual cognitive forms and relations, but the emotional effect is largely incorporated into the musical message by a number of expressive changes in the sound that are interpreted culturally by the listener in very much the same way as are other nonverbal modes of communication (e.g., gestures, mimicry). As in speech, the emotional content of music depends greatly on dynamic changes, contour emphasis, agogic accents, vibrato, rubato, changes in intensity and speed, and even on tasteful exploitation of the inaccuracies and idiomatic limitations of the instruments of expression. All these factors add to the cognitive musical structure an intentiality that permits the listener to detect an emotional meaning (Blumenthal, 1977). Notice that in this process of musical performance, the discrepancy between the abstract musical model and its realization can be substantial (Deutsch, 1978).

Music differs radically from language and speech in the following way: linguistic complexity, in phonological lexical or syntactic terms, is directly related to the effect of the communicative message. Full production of the verbal message is achieved by complete programming of all the cognitive linguistic structures. Furthermore, aside

from rules that regulate the temporal flow in which a linguistic message needs to be encoded (syntax), language is largely a stable referential symbolic system that represents our knowledge of the world. Lexical items (words) are symbols that exist not only by themselves in phonological terms, but also as representatives for their referents. So far as these lexical-referential linkages are fixed and the combinatory rules (syntax) are constant and universal for all languages, one can predict that its computations are made through a similarly stable series of psychological processes. From this we would predict some regularity in the corresponding physiological substrates of these processes (Marin, 1976; Marin & Gordon, 1979a,b).

This is not the case in music: Here we are dealing with a system of communication in which each piece of music consists of a few nonreferential items that are combined according to the prevailing stylistic "syntactic" rules of harmony, melody, timbre, rhythm, or musical form. Music is thus a game of combinatory acoustical constructs that the brain of the composer can conceive and that the listener should be able to learn or discover. For the latter, the intelligence of the communication resides in processing the musical message in such a way as to gain conscious or unconscious access to the rules and its forms (Rozin, 1976). Contemporary music, by making use of new combinatory rules and the great novelty of its acoustical units and clusters, obliges the listener to expend greater effort in perceptual preparation and learning. The cognitive challenge in these circumstances require perceptual learning, and it is similar to that which Warren (1974) has described for other nonverbal sound sequences (see Chapter 4).

If neuropsychologists studying language have found it increasingly difficult to confirm the fixed anatomical localizations predicted by classical neurological models, one should anticiapte that the instability of formal musical units and their combinatory rules might make such precise localizations essentially impossible. These fixed anatomical representations are concepts of the past, because the progress that is being made in the structural psychology of music (reviewed extensively in this volume) offers new approaches for the study of the neurological basis of perception of musical structures.

A review of the types of disorder of musical function due to localized cortical lesions reveals that the traditional clinical classifications are too broad. Grouping symptoms of differing complexity under the same nosological clinical heading does not facilitate an understanding of the neuropsychological processes involved in musical function. For this purpose, musical deficiencies could instead be arranged in a hierarchical order. By borrowing some basic principles from the neuropsychology of language and the psychology of music to establish this order, we can distinguish disorders of psychological, categorical, perceptual, lexicosymbolic, and programmative types. Such a classification would distinguish:

1. *Disorders at the acoustical-psychophysical level:*
   perception of stimulus onset and offset
   perception of changes in intensity
   perception of changes in pitch
   sound localization



perception of arousal, orienting response, habituation
perception of sound spectra or timbre
perception of temporal auditory features (Divenyi & Efron, 1979; Hirsch, 1959)

2. *Disorders of precategorical auditory perception:*
   perception of transition changes (Tallal & Newcomber, 1978)
   interval perception
   chord perception (Gordon, 1970)
   pitch discrimination
   intensity discrimination, agogic, and dynamic changes
   visual perception (for music reading)
   visual spatial orientation (for music reading)
   timbre discrimination
3. *Disorders of categorical perception of musical elements:*
   categorical identification of intervals (Burns & Ward, 1973; Siegel & Sopo, 1975)
   categorical identification of chords (Deutsch, 1969; Locke & Kellar, 1973)
   perception of melodies (Ward, 1970)
   categorical perception of octaves and semitonal subdivision of octaves
   scale perception
   octave equivalence
   perception of rhythmic patterns
4. *Disorders of perception of musical configurations:*
   melody perception (Dowling, 1979), identification of motives and themes chord articulation
   melodic cluster perception (Deutsch, 1978), melodic transformation (Dowling, 1972)
   tonal and modal changes, modulation
   timbre recognition and identification of instruments
   rhythmic discrimination and motor reproductions of rhythm
5. *Symbolic musical processes:*
   apprehension of meaning of musical symbols
   naming of notes, solfeggio
   identification and naming of intervals, chords, and rhythms
   singing with vocalization
6. *Learning complex musical perceptual or executive functions:*
   harmonic spatial orientation (in the sense of Schenker), harmonic pivots, harmonic rhythms (Piston, 1948), sequences, and harmonic changes
   thematic melodic and rhythmic transformations (Dowling, 1972), thematic hierarchy and developments
   recognition of musical form
   stylistic characteristics of melodies, rhythm, use of harmonic functions, use of instruments
   recognition of characteristics of individual composers
   recognition of individual musical
   tactile-praxic loop automatisms (Sudnow, 1978)

This tentative outline would not serve the pragmatic goals of the clinician who searches for a nosology of clinical or anatomical predictive value. A neurological grid systemization would only offer a tentative frame for a process oriented study of musical abnormalities.

## REFERENCES


Alajouanine, T., & Lhermitte, F. Les désorganizations des activités expressives du langage dan l'apahsie. Report at the *VIIth International Congress of Neurology, Rome*, 1961.

Alajouanine, T., & Lhermitte, F. Aphasia and physiology of speech. In *Disorders of communications*, Proceedings of the Association Research Nervous and Mental Disorders, Vol. 42. Baltimore, Maryland: Wilkins, 1964.

Albert, M. L., Sparks, R., von Stockert, T., & Sax, D. A case study of auditory agnosia: Linguistic and non-linguistic processing. *Cortex*, 1972, *8*, 427–443.

Albert, M. L., & Bear, D. Time to understand: a case study of word deafness with reference to the role of time in auditory comprehension. *Brain*, 1974, *97*, 373–384.

Arnold, G. Cortikale Hörstorung bei Leitungsaphasie. *Monatschrift Obrenheilkunde*, 1943, *79/80*, 11–27.

Assal, G. Aphasie de Wernicke chez un pianiste. *Revue Neurologique*, 1973, *29*, 251–254.

Barbizet, J. Role de l'hémisphère droit dans les perceptions auditives. In J. Barbizet, M. Ben Hamida, & P. Duizabo (Eds.), *Le monde de l'hémiplégie gauche*. Paris: Masson, 1972.

Barbizet, J., Duizabo, Ph., Enos, G., & Fuchs, D. Reconnaisance de messages sonores: bruits familiers et airs musicaux familiers lors de lésions cérébrales unilatérales. *Revue Neurologique*, 1969, *121*, 624–630.

Barrett, A. A case of pure word-deafness with autopsy. *Journal of Nervous and Mental Disease*, 1910, *37*, 73–92.

Basso, A., Casati, G., & Vignolo, L. A. Phonemic identification defect in aphasia. *Cortex*, 1977, *13*, 84–95.

Benton, A. L. The Amusias. In M. Critchley & R. A. Henson (Eds.), *Music and the brain*. Springfield, Illinois: Thomas, 1977.

Berlyne, D. E. *Aesthetics and psychobiology*. Appleton, New York, 1971.

Bever, T. G., & Chiarello, R. J. Cerebral dominance in musicians and non musicians. *Science*, 1974, *185*, 537–539.

Bever, T. G. Broca and Lashley were right: Cerebral dominance is an accident of growth. In D. Caplan (Ed.), *Biological studies of mental processes*. Cambridge, Massachusetts: MIT Press, 1980.

Blumenthal, A. L. *The process of cognition*. Englewood Cliffs, New Jersey: Prentice-Hall, 1977.

Blumstein, S. E. *A phonological investigation of aphasic speech*. The Hague: Mouton, 1973.

Blumstein, S. E. The perception of speech in pathology and ontogency. In A. Caramazza & E. B. Zurif (Eds.), *The acquisition and breakdown of language: Parallels and divergencies*. Baltimore, Maryland: Johns Hopkins University Press, 1977.

Bouillaud, J. B. Sur la faculté du langage articulé. *Bulletin de l'Academie de Medicine*, 1865, *30*, 752–768.

Botez, M. I., & Wertheim, N. Expressive aphasia and amusia. *Brain*, 1959, *82*, 186–202.

Brazier, M. Du trouble des facultés musicales dans l-aphasie. Etude sur les représentations mentales des sons et des symboles musicaux. *Revue de Philosophie*, 1892, *34*, 337–368.

Burns, E. M., & Ward, W. D. Categorical perception of musical intervals. *Journal of American Acoustical Society of America*, 1973, *54*, 596.

Cramer, K. Zur Lehre der Aphasie. *Archiv fur Psychiatrie & Nervenheilkunde*, 1891, *22*, 141–160.

Cutting, J. E., & Rosner, B. S. Categories and boundaries in speech and music. *Perception & Psychophysics*, 1974, *16*, 564–570.

Dalin, O. Betattelse om en dumbe som kan slunga. K. Swenka Wetensk. *Academia Hand inguar*, 1745, *6*, 114 (cited by Benton, 1977).

Déjèrine, J. Sur la localisation de la cécité verbale avec intégrité de l'écriture spontanée et sous dictée; cécité verbale pure. *Semaine médicale*, 1892, , 88–89. (a)

Dejerine, J. Des variétés de cécité verbale. *Memoires de la Societé de Biologie*, 1892, 27, 1. (b)

Deutsch, D. The Psychology of Music. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception*, Vol. X, Perceptual Ecology. New York: Academic Press, 1978. Pp. 191-224.

Deutsch, D. Memory and attention in music. In M. Critchley & R. A. Henson (Eds.), *Music and the brain*. London: Heinenmann, 1977.

Divenyi, P. L., & Efron, R. Spectral versus temporal features in dichotic listening. *Cortex*, 1979.

Dorgeuille, C. *Introduction a l'étude des amusics*. Thesis, Faculty of Medicine of the University of Paris, France, 1966.

Dowling, W. J. Recognition of melodic transformations: Inversion, retrograde and retrograde-inversion. *Perception & Psychophysics*, 1972, *12*, 417-421.

Dowling, W. J. Scale and Contour: Two components of a theory of memory for melodies. *Psychological Review*, 1978, *35*, 341-354.

Dupre, E., & Nathan, M. *Le langage musical: étude médicopsychologique*. Paris: Alcan, 1911.

Edgren, J. G. Amusie (musikalische Aphasie). *Deutsche Zietschrift der Nervenheilkunde*, 1895, *6*, 1-64.

Eimas, P. D., Signeland, E. R., Jusczyk, P., & Vigorito, J. Speech perception in infants. *Science*, 1971, *171*, 303-306.

Eimas, P. D., & Corbit, J. D. Selective adaptation of linguistic feature detectors. *Cognitive Psychology*, 1973, *4*, 99-109.

Eimas, P. D. Auditory and linguistic processing of cues for place of articulation by infants. *Perception & Psychophysics*, 1974, *16*, 513-521.

Feuchtwanger, E. *Amusie. Studien zur Psychologie der akustischen Wahrnehmung und Vorstellung and ihrer Strukturgebiete besonders in Musik und Sprache*. Berlin: Springer, 1930.

Finkelnburg, F. C. Asymbolie. *Berliner klinishe Wochenschrift*, 1870, 7, 449-451.

Geschwind, N., Quadfasel, F., & Segarra, J. Isolation of the speech area. *Neuropsychologia*, 1965, *6*, 327-340.

Geschwind, N., & Levitsky, W. Human brain: left-right asymmetries in temporal speech region. *Science*, 1968, *161*, 186.

Gordon, H. W. Hemispheric asymmetries in the perception of musical chords. *Cortex*, 1970, *6*, 387-398.

Gordon, H. W. Auditory specialization of the right and left hemispheres. In M. Kinsbourne & W. L. Smith (Eds.). *Hemispheric disconnection and cerebral function*. Springfield, Illinois: Thomas, 1974.

Gordon, H. W. Left hemisphere dominance for rhythmic elements in dichotically-presented melodies. *Cortex*, 1978, *14*, 58-70.

Gowers, W. R. On a case of simultaneous embolism of central retrieval and middle cerebral arteries. *Lancet*, 1875, *2*, 794-796.

Grison, B. Étude sur les alterations musicales an cours des lesions hémisphèriques. Doctoral Thesis. University of Paris, Faculty of Medicine de Creteil, 1972.

Head, H. *Aphasia and kindred disorders of speech*. London and New York: Cambridge University Press, 1926.

Hemphill, R. E., & Stengel, E. A study on pure word-deafness. *Journal of Neurology and Psychiatry*, 1940, *3*, 251-262.

Henschen, S. E. *Klinische und anatomische Beiträge zur Pathologie des Gehirns*. Teil 5: Ueber Aphasie, Amusie und Akalkulie. Stockholm, Nordiska Bokhandeln, 1920.

Henschen, S. E. On the function of the right hemisphere of the brain in relation to the left in speech, music and calculation. *Brain*, 1926, *49*, 110-126.

Hirsh, I. J. Auditory perception of temporal order. *Journal of the Acoustical Society of America*, 1959, *31*, 759-767.

Jeger, J., Weikers, N. J., Sharbrough, F. W., & Jeger, S. Bilateral lesions of the temporal lobe: A case study. *Acta Oto-Larygngologica Supplement*, 1969, *258*.

Jeger, J., Lovering, L., & Wertz, M. Auditory disorder following bilateral temporal lobe insult: Report of a case. *Journal Speech and Hearing Disorder*, 1972, *37*, 523-535.

Jellinek, A. Zur Phäna menologie der Amusie. *Jahrbuch fur Psychiatry & Neurologie*, 1933, *50*, 115-127.

Jellinek, A. Amusia. *Folia Phonetica*, 1956, *8*, 124-149.

Johnson, P. R. Dichotically-stimulated ear differences in musicians and nonmusicians. *Cortex*, 1977, *13*, 385-389.

Jossmann, P. Die Beziehungen der motorischen Amusie zu den apraklischen Storungen. *Monatsbrift für psychiatrie und Neurologie*, 1927, *63*, 239-274.

Klein, R., & Harper, J. The problem of agnosia in the light of a case of pure word deafness. *Journal of Mental Sciences*, 1956, *102*, 112-120.

Kleist, K. Gehirupathologische und localisatorische Ergebrisse über Hörstörungen, Geräuschtaubheiten und Amusien. *Monatscbrift fur Psychiatrie und Neurologie*, 1928, *68*, 853-860.

Kleist, K. *Gebirnpatologie*. Leipzik: Johann Ambrosius Barth, 1934.

Kleist, K. *Sensory aphasia and amusia, tbe Myeloarchitectonic Basis* (English translation by F. J. Fisch & J. B. Stanton). Oxford: Pergamon, 1962.

Knoblauch, A. Ueber Störungen der musikalischen Leistungs-fähigkeit infolge von Gehirnläsionen. *Deutsches Archiv für Klinische Medizin*, 1888, *43*, 331-352.

Kogerer, H. Worttaubheit, Melodientaubheit, Gebärdenagnosie. *Zeitschrift für Neurologie und Psychiatrie*, 1924, *92*, 469-483.

Kohl, G. G., Tschabitscher. Uber einen Fall von Amusie. *Wiener Zeitschrift für Nervenheilkunde*, 1952, *6*, 219-230.

Laignel-Lavastine, M. M., & Alajouanine, T. Un cas d'agnosie auditive. *Revue Neurologique*, 1921, *37*, 194-198.

Lenneberg, E. H. *Biological Foundation of Language*. New York: Wiley, 1967.

Lenneberg, E. H. In search of a dynamic theory of aphasia. In E. H. Lenneberg and E. Lenneberg (Eds.), *Foundation of language development: A multidisciplinary approach*, Vol. 2. New York: Academic Press, 1975.

Lischtheim, M. L. On aphasia. *Brain*, 1885, 7, 433-484.

Locke, S., & Kellar, L. Categorical perception in a non-linguistic mode. *Cortex*, 1973, *9*, 355-369.

Luria, A. R., Tsvetkova, L. S., & Futer, D. S. Aphasia in a composer. *Journal of Neurological Science*, 1965, *2*, 288-292.

Mahoudeau, D., Lemoyne, J., Dubrisay, J., & Caraes, J. Sur un cas d'agnosie auditive. *Revue Neurologique*, 1956, *95*, 57.

Mahoudeau, D., Lemoyne, J. F., Foncin, J. F., & Dubrisay, J. Considérations sur l'agnosie auditive. *Revue Neurologique*, 1958, *99*, 454-471.

Marchand, M. L. Surdi-mutité par lésion symétrique du lobe temporal. *Bulletin de la Societé Anatomique de Paris*, 1904, *6*, 473-475.

Marin, O.S.M. Neurobiology of Language: An overview. *Annals of the New York Academy of Sciences*, 1976, *280*, 900-912.

Marin, O.S.M., & Gordon, B. Neuropsychologic Ascpects of Aphasia. In H. R. Tyler & D. M. Dawson (Eds.), *Current neurology, Vol. 2*. Boston, Massachusetts: Houghton Mifflin, 1979. (a)

Marin, O.S.M., & Gordon, B. The production of language and speech from a clinical neuropsychological perspective. In G. E. Stelmach (Ed.), *Tutorials in motor behavior*. NATO Advanced Study Institute of Motor Learning and Control, June, 1979, Sénanque, France. Amsterdam: North Holland Publ., 1979. (b)

Marin, O.S.M., Schwartz, M. F., & Saffran, E. M. Origins and distribution of language. In M. S. Gazzaniga (Ed.), *Neuropsychology*, Vol. 2. New York: Plenum Press, 1979.

Marler, P. An ethological theory of the origin of vocal learning. *Annals of the New York Academy of Sciences*, 1976, *280*, 386-395.

Meyer, L. B. *Emotion and meaning in music*. Chicago, Illinois: University of Chicago Press, 1956.

Miller, C. L., & Morse, P. A. The heart of categorical speech discrimination in young infants. *Journal of Speech & Hearing Research*, 1976, *19*, 578-589.

Misch, W. Über Corticale Taubheit. *Zeitschrift der Neurologie & Psychiatrie*, 1928, *115*, 567-573.

Morse, P., & Snowdon, C. An investigation of categorical speech discrimination by Rhesus monkeys. *Perception & Psychophysics*, 1975, *17*, 9-16.

Neisser, V. *Cognitive psychology*. New York: Appleton, 1967.

Nielsen, J. M. Agnosia. In A. B. Baker (Ed.), *Clinical neurology*, Vol. 1. New York: Harper, 1962.

Nottebohm, F. A. A zoologist's view of some language phenomena with particular emphasis on vocal

learning. In E. H. Lenneberg & E. Lenneberg (Eds.), *Foundations of language development*, Vol. 1. New York: Academic Press, 1975.

Palmer, S. E. Hierarchical structure in perceptual representation. *Cognitive Psychology*, 1977, *9*, 441-474.

Piston, W. *Harmony* (2nd ed.) New York: Norton, 1948.

Pötzl, O. Zur pathologie der Amusie. *Zietschrift fur Neurologie & Psychiatrie*, 1939, *165*, 187-194.

Pötzl, O., & Uiberall, H. Zur Pathologie der Amusie. *Wiener Klinical Wochenschrift*, 1937, *50*, 770-775.

Pribram, K. H., & Melges, F. T. Psychophysiological basis of emotion. In P. J. Vinken and G. W. Bruyn (Eds.). *Handbook of clinical neurology*, Vol. 3, Disorders of higher nervous activity. Amsterdam: North Holland Publ., 1969.

Proust, A. De l'aphasia. *Archives Generales Medicine*, 1872, *1*, 147.

Proust, A. Archives generales de Medicine. Cited by Henschen (1920), 1866.

Quensel, F., & Pfeifer, R. A. Über reine sensorische Amusie. *Zeitschrift für der Gesamte Neurologie & Psychiatrie*, 1923, *81*, 311-330.

Reinhold, M. A case of auditory agnosia. *Brain*, 1948, *73*, 203-223.

Reinhold, M. An analysis of speech perception in word deafness. *Brain & Language*, 1976, *3*, 209-228.

Richter, H. E. Akustischer Funktionswandel bei Sprachtaubheit. *Archiv Psychiatrie & Nervenkrankheiten*, 1957, *196*, 99-113.

Rozin, P. The evolution of intelligence and access to the cognitive unconscious. In *Progress in Psychobiology & Physiological Psychology*, Vol. 6. New York: Academic Press, 1976.

Saffran, E. M., Marin, O.S.M., & Yeni-Komshian, G. H. An analysis of speech perception in word deafness. *Brain & Language*, 1976, *3*, 209-228.

Schneider, G. E. Two visual systems. *Science*, 1969, *163*, 895-902.

Schwartz, M. F., Marin, O.S.M., & Saffran, E. M. Dissociations of language functions in dementia: A case study. *Brain & Language*, 1979, 7, 277-306.

Shallice, T., & Warrington, E. K. The possible role of selective attention in acquired dyslexia. *Neuropsychologia*, 1977, *15*, 31-41.

Siegel, W., & Sopo, R. Tonal intervals are perceived categorically by musicians with relative pitch. *Journal of the Acoustical Society of America*, 1975, *57*, 511.

Simon, H. A. The architecture of complexity. *Proceedings of the American Philosphical Society*, 1967, *106*, 467-482.

Sinnott, J. M., Beecher, M. D., Moody, D. B., & Stebbins, W. C. Speech sound discrimination by monkeys and humans. *Journal of the Acoustical Society of America*, 1976, *60*, 687-695.

Souques, A., & Baruk, H. Autopsie d'un cas d'amusie (avec aphasie) chez un professeur de piano. *Revue Neurologique*, 1930, *37*, 545-556.

Spreen, O., Benton, A. L., & Fincham, R. W. Auditory agnosia without aphasia. *Archives of Neurology*, 1965, *13*, 84-92.

Studdert-Kennedy, M. The perception of speech. In T. Sebeock (Ed.), *Current trends in linguistics*, Vol. 12. The Hague: Mouton, 1974.

Sudnow, D. *Ways of the band*. Cambridge, Massachusetts: Harvard University Press, 1978.

Tallal, P., & Newcombe, F. Impairment of auditory perception and language comprehension in dysphasia. *Brain & Language*, 1978, *5*, 13-24.

Ulrich, G. Interhemispheric functional relationships in auditory agnosia. *Brain & Language*, 1978, *5*, 286-300.

Ustvedt, H. I. Über die untersuchunghen musikalischen Functionen bei Patienten mit Gehir leiden, besonders bei Patienten mit Aphasie. *Acta Medica Scandinava Suppl.* 1937, *86*.

Vignolo, L. A. Auditory agnosia: a review and report of recent evidence. In A. L. Benton (Ed.), *Contributions to clinical neuropsychology*. Chicago, Illinois: Aldine, 1969.

Walthard, L. Bemerkungen zum Amusie-Problem. *Schweizer Archiv fur Neurologie & Psychiatrie*, 1927, *20*, 295-315.

Ward, W. D. Musical perception. In J. Y. Tobias (Ed.), *Foundations of modern auditory theory*, Vol. 1. New York: Academic Press, 1970. Pp. 407-447.

Warren, R. M. Auditory temporal discrimination by trained listeners. *Cognitive Psychology*, 1974, *6*, 237-256.

Waters, R., & Wilson, W. Speech perception by Rhesus monkeys: The voicing distinction in synthesized labial and velar stop consonants. *Perception & Psychophysics*, 1976, *19*, 285–289.
Weisenburg, T. H., & McBride, K. E. *Aphasia*. New York Commonwealth Fund, 1935.
Weiskrantz, L., Warrington, E. K., Sanders, M. D., & Marshall, J. Visual capacity in the hemianopic field following a restricted occipital ablation. *Brain*, 1974, *97*, 709–728.
Wertheim, N., & Botez, M. Plan d'investigation des fonctions musicales. *L'Encèphale*, 1959, *48*, 246–254.
Wertheim, N., & Botez, M. I. Receptive Amusia: A clinical Analysis. *Brain*, 1961, *84*, 19–30.
Wertheim, N. Disturbances of the musical functions. In L. Halpern (Ed.), *Problems of dynamic neurology*. Jerusalem: Jerusalem Press, 1963.
Wertheim, N. The amusias. In P. J. Vinken & G. W. Bruyn (Eds.), *Handbook of clinical neurology*, Vol. 4. Amsterdam: North-Holland Publ., 1969.
Wertheim, N. Is there an anatomical localization for musical faculties? In Critchley, M. & Henson, R. A. (Eds.), *Music and the brain*. Springfield, Illinois: Thomas, 1977.
Yates, A. J. Delayed auditory feedback. *Psychological Bulletin*, 1963, *60*, 213–232.
Zatorre, R. J. Recognition of dichotic melodies by musicians and non-musicians. *Neuropsychologia*, 1979, *17*, 607–617.



# 16

1982

# Music Performance

*John A. Sloboda*


I. Introduction ..... 479
II. The Nature of Performance Plans ..... 480
III. Acquisition of Performance Plans ..... 483
A. Memorization ..... 483
B. Improvisation ..... 483
C. Transcription ..... 484
IV. The Role of Feedback in Performance ..... 488
V. Social Factors in Performance ..... 491
VI. Summary ..... 494
References ..... 494


## I. INTRODUCTION

This review examines four central questions about music performance. First, how do we best characterize what it is that someone knows when he or she is able to perform a piece of music? Second, how does this person acquire that knowledge? Third, what use does the performer make of perceptual feedback in controlling his or her performance? Fourth, how is music performance affected by social and situational factors? These questions are central for two reasons. First, they relate to all varieties of music performance, regardless of music type, production mode, or cultural setting. Second, they reflect general psychological concerns that are fundamental to the study of any human skill. Several authors (e.g., Martin, 1972; Shaffer, 1976) have argued that music performance can be seen as a member of a family of skills that includes speech and typing, and whose underlying mechanisms share important features.


Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

## II. THE NATURE OF PERFORMANCE PLANS

Whenever someone performs a piece of music, he or she is translating a mental representation or plan of the music into action. We all have such plans. They allow us to sing or hum tunes we know as we go about our daily affairs. What kind of plan is necessary to execute these everyday acts? One possibility would be a list of actions in a specific order and of specific durations. A gramophone record could be seen as one embodiment of such a plan. Each execution of that plan by the same reproduction equipment results in exactly the same performance. Most human beings, however, do not perform music in exactly the same way on each occasion. They are capable of effecting substantial, systematic, and controlled variation in performance. For instance, most people are able to sing a known tune from any starting note within their range (Attneave & Olson, 1971) and are able to comply with requests to sing faster or louder. In addition, someone who knows a piece of music well enough to sing it can usually, without relearning, thereby whistle, hum, or transfer it to an instrument on which he or she is competent, and, if literate, write it down in musical notation.

These observations suggest that a performer's basic plan is something more abstract than a list of actions to be executed and that this plan can be realized in an indefinite number of ways. It seems that the diversity of outputs that can result when a person decides to perform a tune closely resembles the diversity of inputs from which a person can recognize the same tune. Thus, performance theories are constrained in similar ways to theories of recognition (D. Deutsch, 1969). Given these constraints, we could take the performer's plan to be a list of items in which the pitch, duration, and intensity of each note is specified *relative* to other notes. Before such a plan could result in action, absolute values of these dimensions would need to be assigned to the plan, and the appropriate motor actions would then be determined by reference to knowledge about the specific performance mode being used (voice, instrument, etc.).

The performance variations described so far have a common property, that of *context-free* application. In other words, each variation applies in the same way to each note. When a tune is sung faster, each note is shortened by the same proportion; when a tune is sung in a different key, the pitch of each note is raised or lowered by the same proportionate amount, and so on. Some of the most significant variations in music performance are, however, highly *context-sensitive*. These variations are often called *expressive* and include such techniques as rubato, pitch fluctuation, variations in attack, and timbre. It is these variations that prevent performances from appearing "dead-pan," and performers are generally able to vary the degree of expressiveness to suit the requirements of a situation.

For instance, a professional musician practicing a piece to improve his or her technical grasp of it will probably use less expression than in concert, even though overall speed and intensity remain the same. The point about expressive variations is that they cannot be applied either to all notes in a piece of music or to a randomly chosen subset of notes.

Rubato is a deviation from strict tempo, but not any deviation comprises true rubato.

It matters both which notes rubato is applied to and also what kind of deviation is effected (i.e., whether the note comes earlier or later than it should and by how much). The classical conception of rubato is of a variation in tempo that, while bending the basic metrical pulse, does not destroy it (Schoenberg, 1964, pp. 25–40). Thus, for instance, rubato is probably more effectively applied to subsidiary stresses rather than main stresses. Similarly, expressive sharpening or flattening should probably be avoided on notes that are crucial for determining the tonality of a passage. A performance plan that is adequate as a basis for effective expressive variation must therefore contain information about the rhythmic and tonal structure of the music. In such a plan individual notes cannot be thought of as simply separate items standing in some fixed relationship to other individual notes. Rather they must be thought of as elements within superordinate structures whose nature determines the functions of the items. It is hard to see how a musician could perform expressively, except by rote-learning a set of expressive "tricks" for each new piece, unless his or her knowledge of the music were couched in terms of these structures. Sight-reading an unedited score expressively would certainly be impossible.

The view that abstract structural representation is fundamental to musical cognition has motivated work in artificial intelligence by Longuet-Higgins (1972, 1976), Steedman (1977), and Winograd (1968); and some empirical consequences of this view have been explored by Martin (1972) and Shaffer (1976).

Shaffer (1976) discusses keyboard playing in particular, when the two hands play separate notes simultaneously. Keyboard music (e.g., a Bach fugue) is often best described in terms of two or more independent but temporally overlapping, melodic lines whose structures are to a certain extent dissociated from one another, although they will share superordinate features such as meter and tonality. If a performer organizes his or her playing on the basis of a plan that reflects these structures, this performer should, for instance, be able to apply expressive variation to the several lines independently. Certainly expert performers give the impression of being able to do this, while novices tend to apply variation indiscriminately to all simultaneously played notes. Shaffer (1981) presents the first results of a research program examining these specific issues.

Empirical studies on performance variations of this kind are few despite the long availability of the technical means for accurately recording such variations (Seashore, 1938). A difficulty with Seashore's (1938) methods is that they produce visual records by mechanical means. These records must be analyzed by eye before any useful quantitative data can be obtained. More recently methods have been developed of transferring performance details directly to computer, thus making complex analysis much easier (Liberman, Olive, & Zukofsky, 1977; Michon, 1974; Shaffer, 1981). Michon (1974) reports a finding of some interest. He recorded four performances of Eric Satie's *Vexations*, a piece in which the same short motif is played 840 times in succession. He found considerable within-subject variation in the speed at which the motif was played. There was also considerable deviation from strict metricality. He found a "clearcut relationship" between speed and the nature of the deviation from

metricality, although the relationship is not described in detail. The result shows, however, that alterations in speed are *not* achieved simply by altering the absolute durations of notes whose relative durations are specified in the performance plan. Rather, relative durations must be altered for different performing speeds if the same *rhythmic* pattern is to be communicated.

Gabrielsson (1973) analyzed the performances of three expert musicians asked to perform notated rhythms with strict metricality. He discovered certain regular deviations from strictness. For instance, when two eighth-notes appeared in a sequence of quarter-notes, the duration of the first note was always shorter than the second. Liberman *et al.* (1977) argue that such deviations are inevitable if a performer is to communicate rhythmic information to a listener. How else, for example, is a listener to tell whether a sequence of six equally intense notes is to be thought of as two groups of three notes or as three groups of two notes? Notation often answers such questions for a music reader (Gregory, 1978; Sloboda, 1978b), but the reader-performer must translate his or her perceptions into distinctive actions if a listener is to achieve the same perception. The communicative effectiveness of such actions by pianists of varying standards is currently being studied by this author.

Although many facts about music performance demand the existence of a highly abstract performance plan, it would be a mistake to look for such a plan in every instance of music performance. Sometimes the degree of performance variation is less than that that an abstract plan would allow. For instance, a musician competent on two instruments (say flute and oboe) need not always be able to transfer a piece of music from one instrument to the other without some specific practice, in which fingerings were worked out, and so on. Sometimes instrumental performers seem to learn complex fast passages as highly integrated movement patterns that "fall to pieces" if taken at a different speed. Pianists experience considerable difficulties if asked to transpose a piece learned in one key to a new key (Moore, 1962, Chapter 26). Their knowledge of specific motor actions seems to interfere with their transposition efforts. One could multiply such examples indefinitely. What they suggest is a hierarchy of levels of abstraction such that, according to circumstance or accomplishment, performance plans can exist at a higher or lower level in the hierarchy.

In general, what we would call "musicianship" seems to entail the ability to mobilize the higher, more abstract levels in a wide variety of circumstances. "Technique" would, then, be the elimination of "noise" in the transformation of the plan through successively lower levels. For instance, rubato is only possible for performers who both know where they want variation and also have achieved fine enough control over the timing of finger movements to effect this variation accurately.

The above considerations make unclear the extent to which musicological descriptions of musical structure (e.g., Cooper & Meyer, 1960; Forte, 1962) also describe psychological structures in the perception and performance of music (Laske, 1975; Perkins & Howard, 1974). They should certainly be able to provide a rich source of hypothesis to direct future empirical research. It must not, however, be assumed that they have psychological reality for all performers in all situations.

## III. ACQUISITION OF PERFORMANCE PLANS

The ways in which performance plans are acquired can be divided into three broad categories: memorization, improvisation, and transcription.

### A. Memorization

Very often memorization is a seemingly passive process. One hears a tune several times, and then one is able to reproduce it. Most people have experienced the phenomenon of having a snatch of melody "running around in the head." It is sometimes such an inconsequential fragment that they wish they could forget it. On other occasions, the memorization is more deliberate: conscious strategies are invoked to commit the music to memory. In the case of a lengthy work, such a process may take a musician many months. This process has not been studied in depth by psychologists, even though it is of intense interest as an example of a highly complex cognitive achievment. Nonetheless, there exists some incidental data on errors and transformations in musical memory that shed a little light on the memorization process. For instance, ethnomusicologists have shown that tunes handed on by oral tradition suffer significant transformations over geographical and temporal distance (Cohen & Cohen, 1973; Kolinski, 1969; Seeger, 1966; Sundberg & Lindblom, 1976). Such studies show that contour and overall tonality are often preserved while chromatic notes and passing notes are often removed or added. Furthermore, tunes that migrate from one culture to another are often transformed so as to conform with special harmonic and other formal conventions of that culture. It seems that Bartlett (1932) might well have used serial reproduction of tunes rather than of stories to demonstrate the thesis that memory is a constructive process in which people bring their internal schemata to bear on material presented to them (Dowling, 1978).

Errors that show such transformations occurring in the individual can be found in the performances by young children of songs they have learned. For instance, they frequently make the intervals smaller than they should be, while preserving rhythm and contour (Moog, 1976), although it is not clear how many of these errors are attributable simply to poor voice control. A rather different example of a memorization error will be well known to musically literate churchgoers. The hymn tune "St. Denio" contains a line that, in standard hymn books, is notated as shown in Fig. 1(a). Arguably, every competent church organist therefore plays it this way. Every congregation, however, sings it as in Fig. 1(b), unless they have been specifically drilled not to. It seems that their cognition of the melodic structure is such that they "infer" the added passing note (marked *).

### B. Improvisation

Improvisation is analogous to spontaneous speech. In improvisation a novel piece of music is generated according to an underlying set of rules that constrain the final form

Biblioteca de Psicología - UAB

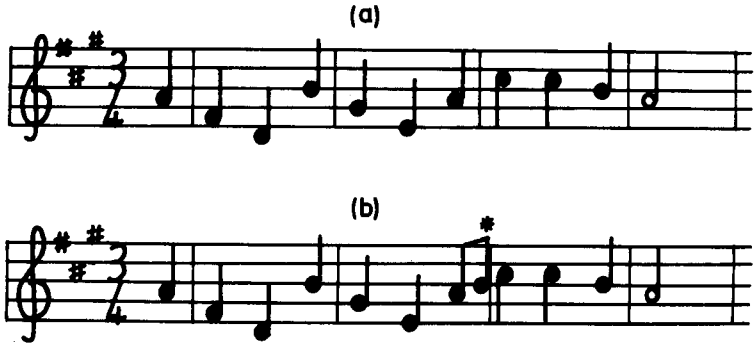



**Fig. 1.** The hymn "St. Denio" (a) as written and (b) as usually sung.

to a greater or lesser degree. This characterization seems to fit musical improvisation no matter what cultural setting is involved. It applies equally to Western jazz (Ostransky, 1977, Chapter 3), the singing of Yugoslavian epic poetry (Lord, 1960; Parry, 1971) and Indian sitar music(Menuhin, 1976, pp. 257–260). Excellence in improvisation results from having "at one's fingertips" a large repertoire of procedures or options for accomplishing some end result within a limited time. In this respect it resembles fluent public speaking, or rapid mental calculation (Hunter, 1979). In such performances, one can often not know the best step to take unless one has determined the result of the previous step. For instance, one may not be able to choose an appropriate closing sequence in a jazz improvisation until one knows where the previous sequence will lead and how much time is available to get "back to base." In mental calculation one can often only take a certain short cut if one can determine that the result of a certain operation will have a certain property, say that of being divisible by 11. The door-to-door sales representative must find rapid answers to novel objections, and so on. Thus, it is clearly not enough for an improviser to know how his or her performance must be structured (e.g., that he or she must negotiate a given chord sequence in a given time); the improviser must have rapid access to a large and well-organized body of knowledge.

Typically, the mediocre improviser solves "real time" problems by choosing "moves" from a restricted and unadventurous set of options, producing a result that is repetitive and uninventive. Even the expert improviser will have a distinctive "style" that reflects the way his or her improvisatory repertoire is chosen from the infinitely large set of possible options (Stewart, 1973).

As with memorization, the complexity of improvisatory skill seems to have inhibited experimental investigation and psychological theorizing. Most workers have chosen to examine a more tractable task, that of transcription.

### C. Transcription

Transcription is the name given to the family of activities in which presentation of musical material in one mode (auditory or visual) is immediatly followed by produc-

tion in another mode (written, vocal, or instrumental). Transcription from auditory input usually occurs only within the context of formal tests of musical attainment. For instance, most college music students are examined in their ability to write down or sing a melody played over once or twice. It is also widely used as a method of teaching songs in the classroom: the teacher will sing a phrase of the song and get the class to sing it back (Franklin, 1972, pp. 89–97). Nonetheless, it is not used to any great extent in real performing situations. Transcription from written or printed text, however, is widely used in music performance. Many contemporary professional musicians perform from text to the near exclusion of any other form of performance.

No definite line can be drawn between transcription and memorization since there is always *some* time lag in transcription between input and output. Also many music readers use printed text to augment existing long-term knowledge of the music acquired from previous exposure to it. It is not uncommon to observe professional performers glance at their music only once or twice each page. In such cases, the text seems to be assisting recall rather than providing new information.

The auditory transcription task approximates to the standard short-term memory paradigm in experimental psychology. Most published research on musical short-term memory, however, has used recognition, rather than recall, measures. This literature has been reviewed by Deutsch (1977) and Dowling (1978) and goes a considerable way towards establishing the salient dimensions along which listeners discriminate and classify melodies. It does not necessarily provide a basis for understanding how listeners asked to recall melodies will memorize them. Unpublished recall data obtained by this author suggests that unskilled adults seldom retain the contour, pitch ratios, or tonality of the original in any simple way. Rather, fragments of a melody are distorted and recombined in complex ways to make up a response that is often, on any simple analysis, quite unlike the original. A method is needed of capturing unfocused intuitions about recall in a formal description procedure before any real sense can be made of such data. The recent explosion of methods for analyzing prose recall (e.g., Bower, 1976; Kintsch, 1976; Harris & Terwogt, 1979) offers some encouragement in this direction.

Transcription from text is usually continuous so that the performer is reading one note while executing another note read earlier. Typically, a good sight-reader will execute a note about 2 sec after reading it. This estimate is derived from studies of the eye–hand span (EHS) in proficient instrumental performance (Sloboda, 1974a, 1977). In these studies subjects were asked to read melodies at predetermined speeds. On each trial the text was removed at a precise point in the performance known in advance to the experimenter only. Subjects had to continue playing until they had executed all the music read up to that point in time. The EHS of each subject was obtained by counting the number of notes played after the text was removed. The EHS of good readers was typically six or seven notes while that of poorer readers was only three or four notes. In addition, EHS was affected by the nature of the material being read. It decreased for meaningless (atonal) material and showed a tendency to expand or contract to coincide with phrase boundaries. It was hypothesized, therefore, that the greater span of fluent readers is related to the ability to perceive

superordinate structures in the text. If, indeed, knowledge of such structures is necessary for forming adequate performance plans, it follows that psychologically effective units in the melodies used must comprise six notes or less on average since the reader would need to identify at least one whole unit before starting to perform it.

Evidence confirming the importance of structural units in reading comes from studies of eye movements in piano sight-reading carried out by Weaver and his associates (Van Nuys & Weaver, 1943; Weaver, 1943). They showed that the sequence of fixations is to some extent determined by the nature of the music. Contrapuntal music, in which horizontal sequences of notes are structurally significant, elicited more horizontal fixation sequences. Homophonic music, in which vertical chords are structurally significant, elicited more vertical fixation sequences. Thus, proficient readers tend to take in a complete musical unit in a series of successive fixations rather than sample from two or more units concurrently. This is an optimal strategy for performance if it is important to identify the next structural unit as quickly as possible.

Transcription errors also reveal something about the performance plan. Although some errors are clearly due to "peripheral" factors (indistinct text, tremor, miscalculated leaps, etc.) a substantial body of sight-reading errors can be explained only by recourse to structural considerations. Wolf (1976) describes one case in which a misprint in a frequently played piano score went unnoticed for many years. Most readers unconsciously substituted the musically correct note for the misprint. It was only when a poor reader actually played what was printed and was criticized for so doing, that the misprint was brought to light. Sloboda (1976b) studied this "proofreaders' error" in an experiment where little-known, highly conventional classical keyboard pieces were rewritten so as to introduce a large number of deliberate notational pitch errors. These errors were distributed equally between the two staves and between various positions in musical phrases. Subjects were proficient readers whose instructions were to play the pieces exactly as written. Not a single subject was able to comply with this request. All restored some errors to their original pitch without noticing it. In addition, on a second playing of the same pieces, subjects made even more unconscious restorations. So, it seems that sight-readers' performance plans are not simple transcriptions of the printed page; rather they also take account of musical "grammar" in order to allow a musically correct performance to result from a musically incorrect score. Another finding was that most restorations occurred in the middle of musical phrases. Exact transcription was most likely at beginnings and ends of phrases. It may be argued that, in this type of music, beginnings and ends of phrases carry most information, whereas middles are, to a certain extent, determined by the surrounding context and the rules of music. Thus, proficient readers seem to have developed selective strategies for transcribing text that correspond to structural constraints in the music.

Written transcription of pitch notation has been studied in several experiments by Sloboda (1976a, 1978a). Subjects were presented with briefly exposed arrays of pitch symbols on staves that they then had to attempt to reproduce onto blank staves. Since the basic transcription skill can be taught in a few minutes, this procedure allowed the comparison of music readers with musical illiterates. Two main results were ob-

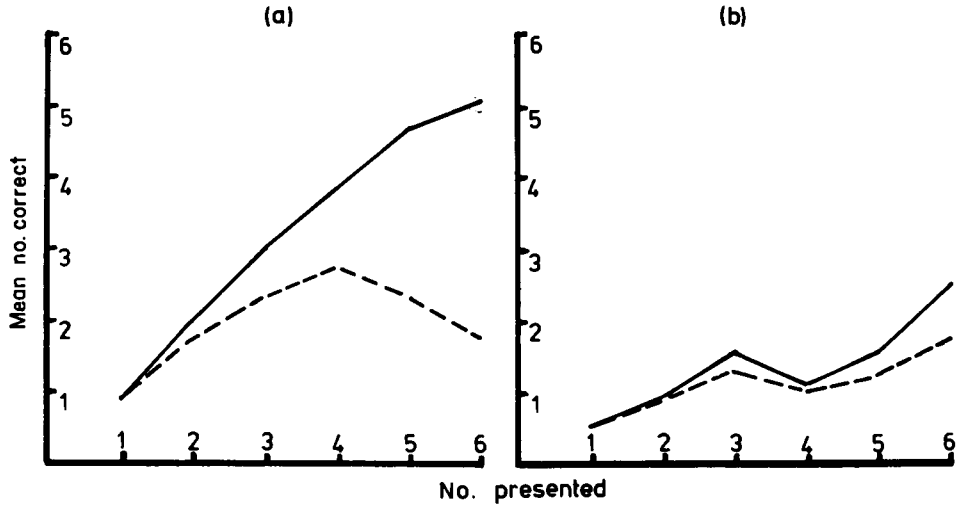



**Fig. 2.** Mean number of notes correctly reported as a function of number of notes displayed (a) at an exposure duration of 2 sec and (b) at an exposure duration of 20 msec. Full lines are musicians; dashed lines are nonmusicians. Adapted from Sloboda (1974a).

tained. When exposure duration was as long as 2 sec, a gross superiority in exact recall was shown by the musicians (Fig. 2a). For instance, when six items were presented, many musicians recalled all six correctly while many nonmusicians recalled one or none correctly. Second, when exposure duration was as short as 20 msec, there was no difference between subject groups in the number of notes correctly recalled [Fig. 2(b)], but musicians were significantly better at preserving the contour of the stimulus pattern. These results suggest that musicians are employing two types of coding procedure not available to nonmusicians. First is a rapid identification of contour (which could be useful in performance, for example, when planning direction of hand movement). Second is a slower identification of the exact position of component notes that allows access to abstract musical codes for holding the notes in memory. These processes allow musicians to overcome memory limitations that prohibit nonmusicians from transcribing much more than one note at a time.

Indirect confirmation that pitch symbols are coded in different ways by musicians and nonmusicians comes from laterality studies by Oscar-Berman, Blumstein, & De Luca (1976) and Salis (1980). In both studies musicians displayed right visual field superiority for report of tachistoscopically presented musical notation. In contrast, report of dot patterns (which would include pitch symbols as seen by nonmusicians) typically shows left-field superiority (McKeever & Hulig, 1970; Salis, 1977). Thus, a shift in hemispheric involvement seems to result from the processing of visual patterns in a musical mode, suggesting mobilization of different cognitive mechanisms. This resembles shifts found when visual material is coded verbally (Cohen, 1977, Chapter 8). Other analogies between language and music reading are discussed more fully by Sloboda (1974a, 1978b).

## IV. THE ROLE OF FEEDBACK IN PERFORMANCE

Performers of music intend to produce a specific auditory experience in listerners. They, therefore, can be expected to check that the results of their actions are in conformity with their intentions and, if not, to take corrective action. How and to what extent do they do this? A major technique used in attempting to answer this question involves disrupting normal feedback by providing delayed auditory feedback (DAF). For instance, J. A. Deutsch and Clarkson (1959) showed that auditory feedback has a crucial role in determining the nature of the vibrato that occurs when subjects are required to sing a sustained note. Both the amplitude and the frequency of pitch change is progressively increased when auditory feedback is delayed by amounts from .10 sec up to .50 sec: They postulate that the vibrato is in part the result of the functioning of a control loop serving to keep the voice on a particular note (see also Clarkson & Deutsch, 1966). When DAF is absent, pitch traverses the distance between its highest and lowest value (one half-cycle) about every .10 sec. If each reversion towards the central pitch is the result of a control adjustment, as Deutsch and Clarkson argue, these adjustments must be made every .10 sec. Rostron (1976, 1978) has argued that such control of the vibrato should be increasingly disrupted by DAF of up to .10 sec and should show no systematic effect with longer delays. He failed to find any effect of DAF below .10 sec and argues that the data are more consistent with an "envelope" theory of pitch control in which less frequent pitch adjustments (about every .50 sec) are superimposed on vocal tremor produced independently of the control loop (e.g., by "physiological wobble" or intensity fluctuations).

However, as Clarkson and Deutsch (1978) point out, the conditions under which DAF is presented are crucial for determining whether the subject will actually use it to control his or her performance. For instance, subjects are less likely to do so if they can still hear their own immediate auditory feedback (IAF) or if the DAF is distorted. Since different methodologies were used in the two laboratories, it is hard to know whether DAF was functioning equivalently in the two situations. However, an observation by Rostron (1978) suggests that the dispute may not have great relevance to most music performance situations. When a sung note is initiated, pitch fluctuation is extreme for the first .50 sec until it reaches the "steady state" studied above. In much vocal music many notes do not last significantly longer than .50 sec, and some last for much less time. Since these initial fluctuations appear to settle rapidly on the intended note, singers must be making control corrections considerably earlier than .50 sec after note onset. Thus, the dispute outlined above turns on whether a rapid control loop is switched off after about .50 sec and not on whether the control loop exists.

A more realistic performance situation has been studied by Gates, Bradshaw, and Nettleton (1974). Keyboard players learned to perform pieces as fast as they were able under normal feedback conditions. They were then required to play the pieces with IAF and DAF presented simultaneously. It was found that total playing time increased for all delays between .10 and 1.05 sec, although the maximum increase occurred at about .20 sec. One cause of the increased playing time was the tendency

to repeat individual notes or to insert an extra note that, for instance, extended a scale passage one note beyond its correct conclusion. The fact that insertions and repetitions never comprised more than a single note suggests that auditory feedback is, in this instance, applied to the level of performance organization in which motor commands for playing individual notes are specified rather than to higher levels in which performance units or groups are specified. If the latter were the case, one might expect repetitions or disruptions of whole groups of notes.

Gates and Bradshaw (1974) had subjects play on the keyboard of an electronic instrument whose speakers could be switched off to eliminate IAF. This condition produced performance indistinguishable from one in which IAF was provided, although both these conditions were superior to conditions in which DAF was present. This is an important result since, if it does not matter whether or not subjects obtain IAF when DAF is absent, the disruption caused by DAF can hardly be due *simply* to interference with IAF. The situation must also make it impossible to use whatever cues are used in the absence of IAF (i.e., kinesthetic or visual feedback). One possible explanation for the effect would be that "auditory capture" operates such that whenever a discrepancy arises between feedback from hearing and some other modality, hearing dominates. It would be interesting to know whether vision dominates over touch in the absence of auditory feedback. This could be tested by presenting delayed visual feedback through a television monitor. There is reason to suppose that vision *is* important for performance, even when auditory feedback is available.

An indirect argument in favor of the importance of vision is that performers who cannot inspect the movements responsible for sound production (i.e., singers) seem to be at a particular disadvantage to other performers (i.e., instrumentalists) when it comes to the acquisition of technique. Singing teachers are more prone than most to resort to prescriptions (e.g., "try to sing more in the head") that turn out to have no meaningful physical or physiological basis (Sundberg, 1974). These prescriptions may, nevertheless, be necessary psychological aids for the singer in conceptualizing his or her actions.

More direct evidence for the importance of vision is provided by studies that show that augmented visual feedback can have quite striking effects on the acquisition of technical skills. For instance, Tucker, Bates, Frykberg, Howarth, Kennedy, Lamb, and Vaughan (1977) describe a computer based interactive aid for musicians that allows a passage played on the keyboard of an electronic organ to be displayed visually in a modified notation that reflects the exact duration of the notes played. With the aid of this feedback, a pianist of concert standard improved his performance of a triple trill quite dramatically within minutes (Fig. 3). Similarly, Basmajian and Newton (1974) showed that with visual electromyographic biofeedback, wind players learned within minutes to suppress or activate specific parts of the buccinator muscle in the cheek. Singers can also be helped in a similar way (Fourcin & Abberton, 1971).

One important aspect of performance that musicians must control is the intensity of sound produced. Composers do not often specify absolute values for intensity, but they establish relative levels of intensity through the use of conventional dynamic instructions. The most frequently used, in order from softest to loudest, are pianis-

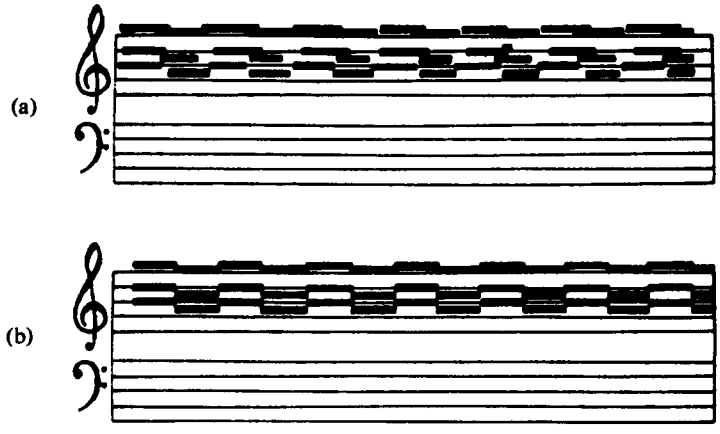



**Fig. 3.** Triple trill played by concert-standard pianist (a) when first played and (b) as played within minutes of visually and aurally interacting with the system. From Tucker *et al.* (1977).

simo, piano, mezzopiano, mezzoforte, forte, and fortissimo. Such a classification demands that a performer be able to perform with at least six discriminable levels of intensity. Patterson (1974) describes some studies of the intensity ranges of professional symphony orchestra players. It appears that intensity changes of as much as 5 decibels go unnoticed in musical performance; therefore, performers need to span an intensity range of at least 30 decibels if they are to get the six basic dynamic levels across to listeners. Such a range is certainly within the capability of most instruments. However, in a sample of professional bassoonists asked to produce their loudest and softest comfortable notes, the largest difference found was 17 decibels, a range capable of encompassing only three dynamic levels. In defense of these players, it has sometimes been claimed that performers achieve perceptible changes between dynamic levels by altering the quality (timbre) rather than the intensity of their sound, at least in some situations. This could not be the case in the studies described by Patterson since listeners were unable to distinguish between notes produced under differing dynamic instructions when playback intensity was equal for each note. Thus, it appears that professional players may not be making effective dynamic changes at points where the composer demands it. This suggests *either* that their own use of auditory feedback is more sensitive than that of most listeners, that they are using inappropriate feedback (such as how hard they feel themselves to be blowing), or that they are simply not using feedback at all. Lack of dynamic range among performers of all standards is such a common phenomenon that it seems unlikely that the difficulty is technical. It could be that, for some reason, intensity is not a highly salient dimension of feedback for performers.

A final type of feedback relates to ensemble playing, where performers need to keep their playing in synchrony with that of other performers. In rigidly metrical music, this problem is, maybe, not serious, but in music in which the speed fluctuates (as in a

rallentando) the performer needs to make complex predictions about when the next note of a colleague is going to come on the basis of previous notes. For example, a pianist accompanying a singer may well find that the singer introduces different degrees of tempo fluctuations in successive performances of the same song. The expert accompanist is rarely "thrown" by this. Experimental studies of synchronization are at present being carried out by Vos.[1] In these studies a human subject and a computer attempt to tap in synchrony with one another. By varying the models on which the computer bases its attempts at synchrony, it is hoped to discover which one "fits" the human performer most precisely.

## V. SOCIAL FACTORS IN PERFORMANCE

Very often a musician is performing with or for other people. The performance situation and the musician's perception of that situation can have a profound effect on the performance itself. Perhaps the most well-known example of this is the phenomenon of "stage fright," an unpleasant state of nervousness experienced before and during public performance by many musicians. This state can have physiological correlates that are highly detrimental to the skilled motor control required in many performance situations (e.g., muscle tremor, excessive sweating). Bochkaryov (1975) studied contestants in international music competitions and found that although the less successful contestants displayed more anxious behavior and reported feeling more nervous than successful contestants, they did not appear to show more arousal (as measured by galvanic skin response, temperature, and heart rate). Rather, the distinctive feature of the successful contestants was a heightened arousal *during* the performance as compared to just prior to performance. In contrast, less successful contestants showed little difference in before performance and during performance measures. Bochkaryov hypothesizes that successful contestants are those who are able to mobilize arousal specifically for their performance, whereas the less successful suffer from a nonspecific arousal that they cannot control and that manifests itself as stage fright.

Since the physiological correlates of stage fright seem mainly a result of the action of the sympathetic nervous system and since these very correlates often confirm the performer's feeling of impending failure thus increasing anxiety further, it has occurred to some that the circle might be broken by the use of drugs that inhibit the action of the sympathetic nervous system while leaving central nervous system functions unimpaired. A suitable drug for this purpose would be one of the $\beta$-adrenoceptor-blocking drugs such as oxprenolol (Turner & Hedges, 1972). James, Griffith, Pearson, and Newbury (1977) administered either oxprenolol or a placebo at random to 24 music students just before a stressful public performance. On the

[1]Vos, P. G. Synchronization in duo tapping tasks. Paper in preparation. Psychologisch Labatorium, Katholieke Universiteit, Erasmuslaan 16, Nijmegen, The Netherlands.

Biblioteca del Real Conservatorio Superior

second day of testing, those who had received the placebo on the first day now received the drug and vice versa. The performances were rated by two professional musicians who did not know which subjects had received the drug. It was found that oxprenolol caused a significant improvement in performance over the placebo condition. It also caused a significant fall in pulse rate measured immediatly before performance. Finally, the subjects who received the drug reported feeling calmer before the recital and happier with their performance.

It appears, then, that oxprenolol has a double effect on performance. One is purely physiological: it affects mechanisms associated with detrimental phenomena such as tremor. The other is cognitive (and resembles the effect described by Schacter and Singer (1966) but in reverse). Because performers perceive themselves to be less "trembly" than usual, they conclude that they are less worried about their performance and are able to set about the performance with greater confidence. The existence of drugs such as oxprenolol is bound to raise ethical problems for performers, especially in competitive situations, since it can be argued that those who resort to drugs have an unfair advantage over those who attempt to control nervousness consciously. On the other hand, there may be constitutional factors that make the control of nervousness easier for one performer than for another. In this case, the naturally placid performer has the unfair advantage.

In many performing situations there is not just one performer "pitted against" a more or less threatening audience; there is instead a group of performers involved in a cooperative effort, and, so, the performers themselves can have an effect on one another. The focal point of many Western performing groups is a conductor who rehearses the performers and provides cues for timing, intensity, and expressive playing during the performance itself. A study of the attitudes of professional symphony orchestra players towards their conductors by Faulkner (1973) illustrates the need of a conductor to communicate effectively with and gain the respect of his or her players. Players are very rapidly aware when a conductor is inadequate (although less able to characterize that inadequacy precisely) and respond with sloppy playing and lowered work effort. Effective conductors use a rich array of verbal and nonverbal gestures to guide and respond to their players. Razhnikov (1973) has described and quantified some of the major rehearsal techniques used by a small sample of professional conductors, but little attempt has been made to articulate the nature of the two-way interaction between conductor and players. That a two-way interaction is essential is evident from a remark of one of Faulkner's (1973) subjects:

> They (the poorer conductors) don't communicate. They are of no use to us. And then maybe they're too distant, you can never tell whether you're doing it right or wrong; they never respond when you respond to them. I guess after a while you stop watching them, you stop paying attention to what they're doing.

One of the only studies to examine conductor behavior experimentally (Yarbrough, 1975) yields results that appear to be at odds with the perceptions of Faulkner's (1973) subjects. In Yarbrough's experiment a variable named "conductor magnitude" was manipulated. The "high-magnitude" conductor maintained eye contact with per-

formers, used expressive gestures and facial expressions of approval and disapproval while the "low-magnitude" conductor never looked at individuals and used no expressive gestures or expressions. Each conductor took the same group of singers through an equivalent amount of material chosen from different parts of the same work. It was found that although performers preferred the high-magnitude conductor, there was no effect of magnitude on attentiveness or on performance of the music. Yarbrough concludes that the conductor "is not the source of reinforcement that maintains the appropriate behaviour. Instead it is suggested that music itself functions as the reinforcement which maintains appropriate behaviour."

The reason for this apparent discrepancy is probably that neither the subjects nor the situations in the two studies are really comparable. Yarbrough's subjects were college students in amateur choirs who attended by choice and who were learning new music under a conductor whom they were required to endure for a few minutes only. On the other hand, Faulkner's typical subject was a seasoned professional whose presence at a rehearsal was a contractual obligation, playing music he or she had probably played many times before, under a conductor whom he or she would probably have to endure for many months. It is hardly surprising that the motivations and effective reinforcements for the two groups were rather different. In addition, it is a common feature of amateur singers that they fail to attend to nonverbal cues from a conductor, especially when learning new music. They tend to "bury their heads" in the score and have not learned the trick, second nature to professionals, of looking at the conductor "over the top of their copy." In such circumstances, it is hardly surprising that conductor behavior has so little effect on them.

One of the most elegant demonstrations of the effect of social factors on music performance is a study by Weick, Gilfillian, and Keith (1973). Two jazz bands of equivalent standard were asked to rehearse two new pieces (A and B). These pieces were by the same arranger and were of similar difficulty and style. The crucial experimental manipulation was the description of each piece given to performers prior to the rehearsal. Players were told that they had been invited to try out two pieces that a major publishing company was thinking of promoting. Each piece was accompanied by one of two press releases incorporating a "serious" or a "nonserious" description of the arranger. The serious release imputed high prestige and relevant experience to the arranger while the nonserious release imputed low prestige and little relevant experience to the arranger. In one band, piece A was accompanied by the serious press release and piece B by the nonserious one. In the other band, the press releases were exchanged. The results showed that for both bands, the nonserious piece was less well performed, liked, and remembered. This finding suggests that a significant proportion of the hostility displayed by performers and listeners alike to, for instance, certain styles of contemporary music stems, not from the music itself, but from the perceptions of the credibility of such styles established before hearing the music. And, of course, if performers, through their own attitudes towards a piece of music, perform it badly, the listener is provided with an objective reason for rejecting the music. In other words, performers' evaluations of music are liable to act as self-fulfilling prophecies.

## VI. SUMMARY

An adequate theory of music performance must account both for the range of variations in output that a performer can display on successive occasions of executing the same piece of music and also for the limitations on these variations that specific situations impose. The performance plans that most people formulate must be couched, at least in part, in abstract tonal and rhythmic form rather than in terms of specific motor sequences or even sequences of items related by relative pitch and duration. This has consequences for the way in which performance plans are acquired. For instance, memorization and transcription errors can often be accounted for by reference to abstract structural properties of the music. Musical text appears to be read in structural "chunks," and improvisation displays structural regularity. As in all purposive sequential behavior, music performers make frequent checks that their intentions are actually being realized. Delayed auditory feedback disrupts this process while various types of augmented feedback improve performance. Performance plans are not executed in isolation from the rest of the performer's mental apparatus. Optimum levels of attention and arousal are determined by the performer's situation and perceptions of his or her situation. To achieve the best possible performance, a musicians need to avoid the excessive arousal of stage fright while having a positive enough view of their coperformers and the music being performed to allocate adequate attention and effort to their performance.

## REFERENCES


Attneave, F., & Olson, R. K. Pitch as a medium: A new approach to psychophysical scaling. *American Journal of Psychology*, 1971, *84*, 147–166.

Bartlett, F. C. *Remembering*. London & New York: Cambridge University Press, 1932.

Basmajian, J. V., & Newton, W. J. Feedback training of parts of buccinator muscle in man. *Psychophysiology*, 1974, *11*, 92.

Bochkaryov, L. L. The psychological aspects of musicians' public performance. *Voprosy Psikhologii*, 1975, *21*, 68–79.

Bower, G. H. Experiments on story understanding and recall. *Quarterly Journal of Experimental Psychology*, 1976, *28*, 511–534.

Clarkson, J. K., & Deutsch, J. A. Effect of threshold reduction on the vibrato. *Journal of Experimental Psychology*, 1966, *71*, 706–710.

Clarkson, J. K., & Deutsch, J. A. Pitch control in the human voice: a reply to Rostron. *Quarterly Journal of Experimental Psychology*, 1978, *30*, 167–169.

Cohen, A., & Cohen, N. Tune evolution as an indicator of traditional musical norms. *Journal of American Folklore*, 1973, *86*, 37–47.

Cohen, G. *The psychology of cognition*. London & New York: Academic Press, 1977.

Cooper, G. W., & Meyer, L. B. *The rhythmic structure of music*. Chicago, Illinois: Chicago University Press, 1960.

Deutsch, D. Music recognition. *Psychological Review*, 1969, *76*, 300–307.

Deutsch, D. Memory and attention in music. In M. Critchley & R. A. Henson (Eds.), *Music and the brain*. London: Heinemann, 1977.

Deutsch, J. A., & Clarkson, J. K. Nature of the vibrato and the control loop in singing. *Nature (London)*, 1959, *183*, 167–168.

Dowling, W. J. Scale and contour: Two components of a theory oa memory for melodies. *Psychological Review*, 1978, *85*, 341-354.

Faulkner, R. R. Orchestra interaction: Some features of communication and authority in an artistic organization. *The Sociological Quarterly*, 1973, *14*, 147-157.

Forte, A. *Tonal harmony in concept and practice*. New York: Holt, 1962.

Fourcin, A. J., & Abberton, E. First applications of a new laryngograph. *Medical and Biological Illustration*, 1971, *21*, 172.

Franklin, E. *Music Education: P ychology and Method*. London: Harrap, 1972.

Gabrielsson, A. Performance of rhythm patterns. *Scandinavian Journal of Psychology*, 1974, *15*, 63-72.

Gates, A., & Bradshaw, J. L. Effects of auditory feedback on a musical performance task. *Perception & Psychophysics*, 1974, *16*, 105-109.

Gates, A., Bradshaw, J. L., & Nettleton, N. C. Effect of different delayed auditory feedback intervals on a music performance task. *Perception & Psychophysics*, 1974, *15*, 21-25.

Gregory, A. H. Perception of clicks in music. *Perception & Psychophysics*, 1978, *24*, 171-174.

Harris, P. L., & Terwogt, M. M. A network account of synoptic processes in story recall. In M. M. Gruneberg, P. E. Morris & R. Sykes (Eds.), *Proceedings of the international conference on practical aspects of memory*. New York: Academic Press, 1979.

Hunter, I. M. L. Memory in everyday life. In P. Morris & M. Gruneberg (Eds.), *Applications of memory*. New York: Academic Press, 1979.

James,.I. M., Griffith, D.N.W., Pearson, R. M., & Newbury, P. Effect of Oxprenolol on stage-fright in musicians. *The Lancet*, 1977, *2*, 952-954.

Kintsch, W. Memory for prose. In C. N. Cofer (Ed.), *The structure of human memory*. San Francisco, California: Freeman, 1976.

Kolinski, M. "Barbara Allen": Tonal versus melodic structure, part II. *Ethnomusicology*, 1969, *13*, 1-73.

Laske, O. E. On psychomusicology. *International Review of the Aestbetics and Sociology of Music*, 1975, *6*, 269-281.

Liberman, M. Y., Olive, J. P., & Zukofsky, P. Studies of metric patterns. *Journal of the Acoustical Society of America*, 1977, *62*(supplement no. 1), 44.

Longuet-Higgins, H. C. Making sense of music. *Proceedings of The Royal Institution of Great Britain*, 1972, *45*, 87-105.

Longuet-Higgins, H. C. Perception of melodies. *Nature (London)*, 1976, *263*, 646-653.

Lord, A. B. *The singer of tales*. Cambridge, Massachusetts: Harvard University Press, 1960.

Martin, J. G. Rhythmic (heirarchical) versus serial structure in speech and other behavior. *Psychological Review*, 1972, *79*, 487-509.

McKeever, W. F., & Huling, M. D. Right hemisphere superiority in graphic production of briefly viewed dot figures. *Perceptual & Motor Skills*, 1970, *31*, 201-202.

Menuhin, Y. *Unfinished Journey*. London: Macdonald and Jane's, 1976.

Michon, J. A. Programs and "Programs" for sequential patterns in motor behaviour. *Brain Research*, 1974, *71*, 413-424.

Moog, H. *The musical experience of the pre-school child*. London: Schott, 1976.

Moore, G. *Am I too loud? The memoirs of a piano accompanist*. London: Hamish Hamilton, 1962.

Oscar-Berman, M., Blumstein, S., & DeLuca, D. Iconic recognition of musical symbols in the lateral visual fields. *Cortex*, 1976, *12*, 241-248.

Ostransky, L. *Understanding jazz*. Englewood Cliffs, New Jersey: Prentice-Hall, 1977.

Parry, M. *The making of homeric verse*. London and New York: Oxford University Press, 1971.

Patterson, B. Musical Dynamics. *Scientific American*, 1974, *233*, 78-95.

Perkins, D. N., & Howard, V. A. Toward a notation for rhythm perception. *Interface*, 1976, *5*, 69-86.

Razhnikov, V. G. The formation and the reproduction of conductor's intention. *Voprosy Psikhologii*, 1973, *19*, 55-66.

Rostron, A. B. Pitch control in the human voice. *Quarterly Journal of Experimental Psychology*, 1976, *28*, 305-310.

Rostron, A. B. What can you do with your larynx? A rejoinder to Clarkson and Deutsch. *Quarterly Journal of Experimental Psychology*, 1978, *30*, 171-174.

Salis, D. L. Laterality effects with visual perception of musical chords and dot patterns. *Perception & Psychophysics*, 1980, *28*, 284-294.

Schacter, S., & Singer, J. E. Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, 1962, *69*, 379-399.

Schoenberg, H. C. *The great pianists*. London: Gollancz, 1964.

Seashore, C. *The psychology of music*. New York: McGraw Hill, 1938.

Seeger, C. Versions and variants of the tunes of "Barbara Allen." *Selected Reports, UCLA Institute of Ethnomusicology*, 1966, *1*, 120-167.

Shaffer, L. H. Intention and Performance. *Psychological Review*, 1976, *83*, 375-393.

Shaffer, L. H. Performances of Chopin, Bach, and Bartok: Studies in motor programming. *Cognitive Psychology*, 1981, *13*, 326-376.

Sloboda, J. A. Music reading and prose reading: Some comparisons of underlying perceptual processes. Unpublished Ph.D. Thesis, University of London, 1974. (a)

Sloboda, J. A. The eye-hand span: An approach to the study of sight reading. *Psychology of Music*, 1974, *2*, 4-10. (b)

Sloboda, J. A. Visual perception of musical notation: registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology*, 1976, *28*, 1-16. (a)

Sloboda, J. A. The effect of item position on the likelihood of identification by inference in prose reading and music reading. *Canadian Journal of Psychology*, 1976, *30*, 228-236. (b)

Sloboda, J. A. Phrase units as determinants of visual processing in music reading. *British Journal of Psychology*, 1977, *68*, 117-124.

Sloboda, J. A. Perception of contour in music reading. *Perception*, 1978, 7, 323-331. (a)

Sloboda, J. A. The psychology of music reading. *Psychology of music*, 1978, *6*, 3-20. (b)

Steedman, M. J. The perception of musical rhythm and metre. *Perception*, 1977, *6*, 555-570.

Stewart, M. L. Structural development in the jazz improvisational technique of Clifford Brown. Unpublished Ph.D. Thesis, University of Michigan, 1973.

Sundberg, J. Articulatory interpretation of the "singing formant." *Journal of the Acoustical Society of America*, 1974, *55*, 838-844.

Sundberg, J., & Lindblom, B. Generative theories in language and music descriptions. *Cognition*, 1976, *4*, 99-122.

Tucker, W. H., Bates, R. H. T., Frykberg, S. D., Howarth, R. J., Kennedy, W. K., Lamb, M. R., & Vaughan, R. G. An interactive aid for musicians. *International Journal of Man-Machine Studies*, 1977, *9*, 635-651.

Turner, P., & Hedges, A. An investigation of the central effects of Oxprenolol. In D. M. Burley, J. H. Frier, R. K. Rondel, & S. H. Taylor (Eds.), *New perspectives in beta-blockade*. Horsham, England: Ciba Laboratories, 1973.

Van Nuys, K., & Weaver, H. E. Memory span and visual pauses in reading rhythms and melodies. *Psychological Monographs*, 1943, *55*, 33-50.

Weaver, H. E. A study of visual processes in reading differently constructed musical selections. *Psychological Monographs*, 1943, *55*, 1-30.

Weick, K. E., Gilfillian, D. P., & Keith, T. A. The effect of composer credibility on orchestra performance. *Sociometry*, 1973, *36*, 435-462.

Winograd, T. Linguistics and the computer analysis of tonal harmony. *Journal of Music Theory*, 1968, *112*, 3-49.

Wolf, T. A cognitive model of musical sight reading. *Journal of Psycholinguistic Research*, 1976, *5*, 143-171.

Yarbrough, C. Effect of magnitude of conductor behavior on students in mixed choruses. *Journal of Research in Music Education*, 1975, *23*, 134-146.

# 17

# Social Interaction and Musical Preference

*Vladimir J. Konečni*


I. Introduction ..... 497
II. Effects of Social Stimulation on Aesthetic Choice ..... 502
III. Effects of Information Load and Arousing Nonsocial Stimulation on Aesthetic Choice .... 505
IV. Effects of Listening to Melodies Differing in Complexity on Emotional States and Social Behavior ..... 507
V. Listeners' Sequencing and "Chunking" of Musical Materials and the Use of Music for Mood Optimization ..... 511
References ..... 516


## I. INTRODUCTION

Studies of aesthetic preference and of the process of appreciation have always represented a major aspect of experimental aesthetics and of the psychology of music and art (e.g., Child, 1969). Experimental work on, and theorizing about, preference for both authentic pieces of music and visual art, on one hand, and for stimulus patterns synthesized specifically for experimental purposes, on the other, have been going on at least since Fechner's (1876) *Vorschule der Ästhetik*, and have had a major boost in the 1960s and 1970s through Berlyne's "new experimental aesthetics" (Berlyne, 1971, 1974) and other influential developments (e.g., Arnheim, 1966; Farnsworth, 1969; Francès, 1958, 1968; Kreitler & Kreitler, 1972; Meyer, 1956). Numerous studies of musical preference have been carried out both early in the century (e.g., Gilliland & Moore, 1924; Moore, 1914) and more recently (e.g., Bragg & Crozier, 1974; Crozier, 1974; Heyduk, 1975; Konečni & Sargent-Pollock, 1976; Overmier, 1962, Simon & Wohlwill, 1968; Steck & Machotka, 1975; Vitz, 1966), and


Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

there have been several good reviews and theoretical integrations, which, though greatly varying in breadth, are all helpful for a better understanding of preference (e.g., Berlyne, 1971; Child, 1978; Madsen, 1973; Meyer, 1956; Walker, 1973).

The sheer size of the literature on music preference sometimes obscures the existence of a serious imbalance in terms of the amount of experimental and theoretical attention that has been devoted to various topics that can be reasonably regarded as falling within this area. For example, a great deal of effort seems to have been invested in discovering the exact shape of the function relating preference (in terms of ratings of pleasingness and interestingness, listening time, and various choice measures) to the complexity of authentic and synthesized musical materials (e.g., Berlyne, 1971; Steck & Machotka, 1975; Heyduk, 1975). By comparison, other issues have been almost completely ignored. For example, the vast majority of research studies and most of the theoretical attempts have treated aesthetic preference and choice as if they, and the process of appreication itself, normally occur in a social, emotional, and cognitive vacuum, as if they were independent of the contexts in which people enjoy aesthetic stimuli in daily life. In contrast, one of the main contentions of the present chapter is that a thorough understanding of aesthetic behavior cannot be achieved without examining how it changes as a function of its immediate social and nonsocial antecedents, concurrent cognitive activity, and resultant emotional states (cf. Cantor & Zillman, 1973). Before I examine these points in more detail, it seems useful to speculate about the reasons for the fact that social, emotional, and cognitive context factors have been so consistently ignored.

One possible reason is that psycho-aestheticians have underestimated or failed to grasp the significance of the extent to which music appreciation has been radically altered by the technological and social changes in the twentieth century. Consider first the conditions prevailing in the eighteenth and nineteenth centuries, when non-folk music was performed almost exclusively in the salons of the wealthy and later in the concert halls and opera houses accessible only to the privileged few. Enjoyment of music was a special occasion, something carefully planned in advance, each performance a unique, fleeting event. Contrast that with our times. There is wide availability of relatively low-cost, high-quality equipment for the recording and reproduction of music. There are excellent recordings of an astonishing range of musical pieces. In most large cities of the Western world, there are numerous radio stations specializing in every conceivable type of music and making it available on a 24-hour basis. Consider the amount of time in an average day that so many people spend listening to music—in their homes, offices, and automobiles (even while walking or roller skating!).

As a function of all of these factors taken collectively, a veritable revolution in music appreciation has occurred. Music of all types has become a major part of the lifestyle of a very large number of people.[1] Gone are the days when only the elite could hear high-quality music, while the rest had to await weddings and harvest

[1]Although perhaps not quite to the same extent, the advent of television has analogously revolutionized the appreciation of theatrical performances, opera, and ballet, and vastly increased the number of viewers.

festivities to hear any music at all (unless they produced it themselves or were content with bird songs). This situation has been replaced by the penetration of music into every corner of people's lives, literally and metaphorically. Entire generations of youngsters have been brought up on rock-'n'-roll. Social mini-movements and subcultures revolve around music and its association to other mood-optimizers, such as drugs and alcohol.[2] The most frequent, prototypical situations in which people listen to music have shifted from specialized locations, such as opera houses and concert halls, into the informal settings like the home and the automobile.

Perhaps related to this notable disregard for the technological and social changes affecting music appreciation are the prevailing elitist views of what constitutes "serious" music and, especially, which types of music are worthy of serious attention by psycho-aestheticians. Commonly ignored in both the experimental and theoretical work in the psychology of music are even the best examples of the jazz, rock-'n'-roll, and rhythm-and-blues idioms. The typically given reasons why these forms of music are not worthy of study are quite feeble when one considers the imprecision and arbitrariness of the existing formal definitions of what constitutes "serious music," the extent of the disagreement in value judgments among the critics, aestheticians and aesthetes themselves, and the extent to which all definitions and evaluations are culture- and period-bound. Above all, the exclusion of, say, rock-'n'-roll from the pool of music worthy of scientific attention blatantly ignores the obvious fact that it brings a great deal of pleasure to an immense number of people. These people's aesthetic preference and choice are presumably at least as important as those of the minority preferring "serious" music—if one is interested in developing a general theory of music appreciation. Moreover, major aspects of the aesthetic experience (in terms of both the factors affecting choice behavior and the hedonic effects of the music) may well be more similar for the appreciators of Palestrina, Stravinsky, and Led Zeppelin, respectively, than is generally believed.[3]

What seems needed is a broader perspective on music appreciation, one that transcends the narrow, elitist, and arbitrary definitions of what constitutes good and serious music and also takes into account the reality of music appreciation in our time. From this perspective, one of the most important by-products of the dramatic changes in the dominant modes and locations of music appreciation is the fact that music is nowadays so frequently enjoyed in a great variety of *social* contexts. More-or-less

[2]The following quotation from a film review by Carol Olten (the *San Diego Union*, 23 Oct. 1980) illustrates this point well, if casually:

> It is doubtful if anyone will truly define how or why rock 'n' roll or the entire area of pop music began to change the fabric of our lives. A groupie—probably a third-generation groupie—in *One-Trick Pony* . . . comes close to some explanation. "Rock 'n' roll kept me sane when I was a child. We were always moving around the country changing houses, changing schools. But on the road I could always turn on the radio and there would be the same rock 'n' roll song playing. It kept me sane, I tell you, sane." Softly, she begins to hum the first few lines of "Me and Bobby McGhee": "From the Kentucky coal mines to the California sun . . ."

[3]Introspections by a subject in an informal study ($N = 1$, myself, and an appreciator, incredibly, of both Bach and the Rolling Stones) have "confirmed" this.

Biblioteca Conservatorio Superior de Música - MADRID

active listening to music has become fully imbedded in the stream of daily life of ordinary men and women. People listen to music while working, talking, eating, engaging in sexual intercourse. That this fact has been ignored by psycho-aestheticians, and that they have continued to think of preference as a process largely unrelated to social situations, is quite remarkable. What music does to people at different times, why they choose to listen to it so much, why they choose a particular type of music while engaged in a particular activity—all of these are important and unanswered questions. They cannot be answered by "speculative aesthetics" (Berlyne, 1971), by philosophical treatments of the "social function" of music and art, or by the hitherto popular types of studies of aesthetic preference. One needs to analyze preference with at least some reference to the typical situations in which music is appreciated, in social contexts, in the presence of friends, lovers, and family members, in the stream of daily life.

Whereas an average person probably rarely goes to a concert of any kind and probably never attends an operatic performance, ordinary days are filled with countless aesthetic micro-episodes—numerous conscious decisions to listen to some type of music by turning on the radio, putting on a record, and so on, and then proceeding to listen. One of the main contentions of this chapter is that the nature of these aesthetic choices, which music people decide to listen to, and for how long are to a very high extent affected by the social context in which the listening occurs. The social activity the listeners are engaged in, whom they are with, and what emotions and moods they are experiencing as a function of the social stimulation emanating from others in their environment are very likely to affect the type of music that is chosen. Furthermore, effects on choice among musical pieces (including computer-generated "melodies" synthesized for research purposes) differing on various stimulus dimensions (e.g., complexity, novelty, surprisingness) can also be reasonably expected to be a function of the nature, difficulty, information-processing requirements, and arousingness of tasks and activities in which a person is concurrently engaged. In short, what I am suggesting is that musical preference and choice may be highly affected by the social, emotional, and cognitive factors leading to, and in, the listening situation.

This point logically leads to another issue largely ignored in the psychology of music. Despite a considerable amount of interesting theoretical work on the relationship between music appreciation and emotion (e.g., Langer, 1942; Meyer, 1956, 1957), there have been few experimental investigations of the very real possibility that not only the listeners' emotional states but also their social behavior and their treatment of others in numerous everyday social micro-episodes may be perceptibly and differentially affected by music. That both subtle and major changes in social behavior may occur as a function of listening to music is another contention of the present chapter.

The two basic propositions made so far are complementary and can be related to each other within the same theoretical framework as components of a prototypical "aesthetic episode." The model assumes that music, and aesthetic stimuli in general—specialized and highly valued as they are—are simply another aspect of a person's acoustic (or visual) environment and that they are chosen largely for the

purpose of mood- and emotion-optimization. The model regards a person as being engaged in a constant exchange with the social and nonsocial environment, of which the acoustic stimuli are a part. The social behavior of others—and I am referring to the ordinary, everyday behavioral sequences unambiguously interpreted as indicating social support, love, antagonism, and so on—is assumed to have a profound effect on a person's emotional states, which, in turn, affect aesthetic choices, including the choice of music, that a person will make in a given situation.[4] The degree of enjoyment of the chosen piece presumably varies as a function of the concurrent social and nonsocial micro-environmental conditions (which also may affect the probability of that particular piece being chosen in the future).[5] Listening to music is further assumed to produce changes in the listener's emotional state and thereby affect his or her behavior toward others in the situation. Since social behavior is by definition interactive, it is safe to assume that the behavior directed toward the listener by others will also change, leading to a further modification in the listener's emotional state, and possibly to different subsequent musical choices. The model thus contains a feedback-loop feature representing the ongoing nature of a person's interaction with the social and musical environment—a series of aesthetic episodes mediated by changes in emotional state and mood.

I do not mean to imply that the factors discussed above are likely to override completely the basic individual differences in the preference for different types of music produced by, say, early exposure (see footnote 4). However, these factors can perhaps account for a respectable amount of variance in the choice behavior within the broad preference domains (e.g., classical versus rock-'n'-roll) and also increase the understanding of the reasons for the occasional choice switches across these domains.

Also the aesthetic episodes and musical choices I have been discussing clearly may well be far removed on several dimensions from the profound effects of hearing for the first time a great piece of music, be it Mozart's *Don Giovanni* or the Who's *Tommy*. But how often do such experiences occur, and how many people have them? Music clearly plays an important role in human life, but one is not likely to understand this role and develop a reasonably general theory of music preference, choice, and enjoyment by consistently ignoring the music-related behaviors of the vast majority of listeners. Indeed, a good argument can be made that further progress in the psychology of music hinges on the success of investigations of *mundane* types of aesthetic appreciation and choice, those preceded and accompanied by routine social behaviors and common emotional states.

[4]I am not denying the importance of other factors that may affect aesthetic choice, including availability and the appropriateness of listening to a particular piece of music in a given situation and, of course, a person's basic liking ("musical taste") for different types of music that is presumably affected by a host of factors in the person's upbringing, exposure, cultural conditions, musical education, peer pressure, and so on.

[5]This type of context effect in which aspects of the social environment are assumed to affect the basic liking (preference for a piece of music has been discussed by others—e.g., Child, 1978) and is, of course, a part of the folklore surrounding music preference. Child's example was of a not particularly musically inclined couple who attended a concert while falling in love and found the music divine.

My collaborators and I have so far carried out a number of experiments relevant to the various stages of aesthetic episodes. We have used different experimental paradigms, induced a variety of emotional states in our subjects, and explored the subjects' preferences and choice behavior for pieces of music ranging from computer-generated melodies to authentic musical pieces from the eighteenth, nineteenth, and 20th century, as well as contemporary rock-'n'-roll. These experiments and the analogous studies using visual aesthetic stimuli (e.g., Konečni & Sargent-Pollock, 1977; Sargent-Pollock & Konečni, 1977) should be regarded as only the first tentative steps toward the elucidation of the various components of aesthetic episodes. For one thing, although they examined the relationship between music-related behaviors and some of the social, emotional, and cognitive factors discussed above, they have so far all been laboratory experiments. Putting some of the ideas presented so far to the test, however, should facilitate subsequent field investigations, both those paralleling the laboratory studies and others designed to collect normative data. Such field investigations could perhaps be profitably organized around the basic question of how people's everyday aesthetic choices are influenced by the setting (house, automobile), emotional state (elation, anxiety, depression), information load (conversation, manual work, reading, eating), physical state (fatigue, illness), the type of other people present (family members, other intimates, co-workers), even by atmospheric conditions.

The remainder of this chapter will consist of a brief review of some of our experiments. With a few exceptions, these studies have been previously published; so the main purpose here is to convey the flavor of the research and present some of the conclusions.

## II. EFFECTS OF SOCIAL STIMULATION ON AESTHETIC CHOICE

In one study (Konečni, Crozier, & Doob, 1976; also see Konečni, 1979, Section II,A), the subjects were, in the first part of the experiment, repeatedly insulted in the course of working on a task by an accomplice of the experimenter, posing as a subject. Pilot work had shown that this procedure significantly increases various psychophysiological indices of arousal and that, moreover, subjects invariably report themselves as experiencing considerable anger. In the second part of the experiment, seemingly quite unrelated to the first, the subjects, now alone in the room, chose on each of many trials to listen to 10 sec of one of two types of melodies. Both types of melodies were computer-generated; they differed considerably from each other in both objective complexity (or "uncertainty," measured in information-theory terms) and subjective complexity (measured by verbal ratings in pilot studies). The two particular complexity levels were chosen because the normally aroused subjects, experiencing no pronounced emotion, chose to listen to them about equally in pilot studies. Indeed, in the experiment proper, a control group of subjects who had not been insulted by the accomplice showed roughly equal preference for the two types of melodies over trials.

In sharp contrast, subjects who had been insulted chose the simpler of the two types of melodies about 70% of the time, shunning the more complex type. In the third experimental condition, insulted subjects were given an opportunity to retaliate behaviorally against the accomplice prior to the choice measure. In line with predictions from prior research (e.g., Hokanson & Shetler, 1961; Konečni, 1975a), which had shown that retaliation by angry people may reduce arousal level, these subjects' choice behavior was very similar to that of the noninsulted controls.

The pattern of results that was obtained had been predicted on the basis of a great deal of prior research relating arousal, anger, aggression, and collative variables, such as complexity. From the perspective of the present chapter, however, the findings are important insofar as they show that a socially induced change in a listener's emotional state may strongly affect that person's aesthetic choice (i.e., choice between musical stimuli varying in complexity—the first stage of the proposed prototypical aesthetic episode. Moreover, to the extent that the insulted subjects' preference for the simpler type of melody was reduced, brought down to "normal," by retaliation (presumably because this behavior had decreased the level of arousal and the degree of anger), the finding in that condition is also relevant for the feedback-loop aspect of the proposed model regarding the relationship between overt behavior and internal states. A person's socially induced emotional state affects his or her social behavior toward other people; in addition, the execution of actions directed at social targets affects the actors emotional state, which, in turn, regulates subsequent behavior, and so on.

Aesthetic factors fit in this feedback loop in two related ways. First, to the extent that choice between musical stimuli varying on certain dimensions (such as complexity) is affected by fluctuations in emotion and arousal, it follows that the performance or nonperformance of behaviors (such as aggression) that are seemingly totally unrelated to musical preference would turn out to be very relevant (because such behaviors seem to affect emotions and arousal). Second, insofar as listening to musical stimuli varying in complexity may itself differentially affect arousal and emotional state, it follows that music would play an important part in the behavior/internal state/behavior causal sequence.

Some of these hypotheses were further investigated in another experiment involving simple and complex computer-generated melodies, anger, and retaliatory aggression (Konečni, 1979, Section II,B). After being either insulted by an accomplice of the experimenter or neutrally treated by this person, the subjects participated in another task involving the same accomplice in the course of which, on each of many trials, they had the choice of either punishing or rewarding the accomplice by pressing the appropriate button. Unlike the situation in the experiment described earlier, the *aesthetic* consequences *for the subjects* were inextricably linked, on each trial, on their punishment versus reward choice. In the study as a whole, across all seven experimental conditions, there were three possible aesthetic consequences for the subjects: a simple melody, a complex melody, and silence. Different permutations of these three consequences were operative in different experimental conditions. For example, in one of the conditions, whenever a subject pressed the punishment button (supposedly physically hurting the accomplice), he or she heard—for as long as the button was

pressed—an example of the simple melody, whereas whenever the subject pressed the reward button, he or she heard—for as long as the button was pressed—an example of the complex melody. In another condition the pairing of the aesthetic consequences with the two behaviors in question was reversed. In still another condition pressing the punishment button yielded silence; this pairing was reversed in a further condition, and so on.

In making the predictions, we found several reasonable assumptions, based on other studies in my laboratory and elsewhere, helpful. One assumption was that in comparison to the noninsulted subjects, the insulted ones would be significantly more inclined to press the punishment button. Another assumption, more relevant to one of the components of the hypothesized aesthetic episode, was that angry subjects would prefer simple to complex melodies both in absolute terms and in comparison to the noninsulted subjects and that their preference for the no-melody (silence) consequence would fall in between.

When these assumptions are put together with other ideas presented so far, quite precise predictions can be made concerning the subjects' aesthetic-choice behavior (i.e., punishment versus reward behavior, since the aesthetic consequences for oneself and behavioral consequences for the social target, the accomplice, where inextricably linked in this experiment). The clearest preference for simple melodies (i.e., the least preference for complex melodies) should be displayed by the insulted subjects for whom listening to simple melodies is linked with the punishment of the accomplice, whereas listening to complex melodies is linked with the reward of the accomplice. At the other extreme should be the insulted subject for whom exposure to simple melodies is linked with the reward of the accomplice, and listening to complex melodies with the punishment of the accomplice. In the former condition, the two motivations (to listen to simple melodies and to punish when angry—an internal and an external consequence of the same behavior) were experimentally made maximally congruent whereas in the latter condition they were maximally divorced. The exact pattern of data for the experiment as a whole would, of course, depend on the relative strengths of the two motives (i.e., their relative strengths as manipulated in this particular experiment) and on the degree of preference for the "silence" consequence, relative to the other two consequences, for the insulted and neutrally treated subjects, respectively.

On the whole the findings closely followed the predictions and thus gave a considerable amount of support to the underlying assumptions that had been based on the previous studies and, more importantly, on the feedback aspect of the aesthetic-episode formulation. The choice between melodies differing in complexity does indeed seem to be affected both by the socially induced emotional states and by the feedback effects on such states from overt actions directed at social targets.

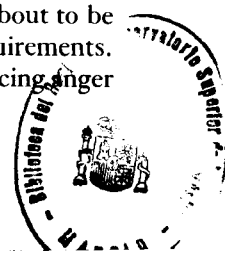
Another way of looking at the results would lead one to stress, perhaps with some surprise, the extent to which the performance of a socially (and presumably biologically) very important behavior—retaliatory aggression—can be modulated by the seemingly subtle differences in the musical consequences for the listener (i.e., aggressor). For example, whereas the insulted subjects who heard a simple melody whenever they

punished the accomplice (and a complex melody whenever they rewarded that person) chose the punishment button on 26.6 of the 50 trials on the average, the mean for the equally insulted subjects who heard a complex melody whenever they pressed the punishment button (and a simple melody whenever they rewarded the accomplice) was only 18.8—and this was a statistically highly significant difference. The difference between these two groups in terms of the mean duration of punishments to the accomplice was no less than about 3.8:1.

## III. EFFECTS OF INFORMATION LOAD AND AROUSING NONSOCIAL STIMULATION ON AESTHETIC CHOICE

Just as the social aspects of music appreciation have been largely neglected in the psychology of music so far, so has the "cognitive context" in which choice and listening occur. Yet, it would seem clear that in many situations involving listening to music, not all of a listener's attention or processing capacity (Broadbent, 1958; Moray, 1969; Treisman, 1964) is devoted solely to the appreciation of the chosen piece of music. As often as not the person may be concurrently engaged in other intellectual activities or motor tasks. In addition to being intrinsically interesting, the issue of how various intellectual and motor activities affect aesthetic choice seems worthy of investigation for another important reason.

The reason is that there is a possibility that the effect of the socially induced emotional states on aesthetic choice (observed in the experiments discussed so far) is mediated by fluctuations in arousal level. If so, a further question arises. How do gross changes in the arousal system affect the subtle operations presumably involved in aesthetic choice? The missing link may be found through an extension of the views that high levels of arousal impair cue selection and lead to a "narrowing" of attention (e.g., Broadbent, 1971; Easterbrook, 1959). After all, in the experiments described so far as well as in real-life choice situations involving different pieces of music, people do not simply choose one piece over another but actually then proceed to listen to the chosen piece for a time. It is reasonable to assume that listening to a piece of music requires cognitive work; it requires that the components of which the piece of music consists be analyzed and processed, and that their meaning be extracted (e.g., Berlyne, 1971; Günzenhauser, 1962; Moles, 1958; Morris, 1957; Perkins and Leondar, 1977). In the experiments discussed earlier, subjects were well aware before each choice of the information-processing demands of the two types of melodies. In other words, they had a fairly good idea of how much their processing capacity would be taxed in the 10 sec following a particular choice. Similarly, in daily life, when people choose between various pieces of music or between radio stations, they are aware from prior experience of the general characteristics of the piece of music about to be listened to, including its complexity and information-processing requirements. Therefore, it could be hypothesized that the experimental subjects experiencing anger



shun complex melodies because these melodies' information content exceeds the subjects' currently available processing capacity (which had been reduced by the elevation of arousal characteristic of anger and other pronounced emotional states).

It therefore seemed of interest to us to examine within the same experiment the effects on aesthetic choice of both high arousal per se (induced by an information-free loud tone) and of cognitive tasks carried out simultaneously with the listening to the chosen melody on each trial where the tasks varied in (a) the processing effort demanded and (b) arousingness (Konečni & Sargent-Pollock, 1976).

Prior to each block of choice trials (the previously described simple and complex melodies were the choice alternatives), some subjects were exposed to bursts of a very loud squarewave tone, which considerably raised their arousal level (measured by the common psychophysiological indices), especially in the first half of the choice trials, whereas other subjects heard bursts of the same tone at a much lower loudness level.

The second experimental variable was the type of task on which the subjects worked while listening to the chosen melody on each trial. During each block of choice trials (between two exposures to the tone), one group of subjects worked continuously on the "digit-symbol" task (which requires the continuous processing of information with little fluctuation in processing difficulty over time), pausing only to press one of the two choice buttons every 10 sec and thus hear either a simple or a complex melody on that trial. Because the digit-symbol task involves not only the processing of information but also rapid writing—which by itself may raise arousal level—a number of additional conditions were included. In one of these ("contour-tracing"), subjects were asked to trace, as fast as possible, the contours of the symbols that had already been written in on the worksheets given to them. Thus, these subjects engaged in at least as much rapid writing as the digit-symbol group but were required to process far less information. In another condition, subjects were asked to watch slides of paintings closely during the choice trials and to memorize as many details as possible for a supposed subsequent test. The active storage of information involved in this task presumably required a great deal of processing effort but contained few elements of "conceptual conflict" (Berlyne, 1960), which the digit-symbol task—a series of mini-problems—may contain and which may itself raise arousal level (Blatt, 1961). In a further condition, subjects saw the same slides but without the memorization instructions, meaning that they would probably process less information during the choice trials. In the final control condition, subjects did not work on any task while choosing and listening to the melodies.

The results of the experiment were quite clear. The subjects' choice behavior was strongly affected by the task in which they were engaged while choosing and processing the melodies. Being required to work on the digit-symbol and slides-memorization tasks, both of which involved a considerable amount of information processing (and thus presumably decreased the subjects' processing capacity), sharply reduced the choice of complex melodies. The contrast of these two conditions with the remaining three accounted for virtually all of the variance due to the type of task. More processing capacity was presumably allocated to these congitively more demanding tasks (cf. Kahneman, 1973); since complex stimuli are more difficult to process than simple

stimuli and given that the option existed, subjects chose to listen to the less demanding additional stimulation, simple melodies.

If we turn to the arousal aspects of the results, several observations are worthy of mention. First of all, being exposed to a loud tone prior to a block of choice trials significantly increased the level of arousal and decreased the percentage of choices of complex melodies in comparison to the control condition. However, the arousal aspects of the cognitive tasks themselves were quite irrelevant. Even though the slides-memorization subjects were physiologically far less aroused than were the rapidly writing digit-symbol subjects, they chose as few complex melodies. Furthermore, the digit-symbol subjects were no more aroused physiologically than were the contour-tracing ones, but they chose significantly fewer complex melodies. These and other aspects of the findings clearly showed that the processing-capacity factors, unmediated by the arousal-level fluctuations, had a direct and powerful effect on the choice between melodies. This finding does not prove that the effect on choice of high arousal (due to a loud tone or anger) is itself mediated by a decrease in processing capacity, but it does make such an explanation tenable.

Experiments discussed so far show that the socially induced emotional states, nonsocially-induced high arousal (due to a loud tone), and processing requirements of the concurrently performed cognitive tasks, all affect choice between melodies differing in complexity. In addition, they suggest a plausible mediational chain, one which runs from the actions of other people directed at the listener, via the listener's interpretation of these actions, his or her labelling of the emotional state, and arousal fluctuations, to the change in processing capacity and its effect on choice behavior.[6]

## IV. EFFECTS OF LISTENING TO MELODIES DIFFERING IN COMPLEXITY ON EMOTIONAL STATES AND SOCIAL BEHAVIOR

In the introductory section I suggested that any treatment of aesthetic appreciation that aspires to some degree of completeness must deal not only with the antecedents of

[6]Perhaps we are now a step closer to understanding why a person enraged by someone, thinking hard about the details of the encounter and simultaneously maneuvering his automobile through noisy, rush-hour traffic, seems likely to tune his car radio away from his favorite classical station playing, say, Stravinsky's *Rite of Spring* (though it seems already a relatively tame, traditional piece) and tune in, volume reduced, to the local "mellow" station or switch off the radio altogether.

It is also of considerable interest to explore the effects on musical preference of variables that fall, on the social–nonsocial dimension, somewhere between impersonal bursts of an aversive squarewave tone and verbal insults from a stranger in a face-to-face situation. In an experiment designed by Sigrid Flath-Becker and presently being carried out by her in my laboratory, subjects are differentially aroused by being repeatedly and sternly told by the experimenter to be faster while working on demanding cognitive tasks and subsequently informed that they did either very poorly or very well in different experimental conditions. (There are several additional control conditions.) Flath-Becker is investigating the effect of such manipulations on subjects' preference for different rhythmic structures (regular, ostinato, syncopated, complex) in piano (Bach, Debussy, Bartok, Schönberg, respectively), orchestral (Bach, Ravel, Bartok, Schönberg), and percussion (Baker, Fink, Fink, Cage) compositions.

preference and choice but also with the consequences of the choice and exposure to aesthetic materials. Yet, this issue has been almost completely ignored in the psychology of art [with the honorable exception of a vintage study by Gilliland and Moore (1924)]. The prevalent, if unstated, attitude has been that the domain of interest does not extend beyond the period of exposure (unless one takes seriously the metaphysical speculations on the enrichment of the soul and spirit and the taming of savage beasts by music). However, to the extent that music may affect moods and emotions (a point that everyone readily agrees with) and given the social context in which music is so often appreciated, it would be very surprising if people's exposure to aesthetic stimuli, including music, would not have quite considerable effects on their social behavior. Paradoxically, this would also suggest that perhaps some of the most important effects of music (in terms of the sheer frequency of their occurrence in the life of average people) may be quite transient in that they are primarily mediated by emotional changes that ordinarily dissipate quite quickly.

Unfortunately, it is extremely difficult to study the ways in which the quality of people's interactions with their intimates is affected by the type of music they listen to in their homes and automobiles, at work, or on a picnic. Perhaps this will be possible in the future. So far, however, I have been forced to remain in the laboratory, but the inclusion of many of the features of the experiments to be described in this section has been inspired by the broader picture of music appreciation in social contexts.

Does listening to melodies varying in complexity differentially affect the amount of subsequent aggressive behavior? If so, does the effect of the melodies combine with that of anger? Also, does the loudness level at which the melodies are presented play a part in determining the amount of aggression that ensures? These were among the questions dealt with in an experiment by Konečni (1975b; also see Konečni, 1979, Section III,A).

In the first part of the experiment, some of the subjects were treated in an insulting, rude way by the accomplice of the experimenter whereas others were treated neutrally. All subjects then had an opportunity to administer supposedly painful "electric shocks" to the accomplice on each of many trials. With the exception of a control group that heard no melodies at any point in the experiment, the subjects listened to a 10-sec computer-generated melody on each trial while making the decision whether or not to shock the accomplice. Some subjects listened to a simple melody played at a comfortable listening level on each of the trials whereas others also heard simple melodies on all trials except that they were played at a very high listening level. Still other groups of subjects heard a complex melody on each trail, either at a comfortable or very high listening level, depending on the experimental condition. This experimental design allowed a detailed comparison of the effects on aggressive behavior of three quite different experimental treatments each of which, however, independently produced differences in arousal level. Pilot work had shown that listening to complex melodies in comparison to simple ones may raise the level of arousal; however, listening to these complex melodies is not, by any criterion, aversive for the normally aroused subjects. On the other hand, listening to melodies at a very high listening level (close to 100 dB-A) in comparison to listening to these melodies at a comfortable

listening level is both arousing and aversive. Finally, prior work had shown that being rudely insulted by an accomplice in a face-to-face situation is arousing, aversive, and leads subjects unequivocally to label themselves angry.

On the basis of such information, rather precise predictions could be made. It was hypothesized that the melodies themselves—whether complex or simple and whether heard at a high or low listening level—would have little, if any, effect on aggressive behavior. The subjects would be differentially aroused and find the experimental situation differentially aversive, but this would presumably not be quite enough to produce differences in a highly important, high-consequence social behavior, such as aggression. In contrast, it was predicted that the aggressive behavior of subjects who had been made angry by the accomplice's insults would be further augmented by the melodies they were forced to listen to while making the choice of whether or not to punish the accomplice on each trial. The more arousing and/or aversive the melodies, the greater the increment over the baseline established by anger alone. The highest amount of aggression was thus expected to occur in the condition in which angry subjects listened to complex melodies at the high listening level on each trial.

The pattern of results that was obtained largely corroborated these predictions with some interesting exceptions. Subjects who had not been insulted generally displayed a low level of aggressive behavior that varied little as a function of the type of melody and loudness level. The one exception to this was the noninsulted subjects who had been repeatedly exposed to complex melodies at the high listening level: these subjects' aggressive behavior was only somewhat less pronounced, and statistically not different, from the amount of aggression displayed by the insulted controls who had heard no melodies whatsoever. It would thus seem that there is at least some tendency for the arousing complex melodies when they are played at a very high listening level to incline the subjects toward aggression. One could well imagine a situation in which people aroused by loud complex music would tend to overreact to relatively slight additional provocations that would ordinarily be brushed aside.

The pattern of results for the insulted subjects was quite different. As predicted, angry people who had been exposed to loud complex melodies displayed the greatest amount of aggression in the experiment, significantly more than the angry subjects who had heard no melodies. The insulted subjects who had heard complex melodies at a low listening level and those who had heard simple melodies at a high listening level displayed an intermediate amount of aggression. The one result that did not fit well in this general pattern but was not altogether unanticipated [on the basis of a study using visual aesthetic stimuli by Konečni and Sargent-Pollock (1977)] was the amount of aggression displayed by the insulted subjects who had repeatedly heard simple melodies at a low listening level. These people's aggressive behavior was significantly lower than that of the insulted subjects who had heard no melodies at all; it was, in fact, lower (though not significantly so) even than the aggressive behavior of some noninsulted subjects (those who had heard complex melodies at a high listening level). It thus seems that simple, soft melodies may have been actively soothing—perhaps by virtue of reducing the level of arousal and the degree of anger faster than was the case in the condition involving insulted subjects who had heard no melodies.

The experiment was thus a tentative demonstration that the exposure to melodies varying in complexity could have a differential effect on an important (anti)social behavior—both directly, insofar as it may combine with the effects of anger, and indirectly, in that it may raise arousal level and create a disposition to aggress. It was also shown that the loudness at which the melodies are heard is important, but it remains to be seen whether loudness would have a similar effect with authentic musical pieces and nonmusical acoustic patterns.

The apparently soothing effect of simple melodies presented at a comfortable listening level was tested by a different procedure in another experiment (Konečni, 1979, Section III,B). In the key condition of this complex experiment, subjects were rudely treated by the experimenter as soon as they walked into the laboratory, blamed for "being late" in a hostile and arrogant manner, and told that being late for important appointments implied irreponsibility and immaturity.[7] Next, subjects listened to simple melodies for three minutes continuously. The hypothesized effect of listening to simple melodies at this point was that the melodies would accelerate the decrease in the subjects' anger over and above the rate of recovery that would have been produced by homeostatic processes acting alone. Following the listening period, subjects rested for 15 minutes, after which they were insulted again by a different person, an accomplice of the experimenter using a different angering procedure.

In the final part of the experiment, the dependent measures were collected. These were the number and duration of simple-melody choices to which the subjects decided to listen, and the number and duration of punishing blasts of noise supposedly delivered to the accomplice in the context of a task. Subjects were told that every time a light of a particular color came on, they could press a button, which would expose them to a simple melody, and that the melody would go on for as long as the button was pressed. They were also told that when a light of a different color came on, they could press another button for as long as they wished, which would deliver a blast of noise to the accomplice (a plausible pretext was devised and presented to subjects).

There were several sets of control conditions that can be characterized by the following: (a) no listening to melodies following the first anger induction; (b) no initial anger induction; and (c) no second induction of anger. (There were additional conditions that are of little interest from the perspective of the present chapter.)

The major experimental hypothesis was that the subjects whose recovery from a high degree of anger was accelerated in the first part of the experiment by listening to the soothing melodies—presumably a relatively favorable state of affairs in that people can be reasonably expected to prefer a fast rate of diminution of a pronounced negative emotional state—would be particularly prone to listen to simple melodies following the second anger induction. In comparison to various control groups, the simple melodies had, for these subjects, proved their usefulness in alleviating a negative emotional state. In contrast, the same melodies had accomplished no such outcome for

[7]The reader may assume that at the end of all experimental sessions involving the insult procedure, subjects are thoroughly debriefed about the purpose of the experiment, the nature of the procedures, and reassured about their performance.

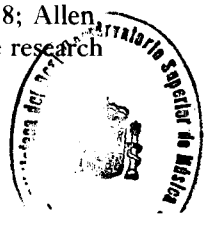
subjects who had listened to them in the first part of the experiment without being angry. Similarly, subjects who had been made angry in the first part of the experiment, but not given the opportunity to listen to the simple melodies at that time, could not have been as sure, following the second anger induction, of the melodies' positive impact on anger. Finally, subjects who were not angered the second time, just before the collection of the dependent measure, presumably did not experience at that time a negative emotional state that needed to be diminished.

The actual results confirmed these hypotheses, in terms of both the frequency and duration of listening to simple melodies and the frequency and duration of blasts of noise to the accomplice (the choice of melodies and the decision to deliver a blast of noise to the accomplice were independent, nonredundant measures). This experimental outcome and the data from an analogous experiment involving the emotion of fear (Konečni, 1979, Section III,C) suggest that listening to simple melodies, paired with negative emotional states, can have a powerful effect on the subjects' behavior at the time of the second anger induction. Subjects who had experienced that simple melodies can successfully decrease their anger (faster than simply the passage of time) resorted to listening to such melodies a great deal when angered again. These subjects were less prone to engage in aggression at that time.

In summary, the experiments presented in this section, taken together, seem to suggest—even if with regard to a very restricted range of musical materials and social behaviors—that listening to melodies varying on certain fundamental stimulus dimensions may differentially affect subsequent behavior of the listeners toward other people and that they may learn to seek actively the melodies with certain properties to alleviate negative emotional states. This question, the use of musical materials for the purpose of mood optimization, is further pursued in the next section.

## V. LISTENERS' SEQUENCING AND "CHUNKING" OF MUSICAL MATERIALS AND THE USE OF MUSIC FOR MOOD OPTIMIZATION

Given a range of musical materials to which they are asked to listen, how do people plan their "musical environment"—that is, how do they program their exposure to the materials in terms of the listening sequence and the duration of the individual "chunks"? If there is some truth to the idea that music is consciously used for the purpose of mood optimization or an emotional boost, would this extend to the active use of certain pieces of music (characterized by a given level of pleasingness, soothingness, and so on) in order to offset the effects of an aversive event (e.g., exposure to a very loud tone)? In other words, what type of music do people choose to listen to immediately after exposure to aversive auditory stimulation? These and related questions were systematically investigated in a series of previously unpublished experiments conducted in my laboratory (Allen, Hammerbeck, & Konečni, 1978; Allen, Breckler, & Konečni, 1980), in which an entirely different aesthetic-choice research paradigm was used.



In the main experiment subjects were first acquainted with the range of musical materials by hearing for 30 sec an example of each of the following: (1) eighteenth century "serious" music; (2) twentieth century "serious" music; (3) "mellow" rock'-n'-roll; (4) "hard" rock'-n'-roll; (5) computer-generated simple melodies (similar to those in the experiments previously described); (6) computer-generated complex melodies; and (7) intermittent bursts of a highly aversive 350-Hz squarewave tone at 95 dB-A. The musical selections that the subjects heard had undergone a considerable amount of pretesting. From a much larger pool of musical pieces, three pieces were chosen to represent each of the four types of music, such that within each type the pieces were homogenous with regard to the ratings (by pilot subjects) of arousingness, soothingness, pleasingness, interestingness, and complexity. In addition, the eighteenth century selections (by Bach, Haydn, and Vivaldi) were rated as highly similar to the "mellow" rock selections (by Genesis, Mott the Hoople, and Linda Ronstadt) on the arousingness, soothingness, and complexity scales. Also, these two types of music had been rated as far less arousing and complex and more soothing than the twentieth century "serious" pieces (by Bartok, Prokofiev, and Skriabin) and the "hard" rock selections (by Montrose, Outlaws, Scorpions) while the latter two types had been rated as equal to each other on these three dimensions. However, on the dimensions of pleasingness and interestingness, the four types of music (or, rather, the particular selections that were decided upon on the basis of pretesting) had been rated as fairly comparable to each other, although there was a certain preference, on the pleasingness scale, for "mellow" rock, followed by the eighteenth century "serious" music—a very stable finding that held across a wide variety of selections and could not be eliminated, at least for the population studied (UCSD students).

In the pretest, subjects also rated the computer-generated melodies and the loud squarewave tone. The simple and complex melodies were rated both significantly less pleasing and less interesting than any of the authentic musical pieces but significantly more pleasing and interesting (by about 4 points on a 10-point scale) than the squarewave tone. The mean ratings of the aversive squarewave tone on both the arousingness and soothingness scales were at the very extremes of the scales, at the expected scale ends.

In the main experiment, after they heard a 30-sec example of each of the seven types of auditory stimuli, subjects were seated in front of a console on which seven keys were arranged in a circle and told that they would have to listen to two minutes of each of the seven types of stimuli. Every 15 sec subjects were to choose one of the seven (and keep track of the choices on a tally-sheet provided) until all seven types had been heard for two minutes. The order in which subjects were to hear the different types of stimuli was completely up to them. At one extreme, subjects could choose the same type eight times in a row, thus exhausting that type, and then move on to the next type, and so on. At the other extreme, subjects could switch from type to type every 15 sec in a more-or-less random fashion.

Subjects heard the stimuli over headphones with musical selections and computer-generated melodies at 73 dB-A and the squarewave tone at 95 dB-A. The tracks with each of the four types of music were prepared in such a way that four minutes of each

of three selections within a type were recorded in a random succession for a total of 36 minutes per track.

There were also a number of variations of this basic experimental condition. In one of these, subjects were asked to write a hypothetical program of 15-sec choices without actually listening to any of the stimuli after making the choices. In addition, there were experimental conditions involving only six (no squarewave tone), five (no simple or complex computer-generated melodies), and four (no squarewave tone or simple and complex melodies) types of stimuli. The basic seven-types condition and these additional three with fewer types of stimuli can be thought of as a $2 \times 2$ design with the squarewave tone (present or absent) and simple and complex melodies (present or absent) as the two factors.

The results were interesting and will be briefly summarized here. In the five-types and seven-types conditions, both of which involved the aversive squarewave tone, there was a very strong tendency for subjects to choose the loud squarewave tone very early in the session. The aversive tone was heard in relatively short "runs" (in terms of sequential 15-sec choices), interspersed with runs of both simple and complex computer-generated melodies (when these, as well as the tone, were available in the seven-types condition). These chunks of exposure to various stimuli of relatively low pleasingness early in the session were followed by longer runs of the type of music that the subject liked *second* best, which for the majority of subjects in the experiment proper, as in the pretest, was the soothing, nonarousing eighteeth century "serious" music. (Since the ratings on the five dimensions described above were available from subjects in the main experiment, it was possible to do the various analyses on the basis of each subject's individual order of preference for the various types of music, as opposed to the group means from the pretest.)

Thus, in what seems to have been a disconfirmation of what one would reasonably predict from most versions of the general learning theory, subjects chose to get the aversive stimulation over with early in the session, rather than delay the aversive experience as long as possible. They "chose to suffer" in small "doses" and apparently made the whole experience more palatable by listening immediately afterward to the type of music that pleased them considerably (though not their most favorite type). These results gave further support to the notion that people actively seek different types of music at different times in order to optimize their mood or, as in the present case, to offset the impact of aversive stimulation.

In the condition with six available types of auditory stimulation (no aversive squarewave tone), an analogous pattern of results was obtained. Subjects tended first to listen to a great deal of both types of computer-generated melodies (thus, again getting the exposure to the least liked stimulation over with early in the session), which was interspersed with chunks of the second best-liked music. However, for this type of choice behavior and listening strategy to occur, a certain threshold of aversiveness apparently has to be exceeded; namely, in the four-types condition, in which the squarewave tone and both type of computer-generated melodies were absent, subjects tended to begin the session by a considerable amount of listening to the second best-liked music, rather than by listening to the less pleasing types. (In the

pretest the least liked type of authentic music, the twentieth century "serious" pieces, still received a relatively high mean rating of 6.2 on a 10-point pleasingness scale in comparison to 3.4 for complex computer-generated melodies and 7.1 for the most pleasing music, "mellow" rock.)

In all the conditions, whether there were four, five, six, or seven types of stimulation available, the majority of the subjects reserved the end of the session for long runs of listening to the most pleasing music, which was—both in the pretest and in the experiment proper—"mellow" rock. With the proverbial liver and spinach eaten—in small bites—subjects helped themselves to mouthfuls of chocolate mousse at the end. (An irresistible, though imprecise, analogy in that the spinach/mousse sequence is influenced by factors other than preference.) This particular mood-optimization strategy was convenient for us as experimenters concerned with the subjects' welfare: it insured that subjects left the experiment happy.

The middle part of the session tended to be filled with listening to the two least preferred types of music, "hard" rock and twentieth century "serious" pieces. Even these types of music, however, were listend to in longer runs than were the squarewave tone and computer-generated melodies.

Perhaps the most parsimonious explanation of the subjects' overall listening strategy links mood-optimizing to subjects' efforts to keep arousal level within a certain range and avoid getting overaroused, while experiencing a controlled variety of arousal fluctuations through a sequence composed of: arousing squarewave tone; soothing eighteenth century "serious" music; arousing "hard" rock and twentieth century "serious" pieces; and, finally, soothing "mellow" rock—an alternating arousing/soothing sequence vaguely reminiscent of the order of fast and slow movements in certain musical compositions (cf. Berlyne, 1971, for a discussion of arousal-related musical "devices" used by composers).

It is somewhat surprising that subjects almost never chose to listen to a full two minutes of the same type of music, even in the 4-types condition. Since the selections within each type were all recorded in four-minute segments, there would have been a good chance for a subject to hear a pleasant piece of music uninterrupted by either other types of stimuli or to change from one to another selection within a type. It is possible that the nature of the experimental procedure favored relatively frequent alternations between types. Also, the fact that a choice had to be made every 15 sec, even if of the same type of music repeatedly, may have disturbed the subjects' perception of continuity of a piece. Interestingly enough, if the latter reason is important, the subjects were apparently unable to predict that this is how they would feel; namely, in the previously mentioned experimental variation in which subjects prepared a hypothetical listening program without actually listening to any auditory stimulation, there were far longer runs of the music of the same type.

The experiments described in this section seem to provide additional information about the use of music for the purpose of mood optimization. Moreover, the major finding has now been replicated with visual nonartistic stimuli. In the course of the experimental session, subjects had to view five different sets of slides that vastly differed in rated pleasingness. Although the subjects had to view all slides within all

sets for a fixed length of time, the order in which they viewed the slides was completely up to them. Subjects generally viewed the least pleasing slides (those belonging to the set of gory slides of accident victims) early in the session, interspersed with the second most-liked slides, those from the set of humorous cartoons. As was the case in the experiment involving musical stimuli, these subjects—who were male—left the most pleasing slides, of *Playboy* nudes, for the end. Again, it seemed to us that subjects left the laboratory quite content.

In summary, it seems that both the idea that music is used for the purpose of mood optimization and a theoretical framework with feedback features based on the concept of an aesthetic episode could be of heuristic value. The experimental steps taken so far have been tentative and relatively narrow in scope, but they can perhaps help develop a psychology of music that recognizes the role of social, emotional, and cognitive factors in music appreciation as well as the conditions under which music of all kinds is enjoyed in daily life.

## REFERENCES


Allen, R. B., Hammerbeck, J. A., & Konečni, V. J. *Planning an auditory aesthetic environment: Listeners' sequencing of musical and non-musical stimuli.* Paper presented at the 58th Annual Convention of the Western Psychological Association, San Francisco, California, 1978.

Allen, R. B., Breckler, S., & Konečni, V. J. *Sequencing and chunking of musical and visual stimuli.* Unpublished manuscript, University of California, San Diego, California, 1980.

Arnheim, R. *Toward a psychology of art.* Berkeley, California: University of California Press, 1966.

Berlyne, D. E. *Conflict, arousal and curiousity.* New York: McGraw-Hill, 1960.

Berlyne, D. E. *Aesthetics and psychobiology.* New York: Appleton, 1971.

Berlyne, D. E. The new experimental aesthetics. In D. E. Berlyne (Ed.), *Studies in the new experimental aesthetics.* New York: Wiley, 1974.

Blatt, S. J. Patterns of cariac arousal during complex mental activity. *Journal of Abnormal and Social Psychology,* 1961, *63,* 272-282.

Bragg, B.W.E., & Crozier, J. B. The development with age of verbal and exploratory varying in uncertainty level. In D. E. Berlyne (Ed.), *Studies in the new experimental aesthetics.* New York: Hemisphere Publ., 1974.

Broadbent, D. E. *Perception and communication.* Oxford: Pergamon, 1958.

Broadbent, D. E. *Decision and stress.* New York: Academic Press, 1971.

Cantor, J. R., & Zillmann, D. The effect of affective state and emotional arousal on music appreciation. *Journal of General Psychology,* 1973, *89,* 97-108.

Child, I. L. Esthetics. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (Vol. 3) (2nd ed.). Reading Massachusetts: Addison-Wesley, 1969.

Child, I. L. Aesthetic theories. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 10). New York: Academic Press, 1978. Pp. 111-131.

Crozier, J. B. Verbal and exploratory responses to sound sequences varying in uncertainty level. In D. E. Berlyne (Ed.), *Studies in the new experimental aesthetics.* New York: Wiley, 1974.

Easterbrook, J. A. The effect of emotion on cue utilization and the organization of behavior. *Psychological Review,* 1959, 66, 183-201.

Farnsworth, P. R. *The social psychology of music* (2nd ed.). Ames, Iowa: Iowa State University Press, 1969.

Fechner, G. T. Vorschule der Ästhetic. Lipzig: Breitkopf & Härtel, 1876.

Francès, R. Le perception de la musique. Paris: Vrin, 1958.

Francès, R. *Psychologie de l'esthétique* Paris: Presses Universitaires de France, 1968.


Biblioteca del Real Conservatorio Superior de Música de Madrid

Gilliland, A. R., & Moore, H. T. The immediate and long-time effects of classical and popular phonograph selections. *Journal of Applied Psychology*, 1924, *8*, 309–323.

Gunzenhäuser, R. *Ästhetisches Mass and Ästhetische Information*. Quickborn bei Hamburg, West Germany: Schnelle, 1962.

Heyduk, R. G. Rated preference for musical compositions as it relates to complexity and exposure frequency. *Perception & Psychophysics*, 1975, *17*, 84–91.

Hokanson, J. E., & Shetler, S. The effect of overt aggression on psysiological arousal level. *Journal of Abnormal and Social Psychology*, 1961, *63*, 446–448.

Kahneman, D. *Attention and effort*. New York: Prentice-Hall, 1973.

Konečni, V. J. Annoyance, type and duration of postannoyance activity, and aggression: The "cathartic effect." *Journal of Experimental Psychology: General*, 1975, *104*, 76–102. (a)

Konečni, V. J. The mediation of aggressive behavior: Arousal level vs anger and cognitive labeling. *Journal of Personality and social Psychology*, 1975, *32*, 706–712. (b)

Konečni, V. J. Determinants of aesthetic preference and effects of exposure to aesthetic stimuli: Social, emotional and cognitive factors. In B. A. Maher (Ed.), *Progress in experimental personality research* (Vol. 9). New York: Academic Press, 1979. Pp. 149–197.

Konečni, V. J., & Sargent-Pollock, D. Choice between melodies differing in complexity under divided-attention conditions. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, *2*, 347–356.

Konečni, V. J., & Sargent-Pollock, D. Arousal, positive and negative affect, and preference for Renaissance and 20th century paintings. *Motivation & Emotion*, 1977, *1*, 75–93.

Konečni, V. J., Crozier, J. B., & Doob, A. N. Anger and expression of aggression: Effects on aesthetic preference. *Scientific Aesthetics/Sciences de l'Art*, 1976, *1*, 47–55.

Kreitler, H., & Kreitler, S. *Psychology of the arts*. Durham, North Carolina: Duke University Press, 1972.

Langer, S. *Philosophy in a new key*. Cambridge, Massachusetts: Harvard University Press, 1942.

Madsen, K. B. Patterns of preference and models of motivation. In D. E. Berlyne & K. B. Madsen (Eds.), *Pleasure, reward, preference*. New York: Academic Press, 1973.

Meyer, L. B. *Emotion and meaning in music*. Chicago, Illinois: University of Chicago Press, 1956.

Meyer, L. B. Meaning in music and information theory. *Journal of Aesthetics and Art Criticism*, 1957, *15*, 412–424.

Moles, A. *Thèorie de l'information et perception esthètique*. Paris: Flammarion, 1958.

Moore, H. T. The genetic aspects of consonance and dissonance. *Psychological Monographs*, 1914, *17* (2, whole No. 73).

Moray, N. *Attention*. London: Hutchinson Educational, 1969.

Morris, C. Significance, signification and painting. In R. Lepley (Ed.), *The language of value*. New York: Columbia University Press, 1957.

Overmier, J. B. Auditory pattern preference as a function of informational context. Unpublished M. A. thesis, Bowling Green State University, 1962.

Perkins, D., & Leondar, B. *The arts and cognition*. Baltimore, Maryland: Johns Hopkins University Press, 1977.

Sargent-Pollock, D., & Konečni, V. J. Evaluative and skin-conductance responses to Renaissance and 20th century paintings. *Behavior Research Methods and Instrumentation*, 1977, *9*, 291–296.

Simon, C. R., & Wohlwill, J. F. An experimental study of the role of expectation and variation in music. *Journal of Research in Music Education*, 1968, *16*, 227–238.

Steck, L., & Machotka, P. Preference for musical complexity: Effects of context. *Journal of Experimental Psychology: Human Perception & Performance*, 1975, *1*, 170–174.

Treisman, A. Selective attention in man. *British Medical Bulletin*, 1964, *20*, 12–16.

Vitz, P. C. Affect as a function of stimulus variation. *Journal of Experimental Psychology*, 1966, *71*, 74–79.

Walker, E. L. Psychological complexity and preference: A hedgehog theory of behavior. In D. E. Berlyne & K. B. Madsen (Eds.), *Pleasure, reward, preference*. New York: Academic Press, 1973.

# 18

# New Music and Psychology

*Robert Erickson*

I. Introduction ..... 517
II. Music Theory and Music ..... 519
III. Understanding Tonality ..... 520
IV. Music and Perceptual Streaming ..... 523
V. Fused Sounds in Music ..... 529
VI. Music Theory and Experimental Science ..... 534
References ..... 535

## I. INTRODUCTION

There has always been a problem of communication between disciplines. The languages often seem incommensurable and the goals different. When music is involved there is a double barrier, a large one between musical practice and music theory and another between music theory and those disciplines that touch music at some point or other—acoustics, psychoacoustics, cognitive psychology, and others. Nevertheless, developments in these fields over the past 20 years have led scientific investigators increasingly toward musical problems. Theoretically minded musicians are turning to experimental psychology for useful concepts and better underpinnings for their theoretical constructs.

The first wave of musical interest came with the electronic technology that followed World War II. The tape recorder, electronic generation of sound by means of oscillators, and its manipulation by filters, gates, and modulators—were quickly taken up by composers in Europe, Asia, and America. Problems of electronic music and music reproduction provided the initial common ground for new music developments in psychophysics, psychology, and other disciplines. Periodicals such as *Die Reihe* and *Gravesaner Blätter* provided a forum in which possible musical uses for an approach through information theory were explored by W. Meyer-Eppler and A.A.


Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-213560-1

altogether another thing to do it smoothly at moderate and fast tempos. Theoretical ideas—say Rameau's theories about the roots of triads, or those of Boethius about musical numbers, or Hindemith's groupings of more and less dissonant chords—are like knowing when to pass the thumb in that scale—not exactly irrelevant, but very distant from the composing process. The chasm between music theorizing and music making has not been narrowed at all since Plato. It has probably widened, largely because theory has increasingly been perceived as prescriptive. The purpose of a proper music theory should be to help us to think more clearly about musical problems and only that. An important arena is the common ground between the physical world of sound and our perception, cognitive processing and mental organizing of ordered sounds. Theory cannot tell us what to do, but it can help us explain to ourselves what we have done.

Theorizing about music has been going on for so long, 2000 or 3000 years, that it is no surprise that it is a very confusing body of knowledge. Theorists of any particular time have written on topics of local interest, usually promoting them to the level of universal validity. The history of music theory can be read as a tale of ingenious special pleading for this or that musical style. That must be changed. We need a music theory that is not style bound, that is valid for world music, not its European, Chinese, or Indonesian varieties. We need to know what things in music are a matter of culture and what things belong to our common humanity. The tools of modern cognitive psychology seem appropriate for this task.

## III. UNDERSTANDING TONALITY

Much of the music theory of the past 100 years has been built upon a theory of tonality, "loyalty to a tonic." Early in this century a controversy arose over a new kind of music composed by Schoenberg, Berg, and Webern, dubbed "atonal" by polemicists. The controversy has not subsided. As Salzer (1952) wrote:

> Today's musical crisis centers on the problem of tonality. Tonality, new tonality, atonality, polytonality, twelve-tone music, new-classicism, impressionism—all these terms may symbolize various and often conflicting currents, but they vitally concern the substance of musical language (p. 5). . . . Our period, however, is completely at odds about basic conceptions of musical utterance and coherence (p. 6).

Theoretical ideas about tonality are bound up with stylistic ideas about music and what music should be. Tonality has been understood to require chords and chord progression. The edifice of chord construction rested upon appeals to the harmonic series, the "chord of nature." Schoenberg was concerned to show that his music was "pantonal" rather than "atonal," a word he considered meaningless; and he used the harmonic series (Schoenberg, 1934) to sketch an evolution in musical perception from the simpler diatonicism found in the lower harmonics to the chromatic scale that could be extracted from higher harmonics. "It is certain that the more perceptible overtones sound more familiar to the ear than those it hears but faintly; these last

therefore remain strange to it. For that reason the chromatic scale is a somewhat more complicated tonal form than the major" (p. 171).

He also suggested that tonality is a product of art rather than a product of the nature of sound. (One must infer that the harmonic series is somehow only important for scale construction and for the construction of chords, unless he is changing his ground here) (p. 175):

> Now then, since tonality is not something which the composer unconsciously achieves, which exists without his contribution and grows of itself, which would be present even if the composer willed the opposite; since, in a word, tonality is neither a natural nor automatic consequence of tone combinations and therefore cannot claim to be the automatic result of the nature of sound and so an indispensible attribute of every piece of music, we shall probably have to define tonality as the art of combining tones in such successions and such harmonies or successions of harmonies, that the relation of all events to a fundamental tone is made possible.

Schoenberg is certainly right in saying that a tonality is constructed, but he *has* shifted his ground, because earlier in the paper, leading up to his discussion of overtones, he had been quite clear about relationships between tones: "How, after all, can two tones be joined one with another? My answer is that such a juxtaposition of tones, if a connection is to be brought about from which a piece of music may be the result, is only possible because a relation already exists between the tones themselves" (p. 169). He continues: "To elucidate the relationship between tones one must first recall that every tone is a compound sound, consisting of a fundamental tone (the strongest sounding one) and a series of overtones. We may now make the statement, and to a great extent test and prove it, that all musical phenomena can be referred to the overtone series so that all things appear to be application of the more simple and more complex relationships of this series" (pp. 169–170). The implication here is tbat atonal music, highly chromatic, can be referred to the upper reaches of the harmonic series. Intervallic relationships are therefore validated by the harmonic series; the creation of a tonality is a matter of art.

Working in Vienna at the same time, Heinrich Schenker used the harmonic series to serve a somewhat different theory. Schenker's ideas and his method of analysis by reductive graphing of music have gained wide support in the United States. Modern definitions of tonality tend to cluster around his notion of a triad unfolding in time. Here is a short description of his essential theory by Kalib (1975) (p. 3).

> Stated in most succinct, nut-shell form, Schenker views the musical work of art as the stage-upon-stage artistic unfolding of the triad of nature in time. The tone upon which this triad is based is viewed as reproducing, in the sense of procreating, other tones in the order in which they occur in the overtone series. Due to the limitations of the ear in perceiving more distant overtones, music, as a human art, avails itself of no more than the first five partials of the overtone series, or composites thereof.

Both Schoenberg and Schenker appear to make chords a requirement of tonality. Each makes statements about the relation between the perceptibility of harmonics and musical structures. Each believed that ancient music and medieval monophonic and polyphonic music (modal music) was somehow deficient and primitive compared to

the developed systems of eighteenth and nineteenth century—chiefly German—music. The basic idea of tonality, loyalty to a tonic, is dealt with in the following way: the root of the unfolding triad is for Schenker the tonic of the composition and the fundamental of the harmonic series. Schoenberg leaves more room for maneuver by deriving the triads on I, IV, and V from the harmonic series (p. 170):

> The origin of the main fundamental tones is explained by the fact that each one occurs as the third overtone of the one lying a fifth below it. So that C is the third overtone of F, just as G is the third overtone of C. In this manner G:C = C:F; And it is evident that C attracts the tones related to it through G, just as F and its related tones do with the complex of C.

For both theorists it appears that tonality rests ultimately upon an appeal to the harmonic series; and if we ask, "Why do we feel loyalty to a tonic?"—the ultimate answer would presumably be that it is the fundamental of the harmonic series, the "chord of nature."

This answer is unsatisfying from a world music standpoint. There are cultures whose music appears to have little relationship to the harmonic series. Worldwide there is a multiplicity of scales and harmonies hardly expressible in the small number ratios that might relate them to the harmonic series, even roughly equidistant scalar schemes, and harmonies that owe no allegiance to the triad, that entity so prominent in the thinking of Schenker and Schoenberg. Yet many of these musics exhibit the characteristic of loyalty to a tonic. Some analysts are finding tonal centers in the very music by Schoenberg, Berg, and Webern that was previously described as atonal. Modern discussions, such as that in the second edition of the Harvard *Dictionary of Music*, have tried to take account of this fact:

> TONALITY. Loyalty to a tonic, in the broadest sense of the word. One of the most striking phenomena of music is the fact that, throughout is evolution—in non-Western cultures, in Gregorian chant, and in harmonized music—practically every single piece gives preference to one tone (the tonic) making this the tonal center to which all other tones are related.

Travis (1959), working within the tradition of Schenker's formulations, has proposed a view that could encompass both non-Western musics and some aspects of atonal and other contemporary music: "Music is tonal when its motion unfolds through time a particular tone, interval, or chord. It is this tone, interval, or chord, called the tonic, which identifies the tonality" (p. 261).

A final definition in this long parade, couched in up-to-date music-theoretic language, is by Berry (1976). It specifically allows for tonal music of all times and places, and has the additional merit of leaving the relation between system and perception ready for empirical research (p. 27):

> Tonality may be broadly conceived as a *formal system in which pitch content is perceived as functionally related to a specific pitch-class or pitch-class-complex of resolution*, often preestablished and preconditioned, as a basis for structure at some understood level of perception. The foregoing definition of tonality is applicable not just to the "tonal period" in which the most familiar conventions of tonal function are practiced (roughly the eighteenth and the nineteenth centuries), but through earlier modality and more recent freer tonal applications as well. (The terms *pitch class* and *pitch-class-complex* are used to denote pitch independent of specific registral occurrence, or a complex of such pitches generically understood).

All of these definitions posit the idea of a pitch center of some sort as basic. Cognitive psychologists could use empirical methods to investigate the question: Is there centering? If it can be established that centering is one of our cognitive strategies then, in my view, music theorists would have a firmer base for definitions of tonality than those that require adherence to the limited and limiting harmonic series. There are easy starting points for experiments in monophonic music: beginning and ending pitches; long notes ornamented by other notes, especially upper and lower neighbors; repetition of significant pitches; upper and lower extrema, among others. Ethnomusicologists have accomplished an enormous amount of analysis of the melodies of various peoples by counting pitch occurrences, extracting weighted scales, and submitting large masses of material to statistical analyses. If cognitive psychologists worked from the other end, we would soon have a body of musical and psychological information about monophonic music that could provide a floor upon which we could build (and test) theories of monophonic, homophonic, polyphonic, and heterophonic music.

Work by Deutsch (1975a), who examined recognition of the pitch of a tone following a sequence of interpelated tones, has shown that repetition of the tone results in substantially improved memory. Such laboratory investigation provides converging evidence for the concept of centering.

A further concept of importance is the more ready processing of pitch combinations that are proximal. Van Noorden (1975) has suggested the idea of pitch motion detectors, hypothesizing that these operate on pitch combinations that are relatively closely related. Deutsch (1979), using her paradigm described above compared the pitches of two tones that were separated by six interpolated tones and found that error rates were lower when the interpolated sequences were composed of smaller melodic intervals, and argued that such sequences formed a more effective framework of pitch relationships to which the test tones could be anchored. Now melodies the world over appear to use a great many narrow intervals and fewer skips, and it is likely that this reflects a musical and perceptual universal. Is there a limit to the size of an interval that can be perceived as a step? Studies of centering would find that information useful. It might also prove fruitful to look into the finding (Bilsen, 1977; Fastl, 1971) that the pitch of bandpass noise is correlated with the center frequency when the bandwidth is small, roughly less than 1/5 of an octave, whereas for larger bandwidths pitches are correlated with the low and high cutoff frequencies. Upper and lower neighbors of a tone in tempered tuning will fill up a minor or a major third, 300 and 400 cents, respectively. A fifth of an octave is 240 cents, bigger than our common whole step of 200 cents but not quite our minor third. It is poised between being a step and being a skip.

## IV. MUSIC AND PERCEPTUAL STREAMING

The phenomenon of perceptual channeling is fundamental to an understanding of melody, and it goes a long way toward demonstrating perceptual phenomena that must underly melodic patterning and certain kinds of polyphony, both old and new.

The formation of perceptual streams or channels was first investigated by Miller and Heise (1950). They found that a rapid trill broke into two separate streams when the frequency distance exceeded about 1/7 of an octave in musical ranges. They named the region of transition the *trill threshold.* Warren, Obusek, Farmer, and Warren (1969) found that in listening to a tape loop of four sounds (40 Hz square wave, 1000 Hz sine wave, the vowel "ee," white noise burst) played at a rate of 200 msec for each of the four items, their subjects had great difficulty in judging the order of the sounds. The four different sounds tended to form separate perceptual channels unless played very slowly. Bregman and Campbell (1971) suggested that stream formation is a primary auditory phenomenon, and Bregman and his associates have performed a number of later experiments relating to the perception of melodic matters in quasi-musical situations. Van Noorden (1975) related phenomena of fission and temporal coherence to music and to certain rules of counterpoint and musical composition stemming from the Classic and Baroque eras.

Bregman and Campbell (1971) defined a stream as "a sequence of auditory events whose elements are related perceptually to one another, the stream being segregated perceptually from other co-occuring events. We assume that attention cannot be paid to more than one such stream at a time, i.e., that the *apparent* simultaneous streams produced by this process have the same properties as actual simultaneous streams sent to separate ears."

Van Noorden (1975) made it clear that, within limits, our listening strategy is important, and he determined those limits experimentally (p. 10).

> The *temporal coherence boundary* is the boundary between temporal coherence and fission when the observer is trying to hear temporal coherence. The *fission boundary* is the boundary between temporal coherene and fission when the observer is trying to hear fission. . . . At tone intervals above the temporal coherence boundary fission is observed no matter what the observer's attentional set, and at tone intervals below the fission boundary there is always temporal coherence; we may thus speak here of *inevitable* fission and *inevitable* temporal coherence, respectively.

The connection between auditory stream segregation and music can be shown vividly in certain movements of sonatas for solo instruments by J.S. Bach, where the single string of tones often breaks up into two, sometimes three, lines of melody, as shown in Fig. 1. The music for solo violin is shown on the top staff, and on the three lower staves, I have parsed it into its streams.

Bach's Finale from his C Major solo violin sonata displays three streams, from the beginning of the excerpt to measure 18, and two streams, in measures 19 through 22. The streams are well maintained in perception, even when there are quite long times between events in a stream; streaming is further facilitated by well-defined pitch registers for each of the streams.

The changes in texture from one to two to three streams are particularly interesting. For example, the downward scale passage on the last two quarter note beats of measure 18 dissolves the earlier established three voice scheme ("erases" it) and clears the perceptual field for the setting up of a two voice texture from measure 19 to the end of the example.

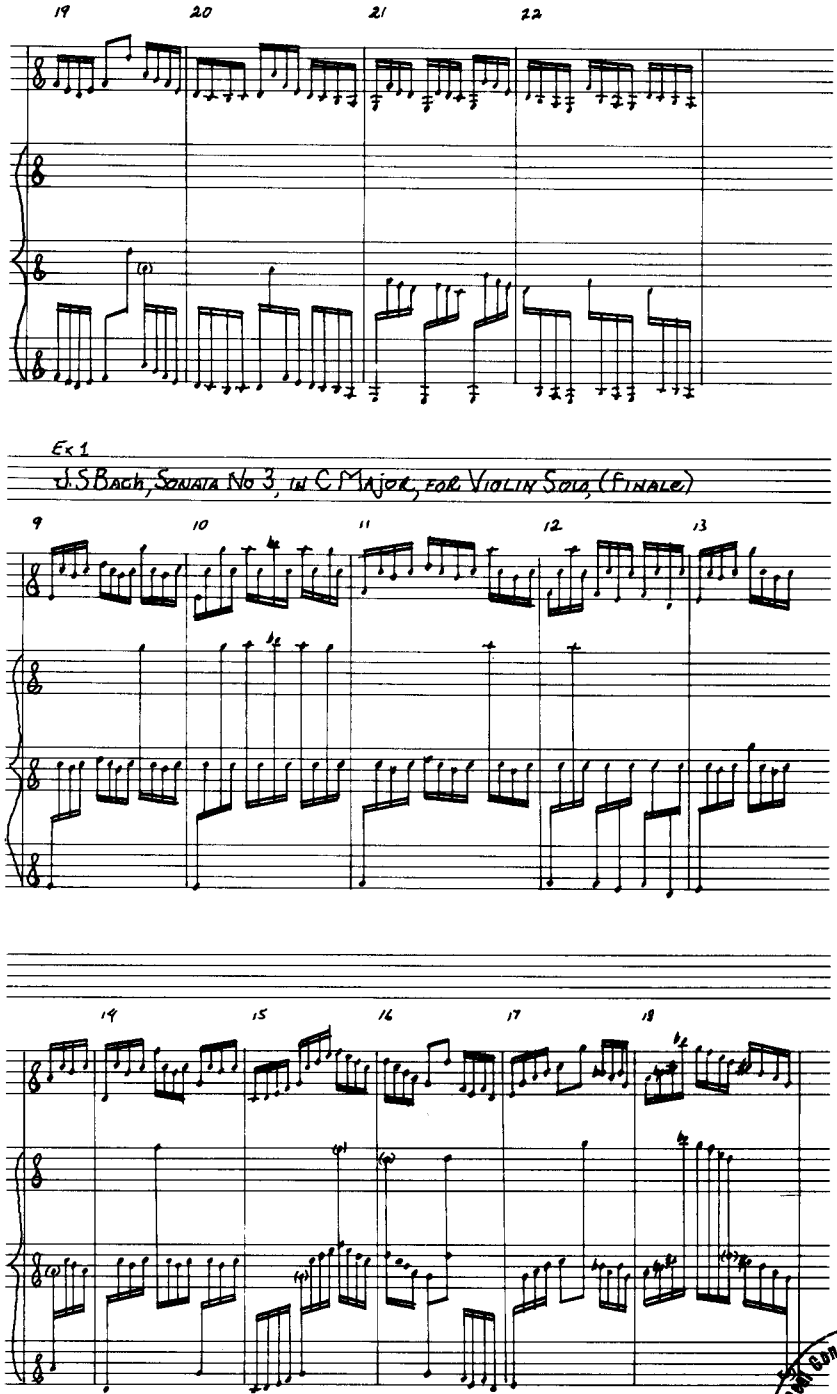



Fig. 1

Biblioteca del Real Conservatorio Superior de Música de Madrid

My layout of the music sometimes shows a particular pitch belonging to two channels simultaneously (measure 16) to indicate a real ambiguity of channel. More often pitches shown in two channels will have one marked in parenthesis to show an ambiguity more easily resolved in listening. Especially at these points, the way we listen (our "set") can affect our perception of the music; viewed musically, these are pivots for changes between one channel and two, or between two channels and three.

The product of this sort of compound melodic line, as it is usually called by musicians, is polyphonic music. A two-part or three-part texture is easily heard. It is probably impossible for one to hear this as music without breaking the string of pitches into streams. Moreover, one hears an implied chordal skeleton with a harmonically meaningful bass line, and normal voice leading in all the virtual voices.

In none of his solo sonatas does Bach use more than three streams. Is three the upper limit? Informal experiments carried out by students in my timbre seminar at the University of California, San Diego, indicate that it is very difficult for musicians to maintain more than three in perception, when the material is disposed in alternating schemes such as those in the Bach solo sonata. Would a wider pitch range facilitate formation of more streams? At furthest stretch, using all musical resources of pitch, timbre, melodic and rhythmic patterning, how many simultaneous streams can we attend to, in the sense of being aware of their identity as streams? More than three? If so, would additional perceptual strategies be involved? Throughout the Bach excerpt in Fig. 1, there are surely musical grounds for "vertical" listening, especially at measures 21 through 22, where we can easily hear the music in terms of a dominant chord. With four or more streams, vertical listening may be even more likely because possible elements of melodic lines would be perceived more easily as elements of a vertical configuration.

Some music is directly related to the rapidly alternating tones design used by Van Noorden and Bregman and his associates, and most other investigators. But more often, the situation in music is mixed, with slower alternation, with some pitches held over to create real rather than virtual polyphony, and with other sorts of stream formation mixed together with those aspects of stream formation relating to pitch. Nevertheless, the model is of great value for our understanding of traditional music concepts. It can extend and even change some of them, and in the case of some contemporary music, this kind of music may be a necessity. The term in music for alternation of this sort is *hocket*. Traditionally, it means the breaking up of a melodic line between two or more singers, each singing a single note or a very few notes; a hocket among instruments is thought of as the same sort of distribution among instruments. It is a very general term, but its root notion is that one participant sings or plays while the other rests. In that way, the melody may be distributed among a group of performers.

There are not very many examples of written hocket in Western medieval music, but it is thought by some scholars that liturgical singers introduced hocketed passages into music notated otherwise. Beethoven's late quartets have many hocketed passages, some of dazzling complexity. In our century a most original development of hocket can be found in the music of Anton Webern. Hocket is common in the music of various peoples of Africa and Indonesia and, in fact, can be found in one form or

another all over the world. Performance of hocket is difficult for Western musicians. Our whole system of notation and performance developed toward different musical goals. The performance of hocket appears to need a highly developed sense of community, in which each performer knows all the parts intimately and is able to contribute his or her single note at exactly the right moment. Hockets performed by BaBenzélé and Ituri forest pygmies show how complex communal performance can be.

In contemporary music the idea of hocket appears most usually in relation to *klangfarbenmelodie*. Webern distributed melodic elements among many instruments to bring out the motivic detail of the music. His orchestration of the Bach "Ricercar" from the Musical Offering (Erickson, 1975) looks in the score like a partially completed jigsaw puzzle but when heard it makes very good sense indeed, even though individual instruments often play only a few notes (but always a significant motivic element) at a time. All of Webern's music makes use of this breaking up of the traditional linking between instrument and melodic line and makes the following questions of more than passing interest. When do we follow timbre, and when do we follow pitch when the link between single melody/single sound source is broken?

In order to learn more about this question, I made an informal experiment (Erickson, 1974) using six instruments and a repeating melodic sequence. I reasoned that if, for example, I had a six-tone sequence performed by five players with each instrument playing a single note in the manner of a hocket, each pitch of the melodic pattern would eventually be played by a different instrument, averaging out, so to speak, qualitative differences of individual instruments in relation to the various pitches of the repeating melodic sequence. The instruments used were those I had available to me: flute, clarinet, saxophone, bassoon, trumpet, and marimba. The marimba is a foreigner among the other instruments, all winds. Nevertheless, I felt that I could at least make gross judgments about the effects of pitch versus timbre on the formation of perceptual streams. In Fig. 2 the repeating melodic figure is six tones long, printed at the top of the score. The individual parts for the instruments are below, transposed for the transposing instruments.

Clearly, one is able to listen to this delicious confusion in more than one way:

1. One may follow the tune through its changes of timbre.
2. One may begin to form perceptual streams on a pitch basis. In this kind of listening the C/B♭ patterns of the high line and the E♭/D♭ patterns of the low line are clearest.
3. One may follow the line of a single instrument (marimba is easy; clarinet is more difficult).
4. One may listen—and this is most likely—in a mixed manner, using (1) or (2) or (3), depending upon the detailed musical situation at any particular moment and depending also upon ones "set."

With fewer instruments one might expect the channeling to be stronger because each instrument is heard with fewer rests between appearances. Therefore, some segments of my LOOPS experiment employ only three instruments. It might be very slightly easier to follow an individual instrumental line when three instruments are employed, but not much. The following of an instrument appears to depend here

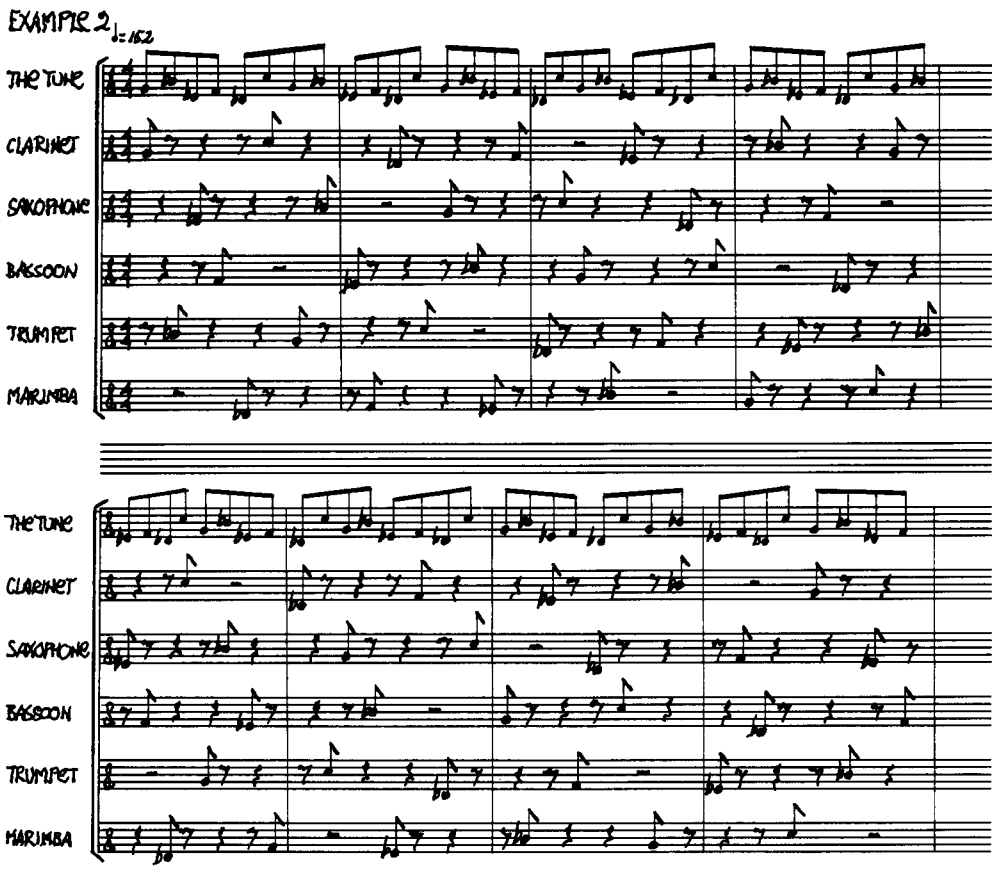



**Fig. 2**

more upon the special characteristics of the instrument, especially its attack quality, in relation to the total group or subgroup, surely reflecting the importance of context to any understanding of how timbre works in music.

Melodic sequences of from 5 to 13 pitches were used, over spans of up to three octaves, and a variety of schemes were devised to cycle the six instruments over these patterns. Rigid schemes were sometimes modified by compositional intuitions, and certainly LOOPS can be thought of as an experiment only in the vague sense that it was not conceived as a musical composition. Nevertheless, it does answer the fundamental musical question: Is this sort of disposition of material worth bothering about for music? And the answer is certainly yes—difficult, very difficult for the players, but worth the effort.

In the light of the LOOPS experiment, it appears that the answer to the question, can a melodic figure be preserved while undergoing radical changes of instrumental timbre is yes—but. The *but* includes matters such as the total range of the melodic pattern, the tempo, the particular instruments involved, their timbres at the specified

pitches, and the type of articulation of the attack and decay. It is easy, by allowing only slight overlaps of sound, to turn a precariously sequential melodic formation into something clearly polyphonic.

How strong is the effect of the timbre pattern versus the melodic pattern? Strong—but no general statement is possible. The subpatterns produced by the competition between pitch channeling and timbre channeling are local effects but controllable and full of compositional potential, not least in the area of rhythm. It is of great interest that in spite of the meter of the counting process, the patterning is chiefly a result of tonic accent, loudness of the various notes in an instrument's repertory, and articulation of attack and decay. Timbral distinctiveness or vividness (Dowling, 1973) appears important, perhaps crucial, to the formation of the subgroupings and, therefore, to the rhythm and the higher levels of the musical organization.

Why is the number of effective perceptual categories often less than six, even when all instruments are playing? This certainly could be investigated experimentally, but it is of musical interest, too, because confusions among clarinet, saxophone, and bassoon in certain contexts of LOOPS mean that there are different perceptual contrast relationships in different musical situations. The distinctiveness of a timbre, and therefore its contrast potential, is different in different registers (or even different pitches not very far apart) in a nonsimple way. We cannot think merely in terms of gross contrast—clarinet, trumpet, saxophone, etc.—but must always consider the timbre of the instrument at whatever particular pitch it is playing. Now, if certain instruments can be composed in such a way that either they can be made to tend toward homogeneity or confusability of sound or toward diversity and distinctiveness of sound, there is a possibility for structural interplay between timbre and pitch that could be of use to composers today.

If all this seems inconclusive, part of the reason is certainly the vagueness of the concept of timbre. We have come a long way from the time when timbre was identified chiefly with the spectrum of a stationary sound. We can now analyze time varying complex individual instrument sounds and synthesize convincing replicas through the computer music technology started by Mathews and Pierce (Grey, 1975; Moorer, 1975). These replicas and the potential for their controlled modification by the experimenter should make possible an approach to musical timbre and its perception that could benefit both music and science. Timbre itself has been a tangle of puzzles. Timbre in music may need many different models. That of streaming is certainly important, exactly because principles of streaming can be applied to the understanding of both timbre and pitch. We may be able to learn more about what we call pitch by learning more about timbre.

...ca del Real Conservatorio Superior de ...

## V. FUSED SOUNDS IN MUSIC

If we ask questions about how we are able to recognize the various instruments in an ensemble, certain musical/perceptual ways of thinking about timbre are uncovered. Helmholtz (1859–1954) was close to Gestalt concepts in this passage (pp. 59–60):

> Now there are many circumstances which assist us first in separating the musical tones arising from different sources, and secondly, in keeping together the partial tones of each separate source. Thus when one musical tone is heard for some time before being joined by the second, and then the second continues after the first has ceased, the separation in sound is facilitated by the succession of time. We have already heard the first musical tone by itself, and hence know immediately what we have to deduct from the compound effects for the effect of this first tone.... All these helps fail in the resolution of musical tones into their constituent partials. When a compound tone commences to sound, all its partial tones commence with the same comparative strength; when it swells, all of them generally swell uniformly; when it ceases, all cease simultaneously.... Hence we have no reason to be surprised that the resolution of a compound tone into its partials is not quite so easy for the ear to accomplish...

These remarks have been a springboard for several investigators. Van Noorden (1975) showed that in a fast alternating sequence of a pure tone and a complex tone of the same pitch but without any contiguous frequency components among its partials, we can hear two streams. Temporal coherence is heard when the frequency components of complex tones are contiguous. Noncontiguous frequency components produce fission. A musician would say that we are making the two streams on the basis of timbre differences.

Rasch (1977, 1978) took up another Helmholtz cue: that we are able to distinguish the instruments because they never play in absolute synchrony. In a series of elegant experiments, he found that asynchronies of up to 30 msec were still perceived as simultaneous and that this "asynchronization is an important factor in the perception of the various parts or voices in music."

Bregman and Pinker (1978) and Dannenbring and Bregman (1978) responded to Helmholtz's remarks from another viewpoint:

> There are two principles embedded in Helmholtz's speculation: they could be referred to as "continuation" and "common fate." Both of these are principles that have been studied by the Gestalt psychologists, e.g. Koffka (1935). The first principle states that if an element of a perceptual array forms a simple continuation of another series, it should be treated as part of that series. In audition, if a set of frequency components, not yet assigned to stream A, are a good continuation of components of a stream, they should be assigned to that stream.
>
> The second principle is that of "common fate," sets of sensory elements that change in parallel ways will be perceived as whole and distinct entities.... In audition, the principle might apply to the movement of pure tone components over time. If two of them change in amplitude together (e.g., come on at the same time) or are frequency modulated together, they should fuse into a single stream with a timbre determined by the set of fused components (Bregman & Pinker, 1978).

The method they used in some of their experiments, in which the listener makes judgments about "richness," may be open to criticism; additional experiments (Dannenbring & Bregman, 1978) explore other aspects such as intensity and onset/offset asynchrony. There is, therefore, considerable experimental data to support their striking assertion (Bregman & Pinker, 1978): "Timbre seems to be a perceptual description of a stream, not of an acoustic waveform." They conclude that "sequential and simultaneous grouping effects can compete... frequency separation is a strong determinant of the sequential effect... onset/offset synchrony is an important deter-

minant of the simultaneous effect (perceptual fusion)." They view the distinction between the sequential and the simultaneous as applying to music, in which the distinction is expressed as that between melody and timbre; and they view that the two together play their role in ordinary listening situations, parsing the input waveform to find the meaningful sources: "The sequential effect hooks frequency components into streams. The simultaneous effect, by selecting which harmonics to fuse, determines the acoustic quality (timbre) of each moment of each stream." A related experiment on the effects of stimulus asynchrony on streaming was performed by Deutsch (1979). She used a paradigm in which melodic sequences were distributed between ears and presented these with a drone pitch in several conditions. She found that onset/offset asynchrony produces less input to the two ears and that such asynchrony resulted in an increased tendency to treat this input as emanating from different sources.

In another experiment Deutsch (1975b) presented two simultaneous sequences of tones, one to each ear. The sequences formed ascending and descending scales such that when a note from the ascending scale was in one ear, a note from the descending scale was in the other ear and successive notes in each scale alternated from ear to ear. She found that this sequence was perceptually reorganized by the listeners so that (in the majority of cases) a sequence corresponding to the higher tones appeared to be coming from one earphone and a sequence corresponding to the lower tones from the other. Butler (1979) studied the effect of presenting the Deutsch scale sequence to students in a school of music through speakers as well as through earphones and found that the same perceptual reorganization occurred. He also studied the effect of introducing differences in timbre between the stimuli presented through the different earphones. Although all listeners heard an overall change of quality that pervaded the sequences, this seemed to emanate from both earphones or both speakers.

If further experimental work supports these hypotheses, musicians will have ways of thinking about timbre that go beyond those of psychophysics. The basic discoveries of psychophysics are solidly in place, and what is needed for theoretical thought about music is more work that deals with the ways in which we mentally process musical information. Experiments with fusion of vertical elements and onset and offset may give us further insight into the meaning of simultaneity in music. Musicians know that the idea of "together" is a very complex one: the onset characteristics of the instruments in an orchestra are vastly different; players must compensate for these differences and consider the further the factor is irreducible human variability; sounding "together" is a matter of context, musical context, with important components contributed by tempo, isolation by silence, register, pitch variability, vibrato, among others. There is also the fascinating fact for musicians that the bigger the ensemble the wider the "together" time window. Why should this be true?

Anything we can learn about fusion processes in audition will be especially valuable to music theory. Much contemporary music uses fused masses of sound for its musical effect—for example, György Ligeti's *Atmospheres*, in which there are no melodies, no profiled rhythms, no "voices." What is presented is a subtle play of timbre mixtures and changing textures. Textures here are good analogies to those of visual perception,

but music theory is poor in concepts dealing with textures of this sort. The usual musical distinctions are between monophonic, polyphonic, and homophonic types with additional distinctions as to the number of voices and the relations, primarily rhythmic and harmonic, between important foreground and background events.

For new music since the 1950s, two ideas about musical texture, massing and layering, are especially prominent. Dense agglomerations of pitches can sometimes be perceived as totalities, textures, sound blocks. They can be contrasted with each other sequentially, or they can be superimposed in various ways. They are objects in the sense that they have perceived beginnings and endings and are disposed cohesively by means of pitch, timbre, tempo and articulation. Nevertheless, their integrity is likely to be perceptually precarious, and that is what makes them interesting. In discussing the musical consequences for 1950s serialism, Ligeti (1965) introduced the notion of permeability: "Let us take an illuminating analogy: playing with plasticine. The distinct lumps of the various colours gradually become more dispersed the more you knead the stuff; the result is a conglomeration in which patches of the colours can still be distinguished, whereas the whole is characterized by lack of contrast." He discusses how a musical organization of these global configurations might be accomplished with a serialist ideology without an accompanying greying out of contrast:

> The function of shaping the form, which was once restricted to individual melodic lines, motifs, or chordal shapes, has been handed over in serial music to more complex categories, such as Groups, Structures or Textures, and, because of this, the way these are woven now takes over a very eminent role in the compositional design. It is possible to distinguish various "aggregate-conditions" of the material. One can see most clearly how such conditions articulate the form in compositions where the diverse types of "weave" are accompanied by considerable differences in timbre and density, and are thus even more clearly differentiated. In Stockhausen's "Gruppen," for example, the backbone of the form is given by contrasted types—hacked, pulverized, melted, highly condensed—and their gradual transformations and mixtures one with another. In this method of composition it is vitally important to pay attention to the available degrees of permeability. The two extreme types enjoy exceptionally good mutual permeability: a dense, gelatinous, soft and sensitive material can be penetrated *ad libitum* by sharp hacked splinters. Their mutual indifference is so great that the layers can get considerably "out" in time, and enjoy fields of inexactitude of considerable latitude. . . . "Soft" materials are less permeable when combined with each other, and there are places in Stockhausen's "Gruppen" of an opaque complexity beyond compare.

Cone (1972) has presented an interesting view of the sudden breaks and interruptions characteristic of Stravinsky's music from the time of *Rite of Spring* onward.

> On examination, the point of interruption proves to be only the most immediately obvious characteristic of a basic Stravinskyan technique comprising three phases, which I call stratification, interlock, and synthesis. By stratification I mean the separation in musical space of . . . musical areas, juxtaposed in time; the interruption is the mark of this separation. The resultant layers of sound may be differentiated by glaring contrast. . . . The effect may be much more subtle. . . . In almost every case, however, there is at least one element of connection between successive levels . . .

Interrupted streams characterized by register, rhythm, timbre, and sometimes by tempo can be understood as layers. Cone's concepts, interlock and synthesis, are more difficult to relate to primary cognitive processes and may belong to different analytical

levels. In Stravinskian strata pitch separation of each stratum is essential; though melodic, rhythmic and timbral organizations certainly facilitate their formation. The rather strict isolation by register makes them easy to hear and to follow throughout the interruptions.

Music theory does not have a very sharply formulated conception of what an object might be, probably because it has been concerned largely with melodic and harmonic processes in which the notion of an object can become very slippery. Nevertheless, it would be of value for an understanding of what Ligeti calls "types of material" and what other writers have referred to as "sound blocks" to look into perceptual aspects of musical phenomena that may be related to our object perception in the visual domain.

Informal experiments, many of them in seminars at the University of California, San Diego, have uncovered a few criteria that lead to agreement among musicians. Isolated sounds (surrounded by silence) can most easily be heard as objects. Compact sounds are good candidates: the beginning of the sound should be clearly defined so that it does not allow the listener to break it apart into two or more components. The end of the sound must disappear smoothly; otherwise, a listener will perceive a break and therefore another object. The law of common fate is involved here. Things that begin and end together have a good chance of being perceived as belonging together. If there are perceptible interruptions, the sound may break apart into two or more objects. If the sound lasts very long, musicians are likely to listen to component frequencies, noise components, beats, and other elements; and the object will lose its holistic identity.

Experiments were made with the sorts of concatenations that Ligeti referred to as aggregates and textures, asking the question, when can an aggregate or a texture be an object; further informal experiments demonstrated that an object could glide in pitch without losing its single object quality. These informal experiments and demonstrations—carried out by means of electronic generation, electronic manipulation, tape speed changes, and shaping with filters—as well as with musical instruments indicate that if scientific studies were to proceed along this line, the results would be of value for musicians and might have some interest for those who are studying the fusion of sounds in the real world.

Similar informal experiments and demonstrations were carried out in relation to layering and stratification, asking the question, how many layers can we hear at the same time. Remembering that melodic fission appeared to offer no musical examples of more than three streams, we wondered if layers of sounds had a similar limitation.

Segments of layered sound were presented in units from 2 to 3 sec long. Some of these had 6 or 7 levels known to the experimenter, but musicians were rarely able to hear more than 3 or 4 layers, even when the layers were distributed across seven octaves of pitch. Apparently we fuse the sounds when times are short; if more time is available, we follow streams of sound. All musicians who heard these sounds were sure that they could hear more layers if the sounds extended longer in time.

Helmholtz mentioned in the passage quoted earlier in this essay that we are helped in our identification of instruments by the fact that they rarely begin together. Com-

posers have been keenly interested since the time of Berlioz, if not before, in composing for instruments in such a way that the sound of the individual instruments are lost in an ensemble timbre. From a musical point of view we may assume that whether we are likely to hear a compound ensemble sound as a single fused timbral entity or as discrete pitches of a chord depends overwhelmingly on the musical situation: how long the sound lasts, whether players begin and end together, the pitch disposition, whether it can be more easily understood as a vertical formation or as individual instrumental parts, whether the component instrumental sounds are steady or unsteady, whether the instruments are balanced in loudness, the characteristics of the hall, and the placement of the instruments. In spite of these hurdles musicians are sometimes able to produce ensemble timbres that sometimes fuse for listeners in a concert hall. It is as though one were hearing an unknown new instrument rather than a group. Brilliant instances are to be heard in the music of Varese, but composers of instrumental music and composers of computer music are equally interested in this musical possibility.

Texture may be related to musical pitch perception, too. The choric effect, the difference between the sound of one violin and a section of 16 or more, can best be thought of as a texture. Investigators have found the choric effect to be acoustically complex and do not know enough about it; but it is musically remarkable that the same passage played by single instruments of an ensemble, say a chamber group of single string instruments, is harder to get acceptably in tune than the same passage played by multiples of those same instruments. Is it because listeners are able to select the "in-tune-ness" from the section sound, which must in fact be a pitch band of substantial width, and are unable to use that strategy with single instruments?

## VI. MUSIC THEORY AND EXPERIMENTAL SCIENCE

I believe that the answers to questions such as this can be addressed with the methods and conceptual framework of experimental cognitive psychology. They are mostly fundamental questions about perceptual processes and strategies, and it is on that level that they intersect with music.

There cannot be a very large number of classes of strategy for the processing of auditory information. The work of the last 15 years indicates that there may be only a few. Viewed from outside, modern cognitive psychology appears to be giving the rules of Gestalt psychology a solid empirical base. These findings intersect with music consistently. Music can be well understood in all its dimensions in terms of grouping—the dividing of the sound world. The complexity comes not from a multiplicity of strategies but from interactions between pitch, timbre, rhythm, texture, and other dimensions, especially interactions on the microlevel that has been so little addressed in music theory.

The problem for music theory is to discover the musical level or levels at which these processes function. I have tried to confine my remarks to those levels where a definite connection between music and perceptual processing can be demonstrated. It

is possible, even probable, that relationships between higher levels of cognition and musical organization can be shown by experiment and supported by empirical work with pseudomusical models. If that is to take place, it appears to me that it can succeed only if there is continuous close interaction between musicians and scientific investigators. There are many barriers to cross before this can be accomplished—the difficulty of communication across disciplines, the belief on the part of musicians that nothing of value to music can be learned from science, the skepticism of scientists about the vague and confusing terminology of music, especially contemporary music, and the scientist's own prejudices about what music is. In reading scientific literature that refers to music, I have found over and over that the implied definition of music within which the writer has conducted his or her discussion has been so primitive, so narrow, that it could only have been formed in childhood, probably during a period of study of some instrument, commonly the piano. If interaction with sophisticated musicians can be of value in helping scientists to expand and refine their ideas about what music has been and what it could be, it will at the same time help to erase the impression of many musicians that scientists are people who wear white coats and produce prescriptive final answers.

Both scientists and musicians may feel that my expectations from science for music are too modest. I do not think that the findings of cognitive psychology are in themselves music theory. They are (should be) the substrate upon which theories of various aspects of music can be erected. A music theory that does not take into account our perceptual processing and perceptual strategies cannot possibly be related to any real-world musical activity no matter how ingenious its superstructure.

Neither science nor music theory is of any great value to the practicing musician when performing or composing. Theory is a separate and distinct world that exists to satisfy our curiosity. Making music is a unitary process. Imagination, invention, and memories of sounds are mixed with thinking, and the thinking is not like theoretical thinking or the thinking that goes into science. That is why prescriptive music theory is so annoying to those musicians who actually make music. Consequently, I see the value of science as providing the substrate of ideas upon which a proper music theory may be erected. The task of that theory is to rationalize what musicians do.

## REFERENCES

Babbitt, M. The synthesis, perception and specification of musical time. *Journal of the International Folk Music Council*, 1964, *16*, 92–95.

Berry, W. *Structural functions in music*. Englewood Cliffs, New Jersey: Prentice-Hall, 1976.

Bilsen, F. A. Pitch of noise signals: Evidence for a "central spectrum." *Journal of the Acoustical Society of America*, 1977, *61*, No. 1, 150–161.

Bregman, A. S., & Campbell, J. Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 1971, *89*, No. 2, 244–249.

Bregman, A. S., & Pinker, S. Auditory streaming and the building of timbre. *Canadian Journal of Psychology*, 1978, *32*, 20–31.

Butler, D. A further study of melodic channeling. *Perception & Psychophysics*, 1979, *25*, 264–268.

Cone, E. T. Stravinsky: The progress of a method. In B. Boretz & E. Cone (Eds.), *Perspectives on Scheonberg and Stavinsky*. New York: Norton, 1972.

Dannenbring, G. L., & Bregman, A. S. Streaming vs. fusion of sinusoidal components of complex tones. *Perception & Psychophysics*, 1978, *24*, No. 4, 369-376.

Deutsch, D. Facilitation by repetition in recognition memory for tonal pitch. *Memory & Cognition*, 1975, *3*, 263-266. (a)

Deutsch, D. Two-channel listening to musical scales. *Journal of the Acoustical Society of America*, 1975, *57*, 1156-1160. (b)

Deutsch, D. Binaural integration of melodic patterns. *Perception & Psychophysics*, 1979, *25*, No. 5, 399-405.

Dowling, W. J. The perception of interleaved melodies. *Cognitive Psychology*, 1973, *5*, 322-337.

Erickson, R. LOOPS; An informal timbre experiment. Center for Music Experiment, University of California, San Diego (with tape of musical examples). Q-037, U.C.S.D., La Jolla, California 92093, 1974.

Erickson, R. *Sound structure in music*. Berkeley, California: University of California Press, 1975.

Fastl, H. Ueber Tonhohenempfindungen bei Rauschen. *Acustica*, 1971, *25*, 350-354.

Grey, J. An exploration of musical timbre. Report No. STAN-M-2, Department of Music, Stanford University, 1975.

Helmholtz, H.L.F. *On the sensations of Tone* (Translated by A. J. Ellis). New York: Dover, 1954.

Kalib, S. *Thirteen Essays from the Three Yearbooks "Das Meisterwerk in der Musik"*, 1975, *1*, 73-30, 626.

Ligeti, G. Metamorphosis of musical form. *Die Reihe*, 1965, 7, 5-19.

Mathews, M. V. The digital computer as a musical instrument. *Science*, 1963, *142*, 553-557.

Miller, G. A., & Heise, G. A. The trill threshold. *Journal of the Acoustical Society of America*, 1950, *22*, 637-638.

Moorer, J. A. The synthesis of complex audio spectra by means of discrete summation formulae. Report No. STAN-M-5, Department of Music, Stanford University, 1975.

Rasch, R. A. The perception of simultaneous notes such as in polyphonic music. Abstract, p. 15. *IRCAM Symposium sur la Psychoacoustique Musicale, Paris*, July 1977.

Rasch, R. A. The perception of simultaneous notes such as in polyphonic music. *Acustica*, 1978, *40*, No. 1, 21-31.

Salzer, F. *Structural hearing*. Vol. 1. New York: Charles Boni, 1952.

Schoenberg, A. Problems of harmony. *Modern Music*, 1934, *11*, No. 4, 167-187. Translated by A. Weiss.

Travis, R. Towards a new concept of tonality? *Journal of Music Theory*, 1959, *3*, No. 2, 257-284.

Van Noorden, L.P.A.S. Temporal Coherence in the Perception of Tone Sequences. Published Ph.D. thesis, Institute for Perception Research, Eindhoven, Holland, 1975.

Warren, R. M., Obusek, C. J., Farmer, R. M., and Warren, R. P. Auditory sequences: Comparisons of patterns other than speech or music. *Science*, 1969, *164*, 586-587.

# Index


## A

Absolute pitch
  accuracy, 438–440
  and heredity, 434
  and imprinting, 435
  and information transfer, 246, 440–443
  and learning, 434, 445–447
  and relative pitch, 344, 432, 337–438, 448
  and the piano, 436–437
  and timbre, 436–437
  and unlearning, 435
  disabilities relating to, 457
  octave errors, 272, 439, 440–441
  quasi-absolute pitch, 438, 445, 446
  stability, 444–445
  the value of, 447–449
Accent, *see also,* Loudness
  in musical performance, 173
  relation to rhythmic groups, 157, 159
Aesthetics, 318ff
  aesthetic episode, 500ff
  and culture, 317–320
  and emotion, 500–502, 507–511
  and mood optimization, 513–515
  and musical idioms, 498–500
  and technology, 498–499
  choice and attention, 505–507
  choice and social factors, 499–501, 502–505
  effects of complexity on emotion and behavior, 507–511
  music appreciation testing, 401–402
  phrasing and emotion in singing, 91–94
  planning of music sequences, 511–515
Agnosia, relation to amusia 456, 462–466
Amusia, 454ff
Aphasia, relation to amusia, 455–462, 466–468
Archetypes, 317ff
Arrhythmia, 150, 165–167
Attention, *see also* Channeling
  automaticity, 458–459, 467
  dual task, 218–220
  effects on aesthetic choice, 505–507
  shifts in, 215–216
  voluntary attention and grouping, 127–130

## B

Bach, 171ff, 524–526
Bartok, 172
Basilar membrane, 52, 260
Beat, 182ff
Beats, 14–16, 91, 255–256, 348
Beethoven, 119, 171, 172, 325–326, 327ff, 526
Berg, 520
Berlioz, 534
Bottom-up processes, 272, 278
Boulez, 50

## C

Categorical perception, 250–254, 349, 384, 396–397, 432, 465–466
Channeling, *see also* Grouping
  of rapid sequences of tones, 43, 48, 118ff, 523ff
  of sequences of simultaneous tones, 526, 533
Chopin, 172, 175
Chords, 273, 377, 394, 400–401, *see also* Harmony, inversions
  and tonality, 520–523
  disabilities relating to, 457
  physiological substrata, 274–277
Circular tones, 54, 263–264, 352, 358, 361, 370–372
Cochlea, 4
Combination tones, 8, 17–19
Composition, 514
Conducting, 492–493
Consonance, 19–21, 348, 376, 400
  of complex tones, 20–21
  sensory consonance, 19–21, 259–262
  small integer ratios, 255–256
Critical band, 5, 6, 15, 19, 28–29, 76, 259

## D

Debussy, 50, 172
Definition, 137, 141–142, 145–146
Development
  comparisons with adults, 416, 418
  of absolute pitch, 435
  of interval sense, 262


Biblioteca del Real Conservatorio Superior de Música - MADRID

of musical ability, 392–393, 394–395, 399, 400–401
of rhythm, 161, 415, 416
Direct sound, 135
Dissonance, 19–21, 255–256, 348, 376, 400
of complex tones, 260
sensory dissonance, 19–21, 259–262
Duration, *see also* Intervals of time
accuracy and pitch, 518
and channel capacity, 168
and emotion, 93
and grouping, 159–160
and rhythmic patterns, 165–167
comparisons with interval fractions, 220–223
of interval markers, 216–217
of notes, 482
subjective duration, 183–184

## E

Echoes, 104, 114, 136
Echogram, 138–140
Emotion, 91–94, 470, 501, 507–511
Euclidean composition, 357, 367–369, 372

## F

Factor analysis, 137, 140, 141, 143, 174, 176, 393–396
Feedback, 403–404
and absolute pitch, 446, 447
and vibrato, 488
delayed auditory feedback, 404, 488–489
immediate auditory feedback, 488–489
in judgment and production, 200–203, 225–226
social feedback, 501, 503, 504
Fission boundary, 118, 120, 128, 524
Formants, 27–28, 36, 61–63, 66–69, 69–71, 74, 79, 82–84
Fourier analysis, 4–5, 26–27, 44, 50–52
Fraction names, 185
Fractional, intervals 184
Frequency, *see* Pitch
Frequency analysis by the ear, 4–5, 26–27, 52
Frequency glide, 54, 124, 127

## G

Gestalt psychology, 7, 100–101, 116, 158, 161, 282, 529–530, 534
common fate, 7, 101, 107, 530, 533
good continuation, 100–101, 102–103, 120, 124, 161, 290, 302, 530
proximity, 100, 102, 105ff, 118, 161, 282, 302, 290–291
similarity 100, 124–125, 161
Gregorian chant, 175
Grouping
and kinesthetic response, 158
and pause, 159–160, 168
and repetition, 120–122
and tonality, 304–311
by amplitude, 126, 162
by frequency, 102–104, 109–110, 112, 118–123ff, 160–162, 527
by spatial location, 102, 105–107, 108ff
by timbre, 124–125, 128, 160, 527, 529–534
duration of groups, 157–158
fission boundary, 118, 120, 128, 524
handedness correlates, 114–116
melodic invariance under timbral change, 526–529
pitch versus timbre, 529, 530
scale illusion, 102–104, 114–115, 127–128, 531
temporal coherence boundary, 119–120, 128, 524
temporal relationships, 104–108, 113–114, 122–123, 126–127, 304–311
the octave illusion, 108–116, 122, 128

## H

Harmonic series, 520–523
Harmonics, 4–5, 7–10, 13–14, 26–29, 60, 255–256, 347
Harmony, 243, *see also* Chords, inversions
and categorical perception, 251
and culture, 320
choric effect, 534
harmonic map 376–378
testing for harmonic appreciation, 400–401
Haydn, 318–319, 327ff
Helix of pitch, 274, 286, 352–354, 361, 373, 432–434
Hocket, 128, 526–529

## I

Idiot–savants, 406
Illusions
octave, 108–116, 122, 128
scale, 102–104, 114–115, 127–128, 531
Improvisation, 483–484
Indirect sound, 136ff
Indirect/direct ratio, 140, 143, 144
INDSCAL, 365–367
Interaural differences, 101–118, 144
Intervals

of pitch, *see also* Octave, pitch
absolute identification, 260–261
adjustment procedure, 249–250, 260
and category-scaling, 247–249, 251, 260
and context-coding, 245–246, 253–254
and culture, 302–303
and information transfer, 245–246
and loudness, 518
children's perception of, 418–419
comparison with speech discrimination, 251ff
contextual independence, 247
critical band, 5, 6, 15, 19
discrimination, 250–254, 260–261
frequency ratio detectors, 260–261
harmonic intervals, 243, 246, 251, 258–259, 260, 273, 274, 276, 278, 300–301, 348
interval and chord equivalence, 273
interval class, 273, 278–282, 286
interval size and tonality, 523
interval succession, 518
learned versus innate categories, 261–262
magnitude estimation, 249
melodic intervals, 243, 246, 274–249, 250–251, 258–259, 261, 273, 277, 278, 302, 347, 348
memory for, 300–304
natural intervals, 255–262
size of intervals, 302–304
small frequency ratios, 347
successive-interval, array 285
transposition, 274–275, 279, 285, 287ff, 346, 414, 415–416, 417–418, 423–424
of time
and spontaneous tapping, 157
dependence on duration versus fraction, 220–223
full-concatenation model, 208, 213
imitation of, 207ff
judgment of, 187ff, 215–223, 457
onset and offset marking, 216–217
partial-concatenation model, 213
production of, 198ff, 224–229, 457
Inversion, 283–284, 286, 425
Inversions of chords, 260, 273, 355–356, 361, 379, 383
interference in memory, 301
physiological substrata, 274–277

## J

Judgment function, 185, 187

## K

Kinesthetic perception, 398, 403–404

## L

Lateralization, 467–468
Ligeti, 531–533
Linear predictive coding, 37, 46–47, 51
Liszt, 175
Loudness, 10–12
absolute loudness, 443
and channeling, 110ff
and classification of melody, 338
and duration, 160
and emotion, 93
and formant tuning, 63–66
and grouping, 162
and room acoustics, 136–137, 140
and spectrum, 12
children's perception of, 419
in performance, 489–490
phone scale, 11
sone scale, 11
thresholds, 463

## M

Melody, *see also* Grouping
and culture, 421–423
and interval class, 278–281
and phase-shifted tones, 116–117
and tonality, 399
ascending melody, 246, 250
changing-note melodies, 325–326, 326ff
channeling of rapid sequences, 43, 48, 118ff, 523ff
concept-identification experiments, 326–340
contour, 111–112, 277, 279, 414, 415, 416–418, 419–420, 421–423, 424–427
descending melody, 246, 250
difficulties in analyzing, 320–322
disabilities relating to, 457, 459, 462, 464
gap-filled melodies, 322–324, 326ff
handedness correlates, 114–116
hierarchical structure, 290–291, 320–322, 337–338
hocket, 526–529
interdependence of features, 422–423
melodic space, 375–376
melody perception in childhood, 416–421
melody perception in infants, 415–416
memory for, 118–119, 279–281, 301–302, 304–311, 397–398, 423–427, 485
recognition of, 278–280
semitone versus tonal representation, 424–425, 426
testing of melodic ability, 394ff
the scale illusion, 102–104, 114–115, 127–128, 531
transposition, 116

two-channel listening, 101–118
Memorization, 406, 483
Memory
and interval size, 245–246
and lateral inhibition, 297–298
and organization of a sequence, 158, 294–300
for absolute pitch, 292
for hierarchical tonal sequences, 304–311
for interval sizes, 424–425
for melody, 118–119, 279–281, 301–302, 304–311, 397–398, 423–427, 485
for melody in children, 416–417, 419–420, 420–421
for pitch, 245–246, 292–304, 397–398, 437–438
for waveform data, 423
interactions between pitches, 292–300
interactions between systems, 301–304
interference and octaves, 272
octave generalization effects, 298–299
Meter, 171–173
Modulation transfer function, 142
Morse code, 168, 169
Mozart, 318–320, 322, 324, 327ff
Multidimensional scaling, 14, 33, 47–49, 138, 174, 287, 356, 365–367, 372
Music and psychoacoustics, 517ff
Music appreciation, *see* Aesthetics
Music reading, disabilities, of 457, 459–461
Music theory and psychology, 517ff, 534–535
Music and culture, 520, 522, 526–527
Musical ability, and intelligence 394, 404–408
Musical achievement, 391–393
Musical aptitude, 391–393
Musical training, 419–420, 427

## O

Octave, 346ff, *see also* Pitch, intervals of pitch
adjustment procedure, 249–250
and culture, 244, 263, 272
and sensory consonance, 262
octave errors, 439
octave generalization, 246, 262–264, 272, 274–277, 278–281, 298–299, 346ff, 426
physiological substrata, 274–277
the octave illusion, 108–116, 122, 128

## P

Paisiello, 119
Palestrina, 499
Partials, *see* Harmonics
Perceptual restoration, 127
Performance, 223, 480ff
Personality, 408
Pitch, *see also* Octave
alphabets and hierarchies, 287–291, 304–311
analogy with vision, 274, 282–283
and bandwith, 523
and duration, 518
and frequency, 6–10, 243, 344–345, 369
and inharmonic partials, 9
as a morphophoric medium, 380–381
children's understanding of tonality, 415, 416, 417–419, 420–421
chroma, 246, 262–264, 272, 273–277, 278–282, 283–287, 298–299, 426–427, 432, 439, 433, 352, 359–364, 365–369, 371, 374
circle of fifths, 287, 320, 359–364, 365–369, 371, 382
circular tones, 54, 263–264, 352, 358, 361, 370–372
cognitive-structural approach to, 346, 349–350, 372–374
combination tones, 8, 17–19
component structures, 357–361
disabilities relating to, 457ff
discrimination, 244–246, 396, 399–400
discrimination in infants, 416
dominance region, 8–9
effect on time interval judgments, 215–216
feature abstraction, 272ff
helix of pitch, 274, 286, 352–354, 361, 373, 432–434
low pitch, 7–10, 349
mel scale, 6–7, 345, 433
memory for, 245–246, 292–304, 397–398, 437–438
microtones, 244, 248–249, 439
multi-dimensional representations of, 286–287, 347–369, 369–384
of complex tones, 7–10, 262–264, 345, 348, 371
pitch accuracy in singing, 89–91
pitch class, 272, 273–277, 278–282, 283–287, 298–299, *see also* Chroma
pitch height, 246, 262–264, 272, 352, 357ff, 365–369, 370, 371, 432, 439
pitch-set, 285
relation to rhythmic groups, 157, 160–161, 163
spatial representation of, 350–369
testing of pitch perception, 396–397ff
transcription of pitch notation, 486–487
variance of pitch and duration, 518
vibrato, 35, 54, 84–89, 488
volley theory of perception, 274
Pitch-synchronous, analysis, 31

Plans in performance, 480–487
Plato, 150–151
Polyphony, 400–401, 521–522, 523ff
Polyrhythm, 170
Power functions, 205–207
Preference, 498ff
  and complexity, 498, 502–504, 505–507
  for frequency ratios, 347
Production function, 185
Psychometric function, 3, 188–189, 194–198, 233, 234–236
Psychophysical methods, 2–4, 137–140, 231–237
Psychophysical scaling, 11–12, 143, 187ff, 345
Pythagoras, 255, 347

## R

Reaction time, 154–155, 203–205
Residue, 8
Retrogression, 283–284, 286, 425
Reverberation, 104, 114, 135–136, 141, 143
Rhythm, *see also* Grouping
  and arrhythmia, 150, 165–167
  and biological clocks, 152
  and culture, 380
  and development, 415, 416
  and lateralization, 467–468
  and melody, 278
  and movement, 150, 152ff, 174, 403–404
  and runs, 163–164
  and spatial patterns, 167–168
  and spontaneous tempo, 151–156
  disabilities of, 462, 463
  musical rhythms, 170–174
  polyrhythm, 170
  subjective rhythmization, 155–156
  testing of rhythmic ability, 393, 394, 402–403, 404, 405
Rock music, 499
Room acoustics, 136ff
Roughness, 14–16, 32–33, 53, 76, 255, 259–260, 348
Scales, 378–380
  and culture, 243–244, 257–258, 261, 379–380, 421–423
  and temperament, *see* Tuning
  chromatic scale, 243, 244, 287, 306, 323, 376, 379, 422, 520–521
  diatonic scale, 102, 243, 257, 287, 288, 306, 323, 362, 376, 378–380, 382, 383–384, 425, 520–521
  discrepancy between step size and JND, 244–246
  natural scales, 376, 378–379, 399
  pentatonic scale, 376, 378–379, 399
  scale illusion, 102–104, 114–115, 127–128, 531
  whole-tone scale, 375, 379
Schemata, 319
Schenker, 291, 521–522
Schoenberg, 50, 277, 283–284, 520–522
Schubert, 327ff
Shadowing, 129–130
Sight-reading, 486
Singing
  acoustics, 60–62
  and heredity, 393
  and masking by orchestra, 63–66, 69, 70–71
  and vowel intelligibility, 66–69, 72–73, 83–84
  coloratura, 89–90
  comparison of alto and tenor, 75–76, 77
  comparison of bass, baritone, and tenor, 73–75
  comparison of male and female voices, 73–76
  disabilities of, 457, 458, 461, 463
  falsetto, 79–82
  female singing, 62–69, 78–79
  in childhood, 416–417
  loudness, 63–66, 69, 77
  male singing, 67, 69–73, 79–82
  phonation, 76–82
  phrasing and emotion, 91–94
  pitch accuracy, 89–91
  range of fundamental frequency, 62, 69, 73
  register, 78–82
  resonatory aspects, 62–76
  respiration, 60
  the singer's formant, 69–73, 75
  vibrato, 80, 82–89
  vocal effort and pitch, 77–82
  vocal play in infants, 415–416
  voice classification, 73–76
  voice synthesis, 36–37, 43–44
Social factors in performance, 491–493
Space, 109, 114–115
Spaciousness, 142–146
Stockhausen, 532
Stravinsky, 172, 499, 532–533
Streaming, *see* Channeling
Synchronization, 154–155, 170, 198–199, 490–491
Synthesized music, 29ff, 517–519
  and musical phrases, 42–44
  signal representations, 50–54
Synthetic sound fields, 138–139

## T

Tchaikowsky, 104
Tempo

Biblioteca del Real Conservatorio Superior de Música - MADRID

and classification of melody, 338
and patterns in time, 164
and rhythmic organization, 151
and temporal coherence, 119–120
preferred tempo, 153–154ff
rubato, 489–481
spontaneous tempo, 153–154, 168
Temporal coherence boundary, 120, 128, 524
Texture, 531–534
Timbre 12–14, 26ff
acoustic modeling, 41–42
additive synthesis, 30–36, 37–39, 53
and classification of melody, 338
and frequency shift, 27–29
and grouping, 124–125, 128
and intensity, 32
and musical phrases, 42–44
and pitch registers, 529
and reflections, 145
and room acoustics, 29, 42
and the scale illusion, 103
and trills, 518
attack transients, 30, 48, 53
brightness, 28, 43, 48
cross-synthesis, 36–37, 47
disabilities relating to, 464
distinguishing instruments, 531–534
inharmonicity of partials, 35
intensity, 530–531
nonlinear distortion, 32
onset/offset asynchrony, 530–531
onsets, 13–14, 32–33, 35, 42, 518
presence, 28
recognition of instruments and distortion, 29
roughness, 14–16, 32–33, 53, 76, 255, 259–260, 348
subtractive synthesis, 39–41
the classical view, 26–29
the effect of context, 42–44
the use of analysis–synthesis models, 45–47
timbral richness, 84
timbral space, 14, 47–49
voice synthesis, 36–37, 43–44
Time delays, 136–137, 141–142
Timing, process model of, 212–215
Tonality, 243, 303, 322, 326, 378–379, 381–384, 399–400, 424–425, 520–523
Tone deafness, 245, 431
Top-down processes, 272, 278
Transcription, 484–487
Transposition, 274–275, 279, 285, 287ff, 376, 414, 415–416, 417–418, 423–424
Trill threshold, 524
Tuning
equal temperament, 242–243, 257ff
intonation in performance, 258–259, 261
just intonation, 242–243, 250, 256–257ff, 354
Pythagorean tuning, 242–243, 256–257ff
Twelve-tone composition, 281–282, 283–284, 520

## U

Ursatz, 291, 320

## V

Varese, 50, 534
Verbal deafness, 464–466
Vibrato, 35, 54, 84–89, 488
Voice source, 60, 77ff

## W

Wagner, 64–65, 141, 175
Webern, 520, 526–527

## Y

Yankee Doodle effect, 278ff, 339, 426

**ACADEMIC PRESS**
**SERIES IN COGNITION AND PERCEPTION**

*SERIES EDITORS:*
**Edward C. Carterette**
**Morton P. Friedman**
*Department of Psychology*
*University of California, Los Angeles*
*Los Angeles, California*

**Stephen K. Reed:** *Psychological Processes in Pattern Recognition*

**Earl B. Hunt:** *Artificial Intelligence*

**James P. Egan:** *Signal Detection Theory and ROC Analysis*

**Martin F. Kaplan and Steven Schwartz (Eds.):** *Human Judgment and Decision Processes*

**Myron L. Braunstein:** *Depth Perception Through Motion*

**R. Plomp:** *Aspects of Tone Sensation*

**Martin F. Kaplan and Steven Schwartz (Eds.):** *Human Judgment and Decision Processes in Applied Settings*

**Bikkar S. Randhawa and William E. Coffman:** *Visual Learning, Thinking, and Communication*

**Robert B. Welch:** *Perceptual Modification: Adapting to Altered Sensory Environments*

**Lawrence E. Marks:** *The Unity of the Senses: Interrelations among the Modalities*

**Michele A. Wittig and Anne C. Petersen (Eds.):** *Sex-Related Differences in Cognitive Functioning: Developmental Issues*

**Douglas Vickers:** *Decision Processes in Visual Perception*

**Margaret A. Hagen (Ed.):** *The Perception of Pictures, Vol. 1: Alberti's Window: The Projective Model of Pictorial Information, Vol. 2 Dürer's Devices: Beyond the Projective Model of Pictures*

**Graham Davies, Hadyn Ellis and John Shepherd (Eds.):** *Perceiving and Remembering Faces*

**Hubert Dolezal:** *Living in a World Transformed: Perceptual and Performatory Adaptation to Visual Distortion*

**Gerald H. Jacobs:** *Comparative Color Vision*



THE PSYCHOLOGY OF MUSIC

DEUTSCH



ACADEMIC PRESS

ISBN 0-12-213560-1